CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Forecasting Battles in High Frequency Data: Classical, Machine Learning, and Deep Learning Methods

by

Muhammad Arslan

A thesis submitted in partial fulfillment for the degree of Master of Science

in the

Faculty of Management & Social Sciences
Department of Management Sciences

2021

Dedicated to My Parents

CERTIFICATE OF APPROVAL

Forecasting Battles in High Frequency Data: Classical, Machine Learning, and Deep Learning Methods

by

Muhammad Arslan

(MMS191017)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Sumayya Chughtai	I.I.U Islamabad
(b)	Internal Examiner	Dr. Saira Ahmed	C.U.S.T Islamabad
(c)	Supervisor	Dr. Arshad Hassan	C.U.S.T Islamabad

Dr. Arshad Hassan
Thesis Supervisor
February, 2021

Dr. Mueen Aizaz Zafar
Head
Dept. of Management Sciences
February, 2021

Dr. Arshad Hassan
Dean
Faculty of Management & Social Sciences
February, 2021

Author's Declaration

I, **Muhammad Arslan** hereby state that my MS thesis titled "**Forecasting Battles in High Frequency Data: Classical, Machine Learning, and Deep Learning Methods**" is my own work and has not been submitted previously by me for taking any degree from Capital University of Science and Technology, Islamabad or anywhere else in the country/abroad.

At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw my MS Degree.

(Muhammad Arslan)

Registration No: MMS191017

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Forecasting Battles in High Frequency Data: Classical, Machine Learning, and Deep Learning Methods**" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Capital University of Science and Technology towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS Degree, the University reserves the right to withdraw/revoke my MS degree and that HEC and the University have the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized work.

(Muhammad Arslan)
Registration No: MMS191017

Acknowledgement

All praises the ALMIGHTY ALLAH, The creator of universe, the most benevolent, and the only to be praised, whose blessing and exaltations flourished my thoughts and enabled me to improve my knowledge up to this stage. Special praise and salutations on the last prophet HAZRAT MOHAMMAD (P.B.U.H.), who enables us to recognize our creator.

I would like to take this opportunity to pay my gratitude to my honorable supervisor, Prof. Dr. Arshad Hassan for his supervision, illustrious advice, guidance, and support during my course and research work. I feel lucky to have such supervisor who cares for my work and who responded my every irritating query. He is truly an inspiration for me and have no words to express my wishes. Further I would like to express gratitude to Dr.Wajid Shakeel Ahmed (Assistant Professor COMSATS University, Islamabad) for his generous time, support and countless help in my analysis section especially for developing code on python. I also want to acknowledge the "Dukas Copy" and "Jforex" for making free access to the high frequency data.

This was really a tough journey with endless hopes would never be possible with help and support of family, teachers, mentors and friends. I am extremely grateful to my parents and siblings for encouraging, supporting me especially financially. Insha Allah will make you guys proud again and again. In the end I would like to take this opportunity to say thanks to my friends, fellows and best wishers for their support and prayers.

Muhammad Arslan

Abstract

Forecasting not only plays a crucial role in developing business operations plan within an organization but also play a vital rule for investors while putting their eggs in one bucket. Financial markets are the complicated in a sense of predicting. Many factors influence them at a time which include the financial or economic factors and other factors e.g., political and government policies and market expectations. Consequently, these factors make financial markets complex while forecasting the financial time series. The objective of this study is to find out best forecasting method out of classical, machine and deep learning techniques with inclusion of a new Hybrid GARCH-LSTM method in forecasting family. Along with this study postulates that data frequency plays a vital role in forecasting and based on empirical results propose best data frequency for financial time series forecasting.

Study employs intraday data of three different time intervals and 16 stock market indices. Results of this study provides a strong evidence regarding deep learning method (i.e., LSTM, ARIMA-LSTM and GARCH-LSTM) outperform the other forecasting methods. However, proposed method provides mix evidence of effectiveness in different markets. In hourly and 10-minute data frequency proposed GARCH-LSTM model outperform the other deep learning methods. In the meanwhile, study finds 10 minutes frequency is better to be used for forecasting purpose as it captures the most market patterns with relatively low noise.This study not only equips the investors and investment institutions to get better forecasting results for stock market indices. But also helps policymakers to choose the best forecasting model to forecast the other dimensions of the economy.

Keywords: Forecasting, Modern Financial Time Series, classical, Machine learning, Deep learning, GARCH-LSTM.

Contents

List of Figures

List of Tables

Abbreviations

ANN	Artificial Neural networks
ARCH	Autoregressive Conditional Heteroskedasticity
ARFIMA	Autoregressive Fractionally Integrated Moving Average
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
CNN	Convolutional neural networks
EMH	Efficient Market Hypothesis
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MSE	Mean Squared Error
PSO	Particle Swarm Optimization
RMSE	Root Mean Squared Error
RNN	Recurrent Neural network
SVM	Support Vector Machine

Chapter 1

Introduction

1.1 Background of the Study

Financial markets are the imperative part of financial system. That perform a critical function by the facilitating the organizations for the saving and investment purposes. Stock markets considered as the performance indicators for the investors. This gives the signal of future investment or future moves for the investors. Stock market indices depict the overall performance of a specific stock market. Since the beginning, investors have a key interest in forecasting the trend in the prices of the financial instruments. Similarly, right on other hand academician also have interest to explore the future market movements. The uncertain environment creates the baseline for the forecasting. although it does not eliminate the uncertainty complete but reduce significantly (Krzysztofowicz, 2001).

Financial markets are the complicated in a sense of predicting. Many factors influence them at a time not only the financial or economic factors other factors e.g., political and government policies and market expectations. These factors makes financial markets complex and complicated in term of application in modern time series forecasting (Falinouss, 2007; Kumar et al., 2016). Because of the chaotic market, stationarity issues, noise in data and dynamics environment contributes in complexity (De Gooijer and Hyndman, 2006; Marszałek and Burczyński, 2014). Over the decade's researcher had developed many of methods that helps to forecast the markets (Arévalo et al., 2016). During 19th century the world was categorized on the idea of deterministic. Since Yule (1927) contribution in notion of stochasticity in time series a number of model has been introduced in literature that has properties to deal with the estimation of parameter, Identification, model checking and forecasting.

De Gooijer and Hyndman (2006) highlight some classic work on field of time series forecasting almost three decades from 1986 to 2005. Exponential smoothing method use the exponential window technique to smoothen the time series data. This exponential function is used to decrease the weights of past observations over the time period. As financial markets are full of noise and assumption of randomness this exponential smoothing technique help to achieve optimal forecast. This past history has information and as, Marszałek and Burczyński (2014) mention information has a huge impact on financial markets and does not follows the random walk pattern. Whereas, exponential smoothing techniques follows the parameters of no seasonality and trend in data (Kwon et al., 2015).

Earlier work on AR (Auto Regressive) and MA (Moving Average) provide the foundation of linear forecasting based on parameters. **?** integrate the classic AR and MA model which is known as ARIMA (Auto Regressive integrated moving average) a generalized class of ARMA model that is a widely used foresting technique. The AR term explain the relationship between the past observation with the current observation. Which tell us about the existence of pattern in other words explain the randomness in series. MA explains the informational impact on current price. But as we mentioned, the stationarity issue in time series data, previously available ARMA model has the limitations if data is non stationary. With the integration now a non-stationary and linear data can be used for forecasting at d number of levels. Through the integration ARIMA reduce the seasonality if trend exists.

ARIMA works on a simple mechanism of using series own lag values, lag values of error terms to forecast the forehead value and denoted with three terms p, q and d term. Whereas, p is the AR order, q is the ma order and d refers to the level of differencing. Similarly, it removes the random movements in data. Because of parametric nature, these models cannot handle the non-linearity phenomenon. Due to dealing with seasonality and trends ARIMA model is considered as alternative to exponential smoothing model(Brownlee, 2018). In modern time series analysis ARIMA has built a huge influence both in theoretically and practically. The development of computational power gives popularity to this linear parametric model of forecasting in different areas.

Volatility is the major problems with these parametric and semi parametric models. ARIMA or ARFIMA does not consider the volatility in market. In market we notice the thick tale, volatility clustering and leverage effects so these models are not able to handle these issues. On other hands returns are not the only focus we have to consider the volatility as well. Many of studies have verified that the financial time series data is time varying along with characteristics of non-linearity and stochastic volatility. Statisticians define volatility as the measure of dispersion on returns of given set of financial assets or financial markets indices. Mostly refer as the amount of uncertainty which relates to the change in the value of specific asset or portfolio. When it comes to dispersion, higher the dispersion from the mean value refer as the highly volatile security which has a dramatic potential of value change or spread over the large values. On the other hand, a lower the dispersion rate refers as the low volatile assets which means that the value of security does not fluctuate dramatically. Sometimes volatility is considered as a risk but it's not same as risk. Usually interpreted as the uncertainty and used as an input by the decision maker while creating the portfolio (Poon and Granger, 2003).

Volatility refers to measure of probable variation or the movement in a certain economic variable or part of that variable, which mostly estimated on observed realization of random variable over the past time period known as realized volatility (Wolf, 2005). In simple words realized volatility is measured by the standard deviation of an economic variable over the past time frame. This volatility is linked with the information inflow to the market. Market adjust the prices as information release sometimes it takes few minutes to adjust but considered more volatile for next few hours to adjust (Ederington and Lee, 1993; Frino and Hill, 2001; Smales, 2013, 2015). This inflow of information categories as good news or bad news that affect the market and the disturb the variance of market. Increase in volatility will raise the expectation level thus raise in expected return where as stock prices get the bullish trend, bad news creates more volatility then the good news in market (Baur and Dimpfl, 2018; Glosten et al., 1993; Goudarzi and Ramanarayanan, 2011; Nelson, 1991). This information inflow creates the environment of uncertainty and this uncertainty is linked with the volatility positively. Announcements of any event in market cause the uncertainty and the intensity on that news/event depends upon the nature of that event. If the news nature is known and intensity is known as well then, the event cause high uncertainty before the specific event and a high change in the prices of assets is expected (Bouoiyour and Selmi, 2018; Kolaric and Schiereck, 2016).

Based on this information flow two kind of opportunist's trade in market, the news watchers and momentum traders. New watchers creates the tendency of underreact for prices in short run and this underreaction creates the way for the momentum traders to make money by chasing these trends (Hong and Stein, 1999). The most common techniques used by the professionals and academicians to capture and forecast the phenomenon of volatility are ARCH (Auto Retrogressive Conditional Heteroskedasticity) which was developed by Engle (1982) and generalization of ARCH is known as GARCH. These models are referring as the conventional models to measures the volatility with in the data by taking the assumption of unconditional heteroskedasticity. That is why known as the non-parametric quantitative models in statistics. A number of studies has been conducted that provides the evidence of effectiveness of GARCH based model for forecasting (Brooks, 1998; Brooks et al., 2001; Franses and Ghijsels, 1999). Caner and Önder* (2005) defines the sources of volatility. Like financial news are also an important sources of volatility which is ignored by the most conventional time series model ARCH and GARCH (Luss and d'Aspremont, 2015). Over the past few decades many of the researchers have modified the methods from standard GARCH model to multivariate GARCH to handle different sources of volatility.

With the development of new technologies demands are significantly increasing. Now businesses need more advance business management tools and techniques to predict the future market demand and business decisions. These decisions are the highly influenced by the inflow of information which can be referred as data (Floridi, 2005). While estimation and sampling, counting and measurement error must be taken into account for the validation of data. And identifying the data pattern has never been easy. Under such environment quantitative forecasting with simple model is really a big challenge for the analyst working in financial markets (Ding et al., 1993; Giot et al., 2010; Li et al., 2020; Slim and Dahmene, 2016). Over the past three decades technological development in machine learning methods has introduced many of forecasting techniques that outperformed the predictive power of the classic methods. Recently numerous machine learning methods such as Artificial Neural networks (ANN), Recurrent Neural network (RNN), Support vector machine and others have gain tremendous popularity (Cavalcante et al., 2016; Kumar et al., 2016).

ANN has been the part of time series analysis since late 1980's. Which is inspired by the biological brain structure of compiled neurons (Arévalo et al., 2016). ANN use the dependent variable as input and filter this input through the hidden layers that consist of further hidden units and process this for the output variable. These ANN models have the ability to extract the hidden inside informational pattern from data. Although ANN has been used rigorously in other fields but in financial field was introduced in early 1990's (Arévalo et al., 2016; Kaastra and Boyd, 1996; Trippi and Turban, 1992). Many of studies has incorporated the ANN to predict the performance of stock market and declare it as better method for nonlinear data (Choi et al., 1995; Hippert et al., 2001; Qi, 1999; Quah and Srinivasan, 1999; Zhang et al., 1998).

ANN often face challenges that also known as the weakness of ANN. First, sometimes ANN miss the bigger picture and start coverage on local minima. Second, training period of pattern. If training period goes long there are chance of noise that ANN start considering as a part of pattern. It also faces the issue of overfitting of model due a huge number of parameters. This overfitting leads the inconsistency in performance results of ANN model (Kim, 2003). Similarly generalization issues, appropriate network architecture even author like Lu and Wang

(2005) mentioned ANN is more attracted by the environmental researchers.

These limitations are solved by the SVM(Support Vector Machine) developed by Vapnik (1995) that is used to separate and label the data and known for its simplicity. By using the Structural risk minimizing principle, SVM minimize the Upper bound of generalization error. Machine learning was initially informally defined by Samuel (1959) "field of study that gives computers the ability to learn without being explicitly programmed". More specifically ability of systems that lean and improve automatically from their experiences using artificial intelligence. Usually, machine learning is separated in two major categories: (a) Supervised Machine learning which refers as predefined results has been labeled with the training data set which supervise the algorithm with already known answers. (b) Unsupervised learning referred as untold or unsupervised algorithm in which previous knowledge about the data is not given neither trained nor labeled (Schrider and Kern, 2018).

SVM is a classifier or supervised learning techniques which is used label or classify the data. Discriminant analysis, Naive Bayes, Nearest Neighbors, linear regression and neural networks are also the part of Supervised machine learning. A number of studies adopted SVM techniques to predict the financial time series and has found better predictor based on its properties (Kim, 2003; Tay and Cao, 2001, 2002).

Traditional neural networks have a shortcoming of thinking from scratch. Like being a human when you start reading a line, under the next word based on the previous word you will not start think from start. This is later solved by the Recurrent neural networks (RNN) that allows output of previous step as input for current step. Like in ANN the inputs and outputs works independently whereas in RNN out or previous task needed to perform the next step. Working with traditional neural network problem of long-term dependencies arise. When the past information and the targeted future information has a large distance introduce by the (Hochreiter and Schmidhuber, 1997). Prices in stock markets are reflection of all current and previous information, and with simple neural networks the earlier information may lost.

RNN often face the vanishing gradient problem while using the certain activation functions, the loss function gradient approaches to zero and this make a network difficult to train. These gradient functions are the algorithm used to estimate the coefficient values that minimize the loss function. If during the training of a deep learning network gradient goes to zero or very small then this problem arises. To deal with this vanishing gradient problem, Hochreiter and Schmidhuber (1997) propose long short-term memory network. Sophisticated LSTM network gives some astonishing results in sequence learning tasks. LSTM network extend the learning memory that help network to remember important pattern of information in network. This network can read, write and delete the information which is stored in memory. LSTM work on three gates the input gate to let information in or stop the new information, the forget gate which delete the unimportant information and output gate. Yan and Ouyang (2018) study found that LSTM is a better method to predict the daily returns of shanghai composite index.

1.2 Motivation of the Study

Stock markets has been evolving our years. This advancement eases the processes for both regulator as well as for the investor. But every picture has two sides. Induction of latest technologies widen the scope and capacity of information collection and information processing. In the meanwhile, it creates the hurdles for investors and institution to process this information in more meaningful way. Now the information in shape of data is even available for in high frequency (small tick). Prices reflect the information as EMH states. It means a person has greater and quicker access to information and then process that information can easily make profit form it vice versa. Now financial environment is highly uncertain and volatile. There are multiple estimations methods exists in literature ranging from simple correlation-based methods to parametric and from parametric methods to non-parametric methods. These parametric methods are generally better than the simple correlational methods but has complex properties to compute (Zhou, 1996). Similarly, non-parametric models are even more complex to compute but with better predictive power.

Methods like ARCH, GARCH, EGARCH and other members of ARCH family is now widely used methods to capture the phenomenon of volatility. ARIMA (Auto regressive integrated moving average) has been used for forecasting of volatility. Similarly fuzzy logics by Zadeh (1965) are now combined with GARCH known as FGARCH now used to capture the uncertainty and then volatility from specific data set. In the field of statistics, it is very well said by Box (1979) " All model are wrong, but some are useful". This creates a question in the mind of satiations and gurus of mathematical finance. All these models are known as the conventional models from ARCH family. Over the past few years computing is one of the most prominent emerged phenomena in the body of knowledge because it opens the door of whole new world in each and every section of the market due to informational transmission. Because now the process of arrival of information, informational spillover and its adjustment becomes more efficient. further information travel with the speed of electricity around the globe. Information is most affective source which can spread rapidly that pointedly affect the individual behavior and decision making (Fu et al., 2017; Qu and Dan, 2014).While estimating the volatility crucial part of data set is determining the frequency because volatility pattern changes with the frequency of the data set.

Previously available models have limitations to deal such high frequency information. To compute high frequency information more advanced and technical models are required due to the characteristics of high frequency information. Presence of high noise characteristics and semi strong form of efficient market it is enormously difficult to predict in financial time series. Market anomalies can affect the expected returns on market. More than 100 anomalies has been identified in literature that can be helpful to beat the market return by using the some specific financial models but these models has the limitations dealing with non-leaner dependency (Shen et al., 2018). Many of studies verify that the financial time series data is time varying along with characteristics of non-linearity and stochastic volatility. Under such environment quantitative forecasting with simple model is really a big challenge for the analyst working in financial markets (Ding et al., 1993; Giot et al., 2010; Li et al., 2020; Slim and Dahmene, 2016)

Machine learning algorithms make it easier. From linear regression to logistic regression, Gradient decent techniques, SVM (Support vector machine), to ANN along with Long Short-term Memory (LSTM) are commonly used machine learning algorithms by the data scientists. Many of the studies has been done previously to improve the predictive power of ARIMA model. Van Gestel et al. (2006); Kumar and Thenmozhi (2014), and many others used modified ARIMA SVM models to predict the stock index returns. de Oliveira and Ludermir (2014) use the PSO (particle swarm optimization) technique to further modify the ARIMA-SVM model. Similarly Wang et al. (2012b) use fourier and seasonal ARIMA based on error correction approach to predict the electricity demand in china and mentioned how this prediction is helpful for both consumers and producers. Kristjanpoller R and Michell V (2018) used GACRH and modified GARCH-ANN and MS-FNN-GACRG the combination of soft computing and econometric techniques performed astonishing. Li et al. (2020) used both ARIMA SVM and ARIMA LSTM hybrid methods to predict the stock market returns of China. So, there is a list of Machine learning methods available to estimate and forecast the volatility in market.

Another problem is the frequency of data. Now the computing power, and development of science and technology open a new era of big data that has improved the storage capacity and computing power significantly. It is well fact that high frequency data hold more market information then the low frequency data (Andersen and Bollerslev, 1997; Andersen et al., 2003). Estimation and prediction power of model also enhance which gives more space and possibilities to design more suitable business or investment strategies. This high frequency analysis can entirely change the investment philosophy. Aldridge and Krawciw (2017) mention in 2016 around 10-40% of us equity is initiated on the basis of high frequency trading. Further the numbers vary from 0-100% in short term intraday trading. Another report from bloomberg mentioned that in 2010 high frequency trade "accented more than 60% of US equity volume". Barrales (2012) mention that till 2012 around 50% - 80% equity of US, 40% of europe, 12% of asia trade are based on high frequency transactions. Now interest of academic research and practitioners is to explore the dynamics of high frequency data for improving the predictability in order to maximize the returns and minimize the risk against the market volatility. In the body of knowledge mix studies are available on high frequency data to predict the market behavior in future while simultaneously which model is to be used for forecasting.

1.3 Research Gap

Forecasting is not a just a single function and selecting the best forecasting method. It is a continues developing technique and continuous dynamic process Golden et al. (1994). Emergence of technology, Data Science, mathematical and computational finance create a whole new world of FinTech. Now these new implications of technology cannot be ignored. Now the prices of stocks are nonlinear in relation as well as non-stationary. So reliability and accuracy of forecasting models is difficult and under a big question mark (Cao et al., 2019). In a review paper of time series forecasting, De Gooijer and Hyndman (2006) mention " among the 44 studied on time series analysis there is no clear winner between the classical ARIMA, ARFIMA and historical volatility based GARCH models. Many of studies are still comparing the classical and Ml methods to give a verdict who perform better. A study by Papacharalampous et al. (2019) compare 11 stochastic and 9 Machine learning methods, Khairalla and AL-Jallad (2017) mention statistical methods like ARIMA, ARFIMA, ARCH, GARCH, ARMA and machine learning methods like ANN. Results of study are shocking ARIMA model outperform the neural network based ANN-ARIMA model and ANN exponential smoothing method. Machine learning methods like SVM and ANN has the ability to find the hidden layers of patterns and complex non linearity. These ML methods are rarely applied by the financial economists.

Empirical results show these ML methods outperform the classic econometric models (Chang et al., 2009; Gong et al., 2016; Hsu et al., 2016; Kewat et al., 2017;

Żbikowski, 2015). Some advance methods also been developed such as CNN (Convolutional neural networks) as the extension of ANN compared with SVM type methods cannot beat the SVM based models (Dingli and Fournier, 2017) . On the other hand, Zhang et al. (2016) criticize the SVM because of its limitations on time correlation at two points this raise the question of efficiency in real time application. Emerging machine learning techniques also create confusion with in the Investors as well as academicians. Many of these studies to forecast the time series like Alzahrani et al. (2017); Cadenas and Rivera (2010); Cadenas et al. (2016); Yaseen et al. (2019) perform in energy sector and suggest the use of hybrid models known as deep learning methods of forecasting. Sezer et al. (2020) conduct a survey between conventional machine learning methods like ANN, SVM and Deep learning methods like RNN, Hybrid version of LSTM in financial sectors. The study suggests that the deep learning-based methods are emerging methods and they are creating opportunities for researcher. Many of studies support the hybrid methods of Classic and machine learning methods. Another study by Li et al. (2020) use hybrid methods to evaluate Chinese stock market and also suggest to improve the GARCH based model of forecasting using the machine learning methods. This creates another confusion either independent models work best or the hybrid methods. So, there is a need to evaluate the commonly used classical and machine learning methods.

Similarly, now the computing power and development of computational science and technology opens a new era of big data that has improved the storage capacity and computing power significantly. It is well fact that high frequency data hold more market information then the low frequency data (Andersen and Bollerslev, 1997; Andersen et al., 2003). Studies on high frequency data does not define what and which frequency is better to predict the future market behavior as well as which method is good to evaluate either the conventional methods or the machine-based method are better predictors in that particular frequency. Matias and Reboredo (2012) use 5, 10, 20, 30, and 60 minutes data intervals to evaluates different forecasting techniques but the focus of study was not the data frequency but the forecasting techniques. All of these above-mentioned studies mainly based on low frequency data probably monthly, weekly or daily. A few of these used the high frequency data like Li et al. (2020) but all of these only consider either one market or 30 minute or 10 minute interval of time for analysis. All of these studies are conducted on one single time series data or single financial market. A comparative study in across market is missing in literature. Purpose of this study is to fill the gap in finance literature and provide an adequate technique to forecast the financial markets returns for investor.

1.4 Problem Statement

Numerous studies have been found based on working of financial time series forecasting with classical forecasting techniques, machine learning techniques, deep learning techniques and now hybrid deep learning techniques. A list of comparative studies can be easily found in literature some studies like (De Gooijer and Hyndman, 2006; Khairalla and AL-Jallad, 2017; Papacharalampous et al., 2019). Recently development of Hybrid techniques emerged that combines the classical econometric models with soft computing methods. Some recent studies like Livieris et al. (2020); Kulshreshtha (2020); Li et al. (2020) improve deep leaning methods with the help of conventional econometric and machine learning techniques. Empirical testing of these proposed works is done on single timeseries data set. Further they tried to incorporate different econometric techniques and suggested in future directions to incorporate other techniques as well. In this study, test the pre-defined Hybrid method with inclusion of some other econometric models on multiple financial time series from different origins. Further this study plan to incorporate GARCH model with hybrid technique to improve to predictivity of LSTM deep learning technique.

1.5 Research Questions

This research will answer the following questions:

- Are classical models successful in forecasting price behavior in financial markets?
- How do machine leaning based models perform in forecasting price trends in financial markets?
- Do machine learning based models outperform classical models of forecasting?
- Does GARCH based LSTM model outperform other Deep learning, machine learning and classical methods?
- What is appropriate frequency for forecasting return behavior?

1.6 Research Objectives for This Study

Objectives of the study are as follows:

- To identify an appropriate forecasting method for forecasting price trends.
- To identify an appropriate frequency for optimal forecasting.
- To propose a forecasting model based on GARCH-LSTM.

1.7 Significance of Study

With the development of technologies practitioners and academicians needs cohesive computing techniques to evaluate, analyze and forecast the time series. As the goal of computational learning theory to design and evaluate the meaningful algorithms. This study not only contribute the novel GARCH based LSTM model algorithm but also validate the previously improved Hybrid machine learning algorithms as well as classical forecasting techniques (Li et al., 2020). On other hand this study contribute in both statical and mathematical finance by developing and validating the statistical methods of forecasting. As data frequency brings the effective results this study contribute statistically validated frequency to be used for analysis.

Due to higher returns stock markets become a common channel for investor and for academicians. Eye on future movement is in interest of both (Wang et al., 2012a). It is obvious one who is currently working in financial markets or with financial markets needs a better understanding of both classical and machine learning methods. As the objective of study, this study propose a novel model of forecasting along with other forecasting techniques that will be beneficial for the field of finance. These financial markets are the driving force for the economic development. With batter tools in hand, it will reduce the element of uncertainty from these markets as well as other financial markets.

Based on empirical work, this help both the policy regulators, institutions and individual investors with the understanding of most advanced forecasting tools and techniques. For policy regulator this help to forecast the economic indicator and set the economic direction in a more efficient path. In the meanwhile, for the investment institution and individual investors efficiency and lower rate of error is the most important subject. This study give them a tool which help them to forecast their investment returns more efficiently and help them to reallocate their invested funds for batter outcomes.

1.8 Plan of Study

The study is planned as follows. Chapter two of the study discusses the theoretical background of efficient market hypothesis and computational learning theory along with the empirical findings of different forecasting techniques, followed by the data sampling and empirical methodology in chapter three. Chapter four discusses the results and findings of the study and in the last chapter five concludes the findings of the study along with the recommendations and limitations.

Chapter 2

Literature Review

This chapter covers the theoretical and empirical studies in the domain of forecasting using classical forecasting methods i.e., (ARIMA, ARFIMA, and GARCH) along with machine learning and deep learning methods i.e., (SVM, LSTM, hybrid ARIMA-SVM, and ARIMA-LSTM).

2.1 Theoretical Background

Market efficiency refers to the degree at which the prices of assets reflects the all available information floating in market. Fama (1970) defines the efficient as a market that fully reflect the information. Literature mostly addresses three important kind of market efficiency known as allocative efficiency, operational efficiency and informational efficiency. Operational efficiency means how well an organization or individual execute the function in term to achieve the cost minimization goal. Allocative efficiency measures the allocation of resources to achieve the maximum benefits. Informational efficiency measures the how rapidly the market adjusts the new information into prices. In real economy, prices of securities are the result of proceed information. Because of more informative prices this help in better informed investment and financing decision (Boehmer and Kelley, 2009). This information pricing classify the market participants running in race of buy and selling of assets. There are three major player bid in stock market: (a) Rational investors who made decisions based on information (Daniel and Titman, 1999), they just use the fair price policy use there analysis either fundamental or technical analysis and fix the bid regardless of market ups and down, (b) Irrational investors who hold the optimism and pessimism biases in their behavior (Prosad et al., 2015). Based on this behaviors we can divide them into two categories further.

Both optimistic and pessimistic investors offset the each other's effect and hence market achieve the efficiency, (c) market Arbitragers are the third common type of investors who takes the miss pricing benefit arises due to irrational investors by taking long and short positions and so market efficiency move back to equilibrium. Tobin (1984) identify other types of efficiency as well. (1) Informational arbitrage efficiency market is the market where it is almost impossible for a person to gain advantage from trading on the publicly available information. Only insider can make money in efficient market. Anything we know about any particular security market already discounted it. (2) Fundamental Valuation efficiency refers to the measure of how accurate prediction of returns on a specific investment. Investment decisions by the investor are performed under the assumption that the price of specific security reflect the maximum of available information in market. The idea behind is the assets market that has good regulation, market makers and developed markets that has a lot depth so then the prices we see are perfect indicators of true value. Efficient market efficiently incorporate the all information about the value of assets. Based on this assumption, fama presented the efficient market hypothesis (EMH). Gibson (1888) first time defined the efficient market hypothesis although it doesn't call it in his book. "When the shares are publically known in open market the value which acquire there may be regarded as the judgment of the best intelligence concerning them". An interesting part of that book was the information flow. It says the information in today's advance world travel with the speed of light. This seems that the book isn't written in 1890's sounds like written in 1990's. We cannot beat the market or get the access return from the market because it already adjust the information in prices. According to them regulators of financial markets has shown a great faith on EMH and therefore, they feels that their primary mission is to regulate the flow of information just to make sure to give an equal opportunity to everyone by providing the same field to play.

Coming back to Fama's EMH, which divide the market into three forms according to the efficiency of market. First one is the weak form of market. Under this hypothesis market prices of assets only reflect the historically available information. Under weak form market only consider the past stock prices it does not use any other information alike earning, merger and acquisition information etc. In the weakest form of efficiency, information about the stock prices is easily and widely available to everyone. If there is a possibility to outperform then the market, then everyone in the market would do so. If market hold weak form arbitraging possibility arise along with the probability of loss. But in Weak form of efficient market information impounded and the trader with the skills enough to exploit the market get the maximum of return (Sensoy et al., 2017).

Second form of market Efficiency is the semi strong form of efficient market hypothesis. When a market reflects the all-available historical prices information and publicly available information like publicly available accounting statements, dividend announcements etc. Semi strong form of efficient market holds fewer arbitraging possibilities. Stocks that are sensitive to the information adjust quickly in prices. This Market achieve the strong form of efficiency when it reflects the all past historical information, public information and all other private information. Private information refers to information that is not available to public at any level. As all public and private information is available to every market player then possibilities for arbitragers become zero. So it's nearly impossible for to make an abnormal profit for everyone (Altin, 2015). When prices are connected with information then forecasting becomes difficult because of the random behavior of the market. Fabozzi et al. (2013) define Efficient market hypothesis as "Publicly available relevant information about the issuers will leads to correct pricing of freely traded securities in properly function markets".

2.2 Computational Learning Theory

Computational learning theory has the goal to understand the machine learning algorithm and determine what is learnable. For the learning task it will help us to determine how much data is sufficient for the training of a particular algorithm as well as the required resources. Association for computational learning defines computational learning theory is as "Learning Theory is a research field devoted to studying the design and analysis of machine learning algorithms. In particular, such algorithms aim at making accurate predictions or representations based on observations". Goldman (1999) defines the computational learning theory as a branch of theoretical computing which discuss the design of computer programs and their ability to lean and identify the computing limitation with machine. Computation learning theory helps to raise the questions and answer the question on the performance of learning algorithm

Forecasting is one of the most critical part of decision-making process (Wan et al., 2014). The term forecasting can be used as a process of creating forecast that will answer the questions *"what, where, how long and how large"* and explain them as well. This forecasting process must be reliable, flexible so can be improved as well as must be transparent and cost effective (Haloub, 2013; Kucharavy and De Guio, 2005). These forecasts play a rule of input for the decision making and organizations used them to plan and shape the future. There are a number of techniques that used to forecast and all of these techniques are design by taking a common goal of accuracy and unbiasedness.

A number of studies had proven the importance of forecasting in different fields of study and how practitioners use the forecasting techniques to forecast the future events. Studies like Li et al. (2015) on traffic management ,Armstrong et al. (2015) on sales forecasting,Mirakyan et al. (2017) forecast next day electricity price,Kikuchi et al. (2018) on wind field forecasting, Ehteram et al. (2019) on solar radiation forecasting and all of these studies highlighted the how importance of forecasting tools and technique in respective fields. Similarly, the finance field always showed a keen interest in forecasting topic and explored very successful prediction of financial time series (Sezer et al., 2020). Studies like Kodogiannis and Lolis (2002) use fuzzy based techniques to forecast the currency exchange rates, Kim (2003) predict the stock market returns of Korea composite stock price index, Al Wadia and Ismail (2011) predict Amman stock market returns, Ticknor (2013) capture the forecast behavior of daily prices of Microsoft Corp,Khashei and Hajirahimi (2017) work on financial markets forecast. Development of smart Financial forecasting models help to improves the accuracy of forecasting that ultimately helped the financial markets in risk management in more affective way (Niu et al., 2020). These aforementioned studies used multiple forecasting methods to forecast the respective time series and evaluate these models accordingly. this study further discus forecasting techniques more deeply in three subsections.

2.3 Classical Forecasting Techniques

Numerous of studies has been conducted to forecast the performance of stock markets with classical forecasting techniques. The idea behind was to achieve the best outcome using the minimum of input with least complex model. The prominent random walk theory was explained by Van Horne and Parker (1967), state that stock price move independently and does not follow any pattern. Further they mentioned random walk only applies on an efficient market. A number of studies has been performed by using random walk method. Earlier studies like Prescott and Stegnos (1989) on gold prices forecasting, White (1988) on IBM stock prediction using neural networks and random walk model like the study of Diebold and Nason (1990) on no non-parametric models with random walk for forecasting evaluation of forex rates. Ironically results show the random walk work batter then the non-parametric models. Tyree and Long (1995) also studies set an experiment on random walk model on financial time series data. Regardless of neural network model which uses the hidden layer didn't perform batter then the random walk model. Results show random walk model produce more accurate results then the neural network. Literature show a mix evidence of random walk models performance over other linear and nonlinear methods. Adhikari and Agrawal

(2014) Show that improved ANN based method outperform the forecasting ability of random walk in forex and stock market.

Initially exponential smoothing method used ad hoc technique to explore the univariate time series analysis. During 1960's an American economist lays down the statistical foundation of simple exponential smoothing to deal with forecasting of time series for a random walk and the noise purposes in data structure (Muth, 1960). Literature show a number of exponential smoothing model variations has been developed by from discontinuities to deal with one or more limitation in forecasting by (Rosas and Guerrero, 1994; Williams and Miller, 1999). Similarly, in early 2000's renormalization equation developed by Archibald and Koehler (2003) reports more accurate forecasting results. Lasek et al. (2016) highlight the various classical and ANN model for forecasting of restaurants sales and mentioned a survey in which exponential smoothing perform batter then the other econometric methods.

Famous study of Box and Jenkins (1970) revolutionize the timeseries forecasting by integrating the old classic ARMA model to deal with stationarity issue. Afterword hundreds and thousands of studies has been published that employ classic ARIMA model to forecast time series and found as best method to predict the time series. Other studies like Poulos et al. (1987) studied and also support the ARIMA model over other forecasting models available at that time. Hein and Spudeck (1988) use RW and ARIMA model to forecast the daily federal fund rate and found ARIMA model as a best predictor. Time series technique has the unique feature of flexibility and takes very few assumptions to perform and no priori probability is needed to analyze the data failure. Ho and Xie (1998) presented the mechanical system failure as example and tried to forecast using ARIMA method of forecasting to forecast the system failure. The study use traditional Duane model in comparison of ARIMA method and found ARIMA as a practical substitute of traditional methods with more accurate forecasting results.

Granger and Joyeux (1980) highlight the problem of infinite variance when the series does not need differencing. Deficiency of unit root tests sometimes couldn't differentiate a stationary series and a nonstationary series I (1) and mostly times series are first differenced. Sometimes due to long range dependence times series cannot be classified as I(1) and known as the long memory. This long memory process is known as the stationary process and ARFIMA Autocorrelation function decline slowly then the short memory of ARMA model. This long memory appears when integration parameter d of ARIMA models process the fraction and has a value ($0 < d$) allows the generalization of ARIMA model. Andersson (2000) highlight the importance of AR and MA term order which can vigorously affect the computed memory. ARFIMA (Auto Regressive fractionally integrated Moving Average) has been used in various studies to capture this issue. Seasonality creates the fraction and available ARIMA model cannot handle this issue. ARFIMA model with introduction of fraction can deal more sophisticatedly with the seasonality.

Studies like Mobarek and Keasey (2000) use ARIMA, AR and Random Walk (RW) techniques to capture the market efficiency of Dhaka stock exchange and found ARIMA model outperformed the AR and RW methods. Tularam and Saeed (2016) conducted a comparative study of three classic method of forecasting Exponential smoothing, Halt winter and ARIMA method to forecast the oil pricing. Results show that holt winter performed batter than the exponential smoothing. Whereas, ARIMA (2,1,2) outperform the both exponential smoothing and holt winter model. For policy makers and industry marketing experts, the suggest that the ARIMA to forecast the oil prices.

In real world, autocorrelation decay in time series is relatively slow and to handle this long range dependency Granger and Joyeux (1980) and Hosking (1981) come up with the idea to use fractional difference of series. ARFIMA model deals with the seasonality and long memory.Koop et al. (1997) used Bayesian techniques to use ARFIMA because criticism of impulse response over infinity. Other studies like Taylor (2000) and Sibbertsen (2004) also validate the long memory and ARFIMA model use. A study by Assaf (2006) found that the long memory and propose ARFIMA as a better method to address this issue over the ARMA model. Wang et al. (2009) conduct an empirical study on characteristics of univariate time series by using over 300 time series data sets and four different forecasting methods. Study consider the long rang dependency of data and founf ARFIMA as good method to study the long run dependence. The study used ARFIMA (0, d,0) method defined as fast and accurate and compute the hurst component of raw data using (d+0.5).

Franses and Ghijsels (1999) and Sabbatini and Linton (1998) found GARCH as an effective model to forecast the stock market returns because of its ability to deal the high and low volatility persistence. Gokcan (2000) examine the linear and nonlinear GARCH model for forecasting in seven emerging market data. Linear GARCH model perform batter then non-linear E-GARCH model regardless of substantially skewed return series. Conditional estimation of with in sample GARCH model proves a batter tool than the E-GARCH. Also, out of sample volatility estimation by GARCH show better results than the E-GARCH. Characteristics of high frequency data includes high kurtosis and lagged variance cluster. Due to outlier's presence in returns, higher kurtosis is common. To deal with this phenomenon Charles and Darné (2005) apply Additive outlier procedure to capture the outliers in daily data of three stock markets. After applying Additive outlier procedure GARCH model is used for forecasting.

Wei et al. (2010) report that GARCH based model are capable to handle the long memory and asymmetric volatility. Therefore, there are the better models to forecast especially volatility forecasting over long time horizon. Corrêa et al. (2016) apply GARCH, ARIMA-GARCH and newly introduced WARIMAX-GARCH in their study. According to the study exogenous variables are important in term of running any casual analysis. These exogenous variables also improve the forecasting ability. Wavelet based ARIMA-GARCH model outperform the traditional GARCH and ARIMA-GARCH forecasting ability. Kristjanpoller R and Michell V (2018) examine the performance of GARCH and ANN models using forex, commodity and interest rates data. Study apply different variants of GARCH method and choose best fitted on basis of MSE and MAPE. While describing the historical volatility effect, GARCH volatility is noted to be low with low external factors. Study conclude the econometric method perform batter with soft computing techniques. These combination works more efficiently for forecasting of financial time series.

2.4 Machine and Deep Learning Forecasting Techniques

Zhang et al. (1998) highlight the application of ANN in different field of study and explain the implementation of ANN classifier to predict the time series. The study found that ANN has the strong abilities to recognize and classify a pattern. Further they said, ANN being a nonlinear method rule out the ARIMA which is a linear statistical method of forecasting. Various studies integrate different clsscial nad machine learning techniques to generate more accuracy in forecasting. A study by Tealab (2018) perform a systematic literature review from year 2006-2016 on advancement of Artifical neural network methods for time series forecasting. The study found 17 studies that develop a new neural network to deal with phenomenon. Some of these studies are:Medeiros et al. (2006) integrated AR and ANN, Li et al. (2008) propose AR and GRNN which out perform traditional neural networks, Alavi and Gandomi (2011) integrate the ANN with simulated annealing(SA) to deal with earthquake, Wang et al. (2013) propose a Hybrid model of classical ARIMA forecasting model and ANN,Chen et al. (2015) used wavlet with SVR and Wang et al. (2016) propose a hybrid model with Elman recurrent neural networks and stochastic time effective function.

Tay and Cao (2001) use gaussian kernel based SVM with the objective of SVM fixability in financial time series and compare the results with a multi layers Back-propagation neural network technique. Results show that SVM outperform the other Back-propagation neural network because of risk minimizing principal that worked upper bound generalization error whereas back-propagation based on empirical risk minimizing principal. Another reason behind the success of SVM is its free parameters rather than the back-propagation parameters. That make back-propagation as a weaker method then the SVM.

Racine (2001) study the relation between the nonlinear and linear forecasting methods and finds that linear methods such as ARIMA couldn't capture the non-linear pattern of complex world problem. While nonparametric models based on neural network methods has more capability to capture the nonlinear pattern. Inherent characteristics of financial time series i.e., non-stationary, non-linear, noise, and chaotic dynamic system makes it a complex time series to handle.Tay and Cao (2002) modifies SVM model to deal with the non-stationary characteristic of financial time series based on idea of discounted least squares. They named their proposed model as C-ascending support vector machine(C-ASVM). Empirical testing of model is done on future contract and results shows that C-ASVM perform batter in simulated environment which shows the effectiveness of model in structural change of financial time series. This C-ASVM use less support vector then the typical SVM.Kim (2003) studies the stock market predictions and conclude that artificial intelligence-based methods perform better in forecasting of stock market. it further found that machine learning classifier SVM is promising technique to forecast as financial time series known as a challenging task for forecasting techniques.

Atsalakis and Valavanis (2009) conduct a survey based on 100 previous studies and finds that these studies use classic methods like ARIMA to ML methods like neural networks. Survey indicate that the soft computing techniques based methods (ML methods) outperform the old classic methods with better forecasting accuracy.Vaisla and Bhatt (2010) analyze the performance of stock market prices applied both Statistical and neural networks-based methods to forecast the daily stock market prices. Forecasting ability of neural networks is better than the statistically built methods because of the complexity in data structure. Further study suggest that neural networks are the better choice as alternative forecasting tool.Wang et al. (2017) propose a hybrid GARCH model based on ARIMA and SVM. Proposed hybrid model uses estimate the variance from data and then apply ARIMA and SVM to forecast. The study recommends that the proposed model has more reliable and accurate forecasting abilities.

Large data set creates the problems for the application of complex non linear methods that mostly creats the overfitting problem during the training of model.Busseti et al. (2012) use deep learning method to forecast the electricity demand and outperformed the other machone machine learning programs. Another study by Chen et al. (2015) perform deep learning based LSTM model on chinese stock market returns. They transform the data into 30 days sequence, learning rate has been set for 10. Results shows that accuracy achieved through the given parameters still can be improved. in short further suggest to use MACD and other features to deal with volatility stocks types and time window. Sagheer and Kotb (2019) examine the performance of proposed deep long-short term memory in oil production forecasting. Time series forecasting is a technical task in which traditional econometric models fails to perform due to complexity in both data and model. Above proposed method follows the same genetic of RNN architecture. Fair evaluation of two real oil field shows that proposed LSTM outperform the traditional and soft computing techniques.

Siami-Namini and Namin (2018) conducted an emerical study and found that Machine leaning baced LSTM methods outclass the traditional methods such as ARIMA. LSTM reduces the error rate around 84%-87% then the then the ARIMA. As well as the traing time numbers had also no significant affact on the trained forecast model. They perform a comprehensive study on classical forecasting method ARIMA and deep learning method LSTM using monthly data set of NIKKIE, HIS, IXIC, S&P 500 and DJI stock market indices. The emperical study shows that deep learning based LSTM method outperform the tradational economatric model ARIMA. Further they suggest that the number of "*epochs*" does not affect the model performance and it works randomly. Yan and Ouyang (2018) used LSTM to capture the non linearity, non stationary and sequencial correlation of time series. the Study has found that LSTM have show better prediction then the SVM. As well LSTM show better prediction and better dynamic forcasting in financial time series.

Lahmiri and Bekiros (2019) perform first empirical study on analysis of cryptocurrency using deep learning algorithm (LSTM) with comparison of GRNN (generalized regression neural network). Study provide that LSTM outperform the GRNN to predict the digital currency (bit coin). Because of LSTM ability to memorize the short and long information. Further study conclude LSTM as a better technique to find the fractal patterns.Merrill Eric (2019) state that forecasting algorithms with the characteristics of data richness and most advanced technological support improved the efficiency of the forecasting process and help finance professional to get rid of old repetitive activities. Further forecasting algorithms are based on statistical models that are developed by the gurus of both statistics and data science field supported by Machine learning tools. These tools and techniques help in data collecting, structuring, handling and storage in more speedily, efficiently and affordable.

Sezer et al. (2020) state that financial sectordevelop popularity machine and deep learning and develop model. Study underline that more then half of deep learning implications in financial sector are focused on the prediction of under lying assets. But "*Stock price forecasting, Index prediction, forex price prediction, commodity price prediction, bond price forecasting, volatility forecasting, cryptocurrency price forecasting*" are also investigated in literature. Li et al. (2020) combine the ARIMA model with machine learning methods like SVM and LSTM to improve the forecasting power. With hybrid methods, study finds an exceptional improvement in ARIMA model. ARIMA-LSTM possess better performance and likely to less complex model while computing. Livieris et al. (2020) Studies the gold price volatilities and said accurate price predition creates the opportunity to gain from the fluctuation of prices. Study combine the two different deep learning methods. First, Convolutional neural networks with benefits of internal representation and ability of extracction of valuable information. Second LSTM which has long and short memory advantage. CNN-LSTM results shows a significant improvement in forecasting ability of LSTM method.

2.5 Time Frequency and Forecasting

Numerous studies highlight the imoratance and disadvantages of time frequency in data analysi. Zhou (1996) discuss the importance of data frequency in his study. By comparing the real time difference, author conclude that low frequency and high frequency data can be only differentiate at level of noise existence. In low frequency, noise does not seems a significant impact whereas, in high frequency data noise become very significant. Any sort of data validation can remove the bad data from data set but not able to remove the noise. Further is found that negative autocorrelation also increases in forex as we increase the data frequency. Liu (2009) examine the performance of portfolio using low frequency(daily) and High frequency(intraday) data for covariance estimation. Empirical results shows that the performance of a portfolio depends on the restructure of portfolio more than the frequency. If a manager rebalances his/her portfolio monthly using yearly data estimation windows than daily data frequency is more reliable then the intraday. However, if manager rebalance the portfolio on daily basis then intraday data is more useful. Matias and Reboredo (2012) examine forecasting techniques on different time intervals in intraday data. The findings of the study shows that linear method AR is only surpassed by the supervised machine learning techniques SVM and KR in 5 minutes and 10 minutes data frequency. In medium (30 minutes) and long (60 minutes) time horizon STAR-,GARCH and MLP respectively outperform the other methods. Another study by Rounaghi and Nassir Zadeh (2016) uses an intra-time frequency to compare the monthly and yearly forecast results produced by using ARMA forecasting technique. Study used London stock market index and S&P 500 market index to check the long memory. The finding of the study shows monthly data exhibit is more volatile and outperform the yearly data.

2.6 Research Hypothesis

H1: Classical forecasting methods outperform Machine learning and Deep learning methods in forecasting price behavior in financial markets.

H2: Machine learning methods outperform Classical and Deep learning forecasting methods in forecasting price behavior in financial markets.

H3: Deep learning methods outperform Classical and machine learning methods in forecasting price behavior in financial markets.

H4: Hybrid GARCH-LSTM Outperform the Classical, Machine leaning and Deep learning in forecasting price behavior in financial markets.

H5: Data frequency significantly affect the forecasting of price behavior in financial markets.

Chapter 3

Research Methodology

This Chapter of the study discusses a brief detail about the sample and empirical methodology to achieve the object of study. Therefore, the first section of this chapter discusses the population and sample for an empirical test of forecasting models. The second section of the chapter debate the empirical methodology of selected classical, machine learning and deep learning methods. In the third section forecasting, evolution indicators are presented. Last but not the least, Environment setting has been discussed in 4th section of this study. Which discribes the used analytical programs and libraries for analysis purpose.

3.1 Population and Sample of Study

This study revolves around the global stock markets to find out the best model to forecast and the appropriate frequency of data to forecast the market. To achieve the objectives, this studies applies different methods and time frequencies of data to produce a generic results applies on all sort of stock markets. For this purpose, study examined 16 stock indices given in table 3.1. Selection of these markets are with subject to the availability of high frequency data. As for the frequency, Andersen et al. (2003) and many other authors believe that high frequency data hold more information. Based on objective obtained data follows three frequencies: hourly data, 10 minutes data and 5 minutes data interval. The study used actual

Table 3.1: Stock Market Indices

Index	Origin of index	Market	Name of Index
ASX	Asia / Pacific	Australian Securities Exchange	S&P/ASX200
A50	Asia / Pacific	Shanghai Stock Exchange	FTSE China A50 Index
EUS	Europe	European stocks	EURO STOXX 50
CAC	Europe	French stock market	CAC 40
DAX	Europe	Frankfurt Stock Exchange	DAX 30
HSI	Asia / Pacific	Hong Kong stock market	Hang Seng 40 Index
NIFTY	Asia / Pacific	National Stock Exchange India	NIFTY 50
NIKKEI	Asia / Pacific	Tokyo Stock Exchange	NIKKEI 225
AEX	Europe	Euronext Amsterdam	AEX 25
WIG	Europe	Warsaw Stock Exchange	WIG 20
SSGF	Asia / Pacific	Singapore stock market	Singapore blue chips cash index
IBX	Europe	Bolsa de Madrid	IBX 35
SMI	Europe	Swiss Market Index	SMI 20
FTSE100	Europe	London Stock Exchange	FTSE100
DJI30	America	United States	DJI 30
SNP500	America	United States	S & P 500

prices data for models training and testing. All the data has been obtained from the data base of "Dukascopy" a Swiss bank for the past three years starts from 02-07-2017 till 30-06-2020.

3.2 Econometric and Soft Computing Methods

This study is based on three major classes of forecasting tools known as the classical methods (i.e. ARIMA, ARFIMA and GARCH (p,q)), similarly advance machine learning and Deep learning methods (i.e., SVM and LSTM) and some hybrid tools which are the combination of classical and machine learning and Deep learning methods (i.e. ARIMA-SVM, ARIMA-LSTM and GARCH-LSTM). For hybrid models studies like Kristjanpoller and Hernández (2017); Li et al. (2020) use three stage machanism to compute. In first stage residual series based on ARIMA and GARCH model has been generated. In second stage residual series from SVM and LSTM model has been generated. In third stage resildus from clsscial, machine learning and deep learning has been combined with orignal data series to generate a hybrid series to estimate ARIMA-SVM, ARIMA-LSTM, and GARCH-LSTM. Econometric models for classic, machince learning and deep learning are presented below.

3.3 Classical Methods

For the classical forecasting methods, this study uses the ARIMA, ARFIMA, and GARCH models to examine the forecasting ability of these classical methods. The empirical methodology of these models are presented below.

3.3.1 Autoregressive Integrated Moving Average

Auto regressive integrated moving average known as ARIMA (p,d,q) is one of the important method in classical forecasting techniques. This method is the combination of AR and MA term and where d is the difference term and p and q are the ARMA delay parameters.

$$y_t = \varphi_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + ... + \varphi_p y_{t-p} + \in_t - \theta_1 \in_{t-1} + \theta_2 \in_{t-2} + ... + \theta_q \in_{t-q} \quad (3.1)$$

Where

$$\begin{cases} \varphi_i \neq 0, \theta_q \neq 0 \\ E(\in_t) = 0, \mathrm{var}(\in_t) = \sigma_u^2 \\ E(\in_t \in_S) = 0, s \neq t \\ E(y_s \in_t) = 0, \vee_s < t \end{cases}$$

Y_t is the representation of independent variable at time t , y_{t-1} and y_{t-1} are the lagged dependent; represent the white noise and is the self-regressive parameter. As times series are not the stationary sequence (Visser, 2011). We need to convert it to stationary form through d-order difference. This ARIMA(p,d,q) model is given below.

$$Y_t * = (1 - B)^d Y_t \quad (3.2)$$

3.3.2 Autoregressive Fractionally Integrated Moving Average

Autoregressive fractionally integrated moving average(ARFIMA) are the methods to deal with the short memory issues. ARFIMA used to establish parameters to deal with long term correlation in data. A general expression of ARFIMA is written as:

$$\phi\,(B)\,y_t = \Theta(B)(1-B)^{-d}\varepsilon_t \tag{3.3}$$

Where $\phi(B) = 1\phi_1 B - \cdots - \phi_p and \Theta(B) = 1 + \theta B + \cdots + \theta_q B^q$ are AR and MA operators, while $\phi(B)$ and $\Theta(B)$ have no common roots. B as a backshift and $(1-B)^{-d}$ is the fractional term.

3.3.3 Generalized Autoregressive Conditional Heteroskedasticity

Generalized Auto regressive conditional Heteroscedasticity (GARCH) belongs to ARCH family and known for the calculation of variance (volatility) in time series. A general expression of GARCH variance equation is as:

$$R_t = \alpha_o + \alpha_1 R_{t-1} + \mu_t, \mu_t \sim N\{0, \sigma_t^2\} \tag{3.4}$$

$$\sigma_t^2 = \alpha_o + \alpha_1 \mu_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{3.5}$$

Whereas,R_t referred as lag returns of respective series, σ_t^2 referred as the conditional variance μ_{t-1}^2 referred as squared unexpected returns at time: $t-1$.

3.4 Machine and Deep Learning Methods

For the machine and deep learning forecasting methods, this study uses the SVM, LSTM, and hybrid LSTM models to examine the forecasting ability of these machine and deep learning methods. The empirical methodology of these models are presented below.

3.4.1 Support Vector Machine

Support Vector MAchne (SVM) a generalized linear classifier method used to classify the infinite observations. A sample training start from $\{(k_i y_i), i = 1, 2, 3, \ldots, I\}$ so regression model will be $f(k) = (\psi \times \vartheta(k)) + b$ so corresponding optimization will be:

$$min \left(\frac{1}{2} \parallel \psi \|^2 + P \sum_{i=1}^{l} (\wp_i + \wp_i{}^*) \right) \tag{3.6}$$

$$\left\{ \begin{array}{l} y_i - (\psi * \vartheta(k_i)) - b \leq \varepsilon + \wp_i, \\ (\psi * \vartheta(k_i)) + b - y_i \leq \varepsilon + \wp_i{}^*, \\ \wp_i, \wp_i{}^* \geq 0, \\ i = 1, 2, ..., I \end{array} \right\} \tag{3.7}$$

Whereas$\wp_i$ and $\wp_i^*$ are the slack variables added to for optimization problem, P refers as penalty function. The whole program is converted in dual problem as

$$\max \Bigg(-\frac{1}{2} \sum_{i,j=1}^{l} (q_i - q_i{}^*)(q_l - q_l{}^*) \left(\wp(k_i) * \wp(k_l)\right) + \sum_{i=1}^{I} y_i (q_i - q_i{}^*) - \varepsilon \sum_{i=1}^{I} (q_i + q_i{}^*) \Bigg) \tag{3.8}$$

$$\left\{ \begin{array}{l} \sum_{i=1}^{I} (q_i + q_i{}^*) = 0 \\ q_i, q_i{}^* \in [0, P] \\ i = 1, 2, 3, ..., I \end{array} \right\} \tag{3.9}$$

This defines the kernel function of inward product of high dimensional feature as $R(k_i, k) = (\wp(k_i), \wp(k))$, and a quadratic programing will be as:

$$Quadraticmapping f(k) = \sum_{i=1}^{I} (q_i - q_i^*)(\wp(k_i), \wp(k)) \tag{3.10}$$

3.4.2 Long Short-Term Memory

Long- and short-term memory (LSTM) networks are the machine learning algorithm used to process and forecast important events of long intervals in time series. Based on RNN, LSTM add a processor to judge whether the information is useful or not. Works on three important gates to judge the information mainly the forget gate j, input gate p and output gate k, so written expression will be as follow:

$$Q_t = f_t Q_{t-1} + j_t \tanh(\omega_{xQ} e_t + \omega_{hQ} h_{t-1} + a_Q) \tag{3.11}$$

Whereas: Q_t is the cell state at t. f_t is present forget gate, $Q_t - 1$ is previous cell state at time $t-1$, is present input gate, x_t present the token for input sentence, $h_t - 1$ are hidden layer, $\omega_x y$ as weight between x to y and a_Q refers as offset amount. Forget gate apply the neglected information, that remain input and hidden layer determine the future of information. The new expression will be as:

$$f_t = \sigma(\omega_{xf} x_t + \omega_{hf} h_{t-1} + \omega_{Qf} Q_{t-1} + a_f) \tag{3.12}$$

Whereas f_t refers as input gate, $h_t - 1$ refer as hidden layer; Input gate J_t applies the updated information as:

$$j_t = \Omega(\omega_{xj} x_t + \omega_{hj} h_{t-1} + \omega_{Qj} Q_{t-1} + aj) \tag{3.13}$$

At the end output gate k_t applies the updated information as:

$$k_t = \Omega(\omega_{xk} x_t + \omega_{ho} h_{t-1} + \omega_{Qk} Q_{t-1} + a_k) \tag{3.14}$$

Whereas hidden layer output will be:

$$h_t = k_t \tanh(Q_j) \tag{3.15}$$

3.5 Evaluation Criteria

In order to evaluate following of the parameters will be used. whereas; R_i is the forecasted value and P_i is the actual value,

Mean Squared Error (MSE):

$$MSE = \frac{1}{n} \sum (R_i - P_i)^2 \tag{3.16}$$

Mean Absolute Error (MAE):

$$MAE = \frac{1}{n} \sum |R_i - P_i| \tag{3.17}$$

Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{MSE} \tag{3.18}$$

Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{100}{n} \sum \frac{||R_i - P_i||}{p_i} \tag{3.19}$$

3.6 Setting of Environments

Python has been used to as programing language to set environments and Jupiter note book as IDE for python. Following packages has been used: Numpy for mathematical functions, Matplotlib for plotting of results, Pandas for Data manipulation and analysis, Sklearn for stat models, Keras and Tensorflow for Machine learning operations, stats models, Pmdarima for auto arima function and arch library for garch in python. R Studio is also been used and rugarch package has been used to compute the GARCH and ARFIMA model estimation and forecasting.

Chapter 4

Results and Discussion

This chapter has been divided into 4 sections. First three sections represent the models estimation, forecasting and comparison of each frequency and section four represent the comparison between the frequencies.

4.1 Section I (Hourly Data Frequency)

In this Section, all of the classical, machine learning, and deep learning methods are examined using hourly data frequency. This section of the study is further divided into five sub-sections. Section one and two deal with the data descriptives and pre-processing. Section two and three discuss the estimated results of all models and the section discuss the models comparison.

4.1.1 Descriptive Statistics

Table 4.1 depicts the data summary for 16 indices this study has used to analyze the aforementioned objectives. Mean, median, standard deviation, skewness, kurtosis and jarque-Bera statistics. Mean values shows the average indices points for the selected time period. ASX, A50, EUS, CAC, DAX, HIS, NIFTY, NIKKEI, AEX, WIG, SSGF, IBX, SMI, FTSE 100, DJI30 and SNP500 has the mean values of 6124.38, 12724.5, 3399.07, 5305.97, 12198.4, 27722.3, 10931.2, 21720.1, 549.703,

Table 4.1: Descriptive Statistics (Hourly Data)

	Mean	**Median**	**Std. Dev.**	**Skew***	**Kurt***	**J.B***
ASX	6124.38	6071.47	442.23	0.07	2.88	24.03
A50	12724.50	12875.00	1063.55	-0.34	2.07	846.60
EUS	3399.07	3443.67	244.82	-1.11	4.72	3569.13
CAC	5305.97	5358.27	393.22	-0.75	4.38	1894.65
DAX	12198.40	12313.20	886.67	-1.03	4.60	5075.15
HSI	27722.30	27791.00	2088.70	-0.08	2.84	23.89
NIFTY	10931.20	10919.50	860.86	-0.82	3.97	1817.13
NIKKEI	21720.10	21801.00	1345.69	-0.56	3.39	1064.20
AEX	549.70	552.34	34.03	-0.61	4.79	1830.47
WIG	2184.95	2238.30	236.84	-1.53	5.27	3624.63
SSGF	361.79	364.94	30.34	-1.06	4.09	3352.84
IBX	9302.79	9411.28	943.24	-1.40	4.93	4518.79
SMI	9400.88	9281.97	620.47	0.68	2.90	837.93
FTSE*	7190.13	7331.47	518.58	-1.94	6.64	20314.18
DJI*	25204.30	25309.10	1886.29	-0.20	3.04	126.44
SNP*	2806.90	2793.97	220.71	0.35	2.82	381.60

**Skew = Skewness, *Kurt = kurtosis, *J.B = Jargue-Bera,*FTSE = FTSE100, *DJI = DJI30, *SNP = SNP500*

2184.95, 361.788, 9302.79 ,9400.88 ,7190.13, 25204.3, 2806.9 respectively. Similarly, standard deviations explain the data desperation from its mean. To measure the location of data we generally interpret the skewness and kurtosis figures. Skewness explain the left are right skewed of data observations. In our data set A50, EUS, CAC, DAX, HIS, NIFTY, NIKKEI, AEX, WIG20, SSGF30, IBX30, FTSE 100, and DJI 30 has negative values (-0.33776 for SSE) which referred as left skewed data. Whereas ASX, SMI and S&P 500 are right tailed with positive skewness values (0.0704, 0.6780, 0.3451) respectively. ASX, SSE, HIS, SMI, DJI30, S&P500 has the platykurtic Behavior and rest of sample has the leptokurtic behavior having kurtosis values greater than 3. Jarque-Bera normality test indicates the none of variable is normally distributed.

4.1.2 Data Pre-Processing

Data diagnostic is the primary part of data analysis. Diagnostics of data set before estimation of any model either based on linear regression or generalized linear determines the model fitting on that specific data set. It is encouraging to have a data look before applying any sort of statistical model Granger and Joyeux (1980). Time series data is not normally distributed mostly. Before proceed any of statistical method stationary of data check is essential to be establish. Horváth et al. (2014). Similarly, for the estimation of seasonality and trend with in data is also needed to be check. Augmented Dickey fuller (ADF) test has been applied on data to check the stationarity and seasonality in our data. It has been found that data is not normally distributed. P-value of ADF test for all indices reported in table 4.2 Data Processing (Hourly Data). Figure Appendix A-1 in appendix shows the lag correlation results with three lags.

TABLE 4.2: Data Processing (Hourly Data)

	Stationarity			**ARCH Lm Test**
	Level	**1st Diff.**	**2nd Diff.**	**p-value**
ASX	0.1707	0.0000	-	0.0000
SSE	0.1845	0.0000	-	0.0000
EUS	0.1593	0.0000	-	0.0000
CAC	0.1929	0.0000	-	0.0000
DAX	0.1495	0.0000	-	0.0000
HSI	0.2806	0.0000	-	0.0000
NIFTY	0.2496	0.0000	-	0.0000
NIKKEI	0.1524	0.0000	-	0.0000
AEX	0.0928	0.0000	-	0.0000
WIG	0.7332	0.0000	-	0.0000
SSGF	0.7327	0.0000	-	0.0000
IBX	0.7221	0.0000	-	0.0000
SMI	0.1285	0.0000	-	0.0000
FTSE	0.4683	0.0000	-	0.0000
DJI	0.0561	0.0000	-	0.0000
SNP	0.1406	0.0000	-	0.0000

Before setting up the python environment and start computing machine learning techniques it also prerequisite to preprocess the data as well. For both SVM and LSTM methods, it is compulsory to scale the data set. Otherwise, unscaled data create problems for model fitness. Existence of Large ranges in data will probably reduce the learning speed, convergence of networks and also reduce the accuracy of model (Provost and Aronis, 1996). So before proceeding data sets has been scaled using minimax scaler. Which transform our data in range from 0 to 1.

4.1.3 Classical Forecasting Methods

This analysis section has been further divided into two sessions: Model Estimation and selection, and Forecasting.

4.1.3.1 Models Estimation and Selection

After the preprocessing of data model estimation and selection is one of the primary tasks in classical forecasting methods. Auto ARIMA function has been used where maximum p order has been set as 36, similarly maximum q order has been set as 36. For seasonality check seasonal has been set true, Maximum SAR and SMA order is set as 36 as well. Stepwise model has also been set true to compute best fitted ARIMA model on given stock indices. For ARFIMA selected p and q order with the term of d is used. Both of models has been computed using log likelihood method. Based on ARCH LM test results on table 4.2, standard GARCH (1,1) model has been used as our third model of conventional model. Table 4.3 depicts the ARIMA model estimation statistics and table 4.4 explains the ARFIMA and GARCH estimation. For the estimation 70% data observations has been used as training window and 30% data observations has been used as test window.

Results presented in table 4.3 exhibit existence of random walk of prices in majority of series. ARIMA order (0,1,0) has been found in EUS, CAC, DAX, HIS, NIKKEI, AEX, WIG, IBX, SMI, SNP500 stock index based on AIC values of 52654.627, 58893.726, 114468.8, 94201.047, 132215.095, 20621.113, 29329.816, 60471.088, 65612.922, and 75173.838 along with BIC values 52661.57, 58900.668, 114476.242, 94208.054, 132222.534, 20627.901, 29336.16, 60477.874, 65619.86, and 75181.277 respectively for all given stock indices. Results clearly states no AR and MA term for respective stock indices. Similarly, for seasonality stepwise model shows no SAR and SMA term. Table also present the variance coefficient values 56.9472, 129.9876, 526.102, 6064.7398, 2166.8422, 1.3609, 62.4884, 604.1174, 323.4789, and 23.1811 for given stock market indices under hourly data frequency. KDE plot of residuals for estimated ARIMA model also has been presented in the appendix.

Table 4.3: ARIMA Estimation

	ASX	**A50**	**EUS**	**CAC**	**DAX**	**HSI**	**NIFTY**	**NIKKEI**	**AEX**	**WIG**	**SSGF**	**IBX**	**SMI**	**FTSE**	**DJI**	**SNP**
Method								Log Likelihood								
Model	(1,1,1)	(5,1,0)	(0,1,0)	(0,1,0)	(0,1,0)	(0,1,0)	(1,1,0)	(0,1,0)	(0,1,0)	(0,1,0)	(1,1,1)	(0,1,0)	(0,1,0)	(0,1,2)	(2,1,0)	(0,1,0)
AIC	86794	108618	52655	58894	114469	94201	77183	132215	20621	29330	21861	60471	65613	92929	131980	75174
BIC	86816	108662	52662	58901	114476	94208	77197	132223	20628	29336	21882	60478	65620	92952	132003	75181
AR	0.776	0.014	-	-	-	-	-0.03	-	-	-	0.747	-	-	-	-0.023	-
	[0.000]	[0.029]	-	-	-	-	[0.000]	-	-	-	[0.000]	-	-	-	[0.000]	-
MA	-0.805	-	-	-	-	-	-	-	-	-	-0.768	-	-	-0.02	-	-
	[0.000]	-	-	-	-	-	-	-	-	-	[0.000]	-	-	[0.01]	-	-
SAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SMA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sigma2	81.0	1371.3	56.9	130.0	526.1	6064.7	526.1	2166.8	1.4	62.5	0.5	604.1	323.5	132.9	2128.3	23.2
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Obs	12002	10795	7654	7644	12575	8158	8479	12570	6555	4207	9892	6544	7615	12026	12569	12569

**Obs = Observations, *p < 0.05*

Estimated results for A50 stock market index shows the ARIMA order (5,1,0) as the best-fitted model. Similarly, NIFTY and DJI30 shows ARIMA(1,10 and 2,1,0 respectively) as best fitted model known as differenced first order autoregressive model (Nau, 2014). All of these models have been selected on basis of lowest AIC values (108617.895, 77182.645, and 131980.25) and BIC values (108661.616, 77196.735, and 132002.566) respectively for A50, Nifty and DJI30 stock market Indices. AR term have a coefficient values 0.0138, -0.0292, and -0.0232 with significant statistical p-value. This means the existence of a pattern in the all these market indices with lag values. Estimated results do not show any of MA terms which means any delayed information adjustment. No SAR and SMA have been identified which refers to no seasonality in data. KDE plot of residuals shown in figure A-1 in Appendix-A validates smoothen the data after applying the model.

For FTSE market index, based on AIC and BIC values (92929.317and 92951.501) respectively, ARIMA (0,1,2) model has been identified as best fitted model for FTSE market index. Also referred as Damped Holt's model. Figure A-2 in appendix indicates the model fitness which smoothen the KDE of residuals. Further MA term of model has the coefficient value of -0.0172 with significant p-value (0.01). This shows the delayed adjustment of information and overly price information as well. the model showed no AR, SAR and SMA terms which indicates the no auto regressive and seasonality pattern exists. Further Sigma2 has the coefficient values of 132.9034 with significant p-values indicates the existence of variance in mean equation of model.

For ASX and SSGF market indices on basis of their AIC values (86794.037 and 21860.805 respectively) and BIC values (86816.216 and 21882.403 respectively) ARIMA (1,1,1) model has been identified as the best-fitted model for both indices. AR term for both models has a significant p-value with coefficient values of 0.7763 and 0.7469. This means the existence of patterns in financial series and significant relation with own lag values. MA term also have significant p-values with coefficient values of -0.8047 and -0.7684 respectively. The negative coefficient value shows the overly priced informational adjustment. Whereas, Stepwise models shows no SAR and SMA seasonality patterns in hourly minutes of data frequency in both FTSE and DJI30 index. KDE plot of residuals shows the model fitness. Significant values of Sigma2 with coefficient values of 22.528 and 359.091 for both markets show the presence of variance in markets.

table 4.4 presents the ARFIMA models specifications for all the selected stock market indices. For the reference of p and q order of model predefined AR and MA order given in ARIMA estimation has been used as p and q order for ARFIMA model. For the stock market indices EUS, CAC, DAX, HIS, NIKKEI, AEX, WIG, IBX, SMI, and SNP500 have computed AIC values -9.35, -9.41, -9.68, -8.96, -9.44, -7.44, -8.48, -9.09, -9.59, and -9.81 respectively for all stock markets. Estimated d parameter values -0.0149, -0.0065, -0.0082, 0.0012, 0.0003, -0.0394, -0.0147, -0.0206, -0.0113, and -0.0092 respectively for all aforementioned markets. EUS, AEX, and IBX are only markets that's follows the intermediate memory (StataCorp, 2019). Other remaining stock market indices have insignificant p-values and on basis of this, we conclude that ARFIMA does not fit for these markets.

For stock market indices A50, NIFTY, and DJI30, based on previously identified p and q computed ARFIMA model has the AIC values -8.78, 9.48, and -9.72 respectively. The estimated d parameter has the values of 0.0114, 0.0153, and 0.0118. Whereas, A50 is the only stock market with statistically significant p-value. Estimated d parameter shows an intermediate memory in autocovariance function of series exists (StataCorp, 2019). AR term has the coefficient value of-0.0179 with a significant p-value of (0.003) which indicates the persistence of autocorrelation with lag values and a pattern exists in A50 stock market. For ASX and SSGF market indices AIC values -10.15 and -9.61 respectively mentioned in table 4.4. Estimated d parameter values are 0.0262 and 0. The insignificant results for ASX and SSGF market indices clearly shows the miss fit for ARFIMA for respective stock markets under hourly data frequency. ARFIMA model for FTSE also has same p, d and q order identified in ARIMA model. Computed Model has the AIC value of -10.04. Coefficient of ARFIMA term d (See table 4.4) has the value of -0.020 with p-value of (0.0001). FTSE with significant coefficient p-value shows the intermediate memory presence.

On the presence of heteroscedasticity condition GARCH(1,1) model also been computed on all given stock market indices represented in table 4.4. For the non-parametric model in the classical forecasting method GARCH (1,1) has been estimated for hourly time interval data. ASX, A50, EUC, CAC, DAX, HSI, NIFTY, and NIKKEI market indices have the lag coefficient values of -0.024, 0.001, -0.019, -0.011, -0.015, 0.008, -0.049, and -0.014 with significant p-values for ASX, EUS, and NIFTY. Results indicate the inverse relation with lag returns of these stock markets. Remaining stock market indices has insignificant relation with its lag returns. α_1 have a coefficient values of 0.013, 0.135, 0.032, 0.027, 0.379, 0.126, 0.181, and 0.025 for aforementioned stock indices respectively with significant p-value. Similarly, β_1 coefficients values 0.985, 0.508, 0.966, 0.969, 0.377, 0.517, 0.449, and 0.970 with significant p-value. These statistical values show the existence of the ARCH effect and persistence of ARCH effect as well as the present price behavior is the result of change in past prices which significantly transmitted into current prices behavior.

For stock indices AEX, WIG, SSGF, IBX, SMI, FTSE, DJI30, and SNP500 estimated for the GARCH (1,1) model has been presented in table 4.4. Estimated AIC values have been presented as well. Aforementioned market indices have the lag coefficient values of -0.007, -0.006, -0.052, -0.008, -0.021, -0.024, 0.008, and 0.006. SSGF, SMI, and FTSE are the only stock indices in the list which has significant inverse relationship between the lag returns and prices of these stock market indices. Coefficient values of α_1 indicate the existence of the ARCH effect in markets. Coefficient values 0.022, 0.071, 0.187, 0.022, 0.028, 0.019, 0.030 , and 0.039 with significant p-values support our argument of volatility existence. Similarly,β_1 coefficients show the persistence of GARCH in data. Coefficient values of GARCH 0.973, 0.286, 0.158, 0.965, 0.969, 0.978, 0.970, and 0.961 with significant p-value validate the argument of GARCH persistence in given data set.

Table 4.4: ARFIMA and GARCH Estimation

	ASX	**A50**	**EUS**	**CAC**	**DAX**	**HSI**	**NIFTY**	**NIKKEI**	**AEX**	**WIG**	**SSGF**	**IBX**	**SMI**	**FTSE**	**DJI**	**SNP**
								AFRIMA Statistics								
Method								Log Likelihood								
Model	(1,1,1)	(5,1,0)	(0,1,0)	(0,1,0)	(0,1,0)	(0,1,0)	(1,1,0)	(0,1,0)	(0,1,0)	(0,1,0)	(1,1,1)	(0,1,0)	(0,1,0)	(0,1,2)	(2,1,0)	(0,1,0)
AIC	-10.15	-8.78	-9.35	-9.41	-9.68	-8.96	-9.48	-9.44	-7.44	-8.48	-9.61	-9.09	-9.59	-10.04	-9.72	-9.81
d	0.026	0.011	-0.015	-0.007	-0.008	0.001	0.015	0.000	-0.039	-0.015	0.004	-0.021	-0.011	-0.020	0.012	-0.009
	[0.138]	[0.007]	[0.022]	[0.305]	[0.076]	[0.856]	[0.128]	[0.945]	[0.000]	[0.165]	[0.873]	[0.004]	[0.081]	[0.001]	[0.112]	[0.159]
AR	0.738	0.018	-	-	-	-	-0.049	-	-	-	0.723	-	-	-	-0.030	-
	[0.00]	[0.00]	-	-	-	-	[0.00]	-	-	-	[0.00]	-	-	-	[0.00]	-
MA	-0.79	-	-	-	-	-	-	-	-	-	-0.75	-	-	-0.01	-	-
	[0.00]	-	-	-	-	-	-	-	-	-	[0.00]	-	-	[0.00]	-	-
								GARCH Statistics								
Method								Maximum Likely hood ARCH								
AIC	-12258	4866	-2158	-2735	-8045	1385	-2864	-5818	176	3008	-4500	432	-3832	-11891	-12583	-13969
Lag	-0.024	0.001	-0.019	-0.011	-0.015	0.008	-0.049	-0.014	-0.007	-0.006	-0.052	-0.008	-0.021	-0.024	0.008	0.006
	[0.003]	[0.958]	[0.028]	[0.266]	[0.148]	[0.574]	[0.000]	[0.090]	[0.566]	[0.740]	[0.000]	[0.50]	[0.023]	[0.001]	[0.256]	[0.429]
α_1	0.013	0.135	0.032	0.027	0.379	0.126	0.181	0.025	0.022	0.071	0.187	0.022	0.028	0.019	0.030	0.039
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
β_1	0.985	0.508	0.966	0.969	0.377	0.517	0.449	0.970	0.973	0.286	0.158	0.965	0.969	0.978	0.970	0.961
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
OBS	12002	10795	7654	7644	12575	8158	8479	12570	6555	4207	9892	6544	7615	12026	12569	12569

4.1.3.2 Forecasting of Models

Figure A-3 below exhibit the forecasting of estimated trained and test ARIMA model for ASX stock market index. Forecasting accuracy indicators presented in table 4.5, MSE (518320.9447), RMSE (9.979912), MAE (497.0621), and MAPE (8.8227). Results shows the sudden drop in forecasted windows due failed the ARIMA forecasting method to predict in such case as arrival of Covid-19 in this time frame. Figure A-4 exhibit the both actual and forecasted ARFIMA series with Estimated MSE (683250.04), RMSE (826.59), MAE (549.43), and MAPE (9.84). For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in figure A-5. Forecasting Accuracy indicators are MSE (571.61), RMSE (23.91), MAE (13.12), and MAPE (0.23).

Figure A-3 in appendix exhibit the forecast of estimated trained and test ARIMA model for A50 market index. Forecasting accuracy indicators presented in table 4.5 are MSE (613288.0788), MAE (685.9897), MAPE (4.9699), and RMSE (783.1271). Results shows the train and test data set with in defined range upper and lower bound that shows the perfection of ARIMA for A50 market index. ARFIMA estimated accuracy indicators presented in Table 4.5 are, MSE (978986.57), MAE (794.07), MAPE (6.37), and RMSE (989.44). Figure A-4 in appendix exhibit the ARFIMA forecasted and actual series for A50 stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are Estimated MSE (977480.63), MAE (817.08), MAPE (6.13), and RMSE (988.68) for A50 market index and presented in table 4.5.

Figure A-3 in appendix exhibit the forecasted results for estimated trainning window and testing window of ARIMA model for A50 market index. Forecasting accuracy indicators presented in table 4.5 are MSE (613288.0788), MAE (685.9897), MAPE (4.9699), and RMSE (783.1271). Results shows the train and test data set with in the defined range upper and lower bound that shows the perfection of ARIMA for A50 market index.

Table 4.5: Classical Forecasting Methods (Hourly Data)

	ASX	**A50**	**EUS**	**CAC**	**DAX**	**HSI**	**NIFTY**	**NIKKEI**	**AEX**	**WIG**	**SSGF**	**IBX**	**SMI**	**FTSE**	**DJI**	**SNP**
							Forecasting indicators for ARIMA									
MSE	518321	613288	147440	419392	1777831	5780628	2259523	4601531	2825	40367	1942	743262	356021	714999	5355411	29027
MAE	497.06	685.99	335.61	581.68	1149.43	1918.54	1169.65	1848.69	39.20	128.19	18.56	657.01	455.58	584.97	1897.17	130.66
MAPE	8.82	4.97	10.27	11.27	9.66	7.79	12.11	8.38	7.61	6.82	5.53	7.17	4.55	9.73	7.41	4.63
RMSE	719.95	783.13	383.98	647.60	1333.35	2404.29	1503.17	2145.12	53.15	200.92	44.07	862.13	596.68	845.58	2314.18	170.37
							Forecasting indicators for ARFIMA									
MSE	683250	978987	144882	424867	1793454	10825561	3714293	4653055	2297	110054	1440	1096498	382891	690989	6837399	57306
MAE	549.43	794.07	331.82	587.02	1155.66	2816.96	1340.41	1866.76	36.93	228.48	29.50	905.33	469.89	578.86	2048.00	189.94
MAPE	9.84	6.37	10.06	11.39	9.63	9.79	14.33	8.47	7.00	13.98	9.63	11.43	4.80	9.61	8.31	6.44
RMSE	826.59	989.44	380.63	651.82	1339.20	3290.22	1927.25	2157.09	47.93	331.74	37.95	1047.14	618.78	831.26	2614.84	239.39
							Forecasting indicators for GARCH									
MSE	973337	977481	210077	767057	2274734	26170970	4822289	5145657	7635	105208	1752	1415933	652481	1322141	20604064	115823
MAE	626.73	817.08	373.28	678.66	1179.00	4159.28	1509.69	1839.47	58.49	222.70	31.26	954.25	566.59	757.71	2898.19	234.60
MAPE	11.35	6.13	12.05	14.18	10.57	15.67	16.17	8.88	11.62	13.64	31.26	12.56	5.92	12.76	12.21	8.31
RMSE	986.58	988.68	458.34	875.82	1508.22	5115.76	2195.97	2268.40	87.38	324.36	41.86	1189.93	807.76	1149.84	4539.17	340.33

ARFIMA estimated accuracy indicators presented in Table 4.5 are MSE (978986), MAE (794.07), MAPE (6.37), and RMSE (989.44) for A50 stock market index. Figure A-4 in appendix exhibit the ARFIMA forecasted and actual series for A50 stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are Estimated MSE (977480.63), MAE (817.08), MAPE (6.13), and RMSE (988.68).

Figure A-3 in appendix exhibit the forecast of estimated trained and test ARIMA model for EUS stock market index. Forecasting accuracy indicators presented in Table 4.5 are MSE (147439.81), MAE (335.61), MAPE (10.27), and RMSE (383.98) Lower and upper bound shows the sudden drop in forecasted windows failed the ARIMA forecasting method to predict in such case as arrival of Covid-19 in this time frame. ARFIMA estimated accuracy indicators presented in Table 4.5 are 144881.79, 331.82, 10.06, 380.63 and 0.0001 respectively. Figure A-4 in appendix exhibit the forecasted and actual series for EUS stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are MSE (210076.66), MAE (373.28), MAPE (12.05), and RMSE (458.34).

Figure A-3 in appendix exhibit the forecast of estimated trained and test ARIMA model for CAC market index. Forecasting accuracy indicators presented in Table 4.5 are MSE (419391.59), MAE (581.68), MAPE (11.27), and RMSE (647.60). Results shows the train and test data set with in defined range upper and lower bound that shows the perfection of ARIMA for CAC market index. ARFIMA estimated accuracy indicators presented in Table 4.5 are, MSE (424867.23), MAE (587.02), MAPE (11.39),and RMSE (651.82).

Figure A-4 in appendix exhibit the ARFIMA forecasted and actual series for CAC stock market. Forecasted results are shown in Figure A-5 in appendix for GARCH model. The model choose 100 observations then 20 observations are used for the training purpose and the next 1 observation is forecasted considering it as an out of sample observation and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are Estimated MSE (767057.34), MAE (678.66), MAPE (14.18), and RMSE (875.82)

Figure A-3 in appendix exhibit the forecast of estimated trained and test ARIMA model for DAX stock market index. Forecasting accuracy indicators presented in Table 4.5 are MSE (1777831.22), MAE (1149.43), MAPE (9.66), and RMSE (1333.35). Lower and upper bound shows the sudden drop in forecasted windows failed the ARIMA forecasting method to predict in such case as arrival of Covid-19 in this time frame. ARFIMA estimated accuracy indicators presented in Table 4.5 are 1793453.96, 1155.66, 9.63, and 1339.20 respectively. Figure A-4 in appendix exhibit the ARFIMA forecasted and actual series for DAX stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are MSE (2274733.60), MAE (1179.00), MAPE (10.57), and RMSE (1508.22).

Forecasting accuracy indicators presented in Table 4.5 are MSE (5780628.46), MAE (1918.54), MAPE (7.79), and RMSE (2404.29). Results shows the train and test data set with in defined range upper and lower bound that shows the perfection of ARIMA for HSI market index. ARFIMA estimated accuracy indicators presented in Table 4.5 are, MSE (10825560.81), MAE (2816.96), MAPE (9.79), and RMSE (3290.22). Figure A-4 in appendix exhibit the ARFIMA forecasted and actual series for HSI stock market. Forecasted results are shown in Table 4.5. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are Estimated MSE (26170969.68), MAE (4159.28), MAPE (15.67), and RMSE (5115.76).

Forecasting accuracy indicators presented in Table 4.5 are MSE (2259523.37), MAE (1169.65), MAPE (12.11), and RMSE (1503.17). Lower and upper bound shows the sudden drop in forecasted windows failed the ARIMA forecasting method to predict in such case as arrival of Covid-19 in this time frame. ARFIMA estimated accuracy indicators presented in Table 4.5 are 3714292.56, 1, 340.41, 14.33, 1927.25, and 0.000961 respectively. Figure A-4 in appendix exhibit the forecasted and actual series for NIFTY stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are MSE (4822288.63), MAE (1509.69), MAPE (16.17), and RMSE (2195.97).

Forecasting accuracy indicators presented in Table 4.5 are MSE (4601531.01), MAE (1848.69), MAPE (8.38), and RMSE (2145.12). Results shows the train and test data set with in defined range upper and lower bound that shows the perfection of ARIMA for NIKKIE market index. ARFIMA estimated accuracy indicators presented in Table 4.5 are, MSE (4653054.52), MAE (1866.76), MAPE (8.47), and RMSE (2157.09). Forecasted results are shown in figure A-4 appendix. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are Estimated MSE (5145656.71), MAE (1839.47), MAPE (8.88), and RMSE (2268.40).

Figure A-3 in appendix exhibit the forecast of estimated trained and test ARIMA model for AEX stock market index. Forecasting accuracy indicators presented in Table 4.5 are MSE (2825.27), MAE (39.20), MAPE (7.61), and RMSE (53.15). Lower and upper bound shows the sudden drop in forecasted windows failed the ARIMA forecasting method to predict in such case as arrival of Covid-19 in this time frame. ARFIMA estimated accuracy indicators presented in Table 4.5 are 2297.13, 36.93, 7.00, and 47.93 respectively. Figure A-4 in appendix exhibit the ARFIMA forecasted and actual series for AEX stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are MSE (7635.30), MAE (58.49), MAPE (11.62), and RMSE (87.38).

Forecasting accuracy indicators for WIG presented in Table 4.5 are MSE (140252), MAE (265.23), MAPE (16.12), and RMSE (337.50). Results shows the train and test data set with in defined range upper and lower bound that shows the perfection of ARIMA for WIG market index. ARFIMA estimated accuracy indicators presented in Table 4.5 are, MSE (110053.55), MAE (228.48), MAPE (13.98), and

RMSE (331.74). Figure A-4 in appendix exhibit the ARFIMA forecasted and actual series for WIG stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are Estimated MSE (105208.24), MAE (222.70), MAPE (13.64), and RMSE (324.36).

Forecasting accuracy indicators fpr SSGF stock index presented in Table 4.5 are MSE (1942.50), MAE (31.92), MAPE (10.60), and RMSE (44.07). Lower and upper bound shows the sudden drop in forecasted windows failed the ARIMA forecasting method to predict in such case as arrival of Covid-19 in this time frame. ARFIMA estimated accuracy indicators presented in Table 4.5 are 1439.88, 29.50, 9.63, and 37.95 respectively. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are MSE (1752.46), MAE (31.26), MAPE (31.26), and RMSE (41.86).

Forecasting accuracy indicators for IBX stock index presented in Table 4.5 are MSE (1435458.15), MAE (959.24), MAPE (12.64), and RMSE (1198.11). Results shows the train and test data set with in defined range upper and lower bound that shows the perfection of ARIMA for IBX market index. ARFIMA estimated accuracy indicators presented in Table 4.5 are, MSE (1096497.99), MAE (905.33), MAPE (11.43), and RMSE (1047.14). Figure A-4 in appendix exhibit the ARFIMA forecasted and actual series for IBX stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are Estimated MSE (1415933.40), MAE (954.25), MAPE (12.56), and RMSE (1189.93).

Figure A-3 in appendix exhibit the forecast of estimated trained and test ARIMA model for SMI stock market index. Forecasting accuracy indicators presented in Table 4.5 are MSE (356021.12), MAE (455.58), MAPE (4.55), and RMSE (596.68). Lower and upper bound shows the sudden drop in forecasted windows failed the ARIMA forecasting method to predict in such case as arrival of Covid-19 in this time frame. ARFIMA estimated accuracy indicators presented in Table 4.5 are

382890.67, 469.89, 4.80, and 618.78 respectively.For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are MSE (652481.23), MAE (566.59), MAPE (5.92), and RMSE (807.76).

Forecasting accuracy indicators presented in Table 4.5 are MSE (714998.67), MAE (584.97), MAPE (9.73), and RMSE (845.58)). Results shows the train and test data set with in defined range upper and lower bound that shows the perfection of ARIMA for FTSE market index. ARFIMA estimated accuracy indicators presented in Table 4.5 are, MSE (690989.36), MAE (578.86), MAPE (9.61), and RMSE (831.26). For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are Estimated MSE (1322141.22), MAE (757.71), MAPE (12.76), and RMSE (1149.84

Figure A-3 in appendix exhibit the forecast of estimated trained and test ARIMA model for DJI30 stock market index. Forecasting accuracy indicators presented in Table 4.5 are MSE (5355410.53), MAE (1897.17), MAPE (7.41), and RMSE (2314.18). Lower and upper bound shows the sudden drop in forecasted windows failed the ARIMA forecasting method to predict in such case as arrival of Covid-19 in this time frame. ARFIMA estimated accuracy indicators presented in Table 4.5 are 6837398.68, 2048.00, 8.31, and 2614.84 respectively. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are MSE (20604064.29), MAE (2898.19), MAPE (12.21), and RMSE (4539.17).

Figure A-3 in appendix exhibit the forecast of estimated trained and test ARIMA model for SNP 500 market index. Forecasting accuracy indicators presented in Table 4.5 are MSE (73398.15), MAE (225.35), MAPE (7.40), and RMSE (270.92). Results shows the train and test data set with in defined range upper and lower bound that shows the perfection of ARIMA for SNP 500 market index. ARFIMA estimated accuracy indicators presented in Table 4.5 are, MSE (57306.14), MAE (189.94), MAPE (6.44), and RMSE (239.39). Figure A-4 in appendix exhibit the

ARFIMA forecasted and actual series for SNP500 stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure A-5 in appendix. Forecasting Accuracy indicators are Estimated MSE (115822.81), MAE (234.60), MAPE (8.31), and RMSE (340.33).

The study applies three dynamic classical models to forecast our selected stock market indices. After the estimation of these models, all of the sample market indices are forecasted one by one. Results of the analysis show that the classical ARIMA model performs better in forecasting the hourly data for stock market indices.

4.1.4 Machine Learning and Deep Learning Methods

This analysis section deals with the machine learning and deep learning analysis. Section has been further divided into two subsessions: Model Estimation and Forecasting.

4.1.4.1 Model Estimations

For machine learning 70:30 ratio has been used to train and test the data set. There are number of parameters that optimize the learning ability of machine. Each stock market has its own dimensions and volatility sequences. Similarly, each frequency has its own dimensions. Different machine learning optimizers are currently available and has presence in literature. To make it simple and comparable we choose and standardize the optimizers for each stock market.

For supervised learning, this study use SVM. Kernel function is one of the important parameters in regression-based forecasting. A range of kernel functions is available with scikit-learn library. Linear kernel functions are the simplest kernel function with machine learning available. Which is used to recognize the pattern in data. It works same as principal component analysis (PCA) that designed to identify the unknown trends. Polynomial kernel function is another kernel function known as non-stationary kernel. This is used to find the high dimensions in data by adding some new dimensions in data. Some other kernel functions are Radial Basis kernel (RBF) a Gaussian based kernel function and sigmoid kernel function comes from neural networks field and often refers as multilayer perception kernel. RBF is known as highly local kernel with capabilities of dealing with small and large data samples. It has the ability to map the data sample into higher dimensions as well as has the characteristics of lower parameters requirements then polynomial kernel (Li et al., 2020). On basis of these characteristics and literature support Kim (2003); Min and Lee (2005); Li et al. (2020) prefer RBF as kernel function of SVM. Other parameters like c, gamma and epsilon has been set as default (gamma = auto). For SVM, we have taken time variable as an input which is the bases for the regression analysis in order to fit the regression model.

Typical machine learning work on pre-define rules and data features. Best thing about deep learning methods is the data features can be identified from raw data. For LSTM this study use one input (closing prices of Indices) to predict the output value. Neural layer is important in LSTM so this study uses 2 layers of neurons. First layer contains the 64 neurons and 20 dense neurons. The second layer 32 neurons and 1 dense neuron to generate the output. To prevent the over fitting 20% nodes has been used as drop (discard rate) out rate (Li et al., 2020). Activation functions play a vital rule in optimization of neural networks so refer as most crucial component. Output of neural networks, its accuracy and efficiency are all depends on the activation function. Plenty of layers activation function are available with keras library. Range from relu (rectified linear unit) that allows to use non zero thresholds. Sigmoid non-linear activation function, that take any number from input on x axis and then transform that input number into a scaler output range from 0 to 1 a bounded output. Similarly other non-linear activation function "tanh' which is also like logistic sigmoid function but has the range from -1 to 1. That makes tanh more better activation function then the sigmoid activation function. Some other activation functions are "softmax, softplus, softsign, selu, elu and some exponential functions" also used worldwide.

Based on best tuned optimized results, tanh activation function perform better than other function and has been standardized for further analysis for this study which is also been used by a recent study on stock market predictions by (Li et al., 2020). To reduce the losses weights and learning rates change is a common practice in deep learning. Optimization algorithms used to change these attributes like learning rate and weights of data. These algorithums ranges from some basic optimizers like "Gradient decent, stocostic gradient decent, minibatch gradient decent" to advance optimizer like "adam and RMSprop". Adam function works on momentums or first and second order. Fast rolling and jumps over the minimum increase the velocity. Adam takes these steps carefully and use exponential decaying average while computing the loss. RMSprop is another optimizer that work like gradient decent but with momentum. By restricting the fluctuation vertically this increases the learning rate and speed up the loss computation process. Based on these characteristics we used RMSprop as optimizer function and MSE to define the loss. Batch size has been set 256 with buffer size set on basis of sample data. According to google tensorflow core v2.3.0 documentation buffer size for data elements 10000 must be 1000. Evaluation intervals in each epoch has been set as 200 and number of epochs(iterations) has been set as 100. For Hybrid methods we use three steps working. In step one computed identified ARIMA and GARCH model residual series. In step two we compute the residual series from machine and deep learning methods. In final step we combine the residual series with original series and compute the final results.

4.1.4.2 Forecasting of Models

Table 4.6 explain the computed errors from selected models for each stock market. For ASX stock market SVM MSE, MAE, MAPE and RMSE values are 5121.81, 48.38, 9.84, 71.57 respectively. For Hybrid SVM accuracy indicators MSE, MAE, MAPE and RMSE values are 383506.12, 490.09, 8.08 and 619.28 respectively which worsen the results indicates simple SVM perform better in machine learning regression. Figure 4.1 exhibit the actual and forecasted SVM and ARIMA-SVM models for ASX.

The Deep learning methods LSTM, and Hybrid LSTM included ARIMA and GARCH results presented in Table 4.6 for all selected stock markets. For ASX stock market index LSTM estimated model error's MSE, MAE, MAPE and RMSE values are 1877.96, 31.10, 0.51 and 43.34 respectively which outperform the machine learning. For hybrid ARIMA-LSTM method computed error's MSE, MAE, MAPE and RMSE values are 1071.24, 20.98, 0.36 and 32.73. Similarly for hybrid GARCH-LSTM method computed error's MSE, MAE, MAPE and RMSE are 1875.46, 30.03, 0.50 and 43.31 respectively. Figure 4.2 exhibit the train, test and real series of ASX. Where green and red colored series represent the trained and actual test time series. Purple, orange and magenta color represent the LSTM, ARIMA-LSTM and GARCH-LSTM respectively tests results. ARIMA-LSTM is the closest to actual time series.

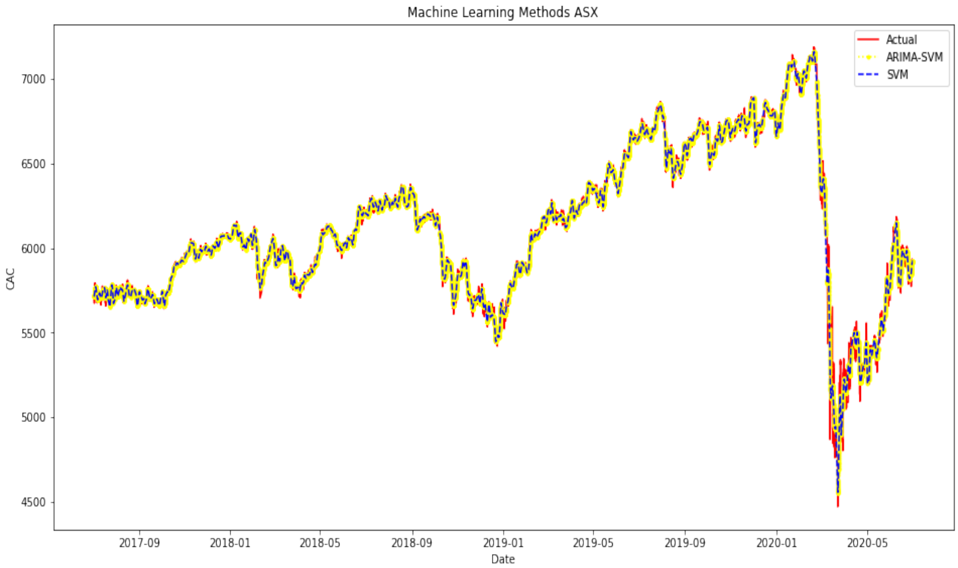


FIGURE 4.1: Figure 4.1: ASX Machine Learning Forecasting (Hourly Data)

For A50 stock market Index, SVM model evaluation criteria MSE, MAE, MAPE and RMSE has values 1160068.88, 919.13, 7.42, and 1077.06 respectively. For Hybrid SVM accuracy indicators MSE, MAE, MAPE and RMSE has the values 1121935.66, 900.60, 7.26, and 1059.21 respectively which improve the results indicates simple ARIMA-SVM. Willmott and Matsuura (2005) suggest MAE as better predictor indicator over RMSE. Based on MAE and MAPE results, ARIMA-SVM perform better than the SVM in machine learning regression. Figure 4.3 exhibit the actual and forecasted SVM and ARIMA-SVM models for A50, which shows that Hybrid ARIMA-SVM out perfomr the SVM model.

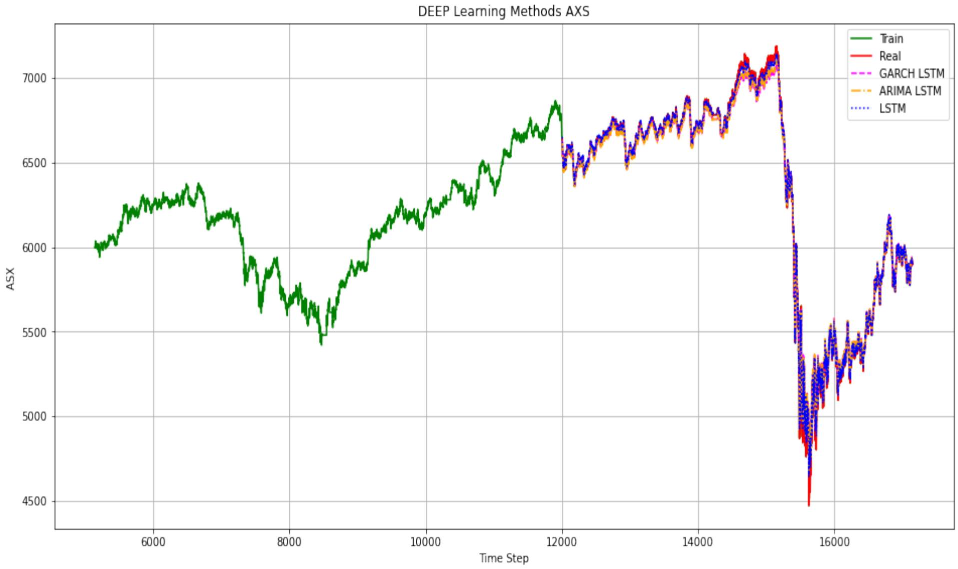


FIGURE 4.2: Figure 4.2: ASX DEEP Learning Forecasting (Hourly Data)

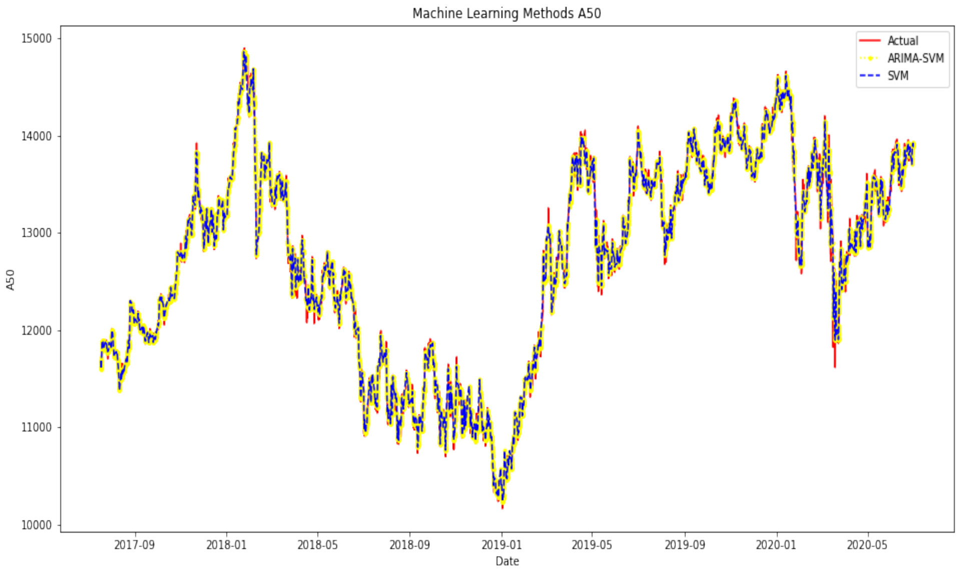


FIGURE 4.3: A50 Machine Learning Forecasting (Hourly Data)

For Deep learning methods LSTM, Hybrid ARIMA-LSTM and GARCH-LSTM based results for A50 stock market index has been reported in Table 4.6. For A50 stock market index LSTM estimated model have MSE, MAE, MAPE and RMSE values as 7957.30, 81.14, 0.60, and 89.20 respectively which outperform the machine learning. For hybrid ARIMA-LSTM computed model have MSE, MAE, MAPE and RMSE values as 6094.01, 68.15, 0.50 and 78.06 respectively. Similarly, for hybrid GARCH-LSTM model have MSE, MAE, MAPE and RMSE values as 3803.69, 49.84, 0.37 and 61.67 respectively. Figure 4.4 exhibit the train, test and real series of A50. Where green and red colored series represent the trained and actual test time series. Purple, orange and magenta color represent the LSTM, ARIMA-LSTM and GARCH-LSTM respectively. Test results shows Proposed GARCH-LSTM is the closest to actual time series.

FIGURE 4.4: A50 Deep Learning Forecasting (Hourly Data)

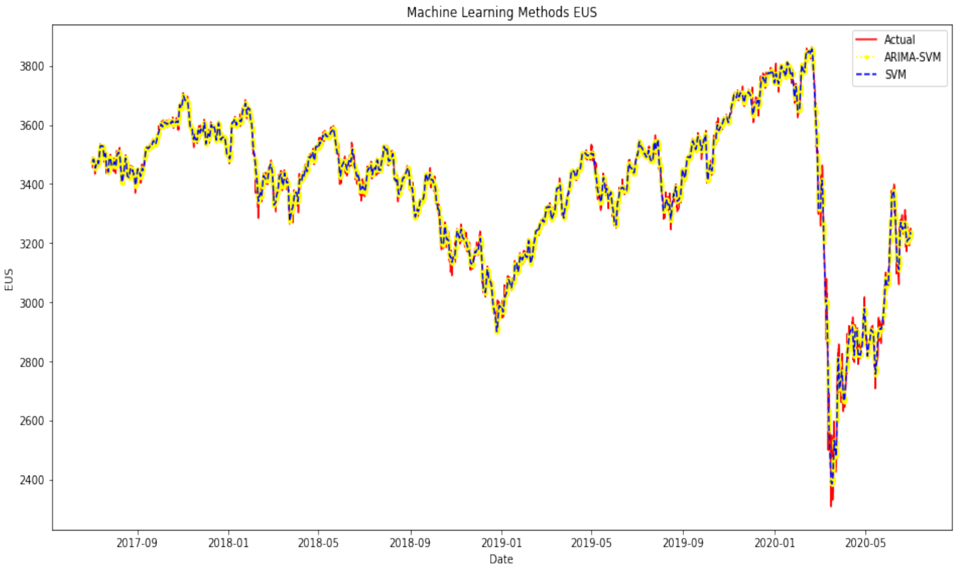


FIGURE 4.5: EUS Machine Learning Forecasting (Hourly Data)

For EUS stock market Index SVM, model evaluation criteria MSE, MAE, MAPE and RMSE has values 58765.89, 180.00, 5.51 and 242.42 respectively. For Hybrid SVM accuracy indicators MSE, MAE, MAPE and RMSE has the values 60355.83, 174.13, 5.41 and 245.67 respectively which improve the MAE and MAPE.

Table 4.6: Machine and Deep Learning Forecasting

	ASX	A50	EUS	CAC	DAX	HSI	NIFTY	NIKKEI	AEX	WIG	SSGF	IBX	SMI	FTSE	DJI	SNP
								SVM								
MSE	191688	1160069	58766	154740	775939	4354805	48415	1788898	1193	55742	927	878858	386118	275918	3621003	49849
MAE	351.80	919.13	180.00	285.32	668.63	1643.51	110.05	1073.40	25.34	164.12	21.68	656.64	510.84	361.92	1481.35	177.50
MAPE	5.75	7.42	5.51	5.59	5.69	6.01	1.20	5.05	4.73	8.36	6.32	937.47	5.38	5.43	5.98	6.35
RMSE	437.82	1077.06	242.42	393.37	880.87	2086.82	220.03	1337.50	34.54	236.10	30.44	937.47	621.38	525.28	1902.89	223.27
								LSTM								
MSE	1878	7957	2040	4530	7746	27421	25719	17217	16	5625	41	42772	16219	45474	39937	1161
MAE	31.10	81.14	30.09	46.28	59.10	101.35	85.04	97.95	2.65	44.90	3.65	125.03	93.54	109.63	144.76	27.00
MAPE	0.51	0.60	1.01	0.92	0.52	0.42	0.92	0.46	0.50	2.84	1.25	1.79	0.90	1.92	0.56	0.87
RMSE	43.34	89.20	45.16	67.31	88.01	165.59	160.37	131.21	4.04	75.00	6.43	206.81	127.35	213.25	199.84	34.08
								ARIMA-SVM								
MSE	192349	1121936	60356	157444	788970	4307237	746738	1737034	1197	58840	341	873790	386953	284149	3562240	41856
MAE	353.62	900.60	174.13	281.05	659.33	1624.73	664.58	1050.55	25.28	155.61	8.99	634.70	488.02	313.48	1456.96	154.56
MAPE	5.79	7.26	5.41	5.56	5.67	6.02	6.29	5.04	4.73	8.20	7.48	7.69	5.22	5.18	6.01	6.34
RMSE	438.58	1059.21	245.67	396.79	888.24	2075.39	864.14	1317.97	34.59	242.57	18.47	934.77	622.06	533.06	1887.39	204.59
								ARIMA-LSTM								
MSE	1071	6094	2255	1669	14456	24226	21808	26492	14	5411	29	63310	12779	37013	30155	960
MAE	20.98	68.15	30.88	23.86	75.77	98.85	91.808	128.18	2.15	43.84	3.41	169.03	84.10	113.73	122.24	23.74
MAPE	0.36	0.50	1.04	0.50	0.71	0.40	0.974	0.60	0.42	2.77	1.16	2.34	0.82	1.94	0.49	0.77
RMSE	32.73	78.06	47.49	40.85	120.23	155.65	147.677	162.76	3.80	73.56	5.36	251.62	113.05	192.39	173.65	30.99
								GARCH-LSTM								
MSE	1875	3804	706	2817	3997	17673	208749	57026	24	4819	35	21381	12388	35550	30279	870
MAE	30.03	49.84	15.49	34.00	42.21	98.15	212.86	183.49	3.49	43.47	3.56	91.78	75.08	96.84	104.03	20.14
MAPE	0.50	0.37	0.50	0.68	0.38	0.39	2.4023	0.88	0.64	2.72	1.21	1.30	0.73	1.70	0.42	0.65
RMSE	43.31	61.67	26.57	53.07	63.22	132.94	456.89	238.80	4.87	69.42	5.90	146.22	111.30	188.55	174.01	29.49

Studies (Willmott and Matsuura (2005); Willmott et al. (2009); Gonzalez-Vidal et al. (2019)) prefer MAE over RMSE while comparing the accuracy of models. On basis of literature, we can infer that ARIMA-SVM perform well than the SVM in machine learning regression. Figure 4.5 exhibit the actual and forecasted SVM and ARIMA-SVM models for EUS.

Figure 4.6: EUS Deep Learning Forecasting (Hourly Data)

For Deep learning methods LSTM, Hybrid ARIMA-LSTM and GARCH-LSTM computed results for EUS stock market index has been reported in Table 4.6. For LSTM estimated model have MSE, MAE, MAPE and RMSE values as 2039.54, 30.09, 1.01 and 45.16 respectively which outperform the machine learning. For hybrid ARIMA-LSTM computed model have MSE, MAE, MAPE and RMSE values as 2255.00, 30.88, 1.04 and 47.49 respectively. Similarly, for hybrid GARCH-LSTM model have MSE, MAE, MAPE and RMSE values as 705.86, 15.49, 0.50 and 26.57 respectively. Figure 4.6 exhibit the train, test and real series of EUS. Where green and red colored series represent the trained and actual test time series. Purple, orange and magenta color represent the LSTM, ARIMA-LSTM and GARCH-LSTM respectively. Test results indicate that GARCH-LSTM is the closest to actual time series.

For the CAC stock market, SVM model forecasting accuracy indicators have the values as 154739.60, 285.32, 5.5885, and 393.37 respectively for MSE, MAE,

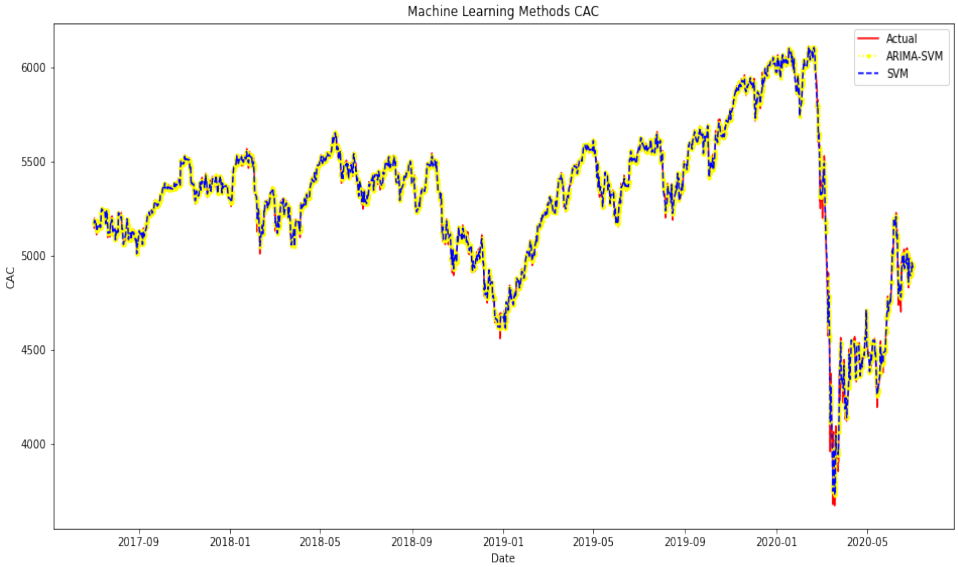


FIGURE 4.7: CAC Machine Learning Forecasting (Hourly Data)

FIGURE 4.8: CAC Deepe Learning Forecasting (Hourly Data)

MAPE, and RMSE. For The hybrid ARIMA-SVM accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 157443.54, 281.05, 5.5594, and 396.79 respectively, which improves the results of machine learning methods with the hybrid technique. Improved MAE and MAPE results indicate that ARIMA-SVM performs better in machine learning regression than the simple SVM model. Figure 4.7 exhibit the actual and forecasted SVM and ARIMA-SVM models for CAC.

The The deep learning methods LSTM, and the hybrid LSTM along with ARIMA and GARCH results are presented in Table 4.6 for all selected stock markets. For the CAC stock market index, LSTM estimated model accuracy indicators have the MSE, MAE, MAPE, and RMSE values 2040.03, 34.94, 0.70, and 45.17 respectively which outperform the machine learning models. For the hybrid, ARIMA-LSTM MSE, MAE, MAPE, and RMSE are 1668.90, 23.86, 0.50, and 40.85. Similarly, for the hybrid GARCH-LSTM method have the values of 4915.89, 49.30, 1.02, and 70.11 for MSE, MAE, MAPE, and RMSE respectively. Figure 4.8 exhibit the train, test, real and forecasted series of CAC. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta colore represent the results of LSTM, ARIMA-LSTM, and GARCH-LSTM model respectively. Results show that ARIMA-LSTM has lowest forecasting accuracy indicators and is close to actual time series so it is the best fitted deep learning method.

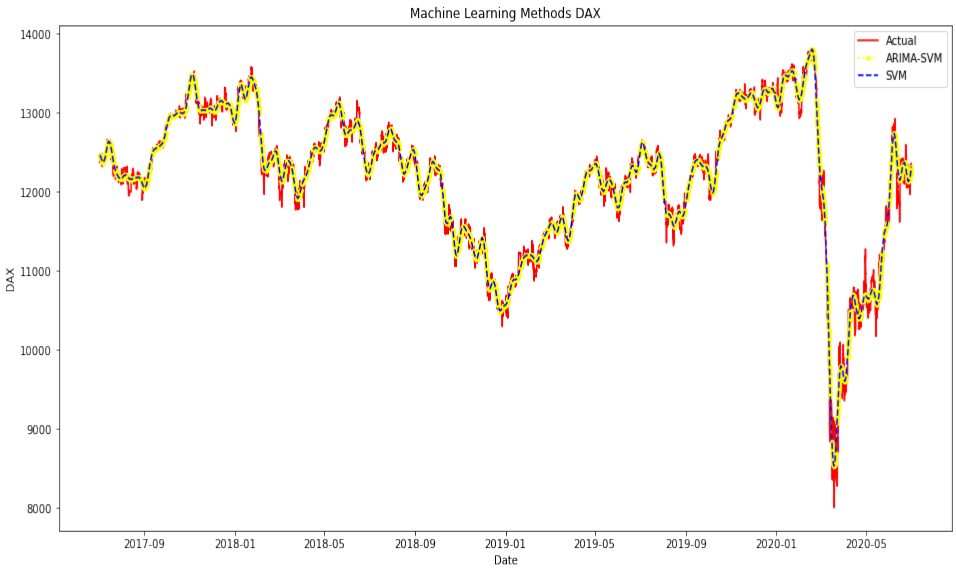


Figure 4.9: DAX Machine Learning Forecasting (Hourly Data)

For DAX stock market Index SVM, model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have values 775939.29, 668.63, 5.6920, and 880.87 respectively. The hybrid ARIMA-SVM model, forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have the values 788970.04, 659.33, 5.6674, and 888.24 respectively, which improve MAE, and MAPE values. Based on the results ARIMA-SVM model is a better predictor than the Simple SVM method for the DAX stock

FIGURE 4.10: DAX Deep Learning Forecasting (Hourly Data)

market index. Figure 4.9 exhibit the actual and forecasted SVM and ARIMA-SVM models for DAX.

For Deep learning methods, LSTM hybrid ARIMA-LSTM, and GARCH-LSTM are used and results for DAX stock market index are reported in Table 4.6. For the DAX stock market index, LSTM estimated model have MSE, MAE, MAPE, and RMSE values as 7746.38, 59.10, 0.52, and 88.01 respectively outperform the machine learning models. The hybrid ARIMA-LSTM model MSE, MAE, MAPE, and RMSE have the values as 14456.40, 75.77, 0.71, and 120.23 respectively. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 3997.23, 42.21, 0.38, and 63.22. Figure 4.10 exhibit the train, test, actual and forecasted series for DAX. Test results shows that GARCH-LSTM with the lowest forecasting indicators values and closest to the actual time series is the best model for the DAX stock market index.

For HSI stock market Index, SVM forecasting model accuracy indicators have the values as 4354805.34, 1643.51, 6.0133, and 2086.82 respectively for MSE, MAE, MAPE, RMSE. The hybrid ARIMA-SVM forecasting model, MSE, MAE, MAPE, and RMSE have the values as 4307237.45, 1624.73, 6.0239, and 2075.39 respectively. Based on MSE, MAE, and RMSE forecasting indicators values, we can

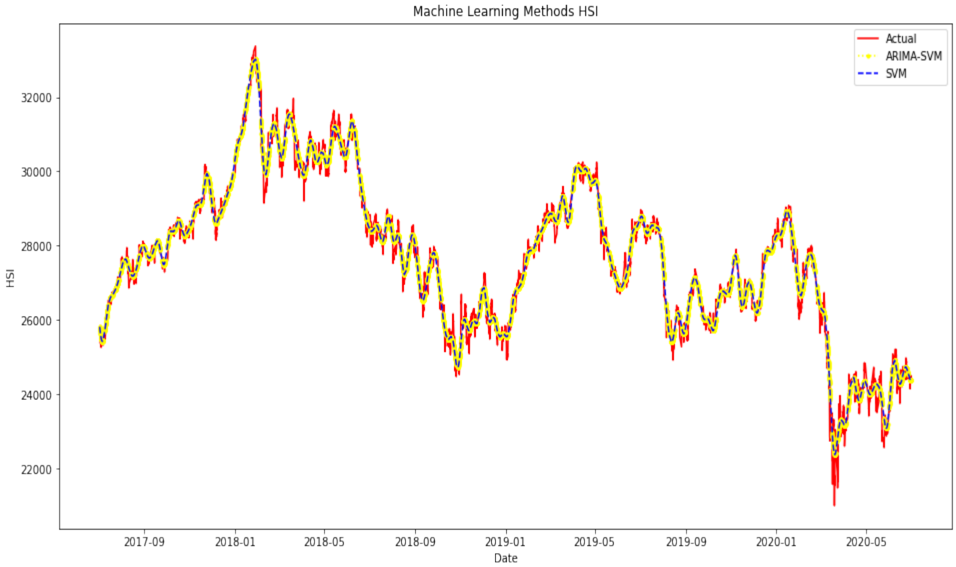


FIGURE 4.11: HSI Machine Learning Forecasting (Hourly Data)

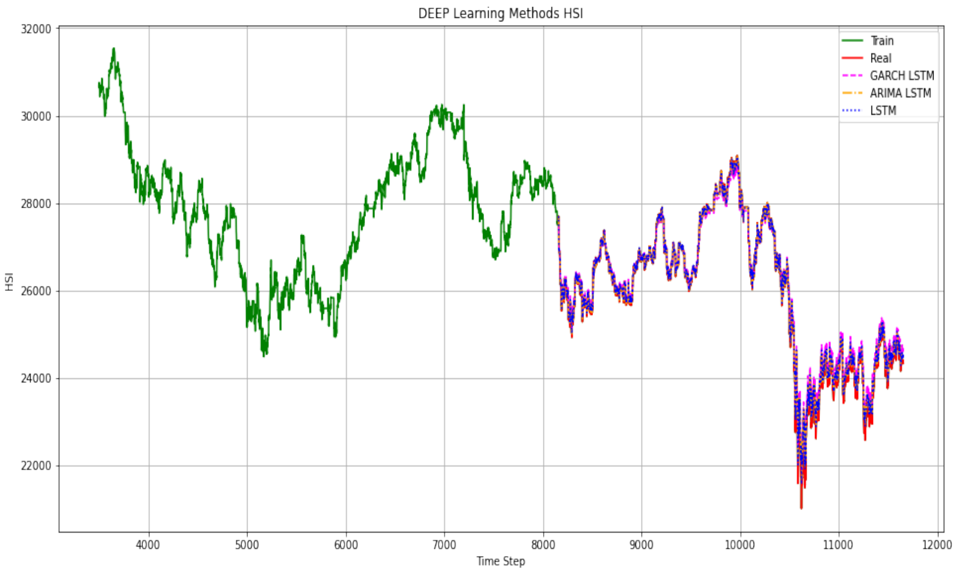


FIGURE 4.12: HSI Deep Learning Forecasting (Hourly Data)

infer that SVM performs better than the ARIMA-SVM in machine learning regression. Figure 4.11 exhibit the actual and forecasted SVM and ARIMA-SVM models for the HSI Stock market index.

For LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM deep learning methods are used for forecasting the HSI stock market index results are presented Table 4.6. For HSI stock market index, LSTM model have forecasting accuracy inidcatiors values as 27420.73, 101.35, 0.42, and 165.59 respectively for MSE,MAE,MAPE and RMSE. Which outperform the conventional machine learning models and the hybrid ARIMA-SVM models. For the hybrid, ARIMA-LSTM model forecasting MSE, MAE, MAPE, and RMSE have the values as 24226.33, 98.85, 0.40, and 155.65 respectively. Similarly, for the hybrid GARCH-LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values 17673.50, 98.15, 0.39, and 132.94 respectively. Figure 4.12 exhibit the train, test, and predicted series of HSI. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta color represents LSTM, ARIMA-LSTM, and GARCH-LSTM respectively. Results shows that GARCH-LSTM being closest to actual the time series is the best fitted deep learning method.

Table 4.6 represents the computed model's forecasting accuracy indicators values for the NIFTY stock market index. SVM model forecasting accuracy indicators have the values as 746593.18, 664.57, 6.2905, and 864.06 respectively for MSE, MAE, MAPE, and RMSE. For The hybrid ARIMA-SVM, accuracy indicators have the values as 746738.28, 664.58, 6.2877, and 864.14 respectively for MSE, MAE,MAPE and RMSE. Which worsen the results of forecasting accuracy indicators for the hybrid ARIMA-SVM. SVM performs better in machine learning regression than the hybrid ARIMA- SVM model. Figure 4.13 exhibit the actual and forecasted SVM and ARIMA-SVM models for NIFTY.

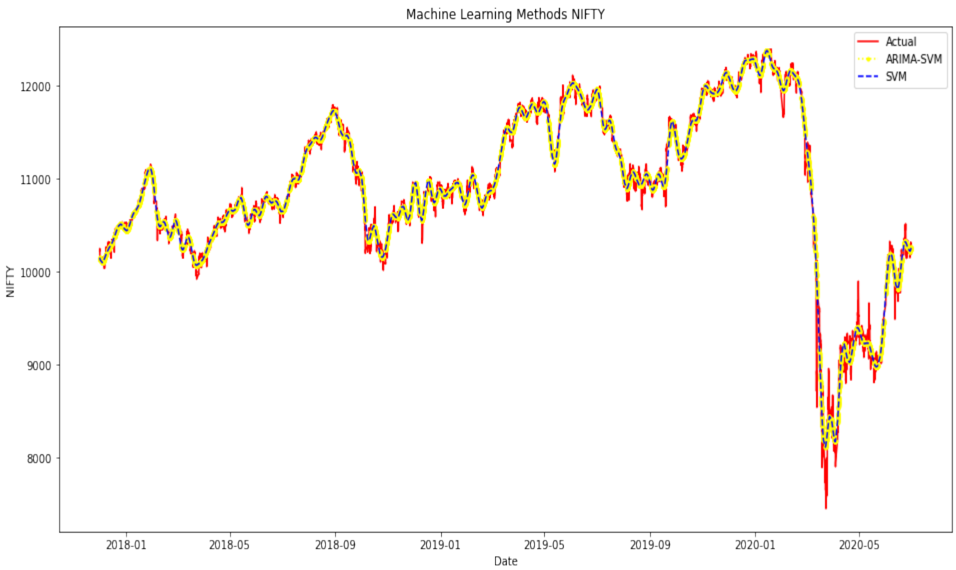


FIGURE 4.13: NIFTY Machine Learning Forecasting (Hourly Data)

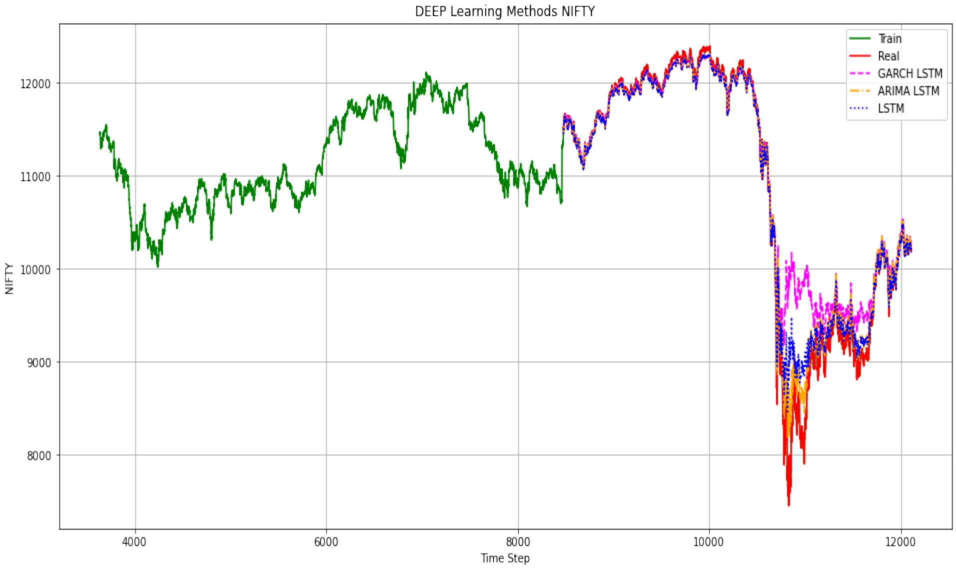


FIGURE 4.14: NIFTY Deep Learning Forecasting (Hourly Data)

For Deep learning methods, LSTM and the hybrid LSTM, including ARIMA and GARCH results are presented in Table 4.6 for the NIFTY Stock Market index. For the NIFTY stock market index, LSTM model have the MSE, MAE, MAPE, and RMSE values as 25718.55, 85.04, 0.92, and 160.37 respectively, which outperform the machine learning models. For the hybrid, ARIMA-LSTM MSE, MAE, MAPE, and RMSE are 21808.414, 91.808, 0.974, and 147.677 respectively. Similarly, for the hybrid GARCH-LSTM method have the values of 208749.07, 212.86, 2.4023, and 456.89 for MSE, MAE, MAPE, and RMSE respectively. Figure 4.14 exhibit the train, test, real and forecasted series of NIFTY. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta colore represent the results of LSTM, ARIMA-LSTM, and GARCH-LSTM model respectively. Results show that ARIMA-LSTM has lowest forecasting accuracy indicators and is close to actual time series so it is the best fitted deep learning method.

For NIKKEI stock market Index, SVM model have MSE, MAE, MAPE and RMSE values are 1788898.47, 1073.40, 5.0463, and 1337.50 respectively. For The hybrid ARIMA-SVM model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have the values 1737034.30, 1050.55, 5.0441, and 1317.97 respectively, which improves the results of forecasting accuracy indicators for the hybrid

ARIMA-SVM. ARIMA-SVM performs better in machine learning regression than the hybrid SVM model. Figure 4.15 exhibit the actual and forecasted SVM and ARIMA-SVM models for NIKKEI.

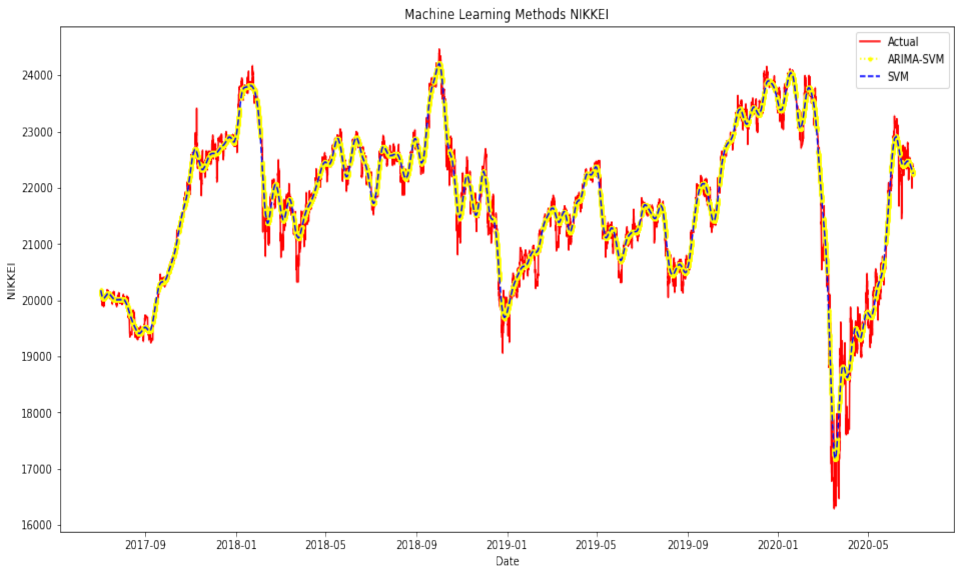


FIGURE 4.15: NIKKEI Machine Learning Forecasting (Hourly Data)

FIGURE 4.16: NIKKEI Deep Learning Forecasting (Hourly Data)

For Deep learning methods, results of LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM method for the NIKKEI stock market index have been reported in Table 4.6. For the NIKKEI stock market index, LSTM estimated model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 17217.08, 97.95, 0.46, and 131.21, respectively outperform the machine learning models. The hybrid ARIMA-LSTM model MSE, MAE, MAPE, and RMSE have the values as 26491.65, 128.18, 0.60, and 162.76, respectively. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 57025.86, 183.49, 0.88, and 238.80. Figure 4.16 exhibit the train, test, and entire series NIKKEI. Test results shows that LSTM with the lowest forecasting indicators values, is the best model for the NIKKEI stock market index.

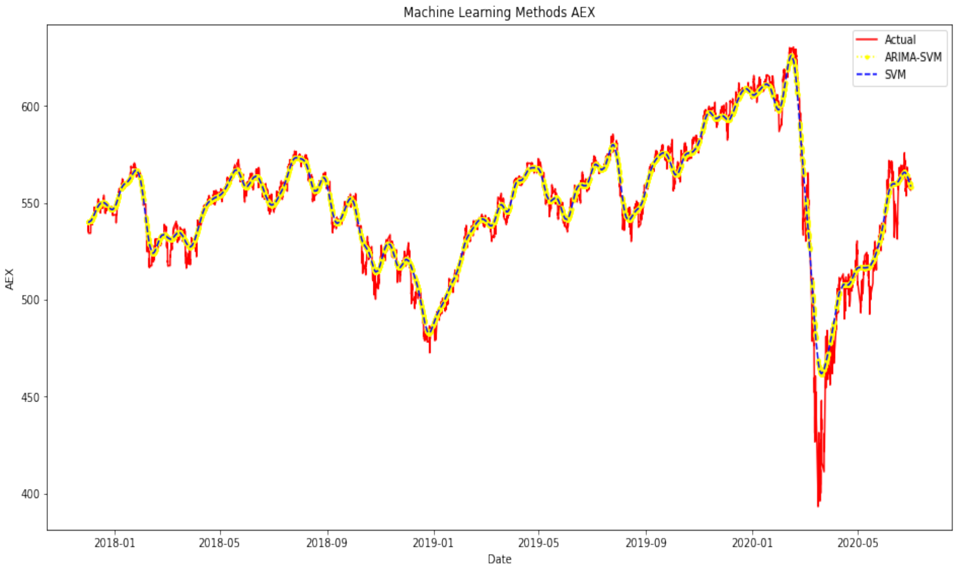


FIGURE 4.17: AEX Machine Learning Forecasting (Hourly Data)

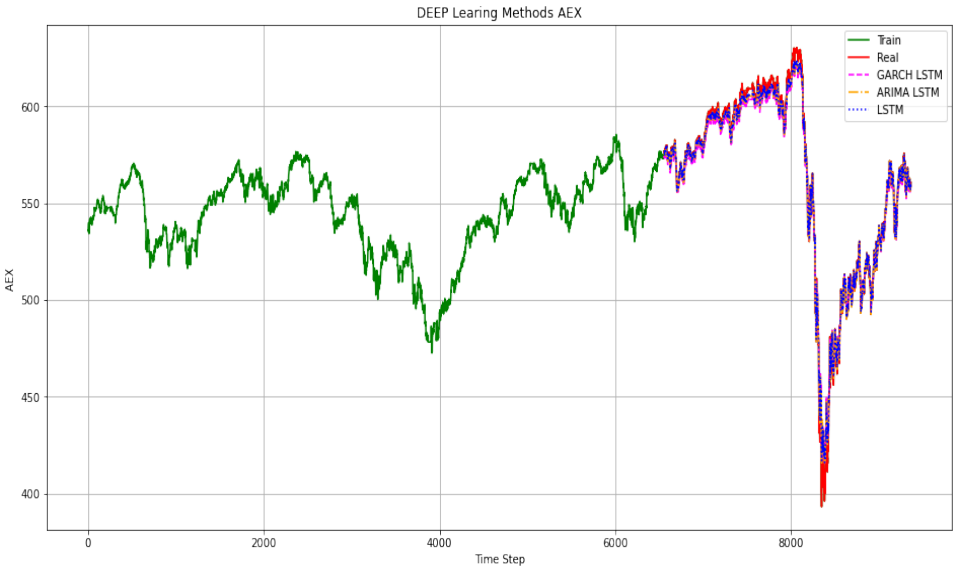


FIGURE 4.18: AEX Deep Learning Forecasting (Hourly Data)

For the AEX stock market Index, SVM forecasting model accuracy indicators have the values as 1192.73, 25.34, 4.7282, and 34.54 respectively for MSE, MAE, MAPE, RMSE. For the hybrid ARIMA-SVM forecasting model, MSE, MAE, MAPE, and RMSE have the values as 1196.72, 25.28, 4.7302, and 34.59 respectively. which worsen the results of forecasting accuracy indicators for the hybrid ARIMA-SVM. SVM performs better in machine learning regression than the hybrid ARIMA-SVM model. Figure 4.17 exhibit the actual and forecasted SVM and ARIMA-SVM models for the AEX Stock market index.

For Deep learning methods, LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM estimated results for the AEX stock market index have given in Table 4.6. For the AEX stock market index LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 16.34, 2.65, 0.50, and 4.04 respectively, which outperform the conventional machine learning models and the hybrid ARIMA-SVM models. For the hybrid, ARIMA-LSTM model forecasting MSE, MAE, MAPE, and RMSE have the values as 14.44, 2.15, 0.42, and 3.80 respectively. Similarly, for the hybrid GARCH-LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values 23.76, 3.49, 0.64, and 4.87 respectively. Figure 4.18 exhibit the train, test, and predicted series of AEX. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta color represents LSTM, ARIMA-LSTM, and GARCH-LSTM respectively. Results shows that LSTM model is the best fitted deep learning method for AEX market index.

For the WIG stock market, SVM model forecasting accuracy indicators i.e., MSE, MAE, MAPE, and RMSE have the values 55742.24, 164.12, 8.3628, and 236.10 respectively. For The hybrid ARIMA-SVM accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 58839.59, 155.61, 8.1984, and 242.57 respectively, which improves the results of machine learning methods with the hybrid technique. Results indicate that ARIMA-SVM performs better in machine learning regression than the simple SVM model. Figure 4.19 exhibit the actual and forecasted values of WIG stock market index, where SVM and ARIMA-SVM models are used and results shows ARIMA-SVM outperform SVM model .

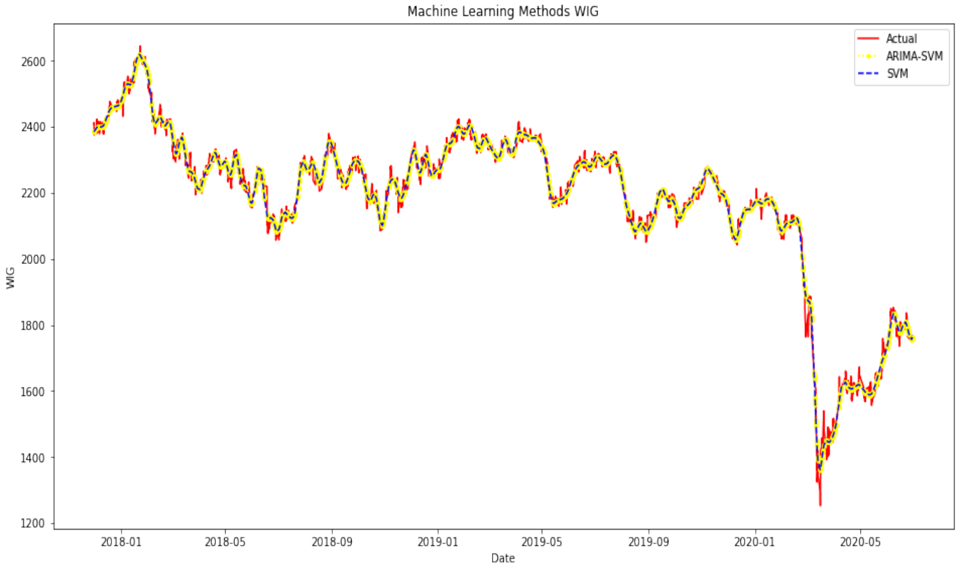


FIGURE 4.19: WIG Machine Learning Forecasting (Hourly Data)

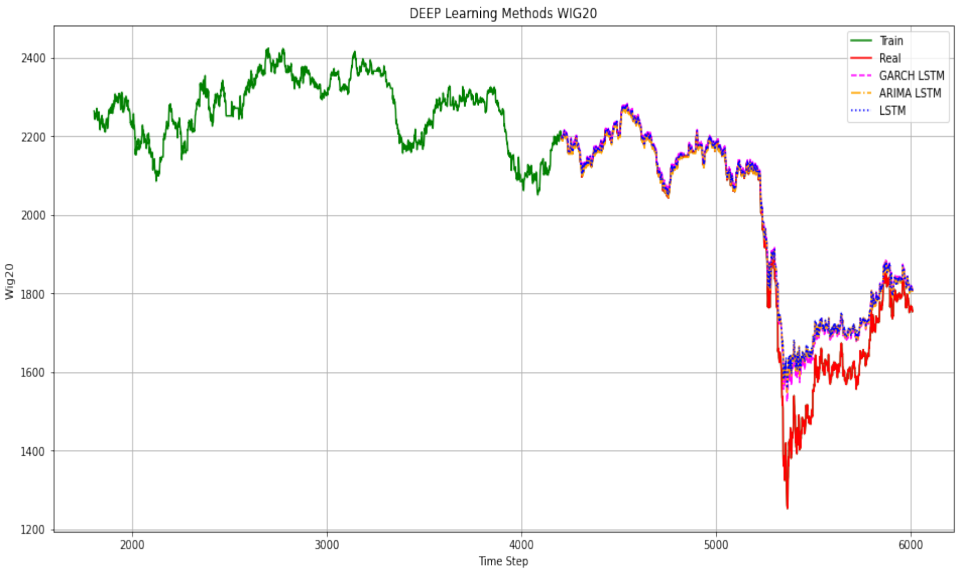


FIGURE 4.20: WIG Deep Learning Forecasting (Hourly Data)

The results of Deep learning methods i.e., LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM are presented in Table 4.6 for all selected stock markets. For the WIG stock market index, LSTM estimated model accuracy indicators have the MSE, MAE, MAPE, and RMSE values 5624.62, 44.90, 2.84, and 75.00 respectively, which outperform the machine learning models. For the hybrid, ARIMA-LSTM MSE, MAE, MAPE, and RMSE are 5410.54, 43.84, 2.77, and 73.56. Similarly, for the hybrid GARCH-LSTM method have the values of 4819.46, 43.47, 2.72, and 69.42 for MSE, MAE, MAPE, and RMSE respectively. Figure 4.20 exhibit the train, test, real and forecasted series of WIG. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta colore represent the results of LSTM, ARIMA-LSTM, and GARCH-LSTM model respectively. Results show that proposed GARCH-LSTM with the lowest forecasting accuracy indicators is the best fitted deep learning method.

For the SSGF stock market Index, SVM model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have values 926.63, 21.68, 6.3192, and 30.44 respectively. For The hybrid ARIMA-SVM model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have the values 341.01, 8.99, 7.4844, and 18.47 respectively. Based on the results, and ARIMA-SVM model is a better predictor than the Simple SVM method for the SSGF stock market index. Figure 4.21 exhibit the actual and forecasted SVM and ARIMA-SVM models for SSGF.

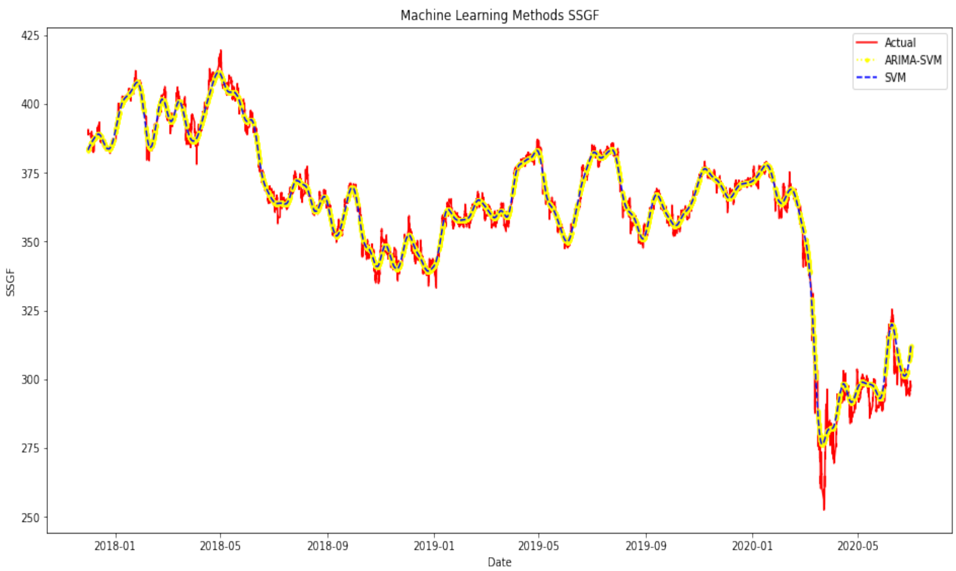


FIGURE 4.21: SSGF Machine Learning Forecasting (Hourly Data)

For Deep learning methods, LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM estimated results for SSGF stock market index have been reported in Table 4.6. For the SSGF stock market index, LSTM estimated model evaluation criteria (MSE, MAE, MAPE, and RMSE values 41.34, 3.65, 1.25, and 6.43 respectively) outperform the machine learning models. The hybrid ARIMA-LSTM computed model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 28.76, 3.41, 1.16, and 5.36 respectively. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 34.77, 3.56, 1.21, and 5.90. Figure 4.22 exhibit the train, test, actual and forecasted series for SSGF. Test results shows that ARIMA-LSTM with the lowest forecasting indicators values, is the best model for SSGF stock market index.

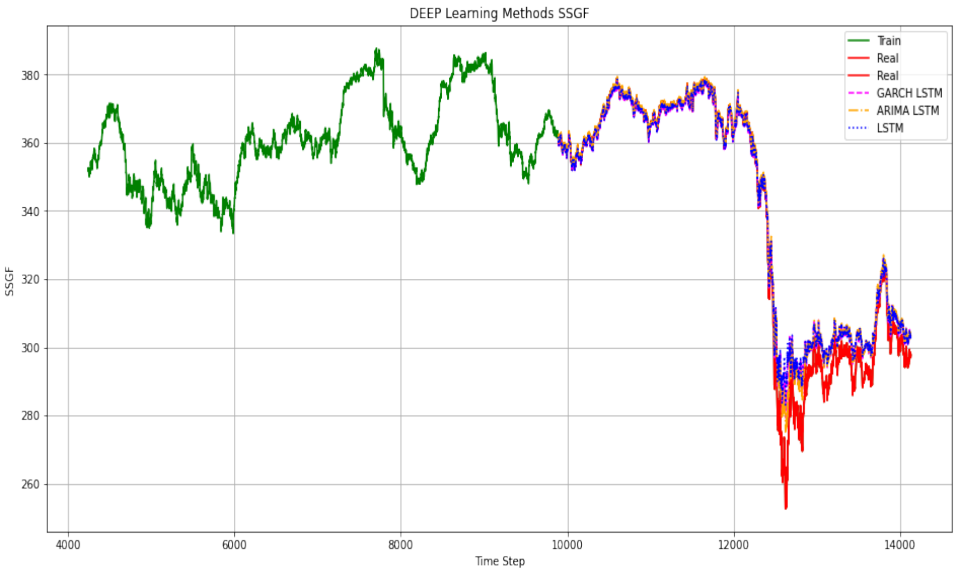


FIGURE 4.22: SSGF Deep Learning Forecasting (Hourly Data)

For the IBX stock market Index, SVM forecasting model accuracy indicators have the values as 878858.16, 656.64, 937.4744, and 937.47 respectively for MSE, MAE, MAPE, RMSE. For the hybrid ARIMA-SVM forecasting model, MSE, MAE, MAPE, and RMSE have the values as 873789.63, 634.70, 7.6916, and 934.77 respectively. Based on improved forecasting accuracy indicators values, we can infer that ARIMA-SVM performs better than SVM in machine learning regression. Figure 4.23 exhibit the actual and forecasted SVM and ARIMA-SVM models for the IBX Stock market index.

For Deep learning methods, LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM are used and results for IBX stock market index are represented in Table 4.6. For IBX stock market index, LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 42772.20, 125.03, 1.79, and 206.81respectively, which outperform the conventional machine learning models and the hybrid ARIMA-SVM models. For the hybrid, ARIMA-LSTM model forecasting MSE, MAE,

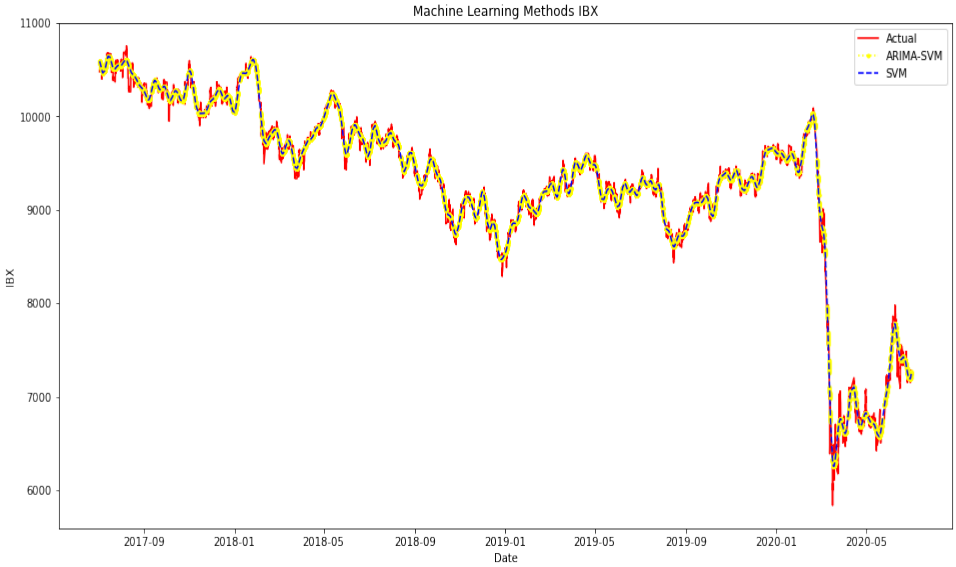


FIGURE 4.23: IBX Machine Learning Forecasting (Hourly Data)

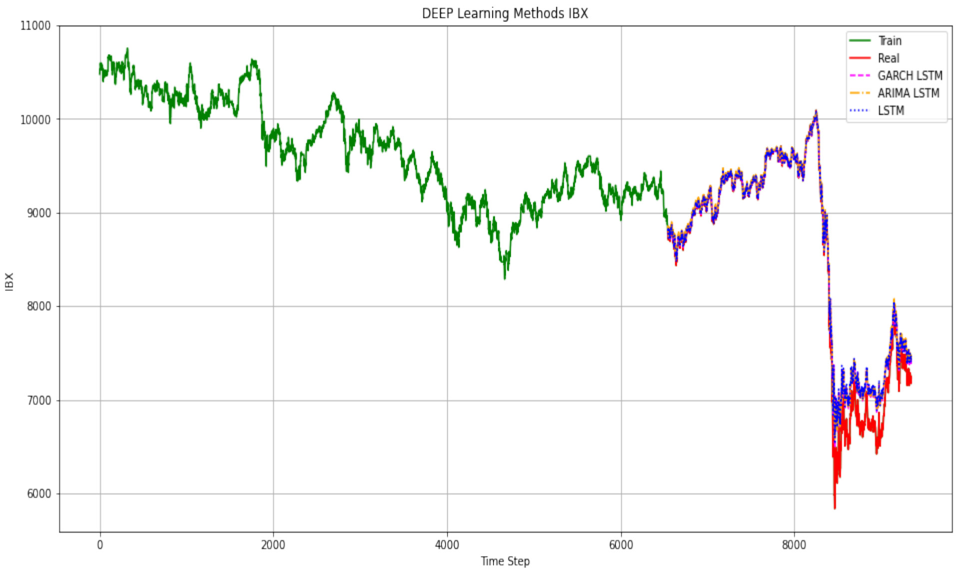


FIGURE 4.24: IBX Deep Learning Forecasting (Hourly Data)

MAPE, and RMSE have the values as 63310.42, 169.03, 2.34, and 251.62 respectively. Similarly, for the hybrid GARCH-LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values 21381.18, 91.78, 1.30, and 146.22 respectively. Figure 4.24 exhibit the train, test, and forecasted series of IBX. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta color represents LSTM, ARIMA-LSTM, and GARCH-LSTM respectively. Results shows proposed GARCH-LSTM is closest to actual the time series so it is the best fitted deep learning method.

For SMI stock market Index SVM, model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have values 386118.39, 510.84, 5.3799, and 621.38 respectively. For the hybrid ARIMA-SVM model forecasting accuracy indicators, MSE, MAE, MAPE, and RMSE have values as 386953.18, 488.02, 5.2188, and 622.06 respectively improve MAE, and MAPE values. Based on the results, and ARIMA-SVM model is a better predictor than the Simple SVM method for the SMI stock market index. Figure 4.25 exhibit the actual and forecasted values of SVM and ARIMA-SVM models for SMI.

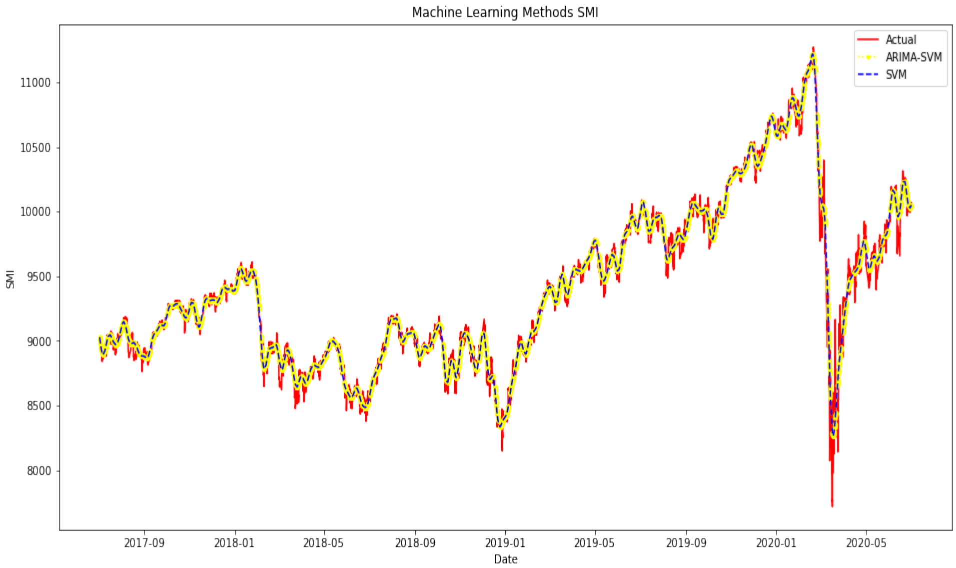


FIGURE 4.25: SMI Machine Learning Forecasting (Hourly Data)

For Deep learning methods, LSTM, hybrid ARIMA-LSTM ARIMA-LSTM, and GARCH-LSTM are used and results for the SMI stock market index are reported in Table 4.6. For the SMI stock market index, LSTM estimated model evaluation criteria (MSE, MAE, MAPE, and RMSE) have value 16218.72, 93.54, 0.90, and 127.35 respectively. Which outperform the machine learning models. The hybrid ARIMA-LSTM model MSE, MAE, MAPE, and RMSE have the values as 12779.40, 84.10, 0.82, and 113.05 respectively. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 12388.23, 75.08, 0.73, and 111.30. Figure 4.26 exhibit the train, test, actual and forecasted series for SMI. Test results shows that the proposed GARCH-LSTM with the lowest forecasting indicators values, is the best model for the SMI stock market index.

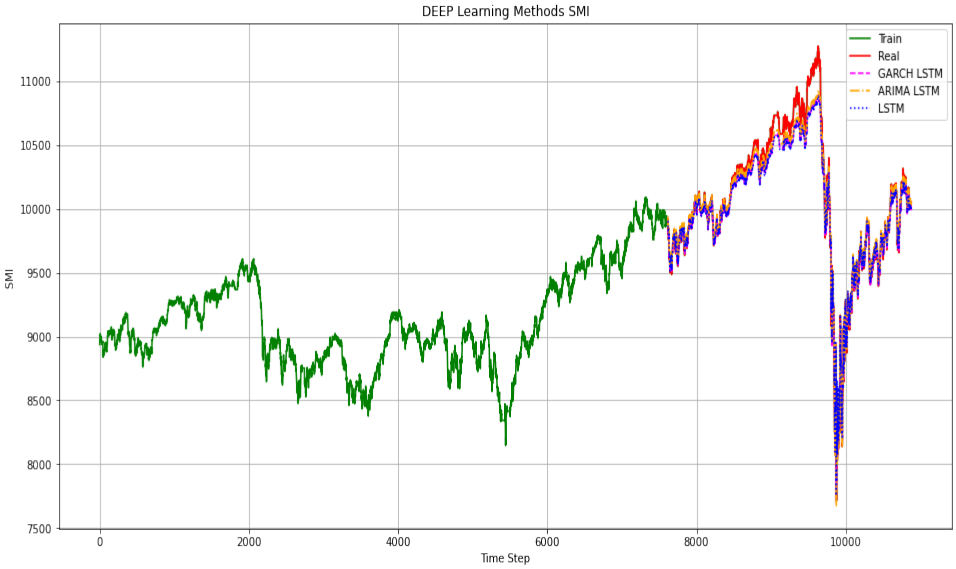


FIGURE 4.26: SMI Deep Learning Forecasting (Hourly Data)

For FTSE stock market Index, SVM model forecasting accuracy indicators (MSE, MAE, MAPE, and RMSE) have the values 275918.20, 361.92, 5.4346, and 525.28 respectively. For the hybrid ARIMA-SVM forecasting model, MSE, MAE, MAPE, and RMSE have the values as 284149.42, 313.48, 5.1777, and 533.06 respectively. Based on forecasting accuracy indicators values, we can infer that ARIMA-SVM performs better than the SVM in machine learning regression. Figure 4.27 exhibit the actual and forecasted values of SVM and ARIMA-SVM models for the FTSE Stock market index.

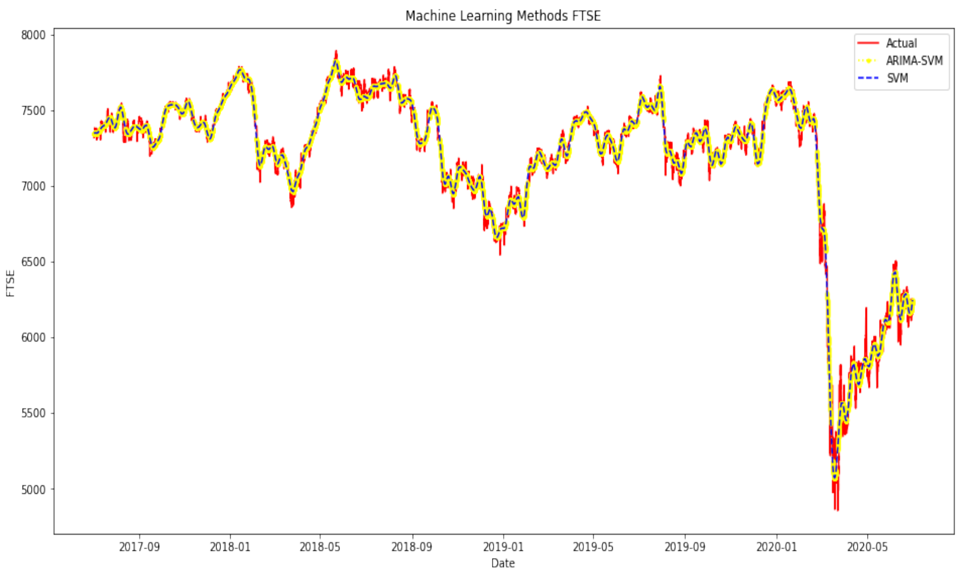


FIGURE 4.27: FTSE Machine Learning Forecasting (Hourly Data)

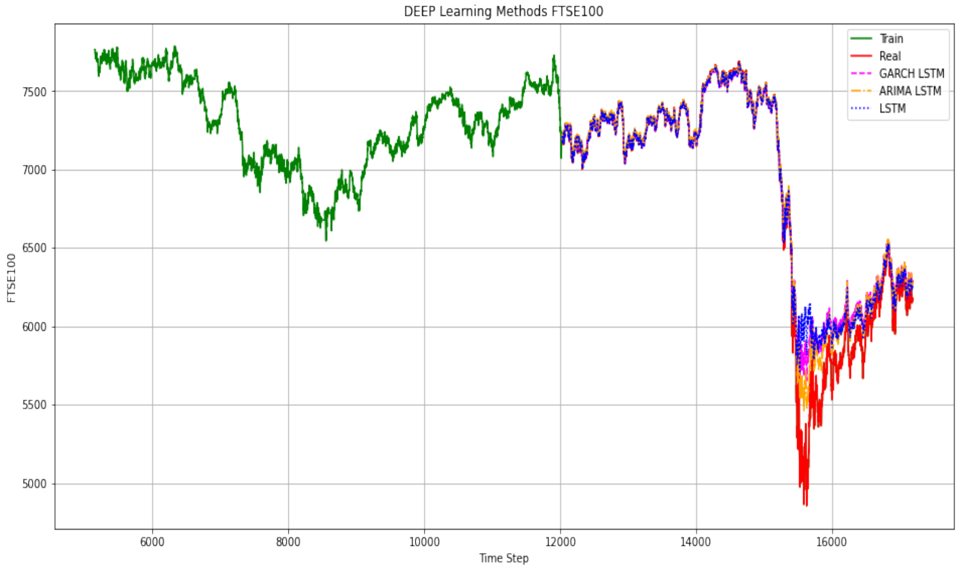


FIGURE 4.28: FTSE Deep Learning Forecasting (Hourly Data)

For Deep learning methods, LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM are used and computed results for the FTSE stock market index have given in Table 4.6. For FTSE stock market index LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 45474.31, 109.63, 1.92, and 213.25 respectively, which outperform the conventional machine learning models and the hybrid ARIMA-SVM models. For the hybrid, ARIMA-LSTM model forecasting MSE, MAE, MAPE, and RMSE have the values as 37013.00, 113.73, 1.94, and 192.39 respectively. Similarly, for the hybrid GARCH-LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values 35550.13, 96.84, 1.70, and 188.55 respectively. Figure 4.28 exhibit the train, test, and predicted series of FTSE. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta color represents LSTM, ARIMA-LSTM, and GARCH-LSTM respectively. Results shows that GARCH-LSTM model is closest to actual the time series so it is the best fitted deep learning method.

For DJI30 stock market Index SVM, model have forecasting accuracy indicators MSE, MAE, MAPE, and RMSE as 3621003.30, 1481.35, 5.9825, and 1902.89 respectively. The hybrid ARIMA-SVM model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have the values 3562240.35, 1456.96, 6.0070, and 1887.39 respectively, which improve the results indicates simple ARIMA-SVM.

MSE, RMSE, and MAE values show that the ARIMA-SVM model is a better predictor than the SVM for the DJI30 stock market. Figure 4.29 exhibit the actual and forecasted SVM and ARIMA-SVM models for DJI30.

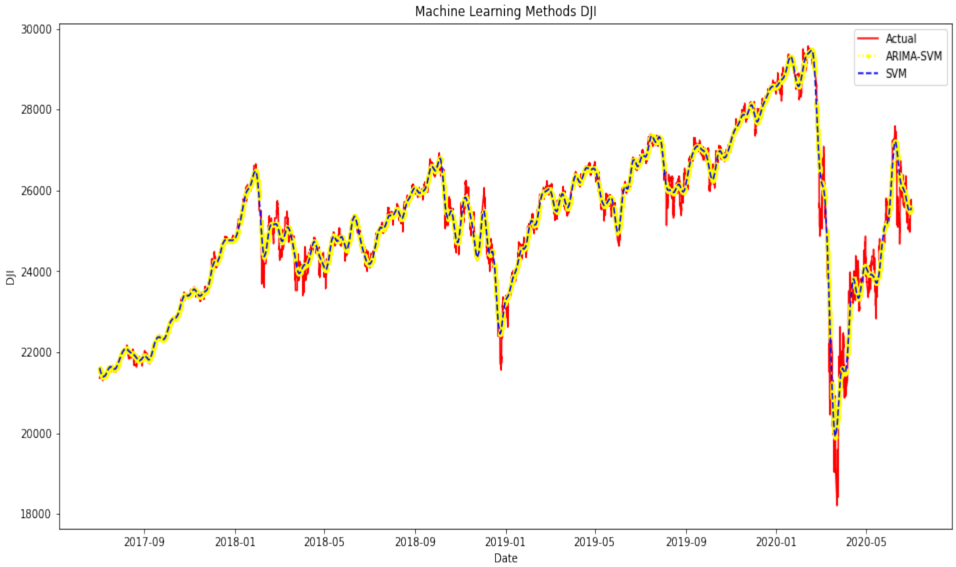


FIGURE 4.29: DJI Machine Learning Forecasting (Hourly Data)

For Deep learning methods, LSTM, hybrid ARIMA-LSTM ARIMA-LSTM, and GARCH-LSTM are used and results for DJI30 stock market index have been reported in Table 4.6. For the DJI30 stock market index, LSTM estimated model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 39937.35, 144.76, 0.56, and 199.84 respectively. Which outperform the machine learning models. The hybrid ARIMA-LSTM computed model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 30154.95, 122.24, 0.49, and 173.65 respectively. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 30279.12, 104.03, 0.42, and 174.01. Figure 4.30 exhibit the train, test, and predicted series for DJI30. Test results shows proposed GARCH-LSTM model has the lowest forecasting indicators values and closest to actual time series, so it is the best model for DJI30 stock market index.

For the SNP500 stock market, SVM model forecasting accuracy indicators have the values as 49849.27, 177.50, 6.3455, and 223.27 respectively for MSE, MAE, MAPE, and RMSE. For The hybrid ARIMA-SVM accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 41856.34, 154.56, 6.3379, and 204.59 respectively,

FIGURE 4.30: DJI Deep Learning Forecasting (Hourly Data)

which improves the results of machine learning methods with the hybrid technique. Results indicate that ARIMA-SVM performs better in machine learning regression than the simple SVM model. Figure 4.31 exhibit the actual and forecasted SVM and ARIMA-SVM models for SNP500.

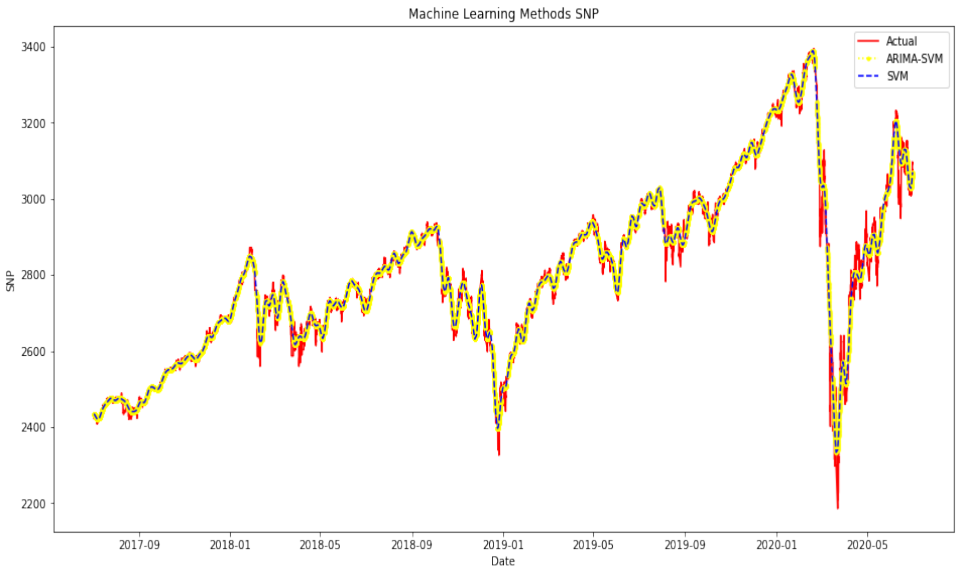


FIGURE 4.31: SNP Machine Learning Forecasting (Hourly Data)

The The deep learning methods LSTM, and the hybrid LSTM along with ARIMA and GARCH results are presented in Table 4.6 for all selected stock markets. For the SNP500 stock market index, LSTM estimated model have accuracy indicators

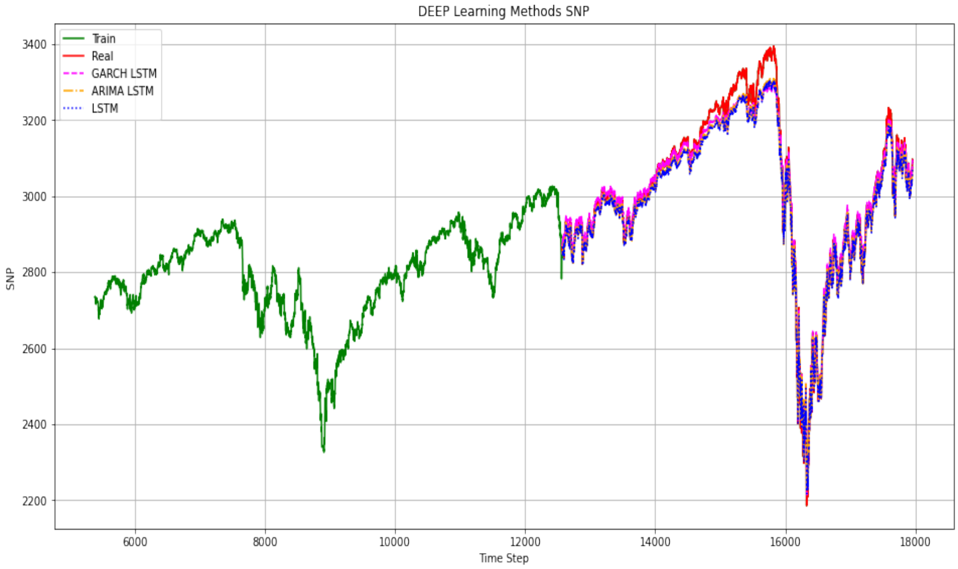


FIGURE 4.32: SNP Deep Learning Forecasting (Hourly Data)

as MSE, MAE, MAPE, and RMSE with the values 1161.13, 27.00, 0.87, and 34.08 respectively, which outperform the machine learning models. For the hybrid, the ARIMA-LSTM method computed forecasting accuracy indicators have the MSE, MAE, MAPE, and RMSE values of 960.28, 23.74, 0.77, and 30.99. Similarly, for the hybrid GARCH-LSTM method have the values of 869.78, 20.14, 0.65, and 29.49 for MSE, MAE, MAPE, and RMSE respectively.

Figure 4.32 exhibit the train, test, real and forecasted series of SNP500. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta colore represent the LSTM, ARIMA-LSTM, and GARCH-LSTM results respectively. Results shows that proposed GARCH-LSTM with the lowest forecasting accuracy indicators is the best fitted deep learning method.

This section concludes the estimation and forecasting of machine learning and deep learning methods to answers the defined research questions in chapter one. The study use SVM, hybrid ARIMA-SVM, LSTM, hybrid ARIMA-LSTM, and hybrid GARCH-LSTM models to forecast the stock market indices. Results shows a mixed result between the battle of deep learning methods. Proposed GARCH-LSTM and ARIMA-LSTM show better forecasting results for hourly data frequency then the other SVM, LSTM and hybrid ARINA-LSTM methods.

4.1.5 Discussion

This section discuss the results of classical, machine learning, and deep learning methods where hourly data is used for analysis. EMH states that the prices are linked with information. As information arrives randomly in market so, stock markets follows the random walk pattern (Copeland, 1976; Längkvist et al., 2014). However, their are multiple studies that provide evidence about the weak-form efficiency of the market. These studies conclude that prices are predictable by using historical prices or previous information. Therefore market players identify and react according to the market patterns. The first objective is to find out best performing classical forecasting model forecasting. Results in Table 4.3 shows that information has a minimal impact on current stock market prices. Which shows these markets has followed the random walk pattern. For the generalization of ARFIMA model which has the coefficient value with in -0.5 < d< 0.5 specified by Baillie (1996) none of our sample market achieved. Below generalized d parameter value a minimal impact of fractional differencing on current prices of markets and consistent with studies of (Floros et al., 2007; Reisen et al., 2001). Which also support the EMH stance of randomness in markets. This clearly provides that classical models fail to identify the pattern in hourly prices for sample countries during the sample period.

Table 4.5 and Table 4.6, compare model on basis of forecasting accuracy of modified deep learning, machine learning and classical forecasting method proposed by various studies i.e.,(Alam et al., 2020; Bhardwaj and Swanson, 2006; Kewat et al., 2017; Li et al., 2020; Papacharalampous et al., 2019).This study introduce a GARCH based deep learning method to analyze the behavior stock market indices. Above stated Tables answer the questions ”Are classical models successful in forecasting price behavior in financial markets? How do machine leaning based models perform in forecasting price trends in financial markets? and Do machine learning based models outperform classical models of forecasting”. Table 4.7 present each model ranking wise based on analysis in above section.

Results shows that for the stock market indices A50, EUS, DAX, HIS, WIG, IBX, SMI, FTSE, DJI30, and SNP500 GARCH-LSTM out perform the other deep learning, machine learning and classical forecasting methods suggested by the (Li et al., 2020). For CAC, NIFTY, AEX and SSGF stock markets ARIMA-LSTM consistent with Alam et al. (2020); Kulshreshtha (2020); Li et al. (2020) statement of improving prediction power with hybrid ARIMA and LSTM is the most suitable model. Where Nikkei is only stock market index, which is better predictable with LSTM method in comparison to the hybrid ARIMA and GARCH lstm method. For stock market indices ASX, CAC, and SSGF GARCH-LSTM performed as second-best method to forecast the respective stock market. Whereas, A50, HIS, NIKKEI, WIG, SMI, DJI30 and SNP500 have ARIMA as second best fitted forecasting method. Nine out of 16 markets show the LSTM as 3rd best fitted model to forecast financial time series. In classical forecasting techniques we witnessed ARIMA as best fitted candidate for the classical forecasting techniques for high frequency data. ARIMA model perform better than the GARCH in out of sample forecasting (Crawford and Fratantoni, 2003). The combined GARCH-LSTM model help to overcome the weakness of GARCH model in out of sample forecasting.

The findings of the study reveal that, Deep learning-based methods outperform the conventional machine learning and classical forecasting technique which are consistent with the finds (Choi, 2018; Fang and Yuan, 2019; Fischer and Krauss, 2018; Kulshreshtha, 2020; Li et al., 2020; Temür et al., 2019; Yan and Ouyang, 2018). Aforementioned forecasting indicators MSE, RMSE, MAE and MAPE values in section 4.1.3.2 and section 4.1.4.2 clearly provide that hybrid/deep learning-based methods performs better than other methods that is align with our research objectives and accept our hypothesis H3 of deep learning outperform conventional machine learning and classical forecasting methods.

Table 4.7: Models Comparison Hourly Data

	Model Rank							
	1st RANK	**2nd RANK**	**3rd RANK**	**4th RANK**	**5th RANK**	**6th RANK**	**7th RANK**	**8th RANK**
ASX	ARIMA-LSTM	GARCH-LSTM	LSTM	SVM	ARIMA-SVM	ARIMA	ARFIMA	GARCH
A50	GARCH-LSTM	ARIMA-LSTM	LSTM	ARIMA-SVM	SVM	ARIMA	GARCH	ARFIMA
EUS	GARCH-LSTM	LSTM	ARIMA-LSTM	ARIMA-SVM	SVM	ARFIMA	ARIMA	GARCH
CAC	ARIMA-LSTM	GARCH-LSTM	LSTM	ARIMA-SVM	SVM	ARIMA	ARFIMA	GARCH
DAX	GARCH-LSTM	LSTM	ARIMA-LSTM	ARIMA-SVM	SVM	ARIMA	ARFIMA	GARCH
HSI	GARCH-LSTM	ARIMA-LSTM	LSTM	ARIMA-SVM	SVM	ARIMA	ARFIMA	GARCH
NIFTY	ARIMA-LSTM	LSTM	GARCH-LSTM	SVM	ARIMA-SVM	ARIMA	ARFIMA	GARCH
NIKKEI	LSTM	ARIMA-LSTM	GARCH-LSTM	ARIMA-SVM	SVM	ARIMA	ARFIMA	GARCH
AEX	ARIMA-LSTM	LSTM	GARCH-LSTM	SVM	ARIMA-SVM	ARFIMA	ARIMA	GARCH
WIG	GARCH-LSTM	ARIMA-LSTM	LSTM	ARIMA-SVM	SVM	ARIMA	GARCH	ARFIMA
SSGF	ARIMA-LSTM	GARCH-LSTM	LSTM	ARIMA-SVM	SVM	ARIMA	ARFIMA	GARCH
IBX	GARCH-LSTM	LSTM	ARIMA-LSTM	ARIMA-SVM	SVM	ARIMA	ARFIMA	GARCH
SMI	GARCH-LSTM	ARIMA-LSTM	LSTM	ARIMA-SVM	SVM	ARIMA	ARFIMA	GARCH
FTSE	GARCH-LSTM	LSTM	ARIMA-LSTM	ARIMA-SVM	SVM	ARFIMA	ARIMA	GARCH
DJI	GARCH-LSTM	ARIMA-LSTM	LSTM	ARIMA-SVM	SVM	ARIMA	ARFIMA	GARCH
SNP	GARCH-LSTM	ARIMA-LSTM	LSTM	ARIMA-SVM	SVM	ARFIMA	ARIMA	GARCH

4.2 Section II (10 Minutes Data Frequency)

In this Section, all of the classical, machine learning, and deep learning methods are examined using 10 minutes data frequency. This section of the study is further divided into five sub-sections. Section one and two deal with the data descriptives and pre-processing. Section two and three discuss the estimated results of all models and the section discuss the models comparison.

4.2.1 Descriptive Statistics

TABLE 4.8: Descriptive Statistics (10 Minutes Data)

	Mean	**Median**	**Std. Dev.**	**Skew**	**Kurt**	**J.B**
ASX	6124.42	6071.47	442.21	0.07	2.88	132.47
A50	12724.12	12872.98	1063.57	-0.34	2.07	4821.81
EUS	3399.05	3443.37	244.87	-1.11	4.72	21473.77
CAC	5305.92	5358.28	393.27	-0.75	4.38	11406.52
DAX	12198.24	12313.21	887.06	-1.03	4.61	29678.93
HSI	27722.04	27788.17	2089.09	-0.08	2.84	137.06
NIFTY	10931.03	10919.25	861.03	-0.82	3.98	10646.35
NIKKEI	21719.81	21802.97	1346.10	-0.57	3.40	6259.17
AEX	549.70	552.36	34.04	-0.61	4.79	11052.93
WIG	2185.08	2238.44	236.83	-1.53	5.28	19780.96
SSGF	361.79	364.93	30.34	-1.06	4.09	19655.34
IBX	9302.83	9410.91	943.23	-1.40	4.93	27145.86
SMI	9400.80	9280.97	620.46	0.68	2.90	5023.55
FTSE	7190.07	7331.45	518.69	-1.94	6.65	122028.20
DJI	25203.80	25305.32	1886.91	-0.21	3.05	743.35
SNP	2806.84	2794.15	220.76	0.34	2.82	2218.91

Table 4.8 depicts the data summary of 10-minutes data frequency. Mean values shows the average indices points for the selected time period. ASX, A50, EUS, CAC, DAX, HIS, NIFTY, NIKKEI, AEX, WIG, SSGF, IBX, SMI, FTSE, DJI 30 and SNP 500 has the mean values of 6124.42, 12724.12, 3399.05, 5305.92, 12198.24, 27722.04, 10931.03, 21719.81, 549.70, 2185.08, 361.79, 9302.83, 9400.80, 7190.07, 25203.80, and 2806.84 respectively. Similarly, standard deviations explain the data desperation from its mean. To measure the location of data we generally interpret

the skewness and kurtosis figures. Skewness explain the left are right skewed of data observations. In our data set A50, EUS, CAC, DAX, HIS, NIFTY, NIKKEI, AEX, WIG, SSGF, IBX, FTSE, and DJI 30 has negative values (-0.34 for A50) which referred as left skewed data. Whereas ASX, SMI and SNP 500 are right tailed with positive skewness values (0.07, 0.68, 0.34) respectively. ASX, A50, HIS, SMI, DJI30, and SNP500 have the platykurtic Behavior and rest of sample has the leptokurtic behavior having kurtosis values greater than 3. Jarque-Bera normality test indicates the none of variable is normally distributed.

4.2.2 Data Pre-Processing

Before proceeding to our 10-minute frequency data again preprocessing has been done to estimate the aforementioned models of forecasting. To check the stationarity of data ADF has used. Results show the non-normality of data with insignificant p values presented in the table 4.9. After the first difference results showsthe significant p-values of the ADF test for all series. Similarly, the ARCH-LM test has also been applied on series to check whether data has the ARCH effect. Results are given in the table 4.9 show existence of the ARCH effect in data. Figure B-1 in appendix KDE plot also depict the data distribution outlook.

4.2.3 Classical Forecasting Methods

This analysis section has been further divided into two sessions: Model Estimation and selection, and Forecasting.

4.2.3.1 Models Selection and Estimation

Based on predefined tool and parameters all of selected classical method has been applied on 10 minutes data frequency of all series. Table 4.10 depicts the estimated results of the ARIMA model for all the data sets. Table 4.11 and table 4.11 shows the estimated results for ARIMA and GARCH models. 70:30 ratios have been used for training and test data.

Table 4.9: Data Processing (10 Minutes Date)

	Stationarity			ARCH Lm Test
	Level	1st Diff.	2nd Diff.	p-value
ASX	0.1435	0.0000	-	0.0000
A50	0.1903	0.0000	-	0.0000
EUS	0.1593	0.0000	-	0.0000
CAC	0.1942	0.0000	-	0.0000
DAX	0.1465	0.0000	-	0.0000
HSI	0.2801	0.0000	-	0.0000
NIFTY	0.2524	0.0000	-	0.0000
NIKKEI	0.1374	0.0000	-	0.0000
AEX	0.0911	0.0000	-	0.0000
WIG	0.7404	0.0000	-	0.0000
SSGF	0.6978	0.0000	-	0.0000
IBX	0.7167	0.0000	-	0.0000
SMI	0.1379	0.0000	-	0.0000
FTSE	0.4483	0.0000	-	0.0000
DJI	0.0532	0.0000	-	0.0000
SNP	0.1272	0.0000	-	0.0000

Estimated results shows random walk of prices (ARIMA order (0,1,0) has been found in A50, CAC, DAX, HIS, NIKKEI, WIG, IBX, and SMI stock index based on AIC values of 511381, 234125, 270717, 537436, 451959, 364931, 642242, 54021, 119472, 31358, 292095, and 310943 respectively. None of the above-mentioned stock indices found any AR and MA term, as well as no seasonality (SAR and SMA), has been found in stepwise model computation. Sigma2 value has the coefficient values of 238.242, 9.587, 21.430, 89.819, 1031.088, 93.899, 376.701, 0.231, 10.776, 0.101, 99.689, and 52.887 respectively for A50, CAC, DAX, HIS, NIKKEI, WIG, IBX, and SMI stock index with significant p-values. KDE plot of residuals for estimated ARIMA model also has been present in the Figure B-2 appendix for the aforementioned series. This verifies the model's fitness.

ASX stock market index shows that the ARIMA order (2,1,0) as the best-fitted model with AIC value of 371431 and BIC value of 371458. AR term has a coefficient value of 0.0012 with a significant p-value. This means the existence of a pattern in the ASX market index. Estimated results do not show any of MA terms which means any delayed information adjustment. No SAR and SMA have been identified which refers to no seasonality in data. KDE plot of residuals shown in figure B-3 in appendix validates smoothen the data after applying the model.

Based on estimated AIC values 234125, 364931, 54021 and BIC values 234142, 364948, and 54038 for EUS, NIFTY, and AEX stock market indices ARIMA (0,1,1) has been identified as the best-fitted model for respective stock market indices. Estimated results given in table 11 show the coefficient value of MA term 0.023, 0.031, and 0.020 with significant p-value respectively for the aforementioned stock markets. This shows the delayed adjustment of information. Results show significant existence of variance with a significant p-value of sigma2 with coefficient values of 9.587, 93.899, and 0.231 respectively. KDE Plot of residuals also been presented in appendix (B-3) that shows the model fitness.

Based on results, ARIMA (2,1,1) model has been identified for SSGF stock market indices on basis of minimum AIC value (31358) and BIC value (31394). AR term has the coefficient value of 0.044 with a significant p-value (0.0000) which indicates a positive relationship with lag values and SSGF follows the patterns. Whereas the MA term with a coefficient value of 0.955 and a significant p-value indicates the positive relationship with lag values of the error term. Results shows that delayed informational adjustment in SSGF. There is no seasonality (SAR and SMA) term that has been identified. Sigma2 coefficient values of 0.10 with significant p-values indicate the existence of variance in the mean equation of the model. Figure B-3 in the appendix Also indicates the model fitness which smoothens the KDE of residuals.

For stock market indices FTSE and DJI 30 Based on AIC values (429538 and 638734 respectively) and BIC values (429566 and 638771 respectively) ARIMA (1,1,1) model has been identified as the best-fitted model for both indices. AR term for both models has a significant p-value with coefficient values of 0.577 and -0.972. This means the existence of patterns in financial series and significant relation with its own lag values. MA term also significant p-values with coefficient values of -0.592 and 0.968 respectively. FTSE stock market index with negative coefficient value shows the overly price adjustment. Whereas, DJI 30 has a significant coefficient value refers as information adjustment.

TABLE 4.10: ARIMA Estimations

	ASX	**A50**	**EUS**	**CAC**	**DAX**	**HSI**	**NIFTY**	**NIKKEI**	**AEX**	**WIG**	**SSGF**	**IBX**	**SMI**	**FTSE**	**DJI**	**SNP**
Method								Log Likelihood								
Model	(2,1,0)	(0,1,0)	(0,1,1)	(0,1,0)	(0,1,0)	(0,1,0)	(0,1,1)	(0,1,0)	(0,1,1)	(0,1,0)	(2,1,1)	(0,1,0)	(0,1,0)	(1,1,1)	(1,1,1)	(2,1,2)
AIC	371431	511381	234125	270717	537436	451959	364931	642242	54021	119472	31358	292095	310943	429538	638734	308908
BIC	371458	511390	234142	270726	537445	451968	364948	642251	54038	119480	31394	292103	310952	429566	638771	308954
AR	-0.012	-	-	-	-	-	-	-	-	-	0.044	-	-	0.577	-0.972	0.907
	[0.000]	-	-	-	-	-	-	-	-	-	[0.000]	-	-	[0.000]	[0.000]	[0.000]
MA	-	-	-0.023	-	-	-	-0.031	-	-0.020	-	-0.955	-	-	-0.592	0.968	-0.901
	-	-	[0.000]	-	-	-	[0.000]	-	[0.000]	-	[0.000]	-	-	[0.000]	[0.000]	[0.000]
SAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SMA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sigma2	14.843	238.242	9.587	21.430	89.819	1031.088	93.899	376.701	0.231	10.776	0.101	99.689	52.887	22.528	359.091	3.974
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Obs	67101	61530	45923	45864	73264	46231	49449	73238	39333	22909	57944	39261	45687	72160	73238	73238

Stepwise models shows no SAR and SMA seasonality patterns in 10 minutes of data frequency in both FTSE and DJI30 index. KDE plot of residuals shows the model fitness. Significant values of Sigma2 with coefficient values of 22.528 and 359.091 for both markets show the presence of variance in markets.

SNP 500 market index has estimated results presented in table 4.10 show the ARIMA (2,1,2) order as best fitted model based on AIC value (308908) and BIC value (308954). Both AR and MA terms have significant p values with coefficients value of 0.907 and -0.901 respectively. Both presences of autocorrelation in lag value of SNP 500 and negative MA term mean the information is overly priced in 10-minute data frequency. No seasonality pattern has been identified in the respective stock market index. A significant Sigma2 value with a coefficient value of 3.974 shows the existence of variance in SNP 500 market index. KDE plot of residuals in B-3 in appendix also validates the model fitness.

Based on results presented in table 4.11 for ARFIMA model estimation for all the stock markets indices. All of the markets follow the same p and q order for model estimation resulted in ARIMA model estimation given in table 4.10. For the stock market indices A50, CAC, DAX, HIS, NIKKEI, WIG, IBX, and SMI computed AIC values has been -10.53, -11.13, -11.21, -11.44, -10.73, -11.22, -11.19, -10.15, -9.62, -11.28, -10.89, and -11.41 respectively. Estimated d parameter values 0.0032, -0.0084, -0.0059, -0.0046, 0.0034, -0.0210, -0.0163, -0.0150, 0.0128, -0.0202, -0.0067, and -0.0047 respectively for all aforementioned markets.

A50 and WIG process long memory in autocovariance function. Whereas CAC, DAX, NIKKEI, WIG, and IBX follows the intermediate memory (StataCorp, 2019) exist in respective markets. HIS and SMI both have insignificant p-values and on basis of this, we can conclude that ARFIMA does not fit in these two markets.

ASX stock market index which identified ARIMA (2,1,0) based on this p and q order ARFIMA model has been computed with an estimated AIC value of -11.85. The estimated d parameter has the values of -0.0120 with a significant p-value of (0.0044). Results of classical ARIMA model estimation shows the intermediate memory presence in the ASX Stock market index.

TABLE 4.11: ARFIMA and GARCH Estimations(10 Minutes Data)

	ASX	A50	EUS	CAC	DAX	HSI	NIFTY	NIKKEI	AEX	WIG	SSGF	IBX	SMI	FTSE	DJI	SNP
								AFRIMA Statistics								
Method								Log Likelihood								
Model	(2,1,0)	(0,1,0)	(0,1,1)	(0,1,0)	(0,1,0)	(0,1,0)	(0,1,1)	(0,1,0)	(0,1,1)	(0,1,0)	(2,1,1)	(0,1,0)	(0,1,0)	(1,1,1)	(1,1,1)	(2,1,2)
AIC	-11.85	-10.53	-11.13	-11.21	-11.44	-10.73	-11.22	-11.19	-10.15	-9.62	-11.28	-10.89	-11.41	-11.80	-11.49	-11.55
d	-0.0120	0.0032	-0.0084	-0.0059	-0.0046	0.0034	-0.0210	-0.0163	-0.0150	0.0128	-0.0202	-0.0067	-0.0047	-0.0060	0.0029	0.0073
	[0.0044]	[0.019]	[0.039]	[0.006]	[0.001]	[0.133]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]	[0.028]	[0.055]	[0.342]	[0.001]	[0.142]
AR	-0.0067	-	-	-	-	-	-	-	-	-	-0.64	-	-	0.49	-0.97	0.50
	[0.039]	-	-	-	-	-	-	-	-	-	[0.026]	-	-	[0.007]	[0.000]	[0.000]
MA	-	-	-0.01	-	-	-	-0.02	-	0.01	-	0.61	-	-	-0.50	0.97	-0.51
	-	-	[0.006]	-	-	-	[0.000]	-	[0.000]	-	[0.032]	-	-	[0.006]	[0.000]	[0.000]
								GARCH Statistics								
Method								Maximum Likelihood ARCH								
AIC	-234012	-100752	-94895	-2735	-186495	1385	-106444	-164232	176	-24879	-4500	-68090	-3832	-206555	-12583	-214670
Lag	-0.038	-0.023	-0.045	-0.031	-0.020	-0.017	-0.031	-0.0235	-0.051	-0.024	-0.083	-0.016	-0.029	-0.020	-0.020	-0.030
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]
α_1	0.056	0.361	0.164	0.166	0.215	0.097	0.165	0.036	0.182	0.193	0.187	0.005	0.173	0.237	0.233	0.241
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
β_1	0.924	0.581	0.743	0.743	0.729	0.850	0.656	0.948	0.789	0.770	0.657	0.993	0.730	0.705	0.709	0.713
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Obs	67101	61530	45923	45864	73264	46231	49449	73238	39333	22909	57944	39261	45687	72160	73238	73238

AR term has the coefficient value of -0.0067 with a significant p-value of (0.039) which indicates the persistence of autocorrelation with lag values. For EUS, NIFTY, and AEX market indices AIC values -11.13, -11.22, and -10.15 respectively mentioned in table 4.11.Estimated d parameter values are -0.0084, -0.0210, and -0.0150 with significant p-values respectively. Estimated coefficient values also show the intermediate memory presence. Significant p-values of MA term coefficient values -0.01, -0.02, and 0.01 indicates the informational adjustment in respective markets and overly priced information in EUS and NIFTY stock market indices.

For the remaining Stock markets indices SSGF, FTSE, DJI 30 and SNP 500 followed the same p and q order presented in table 4.11 with AIC values of -11.28, -11.80, -11.49, and -11.55 respectively for all indices. Estimated d parameter has the values of -0.0202, -0.0060, 0.0029, 0.0073 respectively. FTSE and SNP 500 both have insignificant d parameters p-values. AR term has significant p-values with a coefficient value of -0.64, 0.49, -0.97, and 0.50 indicates the autocorrelation with lag values of series. Similarly, significant p-values of MA term with coefficient values of 0.61, -0.50, 0.97, and -0.51 indicates the adjustment of shocks in prices respectively. A negative sign with FTSE and SNP500 indicates the overly priced information in price.

For the non-parametric model in the classical forecasting method GARCH (1,1) again has been estimated in 10-minute time interval data. Estimated AIC values for all market has been given in table 4.11. ASX, A50, EUC, CAC, DAX, HSI, NIFTY, and NIKKEI market indices have the lag coefficient values of -0.038, -0.023, -0.045, -0.031, -0.020, -0.017, -0.031, and -0.0235 with significant p-values respectively. Results indicate the inverse relation with lag returns of these stock markets. α_1 have a coefficient values of 0.056, 0.361, 0.164, 0.166, 0.215, 0.097, 0.165, and 0.036 for aforementioned stock indices respectively with significant p-value. Similarly, β_1 coefficients values 0.924, 0.581, 0.743, 0.743, 0.729, 0.850, 0.656, and 0.948 with significant p-value. These statistical values show the existence of the ARCH effect and persistence of GARCH in respective stock market indices.

For stock indices AEX, WIG, SSGF, IBX, SMI, FTSE, DJI30, and SNP500 estimated for the GARCH (1,1) model has been presented in table 4.11. Estimated AIC values have been presented as well. Aforementioned market indices have the lag coefficient values of -0.051, -0.024, -0.083, -0.016, -0.029, -0.020, -0.020, and -0.030. The significant p-value of these coefficients indicates the existence of an inverse relationship between the lag returns and prices of these stock market indices. Coefficient values of α_1 indicate the existence of the ARCH effect in markets. Coefficient values 0.182, 0.193, 0.187, 0.005, 0.173, 0.237, 0.233, and 0.241 with significant p-values support our argument of volatility existence. Similarly, β_1 coefficients show the persistence of GARCH in data. Coefficient values of GARCH 0.770, 0.657, 0.993, 0.730, 0.705, 0.709, and 0.713 with significant p-value validate the argument of GARCH persistence in given data set.

4.2.3.2 Models Forecasting

Model estimation has been in done previous section. All of the models have been trained using 70% of data. And 30% data has been used in this section of analysis to forecast the all above estimated models. Forecasted results has been presented in table 4.12 and presentation of respective stock market.

Figure B-3 in appendix exhibit the forecast of estimated trained and test ARIMA model for ASX stock market index. Forecasting accuracy indicators (MSE, MAPE, RMSE, and MAE) results indicates the forecasting accuracy of model in Table 4.12. MSE (515222), RMSE (717.790), MAE (496.776), and MAPE (8.81229). Results shows the sudden drop in forecasted windows due failed the ARIMA forecasting method to predict in such case as arrival of Covid-19 in this time frame. Whereas, Figure B-4 in appendix also exhibit the both actual and forecasted ARFIMA series with Estimated MSE (284599), RMSE (533.478), MAE (422.974), and MAPE (6.69259). For the third non-parametric conventional model GARCH we used 100 observations as out of sample as estimation window using R ruGARCH library. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in Figure B-5 in appendix.

Table 4.12: Classical Forecasting (10 Minutes Data)

	ASX	A50	EUS	CAC	DAX	HSI	NIFTY	NIKKEI	AEX	WIG	SSGF	IBX	SMI	FTSE	DJI	SNP
	Forecasting indicators for ARIMA															
MSE	515222	585996	147982	419535	1763065	5859755	2184277	4638649	2822	140236	1895	1434786	356242	708319	6593601	57742
MAE	496.78	669.35	336.61	581.71	1141.82	1933.65	1171.34	1855.81	39.21	265.15	31.82	959.29	455.59	584.50	2011.05	191.12
MAPE	8.81	4.85	10.29	11.27	9.61	7.85	12.06	8.40	7.61	16.11	10.54	12.64	4.55	9.72	8.15	6.48
RMSE	718	766	385	648	1328	2421	1478	2154	53	374	44	1198	597	842	2568	240
	Forecasting indicators for ARFIMA															
MSE	284599	428363	147050	421155	1838378	7950877	1926944	4358512	1147	42903	1502	1099780	383238	722039	9594271	67331
MAE	422.97	557.56	335.53	582.01	1123.32	2249.32	1140.77	1619.77	24.95	173.81	29.57	915.54	469.89	675.37	2384.13	206.7
MAPE	6.69	4.09	10.14	11.32	9.96	9.14	11.42	7.97	4.67	8.15	9.70	11.48	4.80	11.09	9.89	7.14
RMSE	533	654	383	649	1356	2820	1388	2088	34	207	39	1049	619	850	3097	259
	Forecasting indicators for GARCH															
MSE	1547213	15015982	390146	909647	4147353	39164066	2551607	11325709	13522	30236	5299	1509787	2293583	1518021	20906897	115130
MAE	875.40	3046.89	433.51	698.18	1518.21	5303.28	1221.85	2397.91	95.23	127.48	52.57	973.19	1116.28	877.90	3153.34	261.1
MAPE	15.52	22.79	14.55	14.79	14.05	21.27	12.54	12.00	17.74	6.50	17.46	12.88	11.57	14.82	13.32	9.20
RMSE	1244	3875	625	954	2037	6258	1597	3365	116	174	73	1229	1514	1232	4572	339

Forecasting Accuracy indicators are MSE (1547213), RMSE (1243.870), MAE (875.399), and MAPE (15.51600). Based on results and minimum values of evaluation criteria of models ARFIMA model with lowest MSE, MAE, MAPE and RMSE values has better prediction power then the ARIMA and GARCH model in 10-minute time frequency for ASX stock market index.

For the stock market A50 index, both trained and test series have been shown in Figure B-3 in appendix for our first conventional classical forecasting model ARIMA. The lower and upper bound shows the perfection of ARIMA for the A50 market index. Which is the result of Chinese government policies to counter the economic effect due to covid-19 (Hofman et al., 2020). Forecasting Indicators MSE, MAE, MAPE, and MAPE have the values of 585996, 669.345, 4.85104, and 765.503 respectively presented in Table 4.12. ARFIMA estimated accuracy indicators MSE, MAE, MAPE, and RMSE presented in Table 4.12. Computed forecasting accuracy indicators have the values of, MSE (428363), MAE (557.561), MAPE (4.08930), and RMSE (654.494). Figure B-4 in appendix exhibit the ARFIMA forecasted and actual series for the A50 stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in Figure B-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 15015982, 3046.892, 22.78806, and 3875.046. Computed values of Accuracy indicator of ARFIMA model declare the winner within the classical forecasting techniques for 10-minute data frequency.

Forecasting Indicators of classical Models for EUS stock market index has been presented in Table 4.12. Estimated MSE, MAE, MAPE, and RMSE has the statistical values of 147982, 336.609, 10.28767, and 384.685 respectively for the EUS stock market index. Similarly, a Graphical representation has been given in figure B-3 in appendix. Which shows a sudden fall decreases the ARIMA model prediction power in the testing phase. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 147050, 335.530, 10.13763, and 383.471 Respectively for the EUS stock market index given in Table 4.12. Figure B-4 in appendix depicts the visual outlook of the actual and forecasted series. Our

3rd classical model GARCH, computed forecasting indicators has been presented in Table 4.12. The model choose 100 observations to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure B-5 in appendix. Computed forecasted indicators MSE, MAE, MAPE, and RMSE have the values of 390146, 433.513, 14.55010, and 624.616 respectively for the EUS stock market index. The aforementioned evolution indicators result shows the ARIMA model having the lowest indicator values within the classical forecasting models. That makes ARIMA as best fitted forecasting method in classical forecasting models for EUS stock market index under 10-minute data frequency.

For the stock market index, CAC, both trained and test series have been shown in Figure B-3 in appendix for the predefined ARIMA model. Similarly, predefined lower and upper bound also show the ARIMA model forecasting strength for the CAC market index. Forecasting accuracy indicators, MSE, MAE, MAPE, and RMSE have the values as 419535, 581.713, 11.27040, and 647.715 respectively for the CAC stock market index. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 421155, 582.014, 11.32391, and 648.964 respectively for the CAC stock market index given in Table 4.12. Figure B-4 in appendix depicts the visual outlook of the actual and forecasted series. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in Figure B-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 15015982, 3046.892, 22.78806, and 3875.046. Based on computed values of the Accuracy indicator of All models ARIMA model has the lowest values of these evaluation criteria. Which declares the ARIMA as the winner in the classical forecasting techniques for the CAC stock market index.

Figure B-3 in appendix exhibit the forecast of estimated trained and test ARIMA model for the DAX stock market index. Results of forecasting accuracy indicators (MSE, MAPE, RMSE, and MAE) are presented in Table 4.12. MSE (1763065), RMSE (1327.805), MAE (1141.816and MAPE (9.60626). The Results show the sudden drop in forecasted windows due to failed the ARIMA forecasting method to predict in such case as the arrival of Covid-19 in this time frame. Whereas, figure B-4 in appendix also exhibit the both actual and forecasted ARFIMA series with Estimated MSE (1838378), RMSE (1355.868), MAE (1123.320), and MAPE (9.95661). For the third non-parametric conventional model GARCH we used 100 observations as out of sample as estimation window using R ruGARCH library.For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure B-5 in appendix. Forecasting Accuracy indicators are MSE (4147353), RMSE (2036.505), MAE (1518.206), and MAPE (14.05259). Based on results and minimum values of evaluation criteria of models ARFIMA model with the lowest MSE, MAE, MAPE, and RMSE values has better prediction power than the ARIMA and GARCH model in 10-minute time-frequency for the DAX stock market index.

For the stock market HSI index, both trained and test series have been shown in figure B-3 in appendix for our first conventional classical forecasting model ARIMA. The lower and upper bound shows the perfection of ARIMA for the HSI market index. Which is the result of Hong Kong government policies to counter the economic effect due to covid-19. Forecasting Indicators MSE, MAE, MAPE, and RMSE have the values of 5859755, 1933.652, 7.85193, and 2420.693 respectively presented in Table 4.12. ARFIMA estimated accuracy indicators MSE, MAE, MAPE, and RMSE presented in Table 4.12. Computed forecasting accuracy indicators have the values of, MSE (7950877), MAE (2249.323), MAPE (9.14002), and RMSE (2819.730). Figure B-4 in appendix exhibit the ARFIMA forecasted and actual series for the HSI stock market.

For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure B-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 39164066, 5303.279, 21.27121, and 6258.120. Computed values of Accuracy indicator of ARFIMA model declared the winner within the classical forecasting techniques for 10-minute data frequency.

Forecasting Indicators of classical Models for NIFTY stock market index has been presented in Table 4.12. Estimated MSE, MAE, MAPE, and RMSE has the statistical values of 2184277, 1171.338, 12.06213, and 1477.930 respectively for the NIFTY stock market index. Similarly, a Graphical representation has been given in figure B-3 in appendix. Which shows a sudden fall decreases the ARIMA model prediction power in the testing phase. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 1926944, 1140.769, 11.41932, and 1388.144 Respectively for the Nifty stock market index given in Table 4.12. Figure B-3 in appendix depicts the visual outlook of the actual and forecasted series. Our 3rd classical model GARCH, computed forecasting indicators has been presented in Table 4.12. The model choose 100 observations to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure B-3 in appendix. Computed forecasted indicators MSE, MAE, MAPE, and RMSE have the values of 2551607, 1221.846, 12.54059, and 1597.375 respectively for the NIFTY stock market index. The aforementioned evolution indicators result shows the ARIMA model having the lowest indicator values within the classical forecasting models. That makes ARIMA as best fitted forecasting method in classical forecasting models for NIFTY stock market index under 10minute data frequency.

For the stock market index, NIKKEI both trained and test series have been shown in figure B-3 in appendix for the predefined ARIMA model. Similarly, predefined lower and upper bound also show the ARIMA model forecasting strength for the NIKKEI market index. Forecasting accuracy indicators, MSE, MAE, MAPE, and RMSE have the values as 4638649, 1855.813, 8.40498, and, 2153.752 respectively for the NIKKEI stock market index. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 4358512, 1619.770, 7.96509, and 2087.705 respectively for the NIKKEI stock market index given in Table 4.12. Figure B-4 in appendix depicts the visual outlook of the actual and forecasted series. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in Figure B-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 11325709, 2397.913, 12.00205, and 3365.369. Based on computed values of the Accuracy indicator of All models ARFIMA model has the lowest values of these evaluation criteria. Which declares the ARFIMA as the winner in the classical forecasting techniques for the NIKKEI stock market index.

Figure B-3 in appendix exhibit the forecast of estimated trained and test ARIMA model for AEX stock market index. Results of forecasting accuracy indicators (MSE, MAPE, RMSE, and MAE) are presented in Table 4.12. MSE (2822), RMSE (53.123), MAE (39.213), and MAPE (7.60867). The Results show the sudden drop in forecasted windows due to failed the ARIMA forecasting method to predict in such case as the arrival of Covid-19 in this time frame. Whereas, Figure B-4 in appendix also exhibit the both actual and forecasted ARFIMA series with Estimated MSE (1147), RMSE (33.872), MAE (24.948), and MAPE (4.66897). For the third non-parametric conventional model GARCH we used 100 observations as out of sample as estimation window using R ruGARCH library. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for for the display of forecasting in Figure B-5 in appendix. Forecasting Accuracy indicators are MSE (13522), RMSE (116.286), MAE (95.234), and MAPE (17.74248). Based on results and minimum values of evaluation criteria of models ARFIMA model with the lowest MSE, MAE, MAPE, and RMSE values have better prediction power than the ARIMA and GARCH model in 10-minute time-frequency for the AEX stock market index.

For the stock market WIG index, both trained and test series have been shown in Figure B-3 in appendix for our first conventional classical forecasting model ARIMA. The Results show the sudden drop in forecasted windows due to failed the ARIMA forecasting method to predict in such case as the arrival of Covid-19 in this time frame. Forecasting Indicators MSE, MAE, MAPE, and MAPE have the values of 140236, 265.146, 16.11161, and 374.481respectively presented in Table 4.12. ARFIMA estimated accuracy indicators MSE, MAE, MAPE, and RMSE presented in table 4.12. Computed forecasting accuracy indicators have the values of, MSE (42903), MAE (173.807), MAPE (8.14845), and RMSE (207.130). Figure B-4 in appendix exhibit the ARFIMA forecasted and actual series for the WIG stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in Figure B-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 30236, 127.479, 6.49796, and 173.885. Computed values of Accuracy indicator of GARCH model declared the winner within the classical forecasting techniques for 10 minute data frequency.

Forecasting Indicators of classical Models for SSGF stock market index has been presented in Table 4.12. Estimated MSE, MAE, MAPE, and RMSE has the statistical values of 1895, 31.816, 10.54295, and 43.533 respectively for the SSGF stock market index. Similarly, a Graphical representation has been given in figure B-3 in appendix. Which shows a sudden fall decreases the ARIMA model prediction power in the testing phase. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 1502, 29.569, 9.69588, and 38.760 Respectively for the SSGF stock market index given in Table 4.12. Figure B-4 in appendix depicts the visual outlook of the actual and forecasted series. Our 3rd classical model GARCH, computed forecasting indicators has been presented in Table 4.12. The model choose 100 observations to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in Figure B-5 in appendix. Computed forecasted indicators MSE, MAE, MAPE, and RMSE have the values of 5299, 52.567, 17.45883, and 72.791 respectively for the SSGF stock market index. The aforementioned evolution indicators result shows the AFRIMA model having the lowest indicator values within the classical forecasting models. That makes ARFIMA as best fitted forecasting method in classical forecasting models for SSGF stock market index under 10minute data frequency.

For the stock market index, IBX both trained and test series have been shown in figure B-3 in appendix for the predefined ARIMA model. Similarly, predefined lower and upper bound also show the ARIMA model forecasting strength for the IBX market index. Forecasting accuracy indicators, MSE, MAE, MAPE, and RMSE have the values as 1434786, 959.288, 12.64001, and 1197.826 respectively for the IBX stock market index. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 1099780, 915.537, 11.48247, and 1048.704 respectively for the IBX stock market index given in Table 4.12. Figure

B-4 in appendix depicts the visual outlook of the actual and forecasted series. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in Figure B-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 1509787, 973.194, 12.88215, and 1228.734. Based on computed values of the Accuracy indicator of all models ARFIMA model has the lowest values of these evaluation criteria. Which declares the ARFIMA as the winner in the classical forecasting techniques for the IBX stock market index.

Forecasting Indicators of classical Models for SMI stock market index has been presented in Table 4.12. Estimated MSE, MAE, MAPE, and RMSE has the statistical values of 356242, 455.586, 4.55302, and 596.860 respectively for the SMI stock market index. Similarly, a Graphical representation has been given in figure B-3 in appendix. Which shows a sudden fall decreases the ARIMA model prediction power in the testing phase. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 383238, 469.892, 4.80433, and 619.062 respectively for the SMI stock market index given in above Table. Figure B-4 in appendix depicts the visual outlook of the actual and forecasted series. Our 3rd classical model GARCH, computed forecasting indicators has been presented in table.

Similarly forecasted series For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in Figure B-5 in appendix. Computed forecasted indicators MSE, MAE, MAPE, and RMSE have the values of 2293583, 1116.276, 11.57325, and 1514.458 respectively for the SMI stock market index. The aforementioned evolution indicators result shows the ARIMA model having the lowest indicator values within the classical forecasting models. That makes ARIMA as best fitted forecasting method in classical forecasting models for SMI stock market index under 10minute data frequency.

For the stock market FTSE index, both trained and test series have been shown in figure B-3 in appendix for our first conventional classical forecasting model ARIMA. The Results show the sudden drop in forecasted windows due to failed the ARIMA forecasting method to predict in such case as the arrival of Covid-19 in this time frame. Forecasting Indicators MSE, MAE, MAPE, and MAPE have the values of 708319, 584.502, 9.71500, and 841.617 respectively presented in Table 4.12. ARFIMA estimated accuracy indicators MSE, MAE, MAPE, and RMSE presented in table 4.1.6. Computed forecasting accuracy indicators have the values of, MSE (722039), MAE (675.374), MAPE (11.08673), and RMSE (849.729). Figure B-4 in appendix exhibit the ARFIMA forecasted and actual series for the FTSE stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in Figure B-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 1518021, 877.899, 14.82025, and 1232.080. Computed values of Accuracy indicator of ARIMA model declared the winner within the classical forecasting techniques for 10-minute data frequency.

Forecasting Indicators of classical Models for DJI30 stock market index has been presented in Table 4.12. Estimated MSE, MAE, MAPE, and RMSE has the statistical values of 6593601, 2011.048, 8.14986, and 2567.801 respectively for the DJI30 stock market index. Similarly, a Graphical representation has been given in figure B-3 in appendix. Which shows a sudden fall decreases the ARIMA model prediction power in the testing phase. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 9594271, 2384.127, 9.88539, and 3097.462 respectively for the DJI30 stock market index given in Table 4.12. Figure B-4 in appendix depicts the visual outlook of the actual and forecasted series. Our 3rd classical model GARCH, computed forecasting indicators has been presented in table 4.14 The model choose 100 observations to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure B-5 in appendix. Computed forecasted indicators MSE, MAE, MAPE, and RMSE have the values of 20906897, 3153.342, 13.32388, and 4572.406 respectively for the DJI30 stock market index. The aforementioned forecasting accuracy indicators result shows the ARIMA model having the lowest indicator values within the classical forecasting models. That makes ARIMA as best fitted forecasting method in classical forecasting models for DJI30 stock market index under 10minute data frequency.

For the stock market index SNP500, both trained and test series have been shown in figure B-3 in appendix for the predefined ARIMA model. Similarly, predefined lower and upper bound also result show the sudden drop in forecasted windows due to failed the ARIMA forecasting method to predict in such case as the arrival of Covid-19 in this time frame. Forecasting accuracy indicators, MSE, MAE, MAPE, and RMSE have the values as 57742, 191.116, 6.48323, and 240.296 respectively for the SNP500 stock market index. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 67331, 206.670, 7.13741, and 259.482 respectively for the SNP500 stock market index given in Table 4.12. Figure B-5 in appendix depicts the visual outlook of the actual and forecasted series.

For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure B-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 115130, 261.080, 9.20419, and 339.308. Based on computed values of the Accuracy indicator of all models ARIMA model has the lowest values of these evaluation criteria. Which declares the ARIMA as the winner in the classical forecasting techniques for the SNP 500 stock market index.

The study applies three dynamic classical models (ARIMA, ARFIMA, and GARCH) to forecast our selected stock market indices. After the estimation of these models, all of the sample market indices are forecasted one by one. Results of the analysis show that the classical ARIMA model performs better in forecasting the 10 minutes data for the stock market indices.

4.2.4 Machine and Deep Learning Methods

This analysis section deals with the machine learning and deep learning analysis. Section has been further divided into two subsessions which deals with the model Eetimation and model Forecasting to achieve the objective of the study.

4.2.4.1 Models Estimation

For modal estimation under 10-minutes data frequency, predefined functions in section I, machine and deep learning methods estimation has been used for both SVM and LSTM. For hybrid methods again pre-decided functions and libraries has been used.

4.2.4.2 Models Forecasting

Table 4.13 enlightens the computed model's selection criteria's indicators values for each stock market for both machine learning and deep learning methods. For the ASX stock market, SVM model forecasting accuracy indicators have the values as 194124, 352.503, 5.778, and 440.596 respectively for MSE, MAE, MAPE, and RMSE. For The hybrid ARIMA-SVM accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 193294, 350.788, 5.750, and 439.652 respectively. Which improves the results and indicate that ARIMA-SVM performs better in machine learning regression than the simple SVM model. Figure 4.33 exhibit the actual and forecasted SVM and ARIMA-SVM models for ASX.

The Deep learning methods LSTM and The hybrid LSTM included ARIMA and GARCH results presented in table 4.13 for selected stock markets. For the ASX stock market index, LSTM estimated model accuracy indicators have the MSE, MAE, MAPE, and RMSE values 22671, 53.452, 0.932, and 150.567 respectively, which outperform the machine learning models. For the hybrid, the ARIMA-LSTM method computed forecasting accuracy indicators have the MSE, MAE, MAPE, and RMSE values of 2088, 32.846, 0.542, and 45.700. Similarly, for the hybrid GARCH-LSTM method have the values of 2174, 34.300, 0.548, and 46.629 for MSE, MAE, MAPE, and RMSE respectively. Figure 4.34 exhibit the train, test, real and forecasted series of ASX. Where green and red-colored series represent the trained and actual test time series, Purple, orange, and magenta colore represent the LSTM, ARIMA-LSTM, and GARCH-LSTM results respectively. Results shows that ARIMA-LSTM has lowest forecasting accuracy indicators and is close to actual time series so it is the best fitted deep learning method.

TABLE 4.13: Machine and Deep learning Forecasting(10 Minutes data)

	ASX	A50	EUS	CAC	DAX	HSI	NIFTY	NIKKEI	AEX	WIG	SSGF	IBX	SMI	FTSE	DJI	SNP
								SVM								
MSE	194124	1129494	61592	156642	778201	4372388	744877	1807876	1169	56301	934.56	913678	384024	269138	3543356	48502
MAE	352.50	905.50	183.83	286.70	667.86	1651.11	660.25	1073.24	25.23	163.87	21.70	671.48	507.48	362.01	1465.84	174.96
MAPE	5.778	7.284	5.659	5.610	5.693	6.037	6.2580	5.039	4.697	8.376	6.359	7.895	5.346	5.403	5.916	6.25
RMSE	440.60	1062.78	248.18	395.78	882.16	2091.03	863.06	1344.57	34.19	237.28	30.57	955.87	619.70	518.78	1882.38	220.23
								LSTM								
MSE	22671	1657	7761	2434	27743	39615	34134	6370	10	12704	50	58736	22415	247268	11788	736
MAE	53.45	33.11	30.96	28.11	51.02	92.69	89.22	62.46	1.94	61.14	4.50	137.27	90.35	220.75	82.35	20.44
MAPE	0.932	0.242	1.104	0.564	0.513	0.387	0.979	0.289	0.372	3.953	1.530	1.997	0.853	3.950	0.319	0.652
RMSE	150.57	40.71	88.10	49.33	166.56	199.03	184.75	79.81	3.20	112.71	7.09	242.36	149.72	497.26	108.57	27.13
								ARIMA-SVM								
MSE	193294	1170780	63259	158780	792545	4205017	744874	1816964	1170	58850	935.19	928438	400665	210308	3142900	15349
MAE	350.79	903.85	178.08	283.03	656.94	1607.18	660.25	1070.88	25.20	155.30	21.60	661.02	500.07	220.13	1328.17	72.56
MAPE	5.750	7.379	5.561	5.586	5.661	6.048	6.2583	5.050	4.696	8.192	6.342	7.900	5.195	5.319	5.946	7.075
RMSE	439.65	1082.03	251.51	398.47	890.25	2050.61	863.06	1347.95	34.20	242.59	30.58	963.55	632.98	458.59	1772.82	123.89
								ARIMA-LSTM								
MSE	2088	1305	421	1814	4092	26829	37535	39982	6	6484	45	62945	39767	67807	21414	628
MAE	32.85	29.87	9.49	26.68	48.67	95.61	155.68	96.65	1.58	47.47	3.80	129.75	166.41	103.64	101.00	17.33
MAPE	0.542	0.222	0.325	0.526	0.433	0.397	1.545	0.499	0.301	3.018	1.303	1.912	1.606	1.867	0.376	0.556
RMSE	45.70	36.12	20.51	42.59	63.97	163.80	193.74	199.96	2.50	80.52	6.67	250.89	199.42	260.40	146.34	25.07
								GARCH-LSTM								
MSE	2174	1124	424	1999	6614	6639	14982	4826	25	28593	299	14891	5868	31190	11693	982
MAE	34.30	23.73	11.79	32.40	28.51	49.26	84.36	44.61	3.42	95.78	8.38	76.93	49.34	80.37	74.51	20.48
MAPE	0.548	0.176	0.375	0.587	0.275	0.198	0.868	0.219	0.624	6.147	2.935	1.069	0.476	1.432	0.288	0.647
RMSE	46.63	33.53	20.59	44.71	81.33	81.48	122.40	69.47	5.03	169.09	17.30	122.03	76.60	176.61	108.13	31.34

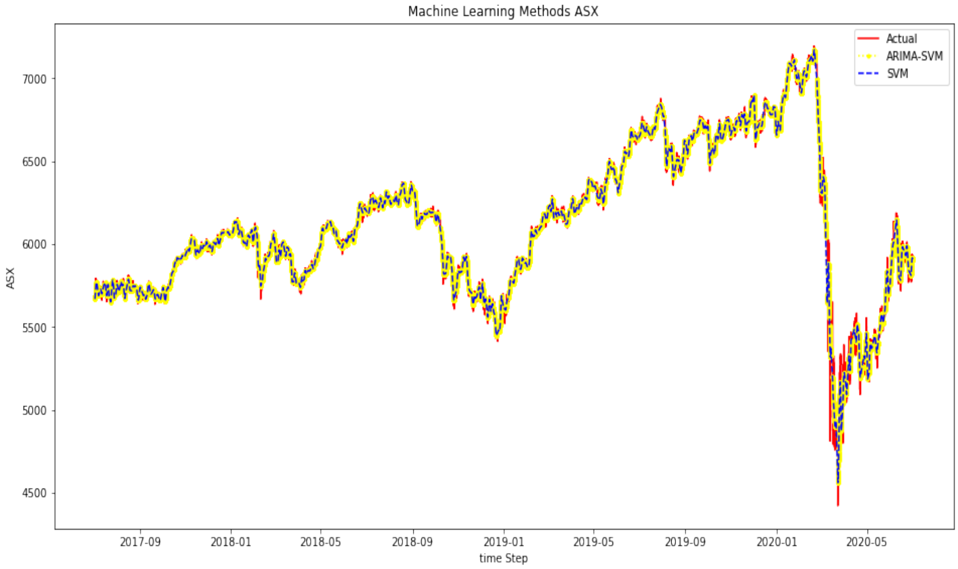


FIGURE 4.33: ASX Machine Learning Forecasting (10 Minutes Data)

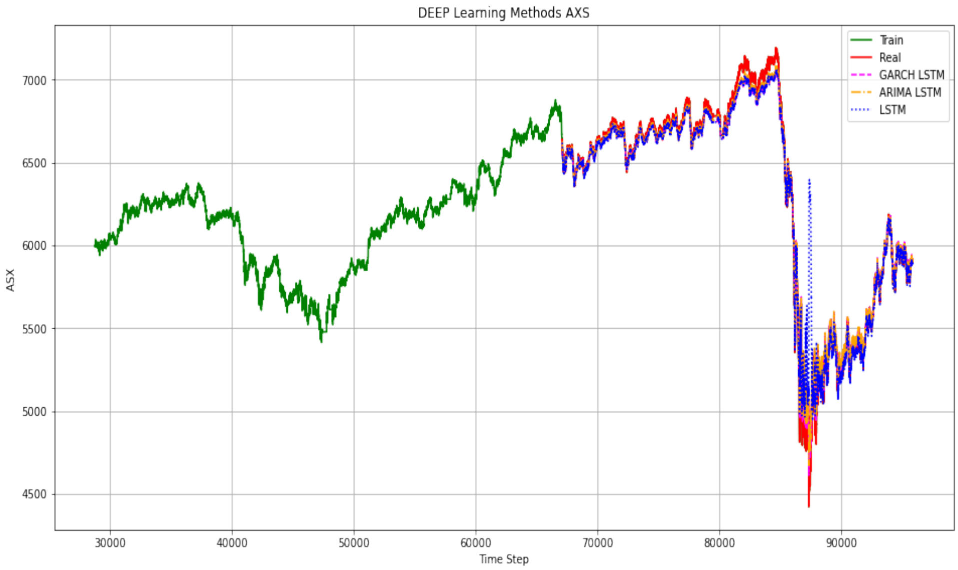


FIGURE 4.34: ASX Deep Learning Forecasting (10 Minutes Data)

For the A50 stock market Index, SVM model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have values 1129494, 905.500, 7.284, and 1062.776 respectively. For The hybrid ARIMA-SVM model, forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have the values 1170780, 903.854, 7.379, and 1082.026 respectively, which improve the results indicates simple ARIMA-SVM. MSE, RMSE, and MAE values show the ARIMA-SVM model is a better predictor but based on the percentage-based method, MAPE leads to better results for the Simple SVM method to be the better-predicting models for the A50 stock market.

Figure 4.35 exhibit the actual and forecasted SVM and ARIMA-SVM models for A50.

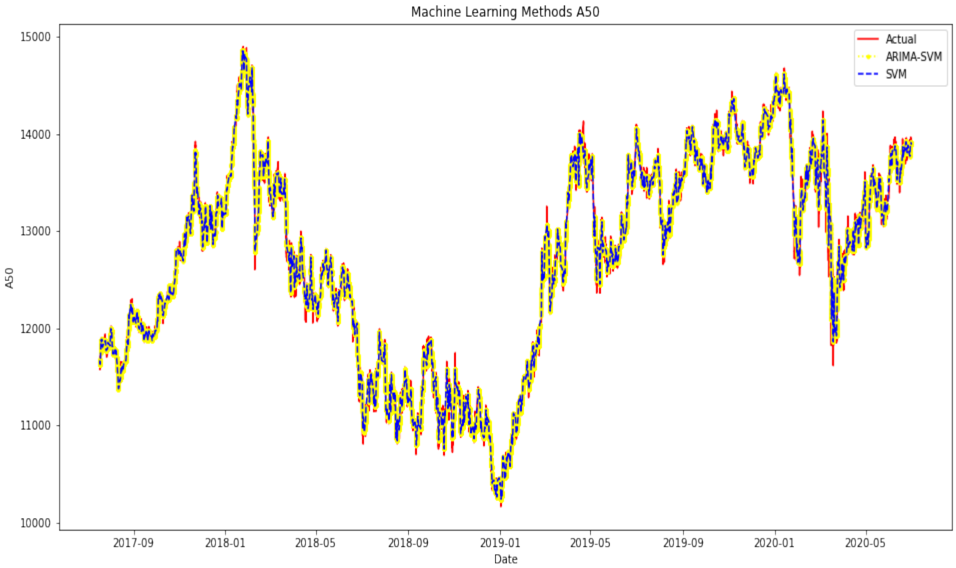


FIGURE 4.35: A50 Machine Learning Forecasting (10 Minutes Data)

FIGURE 4.36: A50 Deep Learning Forecasting (10 Minutes Data)

For Deep learning methods, LSTM, hybrid ARIMA-LSTM ARIMA-LSTM, and GARCH-LSTM based results for the A50 stock market index have been reported in table 4.13. For the A50 stock market index, LSTM estimated model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 1657, 33.112, 0.242, and 40.712 respectively outperform the machine learning models. The hybrid ARIMA-LSTM computed model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 1305, 29.873, 0.222, and 36.120, respectively. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 1124, 23.734, 0.176, and 33.532. Figure 4.36 exhibit the train, test, and forecasted series for A50. Test results shows proposed GARCH-LSTM with the lowest forecasting indicators values and closest to the actual time series is the best model for the A50 stock market index.

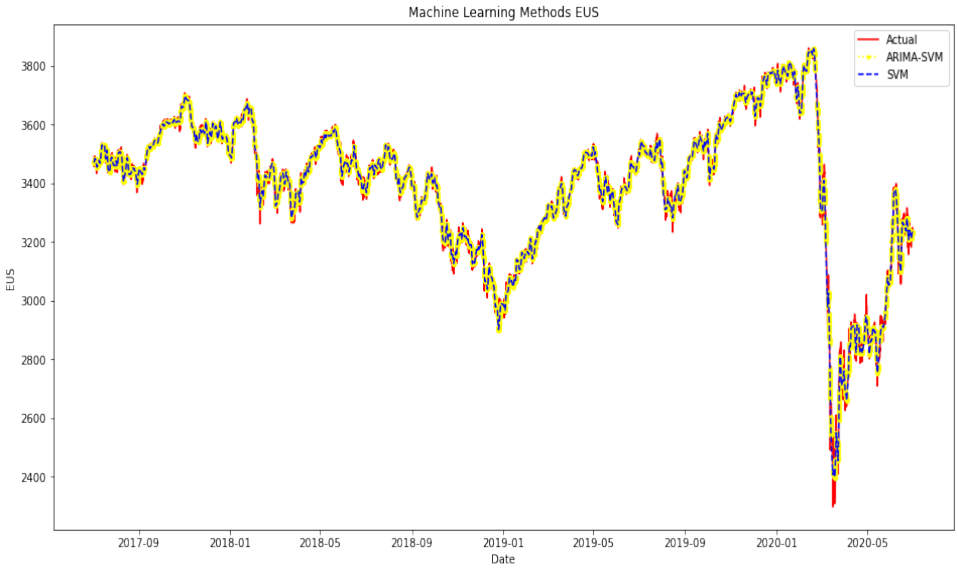


FIGURE 4.37: EUS Machine Learning Forecasting (10 Minutes Data)

For EUS stock market Index SVM, model forecasting accuracy indicators (MSE, MAE, MAPE, and RMSE) have the values 61592, 183.826, 5.659, and 248.178 respectively. For the hybrid ARIMA-SVM forecasting model, MSE, MAE, MAPE, and RMSE have the values as 63259, 178.078, 5.561, and 251.513 respectively. Based on Willmott and Matsuura (2005) suggestions, improved MAE and MAPE forecasting indicators values, we can infer that ARIMA-SVM performs well than the SVM in machine learning regression. Figure 4.37 exhibit the actual and forecasted SVM and ARIMA-SVM models for the EUS Stock market index.

For Deep learning methods, LSTM, hybrid ARIMA-LSTM ARIMA-LSTM, and GARCH-LSTM computed results for the EUS stock market index have given in Table 4.13. For EUS stock market index LSTM model reports MSE, MAE, MAPE,

FIGURE 4.38: EUS Deep Learning Forecasting (10 Minutes Data)

and RMSE values as 7761, 30.960, 1.104, and 88.098 respectively, which outperform the conventional machine learning models and the hybrid ARIMA-SVM models. For the hybrid, ARIMA-LSTM model forecasting MSE, MAE, MAPE, and RMSE have the values as 2255.00, 30.88, 1.04, and 47.49 respectively. Similarly, for the hybrid GARCH-LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values 705.86, 15.49, 0.50, and 26.57 respectively. Figure 4.38 exhibit the train, test, and predicted series of EUS. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta color represents LSTM, ARIMA-LSTM, and GARCH-LSTM respectively. Results shows that Proposed GARCH-LSTM being closest to actual the time series is the best fitted deep learning method.

For the CAC stock market, SVM model forecasting accuracy indicators have the values as 156642, 286.701, 5.610, and 395.780 respectively for MSE, MAE, MAPE, and RMSE. For The hybrid ARIMA-SVM accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 158780, 283.025, 5.586, and 398.473 respectively, which improves the results of machine learning methods with the hybrid technique. Results indicate that ARIMA-SVM performs better in machine learning regression than the simple SVM model. Figure 4.39 exhibit the actual and forecasted SVM and ARIMA-SVM models for CAC.

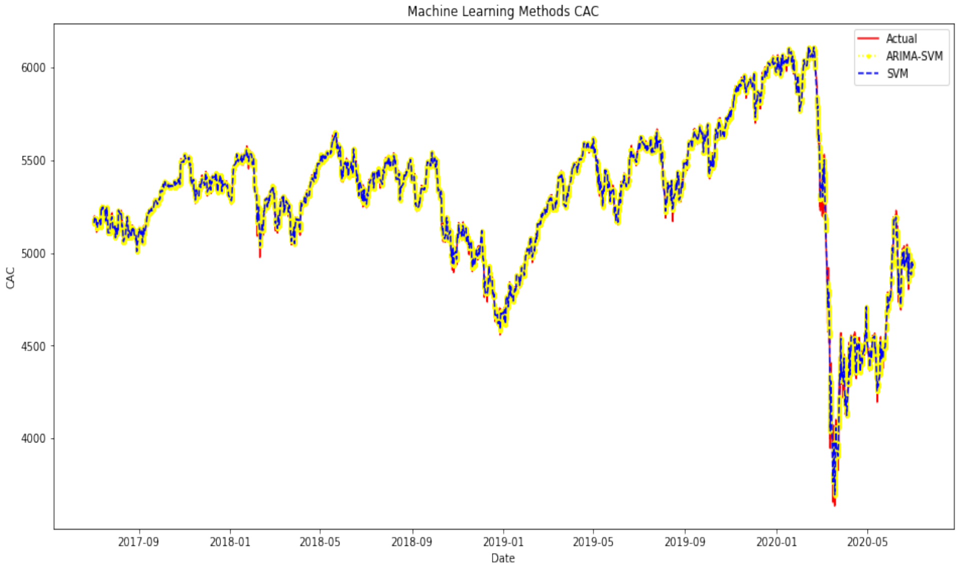


FIGURE 4.39: CAC Machine Learning Forecasting (10 Minutes Data)

The Deep learning methods LSTM, and the hybrid LSTM along with ARIMA and GARCH results are presented in table 4.13 for all selected stock markets. For the CAC stock market index, LSTM estimated model accuracy indicators have the MSE, MAE, MAPE, and RMSE values 2434, 28.107, 0.564, and 49.331 respectively which outperform the machine learning models. For the hybrid, ARIMA-LSTM MSE, MAE, MAPE, and RMSE are 1814, 26.682, 0.526, and 42.588. Similarly, for the hybrid GARCH-LSTM method have the values of 1999, 32.395, 0.587, and 44.709 for MSE, MAE, MAPE, and RMSE respectively. Figure 4.40 exhibit the train, test, real and forecasted series of CAC. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta colore represent the LSTM, ARIMA-LSTM, and GARCH-LSTM results respectively. Results show that ARIMA-LSTM has lowest forecasting accuracy indicators and is close to actual time series so it is the best fitted deep learning method.

For DAX stock market Index SVM, model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have values 778201, 667.860, 5.693, and 882.157 respectively. The hybrid ARIMA-SVM model, forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have the values 792545, 656.940, 5.661, and 890.250 respectively, which improve MAE, and MAPE values. Based on the results ARIMA-SVM model is a better predictor than the Simple SVM method for the DAX stock

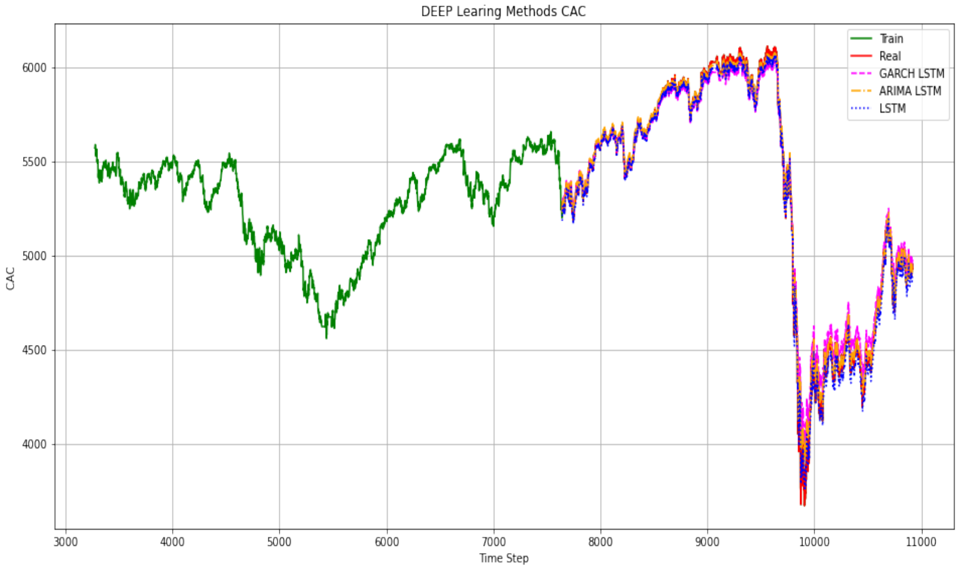


FIGURE 4.40: CAC Deep Learning Forecasting (10 Minutes Data)

market index. Figure 4.41 exhibit the actual and forecasted SVM and ARIMA-SVM models for DAX.

For Deep learning methods, LSTM hybrid ARIMA-LSTM, and GARCH-LSTM computed results for DAX stock market index are reported in table 4.13. For the DAX stock market index, LSTM estimated model have MSE, MAE, MAPE, and RMSE values as 27743, 51.021, 0.513, and 166.564 respectively outperform the machine learning models. The hybrid ARIMA-LSTM model MSE, MAE, MAPE, and RMSE have the values as 4092, 48.666, 0.433, and 63.970 respectively. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 6614, 28.506, 0.275, and 81.328. Figure 4.42 exhibit the train, test, actual and forecasted series for DAX. Based on literature improved MAE and MAPE values proposed GARCH-LSTM with the lowest forecasting indicators values and closest to the actual time series is the best model for the DAX stock market index.

For HSI stock market Index, SVM forecasting model accuracy indicators have the values as 4372388, 1651.108, 6.037, and 2091.025 respectively for MSE, MAE, MAPE, RMSE. The hybrid ARIMA-SVM forecasting model, MSE, MAE, MAPE, and RMSE have the values as 4205017, 1607.180, 6.048, and 2050.614 respectively. Based on MSE, MAE, and RMSE forecasting indicators values, we can infer that

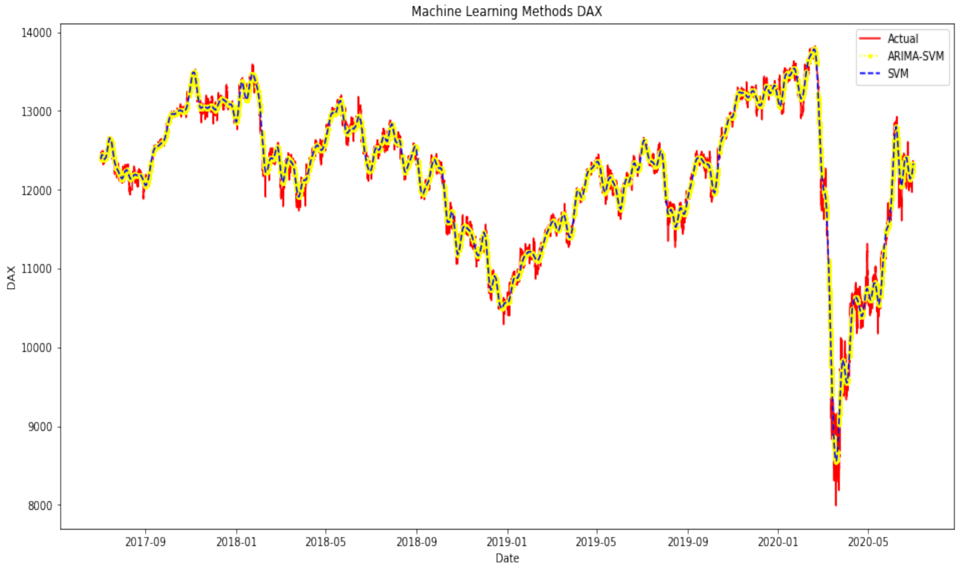


FIGURE 4.41: DAX Machine Learning Forecasting (10 Minutes Data)

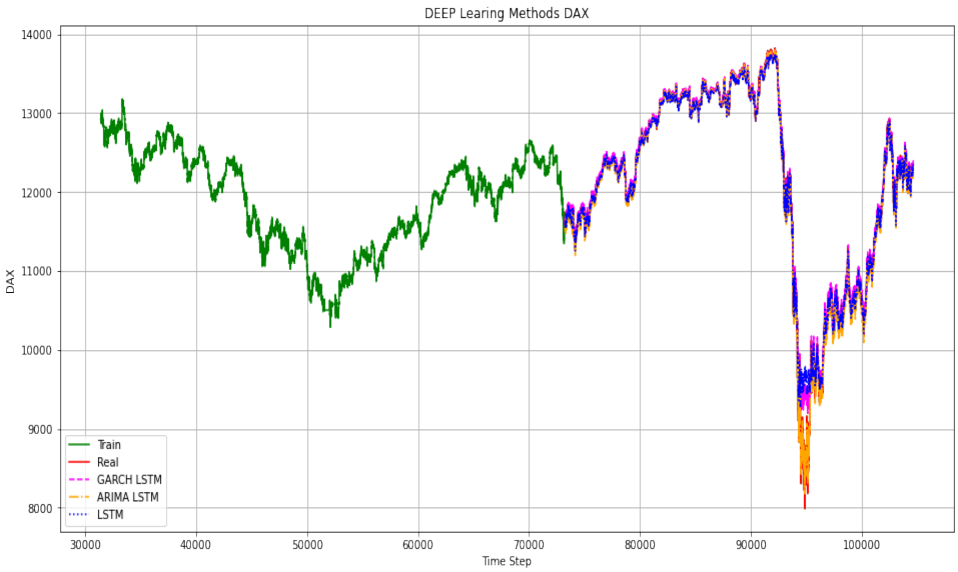


FIGURE 4.42: DAX Deep Learning Forecasting (10 Minutes Data)

ARIMA-SVM performs better than the SVM in machine learning regression. Figure 4.43 exhibit the actual and forecasted SVM and ARIMA-SVM models for the HSI Stock market index.

For LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM deep learning methods are used for forecasting the HSI stock market index results are presented Table 4.13. For HSI stock market index, LSTM model have forecasting accuracy inidcatiors values as 39615, 92.695, 0.387, and 199.034 respectively for MSE,MAE,MAPE and RMSE. Which outperform the conventional machine learning models and the

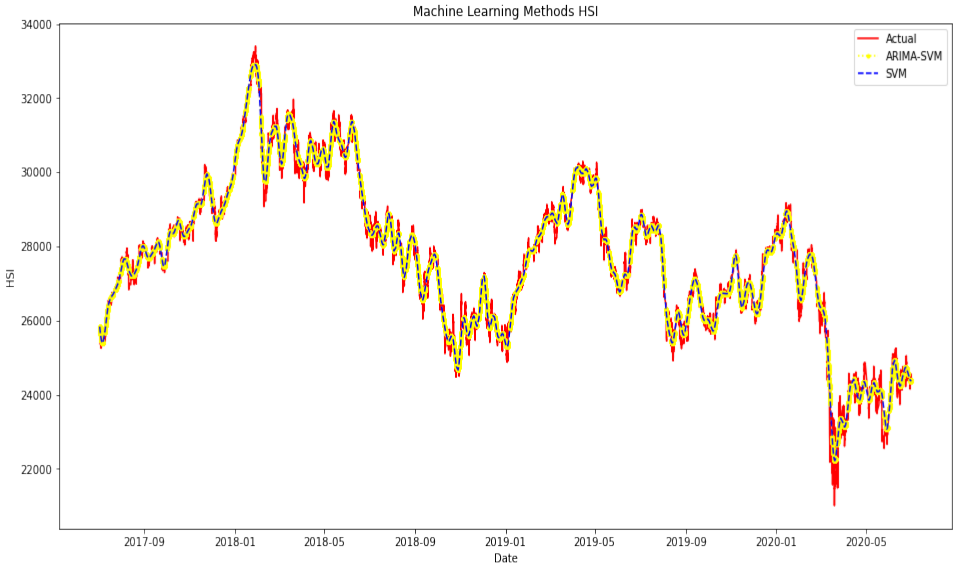


FIGURE 4.43: HSI Machine Learning Forecasting (10 Minutes Data)

hybrid ARIMA-SVM models. For the hybrid, ARIMA-LSTM model forecasting MSE, MAE, MAPE, and RMSE have the values as 26829, 95.612, 0.397, and 163.797 respectively. Similarly, for the hybrid GARCH-LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values 6639, 49.264, 0.198, and 81.481 respectively. Figure 4.44 exhibit the train, test, and predicted series of HSI. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta color represents LSTM, ARIMA-LSTM, and GARCH-LSTM respectively. Results shows that Proposed GARCH-LSTM being closest to actual the time series is the best fitted deep learning method.

Table 4.13 enlightens the computed model's forecasting accuracy indicators values for the NIFTY stock market index. SVM model forecasting accuracy indicators have the values as 744877, 660.247, 6.2580, and 863.062 respectively for MSE, MAE, MAPE, and RMSE. For The hybrid ARIMA-SVM, accuracy indicators have the values as 744874, 660.251, 6.2583, and 863.061 respectively for MSE, MAE,MAPE and RMSE. Which worsen the results of forecasting accuracy indicators for the hybrid ARIMA-SVM. SVM performs better in machine learning regression than the hybrid ARIMA- SVM model. Figure 4.45exhibit the actual and forecasted SVM and ARIMA-SVM models for NIFTY.

For Deep learning methods, LSTM and the hybrid LSTM, including ARIMA and

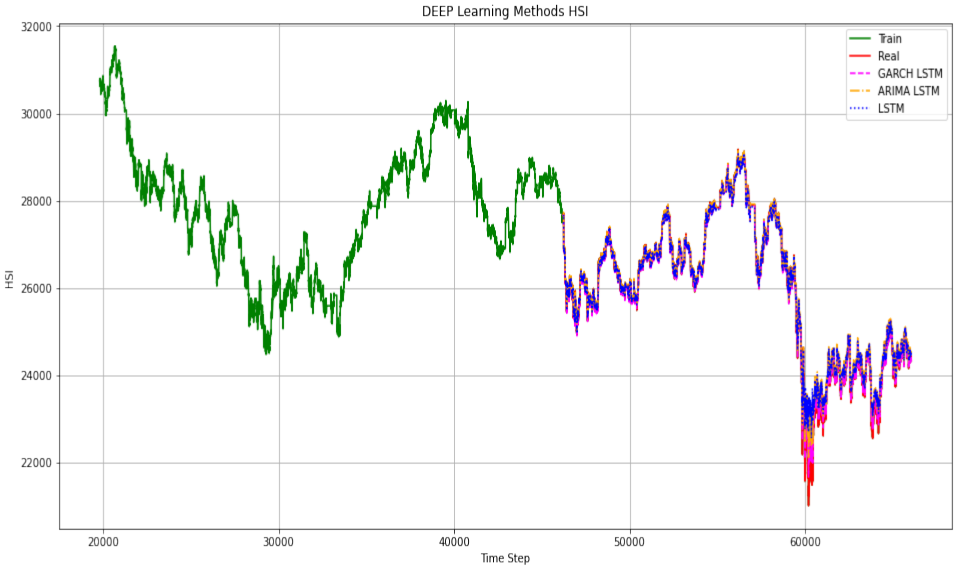


FIGURE 4.44: HSI Deep Learning Forecasting (10 Minutes Data)

GARCH results are presented in table 15 for the NIFTY Stock Market index. For the NIFTY stock market index, LSTM model have the MSE, MAE, MAPE, and RMSE values as 34134, 89.221, 0.979, and 184.754 respectively, which outperform the machine learning models. For the hybrid, ARIMA-LSTM MSE, MAE, MAPE, and RMSE are 37535, 155.683, 1.545, and 193.740. Similarly, for the hybrid GARCH-LSTM method have the values of 14982, 84.356, 0.868, and 122.400 for MSE, MAE, MAPE, and RMSE respectively. Figure 4.46 exhibit the train, test, real and forecasted series of NIFTY. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta colore represent the results of LSTM, ARIMA-LSTM, and GARCH-LSTM model respectively. Results show that proposed GARCH-LSTM with the lowest forecasting accuracy indicators is the best fitted deep learning method.

For NIKKEI stock market Index, SVM model have MSE, MAE, MAPE and RMSE values are 1807876, 1073.239, 5.039, and 1344.573 respectively. For The hybrid ARIMA-SVM model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have the values 1816964, 1070.876, 5.050, and 1347.948 respectively, which worsen the results of forecasting accuracy indicators for the hybrid ARIMA-SVM. SVM performs better in machine learning regression than the hybrid ARIMA-

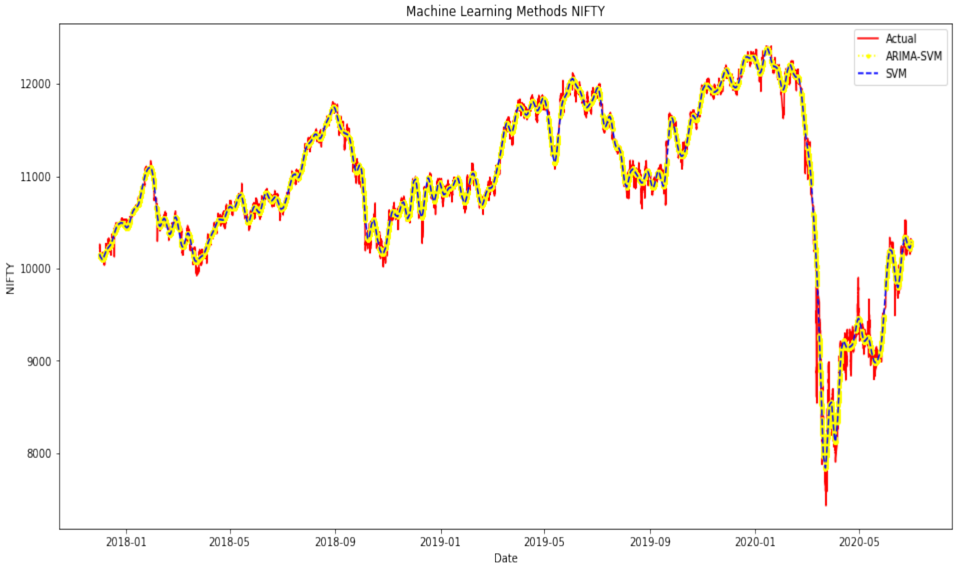


FIGURE 4.45: NIFTY Machine Learning Forecasting (10 Minutes Data)

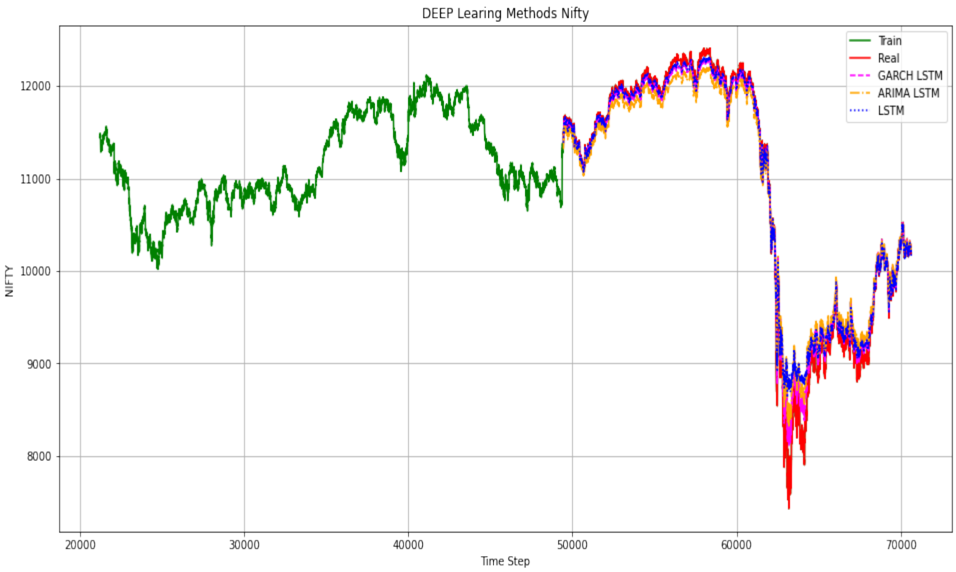


FIGURE 4.46: NIFTY Deep Learning Forecasting (10 Minutes Data)

SVM model. Figure 4.47 exhibit the actual and forecasted SVM and ARIMA-SVM models for NIKKEI.

For Deep learning methods, results of LSTM, hybrid ARIMA-LSTM ARIMA-LSTM, and GARCH-LSTM method for the NIKKEI stock market index have been reported in Table 4.6. 4.13. For the NIKKEI stock market index, LSTM estimated model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 6370, 62.455, 0.289, and 79.812, respectively outperform the machine learning models. The hybrid ARIMA-LSTM model MSE, MAE, MAPE, and RMSE have the values

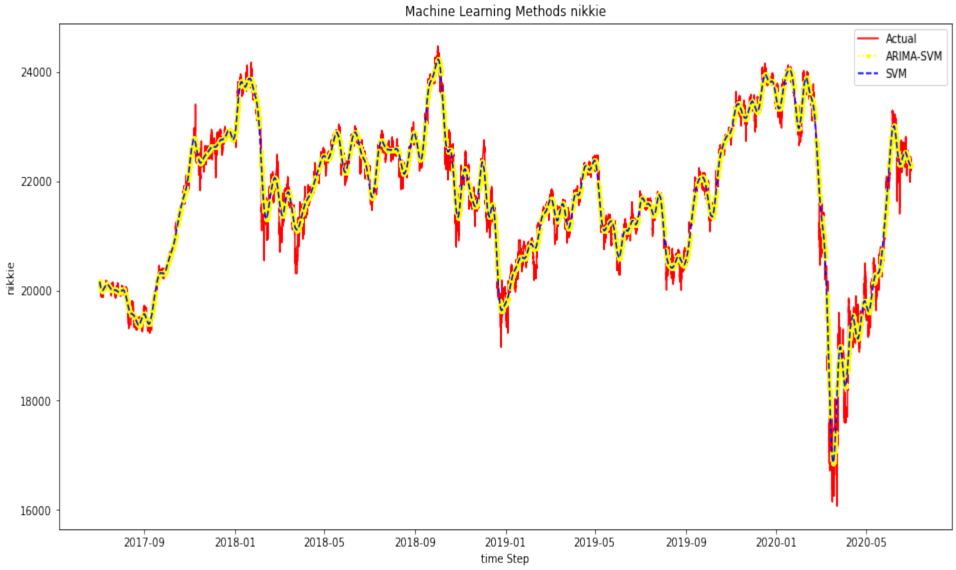


FIGURE 4.47: NIKKEI Machine Learning Forecasting (10 Minutes Data)

as 39982, 96.654, 0.499, and 199.956, respectively. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 4826, 4.608, 0.219, and 69.466. Figure 4.48 exhibit the train, test, and entire series NIKKEI. Test results shows proposed GARCH-LSTM with the lowest forecasting indicators values and closest to the actual time series is the best model for the NIKKEI stock market index.

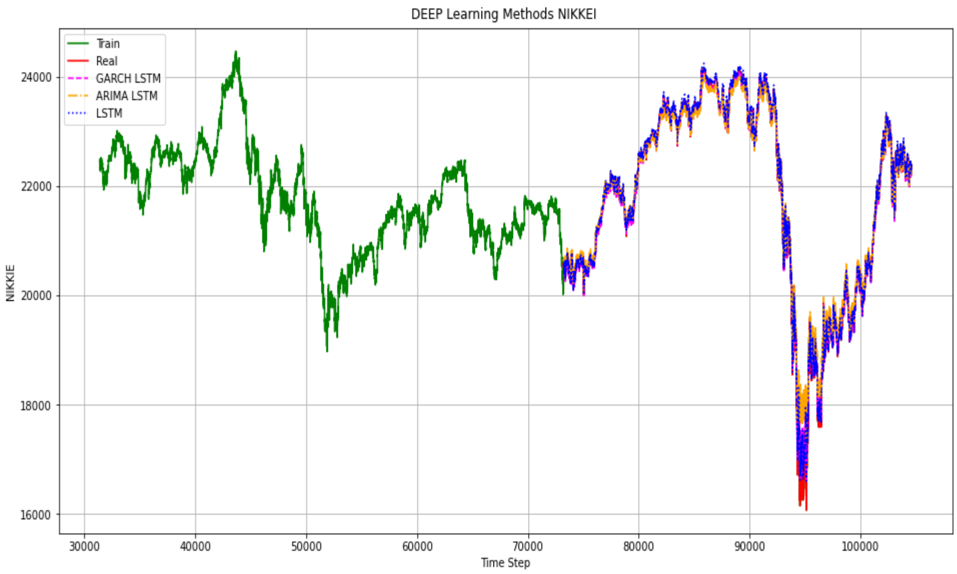


FIGURE 4.48: NIKKEI Deep Learning Forecasting (10 Minutes Data)

For the AEX stock market Index, SVM forecasting model accuracy indicators have the values as 1169, 25.233, 4.697, and 34.192 respectively for MSE, MAE, MAPE, RMSE. For the hybrid ARIMA-SVM forecasting model, MSE, MAE, MAPE, and RMSE have the values as 1170, 25.202, 4.696, and 34.200 respectively. Based on literature improved MAE and MAPE forecasting indicators values, we can infer that ARIMA-SVM performs better than the SVM in machine learning regression. Figure 4.49 exhibit the actual and forecasted SVM and ARIMA-SVM models for the AEX Stock market index.

For Deep learning methods, LSTM, hybrid ARIMA-LSTM ARIMA-LSTM, and GARCH-LSTM computed results for the AEX stock market index have given in Table 4.13. For the AEX stock market index LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 10, 1.936, 0.372, and 3.202 respectively, which outperform the conventional machine learning models and the hybrid ARIMA-SVM models. For the hybrid, ARIMA-LSTM model forecasting MSE, MAE, MAPE, and RMSE have the values as 6, 1.581, 0.301, and 2.499 respectively.

Similarly, for the hybrid GARCH-LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values 25, 3.422, 0.624, and 5.032 respectively. Figure 4.50 exhibit the train, test, and predicted series of AEX. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta color represents LSTM, ARIMA-LSTM, and GARCH-LSTM respectively. Results shows that ARIMA-LSTM being closest to actual the time series is the best fitted deep learning method.

For the WIG stock market, SVM model forecasting accuracy indicators i.e., MSE, MAE, MAPE, and RMSE have the values 56301, 163.868, 8.376, and 237.278 respectively. For The hybrid ARIMA-SVM accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 58850, 155.295, 8.192, and 242.591 respectively, which improves the results of machine learning methods with the hybrid technique. Results indicate that ARIMA-SVM performs better in machine learning regression than the simple SVM model. Figure 4.51 exhibit the actual and forecasted values of WIG stock market index, where SVM and ARIMA-SVM models are used.

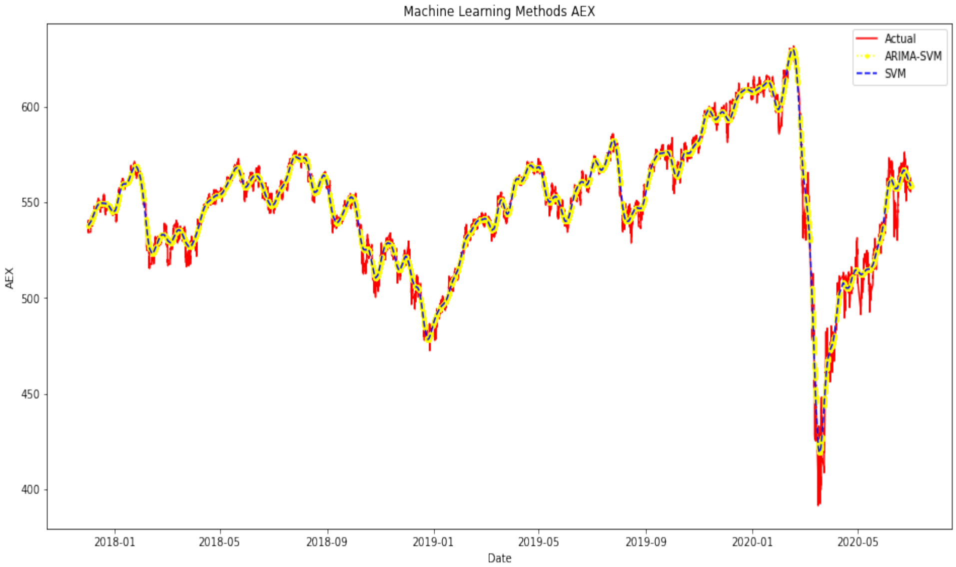


FIGURE 4.49: AEX Machine Learning Forecasting (10 Minutes Data)

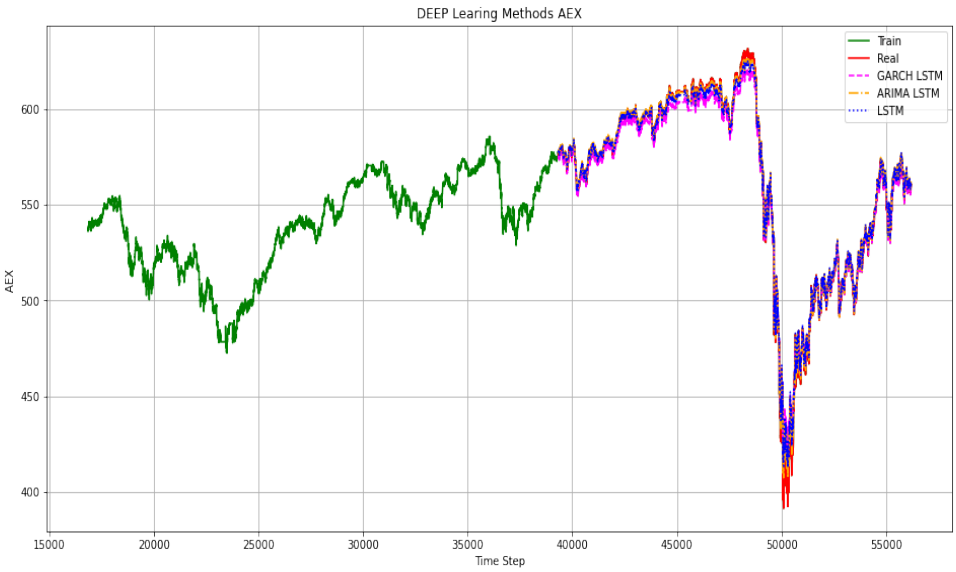


FIGURE 4.50: AEX Deep Learning Forecasting (10 Minutes Data)

The results of Deep learning methods i.e., LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM are presented in Table 4.13 for all selected stock markets. For the WIG stock market index, LSTM estimated model accuracy indicators have the MSE, MAE, MAPE, and RMSE values 12704, 61.143, 3.953, and 112.710 respectively, which outperform the machine learning models. For the hybrid, ARIMA-LSTM MSE, MAE, MAPE, and RMSE are 6484, 47.468, 3.018, and 80.525. Similarly, for the hybrid GARCH-LSTM method have the values of 28593,

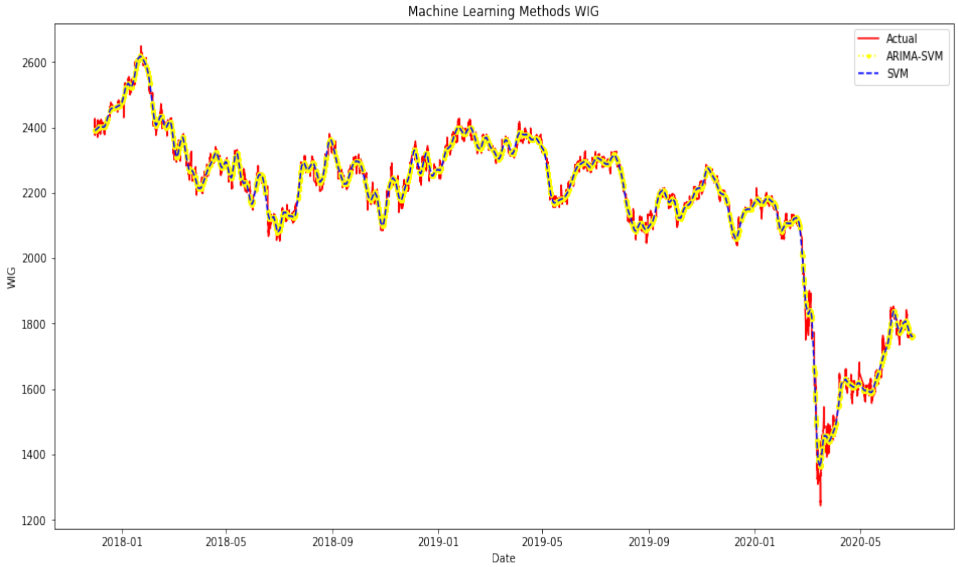


FIGURE 4.51: WIG Machine Learning Forecasting (10 Minutes Data)

95.783, 6.147, and 169.095 for MSE, MAE, MAPE, and RMSE respectively. Figure 4.52 exhibit the train, test, real and forecasted series of WIG. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta colore represent the results of LSTM, ARIMA-LSTM, and GARCH-LSTM model respectively. Results show that ARIMA-LSTM has lowest forecasting accuracy indicators and is close to actual time series so it is the best fitted deep learning method.

FIGURE 4.52: WIG Deep Learning Forecasting (10 Minutes Data)

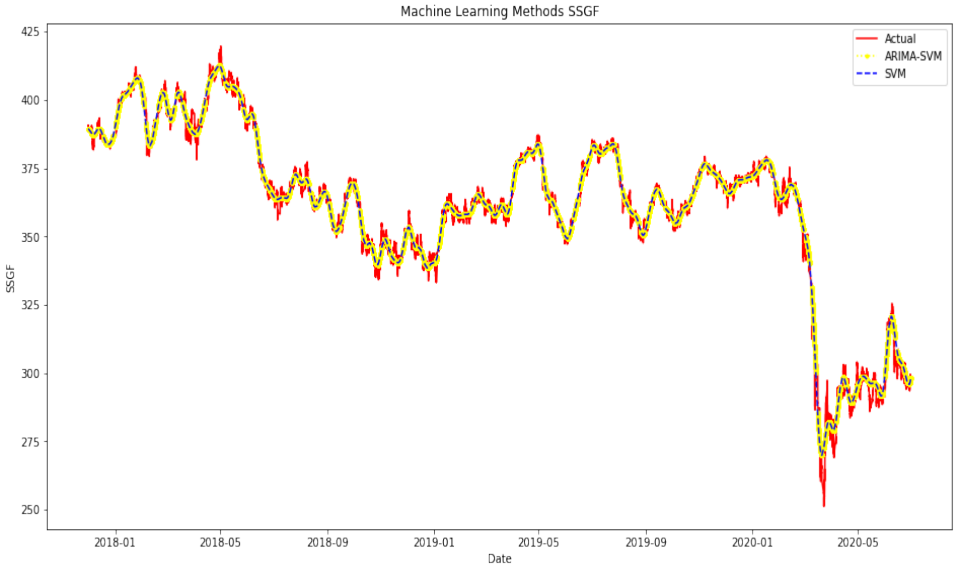


FIGURE 4.53: SSGF Machine Learning Forecasting (10 Minutes Data)

For the SSGF stock market Index, SVM model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have values 934.56, 21.704, 6.359, and 30.571 respectively. For The hybrid ARIMA-SVM model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have the values 935.19, 21.600, 6.342, and 30.581 respectively, which improve MAE, and MAPE values. Based on the results, and ARIMA-SVM model is a better predictor than the Simple SVM method for the SSGF stock market index. Figure 4.53 exhibit the actual and forecasted SVM and ARIMA-SVM models for SSGF.

For Deep learning methods, LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM estimated results for SSGF stock market index have been reported in Table 4.13. For the SSGF stock market index, LSTM estimated model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 50, 4.499, 1.530, and 7.091, respectively outperform the machine learning models. The hybrid ARIMA-LSTM computed model evaluation criteria have values as 45, 3.801, 1.303, and 6.672 respectively for MSE, MAE, MAPE, and RMSE. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 299, 8.377, 2.935, and 17.297. Figure 4.54 exhibit the train, test, actual and forecasted series for SSGF. Test results shows that ARIMA-LSTM with the lowest forecasting indicators values, is the best model for SSGF stock market index.

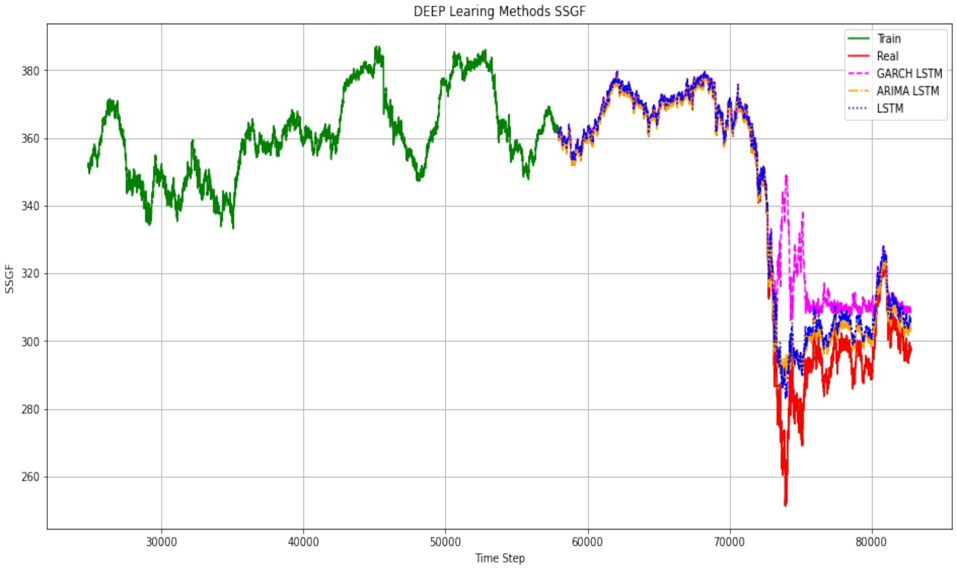


FIGURE 4.54: SSGF Deep Learning Forecasting (10 Minutes Data)

For the IBX stock market Index, SVM forecasting model accuracy indicators have the values as 913678, 671.482, 7.895 955.865 respectively for MSE, MAE, MAPE, RMSE. For the hybrid ARIMA-SVM forecasting model, MSE, MAE, MAPE, and RMSE have the values as 928438, 661.017, 7.900, and 963.555 respectively. Based on MSE, MAE, and RMSE forecasting indicators values, we can infer that SVM performs better than the hybrid ARIMA-SVM in machine learning regression. Figure 4.55 exhibit the actual and forecasted SVM and ARIMA-SVM models for the IBX Stock market index.

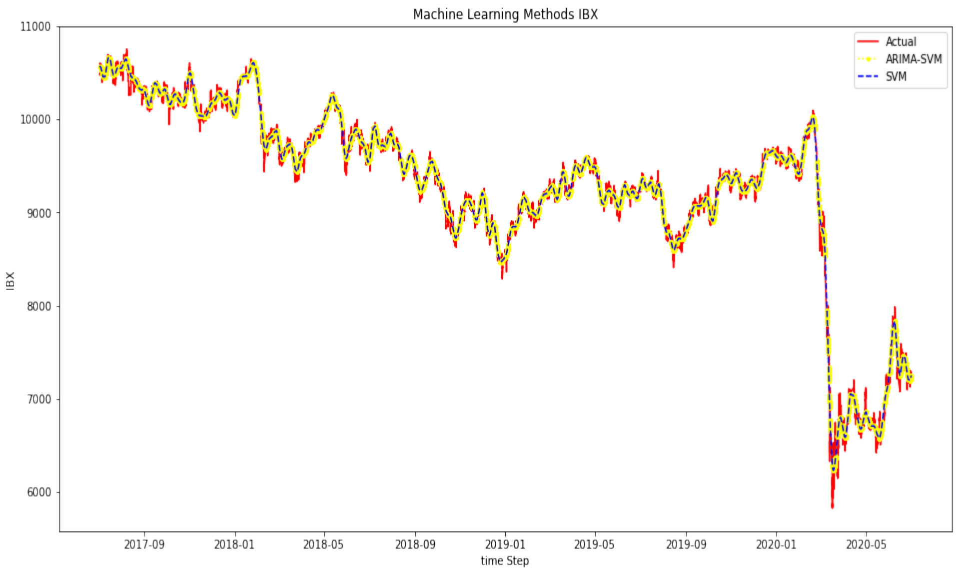


FIGURE 4.55: IBX Machine Learning Forecasting (10 Minutes Data)

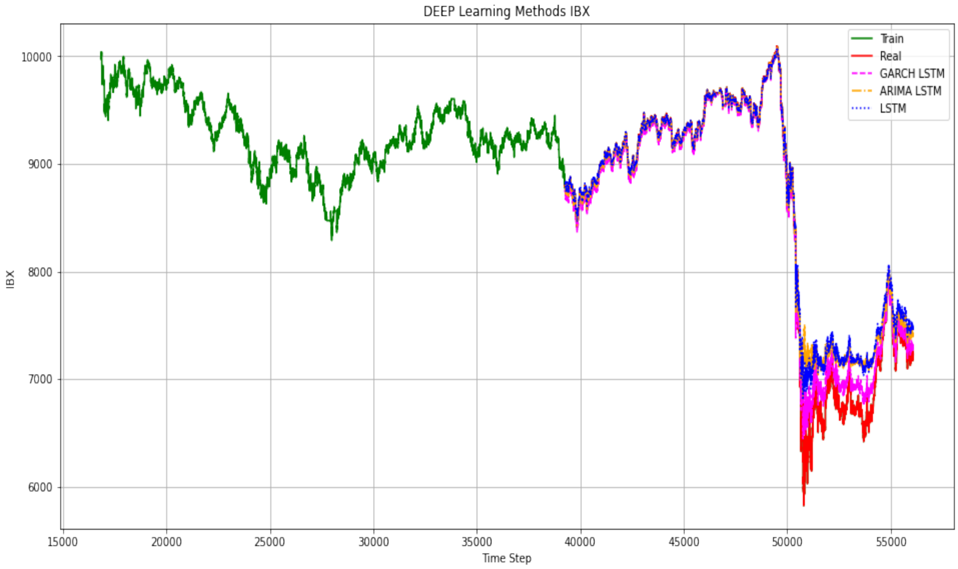


Figure 4.56: IBX Deep Learning Forecasting (10 Minutes Data)

For Deep learning methods, LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM are used and results for IBX stock market index are represented in Table 4.13. For IBX stock market index, LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 58736, 137.27, 1.99, and 242.36 respectively, which outperform the conventional machine learning models and the hybrid ARIMA-SVM models. For the hybrid, ARIMA-LSTM model forecasting MSE, MAE, MAPE, and RMSE have the values as 62945, 129.753, 1.912, and 250.888 respectively. Similarly, for the hybrid GARCH-LSTM model accuracy indicators have the values 14891, 76.935, 1.069, and 122.028 respectively for MSE, MAE, MAPE, and RMSE. Figure 4.56 exhibit the train, test, and forecasted series of IBX. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta color represents LSTM, ARIMA-LSTM, and GARCH-LSTM respectively. Results shows that Proposed GARCH-LSTM is closest to actual the time series so it is the best fitted deep learning method.

For SMI stock market Index SVM, model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have values 384024, 507.478, 5.346, and 619.697 respectively. For the hybrid ARIMA-SVM model forecasting accuracy indicators, MSE, MAE, MAPE, and RMSE have value as 400665, 500.068, 5.195, 632.981 respectively improve MAE, and MAPE values. Based on the results, and ARIMA-SVM model is a better predictor than the Simple SVM method for the SMI stock market index. Figure 4.57 exhibit the actual and forecasted value of SVM and ARIMA-SVM models for SMI.

For Deep learning methods, LSTM, hybrid ARIMA-LSTM ARIMA-LSTM, and GARCH-LSTM are used and results for the SMI stock market index are reported in Table 4.13. For the SMI stock market index, LSTM estimated model evaluation criteria (MSE, MAE, MAPE, and RMSE) have values 22415, 90.346, 0.853, and 149.717 respectively. which outperform the machine learning models. The hybrid ARIMA-LSTM model MSE, MAE, MAPE, and RMSE have the values as 39767, 166.412, 1.606, and 199.416 respectively. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 5868, 49.344, 0.476, and 76.603. Figure 4.58 exhibit the train, test, actual and forecasted series for SMI. Test results shows that the proposed GARCH-LSTM with the lowest forecasting indicators values, is the best model for the SMI stock market index.

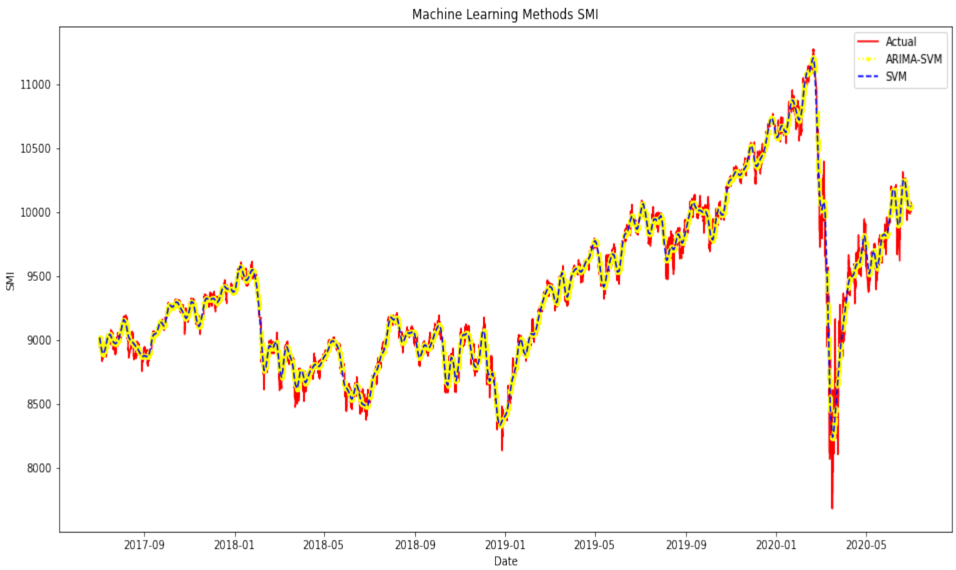


FIGURE 4.57: SMI Machine Learning Forecasting (10 Minutes Data)

For FTSE stock market Index, SVM model forecasting accuracy indicators (MSE, MAE, MAPE, and RMSE) have the values 269138, 362.012, 5.403, and 518.785 respectively. For the hybrid ARIMA-SVM forecasting model, MSE, MAE, MAPE, and RMSE have the values as 210308, 220.133, 5.319, and 458.593 respectively. Based on forecasting accuracy indicators values, we can infer that ARIMA-SVM

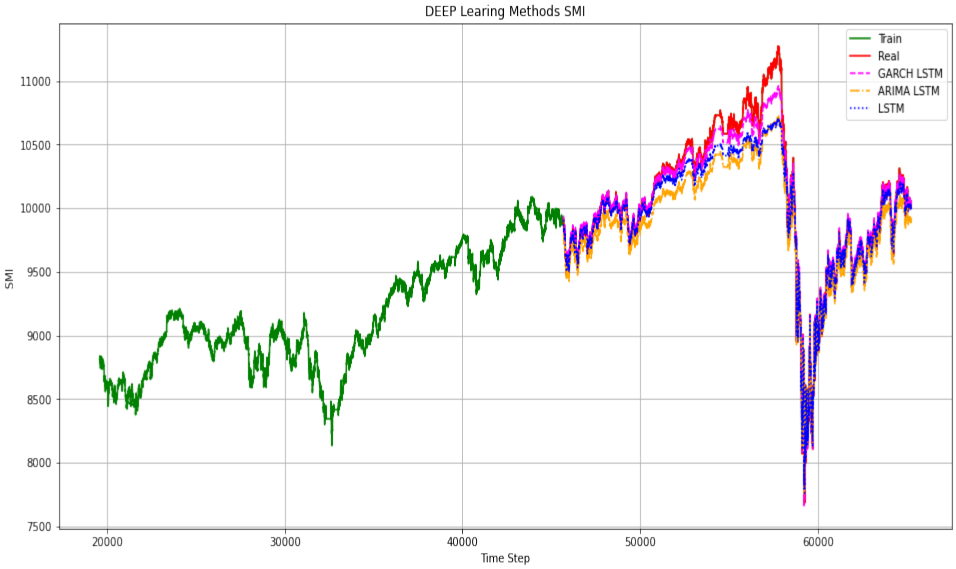


FIGURE 4.58: SMI Deep Learning Forecasting (10 Minutes Data)

performs better than the SVM in machine learning regression. Figure 4.59 exhibit the actual and forecasted values of SVM and ARIMA-SVM models for the FTSE Stock market index.

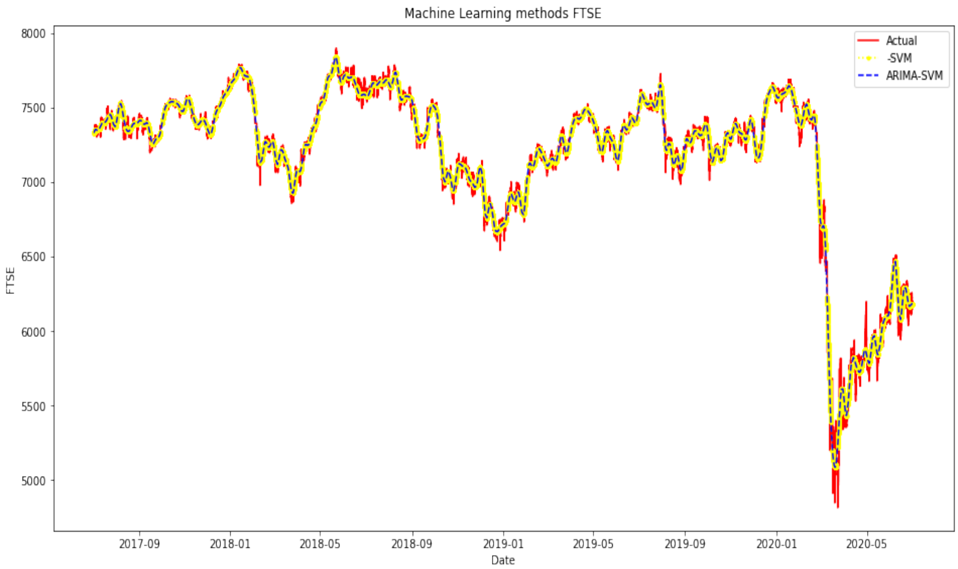


FIGURE 4.59: FTSE Machine Learning Forecasting (10 Minutes Data)

For Deep learning methods, LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM are used and computed results for the FTSE stock market index have given in Table 4.13. For FTSE stock market index LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 247268, 220.746, 3.950, and 497.261

FIGURE 4.60: FTSE Deep Learning Forecasting (10 Minutes Data)

respectively, which outperform the conventional machine learning models and the hybrid ARIMA-SVM models. For the hybrid, ARIMA-LSTM model forecasting MSE, MAE, MAPE, and RMSE have the values as 67807, 103.639, 1.867, and 260.398 respectively. Similarly, for the hybrid GARCH-LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values 31190, 80.370, 1.432, and 176.606 respectively. Figure 4.60 exhibit the train, test, and predicted series of FTSE. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta color represents LSTM, ARIMA-LSTM, and GARCH-LSTM respectively. Results shows that GARCH-LSTM model is closest to actual the time series so it is the best fitted deep learning method.

For DJI30 stock market Index SVM, model have forecasting accuracy indicators MSE, MAE, MAPE, and RMSE as 3543356, 1465.840, 5.916, and 1882.380 respectively. The hybrid ARIMA-SVM model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have the values 3142900, 1328.169, 5.946, and 1772.823 respectively, which improve the results indicates simple ARIMA-SVM. MSE, RMSE, and MAE values show that the ARIMA-SVM model is a better predictor than the SVM for the DJI30 stock market. Figure 4.61 exhibit the actual and forecasted SVM and ARIMA-SVM models for DJI30.

For Deep learning methods, LSTM, hybrid ARIMA-LSTM ARIMA-LSTM, and

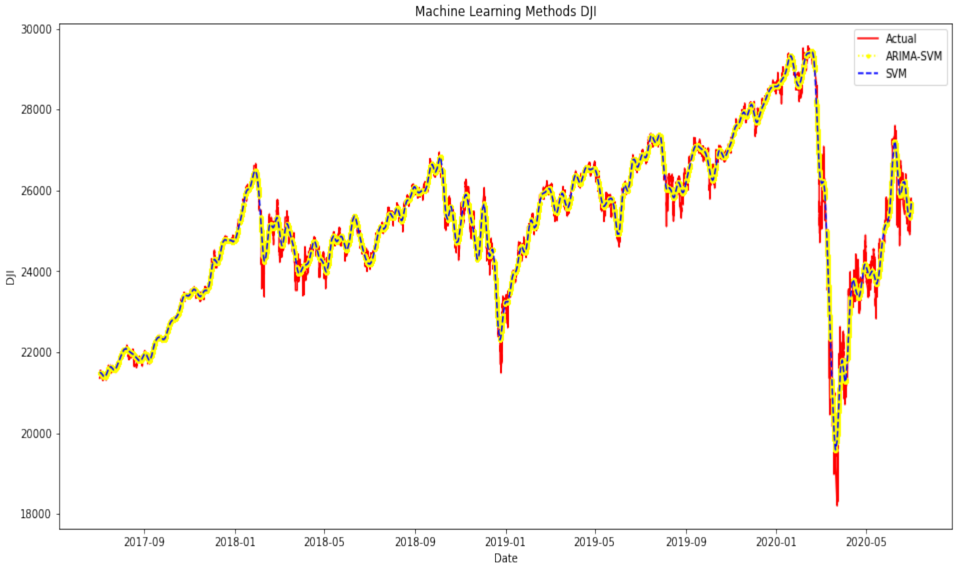


Figure 4.61: DJI Machine Learning Forecasting (10 Minutes Data)

GARCH-LSTM are used and results for DJI30 stock market index have been reported in Table 4.13. For the DJI30 stock market index, LSTM estimated model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 11788, 82.355, 0.319, and 108.570 respectively outperform the machine learning models. The hybrid ARIMA-LSTM computed model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 21414, 100.998, 0.376, and 146.336 respectively. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 11693, 74.515, 0.288, and 108.133. Figure 4.62 exhibit the train, test, and predicted series for DJI30. Results shows that proposed GARCH-LSTM has the lowest forecasting indicators values and closest to actual time series, so it is the best model for DJI30 stock market index.

For the SNP500 stock market, SVM model forecasting accuracy indicators have the values as 48502, 174.96, 6.25, and 220.23 respectively for MSE, MAE, MAPE, and RMSE. For The hybrid ARIMA-SVM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 15349, 72.555, 7.075, and 123.891 respectively, which improves the results of machine learning methods. Results indicate that ARIMA-SVM performs better in machine learning regression than the simple SVM model. Figure 4.63 exhibit the actual and forecasted SVM and ARIMA-SVM models for SNP500.

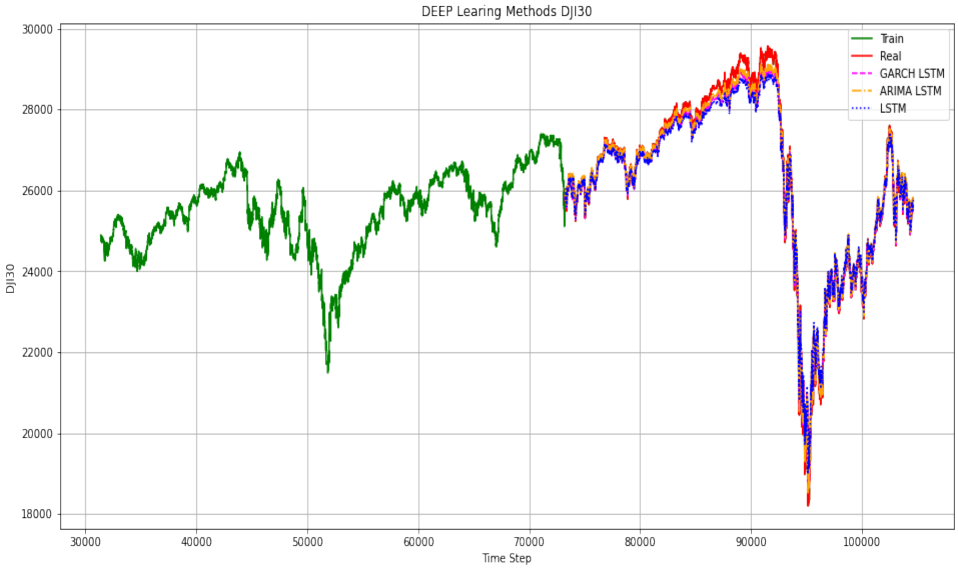


FIGURE 4.62: DJI Deep Learning Forecasting (10 Minutes Data)

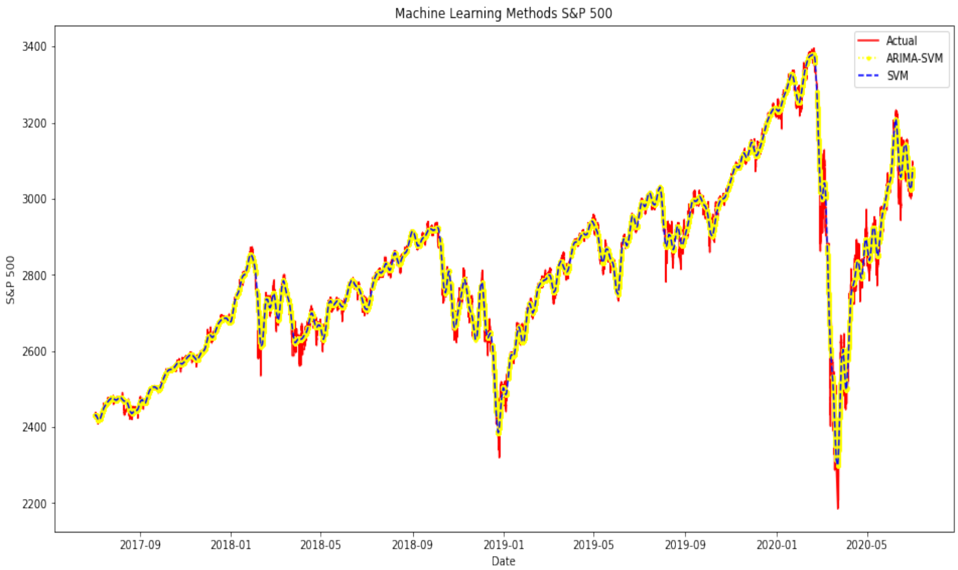


FIGURE 4.63: SNP Machine Learning Forecasting (10 Minutes Data)

The Deep learning methods LSTM, and the hybrid LSTM included ARIMA and GARCH results presented in table 4.13 for all selected stock markets. For the SNP500 stock market index, LSTM estimated model have accuracy indicators as MSE, MAE, MAPE, and RMSE with the values 736, 20.443, 0.652, and 27.127 respectively, which outperform the machine learning models. For the hybrid, the ARIMA-LSTM method computed forecasting accuracy indicators have the MSE, MAE, MAPE, and RMSE values of 628, 17.332, 0.556, and 25.066. Similarly, for the hybrid GARCH-LSTM method have the values of 982, 20.477, 0.647, and

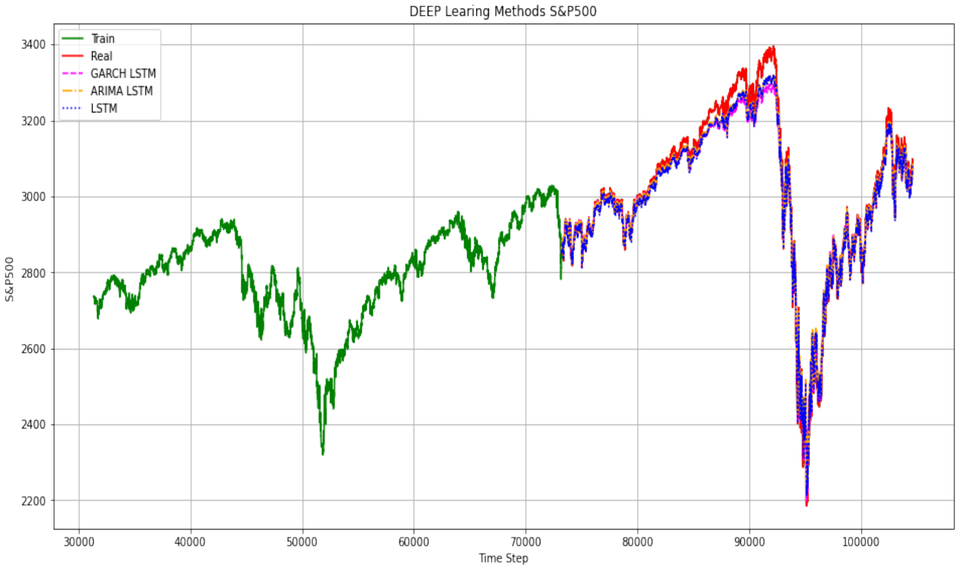


FIGURE 4.64: SNP Deep Learning Forecasting (10 Minutes Data)

31.342 for MSE, MAE, MAPE, and RMSE respectively. Figure 4.64 exhibit the train, test, real and forecasted series of SNP500. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta colore represent the LSTM, ARIMA-LSTM, and GARCH-LSTM results respectively. Results show that ARIMA-LSTM has lowest forecasting accuracy indicators and is close to actual time series so it is the best fitted deep learning method.

This section concludes the estimation and forecasting of machine learning and deep learning methods. The study use SVM, ARIMA-SVM, LSTM, ARIMA-LSTM, and GARCH-LSTM model to forecast the stock market indices. Results show a mixed result between the deep learning battle. GARCH-LSTM and ARIMA-LSTM show better forecasting results for 10 minutes data frequency.

4.2.5 Discussion

Classical, machine learning, and deep learning methods are used to examin the ten minutes data frequency. Results given in table 4.10 show that information impact increase while increase in frequency. Market randomness decreases with the identification of autoregressive and moving average in more markets then the previously analyzed data in section 4.1.3.1. As we increase the frequency dependencies on information has been increased. Similarly, for long memory phenomenon (d parameter) presence increase with more number of statistical significant p-values but still a clear conclusion about the presense of long memory can not be taken, however traces of some long-run memory are observed. The study results of ARFIMA model are not align with the proposed criteria of (Baillie, 1996). Though the study results are align with the previous literature of using ARFIMA i.e., (Floros et al., 2007; Reisen et al., 2001).

Based on results provided in table 4.12, and table 4.13, The study compare various model on basis of forecasting accuracy indicators. This part of research analyze the selected models on different data frequencies. In order to answer the research questions, modified deep learning, machine learning and classical forecasting method has been used which are previously used by or proposed by various studies mentioned in previous section 4.1.5 of discussion. As above mentioned, GARCH-LSTM model also been tested for forecasting with 10-minute data frequency for each stock market index. Table 4.14 answer the questions i.e., "Are classical models successful in forecasting price behavior in financial markets? How do machine leaning based models perform in forecasting price trends in financial markets? Do machine learning based models outperform classical models of forecasting?".

Results shows that our proposed GARCH-LSTM a Deep learning method out perform the other deep learning, machine learning and classical forecasting methods suggested by the Li et al. (2020), in forecasting A50, DAX, HIS, NIFTY, NIKKEI, IBX, SMI, FTSE, and DJI30. For ASX, EUS, CAC, AEX, WIG, SSGF and SNP500 stock markets ARIMA-LSTM perfectly fit in with these markets and is consistent with ARIMA-LSTM performance already explored by various studies (Alam et al., 2020; Kulshreshtha, 2020; Li et al., 2020). DJI with only stock market which works better with LSTM method then the hybrid ARIMA and GARCH-LSTM method in 10-minute frequency. For stock market ASX and EUS GARCH-LSTM performed as second-best method to forecast the respective stock market. Whereas, A50, DAX, IBX, and FTSE shows ARIMA-LSTM as second best fitted forecasting method. In classical forecasting techniques we witnessed ARIMA as best fitted candidate for EUS, CAC, DAX, HIS, SMI, FTSE, DJI30, and SNP500 stock market indices for 10-minute date frequency data. Because of ARIMA model strengthen in performing better than the GARCH in out of sample forecasting (Crawford and Fratantoni, 2003).

Results showed Deep learning-based methods outperform the conventional machine learning and classical forecasting technique which are consistent with the finds of other studies (Choi, 2018; Fang and Yuan, 2019; Fischer and Krauss, 2018; Kulshreshtha, 2020; Li et al., 2020; Temür et al., 2019; Yan and Ouyang, 2018). Aforementioned forecasting accuracy indicators MSE, RMSE, MAE and MAPE clearly provide that hybrid/deep learning-based methods performs better than other methods that is inline with research objectives so, accept our hypothesis H3 of study deep learning outperform conventional machine learning and classical forecasting methods is accepted.

TABLE 4.14: Models Comparison (10 Minutes Data)

	Model Rank (10 MINT)							
	1st RANK	**2nd RANK**	**3rd RANK**	**4th RANK**	**5th RANK**	**6th RANK**	**7th RANK**	**8th RANK**
ASX	ARIMA-LSTM	GARCH-LSTM	LSTM	SVM	ARIMA-SVM	ARFIMA	ARFIMA	GARCH
A50	GARCH-LSTM	ARIMA-LSTM	LSTM	ARIMA-SVM	SVM	ARFIMA	ARIMA	GARCH
EUS	ARIMA-LSTM	GARCH-LSTM	LSTM	ARIMA-SVM	SVM	ARIMA	ARIMA	GARCH
CAC	ARIMA-LSTM	LSTM	GARCH-LSTM	ARIMA-SVM	SVM	ARIMA	ARFIMA	GARCH
DAX	GARCH-LSTM	ARIMA-LSTM	LSTM	ARIMA-SVM	SVM	ARIMA	ARFIMA	GARCH
HSI	GARCH-LSTM	LSTM	ARIMA-LSTM	ARIMA-SVM	SVM	ARIMA	ARFIMA	GARCH
NIFTY	GARCH-LSTM	LSTM	ARIMA-LSTM	SVM	ARIMA-SVM	ARFIMA	ARFIMA	GARCH
NIKKEI	GARCH-LSTM	LSTM	ARIMA-LSTM	SVM	ARIMA-SVM	ARFIMA	ARFIMA	GARCH
AEX	ARIMA-LSTM	LSTM	GARCH-LSTM	ARIMA-SVM	SVM	ARFIMA	ARIMA	GARCH
WIG	ARIMA-LSTM	LSTM	GARCH-LSTM	ARIMA-SVM	SVM	GARCH	ARFIMA	ARIMA
SSGF	ARIMA-LSTM	LSTM	GARCH-LSTM	SVM	ARIMA-SVM	ARFIMA	ARFIMA	GARCH
IBX	GARCH-LSTM	ARIMA-LSTM	LSTM	SVM	ARIMA-SVM	ARFIMA	ARFIMA	GARCH
SMI	GARCH-LSTM	LSTM	ARIMA-LSTM	ARIMA-SVM	SVM	ARIMA	ARFIMA	GARCH
FTSE	GARCH-LSTM	ARIMA-LSTM	LSTM	ARIMA-SVM	SVM	ARIMA	ARIMA	GARCH
DJI	GARCH-LSTM	LSTM	ARIMA-LSTM	ARIMA-SVM	SVM	ARIMA	ARFIMA	GARCH
SNP	ARIMA-LSTM	LSTM	GARCH-LSTM	ARIMA-SVM	SVM	ARIMA	ARIMA	GARCH

4.3 Section III (5 Minutes Data Frequency)

In this Section, all of the classical, machine learning, and deep learning methods are examined using 5 minutes data frequency. This section of the study is further divided into five sub-sections. Section one and two deal with the data descriptives and pre-processing. Section two and three discuss the estimated results of all models and the section discuss the models comparison.

4.3.1 Descriptive Statistics

Table 4.15: Descriptive Statistics (5 Minutes Date)

	Mean	**Median**	**Std. Dev.**	**Skew**	**Kurt**	**J.B**
ASX	6124.40	6071.47	442.21	0.07	2.88	263.48
A50	12724.08	12872.97	1063.59	-0.34	2.07	9643.02
EUS	3399.04	3443.27	244.87	-1.11	4.72	42956.04
CAC	5305.91	5358.28	393.28	-0.75	4.38	22814.47
DAX	12198.20	12313.41	887.09	-1.03	4.61	59127.89
HSI	27721.95	27787.97	2089.11	-0.08	2.84	270.84
NIFTY	10931.01	10919.50	861.01	-0.82	3.98	21292.81
NIKKEI	21719.74	21802.47	1346.14	-0.57	3.40	12477.09
AEX	549.70	552.36	34.04	-0.61	4.79	22112.77
WIG	2185.07	2238.80	236.84	-1.53	5.28	39167.40
SSGF	361.79	364.92	30.34	-1.06	4.09	39314.15
IBX	9302.84	9411.01	943.21	-1.40	4.93	54290.84
SMI	9400.78	9280.97	620.47	0.68	2.90	10044.51
FTSE	7190.06	7331.40	518.69	-1.94	6.64	244033.60
DJI	25203.74	25305.07	1886.93	-0.21	3.05	1479.66
SNP	2806.83	2794.17	220.76	0.34	2.82	4423.48

Table 4.15 depicts the data summary of 5-minutes data frequency. Mean values shows the average indices points for the selected time period. ASX, A50, EUS, CAC, DAX, HIS, NIFTY, NIKKEI, AEX, WIG, SSGF, IBX, SMI, FTSE, DJI 30 and SNP 500 has the mean values of 6124.40, 12724.08, 3399.04, 5305.91, 12198.20, 27721.95, 10931.01, 21719.74, 549.70, 2185.07, 361.79, 9302.84, 9400.78, 7190.06, 25203.74, and 2806.83 respectively. Similarly, standard deviations explain the data desperation from its mean. To measure the location of data we generally interpret the skewness and kurtosis figures. Skewness explain the left are right skewed of

data observations. In our data set A50, EUS, CAC, DAX, HIS, NIFTY, NIKKEI, AEX, WIG, SSGF, IBX, FTSE, and DJI 30 has negative values (-0.34 for A50) which referred as left skewed data. Whereas ASX, SMI and SNP 500 are right tailed with positive skewness values (0.07, 0.68, 0.34) respectively. ASX, A50, HIS, SMI, DJI30, and SNP500 have the platykurtic Behavior and rest of sample has the leptokurtic behavior having kurtosis values greater than 3. Jarque-Bera normality test indicates the none of variable is normally distributed.

4.3.2 Data Pre-Processing

The first step into in-depth analysis preprocessing an essential part of the analysis for 5-minute data frequency. To check the stationarity of data ADF has been used. Results show the non-normality of data with insignificant p values presented in the table 4.16.

TABLE 4.16: Data Pre-Processing (5 Minutes Data)

	Stationarity			**ARCH Lm Test**
	Level	1st Diff.	2nd Diff.	p-value
ASX	0.1418	0.0000	-	0.0000
A50	0.1868	0.0000	-	0.0000
EUS	0.1416	0.0000	-	0.0000
CAC	0.1880	0.0000	-	0.0000
DAX	0.1410	0.0000	-	0.0000
HSI	0.2779	0.0000	-	0.0000
NIFTY	0.2520	0.0000	-	0.0000
NIKKEI	0.1270	0.0000	-	0.0000
AEX	0.0814	0.0000	-	0.0000
WIG	0.7366	0.0000	-	0.0000
SSGF	0.6660	0.0000	-	0.0000
IBX	0.7023	0.0000	-	0.0000
SMI	0.1290	0.0000	-	0.0000
FTSE	0.4432	0.0000	-	0.0000
DJI	0.0529	0.0000	-	0.0000
SNP	0.1245	0.0000	-	0.0000

After the first difference, results shows the significant p-values of the ADF test for all series. None of the series has an insignificant p-value of ADF test after applying the first difference. Similarly, the ARCH-LM test has also been used to check whether data has the ARCH effect. Results are given in the table 4.16 show existence of the ARCH effect in data. Figure C-1 in appendix KDE plot also depict the data distribution outlook for 5-minute data frequency.

4.3.3 Classical Forecasting Methods

This analysis section has been further divided into two sessions: Model Estimation and selection, and Forecasting.

4.3.3.1 Models Selection and Estimation

Predefined tools and parameters have been used to for the selection and estimation of classical methods for 5-minute data frequency. Table 4.17 illustrates the estimated results of the ARIMA model for all the data sets. Table 4.18 shows the computed results for ARFIMA and GARCH models. Again 70:30 ratios have been used for training and test data.

Estimated results shows a random walk of prices (ARIMA order (0,1,0) has been only found in DAX and HIS stock index based on AIC values (970551 and 831651) and BIC values (970561 and 831661) respectively. None of those mentioned above stock indices found any AR and MA term and no seasonality (SAR and SMA), has been found in stepwise model computation. Sigma2 value has the coefficient values of 45.1745 and 525.1620 respectively for DAX and HIS stock index with significant p-values. KDE plot of residuals for estimated ARIMA model also has been shown in the figure C-2 appendix for the series mentioned above. That verifies the model's fitness.

A50, CAC, wig 20, and SMI stock market indices shows that the ARIMA order (3,1,0 for A50, 2,1,0 for wig 20 and 1,1,0 for rest of mentioned markets) as the best-fitted model with AIC values (939277, 478860, 206612, and 561153) and BIC values (939316, 478879, 206639, and 561172) respectively for given markets. AR term has a coefficient values-0.0108, -0.0135, 0.0151, and -0.0236 with a significant p-value for particular stock markets. That means the existence of a pattern in these markets and returns has both positive and negative linked with their lag values. Estimated results do not show any of MA terms, which means any delayed information adjustment. No SAR and SMA have been identified, which refers to no seasonality in data. KDE plots of residuals have been shown in figure C-2 appendix validates smooth the data after applying the model and indicates the model fitness for respected A50, CAC, Wig 20 and SMI stock market indices.

Based on results, AIC values 407510, 1182807, -5637, 532466 and BIC values 407529, 1182827, -5617, and 532485 for EUS, NIKKEI, SSGF and IBX 30 stock market indices ARIMA (0,1,1) has been identifies as the best-fitted model for respective stock market indices. Estimated results given in table 4.17 show the coefficient value of MA term -0.0331, -0.0193, -0.0988, and -0.0277 with significant p-value respectively for the aforementioned stock markets. MA term coefficient vales shows the delayed and overly priced adjustment of information for all respective markets. Results show significant existence of variance with a significant p-value where, sigma2 with coefficient values of 4.9482, 193.9165, 0.0558, and 51.5850 respectively. KDE Plot of residuals for respected stock market indices (EUS, NIKKEI, SSGF, and IBX) also been presented in appendix (figure C-2) that shows the model fitness.

Results shows that ARIMA (2,1,1) model has been identified for FTSE stock market index on basis of minimum AIC value (760881) and BIC value (760921). AR term has the coefficient value of 0.0104 with a significant p-value (0.0000) which indicates a positive relationship with lagged returns values and FTSE current prices follows the patterns that exist in market. Whereas the MA term with a coefficient value of -0.822 and a significant p-value indicates the positive relationship with lagged values of the error term. Also showed the delayed and overly prices informational adjustment in FTSE stock market index.

There is no seasonality (SAR and SMA) term that has been identified to represent the seaonality in the FTSE stock market index. Sigma2 coefficient values of 11.4077 with significant p-values indicate the existence of variance in the mean equation of the model. Figure C-2 in the appendix Also indicates the model fitness which smoothens the KDE of residuals.

Table 4.17: ARIMA Estimation

	ASX	A50	EUS	CAC	DAX	HSI	NIFTY	NIKKEI	AEX	WIG	SSGF	IBX	SMI	FTSE	DJI	SNP
Method								Log Likelihood								
Model	(1,1,2)	(3,1,0)	(0,1,1)	(1,1,0)	(0,1,0)	(0,1,0)	2,1,4	0,1,1	1,1,1	2,1,0	1,1,0	0,1,1	1,1,0	2,1,1	1,1,1	1,1,1
AIC	653197	939277	407510	478860	970551	831651	663222	1182807	58073	206612	-5637	532466	561153	760881	1172683	517301
BIC	653237	939316	407529	478879	970561	831661	663298	1182827	58101	206639	-5617	532485	561172	760921	1172712	517331
AR	0.6801	-0.011	-	-0.014	-	-	0.6846	-	-0.198	0.0151	-	-	-0.02	0.0104	-0.759	-0.253
	[0.000]	[0.000]	-	[0.000]	-	-	[0.000]	-	[0.000]	[0.000]	-	-	[0.000]	[0.000]	[0.000]	[0.000]
MA	0.0171	-	-0.033	-	-	-	0.0153	-0.0193	0.145	-	-0.099	-0.028	-	-0.822	0.7484	0.2381
	[0.001]	-	[0.000]	-	-	-	[0.000]	[0.000]	[0.001]	-	[0.000]	[0.000]	-	[0.000]	[0.000]	[0.000]
SAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SMA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sigma2	7.61	120.88	4.95	10.83	45.17	525.16	47.85	193.92	0.12	5.57	0.06	51.59	27.20	11.41	180.91	2.03
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Obs	134202	123060	91846	91728	145983	91375	98898	145931	78666	45351	115887	78523	91375	144320	145931	145931

For stock market indices AEX, DJI30 and SNP500 results shows that the AIC values (58073, 1172683 and 517301 respectively) and BIC values (58101, 1172712 and 517331 respectively) ARIMA (1,1,1) model has been identified as the best-fitted model for respective indices. AR term for both models has a significant p-value with coefficient values of -0.1977, -0.7593 and -0.2526. This means the existence of patterns in financial series and significant relation with its own lag values. MA term also significant p-values with coefficient values of 0.1454, 0.7484 and 0.2381 respectively. Significant coefficient values refers as information adjustment in series. Stepwise models shows no SAR and SMA seasonality patterns in 5 minutes of data frequency in AEX, DJI30 and SNP500 market indices. KDE plot of residuals shows the model fitness. Significant values of Sigma2 with coefficient values of 0.1225, 180.9142, and 2.0278 for respective markets show the presence of variance in markets.

Both ASX and NIFTY market indicis have estimated results presented in table 4.17 shows that the ARIMA (1,1,2) for ASX and ARIMA(2,1,4) order as best fitted model based on AIC values (653197 and 663222) and BIC value (653237 and 663298)respectively. Both AR and MA terms have significant p values with coefficients values (0.6801 and 0.6846) for AR term and (0.0171 and 0.0153) for MA term respectively. Statistical values illustrate the presences of autocorrelation in lag value of both markets and MA term explains the informational adjustment of data. No seasonality pattern has been identified in the respective stock markets. A significant Sigma2 value with a coefficient values 7.6094 and 47.8532 shows the existence of variance in both markets. KDE plot of residuals represented in figure C-2 of appendix also indicates the validation of the model fitness.

Table 4.18 depicts the ARFIMA model estimations for all the data sample for 5-minute data frequency. All of the markets follow the same p and q order for model estimation resulted in ARIMA model estimation given in table 4.17. For the stock market indices DAX and HIS, computed AIC values are -12.1403 and -11.4052 respectively. Estimated d parameter have the values -0.0042 and 0.0124 respectively for all aforementioned markets. With significant statistical value DAX follows the intermediate memory. Whereas, HSI process long memory in autocovariance function.

For A50, WIG, and IBX 30, SMI stock market indices identified ARIMA order (3,1,0), (2,1,0), (0,1,1), and (0,1,1), based on these p and q order ARFIMA model has been computed with an estimated AIC value of -11.2109, -11.5491 -10.8971, and -12.0723. The estimated d parameter has the values -0.0007 0.0055, 0.0006, and 0.0014 with an insignificant p-value. Based on statistical p-value ARFIMA model does not fit for these stock markets. For CAC market indices has the AIC value -10.8363 mentioned in table 4.18. Estimated d parameter has the value are 0.0070 with significant p-values respectively. Estimated coefficient values also show the intermediate memory presence. Significant p-values of AR term coefficient has values -0.0249 indicates the autocorrelation with lag values of series.

For stock market indices EUS, NIKKEI, and SSGF market indices has the AIC values -11.7943, -11.2690, and -11.8789 respectively presented in table 4.18. Estimated d parameter values are -0.0076, -0.0115, and -0.0095 with significant p-values respectively. Estimated coefficient values also shows the long memory in auto correlation function and intermediate memory presence. Significant p-values of MA term coefficient values -0.0249, -0.0033, and -0.0848 indicates the overly priced information in markets. For the remaining Stock markets indices ASX, NIFTY, AEX, FTSE, DJI 30 and SNP 500 followed the same p and q order presented in table 4.18 with AIC values of -12.5171, -11.8776, -11.8295, -11.4735, -12.1877, and -11.3367 in table 4.18 respectively for all indices.

Estimated d parameter has the values of -0.0321, -0.0209, -0.0069, -0.0043, 0.0040, and -0.0180 respectively with significant statistical p-value. Based on estimated results intermediate memory persist in series. AR term has significant p-values with a coefficient value of 0.8533, -0.2992, -0.0107, and-0.7408, indicates the autocorrelation with lag values of series except NIFTY and SNP 500. Similarly, significant p-values of MA term with coefficient values of 0.0066, 0.0098, 0.2553, 0.3658, and 0.7280 indicates the adjustment of shocks in prices respectively except insignificant coefficient value of SNP 500 index.

For the non-parametric model in the classical forecasting method GARCH (1,1) again has been estimated in 5-minute time interval data. Estimated AIC values results for all market has been given in table 4.18. ASX, A50, EUC, CAC, DAX, HSI, NIFTY, and NIKKEI market indices have the lag coefficient values of -0.0426, -0.0320, -0.0482, -0.0215, -0.0099, -0.0178, -0.0222, and -0.0308 with significant p-values respectively. Results indicate the inverse relation with lag returns of these stock markets. α_1 have a coefficient values 0.0496, 0.2773, 0.1424, 0.1398, 0.1498, 0.0533, 0.1486, and 0.0218 for aforementioned stock indices respectively with significant p-value. Similarly, β_1 coefficients values 0.9094, 0.6921, 0.8054, 0.8124, 0.8115, 0.9196, 0.7357, and 0.9732 with significant p-value. These statistical values show the existence of the ARCH effect and persistence of GARCH in data.

For stock indices AEX, WIG, SSGF, IBX, SMI, FTSE, DJI30, and SNP500 estimated for the GARCH (1,1) model has been presented in table 4.18. Estimated AIC values have been presented as well. Aforementioned market indices have the lag coefficient values of -0.0593, -0.0300, -0.1107, -0.0281, -0.0353, -0.0206, -0.0316, and -0.0388. The significant p-value of these coefficients indicates the existence of an inverse relationship between the lag returns and prices of these stock market indices. Coefficient values of α_1 indicate the existence of the ARCH effect in markets. Coefficient values 0.1071, 0.0923, 0.1410, 0.0038, 0.1441, 0.0588, 0.1897, and 0.2051 with significant p-values support our argument of volatility existence. Similarly, β_1 coefficients show the persistence of GARCH in data. Coefficient values of GARCH 0.7553, 0.7822, 0.7711, 0.9949, 0.8525, 0.9333, 0.7592, and 0.7525 with significant p-value validate the argument of GARCH persistence in given data set.

Table 4.18: ARFIMA and GARCH Estimation (5 Minutes Data)

	ASX	**A50**	**EUS**	**CAC**	**DAX**	**HSI**	**NIFTY**	**NIKKEI**	**AEX**	**WIG**	**SSGF**	**IBX**	**SMI**	**FTSE**	**DJI**	**SNP**
								AFRIMA Statistics								
Method								Log Likelihood								
Model	(1,1,2)	(3,1,0)	(0,1,1)	(1,1,0)	(0,1,0)	(0,1,0)	(2,1,4)	(0,1,1)	(1,1,1)	(2,1,0)	(1,1,0)	(0,1,1)	(1,1,0)	(2,1,1)	(1,1,1)	(1,1,1)
AIC	-12.52	-11.21	-11.79	-10.84	-12.14	-11.41	-11.88	-11.27	-11.83	-10.90	-11.88	-11.55	-12.07	-11.47	-12.19	-11.34
d	-0.032	-0.001	-0.008	0.007	-0.004	0.012	-0.021	-0.012	-0.007	0.006	-0.009	0.001	0.001	-0.004	0.004	-0.018
	[0.000]	[0.470]	[0.0052]	[0.000]	[0.000]	[0.047]	[0.000]	[0.000]	[0.0214]	[0.3310]	[0.0001]	[0.859]	[0.6006]	[0.000]	[0.000]	[0.000]
AR	0.8533	-0.01	-	-0.02	-	0.63	0.00	0.00	-0.30	0.01	-	-	-0.03	-0.01	-0.74	0.18
	[0.000]	[0.000]	-	[0.000]	-	[0.000]	[0.9383]	[0.000]	[0.000]	[0.0157]	-	-	[0.000]	[0.000]	[0.000]	[0.3198]
MA	0.01	-	-0.02	-	-	-0.64	0.01	0.00	0.26	-	-0.08	-0.03	-	0.37	0.73	-0.18
	[0.067]	-	[0.000]	-	-	[0.000]	[0.000]	[0.0034]	[0.000]	-	[0.000]	[0.000]	-	[0.000]	[0.000]	[0.3125]
								GARCH Statistics								
Method								Maximum Likelihood ARCH								
AIC	-463722	-295392	-252632	-262244	-475805	-215252	-283120	-164232	176	-32521	-329621	-40621	-276005	-517178	-515004	-532162
Lag	-0.04	-0.03	-0.05	-0.02	-0.01	-0.02	-0.02	-0.03	-0.06	-0.03	-0.11	-0.03	-0.04	-0.02	-0.03	-0.04
	[0.000]	[0.000]	[0.000]	[0.000]	[0.0004]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
α_1	0.050	0.277	0.142	0.140	0.150	0.053	0.149	0.022	0.107	0.092	0.141	0.004	0.144	0.059	0.190	0.205
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
β_1	0.909	0.692	0.805	0.812	0.812	0.920	0.736	0.973	0.755	0.782	0.771	0.995	0.852	0.933	0.759	0.752
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Obs	1E+05	1E+05	91846	91728	1E+05	91375	98898	1E+05	78666	45351	1E+05	78523	91375	1E+05	145931	1E+05

4.3.3.2 Models Forecasting

After the model estimation in previous section by using 70% of data for training. This section has used 30% data to forecast the all above estimated models. Forecasted results has been presented in table 4.19 and presentation of respective stock market.

Figure C-3 in appendix exhibit the forecast of estimated trained and test ARIMA model for ASX stock market index. Forecasting accuracy indicators (MSE, MAPE, RMSE, and MAE) results indicates presented in table 4.19 as MSE (515281), RMSE (717.83), MAE (496.800), and MAPE (8.81285). Results shows the sudden drop in forecasted windows due failed the ARIMA forecasting method to predict in such case as arrival of Covid-19 in this time frame. Whereas, figure C-4 in appendix also exhibit the both actual and forecasted ARFIMA series with Estimated MSE (266004), RMSE (515.76), MAE (299.601), and MAPE (5.24303). For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used and exhibit in figure C-5 in appendix. Forecasting Accuracy indicators are MSE (2810448), RMSE (1676.44), MAE (1318.582), and MAPE (21.83477). Based on results and minimum values of evaluation criteria of models ARFIMA model with lowest MSE, MAE, MAPE and RMSE values has better prediction power then the ARIMA and GARCH model in 5-minute time frequency for ASX stock market index.

For the stock market A50 index, both trained and test series have been shown in figure C-3 in appendix for the first conventional classical forecasting model ARIMA. The lower and upper band shows the perfection of ARIMA for the A50 market index. Forecasting Indicators, MSE, MAE, MAPE, and MAPE have the values of 586061, 669.361, 4.85114, and 765.55 respectively presented in table 4.19. ARFIMA estimated accuracy indicators MSE, MAE, MAPE, and RMSE presented in table 4.19. Forecasting accuracy indicators have the values as, MSE (968121), MAE (787.106), MAPE (6.31806), and RMSE (983.93). Figure C-4 in appendix exhibit the ARFIMA forecasted and actual series for the A50 stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure

C-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 78257307, 7180.243, 55.90998, and 8846.32. Computed values of Accuracy indicator of ARIMA model declared the winner within the classical forecasting techniques for 5-minute data frequency.

Forecasting Indicators of classical Models for EUS stock market index has been presented in table 4.19. Estimated MSE, MAE, MAPE, and RMSE has the statistical values of 147994, 336.622, 10.28810, and 384.70 respectively for the EUS stock market index. Similarly, a Graphical representation has been given in figure C-3 in appendix. Which shows a sudden fall decreases the ARIMA model prediction power in the testing phase. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 58244, 180.712, 5.42000, and 241.34 respectively for the EUS stock market index given in table 4.19. Figure C-4 in appendix depicts the visual outlook of the actual and forecasted series.

The model choose 100 observations to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure C-5 in appendix. Estimated forecasted indicators MSE, MAE, MAPE, and RMSE have the values of 4446481, 1741.140, 52.92349, and 2108.67 respectively for the EUS stock market index. The aforementioned evolution indicators result shows the ARFIMA model having the lowest indicator values within the classical forecasting models. That makes ARFIMA as best fitted forecasting method in classical forecasting models for EUS stock market index under 10minute data frequency.

For the stock market index, CAC, both trained and test series have been shown in figure C-3 in appendix for the predefined ARIMA model. Similarly, predefined Lower and upper band show the ARIMA model forecasting strength for the CAC market index. Forecasting accuracy indicators, MSE, MAE, MAPE, and RMSE have the values as 419546, 581.710, 11.27041, and 647.72 respectively for the CAC stock market index. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 216137, 371.582, 6.91595, and 464.91 respectively for the CAC stock market index given in table 4.19. Figure C-4 in appendix depicts the visual outlook of the actual and forecasted series.

TABLE 4.19: Classical Models Forecasting (5 Minutes Data)

	ASX	A50	EUS	CAC	DAX	HSI	NIFTY	NIKKEI
				Forecasting indicators for ARIMA				
MSE	515281	586061	147994	419546	1771037	5862127	2843839	4650414
MAE	496.800	669.361	336.622	581.710	1145.685	1934.112	1267.658	1858.107
MAPE	8.813	4.851	10.288	11.270	9.633	7.854	13.271	8.414
RMSE	717.830	765.546	384.700	647.724	1330.803	2421.183	1686.369	2156.482
				Forecasting indicators for ARFIMA				
MSE	266004	968121	58244	216137	832977	7340171	973031	2489099
MAE	299.601	787.106	180.712	371.582	726.252	2192.106	620.194	1246.840
MAPE	5.243	6.318	5.420	6.916	5.982	7.901	6.116	5.741
RMSE	515.756	983.931	241.338	464.905	912.676	2709.275	986.423	1577.688
				Forecasting indicators for GARCH				
MSE	2810448	78257307	4446481	2254691	54515452	271878547	5221993	111975973
MAE	1318.582	7180.243	1741.140	1182.378	6076.148	12660.730	1625.901	8321.731
MAPE	21.835	55.910	52.923	23.167	51.477	48.059	15.732	38.911
RMSE	1676.439	8846.316	2108.668	1501.563	7383.458	16488.740	2285.168	10581.870

Table No 4.19b Classical Models Forecasting (5 Minutes Data)

	AEX	WIG	SSGF	IBX	SMI	FTSE	DJI	SNP
				Forecasting indicators for ARIMA				
MSE	2822	74974	1891	1431439	236893	708351	5355217	73376
MAE	39.215	189.869	31.809	960.121	381.179	584.503	1896.115	225.228
MAPE	7.609	11.578	10.539	12.641	4.089	9.715	7.406	7.395
RMSE	53.126	273.813	43.488	1196.428	486.716	841.636	2314.134	270.880
				Forecasting indicators for ARFIMA				
MSE	1362	106496	633	1423287	278814	413546	3823257	29376
MAE	25.765	224.129	18.574	1027.035	412.816	542.129	1548.837	132.188
MAPE	4.926	13.728	5.532	13.532	4.494	7.494	6.220	4.681
RMSE	36.906	326.337	25.150	1193.016	528.028	643.076	1955.315	171.395
				Forecasting indicators for GARCH				
MSE	6066	42702	22172	2872154	15705528	2229742	26953832	172529
MAE	58.366	132.459	123.826	1357.945	3358.575	1134.026	3719.592	302.854
MAPE	11.704	7.060	36.322	16.110	35.202	16.862	14.740	10.671
RMSE	77.883	206.645	148.903	1694.743	3963.020	1493.232	5191.708	415.367

For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure C-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 2254691, 1182.378, 23.16659, and 1501.56. Based on computed values of the Accuracy indicator of All models ARFIMA model has the lowest values of these evaluation criteria. Which declares the ARFIMA as the winner in the classical forecasting techniques for the CAC stock market index.

figure C-3 in appendix exhibit the forecast of estimated trained and test ARIMA model for the DAX stock market index. Results of forecasting accuracy indicators are presented in Table 4.19, where MSE (1771037), RMSE (1330.80), MAE (1145.685) and MAPE (9.63254). Results show the sudden drop in forecasted windows due to failed the ARIMA forecasting method to predict in such case as the arrival of Covid-19 in this time frame. Whereas, figure C-4 in appendix also exhibit the both actual and forecasted ARFIMA series with Estimated MSE (832977), RMSE (912.68), MAE (726.252), and MAPE (5.98228). For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for for the display of forecasting in figure C-5 in appendix. Forecasting Accuracy indicators are MSE (54515452), RMSE (2036.505), MAE (7383.46), and MAPE (51.47657). Based on results and minimum values of evaluation criteria of models ARFIMA model with the lowest has better prediction power than the ARIMA and GARCH model in 5-minute time-frequency for the DAX stock market index.

For the stock market HSI index, both trained and test series have been shown in figure C-3 in appendix for our first conventional classical forecasting model ARIMA. The lower and upper band shows the perfection of ARIMA for the HSI market index. Which is the result of Hong Kong government policies to counter the economic effect due to covid-19. Forecasting Indicators MSE, MAE, MAPE, and RMSE have the values of 5862127, 1934.112, 7.85384, and 2421.18 respectively presented in table 4.19. ARFIMA estimated accuracy indicators MSE, MAE, MAPE, and RMSE presented in table 4.19. Computed forecasting accuracy indicators have the values of, MSE (7340171), MAE (2192.106), MAPE (7.90131), and RMSE (2709.28). Figure C-4 in appendix exhibit the ARFIMA forecasted and actual series for the HSI stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure C-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 271878547, 12660.730, 48.05946, and 16488.74. Computed values of Accuracy indicator of ARIMA model declared the winner within the classical forecasting techniques for 5-minute data frequency.

Forecasting Indicators of classical Models for NIFTY stock market index has been presented in table 4.19. Estimated MSE, MAE, MAPE, and RMSE has the statistical values of 2843839, 1267.658, 13.27134, and 1686.37 respectively for the NIFTY stock market index. Similarly, a Graphical representation has been given in figure C-3 in appendix. Which shows a sudden fall decreases the ARIMA model prediction power in the testing phase. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 973031, 620.194, 6.11624, and 986.42 respectively for the Nifty stock market index given in table 4.19. Figure C-4 in appendix depicts the visual outlook of the actual and forecasted series. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure C-5 in appendix. Computed forecasted indicators MSE, MAE, MAPE, and RMSE have the values of 5221993, 1625.901, 15.73155, and 2285.17 respectively for the NIFTY stock market index. The aforementioned evolution indicators result shows the ARFIMA model having the lowest indicator values within the classical forecasting models. That makes ARFIMA as best fitted forecasting method in classical forecasting models for NIFTY stock market index under 5-minute data frequency.

For the stock market index, NIKKEI both trained and test series have been shown in figure C-3 in appendix for the predefined ARIMA model. Similarly, predefined Lower and upper band at 95% confidence interval also show the ARIMA model forecasting strength for the NIKKEI market index. Forecasting accuracy indicators, MSE, MAE, MAPE, and RMSE have the values as 4650414, 1858.107, 8.41424, and 2156.48 respectively for the NIKKEI stock market index. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE 2489099, 1246.840, 5.74119, and 1577.69 respectively for the NIKKEI stock market index given in table 4.19. Figure C-4 in appendix depicts the visual outlook of the actual and forecasted series. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure C-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 111975973, 8321.731, 38.91135, and 10581.87. Based on computed values of the Accuracy indicator of All models ARFIMA model has the lowest values of these evaluation criteria. Which declares the ARFIMA as the winner in the classical forecasting techniques for the NIKKEI stock market index.

Figure C-3 in appendix exhibit the forecast of estimated trained and test ARIMA model for AEX stock market index. Results of forecasting accuracy indicators (MSE, MAPE, RMSE, and MAE) are presented in Table 4.19b. MSE (2822), RMSE (53.13), MAE (39.215), and MAPE (7.60915). The lower and upper band is also set at 95% confidence interval to the test and train data ratio. Results show the sudden drop in forecasted windows due to failed the ARIMA forecasting method to predict in such case as the arrival of Covid-19 in this time frame. Whereas, figure C-3 in appendix also exhibit the both actual and forecasted ARFIMA series with Estimated MSE (1362), RMSE (36.91), MAE (25.765), and MAPE (4.92640). For the third non-parametric conventional model GARCH we used 100 observations as out of sample as estimation window using R ruGARCH library. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for for the display of forecasting in figure C-5 in appendix. Forecasting Accuracy indicators are MSE (6066), RMSE (77.88), MAE (58.366), and MAPE (11.70398). Based on results and minimum values of evaluation criteria of models ARFIMA model with the lowest MSE, MAE, MAPE, and RMSE values have better prediction power than the ARIMA and GARCH model in 5-minute time-frequency for the AEX stock market index.

For the stock market WIG index, both trained and test series have been shown in figure C-3 in appendix for our first conventional classical forecasting model ARIMA. The lower and upper band is also set at 95% confidence interval to the test and train data ratio. Results show the sudden drop in forecasted windows due to failed the ARIMA forecasting method to predict in such case as the arrival of Covid-19 in this time frame. Forecasting Indicators MSE, MAE, MAPE, and MAPE have the values of 74974, 189.869, 11.57775, and 273.8129 respectively presented in table 4.19b. ARFIMA estimated accuracy indicators MSE, MAE, MAPE, and RMSE presented in table 4.19b. Computed forecasting accuracy indicators have the values of, MSE (106496), MAE (224.129), MAPE (13.72845), and RMSE (326.34). Figure C-3 in appendix exhibit the ARFIMA forecasted and actual series for the WIG stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure C-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 42702, 132.459, 7.05977, and 206.64. Computed values of Accuracy indicator of GARCH model declared the winner within the classical forecasting techniques for 5-minute data frequency.

Forecasting Indicators of classical Models for SSGF stock market index has been presented in table 4.19b. Estimated MSE, MAE, MAPE, and RMSE has the statistical values of 1891, 31.809, 10.53863, and 43.49 respectively for the SSGF stock market index. Similarly, a Graphical representation has been given in figure C-3 in appendix. Which shows a sudden fall decreases the ARIMA model prediction power in the testing phase. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 633, 18.574, 5.53250, and 25.15 respectively for the SSGF stock market index given in table 4.19b. Figure C-3 in appendix depicts the visual outlook of the actual and forecasted series. The model choose 100 observations to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure C-5 in appendix. Computed forecasted indicators MSE, MAE, MAPE, and RMSE have the values of 22172, 123.826, 36.32193, and 148.90 respectively for the SSGF stock market index. The aforementioned evolution indicators result shows the AFRIMA model having the lowest indicator values within the classical forecasting models. That makes ARFIMA as best fitted forecasting method in classical forecasting models for SSGF stock market index under 5-minute data frequency.

For the stock market index, IBX both trained and test series have been shown in figure C-3 in appendix for the predefined ARIMA model. Similarly, predefined Lower and upper band at 95% confidence interval also show the ARIMA model forecasting strength for the IBX market index. Forecasting accuracy indicators, MSE, MAE, MAPE, and RMSE have the values as 1431439, 960.121, 12.64146, and 1196.43 respectively for the IBX stock market index. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 1423287, 1027.035, 13.53181, and 1193.016 respectively for the IBX stock market index given in table 4.19b. Figure C-4 in appendix depicts the visual outlook of the actual and forecasted series. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in Figure C-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 2872154, 1357.945, 16.11003, and 1694.74. Based on computed values of the Accuracy indicator of all models ARIMA model has the lowest values of these evaluation criteria. Which declares the ARIMA as the winner in the classical forecasting techniques for the IBX stock market index.

Forecasting Indicators of classical Models for SMI stock market index has been presented in table 4.19b. Estimated MSE, MAE, MAPE, and RMSE has the statistical values of 236893, 381.1787, 4.08905, and 486.7162 respectively for the SMI stock market index. Similarly, a Graphical representation has been given in figure C-3 in appendix. Which shows a sudden fall decreases the ARIMA model prediction power in the testing phase. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 278814, 412.816, 4.49388, and 528.03 respectively for the SMI stock market index given in table 4.19b. Figure C-4 in appendix depicts the visual outlook of the actual and forecasted series. Our 3rd classical model GARCH, computed forecasting indicators has been presented in table 4.19b. The model choose 100 observations to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure C-5 in appendix. Computed forecasted indicators MSE, MAE, MAPE, and RMSE have the values of 15705528, 3358.575, 35.20173, and 3963.02 respectively for the SMI stock market index. The aforementioned evolution indicators result shows the ARIMA model having the lowest indicator values within the classical forecasting models. That makes ARIMA as best fitted forecasting method in classical forecasting models for SMI stock market index under 5-minute data frequency.

For the stock market FTSE index, both trained and test series have been shown in figure C-3 in appendix for our first conventional classical forecasting model ARIMA. The lower and upper band is also set at 95% confidence interval to the test and train data ratio. Results show the sudden drop in forecasted windows due to failed the ARIMA forecasting method to predict in such case as the arrival of Covid-19 in this time frame. Forecasting Indicators MSE, MAE, MAPE, and MAPE have the values of 708351, 584.503, 9.71508, and 841.64 respectively presented in table 4.19b. ARFIMA estimated accuracy indicators MSE, MAE, MAPE, and RMSE presented in table 4.19b. Computed forecasting accuracy indicators have the values of, MSE (413546), MAE (542.129), MAPE (7.49407), and RMSE (643.08). Figure C-4 in appendix exhibit the ARFIMA forecasted and actual series for the FTSE stock market. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure C-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 2229742, 1134.026, 16.86205, and 1493.23. Computed values of Accuracy indicator of ARFIMA model declared the winner within the classical forecasting techniques for 5-minute data frequency.

Forecasting Indicators of classical Models for DJI30 stock market index has been presented in table 4.19b. Estimated MSE, MAE, MAPE, and RMSE has the statistical values of 5355217, 1896.115, 7.40563, and 2314.13 respectively for the DJI30 stock market index. Similarly, a Graphical representation has been given in figure C-3 in appendix. Which shows a sudden fall decreases the ARIMA model prediction power in the testing phase. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE 3823257, 1548.837, 6.22008 and 1955.32 respectively for the DJI30 stock market index given in table 4.19b. figure C-4 in appendix depicts the visual outlook of the actual and forecasted series. Our 3rd classical model GARCH, computed forecasting indicators has been presented in table 4.19b. The model choose 100 observations to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure C-5 in appendix. Computed forecasted indicators MSE, MAE, MAPE, and RMSE have the values of 26953832, 3719.592, 14.74031, and 5191.71 respectively for the DJI30 stock market index. The aforementioned forecasting accuracy indicators result shows the ARFIMA model having the lowest indicator values within the classical forecasting models. That makes ARFIMA as best fitted forecasting method in classical forecasting models for DJI30 stock market index under 10minute data frequency.

For the stock market index SNP500, both trained and test series have been shown in figure C-3 in appendix for the predefined ARIMA model. Similarly, predefined Lower and upper band at 95% confidence interval also result show the sudden drop in forecasted windows due to failed the ARIMA forecasting method to predict in such case as the arrival of Covid-19 in this time frame. Forecasting accuracy indicators, MSE, MAE, MAPE, and RMSE have the values as 73376, 225.228, 7.39485, and 270.88 respectively for the SNP500 stock market index. For the ARFIMA model, Computed indicators values of MSE, MAE, MAPE, and RMSE are 29376, 132.188, 4.68075, and 171.40 respectively for the SNP500 stock market index given in table 4.19b. figure C-4 in appendix depicts the visual outlook of the actual and forecasted series. For GARCH model 100 observations used to train the model and 20 observations ahead forecasting windows has been used for presentation of forecasting in figure C-5 in appendix. Forecasting Accuracy indicators MSE, MAE, MAPE, and RMSE has the values of 172529, 302.854, 10.67072, and 415.37. Based on computed values of the Accuracy indicator of all models ARFIMA model has the lowest values of these evaluation criteria. Which declares the ARFIMA as the winner in the classical forecasting techniques for the SNP 500 stock market index.

The study applies three dynamic classical models to forecast our selected stock market indices. After the estimation of these models, all of the sample market indices are forecasted one by one. Results of the analysis show that the classical ARIMA model performs better in forecasting the 5 minutes data for stock market indices.

4.3.4 Machine and Deep Learning Methods

This analysis section deals with the machine learning and deep learning analysis. Section has been further divided into two subsessions: Model Estimation and Forecasting.

4.3.4.1 Models Estimation

For modal estimation under 5-minutes data frequency, predefined functions in section I and section II machine and deep learning methods have been used for the estimation. For hybrid methods again pre-decided functions and libraries has been used.

4.3.4.2 Models Forecasting

Table 4.20 enlightens the computed model's selection criteria's indicators values for each stock market for both machine learning and deep learning methods. For the ASX stock market, SVM model forecasting accuracy indicators have the values as 195010, 352.946, 5.783, and 441.600 respectively for MSE, MAE, MAPE, and RMSE. For The hybrid ARIMA-SVM accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 195140, 353.194, 5.793, and 441.746 respectively. Which worsen the results, indicates that simple SVM performs better in machine learning regression than the ARIMA-SVM model. Figure 4.65 exhibit the actual and forecasted SVM and ARIMA-SVM models for ASX.

The Deep learning methods LSTM and the hybrid LSTM included ARIMA and GARCH results presented in table 4.20 for all selected stock markets. For the ASX stock market index, LSTM estimated model accuracy indicators have the MSE, MAE, MAPE, and RMSE values 5337, 58.008, 0.904, and 73.055 respectively, which outperform the machine learning models. For the hybrid, the ARIMA-LSTM method computed forecasting accuracy indicators have the MSE, MAE, MAPE, and RMSE values of 4717, 48.171, 0.749, and 68.682. Similarly, for the hybrid GARCH-LSTM method have the values of 13467, 41.455, 0.775, and 116.048 for MSE, MAE, MAPE, and RMSE respectively.

Figure 4.66 exhibit the train, test, real and forecasted series of ASX. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta colore represent the results of LSTM, ARIMA-LSTM, and GARCH-LSTM model respectively. Results show that ARIMA-LSTM has lowest forecasting accuracy indicators and is close to actual time series so it is the best fitted deep learning method.

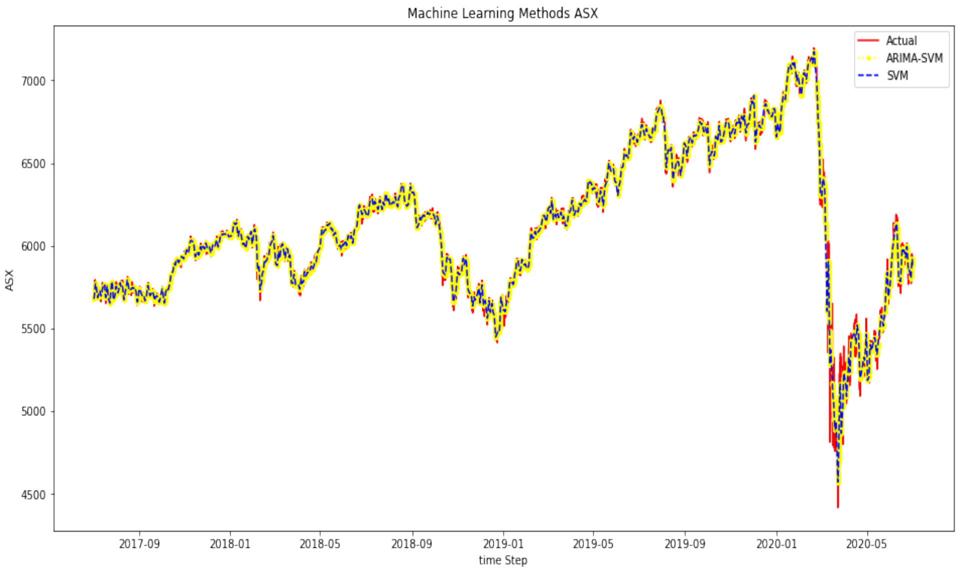


FIGURE 4.65: ASX Machine Learning Forecasting (5 Minutes Data)

Table 4.20: Machine and Deep learning Forecasting (5 Minutes Data)

	ASX	A50	EUS	CAC	DAX	HSI	NIFTY	NIKKEI	AEX	WIG	SSGF	IBX	SMI	FTSE	DJI	SNP
								SVM								
MSE	195010	1131220	59120	152986	779752	4375057	752250	1795466	1137	57236	920	875122	385897	266505	3550469	48606
MAE	352.95	904.71	180.41	284.09	669.04	1656.68	663.47	1066.15	25.13	165.69	21.48	656.35	509.90	361.39	1468.32	175.54
MAPE	5.783	7.281	5.536	5.545	5.699	6.046	6.284	5.007	4.650	8.479	6.283	7.676	5.370	5.388	5.928	6.274
RMSE	441.60	1063.59	243.15	391.13	883.04	2091.66	867.32	1339.95	33.72	239.24	30.34	935.48	621.21	516.24	1884.27	220.47
								LSTM								
MSE	5337	4986	325	38985	26417	29515	37755	4542	30	6086	343	43832	12939	33139	40228	801
MAE	58.008	67.651	11.38	85.858	87.761	146.467	111.104	49.159	2.470	44.635	10.043	129.624	82.926	75.878	135.770	22.435
MAPE	0.904	0.498	0.37	1.915	0.851	0.587	1.190	0.241	0.493	2.830	3.499	1.844	0.794	1.362	0.496	0.718
RMSE	73.055	70.612	18.03	197.445	162.533	171.800	194.306	67.391	5.505	78.015	18.514	209.360	113.749	182.042	200.569	28.309
								ARIMA-SVM								
MSE	195140	1151417	60202	154345	792739	4377786	752043	1810087	1133	59621	920	885737	401312	155080	2784442	3901
MAE	353.19	894.93	175.32	280.75	656.88	1653.72	663.55	1065.15	24.97	157.35	21.41	643.05	502.57	157.83	1205.46	28.00
MAPE	5.793	7.293	5.444	5.519	5.658	6.062	6.292	5.029	4.634	8.296	6.268	7.651	5.224	2.609	4.901	0.957
RMSE	441.75	1073.04	245.36	392.87	890.36	2092.32	867.20	1345.39	33.67	244.17	30.34	941.14	633.49	393.80	1668.66	62.46
								ARIMA-LSTM								
MSE	4717	11242	418	2511	5535	8821	25739	13155	143	9001	33	20028	13722	223010	9862	989
MAE	48.171	103.041	15.33	39.413	58.051	64.701	117.755	92.452	4.382	51.982	3.790	95.735	79.595	211.335	76.491	22.281
MAPE	0.749	0.755	0.45	0.750	0.491	0.262	1.202	0.426	0.890	3.347	1.280	1.326	0.758	3.790	0.291	0.709
RMSE	68.682	106.027	20.45	50.115	74.396	93.922	160.434	114.695	11.949	94.875	5.702	141.520	117.141	472.239	99.307	31.441
								GARCH-LSTM								
MSE	13467	503	1410	730	7286	15133	99207	18120	31	22132	211	6486	55023	77822	25215	4155
MAE	41.455	14.938	20.71	19.505	54.477	115.746	111.547	93.456	2.628	78.418	7.967	55.332	155.496	130.281	95.113	52.410
MAPE	0.775	0.111	0.68	0.370	0.505	0.448	1.294	0.442	0.516	5.099	2.754	0.735	1.472	2.306	0.367	1.671
RMSE	116.048	22.435	37.55	27.026	85.356	123.015	314.971	134.612	5.560	148.77	14.522	80.536	234.57	278.967	158.792	64.461

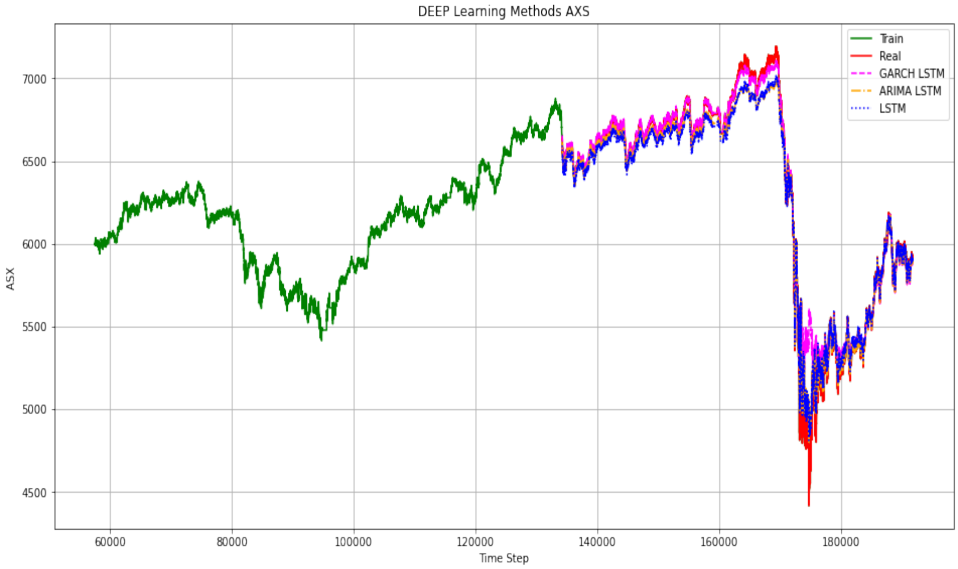


FIGURE 4.66: ASX Deep Learning Forecasting (5 Minutes Data)

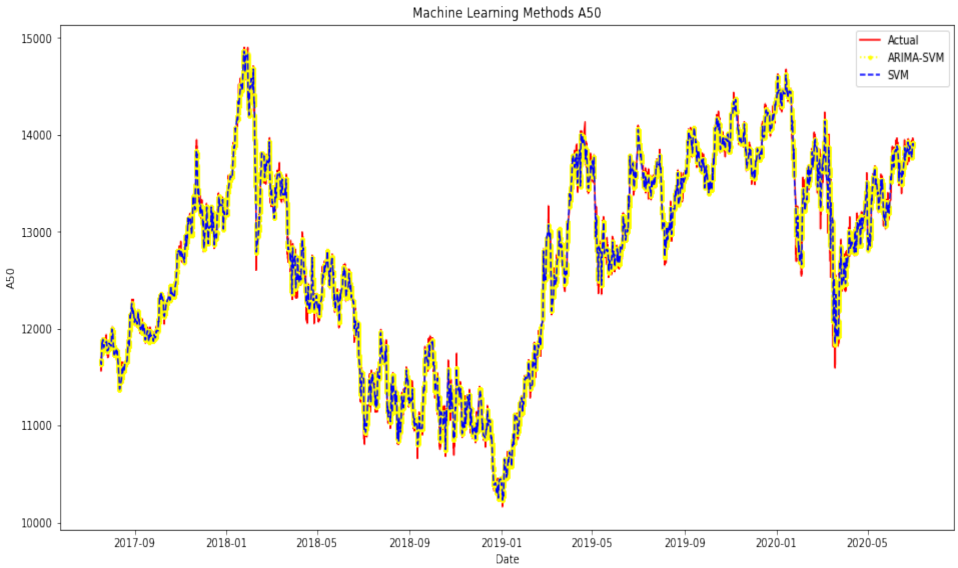


FIGURE 4.67: A50 Machine Learning Forecasting (5 Minutes Data)

For the A50 stock market Index, SVM model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have values 1131220, 904.706, 7.281, and 1063.588 respectively. For The hybrid ARIMA-SVM model, forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have the values 1151417, 894.927, 7.293, and 1073.041 respectively, which worsen the results of ARIMA-SVM. MSE, RMSE, and MAE values show the Simple SVM model is a better predictor than the Hybrid ARIM-SVM method for the A50 stock market. Figure 4.67 exhibit the actual and forecasted SVM and ARIMA-SVM models for A50.

For Deep learning methods, LSTM, hybrid ARIMA-LSTM ARIMA-LSTM, and GARCH-LSTM computed results for the A50 stock market index have been reported in table 4.20. For the A50 stock market index, LSTM estimated model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 4986, 67.651, 0.498, 70.612 respectively outperform the machine learning models. The hybrid ARIMA-LSTM computed model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 11242, 103.041, 0.755, and 106.027, respectively. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 503, 14.938, 0.111, and 22.435. Figure 4.68 exhibit the train, test, and forecasted series for A50. Test results shows Proposed GARCH-LSTM with the lowest forecasting indicators values and closest to the actual time series is the best model for the A50 stock market index.

FIGURE 4.68: A50 Deep Learning Forecasting (5 Minutes Data)

For EUS stock market Index SVM, model forecasting accuracy indicators (MSE, MAE, MAPE, and RMSE) have the values 59120, 180.407, 5.536, and 243.146 respectively. For the hybrid ARIMA-SVM forecasting model, MSE, MAE, MAPE, and RMSE have the values as 60202, 175.320, 5.444, and 245.361 respectively. Based on forecasting accuracy indicators values Simple SVM model outperform ARIMA-SVM in machine learning methods. Figure 4.69 exhibit the actual and forecasted SVM and ARIMA-SVM models for the EUS Stock market index.

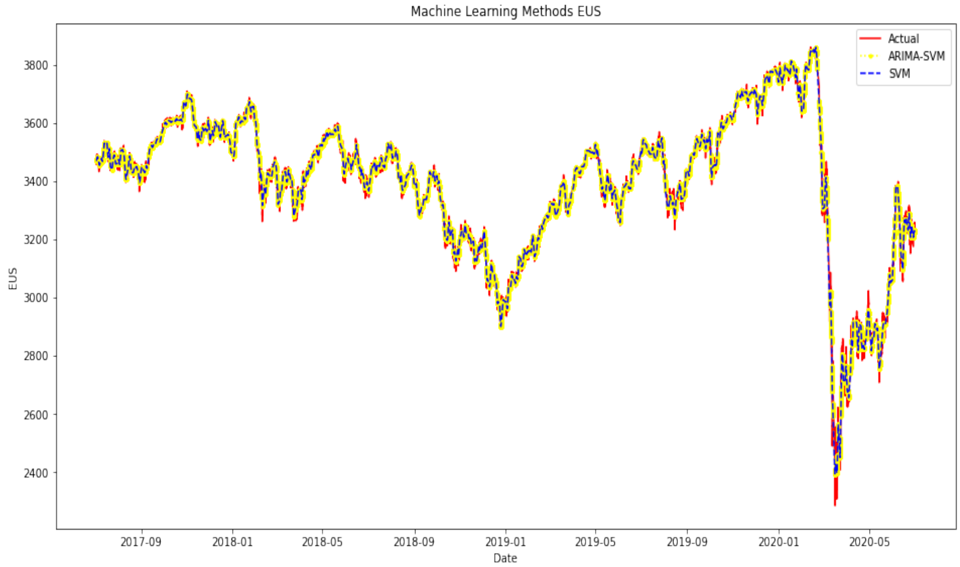


FIGURE 4.69: EUS Machine Learning Forecasting (5 Minutes Data)

FIGURE 4.70: EUS Deep Learning Forecasting (5 Minutes Data)

For Deep learning methods, LSTM, hybrid ARIMA-LSTM ARIMA-LSTM, and GARCH-LSTM computed results for the EUS stock market index have given in table 4.20. For EUS stock market index LSTM model reports MSE, MAE, MAPE, and RMSE values as 325, 11.38, 0.37, and 18.03 respectively, which outperform the conventional machine learning models and the hybrid ARIMA-SVM models. For the hybrid, ARIMA-LSTM model forecasting MSE, MAE, MAPE, and RMSE have the values as 418, 15.33, 0.45, and 20.45 respectively. Similarly, for the hybrid GARCH-LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values 1410, 20.71, 0.68, and 37.55 respectively. Figure 4.70 exhibit the train, test, and predicted series of EUS. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta color represents LSTM, ARIMA-LSTM, and GARCH-LSTM respectively. Results shows that LSTM being closest to actual the time series is the best fitted deep learning method.

For the CAC stock market, SVM model forecasting accuracy indicators have the values as 152986, 284.093, 5.545, and 391.134 respectively. For The hybrid ARIMA-SVM accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 154345, 280.748, 5.519, and 392.868 respectively for MSE, MAE, MAPE, and RMSE, which improves the results of machine learning methods with the hybrid technique. Results indicate that ARIMA-SVM performs better in machine learning regression than the simple SVM model. Figure 4.71 exhibit the actual and forecasted SVM and ARIMA-SVM models for CAC.

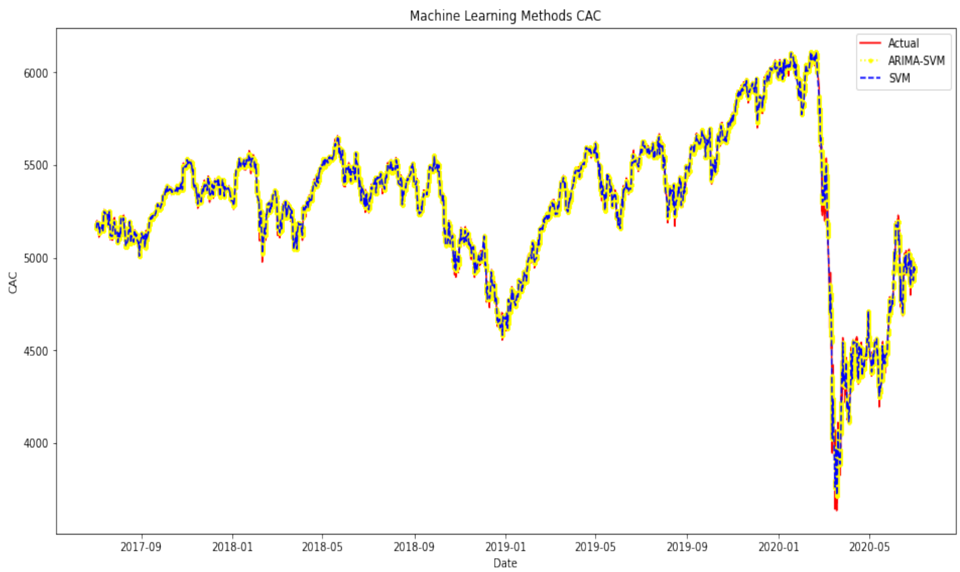


FIGURE 4.71: CAC Machine Learning Forecasting (5 Minutes Data)

The Deep learning methods LSTM, and the hybrid LSTM included ARIMA and GARCH results presented in table 4.20 for all selected stock markets. For the CAC stock market index, LSTM estimated model accuracy indicators have the MSE, MAE, MAPE, and RMSE values 38985, 85.858, 1.915, and 197.445respectively which outperform the machine learning models. For the hybrid, ARIMA-LSTM

MSE, MAE, MAPE, and RMSE are 2511, 39.413, 0.750, and 50.115. Similarly, for the hybrid GARCH-LSTM method have the values of 730, 19.505, 0.370, and 27.026 for MSE, MAE, MAPE, and RMSE respectively. Figure 4.72 exhibit the train, test, real and forecasted series of CAC. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta colore represent the LSTM, ARIMA-LSTM, and GARCH-LSTM results respectively. Results show the proposed GARCH-LSTM with the lowest forecasting accuracy indicators and close to actual time series as the best fitted deep learning method.

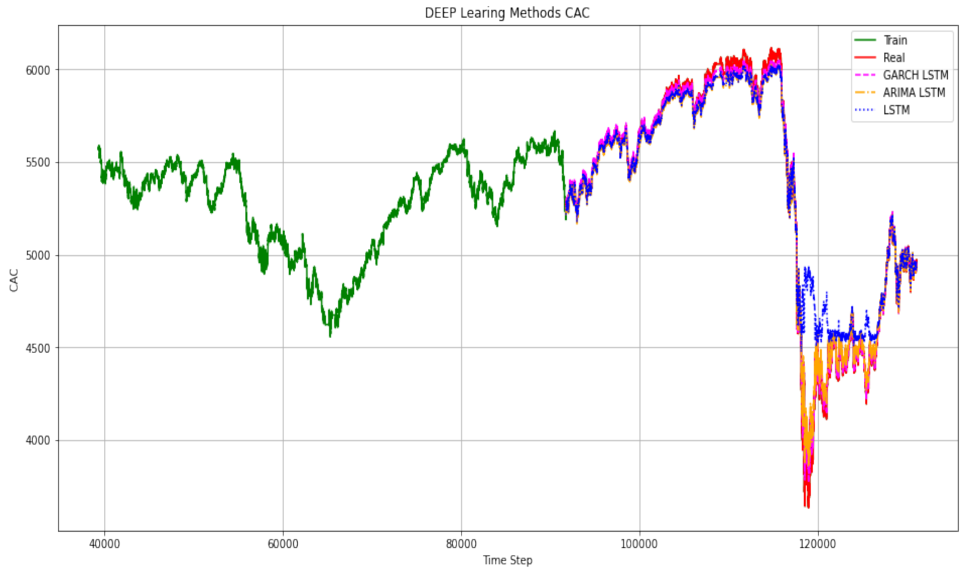


FIGURE 4.72: CAC Deep Learning Forecasting (5 Minutes Data)

For DAX stock market Index SVM, model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have values 779752, 669.041, 5.699, and 883.036 respectively. The hybrid ARIMA-SVM model, forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have the values 792739, 656.884, 5.658, and 890.359 respectively, which improve MAE, and MAPE values. Based on the results ARIMA-SVM model is a better predictor than the Simple SVM method for the DAX stock market index. Figure 4.73 exhibit the actual and forecasted SVM and ARIMA-SVM models for DAX.

For Deep learning methods, LSTM hybrid ARIMA-LSTM, and GARCH-LSTM computed results for DAX stock market index are reported in table 4.20. For the DAX stock market index, LSTM estimated model have MSE, MAE, MAPE,

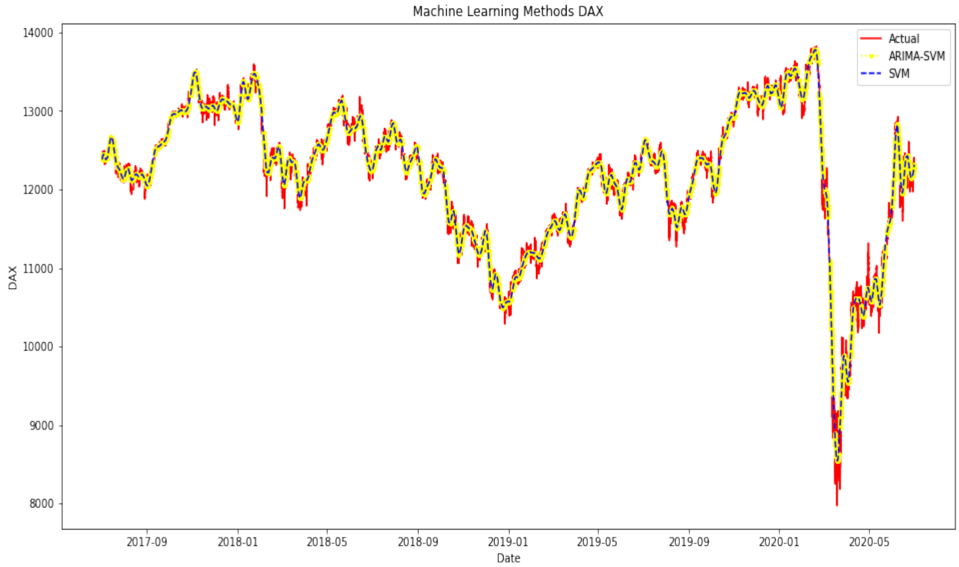


FIGURE 4.73: DAX Machine Learning Forecasting (5 Minutes Data)

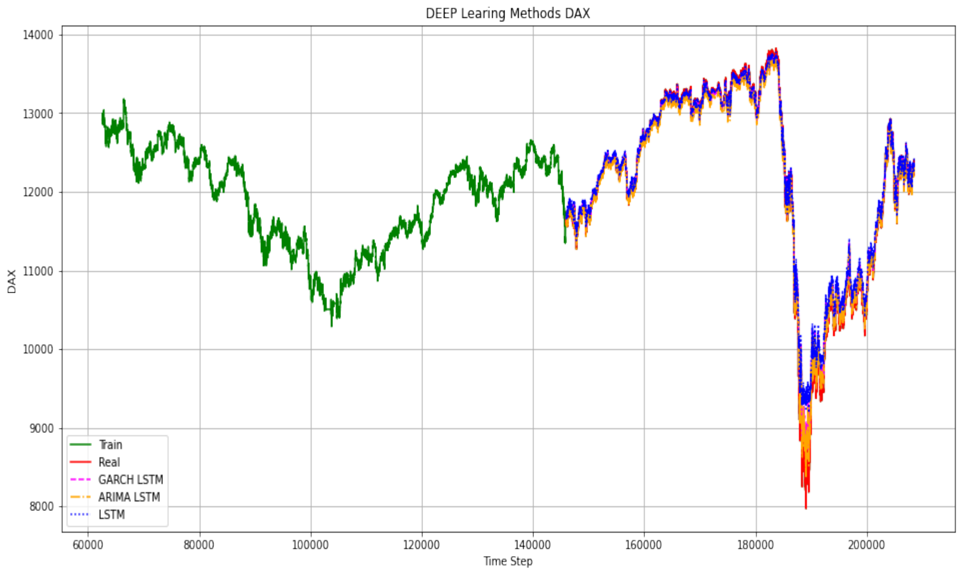


FIGURE 4.74: DAX Deep Learning Forecasting (5 Minutes Data)

and RMSE values as 26417, 87.761, 0.851, and 162.533 respectively outperform the machine learning models. The hybrid ARIMA-LSTM model MSE, MAE, MAPE, and RMSE have the values as 5535, 58.051, 0.491, and 74.396 respectively. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 7286, 54.477, 0.505, and 85.356. Figure 4.74 exhibit the train, test, actual and forecasted series for DAX. Test results shows ARIMA-LSTM with the lowest forecasting indicators values and closest to the actual time series is the best model for the DAX stock market index.

For HSI stock market Index, SVM forecasting model accuracy indicators have the values as 4375057, 1656.677, 6.046, and 2091.664 respectively for MSE, MAE, MAPE, RMSE. The hybrid ARIMA-SVM forecasting model, MSE, MAE, MAPE, and RMSE have the values as 4377786, 1653.722, 6.062, and 2092.316 respectively. Based on MSE, MAE, and RMSE forecasting indicators values, we can infer that SVM performs better than the ARIMA-SVM in machine learning regression. Figure 4.75 exhibit the actual and forecasted SVM and ARIMA-SVM models for the HSI Stock market index.

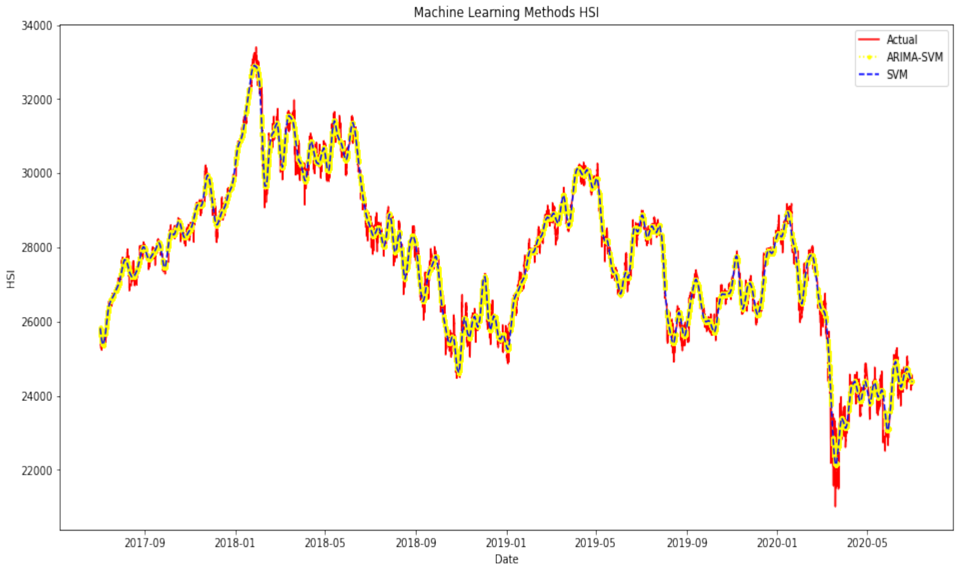


FIGURE 4.75: HSI Machine Learning Forecasting (5 Minutes Data)

For LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM deep learning methods are used for forecasting the HSI stock market index results are presented Table 4.20. For HSI stock market index, LSTM model have forecasting accuracy inidcatiors values as 29515, 146.467, 0.587, and 171.800 respectively for MSE,MAE,MAPE and RMSE. Which outperform the conventional machine learning models and the hybrid ARIMA-SVM models. For the hybrid, ARIMA-LSTM model forecasting MSE, MAE, MAPE, and RMSE have the values as 8821, 64.701, 0.262, and 93.922 respectively. Similarly, for the hybrid GARCH-LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values 15133, 115.746, 0.448, and 123.015 respectively. Figure 4.76 exhibit the train, test, and predicted series of HSI. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta color represents LSTM, ARIMA-LSTM, and GARCH-LSTM respectively. Results shows that ARIMA-LSTM being closest to actual the time series is the best fitted deep learning method.

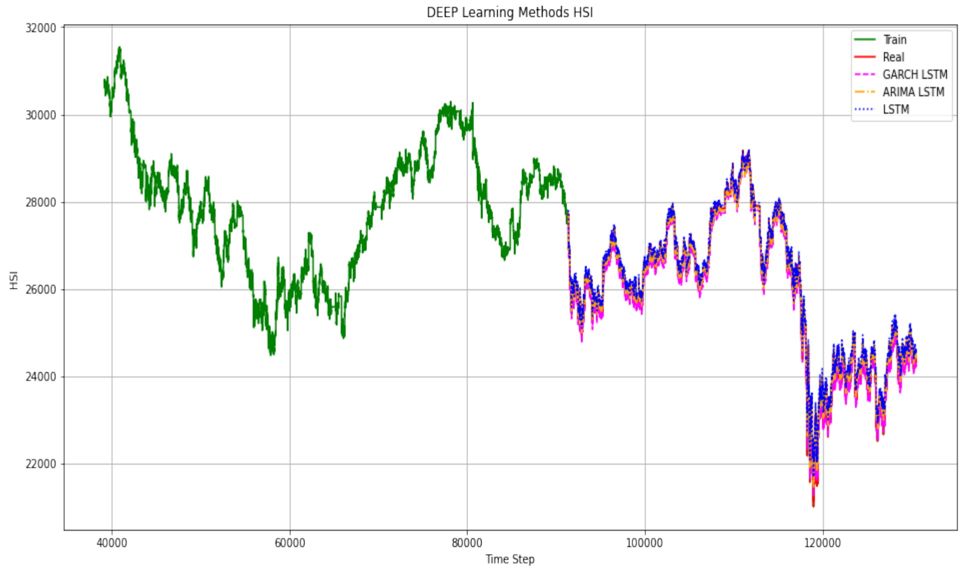


FIGURE 4.76: HSI Deep Learning Forecasting (5 Minutes Data)

Table 4.20 represents the computed model's forecasting accuracy indicators values for the NIFTY stock market index. SVM model forecasting accuracy indicators have the values as 752250, 663.473, 6.284, and 867.324 respectively for MSE, MAE, MAPE, and RMSE. For The hybrid ARIMA-SVM, accuracy indicators have the values as 752043, 663.547, 6.292, and 867.204 respectively for MSE, MAE,MAPE and RMSE. Which worsen the results of forecasting accuracy indicators for the hybrid ARIMA-SVM. SVM performs better in machine learning regression than the hybrid ARIMA- SVM model. Figure 4.77 exhibit the actual and forecasted SVM and ARIMA-SVM models for NIFTY.

For Deep learning methods, LSTM and the hybrid LSTM, including ARIMA and GARCH results are presented in table 4.20 for the NIFTY Stock Market index. For the NIFTY stock market index, LSTM model have the MSE, MAE, MAPE, and RMSE values as 392731.410, 284.084, 3.201, and 626.683 respectively, which outperform the machine learning models. For the hybrid, ARIMA-LSTM MSE, MAE, MAPE, and RMSE are 27293.708, 103.412, 1.073, and 165.208. Similarly, for the hybrid GARCH-LSTM method have the values of 16398.351, 82.718, 0.871,

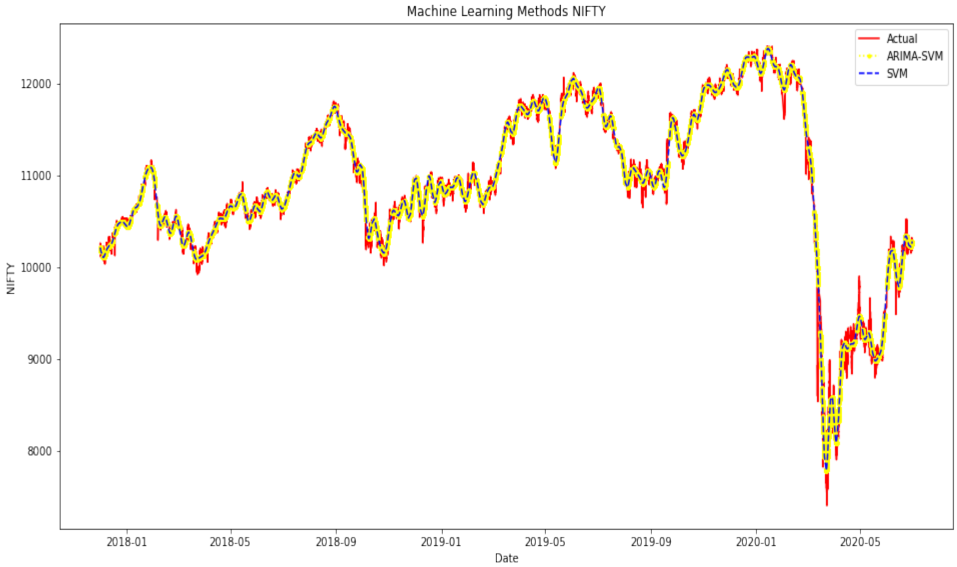


FIGURE 4.77: NIFTY Machine Learning Forecasting (5 Minutes Data)

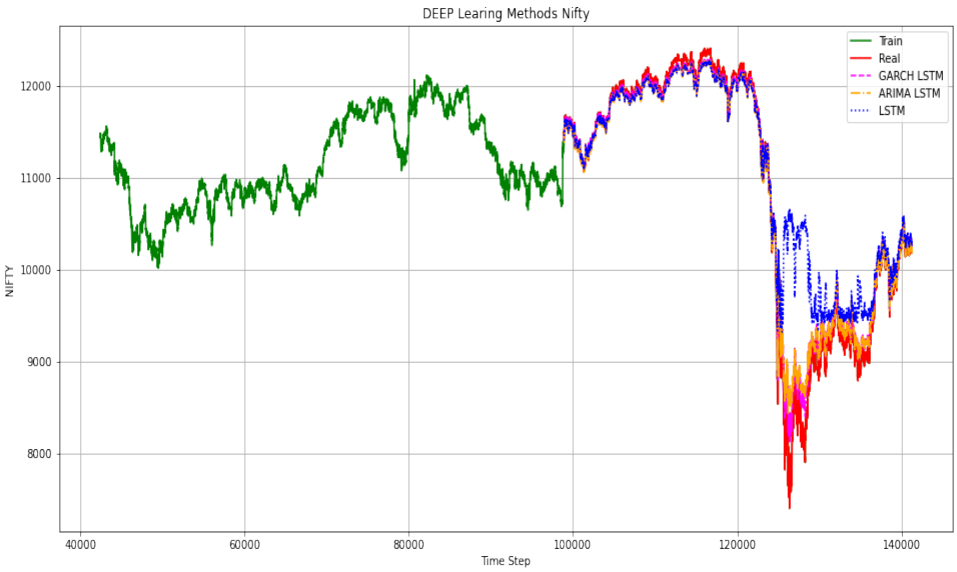


FIGURE 4.78: NIFTY Deep Learning Forecasting (5 Minutes Data)

and 128.056 for MSE, MAE, MAPE, and RMSE respectively. Figure 4.78 exhibit the train, test, real and forecasted series of NIFTY. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta colore represent the results of LSTM, ARIMA-LSTM, and GARCH-LSTM model respectively. Results show that proposed GARCH-LSTM with the lowest forecasting accuracy indicators is the best fitted deep learning method.

For NIKKEI stock market Index, SVM model have MSE, MAE, MAPE and RMSE values are 1795466, 1066.153, 5.007, and 1339.950 respectively. For The hybrid

ARIMA-SVM model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have the values 1810087, 1065.151, 5.029, and 1345.395 respectively, which worsen the results of forecasting accuracy indicators for the hybrid ARIMA-SVM. SVM performs better in machine learning regression than the hybrid ARIMA-SVM model. Figure 4.79 exhibit the actual and forecasted SVM and ARIMA-SVM models for NIKKEI.

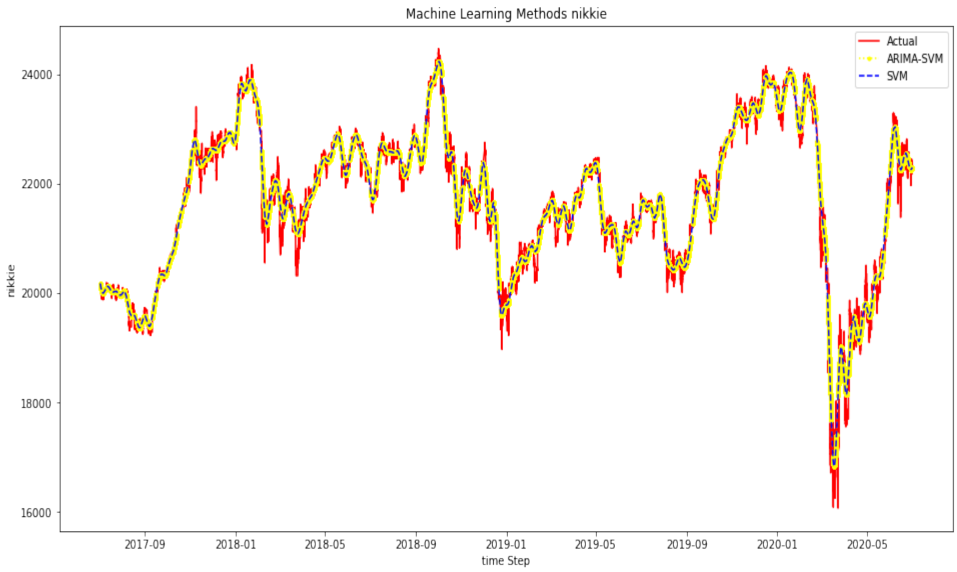


Figure 4.79: NIKKEI Machine Learning Forecasting (5 Minutes Data)

For Deep learning methods, results of LSTM, hybrid ARIMA-LSTM ARIMA-LSTM, and GARCH-LSTM method for the NIKKEI stock market index have been reported in table 4.20. For the NIKKEI stock market index, LSTM estimated model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 4542, 49.159, 0.241, and 67.391, respectively outperform the machine learning models. The hybrid ARIMA-LSTM model MSE, MAE, MAPE, and RMSE have the values as 13155, 92.452, 0.426, and 114.695, respectively. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 18120, 93.456, 0.442, and 134.612. Figure 4.80 exhibit the train, test, and entire series NIKKEI. Test results shows LSTM with the lowest forecasting indicators values and closest to the actual time series is the best model for the NIKKEI stock market index.

For the AEX stock market Index, SVM forecasting model accuracy indicators have the values as 1137, 25.134, 4.650, and 33.715 respectively for MSE, MAE, MAPE,

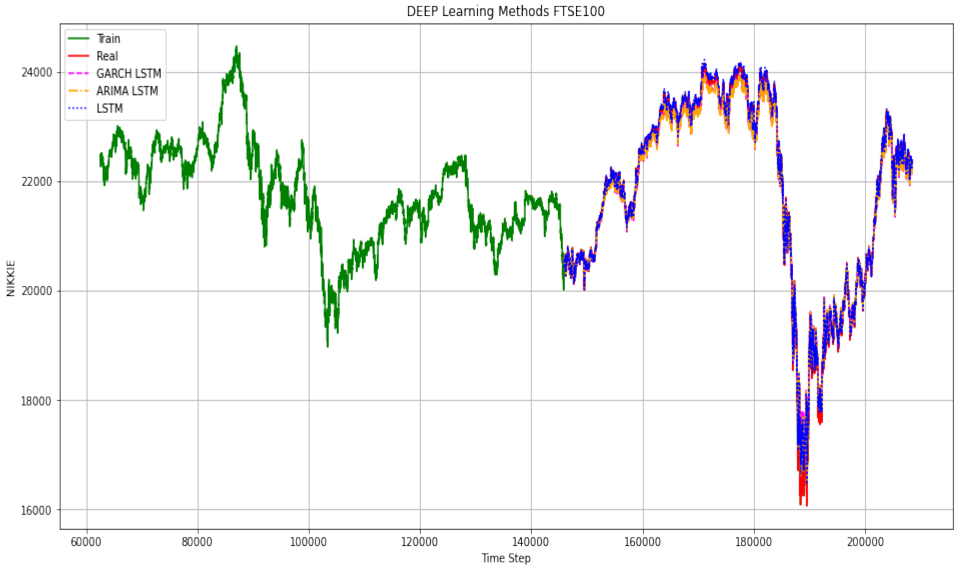


FIGURE 4.80: NIKKEI Deep Learning Forecasting (5 Minutes Data)

RMSE. For the hybrid ARIMA-SVM forecasting model, MSE, MAE, MAPE, and RMSE have the values as 1133, 24.965, 4.634, and 33.667 respectively. Based on literature improved MAE and MAPE forecasting indicators values, we can conclude that ARIMA-SVM performs better than the SVM in machine learning regression. Figure 4.81 exhibit the actual and forecasted SVM and ARIMA-SVM models for the AEX Stock market index. Where yellow color shows the Forecasted values of ARIMA-SVM, Red color shows actual and blue color shows the forecasted values of SVM model.

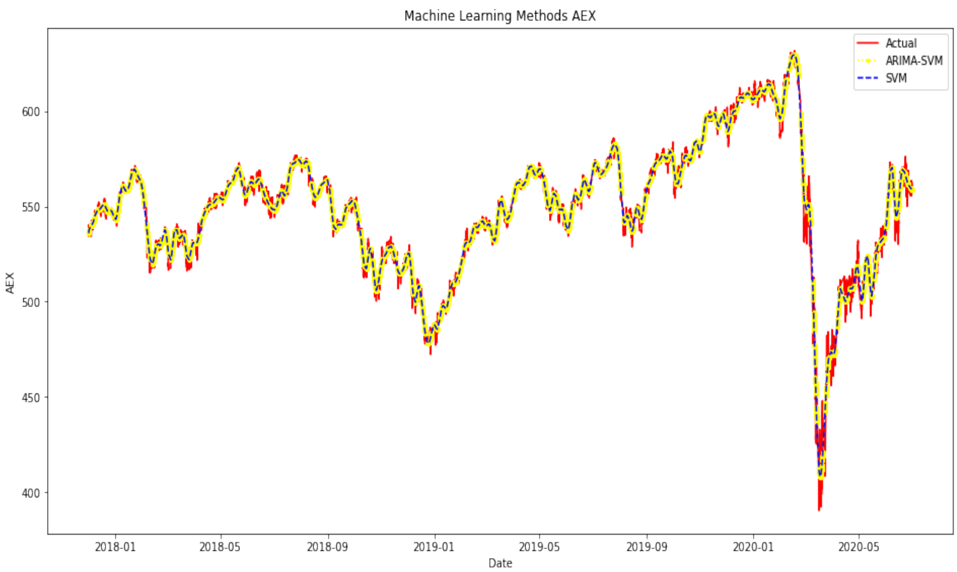


FIGURE 4.81: AEX Machine Learning Forecasting (5 Minutes Data)

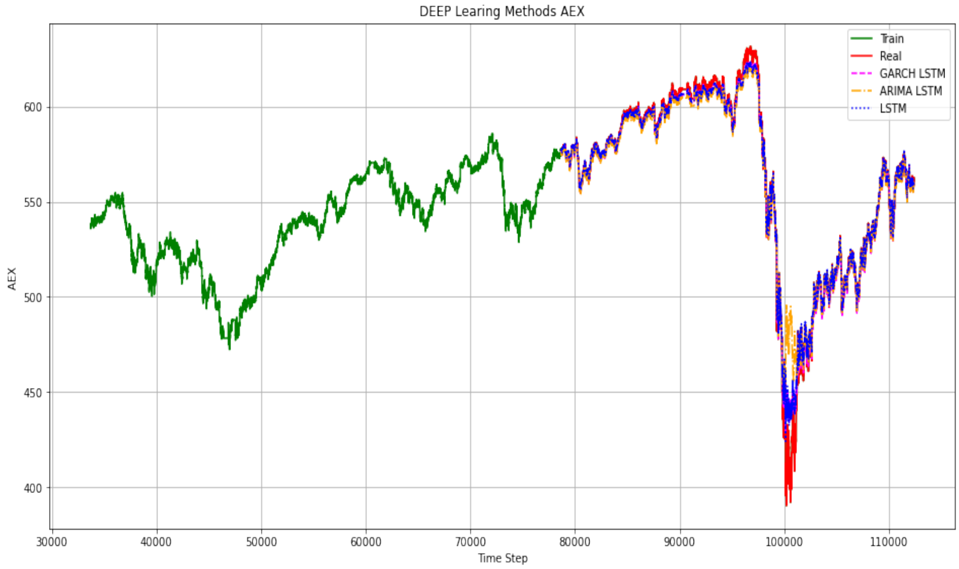


FIGURE 4.82: AEX Deep Learning Forecasting (5 Minutes Data)

For Deep learning methods, LSTM, hybrid ARIMA-LSTM ARIMA-LSTM, and GARCH-LSTM computed results for the AEX stock market index have given in table 4.20. For the AEX stock market index LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 30, 2.470, 0.493, and 5.505 respectively, which outperform the conventional machine learning models and the hybrid ARIMA-SVM models. For the hybrid, ARIMA-LSTM model forecasting MSE, MAE, MAPE, and RMSE have the values as 143, 4.382, 0.890, and 11.949 respectively. Similarly, for the hybrid GARCH-LSTM model accuracy indicators have the values as 31, 2.628, 0.516, and 5.560 respectively for MSE, MAE, MAPE, and RMSE . Figure 4.82 exhibit the train, test, and predicted series of AEX. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta color represents LSTM, ARIMA-LSTM, and GARCH-LSTM respectively. Results shows that LSTM being closest to actual the time series is the best fitted deep learning method.

For the WIG stock market, SVM model forecasting accuracy indicators i.e., MSE, MAE, MAPE, and RMSE have the values 57236, 165.687, 8.479, and 239.240 respectively. For The hybrid ARIMA-SVM accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 59621, 157.355, 8.296, and 244.174 respectively,

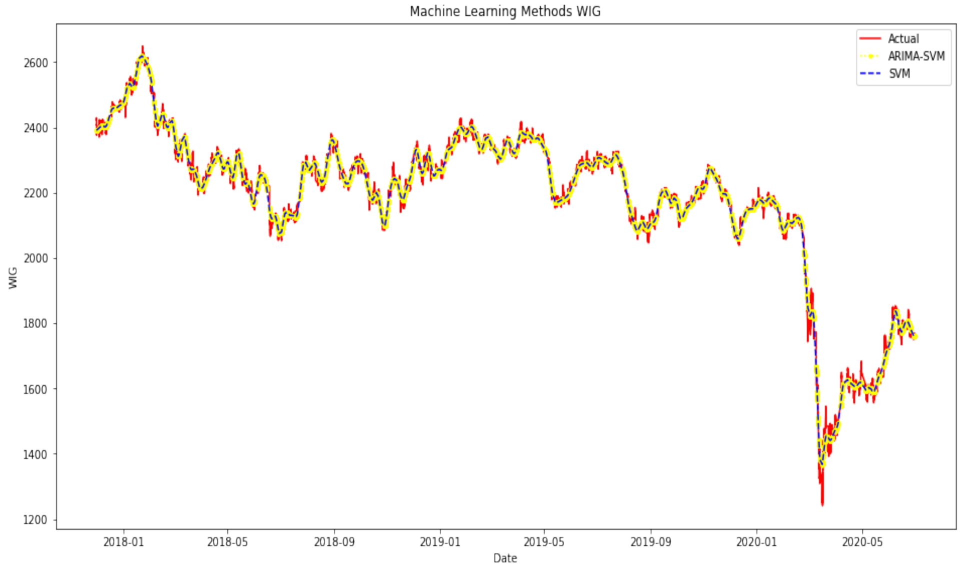


FIGURE 4.83: WIG Machine Learning Forecasting (5 Minutes Data)

which improves the results of machine learning methods with the hybrid technique. Results indicate that ARIMA-SVM performs better in machine learning regression than the simple SVM model. Figure 4.83 exhibit the actual and forecasted values of WIG stock market index, where SVM and ARIMA-SVM models are used.

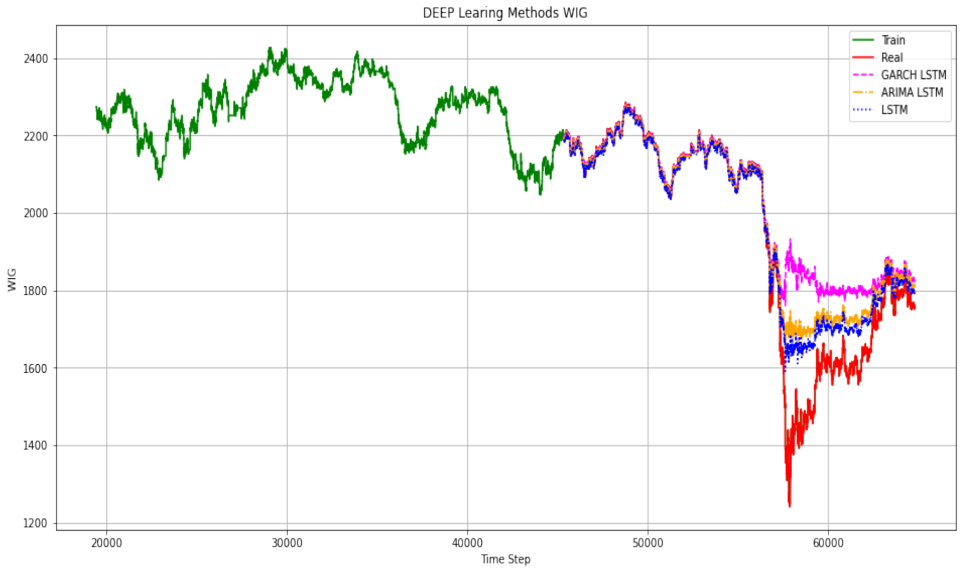


FIGURE 4.84: WIG Deep Learning Forecasting (5 Minutes Data)

The results of Deep learning methods i.e., LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM are presented in Table 4.20 for all selected stock markets. For the

WIG stock market index, LSTM estimated model accuracy indicators have the MSE, MAE, MAPE, and RMSE values 6086, 44.635, 2.830, and 78.015 respectively, which outperform the machine learning models. For the hybrid, ARIMA-LSTM MSE, MAE, MAPE, and RMSE are 9001, 51.982, 3.347, and 94.875. Similarly, for the hybrid GARCH-LSTM method have the values of 22132, 78.418, 5.099, and 148.770 for MSE, MAE, MAPE, and RMSE respectively. Figure 4.84 exhibit the train, test, real and forecasted series of WIG. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta colore represent the results of LSTM, ARIMA-LSTM, and GARCH-LSTM model respectively. Results show the LSTM with the lowest forecasting accuracy indicators and close to actual time series as the best fitted deep learning method.

For the SSGF stock market Index, SVM model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have values 920, 21.485, 6.283, and 30.339 respectively. For The hybrid ARIMA-SVM model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have the values 920, 21.405, 6.268, and 30.340 respectively, which improve MAE, and MAPE values. Based on the results, and ARIMA-SVM model is a better predictor than the Simple SVM method for the SSGF stock market index. Figure 4.85 exhibit the actual and forecasted SVM and ARIMA-SVM models for SSGF.

For Deep learning methods, LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM estimated results for SSGF stock market index have been reported in Table 4.20. For the SSGF stock market index, LSTM estimated model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 343, 10.043, 3.499, and 18.514, respectively outperform the machine learning models. The hybrid ARIMA-LSTM computed model evaluation criteria have values as33, 3.790, 1.280, and 5.702 respectively for MSE, MAE, MAPE, and RMSE. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 211, 7.967, 2.754, and 14.522. Figure 4.86 exhibit the train, test, actual and forecasted series for SSGF. Test results shows that ARIMA-LSTM with the lowest forecasting indicators values, is the best model for SSGF stock market index.

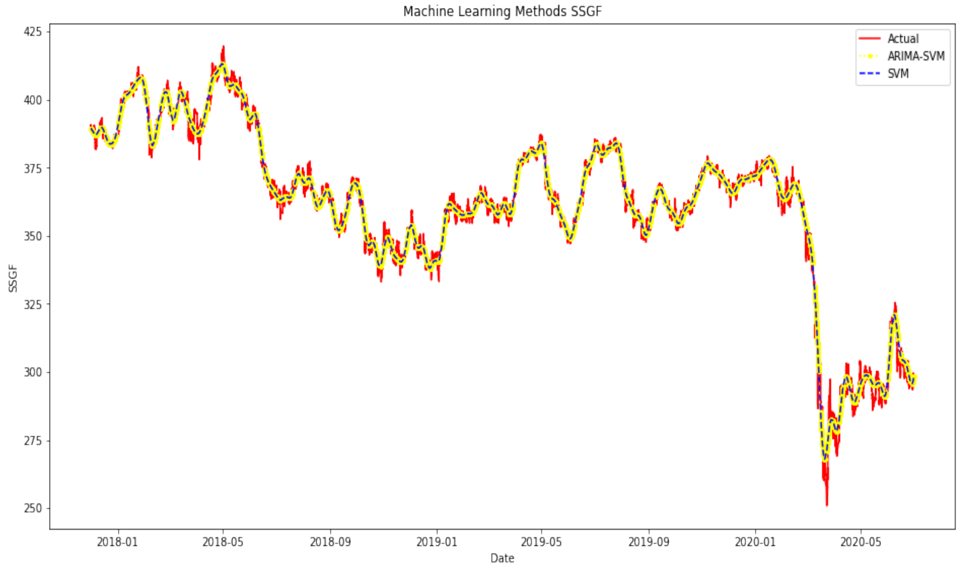


FIGURE 4.85: SSGF Machine Learning Forecasting (5 Minutes Data)

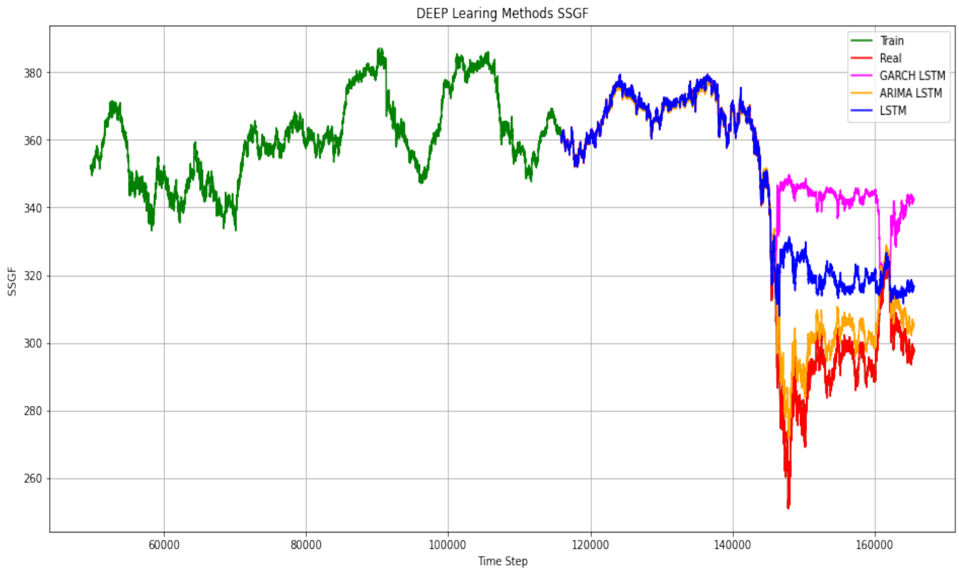


FIGURE 4.86: SSGF Deep Learning Forecasting (5 Minutes Data)

For the IBX stock market Index, SVM forecasting model accuracy indicators have the values as 875122, 656.346, 7.676, and 935.480 respectively for MSE, MAE, MAPE, RMSE. For the hybrid ARIMA-SVM forecasting model, MSE, MAE, MAPE, and RMSE have the values as 885737, 643.051, 7.651, and 941.136 respectively. Based on MAE, and MAPE forecasting indicators values, we can infer that ARIMA-SVM performs better than SVM in machine learning regression. Figure 4.87 exhibit the actual and forecasted SVM and ARIMA-SVM models for the IBX Stock market index.

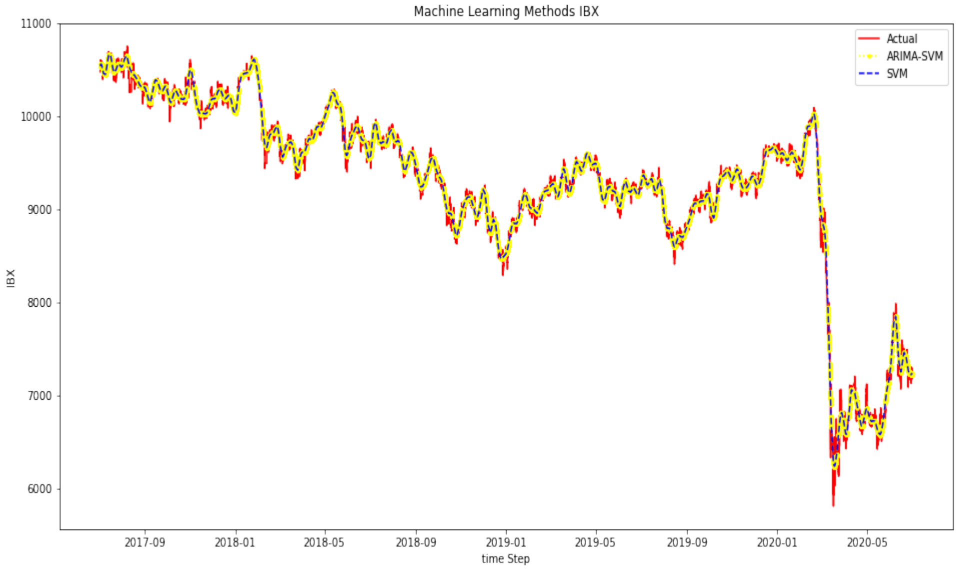


FIGURE 4.87: IBX Machine Learning Forecasting (5 Minutes Data)

For Deep learning methods, LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM are used and results for IBX stock market index are represented in Table 4.20. For IBX stock market index, LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 43832, 129.624, 1.844, and 209.360 respectively, which outperform the conventional machine learning models and the hybrid ARIMA-SVM models. For the hybrid, ARIMA-LSTM model forecasting MSE, MAE, MAPE, and RMSE have the values as 20028, 95.735, 1.326, and 141.520 respectively. Similarly, for the hybrid GARCH-LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values 6486, 55.332, 0.735, and 80.536 respectively. Figure 4.88 exhibit the train, test, and forecasted series of IBX. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta color represents LSTM, ARIMA-LSTM, and GARCH-LSTM respectively. Results shows that Proposed GARCH-LSTM is closest to actual the time series so it is the best fitted deep learning method.

For SMI stock market Index SVM, model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have values 385897, 509.899, 5.370, and 621.206 respectively. For the hybrid ARIMA-SVM model forecasting accuracy indicators, MSE, MAE, MAPE, and RMSE have values as 401312, 502.572, 5.224, and 633.492

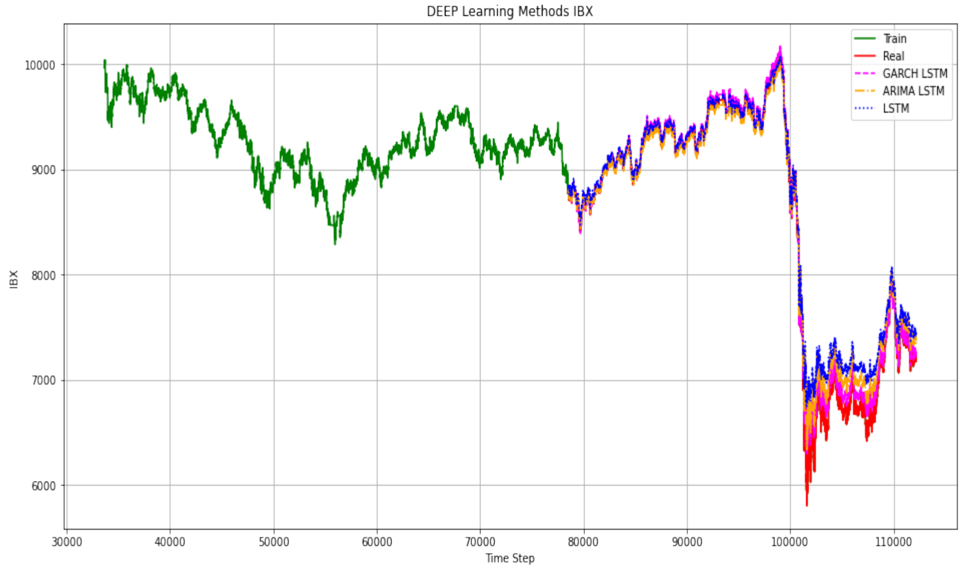


FIGURE 4.88: IBX Deep Learning Forecasting (5 Minutes Data)

respectively improve MAE, and MAPE values. Based on the results, and ARIMA-SVM model is a better predictor than the Simple SVM method for the SMI stock market index. Figure 4.89 exhibit the actual and forecasted values of SVM and ARIMA-SVM models for SMI.

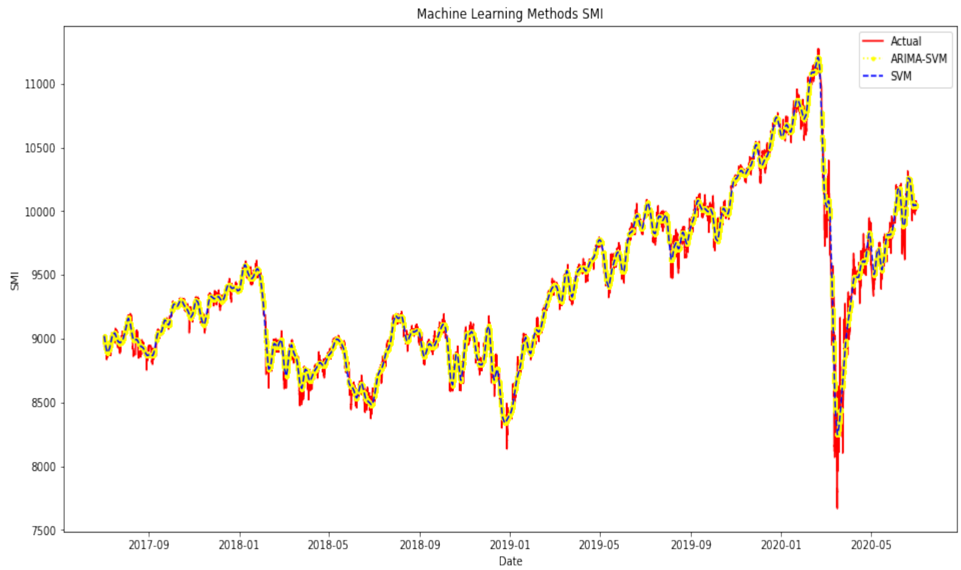


FIGURE 4.89: SMI Machine Learning Forecasting (5 Minutes Data)

For Deep learning methods, LSTM, hybrid ARIMA-LSTM ARIMA-LSTM, and GARCH-LSTM are used and results for the SMI stock market index are reported in Table 4.20. For the SMI stock market index, LSTM estimated model evaluation

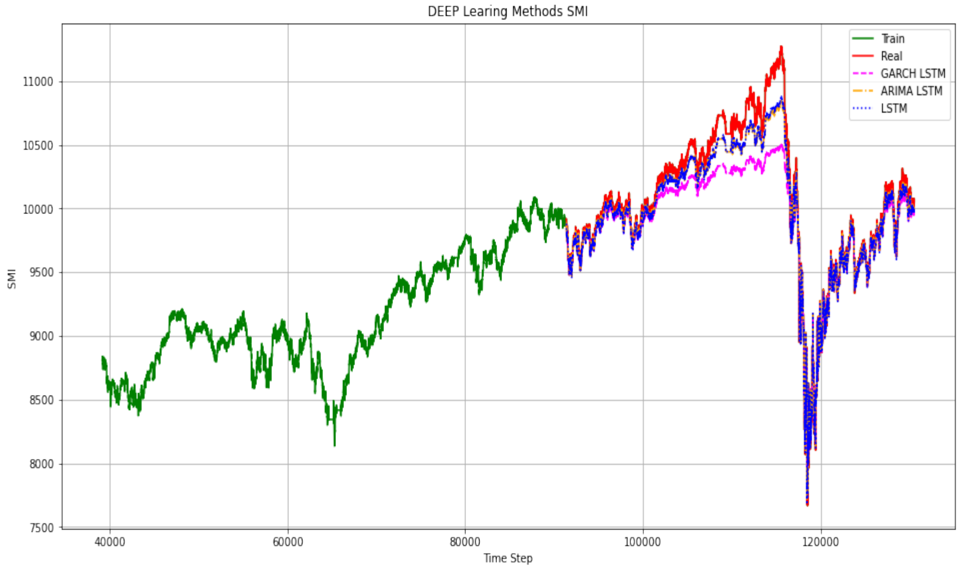


FIGURE 4.90: SMI Deep Learning Forecasting (5 Minutes Data)

criteria (MSE, MAE, MAPE, and RMSE) have values 12939, 82.926, 0.794, and 113.749 respectively. Which outperform the machine learning models. The hybrid ARIMA-LSTM model MSE, MAE, MAPE, and RMSE have the values as 13722, 79.595, 0.758, and 117.141 respectively. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 55023, 155.496, 1.472, and 234.570. Figure 4.90 exhibit the train, test, actual and forecasted series for SMI. Test results shows that the ARIMA-LSTM with the lowest forecasting indicators values, is the best model for the SMI stock market index.

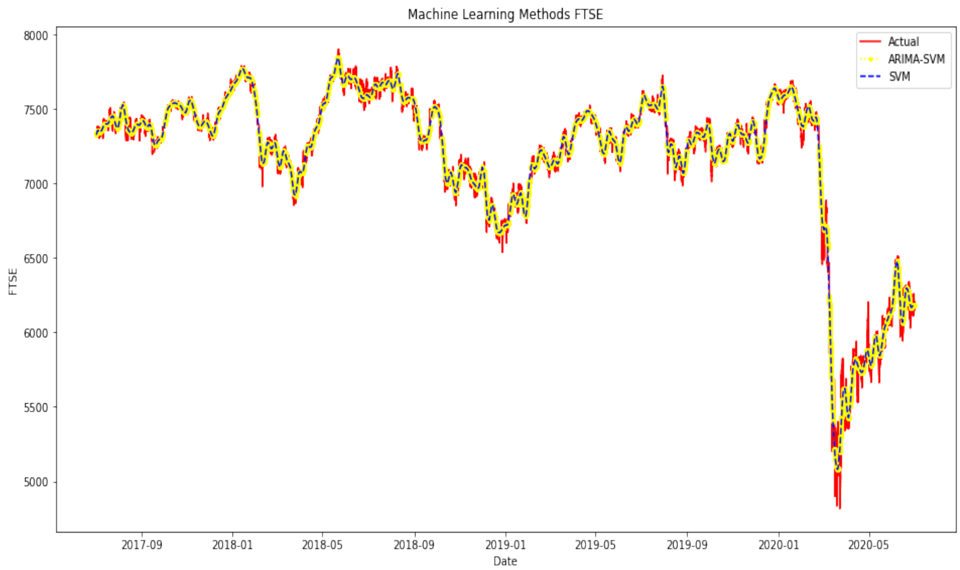


FIGURE 4.91: FTSE Machine Learning Forecasting (5 Minutes Data)

For FTSE stock market Index, SVM model forecasting accuracy indicators (MSE, MAE, MAPE, and RMSE) have the values 266505, 361.385, 5.388, and 516.242 respectively. For the hybrid ARIMA-SVM forecasting model, MSE, MAE, MAPE, and RMSE have the values as 155080, 157.826, 2.609, and 393.802 respectively. Based on forecasting accuracy indicators values, we can infer that ARIMA-SVM performs better than the SVM in machine learning regression. Figure 4.91 exhibit the actual and forecasted values of SVM and ARIMA-SVM models for the FTSE Stock market index.

For Deep learning methods, LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM are used and computed results for the FTSE stock market index have given in table 4.20. For FTSE stock market index LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 33139, 75.878, 1.362, and 182.042 respectively, which outperform the conventional machine learning models and the hybrid ARIMA-SVM models. For the hybrid, ARIMA-LSTM model forecasting MSE, MAE, MAPE, and RMSE have the values as 223010, 211.335, 3.790, and 472.239 respectively. Similarly, for the hybrid GARCH-LSTM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values 77822, 130.281, 2.306, and 278.967 respectively. Figure 4.92 exhibit the train, test, and predicted series of FTSE. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta color represents LSTM, ARIMA-LSTM, and GARCH-LSTM respectively. Results shows that LSTM being closest to actual the time series is the best fitted deep learning method.

For DJI30 stock market Index SVM, model have forecasting accuracy indicators MSE, MAE, MAPE, and RMSE as 3550469, 1468.319, 5.928, and 1884.269 respectively. The hybrid ARIMA-SVM model forecasting accuracy indicators MSE, MAE, MAPE, and RMSE have the values 2784442, 1205.459, 4.901, and 1668.665 respectively, which improve the results indicates simple ARIMA-SVM. MSE, RMSE, and MAE values show that the ARIMA-SVM model is a better predictor than the SVM for the DJI30 stock market. Figure 4.93 exhibit the actual and forecasted SVM and ARIMA-SVM models for DJI30.

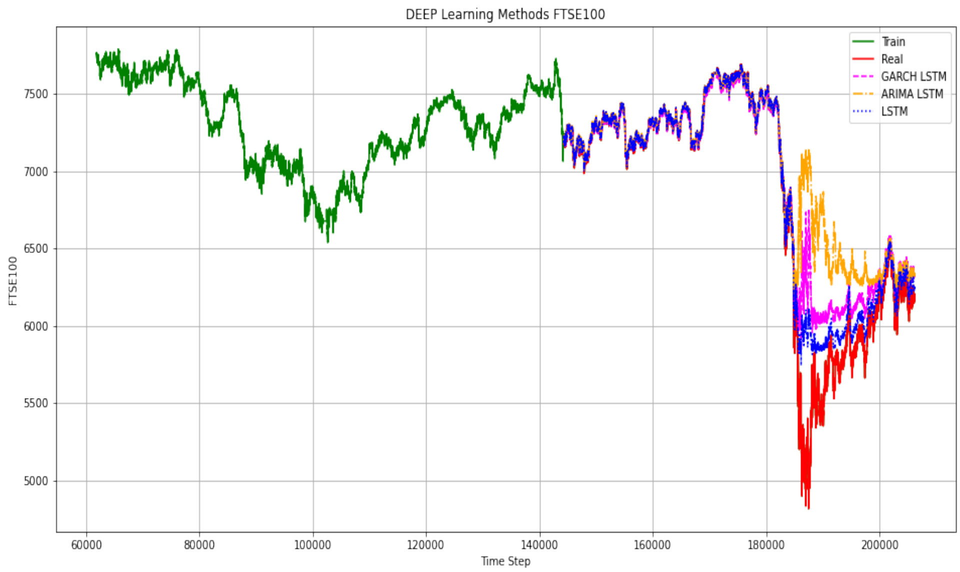


FIGURE 4.92: FTSE Deep Learning Forecasting (5 Minutes Data)

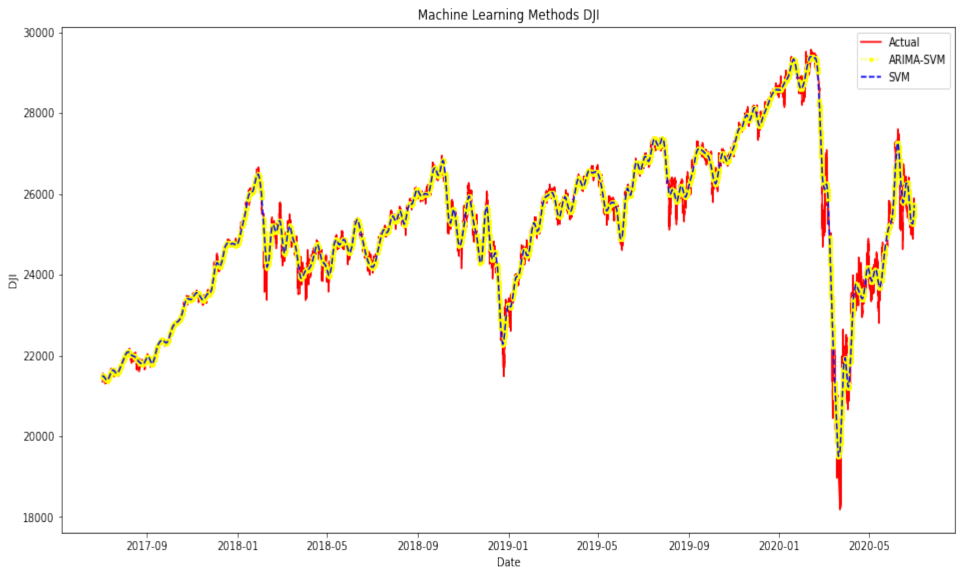


FIGURE 4.93: DJI Machine Learning Forecasting (5 Minutes Data)

For Deep learning methods, LSTM, hybrid ARIMA-LSTM, and GARCH-LSTM models are used and results for DJI30 stock market index have been reported in table 4.20. For the DJI30 stock market index, LSTM estimated model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 40228, 135.770, 0.496, and 200.569 respectively. Which outperform the machine learning models. The hybrid ARIMA-LSTM computed model evaluation criteria (MSE, MAE, MAPE, and RMSE) have 9862, 76.491, 0.291, and 99.307 respectively. Similarly, the hybrid GARCH-LSTM model have MSE, MAE, MAPE, and RMSE values as 25215,

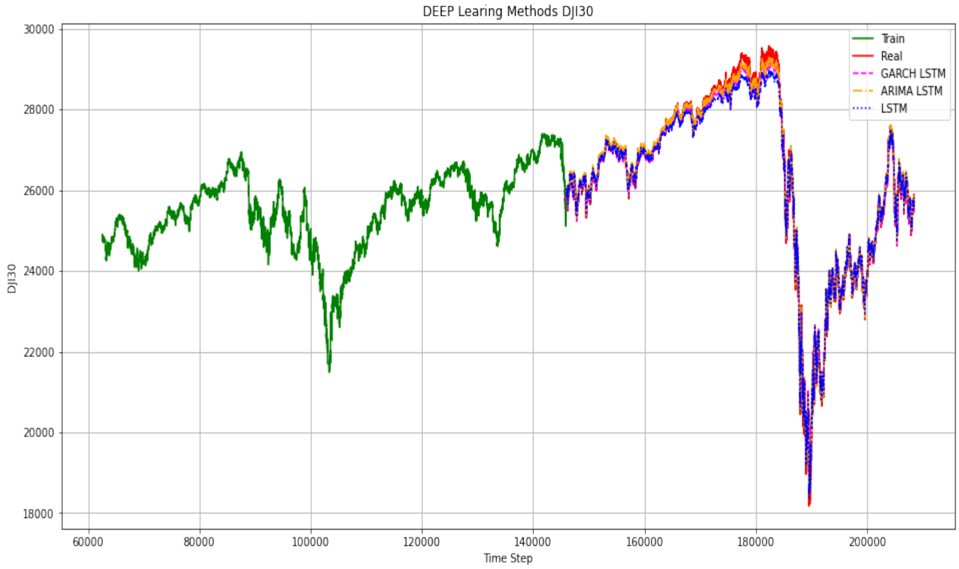


FIGURE 4.94: DJI Deep Learning Forecasting (5 Minutes Data)

95.113, 0.367, and 158.792. Figure 4.94 exhibit the train, test, and predicted series for DJI30. Test results shows that ARIMA-LSTM has the lowest forecasting indicators values and closest to actual time series, so it is the best model for DJI30 stock market index.

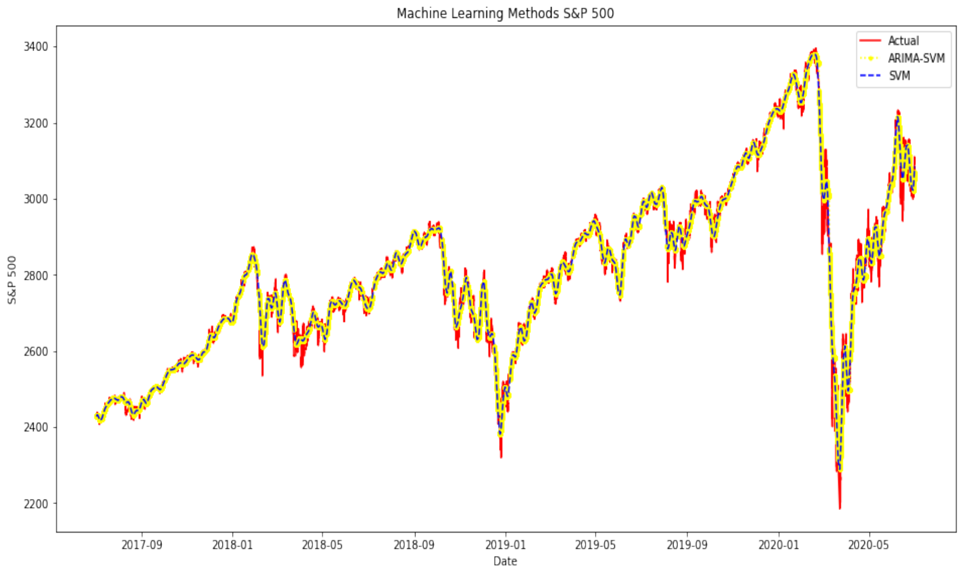


FIGURE 4.95: SNP Machine Learning Forecasting (5 Minutes Data)

For the SNP500 stock market, SVM model forecasting accuracy indicators have the values as 48606, 175.543, 6.274, 220.467 respectively for MSE, MAE, MAPE, and RMSE. For The hybrid ARIMA-SVM model accuracy indicators MSE, MAE, MAPE, and RMSE have the values of 3901, 27.995, 0.957, and 62.456 respectively, which improves the results of machine learning methods. Results indicate that ARIMA-SVM performs better in machine learning regression than the simple SVM model. Figure 4.95 exhibit the actual and forecasted SVM and ARIMA-SVM models for SNP500.

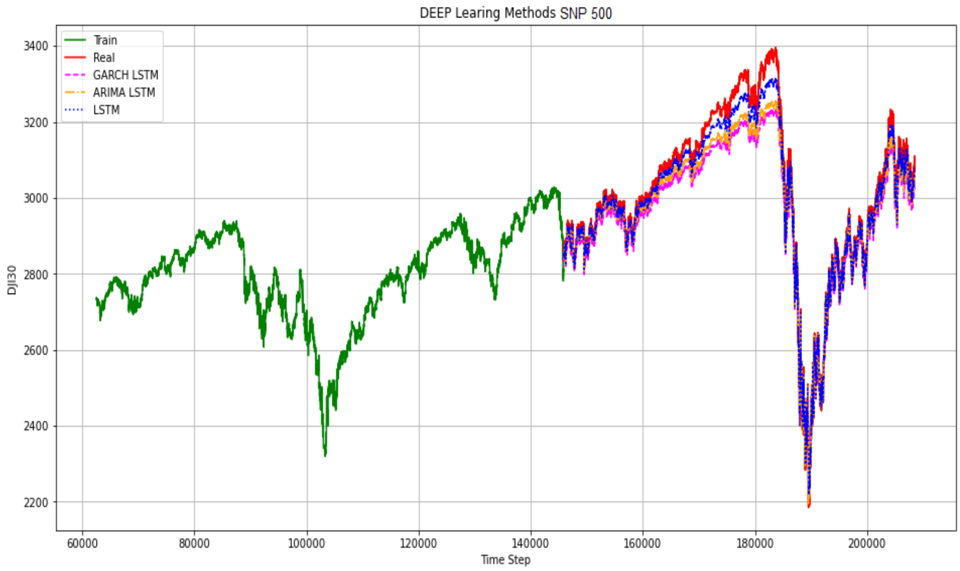


FIGURE 4.96: SNP Deep Learning Forecasting (5 Minutes Data)

The The deep learning methods LSTM, and the hybrid LSTM along with ARIMA and GARCH results are presented in Table 4.20 for all selected stock markets. For the SNP500 stock market index, LSTM estimated model accuracy indicators have the MSE, MAE, MAPE, and RMSE values 801, 22.435, 0.718, and 28.309 respectively, which outperform the machine learning models. For the hybrid, the ARIMA-LSTM method computed forecasting accuracy indicators have the MSE, MAE, MAPE, and RMSE values of 989, 22.281, 0.709, and 31.441. Similarly, for the hybrid GARCH-LSTM method have the values of 4155, 52.410, 1.671, and 64.461 for MSE, MAE, MAPE, and RMSE respectively.

Figure 4.96 exhibit the train, test, real and forecasted series of SNP500. Where green and red-colored series represent the trained and actual test time series, purple, orange, and magenta colore represent the LSTM, ARIMA-LSTM, and GARCH-LSTM results respectively. Results show that ARIMA-LSTM has lowest forecasting accuracy indicators and is close to actual time series so it is the best fitted deep learning method.

This section concludes the estimation and forecasting of machine learning and deep learning methods. The study use SVM, ARIMA-SVM, LSTM, ARIMA-LSTM, and GARCH-LSTM model to forecast the stock market indices. Results show a mixed result between the deep learning battle. GARCH-LSTM and ARIMA-LSTM show better forecasting results for 5 minutes data frequency.

4.3.5 Discussion

This section discuss the results of classical, machine learning, and deep learning method where 5 minutes data is used for analysis. The first objective is to find out best performing classical forecasting model forecasting. Classical model estimation showed in table 4.17 that past prices has the significant impact on current prices. Which support the EMH stance of past prices behavior can transmit any profitable information in future price moment (Strozzi and Zaldivar, 2005). Results show the randomness in markets decreased dramatically with the increase of data frequency. For the generalization of ARFIMA model coefficient value specified by Baillie (1996), none of our sample market achieved. Below generalized d parameter value a minimal impact of fractional differencing on current prices of markets and consistent with studies of (Floros et al., 2007; Reisen et al., 2001).

Table 4.19 and Table 4.20, compare model on basis of forecasting accuracy indicators of modified deep learning, machine learning and classical forecasting method proposed by various aforementioned studies (Alam et al., 2020; Bhardwaj and Swanson, 2006; Kewat et al., 2017; Li et al., 2020; Papacharalampous et al., 2019).GARCH based deep learning method also used to analyze the behavior stock market indices for 5 minutes frequency. Above stated Tables answer the questions ”Are classical models successful in forecasting price behavior in financial markets? How do machine leaning based models perform in forecasting price trends in financial markets? and Do machine learning based models outperform classical models of forecasting?" Below table 4.21 is design to present the best fitted model of forecasting for 5-minute data interval.

Results shows that for the stock market indices A50, CAC, and IBX, GARCH-LSTM outperform the other deep learning, machine learning and classical forecasting methods suggested by the (Li et al., 2020). For ASX, DAX, HIS, SSGF, SMI, DJI30, and SNP500 stock markets indices ARIMA-LSTM consistent with (Alam et al., 2020; Kulshreshtha, 2020; Li et al., 2020) statement of improving prediction power with hybrid ARIMA and LSTM. Ironically EUS, NIFTY, NIKKEI, AEX, WIG, and FTSE stock market indices results shows LSTM as batter method to predict stock market prices then the improved ARIMA-LSTM and GARCH-LSTM methods. GARCH-LSTM appears second best method for the stock market indices ASX, DAX, NIFTY, AEX, SSGF, FTSE, and DJI30 which might be perform batter with the improvement and changes in activation functions, batch and evaluation intervals number. Whereas, EUS, CAC, NIFTY, NIKKEI, WIG, and IBX shows ARIMA as second best fitted forecasting method. Eight out of 16 markets show the LSTM as 3rd best fitted model to forecast financial time series. In classical forecasting techniques we witnessed that long memory method ARFIMA as best fitted candidate for the classical forecasting techniques for high frequency data.

Again, results shows Deep learning-based methods outperform the conventional machine learning and classical forecasting technique which are consistent with the of finds (Choi, 2018; Fang and Yuan, 2019; Fischer and Krauss, 2018; Kulshreshtha, 2020; Li et al., 2020; Temür et al., 2019; Yan and Ouyang, 2018). Aforementioned forecasting indicators MSE, RMSE, MAE and MAPE values in section 4.3.3.2 and section 4.3.4.2 clearly provide that hybrid/deep learning-based methods performs better than other methods that is align with our research objectives and accept our hypothesis H3 of deep learning outperform conventional machine learning and classical forecasting methods.

Table 4.21: Models Comparison (5 Minutes Data)

	Model Rank							
	1st RANK	**2nd RANK**	**3rd RANK**	**4th RANK**	**5th RANK**	**6th RANK**	**7th RANK**	**8th RANK**
ASX	ARIMA-LSTM	GARCH-LSTM	LSTM	SVM	ARIMA-SVM	ARFIMA	ARIMA	GARCH
A50	GARCH-LSTM	LSTM	ARIMA-LSTM	SVM	ARIMA-SVM	ARIMA	ARFIMA	GARCH
EUS	ARIMA-LSTM	GARCH-LSTM	LSTM	SVM	ARIMA-SVM	ARFIMA	ARIMA	GARCH
CAC	GARCH-LSTM	ARIMA-LSTM	LSTM	ARIMA-SVM	SVM	ARFIMA	ARIMA	GARCH
DAX	ARIMA-LSTM	GARCH-LSTM	LSTM	ARIMA-SVM	SVM	ARFIMA	ARIMA	GARCH
HSI	ARIMA-LSTM	GARCH-LSTM	LSTM	SVM	ARIMA-SVM	ARIMA	ARFIMA	GARCH
NIFTY	LSTM	ARIMA-LSTM	GARCH-LSTM	SVM	ARIMA-SVM	ARFIMA	ARIMA	GARCH
NIKKEI	LSTM	ARIMA-LSTM	GARCH-LSTM	SVM	ARIMA-SVM	ARFIMA	ARIMA	GARCH
AEX	LSTM	GARCH-LSTM	ARIMA-LSTM	ARIMA-SVM	SVM	ARFIMA	ARIMA	GARCH
WIG	LSTM	ARIMA-LSTM	GARCH-LSTM	SVM	ARIMA-SVM	GARCH	ARIMA	ARFIMA
SSGF	ARIMA-LSTM	LSTM	GARCH-LSTM	ARIMA-SVM	SVM	ARFIMA	ARIMA	GARCH
IBX	GARCH-LSTM	ARIMA-LSTM	LSTM	ARIMA-SVM	SVM	ARIMA	ARFIMA	GARCH
SMI	ARIMA-LSTM	LSTM	GARCH-LSTM	ARIMA-SVM	SVM	ARIMA	ARFIMA	GARCH
FTSE	LSTM	GARCH-LSTM	ARIMA-LSTM	ARIMA-SVM	SVM	ARFIMA	ARIMA	GARCH
DJI	ARIMA-LSTM	GARCH-LSTM	LSTM	ARIMA-SVM	SVM	ARFIMA	ARIMA	GARCH
SNP	ARIMA-LSTM	LSTM	GARCH-LSTM	ARIMA-SVM	SVM	ARFIMA	ARIMA	GARCH

4.4 Section IV (Data Frequency Comparison)

The last objective of the study is to find out the best data frequency for forecasting purpose. Previously many of the studies has been conducted on to evaluate the different forecasting methods i.e., (Alam et al., 2020; Fang and Yuan, 2019; Khairalla and AL-Jallad, 2017; Li et al., 2020; Sezer et al., 2020; Wang et al., 2017; Yan and Ouyang, 2018). These studies used different time intervals of data to evaluate classical, machine learning and hybrid deep learning techniques. Studies like **?** used daily data to evaluate deep learning techniques on DAX, HSI and SNP stock market indices. Shintate and Pichl (2019) used 30 minutes time interval data to evaluate ML forecasting techniques on DAX and SNP500. Other studies like Cadenas et al. (2016); Ersan et al. (2019); Moews et al. (2019) used hourly data to evaluate different forecasting techniques. Similarly Li et al. (2020) used 5 minutes data frequencies to explore the improved ARIMA-LSTM model of forecasting. Chong et al. (2017) also used 5 minutes data interval for forecasting. Studies like Arévalo et al. (2016); Long et al. (2019) used 1 minute data to evaluate respective techniques.

Financial literature is still missing to find out appropriate frequency that must be used to forecast the financial time series. To answer this, study compare three different time frequencies in 16 different stock markets. The study examines the hourly data frequency, 10-minute data frequency, and 5-minute data frequency (See Table 4.22). The results shows that 56% of sample markets are best forecasted by using 10 minutes of data frequency (Figure 4.97). Hourly data frequency perform better in 19% of sample markets. Five minutes data frequency perform better in 25% of the stock markets. Figure 4.97 present the comparison across the data frequencies. Stock market indices EUS, DAX, HIS, NIFTY, NIKKEI, AEX, SMI, DJI30, and

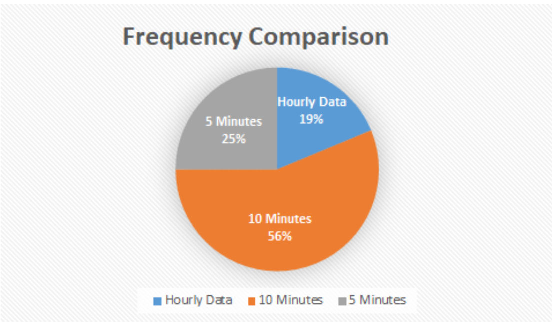


FIGURE 4.97: Frequency Comparison

SNP500 are better forecasted by using 10 minutes data frequency in comparison to the hourly and 5 minutes data frequency. ASX, WIG and SSGF are better forecasted by using hourly data frequency. Whereas 5-minute data frequency is more reliable in the case of A50, CAC, IBX, and FTSE stock market indices.

TABLE 4.22: Frequency Comparison

	Hourly Data			**10 minutes Data**			**5 minutes Data**		
	MAE	**MAPE**	**RMSE**	**MAE**	**Mape**	**RMSE**	**MAE**	**MAPE**	**RMSE**
ASX	20.981	0.355	32.730	32.846	0.542	45.700	48.171	0.749	68.682
A50	49.842	0.366	61.674	23.734	0.176	33.532	14.938	0.111	22.435
EUS	15.492	0.504	26.568	9.490	0.325	20.511	11.375	0.367	18.032
CAC	23.858	0.496	40.852	26.682	0.526	42.588	19.505	0.370	27.026
DAX	42.214	0.380	63.224	28.506	0.275	81.328	58.051	0.491	74.396
HSI	98.146	0.389	132.942	49.264	0.198	81.481	64.701	0.262	93.922
NIFTY	91.808	0.974	147.677	84.356	0.868	122.400	111.104	1.190	194.306
NIKKEI	97.953	0.461	131.214	44.608	0.219	69.466	49.159	0.241	67.391
AEX	2.151	0.417	3.800	1.581	0.301	2.499	2.470	0.493	5.505
WIG	43.470	2.716	69.422	47.468	3.018	80.525	44.635	2.830	78.015
SSGF	3.408	1.155	5.363	3.801	1.303	6.672	3.790	1.280	5.702
IBX	91.783	1.295	146.223	76.935	1.069	122.028	55.332	0.735	80.536
SMI	75.077	0.726	111.302	49.344	0.476	76.603	79.595	0.758	117.141
FTSE	96.843	1.696	188.547	80.370	1.432	176.606	75.878	1.362	182.042
DJI	104.028	0.423	174.009	74.515	0.288	108.133	76.491	0.291	99.307
SNP	20.142	0.652	29.492	17.332	0.556	25.066	22.281	0.709	31.441

Although high frequency hold more market information than the low frequency relatively (Andersen et al., 2003). But as the data frequency increases it will create more noise at the end of the day in data Corradi and Distaso (2006). In low frequency data noise problem plays a limited role but has a significant impact on high-frequency data. Some validation schemes might delete the bad data but has nothing to do with noise of data. Micro activities , rumors, disinformation and delay in forecast are the some sources of noise in data (Zhou, 1996). These financial forecasting models are design to forecast the presence of pattern within the data set not the noise. As we move toward the high frequency more noise reduces the forecasting power of any model.

To answer the question “appropriate frequency for forecasting return behavior” we come up with a conclusion based on (figure 4.97) 10 minutes frequency is more reliable for forecasting purpose. That leads to the acceptance of null hypothesis that data frequencies affect the forecasting of pricing behavior of financial markets.

Chapter 5

Conclusion

The complexity of Financial Time series due to the dynamic environment with high noise, non-linear structure, stationarity problems, time varying and chaotic system makes tough to forecast. These Characteristic influence the forecasting ability of forecasting techniques more specifically the classical forecasting techniques (Li et al., 2020; Marszałek and Burczyński, 2014; Reisen et al., 2001). In period of chaos uncertainty increase massively. Today's markets are more volatile than the past. Literature support the GARCH based methods can handle these issues but the limitations of GARCH that with the reference to out of sample forecasting still raise the questions (Crawford and Fratantoni, 2003). Machine learning and deep learning methods capture the all phenomenon's but for the search of better do not stop. The prime objective of the study is to find the best forecasting technique in selected stock markets. Time plays an important part hance data frequency play a significant rule in performance of any model. Therefore another objective of the study is linked with selection of best time frequency for the stock market forecasting.

This study compare some prominent classical forecasting techniques along with the soft computing i.e., machine learning and deep learning techniques. It also, proposed another combination of Classical GARCH model with Deep learning LSTM method that covered the limitations of GARCH and strengthen the LSTM. In order to test our hypothesis stock markets indices data has been collected for three different frequencies from July 2017 to 2020. All of the tests have been applied on raw data. Empirical results validate the researcher concern over the performance of classical econometric techniques and machine learning techniques over Deep learning-based method. Deep learning-based Hybrid methods outperform all of the tested techniques. In hourly data analysis 63% of sample indices showed that GARCH-LSTM improves the forecasting ability. Similarly, in 10 minutes data frequency results shows 56% of sample indices are forecastable with GARCH-LSTM method. In 5 minutes data frequency, GARCH did not perform as expected. Based on empirical results the study conclude that the Hybrid GARCH-LSTM model is best fitted model to forecast the stock market indices.

To answer the question of best time interval, this study compare the forecasting of thre three different time frequencies. In selected equity markets we noticed that with the increase in frequency mean median, skewness decreases. Whereas, standard deviation and kurtosis of data increase as frequency of data increases. Empirical results show that 10 minutes data frequency perform batter then the 5 minutes and hourly data frequencies. Results satisfy our research objective of appropriate frequency to use. Based on empirical results this study suggests users to use 10-minute data frequency rather than using more high frequency will creates the noise related issues that ultimately reduces the predictive power of models. Noise in data plays a critical role. As we move towards the higher frequency noise element with in the data increase. That reduce the predictive power of forecasting methods. Because we can predict the pattern with in the data but not the noise.

5.1 Recommendations

At the end we would like to suggest on the basis of the empirical testing of our model to the practitioners and academicians to use the Hybrid GARCH-LSTM model as it combines the GARCH model benefits with LSTM a deep learning technique. This unique model helps to cater the uncertainty of market and produce better results than the ARIMA and based hybrid methods and standard classical, machine learning and deep learning techniques. Further this study used a standard parameter i.e., Batches size, epochs, neurons numbers and activation function etc., in tuning of model for each market and frequency. Results can be improved by using and changing objectives for further research, some other functions and parameters like increase in batch size or activation function.

5.2 Limitations and Future Direction

Every market have different dynamics and they needs different parameters to handle these dynamics. To make it comparable the study used a standard tuning procedure for each market. Initially this study was design to compare four different data frequencies. 1 minute's data frequency was initially part of this study. To handle such a huge observation, we faced computational limitations. Such huge observations need more fast and up-to-date computers for computations. With increase of computational capacity more high frequency data can be handled and produced results may be more reliable. For future direction the study recommend to used convolutional neural network wavelet data denoising technique to denoise the data. Further we recommend to use other variants of GARCH family techniques like EGARCH, GJAR-GARCH or DCC-GARCH with some exogenous variables.

Bibliography

Adhikari, R. and Agrawal, R. K. (2014). A combination of artificial neural network and random walk models for financial time series forecasting. *Neural Computing and Applications*, 24(6):1441–1449.

Al Wadia, M. and Ismail, M. T. (2011). Selecting wavelet transforms model in forecasting financial time series data based on arima model. *Applied Mathematical Sciences*, 5(7):315–326.

Alam, R., Arnob, R. I., and Alam, A. E. (2020). An arima-lstm correlation coefficient based hybrid model for portfolio management of dhaka stock exchange. In *International Conference for Emerging Technologies in Computing*, volume 332, pages 214–226.

Alavi, A. H. and Gandomi, A. H. (2011). Prediction of principal ground-motion parameters using a hybrid method coupling artificial neural networks and simulated annealing. *Computers & Structures*, 89(23-24):2176–2194.

Aldridge, I. and Krawciw, S. (2017). *Real-time risk: What investors should know about FinTech, high-frequency trading, and flash crashes.* John Wiley & Sons.

Altin, H. (2015). Efficient market hypothesis, abnormal return and election periods. *European Scientific Journal*, 11(34).

Alzahrani, A., Shamsi, P., Dagli, C., and Ferdowsi, M. (2017). Solar irradiance forecasting using deep neural networks. *Procedia Computer Science*, 114:304–313.

Andersen, T. G. and Bollerslev, T. (1997). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *The journal of Finance*, 52(3):975–1005.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625.

Andersson, M. K. (2000). Do long-memory models have long memory? *International Journal of Forecasting*, 16(1):121–124.

Archibald, B. C. and Koehler, A. B. (2003). Normalization of seasonal factors in winters' methods. *International Journal of Forecasting*, 19(1):143–148.

Arévalo, A., Niño, J., Hernández, G., and Sandoval, J. (2016). High-frequency trading strategy based on deep neural networks. In *International conference on intelligent computing*, pages 424–436. Springer.

Armstrong, J. S., Green, K. C., and Graefe, A. (2015). Golden rule of forecasting: Be conservative. *Journal of Business Research*, 68(8):1717–1731.

Assaf, A. (2006). Dependence and mean reversion in stock prices: The case of the mena region. *Research in International Business and Finance*, 20(3):286–304.

Atsalakis, G. S. and Valavanis, K. P. (2009). Surveying stock market forecasting techniques-part ii: Soft computing methods. *Expert Systems with applications*, 36(3):5932–5941.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of econometrics*, 73(1):5–59.

Barrales, E. O. (2012). Lessons from the flash crash for the regulation of high-frequency traders. *Fordham J. Corp. & Fin. L.*, 17:1195.

Baur, D. G. and Dimpfl, T. (2018). The asymmetric return-volatility relationship of commodity prices. *Energy Economics*, 76:378–387.

Bhardwaj, G. and Swanson, N. R. (2006). An empirical investigation of the usefulness of arfima models for predicting macroeconomic and financial time series. *Journal of Econometrics*, 131(1):539–578.

Boehmer, E. and Kelley, E. K. (2009). Institutional investors and the informational efficiency of prices. *The Review of Financial Studies*, 22(9):3563–3594.

Bouoiyour, J. and Selmi, R. (2018). Are uk industries resilient in dealing with uncertainty? the case of brexit. *European Journal of Comparative Economics*, 15(2):277–292.

Box, G. E. (1979). Robustness in the strategy of scientific model building. In *Robustness in statistics*, pages 201–236. Elsevier.

Box, G. E. and Jenkins, G. M. (1970). Time series analysis forecasting and control. Report, Wisconsin univ madison dept of statistics.

Brooks, C. (1998). Predicting stock index volatility: can market volume help? *Journal of Forecasting*, 17(1):59–80.

Brooks, C., Burke, S. P., and Persand, G. (2001). Benchmarks and the accuracy of garch model estimation. *International Journal of Forecasting*, 17(1):45 – 56.

Brownlee, J. (2018). *Deep learning for time series forecasting: Predict the future with MLPs, CNNs and LSTMs in Python*. Machine Learning Mastery.

Busseti, E., Osband, I., and Wong, S. (2012). Deep learning for time series modeling. *Technical report, Stanford University*, pages 1–5.

Cadenas, E. and Rivera, W. (2010). Wind speed forecasting in three different regions of mexico, using a hybrid arima–ann model. *Renewable Energy*, 35(12):2732–2738.

Cadenas, E., Rivera, W., Campos-Amezcua, R., and Heard, C. (2016). Wind speed prediction using a univariate arima model and a multivariate narx model. *Energies*, 9(2):109.

Caner, S. and Önder*, Z. (2005). Sources of volatility in stock returns in emerging markets. *Applied Economics*, 37(8):929–941.

Cao, J., Li, Z., and Li, J. (2019). Financial time series forecasting model based on ceemdan and lstm. *Physica A: Statistical Mechanics and its Applications*, 519:127–139.

Cavalcante, R. C., Brasileiro, R. C., Souza, V. L., Nobrega, J. P., and Oliveira, A. L. (2016). Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications*, 55:194–211.

Chang, P.-C., Liu, C.-H., Lin, J.-L., Fan, C.-Y., and Ng, C. S. (2009). A neural network with a case based dynamic window for stock trading prediction. *Expert Systems with Applications*, 36(3):6889–6898.

Charles, A. and Darné, O. (2005). Outliers and garch models in financial data. *Economics Letters*, 86(3):347–352.

Chen, X., Jin, S., Qin, S., and Li, L. (2015). Short-term wind speed forecasting study and its application using a hybrid model optimized by cuckoo search. *Mathematical Problems in Engineering*, 2015:18.

Choi, H. K. (2018). Stock price correlation coefficient prediction with arima-lstm hybrid model. *arXiv preprint arXiv:1808.01560*.

Choi, J., Lee, M., and Rhee, M. (1995). Trading s&p 500 stock index futures using a neural network. In *Proceedings of the third annual international conference on artificial intelligence applications on wall street*, pages 63–72.

Chong, E., Han, C., and Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, 83:187–205.

Copeland, T. E. (1976). A model of asset trading under the assumption of sequential information arrival. *The Journal of Finance*, 31(4):1149–1168.

Corradi, V. and Distaso, W. (2006). Semi-parametric comparison of stochastic volatility models using realized measures. *The Review of Economic Studies*, 73(3):635–667.

Corrêa, J. M., Neto, A. C., Júnior, L. T., Franco, E., and Faria Jr, A. (2016). Time series forecasting with the warimax-garch method. *Neurocomputing*, 216:805–815.

Crawford, G. W. and Fratantoni, M. C. (2003). Assessing the forecasting performance of regime-switching, arima and garch models of house prices. *Real Estate Economics*, 31(2):223–243.

Daniel, K. and Titman, S. (1999). Market efficiency in an irrational world. *Financial Analysts Journal*, 55(6):28–40.

De Gooijer, J. G. and Hyndman, R. J. (2006). 25 years of time series forecasting. *International journal of forecasting*, 22(3):443–473.

de Oliveira, J. F. L. and Ludermir, T. B. (2014). A distributed pso-arima-svr hybrid system for time series forecasting. In *2014 IEEE international conference on systems, man, and cybernetics (SMC)*, pages 3867–3872. IEEE.

Diebold, F. X. and Nason, J. A. (1990). Nonparametric exchange rate prediction? *Journal of international Economics*, 28(3-4):315–332.

Ding, Z., Granger, C. W., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of empirical finance*, 1(1):83–106.

Dingli, A. and Fournier, K. S. (2017). Financial time series forecasting-a deep learning approach. *Int. J. Mach. Learn. Comput*, 7(5):118–122.

Ederington, L. H. and Lee, J. H. (1993). How markets process information: News releases and volatility. *The Journal of Finance*, 48(4):1161–1191.

Ehteram, M., Ahmed, A. N., Fai, C. M., Afan, H. A., and El-Shafie, A. (2019). Accuracy enhancement for zone mapping of a solar radiation forecasting based multi-objective model for better management of the generation of renewable energy. *Energies*, 12(14):2730.

Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of variance of united kingdom inflation. *Econometrica*, 50:987–1008.

Ersan, D., Nishioka, C., and Scherp, A. (2019). Comparison of machine learning methods for financial time series forecasting at the examples of over 10 years of daily and hourly data of dax 30 and s&p 500. *Journal of Computational Social Science*, 3:1–31.

Fabozzi, F. J., Modigliani, F., and Jones, F. J. (2013). *Foundations of financial markets and institutions*. Pearson Education.

Falinouss, P. (2007). Stock trend prediction using news articles: a text mining approach. Master's thesis. Validerat; 20101217 (root).

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2):383–417.

Fang, X. and Yuan, Z. (2019). Performance enhancing techniques for deep learning models in time series forecasting. *Engineering Applications of Artificial Intelligence*, 85:533–542.

Fischer, T. and Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2):654–669.

Floridi, L. (2005). Is semantic information meaningful data? *Philosophy and research, phenomenological*, 70(2):351–370.

Floros, C., Jaffry, S., and Lima, G. V. (2007). Long memory in the portuguese stock market. *Studies in Economics and Finance*, 24(3):220–232.

Franses, P. H. and Ghijsels, H. (1999). Additive outliers, garch and forecasting volatility. *International Journal of forecasting*, 15(1):1–9.

Frino, A. and Hill, A. (2001). Intraday futures market behaviour around major scheduled macroeconomic announcements: Australian evidence. *Journal of Banking & Finance*, 25(7):1319–1337.

Fu, L., Song, W., and Lo, S. (2017). A fuzzy-theory-based method for studying the effect of information transmission on nonlinear crowd dispersion dynamics. *Communications in Nonlinear Science and Numerical Simulation*, 42:682–698.

Gibson, G. R. (1888). *The stock exchanges of London, Paris, and New York: a comparison*. GP Putnam.

Giot, P., Laurent, S., and Petitjean, M. (2010). Trading activity, realized volatility and jumps. *Journal of Empirical Finance*, 17(1):168–175.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, 48(5):1779–1801.

Gokcan, S. (2000). Forecasting volatility of emerging stock markets: linear versus non-linear garch models. *Journal of Forecasting*, 19(6):499–504.

Golden, J., Milewicz, J., and Herbig, P. (1994). Forecasting: Trials and tribulations. *Management Decision*, 32(1):33–36.

Goldman, S. (1999). Computational learning theory. In *Algorithms and theory of computation handbook: general concepts and techniques*.

Gong, X., Si, Y.-W., Fong, S., and Biuk-Aghai, R. P. (2016). Financial time series pattern matching with extended ucr suite and support vector machine. *Expert Systems with Applications*, 55:284–296.

Gonzalez-Vidal, A., Jimenez, F., and Gomez-Skarmeta, A. F. (2019). A methodology for energy multivariate time series forecasting in smart buildings based on feature selection. *Energy and Buildings*, 196:71–82.

Goudarzi, H. and Ramanarayanan, C. (2011). Modeling asymmetric volatility in the indian stock market. *International Journal of Business and Management*, 6(3):221.

Granger, C. W. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of time series analysis*, 1(1):15–29.

Haloub, R. (2013). *Assessment of forecasting management in international pharmaceutical companies: A grounded theory study*. Thesis, University of Huddersfield.

Hein, S. E. and Spudeck, R. E. (1988). Forecasting the daily federal funds rate. *International Journal of Forecasting*, 4(4):581–591.

Hippert, H. S., Pedreira, C. E., and Souza, R. C. (2001). Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on power systems*, 16(1):44–55.

Ho, S. and Xie, M. (1998). The use of arima models for reliability forecasting and analysis. *Computers & industrial engineering*, 35(1-2):213–216.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.

Hofman, B., Yong, S., and Yao, L. (2020). Implications of the covid-19 outbreak of china's economy. Technical report, Tech. rep. East Asian Institute National University of Singapore.

Hong, H. and Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of finance*, 54(6):2143–2184.

Horváth, L., Kokoszka, P., and Rice, G. (2014). Testing stationarity of functional time series. *Journal of Econometrics*, 179(1):66–82.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, 68(1):165–176.

Hsu, M.-W., Lessmann, S., Sung, M.-C., Ma, T., and Johnson, J. E. V. (2016). Bridging the divide in financial market forecasting: machine learners vs. financial economists. *Expert Systems with Applications*, 61:215–234.

Kaastra, I. and Boyd, M. (1996). Designing a neural network for forecasting financial. *Neurocomputing*, 10:215–236.

Kewat, P., Sharma, R., Singh, U., and Itare, R. (2017). Support vector machines through financial time series forecasting. In *2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)*, volume 2, pages 471–477. IEEE.

Khairalla, M. and AL-Jallad, N. T. (2017). Hybrid forecasting scheme for financial time-series data using neural network and statistical methods. *International Journal of Advanced Computer Science and Applications*, 8(9):319–327.

Khashei, M. and Hajirahimi, Z. (2017). Performance evaluation of series and parallel strategies for financial time series forecasting. *Financial Innovation*, 3(1):24.

Kikuchi, R., Misaka, T., Obayashi, S., Inokuchi, H., Oikawa, H., and Misumi, A. (2018). Nowcasting algorithm for wind fields using ensemble forecasting and aircraft flight data. *Meteorological Applications*, 25(3):365–375.

Kim, K.-j. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, 55(1):307–319.

Kodogiannis, V. and Lolis, A. (2002). Forecasting financial time series using neural network and fuzzy system-based techniques. *Neural computing and applications*, 11(2):90–102.

Kolaric, S. and Schiereck, D. (2016). Are stock markets efficient in the face of fear? evidence from the terrorist attacks in paris and brussels. *Finance Research Letters*, 18:306–310.

Koop, G., Ley, E., Osiewalski, J., and Steel, M. F. (1997). Bayesian analysis of long memory and persistence using arfima models. *Journal of Econometrics*, 76(1):149–170.

Kristjanpoller, W. and Hernández, E. (2017). Volatility of main metals forecasted by a hybrid ann-garch model with regressors. *Expert Systems with Applications*, 84:290–300.

Kristjanpoller R, W. and Michell V, K. (2018). A stock market risk forecasting model through integration of switching regime, anfis and garch techniques. *Applied Soft Computing*, 67:106–116.

Krzysztofowicz, R. (2001). The case for probabilistic forecasting in hydrology. *Journal of Hydrology*, 249(1):2–9.

Kucharavy, D. and De Guio, R. (2005). Problems of forecast. In *ETRIA TRIZ Future 2005*, pages 219–235, Graz, Austria.

Kulshreshtha, S. (2020). An arima-lstm hybrid model for stock market prediction using live data. *Journal of Engineering Science and Review, Technology*, 13(4).

Kumar, D., Meghwani, S. S., and Thakur, M. (2016). Proximal support vector machine based hybrid prediction models for trend forecasting in financial markets. *Journal of Computational Science*, 17:1–13.

Kumar, M. and Thenmozhi, M. (2014). Forecasting stock index returns using arima-svm, arima-ann, and arima-random forest hybrid models. *International Journal of Banking, Accounting and Finance*, 5(3):284–308.

Kwon, H., Lyu, B., Tak, K., Lee, J., and Moon, I. (2015). Optimization of petrochemical process planning using naphtha price forecasting and process modeling. In *Computer Aided Chemical Engineering*, volume 37, pages 2039–2044. Elsevier.

Lahmiri, S. and Bekiros, S. (2019). Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos, Solitons & Fractals*, 118:35–40.

Längkvist, M., Karlsson, L., and Loutfi, A. (2014). A review of unsupervised feature learning and deep learning for time-series modeling. *Pattern Recognition Letters*, 42:11–24.

Lasek, A., Cercone, N., and Saunders, J. (2016). *Chapter 17 - Smart restaurants: survey on customer demand and sales forecasting*, pages 361–386. Morgan Kaufmann, Boston.

Li, W., Luo, Y., Zhu, Q., Liu, J., and Le, J. (2008). Applications of ar*-grnn model for financial time series forecasting. *Neural Computing*, 17(5-6):441–448.

Li, X., Lv, Z., Hu, J., Zhang, B., Yin, L., Zhong, C., Wang, W., and Feng, S. (2015). Traffic management and forecasting system based on 3d gis. In *2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, pages 991–998. IEEE.

Li, Z., Han, J., and Song, Y. (2020). On the forecasting of high-frequency financial time series based on arima model improved by deep learning. *Journal of Forecasting*, n/a(n/a).

Liu, Q. (2009). On portfolio optimization: How and when do we benefit from high-frequency data? *Journal of Applied Econometrics*, 24(4):560–582.

Livieris, I. E., Pintelas, E., and Pintelas, P. (2020). A CNN-LSTM model for gold price time-series forecasting. *Neural Computing and Applications*, 32(23):17351–17360.

Long, W., Lu, Z., and Cui, L. (2019). Deep learning-based feature engineering for stock price movement prediction. *Knowledge-Based Systems*, 164:163–173.

Lu, W.-Z. and Wang, W.-J. (2005). Potential assessment of the support vector machine method in forecasting ambient air pollutant trends. *chemosphere*, 59(5):693–701.

Luss, R. and d'Aspremont, A. (2015). Predicting abnormal returns from news using text classification. *Quantitative Finance*, 15(6):999–1012.

Marszałek, A. and Burczyński, T. (2014). Modeling and forecasting financial time series with ordered fuzzy candlesticks. *Information sciences*, 273:144–155.

Matias, J. M. and Reboredo, J. C. (2012). Forecasting performance of nonlinear models for intraday stock returns. *Journal of Forecasting*, 31(2):172–188.

Medeiros, M. C., Teräsvirta, T., and Rech, G. (2006). Building neural network models for time series: a statistical approach. *Journal of Forecasting*, 25(1):49–75.

Merrill Eric, E. S. (2019). Forecasting in a digital world.

Min, J. H. and Lee, Y.-C. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications*, 28(4):603–614.

Mirakyan, A., Meyer-Renschhausen, M., and Koch, A. (2017). Composite forecasting approach, application for next-day electricity price forecasting. *Energy Economics*, 66:228–237.

Mobarek, A. and Keasey, K. (2000). Weak-form market efficiency of an emerging market: Evidence from dhaka stock market of bangladesh. In *ENBS Conference held on Oslo*, pages 1–30.

Moews, B., Herrmann, J. M., and Ibikunle, G. (2019). Lagged correlation-based deep learning for directional trend change prediction in financial time series. *Expert Systems with Applications*, 120:197–206.

Muth, J. F. (1960). Optimal properties of exponentially weighted forecasts. *Journal of the american statistical association*, 55(290):299–306.

Nau, R. (2014). Statistical forecasting: notes on regression and time series analysis.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pages 347–370.

Niu, T., Wang, J., Lu, H., Yang, W., and Du, P. (2020). Developing a deep learning framework with two-stage feature selection for multivariate financial time series forecasting. *Expert Systems with Applications*, 148:113237.

Papacharalampous, G., Tyralis, H., and Koutsoyiannis, D. (2019). Comparison of stochastic and machine learning methods for multi-step ahead forecasting of hydrological processes. *Stochastic environmental research and risk assessment*, 33(2):481–514.

Poon, S.-H. and Granger, C. W. (2003). Forecasting volatility in financial markets: A review. *Journal of economic literature*, 41(2):478–539.

Poulos, L., Kvanli, A., and Pavur, R. (1987). A comparison of the accuracy of the box-jenkins method with that of automated forecasting methods. *International Journal of Forecasting*, 3(2):261–267.

Prescott, D. and Stegnos, T. (1989). Do asset markets overlook exploitable non-linearities. *The case of gold. University of Guelph, Ontario.*

Prosad, J. M., Kapoor, S., and Sengupta, J. (2015). Exploring optimism and pessimism in the indian equity market. *Review of Behavioral Finance.*

Provost, F. J. and Aronis, J. M. (1996). Scaling up inductive learning with massive parallelism. *Machine Learning*, 23(1):33–46.

Qi, M. (1999). Nonlinear predictability of stock returns using financial and economic variables. *Journal of Business & Economic Statistics*, 17(4):419–429.

Qu, Y. and Dan, Y. (2014). Modeling the evacuation behavior considering the effect of dangerous source. *Procedia-social and behavioral sciences*, 138(14):800–810.

Quah, T.-S. and Srinivasan, B. (1999). Improving returns on stock investment through neural network selection. *Expert Systems with Applications*, 17(4):295–301.

Racine, J. (2001). On the nonlinear predictability of stock returns using financial and economic variables. *Journal of Business & Economic Statistics*, 19(3):380–382.

Reisen, V., Abraham, B., and Lopes, S. (2001). Estimation of parameters in arfima processes: A simulation study. *Communications in Statistics - Simulation and Computation*, 30(4):787–803.

Rosas, A. L. and Guerrero, V. M. (1994). Restricted forecasts using exponential smoothing techniques. *International Journal of Forecasting*, 10(4):515–527.

Rounaghi, M. M. and Nassir Zadeh, F. (2016). Investigation of market efficiency and financial stability between s&p 500 and london stock exchange: Monthly and yearly forecasting of time series stock returns using arma model. *Physica A: Statistical Mechanics and its Applications*, 456:10–21.

Sabbatini, M. and Linton, O. (1998). A garch model of the implied volatility of the swiss market index from option prices. *International Journal of Forecasting*, 14(2):199–213.

Sagheer, A. and Kotb, M. (2019). Time series forecasting of petroleum production using deep lstm recurrent networks. *Neurocomputing*, 323:203–213.

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of research development*, 3(3):210–229.

Schrider, D. R. and Kern, A. D. (2018). Supervised machine learning for population genetics: A new paradigm. *Trends in Genetics*, 34(4):301–312.

Sensoy, A., Fabozzi, F. J., and Eraslan, V. (2017). Predictability dynamics of emerging sovereign cds markets. *Economics Letters*, 161:5–9.

Sezer, O. B., Gudelek, M. U., and Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning : A systematic literature review: 2005 2019. *Applied Soft Computing*, 90:106181.

Shen, G., Tan, Q., Zhang, H., Zeng, P., and Xu, J. (2018). Deep learning with gated recurrent unit networks for financial sequence predictions. *Procedia computer science*, 131:895–903.

Shintate, T. and Pichl, L. (2019). Trend prediction classification for high frequency bitcoin time series with deep learning. *Journal of Risk and Financial Management*, 12(1):17.

Siami-Namini, S. and Namin, A. S. (2018). Forecasting economics and financial time series: Arima vs. lstm. *arXiv preprint arXiv:1803.06386*.

Sibbertsen, P. (2004). Long memory versus structural breaks: An overview. *Statistical Papers*, 45(4):465–515.

Slim, S. and Dahmene, M. (2016). Asymmetric information, volatility components and the volume–volatility relationship for the cac40 stocks. *Global Finance Journal*, 29:70–84.

Smales, L. A. (2013). Impact of macroeconomic announcements on interest rate futures: High-frequency evidence from australia. *Journal of Financial Research*, 36(3):371–388.

Smales, L. A. (2015). Asymmetric volatility response to news sentiment in gold futures. *Journal of International Financial Markets, Institutions and Money*, 34:161 – 172.

StataCorp (2019). Stata time-series reference manual.

Strozzi, F. and Zaldivar, J. (2005). Non-linear forecasting in high-frequency financial time series. *Physica A: Statistical Mechanics and its Applications*, 353:463–479.

Tay, F. E. H. and Cao, L. (2001). Application of support vector machines in financial time series forecasting. *Omega*, 29(4):309–317.

Tay, F. E. H. and Cao, L. J. (2002). Modified support vector machines in financial time series forecasting. *Neurocomputing*, 48(1):847–861.

Taylor, S. J. (2000). Consequences for option pricing of a long memory in volatility. In *EFA 2001 Barcelona Meetings, EFMA 2001 Lugano Meetings*.

Tealab, A. (2018). Time series forecasting using artificial neural networks methodologies: A systematic review. *Future Computing and Informatics Journal*, 3(2):334–340.

Temür, A. S., Akgün, M., and Temür, G. (2019). Predicting housing sales in turkey using arima, lstm and hybrid models. *Journal of Business Economics and Management*, 20(5):920–938.

Ticknor, J. L. (2013). A bayesian regularized artificial neural network for stock market forecasting. *Expert Systems with Applications*, 40(14):5501–5506.

Tobin, J. (1984). On the efficiency of the financial-system. *Lloyds Bank Annual Review*, 2(153):1–15.

Trippi, R. R. and Turban, E. (1992). *Neural networks in finance and investing: Using artificial intelligence to improve real world performance.* McGraw-Hill, Inc.

Tularam, G. A. and Saeed, T. (2016). Oil-price forecasting based on various univariate time-series models. *American Journal of Operations Research*, 6(3):226–235.

Tyree, E. W. and Long, J. (1995). Forecasting currency exchange rates: neural networks and the random walk model. In *City University Working Paper, Proceedings of the Third International Conference on Artificial Intelligence Applications*. Citeseer.

Vaisla, K. S. and Bhatt, A. K. (2010). An analysis of the performance of artificial neural network technique for stock market forecasting. *International Journal on Computer Science and Engineering*, 2(6):2104–2109.

Van Gestel, T., Espinoza, M., Baesens, B., Suykens, J. A., Brasseur, C., and De Moor, B. (2006). A bayesian nonlinear support vector machine error correction model. *Journal of forecasting*, 25(2):77–100.

Van Horne, J. C. and Parker, G. G. (1967). The random-walk theory: an empirical test. *Financial Analysts Journal*, 23(6):87–92.

Vapnik, V. N. (1995). T he nature of statistical l earning t heory. *New York: Springer2V erlag.*

Visser, M. P. (2011). Garch parameter estimation using high-frequency data. *Journal of Financial Econometrics*, 9(1):162–197.

Wan, C., Xu, Z., Wang, Y., Dong, Z. Y., and Wong, K. P. (2014). A hybrid approach for probabilistic forecasting of electricity price. *IEEE Transactions on Smart Grid*, 5(1):463–470.

Wang, J., Wang, J., Fang, W., and Niu, H. (2016). Financial time series prediction using elman recurrent random neural networks. *Computational intelligence neuroscience*, 2016.

Wang, J.-J., Wang, J.-Z., Zhang, Z.-G., and Guo, S.-P. (2012a). Stock index forecasting based on a hybrid model. *Omega*, 40(6):758–766.

Wang, L., Zou, H., Su, J., Li, L., and Chaudhry, S. (2013). An arima?ann hybrid model for time series forecasting. *Systems Research Behavioral Science*, 30(3):244–259.

Wang, P., Zhang, H., Qin, Z., and Zhang, G. (2017). A novel hybrid-garch model based on arima and svm for pm2.5 concentrations forecasting. *Atmospheric Pollution Research*, 8(5):850–860.

Wang, X., Smith-Miles, K., and Hyndman, R. (2009). Rule induction for forecasting method selection: Meta-learning the characteristics of univariate time series. *Neurocomputing*, 72(10):2581–2594.

Wang, Y., Wang, J., Zhao, G., and Dong, Y. (2012b). Application of residual modification approach in seasonal arima for electricity demand forecasting: A case study of china. *Energy Policy*, 48:284–294.

Wei, Y., Wang, Y., and Huang, D. (2010). Forecasting crude oil market volatility: Further evidence using garch-class models. *Energy Economics*, 32(6):1477–1484.

White, H. (1988). Economic prediction using neural networks: The case of ibm daily stock returns. In *ICNN*, volume 2, pages 451–458.

Williams, D. W. and Miller, D. (1999). Level-adjusted exponential smoothing for modeling planned discontinuities. *International Journal of Forecasting*, 15(3):273–289.

Willmott, C. J. and Matsuura, K. (2005). Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate research*, 30(1):79–82.

Willmott, C. J., Matsuura, K., and Robeson, S. M. (2009). Ambiguities inherent in sums-of-squares-based error statistics. *Atmospheric Environment*, 43(3):749–752.

Wolf, H. (2005). Volatility: definitions and consequences. *J, Aizenman and B. Pinto Managing economic volatility and crises: A practitioner's guide*, pages 45–64.

Yan, H. and Ouyang, H. (2018). Financial time series prediction based on deep learning. *Wireless Personal Communications*, 102(2):683–700.

Yaseen, Z. M., Sulaiman, S. O., Deo, R. C., and Chau, K.-W. (2019). An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. *Journal of Hydrology*, 569:387–408.

Yule, G. U. (1927). Vii. on a method of investigating periodicities disturbed series, with special reference to wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 226(636-646):267–298.

Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8(3):338 – 353.

Żbikowski, K. (2015). Using volume weighted support vector machines with walk forward testing and feature selection for the purpose of creating stock trading strategy. *Expert Systems with Applications*, 42(4):1797–1805.

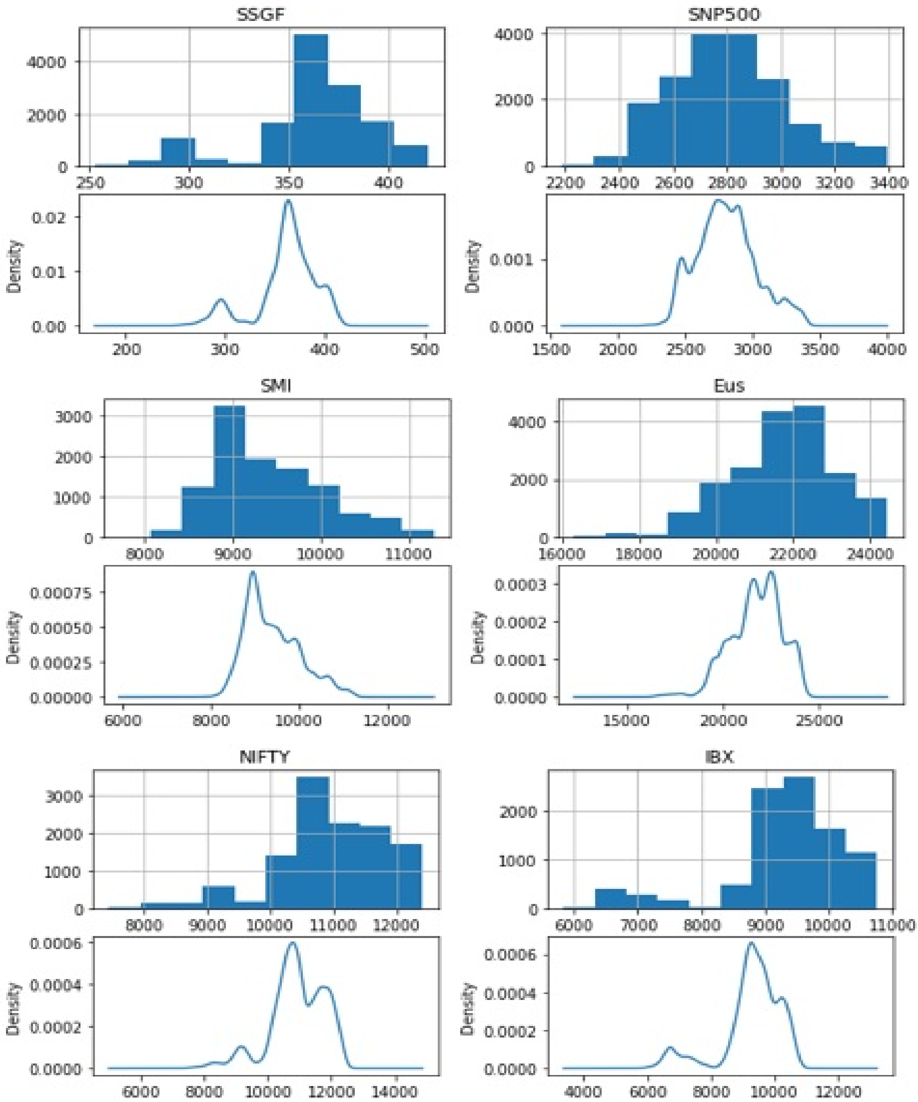
Zhang, G., Patuwo, B. E., and Hu, M. Y. (1998). Forecasting with artificial neural networks:: The state of the art. *International journal of forecasting*, 14(1):35–62.

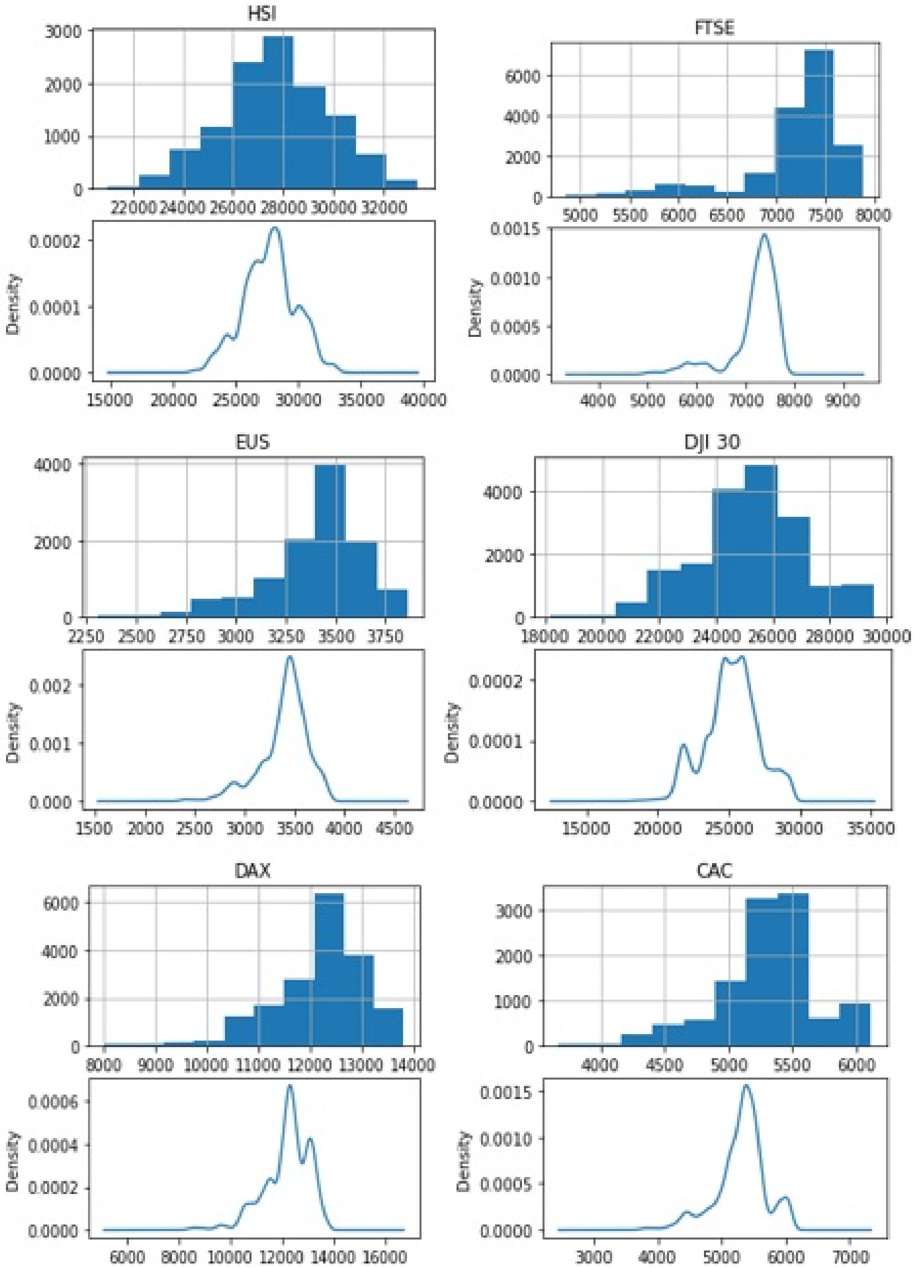
Zhang, G., Zhang, X., and Feng, H. (2016). Forecasting financial time series using a methodology based on autoregressive integrated moving average and taylor expansion. *Expert Systems*, 33(5):501–516.

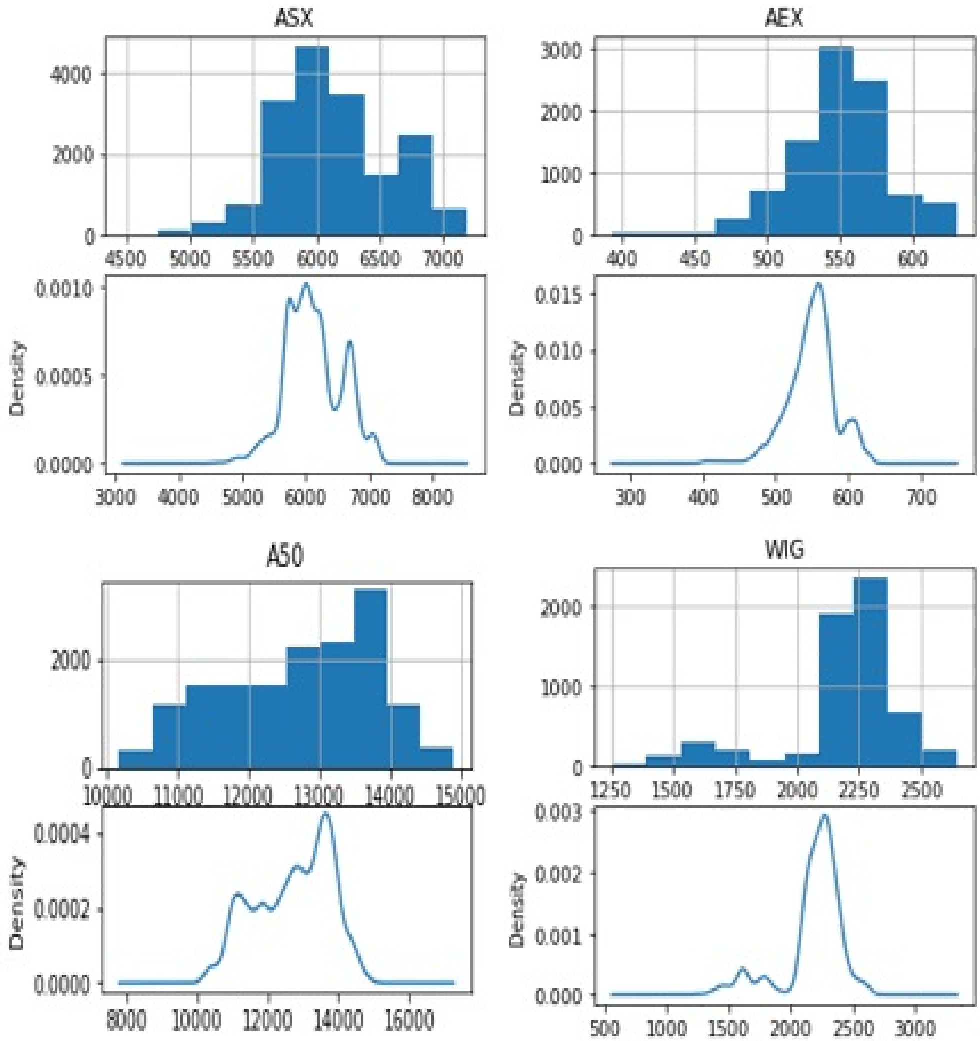
Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates. *Journal of Business & Economic Statistics*, 14(1):45–52.

Appendix A

Appendix A-1: KDE

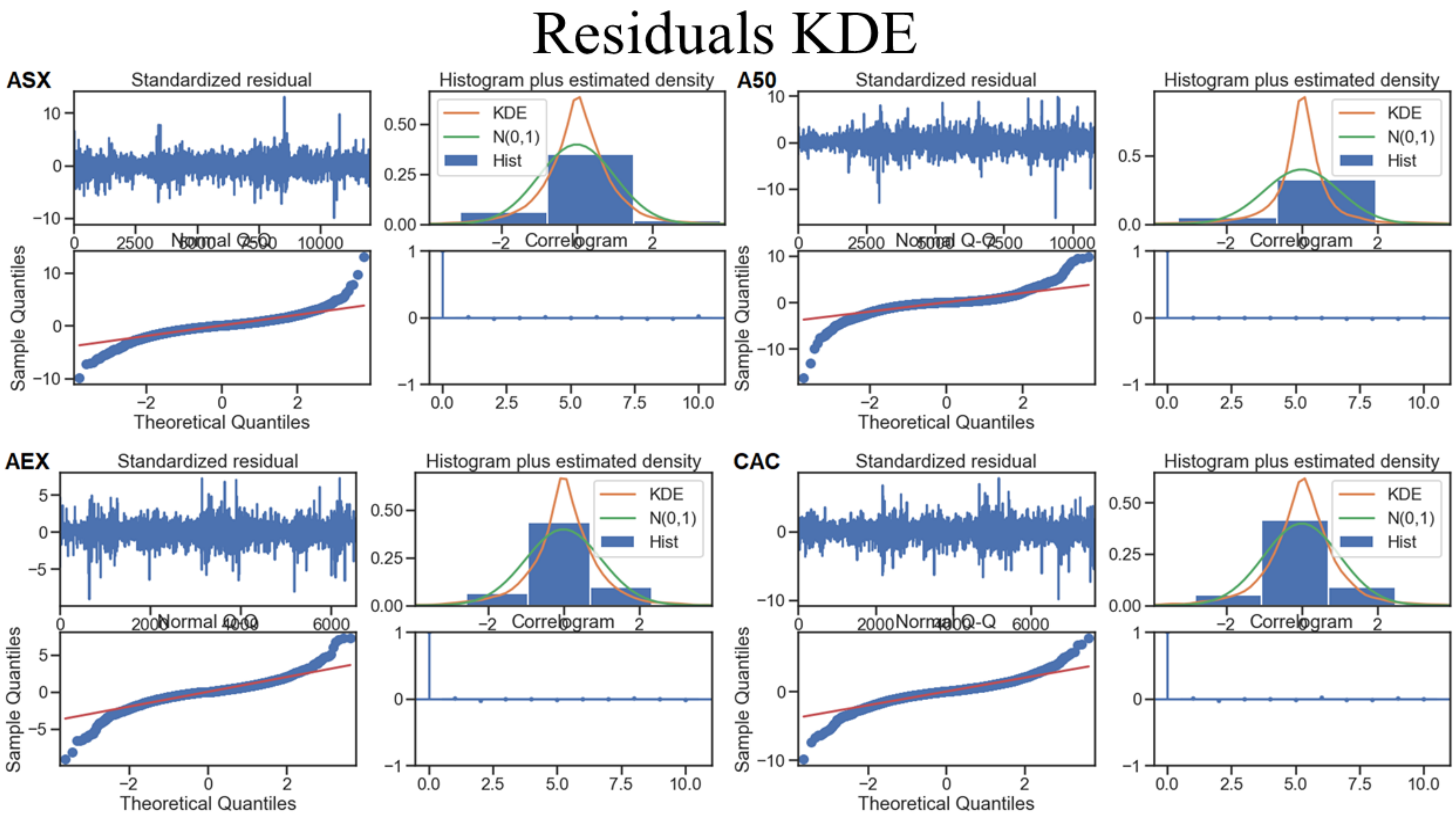


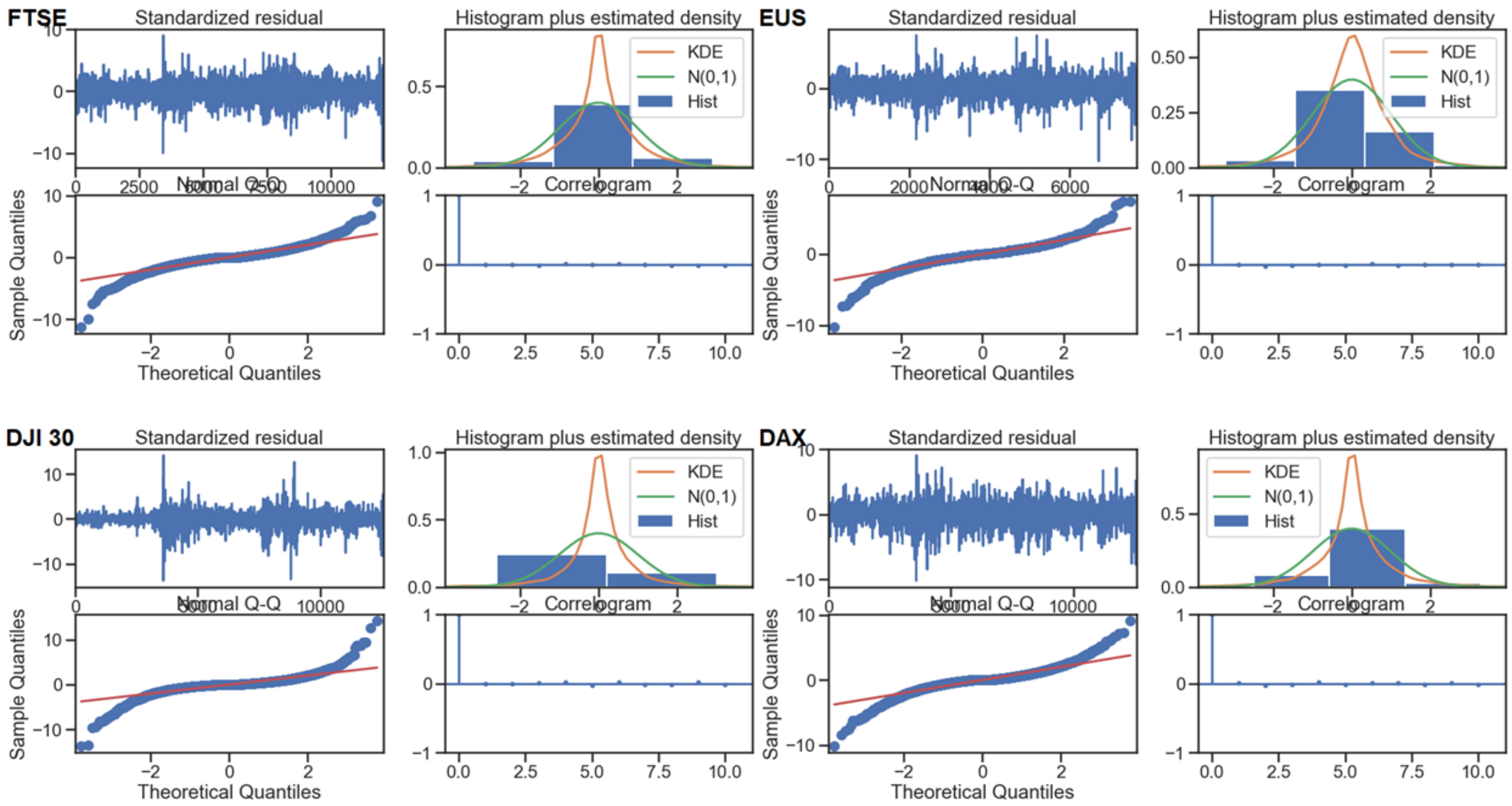

HSI
FTSE
Density
EUS
DJI 30
DAX
CAC

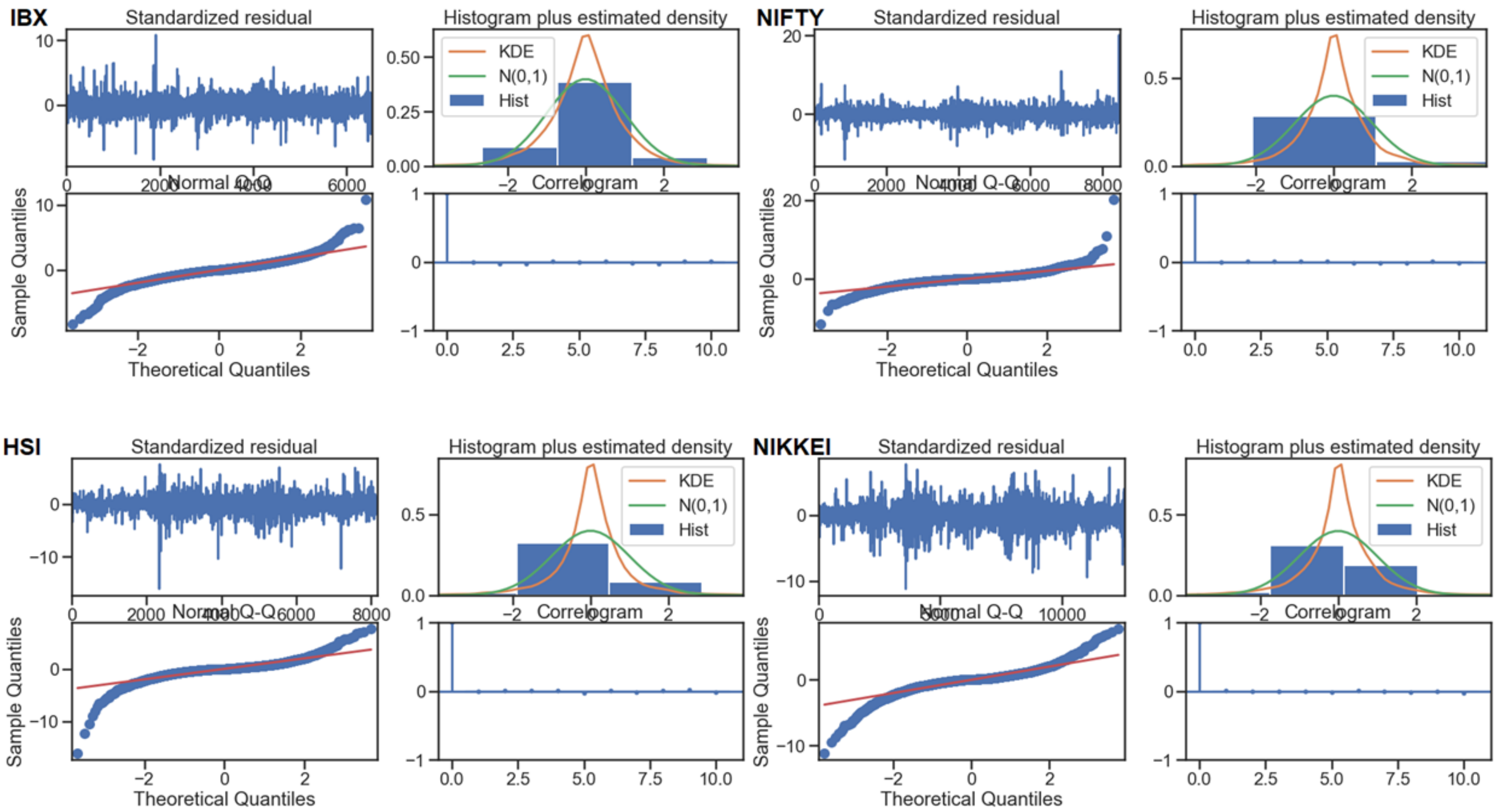


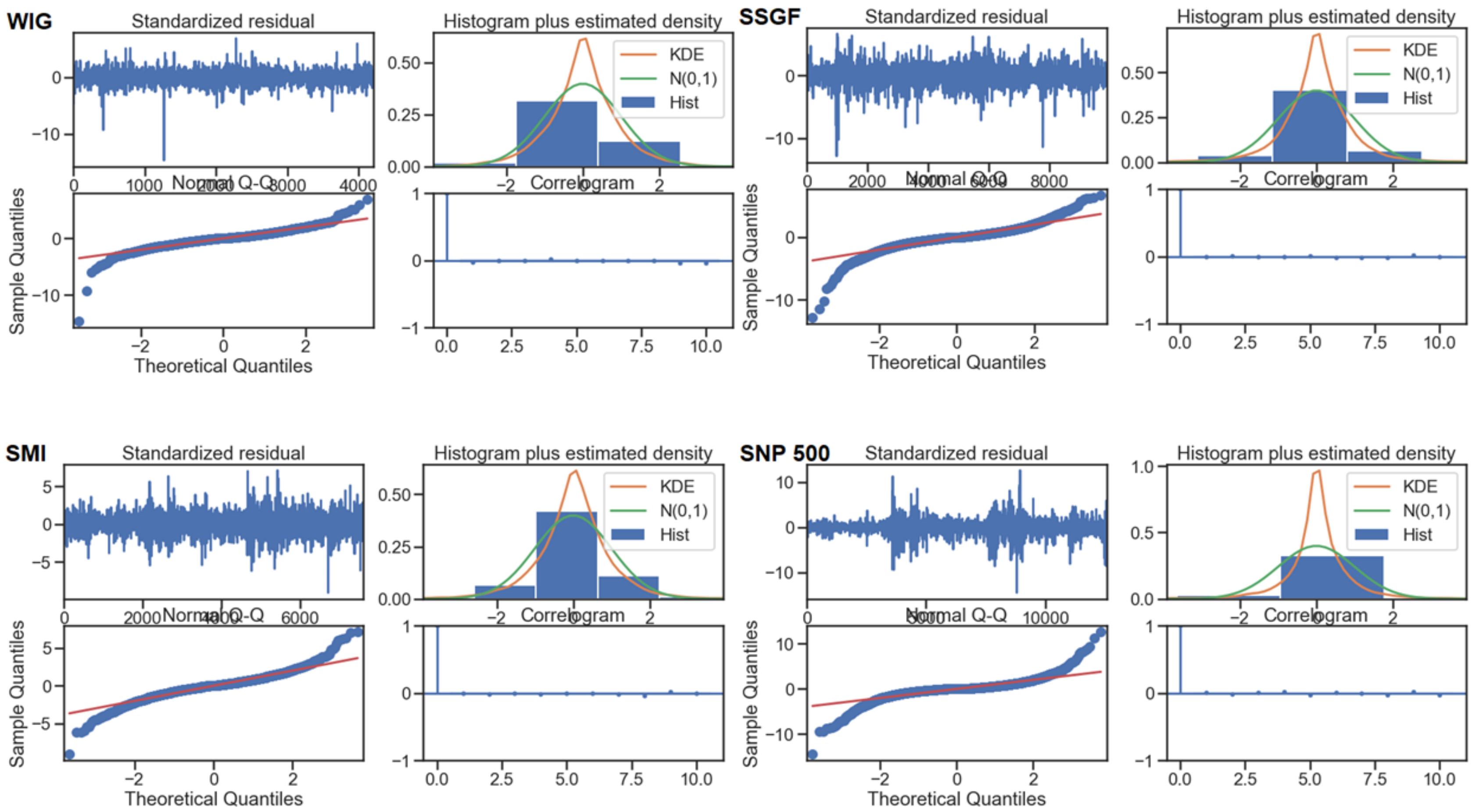
ASX
AEX
A50
WIG
Density

Appendix A-2: KDE of ARIMA models residuals

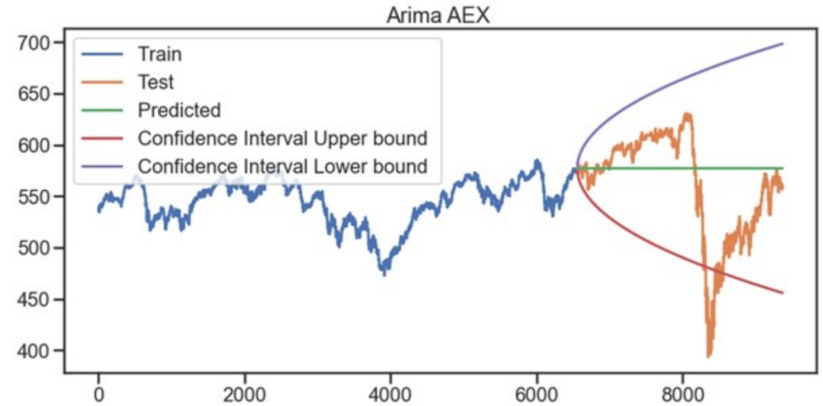


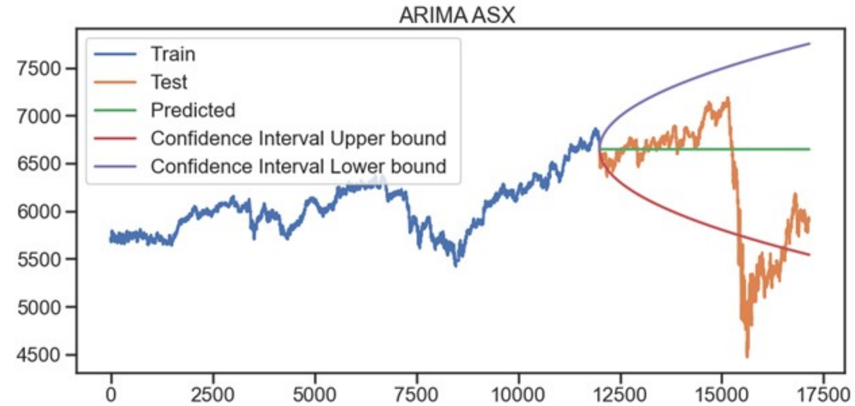

FTSE
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
EUS
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
DJI 30
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
DAX
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles

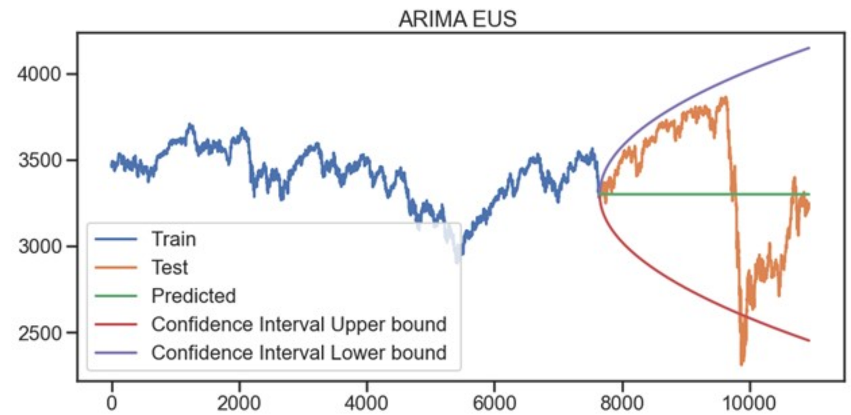

IBX
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
NIFTY
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
HSI
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
NIKKEI
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles

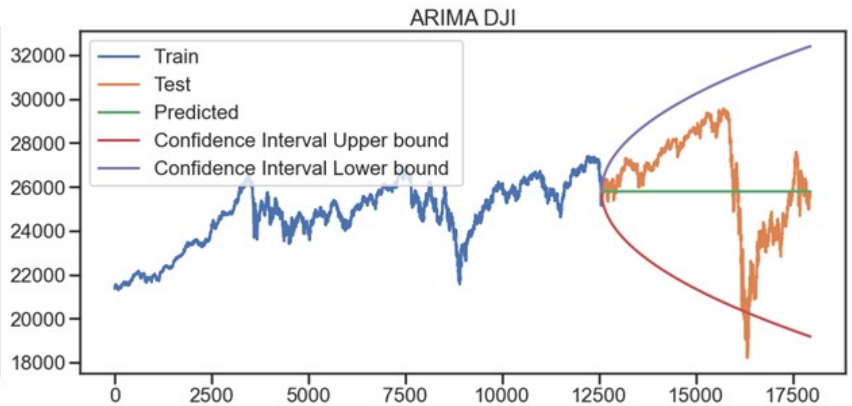

WIG
SSGF
SMI
SNP 500
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles

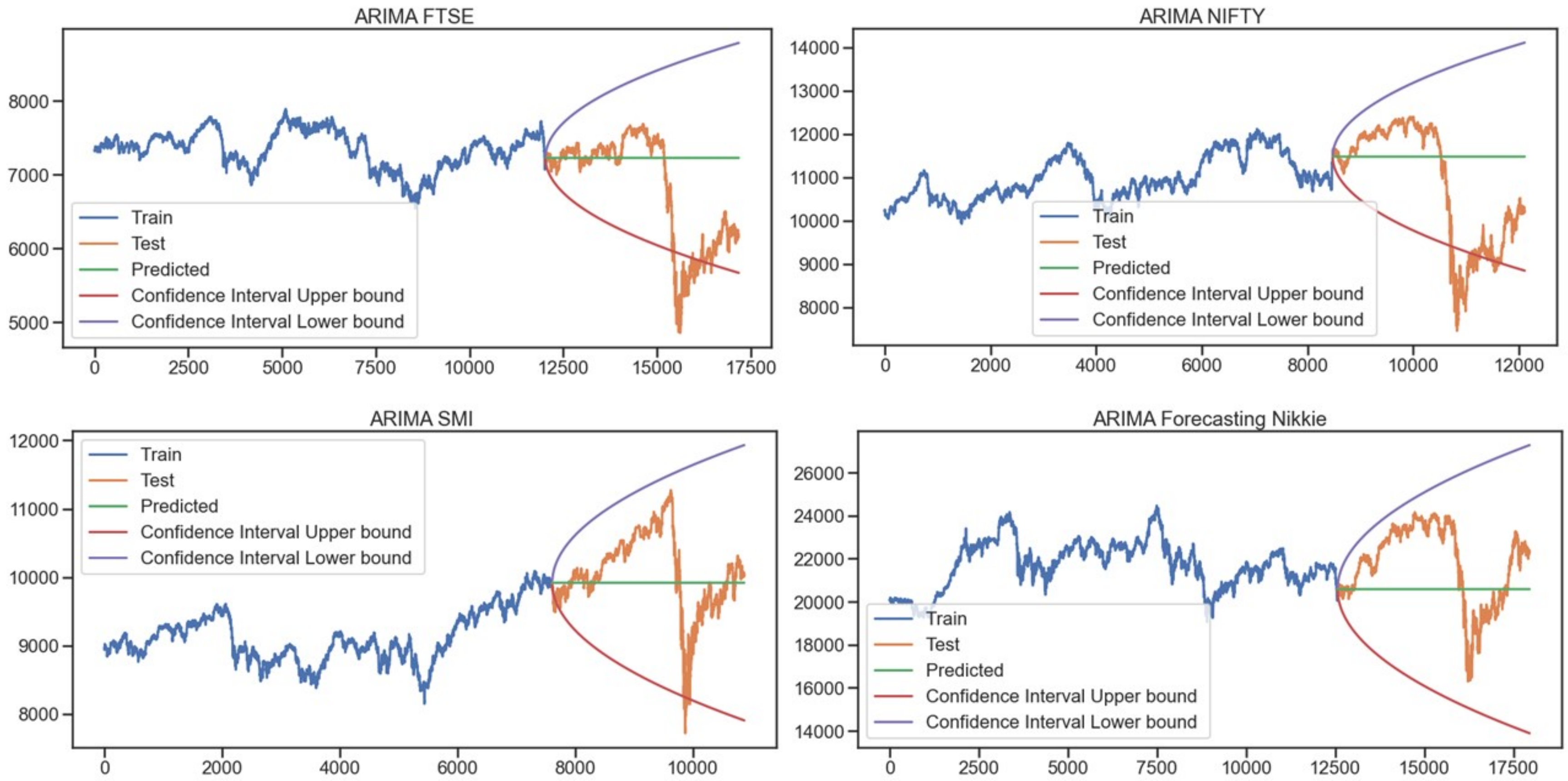
Appendix A-3: ARIMA Forecasting

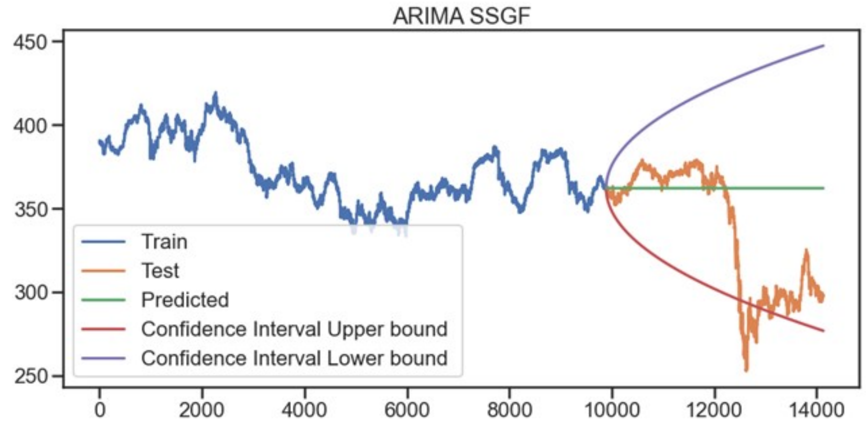


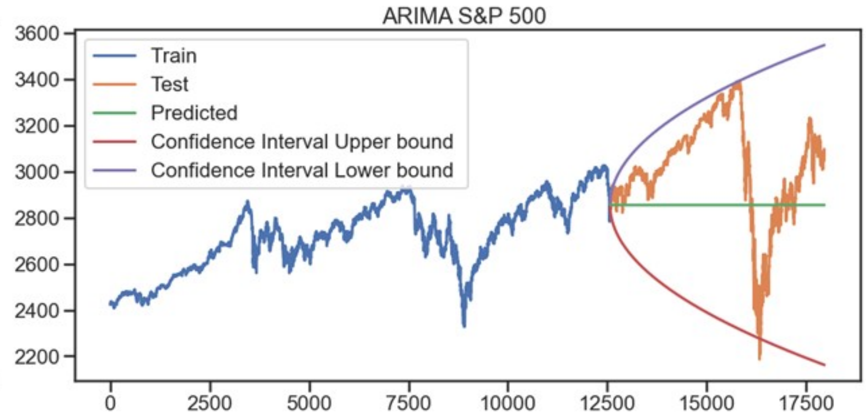


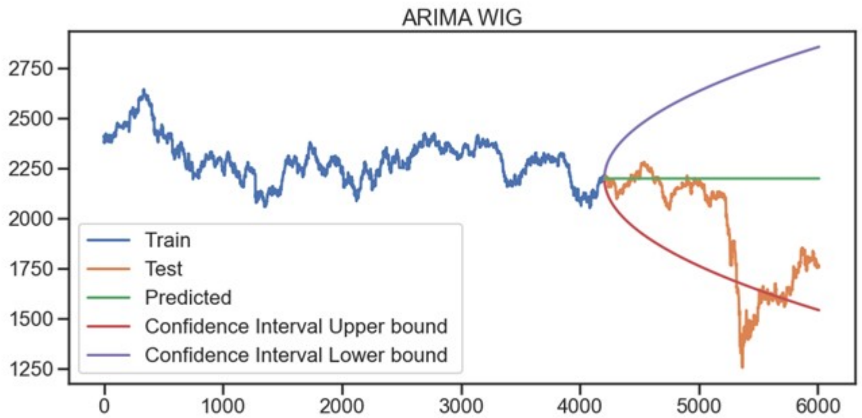


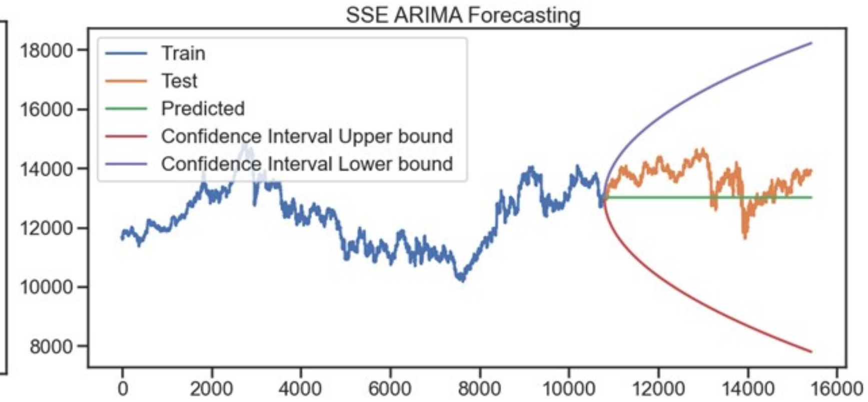


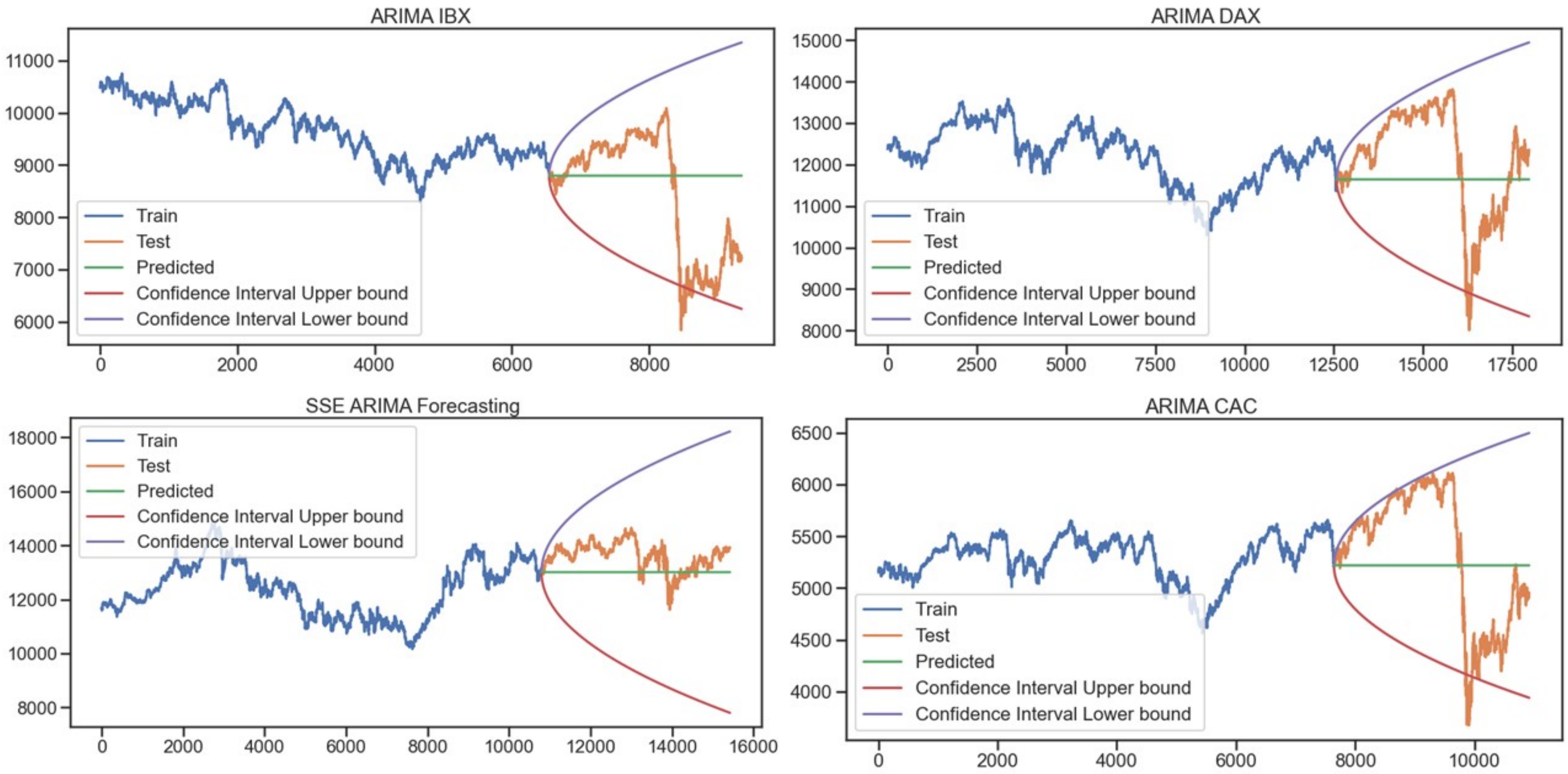

ARIMA FTSE
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
8000
7000
6000
5000
0
2500
5000
7500
10000
12500
15000
17500
ARIMA NIFTY
14000
13000
12000
11000
10000
9000
8000
0
2000
4000
6000
8000
10000
12000
ARIMA SMI
12000
11000
10000
9000
8000
ARIMA Forecasting Nikkie
26000
24000
22000
20000
18000
16000
14000


ARIMA SSGF
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
450
400
350
300
250
0
2000
4000
6000
8000
10000
12000
14000


ARIMA S&P 500
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
3600
3400
3200
3000
2800
2600
2400
2200
0
2500
5000
7500
10000
12500
15000
17500

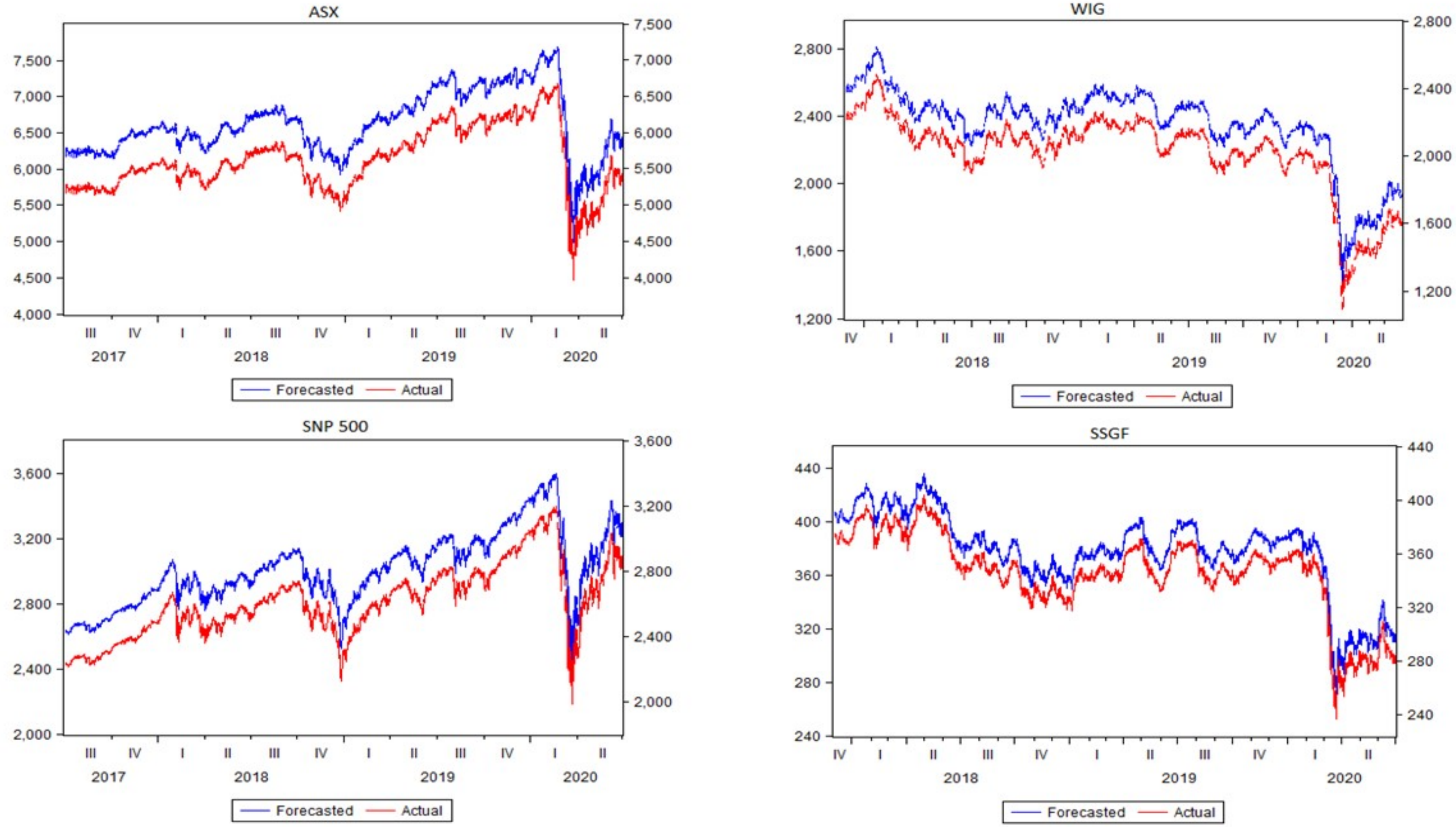

ARIMA WIG
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
2750
2500
2250
2000
1750
1500
1250
0
1000
2000
3000
4000
5000
6000

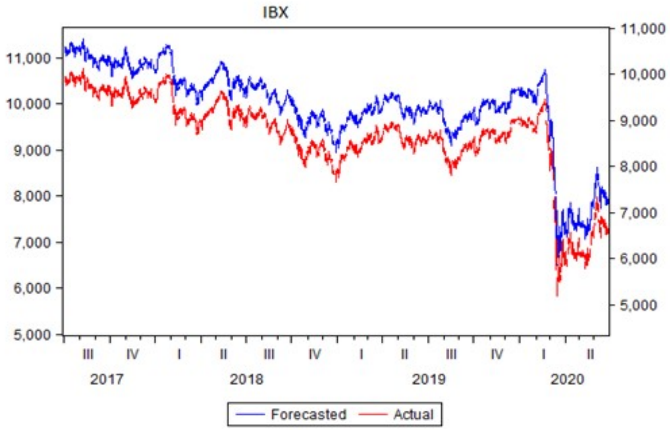

SSE ARIMA Forecasting
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
18000
16000
14000
12000
10000
8000
0
2000
4000
6000
8000
10000
12000
14000
16000

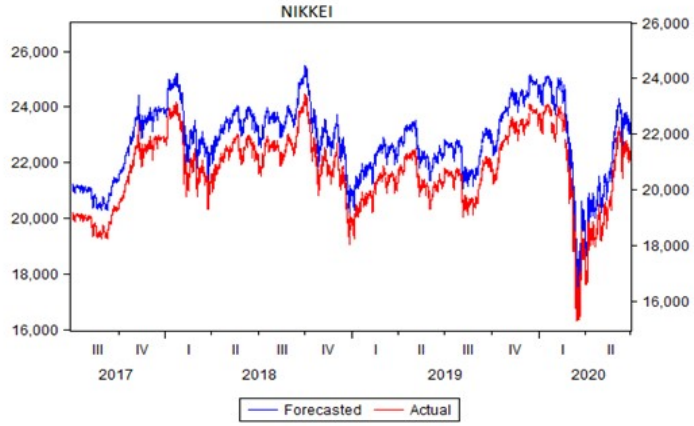

ARIMA IBX
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
11000
10000
9000
8000
7000
6000
0
2000
4000
6000
8000
ARIMA DAX
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
15000
14000
13000
12000
11000
10000
9000
8000
0
2500
5000
7500
10000
12500
15000
17500
SSE ARIMA Forecasting
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
18000
16000
14000
12000
10000
8000
0
2000
4000
6000
8000
10000
12000
14000
16000
ARIMA CAC
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
6500
6000
5500
5000
4500
4000
0
2000
4000
6000
8000
10000

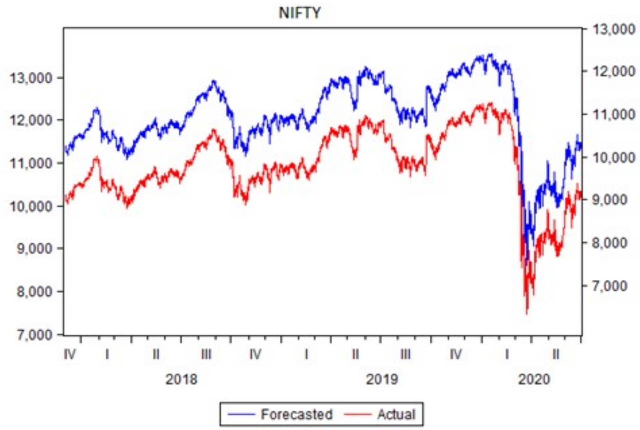
Appendix: A-4: ARFIMA Forecasting

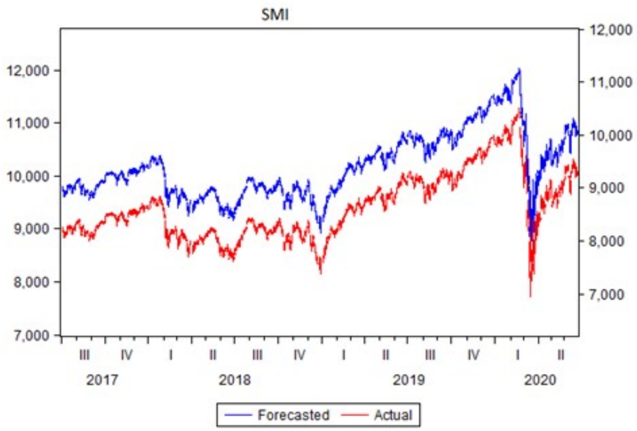
ARFIMA Model

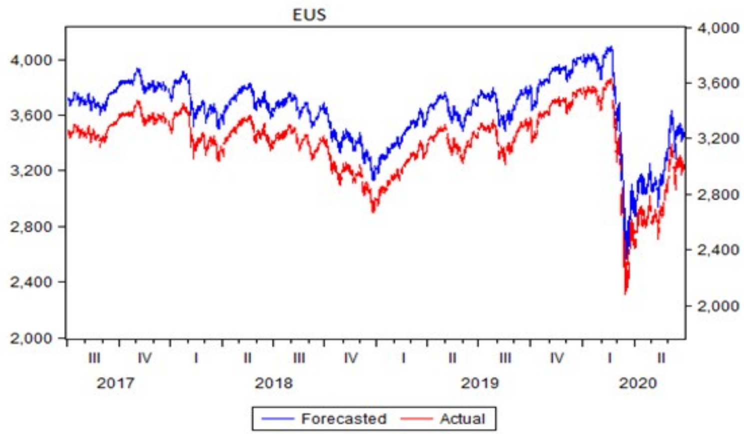


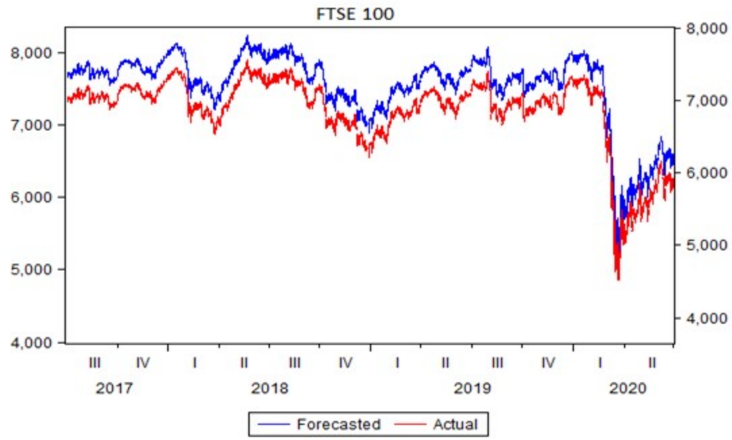

IBX
2017
2018
2019
2020
Forecasted
Actual

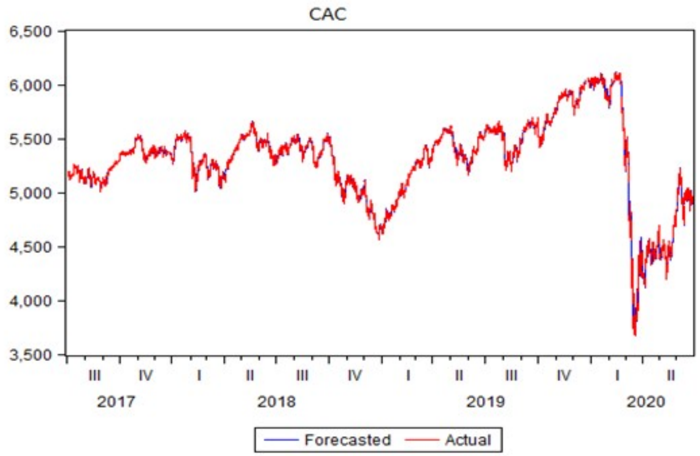

NIKKEI
2017
2018
2019
2020
Forecasted
Actual

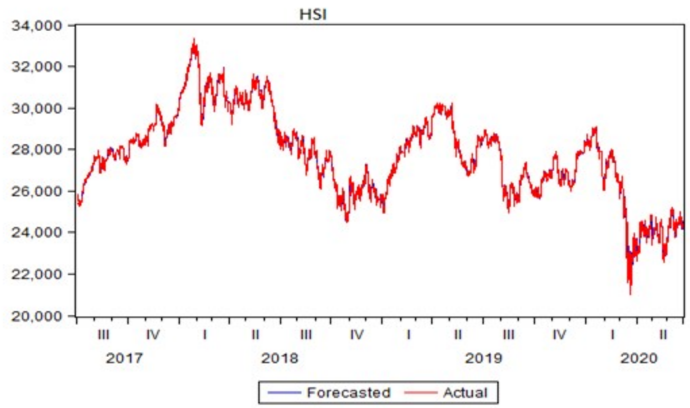

NIFTY
2018
2019
2020
Forecasted
Actual

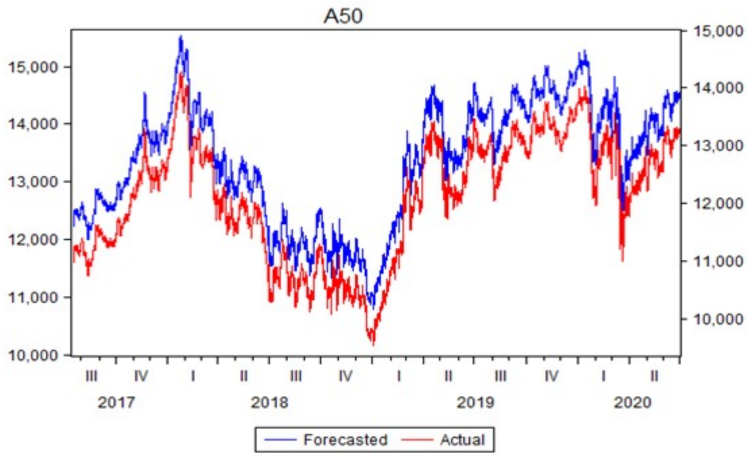

SMI
2017
2018
2019
2020
Forecasted
Actual

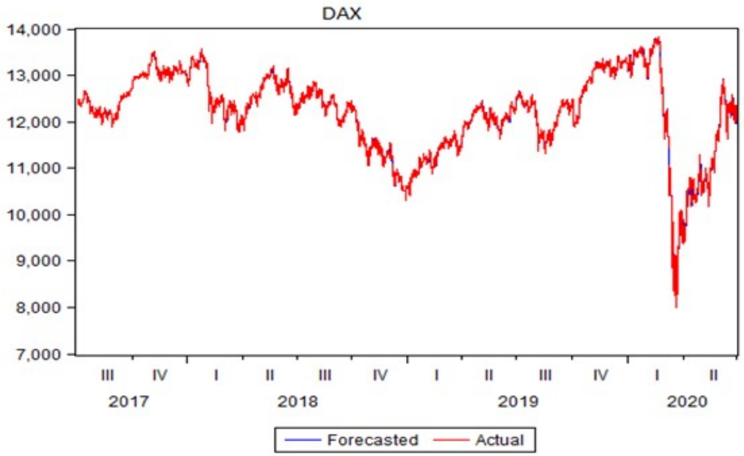

EUS
4,000
3,600
3,200
2,800
2,400
2,000
III IV I II III IV I II III IV I II
2017 2018 2019 2020
Forecasted Actual

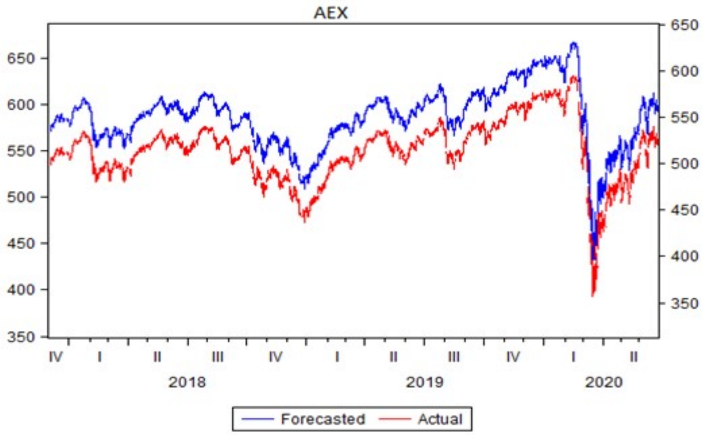

FTSE 100
8,000
7,000
6,000
5,000
4,000
III IV I II III IV I II III IV I II
2017 2018 2019 2020
Forecasted Actual

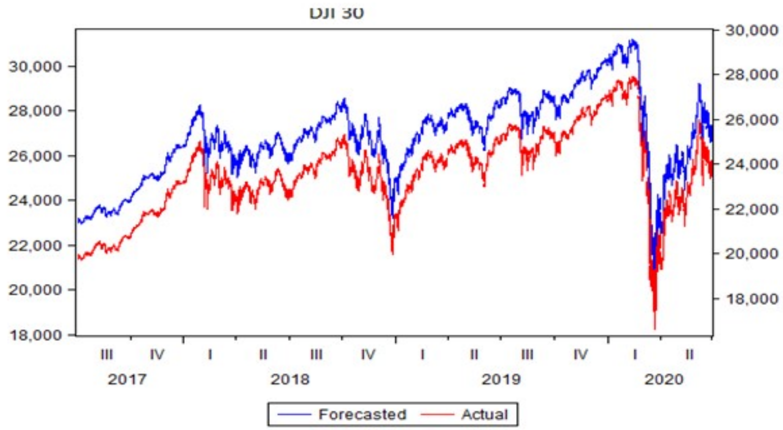

CAC
6,500
6,000
5,500
5,000
4,500
4,000
3,500
III IV I II III IV I II III IV I II
2017 2018 2019 2020
Forecasted Actual


HSI
34,000
32,000
30,000
28,000
26,000
24,000
22,000
20,000
III IV I II III IV I II III IV I II
2017 2018 2019 2020
Forecasted Actual

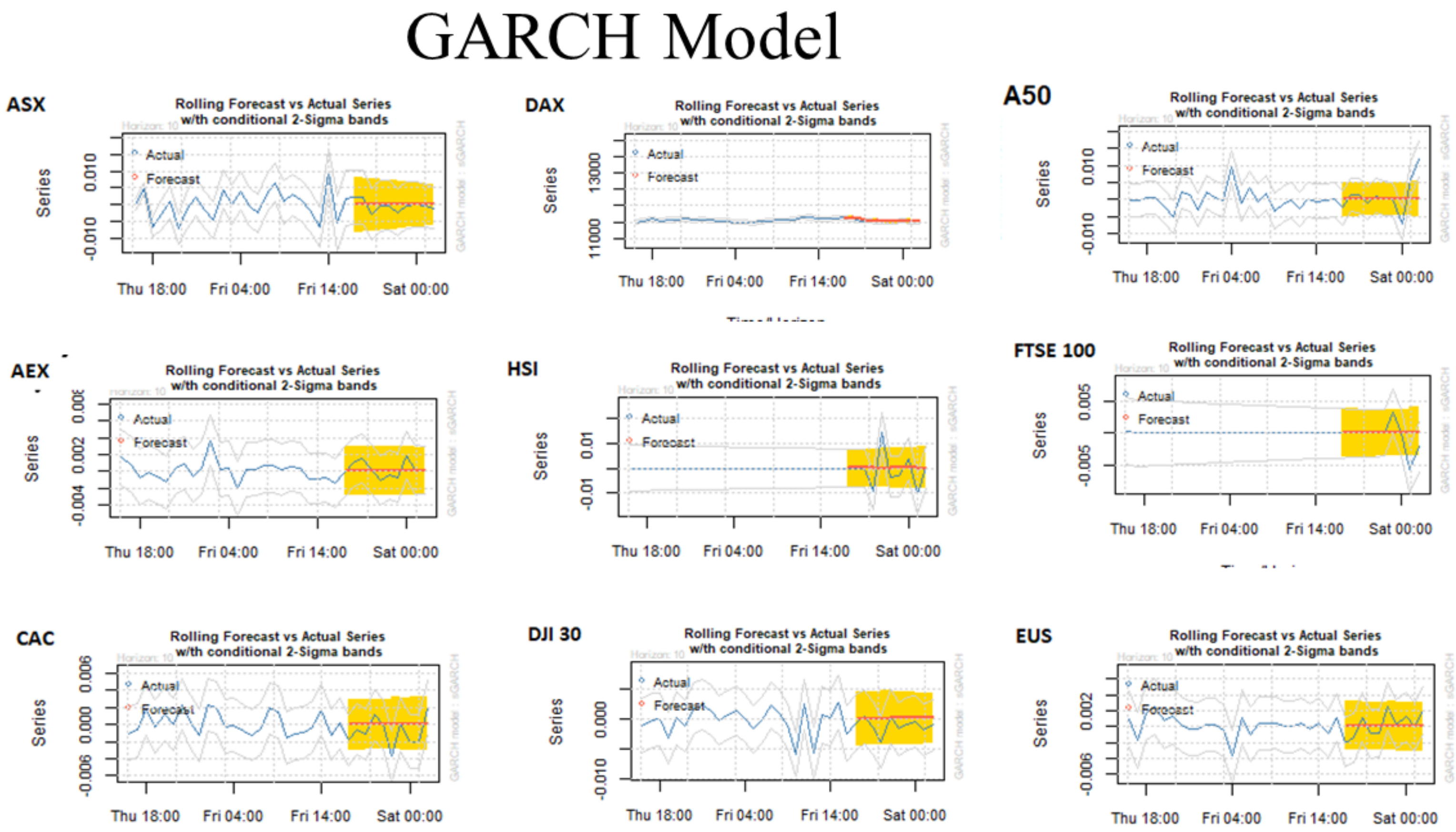

A50
15,000
14,000
13,000
12,000
11,000
10,000
15,000
14,000
13,000
12,000
11,000
10,000
III IV I II III IV I II III IV I II
2017 2018 2019 2020
Forecasted
Actual


DAX
14,000
13,000
12,000
11,000
10,000
9,000
8,000
7,000
III IV I II III IV I II III IV I II
2017 2018 2019 2020
Forecasted
Actual

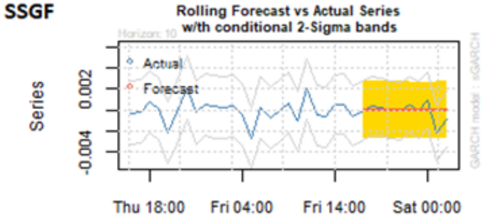

AEX
650
600
550
500
450
400
350
650
600
550
500
450
400
350
IV I II III IV I II III IV I II
2018 2019 2020
Forecasted
Actual


DJI 30
30,000
28,000
26,000
24,000
22,000
20,000
18,000
30,000
28,000
26,000
24,000
22,000
20,000
18,000
III IV I II III IV I II III IV I II
2017 2018 2019 2020
Forecasted
Actual

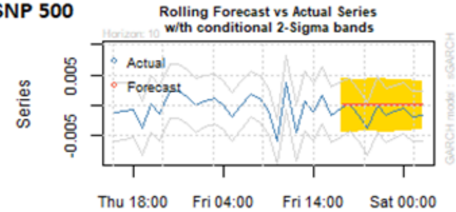
Appendix: A-5: GARCH Forecasting



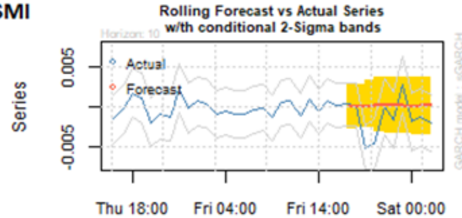
SSGF



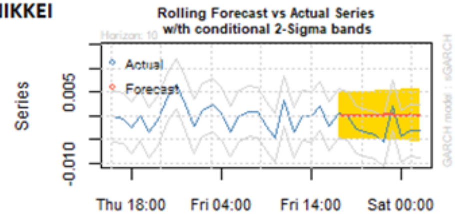
SNP 500



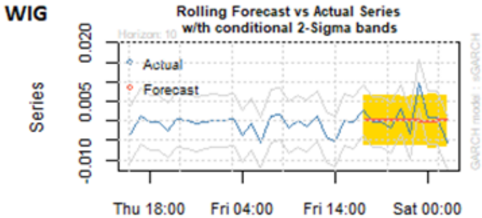
SMI



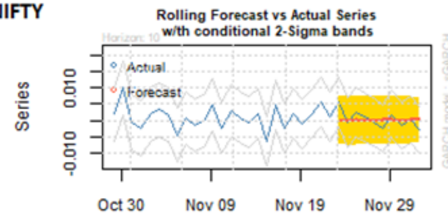
NIKKEI



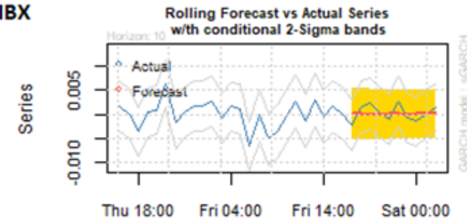
WIG



NIFTY

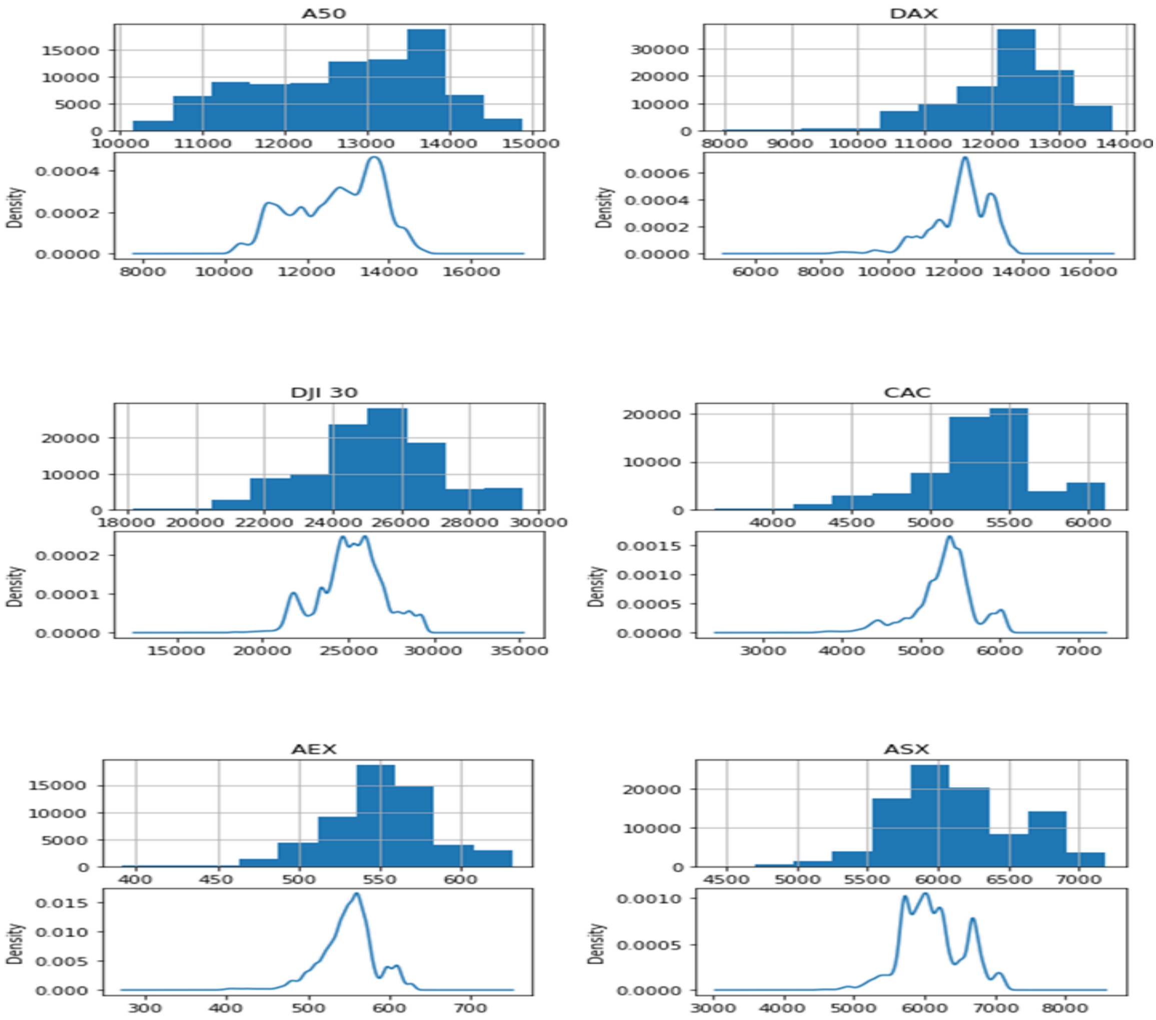


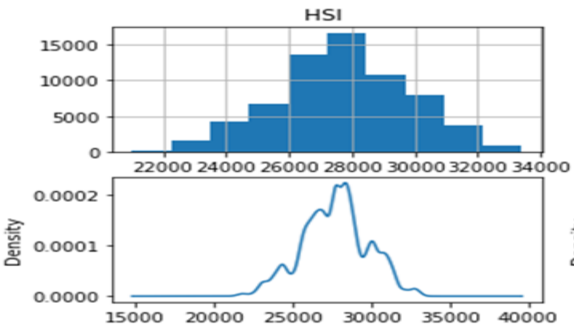
IBX

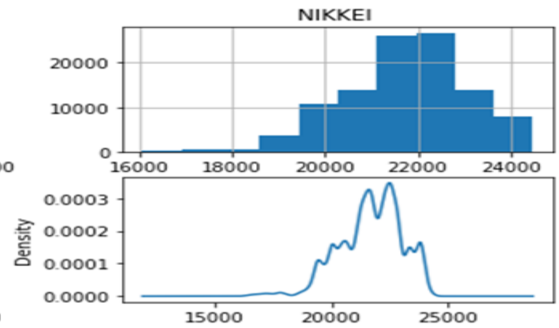


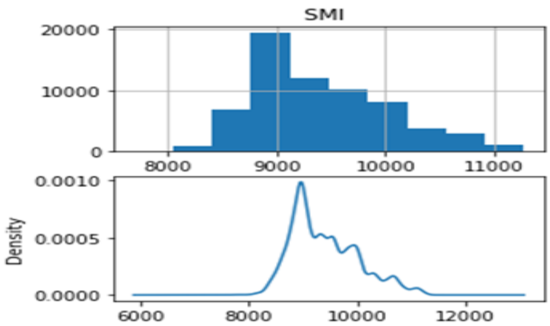
Appendix B

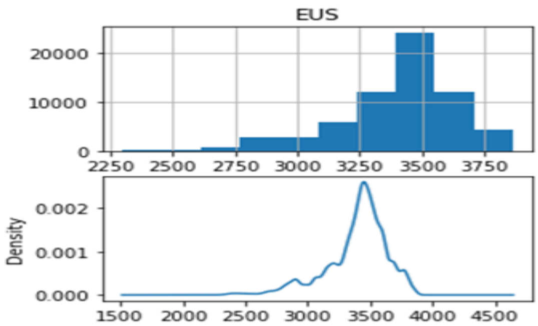
Appendix B-1: KDE

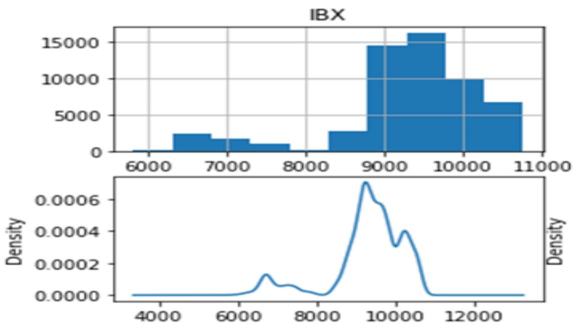


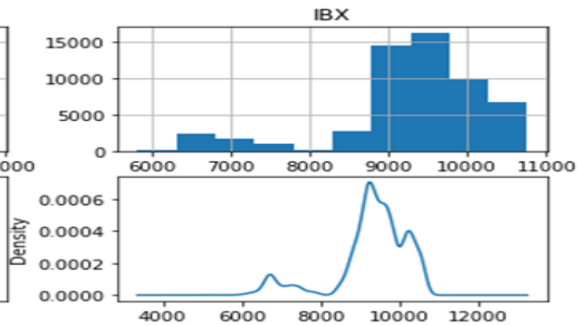

HSI
Density

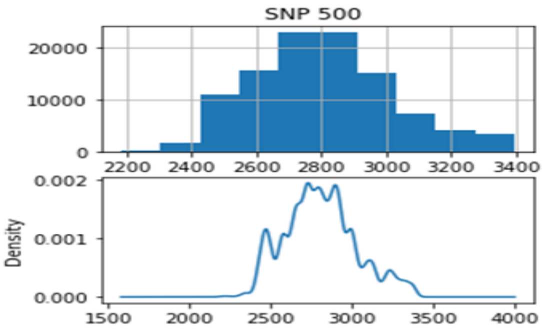

NIKKEI
Density

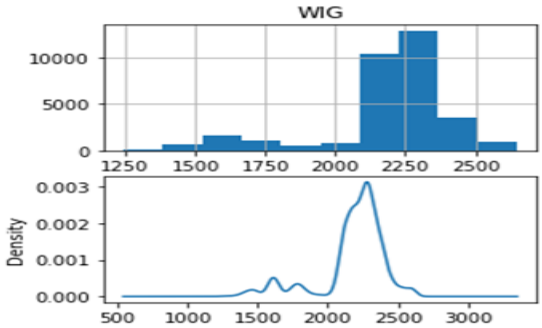

SMI
Density

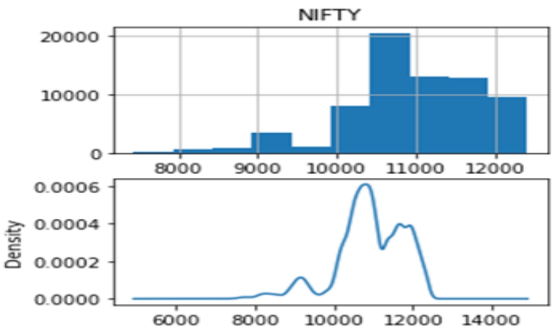

EUS
Density

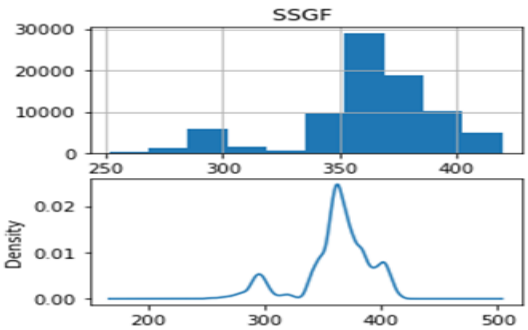

IBX
Density


IBX
Density

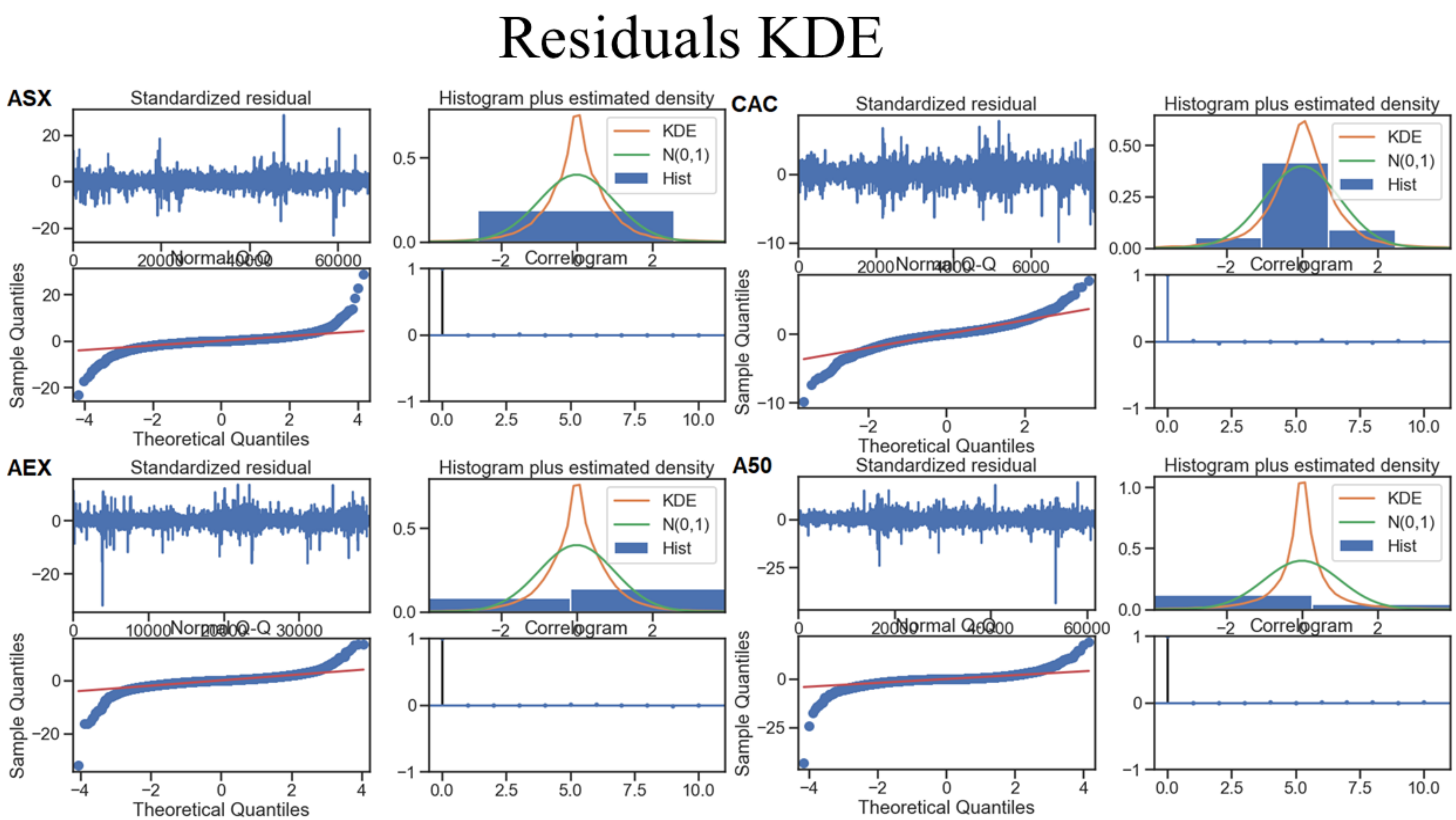

SNP 500
20000
10000
0
2200 2400 2600 2800 3000 3200 3400
0.002
0.001
0.000
Density
1500 2000 2500 3000 3500 4000

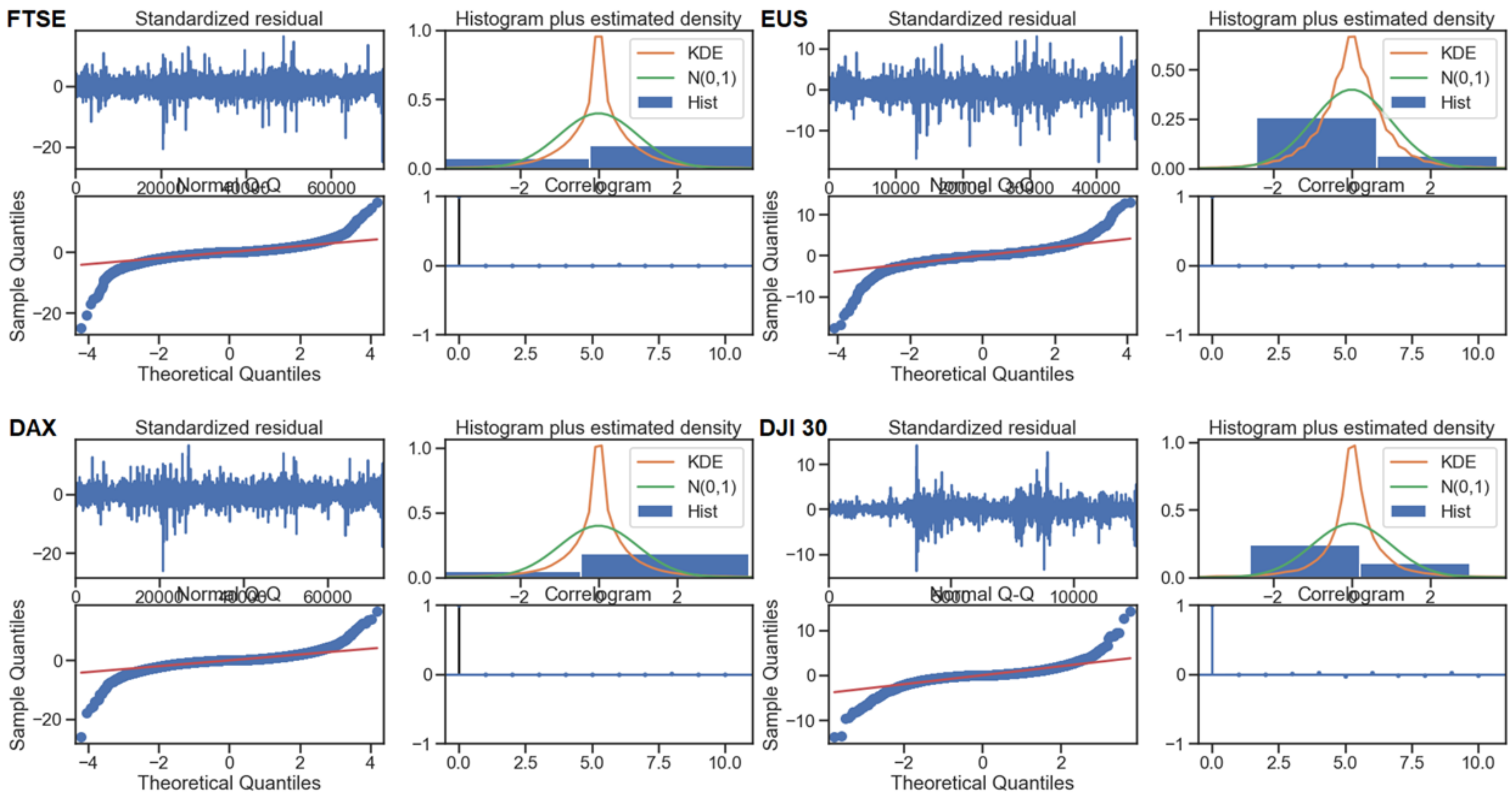

WIG
10000
5000
0
1250 1500 1750 2000 2250 2500
0.003
0.002
0.001
0.000
Density
500 1000 1500 2000 2500 3000

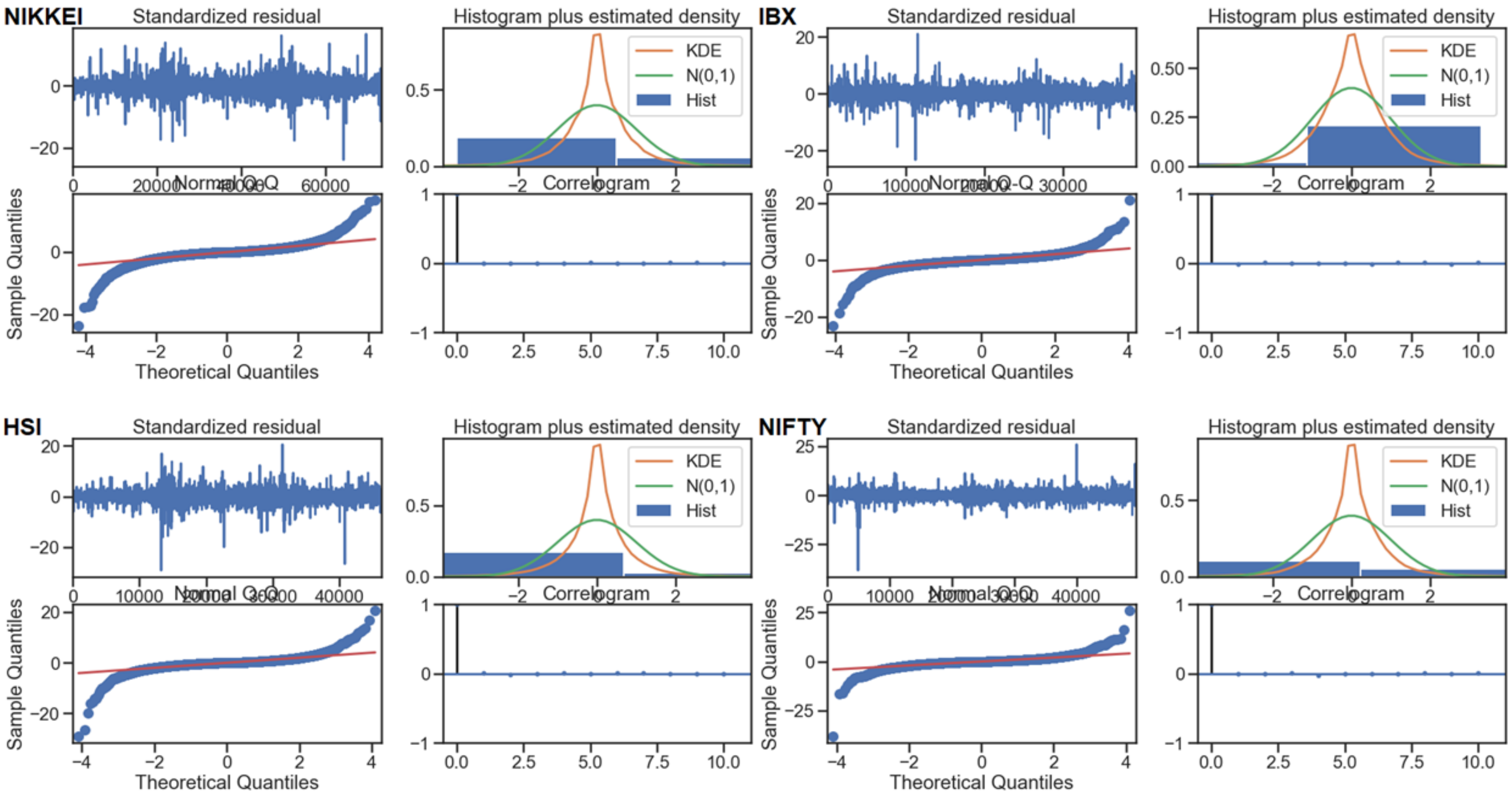

NIFTY
20000
10000
0
8000 9000 10000 11000 12000
0.0006
0.0004
0.0002
0.0000
Density
6000 8000 10000 12000 14000

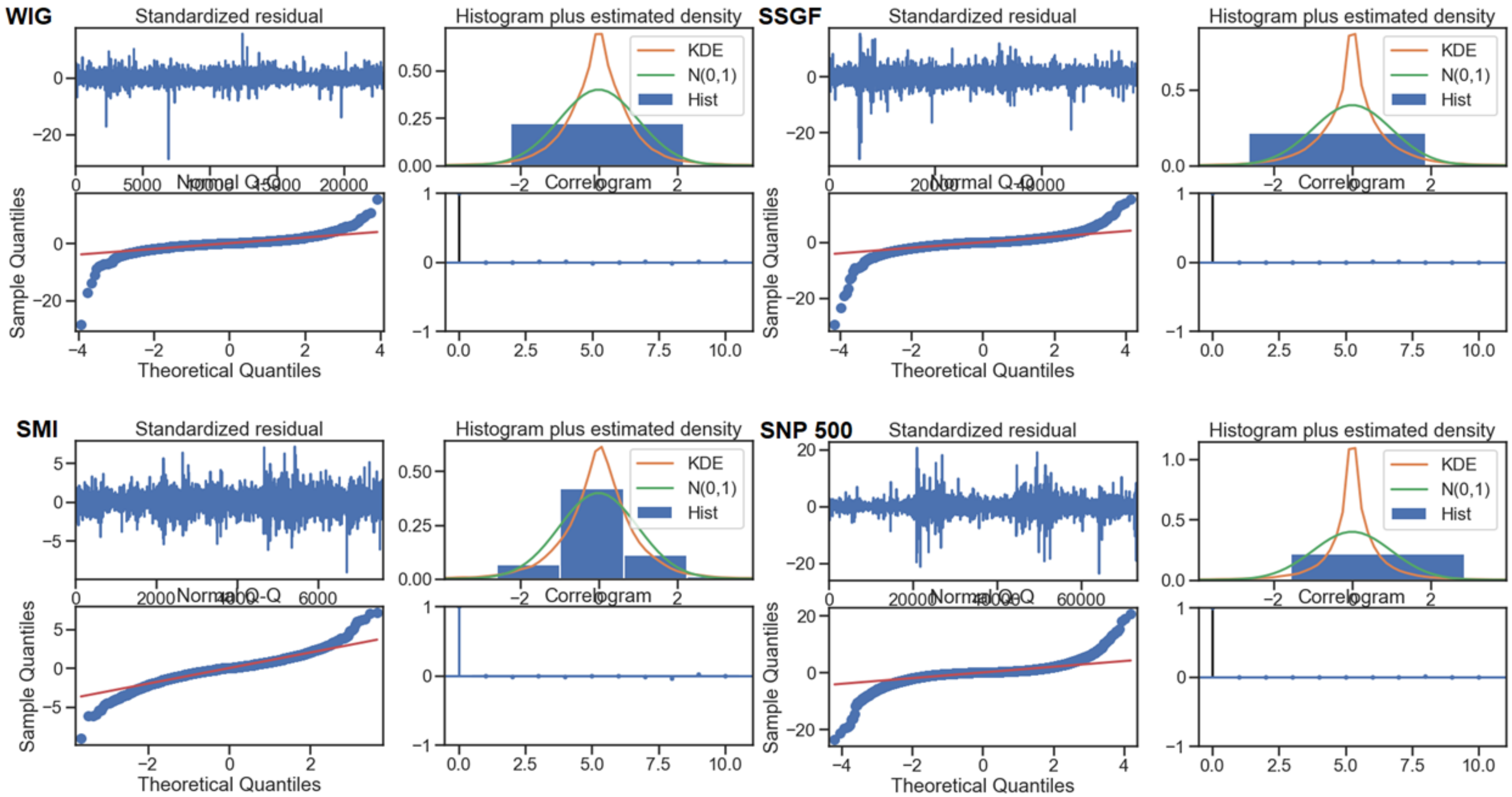

SSGF
30000
20000
10000
0
250 300 350 400
0.02
0.01
0.00
Density
200 300 400 500

Appendix B-2: KDE of ARIMA models residuals

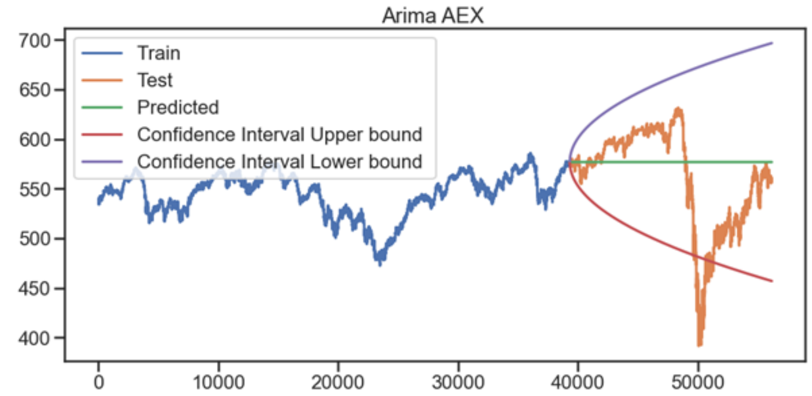


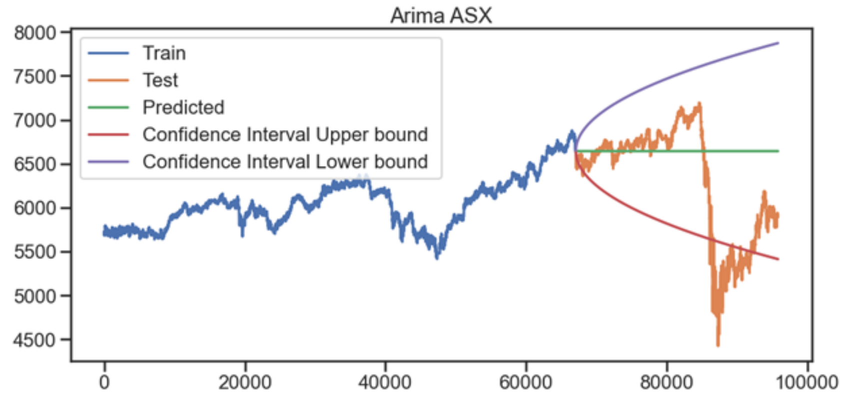

FTSE
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
EUS
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
DAX
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
DJI 30
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles

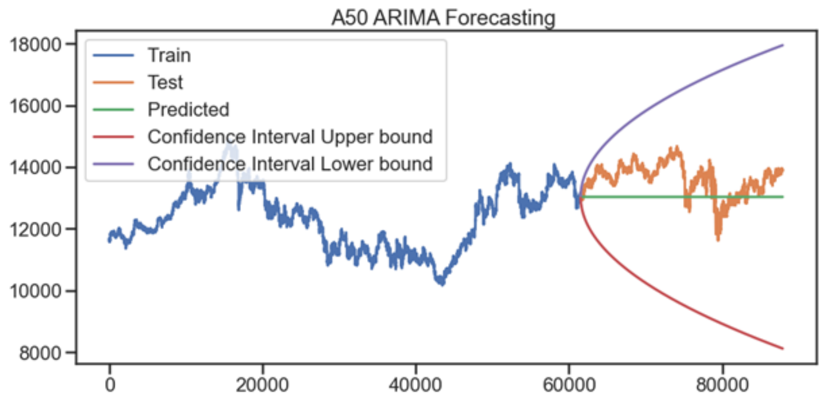

NIKKEI
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
IBX
HSI
NIFTY

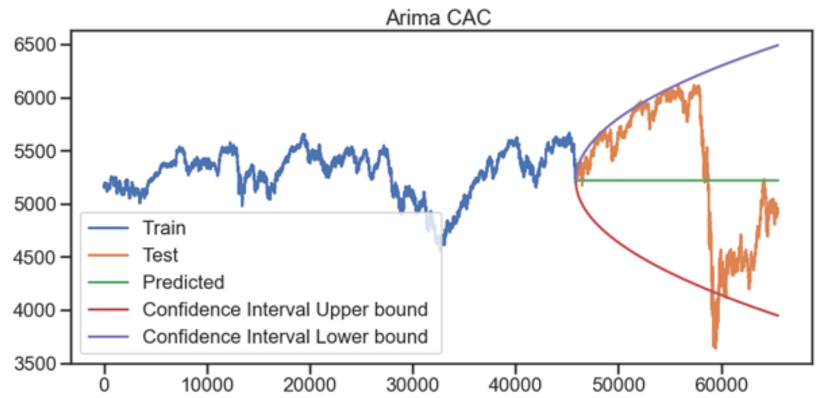

WIG
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
SSGF
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
SMI
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
SNP 500
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles

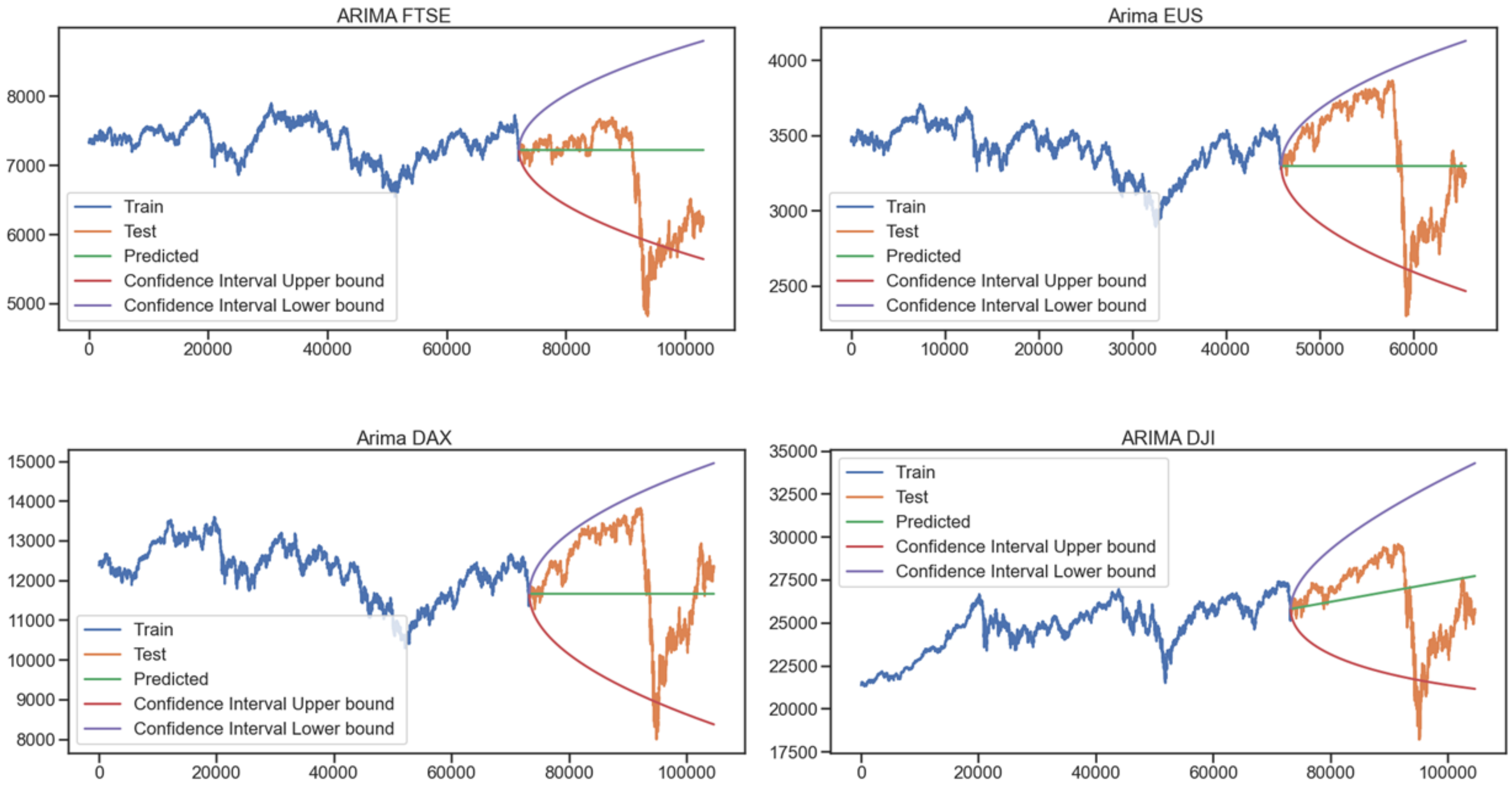
Appendix B-3: ARIMA Forecasting

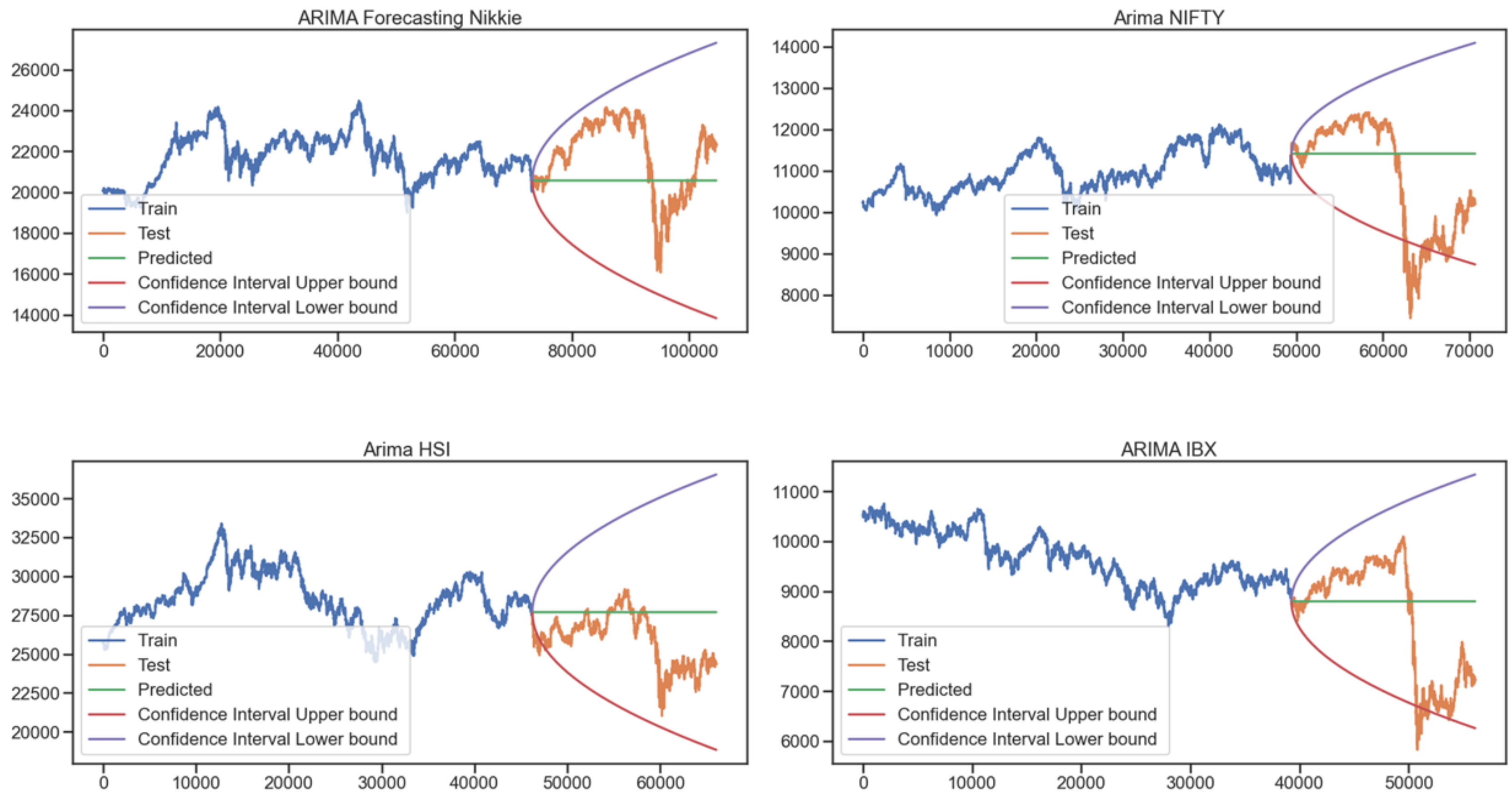


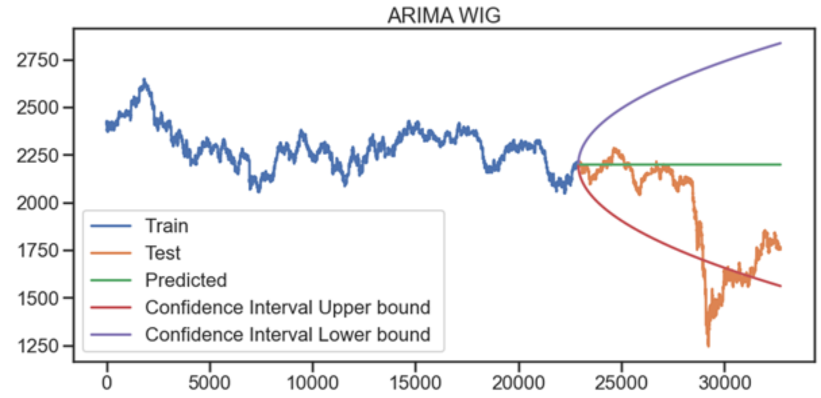


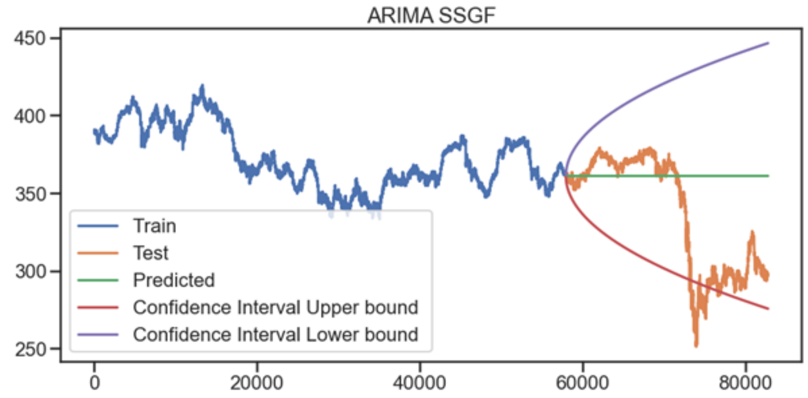


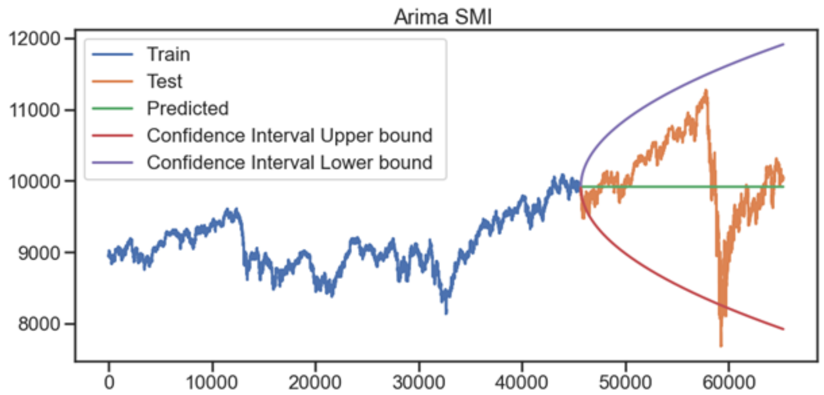


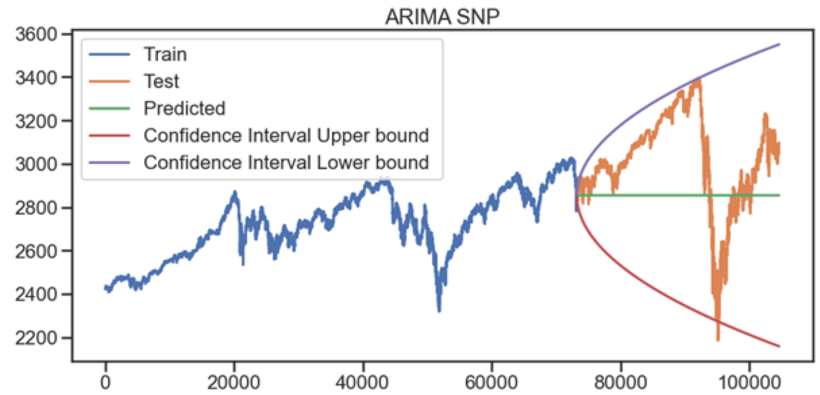

ARIMA FTSE
Arima EUS
Arima DAX
ARIMA DJI
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound


ARIMA Forecasting Nikkie
Arima NIFTY
Arima HSI
ARIMA IBX
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound

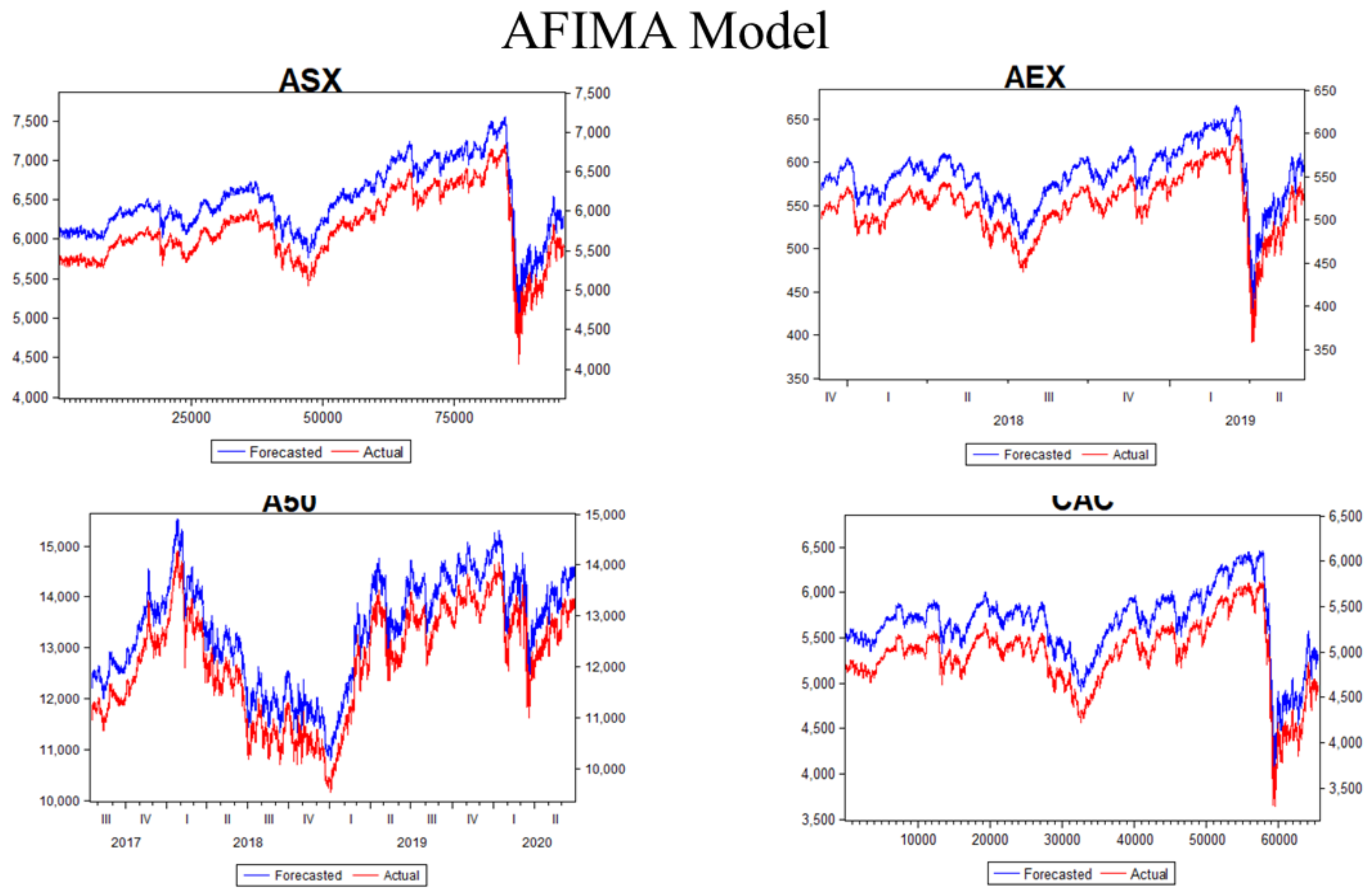

ARIMA WIG
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
2750
2500
2250
2000
1750
1500
1250
0
5000
10000
15000
20000
25000
30000


ARIMA SSGF
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
450
400
350
300
250
0
20000
40000
60000
80000

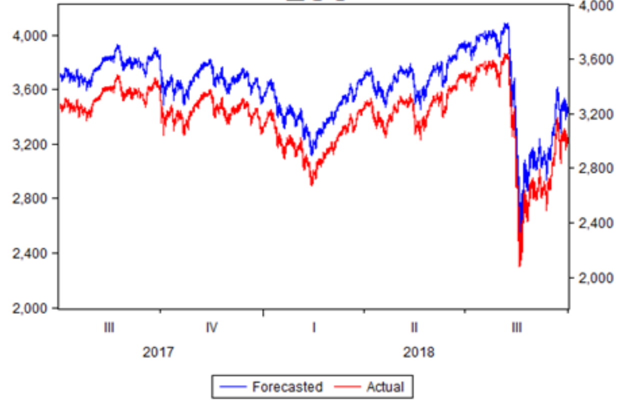

Arima SMI
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
12000
11000
10000
9000
8000
0
10000
20000
30000
40000
50000
60000


ARIMA SNP
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
3600
3400
3200
3000
2800
2600
2400
2200
0
20000
40000
60000
80000
100000

Appendix: B-4: ARFIMA Forecasting

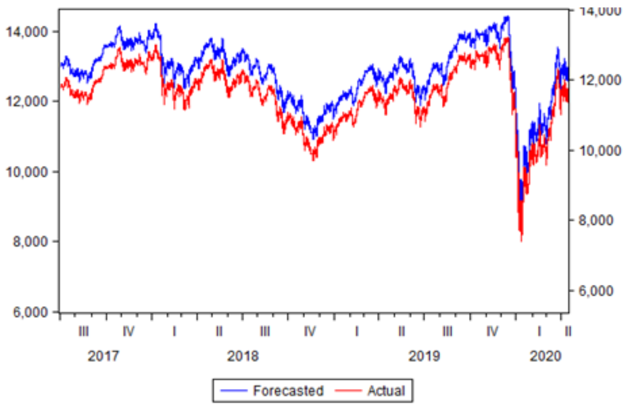


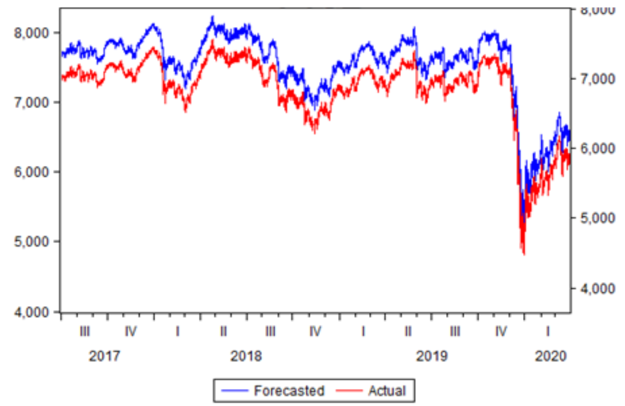
EUS

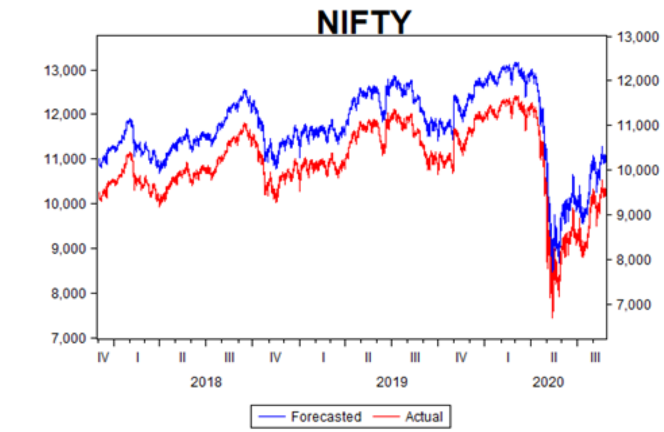

4,000
3,600
3,200
2,800
2,400
2,000
III
IV
I
II
III
2017
2018
Forecasted
Actual

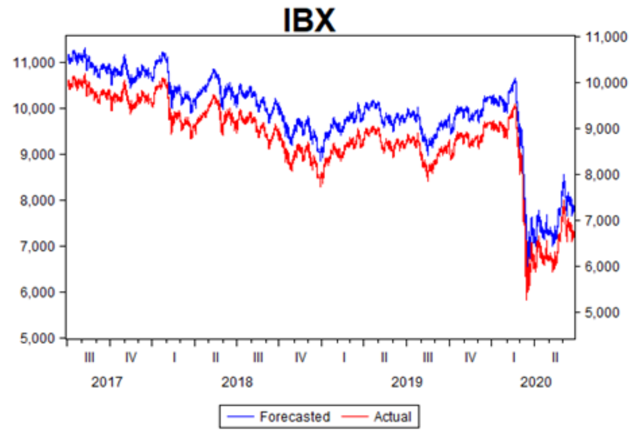
DJI 30

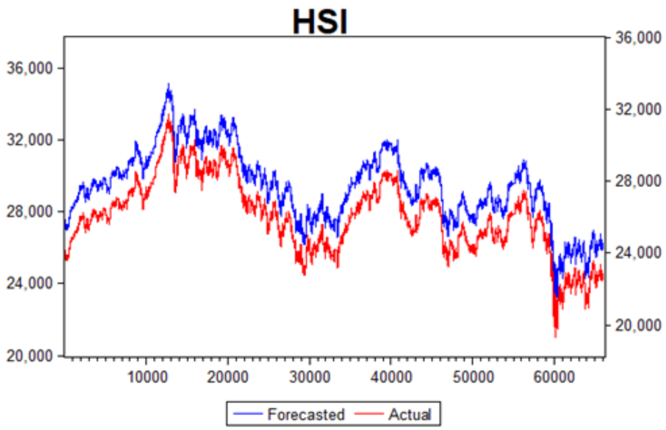

30,000
28,000
26,000
24,000
22,000
20,000
18,000
2017
2018
2019
2020
Forecasted
Actual

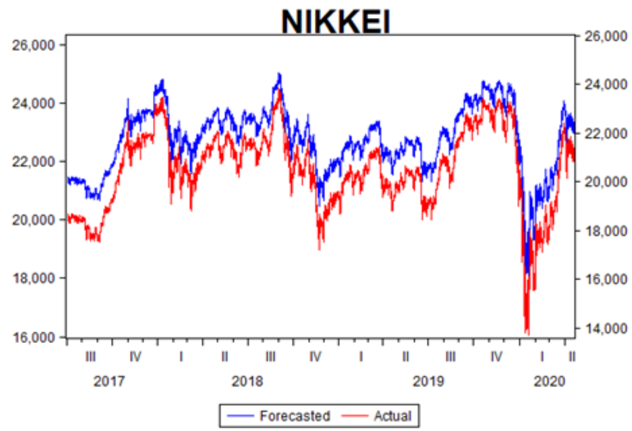

14,000
12,000
10,000
8,000
6,000
2017
2018
2019
2020
Forecasted
Actual

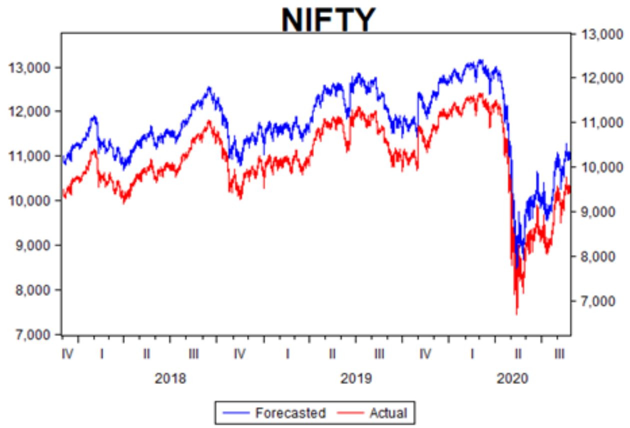

8,000
7,000
6,000
5,000
4,000
2017
2018
2019
2020
Forecasted
Actual

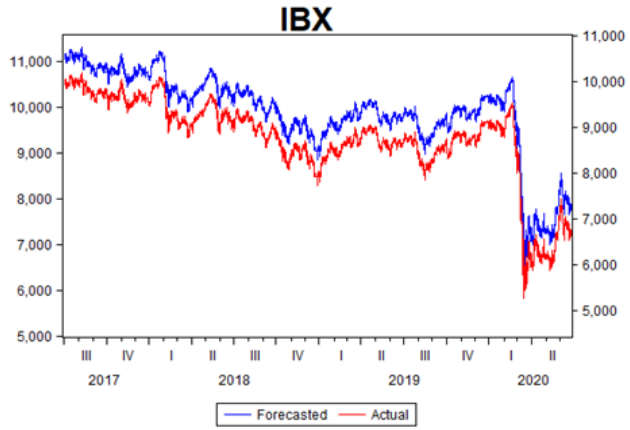

NIFTY
13,000
12,000
11,000
10,000
9,000
8,000
7,000
IV I II III IV I II III IV I II III
2018
2019
2020
Forecasted
Actual

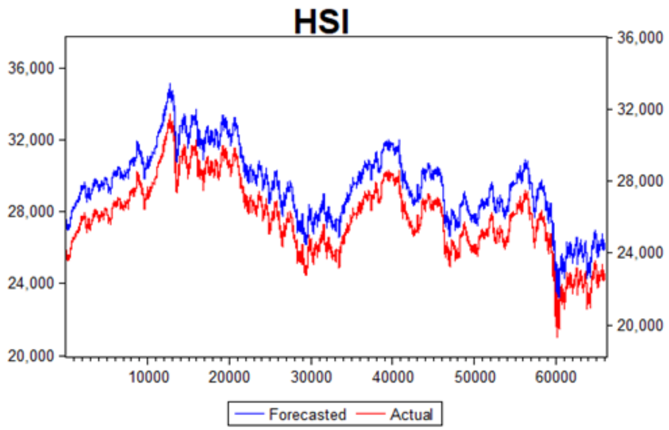

IBX
11,000
10,000
9,000
8,000
7,000
6,000
5,000
III IV I II III IV I II III IV I II
2017
2018
2019
2020
Forecasted
Actual

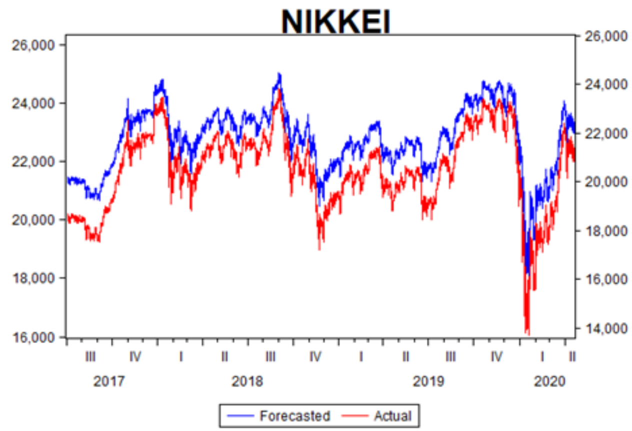

HSI
36,000
32,000
28,000
24,000
20,000
10000
20000
30000
40000
50000
60000
Forecasted
Actual


NIKKEI
26,000
24,000
22,000
20,000
18,000
16,000
14,000
III IV I II III IV I II III IV I II
2017
2018
2019
2020
Forecasted
Actual


NIFTY
13,000
12,000
11,000
10,000
9,000
8,000
7,000
IV I II III IV I II III IV I II III
2018 2019 2020
Forecasted Actual

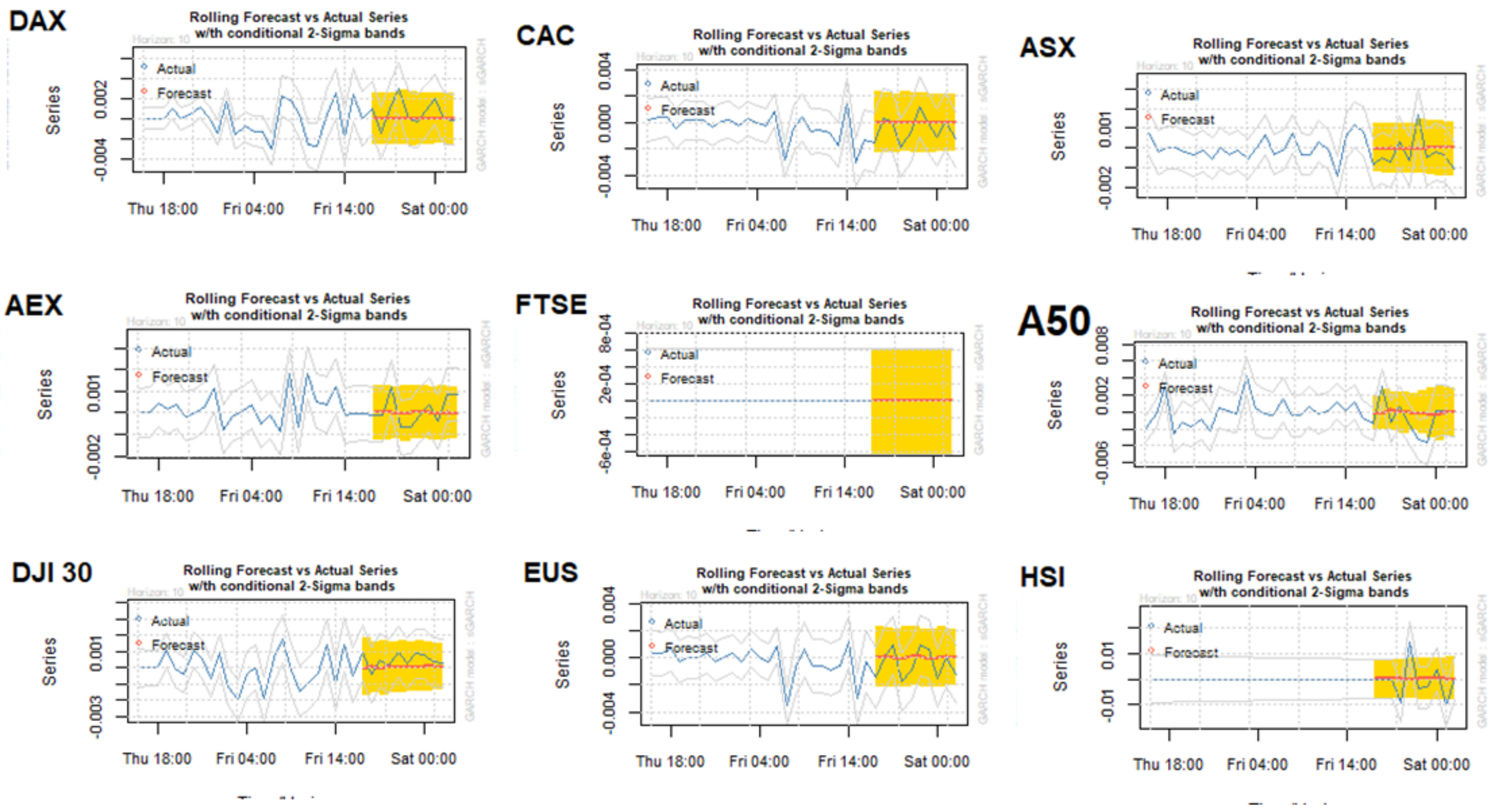

IBX
11,000
10,000
9,000
8,000
7,000
6,000
5,000
III IV I II III IV I II III IV I II
2017 2018 2019 2020
Forecasted Actual


HSI
36,000
32,000
28,000
24,000
20,000
10000 20000 30000 40000 50000 60000
Forecasted Actual

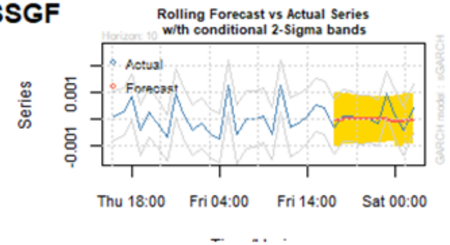

NIKKEI
26,000
24,000
22,000
20,000
18,000
16,000
14,000
III IV I II III IV I II III IV I II
2017 2018 2019 2020
Forecasted Actual

Appendix: B-5: GARCH Forecasting

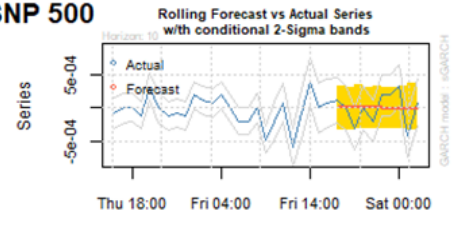
GARCH Model



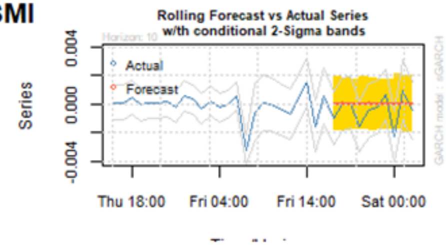
SSGF



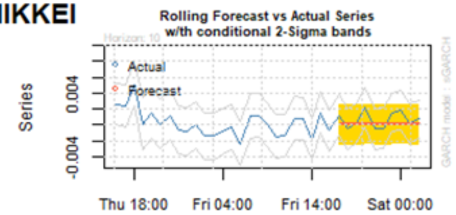
SNP 500



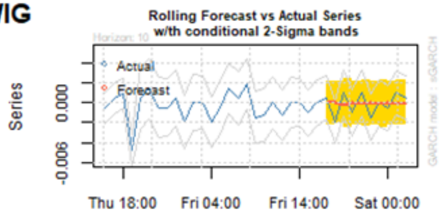
SMI



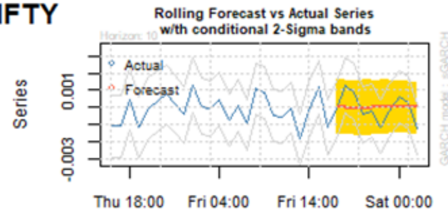
NIKKEI



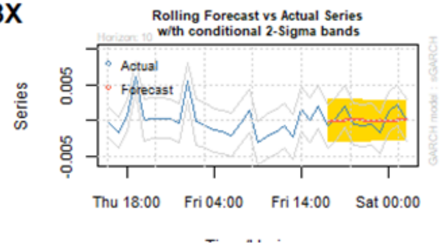
WIG



NIFTY

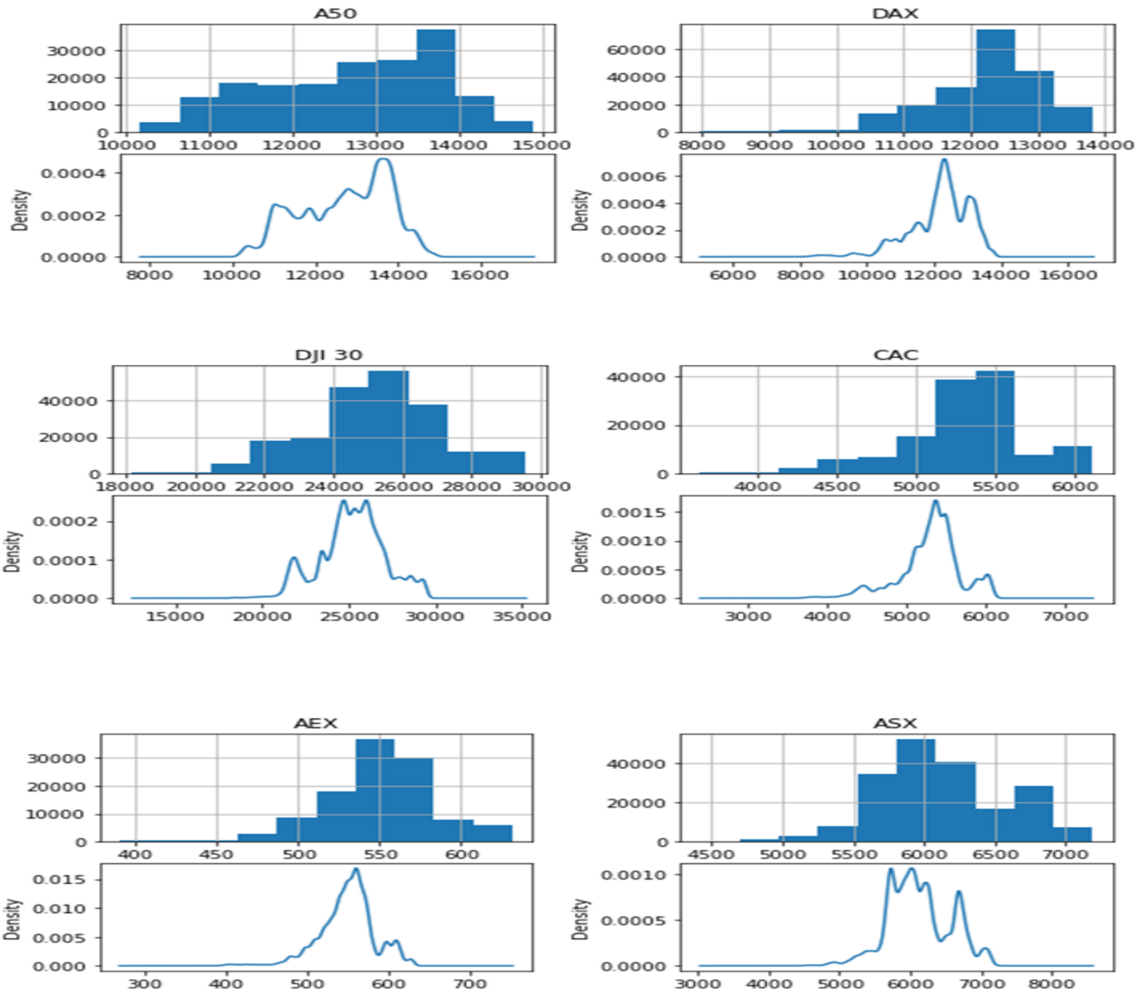


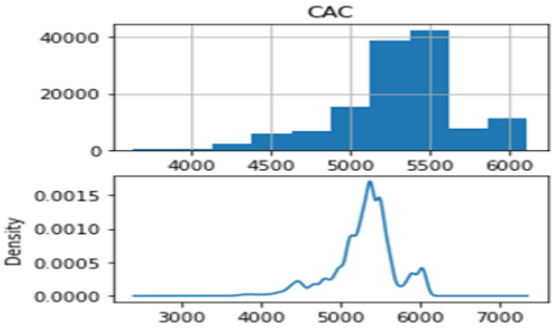
IBX

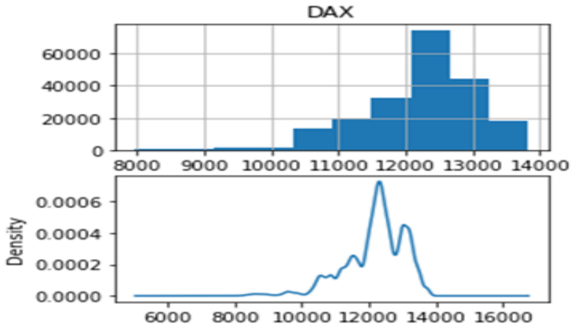


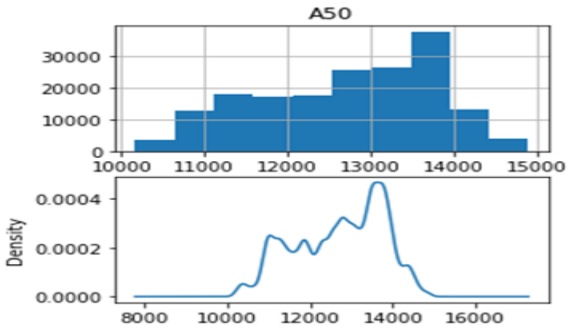
Appendix C

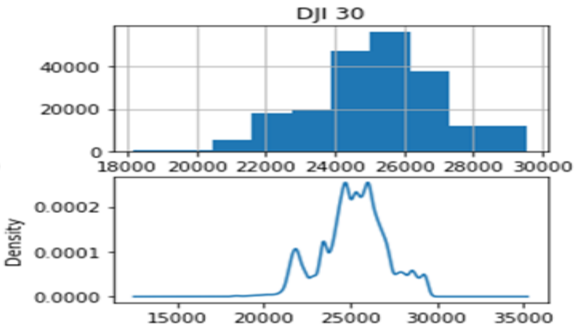
Appendix C-1: KDE

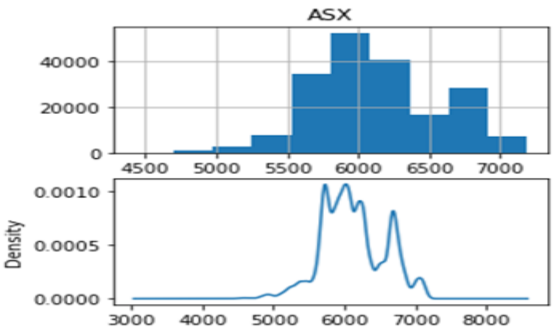


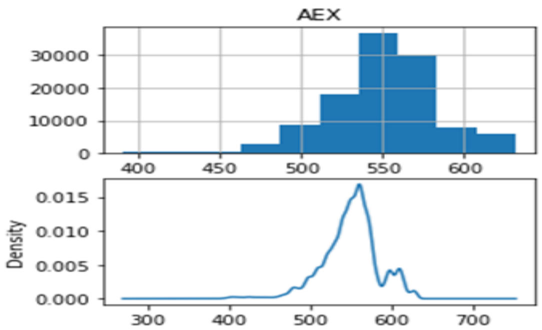

CAC
Density


DAX
Density

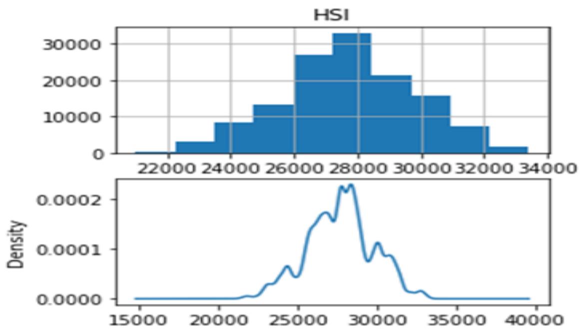

A50
Density

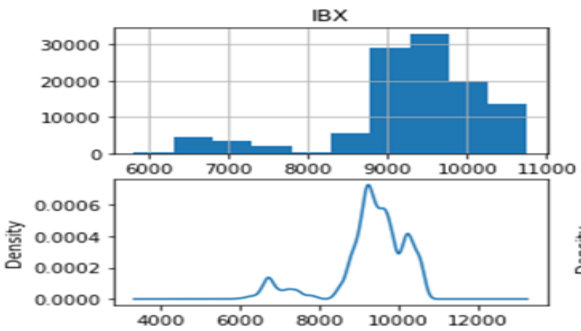

DJI 30
Density

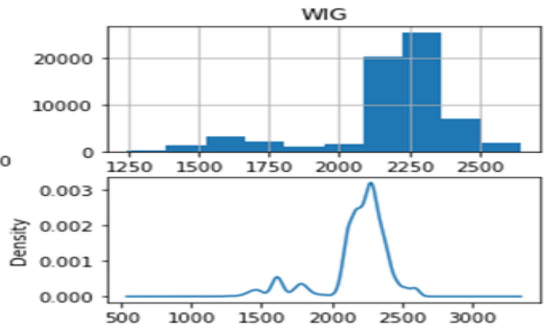

ASX
Density


AEX
Density

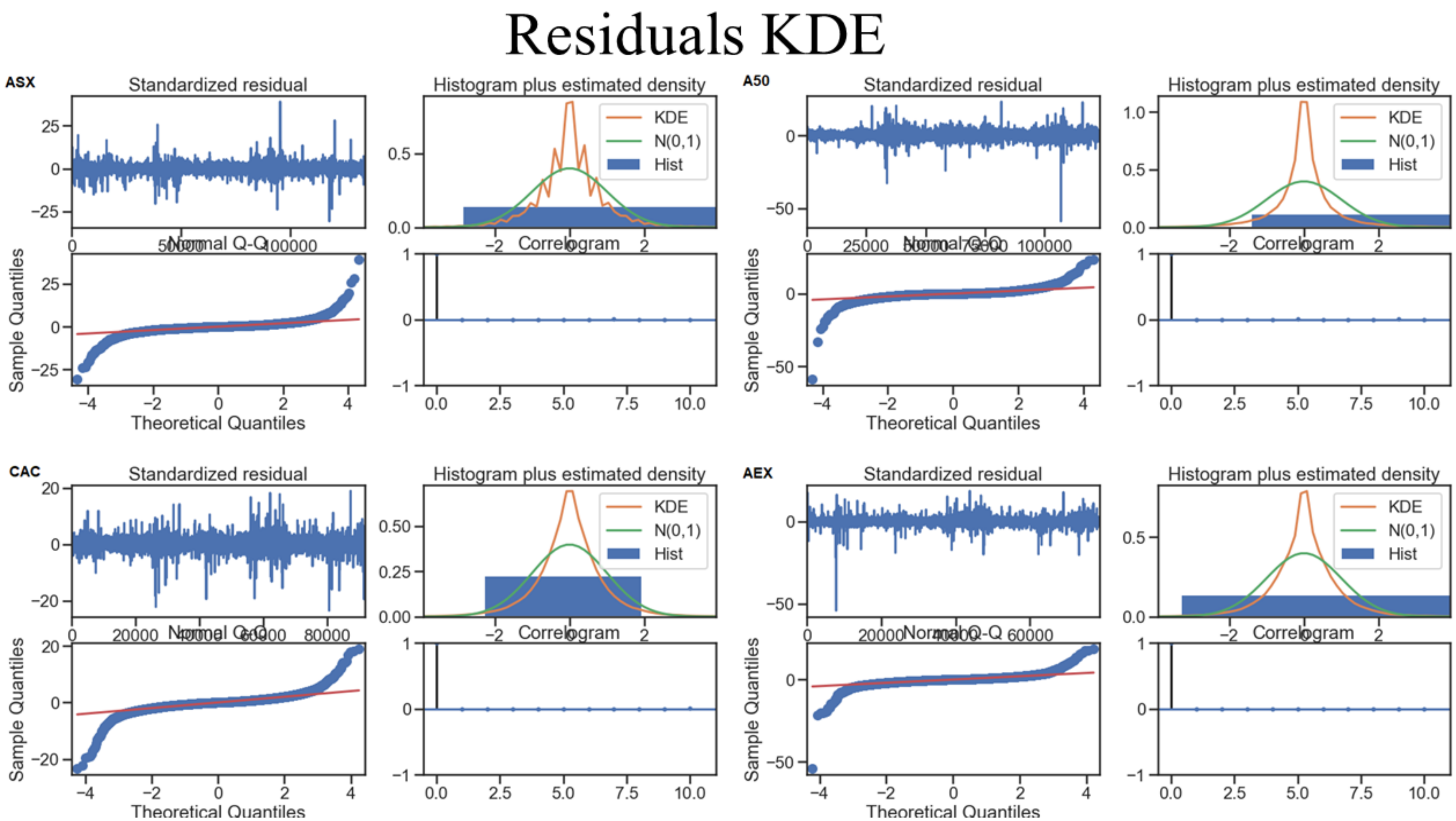

SNP
Density

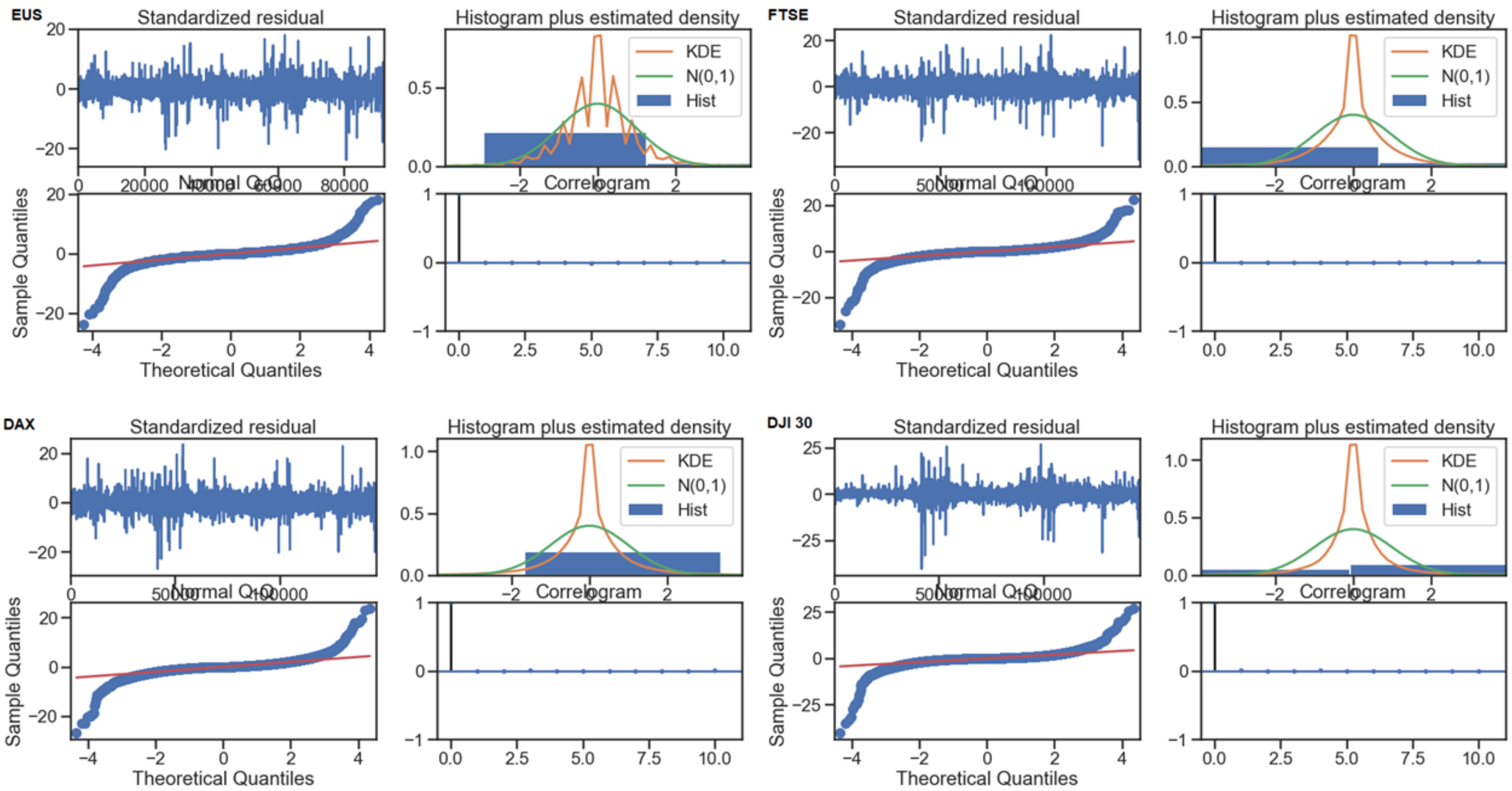

HSI
Density

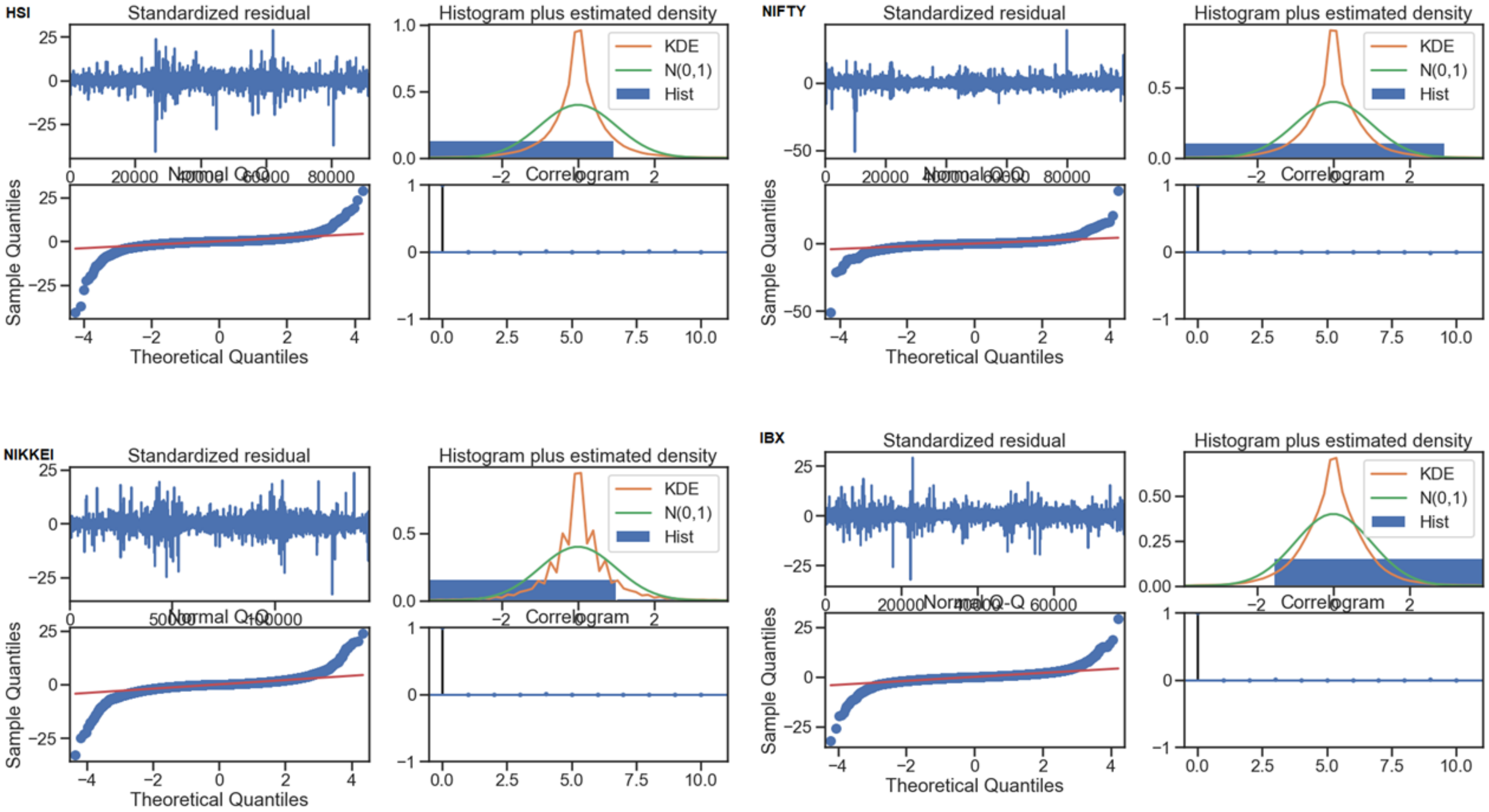

IBX
Density

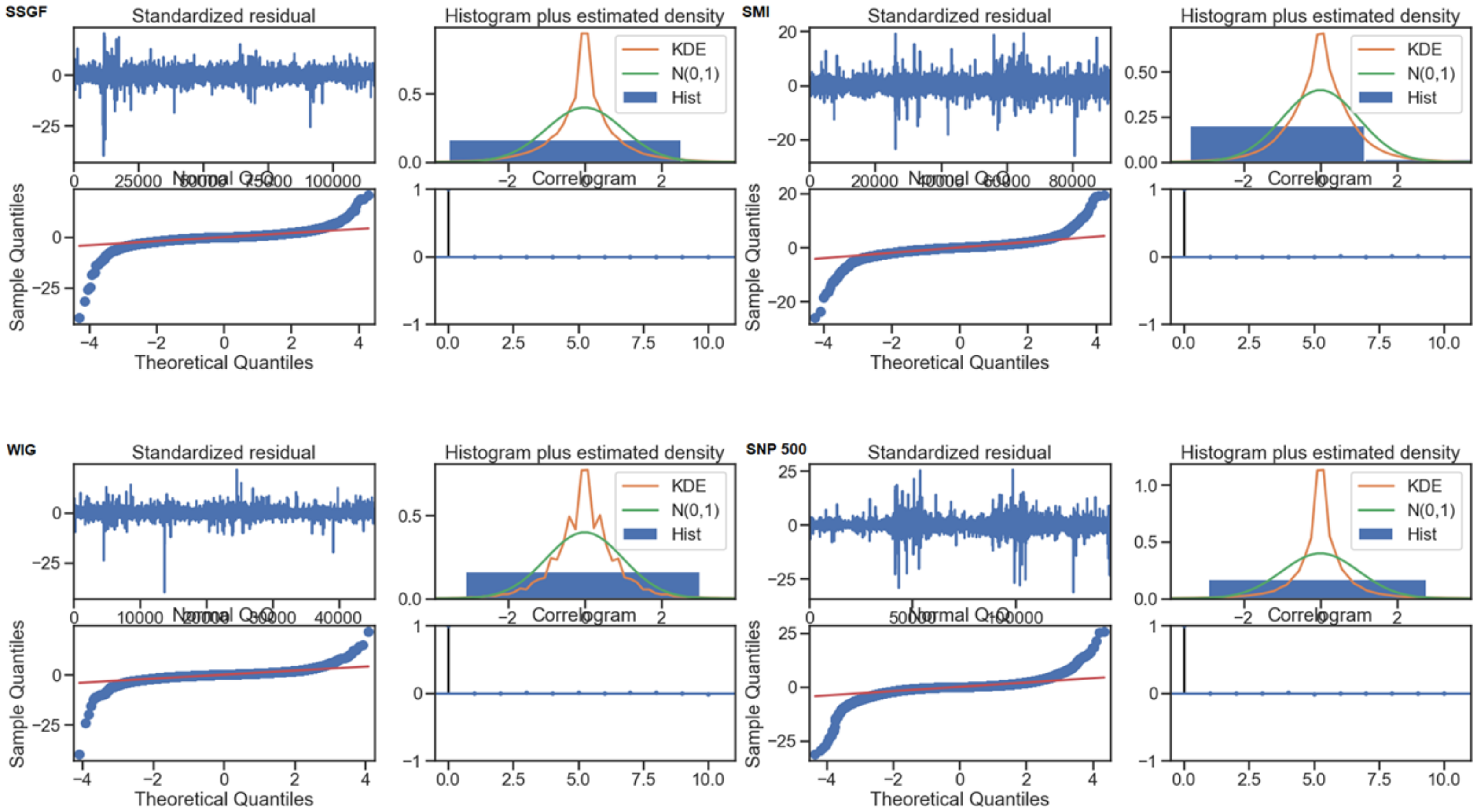

WIG
Density

Appendix C-2: KDE of ARIMA models residuals

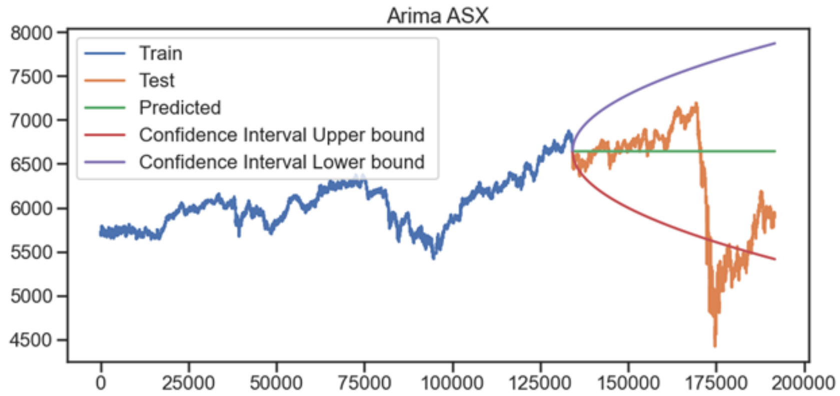


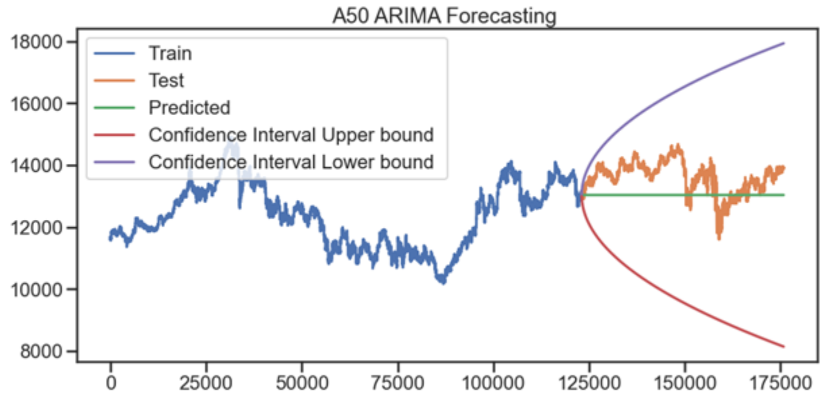

EUS
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
FTSE
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
DAX
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
DJI 30
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles

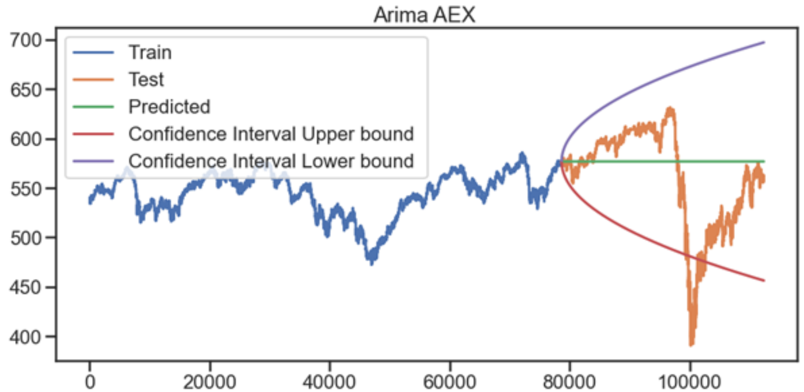

HSI
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
NIFTY
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
NIKKEI
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles
IBX
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Correlogram
Sample Quantiles
Theoretical Quantiles

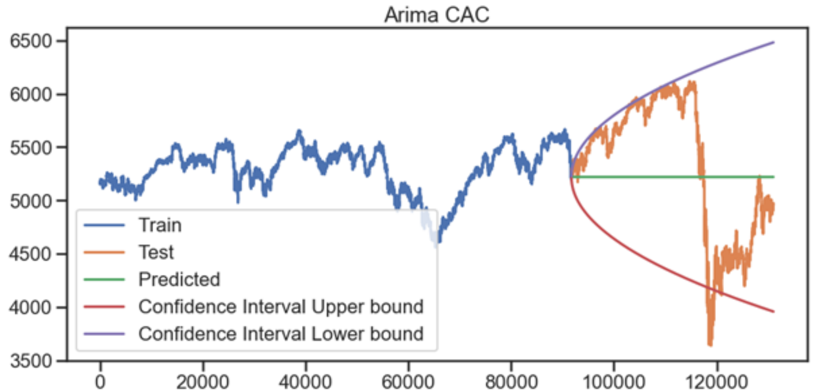

SSGF
Standardized residual
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Sample Quantiles
Theoretical Quantiles
Correlogram
SMI
Standardized residual
Histogram plus estimated density
Normal Q-Q
Correlogram
WIG
Standardized residual
Histogram plus estimated density
Normal Q-Q
Correlogram
SNP 500
Standardized residual
Histogram plus estimated density
Normal Q-Q
Correlogram

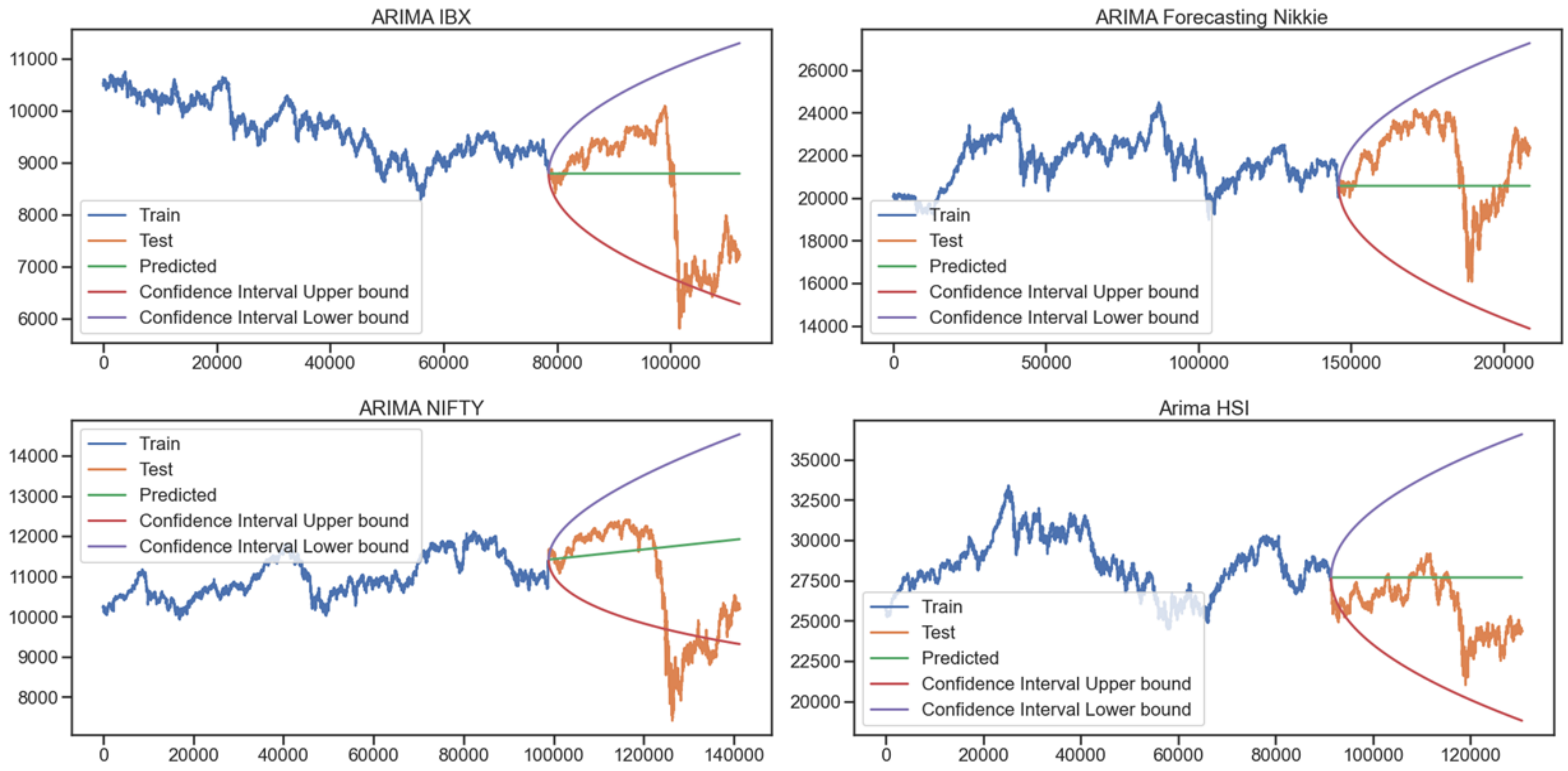
Appendix C-3: ARIMA Forecasting

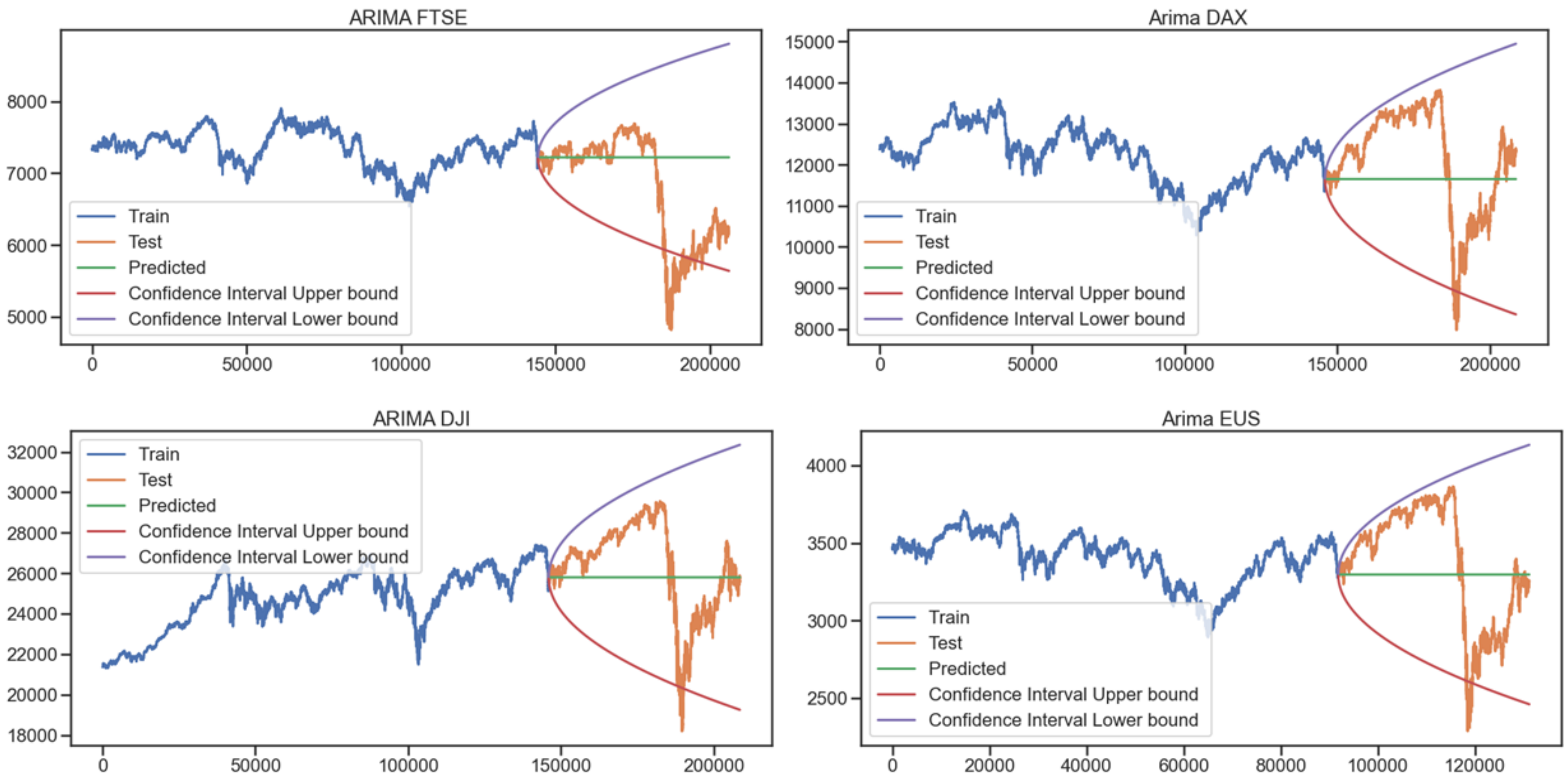




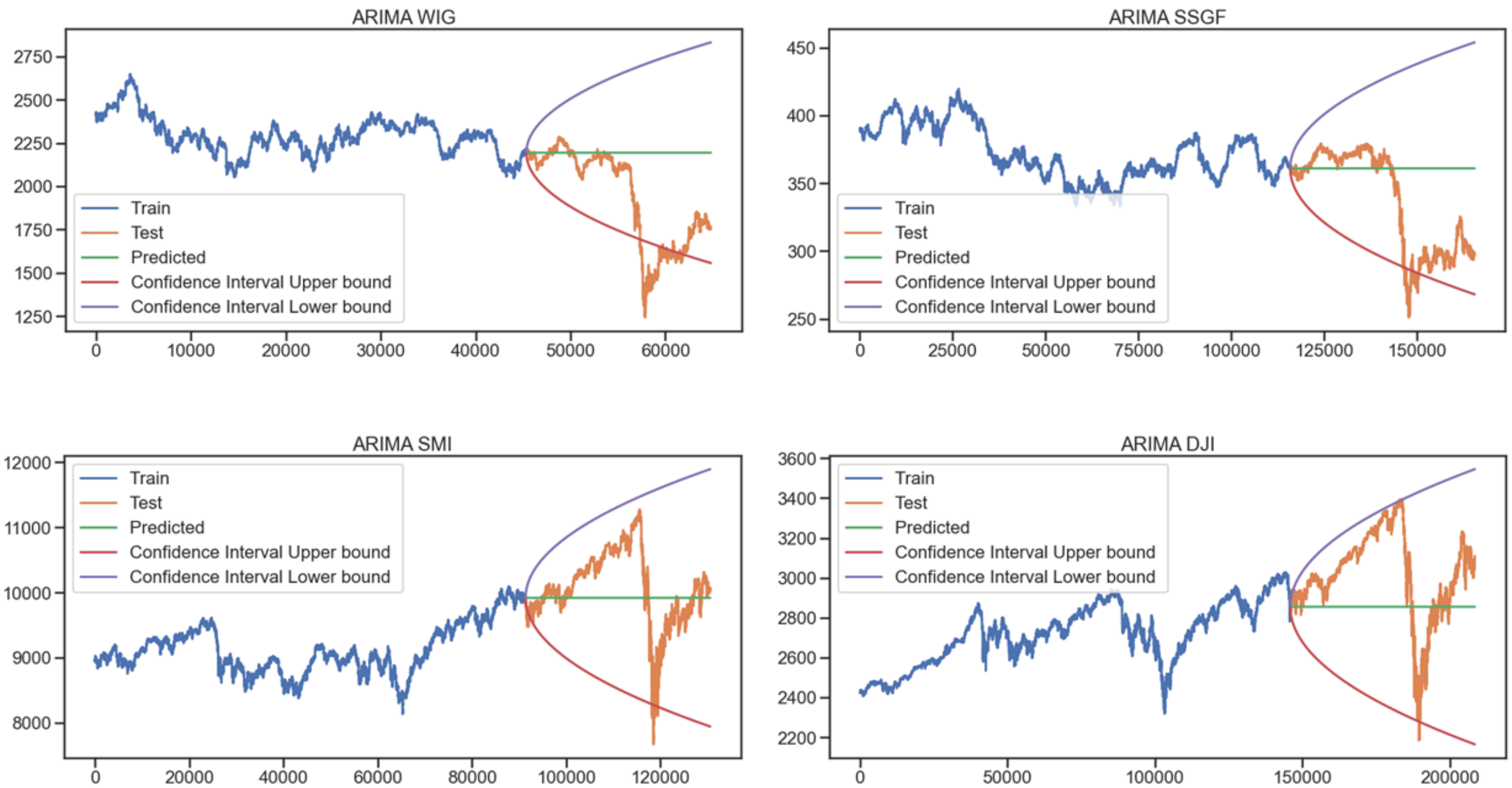





ARIMA IBX
11000
10000
9000
8000
7000
6000
0
20000
40000
60000
80000
100000
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
ARIMA Forecasting Nikkie
26000
24000
22000
20000
18000
16000
14000
0
50000
100000
150000
200000
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
ARIMA NIFTY
14000
13000
12000
11000
10000
9000
8000
0
20000
40000
60000
80000
100000
120000
140000
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
Arima HSI
35000
32500
30000
27500
25000
22500
20000
0
20000
40000
60000
80000
100000
120000
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound

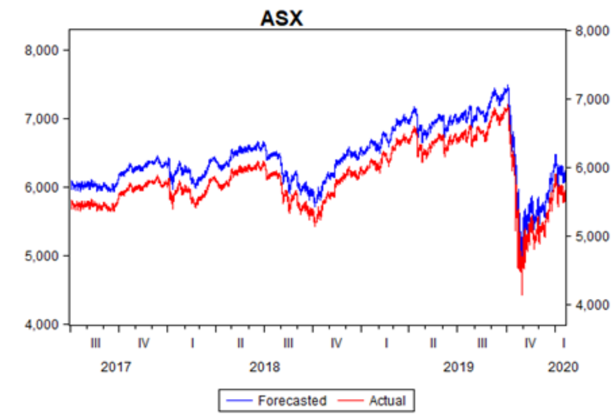


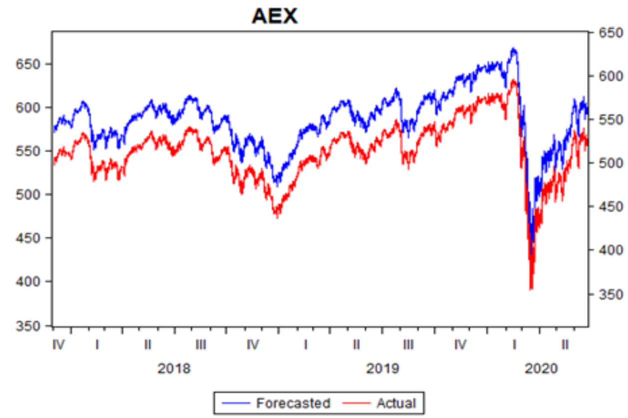
ARIMA FTSE
8000
7000
6000
5000
0
50000
100000
150000
200000
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
Arima DAX
15000
14000
13000
12000
11000
10000
9000
8000
0
50000
100000
150000
200000
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
ARIMA DJI
32000
30000
28000
26000
24000
22000
20000
18000
0
50000
100000
150000
200000
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
Arima EUS
4000
3500
3000
2500
0
20000
40000
60000
80000
100000
120000
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound

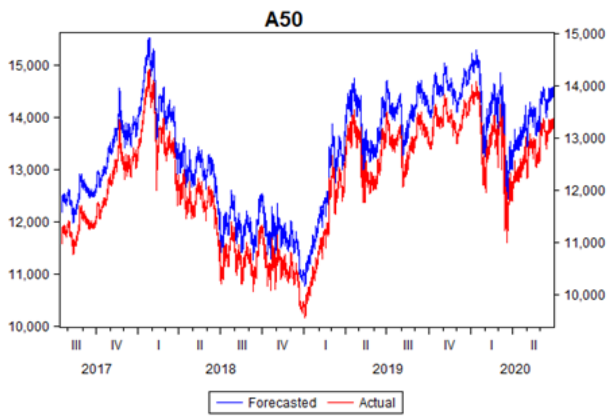

ARIMA WIG
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
2750
2500
2250
2000
1750
1500
1250
0
10000
20000
30000
40000
50000
60000
ARIMA SSGF
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
450
400
350
300
250
0
25000
50000
75000
100000
125000
150000
ARIMA SMI
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
12000
11000
10000
9000
8000
0
20000
40000
60000
80000
100000
120000
ARIMA DJI
Train
Test
Predicted
Confidence Interval Upper bound
Confidence Interval Lower bound
3600
3400
3200
3000
2800
2600
2400
2200
0
50000
100000
150000
200000

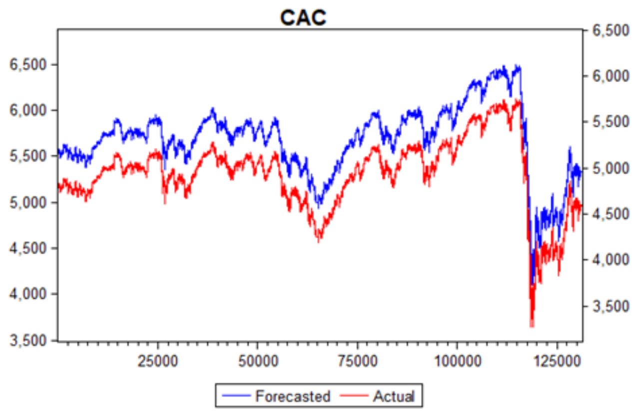
Appendix: C-4: ARFIMA Forecasting

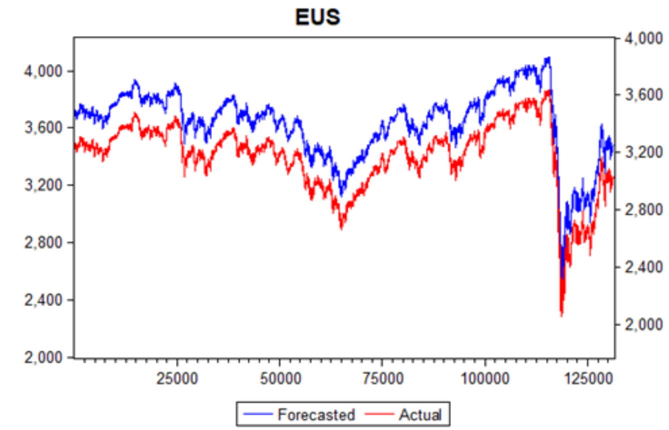
AFIMA Model

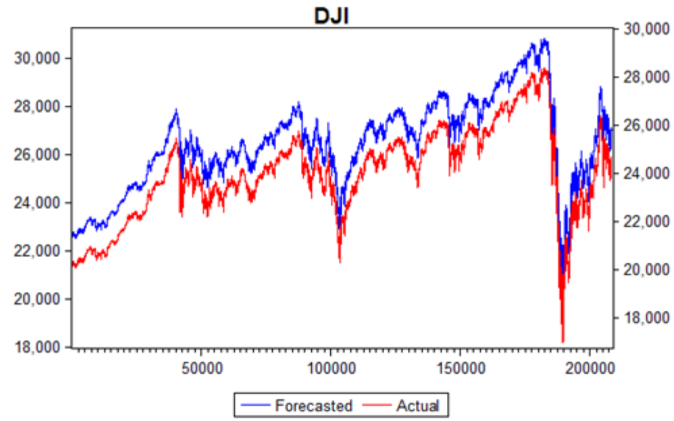


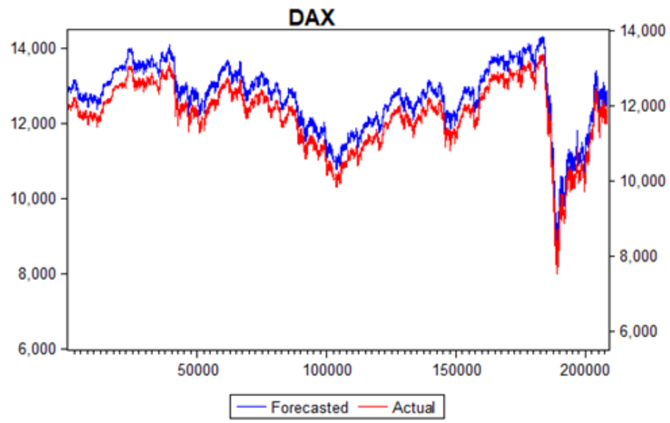


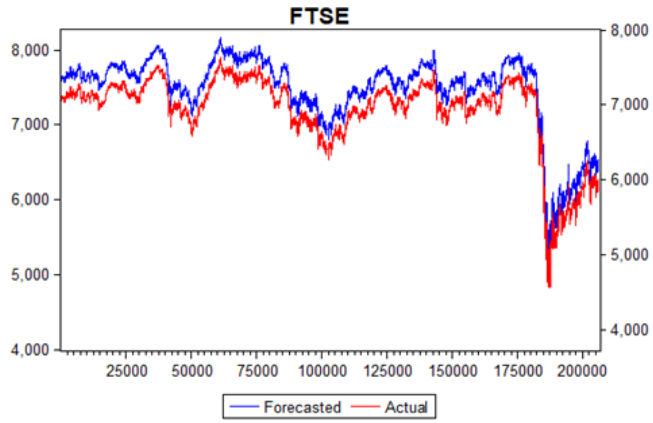


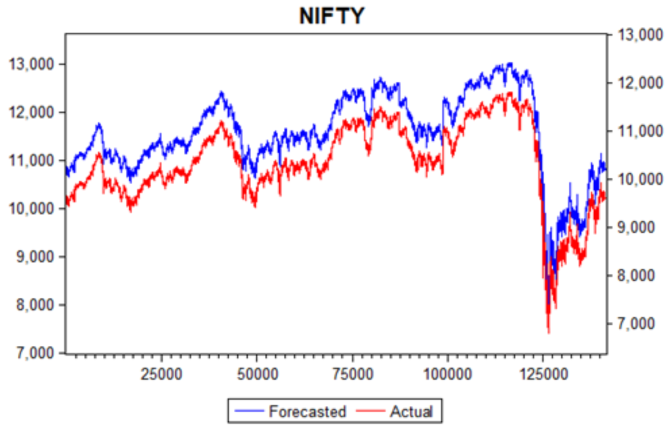


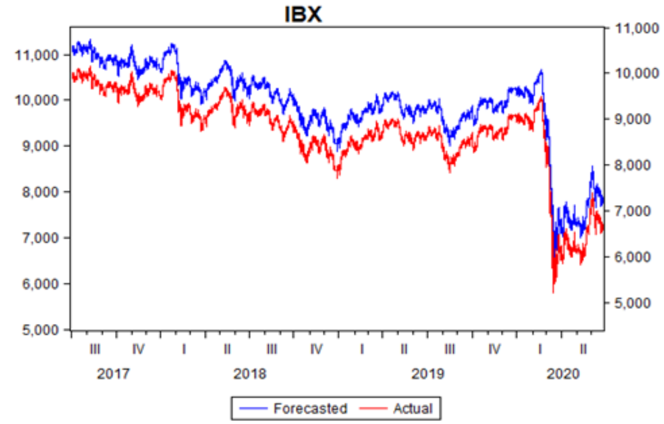

EUS
Forecasted
Actual

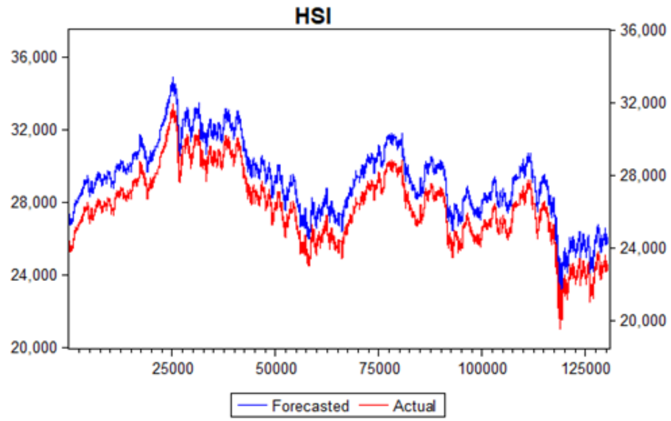

DJI
Forecasted
Actual

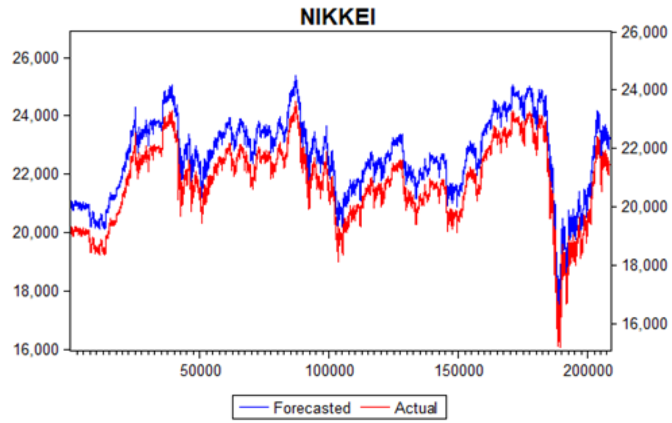

DAX
Forecasted
Actual

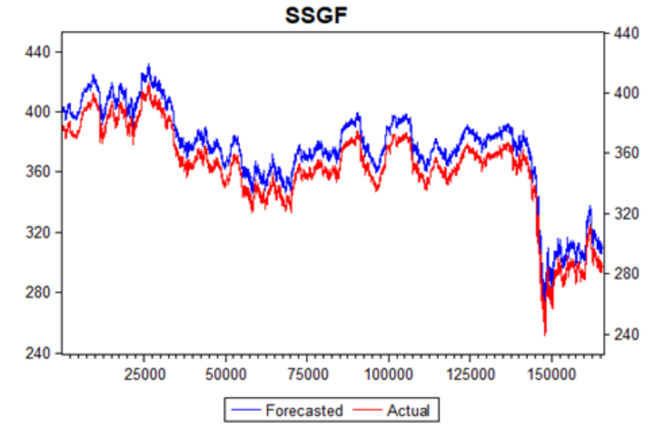

FTSE
Forecasted
Actual

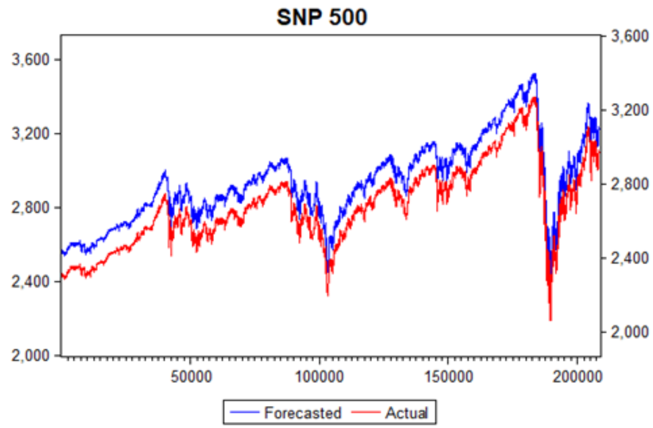

NIFTY
13,000
12,000
11,000
10,000
9,000
8,000
7,000
25000
50000
75000
100000
125000
Forecasted
Actual

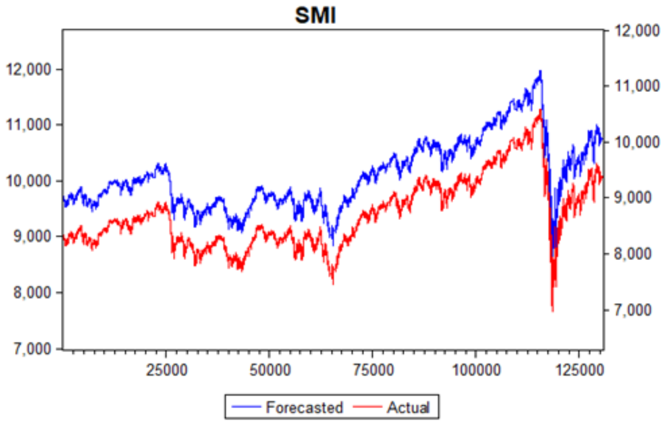

IBX
11,000
10,000
9,000
8,000
7,000
6,000
5,000
III IV I II III IV I II III IV I II
2017
2018
2019
2020
Forecasted
Actual

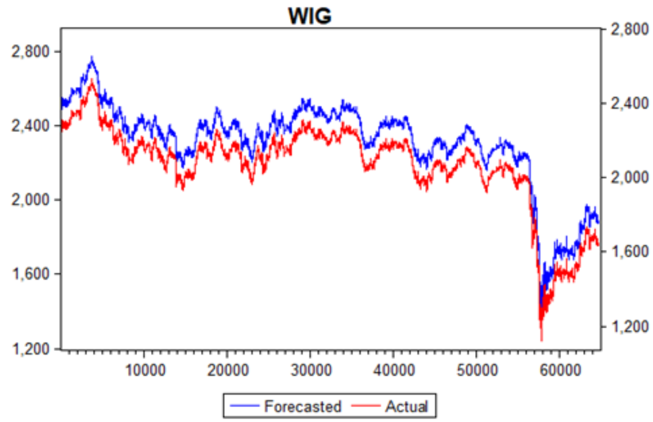

HSI
36,000
32,000
28,000
24,000
20,000
25000
50000
75000
100000
125000
Forecasted
Actual


NIKKEI
26,000
24,000
22,000
20,000
18,000
16,000
50000
100000
150000
200000
Forecasted
Actual


SSGF
440
400
360
320
280
240
25000 50000 75000 100000 125000 150000
Forecasted Actual

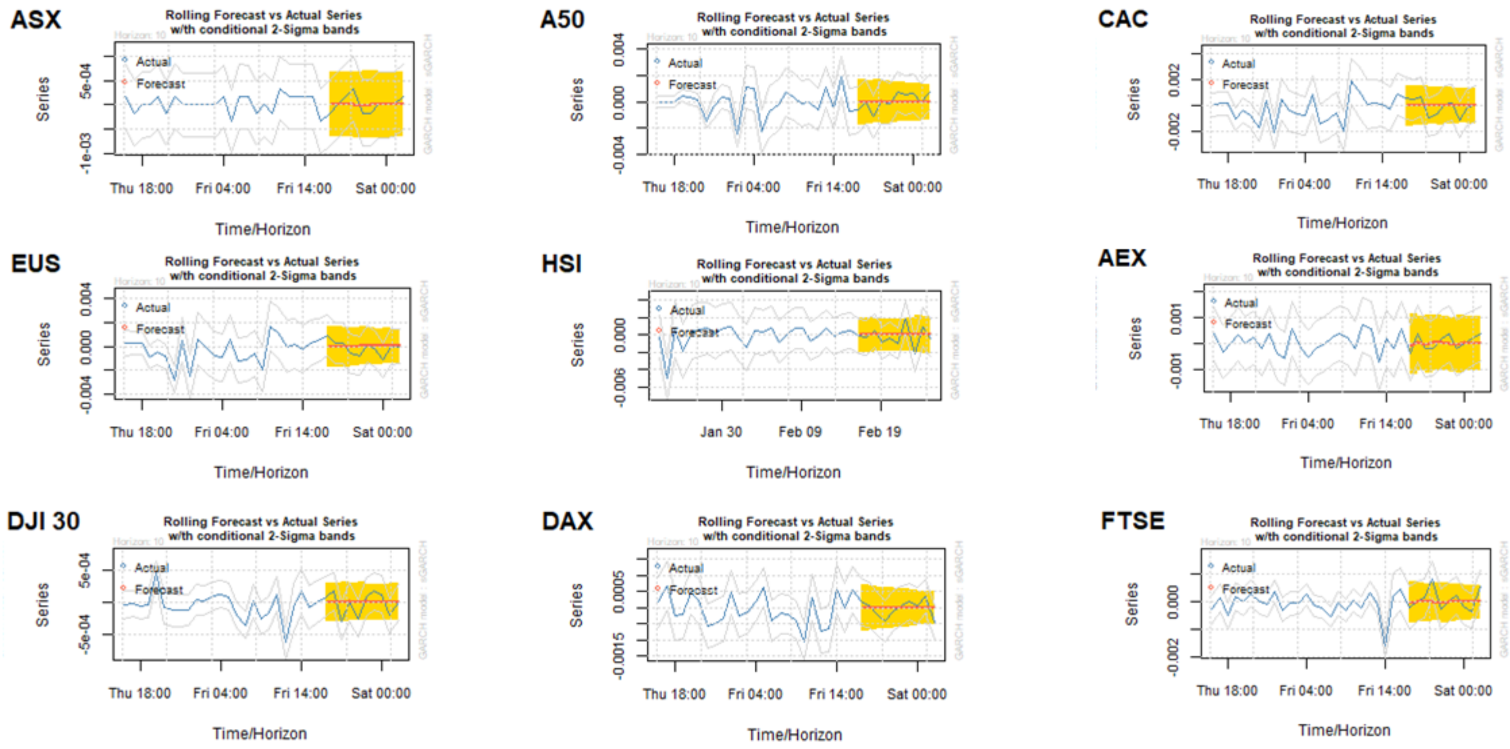

SNP 500
3,600
3,200
2,800
2,400
2,000
50000 100000 150000 200000
Forecasted Actual


SMI
12,000
11,000
10,000
9,000
8,000
7,000
25000 50000 75000 100000 125000
Forecasted Actual

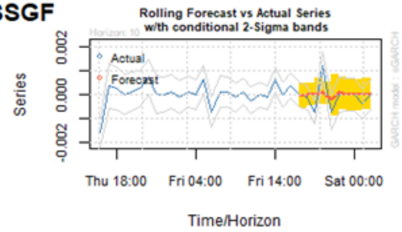

WIG
2,800
2,400
2,000
1,600
1,200
10000 20000 30000 40000 50000 60000
Forecasted Actual

Appendix: C-5: GARCH Forecasting

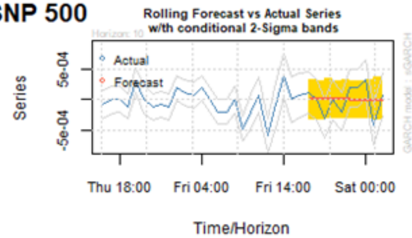
GARCH Model



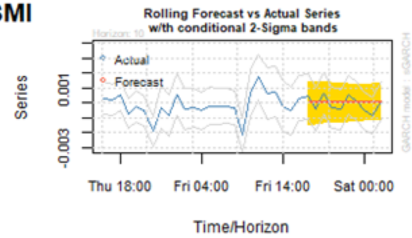
SSGF



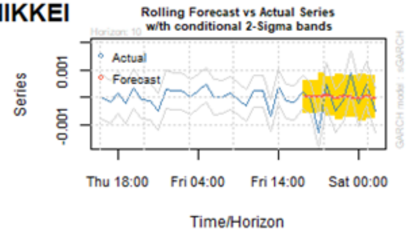
SNP 500



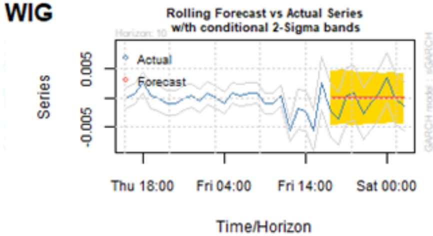
SMI



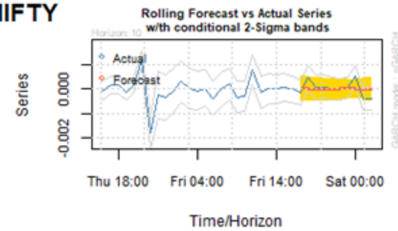
NIKKEI



WIG



NIFTY



IBX