Ministry of higher education scientific and research

University of Djillali BOUNAAMA - Khemis Miliana

Faculty of science and technology

Department of Electrical Engineering

Master Degree's dissertation

Option

INDUSTRIAL ELECTRICAL ENGINEERING

Theme

Identification of PV Fault classes using intelligent method KNN (K-Nearest neighbors)

Prepared By:

Taane Zahreddine

Godfrey Benjamin Zulu

Supervisor:

Dr. C. Kara Mostefa Khelil

Dr. I. Hamzaoui

2022 / 2023

Dedication by Taane Zahreddine

I dedicate my dissertation work to my family and friends.

A special feeling of gratitude to my loving parents Taane Abdelghani and Cherrak Radia whose words of encouragements and push for tenacity ring in my ears. My brothers and sisters have never left my side and are very special.

Lastly I dedicate this work and give special thanks to my friend Baghdad Ali for his morale encouragement and for all the support given during this year.

Acknowledgement by Taane Zahredine

In the accomplishment of this work successfully, many people have best owned upon me their blessings and the heart pledged support, this time I am utilizing to thank all the people who have been concerned with this project.

Primarily I would thank God for being able to complete this project with success. Then I would like to thank Dr.C. Kara Mostefa Khelil and Dr.I.Hamzaoui, whose valuable guidance has been the ones that helped me patch this work and make it full proof success.

DEDICATION BY Godfrey Zulu

I would like to dedicate this humble work to my God Jehovah and to my parents Eunice Bwalya Chayo and Paul Benjamin Zulu who gave me life and supported me throughout my life. Am forever grateful to be your son and I shall forever remain indebted to you. My success today is due to your kindness, motivation and prayers that I am where I am today.

Am not forgetting my younger brother Arpachshad Benjamin Zulu and my uncle Hananiah Shadric Chayo for all the support given during the six years I have be away from home. And lastly I dedicate this work to all who might find it insightful and useful in the academic arena.

ACKNOWLEDGEMENTS BY Godfrey Zulu

We would like to express my heartfelt gratitude and respect to Dr. C. Kara Mostefa Khelil and Dr.I.Hamzaoui for their guidance, assistance, support and encouragement while preparing our dissertation. Her sage advice and useful suggestions have enabled us to complete our dissertation.

I am equally thankful to the almighty Jehovah for giving me an opportunity to have higher education, and whose blessings I anticipate shall continue pouring on me. I am truly thankful to all the professors of the department of electrical engineering at Djilali Bounaama de Khemis Miliana for their assistance throughout the many years I spent here. I am forever indebted to my mother and father for continually inspiring me and whose prayers and advice helped me sail through my six years of university.

ABSTRACT:

Throughout many developing nations of our humble planet, renewable energy is a hot topic. Every country at this very moment is trying to move away from fossil fuels like petrol to complete renewable energy sources especially Photovoltaic systems.

The reliability and efficiency of photovoltaic energy systems is now a frequent topic of discussion. Like all systems of production, photovoltaic systems are subject to failures and defects in their normal operating functions with regards to the amount of power output. These systems break down and deteriorate during their lifetime. This is why a system of diagnostic is required whose many objectives is to provide indicators with the given valuables like temperature, solar irradiation, voltage and current output to detect the faults and thus maintain the energy production at optimum.

The work in progress relates to the diagnostic of faults in the PV systems using artificial intelligent methods particularly the K-nearest Neighbors algorithm.

Keywords: Photovoltaic generator, faults, detection, diagnostic, identification, KNN

ملخص :

في العديد من البلدان النامية على كوكبنا المتواضع ، تعد الطاقة المتجددة موضوعًا ساخنًا. تحاول كل دولة الآن الابتعاد عن الوقود الأحفوري مثل النفط لتكملة مصادر الطاقة المتجددة ، وخاصة الأنظمة الكهروضوئية.

أصبحت موثوقية وكفاءة أنظمة الطاقة الكهروضوئية الآن موضوعًا متكررًا للمناقشة. مثل جميع أنظمة التوليد ، تخضع أنظمة الطاقة الكهروضوئية للفشل والأعطال في وظائف التشغيل العادية فيما يتعلق بكمية الطاقة المنتجة. تتعطل هذه الأنظمة وتتدهور خلال فترة تشغيلها. هذا هو السبب في أن نظام التشخيص ضروري ، وتتمثل أهدافه العديدة في توفير مؤشرات مع القيم المعطاة مثل درجة الحرارة ، وضوء الشمس ، والجهد ، وتيار الخرج لاكتشاف الأعطال وبالتالي الحفاظ على إنتاج الطاقة.

يركز العمل الجاري على تشخيص الأعطال في الأنظمة الكهروضوئية باستخدام طرق الذكاء الاصطناعي ، ولا سيما خوارزمية K- الجيران الأقرب.

الكلمات المفتاحية: المولد الكهروضوئي ، الأعطال ، الكشف ، التشخيص ، KNN

Résumé :

Dans de nombreux pays en développement de notre humble planète, les énergies renouvelables sont un sujet brûlant. Chaque pays essaie en ce moment même de s'éloigner des combustibles fossiles comme le pétrole pour compléter les sources d'énergie renouvelables, en particulier les systèmes photovoltaïques.

La fiabilité et l'efficacité des systèmes photovoltaïque est maintenant un sujet de discussion fréquent. Comme tous les systèmes de production, les systèmes photovoltaïques sont exposés à des défaillances et à des défauts dans leur principe de fonctionnement normal en ce qui concerne la quantité de puissance produite. Ces systèmes tombent en panne et se détériorent pendant leur durée de vie. C'est pourquoi un système de diagnostic est nécessaire dont les nombreux objectifs sont de fournir des indicateurs avec les valeurs données comme la température, l'ensoleillement, la tension et le courant de sortie pour détecter les défauts et ainsi maintenir la production d'énergie à son optimum.

Le présent travail se focalise sur le diagnostic des défauts dans les systèmes PV à l'aide de méthodes intelligentes artificielles, en particulier l’algorithme des K plus proches voisins.

Mots clés : Générateur photovoltaïque, défauts, détection, identification, diagnostique, KNN.

ABREVIATIONS LIST

PV: Photovoltaic

KNN: K-nearest Neighbor

AC: Alternating Current

DC: Direct Current

AGM: Absorbed Mat Sealed Lead Acid

GEL: Gelled Electrolyte Sealed Acid

FDD: Fault detection and Diagnosis

PVS: Photovoltaic systems

PVA: Photovoltaic arrays

MPP: Point of Maximum Power

FLC: Fuzzy logic control

GPV: Photovoltaic Generator

ANN: Artificial Neuron Network

BNN: Bayesian Neural Network

MPPT: Maximum power point tracker

SVMs: Support vector machines

T: Temperature

V: Voltage

List of contents

LIST OF FIGURES

LIST OF TABLES

GENERAL INTRODUCTION

GENERAL INTRODUCTION

Since the dawn of civilization, mankind has looked at the sky and wondered if it's possible to obtain energy from the stars especially our nearest star, the sun. So as we progressed as a civilization, we begun to use different energy sources to help with our daily energy consumption. Among these energy sources were fossil fuels like petrol, diesel, hydro carbons and natural gas.[1]

For a time being we thrived on fossil fuels and hydro carbons until we discovered they were not the best of energy sources. They polluted and still continue to pollute our environment by the emission of carbon dioxide (Co_2) which causes the greenhouse effect. [1]

Since 1990s, Man has worked hard to discover various ways we to harness the solar energy from the sun and the most prominent way he has done this is through Photovoltaic systems. Today PV systems are everywhere and continue to be a subject of discussion in the aspect of renewable energy. [2]

Despite the PV systems having high efficiency rate, ideal for the environment and ease of use, they are often accompanied by system defaults which may not always be detected on time hence the use of Artificial intelligence algorithms like KNN, Naïve Bayes and probabilistic methods to classify different faults. In our work, we will focus more on KNN algorithm and how we can apply it to categorize different faults detected in PV systems at last give an example of a Matlab simulation of the system. [3]

Our work includes four chapters that explain in detail the PV class systems, PV faults, different methods used to detect faults, the KNN algorithm and the conclusion from the simulation using a PV setup of the parameters of our choosing in comparison to already existing ones.

CHAPTER ONE

PV SYSTEMS

I.1 Introduction

Energy is an important aspect in our modern society. Without energy, it is nearly impossible to advance in any aspect as a civilization. While preparing this thesis, we considered Algeria and Zambia. These countries are among the countries in Africa considering taking a step into prioritizing solar energy into their respective energy sectors.

A photovoltaic system, also PV system or solar power system, is an electric power system designed to supply usable solar power by means of photovoltaic. It consists of an arrangement of several components, including solar panels to absorb and convert sunlight into electricity, a solar inverter to convert the output from direct to alternating current, as well as mounting, cabling, and other electrical accessories to set up a working system. It may also use a solar tracking system to improve the system's overall performance and include an integrated battery.

PV systems convert light directly into electricity, and are not to be confused with other solar technologies, such as concentrated solar power or solar thermal, used for heating and cooling. [4]

I.2 The different types of photovoltaic systems

I.2.1 Installation of autonomous PV systems

The role of autonomous systems is to power or feed one or many consumers situated in an isolated zone of the electrical network. The photovoltaic field can need the direct electrical energy to power the receptors (lighting, domestic equipment etc.)

The regulation system is a battery of accumulators which allows to store electrical energy which can later be utilized in the absent of solar energy. The batteries are used to store electrical energy in form of chemical energy. They convert back the energy into electrical form whenever it is needed. The charge controller`s principal function is to protect the battery from overload and high discharge. It is an essential element in the lifespan of the battery. In the isolated sites, we can use receptors that work on alternating current. In such a case, the installation contains an inverter.

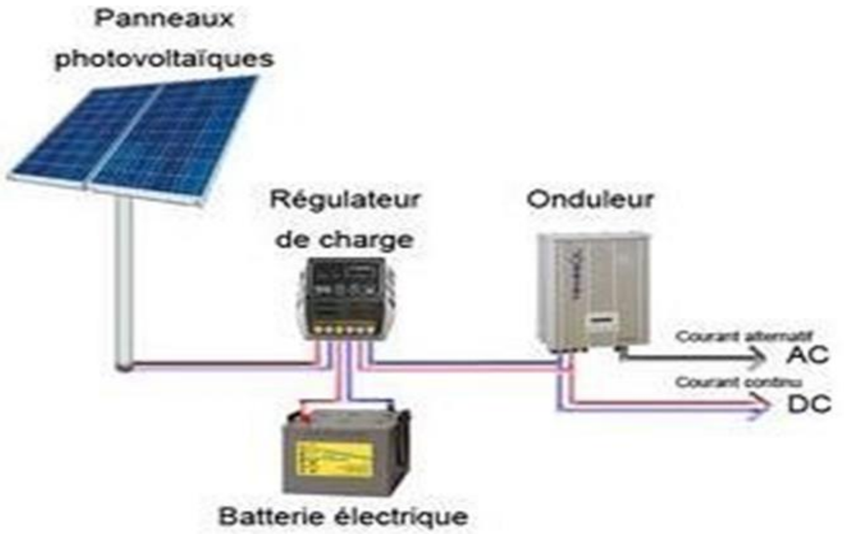

Figure I.1: Example of an autonomous PV system [4]

I.2.2 PV system connected to the power grid

Such a system is installed on a site that is connected to the electrical supply. (Sonelgaz in Algeria and Zesco in Zambia) This is generally applied in domestic or industrial installations that resort to renewable energy which is as a result of abundant sunshine in that location e.g. the Sahara desert in Algeria.

A photovoltaic generator is connected to the electrical power grid which implies the need to store the electrical energy generated is not necessary and saves a lot of money from resulting problems of finding means of storage of the power. It is therefore the grid which acts as the energy storage. There are two ways the PV electrical energy is connected to the grid:

- Either inject the entire produced energy into the grid
- Or inject surplus of the produced energy into the grid

Two energy meters are necessary:

A meter that quantifies the amount of energy bought and used by the consumers and the meter that measures the energy transmitted back to the electrical grid when the production becomes greater than consumption. The third meter is added in the event were the energy produced is injected entirely to the power grid.

The power inverter is used to convert the direct current (DC) of the solar panels into alternative current (AC) and this has to be homologue or equivalent to the standards of the electrical distributor company of the electricity which will receive the current at the hope of making it sinusoidal. [5]

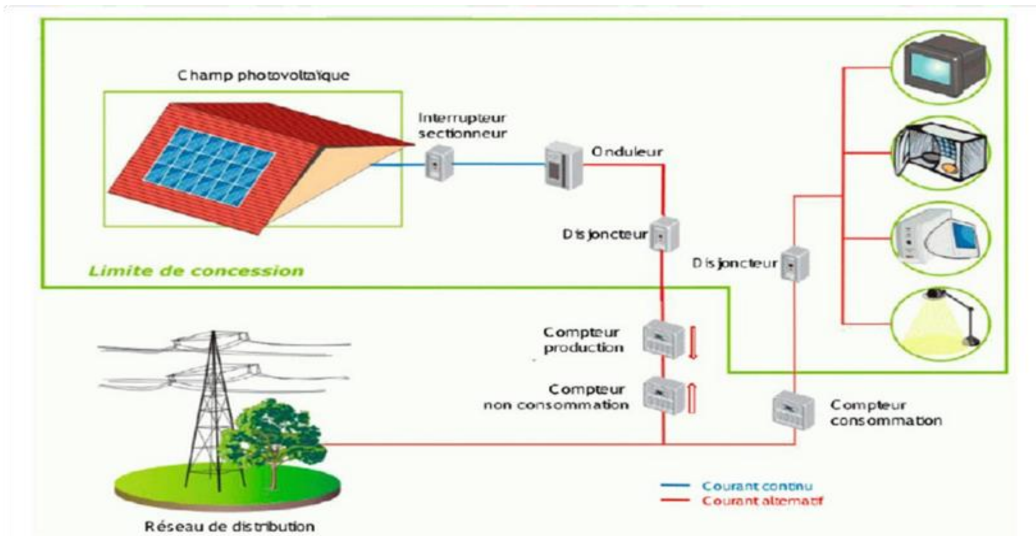

Figure I.2: Example of PV system connected to the power grid [5]

I.2.3 Photovoltaic hybrid installation

This is a system which comprises of groups of different energy sources of different nature such as wind farm, diesel generator, central cogeneration and PV generator. This type of installation is used when the PV generator cannot provide all the necessary energy required. [6]

Figure I.3: Example of photovoltaic hybrid installation [6]

I.3 Study of the components that make up the PV generator

I.3.1 The Photovoltaic Generator

The main element of the system that converts sunlight into electrical energy is PV Generator. In 1950s, PV Generator had about 4% efficiency and today it is about 15% system efficiency. Its efficiency is calculated according to 1000 W/m2 of energy is emitted by the sun and in Turkey this value has more efficiency and it is about 1300W/m2. [7]

In addition, the sunshine duration for Turkey conditions has more experienced variability. In winter, it is approximately 5 hours, 7 hours in the fall and summer is around 11 hours. To get maximum efficiency, the system needs optimum environments and conditions. However, to get more efficiency from this system, the placement of PV panels is important factor. [7]

I.3.2 Charge Regulator

You don’t need a charge controller if you don’t have batteries for storage–i.e., if your electricity goes straight into the grid and isn’t a hybrid system. The charge controller protects your batteries by monitoring how much they’re being charged and slowing or even halting flow to avoid overcharging. This helps the batteries last longer so that they can efficiently store what you generate. The charge regulator from the PV panel current (energy) allows the transmission to the battery, the battery fully discharged. Charge regulators are chosen according to the system used (12A, 24A, 48A or 10V, 20V, 40V, 60V) and it must be stable to the maximum current. [8]

I.3.3 The inverters

Inverters convert the DC current from solar panels and batteries into AC for use by a house or electrical grid. If the electricity is to be used by the grid, you need a special grid-tie inverter. The grid-tie inverter matches the phase of the electricity you generate to the phase of the grid. Without it, the mismatched phases would effectively cancel one another out and be useless. [9]

The grid-tie inverter also acts as a safety device; if there is a power cut, it will immediately switch off the solar panel supply so that the electricity you generate doesn’t harm a worker trying to fix the power lines. [9]

I.3.4 The Battery System

Generally, it is used off-grid or hybrid photovoltaic with using of a bank of deep-cycle batteries to store electricity from solar panels. Electrical energy is produced by solar panels during the day/night without interruption is possible to use the proper battery bank with its establishment. With serial or parallel connection of batteries, system can work 12, 24, 48 or 120V voltage. It is generally three types such as Flooded Lead Acid, Absorbed Glass Mat Sealed Lead acid (AGM) and Gelled Electrolyte Sealed Lead Acid (GEL). But in photovoltaic systems, flooded lead acid is preferred. This type of batteries does not need to add distilled water and electrolyte, charge-discharge kinetics of long lived and very productive. [9]

Furthermore, such products are available in different locations at the same time is ideal for applications that require high-performance solutions. Batteries should be directed to the system to be established during the selection of it. For example, during a day, a gel battery-powered lamps should be chosen for street lighting or garden. Because acid batteries do not need appropriate choice due to make a care four or six months (the addition of distilled water). Flooded lead acid batteries can be used if there are only solar applications and if it will be maintained. [9]

Detecting faults in PV is important for the overall efficiency and reliability of a solar power plant. Ground faults, series and parallel arc faults, high resistance connections, soiling, and partial shadowing need to be detected. [9]

CHAPTER TWO

PV FAULTS, DETECTION AND DIAGNOSIS

II.1 Introduction

Faults in any components (modules, connection lines, converters, inverters, etc.) of photovoltaic (PV) systems (stand-alone, grid-connected or hybrid PV systems) can seriously affect the efficiency, energy yield as well as the security and reliability of the entire PV plant, if not detected and corrected quickly. In addition, if some faults persist (e.g. arc fault, ground fault and line-to-line fault) they can lead to risk of fire. Fault detection and diagnosis (FDD) methods are indispensable for the system reliability, operation at high efficiency, and safety of the PV plant. In this paper, the types and causes of PV systems (PVS) failures are presented, then different methods proposed in literature for FDD of PVS are reviewed and discussed; particularly faults occurring in PV arrays (PVA). Special attention is paid to methods that can accurately detect, localize and classify possible faults occurring in a PVA. The advantages and limits of FDD methods in terms of feasibility, complexity, and cost-effectiveness and generalization capability for large-scale integration are highlighted. [10]

Improving the efficiency of photovoltaic (PV) systems has gained priority in current research due to the large volumes of PV panels installed. Moreover, the remarkable efforts made to investigate different methods of diagnosing PV failures have multiplied, giving additional impetus to research on the efficiency of PV systems. However, most of these methods are limited in the number of faults that can be identified; some are expensive and complex, and others require huge amounts of data to train. In this paper, a simple and robust multivariate statistical analysis method is proposed for the diagnosis and identification of faults in a PV system. [11]

II.2 Fault classification

In this research, we are going to discuss the most common occurring faults in the photovoltaic installation. The faults chosen are classified according to their origin, intrinsic or extrinsic in the PV system.

To better understand the faults, we will elaborate more on the intrinsic and extrinsic faults by grouping them in a table by type of fault, its consequence and its degree of impact (weak, average, strong), also its phase of origin (Fabrication stage, installation, during process of use) .

II.2.1 Intrinsic faults:

Table II.1: Intrinsic faults and their consequences [12]

FAULTS	CONCEQUENCE	CRIT	OCC
Inversion of output links	Incorrectly wired module, reduced performance	60 %	60 %
Bad orientation and inclination of the modules	Shady operating area which leads to reduced performance	40 %	60 %
Galvanic couple due to the mixing of the module junction	Corrosion	40 %	60 %
Bad module ventilation	Heating	40 %	40 %
Badly fixed module	Module displacement Reduced performance	40 %	40 %
Unwired module	Reduced performance	40 %	40 %
Cracking	Loss of its lightness nature Cell deterioration Decrease in its robust and performance	60 %	20 %
Rust by water infiltration	Cell deterioration, Loss of lightness	60 %	20 %
Poor insulation between modules and the inverter	Short-circuit Destruction of the module Fire	60 %	20 %
Humidity penetration	Hot spot Leakage current Reduced resistance of the short circuit current Corrosion Loss of adhesion and insulation	60 %	20 %
Difference in module performances	Decrease in the performance of light intensity	20 %	60 %
Cable gland plugs missing on the connection box	Water penetration Corrosion of the connections	20 %	60 %
Connection box put upside down	Water enters the box by plugs	20 %	60 %
Increase in the series resistance due to thermal cycle	Reduced performance	40 %	60 %
Deterioration of anti-reflective layer	Reduced performance	40 %	60 %
Inclination of the modules too low	Stagnation of water Earth deposit	40 %	60 %

	Mushroom overgrowth Sealing problem		
Degradation of interconnections	Deterioration of joints Reduced performance Augmentation or resistance and heat	40 %	20 %
Inadequate mechanical support for the modules	Less mechanical effort to support the modules hence parts start separating	40 %	20 %
Bad mechanical resistance for the support modules	Deformation of the support system	40 %	20 %
Diffusion of phosphorous to the surface	Loss of encapsulation adhesion	40 %	20 %
Important leakage current	Heating problem	40 %	20 %
Overheating of the modules by the connection box	Detachment of the components and short-circuits of the system. Destruction of the insulation materials And reduced performance	40 %	20 %

II.2.2 Extrinsic faults:

Table II.2: Extrinsic faults and their effects [12]

Fault	**Consequence**	**Critic**	**Occur**
Accumulation of soil, snow and other substances	Loss of power	60%	60%
Degradation of the modules by vandalism	Reduced performance, Lack of proper functioning of the system	60%	40%
Theft of modules	Functioning of the PV system affected	60%	40%
Deterioration of the sealing joints	Sealing losses Deterioration of cells	60%	20%
Deformation of the frame modules	Infiltration of water	60%	20%
Corrosion of the frame for modules	Deterioration of cells	60%	20%
Delamination	Poor performance Overheating	60%	20%
Lightening	Deterioration of the cells	60%	20%
Storms	Torn and broken module	60%	20%
Structure weaknesses to the wind	Torn and broken module	60%	20%
Lightning strikes on	Destruction of modules	60%	20%
Partial shade	Hot spot	60%	20%

	Deterioration of cells		
Degradation of the encapsulation due to ultraviolet	Does not absorb radiation properly Reduced performance	60%	20%
Degradation due to light intensity	Destruction of diodes Reduced performance Voltage sage	40%	20%
Degradation due to heat	Reduced performance, Deterioration of joints and overheating	40%	20%
Nests of insects or birds on the modules	Reduced performance	40%	20%

II.3 Mismatch and shading faults

The mismatch and shading faults are the most frequent occurring faults in the PV systems. We are going to discuss these commonly occurring faults in depth.

II.3.1 Definition:

The mismatch fault is a fault caused by cell grouping which has the non-identic characteristic of I-V. Any change in the characteristic if I-V will cause tremendous amount of problems.

The shading fault problem is a specific type of mismatch fault because its presence signifies the reduction of solar radiation received by the solar cells. The change in the parameters affect two principal factors.

Firstly, the cells can have different physical properties caused by the fabrication tolerance, only the tolerance of power output of the cells are fixed by the fabricator and can vary from +/-3% and +/-5% depending on the fabricator.[13]

Secondly, the PV cell modules can be exposed to different working conditions caused by different faults. The parameters affected in this instance can be represented in the table below:

Table II.3: Impact of different faults on the parameters of the PV cell [14]

NATURE OF FAULTS	PARAMETERS AFFECTED
Torn or broken modules Shading caused by tree leafs, manure, sand, snow, pollution etc.	Variation of optimal current I
Cell heating	Variation in temperature T
Degradation of the interconnections Cracking Corrosion of connections between cells	Variation in the series resistance R
Different module performances Deterioration of the cells Penetration of humidity	Variation of all the parameters of the cells.

II.4 Commonly occurring faults in PV site systems and installation

We have discussed different types of faults in the previous paragraphs and now it's time to talk about commonly occurring problems in PV installations. All the problems mentioned above are experienced in PV systems but not always and often. For example, it's not always it snows and corrosion doesn't happen in a single night, so while these maybe faults encountered, they may not be the most occurring of them all. Take an example of a car, a car experiences a lot of faults but a good driver knows where exactly to look for faults in an event that a car suddenly broke down, similarly a good engineer will start the diagnostic of PV faults with the following commonly occurring faults in the systems. [15]

II.4.1 Current faults in PV systems (I)

The faults in the PV system can be described as temporally or permanent. The temporally faults are caused by shading and fouling of the solar cell modules. The permanent faults in the module are:

- The delamination
- The bubbles and water drop on the surface of the cells
- The yellowing of the cells due to radiation
- Scratches and burnt cells

The permanent faults are eliminated by replacing or repairing the destroyed or affected modules. The most serious and dangerous faults in a photovoltaic system are caused by short-circuits, between the lines, grounding and arc fault. Other factors which might minimize the power output during production is a point of maximum power (MPP), the power losses by joule effect in the cables and faulty equipment. The faults in the photovoltaic system can therefore be classified as the module faults, channels or grid according to the PV components involved. [15]

II.4.2 Hot spot faults

The hot spot faults occur in individual solar modules when they are shaded or broken by mechanical stress. These cells produce far less electrical current compared to the other cells not affected and can be polarized in the inverse direction which results in the production of heat by joule effect in the course of production.

This phenomenon affects the cells made from silicon crystalline and generally results in the fouling, shading of the damaged cells or diodes of the bypass damaged. The hot spot points release energy which increases the temperature of the surface and consequently the hot spot faults are diagnosed and analyzed by thermal and infra-red. If the hot spot faults persist, it can damage the solar cells and the bypass diodes and provokes short circuit faults. [15]

II.4.3 Degradation

The degradation of the solar modules facilitates the reduction in the power output as time passes. The degradation faults can be identified by reading the characteristic I-V of the module. [15]

II.4.4 Partial shade

The shading faults occur when certain parts of PV solar module receive less radiation compared to the rest of the module due to obstructions and shadows. Shading can be diagnosed by looking for unexpected current drops. A shadowing effect gives similar results to open circuit strings but are most often temporally. [16]

II.4.5 Open circuit faults

The open circuit faults are reference the faults of interconnection in the sub-systems of the PV generator or module. It will equally include the disconnections of the cells of the module, the chain of modules or the chain of the PV electrical grid. [16]

The diagnostic in the PV grid can be done by inspecting the voltage and current indicators. The voltage of the PV grid remains constant, however, the fault results in the current drop. The open circuit faults can be caused by the damaged cells, defective diodes and wiring faults. [16]

II.4.6 Short circuit faults

All like the open circuit faults, the short circuit faults can be produced in the different sub-systems of the PV installation. The modules having the short circuit in the chain of production experiences a significant voltage drop in the grid such that the current of the lines increase exponentially. The same effect is produced when the short circuit is produced between two branches of the system. An experimental study shows that the short circuit faults between the modules has harmful effects on the output voltage of the system as the short circuit of the strings. [17]

II.4.7 Ground faults

The ground faults are considered to be the most commonly occurring faults in the PV systems. The faults result from the accidental electrical short circuit between the electrical conductor and the ground. This fault is principally caused by the wiring insulation. The grounding faults can cause serious harm and risks for the security of the workers in an event of the electrical arcs of DC currents generated at the point of failure on the system, the electrical shocks due to ground faults results in less voltage compared to the nominal voltage and risk fires. [17]

II.4.8 Arc faults

The involuntary passage of current in the air or in another dielectric is known by the name arc fault. The fault arcs can be produced by the discontinuity between two electrical conductors having different potential differences. Arc faults in photovoltaic system can risk serious dangers to the installation. [17]

II.4.9 Line to Line faults

The line-to-line fault designates the short circuit faults between the conductors of the PV system. The line-to-line faults can be caused by fault insulation of the wires and mechanical damage. [18]

II.5 Methods of detection and diagnostic of PV systems

Fault analysis and fault detection are important to the efficiency, safety and reliability of solar photovoltaic (PV) systems. Despite the fact that PV systems have no moving parts and usually require low maintenance, they are still subject to various fault conditions. Especially for PV arrays (dc side), it is difficult to shut down PV modules completely during faults, since they are always energized by sunlight in daytime. Furthermore, conventional series-parallel PV configurations increase voltage and current ratings, leading to higher risk of large fault currents or dc arcs.

Once PV modules are electrically connected, any fault among them can affect the entire system performance. This means the PV system is only as robust as its weakest link (e.g., the faulted PV components). In a large PV array, it may become difficult to properly detect or identify a fault, which can remain hidden in the PV system until the whole system breaks down. In addition, conventional series-parallel PV configurations increase voltage and current ratings, leading to higher risk of large fault currents or dc arcs.

There are four methods used for diagnostic of PV systems in the industry:

- Method based on the analysis of current and volt (electrical method)
- Method based on the analysis of other parameters like the I-V curve (also called non electrical method)
- Literature Method [19]
- Artificial Intelligence methods [20]

II.5.1 Non electrical methods

There exist many non-electrical methods, destructive or non-destructive for the diagnostic of PV faults in the module. The main principal fault we can give much attention to is cell cracking. We can cite the methods as follows: mechanical bending tests, imagery by photoluminescence, electroluminescence and the test of thermography. For the diagnostic of PV modules, the method of imagery (thermal camera) infra-red is widely used. [21]

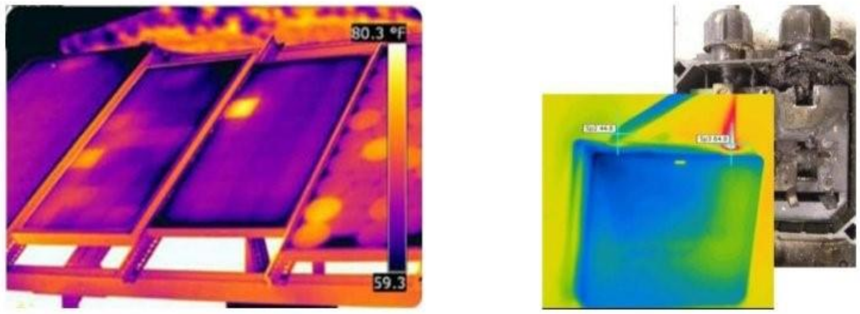
(a) cellule court-circuitée (b) corrosion dans la boîte de jonction

Figure II.1: Certain examples of the detection of PV faults using thermal camera. [21]

There have been most successes in the localization and detection of PV faults using the thermal camera which are noted as: current leakage in the PV, increase in the resistance of the connection between the modules, abnormal heating of the cells, and conduction of the bypass diode. This method can equally be applied for the connections in the junction box and the functionality of the anti-reverse diode. [21]

II.5.2 Electrical methods

Using electrical methods, the most considered parameters are:

- The current output of the GPV
- The voltage on the terminals of the GPV
- The insulation resistance between the positive and negative terminals of the GPV.

It is also possible to add additional parameters like the ambient temperature of the site and sunshine radiation in the electrical measures. The measures on the AC side are important because they are directly related to the energy which will be sold. It is necessary to take note of:

- The AC currents
- The AC voltage
- The frequency
- Impedance of the electrical grid as seen by the inverter

Out of these parameters written, it becomes much easier to deduce the following:

- ❖ DC instantaneous power
- ❖ AC instantaneous power
- ❖ Electrical energy produced on different periods (depending on the capacity of the storage system) on the side of both AC and DC.

We often add the following:

- The functioning duration of the inverter
- The date when put in service
- The CO2 not released in the atmosphere (which is rather economical to the environment). [22]

II.5.3 Literature methods

The different methods proposed in the literature method type of detection and localization of PV faults are as follows:

a. Reflectometry method

The reflectometry method is a diagnostic method used to send a signal in the system or on the diagnostic side. This signal is propagated according to the law of propagation in in the medium in question and it encounters the discontinuity, and part of its energy is re-transmitted to the point of injection. Analysis of the signal allows us to deduce the information on the system or medium been considered.

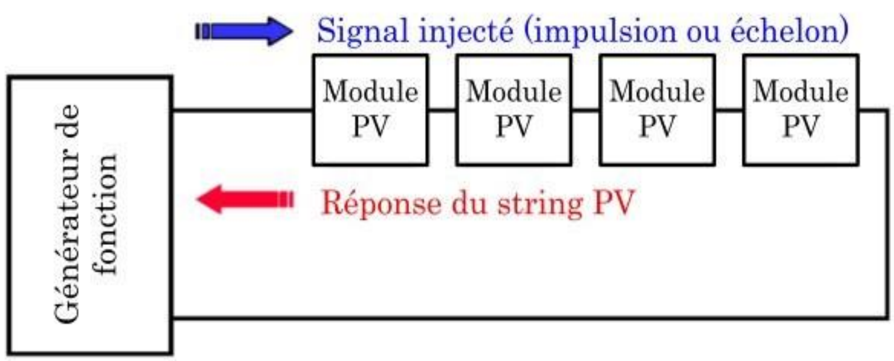

Figure II.2: Principal reflectometry method for the detection and localization of PV faults in a string. [23]

b. Analysis of power and energy produced

The power or energy measured is compared to the expected output and when there is an important deviation, we can be certain that there is surely a faulty.

The suggested analysis consists of generating the supplementary attributes on the power drop or energy produced such as: the duration, the amplitude, the frequency and the drop instances. These same attributes are equally predetermined for the different faults considered in our study.

In the course of their comparison, the fault whose value attributes are considered similar or close to those measured are considered as the faults responsible for the drop in the power output. [23]

II.5.4 Detection of PV faults by means of artificial intelligence methods

a. Introduction

Various factors, including maximum power point tracking error, environmental effects like shading and dust or snow buildup on the PV surface, wiring losses and aging, and malfunction in other PV components like the power conditioner unit and the inverter can all have an impact on how well a PV system operates. According to a monitoring study in [16], faults may cause a PV system to generate roughly 18.9 percent less power annually. In order to continually analyze the current, voltage, and output power characteristics of a PV system and find both existing and emerging defects, proper techniques have to be developed especially ones based on artificial intelligence like the following of the many examples.

b. Fuzzy Logic method Control Implementation

Among several renewable energy resources, Solar has great potential to solve the world's energy problems. With the rapid expansion and installation of PV system worldwide, fault detection and diagnosis has become the most significant issue in order to raise the system efficiency and reduce the maintenance cost as well as repair time.

The Fuzzy Control Implementation (FLC) is one of the modern artificial intelligence techniques used in fault diagnosis in the PV systems. The architecture of the implementation is based on the Max-Min arrangement procedure with a centroid type for the defuzzification. [24]

c. Artificial Neural Networks

Artificial neural networks, a pivotal technique of artificial intelligence, have been developed and applied in many fields including the fault diagnosis of PV systems, due to their strong self-learning ability, good generalization performance, and high fault tolerance. Artificial neural networks (ANN) are type of machine learning algorithms that are commonly used for PV fault diagnostic and detection. ANN`s can be trained on large datasets of PV system performance data to recognize patterns associated with various types of faults. [25]

d. Deep learning

Deep learning is a type of machine learning that uses neural networks with multiple layers to learn complex patterns in data. Deep learning has been shown to be effective in fault diagnosis and detection in PV systems particularly in cases where there are large amounts of data available. [26]

e. Genetic algorithms

Genetic algorithms are a type of optimization algorithms that can be used to optimize the parameters of a diagnostic systems. Genetic algorithms can be used to find the optimal set of parameters for a diagnostic system that can accurately diagnose faults in a PV system. [27]

f. Support Vector Machines (SVMs)

The support vector machines are another machine learning algorithms commonly used for PV fault diagnosis and detection. SVMs are particularly effective in cases where there are multiple types of faults that need to be distinguished from one another. [28]

g. Decision Trees

Decision Trees are a type of machine learning algorithm that can be used to create a model of the decision making process used in fault diagnosis and detection. Decision trees can be used to identify the most likely cause of a fault based on the observed system performance. [29]

h. Operating point analysis

Other comparison between the power and energy produced to that which is expected, the comparison of the point of actual maximum power (current and voltage corresponding to the maximum power) to that which is expected can carry much information on the state of the PV system. The rational comparison between these currents and these voltages gives the two pairs of binary values (0 and 1). Depending on the comparison of these two pairs of values, the nature of the PV fault can be identified. The four families of problems are as follows:

- Faulty modules in a string
- Faulty string
- The family of non-discriminant faults: shading, MPPT error and old modules
- False alarms. [30]

These are just some of the AI methods used for fault detection and diagnosis.

i. The Bayesian Neural Networks

Bayesian neural network (BNN) combines ANN with Bayesian implication. Basically speaking, at BNN level, the treatment of both weights and outputs as variables and control over-fitting. The final goal of BNN is to quantify the uncertainties presented by the models, this approach employs the statistical methodology where the whole data has a probability distribution attached to it, In user interface design software, variables tend to take a specific value will turn the same result at every access to the dedicated variable. In comparative way, the Bayesian world can own similar entities as well-known as random variables that will present a various value at any moment you access it. In other terms, the historical data describe the prior information of the overall manner with each variable giving its own statistical properties which vary with time. [31]

Basically, Bayesian neural networks focus on marginalization comparing to other ANNs, they estimate by maximum a posteriori or predictive distribution. In addition, they depend on Markov Chain Monte Carlos, Variation in Inference, and Normalizing Flows technics. Bayesian neural network are useful, in the area where data are rare, they have the capacity to obtain better results for a large number of labor as well as they can estimate the uncertainties in predictions. [32]

II.6 Analysis of the static characteristic

Deformation of the current-voltage graph characteristic can be provoked by changing the working conditions (sunshine or temperature) or by the appearance of one or many faults in the PV system.

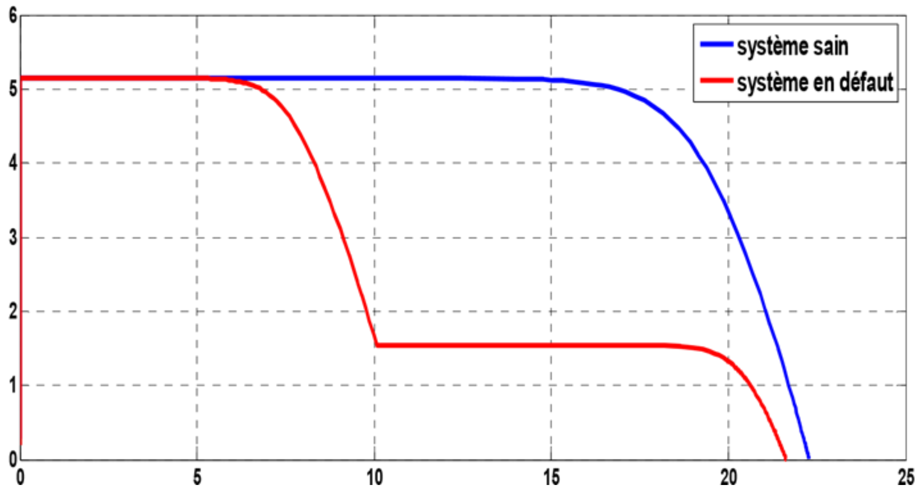

Figure II.3: Appearance of the faulty graph characteristic of I-V. [34]

Figure II.3 shows the faulty graph of I-V characteristic (shading of the module which consists of 36 solar cells at 50 %) compared to the normal working conditions of the normal module. By exploiting the information on the I-V characteristic, the detection and localization of faults can be realized. [33]

II.7 Symptoms of the PV system

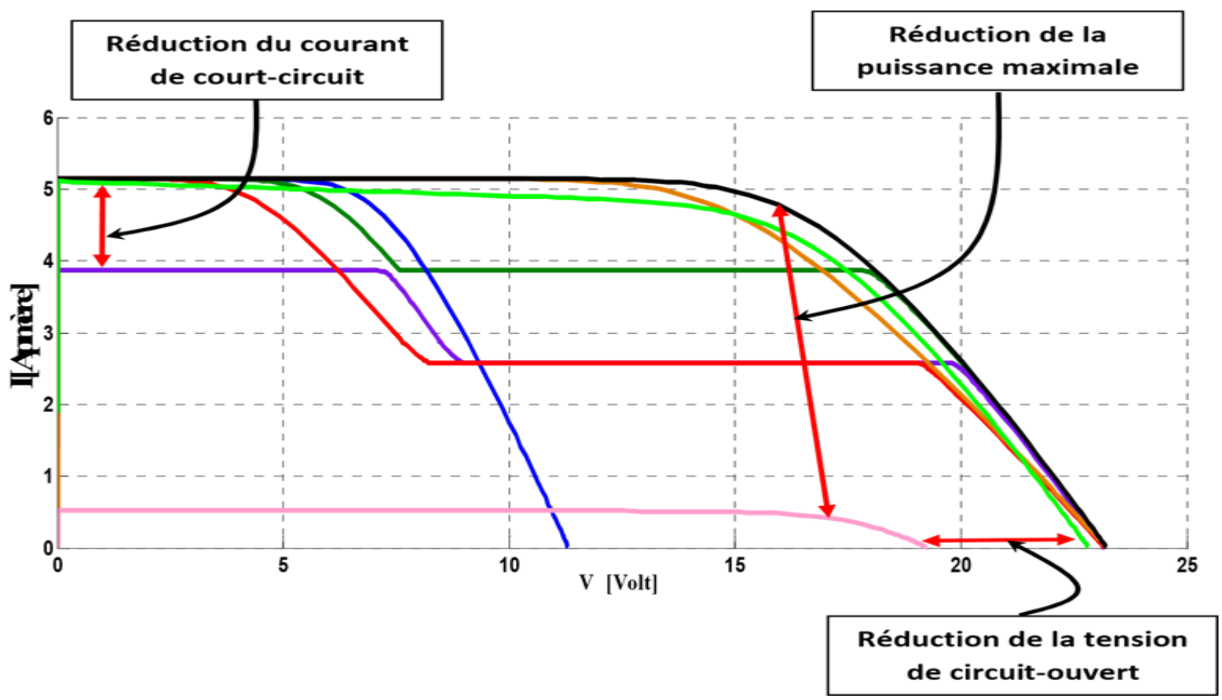

Figure II.4: Symptoms of the I-V characteristic [34]

The figure II.4 shows us how to identify faulty PV modules. From the graph we can extract information such as current, voltage and power. A normal functioning PV has a particular curve on the graph which we can compare to the faulty one as shown in the figure above. So, in solar installations and solar farms, various soft wares such as MATLAB, PSIM, Scada and many more can be used to identify changes in the course of normal working conditions of the PV system. [34]

CHAPTER THREE

THE KNN (K-Nearest Neighbor) AlGORITHM

III.1 Introduction

K-nearest neighbors (KNN) is a type of supervised learning algorithm used for both regression and classification. KNN tries to predict the correct class for the test data by calculating the distance between the test data and all the training points. Then select the K number of points which is closest to the test data. The KNN algorithm calculates the probability of the test data belonging to the classes of 'K' training data and class holds the highest probability will be selected. In the case of regression, the value is the mean of the 'K' selected training points. In this chapter, we are going to learn about the following:

- The key concepts that define nearest neighbor classifiers, and why they are considered lazy learners
- Method to measure the similarity of two examples using distance
- How to apply popular nearest neighbor classifier called K-NN

This algorithm is used in many other fields not just in medicine and engineering. We can use this method to classify many related substances to see and observe were they fit best. A good example of this is in solving the debate whether tomato is a fruit or a vegetable. When KNN was applied, it was discovered that tomato is actually a fruit because it has more features of a fruit than a vegetable. [35]

III.2 Understanding nearest neighbor classification

In a single sentence, nearest neighbor classifiers are defined by their characteristic of classifying unlabeled examples by assigning them the class of similar labelled examples. Despite the simplicity of this idea, nearest neighbor methods are extremely powerful. They have been used successfully for:

- Computer vision applications, including optical character recognition and facial recognition in both still images and video
- Predicting whether a person will enjoy a movie or music recommendation
- Identifying patterns in genetic data, perhaps to use them in detecting specific proteins or diseases.

In general, nearest neighbor classifiers are well suited for classification tasks where relationships among the features and the target classes are numerous, complicated or extremely difficult to understand yet the items of similar class tend to be fairly homogenous. Another way of putting it would be to say that if a concept is difficult to define, but you know it when you see it, then nearest neighbor might be appropriate. On the other hand, if the data is noisy and thus no clear distinction exists among the groups, the nearest neighbor algorithm may struggle to identify the class boundaries.

III.3 The strengths and weaknesses of KNN Algorithm

The nearest neighbors approach to the classification is exemplified by the k-nearest neighbor's algorithm (KNN). Although this perhaps one of the simplest machine learning algorithms, it is still used widely. The strengths and weaknesses of this algorithm are as follows:

Table III.1: The strengths and weaknesses of KNN Algorithm [36]

STRENGTHS	WEAKNESSES
• Simple and effective • Makes no assumptions about the underlying data distribution • Fast training phase • High accuracy of prediction • Quick calculation time • Can be used for multi class problems as well	• Requires selection of an appropriate k • Slow classification phase • Does not produce a model limiting to understand how features are related to class • Nominal features and missing data require additional processing • Computationally expensive and requires high memory as the algorithm stores all the training data • The algorithm gets slower as the variable increases

The KNN algorithm gets its name from the fact that it uses information about an example`s k-nearest neighbors to classify unlabeled examples. The variable term implying that any number of nearest neighbors could be used. After choosing k, the algorithm requires a training dataset made up of examples that have been classified into several categories, as labelled by a nominal variable. Then for each unlabeled record in the dataset, k-NN identifies k records in the training data that are the ``nearest`` in the similarity. The unlabeled test instance is assigned the class of the majority of the k nearest neighbors. [35]

III.4 How KNN works

The KNN can be explained of the basis of the following algorithm:

- Step 1: Select the number of k neighbors
- Step 2: Calculate the distance of K number of neighbors
- Step 3: Take the K nearest neighbors as per the calculated distance.
- Step 4: Among these k neighbors, count the number of the data points in each category.
- Step 5: Assign the new data points to that category for which the number of the neighbor is maximum.
- Step 6: Our model is ready.

Suppose we have a new data point and we need to put it in the required category. Consider the below image:

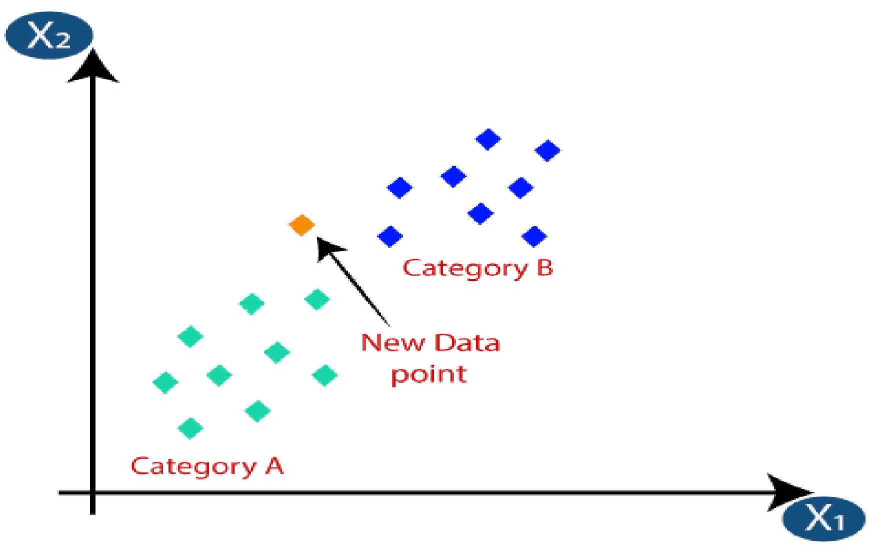

Figure III.1: Data distribution on the x-y plane [35]

- Firstly, we will choose the number of neighbors, so we will choose the k=5.

- Next, we will calculate the Euclidean distance between the data points. The Euclidean distance is the distance between two points, which we have already studied in geometry. It can be calculated as:

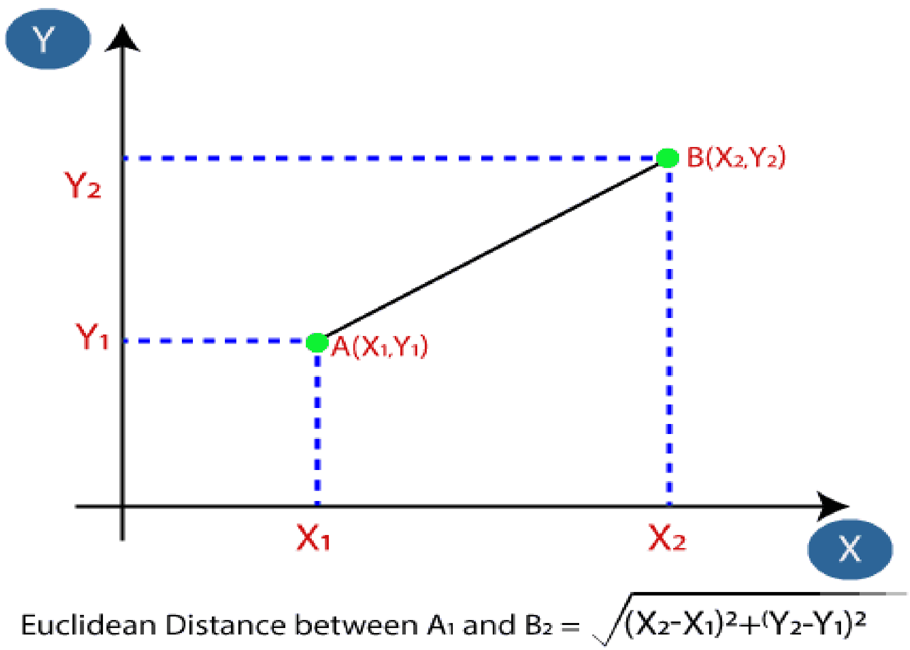

Figure III.2: Euclidean Distance [35]

- By calculating the Euclidean distance we got the nearest neighbors, as three nearest neighbors in category A and two nearest neighbors in category B. Consider the below image:

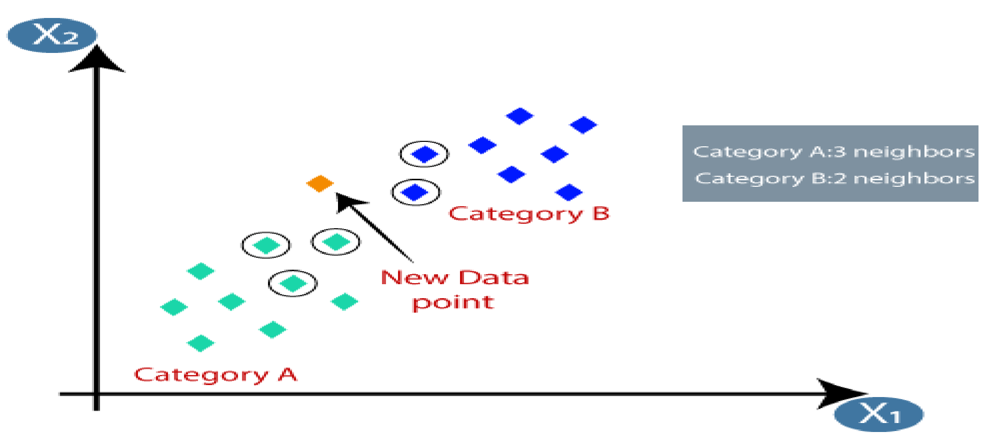

Figure III.3: Data distribution and manipulation [35]

- As we can see the 3 nearest neighbors are from category A, hence this new data point must belong to category A.

III.4.1 Choosing the value of K

The decision of how many neighbors to use for KNN determines how well the model will generalize to future data. The balance between over fitting and under fitting the training data is a problem known as bias-variance trade off. Choosing a large K reduces the impact or variance caused by noisy data, but can bias the learner so that it runs the risk of ignoring small but important patterns.

K value indicates the count of the nearest neighbors. We have to compute distances between test points and trained labels points. Updating distance metrics with every iteration is computationally expensive, and that's why KNN is a lazy learning algorithm.

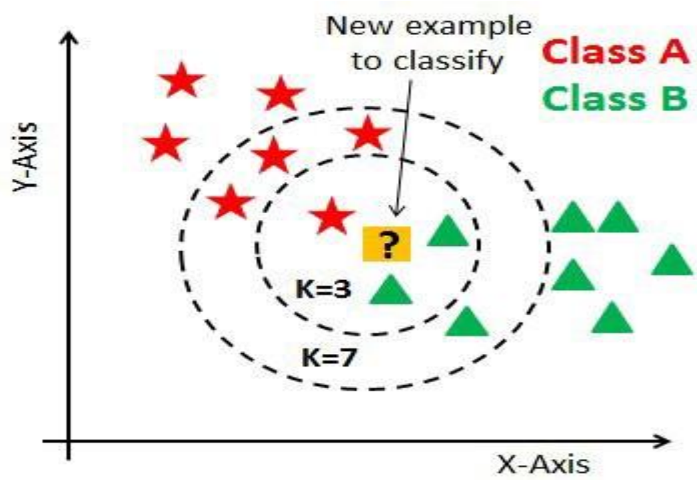

Figure III.4: Use of KNN [35]

- As you can verify from the above image, if we proceed with K=3, then we predict that test input belongs to class B, and if we continue with K=7, then we predict that test input belongs to class A.
- That's how you can imagine that the K value has a powerful effect on KNN performance.

III.4.2 How to select the optimal K value

- There are no pre-defined statistical methods to find the most favorable value of K.
- Initialize a random K value and start computing.
- Choosing a small value of K leads to unstable decision boundaries.

- The substantial K value is better for classification as it leads to smoothening the decision boundaries.
- Derive a plot between error rate and K denoting values in a defined range. Then choose the K value as having a minimum error rate.

Now you will get the idea of choosing the optimal K value by implementing the model.

III.4.3 Calculating distance

The first step is to calculate the distance between the new point and each training point. There are various methods for calculating this distance, of which the most commonly known methods are — Euclidian, Manhattan (for continuous) and Hamming distance (for categorical).

a. Euclidean Distance: Euclidean distance is calculated as the square root of the sum of the squared differences between a new point (x) and an existing point (y).

$$euclidean = \sum_{i=1}^{k}(x1 - y1)^2 \text{......................... (Equation III.1)}$$

b. Manhattan Distance: This is the distance between real vectors using the sum of their absolute difference.

$$manhattan = \sum_{i=1}^{k}|(x1 - y1)|^2 \text{.................. (Equation III.2)}$$

CHAPTER FOUR

SIMULATION OF KNN ALGORITHM USING MATLAB

IV.1 Introduction

MATLAB makes machine learning easy. With tools and functions for handling big data, as well as apps to make machine learning accessible, MATLAB is an ideal environment for applying machine learning to your data analytics. With MATLAB, engineers and data scientists have immediate access to prebuilt functions, extensive toolboxes, and specialized apps for classification and regression , and use data to make better decisions.[37]

IV.2 The parameters of our simulation

In order for us to better understand the program of our simulation, we need to know the parameters in question or rather the values we are trying to change.

The parameters in question are therefore temperature, sunlight radiation, voltage and current output.

Now we are going to be using values of temperature, irradiation, voltage and current output in our KNN algorithm to create a predictive program for our PV system.

IV.3 The KNN program realized in MATLAB script

The basis of our work in to test a KNN PV program that detects and identifies faults based on the data of the system. Instead of using the MATLAB KNN tool box, we decided to code our own on MATLAB script in order to make the program more robust and so it can understand the different values been provided.

We have therefore used data science to classify and organize the data given for solar radiation, temperature, current and voltage in form of codes. Now the essence of our program is to give the code once a certain value is asked about its position in the data codes. The data is arranged in terms of training for five days, validation and test data for three days.

Code 1	Code 2	Code 3	Code 4	Code 5
Open circuit fault for current	Normal operating for current	Normal operating for voltage	5 PV modules short circuit fault	10 PV modules short circuit fault

IV.4 The graphs of the simulation

IV.4.1 The training phase

a. Input climate data

We first run the MATLAB script KNN program for temperature and solar irradiation to better understand the relationship between the two variables, and the following are the graphs we obtained:

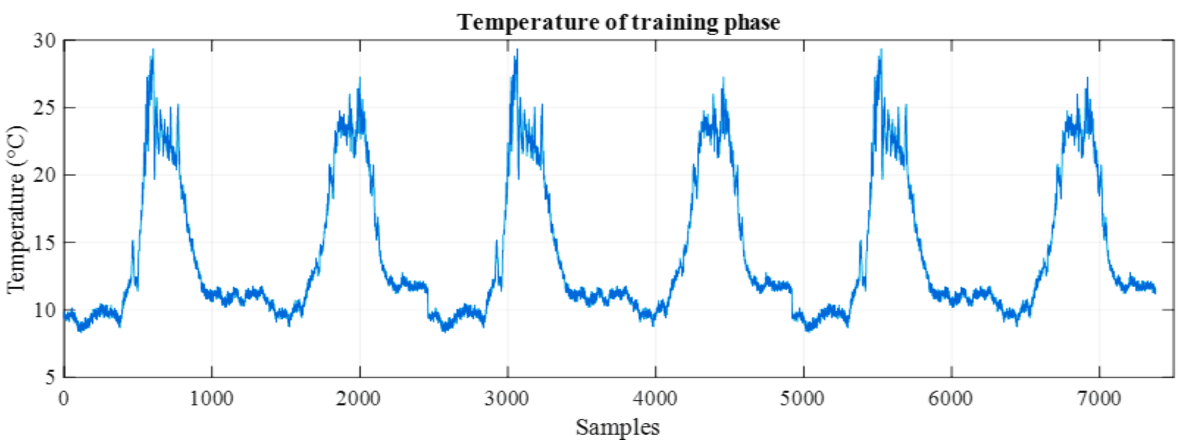

Figure IV.1: The graph of training phase for temperature

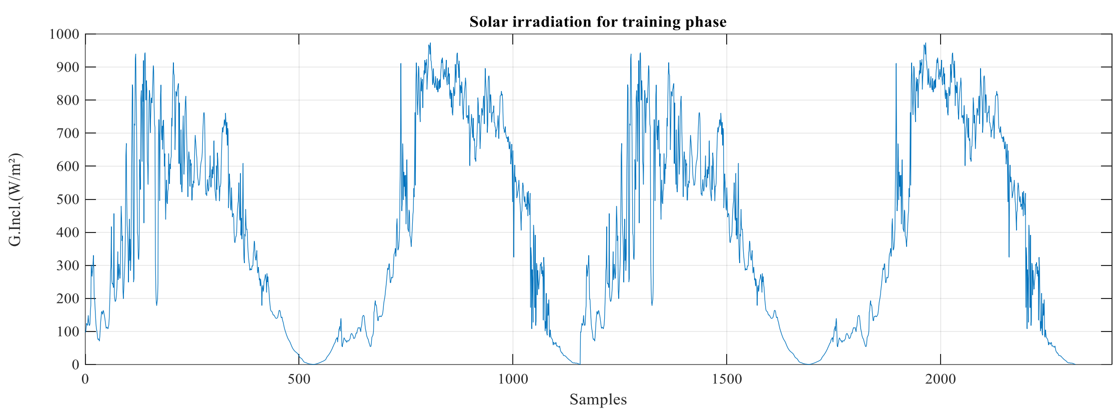

Figure IV.2: The graph of training phase for solar irradiation

In figure IV.1 and IV.2 we get to see the graphs of input climate data for training phase. From the first graph we see the graph of temperature recorded on certain days by 7380 samples the maximum is around 30 (°C). On the other side the maximum or rather the optimum solar irradiation for the PV system given by 2316 samples is 1000 (W/m²).

b. Input electrical data

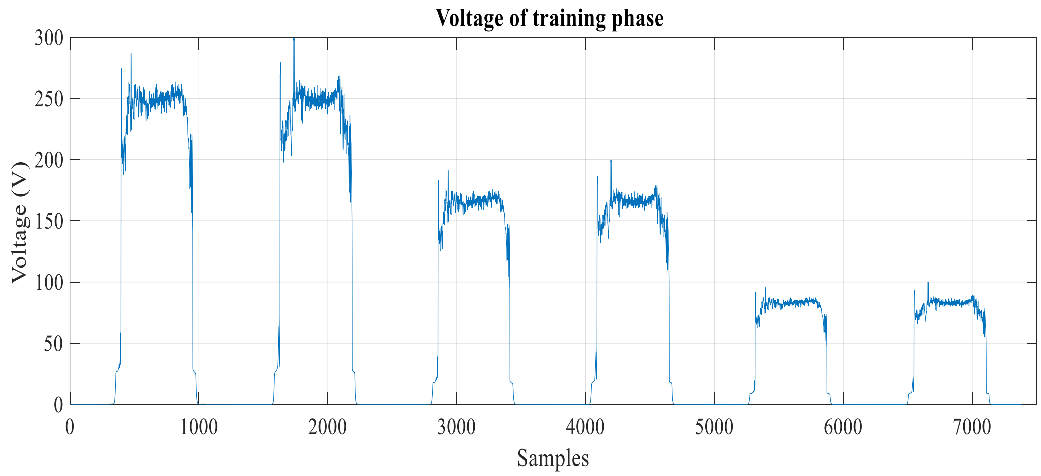

Figure IV.3: The graph of training phase for voltage

The current program was realized using the values of current in our sorted data and the following graphs were obtained:

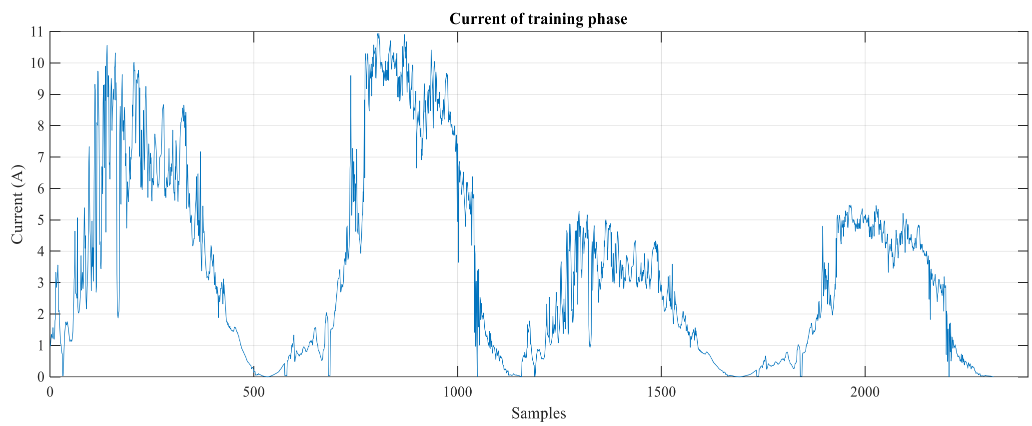

Figure IV.4: The graph of training phase for current

figure IV.3 and IV.4 represent the input electrical data in the training phase where we notice in the first figure that the voltage is at its highest levels (around 300 (V)) in the first two days and gradually decreases with the passage of days . The same thing happens with the current curve ,as it decreases after the first two days by about the half ,from around 11 (A) to 5 (A).

c. Output data (voltage classification)

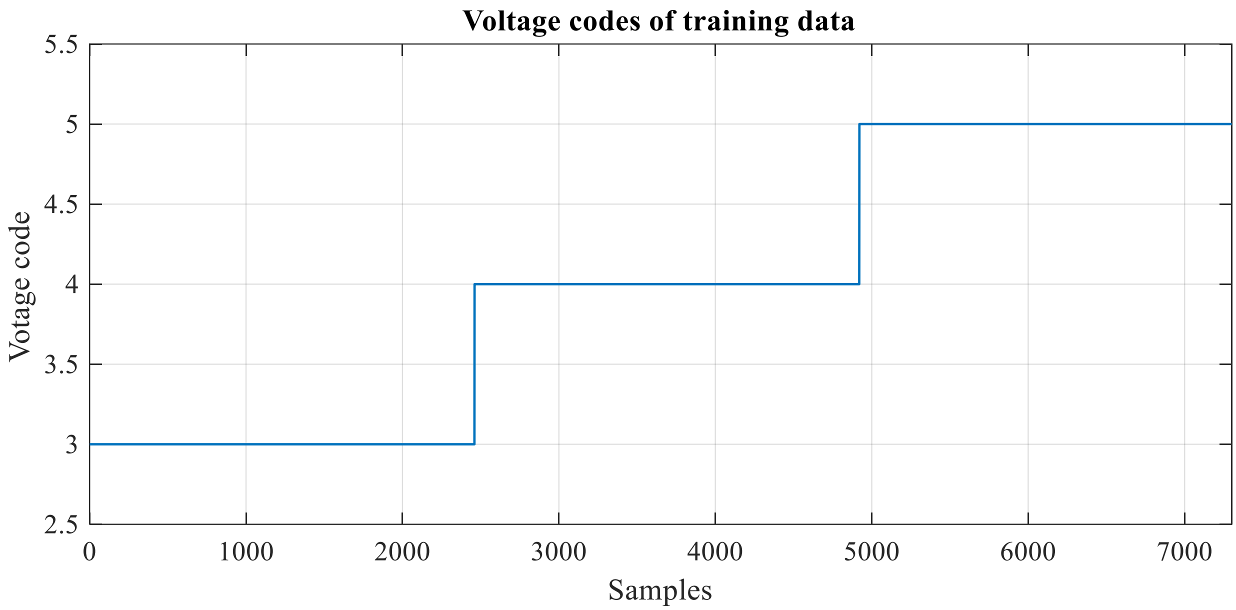

Figure IV.5: The graph of training voltage classification

In figure IV.5 we observe that in the training phase, we had roughly 7380 samples in form of given values as shown on the x-axis .

d. Output data (current classification)

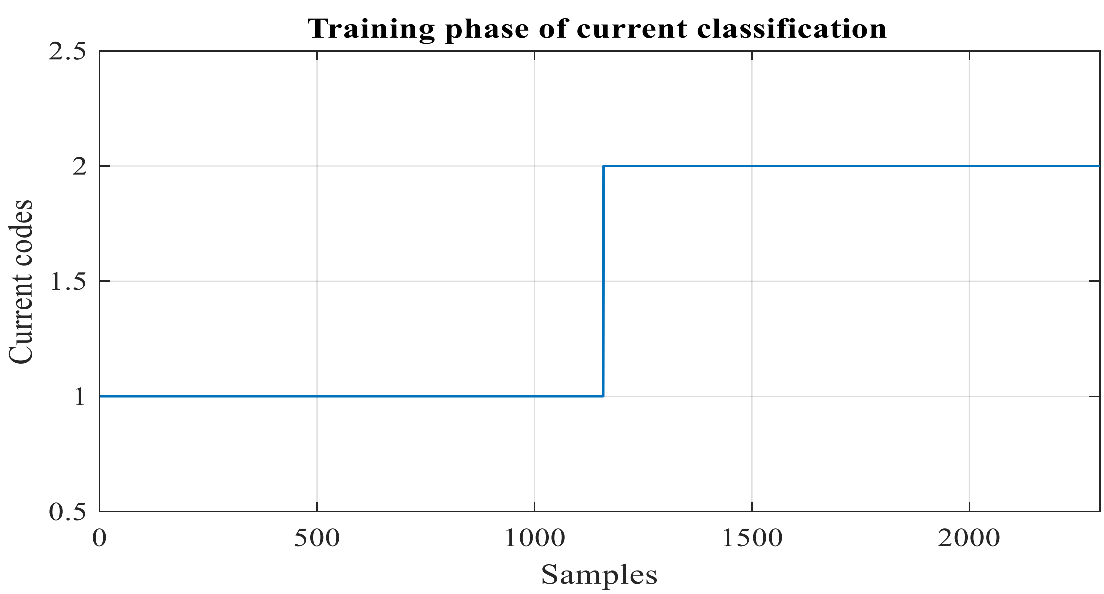

Figure IV.6: The graph of training current classification

Our figure IV.6 displays the codes of the samples for our current classification. For current classification we have two codes, code 1 and code 2. We observe that 1158 samples belong to code 1 and 1158 to code 2 .

IV.4.2 The validation phase

The values tested in program simulation above were for the test data. As mentioned before, our program KNN works with training, validation and test valuables. We have so far done the simulation for training data and now we shall move to validation data.

a. Input climate data

Upon running the program, following graphs were obtained:

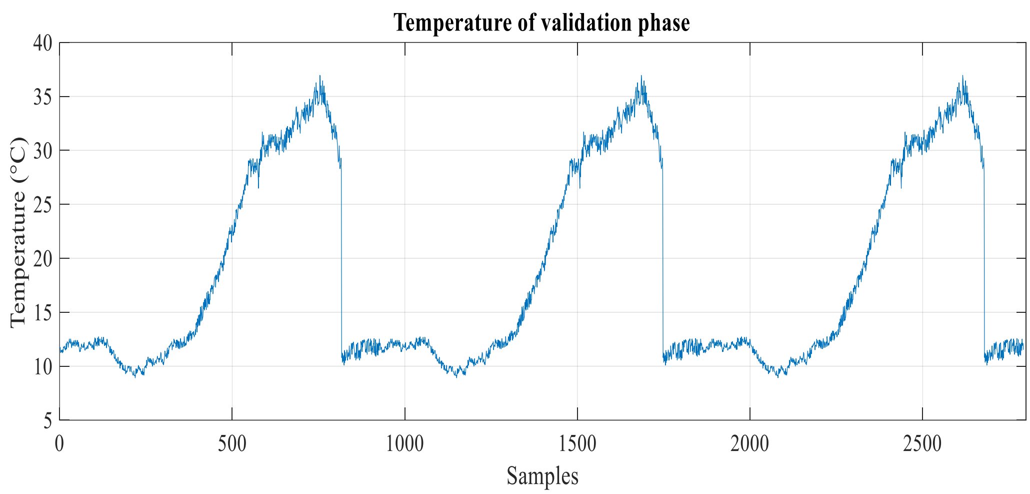

Figure IV.7: The graph of temperature under the validation phase

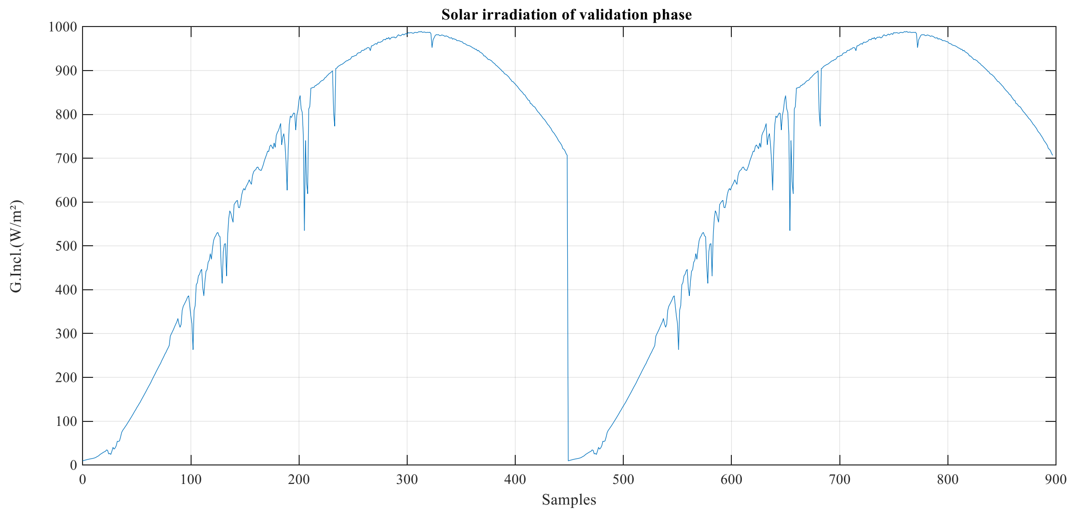

Figure IV.8. The graph of solar irradiation under the validation phase

The figures IV.7 and IV.8 represent the graphs of input climate data under validation phase. The highest recorded temperature shown on figure IV.7 is around 37 (°C) and for solar irradiation is 1000 (W/m²) on figure IV.8.

b. Input electrical data

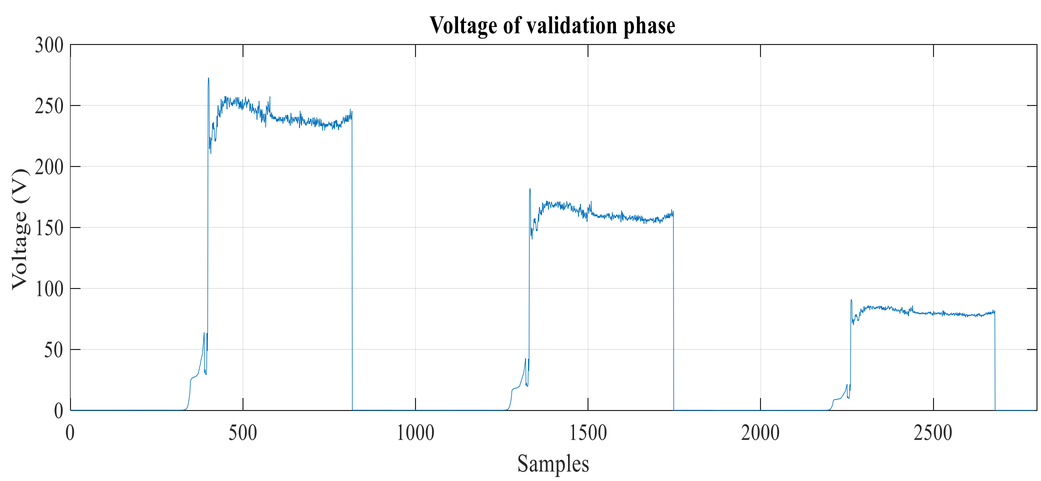

Figure IV.9: The graph of voltage under validation phase

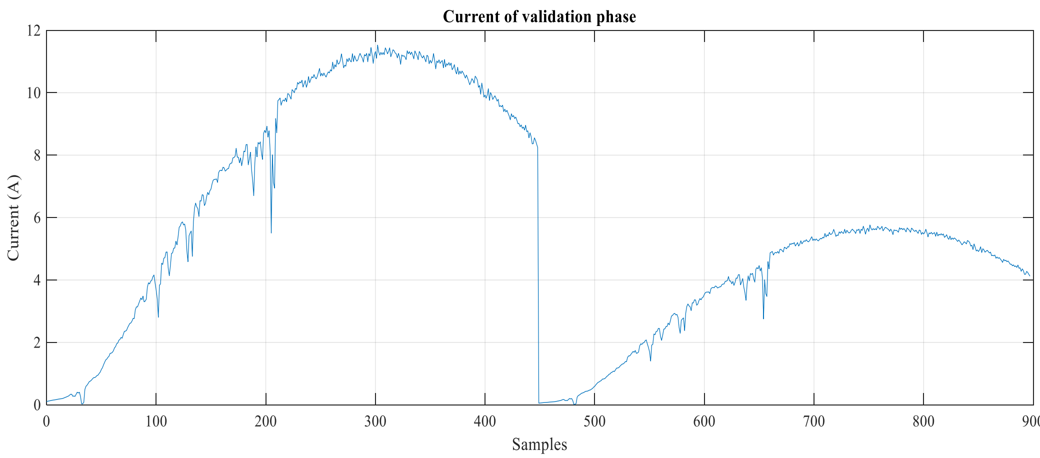

Figure IV.10: The graph of current under validation phase

The figures IV.9 and IV.10 represent input electrical data under validation phase ,The highest recorded voltage in the first graph is around 270 (V) and from the second we can see that the maximum recorded current is 11,5 (A).

c. Output data (voltage classification)

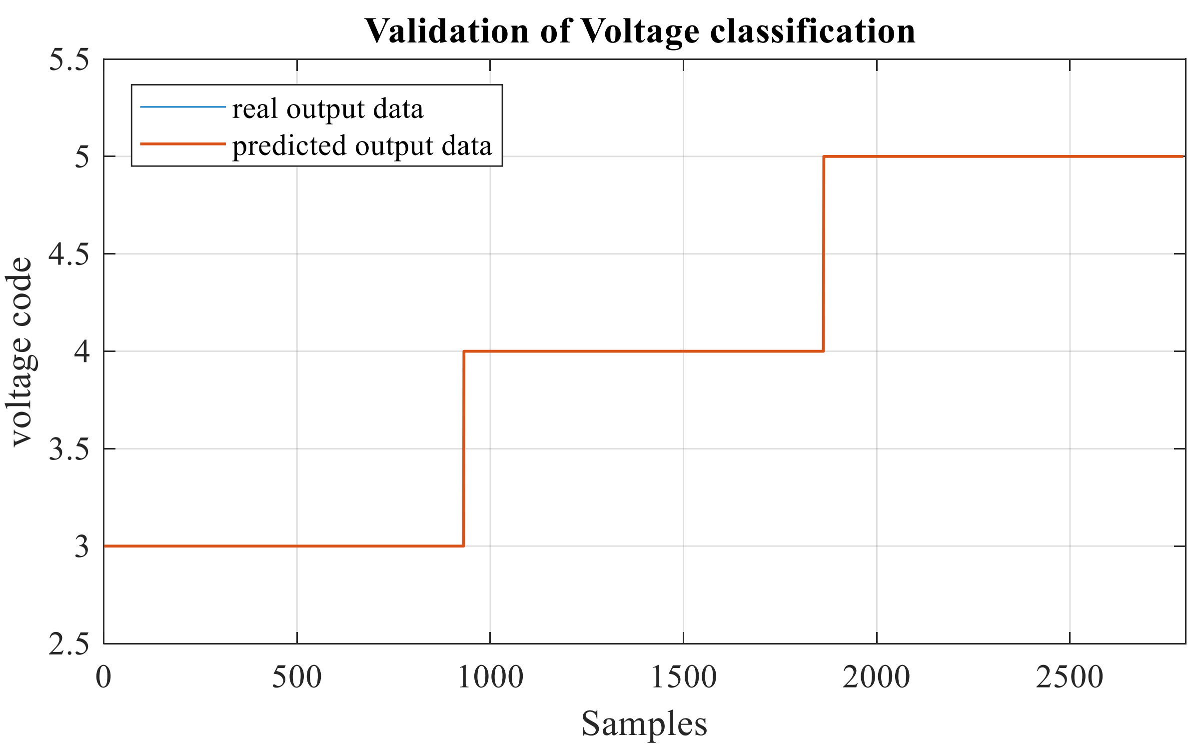

Figure IV.11: The graph of validation voltage classification

Figure IV.11 represent the different voltage codes.

In our graph we have the real and the predicted data and they equal each other hence only one is shown, this is because the accuracy of our classification is around 100 %.

d. Output data (current classification)

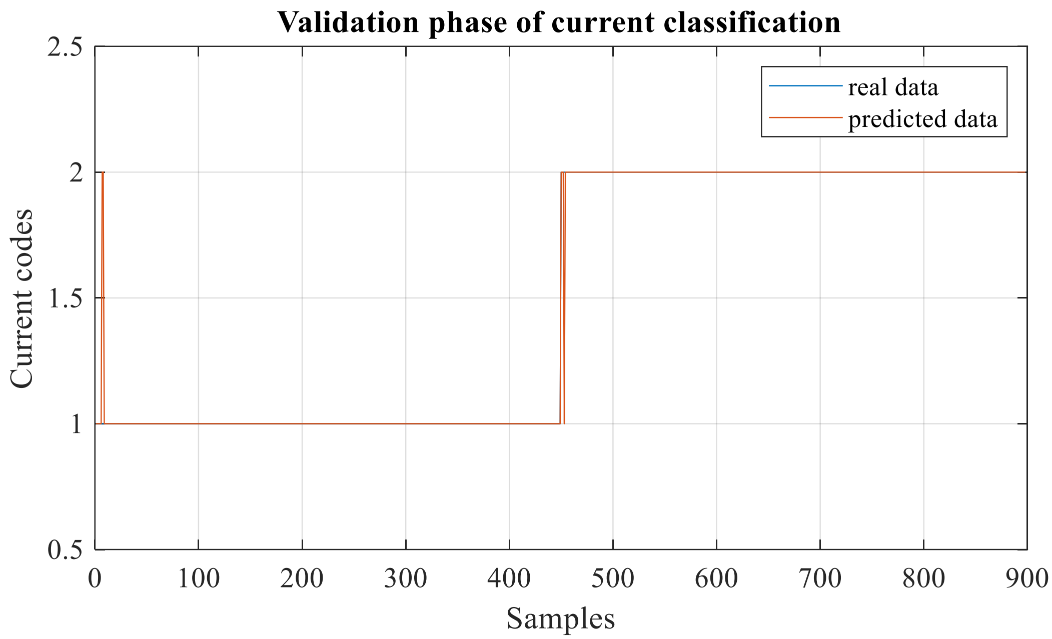

Figure IV.12: The graph of validation current classification

In figure IV.12 shows the output codes for current classification under validation phase.

Our figure contains two graphs: the real and predicted data and we can see the difference between them is minimal because accuracy is around 97 %.

IV.4.3 The testing phase

a. Input climate data

The following are the graphs obtained from the validated data for testing (Temperature and solar irradiation):

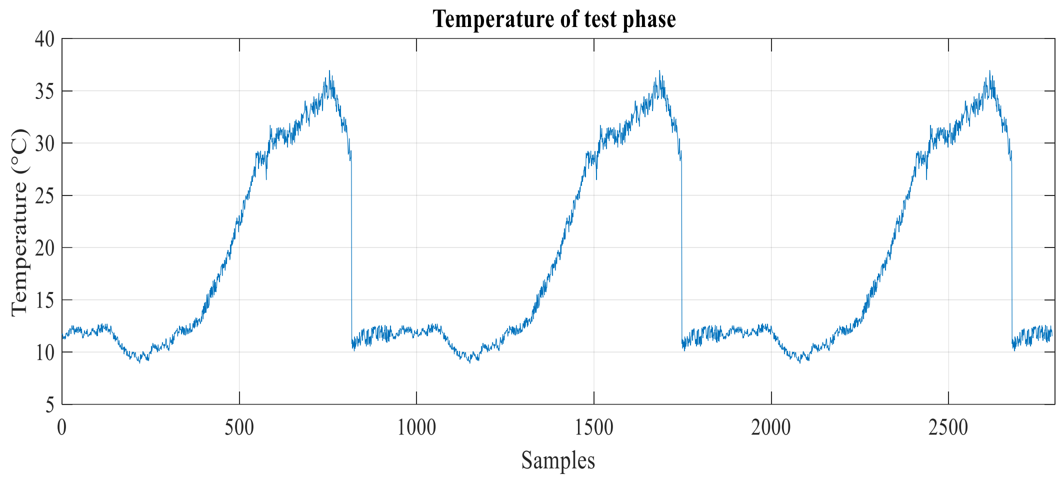

Figure IV.13: The graph of temperature under test phase

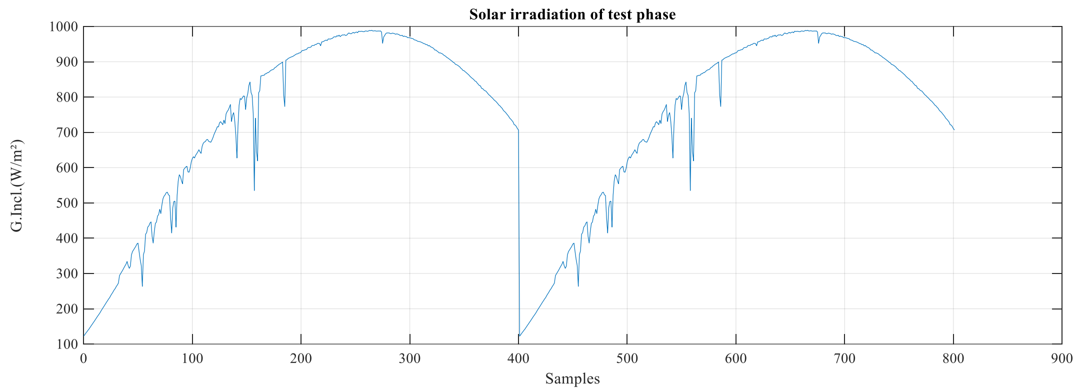

Figure IV.14: The graph of solar irradiation under test phase

Figure IV.13 and figure IV.14 shows the graphs of climate data in the test phase. In figure IV.13, the temperature recorded is 40 (°C) and that represents the overheating faults in PV arrays. Remember the optimum temperature of the PV systems is 25 (°C) and above this value the efficiency of the systems reduces sporadically. And figure IV.14 shows 1000 (W/m²) as the maximum recorded value representing optimum operation.

b. Input electrical data

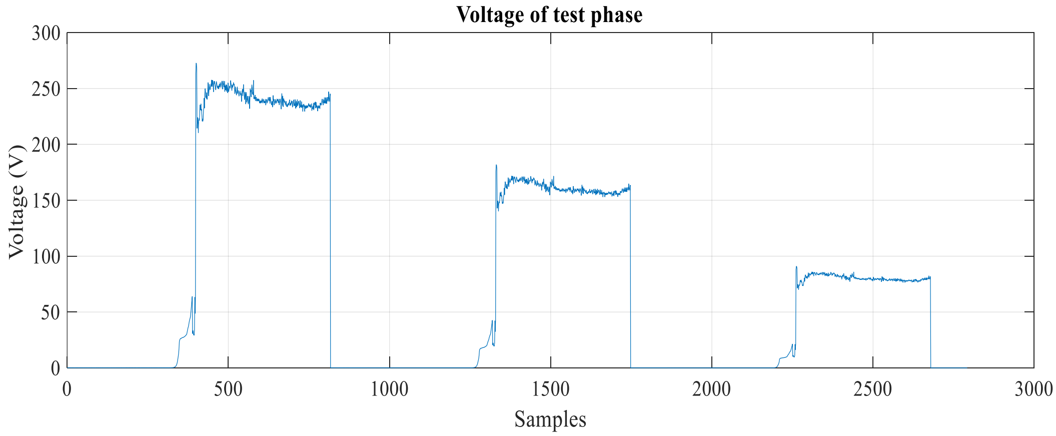

Figure IV.15: The graph of voltage under test phase

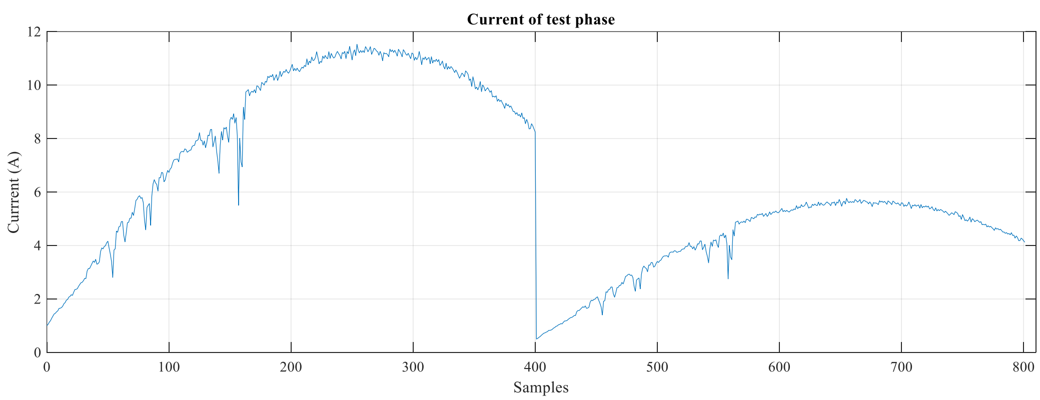

Figure IV.16: The graph of current under test phase

Figure IV.15 and IV.16 depicts the input electrical data. Figure IV.15 shows the recorded voltage is 270 (V) and slowly reduces in value as the week progresses to around 75 (V). For figure IV.16 the current recorded is 11.5 (A) and also it’s reduces to around 6 (A) during the second day.

c. Output data (voltage classification)

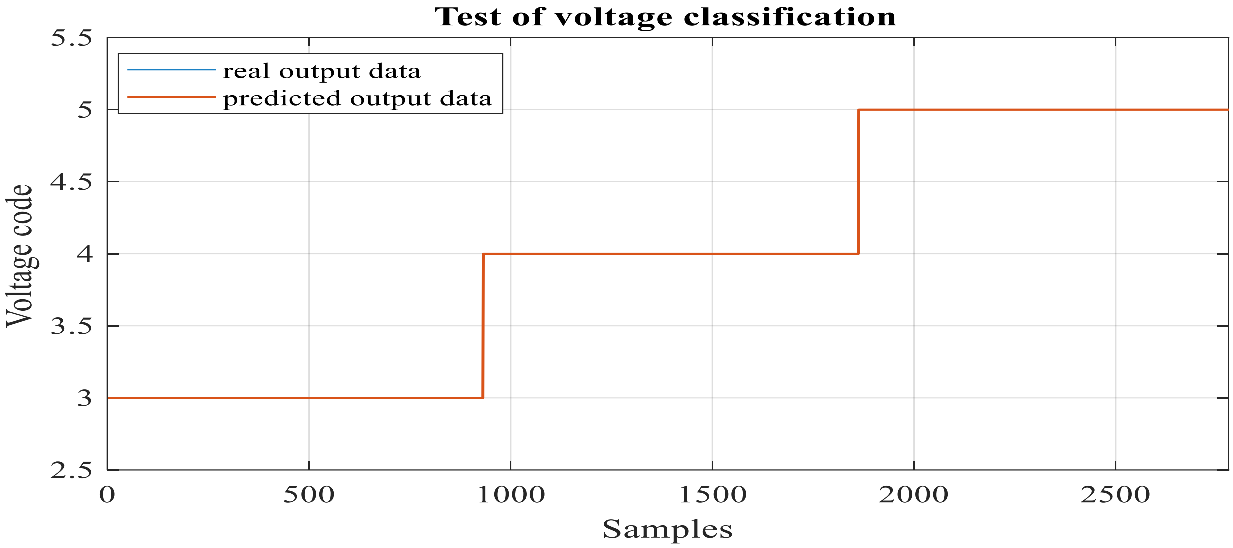

Figure IV.17: The graph of testing voltage classification

In the figure IV.17, can see that the real data is equal to the predicted data as the accuracy is around 100%.

d. Output data (current classification)

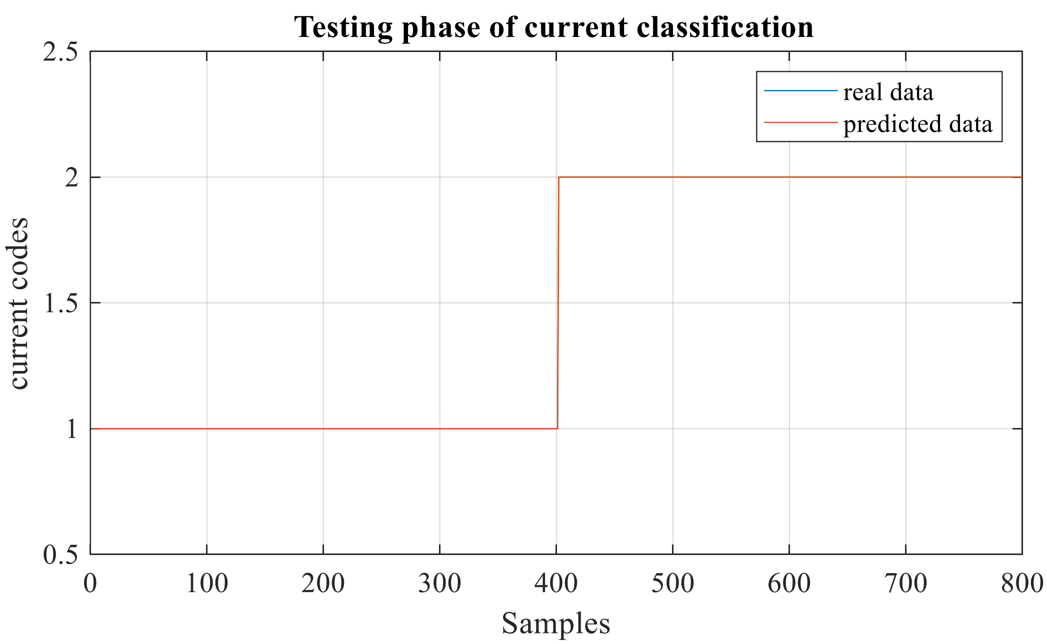

Figure IV.18: The graph of testing current classification

In our final figure of Figure IV.18, we notice that our range of samples fluctuated between code 1 and code 2. This indicates a steady rise of current with respect to solar irradiation.

IV.5 Output of KNN program:

IV.5.1 Training phase:

a. Voltage output classification using KNN

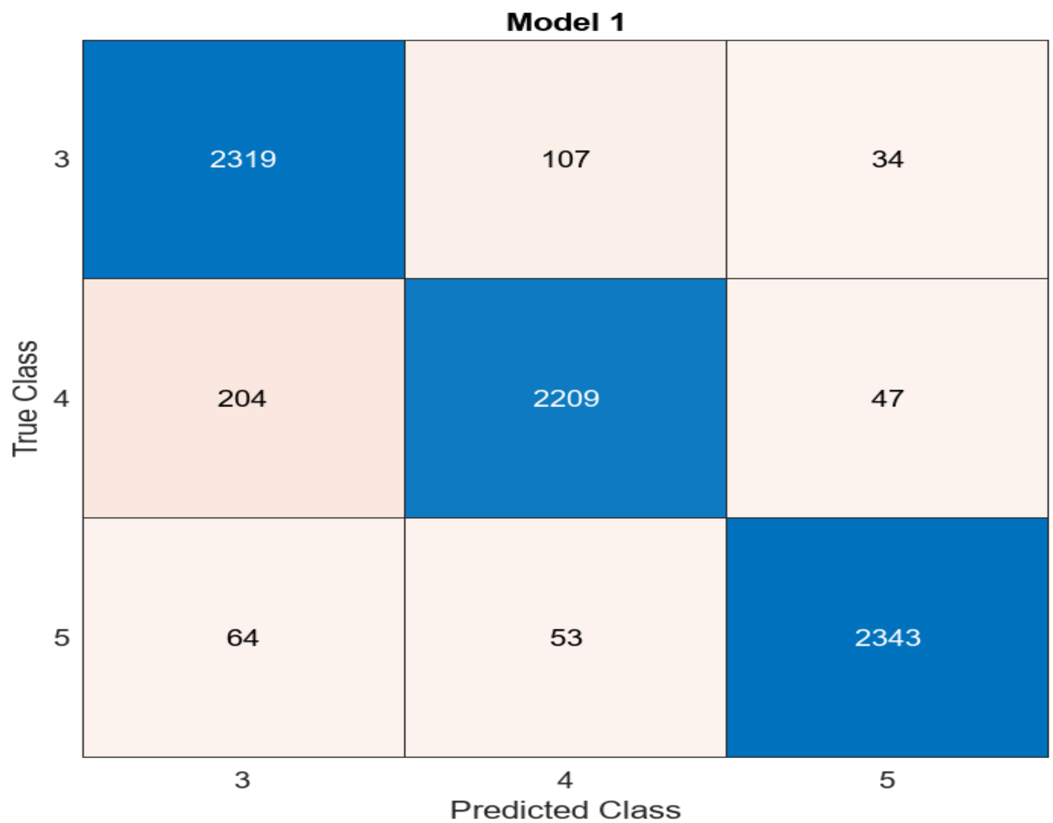

Figure IV.19: Training voltage output classification using KNN with confusion matrix

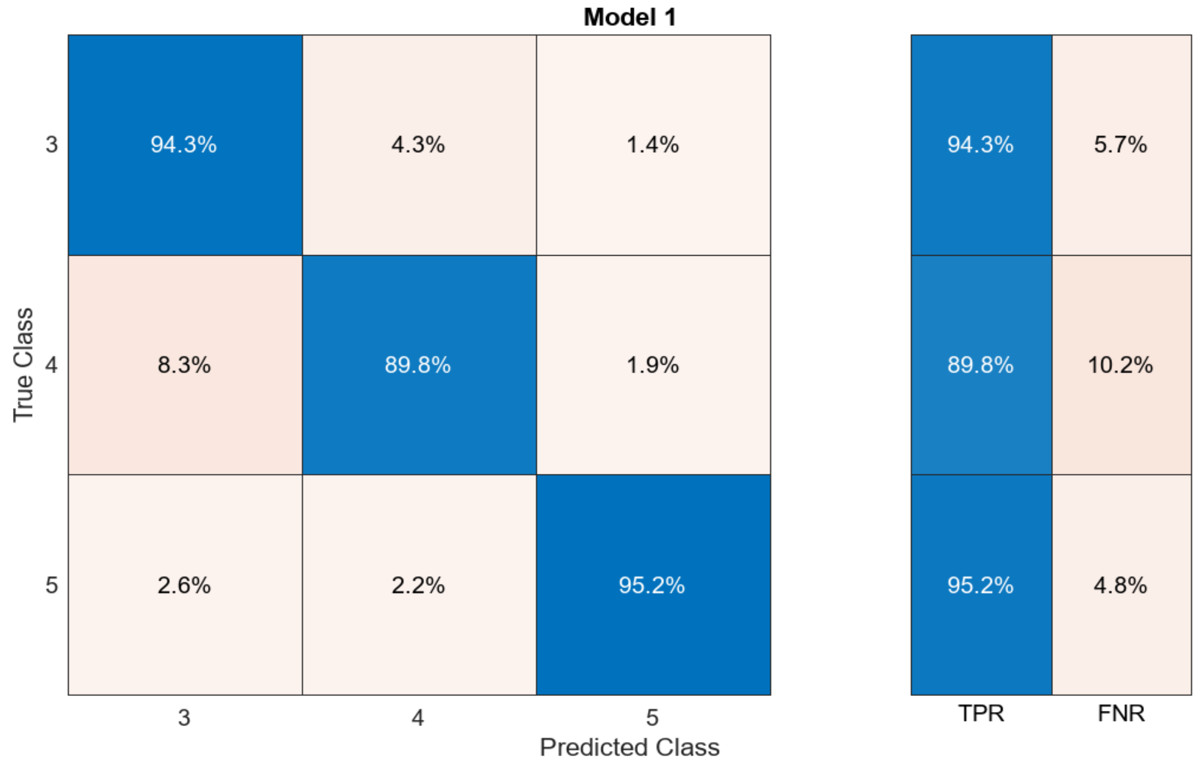

Figure IV.20: Training voltage output classification using KNN with confusion matrix using true positive rate plot

Figures IV.19 and IV.20 shows the confusion matrix before the implementation of the KNN. In figure IV.19, we get to see the two classes, the observed or real class and the predicted class. It further shows the number of samples in each class and figure IV.20 shows the accuracy of each class in percentage form.

b. Current output using KNN

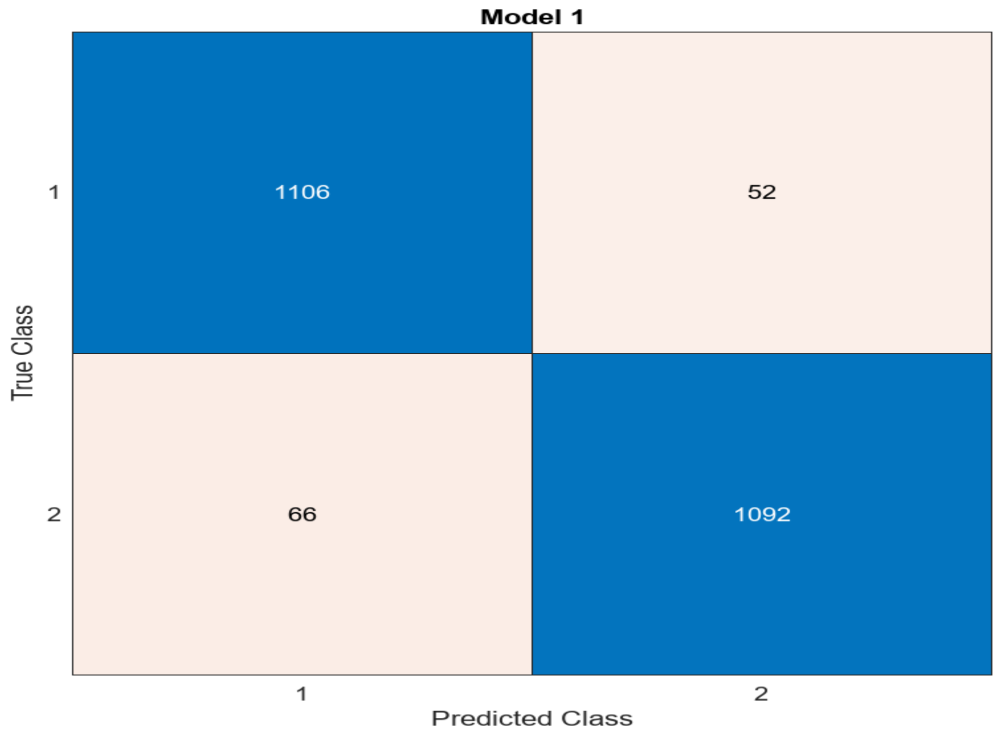

Figure IV.21: Training current output classification using KNN with confusion matrix

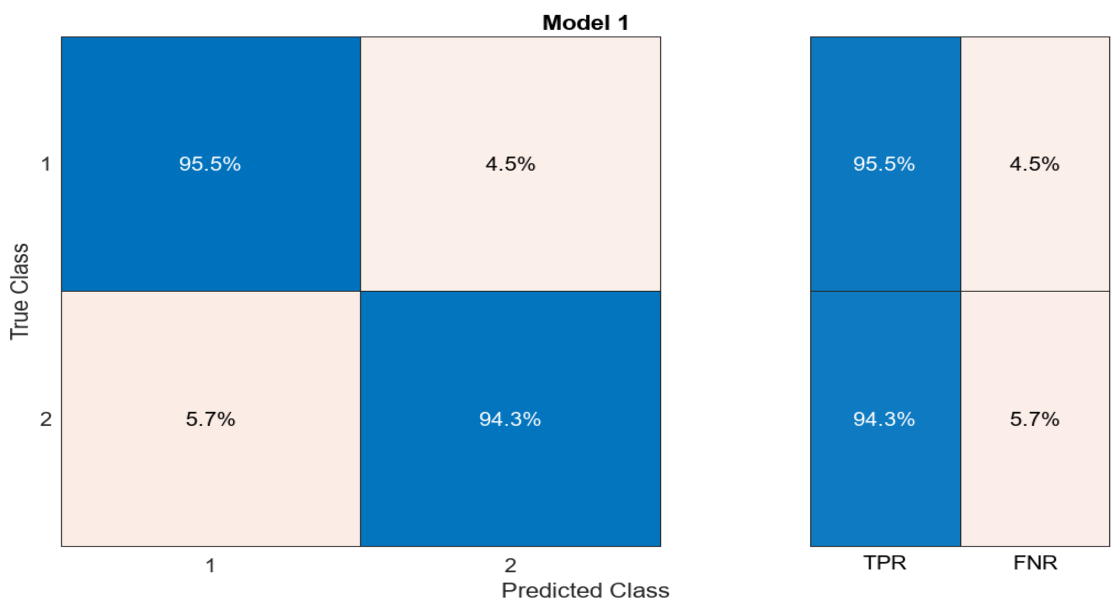

Figure IV.22: Training current output classification using KNN with confusion matrix using true positive rate plot

Figure IV.21 and IV.22 shows the confusion matrix for the current output. Figure IV.21 displays the amount or number of samples in each class and figure IV.22 tells us of the accuracy of each class. After this the program automatically calculates and gives us the overall efficiency or accuracy of 97.2% an indication that our simulation is accurate.

IV.5.2 Validation phase

a.Voltage Output classification using KNN

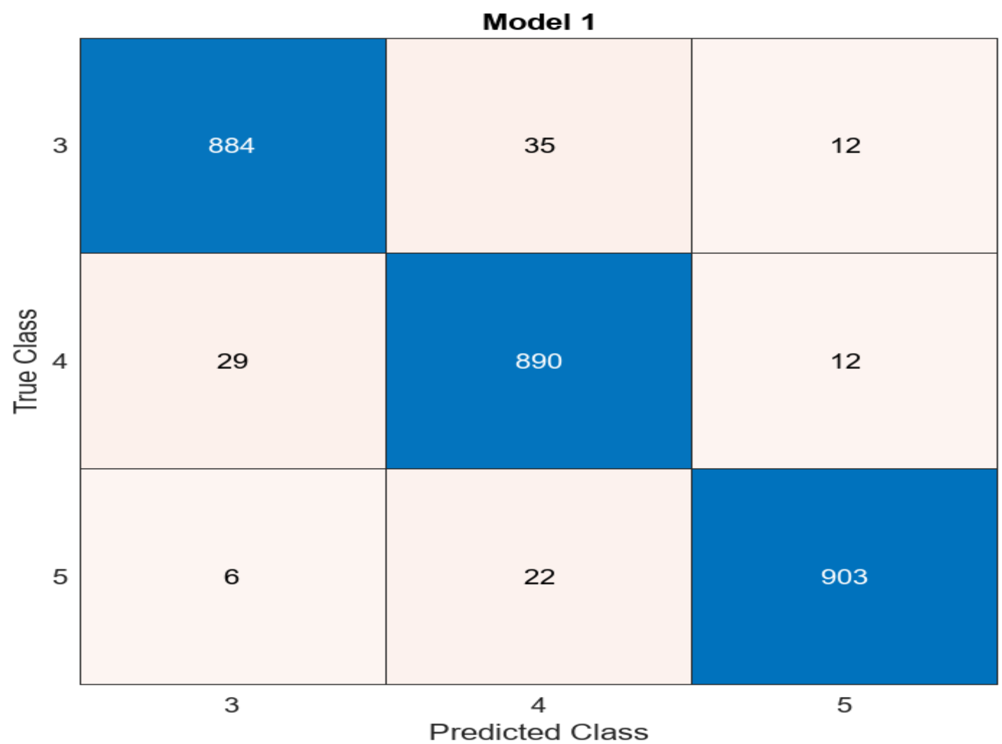

Figure IV.23: Validation voltage output classification using KNN with confusion matrix

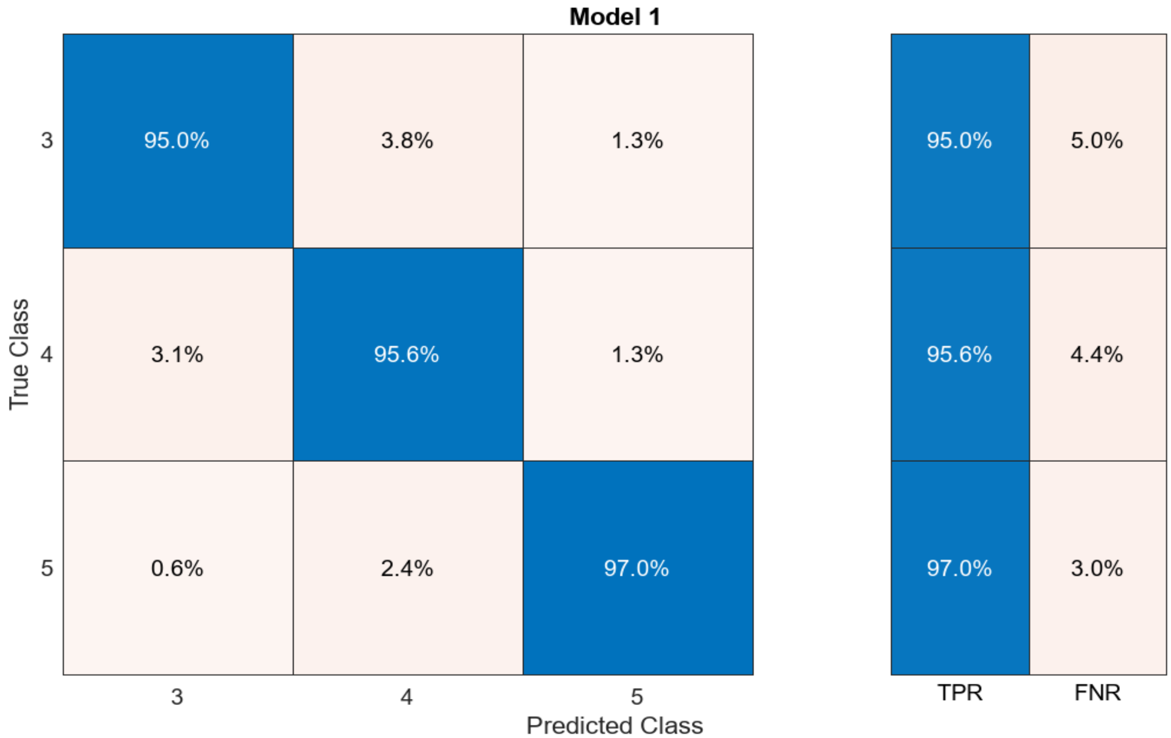

Figure IV.24: Validation voltage output classification using KNN with confusion matrix using true positive rate plot

As with all our confusion matrices, figures IV.23 and IV.24 represent the number of samples and their accuracy prediction. Only the accuracy differs.

b. Current output using KNN

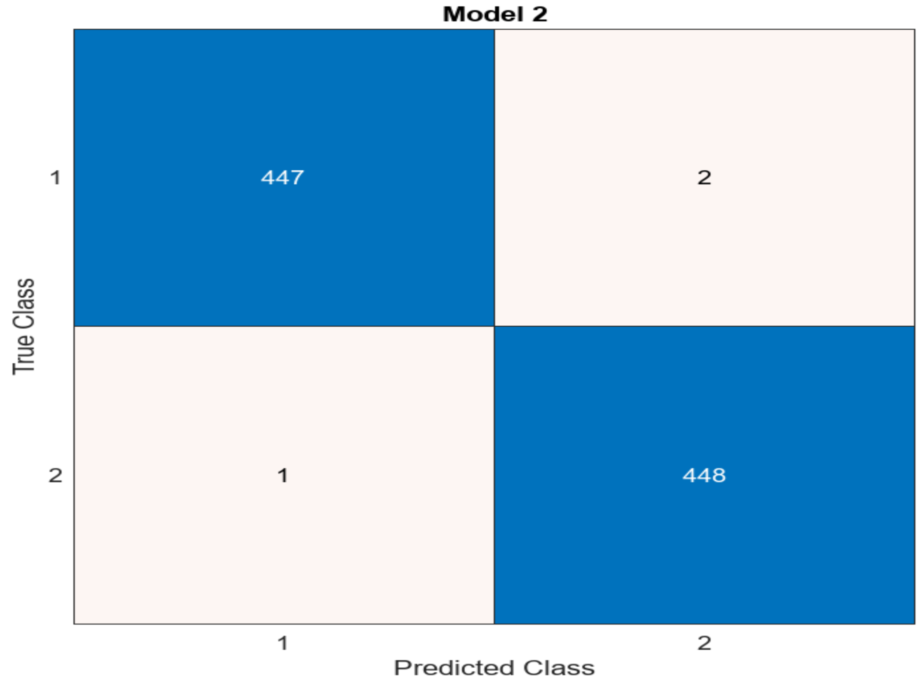

Figure IV.25: Validation current output classification using KNN with confusion matrix

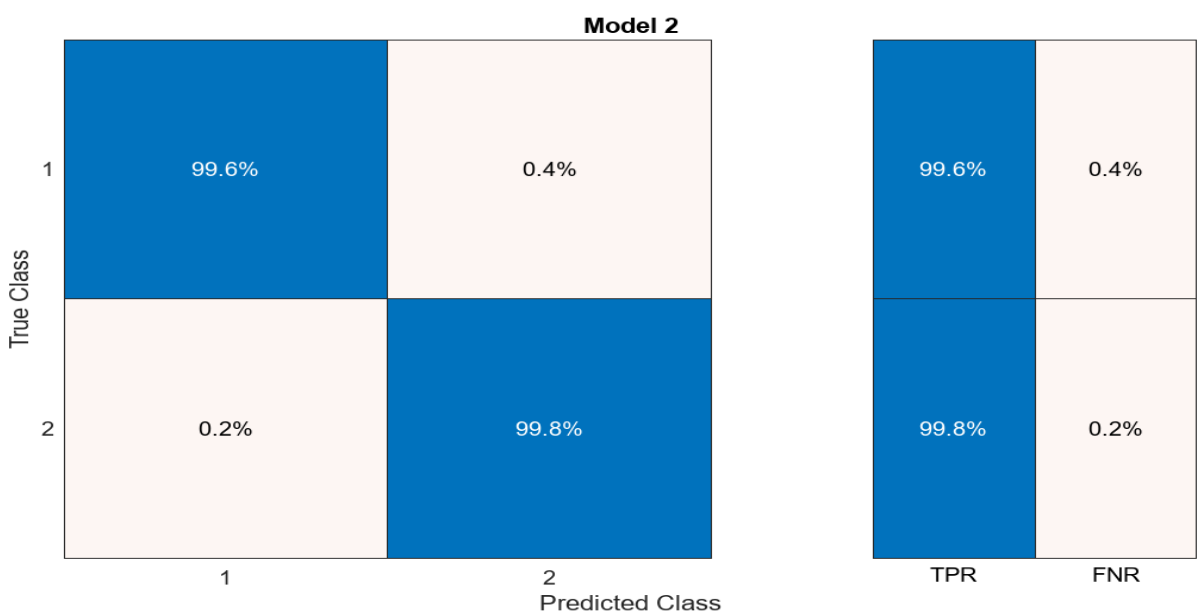

Figure IV.26: Validation current output classification using KNN with confusion matrix using true positive rate plot

The two figures IV.25 and IV.26 show the confusion matrices and the accuracy of the output data under validation.

IV.5.3 Test phase

a. Voltage output classification using KNN

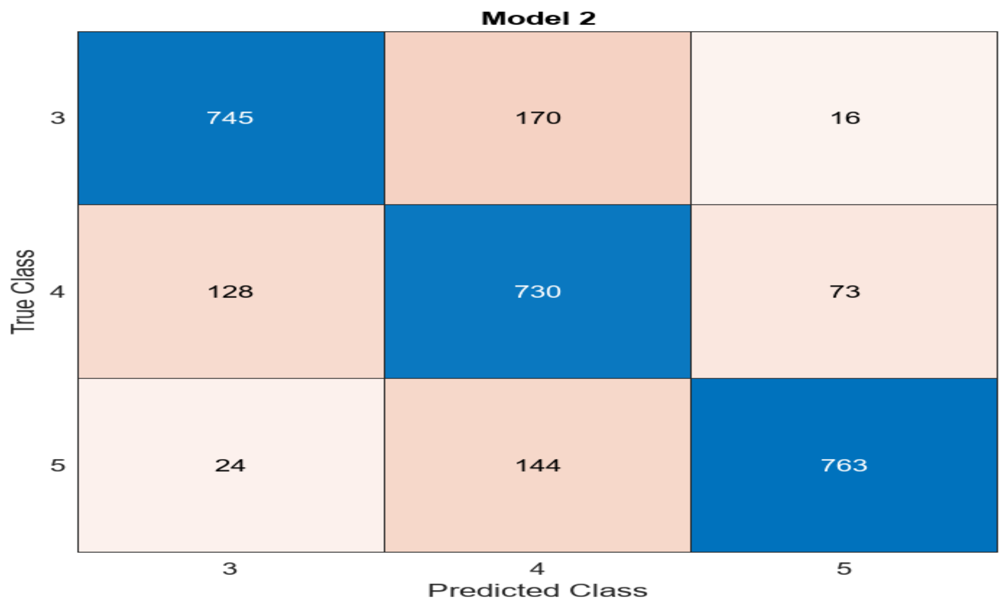

Figure IV.27: Testing voltage output classification using KNN with confusion matrix

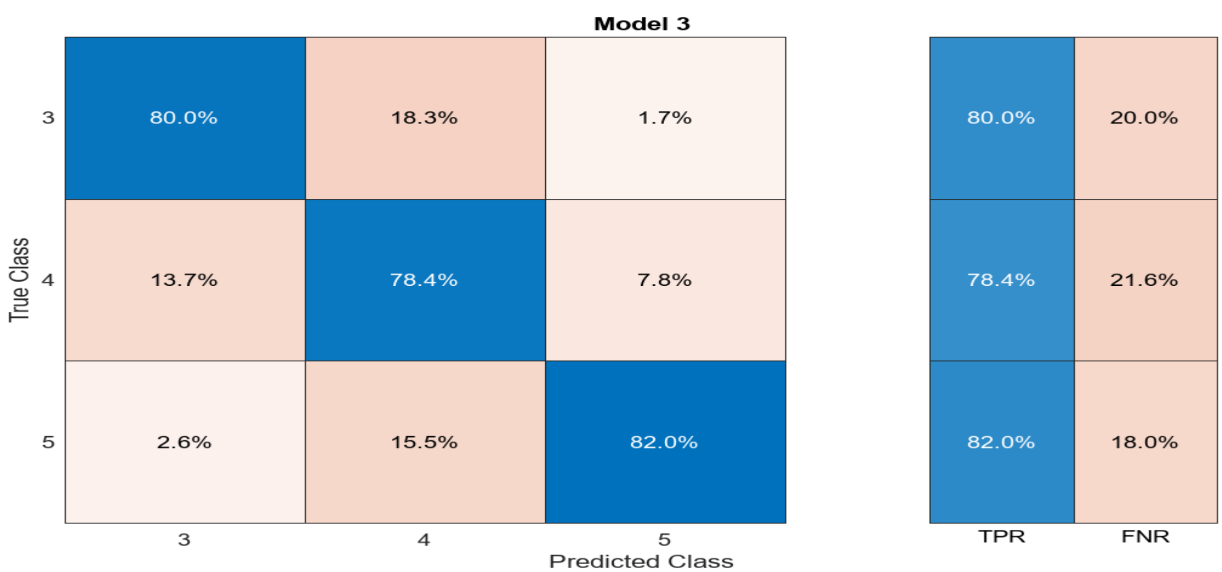

Figure IV.28: Testing voltage output classification using KNN with confusion matrix using true positive rate plot

The two figures show the confusion matrices of the voltage output data and the accuracy of individual sample codes.

b. Current output using KNN

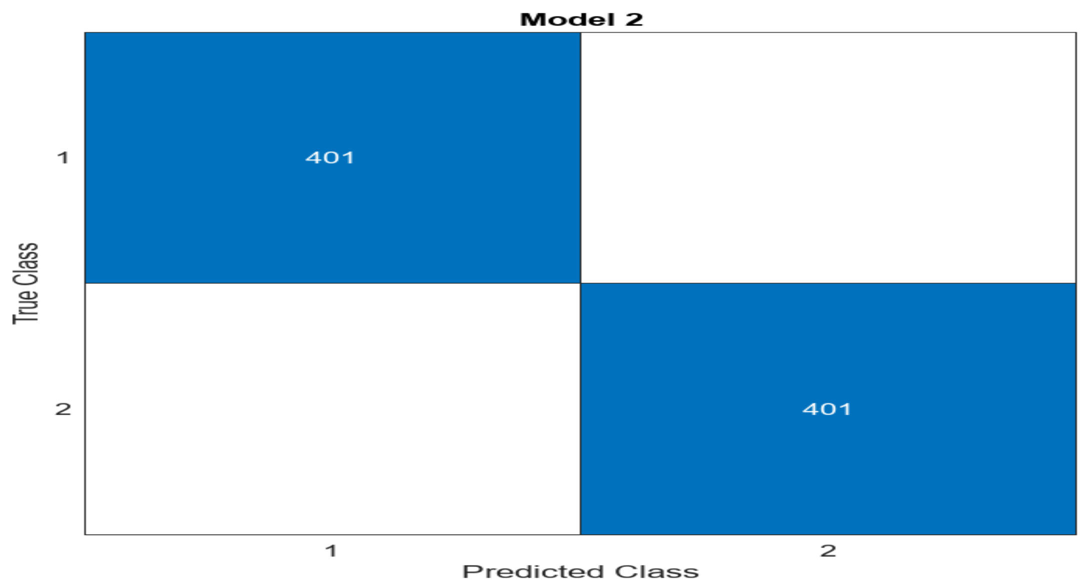

Figure IV.29: Testing KNN output for current with confusion matrix

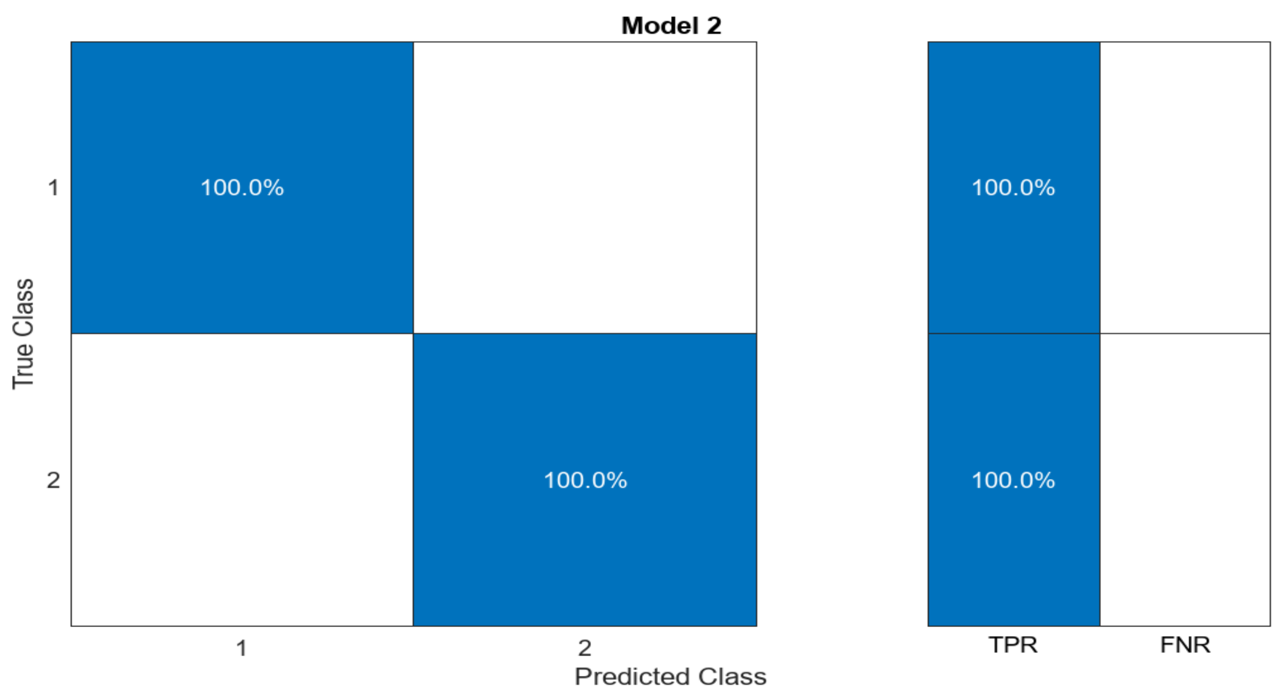

Figure IV.30: Testing KNN output current classification with confusion matrix using true positive rate plot

Figure IV.29 and Figure IV.30 displays the 802 samples in each confusion matrix and the second shows the accuracy 100% of the samples in each class that means we get the ideal classification.

GENERAL CONCLUSION

GENERAL CONCLUSION

Renewable energy is one of the most interesting and exciting fields of modern science and engineering. As the world population increases, the demand for energy becomes more and more challenging. Renewable energy seems to be the partial answer to that challenge hence why it's imperative to consider the best method of detecting problems in solar panels considering they are the most widely used source of renewable energy.

In this dissertation, we briefly talked about the different types of Photovoltaic systems and proceeded to elaborate more on the problems usually faced with engineers and technicians during maintenance of these systems. Which is why in chapter three we elaborated on how we can categorize and detect PV faults using data science and one of Artificial Intelligent (AI) algorithms KNN.

We believe our work has shed more light on the faults mainly encountered in PV systems and how to detect and categorize them. We equally believe that renewable energy technologies will continue to improving as we uncover more knowledge about the field and its integration with AI. We therefore hope and sincerely wish this will be helpful to all out there who are enthusiastic about the faults and causes of PV systems.

REFERENCE

REFERENCE

[1] S. Djeriou. A, Simulation d'un système photovoltaïque alimentant une machine asynchrone.

Mémoire de Magister, Université Farhat Abbas -Setif-, 2018.

[2] M.Jean-François REYNAUD. Recherches d'optimums d'´energie pour charge/décharge d'une batterie à technologie avancée dédiée à des applications photovoltaïques. Thèse de Doctorat. Université de TOULOUSE, 2011.

[3] R. Khenfer. Détection et isolation de défauts combinant des méthodes à base de données appliquées aux systèmes électro-énergétiques. Thèse de doctorat, Université Farhat Abbas -Setif-, 2018.

[4] H. Bouzeria. Modélisation et commande d'une chaine de conversion photovoltaïque. Thèse de Doctorat, Université de Batna 2 -Batna-, 2016.

[5] M. PETIBON Stéphane. Nouvelles architectures distribuées de gestion et de conversion de l'énergie pour les applications photovoltaïques. Thèse de Doctorat. Université de Toulouse, 2009.

[6] H. Bouzeria. Modélisation et commande d'une chaine de conversion photovoltaïque. Thèse de Doctorat, Université de Batna 2 , 2016.

[7] M Miwa, S. Yamanaka, H. Kawamura and H. Ohno, "Diagnosis of a Power Output of PV system",2018

[8] D. Sera, R. Teodorescu and P. Rodriguez, "Photovoltaic module diagnostics by series resistance monitoring and temperature and rated power estimation" in Industrial Electronics, 2008. IECON 2008. 34th Annual Conference of IEEE Orlando, FL 2008, pp. 2195 – 2199.

[9] http://www.batidepot.fr/solaire/photovoltaique/cellulephotovoltaique/cellulephotovoltaique.html

[10] Long Bunne, "Détection et localisation de défauts dans un système photovoltaïque",Thèse de doctorat, Université de Grenoble, Novembre 2011

[11] **S**. SPATARU, D. SERA, T. KEREKES, R. TEODORESCU, "Diagnostic method for Photovoltaic systems based on light I–V measurements", Solar Energy, Elsevier, Vol. 119,September 2015, pp. 29-44.

[12] S.Saravanan, R. S. Kumar, A. Prakash, T. Chinnadurai, R. Tiwari, N. Prabaharan, et al., "Photovoltaic array reconfiguration to extract maximum power under partially shaded conditions", in Distributed Energy Resources in Microgrids. 2019, Elsevier. p. 215-241.

[13] H.Bouzeria. Modélisation et commande d'une chaine de conversion photovoltaïque. Thèse de Doctorat, Université de Batna 2 -Batna-, 2016.

[14]. Arani, M.S.; Hejazi, M.A. The comprehensive study of electrical faults in PV arrays. J. Electr. Comput. Eng. 2016, 2016, 8712960

[15] M. K. Alam, F. H. Khan, J. Johnson,J. Flicker, PV arc-fault detection using spread spectrum time domain reflectometry (SSTDR). in 2014 IEEE energy conversion congress and exposition (ECCE). 2014. IEEE.

REFERENCE

[16] D. Sera. Real-time modelling, diagnostics and optimised MPPT for residential PV systems. Doctoral Thesis, Institute of energy technology-Alborg university, Denmark, 2009.

[17] Firth, S.K.; Lomas, K.J.; Rees, S.J. A simple model of PV system performance and its use in fault detection. Sol. Energy 2010, 84, 624–635.

[18] V. Tamrakar, S. Gupta,Y. Sawle. Single-diode and two-diode PV cell modeling using Matlab for studying characteristics of solar cell under varying conditions. Electrical & Computer Engineering: An International Journal. Vol 4, pp. 67-77, 2015.

[19] W. Rezgui. Système intégré pour la supervision et le diagnostic des défauts dans les systèmes de production d'énergies: les installations photovoltaïque. Thèse de Doctorat, Université de Batna 2 -Batna-, 2015.

[20] S. R. Madeti,S. N. Singh. A comprehensive study on different types of faults and detection techniques for solar photovoltaic system. Solar Energy. Vol 158, pp. 161-185, 2017;

[21] R. Platon, J. Martel, N. Woodruff,T. Chau. Online fault detection in PV systems. IEEE Transactions on Sustainable Energy. Vol 6, pp. 1200-1207, 2015;

[22] Cherifa Kara Mostefa Khelil, Badia Amrouche, Abou Soufiane Benyoucef, Kamel Kara and Aissa Chouder, new intelligent fault diagnosis(IFD) approach for grid-connected photovoltaic systems,2020

[23] T. Pei,X. Hao. A Fault Detection Method for Photovoltaic Systems Based on Voltage and Current Observation and Evaluation. Energies. Vol 12, pp. 1712, 2019;

[24] Sayed Zaki, Fault detection and diagnosis of photovoltaic system using fuzzy logic control

January 2019 E3S web of conferences by Cairo University and Zhu north china electric power university.

[25] [MEL]Mellit A, Benghanem M, Hadj Arab A, Guessoum A, Modeling ofsizing the photovoltaic system parameters using artificial neural network. In: Proceedings of IEEE conference on control application, 2003.

[26] Brett Lantz Discover how to build machine learning algorithms, prepare data and dig deep into data prediction. 2015, page 65-78

[27] Harsh Bhasin, Surbhi Bhatia, Application of Genetic Algorithms in Machine learning Computergard.com Faridabad, India * Amity University, Noida, Uttar Pradesh, India

[28] Tutorial on Support Vector Machine (SVM) Vikramaditya Jakkula, School of EECS, Washington State University, Pullman 99164.

[29] Machine Learning: Decision Trees CS540 Jerry Zhu University of Wisconsin-Madison

[30] T. Pei,X. Hao. A Fault Detection Method for Photovoltaic Systems Based on Voltage and Current Observation and Evaluation. Energies. Vol 12, pp. 1712, 2019;

[31] C Kara Mostefa Khelil, Amrouche and K kara, fault detection and diagnosis of GCPV systems using Bayesian neural network, university of Blida department of renewable energy technologies.

REFERENCE

Page 4 Journal of Physics: Conference Series, Volume 2208, The 4th International Symposium on Hydrogen Energy and Energy Technologies 2021 (HEET 2021) 19/11/2021 - 20/11/2021 Online

[32] Christopher M B 1997 Bayesian Neural Networks J of Brazilian Computer Society 10 1590.

[33] G. M. Tina, F. Cosentino,C. Ventura, Monitoring and diagnostics of photovoltaic power plants. In Renewable Energy in the Service of Mankind Vol II. 2016, Springer. p. 505-516.

[34] Mahmoud Dhimish, fault detection and performance analysis of photovoltaic installations, Doctoral thesis at university of Hudderfield page 23

[35] https://medium.com/swlh/k-nearest-neighbor-ca2593d7a3c4

[36] Brett Lantz Discover how to build machine learning algorithms, prepare data and dig deep into data prediction. 2015, page 69

[37]https://www.mathworks.com/discovery/machine-learning.html#machine-learning-with-matlab