Republic of Iraq
Ministry of Higher Education and Scientific Research
University of Anbar
College of Computer Science and Information Technology
Department of Computer Science

Electricity-Theft Detection in Smart Grids based on Deep Learning

A Thesis
Submitted to the Department of Computer Science - College of Computer Science and Information Technology, University of Anbar as a Partial Fulfillment of the Requirements for Master Degree of Science in Computer Science

By

Noor Mahmoud Ibrahim
Bachelor of Computer Science

Supervised By

Prof. Dr. Sufyan T. Faraj Al-Janabi

Prof. Dr. Belal Al-Khateeb

1442A.H. *2021A.D.*

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

﴿يَرْفَعِ ٱللَّهُ ٱلَّذِينَ ءَامَنُواْ مِنكُمْ وَٱلَّذِينَ أُوتُواْ ٱلْعِلْمَ دَرَجَٰتٍۚ وَٱللَّهُ بِمَا تَعْمَلُونَ خَبِيرٌ ﴿١١﴾﴾

صدق الله العظيم

سورة المجادلة رقم الآية ﴿١١﴾

إسم الطالبة: نور محمود إبراهيم
كلية علوم الحاسوب وتكنولوجيا المعلومات - قسم علوم الحاسبات
عنوان الرسالة: إكتشاف سرقة الكهرباء في الشبكات الذكية بإستخدام التعلم العميق.

عنوان البحث			
Electricity-Theft Detection in Smart Grids based on Deep Learning			
Scopus / Q3	نوع المجلة	Bulletin of Electrical Engineering and Informatics	جهة النشر
ISSN: 2302-9285	رقم المجلة	Accepted	حالة النشر
			رابط البحث

إسم وتوقيع المشرف
Prof. Dr. Sufyan T. Faraj Al-Janabi

إسم وتوقيع رئيس القسم
Assist. Prof. Dr. Wesam Mohammed Jasim

إسم وتوقيع المشرف المشارك
Prof. Dr. Belal Al-Khateeb

Supervisor's Certification

We certify that we read this thesis entitled *"**Electricity-Theft Detection in Smart Grids based on Deep Learning**"* that was carried out under our supervision at the Department of Computer Science of the University of Anbar, by the student *"**Noor Mahmoud Ibrahim**"* and that in our opinion it meets the standard of a thesis for the degree of Master of Science in Computer Science.

Signature:
Name: ***Prof. Dr. Sufyan T. Faraj Al-Janabi***
Date: / /2021

Signature:
Name: ***Prof. Dr. Belal Al-Khateeb***
Date: / /2021

Certification of the Examination Committee

*We the examination committee certify that we have read this thesis entitled "**Electricity-Theft Detection in Smart Grids based on Deep Learning**" and have examined the student "**Noor Mahmoud Ibrahim**", in its contents and what is related to it, and that in our option it is adequate to fulfill the requirements for the degree of **Master of Computer Science**.*

Signature:
*Name: **Prof. Dr. Azmi Tawfeeq Hussein*** ***(Chairman)***
Date: / / 2021

Signature:
*Name**: Assist. Prof. Dr. Jane Jaleel Stephan*** ***(Member)***
Date: / / 2021

Signature:
*Name: **Lecturer. Dr**. **Atheer Bassel AbdulKareem*** ***(Member)***
Date: / / 2021

Signature:
*Name: **Prof. Dr. Sufyan T. Faraj Al-Janabi*** ***(Supervisor)***
Date: / / 2021

Signature:
*Name: **Prof. Dr. Belal Al-Khateeb*** ***(Supervisor)***
Date: / / 2021

Approved by the Dean of the College of Computer Science and Information Technology, University of Anbar.

Signature:
*Name: **Prof. Dr. Salah Awad Salman***
*Title: **Dean of the College***
Date: / / 2021

Acknowledgements

First, all praises to Allah Almighty, who enabled me to complete this work successfully.

I wish to express my deep respect and thank to my supervisors Prof. Dr. Sufyan T. Faraj Al-Janabi and Prof. Dr. Belal Al-Khateeb for their appreciable advice, important comments, and support during the research.

Special thanks to all my teachers in the College of Computer Science and Information Technology" for everything.

I am grateful to the staff of the College of Computer Science and Information Technology.

Finally, I express my deep gratitude for the great support I had from my father, my mother, and my brothers. Without their endless patience and continuous support, this thesis would not have seen the light.

My thanks for all...

Noor Mahmoud Ibrahim

2021

Dedication

This thesis is dedicated to:

My parents

My supervisors

My teachers

My brothers

and my friends

Noor Mahmoud Ibrahim

2021

Student name: ***Noor Mahmoud Ibrahim***

Thesis title: ***Electricity-Theft Detection in Smart Grids based on Deep Learning***

Abstract

Electricity theft is a major concern for the utilities. With the advent of smart meters, the frequency of collecting household energy consumption data has increased, making it available for advanced data analysis, which was not possible earlier. Indeed, using Smart Grid (SG) networks, which are recently upgraded networks of connected objects, can greatly improve the reliability, efficiency, and sustainability of the traditional energy infrastructure.

The SG infrastructure produces a massive amount of data, including the power consumption of individual users. Utilizing this data, machine learning, and deep learning techniques can accurately identify electricity theft users. This thesis presents a Convolutional Neural Network (CNN) based model for automatic electricity theft detection that can achieve high performance classification and detection.

The work considers experimentation to find the best configuration of the sequential model (SM) for classification, beginning with two layers and ending with four layers. The best performance has been obtained in two layers' architecture with the first layer consists of 128 nodes and the second layer is 64 nodes, where the accuracy reached up to 0.92. This enables the design of high-performance electricity signals' classifier that can be applied several applications.

Designing electricity signals classifiers has been achieved using CNN and the data extracted from electricity consumption dataset using SM. In addition, the Blue Monkey (BM) algorithm is exploited to reduce the number of features in the dataset, where these values are used to build models with high performance. In this respect, the emphasis of this thesis has been on reducing the required number features in the dataset in order to achieve a high performance electricity signals' classifier model. The experiments have justified the high performance of the proposed systems, where combining both the CNN and BM algorithms requires only 666 features

compared to 1035 features using CNN alone. This demonstrates the superiority of the CNN and BM model over the CNN model in terms of reducing the features of the model while the accuracy remaining the same.

Keywords: Smart Grid (SG), Deep Learning (DL), Convolutional Neural Network (CNN), sequential model (SM), Blue Monkey Algorithm (BM), electricity consumption dataset.

Contents

List of Tables

List of Figures

Abbreviations

ANN	Artificial Neural Networks
AMI	Advanced Metering Infrastructure
AMR	Automatic Meter Reading
ABC	Artificial-Bee-Colony
AUC	Area Under Curve
ANOVA	ANalysis Of VAriance
AMIIDS	Advanced Metering Infrastructure Intrusion Detection System
BM	Blue Monkey algorithm
BBO	Biogeography-Based Optimizer
BDR	Bayesian Detection Rate
CNN	Convolutional Neural Networks
CFBETD	Consuming Form-Based Energy Theft Detector
DL	Deep Learning
DNN	Deep Neural Networks
DG	Distributed Generation
DR	Detection Rate
DS	Distributed Storage
DT	Decision Trees
DSR	Demand Side Response
DSM	Demand Side Management
DSO	Distribution System Operators
ETD	Electricity Theft Detection
FACTS	Flexible AC Transmission System
GSA	Gravitational Search Algorithm
GAM	Generalized Additive Model
HVDS	High Voltage Detection System
HAN	Home Area Network
HV	High Voltage
KPCA	Kernel Principal Component Analysis
K-NN	K-Nearest Neighbour
LSTM	Long-Short Term Memory
LV	Low Voltage
MV	Medium Voltage
NN	Neural Networks
NaN	Non-Numeric
NTL	Non-Technical Losses
OPF	Optimum Path Forrest
PLC	Power Line Communication
PMU	Phasor Measurement Unit
PSO	Particle Swarm Optimization
PCA	Principal Component Analysis
RNN	Recurrent Neural Networks
RF	Random Forests
ROC	Receiver Operating Curve
RTU	Remote Technical Unit
RTU	Remote Terminal Units

RFID	Radio Frequency Identification
ReLU	Rectified Linear Unit
SCADA	Supervisory Control and Data Acquisition
SOM	Self-Organizing Map
SVM	Support Vector Machine
SGCC	State Grid Corporation of China
SM	Sequential Model
SG	Smart Grid
TL	Technical Losses
WAMS	Wide Area Measurement System

Chapter One

General Introduction

Chapter One:

General Introduction

1.1 Introduction

The grid of electrical energy is one of the most essential and complicated artificial schemes in the new society. With the most recent advancements in observing, communication, control, and sensing, the inheritance energy grid is currently changed alongside the trip to a smart grid. Smart Grid (SG) is the ever-growing dispersion of renewable and divided source of power, which is intended to attain flexibility, self-healing, effectiveness, and sustainability. The idea of SG is being recognized over the application of pretend infrastructure covering the inheritance power grid [1]. The cyber-infrastructure allows the group and study of data from lots of different dispersed endpoints, for example, units of determination of phasor, smart meters, and breakers of the circuit.

Usually, these grids contain some improvements that will develop the dependability, effectiveness, and the delivery of continuous source of energy to households and industries. Besides, SG contains different resources of renewable energy such as (power of wind, solar, and others), distributed storage (DS), and distributed generation (DG) [2–6]. The term of the system of smart metering explains a smart electric instrument that determines the data of using of energy, providing more accurate details than a conventional meter, and drives and obtains data by two-way connection [7]. Consequently, grids of smart metering operate with smart sensors permitting companies to run and regulate the SG, supplied with the technology of communication and information [8].

Electrical energy has become essential in a human's life. Losses of electrical energy regularly happen for the duration of production, distribution, and transition of electrical energy. The losses of electrical energy can generally be classified into Non-technical losses (NTLs) and technical losses (TLs) [9]. Electricity-theft is one of the most serious NTLs.

There is a large group of investigations on detecting electricity-theft. Traditional ways of detection of electricity-theft contain physically examination problematical meter set up or disconfirmation, associating the irregular meter readings with the

regular ones and observing a line of the transition of the by-passed power. These ways are ineffective, tremendously time-consuming, and costly. The presence of SGs brings chances in resolving of electricity-theft. SGs are comprised of conventional networks of power, grids of communications linking smart devices (for example, smart sensors and meters) in networks, and calculating services to sense and regulate networks [10]. Information and energy move in smart networks attach companies of service and employers. In this way, smart sensors or meters may collect a variety of data, such as network status information, using electrical energy, funding information, and cost of electrical energy [11].

As a result, the emphasis of this thesis is on proposing an effective technique for detecting electricity-theft in order to address all the concerns raised above. Specifically, a Convolutional Neural Networks (CNN) have been initially suggested with a recently proposed nature-inspired metaheuristic optimization algorithm called the Blue Monkey (BM) algorithm model to recognize the thieves of electricity and study the data of consumption of electricity. The CNN part consists of several convolutional layers, a pooling layer and a completely connected layer. Principally, the CNN component can capture the periodicity of electricity consumption data. This model combines the strength of the CNN component and the BM algorithm to aid in the detection of electricity-theft. The primary study donations of this thesis can be reviewed as:

- Originally, this work suggests a deep algorithm model consisting of CNN and the BM algorithm to examine electricity-theft in smart networks. To the best of our knowledge, it is the first research to suggest such a deep algorithm model (Mixing CNN with the BM algorithm) and carry it out to study electricity-theft in smart networks.
- Wide-scale experiments have been conducted on a huge accurate electricity consuming dataset. Furthermore, the results of these experiments show that our model (CNN & BM) outperforms several other existing models.
- The proposed model has several merits, including the simplification of the novel knowledge brought via the CNN model and the accurateness in detection electricity-theft.

1.2 Methods of Power Theft in Power System

Losses of non-technical and technical nature are happening solitary in distribution and transition but not in production. It is very problematic to discover non-technical losses happening in the system. There are several techniques of electricity-theft (electrical energy) [12]. Some of them are straight attaching from the line, inserting foreign substances into the meter, digging punctures in the electro-mechanical energy meter. Electromagnetic meters are hardened via putting a too viscous fluid, inserting film, and utilizing rigid magnets, for the interruption of the disc. The electricity-theft is occurred via showing meter to mechanical shock and utilizing the external phase beforehand meter stations. As a consequence, subscribers get free energy without any record.

In other methods, it is possible to substitute the chain of energy at the meter connector box, and the amperage doesn't permit over the present coil of the meter; therefore, the meter doesn't record the consumption of energy. Electricity-theft becomes an infamous problematic in power systems. There are different control approaches for the electricity-theft, but it is still not easy to decrease or eliminate the problem. The most common types of electricity-theft are [12]:

- Inserting foreign substances into the meter.
- Direct attaching from the line.
- Rearranging energy meter reading.
- Varying the stations of leaving and arriving at the meter.
- Utilizing rigid magnets as neodymium magnets.
- Digging punctures into electromechanical energy meter.
- Destructive the pressure coil of the meter.
- Inappropriate or illegitimate standardization of energy meters.
- Putting a too viscous fluid.
- Injection film.
- Revealing the meter to mechanical shock.

The employed techniques for controlling power theft include [12]:

- Recognition and Detection based on a system of High Voltage Detection System (HVDS).

- Utilizing neural grids/ model of Support Vector Machine (SVM).
- Utilizing smart meter/nefarious meter insertion.
- Advanced metering infrastructure (AMI).
- Electricity-theft Power control Automatic Meter Reading (AMR) by a system of Power Line Communication (PLC).
- Intelligent modelling system for detection of line losses in allocation scheme of power.

The comparison of the above-mentioned controlling methods in terms of system reliability, economy and system efficiency is shown in Table (1.1) [12].

Table (1.1): Comparison of control methods in terms of system reliability, economy, and system efficiency [12].

No.	Ways of Controlling	Accuracy of System	Economy	Efficiency of System
1	Detection & Identification according to System of HVDS	Normal	Less	Less
2	Using a Neural Network	Good	Mathematical model cost less	Moderate
3	Using Smart Meter	Perfect (complete)	Extraordinary	Extraordinary
4	AMI(Advance Metering Infrastructure)	Perfect(complete)	Moderate	Moderate
5	Using PLC (Power Line Communication)	Good	Normal	Extraordinary
6	Intelligent System	Perfect(complete)	Extraordinary	Extraordinary

1.3 Smart Grid (SG)

The electricity network is considered as an SG that can smartly mix the actions of whole employers linked to its producers, customers and those that do both to proficiently provide maintainable, financial, and safe sources of electricity [13]. A smart network uses advanced products and facilities together with technologies of communication, control, smart observing, and self-healing to [14]:

- Permit customers to show a portion in improving the system procedure.

- Preferable service the procedure and connection of producers of wholly technologies and sizes.
- Meaningfully decrease the ecological influence of the entire scheme of the source of electricity.
- Deliver customers with superior information and source selection.
- Supply improved dependability levels and source security.

Therefore, the primary purposes of the SG are [14]:

- Deliver a user-centric method and permit novel facilities to arrive at the market.
- Deliver availability to an opened market and foster antagonism.
- Allow request side partaking (Demand Side Response (DSR), Demand Side Management (DSM).
- Allow distributed generation and employment of sources of renewable energy.
- Confirm best utilizing of a central production.
- Study the features of the group.
- Preserve source security, interoperability and confirm integration.
- Found novelty as an inexpensive driver for the renewal of electricity networks.
- Consider suitably the influence of ecological limits.
- Notify the political and controlling features.

The Smart Grid components are Supervisory Control and Data Acquisition (SCADA), Phasor Measurement Unit (PMU), Flexible AC Transmission System (FACTS), Advanced Conductors devices, electric power generators, electric power substations, transmission and distribution lines, controllers, smart meters, collector nodes, and distribution and transmission control centers. Figure (1.1) depicts the key components of SG [15].

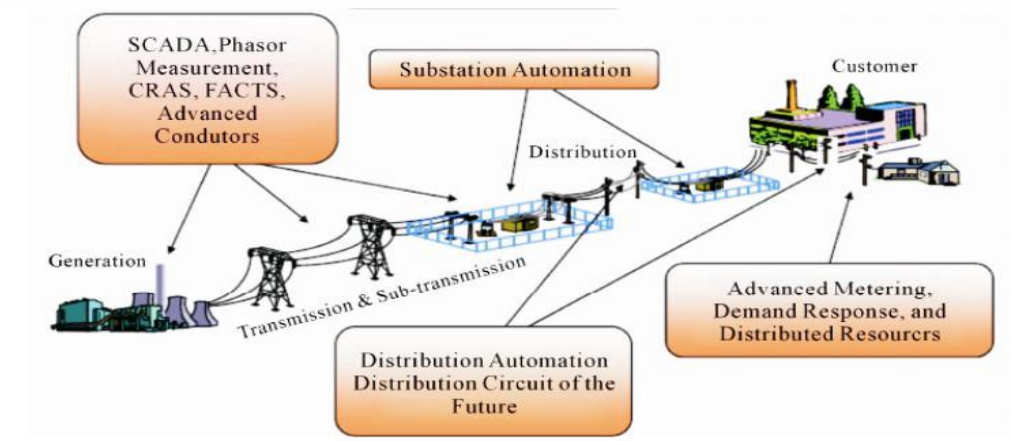

Figure (1.1): Components of the smart grid [15].

1.4 Detection of Electricity Theft based on CNN

Electricity-theft can be damaging to power network sources and cause financial losses. Smart networks can assist in resolving the problems of electricity-theft possessing the obtainability of massive data produced from smart networks. The data examination on the smart networks data is useful in detecting of electricity-theft due to the irregular pattern of electrical energy consumption of thieves of energy. Nevertheless, the current approaches have poor accurateness of detection of electricity-theft as most of them were based on one dimensional (1-D) data of electricity consumption and failed to arrest the electricity consuming periodicity [16].

Thus, it is more prudent to propose an enhanced technique of detection of electricity-theft based on the model of Convolutional Neural Networks (CNN) to treat all worries mentioned above. The CNN can exactly recognize the non-periodicity of electricity-theft and the periodicity of regular electricity using based on two-dimensional (2-D) data of electricity consumption. Consequently, the model of CNN can attain great working in the detection of electricity-theft. The CNN is made up of several convolutional layers, a pooling layer and an entirely attached layer.

1.5 Related Works

In this section, a thorough literature review on some related issues of the Smart Grid (SG) is presented. Firstly, some critical works on SGs are mentioned. Secondly, a survey on electricity theft detection techniques is given. Then, some related works on deep neural networks are reviewed. Next, a survey on the deployment of (CNNs) for detection of electricity theft is given. Finally, the related work on the BM algorithm is introduced.

A. Smart Grids

Lately, matters of privacy and security have been the issues of comprehensive research since the national economy, public security, and safety depend significantly on the grids of energy. Though weaknesses of privacy and security are always being appeared in the protocols, technologies of the grid, and devices utilized in the importance of fears to system-level safety, the systems of energy, fears to privacy are not continuously completely understood in grids of SG metering, and threats or theft by facilities. Next, we explain new survey papers in this domain and indicate exceptional contributions and the distinctive features of them. These contain some survey papers that have been shown on the matters of privacy and security in the field of SG.

In 2011, Line *et al.* compared the security requirements between SG communication network and telecommunication networks [17]. Then they listed the overall cyber security challenges, for example, trust models, connectivity, management of security, the privacy of consumers, software vulnerabilities, and human factors. Solutions to these challenges were also proposed.

In 2012, Deng and Shukla surveyed the vulnerabilities and countermeasures, especially for the transmission subsystem within SG [18]. They focused on the point of weaknesses of technology of Phasor Measurement Units (PMUs) and Wide Area Measurement System (WAMS). They divided the attacks into four group traffic analysis attack, denial of service attack, malicious data injection attack, and great-level implementations attack. Those authors presented the basic of PMU, case approximation with PMU, and how that can be utilized to inverse attack.

In 2013, Wang and Lu examined challenges of security in the grid of SG, containing Home Area Networks (HANs), Advanced Metering Infrastructures (AMIs), subsystems of distribution and transmission [19]. They showed the

necessities of security and estimated network fears with matter studies. The study principally considered cryptographic countermeasures containing verification and managing of the key in different fields of SG. Their paper contained detailed logical study together with some conventional protocols (e.g., distributed network protocol) in the fields of energy. However, since 2013, techniques of comprehensive new and progressive safety have been presented, and those must be discovered.

In 2013, Baig and Amoudi classified the SG cyber-attacks and countermeasures through five categories: Supervisory Control and Data Acquisition (SCADA), Injection of Data and Replay Attacks, Smart Meter Attacks, Network-based Attacks, and Physical Layer Attacks, which span home area networks, grids of the neighborhood, and extensive area grids [20].

In 2014, Komninos *et al.* presented SG and smart home safety study [21]. Those authors generally assumed the communication amid the environments of SG and smart home are categorized their hazards of safety. The paper studied some representative fears and estimated theoretical influences from SG to smart home and conversely. They delivered a review of the presented literature as the countermeasures of safety and contained the SG's current doings from 2009 to 2013. Komninos et al. studied several papers from the point of view of safety countermeasures, including privacy, the critical study of these systems was not explained.

In 2014, Mohassel *et al.* explained a study on (AMI) advanced metering infrastructure [22]. They studied the main ideas of AMI. They showed the physical and cyber safety challenges containing privacy briefly. Their paper included partial but necessities of security and privacy in the grid of AMI. Nevertheless, those authors do not contain detailed threat model, and explanation on modern schemes of security and no expressed the privacy maintaining systems.

B. Electricity Theft Detection (ETD)

The current methods are studied of electricity theft detection in the literature, which use consuming data of smart meters to discover deceitful consumers. Observing of load profiles of consumers for marks of electricity-theft in conventional power schemes has attracted the concerns of academics to this point.

In 2010, Nagi *et al.* applied a data removal technique alongside with Support Vector Machine (SVM) classifier to detect irregular manners [23]. The average daily consumption of consumers over two years was estimated, and the long-period trend in consumption of energy was utilized to detect deceitful consumers. This technique can detect unexpected variations in load profile. Also, the detection, the delay is around two years.

In 2011, Angelos *et al.* utilized six months using reports, five characteristics containing maximum consuming, mean consumption, inspection remarks summation, standard deviance, and the mean consumption of the neighborhood to produce a usual form of consumption of power for every consumer. K-means based fuzzy clustering was achieved to collection consumers with the same profiles. A categorization of fuzzy was then completed, and Euclidean spaces to the group centers were determined. Customers with ample spaces to the cluster centers were assumed potential cheats. Gathering the consumers and depending on long-period measurements limited the accurateness of this ETDS and produced long detection delay. Possessing more detailed metering info in Advanced Metering Infrastructures, Consuming Form-Based Energy Theft Detector (CFBETD) may deliver a much better working with a much shorter delay [24].

In 2011, Depuru *et al.* incorporated a neural network model to calculate factors of Support Vector Machine to decrease the time of training of the classifier, and a data encoding technique was projected to develop the classifier speed and effectiveness. Their way is only active in detecting electricity-theft attacks that produce in zero using reports since in one-step of the encoding process, the metering data is changed into double values. So, the suggested method of classification can't detect a widespread choice of kinds of attack [25].

C. Deep Learning and Convolutional Neural Networks (CNNs)

This subsection, presents a survey on some earlier work related to CNNs.

In 2013, Abdel-Hamid et al. explained the CNN innovative in the variability of domains associated with forming appreciation from image treating to voice recognition [26]. The most advantageous feature of CNN is decreasing the factors number in Artificial Neural Networks (ANNs). This attainment has encouraged both designers and academics to approximate bigger models to resolve difficult jobs, which wasn't probable with classic Artificial Neural Networks.

In 2016, Mallat extended earlier presented tools to progress a mathematical framework to analyze the properties of general CNN architectures [27]. At a significant level, the extension was attained via substituting the requirement of invariants and contractions to translations via contractions along with adaptive collections of local symmetries. Additional, the wavelets were substituted via adapted filter weights same to deep learning models.

In 2017, Albawi et al. described the term Deep Learning or Deep Neural Network that denotes to Artificial Neural Networks (ANN) with several layers [28]. Over the final rare decades, it has been measured one of the most influential apparatuses and has become very prevalent in the literature, as it is capable of treating a massive quantity of data. The attention in possessing deeper unobserved layers has newly initiated to exceed working of classical approaches in various domains, particularly in form recognition. One of the most general deep neural grids is Convolutional Neural Networks. It had taken this term from linear mathematical process amid matrixes named convolution. Convolutional Neural Networks have several layers; containing assembling layer, convolutional layer, ultimately linked layer, and non-linearity layer.

In 2017, Xu et al. presented the prime variance amid regular neural grids and convolutional neural networks, where convolutional neural networks have an automatic feature extractor, which comprises of a complication layer and a down sampling layer (or pooling layer) [29]. A complication layer contains a pair of feature charts, and each has some neurons. Generally, the factors of the complication kernel are adjusted arbitrarily, i.e., utilizing it identified initialization procedures and would be modified for the duration of the stage of training.

D. Convolutional Neural Networks for Electricity Theft Detection

In this subsection, an emphasis is given to review some recent papers that have been conducted on using CNNs for electricity theft detection.

In 2012, Krizhevsky et al. explored the use of CNNs for the task of detection [30]. Motivated via the numerical model method, the periodicity of consecutive data is of considerable significance for the classifier, and the series might have monthly, weekly, yearly periodicity or seasonal. Aimed at the detection of electricity-theft, the form of electrical energy consumption is very noticeable for various users. So, an adequate explanation of the periodicity can be beneficial to develop the

accurateness of the detection of electricity-theft. Concretely, they suggested adjusting the multi-scale Dense Net, which can automatically capture the short-term and extensive-term periodic characteristics of the consecutive data.

In 2016, Bhat et al. investigated three deep learning methods for detection of electricity-theft, specifically, CNNs, Long-Short Term Memory (LSTM), recurrent neural networks (RNNs), and loaded autoencoders [31]. Nevertheless, the detectors working was examined by utilizing synthetic data, which didn't permit a reliable valuation of the performance of detector associated with shallow architectures.

Furthermore, the working of the suggested detectors was examined in contradiction of only two kinds of attacks, i.e., bypass attack (decreasing the reported consuming of energy to 0) and partial decrease attack (dropping the reported consuming of energy via some fraction [31].

E. Blue Monkey Algorithm (BM)

This subsection highlights previous work related to the Blue Monkey (BM) algorithm. **In 2019, M. Mahmood and B. Al-Khateeb** introduced a Forty-three of well-recognized trial functions, which utilized in the optimization area are utilized as standard to examine the BM algorithm [32]. Additionally, confirmation of BM via a relative working examine with Gravitational Search Algorithm (GSA), Artificial-Bee-Colony (ABC), Particle Swarm Optimization (PSO), and Biogeography-Based Optimizer (BBO). The attained outcomes established that BM algorithm is modestly associated with the selection of metaheuristic algorithms. BM is capable of joining towards the worldwide optimum over difficulties of optimization too.

1.6 Problem Statement

Electricity-theft is considered as an illegal manner of theft electrical energy from networks of power. This harmful manner can be completed via bypassing the electricity meter, hacking the meter, or interfering the meter reading. The data-driven methods of detection of electrical energy theft have taken broad consideration in recent times because of electricity-theft can produce in the irregular forms of consuming of electrical energy and the obtainability of smart-meter readings and consuming of electrical energy data from smart grids.

The irregularity of consuming electricity data for energy thieves is discovered, which can be possibly caught via machine learning tools. A preliminary examination is carried out on consuming of electrical energy data. Afterwards statistically examining the consumption of electrical energy data of both thieves of energy and usual consumers, can find that the electricity consumption data of energy thieves are typically less non- frequent or frequent, associated with that of usual consumers. This monitoring can facilitate to classify the irregular using of electricity and the periodicity of the electricity consumption.

Nevertheless, it is challenging to examine the periodicity of the electricity consuming data because of many reasons as:

1) It is problematic to study the periodicity of the electricity consuming data because it is 1-D time series data with an enormous size,
2) The electricity consumption data is frequently incorrect and loud,
3) Several traditional methods of data investigation, for example, ANN and Support Vector Machine (SVM) can't be straight carried out to the consuming of electricity data because of the calculation difficulty and the restricted simplification ability. To face the above challenges, has been suggested using CNN.

1.7 Aim of Thesis

The main objectives of this work are like this:

- Design a deep learning-based system for the detection of electricity-theft in smart grids to reduce the theft of electrical energy and reduce the abnormal consumption of electricity.
- Conducting experiments to test the best configuration of the sequential model for electricity theft detection to choose the best configuration of the CNN.
- Employ an optimization algorithm (BM) to reduce the extracted features in order to speed up the performance of the designed system.

1.8 Contributions

Using the Blue Monkey (BM) algorithm to reduce the selected features in order to speed up the designed system, and use reduced features set to build and train models of CNN theft detection to get better performance.

1.9 Thesis Structure

The remaining chapters of this thesis are arranged as follows:

Chapter Two provides the background of Electricity Theft Detection (ETD) systems that use CNN.

Chapter Three offers a full description of the proposed system in terms of algorithms, as well as measures used to implement the approaches of ETD system and achieve the desired goal.

Chapter Four presents the results are obtained through experimentation with the proposed approaches. Indeed, those results are discussed.

Chapter Five presents the main conclusions of this research in addition to some suggestions for future works.

Chapter Two

Theoretical Background

Chapter Two:

Theoretical Background

2.1 Introduction

Smart grids bring chances in resolving of electricity-theft. Smart grids comprise of communications grids, conventional power networks joining smart devices like smart sensors and meters in networks and calculating services to sense and regulate grids. Information and energy moves in SGs attach employers and usefulness firms. In this behavior, smart sensors or meters can receive data, for example, status information of grids, using electricity, financial information and electricity fee. The data of SGs is supportive for us to project systems of request-response managing, estimate the fee of electricity and timetable the electrical energy in more gainful method [16].

The electricity loss is a significant problematic challenged via firms of power wholly over the world. Regularly, losses happen for the duration of electricity distribution, generation, and transmission. The losses of electricity can be usually classified into losses of Non-technical and technical [9]. One of the significant Non-technical losses is electricity-theft. This lousy behavior contains bypassing the electricity meter, hacking the meter, or tampering the meter reading [33]. Electricity-theft can produce in the weighty load of electrical schemes, the flowing electricity, the hazards to public security and the massive proceeds loss of power firm, for example, electrical and fires shocks. Several methods for detection of Non-technical losses (NTLs) have appeared which can be categorized into three main classes: Data-oriented methods, network-oriented methods, and hybrid methods according to each technique's learning method. However, the most common and most promising of these techniques is ETD based on Deep Learning.

2.2 Types of Energy Losses

The energy loss in distribution and transmission in electrical energy is a significant problematic challenged via firms of power over the world. Generally, the losses of energy are categorized into losses of Non-technical and technical [9].

The technical loss is ingrained to the electricity transference, which is produced via interior activities in the parts of the power scheme, for example, the converters and transition liner [34]. The Non-technical loss is definite as the variance amid technical losses and overall losses, which is mainly produced via electricity-theft that happens over physical attacks as meter reading tampering, line tapping, or meter breaking [35].

These manners of electricity fraudulence might bring around the income loss of power firms as the losses produced via electricity-theft are computed approximately $4.500 million yearly in the USA [36]. It is estimated that international utility corporation loses more than 20000 million annually in the formula of electricity-theft [37]. Additionally, electricity-theft manners can impact the security of the power system. For example, the weighty load of electric systems produced via electricity-theft might cause fires, which intimidate the safety of people. Consequently, the correct electricity-theft detection is essential for the safety of the power grid and stableness. With the application of the advanced metering infrastructure in SGs, services of power attained huge quantities of consumption of electricity data at a significant smart meter's frequency, which is useful for electricity-theft detection [38, 39].

Nevertheless, the grid of advanced metering infrastructure unlocks the door for several novel electricity-theft attacks. These attacks in the advanced metering infrastructure can be thrown in different ways, for example, cyber-attacks and digital tools. The significant ways of detection of electricity-theft contain humanly observing illegal line diversions, associating hateful meter records with the generous ones, and checking complicated apparatus or hardware. Nevertheless, these approaches are expensive and consumption of time tremendously for the duration of complete confirmation of whole meters in a scheme. These manual methods can’t avert cyber-attacks too. To resolve the difficulties stated above, several methods have been suggested in the preceding years. These approaches are principally classified into models of artificial-intelligence-based, state-based, and game theory based [40].

2.3 Methods of Detection of Non-Technical Loss

Electricity-theft has been a primary matter for several years. Distribution System Operators (DSOs) have been testing to perceive electricity-theft; nevertheless, the phenomenon maintains, whereas modest meter check ways can't sufficiently recognize maximum states of fraudulence [41].

In this section, the utmost latest and features research directions on NTL revealing are studied with their main features in brief. NTL uncovering systems are prearranged in three big groups: network-oriented, hybrids, and data-oriented. Relied on the prime perception behindhand detection of NTL, ways of data and network-oriented are more categorized to subgroups. Apart from classifying the different ways, the researchers focusing on size, types of data, algorithms, features estimation metrics, and reply times of detection system of NTL.

2.3.1 Classification of Ways of Non-Technical Loss Detection

According to a review of scientific papers on detection NTL, there is no single conventional procedure keeping an eye on for identifying fraudulence. Researchers assume many ways from various domains of knowledge with the utmost mutual ones, in addition to distribution network analysis, there are anomaly detection, machine learning, and cyber-security.

The different systems of detection of NTL are prearranged in three big groups: network-oriented, data-oriented, and hybrids. What differentiates data-oriented from ways of network-oriented is the utilizing of data of power grid (such as topology or measurements of the network). Ways of data-oriented make utilize of customer associated data solitary (such as type of consumer, consumption of energy). Hybrids are ways that utilize data from the two groups. Figure (2.1) shows these prime groups.

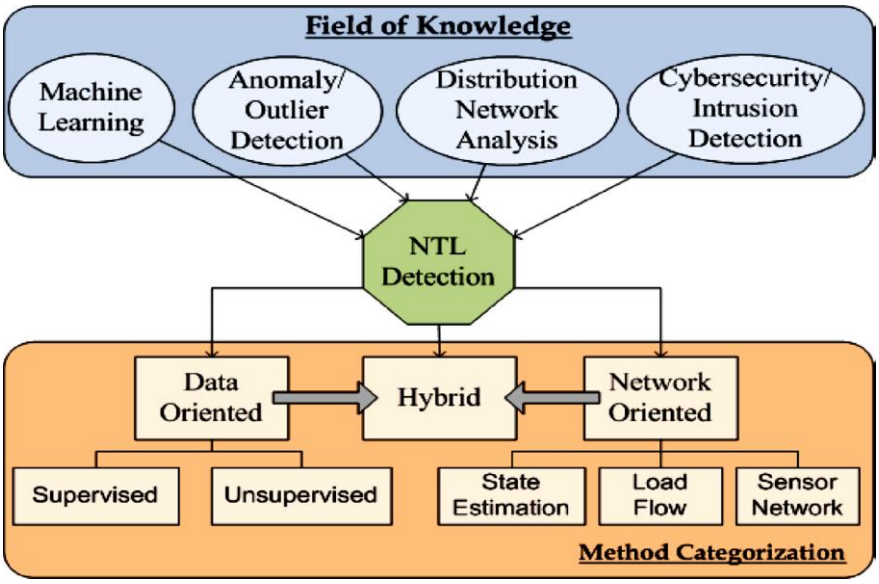

Figure (2.1): NTL detection methods categorization [41].

Data-oriented techniques are further classified to unsupervised and supervised. Ways that make no utilizing of labels are unsupervised, whereas ways that make utilizing of both labels (recognized positive/fraud and negative/not-fraud classes) are supervised. However, unsupervised methods does not use labels. Methods with single label are classified as unsupervised and typically fall under the unsupervised anomaly detection domain. These methods are applied when one of the two classes (e.g. fraud class) consists of tiny samples. Various fraud detection applications can found, for example (credit card fraud) apart from NTL detection. Both labels are known in this state; however, the lack of positive label (fraud) prevent supervised learning methods utilization.

Because they are founded on the analysis of network and the physical principles that define such schemes, network-oriented techniques usually disregard labels. These ways are divided relying on the prime perception/algorithm utilized, i.e. assessment of the state, flow of load, or exceptional sensors for detecting of fraudulence.

Hybrid techniques use conceptions from entire classes stated above. Such as, an estimation of the state way might be utilized on the level of Medium Voltage (MV) to detect NTL at Medium Voltage (MV)/ Low Voltage (LV) converter level. A way of arrangement of supervised can be utilized for pinpointing NTL at the customer level, afterwards detecting portions of the network with NTLs.

Briefing, the factors related to the detection of NTL are [41]:

- **Concept and Category:** The group and subgroup a particular work follows Figure (2.1).
- **Algorithm(s):** The prime algorithms utilized for detecting of NTL. In extreme cases, more than one algorithm can be utilized. For every work, even if they are solitarily utilized for comparing researches, nevertheless, wholly algorithms stated are arranged in a list.
- **Type of data (s):** The data requisite in various ways. This is a precarious factor when scheming a way of detection of NTL or selecting from current ones.
- **Size of the dataset:** Size of data set is considered small if it contains less than 100 customers, medium for 1000 to 10000 customers, and big for more than 10000 customers. The size of the data set is significant, as it delivers feedback on the scalability of systems of detection of NTL.
- **Features:** In several cases, raw data are first treated to extract features to be utilized for arrangement. There is no hint of which features must be utilized, though several studies utilize features for detecting of NTLs. It is possible to summarize features and related them with the type of data and algorithms, therefore making it is easier to select suitable features either utilizing field proficiency or feature choice algorithms. Table (2.1) shows the most important features used for detection of NTL.
- **Metrics:** Working metrics are utilized to evaluate the working of ways of detection NTL under different conditions and to relate systems. Some of the metrics are stated in the literature. It is possible to deliver a filled list of metrics, below an exclusive identifier, together with the aim for must (or mustn't) be utilized. Table (2.2) displays the descriptions of such metrics.
- **Response time:** It is the time needed for a system of detection of NTL to respond if a customer perpetrates fraudulence. This mustn't be mixed up with the time of classifier to generate an outcome specifying the comparative input data (which is hugely relied on device and coding). The time of response relies on the time required to attain the data of input.

More detailed information about algorithms, dataset size, and response time for various data-oriented, network-oriented, and hybrid techniques for ETD can be seen in Appendix A.

Table (2.1): Classifications of prime features utilized for detection of NTL [41].

Feature	Characterization
Standard Deviation, Average, Min/Man,	Typical statistics estimated for a definite time.
The factor of Energy / Power	This factor is definite as the mean of active (kW) to consuming of reactive energy (kVAr). Measurements of Immediate energy are requisite for this estimation. Great-solution (equal / less to 15.0 min) data should be utilized for a good assessment. The factor of power is the reactive energy (kVArh) expended in an interval of time to the active power (kWh) expended in the identical interval.
Factor of Load	The ratio amid the mean active power consuming (kW.h) to the extreme active power consuming (kW.h) for a definite interval of time (such as 30 days).
Streaks	The number of times the consuming curve reaches exceeding and underneath an average line (definite as a moving mean of the consuming curve).
Daily consuming to contractedenergy	The summation of consuming of active energy in an interval (kW.h) to the tapered power (kW)
Pearson coefficient	The Pearson coefficient of the consuming of the curve of active energy in a definite (usually big) period of time. The measuring of the Pearson coefficient shows how well a linear equation defines the relationship amid time and consuming of active energy.
Billed-consuming power coefficient	The variance of power billed (kW.h) to expended active energy (kW.h) to the contracted power (kW).
Predicted kWh	A calculation of the consuming of active energy (kW.h) specified via several models of prediction or the variance of this calculation and the calculated value.
Wavelet coefficients	The variance of the Wavelet coefficients estimated from the curve of consuming to be categorized and the Wavelet coefficients of preceding year's consuming curves.
Coefficients of Fourier	The variance of the coefficients of Fourier determined from the consuming curve to be categorized and the coefficients of Fourier of preceding year's consuming curves. Furthermore, the stage of the 1st five coefficients of Fourier of the consuming of the curve of active energy can be utilized.
Coefficients of Polynomial fit	The variance of the polynomial coefficients that greatest fits the consuming curve to be ordered and the polynomial coefficients that greatest fits preceding years' consuming curves.
Euclidean distance to mean consumer	The Euclidean distance of consuming of active energy curve to a consuming curve estimated as the average consumption of whole customers in the set of data.
The slope of the curve of consumption	The slope of the linear equation that greatest adjusts the curve of consuming of active energy time series.

Parts of Principal Component Analysis	The parts that are estimated from (PCA) Principal Component Analysis or (KPCA) Kernel Principal Component Analysis on the curves of consuming active energy. Not entirely of the components required to be utilized. The average of definite components might be utilized.
Fractional arrangement dynamic errors	Features that explain the variance amid summarized meter utilizing and actual time-consuming time sequence.
Ratio of Mismatch	The variance amid consuming estimated in the transformer of MV/LV and the summation of smart meter estimations and calculated technical losing to the minimal substation power.
Rates of Seasonal consuming	Overall consumer consuming (kW.h) in a definite season (such as winter) to the mean consuming of consumers on the similar substation at the identical season (such as winter).Overall consumer consuming (kW.h) in a definite season (such as winter) to consuming (kW.h) of a different season (such as summer).
Coefficients of Transform of Discrete Cosine	The k significant coefficients of Converting of Discrete Cosine.
Consuming drop associated to preceding	A decreasing of x% in consuming for the duration of an interval of time of length T in contrast to a previous period of time of the similar length or associated to the mean.
Calculated readings	Some of the meter readings that are calculated via effectiveness because of incapability to enter the meter.

Table (2.2): List of metrics utilized to estimate ways of detection of NTL [41].

Metric	Definition
Accurateness	Accuracy = TP+TN/TP+TN+FP+FN
Detection rate (DR)	DR = TP/TP+FN
Precision	Precision = TP/TP+FP
FPR	FPR = FP/FP+TN
TNR	TNR = TN/FP+TN
FNR	TNR = FN/FN+TP
F1 score	F1score = 2TP/2TP+FP+FN
(Area Under Curve)	(AUC) The area under the (Receiver Operating Curve) ROC of the double classifier.
Rate of Recognition	Rec.Rate = 1 − 0.5 (FP/N + FN/P)
Rate of Bayesian Detection	BDR = P(I) · DR/P(I) · DR+P(¬I) · FPR
Support	According to rule founded schemes. Definite as the data number on which a rule carried out to the overall of data number.
Time of Training (s)	This time (s) is necessary to train a system of detection of NTL.
Time of Arrangement (s)	The time (s), it takes a system of detection of NTL to categorize a case in point.

Undetected attack cost	Definite as the cost of the foulest probable unnoticed attack.
Mismatch of energy balance	Definite as the variance amid the totality of customer level active energy and substation level active energy
Rise of the average bill	Definite as the rise of the mean bill if the NTLs were dispersed amongst whole customers.
Cost of Normalized labour	Definite as the fee for studying wholly states categorized as NTL via the system of detection.
Index of Anomaly coverage	Definite as the ratio amid irregular customers underneath RTUs and the whole number of irregular customers.
Cost of RTU	Definite as the overall fee of obtaining RTUs
Deviance of Minutest detected	Definite as the minutest deviance (from a pre-identified usual profile) that can be detected.
Reduction in robbed of electricity	The reduction of robbed electricity once a definite FDS is exercised.

2.3.2 Descriptions and Classification of Data Types

In this part, the different types of data have been utilized in literature are prearranged in broad groupings. The prime aim for this classification is to confirm that investigators are not limited to definite types of data to choose their algorithm, but they are capable of selecting their system of detection of NTL relying on the available data. Figure (2.2) shows the data type of the pyramid.

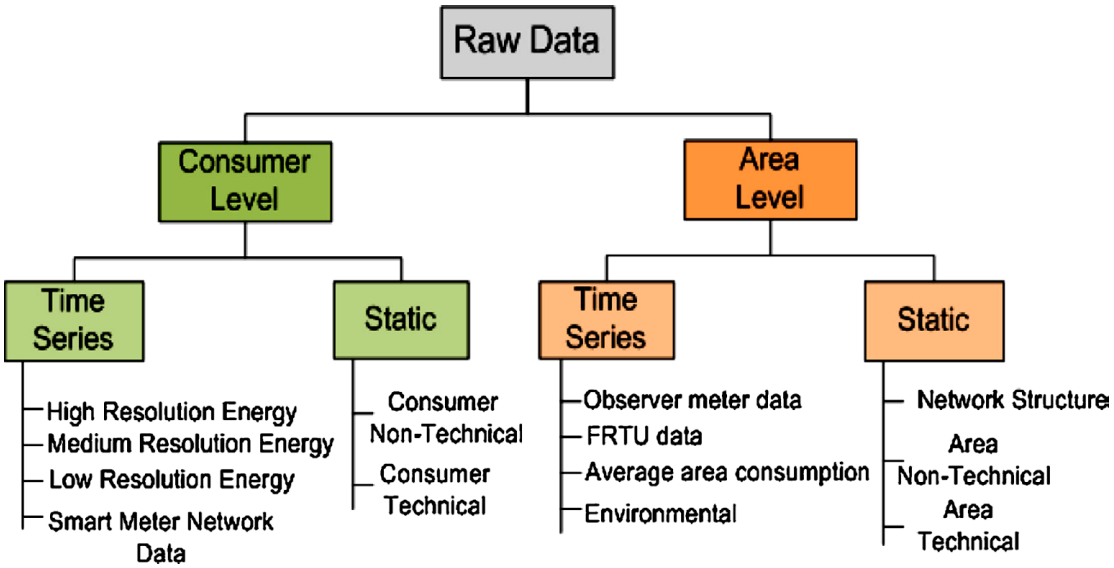

Figure (2.2): Data kind classification for implementations of detection NTL [41].

In the beginning, data prearranged relying on the place of their physical resource. Data relating to a region such as (topology of the network) are categorized as "Level of Area" data, whereas data relating to singular customers such as (estimations of active energy) are categorized as "Level of Customer "data. In the case of data belonging to the two groups mentioned above, it can be extra categorized as a series of time and static data. Data can then be ordered in extra granular groups, as shown in Table (2.3).

Table (2.3): Data used for detection of NTL [41].

Level of Customer	Series of Time	Great resolution Power	estimations of energy active/reactive with a resolution of time equivalent or lesser than 10.0 min.
		Moderate-resolution Power	estimations of energy active/reactive with a resolution of time amid 15.0 min. and 60 min.
		Little resolution Power	estimations of energy active/reactive with a resolution of time of 30 days or extra
		Data of network of Smart meter	non-power data of network of Smart meter (voltage, alarms line resistance, amperage)
	Static	Customer practical	data supplying practical features of the substructure of customer installed power(kW), (request contracted (kW), level of voltage, converter of power (kVA), applications number, stages number, remote system usage for heating of space.
		Customer non-practical	Data expressing the behavior of the customer, e.g. review remarks, geographic region, the action of finance.
Level of Area	Series of Time	Spectator meter data	measurements of power, voltage, and amperage of a meter mounted on the side of LV of the secondary converter of the network of distribution to deliver overall feeder estimations
		Data of Remote technical unit (RTU)	power, voltage, and amperage from RTUs set up in the network of MV or LV
		Average of consuming of area	Mean consuming of the observed region
		Ecological	generally temperature, but also might contain other parameters
	Static	Construction of Network	The topology of the network of LV or MV (may contain length and type of line). structure of Network associated data, such as the converter to which a customer is associated or the practical damages fraction
		Region practical	Data that describe a region from a practical point of sight (fraction of atypical customer per converter, number of converters in the region, a fraction of atypical customers in the region)

		Region non-practical	Data that describe a region from a typical/financial point of opinion (fraction of residences with a group of waste, mean earnings, a fraction of borrowed residences, a fraction of literates, activities of movement opposite to fraudulence in the region, fraction of residences with water, mean number of residents, fraction of residences with roadways)

2.3.3 Algorithms Utilized in ETD Systems

Systems for detecting fraudulent differ; in the meantime, they employ various data in various ways. Some systems show a modest construction, whereas others are extra difficult. NTL detection methods are classified as Data-Oriented, Hybrids, or Network Oriented. Each method can include several algorithms that are at the core of the fraud detection process.

A. Data-Oriented Techniques

These techniques are exclusively founded on machine learning and data analysis methods. They can be categorized into unsupervised and supervised. Figure (2.3) shows the unsupervised and supervised cases by the following aspects [41].

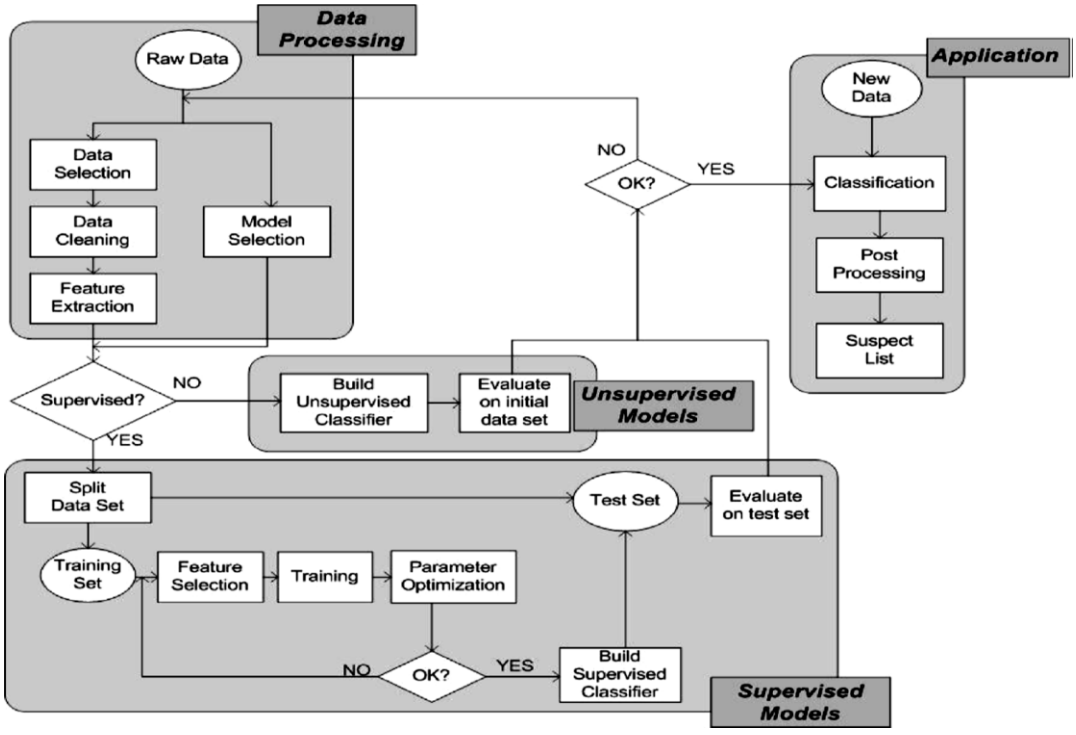

Figure (2.3): Data-oriented methods outline [41].

- *Choice of model and processing of data:* Assuming a group of raw data, the model utilized for detecting of NTL must be selected. The obtainability of categorized data orders the ways of selection, whether unsupervised or supervised, whereas data superiority/variability orders the algorithm to be utilized. The selection of the algorithm might reject several portions of the raw data (level of choice of data). The subsequent level contains cleaning of data (generally in the procedure of discovery of knowledge) and if essential, feature abstraction too.
- *Modelling:* This procedure is various for models of unsupervised and supervised. Unsupervised models don't utilize categorized data in the level of training, but solitary for assessment. Supervised techniques divide the group of data into trial and training. Afterwards describing the group of training (usually with cross-validation) choice of feature is regularly utilized for preparing the model. Choice of this factor makes use of metrics that can be estimated because of the obtainability of the label.
- *Application:* Modern data (not being a part of the group of "Raw Data") are utilized for confirming procedure and working of the model. Outcomes of the arrangement are extra managed for generating a doubtful list (a list includes the possibility of every customer obligating fraudulence). The level can be a part of the procedure of pilot of the model of NTL detection or its model. The procedure of Pilot on real-life places is of great significance in the state of response via physical meter inspections is obtainable.

The supervised methods include [41]:

- Optimum Path Forrest (OPF).
- Rule induction.
- Artificial Neural Network (ANN).
- Decision trees (DT).
- Nearest neighbor (k-NN).
- Support Vector Machine (SVM).
- Bayesian classifiers.
- Generalized Additive Model (GAM).

While the unsupervised methods include [41]:

- Self-Organizing Map (SOM).
- Clustering algorithms.
- Expert systems.
- Statistical control.
- Regression models.
- Outlier detection.
- Game-theoretic approaches.

B. Network-Oriented Techniques

Network-Oriented techniques pull data attained from sensors of the grid of distribution (next to smart meters) and take benefit of the physical rules that manage the entire electric network, to detect fraudulence. To one side from sensor data, they make utilize of network associated data, as the topology of network and customer converter/phase connectivity. Several studies utilize tools of the flow of power to measure the NTL size and to recognize its resource via examination the energy balance with a spectator meter. Additionally, many methods utilize assessment of the state of distribution and detection of insufficient data, as these methods tend to be more exact, even if not continuously probable to the appliance. The usage of sensors dedicated to detecting fraudulence is planned too. Algorithms of sensor placement have been explored, to compute the smallest sensors number and their place in the grid confirming detection of fraudulence. This category includes [41]:

i. Load flow approach
ii. State estimation approach
iii. Sensor network approach

C. Hybrid Methods

For detecting NTL with greater accurateness, Hybrid methods assume a grouping of techniques and algorithms termed above. In this respect, one possible direction is to utilize spectator meters organized with SVMs. The SVM production is substantiated with the spectator meter utilized to estimate the active power balance of the relation network. The algorithm calculates the active power balance discrepancy and the technical losses of the network. Suppose the discrepancy surpasses the predefined beginning, and the SVM generates a positive yield (or

number of positive yields as the system categorizes every day consuming as fraudulence or not). In that case, the customer is categorized as malicious and should be more examined [42].

Another direction is to combine Remote Terminal Units (RTUs) to discover NTL. The grid of distribution is primarily distributed in sub-networks relying on obtainability of RTU and dependability. The planned background identifies sub-networks with NTLs utilizing estimations from RTU sand smart meters and managing power of the network of distribution runs to estimate technical losses. In the case of divergence, the ratio surpasses a definite beginning, and meter damaging is supposed [43].

A various method is utilizing state assessment and analysis of variance (ANOVA). Smart metering data (power and voltage), RTU data (voltage angle and size of the High Voltage/Medium Voltage (HV/MV) substation secondary) and construction of network are requisite. An estimator of distribution state is applied by utilizing the collected smart meter consuming (per Medium Voltage/Low Voltage (MV/LV) transformer) as pseudo-measurement. The normalized residual trial is utilized for localizing irregular consuming at LV converter stage. Consequently, the problem of detection of NTL turns into a problem of detection of insufficient data, where significant overweight errors show possible NTL. The outcomes from ANOVA can be fed back to the state assessment unit for substituting better estimations instead of insufficient data [44].

In [45] the opposite process is projected, where irregularity detection (unsupervised outlier detection founded on a Gaussian distribution) is first carried out for computing the irregularities density (i.e. how often and to what amount fraudulence happens) per converter. To regulate the weight matrix of the estimator of state, this density is utilized, which computes transformer loading by utilizing load forecasts as pseudo-estimations. Losses of non-technical and technical types can be then calculated at the level of the converter.

2.4 Matters of ETD Progressive Metering Substructure in SGs

In the new society, the grid of power has converted to be a requisite. The conventional grid of power, which is unexpectedly still based on the projects more than one hundred years ago, can no longer be appropriate for society in the current days [46]. With the improvement of communication systems and information

technology, several countries have been reforming the old system of power into the smart grid, which is introduced with a two-way method of communication, the response of request real-time, outstanding dependability, sanctuary, and self-handling.

Inside the SG, the Advanced Metering Infrastructure (AMI) shows a significant role and is related to the everyday life of society most nearly [47]. The AMI updates the system of metering of electricity via by replacing mechanical meters with smart meters, which deliver two-way communications amid energy consumers and Utility Corporation. People cannot only read the meter data distantly with the AMI, but do several customized control and applied fine-coarse request-reply too [48]. Additionally, the real-time data gathered from the smart meters can develop the dependability of the grid of distribution by averting line overcrowding and production overloads [49]. The utility corporation can deliver dynamical electricity fee and quicker identification of outage thanks a lot to the AMI. Therefore, AMI has tempted excessive attentiveness from several participants, controllers, containing utility corporation, energy marketplaces.

Technologies of AMI are quickly surpassing the conventional technologies of reading of meter. Many smart meters are prepared in the domiciliary entirely over the world, for example, above 4.7 million smart meters utilized for advertising and other targets in Ontario, Canada [50].

2.4.1 Security Requirements

The AMI is a hierarchical construction and consist of many various networks interconnecting with each other. This can be depicted, as shown in Figure (2.4).

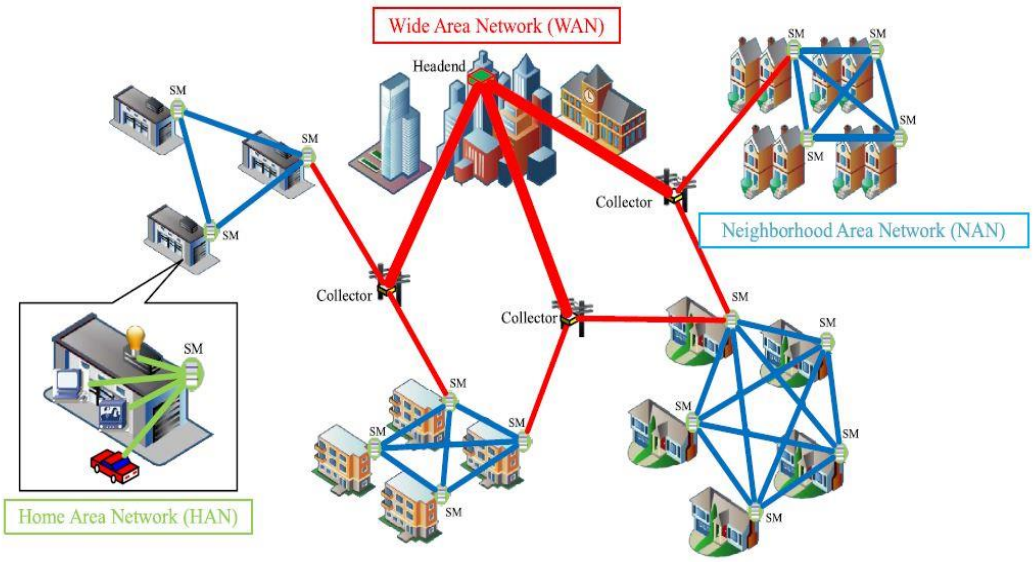

Figure (2.4): A simple AMI architecture [9].

Various shareholders in the AMI may have their specific requirements of security. Such as the consumers care about their secrecy of information and regular using of electrical energy; whereas the purpose of utility corporation to avoid the consumers from electricity-theft and deliver steady energy source.

Generally, those sensitive matters that want to be preserved in AMI can be categorized as [9]:

- *Controller data:* The command of control should be collected and applied via the smart meters totally and fittingly.
- *Smart meter data:* The smart meters data must not be reached via every illegal person.
- *Information of Bill:* The fee of electrical energy and the paid bill must not be operated in illegal persons.
- *Personal information of Customer:* The information contains a summary of daily using of electricity, information about credit card of consumer.

Therefore, the necessities of security for AMI can be categorized as follows [9]:

1. **Integrity:** Data transferred in AMI should be accurate and fittingly reflect the resource data deprived of any illegal handling.
2. **Privacy:** The persons can't conclude any secretive information from the available metering data.

3. **Confidentiality:** Sensitive information must solitary be gain access to via legal persons.
4. **Non-repudiation:** The persons can't reject reception everything, for example, price of converted electrical energy, that they have collected; and can't explain that they have directed some data, i.e., the quantity of electrical energy they have expended, which they don't drive.
5. **Availability:** Data in AMI must be available via official entities when they want the data.

2.4.2 Methods of ETD

Classification of techniques of detection for electricity-theft are presented in advanced metering infrastructure. The technique of ETD in AMI is categorized into three categories, relying on the plans of detection utilized in the working. These are the game theory-based, state-based, classification-based techniques, as shown in Figure (2.5).

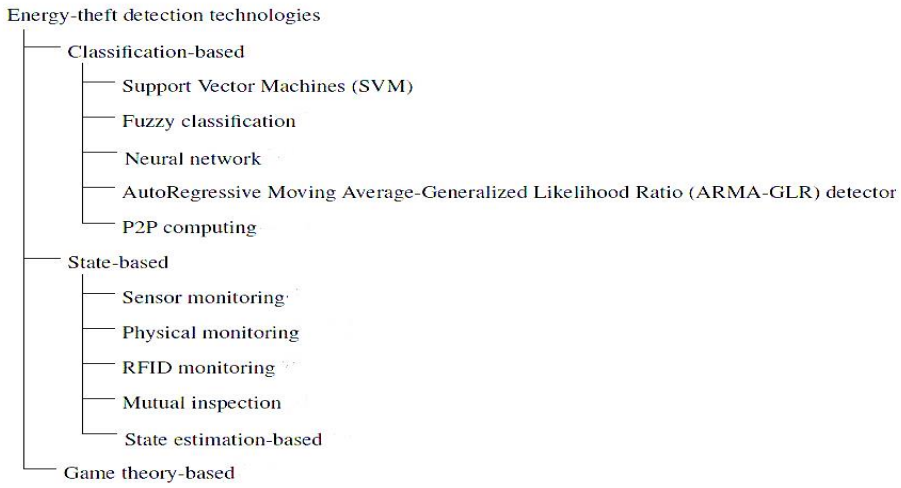

Figure (2.5): ETD methods in AMI [9].

A. Classification-based detection methods

Among the methods of electricity-theft detection are the classification-based detection methods that are defined as the categorization of load profile of electrical energy consumption of a consumer or a set of consumers during an interval of time, are considered of the most extensively utilized methods.

The primary process for classification-based detection of electricity-theft comprises of seven portions that are exposed in Figure (2.6). They include optimization of

factor and classifier training, data acquisition, extraction of feature, categorization, processing of data post, suspected consumer list production, and data preprocessing [9].

The prime concept of this procedure is to identify irregular manners of using energy from whole forms of using of energy founded on a trying dataset, including cases of the attack and normal class.

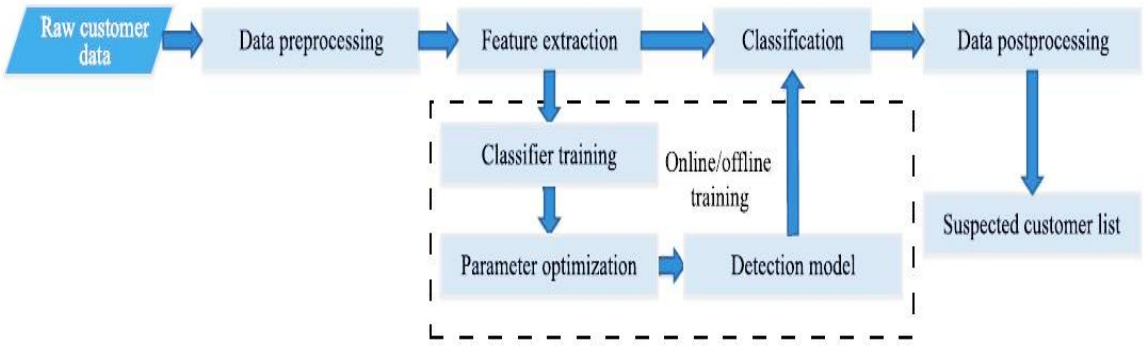

Figure (2.6): Main process for classification-based energy-theft detection [9].

B. Methods of state-based detection

State-based detection utilizes observing state to develop the rate of detection. The observing state can be resulting from networks of wireless sensor RFID, Advanced Metering Infrastructure, typical inspection, etc. Meanwhile networks of wireless sensor are simple to apply and inexpensive, they are general to help detecting electricity-theft [9].

In this respect, people investigated an Advanced Metering Infrastructure Intrusion Detection System (AMIIDS). It utilizes a mixture of information to fuse the sensors and data of consuming from a smart meter to more exactly detection of electricity-theft [51]. Indeed, it was shown that smart meters physical attack could be extended to a network attack via the addition of incorrect data. As a response, it is possible to use a customer attack model that reduces the number of compromised meters deprived of being exposed via preserving a snowballing load at the point of accumulation to which several homes are coupled [52].

Other proposals considered using schemes that apply the technology of Radio Frequency Identification (RFID) to assist the source companies of electrical energy to cope with their ammeter account managing and avoid electricity-theft [53].

Another state-based detection method is to use a paradigm change from the traditional technique of recognizing the illegitimate customer, via physical monitoring of the distribution feeder or assessment of the form of a load of wholly consumers. Suppose the calculated non-technical losses exceed 5.0% of the energy of distributed. In that case, the exterior control position will direct a controller sign to the smart meter Internal Control Station to cut the electrically powered deliver to the real consumers [54], [55].

C. Game theory-based detection methods

These methods have recently proposed and deliver a modern perception to resolve the matter of electricity-theft [56, 57]. It has shown that the electricity-theft and losses of combat as a non-zero summation Stackelberg game with an unfettered distributor. The distributor works as a frontrunner, and the consumers work as a supporter. The distributor can organize AMIs to develop the effectiveness of billing and the observing, therefore decrease the overall amount of non-technical losses because of theft. The effectiveness of stolen electrical energy detection rises with the equal employment of AMIs [56].

Another proposal has introduced to express the problem of detection of electricity-theft as a game amid the electrical energy thief and the distributor. For electricity thieves, they need to reduce the like cover of being detected to theft an amount of electricity which is definite previously. They can attain it via varying their function of the probability density of using electricity for the duration of the interval of measurement.

Oppositely, the distributor desires to capitalize on the possibility of detection of energy theft and regulate the optimum employment acquired via setting up of AMI. The Nash balance of the game is initiated as a function of probability density that protectors and assailants should select, so that direct measurement of AMI. However, the methods of detection based on game theory are not well-developed until now [57].

D. The comparison of methods

As the improvement of AMI offers different technologies that can be utilized to detect the electricity-theft so that it deserves much thanking. It is possible to show the assessment of the three types of systems of ETD mentioned above. Table (2.4)

shows the result of the assessment, from the point of vision of the rate of detection, procedure, cost, and false positive.

These methods of detection for electricity-theft possess their exclusive characteristics as follows [9]:

- The systems of detection based on game theory deliver a novel viewpoint to resolve the energy-theft. The electricity theft detection problem is expressed as a game amid the electricity thief and the utility of electric.
- The methods of detection based on the state can decrease the false positive and the rate of detection by the assistance of definite devices.
- Regularly schemes of detection based on classification take benefit of the data of consumption of energy gathered from the AMI. Technologies of data mining and machine learning are utilized to produce a worthy classifier founded on several example datasets.

Table (2.4): Assessment of systems of detection of energy [9].

System	Procedure	Rate of detection	False-positive	Cost
Classification- based	Artificial intelligence/ Machine learning	Moderate	Moderate	Moderate
State-based	Observing of state via definite device	Extraordinary	Little	Extraordinary
Game theory-based	Game Theory	Extraordinary	Moderate	Little

2.5 Deep Learning and Neural Networks

Deep learning (DL) is a machine learning research area that is founded on a specific kind of learning mechanism. The characterization of DL is according to the effort to generate a learning model at many levels, and the most profound levels take as input the outputs of preceding levels, converting them and continuously abstracting more. This vision on the levels of learning is motivated via the method; the brain processes information and learns, replying to exterior effects [58].

Neural networks are a set of algorithms stacked up together in a manner analogous to the human brain [59]. These networks interpret data through machine perception,

labelling of data or by clustering the data. The patterns recognized by these networks are structured in a vector into which the data of any kind be it images, sound, text, must be translated. The layers in the neural network are made up of nodes which mimic the functioning of a neuron in the human brain. These nodes are nothing but the area where computations happen. A node combines the inputs from the data with the weights which amplifies or dampens the input, hence giving the input a significance value with regards to the task which the network wants to learn (See Figure (2.7)).

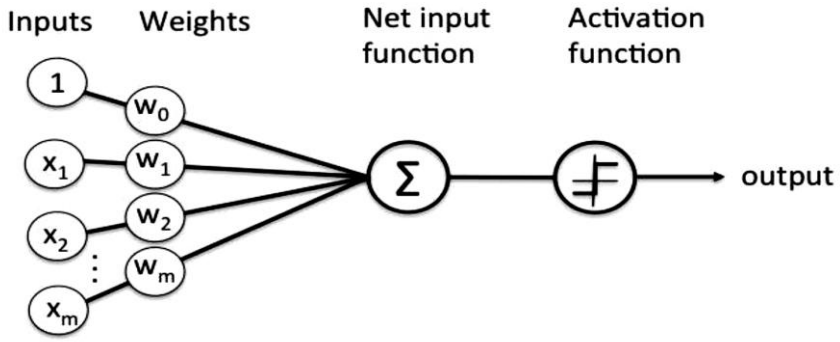

Figure (2.7): Basic Neural Network Structure [59].

CNNs in essence, are neural networks that use the convolution operation (in place of a completely connected layer) as one of its layers [60]. CNN's are a beneficial technology that has relied on difficulties where the data of input on which forecasts are to be prepared has a recognized grid-like topology like a time series (a 1-D grid) or an image (a 2-D grid) [61].

Currently, CNNs has controlled the machine vision space. The CNN comprises of a layer of output, layer of input, and several invisible layers. Usually, the invisible layers comprise of pooling layers, convolutional layers, and layers of normalization and wholly attached layers (Rectified Linear Unit-ReLU). Extra layers can be utilized for further complicated simulations, as shown in Figure (2.8).

The CNN construction has exposed excellent working in several computer vision and machine learning tasks. CNN trains and expects at an abstract level. This model of CNN is utilized lengthily in new implementations of machine learning because of its continuing record-breaking efficiency. Linear algebra is the principle of how these CNNs operate. Multiplication of matrix-vector is at the core of how data and weights are signified.

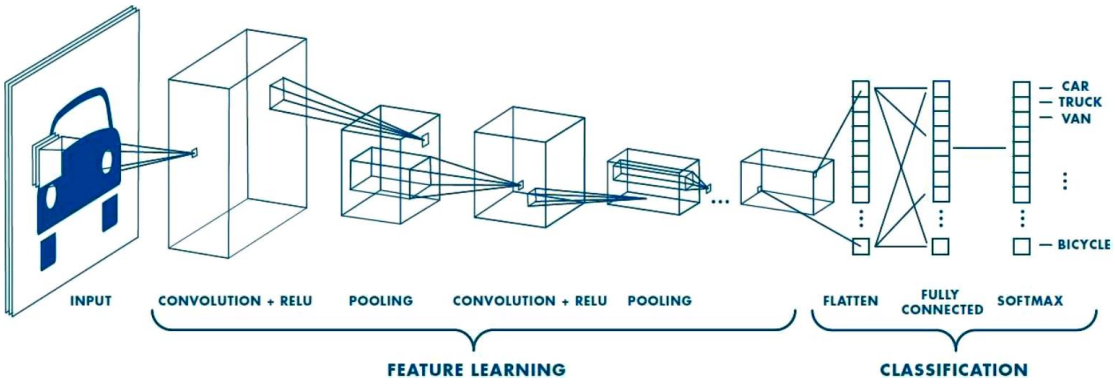

Figure (2.8): Typical CNN architecture [62].

2.6 The Blue Monkey Algorithm

The Blue Monkey (BM) algorithm is a recent algorithm development of metaheuristic founded on the working of blue monkey groups in nature. The BM algorithm classifies how many males in one collection. Usually, external the season of the breeding, the collections of BMs have solitary one mature male like other woodland guenons. Furthermore, it associates patas monkeys (Erythrocebus patas) [32]. The mathematical model and the motivation of the BM method are discussed in the following subsections.

A. Division of collection

The blue monkey algorithmic program simulates the behavior of the Blue Monkey in nature. Each group of the unit of monkey' region wanted to move through the area of search, all mentioned for modelling as communications. The Monkeys once being classified into teams who initiate to try to find locations of nutrition at long spaces region and more powerful monkey not amongst the conventional visibility choice. The male possesses slight to no communication with the young others. The young males must leave as quickly as probable, so that become more effective, Due to the regional nature of the male. They will go into a fight with the prevailing male of another family. If they win to lose that male, they can be the frontrunners of this family; therefore, they can suggest deliveries of food, location to live and socialization for young males. Usually, the collections of BMs possessing one male and a large number of babies and females [63].

B. Location update

The update location of every BM in the collection relies on the superlative location of BM in that collection, and this manner is defining via the subsequent calculations:

$$\mathbf{Rate_{i+1} = (Rate_i * 0.70) - (W_i - W_{leader}) * rand * (X_{best} - X_i)} \quad \mathbf{2.1}$$

$$\mathbf{X_{i+1} = Rate_{i+1} * rand + X_i} \quad \mathbf{2.2}$$

Where:

X_{best}: is the location of leader and rand is an arbitrary number amid [0.0,1.0]

X: is the location of monkey

W_i: is the weight of monkey at which wholly weights are arbitrary numbers amid [4.0, 6.0]

Rate: is the rate of power of monkey

W_{leader}: the weight of the leader

Furthermore, to modernize the kids of blue monkey, the following equations are utilized:

$$\mathbf{Rate^{ch}_{(i+1)} = (0.70 * Rate^{ch}_i) + (W^{ch}_{leader} - W^{ch}_i) * rand * (X^{ch}_{best} - X^{ch}_i)} \quad \mathbf{2.3}$$

$$\mathbf{X^{ch}_{(i+1)} = X^{ch}_i + Rate^{ch}_{(i+1)} * rand} \quad \mathbf{2.4}$$

Where:

X^{ch}: is the location of the child

W^{ch}_{leader}: is the weight of leader child

W^{ch}_i: is the weight of the child at which wholly weights are arbitrary numbers amid [4.0, 6.0]

$Rate^{ch}$: is the power rate of child

X^{ch}_{best}: is the location of leader child and "rand" signifying a random number amid [0.0,1.0]. The location must be modernized in every repetition.

C. Algorithm: Blue Monkey Optimization

1- Initialize the blue monkey and children population bi (i=1…n).

2- Initialize Power Rate (***Rate)*** and Weight (***W***), where (Rate ε [0, 1]), (W ε [4, 6]).

3- Distribute the blue monkeys randomly into teams (T), while all children in one team.

4- Calculate the fitness of children and all blue monkeys in each group.

5- For each group, select the worst value and the best value of fitness and store it in Current Best. While children select the best fitness.

6- t=1.

7- While (t≤ maximum number of iterations)

8. Swapping the worst fitness in each group by the best fitness in children group.

9- Update ***Rate*** and **X** position of all blue monkeys in each group by Equations 1 and 2.

10- Update ***Rate*** and **X** position of children by Equations 3 and 4.

11- Update the fitness of all blue monkey and children.

12- Update Current Best:

if New Best is better than Current Best Then Current Best=New Best.

13- t=t+1.

14- End While.

15- Return the optimal blue monkey.

2.7 Summary

In this chapter, the energy losses are typically classified into technical losses (TLs) and Non-technical losses (NTLs) have been discussed. One of the main parameters of the NTLs in networks of distribution is electricity-theft. The most common methods of NTLs detection are explained and categorized into three groups: Data-Oriented ways, Network-Oriented methods, and Hybrid methods. Each category can further be classified into many techniques. These methods have been explained. Next, related ETD issues in the AMI structures are also explored. Furthermore, DL and CNNs are briefly reviewed. Finally, the BM algorithm has been explained.

Chapter Three

The Proposed Electricity Theft Detection System

Chapter Three:

The Proposed Electricity Theft Detection System

3.1 Introduction

This chapter focuses on the design and implementation aspects of the proposed electricity theft detection system. A realistic electricity consumption dataset released by State Grid Corporation of China is used to train and test the models. This work is intended to identify electricity theft from the power consumption pattern of users, utilizing CNN-based deep learning and Blue Monkey techniques. This classifier model is trained utilizing a dataset consisting of daily power consumption data of both normal and fraudulent users in a supervised manner by several steps. First, the data is prepared by a data-preprocessing algorithm to train the model. The preprocessing step also involves synthetic data generation for better performance. At the next step, the proposed model is hyper-tuned and finally, the optimized model is evaluated via the test data. The overall system is depicted in Figure (3.1).

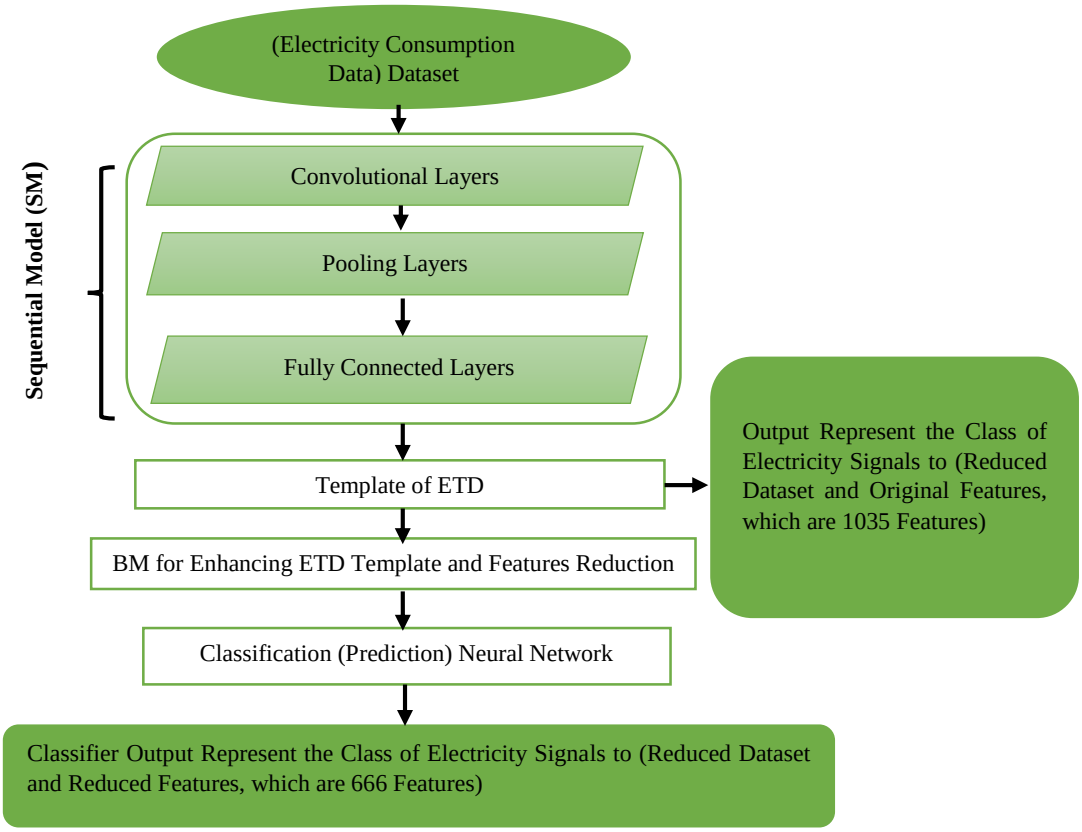

Figure (3.1): Architecture of the Proposed Model (CNN & BM).

3.2 Electricity Consumption Data

The research is performed on a series of real consumer electricity usage data, made accessible by the State Grid Corporation of China (SGCC). The Meta data information for this dataset is presented in Table 3.1. This dataset consists of 42,372 rows and 1,035 columns. The first column includes costumers' ID, and the second column includes pointer of prediction called "**Flag**" while the days' columns start from the third column up to the column (1,035). The Meta data types in the dataset are set of characters, numbers and missing or erroneous values called non-numeric (NaN). The numbers and missing or erroneous values represents the amount of electricity consumption (electricity signals) for each consumer for more than two years. In addition, the Meta data in the flag column are (zero and one) and it is referring to type of consumers (normal or thief), where the numbers of zeros in "Flag" column represents the normal consumer of electricity and the total number of them is (38,757). While the numbers of one in "Flag" column represents the thieves and the total number of them is (3,615). Finally, this means that the number (42,372) represents electricity consumers' data on electricity usage within 1,035 days (from Jan. 1, 2014 to Oct. 31, 2016), as shown in table (3.1).

Table (3.1): Metadata information of the electricity theft dataset.

Description	Value
Time window of data collection	1st January 2014 – 31th October 2016
Total number of consumers	42372
Number of normal users	38757
Number of aberrant user or electricity thieves	3615

In fact, Figure (3.2 (a)) gives an indication of energy utilization of usual use by a consumer in a month (i.e., August 2016). We observe that the data on electricity usage fluctuates day by day. From this 1-D data, it is difficult to catch the main features of electricity thieves and regular customers. However, will noted that the electricity usage of this consumer is seasonal if we plot the data in a 2-D way every week as seen in Figure (3.2 (b)), in which the electricity consumption peaks every week on day 3, while it often on day 5 in every week reaches the bottom (the exception is on the 2nd week, when there is the lowest consumption on day 6). In fact, we can have comparable results for the entire dataset (i.e., electricity

consumption data with 1,035 days). Showed only an excerpt of data from the entire dataset without too many repetitions. If aligned the energy usage data of all the 35 months together, will find that there is a level for most normal customers.

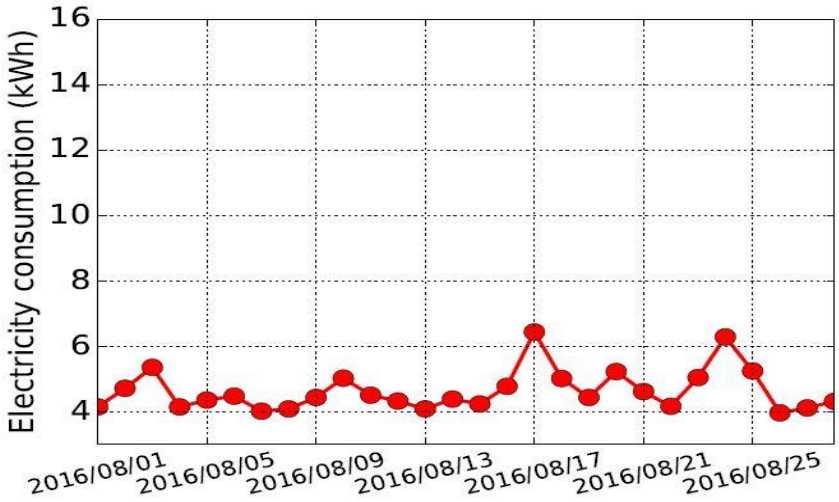

Figure (3.2 (a)): Electricity consumption (kWh) by date.

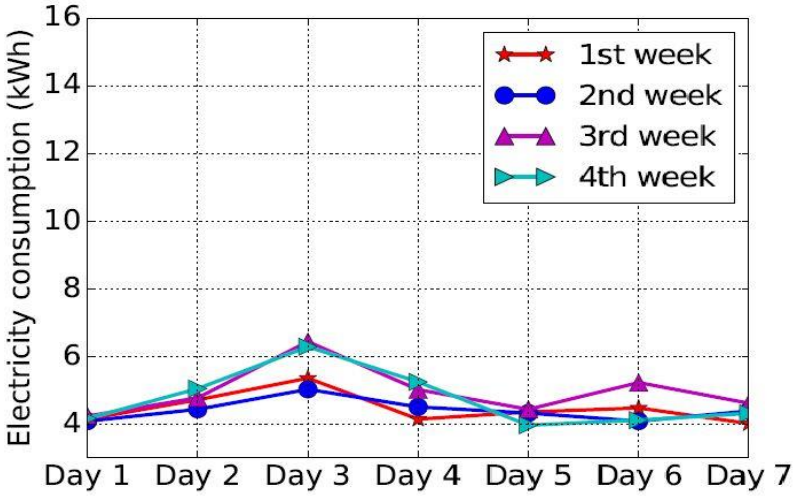

Figure (3.2 (b)): Electricity consumption (kWh) by week.

Figure (3.2): An example of electricity consumption of normal usage.

Figure (3.3), on the other hand, gives an example of the energy usage of a month-long electricity robbery. In addition, electricity consumption by date is mapped (as shown in Figure (3.3 (a))) and electricity consumption by week (as shown in Figure (3.3 (b))) in a similar way to Figure (3.2). As seen in Figure (3.3), found that the use of energy fluctuates regularly in the first two weeks (i.e., week 1 and week 2). For e.g., on day 3 and on day 6 in every week, electricity usage hits its peak.

However, as of the third week, there was a distinct loss of electricity usage and electricity consumption stayed at a low level after that.

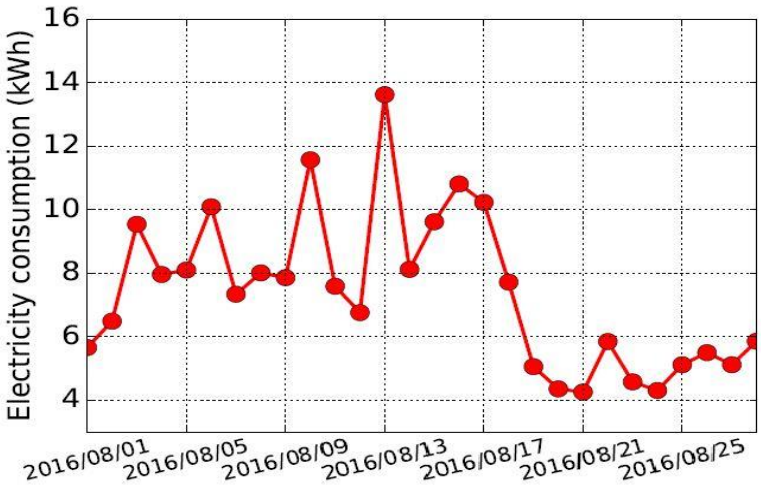

Figure (3.3 (a)): Electricity consumption (kWh) by date.

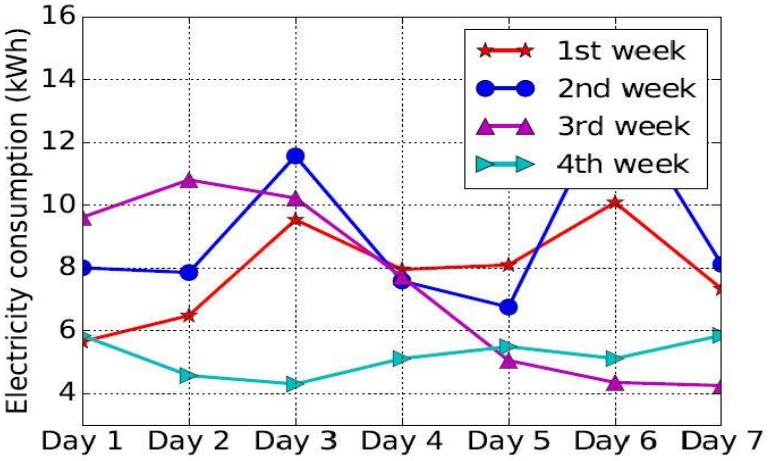

Figure (3.3 (b)): Electricity consumption (kWh) by Week.

Figure (3.3): An example of electricity consumption of electricity theft.

Electricity consumption data is generally acquired through smart meters or various sensors located at the user end. The data is then aggregated to any central location through a data communication network. In this scenario, there is a possibility of smart meter failure, sensor malfunctioning, or faults in data transmission and the storage server.

It is inherent that missing or erroneous data will be present in the electricity consumption datasets. In this dataset, numerous missing values are found. If those missing instances are just discarded, the size of the dataset shrinks considerably,

and thus reliable analysis becomes difficult. To avoid downsizing the dataset, the missing values are replaced with zeros to get rid of null or non-numeric values (NaN), because the neural network accepts numbers only, and these values are not defined, so these values are converted them into zeros until the neural network understands them.

The given dataset of electricity consumption passed in various stages of modifying to reduce it to be used in building operations of electricity theft detection templates using various algorithms. These stages are shown as follows:

1. Generating new dataset by replacing all null and Nan values in original dataset with zero.
2. Splitting new dataset into two parts, one part used for training (80%) and the other part used for testing (20%).
3. Reducing new dataset by dropping location and flag columns from new dataset. The reason is to reduce the complexity and the time as those two attributes will not be used in the proposed system.

3.3 Building of Electricity Theft Detection (ETD) Model

The proposed Electricity Theft Detection model can be summarized as follows: in the first step is the dataset passed in several modify operations to reduce it as discussed in (section (3.2)) then SM have been build using Algorithm (3.1). The third step is to build prediction model (ETD model), and this can be done by two operations. The first operation is by using SM, which described in Algorithm (3.2). The second operation using BM algorithm (see Algorithm (3.3)).

The input is Sequential Model with the reduced dataset and the output is the model of electricity theft detection with its accuracy and loss, where this algorithm consists of a set of fully connected layers, convolution layers and soft max layer to train and test the dataset (electricity consumption data).

3.3.1 Sequential Model (SM)

Sequential model is appropriate for a plain stack of layers where each layer has exactly one input tensor and one output tensor.

A Sequential model is not appropriate when:

- The model has multiple inputs or multiple outputs.
- Any of layers has multiple inputs or multiple outputs.
- You need to do layer sharing.
- You want non-linear topology (e.g., a residual connection, a multi-branch model).

In Algorithm (3.1), the input to this algorithm is the dataset and the output is the reduced dataset. The first step in this algorithm is defining the input shape to be compatible with the SM. After that, there are two cases to use SM: the first case is predicting electricity signals using original fully connected layers of the SM, this can be done by sending the electricity signals to the SM, which it in turn will return the classification of this electricity signals. The second case is using array and will use to build and train a given dataset.

Algorithm (3.1): Sequential Model (SM)
Input: Dataset.
Output: Reduced Dataset.

Step1: Define input shape entering to SM.
Step2: Define the number of Convolution layers and padding of each layers.
Step3: Define the number of fully connected layers and size of each layer.
Step4: Define the size of soft-max layer.
Step5: Define array to store the downloaded dataset.
Step6: For I=1 to number of convolution layers,
Execute Steps 7 to 13.
Step7: Apply three convolution operations on the input; the three operations of Convolution are 2D (3*3).
Step8: Apply flatten operation on the input.
Step9: Apply six activation operations on the input; five operations of activation is (Relu), final operation of activation is (SoftMax).
Step10: Apply four dropout operations on the input; two operations of dropout are (0.25) and the other two operations is (0.5).
Step11: Apply three Dense operations on the input, A=dense (layer1), B=dense (layer2), C=dense (2).
Step12: Reduced Dataset=concatenate (metrics)
Step13: Return (Reduced Dataset)

In addition, there are three types of convolution layers 2D (3*3). The three convolution layers (3*3) will be link with fully connected layers (activation, dropout), then linking the other fully connected layers (dense, activation, dropout) with each other. Finally, the resulting metrics from above operations used to generate the reduced dataset, as shown in Figure (3.4).

In this algorithm, used three fully connected layers (dense) which are (layer1, layer2, output) , where the value of (layer 1) is 128 nodes and the value of (layer 2) is 64 nodes, while the (output) value is represents the customers type either thief or normal customer depending on the dataset.

In addition, Applying four fully connected layers (dropout), where the values of those layers are selected after trying many possible values, it was found that the best values for them are (0.25 for the first two layers and 0.5 for the last two layers).

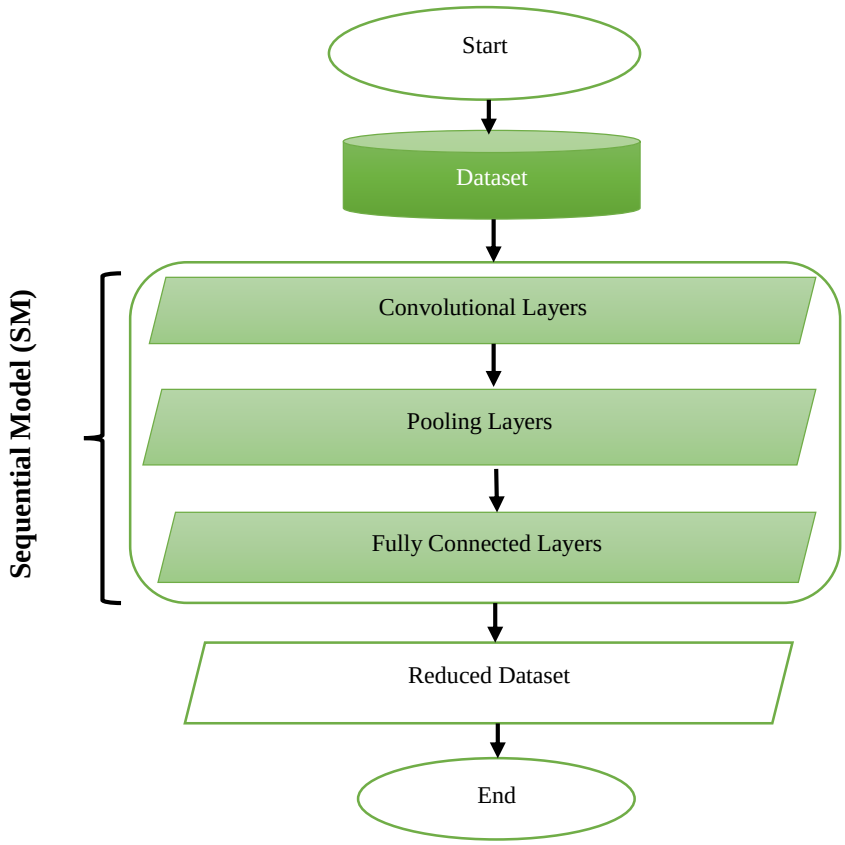

Figure (3.4): Sequential Model (SM).

The Electricity Theft Detection (ETD) model using SM is shown in **Algorithm (3.2**): the input is SM with the reduced dataset and the output is the model of electricity theft detection with its accuracy and loss. Where the dataset is divided into two parts, the first part is used for training, which are (80%) of the dataset and the second part is used for testing, which are (20%) of the dataset.

This algorithm is used to test the best configuration of neural in terms of number of layers and parameters, beginning with two layers and ending with four layers. The maximum dimension of layer is 128 nodes and the minimum one is 16 nodes. The best architecture is obtained with two layers, where the first layer contains 128 nodes, and the second layer contains 64 nodes as shown later in Chapter 4. This system using SM is described in Figure (3.5).

Algorithm (3.2): Building of Electricity Theft Detection (ETD) Template using SM.

Input: SM with the reduced dataset.

Output: The model of electricity theft detection with its accuracy and loss.

Step1: Generate Template.

Step2: Define sequential model.

Step3: Define the size of the input layer.

Step4: Define the dimension of each layer.

Step5: Create the layers using dimensions in step 4 and each lower layer dimension should be less than or equal to the dimension of the layer above it.

Step6: Define the dimension of output layer.

Step7: Define the optimizer used and its parameters (in this case, the Adam optimizer is used).

Step8: Train the model using training values.

Step9: Test the model using testing values.

Step10: Evaluate the model using testing values, then using the evaluated model to generate the score (accuracy, loss).

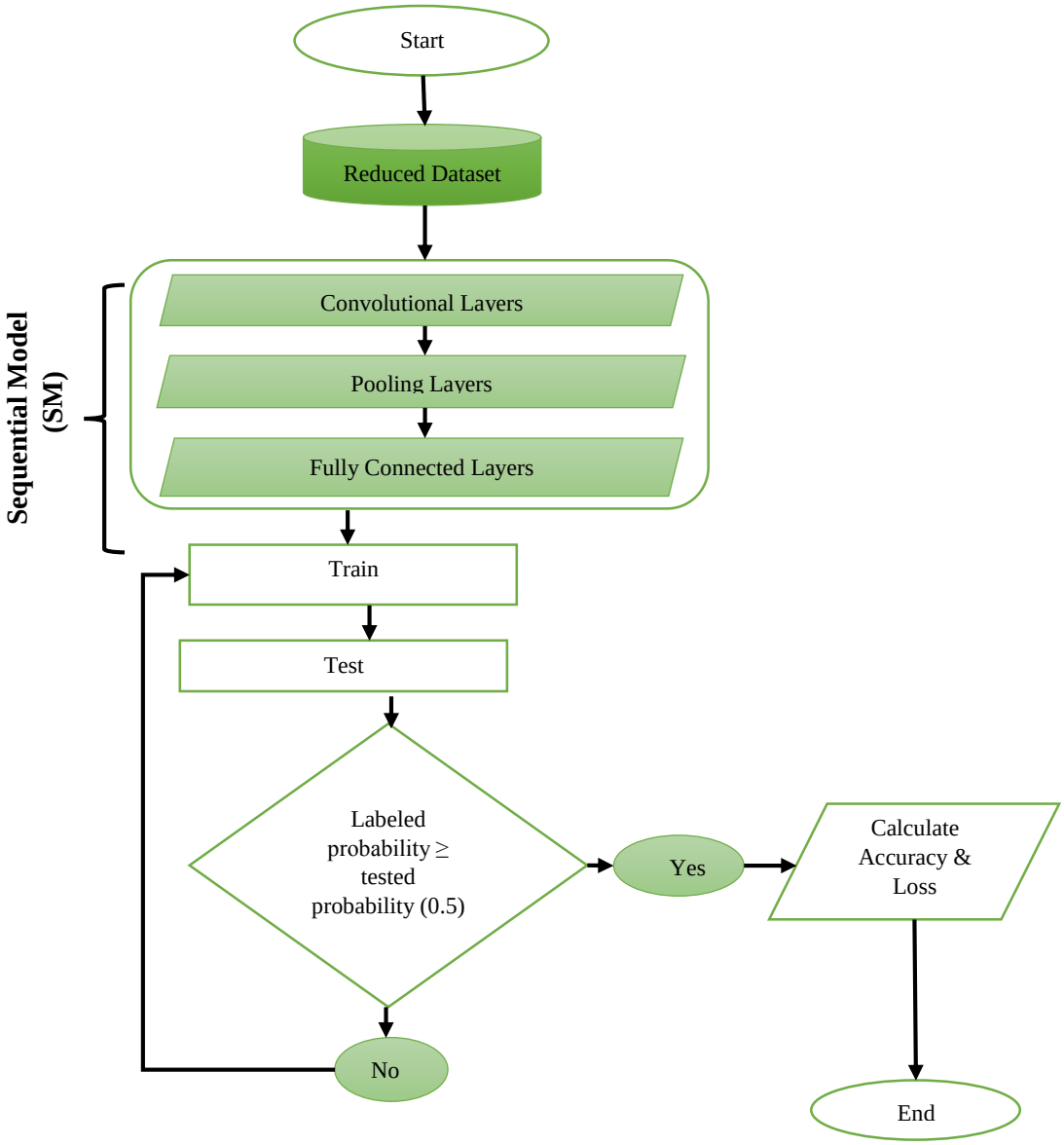

Figure (3.5): Building of Electricity Theft Detection (ETD) template using SM (CNN).

3.3.2 Blue Monkey Algorithm (BM)

BM represented as a function to enhance Electricity Theft Detection (ETD) template and return the solution of best location as describe in Figure (3.6). Where the input to this function is the electricity theft detection (ETD) template. The BM algorithmic program mimics behavior of the Blue Monkey. BM is a set of solutions for parents and children each one of parents and children has random values. The algorithm (3.3) describes the steps of BM.

Algorithm (3.3): Steps of BM.

Input: Electricity theft detection (ETD) template.

Output: Best template of electricity theft detection (ETD) template with its accuracy and loss.

Step1: Generate random template for population of BM (Parents, Childs).

Step2: Generate rand (uniform random).

Step3: Initialize location of monkey (X) and power rate of monkey (rate) and weight of monkey (W), where (rate $\in$[0, 1], W $\in$[4, 6]) for each solution (Parents and Childs).

Step4: Distribute the BM randomly into one team (t).

Step5: Calculate fitness for each solution (Parents, Childs).

Step6: For the Parents, select worst and best value of fitness and store it in current best, while Childs select the best fitness.

Step7: Set the number of iterations.

Step8: For i=1 to number of iterations,
Execute Steps (9 to 14) for each solution.

Step9: Update rate, X location of all BM (Parents) by Equations (2.1 and 2.2).

Step10: Update rate, X location of Childs by Equations (2.3 and 2.4).

Step11: Update fitness for the (Parents, Childs).

Step12: Update current best.

Step13: If the best location of Childs is better than the best location of the Parents then the best solution = current best location of Childs, else best solution = current best location of Parents.

Step14: Return the optimal BM (Best solution in the population).

In Algorithm (3.3), the input to this algorithm is a set of solutions each one represents the template of reducing the dataset and the ETD. The number of solutions used are 10 solutions, each solution has length of 1035 values generated randomly using zeros and ones. The rate and location can be calculated as mentioned in Section (2.6).

This template will be used in two steps: the first step is to reduce the dataset according to modify function, where the input to this function is the template from BM and original dataset (in case of building model). The output is a new dataset, which is less than original dataset. Then building model that has input shape equal to the size on new dataset. The second used is when there is a new electricity signal

to classify, it should reduce the values of the Electricity Signal according to the same template, so the Electricity Signal can be classified using this model. The important goal of building BM template is to reduce the number of features in the given dataset.

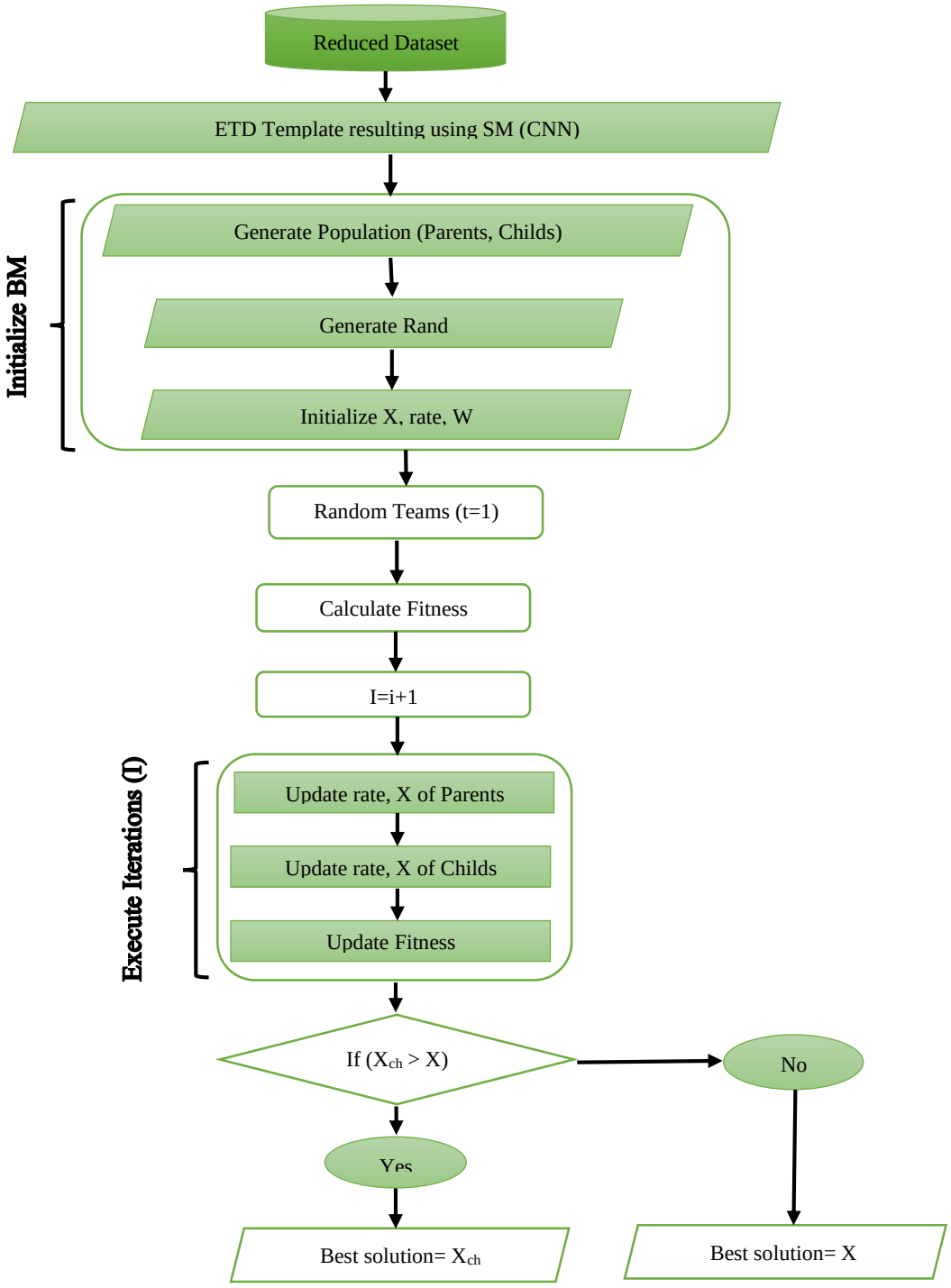

Figure (3.6): BM algorithm.

3.4 Calculating Accuracy using Fitness Value

The input is dataset, and the output is accuracy. In this part, the fitness is calculated and will be used to modify the dataset to produce the reduced dataset. The reduced dataset will be used for train and test. The resulting values from fitness, train and test used to evaluate the accuracy as shown in Algorithm (3.4).

Algorithm (3.4): Calculate Accuracy using Fitness Value.
Input: Dataset.
Output: Accuracy.
Step1: Define fitness (f) and calculate it.
Step2: Modify dataset using fitness value to reduce it.
Step3: Splitting reduced dataset to training (80%) and testing (20%).
Step4: Evaluate the accuracy of dataset.
Step5: Return (Accuracy).

3.5 Summary

The whole approach is working as follows:
The first step is to build SM using Algorithm (3.1). Then before building SM, the given dataset passed to several modify operations to reduce it. The second step is to build prediction model (ETD template), and this can be achieved by two operations. The first operation is by using SM. The second operation is by using BM algorithm.
The resulting template from BM is for enhance ETD Template and Features Reduction. The purpose of reduction process of dataset and features is to enhance the performance of prediction model.
For classifying a new electricity signal, the electricity signal at first is converted to array, then reducing the data of this array using the same template used in building the prediction model by sending electricity signal array and template to modify function. The output of modify function is the electricity signal after reducing its dataset according to the template.

Chapter Four

Results and Discussion

Chapter Four:

Results and Discussion

4.1 Introduction

One of the world's most important issues is the classification of Electricity signals, which has a wide range of practical applications. In this chapter, the proposed solution is tested to obtain and discuss the findings that demonstrate the system's efficacy. There are five parts of experiments conducted using electricity consumption dataset. The **first part** is the results of testing electricity signals classifier configuration. The **second part** represents the testing results of two layers (selected from the first part) on some number of nodes and selecting the best configuration. **The third part** represents the results of applying BM with best configuration of two selected layers, while the **fourth part** represents the results of accuracy and loss using CNN and BM model. The **final part** represents comparing results of loss and accuracy resulting from CNN and BM model with results of loss and accuracy resulting from the CNN model. The tests have been done using python programming language, under windows 7 PC with AMD E2-1800 (1.70 GHz) CPU and (4.00 GB) RAM.

4.2 Configuration of Classifier Part Experiments

The configuration of fully connected layers in term of number of layers and nodes was tested on several models beginning with two layers and ending with four layers. The maximum dimension of each layer is 128 nodes while the minimum dimension is 16 nodes.

In **Table (4.1)**, each row in this table represents the complete model configurations and results obtained from this model, while the columns represent the following: the first column is model number that represents the sequence of the model in the experiment, the second column represents the number of fully connected layers in the model, which ranges from two to four layers, while the third column represents the number of nodes in each layer and it is between (16-128). The fourth column contains the best accuracy of the selected architecture. The fifth column contains worst accuracy of the selected architecture. The sixth column represents the average

accuracy of the model. The seventh column contains the average training loss, where this value equal to the difference between true label and predicted label of the electricity signals inside training, which should be minimized as much as possible. The last column represents the time consumed for training.

Table (4.1): Settings and Results of Network Consists of Two-Four Layers.

No.	No. of Layers	No. of Nodes	Best Accuracy	Worst Accuracy	Average Accuracy	Average Loss	Average Time (Second)
1	2	128-128	0.918112094	0.916578174	0.91716815	0.266775871	13.1
2	2	64-64	0.91764009	0.914926231	0.915740407	0.283764297	10
3	2	32-32	0.915634215	0.915162265	0.91520946	0.285755402	10
4	2	16-16	0.915162265	0.915044248	0.915150464	0.295629337	10
5	2	128-64	0.918230116	0.916460156	0.917297935	0.26462948	10
6	2	128-32	0.91764009	0.915988207	0.916896755	0.276600096	10
7	2	128-16	0.916578174	0.915044248	0.915598828	0.277276114	10
8	2	64-32	0.91716814	0.914808273	0.915905619	0.281315494	10.1
9	2	64-16	0.91716814	0.915162265	0.916176987	0.270790696	10
10	2	32-16	0.915162265	0.91492623	0.915126864	0.296932735	10
11	3	128-128-128	0.918938041	0.916106224	0.91719175	0.255808522	10
12	3	64-64-64	0.916578174	0.915162265	0.915882021	0.272322157	10.1
13	3	32-32-32	0.915162265	0.915162265	0.915162265	0.295278683	10
14	3	16-16-16	0.915162265	0.915162265	0.915162265	0.291884622	10
15	3	128-64-32	0.917404115	0.916578174	0.916849566	0.265754673	10.1
16	3	64-32-16	0.915516198	0.915162265	0.915233052	0.283388585	10
17	4	128-128-128-128	0.916932166	0.915516198	0.916318583	0.267135677	10
18	4	64-64-64-64	0.916578174	0.914218307	0.915339243	0.287885103	10.1
19	4	32-32-32-32	0.915162265	0.915162265	0.915162265	0.293846384	10
20	4	16-16-16-16	0.915162265	0.915162265	0.915162265	0.290049273	10
21	4	128-64-32-16	0.91716814	0.915516198	0.916294992	0.271532404	10

In **Table (4.1)**, the selected architecture is the architecture that has two layers as it has the best average accuracy and the best consumed time.

4.3 Two Layers Experiments

This section explains the results presented in **Table (4.1**) for two layers in details and shows how to find the accuracy; loss and the time spent in performing each operation as shown below:

In **Table (4.2)**, the rows represent ten training rounds of the proposed model. The columns are as follow: the first column represents the number of nodes in each layer, which are 128 as well as 128 nodes for the first, and the second layer, respectively. The second column represents the number of training experiment. The third, fourth, and fifth columns contain the average loss, average accuracy and average time of each model respectively (as mentioned in Section (4.2)), where the average loss is equal to **0.2667759**, the average accuracy is equal to **0.9171682** and the average time is equal to **13.1** seconds. The last two columns represent the worst and best accuracy for ten training rounds in which the worst accuracy is equal to **0.9165782** and the best accuracy is equal to **0.918112094**.

Table (4.2): Two Layers Model (128-128).

No. of Nodes	Training No.	Loss	Accuracy	Time	Average Loss	Average Accuracy	Average Time (Second)
128-128	1	0.262991395	0.917640118	23	0.2667759	0.9171682	13.1
	2	0.255323413	0.918112094	24			
	3	0.249847542	0.91740413	11			
	4	0.263680607	0.916696191	11	Worst Accuracy	Best Accuracy	
	5	0.278808653	0.917758107	11			
	6	0.279120803	0.916578174	10			
	7	0.26984337	0.916696191	11	0.9165782	0.918112094	
	8	0.254106194	0.917286158	10			
	9	0.288878024	0.916814148	10			
	10	0.265158713	0.916696191	10			

In the **Table (4.2)**, the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

In **Table (4.3)**, the rows represent ten training rounds of the model. The columns are as follow: the first column represents the number of nodes in each layer, which are 64 and 64 nodes for the first, and the second layer, respectively. The second column contains the number of layers, which are two layers. The third column represents the sequence of each training round. The fourth, fifth and sixth columns represents the loss, accuracy and time of each model respectively (as mentioned in Section (4.2)). While the seventh, eighth and ninth columns represents the average of loss, accuracy and time for ten training rounds respectively, where the average loss is equal to **0.2837643**, the average accuracy is equal to **0.9157404** and the

average time is equal to **10** seconds. The last two columns represent the worst and best accuracy for ten training rounds in which the worst accuracy is equal to **0.9149262** and the best accuracy is equal to **0.91764009**.

Table (4.3): Two Layers Model (64-64).

No. of Nodes	No. of Layers	Training No.	Loss	Accuracy	Time	Average Loss	Average Accuracy	Average Time (Second)
64-64	2	1	0.288906723	0.91539824	10	0.2837643	0.9157404	10
		2	0.257616371	0.916342199	10			
		3	0.257586628	0.91764009	10			
		4	0.284564883	0.915044248	10	Worst Accuracy	Best Accuracy	
		5	0.300418824	0.915044248	10			
		6	0.296142191	0.915280223	10			
		7	0.257039577	0.916578174	10	0.9149262	0.91764009	
		8	0.331813544	0.914926231	10			
		9	0.258123368	0.916224182	10			
		10	0.305430859	0.914926231	10			

In **Table (4.3)**, the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

In **Table (4.4)**, the rows represent ten training rounds of the model. The columns are as follow: the first column represents the number of nodes in each layer, which are 32 and 32 nodes for the first, and the second layer, respectively. The second column contains the number of layers, which are two layers. The third column represents the sequence of each training round. The fourth, fifth and sixth columns represents the loss, accuracy and time of each model respectively (as mentioned in Section (4.2)). While the seventh, eighth and ninth columns represents the average of loss, accuracy and time for ten training rounds respectively, where the average loss is equal to **0.2857554**, the average accuracy is equal to **0.9152095** and the average time is equal to **10** seconds. The last two columns represent the worst and best accuracy for ten training rounds in which the worst accuracy is equal to **0.9151623** and the best accuracy is equal to **0.915634215.**

Table (4.4): Two Layers Model (32-32).

No. of Nodes	No. of Layers	Training No.	Loss	Accuracy	Time	Average Loss	Average Accuracy	Average Time (Second)
32-32	2	1	0.258463979	0.915634215	10	0.2857554	0.9152095	10
		2	0.263928115	0.915162265	10			
		3	0.303429216	0.915162265	10			
		4	0.305112481	0.915162265	10	Worst Accuracy	Best Accuracy	
		5	0.265692145	0.915162265	10			
		6	0.289280593	0.915162265	10			
		7	0.277891457	0.915162265	10	0.9151623	0.915634215	
		8	0.307486266	0.915162265	10			
		9	0.289243102	0.915162265	10			
		10	0.297026664	0.915162265	10			

In **Table (4.4),** it can be noticed that the accuracy is decreased compared to the results in **tables (4.2),** (**4.3),** and **tables (4.6 - 4.10)** of this experiment. The reason is that the size of the layers cannot cover the important features in the Electricity signals, or the features are expanded on the size of more than 32 nodes.

In **Table (4.5)**, the rows represent ten training rounds of the model. The columns are as follow: the first column represents the number of nodes in each layer, which are 16 and 16 nodes for the first, and the second layer, respectively. The second column contains the number of layers, which are two layers. The third column represents the sequence of each training round. The fourth, fifth and sixth columns represents the loss, accuracy and time of each model respectively (as mentioned in Section (4.2)). While the seventh, eighth and ninth columns represents the average of loss, accuracy and time for ten training rounds respectively, where the average loss is equal to **0.2956293**, the average accuracy is equal to **0.9151505** and the average time is equal to **10** seconds. The last two columns represent the worst and best accuracy for ten training rounds in which the worst accuracy is equal to **0.9150442** and the best accuracy is equal to **0.915162265**.

Table (4.5): Two Layers Model (16-16).

No. of Nodes	No. of Layers	Training No.	Loss	Accuracy	Time	Average Loss	Average Accuracy	Average Time (Second)
16-16	2	1	0.304809988	0.915162265	10	0.2956293	0.9151505	10
		2	0.28919518	0.915162265	10			
		3	0.289177895	0.915162265	10			
		4	0.290117979	0.915044248	10	Worst Accuracy	Best Accuracy	
		5	0.289201409	0.915162265	10			
		6	0.294315189	0.915162265	10			
		7	0.340786278	0.915162265	10	0.9150442	0.915162265	
		8	0.291540414	0.915162265	10			
		9	0.289279938	0.915162265	10			
		10	0.277869105	0.915162265	10			

In **Table (4.5),** it can be noticed that the accuracy is decreased compared to the results in **tables (4.2), (4.3),** and **tables (4.6 - 4.10)** of this experiment. The reason is that the size of the layers cannot cover the important features in the Electricity signals, or the features are expanded on the size of more than 16 nodes.

In the **Table (4.6**), the rows represent ten training rounds of the model. The columns are as follow: the first column represents the number of nodes in each layer, which are 128 and 64 nodes for the first, and the second layer, respectively. The second column contains the number of layers, which are two layers. The third column represents the sequence of each training round. The fourth, fifth and sixth columns represents the loss, accuracy and time of each model respectively (as mentioned in Section (4.2)). While the seventh, eighth and ninth columns represents the average of loss, accuracy and time for ten training rounds respectively, where the average loss is equal to **0.2646295**, the average accuracy is equal to **0.9172979** and the average time is equal to **10** seconds. The last two columns represent the worst and best accuracy for ten training rounds in which the worst accuracy is equal to **0.9164602** and the best accuracy is equal to **0.918230116**.

Table (4.6): Two Layers Model (128-64).

No. of Nodes	No. of Layers	Training No.	Loss	Accuracy	Time	Average Loss	Average Accuracy	Average Time (Second)
128-64	2	1	0.265759706	0.916460156	10	0.2646295	0.9172979	10
		2	0.254469693	0.91716814	10			
		3	0.260951936	0.916578174	10			
		4	0.253114253	0.917286158	10	Worst Accuracy	Best Accuracy	
		5	0.260221213	0.918112099	10			
		6	0.261828154	0.91716814	10			
		7	0.267765313	0.917050123	10	0.9164602	0.918230116	
		8	0.305615753	0.916932166	10			
		9	0.258682787	0.917994082	10			
		10	0.257885993	0.918230116	10			

In **Table (4.6)**, the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

In **Table (4.7)**, the rows represent ten training rounds of the model. The columns are as follow: the first column represents the number of nodes in each layer, which are 128 and 32 nodes for the first, and the second layer, respectively. The second column contains the number of layers, which are two layers. The third column represents the sequence of each training round. The fourth, fifth and sixth columns represents the loss, accuracy and time of each model respectively (as mentioned in section (4.2)). While the seventh, eighth and ninth columns represents the average of loss, accuracy and time for ten training rounds respectively, where the average loss is equal to **0.2766001**, the average accuracy is equal to **0.9168968** and the average time is equal to **10** seconds. The last two columns represent the worst and best accuracy for ten training rounds in which the worst accuracy is equal to **0.9159882** and the best accuracy is equal to **0.91764009**.

Table (4.7): Two Layers Model (128-32).

No. of Nodes	No. of Layers	Training No.	Loss	Accuracy	Time	Average Loss	Average Accuracy	Average Time (Second)
128-32	2	1	0.282110214	0.917286158	10	0.2766001	0.9168968	10
		2	0.255521417	0.917050123	10			
		3	0.280097693	0.916932166	10			
		4	0.390093982	0.91764009	10	Worst Accuracy	Best Accuracy	
		5	0.253987938	0.916460156	10			
		6	0.261899799	0.915988207	10			
		7	0.257771403	0.916932166	10	0.9159882	0.91764009	
		8	0.26001507	0.916578174	10			
		9	0.259834439	0.917404115	10			
		10	0.264669001	0.916696191	10			

In **Table (4.7**), the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

In **Table (4.8)**, the rows represent ten training rounds of the model. The columns are as follow: the first column represents the number of nodes in each layer, which are 128 and 16 nodes for the first, and the second layer, respectively. The second column contains the number of layers, which are two layers. The third column represents the sequence of each training round. The fourth, fifth and sixth columns represents the loss, accuracy and time of each model respectively (as mentioned in Section (4.2)). While the seventh, eighth and ninth columns represents the average of loss, accuracy and time for ten training rounds respectively, where the average loss is equal to **0.2772761**, the average accuracy is equal to **0.9155988** and the average time is equal to **10** seconds. The last two columns represent the worst and best accuracy for ten training rounds in which the worst accuracy is equal to **0.9150442** and the best accuracy is equal to **0.916578174**.

Table (4.8): Two Layers Model (128-16).

No. of Nodes	No. of Layers	Training No.	Loss	Accuracy	Time	Average Loss	Average Accuracy	Average Time (Second)
128-16	2	1	0.26906684	0.916578174	10	0.2772761	0.9155988	10
		2	0.269400626	0.915044248	10			
		3	0.384497672	0.915634215	10			
		4	0.263178349	0.916342199	10	Worst Accuracy	Best Accuracy	
		5	0.265783131	0.915988207	10			
		6	0.264063776	0.915162265	10			
		7	0.279936105	0.915162265	10	0.9150442	0.916578174	
		8	0.260340452	0.915280223	10			
		9	0.258155704	0.915634215	10			
		10	0.258338481	0.915162265	10			

In **Table (4.8**), the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

In **Table (4.9)**, the rows represent ten training rounds of the model. The columns are as follow: the first column represents the number of nodes in each layer, which are 64 and 32 nodes for the first, and the second layer respectively. The second column contains the number of layers, which are two layers. The third column represents the sequence of each training round. The fourth, fifth and sixth columns represents the loss, accuracy and time of each model respectively (as mentioned in Section (4.2)). While the seventh, eighth and ninth columns represents the average of loss, accuracy and time for ten training rounds respectively, where the average loss is equal to **0.2813155**, the average accuracy is equal to **0.9159056** and the average time is equal to **10.1** seconds. The last two columns represent the worst and best accuracy for ten training rounds in which the worst accuracy is equal to **0.9148083** and the best accuracy is equal to **0.91716814**.

Table (4.9): Two Layers Model (64-32).

No. of Nodes	No. of Layers	Training No.	Loss	Accuracy	Time	Average Loss	Average Accuracy	Average Time (Second)
64-32	2	1	0.287540257	0.91539824	10	0.2813155	0.9159056	10.1
		2	0.276144713	0.916578174	11			
		3	0.276388824	0.915988207	10			
		4	0.270792246	0.914808273	10	Worst Accuracy	Best Accuracy	
		5	0.30134666	0.916106224	10			
		6	0.282411814	0.915162265	10			
		7	0.270164937	0.91716814	10	0.9148083	0.91716814	
		8	0.2918697	0.916106224	10			
		9	0.280879855	0.915162265	10			
		10	0.275615931	0.916578174	10			

In **Table (4.9)**, the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

In **Table (4.10**), the rows represent ten training rounds of the model. The columns are as follow: the first column represents the number of nodes in each layer, which are 64 and 16 nodes for the first, and the second layer, respectively. The second column contains the number of layers, which are two layers. The third column represents the sequence of each training round. The fourth, fifth and sixth columns represents the loss, accuracy and time of each model respectively (as mentioned in Section (4.2)). While the seventh, eighth and ninth columns represents the average of loss, accuracy and time for ten training rounds respectively, where the average loss is equal to **0.2707907**, the average accuracy is equal to **0.916177** and the average time is equal to **10** seconds. The last two columns represent the worst and best accuracy for ten training rounds in which the worst accuracy is equal to **0.9151623** and the best accuracy is equal to **0.91716814**.

Table (4.10): Two Layers Model (64-16).

No. of Nodes	No. of Layers	Training No.	Loss	Accuracy	Time	Average Loss	Average Accuracy	Average Time (Second)
64-16	2	1	0.257065147	0.91587019	10	0.2707907	0.916177	10
		2	0.271605045	0.916814148	10			
		3	0.261602074	0.916342199	10			
		4	0.276430458	0.915162265	10	Worst Accuracy	Best Accuracy	
		5	0.271416187	0.916342199	10			
		6	0.266020715	0.91587019	10			
		7	0.265891701	0.916460156	10	0.9151623	0.91716814	
		8	0.265506864	0.91716814	10			
		9	0.27846238	0.916460156	10			
		10	0.293906391	0.915280223	10			

In **Table (4.10)**, the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

In **Table (4.11)**, the rows represent ten training rounds of the model. The columns are as follow: the first column represents the number of nodes in each layer, which are 32 and 16 nodes for the first, and the second layer, respectively. The second column contains the number of layers, which are two layers. The third column represents the sequence of each training round. The fourth, fifth and sixth columns represents the loss, accuracy and time of each model respectively (as mentioned in Section (4.2)). While the seventh, eighth and ninth columns represents the average of loss, accuracy and time for ten training rounds respectively, where the average loss is equal to **0.2969327**, the average accuracy is equal to **0.9151269** and the average time is equal to **10** seconds. The last two columns represent the worst and best accuracy for ten training rounds in which the worst accuracy is equal to **0.9149262** and the best accuracy is equal to **0.91516227**.

Table (4.11): Two Layers Model (32-16).

No. of Nodes	No. of Layers	Training No.	Loss	Accuracy	Time	Average Loss	Average Accuracy	Average Time (Second)
32-16	2	1	0.275003821	0.915162265	10	0.2969327	0.9151269	10
		2	0.277013129	0.91504425	10			
		3	0.310250549	0.91516227	10			
		4	0.316424083	0.91516227	10	Worst Accuracy	Best Accuracy	
		5	0.280209997	0.91516227	10			
		6	0.300594826	0.91516227	10			
		7	0.290422634	0.91516227	10	0.9149262	0.91516227	
		8	0.297431447	0.91516227	10			
		9	0.274115662	0.91492623	10			
		10	0.347861201	0.91516227	10			

In **Table (4.11),** it can be noticed that the accuracy is decreased compared to the results in **tables (4.2),** (**4.3),** and **tables (4.6 - 4.10)** of this experiment. The reason is that the size of the layers cannot cover the important features in the Electricity signals, or the features are expanded on the size of more than 32 and 16 nodes.

In **tables (4.2 - 4.11**), it can be noticed that using different size of layers is better than using same size because the average of accuracy in different sizes is better than the others. In addition, when the sizes of layers are less than 64 the average of the accuracy is decreased because this size cannot cover all possible features.

The results showed that the best configuration in two layers was as follows: the first layer consists of 128 nodes and the second layer is made up of 64 nodes, because the model has highest accuracy using these configurations. Therefore, there is no need to select the layers with high number of nodes because it will increase the complexity without enhancing the accuracy.

4.4 Applying BM with Best Configuration of Two Layers (128-64) in CNN Model

After testing all configuration of classifier part (Fully Connected Layers), it has been found that the best configuration is the two fully connected layers, where the first layer has 128 nodes, and the second layer has 64 nodes. This configuration has been selected to build the proposed classifiers (to reduce the features) with best

configuration of two layers using BM. In this section, showed the number of iterations (10 iterations) with numbers of solutions (10 solutions) used in BM template as follow.

Table (4.12), describing **Iteration 1**, the rows in this table represents twenty training rounds of the model. The columns of this table are as follow: The first, second, third, and sixth columns represent the sequence, loss, accuracy and time of the model respectively for each training round,

Where the average loss is equal to **0.276244349**, the average accuracy is equal to **0.918058997** and the average time is equal to **10** seconds. The worst accuracy is equal to **0.914336264** and the best accuracy is equal to **0.920354009.**

Table (4.12): Iteration 1 of Two Layers Model (128-64) using BM.

Iteration 1						
Model No.	**Loss**	**Accuracy**	**Time**	**Average Loss**	**Average Accuracy**	**Average Time (Second)**
1	0.262136159	0.916932166	10	0.276244349	0.918058997	10
2	0.298456314	0.915988207	10			
3	0.276000908	0.919292033	10			
4	0.266682555	0.918820083	10			
5	0.289418873	0.91716814	10			
6	0.299188606	0.914336264	10			
7	0.270489362	0.916814148	10			
8	0.260285954	0.91716814	10	Worst Accuracy	Best Accuracy	
9	0.265519834	0.918112099	10			
10	0.255108027	0.918348074	10			
11	0.260740122	0.920354009	10			
12	0.343229401	0.918112099	10			
13	0.265137493	0.916932166	10			
14	0.265496415	0.919763982	10	0.914336264	0.920354009	
15	0.33608697	0.918230116	10			
16	0.252497633	0.919528008	10			
17	0.263329231	0.918938041	10			
18	0.260608372	0.919881999	10			
19	0.25374069	0.91941005	10			
20	0.280734053	0.917050123	10			

In **Table (4.12**), the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

Table (4.13), describing **Iteration 2**, the rows in this table represents twenty training rounds of the model with the same description for table (4.12). The average loss is equal to **0.300249134**, the average accuracy is equal to **0.916961652** and the average time is equal to **10** seconds. The worst accuracy is equal to **0.906312704** and the best accuracy is equal to **0.920117974.**

Table (4.13): Iteration 2 of Two Layers Model (128-64) using BM.

Iteration 2						
Model No.	**Loss**	**Accuracy**	**Time**	**Average Loss**	**Average Accuracy**	**Average Time (Second)**
1	**0.256584632**	**0.917286158**	**10**	**0.300249134**	**0.916961652**	**10**
2	**0.262147619**	**0.918112099**	**10**			
3	**0.366652967**	**0.917994082**	**10**			
4	**0.369725381**	**0.919646025**	**10**			
5	**0.272879501**	**0.917286158**	**10**			
6	**0.271322374**	**0.918938041**	**10**			
7	**0.260333325**	**0.917404115**	**10**			
8	**0.272804955**	**0.917758107**	**10**	**Worst Accuracy**	**Best Accuracy**	
9	**0.259746232**	**0.920117974**	**10**			
10	**0.264320086**	**0.919292033**	**10**			
11	**0.345792039**	**0.912684381**	**10**			
12	**0.263699556**	**0.917758107**	**10**			
13	**0.392895122**	**0.906312704**	**10**			
14	**0.304622027**	**0.91764009**	**10**	**0.906312704**	**0.920117974**	
15	**0.266164093**	**0.919292033**	**10**			
16	**0.496647584**	**0.918466091**	**10**			
17	**0.277987163**	**0.914336264**	**10**			
18	**0.261000842**	**0.917758107**	**10**			
19	**0.280470004**	**0.911740422**	**10**			
20	**0.25918719**	**0.91941005**	**10**			

In **Table (4.13)**, the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

Table (4.14), describing **Iteration 3**, the rows in this table represents twenty training rounds of the model with the same description for table (4.12). The average loss is equal to **0.30032918**, the average accuracy is equal to **0.918088496** and the

average time is equal to **10** seconds. The worst accuracy is equal to **0.916578174** and the best accuracy is equal to **0.920235991.**

Table (4.14): Iteration 3 of Two Layers Model (128-64) using BM.

Iteration 3						
Model No.	**Loss**	**Accuracy**	**Time**	**Average Loss**	**Average Accuracy**	**Average Time (Second)**
1	**0.26269831**	**0.916932166**	**10**	**0.30032918**	**0.918088496**	**10**
2	**0.424173199**	**0.917758107**	**10**			
3	**0.286240864**	**0.919056058**	**10**			
4	**0.26597951**	**0.917758107**	**10**			
5	**0.272752143**	**0.918938041**	**10**			
6	**0.273075696**	**0.918112099**	**10**			
7	**0.3682971**	**0.919292033**	**10**			
8	**0.29024347**	**0.91764009**	**10**	**Worst Accuracy**	**Best Accuracy**	
9	**0.438468586**	**0.916578174**	**10**			
10	**0.263626807**	**0.919174016**	**10**			
11	**0.265620796**	**0.918702066**	**10**			
12	**0.365818325**	**0.918702066**	**10**			
13	**0.270651705**	**0.918112099**	**10**			
14	**0.321038176**	**0.916696191**	**10**	**0.916578174**	**0.920235991**	
15	**0.259300513**	**0.918584049**	**10**			
16	**0.266732829**	**0.918230116**	**10**			
17	**0.277736449**	**0.920235991**	**10**			
18	**0.306965165**	**0.917050123**	**10**			
19	**0.265823845**	**0.917522132**	**10**			
20	**0.261340107**	**0.916696191**	**10**			

In **Table (4.14)**, the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

Table (4.15), describing **Iteration 4**, the rows in this table represents twenty training rounds of the model with the same description for table (4.12). The average loss is equal to **0.308467002**, the average accuracy is equal to **0.918283185** and the average time is equal to **10** seconds. The worst accuracy is equal to **0.91539824** and the best accuracy is equal to **0.920707941.**

Table (4.15): Iteration 4 of Two Layers Model (128-64) using BM.

Iteration 4						
Model No.	**Loss**	**Accuracy**	**Time**	**Average Loss**	**Average Accuracy**	**Average Time (Second)**
1	**0.273442773**	**0.917050123**	**10**	**0.308467002**	**0.918283185**	**10**
2	**0.293746553**	**0.916106224**	**10**			
3	**0.263142323**	**0.918938041**	**10**			
4	**0.253592572**	**0.919528008**	**10**			
5	**0.260511446**	**0.920707941**	**10**			
6	**0.270317291**	**0.918702066**	**10**			
7	**0.599808542**	**0.919056058**	**10**			
8	**0.267110954**	**0.917994082**	**10**	**Worst Accuracy**	**Best Accuracy**	
9	**0.257338264**	**0.918348074**	**10**			
10	**0.256245187**	**0.91941005**	**10**			
11	**0.284786543**	**0.918938041**	**10**			
12	**0.27078401**	**0.919056058**	**10**			
13	**0.263804258**	**0.918348074**	**10**			
14	**0.256071756**	**0.918112099**	**10**	**0.91539824**	**0.920707941**	
15	**0.59147545**	**0.918466091**	**10**			
16	**0.29650002**	**0.919292033**	**10**			
17	**0.295266796**	**0.916342199**	**10**			
18	**0.267126147**	**0.91539824**	**10**			
19	**0.385551135**	**0.917994082**	**10**			
20	**0.262718027**	**0.917876124**	**10**			

In **Table (4.15)**, the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

Table (4.16), describing **Iteration 5**, the rows in this table represents twenty training rounds of the model with the same description for table (4.12). The average loss is equal to **0.302507543**, the average accuracy is equal to **0.918660757** and the average time is equal to **10** seconds. The worst accuracy is equal to **0.916106224** and the best accuracy is equal to **0.920589983.**

Table (4.16): Iteration 5 of Two Layers Model (128-64) using BM.

Iteration 5						
Model No.	**Loss**	**Accuracy**	**Time**	**Average Loss**	**Average Accuracy**	**Average Time (Second)**
1	**0.260627226**	**0.919174016**	**10**	**0.302507543**	**0.918660757**	**10**
2	**0.772995302**	**0.919174016**	**10**			
3	**0.256456371**	**0.918938041**	**10**			
4	**0.284027126**	**0.917994082**	**10**			
5	**0.271899974**	**0.919646025**	**10**			
6	**0.292941332**	**0.91764009**	**10**			
7	**0.275698556**	**0.916578174**	**10**			
8	**0.354190652**	**0.91764009**	**10**	**Worst Accuracy**	**Best Accuracy**	
9	**0.263876286**	**0.917404115**	**10**			
10	**0.259846044**	**0.919763982**	**10**			
11	**0.272589616**	**0.918938041**	**10**			
12	**0.263807164**	**0.919528008**	**10**			
13	**0.260979471**	**0.919174016**	**10**			
14	**0.260379289**	**0.918584049**	**10**	**0.916106224**	**0.920589983**	
15	**0.252775537**	**0.920589983**	**10**			
16	**0.267134286**	**0.918938041**	**10**			
17	**0.317843815**	**0.920000017**	**10**			
18	**0.255102373**	**0.918466091**	**10**			
19	**0.318131085**	**0.918938041**	**10**			
20	**0.288849355**	**0.916106224**	**10**			

In **Table (4.16)**, the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

Table (4.17), describing **Iteration 6**, the rows in this table represents twenty training rounds of the model with the same description for table (4.12). The average loss is equal to **0.29349886**, the average accuracy is equal to **0.917740414** and the average time is equal to **10** seconds. The worst accuracy is equal to **0.912920356** and the best accuracy is equal to **0.920235991.**

Table (4.17): Iteration 6 of two Layers Model (128-64) using BM.

Iteration 6						
Model No.	**Loss**	**Accuracy**	**Time**	**Average Loss**	**Average Accuracy**	**Average Time (Second)**
1	**0.255899991**	**0.918348074**	**10**	**0.29349886**	**0.917740414**	**10**
2	**0.262030279**	**0.918938041**	**10**			
3	**0.265712581**	**0.918702066**	**10**			
4	**0.256796517**	**0.920235991**	**10**			
5	**0.289303321**	**0.918466091**	**10**			
6	**0.272421877**	**0.918820083**	**10**			
7	**0.637002913**	**0.919528008**	**10**			
8	**0.286742511**	**0.918112099**	**10**	**Worst Accuracy**	**Best Accuracy**	
9	**0.266422568**	**0.917522132**	**10**			
10	**0.288294993**	**0.912920356**	**10**			
11	**0.276304597**	**0.916460156**	**10**			
12	**0.270508157**	**0.916578174**	**10**			
13	**0.260472651**	**0.918112099**	**10**			
14	**0.34786384**	**0.914690256**	**10**	**0.912920356**	**0.920235991**	
15	**0.290259311**	**0.917876124**	**10**			
16	**0.266339258**	**0.918348074**	**10**			
17	**0.269448951**	**0.917522132**	**10**			
18	**0.268741278**	**0.918230116**	**10**			
19	**0.272670062**	**0.918348074**	**10**			
20	**0.266741538**	**0.917050123**	**10**			

In **Table (4.17**), the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

Table (4.18), describing **Iteration 7**, the rows in this table represents twenty training rounds of the model with the same description for table (4.12). The average loss is equal to **0.30776233**, the average accuracy is equal to **0.917834809** and the average time is equal to **10** seconds. The worst accuracy is equal to **0.911268413** and the best accuracy is equal to **0.920000017.**

Table (4.18): Iteration 7 of Two Layers Model (128-64) using BM.

Iteration 7						
Model No.	**Loss**	**Accuracy**	**Time**	**Average Loss**	**Average Accuracy**	**Average Time (Second)**
1	**0.500642313**	**0.919056058**	**10**	**0.30776233**	**0.917834809**	**10**
2	**0.29355611**	**0.919174016**	**10**			
3	**0.380159939**	**0.918466091**	**10**			
4	**0.417346588**	**0.917994082**	**10**			
5	**0.261931958**	**0.918230116**	**10**			
6	**0.280580217**	**0.919056058**	**10**			
7	**0.259767462**	**0.919646025**	**10**			
8	**0.25704296**	**0.916342199**	**10**	**Worst Accuracy**	**Best Accuracy**	
9	**0.300944026**	**0.911268413**	**10**			
10	**0.26527152**	**0.917522132**	**10**			
11	**0.312285964**	**0.919056058**	**10**			
12	**0.352429105**	**0.91764009**	**10**			
13	**0.26729697**	**0.916696191**	**10**			
14	**0.281900958**	**0.918702066**	**10**	**0.911268413**	**0.920000017**	
15	**0.267868711**	**0.91716814**	**10**			
16	**0.272329986**	**0.918938041**	**10**			
17	**0.289064607**	**0.915516198**	**10**			
18	**0.350622295**	**0.918938041**	**10**			
19	**0.273199571**	**0.920000017**	**10**			
20	**0.271005346**	**0.917286158**	**10**			

In **Table (4.18**), the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

Table (4.19), describing **Iteration 8**, the rows in this table represents twenty training rounds of the model with the same description for table (4.12). The average loss is equal to **0.288925329**, the average accuracy is equal to **0.917557526** and the average time is equal to **10** seconds. The worst accuracy is equal to **0.904070795** and the best accuracy is equal to **0.919763982.**

Table (4.19): Iteration 8 of Two Layers Model (128-64) using BM.

Iteration 8						
Model No.	**Loss**	**Accuracy**	**Time**	**Average Loss**	**Average Accuracy**	**Average Time (Second)**
1	**0.28800324**	**0.918348074**	**10**	**0.288925329**	**0.917557526**	**10**
2	**0.269769874**	**0.917994082**	**10**			
3	**0.268389145**	**0.919763982**	**10**			
4	**0.266029415**	**0.91941005**	**10**			
5	**0.349088648**	**0.91941005**	**10**			
6	**0.253757812**	**0.919292033**	**10**			
7	**0.270642976**	**0.91941005**	**10**			
8	**0.292369911**	**0.916342199**	**10**	**Worst Accuracy**	**Best Accuracy**	
9	**0.389372155**	**0.918820083**	**10**			
10	**0.275913555**	**0.917994082**	**10**			
11	**0.275103527**	**0.919056058**	**10**			
12	**0.253536955**	**0.917758107**	**10**			
13	**0.267949272**	**0.918584049**	**10**			
14	**0.337432662**	**0.904070795**	**10**	**0.904070795**	**0.919763982**	
15	**0.280432781**	**0.91941005**	**10**			
16	**0.275843732**	**0.919292033**	**10**			
17	**0.265476515**	**0.917876124**	**10**			
18	**0.26442283**	**0.918702066**	**10**			
19	**0.319655773**	**0.913038373**	**10**			
20	**0.315315793**	**0.916578174**	**10**			

In **Table (4.19)**, the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

Table (4.20), describing **Iteration 9**, the rows in this table represents twenty training rounds of the model with the same description for table (4.12). The average loss is equal to **0.319363779**, the average accuracy is equal to **0.918123892** and the average time is equal to **10** seconds. The worst accuracy is equal to **0.91587019** and the best accuracy is equal to **0.919763982.**

Table (4.20): Iteration 9 of Two Layers Model (128-64) using BM.

Iteration 9						
Model No.	**Loss**	**Accuracy**	**Time**	**Average Loss**	**Average Accuracy**	**Average Time (Second)**
1	**0.424743643**	**0.918466091**	**10**	**0.319363779**	**0.918123892**	**10**
2	**0.505186644**	**0.919763982**	**10**			
3	**0.285795381**	**0.917050123**	**10**			
4	**0.293457204**	**0.918112099**	**10**			
5	**0.524328275**	**0.916578174**	**10**			
6	**0.295424113**	**0.919056058**	**10**			
7	**0.276254397**	**0.918938041**	**10**			
8	**0.257465389**	**0.918820083**	**10**	**Worst Accuracy**	**Best Accuracy**	
9	**0.263723638**	**0.917522132**	**10**			
10	**0.257740105**	**0.918702066**	**10**			
11	**0.267956655**	**0.918938041**	**10**			
12	**0.2572897**	**0.918938041**	**10**			
13	**0.315698198**	**0.91716814**	**10**			
14	**0.267901669**	**0.918938041**	**10**	**0.91587019**	**0.919763982**	
15	**0.313886412**	**0.91587019**	**10**			
16	**0.313779082**	**0.918230116**	**10**			
17	**0.274871416**	**0.917286158**	**10**			
18	**0.286600947**	**0.916224182**	**10**			
19	**0.27421232**	**0.919174016**	**10**			
20	**0.430960388**	**0.918702066**	**10**			

In **Table (4.20**), the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

Table (4.21), describing **Iteration 10**, the rows in this table represents twenty training rounds of the model with the same description for table (4.12). The average loss is equal to **0.324672409**, the average accuracy is equal to **0.910678467** and the average time is equal to **10** seconds. The worst accuracy is equal to **0.797286153** and the best accuracy is equal to **0.918938041.**

Table (4.21): Iteration 10 of Two Layers Model (128-64) using BM.

Iteration 10						
Model No.	**Loss**	**Accuracy**	**Time**	**Average Loss**	**Average Accuracy**	**Average Time (Second)**
1	**0.265011785**	**0.917758107**	**10**	**0.324672409**	**0.910678467**	**10**
2	**0.27306657**	**0.918112099**	**10**			
3	**0.29941101**	**0.912684381**	**10**			
4	**0.272518669**	**0.916932166**	**10**			
5	**0.265705152**	**0.915280223**	**10**			
6	**0.300083631**	**0.918938041**	**10**			
7	**0.656314497**	**0.797286153**	**10**			
8	**0.28413875**	**0.913982272**	**10**	**Worst Accuracy**	**Best Accuracy**	
9	**0.264199569**	**0.918112099**	**10**			
10	**0.30869456**	**0.918702066**	**10**			
11	**0.262819471**	**0.918702066**	**10**			
12	**0.318708567**	**0.915044248**	**10**			
13	**0.270718829**	**0.916106224**	**10**			
14	**0.313190766**	**0.918584049**	**10**	**0.797286153**	**0.918938041**	
15	**0.261128668**	**0.918938041**	**10**			
16	**0.514016596**	**0.913274348**	**10**			
17	**0.266173484**	**0.916578174**	**10**			
18	**0.524251643**	**0.917758107**	**10**			
19	**0.276477728**	**0.918348074**	**10**			
20	**0.296818235**	**0.912448406**	**10**			

In **Table (4.21)**, the values of accuracy and time are convergent, and this indicates the presence of stability in the network.

The detection accuracy with varying number of iterations starting from (1 to 10) is shown in Figure (4.1). Each value represents the best-obtained accuracy for various iterations.

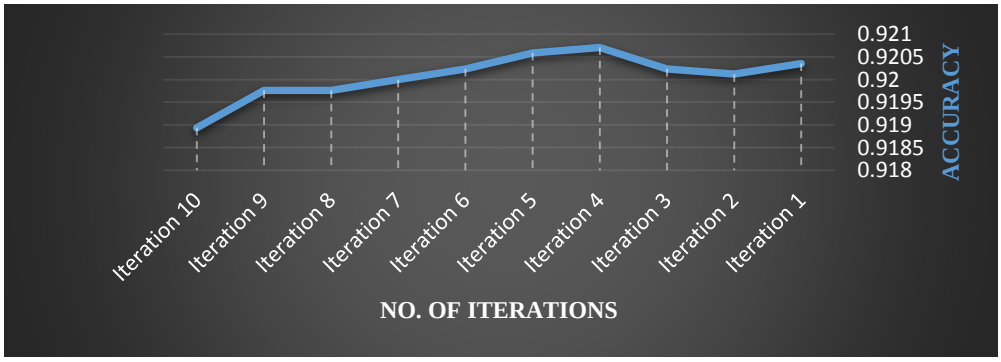

Figure (4.1): The detection accuracy with varying the number of iterations.

4.5 Results of Accuracy and Loss Using CNN and BM Model

This section describes the results of accuracy and loss using CNN and BM model, after describing the number of iterations, which are ten iterations, and the number of solutions, which are ten solutions in Section (4.4). This model consists of two layers, which are 128 and 64 nodes, as shown in **Table (4.22).**

Table (4.22): Results of Accuracy and Loss for two Layers Model (128-64) using CNN & BM.

No. of Nodes	No. of Layers	No. of Iteration	No. of Solution	Accuracy	Loss	Time (Second)
128-64	2	10	10	0.916932166	0.265327107	8

In **Table (4.22)**, the columns of this table are as follow: the first column represents the number of nodes in each layer, which are 128 and 64 nodes for the first, and the second layer respectively. The second column contains the number of layers, which are two layers. The third column represents number of iterations, which are ten iterations. The fourth column represents number of solutions, which are ten solutions. The fifth, sixth and seventh columns represents the loss, accuracy and time of the model. Where the loss is equal to **0.265327107**, while the accuracy is equal to **0.916932165** and the time is equal to **8** seconds. The aim of combining BM algorithm with CNN is to reduce features that will helps in reduction of time and complexity for execution process.

4.6 Comparing Results of Loss and Accuracy

This section presents comparing the results of loss and accuracy that are resulting from CNN and BM model with results of loss and accuracy resulting from CNN model without BM, as shown in **Table (4.23)**.

Table (4.23): Comparing Results of Loss and Accuracy of Two Layers Model (128-64).

	CNN	CNN & Blue Monkey (BM)
No. of Nodes	**128-64**	**128-64**
No. of Layers	**2**	**2**
Accuracy	**0.92**	**0.92**
Loss	**0.26462948**	**0.265327107**
time	**10 seconds**	**8 seconds**
Features	**1035**	**666**

In **Table (4.23**), there are three columns are as follow: the first column represents names of Inputs and outputs, the second column represents the results obtained from CNN algorithm and the third column represents the results obtained from combining CNN algorithm with BM algorithm.

While the rows of this table are as follow: the first row represents name of algorithms used in this work. The second row represents the number of nodes for each algorithm, where are used same nodes in both algorithms (128 and 64 nodes). The third row represents number of layers for each algorithm, which are two layers for both algorithms. The fourth row shows the accuracy ratio, where are the ratio obtained in the both algorithms are (**0.92**). The fifth row shows the loss ratio, where the result obtained from the CNN algorithm is (**0.26462948**) while the result obtained from the CNN and BM algorithms is (**0.265327107**). The sixth row shows the time spent in implementation of operations, where the average is (**10 seconds**) for the CNN algorithm and (**8 seconds**) for the CNN and BM algorithms. The seventh row shows number of features in both algorithms, where result obtained using CNN algorithm is (**1035 feature**), while the result obtained using CNN and BM algorithms is (**666 feature**).

Finally, it can be noted that the superiority of the CNN and BM model over the CNN model in terms of reducing the features of the model while the accuracy remaining the same.

The most important benefit of reducing features it can be useful in terms of reducing time. In addition, if the accuracy is improving, then the features reduction can be useful in the process of eliminating the contradiction between the features.

4.7 The Comparison of the Proposed Approach with Some Other Approaches

This section presents the comparison of the proposed approach with some other approaches for electricity theft detection in smart grids in terms of the applied method, number of customers, and accuracy, as shown in **Table (4.24)**.

Table (4.24): The comparison between the proposed approach with some other approaches.

#	Reference	Methods	Number of Customers	Accuracy (%)
1	[64]	Support Vector Machines (SVM)	36176	60
2	[65]	Convolutional Neural Network-Long Short Term Memory (CNN-LSTM)	17120	89
3	[66]	Combination functions (SVM, OPF, C4,5 tree)	NA	86.20
4	[67]	Regression	30	78
5	[68]	SVM-based fraud detection model (FDM) with the introduction of a fuzzy inference system (FIS-SVM-FIS)	36176	72
6	[69]	Fuzzy logic	NA	55
7	[70]	Fuzzy classification	NA	74.50
8	[71]	Neural networks (NN)	NA	83.5
9	[72]	Neuro-fuzzy	4159	68.2
10	[31]	Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) and Stacked Autoencoder.	12,180	96.9
11	Proposed work	Convolutional Neural Networks (CNN), Blue Monkey (BM)	42372	92

The results in **Table (4.24)** shows the superiority of the proposed (CNN and BM) method over other methods in terms of accuracy with many customers (42372 customers) compared with the obtained accuracy and number of customers that are used by other methods.

Chapter Five

Conclusions and Future Work

Chapter Five:

Conclusions and Future Work

5.1 Introduction

After designing the electricity signals classifier that has high performance and making experiments for testing the results, some conclusions and future works are mentioned in this chapter.

5.2 Conclusions

The most important conclusions of this thesis are:

1. Supervised learning techniques are better than other techniques because there are labeled data that makes training of models has high performance, where using labels in electricity consumption dataset made the building of models has high performance, as described in section (**2.3.1)**.

2. Pre-trained models have high power in addressing electricity consumption data because these models are trained using big datasets and powerful computers and when extracting the data of dataset using normal CNN the accuracy is too low comparing to addressing electricity consumption data using SM, as described in section (**4.2)** and table **(4.1)** shows results for different nodes.

3. The performance of building of models and new electricity signals classification is increased by reducing dataset as described in section **(3.2.2)**.

4. Increasing fully connected layers and nodes in each layer lead to increasing the delay, without much affecting the accuracy, in the models in term of training and classifying, as described in section **(4.3)**.

5. Using an optimization algorithm (the BM algorithm) leads to reducing the extracted features to speed up the performance of the designed system, as described in section **(4.6)** and table **(4.23)**.
6. The important results obtained can be explained as follows: The model consists of two layers, which are 128 and 64 nodes, while number of

iterations, which are ten iteration and number of solutions, which are ten solution, as shown in table **(4.22)**. The final loss resulting from this model is **0.265327107** and the final accuracy resulting from this model is **0.916932165** with an amount of time **8** seconds. In addition, the features were (**1035 feature**) before using BM algorithm, while become (**666 feature**) after using BM algorithm.

5.3 Future Works

There is a list of future works that can be applied in several directions, some of them are:

1. There is stillroom to improve the results accuracy through implementing our methods with wide convolutional neural network.
2. Combining Random forest with CNN may be achieving better results in terms of accuracy and efficiency.
3. Combining CNN with LSTM can, in theory, help achieving better results in terms of accuracy and efficiency.

References

References

[1] S. Tan, D. De, W.-Z. Song, J. Yang, and S. K. Das, "Survey of security advances in smart grid: A data driven approach," IEEE Commun. Surv. Tutorials, vol. 19, no. 1, pp. 397–422, 2017.

[2] D. Alahakoon and X. Yu, "Smart electricity meter data intelligence for future energy systems: A survey," IEEE Transactions on Industrial Informatics, vol. 12, no. 1, pp. 425–436, 2016.

[3] M. Ehsani, Y. Gao, S. Longo, and K. Ebrahimi, Modern electric, hybrid electric, and fuel cell vehicles. CRC Press, 2018.

[4] M. H. Rehmani, M. Reisslein, A. Rachedi, M. Erol-Kantarci, and M. Radenkovic, "Integrating renewable energy resources into the smart grid: Recent developments in information and communication technologies," IEEE Transactions on Industrial Informatics, vol. PP, no. 99, pp. 1–1, 2018.

[5] J. Wu, S. Rangan, and H. Zhang, Green communications: theoretical fundamentals, algorithms, and applications. CRC Press, 2016.

[6] K. Hamedani, L. Liu, R. Atat, J. Wu, and Y. Yi, "Reservoir computing meets smart grids: attack detection using delayed feedback networks," IEEE Transactions on Industrial Informatics, vol. 14, no. 2, pp. 734– 743, 2018.

[7] "Directive 2012/27/eu of the european parliament and of the council of 25 October 2012 on energy efficiency, amending directives 2009/125/ec and 2010/30/eu and repealing directives 2004/8/ec and 2006/32/ec text with eea relevance." [Online]. Available: https: //eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32012L0027

[8] J. Wu, S. Guo, H. Huang, W. Liu, and Y. Xiang, "Information and communications technologies for sustainable development goals: State-of-the-art, needs and perspectives," IEEE Communications Surveys & Tutorials, 2018.

[9] R. Jiang, R. Lu, Y. Wang, J. Luo, C. Shen, and X. S. Shen, "Energy-theft detection issues for advanced metering infrastructure in smart grid," Tsinghua Science and Technology, vol. 19, no. 2, pp. 105–120, 2014.

[10] H. Jiang, K. Wang, Y. Wang, M. Gao, and Y. Zhang, "Energy big data: A survey," IEEE Access, vol. 4, pp. 3844–3861, 2016.

[11] X. Yu and Y. Xue, "Smart grids: A cyber-physical systems perspective," Proceedings of the IEEE, vol. 104, no. 5, pp. 1058–1070, 2016.

[12] B. Saikiran and R. Hariharan, "Review of methods of power theft in Power System," Int. J. Sci. Eng. Res., vol. 5, no. 11, pp. 276–280, 2014.

[13] European Technology Platform, "SMART GRIDS"— Strategic Deployment Document for Europe's Electricity Networks of the Future, September 2008. http://www.smartgrid.eu/documents/smartgrids_SDD_Final_April2010.pdf

[14] T. Vijayapriya and D. P. Kothari, "Smart grid: an overview," Smart Grid Renew. Energy, vol. 2, no. 4, pp. 305–311, 2011.

[15] S. Rahman, "The Smart Grid and Its Impact on the Integration of Distributed Energy Resources," Southeast University, Nanjing, 2 April 2009.

[16] Z. Zheng, Y. Yang, X. Niu, H.-N. Dai, and Y. Zhou, "Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids," IEEE Trans. Ind. Informatics, vol. 14, no. 4, pp. 1606–1615, 2017.

[17] M. B. Line, I. A. Tøndel, and M. G. Jaatun, "Cybersecurity challenges in Smart Grids," in 2011 2nd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies, 2011, pp. 1–8.

[18] Y. Deng and S. Shukla, "Vulnerabilities and Countermeasures–A Survey on the Cyber Security Issues in the Transmission Subsystem of a Smart Grid," J. Cyber Secur. Mobil., vol. 1, no. 2, pp. 251–276, 2012.

[19] W. Wang and Z. Lu, "Cybersecurity in the smart grid: Survey and challenges," Comput. networks, vol. 57, no. 5, pp. 1344–1371, 2013.

[20] Z. A. Baig and A.-R. Amoudi, "An analysis of smart grid attacks and countermeasures," J. Commun., vol. 8, no. 8, pp. 473–479, 2013.

[21] N. Komninos, E. Philippou, and A. Pitsillides, "Survey in smart grid and smart home security: Issues, challenges and countermeasures," IEEE Commun. Surv. Tutorials, vol. 16, no. 4, pp. 1933–1954, 2014.

[22] R. R. Mohassel, A. Fung, F. Mohammadi, and K. Raahemifar, "A survey on advanced metering infrastructure," Int. J. Electr. Power Energy Syst., vol. 63, pp. 473–484, 2014.

[23] J. Nagi, K. S. Yap, S. K. Tiong, S. K. Ahmed, and M. Mohamad, "Nontechnical loss detection for metered customers in power utility using support vector machines," IEEE Trans. Power Deliv., vol. 25, no. 2, pp. 1162–1171, 2009.

[24] E. W. S. Angelos, O. R. Saavedra, O. A. C. Cortés, and A. N. de Souza, "Detection and identification of abnormalities in customer consumptions in power distribution systems," IEEE Trans. Power Deliv., vol. 26, no. 4, pp. 2436–2442, 2011.

[25] S. S. S. R. Depuru, L. Wang, V. Devabhaktuni, and P. Nelapati, "A hybrid neural network model and encoding technique for enhanced classification of energy consumption data," in 2011 IEEE Power and Energy Society General Meeting, 2011, pp. 1–8.

[26] O. Abdel-Hamid, L. Deng, and D. Yu, "Exploring convolutional neural network structures and optimization techniques for speech recognition.," in Inter speech, 2013, vol. 11, pp. 73–75.

[27] S. Mallat, "Understanding deep convolutional networks," Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., vol. 374, no. 2065, p. 20150203, 2016.

[28] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in 2017 International Conference on Engineering and Technology (ICET), 2017, pp. 1–6.

[29] K. Xu, P. Roussel, T. G. Csapó, and B. Denby, "Convolutional neural network-based automatic classification of midsagittal tongue gestural targets using B-mode ultrasound images," J. Acoust. Soc. Am., vol. 141, no. 6, pp. EL531–EL537, 2017.

[30] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Image net classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097–1105.

[31] R. R. Bhat, R. D. Trevizan, R. Sengupta, X. Li, and A. Bretas, "Identifying nontechnical power loss via spatial and temporal deep learning," in 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), 2016, pp. 272–279.

[32] M. Mahmood and B. Al-Khateeb, "The blue monkey: A new nature inspired metaheuristic optimization algorithm," Period. Eng. Nat. Sci., vol. 7, no. 3, pp. 1054–1066, 2019.

[33] S. McLaughlin, B. Holbert, A. Fawaz, R. Berthier, and S. Zonouz, "A multi-sensor energy theft detection framework for advanced metering infrastructures," IEEE Journal on Selected Areas in Communications, vol. 31, no. 7, pp. 1319–1330, 2013.

[34] J. P. Navani, N. K. Sharma, and S. Sapra, "Technical and nontechnical losses in power system and its economic consequence in Indian economy," International Journal of Electronics and Computer Science Engineering, vol. 1, no. 2, pp. 757–761, 2012.

[35] S. McLaughlin, B. Holbert, A. Fawaz, R. Berthier, and S. Zonouz, "A multi-sensor energy theft detection framework for advanced metering infrastructures," IEEE Journal on Selected Areas in Communications, vol. 31, no. 7, pp. 1319–1330, 2013.

[36] P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," IEEE Security & Privacy Magazine, vol. 7, no. 3, pp. 75–77, 2009.

[37] T. B. Smith, “Electricity theft: a comparative analysis,” Energy Policy, vol. 32, no. 1, pp. 2067–2076, 2004.

[38] J. I. Guerrero, C. Le´on, I. Monedero, F. Biscarri, and J. Biscarri, “Improving knowledge-based systems with statistical techniques, text mining, and neural networks for nontechnical loss detection,” Knowledge-Based Systems, vol. 71, no. 4, pp. 376–388, 2014.

[39] C. C. O. Ramos, A. N. Souza, G. Chiachia, A. X. Falcم‌o, and J. P. Papa, “A novel algorithm for feature selection using harmony search and its application for non-technical losses detection,” Computers & Electrical Engineering, vol. 37, no. 6, pp. 886–894, 2011.

[40] P. Glauner, J. A. Meira, P. Valtchev, R. State, and F. Bettinger, “-e challenge of non-technical loss detection using artificial intelligence: a surveyficial intelligence: a survey,” International Journal of Computational Intelligence Systems, vol. 10, no. 1, pp. 760–775, 2017.

[41] G. M. Messinis and N. D. Hatziargyriou, “Review of non-technical loss detection methods,” Electr. Power Syst. Res., vol. 158, pp. 250–266, 2018.

[42] P. Jokar, N. Arianpoo, V.C.M. Leung, Electricity theft detection in AMI using customers’ consumption patterns, IEEE Trans. Smart Grid 7 (2016) 216–226,http://dx.doi.org/10.1109/TSG.2015.2425222.

[43] Y. Guo, C.W. Ten, P. Jirutitijaroen, Online data validation for distribution operations against cyber tampering, IEEE Trans. Power Syst. 29 (2014)550–560, http://dx.doi.org/10.1109/TPWRS.2013.2282931.

[44] S.-C. Huang, Y.-L. Lo, C.-N. Lu, Non-technical loss detection using state estimation and analysis of variance, IEEE Trans. Power Syst. 28 (2013)2959–2966, http://dx.doi.org/10.1109/TPWRS.2012.2224891.

[45] A. Rossoni, R. Trevizan, A. Bretas, D. Gazzana, A. Bettiol, A. Carniato, L. Passos, R. Martin, Hybrid formulation for technical and non-technical losses estimation and identification in distribution networks: application in a Brazilian power system, CIRED 23 Rd Int. Conf. Electr. Distrib. (2015), pp.15–18.

[46] R. Lu, X. Liang, X. Li, X. Lin, and X. Shen, EPPA: An efficient and privacy-preserving aggregation scheme for secure smart grid communications, IEEE Transactions on Parallel and Distributed Systems, vol. 23, no. 9, pp. 1621-1631, 2012.

[47] M. Wen, R. Lu, J. Lei, H. Li, X. Liang, and X. Shen, SESA: An efficient searchable encryption scheme for auction in emerging smart grid marketing, Security and Communication Networks, vol. 7, no. 1, pp. 234-244, 2014.

[48] H. Li, X. Liang, R. Lu, X. Lin, H. Yang, and X. Shen, EPPDR: An efficient privacy-preserving demand response scheme with adaptive key evolution in

smart grid, IEEE Transactions on Parallel and Distributed Systems, vol. PP, no. 99, pp. 1-10, 2013.

[49] X. Liang, X. Li, R. Lu, X. Lin, and X. Shen, UDP: Usage based dynamic pricing with privacy preservation for smart grid, IEEE Transactions on Smart Grid, vol. 4, no. 1, pp. 141-150, 2013.

[50] Independent Electricity System Operator (IESO) Office of the Information and Privacy Commissioner Ontario, Canada, Building privacy into Ontario's smart meter data management system: A control framework, Tech. Rep., http://www.ipc.on.ca/english/Resources/Discussion-Papers/ Discussion-Papers-Summary/?id=1183, May 07, 2012.

[51] S. McLaughlin, B. Holbert, A. Fawaz, R. Berthier, and S. Zonouz, A multi-sensor energy theft detection framework for advanced metering infrastructures, IEEE Journal on Selected Areas in Communications, vol. 31, no. 7, pp. 1319-1330, 2013.

[52] C.-H. Lo and N. Ansari, CONSUMER: A novel hybrid intrusion detection system for distribution networks in smart grid, IEEE Transactions on Emerging Topics in Computing, vol. 1, no. 1, pp. 33-44, 2013.

[53] B. Khoo and Y. Cheng, Using RFID for anti-theft in a Chinese electrical supply company: A cost-benefit analysis, in Proc. IEEE Wireless Telecommunications Symposium (WTS), 2011, pp. 1-6.

[54] S. Depuru, L. Wang, and V. Devabhaktuni, Electricity theft: Overview, issues, prevention and a smart meter based approach to control theft, Energy Policy, vol. 39, no. 2, pp. 1007-1015, 2011.

[55] S. Depuru, L. Wang, V. Devabhaktuni, and N. Gudi, Measures and setbacks for controlling electricity theft, in Proc. IEEE North American Power Symposium (NAPS), 2010, pp. 1-8.

[56] S. Amin, G. A. Schwartz, and H. Tembine, Incentives and security in electricity distribution networks, in Decision and Game Theory for Security, Springer, 2012, pp. 264- 280.

[57] A. A. C´ardenas, S. Amin, G. Schwartz, R. Dong, and S. Sastry, A game theory model for electricity theft detection and privacy-aware control in AMI systems, in Proc. IEEE 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton), 2012, pp. 1830-1837.

[58] A. M. Giancarlo Zaccone, Md. Rezaul Karim, Deep Learning with TensorFlow: Explore neural networks with Python. 2017.

[59] P. Doshi, S. Punktambekar, N. Kini, and S. S. Dhami, “Theft Detection System using Convolutional Neural Network and Object Tracking.” Vol-5 Issue-3 2019, IJARIIE-ISSN(O)-2395-4396.

[60] "A Study on CNN Transfer Learning for Image Classification," *Adv. Intell. Syst. Comput.*, pp. 191–202, 2018.

[61] N. Ketkar and E. Santana, Deep learning with python, vol. 1. Springer, 2017.

[62] "Convolutional Neural Network - MATLAB & Simulink." [Online]. Available: https://www.mathworks.com/solutions/deep-learning/convolutional-neural-network.html. [Accessed: 06-Aug-2019].

[63] S. Mirjalili, A. H. Gandomi, S. Z. Mirjalili, S. Saremi, H. Faris, and S. M. Mirjalili, "Salp Swarm Algorithm: A Bio-Inspired Optimizer for Engineering Design Problems," Adv. Eng. Softw., Vol. 114, Pp. 163–191, 2017.

[64] Nagi, J.; Mohammad, A.; Yap, K.S.; Tiong, S.K.; Ahmed, S.K. Non-technical loss analysis for detection of electricity theft using support vector machines. In Proceedings of the 2008 IEEE 2nd International Power and Energy Conference, Johor Bahru, Malaysia, 1–3 December 2008; IEEE: Piscataway, NJ, USA, 2008; pp. 907–912.

[65] M. Hasan, R. N. Toma, A.-A. Nahid, M. M. Islam, and J.-M. Kim, "Electricity theft detection in smart grid systems: A CNN-LSTM based approach," Energies, vol. 12, no. 17, p. 3310, 2019.

[66] Di Martino, M.; Decia, F.; Molinelli, J.; Fernández, A. Improving Electric Fraud Detection using Class Imbalance Strategies. In Proceedings of the International Conference on Pattern Recognition Applications and Methods (ICPRAM), Vilamoura, Portugal, 6–8 February 2012; pp. 135–141.

[67] Jindal, A.; Dua, A.; Kaur, K.; Singh, M.; Kumar, N.; Mishra, S. Decision tree and SVM-based data analytics for theft detection in smart grid. IEEE Trans. Ind. Inform. 2016, 12, 1005–1016.

[68] Nagi, J.; Yap, K.S.; Tiong, S.K.; Ahmed, S.K.; Nagi, F. Improving SVM-based nontechnical loss detection in power utility using the fuzzy inference system. IEEE Trans. Power Deliv. 2011, 26, 1284–1285.

[69] Nagi, J.; Yap, K.S.; Nagi, F.; Tiong, S.K.; Koh, S.; Ahmed, S.K. NTL detection of electricity theft and abnormalities for large power consumers in TNB Malaysia. In Proceedings of the 2010 IEEE Student Conference on Research and Development (SCOReD), Putrajaya, Malaysia, 13–14 December 2010; IEEE: Piscataway, NJ, USA, 2010; pp. 202–206.

[70] Angelos, E.W.S.; Saavedra, O.R.; Cortés, O.A.C.; de Souza, A.N. Detection and identification of abnormalities in customer consumptions in power distribution systems. IEEE Trans. Power Deliv. 2011, 26, 2436–2442.

[71] Muniz, C.; Figueiredo, K.; Vellasco, M.; Chavez, G.; Pacheco, M. Irregularity detection on low tension electric installations by neural network ensembles. In Proceedings of the 2009 International Joint Conference on Neural Networks, Atlanta, GA, USA, 14–19 June 2009; IEEE: Piscataway, NJ, USA, 2009; pp. 2176–2182.

[72] Muniz, C.; Vellasco, M.M.B.R.; Tanscheit, R.; Figueiredo, K. A Neuro-fuzzy System for Fraud Detection in Electricity Distribution. In Proceedings of the IFSA/EUSFLAT Conference, Lisbon, Portugal, 20–24 July 2009; pp. 1096–1101.

Appendix A

Various data-oriented, network-oriented, and hybrid ETD techniques

Appendix A:

Various data-oriented, network-oriented, and hybrid ETD techniques

Table (A.1): Summary of data-oriented ETD techniques [41].

Ref.	Group	Conception	Procedure	Kinds of Data	Size of Dataset	Features	Metrics	Time of Response
[1]	Oriented of Data	Supervised	Universal Preservative style	Region Technical Non-technical Customer, Region Non-technical,	Big	-	Rate of Hit	365 days
[2]	Oriented of Data	Supervised	SVM	Little Resolution Power, Non-technical Customer	Big	Mean	Rate of Hit, Accurateness	365 days
[3]	Oriented of Data	Supervised	Induction of Rule, SVM	Little Resolution Power, Non-technical Customer	Big	Mean, min./max.	Rate of Hit, Accurateness	365 days
[4]	Oriented of Data	Supervised	SVM, ANN	Little Resolution Power	Moderate	Mean	Rate of Hit, Accurateness	365 days
[5]	Oriented of Data	Supervised	Tree of Decision, SVM, OPF	Little Resolution Power	Big	Mean, Wavelet, coefficients of predicted kWh, coefficients of Fourier, Euclidean distance to Average consumer, coefficients of Polynomial fit, linear fit Slope consuming curve difference,	Rate of Hit, Precision, F _ score, TNR	365 days
[6]	Oriented of Data	Supervised	Tree of Decision, SVM	Little Resolution Power	Big	Readings approved from DSO to other readings, extreme permissible consuming,	The score of F1, Precision, Rate of Hit, Accurateness	365 days

Appendix A: Various data-oriented, network-oriented, and hybrid ETD techniques

						irregularities number, days since update/inspection, postponement of instalment		
[7]	Oriented of Data	Supervised	OPF	Technical Customer, moderate Resolution Power	Big	Min/max, the factor of Load	Time of Arrangement, Accurateness, Time of Training,	30 days
[8]	Oriented of Data	Supervised	OPF	Technical Customer, moderate Resolution Power	Big	Min/max, the factor of energy, the factor of Load	Rate of Recognition, Correctness	30 days
[9]	Oriented of Data	Supervised	OPF, SOM, ,SVM, k-NN ,ANN	Technical Customer , moderate Resolution Power	Big	Min/max, factor of energy , factor of Load	Rate of Recognition ,Correctness	30 days
[10]	Oriented of Data	Supervised	Trees of Decision, Induction of Rule	Technical Customer, moderate Resolution Power, Customer Non-Technical	Big	Min/max, changeability, the factor of energy, daily consuming to contracted energy streaks	Accurateness, Support, exactness	365 days
[11]	Oriented of Data	Supervised	Classifiers of Bayesian, Trees of Decision, Induction of Rule	Technical Customer, Little Resolution Power, Customer Non-Technical	Big	Min/max, changeability, the factor of energy, daily consuming to contracted energy streaks, coefficient of Pearson, Coefficient billed-expended power	Accurateness, Support, exactness, rate of hit	30 days
[12]	Oriented of Data	Supervised	ANN	Technical Customer, Little Resolution Power, Customer Non-Technical	Big	-	FPR, Accurateness, Support, rate of hit	365 days

Appendix A: Various data-oriented, network-oriented, and hybrid ETD techniques

[13]	Oriented of Data	Supervised	KNN, OPF, ANN, SOM, SVM	Technical Customer, moderate Resolution Power	Big	Min/max, parts of PCA factor of load, the factor of power	Time of Training, Accurateness,	30 days
[14]	Oriented of Data	Supervised	SVM	Technical Customer, moderate Resolution Power	Big	Min/max, the factor of load, the factor of power	Accurateness	30 days
[15]	Oriented of Data	Supervised	SVM, the grouping of KNN,	moderate Resolution Power	Big	Max/min, parts of PCA	Accurateness, exactness, rate of hit	30 days
[16]	Oriented of Data	Supervised	ANN	Technical Customer, moderate Resolution Power	Big	Min/max, the factor of load, the factor of power	Time of Training, Accurateness	30 days
[17]	Oriented of Data	Supervised	Classifiers of Bayesian, KNN, SVM, Tree of Decision, ANN	Technical Customer, Little Resolution Power, Customer Non-Technical Ecological, Region of Non-Technical	Big	Max/min, coefficient of billed-consuming Power, coefficient of Pearson, readings number reduction of Consuming associated with the preceding interval,	The score of F1, AUC	30 days
[18]	Oriented of Data	Supervised	OPF	Moderate Resolution Power	Big	coefficients of Transform of Discrete Cosine	Rate of the hit, Exactness, the score of F1	365 days
[19]	Oriented of Data	Supervised	Induction of Rule, SVM	Little Resolution Power	Big	Encoding	Accurateness	30 days
[20]	Oriented of Data	Supervised	ANN	Moderate Resolution Power	Big	-	Accurateness, Time of Classification , Time of Training	30 days

Appendix A: Various data-oriented, network-oriented, and hybrid ETD techniques

[21]	Oriented of Data	Supervised	ANN	moderate Resolution Power	Big	-	FNR, FPR, TPR, Rate of Hit	Hours
[22]	Oriented of Data	Supervised	Tree of Decision ,ANN	moderate Resolution Power, Ecological	-	-	-	30 days
[23]	Oriented of Data	Supervised	Induction of Rule	Technical Customer, Little Resolution Power, Customer Non-Technical	Big	-	-	365 days
[24]	Oriented of Data	Unsupervised & supervised	ANN, Induction of Rule, Systems of Expert,	Technical Customer, Little Resolution Power, Customer Non-Technical	Big	Mean, min/max, standard deviance	-	Months
[25]	Oriented of Data	Unsupervised & supervised	Systems of Expert, SVM	Little Resolution Power, Customer Non-Technical	Big	The slope of Consuming curve, Mean	Accurateness, Rate of Hit, TNR. FNR, FPR, AUC	365 days
[26]	Oriented of Data	Unsupervised & supervised	Classifiers of Bayesian, Clustering	Data of Network of Smart Meter, Great Resolution power	Minor	-	Accurateness, Time of Classification , Time of Training, FNR, FPR	Days
[27]	Oriented of Data	Unsupervised	Clustering	Little Resolution Power, Customer Non-Technical, Mean Region Consuming	Big	Mean, min/max, standard deviance	Exactness, Rate of Hit	Months
[28]	Oriented of Data	Unsupervised	Models of Regression	moderate Resolution Power	Moderate	Mean, predicted kWh , standard deviance	Electrical energy robbed	Hours

Appendix A: Various data-oriented, network-oriented, and hybrid ETD techniques

[29]	Oriented of Data	Unsupervised	Clustering	Little Resolution Power, Mean Region Consuming, Customer Non-Technical	Minor	Mean, min/max, standard deviance	-	Months
[30]	Oriented of Data	Unsupervised	Distribution of Multivariate Gaussian, clustering, OPF	Technical Customer, moderate Resolution Power	Big	Min/max, the factor of load, the factor of power	Accurateness, Rate of Acknowledgment, the score of F1	Months
[31]	Oriented of Data	Unsupervised	Clustering	moderate Resolution Power	Big	parts of PCA	FNR, FPR	Weeks
[32]	Oriented of Data	Unsupervised	System of Expert	moderate Resolution Power	Minor	Arrangement of Fractional of dynamic errors	-	Hours
[33]	Oriented of Data	Unsupervised	(NCG) Non-cooperative game with a system of FOSE	moderate Resolution Power	Minor	Arrangement of Fractional of dynamic errors	-	Hours
[34]	Oriented of Data	Unsupervised	(NCG) Non-cooperative game with a system of FOSE	Moderate Resolution Power	Minor	Arrangement of Fractional of dynamic errors	-	Hours
[35]	Oriented of Data	Unsupervised	Numerical Controller	Little Resolution Power	Moderate	-	FNR, Rate of Hit	Months
[36]	Oriented of Data	Unsupervised	Numerical Controller	Moderate Resolution Power	Moderate	-	Power Balance Mismatch, Rate of Hit, Mean bill rise, standardized labour fee	Days
[37]	Oriented of Data	Unsupervised	SOM	Moderate Resolution Power	Moderate	-	FPR, FNR, TPR, Rate of Hit	Weeks
[38]	Oriented of Data	Unsupervised	The difference of Kullback–Leibler	Moderate Resolution Power	Moderate	-	Electrical energy robbed, Rate of Hit	Weeks

[39]	Oriented of Data	Unsupervised	Models of Regression, Numerical Controller, Local external Factor	moderate Resolution Power	Moderate	-	Undetected attack fee, FPR	Hours

Table (A.2): Summary of network-oriented and hybrid techniques [41].

Ref.	Group	Conception	Procedure	Kinds of Data	Size of Dataset	Features	Metrics	Time of Response
[40]	Oriented of Network	Flow of Load	The flow of the power of the grid of Distribution	Moderate Resolution Power, Data of Observer Meter, data of network of Smart meter, Topology of Network	Moderate	-	-	Hours
[41]	Oriented of Network	Flow of Load	Factor identification for calculation of loss of practical	Moderate Resolution Power, Smart Meter Network, Data of Observer Meter	Minor	-	Rate of Hit	Hours
[42]	Oriented of Network	Flow of Load	Decomposition of individual Value & Stochastic Petri Nets	Moderate Resolution Power, data of FRTU, data of network of Smart meter, mean region consuming, Data of Observer Meter	Minor	-	Smallest detected deviance from typical	Hours
[43]	Oriented of Network	Flow of Load	The distributed answer of linear systems	Moderate Resolution Power, Data of Observer Meter, Topology of Network	Moderate	-	-	Hours
[44]	Oriented of Network	Flow of Load	The probabilistic flow of energy for distribution grid	Little Resolution Power, Data of Observer Meter, Topology of	Big	-	-	Days

Appendix A: Various data-oriented, network-oriented, and hybrid ETD techniques

				Network, Smart Meter Network				
[45]	Oriented of Network	Flow of Load	identification of Voltage sensitivities by small linear squares	Great Resolution Power, Data of Observer Meter, Topology of Network, data of network of Smart meter	Minor	-	-	Hours
[46]	Oriented of Network	Flow of Load	Recursive small Squares for modelling of the manner of meter	Moderate Resolution Power, Data of Observer Meter, Topology of Network	Minor	-	Time of Arrangement, Rate of Hit	Hours
[47]	Oriented of Network	Network of sensor	Cross-Entropy & Conditional Arbitrary domain	Great Resolution Power, Topology of Network	Minor	-	The fee of FRTU, index of coverage of irregularity	Days
[48]	Oriented of Network	Network of sensor	Dynamic Programming	The topology of Network, Smart Meter Network	Moderate	-	The fee of FRTU, index of coverage of irregularity	Days
[49]	Oriented of Network	Network of sensor	Integer Linear Programming	Topology of Network, data of FRTU, Mean region consuming	-	-	fee of FRTU	Days
[50]	Oriented of Network	Network of sensor	algorithms of Tree search	Moderate Resolution Power, Topology of Network, Data of Observer Meter	Moderate	-	Time of Classification , Rate of Hit	Hours
[51]	Oriented of Network	Estimation of State	The distributed answer of Kalman filter	Data of Observer Meter, Topology of Network, data of network of Smart meter	Moderate	-	Accurateness	Minutes

Appendix A: Various data-oriented, network-oriented, and hybrid ETD techniques

[52]	Oriented of Network	Estimation of State	The estimator of state of MV (LV) WLS	Moderate Resolution Power, Topology of Network, Data of Observer Meter, data of FRTU	Minor	-	-	Minutes
[53]	Oriented of Network	Estimation of State	Assessment of the state of LV and MV	Great Resolution Power, Topology of Network, Data of Observer Meter, data of FRTU	Minor	-	-	Minutes
[54]	Oriented of Network	Estimation of State	Assessment of the state of MV WLS and detection of insufficient data	Moderate Resolution Power, Topology of Network, data of FRTU	Minor	-	-	Minutes
[55]	Oriented of Network	Estimation of State	detection of insufficient data and gathering of network	The topology of Network, data of FRTU	-	-	-	Hours
[56]	Oriented of Network	Sensor Network & Estimation of State	The estimator of state of DC, grid located sensor algorithm	The topology of Network, Smart Meter Network data	Moderate	-	Rate of Hit	Hours
[57]	Hybrid	Unsupervised & Estimation of State	The estimator of state of MV and ANOVA	Moderate Resolution Power, data of FRTU, Topology of Network, mean region consuming	Minor	-	-	Hours
[58]	Hybrid	Unsupervised & Estimation of State	ANOVA and assessment of the state of WLS (semidefinite programming)	Great Resolution Power, Data of Observer Meter, Topology of Network, data of network of Smart meter, data of FRTU	Minor	-	-	Minutes

Appendix A: Various data-oriented, network-oriented, and hybrid ETD techniques

[59]	Hybrid	The flow of Load & Supervised	spectator meter with SVM	Moderate Resolution Power, Data of Observer Meter, Topology of Network	Big	-	Rate of Hit, DR-FPR, FPR, Rate of Detection of Bayesian	Days
[60]	Hybrid	The flow of Load & Supervised	Decision Tree with spectator meter, SVM	Moderate Resolution Power, Customer Technical, Topology of Network, Customer Non-Technical, Data of Observer Meter, ecological	Moderate	Predicted kWh	Accurateness, FPR, Rate of Hit, Time of Arrangement	Months
[61]	Hybrid	supervised & Estimation of State	MV and OPF estimation of state	The topology of Network, Little Resolution Power, data of FRTU	Big	-	Exactness, Rate of Hit	Months
[62]	Hybrid	The flow of Load & Supervised	theory of Rough group with the assessment of practical loss	Moderate Resolution Power, Customer Technical, Topology of Network, Customer Non-Technical, Data of Observer Meter	Big	Mean, rates of consuming of season	-	Months
[63]	Hybrid	The flow of Load & un Supervised	Numerical controller with spectator meter	Moderate Resolution Power, Topology of Network	Big	-	-	Months
[64]	Hybrid	Unsupervised & Estimation of State	MV WLS state estimator and Multivariate Gaussian Distribution	Little Resolution Power, data of FRTU, Topology of Network	Big	-	Accurateness, Rate of Hit	Months

Appendix A: Various data-oriented, network-oriented, and hybrid ETD techniques

[65]	Hybrid	The flow of Load & un Supervised	A-star (A*) derivative algorithm and Numerical controller	The topology of Network, data of network of Smart meter, data of FRTU, mean region consuming	Moderate	-	Rate of Hit	Minutes
[66]	Hybrid	Supervised, unsupervised & Sensor Network	SVM with clustering of network, Clustering	Medium Resolution Energy, average area consumption, Observer Meter Data, Network Topology, FRTU data, Moderate Resolution Power, Data of Observer Meter, Topology of Network, data of FRTU, mean region consuming	Moderate	Mean, the ratio of mismatch	FNR, FPR	Days

Appendix A References

[1] L. Faria, J. Melo, A. Padilha-Feltrin, Spatial-temporal estimation fornontechnical losses, IEEE Trans. Power Deliv. 8977 (2015), http://dx.doi.org/10.1109/TPWRD.2015.2469135, 1-1.

[2] J. Nagi, K.S. Yap, S.K. Tiong, S.K. Ahmed, M. Mohamad, Nontechnical lossdetection for metered customers in power utility using support vectormachines, IEEE Trans. Power Deliv. 25 (2010) 1162–1171, http://dx.doi.org/10.1109/TPWRD.2009.2030890.

[3] J. Nagi, K.S. Yap, S.K. Tiong, S.K. Ahmed, F. Nagi, Improving SVM-basednontechnical loss detection in power utility using the fuzzy inference system,IEEE Trans. Power Deliv. 26 (2011) 1284–1285, http://dx.doi.org/10.1109/TPWRD.2010.2055670.

[4] K.S. Yap, S.K. Tiong, J. Nagi, J.S.P. Koh, F. Nagi, Comparison of supervisedlearning techniques for non-technical loss detection in power utility, Int. Rev.Comput. Softw. 7 (2012) 626–636.

[5] M. Di Martino, F. Decia, J. Molinelli, A. Fernández, A novel framework fornontechnical losses detection in electricity companies, in: P. Latorre Carmona,J.S. Sánchez, A.L.N. Fred (Eds.), Pattern Recognit. — Appl. Methods, Springer,Berlin Heidelberg, Berlin, Heidelberg, 2013, pp. 109–120, http://dx.doi.org/10.1007/978-3-642-36530-0 9.

[6] J.P. Kosut, F. Santomauro, A. Jorysz, A. Fernandez, F. Lecumberry, F. Rodriguez,Abnormal consumption analysis for fraud detection: UTE-UDELAR jointefforts, in: 2015 IEEE PES Innov. Smart Grid Technol. Lat. Am. (ISGT LATAM),IEEE, 2015, http://dx.doi.org/10.1109/ISGT-LA.2015.7381272, pp. 887–892.

[7] C.C.O. Ramos, A.N. De Sousa, J.P. Papa, A.X. Falcão, A new approach fornontechnical losses detection based on optimum-path forest, IEEE Trans.Power Syst. 26 (2011) 181–189, http://dx.doi.org/10.1109/TPWRS.2010.2051823.

[8] C.C.O. Ramos, D. Rodrigues, A.N. de Souza, J.P. Papa, On the study ofcommercial losses in brazil: a binary black hole algorithm for theftcharacterization, IEEE Trans. Smart Grid 1 (2016), http://dx.doi.org/10.1109/TSG.2016.2560801.

[9] C.C.O. Ramos, A.N. De Souza, A.X. Falcão, J.P. Papa, New insights onnontechnical losses characterization through evolutionary-based featureselection, IEEE Trans. Power Deliv. 27 (2012) 140–146, http://dx.doi.org/10.1109/TPWRD.2011.2170182.

[10] C. León, F. Biscarri, I. Monedero, J.I. Guerrero, J. Biscarri, R. Millán, Variabilityand trend-based generalized rule induction model to NTL detection in powercompanies, IEEE Trans. Power Syst. 26 (2011) 1798–1807, http://dx.doi.org/10.1109/TPWRS.2011.2121350.

[11] I. Monedero, F. Biscarri, C. León, J.I. Guerrero, J. Biscarri, R. Millán, Detection offrauds and other non-technical losses in a power utility using Pearsoncoefficient, Bayesian networks and decision trees, Int. J. Electr. Power EnergySyst. 34 (2012) 90–98, http://dx.doi.org/10.1016/j.ijepes.2011.09.009.

[12] B.C. Costa, B.L.A. Alberto, A.M. Portela, W. Maduro, E.O. Eler, Fraud detectionin electric power distribution networks using an Ann-basedknowledge-discovery process, Int. J. Artif. Intell. Appl. 4 (2013) 17–23, http://dx.doi.org/10.5121/ijaia.2013.4602.

[13] C.C.O. Ramos, A.N. Souza, G. Chiachia, A.X. Falcão, J.P. Papa, A novel algorithmfor feature selection using Harmony Search and its application fornon-technical losses detection, Comput. Electr. Eng. 37 (2011) 886–894,http://dx.doi.org/10.1016/j.compeleceng.2011.09.013.

[14] D.R. Pereira, M.A. Pazoti, L.A.M. Pereira, D. Rodrigues, C.O. Ramos, A.N. Souza,J.P. Papa, Social-Spider Optimization-based Support Vector Machines appliedfor energy theft detection, Comput. Electr. Eng. 49 (2016) 25–38, http://dx.doi.org/10.1016/j.compeleceng.2015.11.001.

[15] J. No, S.Y. Han, Y. Joo, J. Shin, Conditional abnormality detection based on AMIdata mining, IET Gener. Transm. Distrib. 10 (2016) 3010–3016, http://dx.doi.org/10.1049/iet-gtd.2016.0048.

[16] L.A.M. Pereira, L.C.S. Afonso, J.P. Papa, Z.A. Vale, C.C.O. Ramos, D.S. Gastaldello,A.N. Souza, Multilayer perceptron neural networks training through chargedsystem search and its application for non-technical losses detection, in: 2013IEEE PES Conf. Innov. Smart Grid Technol. (ISGT Lat. Am.), IEEE, 2013, http://dx.doi.org/10.1109/ISGT-LA.2013.6554383, pp. 1–6.

[17] B. Coma-Puig, J. Carmona, R. Gavalda, S. Alcoverro, V. Martin, Fraud detectionin energy consumption: a supervised approach, in: 2016 IEEE Int. Conf. DataSci. Adv. Anal., IEEE, 2016, http://dx.doi.org/10.1109/DSAA.2016.19, pp.120–129.

[18] R.D. Trevizan, A.S. Bretas, A. Rossoni, Non-technical losses detection: a discrete cosine transform and optimum-path forest based approach, in: 2015 NorthAm. Power Symp., IEEE, 2015, http://dx.doi.org/10.1109/NAPS.2015.7335160,pp. 1–6.

[19] S.S.S.R. Depuru, L. Wang, V. Devabhaktuni, R.C. Green, High performance computing for detection of electricity theft, Int. J. Electr. Power Energy Syst.47 (2013) 21–30, http://dx.doi.org/10.1016/j.ijepes.2012.10.031.

[20] A.H. Nizar, Z.Y. Dong, Y. Wang, Power utility nontechnical loss analysis withextreme learning machine method, IEEE Trans. Power Syst. 23 (2008)946–955, http://dx.doi.org/10.1109/TPWRS.2008.926431.

[21] V. Ford, A. Siraj, W. Eberle, Smart grid energy fraud detection using artificial neural networks, in: 2014 IEEE Symp. Comput. Intell. Appl. Smart Grid, IEEE,2014, http://dx.doi.org/10.1109/CIASG.2014.7011557, pp. 1–6.

[22] D. Labate, P. Giubbini, G. Chicco, F. Piglione, Shape: the load prediction and non-technical losses modules, CIRED 23rd Int. Conf. Electr. Distrib. (2015), pp.15–18.

[23] C. León, F. Biscarri, I. Monedero, J.I. Guerrero, J. Biscarri, R. Millán, Integrated expert system applied to the analysis of non-technical losses in powerutilities, Expert Syst. Appl. 38 (2011) 10274–10285, http://dx.doi.org/10.1016/j.eswa.2011.02.062.

[24] J.I. Guerrero, C. León, I. Monedero, F. Biscarri, J. Biscarri, ImprovingKnowledge-Based Systems with statistical techniques, text mining, andneural networks for non-technical loss detection, Knowledge-based Syst. 71(2014) 376–388, http://dx.doi.org/10.1016/j.knosys.2014.08.014.

[25] P. Glauner, A. Boechat, L. Dolberg, R. State, F. Bettinger, Y. Rangoni, D. Duarte,Large-scale detection of non-technical losses in imbalanced data sets, in:2016 IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf., IEEE, 2016,http://dx.doi.org/10.1109/ISGT.2016.7781159, pp. 1–5.

[26] S. McLaughlin, B. Holbert, A. Fawaz, R. Berthier, S. Zonouz, A multi-sensorenergy theft detection framework for advanced metering infrastructures, IEEEJ. Sel. Areas Commun. 31 (2013) 1319–1330, http://dx.doi.org/10.1109/JSAC.2013.130714.

[27] E.W.S. Dos Angelos, O.R. Saavedra, O.A.C. Cortés, A.N. De Souza, Detection andidentification of abnormalities in customer consumptions in powerdistribution systems, IEEE Trans. Power Deliv. 26 (2011) 2436–2442, http://dx.doi.org/10.1109/TPWRD.2011.2161621.

[28] V.B. Krishna, R.K. Iyer, W.H. Sanders, ARIMA-based Modeling and Validationof Consumption Readings in Power Grids, Springer International Publishing,Cham, 2016, http://dx.doi.org/10.1007/978-3-319-33331-1.

[29] T.V. Babu, T.S. Murthy, B. Sivaiah, Detecting unusual customer consumptionprofiles in power distribution systems — APSPDCL, in: 2013 IEEE Int. Conf.Comput. Intell. Comput. Res., IEEE, 2013, http://dx.doi.org/10.1109/ICCIC.2013.6724264, pp. 1–5.

[30] L.A. Passos Júnior, C.C. Oba Ramos, D. Rodrigues, D.R. Pereira, A.N. de Souza,K.A. Pontara da Costa, J.P. Papa, Unsupervised non-technical lossesidentification through optimum-path forest, Electr. Power Syst. Res. 140(2016) 413–423, http://dx.doi.org/10.1016/j.epsr.2016.05.036.

[31] V. Badrinath Krishna, G.A. Weaver, W.H. Sanders, PCA-based method fordetecting integrity attacks on advanced metering infrastructure, Lect. NotesComput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. NotesBioinformatics) (2015), http://dx.doi.org/10.1007/978-3-319-22264-6 5, pp.70–85.

[32] S.-J. Chen, T. Zhan, C. Huang, J. Chen, C. LinNontechnical Loss and, Outagedetection using fractional-order self-synchronization error-based fuzzy petrinets in micro-distribution systems, IEEE Trans. Smart Grid 6 (2015) 411–420,http://dx.doi.org/10.1109/TSG.2014.2345780.

[33] C. Lin, S.-J. Chen, C. Kuo, J. Chen, Non-cooperative game model applied to anadvanced metering infrastructure for non-technical loss screening inmicro-distribution systems, IEEE Trans. Smart Grid 5 (2014) 2468–2469,http://dx.doi.org/10.1109/TSG.2014.2327809.

[34] T.-S. Zhan, C.-L. Kuo, S.-J. Chen, J.-L. Chen, C.-C. Kao, C.-H. Lin, Non-technicalloss and power blackout detection under advanced metering infrastructureusing a cooperative game based inference mechanism, IET Gener. Transm.Distrib. 10 (2016) 873–882, http://dx.doi.org/10.1049/iet-gtd.2015.0003.

[35] J.V. Spiri´c, M.B. Do˘ci´c, S.S. Stankovi´c, Fraud detection in registered electricitytime series, Int. J. Electr. Power Energy Syst. 71 (2015) 42–50, http://dx.doi.org/10.1016/j.ijepes.2015.02.037.

[36] Y. Liu, S. Hu, Cyberthreat analysis and detection for energy theft in socialnetworking of smart homes, IEEE Trans. Comput. Soc. Syst. 2 (2015) 148–158,http://dx.doi.org/10.1109/TCSS.2016.2519506.

[37] J.E. Cabral, J.O.P. Pinto, E.M. Martins, A.M.A.C. Pinto, Fraud detection in highvoltage electricity consumers using data mining, in: 2008 IEEE/PES Transm.Distrib. Conf. Expo., IEEE, 2008, http://dx.doi.org/10.1109/TDC.2008.4517232,pp. 1–5.

[38] V.B. Krishna, K. Lee, G.A. Weaver, R.K. Iyer, W.H. Sanders, F-DETA: aframework for detecting electricity theft attacks in smart grids, in: 2016 46thAnnu. IEEE/IFIP Int. Conf. Dependable Syst. Networks, IEEE, 2016, http://dx.doi.org/10.1109/DSN.2016.44, pp. 407–418.

[39] D. Mashima, A.A. Cárdenas, Evaluating Electricity Theft Detectors in SmartGrid Networks, 2012, http://dx.doi.org/10.1007/978-3-642-33338-5 11, pp.210–229.

[40] P. Kadurek, J. Blom, J.F.G. Cobben, W.L. Kling, Theft detection and smartmetering practices and expectations in the Netherlands, in: 2010 IEEE PESInnov. Smart Grid Technol. Conf. Eur. (ISGT Eur.), IEEE, 2010, http://dx.doi.org/10.1109/ISGTEUROPE.2010.5638852, pp. 1–6.

[41] D.N. Nikovski, Z. Wang, A. Esenther, H. Sun, K. Sugiura, T. Muso, K. Tsuru,Smart meter data analysis for power theft detection, Lect. Notes Comput. Sci.(including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 7988LNAI (2013) 379–389, http://dx.doi.org/10.1007/978-3-642-39712-7 29.

[42] M. Tariq, H.V. Poor, Electricity theft detection and localization in grid-tiedmicrogrids, IEEE Trans. Smart Grid 3053 (2016), http://dx.doi.org/10.1109/TSG.2016.2602660, 1-1.

[43] S. Salinas, M. Li, P. Li, Privacy-preserving energy theft detection in smart grids:a P2P computing approach, IEEE J. Sel. Areas Commun. 31 (2013) 257–267,http://dx.doi.org/10.1109/JSAC.2013.SUP.0513023.

[44] E.A.C. Aranha Neto, J. Coelho, Probabilistic methodology for technical andnon-technical losses estimation in distribution system, Electr. Power Syst.Res. 97 (2013) 93–99, http://dx.doi.org/10.1016/j.epsr.2012.12.008.

[45] S. Weckx, C. Gonzalez, J. Tant, T. De Rybel, J. Driesen, Parameter identificationof unknown radial grids for theft detection, in: 2012 3rd IEEE PES Innov.Smart Grid Technol. Eur. (ISGT Eur.), IEEE, 2012, http://dx.doi.org/10.1109/ISGTEurope.2012.6465644, pp. 1–6.

[46] W. Han, Y. Xiao, A novel detector to detect colluded non-technical loss fraudsin smart grid, Comput. Netw. 117 (2017) 19–31, http://dx.doi.org/10.1016/j.comnet.2016.10.011.

[47] C. Liao, C.W. Ten, S. Hu, Strategic FRTU deployment considering cybersecurityin secondary distribution network, IEEE Trans. Smart Grid 4 (2013)1264–1274, http://dx.doi.org/10.1109/TSG.2013.2256939.

[48] Y. zhou, X. Chen, A. Zomaya, L. Wang, S. Hu, A dynamic programmingalgorithm for leveraging probabilistic detection of energy theft in smarthome, IEEE Trans. Emerg. Top. Comput. (2015), http://dx.doi.org/10.1109/TETC.2015.2484841, 1-1.

[49] L.G. de O. Silva, A.A.P. da Silva, A.T. de Almeida-Filho, Allocation ofpower-quality monitors using the P-median to identify nontechnical losses,IEEE Trans. Power Deliv. 31 (2016) 2242–2249, http://dx.doi.org/10.1109/TPWRD.2016.2555282.

[50] Z. Xiao, Y. Xiao, D.H.-C. Du, Exploring malicious meter inspection inneighborhood area smart grids, IEEE Trans. Smart Grid 4 (2013) 214–226,http://dx.doi.org/10.1109/TSG.2012.2229397.

[51] S.A. Salinas, P. Li, Privacy-preserving energy theft detection in microgrids: astate estimation approach, IEEE Trans. Power Syst (2015) 1–12, http://dx.doi.org/10.1109/TPWRS.2015.2406311.

[52] Lijuan Chen, Xiaohui Xu, Chaoming Wang, Research on anti-electricitystealing method base on state estimation, in: 2011 IEEE Power Eng. Autom.Conf., IEEE, 2011, http://dx.doi.org/10.1109/PEAM.2011.6134972, pp.413–416.

[53] W. Luan, G. Wang, Y. Yu, J. Lin, W. Zhang, Q. Liu, Energy theft detection viaintegrated distribution state estimation based on AMI and SCADAmeasurements, in: 2015 5th Int. Conf. Electr. Util. Deregul. Restruct. PowerTechnol., IEEE, 2015, http://dx.doi.org/10.1109/DRPT.2015.7432350, pp.751–756.

[54] Yuan-Liang Lo, Shih-Che Huang, Chan-Nan Lu, Non-technical loss detectionusing smart distribution network measurement data, in: IEEE PES Innov.Smart Grid Technol., IEEE, 2012, http://dx.doi.org/10.1109/ISGT-Asia.2012.6303316, pp. 1–5.

[55] Y. Liu, Y. Wang, X. Guan, A novel method to detect bad data injection attack insmart grid, in: 2013 Proc. IEEE INFOCOM, IEEE, 2013, http://dx.doi.org/10.1109/INFCOM.2013.6567175, pp. 3423–3428.

[56] C.-H. Lo, N. Ansari, CONSUMER: a novel hybrid intrusion detection system fordistribution networks in smart grid, IEEE Trans. Emerg. Top. Comput. 1 (2013)33–44, http://dx.doi.org/10.1109/TETC.2013.2274043.

[57] S.-C. Huang, Y.-L. Lo, C.-N. Lu, Non-technical loss detection using stateestimation and analysis of variance, IEEE Trans. Power Syst. 28 (2013)2959–2966, http://dx.doi.org/10.1109/TPWRS.2012.2224891.

[58] C. Su, W. Lee, C.-K. Wen, Electricity theft detection in low voltage networkswith smart meters using state estimation, in: 2016 IEEE Int. Conf. Ind.Technol., IEEE, 2016, http://dx.doi.org/10.1109/ICIT.2016.7474800, pp.493–498.

[59] P. Jokar, N. Arianpoo, V.C.M. Leung, Electricity theft detection in AMI usingcustomers' consumption patterns, IEEE Trans. Smart Grid 7 (2016) 216–226,http://dx.doi.org/10.1109/TSG.2015.2425222.

[60] A. Jindal, A. Dua, K. Kaur, M. Singh, N. Kumar, S. Mishra, Decision tree andSVM-based data analytics for theft detection in smart grid, IEEE Trans. Ind.Informatics 12 (2016) 1005–1016, http://dx.doi.org/10.1109/TII.2016.2543145.

[61] R.D. Trevizan, A. Rossoni, A.S. Bretas, D. da Silva Gazzana, R. de PodestaMartin, N.G. Bretas, A.L. Bettiol, A. Carniato, L.F. do Nascimento Passos,Non-technical losses identification using Optimum-Path Forest and stateestimation, in: 2015 IEEE Eindhoven PowerTech, IEEE, 2015, http://dx.doi.org/10.1109/PTC.2015.7232685, pp. 1–6.

[62] J.V. Spiri´c, S.S. Stankovi´c, M.B. Do˘ci´c, T.D. Popovi´c, Using the rough set theoryto detect fraud committed by electricity customers, Int. J. Electr. Power EnergySyst. 62 (2014) 727–734, http://dx.doi.org/10.1016/j.ijepes.2014.05.004.

[63] J.V. Spiri´c, S.S. Stankovi´c, M.B. Do˘ci´c, Determining a set of suspiciouselectricity customers using statistical ACL Tukey's control charts method, Int.J. Electr. Power Energy Syst. 83 (2016) 402–410, http://dx.doi.org/10.1016/j.ijepes.2016.04.035.

[64] A. Rossoni, R. Trevizan, A. Bretas, D. Gazzana, A. Bettiol, A. Carniato, L. Passos,R. Martin, Hybrid formulation for technical and non-technical lossesestimation and identification in distribution networks: application in abrazilian power system, CIRED 23 Rd Int. Conf. Electr. Distrib. (2015), pp.15–18.

[65] J.B. Leite, J.R.S. Mantovani, Detecting and locating non-technical losses inmodern distribution networks, IEEE Trans. Smart Grid 3053 (2016), http://dx.doi.org/10.1109/TSG.2016.2574714, 1-1.

[66] Y. Guo, C.W. Ten, P. Jirutitijaroen, Online data validation for distribution operations against cyber tampering, IEEE Trans. Power Syst. 29 (2014)550–560, http://dx.doi.org/10.1109/TPWRS.2013.2282931.

الخلاصة

تعتبر سرقة الكهرباء مصدر قلق كبير للخدمات العامة. مع ظهور العدادات الذكية، ازداد تواتر جمع بيانات استهلاك الطاقة المنزلية، مما جعل من الممكن تحليل البيانات المتقدم، وهو ما لم يكن ممكنًا من قبل. في الواقع، يمكن أن يؤدي استخدام شبكات الشبكة الذكية (SG)، وهي شبكات مطورة حديثًا من الكائنات المتصلة، إلى تحسين موثوقية وكفاءة واستدامة البنية التحتية التقليدية للطاقة بشكل كبير .

تولد البنية التحتية للشبكة الذكية SG كمية هائلة من البيانات، بما في ذلك استهلاك الطاقة للمستخدمين الفرديين. باستخدام هذه البيانات، يمكن للتعلم الآلي وتقنيات التعلم العميق تحديد مستخدمي سرقة الكهرباء بدقة. تقدم هذه الأطروحة نموذجًا قائمًا على الشبكة العصبية التلافيفية (CNN) للكشف التلقائي عن سرقة الكهرباء التي يمكنها تحقيق تصنيف وكشف عالي الأداء .

يدرس هذا العمل التجريب للعثور على أفضل تكوين للنموذج المتسلسل (SM) للتصنيف، بدءًا من طبقتين وانتهاءً بأربع طبقات. تم الحصول على أفضل أداء في بنية طبقتين حيث تتكون الطبقة الأولى من 128 عقدة والطبقة الثانية 64 عقدة حيث وصلت الدقة إلى (0.92). يتيح ذلك تصميم مصنف إشارات كهربائية عالية الأداء يمكن تطبيقه في العديد من التطبيقات.

تم تصميم مصنفات إشارات الكهرباء باستخدام CNN والبيانات المستخرجة من مجموعة بيانات استهلاك الكهرباء باستخدام (SM) .بالإضافة إلى ذلك، تُستخدم خوارزمية القرد الأزرق (BM) لتقليل عدد الميزات في مجموعة البيانات، حيث تُستخدم هذه القيم لبناء نماذج ذات أداء عالٍ. في هذا الصدد، كان التركيز في هذه الرسالة على تقليل عدد الميزات المطلوبة في مجموعة البيانات من أجل الحصول على نموذج مصنف إشارات الكهرباء عالي الأداء. وقد بررت التجارب الأداء العالي للأنظمة المقترحة، حيث يتطلب الجمع بين خوارزميات CNN و BM 666 ميزة فقط مقارنة بـ 1035 ميزة باستخدام CNN وحدها. يوضح هذا تفوق نموذجي CNN و BMعلى نموذج CNN من حيث تقليل ميزات النموذج بينما تظل الدقة كما هي.

جمهورية العراق
وزارة التعليم العالي والبحث العلمي
جامعة الأنبار
كلية علوم الحاسوب وتكنولوجيا المعلومات
قسم علوم الحاسبات

إكتشاف سرقة الكهرباء في الشبكات الذكية بإستخدام التعلم العميق

رسالة مقدمة إلى:

قسم علوم الحاسبات - كلية علوم الحاسوب وتكنولوجيا المعلومات، جامعة الانبار،

وهي جزء من متطلبات نيل درجة الماجستير في علوم الحاسبات

قدمت من قبل

نور محمود إبراهيم

بإشراف

أ.د سفيان تايه فرج الجنابي

و

أ.د بلال إسماعيل الخطيب

1442 هـ **2021 م**