

DEMOCRATIC AND POPULAR REPUBLIC OF ALGERIA  
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC  
RESEARCH

**University Kasdi Merbah– OUARGLA**

Faculty of New Technologies of  
Information and Communication



Computer Science and  
Communication Department

# Dissertation

**Field:** Mathematics and Computer Science

**Sector:** Computer Science

**Specialty:** Network Administration and Security

**Presented by:**

BOUMADDA Kenza

BOURENANE Nesrine

**Title:**

**Fault detection in photovoltaic power  
converter using machine learning  
algorithm**

**Jury members:**

Dr.KAFI M.Redouane      Supervisor      UKM Ouargla

Mr.MEZATI Messaoud      President      UKM Ouargla

Mr.BEKKARI Fouad      Examiner      UKM Ouargla

Academic Year: 2021 – 2022



# Dedication

**To the one,**

who encouraged me in my career and labored for me, my role model in life, my dear father, may God make you a crown over our heads.

**To the one,**

who gave birth to me and stayed up for me, my nanny and teacher, my dear mother, may God give you a long life. What I am today is a thank you and thanks to your prayers for me I hope you will be proud of me. May God protect you and your health. I love you so much.

**My older sister Sahar,** I wish you a happy life, may God protect you for me.

**My twin brothers Mounir and Imane,**

I wish you success in your studies and reach what you intend to achieve.

**My younger brothers Mohammed Taha and Abd el-Rahmane,**

You are the back, love and happiness of the home. I wish you success and access to the highest ranks. I love you all very much.

**To my dear grandmother,**

I love you so much, may God protect you and prolong your life.

**To my uncle and aunts without exception,**

thank you for your support and prayers for me, I love you so much.

**To my colleague and companion in this work,**

Nesrine, I appreciate your effort and struggle with me in order to write this thesis, may God bless us and preserve our friendship also Don't forget to mention my misses Asma and Maroua, I love you so much.

**To all my family, I love them and may God protect them.**

Kenza Boumadda

**To the soul of my aunt haciba,**

I wish you are here with us, I miss u so much we need u but your soul always with us. I love so much khalto (رحمة الله عليك).

**mom and dad,**

What I am today is thanks to you and thanks to your prayers for me, I hope you are proud of me. May God protect you and keep you healthy. I love you so much.

**My brother Ala,**

There are no words to describe how much I love you, you are my support, my back and my support in everything, I hope to see you always at the top. God save you for me. Love u bro.

**My sister Hadjer,**

My little sister, my friend and my secret hideaway, I wish you success in your studies and reaching what you intend to achieve. I love you.

**Akram, Abdelouhab and Abdelrahmane,**

My three little brothers, you are the back and the love. I wish you success and excellence and reaching the highest ranks. I love you all so much.

**Yema and My aunt Sabrina,**

My dear grandmother, I love you very much, may God protect you and give you long life. Aunt Sabrina, I resemble you in many things. I love you very much, may God protect you for me.

**To Fatima,**

Our new family member, and the piece of sugar in it. I can't forget how much you struggled with me writing my graduation thesis. I love you so much and thank you. May God bless you and my brother.

**To my colleague and companion in this work Kenza,**

My best friend may God bless us and keep our friendship. Thank you for everything. I love so much sweetie. I also do not forget to mention my beasties Marwa and Asma I love u so much.

**To all my family members I love you and may God protect them.**

**Nesrine BOURENANE**

## **Acknowledgment**

We thank Allah the Almighty for having given us the courage and the will to complete this present work.

First of all, we sincerely thank Dr. Kafi Mouhamed Redouane for his supervision and especially for the precious help he gave us throughout the work.

We also thank Dr.Rouabah Boubaker, Henna Hichem and AIADI Oussama for their help to finish our work.

We would also like to thank each of the members of the juries for the honor they do me by accepting to judge our work.

We do not forget before concluding our page of thanks, to address a thank you to the friend Kafi Omar for his patience and his unconditional support to complete our work.

We would like to thank all those who have contributed directly or indirectly to the completion of this work.

**BOUMADDA Kenza.**  
**BOURENANE Nesrine.**

## Abstract

Nowadays, Artificial intelligence applications have increased importance in renewable energies, such as photovoltaic systems, especially in data analysis and fault detection. Therefore, in this paper, a standalone solar photovoltaic system based on multicellular converter with flying capacitor fault detection using machine learning algorithms. Two control strategies; sliding mode and exact linearization controls are used in this paper in order to determine the more robustness and increased accuracy control. Simulation results with MATLAB using K-Nearest Neighbors (KNN) algorithm show that sliding mode control present high accuracy and improved robustness compared with exact linearization control.

**Key words:** Photovoltaic systems, power converter, KNN, sliding mode, exact-linearization mode, fault detection.

## Résumé

De nos jours, les applications d'intelligence artificielle ont une importance accrue dans les énergies renouvelables, telles que les systèmes photovoltaïques, en particulier dans l'analyse des données et la détection des pannes. Par conséquent, dans cette mémoire, un système solaire photovoltaïque autonome basé sur un convertisseur multicellulaire avec détection de défaut de condensateur volant à l'aide d'algorithmes d'apprentissage automatique. Deux stratégies de contrôle ; le mode glissant et les commandes de linéarisation exactes sont utilisés dans cet article afin de déterminer la plus grande robustesse et la commande de précision accrue. Les résultats de simulation avec MATLAB utilisant l'algorithme K-Nearest Neighbors (KNN) montrent que le contrôle par mode glissant présente une précision élevée et une robustesse améliorée par rapport au contrôle de linéarisation exacte.

**Mots clés :** Systèmes photovoltaïques, convertisseur de puissance, KNN, mode glissant, mode de linéarisation exacte, détection des défauts

## ملخص

في الوقت الحاضر، زادت أهمية تطبيقات الذكاء الاصطناعي في الطاقات المتجددة، مثل الأنظمة الكهروضوئية، وخاصة في تحليل البيانات واكتشاف الأخطاء. لذلك، في هذا البحث، نظام كهروضوئي قائم بذاته يعتمد على محول متعدد الخلايا مع مكثفات للكشف عن الأعطال باستخدام خوارزميات التعلم الآلي. استراتيجيتان للمراقبة تم استخدامهما في هذا البحث لتحديد وضع التحكم الأكثر دقة وأعلى متانة: الوضع المنزلق ووضع التحكم الخطي الدقيق.

تظهر نتائج المحاكاة باستخدام ماتلاب وتطبيق خوارزمية الجار الأقرب أن التحكم في الوضع المنزلق يُظهر دقة عالية ومتانة محسّنة مقارنةً بالتحكم في الوضع الخطي الدقيق.

**الكلمات المفتاحية:** الأنظمة الكهروضوئية، محول الطاقة، خوارزمية الجار الأقرب، الوضع المنزلق، الوضع الخطي الدقيق، اكتشاف الأخطاء.



# Table of Contents

|  |            |
|--|------------|
| <b>Dedication.....</b>   | <b>i</b>   |
| <b>Acknowledgment .....</b>  | <b>iii</b> |
| <b>Abstract .....</b>  | <b>iv</b>  |
| <b>Résumé.....</b>   | <b>v</b>   |
| <b>ملخص.....</b>   | <b>vi</b>  |
| <b>Table of Contents.....</b>  | <b>vii</b> |
| <b>List of Figures .....</b>   | <b>x</b>   |
| <b>List of Tables .....</b>  | <b>xii</b> |
| <b>General Introduction .....</b>  | <b>1</b>   |
| <b>1.Chapter 01: FAULT DETECTION OVERVIEW.....</b>                               | <b>3</b>   |
| 1.1 Introduction.....  | 3          |
| 1.2 Fault .....  | 3          |
| 1.3 Types of Faults .....  | 3          |
| 1.3.1 According to the importance of the fault .....                             | 4          |
| 1.3.2 According to the localization of the fault.....                            | 4          |
| 1.3.3 Depending on the temporal aspects .....                                    | 4          |
| 1.3.4 Depending on the way the faults affect to the behavior of the system ..... | 5          |
| 1.4 Fault Detection .....  | 5          |
| 1.5 Benefits of Detecting Fault in System.....                                   | 5          |
| 1.6 Fault Detection Methods .....  | 6          |
| 1.6.1 Data Methods and Signal Models .....                                       | 6          |
| 1.6.2 Process Model Based Methods.....   | 7          |
| 1.6.3 Knowledge Based Methods .....  | 7          |
| 1.7 Conclusion .....   | 8          |

|   |           |
|---|-----------|
| <b>2.Chapter 02: INTRODUCTION TO MACHINE LEARN-ING.....</b>                             | <b>9</b>  |
| 2.1 Introduction.....   | 9         |
| 2.2 The definition of Artificial Intelligence.....                                      | 9         |
| 2.3 Machine learning definition .....   | 10        |
| 2.4 The importance of Machine Learning.....   | 10        |
| 2.5 Machine Learning types.....   | 10        |
| 2.5.1 Supervised learning .....   | 10        |
| Supervised Learning Types.....  | 11        |
| Supervised Learning algorithms .....  | 11        |
| 2.5.2 Unsupervised learning .....   | 14        |
| Unsupervised Learning Type .....  | 14        |
| 2.5.3 Reinforcement learning.....   | 15        |
| 2.6 Application of Machine Learning.....  | 15        |
| 2.6.1 Image recognition.....  | 15        |
| 2.6.2 Smart assistants .....  | 16        |
| 2.6.3 Speech recognition.....   | 16        |
| 2.6.4 Cyber security .....  | 16        |
| 2.6.5 Customer service.....   | 17        |
| 2.7 Conclusion .....  | 17        |
| <br>  |           |
| <b>3.Chapter 03: Machine Learning and Fault-Detection inPhotovoltaic converter.....</b> | <b>18</b> |
| 3.1 Introduction.....   | 18        |
| 3.2 Photovoltaic system.....  | 18        |
| 3.3 Photovoltaic power converter .....  | 19        |
| 3.4 Benefits of Photovoltaic system .....   | 19        |
| 3.5 Fault Detection Using Machine Learning .....  | 19        |
| 3.6 Application of fault detection using machine Learning.....                          | 20        |
| 3.7 Conclusion .....  | 20        |
| <br>  |           |
| <b>4.Chapter 04: Implementation and Results.....</b>                                    | <b>21</b> |
| 4.1 Introduction.....   | 21        |
| 4.2 Work Environment .....  | 21        |
| 4.2.1 Hardware: .....   | 21        |
| 4.2.2 Software:.....  | 21        |
| MATLAB programming language.....  | 21        |
| MATLAB Window .....   | 21        |
| The Use of MATLAB .....   | 22        |
| Advantage and disadvantages of MATLAB .....   | 22        |
| 4.3 System and Data Set .....   | 23        |
| 4.4 Data Analysis.....  | 24        |

|                         |  |           |
|-------------------------|--|-----------|
| 4.5                     | Feature Selection .....  | 27        |
| 4.6                     | Application of KNN algorithm.....                                  | 28        |
| 4.7                     | Comparison between Sliding mode and exact-linearization mode ..... | 30        |
| 4.8                     | Conclusion .....   | 31        |
| <b>Conclusion .....</b> |  | <b>32</b> |
| <b>References.....</b>  |  | <b>33</b> |

## List of Figures

|    |   |    |
|----|---|----|
| 1  | <i>Fault models based on internal fault: faulty component...</i>                    | 4  |
| 2  | <i>Fault models based on faulty form.</i>   | 5  |
| 3  | <i>Fault detection methods.</i>   | 6  |
| 4  | <i>Fault detection with signal models.</i>  | 7  |
| 5  | <i>Process model-based fault detection.</i>   | 7  |
| 6  | <i>Knowledge Based Methods.</i>   | 8  |
| 7  | <i>Artificial Intelligence including Machine learning and Deep learning.....</i>    | 9  |
| 8  | <i>Machine Learning Idea.</i>   | 10 |
| 9  | <i>Supervised Learning Types.</i>   | 11 |
| 10 | <i>Neural network.</i>  | 12 |
| 11 | <i>Example of Linear regression.</i>  | 12 |
| 12 | <i>Example of SVM.</i>  | 13 |
| 13 | <i>Example of KNN.</i>  | 13 |
| 14 | <i>Example of Random Forest algorithm.</i>  | 14 |
| 15 | <i>Example of Clustering algorithm.</i>   | 15 |
| 16 | <i>Image recognition.</i>   | 15 |
| 17 | <i>Smart Assistance.</i>  | 16 |
| 18 | <i>Speech Recognizer.</i>   | 16 |
| 19 | <i>Cyber security using Machine Learning.</i>                                       | 17 |
| 20 | <i>Customer service with Machine Learning.</i>                                      | 17 |
| 21 | <i>Photovoltaic system.</i>   | 18 |
| 22 | <i>Photovoltaic Power Converter system.</i>   | 19 |
| 23 | <i>MATLAB Window</i>  | 22 |
| 24 | <i>Part from the balanced data set.</i>   | 23 |
| 25 | <i>C1 and C2 voltages in the four classes in the sliding control</i>                | 24 |
| 26 | <i>Load current in the four classes in the sliding mode of control</i>              | 25 |
| 27 | <i>C1 and C2 voltages in the four classes in the Exact linearization control.</i>   | 25 |
| 28 | <i>Load current in the four classes in the exact linearization mode of control.</i> | 26 |
| 29 | <i>the feature space on 2D of the two modes of control.</i>                         | 27 |
| 30 | <i>Feature space in 3 dimensions in the two modes of control.</i>                   | 27 |
| 31 | <i>Feature space of the two modes of control.</i>                                   | 28 |
| 32 | <i>part from the code of the split data.</i>  | 29 |
| 33 | <i>the creation of the model in the two modes.</i>                                  | 29 |

|    |  |    |
|----|--|----|
| 34 | <i>part of code to evaluate the model.....</i>                               | 29 |
| 35 | <i>comparison between the sliding mode and exact linearization mode.....</i> | 30 |

## **List of Tables**

|   |  |    |
|---|--|----|
| 1 | Advantage and Disadvantages of MATLAB. ....  | 22 |
| 2 | Table of comparison between the two modes according to the processing<br>time and accuracy ..... | 31 |

## General Introduction

Wind energy, geothermal energy, solar energy and other renewable energies are a source of energy all over the world today. The renewable energy has witnessed tremendous growth and development in recent years, especially solar energy, which has been increasingly adopted for residential, industrial and other applications as photovoltaic energy.[1]

However, the advantages provided by the photovoltaic systems (PV systems), they require continuous maintenance to ensure the reliability of the work efficiently. Therefore, early detection of malfunctions is required to avoid many abnormal events of the system.

The PV Converter system which is considered as a principal part of PV system [2] is one of the systems that can be applied to detect faults and thus reduce the possible damage by comparing the system data in the normal state with that extracted in the abnormal state. Capacitors are also the most common faults in power converters, accounting for 50 % of total defects [3].

Such problems in PV systems can have an impact on the system's efficiency and profitability, as well as the health and safety of workers and community members, if they are not detected promptly [4]. As a result, accurate and timely fault detection in a PV system is critical to avoiding fault progression and minimizing significant productivity losses. However, it is necessary to choose a good control of multicellular power converter in order to maintain its robustness and resistance to faults, among the control modes that they exist, this research uses two modes in order to compare between them; the sliding mode which is more attractive for multicellular converter [5] ,[6],[7],[8],as well as the exact linearization mode which uses a diffeomorphism mapping to change the original nonlinear model into a linear model, and then linear optimal control approaches [9],[10] to transform the control law obtained from the exactly transformed linear system back to the original nonlinear state-space. The need for smarter techniques in detecting and diagnosing faults has encouraged reliance on artificial intelligence methods that rely on machine learning to train models to detect errors and locate them to help maintenance engineers.[11]

The main objective of this thesis is to shed light on the application of machine learning algorithms to detect faults in the PV Converter system.

This document is composed of four chapters.

The first chapter, entitled "Overview of fault detection", presents the definition of the fault, the causes and the consequences of the faults. Also, we expose the technique of the detection of the faults and the methods used.

The second chapter is called "Introduction to machine learning". In this part of the work, we have presented the definition of machine learning terms, its types and its applications. Also, we expose some algorithms and the importance of using these terms in everyday

life.

The third chapter entitled "Machine Learning and Fault-Detection in Photovoltaic converter." explains the definition of our system to also use the relationship between failure detection and machine learning.

Finally, in the last chapter entitled "Implementation and results. », we present the steps and the tools used to apply our work also the result obtained.

We end this document with a conclusion that summarizes the essence of our work and our future work.



# **1. Chapter 01: FAULT DETECTION OVERVIEW**

## **1.1 Introduction**

One of the important goals of all machines or systems is to maintain a very high level of service continuity, and this is done through early detection of possible faults, thus greatly enhancing safety and reliability. The first chapter of this thesis includes an overview of fault detection through: fault definition and its types, fault detection and its various methods and its importance in ensuring the integrity of systems.

## **1.2 Fault**

An error is the loss of the ability of a system or structure to function well or provide its performance in the usual state (an unauthorized deviation of at least one characteristic of the system from the standard, acceptable, and usual state. [12]). It is sufficient to define an error as a change in system performance that results in an unacceptable or satisfactory decrease in quality.[13] The error occurs for several reasons, including:

- Due to corrosion.
- Execution errors.
- Design errors.
- Errors of use or human errors.

All these errors result in deviations in the input and output characteristics of the system from the normal state which leads to deterioration or even loss of the system mission.[14], it also has many consequences, for example: Economic losses, poor performance, waste of raw materials, environmental damage, human damage, low quality, low production, waste of energy.

## **1.3 Types of Faults**

The types of faults can be classified according to several criteria, including their importance, their place of occurrence, the time aspects of their occurrence and how they affect the system.[15]:

### 1.3.1 According to the importance of the fault

- Acceptable departure from the usual state.
- Minor malfunction or catastrophic failure
- The system stops or stops performing its functions for a period due to operating conditions.

### 1.3.2 According to the localization of the fault

- External fault:** Occurs when the interactions between the machine and its surroundings do not align well with the objectives.
- Internal fault:** The internal malfunction is classified depending on the faulty component: system, sensor, actuator.

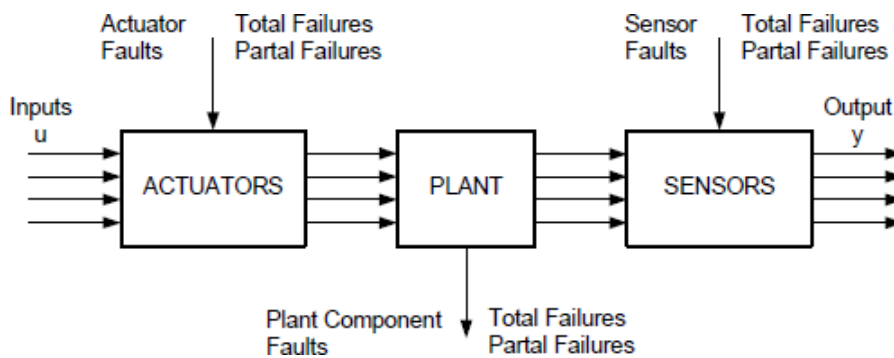


Figure 1. *Fault models based on internal fault: faulty component.*[16]

### 1.3.3 Depending on the temporal aspects

- Abrupt fault:** Sudden malfunction can cause severe damage. Form: step. Example: compensation.
- Incipient or evolutive fault:** This type slowly affects system performance. Shape: ramp, exponential, parabolic. Example: drift.
- Intermittent fault:** Happens with interruptions. Model: pulses.

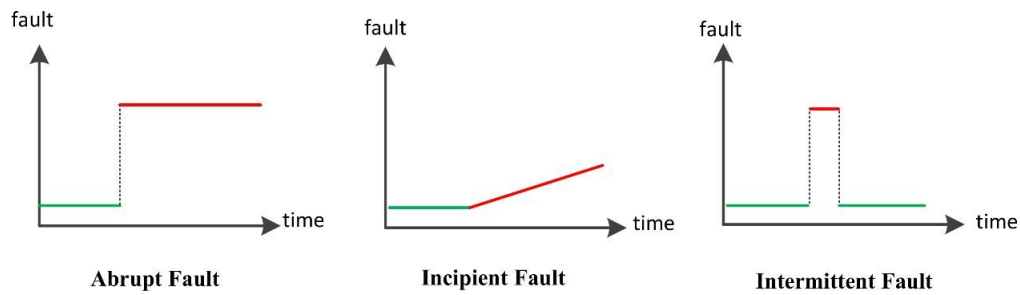


Figure 2. *Fault models based on faulty form.*[17]

### 1.3.4 Depending on the way the faults affect to the behavior of the system

- a. **Additive fault:** Changes occur in the system outputs and they estimate the size of the error and do not occur in the inputs: offsets in sensors and actuators, and disturbances.
- b. **Multiplicative fault:** Changes in output and input occur on the magnitude of the error: sensor gain, degradation, wear, and power loss.

## 1.4 Fault Detection

In the manufacturing sector, any unplanned failure caused by an unexpected machine can cause damage to the motor and it can be very costly. In order to increase the demand for reliability and safety of technical facilities, early detection of faults is a solution that ensures avoiding abnormal progress of events and mitigating damage by monitoring system data during operation and creating waste to determine whether the system operation is normal or not in view of normal system measurements. Fault detection consists of detecting errors in processes, actuators, and sensors using dependencies between different measurable signals. [18],[19].

## 1.5 Benefits of Detecting Fault in System

Due to the importance of detecting faults, it has many benefits, including:[20]

- Troubleshooting, diagnosing and monitoring the machine are very challenging and important topics. Through proper machine monitoring and fault detection plans.
- The safety and reliability of the system can be achieved.
- Avoid possible losses resulting from unplanned downtime.
- Early and initial fault detection is important in saving systems and machinery

maintenance costs, if the fault is detected before it occurs, in whole or in part and as quickly as possible.

## 1.6 Fault Detection Methods

There are three types of fault detection methods:

- Data Methods and Signal Models.
- Process Model Based Methods.
- Knowledge Based Methods

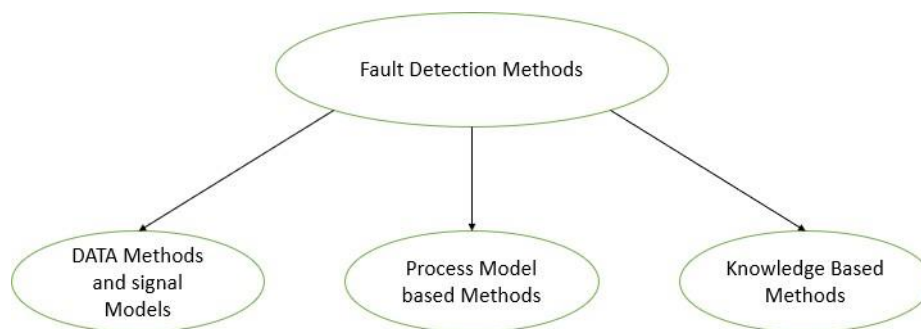


Figure 3. *Fault detection methods.*

### 1.6.1 Data Methods and Signal Models

Data-based methods work by exploiting only the experimental data available to them, but when it comes to changes in the signal associated with malfunctions in a process, the signal analysis feature is applied by comparing the mathematical models of the measured signal with measurements of natural behavior through training intelligence techniques Artificial neural networks such as artificial neural networks.[21],[18],[16] However, the approach of signal processing tools is limited to detection and have huge problems with false alarms because they use statistical tests (mean, or variance) to define detection thresholds.[22]

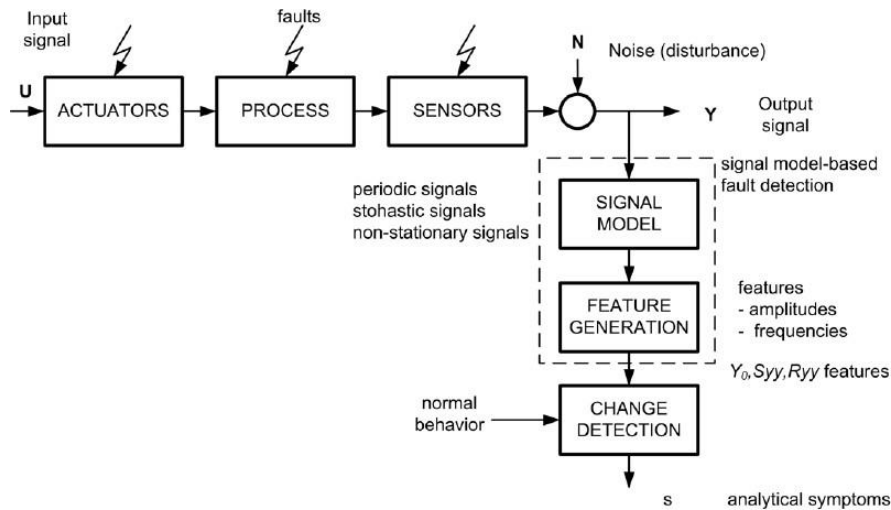


Figure 4. *Fault detection with signal models.*[16]

### 1.6.2 Process Model Based Methods

Model-based methods are considered one of the oldest strategies introduced in 1971 for fault diagnosis, which is based on the concept of analytical iteration by comparing the monitored system outputs with those obtained from an empirical mathematical model.[15],[19]

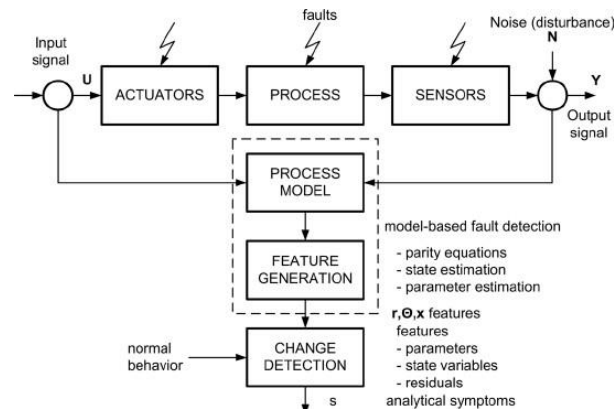


Figure 5. *Process model-based fault detection.*[16]

### 1.6.3 Knowledge Based Methods

Unlike model-based fault diagnosis methods, knowledge-based methods do not rely on a pre-known model of the system, but rather need historical data of the system. It can also use experts' decision to pass judgments. Artificial intelligence techniques have been applied in this method, including: causal analysis, fuzzy logic and expert systems that are used to imitate expert logic when faults are discovered.[19],[16],[17]

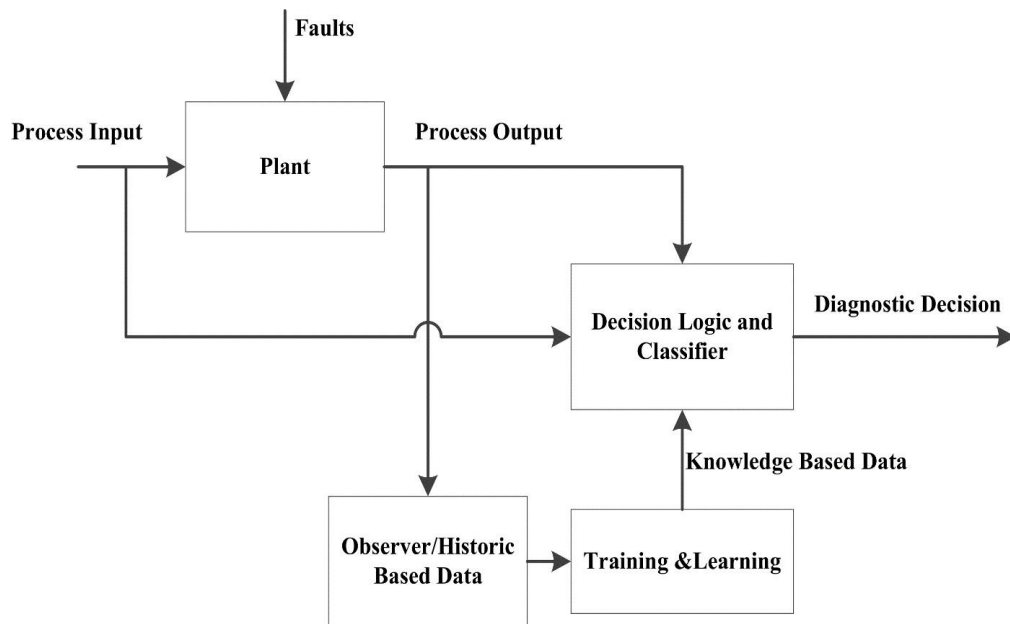


Figure 6. *Knowledge Based Methods.*[17]

## 1.7 Conclusion

In conclusion, early fault detection is very important in the industrial sector for machines and systems, for this machine learning is proposed as a solution to help to detect fault. the following chapter presents an introduction to machine learning.

## 2. Chapter 02: INTRODUCTION TO MACHINE LEARNING.

### 2.1 Introduction

Humans and animals use the most successful learning strategies of observing the actions being performed and trying to repeat them. With the development of technology in the 21st century and the appearance of artificial intelligence technology, especially machine learning, to create intelligent machines with human learning strategies. The second chapter provides an overview of machine learning, its types, algorithms and its importance in human life.

### 2.2 The definition of Artificial Intelligence

Artificial intelligence (AI) can be defined as imitating human intelligence by computer systems, making it perform tasks similar to humans and repeating human thinking and behavior. It also enables artificial intelligence to simulate intelligence that can recognize things, solve problems, and perform other actions that a human can perform. Artificial intelligence includes both machine learning and deep learning.[23]

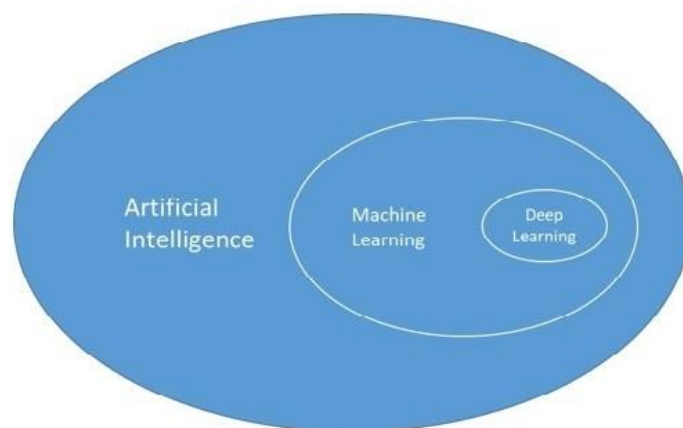


Figure 7. *Artificial Intelligence including Machine learning and Deep learning.*[23]

## 2.3 Machine learning definition

Machine learning (ML) is a tool for converting information into knowledge, by providing computers and machines with data and making them use computational algorithms to convert them into usable models. [24]

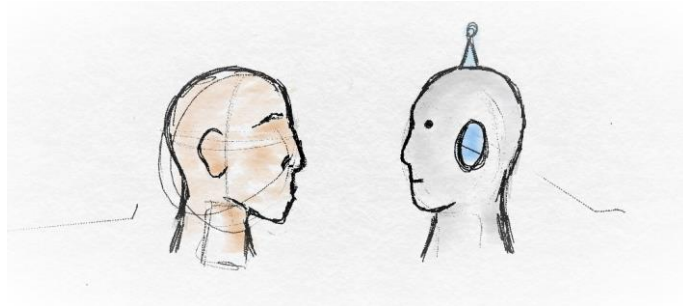


Figure 8. *Machine Learning Idea.*[24]

## 2.4 The importance of Machine Learning

Machine learning is important for many reasons including:[25]

- Allows organizations to see trends in customer behavior and business operating patterns, as well as support the development of new products.
- Machine learning is an essential part of the operations of many of today's leading companies such as Facebook, Google and Uber.
- Machine learning has become a significant competitive differentiator for many companies.

## 2.5 Machine Learning types

Machine learning classifiers fall into three primary categories:

- Supervised learning
- Unsupervised learning
- Reinforcement learning

### 2.5.1 Supervised learning

Supervised learning means to oversee or direct a certain activity and make sure it's done correctly, in this type of learning the machine learns under guidance. In Supervised learning



machines learns by using labeled data, that means the output is already knows to you, the model just needs to map the inputs to the output.[26]

## Supervised Learning Types

Supervised learning can be separated into two types: classification and regression.

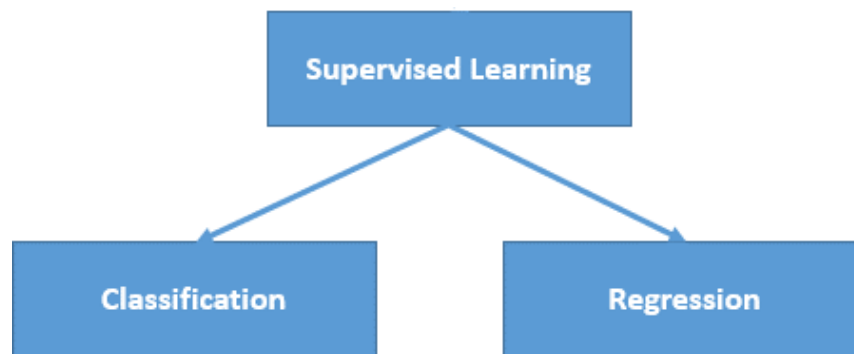


Figure 9. *Supervised Learning Types.*

- **Classification:** It is a process by which a particular set of data, whether structured or unstructured, is classified into a category. Common classification algorithms are [27]:
  - linear classifiers.
  - support vector machines (SVM).
  - k-nearest neighbors (KNN)
  - and random forest which are described in more detail below.
- **Regression:** it is a predictive modeling technique to find the relationship among two or more variable. Regression is primarily used for prediction and causal inference. Linear regression, logistical regression are popular regression algorithms.[24]

## Supervised Learning algorithms

Supervised machine learning techniques have many algorithms and computation methods. Here is a brief explanation of some of the most commonly used learning methods:

- a. **Neural networks:** It consists neurons that are arranged in layers, they take some input vector and convert it into an output. The process involves each neuron taking input and applying a function which is often a non-linear function to it and then pass the output to next layer.[28]
- b. **Naive Bayes:** It is classification algorithm based on Bayes theorem which gives



Figure 10. *Neural network.*[29]

an assumption of independence among predictors. In simple terms a Naïve Bayes classifier assumes that the presence of a feature in a class is unrelated to the presence of any other feature.[30]

- c. **Linear regression :** It is a classification algorithm in machine learning that uses one or more independent variables to determine an outcome. The outcome is measured with a dichotomous variable meaning it will have only two possible outcomes.[31]

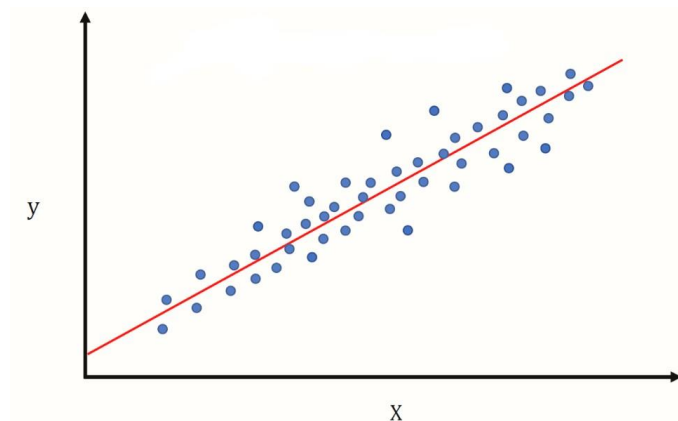


Figure 11. *Example of Linear regression.*[32]

- d. **Logistic regression:** Logistic regression is a classification algorithm, which is used to find out the probability of changing a variable. The nature of the variable to be predicted is dichotomous and is categorized based on the encoded data as either 1 (success/yes) or 0 (fails/no).[33]
- e. **Support vector machine:** It is one of the classification algorithms' known to guarantee its results as well as the simplicity of its use with a great need to know data extraction. Its principle is to separate data into groups using as "simple" limit as possible, such as the distance between different data groups and the limit separating

them is the maximum.[34]



Figure 12. Example of SVM.[35]

f. **K-nearest neighbor:**[36]The K-Nearest Neighbor (k-NN) algorithm is a classification algorithm. It is a simple and easy to implement algorithm. The algorithm receives a set of data labeled with corresponding output values that will be able to train and select the prediction model. This algorithm can then be used on new data to predict its corresponding output values, by:

- Select K number of neighbors.
- Calculate the distance (using the methods of calculating the distance as mentioned in [37].) from the unclassified point to the other points. Classification of the new point within the nearest neighbors.

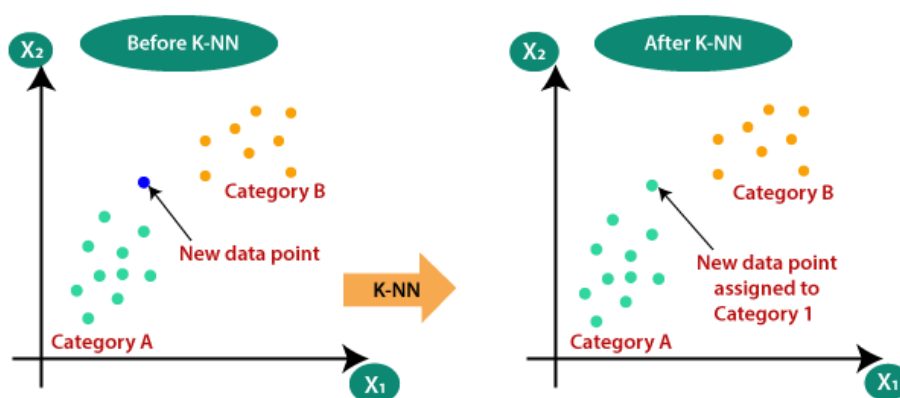


Figure 13. Example of KNN.[38]

g. **Random Forest:** They are also known as random decision forests. One of the classification algorithms whose principle lies in creating many decision trees on samples of data to be reduced to get and predict the best tree and thus get the appropriate solution.[39]

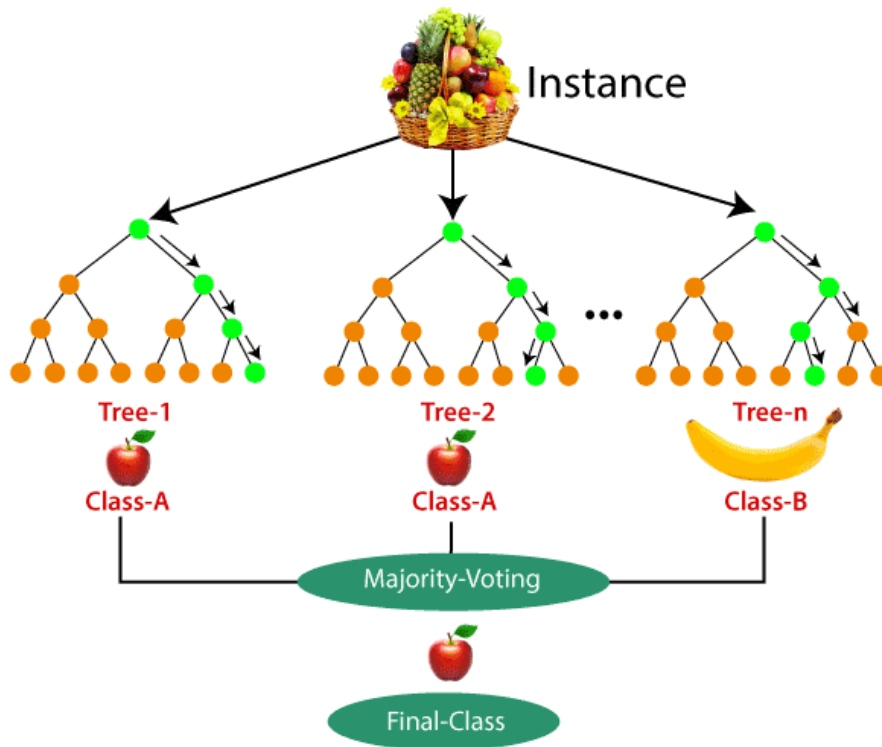


Figure 14. Example of Random Forest algorithm.[40]

## 2.5.2 Unsupervised learning

As its name suggests, it is one of the machine learning methods in which you are not subject to any kind of supervision, and you do not have a supervisor to provide guidance. In this type of machine learning it uses unlabeled data which means there is no static output variable, the model learns from the data it detects patterns and features in the data and returns the output.[26]

### Unsupervised Learning Type

- Clustering:** Clustering is an unsupervised machine learning technique. It collects samples of data and then categorizes them by distance and similarities. Clustering is important because it makes some assumptions about the similarities in existing data to break it down into different groups, but it's also valid.[41]

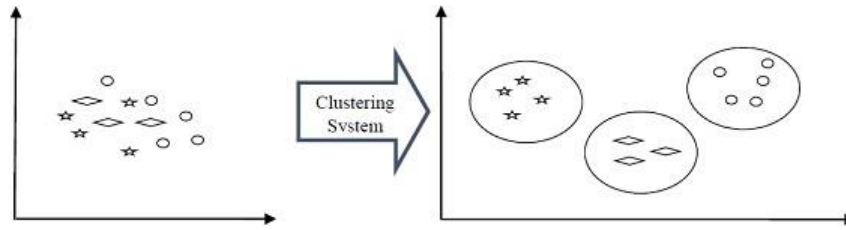


Figure 15. *Example of Clustering algorithm.*[42]

### 2.5.3 Reinforcement learning

Reinforcement learning is one of the machine learning methods that differ from other methods and are rarely used. The principle of this method works on using a worker who is trained in a specific period of time in order to interact with a specific environment, and some strategies are taken on how to interact with the environment as well as monitoring it in order to deduce procedures about its current condition.[26]

## 2.6 Application of Machine Learning

Machine learning technology has witnessed great growth day by day and is used in daily life through its applications. Here are some of the most popular machine learning applications:

### 2.6.1 Image recognition

The neural network takes a ready-made library of images and each pixel it contains is analyzed to discover features and objects in the images. A great example is Clearview's facial recognition technology that analyzes data from social media to get insights into people's faces that they can't get.[43]



Figure 16. *Image recognition.*[44]

### 2.6.2 Smart assistants

Machine learning introduced and developed the feature of intelligent assistants, which gives users the advantage of analyzing voice requests for their personal data, and can complete daily tasks that have been assigned to them and also get used to changing user needs. As an example: Alexa by Amazon for instance use all collected data to improve its pattern recognition skills and be able to address new needs.[45]

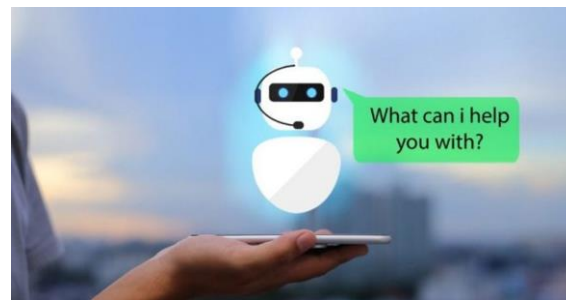


Figure 17. *Smart Assistance*. [46]

### 2.6.3 Speech recognition

Machine learning contributes to the development of this technology skills by adapting to speech styles, where users speak and use short terms as well as colloquial expression in their speech. The importance of machine learning lies in getting used to the different forms of speech by making the system train itself to learn all these different versions, and this is what is difficult for humans to do for millions of types of speech.[47],[48]



Figure 18. *Speech Recognizer*. [49]

### 2.6.4 Cyber security

Machine learning algorithms have an important role in cybersecurity, as they work to discover threats to the system and make it recognize them. The system also compares and analyzes similar cases to take safety and insurance measures.[50]

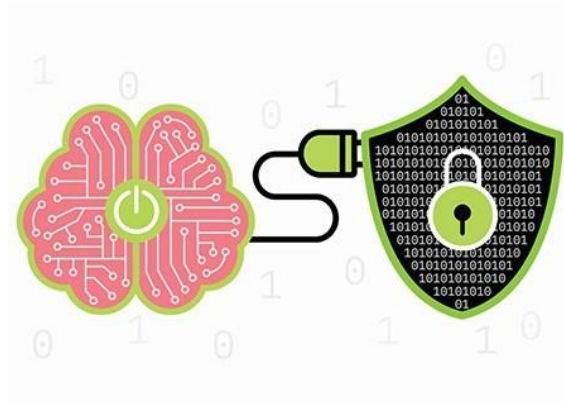


Figure 19. *Cyber security using Machine Learning.*[51]

### 2.6.5 Customer service

Machine learning algorithms analyze customer behavior from this data chat bot developers can know which issues to focus on as soon as several dozens of responses were confirmed the chat bots can learn on their own from daily interactions with clients getting better with each dialogue.[52]



Figure 20. *Customer service with Machine Learning.*[53]

## 2.7 Conclusion

This chapter, give as an overview about machine learning and his types, algorithms and the reason of his important. Next chapter, it will contains the system we work on, also the relation between machine learning and fault detection.

### 3. Chapter 03: Machine Learning and Fault-Detection in Photovoltaic converter.

#### 3.1 Introduction

We have seen in the two previous chapter, an overview about Fault detection and Machine Learning and their importance in human life and the stability of the system. In this chapter, it will contain the relation between the machine learning and fault detection in PV system. We will talk about: PV system definition, fault in PV system, Photovoltaic converter, types of faults in PV system, fault detection based on machine learning and application of fault detection using machine learning.

#### 3.2 Photovoltaic system

Photovoltaic (PV) system is a technology that converts the sun's energy into direct electric current by solar cells using semiconductors when the sun collides with the semiconductors inside the photovoltaic cell, the electrons are released and form an electric current[54]. The photovoltaic system and its application is a profound research project, facing the 21st century, which Collecting:

- Use of green renewable energy.
- Improving the environmental environment.
- Improving the living conditions of people as a whole.
- It will also have great benefits for the economy and politics, as well as for society.

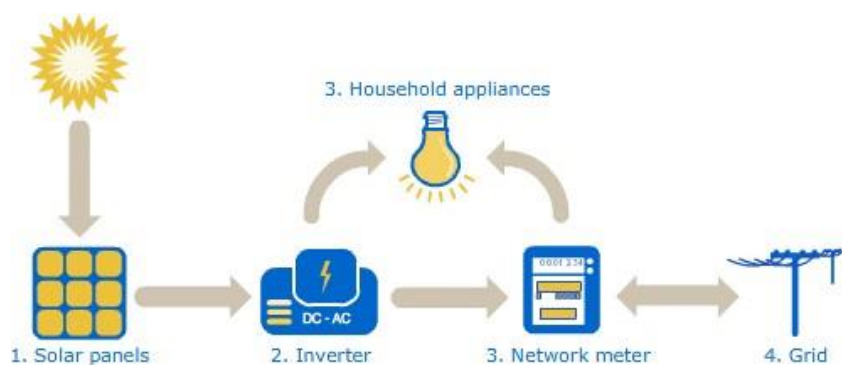


Figure 21. Photovoltaic system.[55]



### 3.3 Photovoltaic power converter

which is considered as a principal part of PV system [2]. A converter is an electrical circuit which accepts a DC input and generates AC output of a different voltage, that through which energy is produced by converting photons to a direct electric current, and then direct current to alternating current depending on the use.[56]

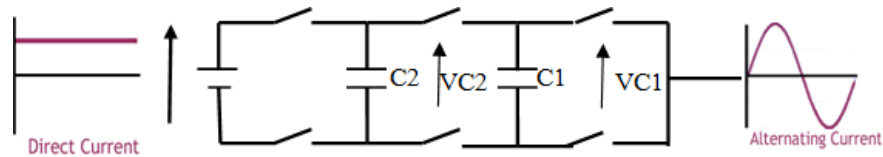


Figure 22. Photovoltaic Power Converter system.

### 3.4 Benefits of Photovoltaic system

There are many benefits to the PV system, including:[57]

- it shifts power generation from large centralized facilities to smaller.
- decentralized generation sites like the roof of your home.
- turning energy consumers into so-called prosumers who produce and consume their own electricity.
- Another benefit is that solar PV uses the most abundant renewable resource on earth estimates show that there is 10,000 times more solar energy coming to the earth's surface than our annual global demand for fossil fuels.

### 3.5 Fault Detection Using Machine Learning

Due to the importance of solar energy systems to generate photovoltaic energy, they need continuous maintenance to ensure efficient operation over time, and therefore it is necessary to develop methods for monitoring and maintaining these systems. The need for more efficient and intelligent strategies other than the traditional methods to discover and diagnose possible malfunctions in the system has necessitated relying on artificial intelligence methods, as they use machine learning methods. Over the years, the use of artificial intelligence has contributed to fault detection and loss prevention, through[11]:

- Quickly detecting defects and locating them.
- The ability to distinguish between different defects.

Among the techniques of artificial intelligence to detect faults, we mention them:

- Artificial neural networks (ANNs).
- Fuzzy Logic (FL).
- K-Nearest neighbor(K-NN) (which was used in this study).

### 3.6 Application of fault detection using machine Learning

Research development in the field of fault finding has been limited to nuclear power plants, aircraft, processing plants, the automobile industry and national defense as important and sensitive areas.[58],[59],[60]. Error detection is established today in many industries. This study includes some of the main categories of applications in which the field of fault detection played a role, including:[16]

- a. **Machines and Engines:** An internal combustion engine signal analysis technology using ANN, engine cylinder fault detection in, and marine diesel engine monitoring is introduced.
- b. **Manufacturing:** Fault detection is an essential part of automated electronics manufacturing systems, particularly in semiconductor manufacturing, and reliable fault detection is critical for maximum productivity.
- c. **Bearings and machines:** Faults are detected in the hydraulic system, and neural networks are used in rotating mechanical systems.
- d. **Aircraft:** Fault detection systems have great applications in the field of critical air engine control systems, in order to achieve a high degree of reliability. The detection of malfunctions of critical flight systems is described. A solution is introduced to find out the errors of the aviation fuel system.
- e. **Automotive systems:** Model-based fault detection adds functionality to the engine's electronic control unit (ECU) for the internal combustion engine. Fault detection for injection, combustion and engine transmission is described. The detection of malfunctions of defective components in the rail suspension is described.

### 3.7 Conclusion

The penultimate chapter of this study included the definition of the photovoltaic system and the benefits that it contributes to daily life. This chapter also included the errors that can occur in the system and how to discover them using machine learning. the last chapter, represent the application of machine learning algorithms on the system to detecting fault.

## 4. Chapter 04: Implementation and Results.

### 4.1 Introduction

After having seen on the previous chapters fault detection and machine learning and their use on a photovoltaic system. In this chapter we will represent tools that we use in this work, in order to visualize the data of the system and to implement the KNN methods in 2 modes.

### 4.2 Work Environment

#### 4.2.1 Hardware:

To carry out this work, we will use a material with characteristics which are as follows:

- **Manufacturer:** Toshiba
- **Rating:** 4.7 Windows performance index
- **Processor:** Intel (R) Core— i3-3217U @ 1.80 GHZ.
- **RAM:** 4.00 GB
- **System type:** 64 bite operating system.

#### 4.2.2 Software:

##### **MATLAB programming language**

MATLAB is a short form of name (matrix laboratory). It is a software which allows you to perform functions like matrix manipulation bloating of function and data implementation of algorithm creation of user interface and interfacing with programs written in other languages like C, C++, Java, Fortran and Python. It is also used by researchers and practitioners in control engineering as well as it allows you image processing and lot more.[61]

##### **MATLAB Window**

- **Current folder:** Shows all files and folders in MATLAB.
- **Workspace:** contains variables created or imported into MATLAB from data files

or other programs.

- **Command space:** is the main window where type, modify, and save commands.
- **Command history:** displays a history of statements you have run in your current and previous MATLAB sessions.

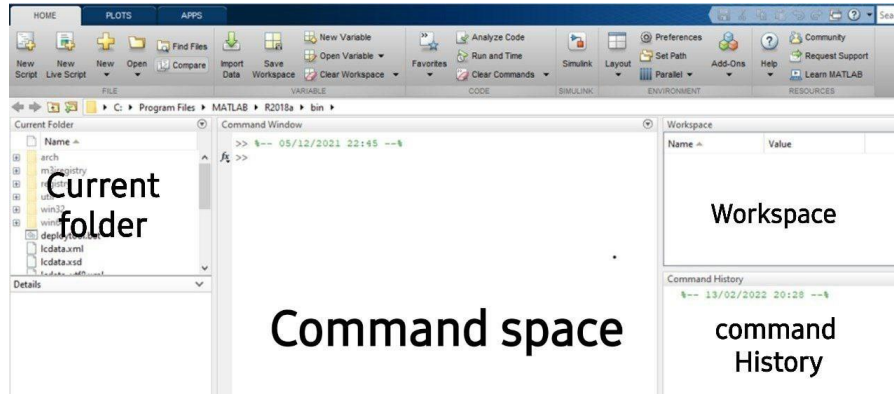


Figure 23. *MATLAB Window.*

## The Use of MATLAB

MATLAB is used in a lot of different ways by lots of people:[62],[63]

- Statistics and machine learning (ML).
- Control systems.
- Curve fitting.
- Electric vehicles designing.
- Deep learning.
- Financial analysis.
- Image processing.
- Signal Processing.
- Mapping, Aerospace, Text analysis, Audio toolbox.

## Advantage and disadvantages of MATLAB

Table 1. Advantage and Disadvantages of MATLAB.[63]

| Advantage                     | Disadvantage   |
|-------------------------------|--|
| Powerful built-in toolboxes   | Must be licensed   |
| Good visualization of results | Costly   |
| Graphical user interface      | It requires fast computer with sufficient amount of Memory |

*Continues...*

Table 1 – *Continues...*

| Advantage                                 | Disadvantage                             |
|---|--|
| Real time interfacing                     | It takes more time to execute operations |
| Auto generate C-code from Simulink models | It is not open-source software           |

### 4.3 System and Data Set

This study included the detection of malfunctions of multi-cell power inverters for solar panels, with a focus on faults in the capacitors. The balanced data (600001) was collected during the period of use of the converter using two methods of control they are: sliding mode and exact linearization mode, which included the data of capacitors voltage and load current. And since the behavior of converters differs from each other during their transition from temporary to permanent mode, this data was previously collected by removing the temporary mode because it does not include all the procedures of the converter.

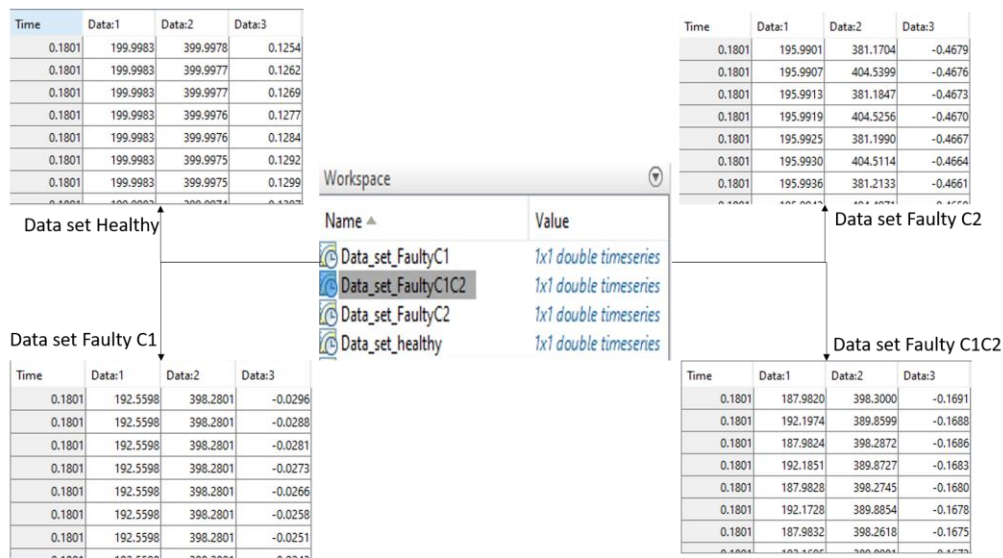


Figure 24. Part from the balanced data set.

By removing the temporary mode, the data will be equal to 220000. The studied system consists of three cells (two condensers) as shown in figure (24), so four possible classes were extracted:

- **Health mode:** It is when there is no error in either of the two capacitors.
- **Fault C1:** When there is an error in capacitor No.1.

- **Fault C2:** When there is an error in capacitor No.2.
- **Fault C1C2:** When an error occurs in both capacitors.

The purpose of providing this data is to create a model that can predict, locate or classify the previously mentioned cases. This form contains three parameters which are VC1 the tension of the first capacitor, VC2 the tension of the second capacitor and the load current.

#### 4.4 Data Analysis

After the data collected from the two capacitors were classified into four cases mentioned in the previous part, which were able to be plotted in terms of time. The following lists of figures in all cases of the both modes show respectively the signal tension of the both capacitors and the load current. The figure 25,26 shows the sliding mode data signal without transient mode, in the same way figure 27,28 shows the data signal of exact linearization.

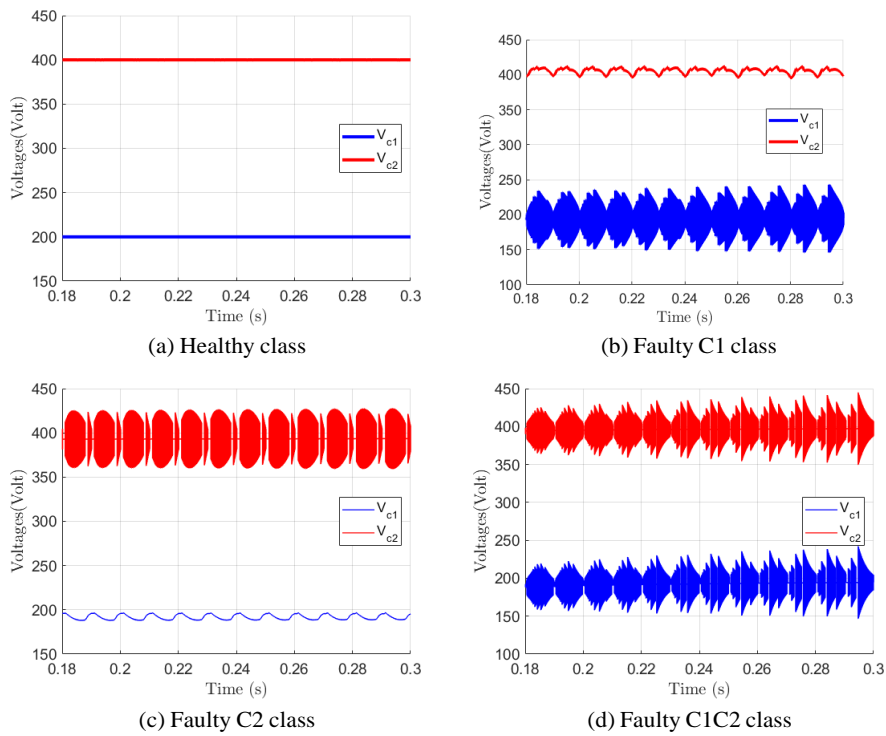


Figure 25. C1 and C2 voltages in the four classes in the sliding control

In the figures (25a) the signal appears in the normal state. However in the remaining classes (25b,25c,25d) the harmonic start to appears in the considering voltage.

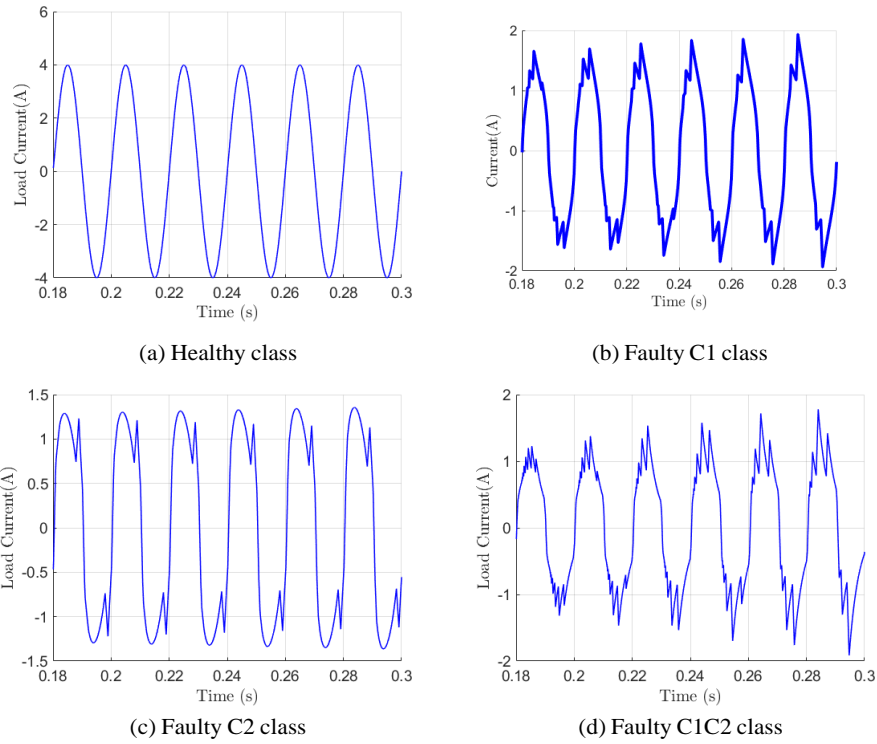


Figure 26. Load current in the four classes in the sliding mode of control

The current appears in its normal state as shown in the figure (26a). But in the other classes the harmonics appear in each peak of signals as the figures (26b, 26c, 26d) shows.

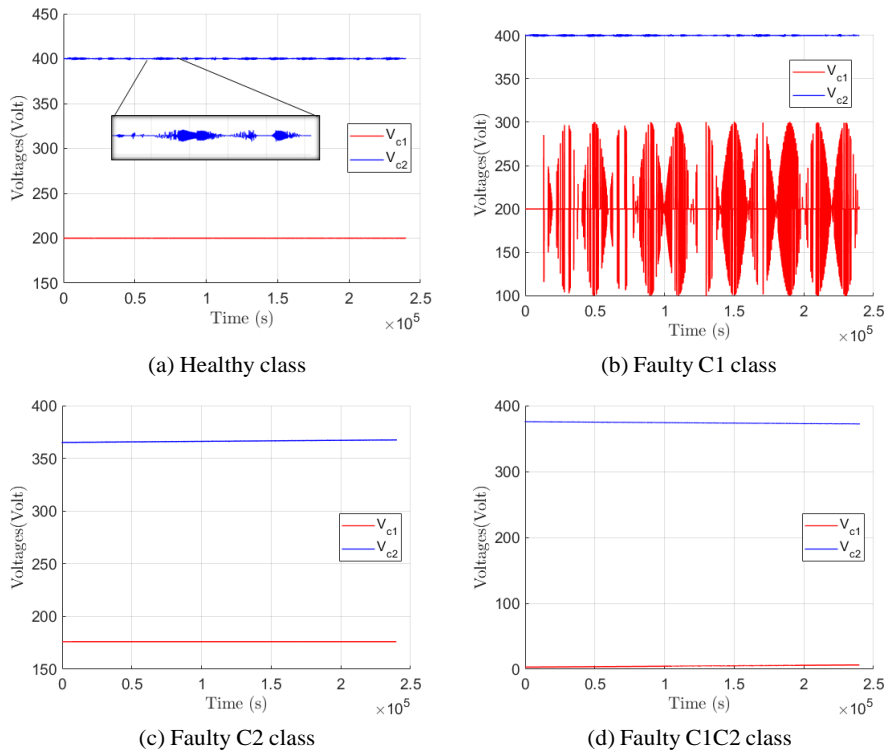


Figure 27. C1 and C2 voltages in the four classes in the Exact linearization control

In the healthy class (27a), the voltages in its normal state with appearance of a little noise in  $v_{c2}$ . In then faulty  $c1$  class (27b), the harmonics start to appear in  $v_{c1}$ . in the remaining classes (27c,27d) there are no harmonics but there is a decrease in the amplitude of the voltages

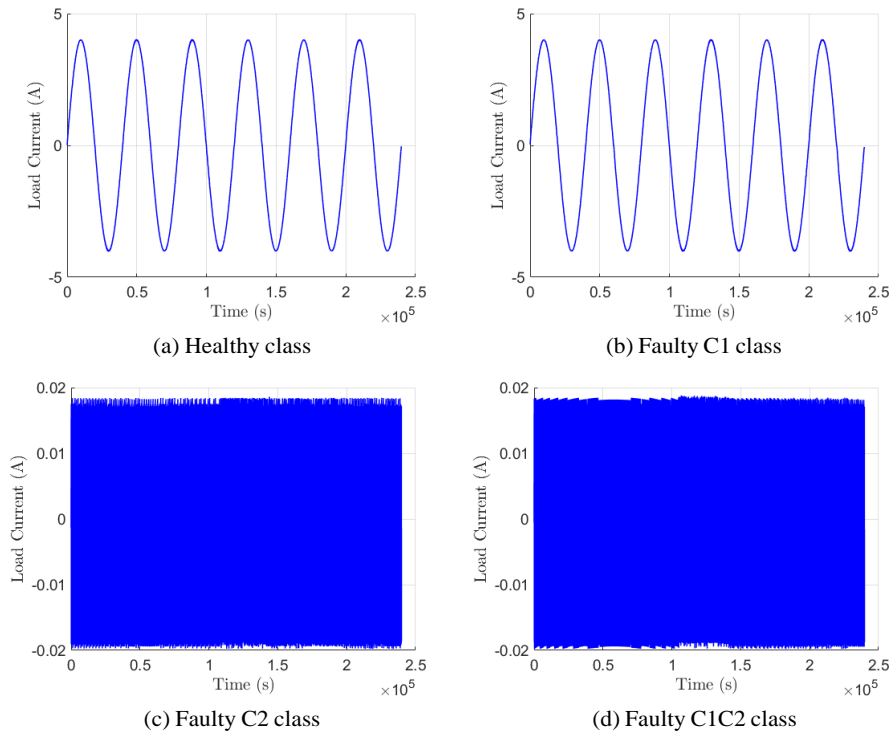


Figure 28. Load current in the four classes in the exact linearization mode of control

In the faulty  $C1$  class, the current appears in the normal state even the event of the failed of the first capacitor as the figure(28b) shows. In the remaining classes(28c,28d) there is a huge distortion of the current which make it useless.

The next step is to draw the feature space in order to follow the data's statistics and understand its characteristics, which allows us to see how all cases are distributed and identify overlapping cases, or cases with similar features, as well as distinguish between significant and non-significant features. To avoid data spacing, the data was normalized between the values of 1 as the highest value and 0 as the lowest value.



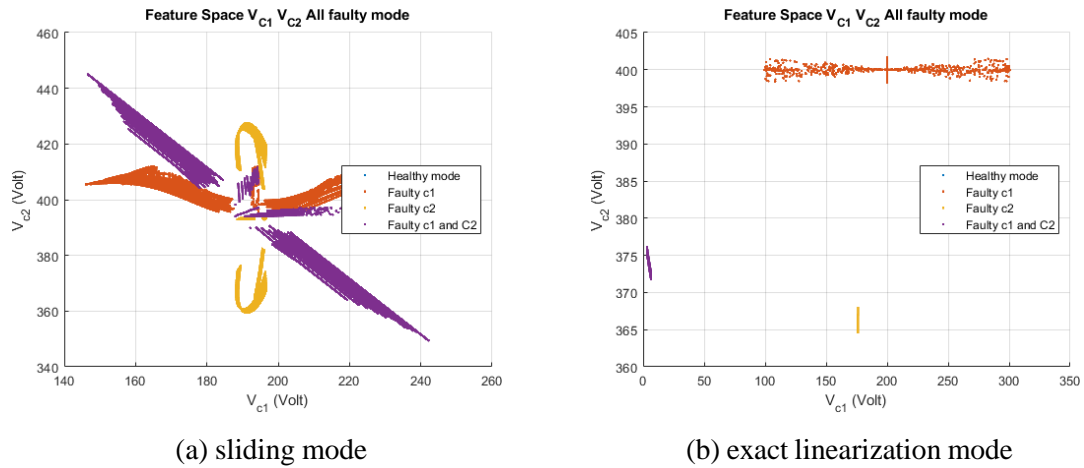


Figure 29. the feature space on 2D of the two modes of control.

As shown in the figures, the feature space appears in two dimensions, which does not allow us to see all cases at a good angle. In this case, another parameter has been added to show the feature space in 3 dimensions which is the current.

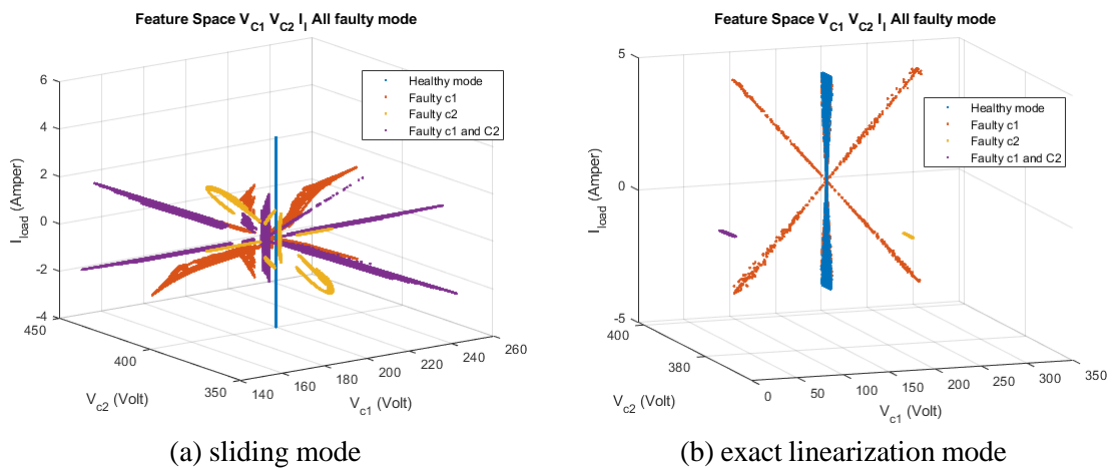


Figure 30. Feature space in 3 dimensions in the two modes of control.

In the phase of drawing a feature space that shows a fine and sparse distribution of possible system situations, it is found that the current changes with the change of the load, which means that the data-dependent model that contains the current data works with only the load data that was used in its construction.

## 4.5 Feature Selection

the goal here is to build a classification model that works with all types of loads, so features that have consistent behavior with all possible loads should be used to avoid building

too many models, and then dispensing with stress, the choice resorted to calculating the standard deviation (as calculated in [64] ) for the first capacitor whose average was measured, add it to the standard deviation of the second capacitor measured by its mean as well, because the importance of the standard deviation is to differentiate between the categories, in addition to the fact that the first and second capacitors differ in the value of the average in each category.

Therefore, changing the values of standard deviations whenever failures occur, as well as measuring these values by an average, ensures that the difference between each category appears. After normalizing the data between values of 1 as the highest value and 0 as the lowest value to avoid data spacing and choose the third dimension [23], the feature space is as follows:

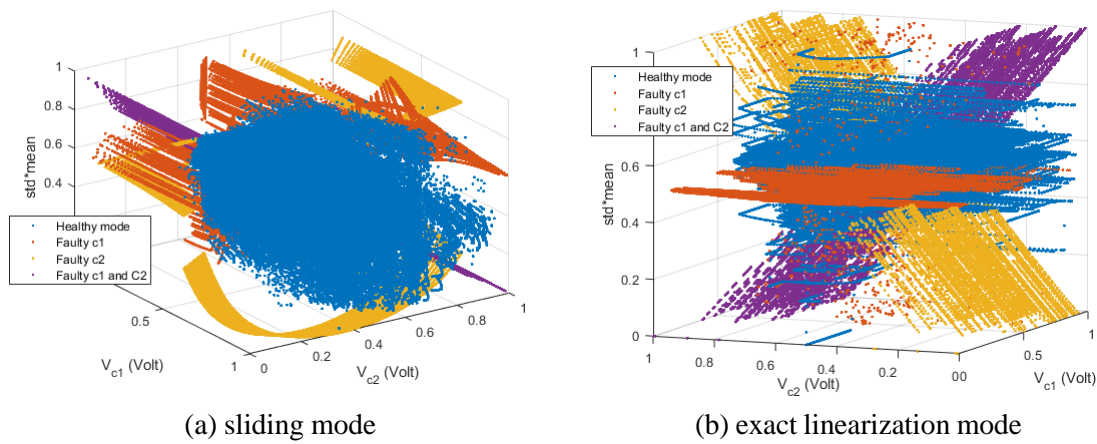


Figure 31. Feature space of the two modes of control.

## 4.6 Application of KNN algorithm

The previous steps focused on preparing the data in both cases “sliding mode and precise linear mode” in order to apply machine learning algorithms to the data. KNN-supervised learning was chosen to apply it to both modes. To apply the algorithms:

- The data was categorized into 4 classes: healthy mode and 3 faults modes.
- labeling the data as well we use the supervised learning, we chose:
  - "0": for the health mode.
  - "1": for the fault C1 mode.
  - "2": for the fault C2 mode.
  - "3": for the fault C1C2 mode.

the next step which is split the data into two part the training set and the testing set:

- training data represent 70% of all the data.
- testing data represent 30% of all the data.

```

##### Split all Data #####
[m n]=size(datasetall);
P = 0.70 ;
idx = randperm(m) ;
Training = datasetall(idx(1:round(P*m)),:);
Testing = datasetall(idx(round(P*m)+1:end),:);

```

Figure 32. part from the code of the split data.

After the split of the data in the two modes, a model has been created using the K-NN method because, it is multiclass algorithm based on the study that was conducted previously in a graduation note, in which it was concluded that K-NN is one of the best algorithms in detecting fault.[14]

```

##### make the model knn #####
mdl=fitcknn(Training,'label','NumNeighbors',2,'Standardize',1)

```

Figure 33. the creation of the model in the two modes.

In order to evaluate the classifier, model. The following metrics have been taken into account: accuracy, recall, specificity, precision, f1-score.

```

#####metrics of model evaluation#####
Tp=0;Tn=0;Fp=0;Fn=0;
test=allTesting.label;
for j=0:3
for i=1:length(allpre)
if((test(i)==j)&&(allpre(i)==j)) Tp=Tp+1; %true positive
end
if((test(i)==j)&&(allpre(i)~=j)) Fn=Fn+1; %false negative
end
if((test(i)~=j)&&(allpre(i)~=j))
if ((test(i))==allpre(i)) Tn=Tn+1; %true negative
else Fn=Fn+1;
end
end
if((test(i)~=j)&&(allpre(i)==j)) Fp=Fp+1; %false positive
end
end
end
Accuracy=(Tp+Tn)/(Tp+Tn+Fp+Fn);
Recall=(Tp)/(Tp+Fn);
Specificity=(Tn)/(Tn+Fp);
Precision=(Tp)/(Tp+Fp);
F1score=2*((Precision*Recall)/(Precision+Recall));
#####

```

Figure 34. part of code to evaluate the model.

## 4.7 Comparison between Sliding mode and exact-linearization mode

Looking at the processes that were applied to the data, as well as applying K-NN machine learning algorithms to the two control modes, and by comparing the load currents in both modes, we find the following:

- The sliding mode appears to reduce fault damage to the load by maintaining the general shape of the current, indicating its robustness against failure with the ability to diagnose the fault.
- On the contrary, we find that in the fine linear control mode, it does not guarantee the safety of the load and does not facilitate the identification of the capacitor in the event of a fault.

As for the comparison of accuracy in both modes, it can be said that the sliding mode classification model is more accurate than the exact linearization mode classification model. The histogram listed below illustrates the behavior of the proposed structure of the diagnostic approach in terms of accuracy, recall, specificity, precision and f1-score:

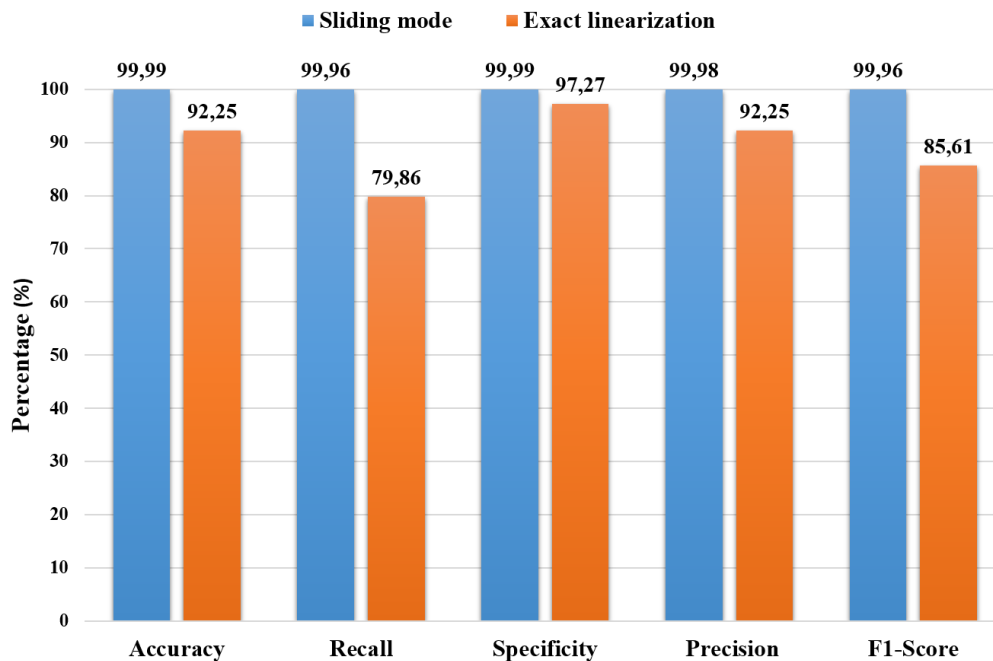


Figure 35. comparison between the sliding mode and exact linearization mode.

Furthermore, in order to compare between the two modes of control (sliding mode and exact linearization mode) in term of accuracy and the processing time. The following table represent the result with changing the K value.

|                             | K-Number           | 1     | 2     | 3     | 4     | 5                   |
|-----------------------------|--------------------|-------|-------|-------|-------|---------------------|
| Sliding mode control        | Accuracy(%)        | 99.99 | 99.98 | 99.99 | 99.98 | <b><u>99.99</u></b> |
|                             | Processing Time(s) | 4.40  | 4.39  | 4.39  | 4.39  | <b><u>4.39</u></b>  |
| Exact linearization control | Accuracy(%)        | 91.42 | 91.91 | 92.13 | 92.19 | <b><u>92.25</u></b> |
|                             | Processing Time(s) | 4.28  | 3.87  | 4.29  | 4.28  | <b><u>4.28</u></b>  |

Table 2. Table of comparison between the two modes according to the processing time and accuracy.

As shown in the above table, the value of K in case equal 5 gives the best results in the two modes of controls.

#### **4.8 Conclusion**

The last chapter of this study, contains tools used to applicant K-NN algorithms for the two modes of control. It also contains the result obtained. Also the comparison between the two modes to know the mode who give better results.

## Conclusion

As part of this memory, early detection of potential faults in photovoltaic transducers was presented using machine learning techniques that were selected based on a previous study, which aimed to find out which of the two system control modes is more efficient and performs better. In order to better understand the topic, the study is divided into four chapters:

- As a first chapter, it included an overview of errors and their types, as well as the technique of early detection of faults.
- The second chapter talks about everything related to machine learning and its related algorithms, types and applications.
- In the third chapter, which deals with the definition of the system on which the study was applied, the errors that can occur in it, as well as the relationship of early detection of faults and machine learning.
- The last chapter, which focused on the work done on the system data in the two control modes, also included the results obtained and the comparison between both modes.

Where this study concluded that the sliding mode provided a more robust control pattern against capacitor failures for the system, and on the other hand, the most accurate and effective classification model in identifying and detecting faults is the sliding mode. Therefore, this study made it possible to diagnose the photovoltaic power converter system in the event of a failure, in order to extend the life of the studied converter and ensure the continuity of solar-electric power supply.

Our future work aims to use deep learning algorithms for better data processing, and to use error prediction and fault tolerance control in order to make the system tolerant of errors so that it can run for as long as possible.

## References

- [1] Kerry THOUBBORON. *Advantages and disadvantages of renewable energy*. 22 November 2021. URL: <https://news.energysage.com/advantages-and-disadvantages-of-renewable-energy>.
- [2] Ehsan Jamshidpour, Philippe Poure, and Shahrokh Saadate. “Photovoltaic systems reliability improvement by real-time FPGA-based switch failure diagnosis and fault-tolerant DC–DC converter”. In: *IEEE Transactions on Industrial Electronics* 62.11 (2015), pp. 7247–7255.
- [3] Ramin Qaedi and Haidar Samet. “Evaluation of various fault detection methods in non-Isolated single switch DC-DC converters”. In: *2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*. IEEE. 2018, pp. 1–5.
- [4] Elyes Garoudja et al. “Statistical fault detection in photovoltaic systems”. In: *Solar Energy* 150 (2017), pp. 485–499.
- [5] Philippe Djondiné, Jean-Pierre Barbot, and Malek Ghanes. “Comparison of sliding mode and petri nets control for multicellular chopper”. In: *International Journal of Nonlinear Science* 25.2 (2018), pp. 67–75.
- [6] Leonardo Amet, Malek Ghanes, and Jean-Pierre Barbot. “Direct control based on sliding mode techniques for multicell serial chopper”. In: *Proceedings of the 2011 American Control Conference*. IEEE. 2011, pp. 751–756.
- [7] Boubakeur Rouabah, Houari Toubakh, and Moamar Sayed-Mouchaweh. “Fault tolerant control of multicellular converter used in shunt active power filter”. In: *Electric Power Systems Research* 188 (2020), p. 106533.
- [8] Boubakeur Rouabah et al. “More Efficient Wind Energy Conversion System Using Shunt Active Power Filter”. In: *Electric Power Components and Systems* 49.4-5 (2021), pp. 321–332.
- [9] Boubakeur Rouabah et al. “Adaptive and exact linearization control of multicellular power converter based on shunt active power filter”. In: *Journal of Control, Automation and Electrical Systems* 30.6 (2019), pp. 1019–1029.
- [10] Qiang Lu, Yuanzhang Sun, and Shengwei Mei. *Nonlinear control systems and power system dynamics*. Vol. 10. Springer Science & Business Media, 2001.

- [11] Carlos Frederico Meschini Almeida Ahmad Abubakar and Matheus Gemignani. *Review of Artificial Intelligence-Based Failure Detection and Diagnosis Methods for Solar Photovoltaic Systems*. 1 December 2021. URL: <https://doi.org/10.3390/machines9120328>.
- [12] Halim Alwi, Christopher Edwards, and Chee Pin Tan. *Fault detection and fault-tolerant control using sliding modes*. Springer, 2011.
- [13] K. Worden and J. M. Dulieu-Barton. *An Overview of Intelligent Fault Detection in Systems and Structures*. March 2004. URL: [https://www.researchgate.net/publication/245381970\\_An\\_Overview\\_of\\_Intelligent\\_Fault\\_Detection\\_in\\_Systems\\_and\\_Structures](https://www.researchgate.net/publication/245381970_An_Overview_of_Intelligent_Fault_Detection_in_Systems_and_Structures).
- [14] MOKHTARI Nouciba BAHLOUL Nousseiba. *Fault detection in photovoltaic power Converter*. Juin 2021.
- [15] Maria Jesús de la Fuente. “Fault Detection and Isolation: an overview”. In: *Dpto. Ingenieria de Sistema'sy Automatico Universiade de Valladolid* (2010).
- [16] Dubravko Miljković. “Fault detection methods: A literature survey”. In: *2011 Proceedings of the 34th international convention MIPRO*. IEEE. 2011, pp. 750–755.
- [17] Alireza Abbaspour et al. “A Survey on Active Fault-Tolerant Control Systems”. In: *Electronics* 9.9 (2020). URL: <https://www.mdpi.com/2079-9292/9/9/1513>.
- [18] Soumaya Yacout. “Fault detection and diagnosis for condition based maintenance using the logical analysis of data”. In: *The 40th International Conference on Computers & Industrial Engineering*. IEEE. 2010, pp. 1–6.
- [19] libe cmms. *Fault Detection And Diagnostics In Equipment Maintenance*. 28 october2021. URL: <https://limblecmms.com/blog/fault-detection-and-diagnostics/#fault-detection>.
- [20] D Divya, Bhasi Marath, and MB Santosh Kumar. “Review of fault detection techniques for predictive maintenance”. In: *Journal of Quality in Maintenance Engineering* ahead-of-print (2022).
- [21] Ndeye Gueye Lo, Jean-Marie Flaus, and Olivier Adrot. “Review of machine learning approaches in fault diagnosis applied to IoT systems”. In: *2019 International Conference on Control, Automation and Diagnosis (ICCAD)*. IEEE. 2019, pp. 1–6.
- [22] Ndeye Gueye Lo et al. “Gear and bearings fault detection using motor current signature analysis”. In: *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. IEEE. 2018, pp. 900–905.



- [23] Ani Madurkar. *Introduction to Artificial Intelligence, Machine Learning, and Deep Learning with Tensorflow*. 29 september 2021. URL: <https://towardsdatascience.com/introduction-to-artificial-intelligence-machine-learning-and-deep-learning-with-tensorflow-b5fa20477e89>.
- [24] Gavin Edwards. *Machine Learning / An Introduction*. 18 november 2018. URL: <https://towardsdatascience.com/machine-learning-an-introduction-23b84d51e6d0#d3ea>.
- [25] quantilus. *Why is Machine Learning Important and How Will It Impact Business*. 1 august 2020. URL: <https://quantilus.com/why-is-machine-learning-important-and-how-will-it-impact-business>.
- [26] Mr. Monteleone. *An Introduction on Artificial Intelligence and Machine Learning*. URL: <https://www.modernanalyst.com/Resources/Articles/tabid/115/ID/5468/An-Introduction-on-Artificial-Intelligence-and-Machine-Learning.aspx>.
- [27] Sotiris B Kotsiantis, Ioannis Zaharakis, P Pintelas, et al. “Supervised machine learning: A review of classification techniques”. In: *Emerging artificial intelligence applications in computer engineering* 160.1 (2007), pp. 3–24.
- [28] Sun-Chong Wang. “Artificial neural network”. In: *Interdisciplinary computing in java programming*. Springer, 2003, pp. 81–100.
- [29] Aurelia F. *Deep Neural Network : qu’est-ce qu’un réseau de neurones profond ?* 6 april 2021. URL: <https://datascientest.com/deep-neural-network#:~:text=Le%20Deep%20Neural%20Network%20ou,inspir%C3%A9s%20par%20le%20cerveau%20humain..>
- [30] Geoffrey I Webb, Eamonn Keogh, and Risto Miikkulainen. “Naive Bayes.” In: *Encyclopedia of machine learning* 15 (2010), pp. 713–714.
- [31] Xiaogang Su, Xin Yan, and Chih-Ling Tsai. “Linear regression”. In: *Wiley Interdisciplinary Reviews: Computational Statistics* 4.3 (2012), pp. 275–294.
- [32] Sai Chandra Nerella. *Linear Regression*. 4July 2021. URL: <https://saichandra1199.medium.com/linear-regression-1e279814e2bb>.
- [33] Vladimir Nasteski. “An overview of the supervised machine learning methods”. In: *Horizons. b* 4 (2017), pp. 51–62.
- [34] Armin Shmilovici. “Support vector machines”. In: *Data mining and knowledge discovery handbook*. Springer, 2009, pp. 231–247.
- [35] Alban T. *SVM, quoi, comment, pourquoi ?* 15 January 2021. URL: <https://datascientest.com/svm>.

- [36] Oliver Kramer. “K-nearest neighbors”. In: *Dimensionality reduction with unsupervised nearest neighbors*. Springer, 2013, pp. 13–23.
- [37] Sebastian Raschka. “STAT 479: Machine Learning Lecture Notes (2018)”. In: *URL [https://sebastianraschka.com/pdf/lecture-notes/stat479fs18/07\\_ensembles\\_notes.pdf](https://sebastianraschka.com/pdf/lecture-notes/stat479fs18/07_ensembles_notes.pdf)*. Citado na pág. viii 38 ().
- [38] Rupika Nimbalkar. *K-Nearest Neighbors Algorithms — Machine Learning*. 12 Jul 2021. URL: <https://medium.com/appengine-ai/k-nearest-neighbors-algorithms-machine-learning-756a7522dccb>.
- [39] Yanli Liu, Yourong Wang, and Jian Zhang. “New Machine Learning Algorithm: Random Forest”. In: *Information Computing and Applications*. Ed. by Baoxiang Liu, Maode Ma, and Jincui Chang. Springer Berlin Heidelberg, 2012.
- [40] Javatpoint. *Random Forest Algorithm*. URL: <https://www.javatpoint.com/machine-learning-random-forest-algorithm..>
- [41] Lior Rokach and Oded Maimon. “Clustering Methods”. In: *Data Mining and Knowledge Discovery Handbook*. Ed. by Oded Maimon and Lior Rokach. Springer US, 2005.
- [42] Tutorials Point. *Clustering Algorithms - Overview*. URL: [https://www.tutorialspoint.com/machine\\_learning\\_with\\_python/clustering\\_algorithms\\_overview.htm](https://www.tutorialspoint.com/machine_learning_with_python/clustering_algorithms_overview.htm).
- [43] Jonathan Leban. *Image recognition with Machine Learning on Python*. 21 May 2020. URL: <https://towardsdatascience.com/image-recognition-with-machine-learning-on-python-image-processing-3abe6b158e9a>.
- [44] Jonathan Leban. *Image recognition with Machine Learning on Python, Image processing*. 21 May 2020. URL: <https://towardsdatascience.com/image-recognition-with-machine-learning-on-python-image-processing-3abe6b158e9a>.
- [45] Prajyot Mane et al. “Smart personal assistant using machine learning”. In: *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*. IEEE. 2017, pp. 368–371.
- [46] Mobinius Editor. *How is AI advantageous as a Smart Assistant?* 6 July, 2020. URL: <https://www.mobinius.com/blogs/advantages-of-ai-as-smart-assistant/>.
- [47] Ali Bou Nassif et al. “Speech recognition using deep neural networks: A systematic review”. In: *IEEE access* 7 (2019), pp. 19143–19165.

- [48] Vineet Vashisht, Aditya Pandey, and Satya Yadav. “Speech Recognition using Machine Learning”. In: *IEIE Transactions on Smart Processing Computing* 10 (June 2021), pp. 233–239. DOI: 10.5573/IEIESPC.2021.10.3.233.
- [49] Camilla Nicolas. *The Complete Beginner’s Guide to Speech Recognition in Python*. 2021. URL: <https://morioh.com/p/57cbf29c7f87>.
- [50] Anand Handa, Ashu Sharma, and Sandeep K Shukla. “Machine learning in cybersecurity: A review”. In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 9.4 (2019), e1306.
- [51] Nirmal Budhathoki. *Data Science for Cyber Security- Opportunities vs. Challenges*. 24 Aug 2019. URL: <https://medium.com/@nirmal.budhathoki/data-science-for-cyber-security-opportunities-vs-challenges-d516a8a01200>.
- [52] Yingzi Xu et al. “AI customer service: Task complexity, problem-solving ability, and usage intention”. In: *Australasian Marketing Journal (AMJ)* 28.4 (2020), pp. 189–199.
- [53] Monica Maria. *5 Effective Ways to Improve Agent Productivity with AI*. 16 Jul 2020. URL: <https://freshdesk.com/customer-service/improve-agent-productivity-with-ai-blog/>.
- [54] Zhiqiang Gao et al. “An overview of PV system”. In: *2016 IEEE International Conference on Mechatronics and Automation*. IEEE. 2016, pp. 587–592.
- [55] al J.M.K.C. Donev. *Energy Education - Photovoltaic system*. 28 april 2020. URL: [https://energyeducation.ca/encyclopedia/Photovoltaic\\_system](https://energyeducation.ca/encyclopedia/Photovoltaic_system).
- [56] Thomas G Habetler and Ronald G Harley. “Power electronic converter and system control”. In: *Proceedings of the IEEE* 89.6 (2001), pp. 913–925.
- [57] Dino Green. *Advantages and disadvantages of Solar Photovoltaic – Quick Pros and Cons of Solar PV*. 19 december 2012. URL: <https://www.renewableenergyworld.com/storage/advantages-and-disadvantages-of-solar-photovoltaic-quick-pros-and-cons-of-solar-pv/#gref>.
- [58] Jianping Ma and Jin Jiang. “Applications of fault detection and diagnosis methods in nuclear power plants: A review”. In: *Progress in nuclear energy* 53.3 (2011), pp. 255–266.
- [59] Jie Chen et al. “Fault detection of aircraft control system based on negative selection algorithm”. In: *International Journal of Aerospace Engineering* 2020 (2020).
- [60] Rolf Isermann. “Model-based fault-detection and diagnosis—status and applications”. In: *Annual Reviews in control* 29.1 (2005), pp. 71–85.

- [61] Marie Postel. “Introduction au logiciel Matlab”. In: *en ligne*. Disponible sur: <https://www.ljll.math.upmc.fr/~postel/matlab/>(consulté le 27/02/2020) (2004).
- [62] Jim Sizemore and John Paul Mueller. *MATLAB for Dummies*. John Wiley & Sons, 2014.
- [63] Erfan. *Introducing MATLAB, its advantages and disadvantages*. 31 may 2021. URL: <https://ded9.com/introducing-matlab-its-advantages-and-disadvantages/>.
- [64] Xiang Wan et al. “Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range”. In: *BMC medical research methodology* 14.1 (2014), pp. 1–13.