King Saud University
College of Computer and Information Sciences
Department of Computer Engineering

CEN 493

A Progress Report for Senior Year Design Project

Reducing the Energy Expenses of Residential Customers by Designing a Deep Learning Method

Submitted by

Abdulaziz Alzahrani 435106315

Abdulaziz Almulhim 435104561

Khalid Ashraf Said 436107353

Supervised by

Dr. Musaed Alhussein

1441AH/2020

Abstract

The power consumption forecasting of residential customers play an important role in the future power grid planning and operation. With the advancement and development of smart metering technologies, the power consumption data of residential customers can be retrieved on hourly basis. Most of the prior works are focused on aggregated power consumption forecasting of residential customers in a community or building level. However, forecasting individual residential customer power consumption is fairly challenging since it relies on customers' lifestyle, number of occupants, weather, calendar information, and so on. Recently, there has been a surge of interest in deep learning methods due to the fact that they have shown to outperform state-of-the-art techniques in several applications. In this project we aim at designing a deep learning method for predicting the power consumption of individual residential customers. The designed method will be evaluated and tested on Smart Grid Smart City (SGSC) database. Individual household power consumption prediction over time allows energy users to evaluate their consumption behaviors and, to shift their energy use to off-peak periods, whenever possible. In this way, the energy users may control their future energy expenses.

ملخص

يلعب التنبؤ باستهلاك الطاقة للعملاء المقيمين دورًا أساسيًا متزايدًا في تخطيط وتشغيل شبكة الطاقة في المستقبل. مقارنةً باستهلاك الطاقة الكلي على مستوى المجتمع ، فإن التنبؤ باستهلاك الطاقة لمستخدم واحد للطاقة يمثل تحديًا كبيرًا بسبب التقلب الشديد وعدم اليقين الذي ينطوي عليه الأمر.

يتيح التنبؤ باستهلاك الطاقة بمرور الوقت لمستخدمي الطاقة تقييم سلوكيات استهلاكهم وتحويل استخدامهم للطاقة إلى فترات خارج أوقات الذروة ، كلما كان ذلك ممكنًا. يتيح التنبؤ الدقيق لاستهلاك الطاقة لمستخدمي الطاقة التحكم في نفقات الطاقة المستقبلية.

يهدف المشروع إلى تصميم طريقة تعلم عميقة للتنبؤ باستهلاك الطاقة للعملاء المقيمين. سيتم تقييم الطريقة المصممة واختبارها على قاعدة بيانات.Smart Grid Smart City (SGSC)

Acknowledgment

First, all praise to Allah, we would like to express our deepest and sincere gratitude to Allah for bestowing his blessing and guidance on us. Second, it is a great honor and privilege to thank our supervisor Dr. Musaed Alhussein for his assistance and guidance and Eng. Syed Irtaza Haider for his technical support.

Table of Contents

Abstract 2

ملخص 3

Acknowledgment 4

List of Abbreviations 9

List of Figures 10

List of Tables 12

Chapter 1 13

Introduction 13

1.1 Project Objective 15

1.1.1 Phase I 15

1.1.2 Phase II 16

1.2 Smart Grid Smart City (SGSC) Database 16

1.3 Scope and Report Organization 16

Chapter 2 18

Background and Related Work 18

2.1 Load Forecasting 18

2.2 Smart Grids vs. Traditional Grids 19

2.3 Classification of Load Forecasting Problems 19

2.4 Important Factors for Short-Term Load Forecasting 20

2.5 Forecasting Accuracy Metrics 20

2.6 Short Term Load Forecasting Techniques 21

2.6.1 Persistence Model 21

2.6.2 Regression Methods 21

2.6.3 Time Series Methods 22

2.6.4 Machine Learning Methods 22

2.6.4.1 Fuzzy logic 22

2.6.4.2 Artificial Neural Network (ANN) 23

2.6.4.3 Support Vector Regression (SVR) 23

2.6.5 Deep learning Methods 24

2.6.5.1 Convolutional Neural Networks (CNN) 24

2.6.5.2 Recurrent Neural Network (RNN) 25

2.6.5.3 Long Short-Term Memory (LSTM) 26

2.7 Related Work 27

2.8 Conclusion 29

Chapter 3 29

Data Analysis and Preparation 29

3.1 SGSC Database 30

3.2 Impact of Customer Behavior on Load Profiles 30

3.3 Fill Missing Data 31

3.4 Evaluation of Daily Load Profiles 31

3.5 Data Preparation for Deep Learning Model 36

3.6 Conclusion 36

Chapter 4 37

LSTM based Deep Learning Framework 37

4.1 Data Split and Test Settings 37

4.2 Preliminary Results and Discussion 37

4.3 Future Work for Second Term 39

Chapter 5 41

Hybrid CNN-LSTM based Deep Learning Framework 41

5.1 Various Layers used in Proposed Model 42

5.1.1 1-D Convolution Layer 42

5.1.2 MaxPooling Layer 44

5.1.3 Rectified Linear Unit (ReLU) 45

5.1.4 Recurrent Neural Network (RNN) and Long-Short Term Memory (LSTM) 45

5.1.5 Dropout Layer 47

5.1.6 Fully Connected Layer 48

5.2 Proposed Hybrid CNN-LSTM Model 48

5.3 Results 52

5.3.1 Single Step Forecast 54

5.3.2 Multiple Step Forecast 58

5.3.3 Cluster Analysis 60

5.4 Conclusion 69

Chapter 6 71

Load Scheduling of Residential Customer for Dynamic Peak Pricing Scheme 71

6.1 Various Trials of Dynamic Peak Pricing (DPP) 73

6.2 Smart Grid Smart City (SGSC) Project 74

6.3 Experimental Setup 80

6.3.1 Residential Electricity Load 80

6.3.2 Load Categorization 81

6.3.3 Load Balancing Techniques 81

6.3.4 Pricing Scheme 83

6.3.5 Algorithms 83

6.3.5.1 Strawberry Search Algorithm (SSA) 83

6.3.5.2 Cuckoo Search Algorithm (CSA) 84

6.3.6 Results 85

6.4 Future Work 90

6.5 Conclusion 90

Appendix: 92

References 93

List of Abbreviations

SG	Smart Grids
DR	Demand Response
EV	Electric Vehicle
ES	Energy Storage
SGSC	Smart Grid Smart City
DBSCAN	Density-Based Spatial Clustering of Application with Noise
STLF	Short-term Load Forecasting
MSE	Mean Square Error
RMSE	Root Mean Square Error
MAD	Mean Absolute Deviation
MAPE	Mean Absolute Percentage Error
ARIMA	Auto Regressive Integrated Moving Average
ANN	Artificial Neural Network
SVR	Support Vector Regression
CNN	Convolutional Neural Network
ReLU	Rectified Linear Unit
RNN	Recurrent Neural Network
LSTM	Long-Short Term Memory
ELM	Extreme Learning Machine
DPP	Dynamic Peak Pricing

List of Figures

Figure 1.1 Aggregated Daily Load Profiles of various Individual Households 14
Figure 1.2 Daily Load Profiles of an Individual Household 15
Figure 2.1 Convolutional Neural Network Architecture 25
Figure 2.2 Basic Building Blocks of Recurrent Neural Network (RNN) 26
Figure 2.3 Basic Building Blocks of LSTM [2] 27
Figure 3.1 Daily Load Profiles of a residential customer (ID 8523058) – Single Major Cluster 33
Figure 3.2 Daily Load Profiles of a residential customer (ID 11462018) – One Major Cluster, One Minor Cluster and Few Outliers 34
Figure 3.3 Daily Load Profiles of a residential customer (ID 8273592) – Two Major Clusters and Several Outliers 34
Figure 3.4 Daily Load Profiles of a residential customer (ID 9012348) – Four Minor Clusters and Several Outliers 35
Figure 3.5 Number of Major Clusters / Minor Clusters and Outliers for few randomly selected households from SGSC database 35
Figure 4.1 Actual Load vs. LSTM forecast for customer ID 8523058 38
Figure 4.2 Actual Load vs. LSTM forecast for customer ID 11462018 38
Figure 4.3 Actual Load vs. LSTM forecast for customer ID 8432046 39
Figure 5.1 Operational example of 1-D convolution. 43
Figure 5.2 Operational example of 1-D MaxPooling. 45
Figure 5.3 Basic structure of RNNs 46
Figure 5.4 Inner Structure of LSTM 47
Figure 5.5 Example of Dropout Layer 48
Figure 5.6 Proposed CNN-LSTM Model 49
Figure 5.7 Proposed CNN-LSTM Model Training Flow 52
Figure 5.8 Energy Consumption Data of 69 Customers from SGSC Database 53
Figure 5.9 Feature Preparation 54
Figure 5.10 Load Forecasting of Customer ID 9393680 using proposed model and LSTM 55

Figure 5.11 Load Forecasting of Customer ID 8257034 using proposed model and LSTM 55
Figure 5.12 Load Forecasting of Customer ID 8181075 using proposed model and LSTM 56
Figure 5.13 Comparison of CNN-LSTM model with the LSTM model 57
Figure 5.14 Silhouette score for grouping residential customers 61
Figure 5.15 Normalized Average Power Consumption of each Cluster 62
Figure 5.16 Cluster_7: Customer ID 8680284 average power consumption during train and test data 67
Figure 5.17 Cluster_3 Consumption Forecast 67
Figure 5.18 Cluster_4 Consumption Forecast 68
Figure 5.19 Cluster_5 Consumption Forecast 68
Figure 5.20 Cluster_7 Consumption Forecast 68
Figure 5.21 Cluster_12 Consumption Forecast 69
Figure 6.1 Key Motivation to participate in the SGSC Trials 75
Figure 6.2 Electricity consumption shift to off-peak time 75
Figure 6.3 Number of Shift able Appliances to Off-Peak Time 75
Figure 6.4 Proportion of appliances customers are unable to shift to off-peak time 76
Figure 6.5 Pricing Design SGSC Database 77
Figure 6.6 Average usage of each control group with DPP during peak events [26] 78
Figure 6.7 Load Shaping Techniques 82
Figure 6.8 Dynamic Peak Pricing Scheme used in SGSC database 83
Figure 6.9 DPP Scheme along with Unscheduled Load (day ahead forecast) 87
Figure 6.10 Electricity cost (Hourly basis) 87
Figure 6.11 Total Cost comparison of unscheduled with SA and CSA 88
Figure 6.12 Unscheduled Load Consumption (Hourly basis) as well as Load Scheduling using SA and CSA 89
Figure 6.13 Peak Average Ratio (PAR) for unscheduled, SA and CSA 89

List of Tables

Table 2-1 Forecasting Accuracy Metrics 20
Table 2-2 Comparison among short-term load forecasting techniques 24
Table 5-1 Configuration of various layers of the proposed model. 50
Table 5-2 Model Training Parameter Setting 51
Table 5-3 Single Step Forecast Comparison of Proposed Method with Various Methods 58
Table 5-4 Multi-Step Forecasting for various look-back and look forward. A comparison of proposed method with the LSTM model. 59
Table 5-5 Comparison of LSTM method, CNN-LSTM method and proposed method with Clustering for each household 64
Table 6 -1 Information about Customers under Study 79
Table 6-2 Parameters for SSA 84
Table 6-3 Parameters for CSA 85
Table 6-4 Appliance Parameters for Simulation 86

Chapter 1

Introduction

Forecasting is a very important function in almost any industry. If we talk about the electric power industry, based on the current technologies, the electricity can't be massively stored. As a result, the electric utilities have to balance the supply and demand. Load forecasting is considered one of the major research fields in power system designing. It is a technique that is used by utilities to balance the supply and demand.

Thanks to the recent advancements in smart grids (SG), the consumers are directly involved in the process of energy consumption. Demand response (DR) is a valuable feature of SGs in which consumers participate in grid management. The load forecasting allows energy users to evaluate their consumption behaviors and, to shift their energy use to off-peak periods, whenever possible.

At the modern time our dependence on the electrical power comes for instance with the new electric vehicles (EVs) and With effective load forecasting, peak load shaving can be achieved through flexibly engaging energy storage (ES) systems or intelligent DR technologies. These electricity suppliers can rely on such information to target the best groups of customers with the highest potential to participate DR programs in the events of power deficiency. In this case, the utilities would be more confident about the demand reduction in incentive-based DR programs, which is important to provide load-balancing reserve and/or hedge market costs.

The electricity consumers are mainly classified into residential, commercial and industrial. Power consumption prediction for residential customers plays an increasingly essential

role in the future power grid planning and operation. According to the annual report of 2011 by the ministry of water and electricity, residential sector constitutes approx. 50% of the total electricity consumption in the Saudi Arabia [1]. Compared to the aggregated power consumption at the community level, forecasting power consumption of a single energy user is fairly challenging due to the high volatility and uncertainty involved. Figure 1.1 shows aggregated daily load profiles of various individual households. As seen in the figure, there is a single major cluster with no outliers. Figure 1.2 shows daily load profiles of an individual household. As seen in the figure, there are two major clusters, two minor clusters and few outliers. It can be inferred from the figures that forecasting an individual household power consumption is challenging due to the high volatility and uncertainty involved compared to the aggregated daily load profiles.

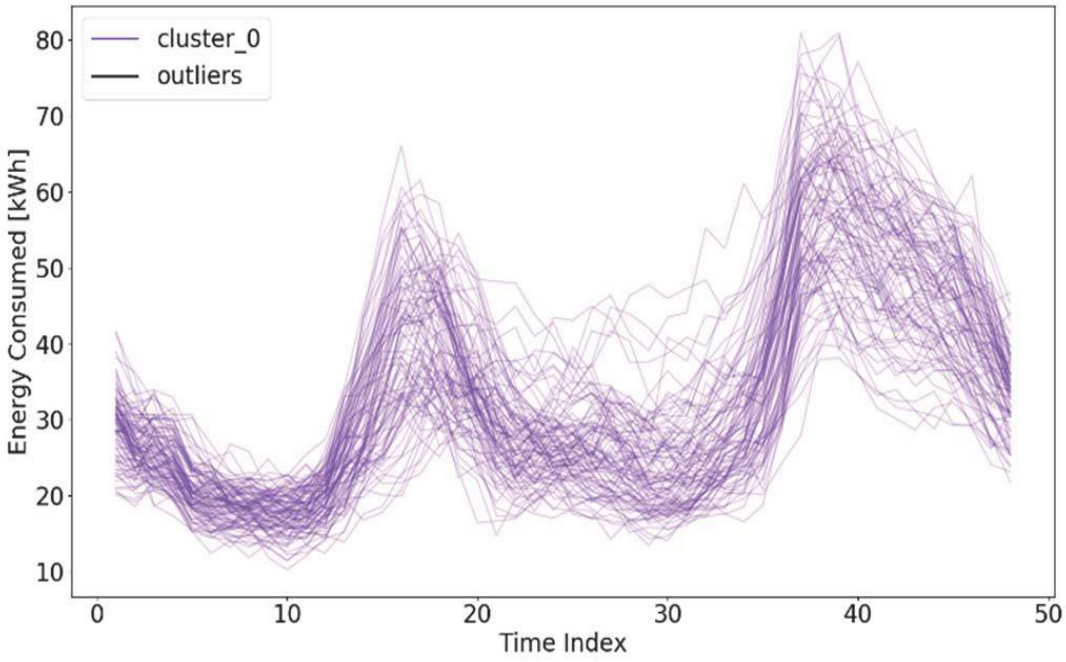

Figure 1.1 Aggregated Daily Load Profiles of various Individual Households

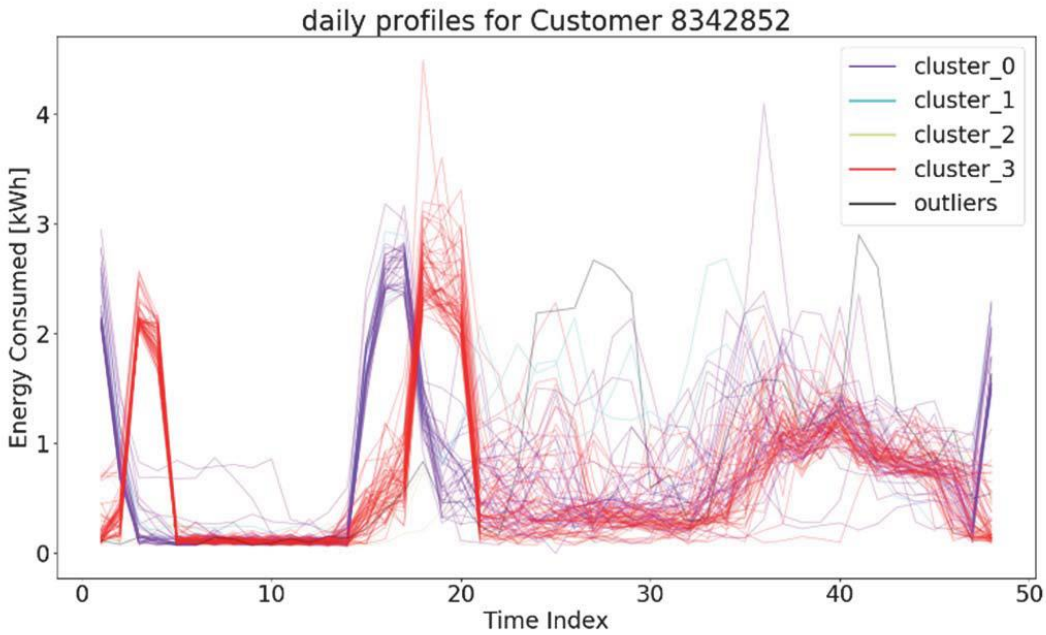

Figure 1.2 Daily Load Profiles of an Individual Household

1.1 Project Objective

The aim of this project is to design a deep learning m method for individual household power forecast of residential customers. A density based clustering technique will be implemented to evaluate the dynamic behavior in the residential power profile. A deep learning method will be proposed to forecast individual residential power consumption.

A breakdown of the tasks and project objectives are listed below for both graduation project classes (CEN492 and CEN493). We completed all the tasks in Phase 1, we will continue working on Phase 2 tasks next semester.

1.1.1 Phase I

- Learn deep learning methods
- Learn Python programming language
- Understand Smart Grid Smart City (SGSC) database
- Use data mining techniques to discover the electricity usage patterns hidden in the data

- Design a deep learning method for power consumption prediction
- Write a report

1.1.2 Phase II

- Implement the designed method using Python.
- Test and evaluate the developed method on SGSC database.
- Write the final report

1.2 Smart Grid Smart City (SGSC) Database

The SGSC project was initiated by the Australian government in 2010. It was the first commercial scale project in Australia. SGSC gathered smart meter data of around ten thousand customers in New South Wale. This database helps to enhance the study of various smart grid technologies. Short-term individual household load forecasting is one of the key research areas that can make use of this database.

The SGSC database is huge and it is quite unrealistic and time consuming to go through each and every customer. For that reason, we choose a subset of SGSC database. The selection criterion for our project is that, we choose those customers that uses a Hot Water System. We end up having 69 customers that can serve as a reasonable data for our project.

1.3 Scope and Report Organization

The report is divided into four chapters. The outline of the report is as follows:

- Chapter 1 introduces the topic, highlights the research problem, discussed about how important the load forecasting for modern time electrical power demand, and what we will try to achieve the requirement and the steps of phase one and phase two.
- Chapter 2 provides the background knowledge about the short-term load forecasting and specifically about individual household load forecasting. This chapter also explores various

machine learning and deep learning methods for short-term load forecasting. Also the related work is discussed in detail.

- Chapter 3 presents the data analysis, visual plotting of the analyzed data, handling missing data and imputation method. This chapter also covers Density-Based Spatial Clustering of Application with Noise (DBSCAN) clustering technique to analyze the individual customer consumption profiles. Also, the data preparation for the deep learning method is discussed in detail.
- Chapter 4 covers the preliminary results of the deep learning method for individual household customers.
- Chapter 5 covers the hybrid deep learning model framework and the forecasting result and the clustering analysis for the individual householders.
- Chapter 6 covers the load scheduling for the individual householders, various trials of Dynamic Peak Pricing (DDP) and the Smart Grid Smart City project(SGSC) and experimental setup and search algorithms.

Chapter 2

Background and Related Work

In this chapter, the background knowledge about load forecasting is given. A comparison between smart grid and traditional grid is presented. Then we discuss about the classification of load forecasting problem. Next, we discuss some of the factors that are important for short-term load forecasting. We present some of the traditional methods as well as machine and deep learning methods for short term load forecasting. In the end, we introduce work related to the individual household load forecasting problem.

2.1 Load Forecasting

In order to ensure the reliable supply and efficient delivery of electricity, there are three main stages: electricity generation, transmission, and distribution. For any electric power supply utility, an accurate planning in all the three stages is very crucial. By applying various forecasting techniques over different forecasting horizons help the utility to correctly plan the resources. With the recent advancements in smart grid technologies, there is a two way communication between utility and a consumer. It has changed the way utility operate and respond to problems.

2.2 Smart Grids vs. Traditional Grids

The smart grid is like an electricity delivery system that originates from the point of generation to the point of consumption. This delivery system is integrated with communications and information technologies for enhanced grid operations, according to the definition by the US Department of Energy [3].

In traditional grid, the consumers have a very little knowledge about their consumption behavior that has a direct impact on the monthly electricity bills. However, with the smart grid, all the electrical devices are connected with the smart meter. The consumers are informed about which device is consuming more energy. Moreover, the consumers may change their consumption behavior to reduce the electricity bills.

2.3 Classification of Load Forecasting Problems

For any utility company, forecast of different time horizons is crucial for different operations. Load forecasting problems are mainly classified into three categories: short-term, medium-term and long-term forecast. The short-term load forecast (STLF) is usually from 1-hour to 1-week ahead. The medium-term forecast is from 1-week to 1-year ahead. The long-term forecast is from 1-year and above.

STLF for individual residential customers plays an important role in the future smart grid operation and planning. Forecasting power consumption of an individual customer is quite challenging because of uncertainty and volatility involved. In this project, we investigate the power consumption of various individual household customers for short-term time horizon.

2.4 Important Factors for Short-Term Load Forecasting

There are several factors that needs to be considered for STLF. One of them is time factor that includes month of the year, week of the month, day of the week, hour of the day, weekday vs. weekend, holidays, and season. Another important factor is weather data. In fact forecasting weather parameters is very important in short-term load forecasting. The two most commonly used weather parameters are temperature and humidity. The time of use pricing is another promising factor that may affects the daily load profiles of consumers. The reason for this is that in some countries, the electricity tariff is different at different time of the day. Thus, consumers may shift there load based on time of use pricing.

2.5 Forecasting Accuracy Metrics

In the literature, several methods exist to measure the forecasting accuracy. The most common metrics are enlisted in Table 2.1. Mean Absolute Percentage Error (MAPE) is the most popular metrics to measure the forecast accuracy. The main problem with MAPE is that for an actual value very close to zero, the error will be very high because of actual value in the denominator. Some researchers are using a modified version of MAPE to tackle this problem.

Table 2-1 Forecasting Accuracy Metrics

Metrics	**Formula**
Mean Squared Error (MSE)	$\frac{1}{n} \cdot \sum_{i=1}^{n} (A_i - F_i)^2$

Root Mean Squared Error (RMSE)	$\sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (A_i - F_i)^2}$
Mean Absolute Deviation (MAD)	$\sum_{i=1}^{n} \lvert A_i - F_i \rvert$
Mean Absolute Percentage Error (MAPE)	$\frac{100}{n} \cdot \sum_{i=1}^{n} \left\lvert \frac{A_i - F_i}{A_i} \right\rvert$

2.6 Short Term Load Forecasting Techniques

In the literature, the researchers have proposed / employed several techniques starting from statistical methods to machine / deep learning methods.

2.6.1 Persistence Model

The simplest but powerful forecasting model is persistence model. In this model, it is assumed that the conditions remained unchanged and hence the value at the future forecasting time step is the same as the actual value at the current time stamp. For STLF, usually it is difficult to beat the persistence model because for short-time it is assumed that the load remains unchanged [4].

2.6.2 Regression Methods

In the early stages of STLF, regression analysis were used. A regression analysis involves some parameters along with a relationship between dependent and independent variable to estimate the forecast. The simplest regression model is linear regression. In [5], authors used linear regression to forecast short-term load in non-residential building.

An extension of linear regression is multiple regression model that involves multiple independent variables. In the past, some studies involve multiple regression model for aggregated STLF [6]. However, to the best of our knowledge, regression methods have not been exploited for the individual residential consumers load forecast.

2.6.3 Time Series Methods

The most used method in time series forecasting method is Auto Regressive Integrated Moving Average (ARIMA). An ARIMA method has been extensively used by researchers to forecast the electricity load [7]. ARIMA models is an extension of Auto Regressive Moving Average (ARMA). ARMA models are used for fixed processes while ARIMA is usually used for non-fixed processes. ARIMA uses the load and time as the input parameters while the load depends on the weather and the time of the day.

ARIMAX is a variant of ARIMA and can be viewed as multiple regression model. It is the most natural method for load forecasting between all classical time series model. In comparing methods to forecast energy up to one day ahead [8] find the result an extended Holt-Winters exponential smoothing method is better than ARIMA in terms of forecast accuracy and neural networks and regression with principal component analysis.

2.6.4 Machine Learning Methods

2.6.4.1 Fuzzy logic

In general, Fuzzy logic uses Boolean logic for digital circuit design, it takes a truth value of 0 or 1 as input under a Boolean logic. Certain qualitative ranges is associated

with an input under Fuzzy logic. For instance transformer load may be high or medium or low. Fuzzy logic inferred outputs from fuzzy inputs that is why Fuzzy logic consider the one of a number of techniques for mapping inputs to outputs.

In [9], the authors proposed fuzzy based rules for peak load forecasting. They achieved comparatively good accuracy.

2.6.4.2 Artificial Neural Network (ANN)

The Artificial Neural Network (ANN) is the depiction of how human brain works. After the development of ANN, it rapidly became one of the most promising forecasting tool. ANN has been extensively explored for electric load forecasting [10][11]. Backpropagation is the most famous ANN architecture. Hong Chen et. Al. in [12] used a three layer ANN model to forecast short-term electric load.

2.6.4.3 Support Vector Regression (SVR)

Support vector regression are the best for solving classification and regression problems. This approach was originated from Vapnik's statistical learning theory. And it's not the same of neural networks, which try to determine a complex function of the input.

Support vector machines produce a nonlinear mapping of the data into a high dimensional space by using so-called kernel function. After that support vector machines use a simple linear function for creating a linear decision boundaries. In [13], the authors used SVR and achieved a good accuracy for electric load forecasting.

Table 2.2 shows a comparison among short term load forecasting techniques.

Table 2-2 Comparison among short-term load forecasting techniques

	Pros	***Cons***
Multiple Regression	Good accuracy , easy to implements	Relies on explanatory variables needs explanatory variables
ARIMA	Good accuracy in short term, Few parameters to estimate	Low accuracy in longer term , high cost to implement
Artificial Neural Network	Good accuracy during normal days	Heavy computation, difficult to interpret

2.6.5 Deep learning Methods

In the previous section, some of the traditional methods were discussed. These traditional methods achieve good results only if there is a linear relationship between electric load and other parameters. Advanced methods such as machine learning methods show promising results compared to traditional methods. However, with the advancements in computational technology, deep learning methods are outperforming machine learning methods because deep learning model rely less on feature engineering. Moreover, due to the availability of large data, and GPUs, the deep learning model can make a complex relationship among various layers within the network with higher computational efficiency. In this section various deep learning methods are briefly described.

2.6.5.1 Convolutional Neural Networks (CNN)

CNN is one of the variants of neural networks, this method is used a lot in computer vision. A CNN model has several layers such as convolution, pooling, normalization, non-linearity, and fully connected layers as shown in Figure 2.1.

The convolution layer consists of set of filters. These filters are convolved with the input to generate the feature map. In order to introduce the non-linearity, rectified linear unit (ReLU) is used. In CNN, pooling layer is used to reduce the dimensionality of the feature maps which in turn reduces the number of learnable parameters. Max pooling is one of the commonly used pooling technique in CNN. A normalization layer normalizes the data after the activation function to reduce the computational load of the network. At the end of the CNN, a fully connected layer is used. Each node in FC layer is connected with each node of the previous layer.

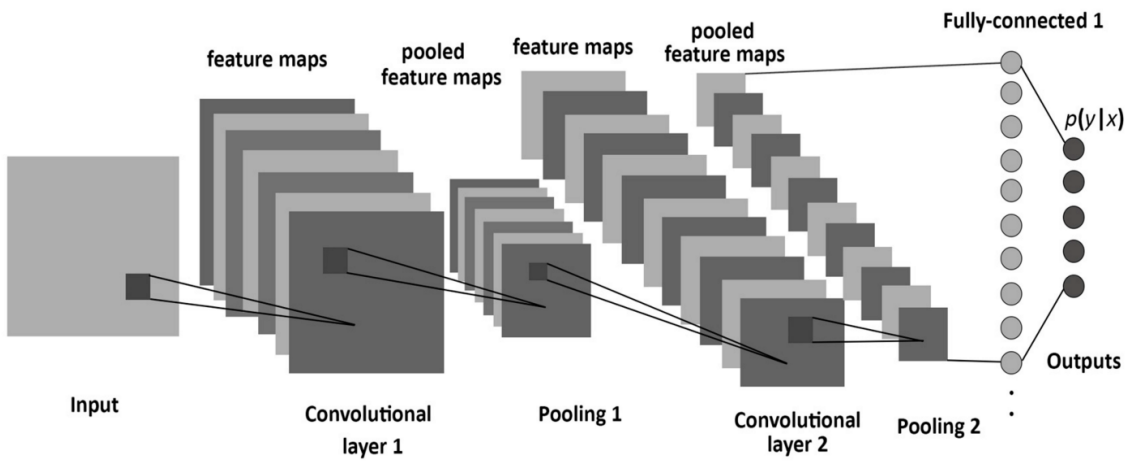

Figure 2.1 Convolutional Neural Network Architecture

In [14], authors proposed a CNN-based STLF for smart grid. They transformed the numerical forecasting problem into an image based problem. They considered various factors such as wind speed, humidity, and temperature to predict the electric load.

2.6.5.2 Recurrent Neural Network (RNN)

RNN is a generalization of feed forward neural network, and it has an internal memory. RNN performs the same function for every input of data, and the output of the current input depends on past one. After the output is coming up, it is copied and

sent it to the recurrent network for making a decision. It considers the current output and the input that it has learned from the last input. In [15], the authors performed a comparative study by applying RNN for short-term load forecasting. The main problem associated with RNN is the vanishing gradient problem. It is quite challenging for RNNs to process very long sequences.

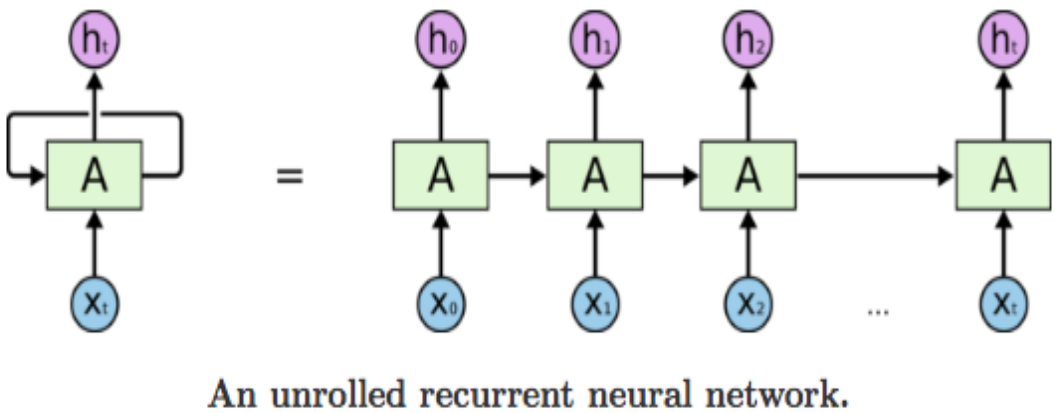

Figure 2.2 Basic Building Blocks of Recurrent Neural Network (RNN)

2.6.5.3 Long Short-Term Memory (LSTM)

LSTMs are designed to avoid the long-term dependency problem and it's a special kind of RNN that got lot of attention recently within the machine learning community.

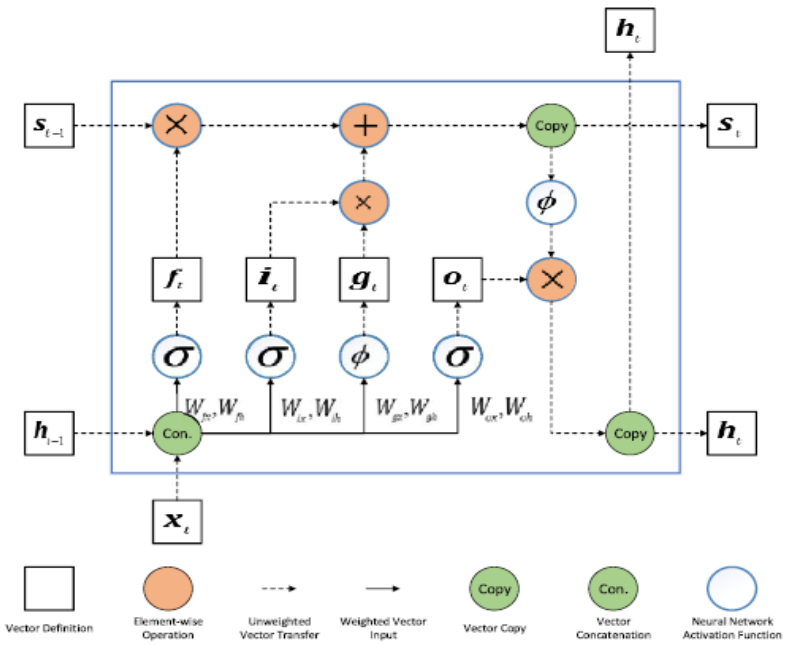

Figure 2.3 Basic Building Blocks of LSTM [2]

LSTM network act as long-term or short-term memory cells and that because of internal contextual state cells that it have.

We modulate the output of the LSTM by the state of these cells. And we need the forecasting of the neural network to depend on the context of input history not only on the very last input. In [2] authors proposed a novel LSTM based short-term load forecasting method for residential customers.

2.7 Related Work

In the past, many researches talk about the load forecasting, for example, Li et al. [16] talk about STLF based on the grid method and the time series fuzzy load forecasting method. Cao et al. [17] talk about ARIMA model and similar day method for intraday load forecasting.

Qingle and Min [18] proposed a neural network based predictor for very short term load forecasting. It takes the load values of previous and the current time steps as input to predict the load value. Zhang et al. [19] used extreme learning machines (ELMs) to predict the total load of the Australian national energy market.

Chaouch's [20] and Ghofrani et al. [21] are the first examples on individual residential customers electric load forecasting. In [20], the authors proposed a functional time series forecasting approach. The more commonly used metric of MAPE was not reported but the daily MAE were reported. The authors in [21] used Kalman filter based approach.

Ryu et al. [22] proposed that we can improve the load forecasting accuracy of industrial consumers by using deep learning based methods.

Mocanu et al. [23] used a conditional restricted Boltzmann machine to predict an individual customer load forecasting. The authors achieved better results compared to the performance of support vector machine and ANN. Marino et al. [23] made the first try for the same load forecasting issue by using LSTM and they compare the results with [23].

In [2] the authors proposed a novel LSTM based load forecasting method for individual household as well as aggregated load. The compared their results with the traditional methods as well as machine learning methods. Their results indicate that the LSTM outperforms all the other rival techniques. They achieved an average MAPE of around 45% for individual household residential customers.

2.8 Conclusion

In this chapter, we provided an overview of short-term load forecasting techniques. We did a comparison in table that shows the pros and cons of various traditional methods in terms of computational cost and accuracy. We then discussed the mechanisms of deep learning methods like the CNN, RNN and LSTM.

Moreover, work related to the aggregated and individual household load forecasting is discussed in detail. Because of the volatile nature of daily load profiles, it is very challenging to accurately forecast the individual household electric load. In the STLF, researchers have targeted mostly for aggregated load forecasting. Very few researches are available that perform forecast at an individual household level. Still there is a lot of room for improvement in terms of MAPE.

Chapter 3

Data Analysis and Preparation

Before developing a deep learning framework for the short-term load forecasting, it is important to visualize the data and perform some exploratory analysis. For the underlying project, we will use SGSC data base. Next, we will discuss about handling data anomalies. We apply DBSCAN clustering technique to analyze the daily load profiles of various individual customers. We study the outliers of various customers load profiles. In the end of this chapter, we prepare the data for the deep learning model.

3.1 SGSC Database

The SGSC project is an Australian project that started in 2010 in New South Wale, with the goal of collecting smart meter data for thousands of residential consumers that helps study various smart grid technologies and predicting short-term load is one of the areas that can benefit from its database. The collected data was sampled at 30 minutes interval with local time stamp.

SGSC is huge database and it is unrealistic to conduct study on all residential consumers, so we will choose a subset of a SGSC database. It is the group of customers who use the hot water system, we will reach 69 clients that can serve as data that can be used in our project.

3.2 Impact of Customer Behavior on Load Profiles

The individual daily load profiles do not have a consistent pattern because of customers' behavior and thus the forecast accuracy is compromised. On the contrary, the aggregated load profile is relatively more predictable because of smooth curve. For residential customers, the lifestyle of customer and the type of major appliances in the household play vital role in the accuracy of short-term load forecast. Some customers have a fixed routine to turn on certain appliances. For example, after using a washing machine they will turn on the clothe dryer which means that electricity consumption in the subsequent hours will be more.

Life patterns and routines have a direct impact on the load profile in the short term. When the customer has a regular life behavior, we will be able to predict his energy use, For example, the customer may have fixed procedures, such as turning on the TV set after 7:00 pm or turning the washing machine at 7:00 am.

3.3 Fill Missing Data

Each of the 69 customers have different length of available data. For all the 69 customers the data of three months starting from 1st June 2013 till 31st August 2013 is available [2]. We consider this three month of data for our study. For some customers data at certain time stamp is missing.

The data that is passed to the deep learning model needs to be complete. If the data is missing then the forecasting method will not work. There are several techniques that can be used to impute the missing data.

For the period that we consider we found some missing reading values so we implement a function to help us find which user have a missing data at what time stamp and then we find the missing data and fill them with the same values from one day earlier. For example, if the value of load is missing for 2nd July 2013 at 09:00 am, then we will use value of load for 1st July 2013 at 09:00 am. In case this data was also missing then we will consider the data from day before.

3.4 Evaluation of Daily Load Profiles

The load profile of individual residential customer is highly volatile due to the customers' behavior and many other factors such as number of electrical appliances, types of electrical appliances, number of residents, weather, and so on. Some customers may have a lifestyle that is more regular, thus it is easy to forecast the electrical load of such customers. In this section we will visualize the daily load profiles of some customers as well as explore the outliers, major/minor clusters within the daily load profiles.

To analyze the daily load profiles of individual residents we apply density based clustering technique i.e. Density-Based Spatial Clustering of Application with Noise (DBSCAN). It

is a well-known clustering technique in machine learning. One obvious advantage of DBSCAN over other clustering technique is it doesn't require number of clusters a prior. Moreover, it can identify outliers in the data.

DBSCAN uses Euclidean distance approach to group set of points that are close to each other. Those points that are in the low density region are marked as outliers. DBSCAN requires two parameters; epsilon and minPoints. The value of epsilon specifies how close the points should be in order to be considered within the cluster. That means if the Euclidean distance between two points is less than or equal to epsilon then the points will be considered within one cluster. The number of minimum points to make the dense region serve as minPoints, the second parameter of DBSCAN.

In this project, we are using residential data of 69 customers from the SGSC database. The data is collected at an interval of 30 minutes. It means daily we have 48 samples that will serve as 48 dimension for DBSCAN. The value of epsilon is 10% of mean daily electric load, whereas minPoints is set 2.

Figure 3.1 shows the daily load profile of a customer (customer ID 8523058). As seen in the figure, there is one major cluster with no outliers. This customer has a regular behavior hence we may forecast the electric load of such customer with better accuracy.

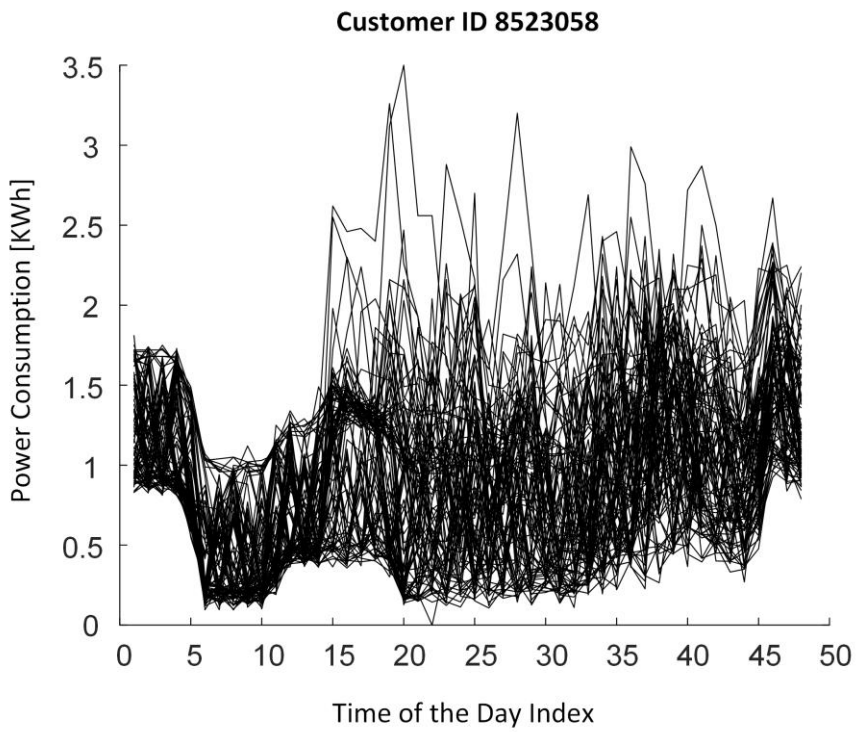

Figure 3.1 Daily Load Profiles of a residential customer (ID 8523058) – Single Major Cluster

Figure 3.2 shows the daily load profile of a customer (customer ID 11462018). As seen in the figure, there is one major cluster, one minor cluster and few outliers. This customer has somewhat regular behavior with few irregular days. Figure 3.3 shows another customers (customer ID 8273592) daily load profile. As seen in the figure, there are two major cluster and several outliers. It is quite challenging to predict the load of such customer. Figure 3.4 shows customers (customer ID 9012348) daily load profile. The customer do not have any major cluster. As seen in the figure, the customer has four minor cluster and several outliers. This customer has an irregular life behavior and it is very challenging to predict the electric load of such customers.

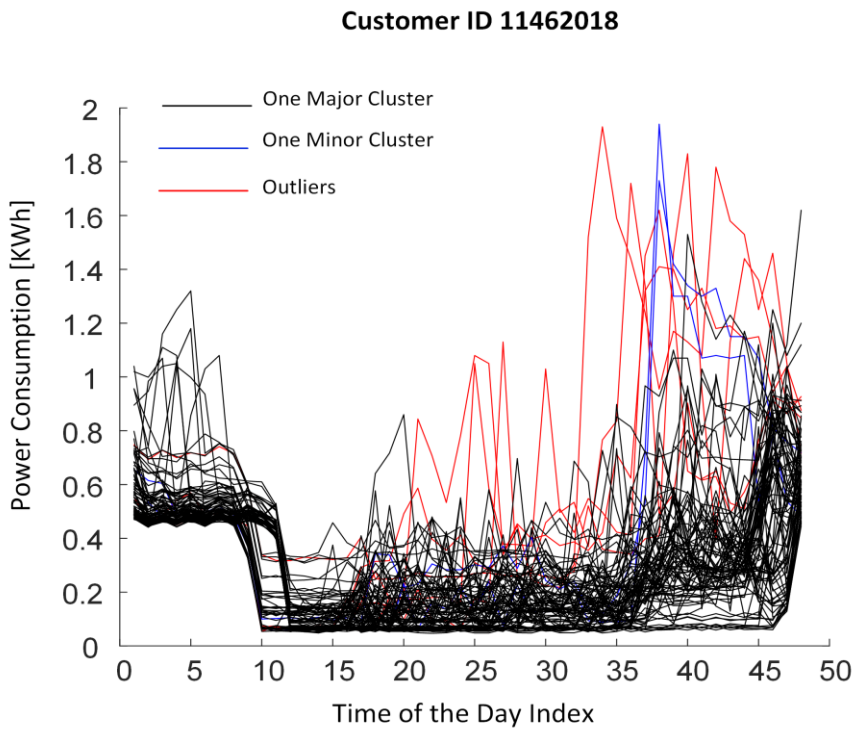

Figure 3.2 Daily Load Profiles of a residential customer (ID 11462018) – One Major Cluster, One Minor Cluster and Few Outliers

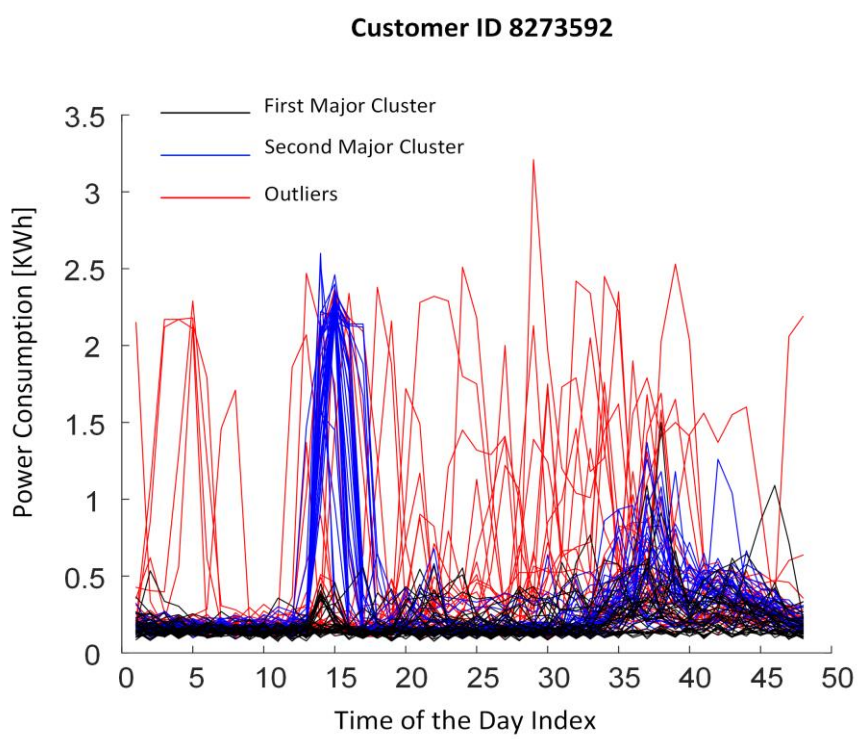

Figure 3.3 Daily Load Profiles of a residential customer (ID 8273592) – Two Major Clusters and Several Outliers

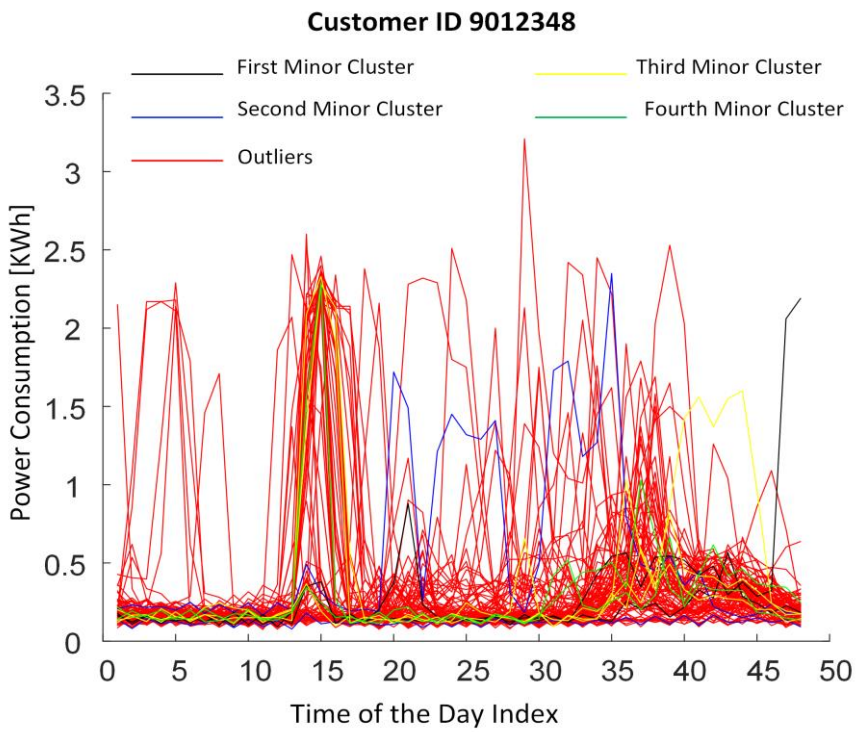

Figure 3.4 Daily Load Profiles of a residential customer (ID 9012348) – Four Minor Clusters and Several Outliers

Figure 3.5 shows the number of major cluster, minor clusters and number of outliers for some of the randomly selected customers out of 69 customers from SGSC database.

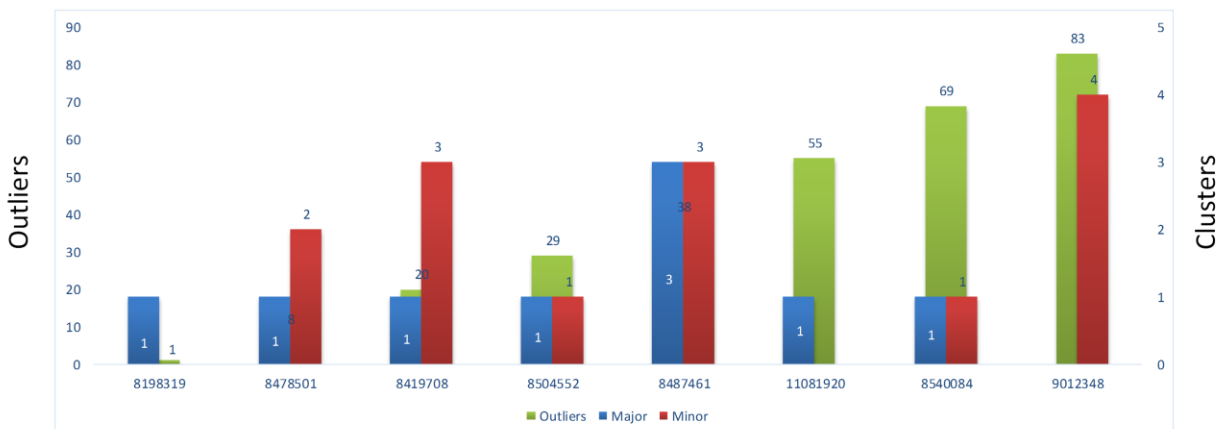

Figure 3.5 Number of Major Clusters / Minor Clusters and Outliers for few randomly selected households from SGSC database

3.5 Data Preparation for Deep Learning Model

In this section we discuss about some of the features that are considered for the preliminary implementation of deep learning LSTM based model.

For 69 customers under study, the data available is from June to August hence we are not considering the seasonal impact as the data only for winter season. Moreover, we are not considering weather parameters such as temperature, pressure, and humidity as they are not available for the location under study (New South Wales Australia).

We consider the following features to prepare the data,

1. Past electric load of J time steps (i.e. $J = 48$ means electric load of past one day will be considered as features)
2. Hour of the day time indices (load is sampled at 30 minutes interval. i.e. for complete day there will be 48 samples. Half-hourly increment will be from 1 to 48)
3. Day of the week indices (increment will be from 0 to 6)
4. Weekend or Weekday(1 / 0)

3.6 Conclusion

In this chapter we present some details about the SGSC database. Then we discuss the impact of customer behavior on the daily load profiles. We applied DBSCAN clustering technique to analyze the daily profiles of various customers. We have seen that some customers have regular behavior thus easy to forecast electric load whereas some customers have several outliers in the daily profiles making the forecast quite challenging. In the end of this chapter, we present the set of features that will be used for the preliminary implementation of deep learning LSTM based model.

Chapter 4

LSTM based Deep Learning Framework

In the first term of this project, we have implemented basic LSTM model proposed by [2] to forecast electric load of individual residential households. In this term, we get an overview about how the prepared data (Section 3.5) is passed to the deep learning method.

4.1 Data Split and Test Settings

We have three months of data starting from 1st June 2013 till 31st August 2013 for all the 69 customers. The data consists of 92 days for each customer. We split the data in training/validation/test as 70/20/10. The deep learning model for each customer was trained with the data starting from 1st June 2013 till 05th August 2013. Each model was validated on the data starting from 06th August 2013 till 22nd August 2013. Finally the developed model was tested for each customer on the respective test data starting from 23rd August 2013 till 31st August 2013.

In the preliminary implementation, hyper parameter optimization is ignored. We used two hidden layers in LSTM network each layer has 20 nodes. We evaluate the model for 150 epochs.

4.2 Preliminary Results and Discussion

In this section we have presented the forecasting accuracy for few randomly selected customers out of 69 customers from SGSC database using LSTM network. We

consider a look back of 48 samples, i.e. 1 day and a look forward of 2 samples, i.e. 1 hour.

Figure 4.1 shows the load forecast of customer ID (8523058). We have achieved a MAPE of 47.01% using LSTM network. As seen in the figure, the forecast is trying to follow the peak.

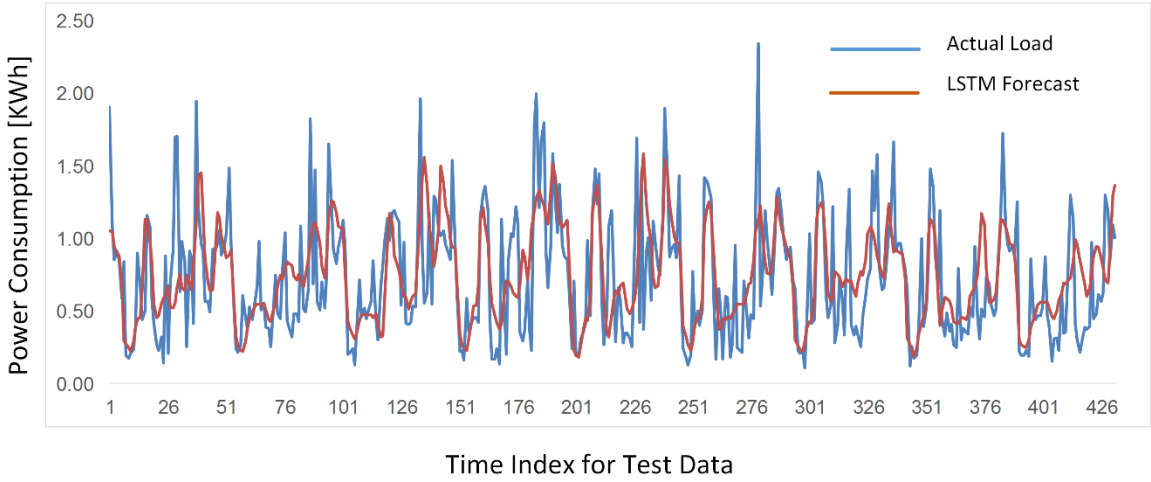

Figure 4.1 Actual Load vs. LSTM forecast for customer ID 8523058

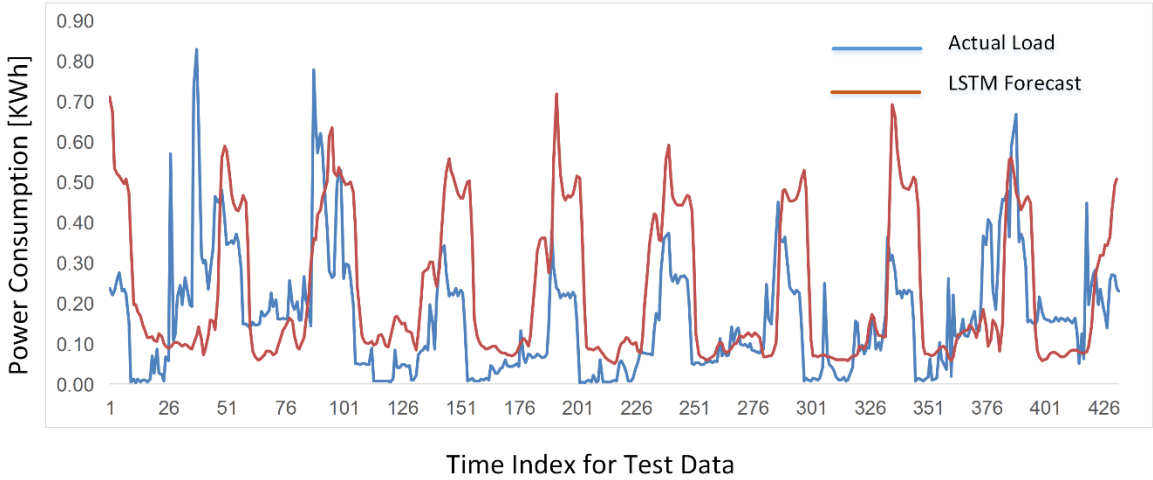

Figure 4.2 Actual Load vs. LSTM forecast for customer ID 11462018

Figure 4.2 shows the load forecast of customer ID (11462018). We have achieved a MAPE of 31.47%. For this customer, the actual power consumption for the test data is below 1kWh in the complete test data. However, in the training phase, the average load is above 1kWh, hence the model forecast is usually higher throughout the test data. As seen in the figure, there are some noise peaks detected by the LSTM model.

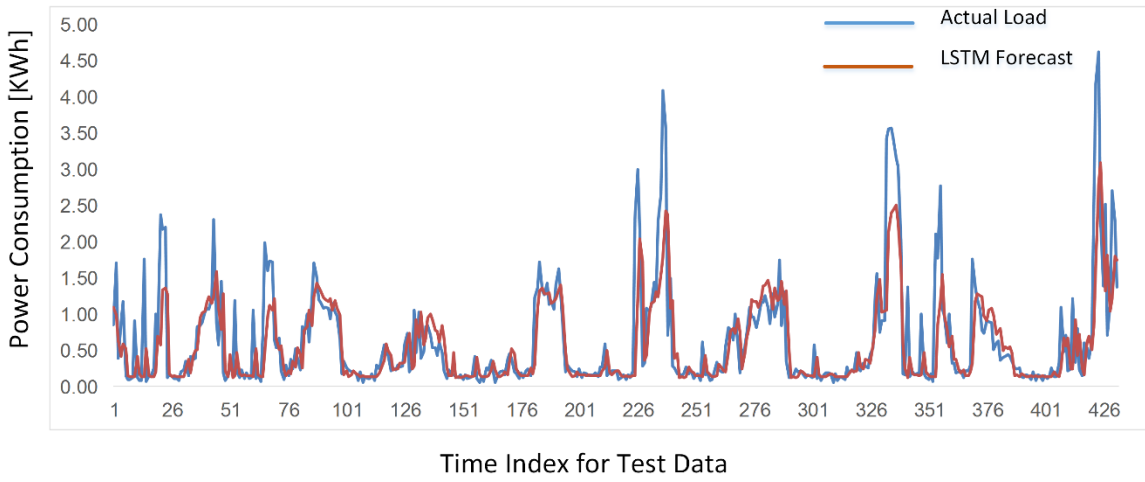

Figure 4.3 Actual Load vs. LSTM forecast for customer ID 8432046

Figure 4.3 shows the load forecast of customer ID (8432046). We have achieved a MAPE of 58.42%. The MAPE of this customer is slightly higher because the forecast is struggling to follow the peaks, whereas the peak power consumption (actual load) is around 4kWh at certain time stamps.

4.3 Future Work for Second Term

In this chapter, we have investigated the LSTM model for individual household electric load forecasting. We concluded that there is still a lot of room for improvement in terms of accurate forecasting.

In the next term, we will further investigate the feature selection and how it will improve the forecasting accuracy. Moreover, we will perform the clustering of individual

households based on average power consumption and group the customers with somewhat similar behavior. Additionally, we will propose and develop a hybrid deep learning model to forecast the electric load of individual household. We will evaluate the proposed model on SGSC database with and without customer clustering.

Chapter 5

Hybrid CNN-LSTM based Deep Learning Framework

The residential load forecasting at an individual level plays an important role in operation and planning of future grids. The aggregated power consumption prediction at the community level is not challenging. On the contrary, the individual load forecasting is quite challenging because of the high uncertainty and volatility. In the past, various studies have been conducted in improving the forecasting accuracy of individual household electric load using various machine learning and deep learning models. However, the results obtained are far from accurate.

In the first term of this project, we have implemented LSTM model proposed by [2] to forecast electric load of individual residential households. The LSTM model outperforms several machine learning and deep learning models for individual household load forecasting.

In this term, we have proposed a hybrid CNN-LSTM model for individual household load forecasting. The idea of utilizing the CNN as well as LSTM layers in developing a hybrid model is that the CNN layers extract the features from the input training data, whereas the LSTM layers will help in finding the long-term dependencies in the training data.

The CNN layers are helpful in extracting the valuable features and finding the local trend as well as pattern in the time-series data. However, a deep learning model that

is based on only CNN layers will not be much helpful because CNN layers are not adapted for long-term dependency problems.

The LSTM networks are extensively used in time-series forecasting problems because of their ability to capture long-term dependencies. However, they only utilize the features and attributes provided in the training dataset. It is very important to carefully select the attributes in the training dataset for the LSTM model to achieve a good forecasting accuracy.

We believe that a hybrid model that utilizes the benefits of both the LSTM layers as well as CNN layers could improve the accuracy of individual household load forecasting.

5.1 Various Layers used in Proposed Model

The most commonly used machine learning and deep learning models for time-series problems are based on the CNNs and the LSTMs. In this project, we have developed a deep learning model that utilizes the CNN layers at the input stage to extract the valuable features as well as learn the internal representation of the training data and LSTM layers in the later stage to find the short-term as well as long-term dependencies. We evaluated the proposed model on SGSC database discussed in Chapter 3.

We first briefly elaborate on various layers used in our deep learning framework, including 1-D convolution, MaxPooling, ReLU, dropout, LSTM, and fully connected.

5.1.1 1-D Convolution Layer

The convolutional layer is the highest significant layer of the CNN based feature extraction block. It is considered as a set of learnable filters which contains

convolution operations. The backpropagation algorithm is applied for enhancing the parameters of the convolution. The receptive field in all of the filters of the convolution layer is the same.

This layer performs the convolution operation. It is a linear operation in which the filter / kernel is multiplied (dot product) with the input. A convolution is an operation where filter is applied on the input data which results in an activation. When the same filter is repeatedly applied on the input data it results in feature map.

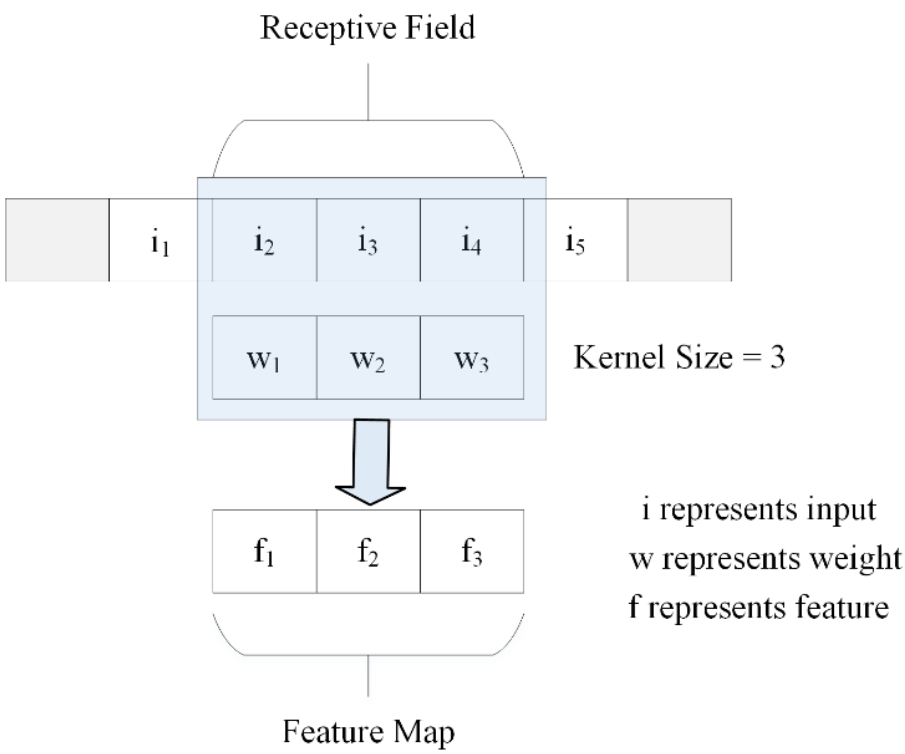

Figure 5.1 Operational example of 1-D convolution.

The operational example of 1D convolution is presented in Figure *5.1*. The weights linked with the kernel are of size 3, i.e., $\{w_1, w_2, w_3\}$, which are shared by the input layer $\{i_1, i_2, i_3, i_4, i_5\}$. The feature map is attained by performing the convolution operation between the inputs and the weights. Specifically, in the example described in Figure *5.1*, the feature f_2 is gained, as shown in (1).

$$f_2 = w_1 * i_2 + w_2 * i_3 + w_3 * i_4 \quad (1)$$

CNNs do not learn a single filter rather multiple filters are used to learn multiple features and attributes in the training data. The convolution layers are not only applied to the input layer rather multiple convolution layers are stacked in the deep learning model to learn low-level as well as high-level features.

5.1.2 MaxPooling Layer

The convolutional layers in a deep learning framework methodically apply the learned filters to the applied input for producing feature maps, which encapsulate the existence of these features in the given input. The convolutional operation is highly effective and piling several convolutional layers in a deep learning framework enable the initial layers (layers close to the input) to learn low-level features in the applied input. The feature map, which is the output of the convolutional layers has a limitation that it keeps track of the precise location of the features in the input. It means that little movements in the location of the feature in the input will lead to a different feature map.

A pooling layer is usually added after the convolutional layer for mitigating the limitation of invariance of the produced feature map, and adding a pooling layer immediately after convolutional layer is a common practice which is applied to order the layers in a deep learning framework. The pooling layer works on each feature map distinctly to produce a new set of pooled feature maps. Typically, the size of the pooling operation is less than that of the feature map. The addition of a pooling layer

in a deep learning framework results in a pooled feature map, which is a summarized version of the detected features.

In our developed deep learning framework, we have added a MaxPooling layer. The MaxPooling is a down-sampling scheme in deep learning. In MaxPooling, a moving window is applied across the given input, where the output is the maximum value in the input stream, which overlaps with the window. Figure *5.2* shows an example of the MaxPooling operation, which is applied to a 1D input stream.

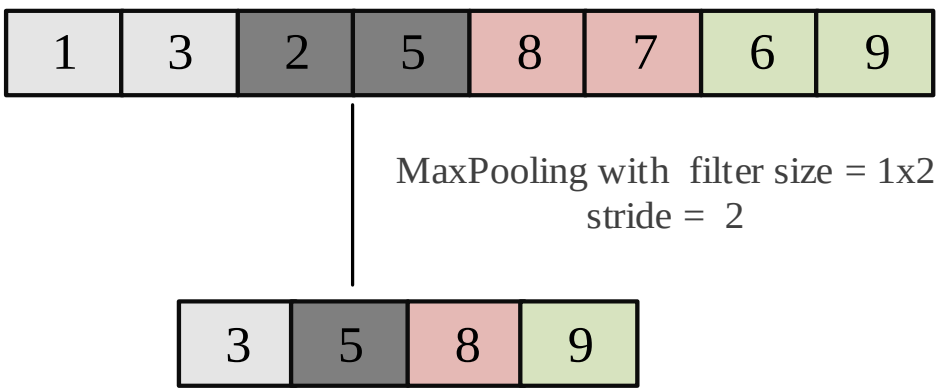

Figure 5.2 Operational example of 1-D MaxPooling.

5.1.3 Rectified Linear Unit (ReLU)

The activation function is applied for enhancing the capability of the model for learning complex structures. The ReLU activation function has been widely implemented by various researchers in their deep learning models to make the network more trainable. It applies the thresholding operation, which is shown in (2).

$$f(z) = \max(0, z) \tag{2}$$

5.1.4 Recurrent Neural Network (RNN) and Long-Short Term Memory (LSTM)

A RNN has shown a strong capability for capturing the hidden correlations in the input data stream. RNNs have been applied in different applications, including time-

series prediction, speech recognition, and natural language processing. RNNs are mainly suitable for solving sequence problems as they work on both the previously acquired information as well as the current input stream. Figure *5.3* shows the mapping from the input node to the output node. The input and outputs are related through (3) and (4).

$$A_t = f(U * X_t + W * S_{t-1}) \tag{3}$$

$$O_t = g(V * S_t) \tag{4}$$

where boxes S's represent memory in the network at different time instances. Letters U symbolize the weight matrices in various layers, V and W. The Ot represents the output while the Xt indicates the input at any time instant t. The f (.) and g(.) indicate the transformation function from the input to the output.

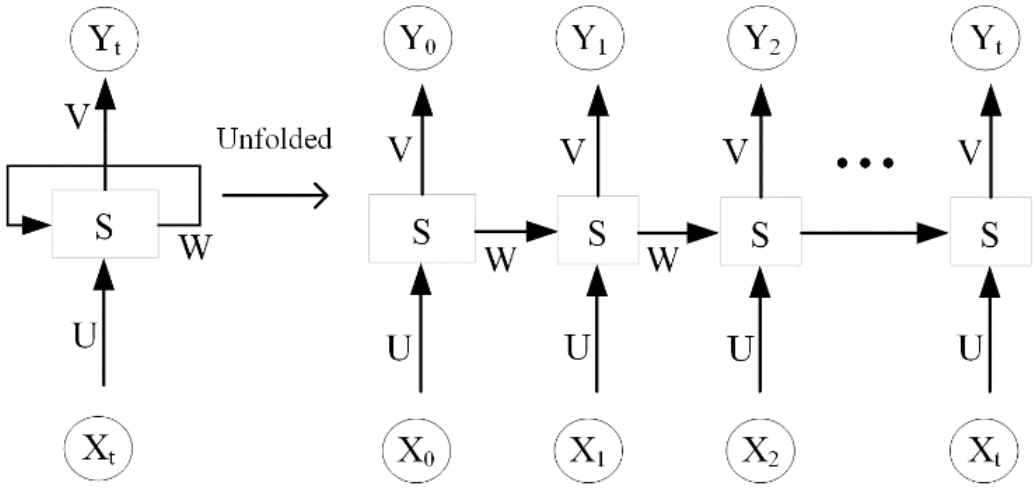

Figure 5.3 Basic structure of RNNs.

It must be noted that the information collected at various instances of time are stored in the hidden state while the output is determined using the current input and values from the previously stored values in memory. The RNN performs excellent for output, which is closely associated with the inputs. The reason is that the previously

stored information is passed to the next node. Theoretically, the RNN is capable of dealing with long dependencies. Nonetheless, for practical applications, the RNN is not capable of storing all the past data for a long time interval due to the gradient vanishing problem. For solving this issue and enhancing the performance of the RNN for practical applications, a variant of RNN called LSTM was explored.

To solve the issues of gradient vanishing of RNN, the LSTM connects short and long-term memory by using a gate control system, as shown in Figure *5.4*. It can be seen that a unit of LSTM is composed of an input gate, a memory cell, a forget gate, and an output gate. The input X_t, at any time instance t, is selectively stored in the memory cell S_t, which is controlled by the input gate. The forget gate can be used for retaining or forgetting the previous values of the memory cell S_{t-1}. Lastly, the output gate determines which value of the memory cell S_t to be added to the output h_t.

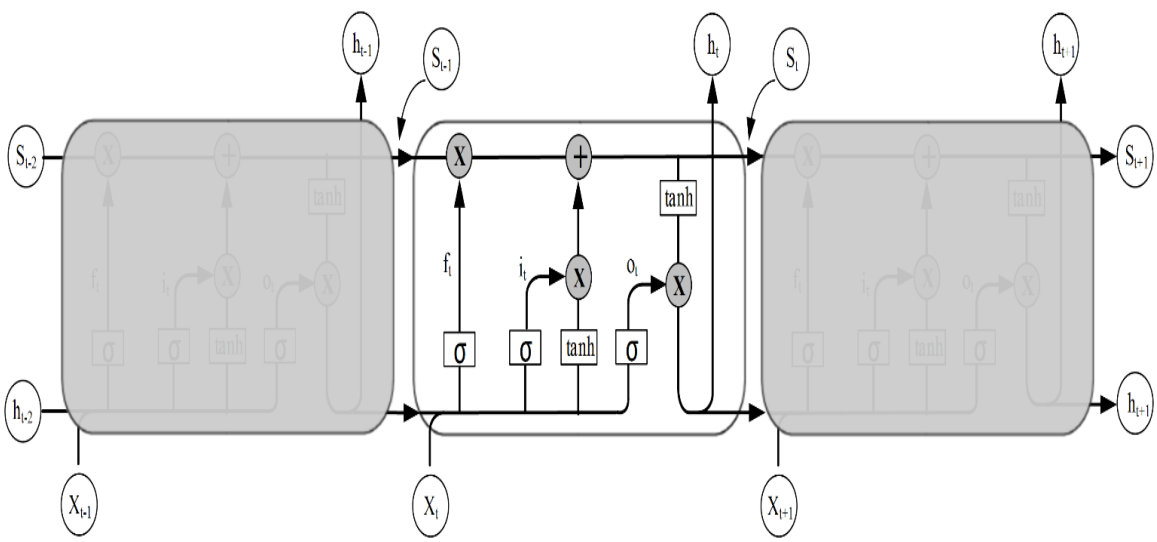

Figure 5.4 Inner Structure of LSTM

5.1.5 Dropout Layer

In the development of any deep learning model, the dropout layer offers a cool way to relieve the overfitting issue. This layer includes the random selection of neurons

and deactivating some of them in the training process. The output of the randomly deactivated neurons is 0. Figure *5.5* shows an example of dropout layer.

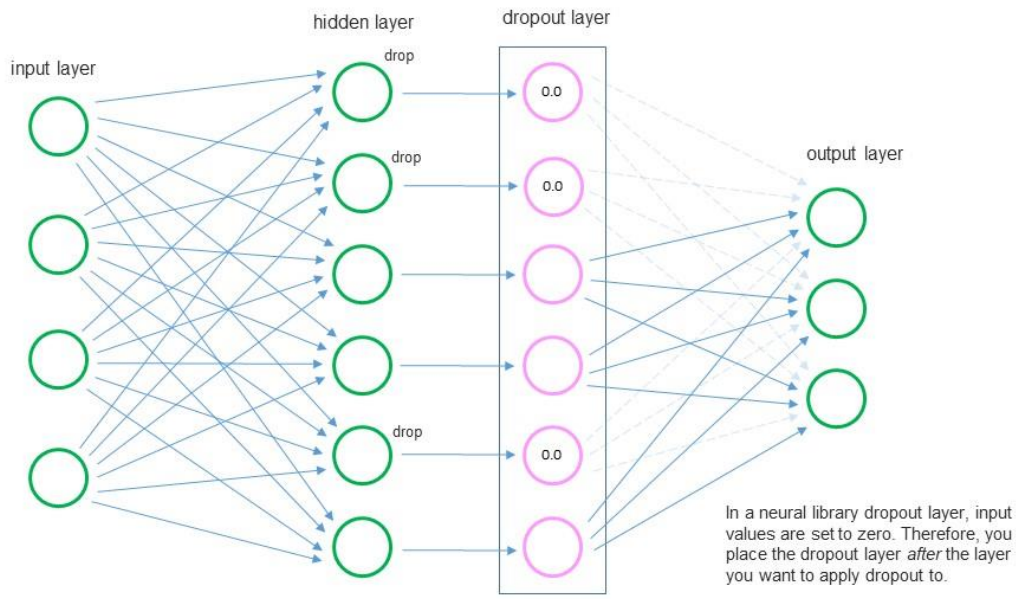

Figure 5.5 Example of Dropout Layer

5.1.6 Fully Connected Layer

The activation function and the bias are used by the fully connected layer for nonlinear mapping from the input to the output. The fully connected layers are typically applied at the end of the deep learning model.

5.2 Proposed Hybrid CNN-LSTM Model

The main architecture of the proposed model is shown in Figure *5.6*. It consists of two building blocks; first, the feature mining block and second, the sequence learning block. The feature mining block consists of four 1D CNN layers. The MaxPooling layer and ReLU layers are incorporated in-between two consecutive convolution layers. When the convolution operation is performed by piling several convolution layers then low-level features are learned by the initial layers. The output of the convolution

layer is a feature map which keeps track of the accurate location of the feature in the input data. A pooling layer is generally incorporated after the convolution operation for mitigating the constraint of feature map invariance whereas activation function is added for improving the model capability to learn complex structures and to introduce non-linearity. We have used MaxPooling layer which down sample the feature map by 2 thus reducing the computation complexity of the model. For the activation function, we have used ReLU which is widely used function.

The dropout layer is very important in deep learning model development as it help solve the over-fitting problem. We have used dropout layer after the feature extraction block as well as after sequence learning block.

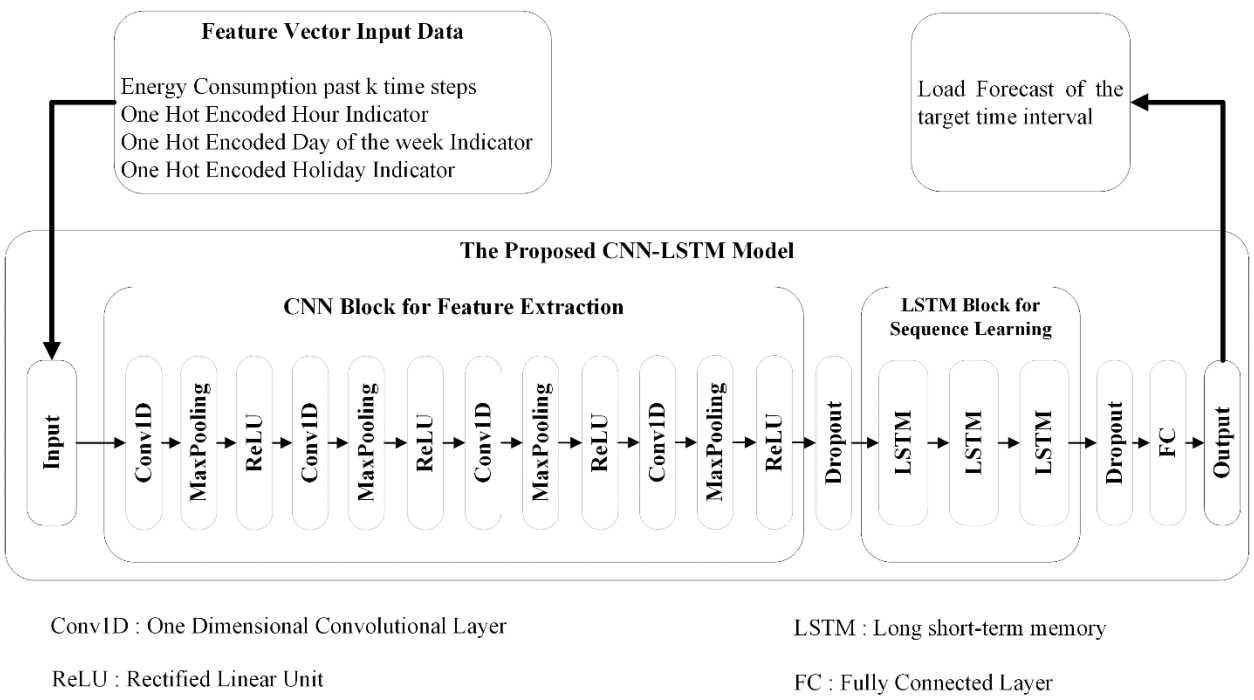

Figure 5.6 Proposed CNN-LSTM Model

For CNN feature extraction block, a pyramid structure is used where number of kernels are more in the lower-level layers that are gradually decreased in the higher-

level layers. Table 5-1 shows the configuration of various layers used in the development of CNN-LSTM model. A kernel size of 64 is selected for the first convolution layer. The kernel size is gradually reduced to 48, 32 and 16 for the second, third and fourth convolution layer. The pyramid structure reduces the number of trainable parameters. For the sequence learning block using LSTM layers, we have used 20 neurons. The return sequence is true for the first two layers of the LSTM whereas return sequence is set to false for the final LSTM layer.

Table 5-1 Configuration of various layers of the proposed model.

Proposed Model		
Conv1D_1	Kernel size	64
	Receptive Field	5
Pooling: Max	-	-
Activation: ReLU	-	-
Conv1D_2	Kernel size	48
	Receptive Field	3
Pooling: Max	-	-
Activation: ReLU	-	-
Conv1D_3	Kernel size	32
	Receptive Field	3
Pooling: Max	-	-
Activation: ReLU	-	-
Conv1D_4	Kernel size	16
	Receptive Field	3
Pooling: Max	-	-
Activation: ReLU	-	-
Dropout1	-	0.25
LSTM1	Hidden Nodes	20
	Return Sequence	True
LSTM2	Hidden Nodes	20
	Return Sequence	True
LSTM3	Hidden Nodes	20
	Return Sequence	False
Dropout2	-	0.25
Fully Connected	Hidden Nodes	20
Output		1 / 2 / 6

We have used Adam optimizer for training our deep learning model. It is a stochastic gradient descent method based on adaptive estimation of first order and second order momentum. Table 5-2 shows the training parameter settings for the developed CNN-LSTM model.

Table 5-2 Model Training Parameter Setting

Parameter	Setting
Optimizer	Adam
Loss Function	Mean Absolute Error (MAE)
Learning Rate	0.001
Learning Rate Adjustment	Monitor = validation loss, patience = 6 Epochs, factor = 0.9, minimum learning rate = 1e-5
Batch Size	256
Epoch	150

In order to train our deep learning model, we split the input data into training, validation and test data as 70%, 20% and 10% respectively. The validation loss is monitored using mean absolute error as a loss function. Figure 5.7 shows the complete flow for training our model. In the start, the training and validation data is loaded. After each epoch, the validation loss is calculated on the validation data and checked if the loss is decreasing. If the validation loss is decreasing then update the weights of the model and save it. If the loss is not decreasing for six successive epochs then the learning rate is decreased by a factor of 0.9 and then increment the epoch. The training stops as the number of epochs exceeds 150.

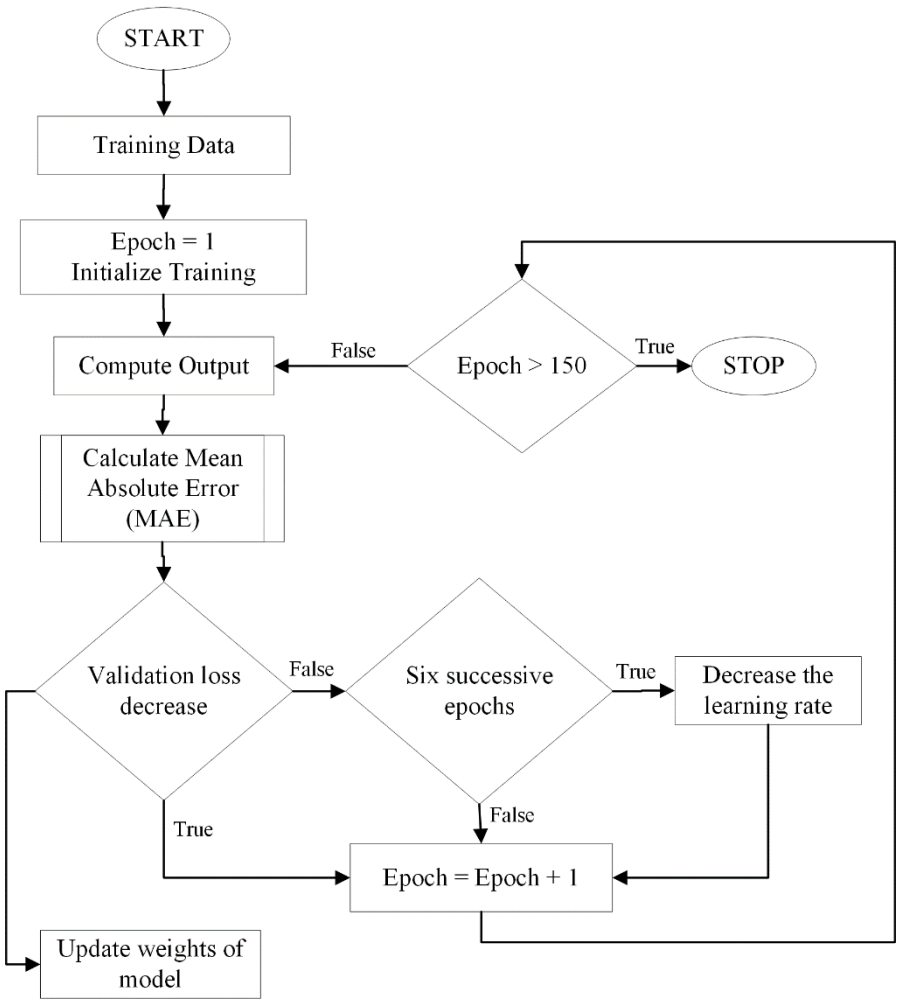

Figure 5.7 Proposed CNN-LSTM Model Training Flow

5.3 Results

In this project, we selected a group of 69 customers out of thousands of customer's data from SGSC project initiated by the Australian Government. The households were selected on the basis of hot water facility. Out of those 69 customers, some customers had energy consumption data of more than one year whereas for some customers the data was for few months. We choose time frame for which all the data was available for all the 69 customers. The energy consumption data was chosen from 1st June 2013 till 31st August 2013. For each household, the energy consumption data

consisted of 92 days, for which we used first 70% of the data for training, next 20% of the data for validation and last 10% of the data for testing the developed model. Figure 5.8 shows the steps in preparing the clean energy consumption data of 69 households.

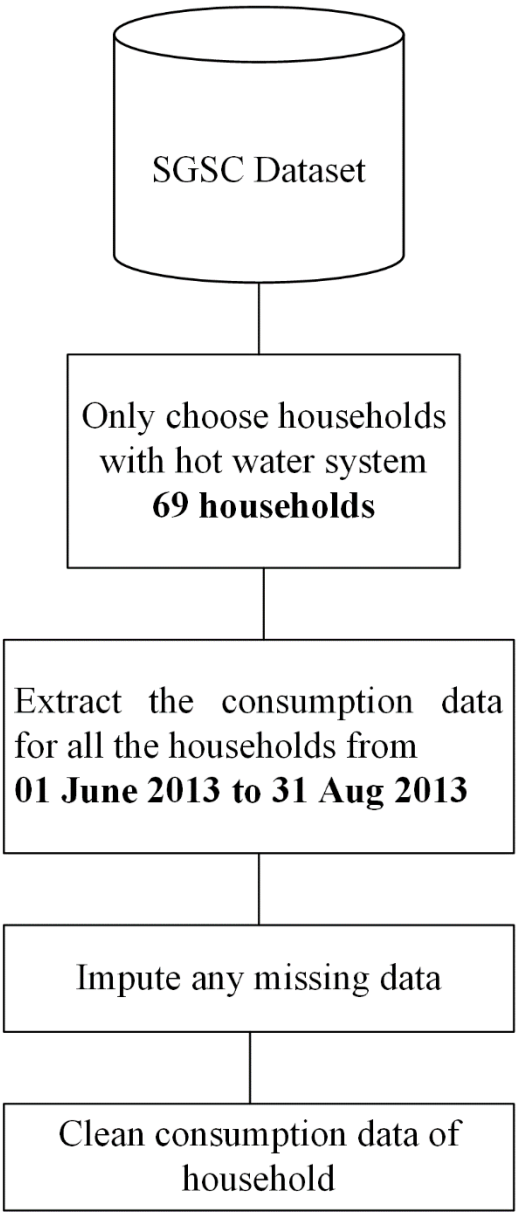

Figure 5.8 Energy Consumption Data of 69 Customers from SGSC Database

Figure 5.9 shows how the feature vector is prepared for the deep learning model. The energy consumption data of past k time steps is min-max normalized. The hour of

the day indicator, the day of the week indicator and weekend indicator, all one hot encoded. We have not considered the seasonal impact as the time frame under study is only for winter season.

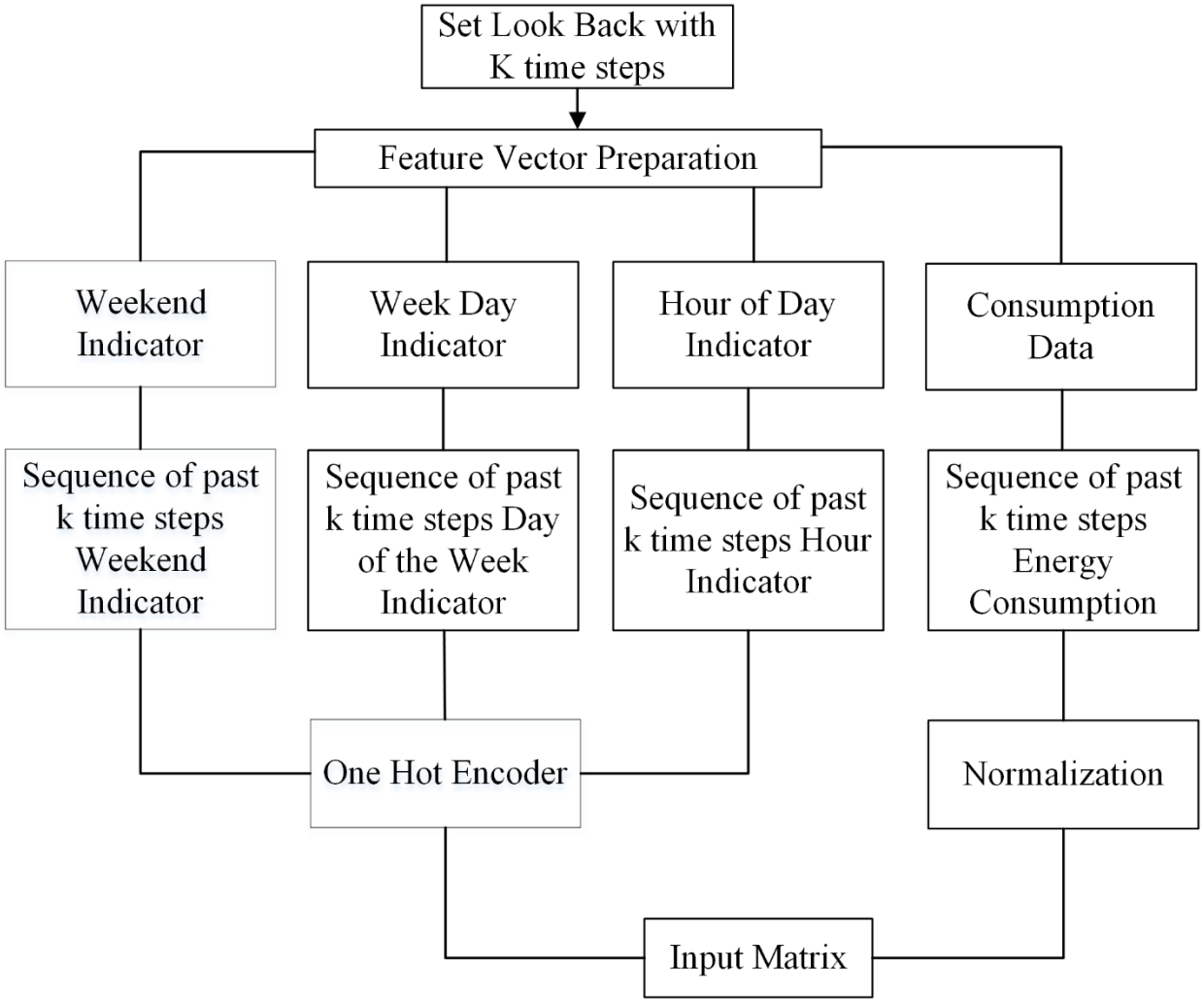

Figure 5.9 Feature Preparation

5.3.1 Single Step Forecast

In this experiment, we investigate the performance of proposed model to forecast the energy consumption at the next time step. In [2], authors compared their LSTM based deep learning model with various machine learning methods such as extreme learning machine, back propagation neural network, k nearest neighbor and input

selection combined with hybrid forecasting, and concluded that the LSTM model outperforms all the mentioned methods. Table 5-3 shows comparison of various machine learning methods with the proposed model. The MAPE of these machine learning methods have been taken from [2]. Figure *5.10* to Figure *5.12* shows the load forecasting of various customers using proposed method as well as LSTM model.

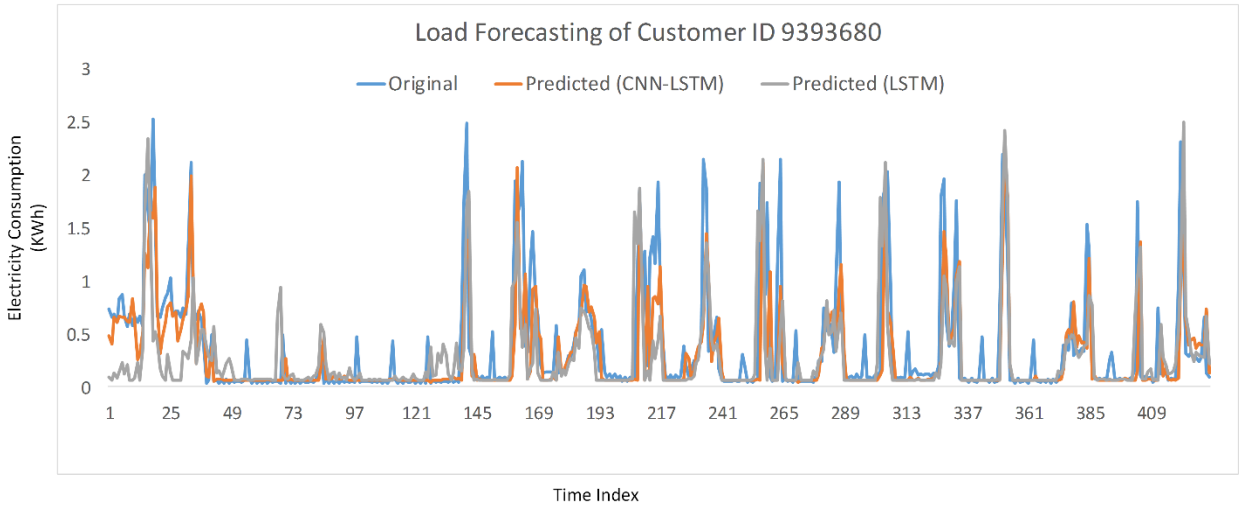

Figure 5.10 Load Forecasting of Customer ID 9393680 using proposed model and LSTM

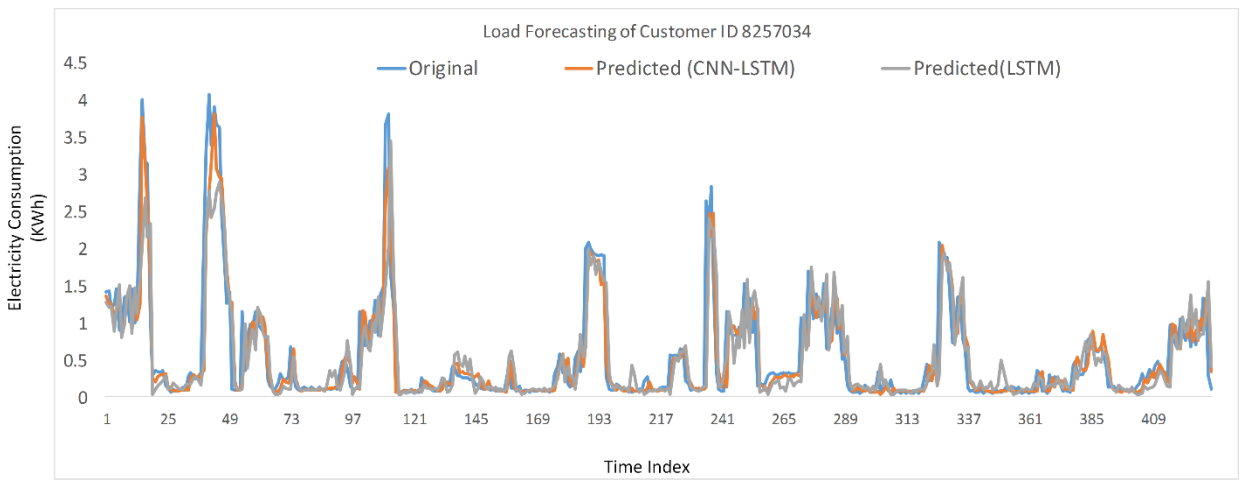

Figure 5.11 Load Forecasting of Customer ID 8257034 using proposed model and LSTM

The x-axis shows the time index corrosponding to various days in the test data whereas y-axis shows the energy consumption in KWh. By carefully analyzing these figures, we can infer that the LSTM model result in unwanted peaks and valleys at

various time indexes. For cusomer ID 9393680, from time index 1 to 35, the forecast made by proposed model is very close to the actual load whereas the prediction made by LSTM is far from actual values. For customer ID 8257034, LSTM model is struggling to follow the peaks whereas CNN-LSTM model is closer to peak values.

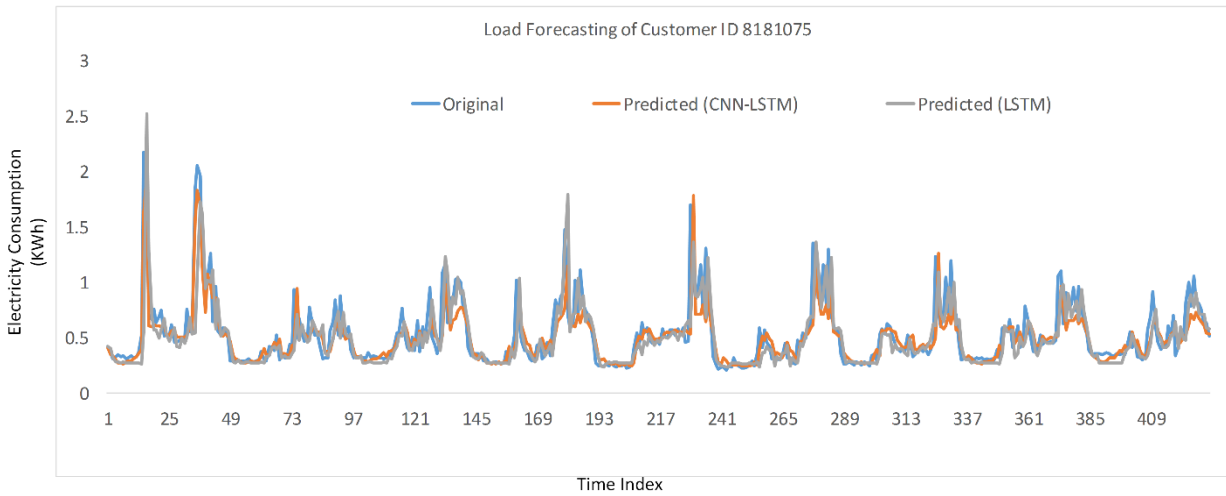

Figure 5.12 Load Forecasting of Customer ID 8181075 using proposed model and LSTM

Figure *5.13* shows comparison of proposed model with the LSTM model. The x axis shows the number of outliers in the training data of household, whereas y axis shows the mean absolute percentage error (MAPE). We can make the following observations from this figure;

1. MAPE of the proposed method is lower than that of the LSTM model for varied number of outliers in the data.

2. The forecasting becomes more challenging as the number of outliers increase. However, the proposed model performs far better than the LSTM model in such scenarios.

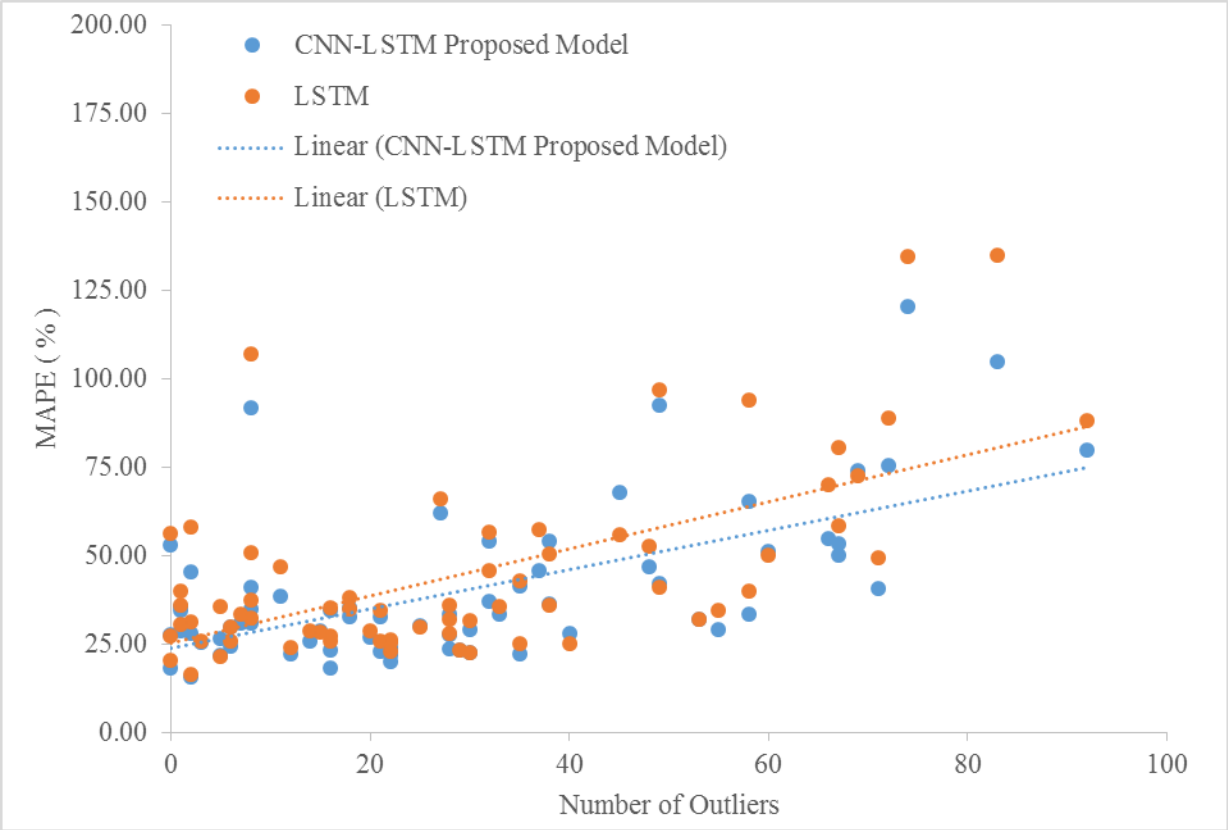

Figure 5.13 Comparison of CNN-LSTM model with the LSTM model

The empirical mean is the most naïve method listed in table below forecasts based on statistical mean when the information such as customer ID, hour of the day, day of the week, weekday or weekend is provided. The empirical mean method achieved an average MAPE of 136%,

MAPE minimization method calculates the probability mass function from the historical data distribution. This method achieved an average MAPE better than many machine learning methods. MAPE minimization method, Back Propagation Neural Network (BPNN), LSTM and the proposed method achieved an average MAPE of below 50%. The extreme learning machine (ELM) is another sophisticated machine learning method that gives much better forecasting accuracy for aggregated load. However, this method fails completely at an individual household level. The

LSTM model achieved an average MAPE of 44.06% for look back of 12 time steps whereas the proposed model achieved best results for look back of 2 time steps with an average MAPE of 40.20%.

Table 5-3 Single Step Forecast Comparison of Proposed Method with Various Methods

Method	**Scenario (Lookback)**	**Individual Forecast Average MAPE (%)**
	2-time steps	**40.20**
Proposed Method	6-time steps	**40.95**
	12-time steps	**42.47**
	2 time steps	44.39
LSTM	6-time steps	44.31
	12-time steps	44.06
	2-time steps	49.62
BPNN	6-time steps	49.04
	12-time steps	49.49
	2-time steps	74.83
KNN	6-time steps	71.19
	12-time steps	81.13
	2-time steps	122.90
ELM	6-time steps	136.49
	12-time steps	123.45
MAPE Minimization	-	46.00
IS-HF	-	96.76
Empirical Mean	-	136.46

5.3.2 Multiple Step Forecast

In this experiment, we test the CNN-LSTM method for various look forward time horizons. If the time horizon was set to 3 hours ahead, then the deep learning model will forecast next six values. The result of multi-step forecasting for different time

horizons is shown in Table 5-4. It is evident from this table that for all the look back and look forward time steps, the CNN-LSTM model achieved better MAPE than the LSTM model. One observation is that the improvement in terms of percentage is slightly decreasing when the lookback is increased from 2 to 12. This happens only for look forward of single time step which is due to the fact that the LSTM model is good in learning from the lag features of training data. In all the remaining cases, the gap between the LSTM model and CNN-LSTM model in terms of MAPE is generally increasing for incremental look ahead time steps. This experiment shows that the performance of CNN-LSTM model is better than the LSTM model not only for single step but for multi-step load forecasting.

Table 5-4 Multi-Step Forecasting for various look-back and look forward. A comparison of proposed method with the LSTM model.

Lookback (no. of time steps)	**Look-forward (no. of time steps)**	**Individual Household Forecast Mean MAPE (%)**	
		CNN -LSTM	**LSTM**
2	1	40.20	44.39
	2	46.31	51.84
	6	59.13	65.14
6	1	40.87	44.31
	2	46.17	53.76
	6	59.35	67.14
12	1	42.07	44.06
	2	47.54	56.38
	6	60.94	69.25

For a look-back of two time steps, the CNN-LSTM model achieved an improvement of 4.19%, 5.53% and 6.01% for look forward of one, two and six time steps. Although

the MAPE achieved are far from accurate but the results achieved are much better than many machine learning and deep learning methods. The reason for such high MAPE is the inconsistency and volatility in daily load profiles of individual households. We have seen in Figure *5.13* that even for households with large number of outliers (high volatility) the CNN-LSTM model achieved much better forecasting accuracy than the LSTM method (which outperforms many machine learning methods).

5.3.3 Cluster Analysis

In this experiment, we try to group the customers based on their consumption pattern, i.e. with similar load profiles. We apply k-mean clustering technique on the training data of 69 customers from SGSC database. For this experiment, we only use the training data of each customer.

We partition a day into four intervals described below,

1. Morning: 06:30 – 09:30
2. Daytime: 09:30 – 15:30
3. Evening: 15:30 – 20:00
4. Night: 20:00 – 06:30

We prepare the data for k-mean cluster to group the customers with similar load profiles during each time interval. We were mostly interested in grouping the customers with similar consumption behavior rather than the amount of actual power consumption during these time intervals.

We used the following features for k-mean clustering;

1. Feature 1 – 4: Relative mean power for each time interval which is the ratio between the sum of average power consumed during each interval to the daily average consumption over the entire data.
2. Feature 5: The average relative standard deviation.
3. Feature 6 – 9: The absolute difference between the average power consumed during the weekend vs average power consumed during weekday score for each partition of day.

With the help of above nine features, we formulate the consumption behavior of households. The average silhouette score was used to find the optimal number of clusters. This score ranges from -1 to +1. A larger positive value indicates that the household is similar to the other member of its own cluster and at the same time very dissimilar to the other members of other clusters. Figure *5.14* shows that the maximum silhouette score achieved is 0.185 for k = 15.

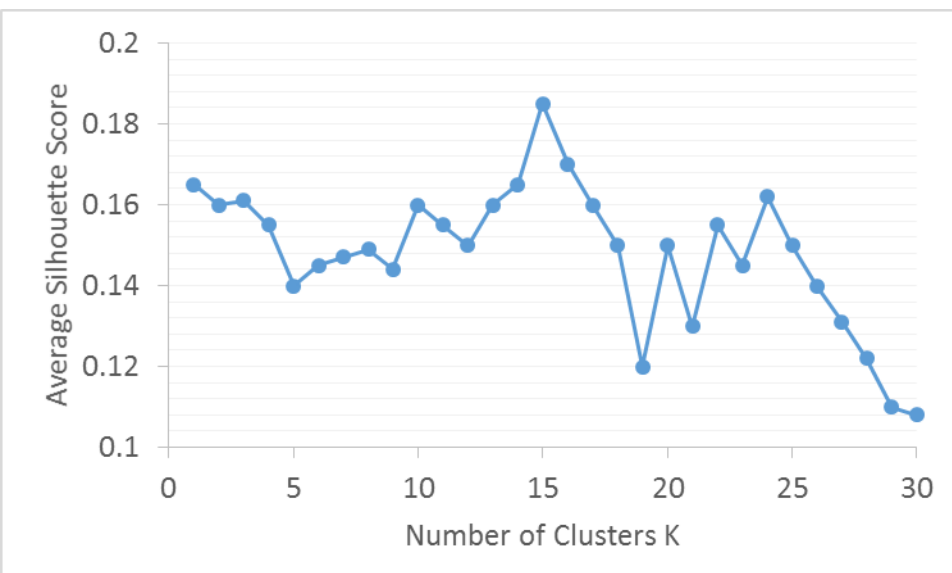

Figure 5.14 Silhouette score for grouping residential customers

The normalized average power consumption of all the 15 clusters is shown in Figure *5.15*.

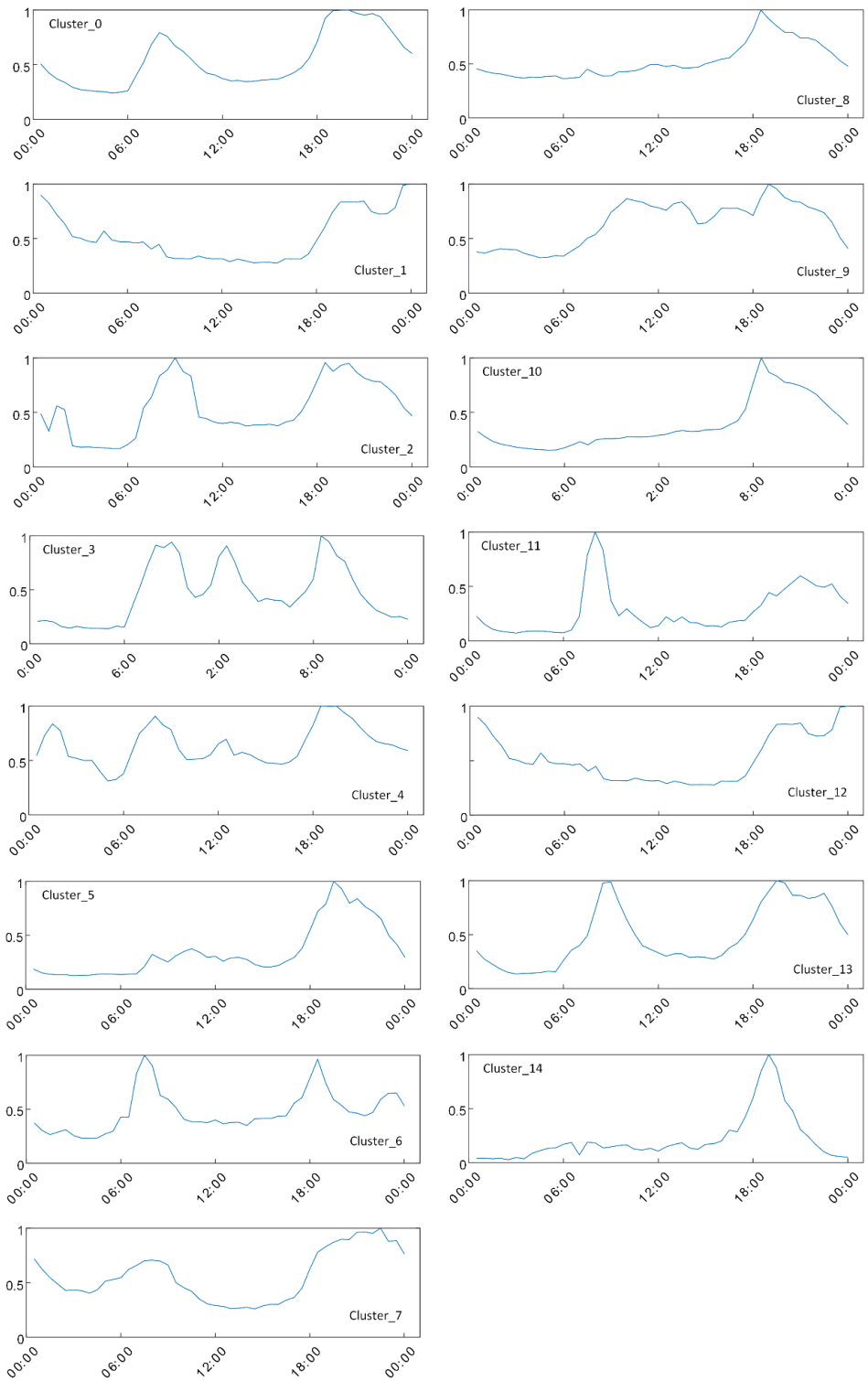

Figure 5.15 Normalized Average Power Consumption of each Cluster

We have used the training and validation data for Figure *5.15*. For each cluster, we have normalized the power consumption from the maximum value within that cluster. We can observe that each cluster is somewhat different from the other cluster. For instance, Cluster_2 and Cluster_3 have three peaks each but they differ in such a way that there is an overnight smaller peak at night for Cluster_2 whereas Cluster_3 don't have a peak at night. Another observation is that Cluster_3 has a distinct peak in the day whereas Cluster_2 don't have peak at day time. Cluster_5, Cluster_6, Cluster_10 and Cluster_14 have one peak each. They all differ in terms of average power consumption during various portions of the day as well as thickness of the peak consumption span.

In Table *5-5*, we provide a one-on-one comparison of LSTM model without grouping, proposed method without grouping and proposed method with grouping. In the clustering experiment, we examine the forecasting improvement by grouping households with similar load profiles. We prepare the training, validation and test data for each cluster by combining the feature vectors of each customer within the cluster. Column 1 of Table *5-5* shows cluster number, column 2 shows customer ID that belongs to each cluster, column 3 shows volatility in the training data, i.e. number of outliers, column 4 shows the MAPE using LSTM model without grouping, column 5 shows the MAPE using CNN-LSTM model without grouping, and column 6 shows the MAPE using CNN-LSTM model with grouping. For some clusters, the number of customers are very few. It is due to the fact that the number of customers under study are only 69 which is not large enough. It is inferred from the table that irrespective of the forecasting method, the MAPE is generally large when the number

of outliers are more. For households with no distinct daily profiles, the CNN-LSTM model without grouping achieved much better results than the LSTM model. The reason for clustering method to not perform generally well for clusters with larger number of outliers is due to the fact that the training data consists of diversified load profiles because of no distinct daily profiles which result in deep learning model to not optimize.

Table 5-5 Comparison of LSTM method, CNN-LSTM method and proposed method with Clustering for each household

Cluster Number	Customer ID	Number of Outliers in Training Data	LSTM without Cluster Analysis	CNN-LSTM without Cluster Analysis	CNN-LSTM with Cluster Analysis
Cluster_0	10598990	46	52.48	47.64	41.56
	11081920	53	34.17	29.12	30.67
	8181075	17	27.60	18.05	24.51
	8184653	29	32.12	28.21	25.15
	8211599	30	45.90	35.71	46.32
	8376656	57	39.17	33.05	43.14
	8451629	13	28.51	26.53	31.45
	8540084	64	73.74	70.11	84.12
	8617151	16	35.69	32.39	32.42
	8673172	17	38.31	33.14	35.95
	8679346	25	67.12	64.59	62.97
Cluster_1	10692972	9	47.36	37.65	37.83
	11462018	06	35.98	26.40	34.51
	9012348	73	135.91	99.31	130.10
	8196669	13	24.01	22.78	22.92
Cluster_2	8196671	07	25.94	24.31	25.41
	8482121	01	31.3	28.11	39.97
	8487461	35	50.77	56.02	61.69
	8342852	02	40.21	33.45	36.21

Cluster_3	10595596	64	70.75	49.31	42.51
	8264534	48	32.07	33.15	39.34
	8519102	42	55.63	62.6	53.43
Cluster_4	8273230	18	25.91	26.3	48.03
	10702066	36	36.3	35.43	47.97
	8504552	29	23.62	24.62	27.91
	8487285	35	25.11	22.97	27.84
	8351602	21	22.94	22.32	25.11
Cluster_5	8145135	64	49.75	44.51	48.32
	8149711	30	56.43	55.19	52.68
	8568209	38	57.21	29.63	68.81
Cluster_6	10509861	25	28.43	23.18	23.41
	8147703	46	41.01	42.09	48.61
	8273592	31	31.75	28.77	28.1
Cluster_7	8680284	01	58.21	50.6	51.87
	8257034	65	80.78	51.93	48.21
	8432046	62	58.11	53.55	64.56
Cluster_8	8257054	21	25.92	20.09	21.56
	8685932	03	16.11	15.32	14.71
Cluster_9	8328122	09	32.11	30.76	28.64
	8308588	22	26.34	24.44	24.88
	8804804	01	36.11	36.58	41.47
	8566459	34	42.75	43.41	42.89
	8496980	05	29.91	29.91	30.41
	8618165	08	33.56	31.07	29.97
	8350006	04	21.54	22.18	31.96
	8176593	31	22.43	22.33	23.39
	8282282	82	81.19	72.57	98.87
Cluster_10	8466525	03	25.52	25.31	27.84
	8478501	09	51.03	42.87	37.95
	8419708	20	28.96	27.95	24.41

	8196621	15	35.49	34.44	34.92
	8459427	00	20.41	18.79	20.98
	8733828	21	29.93	29.54	31.68
	8687500	29	35.73	33.37	40.53
	8347238	23	36.21	33.81	36.29
	8334780	17	25.94	23.54	23.53
	8198267	21	34.47	32.78	34.87
Cluster_11	9393680	64	89.31	76.59	76.59
Cluster_12	8523058	00	56.15	56.38	43.41
	8655993	01	27.43	32.31	27.69
	8198319	01	30.49	28.84	36.71
	8487297	08	107.04	95.31	109.8
Cluster_13	8156517	46	97.27	93.32	145.54
	8661542	15	28.45	28.22	29.94
	8557605	08	37.49	35.87	34.31
	8326944	54	94.02	70.98	74.63
	8196659	40	25.11	28.13	38.54
Cluster_14	8198345	55	50.13	53.22	55.31
	8264534	48	32.31	32.94	39.17
	8291712	69	134.98	121.21	131.39
Average			44.78	40.20	44.91

An interesting observation from the same table is that for some households, the MAPE is very high even though the number of outliers are very few to none. As an example, consider a customer 8680284 in cluster_7 which has only one outlier in the training data, still has a large value of MAPE. If we observe in Figure *5.16*, the average daily load profiles of this customer during the training and test data, we found that in the training data there is a peak event in the late evening whereas no peak in the test data. The model may have predicted the wrong peak which resulted in larger

MAPE. From Figure *5.17* to Figure *5.21* shows the forecasting of various households in each clusters.

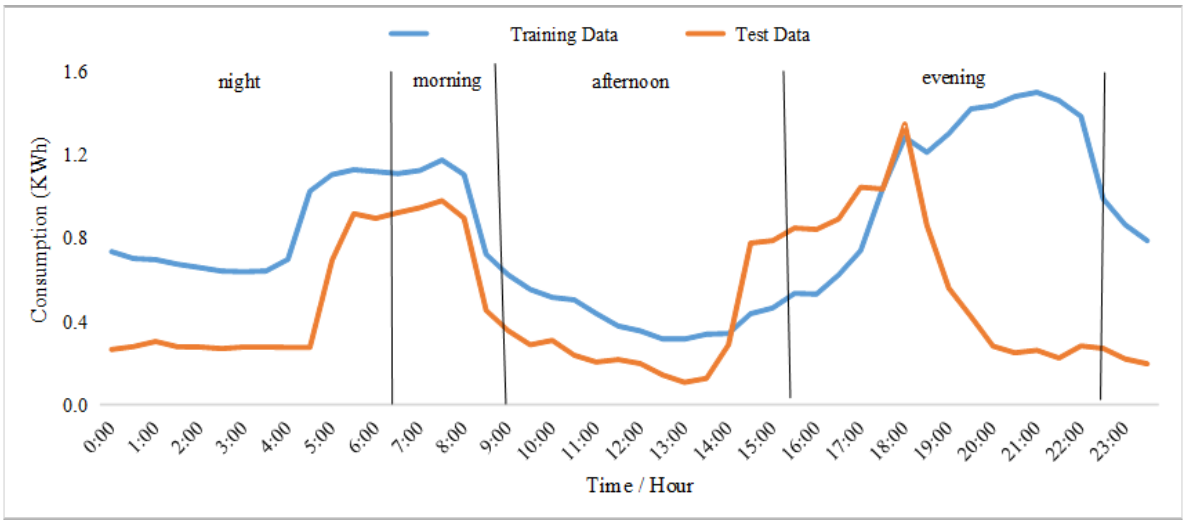

Figure 5.16 Cluster_7: Customer ID 8680284 average power consumption during train and test data

Figure *5.17* shows the forecasting of Cluster_3 households using CNN-LSTM model. We can see that the actual power consumption is very different for all the three households. The proposed model is struggling to follow the peak for customer ID 10595596. However for the remaining two customers the forecast is very close to the actual power consumption.

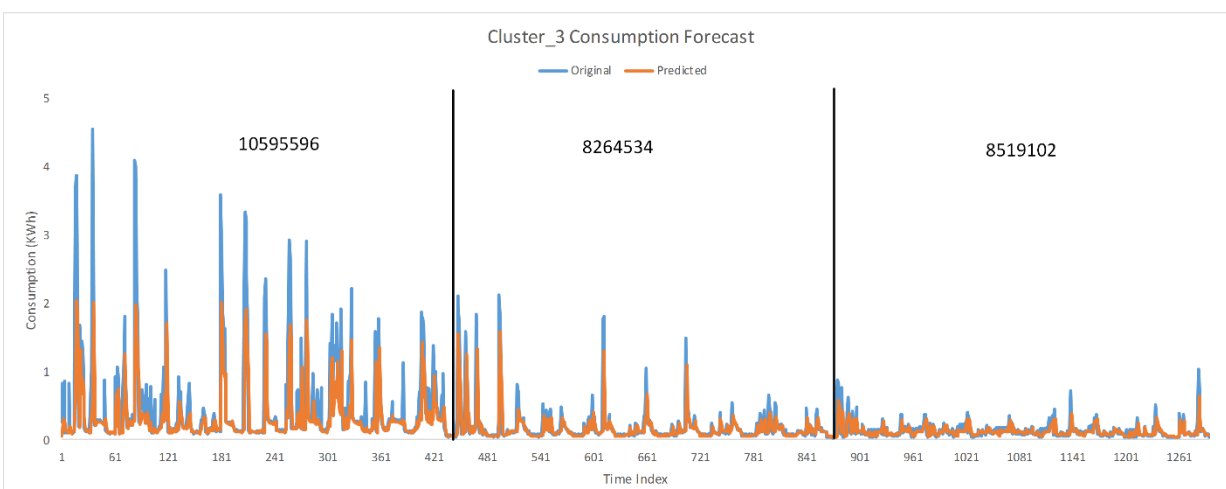

Figure 5.17 Cluster_3 Consumption Forecast

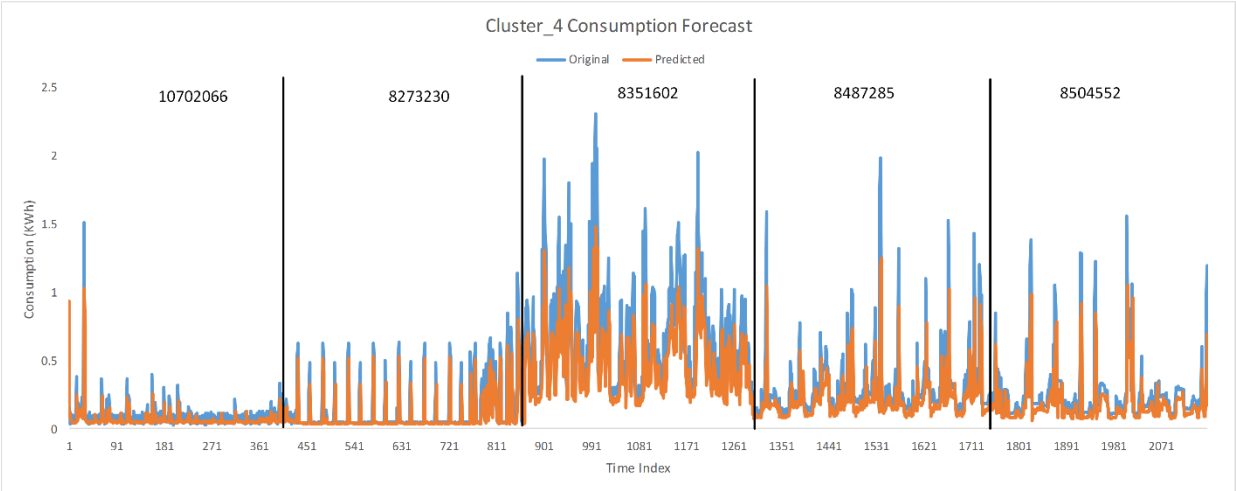

Figure 5.18 Cluster_4 Consumption Forecast

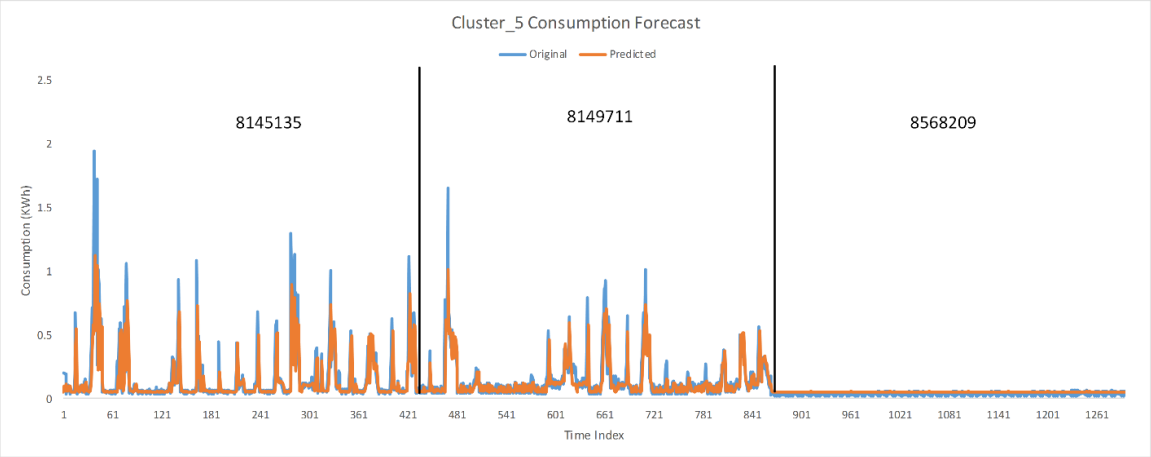

Figure 5.19 Cluster_5 Consumption Forecast

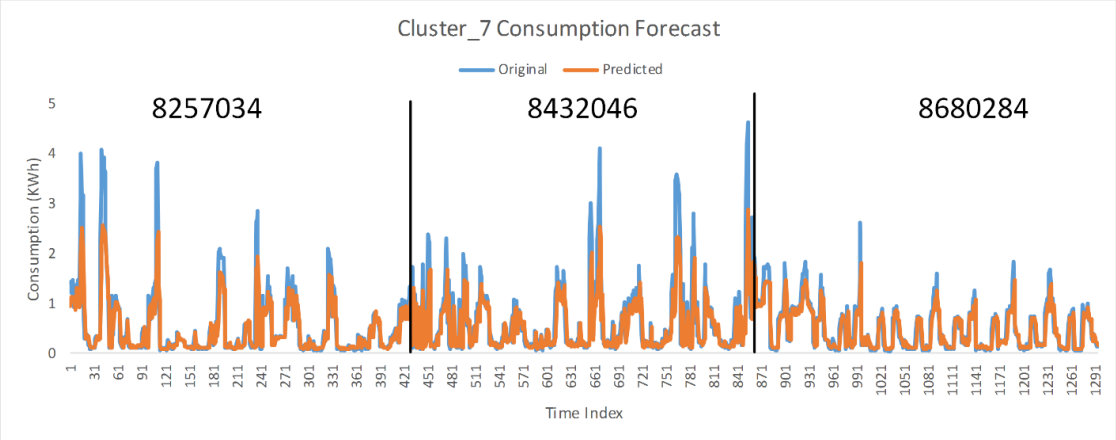

Figure 5.20 Cluster_7 Consumption Forecast

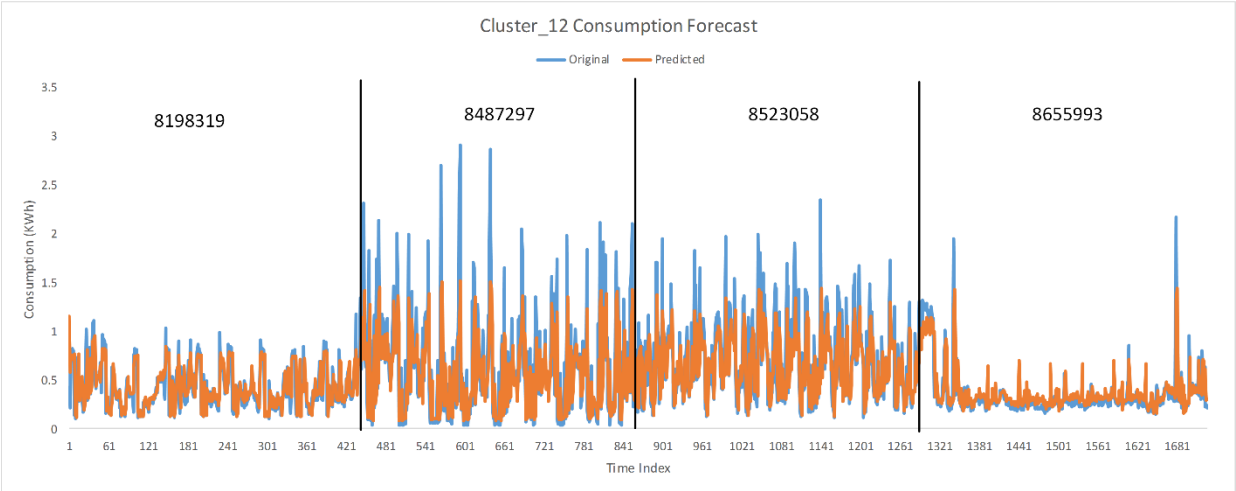

Figure 5.21 Cluster_12 Consumption Forecast

We observed diversity in load forecasting and it happened due to the random power consumptions and for some householders with the power consumption in the training data was below 1kwh. In general, the CNN-LSTM model performed better when it comes to following the peaks. However, for some customers with very high vitality in their daily profiles, the proposed model finds it hard to accurately forecast. Still the results achieved are better than many machine learning methods as well as state-of-the-art deep learning model based on LSTM.

5.4 Conclusion

In this chapter, we proposed a deep learning model based on CNN layers and LSTM layers. We exploit the advantages of both CNN and LSTM by using CNN layers to extract the features from the input training data and LSTM layers to find the long-term dependencies in the data. We found that the CNN-LSTM model achieved much better forecasting accuracy in predicting the load for next few hours. The average MAPE for individual household is generally high

due to the volatile daily profiles and consumption behavior of each household. The load forecasting using CNN-LSTM achieved better results than various machine learning and deep learning methods. We have seen the forecasting result for the 3 months dataset using the proposed hybrid model but, in the future if we can use a dataset that covers the seasonal changes will be more effective and realistic.

This model will be helpful in predicting day-ahead forecast of each residential customer and the utility company will provide a feedback to the customer about their day-ahead forecast consumption using feedback technology installed with smart meters. The utility provider inform the customer day ahead about the forecasted peak event so that the customer may schedule the load from peak-time to off-peak time using load schedulers.

Chapter 6

Load Scheduling of Residential Customer for Dynamic Peak Pricing Scheme

The demand for electricity varies considerably over the period of year/month/day which results in variation in the supply. As the production of electricity varies to satisfy demand, the electricity supplier encounter marginal costs that changes over time. The reason for this is that the electricity demand varies considerably during the day and across days. Moreover, it is inefficient to store the electricity in certain cases. When the demand is less, the base load capacity is adequate. However, when the demand rise, the electricity supplier uses the generating capacity with higher marginal costs to meet the demand. Based on the cost of substitute fuel and the generating method, the marginal cost varies considerably. Usually, the marginal cost is lowest at night and highest in the evening.

In the traditional power grid, the pricing is always fixed, which makes users not interested in how and when they use the energy. But dynamic pricing improves energy consumption and makes users interested in their energy usage behavior regarding the time. However, it is likely that their behavior will not change at sometimes, such as at dinner and lunch, but they will be interested in changing some behaviors such as doing the laundry and charging the electric car late at night. Dynamic pricing will also change the way energy is delivered around the world and make the grid smarter in the long run. The main purpose for dynamic pricing is to smoothen the energy demand curve by shaving the peaks and filling the valleys.

The flat rate tariff is not beneficial for the electricity supplier as it doesn't balance the suppliers marginal cost with their marginal revenue. To address such issue, one of the methods is Time of Use (ToU) pricing. ToU pricing scheme usually have high weekday evening peak rates and reduced off-peak rates during the remainder of week. The demand is less during the high cost period and more in the off-peak period. ToU is priced according to the day, season and type of day (weekday or weekend/holiday). To achieve sustainability and power efficiency through this pricing scheme, higher prices are set during peak times and lower prices during times of low energy demand times. Through this pricing, the user can reduce his power usage consumption cost and reduce the load on the network. This pricing scheme helps to solve the peak demand challenge to some extent however, the overall energy consumption is not reduced.

Thanks to the advent of smart meters that incorporate better technology in terms of data acquisition, data storage, data processing and data communication. The smart meters provide two-way communication between the electricity supplier and consumer. The real time electricity consumption data is provided to the electricity supplier. Moreover, a smart meter can receive a price change notification in real time. Real-time pricing (RTP) is another scheme that can benefit from the smart metering. This pricing scheme helps to solve the peak demand challenge; however, it is less beneficial than Dynamic Peak Pricing (DPP). From the policy makers' perspective, RTP is very complex especially for small electricity users such as residential customers.

6.1 Various Trials of Dynamic Peak Pricing (DPP)

From past many years, DPP trials have been conducted by various electricity providers. This pricing scheme is discussed as a substitute to fixed pricing and ToU pricing schemes. The DPP is achieved when the electricity provider set higher cost of electricity during the peak events which will help in reducing the load during peak times. The DPP sets very high cost for consecutive several hours on extreme weather condition days throughout a year. These days are extreme hot or cold days that are already predicted, where the consumers are informed often 30 mins or 1 day prior. The DPP can be used to encourage reduce electricity consumption during the period of high demand or low supply.

A DPP trial was conducted in California in 2003-2004 where the customers were grouped into two categories. The first category was price only group whereas the second category was price plus technology group (allow the supplier to turn off the air conditioning during peak events). It was found that the customers with price only group reduced their energy consumption by 13% of baseload when the peak event was for five hours. For the customers with technology, the usage was reduced to 41% of baseload. In this trial, the price was increased by 360% during the peak event. In the same experiment, they raise the price during the peak event from $0.5 per KWh to $0.68 per KWh. They found that the household didn't respond to the dynamic peak pricing for such a small increase in cost during peak event.

Another study conducted in Missouri in 2003-2004 where the customers were grouped similar to the previous study. They found the usage drop by 12% for first category and 35% for second category.

Another trial was conducted in Baltimore in summer of 2008 and 2009. It was found that the usage was reduced by 20% for the customers with no feedback technology whereas the usage was reduced by 32.5% for the customers with feedback technology.

In 2014, a study was conducted in Connecticut over a period of two months. They found that the households with in-home-display respond to dynamic peak pricing to reduce the usage by 14% whereas households without in-home-display do not respond to dynamic peak pricing. It was further found that the households with in-home-display shift their load if notified about the dynamic peak event one day prior. The following two observations were made from the above trials. First, when the cost was significantly increased during the peak event time, then the households responded well to the DPP which resulted in significant reduction in usage. Second, by introducing the technology along with the DPP result in further reduction of usage.

6.2 Smart Grid Smart City (SGSC) Project

The SGSC project was initiated by Australian government to demonstrate the trial deployment of smart grids in Newcastle area during period of 2012 – 2014. The Newcastle area was selected because of its versatile demographic characteristics that are representative of Australia. In this trial, 7612 households were selected and were given one of the three pricing tariffs. Figure 6.1 to Figure 6.6 have been taken from [25].

The household participants were asked about why they would like to join the trial. They were allowed to choose more than one category. The participants desire for

the reduce bill was highest (64%) whereas electricity consumption reduction was second (41%) as shown in Figure 6.1.

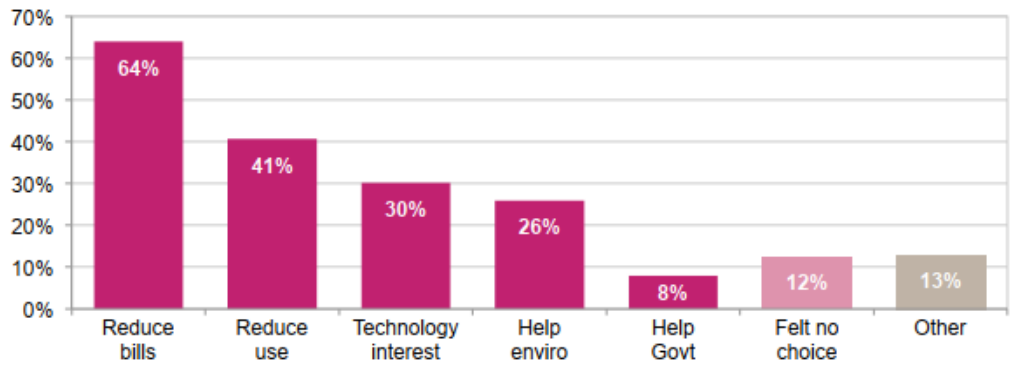

Figure 6.1 Key Motivation to participate in the SGSC Trials

In the trial, the participants were asked if the cost of electricity during 2pm to 8pm will be high during weekdays, would you be able to shift your electricity consumption to an off-peak time. The participant's response is shown in Figure 6.2.

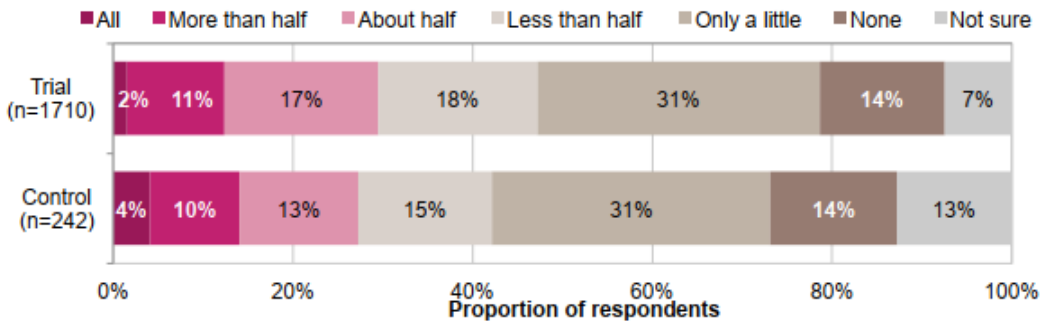

Figure 6.2 Electricity consumption shift to off-peak time

Around 17% of the participants were willing to shift all their appliances to off-peak time whereas 39% were unable to shift at least 1-2 appliances as shown in Figure 6.3

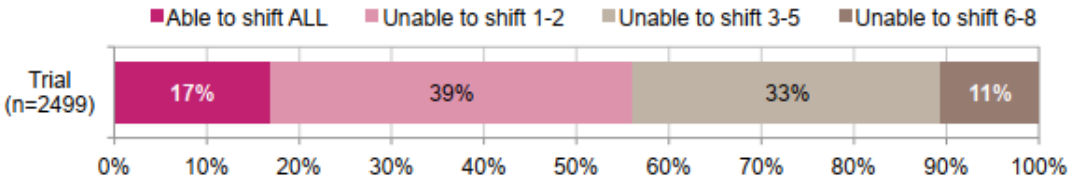

Figure 6.3 Number of Shift able Appliances to Off-Peak Time

The participants were asked which appliances are difficult to be shifted to off-peak time. The participants were allowed to choose more than one option. Around 61% of participants responded that they would not like to shift the refrigerator during the peak event followed by 51% for oven/stove and 41% for T.V. Around quarter participants were unwilling to shift the heat/cool devices as shown in Figure 6.4.

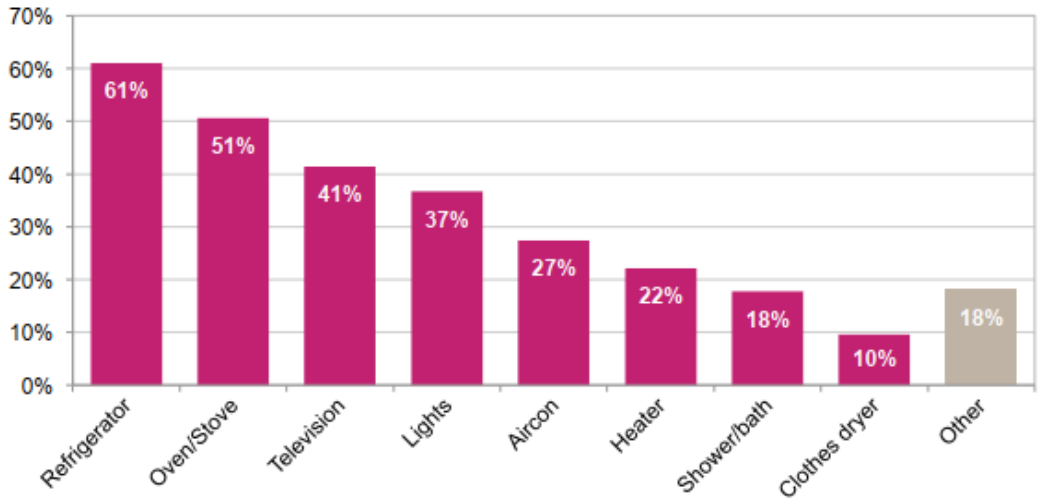

Figure 6.4 Proportion of appliances customers are unable to shift to off-peak time

Before the trial, the households were either on flat tariff or on ToU tariff. The customers on flat tariff were paying the same price all day every day as shown in Figure 6.5. For ToU customers, the price varied according to hour of the day as shown in Figure 6.5.

When the trial begin, some households were kept with same pricing scheme as before so to compare the effect of other two pricing schemes. For the control group with DPP, the cost during the period of 7 am till 10 pm was changed to 24.53c/kWh. During the peak event between 2 pm to 8 pm, the price was increased to 330c/kWh (1200% increase). The DPP was adjusted such that if the customer didn't change the

electricity consumption behavior during the peak event then the annual electricity bill would be cost neutral.

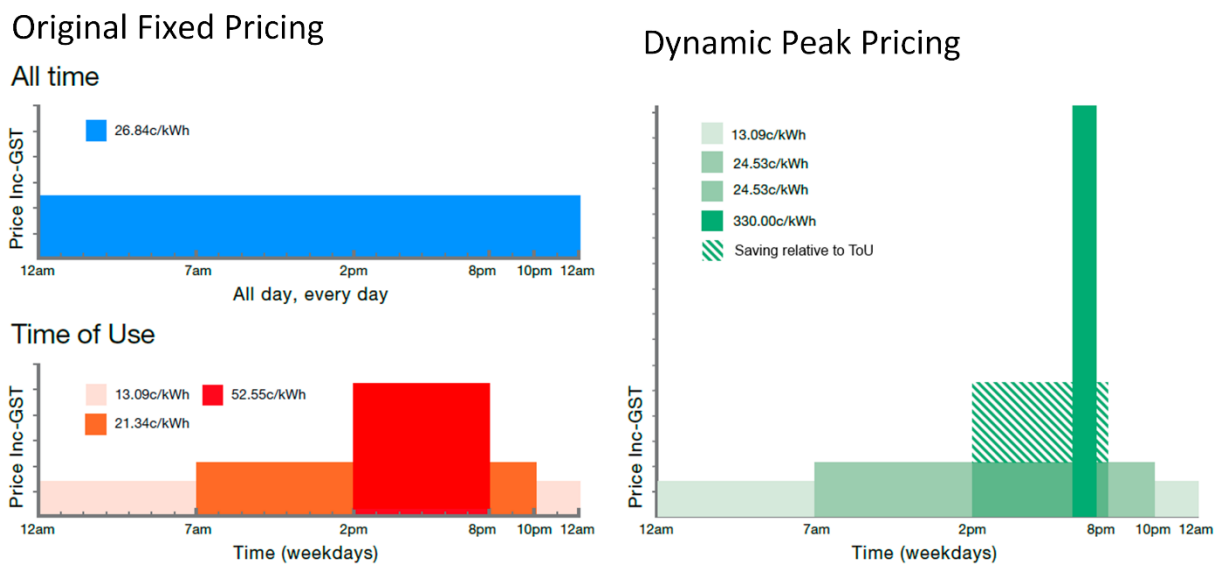

Figure 6.5 Pricing Design SGSC Database

Along with the pricing schemes three different feedback technological trials such as in-home-display, smart plug and online portal were implemented. The customers were randomly assigned with no feedback technology, single feedback technology and multiple feedback technology.

Figure 6.6 shows the average consumption of group of customers on DPP for peak event days [26]. The control group is the baseline group with fixed pricing scheme. DPP is the dynamic peak pricing group without any feedback technology. DPP-IHD is the dynamic peak pricing group with in home display feedback technology. DPP-Portal is the group of customers with online portal as feedback about their consumption behavior. Finally, DPP-Smartplug is the group of customers for which smartplug is connected to various appliances to give information about the

consumption of specific appliance. The gray thick highlighted area is the peak event occurrence on specific days.

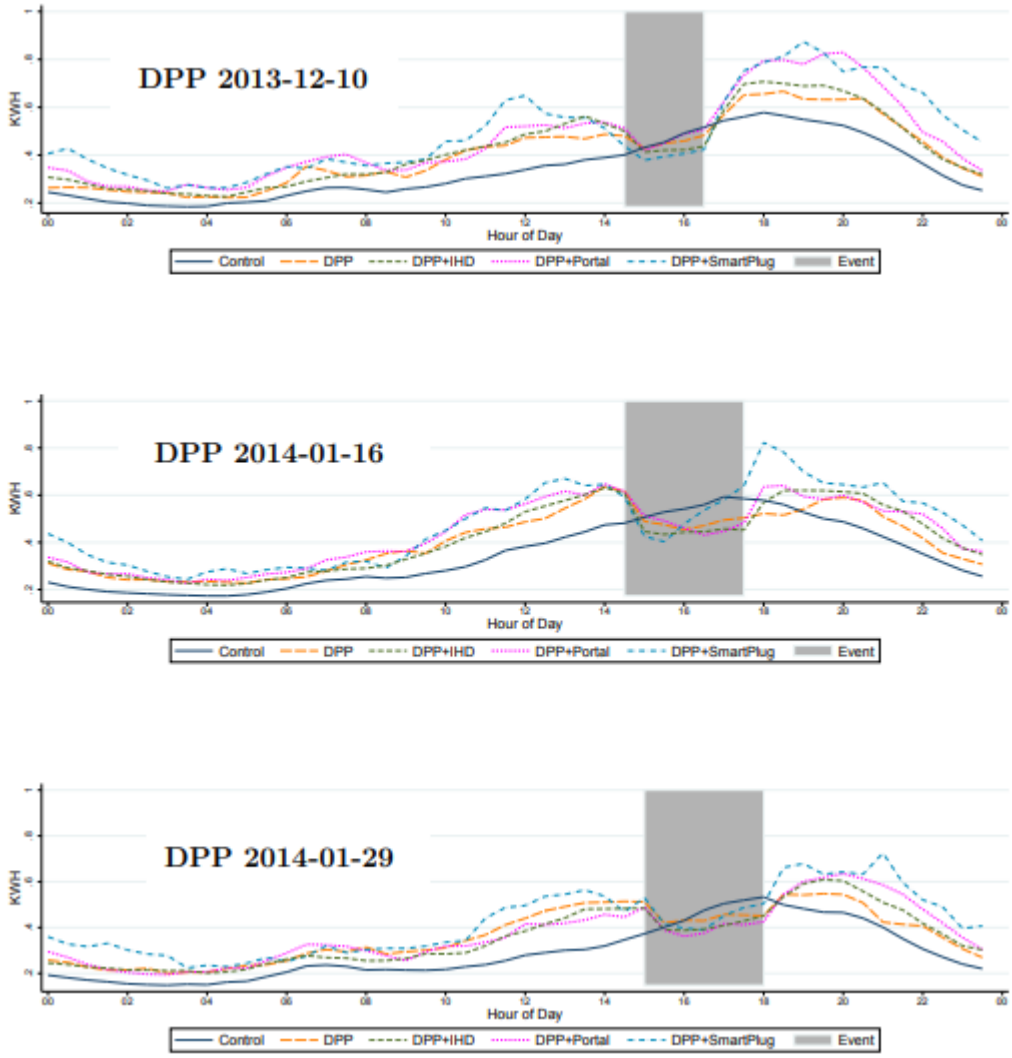

Figure 6.6 Average usage of each control group with DPP during peak events [26]

The following two observations can be made from the above figure; first, DPP group with or without feedback technology reduced their consumption during the peak event time as compare to the fixed pricing scheme (control group), second, load shifting before and after the peak event.

Table 6 -1 Information about Customers under Study

Customer ID	Household Income Group	Electricity Usage Group	Dryer Usage	Reducing Consumption Effort	Aircon Type	Children	Refrigrators	Number of Rooms Heated	PoolPump	Aircon	Is Home During Day
8176593	HI	LOW	LOW	HI	SplitSystem	N	3	3	N	Y	N
8181075	HI	HI	LOW	MED	Ducted	N	1	0	N	Y	N
8196621	HI	LOW	LOW	MED	SplitSystem	N	2	4	N	Y	N
8196659	HI	LOW	LOW	HI	Ducted	Y	2	0	N	Y	N
8196669	HI	MED	LOW	HI	Ducted	N	1	0	Y	Y	N
8196671	HI	MED	NONE	LOW	SplitSystem	N	2	0	N	Y	N
8257034	HI	HI	MED	HI	Ducted	Y	1	0	N	Y	Y
8273592	HI	HI	LOW	MED	Ducted	N	4	0	N	Y	Y
8328122	HI	MED	MED	MED	SplitSystem	Y	1	0	Y	Y	N
8342852	HI	HI	LOW	HI	Ducted	Y	3	1	Y	Y	Y
8733828	HI	MED	NONE	LOW		Y	1	0	N	N	Y
8804804	HI	HI	LOW	HI	Ducted		2	6	Y	Y	Y
9012348	HI	LOW	LOW	NONE		Y	1	0	N	N	Y
9393680	HI	LOW	NONE	LOW		N	1	0	N	N	N
10509861	HI	MED	LOW	HI	SplitSystem	N	3	0	N	Y	Y
10595596	HI	HI	MED	MED	Ducted	Y	1	0	N	Y	Y
10598990	HI	HI	MED	LOW	Ducted		1	0	Y	Y	N
10692972	HI	HI	MED	LOW	Ducted	N	2	0	Y	Y	Y
10702066	HI	LOW							N	N	N
11081920	HI	HI	NONE	HI	Ducted		1	0	N	Y	Y
11462018	HI	LOW	LOW	HI	SplitSystem	N	1	1	Y	Y	N
8145135	LOW	MED	LOW	LOW	SplitSystem	N	1	0	N	Y	Y
8147703	LOW	LOW	LOW	MED		Y	2	0	N	N	N
8198267	LOW	MED	LOW	MED	Ducted	Y	3	0	N	Y	N
8198319	LOW	HI	LOW	MED		N	2	0	Y	N	Y
8198345	LOW	MED	LOW	LOW		N	1	0	N	N	N
8211599	LOW	LOW	LOW	HI	SplitSystem	N	2	0	N	Y	Y
8351602	LOW	HI	MED	MED	SplitSystem	Y	2	0	Y	Y	Y
8376656	LOW	MED	MED	HI	Ducted	N	2	0	Y	Y	Y
8419708	LOW	HI	LOW	LOW	Ducted	N	5	0	N	Y	N
8451629	LOW	MED	LOW	HI	Ducted	Y	2	0	N	Y	N
8459427	LOW	LOW	NONE	NONE			2	1	N	N	Y
8466525	LOW	MED	MED	HI	SplitSystem	Y	3	0	N	Y	Y
8478501	LOW	LOW	LOW	MED	SplitSystem	N	2	0	N	Y	Y
8487285	LOW	LOW	LOW	HI	SplitSystem	N	1	0	N	Y	Y
8487297	LOW	MED	LOW	MED	SplitSystem	Y	1	0	N	Y	Y
8487461	LOW	LOW	LOW	HI	SplitSystem	N	2	0	N	Y	Y
8523058	LOW	HI	LOW	MED	SplitSystem		3	2	Y	Y	Y
8540084	LOW	LOW	NONE	MED	SplitSystem	N	2	0	N	Y	Y
8557605	LOW	HI	LOW	HI		N	1	0	N	N	Y
8566459	LOW	LOW	LOW	MED	SplitSystem	N	2	0	N	Y	Y
8568209	LOW	LOW	LOW	MED		N	1	0	N	N	Y
8655993	LOW	HI	MED	MED	SplitSystem		3	0	Y	Y	Y
8661542	LOW	LOW	HI	LOW	SplitSystem	Y	2	0	N	Y	Y
8673172	LOW	LOW	LOW	MED	SplitSystem	Y	1	0	N	Y	Y
8685932	LOW	HI	LOW	LOW	SplitSystem		2	0	N	Y	Y
8687500	LOW	LOW	MED	LOW		Y	2	0	N	N	Y
8149711	MED	LOW	NONE	MED	SplitSystem	N	2	0	N	Y	Y
8156517	MED	LOW	LOW	MED		Y	1	0	N	N	Y
8184653	MED	LOW	LOW	MED	Ducted	Y	1	0	N	Y	Y
8257054	MED	LOW	LOW	MED	SplitSystem		1	0	N	Y	Y
8264534	MED	LOW	LOW	HI	Ducted	N	1	0	N	Y	Y
8273230	MED	MED	LOW	HI	SplitSystem	N	2	0	Y	Y	Y
8282282	MED	MED	NONE	MED		N	2	0	N	N	Y
8291712	MED	MED	NONE	HI	SplitSystem	N	1	0	N	Y	Y
8308588	MED	HI	LOW	MED	Ducted	N	1	0	N	Y	N
8326944	MED	HI	NONE	NONE	Ducted	Y	2	0	N	Y	Y
8334780	MED	MED	LOW	MED	Ducted	Y	4	0	N	Y	Y
8347238	MED	MED	LOW	MED	SplitSystem	Y	2	0	N	Y	Y
8350006	MED	HI	LOW	MED	SplitSystem	Y	3	0	N	Y	Y
8432046	MED	LOW	LOW	MED	SplitSystem	Y	1	1	N	Y	Y
8482121	MED	LOW	LOW	HI	SplitSystem	N	1	0	N	Y	Y
8496980	MED	LOW	MED	NONE	Ducted	N	1	0	N	Y	Y
8504552	MED	HI	MED	MED	SplitSystem	N	1	0	N	Y	Y
8519102	MED	MED	NONE	HI	SplitSystem	N	1	0	N	Y	Y
8617151	MED	MED	LOW	MED	Ducted	N	1	0	Y	Y	N
8618165	MED	HI	HI	HI	SplitSystem	Y	2	0	N	Y	Y
8679346	MED	LOW	LOW	HI	SplitSystem	Y	1	0	N	Y	Y
8680284	MED	MED	LOW	NONE	SplitSystem	Y	1	0	Y	Y	N

For our project, we consider a subset of SGSC database that comprises of 69 customers with hot water system. Table 6 -1 shows the information about customers

under study. The participants were asked about how much effort they put to reduce the electricity consumption. 23 out of 69 customers responded that they put significantly high effort, whereas 29 customers with medium effort, 11 with low and only 5 customers responded with little to no efforts.

In later sections, we will first discuss about some of the load shaping techniques, then we will discuss about the types of load such as interruptible and non-interruptible. Next we will talk about load scheduling algorithms such as Cuckoo Search and Strawberry Algorithm. In last, we will discuss the experimental setup for load scheduling and some results.

6.3 Experimental Setup

In residential energy management systems, scheduling of different types of loads can only be done after obtaining complete data for the types of load and their characteristics. The utility provides electricity to multiple residential customers with smart meters to record appliance consumption on half-hourly or hourly basis.

6.3.1 Residential Electricity Load

In this experiment we target the reduction in electricity bill of residential customer by load scheduling during the peak event time to off-peak time hence minimizing the peak load as well. A 24-hour window is considered for this experiment where each time slot is of one hour. Each appliance has different power ratings and operating length time (OLT). For this experiment we only consider the electricity imported from the utility. Each residential household is equipped with various appliances that are further categorized into interruptible, un-interruptible and base-load appliances. In this experiment it is assumed that each appliance must complete the OLT.

6.3.2 Load Categorization

We consider a home with multiple appliances, some of which are shiftable and non-shiftable appliances. The shiftable appliances are such appliances that can be shifted to any specific time based on users need and comfort. The interruptible and Un-interruptible appliances falls in the category of shiftable appliances. The un-interruptible appliances have flexible finishing time with certain consumption period such as washing machine and dryer. These appliances may not be interrupted when the execution start. However, these appliances are shiftable to any time slot. On the other hand, the interruptible appliances execution depends on user's choice. For example, vacuum cleaner, microwave and dish washer etc. Finally, we have the base load appliances. These appliances are un-interrupted and non shiftable. For our simulation, we have considered lightening and refrigerator as the base load. The electricity consumption per day of appliances is calculated as,

$$E = \sum_{t=1}^{T} \left(\sum_{a \in A} (\lambda * \alpha) \right)$$

Where α represent the ON / OFF state of the appliance, λ represents the power rating of the appliance and a ∈ A represents an appliance in the list of appliances.

6.3.3 Load Balancing Techniques

The demand side management (DMS) works to help to reduce the electrical consumption in the residential sector and that is done by tracking and monitoring the electricity usage for each householder. There are some techniques to encourage the consumers to use less electricity during on-peak time and to shift the load to the

off-peak time to reduce the electricity cost. Figure 6.7 shows various load shaping techniques.

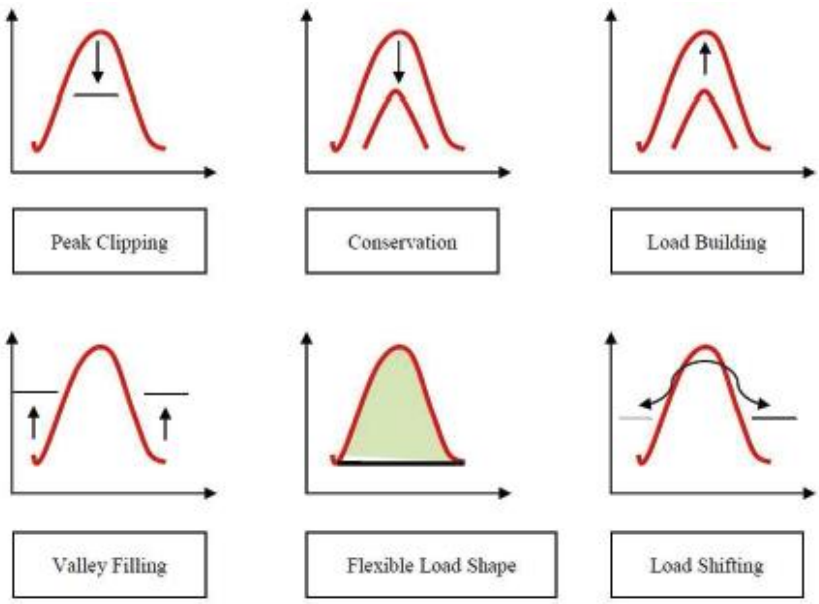

Figure 6.7 Load Shaping Techniques

1. Peak Clipping: Reducing the load during the on-peak time.
2. Conservation: Reducing the load during the entire time.
3. Load Building: Increase the overall power consumption.
4. Valley Filling: Increase the off-peak power usage to decrease the peak at on-peak time.
5. Flexing Load Shape: Making the consumer to reduce the overall power usage by applying different price tariffs.
6. Load Shifting: reducing the power consumption during the on-peak time and increase the power consumption at the off-peak time.

6.3.4 Pricing Scheme

We have considered dynamic peak pricing scheme utilized in the SGSC database trial as shown in Figure 6.8. In this pricing scheme, the price for peak event is increased to 330 cents per KWh.

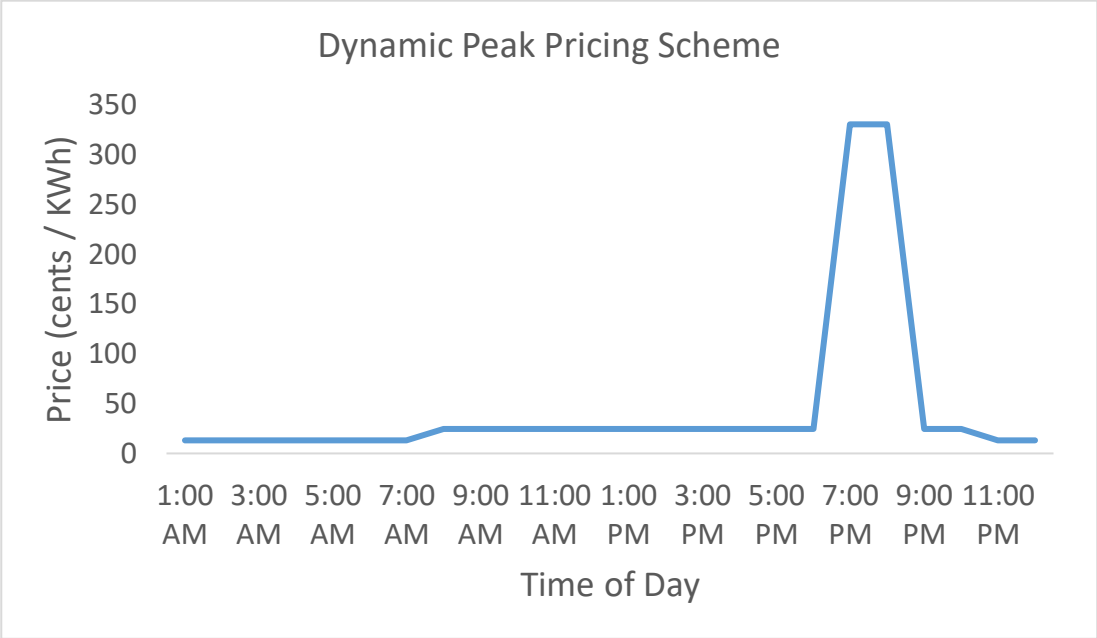

Figure 6.8 Dynamic Peak Pricing Scheme used in SGSC database

6.3.5 Algorithms

In this project, we have implemented two algorithms; strawberry search algorithm (SSA) and cuckoo search algorithm (CSA), for load scheduling to reduce the electricity cost of residential customer as well as to reduce the peak average ratio (PAR). Reduction in the PAR helps utility to preserve the stability and ultimately results in the reduction in electricity cost.

6.3.5.1 Strawberry Search Algorithm (SSA)

The SSA is a nature inspired algorithm. Strawberry plant can model in an effective manner based on three facts including the way strawberry plant propagating by

using their runners which rise randomly to perform their global search for resources; each strawberry parent plant develops its roots and root hairs randomly in order to carry on local search process for resources and finally the strawberry offspring plants have access to richer resources that grow faster and generate more runners and roots.

In the SSA, first we define a certain number of computational agents randomly, each of such agents are imagine as a strawberry mother plant. All the mother plants (strawberry plants) create one root and one runner (daughter plant) at each iteration. In SSA, the roots and runners moves with small and large random steps, respectively. After that the objective function is measured which is cost and PAR reduction, at the points mention by roots and runners, and half of these points are chosen by roulette wheel or elite selection for considering as a mother plant for next iteration while the remaining half will die. This procedure is continuous until a fixed termination condition is hold. Table 6-2 shows set of parameters used in SSA.

Table 6-2 Parameters for SSA

Parameters	Value
Population	100
Length of Runners	50
Length of Roots	10
Number of Iterations	2000
Number of Appliances	12

6.3.5.2 Cuckoo Search Algorithm (CSA)

CSA Algorithm is a Meta-Heuristic Algorithm and it's inspired by some cuckoo species laying their eggs in the nest of other species birds. There are two type of birds

in this algorithm first cuckoo bird second is Host Bird (other species). The behavioral of cuckoo bird is that they don't have nest so they lay their eggs in host bird Nest. If host Bird discovered cuckoo eggs there is two action: the host bird can throw the egg away, or build a new nest. We assume that the discovering probability is 0.25. Table 6-3 shows set of parameters used in SSA. In order to find the global best, Levy flight (random walk) is used to generate the new solution.

The CSA finds the best solution through following rules:

1. Every cuckoo lay only one egg in the randomly chosen nest.

2. Only the high-quality nests with the high-quality eggs are considered for the next generation.

3. Host nests are fixed and the host birds discover the eggs.

So, each egg represent a solution and eggs in nest represent set of solution And Cuckoo egg represent new solution. High quality eggs represent the Best solution near optimal value. New and better solution will replace less fit solution. Host bird discover cuckoo egg represent the worst solution.

Table 6-3 Parameters for CSA

Parameters	Value
Host nests	50
Discovery Rate	0.25
Number of Iterations	2000
Number of Appliances	12

6.3.6 Results

In this experiment, we perform the load scheduling of 12 appliances listed in Table 6-4. The appliances are categorized in three types. In this table, we have provided the

name of appliance, the type of appliance, power rating, possible earliest starting time of appliance, possible least finishing time of appliance and operating length time.

Table 6-4 Appliance Parameters for Simulation

Appliance Category	Name of Appliance	Power (kW)	Possible Earliest Starting Time (h)	Possible Least Finishing Time (h)	operating length time OLT (h)
interruptible	Cooker Hub	3.0	06	10	1
	AC	1.2	08	18	6
	Microwave	1.7	06	10	1
	Computer	0.3	20	24	2
	Water Heater	3.6	23	08	4
	Vacuum Cleaner	0.7	08	13	1
	Electric Car	3.5	20	24	3
	Dish Washer	1.8	20	10	2
Un-interruptible	Washing Machine	1.5	08	12	2
	Spin Dryer	2.5	13	17	1
Base-Load	Lightening	0.84	16	24	8
	Refrigrators	0.3	1	24	24

Since we know that the DPP peak event cost is extremely high, which is 330 cents per KWh. We have to schedule the load such that the cost of electricity as well PAR is reduced. The advantage of individual household load forecasting (Chapter 5) is that it will help the utility as well as customer to forecast the consumption of time ahead and shift it to off peak time. For example, Figure 6.9 shows day ahead forecast of a residential customer ID 10509861 from SGSC database who is using major appliances listed in above table. As shown in figure, there is a peak consumption at on-peak time. If the customer was informed day ahead about his forecast and about peak event. Then the customer could plan to schedule the load accordingly.

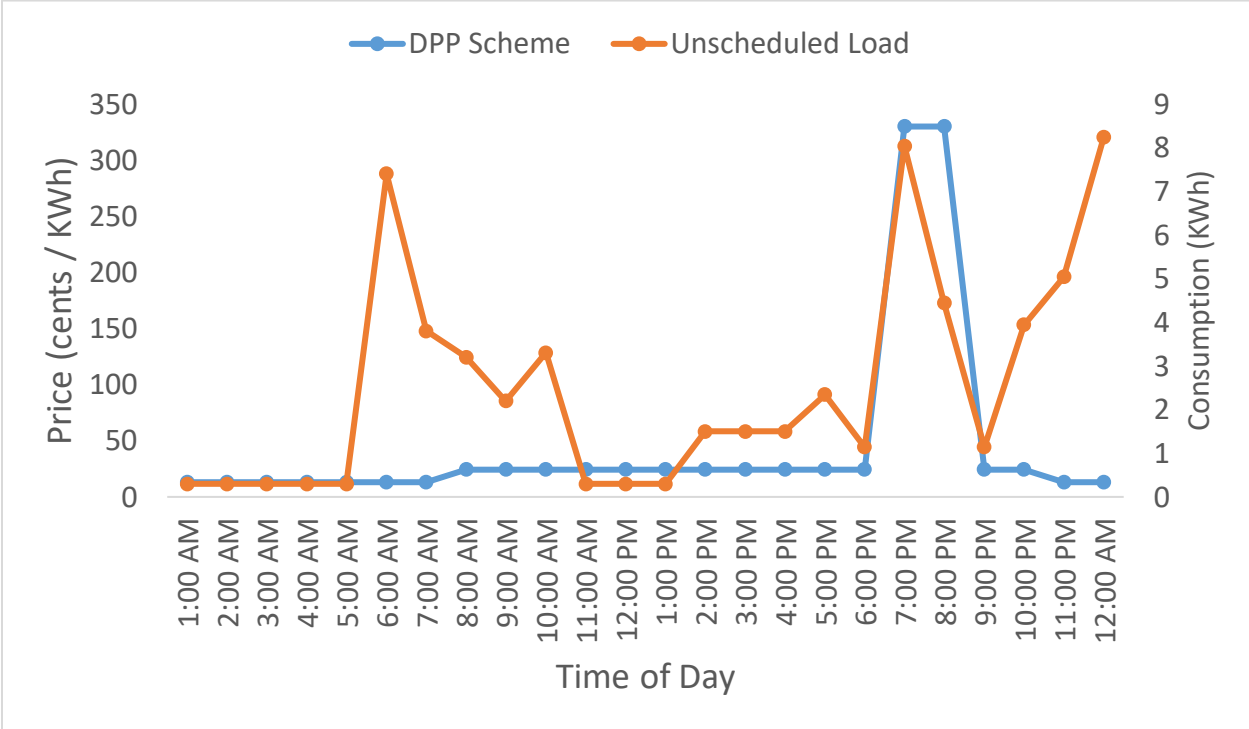

Figure 6.9 DPP Scheme along with Unscheduled Load (day ahead forecast)

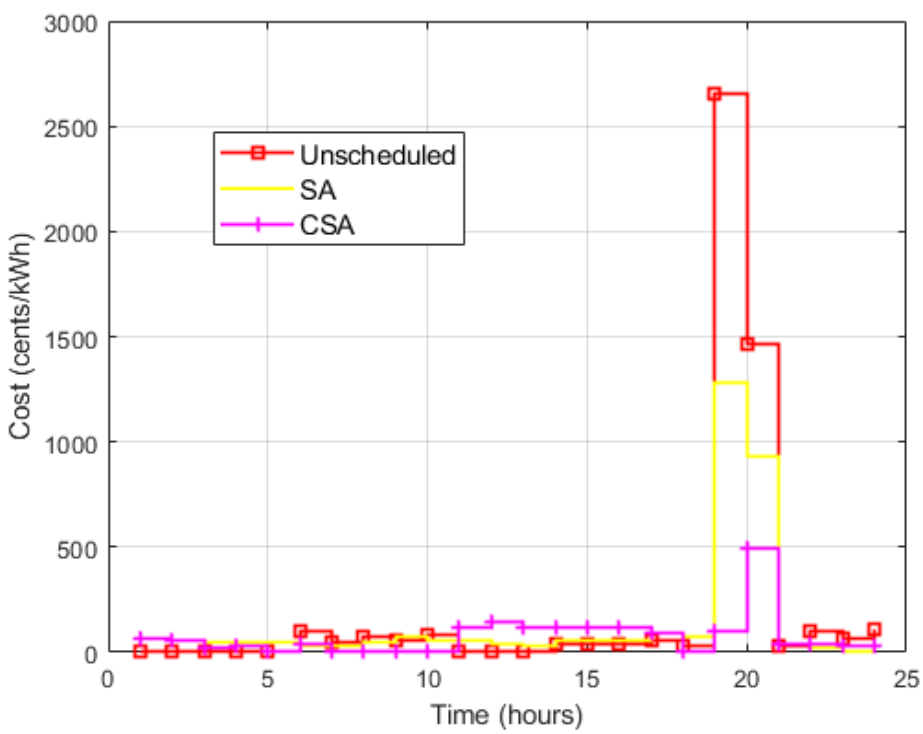

Figure 6.10 Electricity cost (Hourly basis)

Figure 6.10 shows the hourly electricity cost of unscheduled load, scheduled load using SA as well as CSA. As seen in the Figure, the cost during the peak event time is drastically reduced using both search algorithm. However, CSA outperforms SA during the peak event time.

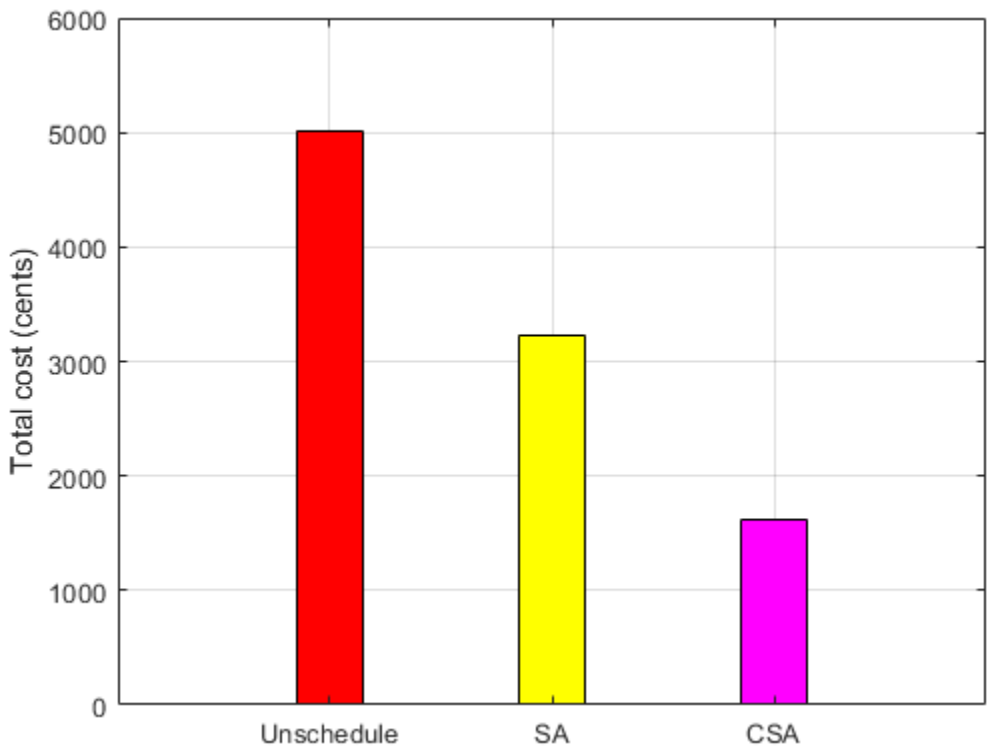

Figure 6.11 Total Cost comparison of unscheduled with SA and CSA

Figure 6.11 shows the total cost of electricity for a single day. It is observed from the figure that CSA and SA both reduces the electricity cost. However, the cost reduction is maximum with CSA which is around 64% whereas for SA is around 38%.

Figure 6.12 shows the unscheduled load as well as load scheduling using SA and CSA. It is observed from the figure that CSA and SA both reduces the peak consumption during the peak event time and shifting the load to off-peak time. In the unscheduled load, the peak consumption is around 8.5kW whereas the peak load is reduced to around 5.2kW and 6kW using SA and CSA respectively.

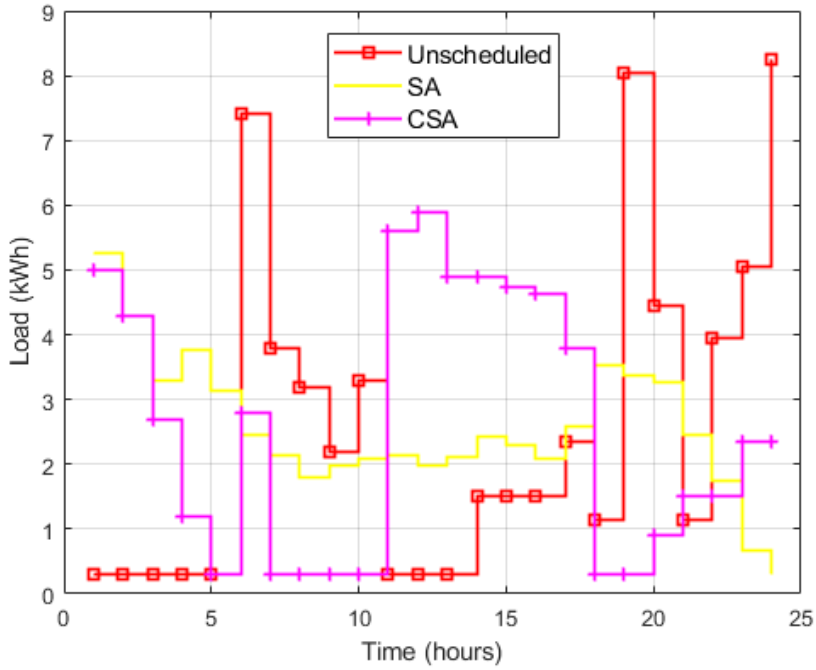

Figure 6.12 Unscheduled Load Consumption (Hourly basis) as well as Load Scheduling using SA and CSA

Both search algorithms efficiently schedule the load while minimizing the PAR as shown in Figure 6.13. PAR can only be reduced by avoiding peak formation in every hour.

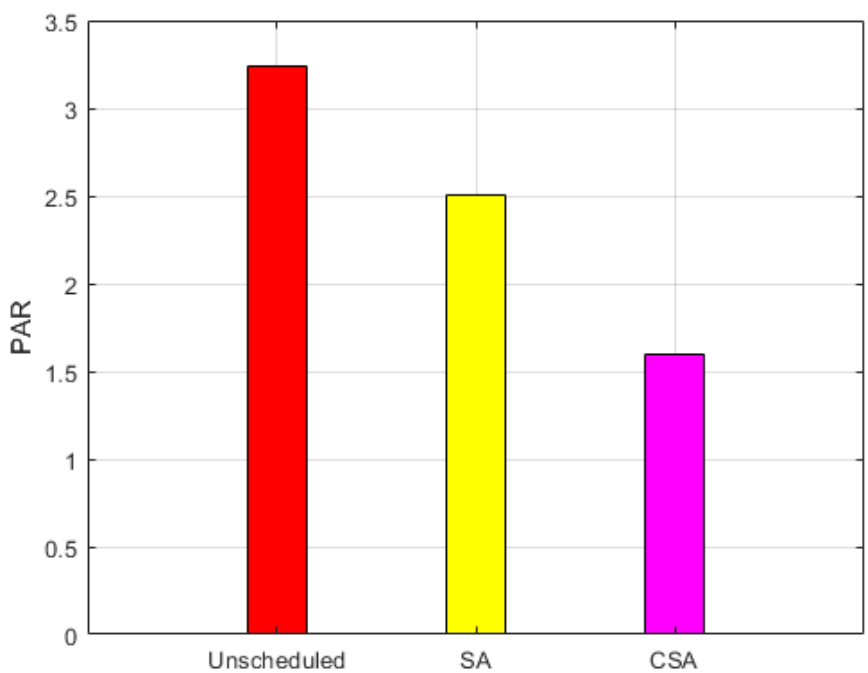

Figure 6.13 Peak Average Ratio (PAR) for unscheduled, SA and CSA

6.4 Future Work

We have focused in our project to analyse the power consumtion from the demand side of the electrical power system but we have also seen that the Smart Grid Smart City will establish two-way communication that can help more in monitoring and controlling the power from the production side. Recently, Saudi Electric Company is installing smart meters in all the residential buildings. It will help in establishing a two way communication between SEC and residential customer. we suggest for the future work in:

- Smart meters for individual house holder.
- Combining an altinative power source (PV panels, wind turbines) to reduce the power import from the grid during the peak event time
- Smart meter with power storage units.

6.5 Conclusion

In the second term of this project, we first developed a hybrid CNN-LSTM model for individual household load forecast. We select the SGSC database where dynamic peak pricing was utilized. The proposed deep learning model will help in forecasting day ahead electricity consumption of residential customer. The utility company will notify the customer about their forecasted consumption behavior as well as peak event one day prior. With the help of CSA and SA, we can reschedule the load to off peak time thus reducing the cost of electricity bill as well as minimizing the PAR which will be beneficial for utility provider. From the results we conclude that the

performance of CSA is better than SA because CSA spends most of the time in global search.

There is always a trade-off between user comfort and electricity cost. We can be more flexible with the possible earliest start time of appliance or possible least end time of appliance mentioned in Table 6-4. However, it would result in more comfort to user but less cost reduction.

Appendix:

Code of ethics:

this project meet the IEEE code of ethics and we picked some rules that related to our project. and we used open dataset and available for education purposes

1. This project hold safety, health and welfare of public, striving to comply with ethical design and sustainable development practice, and disclosing promptly factors that might endanger public.
2. All case, for example, budget, are founded on accessible information.
3. Briberies are dismissed for all situation.
4. The shirking of perceived conflicts of interest whatever point its possible.
5. Taking risk of arising technology for example intelligent system.
6. Understand technology tasks which qualified via training or experience.
7. This project for educational reason any techniques possibly have some mistake or change. To limit blunders. We asked our advisor to improve our insight.
8. We treat with all people fairly, don't founded on religion, national origin and gender.
9. We avoid any sort of harming of others by false or noxious activity.
10. Help of collaborators inside group while in expert advancement to help them in this code of ethic.

References

[1] Saudi Energy Efficiency Center. (2018, March 13). Retrieved from seec.gov.sa/en/blog/buildings

[2] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network," IEEE Transactions on Smart Grid, vol. 10, no. 1, Jan 2019.

[3] What is Smart Grid? Definitions, Perspectives, and Ultimate Goals. Tarbiat Modares University (TMU) 2013.

[4] P. Lusis, K. R. Khalilpour, L. Andrew, A. Liebman, "Short-term residential load forecasting: Impact of Calendar Effects and Forecast Granularity", Applied Energy, vol. 205, pp. 654 – 669, 2017.

[5] Y. K. Penya, C. E. Borges, D. A, and I. Fernandez, "Short-term load forecasting in air-conditioned non-residential buildings", pp. 1359 – 1364, 2011.

[6] N. Charlton, C. Singleton, "A refined parametric model for short term load forecasting", International Journal of Forecast, vol. 30, no. 2, pp. 364 – 368, 2014.

[7] B. Nepal, M. Yamaha, A. Yokoe, and T. Yamaji, "Electricity load forecasting using clustering and ARIMA model for energy management in buildings", Japan Architectural Review, vol. 3, no. 1, pp. 62 – 76, 2020.

[8] N. C. Truong, J. McInerney, L. T-Thank, E. Costanza, and S. D. Ramchurn, "Forecasting multi-appliance usage for smart home energy management", In Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, pp. 2908 – 2914, 2013.

[9] S. J. Kiartzis, A. G. Bakirtzis, J. B. Theocharis, and G. Tsagas, "A fuzzy expert system for peak load forecasting, application to the greek power system", In Electro technical Conference, vol. 3, pp. 1097 – 1100, 2000.

[10] A. Khotanzad, R. A-Rohani, T.-L Lu, A. Abaye, M. Davis, D. J. Maratukulam, "ANNSTLF – A neural network based electric load forecasting system", IEEE Transaction on Neural Network, vol. 8, pp. 835 – 846, 1997.

[11] P. Mandal, T. Senjyu, N. Urasaki, T. Funabashi, "A neural network based several hour ahead electric load forecasting using similar day approach",

International Journal of Electrical Power Energy System, vol. 27, pp. 367 – 373, 2006.

[12] H. Chen, C. A. Canizares, and A. Singh, "ANN based short-term load forecasting in electricity markets", In Power Engineering Society Winter Meeting, 2001, IEEE, vol. 1, pp. 217 – 235, Hong Kong, 2013.

[13] E. A. Feinberg, and D. Genethliou, "Load Forecasting", Applied Mathematics for Restructured Electric Power Systems, Power Electronics and Power Systems, pp. 269 – 285, 2005.

[14] L. Li, K. Ota, and M. Dong, "Everything is Image: CNN-based Short-Term Load Forecasting for Smart Grid", vol. 1, pp. 344 – 351, 2017.

[15] F. M. Bianchi, E. Maiorino, M. C. Kampffmeyer, R. Jenssen, and A. Rizzi, "Recurrent Neural Networks for Short-Term Load Forecasting: An Overview and Comparative Analysis", 2017.

[16] H. Li, Y. Zhao, Z. Zhang, and X. Hu, "Short-term load forecasting based on the grid method and the time series fuzzy load forecasting method," In Proc. Int. Conf. Renewable Power Generation (RPG), Beijing, China, pp. 1–6, 2015.

[17] X. Cao, S. Dong, Z. Wu, and Y. Jing, "A data-driven hybrid optimization model for short-term residential load forecasting," in Proc. IEEE Int. Conf. Comput. Inf. Technol. Ubiquitous Comput. Commun. Dependable Auton. Secure Comput. Pervasive Intell. Comput. (CIT/IUCC/DASC/PICOM), Liverpool, U.K., pp. 283–287, 2015.

[18] P. Qingle and Z. Min, "Very short-term load forecasting based on neural network and rough set," in Proc. Int. Conf. Intell. Comput. Technol. Autom. (ICICTA), Changsha, China, pp. 1132–1135, 2010.

[19] R. Zhang, Z. Y. Dong, Y. Xu, K. Meng, and K. P. Wong, "Short-term load forecasting of Australian national electricity market by an ensemble model of extreme learning machine," IET Generation Transmission and Distribution, vol. 7, no. 4, pp. 391–397, Apr. 2013.

[20] M. Chaouch, “Clustering-based improvement of nonparametric functional time series forecasting: Application to intra-day household-level load curves,” IEEE Trans. Smart Grid, vol. 5, no. 1, pp. 411–419, Jan. 2014.

[21] M. Ghofrani, M. Hassanzadeh, M. Etezadi-Amoli, and M. S. Fadali, “Smart meter based short-term load forecasting for residential customers,” in Proc. North Amer. Power Symp. (NAPS), Boston, MA, USA, pp. 1–5, 2011.

[22] S. Ryu, J. Noh, and H. Kim, “Deep neural network based demand side short term load forecasting,” in Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm), Sydney, NSW, Australia, pp. 308–313, 2016.

[23] E. Mocanu, P. H. Nguyen, M. Gibescu, and W. L. Kling, “Deep learning for estimating building energy consumption,” Sustainable Energy Grids Network, vol. 6, pp. 91–99, Jun. 2016.

[24] D. L. Marino, K. Amarasinghe, and M. Manic, “Building energy load forecasting using deep neural networks,” in Proc. 42nd Annual Conf. IEEE Ind. Electron. Soc. (IECON), Florence, Italy, pp. 7046–7051, 2016.

[25] Langham, E., Downes, J., Brennan, T., Fyfe, J., Mohr, S., Rickwood, P. and White, S. 2014. *Smart Grid, Smart City, Customer Research Report.* Report prepared by the Institute for Sustainable Futures as part of the AEFI consortium for Ausgrid and Energy Australia, June 2014.

[26] Nathan. W, Residential Electricity Usage: How Do Households Respond to Dynamic Peak Pricing Events? Honors Thesis.