CYBERSECURITY FOR NUCLEAR POWER PLANTS

WORKING WITH SIMULATOR'S DATA AND MACHINE LEARNING

ALGORITHMS TO FIND ABNORMALITIES AT NUCLEAR POWER PLANTS

A Thesis

by

MULKIYE SUMEN

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Dilma Da Silva
Co-Chair of Committee,	Shaheen Azim Dewji
Committee Members,	Martin Carlisle
	Eman Hammad
Head of Department,	Scott Schaefer

August 2022

Major Subject: Computer Engineering

ABSTRACT

Cybersecurity has the utmost importance for nuclear power plants (NPPs). Demand for clean and constant energy has increased the need and use of NPPs. Countries want to have and maintain secure NPPs both physically (well-studied area) and digitally. We live in a digital world, and cyber-attacks have skyrocketed in recent years. This study explores the cyber risk for NPPs, digital attacks, potential future attacks, international aspects, and law and policy requirements of cyber protection for nuclear power plants. With the help of data analysis and machine learning algorithms, extra monitoring can be conducted on plants' data. Data monitoring applications require comprehensive data to build models and develop solutions. However, nuclear facilities do not share their data because of security concerns. Plant simulators are heavily used for training people and conducting experiments. In this thesis, we inspect plant simulators to assess their usability by people with a technical background such as cyber experts, information technology technicians, and software developers.

People responsible for protecting digital systems can benefit from the help of data analytic tools and machine learning models to detect abnormalities. We study machine learning models on simulator data to examine their potential in identifying anomalies.

ACKNOWLEDGEMENTS

I would like to thank my committee chairs Dr. Da Silva and Dr. Dewji, and my committee members Dr. Hammad and Dr. Carlisle, for their guidance and support throughout my studies and the covid-19 pandemic.

I also want to thank my virtual classmates and the department faculty and staff for their help in protecting us from the virus.

Finally, I want to thank Deniz and my family for their patience, love, and support.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Professor Dilma Da Silva of the Department of Computer Science and Engineering and Professor Shaheen Azim Dewji of the Department of Nuclear Engineering, Professor Martin Carlisle of the Department of Computer Science and Engineering, and Professor Eman Hammad of the Department Cybersecurity and Computer Science Texas A&M University-Commerce.

All other work conducted for the thesis was completed by the student independently.

Funding Sources

This graduate study was supported by the Republic of Turkey Ministry of National Education.

NOMENCLATURE

APT	Advanced Persistent Threat
BOC	Beginning of Cycle
BWR	Boiling Water Reactor
CISA	Computer Information Sharing Act
CNC	Computer Numerical Control
CPS	Cyber-physical Systems
CRF	Code of Federal Regulations
DCS	Distributed Control System
DoS	Denial of Service
DT	Decision Tree
EOC	End of Cycle
HIPAA	Health Insurance Portability and Accountability Act
HMI	Human-Machine Interface
HTR	Power Pressurizer Heater
I&C	Instrumentation and Control Systems
IAEA	International Atomic Energy Agency
IC	Integrated Circuit
ICS	Industrial Control Systems
IED	Intelligent Electronic Device
IoT	Internet of Things

ISO	International Organization for Standardization
IT	Information Technology
KNN	K-Nearest Neighbors
LR	Logistic Regression
MBK	Integrated Break Flow
MDB	Microsoft Access Database
ML	Machine Learning
MOC	Middle of Cycle
NEA	Nuclear Energy Agency
NEI	Nuclear Energy Institute
NIST	National Institute of Standards and Technology
NPP	Nuclear Power Plants
NPT	Nuclear Nonproliferation Treaty
NRC	Nuclear Regulatory Commission
OS	Operating System
OT	Operational Technology
PCTRAN	Personal Computer Transient Analyzer
PLC	Programmable Logic Controllers
PPC	Plant Process Computer
PPS	Physical Protection System
PWR	Pressurized Water Reactor
RG	Regulatory Guide

RPS	Reactor Protection System
RT	Random Forest
RTU	Remote Terminal Units
SCADA	Supervisory Control and Data Acquisition System
SPDS	Safety Parameter Display System
SSL	Secure Socket Layer
SVM	Support Vector Machine
TLS	Transport Layer Security
USA	United States of America
USB	Universal Serial Bus
WBK	Flow Total Break Entering Reactor Building
WCHG	Flow Charging
WECS	Flow Total Emergency Core Cooling System
WLR	Flow Reactor Cooling System Leak

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

Cybersecurity for nuclear power plants (NPPs) is a crucial subject. A possible accident or incident that can lead to radioactive exposure to the environment is a local and a global problem. Every NPP has its unique design and hardware-software infrastructure, so there are no generic solutions to protect NPPs from cyber-attacks. However, some steps can be taken to decrease the attack surface and increase the protection. Our study examines known and potential problems and practical solutions to them.

Why do we need nuclear energy plants while a possible accident might be too dangerous to life and the environment? Climate change, increase in population, and demand for electricity force countries and organizations to find clean and constant energy sources. Within the currently available energy sources, nuclear energy is one of the best answers to these problems [1]. However, accidents like Chernobyl [2] [3] [4] [5] make people biased toward nuclear power plants. Furthermore, nuclear accidents might cause massive problems for countries, so terrorist organizations might target them to reach their goals. As a result, building impenetrable nuclear structures both physically and digitally is a primary goal of governments and organizations.

In this study, we focus on the digital-cyber protection of NPPs. We explore the feasibility of finding abnormalities in the plant data with the help of data analysis and machine learning algorithms. Our threat model assumes that the attacker has succeeded in passing every control, reaching the plant site, and manipulating data values before the operators monitor the data. If operators who monitor the reactor cannot see the accurate data and, therefore, are not aware of what is happening in the reactor building, they cannot prevent accidents. We propose extra controllers, like software built with machine learning models, to help controllers watch over the reactor.

2. BACKGROUND

2.1. Cybersecurity

This subsection lists some of the issues and concerns related to cybersecurity. Detailed description of these issues can be found in many books and texbooks (for example, in [6], [7], and [8].

Cybersecurity is one of the most crucial matters of this century. Everything is getting digitized and getting vulnerable to cyber-attacks. Cybersecurity is the practice of protecting information (data) and information systems (devices that gather, process, and hold data), networks, systems, and programs from digital attacks and failures. Attacks aim to access, modify, or destroy sensitive information, extort money from users (e.g., ransom), or interrupt normal business processes (e.g., denial of services).

When someone with malicious intentions acts, a problem may occur. Some people search for flaws in systems they can leverage to obtain relevant information. They can use the information to extort money, sell to competitors (for example, intellectual property such as patents or secret formulas), gain access to people's financial accounts, or access critical systems to disrupt services and cause harm to organizations.

The subject of cybersecurity is not just about adversary attacks but also includes system maintenance-related issues. A possible network failure during an upgrade or an emergency like a natural disaster can result in catastrophic damage and the loss of systems and data.

An attack can come from a random person, a hacker, a terrorist group, or an international agency. Attackers can target various systems, for example, automotive systems, wearable health-related devices, intrusion detection and prevention systems, systems that keep track of critical processes, the networks connected to the Internet, or even the closed-network systems. An attack on critical infrastructures such as nuclear power plants, water systems, and power grids can cause environmental damage and even loss of lives.

The three aspects of data are the main focuses of cyber protection: confidentiality, integrity, and availability, known as the CIA triad [9]. Losing one of these can create a vulnerability and result in a cyber accident. With the help of authentication, authorization, accounting (commonly known as the AAA of security), and other security concepts, information can be secured [9].

Confidentiality refers to protecting undisclosed information from unauthorized access. Authorization happens when access is given to the users for specific data or locations. Only authorized users can access information or an area. Confidentiality can be assured by encryption, authentication, access control, and physical security techniques. Authentication, in particular, has a vital role in avoiding the unauthorized disclosure of information. It requires something users know (e.g., username, password), something users are (e.g., fingerprints, eye/retina scan, face ratio), something users have (token with secret keys, identity cards, badges), something users do (speaking, signing), somewhere users are (location), or other user-specific information proof to give access to data.

Integrity prevents data modification without proper authorization or the owner's consent. For example, a malicious compromise might come from malware altering sensitive sensors' values at a nuclear reactor, possibly causing a massive disaster. Thus, building computer systems tools that support integrity is vital. The mechanisms for protecting the confidentiality of information can also help with data integrity. Moreover, there are methods specifically designed to ensure integrity, like accounting, backups, and checksums. In case of a data loss or unwanted change, backup data can be used, and with the help of accounting and checksums, we can detect when a breach of data has occurred. Accounting keeps track of data, network usage, computer usage, building entrance, and logs all this information as proof of access or change. With the help of logged data, an insider threat or data breach can be traced back and found (achieving non-repudiation). Being able to trace a breach also helps to model a protection method, assess weaknesses, and prevent similar actions.

Availability refers to being able to access, use, store and protect the data all the time. Data can be made secure by storing it in a removable hard drive that is kept locked in a safe. However, this approach will not be beneficial for users who want to constantly access and use data, like bank accounts, sensor data, and credential access information. Proper protection of hardware and technical infrastructure and systems must be assured by the previously mentioned CIA security concepts and additional tools such as physical protection and computational redundancy. Physical protections keep information available even under extreme conditions like storms, earthquakes, flooding, fire, or bomb blasts. Backup power generators, water towers, and strong walls can be part of the physical protection. Computational redundancies are computers and memory devices that serve as a backup in case of failures. Companies build data centers in different cities, countries, and even continents to pursue data availability.

Every organization has its unique expectations, and for some organizations, availability and integrity can be more critical than confidentiality. As such, data correctness (integrity) and reaching the data constantly (availability) can be emphasized by the organization. The risk of losing confidentiality, integrity, or availability of information varies. As a result, having well-studied and planned cybersecurity risk assessments and risk management is crucial to organizations.

2.1.1. Cybersecurity Risk Management

According to the International Organization for Standardization (ISO), risk management is the process of identifying, evaluating, and prioritizing risks, then deploying resources in a coordinated and cost-effective manner to reduce the likelihood and impact of unfavorable events or maximize the realization of opportunities [10].

There are many types of methods to calculate cybersecurity risk. Most of them have common steps like deciding on assets, assessing the vulnerability of assets, identifying threats, determining the likelihood of risk, identifying the methods to reduce the risk, and prioritizing risk reduction measures.

2.1.1.1. Assets

Assets (information and information systems) are part of risk assessments. Deciding on the assets and what is going to be protected will be helpful to design systems. While designing security systems, often managing the budget plays a vital role at the first level. Organizations sometimes give up on some security measures to reduce expenses, so identifying the most critical assets can provide better risk management processes and cyber protection.

2.1.1.2. Vulnerabilities

A vulnerability is a weakness of an asset, design, or system that can be exploited to cause harm. Cyber vulnerabilities can result from lacking security policies, outdated software and hardware, accidents, poorly trained employees, and more. Many vulnerabilities remain unknown in deployed systems, which are called Zero-Day vulnerabilities. The National Institute of Standards and Technology (NIST) has a vulnerability database open to the public [11], and there are also other public sources and private companies that supply information. Following these platforms may help conduct broader vulnerability scanning and reduce the number of weaknesses in the systems.

2.1.1.3. Threats

A threat is a way of exploiting a vulnerability of an asset. Anything that can cause harm, loss, damage, or compromise of information systems is perceived as a threat. Threats are external and out of our control; we cannot foresee when an attack comes; meanwhile, vulnerabilities are under our control. We can get rid of vulnerabilities and mitigate the threats. Cyber threats have been increasing with digitalization and getting more sophisticated. Threats vary in format; examples include malware (e.g., ransomware, trojans, worms, etc.), unpatched security vulnerabilities, hidden backdoor programs, superuser or admin user account privileges, phishing (social engineering attacks), Cross-site Scripting (XSS), Denial of Service (DoS), session hijacking and Man-in-the-middle attacks, user-related Credential Reuse (same password everywhere), Thwart attacks, Row Hammer attacks, system failures, third-party software use, rootkits, and physical attacks [12].

Malware is a shortened version of “malicious software;” it is designed to infiltrate systems. Viruses, worms, trojan horses, ransomware, spyware, rootkits, and spams have different properties, but they are all types of malware. Some malware requires user action (like installing or running an application) to infect the system and spread, while others can be activated or replicate themselves without user interaction. Moreover, some malware acts as a helpful program; while doing its job, it also does other things (using processors for other functions or gathering personal data) in the systems without the user's consent. According to AV-TEST (The Independent IT-Security Institute), thousands of new malicious software and potentially unwanted applications are introduced daily. Figure 2.1 depicts the growth of malware over the last decade [13].

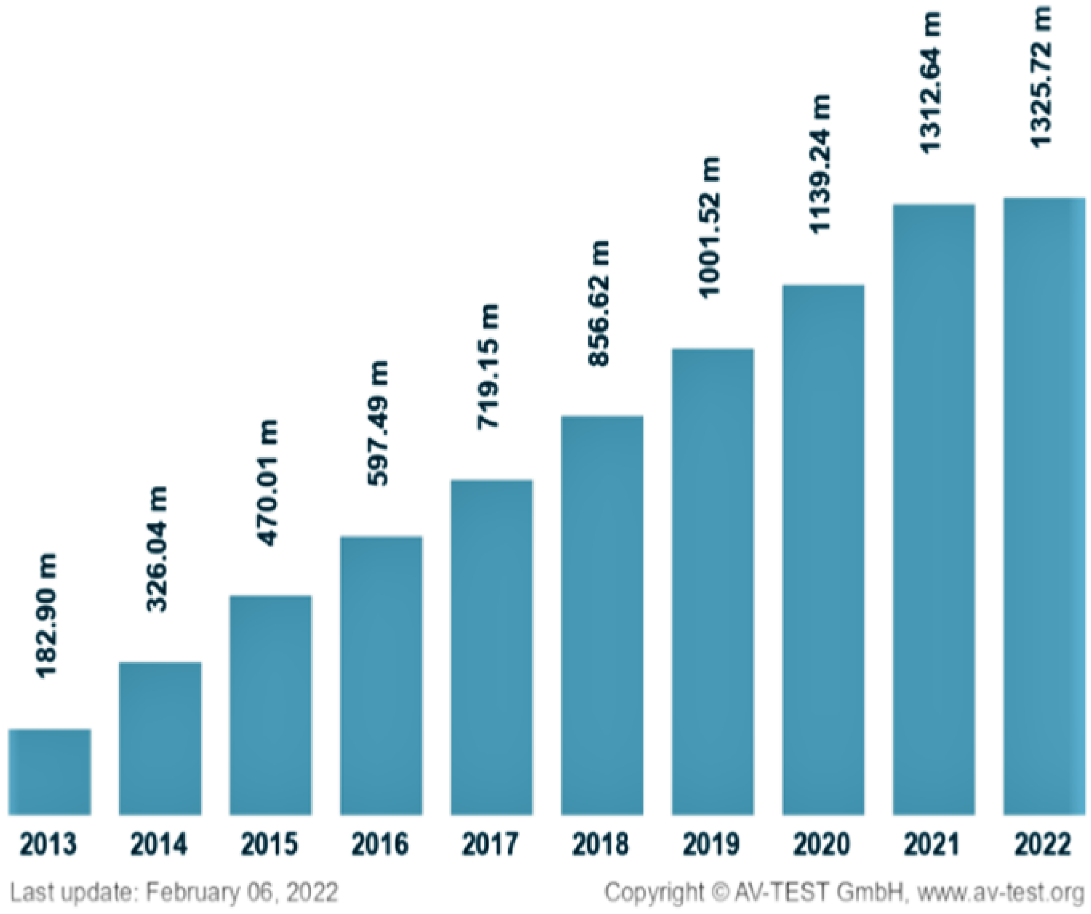


Figure 2.1 AvTest daily chart of malware as of February 6, 2022 [13]

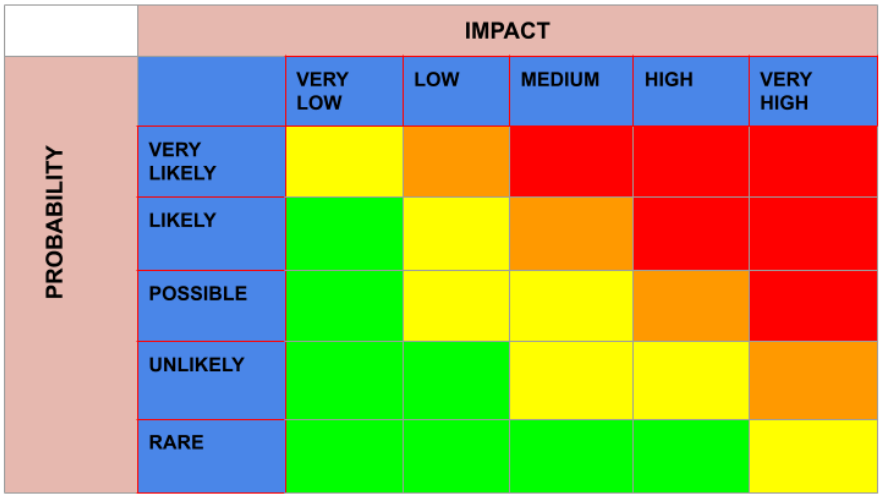
2.1.1.3.1. Threat Actors

Threat actors can be hackers, script kiddies, hacktivists, organizations (e.g., WikiLeaks, Anonymous), advanced persistent threats (APTs), terrorist groups, and intruders [14]. Although some of them have motives (occasionally good and motivated by curiosity), their acts frequently violate the intended usage of the systems. The outcomes can be harmless mischief (developing a virus with no malicious intent) to malicious conduct (stealing or altering information). Threats also can be physical, like natural disasters such as fire and earthquakes.

2.1.1.4. Calculation of Risk

Risk is the potential impact of a threat and the likelihood of that threat occurring. There are ways to calculate risk qualitatively and quantitively. More information can be found in many sources [15][16][17] .

Qualitative analysis uses expert experience, intuition, judgment, and other methods to assign a relative value to the risk (low, medium, high, and critical). Figure 2.2 shows an example of a Qualitative Analysis Matrix.



		IMPACT				
		VERY LOW	LOW	MEDIUM	HIGH	VERY HIGH
PROBABILITY	VERY LIKELY	🟨	🟧	🟥	🟥	🟥
	LIKELY	🟩	🟨	🟧	🟥	🟥
	POSSIBLE	🟩	🟨	🟨	🟧	🟥
	UNLIKELY	🟩	🟩	🟨	🟨	🟧
	RARE	🟩	🟩	🟩	🟩	🟨

Figure 2.2 An example of a Qualitative Analysis Matrix [15]

The quantitative analysis relies on numbers. It includes assigning values to the value of assets, the threat likelihood, the severity of vulnerabilities, and the consequences of a given threat. The risk impact or magnitude of impact estimates the amount of damage. People often prefer to see numbers, especially for financial reasons. Single loss expectancy, annualized rate of occurrence, and annualized loss expectancy are the most common calculations used in determining the magnitude of an impact.

While calculating the cybersecurity risk, it is essential to understand that using only one method might not be enough. Working with a hybrid calculation method can be more effective.

Risk = Vulnerabilities x Threats x Consequences [16] [17]

In this simple calculation, any of the variables cannot be accepted as zero; as a result, the risk is always bigger than zero. The basic understanding of the security experts must be that "a system always carries risk." Even after taking all reasonable precautions, there is still a risk called residual risk. Accepting that there is always a risk is one of the cybersecurity principles. Absolute security is impossible, and security is only as strong as the weakest link [15][16][17].

2.1.1.5. Mitigation of Threats

Once the risk is calculated, four approaches can be taken. Risk can be avoided, transferred, mitigated, or accepted. Risk avoidance involves either stopping the risky activity or choosing a less risky alternative, for example, upgrading an operating system (OS) to a more secure version. Risk transfer is passing the risk to a third party. Accepting the risk (and its consequences) is not making any improvements. This may happen when measures cost more than risk consequences. Mitigation is a strategy to minimize the risk to an acceptable level. Risk calculation includes low, medium, high, or critical levels of risk information. For example, based on this information, critical risk groups can be avoided or mitigated to a lower level [17].

Several basic controls can be used to mitigate threats. Controls are actions, procedures, devices, and techniques that mitigate the vulnerability of systems. Physical controls include security guards, identification cards, alarm systems, locks, and surveillance cameras. Technical controls comprise encryption (symmetric- asymmetric), digital signatures, secure sockets layer (SSL), transport layer security (TLS), digital certificates, smart cards, access control lists, and network authentication. There are administrative (managerial) controls in addition to technological and physical controls. User training, security policies, procedures, contingency, and recovery plans are all covered by administrative controls.

Administrative controls can be broken down into two categories which are procedural and regulatory controls. Procedural controls cover things that organizations choose to do on their own. On the other hand, legal and regulatory controls cover things that organizations must have because of the law. For example, if the organization uses personal health information, it must follow The Health Insurance Portability and Accountability Act (HIPAA).

Finally, even the most secure systems are vulnerable to plugging affected USB thumps, opening malware-infused emails, and connecting compromised computers when the system's users ignore them. As a result, user training plays a significant role in securing systems.

2.1.1.5.1. Threat Intelligence, Sources, and Information Sharing

The Computer Information Sharing Act (CISA) is a USA Federal Law for companies to share attack information and possible defense measures with the government [18]. The information includes cyber-threat indicators like malicious reconnaissance, a method to defeat a security control, a vulnerability, a malicious cyber command, and any other attributes of a cyber threat. Government informs the public or the companies about new findings, so security measures can be taken before the same attack happens on companies' systems.

There are also other platforms that share information about security threats and vulnerabilities. Some organizations do not share information publicly, so they profit from threats. Also, some industries and organizations avoid sharing their cybersecurity-related experiences publicly because of privacy concerns or potential loss of reputation. So, there might be more threats, attacks, and vulnerabilities that remain unknown to the public, and cyber risk might be more significant than the calculated one.

The National Institute of Standards and Technology (NIST) has a broad and well-established cybersecurity framework. The NIST Cybersecurity Framework [19] can be applied to every industry and organization relying on technology, including information technology (IT), industrial control systems (ICS), cyber-physical systems (CPS), or connected devices more generally, including the Internet of Things (IoT) [19].

Continuously monitoring and assessing risk management plans are crucial for cybersecurity. Technology changes too fast, and while systems are getting better at addressing known vulnerabilities, many other vulnerabilities continue to emerge every day.

2.2. Nuclear

2.2.1. History of Nuclear Energy

Nuclear power has a long and noted history [20][21][22]. People began to learn more about the nucleus and its properties when the structure of an atom (composed of a nucleus and an electron cloud) was discovered. Nucleons are made up of protons (positively charged) and neutrons (neutrally charged). When basic particles combine to form an atom, a certain amount of mass is converted into the atom's binding energy, which is necessary to hold the nucleus together. Albert Einstein formalizes this relationship by the equation e = mc2, where "e" is the energy, "m" is the mass, and "c" is the velocity of light in a vacuum. This relation tells us mass can be converted to energy. Fission is splitting something into two or more parts, and nuclear fission is splitting the atom to generate a massive amount of energy, as illustrated in Figure 2.3.

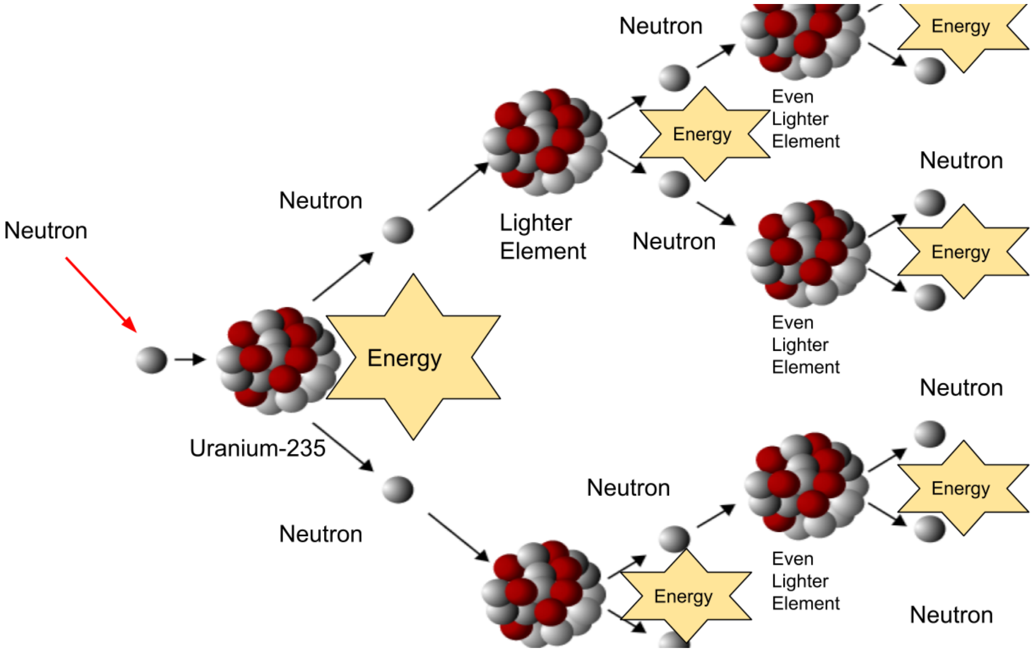


Figure 2.3 Nuclear Fission with neutron bombardment [20]

Fission in nuclear reactors is primarily caused by neutron bombardment. The neutrons produced by this fission reaction can trigger other fission reactions, potentially resulting in a self-propagating chain reaction. Nuclear reactors rely on these chain reactions. Scientist Leo Szilard developed the concept of a nuclear reactor utilizing fission chain reaction in 1933, and this concept was used in building the first atomic bombs later [24].

Albert Einstein, who fled Nazi Germany to the USA, learned that three chemists in Berlin were working on a weapon after using nuclear fission to split the uranium atom during World War II. The split released an abnormal amount of energy capable of powering a massive bomb. Einstein warned President Franklin D. Roosevelt with a letter and said, "a vast nuclear chain reaction involving uranium could lead to the construction of extremely powerful bombs of a new type, the atomic bomb" [21]. Einstein's letter inspired the USA to start the Manhattan Project, a research and development project to produce nuclear weapons (1939). The researchers developed the first atomic bomb in 1945, and two bombs were used on Japan. The first explosion immediately killed an estimated 80,000 people in Hiroshima (August 6,1945), and three days later, a second bomb killed around 40,000 people in Nagasaki. Japan's Emperor Hirohito announced unconditional surrender, citing the devastating power of "a new and most cruel bomb" in the history of humans on August 15, 1945. However, even after surrendering, tens of thousands more people continued to die because of radiation exposure.

Radiation is the transfer of energy through space away from a source. Ionizing radiation, generated through nuclear reaction, has sufficient energy to affect the atoms in living things, posing a health risk by damaging tissue and DNA in genes.

2.2.2. Peaceful Use of Nuclear Energy

During the Second World War, the use of the atomic bomb demonstrated the danger of nuclear weapons and their immediate and long-term consequences to the world. It led to the establishment of national and international organizations to benefit from nuclear energy [22].

President Dwight D. Eisenhower stood in front of the UN General Assembly in New York (1953) and delivered his famous "Atoms for Peace" address. In his speech, President Eisenhower said, "I feel impelled to speak today in a language that in a sense is new—one which I, who have spent so much of my life in the military profession, would have preferred never to use. That new language is the language of atomic warfare." [22]. He proposed the creation of an "international atomic energy agency set up under the aegis of the United Nations" to promote the peaceful uses of nuclear energy. Later in 1957, the International Atomic Energy Agency (IAEA) was founded, an international organization that seeks to promote the peaceful use of nuclear energy and inhibit the use of nuclear for any military purpose, like nuclear weapons or bombs [23]. The Nuclear Energy Agency (NEA), an intergovernmental agency, was established one year later. It facilitates cooperation among countries with advanced nuclear technology infrastructures to seek excellence in nuclear safety, technology, science, environment, and law. It has thirty-four member countries.

The USA Governing Legislation [25]

The fundamental laws governing civilian uses of nuclear materials and facilities are:

- Atomic Energy Act of 1954: the Atomic Energy Commission
- Energy Reorganization Act of 1974: The Nuclear Regulatory Commission (NRC) is established.
- Reorganization Plans,1980

Regarding nuclear waste

- Nuclear Waste Policy Act of 1982
- Low-Level Radioactive Waste Policy Amendments Act of 1985
- Uranium Mill Tailings Radiation Control Act of 1978

On non-proliferation:

- Nuclear Non-Proliferation Act of 1978
- Fundamental Laws Governing the Processes of Regulatory Agencies
- Administrative Procedure Act (5 USC Chapters 5 through 8)
- National Environmental Policy Act

Nuclear Nonproliferation Treaty (NPT)

The Treaty, which serves as the foundation for the non-proliferation and arms control regime, was administered by the IAEA and opened for signatures in 1968. One hundred ninety-eight countries are part of this Treaty. Five-member states are considered nuclear weapons states: the United States, the Russian Federation, the United Kingdom, France, and China. These members manufactured or exploded a nuclear weapon before 1967, i.e., before the Treaty.

Five States are not a party to the NPT: Israel, India, Pakistan, South Sudan, and North Korea (which withdrew in 2003). These countries do not accept the Treaty, meaning that they can develop nuclear weapons.

Today, nuclear, or radioactive material is used in medical, academia, industry, and generating electricity to benefit humanity. Radiation is also valuable for agriculture, archaeology (e.g., carbon dating), space exploration, law enforcement, geology (e.g., mining), and other fields [22]. This study focuses on generating electricity at nuclear power plants.

2.2.3. Nuclear Security

Nuclear security is the concept of the prevention and detection of and response to the theft, sabotage, unauthorized access, illegal transfer, or other malicious acts involving nuclear material, other radioactive substances, or their associated facilities [26]. The nuclear security regime includes various elements and activities, including legislation and regulation, intelligence gathering, assessment of threats to radioactive material, administrative systems, various technical hardware systems, response capabilities, mitigation activities, and associated locations and facilities.

The former Director-General of the IAEA, Mohamed El Baradei, said: “The gravest threat the world faces today, in my opinion, is that extremists could get hold of nuclear or radioactive materials” (2009). He was awarded the Nobel Peace Prize for his efforts to prevent military usage of nuclear energy and ensure its peaceful usage. Previous Presidents of the USA (George W. Bush and Barack Obama) stated that the gravest danger the country faces is nuclear terrorism.

Nuclear terrorism is described as acts of violence and destruction conducted with nuclear weapons or threats of using such weapons to instill fear, get attention, or blackmail. The Government of the United States describes Acts of Nuclear Terrorism as "*Knowingly and unlawfully possesses radioactive material or makes or possesses a device with the intent to cause death or serious bodily injury or substantial damage to property or the environment. Knowingly and unlawfully uses in any way radioactive material or a device, or uses or damages or interferes with the operation of a nuclear facility in a manner that causes the release of or increases the risk of the release of radioactive material, or causes radioactive contamination or exposure to radiation with the intent to cause death or serious bodily injury or with the knowledge that such act is likely to cause death or serious bodily injury, with the intent to cause substantial damage to property or the environment or with the knowledge that such act is likely to cause substantial damage to property or the environment; or with the intent to compel a person, an international organization or a country to do or refrain from doing an act*." [27]

To sum up, nuclear terrorism threats are sabotaging a facility that uses nuclear or radiological material, stealing these materials, fabricating these materials into a nuclear weapon, and stealing a nuclear weapon and detonating it.

The nuclear non-proliferation concept aims to prevent nuclear weapons and nuclear weapon technology from spreading, advance the goal of nuclear disarmament, and promote cooperation in the peaceful uses of nuclear energy. Article 3 of the Non-Proliferation Treaty requires each Non-Nuclear Weapon State to sign a safeguard agreement with the IAEA. The IAEA uses safeguards to protect nuclear installations and materials and tries to verify a State's legal responsibility that nuclear facilities are not misused and that the nuclear material is not diverted from peaceful purposes. States agree with restrictions by signing safeguard agreements [28].

States and member countries are primarily responsible for protecting individuals, society, and the environment from nuclear and radioactive sources within their territory. However, radiation risks may go beyond national borders, requiring an international authority. The IAEA promotes and enhances safety globally by exchanging experience and improving capabilities to control hazards to prevent accidents, respond to emergencies, and mitigate any detrimental outcomes.

2.2.4. Nuclear Power Plants

Nuclear power plants use nuclear reactions to generate electricity. They are classified as critical infrastructure, i.e., essential for the functioning of a society and the economy of countries. In the USA, 20% of the electric power comes from NPPs, while in France, it is over 70% [29] [30][99]. Nuclear energy has become more critical, especially with the increasing concern of climate change. It is a constant energy source, and it produces no greenhouse gas emissions during operations. In terms of carbon footage, constancy, and efficiency, NPPs are more robust than other alternative energy sources [1]. Figure 2.4 contrasts nuclear energy with other electricity sources.

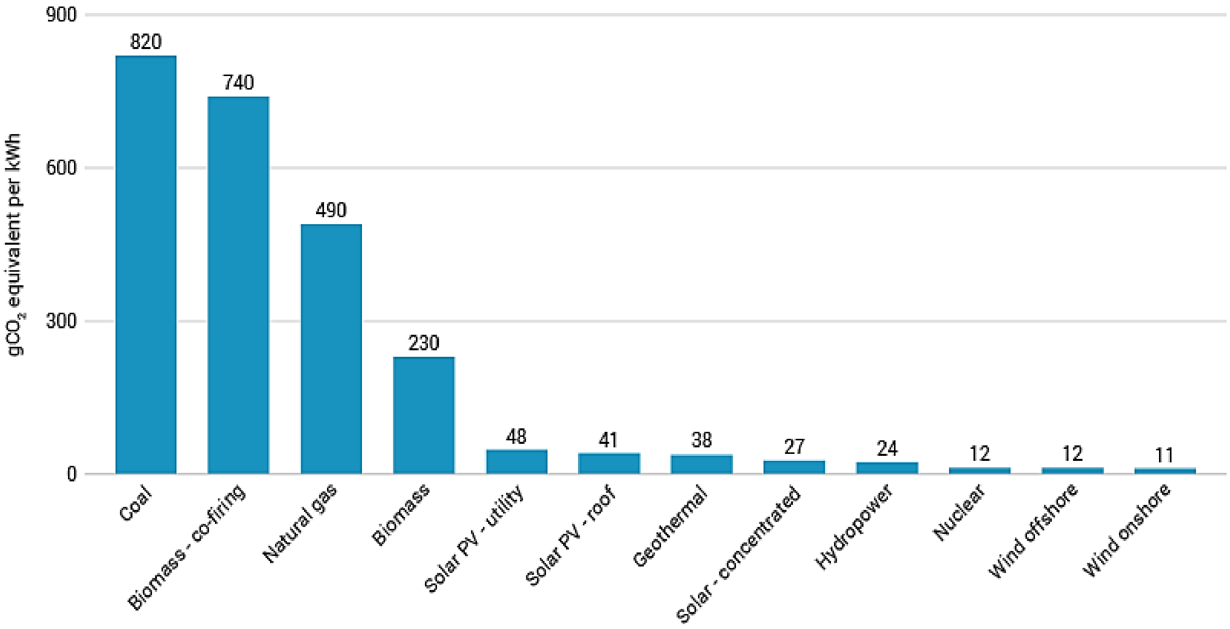


Figure 2.4: Average life-cycle carbon dioxide-equivalent emissions for different electricity generators [31]

2.2.4.1. Nuclear Reactors

A nuclear reactor was first used to generate electricity (sufficient energy to light a bulb) in 1951 in Idaho, USA. The first grid-connected NPP was built in Russia, and it operated from 1954 to 2002. Today, over four hundred plants operate in more than thirty countries, and many new plants are under construction.

A nuclear reactor contains and controls nuclear chain reactions, which produce heat via fission. This heat is utilized to create steam, and then steam is used to turn a turbine to generate power. A nuclear reactor consists of four fundamental systems and components: the fuel, the coolant, the moderator, and the control rods, as well as supplemental structures such as reactor pressure vessel internals and core support plates

[32]. Nuclear reactors are categorized based on their fuel, coolant, and moderator information. Figure 2.5 shows different types of nuclear reactors [32].

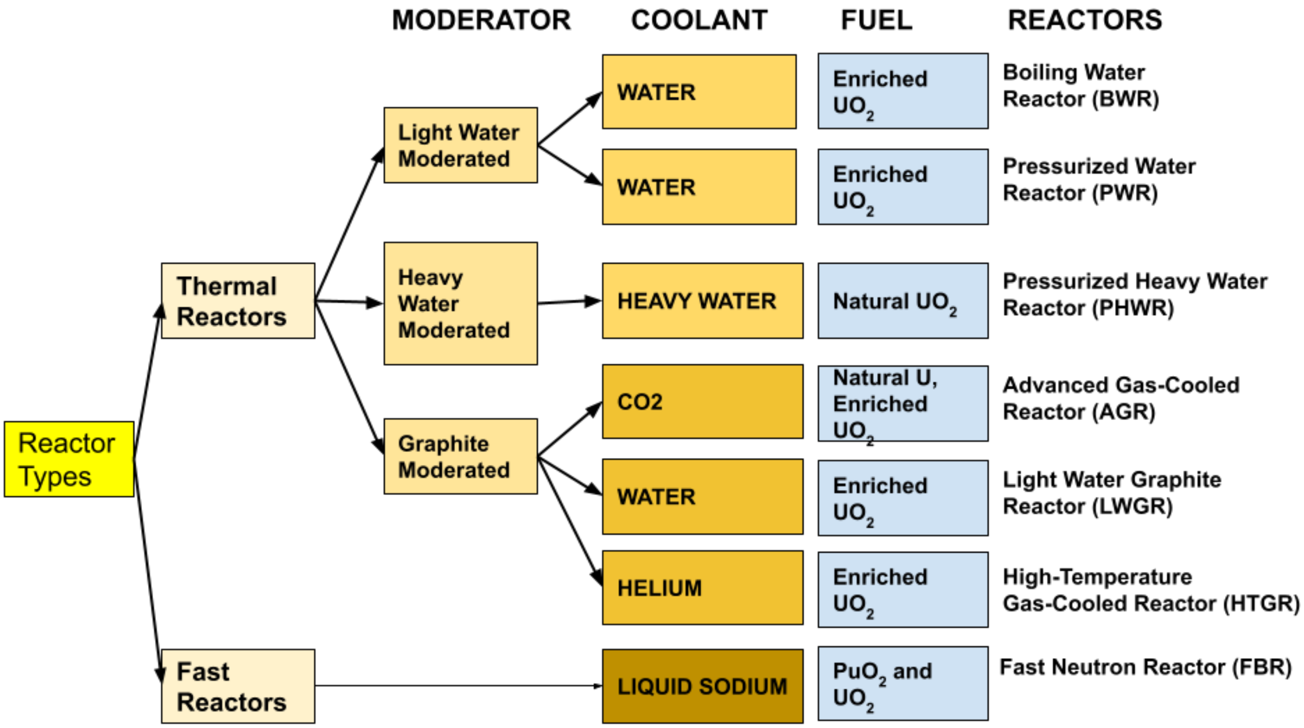


Figure 2.5 Categorization of nuclear reactors

2.2.4.1.1. Pressurized Water Reactor (PWR)

In the pressurized water type, the reactor vessel's core generates heat. The heat is carried to the steam generator by pressurized water in the primary coolant loop. The heated water moves in tubes and goes to the steam generator. Here, the heat vaporizes the water in the secondary loop and produces steam. The steam goes to the main turbine by steamline, and it turns the turbine generator to produce electricity. Water from the primary loop and water from the secondary loop never mix, so radiated water stays in the reactor area [33]. Figure 2.6 depicts this design.

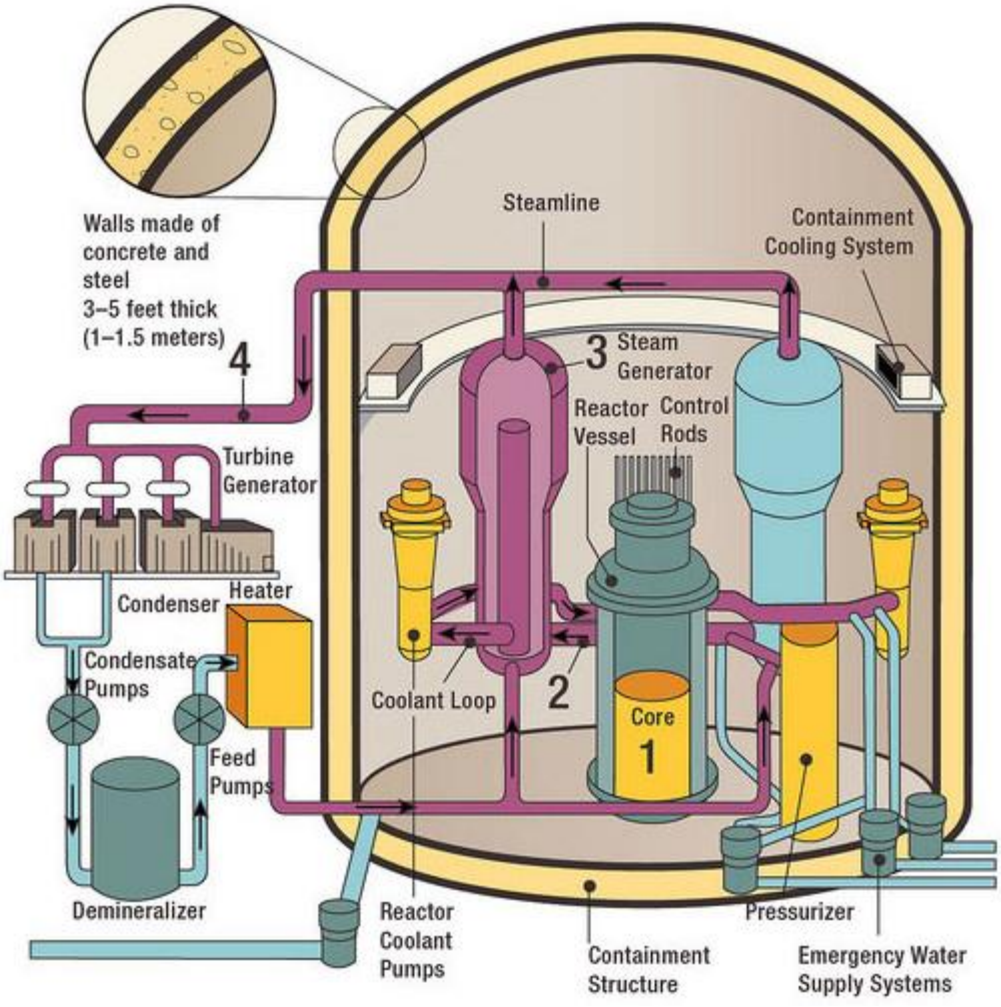


Figure 2.6: Pressurized Water Reactor [33]

2.2.4.1.2. Boiling Water Reactor (BWR)

In the boiling water-based design, the reactor vessel's core generates heat, and the water goes into the reactor vessel, producing a steam-water mixture. The mixture leaves the top of the core and steam, and water separation happens before steam is directed into the turbine. Water is converted to steam, then recycled back into the water by the condenser, to be used again in the heat process, as illustrated in Figure 2.7 [34].

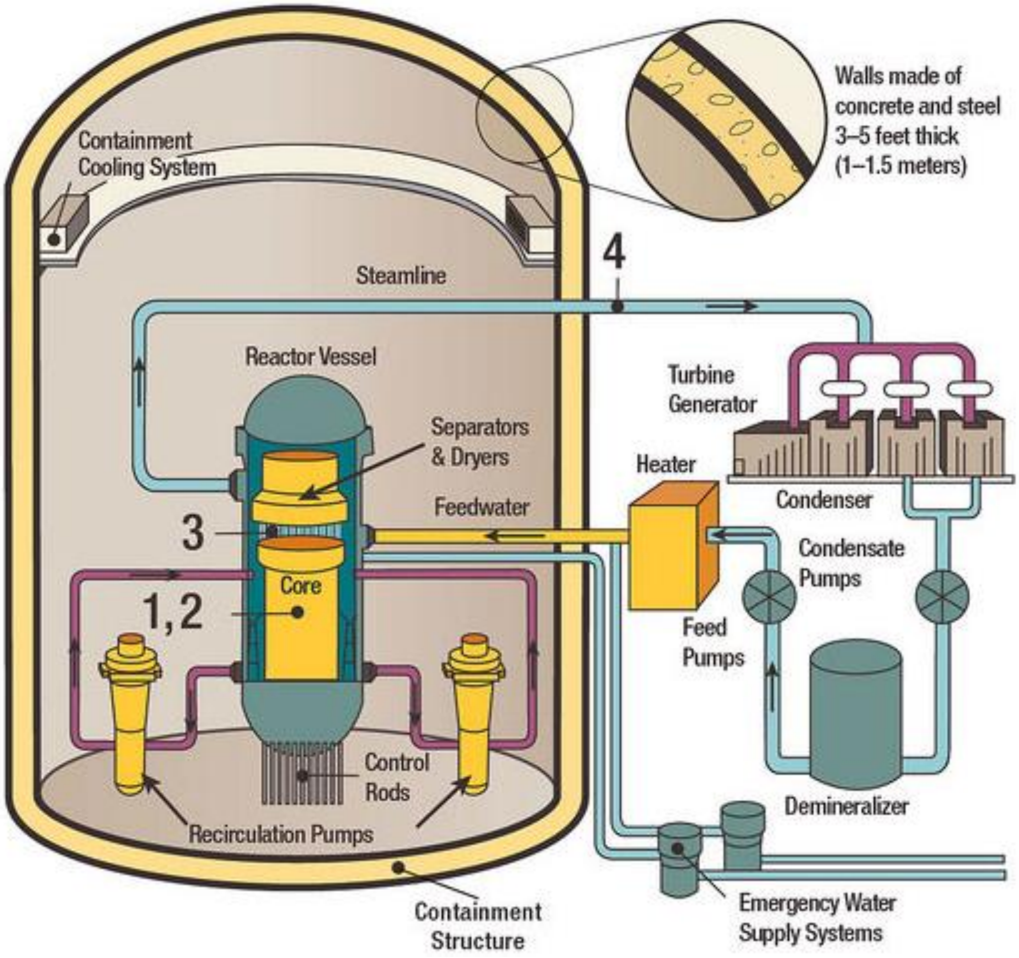


Figure 2.7: Boiling Water Reactor [34]

2.2.4.2. Physical Protection of Nuclear Power Plants

Nuclear power plants are significantly potent structures, and it is difficult to infiltrate them. NPPs adopt “Defense-in-Depth,” a concept used to design physical protection systems that require an adversary to overcome or circumvent multiple obstacles, either similar or diverse, to achieve his objective. Defense-in-depth is achieved by establishing and maintaining various security layers and a robust set of security controls [35]. The IAEA formulation of the defense-in-depth principle specifies five levels of defense surrounding the hazard source (e.g., concerning a nuclear reactor, first-level is cladding, second-level is reactor vessel, third-level containment building), as shown in Figure 2.8, which is designed based on INSAG-10[35].

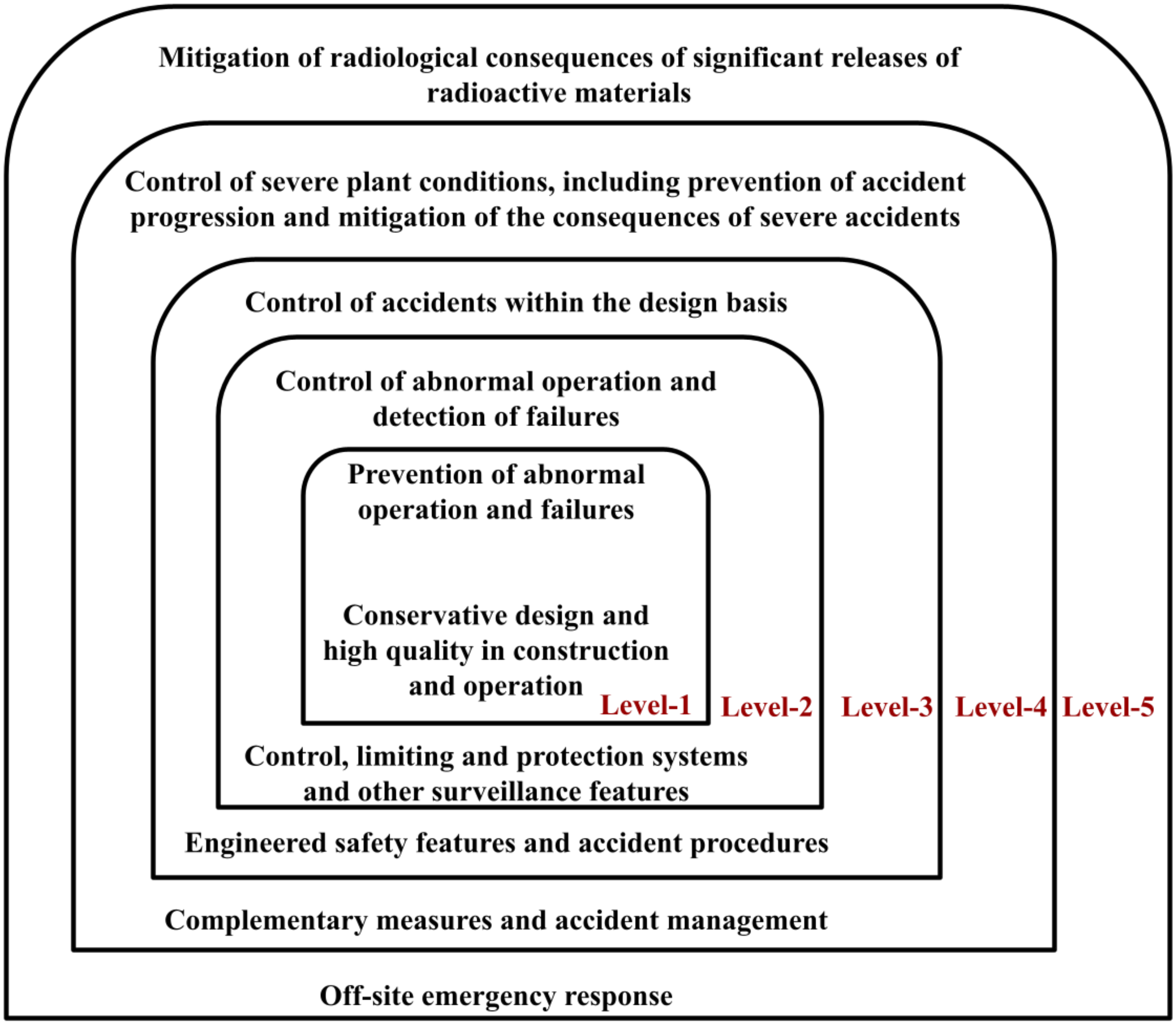

Mitigation of radiological consequences of significant releases of radioactive materials
Control of severe plant conditions, including prevention of accident progression and mitigation of the consequences of severe accidents
Control of accidents within the design basis
Control of abnormal operation and detection of failures
Prevention of abnormal operation and failures
Conservative design and high quality in construction and operation
Level-1
Level-2
Level-3
Level-4
Level-5
Control, limiting and protection systems and other surveillance features
Engineered safety features and accident procedures
Complementary measures and accident management
Off-site emergency response

Figure 2.8 Defense in Depth Approach in Nuclear Security

The IAEA Nuclear Security Series provides an effective physical protection system (PPS) using the defense-in-depth concept. The IAEA defines a PPS as an integrated set of physical protection measures, including people, procedures, and assets. These measures are implemented and sustained using management systems to prevent theft, sabotage, or other malicious acts [35].

PPS creates layered and concentric security areas (e.g., protected areas, limited access areas, inner areas) on identified targets. The key PPS functions are deterrence, detection, delay, and response.

Deterrence is the outer show of strength to try and discourage adversaries from attacking. Examples of physical security deterrence are large fences covered with barbed wire, visibly armed guards, or threatening signs such as “Use of deadly force authorized beyond this point.”

Another key PPS function is detection. Detection is the discovery of criminal or unauthorized acts with the help of an intrusion detection system, motion or infrared sensors, surveillance cameras, lighting, or alarms on doors. Once detection occurs, an alarm is initiated and reported, and assessment starts. An effective assessment helps to understand whether the alarm is valid or a nuisance.

Delay, another essential function of PPS, refers to how the security system hinders the adversary's progress toward its intended target. Locks, heavy doors, walls, and other barriers are tactics to delay an adversary from reaching the target. Delay elements can eventually be defeated, but the delay function is intended to provide time for response measures to be initiated before the adversary completes the malicious act.

The response function consists of the actions taken by security personnel to interrupt and neutralize the adversary from the fulfillment of any malicious act. Guards are part of the response force, but they can also serve several other functions like detecting illicit entry attempts and providing an immediate assessment. They can operate access control points and be trained to detect suspicious behavior in employees. They can also delay adversaries through engagement until more responders arrive, which can be significant if they are in well-secured positions. Their presence on site is also a deterrent to potential adversaries.

Recovery is the ability to resume normal operations after an incident or attempted incident, and it is part of the response function. It checks if normal facility functions are interrupted for a period of time, putting in place additional security measures to prevent future incidents.

Even though the physical security of nuclear power plants is a well-studied area, some cases showed that these methods do not ensure security. The next section covers a few such incidents.

2.2.4.3. Nuclear Security Incidents

Pelindaba, South Africa, 2007

Armed men cut through the chain-link fence surrounding the facility, which stored hundreds of kilograms of weapons-grade uranium, cut off the electricity and some alarms, and stormed the emergency response center holding one employee at gunpoint and shooting another. The intruders were eventually deterred by an (extremely late) response force, turned away, and were never caught[36].

Kleine Brogel, Belgium, 2010

In 2010, a security breach occurred at a Belgian Air Force facility that housed USA nuclear weapons. Six anti-nuclear activists entered the Kleine Brogel Air Base, and, before being arrested, the protestors placed stickers, took pictures, and lingered in the snow-covered base for around 20 minutes. In 2009, a similar issue occurred.[37].

Cruas, France, 2011

Two anti-nuclear activists broke through the Cruas Nuclear Power Plant's fence, evading detection for over 14 hours while uploading footage of their breach to the Internet [38].

Oak Ridge, Tennessee, USA, 2012

Three peace activists entered the perimeter of the Y-12 complex, which houses the facility that stores the US stockpile of highly enriched uranium. The activists marked the storage facility with graffiti and beat it with a hammer before eventually being apprehended by a single security official[39].

2.3. Cybersecurity for Nuclear Power Plants

Nuclear plants are one of the most secure industrial areas in the cyber domain. [40] [41]. Even though they are secure, they are not entirely isolated from attacks. Studies show that the number of cyber incidents at nuclear facilities is increasing [42] [43] [44]. More incidents may have occurred that have not been publicly disclosed or for which the details are classified or unavailable.

The Nuclear Regulatory Commission (NRC) describes a cyber-attack as: “The capability to exploit site computer and communications system vulnerabilities to modify or destroy data and programming code, deny access to systems, and prevent the operation of the computer system and the equipment it controls”[45]. For example, the Stuxnet attack (Iran, 2010) targeted specific process control system components, modified data, and destroyed equipment[46].

NPPs use digitalized systems to have accurate information, achieve better decision-making and efficiency, and manage the power plant's site. Using more digitalized systems might increase the cyber-attack surface. However, using less technology might cause problems too. Studies showed that some incidents that happened might have been avoided with the help of digitalized and automated systems. When Chernobyl (Russia, 1986) happened, the plant was under manual control, and the control computer was disabled [2] [3] [4] [5]. In Fukushima (Japan,2011), pressure relief valves on the wet well were manually operated, staying closed for hours. If the relief valves were automated, there would be no dependence on the operator's judgment, and the seven hours of hesitation to open the valves would have been avoided [5] [47] [48].

In conclusion, both cyber and non-cyber accidents or incidents showed that it is necessary to improve the technology and security measures used at nuclear facilities.

2.3.1. The Technology Used in Nuclear Power Plants

An NPP contains thousands of components and pieces of equipment that have to be operated in a well-coordinated way. This coordination is performed by instrumentation and control (I&C) systems. Detailed information can be found in sources such as [49][50][51].

The purpose of the I&C system is to deliver information to plant personnel and enable and support safe and reliable power generation by controlling the plant processes. The information must be accurate, sufficient, operationally relevant, timely, and dependable. Intervention, unauthorized modification, or delay on the information should be prevented.

In NPPs, components and equipment are divided into layers to facilitate the management of the plant site. Each layer has specific duties, and its protection mechanisms are implemented differently. Layers can be Field Level, Control Level, Supervision Level, Management Level, and Enterprise Level. Figure 2.9 illustrates the relationship between levels which is designed based on IAEA recommendations [35] [51]. Between the layers, communication can be established by different network layers and properties. For example, the corporate network can get information from the plant network but cannot send any data.

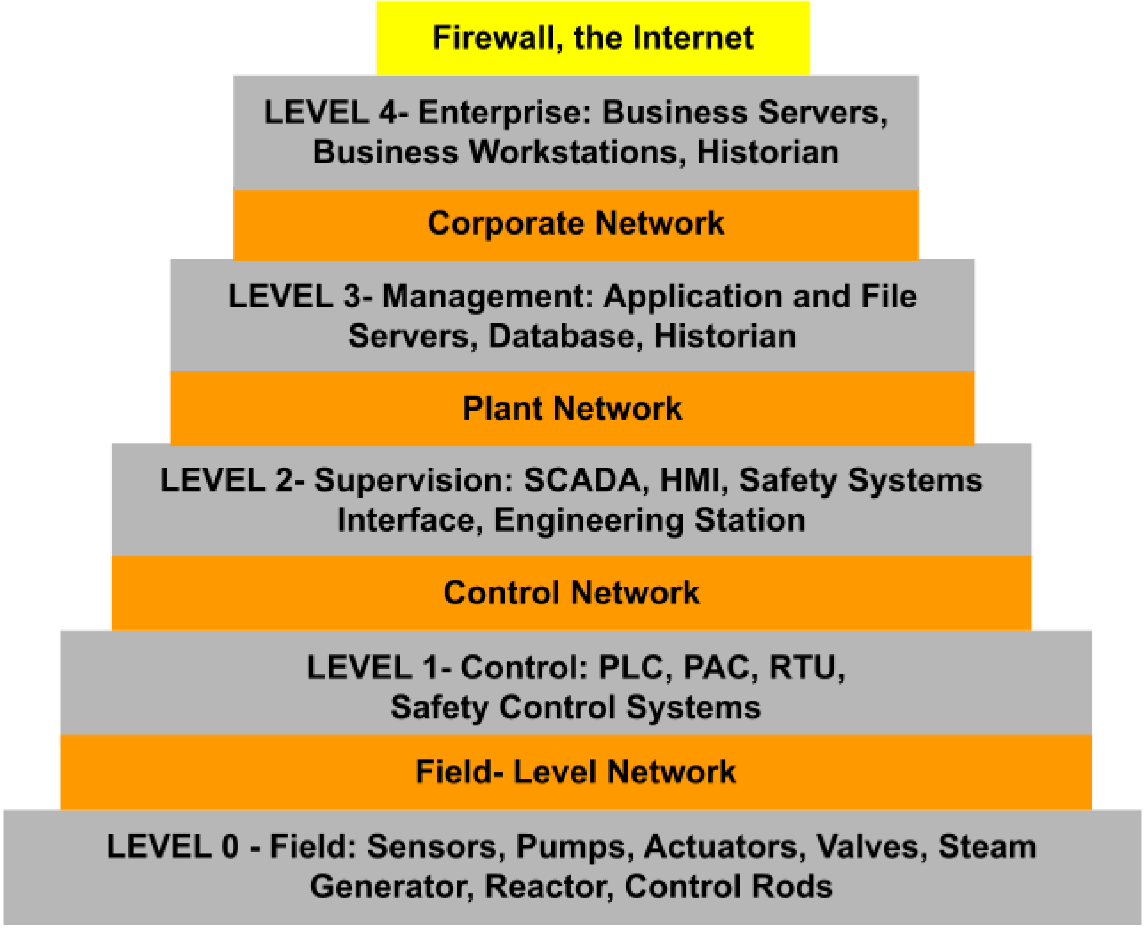

Firewall, the Internet
LEVEL 4- Enterprise: Business Servers, Business Workstations, Historian
Corporate Network
LEVEL 3- Management: Application and File Servers, Database, Historian
Plant Network
LEVEL 2- Supervision: SCADA, HMI, Safety Systems Interface, Engineering Station
Control Network
LEVEL 1- Control: PLC, PAC, RTU, Safety Control Systems
Field- Level Network
LEVEL 0 - Field: Sensors, Pumps, Actuators, Valves, Steam Generator, Reactor, Control Rods

Figure 2.9 Nuclear Power Plant Equipment Fields and Network Layers

Operational Technology (OT) is a software and hardware system that detects or changes data by directly monitoring or controlling industrial equipment, assets, processes, and events. OT includes Supervisory Control and Data Acquisition System (SCADA), Distributed Control System (DCS), Programmable Logic Controllers (PLCs), Computer Numerical Control (CNC), scientific equipment, building management and building automation systems (BMS, BAS), lighting control systems, energy monitoring, transportation, I&C systems, Remote Terminal Units (RTU), Internet of Thing (IoT) devices, embedded systems (e.g., smart instrumentations) and more. OT systems can be required to control engines, converters, and other machines to regulate various process valves such as temperature, pressure, and flow and prevent hazardous conditions.

SCADA is a software and hardware structure for gathering, processing, monitoring, and controlling real-time data and analyzing industrial processes. SCADA is not a specific technology or protocol but an application where data is collected from the plant site to control its systems.

Instruments sense conditions such as temperature, pressure, liquid level, power level, or flow rate. A sensor is a device that measures physical input from its environment and converts it into data that can be analyzed by either a human or a machine.

Operating equipment valves, pumps, motors, and actuators can be controlled automatically or manually to manipulate the systems. An actuator is a machine component responsible for moving and controlling a mechanism or system such as opening a valve. An actuator requires a control signal and a source of energy.

Local processors like PLCs, Intelligent Electronic Devices (IEDs), and Remote Terminal Units (RTUs) communicate with the site's instruments and operating equipment. Human-Machine Interface (HMI) communicates with PLCs and input/output sensors to get and display information for operators to view.

From data gathering at the plant's site, converting the data to human-readable form, processing and monitoring the data, every step has technology in it. As a result, plants are targets of cyber-attacks remotely or locally [42] [43] [44]. The next subsections discuss a few attack examples.

2.3.2. Technology-Related Accidents and Incidents

2003, USA, Ohio, Slammer Worm

A contractor (maintenance personnel not employed by the site) logged into an unsecured network, a breach occurred, and the worm spread from the business network to the plant network. It then found an unpatched Windows server. The plant's Safety Parameter Display System (SPDS) and Plant Process Computer (PPC) were disabled for several hours because of the worm's activity. The SPDS monitors the coolant system's operation, core temperature, radiation levels, and other critical conditions. Fortunately, the plant was not in operation because a hole in the reactor head was being repaired. Another reason there was no harm was that the worm could not attack the analog backups of the SPDS and the PPC. According to the reports, Microsoft released a patch to eliminate the MS-SQL vulnerability almost six months before this attack. However, the systems were not updated [52].

2006, USA, Alabama, Malfunction

The Browns Ferry NPP suffered a plant trip due to the Ethernet-based process control system overflowing the computer data traffic on the plant's internal control system network. The overflow caused the failure of non-safety-related reactor recirculation pumps and the condensate demineralizer controller. It happened during testing and maintenance time. This failure demonstrates that technologic incompatibility can lead to cyber incidents without necessarily involving any malicious intent [53].

2008, USA, Georgia, Update

In the Edwin Irby Hatch NPP, a computer that monitored data from one of the facility's primary control systems was updated. After the update, the computer was rebooted. The reboot also reset the data on the primary control system. This caused safety systems to erroneously read a drop in water reservoirs that cool the plant's radioactive nuclear fuel rods. As a result, automated safety systems triggered a trip [54].

2010, Iran, Natanz, Stuxnet

Stuxnet was a sophisticated attack targeting a specific process control system. It was accomplished via a supply chain attack and a thumb drive without any internet connection. The worm employed Siemens' default passwords to access Windows operating systems running WinCC and PCS7 control systems. These drives were used to power centrifuges used in the concentration of the uranium-235 isotope. Stuxnet altered the electrical current's frequency to the drivers, causing them to switch between high and low speeds. This switching caused almost one thousand centrifuges to fail in an unexpectedly brief time. Stuxnet gave instructions rather than interfering with the PLC, faking rather than disrupting sensor output. This type of attack requires extensive funding and intelligence from a state sponsor [46].

2011, USA, Oak Ridge National Laboratory, A zero-day attack

A lab employee got an email (a spear-phishing email) from the human resources department, directing the employee to click a link (malicious web page), where malware exploited an Internet Explorer vulnerability to download additional code to the victim's machines. The malware was capable of deleting itself if it could not compromise the computer. This attack explored a previously undiscovered vulnerability, i.e., it was a zero-day attack [55].

2014, Japan, Monju NPP, Malware

A software update on a computer at the plant introduced malware to the system. This malware sent data (staff training reports and emails) from the victim's machine to a command and control server that was in another country. The malware also attempted to gain access to a control room computer [56].

2014, South Korea, Hacking

Computer systems at South Korea's nuclear plant operator were hacked. The cyberattack was discovered after a hacker leaked nuclear reactor blueprints online and threatened more leaks unless the reactors were shut down. The hacker had access to reactor blueprints, floor maps, and other internal plant information and shared this information on a Twitter account [57].

2015, Ukraine, Hacking

Hackers attacked Ukraine's power grid, disabling control systems used to coordinate remote electrical substations, leaving people in the capital and western part of the country without power for seven hours. The Ukrainian hack was the first publicly acknowledged case of a cyber-attack successfully causing a power outage [58].

2016, Germany, Gundremmingen, Malware

Malware, including Conficker and W32 Ramnit, was discovered on several computers and removable media at the plant [59].

2017, USA, Burlington, Kansas, Hacking

The Wolf Creek Nuclear Operating Corporation's business network was breached. However, the business network and control network have no communication. Nothing happened at the plant site [60].

2017, Ukraine, Cyber Attacks

A series of powerful cyberattacks using the Petya malware hit websites of Ukrainian organizations, including banks, ministries, newspapers, and electricity firms. The attack also affected Chernobyl's radiation monitoring system. Similar infections were reported in many other countries [61].

2019, India, KudanKulam, Hacking

The NPP's administrative network was breached. A version of the DTrack RAT (Remote Administration Tool) virus had infected the administrative systems. The Dtrack malware can log keystrokes, scan IP addresses, list all available files and running processes, and retrieve browser history on target networks [62].

2020, USA, Hacking

Hacker gained access to the National Nuclear Security Administration systems that keep information on the USA nuclear weapon stockpile. Hackers accessed networks as part of an extensive espionage operation that has affected at least half a dozen federal agencies [63].

2021, Brazil, Ransomware Attack

The attack targeted the Eletrobras (Centrais Elétricas Brasileiras S.A.) power company. It hit the administrative network of its electronuclear subsidiary, which runs two NPPs, Angra1 and Angra2. The affected network was not related to the operational systems of power plants [64].

2.3.3. Incidents That Show Automation Is Needed

Three-Miles Island, Pennsylvania, USA, 1979

A failure in a non-nuclear section of the plant occurred. The failure prevented the main feedwater pumps from sending water to the steam generators, removing the reactor core heat. This caused the plant's turbine generator to be shutting down and the reactor to trip (shutdown). The pressure in the system began to increase and to control the pressure, the pilot-operated relief valve opened, which was located at the top of the pressurizer. The valve is closed when the pressure decreases. So, controllers closed the valve and believed the relief valve was closed (the indicator instruments showed closed). But the valve was stuck open. As a result, operators were unaware that cooling water was going out of the valve, and a loss-of-coolant accident was happening [5] [65] [66].

Chernobyl, Ukraine,1986

A test was conducted under manual control, and all automatic safety systems (emergency protection and emergency core cooling systems) were disabled. After a shift change, inexperienced and uninformed operators were operating the plant under manual control (the control computer was disabled) while their safety controls would have bypassed the accident [2] [3] [4] [5].

Fukushima, Japan, 2011

Pressure relief valves on the wet well were manually operated; they stayed closed for hours. If the system were automated, there would not have been the operator's judgment and 7 hours of hesitation to open the valves; it would have worked automatically. The venting level should decrease whenever pressure increases to a dangerous level (Hydrogen explosion). Moreover, since it was not automated, forgetting to close the pressure safety valve (PSV) meant releasing the additional radioactive vapor into the air from the building. The Fukushima operators did not know the water levels, water/steam ratios, temperatures, and the degrees of meltdown in their reactors or their spent fuel rod storage ponds. If there were well-designed automation, it would have read data accurately and acted quickly [5] [47] [48].

2.3.4. Cybersecurity Regulations, Law, and Policies for Nuclear Power Plants

In the USA, according to the NRC, there are currently 93 NPPs licensed to operate (62 pressurized water reactors and 31 boiling water reactors), which generate about 20% of the nation's electrical use [67].

The USA government has been taking technical security measures for the power plants since the 1970s. The increasing use of digital systems at power reactors pushed the nuclear industry to look for potential cyber threats in 1997. After September 11, 2001, the industry escalated its cybersecurity-related approach further. In 2003, The Nuclear Energy Institute (NEI) established a Cyber Security Task Force and developed guidance documents to help plants adapt cybersecurity programs. Until 2009, US plants were implementing some controls voluntarily; however, Title 10, Section 73.54, "Protection of Digital Computer and Communication Systems and Networks" of the Code of Federal Regulations (CFR) requires that “each licensee currently licensed to operate a nuclear power plant under part 50 of this chapter shall submit a cyber security plan that satisfies the requirements for Commission review and approval" [27].

The NRC issued mandatory cybersecurity requirements at plants for safety, security, and emergency preparedness and published the Regulatory Guide (RG-5.71) in January 2010. The NRC conducts inspections to ensure that licensees are implementing the cybersecurity programs. Still, the NRC does not mandate any cybersecurity program or method but asks about licensees' plans and checks if the plan meets the requirements or not [68]. In 2010, the NEI published a Cyber Security Plan for Nuclear Power Reactors (NEI-08-09). NEI 08-09 describes a defensive strategy that consists of defensive architecture and a set of security controls that are based on the NIST SP 800-82, "Guide to Industrial Control System Security," and NIST SP 800-53, "Recommended Security Controls for Federal Information Systems," standards. By the end of 2013, all plants had completed the first step of their NRC-approved cybersecurity plans, and they needed to complete all expected steps by 2017. CISA updated the Nuclear Sector Cybersecurity Framework Implementation Guidance in May 2020. The Framework offers a flexible way to address cyber security. It can be applied to organizations depending on technology, including IT, ICS, CPS, or connected devices, more generally, including the IoT. It enhances, not replaces, an organization's risk management process, cyber security program, or related framework implementation. Every organization must decide how to implement the Framework specifically to its technology. The Framework can aid organizations in addressing cyber security and assist suppliers that perform physical work on mission-critical equipment (e.g., software updates, firmware replacement, equipment maintenance, refurbishments, and replacements) [19] [59].

2.4. Previous Work For Cybersecurity of Nuclear Power Plants

Some previous work emphasized that automation and digital devices introduce vulnerabilities to nuclear power plant (NPP) systems. NPPs have a vast amount of data, and working with data analytics and machine learning tools could help detect attacks and find abnormalities in the systems [49] [70] [71] [72].

Lee et al. proposed a security measure that detects threats based on the pattern analysis. Security Information and Event Management (SIEM) is a log management system that functions on storing, analyzing, and deleting logs. Lee et al. suggest extracting data with the help of SIEM and using Watson (an artificial intelligence model developed by IBM) can find unknown attacks [73].

Zhang et al. created a cyber-attack detection system that uses supervised and unsupervised machine learning methods on both cyber data and process data. The system was evaluated using data acquired from a real-time testbed with a physical flow-loop facility and a control system subjected to multiple cyber-attacks. The results suggest that the system is capable of detecting cyber-attacks [74].

Lee et al. proposed and developed a network traffic analysis system for a nuclear power plant. The study worked with the TensorFlow big data analysis method. Since quick and accurate detection is crucial for nuclear power plants, studies focused on finding machine learning algorithms that give the highest accuracy and shortest real-time analysis time on data [75].

Allison et al. developed an application that analysis PLC data (where raw data is converted into digital data) to function as a final defender for NPPs[76].

3. CHALLENGES IN SECURING NUCLEAR POWER PLANTS

In Chapter 2, several nuclear power plants (NPPs) incidents were listed, highlighting how advancing NPP's cyber security and automation technology can increase NPP security. This chapter discusses the challenges introduced by the participation of human personnel in the operation and management of NPPs, geopolitical factors, and supply chains.

3.1. Human Factor in Cybersecurity Protection of Nuclear Power Plants

The International Atomic Energy Agency (IAEA) guidebook, Nuclear Power Plant Personnel Training and Its Evaluation, states: "The objectives of safety and reliability cannot be achieved solely by the quality of equipment and hardware, but depend critically also on sufficient numbers of personnel having the necessary qualification and competence to conduct their tasks and responsibilities." [77].

Personnel plays a vital role in the safety and security of NPPs. Operators monitor and control the plant to ensure it is functioning correctly. Test and maintenance personnel help to confirm the equipment is working as expected and restore components when malfunctions occur. Security personnel operates at checkpoints (e.g., personnel entry control, identity check, body check by hand or equipment), following verification procedures consistently, regardless of the person's position or authority.

On the other hand, even the most well-structured secure systems can be compromised because of personnel mistakes. Intentional or unintentional mistakes include leaking information outside of the facility, plugging personal devices into the plant's network, and taking wrong steps (e.g., acting late to report). These mistakes might cause damage at plants [2] [5] [46] [52] [65]. The Nuclear Threat Initiative article states that "security is only as strong as its weakest link" [78]. Non-expert personnel without proper training might be the weakest link even in the most protected and secured nuclear plants. So, training people for their job and regularly reminding them of the significance and sensitivity of their position may have crucial effects on the safety of NPPs.

One of the biggest threats to a nuclear power plant might be an insider's violent, irrational, hazardous act. Insider threats can be internally motivated or externally forced. Threats can be passive and active, and active threats can be violent or non-violent. For example, a person may be unwilling to use force against personnel, or an operator may not report things properly. On the other hand, a violent insider willing to use force against personnel might be irrational or rational. Rational who are not willing to use their life, if they have a problem, they will probably return themselves. Nevertheless, irrational people might kill themselves while harming the systems if they are caught. There are many possible motivations for insiders to jeopardize security, including political, ideological (e.g., fanatical conviction), financial, personal (e.g., revenge, disgruntled employee, or customer), ego (e.g., hackers showing off), mental illness, and coercion [79].

Insider threats can be prevented by monitoring employees and contractors in real-time using access and system logs and evaluating personnel through background checks. Also, under the situations where personnel play prominent roles (e.g., switching a component on and off, changing a value), additional control mechanisms like machine input or another person’s permission might reduce the insider threat.

3.2. Nuclear Power Plant Supply Chain

Power plant operators and managers rely on chains of suppliers of both products and services to produce nuclear energy. These suppliers deliver products and services in all stages of a reactor’s life cycle: design, construction, commissioning, operation, and decommissioning [80]. In recent years, nuclear power plants’ construction and operation have experienced difficulties related to their supply chains. There have been delays in projects and even temporary shutdowns of reactors due to the detection of counterfeit items, technological obsolescence, and licensing to incorporate a greater amount of digital instrumentation and control technologies. Some original equipment manufacturers have left the marketplace in some countries, while operators in other countries have encountered difficulties placing up new, localized supply chains [80]. Different national regulations, standards, and legislation highlight the challenge of coordinating the nuclear supply chain and establishing a worldwide framework that allows the use of high-quality nuclear components. The reliability of supply chain products and services always must be questioned, and quality assurance and police requirements must be satisfied.

Today, some countries rely entirely on other countries to build nuclear power plants. For example, Rosatom, also known as Rosatom State Nuclear Energy Corporation, is building more than thirty new reactors in different countries (e.g., Finland, Hungary, Bulgaria, Uzbekistan, Turkey, and Bangladesh). Some of these countries are entirely new to the nuclear industry, and they do not have nuclear regulations or trained personnel who can work in plants; there is no way these countries can conduct proper inspections on the design and construction phases of the nuclear power plant. The IAEA can inspect independently, but this might not be enough to trust the construction country. For example, in 2015, there was some tension between Turkey and Russia due to a Russian fighter jet being downed by Turkey. Despite the tension, NPP construction continued [81]. It is often not possible to assume trust in the presence of geopolitical risk. A technology provider may implement backdoors, logic bombs, or eavesdropping devices into instruments in the power plant.

3.3. Modernization of Nuclear Power Plants

The average lifespan of NPPs is around forty years [65]. However, instruments and other components have shorter life expectancies. Analog integrated circuit (IC) and measurement systems have operated satisfactorily, but today reactors face challenges due to aging and the obsolescence of components and equipment. Obsolescence is a significant concern due to the lack of spare parts, supplier support, and functional capabilities needed to satisfy current and future policy needs. The aging of the IC systems is another concerning problem that leads to difficulties such as reduced reliability and availability, growing costs to maintain acceptable performance, and the lack of qualified maintenance and engineering personnel. Modernizing or replacing the analog IC systems might be more beneficial (i.e., cost-effective, dependable) than maintaining or increasing the reliability of those systems. In particular, the need for greater reliability and availability may require the capabilities of modern technology that are not possible or applicable with older technologies. Modernizing or replacing old technology provides the opportunity to improve facility performance and human-machine interface functionality, enhance operator performance and reliability, and address hardships in finding professionals with education and experience with older analog technology [50].

Even though digital technology, which includes software-based control logic, provides a flexible, scalable solution to control requirements, introducing programmable components into a system might bring some problems. For example, verification and validation are crucial for plant data. Experience has shown that digital systems need considerable effort to function correctly and not display unintended functionality in any operational mode [50]. Due to licensing and the expected quality assurance, IC systems generally use proprietary equipment with proprietary software; new software and equipment are becoming standardized. This means that the knowledge and experience required to attack IC systems are becoming less specialized, lowering the bar for an attacker to attack NPP industrial control systems successfully.

Moreover, modernization might require retraining operating and maintenance staff due to the introduction of recent technologies. If the change is extensive between the old and new systems, staff might be reluctant to train and learn. In this case, plant managers might consider hiring people with needed skills. However, another problem might arise. The hiring process might take time due to background checks and minimum nuclear-related training of employees (people should be informed about the danger of nuclear security threats and the sensitivity of their job).

The increase in digital technology at NPPs also increases the demand for technology experts. Hence, the development of flexible full-scope plant simulators and well-structured training courses should be prioritized. Experts (e.g., penetration testers, security specialists, cybercrime investigators, security consultants, security engineers, chief information security officer, IT, network specialists) have separate roles and responsibilities in the cyber security domain. Their education and training methods are different, and they are specialized in different fields of cyber security. For these professionals, currently, there are no NPP simulators or proper training mechanisms to understand the structure of NPPs. Information about networks, devices that are used in structures, communication methods between different zones and layers, upgrade and installation policies, SCADA, PLC, HMI, plant-specific devices, maintenance procedures, monitoring processes, and asset protection are not part of training.

3.4. Nuclear Power Plant Cyber Risk Assessments

The common age for a nuclear reactor is between 35 to 60, and there are many active old nuclear plants. Their risk assessments and plans are primarily about physical protection; cybersecurity risk is not part of the assessments. Modernization (change in technology) at old plants introduces new risks and requires updates in risk assessments of NPPs [51][82].

Additionally, new plants must have risk assessment plans, include cybersecurity risk starting from the planning level, and continue to re-assess the risk as long as the plant is active.

NPPs should have and maintain plans that can cover all the requirements of organizations such as NIST, NRC, and IAEA. For example, regulations may ask for the division of network layers and separation of plant and corporate networks, but they do not suggest or mandate specific separation methods. Every plant has a unique design and technology, and there are no general rules or approaches for calculating and considering the cyber security risks. As a result, calculating risk and including cyber risk into the risk assessment is not an easy task.

While estimating the cyber risk, especially for critical infrastructures like nuclear power plants, creating and analyzing worst-case scenarios is essential. Worst-case scenarios for an NPP include attacks performed with the help of an insider (e.g., employees, inspectors, third-party consultants) and damage to the core building that causes external radiation exposure (e.g., to people, to the environment).

The challenges include determining the risk calculation method, deciding what is critical, identifying vulnerabilities, and how dangerous they are [15].

3.5. Nuclear Power Plant Simulators

Previous sections indicated that training plant employees is one of the most crucial parts of cyber protection for nuclear plants. Nevertheless, training programs are limited. Nuclear facilities do not share information about plants' structure (e.g., instruments, hardware-software systems) and data due to safety and security concerns. However, the IAEA has established several different nuclear power simulators [83][84] to support human resource development in nuclear facilities. These simulators mostly mimic the reactor and provide a general idea of its working principles.

In this study, we examine NPP simulators to assess how effective they can be in training cybersecurity and information technology (IT) users who lack domain knowledge in nuclear power plants. We examine the currently available plant simulators to study how dependable they are for training people. Furthermore, we investigate the simulators' capabilities to produce datasets that can be used in security data analysis.

The IAEA has nuclear power plant simulation software available that simulates the behavior of the following reactor types [83]:

- Advanced PWR: Two-Loop Large PWR (Korean-OPR 1000)
- Russian-type PWR (VVER-1000)
- Advanced Passive PWR (AP-600)
- Integral Pressurized Water Reactor (SMR)
- Conventional Boiling Water Reactor with Active Safety Systems (BWR)
- Advanced BWR with Passive Safety Systems (ESBWR)
- Pressurized Heavy Water Reactor (PHWR)
- Conventional Pressurized Heavy Water Reactor (PHWR)
- Advanced PHWR (ACR-700)
- Micro-Physics Nuclear Reactor Simulator

These simulators, designed to run on desktop computers, provide insights and an understanding of the reactor types' designs and operational characteristics. The scope of the simulator programs is limited to providing general response characteristics of specified types of nuclear reactor systems. They are not intended for plant-specific purposes such as design, safety evaluation, licensing, or operator training. These simulation tools do not include functionality to capture and save data or change simulation parameters or data values during execution. Therefore extracting, analyzing data, and conducting experiments with data models is not feasible with these simulators.

Moreover, the simulators are not useful for people responsible for technical areas such as network security, system maintenance, equipment monitoring, and cyber protection. Simulators focus on core buildings, not capturing aspects related to network or monitoring devices. People with technical backgrounds cannot use the simulations to build a base to understand the potential problems and offer solutions to technical problems. In summary, these simulators are not appropriate for training personnel responsible for networking and control technology. The problem that arises from here is how to train and educate IT professionals, cyber security experts, network engineers, or other technical professionals. It can be said that developing a virtual machine mechanism or a new simulator that mimics the nuclear power plant site's SCADA system or hardware-software systems and the reactor's working mechanism is needed.

3.6. Summary

This chapter discussed several aspects of the difficulties in securing NPPS:

- ✓ Humans are often the weakest link.
- ✓ Supply chain management is crucial to protect infrastructure. Hardware-software solutions from suppliers should have quality and reliability assurance.
- ✓ Modernization at plants should be planned ahead of time, taking into consideration the management of human resources and technology compatibility.

- ✓ Implementing a risk assessment plan or cybersecurity framework does not imply that systems are secure. How risk is calculated and managed is vital for NPPs.
- ✓ There is a need for virtual machines or simulators that can be used to practice and learn technical requirements (e.g., networks using router instances or bridged connections, update management, map of devices) for a nuclear power plant. Current plant simulators cover only nuclear reaction monitoring.

4. ATTACK SCENARIO

Nuclear power plants (NPPs) are closed, air-gapped systems, so hacking into the critical internal infrastructure via network attacks seems unlikely. However, as discussed in Chapter 2, incidents such as hacking into the administrative network, infecting systems with worms, and causing damage to the plants' site with malware (e.g., Stuxnet) showed that plants are not entirely isolated from attacks. Moreover, cybersecurity is not only about intended attacks from an adversary via the Internet. A system failure is also subject to nuclear cybersecurity. During maintenance, inspection, or when the plant is under normal operation, a system failure (possibly resulting from an attacker act) may cause an incident.

There are public domains such as the MITRE ATT&CK [85]) that can be used to learn adversary tactics and techniques and develop attack scenarios. In this study, we assume that the attacker's goal is to infiltrate the system, spread through the networks until finding a specific system property, modify the related data slightly, and remain in the system to use this attack in the future. The question is how to reach the plant site, perform the attack and make changes in the system to cause damage while remaining unknown?

4.1. Information Gathering with the Help of an Insider

If the attacker has a specific target (e.g., nuclear power plant), he needs to identify the target's systems, such as network layers, firewalls, operating systems, hardware systems, system properties, and the programming languages that have been used. The attack will be based on this information.

A possible attack scenario is to find an employee to get information about the plant, use the information to develop the attack, and then perform the attack with the employee's help (as depicted in Figure 4.1).

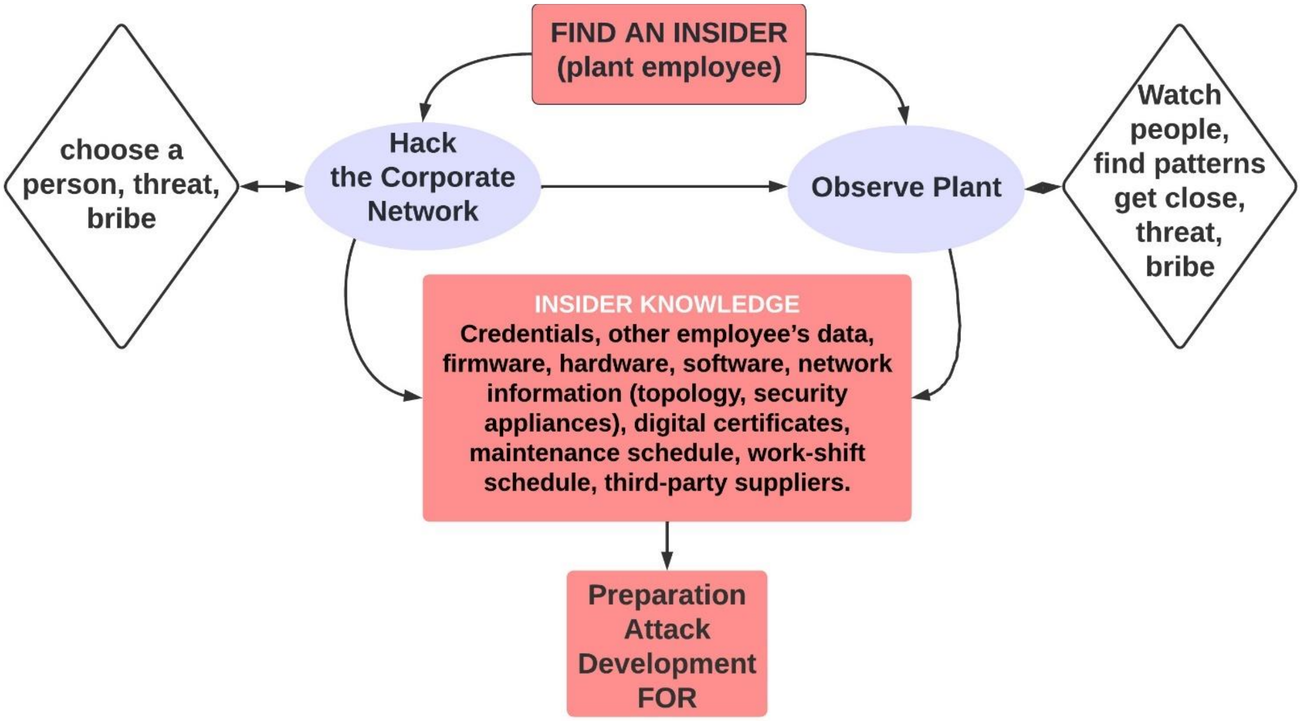

FIND AN INSIDER
(plant employee)
choose a person, threat, bribe
Hack the Corporate Network
Observe Plant
Watch people, find patterns get close, threat, bribe
INSIDER KNOWLEDGE
Credentials, other employee's data, firmware, hardware, software, network information (topology, security appliances), digital certificates, maintenance schedule, work-shift schedule, third-party suppliers.
Preparation Attack Development FOR

Figure 4.1 Finding plant employee to get information

The attacker may hack the corporate network to obtain the plant's design information, employee information, or any piece of information that might lead the attacker further to the next step. In some nuclear power plant attacks, it has been shown that attackers got plant blueprints, floor plans, employee data, and emails [54] [55] [56] [57] [58].

If the attacker could not reach employee information from hacking attempts, s/he can investigate people who work at the plant by watching the workplace. Nuclear plants are built in isolated locations, but their sites are known to the public. The attacker can find the people by observing the workplace and tailing the people or vehicles to reach employees and learn about them.

Once the attacker has the employee's information, s/he can choose a victim (future insider). After deciding on the victim, the attacker might get information by threatening (for example, the victim's life or a family member's life) or obtaining the victim's agreement (for example, through a bribe) to get additional information and take advantage of the insider in future steps.

4.2. Attack Development and Initial Access

The attacker used the victim's knowledge and developed an attack tool. The tool is developed based on an initial access method, as shown in Figure 4.2.

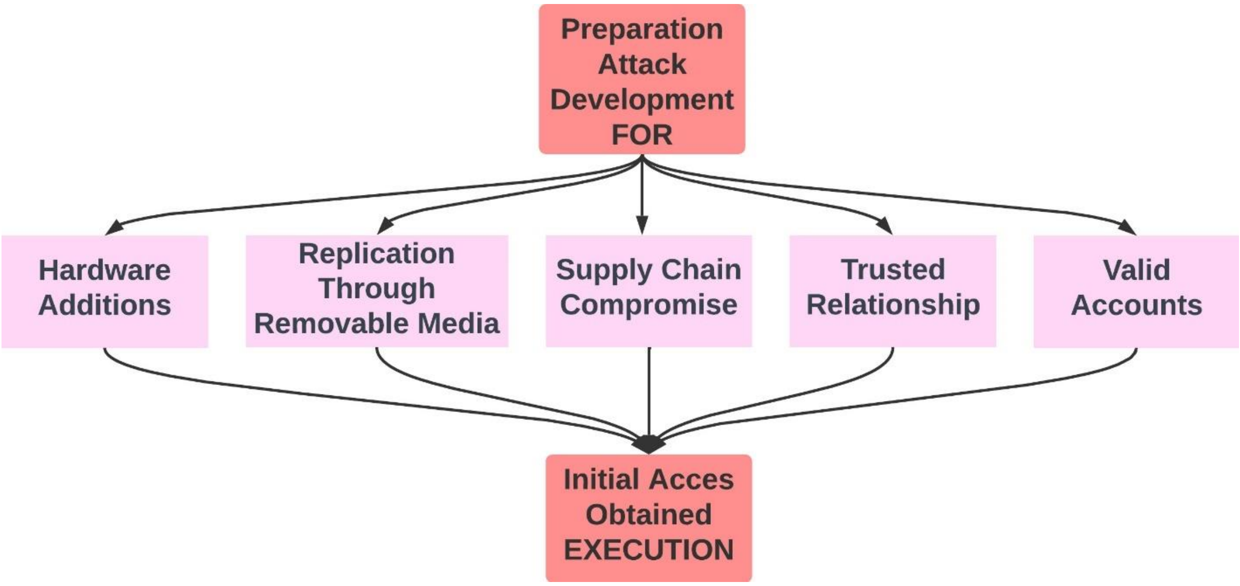

Preparation Attack Development FOR
Hardware Additions
Replication Through Removable Media
Supply Chain Compromise
Trusted Relationship
Valid Accounts
Initial Acces Obtained EXECUTION

Figure 4.2 Attack development and initial access

A hardware addition such as a computer accessory or networking hardware can be developed to enter into a system or network to gain access (e.g., man-in-the-middle attack, eavesdropping).

The attacker may create malware for removable media (such as USB) that runs automatically (with the help of manual manipulation of the media, modification of systems used to initially format the media, or modification to the media's firmware) when it is inserted into a system.

The attacker can take advantage of supply chain products to compromise the plants' security. The attacker might introduce malware to the supply chain’s products (such as development tools, software, and testing mechanisms). This attack can occur at any phase of the power plant management, such as during construction, modernization, or maintenance. Especially in case of changes in suppliers or an urgent need for a supplier service, plant managers might not care about the security of services during maintenance or modification since replacing the hardware/software service to continue production processes is their priority. In this case, the attack would be integrated into the systems with fewer security inspections and with the help of the supplier (for example, during the installation of a new device or maintenance service). The attacker can further attack scale from now on.

The attacker may hack or be part of an organization (third-party) that has access to the power plant or force (threatening) organization personnel. Third-party services may vary as IT services contractors, security providers, infrastructure contractors (e.g., elevators, physical security, HVAC (Heating, Ventilation, and Air Conditioning)).

Access through a trusted third-party may not be protected or receive less inspection than standard access mechanisms. Even though NPP policies are strict, the nature of humans (for example, trusting strangers who present themselves as elevator mechanics) may lessen the security procedures or make people less careful, unintentionally helping the attacker.

The attacker might use valid accounts to access high-level credentials (e.g., network, IT administrator) and bypass administrative access controls. Compromised credentials via valid accounts could give the attacker additional access (e.g., privilege escalation) to specific systems or restricted network locations. The attacker may create a remote access permit to modify protection measures or information with this additional access.

The malware must be executed at least once to infect the system. Most malware tries to install itself on the system during the initial execution. Once the malicious code is installed on the system, it needs to be executed again to perform malicious activities.

4.3. Execution of Malicious Code

The attack execution is the process of running (after initial access) the attacker's malicious programs on the victim's system. The execution can be triggered locally or remotely. Once the attackers' methods are discovered, they cannot run the same codes anymore; as a result, attackers focus on improving and finding new execution methods to run their malicious codes successfully. Execution techniques are becoming more sophisticated, and they are developed as platform-specific or platform-independent execution techniques.

There are many known execution methods, vulnerabilities, and their patches (fixes); however, due to weak update/upgrade maintenance, vulnerabilities are not eliminated from the systems, so attackers can still use known execution techniques. For many large enterprises (e.g., power grids, NPPs), updating and upgrading processes are slow. The enterprises' priority might be non-stop energy generation and answering the demand, so they may choose to extend periodical maintenance time and remain vulnerable. Previous incidents, as discussed in Section 2.3.3, show that patches were already available when the attacks infected the systems. As depicted in Figure 4.3, attackers use some known techniques such as exploiting vulnerabilities in system services, Windows management tools, software development tools, user privileged execution, and shared modules.

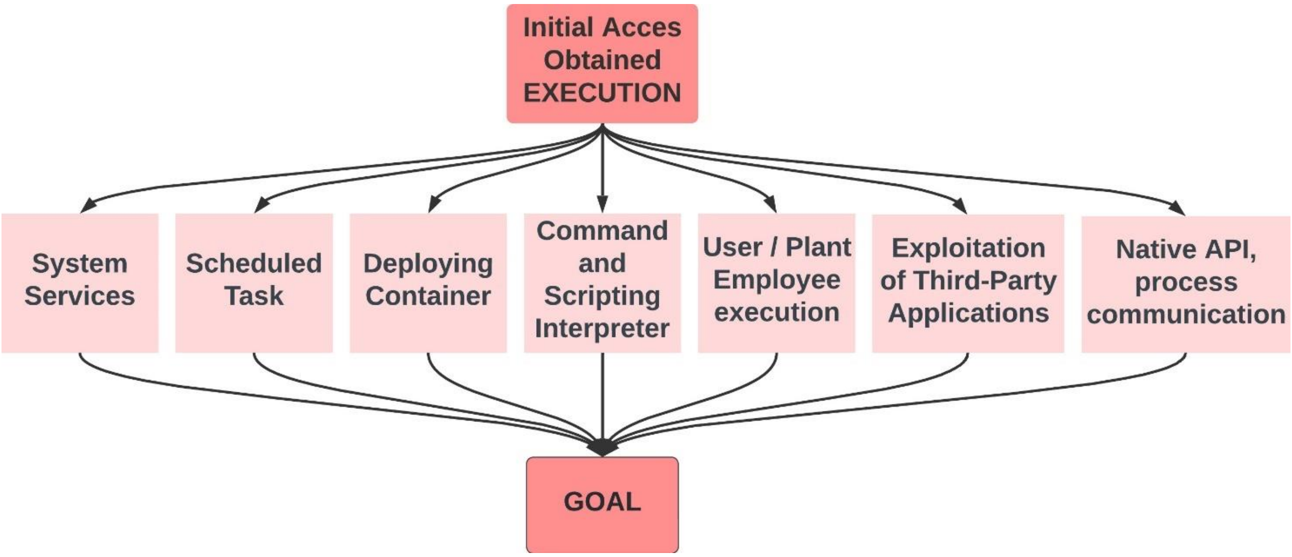

Initial Acces Obtained EXECUTION
System Services
Scheduled Task
Deploying Container
Command and Scripting Interpreter
User / Plant Employee execution
Exploitation of Third-Party Applications
Native API, process communication
GOAL

Figure 4.3 After initial access, execution of malicious code

The goal of the attack will affect the execution method. If the goal is immediate damage, some attack aspects, such as preventing detection, using encryption, selfreplication, and self-modification, may not be necessary. For example, to cause damage to a reactor core building’s appliance, or the monitoring system, the attack can be developed to affect only specific critical devices such as temperature-pressure sensors. The attacker can use removable media or any other additional hardware to integrate his attack into the critical device and cause damage with an insider's help.

On the other hand, the goal of infiltrating systems, creating remote access, gathering information, extracting, or modifying data, and preventing detection for future use would require a complex and tricky attack.

The attacker may use execution methods such as deploying containers, exploitation of third-party applications, and software deployment tools to access and use third-party software to move further through the network. The access may be used to laterally move to other systems, gather information, or cause a specific effect and prevent defense measures. With an insider’s help, the attacker can combine many techniques to avoid detection and spread through the entire system to wait for future users’ input or action.

Recently, the concept of hacking back is a popular subject of cyber security. In case of a cyber-attack, the victim looks for clues and traces back to the attacker, and tries to create an opportunity to hack back. The idea of hacking back might make attackers consider tradeoffs before attacking; however, it might make things more dangerous too. The attacker might have more hidden passive attacks, and fighting back might accelerate the coming attacks. Construction models such as Build-Own-Operate (BOO) and Build-Operate-Transfer (BOT) might be dangerous for nuclear power plants, but they have been accepted by countries (Finland, Turkey, Russia)[81][86]. As George Catlett Marshall Jr. said, “In any event it goes to prove that the friend of today may be the enemy of tomorrow” [87]. The vendors might have hidden back doors or logic bombs implemented into the power plant systems with the idea of hacking back or future use, which cannot be avoided or understood by the customer.

4.4. Attack Goal and Attack Finalization

The attacker's goal affects every phase of the attack, from planning, developing, deploying, and executing to finalizing. Goals may change each step entirely. Figure 4.4 illustrates two possible goals.

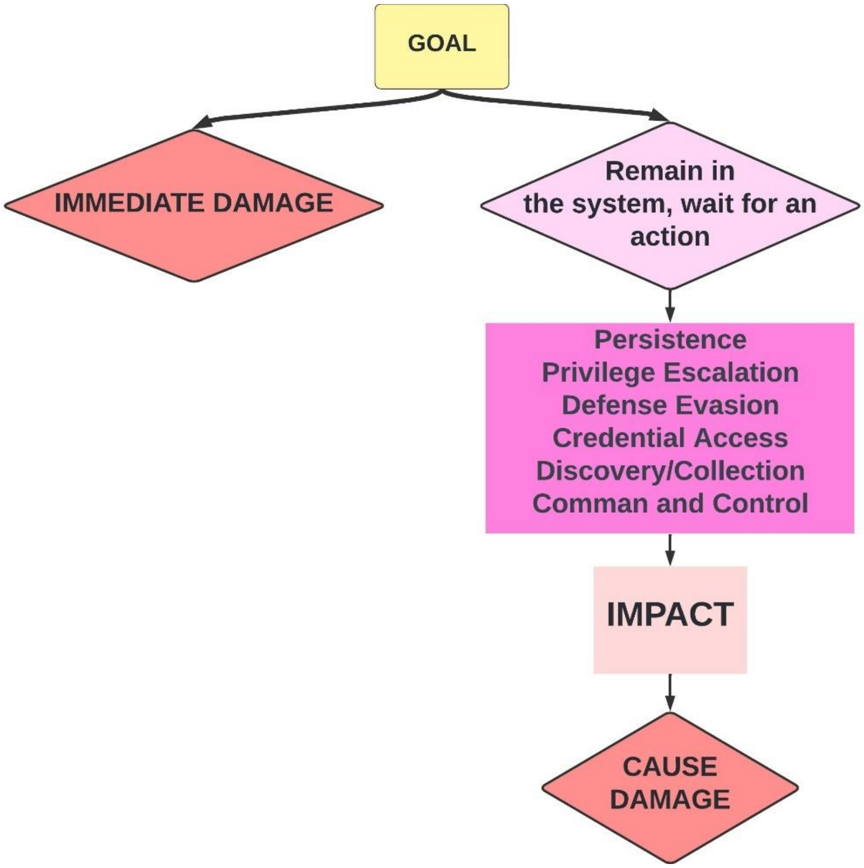


Figure 4.4 Attack goal and attack finalization

In this study, we focus on scenarios where attackers aim to remain in the system for future exploitations. The attacker would consider defense evasion, intrusion detection, and preventive systems, seeking a way to infect with minimum impact on the targeted system, disguising its presence as much as possible to avoid suspicion. Recent studies indicated that enterprise monitoring systems, intrusion detection, and prevention systems also have vulnerabilities, and hacking them is achievable [88] [89] [90]. Knowing what techniques are used (insider's knowledge) helps the attacker develop an attack by using system information and its weaknesses, thereby gaining access, and passing or avoiding control measures. Figure 4.5 shows the attack pattern studied in this work.

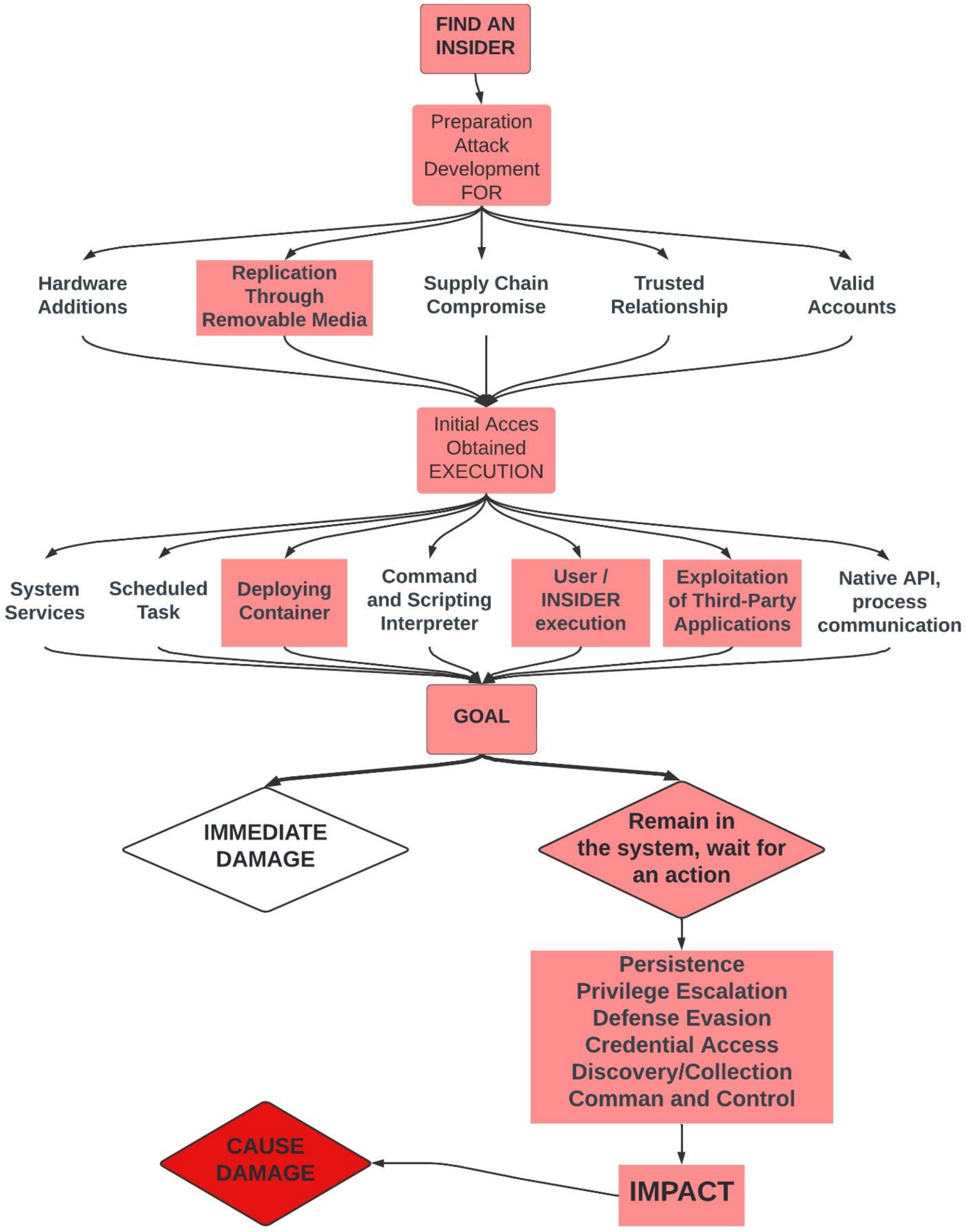


Figure 4.5 Attack pattern of this study

4.5. Summary

In this study, the attack scenario assumes the participation of a plant employee. This assumption simplifies the attack scenario description and analysis, but it does not limit the applicability of our work. The steps the “insider attacker” carries out in this study could be realized by malicious code that manipulates the controller of analog devices.Figure 4.6 presents the scenario explored in this research as a flowchart.

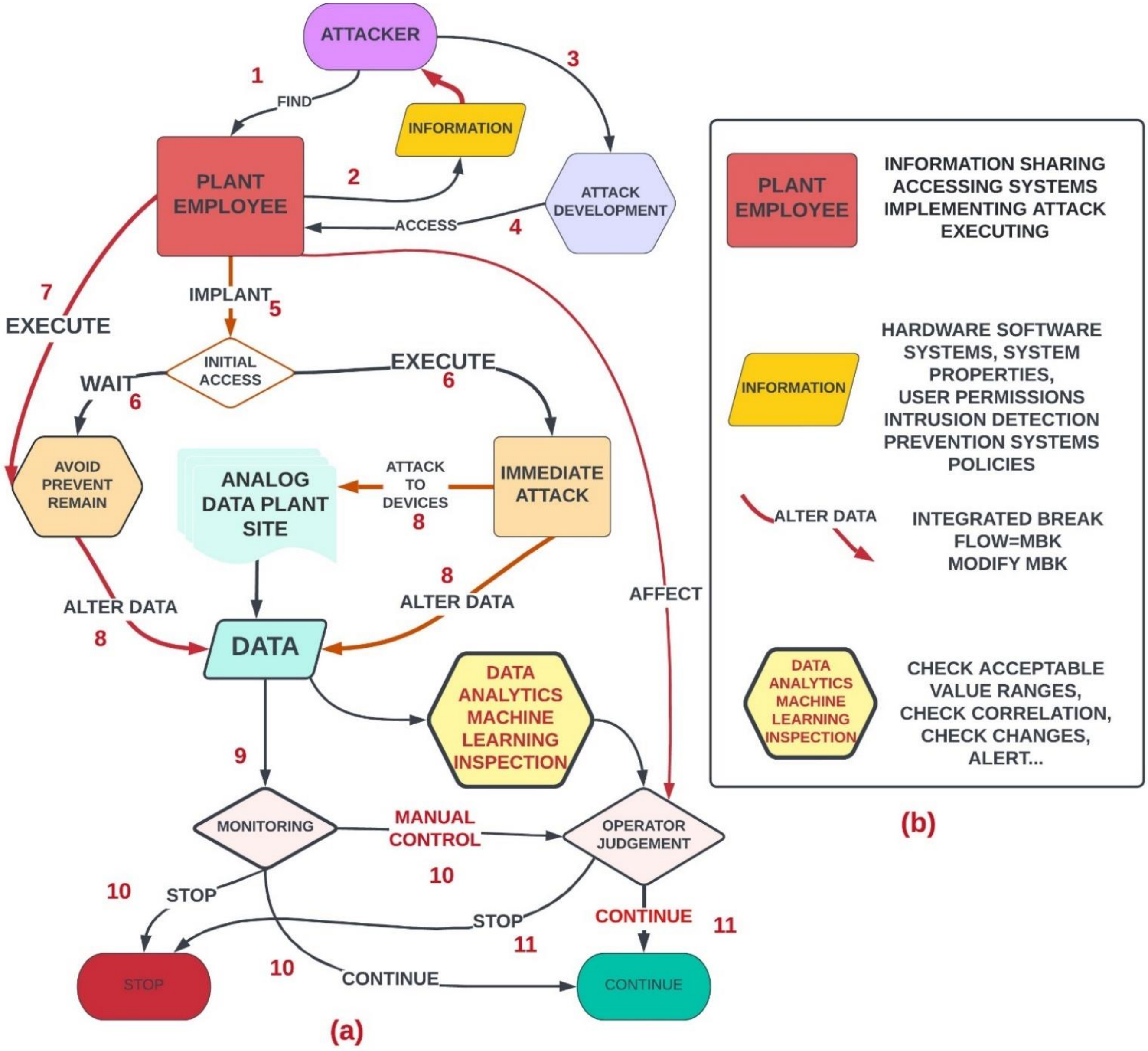


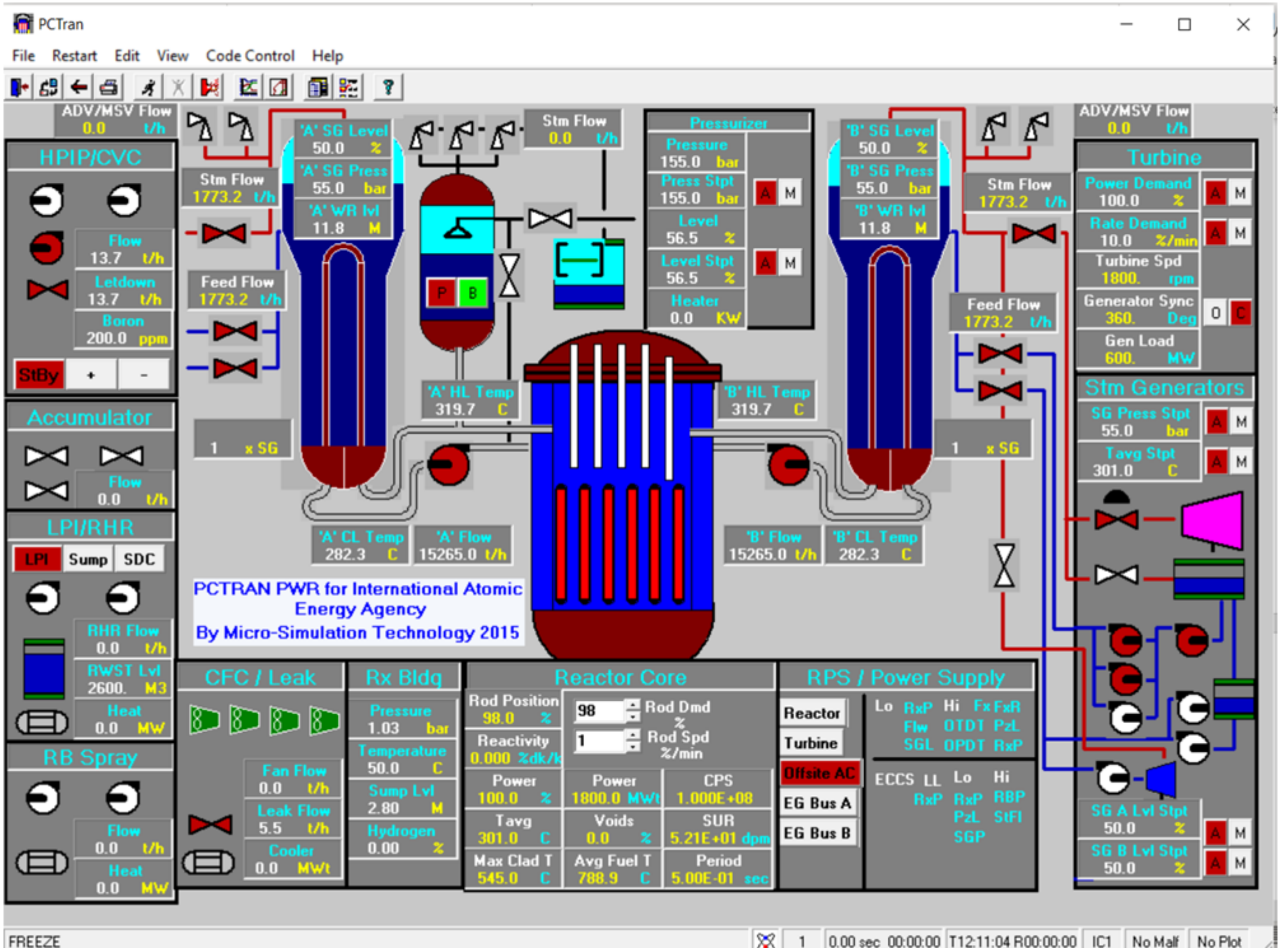
Figure 4.6 Flowchart representation of the attack scenario investigated in the work. Part (a) represents the sequence of actions and activities in the attack. Part (b) identifies the role of the attack actors participating in the attack.

5. EXPERIMENTS WITH THE PCTRAN REACTOR SIMULATOR

In Chapter 5, we discussed the role of simulators in training. We argued that existing simulators were not designed for modeling attacks since they do not capture how data from sensors and actuators is generated or transmitted within the nuclear power plant’s systems. This study explores ways to bypass these limitations by using data analysis to assess the effects of tapering with device data as it is captured or transferred in the system.

5.1. Simulator

In this study, we worked with the IAEA's Personal Computer Transient Analyzer (PCTRAN PWR, version 6.0.4), a two-loop pressurized water reactor simulator. Users can turn on or off valves, pumps, and generators and observe the changes in values. Figure 5.1 shows a snapshot of the tool interface. Red-colored components are operating pumps and open valves, and white-colored components are idle pumps and closed valves [84].



ADV/MSV Flow 0.0 t/h

Stm Flow 1773.2 t/h

Feed Flow 1773.2 t/h

'A' SG Level 50.0 %

'A' SG Press 55.0 bar

'A' WR lvl 11.8 M

Stm Flow 0.0 t/h

1 x SG

'A' HL Temp 319.7 C

'A' CL Temp 282.3 C

'A' Flow 15265.0 t/h

Flow 13.7 t/h

Letdown 13.7 t/h

Boron 200.0 ppm

Flow 0.0 t/h

RHR Flow 0.0 t/h

RWST Lvl 2600. M3

Heat 0.0 MW

Flow 0.0 t/h

Heat 0.0 MW

PCTRAN PWR for International Atomic Energy Agency
By Micro-Simulation Technology 2015

Pressure 155.0 bar

Press Stpt 155.0 bar

Level 56.5 %

Level Stpt 56.5 %

Heater 0.0 KW

'B' SG Level 50.0 %

'B' SG Press 55.0 bar

'B' WR lvl 11.8 M

Stm Flow 1773.2 t/h

Feed Flow 1773.2 t/h

1 x SG

'B' HL Temp 319.7 C

'B' Flow 15265.0 t/h

'B' CL Temp 282.3 C

ADV/MSV Flow 0.0 t/h

Power Demand 100.0 %

Rate Demand 10.0 %/min

Turbine Spd 1800. rpm

Generator Sync 360. Deg

Gen Load 600. MW

SG Press Stpt 55.0 bar

Tavg Stpt 301.0 C

SG A Lvl Stpt 50.0 %

SG B Lvl Stpt 50.0 %

Fan Flow 0.0 t/h

Leak Flow 5.5 t/h

Cooler 0.0 MWt

Pressure 1.03 bar

Temperature 50.0 C

Sump Lvl 2.80 M

Hydrogen 0.00 %

Rod Position 98.0 %

Reactivity 0.000 %dk/k

98 Rod Dmd %

1 Rod Spd %/min

Power 100.0 %

Power 1800.0 MWt

CPS 1.000E+08

Tavg 301.0 C

Voids 0.0 %

SUR 5.21E+01 dpm

Max Clad T 545.0 C

Avg Fuel T 788.9 C

Period 5.00E-01 sec

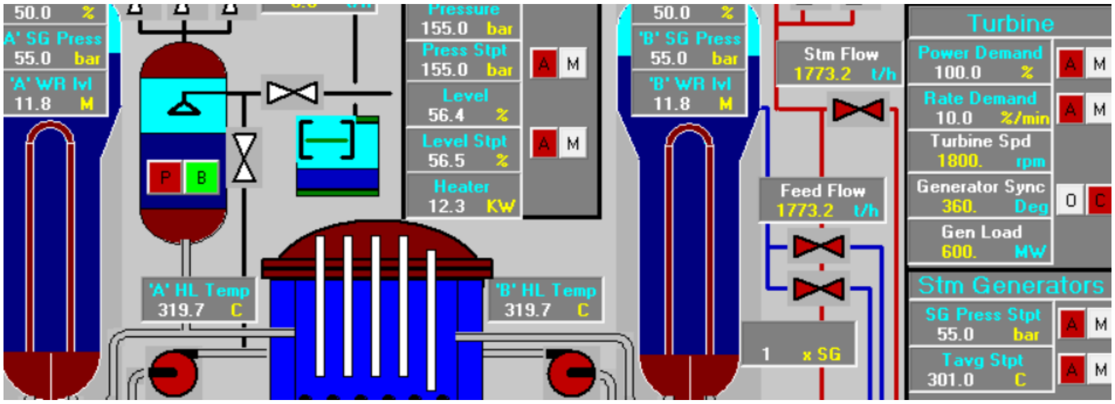
Lo RxP Hi Fx FxR
Flw OTDT PzL
SGL OPDT RxP

ECCS LL Lo Hi
RxP RxP RBP
PzL StFI
SGP

1 | 0.00 sec 00:00:00 | T12:11:04 R00:00:00 | IC1 | No Malf | No Plot

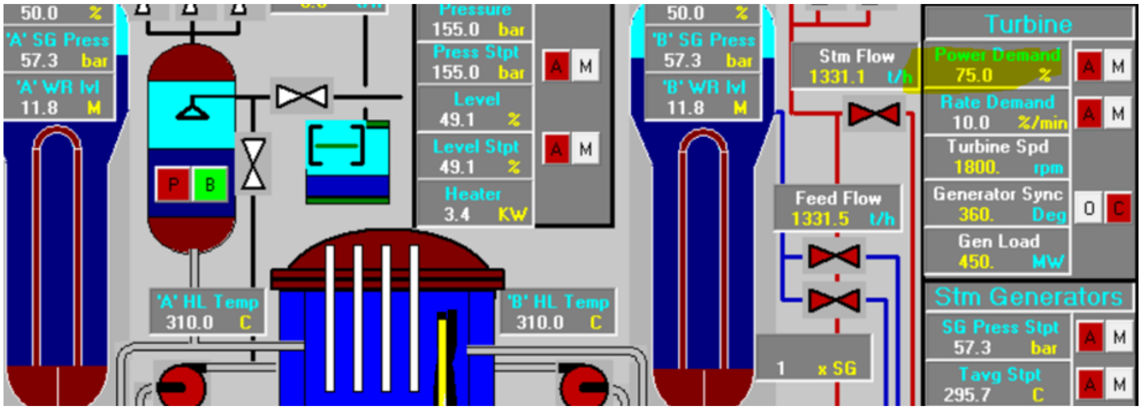
Figure 5.1: Screenshot of PCTRAN simulation interface

The simulator allows users to change values such as turbine power demand and reactor core rod positions. Figures 5.2 and 5.3 show simulator runs with %100 power demand and %75 power demand.



50.0 %
'A' SG Press 55.0 bar
'A' WR lvl 11.8 M
Pressure 155.0 bar
Press Stpt 155.0 bar
Level 56.4 %
Level Stpt 56.5 %
Heater 12.3 KW
'A' HL Temp 319.7 C
'B' HL Temp 319.7 C
50.0 %
'B' SG Press 55.0 bar
'B' WR lvl 11.8 M
Stm Flow 1773.2 t/h
Feed Flow 1773.2 t/h
1 x SG
Turbine
Power Demand 100.0 %
Rate Demand 10.0 %/min
Turbine Spd 1800. rpm
Generator Sync 360. Deg
Gen Load 600. MW
Stm Generators
SG Press Stpt 55.0 bar
Tavg Stpt 301.0 C

Figure 5.2 Screenshot of PCTRAN Simulator; it runs with %100 power demand

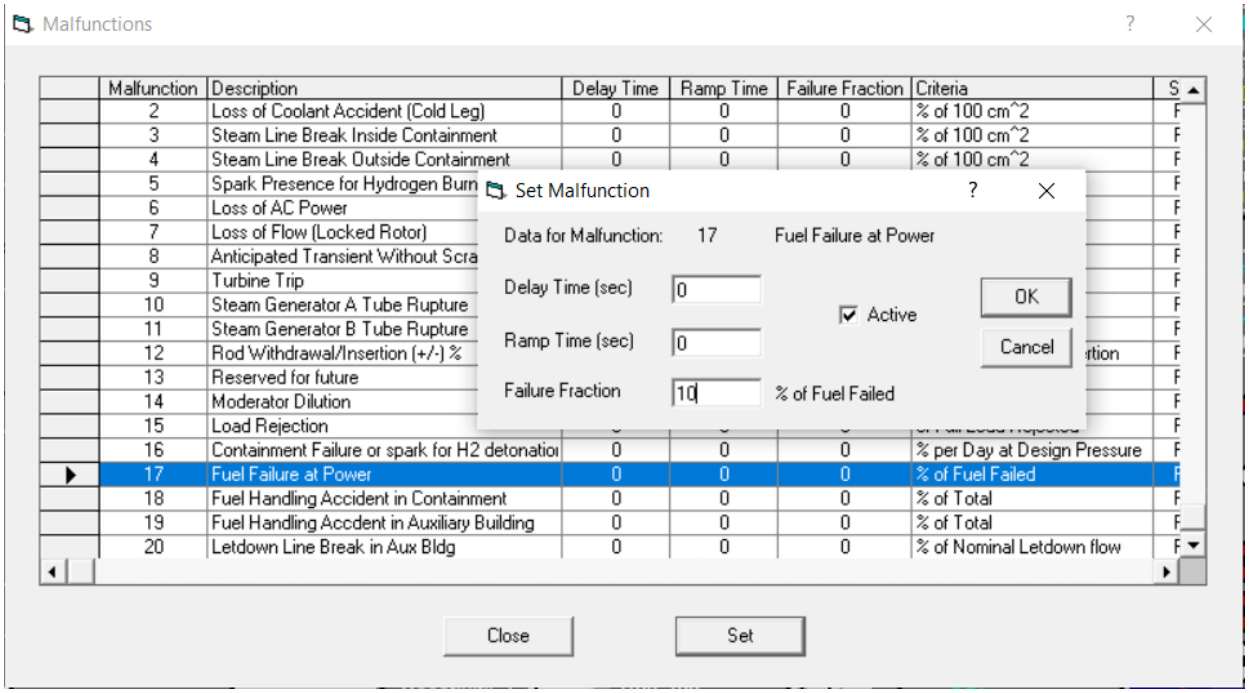


50.0 %
'A' SG Press 57.3 bar
'A' WR lvl 11.8 M
Pressure 155.0 bar
Press Stpt 155.0 bar
Level 49.1 %
Level Stpt 49.1 %
Heater 3.4 KW
'A' HL Temp 310.0 C
'B' HL Temp 310.0 C
50.0 %
'B' SG Press 57.3 bar
'B' WR lvl 11.8 M
Stm Flow 1331.1 t/h
Feed Flow 1331.5 t/h
1 x SG
Turbine
Power Demand 75.0 %
Rate Demand 10.0 %/min
Turbine Spd 1800. rpm
Generator Sync 360. Deg
Gen Load 450. MW
Stm Generators
SG Press Stpt 57.3 bar
Tavg Stpt 295.7 C

Figure 5.3 Screenshot of PCTRAN Simulator; it runs with %75 power demand

The simulator is unrealistic in some ways. For example, in real reactors, several control rods change their position together for a reactor power change; however, in PCTRAN, the control rod icons move one by one to represent a change of rods' position. Still, numerical results and plots are realistic.

There are twenty available malfunctions in PCTRAN. Running the same malfunction with different properties is possible. Users can change the failure fraction or other components of related malfunctions and observe the effect of the changes by running the simulation. Figure 5.4 shows the malfunctions and setting of a malfunction failure fraction to ten percent.



Malfunctions

	Malfunction	Description	Delay Time	Ramp Time	Failure Fraction	Criteria	S
	2	Loss of Coolant Accident (Cold Leg)	0	0	0	% of 100 cm^2	F
	3	Steam Line Break Inside Containment	0	0	0	% of 100 cm^2	F
	4	Steam Line Break Outside Containment	0	0	0	% of 100 cm^2	F
	5	Spark Presence for Hydrogen Burn					F
	6	Loss of AC Power					F
	7	Loss of Flow (Locked Rotor)					F
	8	Anticipated Transient Without Scra					F
	9	Turbine Trip					F
	10	Steam Generator A Tube Rupture					F
	11	Steam Generator B Tube Rupture					F
	12	Rod Withdrawal/Insertion (+/-) %				tion	F
	13	Reserved for future					F
	14	Moderator Dilution					F
	15	Load Rejection					F
	16	Containment Failure or spark for H2 detonatio	0	0	0	% per Day at Design Pressure	F
▶	17	Fuel Failure at Power	0	0	0	% of Fuel Failed	F
	18	Fuel Handling Accident in Containment	0	0	0	% of Total	F
	19	Fuel Handling Accdent in Auxiliary Building	0	0	0	% of Total	F
	20	Letdown Line Break in Aux Bldg	0	0	0	% of Nominal Letdown flow	F

Set Malfunction

Data for Malfunction: 17 Fuel Failure at Power

Delay Time (sec) 0

Ramp Time (sec) 0

Failure Fraction 10 % of Fuel Failed

☑ Active

Figure 5.4 Screenshot of PCTRAN interface for configuring malfunctions

Users can activate some malfunctions or manually cause a reactor trip process and observe it. The Reactor Protection System (RPS) shuts down a PWR power plant when commanded by the operator or specific safety system settings, or setpoints are reached. Some of the crucial parameters, such as pressure inside the pressurizer, reactor coolant flow rate, steam generator water level, etc., are continuously compared to specified safe operation limits. The RPS automatically shuts down the reactor when any parameter exceeds its limit. With a reactor's trip signal, all the control rods are inserted rapidly to absorb neutrons in the reactor and, thus, cease the nuclear fission chain reaction. According to the NRC, SCRAM is "The sudden shutting down of a nuclear reactor, usually by rapid insertion of control rods, either automatically or manually by the reactor operator. Also known as a "reactor trip." Figure 5.5 depicts the PCTRAN interface when a trip has occurred.

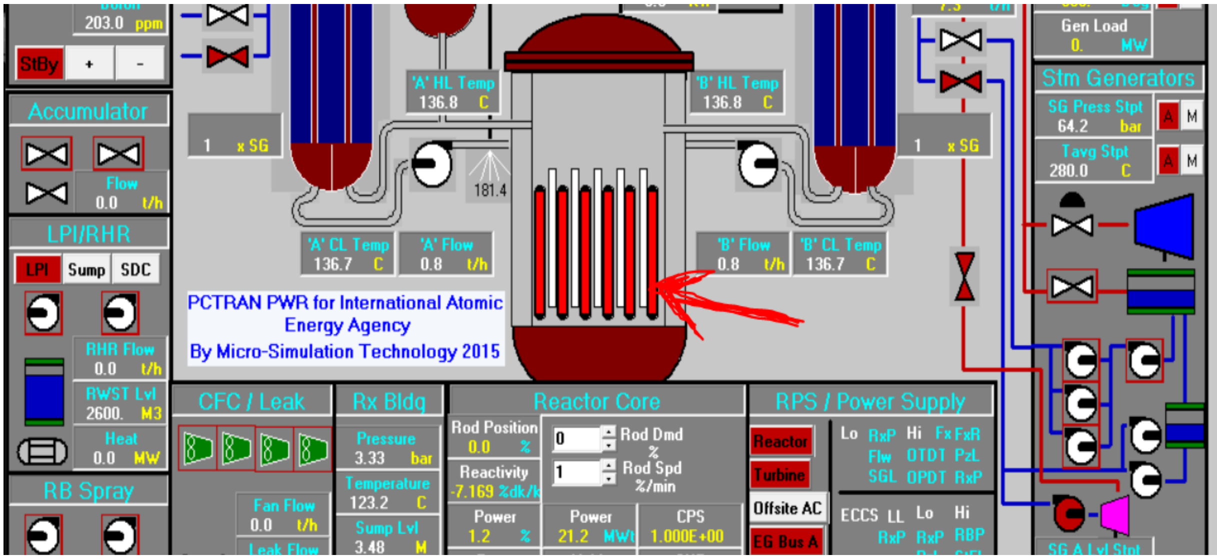


Figure 5.5 Screenshot of PCTRAN. When a reactor trip occurs, control rods are dropped to minimize the nuclear reaction

When a malfunction is activated, some components change their colors, and red boxes appear around some related components. Figures 5.6 and 5.7 show the change in the components.

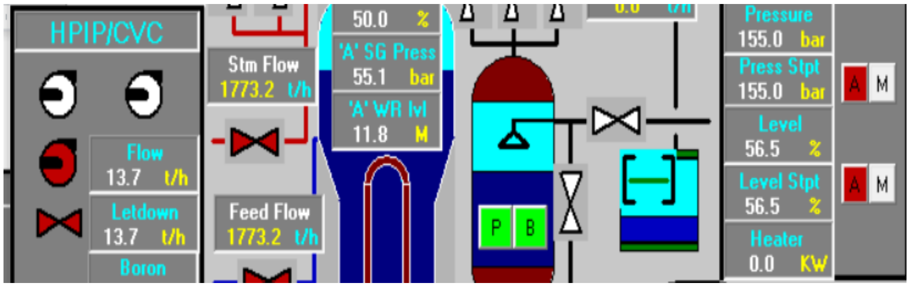


Figure 5.6: Screenshot of PCTRAN Simulator, change in components appearances

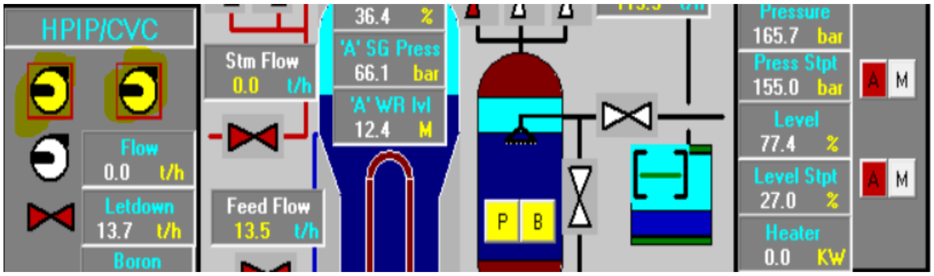
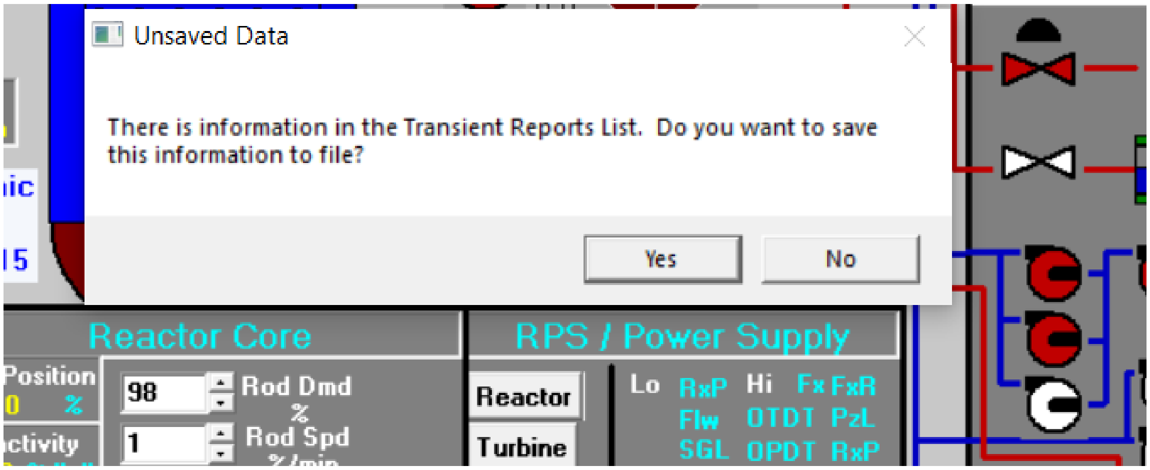


Figure 5.7: Screenshot of PCTRAN Simulator, change in components appearances

After running PCTRAN, all the scenarios, updates, and changes in components, the simulator's state can be saved as data in MS Office's Access database format, as shown in Figure 5.8.


Unsaved Data

There is information in the Transient Reports List. Do you want to save this information to file?

Figure 5.8: Screenshot of PCTRAN, data saving process

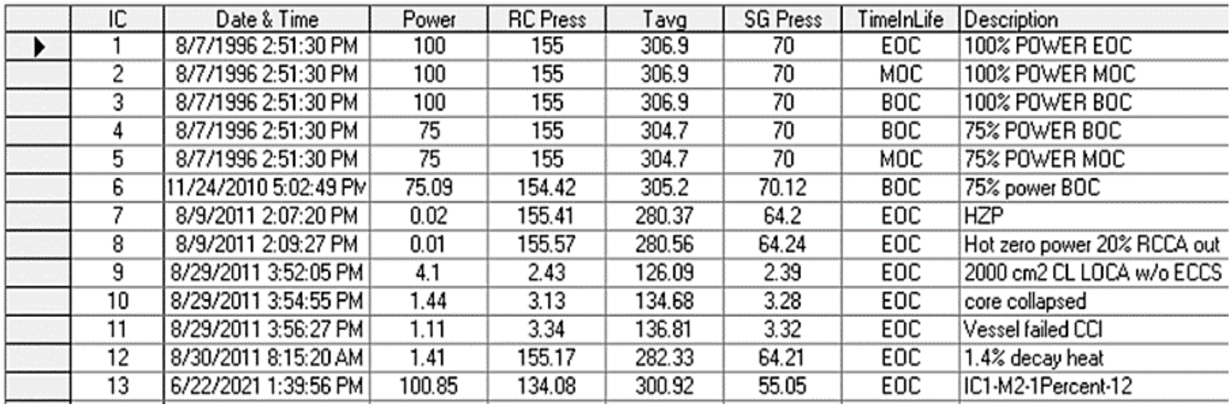
5.2. Experimental Method

Our goal is to explore how feasible it is to use the PCTRAN simulator to investigate new data-driven approaches to detect attacks. For that purpose, we manually changed several PCTRAN component values and continued the simulation to observe the effect of the change. For each of these experiments, we extracted the resulting datasets and analyzed them to assess how well machine learning algorithms could be used to predict behavior.

5.3. Data Extraction

The PCTRAN documentation does not specify ranges of accepted values for its several components. In order to find the accepted value ranges, we ran the simulator with different scenarios. For each scenario, we identified the outcome (for example, the reactor trip within a period of simulated time) and collected datasets capturing the reactor behavior after the change. The reactor properties investigated were fuel information, power demand, rod position, pumps, and valve states (open or closed). Through these manual value changes, we generated large datasets that captured the reactor state in a broad range of situations.

The experimental process can be illustrated by experimentation with one of the studied properties: fuel information. Fuel life cycle information values such as EOC (End of Cycle), MOC (Middle of Cycle), and BOC (Beginning of Cycle) affect the simulation data. Figure 5.9 shows the PCTRAN initial conditions interface with different fuel cycle use (TimeInLife column). We ran many scenarios repeatedly, using different fuel cycle status values while keeping other variables unchanged. As a result, we acquired a broad range of reactor data.



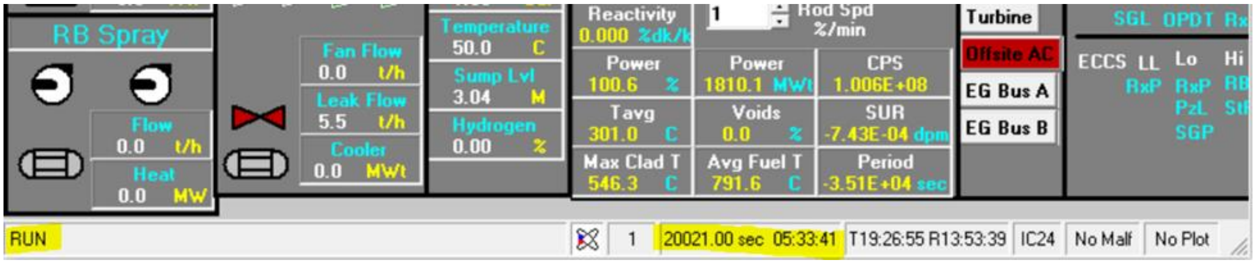
	IC	Date & Time	Power	RC Press	Tavg	SG Press	TimeInLife	Description
▶	1	8/7/1996 2:51:30 PM	100	155	306.9	70	EOC	100% POWER EOC
	2	8/7/1996 2:51:30 PM	100	155	306.9	70	MOC	100% POWER MOC
	3	8/7/1996 2:51:30 PM	100	155	306.9	70	BOC	100% POWER BOC
	4	8/7/1996 2:51:30 PM	75	155	304.7	70	BOC	75% POWER BOC
	5	8/7/1996 2:51:30 PM	75	155	304.7	70	MOC	75% POWER MOC
	6	11/24/2010 5:02:49 PM	75.09	154.42	305.2	70.12	BOC	75% power BOC
	7	8/9/2011 2:07:20 PM	0.02	155.41	280.37	64.2	EOC	HZP
	8	8/9/2011 2:09:27 PM	0.01	155.57	280.56	64.24	EOC	Hot zero power 20% RCCA out
	9	8/29/2011 3:52:05 PM	4.1	2.43	126.09	2.39	EOC	2000 cm2 CL LOCA w/o ECCS
	10	8/29/2011 3:54:55 PM	1.44	3.13	134.68	3.28	EOC	core collapsed
	11	8/29/2011 3:56:27 PM	1.11	3.34	136.81	3.32	EOC	Vessel failed CCI
	12	8/30/2011 8:15:20 AM	1.41	155.17	282.33	64.21	EOC	1.4% decay heat
	13	6/22/2021 1:39:56 PM	100.85	134.08	300.92	55.05	EOC	IC1-M2-1Percent-12

Figure 5.9: Screenshot of PCTRAN simulator running conditions with different fuel cycles (TimeInLife) and properties

First, we ran the simulator under normal conditions, extracting datasets that correspond to normal reactor operation. Next, we create unwanted scenarios such as incidents, malfunctions, or accidents. The resulting datasets can be used to create models to predict unwanted outcomes.

If an accident scenario is not embedded into the simulator system, creating alternative new scenarios is challenging or impossible with most simulators. Without access to the simulator’s internal design and the ability to change its implementation to additional model components, it is not possible to reproduce well-studied accidents/incidents like Fukushima or Chernobyl because they involve changes on reactor characteristics not captured by simulators such as PCTRAN. Also, cyber-related incidents like Stuxnet cannot be simulated without access to the control components impacted by the attack. However, we were able to recreate the Three Miles Island incident scenario because the simulator's user manual shows step-by-step how the accident happened [84]. We also reproduced two more hypothetical cases from the manual, twenty different malfunction cases, and random cases with varying demands of power and fuel quality when a malfunction was running. As a result of these experiments, we gathered many datasets corresponding to abnormal cases.

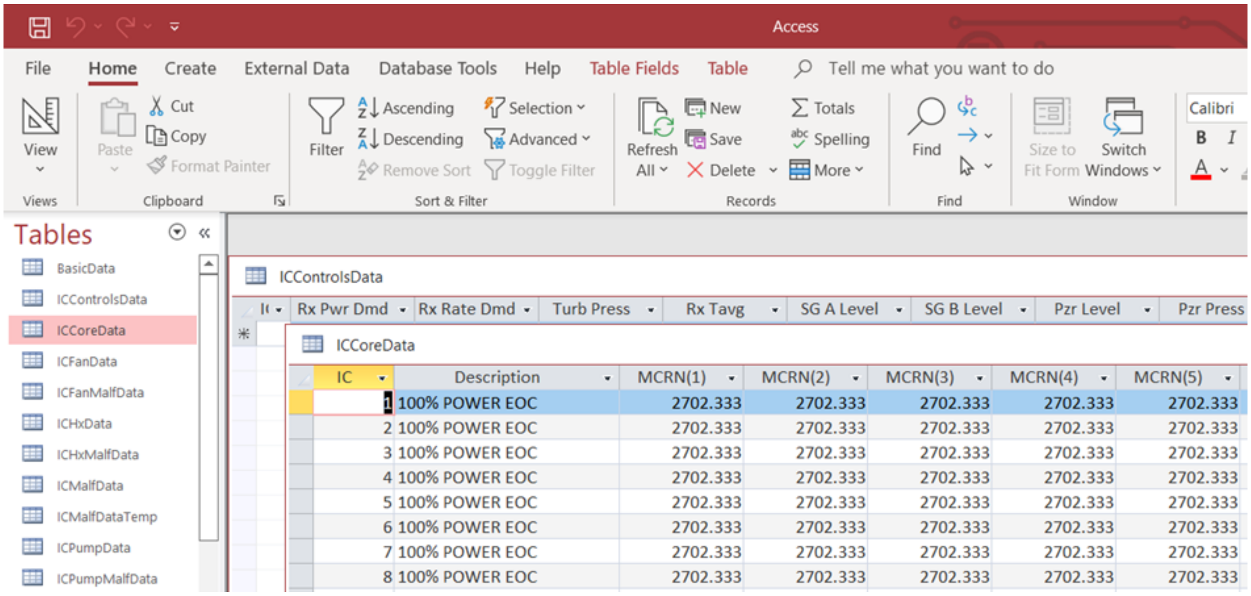
The simulator runs only for a period covering 1000 seconds, but once one execution is completed, users can save the recent run's data as the initial conditions for subsequent execution. Figure 5.10 shows the simulator after a sequence of 20 executions that simulates 20,000 seconds.



RUN | 1 | 20021.00 sec 05:33:41 | T19:26:55 R13:53:39 | IC24 | No Malf | No Plot

Figure 5.10: Screenshot of PCTRAN simulator running information, time, status

PCTRAN simulation data is saved in a Microsoft Access Database (MDB) object. Figure 5.11 shows an example.



ICControlsData

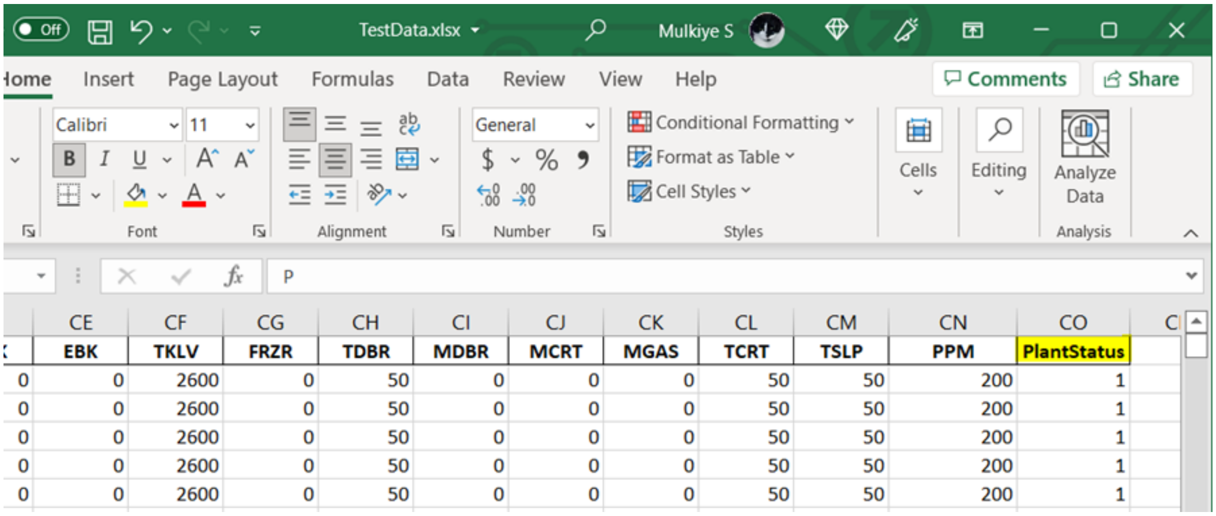
I(	Rx Pwr Dmd	Rx Rate Dmd	Turb Press	Rx Tavg	SG A Level	SG B Level	Pzr Level	Pzr Press

ICCoreData

IC	Description	MCRN(1)	MCRN(2)	MCRN(3)	MCRN(4)	MCRN(5)
1	100% POWER EOC	2702.333	2702.333	2702.333	2702.333	2702.333
2	100% POWER EOC	2702.333	2702.333	2702.333	2702.333	2702.333
3	100% POWER EOC	2702.333	2702.333	2702.333	2702.333	2702.333
4	100% POWER EOC	2702.333	2702.333	2702.333	2702.333	2702.333
5	100% POWER EOC	2702.333	2702.333	2702.333	2702.333	2702.333
6	100% POWER EOC	2702.333	2702.333	2702.333	2702.333	2702.333
7	100% POWER EOC	2702.333	2702.333	2702.333	2702.333	2702.333
8	100% POWER EOC	2702.333	2702.333	2702.333	2702.333	2702.333

Figure 5.11: Saved data - MDB file

After collecting data, we converted the MDB access files to Microsoft Excel files. In the Excel spreadsheet, a new column is created named **PlantStatus** to use for classification; its value was determined by inspecting the property values. Figure 5.12 shows the data in excel format with the PlantStatus column.



	CE	CF	CG	CH	CI	CJ	CK	CL	CM	CN	CO
	EBK	TKLV	FRZR	TDBR	MDBR	MCRT	MGAS	TCRT	TSLP	PPM	PlantStatus
0	0	2600	0	50	0	0	0	50	50	200	1
0	0	2600	0	50	0	0	0	50	50	200	1
0	0	2600	0	50	0	0	0	50	50	200	1
0	0	2600	0	50	0	0	0	50	50	200	1
0	0	2600	0	50	0	0	0	50	50	200	1

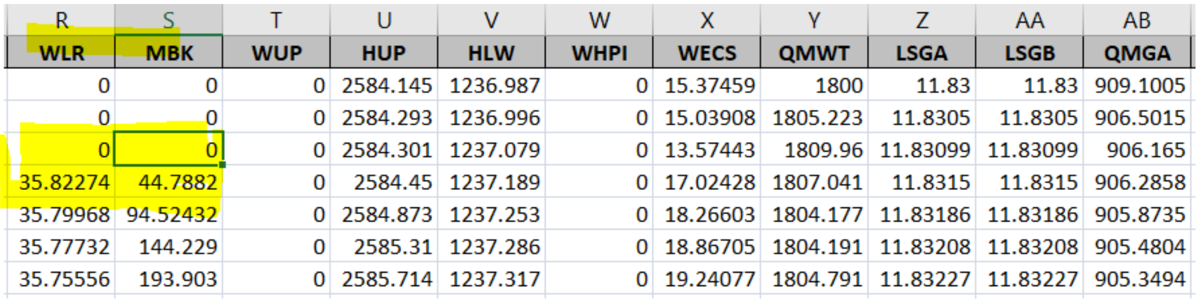
Figure 5.12: Data in excel format with new column PlantStatus

We define three labels for **PlantStatus: 1-Normal, 2-Abnormal, and 3-Trip**. ***Normal*** status means every property value is in its acceptable range, and the plant works as expected. ***Trip*** status means a reactor trip occurs, rods are inserted, the reaction is at the minimum level, and the plant is off. If a plant trip happens, it indicates that there may be a significant problem. Our research introduces the ***Abnormal*** status to capture situations where property values are in their acceptable range, and the plant is operating as expected, but there is a property exhibiting minor value variations that reveal an unusual trend that may lead later on to an undesirable state. The goal is to investigate how data science techniques could be used to detect emerging problematic situations that plant operators are not able to see yet. Using datasets extracted from the simulator, we investigate how effective data-driven methods are in predicting Abnormal states.

If an effective prediction model is available, it can be deployed during operation: NPP data is gathered from the site, converted into readable form, and analyzed immediately to anticipate unusual situations. Operators continue to monitor the system, but they can also leverage information from this new automated smart monitoring system.

5.4. Data Examination

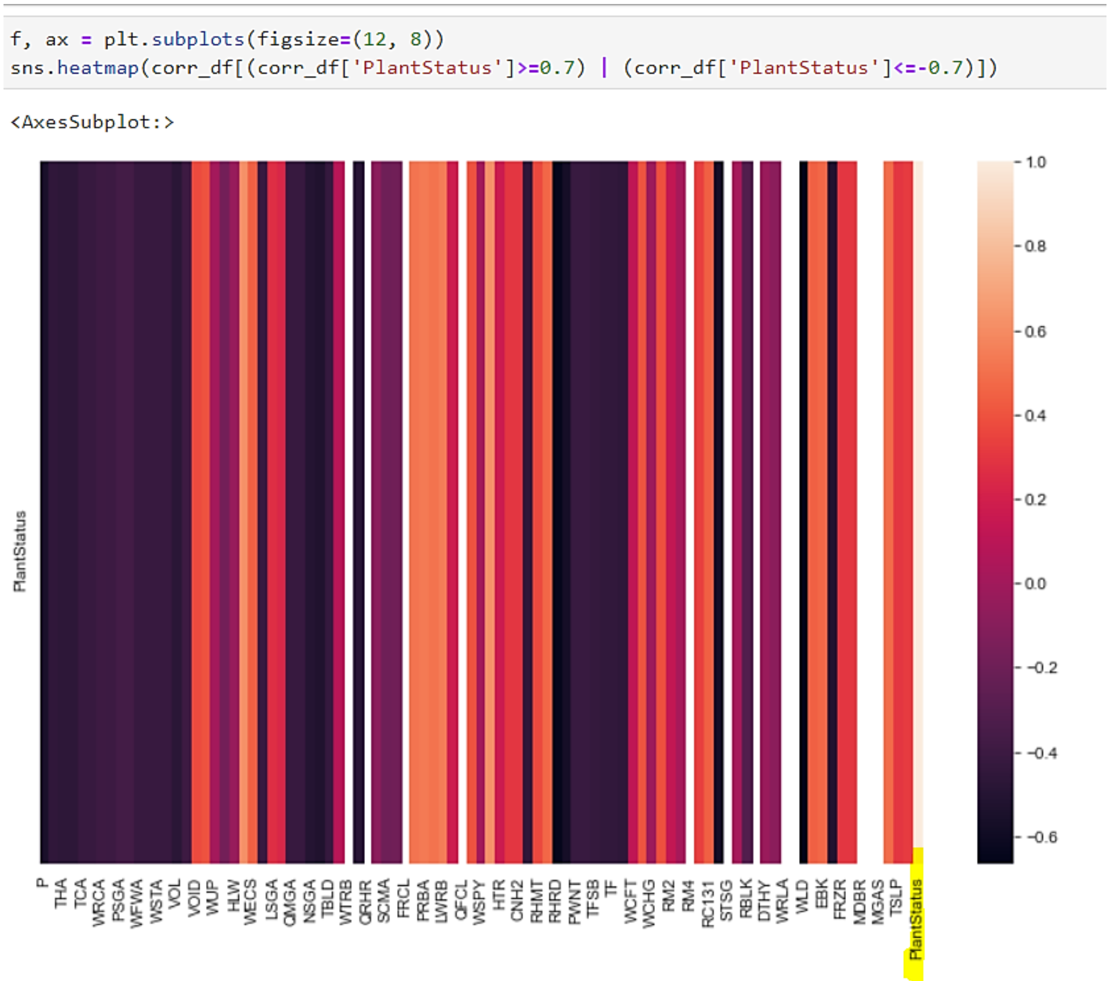
Significant changes in property values on the simulator, such as turning off the cooling system or a radiation monitoring component, result in visible changes in the simulated reactor data. For example, when a malfunction is active (e.g., a small leak in the water flow) on the simulator, it causes visible changes in the values of WLR (Flow Reactor Cooling System Leak) or MBK (Integrated Break Flow). The "simulation operators" can easily spot such significant changes, but it might be hard to notice minor value modifications that may end up leading to significant outcomes. Figures 5.13 illustrate a notable data value change, with both the WLR and MBK properties going from zero to much larger values.



R	S	T	U	V	W	X	Y	Z	AA	AB
WLR	MBK	WUP	HUP	HLW	WHPI	WECS	QMWT	LSGA	LSGB	QMGA
0	0	0	2584.145	1236.987	0	15.37459	1800	11.83	11.83	909.1005
0	0	0	2584.293	1236.996	0	15.03908	1805.223	11.8305	11.8305	906.5015
0	0	0	2584.301	1237.079	0	13.57443	1809.96	11.83099	11.83099	906.165
35.82274	44.7882	0	2584.45	1237.189	0	17.02428	1807.041	11.8315	11.8315	906.2858
35.79968	94.52432	0	2584.873	1237.253	0	18.26603	1804.177	11.83186	11.83186	905.8735
35.77732	144.229	0	2585.31	1237.286	0	18.86705	1804.191	11.83208	11.83208	905.4804
35.75556	193.903	0	2585.714	1237.317	0	19.24077	1804.791	11.83227	11.83227	905.3494

Figure 5.13: Example of notable change in property values.

Identifying which properties have a larger effect on the final status of the plant can be quite useful. With the help of data analytics tools, we can identify correlations between PlantStatus and other features of plant data. Such insights can help design and implement better protection measures for the critical properties of the reactor. Figure 5.14 displays the correlation between properties and PlantStatus for one of the experiments and illustrates how the correlation may vary.



```
f, ax = plt.subplots(figsize=(12, 8))
sns.heatmap(corr_df[(corr_df['PlantStatus']>=0.7) | (corr_df['PlantStatus']<=-0.7)])
```

```
<AxesSubplot:>
```

Figure 5.14: Data Correlation between PlantStatus and plant data properties

5.5. Machine Learning

Russel and Norvig, in their classic book[91], explain machine learning (ML) as "a computer observes some data, builds a model based on the data, and uses the model as both a hypothesis about the world and a piece of software that can solve problems." Machine learning models are designed to be predictive (foresee the future), descriptive (gain knowledge from data), or both.

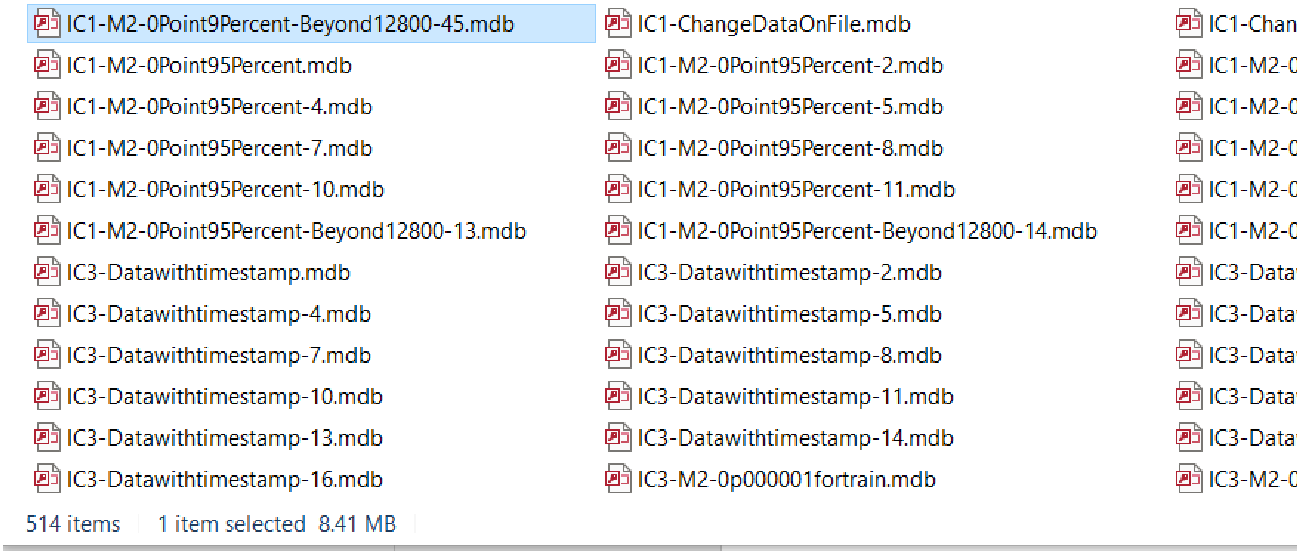
There are different applications of machine learning. For example, the association rule is the interest in learning a conditional probability between distinct aspects of data. Another application is classification, which is a method that takes data and assigns it to separate groups or classes. For example, in this thesis, we classified the reactor data as Normal, Abnormal, or Trip. Regression is another machine learning application. The output is calculated as an actual number based on the given input. Both classification and regression are supervised learning applications that, given input Xi and the corresponding output Yi, identify a method to map the input onto the output. Another ML application is unsupervised learning, which uses only input data to find the input's regularities (density estimation). One method for density estimation is clustering, where the goal is to find clusters or groupings of input data.

If machine learning is to be used to understand critical conditions and decision-making based on predictions, developers or users should ensure that they choose the best model. Model selection is crucial for prediction accuracy. While acquiring data, preparing data, and training a model on the dataset, developers should consider the concepts of underfitting, overfitting, and bias [92].

5.5.1. Training Dataset and Test Dataset

The data analysis in this thesis is conducted using Anaconda Navigator (v 2.0.3) and Jupyter Notebook (v 6.3.0) [93].

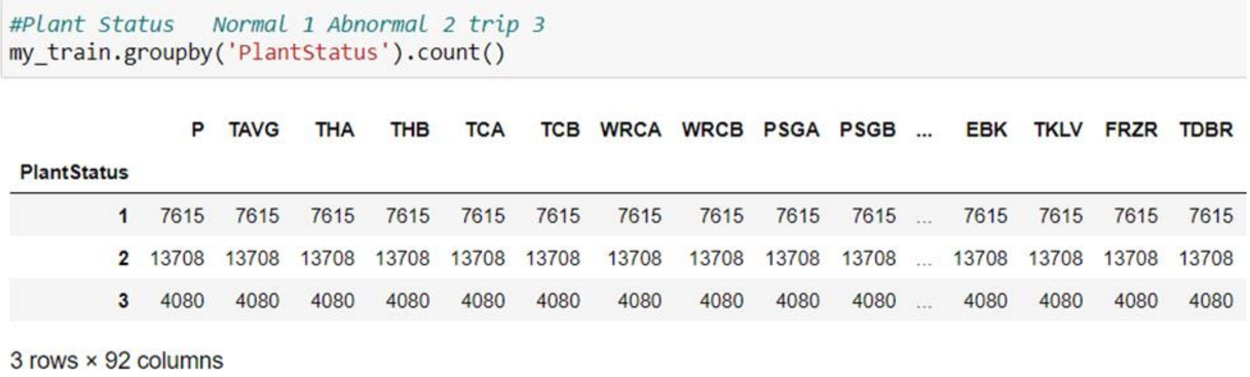
We captured many scenarios and generated a large dataset to train ML models. The scenarios were reactor's vessel failure, turbine trip, Three Miles accident, steam generator tube rupture, small break loss of coolant accident, large break loss of coolant accident, loss of AC power, pre-defined twenty malfunctions, normal run with various levels of energy demands and different rod position demands. While running the simulator, we changed various parameters, such as energy demand reduced from 100 to 45 then, 45 to 75 percent, or malfunctions with various levels of failure fraction (1, 5,10, 20, 50,100 percent). We have over 500 MDB files extracted from the simulator, distributed across several files, as illustrated in Figure 5.15.



```
IC1-M2-0Point9Percent-Beyond12800-45.mdb    IC1-ChangeDataOnFile.mdb                    IC1-Chan
IC1-M2-0Point95Percent.mdb                  IC1-M2-0Point95Percent-2.mdb                IC1-M2-C
IC1-M2-0Point95Percent-4.mdb                IC1-M2-0Point95Percent-5.mdb                IC1-M2-C
IC1-M2-0Point95Percent-7.mdb                IC1-M2-0Point95Percent-8.mdb                IC1-M2-C
IC1-M2-0Point95Percent-10.mdb               IC1-M2-0Point95Percent-11.mdb               IC1-M2-C
IC1-M2-0Point95Percent-Beyond12800-13.mdb   IC1-M2-0Point95Percent-Beyond12800-14.mdb   IC1-M2-C
IC3-Datawithtimestamp.mdb                   IC3-Datawithtimestamp-2.mdb                 IC3-Data
IC3-Datawithtimestamp-4.mdb                 IC3-Datawithtimestamp-5.mdb                 IC3-Data
IC3-Datawithtimestamp-7.mdb                 IC3-Datawithtimestamp-8.mdb                 IC3-Data
IC3-Datawithtimestamp-10.mdb                IC3-Datawithtimestamp-11.mdb                IC3-Data
IC3-Datawithtimestamp-13.mdb                IC3-Datawithtimestamp-14.mdb                IC3-Data
IC3-Datawithtimestamp-16.mdb                IC3-M2-0p000001fortrain.mdb                 IC3-M2-C
514 items    1 item selected  8.41 MB
```

Figure 5.15: Microsoft Access Database (MDB) files; we extracted the data for every different simulator's running scenario and saved it as an MDB file.

Figure 5.16 shows the distribution of PlantStatus labeling for the **training dataset**. The total number of data entries for the training dataset is **25403**.

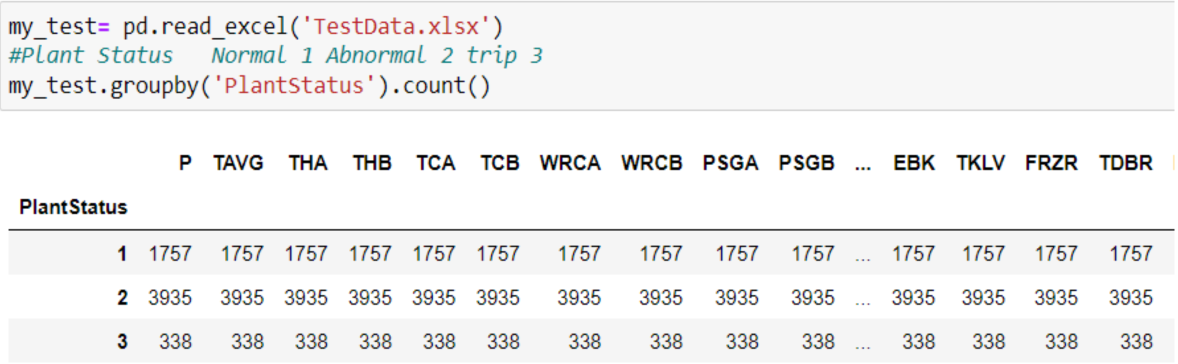


```
#Plant Status   Normal 1 Abnormal 2 trip 3
my_train.groupby('PlantStatus').count()
```

```
                P   TAVG   THA    THB    TCA    TCB   WRCA   WRCB   PSGA   PSGB  ...   EBK   TKLV   FRZR   TDBR
PlantStatus
          1  7615   7615   7615   7615   7615   7615   7615   7615   7615   7615  ...   7615   7615   7615   7615
          2 13708  13708  13708  13708  13708  13708  13708  13708  13708  13708  ...  13708  13708  13708  13708
          3  4080   4080   4080   4080   4080   4080   4080   4080   4080   4080  ...   4080   4080   4080   4080

3 rows × 92 columns
```

Figure 5.16: Training Dataset, PlantStatus: 1: Normal (7615) 2:Abnormal (13708) 3: Trip (4080)

We run the simulator with new scenarios such as different power demand levels, rod position, and failure fractions that are not used for training dataset creation. We created a new dataset for testing to observe how ML models classified the new (unseen) data. Figure 5.17 shows the data distribution for PlantStatus on the **test dataset**. The total number of data entries for the test dataset is **6030**.



```
my_test= pd.read_excel('TestData.xlsx')
#Plant Status   Normal 1 Abnormal 2 trip 3
my_test.groupby('PlantStatus').count()
```

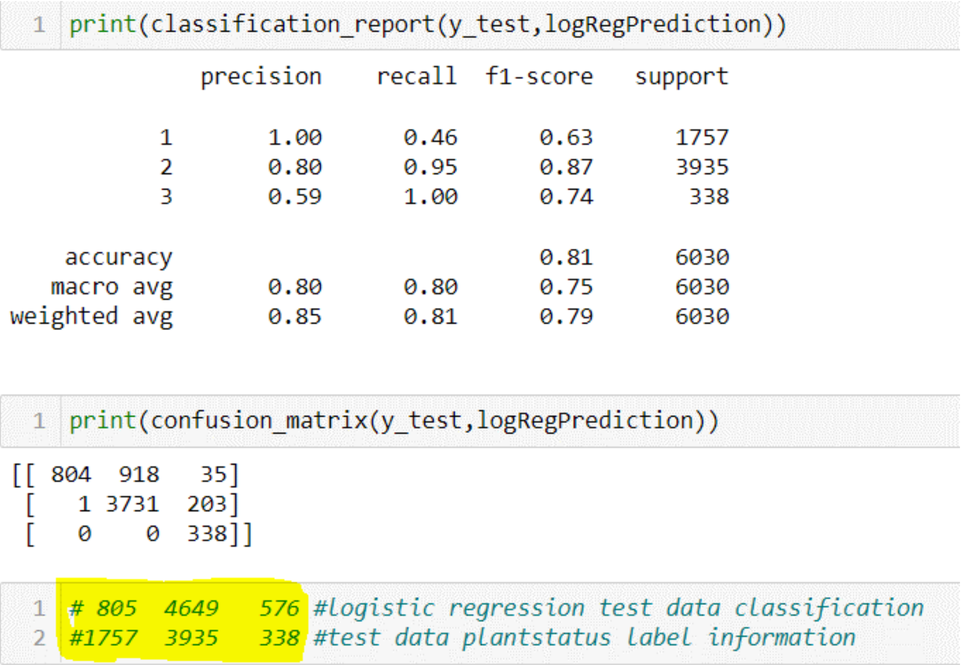
```
               P  TAVG   THA   THB   TCA   TCB  WRCA  WRCB  PSGA  PSGB  ...   EBK  TKLV  FRZR  TDBR
PlantStatus
          1 1757  1757  1757  1757  1757  1757  1757  1757  1757  1757  ...  1757  1757  1757  1757
          2 3935  3935  3935  3935  3935  3935  3935  3935  3935  3935  ...  3935  3935  3935  3935
          3  338   338   338   338   338   338   338   338   338   338  ...   338   338   338   338
```

Figure 5.17: Test Dataset, Distribution of PlantStatus: 1: Normal (1757) 2: Abnormal (3935) 3: Trip (338)

5.5.2. Running Machine Learning Models

We ran some machine learning (ML) models on the test dataset and observed the classification of simulator data. We compared the ML models and their results (i.e., classification and the number of classes).

The first ML model was Logistic Regression (LR), a supervised learning algorithm that is used to predict a dependent categorical target variable. The dependent variable column (PlantStatus) has three different values in this study's data, so multi-linear logistic regression is used. Figure 5.18 shows the results of using LR on the test dataset. The number of data entries classified as Trip is 576, which is bigger than the actual number (338) of Trip cases. Too many (unnecessary) Trip classifications may not be preferable for the operators. The LR showed 81% accuracy, but this can be increased with better model construction.



```
print(classification_report(y_test,logRegPrediction))
```

```
              precision    recall  f1-score   support

           1       1.00      0.46      0.63      1757
           2       0.80      0.95      0.87      3935
           3       0.59      1.00      0.74       338

    accuracy                           0.81      6030
   macro avg       0.80      0.80      0.75      6030
weighted avg       0.85      0.81      0.79      6030
```

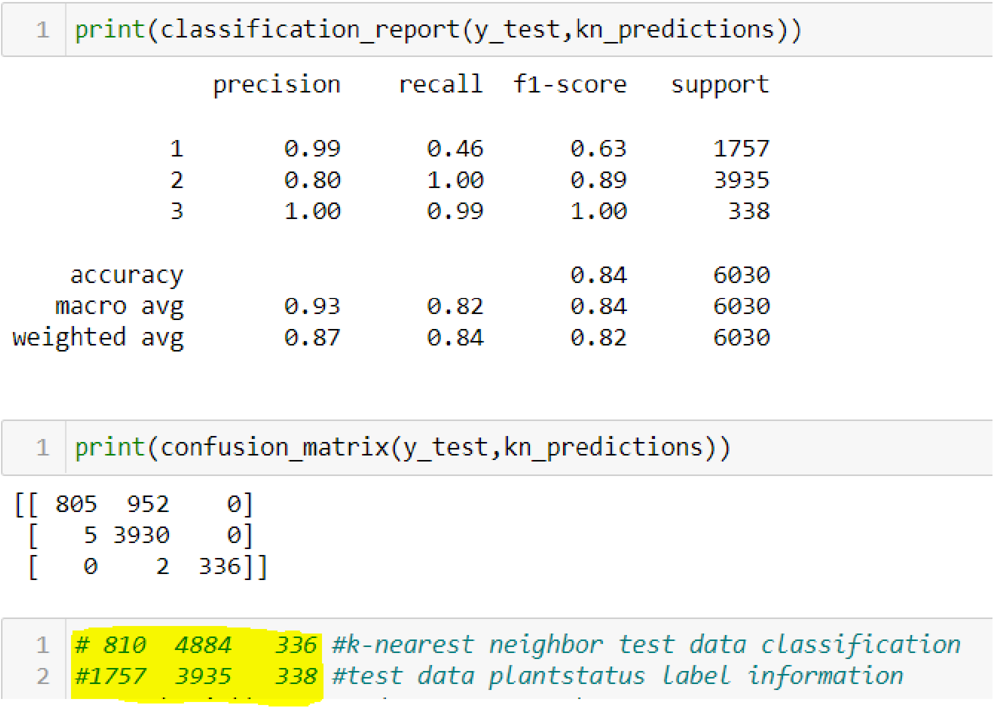
```
print(confusion_matrix(y_test,logRegPrediction))
```

```
[[ 804  918   35]
 [    1 3731  203]
 [    0    0  338]]
```

```
# 805   4649    576 #logistic regression test data classification
#1757   3935    338 #test data plantstatus label information
```

Figure 5.18: Results of Logistic Regression on the test dataset

The second model we used was K-Nearest Neighbors (KNN), a data classification method for estimating the likelihood that a data point will become a member of one group or another based on what group the data points nearest to it belong to. Figure 5.19 shows the results of using KNN. When the number of neighbors is 5, KNN classified two Trip entries as Abnormal; misclassification of Trip cases may not be acceptable by operators since a late plant trip might cause damage to the reactor's core. The KNN showed 84% accuracy, but this can be increased with better model construction.



```
print(classification_report(y_test,kn_predictions))
```

```
              precision    recall  f1-score   support

           1       0.99      0.46      0.63      1757
           2       0.80      1.00      0.89      3935
           3       1.00      0.99      1.00       338

    accuracy                           0.84      6030
   macro avg       0.93      0.82      0.84      6030
weighted avg       0.87      0.84      0.82      6030
```

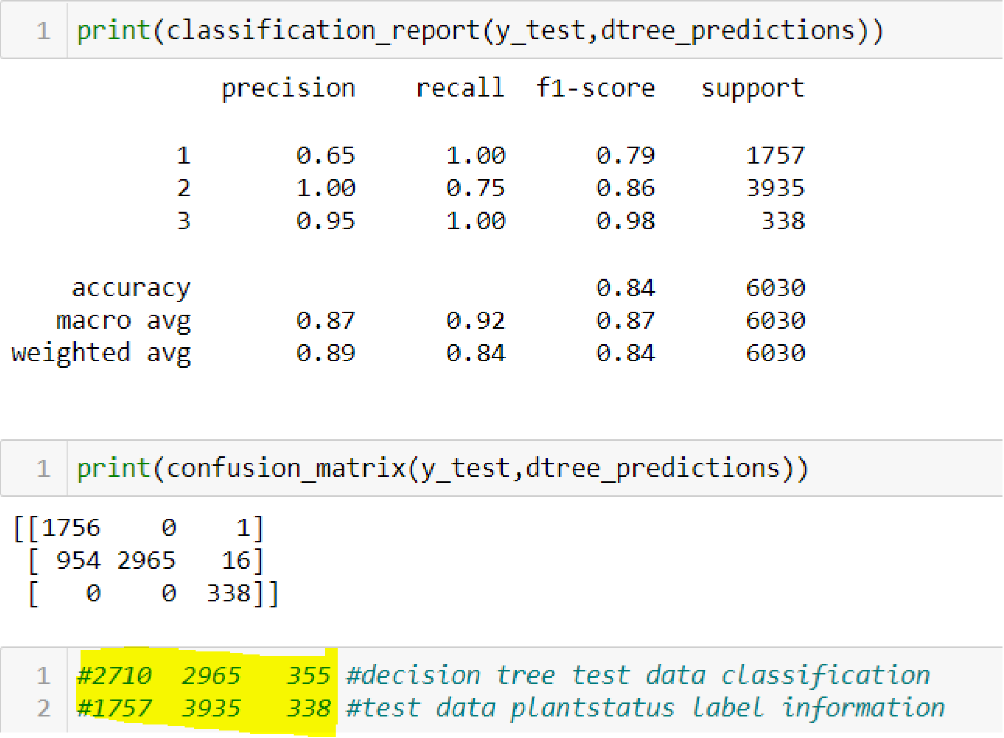
```
print(confusion_matrix(y_test,kn_predictions))
```

```
[[ 805  952    0]
 [    5 3930    0]
 [    0    2  336]]
```

```
# 810  4884   336 #k-nearest neighbor test data classification
#1757  3935   338 #test data plantstatus label information
```

Figure 5.19: Results of K-Nearest Neighbors classification on the test dataset

The third model, Decision Tree (DT), is a supervised learning method used for classification and regression. DT can manage multi-outputs and requires less data preparation (e.g., no need for data normalization). Trees can be visualized, and it is easy to understand and interpret. Figure 5.20 shows the results of DT. In terms of finding every Trip case, the DT model showed better results than LR and KNN. However, many abnormal data entries are classified as normal. The DT showed 84% accuracy, but this can be increased with better model construction.



```
print(classification_report(y_test,dtree_predictions))
```

```
              precision    recall  f1-score   support

           1       0.65      1.00      0.79      1757
           2       1.00      0.75      0.86      3935
           3       0.95      1.00      0.98       338

    accuracy                           0.84      6030
   macro avg       0.87      0.92      0.87      6030
weighted avg       0.89      0.84      0.84      6030
```

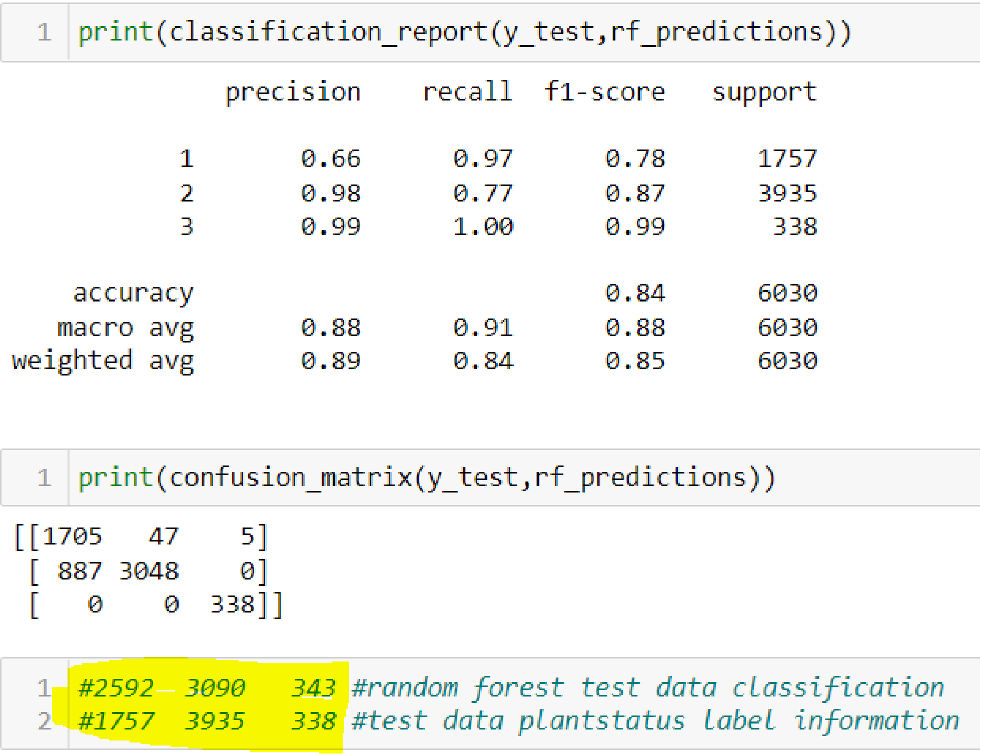
```
print(confusion_matrix(y_test,dtree_predictions))
```

```
[[1756    0    1]
 [ 954 2965   16]
 [   0    0  338]]
```

```
#2710  2965   355 #decision tree test data classification
#1757  3935   338 #test data plantstatus label information
```

Figure 5.20: Results of Decision Tree classification on the test dataset

The fourth model is Random Forest (RF), a supervised machine learning algorithm formed from decision tree algorithms. RF is used for solving regression and classification problems. Figure 5.21 shows the classification results for RF. It successfully found all Trip cases, but it classified Abnormal data entries as Normal data. Also, since there is randomness accuracy score changed after every run. The RF showed 84% accuracy, but this can be increased with better model construction.



```
print(classification_report(y_test,rf_predictions))
```

```
              precision    recall  f1-score   support

           1       0.66      0.97      0.78      1757
           2       0.98      0.77      0.87      3935
           3       0.99      1.00      0.99       338

    accuracy                           0.84      6030
   macro avg       0.88      0.91      0.88      6030
weighted avg       0.89      0.84      0.85      6030
```

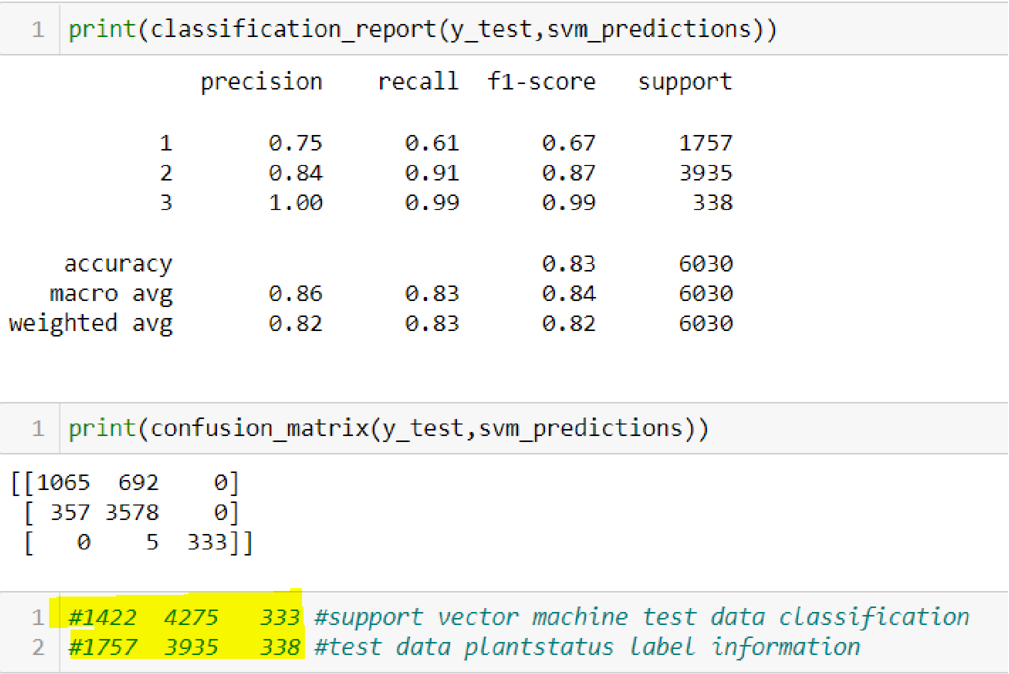
```
print(confusion_matrix(y_test,rf_predictions))
```

```
[[1705   47    5]
 [ 887 3048    0]
 [   0    0  338]]
```

```
#2592   3090     343 #random forest test data classification
#1757   3935     338 #test data plantstatus label information
```

Figure 5.21: Results of Random Forest classification on the test dataset

The fifth ML model used in the first experiment is the Support Vector Machine (SVM). SVMs are a set of supervised learning methods used for classification, regression, and outliers' detection. Figure 5.22 shows the results of using SVM. SVM could not classify every Trip case (333 from 338). The SVM showed 83% accuracy, but this can be increased with better model construction.



```
print(classification_report(y_test,svm_predictions))
```

```
              precision    recall  f1-score   support

           1       0.75      0.61      0.67      1757
           2       0.84      0.91      0.87      3935
           3       1.00      0.99      0.99       338

    accuracy                           0.83      6030
   macro avg       0.86      0.83      0.84      6030
weighted avg       0.82      0.83      0.82      6030
```

```
print(confusion_matrix(y_test,svm_predictions))
```

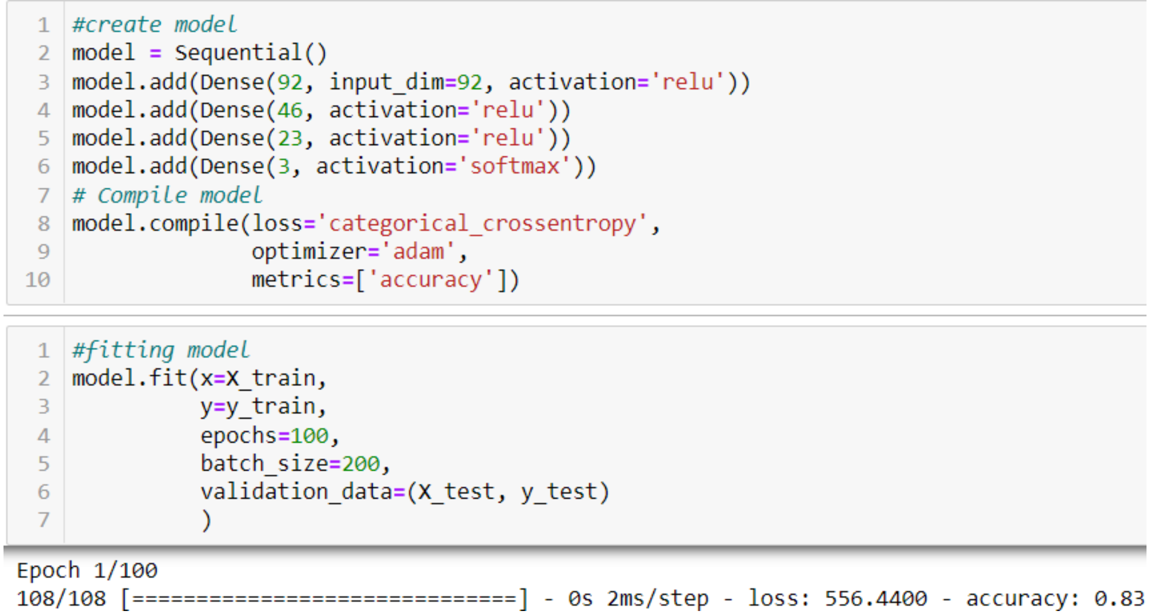
```
[[1065  692    0]
 [ 357 3578    0]
 [   0    5  333]]
```

```
#1422   4275    333 #support vector machine test data classification
#1757   3935    338 #test data plantstatus label information
```

Figure 5.22: Results of Support Vector Machines classification on the test dataset

5.5.2.1. Working with Neural Networks

Artificial neural networks (ANN) simple neural networks are computing systems that are designed based on the brain's working mechanism. An ANN consists of nodes and connections [94]. In this study, we worked with Keras, a neural network application programming interface that runs on the TensorFlow 2 machine learning platform [95]. Keras uses layers, models, optimizers, loss functions, and metrics to build a model. Our data has three classes, so we need a multi-class classification model. Figure 5.23 shows that our data has 92 inputs at the first layer and has three outputs.



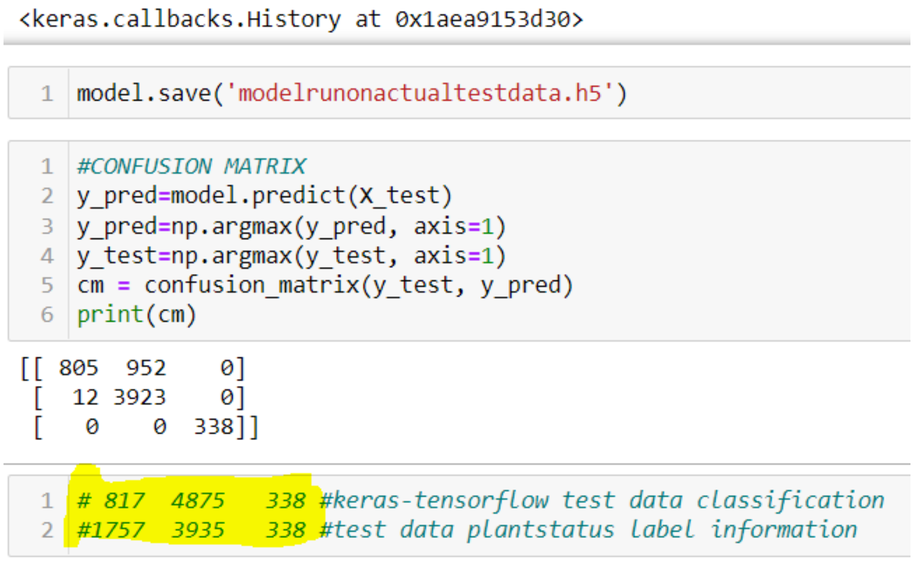
```
#create model
model = Sequential()
model.add(Dense(92, input_dim=92, activation='relu'))
model.add(Dense(46, activation='relu'))
model.add(Dense(23, activation='relu'))
model.add(Dense(3, activation='softmax'))
# Compile model
model.compile(loss='categorical_crossentropy',
              optimizer='adam',
              metrics=['accuracy'])
```

```
#fitting model
model.fit(x=X_train,
          y=y_train,
          epochs=100,
          batch_size=200,
          validation_data=(X_test, y_test)
          )
```

```
Epoch 1/100
108/108 [==============================] - 0s 2ms/step - loss: 556.4400 - accuracy: 0.83
```

Figure 5.23 Keras-TensorFlow model creating and fitting for multi-classification

We ran the model on the test dataset, and it classified every Trip case. Figure 5.24 shows the Keras results on the test dataset.



```
<keras.callbacks.History at 0x1aea9153d30>
```

```
model.save('modelrunonactualtestdata.h5')
```

```
#CONFUSION MATRIX
y_pred=model.predict(X_test)
y_pred=np.argmax(y_pred, axis=1)
y_test=np.argmax(y_test, axis=1)
cm = confusion_matrix(y_test, y_pred)
print(cm)
```

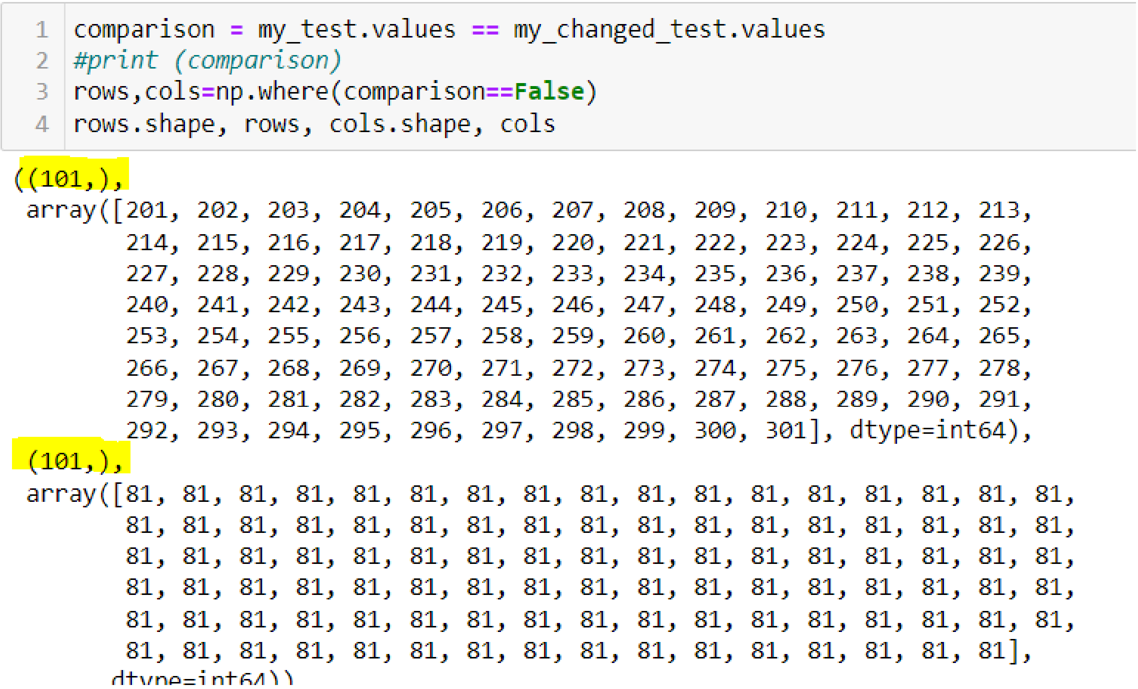
```
[[ 805  952    0]
 [  12 3923    0]
 [   0    0  338]]
```

```
# 817   4875     338 #keras-tensorflow test data classification
#1757   3935     338 #test data plantstatus label information
```

Figure 5.24 Results of Keras-TensorFlow on the test dataset

5.5.3. Working with Artificial Test Dataset

This study aims to detect abnormalities in data, so we wanted to introduce changes artificially into data to observe if the machine learning model catches the changes or not. After working with the actual test dataset, we modified some data properties to create a second artificial test dataset to imitate man in a middle attack or false data injection. We changed the rows from 201 to 301 on the column MBK (Integrated Break Flow). Figure 5.25 shows the difference between the two test datasets.



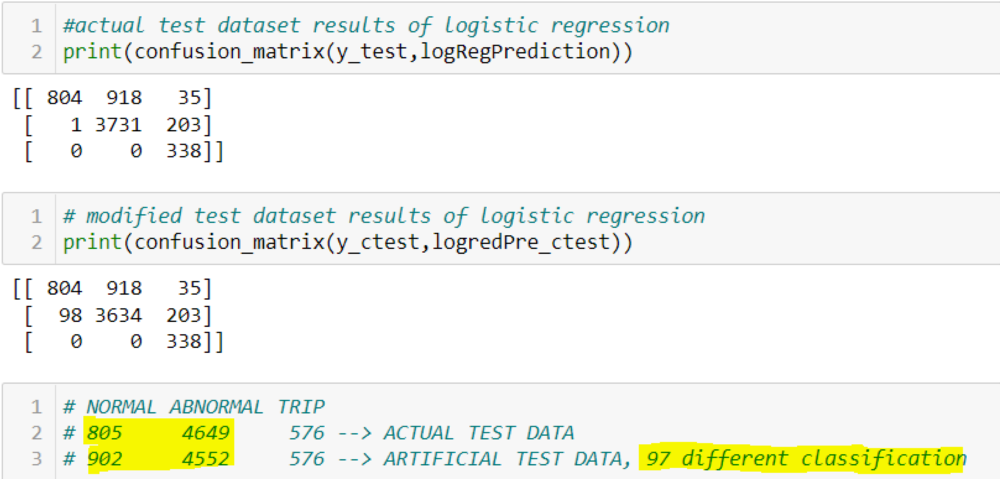
```
comparison = my_test.values == my_changed_test.values
#print (comparison)
rows,cols=np.where(comparison==False)
rows.shape, rows, cols.shape, cols
```

```
((101,),
 array([201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213,
        214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226,
        227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239,
        240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252,
        253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265,
        266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278,
        279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291,
        292, 293, 294, 295, 296, 297, 298, 299, 300, 301], dtype=int64),
 (101,),
 array([81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81,
        81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81,
        81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81,
        81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81,
        81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81,
        81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81, 81],
       dtype=int64))
```

Figure 5.25 The difference between two test datasets

We expected to see different results for the two test datasets and rerun the models.

Results showed that Logistic Regression(LR) created a different number of classes for two test datasets. Between two test datasets, rows from 201 to 301 were changed. LR was able to differentiate 97 of the 101 changes. However, the modified data was classified as normal data while it was abnormal data. Figure 5.26 shows the confusion matrix for actual and artificial test datasets.



```
#actual test dataset results of logistic regression
print(confusion_matrix(y_test,logRegPrediction))
```

```
[[ 804  918   35]
 [   1 3731  203]
 [   0    0  338]]
```

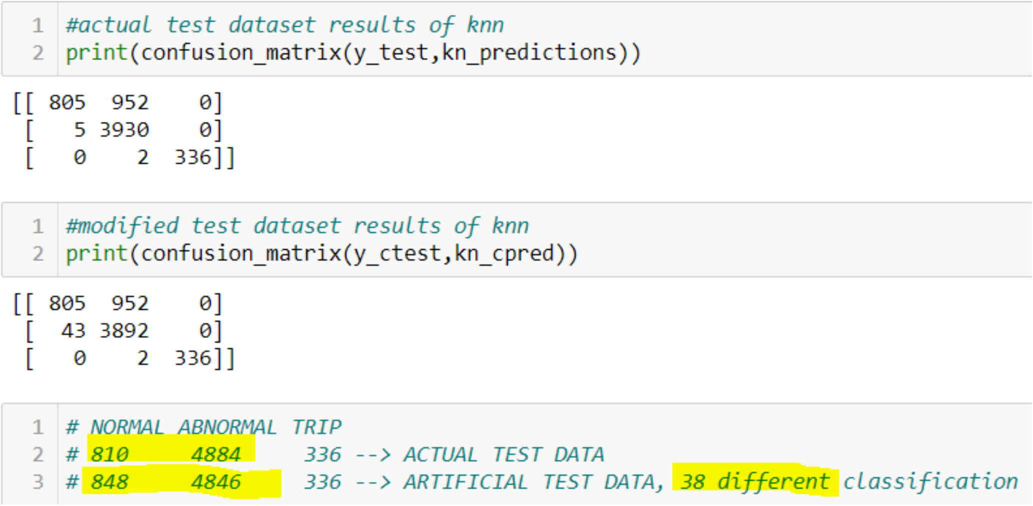
```
# modified test dataset results of logistic regression
print(confusion_matrix(y_ctest,logredPre_ctest))
```

```
[[ 804  918   35]
 [  98 3634  203]
 [   0    0  338]]
```

```
# NORMAL ABNORMAL TRIP
# 805      4649     576 --> ACTUAL TEST DATA
# 902      4552     576 --> ARTIFICIAL TEST DATA, 97 different classification
```

Figure 5.26 Comparison of results in LR on artificial and actual test datasets

Figure 5.27 shows the results of KNN for both test datasets. KNN was able to catch 38 of the 101 changes we introduced, and abnormal data was classified as normal data.



```
#actual test dataset results of knn
print(confusion_matrix(y_test,kn_predictions))
```

```
[[ 805  952    0]
 [   5 3930    0]
 [   0    2  336]]
```

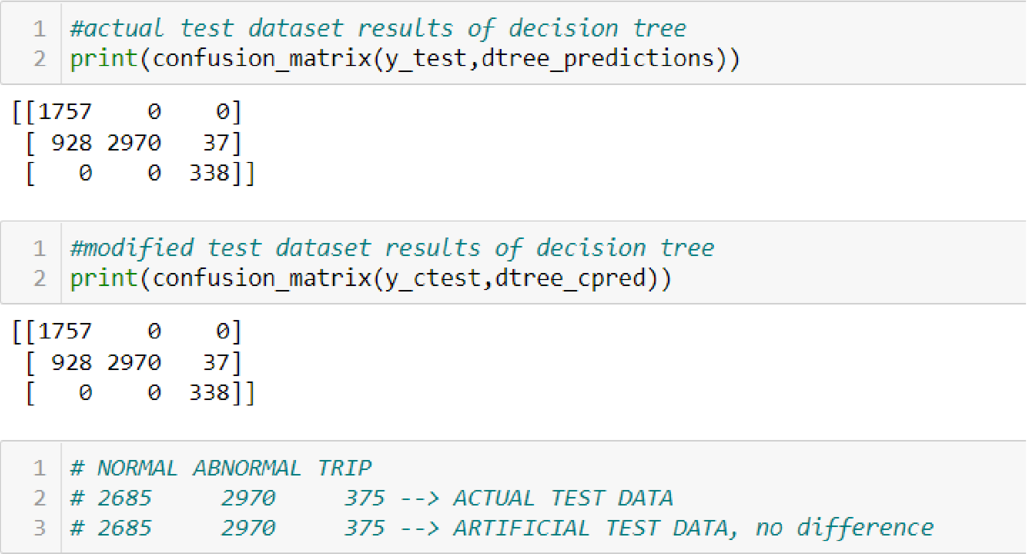
```
#modified test dataset results of knn
print(confusion_matrix(y_ctest,kn_cpred))
```

```
[[ 805  952    0]
 [  43 3892    0]
 [   0    2  336]]
```

```
# NORMAL ABNORMAL TRIP
# 810      4884      336 --> ACTUAL TEST DATA
# 848      4846      336 --> ARTIFICIAL TEST DATA, 38 different classification
```

Figure 5.27 Comparison of results in KNN on artificial and actual test datasets

As shown in figures 5.28 and 5.29, the decision tree and the random forest did not create a different number of classes for both datasets. They could not understand the modified data, but better modeling can improve the results.



```
#actual test dataset results of decision tree
print(confusion_matrix(y_test,dtree_predictions))
```

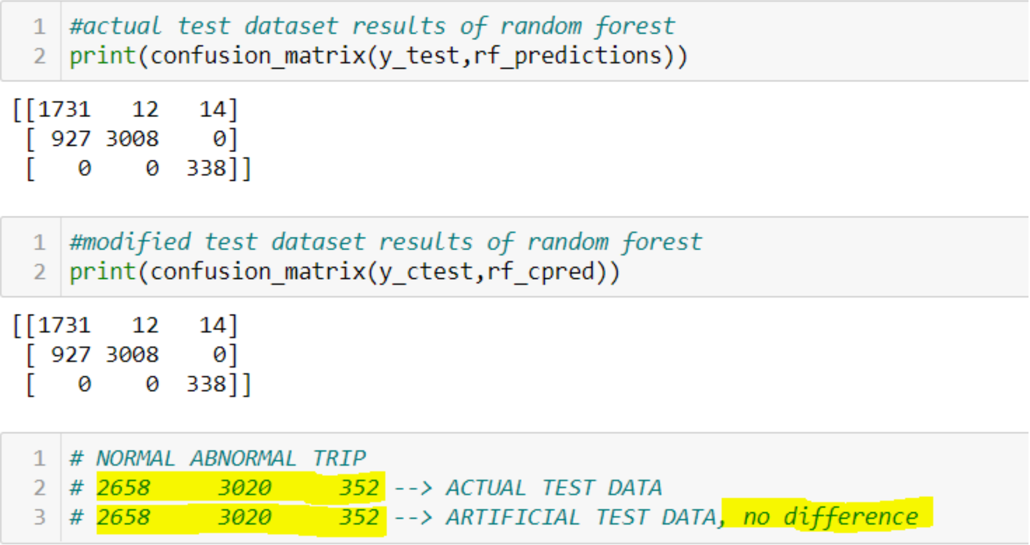
```
[[1757    0    0]
 [ 928 2970   37]
 [   0    0  338]]
```

```
#modified test dataset results of decision tree
print(confusion_matrix(y_ctest,dtree_cpred))
```

```
[[1757    0    0]
 [ 928 2970   37]
 [   0    0  338]]
```

```
# NORMAL ABNORMAL TRIP
# 2685      2970      375 --> ACTUAL TEST DATA
# 2685      2970      375 --> ARTIFICIAL TEST DATA, no difference
```

Figure 5.28 Comparison of results in DT on artificial and actual test datasets



```
#actual test dataset results of random forest
print(confusion_matrix(y_test,rf_predictions))
```

```
[[1731   12   14]
 [ 927 3008    0]
 [   0    0  338]]
```

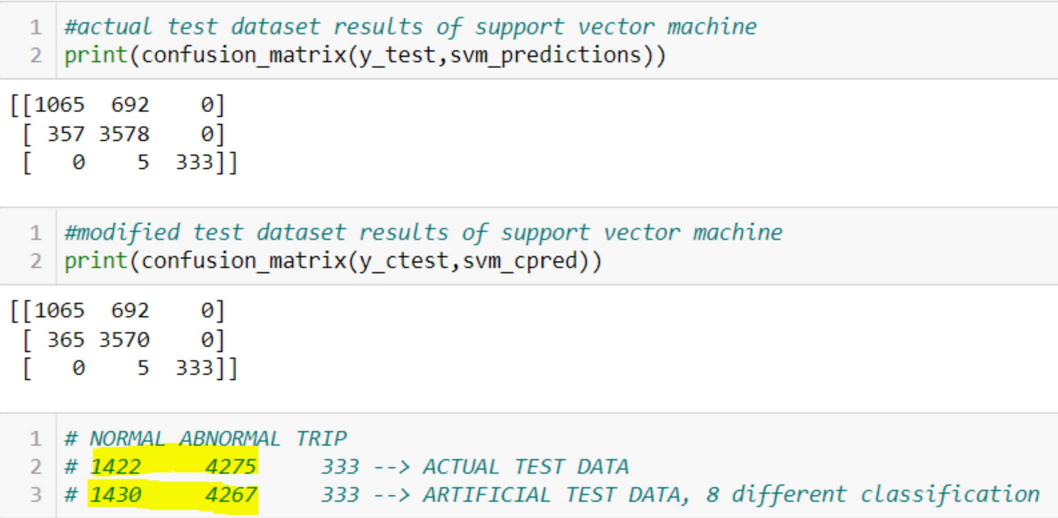
```
#modified test dataset results of random forest
print(confusion_matrix(y_ctest,rf_cpred))
```

```
[[1731   12   14]
 [ 927 3008    0]
 [   0    0  338]]
```

```
# NORMAL ABNORMAL TRIP
# 2658      3020      352 --> ACTUAL TEST DATA
# 2658      3020      352 --> ARTIFICIAL TEST DATA, no difference
```

Figure 5.29 Comparison of results in RF on artificial and actual test datasets

SVM caught eight different classifications out of 101, but better modeling can improve results. Figure 5.30 shows the confusion matrices for both test datasets.



```
#actual test dataset results of support vector machine
print(confusion_matrix(y_test,svm_predictions))
```

```
[[1065  692    0]
 [ 357 3578    0]
 [   0    5  333]]
```

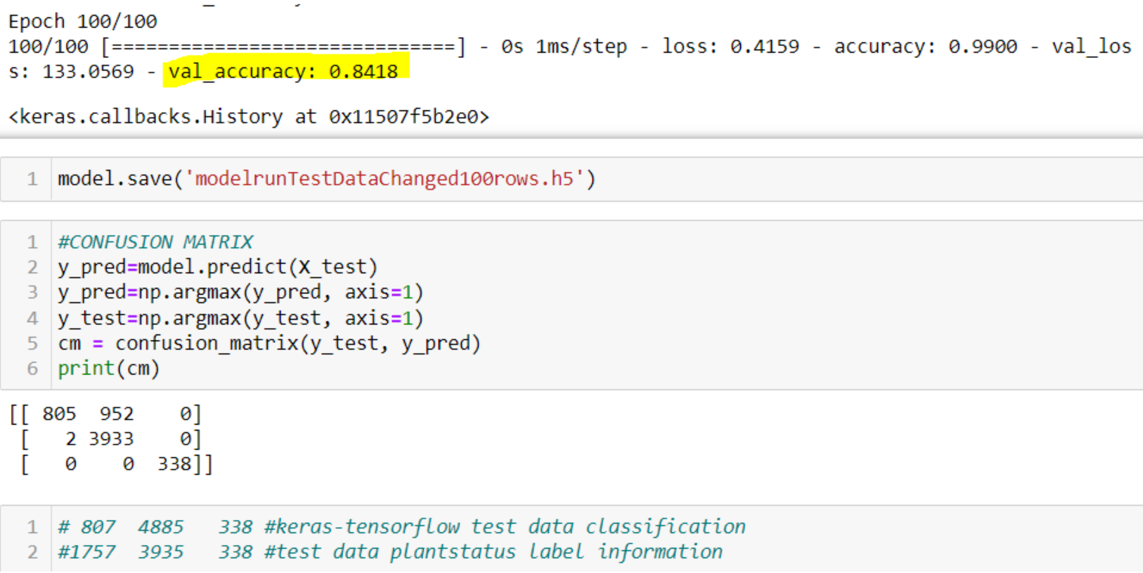
```
#modified test dataset results of support vector machine
print(confusion_matrix(y_ctest,svm_cpred))
```

```
[[1065  692    0]
 [ 365 3570    0]
 [   0    5  333]]
```

```
# NORMAL ABNORMAL TRIP
# 1422      4275      333 --> ACTUAL TEST DATA
# 1430      4267      333 --> ARTIFICIAL TEST DATA, 8 different classification
```

Figure 5.30 Comparison of results in SVM on artificial and actual test datasets

Figure 5.31 shows the results of the Keras API, and it classified eight data entries differently.



```
Epoch 100/100
100/100 [==============================] - 0s 1ms/step - loss: 0.4159 - accuracy: 0.9900 - val_los
s: 133.0569 - val_accuracy: 0.8418

<keras.callbacks.History at 0x11507f5b2e0>
```

```
model.save('modelrunTestDataChanged100rows.h5')
```

```
#CONFUSION MATRIX
y_pred=model.predict(X_test)
y_pred=np.argmax(y_pred, axis=1)
y_test=np.argmax(y_test, axis=1)
cm = confusion_matrix(y_test, y_pred)
print(cm)
```

```
[[ 805  952    0]
 [   2 3933    0]
 [   0    0  338]]
```

```
# 807  4885   338 #keras-tensorflow test data classification
#1757  3935   338 #test data plantstatus label information
```

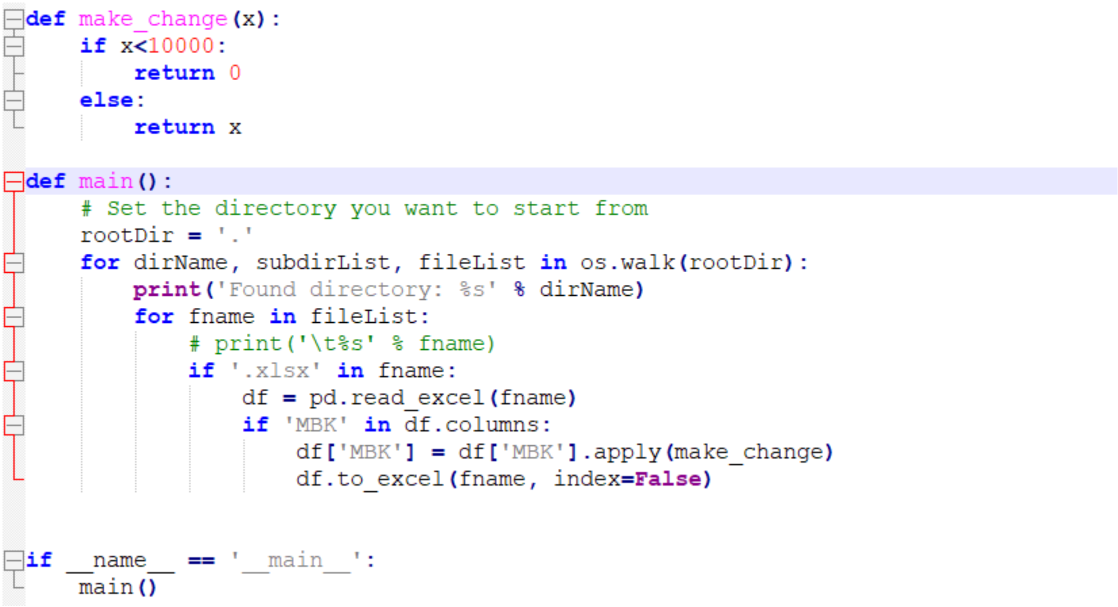
Figure 5.31 Results of Keras on the artificial test dataset

5.6. Attack Scenario

The attacker may be trying to remain in the system without action or take down the system immediately. Based on the attacker's goal, the behavior of the attack would change.

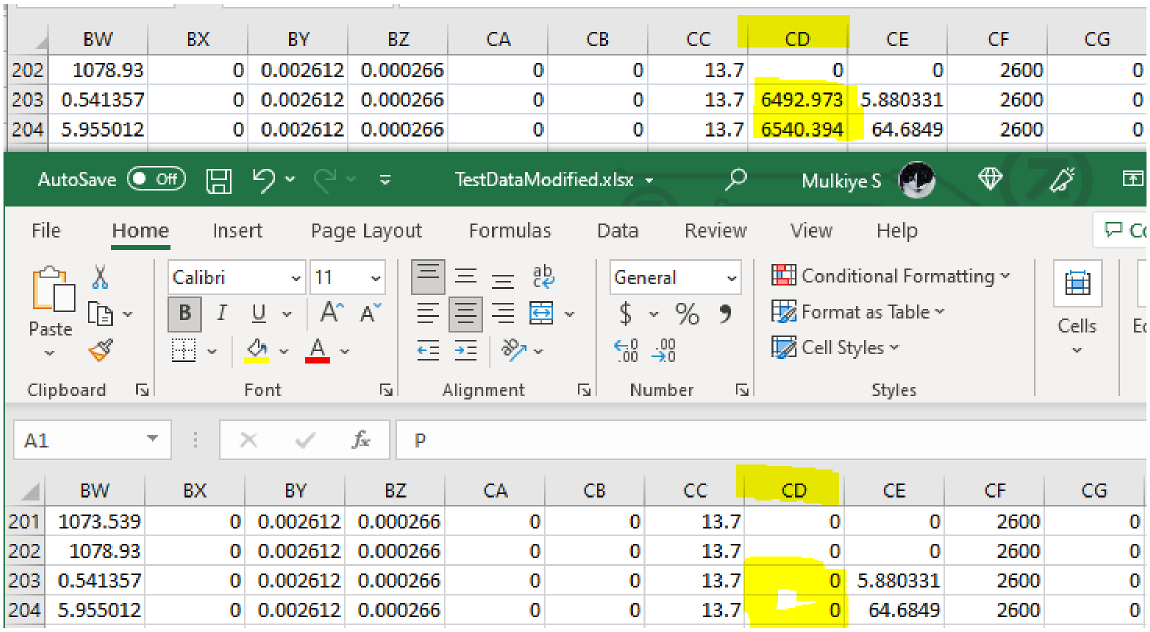
Attackers do not have direct access to the plant systems but may have crucial information about plant data and devices. Thus, the attacker can develop malware that looks for specific information and alter it. Chapter 4 presented a scenario in which the attacker reached the critical internal network using the scenario explained in Chapter 4, and left malware in the systems lurking and looking for its specific target.

The attack scenario we consider in this study is a malware that succeeds in changing the configuration of a device control mechanism such that it alters the component related to the Integrated Break Flow. In our simulated attack, this information is captured by the MBK property. In the simulated attack, the malware is a program (written in a cross-platform language such as Python) that looks for Excel files, searches for the MBK column, and modifies it such that values less than 10,000 are changed to zero; bigger values are not modified. Figure 5.32 shows the malware codes.



```
def make_change(x):
    if x<10000:
        return 0
    else:
        return x

def main():
    # Set the directory you want to start from
    rootDir = '.'
    for dirName, subdirList, fileList in os.walk(rootDir):
        print('Found directory: %s' % dirName)
        for fname in fileList:
            # print('\t%s' % fname)
            if '.xlsx' in fname:
                df = pd.read_excel(fname)
                if 'MBK' in df.columns:
                    df['MBK'] = df['MBK'].apply(make_change)
                    df.to_excel(fname, index=False)


if __name__ == '__main__':
    main()
```

Figure 5.32: Malware that alters MBK column. Figure 5.33 illustrates the result of modifying the MBK column in this way.



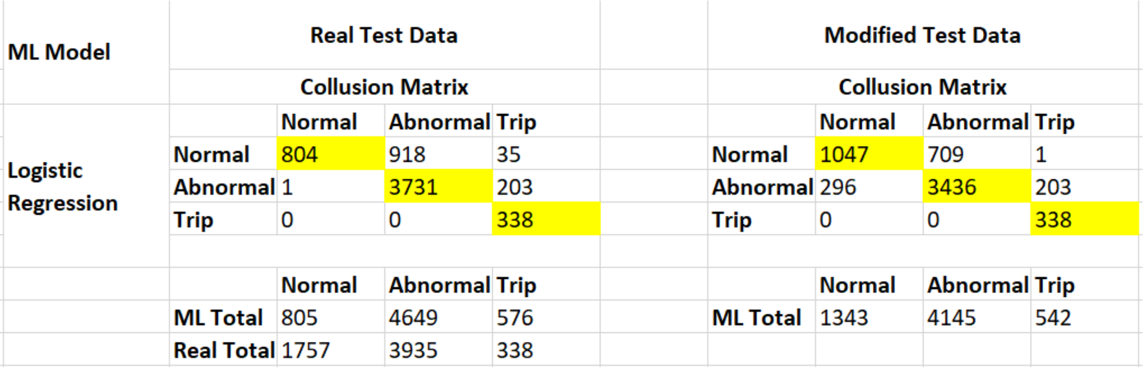
	BW	BX	BY	BZ	CA	CB	CC	CD	CE	CF	CG
202	1078.93	0	0.002612	0.000266	0	0	13.7	0	0	2600	0
203	0.541357	0	0.002612	0.000266	0	0	13.7	6492.973	5.880331	2600	0
204	5.955012	0	0.002612	0.000266	0	0	13.7	6540.394	64.6849	2600	0

A1 | P

	BW	BX	BY	BZ	CA	CB	CC	CD	CE	CF	CG
201	1073.539	0	0.002612	0.000266	0	0	13.7	0	0	2600	0
202	1078.93	0	0.002612	0.000266	0	0	13.7	0	0	2600	0
203	0.541357	0	0.002612	0.000266	0	0	13.7	0	5.880331	2600	0
204	5.955012	0	0.002612	0.000266	0	0	13.7	0	64.6849	2600	0

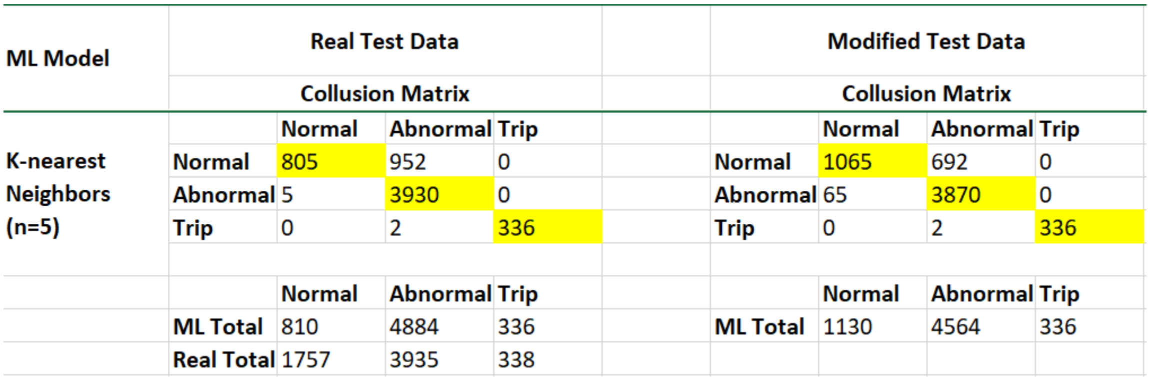
Figure 5.33: MBK column after modification in the simulated attack scenario.

This attack scenario introduces a notable change in the simulated reactor. After running the malware, the ML models should show different results for the original and modified data. Figures 5.34 to 5.39 show the results of applying the ML models on the actual test dataset (collected from the simulator) and the modified test dataset with modified MBK values.



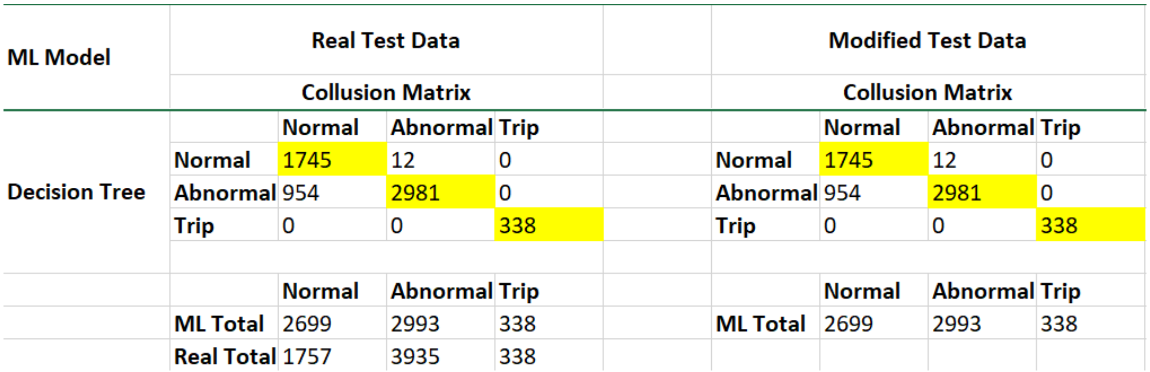
ML Model	Real Test Data					Modified Test Data			
	Collusion Matrix					Collusion Matrix			
		Normal	Abnormal	Trip			Normal	Abnormal	Trip
Logistic Regression	Normal	804	918	35		Normal	1047	709	1
	Abnormal	1	3731	203		Abnormal	296	3436	203
	Trip	0	0	338		Trip	0	0	338
		Normal	Abnormal	Trip			Normal	Abnormal	Trip
	ML Total	805	4649	576		ML Total	1343	4145	542
	Real Total	1757	3935	338					

Figure 5.34: Results of Logistic Regression, contrasting original data with the simulated attack.



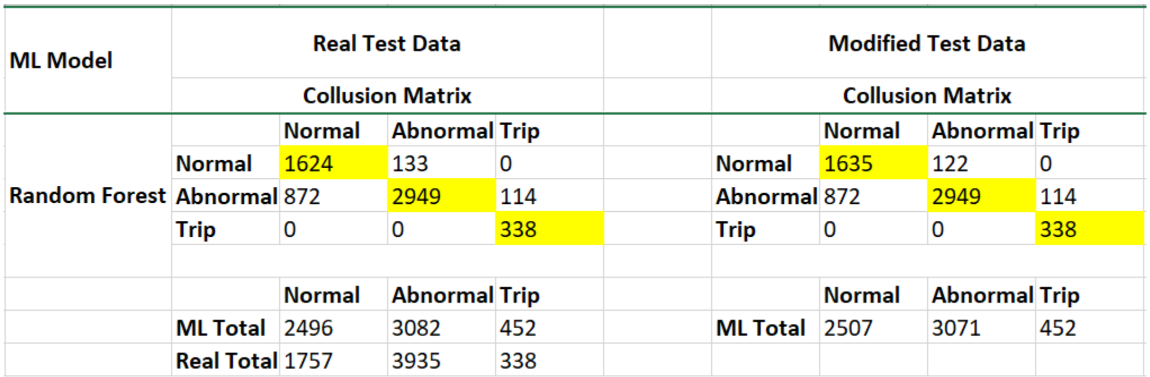
ML Model	Real Test Data					Modified Test Data			
	Collusion Matrix					Collusion Matrix			
		Normal	Abnormal	Trip			Normal	Abnormal	Trip
K-nearest Neighbors (n=5)	Normal	805	952	0		Normal	1065	692	0
	Abnormal	5	3930	0		Abnormal	65	3870	0
	Trip	0	2	336		Trip	0	2	336
		Normal	Abnormal	Trip			Normal	Abnormal	Trip
	ML Total	810	4884	336		ML Total	1130	4564	336
	Real Total	1757	3935	338					

Figure 5.35: Results of K-Nearest Neighbors, contrasting original data with the simulated attack.



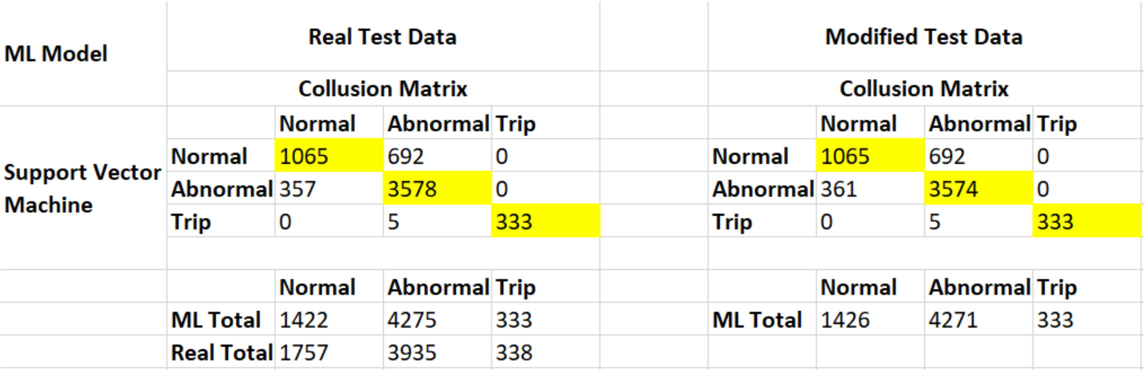
ML Model	Real Test Data					Modified Test Data			
	Collusion Matrix					Collusion Matrix			
		Normal	Abnormal	Trip			Normal	Abnormal	Trip
	Normal	1745	12	0		Normal	1745	12	0
Decision Tree	Abnormal	954	2981	0		Abnormal	954	2981	0
	Trip	0	0	338		Trip	0	0	338
		Normal	Abnormal	Trip			Normal	Abnormal	Trip
	ML Total	2699	2993	338		ML Total	2699	2993	338
	Real Total	1757	3935	338					

Figure 5.36: Results of Decision Tree, contrasting original data with the simulated attack.



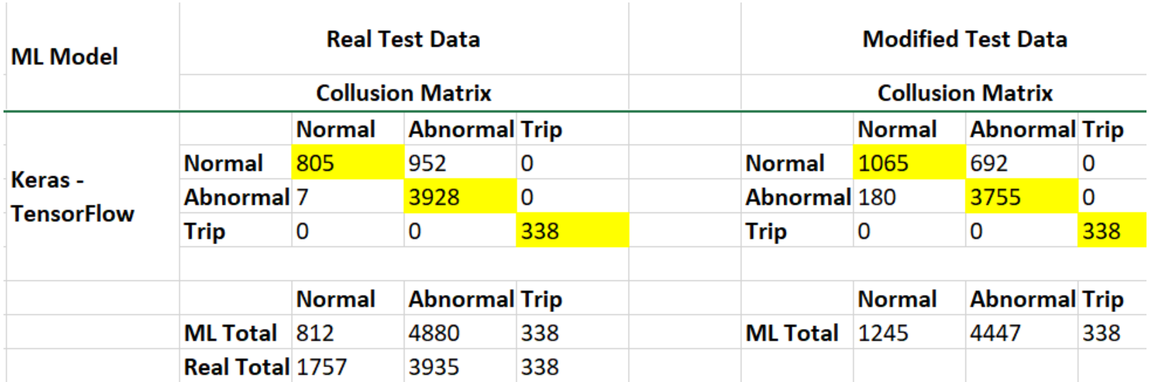
ML Model	Real Test Data					Modified Test Data			
	Collusion Matrix					Collusion Matrix			
		Normal	Abnormal	Trip			Normal	Abnormal	Trip
	Normal	1624	133	0		Normal	1635	122	0
Random Forest	Abnormal	872	2949	114		Abnormal	872	2949	114
	Trip	0	0	338		Trip	0	0	338
		Normal	Abnormal	Trip			Normal	Abnormal	Trip
	ML Total	2496	3082	452		ML Total	2507	3071	452
	Real Total	1757	3935	338					

Figure 5.37: Results of Random Forest, contrasting original data with the simulated attack.



ML Model	Real Test Data					Modified Test Data			
	Collusion Matrix					Collusion Matrix			
		Normal	Abnormal	Trip			Normal	Abnormal	Trip
	Normal	1065	692	0		Normal	1065	692	0
Support Vector Machine	Abnormal	357	3578	0		Abnormal	361	3574	0
	Trip	0	5	333		Trip	0	5	333
		Normal	Abnormal	Trip			Normal	Abnormal	Trip
	ML Total	1422	4275	333		ML Total	1426	4271	333
	Real Total	1757	3935	338					

Figure 5.38: Results of Support Vector Machines, contrasting original data with the simulated attack.

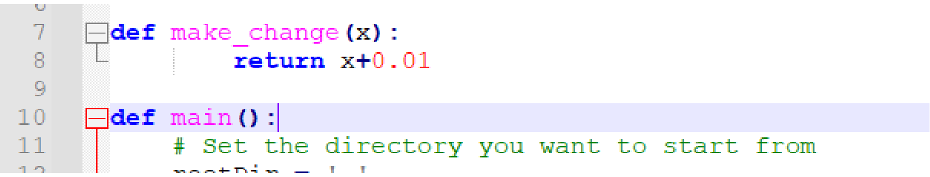


ML Model	Real Test Data					Modified Test Data			
	Collusion Matrix					Collusion Matrix			
Keras - TensorFlow		Normal	Abnormal	Trip			Normal	Abnormal	Trip
	Normal	805	952	0		Normal	1065	692	0
	Abnormal	7	3928	0		Abnormal	180	3755	0
	Trip	0	0	338		Trip	0	0	338
		Normal	Abnormal	Trip			Normal	Abnormal	Trip
	ML Total	812	4880	338		ML Total	1245	4447	338
	Real Total	1757	3935	338					

Figure 5.39: Results of Keras-TensorFlow, contrasting original data with the simulated attack.

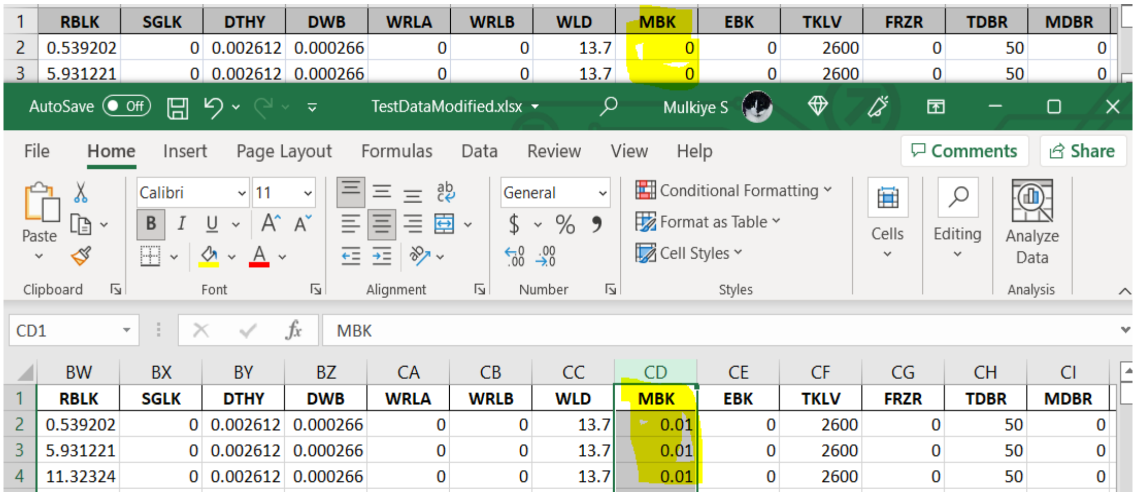
As expected, almost all ML models easily spotted the significant modifications and created different numbers of classes for the plant status as normal and abnormal (which would enable operators to notice that something is amiss).

However, since the attacker’s goal is to remain unnoticed in the system, a more realistic scenario is captured by analyzing the models with slight modifications in the device properties. In this experiment, the malware code adds 0.01 to the MBK values, as shown in Figure 5.40. We did not study the sensitivity of the change in detail since we do not have the information for data sensitivity.



```
def make_change(x):
        return x+0.01

def main():
    # Set the directory you want to start from
```

Figure 5.40: Simulated malware that introduces a minor modification in the MBK values. The result is illustrated in Figure 5.41.



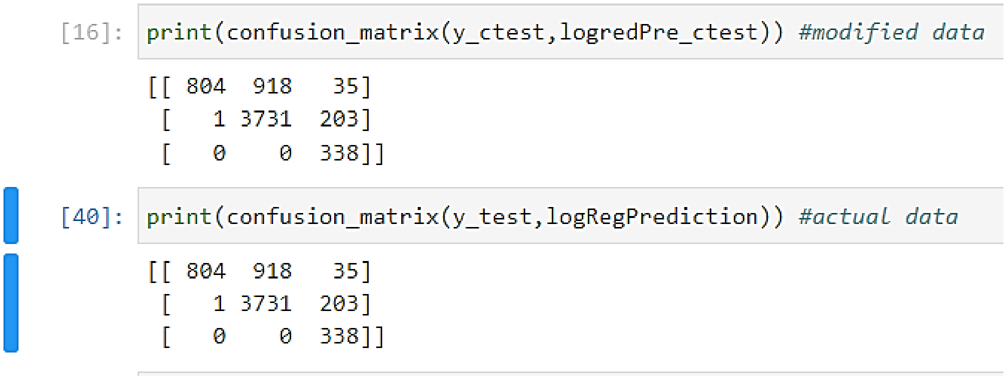
1	RBLK	SGLK	DTHY	DWB	WRLA	WRLB	WLD	MBK	EBK	TKLV	FRZR	TDBR	MDBR
2	0.539202	0	0.002612	0.000266	0	0	13.7	0	0	2600	0	50	0
3	5.931221	0	0.002612	0.000266	0	0	13.7	0	0	2600	0	50	0

CD1 MBK

	BW	BX	BY	BZ	CA	CB	CC	CD	CE	CF	CG	CH	CI
1	RBLK	SGLK	DTHY	DWB	WRLA	WRLB	WLD	MBK	EBK	TKLV	FRZR	TDBR	MDBR
2	0.539202	0	0.002612	0.000266	0	0	13.7	0.01	0	2600	0	50	0
3	5.931221	0	0.002612	0.000266	0	0	13.7	0.01	0	2600	0	50	0
4	11.32324	0	0.002612	0.000266	0	0	13.7	0.01	0	2600	0	50	0

Figure 5.41: MBK column after minor modification

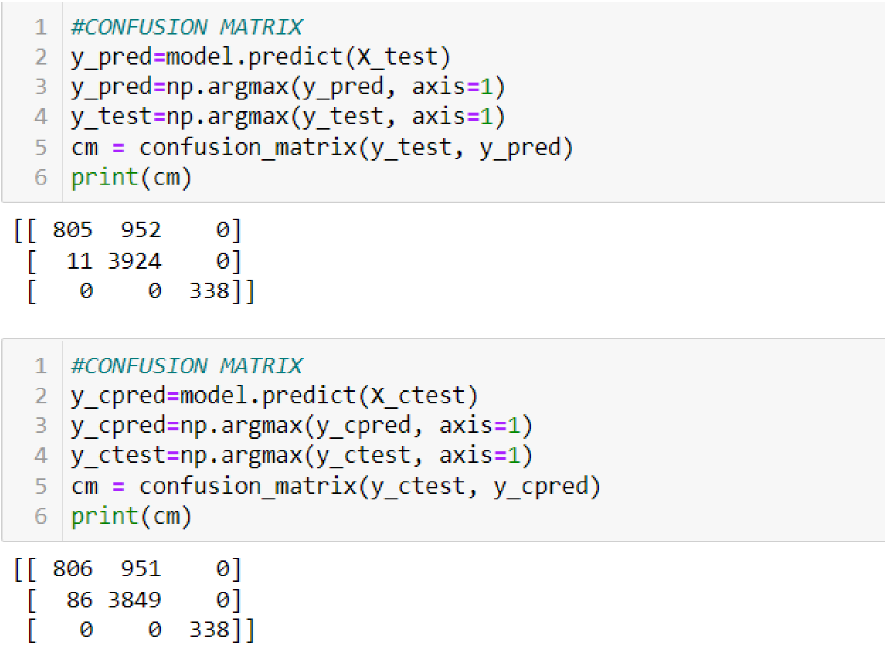
We ran the ML models for this new modified dataset and checked the results on the Confusion Matrix, i.e., the summary of prediction results on a classification. Logistic regression, K-nearest neighbors, decision trees, random forest, and support vector machine models created the same number of classes for both two test datasets. They could not create different results for slight modification. Figure 5.42 shows the results for Logistic Regression on the real and modified test datasets.



```
[16]: print(confusion_matrix(y_ctest,logredPre_ctest)) #modified data

      [[ 804  918   35]
       [   1 3731  203]
       [   0    0  338]]

[40]: print(confusion_matrix(y_test,logRegPrediction)) #actual data

      [[ 804  918   35]
       [   1 3731  203]
       [   0    0  338]]
```

Figure 5.42: Confusion Matrix of LR, contrasting the actual simulated test dataset and modified test dataset with slight modifications in the MBK values.

On the other hand, the Keras-TensorFlow application created a different number of classes for the two test datasets. Figure 5.43 shows the confusion matrix results.



```
#CONFUSION MATRIX
y_pred=model.predict(X_test)
y_pred=np.argmax(y_pred, axis=1)
y_test=np.argmax(y_test, axis=1)
cm = confusion_matrix(y_test, y_pred)
print(cm)
```

```
[[ 805  952    0]
 [  11 3924    0]
 [   0    0  338]]
```

```
#CONFUSION MATRIX
y_cpred=model.predict(X_ctest)
y_cpred=np.argmax(y_cpred, axis=1)
y_ctest=np.argmax(y_ctest, axis=1)
cm = confusion_matrix(y_ctest, y_cpred)
print(cm)
```

```
[[ 806  951    0]
 [  86 3849    0]
 [   0    0  338]]
```

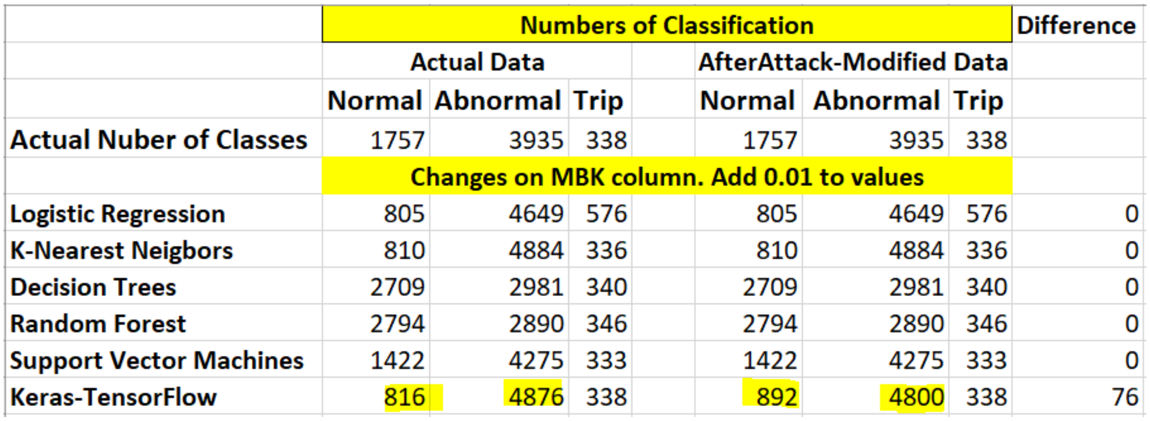
Figure 5.43 Confusion Matrix of Keras-TensorFlow, contrasting the actual simulated test dataset and modified test dataset with slight modifications in the MBK values.

5.7. Discussion

Every SCRAM-plant trip occurrence is significant in the operation of NPPs. It is crucial to have tools that find all trip cases, and no trip status should be missed. On the other hand, is it acceptable to label normal or abnormal data as trip data and cause a trip? Unnecessary or frequent trips to a power plant might cause instability in energy generation and financial damage to the institution. The sensitivity of model design is vital to catch every trip status (i.e., its ability to predict true positives) and reduce the additional classification of trip status (i.e., the number of false negatives).

Furthermore, predicting the Abnormal case is essential. Labeling abnormal data as normal data might be dangerous for a nuclear power plant, missing the opportunity to alert operators for potential danger while labeling Normal data as Abnormal is not as hazardous, but labeling Normal data as Abnormal creates unnecessary work for the plant operators. Studies showed that false alarm affects operators' decision-making and makes operators less aware or concerned about the problem [96]. However, the most critical cyber aspect for nuclear power plants is safety. If there is uncertainty, the worst-case scenario must be considered to avoid big problems. So, if the primary goal is to ensure safety, a higher number of abnormalities should be tolerated.

ML models, except for Keras, could not capture the slight minor changes. This is a success for the attacker, who was able to modify data, remain unnoticed in the system, and wait for an expected condition (e.g., time, user input) so that it can cause severe damage. However, with better model creation and optimization, results can be changed. Figure 5.44 shows the difference between two test datasets focusing on the number of classifications for different machine learning models and Keras API.



	Numbers of Classification							Difference
	Actual Data			AfterAttack-Modified Data				
	Normal	Abnormal	Trip		Normal	Abnormal	Trip	
Actual Nuber of Classes	1757	3935	338		1757	3935	338	
	Changes on MBK column. Add 0.01 to values							
Logistic Regression	805	4649	576		805	4649	576	0
K-Nearest Neigbors	810	4884	336		810	4884	336	0
Decision Trees	2709	2981	340		2709	2981	340	0
Random Forest	2794	2890	346		2794	2890	346	0
Support Vector Machines	1422	4275	333		1422	4275	333	0
Keras-TensorFlow	816	4876	338		892	4800	338	76

Figure 5.44 Numbers of Classes for models after minor changes

The simulated experiments showed that the modified data could go through the monitoring systems and be classified as Normal data, and operators do not have the means to assess the correctness of a prediction quickly. Minor changes in data could indicate an attack.

Identifying unusual patterns in the monitored data from sensors and actuators will help defend against some cyberattack scenarios. If an attacker penetrates the NPP system and then succeeds in controlling an NPP component, analyzing the monitoring data streams may reveal the attacker’s interference before the integrity of the NPP devices is compromised. Detecting anomalies and unintended changes can benefit from the help of an extra inspection tool that runs machine learning models on the monitoring data. Data analytics tools and machine learning models should be part of decision mechanisms with operators' judgment.

6. DISCUSSION AND CONCLUSION

In this work, we explored cybersecurity for nuclear power plants. Our goal was to explore the characteristics of cybersecurity frameworks leading to more secure nuclear power plants and investigate the potential of data analytics to help operators to identify abnormal situations. This chapter summarizes our investigation and derives recommendations based on the studied literature.

As discussed in Chapter 2, even the most secure systems are still vulnerable to cyber-attacks because of people's intentional or unintentional acts. Previous plant incidents indicate that training people is one of the most crucial parts of cyber protection for plants. Still, training is limited in nature, given that nuclear facilities do not share information about plants' structure and their data because of security concerns. As a result, simulators have a massive role in training future power plant workers. Simulators are developed for teaching general usage of nuclear power plants. Learning the basics like water flow, generators, rod position, and power demand is useful, but understanding their technical components is not practical. People who have a background in information technology, security, and computer/electric-electronic fields can get an understanding of how nuclear reactions happen with the currently available simulators. However, the simulators do not convey what kind of monitoring, networking, and control mechanisms are used at nuclear power plants.

6.1. The Role of Data Analytics

This study explored how to use simulators to investigate the potential of conducting data analytics on the NPP's monitored data. We developed a method to bypass the simulator's time limitations to generate datasets corresponding to long periods. By manual manipulation of the simulator's parameters, we generated a representative dataset. The simulator data represents data from many diverse sources. Under real conditions, data from the physical environment (e.g., sensors, valves, generators, pumps) is captured by the monitoring system, and it can be watched/evaluated. In this experiment, data was produced by the simulator through repetitive interaction with the simulator's user interface. The limitations of the simulator make observing the consequences of modified data or changes unfeasible. A change in some values on the simulator ended up with SCRAM/trip or was shown as a malfunction, and further details were unattainable.

Our experiments with extracting data from the simulator revealed that some properties have a crucial effect on the reactor's status. Our analysis, using different machine learning models, showed that for some data modifications, the models are able to detect abnormal and trip situations. This classification can be used to categorize system components based on their effects on the reactor. More importantly, they may allow operators to pursue security measures that can increase NPP safety.

6.2. Overall Cybersecurity Recommendations

Murphy’s law states that “Anything that can go wrong will go wrong!“[97]. The Nuclear power plant administration should accept Murphy’s law, and they should believe security is not guaranteed, but the risk can be mitigated. Cybersecurity risk must be assessed based on worst-case scenarios. The Homeland Security Agency states that the best assessment methodology is the one that promises the highest vulnerability reduction at the lowest cost [82]. While lowering the budget, increasing the security is only possible with a detailed inspection of the systems. With the assistance of experienced personnel, running data analytics tools on whole data (to see which inputs have more effects on the final product of data) will be helpful to decide which systems are the most critical ones. Knowing the most critical systems and their vulnerabilities will lead to good risk management. Aspects to be taken into consideration include:

- NPPs should list or map all the systems and then divide all the systems/assets into groups (e.g., analog systems, converters, carriers, connectors, monitoring, reporting tools, networking, databases, logging tools), create layers, clustering based on different levels of importance.
- NPP should answer the following: What attributes (e.g., confidentiality, integrity, utility, authenticity, non-repudiation) do the assets need? What are security controls (e.g., encryption, hashing, digital signature) required by assets to enable asset attributes? Based on the answer, better risk management can be conducted.
- The Defense in Depth approach emphasizes detecting, preventing, responding, and recovering with different levels of protection. Organizations should apply this approach to the systems and use zone/layer-based networks. Between zones, there should be network firewalls that block the protocols. Any Internet connection to the Industrial Control network should be denied. External storage devices should be blocked, and air gap systems should be hardened. Systems and networks should be monitored for suspicious activity. Activities should be

logged, and extra checks should be conducted on the logged data to ensure security. Table 1 summarizes the controls in this approach.

PREVENT	**DETECT**	**RESPOND/RECOVER**
Blacklists	Anti-virus /anti-spam	Anti-virus-spam
Reputation Systems	Intrusion Detection Systems	Automated response and remediation
Threat Intelligence	Web Application Firewall	Backups
Signature Based Network and Endpoint Methods	Credit Monitoring	Snapshots
Intrusion Prevention Systems	Vulnerability Scanning	Re-imaging
URL-Blockers	Traffic Monitoring	Rollback
Content Filtering	Behavioral Analysis	
Host-Based Firewalls	Anomaly Detection	
File and Disk encryption	Binary Analysis	
Exploit Prevention	Machine Learning	
Sandboxes	Heuristic Detection	
Application whitelisting		
Application control		
File and Disk Encryption		
Access Control List		
User Access Control		
Software Restriction Policy		

Table 1: Example controls for Defense in Depth of Cybersecurity [17] [98]

- Apply the Zero Trust Model that suggests trusting nothing or no user. The risk can be mitigated by distributing the trust. Furthermore, this approach should be applied to every part of the systems and human sources. Hybrid systems(human-machine decision mechanisms) should be used to evaluate, monitor, and make critical decisions.
- The subjects of supply chain assessment, due diligence, trusted foundry, hardware, and software source authentication must be carefully managed, and risk never is accepted as zero.
- Old NPPs should modify their systems by learning from accidents/incidents. For example, Stuxnet showed that the insider threat is real, air gap systems are porous, portable media is dangerous, hard-coded default passwords for devices are unsafe, and digital certificates can be hacked. It also showed that connection between PLCs and computer networks could open digital doors to physical infrastructure. So, PLCs should be assessed for vulnerabilities, and their connectivity to the network should be limited.
- Nuclear power plants can adapt to new job positions where personnel will only work on threat hunting or cyber risk assessment. Threats can be calculated based on the possible answers of who might attack and what their objectives are.
- Personnel should be monitored constantly, and their psychologies and work performance should be evaluated regularly.
- Employee training is one of the most important aspects of power plants' security. Power Plants and educational institutions can work to develop better NPP

simulators that are not only about reactor cores but include the whole plant structure. For example, components from the plant site, connections to PLCs, networking, layering between different zones, and control room can be part of simulators. Virtual twins can be used, but because of security concerns, twins may not be publicly available, so something similar can be developed for training future employees.

REFERENCES

[1] Nuclear Energy Institute, Nuclear Can Help Us Avoid the Worst Case Scenario, August 2021. https://www.nei.org/news/2021/nuclear-can-help-us-avoid-the-worst-case-scenario

[2] World Nuclear Association, Chernobyl 1986, May 2021. https://world-nuclear.org/information-library/safety-and-security/safety-of-plants/chernobyl-accident.aspx

[3] International Atomic Energy Agency, Safety Series No 75-INSAG-7 The Chernobyl Accident: Updating of INSAG-1, 1992, Accessed April 20, 2022. https://www-pub.iaea.org/MTCD/publications/PDF/Pub913e_web.pdf

[4] Nuclear Regulatory Commission, Backgrounder on Chernobyl Nuclear Power Plant Accident, Accessed April 20, 2022. https://www.nrc.gov/reading-rm/doc-collections/fact-sheets/chernobyl-bg.html

[5] Béla Lipták on Safety: Cyber Security and Nuclear Power, Accessed April 20, 2022. https://www.controlglobal.com/assets/wp_downloads/pdf/CT-1611-Bela-Liptak-on-Safety-Cyber-security-and-nuclear-power.pdf.

[6] Du, Wenliang. Computer & Internet Security: A Hands-on Approach, 2019. ISBN:9781733003926, 1733003924

[7] Brooks, R.R. (2013). Introduction to Computer and Network Security: Navigating Shades of Gray (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/b14801

[8] Perlroth, Nicole. 2021. This Is How They Tell Me the World Ends : The Cyberweapons Arms Race. New York: Bloomsbury Publishing USA. ProQuest Ebook Central. ISBN-9781635576054

[9] NIST Special Publication (SP) 800-27 Revision A, Accessed April 20, 2022. https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-27ra.pdf

[10] ISO 31000 Risk Management, Accessed April 20, 2022. https://www.iso.org/iso-31000-risk-management.html

[11] National Vulnerability Database, Accesses April 20, 2022. https://nvd.nist.gov/vuln

[12] The Computing Technology Industry Association (CompTIA), Accesses April 20, 2022. https://www.comptia.org/content/articles/what-is-cybersecurity

[13] AV-TEST (The Independent IT-Security Institute) Malware, February 6, 2022. https://www.av-test.org/en/statistics/malware/

[14] Cyber threat and cyber threat actors, Accessed April 20, 2022. https://cyber.gc.ca/en/guidance/cyber-threat-and-cyber-threat-actors

[15] RICHARD, HUBBARD DOUGLAS W. SEIERSEN. How to Measure Anything in Cybersecurity Risk. S.l.: JOHN WILEY &amp; SONS, 2023.

[16] Security Intelligence, Simplifying Risk Management, Accessed by May 13 2022, https://securityintelligence.com/simplifying-risk-management/

[17] Dion Training,CompTIA® Security+ (SY0-601) training course, Accessed by May 13, 2022 https://www.diontraining.com/comptia-security/

[18] Cybersecurity Information Sharing Act of 2015, https://www.cisa.gov/publication/cybersecurity-information-sharing-act-2015-procedures-and-guidance

[19] Cybersecurity Framework, Accessed April 20, 2022. https://www.nist.gov/cyberframework

[20] The Manhattan Project: Making the Atomic Bomb, Accessed April 20, 2022. https://www.atomicarchive.com/history/manhattan-project/p1s5.html

[21] Einstein's Letter to President Roosevelt - 1939 | Historical Documents. Accessed April 24, 2022. https://www.atomicarchive.com/resources/documents/beginnings/einstein.html.

[22] Atoms for Peace - Wikipedia, Accessed April 24, 2022. https://en.wikipedia.org/wiki/Atoms_for_Peace

[23] International Atomic Energy Agency - Wikipedia, Accessed April 24, 2022. https://en.wikipedia.org/wiki/International_Atomic_Energy_Agency

[24] Atomic Heritage Foundation, Accessed May 10, 2022. https://www.atomicheritage.org/profile/leo-szilard

[25] Nuclear Regulatory Commission, International Organizations, Accessed April 24, 2022. https://www.nrc.gov/about-nrc/ip/intl-organizations.html

[26] International Atomic Energy Agency Nuclear Security Glossary - Terminology Used In IAEA Nuclear Security Guidance, Accessed April 24, 2022. https://www.iaea.org/sites/default/files/21/06/nuclear_security_glossary_august_2020.pdf

[27] Code of Federal Regulations, 10 CFR 73.54 - Protection of Digital Computer and Communication Systems and Networks, Accessed April 24, 2022. https://www.govinfo.gov/app/details/CFR-2012-title10-vol2/CFR-2012-title10-vol2-sec73-54/summary

[28] International Atomic Energy Agency, Safeguards explained, Accessed April 24, 2022. https://www.iaea.org/topics/safeguards-explained

[29] US Energy Information Administration, Electricity Explained, Accessed 24,2022. https://www.eia.gov/energyexplained/electricity/electricity-in-the-us.php

[30] World Nuclear Association, Nuclear Power in France, Accessed April 24, 2022. https://world-nuclear.org/information-library/country-profiles/countries-a-f/france.aspx

[31] How can nuclear combat climate change? Accessed 24, 2022. https://world-nuclear.org/nuclear-essentials/how-can-nuclear-combat-climate-change.aspx#.YmYBamM5Asg.gmail

[32] Advanced Reactors Information System(ARIS), Accessed 24, 2022. https://aris.iaea.org/default.html

[33] Nuclear Regulatory Commission, Pressurized Water Reactors, Accessed 24, 2022. https://www.nrc.gov/reactors/pwrs.html

[34] Nuclear Regulatory Commission, Boiling Water Reactors, Accessed 24, 2022. https://www.nrc.gov/reactors/bwrs.html

[35] International Atomic Energy Agency, DEFENCE IN DEPTH IN NUCLEAR SAFETY - INSAG-10, Accessed 24, 2022. https://www-pub.iaea.org/MTCD/publications/PDF/Pub1013e_web.pdf

[36] Pelindaba - Wikipedia, Accessed 24, 2022. https://en.wikipedia.org/wiki/Pelindaba

[37] Federation of American Scientists, US Nuclear Weapons Site in Europe Breached, Published February 4, 2010, Accessed 24, 2022. https://fas.org/blogs/security/2010/02/kleinebrogel/

[38] The Seattle Times, Breaches at N-plants heighten France's debate over reactors, Published December 21, 2011, Accessed 24, 2022. https://www.seattletimes.com/nation-world/breaches-at-n-plants-heighten-frances-debate-over-reactors/

[39] CBS News, Nun, 84, gets 3 years in prison for breaking into nuclear weapons complex, Published February 18, 2014, Accessed 24, 2022. https://www.cbsnews.com/news/nun-84-gets-3-years-in-prison-for-breaking-in-nuclear-weapons-complex/

[40] Nuclear Regulatory Commission, Backgrounder on Nuclear Security, April 2019, Accessed April 24, 2022. https://www.nrc.gov/docs/ML0500/ML050070043.pdf

[41] Nuclear Energy Institute, Safety: The Nuclear Energy Industry's Highest Priority, June 2015, Accessed 24, 2022. https://www.nei.org/resources/fact-sheets/safety-nuclear-energy-industry-highest-priority

[42] Nuclear Threat Initiative, Nuclear Facilities Face Urgent, Evolving Cyber Threat, December 2016, Accessed 24, 2022. https://www.nti.org/newsroom/news/nuclear-facilities-face-urgent-evolving-cyber-threat/

[43] Federal Bureau of Investigation- Internet Crimes Report, 2020 https://www.ic3.gov/Media/PDF/AnnualReport/2020_IC3Report.pdf

[44] Washington Post, An Indian Nuclear Power Plant Suffered a Cyberattack. Here is What You Need to Know, November 2019 https://www.washingtonpost.com/politics/2019/11/04/an-indian-nuclear-power-plant-suffered-cyberattack-heres-what-you-need-know/

[45] Nuclear Regulatory Commission, Nuclear Energy Institute, Cyber Security Plan for Nuclear Power Reactors, April 2010 https://www.nrc.gov/docs/ML1011/ML101180437.pdf

[46] Langner, Ralph. "To Kill a Centrifuge: A Technical Analysis of What Stuxnet' s Creators Tried to Achieve." (2013). Accessed April 24, 2022. https://www.langner.com/wp-content/uploads/2017/03/to-kill-a-centrifuge.pdf

[47] World Nuclear Association, Fukushima Daiichi Accident, April 2021, Accessed 24, 2022. https://world-nuclear.org/information-library/safety-and-security/safety-of-plants/fukushima-daiichi-accident.aspx

[48] Committee on Lessons Learned from the Fukushima Nuclear Accident for Improving Safety and Security of US Nuclear Plants; Nuclear and Radiation Studies Board; Division on Earth and Life Studies; National Research Council. Lessons Learned from the Fukushima Nuclear Accident for Improving Safety of US Nuclear Plants. Washington (DC): National Academies Press (US); 2014 October 29. Summary. Available from: https://www.ncbi.nlm.nih.gov/books/NBK253923/

[49] International Atomic Energy Agency, Digital Instrumentation and Control Systems for New and Existing Research Reactors, IAEA Nuclear Energy Series No. NR-G-5.1, IAEA, Vienna (2021)

[50] Nuclear Regulatory Commission, Standard Review Plan for the Review of Safety Analysis Reports for Nuclear Power Plants: LWR Edition — Instrumentation and Controls (NUREG-0800, Chapter 7), Accessed 24, 2022. https://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr0800/ch7/index.html

[51] Instrumentation and Controls in Nuclear Power Plants: An Emerging Technologies Update, Accessed by May 13,2022. https://www.nrc.gov/docs/ML0929/ML092950511.pdf

[52] Nuclear Regulatory Commission, Davis-Besse Reactor Vessel Head Degradation Lessons-Learned Task Force Report, 2002 https://www.nrc.gov/reactors/operating/ops-experience/vessel-head-degradation/lessons-learned/lessons-learned-files/lltf-rpt-ml022760172.pdf

[53] U.S. Nuclear Regulatory Commission - Operations Center, Accessed July 2021, https://www.nrc.gov/reading-rm/doc-collections/event-status/event/2011/20110428en.html#en46793

[54] Washington Post, Cyber Incident Blamed for Nuclear Power Plant Shutdown, Published June 5, 2008, Accessed April 24, 2022. https://www.washingtonpost.com/wp-dyn/content/article/2008/06/05/AR2008060501958_pf.html

[55] NBC News, Cyberattack Hits Oak Ridge National Laboratory, Published April 19, 2011, Accessed April 24, 2022. https://www.nbcnews.com/id/wbna42673411

[56] Japan Today, Monju power plant facility PC infected with virus, Published January 7, 2014, Accessed April 24, 2022.

https://japantoday.com/category/national/monju-power-plant-facility-pc-infected-with-virus

[57] Cnet, South Korea nuclear plant hit by hacker, Published December 22, 2014, Accessed April 24, 2022. https://www.cnet.com/tech/services-and-software/south-korea-nuclear-plant-hit-by-hackers/

[58] Ukraine power grid hack - Wikipedia, Accessed April 24, 2022. https://en.wikipedia.org/wiki/Ukraine_power_grid_hack

[59] Reuters, German nuclear plant infected with computer viruses, operator says, Published April 27, 2016, Accessed April 24, 2022. https://www.reuters.com/article/us-nuclearpower-cyber-germany-idUKKCN0XN2OS

[60] New York Magazine, Russian Hackers Infiltrated U.S. Energy Business Networks, Including Nuclear-Power Companies, Published July 9, 2017, Accessed 24, 2022. https://nymag.com/intelligencer/2017/07/russian-hackers-infiltrated-u-s-energy-business-networks.html

[61] 2017 Ukraine ransomware attacks- Wikipedia, Accessed April 24, 2022. https://en.wikipedia.org/wiki/2017_Ukraine_ransomware_attacks

[62] India Today, What is DTrack: North Korean virus being used to hack ATMs to nuclear power plant in India, Published October 30, 2019, Accessed April 24, 2022. https://www.indiatoday.in/india/story/kudankulam-nuclear-power-plant-dtrack-north-korea-atms-1614200-2019-10-30

[63] U.S. Nuclear Weapons Agency Hacked as Part of Massive Cyber-Attack, Published December 17, 2020, Accessed April 24, 2022. https://time.com/5922897/us-nuclear-weapons-energy-hacked/

[64] Reuters, Brazil's Eletrobras says nuclear unit hit with cyberattack, Published February 4, 2021, Accessed April 24, 2022. https://www.reuters.com/business/energy/brazils-eletrobras-says-nuclear-unit-hit-with-cyberattack-2021-02-04/

[65] Nuclear Regulatory Commission, Backgrounder on the Three Mile Island Accident, Accessed April 24, 2022. https://www.nrc.gov/reading-rm/doc-collections/fact-sheets/3mile-isle.html

[66] Jon F. Elliott, The Kemeny Report on the Accident at Three Mile Island, 8 Ecology L. Q. 810 (1980). http://scholarship.law.berkeley.edu/elq/vol8/iss4/11

[67] Nuclear Regulatory Commission, List of Power Reactor Units, Accessed April 24, 2022. https://www.nrc.gov/reactors/operating/list-power-reactor-units.html

[68] Nuclear Regulatory Commission, Regulatory Guide, RG-5.71, January 2010 https://www.nrc.gov/docs/ml0903/ml090340159.pdf

[69] Nuclear Sector: Cybersecurity Framework Implementation Guidance, May 2020, Accessed April 24, 2022. https://www.cisa.gov/sites/default/files/publications/Nuclear_Sector_Cybersecurity_Framework_Implementation_Guidance_FINAL_508.pdf

[70] Poresky, Christopher & Andreades, Charalampos & Kendrick, James & Peterson, Per. (2017). Cyber Security in Nuclear Power Plants: Insights for Advanced Nuclear Technologies. 10.13140/RG.2.2.34430.69449.

[71] Bill Miller and Dale Rowe. 2012. A survey SCADA of and critical infrastructure incidents. In Proceedings of the 1st Annual conference on Research in information technology (RIIT '12). Association for Computing Machinery, New York, NY, USA, 51–56. DOI:https://doi.org/10.1145/2380790.2380805

[72] J. Son, J. Choi, and H. Yoon, "New Complementary Points of Cyber Security Schemes for Critical Digital Assets at Nuclear Power Plants," in IEEE Access, vol. 7, pp. 78379-78390, 2019, Doi: 10.1109/ACCESS.2019.2922335.

[73] Lee, Sangdo, and Jun-Ho Huh. 2019. "An Effective Security Measures for Nuclear Power Plant Using Big Data Analysis Approach." Journal of Supercomputing 75 (8): 4267–94. doi:10.1007/s11227-018-2440-4.

[74] Zhang, Fan, J. Wesley Hines, and Jamie B. Coble. "A Robust Cybersecurity Solution Platform Architecture for Digital Instrumentation and Control Systems in Nuclear Power Facilities." Nuclear Technology 206 (2019): 939 - 950.

[75] Lee, Sangdo, Jun-Ho Huh and Yong Hoon Kim. "Python TensorFlow Big Data Analysis for the Security of Korean Nuclear Power Plants." Electronics 9 (2020): 1467.

[76] PLC-based CYBER-ATTACK detection: A last line of Defence. (n.d.). Retrieved April 20, 2022, from https://conferences.iaea.org/event/181/contributions/15513/attachments/9194/12424/CN278_PLC-based-Detection.pdf

[77] Nuclear Power Plant Personnel Training and its Evaluation A Guidebook, Accessed April 24, 2022. https://www-pub.iaea.org/MTCD/Publications/PDF/trs380_web.pdf

[78] Nuclear Threat Initiative, Cira Mancuso, Nuclear Security is Only as Strong as the Weakest Link: 2020 NTI Index Highlights Cybersecurity and Insider Threat Prevention, August 2020 http://nti.org/7864AP

[79] INTERNATIONAL ATOMIC ENERGY AGENCY, Preventive and Protective Measures Against Insider Threats, IAEA Nuclear Security Series No. 8, IAEA, Vienna (2008). Accessed by May 13, 2022. https://www.iaea.org/publications/7969/preventive-and-protective-measures-against-insider-threats

[80] International Atomic Energy Agency, Management of the nuclear supply chain, Accessed April 24, 2022. https://www.iaea.org/topics/management-systems/management-of-the-nuclear-supply-chain

[81] Akkuyu Nuclear Power Plant, Wikipedia, Accessed 24, 2022. https://en.wikipedia.org/wiki/Akkuyu_Nuclear_Power_Plant#cite_note-14

[82] Homeland Security, Risk Management Fundamentals, April 2011, Accessed April 24, 2022. https://www.dhs.gov/xlibrary/assets/rma-risk-management-fundamentals.pdf

[83] International Atomic Energy Agency, Nuclear Reactor Simulators for Education and Training https://www.iaea.org/topics/nuclear-power-reactors/nuclear-reactor-simulators-for-education-and-training

[84] IAEA, PCTRAN Generic Pressurized Water Reactor Simulator Exercise Handbook, 2019 https://www.iaea.org/publications/13463/pctran-generic-pressurized-water-reactor-simulator-exercise-handbook

[85] MITRE ATT&CK, Accessed 24, 2022. https://attack.mitre.org/

[86] Nuclear Engineering International, NEI Magazine, A new model for nuclear new-build, Accessed by May 13, 2022. https://www.neimagazine.com/features/featurea-new-model-for-nuclear-new-build-8690745/

[87] George C. Marshall, A Dynamic Leader of Transition & Adaptation, Accessed by May 13, 2022. https://digitalcommons.liberty.edu/cgi/viewcontent.cgi?article=1504&context=masters

[88] CISA Adds Two Known Exploited Vulnerabilities to Catalog, February 22, 2022, Accessed 24, 2022. https://www.cisa.gov/uscert/ncas/current-activity/2022/02/22/cisa-adds-two-known-exploited-vulnerabilities-catalog

[89] T. Cheng, Y. Lin, Y. Lai, and P. Lin, Evasion Techniques: Sneaking through Your Intrusion Detection/Prevention Systems, in IEEE Communications Surveys & Tutorials, vol. 14, no. 4, pp. 1011-1020, Fourth Quarter 2012, DOI: 10.1109/SURV.2011.092311.00082.

[90] I. Corona, G. Giacinto, F. Roli, Adversarial attacks against intrusion detection systems: Taxonomy, solutions, and open issues, Inform. Sci. 239 (2013) 201–225.

[91] Russel, Stuart and Norvig, Peter, "Artificial Intelligence: a Modern Approach", Pearson Series in Artificial Intelligence, 2002 (Fourth Edition in 2020), ISBN-13: 978-0134610993

[92] Abu-Mostafa, Yaser S., Malik Magdon-Ismail, and Hsuan-Tien Lin. Learning from data. Vol. 4. New York: AMLBook, 2012. ISBN-13 978-1600490064.

[93] Anaconda Navigator https://docs.anaconda.com/anaconda/navigator/

[94] Artificial Neural Network, accessed May 10, 2022. https://www.search.com.vn/wiki/en/Artificial_Neural_Network

[95] Keras's official website, https://keras.io/about/

[96] Wiczorek R, Meyer J. Effects of Trust, Self-Confidence, and Feedback on the Use of Decision Automation. Front Psychol. 2019;10:519. Published 2019 Mar 12. doi:10.3389/fpsyg.2019.00519

[97] Murphy's law - Wikipedia, Accessed 24, 2022. https://en.wikipedia.org/wiki/Murphy%27s_law

[98] Udemy course, The Complete Cyber Security Course : Hackers Exposed, Nathan House , Accessed by May 13 2022, https://www.udemy.com/course/the-complete-internet-security-privacy-course-volume-1/

[99] Energy Information Administration, Nuclear explained- U.S. nuclear industry, Accessed 24, 2022. https://www.eia.gov/energyexplained/nuclear/us-nuclear-industry.php