



REPUBLIC OF TÜRKİYE
ALTINBAŞ UNIVERSITY
Institute of Graduate Studies
Information Technologies

**ENERGY CONSUMPTION ANALYSIS VIA
DIFFERENT MACHINE LEARNING
ALGORITHMS**

Rwan DARWESH

Master's Thesis

Supervisor

Asst. Prof. Dr. Hakan KOYUNCU

Istanbul, 2023

ENERGY CONSUMPTION ANALYSIS VIA DIFFERENT MACHINE LEARNING ALGORITHMS

Rwan DARWESH

Information Technologies

Master's Thesis

ALTINBAŞ UNIVERSITY

2023

The thesis titled ENERGY CONSUMPTION ANALYSIS VIA DIFFERENT MACHINE LEARNING ALGORITHMS prepared by RWAN DARWESH and submitted on 28/04/2023 has been **accepted unanimously** for the degree of Master of Science in Information Technologies.

Asst. Prof. Dr. Hakan KOYUNCU

Supervisor

Thesis Defense Committee Members:

Asst. Prof. Dr. Hakan KOYUNCU

Department of Computer
Engineering,

Altınbaş University

Asst. Prof. Dr. Abdullahi Abdu
İBRAHİM

Department of Computer
Engineering,

Altınbaş University

Asst. Prof. Dr. Hakan AYDIN

Department of Computer
Engineering,

Topkapı University

I hereby declare that this thesis meets all format and submission requirements for a Master's Thesis.

Submission date of the thesis to Institute of Graduate Studies: ___/___/___

I hereby declare that all information/data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Rwan DARWESH

Signature



ABSTRACT

ENERGY CONSUMPTION ANALYSIS VIA DIFFERENT MACHINE LEARNING ALGORITHMS

DARWESH, Rwan

M.Sc., Information Technology, Altınbaş University,

Supervisor: Asst. Prof. Dr. Hakan KOYUNCU

Date: April / 2023

Pages: 59

The amount of energy used by buildings is increasing as a consequence of increased urbanization and social advancement. Predicting a building's energy needs is essential for promoting sustainable growth and energy efficiency, which in turn reduces energy costs and has a lesser impact on the environment. This research focuses on the topic of applying deep learning (DL) techniques to forecast energy use across time series using actual data. The performance of statistical and DL algorithms was evaluated using data collected in real time from a smart grid installed in an experimental building. Usage of energy in ensemble and single situations was examined using well-known artificial intelligence techniques. The models which combine prediction and optimization approaches are examined in-depth. The thorough comparative analysis demonstrated that the hybrid model was excellent in performance than the single and ensemble models in terms of accuracy.

These models are thought to be suitable for usage and accurate enough to provide predictions, which can help users plan their energy management strategies.

Keywords: Energy Consumption, Artificial Intelligence, Data Mining, Time-Series Forecasting, Machine Learning, Residential Buildings.

TABLE OF CONTENTS

	<u>Pages</u>
ABSTRACT	v
LIST OF FIGURES.....	ix
LIST OF ABBREVIATIONS.....	xi
1. INTRODUCTION	1
1.1 OVERVIEW	1
1.2 RESEARCH SIGNIFICANCE.....	4
1.3 RESEARCH OBJECTIVES	5
1.4 RESEARCH QUESTIONS	5
1.5 STRUCTURE OF THE THESIS.....	5
2. LITERATURE REVIEW.....	7
2.1 OVERVIEW	7
2.2 AN HISTORICAL PERSPECTIVE OF BUILDING ENERGY CONSUMPTION ...	7
2.3 PREDICTIONS OF BUILDING ENERGY DEMAND	8
2.4 ARTIFICIAL NEURAL NETWORK (ANN).....	9
2.5 RECURRENT NEURAL NETWORK (RNN)	9
2.6 LONG SHORT-TERM MEMORY NEURAL NETWORK (LSTM)	10
2.7 ENERGY CONSUMPTION IN BUILDINGS.....	10
2.8 AI METHODS FOR TIME-SERIES ENERGY CONSUMPTION FORECASTING.....	12
2.9 SUMMARY.....	15
3. METHODOLOGY	16
3.1 INTRODUCTION	16
3.2 METHODOLOGY	16
3.3 NETWORK OF LSTM, RNN AND CNN	17

3.4	AMERICAN ELECTRIC POWER (AEP).....	23
3.4.1	Data Loading and Data Exploration	23
4.	RESULTS AND DISCUSSION.....	25
4.1	DATA NORMALIZATION.....	26
4.1.1	Data Visualization Following Normalizing.....	26
4.1.2	Preparation of Data to Implement the RNN Models Training.....	27
4.1.3	LSTM Model	28
4.1.4	Comparison Between RNN and LSTM Model Predictions.....	29
4.2	DOMINION VIRGINIA POWER (DOM).....	30
4.2.1	Data Loading and Data Exploration	30
4.2.2	Data Normalization.....	32
4.2.2.1	Data visualization following normalizing.....	32
4.2.3	Preparation of Data to Implement the RNN Models Training.....	33
4.2.4	LSTM Model of DOM Power Consumption Dataset	34
4.2.5	Comparison Between RNN and LSTM Model Predictions for DOM Power Consumption Dataset	36
4.3	COMMONWEALTH EDISON (COMED)	37
4.3.1	Data Normalization.....	38
4.3.2	Preparation of Data to Implement the RNN Models Training.....	39
4.3.3	LSTM Model of COMED Power Consumption Dataset.....	40
4.3.4	CNN model of COMED Power Consumption Dataset.....	41
4.3.5	Comparison Among RNN, LSTM And CNN Model Predictions for COMED Power Consumption Dataset.....	42
5.	CONCLUSIONS.....	44
	REFERENCES	45
	APPENDIX A.....	50

LIST OF TABLES

	<u>Pages</u>
Table 3.1: Labeled Date Time and American Electric Power (AEP) Dataset with The Unite Of MW.....	23
Table 3.2: Indexed Via Datetime Dataset of American Electric Power (AEP) Dataset with The Unit Of MW.	24
Table 4.1: Labeled Date Time and Dominion Virginia Power (DOM) Dataset with The Unite Of MW.....	31
Table 4.2: Indexed Via Datetime Dataset of Dominion Virginia Power (DOM) Dataset with The Unit Of MW.	31
Table 4.3: LSTM Model Properties for The Dataset of DOM Power Consumption Dataset. .	35

LIST OF FIGURES

	<u>Pages</u>
Figure 2.1: Time-Series Breakdown to Discover Energy Usage Components [43].....	14
Figure 3.1: Structure of Recurrent Neural Network (RNN) and Unrolled Structure of RNN In Time.	18
Figure 3.2: Visualization of RNN.....	18
Figure 3.3: Structure of a Long Short-Term Memory Cell (LSTM).	19
Figure 3.4: Visualization of LSTM	20
Figure 3.4: Structure of a CNN (Convolutional Neural Network)	21
Figure 3.5: CNN Data Preparation.	22
Figure 3.6: Visualization CNN.....	22
Figure 4.1: Pre-Normalization Power Consumption Data.....	25
Figure 4.2: Normalized Power Consumption Data Vs Datetime.	27
Figure 4.3: RNN Model Prediction Vs Real Power Consumption Data.	28
Figure 4.4: LSTM Model Prediction Vs Real Power Consumption Data.....	29
Figure 4.5: Comparison Between RNN and LSTM Model Predictions.....	30
Figure 4.6: Pre-Normalized Dom Power Consumption Data.....	32
Figure 4.7: Normalized DOM Power Consumption Data Vs Date Time.....	33
Figure 4.8: RNN Model Prediction Vs Real Power Consumption Data Of DOM.....	34
Figure 4.9: LSTM Model Prediction Vs Real Power Consumption Data for DOM	36
Figure 4.10: Comparison Between RNN and LSTM Model Predictions.....	37
Figure 4.11: Pre-Normalized Comed Power Consumption Data	38

Figure 4.12: Normalized COMED Power Consumption Data Vs Date Time.....39

Figure 4.13: RNN Model Prediction Vs Real Power Consumption Data Of COMED.....40

Figure 4.14: LSTM Model Prediction Vs Real Power Consumption Data for COMED41

Figure 4.15: CNN Model Prediction Vs Real Power Consumption Data for COMED.42

Figure 4.16: Comparison Among RNN, LSTM And CNN Model Predictions43



LIST OF ABBREVIATIONS

- AI : Artificial Intelligence
- IoT : Internet of Things
- DOM : Dominion Virginia Power
- RNN : Recurrent Neural Networks
- CNN : Convolutional Neural Network



1. INTRODUCTION

1.1 OVERVIEW

Global growth is accelerating at a significant rate, which will inevitably result in higher energy consumption. Population expansion, rising standards of living, and shifting climatic conditions have all contributed to the building sector's rise to the top as the world's greatest user of energy [1]. To be more precise, building construction accounts for around 32% of global energy consumption and as much as 40% in several industrialized nations (including about in Europe and 39% in the United States 40%) [2, 3]. Furthermore, a rise in use of building energy is consequence of the quick development of residential and commercial sectors. Furthermore, in many countries, new laws, and standards for the efficient design of new buildings have been enacted with the goal of attaining building energy conservation. As a result, increasing the efficiency with which buildings use energy has emerged as a pressing concern worldwide. Due to its central role in numerous building energy management tasks, such as optimal operation strategy control [5], fault detection and diagnosis [4], and management of demand-side [6], accurate forecasting of energy consumption for building is increasingly desirable and crucial for enhancing building energy efficiency. It has been demonstrated in the study of Biswas et al. [7] that accurate projections of a building's energy consumption may lead to savings of 10% to 30%, highlighting the crucial role that energy consumption forecasting plays in enhancing building energy efficiency. Predicting a building's energy consumption using current methods may be categorized into three broad fields: the statistical approach, the engineering approach, or the artificial intelligence (AI) based approach [8]. White-box engineering approaches, which depend on complex physical functions and thermodynamic laws, need the creation and solution of several physical equations to estimate the energy behaviors of buildings. In addition, engineering calculations rely on a wide range of building characteristics, such as design specifics, thermal qualities of materials, weather, and occupancy levels, that are not always accessible. Because of these challenges and constraints, engineering-based approaches to estimating building energy usage are seldom used in practice [9]. In case of statistical approach, the authors only link the energy consumption of buildings with key input data (such as weather variables), and various flaws in practice have been observed, the most prominent of which is a lack of precision

and flexibility [10]. In contrast, AI-based methods are data-driven and can function with less specialized expertise. It is the goal of such data-driven approaches to learn from past patterns of energy usage and make predictions about future energy needs. In particular, they make an effort to create prediction models in a supervised way by identifying and generalizing the underlying linear or nonlinear connection between inputs (such as past energy data and weather data) and outputs such as consumption of building energy). Data-driven models are easier to implement because to the abundance of data generated by the Building Automation System and the burgeoning field of data science. Due to their adaptability and effectiveness, data-driven approaches have recently been a focus of investigation [11]. There are three main aspects that determine how well a data-driven model predicts the future, and they are the accuracy of the collected data, the variables chosen for inclusion in the model, and the prediction algorithms used to create the model [12]. Data-driven model creation includes the selection of a suitable prediction method. As a result, scientists have focused their attention on creating more accurate models that use less processing power. Supervised machine learning algorithms, which may be further subdivided into two groups, namely, classical machine learning techniques and deep learning methods, are the most extensively used approaches in predicting building energy usage [22]. Multiple Linear Regression, Support Vector Machines, Artificial Neural Networks, and Decision Trees, as well as its variants, make up the bulk of the classic machine learning library. Neto and Fiorelli, [23] used an ANN method to anticipate the administrative building's energy usage, and their findings demonstrated that an ANN-based model can provide more accurate predictions than simulation software. According to a study by Li et al. [24], SVM outperformed a back-propagation neural network in predicting the hourly cooling demand of a building (BPNN). Predictions of energy consumption in buildings using the DT algorithm developed by Yu et al. [25] have shown an accuracy of up to 92% in testing data. The use of ensemble learning methods has been advocated to increase the stability of models. Predictions of ground source heat pump loads are studied using ensemble models based on data decomposition in Ref. [26]. The results of the predictions demonstrated that the ensemble models significantly improved the accuracy of the predictions. Hourly power usage in buildings was predicted using a method called Ensemble Bagging Trees (EBT) by Wang et al. [9]. The EBT model in this ensemble employed DT as its foundation, and it generated its predictions by averaging the findings

generated by the individual DT models in the ensemble. Based on the findings, the suggested ensemble model is more reliable and accurate than individual prediction models. However, developing these ensemble models needed more time and processing power.

Most of the aforementioned prediction techniques use "shallow" models, which results in weak feature extraction from the raw data. An advancement of ANN, deep learning uses many processing layers to automatically discover appropriate representations of raw inputs, therefore addressing a fundamental shortcoming of conventional machine learning methods [27]. Researchers have also concentrated a lot of efforts to deep learning for use in predicting how much energy a building would need. Berriel et al. [28] utilized three different deep learning algorithms namely Convolutional, Deep Full Connected, and Long Short-Term Memory neural networks to make monthly predictions about the energy usage of buildings. Using a deep brief network and an ensemble method, the researcher was able to provide competitively accurate predictions for cooling load [29]. To predict the long-term power needs of buildings, Rahman [30] created and improved deep Recurrent Neural Network (RNN) models, which he found to be superior to the more traditional 3-layer perception neural network.

Building energy efficiency can have significant economic advantages in addition to environmental advantages. The expenses regarding operations of buildings with effective energy management and technologies are much cheaper. The adoption of energy rules and norms for various types of buildings has increased in a few nations. In addition, several different pieces of software have been developed and see regular usage in the design of energy-efficient structures.

Globally, energy usage was predicted to be 700 quadrillion Btu in 2035, that was 505 British thermal units (btu) in 2008 [1]. Energy norms and requirements for various kinds of buildings have been advanced in a number of nations. These standards establish fundamental needs to guarantee energy-efficient building systems with the goal of lowering energy usage and associated Emissions of carbon dioxide [2]. Energy management systems should be operated according to the proper operational methods to enhance the energy performance of buildings. In addition to managing the variables that affect the energy performance of their buildings, managers must manage and constantly monitor the time series of energy. The continuous control and monitoring of energy use

includes prediction as a key component. Improving building energy efficiency requires accurate energy usage forecasting [3]. A time-series forecasting model learns from past energy consumption usage trends, so if energy consumption is somewhat higher than expected over time, facility managers and energy systems may be alerted to the change.

Considering all of the above-mentioned issues and prospect of predicting consumption of energy of a building, this research presents a comprehensive evaluation of ML techniques for predicting energy use over time. The study used three one of the emerging algorithms like LSTM, RNN and CNN for predicting the energy consumption of a residential building. Finally predicted data was validated comparing with the real data.

1.2 RESEARCH SIGNIFICANCE

Constant increases in energy output are needed to keep up with the demands of rapidly increasing high-density cities and their accompanying commercial and industrial sectors [31]. It is difficult to accurately anticipate a building's energy consumption because of the various variables that affect it, such as the weather, building construction, location, occupancy, etc. Prediction issues around energy use have received considerable attention during the last two decades [32]. In this study, a new model based on the LSTM method for reliable and precise short-term load predictions. The proposed model is tested, and the results are compared to those of the exponential smoothing and the ARIMA models in terms of their Mean Absolute Error (MAE) values. Because of their capacity to memorize vast amounts of information, LSTM models are well-suited for use in time series-based forecasting. Since residential load demand is very variable and unexpected, LSTM performs better than competing models [33]. Hence, this research will focus on the development of a model for the prediction of energy consumption by using different emerging machine learning algorithms using real data. With the use of real-time data collected from a smart grid installed in an experimental building, the efficiency and efficacy of statistical and DL algorithms will be evaluated. Energy usage in single and ensemble situations will be examined using well-known artificial intelligence techniques. The models which combine forecasting and optimization approaches will be examined in-depth.

1.3 RESEARCH OBJECTIVES

In the past, numerous research work was conducted to predict the energy consumption of building in different ways; however, robust ML techniques and wider range of datasets is required for the development of trustworthy knowledge for predicting energy consumption. ML aims to create models that, after learning from specific training datasets, are capable of making quantifiable, accurate predictions. The following objectives are outline below:

- a. To conduct a comprehensive literature review regarding machine learning algorithm for predicting energy consumption of a building.
- b. To assess the effectiveness and usefulness of statistical and ML algorithms using real-time data gathered from a smart grid placed in an experimental building.
- c. To examine usage of energy in single and ensemble situations using well-known artificial intelligence techniques.

1.4 RESEARCH QUESTIONS

In this study, the following questions to be answered in detail.

- a. Which machine learning algorithms will be used for predicting energy consumption in an efficient way?
- b. How will the algorithm perform the forecasting model?
- c. In which way the usefulness and effectiveness of statistical and ML algorithms will be assessed?
- d. How forecasting and optimization approach will be examined?

1.5 STRUCTURE OF THE THESIS

This dissertation will follow the following format to ensure that the aforementioned goal is met:

Chapter 1: This chapter describes about the background in brief, research aim, research objectives, and research questions of the dissertation.

Chapter 2: This chapter covers the rigorous literature review with an overview of the factors affecting the energy consumption, different types of machine learning algorithms and their effectiveness.

Chapter 3: This section describes the whole process of this study from the beginning to end. This chapter outlines the techniques utilized to predict the energy consumption and examine if the methodological operationalization of these for the aforementioned goals was viable. In order to achieve the goals of this research, the tools for method analysis that are now accessible are also briefly presented. The quantitative research method was described here briefly.

Chapter 4: It deals with the analysis and results to show the accuracy and effectiveness of different machine learning algorithms.

Chapter 5: This chapter summarizes and points out the major findings of the current study and suggests some recommendations for future research.

2. LITERATURE REVIEW

2.1 OVERVIEW

This chapter provides a state of the art of building energy performance simulation techniques using various modern information technologies such as machine learning, deep learning, and artificial intelligence. A review of the pertinent literature is first carried out in this chapter which represents the material brief history of energy consumption, tools for the prediction of energy consumption, ANN, RNN and LSTM. Special attention is paid in different modern energy simulation tools for energy performance of the building sector. Finally, this chapter reviews the most significant existing study building energy performance techniques.

2.2 AN HISTORICAL PERSPECTIVE OF BUILDING ENERGY CONSUMPTION

The global economy is expanding at a rapid pace, and this will inevitably need more power. A further factor in the rising demand for energy from buildings is the fast development of new residential and business areas. The ability to predict energy use is now a vital part of any energy consumption estimation process. Environmental advantages aren't the only ones that may be gained from a more energy-efficient structure. Lower utility bills are a major benefit of energy-efficient buildings. Several nations' efforts to enforce energy laws and norms for buildings have been stepped up in recent years. In addition, several programs have been created and are frequently used for planning greener structures. An increasing number of nations are swiftly enforcing different sorts of building energy norms and laws. These rules provide minimum criteria for ensuring that the designs of new buildings are energy efficient, with the goal of decreasing energy consumption and the resulting CO₂ emissions [1]. Energy efficiency at the end-use level has been shown to dramatically cut global energy consumption [2]. Consumers and utilities alike may benefit greatly from even modest reductions in peak demand [3]. Therefore, it is essential to increase building energy efficiency in order to reduce environmental effects and stabilize energy prices. Predicting energy use is essential for optimizing a building's energy performance [4]. If energy consumption during a period is somewhat greater than

expected, the time-series forecasting model may alert facility management and energy systems.

Different prediction methods have been applied during the past 20 years to determine how much energy buildings would use [6]–[10]. They may be divided into three groups: artificial intelligence methods, statistical methods, and engineering approaches. Artificial intelligence (AI) techniques, which include Artificial Neural Networks (ANNs) [8], [11], the Classification and Regression Tree (CART) [12], the Support Vector Machine (SVM) [6], the Linear Regression (LR) [13], the autoregressive integrated moving average (ARIMA) [14]–[16], and the seasonal autoregressive integrated moving average (SARIMA) methods [17] are the most often used among them.

2.3 PREDICTIONS OF BUILDING ENERGY DEMAND

It is important to manage and operate a facility in a way that minimizes energy usage and associated expenses. In brief, we need to maximize efficiency in both our energy use and production to reduce waste. Technology that can estimate future energy needs is crucial for this goal [5]. Firstly, the most popular kind of method is statistical analysis, and they may be used to accurately anticipate energy usage from large data sets. Load modeling and energy consumption forecasting are often performed with multiple regression or probability distribution analysis.

The second most common approach is using energy simulation software to digitally recreate a building's energy consumption pattern by inputting specific information about the structure, modeling the facility, and testing different scenarios. In order to get the virtual energy consumption statistics, a lot of information about the building must be entered, and then the heat exchange will be calculated based on the weather, ventilation, and immersion circumstances, using an internal simulation program. Dynamic energy simulations are often performed using software like EnergyPlus, TRNSYS, Design Builder, and ESP-r. As a last common approach, we have the usage of AI, or the use of intelligent computer systems. As with energy demand forecasting, this approach finds widespread use in the sectors of automation, finance, and IT due to its ability to extrapolate future outcomes from historical data. Large amounts of data may be analyzed in a short period of time by inferring complex relationships between weather and energy data.

Existing statistical analysis techniques need large volumes of data and expert knowledge; hence, models based on the aforementioned technologies have been utilized to predict future energy demands in buildings [6].

2.4 ARTIFICIAL NEURAL NETWORK (ANN)

Artificial neural networks (ANN) are a kind of computing inspired by biological neural networks. By stimulating every neuron in the input node and sending the resulting signals to the output node, we can simulate the connection between input and output. Traditional ANN has a three-layer architecture, with the input, hidden, and output layers all interconnected. These layers include diverse neurons that have an activating role. Connection mode, learning method, and activation function are the three factors used to characterize ANN [7]. Image processing, voice recognition, energy forecasting, data mining, robotic control, and many more are only some of the many applications of ANN [8]. The size of an ANN's hidden layers is selected arbitrarily. For training networks, the size of their hidden layers is often determined through trial and error. During training, the network's erroneous weight is corrected. When the gradient performance is below the threshold or when the target value for the decreased error is reached, training is complete [9]. Despite ANN's computing prowess, it has a number of drawbacks, including model over fitting, susceptibility to random weight optimization, and a high chance of local optimum convergence [10].

2.5 RECURRENT NEURAL NETWORK (RNN)

Recurrent neural networks (RNNs) are a form of deep learning algorithm with a layered architecture modeled after the human brain. This is among the most popular techniques now in use. Convolutional neural networks (CNNs) are also considered deep learning. The primary distinction between CNN and RNN is that the latter takes into account both the present and past inputs in its decision-making process. The term "Artificial Neural Network" is often used to refer to these methods (ANN) [11]. In a recurrent neural network, or RNN, the results of the previous stage are used as input for the current stage. While classic neural networks treat all inputs and outputs as independent variables, there are times when a connection between them is necessary, such as when trying to anticipate the next word in a phrase. So RNN was developed, and its Hidden Layer enabled it to deal

with this problem. The Hidden state is RNN's primary and most crucial characteristic, since it allows the network to retain data about previous sequences [12]. Time-dependent information may be included in time series data. Through the use of feedback connections, a Recurrent Neural Network models the dependence by remembering the value at previous time steps [13].

2.6 LONG SHORT-TERM MEMORY NEURAL NETWORK (LSTM)

Long short-term memory neural networks (LSTMs) are the sophisticated version of Recurrent neural networks (RNNs) and long short-term memory neural networks (LSTMs) are an example of RNNs [14]. The RNN's distinctive properties, such as processing the input as a linked time series, are carried over into the LSTM. Furthermore, the LSTM cell's intricate design addresses the issue of disappearing gradients. The four main parts of the LSTM model's flowchart are the input gates, cell status, forget gates, and the output gates [14]. Using input gates, output gates, and forget gates, the data in the cell status may be saved, modified, and removed. As its name suggests, the forget gate is in charge of selecting which bits of data from the previous stage to throw away and which to keep. This is accomplished through the introductory sigmoid layer. Once the new cell state has been established, it is necessary to decide what data should be kept there. As the input gate layer, the sigmoid function calculates the update's value. The tanh layer then produces a fresh vector of values that may be used to update the state. The next thing to do is combine them into a brand-new status. The LSTM stores information about the current cell state [14]. In this case, it performs better than standard RNNs on lengthier input sequences. Each time step, the forget gate is related to the cell's prior state, and the gate uses that information to determine what information to transmit. Afterwards, it is fused with the input gate to create a fresh storage space for the cell. An output must be generated by the LSTM cell at some point. Above, the cell state are retrieved by using a hyperbolic function called tanh, which filters the value of the cell state between -1 and 1.

2.7 ENERGY CONSUMPTION IN BUILDINGS

Energy resource shortages, supply capacity overshoot, and significant environmental repercussions (global warming, climate change, ozone layer depletion, and others) have already been brought up by the world's fast expanding energy usage [18]. A residential

building's energy usage is mostly influenced by its size and location. The quantity of energy utilized in residential structures is mostly influenced by the weather, architectural style, energy infrastructure, and household income. In the next 20 years, the built environment's energy demand will increase by 34%, on average by 1.5% every year. Homes will use 67% of the total energy in 2030, while non-domestic uses will account for 33% [18]. Building energy management systems have recently experienced fast progress, enabling end users to use power in their facilities more efficiently [19]–[23]. For instance, a case study of the energy and environmental performances of a lighting control system in offices was reported by Agheno et al. [24]. Taking into consideration both the observed yearly electric energy consumption and the parasitic energy consumption, they assessed the outcomes of 10 workplaces in Torino (Italy) in terms of possible energy savings (from 17% to 32%). For households in the United Kingdom, Zhou et al. [25] suggested a technique for real-time energy usage control. The results of the simulation presented that, even with faulty anticipated information, the suggested control strategy can optimize the schedule of usage of household appliances and battery charging/discharging behaviors.

Additionally, Valor et al. (2001) used a population-weighted temperature index [26] to evaluate the correlation between building electrical usage and daily outside temperature in Spain. Their research showed that, although the sensitivity in the cold season is usually greater than in the warm season, the sensitivity of electrical demand to daily outside temperature has grown over time, more so for summer than for winter. The heating and cooling demand functions, which have correlation coefficients of 0.79 and 0.87 and anticipate electricity loads with standard errors of estimation of 4% and 2%, respectively, are defined by the relationship between electricity data and outside temperature [26].

In the last 10 years, a lot of research has been performed on energy optimization in buildings. efficiency use of building energy has been the subject of different studies [27], [28] and numerous research projects have been conducted to find the most energy-efficient structures. Using case-based reasoning, Monfet et al. [10] proposed an unique method for estimating the energy consumption of commercial buildings. Their technique may be applied to building operating systems and utilized to forecast energy demand.

The key factors methodology was introduced by Costa et al. [29] to assist energy managers in choosing the best building operating plan with regard to both consumption of energy and

thermal comfort. A framework for smart grid big data analytics as well as the necessary elements of an energy-saving decision-support system were proposed by Chou et al. [30]. The goal of these studies has been to create prediction platforms and techniques that will increase the effectiveness and dependability of energy management systems.

According to research, end-use energy efficiency may dramatically lower the overall amount of energy consumed worldwide. In a research published by Lee et al., they have mentioned that in the majority of nations, buildings use 40% or more of the total energy and are a major source of greenhouse gas emissions. Buildings must be more energy efficient and use less energy while yet maintaining tenant comfort and safety as environmental sustainability becomes a major worldwide problem [31]. Also, they have explained that the installation of meters in an increasing number of buildings is a significant trend (e.g., smart meters, sensors, BMS, and other systems). Analytics that are based on basic science and math may be used to fully exploit all the information being acquired from equipment in buildings as well as other conventional data, such as utility data. Furthermore, they have presented a simple yet effective mathematical method for building operators, renters, and owners to measure, benchmark, analyze, and simulate energy. The analytics can be useful in locating and evaluating options for energy-saving in buildings. Currently being evaluated in public schools and business office buildings, the analytics technology will eventually be expanded to serve other applications [31].

2.8 AI METHODS FOR TIME-SERIES ENERGY CONSUMPTION FORECASTING

Energy consumption assessment refers to the process of analyzing time series data to derive meaningful statistics and other features of the data. Using a model to forecast future values based on observed past values is known as time series forecasting [32]. An ordered set of values that are measured at regular intervals of time is referred to as a time series. A time series analysis is categorized into two components [33]. The structure and underlying pattern of the observed data are discovered in the first section. To support future predictions, a model is fitted to the data in the second section. Numerous areas of applied science and engineering that use periodic data, such as statistics, pattern recognition, signal processing, energy prediction, econometrics, earthquake prediction, weather forecasting, astronomy, and other subjects, all use time series analysis.

Since buildings are increasingly monitored in real time, time series analysis is frequently used to forecast energy use. Such monitoring can offer historical data that can be analyzed critically and used to anticipate how much energy would be used by buildings. Numerous studies have looked at how energy-efficient buildings are [34]–[37]. Several time series models have been created to predict building energy usage [38]–[40].

The prediction of consumption of electric energy is a multivariate time series prediction issue [33]. The gathered data from sensors is susceptible to error [41]. Using the prediction model and a specified window of many sensor inputs, distinct information may be retrieved to anticipate the power usage [38]. Nevertheless, it is very challenging to anticipate electric energy consumption using traditional forecasting techniques since power consumption has a predictable seasonal pattern and incorporates irregular trend components [42]. Observed electric energy consumption may be broken down into three components: the trend factor, the seasonal factor, and the random factor [14]. Figure 1 depicts a time series breakdown into its three constituent components, all of which relate to the use of electrical power. Consumption of electric energy exhibits a variety of patterns, and it is useful to classify some of the time series' observed characteristics. Decomposition of time series gives helpful visuals for comprehending the difficulty of assessing and estimating energy use in general. We discuss three kinds of time series patterns and concentrate on the irregular trend in the breakdown of the energy consumption time series in Figure 2.1.

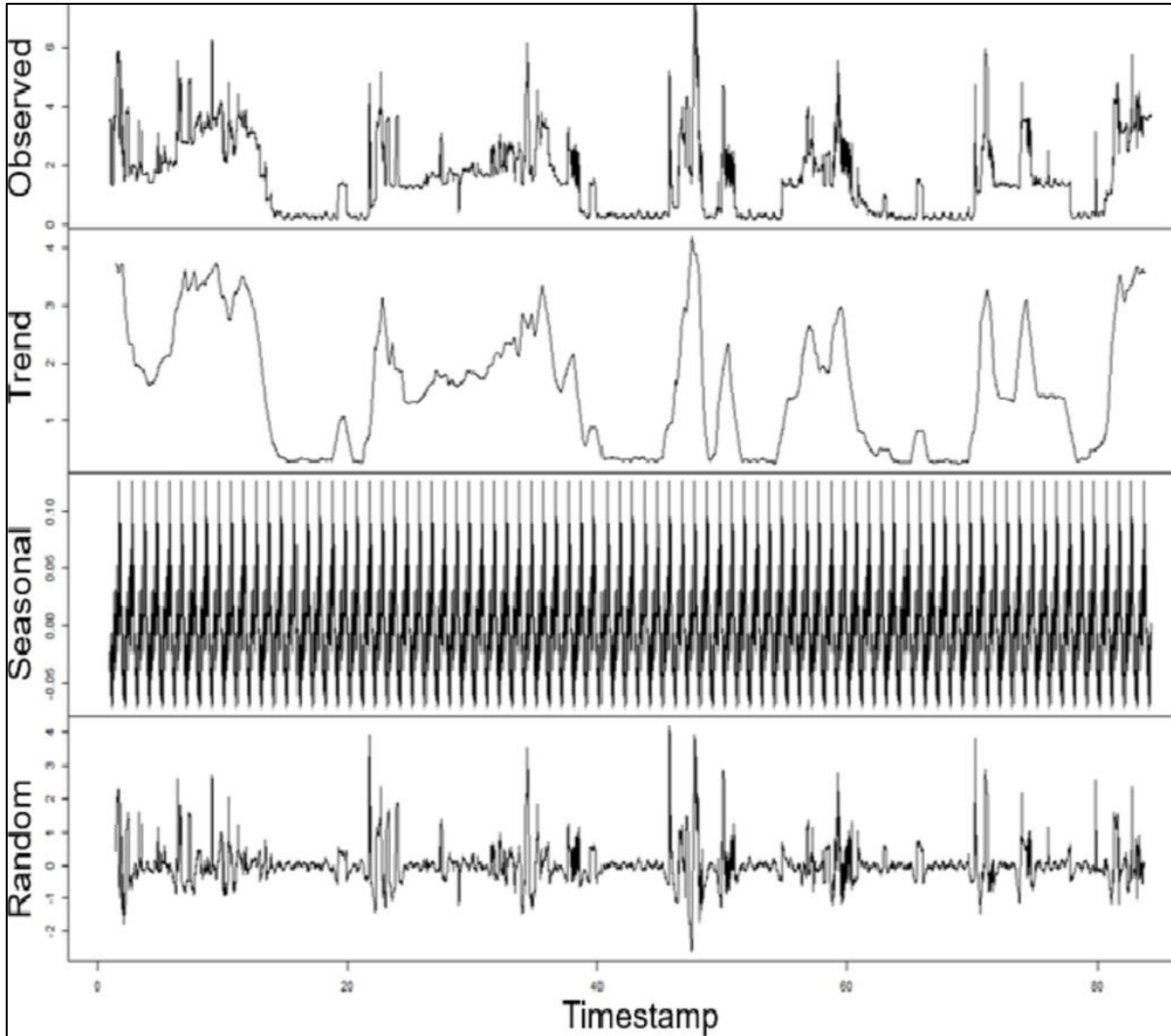


Figure 2.1: Time-Series Breakdown to Discover Energy Usage Components [43]

The effectiveness of machine learning algorithms is often contingent on the data representation. Among them, the deep learning approach seeks to construct a synthesis of several nonlinear transformations, helpful expressions that may provide more abstract and, eventually, more profit [44]. This approach of deep learning has been intensively researched for several applications. Convolutional neural network (CNN) outperforms conventional image recognition algorithms, but recurrent neural network (RNN) excels in voice recognition and natural language processing. CNN learns the weights of the feature maps that comprise each layer, extracts abstract visual elements namely lines, points, and faces from the input data, and maintains the pixel connections for the learning picture [45]. Long sequential information is stored in the RNN's hidden memory for representation,

processing, and storage. It also changes throughout time to guarantee the persistence of time-related information [46].

Recently, several researchers have combined LSTM and CNN models to extract temporal and spatial characteristics. Wang et al. [47] examined emotions in the realm of processing of natural language utilizing the regional CNN-LSTM model with text input. In the realm of speech processing, Sainath et al. [48] proved the noise resilience of several voice search tasks using the CLDNN model that combines LSTM, CNN, and DNN. In the realm of video processing, Ullah et al. [49] created a model that recognizes human activity in video sequences by merging CNN with Bi-directional LSTM. Oh et al. used LSTM and CNN to reliably identify arrhythmias in the ECG [50] in the medical industry. Zhao et al. [51] built the convolutional bi-directional long short-term memory network to anticipate tool wear in the industrial sector using the mechanical health monitoring approach.

The challenge of predicting energy consumption also involves time series data applied to multiple sectors. For time series prediction, the usual technique employs a subset of all characteristics and solely models temporal information.

2.9 SUMMARY

Based on the literature review presented in this chapter, the following points are observed. Most of the available research took into account cutting-edge statistical methods for estimating building energy usage. Numerous studies have been done to create reliable knowledge for forecasting energy consumption in buildings, however this is hindered by a lack of strong ML approaches and a lack of diverse datasets. To this end, ML seeks to develop models that, after being exposed to and learning from designated training datasets, can reliably and quantitatively predict future outcomes. Although there are huge numbers of ML tools available to predict the energy consumption of a building, it requires finding the most effective options among them. The existing literature lacks in proving any tools for claiming it to be most efficient. This study has conducted a comparative study to find out the most effective tool for energy performance analysis.

3. METHODOLOGY

3.1 INTRODUCTION

This chapter shows the measure the effectiveness of ML tools to predict energy consumption of a building. This chapter analysed and compared the working principles of ML tools for predicting energy consumption. It has considered three datasets to measure the effectiveness of the machine learning tools for selecting the best among them.

3.2 METHODOLOGY

The goal of this study is to forecast the energy consumption of a residential building using different machine learning algorithms. The LSTM, RNN and CNN algorithms were used for prediction purpose using real data. Using real-time data collected from a smart grid installed in an experimental building, the efficacy and utility of statistical and ML algorithms were evaluated. Energy usage in single and ensemble situations was examined using well-known artificial intelligence techniques. The models which combine prediction and optimization approaches, is examined in-depth. The thorough comparison shows that the hybrid model performs better than the single and ensemble models in terms of accuracy. These models are thought to be suitable for usage and accurate enough to provide predictions, which can help users plan their energy management strategies. Since the initial 1997 LSTM study [4] was published, additional theoretical and practical research on this sort of RNN have been published, with many describing the remarkable results attained in a broad range of sequential data application fields. LSTM networks have had a significant influence on speech-to-text transcription, language modeling, and machine translation, among other applications [5]. Some academic and industry readers, inspired by the remarkable benchmarks given in the literature, desire to learn about the Long Short-Term Memory network (henceforth "the LSTM network") in order to evaluate its relevance to their own study or practical use-case. Numerous RNN and LSTM network designs are efficiently and production-ready implemented in all of the main open-source frameworks for machine learning. Obviously, some practitioners, even if they are unfamiliar with LSTM/ RNN systems, take advantage of this accessibility and cost-effectiveness and immediately begin research and experimenting. Others want a comprehensive understanding of the working of this beautiful and successful system. The

upside of taking this longer route is that it provides the chance to advance intuitive skills that will serve them well throughout the entire process of adapting an open source module to their research project or business application, from initial data collection and cleaning to troubleshooting and fine-tuning.

In this study, both RNN and LSTM was used to predict the energy consumption provided from different companies with three datasets.

In this thesis, three different method of Convolutional Neural Networks (CNN), Recurrent Neural Network (RNN) and Long Short-Term Memory Network (LSTM) have been utilized to predict the energy consumption of datasets taken from different regions and different companies. Three different datasets were normalized, trained and tested and the accuracies were compared to each other.

This study has outlined all the procedures needed in this kernel for doing time series analysis using time series data and RNN and LSTM models. Furthermore, the performance of the LSTM model with the basic RNN model has been compared in detail. Three different datasets from different companies were selected and used to implement CNN, RNN and LSTM models on them and the results were compared together for all the datasets.

3.3 NETWORK OF LSTM, RNN AND CNN

In conventional neural networks, only the layers above and below each other are linked together, and no nodes within the same layer communicate with one another. Therefore, each input and each output may be treated separately. It's tough to time-dependent exchanges that interact temporarily. This problem may be solved by using an RNN with loops. A typical RNN design (left) and its development over time are shown in Figure 3.1 (right). These concealed units in the network are able to pass on information from one stage to the next because to the feedback loop established between the past and current states. A useful analogy for the unrolled structure is that of many copies of the same network. It is discovered that the output of the last hidden layer is reliant on the input at the moment as well as the output of the layer before it. The capacity to learn temporal correlations over longer periods of time is constrained by the fact that traditional RNN struggle with long-range dependencies where the gradient either expands or disappears

during back propagation [1]. The purpose of an RNN equipped with LSTM cells is to solve the memory retention issue over extended periods of time without developing a dependence problem. Hochreiter et al. presented the first LSTM model in 1997, and Gers et al. enhanced it by adding a forget gate [2]. It is a subclass of RNNs that can learn order dependency, and it has been widely used and improved upon for a wide range of tasks, including sequence prediction.

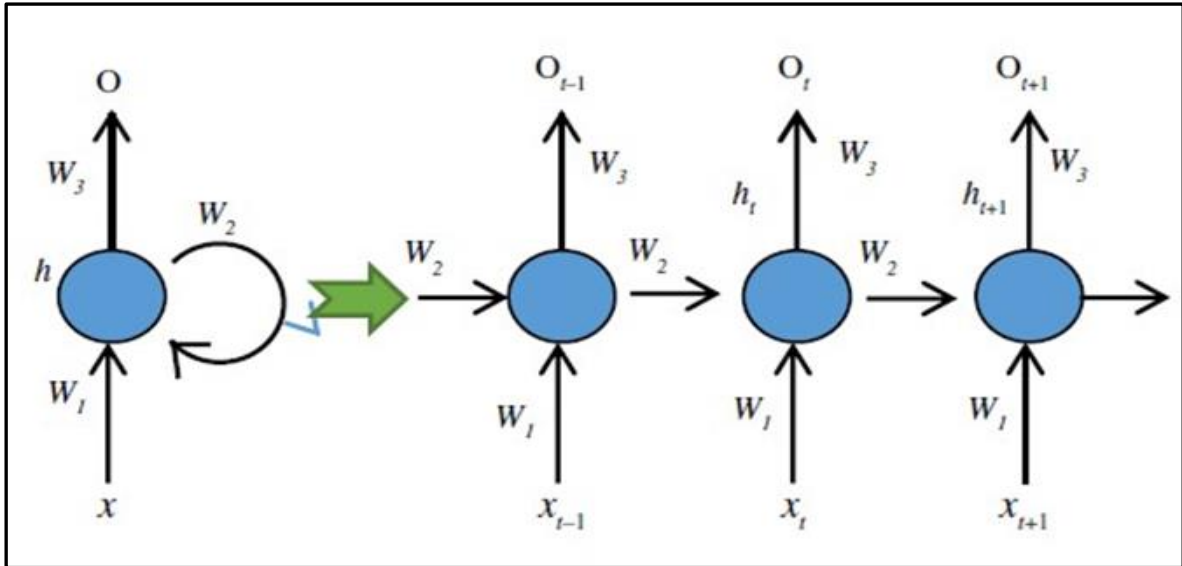


Figure 3.1: Structure of Recurrent Neural Network (RNN) and Unrolled Structure of RNN In Time.

To understand the mechanism of RNN we need to follow the figure 3.2 in this figure 4 neuron are communicating with the other hidden layer and get the expected result.

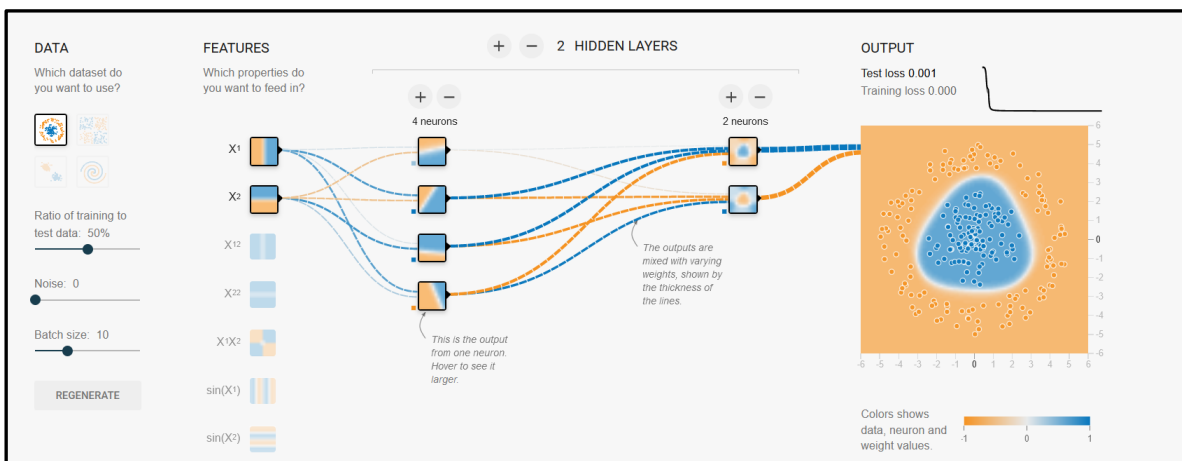


Figure 3.2: Visualization of RNN

A RNN's unit uses a basic nonlinearity to alter both inputs and recurrent units. In contrast to the RNN's outer recurrence, which is seen in Fig. 3.3, the LSTM's structure includes a self-loop. A typical LSTM receives a time series of inputs denoted by $[x_1, x_2, \dots, x_T]$, where $[x_t]$ is the feature dataset applied to the input at time step t , and $[c_t]$ is a memory cell inside the LSTM cell that holds the long-term dependencies. Training an LSTM involves erasing and adding data to memory cells, respectively, using the input gate i_t , output gate O_t , and forget gate f_t . Input gate, forget gate, cell state, output gate, and hidden output state formulae are shown in Fig. 3.3

$$i_t = g(W_{xi} x_t + W_{hi} h_{t-1} + b_i) \quad (3.1)$$

$$f_t = g(W_{xf} x_t + W_{hf} h_{t-1} + b_f) \quad (3.2)$$

$$c_t = f_t c_{t-1} + i_t \tan h(W_{xc} x_t + W_{hc} h_{t-1} + b_c) \quad (3.3)$$

$$o_t = g(W_{xo} x_t + W_{ho} h_{t-1} + b_o) \quad (3.4)$$

$$h_t = o_t \tan h(c_t) \quad (3.5)$$

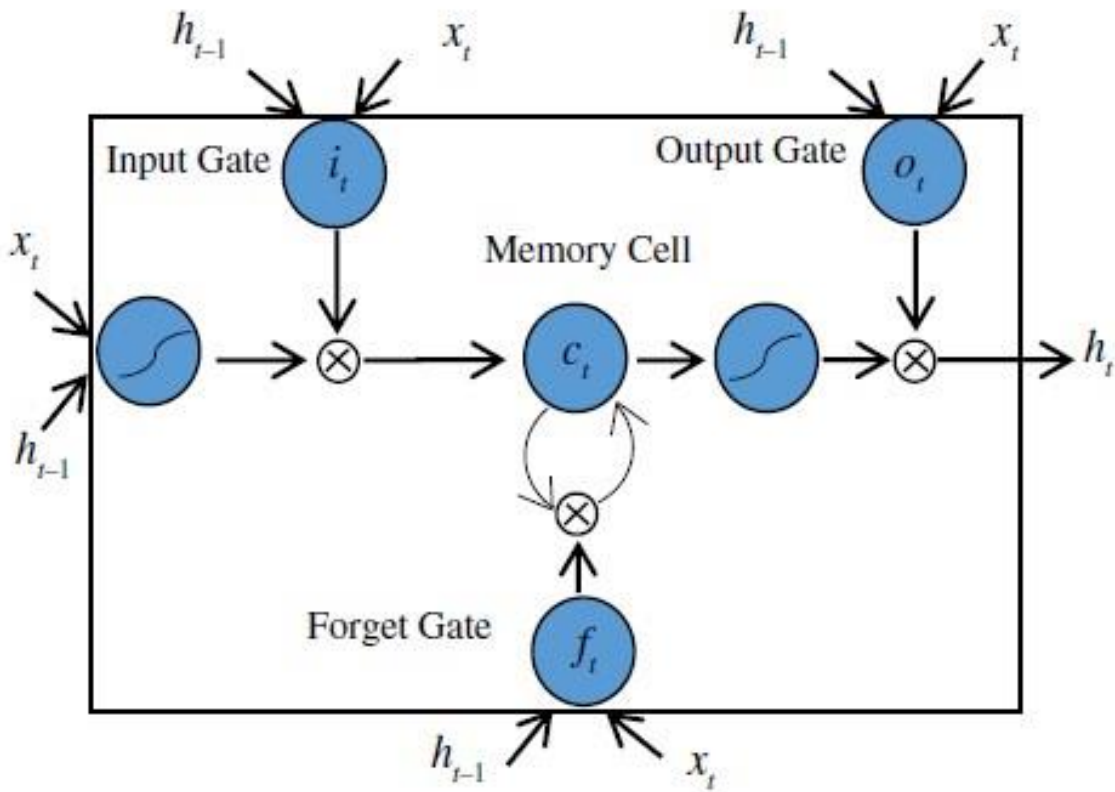


Figure 3.3: Structure of a Long Short-Term Memory Cell (LSTM).

where h_0 is the zero-based starting hidden state and $g()$ is the sigmoid functions. Forget, Memory, Input, and Output Gate at Time Step t are f_t , c_t , i_t , and o_t , respectively. All of these functions are examples of sigmoid functions, and their potential nil values indicate that the data being calculated may be constrained to the next node. The network's transition between two components is represented by the weights of the gates, each of which is denoted by the letter W . W_{hi} stands for the hidden-input gate matrix, whereas W_{xf} for the input-forget gate matrix. Matrix W_{hc} stands for the hidden-memory gates. The addition of bias b to the formula at each gate makes the network more adaptable to new sets of data. To visualize LSTM we need to follow in the figure 3.4 in here the frequency is counted and set it as a frequency identifier.

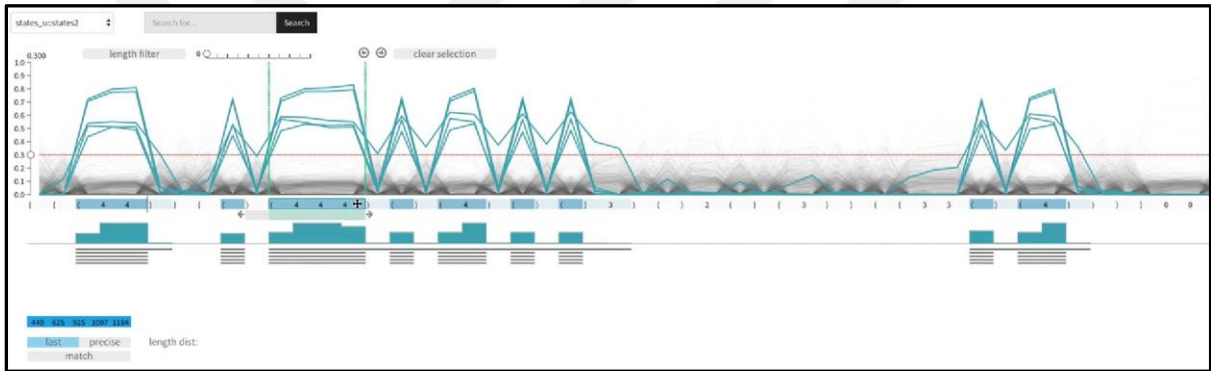


Figure 3.4: Visualization of LSTM

To determine the load power at time step t , a fully linked layer exhibits a linear transformation of the output hidden state h_t . the equation is as follows:

$$P_t = V_{out} h_k + b_y \quad (3.6)$$

where V_{out} is the weight matrix of the fully linked layer, and b_y is the biases.

For CNN It is a type of deep learning neural network that is commonly used in large number of file to make a diction. The key feature of a CNN is the use of convolutional layers, which learn spatial hierarchies of features from input data. These layers are composed of a set of filters that scan the input data, apply a mathematical operation (convolution), and produce a set of output feature maps. The output feature maps are then processed by pooling layers, which reduces the spatial dimensions of the data and increase the model's ability to recognize patterns in the data which is shown in the figure 3.4.

Finally, the output from the pooling layers is passed through fully connected layers, which make predictions based on the learned features.

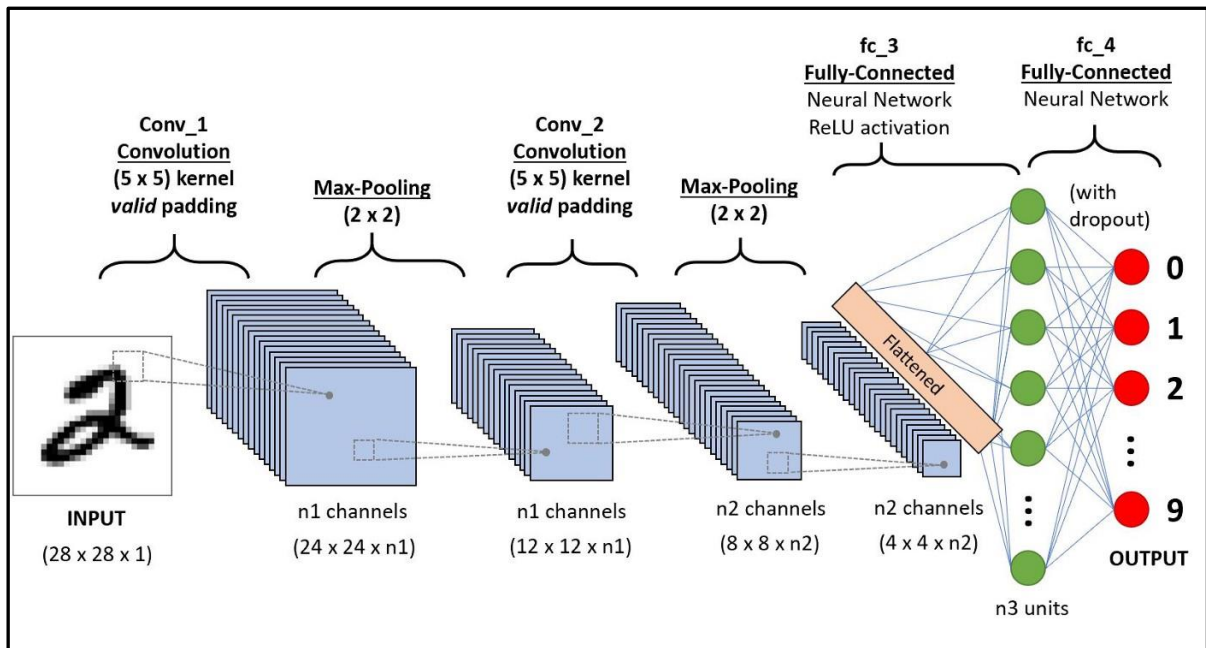


Figure 3.5: Structure of a CNN (Convolutional Neural Network)

Preparing data for a CNN typically involves several steps:

Collecting and labeling data: This step involves gathering a large dataset of csv file that are labeled with the appropriate class or category. The dataset should be diverse and representative of the problem we are trying to solve.

Data preprocessing: This step involves cleaning and preparing the data for use in a CNN. This may include resizing csv file to a consistent size, normalizing csv file, and converting csv file to normal train data.

Data augmentation: This step involves applying various techniques to increase the size and diversity of the dataset. This can include random rotations, flipping, cropping, and zooming of csv file, which can help to increase the robustness of the CNN.

Data splitting: The final step is to split the data into training, validation, and testing sets. This is important to ensure that the CNN is trained and evaluated on different data. Typically, 80% of the data is used for training, 10% for validation, and 10% for testing. After that we have got Figure 3.5.

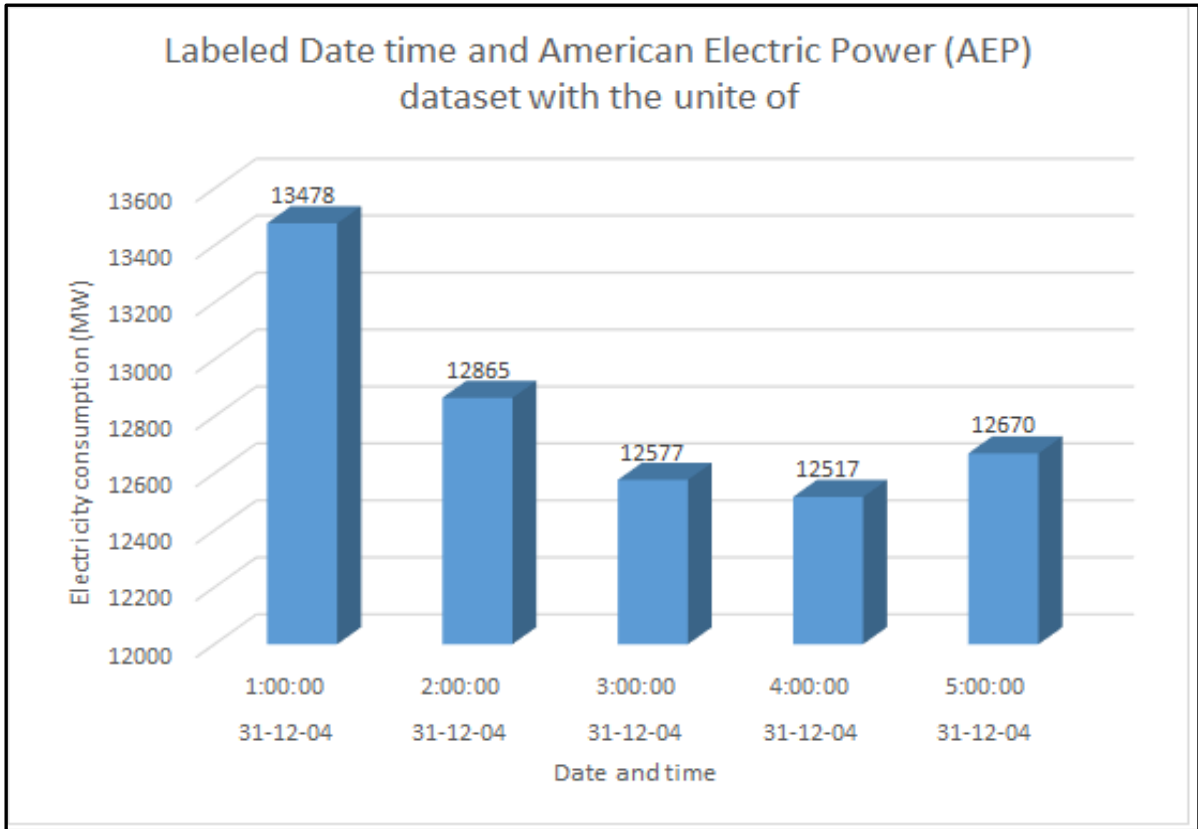


Figure 3.6: CNN Data Preparation.

For visualization CNN we need to see the figure 3.6 in this figure all the dataset is divided into 3 types of main database and passing it to a hidden layer and get the outcomes of the results.

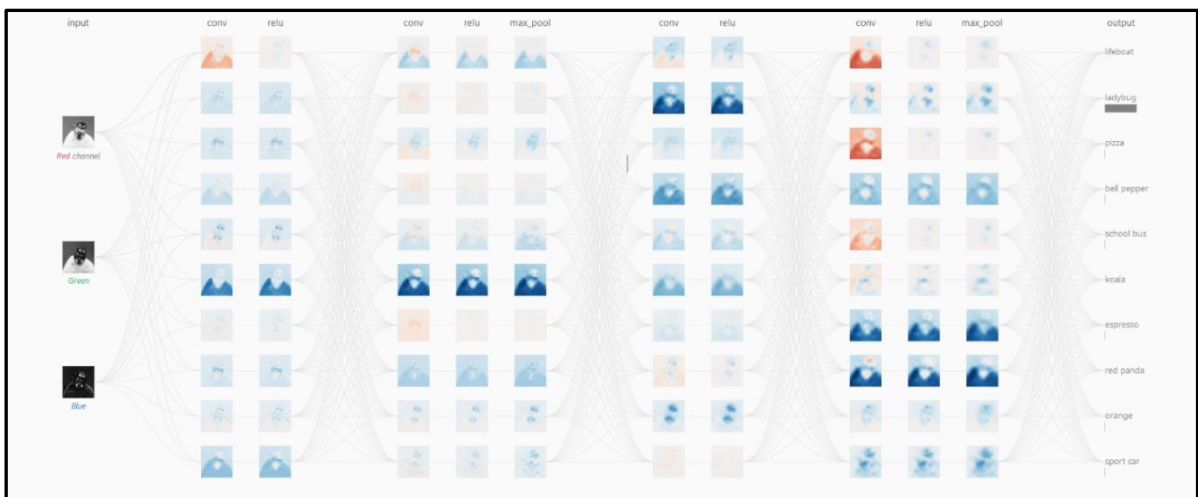


Figure 3.7: Visualization CNN.

It's important to note that the specific data preparation steps will vary depending on the problem we are trying to solve and the type of data we are working with.

3.4 AMERICAN ELECTRIC POWER (AEP)

The first dataset has been taken from American Electric Power (AEP), which contains 121274 hourly consumed energy consumption. Different steps of data loading, exploration, normalization, and implementation of RNN and LSTM models on this dataset have been explained in this section.

3.4.1 Data Loading and Data Exploration

In the first step, all the excel files were uploaded to the google colab library content then we called them to explore the dataset in python. Dataset from various energy consumption companies were uploaded. Table 3.1 shows the labeled Date time and American Electric Power (AEP) dataset with the unite of MW.

Table 3.1: Labeled Date Time and American Electric Power (AEP) Dataset with The Unite Of MW.

Label	Date	Time	Electricity consumption (MW)
0	2004-12-31	01:00:00	13478.0
1	2004-12-31	02:00:00	12865.0
2	2004-12-31	03:00:00	12577.0
3	2004-12-31	04:00:00	12517.0
4	2004-12-31	05:00:00	12670.0

After this section, we changed the labels of 0,1,2, . to the date time of the dataset to make the processing and analysis easier since our problem is a time series problem. Table 3.2 demonstrates the dataset indexed via datetime.

Table 3.2: Indexed Via Datetime Dataset of American Electric Power (AEP) Dataset with The Unit Of MW.

Date	Time	Electricity consumption (MW)
2004-12-31	01:00:00	13478.0
2004-12-31	02:00:00	12865.0
2004-12-31	03:00:00	12577.0
2004-12-31	04:00:00	12517.0
2004-12-31	05:00:00	12670.0

After this step, the dataset was checked to prevent any possible data missing. After checking it was confirmed that there is no missing data.

4. RESULTS AND DISCUSSION

The main focus of this chapter is to analyze and compare the performance of the building energy performance measuring tools based on ML. Building energy prediction has been the subject of much study in the past, but reliable energy prediction requires the use of advanced ML methods and a larger variety of datasets. First part of this chapter was started with the use of carefully selected training datasets, ML attempts to build models with the ability to make precise, measurable forecasts. This study predicted energy consumption using two most promising deep learning algorithms namely LSTM and RNN. In the second part, a comparative analysis was performed using original data that was collected from the field. Before encoding the data, normalization was done using a normalization scale. Finally, the best simulation tools have been selected based on the comparative analysis.

AEP Power Consumption

In the step, in order to visualize the dataset, the datetime vs Consumed energy without any normalization of data was plotted. Fig. 4.1 demonstrates the dataset for different dates in MW in pre-renormalized stage.

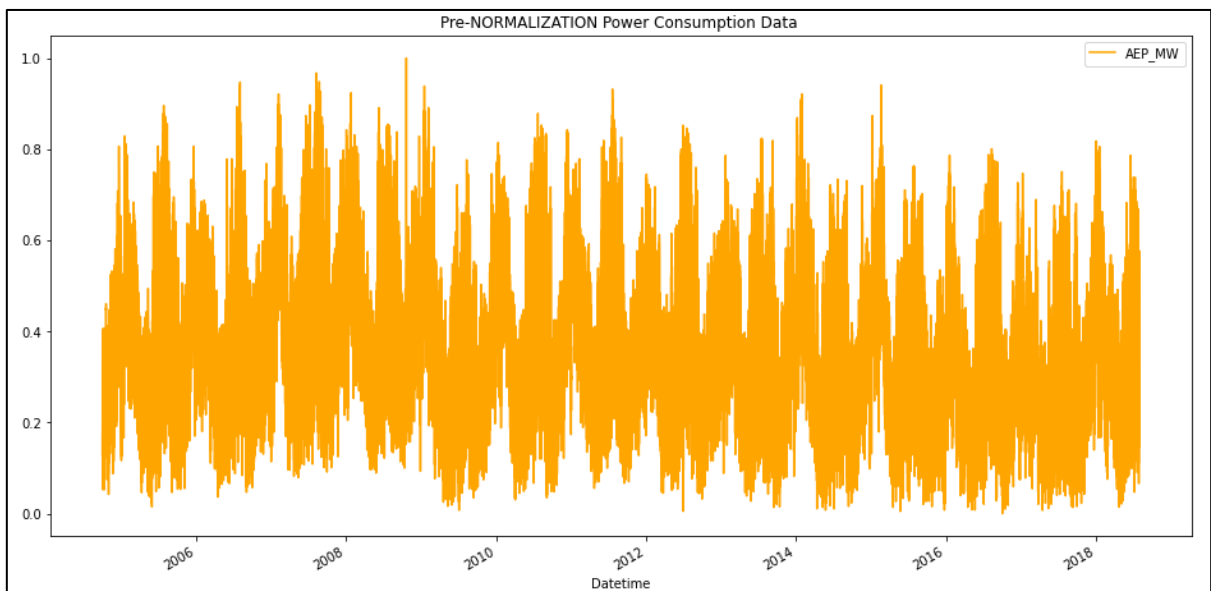


Figure 4.1: Pre-Normalization Power Consumption Data.

4.1 DATA NORMALIZATION

The normalization procedure is a preliminary step in solving any kind of issue. Particularly in the fields of cloud computing, soft computing, etc., normalization plays a vital role in the manipulation of data, such as scaling down or scaling up the range of data before it is utilized in the next step. Common methods of normalizing include the Z-score, Min-Max, and decimal-scaling normalizations [1]. MinMax Scaler is one of the methods of data scaling involves setting the lowest value of a feature to zero and the highest value to one. The data is reduced by MinMax Scaler to fit inside the specified range, which is typically between 0 and 1. It alters information by stretching or compressing some aspects of it. It preserves the original distribution form while adjusting the numbers to fit inside a specified interval. The MinMax scaling is performed using as follows [2]:

$$x_std = (x - x.min(axis=0)) / (x.max(axis=0) - x.min(axis=0))$$

$$x_scaled = x_std * (max - min) + min$$

Where,

min, max = feature_range

x.min(axis=0) : Minimum feature value

x.max(axis=0): Maximum feature value

This may be done with the help of the MinMaxScaler() function defined in the Sklearn preprocessing.

In this study, the data was normalized prior to the processing data for further analysis, in order to normalize dataset, sklearn MinMaxScaler was used. Min-max scaling is similar to z-score normalization in that it will replace every value in a column with a new value using a formula.

4.1.1 Data Visualization Following Normalizing

After normalization, the graph's y-axis shows a change in the range of power consumption in figure 4.2. Its range in the previously presented graph was 0 to 22500. Now that the data have been normalized, we can see that the y-axis data range is 0.0 to 1.0. As Min-Max

scaler was used to normalized data it turned the energy consumption data in the range of 0 to 1 which was previously showing the high values whereas x-axis showed the hourly data. If there is big difference in the two types of data that may conclude the poor prediction results. This necessitated normalizing the data.

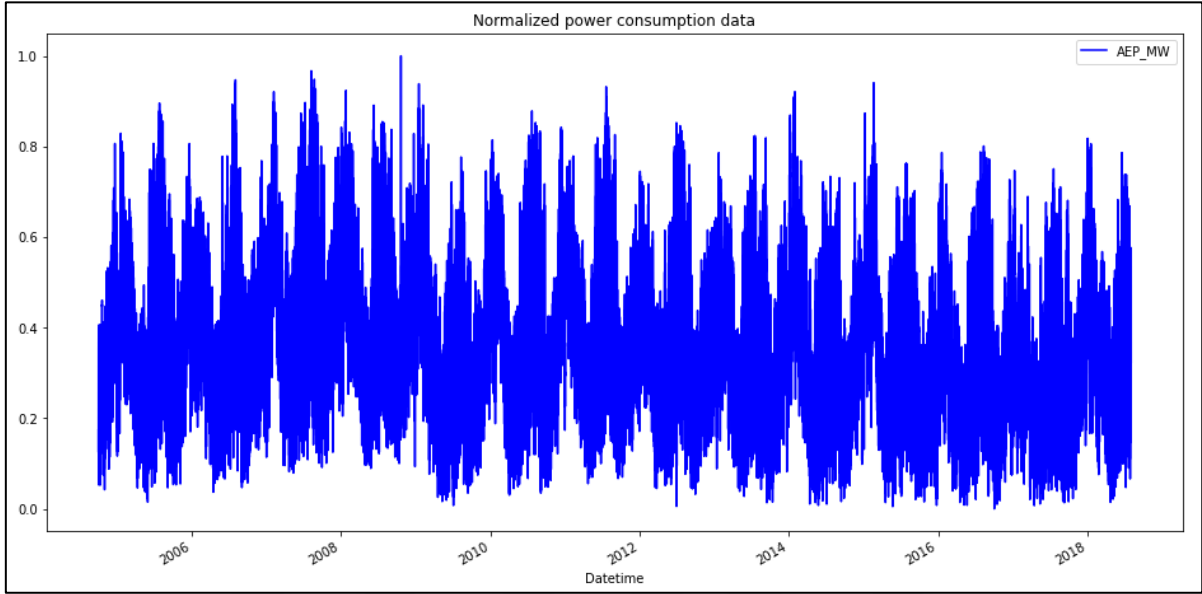


Figure 4.2: Normalized Power Consumption Data Vs Datetime.

4.1.2 Preparation of Data to Implement the RNN Models Training

In order to train dataset for recurrent neural networks (RNN) model, after loading data and using sequential data length, 11273 hours of data were used for the test and first 110000 hours were used for training purpose. By definition, recurrent neural networks are used with sequential data, which without sacrificing generality implies data samples that change along a time axis. The sample values across a limited time frame then explain a whole history of a data sample; hence, if your data are in an N-dimensional space and change over t-time steps, your input representation must have the form (num samples, t, N).

As it can be seen in the graph, the actual numbers can be compared to the projected ones. Since the predicted values are quite close to the real values, the RNN model is doing a decent predicting the sequence, as can be shown. 10 epochs with batch size of 1000 were used for the training and MSE used to find the loss of the results. For RNN model of AEP dataset the R^2 score value of 0.96837 was achieved. The value of R-squared may vary from 0 to 1. In this case, higher R^2 value represents the most reliable predictability of the model.

According to the figure similar behavior of the predicted model mimicked the actual values. The acceptability of predicted model was clarified with the statistical analysis with R^2 value.

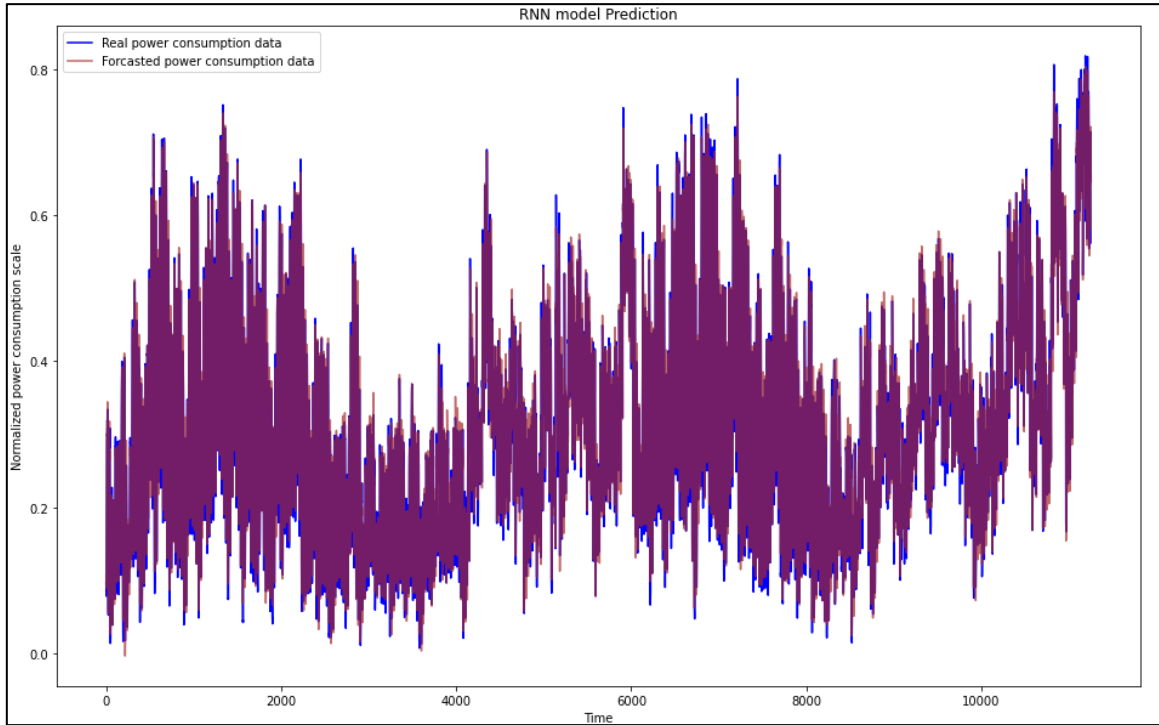


Figure 4.3: RNN Model Prediction Vs Real Power Consumption Data.

4.1.3 LSTM Model

For the LSTM model of AEP dataset, after data sequential, adam optimizer, and MSE for the loss results were used for the training of dataset. With setting epochs as 10 and the batch size of 1000, the results of R^2 score was achieved as 0.96858773. This value was a bit higher than that of RNN model for this dataset. This higher R^2 value clarify the predictability of the model more accurately. In accordance with the figure 4.4, it was observed that wave of predicted energy consumption data followed that similar sequence of real energy consumption data. Moreover, while focusing on the spike of the graph similar behavior of the predicted model followed the actual values. The acceptability of predicted model was clarified with the statistical analysis with R^2 value.

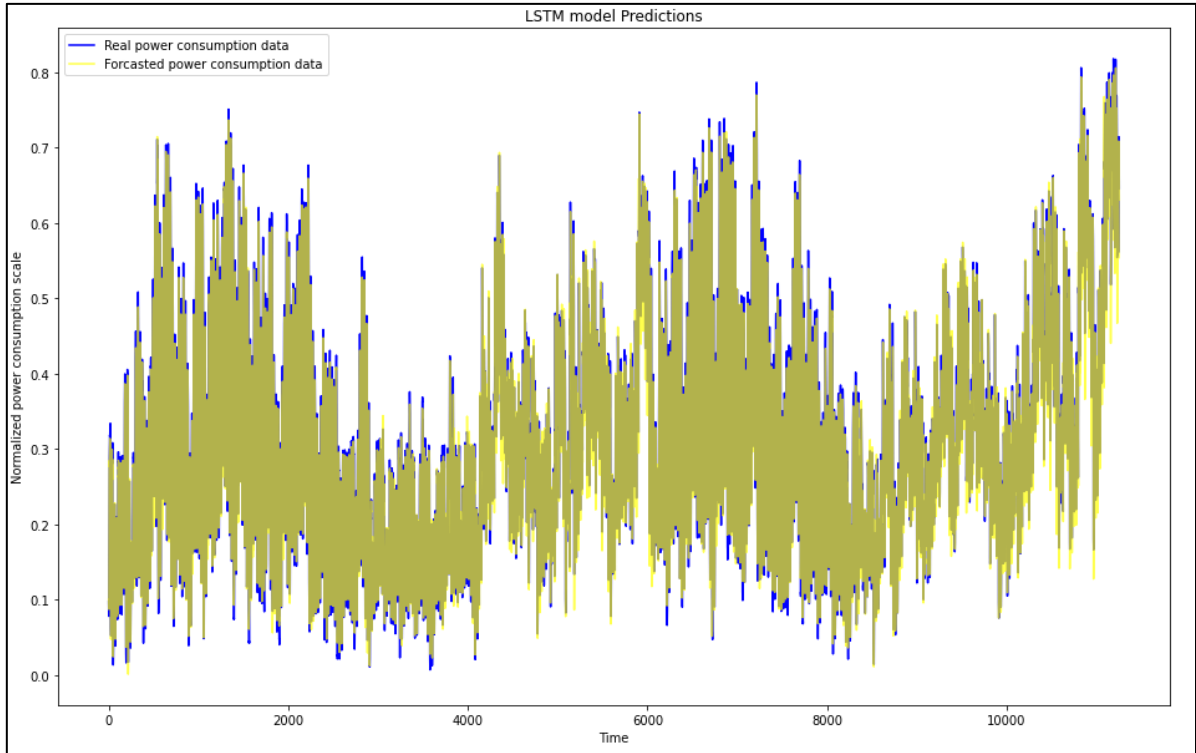


Figure 4.4: LSTM Model Prediction Vs Real Power Consumption Data.

4.1.4 Comparison Between RNN and LSTM Model Predictions

In this section, a comparison between the results of the prediction made by RNN and the results taken from LSTM model have been shown in a single plot to check and compare them. Figure 4.5 represents the comparative analysis where LSTM, RNN and original values were presented with different colors. It can be observed that in different datetimes LSTM and RNN are closer to the original dataset. Both models have successfully predicted the test data with a decent R^2 value (about 0.97) where LSTM model executed higher R^2 value compared to the RNN model. Hence, LSTM prediction model was accurate than RNN prediction model in terms of acceptability. Figure 4.5 demonstrated that wave of both predicted energy consumption model followed that similar sequence of real energy consumption data. Moreover, while focusing on the spike of the graph similar behavior of the predicted model mimicked the actual values. The acceptability of predicted model was clarified with the statistical analysis with R^2 value.

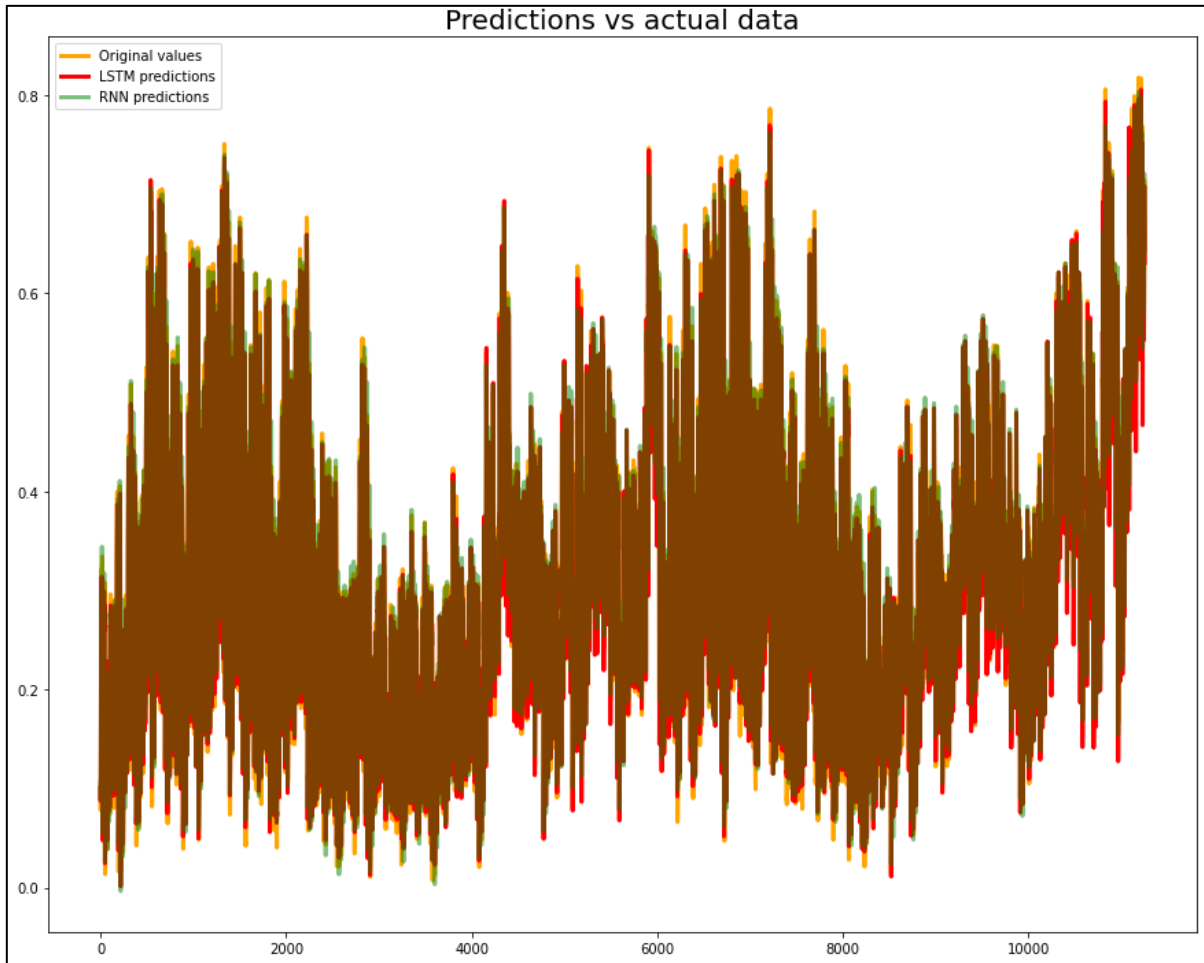


Figure 4.5: Comparison Between RNN and LSTM Model Predictions.

4.2 DOMINION VIRGINIA POWER (DOM)

The second dataset has been taken from Dominion Virginia Power (DOM), which contains 116190 hourly consumed energy consumption. Various steps of the implementation of RNN and LSTM models on this dataset also have been demonstrated in the following section.

4.2.1 Data Loading and Data Exploration

As mentioned in the previous section for the AEP dataset, firstly, all the excel files were uploaded to the google colab library content then we called them to explore the dataset in python. Dataset from various energy consumption companies were uploaded. Table 4.1 shows the labeled Date time and Dominion Virginia Power (DOM) dataset with the unite of MW.

Table 4.1: Labeled Date Time and Dominion Virginia Power (DOM) Dataset with The Unite Of MW.

Label	Date	Time	Electricity Consumption (MW)
0	2005-12-31	01:00:00	9389.0
1	2005-12-31	02:00:00	9070.0
2	2005-12-31	03:00:00	9001.0
3	2005-12-31	04:00:00	9042.0
4	2005-12-31	05:00:00	9132.0

The same procedure has been applied on this dataset as well and after labeling, Table 4.2 demonstrates the dataset indexed via datetime.

Table 4.2. Indexed Via Datetime Dataset of Dominion Virginia Power (DOM) Dataset with The Unit Of MW.

Date	Time	Electricity consumption (MW)
2005-12-31	01:00:00	9389.0
2005-12-31	02:00:00	9070.0
2005-12-31	03:00:00	9001.0
2005-12-31	04:00:00	9042.0
2005-12-31	05:00:00	9132.0

After this step, the dataset was checked to prevent any possible data missing. After checking it was confirmed that there is no missing data. In the next step in order to visualize the dataset, plot the date time vs Consumed energy was plotted without any normalization of data. Fig. 4.6 demonstrates the dataset for different dates in MW. There

was steady increase in energy consumption from 2006 to 2018 along with a high demand of around 2100 MW in between 2014 to 2015. The lowest energy consumption was happened during 2010 and it was around 1000 MW.

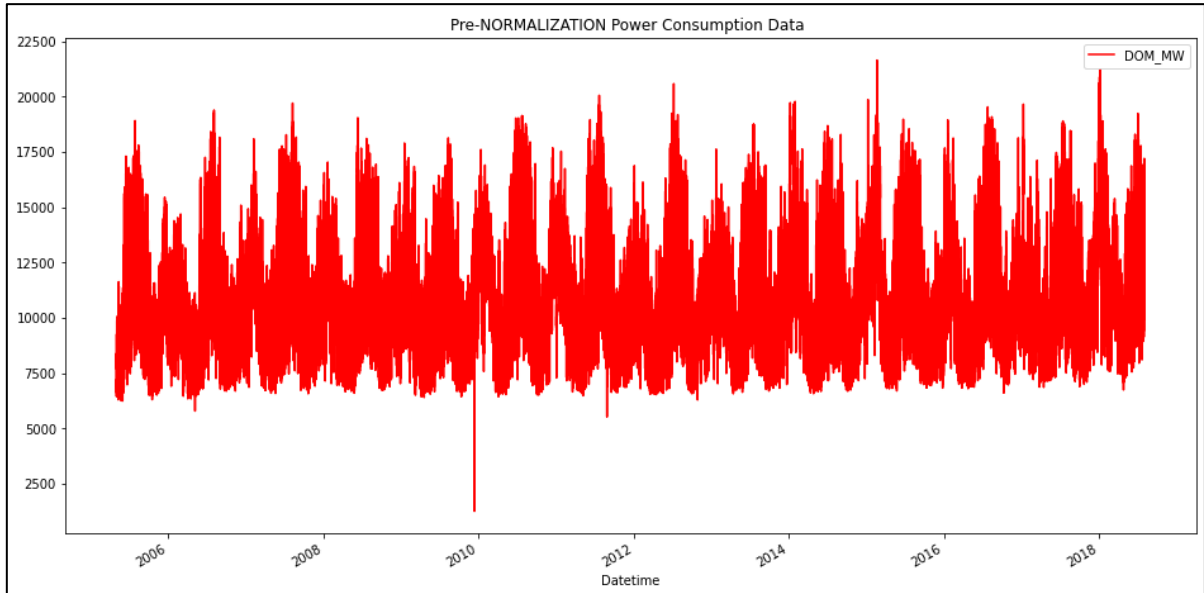


Figure 4.6: Pre-Normalized Dom Power Consumption Data.

4.2.2 Data Normalization

Prior to the processing data for further analysis, the data was normalized. In order to normalize dataset, we used sklearn MinMaxScaler.

4.2.2.1 Data visualization following normalizing

After normalization, the graph's y-axis shows a change in the range of power consumption in figure 4.7. Its range in the previously presented graph was 0 to 22500. Now that the data have been normalized, it can be observed that the y-axis data range is 0.0 to 1.0. As Min-Max scaler was used to normalized data it turned the energy consumption data in the range of 0 to 1 which was previously showing the high values whereas x-axis showed the hourly data. If there is big difference in the two types of data that may conclude the poor prediction results. This necessitated normalizing the data. Lee et al. [31] has also emphasized on the normalization of data when the difference is large among data.

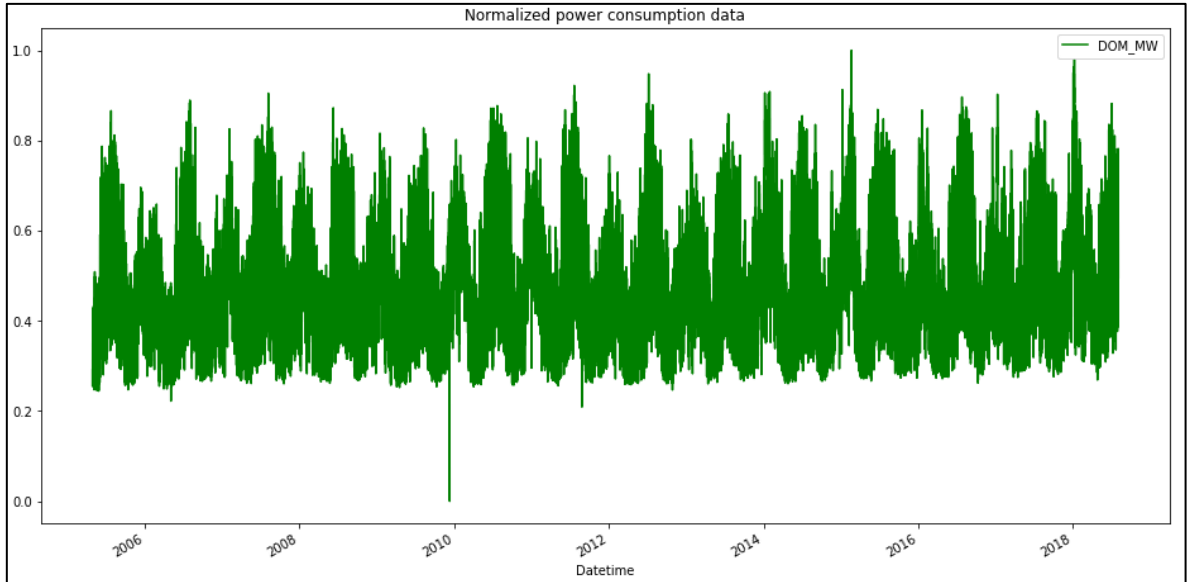


Figure 4.7: Normalized DOM Power Consumption Data Vs Date Time.

4.2.3 Preparation of Data to Implement the RNN Models Training

In order to train dataset for recurrent neural networks (RNN) model, after loading data and using sequential data length, we used 6189 days of data for the test and first 110000 days were used for training purpose. As it can be seen in the graph, we can compare the actual numbers to the projected ones. Since the predicted values are quite close to the real values, the RNN model is doing a decent predicting the sequence, as can be shown. 10 epochs with batch size of 1000 were used for the training and MSE used to find the loss of the results. For RNN model of DOM dataset the R^2 score value of 0.94335387 was achieved which is comparably lower than AEP power data set. Even after it is demonstrating higher prediction rate. In accordance with the figure 4.8, it was observed that wave of predicted energy consumption data followed that similar sequence of real energy consumption data. Moreover, while focusing on the spike of the graph similar behavior of the predicted model followed the actual values. The acceptability of predicted model was clarified with the statistical analysis with R^2 value.

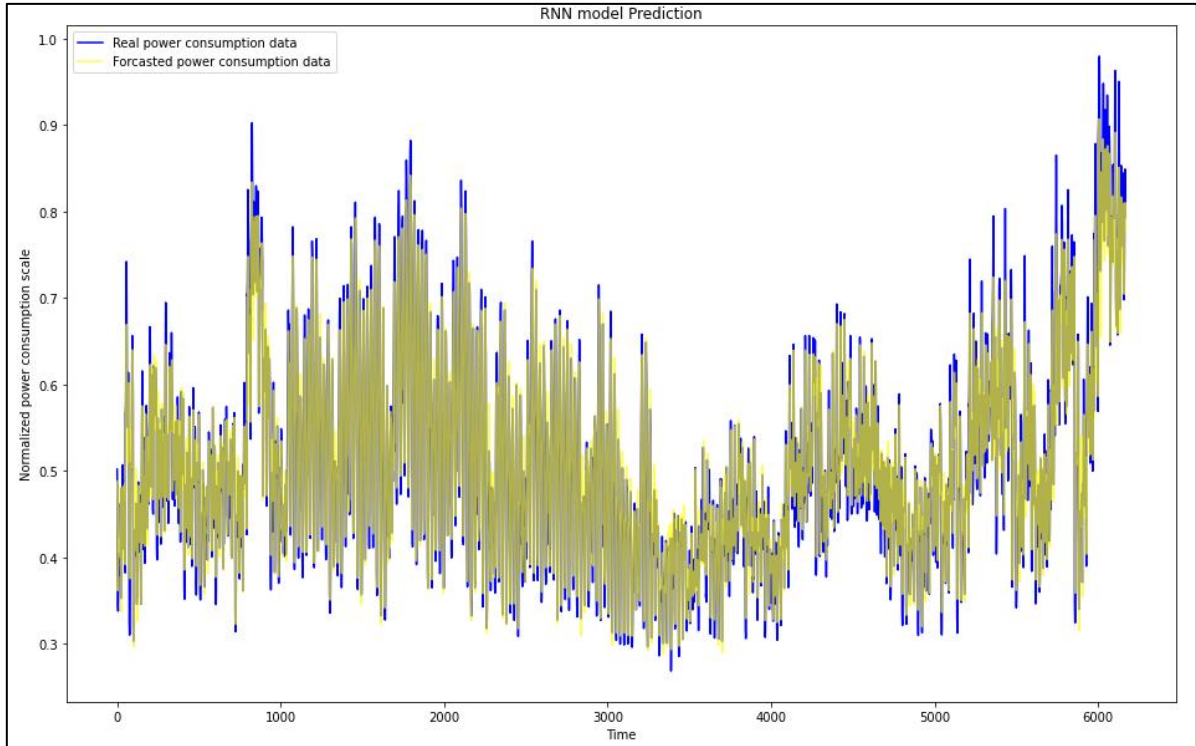


Figure 4.8: RNN Model Prediction Vs Real Power Consumption Data of DOM.

4.2.4 LSTM Model of DOM Power Consumption Dataset

For the LSTM model of DOM dataset, after data sequential, adam optimizer, and MSE for the loss results were used for the training of dataset. Figure 4.9 shows the LSTM prediction result compared with real data. There was little variation between them. However, as time progressed a bit of fluctuation was observed compared to initial stage. With setting epochs as 10 and the batch size of 1000, the results of R^2 score was achieved as 0.94107854. This value was about the same as that of RNN model for this dataset. However, in this case there was a very little variation but not worth of rejecting the prediction result. Figure 4.9 showed that wave of predicted energy consumption model followed that similar sequence of real energy consumption data. Moreover, while focusing on the spike of the graph similar behavior of the predicted model followed the actual values. The acceptability of predicted model was clarified with the statistical analysis with R^2 value.

Table 4.3: LSTM Model Properties for The Dataset of DOM Power Consumption Dataset.

Layer (type)	Output Shape	Param #
lstm_1 (LSTM)	(None, 20, 40)	6720
dropout_4 (Dropout)	(None, 20, 40)	0
lstm_2 (LSTM)	(None, 20, 40)	12960
dropout_5 (Dropout)	(None, 20, 40)	0
lstm_3 (LSTM)	(None, 40)	12960
dropout_6 (Dropout)	(None, 40)	0
dense_2 (Dense)	(None, 1)	41

=====
Total params: 32,681
Trainable params: 32,681
Non-trainable params: 0
=====

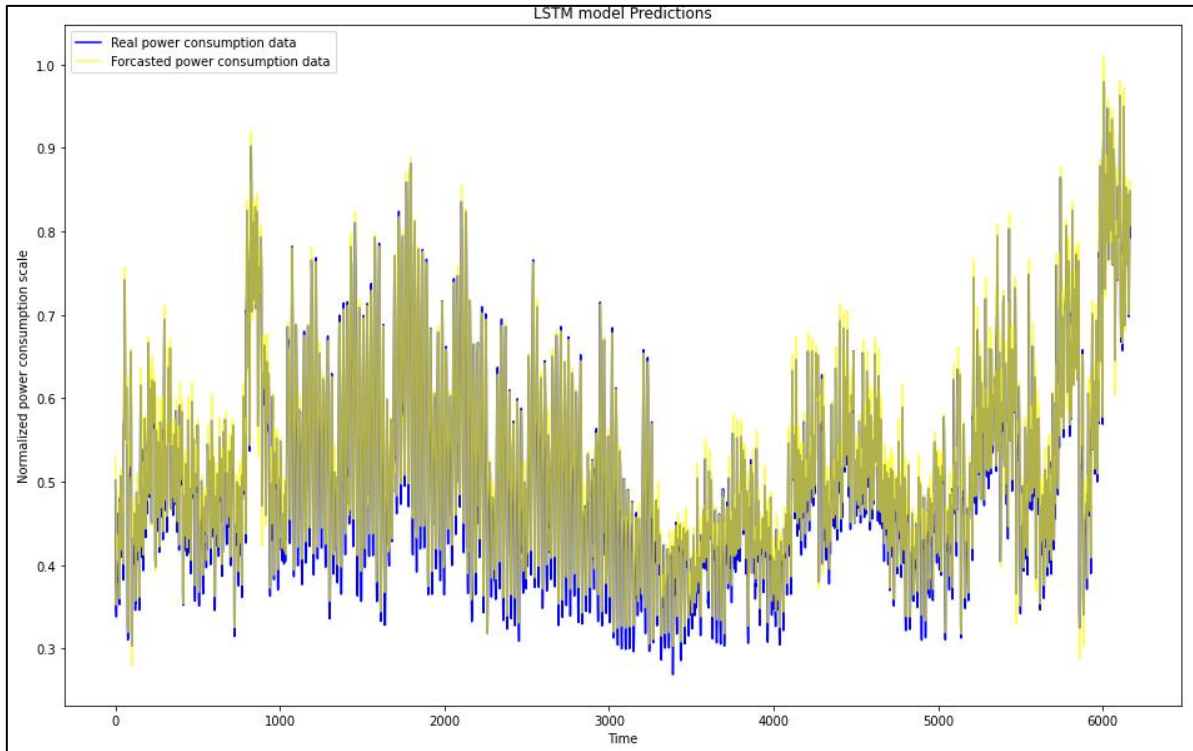


Figure 4.9: LSTM Model Prediction Vs Real Power Consumption Data for DOM.

4.2.5 Comparison Between RNN and LSTM Model Predictions for DOM Power Consumption Dataset

In this section, a comparison between the results of the prediction made by RNN and the results taken from LSTM model have been shown in a single plot to check and compare them. Figure 4.10 represents the comparative analysis where LSTM, RNN and original values were presented with different colors. It can be observed that in different datetimes LSTM and RNN are closer to the original dataset. Both models have successfully predicted the test data with a decent R^2 value (about 0.94) where LSTM model executed near around same R^2 value compared to the RNN model. Hence, both LSTM prediction model and RNN prediction model was accurate in terms of acceptability. Rahman et al. [30] have also ensured that LSTM and RNN could provide more acceptable simulation results. Figure 4.10 demonstrated that wave of both predicted energy consumption model followed that similar sequence of real energy consumption data. Moreover, while focusing on the spike of the graph similar behavior of the predicted model mimicked the actual values. The acceptability of predicted model was clarified with the statistical analysis with R^2 value.

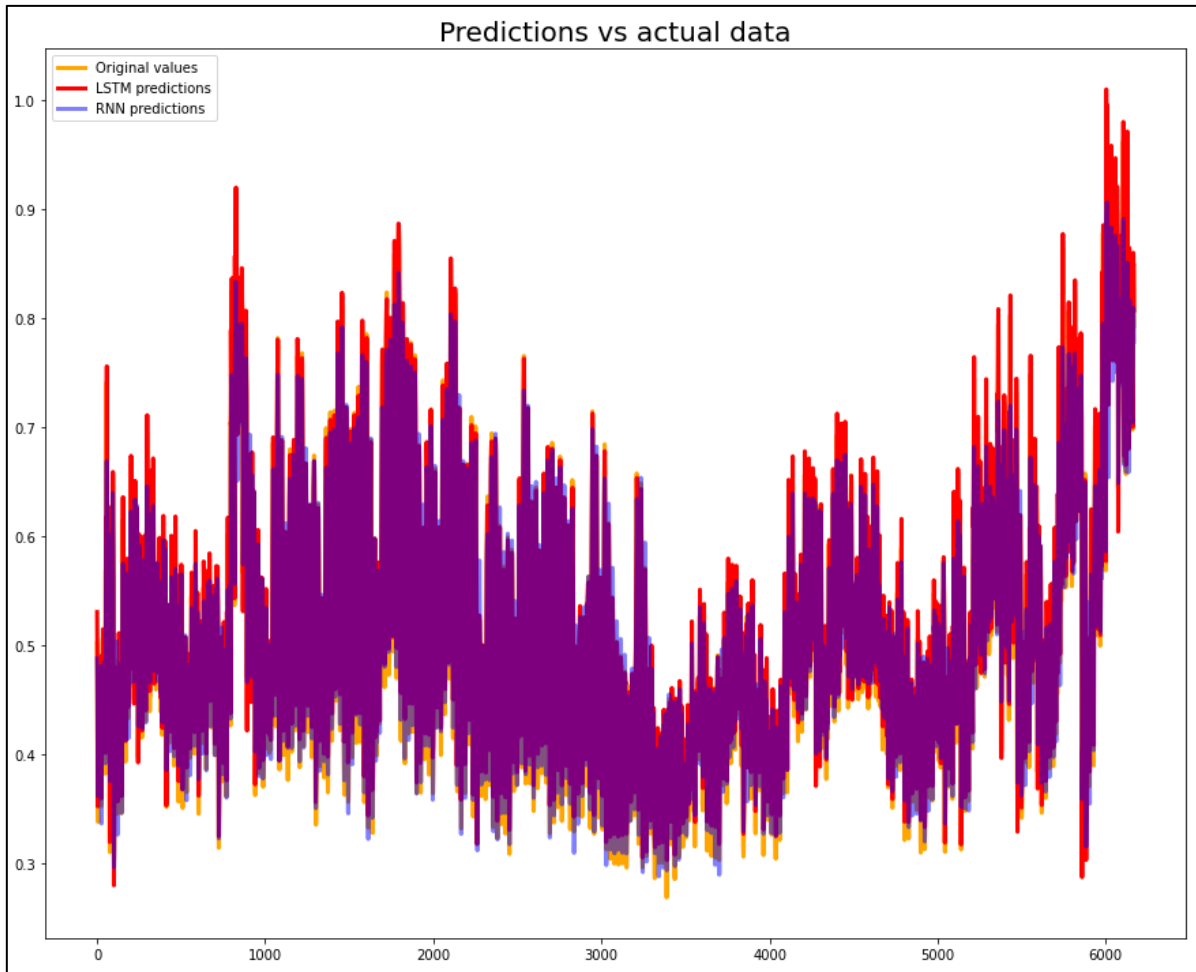


Figure 4.10: Comparison Between RNN and LSTM Model Predictions.

4.3 COMMONWEALTH EDISON (COMED)

This dataset has been taken from Commonwealth Edison (COMED) and it includes 66457 hourly consumed power consumption data. The same process of normalization dataset has been conducted on COMED dataset. Figure 4.11 shows pre-Normalized dataset of COMED and the normalized version is demonstrated in Figure 4.12. As Min-Max scalar was used to normalized data it turned the energy consumption data in the range of 0 to 1 which was previously showing the high values whereas x-axis showed the hourly data. If there is big difference in the two types of data that may conclude the poor prediction results. This necessitated normalizing the data.

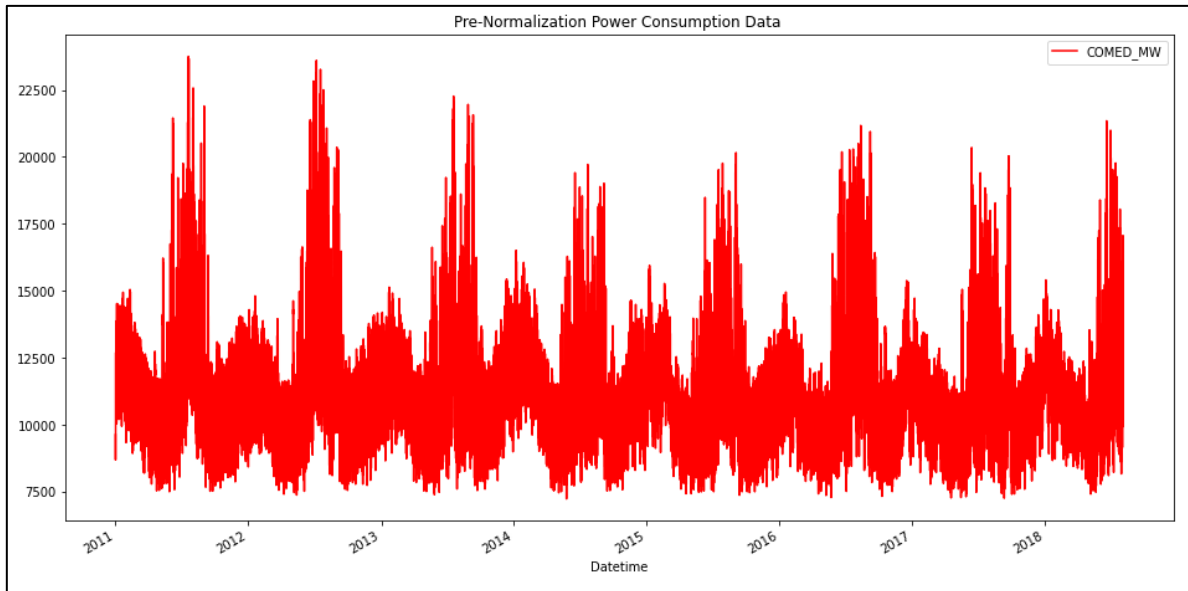


Figure 4.11: Pre-Normalized Comed Power Consumption Data

4.3.1 Data Normalization

The same sklearn MinMaxScaler was implemented to normalize COMED dataset. After normalization, the graph's y-axis shows a change in the range of power consumption figure 4.12. Its range in the previously presented graph was 0 to 22500. Now that the data have been normalized, we can see that the y-axis data range is 0.0 to 1.0. If there is big difference in the two types of data that may conclude the poor prediction results. This necessitated normalizing the data.

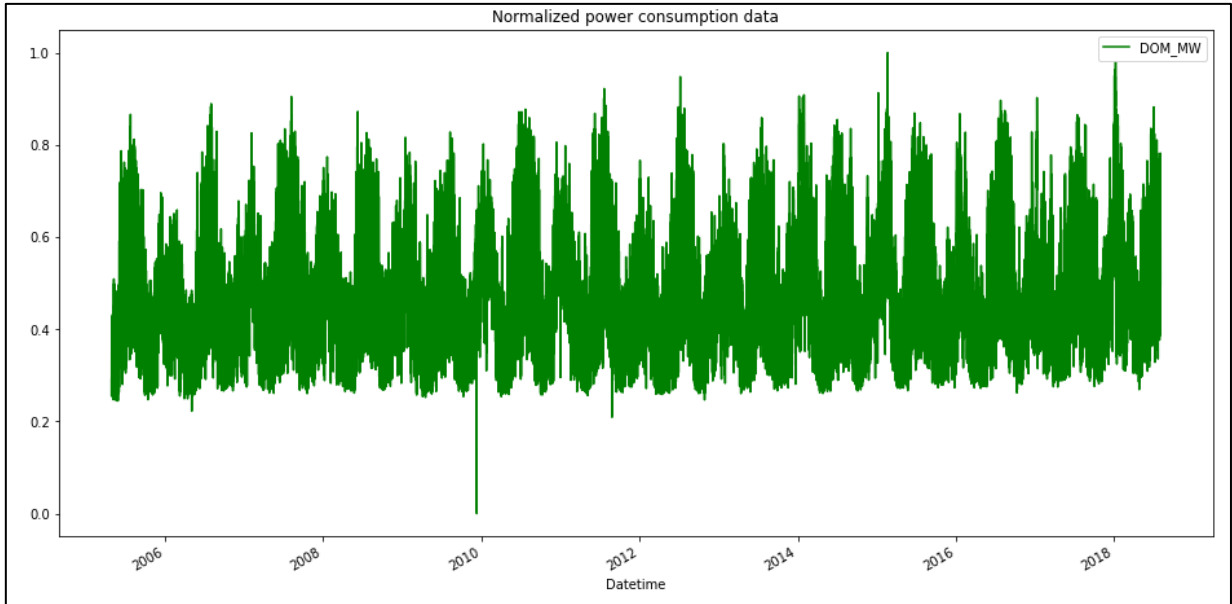


Figure 4.12: Normalized COMED Power Consumption Data Vs Date Time.

4.3.2 Preparation of Data to Implement the RNN Models Training

Figure 4.13 presents the RNN model Prediction vs real power consumption data of COMED. For this dataset to train them for RNN model, after loading data and using sequential data length, we used 6457 hours of data for the test and first 60000 hours were used for training purpose. As it is illustrated in the graph, we can compare the actual numbers to the projected ones. The same observation was observed for this dataset as well, and the predicted values are quite close to the real values, the RNN model is doing an acceptable predicting the sequence, as can be demonstrated. 10 epochs with batch size of 1000 were used for the training and MSE used to find the loss of the results. For RNN model of DOM dataset the R^2 score value of 0.9607209356 was achieved. This higher R^2 represent the most acceptability of RNN model in case of prediction. In accordance with the figure 4.13, it was observed that wave of predicted energy consumption data followed that similar sequence of real energy consumption data. Moreover, while focusing on the spike of the graph similar behavior of the predicted model followed the actual values. The acceptability of predicted model was clarified with the statistical analysis with R^2 value.

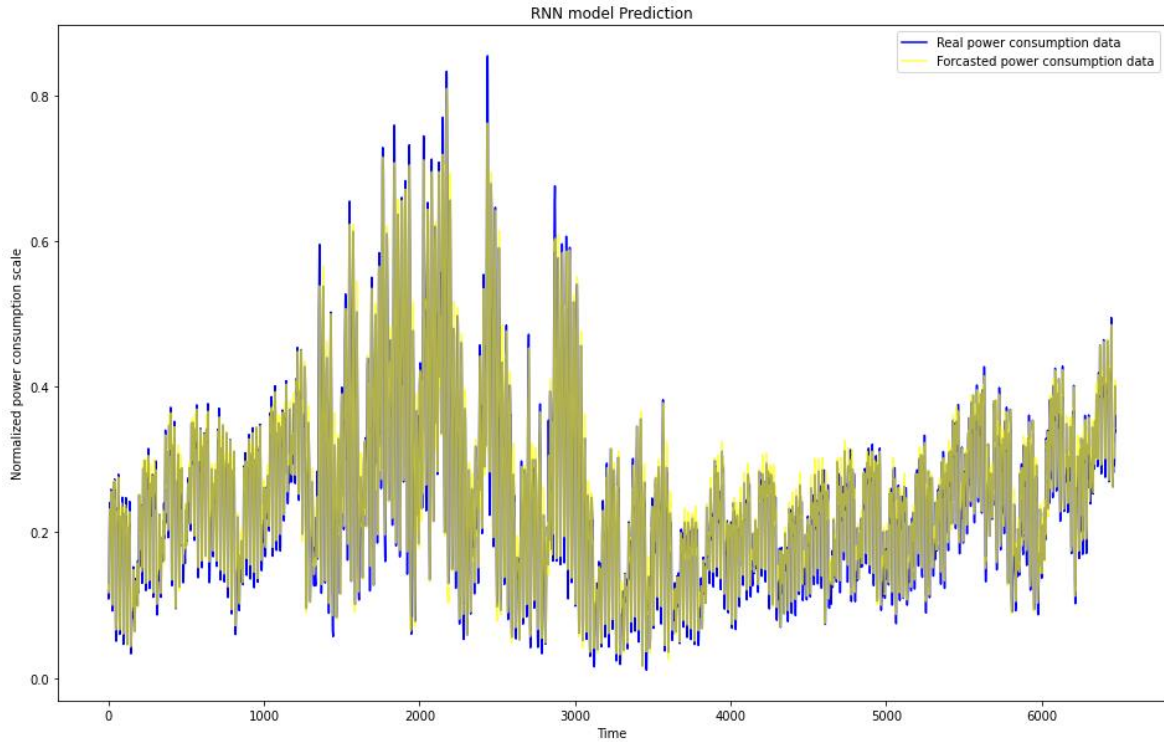


Figure 4.13: RNN Model Prediction Vs Real Power Consumption Data Of COMED.

4.3.3 LSTM Model of COMED Power Consumption Dataset

Figure 4.14 shows the LSTM model Prediction vs real power consumption data for COMED. For the LSTM model of COMED dataset, with setting epochs as 10 and the batch size of 1000, the results of R^2 score was achieved as 0.954822584. This value was lower than that of RNN model for this dataset. The statistical analysis also represents that there was less variation between the original energy consumption data and the predicted energy consumption data from LSTM model. Figure 4.14 showed that wave of predicted energy consumption model followed that similar sequence of real energy consumption data. Moreover, while focusing on the spike of the graph similar behavior of the predicted model followed the actual values. The acceptability of predicted model was clarified with the statistical analysis with R^2 value.

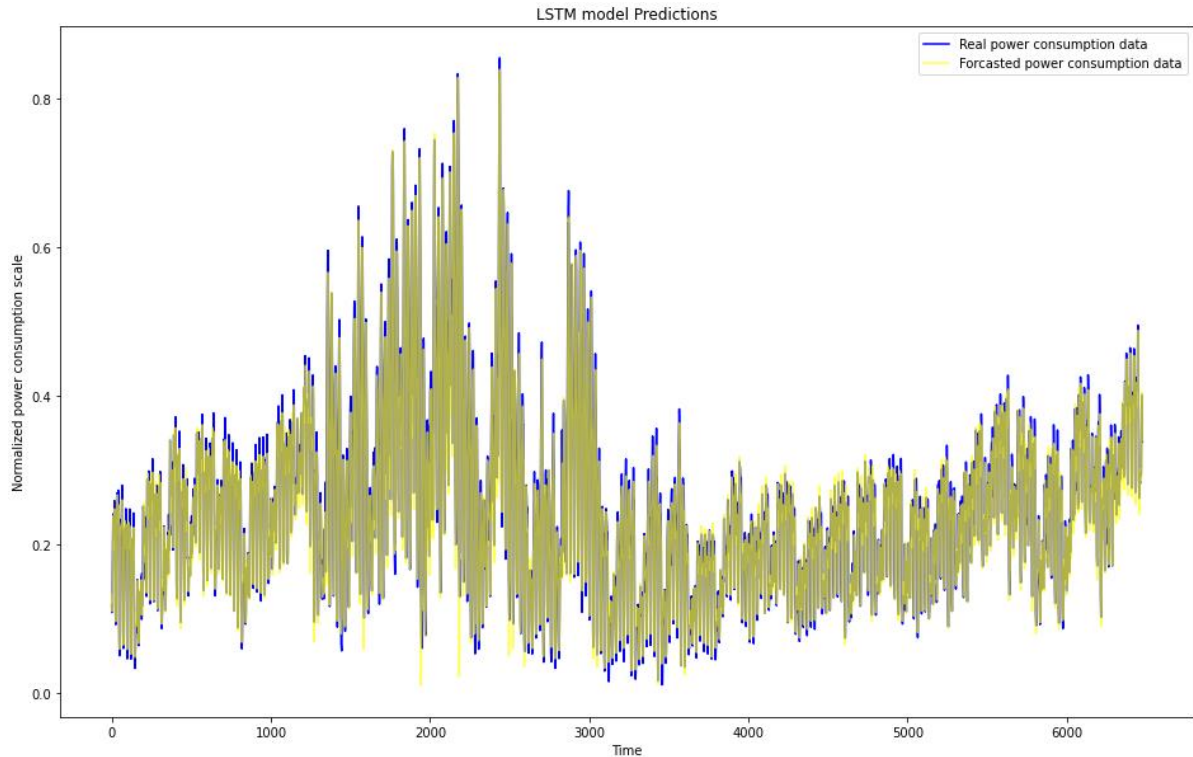


Figure 4.14: LSTM Model Prediction Vs Real Power Consumption Data for COMED

4.3.4 CNN model of COMED Power Consumption Dataset

After loading the data and using sequential data length, it used 6457 hours of data for the test and the first 60000 hours for the training of the CNN model for this dataset. We may compare the actual values to the projected ones, as shown in the graph. The same observation was made for this dataset as well, and the predicted values are very similar to the actual values, showing that the CNN model is doing a respectable job of predicting the sequence. 1000 batches of 10 epochs each were utilized for training, and MSE was used to determine the results' loss. The R^2 score for the DOM dataset's RNN model was 0.9442125489. This higher R^2 represent the most acceptability of CNN model in case of prediction. Figure 4.15 shows the CNN model Prediction vs real power consumption data for COMED. With epochs set to 10 and a batch size of 1000, the CNN model of the COMED dataset got an R^2 score of 0.9442125489. For this dataset, this value was less than the CNN models. The statistical analysis shows that the differences between the actual energy consumption data and the predicted energy consumption data from the CNN model were less significant. The wave of predicted energy consumption model was depicted in Figure 4.15 as following a similar pattern to the real energy consumption data.

Additionally, while concentrating on the graph's spike, comparable behavior of the projected model coincided with the actual values. The statistical study with R^2 value clarified the suitability of the predicted model.

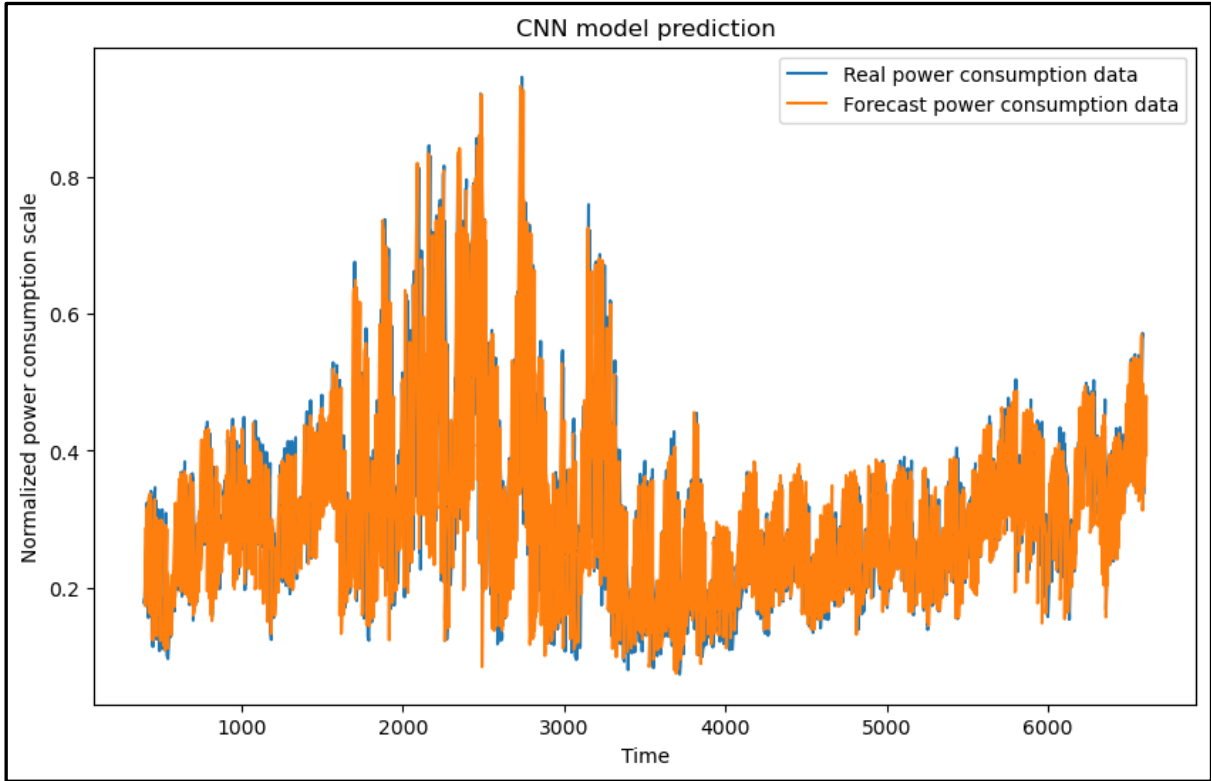


Figure 4.15: CNN Model Prediction Vs Real Power Consumption Data for COMED.

4.3.5 Comparison Among RNN, LSTM And CNN Model Predictions for COMED Power Consumption Dataset

In this section, a comparison between the results of the prediction made by RNN and the results taken from LSTM and CNN model have been shown in a single plot to check and compare them. Figure 4.16 represents the comparative analysis where LSTM, RNN, CNN and original values were presented with different colors. It can be observed that in different date times LSTM and RNN are closer to the original dataset. Both models have successfully predicted the test data with a decent R^2 value (about 0.95) where RNN model resulted higher R^2 value compared to the LSTM and CNN model. Hence, in this case, RNN prediction model was more accurate than LSTM and CNN prediction model in terms of acceptability. Figure 4.15 demonstrated that wave of both predicted energy consumption model followed that similar sequence of real energy consumption data. Moreover, while

focusing on the spike of the graph similar behavior of the predicted model mimicked the actual values. The acceptability of predicted model was clarified with the statistical analysis with R^2 value.

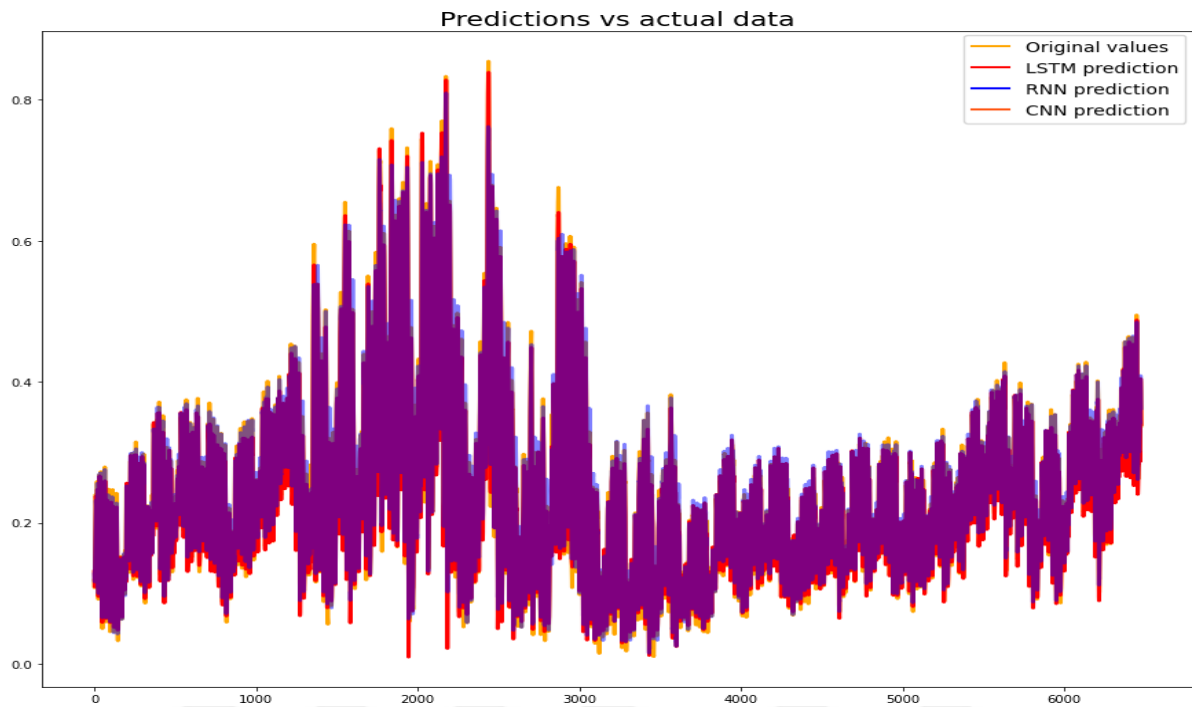


Figure 4.16: Comparison Among RNN, LSTM And CNN Model Predictions.

5. CONCLUSIONS

The prediction of building energy consumption has been the subject of a great deal of study in recent times; nevertheless, reliable information for energy prediction requires the use of strong ML algorithms and a larger variety of datasets. In order to do this, ML seeks to train models on particular datasets so that they can make precise predictions. In this thesis, three different energy consumption datasets from different companies were used to investigate the implementation of RNN and LSTM deep learning algorithms to predict the energy consumption of the data. Python was used to implement the models in Google Colab notebook. All datasets before the processing data for further analysis, the data was normalized using sklearn MinMaxScaler. In all datasets the performances of LSTM and RNN were close to each other. The results revealed that both RNN and LSTM after training the normalized datasets predicted the energy consumption with a good R^2 value of about 0.94 up to 0.97 for LSTM model on different dataset. In case of data collected from AEP showed the highest level of accuracy with R^2 value about 0.96 for both DL method. Above all, LSTM model presented highest performance of predicting energy consumption with accuracy.

REFERENCES

- [1] I. IEA, “World energy outlook 2017 executive summary,” *Energy Policy*. [https://doi.org/10.1016/0301-4215\(73\)](https://doi.org/10.1016/0301-4215(73)), pp. 90024–4, 2017.
- [2] H. Zhao and F. Magoulès, “A review on the prediction of building energy consumption,” *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 3586–3592, 2012.
- [3] J.-S. Chou and N.-T. Ngo, “Time series analytics using sliding window metaheuristic optimization-based machine learning system for identifying building energy consumption patterns,” *Applied energy*, vol. 177, pp. 751–770, 2016.
- [4] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, Art. no. 8, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.
- [5] H. W. Lin and M. Tegmark, “Criticality in formal languages and statistical physics,” *arXiv preprint arXiv:1606.06737*, 2016.
- [6] A. S. Ahmad *et al.*, “A review on applications of ANN and SVM for building electrical energy consumption forecasting,” *Renewable and Sustainable Energy Reviews*, vol. 33, pp. 102–109, 2014.
- [7] M. R. Biswas, M. D. Robinson, and N. Fumo, “Prediction of residential building energy consumption: A neural network approach,” *Energy*, vol. 117, pp. 84–92, 2016.
- [8] I. Khandelwal, R. Adhikari, and G. Verma, “Time series forecasting using hybrid ARIMA and ANN models based on DWT decomposition,” *Procedia Computer Science*, vol. 48, pp. 173–179, 2015.
- [9] Y.-D. Lim, S.-C. Kim, J.-O. Park, and Y.-G. Choo, “A predictive demand of the maximum electric power using chaos-fuzzy,” in *10th IEEE International Conference on Fuzzy Systems.(Cat. No. 01CH37297)*, 2001, vol. 3, pp. 1511–1514.
- [10] D. Monfet, E. Arkhipova, and D. Choinière, “Evaluation of a case-based reasoning energy prediction tool for commercial buildings,” 2013.
- [11] K. Kandananond, “Forecasting electricity demand in Thailand with an artificial neural network approach,” *Energies*, vol. 4, no. 8, pp. 1246–1257, 2011.
- [12] Y. Hu, L. Peng, X. Li, X. Yao, H. Lin, and T. Chi, “A novel evolution tree for analyzing the global energy consumption structure,” *Energy*, vol. 147, pp. 1177–1187, 2018.

- [13] S.-C. Yip, K. Wong, W.-P. Hew, M.-T. Gan, R. C.-W. Phan, and S.-W. Tan, "Detection of energy theft and defective smart meters in smart grids using linear regression," *International Journal of Electrical Power & Energy Systems*, vol. 91, pp. 230–240, 2017.
- [14] P. Chujai, N. Kerdprasop, and K. Kerdprasop, "Time series analysis of household electric consumption with ARIMA and ARMA models," in *Proceedings of the international multiconference of engineers and computer scientists*, 2013, vol. 1, pp. 295–300.
- [15] Z. Tan, J. Zhang, J. Wang, and J. Xu, "Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models," *Applied energy*, vol. 87, no. 11, pp. 3606–3610, 2010.
- [16] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.
- [17] K.-Y. Chen and C.-H. Wang, "A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan," *Expert Systems with Applications*, vol. 32, no. 1, pp. 254–264, 2007.
- [18] L. Pérez-Lombard, J. Ortiz, and C. Pout, "A review on buildings energy consumption information," *Energy and buildings*, vol. 40, no. 3, pp. 394–398, 2008.
- [19] Y.-S. Lee and L.-I. Tong, "Forecasting energy consumption using a grey model improved by incorporating genetic programming," *Energy conversion and Management*, vol. 52, no. 1, pp. 147–152, 2011.
- [20] A. Allouhi, Y. El Fouih, T. Kousksou, A. Jamil, Y. Zeraouli, and Y. Mourad, "Energy consumption and efficiency in buildings: current status and future trends," *Journal of Cleaner production*, vol. 109, pp. 118–130, 2015.
- [21] J. J. Guo, J. Y. Wu, and R. Z. Wang, "A new approach to energy consumption prediction of domestic heat pump water heater based on grey system theory," *Energy and Buildings*, vol. 43, no. 6, pp. 1273–1279, 2011.
- [22] F. Radulovic, M. Poveda-Villalón, D. Vila-Suero, V. Rodríguez-Doncel, R. García-Castro, and A. Gómez-Pérez, "Guidelines for Linked Data generation and publication: An example in building energy consumption," *Automation in Construction*, vol. 57, pp. 178–187, 2015.

- [23] C. Sandels, J. Widén, L. Nordström, and E. Andersson, “Day-ahead predictions of electricity consumption in a Swedish office building from weather, occupancy, and temporal data,” *Energy and Buildings*, vol. 108, pp. 279–290, 2015.
- [24] C. Aghemo, L. Blaso, and A. Pellegrino, “Building automation and control systems: A case study to evaluate the energy and environmental performances of a lighting control system in offices,” *Automation in Construction*, vol. 43, pp. 10–22, 2014.
- [25] S. Zhou, Z. Wu, J. Li, and X. Zhang, “Real-time energy control approach for smart home energy management system,” *Electric Power Components and Systems*, vol. 42, no. 3–4, pp. 315–326, 2014.
- [26] E. Valor, V. Meneu, and V. Caselles, “Daily air temperature and electricity load in Spain,” *Journal of applied Meteorology*, vol. 40, no. 8, pp. 1413–1421, 2001.
- [27] J. Yang, M. Santamouris, S. E. Lee, and C. Deb, “Energy performance model development and occupancy number identification of institutional buildings,” *Energy and Buildings*, vol. 123, pp. 192–204, 2016.
- [28] M. A. M. Daut, M. Y. Hassan, H. Abdullah, H. A. Rahman, M. P. Abdullah, and F. Hussin, “Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review,” *Renewable and Sustainable Energy Reviews*, vol. 70, pp. 1108–1118, 2017.
- [29] A. Costa, M. M. Keane, P. Raftery, and J. O’Donnell, “Key factors methodology—A novel support to the decision making process of the building energy manager in defining optimal operation strategies,” *Energy and buildings*, vol. 49, pp. 158–163, 2012.
- [30] J.-S. Chou and N.-T. Ngo, “Smart grid data analytics framework for increasing energy savings in residential buildings,” *Automation in construction*, vol. 72, pp. 247–257, 2016.
- [31] Y. M. Lee, L. An, F. Liu, R. Horesh, Y. T. Chae, and R. Zhang, “Applying science and mathematics to big data for smarter buildings,” *Annals of the NEW YORK Academy of Sciences*, vol. 1295, no. 1, pp. 18–25, 2013.
- [32] P. J. Brockwell and R. A. Davis, *Introduction to time series and forecasting*. Springer, 2002.

- [33] C. Deb, F. Zhang, J. Yang, S. E. Lee, and K. W. Shah, "A review on time series forecasting techniques for building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 74, pp. 902–924, 2017.
- [34] L. De Boeck, S. Verbeke, A. Audenaert, and L. De Mesmaeker, "Improving the energy performance of residential buildings: A literature review," *Renewable and Sustainable Energy Reviews*, vol. 52, pp. 960–975, 2015.
- [35] A. Azadeh, M. Saberi, and O. Seraj, "An integrated fuzzy regression algorithm for energy consumption estimation with non-stationary data: a case study of Iran," *Energy*, vol. 35, no. 6, pp. 2351–2366, 2010.
- [36] R. Mena, F. Rodríguez, M. Castilla, and M. R. Arahál, "A prediction model based on neural networks for the energy consumption of a bioclimatic building," *Energy and Buildings*, vol. 82, pp. 142–155, 2014.
- [37] K. Li, C. Hu, G. Liu, and W. Xue, "Building's electricity consumption prediction using optimized artificial neural networks and principal component analysis," *Energy and Buildings*, vol. 108, pp. 106–113, 2015.
- [38] N. Arghira, L. Hawarah, S. Ploix, and M. Jacomino, "Prediction of appliances energy use in smart homes," *Energy*, vol. 48, no. 1, pp. 128–134, 2012.
- [39] B. B. Ekici and U. T. Aksoy, "Prediction of building energy consumption by using artificial neural networks," *Advances in Engineering Software*, vol. 40, no. 5, pp. 356–362, 2009.
- [40] H. C. Jung, J. S. Kim, and H. Heo, "Prediction of building energy consumption using an improved real coded genetic algorithm based least squares support vector machine approach," *Energy and buildings*, vol. 90, pp. 76–84, 2015.
- [41] K.-H. Kim and S.-B. Cho, "Modular Bayesian networks with low-power wearable sensors for recognizing eating activities," *Sensors*, vol. 17, no. 12, p. 2877, 2017.
- [42] M. I. Ahmad, "Seasonal decomposition of electricity consumption data," *Review of Integrative Business and Economics Research*, vol. 6, no. 4, p. 271, 2017.
- [43] T.-Y. Kim and S.-B. Cho, "Predicting residential energy consumption using CNN-LSTM neural networks," *Energy*, vol. 182, pp. 72–81, Sep. 2019, doi: 10.1016/j.energy.2019.05.230.

- [44] Y. Bengio, A. Courville, and P. Vincent, “Representation Learning: A Review and New Perspectives,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, Art. no. 8, Aug. 2013, doi: 10.1109/TPAMI.2013.50.
- [45] J. Donahue *et al.*, “Long-term recurrent convolutional networks for visual recognition and description,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 2625–2634.
- [46] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, “LSTM: A search space odyssey,” *IEEE transactions on neural networks and learning systems*, vol. 28, no. 10, pp. 2222–2232, 2016.
- [47] J. Wang, L.-C. Yu, K. R. Lai, and X. Zhang, “Dimensional sentiment analysis using a regional CNN-LSTM model,” in *Proceedings of the 54th annual meeting of the association for computational linguistics (volume 2: Short papers)*, 2016, pp. 225–230.
- [48] T. N. Sainath, O. Vinyals, A. Senior, and H. Sak, “Convolutional, long short-term memory, fully connected deep neural networks,” in *2015 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, 2015, pp. 4580–4584.
- [49] A. Ullah, J. Ahmad, K. Muhammad, M. Sajjad, and S. W. Baik, “Action recognition in video sequences using deep bi-directional LSTM with CNN features,” *IEEE access*, vol. 6, pp. 1155–1166, 2017.
- [50] S. L. Oh, E. Y. Ng, R. San Tan, and U. R. Acharya, “Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats,” *Computers in biology and medicine*, vol. 102, pp. 278–287, 2018.
- [51] R. Zhao, R. Yan, J. Wang, and K. Mao, “Learning to monitor machine health with convolutional bi-directional LSTM networks,” *Sensors*, vol. 17, no. 2, p. 273, 2017.

