LOAD ANALYSIS AND FORECASTING OF THE IRAQI ELECTRICAL DISTRIBUTION GRID USING MACHINE LEARNING TECHNIQUES

By

Noor Ghazi Bdaiwi

Supervisor

Prof. Gheith Ali Abandah

This thesis was submitted in partial fulfillment of the requirements for the master's degree in computer and network engineering

School of Graduate Studies

The University of Jordan

August 8, 2022

COMMITTEE DECISION

This Thesis (Load Analysis and Forecasting of the Iraqi Electrical Distribution Grid Using Machine Learning Techniques) was Successfully Defended and Approved on August 8, 2022.

Examination Committee	**Signature**
Dr. Gheith Abandah, (Supervisor) Professor of Computer Engineering	-------------------------
Dr. Ramzi Saifan, (Member) Associate Professor of Computer Engineering	------------------------
Dr. Mohammad Abdel-Majeed, (Member) Associate Professor of Computer Engineering	------------------------
Dr. Hisham Almasaeid, (Member) Associate Professor of Computer Engineering (Yarmouk University)	-----------------------

AUTHORIZATION

I, Noor Ghazi Bdaiwi,

authorize The University of Jordan to provide copies of my thesis to libraries or establishments or individuals on request, based on the regulations of the University.

Signature:

…………………………

DEDICATION

To my great Father,

To my beloved Mother,

To all my family members,

To all my professors, computer and network engineering experts, and academic staff in my beloved university; The University of Jordan.

Also, this work is dedicated to my beloved Friends,

To every person who have supported me to conduct this work.
Moreover, I do not forget my great country, Iraq,

To all of you,

I am honored to dedicate this work.

Noor Ghazi Bdaiwi

ACKNOWLEDGMENT

In the name of Allah, the Merciful, and blessings and peace be upon Prophet Muhammad, I would like to thank God as he supported me in accomplishing this thesis. Additionally, I am indebted to acknowledge and thank my supervisor, Professor Gheith Ali Abandah, who have guided me during the work of this thesis. Prof. Gheith Ali Abandah gave me valuable effort and time to carry out this study. My acknowledgment would also be to the faculty members of the computer and network engineering department at The University of Jordan, my beloved university. Those doctors and academic experts who have provided me with substantial and valuable support and advice with highly appreciated kindness. Furthermore, I am happy to express my commitment and gratitude to all persons who offered guidance and advice to perform this study. Moreover, I would also like to acknowledge the members of my Examination Committee with the deepest gratitude, as they shared their advice, valuable time, and efforts to read and review my work.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANN	Artificial Neural Networks
ACs	Air Conditioners
ARIMA	Autoregressive Integrated Moving Average () and (
BGC	Al-Basrah Gas Company
BCM	Billion Cubic Meters
CCGT	Combined Cycle Gas Turbines
DL	Deep Learning
GHG	Greenhouse Gas
GPU	Graphics Processing Unit
HFO	Heavy Fuel Oil
LSTM	Long Short-Term Memory
MW	Megawatt
PV	Photovoltaic
RNNs	Recurrent Neural Networks
RMSE	Root Mean Squared Error
T&D	Transmission and Distribution
TPU	Tensor Processing Unit

LOAD ANALYSIS AND FORECASTING OF THE IRAQI ELECTRICAL DISTRIBUTION GRID USING MACHINE LEARNING TECHNIQUES

By

Noor Ghazi Bdaiwi

Supervisor

Prof. Gheith Ali Abandah

ABSTRACT

This work was carried out to investigate and examine the critical contributions and beneficial role of machine learning algorithms in predicting the electrical load distribution in the Iraqi network with high accuracy. Datasets are collected and prepared from the Ministry of Electricity in Iraq for four years (2018, 2019, 2020, and 2021). These datasets are associated with electricity distribution in fifteen Iraqi provinces. Further, two approaches, ARIMA and LSTM, are investigated to forecast the required load for multiple periods in the future. The ARIMA algorithm is used to forecast the required load for one month as a case study in 2022. While the LSTM cells are used to build a deep learning model to predict the required load through a specific time within the four years, relying on the load of past days. The work findings reveal that the LSTM approach a lower RMSE value of 163 and 4% as a NRMSE value when the 14 past days are used to predict one future day with the following parameters: batch size is 128, optimizer function is adam, number of layers is two layers as in LSTM architecture, early stopping, and number of neurons is five neurons. Furthermore, the results indicated that the ARIMA and LSTM algorithms achieve better accuracy in the forecasting of load process of one month of 2022.

CHAPTER ONE

INTRODUCTION

1.1 Research Background

Due to rapid increase in the world populations, the demand on the electrical power has been significantly increasing. The growth in electrical consumption led to high pressure on the electrical grid and its infrastructure to meet all customers' needs. Thus, it is considerably desired to understand customers' load data and its nature to predict its behavior and forecast periods at which consumption will have off-peak and on-peak times to assure high level of safety and efficiency. Load forecasting is remarkably beneficial for managing electrical consumption and achieve better control, monitoring, and planning of power distribution grids based on up-to-date database.

Iraq has developed an energy plan to increase the contribution of renewable energy projects in generating clean electrical power to mitigate the climate change and global warming concerns. This plan was issued in June 2013. The aim was to increase the clean electrical power production capacity from (10 to 12 GW) to 20 GW before 2015. Iraqi government planned to build 8 GW by 2019, nevertheless, just 4 GW of renewable energy projects was installed, leading to an overall capacity of 16.5 GW. Additionally, private diesel generator bought by citizens are considerably noisy, costly, toxic. Also, they cannot meet electrical power demand when ACs are operated. In 2013, the Iraq South Gas Company (51%), Royal Dutch Shell (44%), and Mitsubishi Corporation (5%) established Al-Basrah Gas Company (BGC) to capture associated gas from the Rumaila, Zubair, and West Qurna oil fields to provide electrical power generation (Mills and Salman, 2020). Capturing natural gas has grown to around 7.2 billion cubic meters (BCM) per year since then, accounting for roughly a third of related gas output (26 BCM). Despite this great achievement, about half of related gas is still flared. Iran began selling gas to Iraq in June 2017, via a pipeline that runs through Diyala to Baghdad, with supplies to Basra following in 2018. Nevertheless, the USA opposed these steps and pushed Iraq to stop doing so. Despite these efforts, Iraq made little developments to meet the electrical power demand for its citizens. Electrical power production per capita is considered one of the lowest values compared to the Middle East. The interactions between Iraqi manufacturing companies, transportation, distribution, and marketing became remarkably complicated through the last decades, both commercially and technically. For this reason, it became considerably necessary to utilize load forecasting to increase the reliability and performance of the electrical power distribution network (Ministry of Electricity, 2020).

Figure 1.1: Iraqi population growth over the years.

Source: (IOM, 2020).

Through the last decades, several advanced and state of the art approaches have been used in forecasting electrical power load. The aim is to offer effective, safe, reliable, and continuous electrical power provision for all Iraqi citizens. These innovative methods include artificial neural networks (ANN), and deep learning (DL) (Nalcaci *et al*. 2019; Nyandwi and Kumar, 2020; Wen *et al*. 2020a). In this work, these two methodologies are investigated and implemented.

1.2 Problem Statement

Iraqi electrical power sector has passed through a variety of obstacles and challenges that limit its wide growth and infrastructure expansions through the history. Iraq has been suffering from an electric power shortage for more than 17 years. For instance, the electrical power infrastructure was considerably damaged as of the first Gulf War (between 1990 and 1991). Furthermore, the electrical power infrastructure had extremely destroyed due to theft and civil war because of the USA invasion in 2003. Moreover, Iraq faced several difficult international sanctions that limited the electrical grid infrastructure growth and development. Owing to the significant increase in Iraqi populations during the last decades, the demand on the electrical power had grown (Mills and Salman, 2020). This large demand includes the use of high-consumption electrical appliances such as air conditioners (ACs), water heaters, electric heaters, inefficient lighting. The large demand on Iraqi electrical network caused several uncertainties, disruption, and critical issues in the network. Thus, innovative methods for electrical load forecasting, comprising ANN, DL became remarkably crucial to provide sufficient electrical power and meet the needs of all customers without interruption problems or lack of electrical power performance. Numerous citizens used diesel generators as a solution to the electrical power interruption. However, the broad range use of diesel generators led to considerable pollution issues including the emissions of carbon dioxide, sulfur oxides, nitrogen oxides, and the particulate matter (Ahmed *et al*. 2020; Iraqi Ministry of Electricity, 2010). For these reasons, it is remarkably necessary to utilize electrical loads forecasting systems that assure reliable and safe provision of electrical power for all citizens despite the operation of higher consumption electrical devices.

1.3 Research Significance

This research comes with some state of the art relevances. Through the first relevance, this study will shed light on major Iraqi electrical network limitations and difficulties Iraq faced during the last decades. Through the second relevance, this research will review the significance of ANN and DL in load forecasting and their contributional benefits in helping provide stable, reliable, and safe electrical power production and meet all Iraqi citizens consumption needs despite the use of higher consumption electrical appliances. Through the third relevance, the work will conduct simulations and modelling of a case study in Iraq and apply ANN and DL to facilitate the electrical load power forecasting process. Based on the results obtained from the third relevance, this study will propose some recommendations and key suggestions to help Iraqi decision-makers and corresponding engineering working in the Iraqi electrical power sector make effective modifications and improvements in the Iraqi load forecasting for reliable electrical power provision.

1.4 Research Aim and Objectives

This work is conducted to implement ANN and DL in electrical load forecasting to offer effective, safe, reliable, and continuous electrical power provision for all Iraqi citizens. To achieve the major aim of this research, some objectives are implemented including:

- To shed light on significant Iraqi electrical network limitations and difficulties faced by Iraq through the last decades.
- To review the significance of ANN and DL in the load forecasting and their beneficial role in providing reliable, stable, and safe electrical power production to meet all Iraqi citizens consumption needs despite the use of higher consumption electrical appliances.
- To conduct simulations and modelling of a case study in Iraq and apply ANN and DL to facilitate the electrical load power forecasting process.
- To use the Python code for investigating, modeling, and simulating the Iraqi case study selected.
- To implement long-term load forecasting and compare the accuracy of ANN and DL load forecasting with conventional methods.
- To validate and modify the results obtained and make critical recommendations for decision-makers and corresponding engineering to improve the Iraqi load forecasting for reliable electrical power provision.

1.5 Research Questions

This study is conducted to provide some answers to the research questions, presented in the following paragraphs:

- What are the major Iraqi electrical network limitations and difficulties faced by Iraqi citizens through the last decades?
- How can the use of ANN and DL in the load forecasting mitigate these challenges?
- What are the beneficial impacts and contributional roles of implementing ANN and DL in the load forecasting of electrical load demand?
- To what degree are the ANN and DL accurate in making long-term load forecasting of electrical load demand compared to conventional methods?
- What are major recommendations and key solutions implemented to provide reliable electrical power for all Iraqi citizens?

1.6 Research Hypotheses

This research is achieved to offer the validity and invalidity of some hypotheses, presented in the following paragraphs:

- First Hypothesis - Null part, $H_{1,o}$: “Implementing ANN and DL in the load forecasting would not mitigate challenges and obstacles faced by the Iraqi electrical power sector”.

- First Hypothesis – Alternative part, $H_{1,a}$: “Implementing ANN and DL in the load forecasting can mitigate challenges and obstacles faced by the Iraqi electrical power sector”.

- Second Hypothesis - Null part, $H_{2,o}$: “Implementing ANN and DL in the load forecasting have lower accuracy than other conventional load forecasting methods”.

- Second Hypothesis – Alternative part, $H_{2,a}$: “Implementing ANN and DL in the load forecasting have higher accuracy than other conventional load forecasting methods”.

1.7 Thesis Statement

The statement of this thesis is: *"Load Analysis and Forecasting of the Iraqi Electrical Distribution Grid Using Machine Learning Techniques"*. The thesis will use ANN and DL to make load forecasting to assure that all Iraqi citizens receive sufficient electrical power from the Iraqi electrical power network despite the variation in the electrical demand.

1.8 Thesis Structure

- **Chapter one** provided an illustration of research background, problem statement, research significance, research main goals, research questions, research hypotheses, and thesis statement. The thesis will present more details on this research depending on the following sequence:
- **Chapter Two** presents the ***Literature Review*** of this research. It presents a discussion from recent articles and research publications that address the major significances and critical roles of ANN and DL in making effective load forecasting.
- **Chapter Three** presents the ***Research Methodology*** of this research. It presents the procedure and steps followed by the researcher to conduct this work. It also shows the major dataset defined in the Python code and key variables to make simulations and modelling of the case study using ANN and DL.
- **Chapter Four** presents the ***Results and Discussions*** of the work. It shows the main findings obtained in this work by the simulations and modelling via Python code. It also presents the discussions relating the findings of this work to other scholars.
- **Chapter Five** illustrates the major ***Conclusion and Recommendations*** associated with the numerical findings obtained in the numerical analysis of this work.

CHAPTER TWO

LITERATURE REVIEW

In this chapter, data are illustrated to indicate the major significances and critical roles of ANN and DL in providing practical load forecasting. Also, chapter two presents some challenges and key obstacles faced by the Iraqi electrical power network.

2.1 Major Iraqi Electrical Network Limitations and Difficulties

Averbukh *et al*. (2019) conducted an analysis to explore and determine major challenges and leading obstacles that hinder the broad development of Iraqi grid infrastructure and electrical power network. Their calculation results and analysis revealed that there are some power problems and electrical issues in the Iraqi electrical grid. These affairs include power losses, voltage variation, and lower power quality.

A study conducted by Mills and Salman (2020) aimed to classify critical challenges and major barriers of developing Iraqi electrical network. The authors depended on data collection and analysis from different publications and resources. Their study results indicated that Iraqi power network faces the inability to match the electrical production with consumption load. They reported that the peak demand reached 26 GW in 2019, which is 58% larger than the electrical production rate. However, it is expected that after 2019, the production rate of electrical power would be increased due to the use of new natural gas resources and renewable energy projects according to the energy mix plan enacted in June 2013. Figure 2.1 illustrates the Iraqi generation growth of the electrical power between 2003 and 2020.

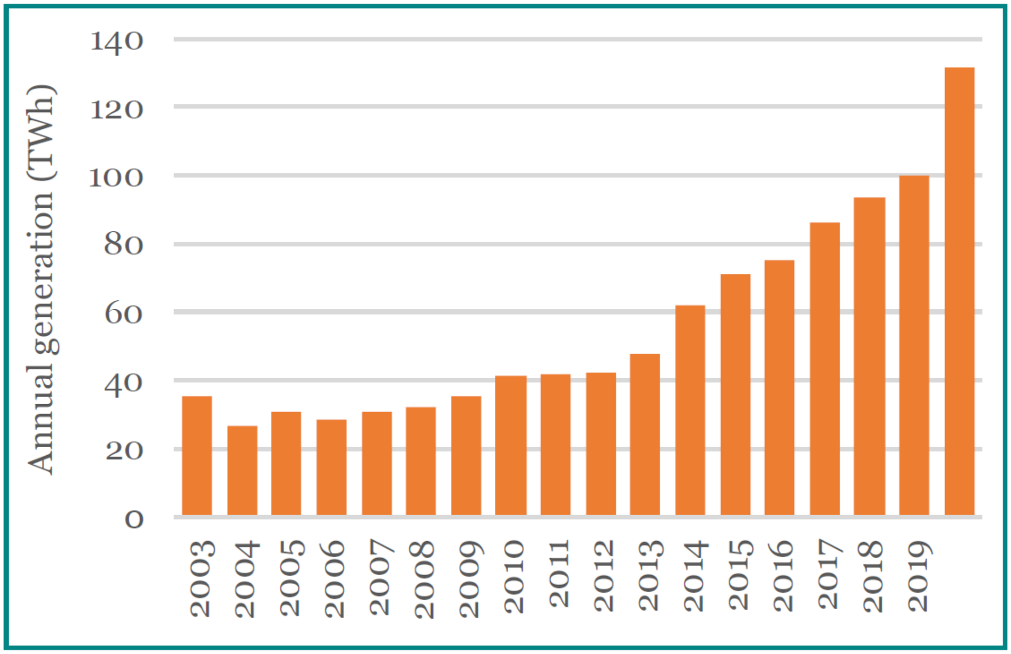

Figure 2.1: The Iraqi generation growth of the electrical power between 2003 and 2020.

(Mills and Salman, 2020).

It is inferred from Figure 2.1 that the electrical power production witnessed a significant increase through the last decades due to the population growth in Iraq. The minimum electrical energy generation was around 25 TWh in 2004, while its value reached approximately 100 TWh in 2019 and about 130 TWh in 2020. The scholars remarked on that some strategic solutions and key methods should be used to solve the shortage of Iraqi electrical production. These methods consist of using small oil generators, heavy fuel oil (HFO), mobile units, rooftop solar Photovoltaic (PV) systems, wind farms, utility gas Combined Cycle Gas Turbines (CCGT), upgraded transmission and distribution (T&D) networks and expanded transmission networks.

Yasen (2016) conducted an analysis to locate major Iraqi network problems. The author depended on Iraqi grid’s stability and performance analysis. He analyzed the frequency, voltage, rotor angle, and energy efficiency. The author’s analysis results revealed that Iraqi electrical network suffer from voltage and frequency instability and fluctuations that influence Iraqi grid’s reliability. Furthermore, the author found that Iraqi power network faces critical obstacles in providing sufficient supply and demand matching. Also, Iraqi electrical gird suffers from losses and in the distribution and transmission systems.

Istepanian (2014) conducted an analysis to investigate several problems existed in the electrical power industry of Iraq. The author depended on a literature review analysis through which numerous articles are revised to address the critical problems and major challenges of Iraqi power network. The conducted literature analysis revealed that Iraq has passed through difficult circumstances and severe conditions that affect the growth and development of its local electrical grid. These challenges include poor political and economic frameworks and the USA invasion, leaving several hard consequences on the network and electrical infrastructure. The solutions of such problems are to consider strategic economic and political frameworks. In addition, the author forecasted the electrical demand of the electrical power in Iraq, as shown in Figure 2.2.

Al-Mahdawi (2021) carried out research to investigate critical solutions that can offer higher performance, stability, and reliability of the Iraqi electrical power grid. The author depended on a comprehensive literature review analysis by which the author reviewed numerous publications to address Iraqi power network issues. The literature analysis revealed that Iraqi electrical network faces several challenges due to the mismatching of supply and demand, lower power quality, and voltage and frequency variations. Therefore, the author explored the critical role of utilizing energy storage technologies including batteries, flywheels, and supercapacitors to mitigate these problems. The author found that these technologies are greatly beneficial in achieving higher stability and reliability to the Iraqi electrical grid.

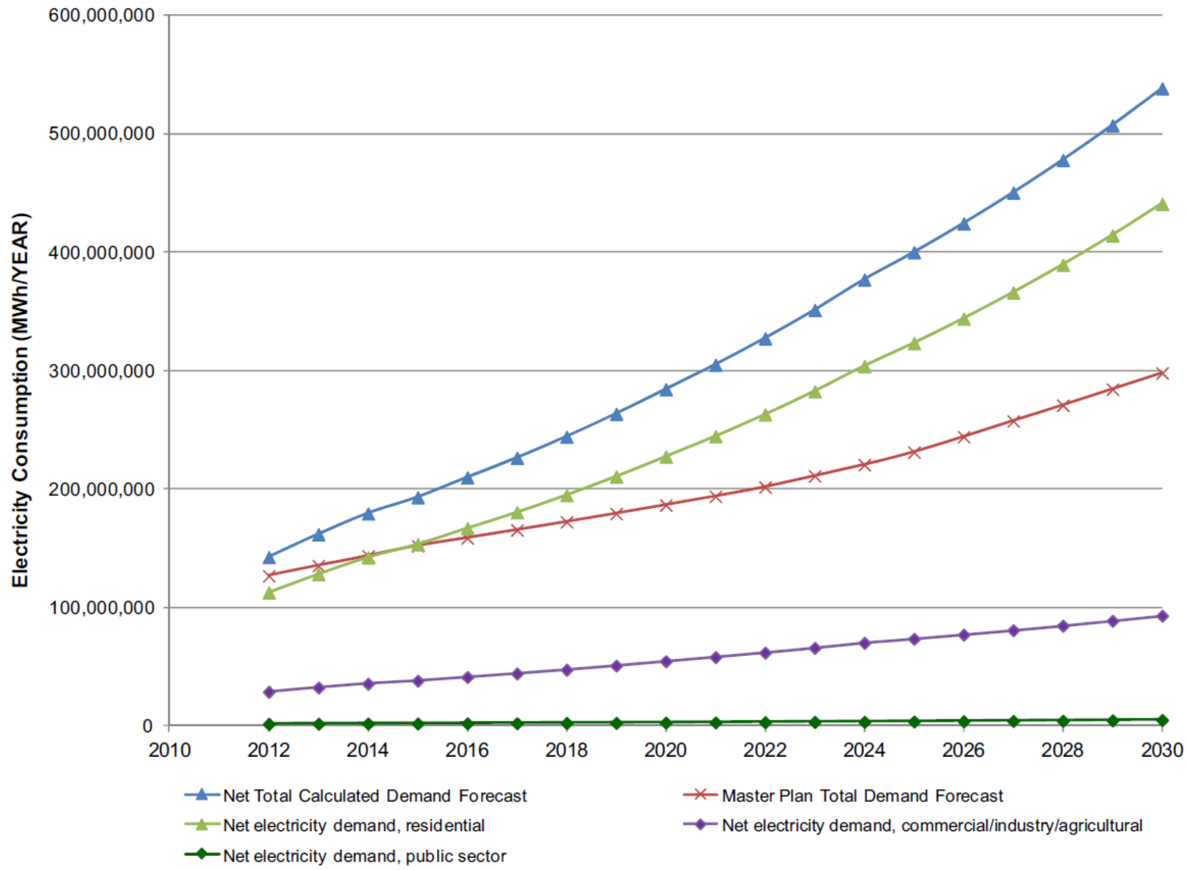

Figure 2.2: The forecasted increase in the electrical power demand in Iraq.

(Istepanian 2014)

As indicated from Figure 2.2, Iraqi power network will witness a significant increase and growth in the electrical power demand due to the populations' growth until 2030.

2.2 Long-Term Electrical Loads Forecasting

Jung *et al*. (2020) conducted an experimental analysis to investigate the critical role of long-term load forecasting for smart cities using Python code. The authors depended on analyzing the monthly loads for a city district according to data transfer learning approach and using several databases from other similar cities (25 numerous districts in Seoul of fourteen years). They considered examining the population, weather data, and calendar. Their experimental results and modelling analysis revealed that their data transfer learning approach along with deep neural networks achieve higher electrical load prediction than other connectional machine learning forecasting methods.

Wen *et al*. (2020b) carried out a study to make a long-term load forecasting of Austin city, Texas, USA. The authors relied on analyzing the hourly load of Austin city (up to 10,000 hours – long-term forecasting) and used Python software to model and implement their forecasting approach. In addition, they compared their deep learning forecasting model to other conventional machine learning methods. Their modelling and analysis results revealed that integrating deep learning into the load forecasting method can greatly improve the prediction accuracy of electrical load related to Austin city.

Ibrahim and Rabelo (2021) executed an analysis to implement a deep learning method to forecast peak loads in Panama City. The authors remarked on the significance of predicting future peak loads due to the significant increase of population. In addition, they mentioned that customers use recently electric vehicles and new applications that need to be considered in the electrical consumption. Hence, deep learning approaches are needed to use to offer higher support levels for load prediction, including short-long term memory and convolutional neural networks. They collected residential electrical consumption and other data between 2004 and 2019 related to Panama City's power network. They used Python software and built a code to validate their study. Their modeling and analysis results revealed that using convolutional neural networks in load forecasting is more accurate than short-long term memory. In addition, the use of convolutional neural networks can accomplish energy management of the grid and assure higher electrical power stability and reliability.

Almazrouee *et al*. (2020) performed a study to explore the major contribution of prophet and holt–winters model in predicting Kuwait load next ten years (long-term prediction between 2020 and 2030). The scholars remarked on the significance of load forecasting due to the rapid population increase through the last decades that contributed to large electrical power consumption. The authors collected realistic electrical load data of Kuwait between 2010 and 2020. Their analysis results revealed that Kuwait will have a maximum peak load of 18,550 MW (using Prophet model), and 19,588 MW (using Holt-Winters model). Furthermore, the utilization of Prophet model and Holt-Winters model can accomplish energy management of the grid and assure higher electrical power stability and reliability.

Jiang *et al*. (2021) conducted research to investigate the considerable role of load forecasting in managing power systems. They applied a medium- and long-term load forecasting for some industrial plants and facilities in Jiangsu, China. Their medium- and long-term load prediction periods are 20 and 40 months. The researchers remarked on the contribution of load prediction in achieving better economic dispatch. The researchers relied on collecting previous load data of thirty industrial facilities and two major industrial plants in Jiangsu, China. Thereafter, a machine learning model is implemented along with load forecasting to improve its accuracy and provide better power measures. Simulations and modelling are used for this purpose. Their modelling and simulations results indicated that using machine learning principles in predicting industrial load is greatly accurate and better than conventional load forecasting methods. Moreover, the employment of machine learning principles in predicting industrial load can accomplish energy management of the grid and assure higher electrical power stability and reliability for industrial facilities.

2.3 Artificial Neural Networks and Deep Learning in Forecasting Electrical Loads

Chitalia *et al*. (2020) conducted a study to investigate the key role of a functional short-term load forecasting method. This approach can offer practical capturing of load variations in several facilities despite their location and type. The authors analyzed nine approaches of clustering and neural networks. Five commercial facilities are tested as case studies located in several areas in (Massachusetts, USA), (Virginia, USA), (Bangkok, Thailand), (Hyderabad-India), and (New York, USA). Their case study analysis revealed that integrating deep learning approaches and neural networks algorithms can achieve between 20% and 45% of better accuracy and effective load forecasting to the building loads of these five case studies.

Nalcaci *et al*. (2019) performed research to examine the critical contribution of the load forecasting according to a long-term basis. They suggested that governments worldwide should consider longer-term load forecasting (2 years) to meet the increasing number of populations, as the electrical consumption is growing annually. In addition, load prediction is significant to enable effective electrical power provision as renewable energy systems, electric vehicles and smart grids are used. They investigated three models using multi-variate adaptive regression splines. They considered some parameters in load forecasting, including temperature, wind speed, humidity rate, and existence of holidays and weekends. Their simulations and investigation results indicated that using artificial neural networks can significantly improve the load forecasting accuracy level and achieve better results. In addition, the use of artificial neural networks can accomplish energy management of the grid and assure higher electrical power stability and reliability.

Kuo and Huang (2018) investigated the ANNs significant role in providing high accuracy of short-term load forecasting. The authors reported the significance of short-term load forecasting in managing smart power grid and achieving higher reliability rate. The researchers stated that linear regression and time series analysis are some conventional methods for load forecasting used in the past that gave less precision. They relied on five ANN algorithms to conduct load prediction. The root mean square and absolute percentage error approaches are used to estimate the algorithms' accuracy. Their analysis results indicated that the accuracy of the root mean square and absolute percentage error approaches reached 11.66% and 9.77%, respectively, indicating great precision potential of load forecasting.

Vanting *et al*. (2021) conducted a study to investigate that key role of DL in achieving accurate electrical load forecasting. The authors depended on a comprehensive literature review through which they reviewed state of the art approaches used to forecast electrical load with higher accuracy rate. Their literature analysis indicated that using DL in load forecasting using long short-term memory network is greatly effective and more accurate than conventional methods. This aspect can help provide more reliable, stable, and safer electrical power for citizens.

Almalaq and Edwards (2017) carried out an analysis to determine major contributions and key significances of DL used for electric load forecasting. The authors depended on a comprehensive literature review by which they revised recent articles and reviewed academic publications that address the beneficial roles of DL approach in accurate load forecasting. Their literature review analysis revealed the importance of using ANN, DL, and artificial intelligence algorithms to provide better load forecasting accuracy for electrical networks and smart grids. Furthermore, results revealed that these innovative approaches can replace the old and conventional methods of load forecasting due to their higher precision and efficiency.

Abdulrahman *et al*. (2021) executed research to investigate the critical benefits and functional impact of integrating DL and ANNs with electrical load forecasting for residential buildings. The authors used a comprehensive literature review approach to address major contributions of DL and ANNs in achieving accurate load forecasting for homes and dwellings. Their literature review analysis revealed that ANNs and DL approaches need has been growing due to the urgent requirement of accurate load forecasting to achieve grid stability and reliability amidst the growing number of populations. In addition, findings indicated that using ANNs and DL approaches in load forecasting for residential facilities is remarkably beneficial due to their higher accuracy compared to conventional methods of load forecasting.

Runge and Zmeureanu (2021) carried out an analysis to address key contributions and major advantages of DL and ANN techniques to the load forecasting of several buildings and facilities. The authors relied on secondary data collection to achieve the goals of their study. They used a comprehensive literature review by which they reviewed several articles and publications that discuss the key role of DL and ANN in load forecasting. Their literature analysis revealed that DL and ANN are vital to provide higher energy efficiency and lower electrical power consumption amidst the rapid increase in population that cause critical environmental issues, incusing climate change and global warming due to the greenhouse gas (GHG) emissions.

Moon *et al*. (2019) performed research to examine the beneficial role of integrating ANNs to the load forecasting approaches to predict the future electrical power of several buildings, helping achieve effective energy management. The authors depended on a comprehensive literature review analysis through which they addressed several characteristics and key strengths of ANNs in managing accurate load forecasting. Their literature review analysis indicated that using ANNs is greatly vital to provide accurate load prediction of buildings connected to the smart grids. Thus, higher efficiency, stability, and power flexibility can be achieved. In addition, results revealed that ANNs help offer practical energy management and good energy performance of the smart grid.

Ozcan *et al*. (2021) carried out a study to investigate major contributions and advantages of ANNs and DL in improving the load forecasting performance of buildings. The authors relied on a comparative analysis by which they compared their model (attention-based dual-stage recurrent neural networks) to conventional load forecasting methods. Their comparative analysis revealed that their model had more accuracy in predicting electrical loads of the investigated buildings connected to the smart grid. Furthermore, the authors found that DL and ANNs can improve the load forecasting performance to achieve better energy management of the smart grid.

2.4 Literature Review Summary

The major results and critical contributions of the literature review publications can be summarized in Table 2.1.

Table 2.1: Major Contributions and Findings of the literature review publications.

#	Author(s) and Year	Major Contributions and Findings
1	Al-Mahdawi (2021)	The literature analysis revealed that Iraqi electrical network face several challenges due to the mismatching of supply and demand, lower power quality, and voltage and frequency variations. Therefore, the author explored the critical role of utilizing energy storage technologies including batteries, flywheels, and supercapacitors to mitigate these problems. The author found that these technologies are greatly beneficial in achieving higher stability and reliability to the Iraqi electrical grid.
2	Ibrahim & Rabelo (2021)	Their modeling and analysis results revealed that using convolutional neural networks in load forecasting is more accurate than short-long term memory. In addition, the use of convolutional neural networks can accomplish energy management of the grid and assure higher electrical power stability and reliability.
3	Jiang *et al*. (2021)	Their modelling and simulations results indicated that using machine learning principles in predicting industrial load is

		greatly accurate and better than conventional load forecasting methods. Moreover, the employment of machine learning principles in predicting industrial load can accomplish energy management of the grid and assure higher electrical power stability and reliability for industrial facilities.
4	Vanting *et al*. (2021)	Their literature analysis indicated that using DL in load forecasting using long short-term memory network is greatly effective and more accurate than conventional methods. This aspect can help provide more reliable, stable, and safer electrical power for citizens.
5	Abdulrahman *et al*. (2021)	Their literature review analysis revealed that ANNs and DL approaches need has been growing due to the urgent requirement of accurate load forecasting to achieve grid stability and reliability amidst the growing number of populations. In addition, findings indicated that using ANNs and DL approaches in load forecasting for residential facilities is remarkably beneficial due to their higher accuracy compared to conventional methods of load forecasting.
6	Runge and Zmeureanu (2021).	Their literature analysis revealed that DL and ANN are vital to provide higher energy efficiency and lower electrical power consumption amidst the rapid increase in population that cause critical environmental issues, incusing climate change and global warming due to the greenhouse gas (GHG) emissions. Additionally, the use of DL and ANN in load forecasting is greatly advantageous as of their higher precision compared to conventional load forecasting methods.

2.5 Chapter Summary

It is indicated from reviewing several articles and recent academic work in numerous research engines and research websites that Iraqi electrical network has passed through a variety of accidents and difficulties that negatively affected its infrastructure and capability to provide sufficient electrical power for Iraqi citizens. In addition, this chapter described diverse significances and beneficial contributions of utilizing ANN and DL in forecasting electrical loads of Iraqi citizens. This aspect is considerably desired to improve the electrical grid efficiency and performance and help meet Iraqi customers loading needs. Following chapter will present the research methodology used to execute this work.

CHAPTER THREE

EXPERIMENTAL SETUP

3.1. Introduction

This chapter presents the theoretical background about the algorithms used, dataset used, Python libraries used, and the evaluation metrics.

3.2. Theoretical Background

The theoretical background is presented for the autoregressive integrated moving average (ARIMA) and long short-term memory (LSTM). In addition, how these models deal with the dataset to accomplish the required task.

3.2.1 ARIMA Algorithm

ARIMA is an acronym for "autoregressive integrated moving average." It's a model used in statistics and econometrics to measure events that happen over a period. The model is used to understand past data or predict future data in a series. It's used when a metric is recorded in regular intervals, from fractions of a second to daily, weekly, or monthly periods. The ARIMA model predicts a given time series based on its own past values (Ediger *et al*, 2007). It can be used for any non-seasonal series of numbers that exhibits patterns and is not a series of random events. For example, sales data from a clothing store would be a time series because it was collected over a period (Ediger *et al*, 2007). One of the key characteristics is the data is collected over a series of constant, regular intervals. A modified version can be created to model predictions over multiple seasons (Nepal *et al*, 2020). The dataset size for each province is 1461 instances and 8 attributes, which are daily data for four years (2018, 2019, 2020, and 2021). First, the ARIMA model reads the dataset in csv format.

We split the into two groups: training and testing datasets. The training dataset is used to build the ARIMA model, and the test dataset is used to assess the performance of this model. The training dataset size is 1,401 instances with 8 attributes, and the rest is for the testing dataset (60 instances with 8 attributes).

After the dataset is split, the model is built based on the training dataset. In the ARIMA model, it takes two parameters: the label of the training dataset and the order. The order parameter takes three values: p = 1, d = 0, and q = 5. The p (order of the autoregressive model) specifies the number of lags used by the autoregressive part of the ARIMA model. The d (degree of differencing) specifies how often the time series values are differentiated. Q (order of the moving-average model) specifies the order of the moving-average part of the model. After the model is built, the test dataset is used to assess the model's performance. The testing dataset used consists of 60 instances that are from the last two months of 2021. The final process of this model is to predict the required load for a future dataset based on the whole dataset of Baghdad. Therefore, we can forecast the required load for the January month of 2022.

3.2.2 LSTM Algorithm

Long short-term memory (LSTM) is a deep learning architecture based on artificial recurrent neural networks (RNNs) (Gers *et al.*, 2012). LSTM has feedback connections, unlike normal feedforward neural networks. It can handle not only individual data points (such as photos), but also complete data streams (such as speech or video). LSTM can be used for tasks like unsegmented, connected handwriting identification, speech recognition, time series and anomaly detection in network traffic or IDSs, for example (intrusion detection systems) (Ghany *et al.*, 2021).

A cell, an input gate, an output gate, and a forget gate make up a typical LSTM unit. The three gates control the flow of information into and out of the cell, and the cell remembers values across arbitrary time intervals. Since there may be lags of uncertain length between significant occurrences in a time series, LSTM networks are well-suited to categorizing, processing, and making predictions based on time series data (Ghany *et al.*, 2021). To solve the vanishing gradient issue that can arise when training conventional RNNs, LSTMs were created. In many cases, LSTM has an advantage over RNNs, hidden Markov models, and other sequence learning techniques due to its relative insensitivity to gap length (Gers *et al*., 2012).

The dataset was split the into two groups: training and testing datasets using train test split function in Python. The training dataset is used to build the LSTM model, and the test dataset is used to assess the performance of this model. The training dataset size is 0.80 of whole instances with 8 attributes, and the rest is for the testing dataset (0.20 of whole instances with 8 attributes). Now we built model based on the training dataset and the rest of the dataset is used to assess the performance of model. Therefore, the size of training dataset is 0.80 of whole instances and the size of testing dataset is 0.20 of whole instances. The architecture of the LSTM model is as follows with several experiments to select the suitable values for batch size, number of neurons, optimizer functions and the number of layers that shown in the next chapter.

3.3 Experimental Setup

3.3.1 Dataset

The dataset that is used in this thesis was collected from the Ministry of Electricity in Iraq for four years (2018, 2019, 2020, and 2021). This dataset is related to electrical distribution in 15 Iraqi provinces, as shown in Table 3.1.

Table 3.1: Dataset Features and description.

Feature	**Description**
Date	Daily Date for four years (2018, 2019, 2020, and 2021)
Year	2018, 2019, 2020, and 2021
Province	1. Al Anbar 2. Al-Qadisiyyah 3. Babil 4. Baghdad 5. Basra 6. Dhi Qar 7. Diyala 8. Karbala 9. Kirkuk

	10. Maysan 11. Muthanna 12. Najaf 13. Nineveh 14. Saladin 15. Wasit
Max Temperature	The Maximum Temperature in °C for each Iraqi province daily
Min Temperature	The Minimum Temperature in °C for each Iraqi province daily
Population	Population for each Iraqi province
Fitted Load*	The Fitted Load for each Iraqi province daily in Megawatt (MW)
Total Load*	The Total Load for each Iraqi province daily in Megawatt (MW)
Label (Required Load)*	The Required Load for each Iraqi province daily in Megawatt (MW)

- **Fitted Load:** It is the actual daily average load that has been prepared for consumers, based on the available generation. It is less than the required load to be achieved, and this is the result of the gap between production and demand.
- **Total Load:** Total of Fitted Load for each Province along 24 hours.
- **Label (Required) Load:** It is the daily average load required to supply electrical energy.

In this thesis, we relied on the required load as a class label and it is forecasted and predicted based on the aforementioned features that change every day. This forecast and predict are done by using two algorithms that mentioned above: ARIMA and LSTM.

Table 3.2 shows the average fitted load, actual average required load, expected average required load, and (b) refers to difference between actual and expected whether positive or negative for 15 Iraq provinces for four years: 2018, 2019, 2020, and 2021.

Table 3.2: Average Fitted Load & Average Required Load for 15 Iraq Provinces.

Province	Average Fitted Load	Average Required Load (a)	b	+/-	Expected Required in 2022
Al-Anbar	466	734	172	+	906
Al-Qadisiyyah	425	546	149	+	695

Babil	638	858	233	+	1091
Baghdad	3273	4322	682	+	5004
Basra	2066	2300	363	+	2663
Dhi Qar	870	1020	295	+	1315
Diyala	608	807	274	+	1081
Karbala	596	762	220	+	982
Kirkuk	630	880	204	+	1084
Maysan	519	643	148	+	791
Muthanna	326	409	87	+	496
Najaf	573	742	173	+	915
Nineveh	759	1365	-33	-	1332
Saladin	500	767	210	+	977
Wasit	579	704	179	+	883
Total	12828	16859	3356	+	20215

Figure 3.1 depicts the fitted load for each Iraqi province over a four-year period in each month. For each year, the average fitted load was calculated for 15 Iraqi provinces.

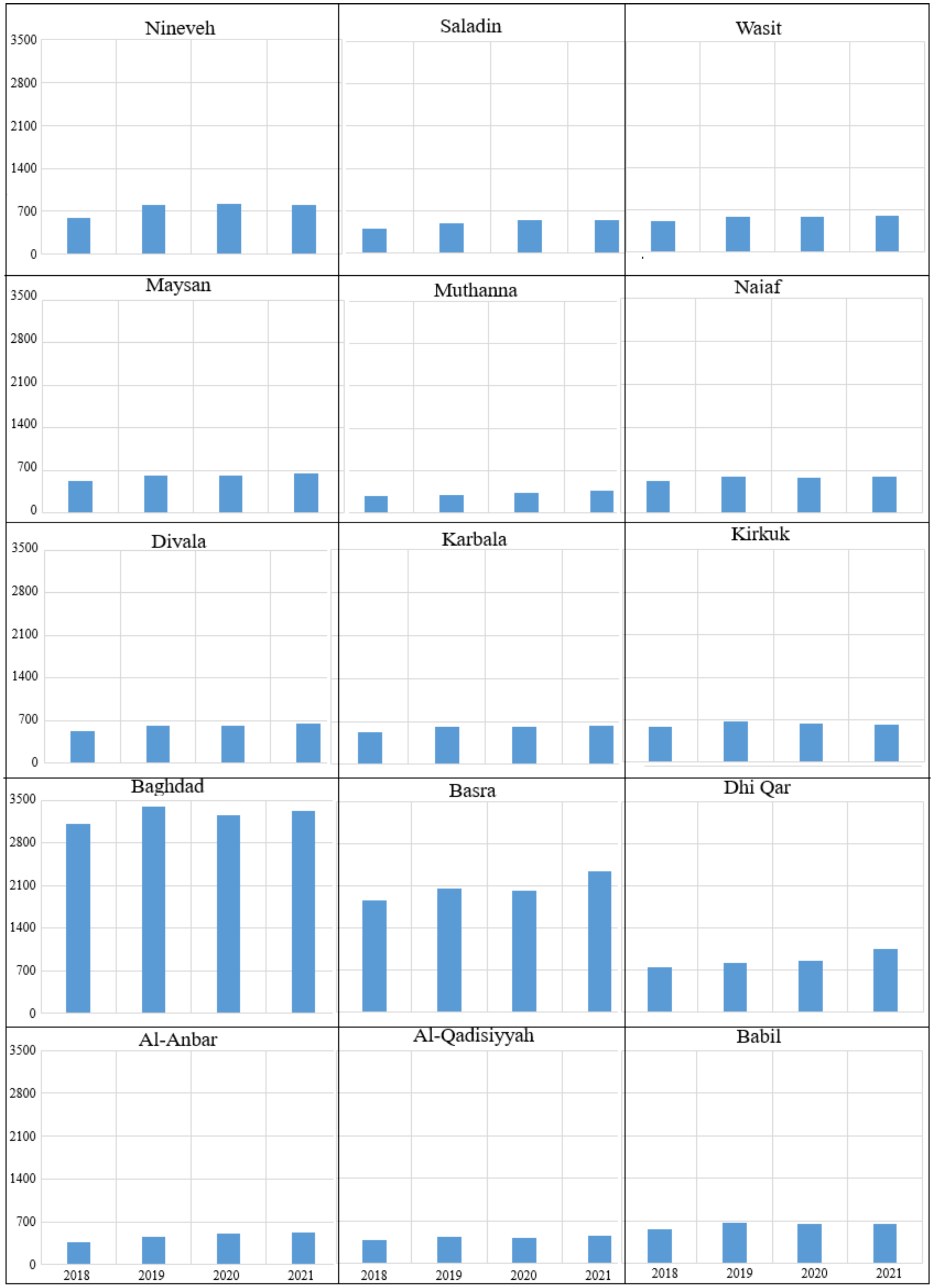

Figure 3.1: Fitted Load in each Iraq Province.

Figures 3.2 depicts the required load for each Iraqi province over a four-year period in each month. For each year, the average required load was calculated for 15 Iraqi provinces.

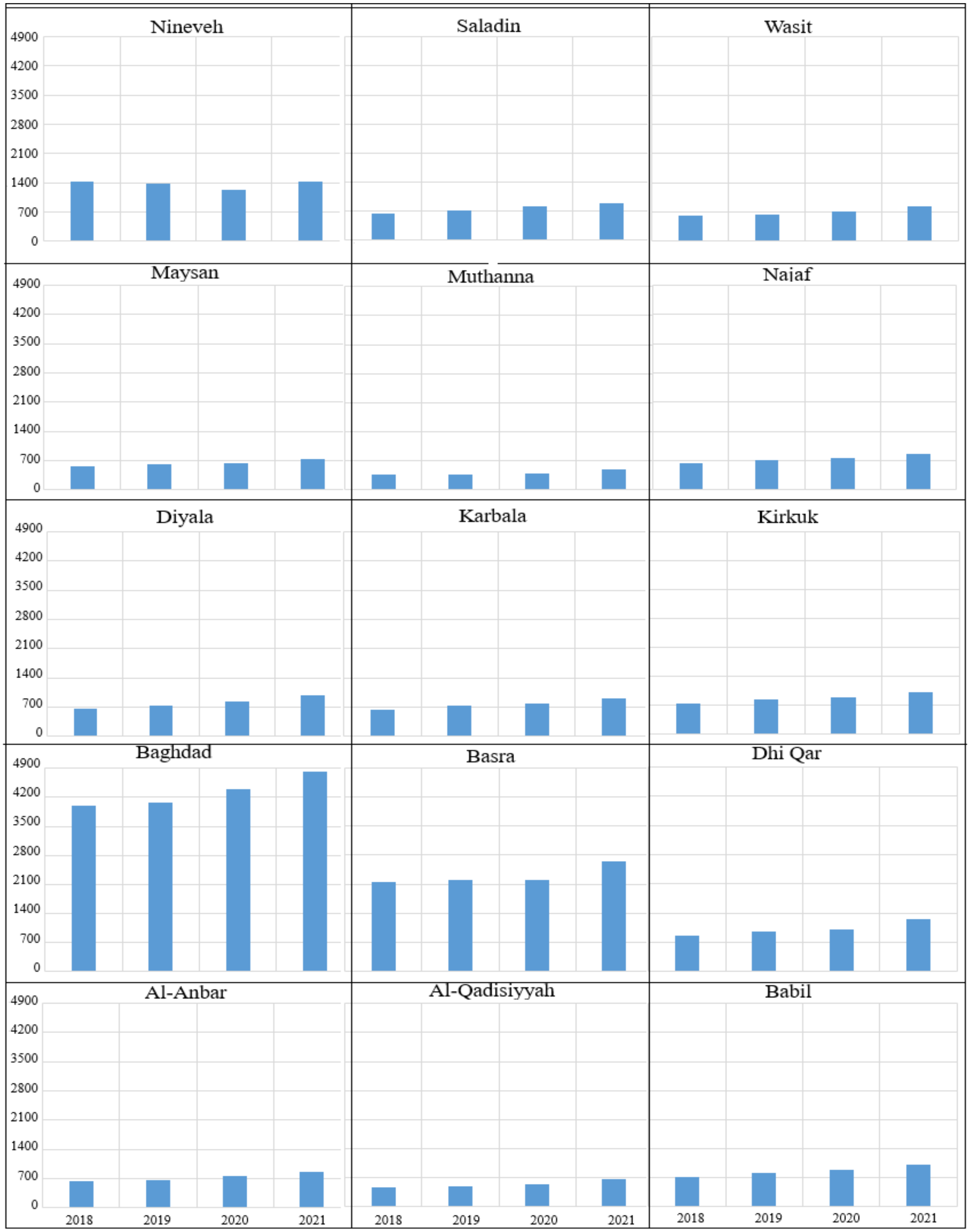

Figure 3.2: Required Load in each Iraq Province.

3.3.2 Dataset Preparation

Figure 3.1 shows the distribution of the required load for Baghdad (class label of the dataset) in 2018, 2019, 2020, and 2021, respectively that include many gaps, which means that the value of required load in many days is zero. While Figure 3.2 shows the distribution of the required load for Baghdad in 2018, 2019, 2020, and 2021, respectively without any gaps. We have filled these gaps by taking the average of required load for previous day and next day.

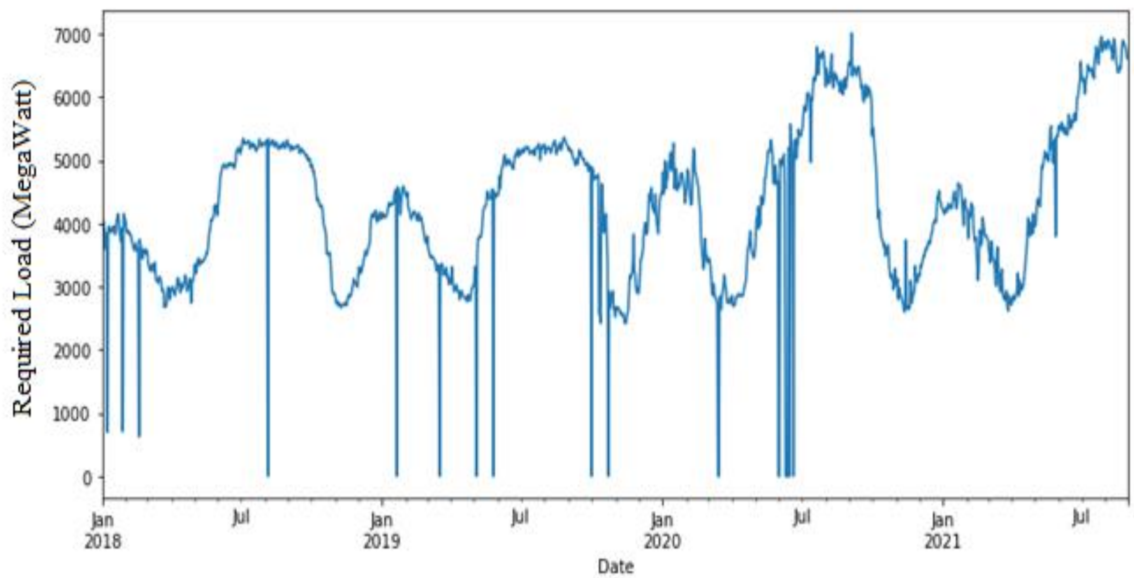

Figure 3.3: Required Label in Four Years with Gaps.

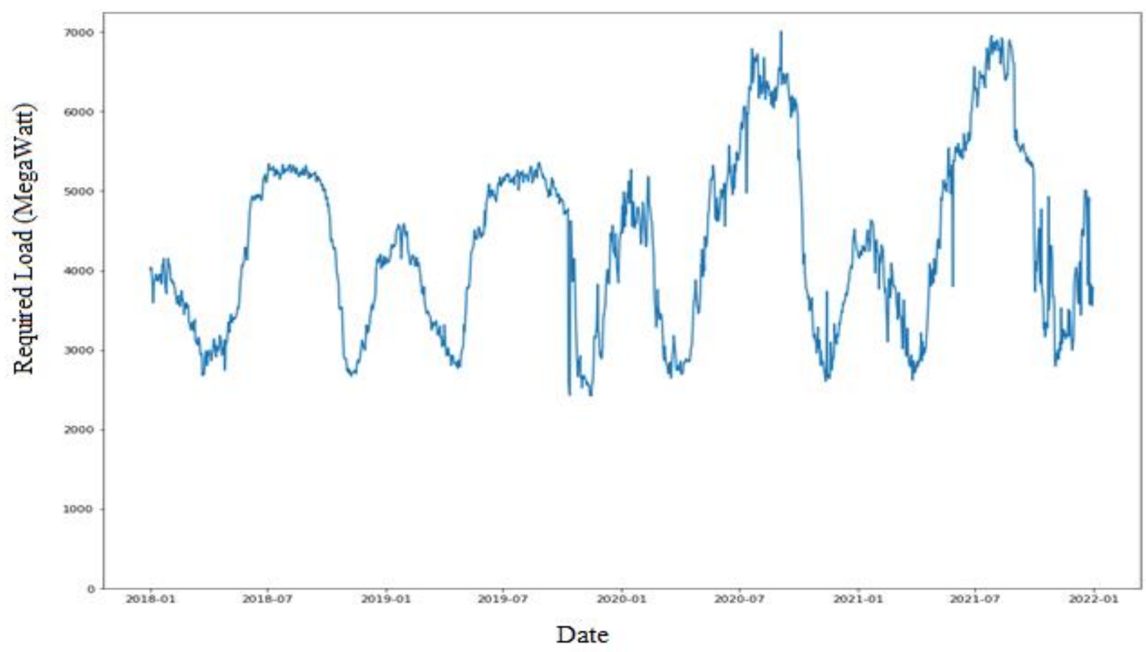

Figure 3.4: Required Label in Four Years with Gaps Removed.

We split the into two groups: training and testing datasets. The training dataset is used to build the ARIMA model, and the test dataset is used to assess the performance of this model. The training dataset size is 1,401 instances with 8 attributes, and the rest is for the testing dataset (60 instances with 8 attributes).

3.3.2 Python Libraries

In our experiments, we used three main libraries with related functions to accomplish the prediction task: Tenserflow, Keras, scikit learn (sklearn).

1. **Keras**

For creating and analyzing deep learning models, Keras is a strong and user-friendly free open source Python library. With only a few lines of code, you can define and train neural network models using this component of the TensorFlow library (Brownlee *et al.*, 2016). To make dealing with image and text data easier and to streamline the coding required to create deep neural networks, Keras includes multiple implementations of widely used neural-network building blocks like layers, objectives, models, activation functions, and optimizers. Convolutional and recurrent neural networks are supported by Keras in addition to traditional neural networks. Other widely used utility layers are supported, including dropout, batch normalization, and pooling. Users of Keras can productize deep models on web-enabled devices, Java Virtual Machines, and smartphones (iOS and Android) (Brownlee *et al*., 2016). Additionally, it permits the deployment of distributed deep learning model training on Graphics Processing Unit (GPU) and Tensor Processing Unit (TPU) clusters. The frameworks supported by Keras are Tensorflow, Theano, PlaidML, MXNet, and CNTK (Microsoft Cognitive Toolkit) (Brownlee *et al*., 2016).

2. TensorFlow

A free and open-source software library for artificial intelligence and machine learning is called TensorFlow. Although it can be applied to many different tasks, deep neural network training and inference are given special attention. The Google Brain team created TensorFlow for use in internal Google research and production (Pang *et al*., 2020). 2015 saw the maiden release under the Apache License 2.0. TensorFlow 2.0, the upgraded version of TensorFlow from Google, was launched in September 2019. Many programming languages, including Python, Javascript, C++, and Java, can use TensorFlow. This adaptability allows for a wide range of applications across numerous industries (Pang *et al*., 2020).

3. Sklearn

Scikit-learn (formerly scikits.learn and also known as sklearn) is a free software machine learning library for the Python programming language. It features various classification, regression and clustering algorithms including support-vector machines, random forests, gradient boosting, k-means and DBSCAN, and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy. It also provided many functions like preprocessing methods (*i.e*., MinMaxScaler, and StandardScaler), or metrics (*i.e.*, accuracy, precision, recall, or mean squared error) (Hao *et al*., 2019).

3.3.3 Evaluation Metrics

This part explained the regression metrics used in our thesis to assess the ARIMA and LSTM algorithms: RMSE, and NRMSE.

Root Mean Square Error or RMSE is a popular regression metric that calculated the error between the actual and prediction label (Almalaq *et al*., 2017). The formula of RMSE is shown in Equation 3.1, where predicted is the required load predicted by the models, actual is the actual required load, and the N is the number of days in the testing dataset:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Predicited_i - Actual_i)^2}{N}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \qquad (3.1)$$

Normalized Root Mean Square Error or NRMSE is a function that normalize the RMSE value. The formula of NRMSE is shown in Equation 3.2 and the calculation process is done by divided the RMSE value over the mean of required load label and then multiplied with 100% (Johnston *et al*. 2018).

$$NRMSE = \frac{RMSE}{Mean\ Label} * 100\% \quad \ldots\ldots\ldots\ldots\ldots\ldots..\ldots\ldots\ldots\ldots. \quad (3.2)$$

3.4 Software Environments

In our experiments, we used two Python environments to predict the required load of 15 Iraqi provinces very well: Google Colab and Anaconda Python.

3.5 Chapter Summary

This chapter presented the dataset description that was collected from the Ministry of Electricity in Iraq for four years (2018, 2019, 2020, and 2021). This dataset relates to electrical distribution for 15 Iraqi provinces and contains nine features as follows: Date, Year, Province, Max Temperature, Min Temperature, Population, Fitted Load, Total Load, and Label (Required Load). Then, it includes visualizations of average fitted and average required loads in each province for the last four years in each month. 2021 has the highest average fitted and average required loads for most Iraqi provinces, and 2018 has the lowest average fitted and average required loads for most Iraqi provinces. In addition, this chapter presented the algorithms used and related Python libraries to predict the required load followed by the evaluating metrics: RMSE and NRMSE.

CHAPTER FOUR

EXPERIMENTAL RESULTS

4.1 Introduction

This chapter presents the experimental results for two algorithms.

4.2 Results

This section presents the experimental results of the required load prediction for the ARIMA and LSTM algorithms in terms of RMSE and NRMSE.

4.2.1 ARIMA Results

Figure 4.1 shows the prediction label by model in the blue line and the actual label in the orange line. In the prediction process, the mean value in the testing dataset is 3644.67 megawatts, and the Root Mean Squared Error (RMSE) is 591.88. While Figure 4.2 shows the forecast results produced by the ARIMA model for future month in 2022.

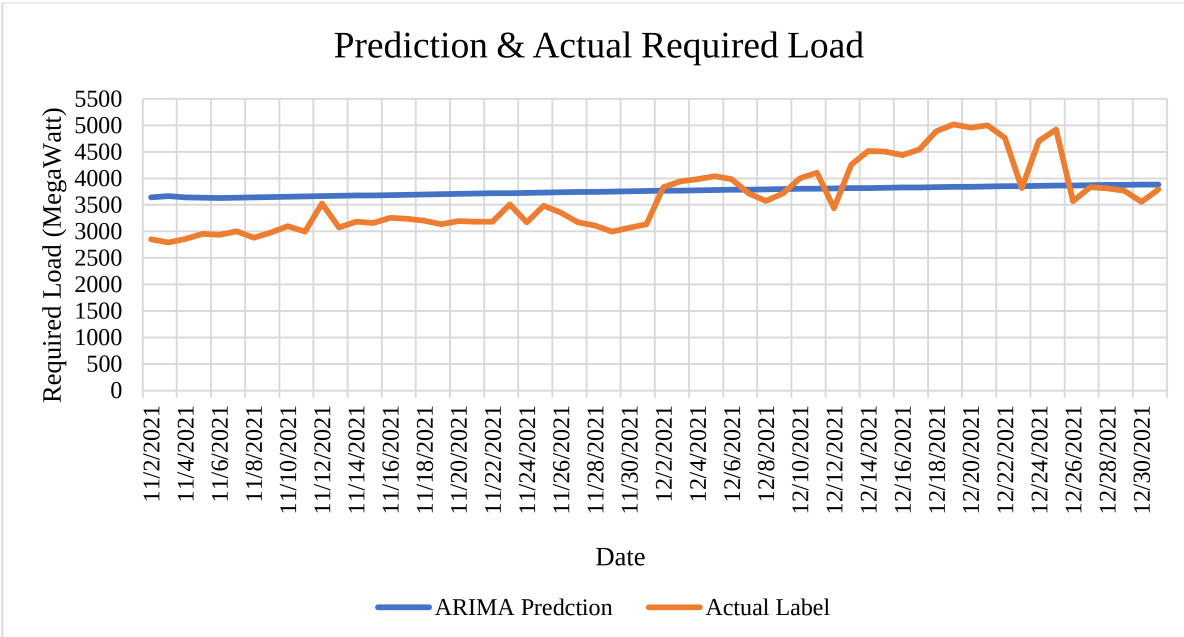

Figure 4.1: ARIMA Prediction and Actual Label of last Two Months in 2021.

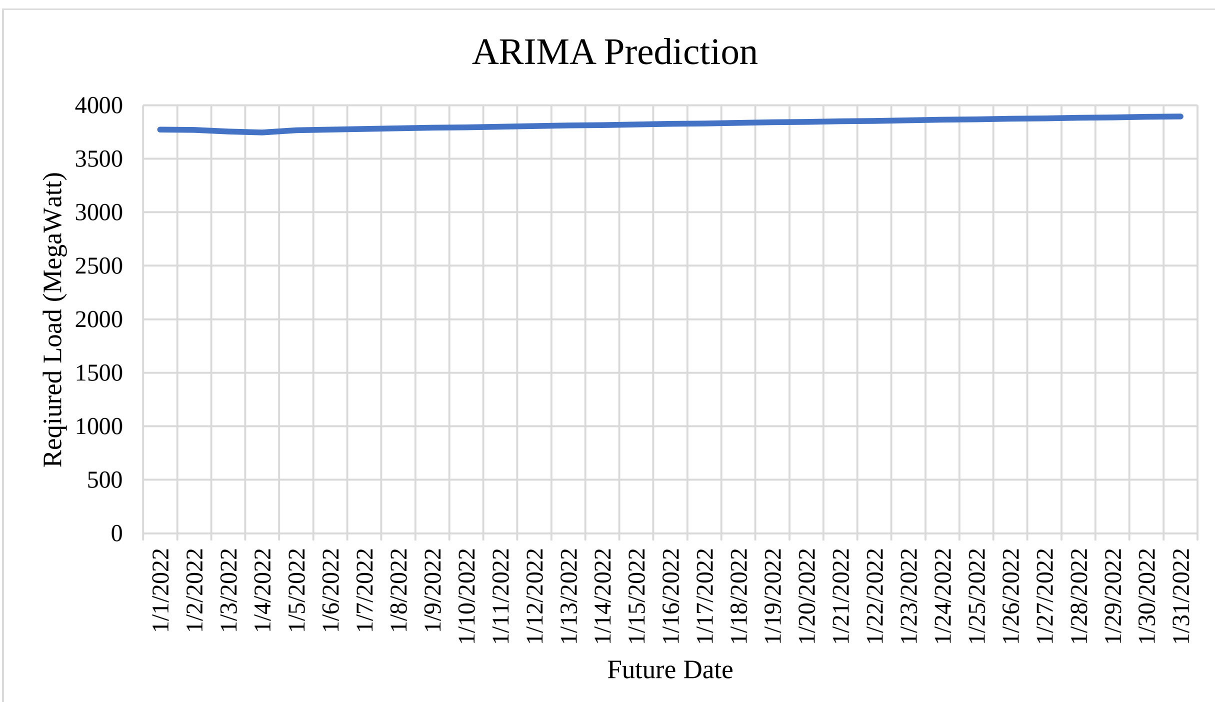

Figure 4.2: ARIMA Prediction of January Month in 2022.

The RMSE values with mean values and the NRMSE for the 15 Iraqi provinces are shown in Table 4.1. The results showed that the Al-Qadisiyyah and Baghdad provinces had the lowest RMSE and NRMSE values. While Basra Province has higher RMSE and NRMSE values than other provinces.

Table 4.1: Mean Value of testing Dataset and RMSE results for 15 Iraqi Provinces.

Province	**RMSE**	**Mean Label**	**NRMSE**
Al-Anbar	134	725	18%
Al-Qadisiyyah	85	535	16%
Babil	184	844	22%
Baghdad	592	4326	14%
Basra	870	2257	39%
Dhi Qar	283	997	28%
Diyala	190	792	24%
Karbala	161	750	21%
Kirkuk	251	871	29%
Maysan	160	631	25%
Muthanna	101	401	25%
Najaf	141	730	19%
Nineveh	448	1358	33%
Saladin	208	763	27%
Wasit	130	690	19%

4.2.2 LSTM Results

To determine the best values for the LSTM architecture, we first relied on base configuration and carried out some modifications on this configuration to know the best values. The base configuration is as following points:

- Number of layers is two: LSTM with 10 neurons and Dense with 1 neuron.
- Number of neurons is 10 neurons.
- Optimizer Function is adam

- Batch Size is 16.
- Early stopping is used.

The RMSE value after determined these values for each parameter of LSTM algorithm is 227. To improve the result, we should change values for each of parameter as following experiments:

- Changing the number of layers from two to one, three, and four layers.
- Changing the number of neurons from 10 to 5, 25, and 40.
- Changing the optimizer function from adam to nadam, SGD, and RMSprop.
- Changing the batch size from 16 to 4, 8, 32, 64, and 128. At each experiment, we fixed three values and change one of them to select the best values.

Figure 4.3 shows the different number of layers: 1, 2 3, and 4 with RMSE value for each of them. The best number of layers that gave the less RMSE is 2, which is 227.

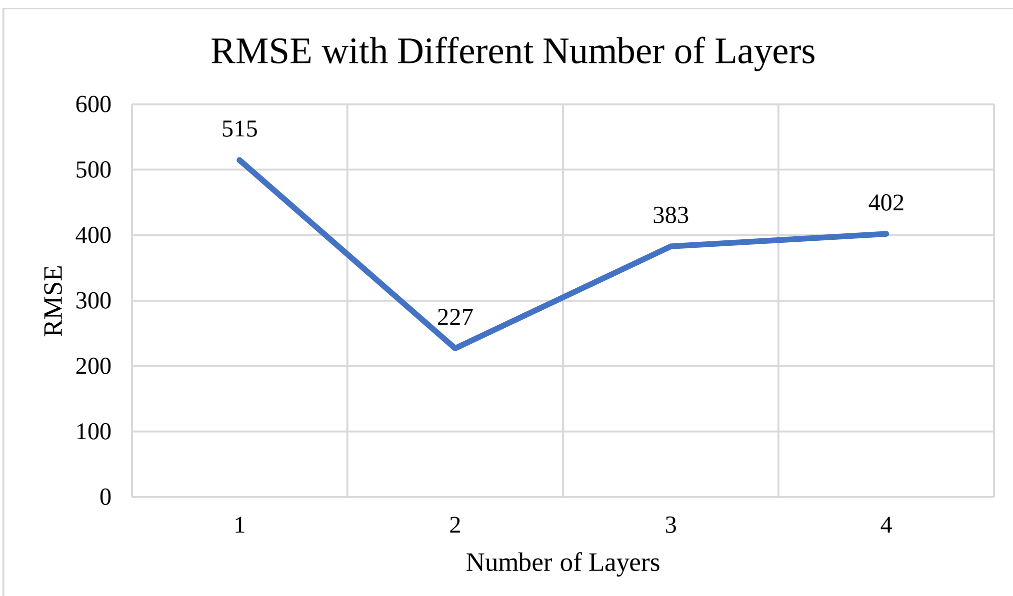

Figure 4.3: RMSE value with different architectures.

Figure 4.4 shows the different number of neurons: 5, 10, 25, and 40 with RMSE value for each of them after the two layers were selected as LSTM architecture. The best number of neurons that gave the less RMSE is 5 neurons that is 193.

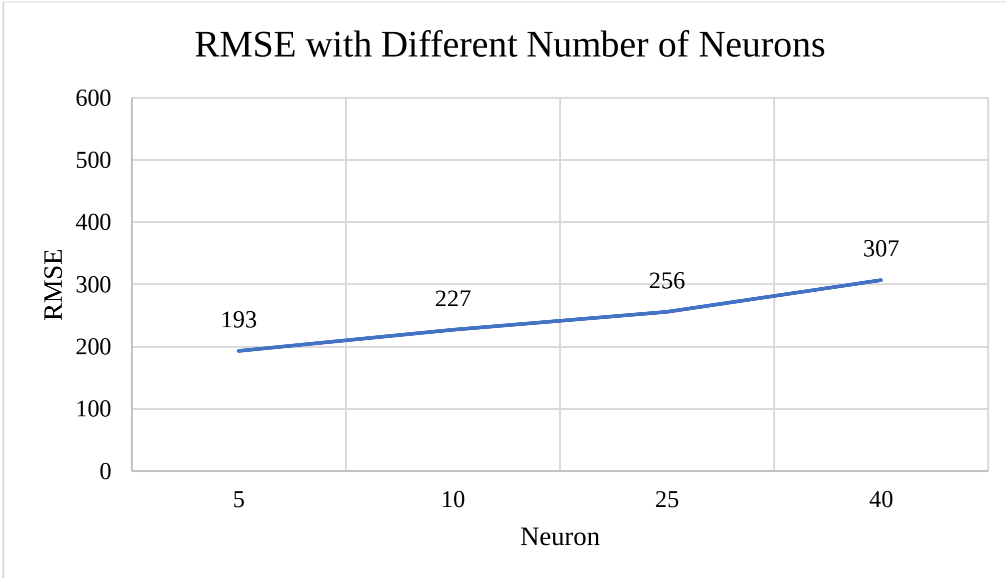

Figure 4.4: RMSE value with different number of Neurons.

Figure 4.5 shows the different optimizer functions: adam, nadam, SGD, and RMSprop with RMSE value for each of them after two layers and 5 neurons were selected. The best optimizer function that gave the less RMSE is adam with 193.

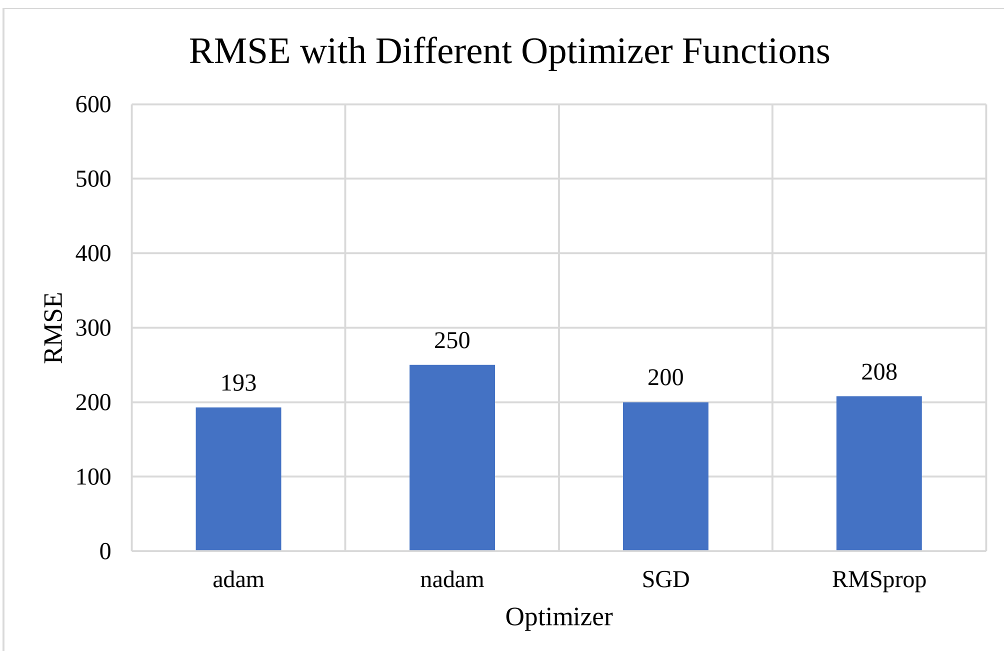

Figure 4.5: RMSE value with different optimizer functions.

Finally, Figure 4.6 shows the different batch sizes: 4, 8, 16, 32, 64, and 128 with RMSE value for each of them after two layers, 5 neurons, and adam as an optimizer function were selected. The best batch size that gave the less RMSE is 128 with 191.

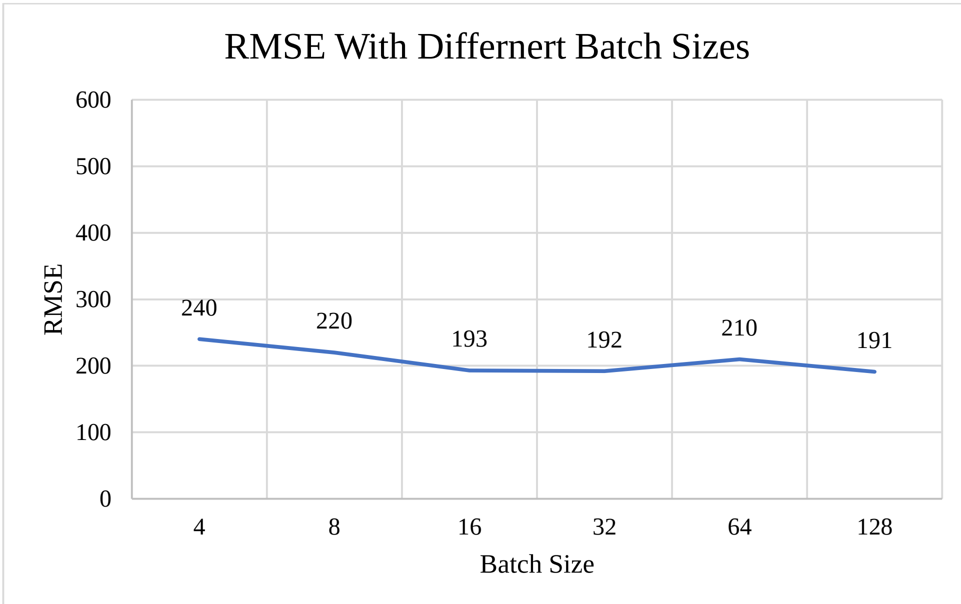

Figure 4.6: Figure 4.6: RMSE value with different Batch Sizes.

Therefore, the best values for the batch size are 128, optimizer function is adam, number of layers is two layers as LSTM architecture and number of neurons is 5 neurons. After that, the best RMSE value is 191 when we select the aforementioned values.

Also, we carried out some experiments to know what are the best sequence that gave the less RMSE, which means that the number of previous days that need to predict the required load for number of future days. Table 4.2 and Figures 4.7 and 4.8 show the RMSE results of five cases when we predict the required load. These five cases are:

- In case number 1, the future day is fixed that is 1, but the previous days are varied: 7, 14, 21, and 28. Therefore, the RMSE results of these days are 191, 163, 192, and 246. The best RMSE is 163 when we predict one day for the required loads based on the previous 14 days.

- In case number 2, the future days are varied, which are 7, 14, 21, 28, and the previous days is 30. Therefore, the RMSE results of these days are: 276, 364, 385, and 475.

- In case number 3, the future days are varied, which are 7, 14, 21, 28, and the previous days is 28. Therefore, the RMSE results of these days are: 336, 356, 376, and 400. We can conclude that using 30 days is best to predict 7 days ahead and using 28 days is the best to predict 14, 21, and 28 days ahead.

Table 4.2: Different sequences in prediction Required Load.

Case	Previous day	Future Day	RMSE	NRMSE
Case #1	7	1	191	4%
	14	1	163	4%
	21	1	192	4%
	28	1	246	6%
Case #2	30	7	297	7%
	30	14	364	8%
	30	21	385	9%
	30	28	475	11%
Case #3	28	7	336	8%
	28	14	356	8%
	28	21	376	9%
	28	28	400	9%

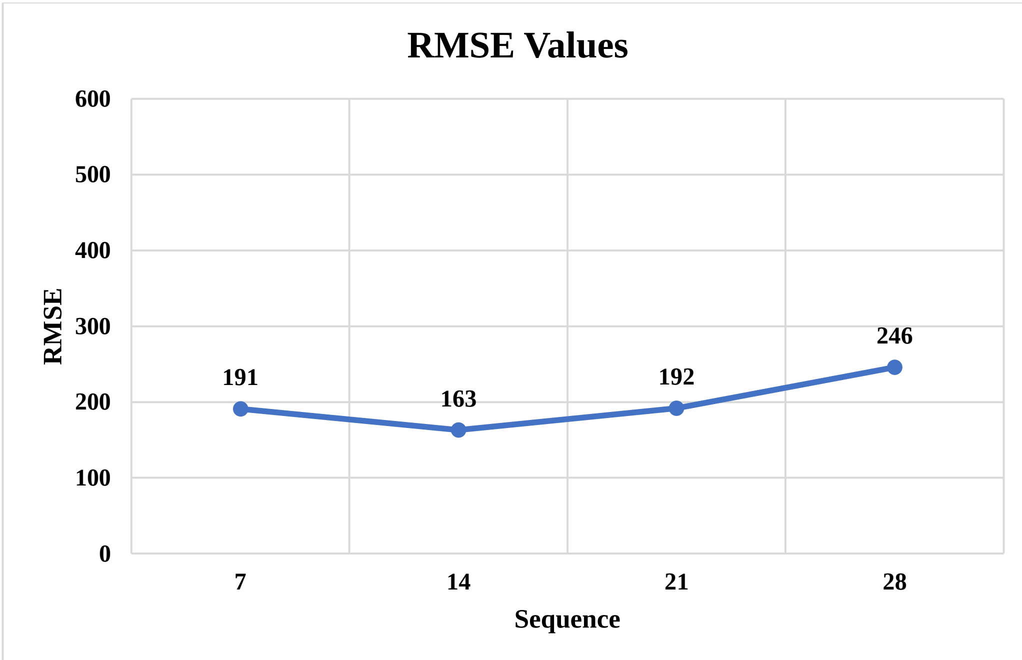

Figure 4.7: RMSE Values with 1 as Future Day.

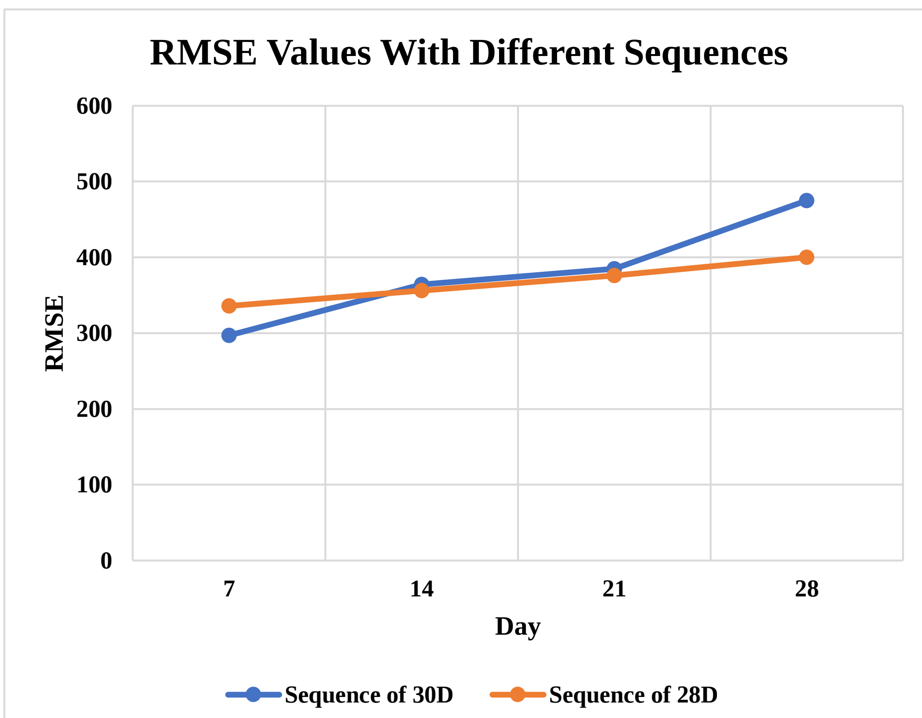

Figure 4.8: RMSE Values Sequence with 30 and 28 previous days.

Figure 4.9 shows the prediction label by model in the orange line and the actual label in the blue line when we choose one day as a number of days that predicted in prediction process based on previous 14 days with the best selected values.

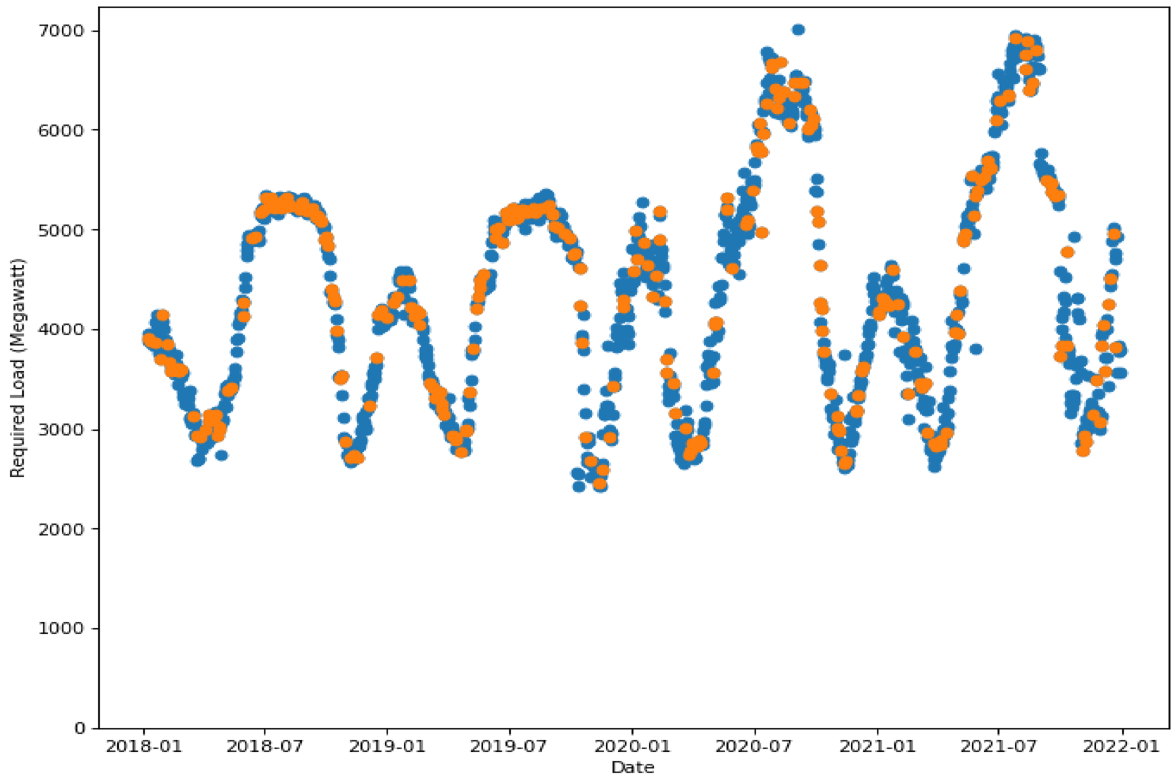

Figure 4.9: Visualization of All Dataset & Testing Dataset.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

The previous chapter illustrated major experimental results obtained from the simulations via Python software package using a numerical code. Chapter five shows major conclusions and describes critical recommendations of this work to improve the research outputs regarding the load analysis and forecasting of the Iraqi electrical distribution grid using a machine learning approach. Furthermore, the research limitations associated with the work in this study are presented.

5.2 Conclusions

This work is carried out to investigate and examine the critical contributions and beneficial role of machine learning algorithms in predicting the electrical load distribution with higher accuracy. Datasets are prepared and collected from the Ministry of Electricity in Iraq for four years (2018, 2019, 2020, and 2021). These datasets are associated with electricity distribution in several fifteen Iraqi provinces. Further, two algorithms, including ARIMA and LSTM, are developed to forecast the required load for an extended period in the future. The ARIMA algorithm is used to forecast the required load for one month as a case study in 2022. While the LSTM algorithm builds a deep learning model to predict the required load through a specific time in four years, relying on past and future days. The major work findings can be summarized in the following paragraphs:

1- The LSTM gave a lower RMSE value of 163 and 4% as a NRMSE value when we selected the 14 past days to predict one future day with the following parameters: batch size is 128, optimizer function is adam, number of layers is two layers as in LSTM architecture, early stopping is used, and number of neurons is five neurons.

2- The ARIMA and LSTM algorithms in the forecasting of load process of one month of 2022.

5.3 Recommendation

Based on the findings obtained from this work, this thesis proposes a set of crucial aspects and critical suggestions that can improve the research outputs attained from the numerical analysis. These recommendations include the following:

I- Using the ARIMA and LSTM algorithms in the forecasting of electrical load distribution of Iraqi network depending on a larger future time scale than the time investigated in this study.

II- Using other intelligent algorithms that depend on machine learning and deep learning to forecast load distribution of the Iraqi grid and compare their accuracy with the algorithms used in this study.

5.4 Research Limitations

Despite the satisfying and successful results obtained from the numerical simulations in this study, there were some barriers and key obstacles that limit the wide implementation of this work, including: The size of the dataset is small, which means that the period of time is short. The LSTM and ARIMA need a large dataset to get the lower RMSE values.

REFERENCES

Ahmed, W. K., Salam, A. Q., Mahdiy, M. T., & Chaichan, M. T. (2020). Environmental Impact of Using Generators in The University of Technology in Baghdad, Iraq. **Journal of Thermal Engineering**, 6(6), 272-281.

Almazrouee, A. I., Almeshal, A. M., Almutairi, A. S., Alenezi, M. R., & Alhajeri, S. N. (2020). Long-term forecasting of electrical loads in Kuwait using prophet and holt–winters models. **Applied Sciences**, 10(16), 5627.

Abdulrahman, M. L., Ibrahim, K. M., Gital, A. Y., Zambuk, F. U., Ja'afaru, B., Yakubu, Z. I., & Ibrahim, A. (2021). A Review on Deep Learning with Focus on Deep Recurrent Neural Network for Electricity Forecasting in Residential Building. **Procedia Computer Science**, 193, 141-154.

Almalaq, A., & Edwards, G. (2017). A review of deep learning methods applied on load forecasting. **In 2017 16th IEEE international conference on machine learning and applications (ICMLA)** (pp. 511-516). IEEE.

Averbukh, M. A., Zhilin, E. V., & Abdulwahhab, M. W. (2019). Problems of energy supply of the main consumers of distribu-tive networks of Iraq. In IOP Conference Series: Materials Science and Engineering (Vol. 552, No. 1, p. 012010). **IOP Publishing**.

Al-Mahdawi, E. (2021). An outlook on deployment the storage energy technologies in iraq. In IOP Conference Series: Earth and Environmental Science (Vol. 779, No. 1, p. 012020). **IOP Publishing**.

Almalaq, A., & Edwards, G. (2017, December). A review of deep learning methods applied on load forecasting. **In 2017 16th IEEE international conference on machine learning and applications (ICMLA)** (pp. 511-516). IEEE.

Brownlee, J. (2016). Deep learning with Python: develop deep learning models on Theano and TensorFlow using Keras. **Machine Learning Mastery**.

Chitalia, G., Pipattanasomporn, M., Garg, V., and Rahman, S. (2020). Robust short-term electrical load forecasting framework for commercial buildings using deep recurrent neural networks. **Applied Energy**, *278, 115410.* doi:10.1016/j.apenergy.2020.115410

Ediger, V. Ş., & Akar, S. (2007). ARIMA forecasting of primary energy demand by fuel in Turkey. **Energy policy**, 35(3), 1701-1708.

Gers, F. A., Eck, D., & Schmidhuber, J. (2002). Applying LSTM to time series predictable through time-window approaches. In Neural Nets WIRN Vietri-01 (pp. 193-200). **Springer**, London.

Ghany, K. K. A., Zawbaa, H. M., & Sabri, H. M. (2021). COVID-19 prediction using LSTM algorithm: GCC case study. **Informatics in Medicine Unlocked**, 23, 100566.

Hao, J., & Ho, T. K. (2019). Machine learning made easy: a review of scikit-learn package in python programming language. **Journal of Educational and Behavioral Statistics**, 44(3), 348-361.

Ibrahim, B., & Rabelo, L. (2021). A Deep Learning Approach for Peak Load Forecasting: A Case Study on Panama. **Energies**, 14(11), 3039.

Istepanian, H. H. (2014). Iraq's electricity crisis. **The Electricity Journal**, 27(4), 51-69.

IOM, (2020). International Organization for Migration. Visited online through the website: http://www.iom-iraq.net/iraq-population.html

Iraqi Ministry of electricity. (2010). Iraq electricity masterplan - final report volume 2: load forecasting.

Jiang, Y., Huang, Q., Zhang, K., Lin, Z., Zhang, T., Hu, X., & Lin, Z. (2021). Medium-long term load forecasting method considering industry correlation for power management. **Energy Reports**, 7, 1231-1238.

Jung, S. M., Park, S., Jung, S. W., & Hwang, E. (2020). Monthly electric load forecasting using transfer learning for smart cities. **Sustainability**, 12(16), 6364.

Johnston, B., & Chazal, P. D. (2018). A review of image-based automatic facial landmark identification techniques. EURASIP **Journal on Image and Video Processing**, 2018(1), 1-23.

Kuo, P. H., & Huang, C. J. (2018). A high precision artificial neural networks model for short-term energy load forecasting. **Energies**, 11(1), 213.

Moon, J., Park, S., Rho, S., & Hwang, E. (2019). A comparative analysis of artificial neural network architectures for building energy consumption forecasting. International **Journal of Distributed Sensor Networks**, 15(9), 1550147719877616.

Mohammad, F., Lee, K. B., & Kim, Y. C. (2018). Short-term load forecasting using deep neural networks. arXiv preprint arXiv:1811.03242.

Mills R., & Salman M. (2020). Powering Iraq: Challenges Facing the Electricity Sector in Iraq. Friedrich Ebert Foundation (Friedrich-Ebert-Stiftung), **Al-Bayan Center for Planning and Studies**.

Ministry of electricity in Republic of Iraq. (2020). Datasheet of electric power demands in the Iraqi governorate over several recent years.

Nalcaci, G., Özmen, A., & Weber, G. W. (2019). Long-term load forecasting: models based on MARS, ANN and LR methods. **Central European Journal of Operations Research**, 27(4), 1033-1049.

Nyandwi, A., & Kumar, D. (2020). Neural Network Approach to Short and Long Term Load Forecasting Using Weather Conditioning. **In 2020 International Conference on Electrical and Electronics Engineering (ICE3)** (pp. 258-263). IEEE.

Nepal, B., Yamaha, M., Yokoe, A., & Yamaji, T. (2020). Electricity load forecasting using clustering and ARIMA model for energy management in buildings. **Japan Architectural Review**, 3(1), 62-76.

Ozcan, A., Catal, C., & Kasif, A. (2021). Energy Load Forecasting Using a Dual-Stage Attention-Based Recurrent Neural Network. **Sensors**, 21(21), 7115.

Pang, B., Nijkamp, E., & Wu, Y. N. (2020). Deep learning with tensorflow: A review. **Journal of Educational and Behavioral Statistics**, 45(2), 227-248.

Runge, J., & Zmeureanu, R. (2021). A review of deep learning techniques for forecasting energy use in buildings. **Energies**, 14(3), 608.

Vanting, N. B., Ma, Z., & Jørgensen, B. N. (2021). A scoping review of deep neural networks for electric load forecasting. **Energy Informatics**, 4(2), 1-13.

Wen, L., Zhou, K., & Yang, S. (2020b). Load demand forecasting of residential buildings using a deep learning model. **Electric Power Systems Research**, 179, 106073.

Wen, Z., Xie, L., Fan, Q., & Feng, H. (2020a). Long term electric load forecasting based on TS-type recurrent fuzzy neural network model. **Electric Power Systems Research**, 179, 106106

Yasen, M. H. (2016). Analyse the Problems of the Iraqi Power System. **Advances in Energy Engineering**, 4.

ABSTRACT IN ARABIC

استخدام تقنيات التعلم الآلي لإجراء عمليات التحليل والتنبؤ بالأحمال الكهربائية في شبكة التوزيع الكهربائية العراقية

إعداد

نور غازي بديوي

المشرف

الاستاذ الدكتور غيث علي عبندة

الملخص

تهدف هذه الرسالة لدراسة الأثر الفاعل والدور المهم لخوارزميات التعلم الآلي في التنبؤ بتوزيع الحمل الكهربائي لشبكة الكهرباء العراقية بدقة عالية. وقد تم جمع مجموعة بيانات وإعدادها من وزارة الكهرباء في العراق لمدة أربع سنوات (2018 و2019 و2020 و2021)، وهي بيانات لشبكة توزيع الكهرباء في خمس عشرة محافظة عراقية. وتم استقصاء منهاجين هما ARIMA و LSTM للتنبؤ بالحمل المطلوب لفترة محددة في المستقبل. تُستخدم خوارزمية ARIMA للتنبؤ بالحمل المطلوب لمدة شهر كدراسة حالة في عام 2022. بينما تبني منهاج LSTM نموذجًا للتعلم العميق للتنبؤ بالحمل المطلوب خلال فترة زمنية محددة اعتماداً على الأيام الماضية والمستقبلية. وقد كشفت نتائج الدراسة أن LSTM أعطت قيمة RMSE أقل من 163 و 4٪ كقيمة NRMSE عندما تم اختيار الأربعة عشر يومًا الماضية للتنبؤ بيوم واحد في المستقبل باستخدام المعلمات التالية: حجم الدُفعة هو 28، نوع المُحسِّن هو adam، عدد الطبقات اثنين من خلايا LSTM، تم استخدام التوقف المبكر، عدد الخلايا العصبية هو خمسة خلايا عصبية للطبقة. وقد أشارت النتائج إلى أن خوارزميات ARIMA و LSTM حققت دقة عالية في التنبؤ بعملية التحميل لمدة شهر واحد من عام 2022.