REPUBLIC OF TÜRKİYE

ALTINBAŞ UNIVERSITY

Institute of Graduate Studies

Information Technology

ELECTRIC DEVICE CLASSIFICATION BY POWER CONSUMPTION FOR OPTIMAL POWER MANAGEMENT IN IOT NETWORK

Zohad Abduljawad M. TAWFWWQ

Master's Thesis

Supervisor

Asst. Prof. Dr. Hakan KOYUNCU

Istanbul, 2023

ELECTRIC DEVICE CLASSIFICATION BY POWER ONSUMPTION FOR OPTIMAL POWER MANAGEMENT IN IOT NETWORK

Zohad Abduljawad M. TAWFWWQ

Information Technology

Master's Thesis

ALTINBAŞ UNIVERSITY

2023

The thesis titled ELECTRIC DEVICE CLASSIFICATION BY POWER CONSUMPTION FOR OPTIMAL POWER MANAGEMENT IN IOT NETWORK prepared by ZOHAD ABDULJAWAD M. TAWFWWQ and submitted on 14/04/2023 has been **accepted unanimously** for the degree of Master of Science in Information Technology.

____
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 IBRAHIM	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.

Zohad Abduljawad M. TAWFWWQ

Signature

DEDICATION

I extend my sincere thanks and gratitude to the Republic of Turkey for its good hospitality and reception to us. I especially thank my Asst. Prof. Dr. Hakan KOYUNCU, who gave me great support and great guidance to show my thesis at the highest level and with this scientific value .. Thank you from the heart, my teacher ... As for my family, allow me to grant this Effort for the soul of the man whose name honored me, my father, and for the great woman who motivated me, strength and determination to reach what I am today My mom, and for the clusters of love, my brothers, thank you, fellow roadmates, colleagues, thank you, wonderful people .. Thank you to everyone who helped me with any information or advice. Thank you. you from the heart.

ABSTRACT

OPTIMIZATION SMART CLASSROOM USING FEATURE SELECTION IN ACO ALGORITHM

TAWFWWQ, Zohad Abduljawad M.

M.Sc., Information Technologies, Altınbaş University,

Supervisor: Asst. Prof. Dr. Hakan KOYUNCU

Date: April / 2023

Pages: 101

Power factor is the ratio of (real) power to total (apparent) power (kVA) consumed by an AC component. It is a measure of how efficiently electrical energy is converted into useful work output. The ideal power factor is 1. Anything less than one means that extra power is required to accomplish the actual task at hand. Household loads vary between inductive loads and capacitive loads. Unbalanced loads lead to a difference in the power factor, and this in turn leads to losses in power. The main objective of this thesis was load balancing for the purpose of eliminating power losses. The work of this thesis was summarized in four stages, the first was the prediction of the amount of the difference between the loads, the second was the prediction of the quality of the overpowering load that causes a deviation inA the power factor, and the third was the determination of the treatment method, and all of this was done through machine learning algorithms that are mentioned in the third and fourth chapters of this work. . The fourth step in this work was to organize a system that works on balancing the loads by alternating operation with a time difference by using the programmed logic controller. The performance of the PLC system, which in turn led to the harmonization of loads.

Keywords: Machine Learning, Power Factor, Load Balance, PLC, GBM.

TABLE OF CONTENTS

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

ABBREVIATIONS

AI	:	Artificial Intelligence
FSK	:	Frequency Shift Keying
IoT	:	Internet of Things
PAN	:	Personal Area Network
OFDM	:	Orthogonal Frequency Division Multiplexing
TDMA	:	Time Division Multiple Access

1. INTRODUCTION

1.1 INTRODUCTION

Electric energy today is one of the most important elements and needs of daily life that are indispensable. Therefore, it has become necessary to find ways to reduce its consumption [1].

And since any decrease in electrical energy leads to major problems worldwide in several respects, the most important of which are economic and commercial, most life facilities depend heavily on electrical energy, so it is very necessary that we have an effective management of electrical energy to prevent energy loss as well as reduce consumption [2]. We must not forget that electricity today has become one of the contributors to raising the standard of living of any society, and this in itself is a very important issue [3].

Technology today and the great development in it led to a great change in many fields. After information technology was at the forefront in terms of its clear large consumption of electrical energy for a long time [4][5], today it has proven to the whole world its ability to give positively in all fields, especially in the field of electric energy. It is considered one of the positive contributors in this field by using several methods and technologies to control and control energy consumption [6][7], and the most important of these technologies is the Internet of Things [8] This technology allowed devices to communicate with each other and work with high efficiency. This technology has entered many fields and many applications such as health, agriculture and electricity, and among its most important applications are smart homes, smart buildings, smart networks, etc. [9]. Recent years The Internet of Things is seen as an important factor in the development of technologies to support access to the 5G ecosystem [10] ,

It is also possible to define IoT as a group of devices connected to each other by smart networks through platforms and interfaces used for a number of transmission protocols such as TCP/IP protocol, etc. [9], and recent years have witnessed the connection of a wide range of different devices estimated at billions of devices, whether they are electronic devices or sense if it supports smart buildings.

The Internet of Things consists of sensors, communication devices, and systems that process data [8].

One of the most important applications of the Internet of Things technology is smart grids. Smart grids did not appear in vain, but appeared to provide unparalleled assistance in the field of energy, i.e. electrical grids. These grids have great independence and the ability to connect with high efficiency [11] ,This technology is fundamentally different from regular grids, as it It has the advantage of remote control as well as it has the feature of automation through two-way communication and it also enjoys higher reliability and very high efficiency as it has the ability to answer many requests by integrating distribution and generation sources at the same time [12].

Machine learning will have an important role in this thesis, as machine learning is a type of artificial intelligence that helps machines learn and simulate human intelligence. Machine learning has its entry into multiple fields, as it has a large number of fields. The most important characteristic of machine learning is that it has two types of models: supervised machine learning and unsupervised machine learning [11]. For the convenience of the user, it is important to provide the best solution to the issue of load problems, reduce bill payment, as well as support and provide convenience for the user [13].

The thesis seeks to classify the electrical loads of a city on the basis of the losses in the loads, and also classify the loads on the basis of the type of load, and then predict the type of electrical load at the level of one house.

1.2 PROBLEM STATEMENT

The urban and technical development and the diversity of devices and services led to the emergence of gaps in providing the amount of electrical energy sufficient to ensure the complete well-being of the population, as this development was accompanied by the innovation of many devices and equipment serving human well-being that require a greater amount of electrical energy, and as the provision of a greater number of engines and generators Electric energy undoubtedly leads to an increase in the problems of pollution and global warming, in addition to wasting the resources of countries.

From this logic, it was necessary to think about reducing the consumption of electrical energy. For this reason, this study examined how to evaluate the consumption of electrical energy by devices based on the Internet of Things (IOT), if it studied the structure of the Internet of Things, its mechanism of work, and how to control through it the consumption of electrical energy by devices to achieve optimal consumption. Through the foregoing, the research problem can be summarized in the following points

a. Is the currently available electrical energy?

b. Sufficient to equip homes and factories in their current condition?

c. Is it possible to provide the amount of electrical energy associated with the rapid urban and technical development?

d. How can devices be divided according to their energy consumption based on the Internet of Things?

e. How can the Internet of Things be used to control the consumption of daily electrical energy by devices?

f. How can artificial intelligence be employed to benefit from it in the automatic control of Internet of Things devices and thus control the expenditures of electrical energy automatically?

1.3 MOTIVATION

The main motive behind the work of this project was to help find advanced solutions to reduce the consumption of electrical energy and provide more energy in the midst of the technological revolution that swept our daily lives. Societies today are in great need to provide ways that contribute to reducing energy consumption. From this perspective, we need therefore to provide Intelligent systems help solve this problem.

This prompted researchers and scientists to search extensively to find the optimal solution, as they developed many smart systems that support smart networks for the purpose of saving more energy and predicting the best use of devices. So the main objective of the project is to create a smart system to classify electrical appliances according to Optimal consumption by relying on the Internet of Things, smart networks, artificial intelligence, and machine learning to reduce energy consumption.

1.4 THESIS OBJECTIVES

In light of the global crisis in the consumption of electrical energy, we need to develop more effective ways to reduce consumption and save electrical energy. The research aims to present a new methodology for reducing energy consumption by classifying electrical appliances according to their optimal energy consumption in effective ways that we have not noticed before, which leads to improved efficiency Network operability and energy efficiency, in addition to reducing costs, providing energy stability, awareness of residents of optimal usage times and methods, and proving the success of the Internet of Things as the best method in helping to manage the electrical network in a wonderful way, and contributing to reducing electrical energy consumption and excelling in dealing with big data and a huge number of sensors connected to each other. . while highlighting the problems as well as challenges open to these approaches and ultimately providing a basis for future work and further analysis.

1.5 CONTRIBUTION

The energy system is almost impossible without the effective use of energy and awareness in use and reducing consumption. The actual contribution of this research lies in proposing a solution to reduce energy consumption by classifying electrical appliances based on the Internet of Things and artificial intelligence, through the use of machine learning algorithms to build an intelligent system. Where IoT technologies will be analyzed in terms of energy efficiency and consumption aspects, and a classification of the measures used to evaluate energy efficiency in IoT networks will be provided by building a complete system for a consumption network connected to each other via the Internet of Things, so that the system records all energy consumption for each device, and records it on Cloud database. Where the algorithm is applied to schedule the pattern of energy consumption by means of the obtained database, and then use the prediction scheme to connect to the best consumption.

1.6 ORGANIZATION

This study consisted of five chapters, the first chapter dealt with a general introduction on the subject of the research and reviewed its problem and the most important goals and contributions, then the second chapter began with a review of previous studies on the topic of the thesis, followed by the third chapter, which talked about the Internet of Things, the power factor, electrical loads, machine learning and its algorithms, either The fourth chapter dealt with the collection of data, the implementation of the system and its results, the programming of the logical controller and the review of the results, while the fifth chapter concluded with a number of conclusions reached by the researcher and the most important recommendations that would develop the work

2. LITERATURE REVIEW

2.1 INTRODUCTION

The world of electricity began since ancient times, and developments began to occur until it reached what it is today. It is today completely different in all respects, whether in terms of power generation or energy distribution, as technology contributed to positive intervention in all aspects of electrical energy by achieving a balance in energy and the environment, but development Which invaded life in all its fields and the needs of the population for energy and the fear of depletion of fossil fuels, which is a major source of energy production, all of this must be found a solution to reduce energy consumption. To control and reduce human consumption of energy [14][15].

The huge waste that affected energy recently, a solution must be found to reduce waste and reduce consumption, as many technologies were proposed that help solve the problem. The main goal of these technologies was to maintain a balance between the user's needs and his comfort each This prompted researchers and scientists to present studies that dealt with the problem and how to solve it, depending on different algorithms, and with the great progress of technology, the biggest contribution was in providing The Internet solutions are things that are characterized by sensing techniques and their ability to connect anything and anywhere, as in Internet networks. Things are distinguished by their ability to succeed in real environments through predicting energy consumption. Forecasting energy consumption has become one of the most important problems faced by generation companies due to the need and large demand for energy as a result of expansion. The large population all over the world, where scientists believe that if consumption is not controlled, we will face a loss and shortage of energy in the next few years. If all of this puts us before one of two options to compensate for the energy shortage, either provide more energy and this requires a lot of material capabilities and time, or reduce consumption energy by taking precautionary measures, as there are possibilities to improve and reduce consumption through predicting consumption, and it is also possible to improve consumption in smart cities, homes, and factories, as it will be eye-catching.

Numerous techniques based on algorithms to improve and reduce a for consumption, such as the earthworm algorithm and the bat algorithm ... etc., but it is necessary to take into account the geographical area of the systems, the environmental difference, and the conditions. Despite all the studies that researchers have presented regarding energy consumption and its improvement, they are still looking to this day for more intelligent and accurate systems that help end the gaps that previous studies were exposed to [16].

2.2 RELATED WORKS

In 2016 Muralitharan, et.al, discusses the implementation of a system and reduces the cost in relation to the electricity bill by balancing the loads through a mechanism that reduces the delay in the implementation of electrical appliances by using the Load Balancing Algorithm for Electrical Appliances, where the proposed algorithm will give an idea about the load management control mechanism in the house, as it also makes a proposal to the service provider To manage all load needs during peak hours. The researcher obtained a cost result using this improvement method estimated at 28.67 compared to the normal average cost without improvement. The cost was estimated at 33.09. As for the waiting time delay (Davg), it will decrease by approximately 10 to 15 Using MOEA. If you help the user to use energy within the specified limit to avoid paying additional costs, this will also help in using more energy at a lower cost. If this method is good for reducing costs, besides reducing the delay time for implementing devices [17].

In 2016 Grigoras, et.al, discussed a study on a phase-dependent shunt to balance the phase of the load (lv) low voltage in a distribution feeder by switching the side branches. The researcher applied this proposed method to an unbalanced (lv) distribution feeder by using information from smart meters. The results of this study I confirmed to reduce the imbalance to approx. (from 1.23 r.u. to 1.02 r.u.) and reducing the losses that occurred in a week to approximately 2.04%, from 6.43% to 4.39% [18].

In 2017 Wasoontarajaroen.et.al., presented an advanced device for IOT. This device measures power, current, voltage, and the energy of a four-phase line inside a laboratory building. The device is damaged from the control unit, the sensor unit, and the Wi-Fi unit. The device is originally one of the low-cost devices. This device was chosen through the data that was He collected it for a week and found the result that the device works well in energy management through that data [19].

In 2017 ALRikabi.et.al., presented a study on a system that manages energy using Arduino mega as a control unit, and the researcher used SIM800L GSM module that communicates in case of dangerous situations and Relay 8 CH electronic switch for devices, where the system controls and manages energy consumption within Iraqi government buildings based on the Internet of Things. The system gives a signal to start the work of the main electrical appliances to work at 7 in the morning and stop working with them at 2 in the afternoon, but when any device remains working from 8 in the evening and before 7 in the morning, the system will give a signal indicating the existence of a danger to the manager. This system has proven to reduce consumption to 50 percent and reduce The risks are also reduced by 50 percent [20].

In 2017 Chen, et.al, dealt with a new study of cloud load balancing by comparing it with previous studies. This proposal provided the possibility of implementing the CLB model with a physical and virtual web load balancing architecture and algorithms. Experiments conducted on physical and virtual servers that support cloud load balancing (clb). The results have proven that the cloud server of any form depends on the proposed architecture. The study balances when users log in at the same time [21].

In 2017 Zheng, et.al. presented a study of a hybrid neural network model with SD and EMD selection Methods for STLF The purpose of sd selection is to determine the feature weights where by using Xgboost and calculate the distance between the specified day and the day that depends on the different measured attributes with different weights and then after all that k SDs are collected or combined into a group One as input data for subsequent prediction based on Xgboos distance. If emd finally it will show the main sd load shapes at high and low frequencies. Discrete neural networks (LSTM) are used to predict future values in approximated low-frequency and high-frequency time series. The method has been compared with LSTM, SD-LSTM, EMD-LSTM, ARIMA, BPNN and SVR models in real-load dependent data obtained from NE-ISO to perform pre-load prediction. One day and one week The results of the comparisons showed that the proposed Xgboost-k-mean method can effectively integrate SDs within one group. Moreover, the EMD-LSTM model has a good ability to accurately predict the nonlinear electrical load time series over a long horizon. The above means that the proposed SD-EMD-LSTM framework could be an alternative approach to STLF[22].

In 2017 Li,et.al., discussed the EMS management system that gives a periodic resemblance to the characteristics of changes in energy loads at different times, according to the data and their dates for the same time load for a previous day, the same load for the last week, an average load for the last 24 hours, meteorological data, and all information about working days and holidays, where an enhancement model for severe scaling was built xgboost With mixing information, the algorithm can predict the direction of a load 24 hours before one day. The results will appear through the prediction of the 24-hour curve of the current network, as this algorithm has proven its advantages compared to KNN baysian in terms of accuracy and speed [23].

In 2018 Motlagh.et.al., presented a study on a control system for modern and old buildings and providing comfort to the residents, based on the Internet of Things, where he presented a test platform for the Internet of Things in which he used sensors to monitor lighting and compared a traditional energy system with a smart energy system, and the smart system proved its effectiveness in old buildings in terms of reducing consumption and providing comfort to its residents [24].

In 2018 Sicchar, et.al., presented a phase load balancing system based on hierarchical petri networks (PNs) that help verify the characteristics of the system. The integrated algorithms let lbc determine the load imbalance in several stages and improve the single-stage consumer test that depends on the level of load imbalance. The performance results of this lbc system were compared With another load, it was less than them in terms of the value of the unbalanced load, about 0.6 watts. LBC was compared and a second test was obtained, so the result of the lowest average load imbalance was 0.1 kW. All this means that the system is efficient [25]. .

In 2018 Lopez-Martin, et.al., dealt with a proposal, which is to modify the PFC controllers by adapting the operation depending on the file that contains the T HDV measurement. This made the PFCs operate either in a current mode with low harmonic distortion T HDI or in a traditional resistor simulated mode as well when the distribution feeders contribute to the regulation of T HDV and P F. To confirm this concept, the modification was applied to a non-sensitive control unit (NLC) for a single-phase Boost rectifier. The experimental results showed that the PFC operates in network-connected mode, CCM mode, and a different mode T HDV [26].

In 2018 Moon, et.al, proposed a hybrid model for predicting short-term loads for higher education institutions such as universities and institutes, using random forests and a multi-layered perceptron to build the prediction model. First, we collected education buildings as an academic group and collected load data over a period of 6 years. These data were collected and then divided. Between training and test groups in terms of training, in which we classified the load data by means of a decision tree according to the similarity of the load patterns. The random forest formations and multi-layer realization were studied, and their prediction performance values were evaluated using the validation group to determine the best and optimal model. Based on this, a Hayden model was built to predict the electrical load By selecting models with better prediction over a serial time series. Finally, using the test, the prediction performance of the daily electrical load was compared for the hybrid model and another common model, as the results showed that the hybrid models outperformed the common models. Also, the effectiveness of LSTM was shown in reflecting the electrical load depending on the day of the week. The decision tree proved its effectiveness in classifying time series through similarity. Using these two models for prediction can complement their weaknesses [27].

In 2019 Mylonas.et.al., discussed a way to save energy using specific methods and steps to be followed systematically based on data taken from the infrastructure of schools, i.e. using a strategy directed towards the energies saved, and the tangible results were that these strategies help in saving energies and environmental friendliness. One of the results of the examples was approximately 15-20% of thrift [28].

In 2019 Gupta.et.al., proposed a control and monitoring system for electrical devices that save energy based on the Internet of Things. The IEEE 802.1 technology was used, and the system proved that it helps reduce energy consumption and increases efficiency, as it is an efficient system for controlling street lighting lamps that need a lot of energy [29].

In 2019 Soykan.et.al., presented a study on predicting loads in the smart grid and a study on the privacy and guarantee of those loads and providing the possibility of a change in the price of electricity. The researcher used the LSTM algorithm and Tensor flow techniques for privacy. The study showed that data privacy can be achieved to varying degrees, in addition to noting a reasonable deterioration in prediction [30].

In 2019 Leninpugalhanthi.et.al., dealt with a study on theft of electric power and its cause in the loss of a huge amount of energy and economic losses. In India, the loss of electricity thefts amounted to approximately 16.2 billion dollars [31].

In 2019 Grigoras, et.al, discusses a method in which Tarsheed proposes to allocate an example for consumers on phases in order to balance the phase load of the electric distribution networks. This method relies on information about the users' electricity consumption obtained from smart meters. This method has proven effective by using an unbalanced real low-voltage distribution network. In this case . The proposed system may be economically feasible in unbalanced networks at a very high level and with the same load, as the results showed a decrease in the initial imbalance from 1.287 to 1.003, as well as a decrease in energy from 20.25 to 7.04 kilowatt-hours, or by about 63.23%. When taking into account these results, it is possible To say that phase load balancing is considered the cheapest way to improve the energy efficiency of low voltage distribution networks [32].

In 2020 Muralidhara, et.al, introduced a smart energy meter to reduce the problem of consumption and measurement, in which he used Arduino uno. This meter made us get a lower cost of energy and control of the meter, as it helped in classifying energy and reducing consumption, as it gives the ability to monitor and record consumed energy data [33].

In 2020 Seyyedabbasi.et.al., presented a system for controlling the electricity consumption of smart homes. The researcher took advantage of the sensor nodes of the Internet of Things and used it for weather forecast sensors inside the house. This system would be one of the useful and effective systems in managing and monitoring energy consumption in smart homes [34].

In 2020 Veeramsetty, et.al., discussed the use of the neural network from machine learning to create a system that predicts load during the hours of the day by relying on the load data for the last 3 hours and downloading the data at the same hour during the last 4 days. The neural network topology has been developed to predict load based on the proposed approach. Using the back propagation algorithm based on Levenberg Marquardt to be able to train the proposed ANN model. The ANN model has been implemented and tested within the MATLAB environment. This method has been proven to predict the load more accurately [35].

In 2020 Cao, et.al., Discuss a study comparing 8 trained machine learning models with daily and weekly data from a hospital in Shanghai. It was found that the daily electrical load predictions of XGBoost, SVR, RF are the most accurate and individual models. In the MAPE run of the data test 9.64%, 9.81% and 10.67% sequence of These can be considered reliable intelligent forecasting models for energy managers/management to perform daily forecasting of future electrical load. By making a comparison between daily and weekly forecasting it is found that performance will improve by using more granular or detailed energy usage data. One possible explanation is that a detailed load profile can reduce com-mode uncertainty. In terms of feature significance, the outside temperature, air pressure and operating status of the central air conditioning system were found to be the top 3 contributing factors [36].

In 2021 Awan.et.al., used the machine learning algorithm in the management of smart buildings for smart cities, where he suggested the use of CDBR based on improving the particle swarm that enjoys its effectiveness by using smart cities, as this algorithm granted a smooth method and a specific pattern using energy, and the result was to reduce the peak to moderate [37].

In 2021 Xiaoyi,et.al., discussed the integration of energy with artificial intelligence with electricity networks through the use of the modda algorithm intended to send a multi-purpose distributor [38].

In 2022 Jasim, et.al, discussed a smart system that measures and monitors energy consumption based on the Internet of Things, where the system gives users the ability to monitor their energy consumption through smart phone applications, all by relying on the Internet of Things. In the end, the system has proven effective in reducing energy consumption and saving it, and thus this leads to saving money as well [39].

In 2022 Bhagwatkar.et.al., discussed a study on collecting data through smart sockets, where he used an Arduino chip and sensors to relay and measure the power voltage in electrical appliances, where he built a wireless network aimed at measuring the energy in electrical appliances in the house. It did not specify any time for the use and control of electrical appliances [40].

In 2022 Thapliyal, et.al., addressed an important topic: load balancing in the cloud, where load balancing is very necessary and important. Storage load balancing was not taken into account. The great development in the cloud called for competition over control and load control strategies. In this system, he proposed a new algorithm for load balancing called the bee colony. From synthetic experiments, we can say that this algorithm achieved good efficiency and maintained network load balancing within the cloud [41].

In 2022 Gamez Medina, et al., discussed a new proposal that predicts the power factor based on relaying on phase-phase currents to facilitate data acquisition, which leads to reducing time and cost for a simple analysis of energy efficiency and quality in consumption facilities. The Random Forest model showed us that it gives a modern result for various sites with various loads. The mean error with the coefficient of determination were acceptable. I have shown results about predicting pf changes using phase currents as input variables in power systems as pf reflects the energy consumption from the network. There is the possibility of changing a more developed model to predict the changes of the pf when the phase flashes do not show us their great correlation due to the specific stabilization conditions and taking into account the presence of renewable energy sources that are connected to the network [42].

2.3 CONCLUSION

The Internet of Things today is the brightest and most important term in the world of technology and more important in relation to networks, where there are many smart devices in data or cities that can communicate with each other with the Internet of Things, meaning in the end we will have a huge amount of data and this in turn will help in the emergence of how much is not With few services and applications provided by the Internet of Things, on the other hand, machine learning is an area no less important than the Internet of Things, as it is an important field that has entered many fields, whether processors or research studies So machine learning and the Internet of things are both very important and strongly interfered with the consumption of electrical energy and the classification of devices, as proven by the experiences of scientists and researchers through the studies that were presented, is also in the process of studying and continuous research to find optimal solutions in reducing the consumption of electrical energy.

3. THEORETICAL BACKGROUND

3.1 INTRODUCTION

This chapter includes the theoretical side of the work. An explanation has been given fully about the sources of alternating current and the electrical circuit, as well as an explanation of the form of alternating current, how to represent it, and how to analyze alternating current. As for electrical loads, the classification of electrical loads has been explained in addition to an explanation of Ohmic, capacitive and inductive loads. Separately. The chapter also includes an explanation of the phase and voltages applied to the LCR series and a phase diagram solution. We have also dealt with the power factor and its correction and the leading and lagging power. The chapter also includes the importance of machine learning and its types, focusing on the types that were used in this research, such as the decision tree, and the explanation of reinforcement was chosen, and the two most important algorithms of this method are LightGBM and CatBoost. PERCEPTRON (MLP) in detail and in the conclusion plc was explained in detail and what the logical control unit means and how it works.

3.2 ELECTRICAL CURRENT SOURCE

Since the beginning of the emergence of this huge science of electricity, the so-called war of currents appeared, i.e. the war between direct current and alternating current. This war has its roots since the beginning of the world of electricity. Alternating current is symbolized by AC and direct current is symbolized by DC. Electrical scientists such as Edison and Tesla each have a point of view in the method of electricity. Edison supported DC, while Tesla supported AC as methods of electricity. At the beginning of those battles, AC won the battle as a result of the invention of transformers, as this method enjoyed its effectiveness, ease, and importance in terms of costs to step up and step down voltages according to its need. Where the alternating current was considered the hero in all branches of electricity such as generation, distribution and transmission. Thus, AC imposed its supremacy and dominance for a long period of time.

Researchers have made continuous and great efforts to find new generation methods that help in inventing renewable resources such as solar panels and fuel cells that produce energy in the form of DC.

The direct current has proven itself in the transmission branch, as it is distinguished from the alternating current due to its low losses. With continuous and advanced research in the relative field, the DC started its way in distribution, as the direct current challenged the alternating current at the level of the distribution branch as a result of providing high-efficiency electrical transformers as well as many electronic loads current operated.

While other researchers presented a model of a continuous current system with different operating voltages and suggested that the voltage of 326 volts is the most appropriate value, taking into account the voltage drop and the amount of energy, while Nilsson and Sannino presented a consideration of the designs of three AC, DC, and another mixed AC/DC, where they proved that the mixed system is the worst. In terms of system losses, they proved that the direct current system is better than the alternating current system only when we consider the semiconductor losses half. With the use of DC as a means of power transmission, many research efforts were made that included generation, distribution, and small networks, where the losses in power transmission, which included conversion losses, were of interest to many researchers. In favor of direct current in its preference over alternating current, along with many successes of DC in various fields, but DC differs from AC in many areas, the most important of which is protection. As a result of the absence of empty points, DC, if used at the distribution level, represents a threat. This is in addition to the most important property, which is the property of calculating efficiency Where the efficiency of the DC-DC converter differs, and thus this will lead to a decrease in the efficiency of the direct current in relation to the alternating current [43].

3.2.1 An Electric Circuit

Voltage, current, and resistance have a major relationship, where the electrical circuit consists of voltage, a type of load, and connectors that give electrons to flow between the load and the voltage sources, so a battery circuit provides a source of voltage, while the wire is used for the connector and provides light for the resistance. In the case of adding the switch to the circuit, there must be a path for the flow The current. When the switch is open, this means that the path will not light because it is incomplete, but when the switch is closed, the path will be complete, allowing electrons to leave the negative terminal and flow through the light towards the positive [43].

3.2.2 AC Current Form

The alternating current usually flows opposite the direct current at first in one direction and then in the opposite direction. The most commonly used and common form of AC is the sine wave, where this wave is known as the signal that does not change its shape or change through a real circuit. For this reason, it is ideal as a test signal When we talk about alternating current or discuss, we must express the current and voltage in terms of peak values, maximum value, peak-to-peak values, or the most effective, average, or instantaneous values, where each value has a meaning used to describe the current and voltage and their different quantities.

$$v(t) = V_P \cos(\omega t + \theta) \tag{2.1}$$

Where (V_P) is the peak voltage, ($\omega = 2\pi f$) is the angular speed expressed in radians per second (rad/s), f is the frequency expressed in Hertz (Hz), t is the time The function (Figure 12-1) starts at a value of (0) at (0^o) and rise smoothly to a maximum of (1) at (90^o) They then fall, just as they rose, back to (0^o) at (180^o) the negative peak is reached three quarters of the way at (270^o). The function then at (360^o) [44].

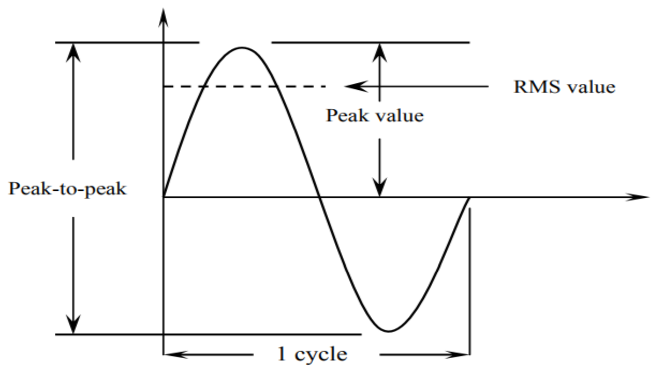

Figure 3.1: Alternative Current Waveform.

3.3 AC CIRCUIT ANALYSIS

The impedance parameters in the previous sections are of great benefit in solving the problems of the alternating current circuit because they give the possibility of making an exchange between most of the network theories developed for the direct current circuits by replacing the resistance with impedances that have a value that is not easy, but rather complex. The following figure will show us the value of each of the impedances (L, R, C) in at a complex level.

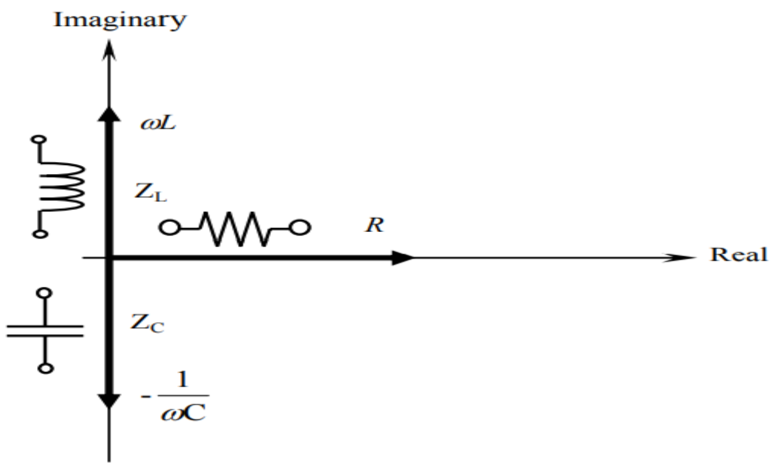

Figure 3.2: RLC Impedance.

All the laws that have been applied to DC circuits apply to AC circuits as well, including Kirchhoff's Law, Ohm's Law and the Laws of Network Analysis. The only qualification is that it is very important to express all variables in a complex form with attention and consideration of magnitude and phase. That all currents and voltages be of the same frequency in order for their phase relationship to remain constant. It is very important to emphasize that although the resistance of the circuit elements is either purely real for the resistors or purely imaginary for the inductors and capacitors. The general definition of the resistance of an arbitrary circuit should give the possibility of having a real part and an imaginary part, where the practical circuits consist of links less than the elements of the circuit and more complex [44].

3.4 LOAD CLASSIFICATION

The residential loads were divided into several sections, DC loads This category is usually for loads that require continuous power to operate, AC loads These loads require AC power to operate, AC and DC loads usually need both types to operate. Where these loads include large heating loads, when the system is simulated with direct current supply, the supply will be direct current, and for AC supply, these loads will be operated on alternating current.

3.4.1 Resistive Loads

There is just a single kind of burden called resistor in the DC circuit. This isn't accurate. AC circuits contain three kinds of burdens, which are opposition, capacitance and inductive. These heaps produce different circuit states from one another. Opposition is characterized as the voltage partitioned by the ongoing in the circuit concerning the impedance, it is the obstruction presented by the sinusoidal current component, that is to say, it is the amount of stage. Any circuit that has unadulterated obstruction makes the section of rotating current and voltage rise together and bomb together. It is conceivable that the current and voltage don't have a similar limit, yet they are in a similar stage. The circuit on an obstruction will create heat. We will see that the voltage and current are in stage with one another in an unadulterated circuit as in the Figure (3.3).

It can't create genuine energy except if the current and voltage are either sure or negative. At the point when the indications of deference for the item increase, it will be positive, however assuming the signs duplicate, something contrary to the item signs will be negative. Since the current and the item are either certain or negative simultaneously. The item is watts, normally it will constantly be positive.

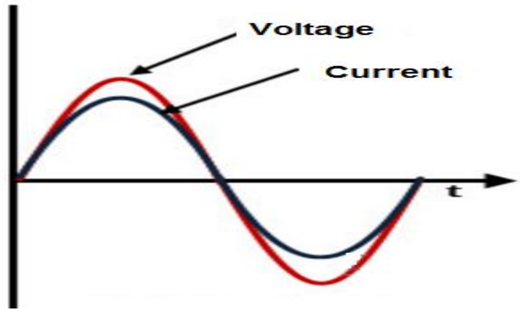

Figure 3.3: Voltage and Current Phase in Resistive Load.

It has an opposition of a similar size as the resistor so its units are ohms. Typically the voltage across the resistor is in stage with the ongoing through the resistor in light of the fact that the stage point is 0 degrees where the 0 stage shift demonstrates the voltage across the resistor is in stage with the ongoing through the resistor having no fanciful part [44].

$$Z_R = R + j0 \tag{3.2}$$

3.4.2 Inductive Loads

The Henry unit (h) is the unit in which the inductance is measured (inductance). The material composition of the coil is dependent on it, such as the cross-sectional area (A), the length, the number of turns of the wire (m) and the amount of permeation of the material in the core (u), all of which are expressed mathematically as follows

$$L = \frac{N^2 \mu A}{l} \tag{3.3}$$

Inductance is defined as the main load within the AC circuits. Due to the presence of the magnetic field, which is constantly changing, then there is some amount of inductance within all AC circuits. When using any load that contains a coil, this means that the circuits contain inductance, as it contains loads such as transformers and generators On a coil of wire, usually when current is flowing through a coil, a magnetic field will be formed around the wire. The decrease in the amount of current will lead to the collapse of the magnetic field. If inductance is present in the circuit the opposite force of the electromagnetic equipment will be sufficient to prevent the current from remaining in phase with the applied voltage [44].

$$X_L = 2\pi f L \tag{3.4}$$

$$Z_L = R + j\theta \tag{3.5}$$

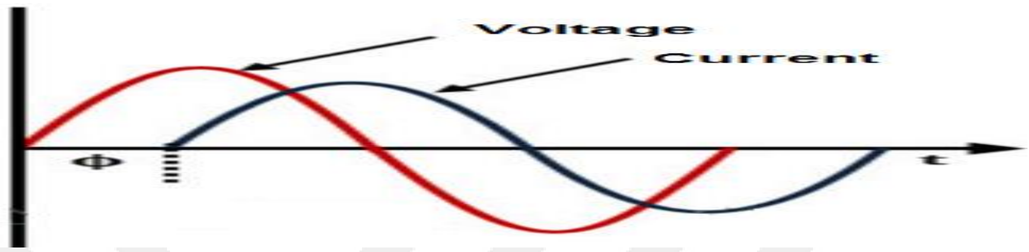

Figure 3.4: Voltage and Current Phase in Inductive Load.

3.4.3 Capactive Loads

The resistance to an adjustment of current in a capacitor does the inverse. It goes against an adjustment of the voltage of the current. To forestall direct current in circuits, capacitors are utilized. They are framed when two clients work one next to the other. The ongoing through the capacitance really relies on how rapidly the voltage changes.

$$i = C\frac{dv}{dt} \tag{3.6}$$

On the off chance that the opposition of a capacitor is an intricate amount that relies upon recurrence, then, at that point, the obstruction of the capacitor changes as a backwards capability of recurrence. At high frequencies, the capacitor fills in as a circuit, and at low frequencies it functions as an open circuit. Condition (3.7)

$$X_C = \frac{1}{2\pi f C} \tag{3.7}$$

indicates that the opposite is provided by the sinusoidal capacitor, where the voltage is inversely proportional to the size of the capacitor (C) and the value of the frequency. The current is shifted 90 degrees before the voltage. In a rectangular shape, the resistance of a capacitor has both an imaginary and a real component.

$$Z_L = R - j\theta \tag{3.8}$$

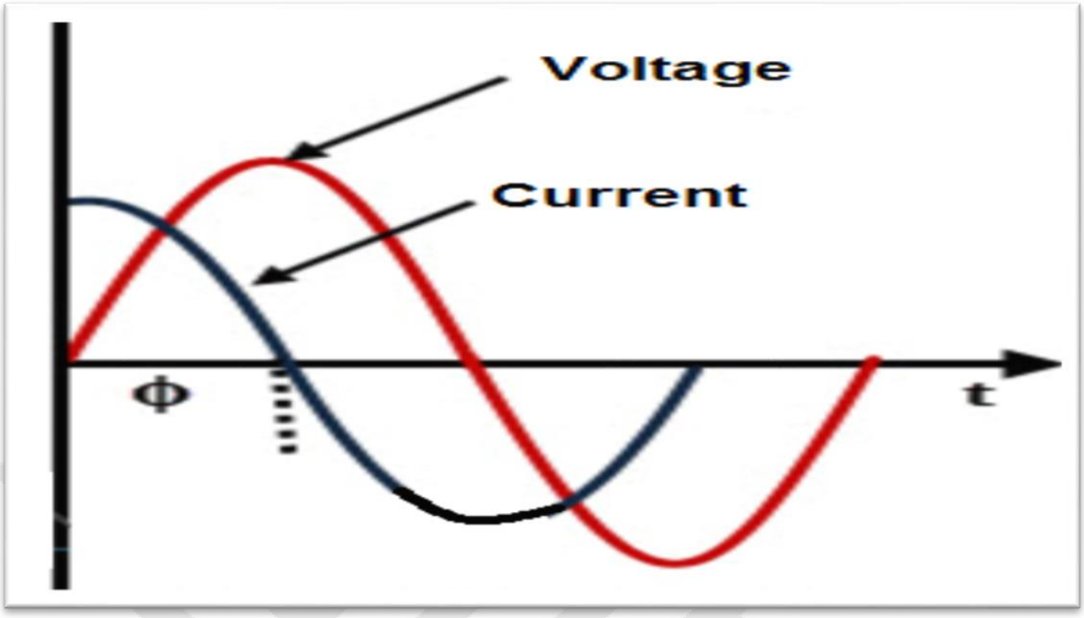

Figure 3.5: Voltage and Current Phase in Capacitive Load.

3.5 ELCETRICAL PHASE

Two sine waves are said to be in phase. To be in phase when these two waves are identical to each other, the two waves must pass at the same time and direction at their minimum and peak points. The term lead and lag is used to describe the phase relationship between two sine waves. Degrees are a measure of the amount by which one or more sine waves lead. different from others. According to Figure (3.6)

$v(t) = V_P \cos(\omega t + \theta)$ occur θ rad, θ degrees seconds, earlier. In this case we say VP sin (ωt + θ) leads VP sin ωt by θ. Also, we may say that VP sin ωt lags $V_P \cos(\omega t + \theta)$ by θ.

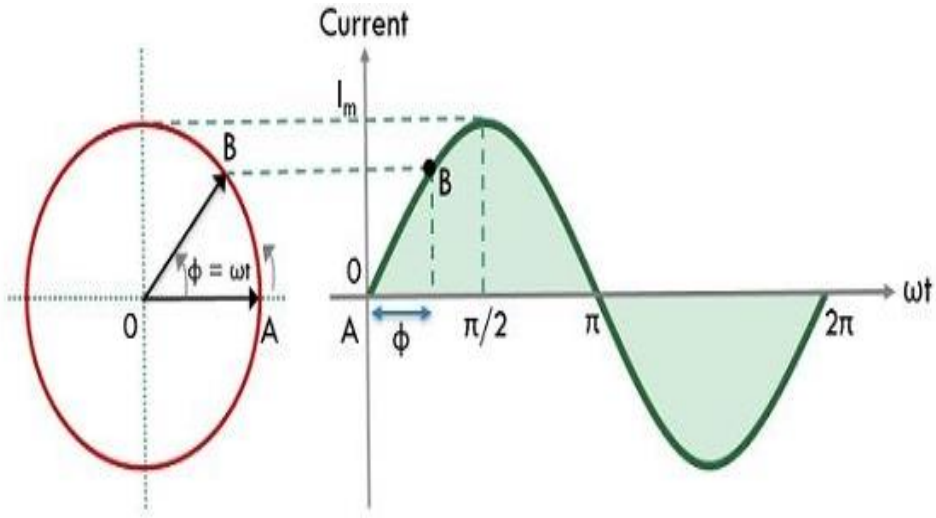

Figure 3.6: Phase Calculation.

So as a rule, one sine wave can perform or contrast by quite a few degrees from another sine wave. Aside from degrees 0 and 360. At the point when the case is 360 here, it is said that the two waves are in stage, and this implies that the two sine waves contrast in stage by 45 degrees. For instance, as a matter of fact, they are out of stage with one another, while two waves are viewed as in stage with one another when those two waves contrast in stage by 360. While characterizing two waves, a stage distinction between them, we will characterize the focuses on the time pivot, for example the two waves address the time pivot moving in a similar heading. While the stage contrast is the quantity of degrees between the focuses. Where the wave that crosses the hub sometime in the not too distant future (to one side) is said time hub) is not quite the same as different waves Beforehand, it was perceived how to decide and address in one condition the size of the sinusoidal sign, its recurrence and its stage shift. Any direct circuit contains inductors, capacitors, and resistors that don't change the shape or recurrence of this sign, however regardless of this, the straight circuit changes the adequacy of the sign (enhancement) and movements its stage (this makes the result signal lead to the info signal and its postponement). The recurrence and capacitance are considered as the two amounts in particular in the manner the circuit influences the sign then we can say that the sign is a straight blend of perplexing sinusoids. The size of the stage (vector) decides the stage and the perplexing number. The vector stage is like the vector concentrated on in science. The accompanying figure (3.6) will show the sinusoidal amounts AC and their portrayal by setting a pivoting vector. At the point when the vector turns, it creates a point. The area of the vector not entirely set in stone on the outer layer of the plane by the created point and the size (length) of the vector. Addressing sinusoidal signs utilizing stage is helpful Very on the grounds that circuit examination regulations, for example, kVL, KCL and more natural investigation apparatuses, for example, series equality, voltage, equal and current are appropriate in the field of stage as they have been concentrated on in DC circuits. We needn't bother with any new strategies to manage any circuit in the field of stage. The main contrast is that The circuit's reaction will be numbers for example stages rather than dc signals. To manage complex numbers without drawing vectors we will require a standard numerical documentation of some sort. There are two principal sorts of documentation for complex numbers which are polar and rectangular [45].

The previous figure represents the magnitude of that sine wave a and b will represent a vector with the end fixed at the origin and rotating counterclockwise representing variable conditions of an alternating current signal [45].

3.6 POWER AND POWER FACTOR

When studying and understanding the load and its characteristics in this process, the understanding of the load within the electrical systems will include an understanding of the term energy and its factor, as the energy consumed by the load will consist of several components, which are the apparent, interactive, active or real power [45].

3.6.1 Power Components

One of the energy components is the real component of the load. It is the component that does a real job. It uses the interactive component of the load to supply the energy stored within an electric or magnetic field [46].

An example of reactive energy is the magnetic current consumed by the electric motor. An example of the device that supplies the reactive energy is the capacitor. Whether the load consumption of reactive energy is first, this in itself is a characteristic of the load itself. The best solution to achieve the relationship between the energy quantities will be by using the power triangle. When he applies alternating current power to the load of a reactive nature, the voltage will be 90 degrees out of phase with the current, while when the instantaneous capacity of the voltage and current doubles, this will lead us to the fact that the resulting wave will represent the instantaneous power of the reactor, so from all of this it will be reached that the average power is zero. This It indicates that energy is not dissipated by reactive loads. Upon observing Figure (3.7), a rule will be created such as the power will represent the real power while the head will represent the reactive power and the string will represent the apparent power, which will be calculated as [47]

$$S = \sqrt{P^2 + Q^2} \tag{3.9}$$

S = P + Q, the equation is incomplete

s = the magnitude of the apparent vap

P = the magnitude of the active force in w

Q=represents the magnitude of the reactive force in var [44]

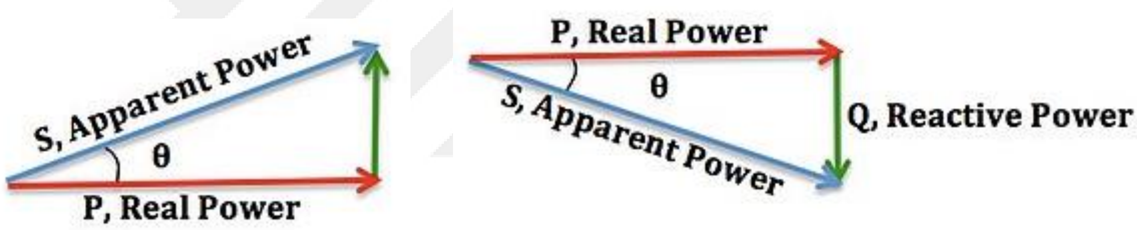

Figure 3.7: Power Factor Calculation for Lagging and Leading State.

3.6.2 Power Factor

To begin with, in order to understand what the strength factor is, one must understand the following:

Power Activities (real power kw) are responsible for the implementation of a useful order in addition to operating the equipment if

$$P = VI\cos(\theta) \tag{3.10}$$

Reactive force (kVAR) All magnetic fluxes are generated and magnetic equipment such as transformers are generated using reactive force

$$Q = VI\ \sin(\theta) \tag{3.11}$$

Apparent power (kVA) is a mixture between active and passive power $S = \sqrt{P^2 + Q^2}$

The power factor represents the relationship between actual power and apparent power

$$Power\ Factor = \frac{Q}{S} \tag{3.12}$$

The phase angle of the load resistance is of great importance in absorbing energy through the load resistance. So the power factor expresses the angle between the voltage phase and the current phase in the electric loads, which range from (0-90), where the Cos power factor represents the angle and its value is between 0 to 1, and each The closer the number to one, the better, and the lower the value of the PF than (1), the load will be supplied with additional current from the source, which leads to a decrease in the efficiency of power transmission and the occurrence of other problems, and if the angle between the phase of the current and the phase of the voltage is 0, then Cos (0) = 1, so the power factor is equal to one in the case of (Ohmic loads).

When the energy factor is low, it leads to many disadvantages, the most important of which are:

a) Losses and burning of conductors and cables.

b) Combustion transformers.

c) Energy loss as we would have wasted energy.

d) An increase in generation in order to compensate for losses.

e) Environmental pollution as a result of gas emissions.

While applying the ability factor in a way that will have many positive effects, including:

a) Environmental benefits, as the less energy consumption is reduced as a result of energy efficiency, this decrease in consumption will lead to a reduction in greenhouse gas emissions.

b) Reducing the burden on electrical materials, such as cables, which will extend the life of this equipment.

c) Low cost of electric bills.

d) Any kVA will minimize losses in equipment and transformers.

e) Reduce voltage drop in long cables.

3.6.3 Power Factor Correction

Power factor correction (PFC) is a very important factor, as power factor correction increases as the reactive power decreases, where the power factor is 1, so the energy consumed in machines goes to provide useful work, such as motors. Figures (3.8) and (3.9) show both the instantaneous power in its waveform and the average power calculated from the current voltage, showing the power factor of 1 and 0 and the delayed power factor [48].

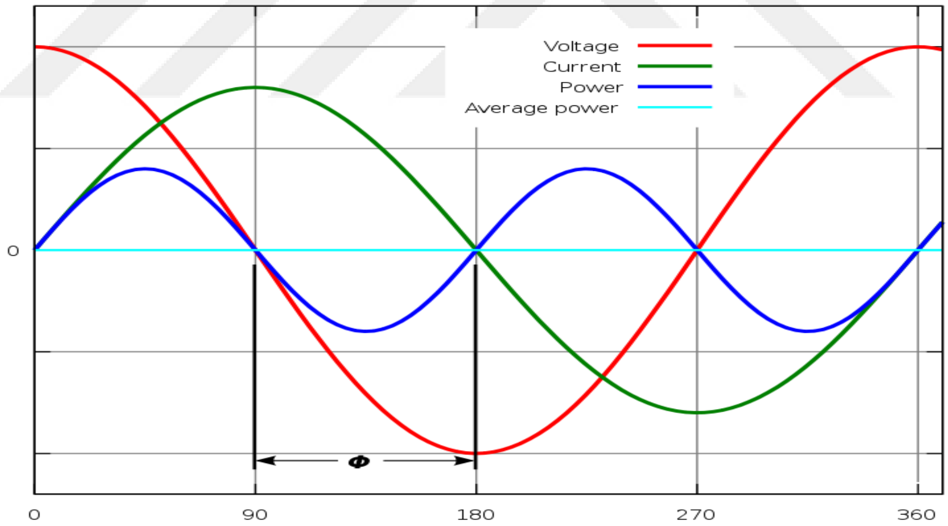

Figure 3.8: Power Factor Waveform

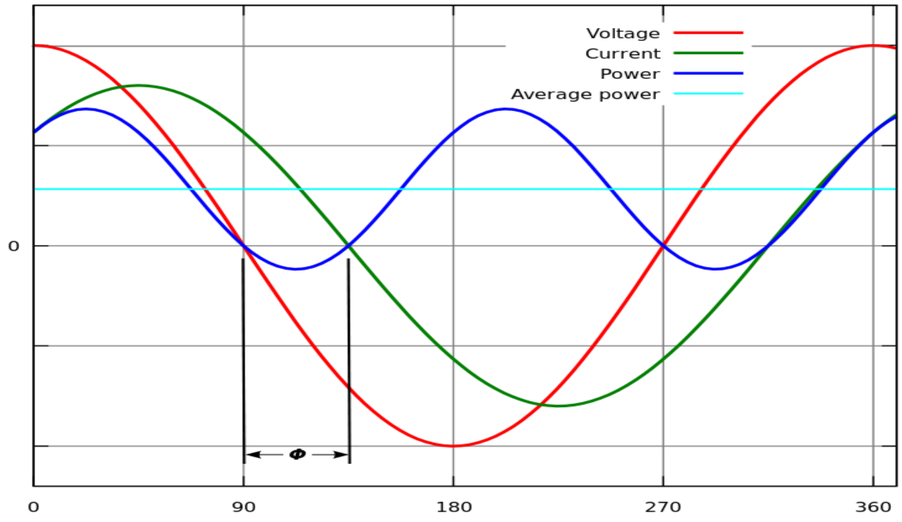

Figure 3.9: Lagging form Power Factor.

3.6.4 Power Factor Correction Equipment's

There are equipment's that help in power factor correction such as

a) Synchronous motors: Synchronous motors are usually used to improve the power factor due to the fact that they produce reactive power in high excitation mode where KVAR is produced by these motors during high excitation and idle process usually applied by PF size correction with the addition of power to regulate the correction magnitude by changing the excitation.

b) Fixed capacitors: These capacitors are characterized by being free of loss and usually acquire a leading current. They are also used in AC motors to support the pf. These capacitors are installed to improve the pf of AC motors.

c) Stage progression: It is fixed with all separate or induction machines and when each device is turned off and on with the motor. The most important characteristic of the connection at this particular stage is the support of the power factor and its achievement in the cables of consumers and all electrical switches until the end point of the motor.

3.7 MACHINE LEARNING

Learning itself, as a general term, means acquiring a new behavioral approach, skills, or recognizing new values. So learning and social learning determine how human learning is. As for machine learning, or machine learning, it is a branch of artificial intelligence that helps machines learn and think on their own, that is Computer hardware systems can learn and improve performance on their own. The term machine learning appeared for the first time in 1959 by the scientist Arthur Samuel, where he defined machine learning as a field of study that gives the computer learning without being properly programmed [49].

In the absence of machine learning algorithms, learning cannot be complete because these algorithms act as an input source for the model to help understand the main features of the data structure. Algorithm learning is often referred to as training models. Machine learning grouped these algorithms into three main types [49].

3.7.1 Supervised Learning

In this type of learning, training sets are provided with their correct outputs. On the basis of these outputs, the algorithm learns how to respond with higher accuracy by comparing the outputs with the input values. Supervised learning is also known as learning from models. Supervised learning includes regression and classification, as it is in the case of classification. The outputs are non-continuous, i.e. discrete, while the outputs are classified as continuous, i.e. not discrete.

3.7.2 Unsupervised Learning

This type of education revolves around recognizing patterns with unlimited data in order to infer rules from them. This type of education is appropriate when unknown data are available here. Training data has not been classified, as this type of learning is considered as a statistic with a learning-based approach.

3.7.3 Reinforcement Learning

This type of learning is medium than machine learning, as it provides algorithms with a response that indicates if the output is correct or not, i.e. the algorithm must explore and exclude various possibilities to obtain correct output [49]. This kind of algorithm consumes a lot of memory space.

Machine learning algorithms are classified according to the type of learning and the type of inputs and outputs as the supervised learning algorithms These algorithms are characterized by the process of learning those algorithms with the training set where the output and input variables are present The goal of all this is to approximate where we can predict the outputs of the data We can collect the problems of supervised learning For supervised regression and classification, as for unsupervised learning algorithms, these algorithms have only inputs, i.e. there are no corresponding outputs, where unsupervised learning problems can be grouped more importantly in relation to extraction output problems. In Figure (3.10), we will show a summary classification of supervised and unsupervised learning algorithms [50].

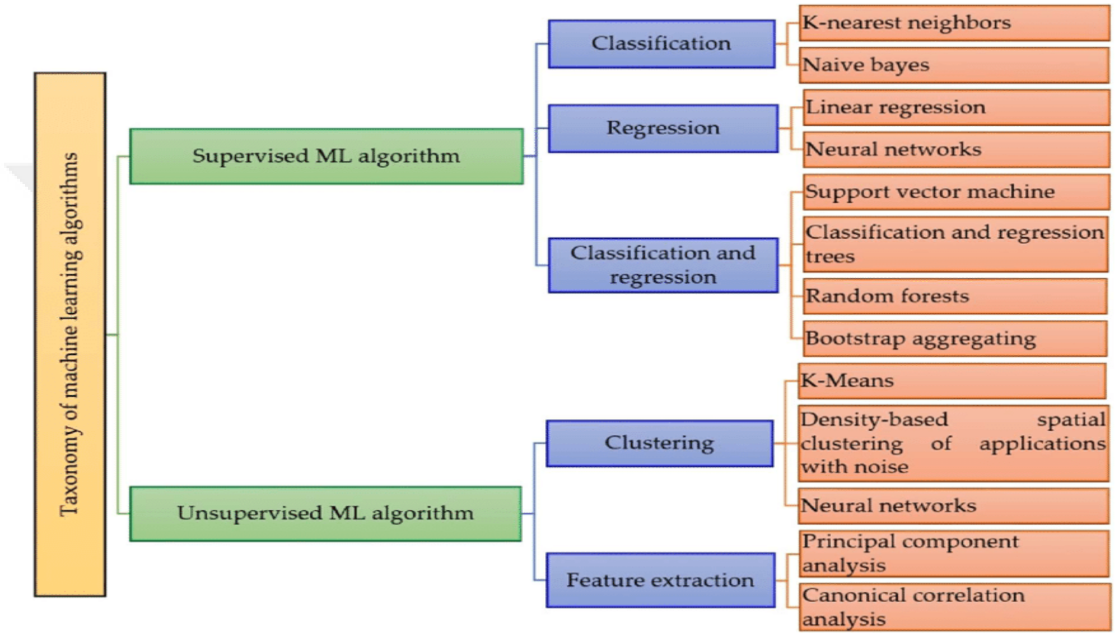

Figure 3.10: Summarized of Supervised and Unsupervised ML Algorithms.

3.8 ALGORITHMS USED IN THIS WORK

More than one algorithm used in this work, they are:

3.8.1 Decision Tree

The great need to obtain huge amounts of useful data that accompanied the development of computer technology called for the urgent need for a way to extract and deduce useful data from huge, unclear and random data. For this reason, we have the decision tree technique. This technique has proven over the years its effectiveness in mining useful data, Since the decision tree is one of the supervised learning techniques, therefore, we can say about the decision tree that it is a model that has a hierarchical method, the main objective of which is to access the required data, whether in classification or regression, both of which help in improvement and increase in efficiency.

The structure of the decision tree is distinguished by its resemblance to the structure of the ordinary tree. Just as the ordinary tree consists of leaves, branches and roots, the same applies to the decision tree, as it consists of roots and leaf nodes, where the root node is considered the origin of all nodes, and each node represents a characteristic of a certain medal, and each branch is a specific decision, and each leaf is a result of the idea. For all of this, it is to create a data tree and find the result of the treatment in each paper [51]. The structure of the decision tree is shown as in Figure (3.11).

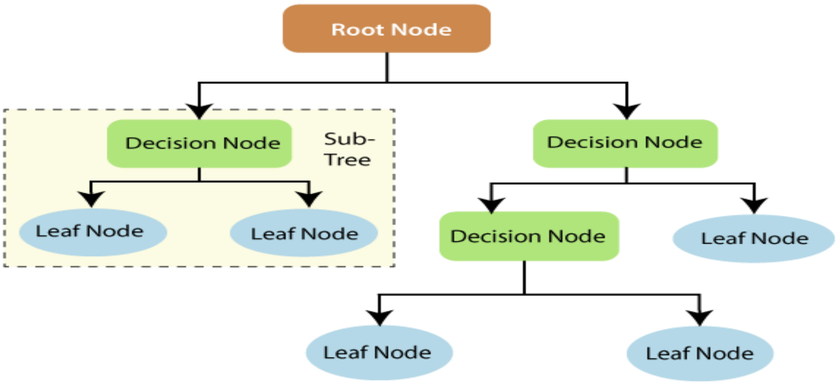

Figure 3.11: Structure of DT.

The work of the decision tree is very similar to human decision-making, as it is easy to understand its work, as this technique helps in solving the problem, whether the inputs are discrete or continuous. For example, each node will represent a feature that will be classified, and each group will know the value that the node can take. The tree will perform a simple analysis on multiple data [51].

3.8.2 Boosting

Support is one of the main techniques for AI and one of the main instances of a choice tree, as it can change over frail students into solid classifiers, as a sort of significant engaging calculation is utilized to decrease difference and predisposition, and simultaneously, any feeble student is a classifier that performs somewhat better compared to Irregular speculating areas of strength for while accomplish high and great precision as they are the center of building calculations for support gatherings. The calculation was examined without precedent for the world in 1990 by SCHAPIRE by responding to an inquiry presented at the time by Kearns and Bold where it was feasible for a gathering of frail students to give areas of strength for one SCHAPIRE's work has had a colossal effect in the realm of AI and measurements. This prompted the huge advancement of a few supporting calculations including AdaBoost and XGBoost. The primary thought behind helping is to incorporate the powerful utilization of learning calculations to altered adaptations of electronic information. Support methods, specifically, utilize input information to prepare the feeble student, process forecasts from the powerless student, select examples of misclassified preparing tests, and coordinate and train the connected frail student with a changed preparation set that incorporates misclassification. Models from the preparation round. The iterative growing experience comprises of fundamental components until Getting a foreordained number of key students, guessing all students, and weighting the key students together. Showing Ds could zero in a greater number of on diminishing predisposition than change. Consequently, reproductions build up essential students with a serious level of both predisposition and low fluctuation, eg choice trunks (a tree with one inward endeavor at a hub) [52].

The wrongly grouped examples put on more weight, which prompts the focal point of the essential student and the student of such examples, and afterward, when the standard is ideal, a few classifiers will be one-sided against explicit examples, which makes them give those examples more weight and afterward right the predisposition. In any case, this iterative learning approach makes how much clamor unacceptable for learning the boisterous information on the grounds that the heaviness of the uproarious examples is many times a lot more noteworthy than those loads that were given to different examples, and this will prompt be constraining the calculation to settle it so it centers vigorously around the loud examples. This prompts overfitting. In spite of this, support based strategies are all among the best calculations and techniques in AI [52].

3.8.3 Lightgbm

The stamping machine called LightGBM is an effective execution of the scaling upgrade calculation. It was created in 2017 by Microsoft researchers[53]. This calculation can be utilized for arrangement, positioning and other AI issues. LightGBM utilizes two new techniques, uneven Angle Based Testing (GOSS) and the other strategy is Select Element Pooling (EFB), which helps and guarantees that the calculation prepares quicker and accomplishes exceptionally high precision. GOSS innovation is a change of Inclination Supporting innovation. Which considers the preparation cases that lead us to a more noteworthy inclination, and that at last prompts speeding up the growing experience and diminishing the computational intricacies of the model. The (GOSS) method explicitly incorporates barring various not a couple, however an enormous number of outstanding guides to compute data securing [54]. The reason behind excluding those samples that are characterized by their small gradations is that large gradient cases are more useful and beneficial in calculating information acquisition (IG), and then the (GOSS) technique achieves a distinct estimation of (IG) with a low sample size [55] at the same time The (EFB) method is the task of selecting the feature through all the mutually exclusive and disparate features, and thus the number of features decreases [56] because many features are almost exclusive in the sparse world space. That is, by that we mean that it has almost zero value at the same time [57].

There are typical examples of exclusive features, including encrypted features that have one run. In addition, the (EFB) technology combines these features to reduce the dimensions of the matrix and the attributes [58]. The (LightGBM) enjoys a significant benefit, which is speed, which frequently prompts an exceptionally effective model. Besides, it is portrayed by its low memory utilization, as it changes over consistent qualities into boxes. Thirdly, it accomplishes high exactness, outflanking the greater part of the support techniques, coming about because of the presentation of both (GOSS) and (EFB) advancements. At long last, this calculation functions admirably while prepared utilizing huge informational indexes, and more rapidly than preparing the (XGBoost) calculation [59]. As for the shortcomings of the (LightGBM) algorithm, it can easily accommodate small training data, as it works better with large data. Dividing the leaves of the tree may lead to an increase in processing as a result of producing trees of a more complex nature. At the same time, LightGBM was applied to various classification problems and achieved more than impressive results [60].

3.8.4 Catboost

CatBoost calculation is an execution of slope supporting which was proposed by Prokhorenkova et al in 2017[61]. This calculation bargains fundamentally and successfully with absolute elements during the preparation stage. One of the most striking enhancements to the CatBoost calculation is its capacity to perform impartial scaling and decrease overfitting. So to appraise the variety array of all models for each increased emphasis. The CatBoost calculation disposes of this model from its utilization to prepare a current model. One more significant improvement in this calculation is the course of consequently changing over absolute highlights into numeric. Clear cut highlights contain a discrete arrangement of values that are frequently not practically identical. Hence, these highlights are not appropriate for building choice trees in their present status, as absolute elements are frequently changed over completely to mathematical prior to handling. Hot coding is quite possibly of the most generally utilized and applied technique on straight out highlights that have low fundamental connections. Where the first characteristics will be supplanted by a parallel variable, and more than that, there is a strategy utilized with all out qualities, which is the covetous objective based insights (Ravenous TS), where the downright characteristics are supplanted with an identical rating worth of normal age [52].

3.8.5 Neural Networks

Artificial neural networks consist of many layers, which are the input layer (units and nodes), two or three layers, or one hidden layer, and the last layer is for results or outputs. The Figure (3.12) will show the structure of the neural network, as it will show us the lines that will connect the cells with each other, where each connection will be associated with what is called the weight, which is intended as a digital connection. The output, hi, from neurons (i) in the hidden layer.

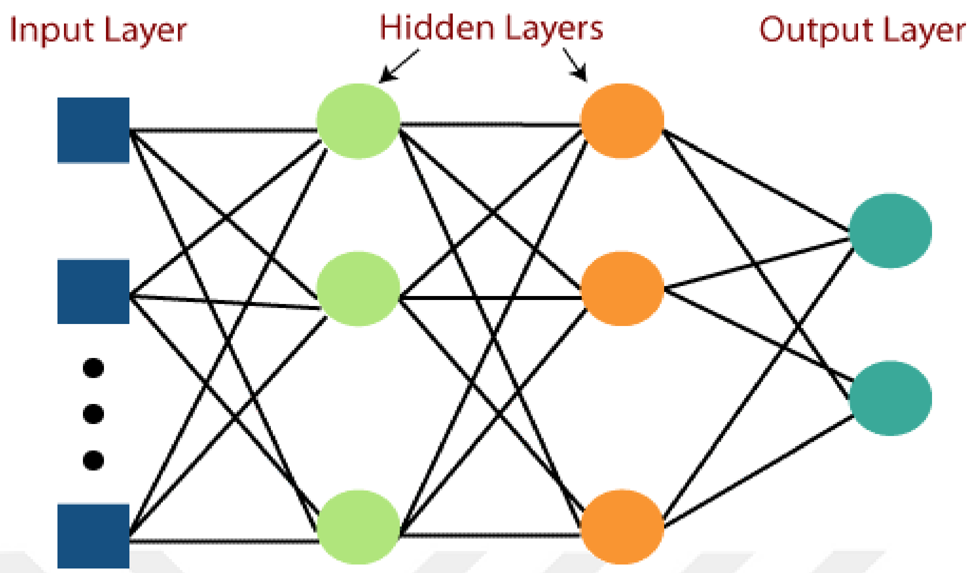

Figure 3.12: Neural Network Architecture.

3.8.6 Multi Layer Perceptron (MLP)

MLP, or what is said about sensory perception, usually consists of several layers, as it represents a neural network that is not linear with multiple layers. Those layers that it consists of are an input layer, another for output, and a hidden layer that can be one or more. The activation function is not any. A hidden neuron would be a nonlinear sigmoid or tan sigmoid. Each neuron is connected to all cells in the layer and is characterized by the function of activation Either each link will be distinguished by its weight factor or its network weight [62].

3.9 PROGRAMMABLE LOGIC CONTROL (PLC)

The logical control unit introduced this unit for the first time in the sixties when the scientist Dick Morley presented the first model for this unit in 1968 after that this unit became famous and spread more widely after the seventies [63]. However, it is preferable to call or use the abbreviation plc on the logical control unit in order to distinguish it from the abbreviation pc of the personal computer. This unit is considered the most necessary units in industrial automation [64].

PLC logic control was designed in the world of programming to take the place of paging systems, as plcs were programmed by “ladder logic”, which is very similar to logic diagrams. This control unit is a special model of a microprocessor-based control unit. This unit contains memory that can be programmed for instructions related to storing and executing functions such as logic, sequences, timings, arithmetic, etc., as in Figure (3.13).

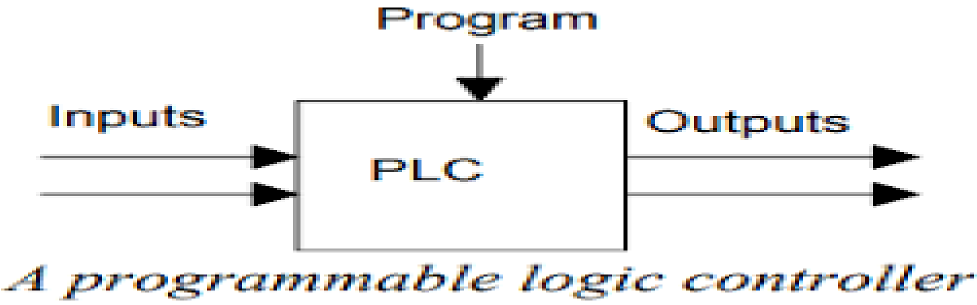

Figure 3.13: Programmable Logic Control.

PLC appeared to control machines and processes, and it can be managed by engineers who have limited knowledge of computers and computing languages. It has a very important feature that is able to change the ladder diagram of the PLC after its operation or construction. This gives the possibilities of using one PLC itself to control multiple different systems after scanning in Each time has many advantages as well, such as not being affected by heat, cold, noise, or spare parts, and the possibility of programming it easily in any language that has an interest in logic and conversion processes. It also has the possibility of arranging outputs and inputs. It also enjoys a very low cost compared to other macro-controller systems. The reason for this is the multiplicity of using the plc in several applications. It only requires changing the program for each application. In the case of using a microcontroller, the components must be changed for the devices themselves [65].

The PLC contains the processing unit, memory, power supply unit, communication interface, output and input interface, where each component has a function as in the following figure (3.14).

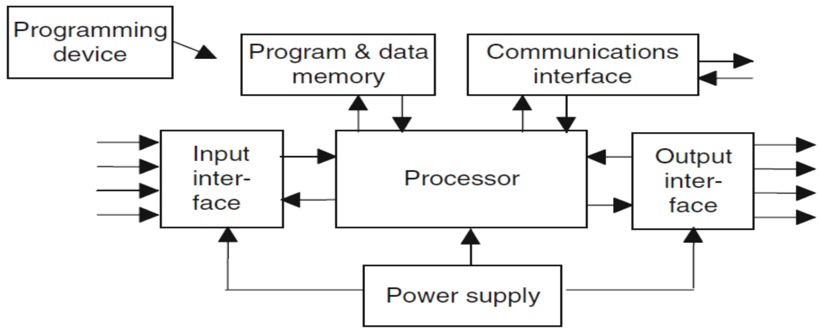

Figure 3.14: PLC Structure.

The working mechanism will be explained in Figure (3.14) in a simplified manner in the following points:

a) The program is put away to demonstrate the info signals and the correspondence between these info and result choices to execute the control that happens in the Focal Handling Unit (computer chip).

b) The power source depends on AC voltage to coordinate current, as it is exceptionally important for the processor and circuits in the info and result units.

c) The program is placed into the memory of the processor using the programming gadget, where it is then moved to the memory of the PLC.

d) The program is put away in the memory unit to be utilized in the control choices made by the chip and the put away information from inputs for processors and result for yield.

e) The information and result unit is the spot from which the processor takes data from an external perspective and imparts that data to outside gadgets.

The correspondence connection point is utilized to take in and send information over correspondence organizations to and from other far off PLCs. With respect to the readiness of the PLC program. The PLC program is executed as a feature of the tedious interaction alluded to as an output. The output begins with the PLC by adding something extra to the computer processor for data sources and afterward executes the application program utilizing the info states. At last, when the program finishes, the computer processor performs inside diagnostics and correspondence errands. The filtering system closes with an update of the result, and afterward it begins once more... The sweep process duration relies upon the quantity of data sources/yields, the output and how much correspondence required and the control program is based on things called guidelines where the directions are characterized as little codes that empower the information sources Results from doing how the programming obtains the outcome. There are various kinds of PLC projects, for example, Capability Block Graph (FBD), Manifest Records (STL) and Rationale Stepping stool Outline (Fellow). Fellow is one of the most famous programming dialects utilized in rationale regulator programming where the rationale stepping stool chart comprises of parts, for example, line graph coordinate components to depict the control on the hard wires. The left upward line for the Fellow addresses the power connector. The result component addresses the return way of the circuit [64].

4. SYSTEM IMPLEMENTATION AND RESULTS

4.1 INTRODUCTION

This chapter included the presentation of system design tools, its own scheme, and evaluation metrics that show the accuracy of the results, and then touched on the design of the practical control part based on programmable logic control (PLC). Then he moved to compare the results of this study with the results of previous researchers in terms of design, results and tools.

4.2 COLAB PLATFORM

Guido Rossum designed the object-oriented programming language Python in 1989. Rapid prototyping of complicated applications is best served by its design.

Colab is the Google Research product Collaboratory, or Colab for short, enables anyone to develop and run Python code through a browser and is especially well suited for machine learning, data analysis, and education. It is a Google-hosted, free cloud service. to promote AI and machine learning research [66].

Google offers a free Jupyter notebook environment called Colaboratory where you can leverage free GPUs and TPUs to address all of these problems. Nearly all of the modules required for data science analysis are present. These tools include Numpy, Scipy, Pandas, among others, but are not limited to them. Even frameworks for deep learning [67].Additionally there is more resource but it is not free I used it for our work in this paper for high Productivity and speedup the result.

4.3 PYTHON LANGUAGE USED LIBRARIES

Python is an object-oriented programming language created by way of skill of Guido Rossum in 1989. It is ideally designed for fast prototyping of problematic applications. It consists of interfaces to many strolling machine calls and libraries. colab.reseaech.google) used to be a gadget for carrying out the examine about Collaboratory, or Colab for its abbreviation Colab for its abbreviation, is a product from Google Research that approves every individual to write and execute Python code via a browser, and is particularly suitable for laptop learning, data analysis, and education. It is a free cloud service hosted with the resource of (Google) to encourage Machine reading and artificial Genius look up. Collaboratory allows the use of free GPUs and TPUs that can treatment all these problems. It consists of almost all modules that can be used for data science analysis. These tools consist of on the other hand are no longer restrained to Numpy, Scipy, Pandas, etc. It moreover consists of deep getting to know frameworks, such as (Tensorflow), (Keras), (Pytorch) and special tools and libraries.

a. TensorFlow: provides a set of tactics for Python mannequin improvement and training, as properly as rapid deployment in the cloud, on-premises, in the browser, or on a device. This library has been used by using Kersa, and Keras permits rapid and simple neural community software prototyping. It helps seamless (GPU) networks, convolutional networks (CNN), redundant networks, as nicely as their clusters.

b. Numpy: It is a Python package deal known as "Scalar Python," which refers to a library with multidimensional array objects and a set of strategies for manipulating matrices.

c. Pandas: Using its powerful facts structures, the Pandas Python library gives a excessive overall performance statistics processing and evaluation tool. The phrase "Panel Data," which refers to econometrics of multidimensional data, is the place the identify "Pandas" originates.

d. Sciklearn: It is Python's most realistic and strong computing device studying package. This library is frequently designed in Python and is based totally on NumPy, SciPy, and Matplotlib. It gives a range of amazing equipment for statistical modeling and laptop learning, together with classification, regression, clustering, and dimension discount with the aid of the consistency interface in Python.

e. Matplot library: As a cross-platform toolkit for drawing 2D diagrams from records in arrays, Matplotlib is one of the most general Python packages for records visualization. NumPy is used with the aid of Matplotlib, which is developed in Python. offers get right of entry to to an object-oriented API that allows the incorporation of processors into software program the usage of Python GUI toolkits like PyQt and Wx Python Tkinter. This API can additionally be used with IPython, Jupyter, and net utility servers.

f. Scipy: By giving the person high-level directions and training for statistics processing and visualization, the SciPy series of mathematical strategies and comfort features substantially enhances the competencies of an interactive Python session. By making use of SciPy, an interactive Python session may additionally compete with structures like MATLAB, IDL, Octave, R-Lab, and others as a facts processing and machine prototyping surroundings (SciLab).

4.4 EVALUATION METRICS

When training a classifier, the assessment scale is crucial to getting the best classifier possible. So selecting the right rating scale is crucial for differentiating and getting the best classifier. In order to enhance the generative classifier, this section has thoroughly analyzed pertinent assessment metrics that are intended to act as discriminators. Generally speaking, precision is a measure that many generative classifiers employ to identify the best solution while training [68]. Accuracy has various drawbacks, including being less informative, less discriminatory, and biased against data from the dominant class. Other measurements that are explicitly intended to define the ideal solution are briefly included in this paragraph as well. Also mentioned are these alternate measures' drawbacks [68].

Confusion matrix for a binary classifier. Actual values are marked True (1) and False (0), and are predicted as Positive (1) and Negative (0). Estimates of the possibilities of classification models are derived from the expressions TP, TN, FP, FN, which exist in the confusion matrix [69].

a. TP (True Positive) - The data of interest in the disarray framework is the genuine positive point (TP) when a positive result is normal and exactly the same thing occurs.

b. FP (False positive) - The data of interest in the disarray lattice is a bogus positive when a positive outcome was normal, and what happened is an adverse outcome. This situation is known as a sort 1 mistake. It resembles the gift of awful premonition.

c. FN (False Negative) - The data of interest in the disarray lattice is bogus negative when an adverse result was normal, and what happened is a positive result. This situation is notable as a kind 2 blunder and is considered as perilous as a sort 1 mistake.

d. TN (True Negative) - The data of interest in the disarray framework is Valid Negative (TN) when an adverse result is normal and the equivalent occurs.

Table 4.1: The Elements of the Evaluation Process (Variables, Definitions, and Equations) [70].

1. Variable	2. Definition	3. Equation
4. Accuracy	the percentage of accurately anticipated data from tests is easily determined by dividing all accurate forecasts by all predictions.	6. $Accuracy = \frac{Tp+TN}{TP+TN+FP+FN}$
7. Precision	8. he proportion of outstanding instances among all anticipated ones from a specific class	. $Precision = \frac{TP}{TP+FP}$
10. Recall	11. the ratio of the total number of occurrences to the proportion of instances that were supposed to be members of a class	12. $Recall = \frac{TP}{TP+FN}$
13. F1-Score	14. The phrase is used to describe a test's accuracy. The maximum F1-score is 1, which denotes outstanding recall and precision, while the lowest F1-score is 0.	5. $F1 - Score = 2 \times \frac{percision \times recall}{Percison+recall}$

4.5 DATASET COLLECTION AND DESCRIPTION

In order to measure the power consumed by the device and to know whether the behavior of the device is inductive or capacitive, the Arduino circuit was connected to a CT sensor to measure the current and a sensor to measure the voltage ACS712 for each device, and then a piece that was connected to the Wi-Fi ESP 8266 for the purpose of connecting the circuit to the internet and from Then save it to an excel file in Google Drive.

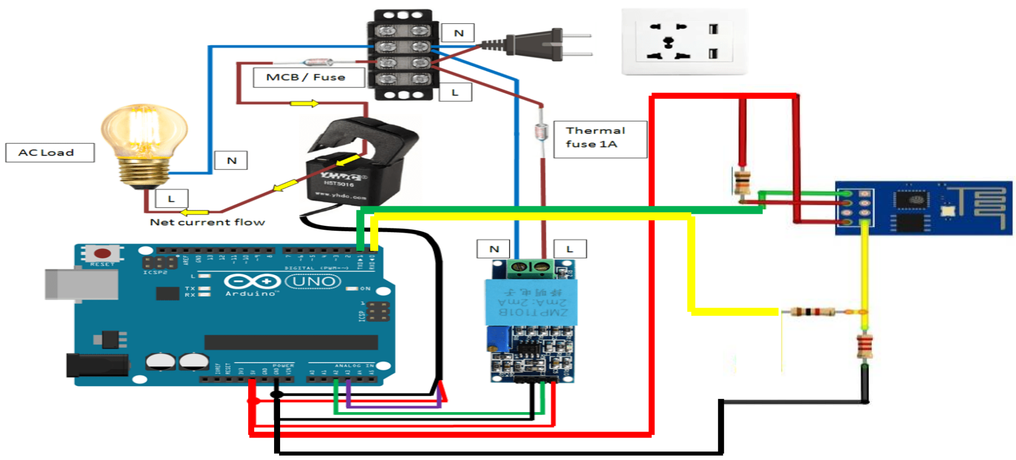

Figure 4.1: Data Set Collection Circuit.

Figure (4.2) shows the data set collection circuit for one device, as the data set included readings for approximately) 12,000 (homes, each house containing)128) devices distributed between inductive and capacitive devices. Table (4.1) Below shows a sample of data set collected.

Table 4.2: Sample of Collected Dataset.

Device1	Device2	Device3	Device4	..	..	..	..	Device 128
779.22	-524.31	524.98	-458.36	..	..	..	..	258.58
258.12	987.25	-789.12	258.36	..	..	..	..	-987.69
-456.25	852.69	357.21	875.69	..	..	..	..	953.36

4.6 SYSTEM BLOCK DIAGRAM

The diagram shows the progress of the process to implement the system, starting with taking data and processing it, then entering that data in several stages and applying several algorithms for classification and prediction on it. At first, the electrical loads were classified according to the percentage of loss in them, and then they were classified according to the type of loads, and thus the type of each load was predicted for each house, if the result Therefore, we have classified loads and a predictor of their type. We have entered these results into the PLC unit to address the problem.

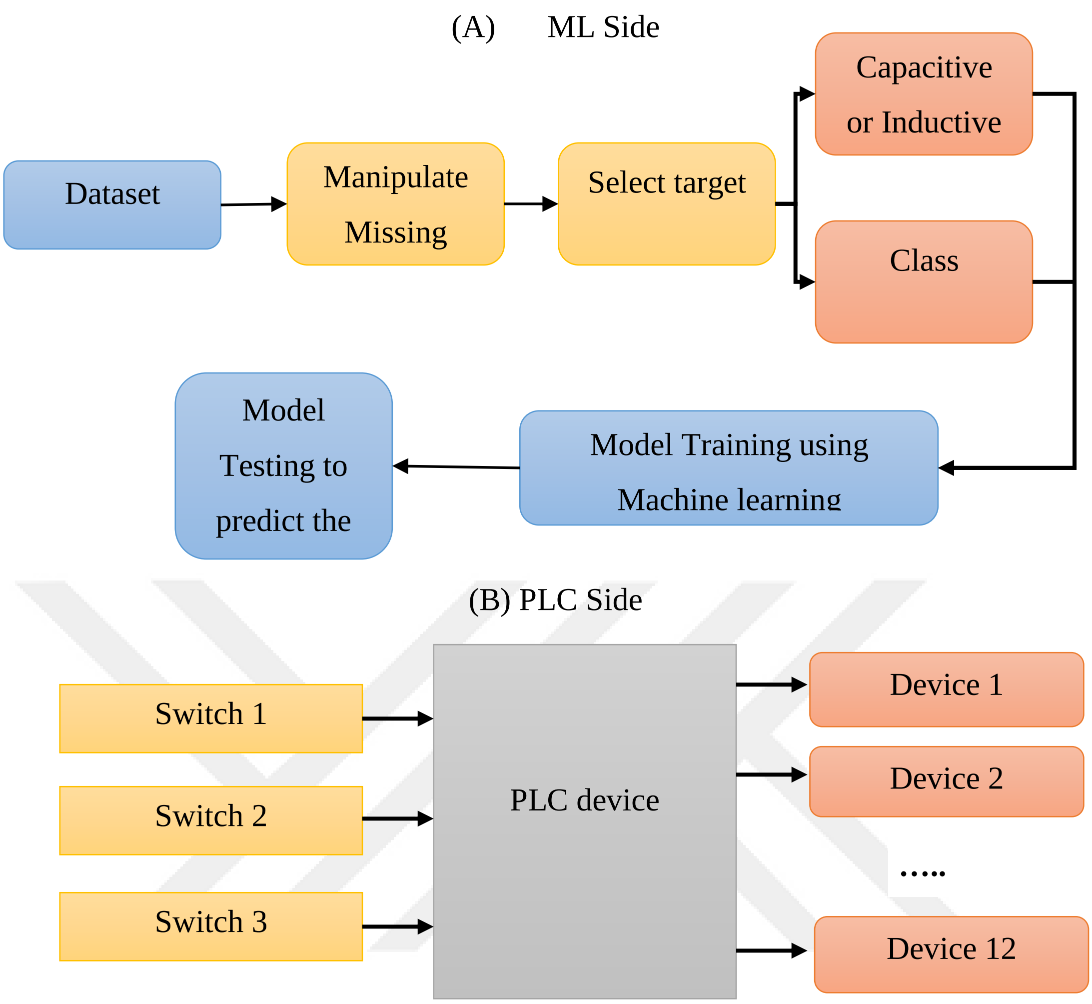

Figure 4.2: System Block Diagram.

4.7 MACHINE LEARNING RESULTS

The system of classifying electrical appliances according to their best consumption was implemented by machine learning, and in particular we used the classifiers of the decision tree in which we used the method of reinforcement, where two algorithms were used for reinforcement, as well as the implementation of a multi-layer algorithm for reception of Neural Network, the system was implemented in three steps.

4.7.1 Load Classification as (Inductive/Capacitive)

The first step included the process of classifying the electrical loads of a city according to the type of load, whether it was capacitive or inductive. The following results were obtained:

4.7.1.1 LGBM classifier results

In the decision tree classifier of the LightGBM algorithm, the best accuracy obtained is (98.75%). The confusion matrix diagram shown in Figure (4.3) and the classification report shown in Table (4.3).

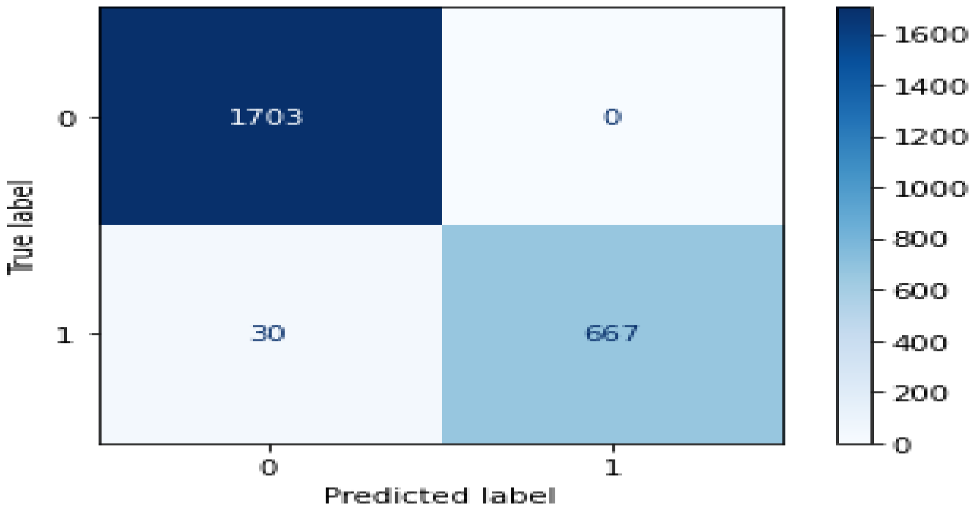

Figure 4.3: Decision Tree Classifier Confusion Matrix Light GBM Algorithm.

Table 4.3: Decision Tree Classification Light GBM Report.

Try	Precision	Recall	F1-Score	Support
0	0.98	1.00	0.99	1703
1	1.00	0.96	0.98	697
accuracy = 98.75%				
macro avg	0.99	0.98	0.98	2400
weighted avg	0.99	0.99	0.99	2400

4.7.1.2 Catboost classifier results

In the decision tree classifier of the CatBoost algorithm, the best obtained accuracy is (98.12%). The confusion matrix diagram shown in Figure (4.4) and the classification report shown in Table (4.4).

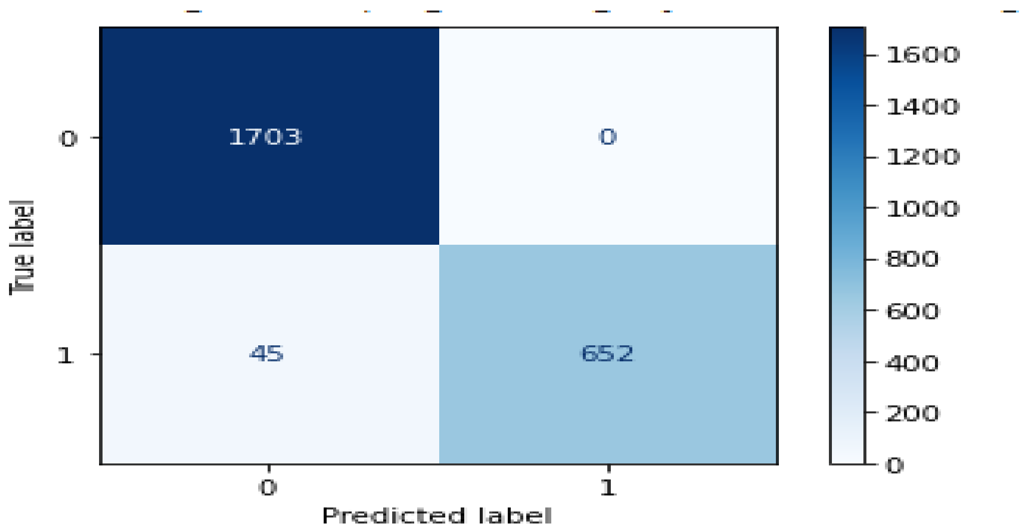

Figure 4.4: Decision Tree Classifier Confusion Matrix Catboost Algorithm.

Table 4.4: Decision Tree Classification CatBoost Report.

Try	Precision	Recall	F1-Score	Support
0	0.97	1.00	0.99	1703
1	1.00	0.94	0.97	697
Accuracy = 98.12				
Macro Avg	0.99	0.97	0.98	2400
Weighted Avg	0.98	0.98	0.98	2400

4.7.1.3 MLP classifier results

In the MLP neural network classifier, the best obtained accuracy is (91.08%). The confusion matrix diagram shown in Figure (4.5) and the classification report shown in Table (4.5).

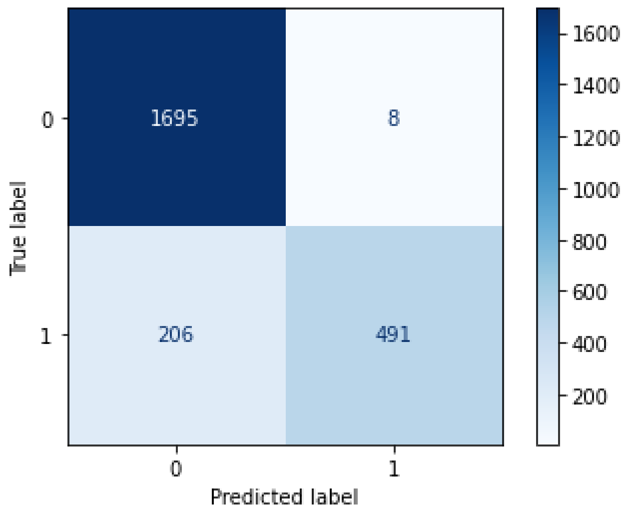

Figure 4.5: Neural Network Classifier Confusion Matrix MLP Algorithm.

Table 4.5: Neural Network Classification MLP Report.

Try	Precision	Recall	F1-Score	Support
0	0.89	1.00	0.94	1703
1	0.98	0.70	0.82	697
Accuracy = 91.08				
Macro avg	0.94	0.85	0.88	2400
Weighted avg	0.92	0.91	0.91	2400

4.7.2 Load Classification Unbalance Classes

The second step included the process of classifying the electrical loads of a city according to the losses incurred in those loads, and the results were as follows

4.7.2.1 LGBM classifier results

In the decision tree classifier of the LightGBM algorithm, the best accuracy obtained is (98.71%). The confusion matrix shown in Figure (4.6) and the classification report shown in the Table. (4.5).

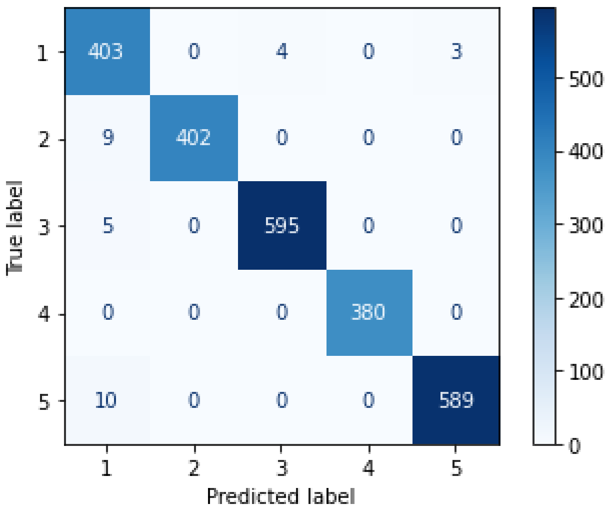

Figure 4.6: Decision Tree Classifier Confusion Matrix Light GBM Algorithm.

Table 4.6: Decision Tree Classification Light GBM Report.

Try	Precision	Recall	F1-Score	Support
1	0.94	0.98	0.96	410
2	1.00	0.98	0.99	411
3	0.99	0.99	0.99	600
4	1.00	1.00	1.00	380
5	0.99	0.98	0.99	599
Accuracy = .98.71				
Macro Avg	0.99	0.99	0.99	2400
Weighted Avg	0.99	0.99	0.99	2400

4.7.2.2 CatBoost classifier results

In the decision tree classifier of the CatBoost algorithm, the best obtained accuracy is (96.29%). The confusion matrix diagram shown in Figure (4.7) and the classification report shown in Table (4.7).

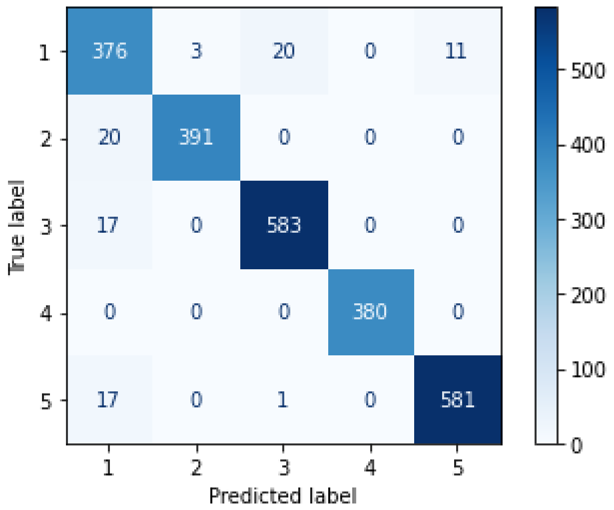

Figure 4.7: Decision Tree Classifier Confusion Matrix CatBoost Algorithm.

Table 4.7: Decision Tree Classification CatBoost Report.

Try	Precision	Recall	F1-Score	Support
1	0.87	0.92	0.90	410
2	0.99	0.95	0.97	411
3	0.97	0.97	0.97	600
4	1.00	1.00	1.00	380
5	0.98	0.97	0.98	599
Accuracy = 96.29				
Macro Avg	0.96	0.96	0.96	2400
Weighted Avg	0.96	0.96	0.96	2400

4.7.2.3 MLP classifier results

In the MLP neural network classifier, the best obtained accuracy is (74.17%). The confusion matrix diagram shown in Figure (4.7) and the classification report shown in Table (4.7).

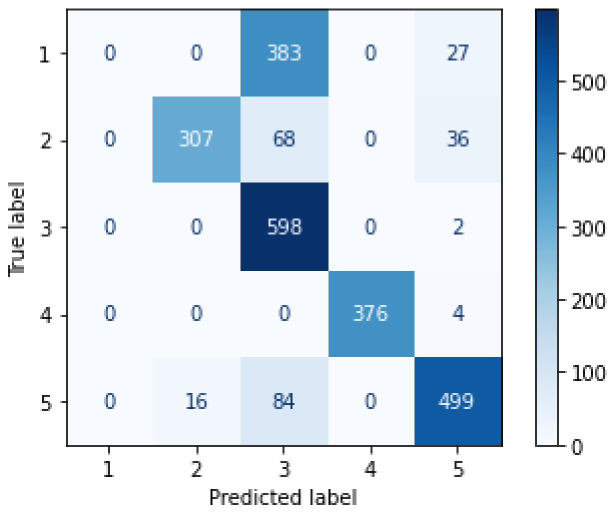

Figure 4.8: Neural Network Classifier Confusion Matrix MLP Algorithm.

Table 4.8: Neural Network Classification MLP Report.

Try	Precision	Recall	F1-Score	Support
1	0.00	0.00	0.00	410
2	0.95	0.75	0.84	411
3	0.53	1.00	0.69	600
4	1.00	0.99	0.99	380
5	0.88	0.83	0.86	599
Accuracy = 74.17				
Macro Avg	0.67	0.71	0.68	2400
Weighted Avg	0.67	0.74	0.69	2400

4.7.3 Predict the Type of Electrical Loads as (Inductive/Capacitive)

The third step included the process of predicting the type of loads for each house separately, whether the load is inductive or capacitive at the level of one house, and the following results were obtained.

4.7.3.1 LGBM prediction results

In the decision tree prediction of the LightGBM algorithm, the best obtained accuracy is (97.75%). The confusion matrix diagram shown in Figure (4.8) and the classification report shown in Table (4.8)

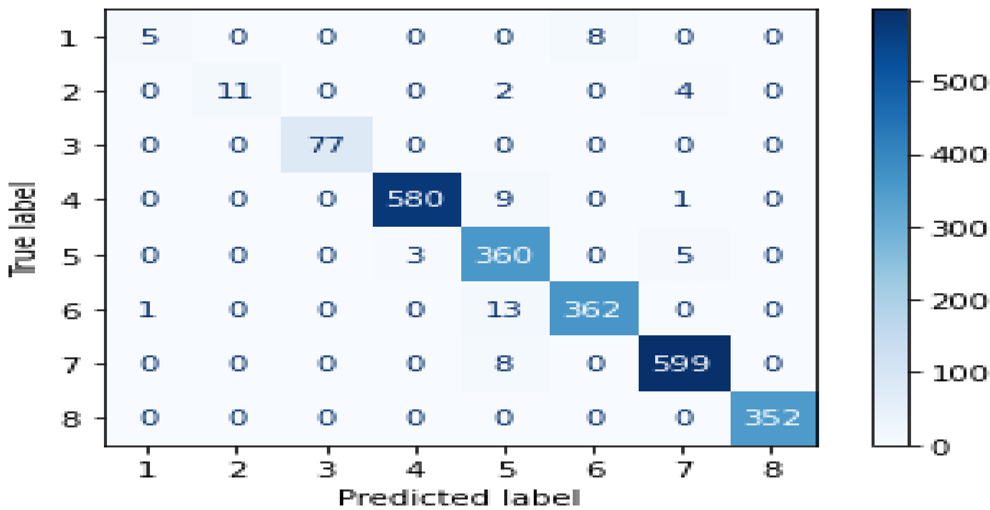

Figure 4.9: Decision Tree Prediction Confusion Matrix Light GBM Algorithm.

Table 4.9: Decision Tree Prediction Light GBM Report.

Try	Precision	Recall	F1-Score	Support
1	0.83	0.38	0.53	13
2	1.00	0.65	0.79	17
3	1.00	1.00	1.00	77
4	0.99	0.98	0.99	590
5	0.92	0.98	0.95	368
6	0.98	0.96	0.97	376
7	0.98	0.99	0.99	607
8	1.00	1.00	1.00	352
Accuracy = 97.75%				
Macro Avg	0.96	0.87	0.90	2400
Weighted Avg	0.98	0.98	0.98	2400

4.7.3.2 CatBoost prediction results

In the decision tree prediction of the algorithm, CatBoost, the best obtained accuracy is (95.71%). The confusion matrix diagram shown in Figure (4.9) and the classification report shown in Table (4.9).

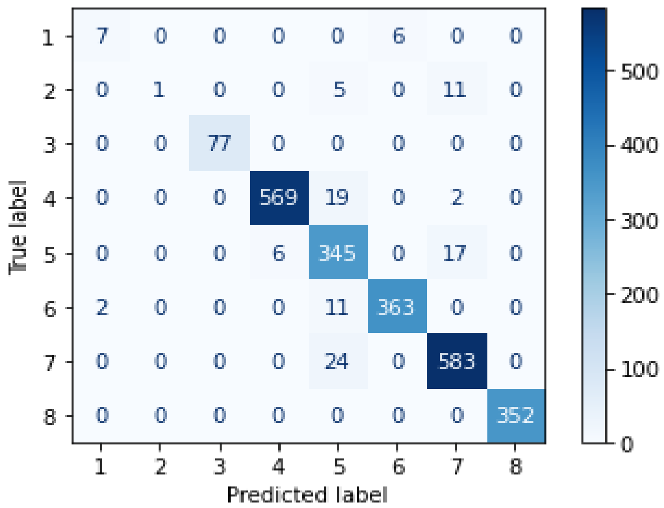

Figure 4.10: Decision Tree Prediction Confusion Matrix Catboost Algorithm.

Table 4.10: Decision Tree Prediction Lightgbm Report.

Try	Precision	recall	f1-score	Support
1	0.78	0.54	0.64	13
2	1.00	0.06	0.11	17
3	1.00	1.00	1.00	77
4	0.99	0.96	0.98	590
5	0.85	0.94	0.89	368
6	0.98	0.97	0.97	376
7	0.95	0.96	0.96	607
8	1.00	1.00	1.00	352
accuracy = 95.71%				
macro avg	0.94	0.80	0.82	2400
weighted avg	0.96	0.96	0.95	2400

4.7.3.3 MLP prediction results

In the decision tree prediction of the algorithm, MLP, the best obtained accuracy is (72.50%). The confusion matrix diagram shown in Figure (4.9) and the classification report shown in Table (4.10).

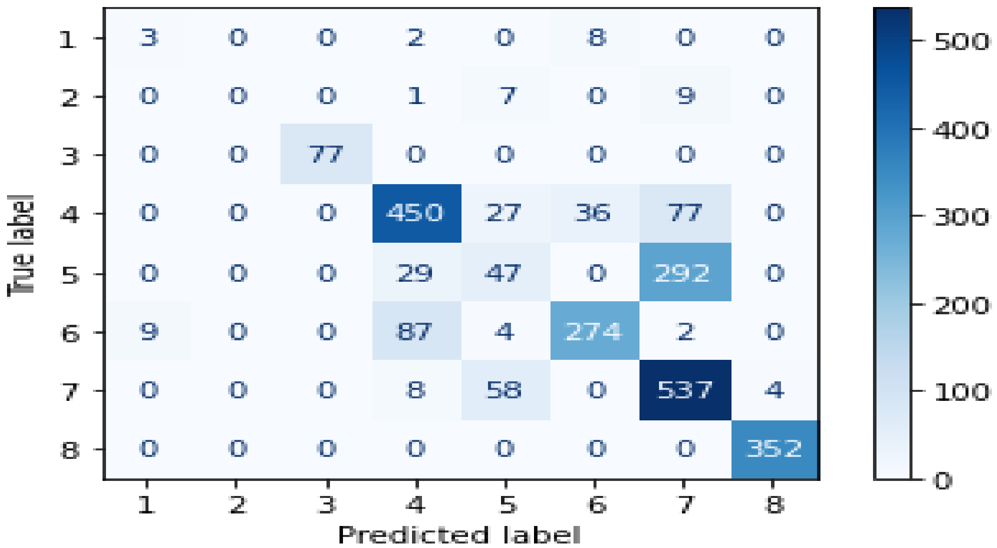

Figure 4.11: Decision Tree Prediction Confusion Matrix MLP Algorithm.

Table 4.11: Decision Tree Prediction MLP Report.

Try	Precision	Recall	f1-score	Support
1	0.25	0.23	0.24	13
2	0.00	0.00	0.00	17
3	1.00	1.00	1.00	77
4	0.78	0.76	0.77	590
5	0.33	0.13	0.18	368
6	0.86	0.73	0.79	376
7	0.59	0.88	0.70	607
8	0.99	1.00	0.99	352
accuracy = 72.50%				
macro avg	0.60	0.59	0.59	2400
weighted avg	0.70	0.72	0.70	2400

4.8 CONTROL SYSTEM IMPLEMENTATION USING PLC

4.8.1 PLC Design

After completing the classification and prediction process using machine learning and its algorithms, it was found that there are five types of load bias, whether the load is inductive or capacitive, and in order to achieve load balance and reach the pure Ohmic load, the device. PLC was used and programmed. Where three inputs of the PLC device (X1, X2, X3) were placed for the purpose of determining the type of load bias as shown in the Table (4.11).

Table 4.12: PLC Input Switches.

X2	X1	X0	Class
0	0	1	Class1
0	1	0	Claas2
0	1	1	Class3
1	0	0	Class4, Class5

PLC choose from DELTA company version (ES2). Using WPLSOFT software to configure PLC. So firstly after opening software as shown in Figure (4.11) the New project selection chooses and PLC device type selected.

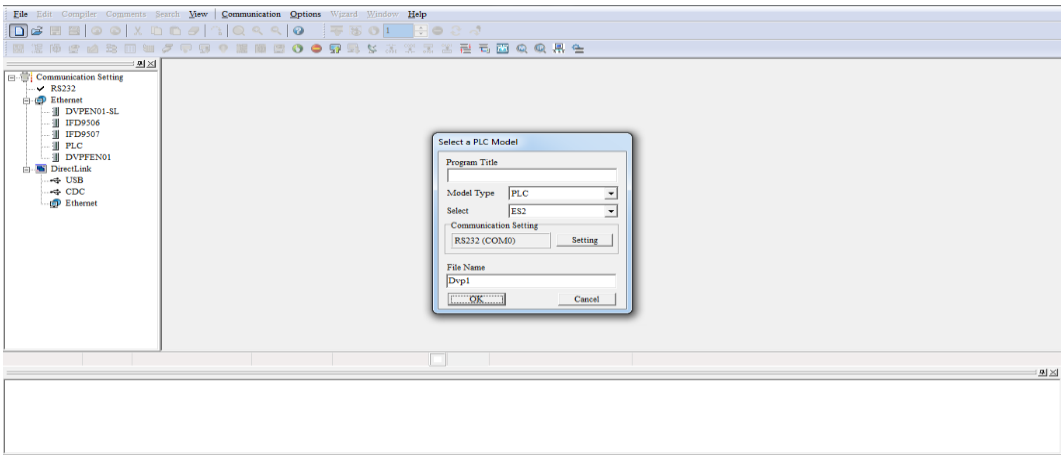

Figure 4.12: WPLSOFT Interface and New Project Creation.

With regard to the first half, twelve devices close in load of the inductive or capacitive type were connected according to the slope of the load, and their inputs were connected to the output ports of the PLC device. The devices were operated by ten devices, leaving two devices in the waiting state, where five Timers, and each timer has a period of time estimated at thirty minutes, as these timers work to put two devices in the waiting state periodically, for example the first and second, and then the third and fourth, and so on, so that the overload is always less than the real one, which leads to load balancing.

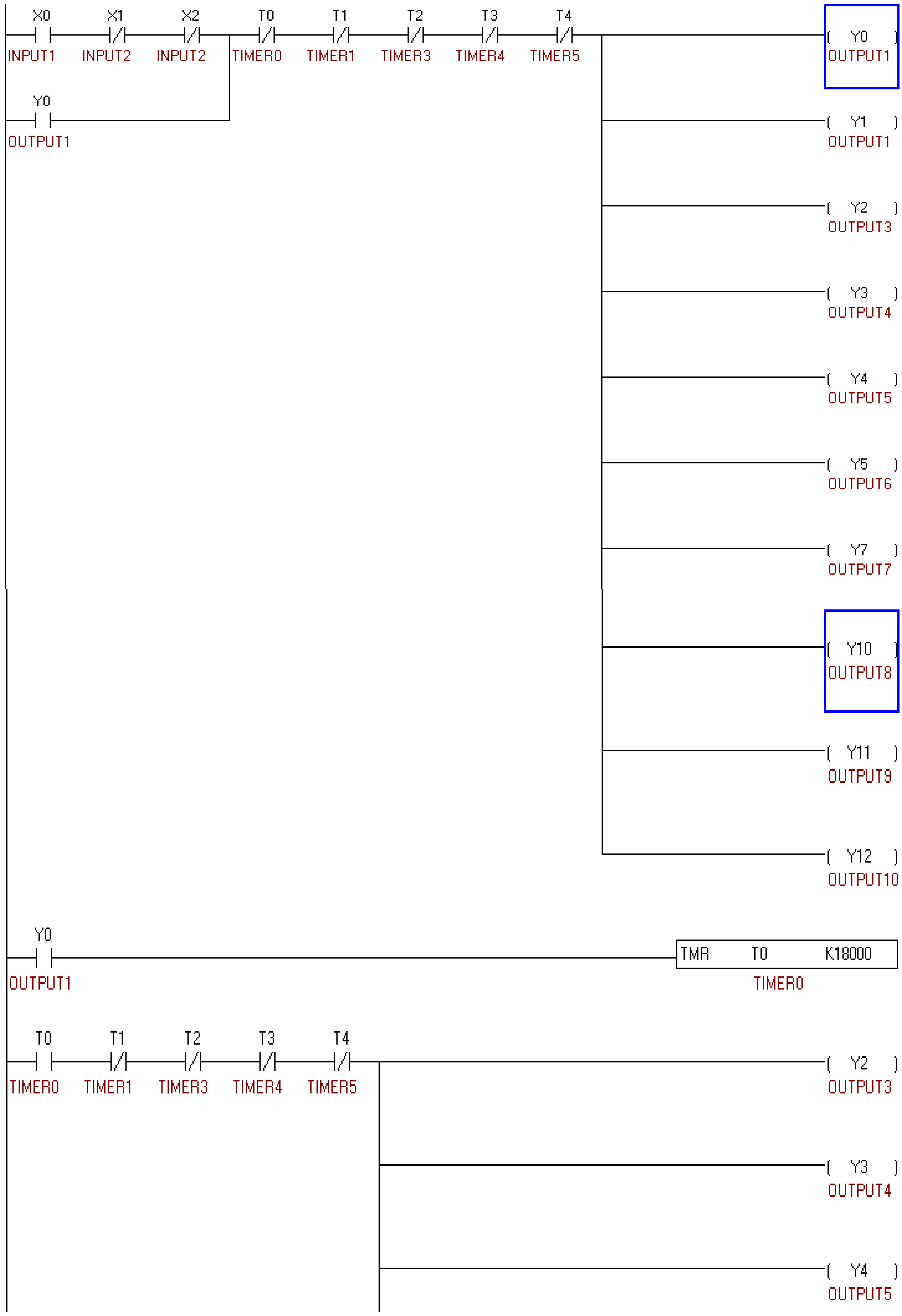

Figure 4.13: Case1 Ladder Diagram.

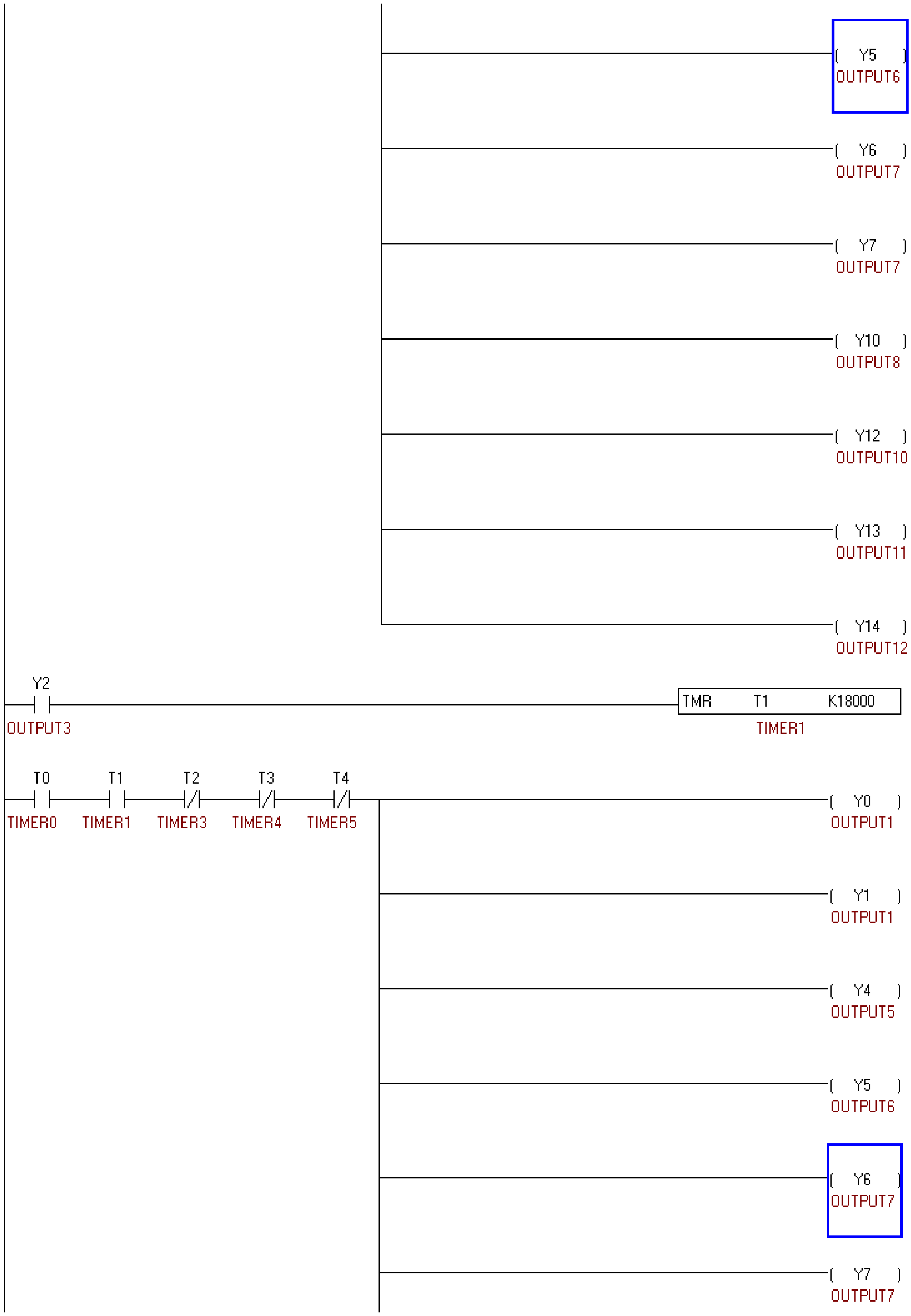

Figure 4.13: Case1 Ladder Diagram"figures continued".

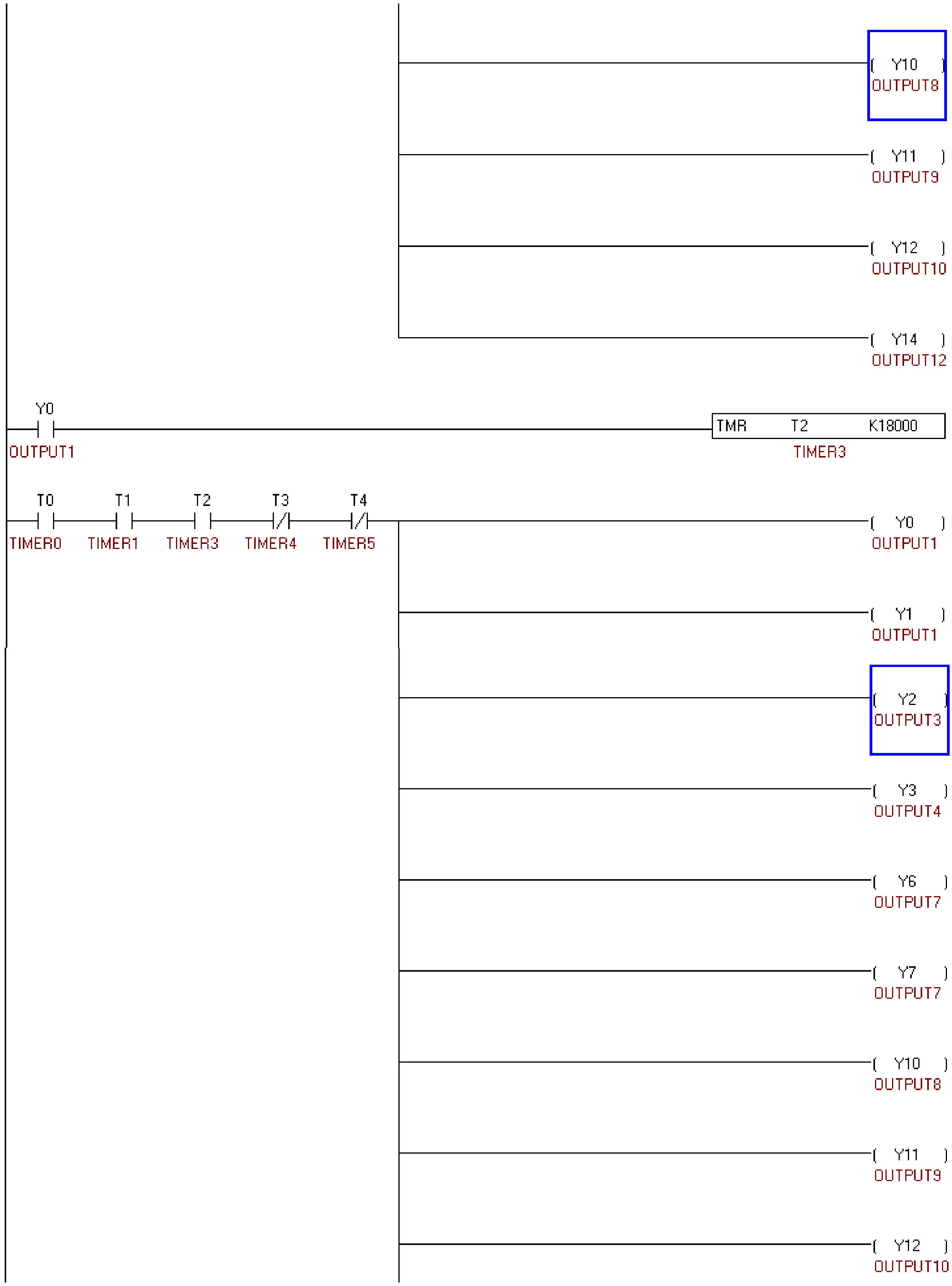

Figure 4.13: Case1 Ladder Diagram"figures continued".

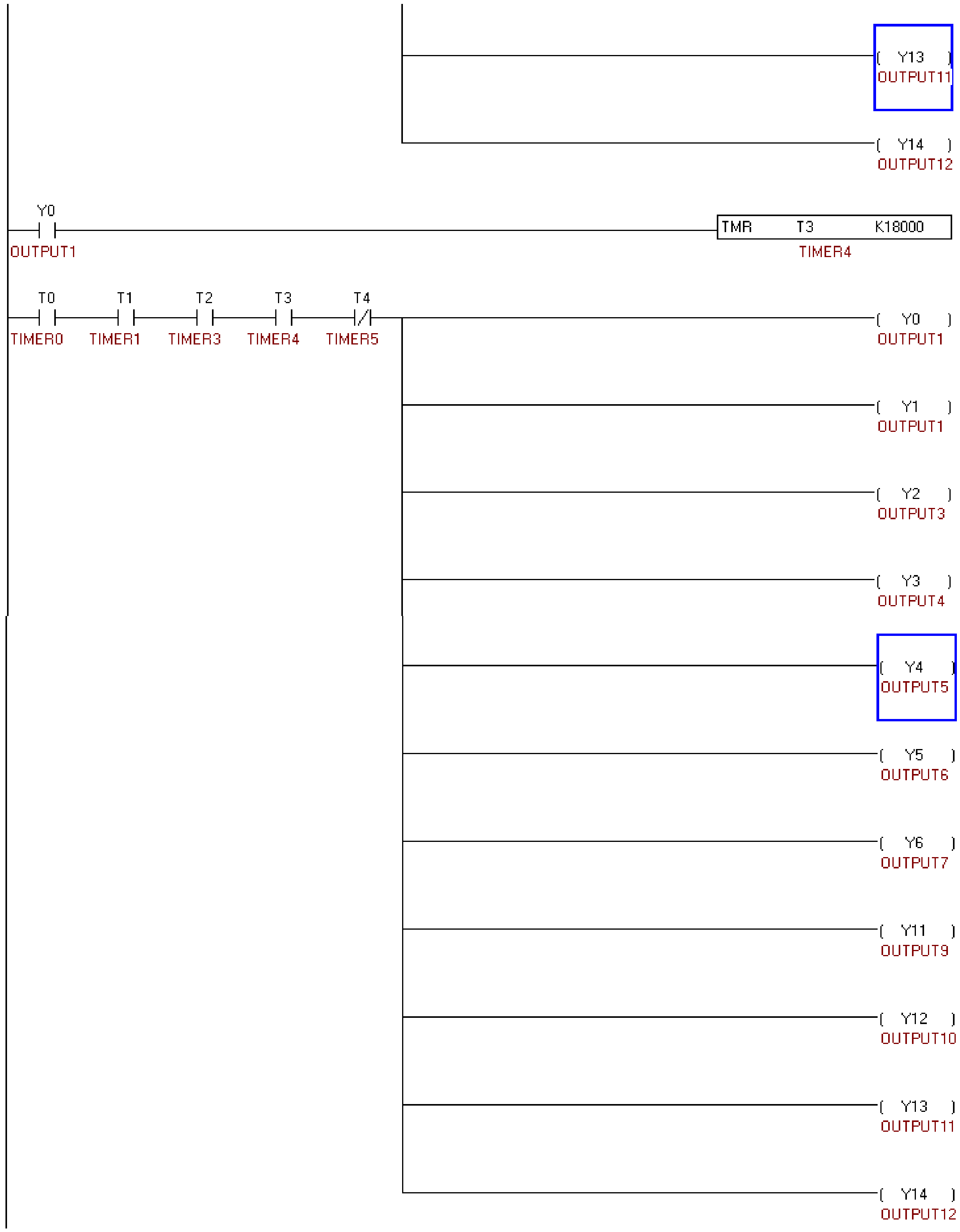

Figure 4.13: Case1 Ladder Diagram"figures continued".

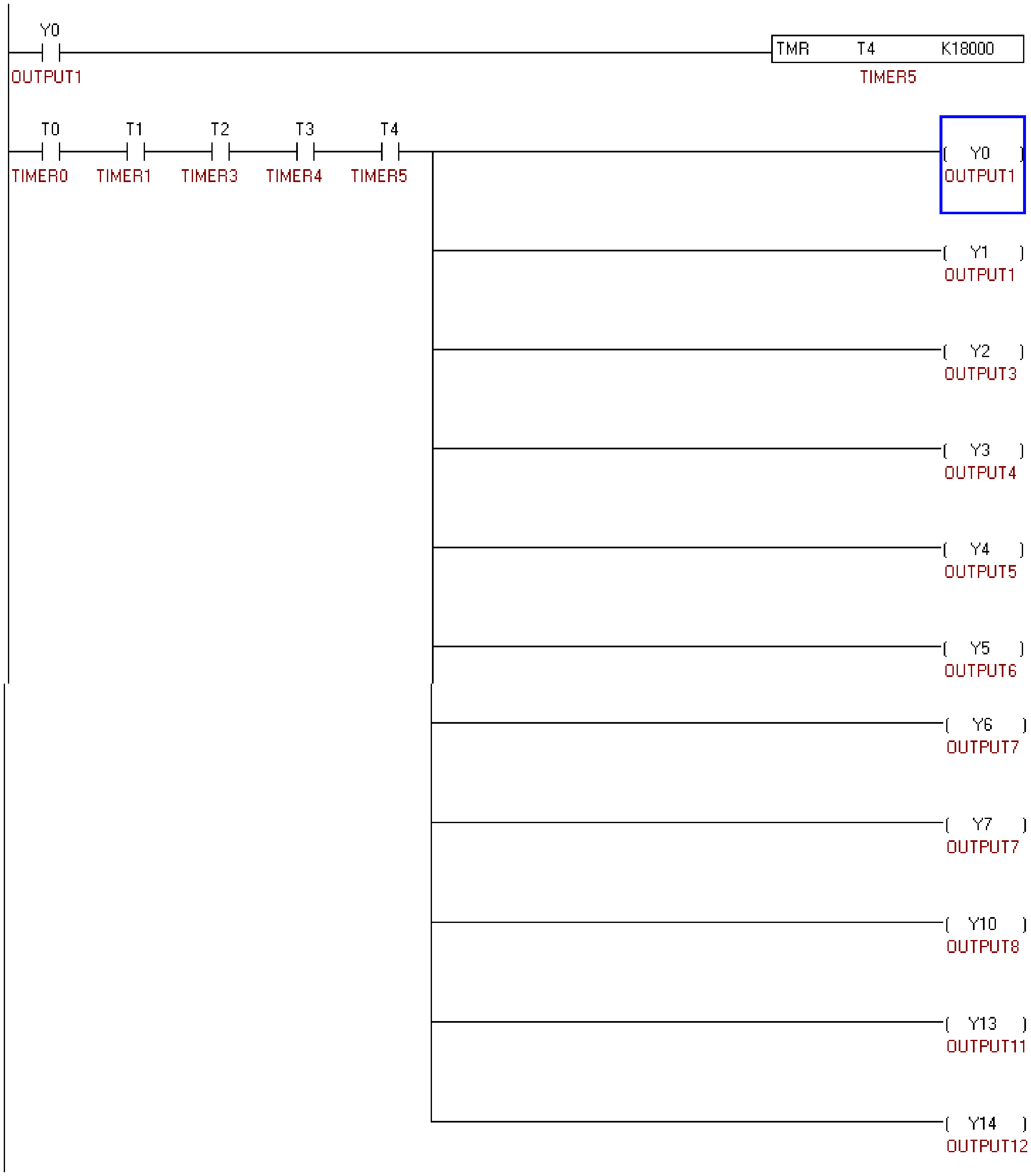

Figure 4.13: Case1 Ladder Diagram"figures continued".

As for the second case, three timers were set, as the load inclination in this case is greater than the first case, so the number of devices that will be in standby mode must be greater for the purpose of matching the load, and from here only three out of twelve devices were placed in the waiting state. Then, with the same idea of making the timers that were mentioned above, they are operated alternately, which leads to matching the load.

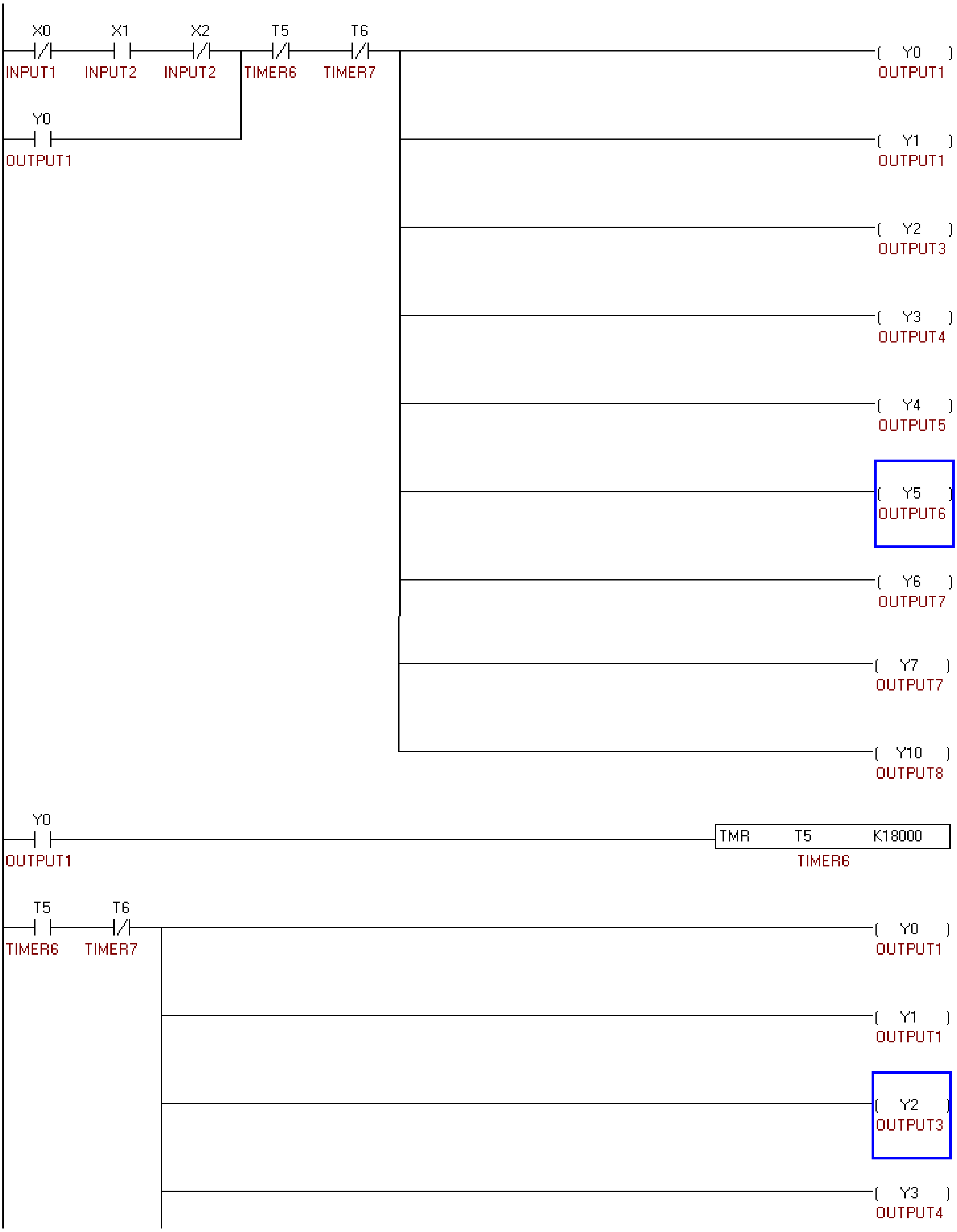

Figure 4.14: Case2 and Case3 Ladder Diagram.

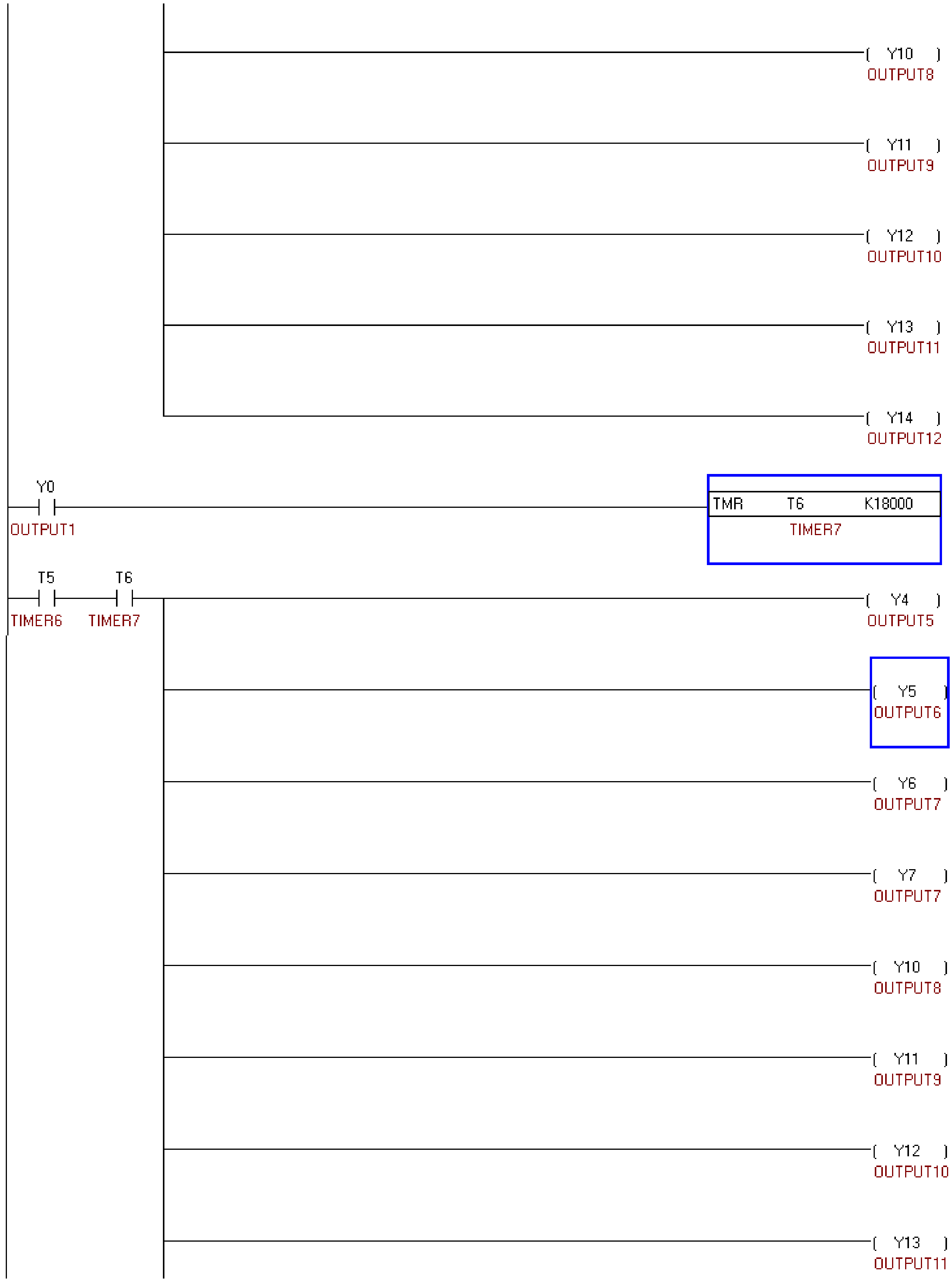

Figure 4.14: Case2 and Case3 Ladder Diagram"figures continued".

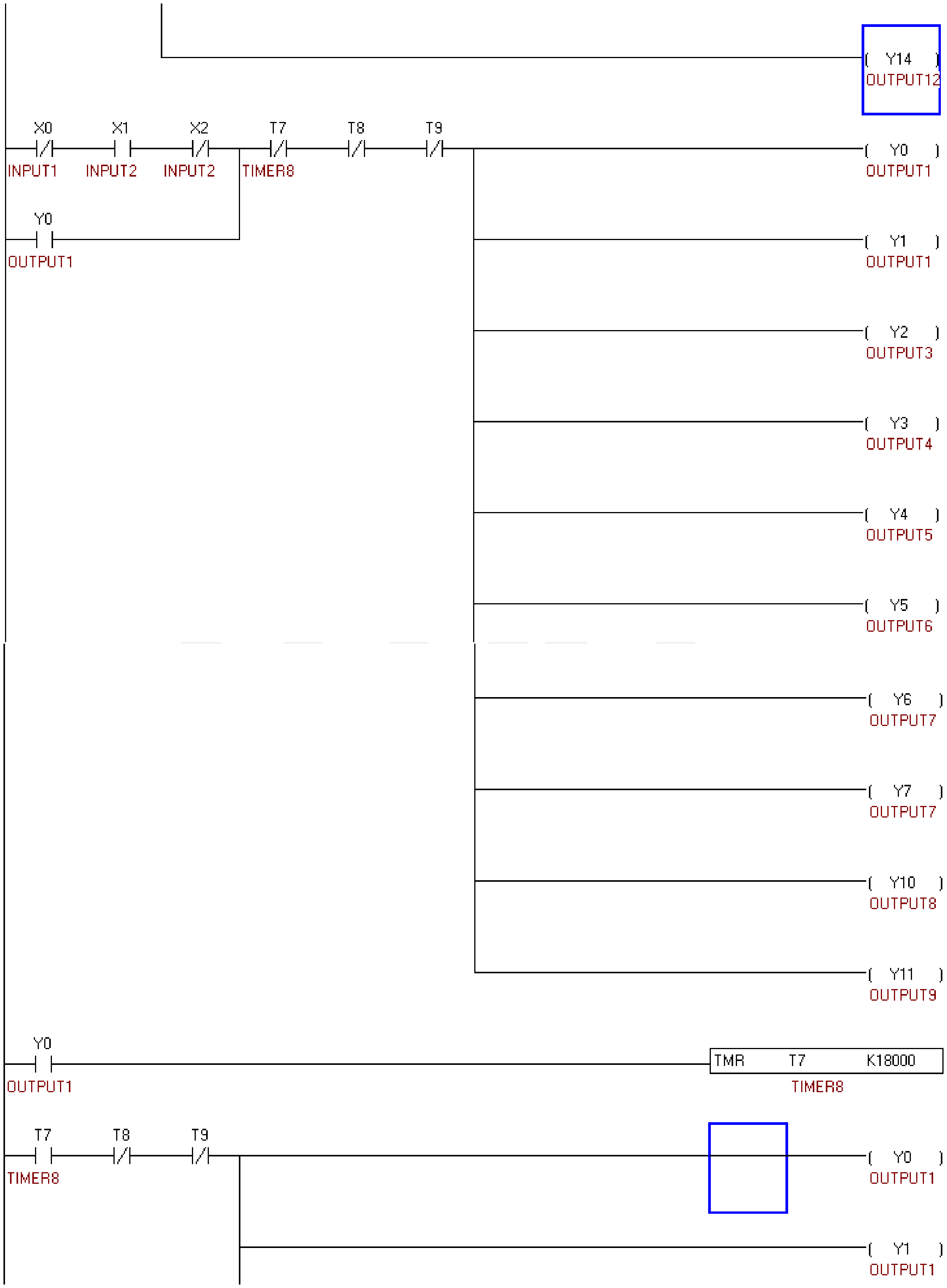

Figure 4.14: Case2 and Case3 Ladder Diagram "figures continued".

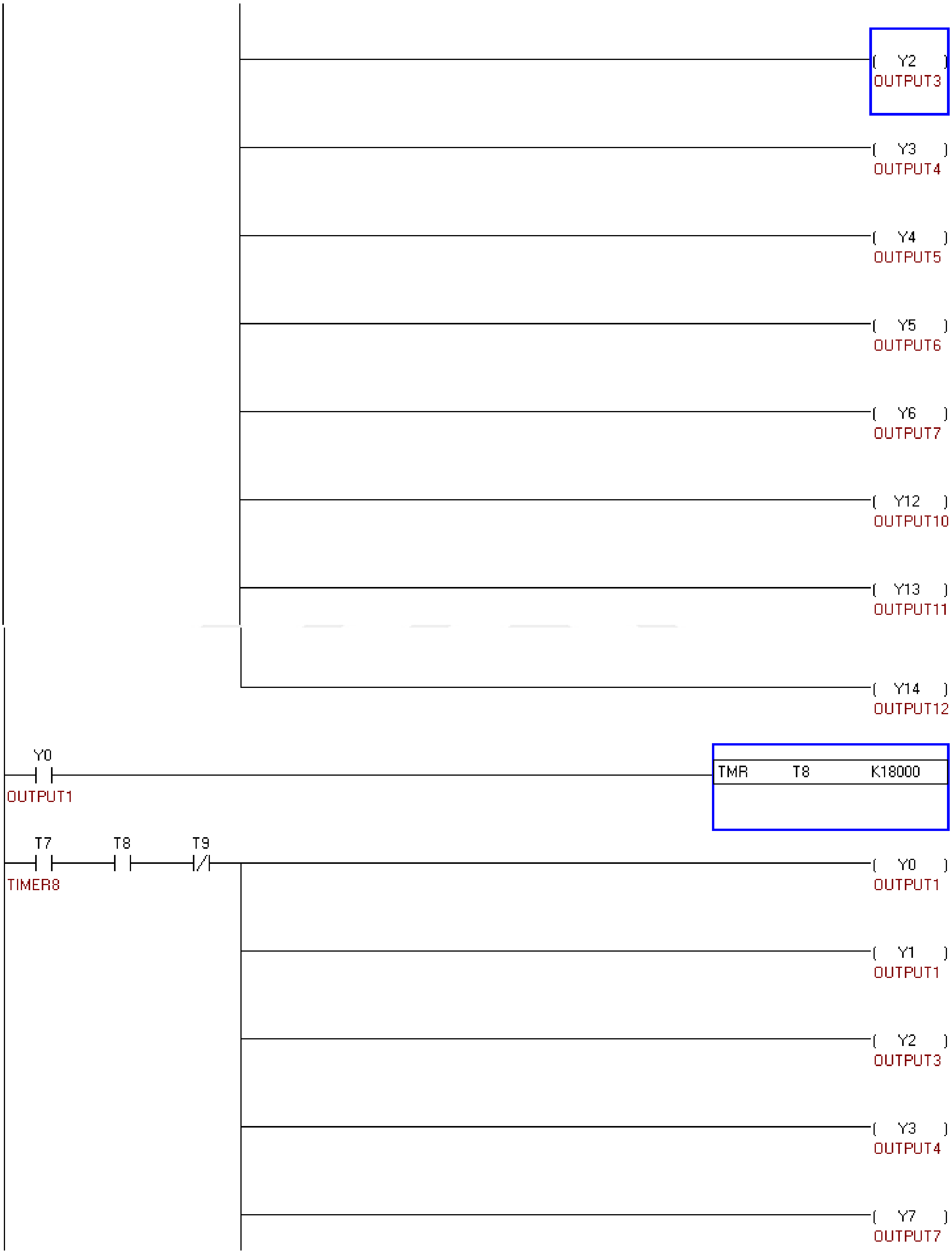

Figure 4.14: Case2 and Case3 Ladder Diagram "figures continued".

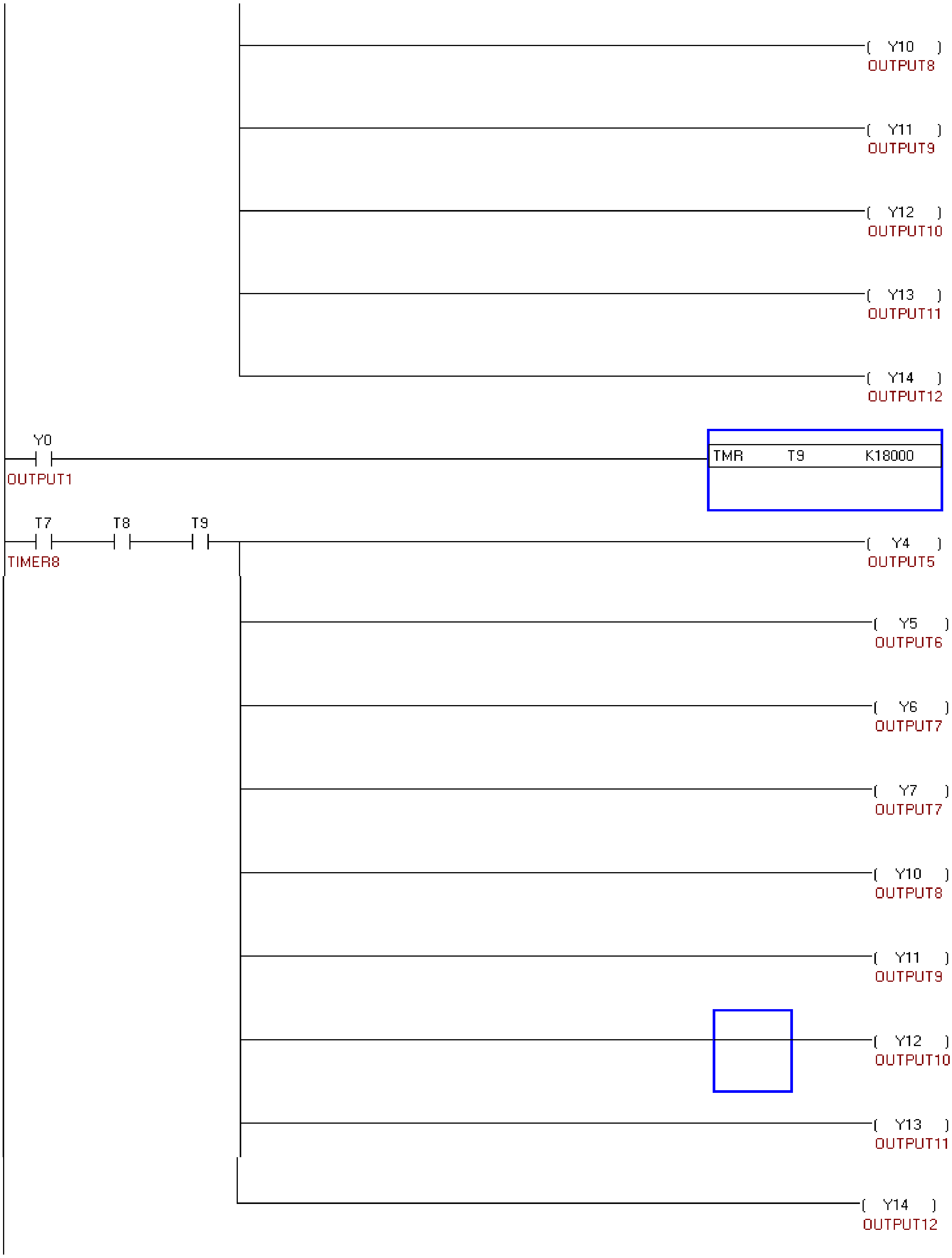

Figure 4.14: Case2 and Case3 Ladder Diagram "figures continued".

With regard to the third case, only two timers were used, and this is the switching off alternately each time by four devices out of twelve devices, and finally with regard to the fourth fourth and fifth cases, they were considered as one case due to the lack of items in case in the data set, where one timer was placed to operate alternately for six devices out of twelve devices He instruction list mode shown in Figure (4.15) Below.

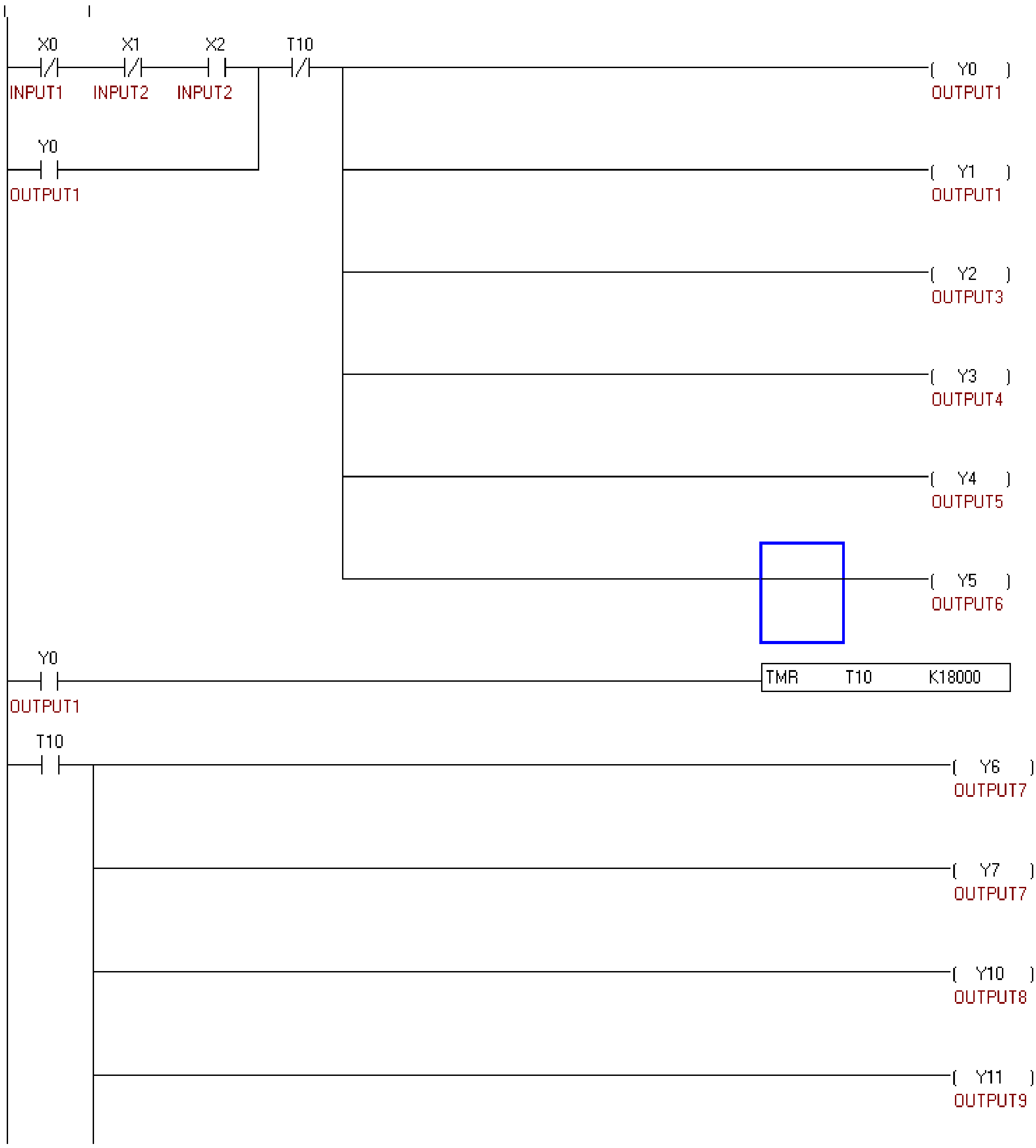

Figure 4.15: Case4 and Case5 Ladder Diagram.

Figure (4.16) Summarized the PLC Operation Mechanism.

000000	LD	X0	000035	OUT	Y10	
000001	ANI	X1	000036	OUT	Y12	
000002	ANI	X2	000037	OUT	Y13	
000003	OR	Y0	000038	OUT	Y14	
000004	ANI	T0	000039	LD	Y2	
000005	ANI	T1	000040	TMR	T1	K18000
000006	ANI	T2	000044	LD	T0	
000007	ANI	T3	000045	AND	T1	
000008	ANI	T4	000046	ANI	T2	
000009	OUT	Y0	000047	ANI	T3	
000010	OUT	Y1	000048	ANI	T4	
000011	OUT	Y2	000049	OUT	Y0	
000012	OUT	Y3	000050	OUT	Y1	
000013	OUT	Y4	000051	OUT	Y4	
000014	OUT	Y5	000052	OUT	Y5	
000015	OUT	Y7	000053	OUT	Y6	
000016	OUT	Y10	000054	OUT	Y7	
000017	OUT	Y11	000055	OUT	Y10	
000018	OUT	Y12	000056	OUT	Y11	
000019	LD	Y0	000057	OUT	Y12	
000020	TMR	T0	000058	OUT	Y14	
000024	LD	T0	000059	LD	Y0	
000025	ANI	T1	000060	TMR	T2	K18000
000026	ANI	T2	000064	LD	T0	
000027	ANI	T3	000065	AND	T1	
000028	ANI	T4	000066	AND	T2	
000029	OUT	Y2	000067	ANI	T3	
000030	OUT	Y3	000068	ANI	T4	
000031	OUT	Y4	000069	OUT	Y0	
000032	OUT	Y5	000070	OUT	Y1	
000033	OUT	Y6	000071	OUT	Y2	
000034	OUT	Y7	000072	OUT	Y3	

Figure 4.16: Sample of Instruction List Mode.

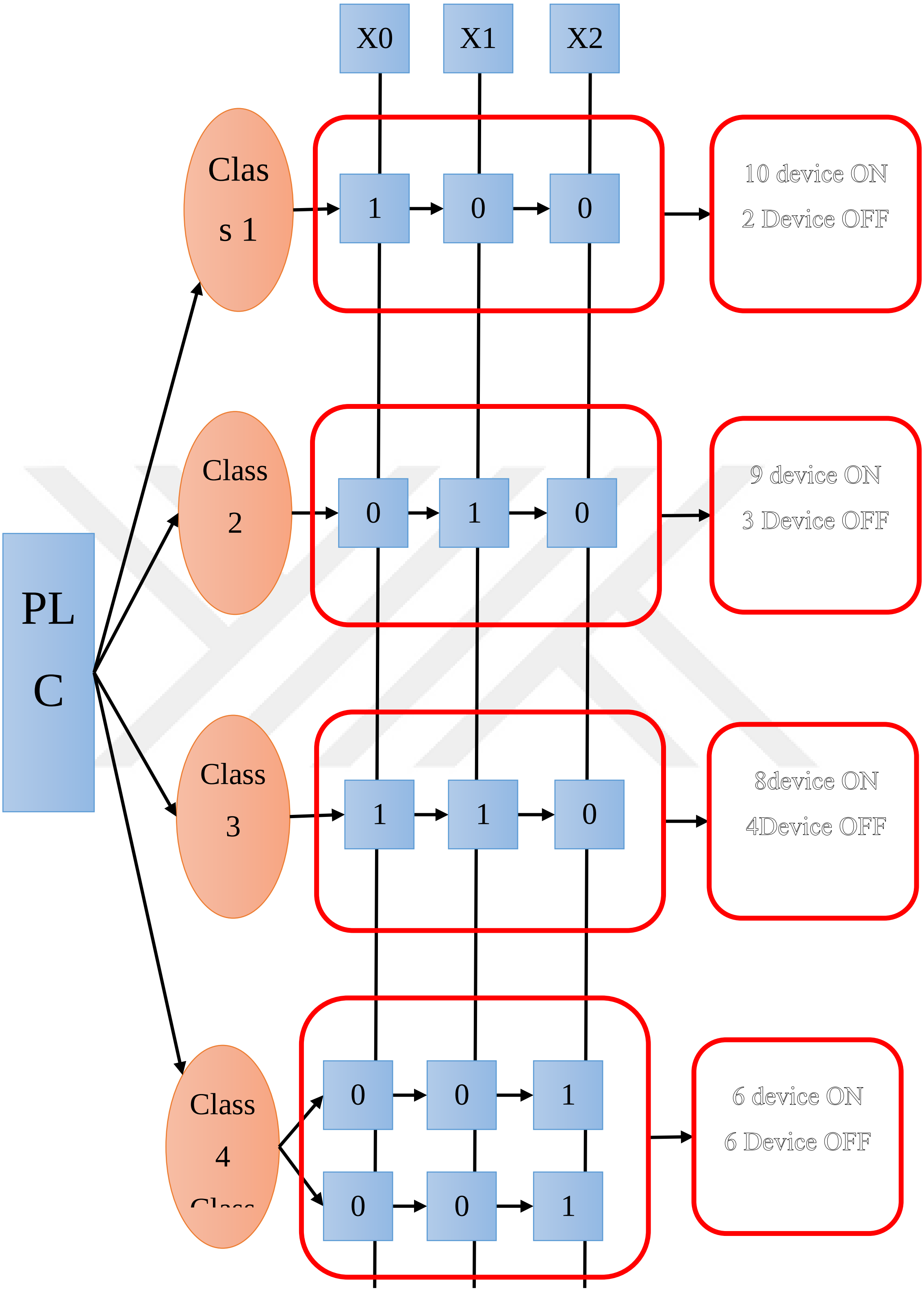

Figure 4.17: PLC Operation Mechanism.

4.9 COMPARISON TO RELATIVE WORKS

The study methods and results compare to relative work and displayed in Table (4.12).

Table 4.13: Relative Work Comparison.

Citation	Method	Results
Muralitharan, et.al, 2016	Load Balancing using MOEA	They decreases both the expense of power and the defer time in executing electrical apparatuses for the buyer.
Wasoontarajaroen.et.al., 2017	smart power meter monitors	Track consumption by mobile phone
ALRikabi.et.al.,2017	To alert devices to start working at 7:00 AM	Send an alert to the manager about the negligence of employees by leaving the devices on when they are not needed
Chen, et.al, 2017	Dynamic attachment scale method	Decrease the likelihood that the server cannot deal with unreasonable computational necessities
Zheng, et.al.2017	LSTM Prediction	Accuracy 82%
Li,et.al., 2017	XGBoost	RMSE 373.22
	KNN	RMSE 1451.57
Sicchar et. al., 2018	Microgrid Cluster Based	Fuzzy Logic Balancing
Present Study	Machine Learning and PLC controlling	CatBoost accuracy 96.81%
		LGBM accuracy 98.52%
		MLP accuracy 73.23%

5. CONCLUSION AND FUTURE WORKS

5.1 CONCLUSION

Electrical loads of all kinds have a great importance in playing roles to reduce or increase the consumption of electrical energy, as each type has characteristics, attributes, and a role in increasing and reducing energy consumption, in addition to the fact that the balance of these loads plays a very large role in consumption and energy efficiency, so technology had to play its role in this The field also where it had a share in solving most of the problems of consumption.

In light of the great development of technology in the field of electrical energy, the Internet of Things has proven this wonderful and very useful technology to be effective in many fields, especially electric energy, as this technology was characterized by its ability to manage electrical systems very effectively and deal with huge data wisely and link the physical parts of any system in a easy way. As for machine learning, it has added a very big qualitative leap in the field of electric energy, as machine learning algorithms impressed us with their great ability to help support and provide electric energy, increase system efficiency, reduce electric energy consumption, and support smart grids.

Machine learning depends mainly on classification and prediction, as its algorithms have the ability to classify very large data and are distinguished by their speed in providing results that have high efficiency and great accuracy. Our proposed system included a mixture of techniques to help classify electrical appliances according to their best consumption of electrical energy. The system was built on Important and sequential steps, due to the importance of finding a solution to load losses that lead to large amounts of energy without benefiting from it, but on the contrary help in energy shortage and reduce its efficiency. LightGBM and the mlp algorithm from Neural Network We have used in our system the Python programming language because it is one of the most important programming languages and has very important advantages such as its ease and containing a set of uncomplicated sentences as it contains sentences written in English and most importantly it is a freely available language as it does not need For the money to get it and its updates are also available, this language has many libraries and a large pal In addition to its high speed in the development of all applications, in addition to the Python language, the Google Colab cloud service helped us with this wonderful platform that helped us edit and download Python codes

This platform also does not need high technical specifications for the computer used in programming and does not need a very large memory

as a first step, we classified the electrical load losses for a city according to certain percentages, and as a second step, we classified the loads according to their type, and then in the third step, we predicted the type of load for each house separately, as these algorithms had the ability to predict and classify within hours only, while in time In itself, this process takes a large human hand, and it may take several months to complete it Then we worked on entering the results into a simulation system through the model logic unit. This proposed system saved us a large amount of effort from the manpower, as it works without the intervention of the human hand. We have obtained excellent results that are more than the expected results. Such systems are recommended for The great contribution in reducing the consumption of electrical energy. If we can say that in this system we focused on important points that have their role and importance in reducing electrical consumption of energy and balancing the electrical load in the electrical network, which are:

a) Calculating the losses in electrical loads at the city level.

b) Classification of the type of electrical loads, are they capacitive loads or are they inductive loads.

c) Predicting the type of load inductive or inductive, at the level of one house.

d) Load balancing work through the PLC logical processing unit.

5.2 FUTURE WORKS

a) Designing a special system for calculating heat loss in loads.

b) Designing a system that performs balancing capacitive loads by specifying their entry into the network in increasing stages and specifying the appropriate time and place to add the capacitor and the places to install it, i.e. making a system in which there are multiple entries with appropriate times and values for the purpose of achieving the greatest possible balance and minimizing losses.

c) Using other algorithms with more devices to make the project more comprehensive.

d) Designing a project that calculates the cost of one house by calculating the loads.

e) Designing an application for Android devices that informs the user of the quantity and type of loads for his home so that he is constantly informed.

REFERENCES

[1] N. A. Hussien, A. A. D. Al Magsoosi, A. A. D. Al Magsoosi, H. T. Salim AlRikabi, and F. T. Abed, "Monitoring the Consumption of Electrical Energy Based on the Internet of Things Applications," *Int. J. Interact. Mob. Technol.*, vol. 15, no. 7, pp. 17–29, 2021, doi: 10.3991/ijim.v15i07.20183.

[2] E. Sulaiman *et al.*, "a Study on Electric Energy Usage," *Energy Environ.*, no. c, 2007.

[3] A. A. Azlina, M. Kamaludin, E. S. Z. E. Abdullah, and A. Radam, "Factors influencing household end-use electricity demand in Malaysia," *Adv. Sci. Lett.*, vol. 22, no. 12, pp. 4120–4123, 2016, doi: 10.1166/asl.2016.8189.

[4] V. L. Chen, M. A. Delmas, S. L. Locke, and A. Singh, "Information strategies for energy conservation: A field experiment in India," *Energy Econ.*, vol. 68, pp. 215–227, 2017, doi: 10.1016/j.eneco.2017.09.004.

[5] E. Gelenbe and Y. Caseau, "The impact of information technology on energy consumption and carbon emissions," *Ubiquity*, vol. 2015, no. June, pp. 1–15, 2015, doi: 10.1145/2755977.

[6] A. De Almeida, P. Fonseca, B. Schlomann, and N. Feilberg, "Characterization of the household electricity consumption in the EU, potential energy savings and specific policy recommendations," *Energy Build.*, vol. 43, no. 8, pp. 1884–1894, 2011, doi: 10.1016/j.enbuild.2011.03.027.

[7] I. Vassileva, *Characterization of household energy consumption in Sweden*. 2012.

[8] Z. Yaoxue, R. Ju, L. Jiagang, X. Chugui, G. Hui, and and L. Yaping, "Chinese J of Electronics - 2017 - Zhang - A Survey on Emerging Computing Paradigms for Big Data.pdf," *Chinese J. Electron.*, vol. 26, no. 1, 2017.

[9] M. Duggan, "Online harassment 2017," *Pew Res. Cent.*, pp. 1–85, 2017.

[10] S. Waechter, B. Sütterlin, and M. Siegrist, "Desired and undesired effects of energy labels - An eye-tracking study," *PLoS One*, vol. 10, no. 7, pp. 1–26, 2015, doi: 10.1371/journal.pone.0134132.

[11] M. L. Tuballa and M. L. Abundo, “A review of the development of Smart Grid technologies,” *Renew. Sustain. Energy Rev.*, vol. 59, pp. 710–725, 2016, doi: 10.1016/j.rser.2016.01.011.

[12] S. C. Kim, P. Ray, and S. S. Reddy, “Features of smart grid technologies: An overview,” *ECTI Trans. Electr. Eng. Electron. Commun.*, vol. 17, no. 2, pp. 169–180, 2019, doi: 10.37936/ecti-eec.2019172.215478.

[13] A. Al-Qarafi *et al.*, “Artificial Jellyfish Optimization with Deep-Learning-Driven Decision Support System for Energy Management in Smart Cities,” *Appl. Sci.*, vol. 12, no. 15, pp. 1–17, 2022, doi: 10.3390/app12157457.

[14] D. A. J. L. Mr. M.Pratapa Raju, “IOT based Online Load Forecasting using Machine Learning Algorithms,” *Third Int. Conf. Comput. Netw. Commun.*, vol. 171, pp. 551–560, 2020.

[15] C. Te Lee, L. B. Chen, H. M. Chu, C. J. Hsieh, and W. C. Liang, “An Internet of Things (IoT)-Based Master-Slave Regionalized Intelligent LED-Light-Controlling System,” *Appl. Sci.*, vol. 12, no. 1, 2022, doi: 10.3390/app12010420.

[16] A. S. Shah, H. Nasir, M. Fayaz, A. Lajis, and A. and Shah, “A Review on Energy Consumption Optimization Techniques in IoT Based Smart Building Environments,” *Information*, vol. 70, no. 108, 2019.

[17] Y. S. R. Sakthivel, K. Muralitharan, “Multiobjective optimization technique for demand side management with load balancing approach in smart grid,” *Neurocomputing*, 2015.

[18] I. H. Al-Kharsan, A. F. Marhoon, and J. R. Mahmood, “Phase Swapping Load Balancing Algorithms, Comprehensive Survey Ibrahim,” *Iraqi J. Electr. Electron. Eng.*, vol. 15, no. 2, 2019.

[19] M. Güçyetmez and H. S. Farhan, “A Real-Time Invoice Based Smart Meter Design with Mobile Application,” *Int. J. Eng. Res. Dev.*, vol. 14, no. 2, 2022.

[20] H. T. S. AlRikabi, K. W. Nasser, and A. H. M.Alaidi, “The Application of Wireless Communication in IOT for Saving Electrical Energy.” pp. 152–160, 2019.

[21] S.-H. K. I. Shang-Liang Chen , Yun-Yao Chen, “CLB: A novel load balancing architecture and algorithm for cloud services,” *Comput. Electr. Eng.*, vol. 1, no. 7, 2016.

[22] H. Zheng, J. Yuan, and and L. Chen, “Short-Term Load Forecasting Using EMD-LSTM Neural Networks with a Xgboost Algorithm for Feature Importance Evaluation,” *Energy Environ.*, vol. 10, no. 1168, 2017.

[23] G. Li, W. Li, X. Tian, and Y. Che, “Short-Term Electricity Load Forecasting Based on the XGBoost Algorithm,” *Hansi*, vol. 7, no. 4274–285, 2017.

[24] and J. H. Naser Hossein Motlagh, Siavash H. Khajavi, Alireza Jaribion, “An IoT-based Automation System for Older Homes: A Use Case for Lighting System Naser.” 2018.

[25] J. R. S. Jose R. Sicchar , Carlos T. Da Costa Jr. and Raimundo C. Oliveira and Werbeston D. Oliveira, “A Load-Balance System Design of Microgrid Cluster Based on Hierarchical Petri Nets,” *Energies*, vol. 11, no. 3245, 2018.

[26] F. J. A. A. P. V´ıctor M. L´opez-Mart´ın, “Power Quality Enhancement in Residential Smart Grids through Power Factor Correction Stages,” *IEEE Trans. Ind. Electron.*, 2018.

[27] M. S. and E. H. Jihoon Moon, Yongsung Kim, “Hybrid Short-Term Load Forecasting Scheme Using Random Forest and Multilayer Perceptron,” *Energies*, vol. 11, no. 3283, 2018.

[28] L. P. J. G. Georgios Mylonas, Dimitrios Amaxilatis, Stelios Tsampas, “A Methodology for Saving Energy in Educational Buildings Using an IoT Infrastructure,” *arXiv:1907.07760v1*, 2019.

[29] A. K. G. R. Johari, “IOT based Electrical Device Surveillance and Control System,” *IEEE*, vol. 978, no. 1, 2019.

[30] J. M. A.-P. and M. Á. Pérez-Juárez, “An Insight of Deep Learning Based Demand Forecasting in Smart Grids,” *sensors*, vol. 23, no. 1467, 2023.

[31] J. B. SJ, K. D. K.Sameera Banu, K.Divyashree, K.Sameera Banu, and S. B. Priyanka M Joshi, "POWER THEFT IDENTIFICATION SYSTEM USING IOT," *Int. J. Creat. Res. Thoughts*, vol. 9, no. 5, 2021.

[32] tiu and Gheorghe Grigoras, , Bogdan-Constantin Neagu, Mihai Gavrilas,, Ion Tris and C. Bulac, "Optimal Phase Load Balancing in Low Voltage Distribution Networks Using a Smart Meter Data-Based Algorithm," *Mathematics*, vol. 8, no. 549, 2020.

[33] R. P. Shishir Muralidhara , Niharika Hegde, "An internet of things-based smart energy meter for monitoring device-level consumption of energy," *Comput. Electr. Eng.*, vol. 45, no. 7906, 2020.

[34] F. K. Amir Seyyedabbasi, "Decrease Electricity Consumption in Rooms with IoT Technology," *Int. J. Inf. Syst. Comput. Sci.*, vol. 9, no. 1, 2020.

[35] V. V. · R. Deshmukh, "Electric power load forecasting on a 33/11 kV substation using artificial neural networks," *SN Appl. Sci.*, vol. 2, no. 855, 2020.

[36] J. W. Lingyan Cao, Yongkui Li , Jiansong Zhang Yi Jiang , Yilong Han, "Electrical load prediction ofhealthcare buildings through single and ensemble learning," *Energy Reports*, vol. 6, pp. 2751–2767, 2020.

[37] S. K. Nabeela Awan1, M. K. I. Rahmani, R. A. , Muhammad Tahir , Nur Alam MD4, and and I. Ullah6, "Machine Learning-Enabled Power Scheduling in IoT-Based Smart Cities," *Comput. Mater. Contin.*, vol. 67, no. 2, 2021.

[38] J. E. S. Tatiana Tucunduva Philippi Cortese, Jairo Filho Sousa de Almeida and A. L. and T. Yigitcanlar, "Understanding Sustainable Energy in the Context of Smart Cities: A PRISMA Review," *Energies*, vol. 15, no. 2382, 2022.

[39] H. T. S. A. R. Nabaa Ali Jasim, "Design and Implementation a Smart System for Monitoring the Electrical Energy based on the Internet of Things," *Wasit J. Eng. Sci.*, vol. 10, no. 2, 2022.

[40] N. D. Bhagwatkar, S. G. Maskar, and P. S. Lewarkar, "Energy Monitoring and Conservation using Smart Plug," *Int. J. Res. Eng. Sci.*, vol. 10, no. 5, pp. 66–71, 2022.

[41] P. D. Nitin Thapliyal, “Load Balancing in Cloud Computing Based on Honey Bee Foraging Behavior and Load Balance Min-Min Scheduling Algorithm,” *Int. J. Electr. Electr. Electron. Res.*, vol. 10, no. 1, pp. 1–6, 2022.

[42] L. C. · S. Y. · F. C. · Z. M. · N. L. J. Qin, “A survey on application of machine learning for Internet of Things,” *Int. J. Mach. Learn. Cybern.*, vol. 9, no. 1399, 2018.

[43] and H. H. Hasan Erteza Gelani , Mashood Nasir , Faizan Dastgeer, “Efficiency Comparison of Alternating Current (AC) and Direct Current (DC) Distribution System at Residential Level with Load Characterization and Daily Load Variation,” *Pakistan Acad. Sci.*, vol. 54, no. 2, pp. 111–118, 2017.

[44] C. Löffler, “Fundamentals of alternating current technology.” pp. 1–38, 2010.

[45] D. P. Mishra, “Fundamentals of Rocket Propulsion,” pp. 1–483, 2017.

[46] B. J. Ware, “POWER FACTOR,” pp. 22–24, 2006.

[47] P. Riese, “Manual of Power Factor Correction,” pp. 1–72, 2014.

[48] N. T.-U. and S.-B. R. Saeid Gholami Farkoush, Chang-Hwan Kim, Ho-Chul Jung, Sanghyuk Lee, “Power Factor Improvement of Distribution System with EV Chargers based on SMC Method for SVC,” *J Electr Eng Technol*, vol. 12, no. 4, pp. 1340–1347, 2017.

[49] M. A. R. Md Abdul Mubeen, “AN OVERVIEW OF MACHINE LEARNING TECHNOLOGIES AND THEIR USE IN E LEARNING,” *J. Eng. Sci.*, vol. 13, no. 9, 2022.

[50] M. H. Alsharif, A. H. Kelechi, K. Yahya, and S. A. Chaudhry, “Machine learning algorithms for smart data analysis in internet of things environment: Taxonomies and research trends,” *Symmetry (Basel).*, vol. 12, no. 1, pp. 1–17, 2020, doi: 10.3390/SYM12010088.

[51] P. P. Harsh H. Patel, “Study and Analysis of Decision Tree Based Classification Algorithms,” *Int. J. Comput. Sci. Eng.*, vol. 610, 2018.

[52] I. D. MIENYE and Y. SUN, “A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects,” *IEEE Access*, vol. 10, 2022.

[53] G. Ke *et al.*, “LightGBM: A highly efficient gradient boosting decision tree,” *Adv. Neural Inf. Process. Syst.*, vol. 2017-December, no. Nips, pp. 3147–3155, 2017.

[54] R. Wang *et al.*, “Power System Transient Stability Assessment Based on Bayesian Optimized LightGBM,” *IEEE Conf. Energy Internet Energy Syst. Integr.*, 2019.

[55] A. Bakhshi, N. Noman, Z. Chen, M. Zamani, and S. Chalup, “Fast Automatic Optimisation of CNN Architectures for Image Classification Using Genetic Algorithm.” 2019.

[56] Y. H. and Y. M. Jinze Song , Yuhao Li , Shuai Liu , Youming Xiong , Weixin Pang, “Comparison of Machine Learning Algorithms for Sand Production Prediction: An Example for a Gas-Hydrate-Bearing Sand Case,” *Energies*, vol. 15, no. 6509, 2022.

[57] and M. H. Yanjun Chen, Kun Liu Yuantao Xie, “Financial Trading Strategy System Based on Machine Learning Yanjun,” *Hindawi Math. Probl. Eng.*, pp. 1–13, 2020.

[58] D. A. McCarty, H. W. Kim, and H. K. Lee, “Evaluation of light gradient boosted machine learning technique in large scale land use and land cover classification,” *Environ. - MDPI*, vol. 7, no. 10, pp. 1–22, 2020, doi: 10.3390/environments7100084.

[59] C. and H. L. Lei Deng1,JuanPan1, Xiaojie Xu1, Wenyi Yang1, “PDRLGB: precise DNA-binding residue prediction using a light gradient boosting machine,” *BMCBioinformatics*, vol. 19, no. 19, 2018.

[60] A. A. T. A. S. J. MALEBARY, “An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine,” *IEEE Access*, vol. 8, 2020.

[61] A. G. Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, “CatBoost: unbiased boosting with categorical features.” 2018.

[62] Q.-J. Zhang, K. C. Gupta, and V. K. Devabhaktuni, “Artificial Neural Networks for RF and Microwave Design—From Theory to Practice,” *IEEE Trans. Microw. Theory Tech.*, vol. 51, no. 4, 2003.

[63] S. B. M. and I. R. P. M Eriyadi, “Automatic metal sorting conveyor machine based on Programmable Logic Controller,” *OP Conf. Ser. Mater. Sci. Eng. Pap.*, pp. 1–8, 2020.

[64] and Z. S. Adam Szcześniak, “Fast Designing Ladder Diagram of Programmable Logic Controller for a Technological Process,” *INTL J. Electron. Telecommun.*, vol. 68, no. 4, pp. 709–714, 2022.

[65] M. M. Lashin, “Different Applications of Programmable Logic Controller (PLC),” *Int. J. Comput. Sci. Eng. Inf. Technol.*, vol. 4, no. 1, pp. 27–32, 2014, doi: 10.5121/ijcseit.2014.4103.

[66] P. Naik and G. Naik, *Conceptualizing Python in Google COLAB*, no. January. 2022.

[67] G. Colaboratory, G. Colabatory, G. Colaboratory, and T. Python, “Google Colab Getting Started Guide,” pp. 3–6, 2021.

[68] Ž. Vujović, “Classification Model Evaluation Metrics,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 6, pp. 599–606, 2021, doi: 10.14569/IJACSA.2021.0120670.

[69] I. C. Dipto, M. A. Rahman, T. Islam, and H. M. M. Rahman, “Prediction of Accident Severity Using Artificial Neural Network: A Comparison of Analytical Capabilities between Python and R,” *J. Data Anal. Inf. Process.*, vol. 08, no. 03, pp. 134–157, 2020, doi: 10.4236/jdaip.2020.83008.

[70] H. M and S. M.N, “A Review on Evaluation Metrics for Data Classification Evaluations,” *Int. J. Data Min. Knowl. Manag. Process*, vol. 5, no. 2, pp. 01–11, 2015, doi: 10.5121/ijdkp.2015.5201.

APPENDIX A

```
import os import random import pandas as pd import numpy from sklearn.model_selection
import train_test_split from sklearn.metrics import plot_confusion_matrix,
ConfusionMatrixDisplay, classification_report import matplotlib.pyplot as plt from
sklearn.tree import export_text from joblib import dump, load from tqdm import tqdm

#from wandb.sklearn import plot_learning_curve from
sklearn.model_selection import GridSearchCV import time from
sklearn.metrics import accuracy_score from sklearn.preprocessing
import StandardScaler, MinMaxScaler from joblib import dump,
load from datetime import datetime from imblearn.over_sampling
import SMOTE from imblearn.under_sampling import
RandomUnderSampler from imblearn.pipeline import Pipeline from
sklearn.preprocessing import LabelEncoder

MODEL_OUTPUT_PATH = "/content/"

SCALER_OUTPUT_PATH = "/content/"

 data = pd.read_csv('/content/PowerQualityDistributionDataset1.csv') data
data=data.dropna()

 data=data.drop(['Unnamed: 0'],axis=1) data

data['PowerQualityState'].value_counts()

LEVELS = ['1', '2', '3','4','5']

data.head()

X= data.drop(['PowerQualityState'], axis=1)

y = data['PowerQualityState']


# Used for the scikit models

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2) scaler =
MinMaxScaler() scaler.fit(X_train)

# Normalized Splitted X Data

X_train_normalized = scaler.transform(X_train)

X_test_normalized = scaler.transform(X_test)

 len(X_train_normalized)
```

```
dump(scaler, SCALER_OUTPUT_PATH +
'MinMaxScaler.joblib')

scaler.scale_

import lightgbm as lgb lgb_model=lgb.LGBMClassifier(boosting_type='goss',
num_leaves=31,                    max_depth=10, learning_rate=0.1,
n_estimators=1000,                    min_split_gain=0.1,
min_child_weight=0.001,random_state=42) start = datetime.now()

#lgb_model =
lgb.LGBMClassifier(random_state=42)
lgb_model.fit(X_train_normalized, y_train) end =
datetime.now() duration=end-start print(duration)

start = datetime.now()
accuracy=lgb_model.score(X_test_normalized, y_test)
end = datetime.now() duration=end-start
print(duration)

print('Accuracy: {0:.2%}'.format(accuracy))

accuracy = lgb_model.score(X_test_normalized, y_test)  print('Accuracy:
{0:.2%}'.format(accuracy))

f = plot_confusion_matrix(lgb_model, X_test_normalized, y_test, cmap='Blues',
display_labels = LEVELS)   plt.show()

f.figure_.savefig("lgb_model_confusion.pdf", bbox_inches='tight')

print(classification_report(y_test, lgb_model.predict(X_test_normalized),
target_names=LEVELS))

dump(lgb_model, MODEL_OUTPUT_PATH +
'lgb_model_classifier.joblib')  predicted_risk =
lgb_model.predict(scaler.transform([[]]))[0]

print("Predicted RISK is: " + prediction_result_mapping[predicted_risk])

from sklearn.ensemble import
AdaBoostClassifier

pip install
catboost

from catboost import
CatBoostClassifier cb_model =
```

```
CatBoostClassifier(num_boost_round=1000,learning_rate=0.08,random_seed=42).fit(X_
train_normalized, y_train)

 cb_train, cb_test = cb_model.score(X_train_normalized, y_train),
cb_model.score(X_test_normalized, y_test)

 print(f"Training Score: {cb_train}")
print(f"Test Score: {cb_test}")

 accuracy = cb_model.score(X_test_normalized, y_test)
print('Accuracy: {0:.2%}'.format(accuracy))

 f = plot_confusion_matrix(cb_model, X_test_normalized, y_test, cmap='Blues',
display_labels = LEVELS)   plt.show()

f.figure_.savefig("catboost classifier_confusion.pdf", bbox_inches='tight')

print(classification_report(y_test,cb_model.predict(X_test_normalized),
target_names=LEVELS))

 dump(cb_model, MODEL_OUTPUT_PATH + 'catboostclassifier_classifier.joblib')
predicted_risk =cb_model.predict(scaler.transform([[3.4,1,2.64]]))[0] predicted_risk

#print("Predicted RISK is: " + prediction_result_mapping[predicted_risk])

 from sklearn.neural_network import MLPClassifier mlp = MLPClassifier()

#mlp = load(MODEL_OUTPUT_PATH + 'MLP_classifier.joblib')

 mlp_params={

'learning_rate': ["constant", "invscaling", "adaptive"],

'hidden_layer_sizes': [(10, 10)],

'solver': ["lbfgs", "sgd", "adam"],

'activation': ["logistic", "relu", "tanh"] }

 # Instanciate the Grid for parameter tuning mlp_grid = GridSearchCV(mlp, mlp_params,
cv=5, scoring='accuracy', n_jobs=-1) start=datetime.now() mlp =
mlp_grid.fit(X_train_normalized, y_train) end = datetime.now() duration=end-start
print(duration) mlp_result = pd.DataFrame(mlp.cv_results_)

mlp_result.head()

 print(mlp.best_estimator_) print(mlp.best_params_)
print('Best Accuracy: {0:.2%}'.format(mlp.best_score_))

 test_scores = mlp.cv_results_['mean_test_score']
```

```
plt.title("Mean accuracy after 10-Fold Cross Validation with multiple param combinations") plt.plot(test_scores) plt.ylabel("Accuracy") plt.show() start = datetime.now() accuracy = mlp.score(X_test_normalized, y_test)  end = datetime.now() duration=end-start print(duration)

print('Accuracy: {0:.2%}'.format(accuracy))

f = plot_confusion_matrix(mlp, X_test_normalized, y_test, cmap='Blues', display_labels = LEVELS)   plt.show()

f.figure_.savefig("mlp_confusion.pdf", bbox_inches='tight')

print(classification_report(y_test, mlp.predict(X_test_normalized), target_names=LEVELS))

dump(mlp, MODEL_OUTPUT_PATH +
'MLP_classifier.joblib')   predicted_risk =
mlp.predict(scaler.transform([[3.4,1,2.64]]))[0]
print("Predicted RISK is: " +
prediction_result_mapping[predicted_risk])
```

APPENDIX B

```
import os import random import pandas as pd import numpy from sklearn.model_selection
import train_test_split from sklearn.metrics import plot_confusion_matrix,
ConfusionMatrixDisplay, classification_report import matplotlib.pyplot as plt from
sklearn.tree import export_text from joblib import dump, load from tqdm import tqdm

#from wandb.sklearn import plot_learning_curve from
sklearn.model_selection import GridSearchCV import time from
sklearn.metrics import accuracy_score from sklearn.preprocessing
import StandardScaler, MinMaxScaler from joblib import dump,
load from datetime import datetime from imblearn.over_sampling
import SMOTE from imblearn.under_sampling import
RandomUnderSampler from imblearn.pipeline import Pipeline from
sklearn.preprocessing import LabelEncoder

MODEL_OUTPUT_PATH = "/content/"

SCALER_OUTPUT_PATH = "/content/"

 data = pd.read_csv('/content/PowerQualityDistributionDataset1 (1).csv') data
data=data.dropna()

 data=data.drop(['Unnamed: 0','PowerQualityState','losses amount','Balance
state'],axis=1) data

data['Type of losses'].value_counts()

type_of_losses = { 'capacitance': 0, 'Inductance': 1}

 data = data.replace({'Type of losses': type_of_losses})

 data

LEVELS = ['0', '1']

data.head() X= data.drop(['Type of losses'], axis=1)

 y = data['Type of losses']

 # Used for the scikit models

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)

 scaler = MinMaxScaler() scaler.fit(X_train)

# Normalized Splitted X Data

X_train_normalized = scaler.transform(X_train)
```

```
X_test_normalized = scaler.transform(X_test)

len(X_train_normalizd)

 dump(scaler, SCALER_OUTPUT_PATH +
'MinMaxScaler.joblib')

 scaler.scale_

 import lightgbm as lgb lgb_model=lgb.LGBMClassifier(boosting_type='goss',
num_leaves=31,                    max_depth=10, learning_rate=0.1,
n_estimators=1000,                    min_split_gain=0.1,
min_child_weight=0.001,random_state=42) start = datetime.now()

#lgb_model =
lgb.LGBMClassifier(random_state=42)
lgb_model.fit(X_train_normalized, y_train) end =
datetime.now() duration=end-start print(duration)

start = datetime.now()

accuracy=lgb_model.score(X_test_normalized, y_test)
end = datetime.now() duration=end-start
print(duration) print('Accuracy:
{0:.2%}'.format(accuracy))

accuracy = lgb_model.score(X_test_normalized, y_test)  print('Accuracy:
{0:.2%}'.format(accuracy))

 f = plot_confusion_matrix(lgb_model, X_test_normalized, y_test, cmap='Blues',
display_labels = LEVELS)

plt.show()

f.figure_.savefig("lgb_model_confusion.pdf", bbox_inches='tight')
print(classification_report(y_test, lgb_model.predict(X_test_normalized),
target_names=LEVELS))

 dump(lgb_model, MODEL_OUTPUT_PATH +
'lgb_model_classifier.joblib')

 predicted_risk = lgb_model.predict(scaler.transform([[]]))[0]

print("Predicted RISK is: " + prediction_result_mapping[predicted_risk])

 from sklearn.ensemble import
AdaBoostClassifier
```

```
from catboost import CatBoostClassifier
cb_model =

CatBoostClassifier(num_boost_round=1000,learning_rate=0.08,random_seed=42).fit(X_
train_normalized, y_train)

 cb_train, cb_test = cb_model.score(X_train_normalized, y_train),
cb_model.score(X_test_normalized, y_test)

 print(f"Training Score: {cb_train}")
print(f"Test Score: {cb_test}")

accuracy = cb_model.score(X_test_normalized, y_test)  print('Accuracy:
{0:.2%}'.format(accuracy))

f = plot_confusion_matrix(cb_model, X_test_normalized, y_test, cmap='Blues',
display_labels = LEVELS)   plt.show()

f.figure_.savefig("catboost classifier_confusion.pdf", bbox_inches='tight')
print(classification_report(y_test,cb_model.predict(X_test_normalized),
target_names=LEVELS))

 dump(cb_model, MODEL_OUTPUT_PATH +
'catboostclassifier_classifier.joblib')  predicted_risk
=cb_model.predict(scaler.transform([[3.4,1,2.64]]))[0] predicted_risk

#print("Predicted RISK is: " + prediction_result_mapping[predicted_risk])

 from sklearn.neural_network import
MLPClassifier mlp = MLPClassifier()

#mlp = load(MODEL_OUTPUT_PATH + 'MLP_classifier.joblib')

mlp_params={

'learning_rate': ["constant", "invscaling", "adaptive"],

'hidden_layer_sizes': [(10, 10)],

'solver': ["lbfgs", "sgd", "adam"],

'activation': ["logistic", "relu", "tanh"] }

# Instanciate the Grid for parameter tuning mlp_grid = GridSearchCV(mlp,
mlp_params, cv=5, scoring='accuracy', n_jobs=-1) start=datetime.now() mlp =
mlp_grid.fit(X_train_normalized, y_train) end = datetime.now() duration=end-start
print(duration) mlp_result = pd.DataFrame(mlp.cv_results_)

mlp_result.head()
```

```
print(mlp.best_estimator_) print(mlp.best_params_)
print('Best Accuracy: {0:.2%}'.format(mlp.best_score_))

test_scores = mlp.cv_results_['mean_test_score']

plt.title("Mean accuracy after 10-Fold Cross Validation with multiple param combinations") plt.plot(test_scores) plt.ylabel("Accuracy") plt.show() start = datetime.now()

accuracy = mlp.score(X_test_normalized, y_test)
end = datetime.now() duration=end-start
print(duration)

print('Accuracy: {0:.2%}'.format(accuracy))

f = plot_confusion_matrix(mlp, X_test_normalized, y_test, cmap='Blues', display_labels = LEVELS)   plt.show()

f.figure_.savefig("mlp_confusion.pdf", bbox_inches='tight')
print(classification_report(y_test, mlp.predict(X_test_normalized), target_names=LEVELS))

dump(mlp, MODEL_OUTPUT_PATH + 'MLP_classifier.joblib')

predicted_risk = mlp.predict(scaler.transform([[3.4,1,2.64]]))[0]
print("Predicted RISK is: " + prediction_result_mapping[predicted_risk])
```

APPENDIX C

```
import os import random import pandas as pd import numpy from sklearn.model_selection
import train_test_split from sklearn.metrics import plot_confusion_matrix,
ConfusionMatrixDisplay, classification_report import matplotlib.pyplot as plt from
sklearn.tree import export_text from joblib import dump, load from tqdm import tqdm

#from wandb.sklearn import plot_learning_curve from
sklearn.model_selection import GridSearchCV import time from
sklearn.metrics import accuracy_score from sklearn.preprocessing
import StandardScaler, MinMaxScaler from joblib import dump,
load from datetime import datetime from imblearn.over_sampling
import SMOTE from imblearn.under_sampling import
RandomUnderSampler from imblearn.pipeline import Pipeline from
sklearn.preprocessing import LabelEncoder

MODEL_OUTPUT_PATH = "/content/"

SCALER_OUTPUT_PATH = "/content/"

 data = pd.read_csv('/content/PowerQualityDistributionDataset1 (2).csv') data
data=data.dropna()

 data=data.drop(['Unnamed: 0','PowerQualityState','losses amount','Type of
losses'],axis=1) data

data['Balance state'].value_counts()

 Balance_state = { '1000': 1,'2000': 2,'3000': 3,'4000': 4,'5000': 5}

 data = data.replace({'Balance state': Balance_state})

data

LEVELS = ['1000', '2000','3000','4000','5000']

 ata.head() X= data.drop(['Balance state'], axis=1)

 y = data['Balance state']

 # Used for the scikit models

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)

 scaler = MinMaxScaler() scaler.fit(X_train)

 # Normalized Splitted X Data

X_train_normalized = scaler.transform(X_train)
```

```
X_test_normalized = scaler.transform(X_test)

 len(X_train_normalized)

 dump(scaler, SCALER_OUTPUT_PATH +
'MinMaxScaler.joblib')

 scaler.scale_

 import lightgbm as lgb lgb_model=lgb.LGBMClassifier(boosting_type='goss',
num_leaves=31,                    max_depth=10, learning_rate=0.1,
n_estimators=1000,                    min_split_gain=0.1,
min_child_weight=0.001,random_state=42) start = datetime.now()

#lgb_model =
lgb.LGBMClassifier(random_state=42)
lgb_model.fit(X_train_normalized, y_train) end =
datetime.now() duration=end-start print(duration)

 start = datetime.now()

accuracy=lgb_model.score(X_test_normalized, y_test)
end = datetime.now() duration=end-start
print(duration) print('Accuracy:
{0:.2%}'.format(accuracy))

 accuracy = lgb_model.score(X_test_normalized, y_test)  print('Accuracy:
{0:.2%}'.format(accuracy))

 f = plot_confusion_matrix(lgb_model, X_test_normalized, y_test, cmap='Blues',
display_labels = LEVELS)

plt.show()

f.figure_.savefig("lgb_model_confusion.pdf", bbox_inches='tight')
print(classification_report(y_test, lgb_model.predict(X_test_normalized),
target_names=LEVELS))

 dump(lgb_model, MODEL_OUTPUT_PATH +
'lgb_model_classifier.joblib')

 predicted_risk = lgb_model.predict(scaler.transform([[]]))[0]

print("Predicted RISK is: " + prediction_result_mapping[predicted_risk])

pip install
catboost

 from catboost import CatBoostClassifier cb_model =
```

```
CatBoostClassifier(num_boost_round=1000,learning_rate=0.08,random_seed=42).fit(X_
train_normalized, y_train)

 cb_train, cb_test = cb_model.score(X_train_normalized, y_train),
cb_model.score(X_test_normalized, y_test)

 print(f"Training Score: {cb_train}")
print(f"Test Score: {cb_test}")

accuracy = cb_model.score(X_test_normalized, y_test)  print('Accuracy:
{0:.2%}'.format(accuracy))

 f = plot_confusion_matrix(cb_model, X_test_normalized, y_test, cmap='Blues',
display_labels = LEVELS)   plt.show()

f.figure_.savefig("catboost classifier_confusion.pdf", bbox_inches='tight')
print(classification_report(y_test,cb_model.predict(X_test_normalized),
target_names=LEVELS))

 dump(cb_model, MODEL_OUTPUT_PATH +
'catboostclassifier_classifier.joblib')  predicted_risk
=cb_model.predict(scaler.transform([[3.4,1,2.64]]))[0] predicted_risk

#print("Predicted RISK is: " + prediction_result_mapping[predicted_risk])

 from sklearn.neural_network import
MLPClassifier mlp = MLPClassifier()

#mlp = load(MODEL_OUTPUT_PATH + 'MLP_classifier.joblib')

 mlp_params={

'learning_rate': ["constant", "invscaling", "adaptive"],

'hidden_layer_sizes': [(10, 10)],

'solver': ["lbfgs", "sgd", "adam"],

'activation': ["logistic", "relu", "tanh"]

}

 # Instanciate the Grid for parameter tuning mlp_grid = GridSearchCV(mlp, mlp_params,
cv=5, scoring='accuracy', n_jobs=-1) start=datetime.now() mlp =
mlp_grid.fit(X_train_normalized, y_train) end = datetime.now() duration=end-start
print(duration) mlp_result = pd.DataFrame(mlp.cv_results_)

mlp_result.head()
```

```
print(mlp.best_estimator_) print(mlp.best_params_)
print('Best Accuracy: {0:.2%}'.format(mlp.best_score_))

test_scores = mlp.cv_results_['mean_test_score']

plt.title("Mean accuracy after 10-Fold Cross Validation with multiple param
combinations") plt.plot(test_scores) plt.ylabel("Accuracy") plt.show() start =
datetime.now()

accuracy = mlp.score(X_test_normalized, y_test)
end = datetime.now() duration=end-start
print(duration)

print('Accuracy: {0:.2%}'.format(accuracy))

f = plot_confusion_matrix(mlp, X_test_normalized, y_test, cmap='Blues',
display_labels = LEVELS)   plt.show()

f.figure_.savefig("mlp_confusion.pdf", bbox_inches='tight')
print(classification_report(y_test, mlp.predict(X_test_normalized),
target_names=LEVELS))

dump(mlp, MODEL_OUTPUT_PATH +
'MLP_classifier.joblib')

predicted_risk = mlp.predict(scaler.transform([[3.4,1,2.64]]))[0]
print("Predicted RISK is: " + prediction_result_mapping[predicted_risk])
```