ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

DEEP NEURAL NETWORK-BASED STEALTHY FALSE DATA INJECTION ATTACK DETECTION ON DER INTEGRATED SYSTEMS

M.Sc. THESIS

Can GÜRKAN

Department of Electrical Engineering

Electrical Engineering Programme

JUNE 2023

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

DEEP NEURAL NETWORK-BASED STEALTHY FALSE DATA INJECTION ATTACK DETECTION ON DER INTEGRATED SYSTEMS

M.Sc. THESIS

Can Gürkan
(504191063)

Department of Electrical Engineering

Electrical Engineering Programme

Thesis Advisor: Prof. Dr. V.M. İstemihan Genç

JUNE 2023

ISTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

DEK ENTEGRE SİSTEMLERİNDE DERİN SİNİR AĞI TABANLI GİZLENMİŞ YANLIŞ VERİ ENJEKSİYON SALDIRISI TESPİTİ

YÜKSEK LİSANS TEZİ

Can GÜRKAN
(504191063)

Elektrik Mühendisliği Anabilim Dalı

Elektrik Mühendisliği Programı

Tez Danışmanı: Prof. Dr. V.M. İstemihan Genç

HAZİRAN 2023

Can GÜRKAN, a M.Sc. student of İTU Graduate School student ID 504191063, successfully defended the thesis/dissertation entitled "DEEP NEURAL NETWORK-BASED STEALTHY FALSE DATA INJECTION ATTACK DETECTION ON DER INTEGRATED SYSTEMS", which he/she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Prof. Dr. V.M.İstemihan Genç** İstanbul Technical University	.....
Jury Members :	**Prof. Dr. Emine Ayaz** Istanbul Technical University	.....
	Asst. Prof. Dr. Cavit Fatih Küçüktezcan Bahcesehir University	.....

Date of Submission : 25.05.2023
Date of Defense : 15.06.2023

To my family,

FOREWORD

First of all, I would like to thank and express my deepest gratitude and appreciation to my thesis advisor Prof. Dr. Veysel Murat İstemihan GENÇ, for his guidance, support and consulting during my master's program. I would want to convey my gratitude to Dr. Mostafa Mohammadpourfard and Necati Aksoy for their support and help in the field of Cyber Security and Machine Learning.

I am deeply grateful to my family for the invaluable mental and physical support, encouragement, guidance and unrequited love. It is great to know and feel that I have your unwavering support in my life.

Haziran 2023 Can Gürkan

TABLE OF CONTENTS

Page

ABBREVIATIONS

ANN **:** Artificial neural network
BDD : Bad data detection
CDBN : Conditional Deep Belief Network
dB : Decibel
DL : Deep learning
DNN : Deep neural network
DoS : Denial of service
EV : Electrical vehicle
ESS : Error sum-of-squares
FDIA : False data injection attack
FN : False Negative
FP : False Positive
ICT : Information and communication technologies
EMS : Energy Management Systems
K-NN : k-Nearest neighbor
LR : Logistic regression
ML : Machine learning
PMU : Phasor measurement unit
RAT : Remote Access Trojan
SCADA : Supervisory Control and Data Acquisition
SE : State estimation
SNR : Signal-to-Noise Ratio
TN : True Negative
TP : True Positive
WACS : Wide area control systems
WAMS : Wide area monitoring systems
WAPS : Wide area protections systems
WGN : White Gaussian Noise

SYMBOLS

$\boldsymbol{\mu}$	: Mean
$\boldsymbol{\sigma}$	: Standard deviation
$\boldsymbol{\mathcal{L}}$	: Likelihood function
$\boldsymbol{\theta}$	: Voltage angle
$\hat{\boldsymbol{z}}$	: Estimated measurements
$\boldsymbol{z_{bad}}$	: Attacked measurements
$\hat{\boldsymbol{x}}$	: Estimated system states
$\hat{\boldsymbol{x}}_{\boldsymbol{bad}}$	: Attacked estimated system states
$\boldsymbol{\Omega}$	: Covariance

LIST OF TABLES

Page

LIST OF FIGURES

Page

DEEP NEURAL NETWORK-BASED STEALTHY FALSE DATA INJECTION ATTACK DETECTION ON DER INTEGRATED SYSTEMS

SUMMARY

The world's rapidly growing population has led to an increase in demand for consumption, which in turn requires an energy production system that is both economically feasible and environmentally sustainable. To meet these requirements, renewable energy resources (RES) such as wind, solar, and thermal energy have been utilized to improve energy efficiency, while also adhering to stringent carbon emission regulations. By using these clean and renewable energy sources, we can achieve a more sustainable and environmentally friendly energy production system for the future. The shift towards more sustainable grid structures has resulted in a greater adoption of distributed energy resources (DERs) which allows for the generation and distribution of energy from multiple small-scale sources, rather than relying on a few large power plants. By deploying distributed energy resources (DERs) in close to consumers, we can strategically leverage on-site generation and reduce utility costs, as it eliminates the need for significant investments in expanding the power system network. This approach allows energy to be generated and consumed locally, reducing transmission losses and enabling greater control and flexibility over energy production and consumption. Consequently, DERs can offer a more cost-effective and efficient solution for meeting energy demands while also helping to reduce greenhouse gas emissions.

Because the demand for energy to growing and power system topologies and strategies rapidly evolving, traditional power systems have become inadequate in meeting the modern society's energy requirements from multiple perspectives. Conventional power networks are designed for unidirectional power flow. Conventional power networks are designed for unidirectional power flow. Additionally, traditional power networks lack the flexibility, resilience, and monitoring and control capabilities needed to effectively manage the modern energy demands. Consequently, to meet the needs of society, smart grids have replaced conventional grids.

The smart grid is a complex cyber-physical system that relies on modern information and communication technologies (ICT), advanced control systems, and the electrical grid. This system is composed of two fundamental layers: the cyber layer and the physical layer. The cyber layer includes various communication, information, and control systems that enable the smart grid to collect and analyze data, monitor performance, and facilitate decision-making. The physical layer of the smart grid consists of the electrical infrastructure that provides power to homes, businesses, and industries. This layer includes transmission and distribution lines, transformers, generators, and other equipment that are crucial for the distribution and management of electricity. Smart Grids are equipped with Remote Terminal Units (RTUs), that collects and transmits field data such as smart meters and sensors to monitor the system and retrieve data, for instance active and reactive powers flows on branches and voltages and voltage angles of buses, using ICT. Along with the benefits of smart grid structure and ICT, it may also cause issues on the grid such as cyber security and system security. Smart electrical power systems are encountered with new challenges: cyber security of the smart grids.

State estimation is a critical process that ensures the secure and reliable operation of power systems by determining the system's operating state based on available measurements. However, recent research has shown that this process can be susceptible to False Data Injection Attacks (FDIAs), where attack vectors are injected into compromised measurements to bypass bad data detection methods. With the increasing penetration of distributed energy resources (DERs), the traditional state estimation process has become more vulnerable to cyber-attacks, exacerbating the risk of successful FDIAs.

In this thesis, we first identify the available measurements and perform state estimation based on the identified measurements. We assume that an attacker targeting the grid compromises the system and gains access to the measurements and data used in state estimation. We also assume that the compromised data is used to launch a False Data Injection Attack (FDIA) on the measurements of the power system.

In this thesis, Deep Learning-based method for detecting cyber-attacks in power systems with a high penetration rate of DERs is examined. The proposed method aims to detect anomalies in measurements, that are used in state estimation, with high detection rate. The proposed approach is evaluated implementing historical hourly load data from the New York Independent System Operator (NYISO) to three IEEE systems; 14 Bus, 30 Bus and 57 Bus. To test the effectiveness of the proposed method, four different system configurations with varying levels of DER penetration were used. To reflect the real-life conditions to the work, proposed method's performance also examined under different noise levels. Proposed Deep Learning-based method's performance is compared with widely-used classification algorithms, k-Nearest Neighbor (KNN) and Logistic Regression (LR).

The results of the study indicates that LR had a higher attack detection rate than k-NN at low noise levels in the IEEE 14 Bus system. However, it was observed that the Deep Learning-based Deep Neural Networks (DNNs) was more accurate than both algorithms at both high and low noise levels. In the 30 Bus system, which has medium complexity among the introduced systems, it was observed that the k-NN algorithm detected more attacks than the LR algorithm at both low and high noise levels. Similar to the 14 Bus system, it was also observed that the DNN algorithm had a higher ability to detect attacks than both classification algorithms. DNN algorithm performed the highest attack detection ability at different noise levels in 57 Bus system. However, it was observed that the attack detection rate dropped to as low as 90% at high noise levels.

The studies have shown that DNNs perform well even in the presence of noisy measurements against False Data Injection Attacks. Although the results are satisfactory, it is possible to achieve higher attack detection rates and performances by using configurations that include different hidden layers, optimizers, loss functions, or completely different algorithms.

DEK ENTEGRE SİSTEMLERİNDE DERİN SİNİR AĞI TABANLI GİZLENMİŞ YANLIŞ VERİ ENJEKSİYON SALDIRISI TESPİTİ

ÖZET

Artan nüfusuyla birlikte tüketim talebinde artış meydana gelmiş ve bunun sonucunda ekonomik olarak mümkün ve çevresel açıdan sürdürülebilir bir enerji üretim sistemi gereksinimi ortaya çıkmıştır. Bu gereksinimleri karşılamak için rüzgar, güneş ve termal enerji gibi yenilenebilir enerji kaynakları (YEK) enerji verimliliğini artırmak için kullanılmakta ve aynı zamanda katı bir şekilde karbon emisyonu düzenlemelerine uyulmasını sağlanmaktadır. Bu temiz ve yenilenebilir enerji kaynaklarını kullanarak, gelecek için daha sürdürülebilir ve çevre dostu bir enerji üretim sistemi elde edilebilmektedir. Daha sürdürülebilir şebeke yapılarına geçiş, birkaç büyük enerji santralinden üretim ihtiyacını karşılamak yerine birden fazla küçük ölçekli kaynaktan enerji üretimine ve dağıtımına izin veren dağıtılmış enerji kaynaklarının (DEK) daha fazla benimsenmesiyle sonuçlanmıştır. Güç sistem ağının genişletilmesine yönelik önemli yatırımlara ihtiyaç, DEK'leri tüketicilere yakın noktalarda kullanarak ortadan kalkmaktadır. Bu yaklaşım, enerjinin yerel olarak üretilip tüketilmesine izin vererek, iletim kayıplarını azaltır ve enerji üretimi ve tüketimi üzerinde daha fazla kontrol ve esneklik sağlamaktadır. Sonuç olarak, DEK'ler enerji taleplerini karşılama konusunda daha ekonomik ve verimli bir çözüm sunabilmekte ve aynı zamanda sera gazı emisyonlarının azaltılmasına yardımcı olabilmektedir.

Enerji talebinin artması ve güç sistemi topolojileri ile stratejilerinin hızla değişmesi nedeniyle geleneksel güç sistemleri, birçok açıdan modern toplumun enerji gereksinimlerini karşılamada yetersiz hale gelmiştir. Geleneksel güç sistemleri şebekeleri tek yönlü güç akışı için tasarlanmıştır. Ayrıca, geleneksel güç sistemleri şebekeleri, modern enerji taleplerini etkili bir şekilde yönetmek için gereken esneklik, dayanıklılık ve izleme ve kontrol yeteneklerine sahip değillerdir. Sonuç olarak, toplumun ihtiyaçlarını karşılamak için akıllı şebekeler geleneksel şebekelerin yerini almıştır.

Akıllı şebeke, modern bilgi ve iletişim teknolojilerine (BİT), ileri kontrol sistemlerine ve elektrik şebekesine dayanan karmaşık bir siber-fiziksel sistemdir. Bu sistemler, iki temel katmandan oluşur: siber katman ve fiziksel katman. Siber katman, akıllı şebekenin veri toplamasını ve analizini, performansı izlemeyi ve karar vermeyi kolaylaştıran çeşitli iletişim, bilgi ve kontrol sistemlerini içerir. Akıllı şebekenin fiziksel katmanı ise, evlere, işletmelere ve endüstriyel bölgelere güç sağlayan elektrik altyapısını içerir. Bu altyapı, iletim ve dağıtım hatları, transformatörler, jeneratörler ve elektrik yönetimi için büyük öneme sahip diğer ekipmanları içerir.

Akıllı şebekeler, akıllı sayaçlar ve sensörler gibi saha verilerini toplamak ve iletmek için Uzak Terminal Birimi (UTB'ler) gibi donanımlara sahiptir ve bilgi ve iletişim teknolojileri kullanarak bir hat üzerindeki aktif ve reaktif güç akışları ile bara gerilimi ve gerilim açıları gibi verileri izleyebilir.

Akıllı şebeke yapısı ve bilgi ve iletişim teknolojilerinin sağladığı faydaların yanı sıra, şebekede siber güvenlik ve sistem güvenliği gibi sorunlara neden olabilmektedir. Akıllı elektrik enerjisi sistemleri, akıllı şebekelerin siber güvenliği gibi yeni sorunlar ortaya çıkmıştır.

Durum tahmini, mevcut ölçümlere dayanarak sistemdeki çalışma durumunu belirleyerek güvenli ve güvenilir bir şekilde güç sistemlerinin çalışmasını sağlayan önemli bir işlemdir. Ancak, son zamanlarda literatürde bulunan çalışmalar, bu sürecin Yanlış Veri Saldırıları (YVS) tarafından yanıltılabileceğini göstermektedir. Yanlış Veri Saldırılarında, saldırganlar siber saldırılar ile erişebildikleri ölçümlere saldırı vektörleri enjekte ederek Yanlış Veri Tespit (YVT) yöntemleri tarafından tespit edilmemeyi amaçlamaktadır. DEK'lerin artan kullanımı ile birlikte, mevcut güç sistemi şebekeleri kompleks hale gelerek geleneksel durum tahmini süreci daha fazla güvenlik açıklarına sahip hale gelmiş ve başarılı YVS riskini artırmıştır.

Bu tezde, öncelikle test sistemlerinde bulunan mevcut ölçümleri belirliyor ve belirlenen ölçümlere dayanarak durum tahmini yapılmaktadır. Şebekeye saldıran bir saldırganın sistemdeki güvenlik açıklarını kullanarak durum tahmininde kullanılan ölçümlere ve verilere erişim sağladığını varsayılarak ele geçirilen verilerin, güç sistemi ölçümlerine YVS gerçekleştirmek için kullanıldığı varsayılmaktadır. Gerçekleştirilen bu saldırıların tespit edilebilmesi için Derin Öğrenme (DÖ) tabanlı Derin Sinir Ağları (DSA) kullanılması üzerine çalışmalar gerçekleştirilmiştir. DEK'lerin yüksek güç üretim oranına sahip olduğu güç sistemlerinde siber saldırıları tespit etmek için DÖ tabanlı bir yöntem incelenmektedir. Önerilen yöntem, durum tahmininde kullanılan ölçümlerdeki anormallikleri düşük hata oranıyla tespit etmeyi hedeflemektedir. Önerilen yaklaşım, New York Independent System Operator (NYISO) tarafından sağlanan geçmişe yönelik saatlik yük verileri kullanılarak üç IEEE sisteminde (14 baralı, 30 baralı ve 57 baralı) uygulanmıştır. Önerilen yöntemin performansını test etmek için, farklı DER güç seviyelerine sahip dört farklı sistem yapılandırması kullanılmıştır. Bunlar; DEK'lerin bulunmadığı, DEK'lerin sistemde bulunan üretim tesislerinin toplam gücünün %10'u, %20'si ve %30'u kadar olduğu senaryolardır. Bununla birlikte, gerçek yaşam koşullarını çalışmaya yansıtmak için önerilen yöntemin performansı farklı gürültü seviyelerinde incelenmiştir. Önerilen Derin Öğrenme tabanlı yöntemin performansı, yaygın olarak kullanılan sınıflandırma algoritmaları olan "k En Yakın Komşu" (k-EYK) ve "Lojistik Regresyon" (LR) ile karşılaştırılmıştır.

Çalışmanın sonuçları, düşük gürültü seviyelerinde LR'nin IEEE 14 Baralı sisteminde k-EYK'den daha yüksek bir saldırı tespit oranına sahip olduğunu göstermektedir. Bununla birlikte, Derin Öğrenme tabanlı derin sinir ağlarının hem yüksek hem de düşük gürültü seviyelerinde her iki algoritmadan da daha yüksek doğruluk oranına sahip olduğu gözlemlenmiştir. Çalışmada kullanılan sistemler arasında orta seviyede kompleks olduğu varsayılan IEEE 30 Baralı sisteminde, k-EYK algoritmasının LR algoritmasına göre hem düşük hem de yüksek gürültü seviyelerinde daha fazla saldırı tespit ettiği gözlemlenmiştir. IEEE 14 Baralı sistemine benzer şekilde, DSA algoritmasının her iki sınıflandırma algoritmasından daha yüksek saldırı tespit yeteneğine sahip olduğu görülmüştür. IEEE 57 Baralı sisteminde farklı gürültü seviyelerinde DSA algoritmasının en yüksek saldırı tespit yeteneğini sergilediği gözlemlenmiştir. Bununla birlikte, yüksek gürültü seviyelerinde saldırı tespit oranının %90'a kadar düştüğü görülmüştür. Çalışmalar sonucunda, gürültülü ölçümler karşısında bile DSA, YVS karşısında iyi performans gösterdiğini göstermiştir.

Sonuçların yüksek doğruluğa sahip olmasına rağmen farklı gizli katmanlar, optimizasyon yöntemleri, kayıp fonksiyonları veya tamamen farklı algoritmalar içeren konfigürasyonlar kullanılarak daha yüksek saldırı tespit oranları ve performansları elde etmek mümkün olabilmektedir.

1. INTRODUCTION

The network of components combined to produce, deliver and use electrical energy is known as an electrical power system which is a core foundation of twenty-first century society and it is indispensable for the functioning of all other crucial infrastructure sectors including health, military, entertainment, education, energy and production.

With the rapid increase of population of the world, consumption demand has increased and this has led need for an economically and environmentally friendly energy production. Wind, solar, and thermal energy have been utilized to improve energy efficiency while adhering to rigorous carbon emission regulations through the use of renewable energy resources (RES) [1]. The 2014 World Energy Outlook Report indicates that the global energy demand is set to grow by 37% by 2040. Furthermore, the move towards more sustainable grid structures has led to an increase in the adoption of distributed energy resources (DER) [1]. Deploying distributed energy resources (DERs) close to consumers can strategically leverage on-site generation and reduce utility costs by avoiding investments in expanding the power system network. This approach enables power generation to be produced locally, eliminating the need for long-distance transmission and reducing energy losses, while also offering cost-effective power delivery to consumers that can be flexibly managed [2]. With higher DER penetration, impacts of catastrophic events, such as power outages, severe weather conditions can be significantly reduced as they can improve the resiliency and reliability of the power grid [3]. By implementing new policies in the major energy markets, the annual investment in clean energy is expected to exceed USD 2 trillion by 2030 and expected to achieve net zero carbon emissions by 2050 [4].

Due to the increasing demand for energy and the rapid evolution of power system topologies and strategies, traditional power systems have become inadequate in meeting modern society's requirements from several perspectives. One of the main shortcomings of these systems is the lack of bi-directional communication capability, which hinders real-time monitoring and control of critical components. Additionally, such systems are not capable of dynamically responding to the increasing and changing

demand of the customers. Alongside the need to accomplish increasing demand for reliable, sustainable, and efficient electrical power, innovations in ICT and the increasing role of these systems in traditional electrical networks have enabled conventional electrical networks to be replaced by smart grids. Smart grids are modern electrical power systems that allow real-time monitoring and control owing to capability of bi-directional communication. Also, it facilitates integration of renewable energy resources and EVs [5]. Smart grids are more efficient, stable, flexible, controllable and resilient to anomalies in the grid compared to conventional electrical grids.

The smart grid's main contributions can be attributed to bidirectional communication, distributed generation and EV integration. It also enables self-healing, self-monitoring, and pervasive control, while supplying high-quality, efficient, and stable electrical energy [6]. Smart grids can be described as sophisticated energy supply networks that aim to achieve the goals of providing economical, secure, and sustainable electricity supply [7].

The smart grids are complex cyber-physical systems that are based on modern ICT and the electrical grid. The cyber-physical system is represented by two layers, the cyber layer, which includes communication, information, and control systems, and the physical layer which comprises the electrical grid infrastructure. As a next-generation power system, the smart grid leverages sensor and measurement technologies, control methods, and decision algorithms to manage and deliver energy in a secure, efficient, and cost-effective manner.

Conventional power systems cover the generation, transmission, distribution and consumption processes. These processes were carried out uni-directionally, but the Smart Grid concept enable the bi-direction communication between control center and equipment (e.g., status of the breaker, sensors of the wind farm, charging status of the EV). A Smart Grid's control center is assembled with SCADA system, that has a function to carry out the analysis, computation, supervision and control of the system in accordance with the feedback information given by the communication network.[8].

Smart Grids are integrated with RTUs, which collects and transmits field data such as smart sensors to monitor the system and retrieve data, for instance powers flows and bus voltages and voltage angles, using ICT [9].

Along with the benefits of smart grid structure and ICT, it may also cause issues on the grid such as cyber security and system security. Traditionally, the secure and reliable operation of power systems has been achieved by segregating the system design, operation, and information infrastructure. Nowadays, as a consequence of integration of ICT and extensive utilization of smart grid concept, power system communication and information systems structures are changed. As a result, smart electrical power systems are encountered with new challenges: cyber security of the smart grids.

1.1 Cyber Security Concept

Cybersecurity encompasses a range of processes and technologies that are specifically developed to safeguard cyber and physical infrastructure data from potential attacks, damages, and unauthorized access [10]. The goal of a cyberattack can be unauthorized access, manipulation of data, disruption of functioning, and economic and political harm.

The significance of cybersecurity in the context of the Smart Grid is crucial due to the interconnection of numerous devices. These devices rely on a complex network infrastructure to facilitate communication, making it imperative to implement robust security measures across these networks using diverse techniques [10]. The focus of the cyber security concept is to prevent, reduce the impact, detect and recover from cyber-attacks.

1.2 Cyber Attack History

The energy sector is a critical infrastructure that is often targeted by cyber-attacks. These attacks can have devastating consequences, including disruption of power grids, theft of sensitive information, and compromise of safety systems.

In 2010, malicious computer worm “Stuxnet” targeted Iran's nuclear power plant. Stuxnet aimed to defect the nuclear power plant by gaining access over SCADA systems. Stuxnet specifically aims at PLCs, that are responsible for automation and control of electrical components’ processes. It functions by targeting Siemens Step 7 software on machines that are using Microsoft Windows [11]. Stuxnet is widely acknowledged to have infiltrated Iranian PLC, acquiring critical data on industrial

systems while inflicting substantial harm by disabling an estimated 990 fast-spinning uranium centrifuges. [12]. Also, it is stated that the Stuxnet worm has infiltrated the system via USB flash drive. The power plant's network system was physically isolated from unsecured networks such as the public internet. Which means, the Stuxnet attack on Iran's nuclear power plant is caused by human-induced vulnerability. After the Stuxnet attack, research studies started to concentrating on detection and mitigation of analogous cyber-attacks.

In 2011 it is reported that a Trojan Horse based malware, which is named BlackEnergy, compromised the numerous vital infrastructures of United States [13].

In 2012, modular computer virus named "Shamoon" targeted the Microsoft Windows operation system. This virus has an ability to spread to other computers over the local network [14]. It aims to become computer unusable by overwriting the master boot record of the computer. It targeted the Saudi Arabia and Qatar's oil companies Saudi Aramco and RasGas [15]. It affected over 30.000 computers and workstations.

In 2013, an act of cyber-attack was directed towards the Korea Hydro and Nuclear Power (KHNP). Malware known as "DarkSeoul" designed to wipe data from the infected computers [16]. The attack affected the internal network of the operator, but not for the operational systems of the nuclear power plant.

In 2013, Jacksonville Electric Authority, JEA, is targeted by distributed denial of service attack which aims to restrict the access to payment systems of the company [17].

In 2014, Havex attack has launched with the aim of gaining access over industrial control systems. Malware known as RAT is used to gain access to the targeted system[18]. It is specifically designed to target the Human-Machine Interface (HMI) software used in industrial control systems to monitor and control processes [19].

In 2015, the electric grid operators experienced their first cyber-attack that interrupts the electric grid operations [19]. The attack was triggered by an employee of the Ukrainian Electrical Distribution Systems Operator, who unwittingly activated a version of the BlackEnergy malware. Malware was hidden in the macro of the attached Excel document to the email [20]. Its aim was to collect critical data regarding the infrastructure and network and to open the way for future cyberattacks. After the data collection and system exploration phase, a cyberattack was conducted to the system.

Control systems operator stated that the pointer of the computer clicked the human-machine interface (HMI) itself and started to switch the state of the breakers of the electrical distribution system. Mouse was uncontrollable by the operator; it was controlled only by the attacker. In addition, administrator credentials were changed in order to prevent the operator from regaining access. Attack only lasted for a few minutes, and after that, additional malware was used to wipe the computer's disks out. Also, to prevent the customers from reaching the distributor to report the fault, a DoS attack was conducted on the call center [21]. At the end, 30 substations were disconnected for several hours and around 230,000 customers were affected by this cyberattack [19].

In 2017, a large-scale ransomware attack "WannaCry" has affected around 220,000 computers located in more than 150 countries [22]. WannaCry ransomware encrypts files on the infected computers and demands crypto currency to decrypt the encrypted files. It mainly targeted the older versions of the Microsoft Windows [23]. The adverse impact of the cyber-attack resulted in substantial data loss for various organizations.

Cyber-attacks have affected many organizations, users and facilities by so far. Figure 1.1 provides a chronological sequence of cyberattacks that have specifically directed the energy sector. Numerous studies have been conducted in academia and industry to develop defense mechanism against cyber-attacks, the density and impact of which have increased in the last decade, and to minimize their effects.

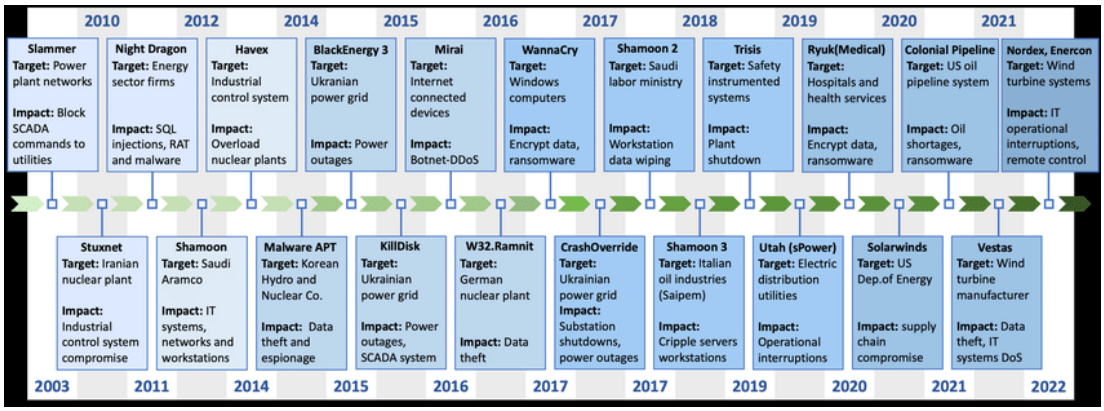


Figure 1.1 : Timeline of cyberattacks of energy sector [2].

1.3 Related Work

Liu et al. [24], conducted a study illustrating how an attacker possessing either complete or partial network information can conduct FDIA to measurements while evading detection by residual calculation-based BDDs. Yuan et al. [25], introduced the different type of FDIA, load redistribution attacks (LRAs). Power outputs of generators and zero-injection buses are not attackable and attack amount is restricted to be stealthy. Also, load and demand must be zero to achieve power balance and force operator to redistribute attack to less cost-effective state.

In practical situations, attaining comprehensive information of network proves to be nearly impractical due to the exceptionally large number of buses and lines comprising the power system network. Thus, studies started to focus on incomplete network information scenarios. In 2010, Sandberg et al. [26] introduced two security indices for state estimators to help to locate measurements which are more vulnerable to manipulate. Esmalifalak et al. [27] introduced an algorithm employing linear independent component analysis to everage system vulnerabilities, showcasing how an attacker, armed solely with access to power flow measurements, can accurately estimate both state topology and power states. Hug et al. [28] proposed a method to determine the number of measurements that an attacker needs to modify to conduct a single stealthy false data injection attack at the RTU level on state estimation.

Liu and Li et al. [29] conducted load redistribution attack with partial information of network demonstrated that obtaining the information of attack location is sufficient to conduct FDIA to measurements in attack area and to evade BDD.

Manandhar et al. [30] [31] since FDIAs can be undetectable by x^2-detector, designed a framework Kalman Filter estimator and using Kalman Filter Estimator with x^2-detector and Euclidian detector. Also, by x^2-detector and cosine similarity matching and used Kalman filter estimation to detect anomalies between actual measurements and estimated values [32].

Chaojun et al [33] introduced a methodology for detecting FDIA in AC SE. Their approach utilizes the Kullback-Leibler distance to measure the dissimilarity between two probability distributions derived from the variations in measurements.

Wang et al. [34] illustrated a cyber-physical attack by applying bad data injection (BDI) attack to smart meters. Cyber and physical aspects of the system inspected

separately by injecting bad data into information system and constructing FDIA to bypass the traditional error detection algorithms in power systems. Mohammadpourfard et al [35] proposed an unsupervised visualization-based anomaly detection method to detect FDIA using principal component analysis to reduce the dimensionality of data.

Kim et al. [36] inspected the man-in-the middle attack, conducted by altering the data sent from meters to the operator, to create a false topology control. Kim et al [37] proposed a learning approach to minimize the likelihood of attacks on multiple meters, with the objective of manipulating state estimation. Chen et al [38] introduced a real-time monitoring mechanism that utilizes the correlation of spatiotemporal network states to detect data integrity attacks. Sou and Sandberg et al. [39] proposed a method to inspect the integrity-compromising attacks by using the constraint cardinality minimization problem. Khanna et al. [40] introduced an algorithm to determine the exposed measurements by an attacker using NYISO load data in IEEE 14 Bus system using AI-based solutions.

Liu et al. [41] proposed an algorithm that detects power grid anomalies from the nominal states which are nuclear norm minimization and low rank matrix factorization. Proposed algorithms are capable of identifying nominal system operation states and maliciously infected data.

Ozay et al. [42] proposed attack methods and distributed state vector estimation methods to handling the sparsity of smart grid networks to detect and estimate the impact of unobservable FDIAs. Attack methods proposed are Distributed Sparse Attacks and Collective Sparse Attacks which are conducted by having limited network information and full network information, respectively. Later that it is shown that the supervised and semi-supervised machine learning algorithms, namely Sparse Logistic Regression, k-Nearest Neighbors and Support Vector Machine, can detect the attacks with higher performances compared to state vector estimation methods [43].

On the other hand, Goldstein et al. [44] handled the anomaly detection by using 19 unsupervised machine learning algorithms on 10 different datasets. These algorithms mainly based on density estimation using the k-Nearest Neighbors or k-means clustering. With the lead of [44], McLaughlin et al. [45] applied that method to detect the FDIA in energy theft.

The utilization of ANNs and DNNs has several main limitations, including their dependency on substantial volumes of training data, their computationally intensive nature, and challenges related to interoperability. Deep neural networks require high-performance hardware to achieve fast computations. With the advancements in computer science and technology, particularly with the development of Graphics Processing Units (GPUs), the requirement for high-performance hardware has been eased. Complex architecture of DNNs may cause interpretability problems. Also, different approaches that are studied to enhance the interpretability of DNNs such as visualization and feature attribution methods, interpretability problem can be overcome.

Ashrafuzzaman et al. [46], inspected the ability of DNNs to detect FDIAs on state estimation of the power grids on IEEE 14 Bus system by comparing DNN with widely-used machine learning. It is demonstrated that DL algorithms exhibit superior accuracy and precision compared to alternative methods.

He et al. [47], used historical data to train CDBN to use the detect FDIAs in real-time in 118 Bus IEEE test system. It is demonstrated that CDBN achieves high detection accuracy, which is more than 98.5%, in the presence of the occasional operation faults and it outperforms the ANN-based and SVM-based detection methods [47]. Similar to this study, Wei et al [48] examined the concerns of integrity of the data related to transient stability of WAMSs and presented a novel cyber-physical strategy that utilized CDBN and distributed control, incorporating real-time data for enhanced performance.

Ayad et al. [49] inspected the use of Recurrent Neural Networks (RNNs) as a machine learning technique to detect the FDIA which is tested by simulating FDIAs using IEEE 30 Bus system. It is demonstrated that RNN-based FDIA detection algorithm can accurately detect anomalies by analyzing the temporal variation in the successive data sequence [49].

2. BACKGROUND INFORMATION

2.1 State Estimation

Smart grids facilitate the integration of renewable energy sources. With this integration, rapid development and integration of energy storage systems, sensors, communication technologies, new computer technologies and distributed automated control is inevitable. Thus, the modern power systems can be acknowledged as cyber-physical systems. The term of the cyber in cyber-physical, stands for the communication/information, computation and control processes of the power system such as SCADA/EMS, WAMS, WAPS, and WACS [50]. Developments of the cyber part eases the operation of the power systems but it also increases the vulnerabilities of the systems from the cyber-security aspect [51].

Cyber-attacks can cause various harmful effects, such as altering the electricity market, degrading grid operations, compromising grid integrity, and leading to equipment damages, destruction, and even widespread blackouts [51]. Since the reliability of the power system are directly affected by the cyber-attacks, it is crucial to secure the reliability of the power system. In order to secure the power systems, techniques and algorithms are developed.

State estimation is a crucial application to provide information of the state of the system and facilitates the efficient operation of the smart grid. State Estimation can be defined as the procedure of aligning the measured data obtained from SCADA or networks. Recently, approaches have been devised to utilize network measurements for determining the network state, known as state estimators. These methods essentially employ weighted least squares techniques to identify the optimal vector that can effectively capture the data scatter [51]. Measurements made in the real-life power system have a margin of error due to disturbances, noises, metering accuracies and analog-to-digital conversions, thus this measured data scatters.

AC state estimation is a technique used to estimate the system state by using AC power flow analysis techniques. Conversely, DC SE is termed as a condensed version of AC

SE that based on only DC power flows. In DC-SE, reactive powers are completely ignored consequently it does not consider the non-linear behavior of the power system.

RTUs sample analog variables from the network and convert them into a digital format. These units collect voltage measurements from all phases of the line, even though only one measurement is required for normal operation conditions. Consequently, RTUs often gather redundant data. However, state estimator has to include all of the measurements to achieve highest estimation accuracy. accuracy. Since the power system measurements contains redundant data, state estimators are used with data-driven techniques to identify the bad data. State estimators also can be used to detect changes in the network configurations. For instance, when one phase of a transmission line is suddenly open circuited, the average power flow on the healthy phases will be much lower than the last estimation, estimated by state estimator. Then the operator will be alerted at the first state estimation cycle from that moment. Measurements and signals may become unavailable due to cyber or physical failure, can be estimated by the state estimator. Main challenges such scenarios include determining the minimum number of measurements needed for state calculation and identifying methods to enhance state estimation through additional measurements.

Recently, electrical networks have grown increasingly intricate, thus ensuring the safe operation of the system more challenging. Prior to conducting any security assessment or implementing corrective measures, it is crucial to obtain a reliable estimation of the current state of the power system. Consequently, the limitation of physical measurements cannot be restricted solely to the quantities necessary for supporting conventional power-flow programs. The conventional power-flow solution relies on just only one of the power and voltage related measurement. In real life, other conveniently measured quantities cannot be used in conventional power flow calculations. State estimation techniques based on weighted least-squares calculations offer a solution to overcome these limitations. [52].

Aim of the SE is to obtain an optimum solution for system states using real life measurements. These measurements can be analogue or continuous quantities. Analogue quantities are converted to digital via analog-to-digital converters, and then converted quantities and digital quantities telemetered to control centers via communication links. The data received at the control center is processed to provide the operator with valuable insights into the system's state. This information enables the

operator to stay informed and make informed decisions regarding system operations. In real life, the processed and telemetered data always contains inaccuracies due to errors and noises. These errors can be assessed statistically, allowing for the quantification of their impact. The estimated values of the measured quantities are then evaluated based on specific accuracy criteria. If the estimated values meet the required level of accuracy, they are considered acceptable. Conversely, if the estimated values surpass the predefined accuracy thresholds, they are deemed unreliable and rejected [52]. Since, due to noises and errors, actual values of these quantities are never known, best possible values of unknown quantities have to be estimated.

2.1.1 Maximum likelihood estimation

The main aim of the state estimation process is to obtain the most likely state of the system regarding to the measured quantities. To achieve that, maximum likelihood estimation (MLE) is used. The assumption is made that the measurement errors adhere to a probability distribution with parameters that are not yet known. Subsequently, the joint probability density function for all measurements can be formulated based on these unknown parameters. To maximize the likelihood function, an optimization can be set and by solving this optimization problem, maximum likelihood estimates of the parameters can be obtained.

The measurement errors stated before are assumed as distributed by Gaussian (Normal) Distribution which has μ mean and σ^2variance. Probability density function of Gaussian (Normal) Distribution is defined as:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left\{\frac{x-\mu}{\sigma}\right\}^2} \quad (2.\ 1)$$

where x is random variable, μ mean and σ^2variance. $f(x)$ is dependent on mean and the variance, then it can be standardized by defining;

$$z = \frac{x-\mu}{\sigma} \quad (2.\ 2)$$

where, expected value and variance of z are 0 and 1, respectively. Thus, the function can be rearranged as;

$$\Phi(z) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-z^2}{2}} \quad (2.\ 3)$$

Plot of probability density function of $f(x)$ is given in Figure 2.1

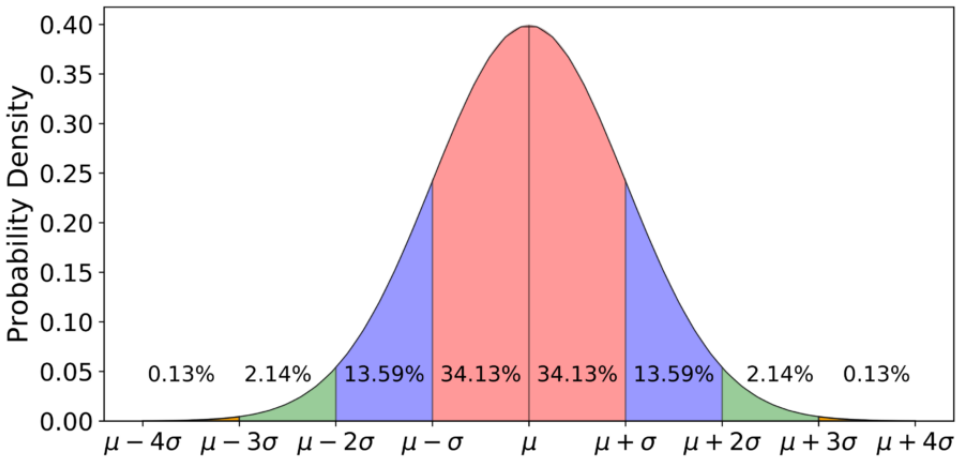


Figure 2.1 : Probability density of Gaussian Distribution [53].

The joint probability density function of m number of independent measurements can be demonstrated as the product of individual probability density function;

$$f_m(z) = f(z_1)f(z_2) \dots f(z_m) \quad (2.4)$$

where, z_i denotes ith measurement.

The defined function $f_m(z)$ is termed as likelihood function (MLF) of the z and it depicts the probability of observing specific set of measurements in the vector z.

MLE aims to maximize the likelihood function by optimizing the probability density function parameters. By using;

$$\mathcal{L} = logf_m(z) = \sum_{i=1}^{m} logf(z_i) = -\frac{1}{2}\sum_{i=1}^{m}\frac{z_i-\mu_i}{\sigma_i} - \frac{m}{2}log2\pi - \sum_{i=1}^{m} log\sigma_i \quad (2.5)$$

optimal parameters can be determined by;

maximizing the $logf_m(z)$

or

minimizing the $\sum_{i=1}^{m}\frac{z_i-\mu_i}{\sigma_i}$

The minimization process to obtain optimal solution can be expressed by, substituting mean μ_i with the expected value of set of measurements $E(z_i)$;

$$r_i = z_i - \mu_i = z_i - E(z_i) \quad (2.6)$$

z_i can be represented as a nonlinear function that establishes a connection between the system state vector x to the measurement i. Furthermore, each residual is squared and

multiplied by the inverse of the error variance associated with the corresponding measurement, effectively incorporating weighting into the calculation. Thus, minimization of Equation (2.6) can be solved by minimizing the weighted sum of squared of the residuals or solving the optimization problem;

$$\sum_{i=1}^{m} W_{ii} r_i^2 \tag{2. 7}$$

where, $W_{ii} = \sigma_i^{-2}$ and $r_i = z_i - h_i(x)$.

2.1.2 Weighted least squares state estimation

Let e denotes the errors in measuring the m number of the analog quantities (measurements) from the power system. In ideal case e is zero. Also, $(m \times 1)$ denotes the state vector to be estimated and x $(n \times 1)$ vector of non-linear functions that contains the state of the measurements. The number m must be greater than n to avoid the redundancy.

In state estimation, measurements are modelled by a set of non-linear equations as a function of the states and the errors as follows;

$$z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix} = \begin{bmatrix} h_1(x_1, x_2, \dots, x_n) \\ h_2(x_1, x_2, \dots, x_n) \\ \vdots \\ h_m(x_1, x_2, \dots, x_n) \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix} = h(x) + e \tag{2. 8}$$

In WLS-SE, it is assumed that errors have zero mean and independent from each other. Thus, measurement error covariance matrix can be formulated by;

$$E(e_i) = 0\,, i = 1,2, \dots, m \tag{2. 9}$$

$$Cov(e) = R = \begin{bmatrix} \sigma_{11} & 0 & \cdots & 0 \\ 0 & \sigma_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{mm} \end{bmatrix} \tag{2. 10}$$

The covariance matrix R, is diagonal matrix.

The solution of the state estimation problem can be formulated as a minimization of the objective function J;

$$J(x) = \sum_{i=1}^{m} \frac{(z_i - h_i(x))^2}{R_{ii}} \tag{2. 11}$$

In Equation 2.10, subscript i and ii denotes the ith entry of related vector and iith entry of the covariance matrix.To minimize the function *J(x)*, the derivative of this function must be set zero and the function *J(x)* can be expressed by using g(x);

$$g(x) = \frac{\partial J(x)}{\partial x} = -H^T(x)R^{-1}[z - h(x)] = 0\,,\ H(x) = \frac{\partial h(x)}{\partial x} \qquad (2.\ 12)$$

H(x) measurement Jacobian matrix with size of $(m\ x\ n)$. Since *g(x)* is nonlinear function, to solve the nonlinear problem, an iterative solution is required. Iterative solution can be obtained by ignoring the higher order terms of the Taylor series expansion of the derivative of the function as follows;

$$x^{k+1} = x^k + [G(x^k)]^{-1}[[H(x^k)]^T[R]^{-1}[z - h(x)]] \qquad (2.\ 13)$$

where,

$$G(x^k) = \frac{\partial\, g(x)}{\partial\, x} = H^T R^{-1} H$$

$H(x^k)$ is measurement Jacobian matrix with size of $(m\ x\ n)$.

(x^k) represents the state vector that is estimated at iteration *k*

$h(.)$ is the measurement function that creates measurements by using x^k

By using equation *(x)* into *(y)*,

$$x^{k+1} = x^k - [G(x^k)]^{-1} g(x^k) \qquad (2.\ 14)$$

Where $G(x^k)$ is gain matrix. In each iteration, the gain matrix is decomposed into its triangular variables, and the forward/backward substitution method is employed to solve the subsequent linear set equation:

$$[G(x^k)]\Delta\, x^{k+1} = H^T R^{-1}(z - h(x^k)) \qquad (2.\ 15)$$

$$\Delta\, x^{k+1} = x^{k+1} - x^k \qquad (2.\ 16)$$

$$\Delta\, x^{k+1} = G(x^k)^{-1} H^T R^{-1}(z - h(x^k)) \qquad (2.\ 17)$$

In power systems, measured values can be of various types. These measurements are mainly containing the power flows and injections and bus voltage magnitudes and phase angles. The state variables can be used to express the measurements. Measured data and state variables must be determined to build measurement function $h(x^k)$.

Assume that the system contains N buses, the state vector will have $(2N-1)$ elements, N voltage magnitudes and $(N-1)$ voltage angles. Voltage angles has one less element than voltage magnitudes since the voltage angle of reference (slack) bus is 0.

The state vector x will be in following form if the Bus 1 is chosen as slack bus;

$$x^T = [\theta_2 \quad \theta_3 \quad \dots \quad \theta_N \quad V_1 \quad V_2 \quad \dots \quad V_N] \tag{2. 18}$$

Consider that the bus i and bus j are the two buses of a complex power system given in Figure 2.2. Power injection at bus i;

$$P_i = V_i \sum_{j \in N} V_j \left(G_{ij} cos\theta_{ij} + B_{ij} sin\theta_{ij}\right) \tag{2. 19}$$

$$Q_i = V_i \sum_{j \in N} V_j \left(G_{ij} sin\theta_{ij} - B_{ij} cos\theta_{ij}\right) \tag{2. 20}$$

Power flow from bus i to bus j;

$$P_{ij} = {V_i}^2 \left(g_{si} + g_{ij}\right) - V_i V_j (g_{ij} cos\theta_{ij} + b_{ij} sin\theta_{ij}) \tag{2. 21}$$

$$Q_{ij} = -{V_i}^2 \left(b_{si} + b_{ij}\right) - V_i V_j (g_{ij} sin\theta_{ij} - b_{ij} cos\theta_{ij}) \tag{2. 22}$$

The Jacobian matrix of the measurements;

$$H = \begin{bmatrix} \frac{\partial P_{injection}}{\partial \theta} & \frac{\partial P_{injection}}{\partial V} \\ \frac{\partial P_{flow}}{\partial \theta} & \frac{\partial P_{flow}}{\partial V} \\ \frac{\partial Q_{injection}}{\partial \theta} & \frac{\partial Q_{injection}}{\partial V} \\ \frac{\partial Q_{injection}}{\partial \theta} & \frac{\partial Q_{injection}}{\partial V} \\ \frac{\partial I_{magnitude}}{\partial \theta} & \frac{\partial I_{magnitude}}{\partial V} \\ 0 & \frac{\partial V_{magnitude}}{\partial V} \end{bmatrix} \tag{2. 23}$$

Power injection at bus i and power flow measurements from bus i to bus j can obtained by:

$$\frac{\partial P_i}{\partial \theta_i} = \sum_{j=1}^{N} V_i V_j \left(-G_{ij} \sin\theta_{ij} + B_{ij} \cos\theta_{ij}\right) - V_i^2 B_{ii} \tag{2. 24}$$

$$\frac{\partial P_i}{\partial \theta_j} = V_i V_j \left(G_{ij} \sin\theta_{ij} - B_{ij} \cos\theta_{ij}\right) \tag{2. 25}$$

$$\frac{\partial P_i}{\partial V_i} = \sum_{j=1}^{N} V_j \left(G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij}\right) + V_i G_{ii} \tag{2. 26}$$

$$\frac{\partial P_i}{\partial V_j} = V_i \left(G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij}\right) \tag{2. 27}$$

$$\frac{\partial Q_i}{\partial \theta_i} = \sum_{j=1}^{N} V_i V_j \left(G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij}\right) - V_i^2 G_{ii} \quad (2.\ 28)$$

$$\frac{\partial PQ_i}{\partial \theta_j} = V_i V_j \left(-G_{ij} \cos\theta_{ij} - B_{ij} \sin\theta_{ij}\right) \quad (2.\ 29)$$

$$\frac{\partial Q_i}{\partial V_i} = \sum_{j=1}^{N} V_j \left(G_{ij} \sin\theta_{ij} - B_{ij} \cos\theta_{ij}\right) - V_i G_{ii} \quad (2.\ 30)$$

$$\frac{\partial Q_i}{\partial V_j} = V_i \left(G_{ij} \sin\theta_{ij} - B_{ij} \cos\theta_{ij}\right) \quad (2.\ 31)$$

$$\frac{\partial P_{ij}}{\partial \theta_i} = V_i V_j \left(g_{ij} \sin\theta_{ij} - b_{ij} \cos\theta_{ij}\right) \quad (2.\ 32)$$

$$\frac{\partial P_{ij}}{\partial \theta_j} = -V_i V_j \left(g_{ij} \sin\theta_{ij} - b_{ij} \cos\theta_{ij}\right) \quad (2.\ 33)$$

$$\frac{\partial P_{ij}}{\partial V_i} = -V_j \left(g_{ij} \cos\theta_{ij} + b_{ij} \sin\theta_{ij}\right) + 2\left(g_{ij} + g_{si}\right) V_i \quad (2.\ 34)$$

$$\frac{\partial P_{ij}}{\partial V_j} = -V_i \left(g_{ij} \cos\theta_{ij} + b_{ij} \sin\theta_{ij}\right) \quad (2.\ 35)$$

$$\frac{\partial Q_{ij}}{\partial \theta_i} = -V_i V_j \left(g_{ij} \cos\theta_{ij} + b_{ij} \sin\theta_{ij}\right) \quad (2.\ 36)$$

$$\frac{\partial Q_{ij}}{\partial V_i} = -V_j \left(g_{ij} \sin\theta_{ij} - b_{ij} \cos\theta_{ij}\right) - 2V_i \left(b_{ij} + b_{si}\right) \quad (2.\ 37)$$

$$\frac{\partial Q_{ij}}{\partial V_j} = -V_i \left(g_{ij} \sin\theta_{ij} - b_{ij} \cos\theta_{ij}\right) \quad (2.\ 38)$$

$$\frac{\partial V_i}{\partial V_i} = 1, \frac{\partial V_i}{\partial V_j} = 0, \frac{\partial V_i}{\partial \theta_i} = 0, \frac{\partial V_i}{\partial \theta_j} = 0 \quad (2.\ 39)$$

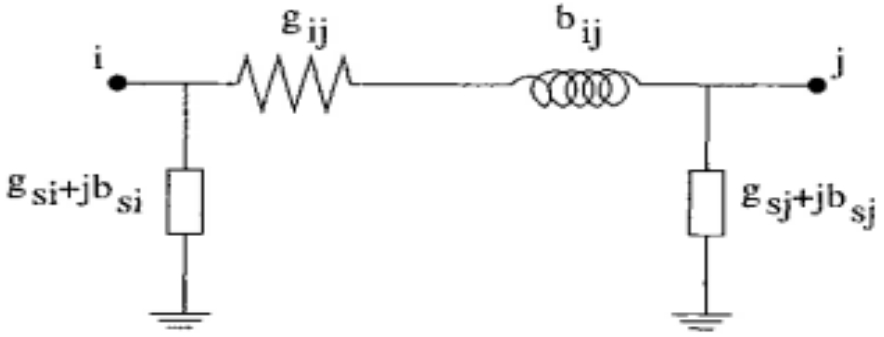


Figure 2.2 : Pi-Model of a network branch.

Gain matrix is built by using the measurement Jacobian matrix and the measurement error covariance matrix R. It is product of transpose of measurement Jacobian matrix, inverse of measurement error covariance matrix and measurement Jacobian matrix;

$$G(x^k) = H^T R^{-1} H \quad (2.\ 40)$$

Gain matrix is numerically symmetric, sparse and generally non-negative. Since H_i is sparse matrix, any mathematical expression using H_i will be also a sparse matrix. Non-zero terms in G can be calculated and stored in sparse form [52].

Representation of a matrix by multiplication of a lower triangular matrix and transpose of this lower triangular matrix is called as Cholesky decomposition. Gain matrix G also can be represented by Cholesky decomposition by;

$$G = L \cdot L^T = \begin{bmatrix} a & 0 & 0 \\ a & a & 0 \\ a & a & a \end{bmatrix} \begin{bmatrix} a & a & a \\ 0 & a & a \\ 0 & 0 & a \end{bmatrix} \quad (2.41)$$

To solve the Normal equation for Δx^k, using Cholesky decomposition on the Eq. (2.41) ;

$$LL^T \Delta x^k = t^k = H^T R^{-1}(z - h(x^k)) \quad (2.42)$$

In forward substitution, it is assumed that $L^T \Delta x^k = u$, and elements of u starting from u_1 by using substitutions in the transformed equation $Lu = t^k$ are obtained. First row provides the solution for the first step u_1 as t_1/L_{11}. Repeating the same procedure for u_2 and the rest of the elements are used to compute the solution for u.

Since u is computed at the forward substitution process, $L^T \Delta x^k$ is used to compute the entries of Δx^k. In back substitution, computing process needs to be started from the bottom row n. Starting from $\Delta x^k(n) = t_n/L_{nn}$, back substitution process is processed to first row.

WLS-SE process can be summarized by;

1- Start iterations by calculating x^k, k=0.

2- Calculate gain matrix $G(x^k)$.

3- Calculate the t^k using Forward/Backward substitution.

4- Decompose $G(x^k)$ and solve for Δx^k.

5- Check for the convergence if Δx^k is convenient, $\Delta x^k \leq \varepsilon$.

- If $\Delta x^k \leq \varepsilon$, end the iterations,
- If $\Delta x^k > \varepsilon$, continue iterations from the Step 2.

2.2 Bad Data Detection

Main function of SE includes bad data detection process, which aim to detect the measurement errors. Measurement errors can be caused by accuracy of the metering devices, analog-to-digital conversions, noises or possible adversary attack. This error can be filtered by state estimation by providing sufficient measurements to the state estimator.

Most of the bad measurements can be detected easily such as negative currents or voltages, larger or smaller than expected values or power demand on non-loaded buses and so forth. Unlikely, not all types of the bad measurements can be detected directly. For that, state estimators use other detection and identification methods to detect and identify bad measurements. Widely used conventional bad data detection methods are based on statistical distributions and residual based detection algorithms and most common algorithms used are "Chi-Squares Distribution BDD" and "Normalized Residuals BDD" respectively.

2.2.1 Chi-squares χ^2 distribution

$X_1, X_2, \dots, X_N$ are independent random variables where each X_i is distributed according to the Standard Normal distribution. A new random variable *Y* defined by;

$$Y = \sum_{i=1}^{N} X_i^2 \tag{2. 43}$$

will have chi (χ^2) distribution with degrees of freedom N. Degrees of freedom N indicates the number of independent variables in the sum of squares. Consider a function $f(x)$, defined in terms of the measurement errors;

$$f(x) = \sum_{i=1}^{m} R_{ii}^{-1} e_i^2 = \sum_{i=1}^{m} \left(\frac{e_i}{\sqrt{R_{ii}}}\right)^2 = \sum_{i=1}^{m} (e_i^N)^2 \tag{2. 44}$$

where, e_i is the ith measurement error, R_{ii} is the diagonal entry of the measurement covariance matrix and m is the total number of measurements. It is assumed that the all measurement errors are normally distributed random variables with zero mean and variance of R_{ii}.

Thus, $f(x)$ will have maximum $(m-n)$ degrees of freedom. Since at least n measurements are required to calculate the power balance equations, at most $(m-n)$ of the measurement errors will be linearly independent [52].

In power systems since there are 2n state variables, which are bus voltage angles and bus voltage angles. Degrees of freedom of the power system is calculated as $(m-(2n-1))$.

Chi (χ^2) distribution probability density function is given in Figure 2.3.

$$P(X \geq x_t) = \int_{x_t}^{\infty} \chi^2\,(u) \cdot du \tag{2. 45}$$

Eq. (2.45) represents the probability of X being larger than the pre-determined threshold x_t.

WLS-SE objective function, $J(x)$, can be used as $f(x)$ for bad data detection test by;

1- Solving the WLS estimation problem and computing the objective function;

$$J(\hat{x}) = \sum_{i=1}^{m} \frac{(z_i - h_i(\hat{x}))^2}{R_{ii}^2} \tag{2. 46}$$

where $\hat{x}$ is estimated state vector with dimension n, $h_i(\hat{x})$ is estimated measurement i, z_i is measured value of the measurement i, R_{ii}^2 is the variance of the error in measurement i and m is the number of measurements.

2- Check the calculated value from the Chi-squares distribution table corresponding detection confidence with probability and corresponding degrees of freedom, $(m-n)$.

3- Compare the value obtained from the Chi-squares distribution table with calculated $J(\hat{x})$. If $J(\hat{x})$ is greater than chi-squares distribution value, then bad data will be suspected. If not, it is accepted that bad data do not exist in measurements.

2.2.2 Normalized residuals

Linearized measurement equations are diagonal matrix relies on the presumption that the measurement errors are not correlated.

$$\Delta z = H\Delta x + e \tag{2. 47}$$

where, E(e)=0, and Cov(e)=R.

WLS estimator of the linearized state vector is;

$$\Delta\hat{x} = (H^T R^{-1} H)^{-1} H^T R^{-1} \Delta z = G^{-1} H^T R^{-1} \Delta z \tag{2. 48}$$

Thus,

$$\Delta\hat{z} = H\Delta\hat{x} = K\Delta z \tag{2. 49}$$

where, $K = HG^{-1}H^T R^{-1}$ is called hat matrix. Using the properties above, residuals can be expressed as:

$$r = \Delta z - \Delta\hat{z} \tag{2. 50}$$

$$r = (I - K)\Delta z = (I - K)(H\Delta x + e) = (I - K)e = S\, e \tag{2. 51}$$

Matrix S is the denoted as residual sensitivity matrix. Since the WLS estimation is relies on the presumption that the measurement errors are distributed according to Normal distribution, mean and covariance of the can be obtained by;

$$E(r) = E(S \cdot e) = S \cdot E(e) = 0 \tag{2. 52}$$

$$Cov(r) = \Omega = E[rr^T] \tag{2. 53}$$

$$Cov(r) = S \cdot E[rr^T] \cdot S^T = SRS^T = SR \tag{2. 54}$$

After calculating S matrix, normalized residuals can be obtained by following formulations:

$$r_i = z_i - h_i(\hat{x}) \tag{2. 55}$$

$$r_i^N = \frac{|r_i|}{\sqrt{\Omega_{ii}}} = \frac{|r_i|}{\sqrt{R_{ii}S_{ii}}} \tag{2. 56}$$

where i is the measurement number, r_i^N is the normalized value of residuals at i^{th} indices and Ω_{ii} is the i^{th} diagonal entry of the sensitivity matrix S. Existence of bad data can be detected by checking largest element in r^N against the threshold value.

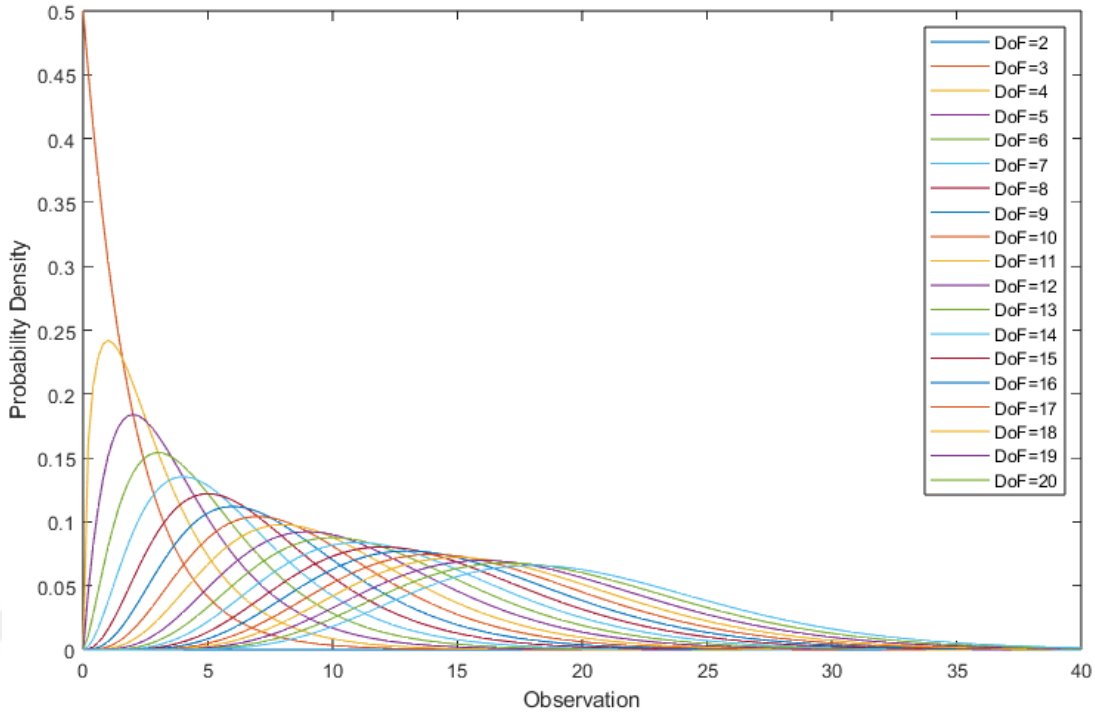


Figure 2.3 : Probability density function of different levels of degrees of freedom of Chi-Squares Distribution

2.3 False Data Injection Attacks

False data injection attacks (FDIAs) aim to inject malicious measurements that deludes the state estimation process being detected by bad data detection algorithms. FDIAs can be conducted by modifying sensor readings, manipulating control signals, or modifying system parameters. FDIAs can have cascading results.

Consider that x, z and e matrices denote the state variables of power system, the meter measurements and measurement errors, respectively and $m \geq n$. It is commonly accepted that the errors of the measurements are Normal (Gaussian) Distributed, with zero mean and σ^2 covariance.

$$x = [x_1 \quad x_2 \quad \dots \quad x_n]^T \tag{2. 57}$$

$$z = [z_1 \quad z_2 \quad \dots \quad z_m]^T \tag{2. 58}$$

$$e = [e_1 \quad e_2 \quad \dots \quad e_m]^T \tag{2. 59}$$

In power system, state vector, x, mainly contains the voltage angles and voltage magnitudes of power systems' buses. Thus x can be represented by a matrix of voltage angles and voltage magnitudes,

$$x = [\theta_1 \quad \theta_2 \quad \dots \quad \theta_n \; V_1 \quad V_2 \quad \dots \quad V_n \,]^T \tag{2. 60}$$

State variables can be obtained by using meter measurements, errors and state function $h(x)$.

$$z = H(x) + e \tag{2. 61}$$

H is the Jacobian Matrix (topology matrix) decided by topology structure of power system and parameter attribute of the circuits which is calculated by using Eq. 2.23. $H(x)$ is the nonlinear vector function of state variables. As it is stated in 2.1, main aim of the state estimation is to obtain the optimal solution of weighted least squares problem for $\hat{x}$ by minimizing the objective function $J(x)$ by iterative approach.

$$J(x) = (z - H(x))^T \, W \, (z - H(x)) \tag{2. 62}$$

where, W denotes the weight matrix, used to denote the variance of measurements which is $W = R^{-1}$. Minimization of objective function for can be represented by;

$$\hat{x} = arg \; \min_x J(x) \tag{2. 63}$$

Most common bad data detection algorithms are explained at 2.2, chi (χ^2) distribution based and residuals-based detection algorithms. Chi (χ^2) test can be formulated by Eq. 2.46 which can be also represented by;

$$J(\hat{x}) = \sum_{i=1}^{m} \frac{(z_i - h_i(\hat{x}))^2}{R_{ii}^2} = \sum_{i=1}^{m} \frac{r_i^2}{\sigma_{ii}^2} \tag{2. 64}$$

Residuals-based BDD simply relies on difference between measured values, z, and estimated values of these measurements $H(\hat{x})$ which can be formulated by,

$$r = z - H(\hat{x}) \tag{2. 65}$$

In both detection algorithms, result is compared with pre-determined threshold and if the result of the algorithm is greater than the threshold then presence of the bad data can be suspected. False data injections attacks can be done by building an attack vector to manipulate the measurements without being detected by BDD algorithms. Since it is not brute forcing the measurements and tries to be undetected, it is also called as "Stealthy False Data Injection Attack".

z_{bad} is manipulated (attacked) measurements by an attacker and $\hat{x}_{bad}$ is the manipulated estimated system variables by set of manipulated measurements z_{bad}. They can be formulated by;

$$z_{bad} = z + a \tag{2. 66}$$

$$\hat{x}_{bad} = \hat{x} + c \tag{2. 67}$$

where, z is unattacked measurements, a is nonzero attack matrix that contains the attack vector, $\hat{x}$ is unattacked system variables by set of unattacked measurements z, c is difference between attacked estimated system variables $\hat{x}_{bad}$ and $\hat{x}$ caused by attack vector a. Suppose that unattacked (original) measurements can pass the BDD, the attacked measurements z_{bad} can also pass the BDD if only attack vector a provides;

$$a = H(c) \tag{2. 68}$$

It can be proven by calculating $\hat{x}_{bad}$ using z_{bad} that provides Eq. 2.68.

$$\begin{aligned}\hat{x}_{bad} &= (H^T(\hat{x})\, W\, H(\hat{x}))^{-1}\, H^T(\hat{x}) W z_{bad} \\ &= (H^T(\hat{x})\, W\, H(\hat{x}))^{-1}\, H^T(\hat{x}) W (z + a) \\ &= \hat{x} \;+\; (H^T(\hat{x})\, W\, H(\hat{x}))^{-1}\, H^T(\hat{x})\, W\, a\end{aligned}$$

If Eq. 2.68 is provided,

$$\begin{aligned}\|z_{bad} - H(\hat{x}_{bad})\| &= \|z + a - H(\hat{x})(\hat{x} \;+\; (H^T(\hat{x})\, W\, H(\hat{x}))^{-1}\, H^T(\hat{x})\, W\, a)\| \\ &= \|z - H(\hat{x}) \;+\; (a \;-\; H(\hat{x})\, (H^T(\hat{x})\, W\, H(\hat{x}))^{-1}\, H^T(\hat{x})\, W\, a)\| \\ &= \|z - H(\hat{x}) \;+\; (H(c) \;-\; H(\hat{x})\, (H^T(\hat{x})\, W\, H(\hat{x}))^{-1}\, H^T(\hat{x})\, W\, H(c))\| \\ &= \|z - H(\hat{x}) \;+\; (H(c) \;-\; H(c))\| \\ &= \|z - H(\hat{x})\,\|\end{aligned} \tag{2. 69}$$

Thus, residuals of the measurements z_{bad} became less than threshold, and z_{bad} can pass the bad data detector. Measurement attack vector, a, is built by using arbitrary nonzero vector, c, to mislead the state estimator without being detected by bad data detector.

2.4 Machine Learning

Algorithms are the key element from the beginning for the computer science. An algorithm is a series of instructions to obtain the intended output from the given input. Despite the significant advancements in computer science, certain tasks still lack algorithms to perform them. Machine Learning (ML) is a technique to perform those tasks without building specific algorithm.

ML algorithms enables computer systems to improve their performance on a specific task over time. In ML, a generic model is adapted to perform a specific task by adjusting its parameters using the training data. With the current state of innovation of technology, the collection and storage of data have become significantly easier. Therefore, ML has been made almost universally applicable by combining it with collected data in almost every field such as classification, detection, recognition, clustering, identification, segmentation, analysis, filtering, perception, control, planning, detection, diagnosis and so forth.

ML has different sets of learning methods and algorithms to process data. These algorithms and learning methods depend on the input data that wanted to be processed and desired output data. Machine learning uses three types of techniques that are supervised learning, unsupervised learning and reinforced learning. Widely used machine learning algorithms in literature are given in Figure 2.4.

- Supervised machine learning algorithms and models' predictions and learning processes are relying on labeled inputs. Labeled dataset indicates that the inputs and outputs of the dataset are known. Main aim of supervised learning is to generate a function that match inputs with labels (output) and to predict the output of new inputs using that function.
- Unsupervised machine learning algorithms and models' predictions and learning processes are based on unlabeled input. Unsupervised

learning methods cannot be directly used for regression or classification problems since the output data is not known. The model acquires valuable properties of the input data. It is expected to find patterns and conclude the valuable properties of the input data instead of directly pointing what ML algorithm expected to learn.

- Reinforced learning can be considered as a hybrid of the supervised and unsupervised learning since it is supervised as in supervised learning and also input output links are not clear as in unsupervised learning. Rather, in reinforced learning the algorithm checks the feedback from the environment after linking the pairs. With the reinforcement arriving from environment, the algorithm completes the learning goal.

2.4.1 Logistic Regression

Logistic Regression is a statistical model based on the probability of an occurrence of an event by function of input variables. Logistic term in Logistic Regression indicates the Logistic function which is S-shaped curve also known as sigmoid curve. Logistic function can be formulated by;

$$f(x) = \frac{L}{1+e^{-k(x-x_0)}} \tag{2. 70}$$

where, x_0 is the x value of the function's midpoint, L is the supremum of the values of function, k is the logistic growth rate or steepness of the curve. Sigmoid curve can be represented by $f(x)$, where L =1, $k = 1$ and $x_0 = 0$ and its graph is given in Figure 2.4.

Logistic Regression can be formulated by using logistic function in form of;

$$p(x) = \frac{1}{1+e^{-(x-\mu)/s}} = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}} \tag{2. 71}$$

where, μ is the midpoint of the curve, s is the scale parameter.

$$\beta_0 = -\mu/s \tag{2. 72}$$

Figure 2.4 : Standard logistic function (Sigmoid Function).

$$\beta_1 = -1/s \tag{2. 73}$$

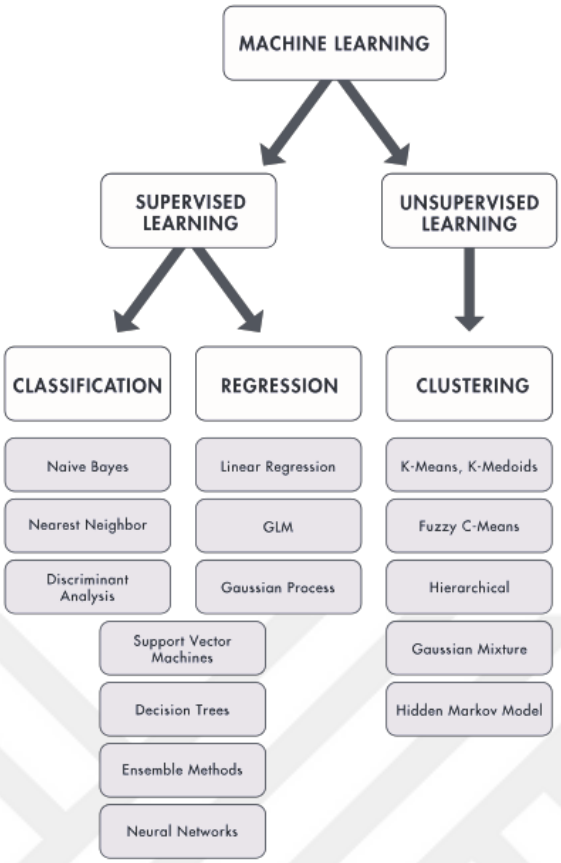
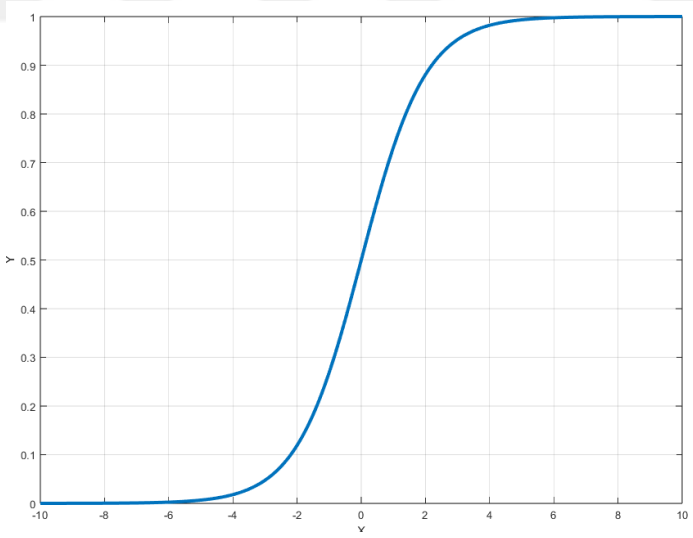


Figure 2.5 : Machine Learning Algorithms [54].



Also, standard logistic function can be defined by using $\sigma: \mathbb{R} \rightarrow (0,1)$;

$$\sigma(t) = \frac{e^t}{1+e^t} = \frac{1}{1+e^{-t}} \qquad (2.\ 74)$$

where $t = \beta_0 + \beta_1 x.$ (2. 75)

In logistic regression model, $p(x)$ indicates the success probability of the Y variable. It also describes the probability of an event Y happening as a function of x variables. Maximizing the likelihood function can be done by iterative process.

For a sample size of observations n, the likelihood function for a logistic regression is;

$$\mathcal{L}(\beta; y, X) = \prod_{i=1}^{n} p_i^{y_i}(1-p_i)^{1-y_i} \quad (2.\ 76)$$

$$= \prod_{i=1}^{n} \left(\frac{e^{X_i\beta}}{1+e^{X_i\beta}}\right)^{y_i} \left(\frac{1}{1+e^{X_i\beta}}\right)^{1-y_i} \quad (2.\ 77)$$

This yields the log likelihood;

$$\ell(\beta) = \sum_{i=1}^{n} [y_i\, log(p_i) + (1-y_i) log(1-p_i)] \quad (2.\ 78)$$

$$= \sum_{i=1}^{n} [y_i X_i \beta - log(1+e^{X_i\beta})] \quad (2.\ 79)$$

where, n is the number of samples in training data and y_i is the label (output) of the i^{th} sample. Graphical representation of the classification using LR algorithm is given in Figure 2.6, where classification is done by using sigmoid boundary.

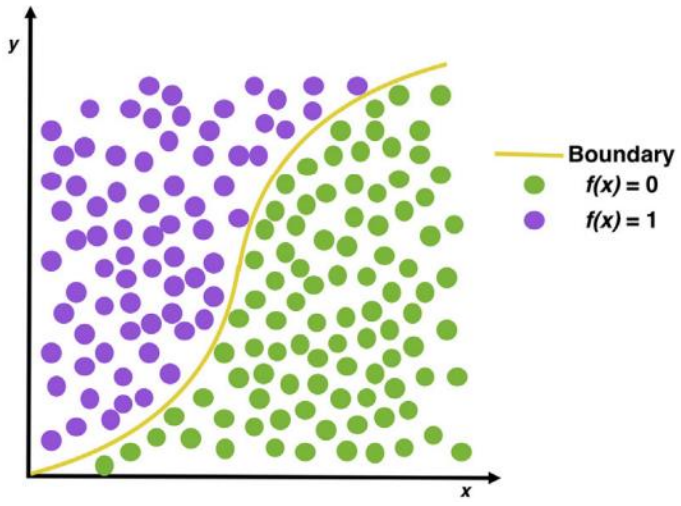


Figure 2.6 : Graphical representation of traditional statistical approaches to Logistic Regression, adapted from [55].

2.4.2 k-Nearest Neighbor

k-Nearest Neighbor (k-NN) is a commonly used classifying algorithm that assigns the samples to class of their nearest neighbors. To classify a new test data, k-NN algorithm finds the "k" nearest neighbors to the system training data and uses the labels of nearest neighbors to weight the unlabeled data. Different distance metrics are used to calculate the distance between the labeled samples and sample wanted to be predicted. Euclidian distance, Manhattan distance, which is also known as L_1norm, are the most common distance metrics to calculate distance between samples.

Euclidian distance between two samples can be calculated by,

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{2. 80}$$

Minkowski distance between two samples can be calculated by,

$$d(x,y) = (\sum_{i=1}^{n}|x_i - y_i|^p)^{\frac{1}{p}} \tag{2. 81}$$

where, $d(x,y) \geq 0$

$d(x,y) = 0$ if $x = y$

$d(x,y) = d(y,x)$

$d(x,z) \leq d(x,y) + d(y,z)$

To determine the class of the unlabeled sample belongs to, distance is calculated and compared with the labeled samples. As a result, the test signal is classified as the same label as the signal with the closest distance. In k-NN, it is crucial to select the number of neighbors to classify the sample, k value, properly.

Classification using k-NN can be done by:

- Selecting k value which represents the number of neighbors to be considered while classifying.
- Calculating the distance of the unlabeled sample with respect to the labeled samples. If the distances are relatively large numbers, normalization can be used.
- The k nearest neighbors of the related distances is considered while labeling. Sample is assigned to the class (label) of k neighbors or neighbors according to the attributes of the sample.

2.5 Deep Neural Networks

Artificial Neural Networks (ANNs) are the type of computer algorithm designed to learn and make predictions based on patterns in data, inspired by the structure and function of the human brain. Traditional machine learning requires manual specification of features for decision-making, while ANNs can learn features autonomously and improve with more data. ANNs require more data, training time, and computing power, but notes that they often outperform classical machine learning methods in complex tasks, which justifies these requirements.

Deep Neural Networks (DNNs) are a type of ANNs which has capability of learning and recognizing complex patterns in data. DNNs include multiple layers of artificial neurons to process input and produce output such as prediction or decision. Each hidden layer of DNNs receives input from the previous layer, process the data and transmits it to the next layer. The main difference between DNN and ANN is the number of hidden layers. While DNNs can be in a multi-layered structure depending on the requirement of the input data, ANNs have generally one or two hidden layers. Deep neural networks can be applied especially in cases where there is no mathematical model or algorithm for the solution of non-linear, multidimensional, noisy, complex, uncertain, incomplete data and problems.

DNNs are widely used for; probabilistic function estimation, classification, pattern recognition, time series analysis, signal filtering, data compression and feature extraction, nonlinear signal processing, nonlinear system modeling, optimization control problems. Simple artificial neuron model (ANN) is given in Figure 2.7, which basically composed by inputs, weights, bias, activation function and output.

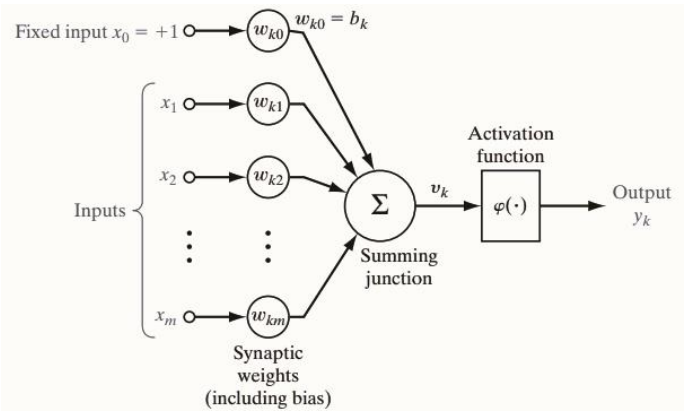


Figure 2.7 : Artificial Neuron Model with bias, adapted from [56].

The Perceptron is a network model consisting of a single artificial neural cell, and it is the mathematical representation of a neuron and it can only generate two inputs, which are zero or one, from multiple input. General perceptron models consist of three parts that are input, computation and output layers. The working principle of perceptrons can be summarized as;

- Multiply the values in the input layer by the weight values,
- Add the bias,
- If the result is greater than the threshold value, output 1, otherwise output 0.

Perceptrons are the type of artificial neuron that constituents the artificial neural network can be mathematically represented by:

$$a_k = \sum_{i=1}^{m} w_{ki} x_k + b_k \tag{2. 82}$$

$$y_k = f(a_k + b_k) \tag{2. 83}$$

where, function *f(·)* is non-linear activation.

Single-layer perceptron models are inadequate for complex operations, consequently multi-layer perceptrons are used to create deep neural network models. Since there are many parameters that effects the performance of the DNNs, such as number of inputs, number of hidden layers, learning rate etc., it is crucial to optimize the learning process to obtain more accurate results. Generic models of the ANN and DNN are given in Figure 2.8. The gradient for this optimization process can be calculated by taking the derivative of the error with respect to the weight and bias. The main aim of this process is to minimize the error.

In this thesis, "Adam" optimization algorithm is used to optimize the weight and bias. Adam is a stochastic optimization technique that utilizes first-order gradients with minimal memory requirements. It calculates adaptive learning rates for various parameters using estimates of the first and second moments of the gradients [58].

Using adaptive learning rates facilitates the minimizing errors [59]. The name Adam is derived from adaptive moment estimation. Algorithm of the Adam optimization algorithm is given in Figure 2.9.

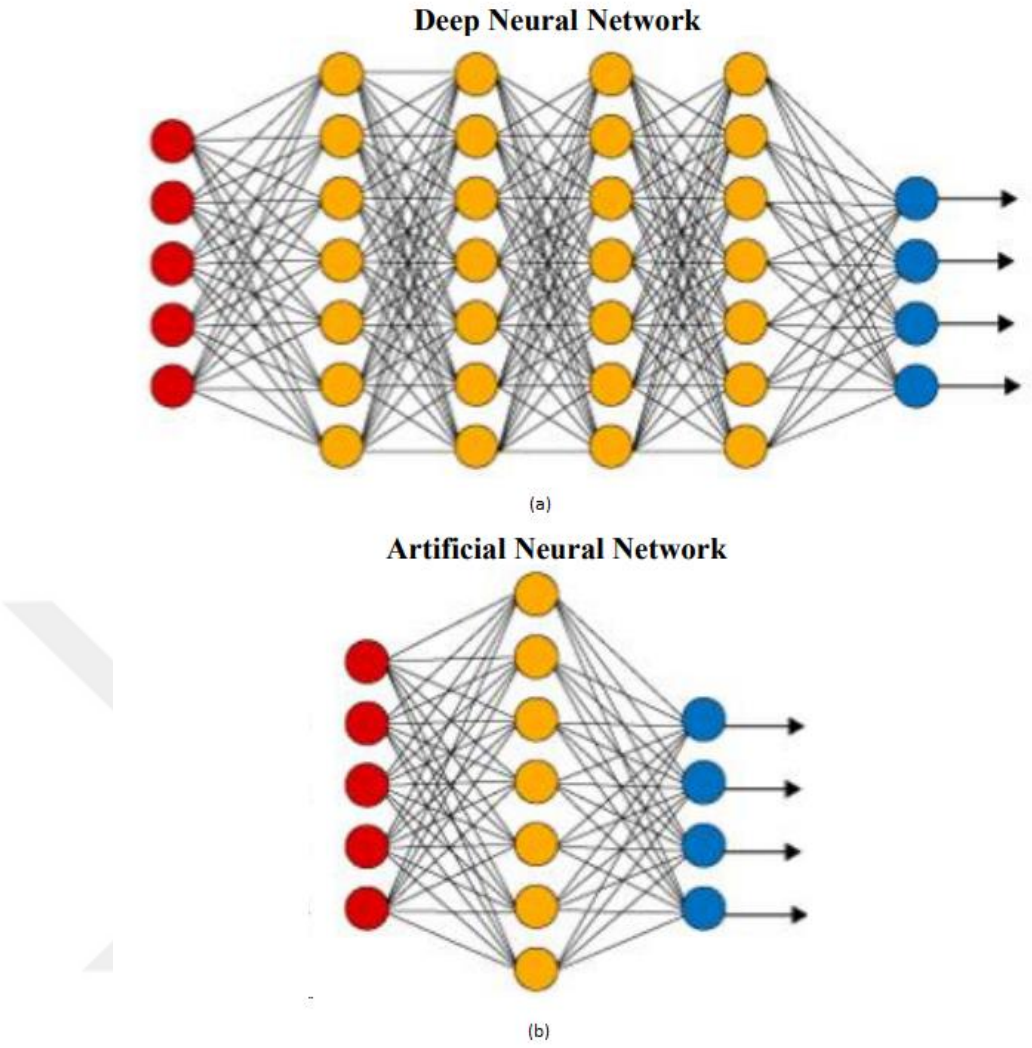


Figure 2.8 : (a) Deep Neural Network, (b) Artificial Neural Network generic models, adapted from [57].

Require: α: Stepsize
Require: $\beta_1, \beta_2 \in [0, 1)$: Exponential decay rates for the moment estimates
Require: $f(\theta)$: Stochastic objective function with parameters θ
Require: θ_0: Initial parameter vector
$m_0 \leftarrow 0$ (Initialize 1st moment vector)
$v_0 \leftarrow 0$ (Initialize 2nd moment vector)
$t \leftarrow 0$ (Initialize timestep)
while θ_t not converged **do**
$t \leftarrow t + 1$
$g_t \leftarrow \nabla_\theta f_t(\theta_{t-1})$ (Get gradients w.r.t. stochastic objective at timestep t)
$m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$ (Update biased first moment estimate)
$v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$ (Update biased second raw moment estimate)
$\widehat{m}_t \leftarrow m_t / (1 - \beta_1^t)$ (Compute bias-corrected first moment estimate)
$\widehat{v}_t \leftarrow v_t / (1 - \beta_2^t)$ (Compute bias-corrected second raw moment estimate)
$\theta_t \leftarrow \theta_{t-1} - \alpha \cdot \widehat{m}_t / (\sqrt{\widehat{v}_t} + \epsilon)$ (Update parameters)
end while
return θ_t (Resulting parameters)

Figure 2.9 : Adam optimization algorithm, retrieved from [58].

Loss functions is a mathematical function that measures difference between predicted output and target. Main aim of the optimization function is to minimize the loss function. In this thesis, binary cross entropy loss function is used alongside the Adam algorithm. It is widely-used binary classification algorithm using cross entropy formula. Binary cross entropy can be formulated by:

$$\text{BCE} = -(ylog(p) + (1-y)log(1-p)) \tag{2. 84}$$

where, y is the target and $p = Pr(y = 1)$.

The output of a layer in a neural network is determined by activation function that transforms weighted sum of inputs. Activation function influences the learning ability of neural network. Typically, the activation function used for all the perceptions in a hidden layer is the same, whereas the activation function for the output layer is dependent on the desired prediction.

Graphical representations of the widely-used activation functions are given in Figure 2.10.

- Sigmoid,

$$f(x) = \frac{1}{1+e^{-ax}} \tag{2. 85}$$

- Arctangent,

$$f(x) = \frac{2}{\pi} arctan(ax) \tag{2. 86}$$

- Rectified Linear Unit (ReLU),

$$f(x) = \begin{cases} 0 & if\ x \leq 0 \\ f(x) & if\ > 0 \end{cases} \tag{2. 87}$$

The learning ability of artificial neural networks is the ability to make accurate predictions by adjusting the weight and bias to the optimum value. Initially, the weight is usually randomly generated, and then the weight and bias are adjusted to obtain the optimal value.

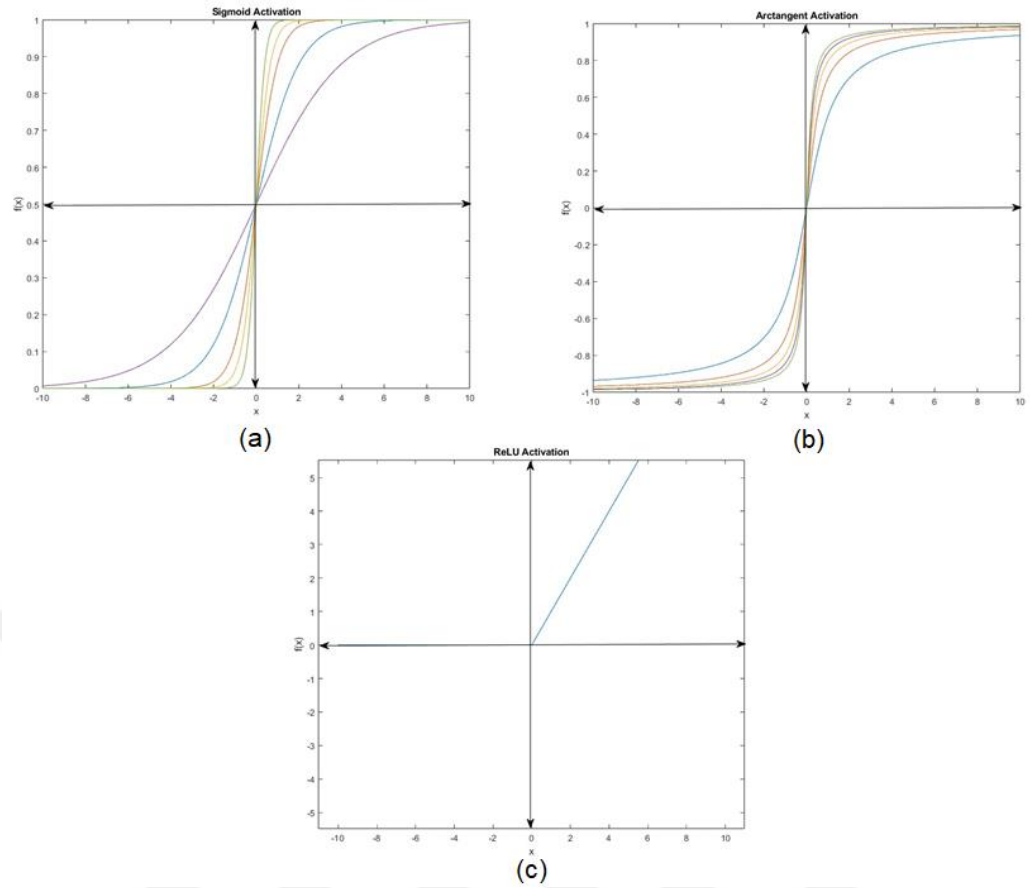


Figure 2.10 : (a) Sigmoid (b) Arctangent (c) ReLU activation functions.

2.6 F-1 Score

F-Score is a widely-used performance index in ML applications especially in binary classification problems. It relies on weighted average of precision and recall. Precision shows that how accurate the model is predicting positive values and recall is useful to measure the strength of a model to predict positive outcomes [60]. Precision and recall can be calculated by using confusion matrix indices. The confusion matrix is a tool commonly used in evaluating the performance of classification models, where it provides a matrix that allows for the comparison of predicted and actual values of a test set. Indices of confusion matrix is given in Figure 2.11. True positives (TP), denotes that the number of observations that are correctly predicted as positive by the algorithm. False positives (FP), denotes that the number of observations that are incorrectly predicted as positive by the algorithm. True negatives (TN), denotes that the number of observations that are correctly predicted as negative by the algorithm. False Negative (FN), denotes that the number of observations that are incorrectly predicted as negative by the algorithm.

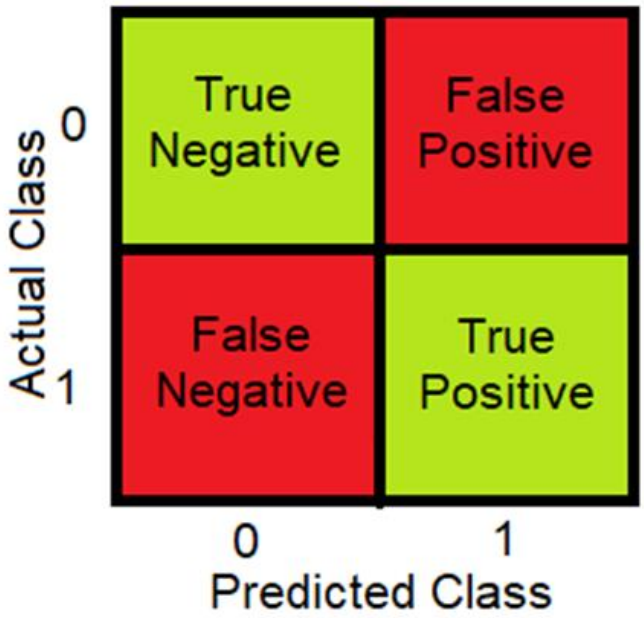


Figure 2.11 : Indices of confusion matrix.

Precision and recall can be calculated by using;

$$Precision = \frac{TP}{TP+FP} \tag{2. 88}$$

$$Recall = \frac{TP}{TP+FN} \tag{2. 89}$$

and Z-Score using precision and recall can be formulated by;

$$F_\beta = (1+\beta^2) \times \frac{Precision \times Recall}{(\beta^2 \times Precision)+Recall} \tag{2. 90}$$

In the given equation, assigning different values to β can determine the significance of each term. Most widely utilized value for β is 1, which is termed as F-1 Score. F score can be maximum value of 1, which is the best score, and minimum value of 0.

3. METHODOLOGY

In the methodology section methods and notations employed for detecting false data injection attacks using deep neural networks are explained.

The process begins with the generation of a dataset derived from the test system, which is subsequently utilized for training the proposed algorithm. Initially, the parameters of the test system, including bus data, generator data, and branch data, are defined to establish the initial conditions. Subsequently, load data corresponding to the test system is selected, comprising both active and reactive power demands of the buses. The acquired load data is then mapped to the load buses of the test system, and normalization is applied regarding to the initial loads of the load buses. Additionally, the generation of the system's generators is scaled proportionally to the normalized loads to maintain the load-generation balance. However, if load data is observed using the test system, which is not obtained by implementing another system’s load data, normalizing and generator scaling processes can be skipped.

Following the definition of the test system and load data, the addition of DERs to the system occurs. This is achieved by determining the buses to which DERs will be added and their power generation contributions. Additionally, the loads of the system's are scaled proportionally to the DERs contribution to maintain the load-generation balance.

Available measurements are defined to initialize the state estimation process. Locations of the measuring devices such as PMUs and wattmeters are defined. At this point, the crucial aspect to be noted is the observability of the system. If an insufficient number of measurements are provided for state estimation, the system becomes unobservable, which directly effects the accuracy and reliability of the state estimation process.

Following the definition of available measurements, state estimation occurs. After state estimation of the system with un-attacked measurements, a decision must be

made regarding the occurrence of a false data injection attack. If FDIA is not conducted, un-attacked measurements are labelled as “0” which indicates the absence of the attack and saved to the data set. On the other hand, if FDIA is conducted, the first step involves obtaining the H matrix from the state estimation process. Subsequently, the attack vector is created using the obtained H matrix. The attack vector is added to the measurements, and the state estimation process is executed once again using the attacked measurements. After state estimation with attacked measurements, inputs and results are inspected by BDD to detect anomalies. If attacked measurements are detected by BDD algorithms, which means the attack can be easily detected by conventional BDD methods, measurements and results are not saved to the dataset. On the other hand, if the attack is able to bypass the BDD algorithms, measurements and results are saved with the label “1” which means attack is conducted to the measurements and not detected by conventional methods. Flowchart of data generation process is given in Figure 3.1.

Dataset is generated by repeating these processes for available load data samples. DNN applications follows the dataset generation processes. Before training DNNs, the configuration of the DNN is determined based on the dataset parameters, which is decisive in determining the performance of the DNN. Configuration of DNNs includes the parameters such as numbers, input and output sizes, activation functions, kernel initializers of hidden layers and output layer, activation functions, loss functions, epochs and batch sizes of the DNNs. Since the configuration of the DNNs are set, DNNs can be trained with generated datasets.

Inspecting only the accuracy of the DNN can be misleading, as the primary objective is to detect FDIAs. Along with the F-1 score and general performance of the DNNs, it is crucial to inspect the rate of false negatives. False negatives indicate that the attack is conducted on the system and the DNN is predicts the attacked measurements as if the attack does not exist. Consequently, the rate of false negatives to the total number of attacked measurements, expressed as a percentage, determines the effectiveness of DNN on attack detection.

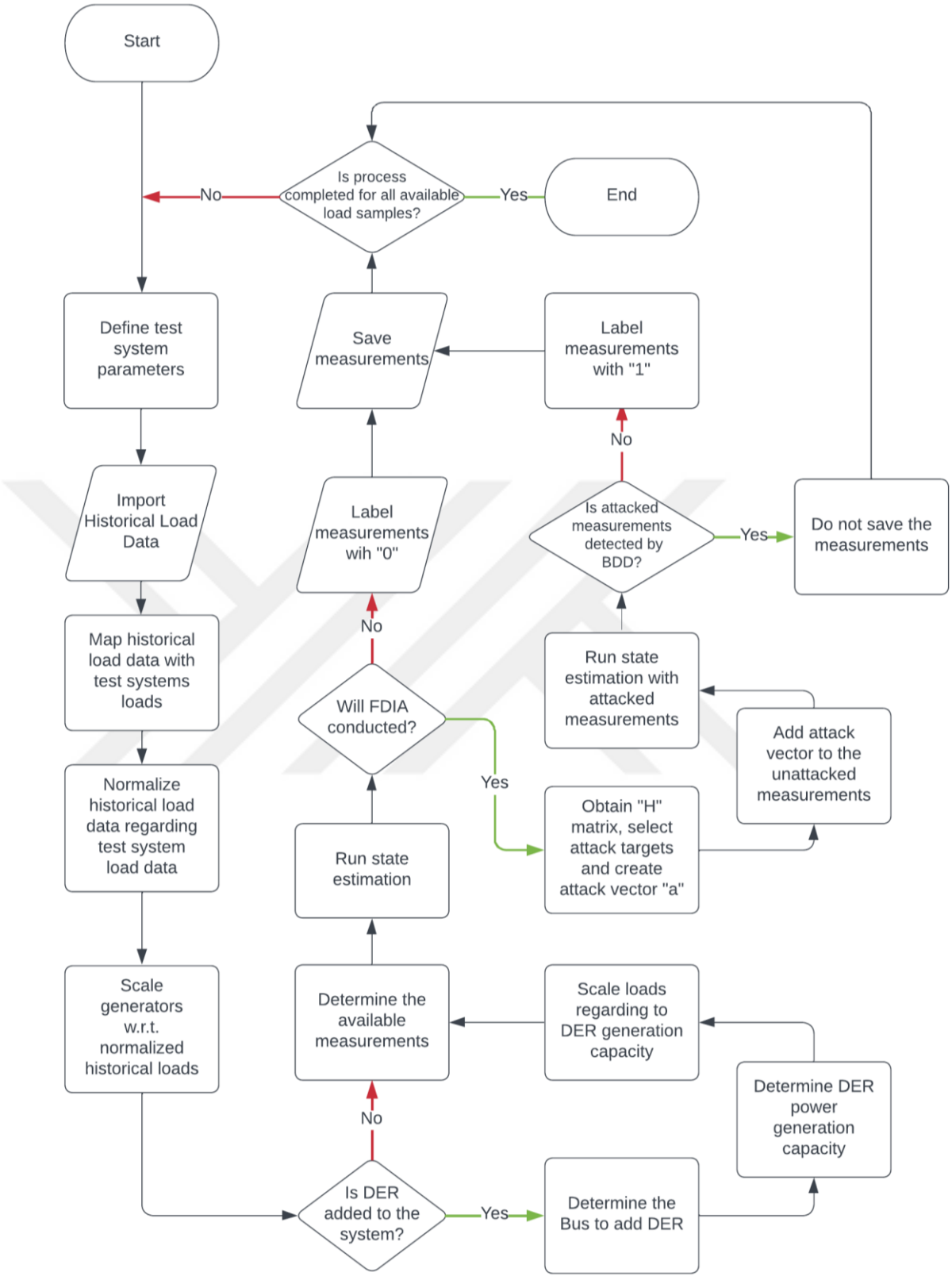


Figure 3.1 : Flowchart of the dataset generation process.

4. TEST SYSTEMS AND SCENARIOS

In this section, we present technical details related to the system where the tests are performed. Within the scope of the performed tests, we create datasets to examine false data injection attacks on systems with a high penetration of distributed energy resources. We measure the ability of three different machine learning algorithms to detect false data injection attacks at different noise levels using these datasets.

4.1 Test Systems and Dataset Generation

The datasets to be created for ML applications was prepared using historical load data of the real system, as done in [61]. The methods used to generate dataset in this thesis are replicating the methods used in [61]. The load data used in the test system are based on the New York Independent System Operator (NYISO) from January 2012 to October 2012 and from June 2013 to April 2016 hourly load. In NYISO zone, there are 11 load regions.

The following steps were followed to create the load model:

- Each bus of IEEE test system bus system is linked with one region of NYISO load region buses. Since there are is no load demand at various buses of IEEE test systems, these buses are not linked with the NYISO loads. For instance, 11 load demanding buses of IEEE 14 bus system and 11 regions of NYISO are linked.
- NYISO load is normalized regarding to the IEEE test system default load data in order to operate around the initial state of IEEE test system.
- If the number of buses in the system being used is greater than the number of NYISO load regions, the NYISO load data is reused but with the addition of a normally distributed random variable during normalization of the data to avoid the linearity.

- The normalized loads are summed and their ratio to the default IEEE test system's loads is found. The ratio found is used to scale the generators to the normalized loads.
- Different power levels of White Gaussian Noise (WGN), which is a common type of noise signal with a stationary power spectral density (PSD) that remains constant across all frequencies and characterized by a Gaussian (normal) distribution that the noise values at each point in time are independent and identically distributed with a mean of zero and a variance of one, are added to the measurements to demonstrate the noises that exists in practical applications.

NYISO load data from 01.01.2012 to 02.01.2012 is given in Figure 4.2. Also, the table containing the 2-hour load data is given in Table 4.1, loads both given in figure and table are in kilowatts.

In total, 30,947 sample of load data obtained for each bus. In case scenarios it is assumed that distributed energy resources are added to the system at different levels with power ratings of 10%, 20%, and 30% of the total generation capacity of the existing generators in the system. Thus, an additional 92,841 data samples are included in the dataset that contains DERs for each test system.

In order to enhance the practical relevance of the tests, attacks are only conducted on a subset of data from January 2012 to October 2012. As power systems are not typically subjected to attacks every hour, it is assumed that the entire dataset is not available to attack.

FDIAs aim to mislead state estimation by manipulating measured quantities, such as active power flow, reactive power flow, power generation, voltage angle, and voltage magnitude. To manipulate these measurements, attack vector, a, need to be built, as discussed in Chapter 2.3. The matrix can be constructed with knowledge of the function of state variables, $H(\cdot)$, and the attack target vector, c. Therefore, the attack target, c. needs to be selected. In 14 Bus system, there are 26, 58 and 112 quantities to attack of IEEE 14 Bus, IEEE 30 Bus and IEEE 57 bus systems, respectively. These quantities include the voltage angles and voltage magnitudes of buses of the test systems except slack bus voltage angle and voltage magnitude.

The voltage angle and voltage magnitude of the slack bus cannot be attacked. The attack vector, c, is generated by randomly selecting the attack target. After selecting the bus, the attack vector scale is chosen by randomly selecting the scale between 0.75 and 1.25 which means attack aims to increase the targeted quantity between 7.5% and 12.5%. Gaussian normal distribution is used to select both the attack target and attack scale. In this thesis, AC-SE (AC state estimation) is used to estimate the operating state of the power system, taking into account the non-linear behavior of the system due to the effects of reactive power.

After selecting the attack target and scale, the false data injection attack (FDIA) is conducted on the measurements. Chi-squares (χ^2) and residual BDD methods are subsequently applied to the attacked measurements. If an attacked measurement can evade detection by both of these algorithms, it is recorded as a successful attack in the dataset and labeled with a "1," indicating that the attack was executed and bypassed the conventional BDD algorithms.

The dataset used in this study consisted of 123,788 samples with distributed energy resources (DER) but no attacks. To augment this dataset with cyber-attack scenarios, an additional 49,344 samples were added. This additional dataset was divided in half, with one half containing un-attacked measurements from January 2012 to October 2012 and the other half containing attacked measurements data from the same time period. But it is not guaranteed that the all attacked measurements are bypassed the BDD algorithms and labeled with a “1”. Consequently, the number of the labeled as “1” in dataset can be differ regarding to detection ability of conventional BDD algorithms. Dataset generation process, iterative solutions of state estimation, FDIA and BDD operations were performed using MATLAB [62]. State estimation iterative process MATLAB code is adapted from MatPOWER [63]. For each test system these processes are repeated and consequently 3 different dataset is generated for 3 IEEE test systems to use for ML and DL applications.

4.1.1 System 1: IEEE 14 Bus system

Single line diagram of IEEE 14 bus system is given in Figure 4.3. It consists of 14 buses, 2 generators (Bus 1 and Bus 2), 3 synchronous condensers (Bus 3, Bus 6 and Bus 8), 11 load buses, 3 transformers and 20 branches. Available measurements for the IEEE 30 Bus system are given in Table 4.2

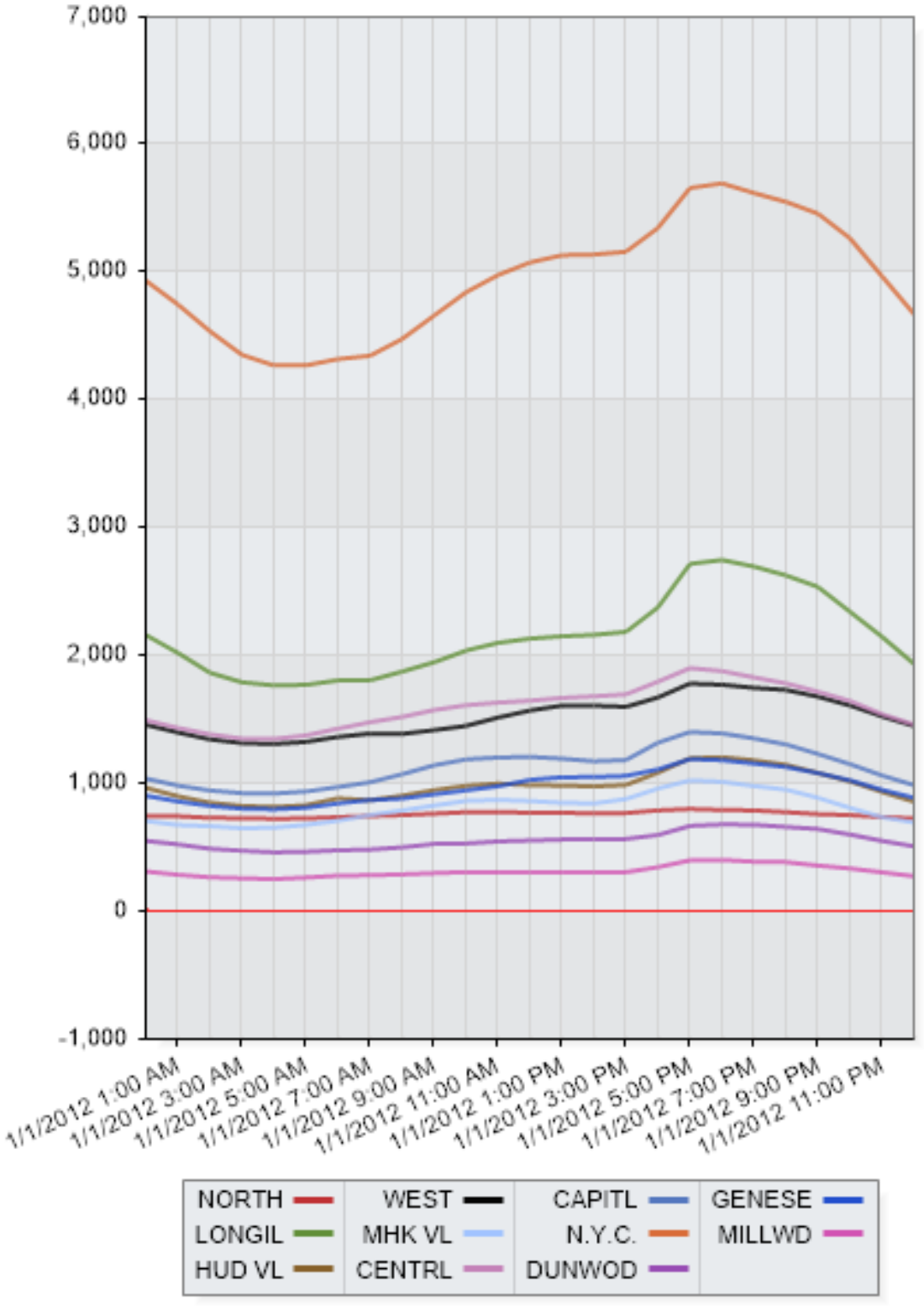


Figure 4.2 : NYISO load data from 01/01/2012 to 02/01/2012 in kilowatts [64].

Table 4.1 : Example NYISO load data for 24 hours.

Date	Load (kW)	Zone
1.01.2012 00:00	10408	CAPITL
1.01.2012 00:00	14973	CENTRL
1.01.2012 00:00	5533	DUNWOD
1.01.2012 00:00	9068	GENESE
1.01.2012 00:00	9691	HUD VL
1.01.2012 00:00	21648	LONGIL
1.01.2012 00:00	7096	MHK VL
1.01.2012 00:00	3135	MILLWD
1.01.2012 00:00	49375	N.Y.C.
1.01.2012 00:00	7478	NORTH
1.01.2012 00:00	14622	WEST
1.01.2012 01:00	9861	CAPITL
1.01.2012 01:00	14375	CENTRL
1.01.2012 01:00	526	DUNWOD
1.01.2012 01:00	8624	GENESE
1.01.2012 01:00	9024	HUD VL
1.01.2012 01:00	20247	LONGIL
1.01.2012 01:00	6778	MHK VL
1.01.2012 01:00	2883	MILLWD
1.01.2012 01:00	47521	N.Y.C.
1.01.2012 01:00	7471	NORTH
1.01.2012 01:00	14023	WEST
1.01.2012 02:00	9477	CAPITL
1.01.2012 02:00	13873	CENTRL
1.01.2012 02:00	4917	DUNWOD
1.01.2012 02:00	8271	GENESE
1.01.2012 02:00	8523	HUD VL
1.01.2012 02:00	18696	LONGIL
1.01.2012 02:00	6687	MHK VL
1.01.2012 02:00	2684	MILLWD
1.01.2012 02:00	45426	N.Y.C.
1.01.2012 02:00	7337	NORTH
1.01.2012 02:00	1348	WEST

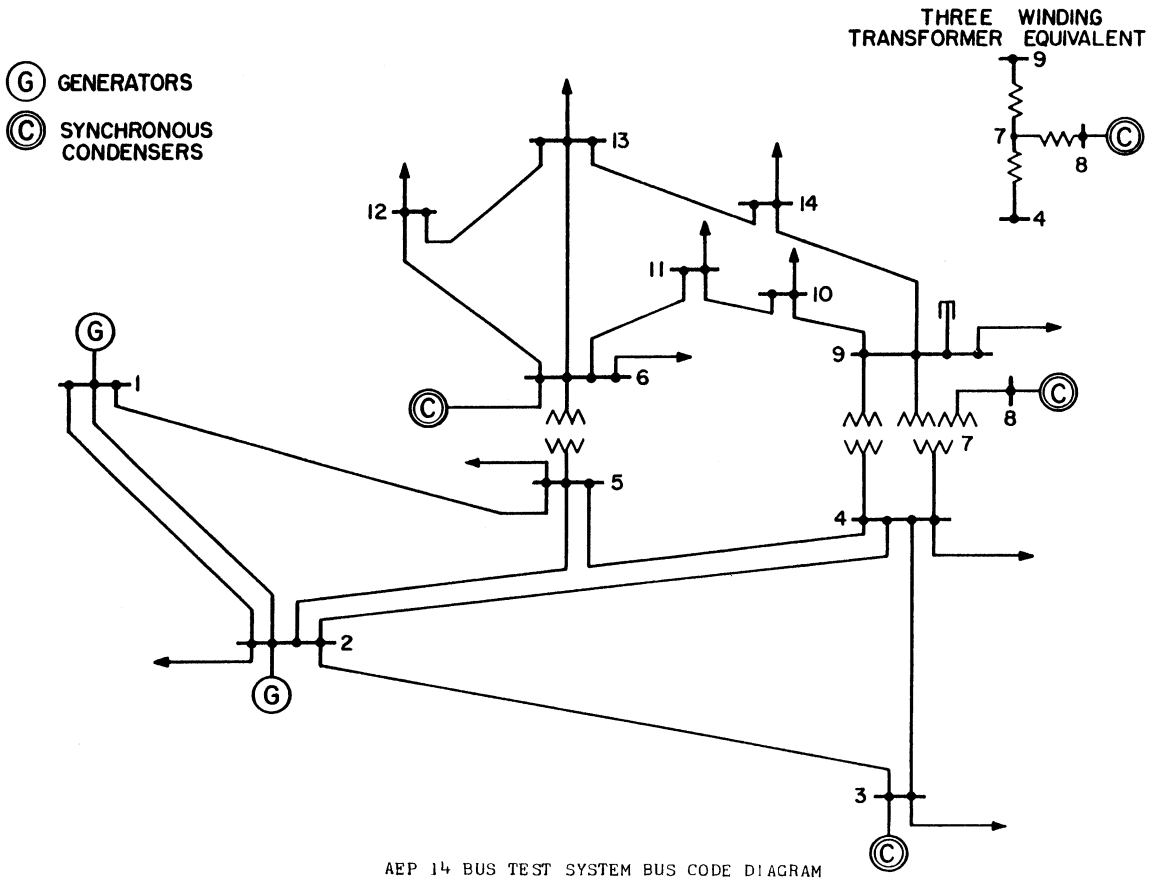


Figure 4.3 : IEEE 14 Bus System, adapted from [65].

Table 4.2 : Available measurements of IEEE 14 Bus System.

Available Measurements	
Active Power From (Branch)	1,3,8,9,10,13,15,16,17,19
Active Power To (Branch)	4,5,7,11
Active Power Generation (Generator)	1,2,3,4,5
Reactive Power From (Branch)	1,3,8,9,10,13,15,19
Reactive Power To (Branch)	4,5,7,11
Reactive Power Generation (Generator)	1,2
Voltage Magnitude (Bus)	2,3,6,8,10,14

4.1.2 System 2: IEEE 30 Bus system

Single line diagram of IEEE 30 bus system is given in Figure 4.4. It consists of 30 buses, 2 generators (Bus 1 and Bus 2), 4 synchronous condensers (Bus 5, Bus 8, Bus 11 and Bus 13), 21 load buses, 9 transformers and 41 branches. Available measurements for the IEEE 30 Bus system are given in Table 4.3.

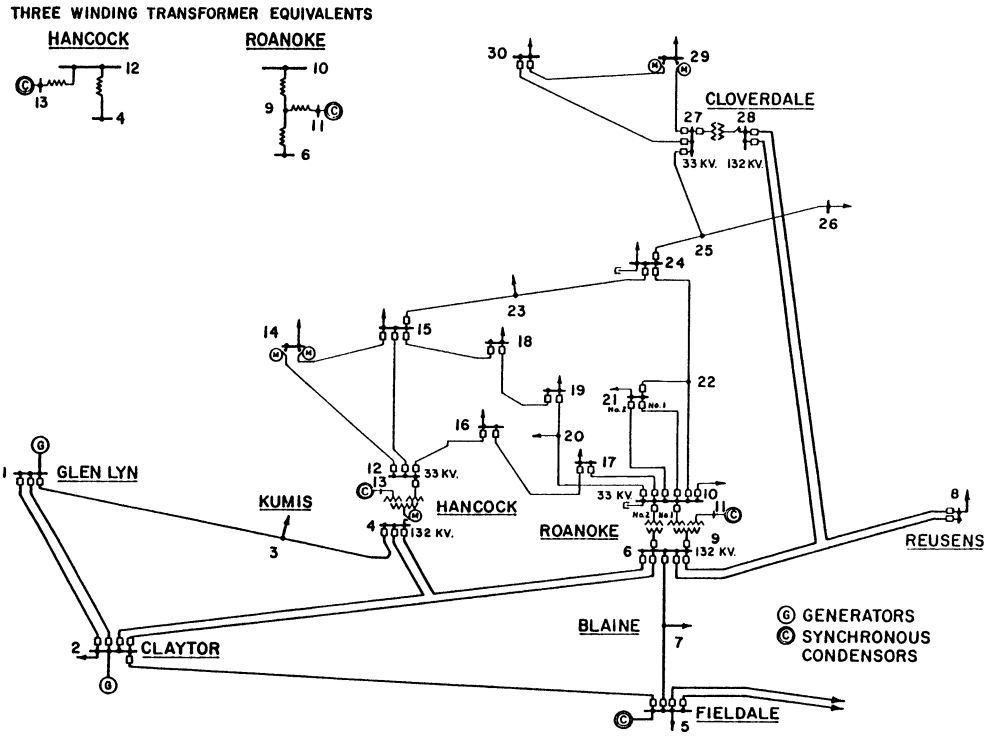


Figure 4.4 : IEEE 30 Bus System, adapted from [66].

Table 4.3 : Available measurements of IEEE 30 Bus System.

	Available Measurements
Active Power From (Branch)	1,2,3,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,22,24,25, 26,27,28,30,34,35,37,38,40,41
Active Power To (Branch)	4,10,11,15,16,18,20,23,33,35,36
Active Power Generation (Generator)	1,2,3,4,5,6
Reactive Power From (Branch)	1,2,3,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,22,24,25, 29,27,28,30,34,35,37,38,40,41
Reactive Power To (Branch)	10,12,16,20,33,36
Reactive Power Generation (Generator)	1,2
Voltage Magnitude (Bus)	1,2,5,8,11,13

4.1.3 System 3: IEEE 57 Bus system

Single line diagram of IEEE 57 bus system is given in Figure 4.5. It consists of 57 buses, 4 generators (Bus 1, Bus 3, Bus 8 and Bus 12), 3 synchronous condensers (Bus 2, Bus 6 and Bus 9), 42 load buses, 3 transformers and 80 branches. Available measurements for the IEEE 57 Bus system are given in Table 4.4.

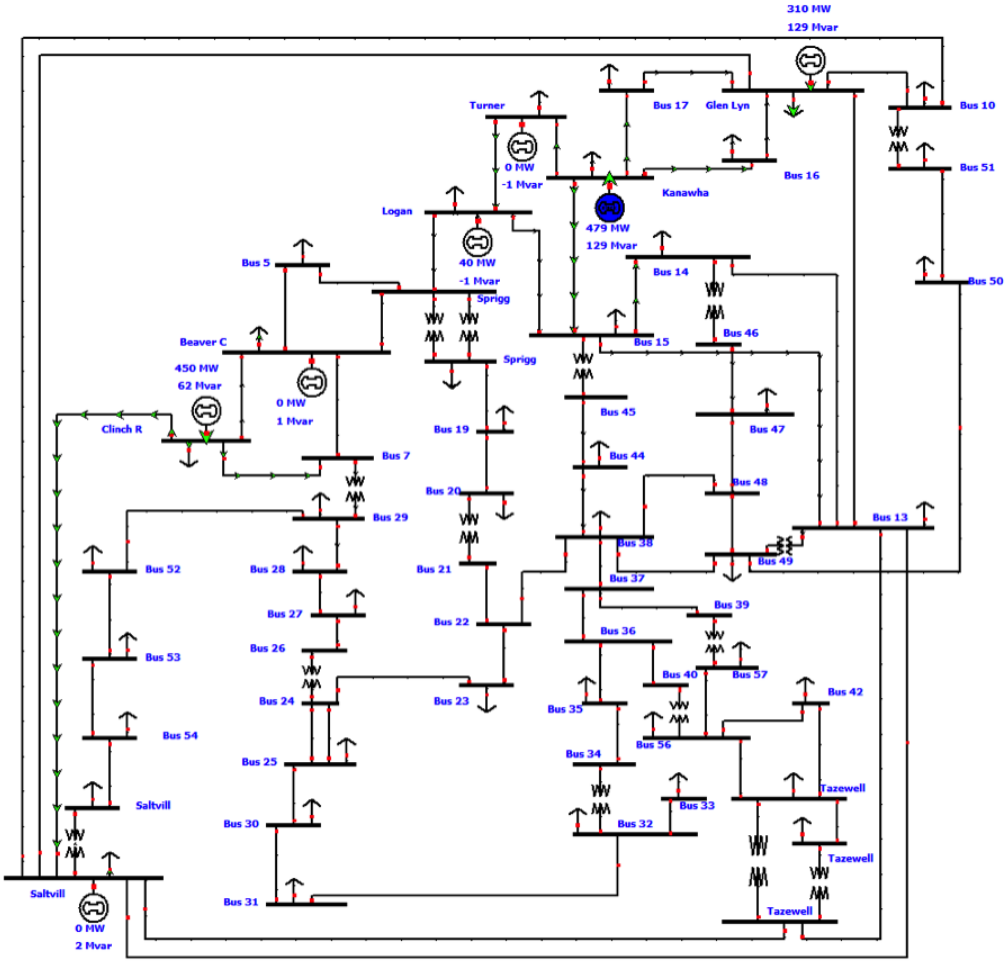


Figure 4.5 : IEEE 57 Bus System, adapted from [67].

Table 4.4 : Available measurements of IEEE 57 Bus System.

	Available Measurements
Active Power From (Branch)	1,2,3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 26, 27, 28, 29, 30, 32, 33, 34, 35, 36, 37, 39, 40, 41, 42,43, 44, 45, 46,47,49, 50, 51, 52, 53, 54, 55, 56, 57,58, 60,61, 62, 63, 64, 65, 68, 69,70,71,72,73, 76, 78, 79, 80
Active Power To (Branch)	2,3,4,5,6,9,10,11,12,19,20,21,31
Active Power Generation (Generator)	1,2,3,4,5,6,7,8
Reactive Power From (Branch)	1, 2, 3, 4, 6, 7, 8, 9, 10, 11,12, 15, 16, 17, 18, 19, 20, 24, 25, 26, 27, 30, 32, 33, 34, 35, 36, 37, 39, 40, 41, 42, 43, 44, 45, 46, 47, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 60, 61, 62, 63, 64, 65, 68, 69, 70, 71, 72, 73, 76, 78, 79, 80
Reactive Power To (Branch)	2,3,4,5,6,9,10,11,12,19,20,21,31
Reactive Power Generation (Generator)	1,2,3,4,5,6,7,8
Voltage Magnitude (Bus)	1,2,3,4,6,8,9,12,15,20,24,28,31,32,36,38,41,47,51,53, 57

4.2 Test Scenarios

Section 4.1 provides a comprehensive description of the dataset generation process. The test scenarios follow a similar outline in terms of the dataset generation process, with the only differences being variations in the levels of DER penetration and the noise levels taken into account during measurements. It is important to note that these variables are controlled with utmost care to ensure that the test scenarios are reflective of real-world conditions and generate reasonable results. In practice, the contribution of distributed energy resources (DERs) to the grid varies instantaneously. For instance, the energy produced by photovoltaic panels, which rely on sunlight, can be affected by clouds and weather conditions throughout the day, while wind turbines are exposed to varying wind speeds. This variability results in DERs producing different levels of power throughout the day.

In this study, three different test systems are utilized, namely the IEEE 14 Bus, IEEE 30 Bus, and IEEE 57 Bus, to conduct various test scenarios. These scenarios aimed to showcase the renewable energy production levels that vary throughout the day, with DER penetrations set at four different levels: 0%, 10%, 20%, and 30%. These percentages represent the ratio of the active power generation from DERs to the total active power generation of the other generators present in the test system. For example, in the IEEE 14 Bus system, there are two generators capable of generating active power, with capacities of 232.4 MW and 40 MW, respectively. Consequently, DERs located at specific buses within the dataset were assigned active power generation values of 0 MW, 27.24 MW, 54.48 MW, and 81.72 MW, corresponding to each load sample obtained from the NYISO historical load data. Additionally, to maintain a load-generation balance, all loads in the system were scaled in proportion to the DER generation scale. Consequently, in this study it is aimed to add DERs with variable power generation to the datasets in order to train and utilize ML algorithms to detect false data injection attacks under different scenarios of DER power generation that can occur in practice. In this context, the test matrix created for this purpose is provided in Table 4.5. These scenarios were created to observe the detection ability of the machine learning algorithms against FDIAs across different test systems.

While measuring quantities in practice, there is often a difference between the actual value and the measured value of the quantity being measured. This is because

measurements are subject to noise, which can result from various factors such as environmental factors, instrumentation, signal processing, signal transmission, and signal source. To reflect these noise's effect, in this study White Gaussian Noise (WGN) is added to the measurements which is a noise signal that follows Normal Gaussian Distribution. Signal-to-Noise Ratio (SNR) denotes the ratio between the power of the desired output signal and the background noise [68]. SNR is used to separate the power of a signal from the power of WGN. Thus, higher SNR values indicates that the signal is more clear and consequently more accurate. Lower SNR values mean a higher error between the measured value and the actual value. In this context, it is assumed that the PMUs measuring voltage magnitude and voltage angle have a SNR value of 70 dB. Additionally, it is assumed that the wattmeters measuring active and reactive power have a SNR value of 5 dB. In the process of creation of test scenarios, scenarios with different WGN levels were included in the study to investigate the effects of measurements with different signal-to-noise ratio (SNR) values on the predictions of the machine learning algorithms.

Table 4.5 : Main test scenarios for IEEE test systems.

Test Number	Test System	DER Integration Levels
1	IEEE 14 BUS	0% - 10% - 20% - 30%
2	IEEE 30 BUS	0% - 10% - 20% - 30%
3	IEEE 57 BUS	0% - 10% - 20% - 30%

In this context, created test matrix is given in Table 4.6. These scenarios were created to observe the effect of different noise levels on the FDIA detection ability of the machine learning algorithms. The methodology of the machine learning algorithms used in this thesis is explained in Chapter 2.4. These algorithms include k-Nearest Neighbor (kNN), Logistic Regression (LR), and Deep Neural Network (DNN). The datasets created for each test scenario were divided into 66.6% for training and 33.3% for testing.

The DNN algorithm consists two hidden layers. Structure of the first layer is changes regarding to the size of the input data, dimension of second layer is 4 and the output layer with dimension of 1. This configuration was utilized to build the DNNs and evaluate their performance. The ReLU activation function was used in these hidden layers, while the sigmoid function was utilized in the output layer. The number of

epochs was set to 10. “ADAM” optimizer is used and binary cross entropy is used as loss function. The learning rate is a crucial parameter that affects both the convergence speed and the resulting accuracy of the model. In this thesis, the learning rate is set to 0.2.

Table 4.6 : Test cases of different SNR values for IEEE test systems**.**

Test Case	Test System	SNR Value of PMUs	SNR Value of Wattmeters
Case 1	IEEE 14 BUS	70	2.5
Case 2	IEEE 14 BUS	70	10
Case 3	IEEE 14 BUS	10	5
Case 1	IEEE 30 BUS	70	2.5
Case 2	IEEE 30 BUS	70	10
Case 3	IEEE 30 BUS	10	5
Case 1	IEEE 57 BUS	70	2.5
Case 2	IEEE 57 BUS	70	10
Case 3	IEEE 57 BUS	10	5

4.2.1 Test 1 : IEEE 14 Bus

In this test set, the IEEE 14 Bus system is used to benchmark the performance of Linear k-Nearest Neighbor (k-NN), Regression (LR) and Deep Neural Network (DNN) algorithms as explained in Chapters 3.1 and 3.2. The dataset includes four different levels of DER contribution, and consists of 41 quantities, including 40 measurements and one label indicating the presence or absence of FDIA.

Since the number of input parameters in this system is lower compared to other systems, it is expected that the accuracy of the algorithms will be highest in this system. In the dataset, there are 131,309 observations, bar graph is given in Figure 4.6, out of which 7,521 have been labeled as "1" and the remaining 123,788 have been labeled as "0". This means that 7,521 out of the 131,309 observations are targeted by FDIA.

The accuracies of the k-NN, LR and DNN algorithms are given in the Table 4.7. From the table, it is observed that the results of all algorithms are satisfactory. However, since the data set is large and the number of observations labeled as "0" in the data set is high, examining only the accuracy can lead to misleading results. For instance, the number of observations labeled as "0" in the data set is 123,788, which constitutes 94.27% of the entire data set. Therefore, if an algorithm labels the entire test set as "0" in the predictions made, a very high accuracy is obtained. Therefore, the confusion

matrix, which is matrix used to evaluate the performance of a classification model that compares the predicted and actual values of the test set, can be used. Confusion matrices of the k-NN, LR and DNN algorithms are given in the Figure 4.7, Figure 4.8 and Figure 4.9 respectively.

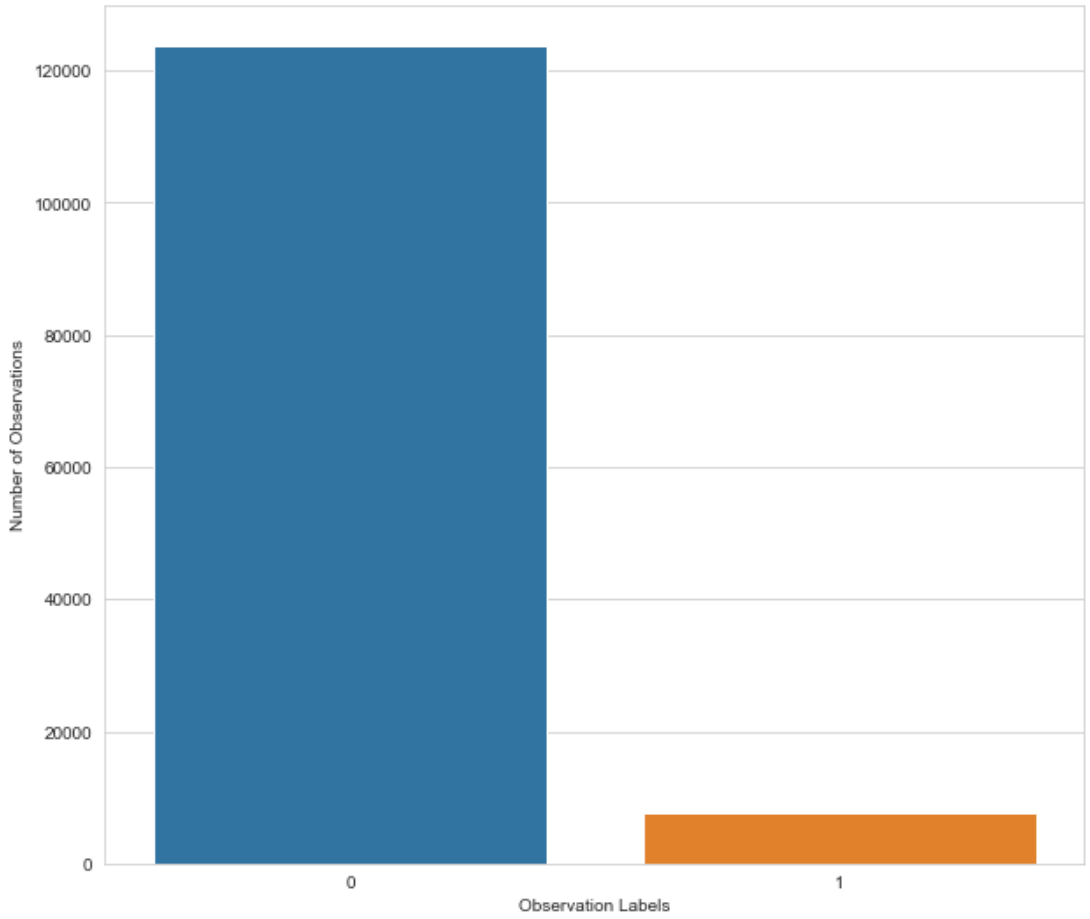


Figure 4.6 : Distribution of observation labels of Test Scenario 1.

Table 4.7 : Accuracies of proposed algorithms.

Algorithm	Accuracy [%]
k-NN	99.30
LR	99.92
DNN	99.99

In k-NN algorithm, the k-value for the k-NN algorithm has been set to 1, which means classification will be performed using the value of the nearest neighbor. The "Minkowski" distance metric was used for geometric distance in the conducted tests. Under these conditions, confusion matrix of k-NN algorithm shows that, the k-NN

algorithm predicted 40798 observations as non-attacked and 2231 observations as attacked which are true predictions. But there are 229 false negatives and 74 false positives. Thus 229 of 2460 attacks are passed the k-NN based FDIA detection algorithm. Which means k-NN based FDIA detection algorithm detected 90.69% of FDIAs successfully. F-1 Score of k-NN algorithm for IEEE 14 Bus system is calculated as 0.9364. In addition to these results, in search of better results, it was observed that when the k-value was set to 1 and the "Euclidean" distance metric was used, the same results were obtained as with the "Minkowski" distance metric. Furthermore, when the k-value was set to 2 and the "Minkowski" distance metric was used, the F-1 score increased to 0.9408, resulting in a 0.47% increase and reducing the number of false positive predictions to 1. However, since the main focus of this study was to detect FDIAs, when the number of false negatives was examined, it was observed to have increased from 229 to 274. Thus, the FDIA detection rate decreased from 90.69% to 88.87%. Therefore, although it had a higher F-1 score, the case where the number of neighbors was set to 1 had higher performance in the scope of this study.

Confusion matrix of LR algorithm shows that, the LR algorithm predicted 40869 observations as non-attacked and 2427 observations as attacked which are true predictions. But there are 33 false negatives and 3 false positives. Thus 33 of 2460 attacks are passed the LR based FDIA detection algorithm. Which means LR based FDIA detection algorithm detected 98.66% of FDIAs successfully. F-1 Score of LR algorithm for IEEE 14 Bus system is calculated as 0.9926.

Confusion matrix of DNN algorithm shows that, the DNN algorithm predicted 40872 observations as non-attacked and 2455 observations as attacked which are true predictions. But there are 5 false negatives and 0 false positives. Thus 5 of 2460 attacks are passed the DNN based FDIA detection algorithm. Which means DNN based FDIA detection algorithm detected 99.8% of FDIAs successfully. F-1 Score of DNN algorithm for IEEE 14 Bus system is calculated as 0.9984. Since the results are satisfactory, no other activation functions, different epoch numbers or different hidden layer structures are tested.

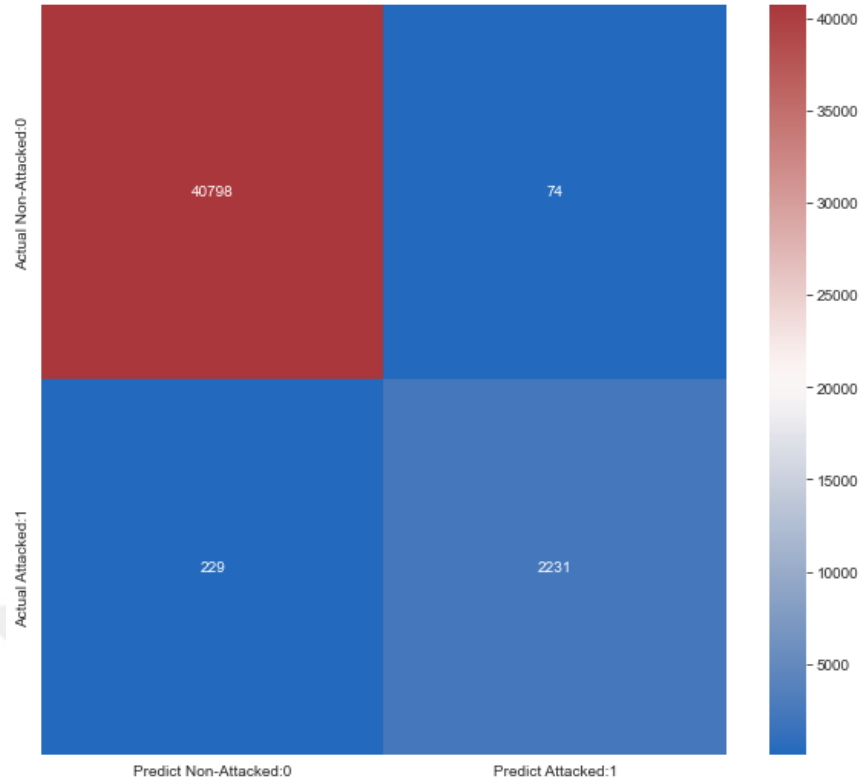


Figure 4.7 : Confusion Matrix of k-NN algorithm for Test Scenario 1, IEEE 14 BUS.

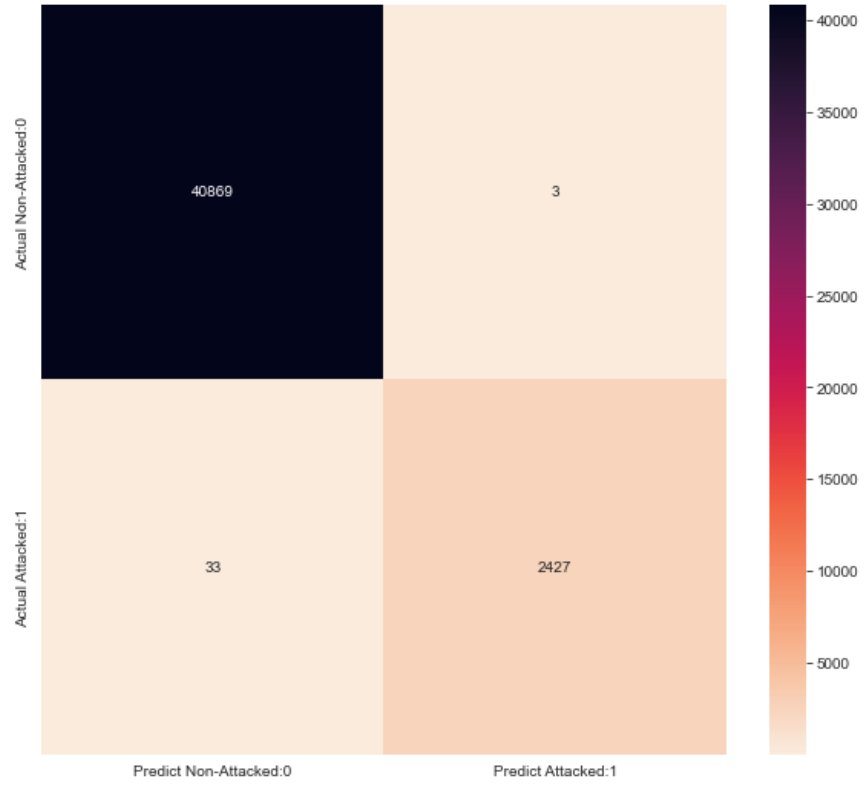


Figure 4.8 : Confusion Matrix of LR algorithm for Test Scenario 1, IEEE 14 BUS.

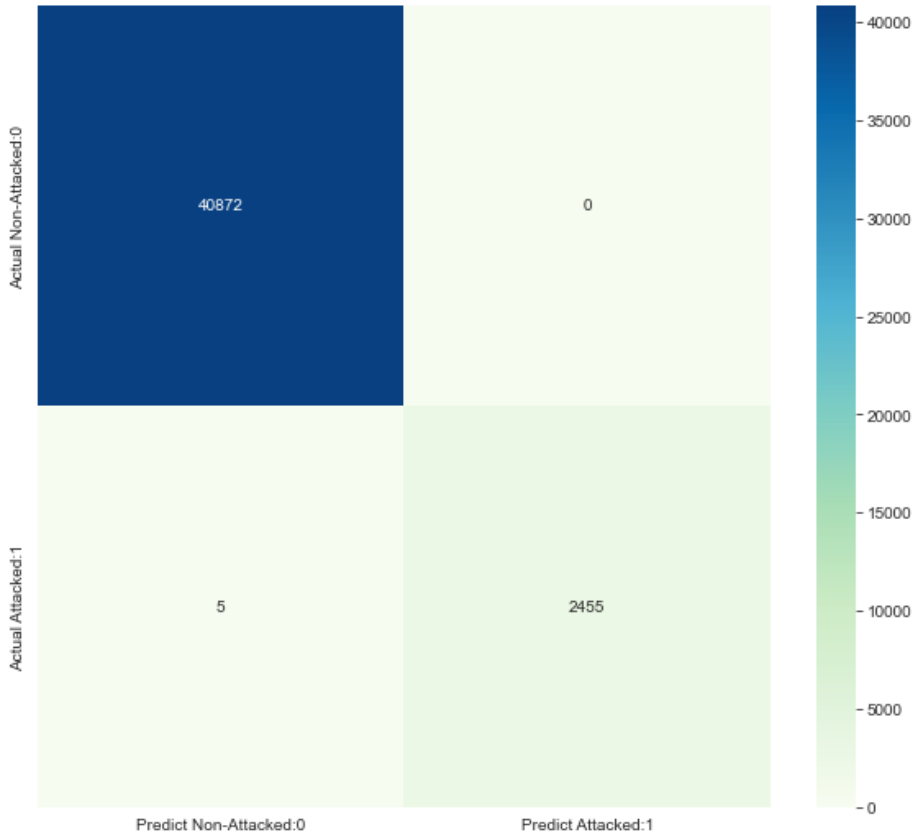


Figure 4.9 : Confusion Matrix of DNN algorithm for Test Scenario 1, IEEE 14 BUS.

All of the results of proposed algorithms are summarized in Table 4.8 and graphical representations are given in Figure 4.10, Figure 4.11 and Figure 4.12.

Table 4.8 : Performances of proposed algorithms for Test System 1, IEEE 14 Bus.

Algorithm	Accuracy [%]	Correct Predictions	False Negatives	False Positives	Attack Detection [%]	F-1 Score
k-NN	99.3	43029	229	74	90.69	0.9364
LR	99.92	43296	33	3	99.92	0.9926
DNN	99.99	43327	5	0	99.99	0.9984

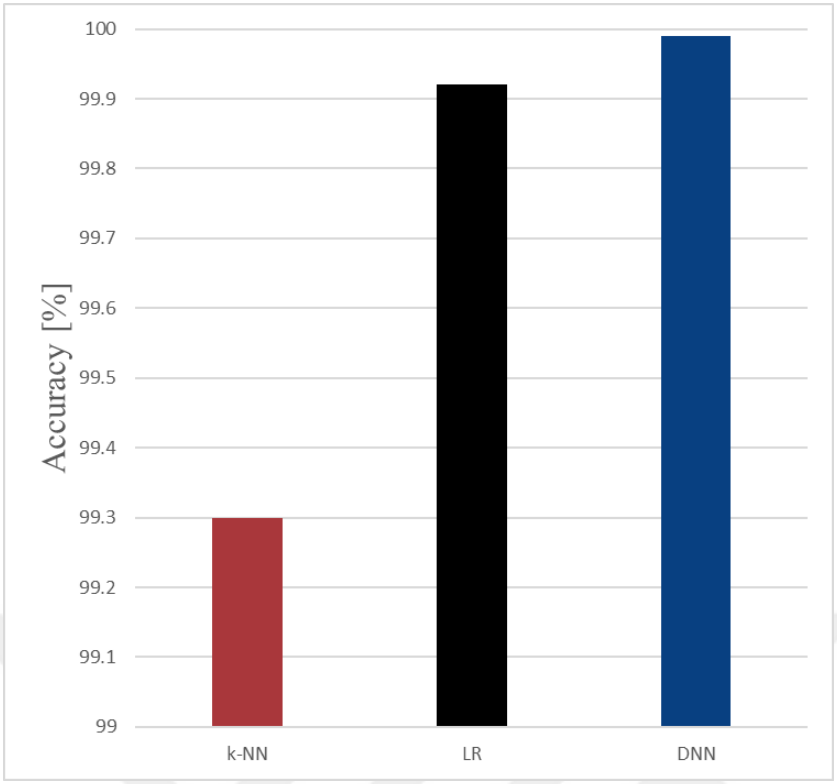


Figure 4.10 : Attack detection rates of proposed algorithms for Test Scenario 1, IEEE 14 BUS.

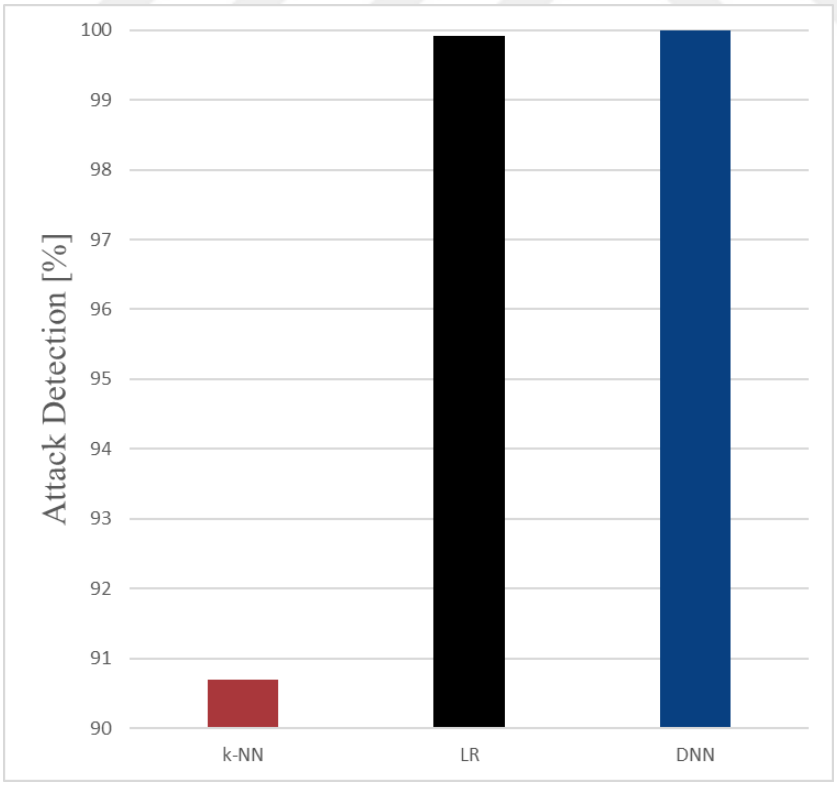


Figure 4.11 : Attack detection rates of proposed algorithms for Test Scenario 1, IEEE 14 BUS.

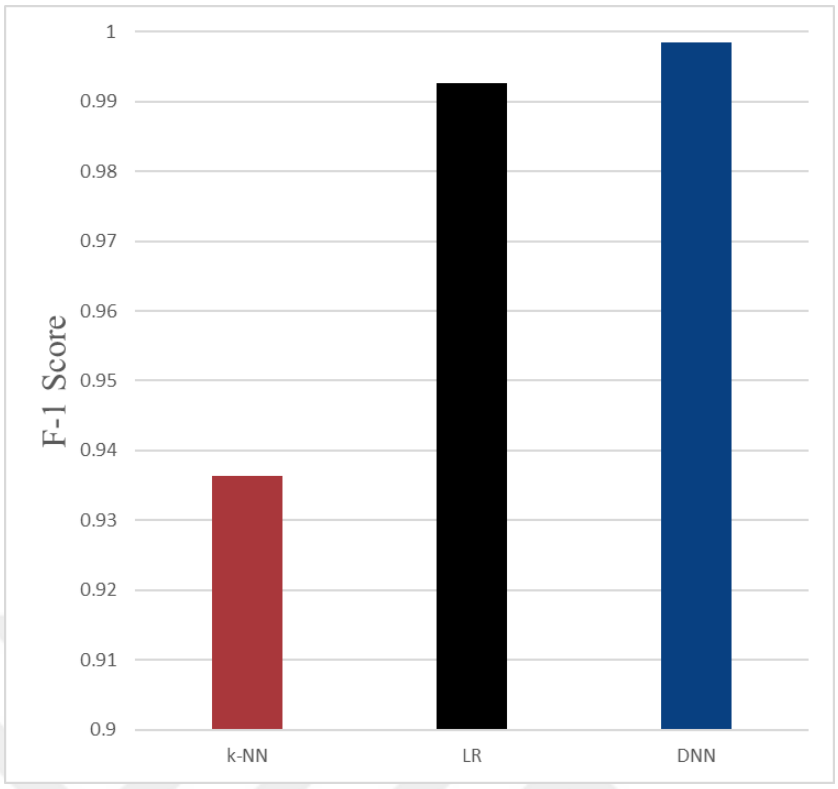


Figure 4.12 : F-1 scores of proposed algorithms for Test Scenario 1, IEEE 14 BUS.

In practice, since power grids and ICT infrastructures are dynamic systems, the noise levels within the system can change instantaneously. Therefore, comparing the performance of ML algorithms under a single noise level would not yield accurate results. The performance of ML algorithms has been evaluated under different noise levels by varying the PMUs and wattmeters SNR levels, according to Table 4.6. Results of the different SNR scenario cases are given in Table 4.9, Table 4.10 and Table 4.11. These tables contain the accuracies, correct prediction rate, false negative and positive rates, attack detection rate and F-1 scores of algorithms under different circumstances.

In Test Case 1, SNR value of the PMUs is not changed and wattmeters SNR value is decreased to 2.5 dB from 5 dB. Thus, there will be more noise in power-related measurements, and it is expected that will lead to a decrease in the performance of the proposed algorithms. When the accuracies are inspected, it is observed that the k-NN algorithm has the biggest accuracy loss, with a decrease of 0.22% compared to the main case. Additionally, it is observed that the DNN has the highest accuracy among the proposed algorithms, with an accuracy of 99.97%. It is observed that the k-NN and LR algorithms has the greatest detection loss, with an 88.56% and 97.19% detection

rate compared to the main case and DNN had the highest FDIA detection rate among the proposed algorithms, with an accuracy of 99.65%. When the F-1 scores were inspected, it was observed that the k-NN algorithm has the lowest F-1 score with 0.9191. Additionally, it is observed that the LR and DNN have the high F-1 scores with a F-1 score of 0.9842 and 0.9973, respectively.

Table 4.9 : Performances of proposed algorithms for Test Case 1 for Test System 1, IEEE 14 Bus.

Algorithm	Accuracy [%]	Correct Predictions [%]	False Negatives [%]	False Positives [%]	Attack Detection [%]	F-1 Score
k-NN	99.08	99.08	0.67	0.24	88.56	0.9191
LR	99.82	99.82	0.17	0.02	97.19	0.9842
DNN	99.97	99.97	0.02	0.01	99.65	0.9973

In Test Case 2, SNR value of the PMUs is not changed and wattmeters SNR value is increased to 10 dB from 5 dB. Thus, there will be less noise in power-related measurements, and it is expected that will lead to an increase in the performance of the proposed algorithms. It is observed that the accuracies of all proposed algorithms are increased compared to base case, as expected. Also, it is observed that the LR and DNN algorithms have performed great detection ability with an 99.29% and 99.65%, respectively, detection rate compared to the main case. When the F-1 scores were inspected, it was observed that the k-NN algorithm has the lowest F-1 score with 0.9671. Additionally, it is observed that the LR and DNN have the high F-1 scores with 0.9964 and 0.998, respectively.

Table 4.10 : Performances of proposed algorithms for Test Case 2 for Test System 1, IEEE 14 Bus.

Algorithm	Accuracy [%]	Correct Predictions [%]	False Negatives [%]	False Positives [%]	Attack Detection [%]	F-1 Score
k-NN	99.62	99.62	0.33	0.04	94.33	0.9671
LR	99.96	99.96	0.04	0.00	99.29	0.9964
DNN	99.97	99.98	0.02	0.00	99.65	0.9980

In Test Case 3, SNR value of the PMUs is decreased to 20 from 70 and wattmeters SNR value is not changed. Since the voltage magnitudes in system are measured by per units, it is expected that this case will have the huge negative impact on FDIA

detection rates and consequently F-1 scores. There will be more noise in voltage measurements, and it is expected that will lead to a decrease in the performance of the proposed algorithms. When the accuracies are inspected, it is noted that the LR algorithm has the biggest accuracy loss, with a decrease of 3.29% compared to the main case. Additionally, it is can be seen that the DNN has the highest accuracy among the proposed algorithms, with an accuracy of 99.93%. It is clear that the DNN algorithm has performed greatest performance, with an 98.67% FDIA detection rate compared to the main case and k-NN and LR algorithms performed 78.14% and 42.67% FDIA detection rate, respectively. When the F-1 scores were inspected, it was observed that the LR algorithm has the lowest F-1 score with 0.596. Additionally, it is observed that the DNN has the high F-1 scores with a score of 0.9927.

Table 4.11 : Performances of proposed algorithms for Test Case 3 for Test System 1, IEEE 14 Bus.

Algorithm	Accuracy [%]	Correct Predictions [%]	False Negatives [%]	False Positives [%]	Attack Detection [%]	F-1 Score
k-NN	98.18	98.18	1.29	0.53	78.13	0.8340
LR	96.63	96.63	3.37	0.00	42.67	0.5960
DNN	99.93	99.91	0.08	0.01	98.67	0.9927

Results of all three cases and main case are summarized in Figure 4.13. It is clearly seen from the results that while LR algorithm has better FDIA detection rate with lower noise levels, k-NN algorithm is outperformed the LR algorithm on high noise levels. However, DNN algorithm has the highest FDIA detection rate at every noise level.

According to the results of the test scenarios, it can be seen that the LR and DNN algorithms have much higher ability to detect FDIA compared to the k-NN algorithm. Nevertheless, due to the relatively smaller size of the test system compared to other systems, the number of variables in the datasets generated within this particular system is limited. Hence, it would not be appropriate conclude inference based solely on the analysis of a single system, and all test systems should be examined.

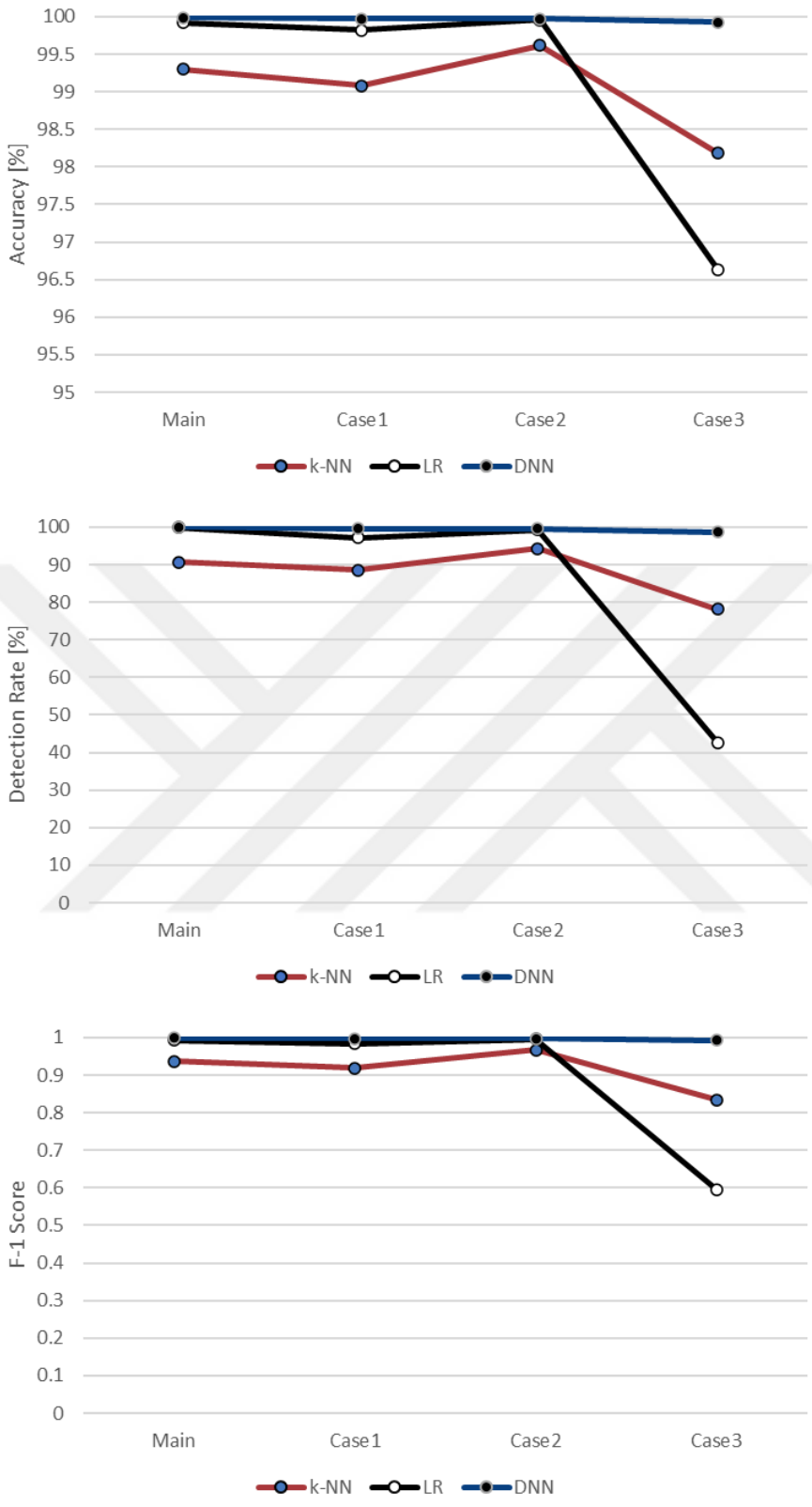


Figure 4.13 : (a) Accuracies, (b) attack detection rates and (c) F1 scores of proposed algorithms for all scenarios, IEEE 14 BUS.

4.2.2 Test 2 : IEEE 30 Bus

In this test scenario, the IEEE 30 Bus system is used to benchmark the performance of machine learning algorithms on more complex system than Test 1. The dataset includes four different levels of DER contribution, and consists of 96 quantities, including 95 measurements and one label indicating the presence or absence of FDIA.

This dataset contains 131,441 observations, bar graph is given in Figure 4.14, out of which 7,653 have been labeled as "1" and the remaining 123,788 have been labeled as "0". This means that 7,653 out of the 131,441 observations are targeted by FDIA.

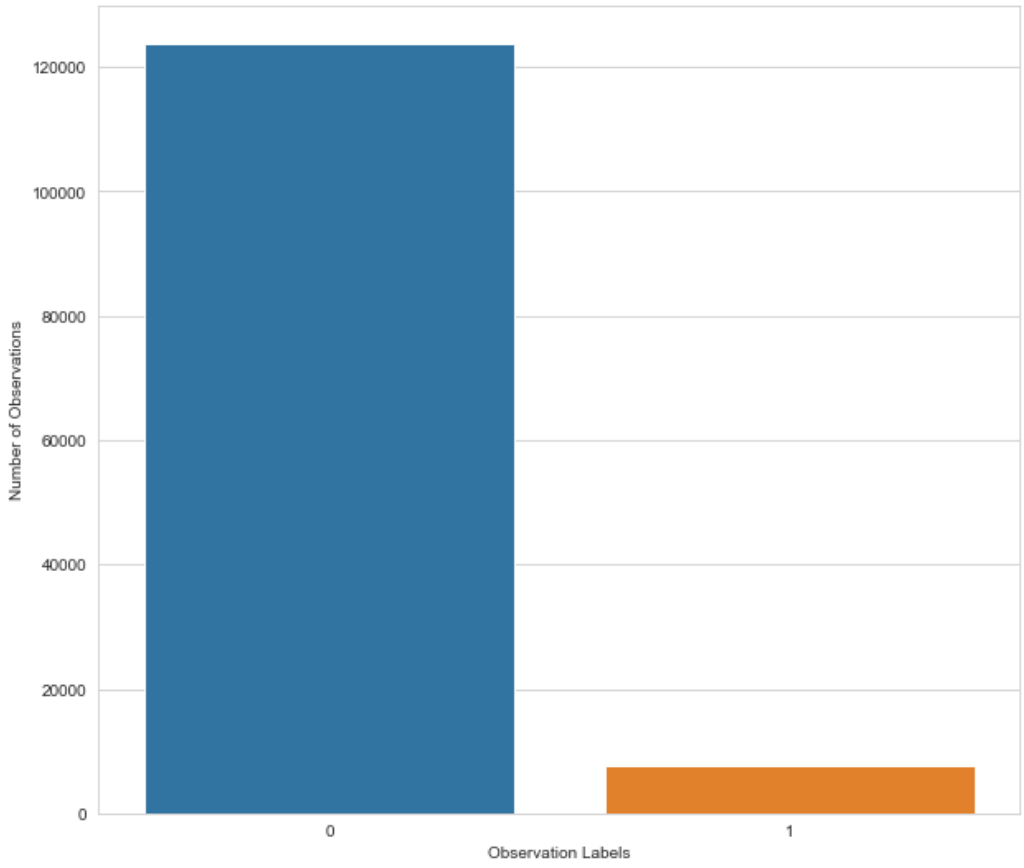


Figure 4.14 : Distribution of observation labels of Test Scenario 2.

Confusion matrices of the k-NN, LR and DNN algorithms are given in Figure 4.15, Figure 4.16 and Figure 4.17, respectively. The setting parameters of ML algorithms are the same as Test Scenario 1 for all algorithms. For k-NN, the value of k is chosen as 1 and "Minkowski" distance metric is used, while for DNN, 2 hidden layers with ReLU activation function and an output layer with sigmoid activation function are used with 10 epoches.

Confusion matrix of k-NN algorithm indicates that, the k-NN algorithm predicted 40769 observations as non-attacked and 2436 observations as attacked which are true predictions. But there are 112 false negatives and 59 false positives. Thus 112 of 2548 attacks are passed the k-NN based FDIA detection algorithm. Which means k-NN based FDIA detection algorithm detected 95.60% of FDIAs successfully. F-1 Score of k-NN algorithm for IEEE 30 Bus system is calculated as 0.9661. Compared to the IEEE 14 Bus system, it is observed that the FDIA detection rate of the k-NN at IEEE 30 Bus system performed an increase of 4.91%. This can be achieved by the ability of the k-NN algorithm to perform data extraction and data classification with a larger number of inputs. Additionally, when using large datasets, the overfitting issue observed in small datasets can be overcome.

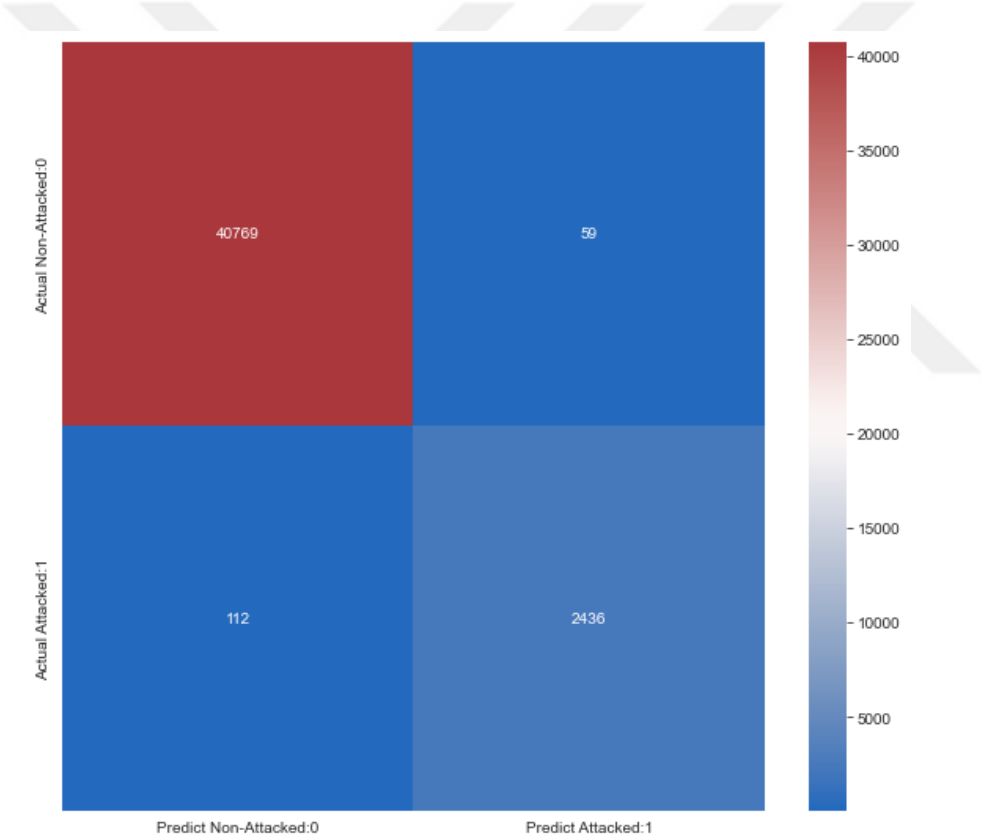


Figure 4.15 : Confusion Matrix of k-NN algorithm for Test Scenario 2, IEEE 30 BUS.

Confusion matrix of LR algorithm shows that, the LR algorithm predicted 40808 observations as non-attacked and 972 observations as attacked which are true predictions. But there are 1576 false negatives and 20 false positives. Which means LR based FDIA detection algorithm detected 38.15% of FDIAs successfully. F-1 Score of LR algorithm for IEEE 14 Bus system is calculated as 0.5492.

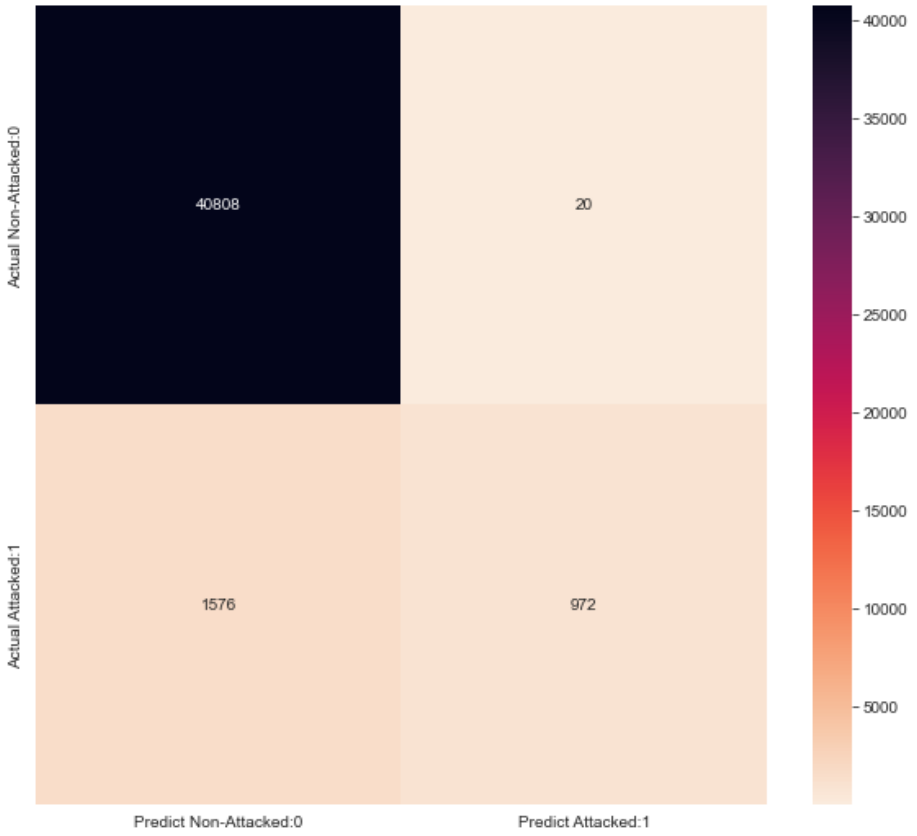


Figure 4.16 : Confusion Matrix of LR algorithm for Test Scenario 2, IEEE 30 BUS.

Compared to the IEEE 14 Bus system, it is observed that the FDIA detection rate of the LR at IEEE 30 Bus system is substantially lowered by 61.65%.

Confusion matrix of DNN algorithm indicates that, the DNN algorithm predicted 40821observations as non-attacked and 2518 observations as attacked which are true predictions. But there are 30 false negatives and 7 false positives. Thus 30 of 2548 attacks are passed the DNN based FDIA detection algorithm. Which means DNN based FDIA detection algorithm detected 98.82% of FDIAs successfully. F-1 Score of DNN algorithm for IEEE 30 Bus system is calculated as 0.9927.

As the results have been deemed satisfactory, no further experimentation was conducted to test different activation functions, epoch numbers, or hidden layer structures.

All of the results of proposed algorithms are summarized in Table 4.12 and graphical representations are given in Figure 4.18, Figure 4.19 and Figure 4.20, respectively.

The performance of ML algorithms has been evaluated under different noise levels by varying the PMUs and wattmeters SNR levels, according to Table 4.6. Results of the

different SNR scenario cases are given in Table 4.13, Table 4.14 and Table 4.15. These tables contain the accuracies, correct prediction rate, false negative and positive rates, attack detection rate and F-1 scores of algorithms under different circumstances.

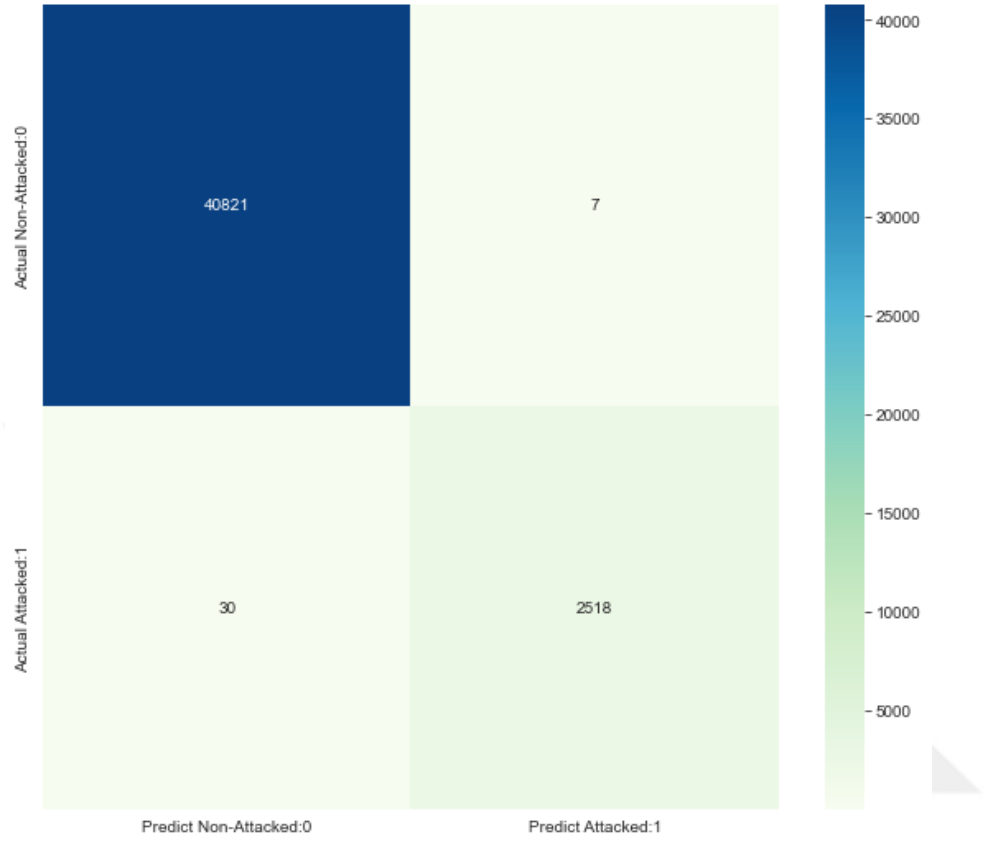


Figure 4.17 : Confusion Matrix of DNN algorithm for Test Scenario 2, IEEE 30 BUS.

Table 4.12 : Performances of proposed algorithms for Test System 2, IEEE 30 Bus.

Algorithm	Accuracy [%]	Correct Predictions [%]	False Negatives [%]	False Positives [%]	Attack Detection [%]	F-1 Score
k-NN	99.61	99.6058	0.25821	0.13602	95.6044	0.9661
LR	96.32	96.3205	3.63335	0.04611	38.1476	0.5492
DNN	99.96	99.9147	0.06916	0.01614	98.8226	0.9927

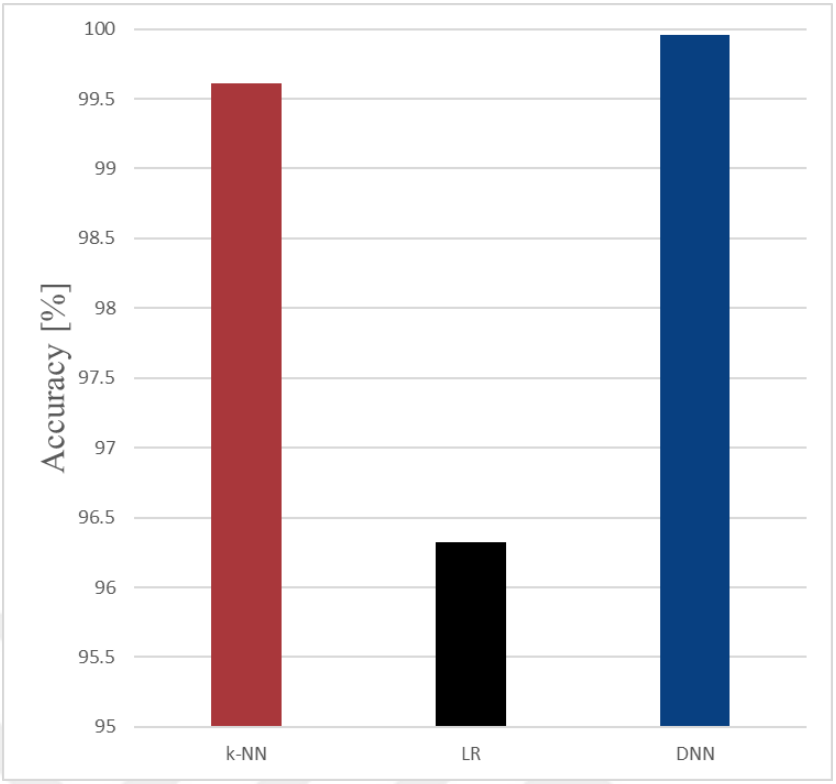


Figure 4.18 : Accuracies of proposed algorithms for Test Scenario 2, IEEE 30 BUS.

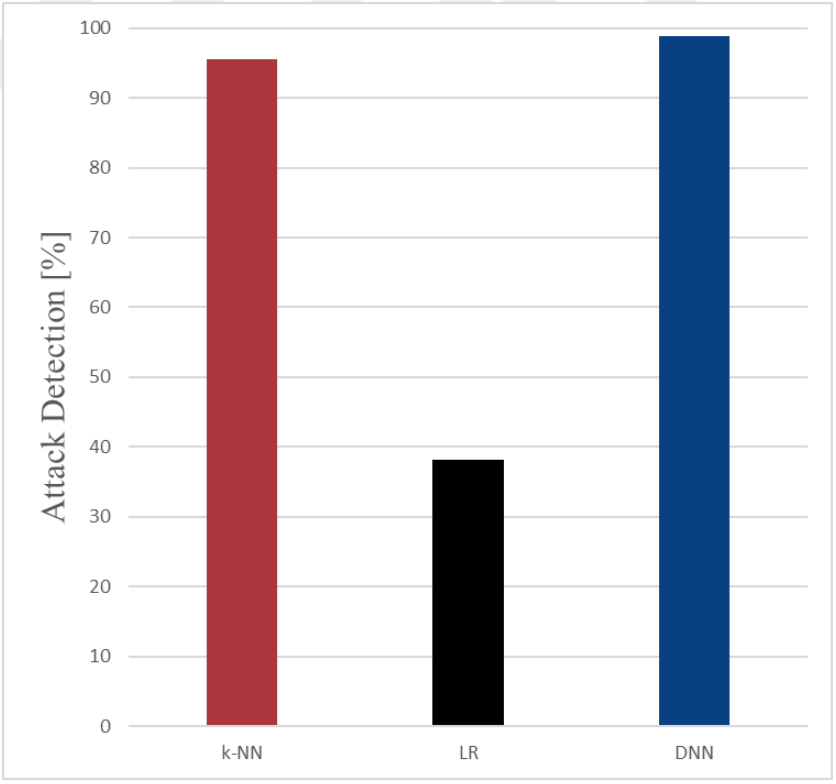


Figure 4.19 : Attack detection rates of proposed algorithms for Test Scenario 2, IEEE 30 BUS.

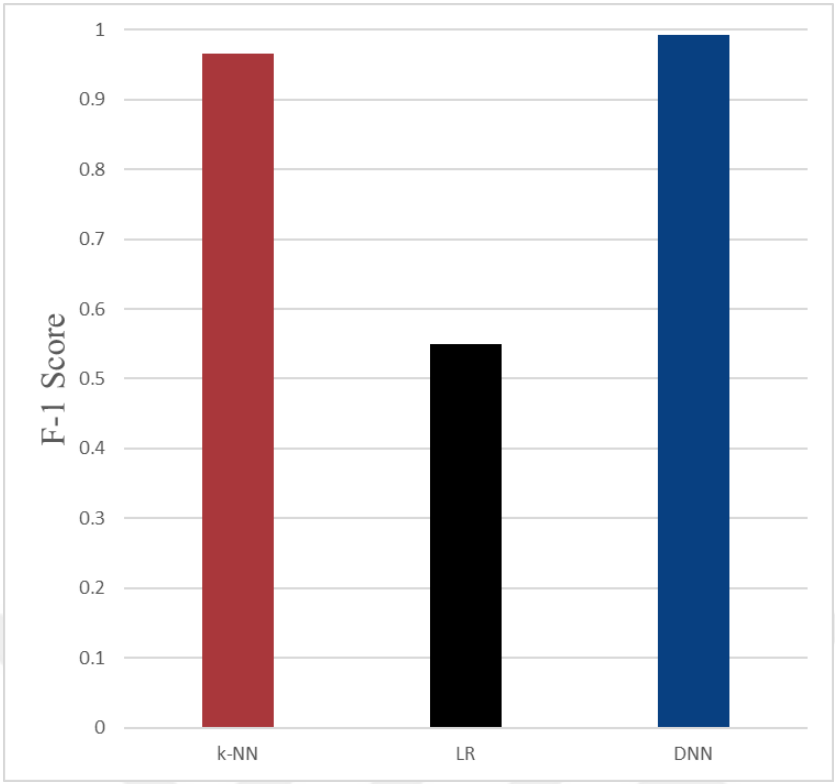


Figure 4.20 : F-1 scores of proposed algorithms for Test Scenario 2, IEEE 30 BUS.

In Test Case 1, SNR value of the PMUs is not changed and wattmeters SNR value is decreased to 2.5 dB from 5 dB. Thus, there will be more noise in power-related measurements, and it is expected that will lead to a decrease in the performance of the proposed algorithms. When the accuracies are inspected, it is observed that the k-NN and LR algorithms have almost same amount of accuracy loss of 0.27% and 0.32%. compared to the main case.

Additionally, it is observed that the DNN has the highest accuracy among the proposed algorithms, with an accuracy of 99.96%. It can be seen that the LR algorithm has the greatest detection loss, with an 32.50%, which was already lower compared to other algorithms, and DNN had the highest FDIA detection rate among the proposed algorithms, with an accuracy of 98.23%.

When the F-1 scores were inspected, it was observed that the LR algorithm has the lowest F-1 score with 0.4882. Additionally, it is observed that the k-NN and DNN have the high F-1 scores with a F-1 score of 0.9426 and 0.9884, respectively.

In Test Case 2, SNR value of the PMUs is not changed and wattmeters SNR value is increased to 10 dB from 5 dB. Thus, there will be less noise in power-related measurements, and it is expected that will lead to an increase in the performance of

the proposed algorithms. It is observed that the accuracies of all proposed algorithms are increased compared to base case, as expected. It is apparent that the DNN algorithm has highest FDIA detection rate among all algorithms with 98.23% detection rate.

Table 4.13 : Performances of proposed algorithms for Test Case 1 for Test System 2, IEEE 30 Bus.

Algorithm	Accuracy [%]	Correct Predictions [%]	False Negatives [%]	False Positives [%]	Attack Detection [%]	F-1 Score
k-NN	99.34	99.34065	0.461084	0.198266	92.15071	0.9426
LR	96	95.99779	3.965326	0.036887	32.49608	0.4882
DNN	99.96	99.86398	0.103744	0.032276	98.23391	0.9884

Table 4.14 : Performances of proposed algorithms for Test Case 2 for Test System 2, IEEE 30 Bus.

Algorithm	Accuracy [%]	Correct Predictions [%]	False Negatives [%]	False Positives [%]	Attack Detection [%]	F-1 Score
k-NN	99.88	99.88473	0.103744	0.011527	98.23391	0.9901
LR	96.82	96.82082	3.13768	0.041498	46.58556	0.6326
DNN	99.99	99.97925	0.018443	0.002305	99.68603	0.9982

In Test Case 3, SNR value of the PMUs is decreased to 20 from 70 and wattmeters SNR value is not changed. Since the voltage magnitudes in system are measured by per units, it is expected that this case will have the huge negative impact on FDIA detection rates and consequently F-1 scores. There will be more noise in voltage measurements, and it is expected that will lead to a decrease in the performance of the proposed algorithms. When the accuracies are inspected, it is noted that the DNN algorithm has the biggest accuracy loss, with a decrease of 0.26% compared to the main case. However, the DNN still has the highest FDIA detection rate with a 94.66% FDIA detection rate compared to the main case and k-NN and LR algorithms performed 93.29% and 35.71% FDIA detection rate, respectively. When the F-1 scores were inspected, it was observed that the LR algorithm has the lowest F-1 score with 0.523.

Results of all three cases and main case are summarized in Figure 4.13. It is clearly seen from the results that DNN algorithm has the highest FDIA detection rate at every noise level.

Table 4.15 : Performances of proposed algorithms for Test Case 3 for Test System 2, IEEE 30 Bus.

Algorithm	Accuracy [%]	Correct Predictions [%]	False Negatives [%]	False Positives [%]	Attack Detection [%]	F-1 Score
k-NN	99.44	99.43517	0.394227	0.170601	93.28885	0.951
LR	96.17	96.173	3.776282	0.050719	35.71429	0.523
DNN	99.7	99.65419	0.313537	0.032276	94.66248	0.9698

According to the results of the test scenarios, it can be seen that the k-NN and DNN algorithms have much higher ability to detect FDIA compared to the LR algorithm on more complex systems. It is noteworthy that the k-NN algorithm had a lower detection rate on the IEEE 14 Bus system compared to other algorithms, with specific numbers indicating a detection rate of varying 78.13% - 94.33%. However, on the more complex IEEE 30 Bus system with a larger number of inputs, the k-NN algorithm outperformed the LR algorithms in FDIA detection, although the DNN algorithm achieved the highest detection rate overall.

4.2.3 Test 3 : IEEE 57 Bus

In this test set, the IEEE 57 Bus system is used to benchmark the performance of proposed algorithms on more complex system in comparison to the IEEE 14 Bus and IEEE 30 Bus test systems. In IEEE 57 bus system, there are 192 measurements available for state estimation process and one label that indicates the available measurements are attacked or not. Compared to IEEE 14 and IEEE 30 bus systems, which have 40 and 95 measurements. Thus, the IEEE 57 Bus system is more complex, making it more challenging to detect FDIAs leading to lower expected FDIA detection rates compared to other test systems.

Prepared dataset contains 131,441 observations with 7653 and 123,788 are labeled as attacked and non-attacked, respectively. Bar graph of label distribution is given in Figure 4.22.

The accuracies of k-NN, LR, and DNN are calculated to be 97.64%, 96.31%, and 99.42%, respectively. However, as explained in section 3.1.1, comparing the performances of the algorithms based solely on their accuracies can be misleading, since the primary aim is to detect FDIAs. Confusion matrices of the algorithms are given in, Figure 4.25 and Figure 4.26.

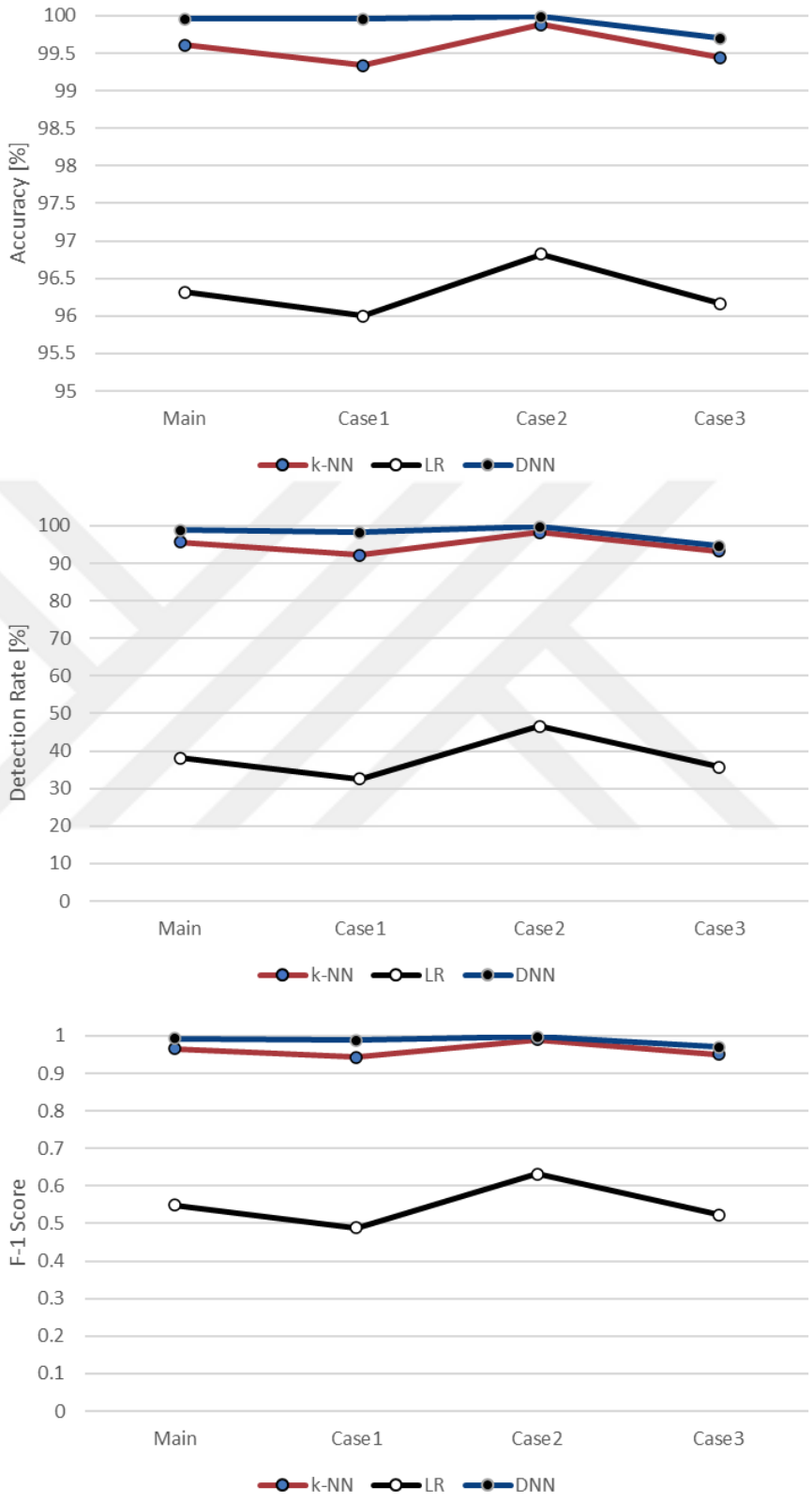


Figure 4.21 : (a) Accuracies, (b) attack detection rates and (c) F1 scores of proposed algorithms for all scenarios, IEEE 30 BUS.

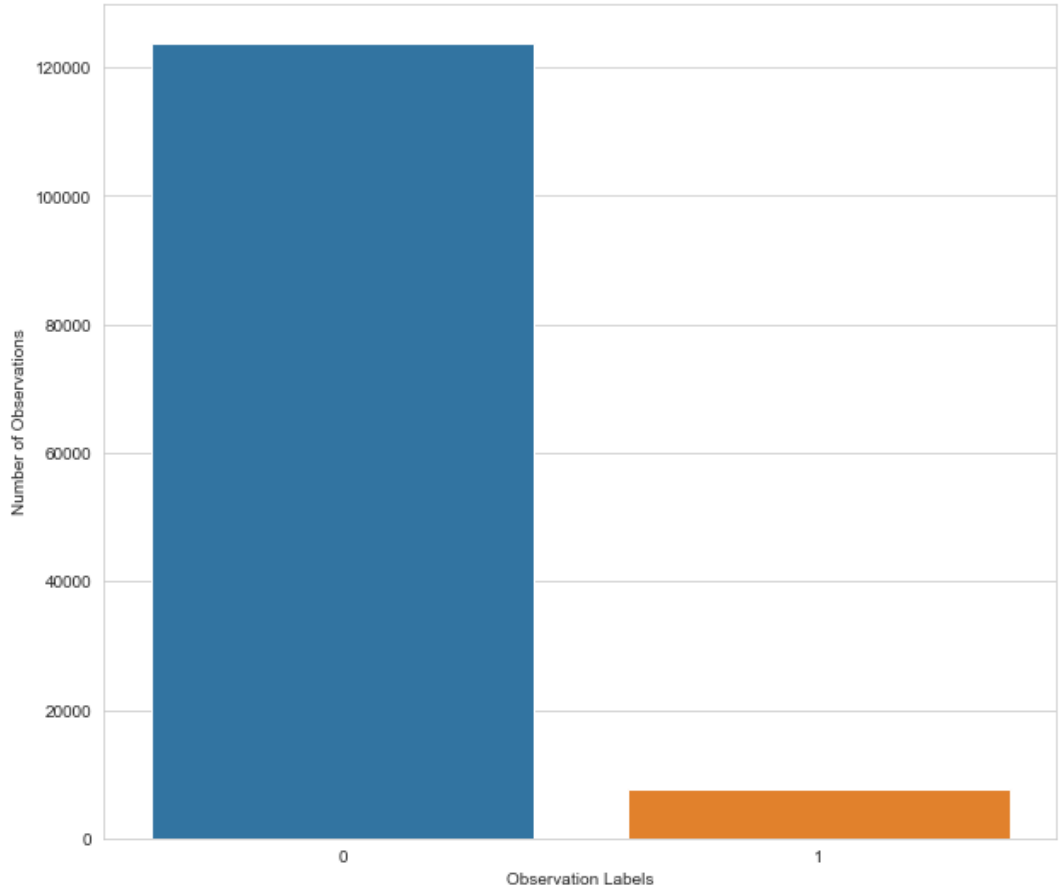


Figure 4.22 : Distribution of observation labels of Test Scenario 3.

Confusion matrix of the k-NN algorithm is given in Figure 4.23 and from the figure it can be clearly seen that numbers of false-negatives are considerably increased compared to “Test Case 2 : IEEE 30 Bus system”. Consequently, the detection rate is decreased from 95.60% to 75.08%. On the other hand, the false positive rate is also increased from 0.14% to 0.90%. F-1 Score of the k-NN algorithm is calculated as 0.789.

Based on the confusion matrix, it is evident that the k-NN algorithm's performance has significantly decreased. Same with other test scenarios, the value of k was set to 1 for the k-NN algorithm. However, given the observed decrease in detection rate, adjusting the k value can increase the algorithm's detection rate. Results for the adjusting k value is given in Figure 4.24. According to results, changing k value effects the detection rate of FDIAs and best detection rate is obtained when k is set to 3 by 76.18%.

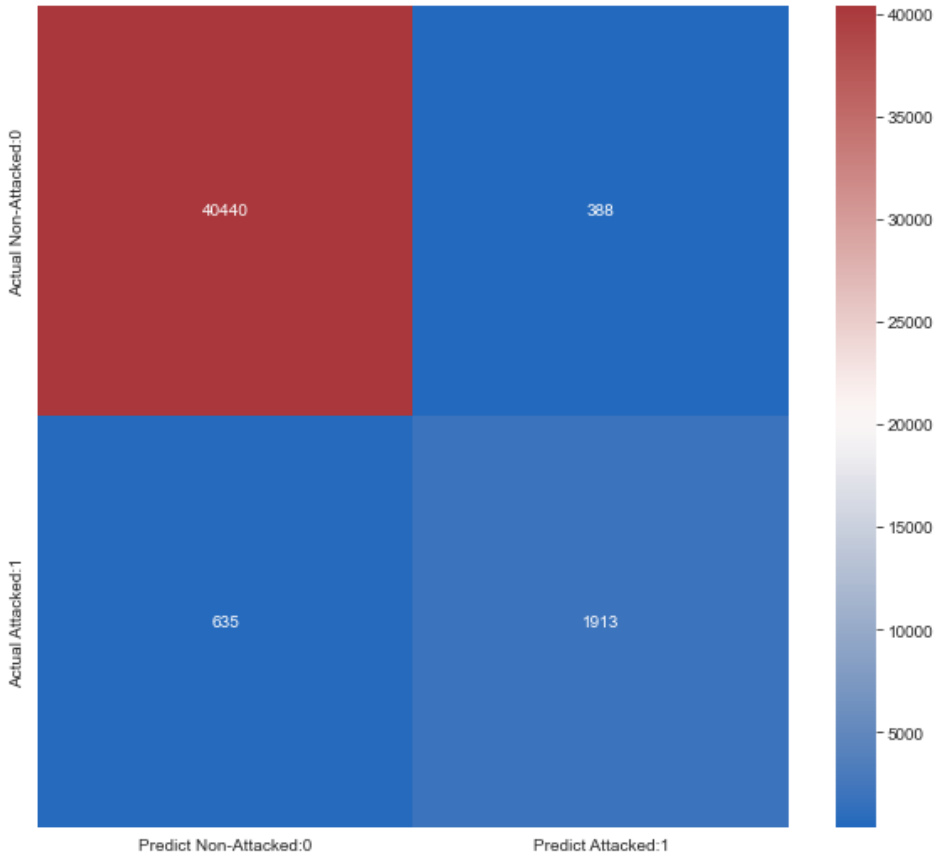


Figure 4.23 : Confusion Matrix of k-NN algorithm for Test Scenario 3, IEEE 57 BUS.

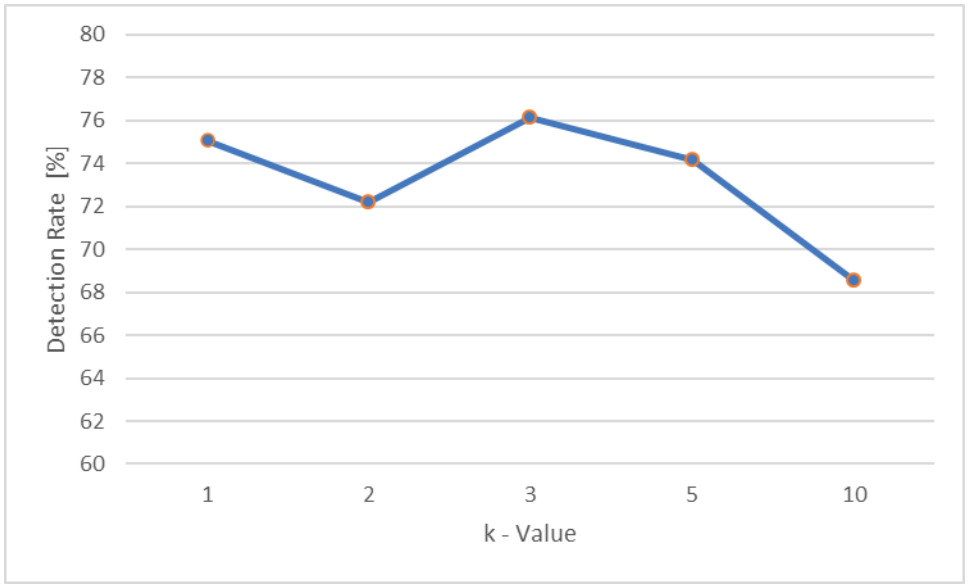


Figure 4.24 : Detection rate of k-NN algorithm for different k values at Test Scenario 3, IEEE 57 BUS.

Confusion matrix of LR algorithm shows that, the LR algorithm predicted 40805 observations as non-attacked and 970 observations as attacked which are true predictions. But there are 1578 false negatives and 23 false positives. Thus 1578 of

2548 attacks are passed the LR based FDIA detection algorithm that means LR based FDIA detection algorithm detected 38.07% of FDIAs successfully, which is almost identical with the “Test Scenario 2 : IEEE 30 Bus”. F-1 Score of LR algorithm for IEEE 14 Bus system is calculated as 0.5479.

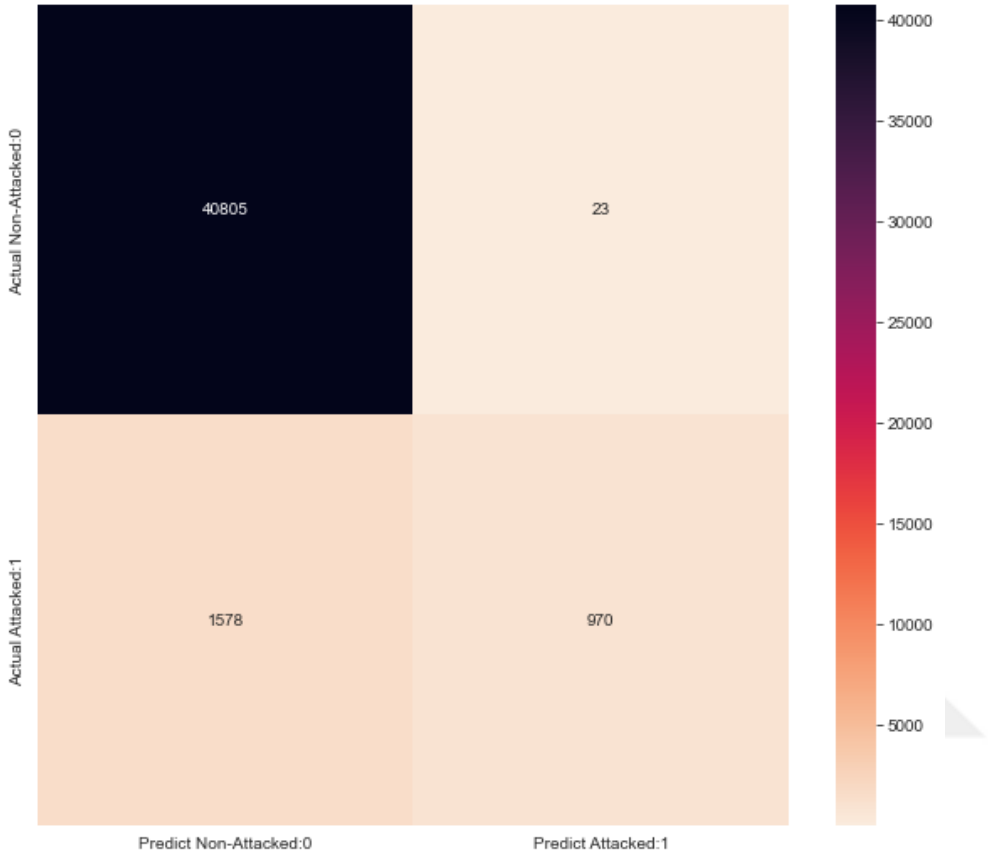


Figure 4.25 : Confusion Matrix of LR algorithm for Test Scenario 3, IEEE 57 BUS.

Confusion matrix of DNN algorithm indicates that, the DNN algorithm predicted 40686 as non-attacked and 2369 observations as attacked which are true predictions. But there are 179 false negatives and 142 false positives. Thus 179 of 2548 attacks are passed the DNN based FDIA detection algorithm. Which means DNN based FDIA detection algorithm detected 92.98% of FDIAs successfully. F-1 Score of DNN algorithm for IEEE 14 Bus system is calculated as 0.9365.

All of the results of proposed algorithms are summarized in Table 4.12 and graphical representations are given in Figure 4.27, Figure 4.28 and Figure 4.29, respectively. The performance of ML algorithms has been evaluated under different noise levels by varying the PMUs and wattmeters SNR levels, according to Table 4.6.

Results of the different SNR scenario cases are given in Table 4.17, Table 4.18 and Table 4.19. These tables contain the accuracies, correct prediction rate, false negative and positive rates, attack detection rate and F-1 scores of algorithms under different circumstances.

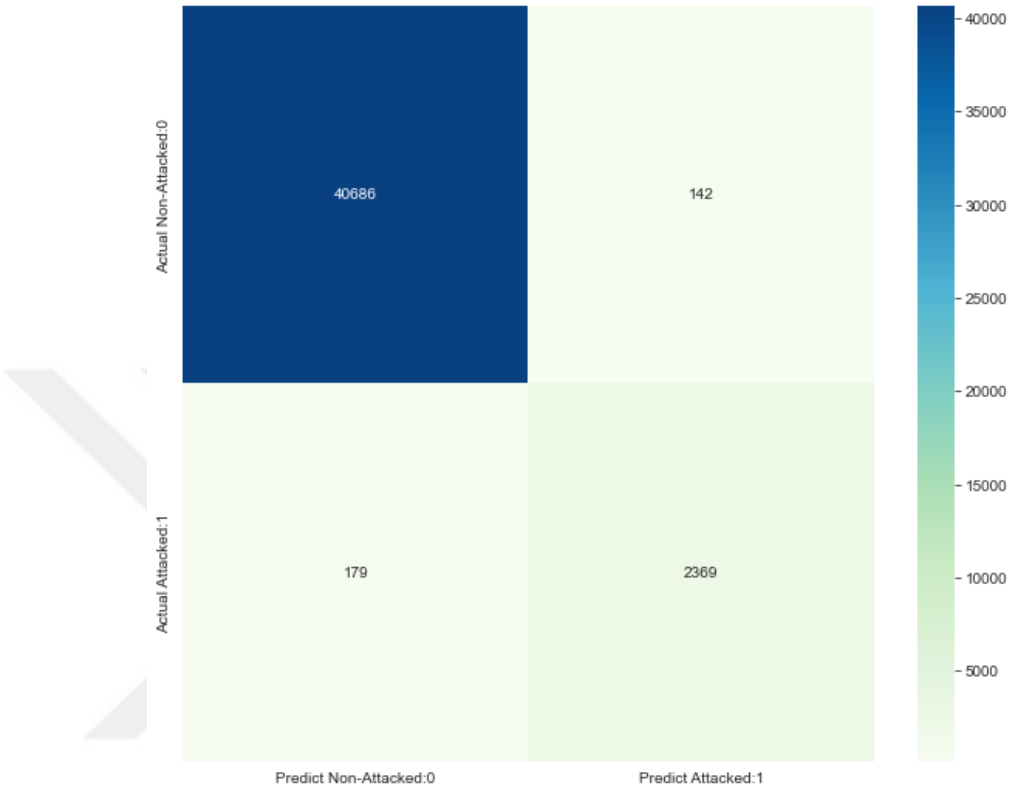


Figure 4.26 : Confusion Matrix of DNN algorithm for Test Scenario 3, IEEE 57 BUS.

Table 4.16 : Performances of proposed algorithms for Test System 3, IEEE 57 Bus.

Algorithm	Accuracy [%]	Correct Predictions [%]	False Negatives [%]	False Positives [%]	Attack Detection [%]	F-1 Score
k-NN	97.64	97.6416	1.46394	0.8945	75.0785	0.789
LR	96.31	96.309	3.63796	0.05302	38.0691	0.5479
DNN	99.42	99.26	0.41267	0.32737	92.9749	0.9365

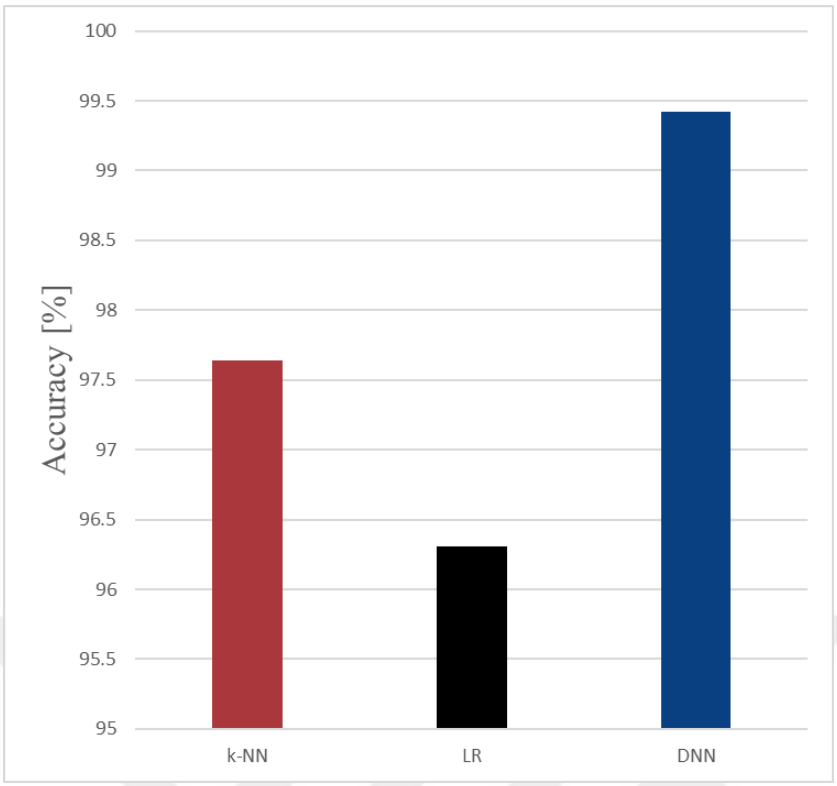


Figure 4.27 : Accuracies of proposed algorithms for Test Scenario 3, IEEE 57 BUS.

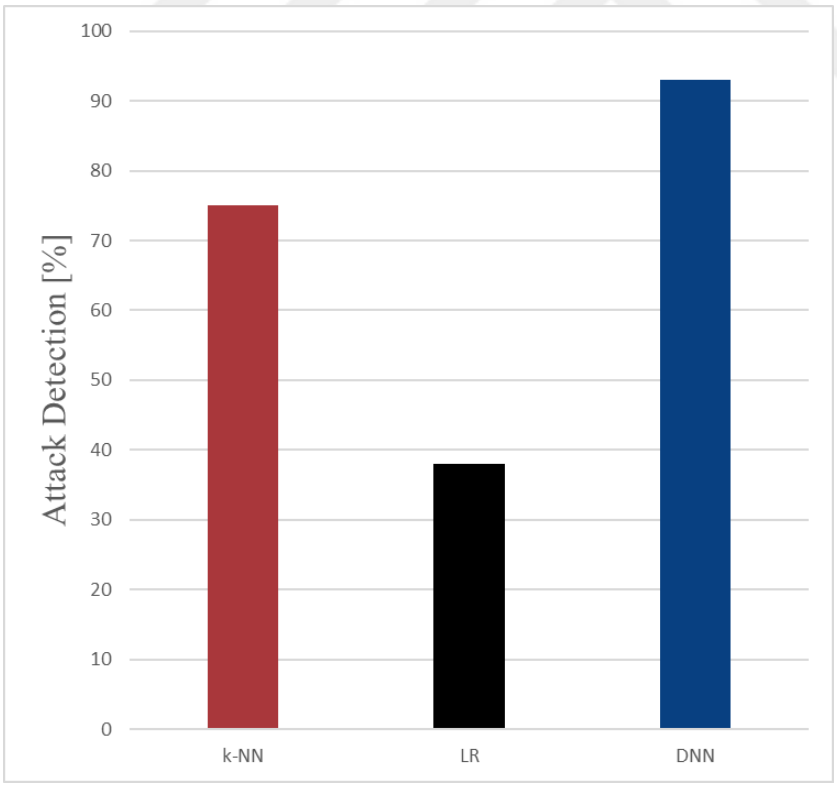


Figure 4.28 : Attack detection rates of proposed algorithms for Test Scenario 3, IEEE 57 BUS.

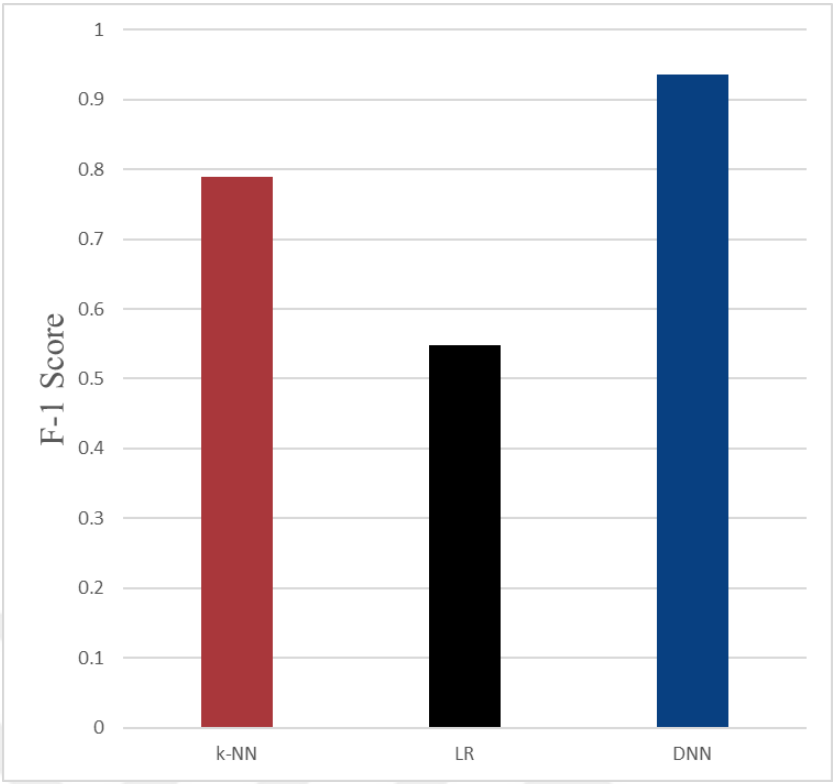


Figure 4.29 : F-1 scores of proposed algorithms for Test Scenario 3, IEEE 57 BUS.

In Test Case 1, SNR value of the PMUs is not changed and wattmeters SNR value is decreased to 2.5 dB from 5 dB. When the accuracies are inspected, it is observed that the LR algorithms has highest accuracy among the proposed algorithms. Additionally, it is observed that the DNN has the highest accuracy loss, with a loss of 5,37%. Despite of accuracy, DNN had the highest FDIA detection rate among the proposed algorithms, with an accuracy of 89.05%. When the F-1 scores were inspected, it was observed that the LR algorithm has the lowest F-1 score with 0.523. Additionally, it is observed that the k-NN and DNN have the high F-1 scores with a F-1 score of 0.692 and 0.9405, respectively.

Table 4.17 : Performances of proposed algorithms for Test Case 1 for Test System 3, IEEE 57 Bus.

Algorithm	Accuracy [%]	Correct Predictions [%]	False Negatives [%]	False Positives [%]	Attack Detection [%]	F-1 Score
k-NN	96.73	96.73091	2.012634	1.256455	64.59854	0.692
LR	96.3	96.29519	3.654094	0.050719	35.72587	0.523
DNN	94.05	99.35909	0.622464	0.018443	89.05109	0.9405

In Test Case 2, SNR value of the PMUs is not changed and wattmeters SNR value is increased to 10 dB from 5 dB. Thus, there will be less noise in power-related measurements, and it is expected that will lead to an increase in the performance of the proposed algorithms. It is observed that the accuracies of all proposed algorithms are increased compared to base case, as expected. It is apparent that the DNN algorithm has highest FDIA detection rate among all algorithms with 97.26% detection rate. In addition, the k-NN algorithm has significantly increased its detection rate to levels of 86.74%. Despite having the lowest noise level in this case, the LR algorithm has a detection rate of only 42.54% and is not recommended to be used in practical system.

Table 4.18 : Performances of proposed algorithms for Test Case 3 for Test System 3, IEEE 57 Bus.

Algorithm	Accuracy [%]	Correct Predictions [%]	False Negatives [%]	False Positives [%]	Attack Detection [%]	F-1 Score
k-NN	98.95	98.94873	0.753873	0.297399	86.73966	0.9037
LR	96.68	96.67558	3.266783	0.057636	42.53852	0.5927
DNN	99.9	99.83401	0.154463	0.011527	97.28305	0.9852

In Test Case 3, SNR value of the PMUs is decreased to 20 from 70 and wattmeters SNR value is not changed. Since the voltage magnitudes in system are measured by per units, it is expected that this case will have the huge negative impact on FDIA detection rates and consequently F-1 scores. There will be more noise in voltage measurements, and it is expected that will lead to a decrease in the performance of the proposed algorithms. Despite the high accuracies, the detection rates have significantly decreased compared to the main scenario. Detection rates of k-NN and LR have decreased to levels of 53.29% and 25.59%, respectively. However, the DNN still has the highest FDIA detection rate with 83.86% FDIA detection rate compared to the k-NN and LR algorithms. When the F-1 scores were inspected, it was observed that the LR algorithm has the lowest F-1 score with 0.4013.

Results of all three cases and main case are summarized in Figure 4.30. According to the results that DNN algorithm has the highest FDIA detection rate at every noise level. Also, it can be seen that the k-NN and DNN algorithms have much higher ability to detect FDIA compared to the LR algorithm on more complex systems. However, on the more complex IEEE 57 Bus system with a larger number of inputs compared to

the IEEE 14 Bus and IEEE 30 Bus systems, the k-NN algorithm outperformed the LR algorithms in FDIA detection, although the DNN algorithm achieved the highest detection rate overall.

Table 4.19 : Performances of proposed algorithms for Test Case 3 for Test System 3, IEEE 57 Bus.

Algorithm	Accuracy [%]	Correct Predictions [%]	False Negatives [%]	False Positives [%]	Attack Detection [%]	F-1 Score
k-NN	95.59	95.58512	2.655847	1.759037	53.28467	0.5785
LR	95.66	95.65889	4.23045	0.11066	25.588	0.4013
DNN	99.3	99.04786	0.917558	0.034581	83.8605	0.9092

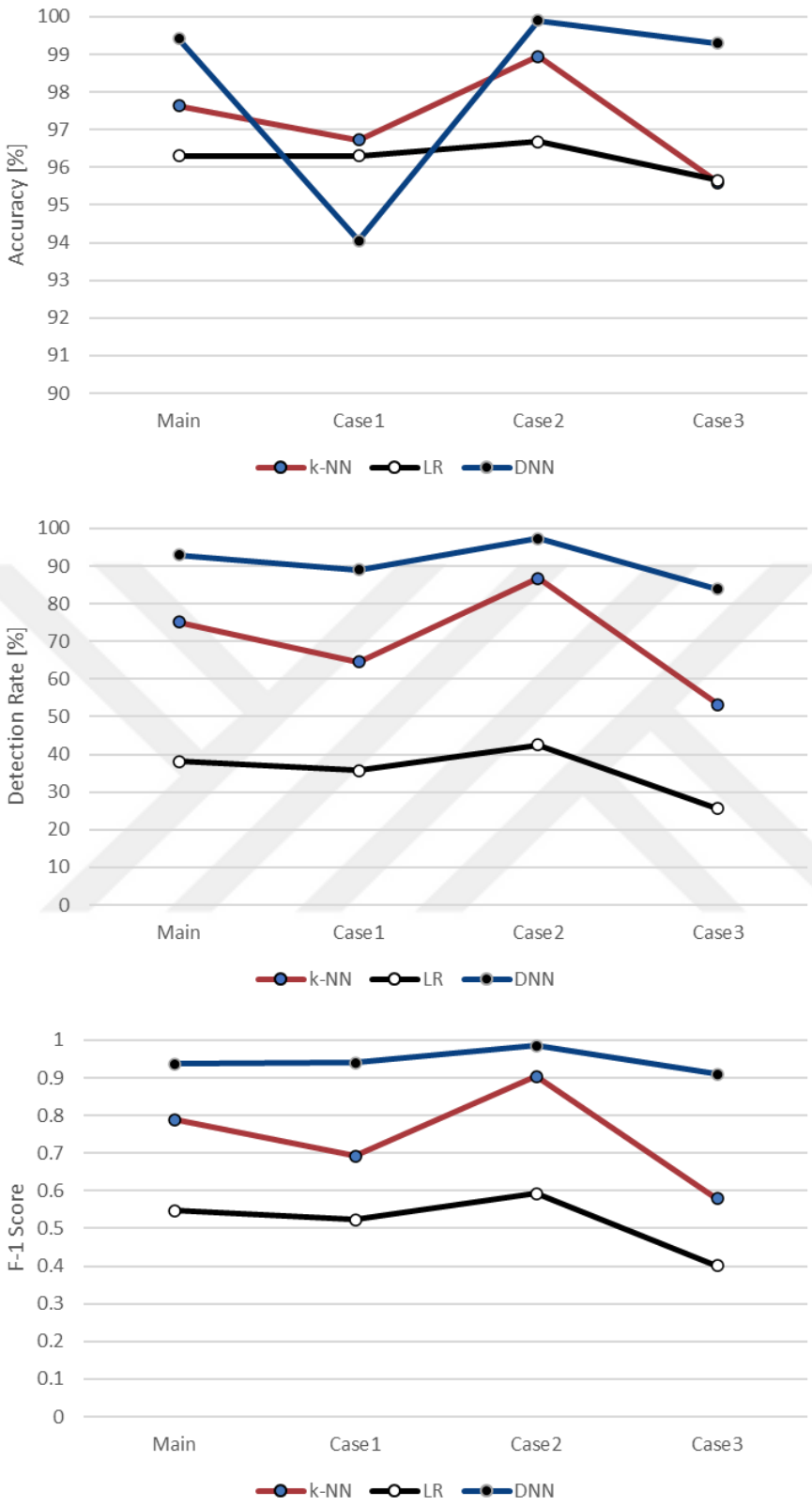


Figure 4.30 : (a) Accuracies, (b) attack detection rates and (c) F1 scores of proposed algorithms for all scenarios, IEEE 57 BUS.

5. CONCLUSIONS

In the preparation stage of the first part of the thesis, NYISO historical hourly load data (January 2012 - October 2012) were used to create the dataset. This load data was integrated into the IEEE systems, and the state estimation was performed for each load data instance and DER penetration level which are 0%, 10%, 20% and 30%. Also, White Gaussian Noise was added to the measurements to reflect the noise that can be present in real life. It is assumed that the H matrix obtained during the state estimation process is compromised by attacker and FDIA was performed using the system state vector. This attack will perform an increment attack on a randomly selected system variable, which can be bus voltage angle or bus voltage magnitude, for each load data instance by a random value between 5% and 12%. It is considered that the system only has traditional BDD methods to prevent these attacks. Measurements that include attacks that cannot be detected by traditional BDD methods are labeled as "1" to indicate the presence of an attack and included in the dataset created for ML and DL applications. Attacks detected by traditional BDD methods are not recorded. In addition, measurements that has not been attacked is labeled as "0", indicating the absence of an attack are directly included in the dataset. However, since the number labeled as attacked and not attacked in the dataset are close, to prevent overfitting, the ratio of the attacked label to the entire dataset was reduced by integrating NYISO hourly load data from June 2013 to April 2016 without any attacks into the system. Thus, a dataset was created where approximately 5% of the data instances are labeled as attacked and 95% as not attacked. There are three main datasets for IEEE test systems; 14 Bus, 30 Bus and 57 Bus. In main cases, the PMUs and wattmeters on the system had 70 dB and 5 dB of SNR values, respectively. Three different cases were added to observe the performances of the proposed algorithms on different noise levels to the PMUs and wattmeters.

Proposed algorithms for ML and DL algorithms are k-Nearest Neighbor (k-NN), Logistic Regression (LR) and Deep Neural Networks (DNN). The SNR values of PMUs and wattmeters for these three cases are 70 dB – 2.5 dB, 70 dB – 10 dB and 20

dB – 5 dB, respectively. It is expected to increase the performances of proposed algorithms at Case 2 which has lower noise. There were 40 measurements available for IEEE 14 Bus system consequently 40 inputs were supplied to ML and DL algorithms. Since the number of inputs and complexity of the system is lower compared to the other test systems, highest detection rates and F-1 scores were observed at tests conducted on IEEE 14 Bus system. The dataset created for IEEE 14 Bus system had 123,788 observations with 94.27% labelled as "0" and 5.73% labelled as "1". k-NN algorithm achieved 99.30% accuracy on the main test case of IEEE 14 Bus system when the k was equal to 1. However, upon inspecting the confusion matrix, it was seen that the k-NN algorithm detected 90.69% of the FDIAs. F-1 Score of the k-NN was calculated as 0.9364. Different values for k were tested and the highest F-1 score achieved when k was equal to 2 by 0.9408. However, upon inspecting the confusion matrix, it was observed that this increment was caused by the FP rates which decreased unlike the FN rate. Since the main aim of this study was to detect FDIAs, it is aimed to obtain lower FNs. LR algorithm obtained a 98.66% FDIA detection rate with a 0.9926 F-1 score which is satisfying result compared to the k-NN. DNN algorithm was achieved 99.8% detection rate and F-1 score of 0.9984 with two hidden layers and 10 epochs. According to the results, it was observed that the DNN achieved highest detection rate among the three algorithms. Besides, to accurately represent PMU and wattmeter measurements obtained from the real system, noise has been added to the measurements at various levels and performances of these three algorithms were observed. Since the PMUs installed in the system measure voltage in per unit and wattmeters measure power quantities in MW, it is worth noting that the effect of noise on PMUs will be more significant than on wattmeters.

When the noise of wattmeters was increased, it is observed that the detection rates of the k-NN, LR and DNN were decreased by 2.13%, 2.73% and 0.34%, respectively. In Case 2, where noises were decreased, it is observed that the detection rate of the k-NN were increased by 3.64% where LR and DNN were decreased by 0.63% and 0.34% compared to the main scenario. In Case 3, where noise of the PMUs was increased, it was observed that the k-NN outperforms the LR by 78.14% detection rate. However, DNN was performed the best detection rate with 98.67%. At every test case performed on IEEE 14 Bus system, DNN achieved highest FDIA detection rate among the proposed algorithms.

There were 95 measurements available for IEEE 30 Bus system consequently 95 inputs were supplied to ML and DL algorithms. In IEEE 30 Bus system, k-NN algorithm achieved 95.60% detection rate of the FDIAs and F-1 Score of 0.9961. Detection rate of k-NN was increased by 5.09% compared to IEEE 14 Bus main test system. This can be achieved by the ability of the k-NN algorithm to perform data extraction and data classification with a larger number of inputs. LR algorithm obtained a 38.15% FDIA detection rate with a 0.5492 F-1 score which is substantially decreased compared to IEEE 14 Bus main test case. Consequently, LR algorithm cannot be used to detect FDIAs on IEEE 30 Bus system since the attack detection rate is too low. DNN algorithm was achieved 98.82% detection rate and F-1 score of 0.9927 with two hidden layers and 10 epochs. According to the results, it was observed that the DNN achieved highest detection rate among the three algorithms.

When the noise of wattmeters was increased, it is observed that the detection rates of the k-NN, LR and DNN were decreased compared to the main scenario by 3.45%, 5.65% and 0.359%, respectively. In Case 2, where noises were decreased, it is observed that the detection rate of the all algorithms were increased to 98.23%, 46.59% and 99.69%. In Case 3, where noise of the PMUs was increased, it was observed that DNN was performed the best detection rate with 94.662% and F-1 Score of 0.9698. At every test case performed on IEEE 30 Bus system, DNN achieved highest FDIA detection rate among the proposed algorithms. It is also observed that the k-NN is outperformed the LR algorithm on the test systems where noise levels are higher and systems are more complex. Based on the results obtained in IEEE 30 Bus system, it can be concluded that the k-NN and LR algorithms are suitable for detecting FDIA in middle-level complexity systems with low noise.

The most complex system among the three system was IEEE 57 Bus system with 192 measurements. In IEEE 57 Bus system, k-NN algorithm achieved 75.08% detection rate of the FDIAs and F-1 Score of 0.789. Detection rate of k-NN was decreased critically by 20.52% compared to IEEE 30 Bus main test system. LR algorithm obtained a 38.07% FDIA detection rate with a 0.5479 F-1 score which is decreased compared to IEEE 30 Bus main test case. In both test systems, LR algorithm was not useful to detect FDIAs since it cannot detect more than 60% of the attacks. DNN algorithm was achieved 92.98% detection rate and F-1 score of 0.9365.

According to the results, it was observed that the DNN achieved highest detection rate among the three algorithms, but it performed great performance loss compared to the less complicated systems with less inputs.

When the noise on measurements that wattmeters measured was increased, it is observed that the detection rates of the k-NN, LR and DNN were decreased compared to 64.60%, 35.73% and 89.05%, respectively. In contrast, when the noises were decreased, it is observed that the detection rate of the all algorithms were increased. k-NN, LR and DNN were achieved 86.74%, 42.54% and 97.283% attack detection rate in this scenario. When the noise of the PMUs was increased, it was observed that DNN was performed the best detection rate with 83.86% and F-1 Score of 0.9092. At every test case performed on IEEE 57 Bus system, DNN achieved highest FDIA detection rate among the proposed algorithms. It is also observed that noise levels effect the detection rate of k-NN more than complexity. LR algorithm on achieves better performance with the systems that have less noise and less complex.

To summarize, results reveals that the DNN is much superior to k-NN and LR algorithm both on complex and noisy systems. As stated before, major drawback of the DNNs are the computation time and computational cost. With the continual advancements in computer science and technology, it is now possible to partially overcome this problem. Compared to DNN, k-NN and LR have lower computational costs, but FDIA detection performance of DNNs makes this drawback negligible. In addition to training time, execution time is an important parameter to investigate the usability of algorithms for stealthy False Data Injection Attack detection. This parameter indicates the time required to classify new data or a data set after training. Based on the conducted tests, the classification process of a new data set is performed by k-NN in an average of 3.4 seconds, by LR in an average of 0.0037 seconds, and by DNN in an average of 0.92 seconds. When considering the attack detection rates and execution times, the DNN algorithm proves to be the most usable algorithm in terms of average speed and accuracy rate. LR can be employed in systems characterized by a small scale and low noise, particularly in situations where an ultra-high decision speed is required.

REFERENCES

[1] **IEA** (2014), World Energy Outlook. (2014), IEA, Paris https://www .iea.org/reports /world-energy-outlook-2014, License: CC BY 4.0.

[2] **Zografopoulos, I., Konstantinou, C., & Hatziargyriou, N. D.** (2022). Distributed energy resources cybersecurity outlook: Vulnerabilities, attacks, impacts, and mitigations. arXiv preprint arXiv:2205.11171.

[3] **Syrmakesis, A. D., Alcaraz, C., & Hatziargyriou, N. D.** (2022). Classifying resilience approaches for protecting smart grids against cyber threats. International Journal of Information Security, 21(5), 1189-1210.

[4] **IEA** (2022), World Energy Outlook. (2022) IEA, Paris https://www. iea.org/reports/world-energy-outlook-2022, License: CC BY 4.0 (report); CC BY NC SA 4.0

[5] **Ma, R., Chen, H. H., Huang, Y. R., & Meng, W.** (2013). Smart grid communication: Its challenges and opportunities. IEEE transactions on Smart Grid, 4(1), 36-46.

[6] **Farhangi, H.** (2009). The path of the smart grid. IEEE power and energy magazine, 8(1), 18-28.

[7] CEA, "The smart grid: a pragmatic approach," Tech. Rep., Canadian Electricity Association, 2010, http://www.electricity.ca/ media/SmartGrid/SmartGridpaperEN.pdf

[8] **Guan, Z., Sun, N., Xu, Y., & Yang, T.** (2015). A comprehensive survey of false data injection in smart grid. International Journal of Wireless and Mobile Computing, 8(1), 27-33.

[9] **Ahmed, M. M., & Soo, W. L.** (2008, December). Supervisory control and data acquisition system (SCADA) based customized remote terminal unit (rtu) for distribution automation system. In 2008 IEEE 2nd International Power and Energy Conference (pp. 1655-1660). IEEE.

[10] **Dileep, G. (2020).** A survey on smart grid technologies and applications. Renewable energy, 146, 2589-2625.

[11] **McDonald, G., Murchu, L. O., Doherty, S., & Chien, E.** (2013). Stuxnet 0.5: The missing link. Symantec Report, 26.

[12] **Creators, W. S. S.** (2013). To Kill a Centrifuge.

[13] **Beibei L**. (2019). Detection of False Data Injection Attacks in Smart Grid Cyber-Physical Systems. [Doctor of Philosophy, Nanyang Technological University].

[14] **Bronk, C., & Tikk-Ringas, E.** (2013). The cyber attack on Saudi Aramco. Survival, 55(2), 81-96.

[15] **Bronk, C., & Tikk-Ringas, E.** (2013). Hack or attack? Shamoon and the Evolution of Cyber Conflict.

[16] **Marpaung, J. A., & Lee, H.** (2013). Dark Seoul Cyber Attack: Could it be worse?. In Conf. Indonesian Stud. Assoc. in Korea.

[17] **Url-1** <https://www.jacksonville.com/story/news/2013/02/19/jea-website-whacked-cyber-attack/15837677007>, date retrieved 10.05.2023.

[18] **Url-2** <https://www.cisa.gov/news-events/ics-alerts/ics-alert-14-176-02a>, date retrieved 10.05.2023.

[19] **Dragos Inc.,** (2017). CRASHOVERRIDE: Analysis of the Threat to Electric Grid Operations.

[20] **Shehod, A**. (2016). Ukraine power grid cyberattack and US susceptibility: Cybersecurity implications of smart grid advancements in the US. Cybersecurity Interdisciplinary Systems Laboratory, MIT, 22, 2016-22.

[21] **Lee, R. M., Assante M. J. and Conway T.** (2016). Analysis of the cyber attack on the Ukrainian power grid. SANS Industrial Control Systems.

[22] **Askarifar, S., Rahman, N. A. A., & Osman, H.** (2018). A review of latest wannacry ransomware: Actions and preventions. J. Eng. Sci. Technol, 13, 24-33.

[23] **Shakir, H. A., & Jaber, A. N.** (2018). A short review for ransomware: pros and cons. In Advances on P2P, Parallel, Grid, Cloud and Internet Computing: Proceedings of the 12th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC-2017) (pp. 401-411). Springer International Publishing.

[24] **Liu, Y., Ning, P., & Reiter, M. K**. (2011). False data injection attacks against state estimation in electric power grids. ACM Transactions on Information and System Security (TISSEC), 14(1), 1-33.

[25] **Yuan, Y., Li, Z., & Ren, K.** (2011). Modeling load redistribution attacks in power systems. IEEE Transactions on Smart Grid, 2(2), 382-390.

[26] **Sandberg, H., Teixeira, A., & Johansson, K. H.** (2010). On security indices for state estimators in power networks. In First workshop on secure control systems (SCS), Stockholm, 2010.

[27] **Esmalifalak, M., Nguyen, H., Zheng, R., & Han, Z.** (2011, October). Stealth false data injection using independent component analysis in smart grid. In 2011 IEEE international conference on smart grid communications (SmartGridComm) (pp. 244-248). IEEE.

[28] **Hug, G., & Giampapa, J. A.** (2012). Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks. IEEE Transactions on smart grid, 3(3), 1362-1370.

[29] **Liu, X., & Li, Z.** (2014). Local load redistribution attacks in power systems with incomplete network information. IEEE Transactions on Smart Grid, 5(4), 1665-1676.

[30] **Manandhar, K., Cao, X., Hu, F., & Liu, Y.** (2014). Detection of Faults and Attacks Including False Data Injection Attack in Smart Grid Using

Kalman Filter. IEEE Transactions on Control of Network Systems, 1(4), 370–379. doi:10.1109/tcns.2014.2357531

[31] **Manandhar, K., Cao, X., Hu, F., & Liu, Y.** (2014, February). Combating false data injection attacks in smart grid using kalman filter. In 2014 International Conference on Computing, Networking and Communications (ICNC) (pp. 16-20). IEEE.

[32] **Rawat, D. B., & Bajracharya, C.** (2015). Detection of false data injection attacks in smart grid communication systems. IEEE Signal Processing Letters, 22(10), 1652-1656.

[33] **Chaojun, G., Jirutitijaroen, P., & Motani, M.** (2015). Detecting False Data Injection Attacks in AC State Estimation. IEEE Transactions on Smart Grid, 6(5), 2476–2483. doi:10.1109/tsg.2015.2388545

[34] **Wang, D., Guan, X., Liu, T., Gu, Y., Sun, Y., & Liu, Y.** (2013, December). A survey on bad data injection attack in smart grid. In 2013 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC) (pp. 1-6). IEEE.

[35] **Mohammadpourfard, M., Sami, A., & Seifi, A. R.** (2017). A statistical unsupervised method against false data injection attacks: A visualization-based approach. Expert Systems with Applications, 84, 242–261. doi:10.1016/j.eswa.2017.05.013

[36] **Kim, J., Tong, L., & Thomas, R. J.** (2014). Data framing attack on state estimation. IEEE Journal on selected areas in communications, 32(7), 1460-1470.

[37] **Kim, T. T., & Poor, H. V.** (2011). Strategic protection against data injection attacks on power grids. IEEE Transactions on Smart Grid, 2(2), 326-333.

[38] **Chen, P.-Y., Yang, S., McCann, J. A., Lin, J., & Yang, X. (**2015). Detection of false data injection attacks in smart-grid systems. IEEE Communications Magazine, 53(2), 206–213. doi:10.1109/mcom.2015.7045410

[39] **Sou, K. C., Sandberg, H., & Johansson, K. H.** (2013). On the Exact Solution to a Smart Grid Cyber-Security Analysis Problem. IEEE Transactions on Smart Grid, 4(2), 856–865. doi:10.1109/tsg.2012.2230199

[40] **Khanna, K., Panigrahi, B. K., & Joshi, A.** (2018). AI-based approach to identify compromised meters in data integrity attacks on smart grid . IET Generation, Transmission & Distribution, 12(5), 1052–1066. doi:10.1049/iet-gtd.2017.0455

[41] **Liu, L., Esmalifalak, M., Ding, Q., Emesih, V. A., & Han, Z.** (2014). Detecting false data injection attacks on power grid by sparse optimization. IEEE Transactions on Smart Grid, 5(2), 612-621.

[42] **Ozay, M., Esnaola, I., Vural, F. T. Y., Kulkarni, S. R., & Poor, H. V.** (2012, November). Distributed models for sparse attack construction and state vector estimation in the smart grid. In 2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm) (pp. 306-311). IEEE.

[43] **Ozay, M., Esnaola, I., Vural, F. T. Y., Kulkarni, S. R., & Poor, H. V.** (2015). Machine learning methods for attack detection in the smart grid. IEEE transactions on neural networks and learning systems, 27(8), 1773-1786.

[44] **Goldstein, M., & Uchida, S.** (2016). A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. PloS one, 11(4), e0152173.

[45] **McLaughlin, S., Holbert, B., Fawaz, A., Berthier, R., & Zonouz, S.** (2013). A multi-sensor energy theft detection framework for advanced metering infrastructures. IEEE journal on selected areas in communications, 31(7), 1319-1330.

[46] **Ashrafuzzaman, M., Chakhchoukh, Y., Jillepalli, A. A., Tosic, P. T., de Leon, D. C., Sheldon, F. T., & Johnson, B. K.** (2018, June). Detecting stealthy false data injection attacks in power grids using deep learning. In 2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC) (pp. 219-225). IEEE.

[47] **He, Y., Mendis, G. J., & Wei, J.** (2017). Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism. IEEE Transactions on Smart Grid, 8(5), 2505-2516.

[48] **Wei, J., & Mendis, G. J.** (2016, April). A deep learning-based cyber-physical strategy to mitigate false data injection attack in smart grids. In 2016 Joint Workshop on Cyber-Physical Security and Resilience in Smart Grids (CPSR-SG) (pp. 1-6). IEEE.

[49] **Ayad, A., Farag, H. E., Youssef, A., & El-Saadany, E. F.** (2018, February). Detection of false data injection attacks in smart grids using recurrent neural networks. In 2018 IEEE power & energy society innovative smart grid technologies conference (ISGT) (pp. 1-5). IEEE.

[50] **Kolosok, I., Korkina, E., Mahnitko, A., & Gavrilovs, A.** (2018). Supporting Cyber-Physical Security of Electric Power System by the State Estimation Technique. 2018 IEEE 59th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON). doi:10.1109/rtucon.2018.8659853

[51] **Chakhchoukh, Y., & Ishii, H.** (2019). Cyber security for power system state estimation. Smart Grid Control: Overview and Research Opportunities, 241-256.

[52] **Abur, A., & Exposito, A. G.** (2004). Power system state estimation: theory and implementation. CRC press.

[53] **Url-3** <https://integratedmlai.com/normal-distribution-an-introductory-guide-to-pdf-and-cdf/>, date retrieved 10.05.2023.

[54] **Url-4** <https://www.mathworks.com/solutions/machine-learning.html?s_tid=srchtitle_machine%20learning_3>, date retrieved 10.05.2023.

[55] **Panesar, S. S., D'Souza, R. N., Yeh, F.-C., & Fernandez-Miranda, J. C.** (2019). Machine Learning Versus Logistic Regression Methods for 2-Year Mortality Prognostication in a Small, Heterogeneous Glioma Database. World Neurosurgery: X, 2, 100012. doi:10.1016/j.wnsx.2019.100012

[56] **Haykin S.** (2009). Neural Networks and Learning Machines. 3rd ed. Hamilton, Ontario, Canada: Pearson Education, Inc.

[57] **Url-5** <https://www.global-engage.com/life-science/deep-learning-in-digital-pathology/>, date retrieved 10.05.2023.

[58] **Kingma, D. P., & Ba, J.** (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

[59] **Bock, S., Goppold, J., & Weiß, M.** (2018). An improvement of the convergence proof of the ADAM-Optimizer. arXiv preprint arXiv:1804.10587.

[60] **Kulkarni, A., Chong, D., & Batarseh, F. A.** (2020). Foundations of data imbalance and solutions for a data democracy. Data Democracy, 83–106. doi:10.1016/b978-0-12-818366-3.00005-8

[61**] Chaojun, G., Jirutitijaroen, P., & Motani, M.** (2015). Detecting false data injection attacks in AC state estimation. IEEE Transactions on Smart Grid, 6(5), 2476-2483.

[62] **The MathWorks Inc.** (2022). MATLAB version: 9.13.0 (R2022b), Natick, Massachusetts: The MathWorks Inc. https://www.mathworks.com

[63] **Zimmerman, R. D., Murillo-Sánchez, C. E., & Thomas, R. J.** (2010). MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. IEEE Transactions on power systems, 26(1), 12-19.

[64] **Url-6** < http://www.energyonline.com/Data/GenericData.aspx?DataId=13> date retrieved 10.05.2023.

[65] **Url-7** <http://labs.ece.uw.edu/pstca/pf14/14bus600.tif> date retrieved 10.05. 2023.

[66] **Url-8** <http://labs.ece.uw.edu/pstca/pf30/30bus600.tif> date retrieved 10.05. 2023.

[67] **Url-9** <https://icseg.iti.illinois.edu/files/2013/10/IEEE57.png> date retrieved 10.05.2023.

[68] **Dinh T., Nguyen T., Phan H., Dau V., Dao D. and Nguyen N.** (2023). Physical Sensors: Thermal Sensors. Encyclopedia of Sensors and Biosensors (First Edition). Elsevier.

[69] **Alpaydın, E.,** (2020). Introduction to Machine Learning (Adaptive Computation and Machine Learning series), MIT Press.

[70] **Aoufi, S., Derhab, A., & Guerroumi, M.** (2020). Survey of false data injection in smart power grid: Attacks, countermeasures and challenges. Journal of Information Security and Applications, 54, 102518. doi:10.1016/j.jisa.2020.102518

[71] **Liu, X., & Li, Z.,** (2017). False data attack models, impact analyses and defense strategies in the electricity grid. The Electricity Journal, 30(4), 35–42. doi:10.1016/j.tej.2017.04.001

[72] **Soltan, S., Mittal, P., & Poor, H. V.** (2019). Line Failure Detection After a Cyber-Physical Attack on the Grid Using Bayesian Regression. IEEE Transactions on Power Systems, 34(5), 3758–3768. doi:10.1109/tpwrs.2019.2910396

[73] **Li, Y., & Wang, Y.** (2019). False Data Injection Attacks with Incomplete Network Topology Information in Smart Grid. IEEE Access, 7, 3656–3664. doi:10.1109/access.2018.2888582

[74] **Bobba, R., Davis, K., Wang, Q., Khurana, H., Nahrstedt, K., & Overbye, T.,** (2010). Detecting False Data Injection Attacks on DC State Estimation.

[75] **Xu, R., Chen, D., & Wang, R.** (2020). Data Integrity Attack Detection for Node Voltage in Cyber-Physical Power System. Arabian Journal for Science and Engineering. doi:10.1007/s13369-020-04813-y

[76] **Kaviani, R. & Hedman, K.,** (2020). An Enhanced Energy Management System Including a Real-Time Load-Redistribution Threat Analysis Tool and Cyber-Physical SCED.

[77] **Mohammadi, F., & Rashidzadeh, R.** (2022). Impact of stealthy false data injection attacks on power flow of power transmission lines—A mathematical verification. International Journal of Electrical Power & Energy Systems, 142, 108293.

[78] **Sridhar, S., & Manimaran, G.** (2011, July). Data integrity attack and its impacts on voltage control loop in power grid. In 2011 IEEE power and energy society general meeting (pp. 1-6). IEEE.

[79] **Liu, S., Chen, B., Zourntos, T., Kundur, D., & Butler-Purry, K.** (2014). A coordinated multi-switch attack for cascading failures in smart grid. IEEE Transactions on Smart Grid, 5(3), 1183-1195.

[80] **Chen, B., Pattanaik, N., Goulart, A., Butler-Purry, K. L., & Kundur, D.** (2015, May). Implementing attacks for modbus/TCP protocol in a real-time cyber physical system test bed. 2015 IEEE International Workshop Technical Committee on Communications Quality and Reliability (CQR) (pp. 1-6). IEEE.

[81] **Chen, B., Butler-Purry, K. L., Goulart, A., & Kundur, D.** (2014, September). Implementing a real-time cyber-physical system test bed in RTDS and OPNET. In 2014 North American Power Symposium (NAPS) (pp. 1-6). IEEE.

[82] **Agrawal, A., Momin, D. M., Syndor, D., & Affijulla, S.** (2020, June). Impact analysis of cyber attack under stable state of power system: voltage stability. In 2020 IEEE Region 10 Symposium (TENSYMP) (pp. 402-405). IEEE.

[83] **Lei, H., Chen, B., Butler-Purry, K. L., & Singh, C**. (2018, May). Security and reliability perspectives in cyber-physical smart grids. In 2018 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia) (pp. 42-47). IEEE.

[84] **Ghafouri, M., Au, M., Kassouf, M., Debbabi, M., Assi, C., & Yan, J.** (2020). Detection and mitigation of cyber attacks on voltage stability monitoring of smart grids. IEEE Transactions on Smart Grid, 11(6), 5227-5238.

[85] **Shapsough, S., Qatan, F., Aburukba, R., Aloul, F., & Al Ali, A. R.** (2015, October). Smart grid cyber security: Challenges and solutions. In 2015 international conference on smart grid and clean energy technologies (ICSGCE) (pp. 170-175). IEEE.

[86] **Chen, B., Butler-Purry, K. L., Nuthalapati, S., & Kundur, D.** (2014, July). Network delay caused by cyber attacks on SVC and its impact on transient stability of smart grids. In 2014 IEEE PES General Meeting| Conference & Exposition (pp. 1-5). IEEE.

[87] **Di Giorgio, A., Giuseppi, A., Liberali, F., Ornatelli, A., Rabezzano, A., & Celsi, L. R.** (2017, July). On the optimization of energy storage system placement for protecting power transmission grids against dynamic load altering attacks. In 2017 25th Mediterranean Conference on Control and Automation (MED) (pp. 986-992). IEEE.

[88] **Chen, B., Mashayekh, S., Butler-Purry, K. L., & Kundur, D.** (2013, July). Impact of cyber attacks on transient stability of smart grids with voltage support devices. In 2013 IEEE Power & Energy Society General Meeting (pp. 1-5). IEEE.

[89] **Chen, B., Butler-Purry, K. L., & Kundur, D.** (2013, September). Impact analysis of transient stability due to cyber attack on FACTS devices. In 2013 North American Power Symposium (NAPS) (pp. 1-6). IEEE.

[90] **Baumeister, T.** (2010). Literature review on smart grid cyber security. Collaborative Software Development Laboratory at the University of Hawaii, 650.

[91] **Amini, S., Pasqualetti, F., & Mohsenian-Rad, H.** (2016). Dynamic load altering attacks against power system stability: Attack models and protection schemes. IEEE Transactions on Smart Grid, 9(4), 2862-2872.

[92] **Anzalchi, A., & Sarwat, A.** (2015, April). A survey on security assessment of metering infrastructure in smart grid systems. In SoutheastCon 2015 (pp. 1-4). IEEE.

[93] **Yang, Y., Littler, T., Sezer, S., McLaughlin, K., & Wang, H. F.** (2011, December). Impact of cyber-security issues on smart grid. In 2011 2nd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies (pp. 1-7). IEEE.

[94] **Monticelli, A**. (2012). State estimation in electric power systems: a generalized approach. Springer Science & Business Media.

[95] **Khaitan, S. K., McCalley, J. D., & Liu, C. C.** (2015). Cyber physical systems approach to smart electric power grid. Springer.

[96] **Freund, Y., & Schapire, R. E.** (1998, July). Large margin classification using the perceptron algorithm. In Proceedings of the eleventh annual conference on Computational learning theory (pp. 209-217).

[97] **Ge, Q.** (2020). Detection of Stealthy False Data Injection Attacks Against State Estimation in Electric Power Grids Using Deep Learning Techniques. Theses and Dissertations. 2504. [Master of Science Thesis, The University of Wisconsin-Milwaukee].

CURRICULUM VITAE

Name Surname : Can Gürkan

EDUCATION :

- **B.Sc. :** 2019, Bahcesehir University, School of Engineering and Natural Sciences, Electrical and Electronics Engineering

PROFESSIONAL EXPERIENCE AND REWARDS:

- 2021 - Currently working as a power system engineer at Lean Power Solutions (LPS ENERJİ VE BİLİŞİM TEKNOLOJİLERİ SAN.TİC.LTD.ŞTİ.)

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Gürkan C.,** Mohammadpourfard M., Aksoy N., Genç V.M.İ. 2023: Deep Learning - Based FDIA Detection with High Penetration Rate of DERs. International Graduate Research Symposium, May 16-18, 2023 Istanbul, Turkey.

OTHER PUBLICATIONS, PRESENTATIONS AND PATENTS:

- Dashti, N., Gökçe, C., Başağaç, Z., Öner, H., Yener, U., Süslüoğlu, B., Önal, G., **Gürkan, C.** and Özçelik, O. ,2022. Modelling and Analysis of Delta-connected Power Transformers Grounded Using Zig-Zag Transformer. In 2022 4th Global Power, Energy and Communication Conference (GPECOM) (pp. 406-413). IEEE.