CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Early Detection of Dementia using Activities of Daily Life Data from Smart Homes

by

Sara Ibrahim

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

in the

Faculty of Computing
Department of Computer Science

2023

This thesis is dedicated to my parents, sister, and friends for their unwavering support. Special thanks to my supervisor, Dr. Umair Rafique, for invaluable guidance and expertise in achieving this milestone.

CERTIFICATE OF APPROVAL

Early Detection of Dementia using Activities of Daily Life Data from Smart Homes

by

Sara Ibrahim

(MCS213030)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Raheel Zafar	NUML, Islamabad
(b)	Internal Examiner	Dr. Nayyer Masood	CUST, Islamabad
(c)	Supervisor	Dr. Umair Rafique	CUST, Islamabad

Dr. Umair Rafique
Thesis Supervisor
September, 2023

Dr. Abdul Basit Siddiqui
Head
Dept. of Computer Science
September, 2023

Dr. M. Abdul Qadir
Dean
Faculty of Computing
September, 2023

Author's Declaration

I, **Sara Ibrahim** hereby state that my MS thesis titled "**Early Detection of Dementia using Activities of Daily Life Data from Smart Homes**" 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.

(Sara Ibrahim)

Registration No: MCS213030

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Early Detection of Dementia using Activities of Daily Life Data from Smart Homes**" 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.

(Sara Ibrahim)
Registration No: MCS213030

Acknowledgement

In the Name of Allah, the Most Merciful, the Most Compassionate, all praise be to Allah, the Lord of the worlds; and prayers and peace be upon Mohammad, His servant, and messenger. First and foremost, I humbly acknowledge my boundless gratitude to Allah, the Ever-Magnificent and Ever-Thankful, for His guidance, help, and blessings. I am steadfastly certain that this work would never have come to fruition without His divine guidance.

I am profoundly grateful to those individuals who diligently collaborated with me from the project's inception to the culmination of the present research, particularly my supervisor, Dr. Umair Rafique. His constant generosity and guidance have been a guiding light through all phases of this research. I would like to take this moment to extend heartfelt gratitude to my friends and colleagues, whose unwavering support has been a driving force throughout the course of my thesis.

Lastly, I express my sincere and wholehearted appreciation to my family for their unwavering support throughout my entire life journey, and especially during the pursuit of my master's degree. Their unconditional love and prayers have provided me with the opportunity to complete this thesis.

(Sara Ibrahim)

Abstract

Dementia is a neuro-degenerative condition that is characterized by cognitive impairment, memory loss, and a decline in daily functioning. The rise in dementia cases in elderly individuals living alone has become a significant global concern. Traditional methods mainly depend on subjective recall-based evaluations, which include questionnaires and observations of changes in behavior by caregivers. Monitoring Activities of Daily Living (ADL) data from sensor-based smart home environments is a promising alternative that can help in diagnosing early dementia.

Current state of the art work tries to solve this problem as anomaly detection where an episode of irregular behavior is considered an anomaly compared to the more frequent behavior (which is considered normal). There are two problems with this approach however.

Firstly, existing approaches struggle to differentiate between various types of anomalies in behavior and often fail to reliably distinguish whether unusual behavior is due to dementia or other factors, such as fever or injury. Secondly, existing techniques assume that the data included both regular and irregular behavior. But what if an individual has dementia from the beginning? In such cases, there might not be any noticeable change because their "normal" behavior is already affected by dementia.

To address these gaps, the proposed approach focuses on isolating anomalies specifically associated with dementia, targeting dementia-specific irregularities. An Autoencoder model, trained on normal data, is employed to establish a baseline for normal behavior. Subsequently, the model is tested with dementia cases developed using medical knowledge about dementia. Two separate experiments were conducted using various attribute combinations on both the Aruba Testbed from the Center for Advanced Studies for Adaptive Systems (CASAS) [1] and the Ordonez Activities of Daily Living dataset from the UCI Repository [2] to assess the model's performance. Autoencoders are particularly well-suited for this task, as they can effectively capture patterns and identify dementia symptoms from activities of daily life data without relying on labeled information. They reconstruct

input data with minimal error, effectively highlighting anomalies. Through the utilization of Autoencoders, the proposed approach demonstrates exceptional performance, achieving accuracy rates above 90% in reconstructing the original data on both datasets. Notably, the model effectively detects various dementia-related scenarios, providing an accuracy rate close to 0%. This occurs because the model, originally trained on normal data, exhibits errors when dementia-specific cases are intentionally introduced during the reconstruction process, causing a drop in accuracy. This showcases its ability to recognize abnormal patterns. The insights gained from this study are helpful in providing valuable foundational information for enhancing the objective diagnosis of dementia.

Contents

List of Figures

List of Tables

Abbreviations

ADL	Activities of Daily Living
AE	Autoencoder
CASAS	Center for Advanced Studies for Adaptive Systems
CDR	Dementia Rating scale
MSE	Mean Squared Error
RE	Reconstruction Error
SED-11Q	Early Dementia-11 Questionnaire

Symbols

r	Reconstruction Error
$\hat{X}$	Reconstructed Vector
X	Original Input Vector
σ	Activation Function
W	Weight Matrix
b	Bias Vector
H	Hidden Representation

Chapter 1

Introduction

1.1 Introduction to the Domain

Every country in the world is experiencing growth in both the size and the proportion of older persons in the population. Older adults compose a larger proportion of the world's population than ever before. According to World Health Organization (WHO), by 2030, 1 in 6 people in the world will be aged 60 years or over [3]. The reduction in ability to perform activities of daily living (ADL) due to older age can adversely affect the quality of living of an elderly person [4]. Additionally, the presence of a disease like dementia can significantly worsen a person's mental abilities and is now a growing concern worldwide. Dementia is a general term for loss of memory, language, problem-solving and other thinking abilities that are severe enough to interfere with daily life [5]. According to WHO [6], [7] there are around 50 million people worldwide living with dementia, and this number is expected to triple by 2050.

1.1.1 Early Detection of Dementia

An early detection in case of dementia is of prime importance [8] [9]. Although there is no current treatment to prevent or cure the disease, early detection can help in implementing measures that reduce or prevent further progression. The

key issue here is that the identification of cognitive decline due to dementia is not dependant on an individual's behaviour at a single point in time, rather it requires monitoring trends and changes in behavior over a period of time [10]. Behavioral changes, such as sleep disturbances, forgetting to perform some activities, an inability to complete tasks, are common indicators of certain types of dementia like Alzheimer's. Hence a change in the resident's behavior, for example forgetting to eat lunch or repeating certain activities, can indicate early signs of dementia.

Traditional methods to diagnose dementia predominantly rely on questionnaires and in-person examinations. The care giver of the elderly person is asked questions about changes in the behavior of the elderly person. One such questionnaire used by [11] is shown in Figure 1.1. Once a person has been identified as showing signs of dementia, further examination can be performed to diagnose the type and severity of dementia.

The biggest drawback of the traditional approach is that it may fail to detect dementia at an early stage. Firstly, there may not be anyone living in close contact with the elderly person to identify signs of dementia. Secondly, even if there is someone, they may not be able to identify subtle changes in the behavior of the elderly person. Hence once the elderly person is taken to the doctor for detailed diagnosis, they may already have passed that early stage of dementia.

1.1.2 Smart Home Environments

An aging population, suffering from issues like dementia, means an increase in the number of people requiring assistive care while a decrease in the number of care givers which can put a huge strain on the health care system. This has led researchers to investigate new approaches to keep older people independent in their own homes for as long as possible [12].Equipping homes with different types of sensors that gather useful data from the surrounding environment is one way to monitor the well-being of the residents [13] [14] (by detecting anomalies). This not only facilitates the identification of anomalies but also enables the implementation of proactive measures to optimize living conditions and promoting a healthier life.

Symptoms of Early Dementia-11 Questionnaire (SED-11Q)

Date(MM/DD/YYYY) / /

Patient Name : Patient ID :

Respondent Name : Relationship

Respondent-completed / Interview by Name:

How has the patient's daily life been for the last month?
Please answer the following questions by circling the appropriate responses
(Exclude any difficulties caused by physical issues, e.g., pain).
Please ask for any help if needed.

He/she talks and asks about the same things repeatedly.	YES	NO	N/A Don't know
He/she has become unable to understand the context of facts.	YES	NO	N/A
He/she has become indifferent about clothing and other personal concerns.	YES	NO	N/A
He/she has begun to forget to turn off the faucet and/or close the door, and/or has become unable to clean up properly.	YES	NO	N/A
When doing two things at the same time, he/she forgets one of them.	YES	NO	N/A
He/she has become unable to take medication under proper management.	YES	NO	N/A
He/she has begun to take a longer time to do work (e.g., household chores), which could be done quickly before.	YES	NO	N/A
He/she has become unable to make a plan.	YES	NO	N/A
He/she cannot understand complex topics.	YES	NO	N/A
He/she has become less interested and willing, and stopped hobbies, etc.	YES	NO	N/A
He/she has become more irritable and suspicious than before.	YES	NO	N/A
TOTAL SED-11Q SCORE			

He/she has delusions, e.g., claims to have had valuables stolen.	YES	NO	N/A
He/she has illusions, e.g., sees something that isn't there.	YES	NO	N/A

If the answer is "yes" to either of these 2 questions, then a more comprehensive medical consultation is recommended.

FIGURE 1.1: A Sample Questionnaire for the Diagnosis of Dementia

Such sensor based environments, also referred to as Smart Homes, can enable elderly individuals to remain self-reliant in their own residences for a longer period

of time [15]. Some commonly used sensors are (See Figure 1.2):

- **Motion sensors:** These can be used to detect movement and monitor activity levels in different areas of the home, such as the bedroom, bathroom, and kitchen.
- **Bed sensors:** These sensors can be placed under the mattress to monitor sleep patterns and detect any changes in sleep duration or quality.
- **Smart door locks:** These locks can be used to monitor when a person enters and leaves the home, providing data on daily routines and changes in behavior.
- **Smart appliances:** Appliances such as refrigerators and stoves can be used to monitor a person's use of the kitchen, including meal preparation and consumption.
- **Environmental sensors:** These can be used to monitor temperature, humidity, and air quality in different areas of the home, providing data on comfort levels and changes in behavior.

FIGURE 1.2: Commonly Used Sensors [16]

Figure 1.3a shows the layout of sensors in a smart home environment known as Aruba testbed. Figure 1.3b shows a sample of data indicating the time a given

sensor changes its state. Changes in the states of a set of sensors in a sequence can be interpreted as a particular activity. For example when motion sensors M003, M005 and M004 detect motion in a sequence, it will indicate the activity of "Bed to Toilet" as shown in the figure. This activity data of the residents of a smart home collected through sensors is what we refer to as Activities of Daily Living (ADL) data.

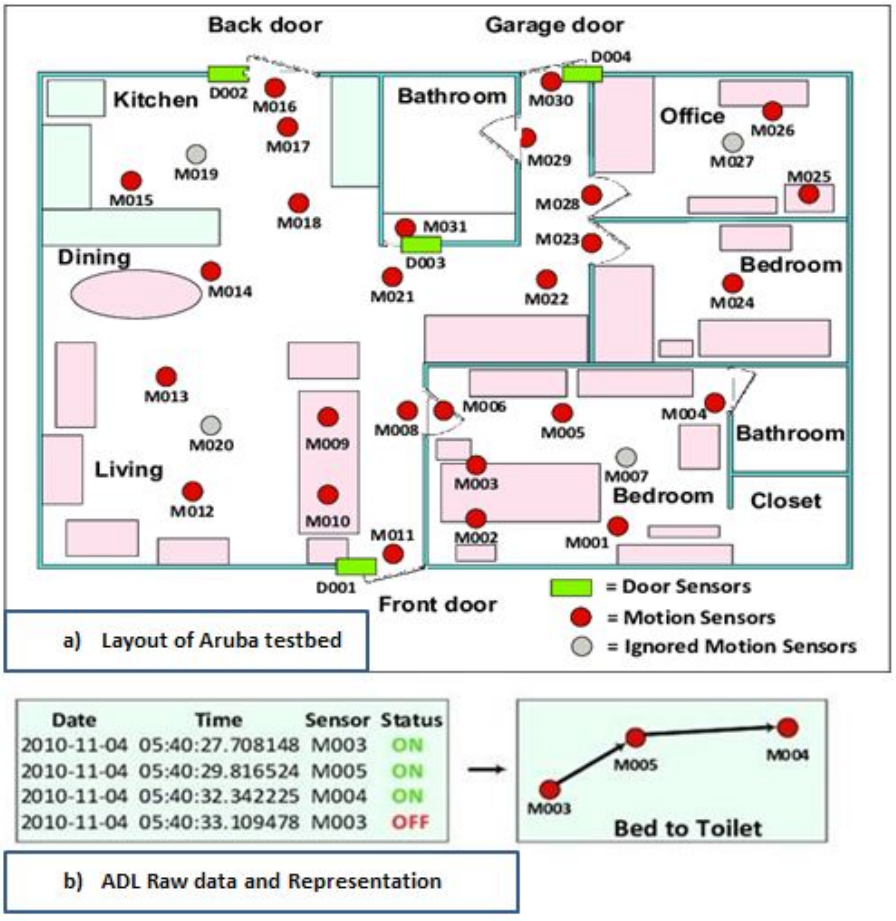

FIGURE 1.3: (a) A Layout of a Smart Home Environment (b) Raw Data and its Interpretation as an Activity[1]

1.2 Using ADL Data for Early Detection of Dementia

ADL data, collected in sensor-based environments (or smart homes) as discussed in the previous section, can be used to identify early signs of dementia. Such ADL

data collected over a period of time provides detailed insights into the daily routine of the resident of a smart home, like which activities they perform in a day, at what times, for how long they perform an activity etc. Whereas a normal person may have a certain daily routine (a sequence of activities they usually perform in a day), it is well known from literature in medical that a dementia sufferer exhibits typical behavior patterns [17]. One such pattern is performing an activity (or a sequence of activities) repeatedly [18] , [19] , [20] because of forgetfulness. For example a dementia sufferer may forget having taken medicine and then repeats it one or more times.

By looking for such patterns in the behavior of the resident of a smart home, we can identify if they are showing early signs of dementia. However, we cannot just look for dementia-specific behavior patterns (for example repetitive behavior) in the data as the number of possible such patterns is huge (possibly infinite). Additionally, even a normal person may exhibit some repetitive behavioral pattern. For example a person may routinely like Watching TV, Eating, Watching TV, Eating, etc. Hence for such an approach to work, we must have some kind of representation of "normal" behavior (which may include some normal repetitive behavior). Then we can identify dementia-specific behavior as being different from this normal behavior (a deviation). The problem is that different people may have different "normal" behavior. They may have different routines and may spend different amounts of time performing different activities. Hence the key issue here is to decide what constitutes "normal" behavior so that we can compare the behavior of a resident against it while looking for dementia-specific behavioural patterns.

One option is to look for deviations in a resident's behavior as compared to their own normal behavior. That means we must have a record of the normal behavior of a resident under consideration and then some instances of dementia-specific "abnormal" behavior which we must be able to identify as being different from their normal behavior. However it is possible that such record of normal behavior may not be available as for example the resident may already be suffering from dementia when we start recording ADL data. Also in practice this is not how initial

assessments regarding dementia are made. A doctor or a caretaker, without having a detailed history of the patient's normal behavior, can still identify dementia specific patterns in their behavior. A doctor or a care giver seems to have an implicit understanding of what constitutes "normal" behavior and can compare the behaviour of the patient against it. Even us, ordinary people not trained in medicine, can identify a person behaving "not in a normal way" even if we see them the first time.

Hence an alternate option is to construct a representation of "normal" behavior and then identify dementia specific behavior as a deviation from this normal behavior. [21] have prepared one such representation of normal behavior after consultation with doctors in the form of a table listing normal start and end times of activities (see Figure 1.4). Any behavior that is different from this normal behavior will be considered not normal. Our argument is that any such rule based representation is very rigid and may not represent the whole spectrum of normal behaviors. To address this problem, we have sought help from Machine Learning.

Activity	Contextual Information					
	Frequency		Duration		Start Time	
	Min	Max	Min	Max	Min	Max
Sleeping	2	3	8 h	12 h	22:30	23:30
Leaving (walk/exercise)	1	2	30 min	60 min	10:30	15:30
Toileting	4	8	1 min	10 min	-	-
Breakfast	1	1	15 min	35 min	9:30	10:30
Lunch	1	1	25 min	50 min	13:00	15:30
Dinner	1	1	25 min	50 min	18:30	21:30
Showering	1	3	10 min	15 min	-	-

FIGURE 1.4: Normal Behaviour Representation used in [21].

1.3 Learning Normal Behavior from ADL Data

The basic idea behind our approach is to use ADL data of smart home residents to develop a representation of their "normal" behavior. We can then compare specific instances of behaviors with this representation to see if they are normal or not. If we had ADL data both for normal behavior and dementia-specific behavior then this was a classical machine learning problem with two classes "normal" and

"showing signs of dementia". However no such data set is publicly available. Actually, there is one recent work that has tried to prepare such a data set [22]. However they have not made their data set public and did not answer our requests to share their data set either. A number of data sets of ADL data from smart homes are publicly available for example Aruba Testbed from CASAS data set, and Ordonez's ADL Activity data set from UCI Repository [1], [2] and have been used in literature quite extensively for different purposes. However, these data sets include ADL data only for "normal" residents. There is no publicly available data set that has ADL data of dementia sufferers.

We have used the above mentioned two data sets to develop a representation of "normal" behavior. As a first step we transformed raw ADL data in the form of "daily routines" where a daily routine represents a sequence of activities performed by a person in a single day. Then we developed an algorithm employing regular expressions to summarise these daily routines focusing on repetitive patterns of activities. This transformation allowed us to capture the repetitive behavior of a person from ADL data and allowed us to create features which are highly relevant for dementia prediction. We then trained an Autoencoder-Decoder model [23] on this data. Autoencoder-Decoder is an unsupervised Machine Learning algorithm which transforms the input to a latent space (encoding part) and then reconstructs it back to the original data (decoding part). That means the input and output of an Encoder-Decoder model is the same and hence no explicit labelling is required (hence it is unsupervised).

During the training, the Encoder-Decoder model "learns" a representation of the data in the form of its parameters (weights of the neural network). The trick here is to train the Encoder-Decoder only on ADL data of "normal" people (who are not suffering from dementia) so that it can learn the representation of the normal behavior.

Once the Autoencoder-Decoder has been trained on this data, it has learnt the representation of "normal" behavior. It can now encode and decode daily routines of normal people with an accuracy close to 100%. However when it is presented with an instance of a daily routine of a dementia sufferer, it must not be able to

reconstruct it correctly producing an error. In other words this behavior pattern is like an anomaly for the model. In this way we can single out scenarios where a person's behavior exhibits signs of dementia. We have discussed the details of our approach in Chapter 3. To evaluate our model we simulated certain scenarios that represented abnormal behaviors specific to individuals with dementia and presented them to our model. Our model was able to single them out as anomalies with 100% accuracy. We discuss the details of our proposed methodology in Chapter 4.

1.4 Problem Statement

The growing number of cases of dementia among elderly population worldwide highlights the need for a cost-effective way to detect dementia at an early stage. The prospect of an increasingly smaller number of caregivers due to aging population has led to the investigation of using smart home environments for collecting ADL data of its residents which can then be used to identify signs of dementia (among other things). Our work focuses on how to use this ADL data to develop a representation of what we consider as "normal" behavior and how to differentiate it from some behavior exhibiting signs of dementia.

1.5 Research Questions

1. Given the particular case of having ADL data only of healthy people, how can the Autoencoder machine learning algorithm be used to develop a representation of "normal" behavior which can then be matched against ADL data of dementia sufferers?

2. Can ADL data of dementia sufferers that includes dementia-specific behavioral patterns be distinguished from ADL data of "normal" people using the selected Autoencoder machine learning method?

In addressing these research questions, this study aims to contribute valuable insights into the application of the Autoencoder machine learning algorithm for constructing 'normal' behavior from ADL data. This construction is intended to facilitate the detection of dementia-specific behavioral patterns within the Activities of Daily Living (ADL) data.

1.6 Objectives of the Research

1. Develop a robust predictive model utilizing machine learning techniques to effectively detect and predict early signs of dementia using ADL data from smart home environments.

2. Evaluate the effectiveness and accuracy of the predictive model through rigorous testing and validation on real-world ADL data from smart homes.

3. Contribute to the field of dementia research by developing a cost-effective and practical approach that empowers caregivers, medical professionals, and individuals to recognize and address potential cognitive deterioration due to dementia, ultimately improving the overall quality of life for the elderly population.

1.7 Organization of Thesis

Chapter 2 presents an extensive review of the existing literature, providing insights into the research landscape. A comparative analysis and survey of existing techniques is presented, highlighting the strengths and weaknesses of various approaches. This assessment helps identify gaps in the current knowledge and sets the stage for the subsequent research.

In Chapter 3 a detailed discussion of the proposed methodology is presented, including a thorough explanation of the selected data sets and models. The rationale behind the choices made is discussed.

Chapter 4 provides a detailed discussion of the experimental setup. The obtained results are presented, analyzed, and discussed in detail. The implications and significance of the findings are thoroughly examined, shedding light on the insights gained from the conducted experiments.

Chapter 5 summarizes the key findings and contributions. Implications of the results are discussed, along with potential avenues for future research and further exploration.

Chapter 2

Literature Review

Within the context of smart homes equipped with sensors (as discussed in the previous chapter), the activities of a resident are recorded in the form of raw sensor activation data. These records simply include which sensor was turned on (i.e. sensor activation) and at what time. Once a sensor is activated due to some activity (for example a motion sensor getting activated due to the movement of the resident), it automatically turns off after a specified amount of time (generally a few seconds). A snapshot of raw sensor data from the Aruba test bed [1] is shown in Figure 2.1. A number of sensors activated in a sequence may represent a certain activity being performed. For example the sequence given in Figure 2.1 represents the activity of going from Bed to Toilet and then returning. In this context the first task is to label the sequences of sensor activations as meaningful activities. This task in literature is known as Activity Recognition [24]. Different researchers have explored different sensor modalities and classification algorithms for the purpose of activity recognition (see for example [25], [26], [27], [28]).

A labelled data set in this context will have a sequence of sensor activations (and de-activations) along with a label indicating the start and the end of an activity. Another way can be to keep start time, end time and label of an activity without the detailed sensor activation data. Activity Recognition is not our focus in this work and we have used the versions of the datasets already labelled with start and end times of activities.

```
2010-11-04 05:40:51.303739 M004 ON
2010-11-04 05:40:52.342105 M005 OFF
2010-11-04 05:40:57.176409 M007 OFF
2010-11-04 05:40:57.941486 M004 OFF
2010-11-04 05:43:24.021475 M004 ON
2010-11-04 05:43:26.273181 M004 OFF
2010-11-04 05:43:26.345503 M007 ON
2010-11-04 05:43:26.793102 M004 ON
2010-11-04 05:43:27.195347 M007 OFF
2010-11-04 05:43:27.787437 M007 ON
2010-11-04 05:43:29.711796 M005 ON
2010-11-04 05:43:30.279021 M004 OFF
```

FIGURE 2.1: A Snapshot of Raw Sensor Data from Aruba Testbed. Column 1 is Date, Columns 2 is Time, Column 3 is Sensor ID and Column 4 is the sensor state.

A list of activities performed in a day forms the routine of the resident. It has been observed that a resident generally has a similar routine from one day to another and a deviation from this routine may indicate some problem with the resident's health. A number of approaches have focused on this issue and in literature it is generally referred to as "Anomaly Detection". Anomaly Detection focuses on detecting irregularities in the daily routine, which may indicate underlying health problems.

2.1 Irregular Behavior Identification as Anomaly Detection

Understanding and identifying irregular behavior patterns has gained increasing significance, particularly in the context of anomaly detection and cognitive decline. This literature review explores a range of studies studies that emphasize the detection of anomalies within daily activities, providing insights into the field of identifying irregular behaviors. Starting with the work of Viorica et al. [29], the challenge of identifying adaptable daily schedules among older individuals is

tackled. Their approach integrates collaborative clustering and the Gap-BIDE algorithm to uncover frequent behavior patterns, while accommodating variations in activity sequence and duration. This pioneering effort seeks to identify irregularities that could potentially indicate underlying health concerns. Similarly, Ahmad Lotfi et al. [30] explore the domain of next activity prediction and anomalous behavior detection, offering valuable insights into irregular behavior identification. Lotfi et al. emphasize the efficacy of recurrent neural networks and the Echo State Network technique in predicting forthcoming sensor activities, showcasing the potential to foresee behavioral sequences. This research, underscores the importance of early anomaly detection in enhancing various contexts, including individual routines and behavior.

With the growing recognition of the significance of irregular behavior patterns, the domain of multi-resident activity monitoring has attracted significant attention. This emphasis on detecting anomalies within shared living environments. Within this domain, a multitude of studies have emerged, each offering diverse and valuable perspectives on the identification of irregular behavior patterns. For instance Fahad et al. [31] introduce a novel approach employing probabilistic neural networks and autoencoders for anomaly detection within smart homes. Their study not only classifies pre-segmented activity instances but also identifies anomalies, enhancing the detection of unexpected behavior patterns in multi-resident contexts. Furthering the exploration of multi-resident activity recognition, Jinghuan et al. [32] adopt a time clustering approach. By extracting features from datasets through de-noising techniques, they employ similarity matching mechanisms to recognize two-resident activities, adding nuanced insights to irregular behavior identification in shared living environments.

The utilization of Hidden Markov models (HMM) for anomaly detection in the literature has also gained significance due to their ability to effectively capture and analyze complex behavioral patterns. Researchers have increasingly recognized the potential of HMM in identifying irregularities, particularly within the context of monitoring Activities of Daily Living (ADL) data and detecting deviations. Notably, both Sanchez et al. [33] and Chifu et al. [34] used the Hidden Markov

models for anomaly detection, each with a distinct focus and approach. Sanchez et al. [33] direct their attention towards human behavior modeling within the domain of welfare technology. By incorporating Hidden Markov models, they enable the recognition of abnormal behavior patterns through the analysis of factors like person location, posture, and time frame rules. This approach empowers their model to accurately detect irregular behavior and offer timely assistance when necessary. Similarly, Chifu et al. [34] pivot towards the recognition of deviations in individual routines, particularly among older individuals. Through the integration of beacon technology and Markov models, they harness Hidden Markov models to uncover irregular behavior patterns within daily routines. This work contributes to a deeper understanding of anomalies that might indicate cognitive impairments or other health-related concerns.

Moving beyond behavior recognition, Ujager et al. [21] introduce the concept of wellness determination through activity monitoring within smart homes. By analyzing abnormal behavioral patterns and generating alarms, this study underscores the importance of identifying irregularities that could potentially lead to emergency situations. In addition to the concept of wellness determination, now many researchers are increasingly directing their focus towards cognitive impairments, utilizing Activities of Daily Living (ADL) data as a means of detecting diseases that often manifest in the routine behaviors. This approach recognizes that subtle variations in daily activities can provide valuable insights into cognitive health. Consequently, studies have emerged with the aim of utilizing ADL data to identify irregular behavior patterns indicative of cognitive decline or other related conditions.

In this context, the works of Arifoglu et al. [35] and Alaghbari et al. [36] stand out as noteworthy contributions. Arifoglu et al. employ Convolutional Neural Networks and LSTM to detect anomalies within ADL data, specifically focusing on early dementia-related behavior patterns. This endeavor highlights the potential to capture deviations indicative of cognitive decline in daily activities. Similarly, Alaghbari et al. present a comprehensive sensor-based deep learning model tailored for monitoring older individuals with cognitive impairments. Their approach,

characterized by deep neural networks and autoencoders, not only detects anomalies but also predicts subsequent activities. This common thread of addressing cognitive impairments underscores the crucial role of irregular behavior identification in flagging potential cognitive decline through comprehensive behavioral analysis.

Among the efforts to utilizing the ADL data for finding irregularities in behaviour, Kwon et al. [22] have conducted a study aimed at objectively evaluating activities of daily living to diagnose early-stage dementia in senior citizens.

However, there were limitations to the study, including a small sample size of data. One notable limitation of Kwon et al.'s approach lies in its inherently restricted framework for identifying patterns of normal and abnormal behavior. By exclusively categorizing individuals into two groups—those classified as normal and those diagnosed with dementia—the model's understanding of "normalcy" and "abnormality" becomes confined to this specific demographic. This limited perspective might hinder the generalizability of their findings across broader populations, as the patterns identified may only hold true within the bounds of their defined groups.

However, an alternative approach that avoids the strict categorization of "Normal" and "Dementia Sufferer" offers a more generalized solution. Instead of imposing predefined labels, their model could be structured to recognize patterns that are common among a diverse range of individuals without solely relying on the demographics of normal versus dementia cases. By training the model on a dataset encompassing a larger pool of normal individuals and incorporating the well-established patterns associated with dementia, the resulting model could transcend demographic variations. This shift would enable the model to focus on detecting deviations from the norm, which are indicative of potential cognitive impairments, without being overly confined by the demographic identity of the individual being assessed.

In summary, the current approach's dependence on specific labels might limit its applicability beyond the studied population. Adopting an approach that utilizes

the the vast dataset of normal individuals and the recognizable dementia patterns for anomaly detection can yield a model that is better suited for diverse demographic groups and is less susceptible to being overly influenced by demographic characteristics.

2.2 Comprehensive Analysis of Existing Anomaly Detection Techniques using ADL Data

The summary of existing work is shown in a tabular form in Table 2.1.

TABLE 2.1: Summary of Existing Techniques

Ref.	**Method**	**Features**	**Datasets**	**Objective**
[29], 2022	Deep learning, GAP-BIDE algorithm, collaborative clustering	ADL timespan, start time, end time	WISDM smartphone and smartwatch Activity and Biometrics Datasets	A single flexible routine composed of mandatory activities, optional activities, alternative variants of activities
[31], 2021	probabilistic neural networks for classifying pre-segmented activity instances and H2O autoencoder for detecting anomalies	the anomalous days based on the number of activities performed per day. the boxplots of the number of Features and the duration of activities.	CASAS dataset Aruba and Melan	anomaly detection

[32], 2020	de-noising method to extract the feature, time clustering to separate activities that occur at the same space but at different times, and similarity matching formula based on Levenshtein Distance for daily activity recognition	ADL timespan, start time, end time	Tulum2010 and Cairo	feature extraction, temporal clustering, and activity recognition
[34], 2022	Markov model for identify the daily routines, entropy rate and cosine functions for similarity between the daily monitored activities in a day and the inferred routine. A Distributed monitoring using Beacons and trilateration for elder activity.	length of the monitored activities and transition probabilities among activities as relevant features	Collected by sensors	Propose a solution for detecting seniors' daily routine deviations from monitored activities, requiring caregivers' intervention.
[30], 2012	Echo State Network (ESN), Back Propagation Through Time (BPTT) and Real Time Recurrent Learning (RTRL)	ADL timespan, start time, end time	Collected by sensors	Anomaly detection and next activity prediction.

[35], 2019	Convolutional Neural Networks (CNNs) to model patterns in activity sequences, LSTM for learning the activity sequences and the behavioral routine	dataset is segmented into time-slices by using a sliding window approach as	Aruba , WSU testbeds from CASAS smart home datasets	detect the following 3 different kinds of anomalies that can be seen in daily-life routines of elderly people with dementia: 1. Repeating activities 2. Disruption in sleep 3. Confusion (getting confused during the activities)
[21], 2019	the lazy associative classifier (LAC) for the frequent behavioral patterns	Start time of activity, sub activity, location, Duration	Ordonez's ADL Dataset	abnormal behavior detection
[36], 2022	deep neural network (DNN), overcomplete-deep autoencoder (OCD-AE) and long shortterm memory (LSTM) network	duration, No. of sensors activated during an activity, the number of times an activity is performed	Aruba and Cairo from CASAS labelled activities dataset	activity recognition, anomaly detection and next activity prediction.

[33], 2020	Hidden Markov model (HMM) to predict a person's behavior	person's location in the house, posture, and time frame rules, to detect abnormal behavior	two datasets, OrdonezA and OrdonezB	abnormal behavior
[22], 2021	Developed a personalized model using a random forest algorithm to reflect individual daily patterns.	Movements and duration of activity.	IoT and 2D LIDAR sensors to collect data from smart homes and developed a personalized model	diagnose and early-stage dementia prediction

2.3 Identified Research Gaps

In the previous section, various existing approaches have used ADL data to identify irregular behavior of a resident, where this irregular behavior is generally seen as an anomaly to their regular behavior. After a detailed analysis of the problem and the existing approaches, two issues with these approaches have been identified. Firstly, these approaches cannot differentiate between different types of anomalies in behavior. An episode of irregular behaviour (the anomaly) could be due to any health condition like fever or an injury and not necessarily the result of the resident's suffering from dementia. The approaches above cannot differentiate between different types of anomalies and hence cannot isolate anomalies due to dementia. Hence, although anomaly detection can help us identify episodes of irregular behavior (which is enough for the resident's wellness determination, for example), we cannot attribute these episodes to dementia. On the other hand, it is well known that a dementia sufferer exhibits certain behavioral patterns (as discussed in Chapter 1) like forgetting to perform some activities they normally perform or repeating certain activities. In other words an episode of irregular

behavior due to dementia has certain characteristics which must be taken into account while identifying signs of dementia using ADL data. Hence an approach that focuses on anomalous behavior only due to dementia, focusing on dementia specific irregularities, is needed.

Secondly, the whole concept of anomaly detection is based on the idea that the ADL data we have of a resident includes both "regular" behavior data as well as "irregular" behavior data (anomalies). Consider a scenario where the resident is suffering from dementia from the start of the recording of ADL data. In this case the resident may not exhibit any "change" in behavior and their "normal" behavior is already affected by dementia. However any approach based on anomaly detection will not be able to identify this case as there is no anomaly (or deviation from a regular pattern of behavior).

Hence a new approach to the detection of dementia based on ADL data is needed. This approach should be able to isolate behavior with signs of dementia (case 1 discussed above) and must also be able to handle the scenarios when the ADL data only includes abnormal behavior (case 2 discussed above).

Chapter 3

Research Methodology

In this chapter, the research methodology of proposed work is discuss, which consists of two main phases. Phase 1 serves as the foundational stage, encompassing vital components such as problem identification, research problem formulation, thorough literature review, and the identification of key research gaps. This phase establishes a solid foundation by offering a comprehensive understanding of the current state of research. This understanding allows us to place our study within the broader context of the field. Moving into Phase 2, our methodology transitions into a more data-driven and action-oriented approach. This phase involves the systematic collection of relevant data, followed by rigorous preprocessing to ensure its quality and consistency. Subsequently, feature engineering techniques were employed to extract meaningful information from the data, enhancing its suitability for analysis. With a well-prepared dataset in hand, the implementation of the chosen model progressed to achieve the research objectives.This structured approach guides the research process and ensures a systematic exploration of the research topic.

The flow of research methodology phases are depicted in Figure 3.1.

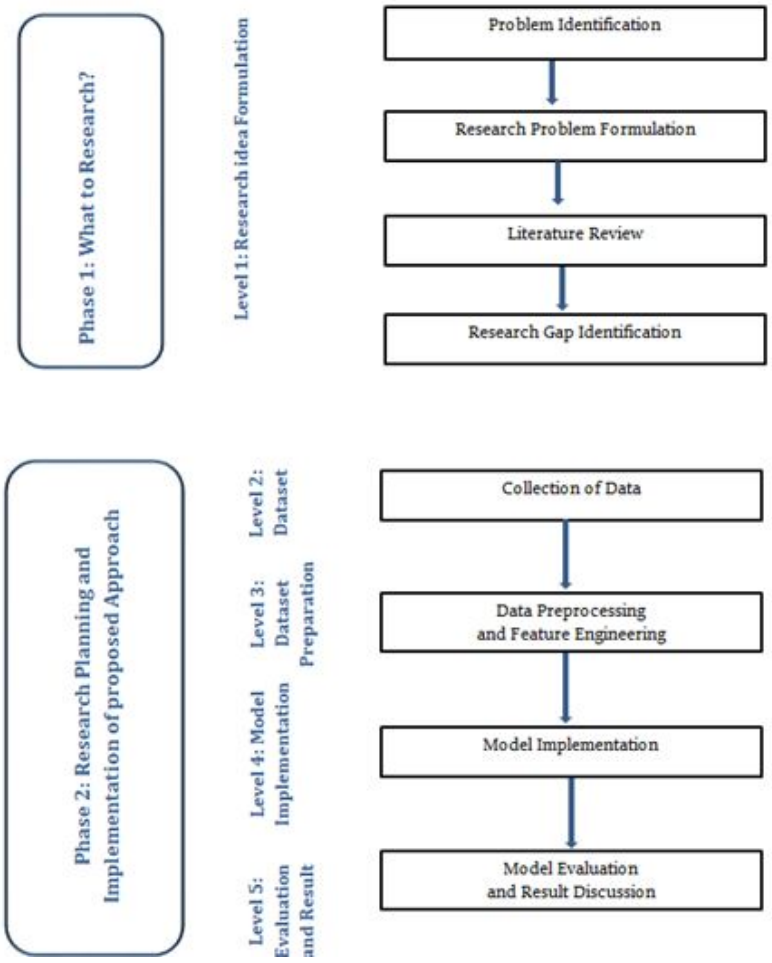

FIGURE 3.1: Research Methodology Framework

3.1 Research Idea Formulation

In this phase of our research methodology, we systematically navigate through four crucial steps: Problem Identification, Research Problem Formulation, Literature Review, and Research Gap Identification. Each step serves as a building block, guiding us from recognizing the problem to refining research objectives and identifying gaps in existing knowledge.

3.1.1 Problem Identification

In the initial phase of the research, the focus is squarely on identifying a compelling problem that demands attention and investigation. The chosen topic was, 'Early Detection of Dementia using Activities of Daily Life (ADL) Data from

Smart Homes'. The reason behind selecting this topic is the alarming rise in cases of dementia among the elderly population. In the context of a growing elderly population, many individuals are affected by conditions such as dementia, leading to an increased demand for assisted care. Meanwhile, the availability of caregivers has become limited, leading to challenges within the healthcare system. This situation has encouraged researchers to seek innovative methods that allow elderly to maintain their independence in their own residences for an extended period. Through a thorough examination of current literature and discussions, the growing need for a cost-effective means of detecting dementia at its early stages was pinpointed. The output of this phase was a clear understanding of the problem statement, outlining the significance of addressing dementia detection using ADL data.

3.1.2 Research Problem Formulation

In the second phase, the research problem was formulated based on the insights gained from problem identification. The focus was on the increasing cases of dementia in the elderly population and the potential of utilizing smart home environments for collecting ADL data. This step allowed for the precise definition of research objectives and the outlining of the study's goals. The output of this step was a well-defined research problem that guided subsequent actions.

3.1.3 Literature Review

Chapter 2 provides an extensive review of the literature related to the detection of irregular behavior patterns in the context of smart homes equipped with sensors. The following is a concise summary of the key points discussed in this literature review.

Within the context of smart homes equipped with sensors, resident activities are recorded as raw sensor activation data, indicating which sensor was turned on and at what time. Sequences of sensor activation's represent activities performed by

residents. Labeling these sequences as meaningful activities is known as Activity Recognition. This literature review primarily focuses on detecting irregular behavior patterns, referred to as Anomaly Detection. Irregular behavior in daily routines may indicate underlying health issues.

Researchers have explored various approaches for identifying irregularities, by employing models like Hidden Markov Models (HMM), Autoencoders, Long Short-Term Memory (LSTM) networks, and Convolutional Neural Networks (CNN) to identify irregular behavior within daily activities. Each of these models serves a distinct purpose: HMMs are effective at capturing complex behavioral patterns, Autoencoders are effective for anomalies detection, dimensionality reduction and data compression , LSTMs are valuable for predicting sensor activities, and CNNs analyze Activities of Daily Living (ADL) data for deviations indicative of irregular behaviour. Despite these efforts, two primary gaps persist: The current studies cannot differentiate between different types of anomalies in behavior. Secondly, the whole concept of anomaly detection is based on the idea that the ADL data we have of a resident includes both "regular" behavior data as well as "irregular" behavior data (anomalies). Hence a new approach to the detection of dementia specific anomalies based on ADL data is needed.

so the output of this step was a synthesized understanding of the existing landscape, enabling us to position our research within the context of previous work and identify gaps that our study could address.

3.1.4 Research Gap Identification

Building upon the insights from the literature review, we identified specific gaps in the existing body of knowledge. The existing literature has highlighted two significant research gaps in the context of using Activities of Daily Living (ADL) data for identifying irregular behavior patterns, particularly related to dementia.

Firstly, current approaches struggle to differentiate between various types of anomalies in behavior, making it challenging to attribute irregular behavior solely to

dementia. A more specific focus on dementia-related irregularities is needed. Secondly, the conventional anomaly detection framework assumes the coexistence of "regular" and "irregular" behavior data, which poses limitations when individuals have dementia from the beginning, as their "normal" behavior is already affected. Addressing these gaps will enhance the effectiveness of using ADL data for dementia detection.

These gaps represented areas where previous research had not fully addressed the nuances of dementia detection using ADL data. By analyzing the limitations and unexplored aspects of prior work, we were able to discover a unique niche for our research. The output of this step was a clear understanding of the research gaps that our study aimed to fill, setting the stage for our research's original contribution.

3.2 Research Planning and Implementation of Proposed Approach

Phase 2 of research methodology involves practical implementation of our proposed work. The proposed methodology has been structured to address two primary research questions (RQs), which are outlined as follows:

1. Given the particular case of having ADL data only of healthy people, how can the Autoencoder machine learning algorithm be used to develop a representation of "normal" behavior which can then be matched against ADL data of dementia sufferers?

2. Can ADL data of dementia sufferers that includes dementia-specific behavioral patterns be distinguished from ADL data of "normal" people using the selected Autoencoder machine learning method?

The objective of RQ1 is how to develop a representation of normal behavior containing data only from healthy patients. To address this question, existing work

on ADL data was examined, along with exploration of various unsupervised ML models and datasets containing sufficient data instances suitable for creating a representation of normal behavior. This representation can then be compared to the daily activity data of dementia sufferers.

The objective of RQ2 is how to create features that enable the selected model to distinguish dementia-specific behavioral patterns from ADL data of "normal" people. To address this question, exploration of the medical literature was conducted to gain insights into the behavior exhibited by dementia patients during ADL. This understanding guides the creation of features that assist the model in recognizing signs of dementia in ADL data.

Phase 2 describes the proposed solution methodology that addresses these two research questions. Details of implementing these two objectives are discussed in Chapters 3 and 4.

The conceptual design of our proposed solution is depicted in Figure 3.2.

3.2.1 Datasets

In response to the research objectives outlined in Chapter 1, this section presents a comprehensive and innovative approach designed to utilizing the data collected from smart home environments. The proposed solution aims to leverage this information, particularly through the utilization of sensors, to effectively monitor individuals' daily activities. By pursuing this approach, the aim is to identify potential signs of dementia

Two publicly available datasets of ADL from smart homes were used. The descriptions of these Datasets are given below in table 3.1. The selection of appropriate datasets is a critical aspect of any research study. The datasets chosen for this research include the Aruba Testbed from CASAS [1] and Ordonezs ADL Activity dataset from UCI Repository [2].

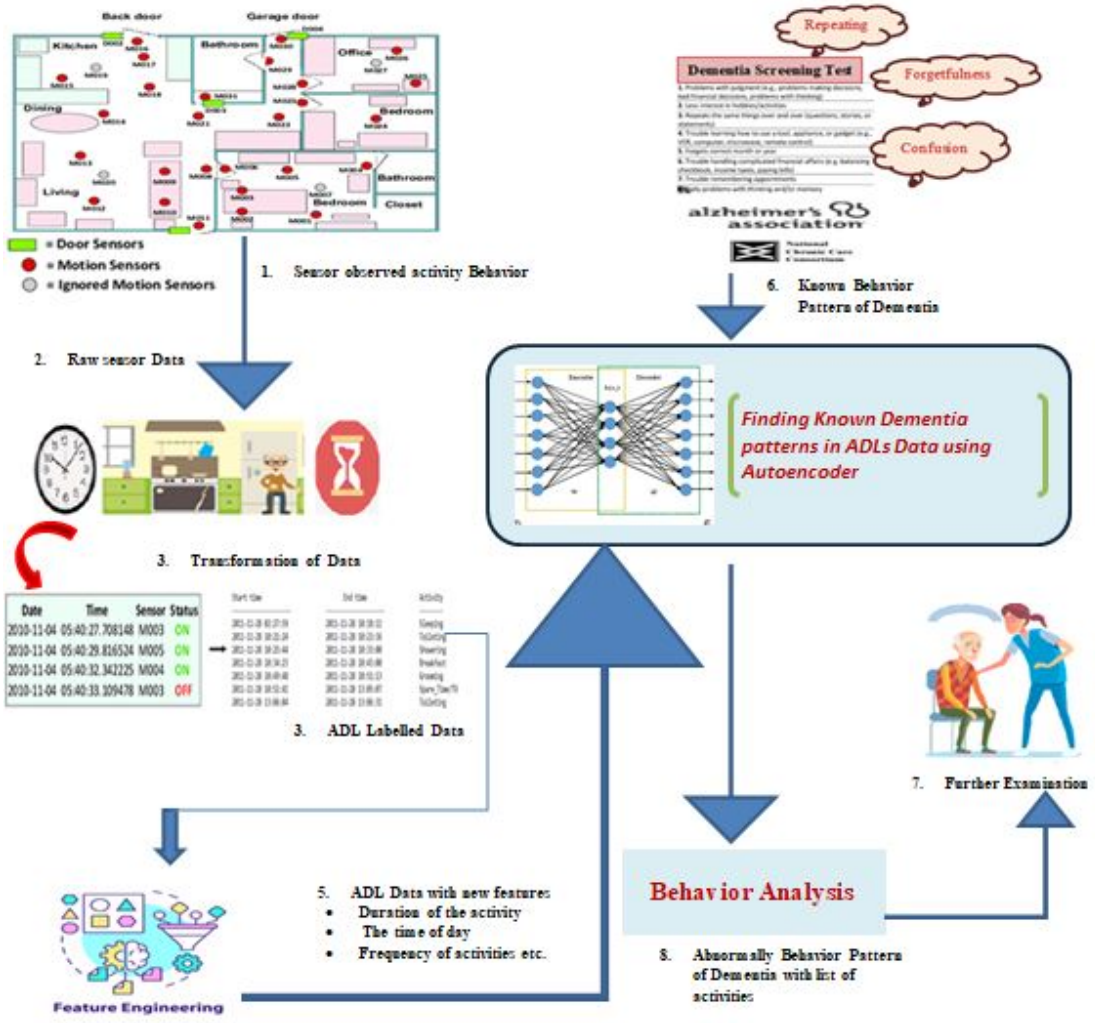

FIGURE 3.2: Conceptual Design of proposed Methodology for Early Dementia detection

The chosen datasets contain ADL data from smart homes, which is relevant to the research question of predicting dementia from ADL data of elderly people in smart homes. Description of each dataset is given below.

3.2.1.1 Aruba Testbed

In Aruba testbed, motion, door and temperature sensors are used. The data is provided as a list of (sensor, time-stamp) sensor measurements. In this dataset, there are 11 daily activities performed by a single user and it spans 224 days. The description of Aruba ADLs and instances is given in Table 3.2.

TABLE 3.1: Activities of Daily Life Data sets

Serial Number	**Datasets**	**Period**	**Activities (1 participant)**	**Instances**
1	Aruba	Nov.04,2010-Jun.11,2011	11	6467
2	Ordonez A	Nov.28,2011-Dec.11,2011	10	182
2	Ordonez B	Nov.11,2012-Dec.02,2012	10	273

TABLE 3.2: Aruba Testbed ADLs Instances

Serial Number	**ADLs**	**Instances**
1	Meal_Preparation	1606
2	Relax	2910
3	Eating	257
4	Work	171
5	Sleeping	401
6	Wash_Dishes	65
7	Bed_to_Toilet	157
8	Enter_Home	431
9	Leave_Home	431
10	Housekeeping	33
11	Resperate	6

3.2.1.2 Ordonez's ADL Dataset

Ordonez's dataset comprises information regarding the ADLs performed by two users on a daily basis in their own homes. 12 different sensors attached on different locations in home to record the ADLs. This dataset is composed by two files of data, each one corresponding to a different user and summing up to 35 days of fully labelled data.

The description of ADLs and sensors for the two subjects is given in Table 3.3. While these datasets may not contain information specifically related to dementia,

TABLE 3.3: ADLs with associated sensors in Ordonez's dataset

Serial Number	**ADLs**	**Installed Sensors**
1	Leaving	Magnetic (main door)
2	Toileting	Passive Infrared (PIR) (basin), flush (toilet)
3	Showering	PIR (showering)
4	Sleeping	Pressure (bed)
5	Breakfast	PIR (cooktop, microwave), electric (toaster), magnetic (fridge, cabinet, cupboard)
6	Lunch	PIR (cooktop, microwave), magnetic (fridge, cabinet, cupboard)
7	Dinner	PIR (cooktop, microwave), magnetic (fridge, cabinet, cupboard)
8	Snack	Electric (microwave, toaster), magnetic (fridge)
9	Spare time/TV	Pressure (seat)
10	Grooming	Magnetic (cabinet)

they provide valuable information about the daily activities of elderly people in a smart home environment. This information can be used to identify patterns of abnormal behavior associated with dementia, such as skipping activities, repetitive behavior, and confusion.

3.2.2 Data Preprocessing and Feature Engineering

The collected data is preprocessed and features are engineered to prepare it for analysis. This involves cleaning, transforming, and selecting relevant features as shown in Figure 3.3.

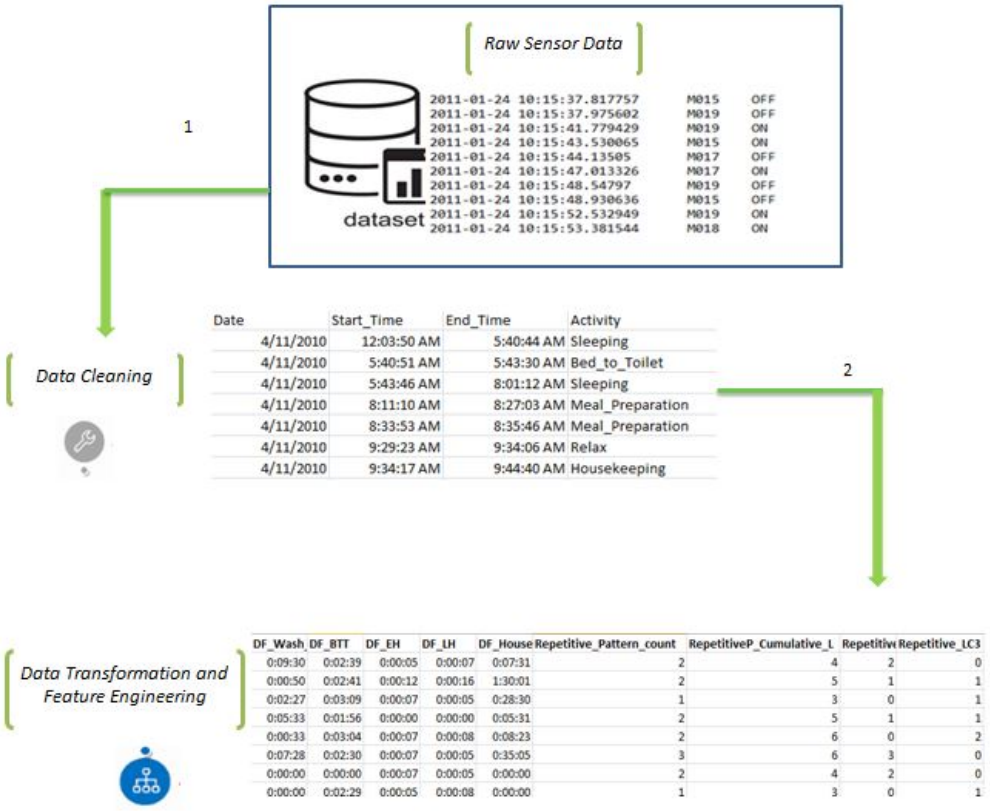

FIGURE 3.3: Data preprocessing and Feature Engineering

3.2.2.1 Data Cleaning and Transformation

In the process of Data Cleaning and Transformation for dementia prediction, several steps are undertaken to filter out the unnecessary information.
The first step involves separating the date and time from the activity start and end labels, establishing a clear distinction between the temporal aspects and the specific activities being performed.

Next, the process continues with the extraction of relevant features from the data, such as duration and frequency count. This entails identifying the key variables or attributes that have a significant impact on dementia prediction. These features serve as the basis for subsequent analysis and modeling. To ensure comparability and consistency across different data points, the extracted data is then normalized. Normalization, a technique that scales the data to a standard range, helps eliminate potential biases arising from varying measurement units or scales. This critical step enhances the reliability of subsequent analyses, promoting a uniform foundation for meaningful comparisons among diverse datasets and facilitating

more accurate insights into the data. Implementing the data cleaning and transformation of data into daily routines aims to refine the dataset by filtering out unnecessary information, extracting relevant features, and normalizing the data. This process lays the foundation for accurate and reliable dementia prediction models.

3.2.2.2 Feature Engineering

Initially, the dataset includes the Date, start time, end time, and activity features. As shown in Figure 3.4. However, in order to improve the effectiveness of dementia

Date	Start_Time	End_Time	Activity
4/11/2010	12:03:50 AM	5:40:44 AM	Sleeping
4/11/2010	5:40:51 AM	5:43:30 AM	Bed_to_Toilet
4/11/2010	5:43:46 AM	8:01:12 AM	Sleeping
4/11/2010	8:11:10 AM	8:27:03 AM	Meal_Preparation
4/11/2010	8:33:53 AM	8:35:46 AM	Meal_Preparation
4/11/2010	9:29:23 AM	9:34:06 AM	Relax
4/11/2010	9:34:17 AM	9:44:40 AM	Housekeeping
4/11/2010	9:48:52 AM	9:53:03 AM	Meal_Preparation
4/11/2010	9:54:59 AM	9:56:27 AM	Meal_Preparation
4/11/2010	9:56:42 AM	9:59:04 AM	Eating
4/11/2010	9:59:48 AM	10:02:49 AM	Eating
4/11/2010	10:03:22 AM	10:04:26 AM	Wash_Dishes
4/11/2010	11:29:08 AM	11:33:48 AM	Housekeeping

FIGURE 3.4: Initial Feature set

detection, it is important to incorporate additional relevant features. So the Duration and frequency count of each activity is computed and data is transformed in the form of daily routines in 1st Transformation as show in Table 3.4 and 3.5. This involves creating a sequence of activities that a person performs in a day. This sequence is very helpful because it shows us the order of activities, how many times each activity happens, and how much time an elderly person spend on each one.

This kind of information is like looking at a person's whole day in one picture.

Days	D_Meal_Preparation	D_Relax	D_Eating	D_Work	D_Sleeping	D_Wash_Dishes	D_Bed_to_Toilet	D_Enter_Home	D_Leave_Home	D_Housekeeping	D_Respirate
4-Nov	1:07:18	1:04:00	0:10:27	0:25:08	7:54:20	0:19:01	0:02:39	0:00:10	0:00:13	0:15:03	0:00:00
5-Nov	0:08:01	0:09:45	0:03:30	1:02:09	13:20:18	0:00:50	0:02:41	0:01:12	0:01:34	1:30:01	0:00:00
6-Nov	0:21:30	3:29:38	0:29:55	0:26:47	4:30:15	0:04:53	0:03:09	0:00:21	0:00:14	0:56:59	0:00:00
7-Nov	0:31:40	8:00:52	1:17:53	0:18:29	13:11:01	0:05:33	0:05:48	0:00:00	0:00:00	0:05:31	0:00:00
8-Nov	0:35:18	7:08:39	0:53:44	0:25:50	2:58:52	0:00:33	0:03:04	0:00:37	0:00:41	0:08:23	0:00:00
9-Nov	0:46:56	4:06:09	0:42:09	1:23:26	8:04:20	0:07:28	0:05:01	0:00:21	0:00:14	1:10:11	0:00:00
10-Nov	0:36:38	4:20:41	0:22:13	0:00:00	12:18:43	0:00:00	0:00:00	0:00:07	0:00:05	0:00:00	0:01:44
11-Nov	0:33:21	4:08:31	0:25:36	0:30:37	9:14:33	0:00:00	0:02:29	0:00:09	0:00:15	0:00:00	0:00:00
12-Nov	0:35:54	6:29:03	0:25:27	1:04:22	8:58:01	0:00:00	0:03:25	0:00:23	0:00:23	0:11:54	0:00:00
13-Nov	0:32:03	5:41:50	0:20:10	2:04:42	10:54:57	0:01:20	0:03:47	0:00:00	0:00:00	0:00:00	0:01:38
14-Nov	0:30:11	5:49:24	0:07:51	0:00:00	0:00:00	0:02:58	0:00:00	0:00:05	0:00:04	0:00:00	0:00:00
15-Nov	0:35:35	2:41:26	0:37:41	0:06:21	2:45:57	0:07:15	0:03:28	0:00:44	0:00:49	0:03:11	0:00:00
16-Nov	0:55:38	8:10:24	0:10:24	1:10:05	8:50:47	0:05:50	0:01:58	0:00:10	0:00:36	0:36:22	0:00:00
17-Nov	0:29:02	3:53:23	0:18:42	0:00:00	6:37:03	0:06:50	0:03:52	0:00:18	0:00:06	0:05:41	0:00:00
18-Nov	0:19:23	5:20:17	0:39:28	0:35:21	9:33:04	0:00:58	0:04:21	0:00:10	0:00:13	0:11:09	0:00:00
19-Nov	0:09:27	8:30:51	0:18:04	0:00:00	6:46:21	0:00:00	0:00:00	0:00:05	0:00:05	0:00:00	0:00:00
20-Nov	0:29:00	5:27:52	0:53:38	0:26:38	8:36:42	0:20:05	0:04:06	0:00:00	0:00:00	0:00:00	0:00:00
21-Nov	0:14:12	5:24:59	0:00:00	1:11:01	6:21:34	0:00:00	0:04:32	0:00:50	0:00:42	0:00:00	0:00:00
22-Nov	0:28:30	10:15:54	0:04:01	0:00:00	0:44:13	0:09:22	0:03:15	0:00:00	0:00:00	0:25:32	0:00:00
23-Nov	0:10:58	9:55:06	0:00:00	0:03:43	8:06:30	0:00:00	0:03:04	0:00:06	0:00:05	0:02:31	0:00:00

TABLE 3.4: 1st Transformation Feature set (1-12)

D_Respirate	FC_Meal_Preparation	FC_Relax	FC_Eating	FC_Work	FC_Sleeping	FC_Wash_Dishes	FC_Bed_to_Toilet	FC_Enter_Home	FC_Leave_Home	FC_Housekeeping
0:00:00	5	5	3	2	2	2	1	2	2	2
0:00:00	3	2	1	3	3	1	1	6	6	1
0:00:00	2	3	3	2	1	2	1	3	3	2
0:00:00	6	9	10	4	4	1	3	0	0	1
0:00:00	5	6	3	3	2	1	1	5	5	1
0:00:00	5	5	4	1	3	1	2	3	3	2
0:01:44	5	4	3	0	1	0	0	1	1	0
0:00:00	8	6	3	3	2	0	1	2	2	0
0:00:00	6	4	3	3	2	0	1	5	5	1
0:01:38	3	4	2	1	2	1	1	0	0	0
0:00:00	5	6	1	0	1	1	0	1	1	0
0:00:00	4	3	2	1	2	1	1	8	8	1
0:00:00	5	7	1	2	2	1	1	2	2	2
0:00:00	2	1	2	0	2	2	2	2	2	1
0:00:00	3	4	2	1	2	1	1	2	2	1
0:00:00	1	2	1	0	1	0	0	1	1	0
0:00:00	3	4	2	2	4	2	2	0	0	0
0:00:00	1	3	0	1	3	0	2	9	9	0
0:00:00	5	6	1	0	1	1	1	0	0	1
0:00:00	1	4	0	1	2	0	1	1	1	1

TABLE 3.5: 1st Transformation Feature set (13-23)

It provides insights when they do different things, how much time they spend on each task, and how their day is divided. This big picture helps us find any unusual patterns or changes in their routine. This is important for spotting any signs of dementia early on. To explore further, let's understand the significance of the frequency count and duration features.
Frequency count tells us how often a person performs a specific activity. For instance, if someone is forgetting to do a task that they used to do regularly, it might be a sign of cognitive decline.
On the other hand, the duration feature reveals how much time a person spends on each activity. This can help us notice if they are taking longer to complete tasks that were once done quickly.

Combining these features into daily routines offers a comprehensive view of a person's behavior. By analyzing the sequence of activities and their frequencies and durations, we can spot any irregularities. These could include sudden changes in the number of times an activity is performed or unexpected variations in the time spent on an activity. Such changes might indicate the onset of dementia or other cognitive issues. Therefore, these features and their arrangement into daily routines hold crucial insights for accurate early detection.

After the first transformation, combining the duration and frequency count features for each activity was done to provide a more comprehensive understanding of the data. For example, calculating the duration of sleeping and dividing it by the frequency count of sleeping. This merging of information allows us to gain insights into the average duration or time spent on each activity, providing a clearer picture of the patterns and behavior observed. Combining these features aims to enhance the interpretability and usefulness of the data. The final transformation Features set shown in Table 3.6.

These features, including frequency count and duration, are certainly valuable. However, incorporation of additional features is necessary that specifically represent signs of dementia, which can manifest through ADL data, such as repetitive behavior.

Days	DF_MP	DF_Relax	DF_Eating	DF_work	DF_Sleeping	DF_Wash_Dishes	DF_BTT	DF_EH	DF_LH	DF_Housekeeping	Repetitive_Pattern_count	RepetitiveP_Cumulative_L	Repetitive_LC2	Repetitive_LC3
4-Nov	0:13:28	0:12:48	0:03:29	0:12:34	3:57:10	0:09:30	0:02:39	0:00:05	0:00:07	0:07:31	2	4	2	0
5-Nov	0:02:40	0:04:53	0:03:30	0:20:43	4:26:46	0:00:50	0:02:41	0:00:12	0:00:16	1:30:01	2	5	1	1
6-Nov	0:10:45	1:09:53	0:09:58	0:13:24	4:30:15	0:02:27	0:03:09	0:00:07	0:00:05	0:28:30	1	3	0	1
7-Nov	0:05:17	0:53:26	0:07:47	0:04:37	3:17:45	0:05:33	0:01:56	0:00:00	0:00:00	0:05:31	2	5	1	1
8-Nov	0:07:04	1:11:27	0:17:55	0:08:37	1:29:26	0:00:33	0:03:04	0:00:07	0:00:08	0:08:23	2	6	0	2
9-Nov	0:09:23	0:49:14	0:10:32	1:23:26	2:41:27	0:07:28	0:02:30	0:00:07	0:00:05	0:35:05	3	6	3	0
10-Nov	0:07:20	1:05:10	0:07:24	0:00:00	12:18:43	0:00:00	0:00:00	0:00:07	0:00:05	0:00:00	2	4	2	0
11-Nov	0:04:10	0:41:25	0:08:32	0:10:12	4:37:17	0:00:00	0:02:29	0:00:05	0:00:08	0:00:00	1	3	0	1
12-Nov	0:05:59	1:37:16	0:08:29	0:21:27	4:29:01	0:00:00	0:03:25	0:00:05	0:00:05	0:11:54	2	5	1	1
13-Nov	0:10:41	1:25:27	0:10:05	2:04:42	5:27:28	0:01:20	0:03:47	0:00:00	0:00:00	0:00:00	0	0	0	0
14-Nov	0:06:02	0:58:14	0:07:51	0:00:00	5:37:11	0:02:58	0:00:00	0:00:05	0:00:04	0:00:00	1	4	0	0
15-Nov	0:08:54	0:53:49	0:18:50	0:06:21	1:22:58	0:07:15	0:03:28	0:00:05	0:00:06	0:03:11	1	7	0	0
16-Nov	0:11:08	1:10:03	0:10:24	0:35:03	4:25:24	0:05:50	0:01:58	0:00:05	0:00:18	0:18:11	2	4	2	0
17-Nov	0:14:31	3:53:23	0:09:21	0:00:00	3:18:31	0:03:25	0:01:56	0:00:09	0:00:03	0:05:41	2	4	2	0
18-Nov	0:06:28	1:20:04	0:19:44	0:35:21	4:46:32	0:00:58	0:04:21	0:00:05	0:00:07	0:11:09	1	2	1	0
19-Nov	0:09:27	4:15:26	0:18:04	0:00:00	6:46:21	0:00:00	0:00:00	0:00:05	0:00:05	0:00:00	0	0	0	0
20-Nov	0:09:40	1:21:58	0:26:49	0:13:19	2:09:10	0:10:02	0:02:03	0:00:00	0:00:00	0:00:00	1	2	1	0
21-Nov	0:14:12	1:48:20	0:00:00	1:11:01	2:07:11	0:00:00	0:02:16	0:00:06	0:00:05	0:00:00	2	9	1	0
22-Nov	0:05:42	1:42:39	0:04:01	0:00:00	0:44:13	0:09:22	0:03:15	0:00:00	0:00:00	0:25:32	1	3	0	1
23-Nov	0:10:58	2:28:46	0:00:00	0:03:43	4:03:15	0:00:00	0:03:04	0:00:06	0:00:05	0:02:31	0	0	0	0
24-Nov	0:03:35	1:24:44	0:12:08	0:00:00	2:39:53	0:00:00	0:02:39	0:00:04	0:00:03	0:10:29	3	6	3	0
25-Nov	0:17:16	0:39:06	0:00:00	0:38:01	2:41:50	0:00:00	0:03:12	0:00:32	0:00:07	0:04:41	3	8	1	2

TABLE 3.6: Final Transformation Feature set

To comprehend dementia signs and behaviors, insights are drawn from various screening tests and methods employed for early dementia detection, such as the Early Dementia Screening Test Questionnaire. Among these methods, the SED-11Q (Symptoms of Early Dementia-11 Questionnaire) [37] is a common and extensively used test developed by the National Institutes of Health Japan.

The SED-11Q is designed as a screening test comprising 11 questions, with some questions being particularly relevant to daily living activities (ADL). Aligning with the approach, the following questions are selectively chosen to portray dementia behaviors. For example, like question 1 deals with the repetitive behavior of a person and question 2 deals with skipping an activity. Similarly, question3 when someone forgets to perform an activity and question4 deals with the confusions behavior as it take longer time to complete a particular task. These behaviors can be observed and detected through monitoring the activities of elderly individuals residing in smart homes.

1. He/ She perform a particular Task repeatedly.
2. When doing two things at the same time, he/she forgets one of them.
3. He/ She has begun to forget to turn off the faucet and/ or close the door, and/or has become unable to clean up properly.
4. He/ She has begun to take a longer time to do work (e.g., household chores), which could be done quickly before.

Additional features were engineered based on above questions, including: "Repetitive Pattern count, RepetitiveP Cumulative L, Repetitive LC2, and Repetitive LC3" . Theses features represent that how many repetitive patterns exist in a day routine and how lengthy these patterns are. The final engineered features include Days, DF_MP, DF_Relax, DF_Eating, DF_work, DF_Sleeping, DF_Wash_Dishes, DF_BTT, DF_EH, DF_LH, DF_Housekeeping, Repetitive_Pattern_count, RepetitiveP_Cumulative_L, Repetitive_LC2, and Repetitive_LC3. The interpretation of each feature is described in Table 3.7.

TABLE 3.7: Features Description for Aruba Dataset

Serial Number	**Feature Notation**	**Feature description**
1	Days	Date information for identifying Routines
2	DF_MP	average duration or time spent on Meal Preparation activity
3	DF_Relax	average duration or time spent on Relax activity
4	DF_Eating	average duration or time spent on Eating activity
5	DF_work	average duration or time spent on Work activity
6	DF_Sleeping	average duration or time spent on Sleeping activity
7	DF_Wash_Dishes	average duration or time spent on Wash Dishes activity
8	DF_BTT	average duration or time spent on Bed to Toilet activity)
9	DF_EH	average duration or time spent on Enter Home activity
10	DF_LH	average duration or time spent on Leave Home activity
11	DF_Housekeeping	average duration or time spent on Housekeeping activity
12	DF_Resperate	average duration or time spent on Resperate activity
13	Repetitive_Pattern_count	Count of Repetitive patterns in a day routine
14	RepetitiveP_Cumulative_L	Cumulative Length Count of Repetitive patterns in a day routine
15	Repetitive_LC2	Cumulative Count of Repetitive patterns with length 2 in a day routine
16	Repetitive_LC3	Cumulative Count of Repetitive patterns with length 3 in a day routine

These features have been carefully selected to capture various aspects of the activities and patterns observed in the dataset. To accurately capture patterns related for the features like Repetitive pattern count and Repetitive length, a technique is required to identify patterns within daily routines. In this study, the regular expressions is employed to achieve this objective.

Specifically, the the following regular expression is utilized shown in Figure 3.5, and the whole Algorithm is given below in Table 3.8.

$$\backslash matches = re.findall(r'(?:|(?<=,))(A\backslash d)(?:,(A\backslash d))$$
$$\backslash b(?:(?<=,)\backslash 1,\backslash 2|(?<=,)\backslash 2,\backslash 1)*', routine\_str)\backslash$$

FIGURE 3.5: Regular Expression for Capturing Repetitive Sequence

This regular expression enables the identification and extraction of patterns by searching for specific sequences of activities.

The regular expression scans the routine string, looking for occurrences where activities are repeated in a specific order. The expression captures these repeated patterns, considering both the forward and backward ordering of the activities. Applying this regular expression enables the effective identification and analysis of patterns within day-to-day routines, providing valuable insights into the repetitive nature of activities.

3.2.3 Model Selection

In this research, the Autoencoder architecture was used as a key tool for a primary objective: signs of dementia detection using ADLs by detecting anomalies in repetitive patterns within day routines. This section elucidates our rationale behind selecting the Autoencoder for these tasks and provides a description of the model's structure.

TABLE 3.8: Algorithm for Detecting Repetitive Patterns in Daily Routines

Algorithm: Capturing Repetitive Pattern from Day Routines
Input: Activities of daily living (ADLs) dataset
Output: Sequence of Patterns, Patterns count, and length of patterns

Steps:

1. Read the Activities of daily living (ADLs) dataset containing Activities.
2. Define the activity_mapping dictionary to map activities to shorter labels.
3. Initialize a dictionary to store routines for each day.
4. Extract days and map activities for each day, storing them in the routines dictionary.
5. Convert routines into a string format for pattern matching.
6. Initialize a pattern_dict dictionary to store repetitive patterns for each day.
7. Apply regular expression pattern matching to find activities that appear consecutively together in the day routines repeatedly.
8. Count the occurrences of repetitive patterns and store them in the pattern_dict dictionary, along with the number of times each pattern repeats.
9. Calculate the length of each repetitive pattern and store it in the pattern_dict dictionary.
10. Display the detected repetitive patterns, their counts, and lengths for each day using the pattern_dict dictionary.

3.2.3.1 Dementia Detection

The main goal was to predict dementia based on various features extracted from preprocessed and feature-engineered data. The features considered for the Autoencoder include:

Days: Representing the temporal dimension, allowing the model to learn patterns over time. *Average Duration or Time Spent of Each Activity:* Capturing activity-related information, which could be significant in detecting early signs of dementia.

Repetitive Patterns Features: We incorporated various features related to repetitive patterns in day routines, such as the number of patterns existing in the routines, cumulative length of patterns. These additions aim to capture and quantify the presence and extent of repetitive behaviors within the data, facilitating a thorough analysis of daily routines.

3.2.3.2 Why Autoencoder?

Autoencoders are a type of neural network designed for unsupervised learning and dimensionality reduction. The selection of an autoencoder for dementia prediction using daily activity data from smart homes is primarily due to the unsupervised nature of ADL data. Autoencoders are well-suited for such scenarios as they have the capacity to learn a representation of 'normal' behavior solely from the input data and subsequently identify dementia-related anomalies. They consist of two main components: an encoder and a decoder, both illustrated in Figure 3.6.

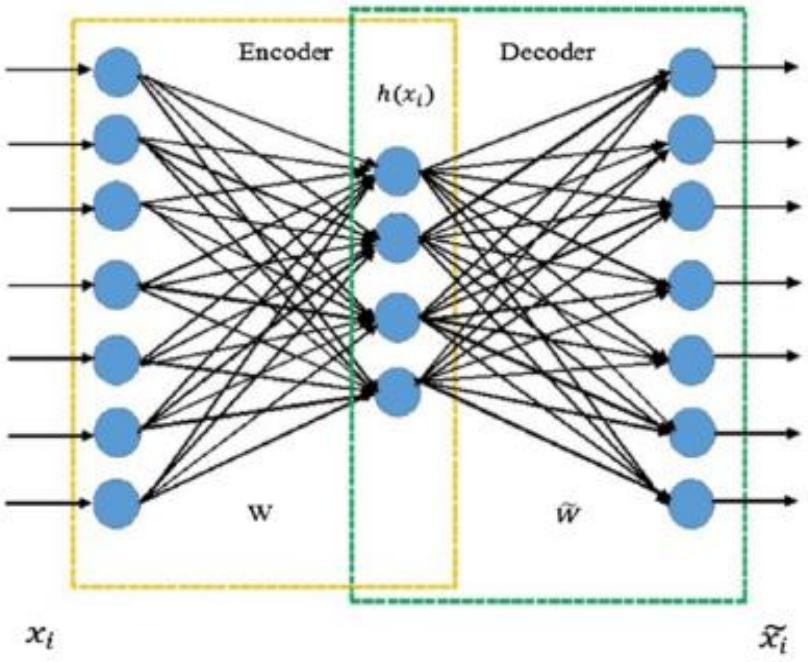

FIGURE 3.6: The Architecture of a sample Autoencoder

The encoder takes the input vector X and transforms it into a hidden representation H using Eq.3.1, where σ denotes an activation function, such as a sigmoid function or rectified linear unit, while W and b represent the weight matrix and bias vector, respectively. Subsequently, this hidden representation H undergoes a transformation operation in the decoder to reconstruct the initial input space using Eq.3.2.

$$H = \sigma(W_{xh}X + b_{xh}) \tag{3.1}$$

$$\hat{X} = \sigma(W_{h\hat{x}}H + b_{h\hat{x}}) \tag{3.2}$$

The reconstruction error (RE), denoted as "r," is determined by the difference between the reconstructed vector $\hat{X}$ and the original input vector X. In other words, RE quantifies how well the Autoencoder is able to reconstruct the input data. The equation for calculating RE is illustrated in Equation.3.3.

$$r = \|X - \hat{X}\| \tag{3.3}$$

The Autoencoder is trained to minimize the reconstruction error (r) using an unsupervised training approach [38]. The goal is to fine-tune the model's parameters and learn a compressed representation of the input data that can accurately reconstruct the original data. The training process of the Autoencoder is depicted in Figure. 3.7, illustrating the flow chart of how the model is updated to achieve its objective of minimizing the reconstruction error.

The reconstruction error (RE) serves as the crucial criterion for detecting anomalies in an Autoencoder (AE). During training, the AE is specifically trained to minimize this reconstruction error, allowing it to learn the intricate relationships between the features present in the input dataset. Once the AE is trained on a particular dataset, it should be able to accurately reproduce the input data when fed with similar data that it has never encountered before.

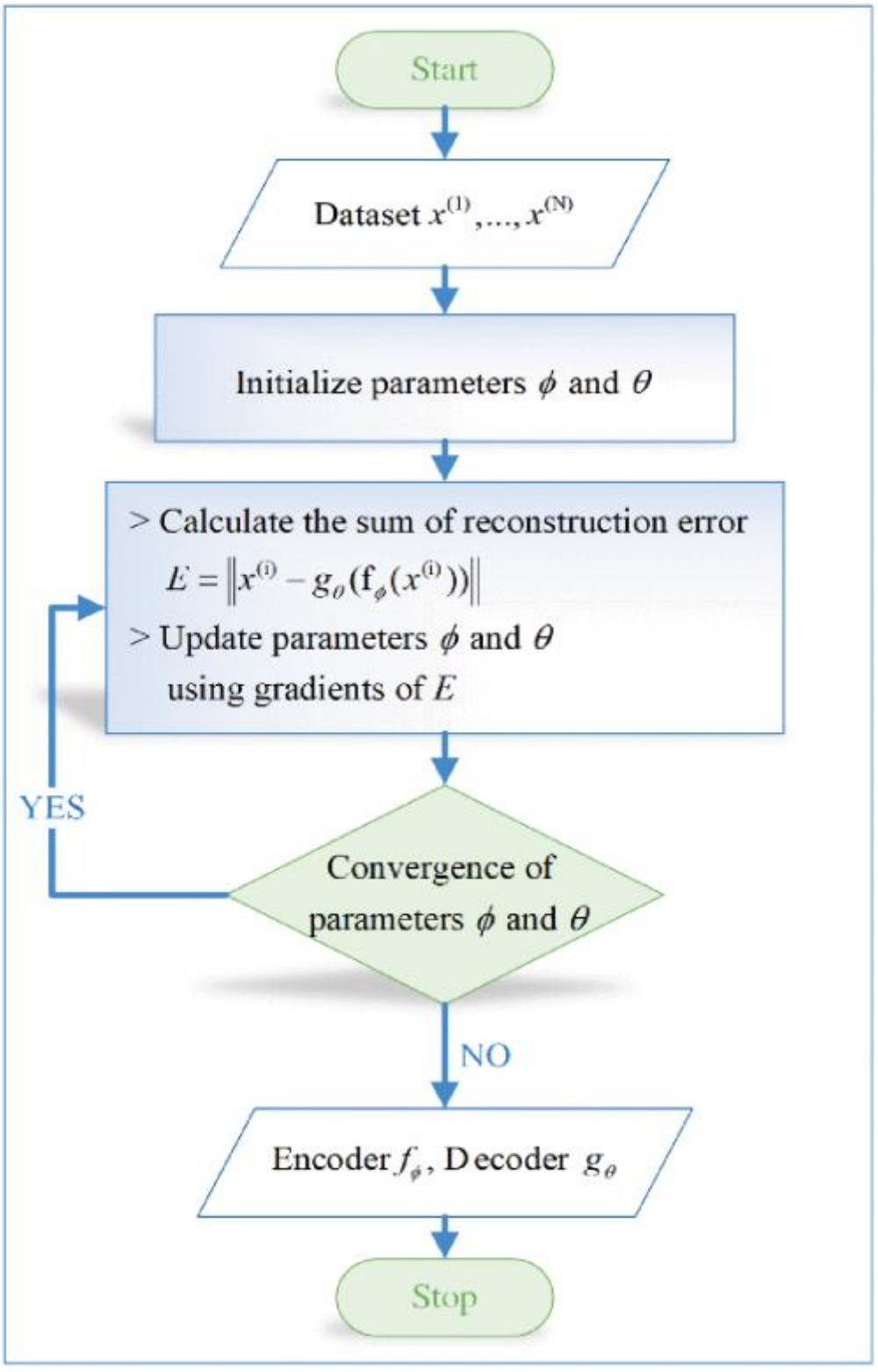

FIGURE 3.7: Autoencoder Training Algorithm

In such cases, the AE will generate a low reconstruction error, indicating a successful reconstruction of normal input data.

However, when the AE is presented with anomalous data that significantly differs from what it has seen during training, it will struggle to reconstruct the input correctly. As a result, the reconstruction error will be notably higher for the anomaly data compared to normal data. This characteristic of the AE forms the foundation for AE-based anomaly detection. By observing the magnitude of the reconstruction error, we can identify potential anomalies.

An input that generates a high reconstruction error is indicative of an anomaly, as it deviates significantly from the patterns seen during training. Figure 3.8, illustrates the flow chart of the AE-based anomaly detection algorithm [38].

This AE-based anomaly detection approach offers a powerful and effective means to identify abnormal patterns and behaviors in various applications.

3.2.4 Model Evaluation

In this section, the performance evaluation of the Autoencoder model for dementia detection on unsupervised Activities of Daily Living (ADLs) data is discussed. The Autoencoder was trained on a set of specific features, and a threshold for anomaly detection was established based on the reconstruction errors. Dealing with unsupervised data and the absence of a confusion matrix with True Positive, False Positive, False Negative, and True Negative, traditional evaluation matrices like Precision and Recall are not applicable. Therefore, the model's performance was assessed using accuracy as the evaluation metric.

The accuracy was determined based on the number of anomalies detected by the Autoencoder. It was calculated as the percentage of data points not identified as anomalies. Additionally, for further model testing, reconstruction was performed using only the cases flagged as anomalies by the Autoencoder, accurately representing the anomalies detected by the model.

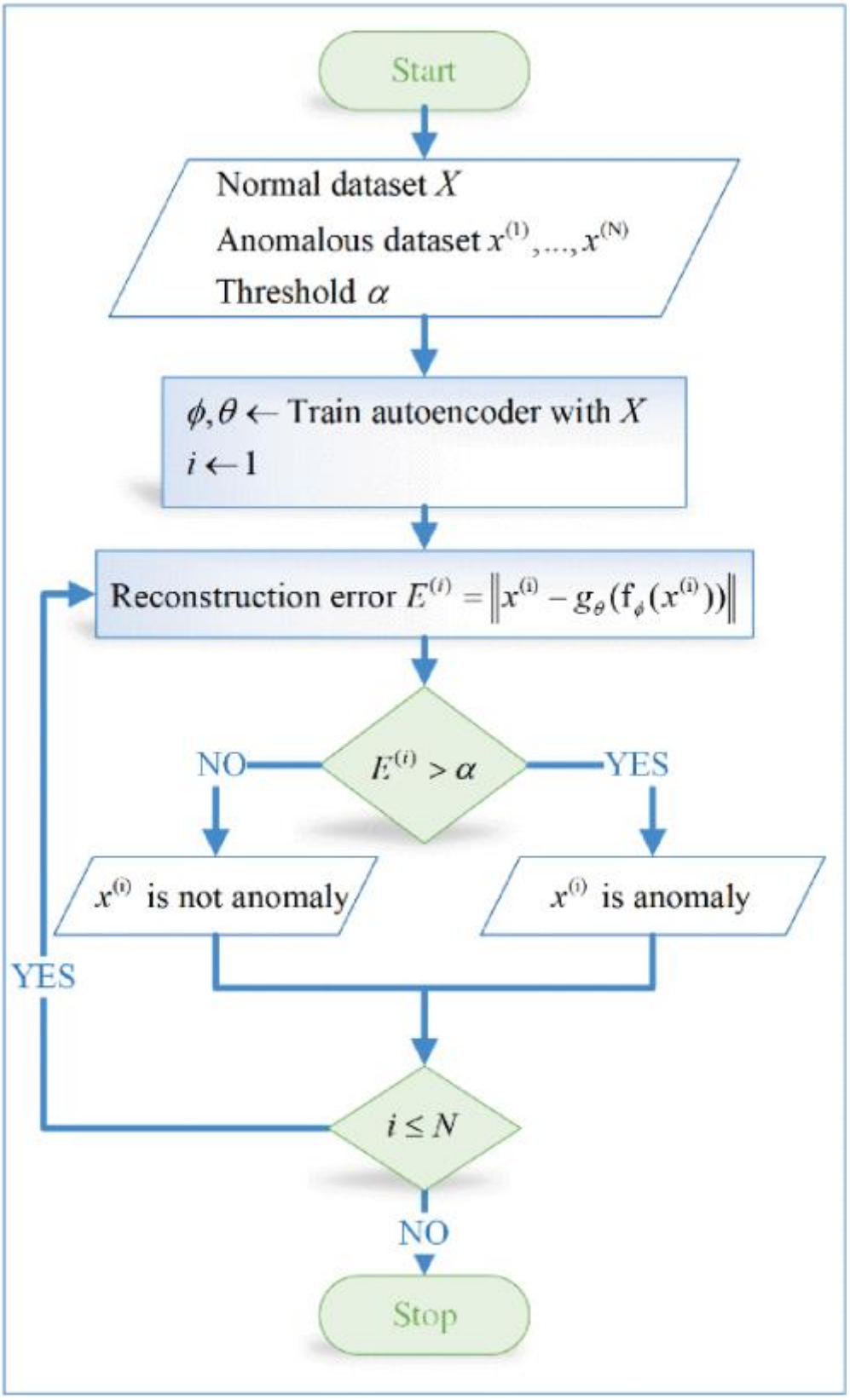

FIGURE 3.8: Autoencoder based Anomaly Detection

Using this approach, insights are gained into the Autoencoder's performance in identifying normal patterns and detecting repetitive anomalies (signs of dementia)

within the unsupervised data. Although precision and recall metrics are unavailable, this method allows us to approximate the overall effectiveness of the model in detecting anomalies without relying on labeled data.

Chapter 4

Experimental Setup and Results

In the previous chapter, the research methodology was discussed in detail, with a focus on addressing specific research questions. In this chapter, the focus will be on the validation of the proposed research questions through various experiments. The research questions are as follows:

1. Given the particular case of having ADL data only of healthy people, how can the Autoencoder machine learning algorithm be used to develop a representation of "normal" behavior which can then be matched against ADL data of dementia sufferers?

2. Can ADL data of dementia sufferers that includes dementia-specific behavioral patterns be distinguished from ADL data of "normal" people using the selected Autoencoder machine learning method?

The two data sets used in the study were Aruba Testbed from CASAS and Ordonezs ADL Activity dataset from UCI Repository. Additionally, the conversion of this ADL data into a format representing the daily routine of the residents and the engineering of new features to capture repetitive behavior were discussed in the previous chapter. In this chapter, the experimental setup and the results of the proposed study are discussed in detail, providing a comprehensive examination of the methodology and outcomes.

4.1 Tools and Technologies

The following tools and libraries were utilized for the implementation, Python used as the base language which allowed us to implement machine learning models and algorithms with ease.

Google Colaboratory: It provides an accessible environment for running Python code, facilitating the exploration and implementation of machine learning algorithms. Its integration with Google Drive allowed seamless access to data and notebooks from anywhere.

TensorFlow: TensorFlow served as the backbone of our model development process. Its powerful and flexible architecture allowed us to define complex neural network architectures, customize optimization strategies, and streamline the training pipeline.

Keras: Keras, an integral part of TensorFlow, provided a high-level API for building and training neural networks. We used it to create model architectures with the help of layers like Input and Dense, which form the building blocks of our Auto-encoder network.

Pandas: Pandas is a powerful data manipulation and analysis library in Python. We utilized Pandas extensively for data loading, cleaning, and transformation. It enabled efficient handling of large datasets and facilitated various data preprocessing tasks.

NumPy: NumPy serves as a fundamental package for scientific computing with Python. It was essential for mathematical operations and array manipulation. In our research, NumPy contributed to tasks involving numerical data processing and manipulation.

Scikit-learn: The Scikit-learn library provides a rich toolkit for machine learning in Python.

Matplotlib: Matplotlib is a widely-used data visualization library in Python. We employed it to generate a diverse range of visual representations, including

line plots, bar charts, and polar plots. Through customization options and a versatile interface, Matplotlib enabled us to effectively showcase patterns, trends, and relationships within our dataset.

4.2 Experimental Setup

The following experimental setup is the same for all the experiments.

As discussed in the previous chapter, an autoencoder has two distinct parts, an encoder and a decoder. The configuration used for the encoder and the decoder is depicted in Figure 3.7. The number of nodes in the input layer of the encoder will be the same as the number of features in the data set (which varies from one experiment to the other as discussed in the next section). The output of the encoder is two dimensional which represents the dimensions of the latent space. The reason behind using two dimensions in the latent space is to be able to visualise these representations in 2-D. As a result the input to the decoder is also two dimensional. The output of the decoder has the same dimension as the input to the encoder. In summary, the encoder transforms the input data to the latent space while the decoder will transform this data from the latent space back to the original as shown in 4.1.

The model was compiled with 'adam' optimizer and Mean Squared Error (MSE) as the loss function. While training an auto encoder, the input and the output are the same and the model is trained to reproduce the original data as well as possible. We used batch size = 1 so that the network can update its weight after each instance. The reason behind this choice is that our input (and output) are multi-dimensional. Computing MSE for a data sample means computing the error for all the features and then taking its mean and optimizing the weights of the network to minimize this error. If we use batch size other than 1, the mean will be computed not only for the set of features but also for the number of instances in the batch and that would affect the learning process by hiding all this information in a single mean value.

```
Model: "model_3"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_3 (InputLayer)        [(None, 14)]              0

 dense_6 (Dense)             (None, 12)                180

 dense_7 (Dense)             (None, 12)                156

 dense_8 (Dense)             (None, 2)                 26

=================================================================
Total params: 362
Trainable params: 362
Non-trainable params: 0
_________________________________________________________________
Model: "model_4"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_4 (InputLayer)        [(None, 2)]               0

 dense_9 (Dense)             (None, 12)                36

 dense_10 (Dense)            (None, 12)                156

 dense_11 (Dense)            (None, 14)                182

=================================================================
Total params: 374
Trainable params: 374
Non-trainable params: 0
```

FIGURE 4.1: Encoder and Decoder Architecture

The auto encoder was trained on the two data sets using different sets of features as discussed in the next section.

The ADL data used for the training of the autoencoder however is only of healthy people (not suffering from dementia), as discussed previously. The idea is that by training the model only on the data of healthy people, it will "learn" the representation of the "normal" behavior. Then when it is presented with data that includes signs of dementia (and hence is different from normal behavior), it will not be able to reconstruct it correctly.

After training the autoencoder, we reconstruct the original data (of healthy individuals) using the trained model. The accuracy of the model on this data shows how well it has learned the representation of normal data. For all the experiments this value is above 90%. Along with finding the accuracy, we also calculate the reconstruction error as MSE. Based on the MSE distribution, we set a threshold for anomaly detection as the mean plus two times the standard deviation of the MSE.

A data sample that has MSE above this threshold will be considered an anomaly. When applying the model on the data of dementia sufferers, an accuracy of close to 0% is desired which indicates that the model was unable to reconstruct the data correctly.

4.3 Abnormal Test Cases

As discussed previously, there is no publicly available ADL dataset of dementia sufferers. Hence, specific test cases were constructed, each representing typical behavior of a dementia sufferer, guided by extensive literature in the field of medicine related to dementia and its symptoms [20], [39], [40], [41]. These test cases will be used to evaluate the proposed approach's effectiveness in detecting dementia-related anomalies. Table 4.1 illustrates the test cases.

4.3.1 Test Case 1:

Dementia patients often exhibit repetitive behaviors as their cognitive functions decline. In this case, the individual tends to perform toileting activities more frequently than usual, as their brain may struggle to communicate properly with the urinary system, leading to more frequent trips to the toilet even when not needed [42]. Additionally, disturbances in sleeping patterns are observed, leading to repeated instances of sleeping activity.

Moreover, meal preparation and eating activities are performed repeatedly as the person's memory deteriorates. They may forget that they have already completed these tasks and repeat them, indicating memory-related challenges.

The day routine for this case is shown in Table 4.2. The activities are denoted by codes: A1 (Meal Preparation), A2 (Relax), A3 (Eating), A4 (Work), A5 (Sleeping), A6 (Wash Dishes), A7 (Bed to Toilet), A8 (Enter Home), A9 (Leave Home), and A10 (Housekeeping).

DF_MP	DF_Relax	DF_Eating	DF_work	DF_Sleeping	DF_Wash_Dishes	DF_BTT	DF_EH	DF_LH	DF_Housekeeping	Repetitive_Pattern_count	RepetitiveP_Cumulative_L	Repetitive_LC2	Repetitive_LC3
0:56:50	3:00:13	0:52:59	0:18:29	1:53:00	0:02:33	0:05:07	0:00:00	0:00:00	0:00:00	3	12	1	1
0:27:49	1:51:02	0:20:26	0:00:00	4:02:10	0:43:52	0:05:01	0:00:21	0:00:14	0:00:00	2	9	0	1
0:12:41	1:21:12	0:07:18	0:00:00	2:11:57	0:14:00	0:01:30	0:00:00	0:00:00	1:10:01	2	10	0	0
0:46:50	1:07:37	0:10:00	0:00:00	4:47:31	0:42:33	0:10:00	0:00:11	0:00:17	0:40:00	3	9	1	1
0:00:00	0:00:00	0:00:53	1:02:09	6:40:09	0:17:16	0:05:40	0:00:00	0:00:00	0:22:30	3	10	0	2

D_Sleeping	D_Toileting	D_Breakfast	D_Spare_Time	D_Leaving	D_Lunch	D_Snack...	FC_Sleeping	FC_Toileting	FC_Showering	FC_Breakfast..	Repetitive_Pattern_count	RepetitiveP_Cumulative_L	Repetitive_LC2.
7:54:20	0:35:39	1:10:27	2:04:00	0:11:12	1:10:27	0:20:00	5	7	4	6	3	12	1
15:20:18	0:02:41	0:28:30	0:09:45	0:00:00	0:28:30	1:00:00	4	9	2	4	4	16	0
4:30:15	0:43:09	2:10:55	5:29:38	2:00:00	2:10:55	0:56:59	3	2	3	5	3	15	0
13:11:01	0:15:48	0:17:53	8:00:55	1:10:00	0:17:53	0:05:31	6	10	4	1	4	13	1
2:58:52	0:03:04	0:53:44	6:08:39	2:30:00	0:53:44	1:08:23	4	12	0	7	3	10	0

TABLE 4.1: Dementia Test Cases

TABLE 4.2: Test Case1

Routine 1: A7, A5, A7, A5, A7, A5, A7, A5, A7, A5, A7, A5, A7, A5, A1, A3, A6, A2, A4, A2, A1, A3, A1, A3, A1, A3, A2, A1, A3, A1, A3, A2.

The day routine shows that the person performs "Bed To Toilet" and "Sleeping" activities (A7 and A5) repeatedly, occurring seven times in sequence.

Similarly, "Meal Preparation" and "Eating" activities (A1 and A3) are performed three times consecutively. This pattern of repetitive activities is common among dementia patients.

4.3.2 Test Case 2:

In Test Case 2, we explore another scenario depicting abnormal behavior in dementia patients. In this case, the individual's day routine showcases a pattern of repeated activities, with an emphasis on meal preparation and relaxation as well as eating and dish_washing activities. Such behavior is common among dementia patients due to the cognitive challenges they face [20].

Dementia patients often struggle with memory and sequencing tasks, leading to the repetition of daily activities. In this instance, the individual may not remember completing certain tasks and perform them again, causing repetitive patterns to emerge.

The day Routine for case 2 is in Table 4.3.

TABLE 4.3: Test Case2

Routine 2: A5, A2, A1, A2, A1, A2, A1, A2, A1, A2, A1, A2, A1, A3, A5, A2, A3, A6, A7, A9, A8, A3, A6, A3, A6, A3, A6, A2.

4.3.3 Test Case 3:

In Test Case 3, we explore another scenario representing abnormal behavior in dementia patients. The day routine for this case is shown in table 4.4. The day routine indicates that the person repeatedly performs "Meal Preparation" and "Eating" activities (A1 and A3) five times consecutively. Similarly, "Wash Dishes" and "Relax" activities (A6 and A2) are repeated three times in sequence [17].

TABLE 4.4: Test Case3

Routine 3: A5, A7, A1, A2, A10, A1, A3, A1, A3, A1, A3, A1, A3, A1, A3, A6, A2, A6, A2, A6, A2, A5, A7, A2, A5.

4.3.4 Test Case 4:

In Test Case 4, we explore another scenario illustrating abnormal behavior in dementia patients. The day routine for this case is shown in table 4.5. The day routine reveals that the person repeatedly performs "Enter Home" and "Leave Home" activities (A8 and A9) three times consecutively. Additionally, "Meal Preparation" and "Relax" activities (A1 and A2) are performed twice in sequence. Moreover, "Sleeping" and "Relax" activities (A5 and A2) are repeated four times in sequence [43].

TABLE 4.5: Test Case4

Routine 4: A5, A7, A5, A1, A2, A1, A3, A6, A10, A9, A8, A9, A8, A9, A8, A1, A2, A1, A2, A7, A2, A7, A2, A7, A2, A7, A2.

4.3.5 Test Case 5:

In Test Case 5, we explore another scenario exemplifying abnormal behavior in dementia patients. The day routine for this case is in table 4.6. The day routine

shows that the person repeatedly performs "Wash Dishes" and "Housekeeping" activities (A6 and A10) four times consecutively. Additionally, "Toileting" and "Eating" activities (A7 and A3) are performed twice in sequence. Moreover, there is a repetitive pattern involving "Toileting" and "Wash Dishes" activities (A7 and A6), occurring four times in sequence [37].

TABLE 4.6: Test Case5

Routine 5: A7, A5, A3, A2, A6, A10, A6, A10, A6, A10, A6, A10, A7, A3, A7, A3, A7, A3, A7, A6, A7, A6, A7, A6, A4, A5.

A significant finding across all cases is the correlation between increased frequency of activities and their impact on the duration of those activities. Dementia patients tend to perform certain tasks more frequently than usual, affecting the time spent on each activity. Moreover, the repetition of activities, while skipping others, contributes to the formation of recognizable patterns in their behavior. The key features that are Frequency Count, Duration, and Repetitive Patterns Count, along with the length of these patterns have proven highly relevant in detecting these repetitive anomalies.

4.4 Experiment 1 for Aruba Dataset

In experiment 1, the final feature set shown in Table 3.6 was utilized. All the features were normalized using the min-max normalization method 4.1 for both normal and abnormal data instances. It scales the data to a common range, typically [0, 1]. Normalizing features to the same scale is essential to prevent some features from dominating others during modeling, especially for algorithms like neural networks and k-means clustering that rely on distance measures or gradient-based optimization. The formula for Min-Max normalization is as follows:

$$X_{\text{normalized}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \tag{4.1}$$

Where: X_normalized is the normalized value of the feature. X is the original value of the feature. X_min is the minimum value of the feature in the dataset. X_max is the maximum value of the feature in the dataset.

The Architecture of autoencoder that used for experiment 1 is shown in Figure 4.1. After training the autoencoder with 200 epochs and a batch size of 1, we passed both the original and abnormal data through the encoder. This process generates embedding for both types of data samples. To visually inspect the distribution of normal and abnormal data in a 2D plot, we create scatter plots shown in Figure 4.2. The embedding for normal data are represented in black color, while the embedding for abnormal data are shown in red color. This visualization helps us understand the clustering and separation of normal and abnormal data points in the reduced-dimensional space. Additionally, multiple clusters are visible in the plot for normal data, depicts the presence of various distinct patterns of normal behavior.

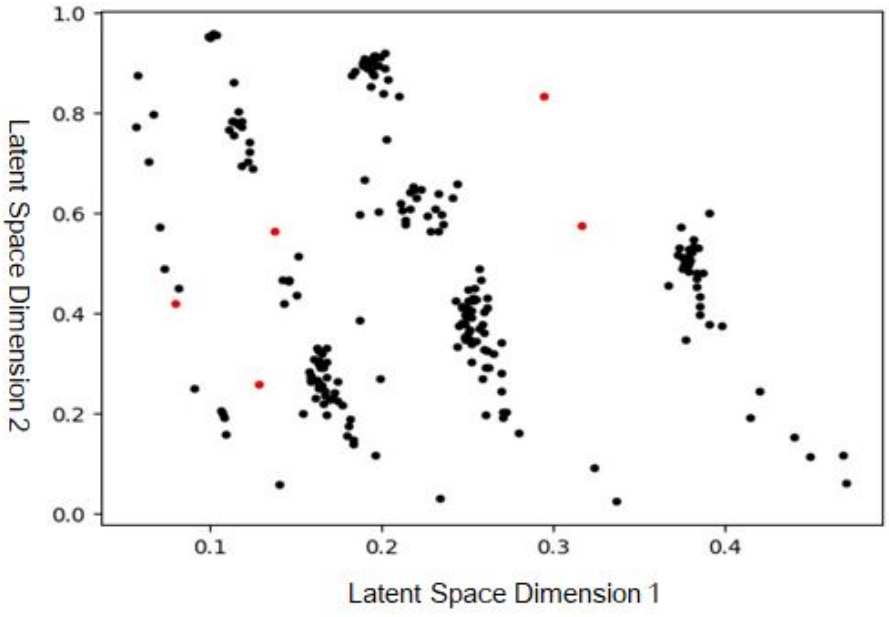

FIGURE 4.2: Two-Dimensional Plot of Embedding for Normal and Abnormal Data Experiment 1 for Aruba

Next, the data is reconstructed from the embedding using the trained Autoencoder's decoder part, achieving an accuracy of 96.82% when reconstructing the same input data on which the autoencoder was trained.

The reconstruction error (MSE) for all features are shown in Figure 4.3.

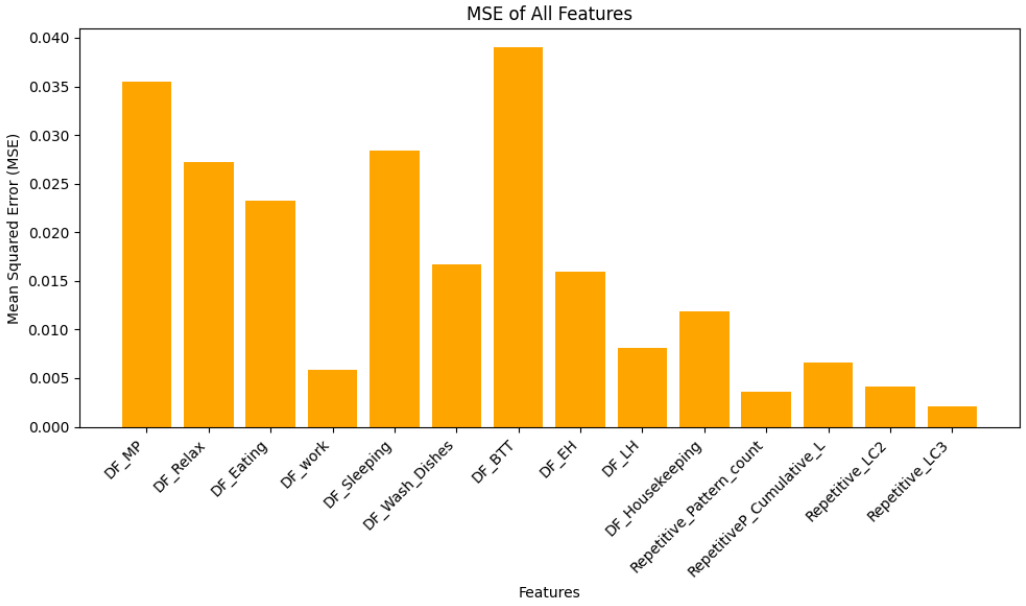

FIGURE 4.3: The reconstruction error (MSE) for all features of Experiment 1 for Aruba

The reconstruction error (Mean Squared Error or MSE) provides insights into how effectively the autoencoder model captures and reproduces different data patterns. Low Mean Squared Error (MSE) values in the reconstruction of features associated with repetitive patterns, such as 'Repetitive_Pattern_count,' 'RepetitiveP_Cumulative_L,' 'Repetitive_LC2,' and 'Repetitive_LC3,' indicate that the autoencoder model effectively captures and reproduces these repetitive data patterns. While features associated with the average duration spent on each activity, such as 'DF_MP,' 'DF_Relax,' 'DF_Eating,' 'DF_Sleeping,' and 'DF_BTT,' display higher MSE values in the reconstruction process. This means that the autoencoder encounters difficulties when encoding and reproducing these features, resulting in higher reconstruction errors.

For enhanced visualization and a deeper understanding of the autoencoder's reconstruction of the original input data, line charts were created to map the input data (original data) onto the reconstructed data. Figures 4.4, 4.5, and 4.6 demonstrate that features with low Mean Squared Error (MSE) values are accurately reconstructed, while those with high MSE values exhibit poorer reconstruction.

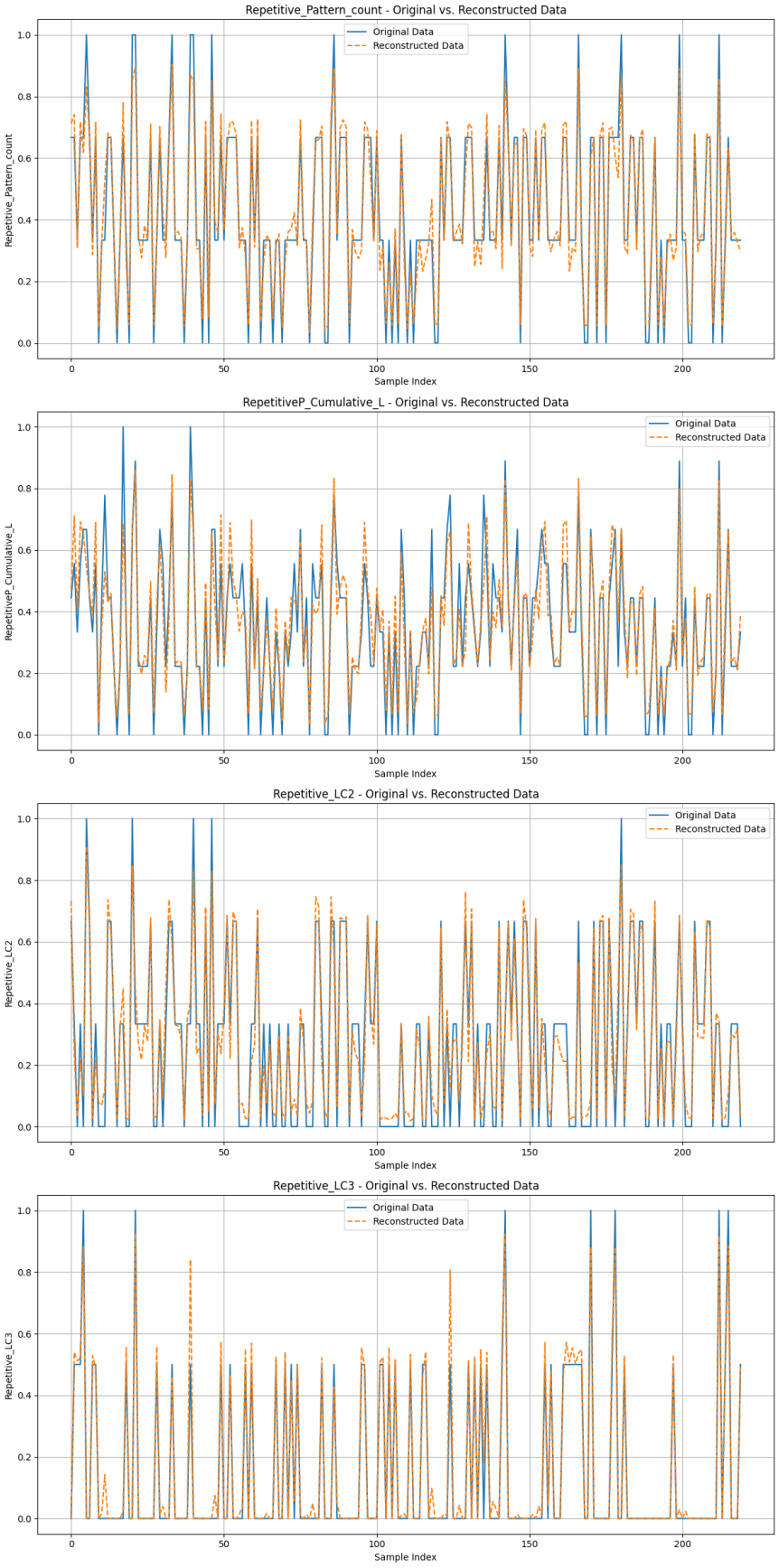

FIGURE 4.4: Repetitive Pattern Feature Mapping: Original vs. Reconstructed Data Experiment 1 for Aruba

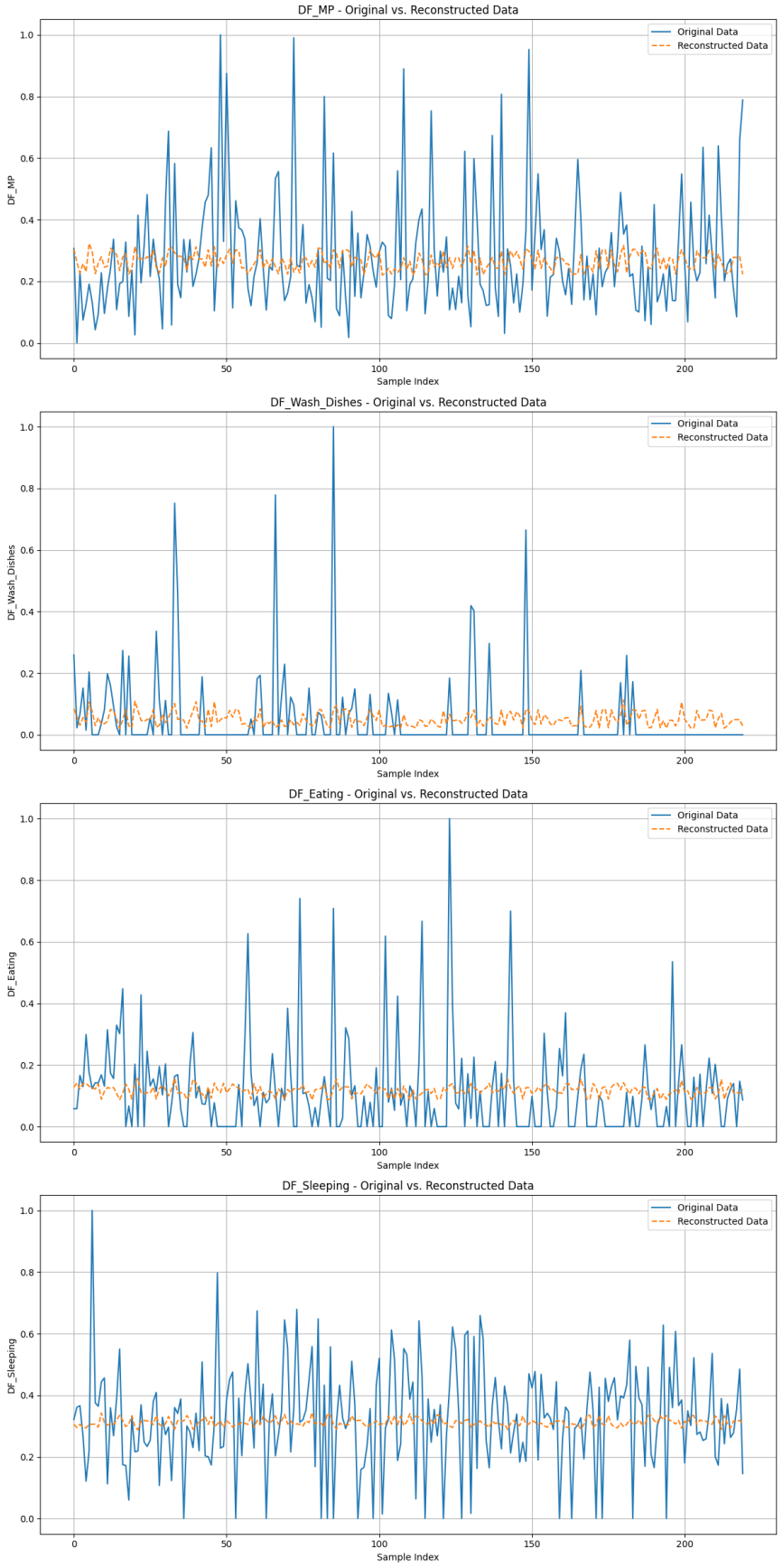

FIGURE 4.5: Average Duration spent on each activity Features: Original vs. Reconstructed Data (1-5) Experiment 1 for Aruba

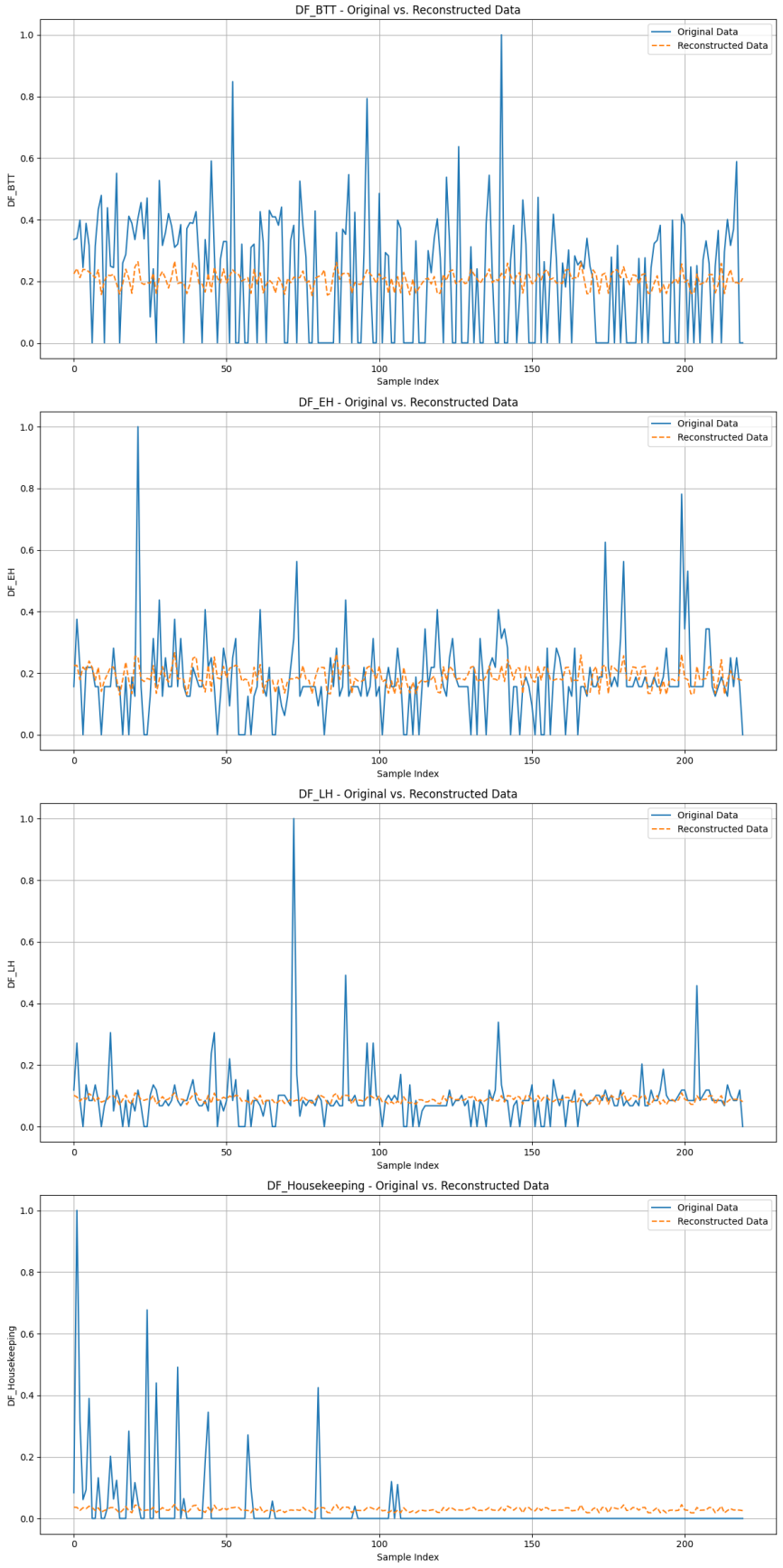

FIGURE 4.6: Average Duration spent on each activity Features: Original vs. Reconstructed Data (6 to 10) Experiment 1 for Aruba

As observed, certain features have been successfully reconstructed, while others exhibit less accurate reconstruction.

In the next step, the trained autoencoder, which has been trained on normal data, is applied to reconstruct abnormal data. This step allows us to assess whether the autoencoder makes mistakes while reconstructing abnormal instances. If mistakes are observed and accuracy drops, it means that the autoencoder can differentiate between normal and abnormal data. All five abnormal cases exhibiting dementia patient behaviors were reconstructed, and all of these cases were successfully detected, resulting in a 0% accuracy.

Figure 4.7 displays data points representing both normal and abnormal instances. Black points denote normal data clusters, blue points represent instances successfully detected as abnormal by the autoencoder, and orange points indicate instances where the autoencoder made mistakes.

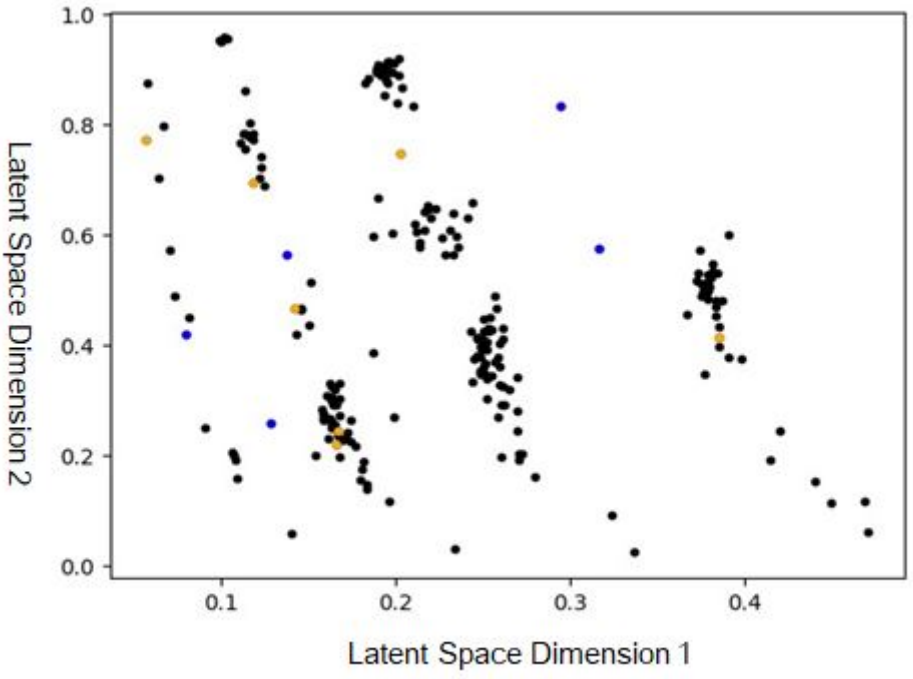

FIGURE 4.7: Embedding Normal and Abnormal data Experiment 1 for Aruba

For this experiment we know that the reconstruction is poor for some DF features. hence the next experiment is to investigate this issue by separating Duration and frequency features.

4.5 Experiment 2 for Aruba Dataset

In Experiment 2, the first transformation feature set, as depicted in Tables 3.4 and 3.5, was utilized as discussed in Chapter 3. All features were normalized in the same manner as in Experiment 1, using the min-max normalization method for both normal and abnormal data instances.

The Architecture of autoencoder that used for experiment 2 is shown in Figure 4.8.

```
Model: "model_24"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_17 (InputLayer)       [(None, 24)]              0

 dense_48 (Dense)            (None, 12)                300

 dense_49 (Dense)            (None, 12)                156

 dense_50 (Dense)            (None, 2)                 26

=================================================================
Total params: 482
Trainable params: 482
Non-trainable params: 0
_________________________________________________________________
Model: "model_25"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_18 (InputLayer)       [(None, 2)]               0

 dense_51 (Dense)            (None, 12)                36

 dense_52 (Dense)            (None, 12)                156

 dense_53 (Dense)            (None, 24)                312

=================================================================
Total params: 504
Trainable params: 504
Non-trainable params: 0
```

FIGURE 4.8: Architecture of autoencoder for Experiment 2 for Aruba

The same batch size and number of epochs from Experiment 1 were used for training the autoencoder. Subsequently, both the original and abnormal data were passed through the encoder. This process generates embedding for both types of data samples. For visual inspection of the distribution of normal and abnormal data in a 2D plot, scatter plots are generated, as shown in Figure 4.9.

The embeddings for normal data are represented in black, while the embeddings for abnormal data are shown in red. This visualization helps us understand the clustering and separation of normal and abnormal data points in the reduced-dimensional space.

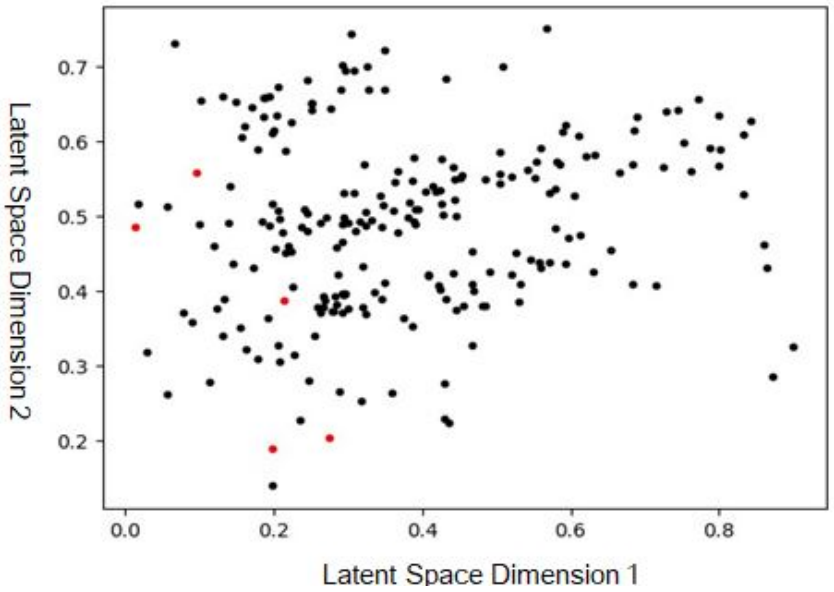

FIGURE 4.9: Two-Dimensional Plot of Embedding for Normal and Abnormal Data for Experiment 2 for Aruba

Next, the data is reconstructed from the embedding using the trained Autoencoder's decoder part, achieving an accuracy of 94.09% when reconstructing the same input data on which the autoencoder was trained.

The reconstruction error (MSE) for all features are shown in Figure 4.10. The reconstruction error (Mean Squared Error or MSE) provides insights into how effectively the autoencoder model captures and reproduces different data patterns. As we can see in Figure 4.10, only a few features have a high MSE value, such as D_Sleeping, FC_Work, FC_Wash_Dishes, and Repetitive_LC3.

However, the overall reconstruction error for experiment 2 is relatively lower than that for experiment 1 features. This indicates that the autoencoder more accurately reconstructs the features for Experiment 2 than for Experiment 1. For enhanced visualization and a deeper insight into the quality of the autoencoder's reconstruction, line charts for Experiment 2 are generated as well.

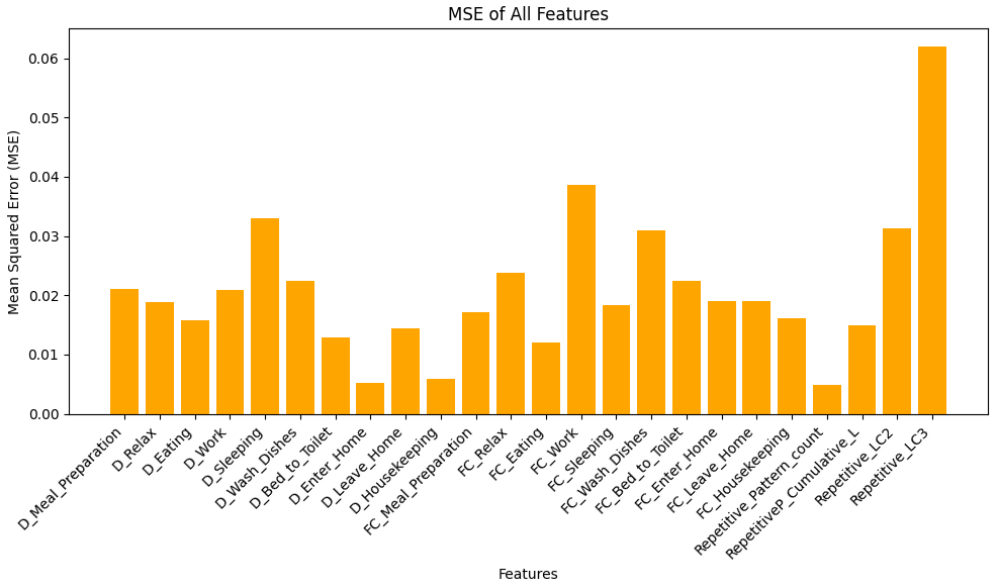

FIGURE 4.10: The reconstruction error (MSE) for all features of Experiment 2 for Aruba

These charts depict the mapping of the input data (original data) onto the reconstructed data, for Experiment 2 features, which include Duration, Frequency count (separately for each activity), and Repetitive patterns. Figures 4.11, 4.12, and 4.13, 4.14, 4.15 illustrate that features with low MSE values are accurately reconstructed, while those with high MSE values show poorer reconstruction.

As observed, most of the features have been successfully reconstructed, while some exhibit less accurate reconstruction.

In the next step, the trained autoencoder, specifically trained on normal data for the Experiment 2 feature set, was applied to reconstruct abnormal data.

This step allowed to assess whether the autoencoder makes mistakes while reconstructing abnormal instances. Five different abnormal cases exhibiting dementia patient behaviors were created using the Experiment 2 feature set and reconstructed using the trained autoencoder. As expected, the autoencoder made mistakes with the abnormal cases, resulting in all of these cases being successfully detected, which led to a 0% accuracy. Figure 4.16 displays data points representing both normal and abnormal instances.

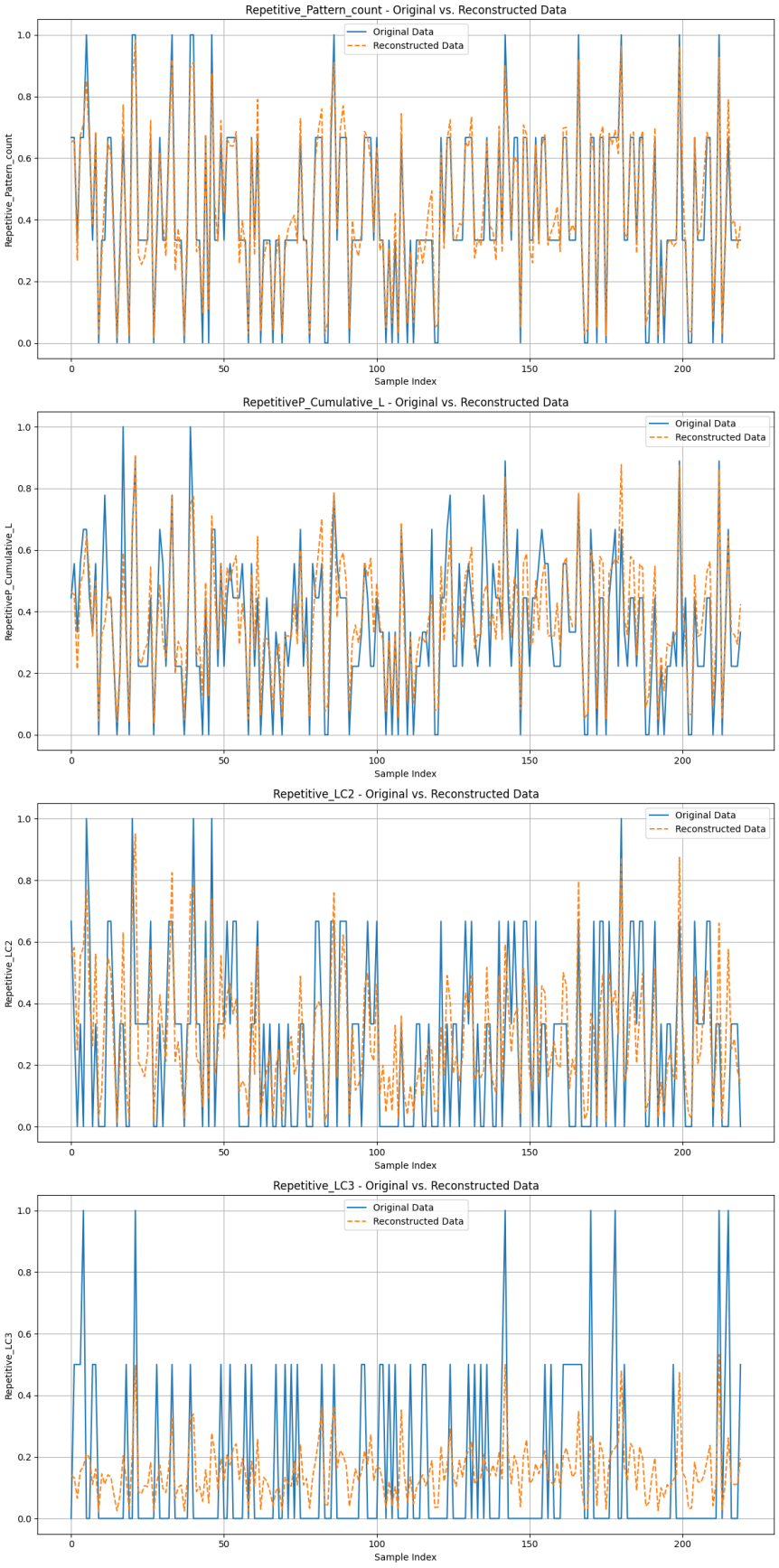

FIGURE 4.11: Repetitive Pattern Feature Mapping (1-4): Original vs. Reconstructed Data for Experiment 2 for Aruba

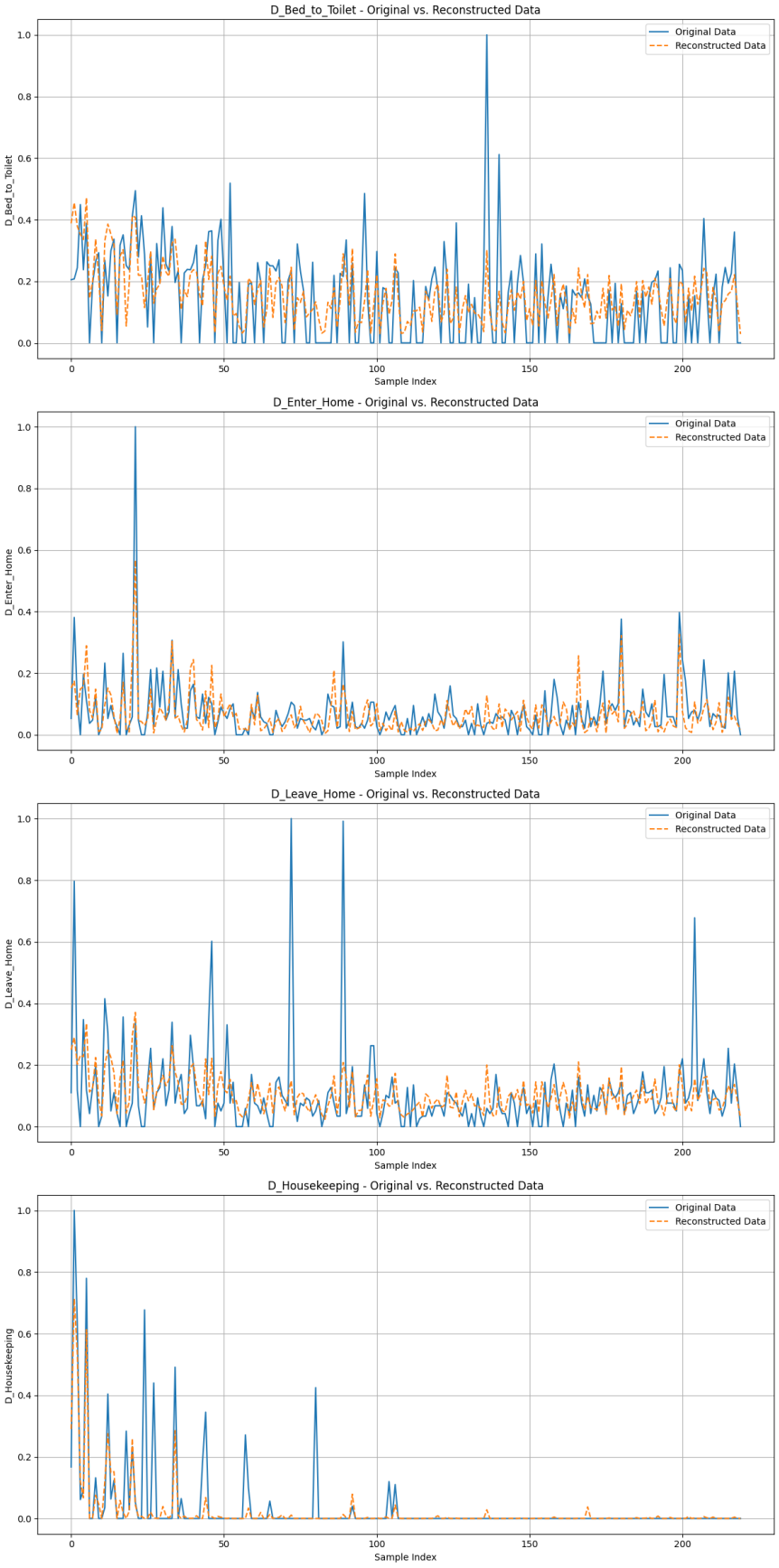

FIGURE 4.12: Duration spent on an activity Features: Original vs. Reconstructed Data (5 to 8) Experiment 2 for Aruba

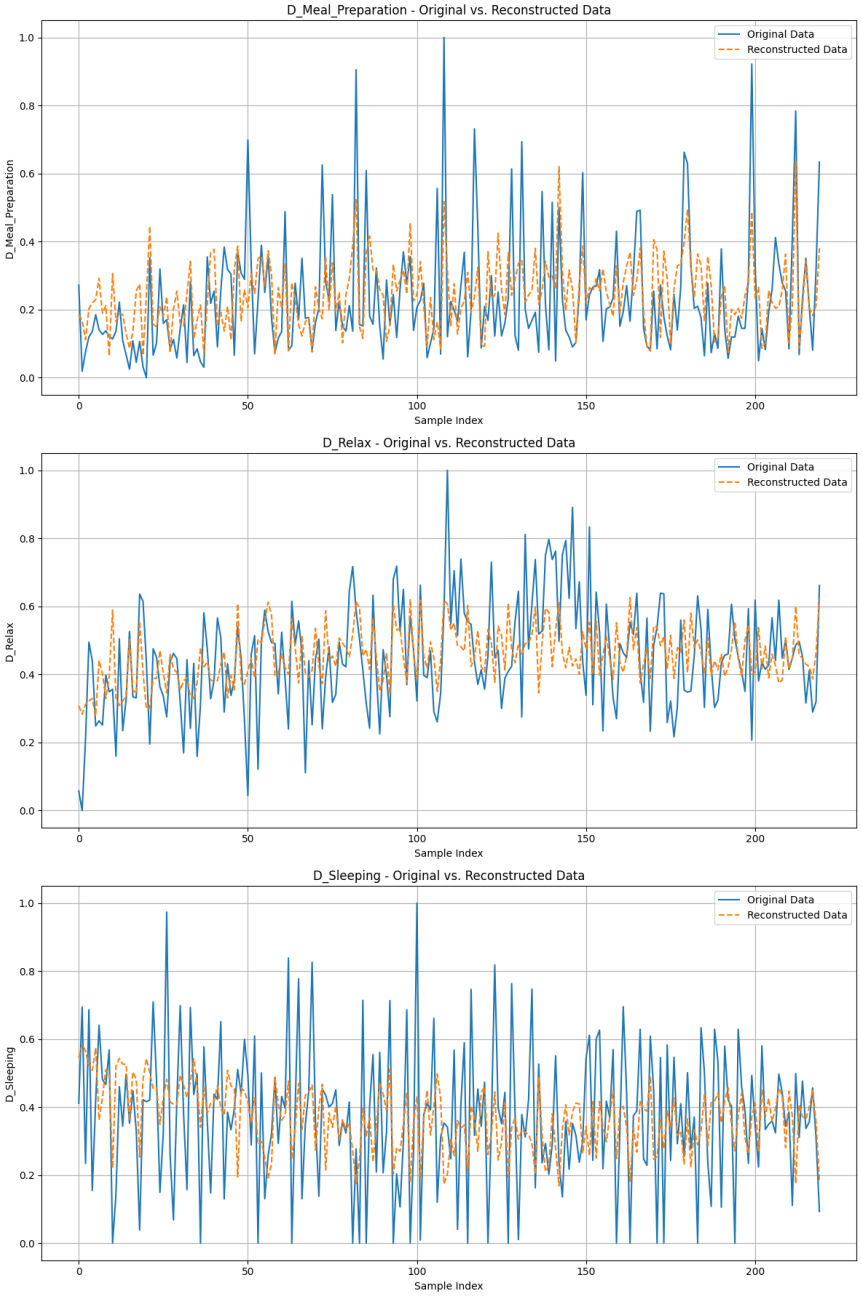

FIGURE 4.13: Duration spent on an activity Features: Original vs. Reconstructed Data (8-12) Experiment 2 for Aruba

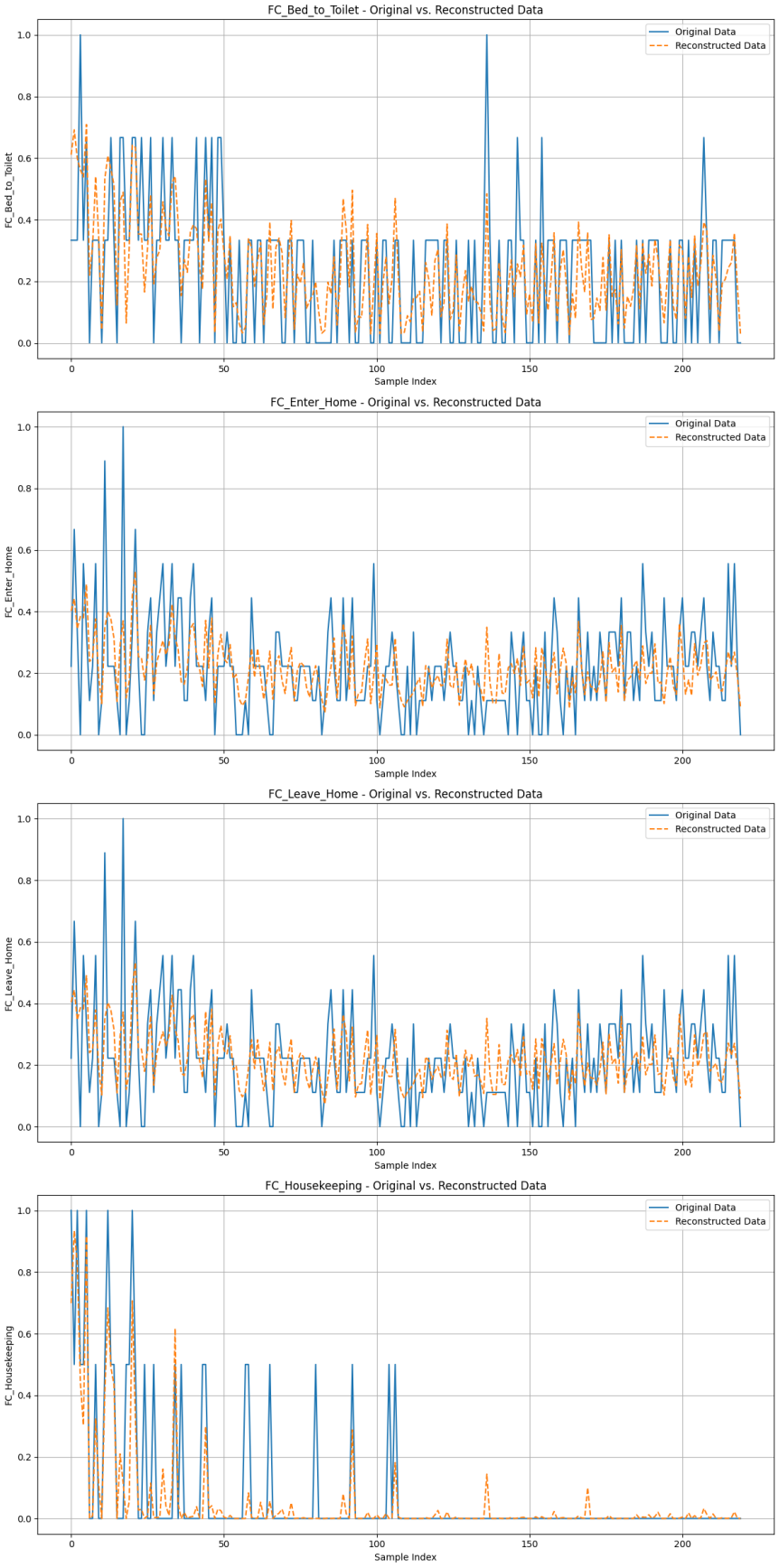

FIGURE 4.14: Frequency Count Features: Original vs. Reconstructed Data (13-16) Experiment 2 for Aruba

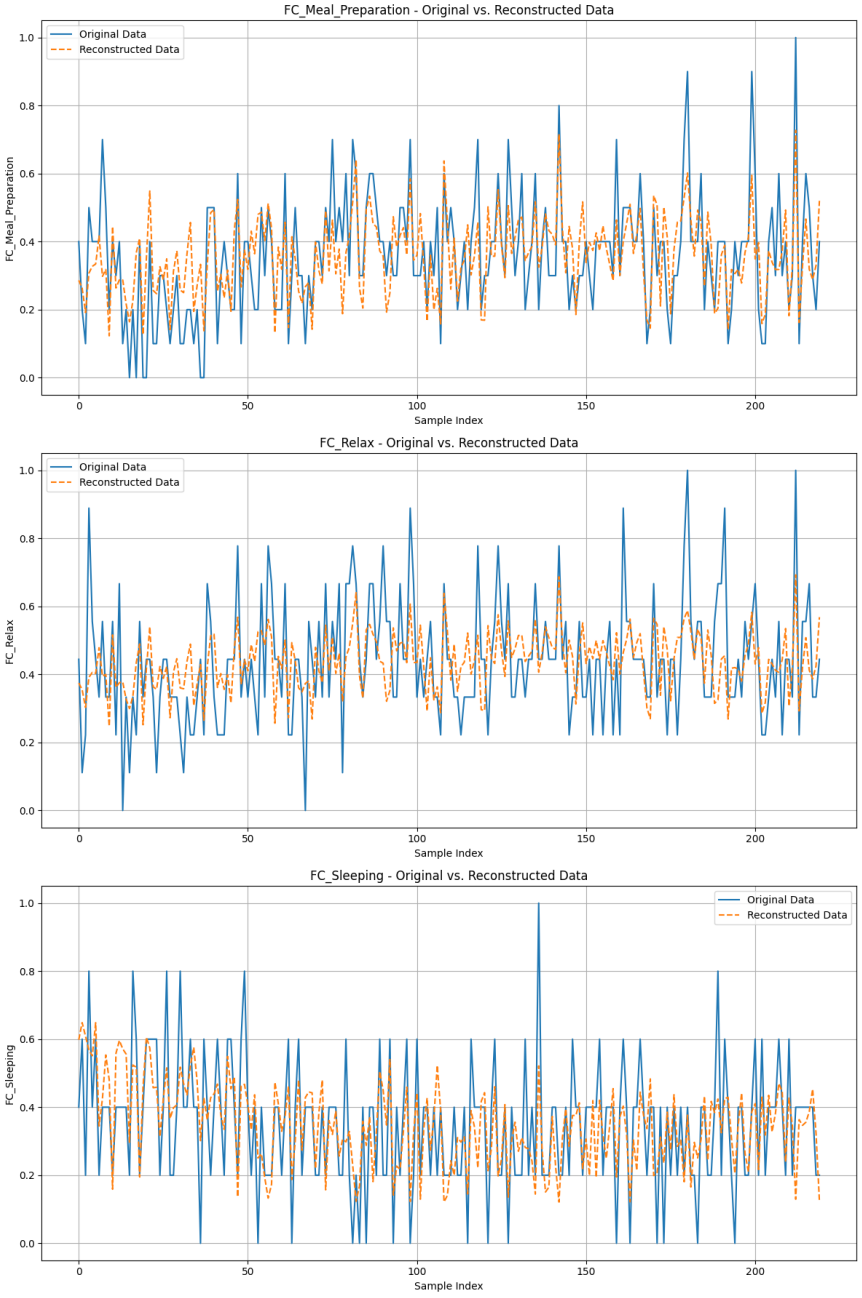

FIGURE 4.15: Frequency Count Features: Original vs. Reconstructed Data (16-20) Experiment 2 for Aruba

Black points denote normal data clusters, blue points represent instances successfully detected as abnormal by the autoencoder, and orange points indicate instances where the autoencoder made mistakes.

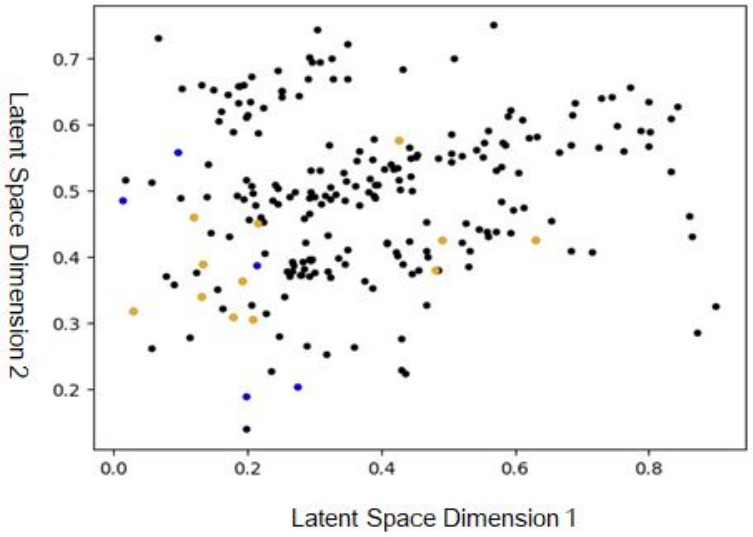

FIGURE 4.16: Embedding Normal and Abnormal data Experiment 2 for Aruba

4.6 Experiment 1 for Ordonez's Dataset

In Experiment 1 with Ordonez's dataset, the same approach was applied as in Experiment 1 with Aruba's dataset.
Specifically, combining the frequency and duration features was achieved by calculating their averages, as illustrated in Table 4.7.

All features were then normalized using the same method applied in Experiment 1 with the Aruba dataset. The architecture of the autoencoder used for this experiment is depicted in Figure 4.17.

After training the autoencoder, both the original and abnormal data are passed through the encoder.
This process generates embedding for both types of data samples. To visually inspect the distribution of normal and abnormal data in a 2D plot, scatter plots are created as shown in Figure 4.18. The embedding for normal data are represented in black color, while the embedding for abnormal data are shown in red color.

DF_Sleeping	DF_Toileting	DF_Showering	DF_Breakfast	DF_Grooming	DF_Spare_Time/TV	DF_Leaving	DF_Lunch	DF_Snack	Repetitive_Pattern_count	RepetitiveP_Cumulative_L	Repetitive_LC2	Repetitive_LC3
7:51:00	0:02:20	0:08:00	0:09:00	0:02:00	3:28:45	0:20:00	0:37:00	0:00:00	0	0	0	0
9:15:00	0:04:45	0:02:00	0:10:00	0:05:00	1:28:50	1:21:00	0:31:00	0:00:00	1	2	1	0
8:45:00	0:00:30	0:04:00	0:13:00	0:01:00	1:32:07	0:19:00	0:21:00	0:00:00	1	3	0	1
9:13:00	0:01:00	0:14:00	0:12:00	0:03:20	1:46:00	1:20:00	0:26:00	0:00:00	0	0	0	0
10:00:00	0:10:20	0:04:00	0:10:00	0:01:15	1:38:17	0:00:00	0:33:00	0:00:00	0	0	0	0
10:09:00	0:01:00	0:04:00	0:09:00	0:01:40	2:30:45	3:30:00	0:00:00	0:00:00	1	2	1	0
10:18:00	0:02:45	0:04:00	0:03:00	0:00:30	1:33:20	2:20:00	0:00:00	0:00:00	0	0	0	0
10:13:00	0:06:00	0:07:00	0:10:00	0:00:20	0:52:43	0:58:30	0:43:00	0:00:00	1	2	1	0
9:25:00	0:02:20	0:13:00	0:05:00	0:01:15	2:15:24	0:00:00	0:35:00	0:00:00	0	0	0	0
9:46:00	0:00:30	0:06:00	0:08:00	0:04:00	1:17:37	1:47:00	0:36:00	0:03:00	2	2	0	0
9:35:00	0:00:00	0:05:00	0:09:00	0:00:12	2:40:15	2:11:00	0:00:00	0:00:00	1	3	0	1
9:33:00	0:02:00	0:04:00	0:08:00	0:02:00	1:17:20	2:02:30	0:00:00	0:00:00	0	0	0	0
8:36:00	0:07:00	0:05:00	0:07:00	0:03:00	2:01:40	4:17:00	0:00:00	0:00:00	0	0	0	0
9:26:00	0:01:30	0:16:00	0:09:00	0:01:00	3:05:40	0:00:00	0:52:00	0:00:00	0	0	0	0
7:37:00	0:00:45	0:08:00	0:10:00	0:01:55	0:37:51	0:46:00	1:06:00	0:10:30	3	7	1	0
7:22:00	0:00:45	00:00:00	0:08:00	0:04:00	1:16:40	1:33:30	0:34:00	0:17:45	2	3	1	0
4:14:00	0:00:20	0:02:00	0:17:00	0:02:20	1:03:45	0:55:40	0:00:00	0:03:00	3	4	1	0
2:01:00	0:00:00	00:00:00	0:10:00	0:01:40	1:23:00	2:48:00	1:02:00	0:10:00	2	3	1	0
9:39:00	0:00:40	0:03:00	0:13:00	0:03:24	2:09:00	2:30:20	0:36:00	0:02:00	1	1	0	0
7:12:00	0:00:30	0:02:00	0:06:00	0:03:45	1:49:40	2:07:00	0:00:00	0:05:00	3	3	0	0
7:18:00	0:01:00	0:07:00	0:07:00	0:05:48	0:52:24	2:06:00	0:07:00	0:04:00	1	1	0	0
9:56:00	0:00:24	0:07:00	0:12:00	0:03:10	2:13:24	1:21:00	0:00:00	0:06:30	2	3	1	0

TABLE 4.7: Duration Frequency combined Features for Ordonez's Experiment1

```
Model: "model_42"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_29 (InputLayer)       [(None, 13)]              0

 dense_84 (Dense)            (None, 12)                168

 dense_85 (Dense)            (None, 12)                156

 dense_86 (Dense)            (None, 2)                 26

=================================================================
Total params: 350
Trainable params: 350
Non-trainable params: 0
_________________________________________________________________
Model: "model_43"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_30 (InputLayer)       [(None, 2)]               0

 dense_87 (Dense)            (None, 12)                36

 dense_88 (Dense)            (None, 12)                156

 dense_89 (Dense)            (None, 13)                169

=================================================================
Total params: 361
Trainable params: 361
Non-trainable params: 0
```

FIGURE 4.17: Architecture of autoencoder for Ordonez' Experiment 1

This visualization helps us understand the clustering and separation of normal and abnormal data points in the reduced-dimensional space.

Next, the data is reconstructed from the embedding using the trained Autoencoder's decoder part, achieving an accuracy of 94.29% when reconstructing the same input data on which the autoencoder was trained.

The reconstruction error (MSE) for all features are shown in Figure 4.19.

The reconstruction error (Mean Squared Error or MSE) provides insights into how effectively the autoencoder model captures and reproduces different data patterns.

Some features exhibit low Mean Squared Error (MSE) values, while others, such as 'DF_Leaving,' 'DF_Lunch,' 'Repetitive_LC3,' 'DF_Snack,' and 'DF_Showering,' display higher MSE values in the reconstruction process.This means that the autoencoder encounters difficulties when encoding and reproducing these features, resulting in higher reconstruction errors.

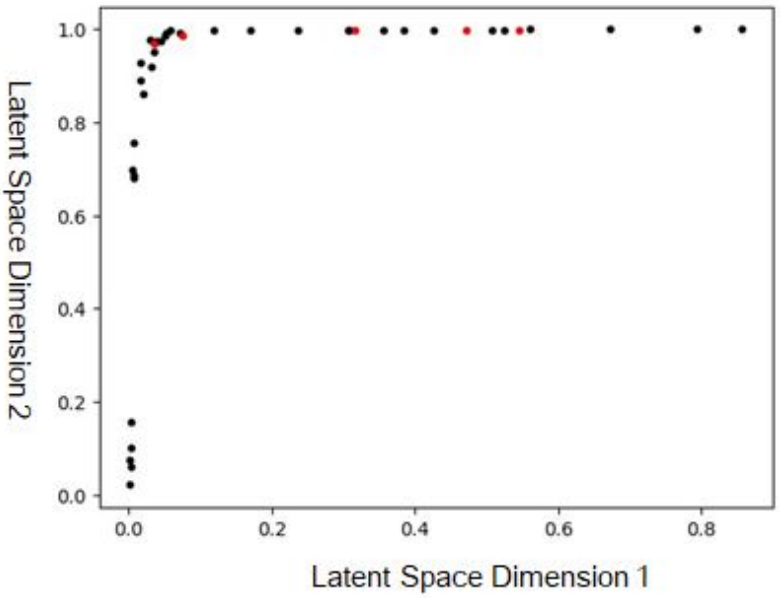

FIGURE 4.18: Two-Dimensional Plot of Embedding for Normal and Abnormal Data for Ordonez's Experiment 1

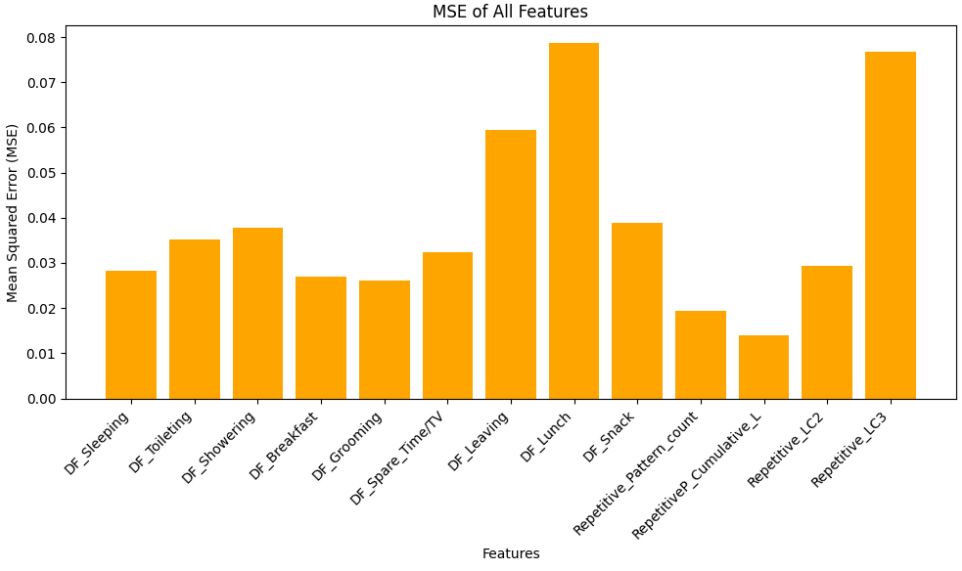

FIGURE 4.19: The reconstruction error (MSE) for all features of Experiment 1 for Ordonez's dataset

For enhanced visualization and a better understanding of how well the autoencoder has reconstructed the original input data, line charts are created to map the input data (original data) onto the reconstructed data. Figures 4.20, 4.21, and 4.22 illustrate that features with low MSE values are accurately reconstructed, while those with high MSE values show poorer reconstruction.

As observed, certain features have been successfully reconstructed, while others exhibit less accurate reconstruction. In the next step, we apply the trained autoencoder, which has been trained on normal data, to reconstruct abnormal data. This step allows us to assess whether the autoencoder makes mistakes while reconstructing abnormal instances. If mistakes are observed and accuracy drops, it means that the autoencoder can differentiate between normal and abnormal data. We tried to reconstruct all five abnormal cases that exhibited dementia patient behaviors, and all of these cases were successfully detected, resulting in a 0% accuracy. Figure 4.23 displays data points representing both normal and abnormal instances. Black points denote normal data instances, blue points represent instances successfully detected as abnormal by the autoencoder, and orange points indicate instances where the autoencoder made mistakes.

It was observed that certain DF features didn't reconstruct well in this experiment. Consequently, another experiment will be conducted to investigate the issue, focusing specifically on examining Duration and Frequency features separately.

4.7 Experiment 2 for Ordonez's Dataset

In Experiment 2, separate frequency and duration features, along with repetitive patterns features, were used as depicted in Table 4.8.

All features were normalized using the same method applied in Experiment 1, utilizing the min-max normalization method for both normal and abnormal data instances. The architecture of the autoencoder used for Experiment 2 is illustrated in Figure 4.24.

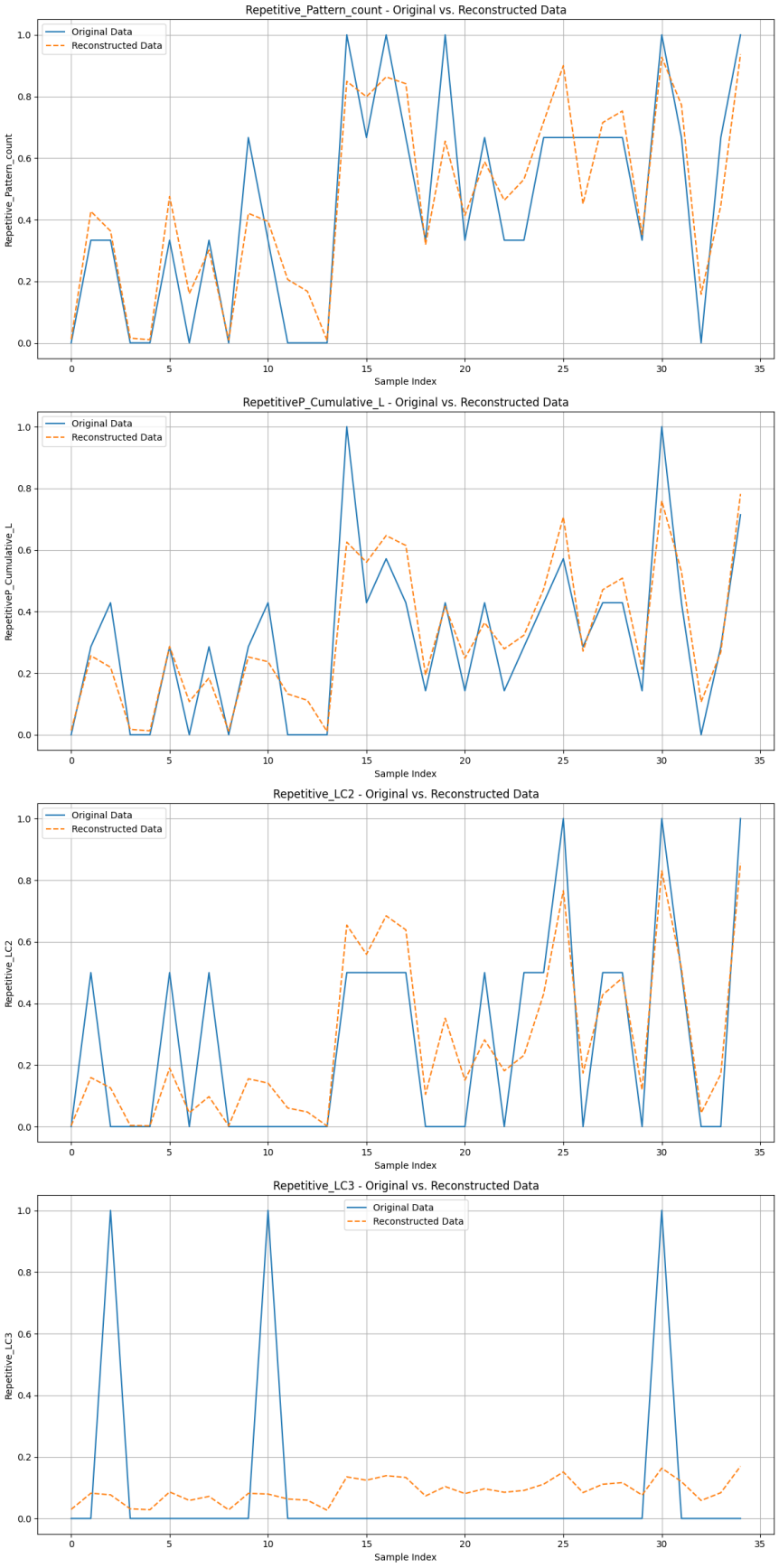

FIGURE 4.20: Repetitive Pattern Feature Mapping: Original vs. Reconstructed Data

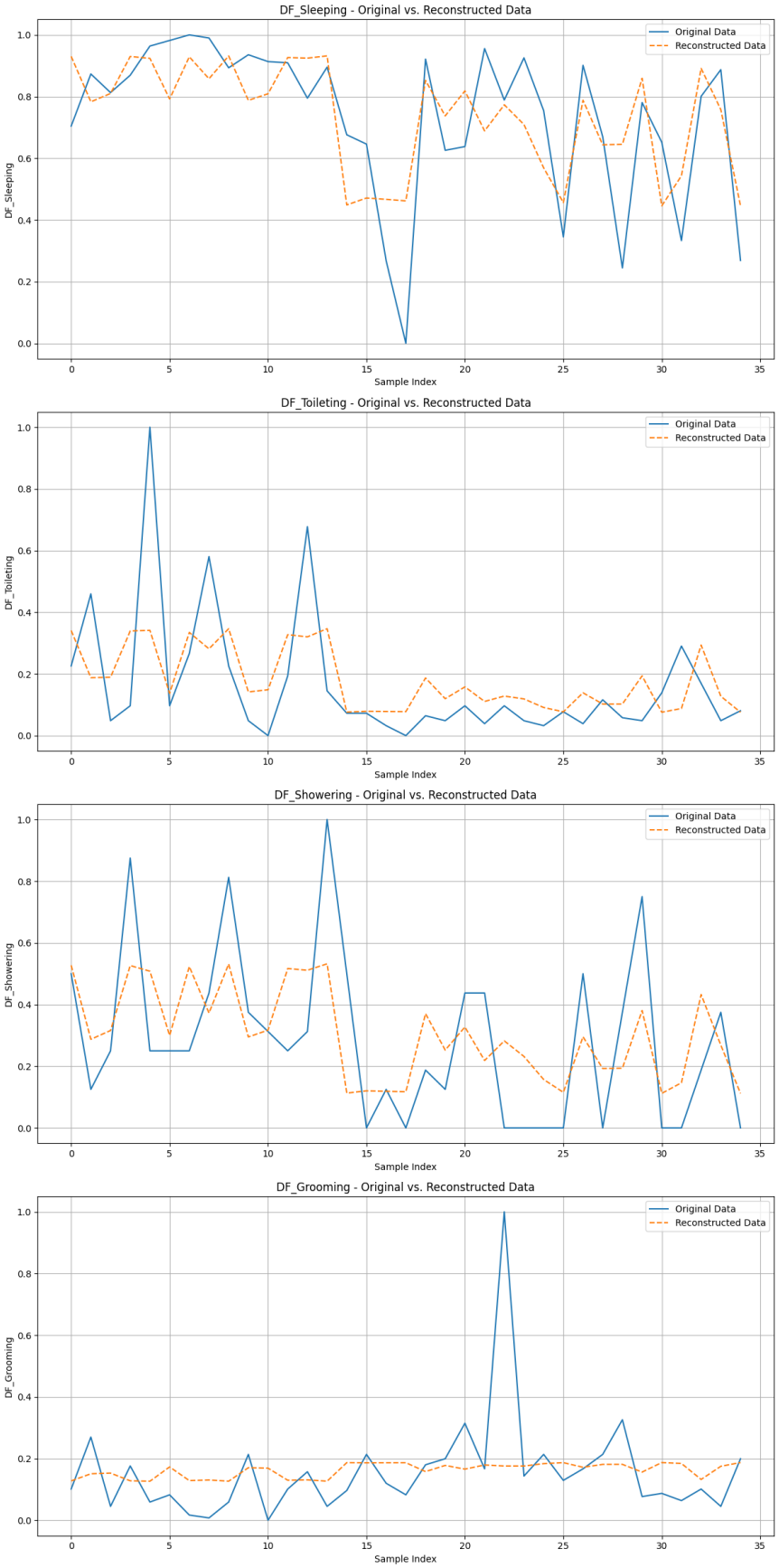

FIGURE 4.21: Average Duration spent on each activity Features: Original vs. Reconstructed Data (1-4) Ordonez' dataset Experiment 1

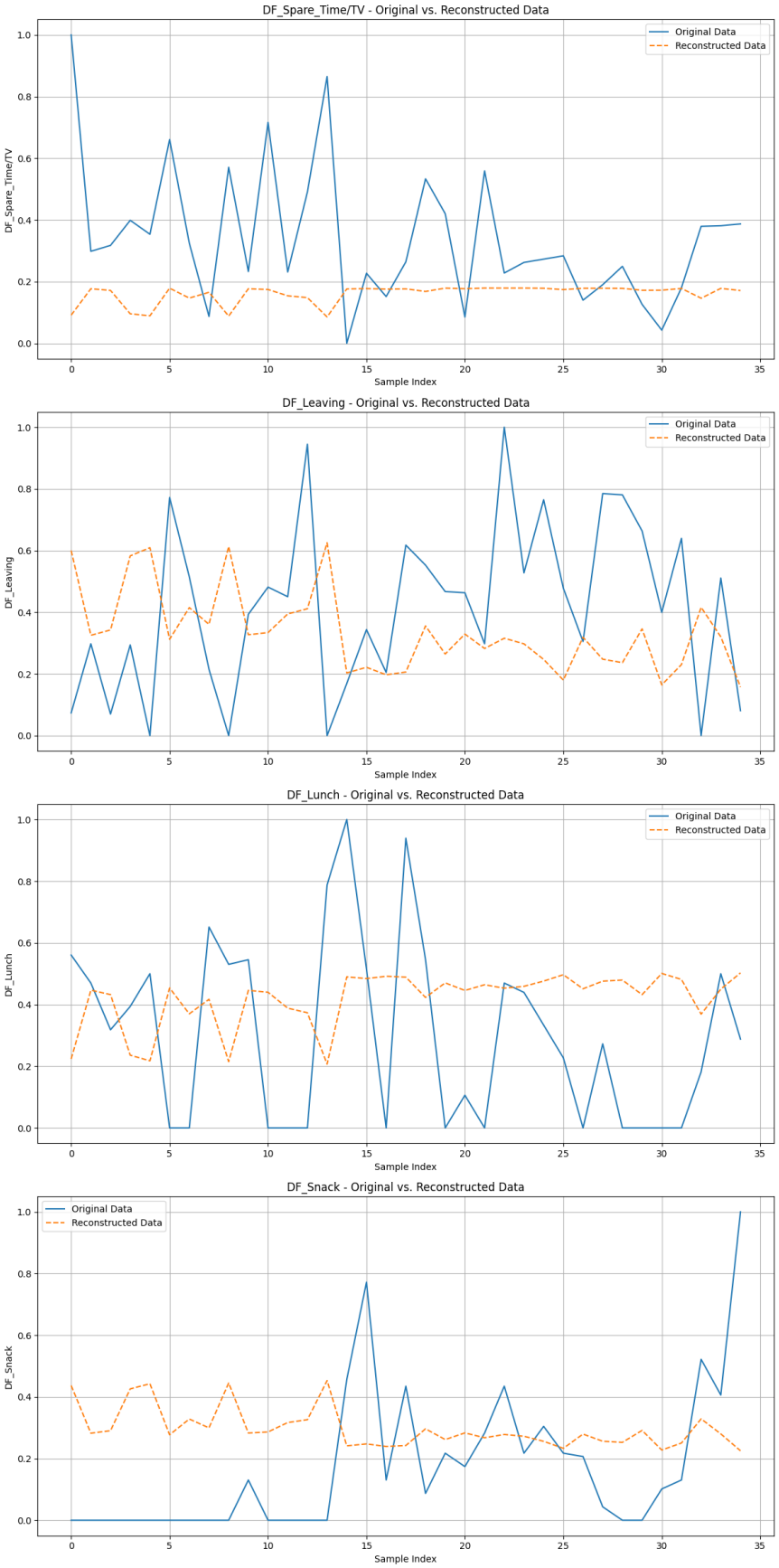

FIGURE 4.22: Average Duration spent on each activity Features: Original vs. Reconstructed Data (5-8) Ordonez' dataset Experiment 1

D_Sleeping	D_Toileting	D_Showering	D_Breakfast	D_Grooming...	FC_Sleeping	FC_Toileting	FC_Showering	FC_Breakfast	FC_Grooming...	Repetitive_Pattern_count	RepetitiveP_Cumulative_L	Repetitive_LC2..
0 days 07:51:0	0 days 00:07:0	0 days 00:08:00	0 days 00:09:0	0 days 00:04:0(	1	3	1	1	2	0	0	0
0 days 09:15:0	0 days 00:19:0	0 days 00:02:00	0 days 00:10:0	0 days 00:15:0(	1	4	1	1	3	1	2	1
0 days 08:45:0	0 days 00:03:0	0 days 00:04:00	0 days 00:13:0	0 days 00:02:0(	1	6	1	1	2	1	3	0
0 days 09:13:0	0 days 00:01:0	0 days 00:14:00	0 days 00:12:0	0 days 00:10:0(	1	1	1	1	3	0	0	0
0 days 10:00:0	0 days 00:31:0	0 days 00:04:00	0 days 00:10:0	0 days 00:05:0(	1	3	1	1	4	0	0	0
0 days 10:09:0	0 days 00:02:0	0 days 00:04:00	0 days 00:09:0	0 days 00:05:0(	1	2	1	1	3	1	2	1
0 days 10:18:0	0 days 00:11:0	0 days 00:04:00	0 days 00:03:0	0 days 00:01:0(	1	4	1	1	2	0	0	0
0 days 10:13:0	0 days 00:30:0	0 days 00:07:00	0 days 00:10:0	0 days 00:02:0(	1	5	1	1	6	1	2	1
0 days 09:25:0	0 days 00:07:0	0 days 00:13:00	0 days 00:05:0	0 days 00:05:0(	1	3	1	1	4	0	0	0
0 days 09:46:0	0 days 00:03:0	0 days 00:06:00	0 days 00:08:0	0 days 00:20:0(	1	6	1	1	5	2	2	0
0 days 09:35:0	0 days 00:00:0	0 days 00:05:00	0 days 00:09:0	0 days 00:01:0(	1	1	1	1	5	1	3	0
0 days 09:33:0	0 days 00:04:0	0 days 00:04:00	0 days 00:08:0	0 days 00:10:0(	1	2	1	1	5	0	0	0
0 days 08:36:0	0 days 00:07:0	0 days 00:05:00	0 days 00:07:0	0 days 00:12:0(	1	1	1	1	4	0	0	0
0 days 09:26:0	0 days 00:03:0	0 days 00:16:00	0 days 00:09:0	0 days 00:03:0(	1	2	1	1	3	0	0	0
0 days 07:37:0	0 days 00:06:0	0 days 00:08:00	0 days 00:10:0	0 days 00:23:0(	1	8	1	1	7	3	7	1
0 days 07:22:0	0 days 00:03:0	0 days 00:00:00	0 days 00:08:0	0 days 00:16:0(	1	4	0	1	4	2	3	1
0 days 08:28:0	0 days 00:01:0	0 days 00:02:00	0 days 00:34:0	0 days 00:14:0(	2	3	1	2	6	3	4	1
0 days 08:04:0	0 days 00:00:0	0 days 00:00:00	0 days 00:10:0	0 days 00:10:0(	4	4	0	1	6	2	3	1
0 days 09:39:0	0 days 00:02:0	0 days 00:03:00	0 days 00:13:0	0 days 00:17:0(	1	3	1	1	5	1	1	0
0 days 07:12:0	0 days 00:02:0	0 days 00:02:00	0 days 00:06:0	0 days 00:30:0(	1	4	1	1	8	3	3	0
0 days 07:18:0	0 days 00:02:0	0 days 00:07:00	0 days 00:07:0	0 days 00:29:0(	1	2	1	1	5	1	1	0
0 days 09:56:0	0 days 00:02:0	0 days 00:07:00	0 days 00:12:0	0 days 00:19:0(	1	5	1	1	6	2	3	1

TABLE 4.8: Feature set used for Ordonez's Experiment 2

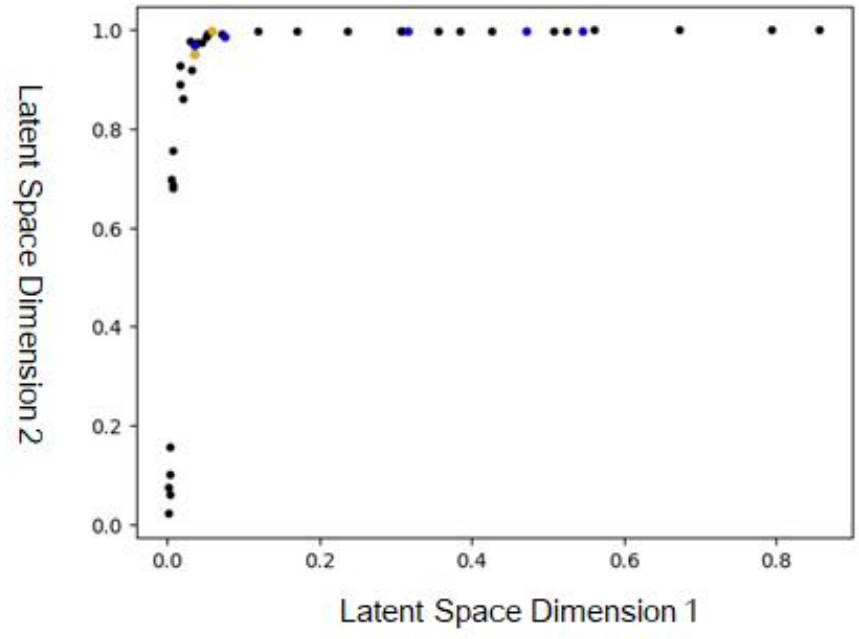

FIGURE 4.23: Embedding Normal and Abnormal data Ordonez' dataset Experiment 1

The same batch size and number of epochs as in Experiment 1 were employed to train the autoencoder. Subsequently, both the original and abnormal data were passed through the encoder. This process generates embedding for both types of data samples. To visually inspect the distribution of normal and abnormal data in a 2D plot, scatter plots are created as shown in Figure 4.25.

The embedding for normal data are represented in black color, while the embedding for abnormal data are shown in red color. This visualization helps us understand the clustering and separation of normal and abnormal data points in the reduced-dimensional space.

Next, the data is reconstructed from the embedding using the trained Autoencoder's decoder part, achieving an accuracy of 97.14% when reconstructing the same input data on which the autoencoder was trained. The reconstruction error (MSE) for all features are shown in Figure 4.26.
The reconstruction error (Mean Squared Error or MSE) provides insights into how effectively the autoencoder model captures and reproduces different data patterns.

```
Model: "model_39"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_27 (InputLayer)       [(None, 22)]              0

 dense_78 (Dense)            (None, 12)                276

 dense_79 (Dense)            (None, 12)                156

 dense_80 (Dense)            (None, 2)                 26

=================================================================
Total params: 458
Trainable params: 458
Non-trainable params: 0
_________________________________________________________________
Model: "model_40"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_28 (InputLayer)       [(None, 2)]               0

 dense_81 (Dense)            (None, 12)                36

 dense_82 (Dense)            (None, 12)                156

 dense_83 (Dense)            (None, 22)                286

=================================================================
Total params: 478
Trainable params: 478
Non-trainable params: 0
```

FIGURE 4.24: Architecture of autoencoder for Experiment 2 for Ordonez's dataset

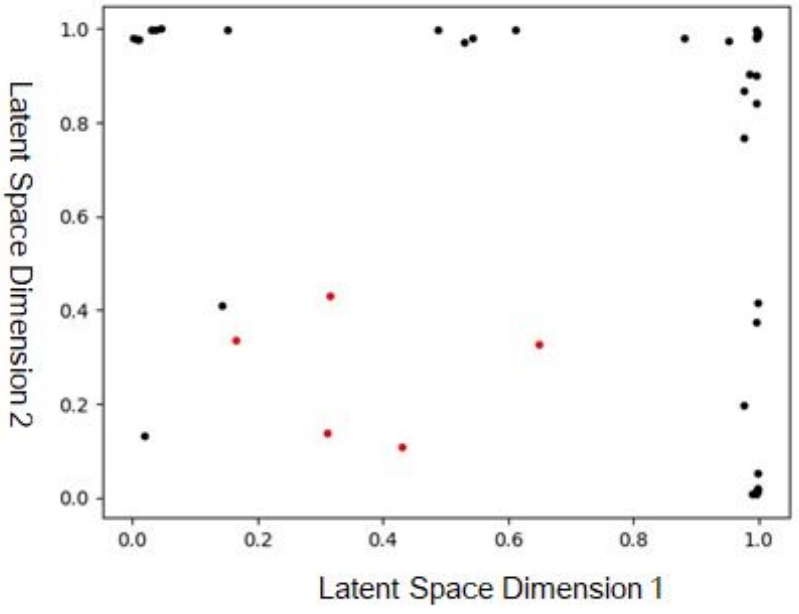

FIGURE 4.25: Two-Dimensional Plot of Embedding for Normal and Abnormal Data for Experiment 2 of Ordonez's dataset

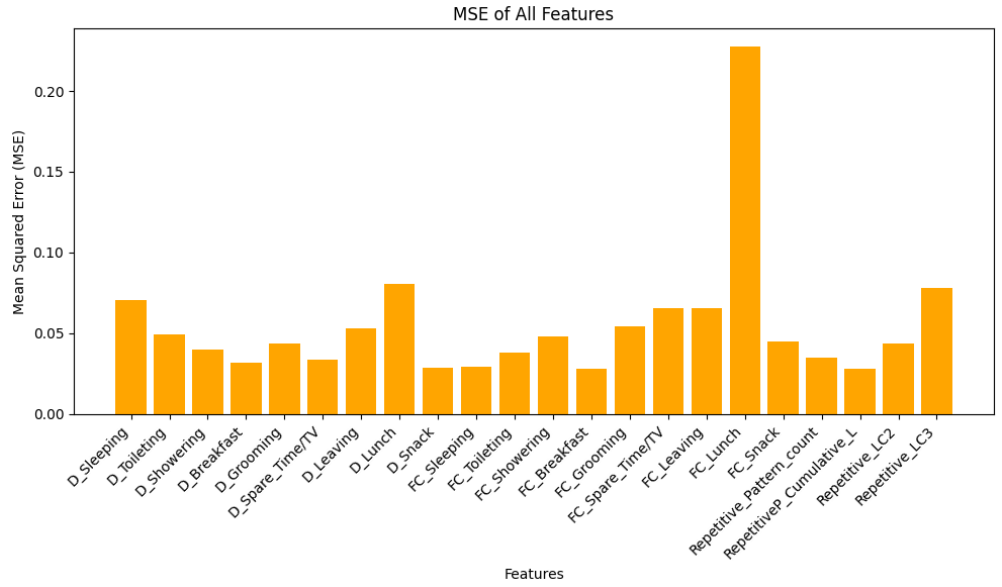

FIGURE 4.26: The reconstruction error (MSE) for all features of Experiment 2 for Ordonez's dataset

As observed in Figure 4.26, only a few features exhibit a high Mean Squared Error (MSE) value, including D_Lunch, FC_Lunch, and Repetitive_LC3. However, the overall reconstruction error for experiment 2 is relatively lower than that for experiment 1 features. This indicates that the autoencoder more accurately reconstructs the features for Experiment 2 than for Experiment 1.

For enhanced visualization and a deeper insight into the quality of the autoencoder's reconstruction, line charts for Experiment 2 are generated as well. These charts depict the mapping of the input data (original data) onto the reconstructed data, for Experiment 2 features, which include Duration, Frequency count (separately for each activity), and Repetitive patterns. Figures 4.27, 4.28, and 4.29, 4.30, 4.31 illustrate that features with low MSE values are accurately reconstructed, while those with high MSE values show poorer reconstruction.

As observed, most of the features have been successfully reconstructed, while some exhibit less accurate reconstruction. Then the trained autoencoder, specifically trained on normal data for the Experiment 2 feature set, was applied to reconstruct abnormal data. This step allowed us to assess whether the autoencoder makes mistakes while reconstructing abnormal instances.

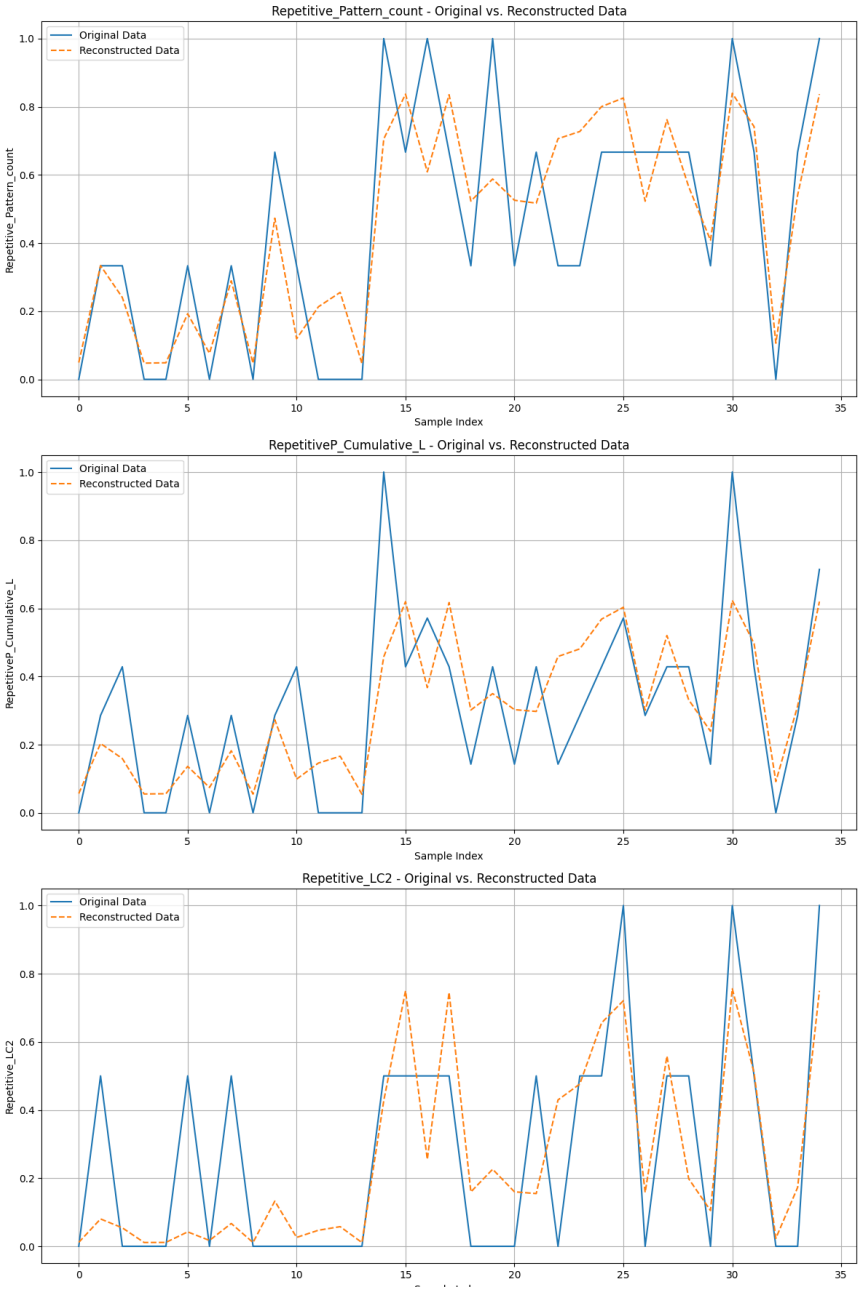

FIGURE 4.27: Repetitive Pattern Feature Mapping: Original vs. Reconstructed Data for Experiment 2

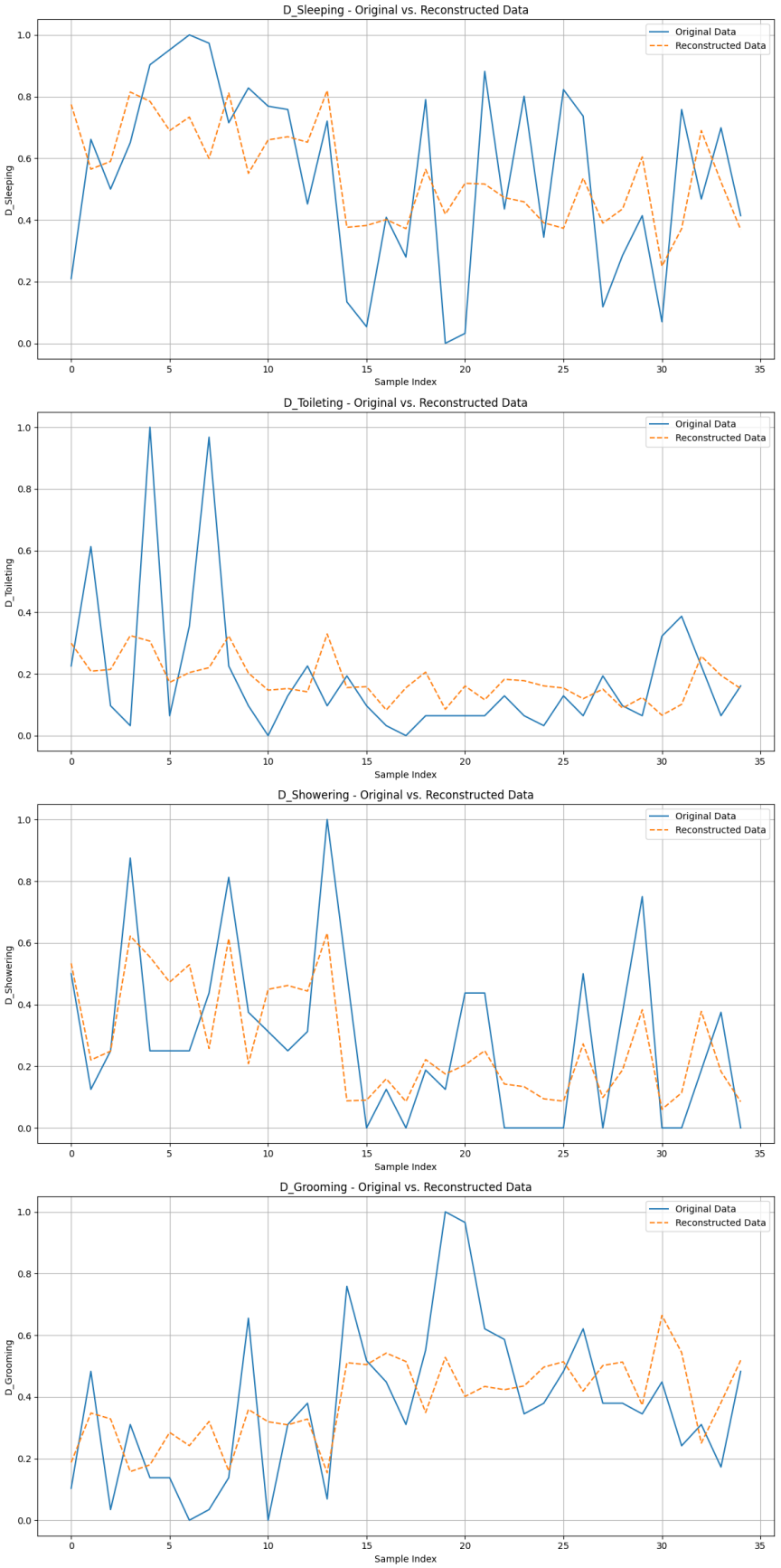

FIGURE 4.28: Duration spent on an activity Features: Original vs. Reconstructed Data (1-5) Ordonez' dataset Experiment 2

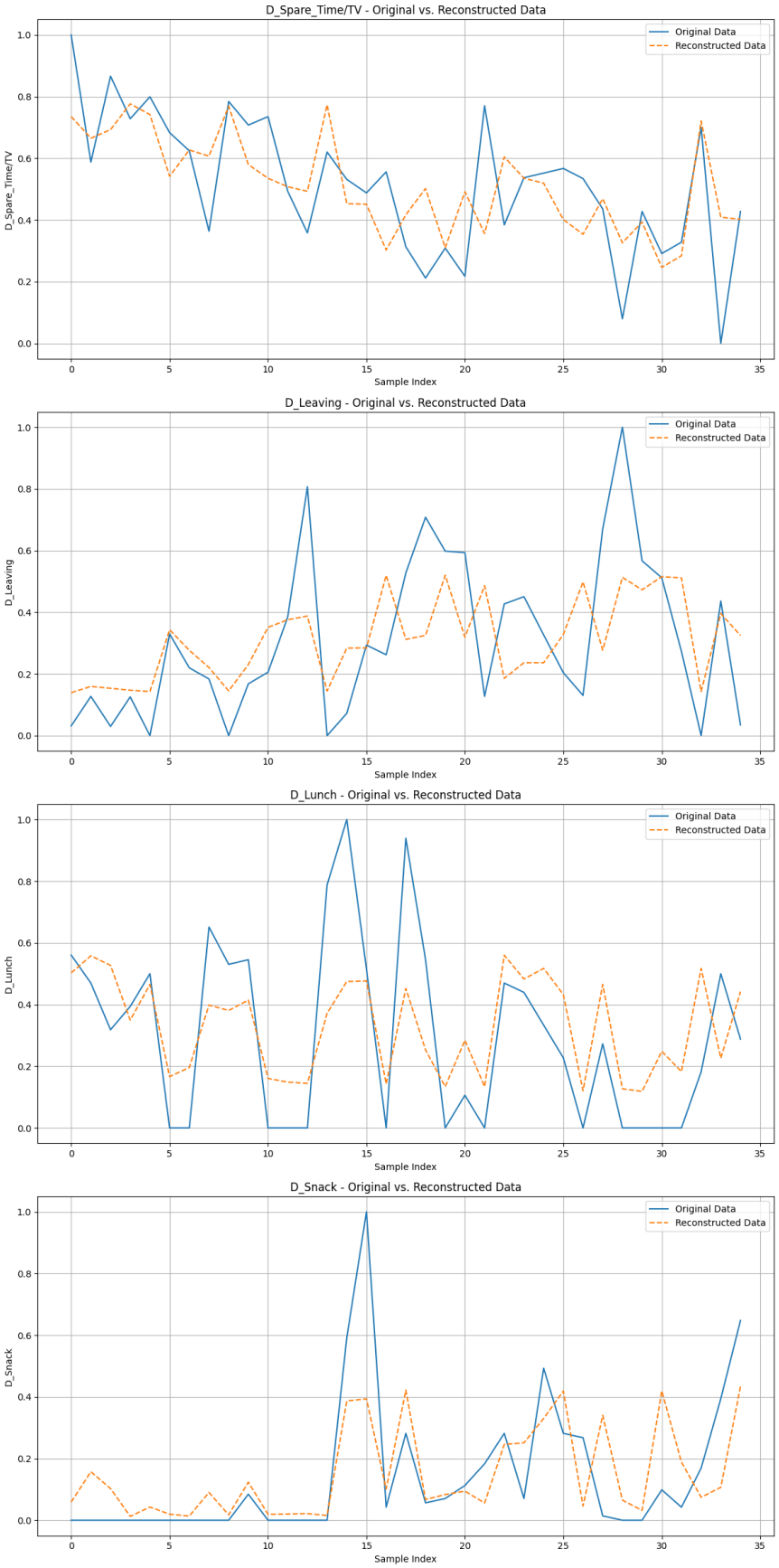

FIGURE 4.29: Duration spent on an activity Features: Original vs. Reconstructed Data (6-9) Ordonez' dataset Experiment 2

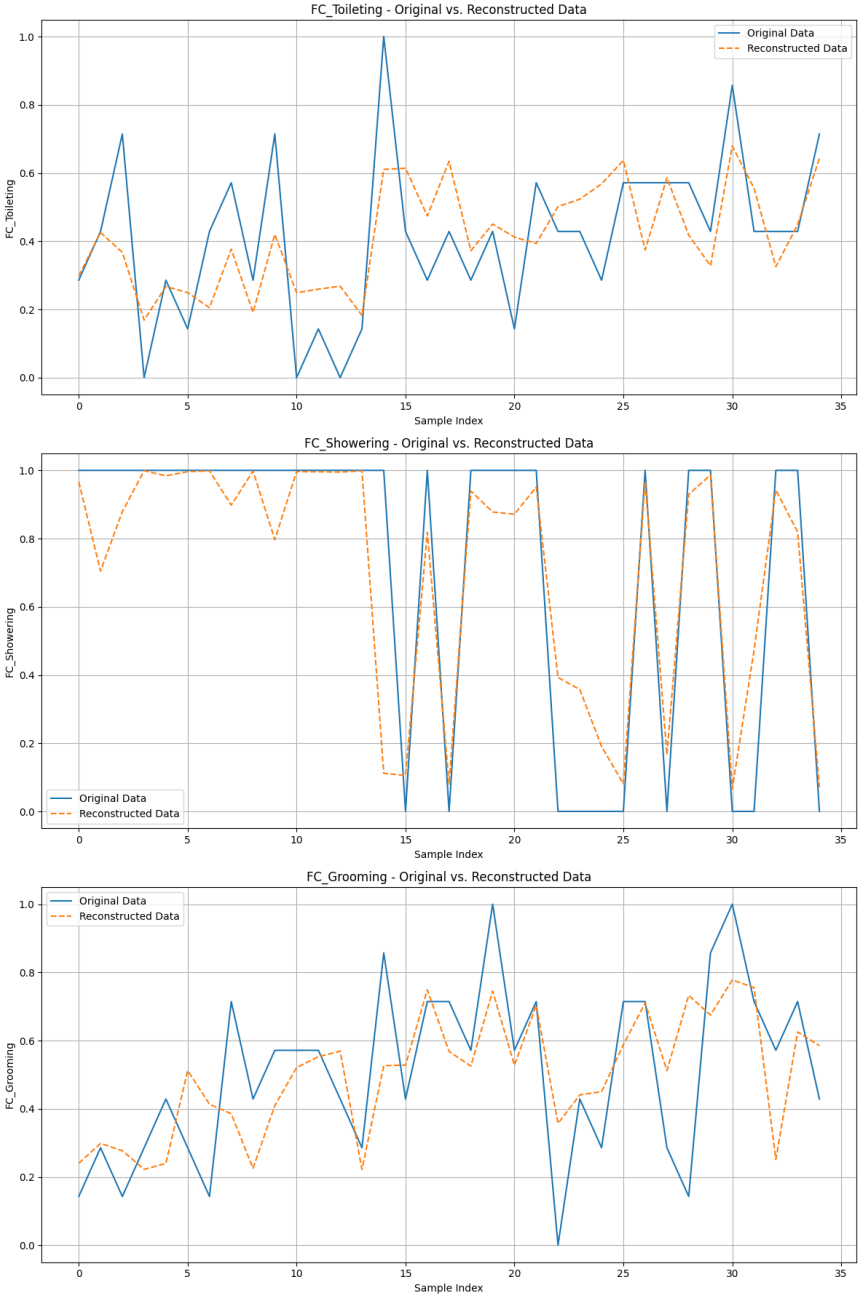

FIGURE 4.30: Frequency Count Features: Original vs. Reconstructed Data (10-13) Ordonez' dataset Experiment 2

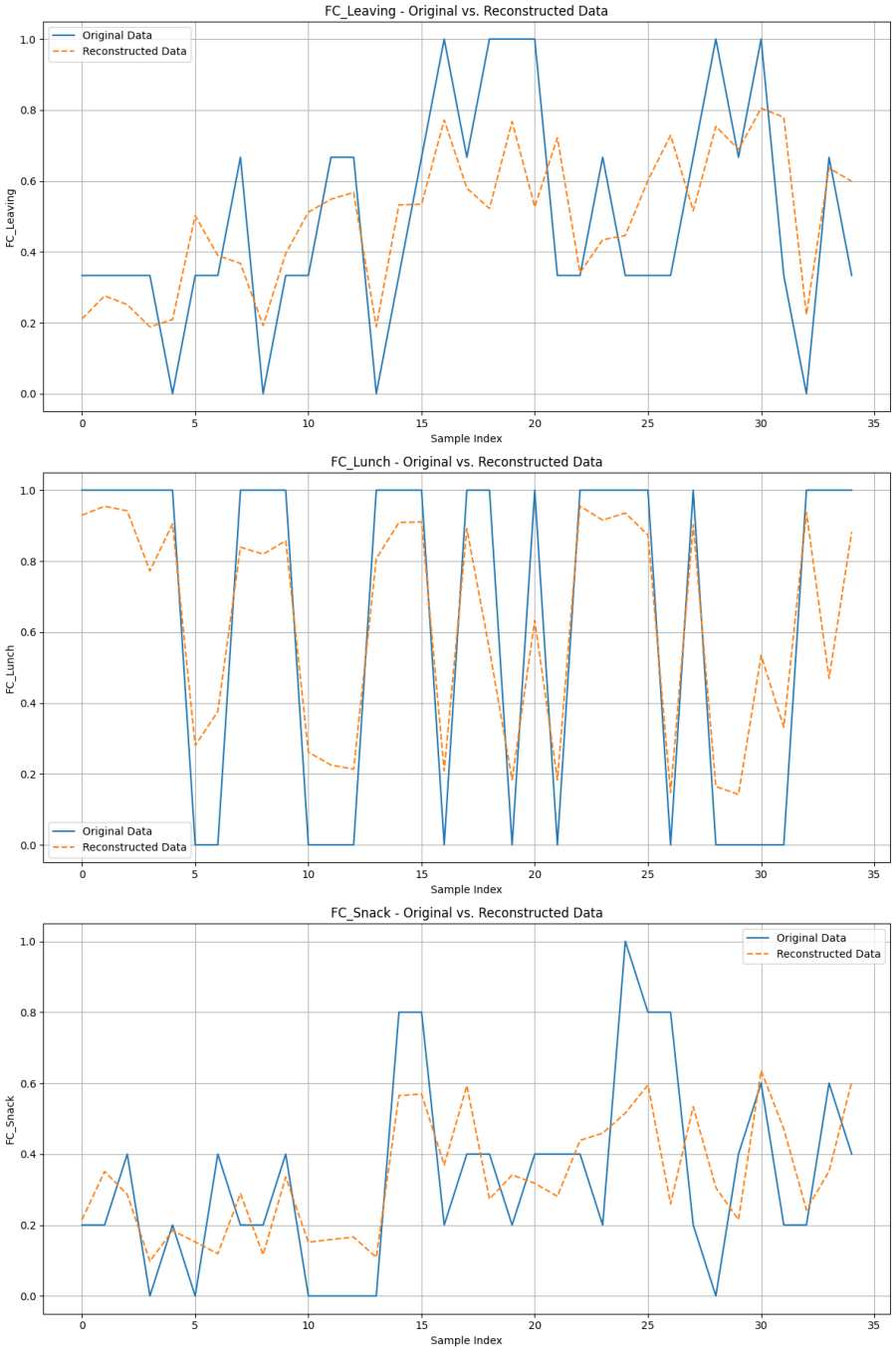

FIGURE 4.31: Frequency Count Features: Original vs. Reconstructed Data (14-17) Ordonez' dataset Experiment 2

Five different abnormal cases exhibiting dementia patient behaviors were created using the Experiment 2 feature set and reconstructed using the trained autoencoder. As expected, the autoencoder made mistakes with the abnormal cases, resulting in all of these cases being successfully detected, which led to a 0% accuracy.

Figure 4.32 displays data points representing both normal and abnormal instances. Black points denote normal data clusters, blue points represent instances suc-

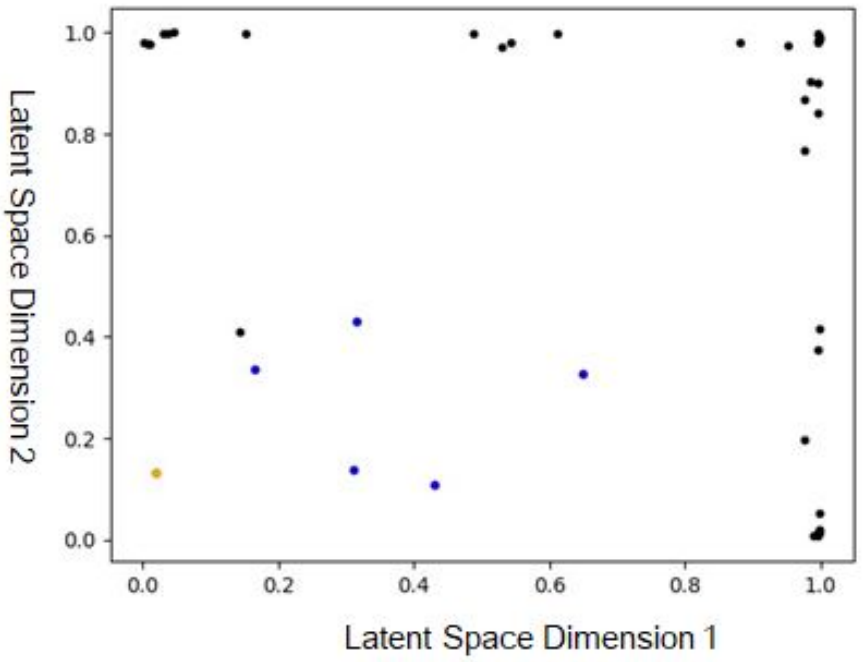

FIGURE 4.32: Embedding Normal and Abnormal data of Ordonez' dataset Experiment 2

cessfully detected as abnormal by the autoencoder, and orange points indicate instances where the autoencoder made mistakes.

4.8 Discussion

In Experiment 1, a similar approach was followed for both the Aruba and Ordonez's datasets.

This involved combining duration and frequency features by averaging the duration spent on each activity, along with the repetitive patterns features. However, during the data reconstruction, it was observed that the majority of the combined duration and frequency features were not accurately reconstructed for both datasets. But the features related to repetitive patterns were successfully reconstructed.

In the case of normal data reconstruction (reproducing the original data on which the autoencoder was trained) in Experiment 1 for the Aruba dataset, an accuracy rate of 96% was achieved. For the Ordonez dataset, the accuracy reached 94.29%. Importantly, all abnormal cases were effectively detected in both datasets during Experiment 1. However, it's important to note that these results may not be entirely reliable, as the combined duration and frequency features were not properly reconstructed on their own data.

To address this issue and gain further insights, Experiment 2 was conducted for both datasets. In this experiment, separate duration and frequency count features were used. The model was retrained, incorporating duration, frequency, and repetitive patterns features, employing the same settings as in Experiment 1. The results yielded some variations between the two datasets: for the Aruba dataset, the accuracy in normal data reconstruction slightly decreased in Experiment 2 compared to Experiment 1 while in Ordonez's dataset accuracy improved. However, by separating the features, significantly improved results were achieved in feature reconstruction. In both datasets, the majority of features were successfully reconstructed in Experiment 2, with only a few exceptions. Remarkably, all abnormal cases were successfully detected in both the Aruba and Ordonez datasets during Experiment 2.

These findings suggest that the features used in Experiment 2 yield more accurate results compared to Experiment 1, particularly in terms of feature reconstruction. The achieved accuracy results are summarized in Table 4.9, offering a comprehensive overview of the evaluation outcomes.

TABLE 4.9: Auto-Encoder Performance Evaluation on ADLs Datasets

Dataset	**Experiment**	**No. of features used for training**	**Accuracy Results for Reconstruction on Normal Data**	**Accuracy Results for Test Cases**
Aruba Testbed	1	14	96.82%	0%
Aruba Testbed	2	24	94.09%	0%
Ordonez's,	1	13	94.29%	0%
Ordonez's	2	22	97.14%	0%

As it can seen from the evaluation results in Table 4.5, the objectives of our Research Questions have been successfully achieved. The objective of RQ1 is to develop a representation of normal behavior using data only from healthy patients by selecting a suitable model. In this table, it is observed that the accuracy results for normal behavior reconstruction in both datasets are above 90%.

While RQ2 focuses on creating features that enable our selected model to distinguish dementia-specific behavioral patterns from ADL data of "normal" people. The created features are highly effective in detecting dementia-specific anomalies, as indicated in Table 4.9, where we achieved 0% accuracy for all experiments as our Autoencoder model, trained on normal data, and when applied to dementia patients' data, it generates errors and provides 0% accuracy, indicating its successful detection of all dementia-related anomalies.

Hence, both research objectives have been successfully achieved by validating the research questions through our experiments. This demonstrates the effectiveness of our chosen model and engineered features in accurately distinguishing dementia-specific anomalies from ADL data of 'normal' individuals, providing valuable insights into dementia detection and anomaly identification.

4.9 Abnormal Data Generation

After obtaining encouraging results from the Autoencoder model, which had been trained on normal data and successfully identified 5 dementia-specific test cases, we proceeded to generate abnormal data to further validate the model performance. The steps for generating abnormal data are illustrated in Figure 4.33.

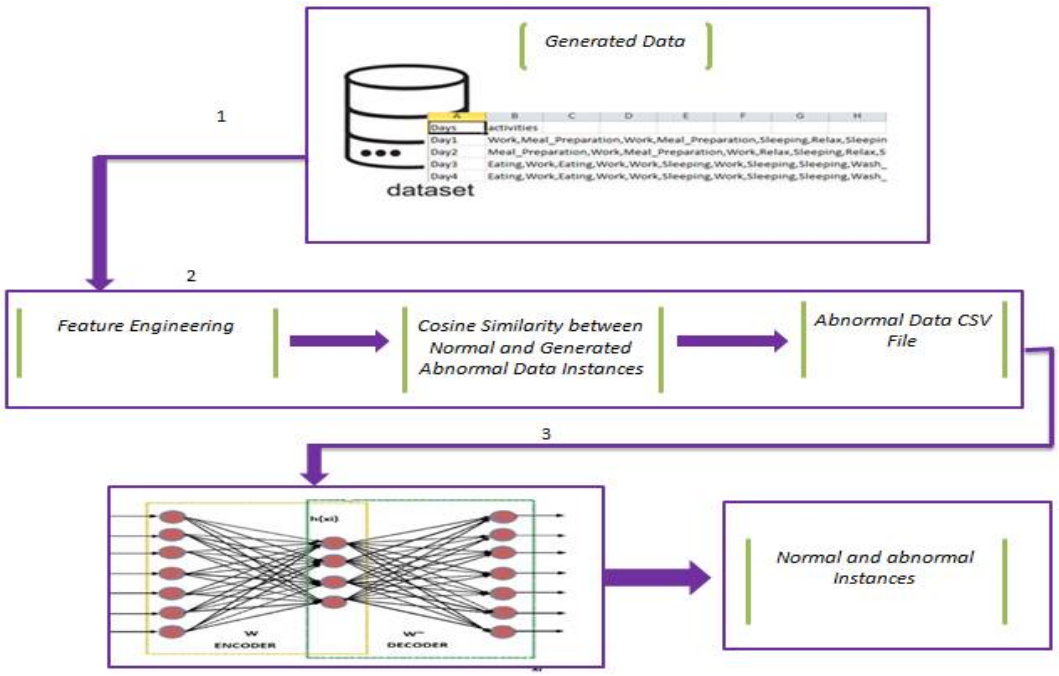

FIGURE 4.33: Abnormal Data Generation

This involved developing a program that randomly introduced dementia-specific repetitive patterns into daily routines, resulting in a wide variety of abnormal routines.

Once the Abnormal dataset was ready, the next step involved engineering features, similar to what was done for the normal data. Prior to applying the Autoencoder, cosine similarity was calculated between the Normal and generated abnormal data. Since patterns were introduced randomly, there was a chance of including Normal Instances in the Abnormal data. To isolate genuine Abnormal instances, cosine similarity was measured with the Normal data, and instances significantly different based on a threshold value of 0.5 were separated into a separate file. These separated instances were truly Abnormal as they significantly deviated from the Normal data.

With this refined Abnormal dataset in hand, the next step involved subjecting the abnormal instances to further validation using the Autoencoder model, originally trained on normal data. As a result, an accuracy rate of 5.35% was achieved on the abnormal data shown in 4.10 , indicating that the Autoencoder correctly identified 95% of the anomalies.

TABLE 4.10: Auto-Encoder Performance Evaluation on Abnormal Datasets

Dataset	**No. of features used for training**	**Accuracy Results for Reconstruction on Normal Data**	**Accuracy Results on Abnormal Data**
Aruba Testbed	14	90.00%	5.35%

The implementation links to the Python code notebooks for our proposed work are accessible below:

1. Experiment 1 for Aruba Dataset (https://tinyurl.com/yw7cfr37)
2. Experiment 2 for Aruba Dataset (https://tinyurl.com/yu3nvhsm)
3. Experiment 1 for Ordonez's Dataset (https://tinyurl.com/5h4jpj3r)
4. Experiment 2 for Ordonez's Dataset (https://tinyurl.com/47anseep)

Chapter 5

Conclusion and Future Work

5.1 Conclusion

In this thesis, an approach for detecting dementia using everyday activity data is presented. Two key challenges encountered in existing research were addressed. Firstly, existing approaches not able to differentiate between various types of anomalies in behavior. Existing methods couldn't reliably tell if the unusual behavior was due to dementia or some other cause like fever or injury. The proposed approach aimed to isolate anomalies specifically associated with dementia, focusing on these dementia specific irregularities. Secondly, existing techniques assumed that the data included both regular and irregular behavior. But what if someone had dementia from the beginning? In such cases, there might not be any noticeable change because their "normal" behavior is already affected by dementia.

To address these gaps, a model (Autoencoder) was trained on normal data to understand what normal behavior looks like. Then, this model was tested with cases developed using medical knowledge about dementia. Two experiments were conducted using different combinations of attributes on the Aruba and Ordonez datasets to assess the model's performance. The results indicated that in Experiment 2, utilizing separate duration and frequency count features led to better performance in accurately reconstructing behavior data. In summary, the deployment of an Autoencoder model, trained on normal data, showcased promising

results in identifying dementia-related anomalies, particularly with the utilization of separate duration and frequency count features.These findings have the potential to improve early dementia detection and care, offering valuable assistance to caregivers and healthcare providers. By enabling timely intervention and providing insights into dementia-related behaviors, our work has the capacity to enhance the quality of life for individuals affected by this condition. Furthermore, our research also provide new insights to researchers, for further investigation and innovation in the field of dementia research.

5.2 Future Directions

The existing work can be extended and enriched in several ways. Firstly, the validation of our approach could be greatly enhanced by obtaining ADL data from actual dementia patients. This would provide a real-world validation of the model's effectiveness in identifying dementia-related irregularities and could serve as a crucial step toward practical implementation. Additionally, expanding the evaluation to encompass diverse dementia-related datasets from various geographical locations and demographics would validate the model's generalizability and applicability across different contexts.

Moreover, One potential area of exploration involves analyzing sub-activities within each activity with the assistance of sensor data could be explored. Sensor data can provide detailed insights into the execution of activities, enabling a more fine-grained understanding of behavioral patterns associated with dementia.

Furthermore, incorporating more features that capture other valuable signs of dementia could further improve the model's accuracy and sensitivity. For instance, features related to sleep patterns, mood fluctuations, or social interactions might offer additional information to strengthen dementia detection. Additionally, domain experts' feedback and clinical validation are crucial to enhancing the model's practicality in dementia detection and supporting early intervention efforts.

Bibliography

[1] "Welcome to CASAS — casas.wsu.edu." https://casas.wsu.edu/datasets/. [Accessed 10-02-2023].

[2] F. J. Ordónez, P. De Toledo, and A. Sanchis, "Activity recognition using hybrid generative/discriminative models on home environments using binary sensors," *Sensors*, vol. 13, no. 5, pp. 5460–5477, 2013.

[3] WHO, "Ageing and health." https://www.who.int/news-room/fact-sheets/detail/ageing-and-health.

[4] V. Fuster, "Changing demographics: a new approach to global health care due to the aging population," 2017.

[5] "What Is Dementia? — alz.org." https://www.alz.org/alzheimers-dementia/what-is-dementia. [Accessed 22-08-2023].

[6] "Dementia — who.int." https://www.who.int/news-room/fact-sheets/detail/dementia. [Accessed 20-02-2023].

[7] A. D. International, "World alzheimer report 2019." https://www.alzint.org/u/WorldAlzheimerReport2019.pdf.

[8] J. Rasmussen and H. Langerman, "Alzheimer's disease – why we need early diagnosis," *Degenerative Neurological and Neuromuscular Disease*, vol. 9, pp. 123–130, 2019.

[9] J. Zenisek, F. Holzinger, and M. Affenzeller, "Machine learning based concept drift detection for predictive maintenance," *Computers & Industrial Engineering*, vol. 137, p. 106031, 2019.

[10] Ó. Belmonte-Fernández, A. Caballer-Miedes, E. Chinellato, R. Montoliu, E. Sansano-Sansano, and R. García-Vidal, "Anomaly detection in activities of daily living with linear drift," *Cognitive Computation*, vol. 12, no. 6, pp. 1233–1251, 2020.

[11] Y. Maki, T. Yamaguchi, and H. Yamaguchi, "Symptoms of early dementia-11 questionnaire (sed-11q): a brief informant-operated screening for dementia," *Dementia and geriatric cognitive disorders extra*, vol. 3, no. 1, pp. 131–142, 2013.

[12] D. Zekri, T. Delot, M. Thilliez, S. Lecomte, and M. Desertot, "A framework for detecting and analyzing behavior changes of elderly people over time using learning techniques," *Sensors*, vol. 20, no. 24, p. 7112, 2020.

[13] B. Friedrich, T. Sawabe, and A. Hein, "Unsupervised statistical concept drift detection for behaviour abnormality detection," *Applied Intelligence*, vol. 53, no. 3, pp. 2527–2537, 2023.

[14] M. Bi, J. Xu, M. Wang, and F. Zhou, "Anomaly detection model of user behavior based on principal component analysis," *Journal of Ambient Intelligence and Humanized Computing*, vol. 7, pp. 547–554, 2016.

[15] C. Azefack, R. Phan, V. Augusto, G. Gardin, C. M. Coquard, R. Bouvier, and X. Xie, "An approach for behavioral drift detection in a smart home," in *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*, pp. 727–732, IEEE, 2019.

[16] "Redirect Notice — google.com." https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.mdpi.com%2F1424-8220%2F15%2F5%2F11953&psig=AOvVaw2yGsMHkvESzK4GOOlXSQ6x&ust=1699171315060000&source=images&cd=vfe&opi=89978449&ved=0CBIQjRxqFwoTCNjTl_7wqYIDFQAAAAAdAAAAABAS. [Accessed 04-11-2023].

[17] "Repetitive behaviour and dementia — alzheimers.org.uk." https://www.alzheimers.org.uk/about-dementia/symptoms-and-diagnosis/symptoms/repetitive-behaviour. [Accessed 17-08-2023].

[18] "Repetition — alz.org." https://www.alz.org/help-support/caregiving/stages-behaviors/repetition. [Accessed 17-08-2023].

[19] J. Bourgeois, M. Laye, J. Lemaire, E. Leone, A. Deudon, N. Darmon, C. Giaume, V. Lafont, S. Brinck-Jensen, A. Dechamps, *et al.*, "Relearning of activities of daily living: a comparison of the effectiveness of three learning methods in patients with dementia of the alzheimer type," *The journal of nutrition, health & aging*, vol. 20, pp. 48–55, 2016.

[20] "Repetition in people with dementia - SCIE — scie.org.uk." https://www.scie.org.uk/dementia/living-with-dementia/difficult-situations/repetition.asp#:~:text=People%20with%20dementia%20often%20repeat,and%20needs%20behind%20the%20repetition. [Accessed 17-08-2023].

[21] F. S. Ujager and A. Mahmood, "A context-aware accurate wellness determination (caawd) model for elderly people using lazy associative classification," *Sensors*, vol. 19, no. 7, p. 1613, 2019.

[22] L.-N. Kwon, D.-H. Yang, M.-G. Hwang, S.-J. Lim, Y.-K. Kim, J.-G. Kim, K.-H. Cho, H.-W. Chun, and K.-W. Park, "Automated classification of normal control and early-stage dementia based on activities of daily living (adl) data acquired from smart home environment," *International journal of environmental research and public health*, vol. 18, no. 24, p. 13235, 2021.

[23] D. Bank, N. Koenigstein, and R. Giryes, "Autoencoders," *arXiv preprint arXiv:2003.05991*, 2020.

[24] D. Bouchabou, S. M. Nguyen, C. Lohr, B. LeDuc, and I. Kanellos, "A survey of human activity recognition in smart homes based on iot sensors algorithms: Taxonomies, challenges, and opportunities with deep learning," *Sensors*, vol. 21, no. 18, p. 6037, 2021.

[25] P. Ariza Colpas, E. Vicario, E. De-La-Hoz-Franco, M. Pineres-Melo, A. Oviedo-Carrascal, and F. Patara, "Unsupervised human activity recognition using the clustering approach: A review," *Sensors*, vol. 20, no. 9, p. 2702, 2020.

[26] S. K. Hiremath, Y. Nishimura, S. Chernova, and T. Plötz, "Bootstrapping human activity recognition systems for smart homes from scratch," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 6, no. 3, pp. 1–27, 2022.

[27] T.-H. Tan, L. Badarch, W.-X. Zeng, M. Gochoo, F. S. Alnajjar, and J.-W. Hsieh, "Binary sensors-based privacy-preserved activity recognition of elderly living alone using an rnn," *Sensors*, vol. 21, no. 16, p. 5371, 2021.

[28] A.-C. P. Patricia, V. Enrico, B. A. Shariq, D.-l.-F. Emiro, P.-M. M. Alberto, O.-C. A. Isabel, M. I. Tariq, J. K. G. Restrepo, and P. Fulvio, "Machine learning applied to datasets of human activity recognition: Data analysis in health care," *Current Medical Imaging*, vol. 19, no. 1, pp. 46–64, 2023.

[29] V. R. Chifu, C. B. Pop, A. M. Rancea, A. Morar, T. Cioara, M. Antal, and I. Anghel, "Deep learning, mining, and collaborative clustering to identify flexible daily activities patterns," *Sensors*, vol. 22, no. 13, p. 4803, 2022.

[30] A. Lotfi, C. Langensiepen, S. M. Mahmoud, and M. J. Akhlaghinia, "Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behaviour," *Journal of ambient intelligence and humanized computing*, vol. 3, pp. 205–218, 2012.

[31] L. G. Fahad and S. F. Tahir, "Activity recognition and anomaly detection in smart homes," *Neurocomputing*, vol. 423, pp. 362–372, 2021.

[32] J. Guo, Y. Li, M. Hou, S. Han, and J. Ren, "Recognition of daily activities of two residents in a smart home based on time clustering," *Sensors*, vol. 20, no. 5, p. 1457, 2020.

[33] V. G. Sánchez, O. M. Lysaker, and N.-O. Skeie, "Human behaviour modelling for welfare technology using hidden markov models," *Pattern Recognition Letters*, vol. 137, pp. 71–79, 2020.

[34] V. R. Chifu, C. B. Pop, D. Demjen, R. Socaci, D. Todea, M. Antal, T. Cioara, I. Anghel, and C. Antal, "Identifying and monitoring the daily routine of seniors living at home," *Sensors*, vol. 22, no. 3, p. 992, 2022.

[35] D. Arifoglu and A. Bouchachia, "Detection of abnormal behaviour for dementia sufferers using convolutional neural networks," *Artificial intelligence in medicine*, vol. 94, pp. 88–95, 2019.

[36] K. A. Alaghbari, M. H. M. Saad, A. Hussain, and M. R. Alam, "Activities recognition, anomaly detection and next activity prediction based on neural networks in smart homes," *IEEE Access*, vol. 10, pp. 28219–28232, 2022.

[37] "Symptoms of Early Dementia-11 Questionnaire (SED-11Q): A Brief Informant-Operated Screening for Dementia — ncbi.nlm.nih.gov." https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3656692/. [Accessed 08-03-2023].

[38] S. Zavrak and M. İskefiyeli, "Anomaly-based intrusion detection from network flow features using variational autoencoder," *IEEE Access*, vol. 8, pp. 108346–108358, 2020.

[39] N. Moheb, K. Charuworn, M. M. Ashla, R. Desarzant, D. Chavez, and M. F. Mendez, "Repetitive behaviors in frontotemporal dementia: Compulsions or impulsions?," *The Journal of neuropsychiatry and clinical neurosciences*, vol. 31, no. 2, pp. 132–136, 2019.

[40] M. Tran, M. Bedard, S. Dubois, B. Weaver, and D. Molloy, "The influences of psychotic symptoms on the activities of daily living of individuals with alzheimer disease: a longitudinal analysis," *Aging & Mental Health*, vol. 17, no. 6, pp. 738–747, 2013.

[41] G. Cipriani, S. Danti, L. Picchi, A. Nuti, and M. D. Fiorino, "Daily functioning and dementia," *Dementia & neuropsychologia*, vol. 14, pp. 93–102, 2020.

[42] "Using the toilet - difficult situations in dementia." https://www.scie.org.uk/dementia/living-with-dementia/difficult-situations/using-the-toilet.asp. [Accessed 20-02-2023].

[43] A. if available, "Is it typical for people with dementia to sleep a lot during the day?," Year of Publication if available. [Accessed 20-02-2023].

Turnitin Originality Report

Early Detection of Dementia using Activities of Daily Life Data from Smart Homes by turnitin
Sara Ibrahim

From CUST Library (MS Thesis)

- Processed on 12-Sep-2023 09:30 PKT
- ID: 2163848962
- Word Count: 15189

Graduate Studies Office
09%
Within Limits

Similarity Index
9%
Similarity by Source

Internet Sources:
7%
Publications:
5%
Student Papers:
3%

sources:

1 1% match (student papers from 30-Mar-2016)
Submitted to UNESCO-IHE Institute for Water Education on 2016-03-30

2 1% match (student papers from 26-Apr-2018)
Submitted to Asian Institute of Technology on 2018-04-26

3 1% match (Internet from 28-Sep-2022)
https://res.mdpi.com/sensors/sensors-19-01613/article_deploy/sensors-19-01613-v2.pdf?attachment=1&filename=

4 < 1% match (student papers from 29-Mar-2016)
Submitted to UNESCO-IHE Institute for Water Education on 2016-03-29

5 < 1% match (Internet from 11-Dec-2021)
http://eprints.bournemouth.ac.uk/32750/1/ARIFOGLU%2C%20Damla_Ph.D._2019.pdf

6 < 1% match ()
Arifoglu, D., Bouchachia, Abdelhamid, "Detection of abnormal behaviour for dementia sufferers using Convolutional Neural Networks.", 'Elsevier BV', 2019

7 < 1% match (Khaled A. Alaghbari, Mohamad Hanif Md. Saad, Aini Hussain, Muhammad Raisul Alam. "Activities Recognition, Anomaly Detection and Next Activity Prediction Based on Neural Networks in Smart Homes", IEEE Access, 2022)
Khaled A. Alaghbari, Mohamad Hanif Md. Saad, Aini Hussain, Muhammad Raisul Alam. "Activities Recognition, Anomaly Detection and Next Activity Prediction Based on Neural Networks in Smart Homes", IEEE Access, 2022

8 < 1% match (Internet from 15-Feb-2023)
https://www.researchgate.net/publication/259631936_Evaluation_of_Anosognosia_in_Alzheimer's_Disease_Using_the_Symptoms_of_11_Questionnaire_SED-11Q

9 < 1% match (Internet from 26-Apr-2023)
https://www.researchgate.net/publication/366981366_A_Review_of_Abnormal_Behavior_Detection_in_Activities_of_Daily_Living

10 < 1% match (Internet from 29-Oct-2020)
https://archive.ics.uci.edu/ml/datasets/Activities+of+Daily+Living+(ADLs)+Recognition+Using+Binary+Sensors

11 < 1% match (Internet from 25-Nov-2020)
http://archive.ics.uci.edu/ml/datasets.php?area=&att=&format=&numAtt=&numIns=&sort=nameUp&task=&type=&view=list

12 < 1% match (Labiba Gillani Fahad, Syed Fahad Tahir. "Activity recognition and anomaly detection in smart homes", Neurocomputing, 2021)