CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Returns Optimization and Returns Prediction in Presence of Fear Sentiments Using Machine Learning Algorithms

by

Hifsa Hussain Raja

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

Returns Optimization and Returns Prediction in Presence of Fear Sentiments Using Machine Learning Algorithms

by

Hifsa Hussain Raja

(MMS191011)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Muhammad Khalid Sohail	BU,Islamabad
(b)	Internal Examiner	Dr. Jaleel Ahmed Malik	CUST Islamabad
(c)	Supervisor	Dr. Nousheen Tariq Bhutta	CUST Islamabad

Dr. Nousheen Tariq Bhutta
Thesis Supervisor
April, 2021

Dr. Lakhi Muhammad
Head
Dept. of Management Sciences
April, 2021

Dr. Arshad Hassan
Dean
Faculty of Management & Social Sci.
April, 2021

Author's Declaration

I, **Hifsa Hussain Raja** hereby state that my MS thesis titled "**Returns Optimization and Returns Prediction in Presence of Fear Sentiments Using Machine Learning Algorithms**" 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.

(Hifsa Hussain Raja)

Registration No: MMS191011

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Returns Optimization in Presence of Fear Sentiments Using Machine Learning Algorithms**" 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.

(Hifsa Hussain Raja)

Registration No: MMS191011

Acknowledgement

With the name of Allah, who is most Beneficent and most Merciful. I would like to thanks my parents for their constant and endless support and prayers. I could never ask anything better than such beautiful people. My sisters (Api Fouzia, Aqsa and Samavia) remained very consoling and encouraging.

I am grateful to my supervisor Dr. Nousheen Tariq Bhutta for her kind guidance in my work. She was very humble with me throughout. I would express my appreciation for Dr. Arshad Hassan for his in-depth knowledge and intellectual input in eradicating my queries.

I also acknowledge my friend Rimsha Ejaz for always being there for me, no matter what. She was very supportive and helpful in the hard time of study. I like to pay respect and graditude to my brother Muhammad Arslan for his help in work.

(Hifsa Hussain Raja)

Abstract

Investor decision making is very critical for return optimization. Use of machine learning help us to predict the returns of the indices's in presence of fear sentiment. We consider the fear sentiment and forecast the returns of the indices's in the period of high sentiment. Machine learning has helped us to do so.Investor decisions involve human error due to presence of sentiments and emotions as human decisions are composite of sentimental and biased behavior. This element can also effect the asset allocation and investment. Machine learning can overcome the sentimental impact in decision making as machine leaning algorithms are mechanical methods with analytical attributes and rational and unbiased decision making. Therefore, this study employs machine learning and non-linear classical methods for prediction, to achieve the objective of the study is to find out the impact of fear sentiment of investor on the returns in developed and emerging markets. Results concludes that machine learning outperform the forecasting of returns and rational assets allocation. This study uses system of equations to test the impact of fear sentiments on returns of developed and emerging markets. GARCH is used to check the persistent of volatility of respective markets .Machine Learning Algorithms are used for rational capital allocation to optimize the returns.

The results show that developed and emerging markets have the impact of fear sentiment on the returns. Emerging markets outperform the developed markets in return maximization. The results tells us that machine learning help us to predict the returns even in the presence of high sentiment in the market. The results depicts that machine learning algorithms make the rational decision making in high fear sentiment because it reallocate its investment accordingly for return optimization. This study facilitates the investors and fund managers to use these machine learning algorithms for booking higher returns in market. It also enables policy makers for developing strategies for macro stabilization. This study helps in return forecasting in fear sentiment and in portfolio optimization.

Keywords: Machine Learning Algorithms, Fear Sentiment, Rational Decision Making, Return Optimization, Economic Policy Uncertainty.

Contents

Author's Declaration iv

Plagiarism Undertaking v

Acknowledgement vi

Abstract vii

List of Figures x

List of Tables xi

Abbreviations xii

1 Introduction 1
1.1 Background of the Study 1
1.2 Theoretical Background 3
1.2.1 Mean Variance Theory 4
1.2.2 Market Efficiency Theory 5
1.2.3 Behavioral Finance 6
1.3 Gap Analysis 7
1.4 Problem Statement 8
1.5 Research Questions 9
1.6 Research Objectives for This Study 10
1.7 Significance of Study 10
1.8 Plan of Study 11

2 Literature Review 12
2.1 Economic Policy Uncertainty 12
2.1.1 Fear Sentiments 13
2.2 Forecasting Models 17
2.3 ML and Return Optimization 21
2.4 Framework 22
2.5 Hypothesis Statements 22

3 Research Methodology **24**
3.1 Population and Sample of the Study . . . 24
3.2 Empirical Methodology . . . 25
3.2.1 Interactive Regression Model . . . 26
3.2.2 System of Equations . . . 27
3.2.3 GARCH Model . . . 27
3.2.4 Machine Learning Model . . . 28
3.3 Data Pre-Processing and Environment Setting . . . 29
3.4 Portfolio Optimization Methodology . . . 30
3.4.1 Processing historical prices . . . 30
3.4.2 Mean-Variance Optimization . . . 30
3.4.3 Short Positions . . . 31
3.4.4 Black-Litterman Allocation . . . 32
3.4.5 Output of the BL model . . . 32
3.4.6 Other Optimizers . . . 32
3.4.7 Hierarchical Risk Parity Algorithm . . . 33
3.4.8 The Critical Line Algorithm . . . 33
3.4.9 Selection of Model . . . 33

4 Results and Discussion **34**
4.1 Descriptive Statistics . . . 34
4.2 Fear Sentiments and Market Returns . . . 36
4.3 Forecasting Methods . . . 42
4.3.1 GARCH Model Estimation . . . 42
4.3.2 Machine Learning Model Estimation . . . 43
4.3.3 Models Forecasting . . . 44
4.4 Return Optimization . . . 58
4.4.1 Scenario Analysis . . . 59
4.5 Discussion . . . 63

5 Conclusion **66**
5.1 Recommendations . . . 67

Bibliography **70**

List of Figures

2.1 The impacts on the stock returns of the market and volatility due to the fear sentiment and the economic policy uncertainty 23

3.1 Interactive Regression for Moderation 26
3.2 Portfolio Optimization 30

4.1 Descriptive Statistics (Developed Markets) 37
4.2 Descriptive Statistics (Developed Markets) 38
4.3 Descriptive Statistics (Emerging Markets) 39
4.4 Descriptive Statistics (Emerging Markets) 40
4.5 GARCH Model Forecasting(Developed Markets) 46
4.6 GARCH Model Forecasting(Developed Markets) 47
4.7 Machine Learning Forecasting(Developed Markets) 48
4.8 GARCH Model Forecasting(Developed Markets) 49
4.9 Machine Learning Forecasting with Fear Sentiment (Developed Markets) 50
4.10 Machine Learning Forecasting with Fear Sentiment (Developed Markets) 51
4.11 GARCH Model Forecasting(Emerging Markets) 52
4.12 GARCH Model Forecasting(Emerging Markets) 53
4.13 Machine Learning Forecasting(Emerging Markets) 54
4.14 Machine Learning Forecasting(Emerging Markets) 55
4.15 Machine Learning Forecasting with Fear Sentiment (Emerging Markets) 56
4.16 Machine Learning Forecasting with Fear Sentiment (Emerging Markets) 57

List of Tables

3.1 Stock Market Indices 26

4.1 Descriptive Statistics 35
4.2 Fear Sentiments and Stock Market Returns 41
4.3 GARCH Estimation 44
4.4 Models Comparison 58
4.5 Return Optimization 59
4.6 Portfolio Reconstruction (Developed Markets) 60
4.7 Portfolio Reconstruction (Emerging Markets) 62

Abbreviations

ANN	Artificial Neural networks
ARCH	Autoregressive Conditional Heteroskedasticity
AI	Artificial Intelligence
API	Application programming interface
CBOE	Chicago Board Options Exchange
EMH	Efficient Market Hypothesis
EPU	Economic Policy Uncertainty
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GFC	Global Financial Crisis
LSTM	Long Short Term Memory
ML	Machine Learning
MSCI	Morgan Stanley Capital International
NPV	Net Present Value
RMSE	Root Mean Squared Error
VIX	Volatility Index

Chapter 1

Introduction

1.1 Background of the Study

The classical theory devises the market with two major classes of investors. The rational investor, who use the fundamental analysis while making the decisions regarding the investments. Whereas, another class of irrational investors, those believe in the heuristic approach. Today's environment is more dynamic and chaotic in nature and one cannot define the rationality term in this environment. Markets stability and efficiency cannot be achieved in presence of human behavior, which is not predictable (Habibah et al., 2017). These behavioral biases are known as the investor's sentiments. And these investors' sentiments represent the overall investor attitude towards a specific financial instrument. Literature also supports the argument of investor's sentiments have an economically significant impact. Zhang (2008) defines sentiments as erroneous beliefs of investors against some future cash flows. In classical finance, sentiments are also known as a myth. It is well-documented fact in the literature that returns of marketable securities are mostly determined by the investor's behavior in a particular market rather than the rationale analysis. This fact was established by numerous studies that investor's sentiments affect the market returns (Fisher and Statman, 2000; Goetzmann et al., 2015; Maghyereh et al., 2020; Schmeling, 2009; Wang et al., 2006).

While explaining market returns, the literature suggests a change in sentiments is more likely effective rather than the sentiments levels. To measure the future market returns and sentiments connections study like Smales (2017) mentioned that market response during the financial distress against the sentiments found to be strong. An investor fear gauge or fear index provides superior information regarding the sentiments and help to foresee the future returns more effectively. Investors sometimes trade on basis of information flow. Which mean as information arrives investor immediately response to information without prior fundamentals. In presence of arbitrager and noise traders in market, they don't rely on fundamental analysis of the market to gain. But they rely on the sentiments and used to take advantage of the market (Verma et al., 2008). This noise trading affects the market trading structure as well.

Previously Verma and Verma (2007) raised questions on this noise trading effect on market risk pattern by "To what extent do individual and institutional investor sentiments impact stock market volatilities? Moreover, if such relationships do exist, are the effects driven by rational risk factors or noise?". The study finds a noteworthy impact on volatility of the financial market due to investor's sentiments. Literature suggests that investor sentiments significantly cause the market lagged variance (Chakraborty and Subramaniam, 2020; Haritha and Rishad, 2020). These sentiments are the results of investor's emotions towards the specific financial instrument and substantially explain the lagged variance in the stock markets. As the sentiments have stronger intensity in variations cause higher volatility in the market (Lee et al., 2002). In a hypothetical world, we may assume that markets are efficient and all of the participants of the market are rational and their decisions making is based on fundamentals. But in real world, as the literature suggests the market is a composite of both rational and irrational investors. So this sentiment is a permanent part of the market. We can not control the effect of sentiments but we can minimize the effect.

Investor in any market comes with the aim of wealth maximization by using the sentiments and rational approaches. Investors use different methods to maximize their returns either by investing in positive NPV projects or by minimizing the risk factor and by applying different investment strategies i.e., speculations, arbitraging, etc. These optimizations could be based on specific techniques or by using the advance mathematical function. In both cases, investors have to deal with the sentiments. For this stock market forecasting is an important part of portfolio creation and optimization. Hence today's stock markets are more volatile and nonlinear in nature (Sim et al., 2019). An increasing volatility phenomenon raises serious issues for the investors to forecast the stock market returns and then optimize the portfolio according to. In literature, there are multiple methods available that used to forecast and optimize the stock market returns. Mainly these methods are divided into two major classes (a) Econometric models and (b) Machine learning-based methods (Athey and Imbens, 2019). These machine learning methods can be briefly described as, allowing computers to predict effectively with experience in the past. And we have seen remarkable progress with the support of the rapid rise in the storage and processing capacity of computer systems (Baştanlar and Özuysal, 2014). To make it useful, machine learning methods learn from experiences and thus boost the performance of systems automatically through execution. In the context of adequate programming, a basic yet efficient rote-learning facility allow investors to spend in future and to earn high returns on the market.

1.2 Theoretical Background

The theoretical support for this study has been given by the three theories of finance. The first theory is the mean-variance theory, efficient market theory, and behavioral theory. The mean-variance theory tells us about the portfolio optimization and resource allocation of the assets. Efficient market theory tells us that stock prices consider all the available information. Any news in market is translated into prices. The behavior finance tells us about the sentiment effect of the investor on the market behavior. The fear sentiment affects the investment of a particular stock if the price is continuously crashing.In behavioral finance ,the prospect theory is used.

1.2.1 Mean Variance Theory

The uncertainty is present in everything specially in business world. To reduce the risk, investment in different types of securities is an important factor for investor. But returns and risk associated with collection of securities still raise questions. Finding the risk associated with the securities was a major challenge in the 1950s. Also an adequate theory that deals with the phenomenon of diversification whole the components of portfolios or combinations of securities with correlated risk, finding efficient and inefficient securities in portfolios, and tradeoff between risk and return was missing in literature (Markowitz, 1999).

Markowitz (1952) presented the theory of risk and return which opens a new chapter of managing investment in a systematic way. Which is a major breakthrough in the field of computational finance. In theory, he proposed the methods for the estimation of risk on certain security and return on these securities. According to mean-variance theory risk on an asset reduce by adding uncorrelated assets in a portfolio. So unsystematic risk will be reduced by doing this. But these securities still have systematic risk left with them. If we define the systematic risk and unsystematic risk, systematic risk refers to the components that are correlated with the securities, and such components cannot be diversified that affect the whole market (Beja, 1972). Further, unsystematic risk is the element of risks that can be diversified or uncorrelated with securities and known as firm-specific risk. The investor tries to reduce the unsystematic risk as well as systematic risk. The unsystematic risk can be reduce whereas the systematic risk can not be reduce after a certain limit even after diversification. In other words, uncertainty or volatility is the systematic risk that every asset has to face. The market is not efficient as we perceive. Today's market volatility is increased more as compared in the past. But uncertainty in the environment creates the risk of loss on the investment, this creates the concern of investors and academic researchers to find out the way to predict this environment more effectively. The estimated statistics of risk enhance the confidence of investors regarding their investment and anticipated future potential loss. Mean-Variance Model can also be used as a way to optimize the portfolio as described in mean-variance theory.

1.2.2 Market Efficiency Theory

Fama proposes an efficient market theory in 1960 which says, markets are efficient and does not follows any patterns. Prices reflect the all present and past information. And this information is easily accessible to every market participant. Markets follows random walk pattern and in random walk patterns we cannot predict the prices of any instrument. During early seventies Fama (1970) revisited his work and proposed a phenomenal work known as Efficient Market Hypothesis(EMH) that built an new era of modern financial markets and gives a solid support to efficient market theory. According to EMH markets can be segregated in three forms of efficiency. The strong form of efficiency is that form of efficiency which includes all the information that is available in public and private information. No technical analysis nor any fundamental analysis or any other insider information can help to forecast the movements of the prices. This means we cannot predict the prices of the stocks on basis of information. Second form of efficiency refers to the semi-strong includes the publicly available information and some form of other information like dividend announcement and stock split. Such markets are informally efficient. The weak form of efficiency proposes that previous prices and equally likely to available every participant in the market. There are three type of market participants exists mainly. The rational investor who uses the fundamental and participate in market. Second type of participants are irrational investor which uses the sentiments to participate and third type of participants are the arbitragers who takes the advantages of miss pricing and move the market into equilibrium.

Another school of thought support the argument that in markets there is no random walk prices phenomenon and prices can be easily predicted Marszałek and Burczyński (2014). This study is getting support from a weak form of efficiency as fear sentiments have reflection in the prices and if affects the stock prices whenever there is some bad news that increases the volatility of the market. The fear sentiments also increase the volatility and volatility affects the stock prices. This means the fear sentiment created by the bad news is translated in the stocks. Whenever the fear sentiment is high, the stock prices will fall. The market will reflect that fear as soon as the information or the news arrives. Supported by Fama argument market will translate all the information which is present in the market.

1.2.3 Behavioral Finance

Complexities in the financial market and market uncertainties have contributed to the emergence of Behavioral Finance, the latest area of finance science. Behavioral Finance is an aspect of behavioral economics, which was created by Kahneman and Tversky (1979) as the product of a theory of prospect. Cross-sectional and time-series patterns on the return on investment of securities that cannot be modelled by any clearly-defined hypothesis is the source of anomalies in the financial market. Behavioral finance is one that includes the psychological factors of an investor in investment.

This modern approach to financial analysis promotes the social and emotional influences impacting the judgment on the investment. This modern approach seeks to show that investors are motivated by psychological influences (e.g., mood, over-confidence, hope, pessimism, and fear). So, due to the impact of investor's psychological factors on investment decisions and strategies theoretical support for this study is based on the major behavioral finance theory which is prospect theory. Prospect theory was proposed by Kahneman and Tversky (1979) to explain the investor/people decision regarding all those choices which contain risk.

According to prospect theory while taking decision people think about the expected utility relative to a reference point instead of the possible outcomes of the decision. This theory concludes that people do not like loss and they are willing to take the risk to avoid loss instead of taking risks for getting equivalent gains. And based on prospect theory categories of peoples are made such as risk-averse and risk-seeking on the basis of probabilities of certainty and possibility effect. Theoretical argument for this study is derived on prospect theory that in a volatile environment investor takes investment decision. Bad news creates more volatility in the market which increases the investor fear sentiments that are related to loss on the investment. According to prospect theory investor do not like loses which makes the investor panic and they take an irrational decision regarding the buying and selling of the assets which affect the returns in the market.

1.3 Gap Analysis

Sentiment analysis with respect to stock market returns has been explored by various researchers over the time period. Studies like Canbaş and Kandır (2009); Fisher and Statman (2000); Liu et al. (2020); Verma and Verma (2007) analyzed relationship of investor sentiments with stock markets returns. Studies used different approaches to analyze the phenomenon. It is well established fact that bad news creates more volatility in market and volatility has inverse relation with stock market returns. Similarly, events like GFC (Global Financial Crisis) or pandemic outbreaks i.e., COVID-19 also creates the element of fear in investor and that leads towards the market volatility (Chen et al., 2020a). During such events economic policies of governments play a crucial rule. Like during current global pandemic Chinese stock market outperform the rest of the world due to Chinese government policies regarding the economy and that reduce the economic stress. Many studies show a closely integration of Chinese stock market with other emerging markets and some develop markets but during current pandemic most of these markets fails to perform because of economic policies of respective governments. These economic policies uncertainty has significant impact not only on volatility of financial instrument but also on the investors behavior (Al-Thaqeb and Algharabali, 2019; Ftiti and Hadhri, 2019). Ultimately these sentiments are translated into the investor's investment decisions and these decision has a significant impact on stock returns (Salisu et al., 2020). Previously many studies explore this relation using methods like VAR, OLS, GARCH based model etc. These studies are successful in establishing the fear sentiments, volatility, news/events effect on stock markets performance.

Many of these studies used across markets analysis to determine the facts. But we all know we cannot change the investor emotions about any market or investment. But after affect can be reduce after incorporating the knowledge. In real world it is more effected that one can use to analyze the performance of stock market in presence of these sentiments. Bad news affects the volatility but moderating role of bad news on volatility and stock returns is still missing in literature as per best of my knowledge. Right on other hand, use of rational tools in irrational world may help to improve the returns on investments. Over the past few decades there are number of models has been introduced to capture the performance and forecast the investor returns to be used for future investment decisions. Forecasting and optimization of returns have been a challenge for the investors throughout the time period. Literature supports the argument of smartly forecasting and then optimization can reduce the risk of losses (Elmachtoub and Grigas, 2020). For forecasting, classical linear to non-linear and then advanced machine learning based algorithm has been intensively used in different studies. Studies like Dingli and Fournier (2017); Kewat et al. (2017); Li et al. (2020b); Sezer et al. (2020) used different econometric and machine learning methods to forecast the performance of stock markets but didn't use these techniques for the investment optimization. These machine learning algorithms outperform the forecasting abilities of models and it is well expected these models will also perform better in presence of fear sentiments. Another study by Ta et al. (2018) suggested to apply the quantitative trading technique to optimize the portfolio by using the machine learning methods. This is what still has remained understudied in the literature to use of machine algorithms while taking an investment decision in presence of fear sentiments. Further, A comprehensive study that combine the aforementioned work in one piece, so it produces the ease for the investors in era of uncertainty.

1.4 Problem Statement

Numerous studies have been carried out on investor's fear sentiments and its impact on stock markets returns across the markets.These studies focus on the behavior of the investor. These studies used different econometric models to measure the presence of investors sentiments and stock markets returns relationship (Balcilar et al., 2017; Canbaş and Kandır, 2009; Corredor et al., 2015; Smales, 2017;

Verma et al., 2008). Recently machine learning methods has been emerged as the non-linear method to deal this complex phenomenon. Studies like i.e., Albulescu (2021) used Machine algorithm to capture the investor sentiments and forecast the respective financial markets. This study is designed to capture the investor fear sentiments along with some important control and moderating variable like economic policy uncertainty and bad news. Further, forecasting of markets for returns maximization either an investor should go for the econometric model like GARCH or Machine learning method in presence of fear sentiments. However, how these methods react in developed markets and in emerging markets too give the confidence to investor for future prediction in presence of fear sentiments.

1.5 Research Questions

This research will answer the following questions:

Research Question 1

What is the impact of fear sentiment on the market returns for developed markets?

Research Question 2

What is the impact of fear sentiment on the market returns for emerging markets?

Research Question 3

How to optimize the returns in the presence of fear sentiments?

Research Question 4

Does machine learning model outperform the econometric model in forecasting in developed markets?

Research Question 5

Does machine learning model outperform the econometric model in forecasting in emerging markets?

Research Question 6

How do ML algorithms outperform returns optimization in developed markets?

Research Question 7

How do ML algorithms outperform returns optimization in emerging markets?

1.6 Research Objectives for This Study

Objectives of the study are as follows:

Research objective 1

To evaluate the fear sentiments in market returns in developed markets.

Research objective 2

To evaluate the fear sentiments in market returns in emerging markets.

Research objective 3

To optimize the returns in the presence of fear sentiments.

Research objective 4

To identify a better prediction model for forecasting in developed markets.

Research objective 5

To identify a better prediction model for forecasting in emerging markets.

Research objective 6

To evaluate the ML algorithms for returns optimization in developed markets.

Research objective 7

To evaluate the ML algorithms for returns optimization in emerging markets.

1.7 Significance of Study

Financial markets are full chaotic environment that's creates and contribute significantly in the volatility dynamics of markets (Chakraborty and Subramaniam, 2020). It is well established fact that we cannot ignore the existence of fear sentiments with in the behavior of investors. The rationality of investors persists in an efficient market with stable environment. But human behavior reacts according to the variations in surrounding environment. In the world of artificial intelligence, presence of methods that can handle the human emotions more realistic can help the prime objective wealth maximization. In the presence of fear sentiments, forecasting and portfolio optimization using the machine learning method enhance the investor decision making power under taking the behavioral aspects. With use of machine learning method investors not only can forecast the performance of the stock markets and their investments but also can use the machine learning optimization algorithm to optimize their returns and restructure and their portfolios. In local contest, Pakistan stock market has significant closely integrated with Chinese stock market (Joyo and Lefen, 2019). As closely integrated so both markets share same properties of sentiments and investment behavior. Similarly, machine learning portfolio optimizers works in a same direction for both of markets.

1.8 Plan of Study

This study is comprised of five major chapters which are the introduction, literature review, methodology, results, and discussion, conclusion, and recommendation. The first chapter deals with the introduction. In this chapter, the research topic is introduced along with its little background and significance. The research question and research objective of the study are briefly explained. The second chapter deals with the literature review. The literature review tries to conclude few studies that are conducted in the background of fear sentiments, classical and ML algorithms for price forecasting, and portfolio optimization. Chapter three of the study consist of information related to population, data, a sample of the study, and econometric models. Chapter four of the study deals with the results and discussion of the econometric models. The last fifth chapter deals with the conclusion of the study and recommendations.The study will follow the above mentioned plan.

Chapter 2

Literature Review

2.1 Economic Policy Uncertainty

After the release of the book, 'The Age of Uncertainty' numerous major media and scholarly activities have highlighted uncertainty as an important topic in the financial environment (Galbraith and Sorel, 1977). There is no question about the significance of economic uncertainty, but a single concept of uncertainty is not agreed upon in the literature. Also, until a few years ago, the impact of economic uncertainty on companies was not studied. The economic policy is associated with the economic risk with unknown government policies and regulatory authorities for the future. The possibility that both corporations and individuals will defer their investment is further enhanced by this phenomenon. According to Baker and Wurgler (2007) concern about government policy peaked after 2008 global crisis due to financial market. Household uncertainty about potential legislative structure of the government, spending, wages, monetary policies, and universal health care. Liu and Zhang (2015) investigate the forecast of stock market volatility from the economic policy uncertainty (EPU). He told that higher EPU lead to a higher change in market volatility. He used different prediction models to forecast the volatility by incorporating economic policy uncertainty and this new variable improves the forecasting results.

Further study motivates many researchers to define indicators of uncertainty, especially concerning economic policy uncertainty, which has resulted in different proxies for uncertainty. Economic policy uncertainty(EPU) has an effect on stock markets in the United States for over 25 years. EPU reduces the stock returns and is high during a period of high volatility (Arouri et al., 2016). Economic uncertainty has especially postponed the prospect of recovery from the financial stagnation and recession as companies and households sources have deferred their investment and decisions of consumption. There is just a long-term consequence of the uncertainty of future policies. The confusion is caused by multiple variables. In the short and long term, different challenges impact confusion. Al-Thaqeb and Algharabali (2019) showed a critical effect of uncertainty on economies, companies, and household consumption and investment. Using these metrics, the aim is to track volatility and catch patterns in fluctuations in government policies and regulations.

2.1.1 Fear Sentiments

There is an impact of bad news on the volatility of the stock return of market. There is no direct relation of bd new on the return but it may affect the volatility of the market. In addition to the dynamic influence of linear and nonlinear data, the long and short-term fluctuations of the financial market make it incredibly difficult for stock prices to forecast (Braun et al., 1995). Period shifts in return volatility and beta are closely linked to how asset prices are balanced. Therefore, it is important to encourage our knowledge of asset valuation to know the properties whether volatility or beta. Braun et al. (1995) Model enables a quadratic volatility reaction to the news, with multiple responses to good and bad news, but minimum volatility will happen when there is no news.

With the development of sentiments investing, scholars began to discuss the problem of calculating sentiments to clarify and forecast asset returns. The DSSW model was one of the first systematic theories that incorporated emotions as a factor influencing returns (Aggarwal, 2019). They stated that emotions contribute to investors' erratic actions, which hinders investors from being completely rational. Fleming et al. (1995) find the proof of a broad contemporary association with substantial unevenness between shifts in the VIX CBOE and performance of stock market. VIX is also used as a market-timing confrontational weapon. Further besides, occasionally the Wall Street Journal publishes on VIX movements. VIX as a method to calculate the negative sentiment of investors has already been used. Since 2000, researchers have used sentiment analysis to empirically analyze the association between the forms of emotions with developments of stock markets; moreover, they have gained attention since the 2008 financial crisis (Loewenstein, 2000). Behavioral finance assumes that arbitrage will not fix the deviation between stock price and valuation created by aggressive investors automatically due to the volatile behavior of investors and restricted arbitrage in the real world (Ricciardi and Simon, 2000).

Skiadopoulos (2004) uses the VIX. The reason of this analysis is the use of negative sentiments of investors through the VIX to forecast potential stock market returns by the use of the equation method and GARCH. Analysis of the sentiments are also widely investigated study fields. Further a significant factor shaping the uncertainty of the prices and returns of shares is market sentiment. However, linear models are often used in current research studies to examine the effect of consumer sentiments on financial stock market returns. Most literature studies believe that there is a linear association between investor sentiments and the stock returns, although some literature suggests that the sentiment of investors has a positive effect on market return. Also, Skiadopoulos (2004) demonstrated that "Changes in the VIX drive variations in the expected returns of the Fama and French three-factor model variables increase with a momentum factor." Another study indicated VIX by the Chicago Board Options Exchange as the world's premier barometer of investor sentiment (CBOE). Dash and Moran (2005) use it as a large signal of investor sentiment for hedge fund returns. We've used the VIX indicator for fear sentiments. The multiple proxies are used by different researchers to calculate sentiments. VIX can used as an indicator of volatility but it measure fear sentiment more accurately so the study uses VIX as sentiment indicator.

Besides, Banerjee et al. (2007) researched the association between VIX with portfolio market returns. The asymmetric return for S&P 100 returns is also recorded and indicates that negative stock index returns are more linked to higher VIX shifts than positive returns. As compared to Nasdaq 100, this asymmetric relationship in S&P 100 is considered to be higher. This helps one to discover the association or relationship between returns from VIX and S&P 50.Others concentrate on the interpretation aspect of misunderstanding. Da et al. (2015) create an index of uncertainty focused on the feelings and fears of investors.FEARS index is determined using data in text form from Internet searches. The FEARS index shows a higher market estimate of returns and uncertainty in the short term, according to the authors.

The CBOE implied VIX and used as a proxy and used for risks in the financial market. VIX, however, only captures business volatility as a market metric, Another research seeks to equate Volume of Google Search Indexes with VIX to illustrate the returns of the S&P 500. Finding shows that VIX is more vigorous stock market return indicator than the other one. In comparison, VIX has a more influential influence on all Google indices of its past values in the vector auto-regression model (Habibah et al., 2017). Behaviour finance is the association between investor sentiments and returns of the financial market. (Smales, 2017). The stock price and return are then calculated by its simple risk and the mispricing induced by the unreasonable sentiment of investors.

Tsai (2017) research examines the positive and negative sentiments of financial market of three main institutional investors and explores relationships and consequences of these forms of sentiments. To test whether investor feelings are infectious among stock owners, similar indices are first determined. Next, he investigated how each one of them diffused in the business, positive and negative emotions are distinguished. Finally, to research the diffusion impact of institutional investor sentiment under varying market performance, dynamic spillover of sentiment is calculated. The findings of his study confirm that the propagation impact of investor sentiment is negligible under good performance of financial markets, where institutional investor is positive. Impact of diffusion of negative sentiments, on the other hand, is important, implying that the contagion of investor sentiment is unstable (Tsai, 2017). The exchanges of investors on the stock market became quicker and convenient with the advent of the social network. Thus, consumer opinions that can impact decisions of their investment and it can be distributed and magnified rapidly across the stock market can be influenced to some degree. Data is taken from a popular China stock market specialist social networking platform named Xueqiu, which can then collected investor sentiment data by semantic analysis. Centered on the thermal optimal route (TOP) approach, suggested complex analysis of the relationship between sentiments of investor and the returns of stock market. The findings of this analysis show that the diffusion impact of the investor's feeling is negligible under favorable market performance if institutional investors are positive. The dissemination of the negative sentiment, on the opposite, suggests an asymmetrical contagion of an investor's sense of thought (Guo et al., 2017). Further the relation between investor sentiment, stock return, and volatility is the subject of many studies. Most papers predict a negative relationship to occur, as high sentiment is argued in one cycle to push markets up above their simple values, and a corresponding downward remedial market change should be noticed.

Ji et al. (2019) explored the spillover of information between WTI yield and investment sentiment indices calculated using the communication method in different trader positions. Their results suggested that the feeling is strongly linked to WTI returns by the form of trader. The spectator sentiment contributes much to the variance of WTI returns over the entire sampling duration among the various sentiment forms. The findings demonstrated that the influence of investor sentiments increase considerably as oil prices descend, in which hedger sentiment plays a major role in knowledge transfer. Sentiment analysis can collect information from multiple sources of content, such as ratings, news, and journals, and then interpret them based on their polarity (Hajiali, 2020). Further addresses how the fear of investors affects the dynamic price of Bitcoin during the pandemic coronavirus. They developed a Google searches based measure for proxy investors' high-frequency fear of coronavirus. Their findings suggest that an increase in the quest of pandemic is associated with an increase in stock market volatility. They investigate the effect of coronavirus fear on the returns of Bitcoin and the trading behavior of investors.

The rising concerns of Coronavirus contribute to the Bitcoin returns that are negative with high volume, suggesting that during a time of market distress Bitcoin is like other financial assets than conventional safe assets like gold. The study indicates that investors do not want to devote Bitcoin capital to decrease their exposure to risks since they cannot have a safe haven during a pandemic of coronavirus (Chen et al., 2020b). While several papers have started to research the nonlinear influence of investor sentiment on stock market returns in recent years, they prefer to use the technique of ordinary least squares regression analysis to study the relationship between the two where the average amount of market returns is shown (He et al., 2020). The quantile regression approach can more fully explain the distribution properties of investor sentiment relative to the ordinary least square regression and catch the effect of stock return on the tail distribution of sentiment (Song et al., 2020).

2.2 Forecasting Models

The growing availability of vast quantities of historical data and the need to predict future behavior correctly in a wide variety of experimental and applied fields involves the identification of robots and appropriate strategies that can infer stochastic dependency between past and future observations. From the 1960s on, linear mathematical techniques such as ARIMA models dominated the forecasting domain. Most recently, machine learning has gained interest and been serious rivals in the forecasting community for traditional predictive models. Then returns are forcasted from machine learning algorithms to do a comparison with the forecast of the GARCH model. In the forecasting of time series, the stock market prediction is a big problem. The stock market is exposed to considerable price fluctuations, implying that common equity investors are at high risk. Portfolio diversification allows particular business risks to be of. Machine learning methods are intended to understand and identify patterns automatically in vast quantities. In the literature, the classification method is complicated due to a wide range of machine learning techniques (Krollner et al., 2010).

In financial markets such as,risk management of funds and hedging techniques, volatility plays a key role. Accurate volatility forecasting is important. GARCH Model family is some models that are meant to illustrate the regularity and the capacity to represent the volatility of financial data and the theoretical and functional significance of fluctuations in time series. In the foreign currency market, the GARCH model Family can typically be applied to test problems such as stock efficiency (Dritsaki, 2018). The fat tails phenomena of a stationary alternative and fluctuating time, generally occurring in markets, is usually represented by the GARCH model Family in the stocks the model family GARCH plays a significant part in the stock market And this is because the family of GARCH models takes into consideration two essential characteristics of the financial data: kurtosis and volatility clustering (Menezes et al., 2019).

Gyamerah (2019) research assesses Bitcoin Returns uncertainty with three GARCH models. The latest litrature enables the modeling of the volatility effects in the Bitcoin return sequence, leptokurtic and distorted distribution. Gaussian distribution sufficiently traces the leptokurtosis and skewness in contrast with the T distribution. It analyzes capacity to estimate variance of the return sequence of Bitcoin for the period between 01 January 2014 and 16 August 2019 in various non-parametrically constructed GARCH models. Non-parametric approaches are extended to GARCH model forms since they do not presume distribution and can catch the kurtosis and fatty tails of the Bitcoin Return Sequence. Bouoiyour and Selmi (2016) studied a hybrid of ARIMA-GARCH model to estimate return volatility of prices of oil. Some other studies have used generalized models of GARCH to explain bitcoin volatility autoregressive conditional heteroskedasticity (Dyhrberg, 2016; Balcilar et al., 2017).

Katsiampa (2017) recently contrasted the fitness of six different GARCH models that have been normally distributed to evaluate bitcoin volatility. Nonetheless, all Bitcoin-style GARCH models, used in previous texts, presumed that inventions were spread in Gaussian, while financial returns typically indicate high tag distribution. The estimation of stock patterns was an interesting subject and scholars from numerous fields studied extensively. Machine learning has been researched widely for its ability to forecast financial markets, a well-established algorithm over a wide variety of applications. It has been stated that common algorithms are very successful in tracing the stock market and helping to optimize benefit from stock options while retaining the low risk. But in much of the above-mentioned literature, data on the same market are primarily obtained from the features chosen for inputs in masters learning algorithms. Such removal of critical information from other agencies makes the forecast more sensitive to local disruptions. Efforts have been made to crack the limits by adding external information through fresh financial news or personal web messages like Twitter. These methods, called sentimental research, refer to a range of main players or active market analysts' attitudes to interplay general investor opinion (Shen et al., 2012).

To forecast stock prices in science, Du et al. (2016) used the models for Bayesian learning (BL). This model is similar to ARIMA model. This model is the same. It learns the attributes of the stock series based on statistical knowledge. Tsantekidis et al. (2017) used a more advanced GBDT to forecast market prices in the conventional machine training sector. This model decide the features of the current stock series, but it is not sufficient to solve serial data problems such as stocks in the GBDT model layout itself. In truth, the BL model does not fit data itself in sequence. Further suggested in their study a model of predicting stock markets based on a CNN encoder. CNN is a highly efficient image input model. Use of encoder, first to encrypt sequence of information, and then CNN are used for training, for sequencing of data. Signal and system this approach is somewhat similar. Sequenced data may be used as time signal for filtering theory and CNN can be used as a convolution filter.

Bao et al. (2017) developed Special algorithm is with long short term memory(LSTM), based on repetitive neural networks. The arrangement of the neural unit makes it highly ideal for sequence data processing, such as inventories. In addition, this technique requires a research autocoder to encrypt the stock sequence and then uses the LSTM training network. Further experiments were based on the core models of deep learning and took note of and incorporated the fundamental features of the stock market. Zhang and Tan (2018) selected current stock selection method based on deep neural networks utilizes past market evidence to estimate projected stock returns rankings. Loewenstein (2000) have developed a system using a deep learning architecture to improve feature representation and use extreme learning machines to predict returns. They concluded that a deep learning feature and intensive learning machines would increase the quality of market effect predictions.

Li et al. (2020a) found that emotional vectors are generated by news item sentiment analysis and sentiment vectors are used to estimate stock prices in the LSTM model. Experiments on the Hong Kong capital exchange have shown positive results. The prediction techniques have the best forecast with lower prediction errors and higher prediction accuracy. In the time series forecast, there are various stochastic models. The most popular method in the single time series data combined in Auto-Regressive (ARMA) and Moving Average MA models is the univariate "Auto-Regressive Moving Average (ARMA)". "ARIMA" is a special kind of ARIMA, where the distinction in the model is taken into consideration. The model is a particular kind. Univariate "ARIMA" The other most popular projections models, which produce univariate ARIMA models and univariate Self-Regression (AR) models by enabling more than one variable to create, are multi-dimensional ARIMA modeling and vector auto-regression (VAR).

Machine learning methods and more recently deep learning algorithms developed new techniques for predictive issues, which are deep and succinct hierarchies in which associations between variables are modeled. In recent years there has been a great deal of interest in their application in a variety of areas including finance in machine-based methods of education such as support vector machines (SVMs) and random forests (RFs) as well as in deep-seated learning algorithms such as the Recurrent neuropathic network (RNN) and long-term memory. Profound learning methods may describe data structure and patterns such as non-linearity and ambiguity in prediction of time series. In addition, LSTM was used to predict time series and to measure the S&P 500's uncertainty in economic and financial performance.(Siami-Namini et al., 2018).

2.3 ML and Return Optimization

Furthermore, optimum asset weights are particularly vulnerable to expected sales estimates. In view of the ever uncertain forecasting of future anticipated returns, optimization will result in unstable sample weights. Noise in return equations in particular would deny all advantages of diversification. For example, DeMiguel et al. (2009) present a higher out-of-sample ratio than the optimally weighted portfolio of Markowitz and some additional optimal portfolios. Secondly, in Markowitz's principle, the estimation of the variance-covariance matrix requires a broad data collection, and the expectation of stable associations of asset returns. Other than that, as asset correlations increase, the matrix becomes predictable that happens during times when diversification is more important and much harder to achieve (De Prado, 2016). In comparison, AI tackles these concerns in two ways. Firstly, it can produce return and risk forecasts that are more accurate and can be used within traditional portfolio development structures than those that are generated by other methods. Second, AI techniques may provide alternative portfolio management approaches in order to create more accurate portfolio weights and custom portfolios that provide improved sample performance than typical linear technology.

The evolutions algorithms with the flexibility to solve more complicated assets allocation problems are another typical AI technique in portfolio construction. AI approaches can be used to perform advanced simple analytics, including text analysis, and to further enhance wealth allocation in financial portfolios. In spite of many challenges to conventional approaches of portfolio optimization, AI techniques often have simpler return and covariance estimates. The equations will then be used in the optimization of traditional portfolios. Furthermore, AI can be specifically used to render portfolios that meet performance targets more closely for asset allocation decisions (Liang et al., 2018).Further investigation indicated that a portfolio manager's decision demands that the funds be spread through a (large) spectrum of assets in a manner that the target portfolio reaches the target under such limitations (e.g., index imitation, Sharpe Ratio optimisation). The mean variance method of Markowitz typically lays down a theoretical base. With the growth of the capital market, the utility of the artificial subjective investment mode is steadily diminished due to the dynamic and varied investment priorities. Taking account of the advancement of data analysis and mathematical techniques, the former subjective investing mode has increasingly been replaced rapidly by a quantitative investment approach that uses data and simulations to create investment strategies. The modern investment model, the collection of inventories of investment value by integrating free-market knowledge with statistical techniques, eliminates the subjective influence of humans to some degree (Chen et al., 2020b). Snow (2020) analyzes Chinese A-share data as a research object from July 2014 until September 2017 and proposes a stock surplus-return projection method, which combines research reports and investor sentiments. Machine learning can be used to maximize stock returns by maximizing the distribution of properties. Reports relevant to computer science in investment management and more narrowly, machine learning strategies. Machine learning will also lead to many of the portfolio creation tasks including idea output, alpha factor design, asset collection, weight optimization, position scale and strategy monitoring.

2.4 Framework

2.5 Hypothesis Statements

H1: High fear sentiment has a negative impact on stock returns of developed markets.

H2: High fear sentiment has a negative impact on stock returns of emerging markets.

H3: Machine learning model outperform the econometric model in forecasting the developed markets

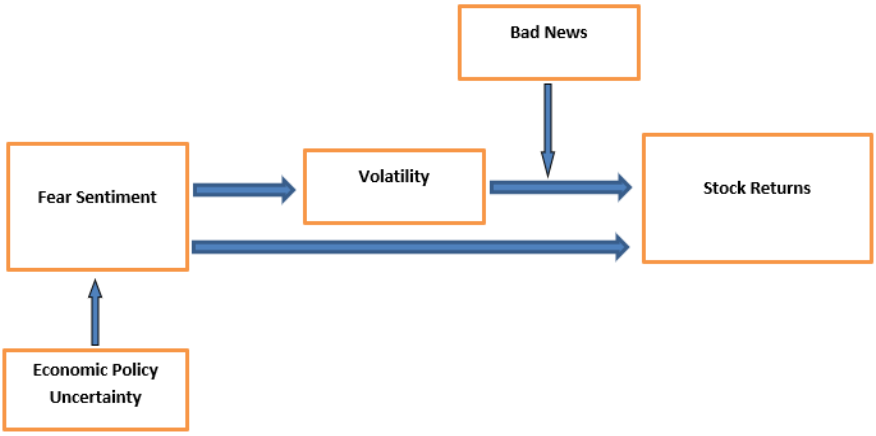


FIGURE 2.1: The impacts on the stock returns of the market and volatility due to the fear sentiment and the economic policy uncertainty

H4: Machine learning model outperform the econometric model in forecasting the emerging markets.

H5: Machine Learning algorithms outperform returns optimization in developed markets.

H6:Machine Learning algorithms outperform returns optimization in emerging markets.

Chapter 3

Research Methodology

This chapter is designed to discuss the details about the population and sample of study along with the empirical methodology to answer research questions and to achieve objective of study. The population and sample of the study have been discussed in first section. In the second section of the study, empirical methodologies along with data pre-processing and environment setting have been discussed.

3.1 Population and Sample of the Study

This study revolves around the investors investing in different stock markets around the globe. To achieve objectives data of the study has been divided into two major classes; developed and emerging class. Due to time limitations this study only considers six countries from both classes and a decision is made on basis of MSCI emerging and developed markets. Further, these countries are selected on their weight in the index presented on the MSCI website. For fear sentiments, multiple proxies data have been used. The economic policy uncertainty index has been used as a control variable for the fear sentiment which is also used by (Albulescu, 2021). For the fear, sentiments volatility index (VIX) is the preferred measure of sentiment in terms of enhancing model fit and to put in explanatory power (Smales, 2017). The VIX estimates the volatility of the market by applying the weighted values of SNP500 to a wide array of striking prices. In fact, the VIX is determined by considering the intermediate points of the SNP500 option offer and request rates in real-time. Which is also known as Chicago Board Options Exchange (CBOE volatility index). Daily prices has been used for all indices which converted into return by using the given formula:

$$R_t = \ln(P_t / P_{(t-1)}) \tag{3.1}$$

Whereas R_t denotes daily returns for given time t, ln represents the natural log, Pt refers as daily prices of respective indices at time t and $P_{(t-1)}$ refers to lagged prices of given indices.

For volatility, the study used standard deviation of past 10 days returns and multiply with the square root days in one single year. However, studies shows that VIX is not an efficient volatility predictor and does not give any extra information regarding future volatility (Becker et al., 2007). Bad news in market also have some relevancy with stock returns. Literature suggests that bad news decrease the return and elevates the volatility (Suleman, 2012). To capture this phenomenon this study used a dummy variable of bad news to capture the moderating effect of bad news on relationship of volatility and financial market returns. For dummy variable computation 1 has been used as a bad news with negative stock returns and 0 refers to no bad news which mean positive returns.We used big data frame of 20 years to provide enough data to machine learning model for trading.Bigger the data frame,better the training and more accurate the prediction will be.

3.2 Empirical Methodology

To achieve the objectives, this study follows the methodology of the two most recent studies of Albulescu (2021); Wang et al. (2020), which used some classical and machine learning methods. The study used 2SLS a system of equations to find the presence of fear sentiments, for forecasting purposes study used the classical GARCH model and Machine learning methods. For results comparison root mean square error has been used.

TABLE 3.1: Stock Market Indices

	Index Name	Country	Time Period
Developed	S&P-500	USA	Jan 2000 to Dec 2020
	Nikkei	Japan	Jan 2000 to Dec 2020
	FTSE-100	UK	Jan 2000 to Dec 2020
	CAC-40	France	Jan 2000 to Dec 2020
	Swiss Market Index(SMI)	Switzerland	Jan 2000 to Dec 2020
	S&P/TSX Composite Index	Canada	Jan 2000 to Dec 2020
Emerging	Shanghai Stock Exchange(SSE)	China	Jan 2000 to Dec 2020
	TWSE	Taiwan	Jan 2000 to Dec 2020
	KOPSI	South Korea	Jan 2000 to Dec 2020
	SENSEX	India	Jan 2000 to Dec 2020
	BOVESPA	Brazil	Jan 2000 to Dec 2020
	FTSE/JSE-40	South Africa	Jan 2000 to Dec 2020

FIGURE 3.1: Interactive Regression for Moderation

Dependent Variable: RT
Method: Least Squares
Date: 03/28/21 Time: 18:49
Sample: 1/01/2000 12/31/2020
Included observations: 7671

Variable	Coefficient	Std. Error	t-Statistic	Prob.
V	-0.038401	0.001021	-37.62362	0.0000
VBN_V	0.092044	0.001030	89.36660	0.0000
C	0.000736	0.000179	4.104187	0.0000

R-squared	0.510171	Mean dependent var	-8.10E-06
Adjusted R-squared	0.510043	S.D. dependent var	0.012092
S.E. of regression	0.008464	Akaike info criterion	-6.705564
Sum squared resid	0.549349	Schwarz criterion	-6.702848
Log likelihood	25722.19	Hannan-Quinn criter.	-6.704633
F-statistic	3993.224	Durbin-Watson stat	1.820060
Prob(F-statistic)	0.000000		

3.2.1 Interactive Regression Model

The significant values of the bad news variable show the moderating relationship with the Returns as shown in fig 3.1.

The significant results show that moderating effect of bad news is present on the relationship of volatility and returns. The interactive regression model also tells that there is no direct relationship or impact of bad news on returns in fact there is a moderating impact of bad news. The moderating relationship is not discussed prior in literature. The prior literature only discuss the direct effect of bad news on volatility. In times of bad news or high sentiments ,volatility is also high.

3.2.2 System of Equations

To measure the fear sentiment existence in stock markets study applies the system of equations as follows:

$$\vartheta_t = \alpha_\circ + \beta_1 \wp_{t-1} + \varepsilon_t \tag{3.2}$$

$$\varphi_t = \alpha_\circ + \beta_1 \vartheta_{t-1} + \varepsilon_t \tag{3.3}$$

$$R_t = \alpha_\circ + \beta_1 \varphi_{t-1} + \beta_2 \eta_{t-1} + \beta_3 \varphi_{t-1} \eta_{t-1} + \varepsilon_t \tag{3.4}$$

Whereas, ϑ refers to the fear sentiments (VIX) of the investor in the market, $\wp_{t-1}$ refers to the economic policy uncertainty (EPU) that exists in the market and t refers to time and t-1 refers to a lagged value. Whereas equation 3, φ_t refers to the volatility of that market. Equation 4 represents the moderation effect. Where R_t refers to the stock returns and η_{t-1} refers to the bad news impact. $\varphi.\eta$ referred as moderating(MOD) effect of bad news on relationship of volatility and financial market returns.

3.2.3 GARCH Model

GARCH model is best known for its nonlinear properties to handle volatility and volatile data. To measure the conditional lagged variance in given stock markets GARCH(1,1) model has been used in this study. Which is proposed by Engle (1982) initially as ARCH and later on Bollerslev (1986) improved and known as a generalized form of ARCH. GARCH use to measure the volatility effect in the market and persistence of that volatility in long run. In this study, we use GARCH to forecast our stock markets and written as follow:

$$K_t = \beta_\circ + \beta_1 K_{t-1} + \psi_t \tag{3.5}$$

$$\Im_t^2 = \beta_\circ + \beta_1 \psi_{t-1}^2 + \beta_2 \delta_{t-1}^2 \quad whereas,\ \beta_\circ > 0, \beta_1 > 0, \beta_2 \geq 0 \tag{3.6}$$

Equation 3.5 represents the mean equation of the GARCH model where K_t refers to returns of given series at time t and $K_(t-1)$ is the lagged returns and refers to the error term which is the uncaptured portion. The second equation of the model shows variance equation of model. Where ψ_{t-1}^2 denotes the ARCH effect and δ_{t-1}^2 denotes the GARCH term. Whereas, β_1 and β_2 refers to change in ARCH and GARCH.

3.2.4 Machine Learning Model

Machine learning methods are prominent algorithm nowadays which have been used by various institutions to forecast the financial time series. Over the past two decades, Artificial Intelligence(AI) emerged significantly. In literature, multiple studies used different machine learning algorithms to forecast and compare the results of these models. Studies show these models outperform the classical econometric models. This study not only used machine learning algorithms to forecast the financial time series but also used to optimize the returns. This study applies RNN (Recurrent Neural Networks) based LSTM (Long short term memory) machine-learning algorithm of Schmidhuber and Hochreiter (1997), to forecast the given time series and then applies ML to optimize the portfolio returns. For information processing, RNN use three gates forget gate as f, R as input gate, and p as output gate.

$$X_t = P_t O_{t-1} + y_t sigmoid(\omega_{cX} s_t + \omega_{jX} j_{t-1} + \mu_X) \tag{3.7}$$

Whereas X is the cell for time t, p is present forget gate with previous time period t-1. For the present token of input sequence c has been denoted with j hidden layer, ω is the weight, and μ is the offset amount.

$$P_t = \mu(\omega_{cj}c_t + \omega_{jP}j_{t-1} + \omega_{XP}X_{t-1} + \phi_P) \tag{3.8}$$

Where P denotes for input gate at t time, j for the hidden layer, for the update of information k has been applied and l_t used as input gate written as:

$$K_t = \delta(\omega_{ck}c_t + \omega_{jk}j_{t-1} + \omega_{Xk}X_{t-1} + \phi_k) \tag{3.9}$$

For the output, gate has been applied and updated information will as:

$$l_t = \delta(\omega_{cl}c_t + \omega_{jo}j_{t-1} + \omega_{Xl}X_{t-1} + \phi_l) \tag{3.10}$$

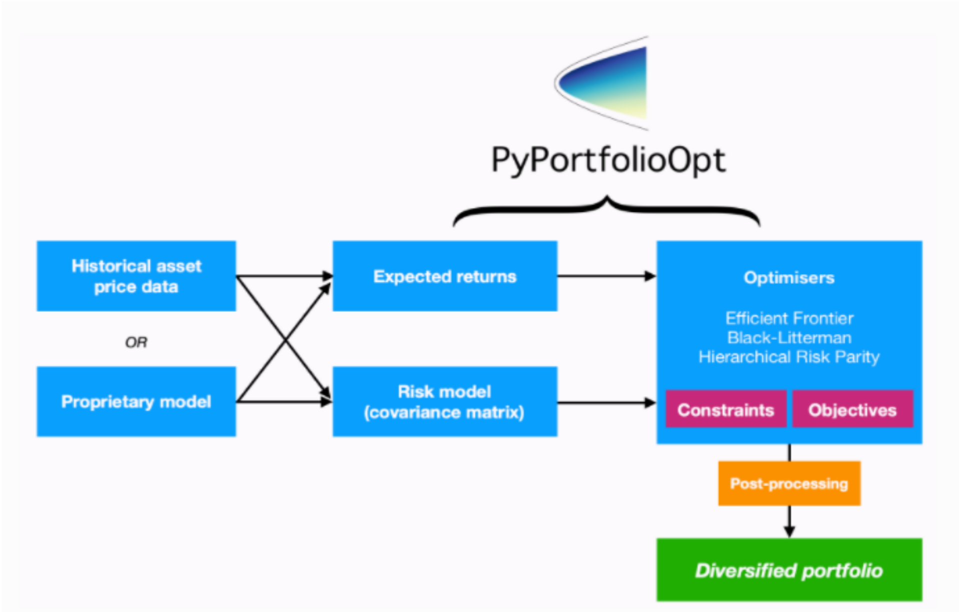
Now the hidden layer output will be as:

$$j_t + l_t sigmoid(X_P) \tag{3.11}$$

3.3 Data Pre-Processing and Environment Setting

Data pre-processing is the essential part of data analysis. To make our data more reliable, this study uses some preprocessing tasks to improve and validate the testing. The study used the ARCH LM test for heteroscedasticity and all of the indices show a significant existence of heteroscedasticity. This allows us to use the GARCH model for forecasting. Similarly, the Minimax scaler has been used to scale our data for machine learning. Python has been used as a programing language with google collab as IDE for python. Some libraries i.e., Numpy, Pandas for data fram, matplotlib for data visualization, Keras and TensorFlow for machine learning, and pyportfolioopt for portfolio optimization our returns.

FIGURE 3.2: Portfolio Optimization



3.4 Portfolio Optimization Methodology

PyPortfolioOpt is made by keeping modularity in mind; the below flowchart shows the functionality and layout of PyPortfolioOpt. This will tells the working of this optimizer. This portfolio optimizer is used to optimize the returns.

3.4.1 Processing historical prices

There are two things which are needed in Mean-variance optimization first is the expected returns of assets, and second one is the covariance matrix or we can say a risk model which will tells us the risk of the asset. PyPortfolioOpt has the methods for estimation of both expected returns and risk models.

3.4.2 Mean-Variance Optimization

Harry Markowitz wrote his classic paper on mean-variance optimization in 1952. This paper changed the concept of portfolio management from an art to the concept of science. Now people started considering it as science, which is more logical and complex then arts. The key secret is that by making combination of assets with different volatiles and expected returns, one can allocate theirs resources. This optimal allocation is made mathematically. If w is the weight vector of stocks with expected returns, then the portfolio return is equal to each stock's weight multiplied by its return. The portfolio risk in terms of the covariance matrix is given. Portfolio optimization can then be regarded as a convex optimization problem, and a solution can be found using quadratic programming. If we denote the target return , the precise statement is given of the long-only portfolio optimization problem. If we change target return, different set of weights is found (i.e a different portfolio) – the collection of these optimal portfolios is called as the efficient frontier. The Sharpe ratio is the portfolio's return in excess of the risk-free rate, per unit risk (volatility).

$$SR = \frac{R_p - R_f}{\sigma} \tag{3.12}$$

It is particularly important because it measures the portfolio returns, adjusted for risk. So in practice, rather than trying to minimize volatility for a given target return as per Markowitz Markowitz (1952), it often makes more sense to just find the portfolio that maximizes the Sharpe ratio. This is implemented as the max Sharpe () method in the Efficient Frontier class. Using the series mu and data frame s. There are two things which are needed in Mean-variance optimization first is the expected returns of assets, and second one is the covariance matrix or we can say a risk model which will tells us the risk of the asset. PyPortfolioOpt has the methods for estimation of both expected returns and risk models.

3.4.3 Short Positions

Short position is the selling position in which the investor borrow the shares of a stock which will decrease in the value as believed by the investor.If we allow taking the short position,we simply start with the negative weights. Through this the neutral portfolios can be generated but due to mathematical reasons these are only available for efficient risk() and efficient return() optimization methods. Pass market neural to be true if you want the a market neural portfolio.

3.4.4 Black-Litterman Allocation

For asset allocation the Black-Litterman(BL) model took an approach called as Bayesian approach. In this approach we combines the prior estimate of returns with views on assets to create an estimate of expected returns.The advantages of this approach will be:

• We can have the views on the subset of assets and BL will take it further by considering the covariance with other assets.

• By using this Black Litterman expected returns ,the results became much more stable. The portfolio formed by this approach are more stable than the ones formed by mean-historical returns.

3.4.5 Output of the BL model

The outputs of this BL model is the covariance metrix and the other output is the posterior estimates of the returns.Now it is suggested that we should all of this BL output into the optimizer. In PyPortfolioOpt, this is available under Black Litter man Model. The Black Litterman Model originates from Base Optimizer so the API is the same as that of the Efficient Frontier. The module of BL contain this BL Model class, which contains the posterior estimates of prior estimates of returns along with the views provided by the investor. The number,type of investment and the amount of investment is decided by the investor. In addition, the functions that we uses calculate the following:

• prior estimate of expected returns based on markets

• parameter of risk-aversion market based

3.4.6 Other Optimizers

There are other optimizer that can also be used. These optimizers also has the access to same API for post and pre-processing.One is Hierarchical Risk Parity Algorithm, other is the Critical Line Algorithm. Both of these optimizer algorithms along with mean-variance model is present in the module of PyPortfolioOpt.

3.4.7 Hierarchical Risk Parity Algorithm

Hierarchical Risk Parity is a novel portfolio optimization method developed by Marcos Lopez de Prado De Prado (2016) . The working of HRP is as follows:
1. A distance matrix is formed based correlation of the assets and it selected from the collection of assets.
2. Tree cluster assets is formed with the help of distance matrix through hierarchical clustering.
3. Minimum variance portfolio is formed on each of the branch of tree.
4. Through iteration at each step mini-portfolios are optimized at every node. There is no covariance matrix as in the mean-variance optimization. There is no requirement of making the covariance matrix which is the plus point of the HRP. The other advantage of HRP is that diverse portfolios can be made that can perform out of sample.

3.4.8 The Critical Line Algorithm

There is another optimizer that can also be used that is called as The Critical Line Algorithm for optimization. This is very useful in apply the linear inequalities. CLA is also designed for portfolio optimization. After some iteration it will definitely converge and is able to derive the efficient frontier.

3.4.9 Selection of Model

We choose the Mean-Variance optimization from PyPortfolioOpt and not the Hierarchical Risk Parity Algorithm and the Critical Line Algorithm. The reason is Mean-Variance optimizer allow the investor to select the assets and index of their own and also the budget of the investment. The prior and posterior comparison of expected return is also very useful in deciding the correct allocation of the assets. We also choose mean-variance model because its Black Litterman Model allows the investors to have the views on the subsets of the stocks,which is missing in others.Another feature of this model is the presence of covariance matrix,which is only present is this. That is why study selected Mean-Variance Model.

Chapter 4

Results and Discussion

4.1 Descriptive Statistics

Table 4.1 represents the Descriptive Statistics of sample data. Mean values for the developed markets shows the average returns for the given time frame. Results show 0.132804, -9.22E-05, -8.10E-06, 4.79E-05, 4.52E-05, and 9.58E-05 as average returns for SNP, FTSE, CAC, Nikkei, SMI, and TSX respectively. Data shows FTSE, CAC both have negative returns. Standard deviation from the mean values of given markets are 0.101816, 0.002526, 0.012092, 0.012219, 0.009673, and 0.009429 respectively. Which shows the dispersion of stock market returns from the average returns. S&P shows the highest dispersion with 10% of standard deviation value. Statistics show the minimum returns for respective markets are -0.127652, -0.044362, -0.130983, -0.12111, -0.101339, and -0.131758. With maximum returns of 0.109572, 0.186043, 0.105946, 0.132346, 0.107876, and 0.112945 respectively for each developed market. Skewness shows the asymmetry of the returns. In our data set of all the values of the skewness are negative except S&P and FTSE (3.293517 and 51.61873) respectively. Negative values show the left-skewed returns. Positive values show the right-skewed returns. Kurtosis shows the tailed ness of the data. All the values of indices are greater than 3 which show the fat tails of the market returns which indicate the leptokurtic behavior of data.

Table 4.1: Descriptive Statistics

	Mean	Min.	Max.	Std. Dev.	Skew*	Kurt*	J.B*
Developed							
SNP	0.132804	-0.1277	0.1096	0.1018	3.2935	20.8572	115790.1
FTSE	-9.22E-05	-0.0444	0.1860	0.0025	51.6187	3872.376	4.79E+09
CAC	-8.10E-06	-0.1310	0.1059	0.0121	-0.2411	13.3172	34096.37
Nikkei	4.79E-05	-0.1211	0.1323	0.0122	-0.4663	13.9776	38795.30
SMI	4.52E-05	-0.1013	0.1079	0.0097	-0.3513	15.6388	51214.12
TSX	9.58E-05	-0.1318	0.113	0.0094	-1.1175	28.7049	212787.2
Emerging							
Sensex	0.0003	-0.141	0.1599	0.0121	-0.4282	18.0590	72716.97
BOVSEPA	0.0003	-0.1599	0.1368	0.01496	-0.4159	14.0544	39278.87
KOPSI	0.0001	-0.1281	0.1128	0.0123	-0.681	14.8441	45430.38
SSE	0.0001	-0.0926	0.0940	0.0126	-0.4412	12.0668	26523.96
JSE	0.0004	-0.1045	0.0791	0.0132	-0.2478	7.5387	4566.051
TWSE	6.74E-05	-0.0994	0.0653	0.0109	-0.3626	10.2378	16912.21
VIX	2.9121	2.2126	4.4151	0.3774	0.6959	3.3518	658.79

**Skew = Skewness, *Kurt = kurtosis, *J.B = Jargue-Bera*

For emerging markets, descriptive statistics again depicted to show the data summary in Table 4.1. In emerging markets mean values show the average returns for Sensex, BOVSEPA, KOPSI, SSE, JSE, and TWSE are 0.000285, 0.000255, 0.000130, 0.000119, 0.000366, 6.74E-05 respectively. With minimum returns of -0.141017, -0.15993, -0.128047, -0.092561, -0.104504, and -0.09936. Data shows emerging markets delivered well in terms of average market returns with no negative return over the selected time frame. Similarly, the Standard deviation from the mean values of given emerging markets is 0.012108, 0.014959, 0.012274, 0.012614, 0.013236, and 0.010938 respectively. Which shows the dispersion of stock market returns from the average returns. BOVESPA shows the highest dispersion with a 1.4% of standard deviation value. Skewness shows the asymmetry and location of the returns. Data shows all of the markets are left-skewed. Negative values show the left-skewed returns. Kurtosis shows the tailedness of the data. All the values of indices's are greater than 3 which show the fat tails of the market returns which indicate the leptokurtic behavior of data. Jarque-Bera statistics show the non-normality of data exists in our sample data. This means that data of our sample is not normal.We check the descriptive statistics of all the indices's of close prices. Figure 4.1 and 4.2 represent the graphical outlook of data and shows the descriptive statistics.

4.2 Fear Sentiments and Market Returns

Table 4.2 represent the estimated results for the fear sentiments of markets returns for both developed and emerging markets. For the S&P, FTSE, CAC, NIKKEI, SMI, and TSX markets coefficient values of EPU 13.3499, 12.2736, -0.2423, 13.3754, -15.8146, and 2.4125 respectively, have significant positive relationship for S&P, FTSE, NIKKEI and SMI and inverse for rest of developed stock markets. Which shows EPU has significant impact on VIX that mean change in EPU affect the VIX in respective directions. Similarly estimated results of VIX shows a significant positive impact on volatility of respective markets. Coefficient values for respective markets are 0.2104, 0.0064, 0.2016, 0.4109, 0.1515, and 0.1709. Which indicates that one percent change in VIX change the volatility by the 21%, 0.6%, 20%, 41%, 15%, and 17% respectively for each stock market. VoL represents the direct impact of volatility on returns of respective markets, and results shows that increase in volatility decreases the overall market returns. Estimated coefficient values -0.0381, -0.0090, -0.0380, -0.0368, -0.0364, and -0.0395 shows a significant inverse relation betwee volatility and stock market returns for respective sample markets. BDN results shows insignificant relation of bad news on markets returns except FTSE. Further results of MoD show a significant moderating relation with coefficient values 0.0939, 0.2296, 0.0910, 0.0897 , 0.0932, and 0.0917 respectively for each sample developed markets.

Results shows that EPU has significant positive relationship with VIX and coefficient values are 2.4530, 1.9102, 10.3200, 10.7781, 3.1517, and 3.1517 for emerging markets SENSEX, BOVESPA, KOPSI, JSE, TWSE respectively. Which shows EPU has significant impact on VIX that mean change in EPU affect the VIX in respective directions. Similarly estimated results of VIX shows a significant positive impact on volatility of respective markets. Coefficient values for respective markets are 0.1464 , 0.1690, 0.1786, 0.0696, and 0.1454 for SENSEX, BOVESPA, KOPSI, TWSE.

Figure 4.1: Descriptive Statistics (Developed Markets)

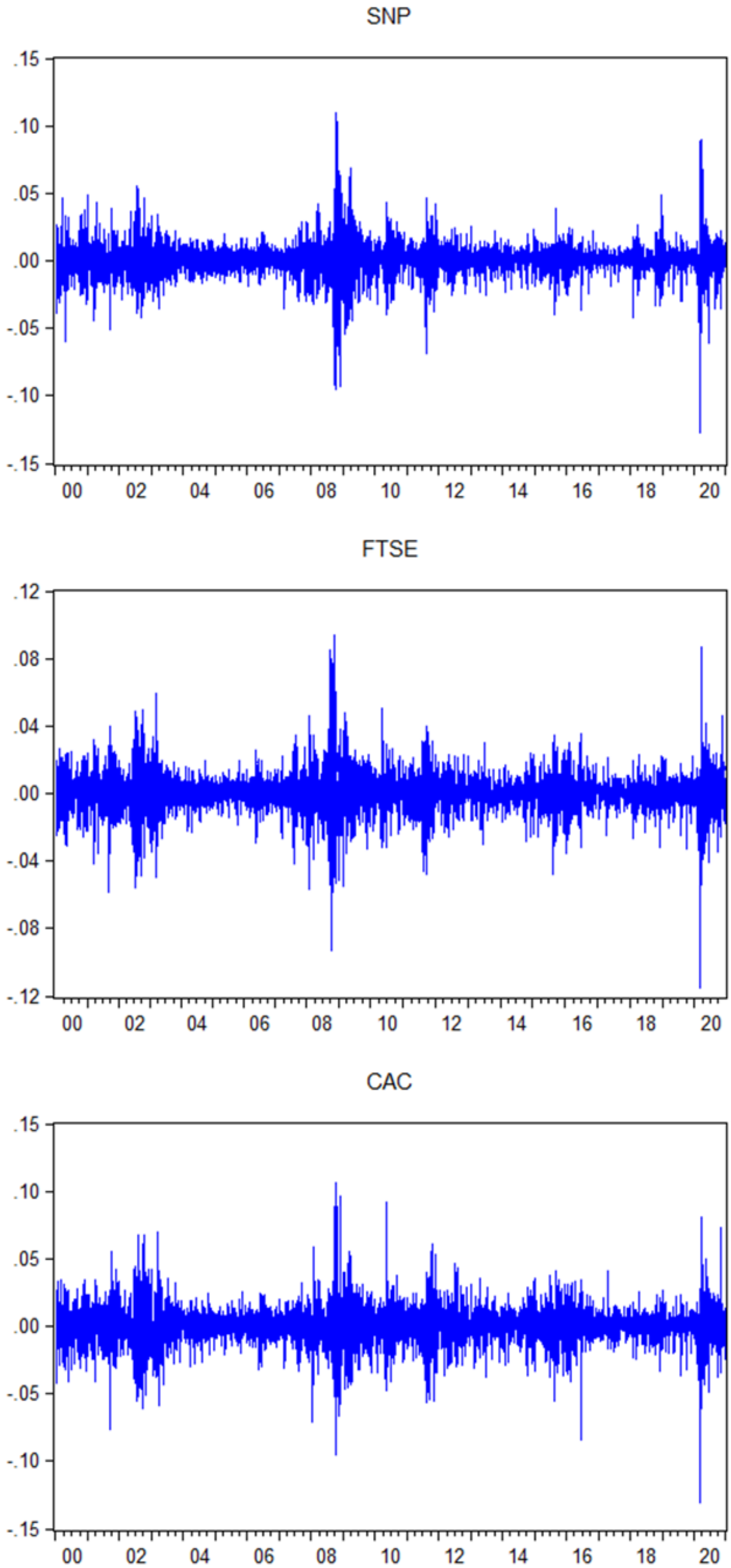


FIGURE 4.2: Descriptive Statistics (Developed Markets)

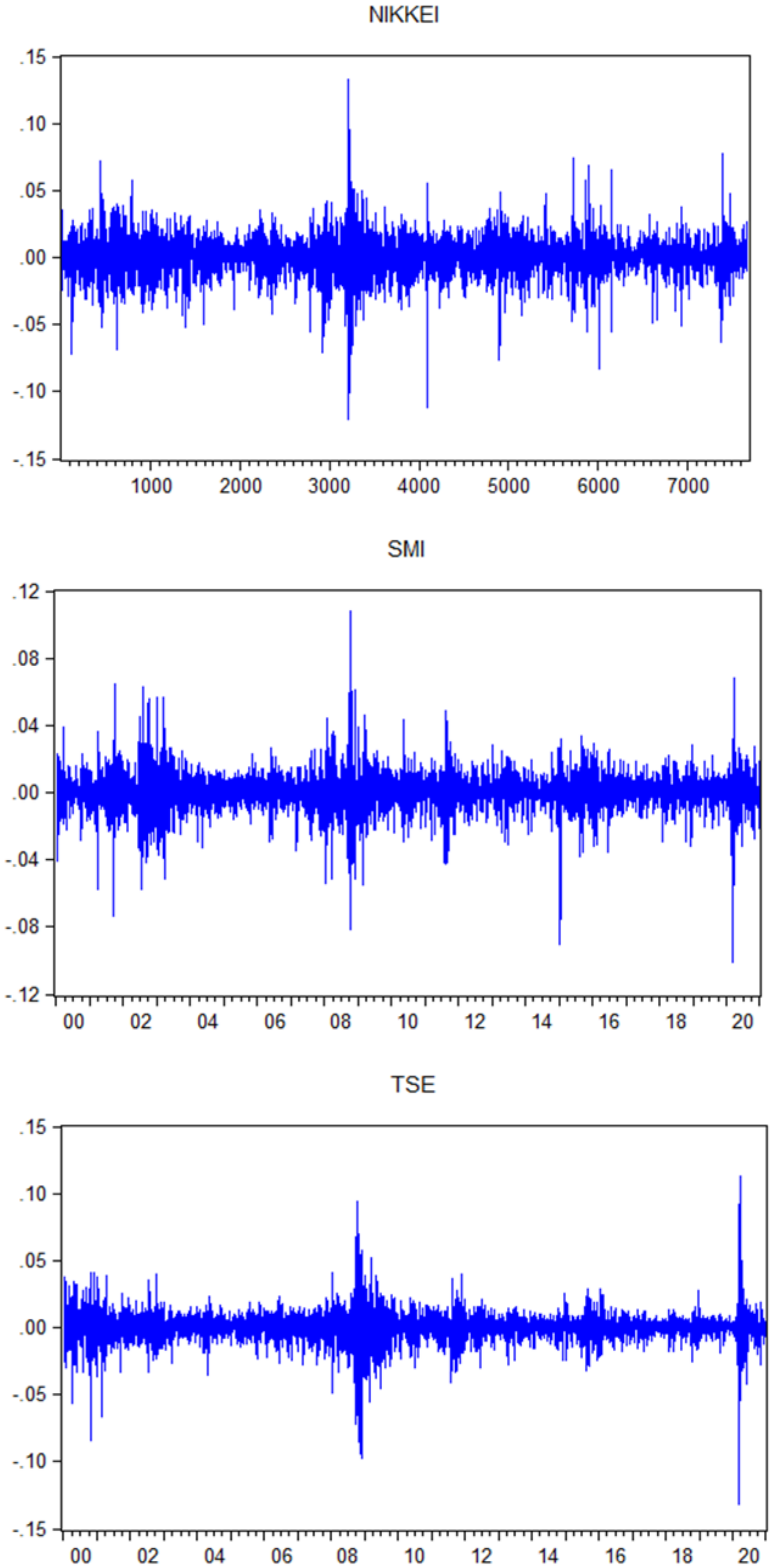


FIGURE 4.3: Descriptive Statistics (Emerging Markets)

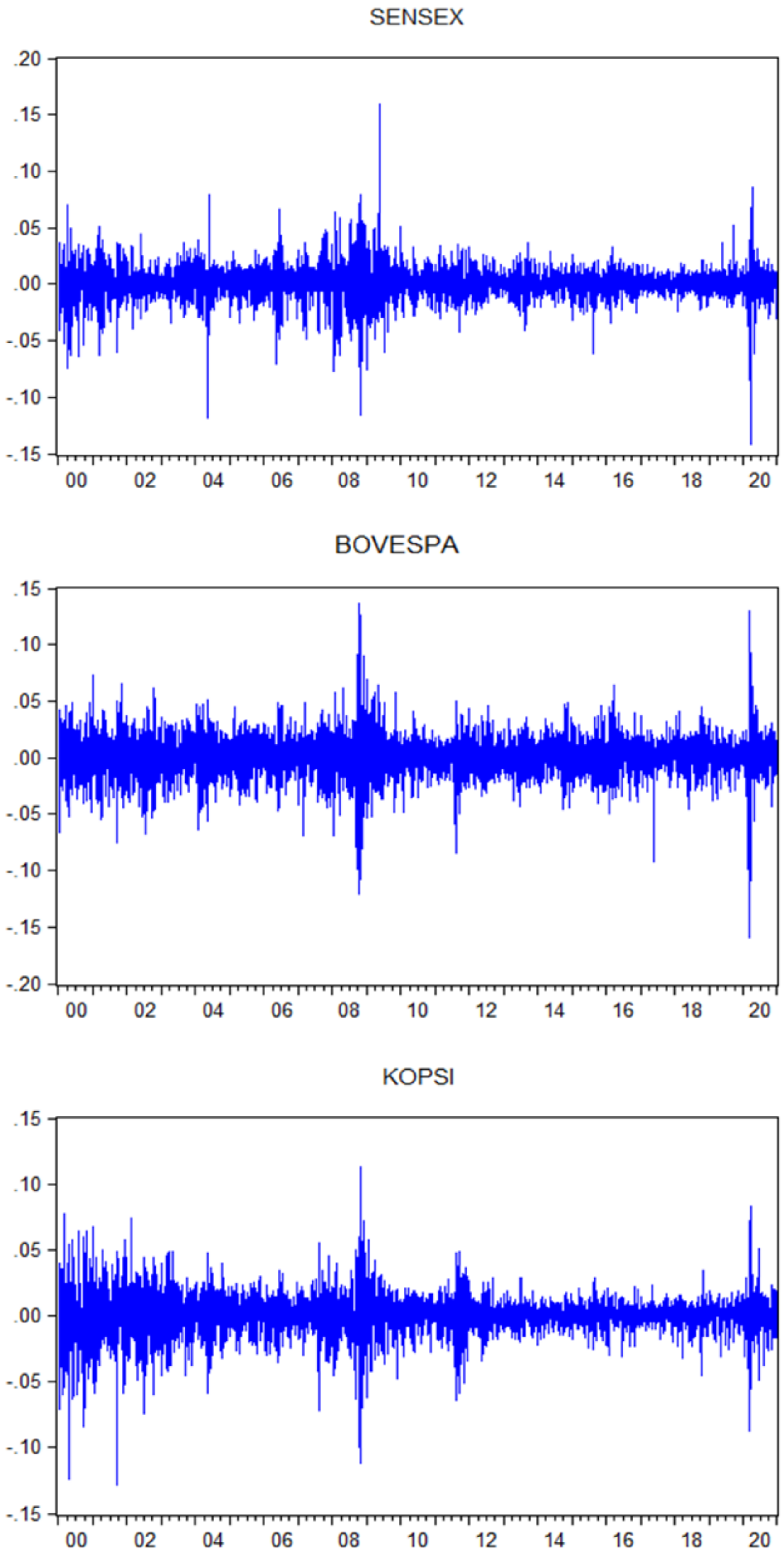


FIGURE 4.4: Descriptive Statistics (Emerging Markets)

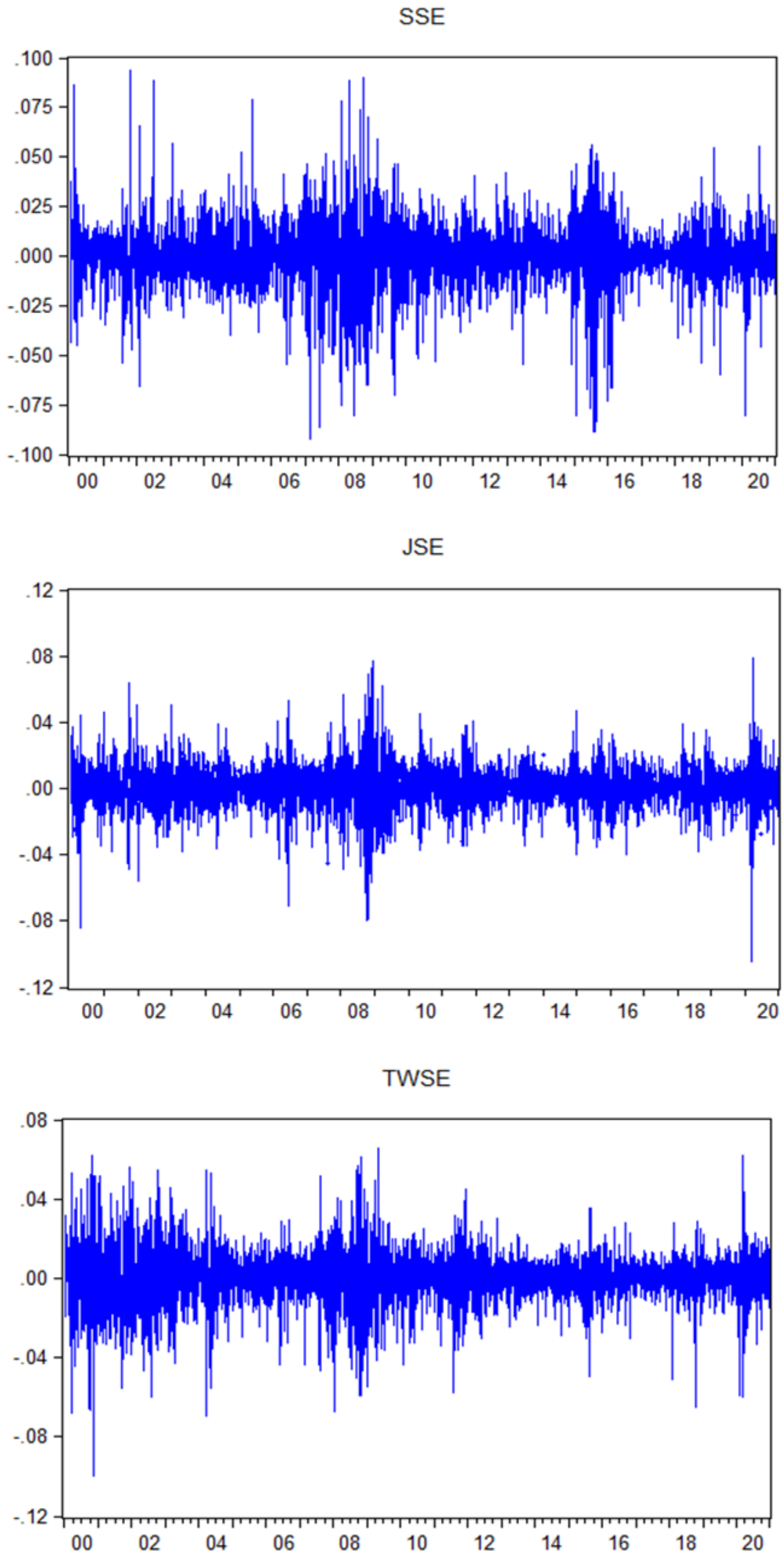


Table 4.2: Fear Sentiments and Stock Market Returns

	S&P	**FTSE**	**CAC**	**NIKKEI**	**SMI**	**TSX**	**SENSEX**	**BOVESPA**	**KOPSI**	**JSE**	**SSE**	**TWSE**
EPU	13.3499	12.2736	-0.2423	13.3754	-15.8146	2.4125	2.4530	1.9102	10.3200	10.7781	3.1517	3.1517
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.0022]	[0.0001]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
VIX	0.2104	0.0064	0.2016	0.4109	0.1515	0.1709	0.1464	0.1690	0.1786	-0.0718	0.0696	0.1454
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
VoL	-0.0381	-0.0090	-0.0380	-0.0368	-0.0364	-0.0395	-0.0345	-0.0391	-0.0375	-0.0503	-0.0372	-0.0362
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
BDN	-0.0004	-0.0303	0.0002	0.0004	-0.0001	0.0001	0.0004	-0.0001	0.0005	0.0007	0.0007	0.0003
	[0.1802]	[0.000]	[0.5141]	[0.2825]	[0.7575]	[0.7225]	[0.3268]	[0.7834]	[0.1691]	[0.1258]	[0.0748]	[0.4184]
MoD	0.0939	0.2296	0.0910	0.0897	0.0932	0.0917	0.0912	0.0941	0.0876	0.0992	0.0854	0.0896
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]

JSE results shows a significant inverse relation of VIX and Vol with coefficient value of -0.0718. VoL represents the direct impact of volatility on returns of respective markets, and results shows that increase in volatility decreases the overall market returns except JSE.Estimated coefficient values -0.0345, -0.0391, -0.0375, -0.0503, -0.0372, and -0.0362 shows a significant inverse relation between volatility and stock market returns for respective sample markets. BDN results shows insignificant relation of bad news on markets returns. Further results of MoD show a significant moderating relation with coefficient values 0.0912, 0.0941, 0.0876, 0.0992, 0.0854, and 0.0896 respectively for each sample emerging markets.The impact of one variable on the other is calculated in table 4.2 is exactly according to the relationships of model in figure 2.1.

4.3 Forecasting Methods

This section of the analysis is consisting of three sessions, econometric GARCH model estimation, machine learning estimation, and forecasting comparison of these models.

4.3.1 GARCH Model Estimation

Results in Table 4.3 represent the GARCH model estimation for all sample market indices. Lagged values of S&P, FTSE, NIKKEI, and SSE shows a significant relationship with returns of respective markets at 95% confidence interval. This gives the information, that returns of respective markets can be predicted by using the lagged returns. Lagged values of CAC, SMI, TSX, shows an insignificant relation with returns and p-values are 0.2768, 0.3909, and 0.5319 respectively at 95% confidence interval in respective markets. So, returns of CAC, SMI and TSX can't be predicted by using lagged returns.

Similarly, ARCH of SNP, FTSE, CAC, NIKKEI, SMI and TSX has a significant and positive coefficient value that are 0.070654, 0.104473, 0.15, 0.057364, 0.075491, and 0.063341 respectively in developed markets. This confirms that the behavior of previous prices affects the present volatility. Then we further consider the co-efficient values of the GARCH in the developed markets.In developed markets GARCH coefficient values of SNP, FTSE, CAC, NIKKEI, SMI and TSX are 0.919998, 0.911695, 0.6, 0.931473, 0.90986 and 0.933425 respectively with statistically significant p-values, that shows the persistence of volatility in current prices. Sum of ARCH and GARCH coefficient for respective developed markets is near to one which shows the persistence of volatility in long run.

Lagged values of SENSEX, BOVESPA, KOPSI, JSE, TWSE shows an insignificant relation with returns and p-value are 0.3525, 0.0817, 0.9315, 0.0974, and 0.9140 respectively at 95% confidence interval for emerging markets.This tells us that the returns of respective market cannot be predicted by using the lagged returns. Whereas ARCH of SENSEX, BOVESPA, KOPSI, SSE, JSE, TWSE have a significant and positive coefficient value of 0.0574, 0.043052, 0.041686, 0.039037, 0.097228 and 0.040439 respectively. This confirms that the behavior of previous prices affects the present volatility. In emerging markets GARCH coefficient values for respective sample markets are 0.936931, 0.94112, 0.955954, 0.95766, 0.88697 and 0.954523 with statistically significant p-values, that shows the persistence of volatility in current prices. Sum of ARCH and GARCH coefficient for sample emerging markets is near to one which shows the persistence of volatility is long run in nature.

4.3.2 Machine Learning Model Estimation

Some estimation functions are the key part of machine learning and its performance i.e., neurons, batch sizes, loss function, and many others. For results estimation environment setting is an essential part of machine learning analysis. This study uses python as a programing language for analysis. Different libraries like KERAS as Tensorflow interface for neural networks. Pandas for data analysis, Numpy for numeric, minimax Scaler to scale the data from 0-1. Matplotlib is used for data plotting and data visualization, PyPortfolioOpt library for portfolio optimization. Machine training plays a vital role in its performance of the model. For training

Table 4.3: GARCH Estimation

	Developed Markets					
	S&P	**FTSE**	**CAC**	**Nikkei**	**SMI**	**TSX**
Model	(1,1)	(1,1)	(1,1)	(1,1)	(1,1)	(1,1)
AIC	-6.3800	-14.8000	-5.8800	-6.2151	-6.8237	-7.0784
Lag	-0.0600	0.1500	-0.0300	-0.0259	0.0103	0.0072
	[0.0000]	[0.0000]	[0.2768]	[0.0332]	[0.3909]	[0.5319]
ARCH	0.0700	0.1000	0.1500	0.0574	0.0755	0.0633
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
GARCH	0.9200	0.9100	0.6000	0.9315	0.9099	0.9334
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
			Emerging Markets			
	Sensex	**BOVSEPA**	**KOPSI**	**SSE**	**JSE**	**TWSE**
Model	(1,1)	(1,1)	(1,1)	(1,1)	(1,1)	(1,1)
AIC	-6.3782	-5.7880	-6.3940	-6.1770	-6.0893	-6.5133
Lag	0.0114	-0.0200	0.0011	-0.0369	0.0251	0.0013
	[0.3525]	[0.0817]	[0.9315]	[0.0015]	[0.0974]	[0.9140]
ARCH	0.0574	0.0431	0.0417	0.0390	0.0972	0.0404
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
GARCH	0.9369	0.9411	0.9560	0.9577	0.8870	0.9545
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]

window study used 80 percent of the sample data and for the testing study used the remaining 20 percent of the data. The reason for this ratio is that the more data you train, the better the results are predicted. This study uses the two layers of neurons with 50 neurons in each layer and 25 dense neurons in the first layer while 1 dense neuron in second layer. For optimizer, function study uses the ADAM as loss function estimator from the KERAS library. Mean square error is used as loss computation and comparison. The batch number has been set as 1 with 1 epoch for iteration purpose. Evaluation intervals, buffer size, and activation function are used as default.

4.3.3 Models Forecasting

Forecasting was done by both of the models that is GARCH and ML. Both can be used for forecasting but the objective of the study is to find a model that has more accuracy. So by using that model to forecast accurately and gain more return which is the ultimate objective of this study. This model forecasting is crucial because it helps to achieve the objective of returns optimization in the future.

S&P500 estimated accuracy indicator for GARCH is RMSE (700.1825) and for ML this value is 57.8967. For FTSE RMSE value for GARCH is 1360.170 and 3.7709 for ML of the same market. GARCH estimated accuracy indicator for CAC is 1494.159 whereas the value is 1.3 for ML for the respective market. NIKKEI estimated RMSE for GARCH is 5499.784 whereas RMSE 9.8250 for ML. The accuracy estimator of GARCH for SMI is 1820.422) whereas ML RMSE is 56.7880. TSX RMSE indicator for GARCH is 6111.605 and for ML RMSE is 1.9425. Figure 4.1 present the actual and forecasted prices for machine learning method for developed markets. In figure blue color represent the training window of machine, whereas red color shows the actual prices and yellow color represent the forecasted prices in testing window of machine.

SENSEX indicator for estimated accuracy for GARCH is 17133.28 on the other hand RMSE (515.9050) for ML. GARCH estimated accuracy indicator for BOVESPA is RMSE (17822.93) whereas 2704.6588 for ML. KOPSI also uses the same indicator with a value of 666.7009 for GARCH and 96.1898 for ML. JSE estimator of accuracy value for GARCH is 22414.07 and for ML 690.1145. SSE market index also depicts a difference between the two values. The GARCH estimator for accuracy is 1048.091 whereas RMSE for ML is 36.0843 for the respective market.

TWSE accuracy estimator for GARCH is RMSE (5325.208) and the RMSE value for ML is 75.2248 for the respective market. There are some markets where ML has also given a relatively high value of indicator but even then, that value is much less than the value of GARCH. There is no single market that has a higher RMSE in ML and lowers in GARCH estimation. Figure 4.2 present the actual and forecasted prices for machine learning method for emerging markets. In figure blue color represent the training window of machine, whereas red color shows the actual prices and yellow color represent the forecasted prices in testing window of machine.

Based on the above results the model that outperforms the forecasting is ML because it has less RMSE values as compared to Econometric model GARCH. The

Figure 4.5: GARCH Model Forecasting(Developed Markets)

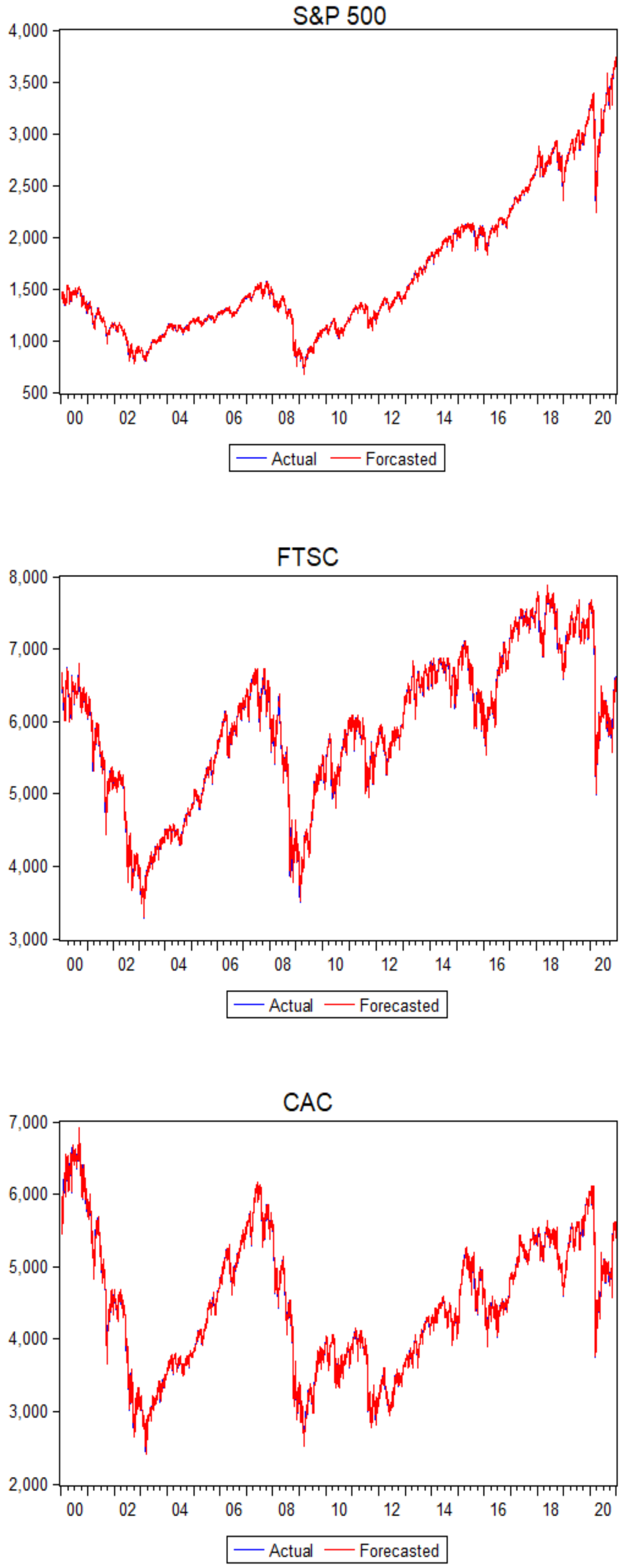


FIGURE 4.6: GARCH Model Forecasting(Developed Markets)

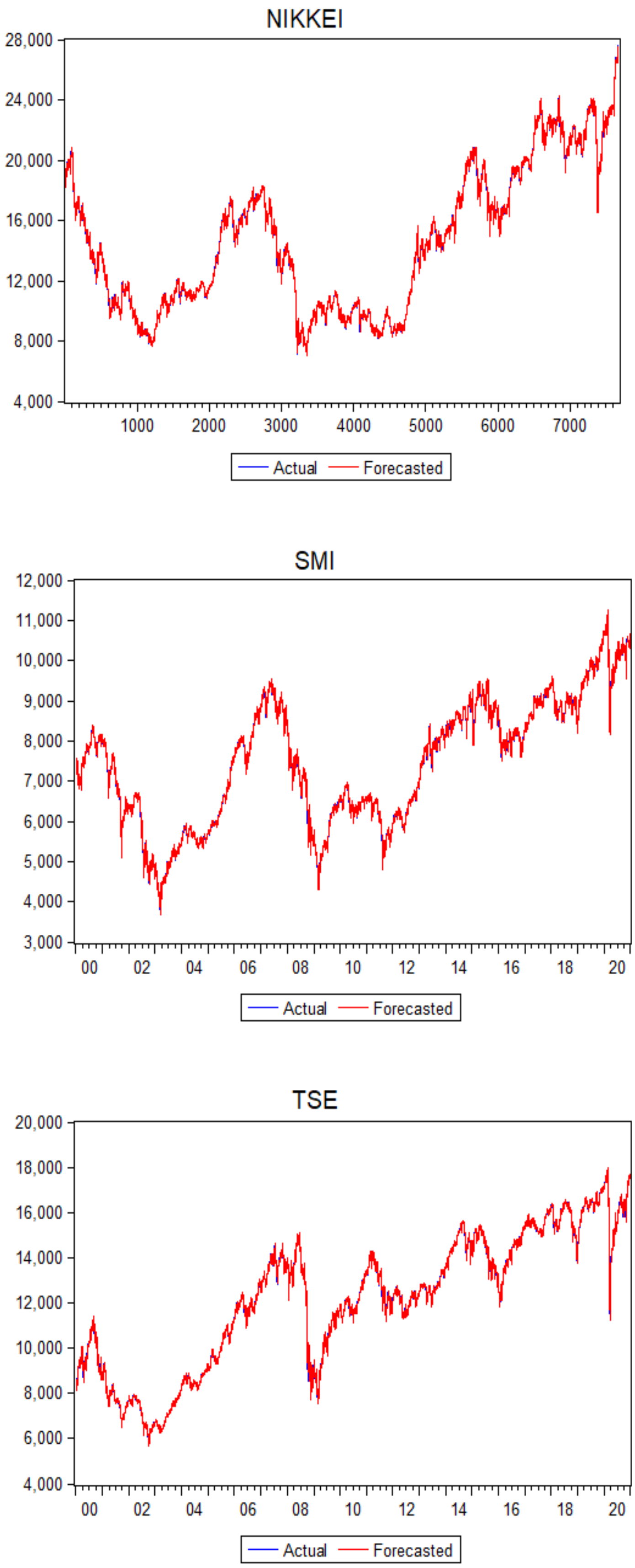


FIGURE 4.7: Machine Learning Forecasting(Developed Markets)

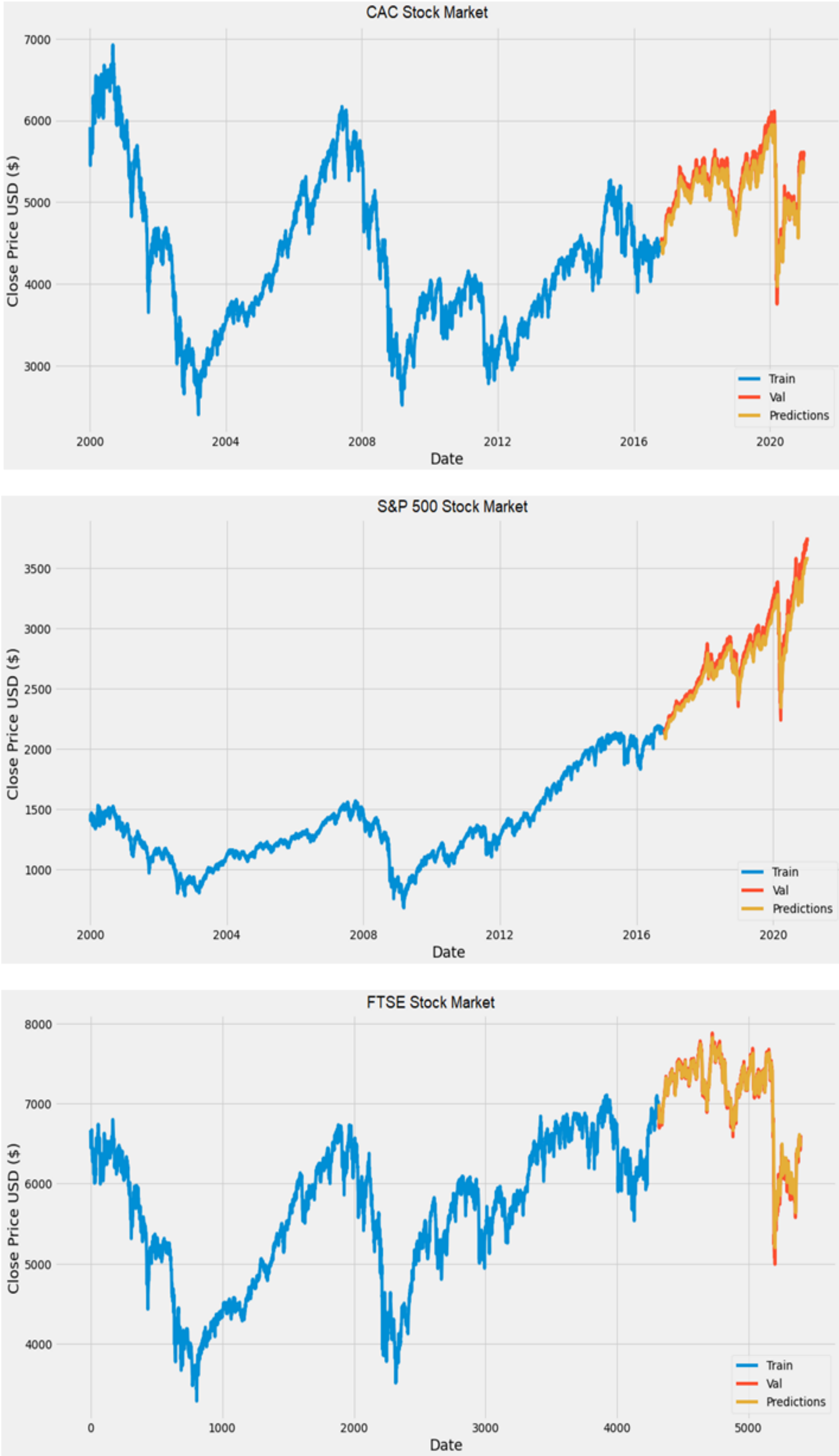


FIGURE 4.8: GARCH Model Forecasting(Developed Markets)

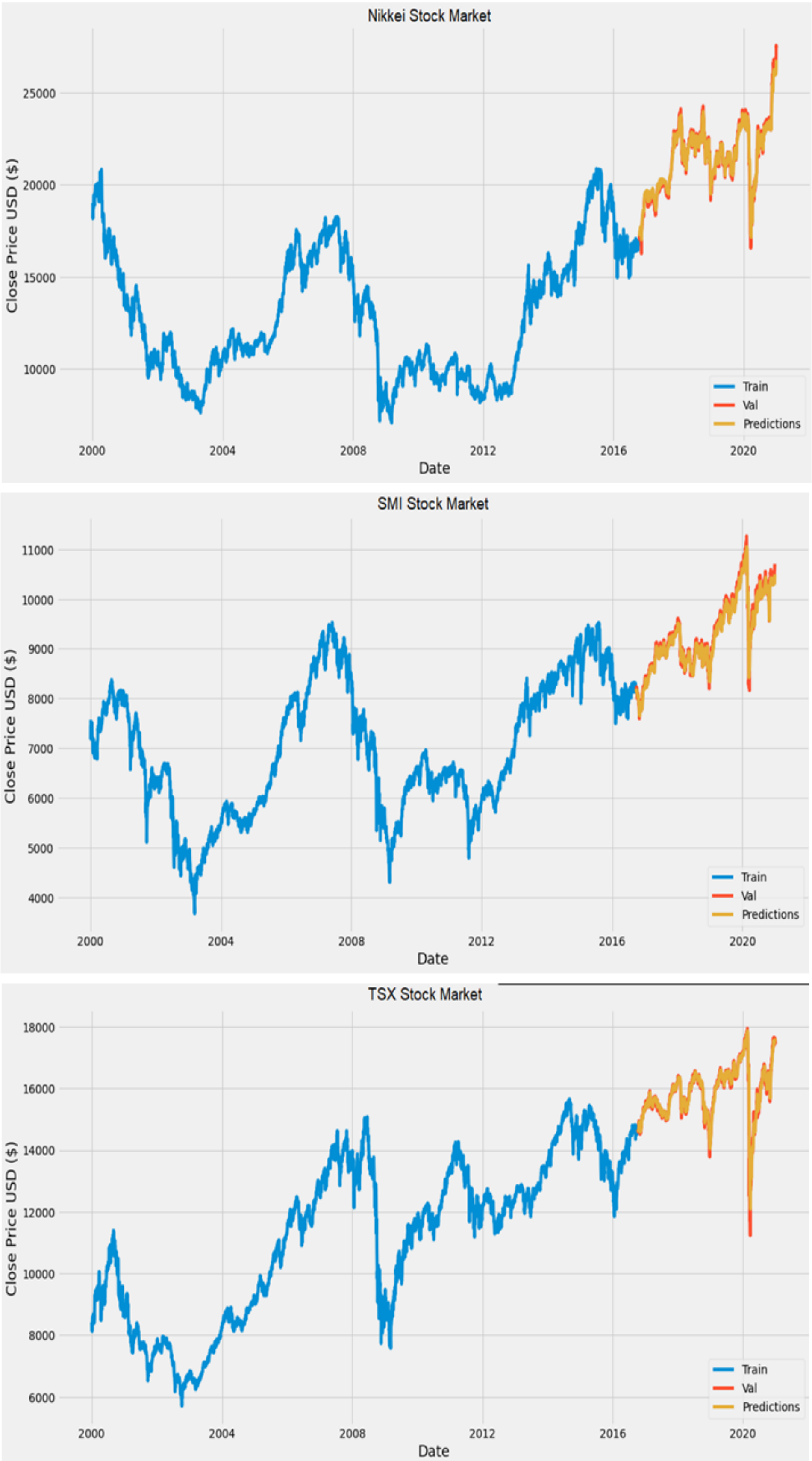


FIGURE 4.9: Machine Learning Forecasting with Fear Sentiment (Developed Markets)

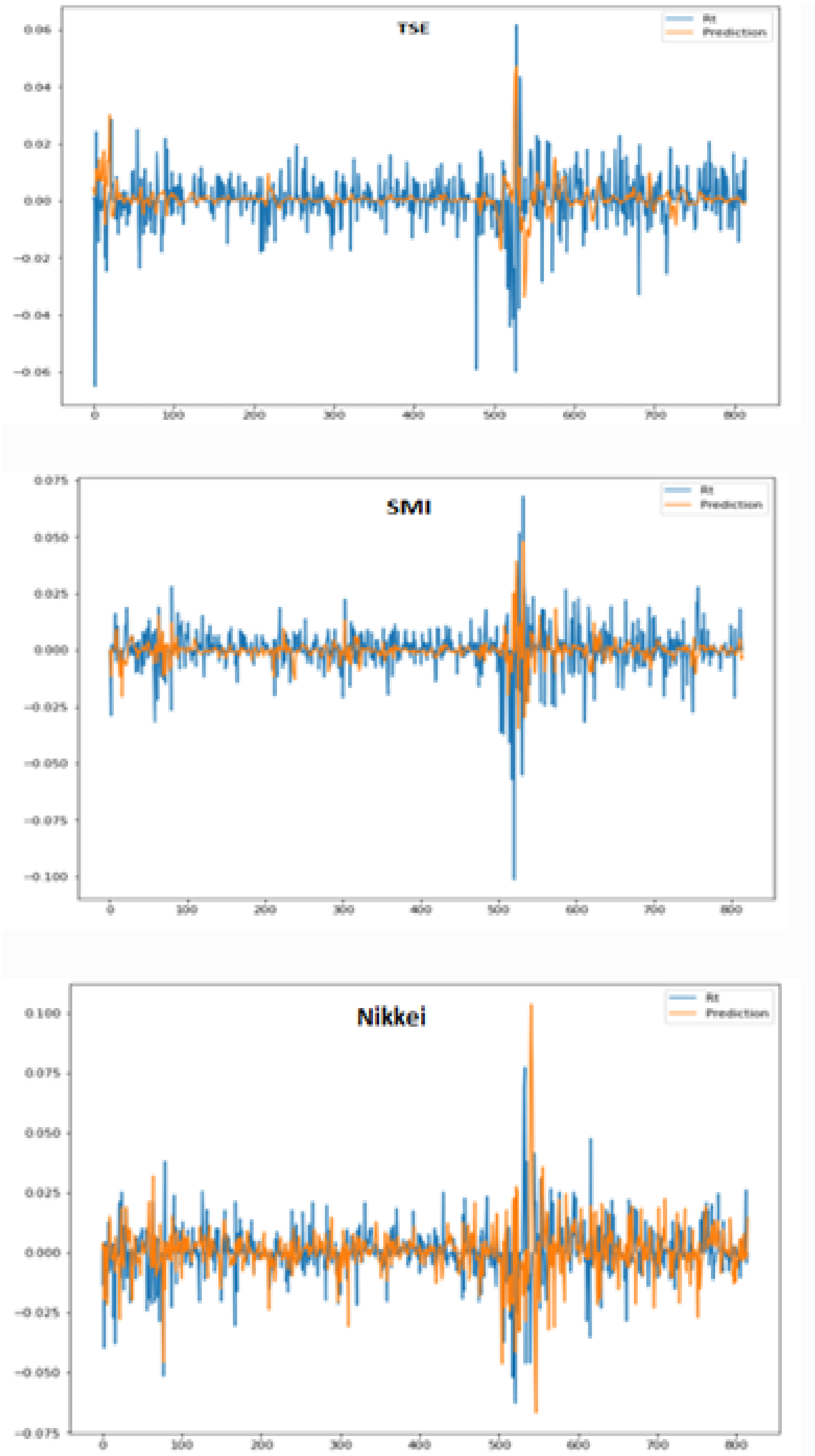


FIGURE 4.10: Machine Learning Forecasting with Fear Sentiment (Developed Markets)

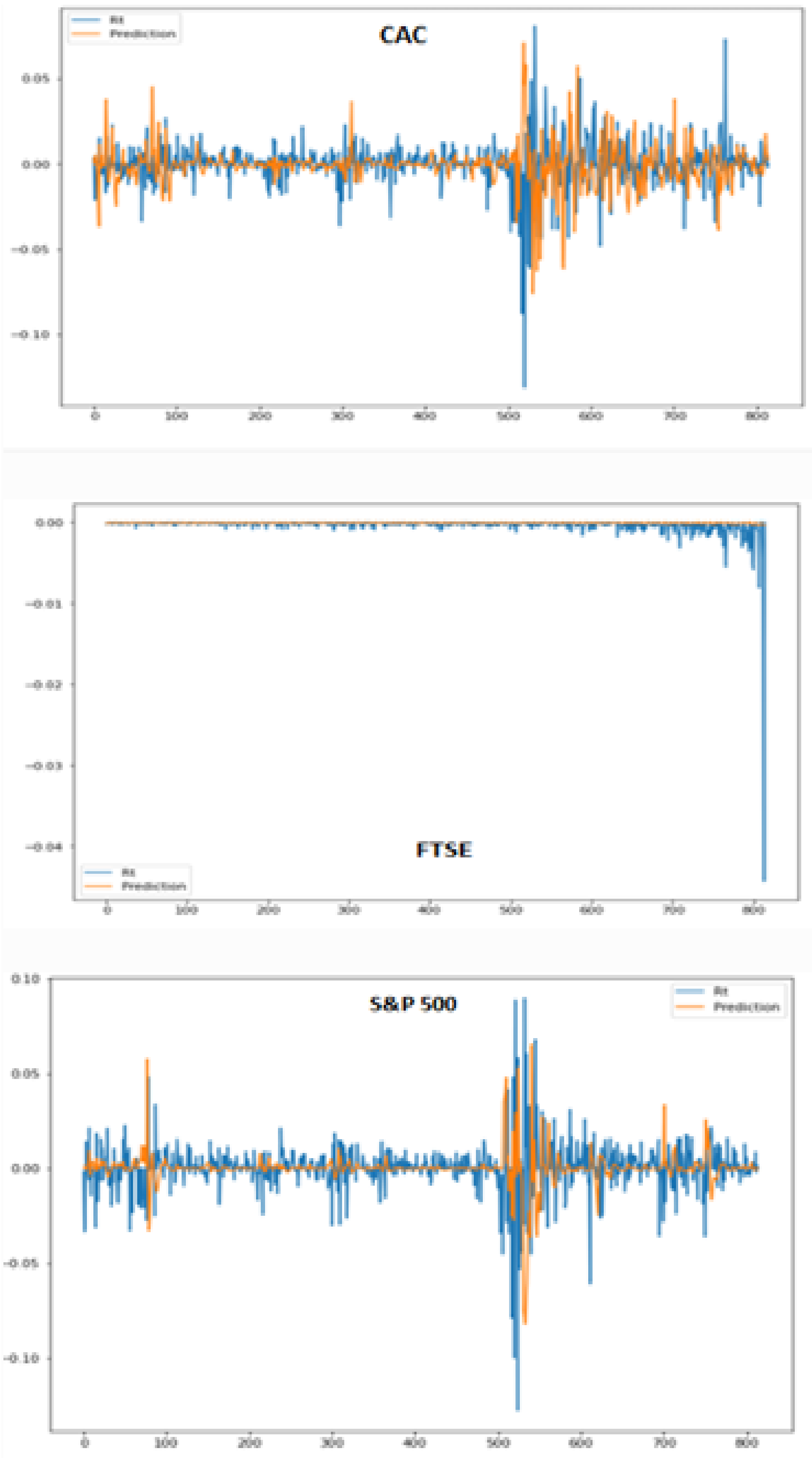


FIGURE 4.11: GARCH Model Forecasting(Emerging Markets)

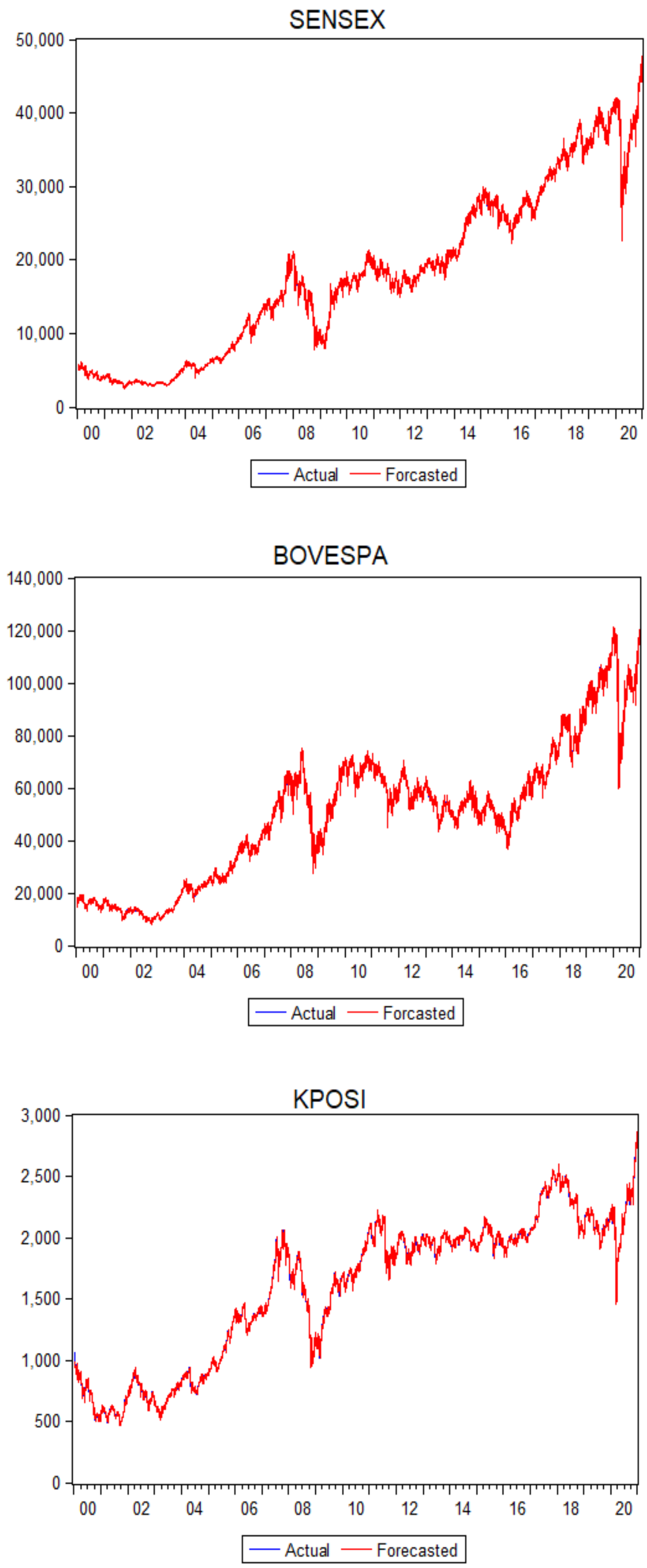


FIGURE 4.12: GARCH Model Forecasting(Emerging Markets)

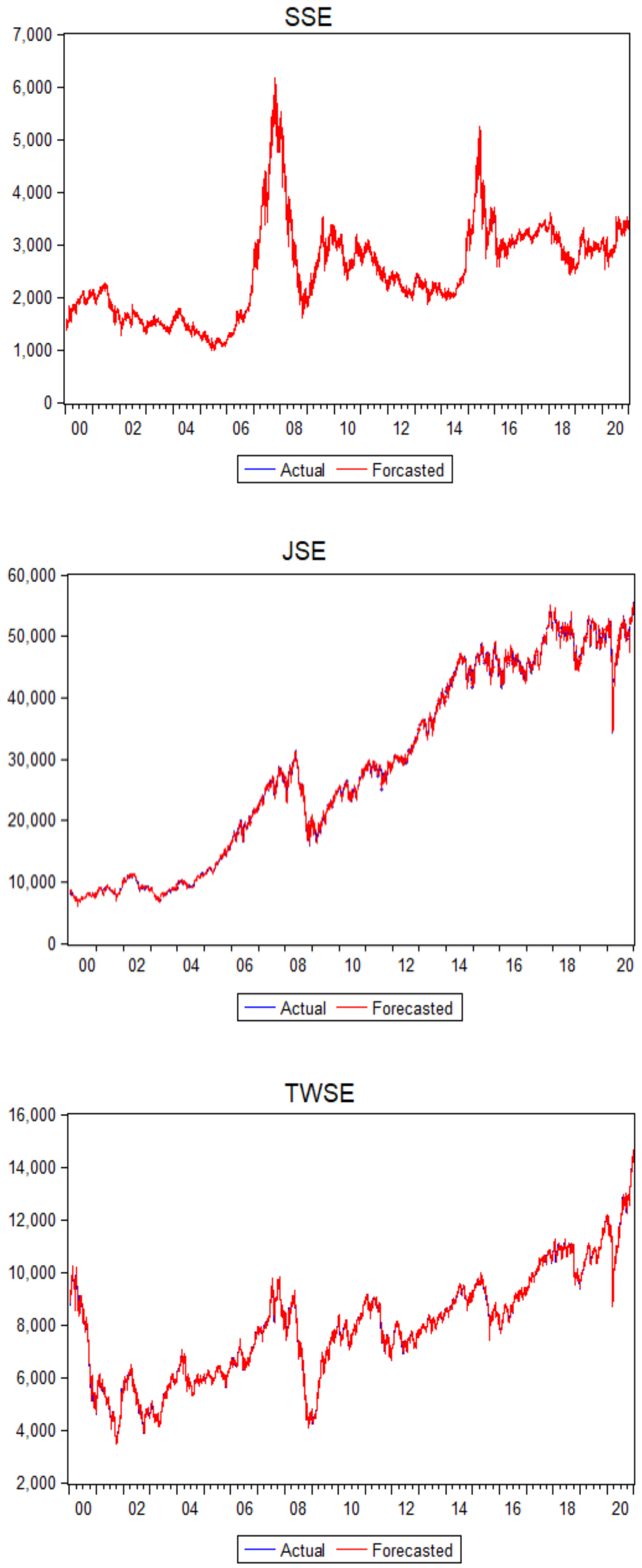


FIGURE 4.13: Machine Learning Forecasting(Emerging Markets)

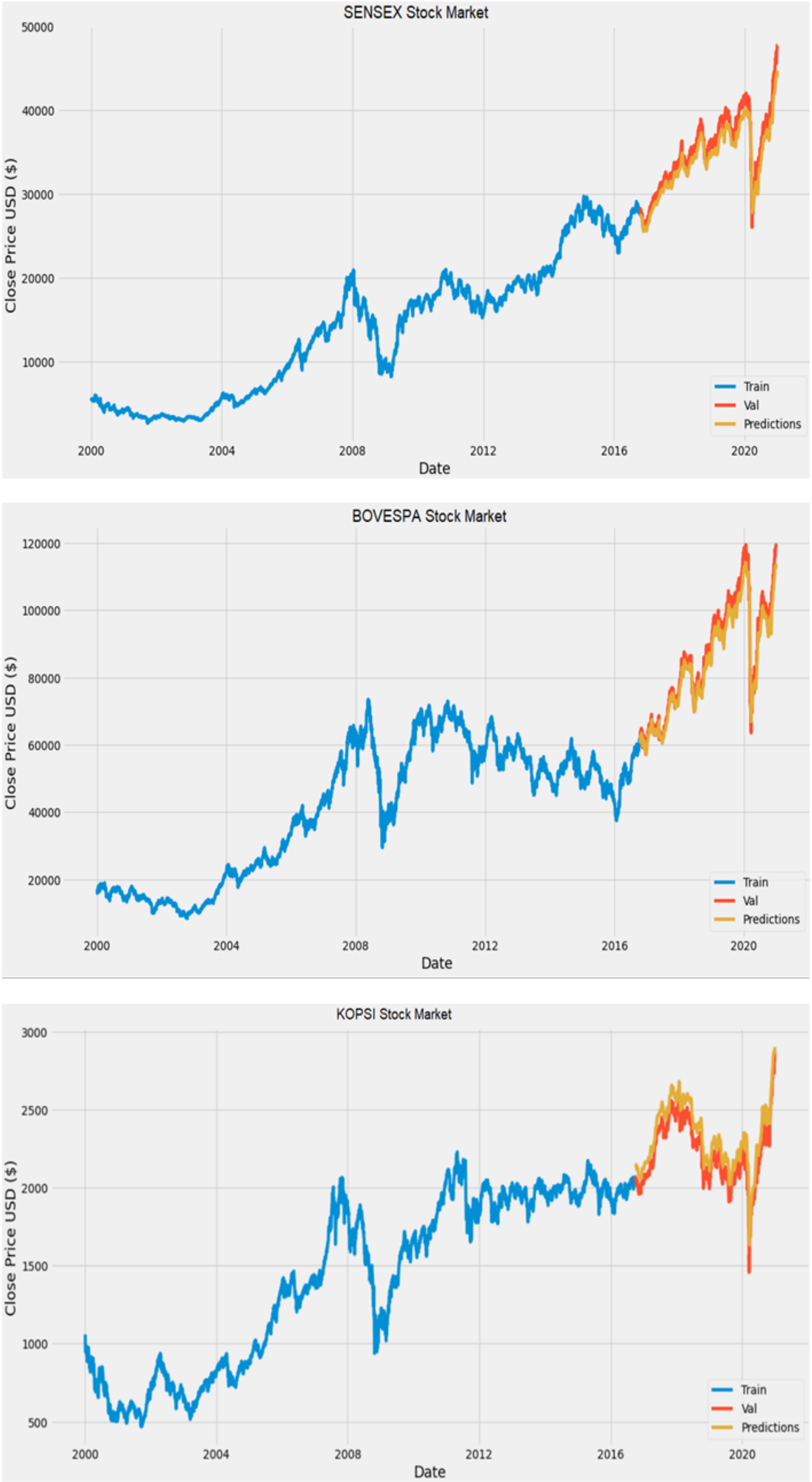


Figure 4.14: Machine Learning Forecasting(Emerging Markets)

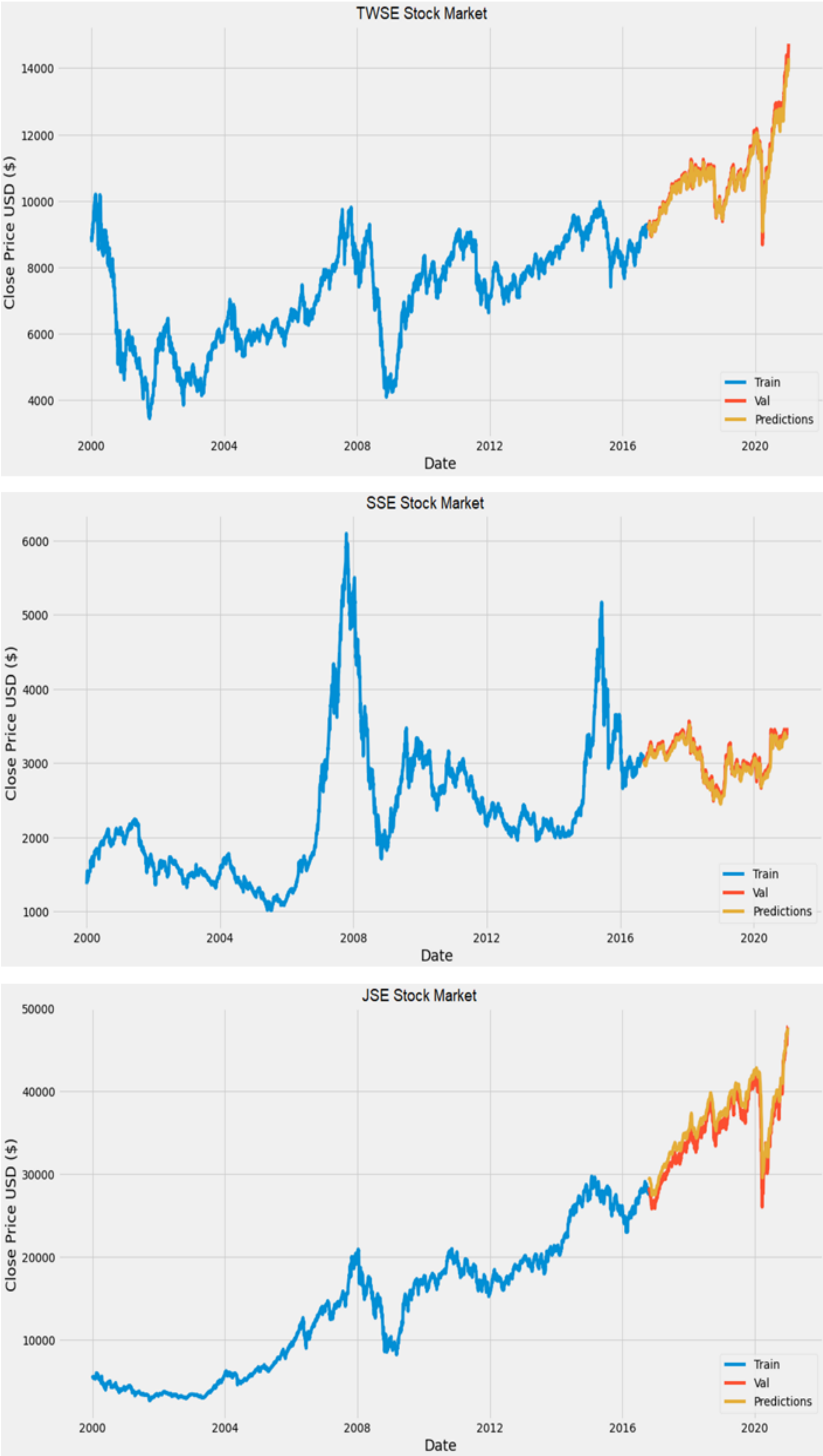


FIGURE 4.15: Machine Learning Forecasting with Fear Sentiment (Emerging Markets)

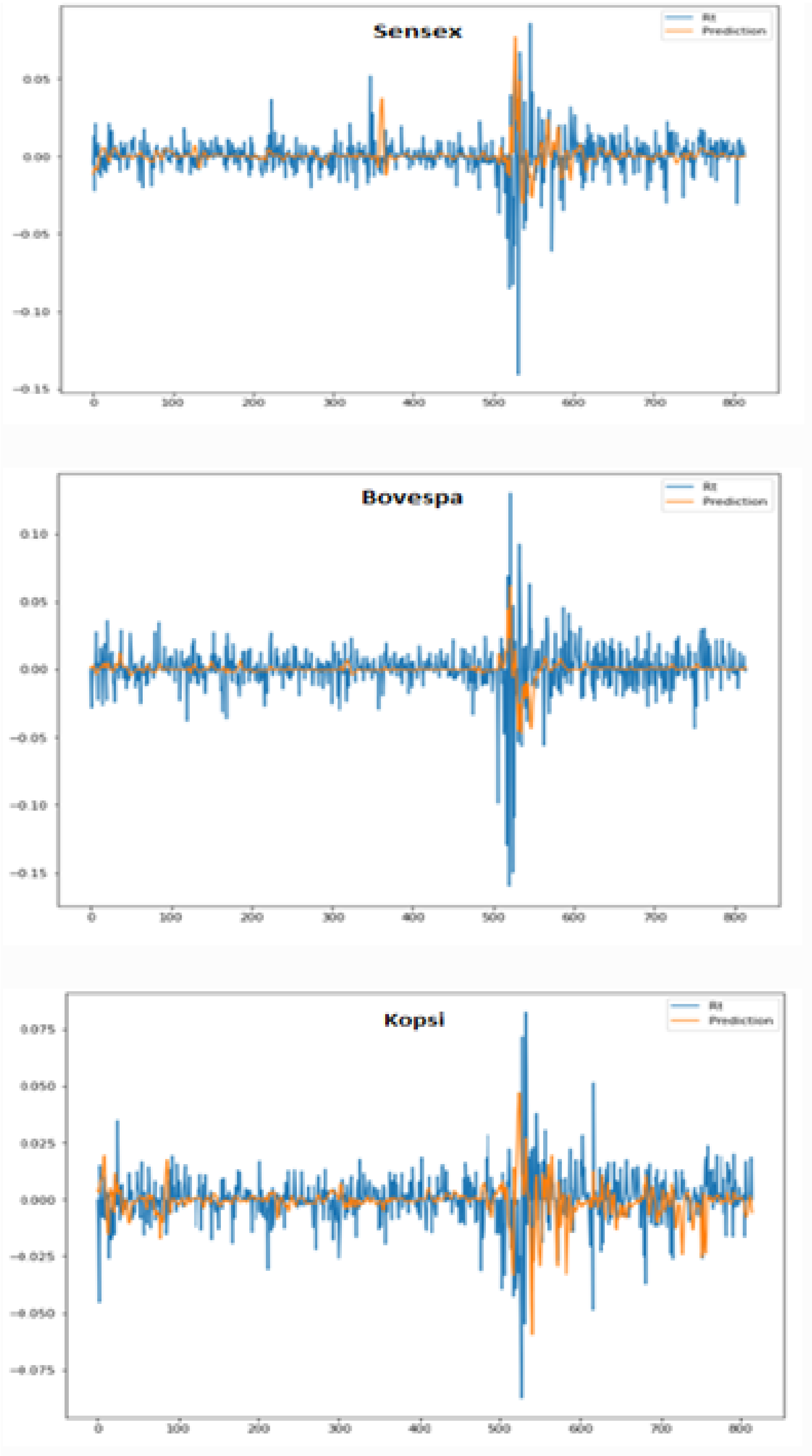


FIGURE 4.16: Machine Learning Forecasting with Fear Sentiment (Emerging Markets)

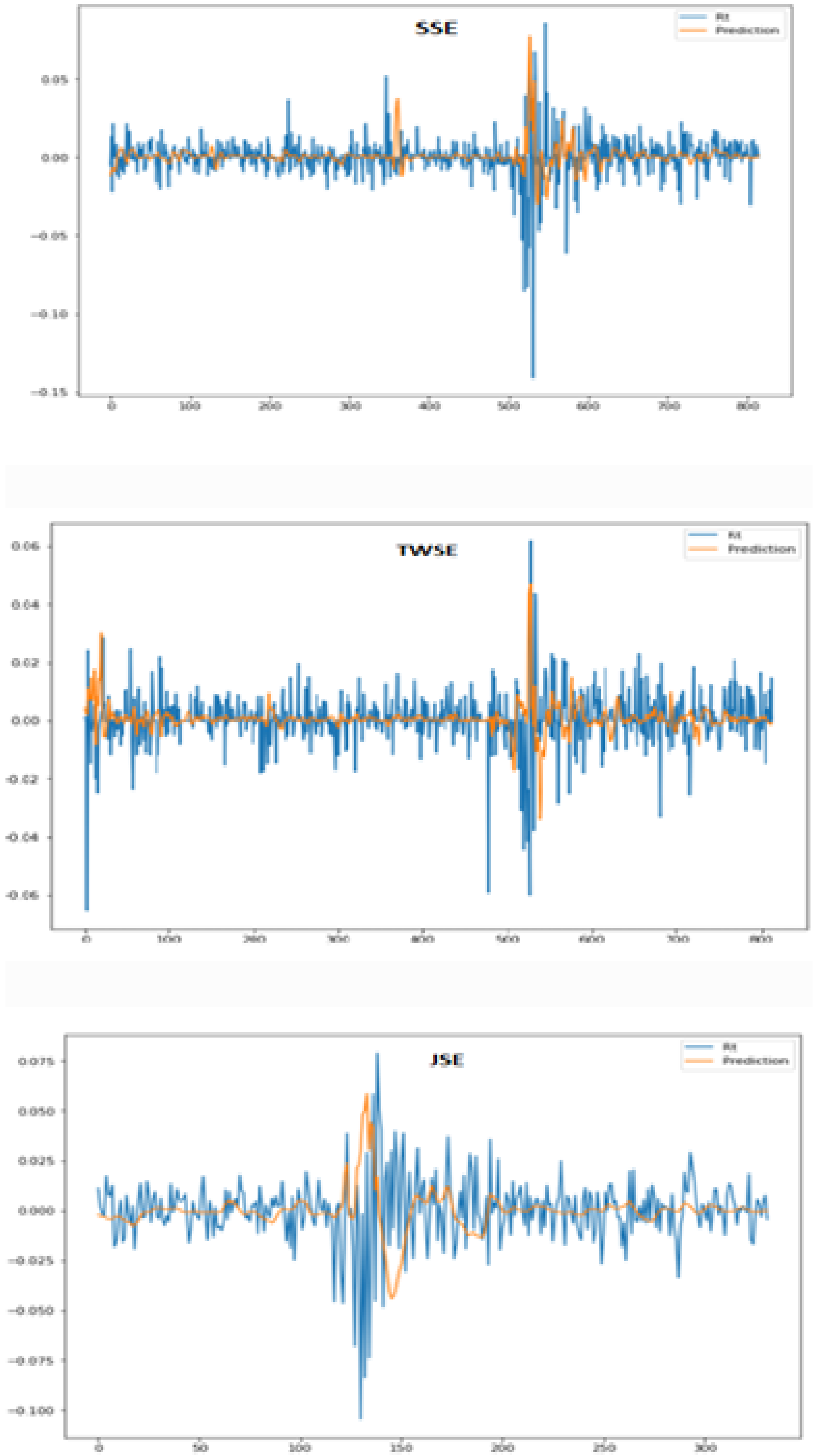


Table 4.4: Models Comparison

Markets		RMSE GARCH	RMSE ML
Developed	**S&P**	700.1825	57.8967
	FTSE	1360.1700	3.7709
	CAC	1494.1590	1.2000
	Nikkei	5499.7840	9.8250
	SMI	1820.4220	56.7880
	TSX	6111.6050	1.9425
Emerging	**SENSEX**	17133.2800	515.9050
	BOVESPA	17822.9300	2704.6588
	KOPSI	666.7009	96.1898
	JSE	22414.0700	690.1145
	SSE	1048.0910	36.0843
	TWSE	5325.2080	75.2248

results have proved that machine learning can forecast better than the other one. The finding of the study shows similar results as proposed by Hsu, Lessmann, Sung, Ma, and Johnson (2016); Kewat et al. (2017); Kulshreshtha (2020) historically. The use of Machine Learning increases the accuracy of future forecasting. The selection of a misfit forecasting tool will lead to the selection of some wrong investment and end up losing the returns so this decision is important and should be taken with care.

4.4 Return Optimization

This Study used machine learning algorithms to increase the expected returns of markets. In developed markets expected returns of the markets are increased by the use of ML. NIKKEI market expected return is an increase from 13% to 28.60% and the volatility is also increased from 21% to 28.50%. The expected returns of FTSE are increased by using ML Algorithms from 4% to 9.10% and the volatility is also increased from 19% to 20.40%. Although volatility is increased but our objective is to optimize the returns which is being fulfilled with the use of ML. CAC expected returns are increased from 14% to 17.5%, whereas volatility of the respective market jumped from 17% to 21.5%. TSX market index increased from 12% to 15.30%. And the volatility of the same markets is showed as 17.9%

Table 4.5: Return Optimization

		Actual Portfolio		Optimized portfolio	
		Exp. Return	**Volatility**	**Exp. Return**	**Volatility**
Developed	**Nikkei**	13.00%	21.00%	28.60%	28.50%
	FTSE	4.00%	19.00%	9.10%	20.40%
	CAC	14.00%	17.00%	17.50%	21.50%
	TSX	12.00%	17.00%	15.30%	17.90%
	SMI	12.00%	18.00%	11.20%	16.20%
	S&P	43.00%	28.00%	41.00%	25.00%
Emerging	**SENSEX**	26.00%	20.00%	39.90%	24.20%
	BOVESPA	6.00%	39.00%	7.00%	40.00%
	KOPSI	26.00%	25.00%	21.00%	24.10%
	JSE	15.00%	3.00%	19.00%	30.00%
	SSE	16.00%	18.00%	36.00%	28.90%
	TWSE	17.00%	18.00%	29.10%	25.40%

previously it was 17% before the use of ML algorithm. With these parameters, the study found less returns of S&P, SMI, and KOPSI with ML.Through this we are maximizing the returns and the selection of the portfolio is made on the basis of the value of Sharpe ratio. By changing the parameters results may improve.

In emerging markets, the expected returns of SENSEX are increased from 26% to 39.9% and the volatility jumps from 20% to 24.2%. In BOVESPA the expected return is also increased from 6% to 7%. The volatility of that specific market is changed from 39% to 40% after using ML. Machine learning algorithms of JSE increase the expected returns from 15% to 19% with a sharpe increase of 3% to 30% in volatility. Machine learning processed the SSE returns and increases the expected returns from 16% to 36%. An instant profit of 20% on the investment. The volatility increase for this specific market is 10% from 18% to 28%. The ML algorithms increased the expected returns of TWSE from 17% to 29% with a change in the volatility from 18% to 25.4%.

4.4.1 Scenario Analysis

The study created a simulated portfolio of given stock market indices by using machine learning and Sharpe ratio has been calculated to evaluate the performance

Table 4.6: Portfolio Reconstruction (Developed Markets)

Markets	**Actual**	**ML**	**Shares**	**S.R***	**Markets**	**Actual**	**ML**	**S.R***	**Shares**
				Developed					
	FB	FB	2			Sony			
	Amazon	Amazon	1			Mitsu	Sony		125
SNP	Apple	Apple	1	1.53	**Nikkei**	Toyota		0.93	
	Tesla	Tesla	7			Honda	Softbank		71
	MS	MS	19			Softbank			
	RDS					Nestle	Nestle		92
	Unilever					Swiss			
FTSE	BP Oil	Unilever	246	0.35	**SMI**	Zurich	Zurich	0.57	11
	B.A.T					Rick	Alcon		20
	HSBC					Alcon			
	L'Oréal					R. Bank	R. Bank		38
	Air Bus	L'Oréal	24			T. Bank			
CAC	Vinci			0.72	**TSX**	Enbridge		0.74	
	AXA					Nova	Railway		80
	LVMH	LVMH	15			Railway			

**FB = FaceBook *MS = Microsoft, *Mitsu = mitsubishi*

of the portfolio. To make it comparable study also computed an actual portfolio of stocks.

Table 4.6 and 4.7 represent the simulated portfolios results for developed and emerging markets respectively. The portfolio is constructed with five companies based on market capitalization in respective markets. Machine Learning Algorithms help to calculate the optimized portfolio.The selection of prior portfolio is done by the investor according to the module of Black Litterman Model. The combination of stocks is according to the views of the investor. This is another advantage of this module as it allows the investor to priorly select it. The ML not only helps in optimization but also tells, which stocks should need to drop and which stocks should need to add to increase the returns and ultimately the Sharpe ratio. All of these are based on the self assumption of the study. Another assumption that the study used is budget, which is set to be $15,000 for each market. This selection of this customized budget is done because of Black Litterman Model because it allows investors to determine it by them.

Sharpe ratio measures the performance of the investment. The above are the

Sharpe ratios of the optimized portfolio of the selected markets both in developed and emerging stock markets. In developed markets, Sharpe ratio of optimized S&P portfolio through machine learning algorithms is 1.52. The actual portfolio that we have designed in stimulations contains stocks of Facebook, Amazon, Apple, Tesla, and Microsoft. The ML optimized portfolio contained 2 shares of Facebook,1 share of Amazon, 19 shares of Apple, 7 shares of Netflix and 19 shares of Microsoft.

The Sharpe ratio of ML optimized portfolio of FTSE is 0.35. The actual portfolio contains the stocks of Royal Dutch Shell, Unilever, BP Oil Industry, British American Tobacco, and HSBC Bank. Machine learning algorithms tell us to buy 246 shares of Unilever and should sell out all the other shares. The Sharpe ratio of the optimized portfolio using ML is 0.72 for CAC. The actual portfolio contains the stocks of L'Oréal, Air Bus, Vinci, AXA. SA-Insurance and LVMH-luxury goods. Machine learning algorithms tell us to buy 24 shares of L'Oréal and 15 shares of LVMH. should sell out all the other shares.

The Sharpe ratio of ML optimized portfolio of NIKKEI is 0.93. The actual portfolio contains the stocks of Toyota Motor Corporation, Mitsubishi UFJ Financial Group, Softbank, Honda Motor, and Sony. Machine learning algorithms tell us to buy 125 shares of Sony and 71 Softbank and should sell out all the other shares. The Sharpe ratio of ML optimized portfolio of SMI is 0.57. The actual portfolio contains the stocks of Nestle N, Zurich Insurance, Swiss com, Rick Mont N, and Alcon. Machine learning algorithms tell us to buy 92 shares of Nestle N and 11 shares of Zurich Insurance and 20 shares of Alcon. All the other stocks are suggested to be sold for optimization. The Sharpe ratio of ML optimized portfolio of TSX is 0.74. The actual portfolio contains the stocks of Royal Bank Canada, Toronto Dominion Bank, Enbridge, Bank Of Nova Scotia, and Canadian National Railway. Machine learning algorithms tell us to buy 38 shares of Royal Bank Canada and 80 shares of Canadian National Railway. And sell out all the other shares. In emerging markets, the Sharpe ratio of optimized portfolio using ML is 0.72 for SENSEX. The actual portfolio contains the stocks of Nestle Indian Limited, Hindustan Unilever Limited, State Bank of India, Bajaj Finance Limited and

Table 4.7: Portfolio Reconstruction (Emerging Markets)

Markets	**Actual**	**ML**	**Shares**	***S.R**	**Markets**	**Actual**	**ML**	**S.R**	**Shares**
				Emerging Markets					
	Nestle					Moutai	Moutai		6
	Unil	Unil	2			ICBC			
SEN	S.Bank			1.57	**SSE**	Ping An		0.56	
	Bajaj	Bajaj	2			BOC	Ping An		60
	Reliance					CMB			
	Petro					Naspers			
	Unibanco					Anglo	Nasp		235
BOV	Bradesco	Bras	2054	0.1	**JSE**	R.mond		1.2	
	Ambev					BHP	BHP		67
	Brasilia					S.Bank			
	Samsung	-	-			NAN	MedTek		2
	Hynix	-	-			MedTek			
KOP	LG	-	-	0.82	**TWSE**	Forms		1.07	
	Naver	-	-			Hon Hai	TSM		105
	Hyundai	-	-			TSM			

Where as S.R = Sharpe Ratio, SEN = SENSEX, BOV = BOVESPA, KOP = KOPSI, UNIL = Unilever, Bras = Brasilia, Nesp = Naspers, MedTek = Media Tek

Reliance Industries Limited. The ML optimized portfolio only contains 2 shares of Hindustan Unilever Limited and 2 shares of Bajaj Finance Limited.

The Sharpe ratio of the optimized portfolio using ML is 0.1 for BOVESPA. The actual portfolio contains the stocks of Petro bras, It au Uni banco, Banco Bradesco, Ambev, and Santander Brasilia. Machine learning algorithms tell us to buy 2050 shares of Santander Brasilia. The Sharpe ratio of ML optimized portfolio of KOPSI is 0.82. The actual portfolio contains the stocks of Samsung Electronics, Hynix, LG Chem, Naver, and Hyundai Motor. Machine learning algorithms tell us not to buy any of stock from respective market. Similarly, the Sharpe ratio of the optimized portfolio using ML is 1.20 for SSE. The actual portfolio contains the stocks of Kweichow Moutai, ICBC, Ping An, Agricultural Bank of China, and China Merchants Bank. Machine learning algorithms tell us to buy 60 shares of Ping An Insurance and 6 shares of Kweichow Moutai within the specified budget of $15000. Others stocks should be sold for optimization purpose.

The Sharpe ratio of the optimized portfolio using ML is 0.56 for JSE. The actual portfolio contains the stocks of Naspers, BHP Group, Richmond N, Anglo American Plc and Standard Bank. Machine learning algorithms tell us to buy 235 shares of Naspers and 67 shares of BHP Group. Other stock should be sold. The Sharpe ratio of ML optimized portfolio of TWSE is 0.57. Machine learning algorithms tell us to buy 105 shares of Taiwan Semiconductor Manufacture (TSM) and 2 shares of TN, all other stocks were sold by the machine during the optimization.

Table 4.6 represents the portfolio optimization. It shows that investor should invest in S&P500 portfolio from the developed markets because its Sharpe ratio is higher than rest of the portfolio of the respective markets. Emerging markets results depicts that investor should restructure all the assets in SENSEX, SSE and TWSE as their Sharpe ratio is higher than other portfolios from the respective markets. Overall emerging markets outperform the developed markets according to the results of the study.

4.5 Discussion

This section discuss results of above-estimated models to answers the study questions. To meet the first two objectives of study results shows that economic policy uncertainty has contributed to the fear sentiments of investors. Because of global diversification an uncertain environment has been developed over the time period for the investors. When an investor enters in global diversification economic polices plays a crucial rule for the investment decision. Economic polices predict the future governments actions that how they will devise their economic policy, whether they want to promote their businesses by decreasing the interest rates or they want to promote consumer for saving by increasing the interest rate. However, when governments promote business it will creates a profitable opportunity for investors. So far this reason, economic policies uncertainty shows a significant positive relationship with investors fear sentiments.

Based on the results of the system of equation fear sentiments shows a significant positive impact on volatility which means investor fear sentiments increase the market volatility (Balcilar et al., 2017). Because in the uncertain market situation investor gets panic and makes an irrational decision regarding the buying and selling of financial assets of their portfolio which makes the financial assets mispriced which will negatively affect their portfolio returns. So, investors fear sentiments will contribute towards the increment in the market volatility (Smales, 2017). According to the results of the study market volatility shows a significant inverse relation with the returns which shows that an increase in volatility will lead to a decrease in returns. Because market volatility is comprised of multiple components such as interest rate risk, commodity markets, and equity risk, etc. The interactive regression model also concludes that bad news shows an insignificant impact on returns but at the same time it moderates the relationship of volatility and returns. Because bad news increases the fear sentiment and this fear sentiment have a significant impact on market volatility in long run (Aggarwal, 2017; Griffith et al., 2020).

To answer the question on returns optimization under the presence of investors fear sentiments forecasting of stock markets help investors get confidence and reduce the fear with the knowledge of future. In this study, the GARCH model is employed for forecasting the volatility of emerging and developed markets. Results of GARCH show that lagged returns of all indices of emerging and developed markets are descriptively insignificant. While Arch is statistically significant which concludes that past price volatility will impact the current price volatility. Based on results GARCH term shows is statistically significant and confirms the persistence of volatility in the long run for both emerging and developed markets. Results presented in table 4.4 shows the comparative results of forecasting performance of GARCH model and Machine learning model. Results shows that machine learning algorithms prediction of returns outperform the GARCH model return prediction in both developed and emerging markets. Results of this study are also align with the results of Hsu et al. (2016); Kewat et al. (2017); Papacharalampous et al. (2019) studies on comparison of machine learning and classical econometric models. This validates the argument of machine learning methods have better predicting ability then the econometric models. Because of the ability to deal with the non-linearity and complex environment. This answer our questions "Does machine learning model outperform the econometric model in forecasting in developed markets? Does machine learning model outperform the econometric model in forecasting in emerging markets?".

Another objective of the study is to use of these machine learning algorithms to optimize the returns of investors. To achieve this objective table 4.5 shows a comparative return maximization of classical mean variance optimization method and advance machine learning method. Results shows machine learning based methods outperform the conventional method of returns optimization. As returns increased the risk which is associated with the respective market also increases. To validate this study built a scenario by taking five stocks from each market and built a portfolio. Results also validate the performance of machine learning optimization. Sharpe ratio show the significant improvement and machine restructure the portfolio as to maximize the overall returns. These machine leaning methods perform equally in developed and developing markets which also full fill the objective of this study. For the return optimization through portfolio structuring, results presented in table 4.6 and 4.7 shows the portfolio optimization along with investment reallocation. The use of Machine learning quantitative trading approach eliminates the irrational impact of fear sentiment by reconstructing of investment. This approach illustrates to reallocate during the financial distress. We used Mean-Variance Model for return and portfolio optimization. Just like Mean-Variance theory ,this model also help us in portfolio optimization and resource allocation. As we can see that there is high sentiment in all markets from the figure 4.4 and 4.6 during 2008 financial crisis and during 2019 covid distress. This shows that market consider and react instantly according to each and every piece of information in market. The presence of high sentiment instantly after the crisis show the behavior of market after any news or the crisis as discribed in behavioral finance.

Chapter 5

Conclusion

This Study explore how the fear sentiments influence the stock market returns. If investor can control sentiments, then investor control the returns but sentiments are not in control by the human. But if returns can be forecasted and invest accordingly then investor will be in a position to optimize the returns which is the ultimate goal of the study. The understanding of relationship of fear sentiments of investor and the stock return is important to understand because fear sentiments pushes away the volatility equilibrium which impacts the stock return of a particular market. The reason of this study is to optimize the returns by forecasting with the help of machine learning.

In first stage of studies,study used the system of equations to evaluate the fear sentiments in market returns in developed markets. The results of this answer research question of this study that what is impact of fear sentiments on stock return in developed market. The study illustrates that fear sentiment has significant impact on the developed markets. Then GARCH is applied to check whether volatility is persistent in nature or not.It was shown that volatility of all the developed market is persistent in nature. In the second stage the study check the fear sentiment impact on the emerging markets so system of equations gives the answer to research question that what is the impact of fear sentiments on emerging markets return. Through system of equations, we evaluate impact of sentiments on returns of the market. It tells that there is a significant impact of fear sentiments on returns of respective markets, which is also the second objective of our study.

In the third stage the study compared two model's for forecasting of developed markets. It uses GARCH and ML algorithms for this. The result shows that ML algorithms have more forecasting power than the other model. This result also answers the question that which is better forecasting model for developed markets. This is how we identify the better forecasting model for prediction. In the fourth stage we have made a comparison of econometric model with the machine learning model to check better prediction model for the emerging markets. We have come to know that machine learning model do better job in the specific markets. We selected the model on the basis of value of accuracy estimator which in this case is RMSE.

In the fourth stage, machine learning algorithms are used for optimizing the returns of the developed markets. In this stage it was shown how the returns are optimized and how we can allocate the resources more effectively in order to achieve objective of returns optimization. Based on results of the return optimization suggested by the machine learning algorithms, the investor can structure it portfolio accordingly. In the fifth stage, study restructure of portfolio of the emerging markets in order to optimize the returns of the respective markets. The ML helps us to identify the better investment opportunities in the markets and how these opportunities will work in the favor of the investor. The customized budget for every investor can also be considered while structuring the portfolio. The study helps the investor to reduce the biasedness from decision making because machine itself is restructuring the portfolio. Machine take the decisions on the basis of optimized returns without considering the emotional quotient impact. Thus it takes the rational decision, even when sentiment in the market is high.

5.1 Recommendations

Based on this study following recommendation are made. All these recommendations are deals with multiple dimension such as the selection of the best forecasting model, return optimization, designing the policies by the fund manager, policy-maker and for the academic researcher. According to the results of the system of equations impact of fear sentiments on both developed and emerging market returns are evaluated and conclude that fear sentiments have a significant positive impact on market returns. So, it's recommended for an investor to make a rational investment decision and also consider the impact of fear sentiment on market returns while adding the securities in their portfolio. The investor should try to add all those market securities which are more efficient and less affected by the fear sentiments (Corredor et al., 2015).

Results of GARCH models show the persistence of volatility in emerging and developed markets and this persistence of volatility is the long run-in nature. It is recommended based on the results of GARCH that investors should be careful while constructing their portfolios because all these markets show the persistence of volatility which decreases the opportunity for portfolio optimization. Based on the comparison of GARCH and ML for the prediction of prices of emerging and developed stock markets results conclude that ML method outperforms the classical GARCH model. So, it is recommended that investors and academic researchers can use the results of the comparison and predict the future prices to create an optimal hedge in the future for profit maximization.

ML algorithms are used in this study for optimizing the returns of emerging and developed markets. Results of ML algorithms provide that how much we can allocate resources for return optimization. Markets will continue to translate the sentiments of investor in the returns. Machine Learning Algorithms are scientific methods having properties of rationality, unbiased and analytical with dispassionate and impassive decision-making abilities. Human decision making is the composite of irrational and biased behavior. So, it is recommended that investor should use the Quantitative trading and machine algorithms for rational decision making in the presence of sentiments.

Further it is recommended that investors can use the results while constructing their portfolios for efficient and effective resource allocation. In the last outcomes of this will be used by the policymakers while designing the strategies of risk management and macro stabilization. The fund manager can use the outcomes of this study while devising their investment strategies, to achieving the objective of profit maximization by efficiently and effectively allocating their resources. The academic researchers can work on the limitations of the study. The future study can incorporate other stock markets to increase sample size of study. In this study, 12 indices are considered. Future work may increase the sample of study. The study can be applied on the whole by considering all the listed stocks of that market as we have chosen top five with high market capitalization.

Bibliography

Aggarwal, D. (2017). Exploring relation between indian market sentiments and stock market returns. *Asian Journal of Empirical Research*, 7(7):147–159.

Aggarwal, D. (2019). Defining and measuring market sentiments: a review of the literature. *Qualitative Research in Financial Markets*.

Al-Thaqeb, S. A. and Algharabali, B. G. (2019). Economic policy uncertainty: A literature review. *The Journal of Economic Asymmetries*, 20:e00133.

Albulescu, C. T. (2021). Covid-19 and the united states financial markets? volatility. *Finance Research Letters*, 38:101699.

Arouri, M., Estay, C., Rault, C., and Roubaud, D. (2016). Economic policy uncertainty and stock markets: Long-run evidence from the us. *Finance Research Letters*, 18:136–141.

Athey, S. and Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11(1):685–725.

Baker, M. and Wurgler, J. J. J. o. e. p. (2007). Investor sentiment in the stock market. *Journal of economic perspectives*, 21(2):129–152.

Balcilar, M., Bonato, M., Demirer, R., and Gupta, R. (2017). The effect of investor sentiment on gold market return dynamics: Evidence from a nonparametric causality-in-quantiles approach. *Resources Policy*, 51:77–84.

Banerjee, P. S., Doran, J. S., and Peterson, D. R. (2007). Implied volatility and future portfolio returns. *Journal of Banking & Finance*, 31(10):3183–3199.

Bao, W., Yue, J., and Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PloS one*, 12(7):e0180944.

Baştanlar, Y. and Özuysal, M. (2014). *Introduction to Machine Learning*, pages 105–128. Humana Press, Totowa, NJ.

Becker, R., Clements, A. E., and White, S. I. (2007). Does implied volatility provide any information beyond that captured in model-based volatility forecasts? *Journal of Banking & Finance*, 31(8):2535–2549.

Beja, A. (1972). On systematic and unsystematic components of financial risk. *The Journal of Finance*, 27(1):37–45.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *ournal of econometrics*, 31(3):307–327.

Bouoiyour, J. and Selmi, R. (2016). Bitcoin: A beginning of a new phase. *Economics Bulletin*, 36(3):1430–1440.

Braun, P. A., Nelson, D. B., and Sunier, A. M. (1995). Good news, bad news, volatility, and betas. *The Journal of Finance*, 50(5):1575–1603.

Canbaş, S. and Kandır, S. Y. (2009). Investor sentiment and stock returns: Evidence from turkey. *Emerging Markets Finance and Trade*, 45(4):36–52.

Chakraborty, M. and Subramaniam, S. (2020). Asymmetric relationship of investor sentiment with stock return and volatility: evidence from india. *Review of Behavioral Finance*, 12(4):435–454.

Chen, C., Liu, L., and Zhao, N. (2020a). Fear sentiment, uncertainty, and bitcoin price dynamics: The case of covid-19. *Emerging Markets Finance and Trade*, 56(10):2298–2309.

Chen, C., Liu, L., and Zhao, N. (2020b). Fear sentiment, uncertainty, and bitcoin price dynamics: The case of covid-19. *Emerging Markets Finance and Trade*, 56(10):2298–2309.

Corredor, P., Ferrer, E., and Santamaria, R. (2015). The impact of investor sentiment on stock returns in emerging markets: The case of central european markets. *Eastern European Economics*, 53(4):328–355.

Da, Z., Engelberg, J., and Gao, P. (2015). The sum of all fears investor sentiment and asset prices. *The Review of Financial Studies*, 28(1):1–32.

Dash, S. and Moran, M. T. (2005). Vix as a companion for hedge fund portfolios. *The Journal of Alternative Investments*, 8(3):75–80.

De Prado, M. L. (2016). Building diversified portfolios that outperform out of sample. *The Journal of Portfolio Management*, 42(4):59–69.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). How inefficient are simple asset allocation strategies. *Review of Financial Studies*, 22(5):1915–1953.

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

Dritsaki, C. (2018). The performance of hybrid arima-garch modeling and forecasting oil price.

Du, B., Zhu, H., and Zhao, J. (2016). Optimal execution in high-frequency trading with bayesian learning. *Physica A: Statistical Mechanics and its Applications*, 461:767–777.

Dyhrberg, A. H. (2016). Bitcoin, gold and the dollar?a garch volatility analysis. *Finance Research Letters*, 16:85–92.

Elmachtoub, A. N. and Grigas, P. (2020). Smart "predict, then optimize".

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

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

Fisher, K. L. and Statman, M. (2000). Investor sentiment and stock returns. *Financial Analysts Journal*, 56(2):16–23.

Fleming, J., Ostdiek, B., and Whaley, R. E. (1995). Predicting stock market volatility: A new measure. *Journal of Futures Markets*, 15(3):265–302.

Ftiti, Z. and Hadhri, S. (2019). Can economic policy uncertainty, oil prices, and investor sentiment predict islamic stock returns? a multi-scale perspective. *Pacific-Basin Finance Journal*, 53:40–55.

Galbraith, J. K. and Sorel, G. (1977). *The age of uncertainty*. Boston: Houghton Mifflin.

Goetzmann, W. N., Kim, D., Kumar, A., and Wang, Q. (2015). Weather-induced mood, institutional investors, and stock returns. *The Review of Financial Studies*, 28(1):73–111.

Griffith, J., Najand, M., and Shen, J. (2020). Emotions in the stock market. *Journal of Behavioral Finance*, 21(1):42–56.

Guo, K., Sun, Y., and Qian, X. (2017). Can investor sentiment be used to predict the stock price? dynamic analysis based on china stock market. *Physica A: Statistical Mechanics and its Applications*, 469:390–396.

Gyamerah, S. A. (2019). Modelling the volatility of bitcoin returns using garch models. *Quantitative Finance and Economics*, 3(4):739–753.

Habibah, U., Rajput, S., and Sadhwani, R. (2017). Stock market return predictability: Google pessimistic sentiments versus fear gauge. *Cogent Economics & Finance*, 5(1):1390897.

Hajiali, M. (2020). Big data and sentiment analysis: A comprehensive and systematic literature review. *Concurrency and Computation: Practice and Experience*, 32(14):e5671.

Haritha, P. and Rishad, A. (2020). An empirical examination of investor sentiment and stock market volatility: evidence from india. *Financial Innovation*, 6(1):1–15.

He, G., Zhu, S., and Gu, H. (2020). The nonlinear relationship between investor sentiment, stock return, and volatility. *Discrete Dynamics in Nature and Society*, 2020.

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.

Ji, Q., Li, J., and Sun, X. (2019). Measuring the interdependence between investor sentiment and crude oil returns: New evidence from the cftc's disaggregated reports. *Finance Research Letters*, 30:420–425.

Joyo, A. S. and Lefen, L. (2019). Stock market integration of pakistan with its trading partners: A multivariate dcc-garch model approach. *Sustainability*, 11(2):303.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Katsiampa, P. (2017). Volatility estimation for bitcoin: A comparison of garch models. *Economics Letters*, 158:3–6.

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.

Krollner, B., Vanstone, B. J., and Finnie, G. R. (2010). Financial time series forecasting with machine learning techniques: a survey. In *ESANN*.

Lee, W. Y., Jiang, C. X., and Indro, D. (2002). Stock market volatility, excess returns, and the role of investor sentiment. *Journal of banking & Finance*, 26(12):2277–2299.

Li, X., Wu, P., and Wang, W. (2020a). Incorporating stock prices and news sentiments for stock market prediction: A case of hong kong. *Information Processing & Management*, 57(5):102212.

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

Liang, Z., Chen, H., Zhu, J., Jiang, K., and Li, Y. (2018). Adversarial deep reinforcement learning in portfolio management. *arXiv preprint arXiv:1808.09940*.

Liu, L. and Zhang, T. (2015). Economic policy uncertainty and stock market volatility. *Finance Research Letters*, 15:99–105.

Liu, Y.-H., Dai, S.-R., Chang, F.-M., Lin, Y.-B., and Lee, N. R. (2020). Does the investor sentiment affect the stock returns in taiwan's stock market under different market states? *Journal of Applied Finance and Banking*, 10(5):41–59.

Loewenstein, G. (2000). Emotions in economic theory and economic behavior. *American economic review*, 90(2):426–432.

Maghyereh, A., Awartani, B., and Abdoh, H. (2020). The effects of investor emotions sentiments on crude oil returns: A time and frequency dynamics analysis. *International Economics*, 162:110–124.

Markowitz, H. (1952). Portfolio selection*. *The Journal of Finance*, 7(1):77–91.

Markowitz, H. M. (1999). The early history of portfolio theory: 1600?1960. *Financial analysts journal*, 55(4):5–16.

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

Menezes, R., Oliveira, l., and Portela, S. (2019). Investigating detrended fluctuation analysis with structural breaks. *Physica A: Statistical Mechanics and its Applications*, 518:331–342.

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.

Ricciardi, V. and Simon, H. K. (2000). What is behavioral finance? *Business, Education & Technology Journal*, 2(2):1–9.

Salisu, A. A., Akanni, L., and Raheem, I. (2020). The covid-19 global fear index and the predictability of commodity price returns. *Journal of Behavioral and Experimental Finance*, 27:100383.

Schmeling, M. (2009). Investor sentiment and stock returns: Some international evidence. *Journal of Empirical Finance*, 16(3):394–408.

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

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, S., Jiang, H., and Zhang, T. (2012). Stock market forecasting using machine learning algorithms. *Department of Electrical Engineering, Stanford University, Stanford, CA*, pages 1–5.

Siami-Namini, S., Tavakoli, N., and Namin, A. S. (2018). A comparison of arima and lstm in forecasting time series. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 1394–1401. IEEE.

Sim, H. S., Kim, H. I., and Ahn, J. J. (2019). Is deep learning for image recognition applicable to stock market prediction? *Complexity*, 2019:10.

Skiadopoulos, G. (2004). The greek implied volatility index: construction and properties. *Applied Financial Economics*, 14(16):1187–1196.

Smales, L. A. (2017). The importance of fear: investor sentiment and stock market returns. *Applied Economics*, 49(34):3395–3421.

Snow, D. (2020). Machine learning in asset management?part 1: Portfolio construction?trading strategies. *The Journal of Financial Data Science*, 2(1):10–23.

Song, H., Peng, D., and Huang, X. (2020). Incorporating research reports and market sentiment for stock excess return prediction: A case of mainland china. *Scientific Programming*, 2020.

Suleman, M. T. (2012). Stock market reaction to good and bad political news. *Asian Journal of Finance & Accounting*, 4(1):299–312.

Ta, V.-D., Liu, C.-M., and Addis, D. (2018). Prediction and portfolio optimization in quantitative trading using machine learning techniques. In *Proceedings of the Ninth International Symposium on Information and Communication Technology*, SoICT 2018, pages 98–105, New York, NY, USA. Association for Computing Machinery.

Tsai, I.-C. (2017). Diffusion of optimistic and pessimistic investor sentiment: an empirical study of an emerging market. *International Review of Economics & Finance*, 47:22–34.

Tsantekidis, A., Passalis, N., Tefas, A., Kanniainen, J., Gabbouj, M., and Iosifidis, A. (2017). Forecasting stock prices from the limit order book using convolutional neural networks. In *2017 IEEE 19th Conference on Business Informatics (CBI)*, volume 1, pages 7–12. IEEE.

Tuyon, J. and Ahmad, Z. (2016). Behavioural finance perspectives on malaysian stock market efficiency. *Borsa Istanbul Review*, 16(1):43–61.

Verma, R., Baklaci, H., and Soydemir, G. (2008). The impact of rational and irrational sentiments of individual and institutional investors on djia and s&p500 index returns. *Applied Financial Economics*, 18(16):1303–1317.

Verma, R. and Verma, P. (2007). Noise trading and stock market volatility. *Journal of Multinational Financial Management*, 17(3):231–243.

Wang, G., Yu, G., and Shen, X. (2020). The effect of online investor sentiment on stock movements: An lstm approach. *Complexity*, 2020:4754025.

Wang, Y.-H., Keswani, A., and Taylor, S. (2006). The relationships between sentiment, returns and volatility. *International Journal of Forecasting*, 22(1):109–123.

Zhang, C. (2008). *Defining, modeling, and measuring investor sentiment*. Thesis.

Zhang, X. and Tan, Y. (2018). Deep stock ranker: A lstm neural network model for stock selection. In *International conference on data mining and big data*, pages 614–623. Springer.