N° Order........../DGM/FT/UMBB/2022

People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research

University M'hamed BOUGARA – Boumerdes

Faculty of Technology

Mechanical Engineering Department

Final Year Project Report Presented in Partial Fulfilment of the Requirements for the Degree of

MASTER

In **Electromechanical Engineering**

Option: **Industrial Maintenance**

Titled

Improving condition-based maintenance of a naval propulsion plants using Ensemble Learning

Presented by

- **Mr. CHEKHCHOUKH Djaafar**

Supervisor

- **Dr. RAHMOUNE Chemseddine**

Promotion 2021-2022

Thanks

First of all, I thank Allah for giving me faith, courage and confidence to be able to continue my studies and reach this level, Al hamdoulillah.

I also thank my family for the sacrifices they made in order for me to complete my studies.

I would like to express my gratitude and my heartfelt thanks to my promoter, Dr. RAHMOUNE Chemseddin who proposed and accepted to direct this project , for his help, his orientations and advice throughout the preparation of this work, with his good humor, his encouragement, and his modesty.

My thanks will also go to each doctor at the University of Boumerdes who responded favorably to all my requests

I also thank the jury members who took their time to review and judge this work.

And finally, a big thank you to all those who contributed directly or indirectly to the success of this project.

Dedications

I *dedicate this modest work to:*

M*y dear father* ***ALI****, who suffered to get me to this level, and who is looking forward to my return with my degree.*

M*y elegant mother, who stayed up and worked with me in every moment of my study and tired for this moment.*

T*o my dear brothers* **ahmed , omar , abderrahim , Mohamed**

T*o my aunts, and my grandmothers.*

T*o all my family, near or far.*

T*o all my college friends.*

T*o all my friends from Boudouaou*

T*o all of my co-workers at Diamond Brosse company*

T*o all the employees of the University of Boumerdes*

Chekhchoukh Djaafar

Table of Contents

Chapter III: Artificial Intelligence and machine learning, Generalities

Chapter IV: Ensemble Learning (Boosting and Stacking) for classification

List of figures

List of tables

ملخص

بسبب التقدم التكنولوجي الرهيب في العقود الأخيرة في مجال الدفاع عن الدول وسياسات الحرب، أصبحت الفرقاطات عاملا هاما في القوات البحرية لكل دول العالم.

يعد توافر وموثوقية محطات الدفع البحري لهذه الفرقاطات من العناصر اللازمة التي يجب التعامل معها لان تكاليف الصيانة تمثل شريحة كبيرة من إجمالي نفقات التشغيل، والتي تؤثر بشكل مباشر وغير مباشر على فعالية هذه السفن أثناء أداء مهامها.

إعتمادا الإستراتجية المتبناة، يمكن أن يختلف تأثير الصيانة على النفقات الإجمالية بشكل ملحوظ.

في هذا العمل، نأخذ في الإعتبار تحسين الصيانة الوقائية المستندة إلى الضروف لفرقاطة مجهزة بمحطة الدفع CODLAG
حيث نستخدم نهج تعليم الآلة لإظهار فعالية أساليب التعليم الجماعي المقترحة للقيام بمعايرتها في تطبيق بحري واقعي.

Abstract

Because of the terrible technological and scientific progress in recent decades in the field of war defense of states and war politics, frigates have become an integral part of the navies of most countries in the world.

Availability and reliability of naval propulsion plants of these Frigates are key elements to cope with because, maintenance costs represent a large slice of total operational expenses, which directly or indirectly affects the performance of ships while fulfilling their missions.
Depending on the adopted strategy, the impact of maintenance on overall expenses can remarkably vary.

In this work we take into consideration an improving condition-based maintenance of a Frigate equipped with a CODLAG naval propulsion plant where we use a machine learning approach to show the effectiveness of the proposed Ensemble Learning methods and to be benchmark them in a realistic maritime application.

Résumé

En raison des terribles progrès technologiques et scientifiques des dernières décennies dans le domaine de la défense militaire des États et de la politique de guerre, les frégates sont devenues partie intégrante des marines de la plupart des pays du monde.

La disponibilité et la fiabilité des centrales de propulsion navale de ces frégates sont des éléments clés auxquels il faut faire face car les coûts de maintenance représentent une part importante des dépenses opérationnelles totales, ce qui affecte directement ou indirectement les performances des navires tout en remplissant leurs missions.

Selon la stratégie adoptée, l'impact de la maintenance sur les dépenses globales peut varier considérablement.

Dans ce travail, nous prenons en considération une amélioration de la maintenance conditionnelle d'une frégate équipée d'une centrale de propulsion navale CODLAG où nous utilisons une approche d'apprentissage automatique pour montrer l'efficacité des méthodes d'apprentissage d'ensemble proposées et les comparer dans une application maritime réal

General introduction

The French Association for Standardization AFNOR is the French organization that represents France at the International Organization for Standardization and the European Committee for Standardization, defines that maintenance includes all actions necessary to retain a system or an item in, or restoring it to, a state in which it can per- form its required functions.

The most common way of inflecting such concepts in practice has been always deployed according to a ‘‘fix it when it breaks’’ approach. However, this has been becoming an unaffordable method since data gathering from the field is ever cheaper and costs related to a breakdown may overcome the asset value.

The smart technology used in factories and maintenance recently has lead to think and develop another new perspective that aims to produce a batter handling and a batter reaction to the faults that we face with any equipment. this would be possible by preventive measures and activities far away from the original thinking based on replacement and reform.

These measures can be distilled into conditional based maintenance, As the maintenance carried out in this category limits us the interruptions and the downtime of the equipments.

The core and essence of this application consists in the adoption of a proactive approach to maintenance in place of more conventional reactive interventions. For this purpose, the use of machine learning and artificial intelligence is necessary.

Chapter I: Frigates and the CODLAG naval propulsion plants

I.1 Introduction

A frigate is synonymous with a medium sized warship that was widely used in the eighteenth and nineteenth centuries. Generally used as an escort vessel for large war and merchant ships, frigates are important to navies all around the world.

The earliest frigates were considered ships that could not stand in battle and hence were used for carrying goods or helping other warships. However, as time passed, the sizes of frigates increased, and they were thus used for patrolling or escorting bigger vessels. The early frigates usually had three masts and were equipped with light armaments.

In this chapter, we will talk about frigates and the CODLAG naval propulsion system and how does it work.

I .2 Frigates

The frigates are vessels designed to achieve the strategic objectives of the National Defense of any nation in the 21st century antisubmarine and especially antiaircraft capabilities which make this vessel the essential platform to guarantee the defense and control of all areas under the sovereignty of a country.

Additionally, they are an indispensable tool in a Nation's Foreign Policy, acting both in conflict operations and in peace and humanitarian missions, as well as in the performance of resolutions. Its high capacities and the flexibility to operate both in coastal areas and on the high seas allow it to act as command and vessel control in operations with allied fleets, providing the required coverage. **[1]**

Figure 1: Modern frigate

Chapter I: Frigates and the CODLAG naval propulsion plants

I.2.1 The history of Frigates

Modern frigates are related to earlier frigates only by name. The term "frigate" was readopted during the Second World War by the British Royal Navy to describe an anti-submarine escort vessel that was larger than a corvette, while smaller than a destroyer. Equal in size and capability to the American destroyer escort, frigates are usually less expensive to build and maintain. **[2]** Anti-submarine escorts had previously been classified as sloops by the Royal Navy, and the Black Swan-class sloops of 1939–1945 were as large as the new types of frigate, and more heavily armed..

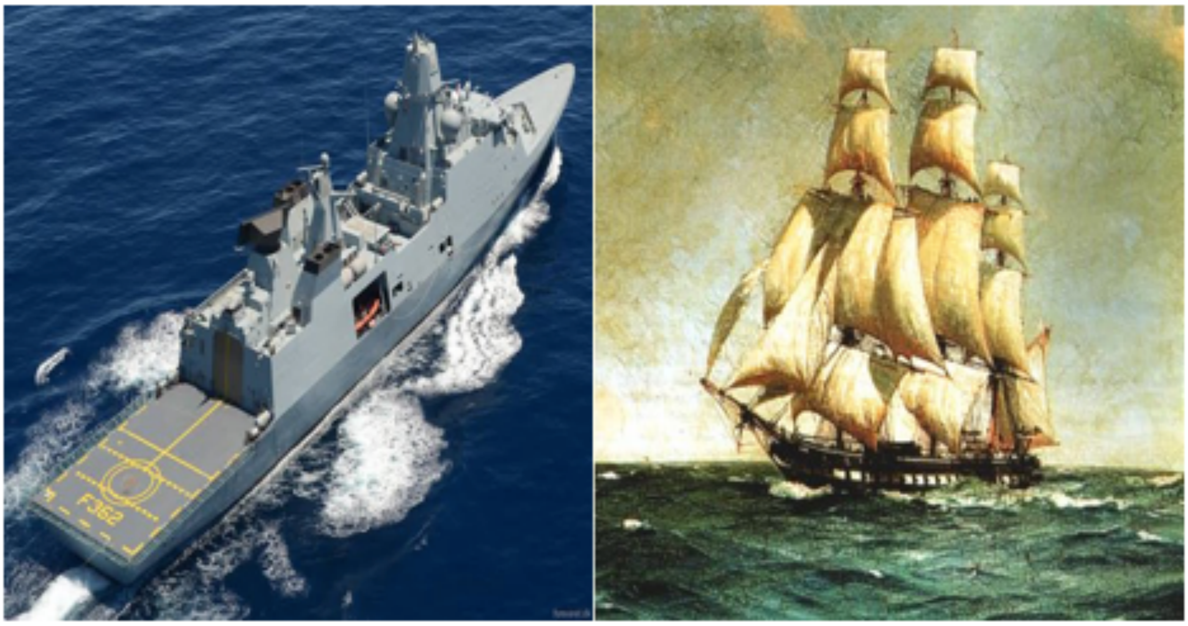


Figure 2:sailing frigate and Modern frigate

The frigate was introduced to remedy some of the shortcomings inherent in the flower-class corvette design: limited armament, a hull form not suited to open-ocean work, a single shaft which limited speed and maneuverability, and a lack of range. The frigate was designed and built to the same mercantile construction standards (scantlings) as the corvette, allowing manufacture by yards unused to warship construction. The first frigates of the River class (1941) were essentially two sets of corvette machinery in one larger hull, armed with the latest Hedgehog anti-submarine weapon.

The frigate possessed less offensive firepower and speed than a destroyer, but such qualities were not required for anti-submarine warfare. Submarines were slow while submerged, and ASDIC sets did not operate effectively at speeds of over 20 knots (23 mph; 37 km/h). Rather, the frigate was an austere and weatherly vessel suitable for mass-construction and fitted with the latest innovations in anti-submarine warfare. As the frigate was intended purely for convoy duties, and not to deploy with the fleet, it had limited range and speed.

It was not until the Royal Navy's Bay class of 1944 that a British design classified as a "frigate" was produced for fleet use, although it still suffered from limited speed. These anti-aircraft frigates, built on incomplete Loch-class frigate hulls, were similar to the United States Navy's destroyer escorts (DE), although the latter had greater speed and offensive armament to better suit them to fleet deployments. The destroyer escort concept came from design studies by the General Board of the United States Navy in 1940, as modified by requirements established by a British commission in 1941, **[3]** prior to the American entry into the war, for deep-water escorts. The American-built destroyer escorts serving in the British Royal Navy were rated as Captain-class frigates. The U.S. Navy's two Canadian-built Asheville-class and 96 British-influenced, American-built Tacoma-class frigates that were originally classified as "patrol gunboats" (PG) in the U.S. Navy. Navy but on 15 April 1943 were all reclassified as patrol frigates (PF).

Figure 3:The first frigate with new meaning of a frigate

I.2.2 Design and Construction

The eighteenth century frigates were all square rigged and had a couple of decks, each meant for a particular purpose. Generally, the upper deck used to carry all the guns and other weapons, and the lower deck was meant for crew accommodations. However, in some ships, the lower deck was used as the gun deck, and an additional deck below it served the purpose of crew accommodations. The early frigates were smaller in size, easy to maneuver, and good for assisting bigger ships during war time. Later on, advanced frigates were made with higher weapon carrying capability and more power along with special designs for less hogging and higher hydrodynamic efficiency. Live oak trees that are found abundantly in America were extensively used at that time to build these ships. In order to prevent hogging, the hull of the frigates was designed in such a way that the weight of the guns come directly on the keel. Diagonal riders, also made from live oak, were used on both the sides, making an angle of forty-five degrees. The riders, which had a width of about two feet and a thickness of about one foot, assisted in maintaining the hull form by preventing hogging, increasing flexibility, and reducing various stresses on the ship. The placing of riders was a multi-layer method, wherein the planks were kept not only on the two faces of the frigate's hull but also horizontally throughout the ribs for making a cross pattern. This particular pattern helped to resist any kind of stresses and thus made the structure stronger and more reliable. **[4]**

I.2.3 The Role of Frigates

In the nineteenth century, frigates with steam propulsion systems were made, which drastically increased their role in the navies around the world. However, as their applications increased, the term frigate was substituted by the terms battleship and destroyer. During the Second World War, some of the most advanced frigates were used in war applications such as anti-submarine vessels, minesweepers, minelayers, and merchant ships escorts.

Some of the modern frigates are also used as anti-aircraft vessels and have diesel, steam, and even nuclear propulsion systems. They are also equipped with various forms of offensive or defensive missiles and are called Guided Missile Frigates. A few of the most recent frigates are equipped with Stealth technology, which allows them to go undetected in an enemy's radar system. In many navies, frigates are also known by the term destroyer. **[5]**

Figure 4:A missile carrier frigate

Chapter I: Frigates and the CODLAG naval propulsion plants

Some of the modern frigates have diesel, steam, and even nuclear propulsion systems. They are also equipped with various forms of offensive or defensive missiles and are called Guided Missile Frigates. A few of the most recent frigates are equipped with Stealth technology, which allows them to go undetected in an enemy's radar system. In many navies, frigates are also known by the term destroyer.

I.3 The CODLAG propulsion system

Combined diesel-electric and gas (CODLAG) is a modification of the combined diesel and gas propulsion system for ships. A variant, called the combined diesel-electric or gas (CODLOG) system, contains the same basic elements but will not allow simultaneous use of the alternative drive sources.

In this work , we are interested in the CODLAG propulsion system. Naval forces in many countries use the CODLAG propulsion system in their frigate or destroyer.

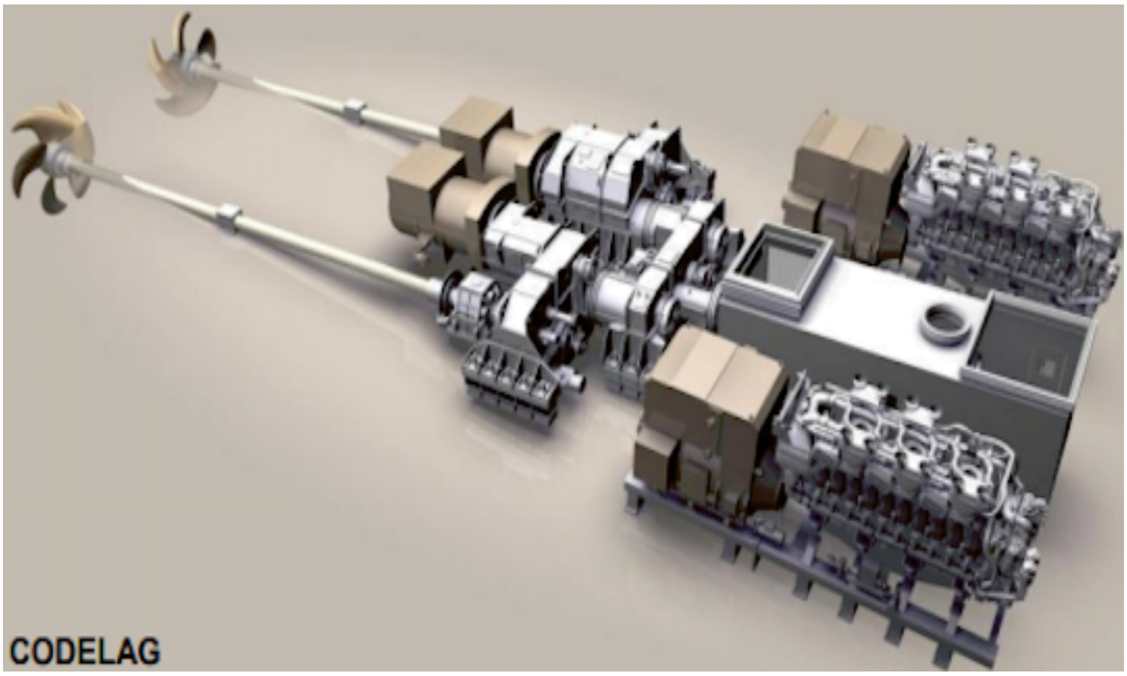


Figure 5:three-dimensional design of the CODLAG

I.3.1 list of some CODLAG ships

_ Type 23 frigate (Royal Navy)

_ FREMM multipurpose frigate (Italian Navy)

_ F125-class frigate (German Navy)

_ Pohjanmaa-class corvette (Finnish Navy) _ GTS Finnjet (Finnish Cruise) **[6]**

I.3.2 How does CODLAG propulsion system works

CODLAG is a hybrid propulsion system. CODLAG uses an electric motor, powered by a diesel generator for cruising or to run the ship silently.

A gas turbine is used when high speed is required, for example during an emergency situation faced by a naval force.

GT has three major components: air compressor, combustion chamber, and turbine. The compressor absorbs air from the atmosphere and increases the pressure, The pressurized air then enters into the combustion chamber. Inside the combustion chamber there are fuel injectors and high intensity spark igniters.

Fuel injectors inject a steady stream of fuel into the chamber.

The injected fuel is mixed with the air entered from the compressor. The injected fuel-air mixture is ignited by the igniters present in the combustion chamber. In this process, very hot and highly pressurized gas is produced.

The produced gas is then expanded through an exhaust nozzle. The resulting hot , high velocity exhaust gas from the combustion chamber drives the turbine. The GT produces mechanical energy to rotate the shaft. This shaft having enormous torque turns the propeller or generator **[7, 8] .**

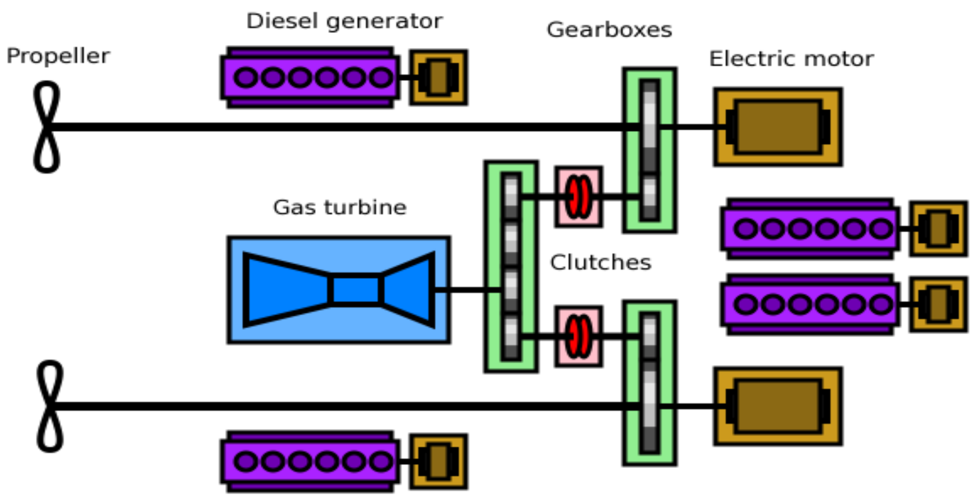


Figure 6:diagram represent the CODLAG

I.4 Gas turbine

A gas turbine is an internal combustion engine, designed to accelerate the flow of gas, which is used to produce a reaction force to propel an object or to produce mechanical energy to exploit it for mechanisms such as propulsion of ships exactly as it was mentioned previously with the CODLAG propulsion system.

Figure 7:Gas Turbine (GT)

I.4.1 The main components of a gas turbine

I.4.1.1 Air compressor

No matter what a gas turbine type you consider, all gas turbines have a compressor that increases input air pressure before entering the combustion engine. The output of the compressor is critical for the overall engine efficiency.

Gas turbine compressors provide the compression component of the thermodynamic cycle of the gas turbine engine. Generally, three specific types of gas turbine compressors are available: an axial compressor, a centrifugal compressor, and a compressor with a mixed flow. **[7] .**

I.4.1.2 Combustion chamber

combustion chamber or combustor is an important component of gas turbine engines, in which the fuel molecules go through the exothermic chemical reaction of combustion, or simply called burning, producing an enormous amount of heat.

The chemical reaction leaves a mixture of different gases that are not only extremely hot because of the nature of the process they went through, but also have a very high pressure as their parent gases came from the compressor. This makes them begging to run out of the chamber and expand in the turbine

I.4.1.3 Turbine

Turbine is where the mechanical power generated. This mechanical power is generated when the hot gases exiting the chamber pass through the turbine blades forcing them to move about their axis of rotation.

I.4.1.4 Shaft

There are two types the single-shaft or two-shaft configuration may be used in gas turbines. The single-shaft structure consists of a shaft which connects a rotating part of the air compressor, gas turbine maker, and power turbine. This design is best suited for applications at constant speeds, such as electrical generators for constant frequencies.

The two-shaft configuration involves the air compressor, the gas supplier, and the power turbine on the second independent shaft. This design offers the speed in a flexible way needed to effectively cover a broader map of the powered system. The gas producer can, therefore, operates at the necessary speed to build the horsepower needed by the powered equipment including centrifugal compressors or pumps. **[8] .**

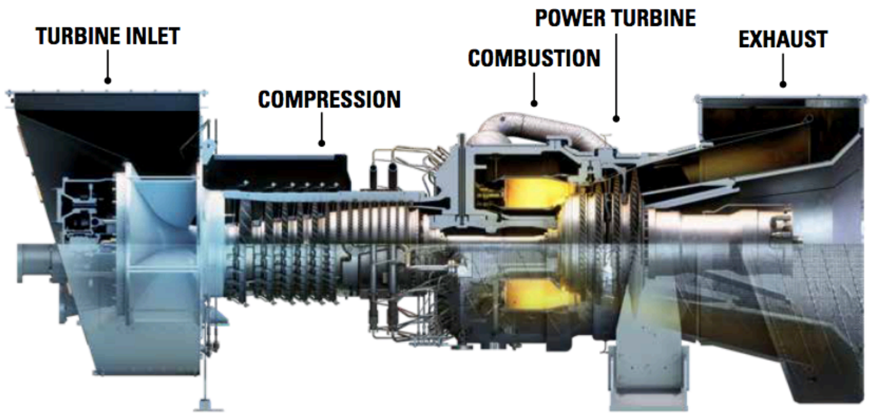


Figure 8:Gas Turbine (GT) main components

I.4.2 Gas Turbine working Principle

Simply a gas turbine, also called a combustion turbine, is a type of continuous and internal combustion engine.

Any gas turbine operates with intake, compression, expansion, and exhaust cycle. As a fundamental of the gas turbine working principle, in each gas turbine type, the compressor first compresses the air and this air is then driven through the combustion chambr where Fuel is continuously burned for high-temperature and high-pressure gas processing. What a gas turbine is doing is expanding the gas generated by the combustor into the turbine and thus generating the rotary energy that is used by the compressor on the preceding stage. There is an output shaft for the remaining energy

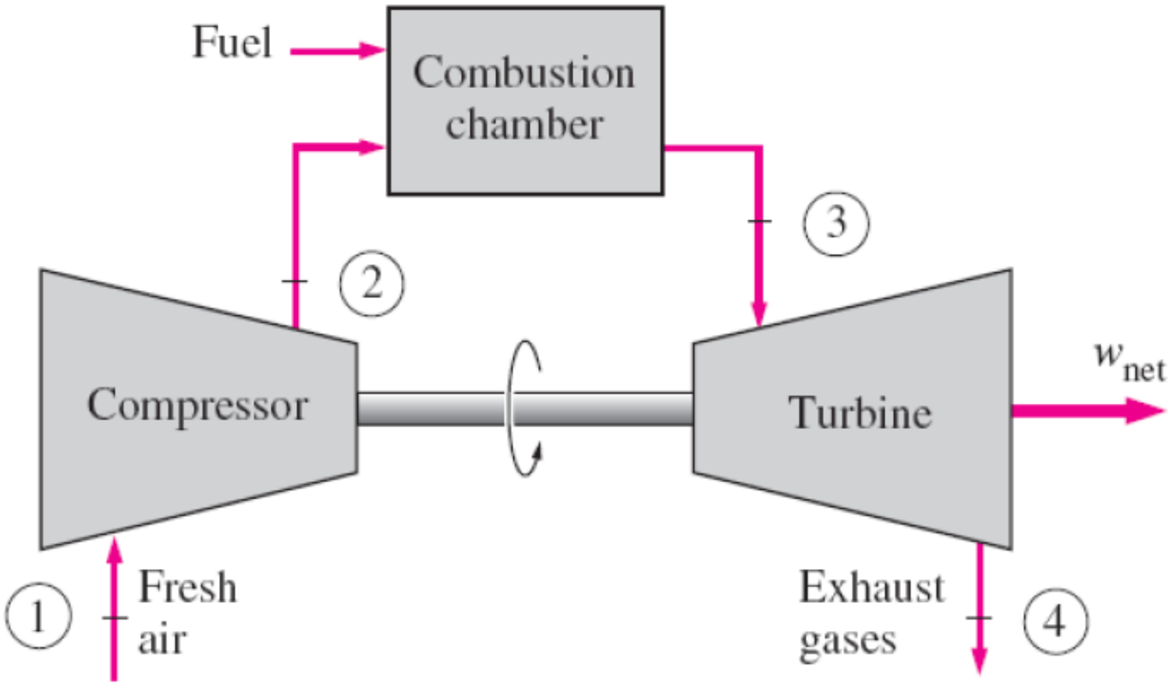


Figure 9:schematic diagram of Gas Turbine (GT) working principle

I.4.3 The main gas turbine cycles

I.4.3.1 ideal cycle (Joule or Baritone cycle)

In its simplest form, the gas turbine works according to the called Joule cycle (or of Baritone) including:

¸Adiabatic compression that consumes mechanical energy;

¸Isobaric combustion;

¸An adiabatic expansion down to ambient pressure which produces energy

Mechanical.

The Joule cycle, as it is well represented in the next figure , comprises two Isentropic processes (adiabatic and reversible) and an isobaric process.

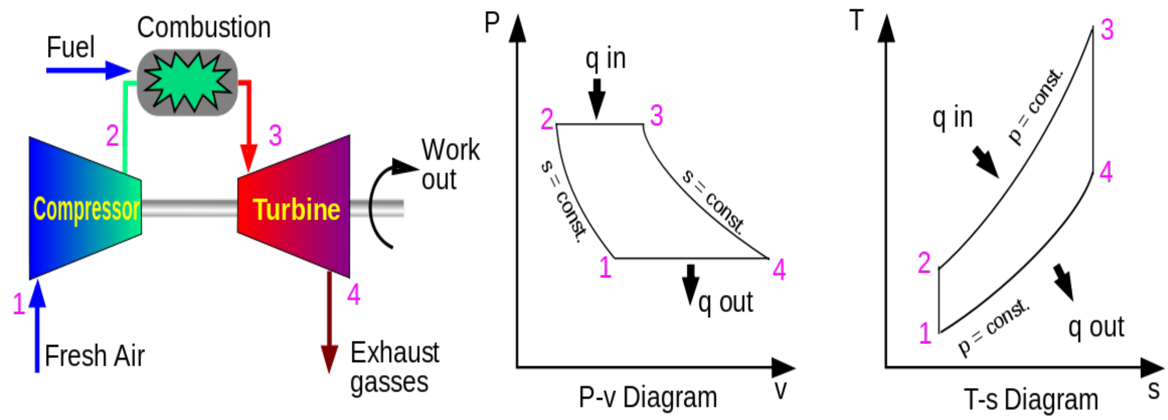


Figure 10:diagram T _S and P_V of baritone cycle of a Gas Turbine (GT)

• 1 to 2: isentropic compression, (Compressor).

• 2 to 3: isobaric combustion, (Combustion chamber).

• 3 to 4: isentropic expansion, (Turbine).

• 4 to 1: isobaric cooling, (Exhaust).

I.4.2. The Real cycle

The real cycle of the simple gas turbine deviates from the ideal cycle both by the irreversibility in the compressor and in the turbine only by the pressure drop in the chamber of combustion and flow channels.

This cycle includes:

¸ Adiabatic compression with entropy increase from (1) to (2);

¸ Combustion with a pressure drop due to pressure drops from (2) to (3);

¸ An adiabatic expansion down to atmospheric pressure with an increase entropy from (3) to (4).

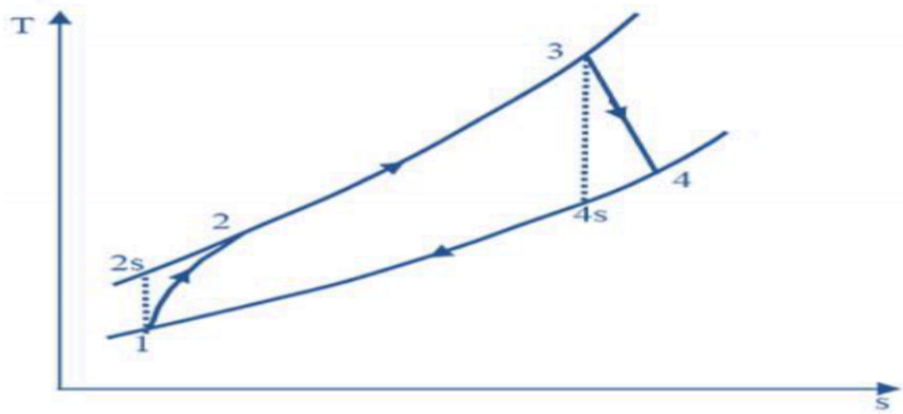


Figure 11:diagram T_S of the Real cycle of a Gas Turbine (GT)

The nest figure shows the different between the ideal baritone cycle and the real cycle in the T-S diagram

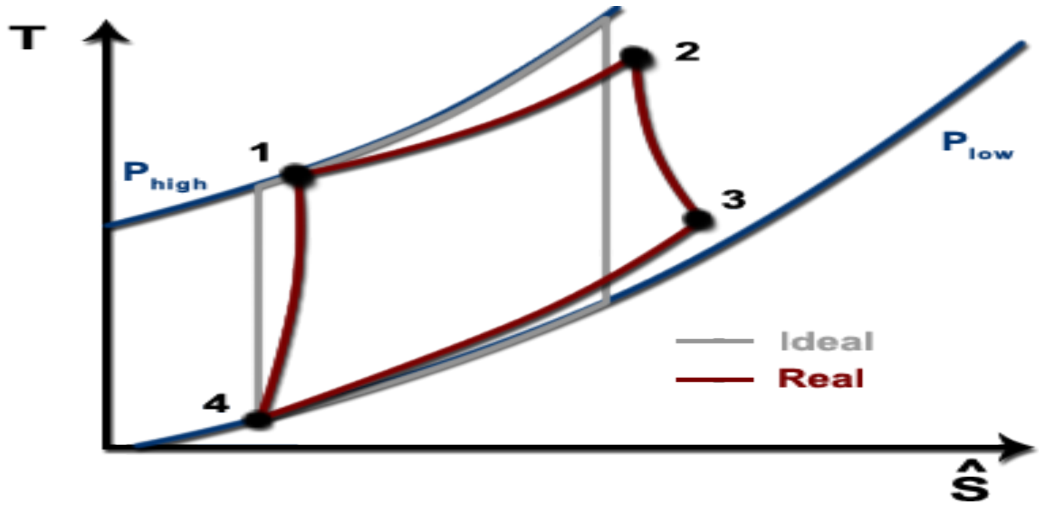


Figure 12:A T_S diagram of Ideal and Real cycle of a Gas Turbine (GT)

I.5 Conclusion

In this chapter we gave generalities on frigates and their history and design. We also discussed CODLAG propulsion systems and its principle of working where we focused on the gas turbine components its cycles.

In the next chapter we will talk little bit about maintenance.

Chapter II: Basic principals about Maintenance and Diagnosis

II.1 Introduction

The word Maintenance appeared in the industrial vocabulary in the 1950s, on the other hand, the concepts of maintenance, such as we know today, in fact go back to the highest antiquity, from the development of man from the first machines.

II.2 Maintenance:

Maintenance is the set of all administrative and management technical actions consists of maintaining or restoring a good in a specified state in which it can perform its required function but it must be specified that this is always done under given conditions and within a specific time interval. **[9]**

II.3 types of maintenance

It's about taking certain steps to help the system work at its best. In short, corrective and curative maintenance solves problems while preventive maintenance avoids them in the first place. This is the major difference between these major types of maintenance.

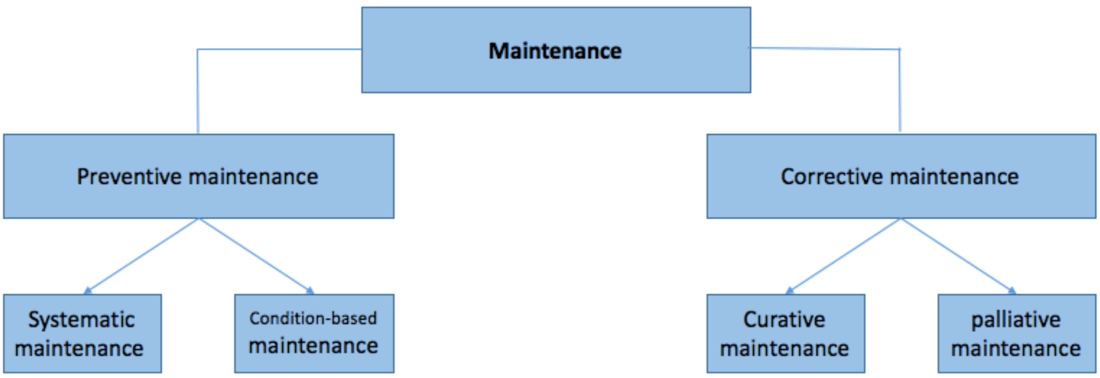


Figure 13:types of maintenance

II.3.1 Preventive maintenance

This type of maintenance is carried out according to predetermined criteria, the objective of which is to reduce the probability of failure of equipment in use, to reduce downtime in the event of overhaul or breakdown and to eliminate the causes serious accidents. It is based on compliance checks, periodic monitoring to identify anomalies and perform simple adjustments without specific tools or production tool shutdown.

Any maintenance intervention includes a financial impact. The cost analysis must highlight a gain in relation to the failures it avoids. It is also at this time that the company considers whether to decide to move on to the stage of corrective maintenance, which is more expensive.

II.3.1.1 Systematic maintenance

It follows a schedule generally established according to time but also according to the quantity of products manufactured, the distance traveled when it comes to transporting goods, or the number of cycles carried out, etc. This frequency of intervention is determined from commissioning or after a complete or partial overhaul.

This maintenance method requires knowledge of the behavior of the equipment, the modes of degradation and the average time of good operation between two failures. It applies to the maintenance of equipment subject to regulations such as lifting devices, fire extinguishers, pressure equipment, conveyors, elevators, etc.; for equipment whose failure risks causing serious accidents (aircraft, trains), equipment with a high failure cost. This is the case of the elements of an automated production line, such as industrial pumps, in the chemical, food, metallurgy or boiler making sectors, etc. Manufacturers also use this method for equipment whose operating costs become abnormally high during their service life: excessive energy consumption, lighting by used lamps, improper ignition and carburetion, etc.

II.3.1.2 Condition-based maintenance

It is, as its name suggests, conditioned on a type of predetermined event: a diagnosis, information from a sensor, a measurement of wear, etc.). This predictive maintenance is therefore maintenance dependent on experience and involving information collected in real time. In short, it aims to identify weak points. These regular and rigorous checks make it possible to highlight the operating condition of the equipment. Depending on the case, it is desirable to put the equipment under surveillance and then to decide on an intervention when a certain threshold is reached.

The computer equipment needed to perform conditional preventive maintenance and analyze the proper functioning of production tools is often expensive, but pays for itself quickly. **[10]**

II.3.2 Corrective maintenance

Corrective maintenance operations take place once the failure is identified. In short, it is a matter of troubleshooting since corrective maintenance is carried out after detection of a breakdown and intended to restore an item to a state in which it can perform a required function.

II.3.2.1 Curative maintenance

This form of maintenance is applied when a malfunction is detected. Its purpose is to repair this anomaly. Unlike the palliative which acts in an emergency, this is a real long-term restoration, in a definitive way. Curative maintenance is a priority repair because production stoppages cause heavy losses for companies.

This intervention improves performance and production conditions. It allows the repair of damaged equipment on site or in the workshop or the installation of a fleet of new machines. This corrective maintenance also intervenes in the compliance of the production tool after the passage of an inspection body which ensures compliance with the legislation in force.

II.3.2.2 palliative maintenance

Its objective is to restore an asset to a specified state: this does not necessarily mean its initial state. For example, if a leak occurs in an industrial pipe or a storage tank, the palliative maintenance operation must ensure that these tools regain their required function, without however restoring their original appearance. It is an action intended to allow a good to function temporarily.

Palliative maintenance is not a planned action and is therefore not part of a maintenance policy. These temporary troubleshooting actions are to be distinguished from the repair that characterizes curative maintenance. Palliative maintenance can be dangerous. The technician must therefore, in agreement with the management of the company, make the balance between the danger that the stoppage of production could represent and that inherent in any palliative maintenance. **[10]**

II.4 The different levels of maintenance

Standard X 60-010 distinguishes 5 degrees of maintenance, classified in an increasing way, according to the complexity of the interventions to be carried out.

II.4.1 First level maintenance

The first level maintenance actions are simple actions necessary for the operation and carried out on elements that are easily accessible, in complete safety, using equipment support integrated into the property.

These are, for example, the adjustments and checks or inspections necessary for operation, the basic preventive maintenance operations, the replacement of consumable items or accessories (fuses, bulbs, etc.).

• This type of operation can be carried out by the operator of the asset with support equipment incorporated into the property and using the instructions for use.

II.4.2 Second level maintenance

• The second level of maintenance concerns actions that require simple procedures and/or support equipment (integrated or external) that is easy to use and implement.

• These are, for example, performance checks, certain adjustments, repairs by exchange standard of sub-assemblies whose replacement is easy.

• This type of maintenance can be carried out by qualified personnel with detailed procedures and the support equipment defined in the maintenance instructions.

• are thus concerned by this level the operations of replacement of parts not entailing overall dismantling of the equipment. It is therefore a work relating to isolated elements or results verification operations such as performance control of delivered equipment.

II.4.3 Third level maintenance

• The third level concerns operations that require complex procedures and/or complex support, use or implementation equipment.

• These are, for example, general adjustments, delicate systematic maintenance operations, repairs by exchange of components.

• These operations require a global approach to the operation of the equipment, i.e. the taking into account several elements, their interactions and their coherence.

II.4.4 Fourth level maintenance

• The 4th level concerns operations whose procedures involve the mastery of a technology and/or the implementation of specialized support equipment.

• These are, for example, specialized repairs, verifications of measuring devices.

II.4.5 Fifth level Maintenance

• Renovation or reconstruction activities whose procedures involve know-how use of specific techniques or technologies, processes and/or equipment for industrial supports.

Chapter II: Basic principals about Maintenance and Diagnosis

II.5 The main maintenance concepts

_ Reliability

Ability of equipment to perform a required function under given conditions, for a given time interval. It is generally assumed that the item is in a condition to perform the required function at the start of its life the given time interval.

The concept of reliability is often translated in practice as the ability of an entity to have a low failure frequency. **[9]**

_ Maintainability

Ability of equipment to be maintained in or restored to a state in which it can perform a required function, when maintenance is performed under given conditions, with prescribed procedures and means. **[9]**

_ Availability

Ability of an entity to be able to perform a required function under given conditions, at a given moment or during a given time interval, assuming that the necessary external means are provided. **[9]**

_ Breakdown

Inability of an entity to perform a required function

_ Failure

Impairment or cessation of an entity's ability to perform a required function

_ Repair

Repair consists of restoring, in a lasting way, with the aim of eliminating or reducing the consequences of obsolescence, wear or disorder, of equipment that no longer performs in acceptable conditions the function which is his.

_ Troubleshooting

Action on broken down equipment, with a view to restoring it to working order, at least temporarily. Depending on the objective, troubleshooting can accommodate temporary results and “non-standard” performance conditions and; in this case, will be followed by repair.

II.6 Diagnostic

II.6.1 Definition of Diagnostic

Analysis of a set of factors or symptoms, aimed at establishing the state of an element or the causes of any observed disorder, in order to choose the measures to be taken to remedy it.

Examination to assess the state of wear of a component, in order to determine the operations maintenance to be performed, or the remaining life. **[9]**

II.6.2 Fault detection procedure

we find in the diagnosis two essential tasks are the observation of the symptoms of failure as well as the identification of the cause of failure using a logical reasoning founded on observations acquired on the system, starting from the information available on its operating condition, the objective is to detect, locate and identify failures that may affect its operating safety.

_ Detection

which consists of using a set of measures (fault indicators) that generate symptoms. We also speak of attributes when referring to the use of classification methods for detection.

_ Localization

involves further determining the failing components or probable cause of failure.

_ fault identification: aims to determine the type (class, size, ...) of a defect (classification).

_ Decision-making: must make it possible to generate, possibly under the control of a human operator, the corrective actions necessary for a return to normal operation of the installation

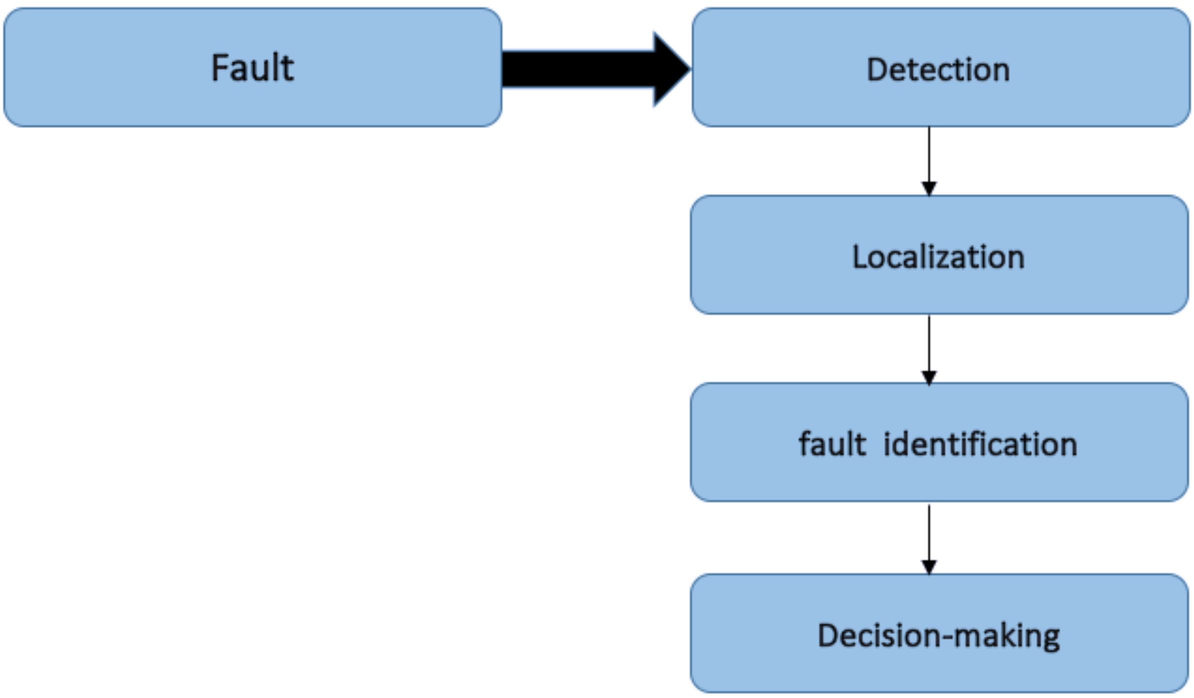


Figure 14:Fault detection procedure

II.6.3 Diagnostic methods

Overall, two main families can be distinguished in the diagnostic methods:

- Methods based on modeling systems or signals.
- Methods based on artificial intelligence

II.6.3.1 Model-based methods

Model-based diagnostic methods can be divided into two parts: methods based on 'mathematical model of the system' and methods based on 'model of the signal.

On the one hand, the methods based on mathematical model aim to represent the dynamics of the system by mathematical equations which are then exploited for the diagnosis the modeling of the dynamic systems often formulated by nonlinear relations which complicate the estimation of the quantities of 'interest. On the other hand, signal model-based methods aim to model the measured signal and exploit its parameters (phase, amplitude, average, etc.) to make the diagnosis.

The principle of diagnosis based on a mathematical model is to compare the current behavior of the system with its expected behavior (healthy state). This comparison generates residual quantities which are used to make the diagnostic decision.

II.6.3.2 Methods based on artificial intelligence

Methods based on artificial intelligence work on the same principle as those based on quantitative models, but the relationship between the descriptors is expressed by qualitative functions. These methods are not based on writing physical laws

Diagnosis with artificial intelligence is based on the use of machine learning algorithms which are subject to two steps:

-The learning stage: In this phase (training or learning), the objective is to find, from the set of experimental data, the characteristics of the behavior of the system which will make it possible to differentiate the states in which can be found the system. This concerns the design and construction of a classification system. It consists of two main phases: a learning phase (supervised or unsupervised) and the assignment of the classes obtained in important functional states.

Recognition phase:

During the second phase, the recognition of new data is carried out. This will identify the functional status of the procedure from online data from sensors and actuators or other information from process variables. The data is classified according to the classifier obtained in the previous phase **[11]**

II.6.4 Redundancy

II.6.4.1 Hardware redundancy

The basic principle of fault diagnosis is based on the notion of redundancy, which provides the system with several different pieces of information on the same variable. Tests will then make it possible to verify the consistency of this information. **[12]**

II.6.4.2 Analytical redundancy

A complement to physical redundancy consists in exploiting the constraints linking the different variables of the system. These constraints can often be expressed in the form of analytical relationships linking known variables (input/output or output/output relationships). These relations are called analytical redundancy relations.

II.7 Objectives of maintenance

Equipments are an important resource which is constantly used for adding value to products. So, it must be kept at the best operating condition. Otherwise, there will be excessive downtime and also interruption of production if it is used in a mass production line. Poor working of equipment
will lead to quality related problems. Hence, it is an absolute necessity to maintain the equipments in good operating conditions with economical cost. Hence, we need an integrated approach to minimize the cost of maintenance. In certain cases, the equipment will be obsolete over a period of time. If a firm wants to be in the same business competitively, it has to take decision on whether to replace the equipment or to retain the old equipment by taking the cost of maintenance and operation into account.

II.8 Conclusion

In this chapter we talked about maintenance types, levels, some maintenance concepts, diagnostic. And maintenance objectives

There are several diagnostic methods for example the analysis of ultrasonic measurements, and oil analysis which based on a modeling of systems or signals, there are new modern methods based on artificial intelligence and machine learning witch we have used in our project.

We will mention in the next chapter the concepts of artificial intelligence and machine learning

Chapter III: Artificial Intelligence and machine learning, Generalities

III.1. Introduction

By the mid-twentieth century, the world had witnessed a development of computing, and researchers began to consider the creation and development of devices that simulate human thought by simulating the neuro system, in conjunction with discoveries in neuroscience and the development of automated control. Artificial intelligence is a new concept that has been coined in the field of computer science. In the summer of 1956, Dartmouth presented the first artificial intelligence conference.

Artificial intelligence has progressed tremendously since then, and man has applied it to a variety of sectors, including medical diagnosis and problem diagnostics in industrial equipment. We define artificial intelligence and its applications in this chapter, as well as one of its branches (machine learning) that is employed in diagnostics.

III.2 Artificial intelligence

Artificial Intelligence would be defined as the set of theories and techniques implemented to produce machines capable of simulating intelligence. They would therefore be computers or machines with programs capable of performance similar to human intelligence, or even amplified by technology. These machines have the :

- Ability to reason
- Ability to process large amounts of data
-The ability to distinguish patterns that humans cannot detect.
- Ability to interact with humans.
- Ability to learn gradually.

Artificial intelligence needs two basic things: a large collection of data and the algorithms that process that data. **[13]**

III.3 Artificial intelligence applications

There are many areas of AI application, often with other computer disciplines.

▪ the transport sector (autonomous vehicles).

▪ Diagnostic systems (medical, industrial).

III.4 Machine learning

Machine learning is a field emerging from artificial intelligence concerned with understanding and replicating the power of human learning through artificial systems. It is a question of a very schematically of designing algorithms and methods allowing to extract relevant information from data, or to learn behaviors from examples.

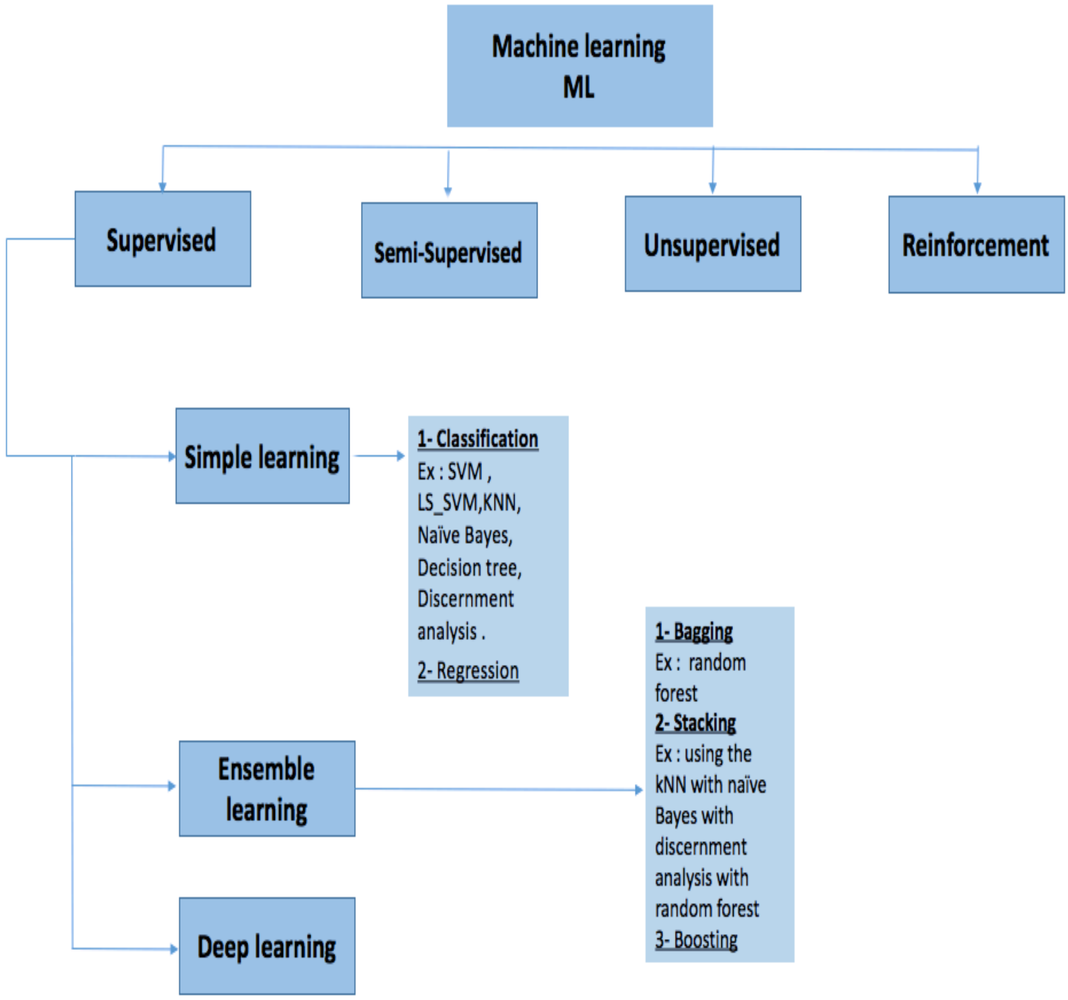


Figure 15:Machine learning's fields

Thus, the main goal of machine learning is to determine the relationship (pattern) between objects and their categories for prediction and knowledge discovery .

The data used by machine learning gives the modeling process are called (training) data, in English Training Data . **[14]**

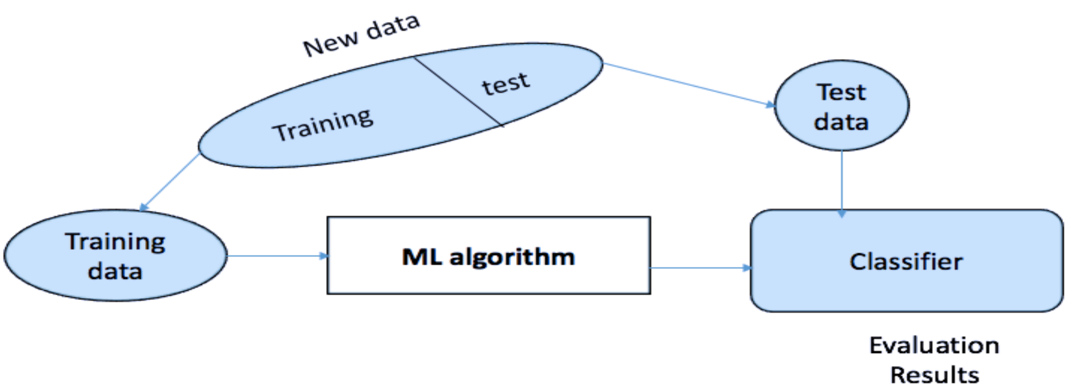


Figure 16:supervised Machine learning working principle

III.4.1 Types of Machine learning

1) - Supervises: we have a set of objects and for each object an associated target value; you have to learn a model capable of predicting the correct target value of a new object.

2) -unsupervised: we have a set of objects without any associated target value; you have to learn a model capable of extracting the regularities present within the objects to better visualize or understand the structure of all the data **[15]**

3)-By reinforcement: we have a set of sequences of decisions (political or strategic) in a dynamic environment, and for each action of each sequence a reward value (the reward value of the sequence is then the sum of the values of the rewards of the actions it implements); you have to learn a model capable of predicting the best decision to make given a state of the environment. **[16]**

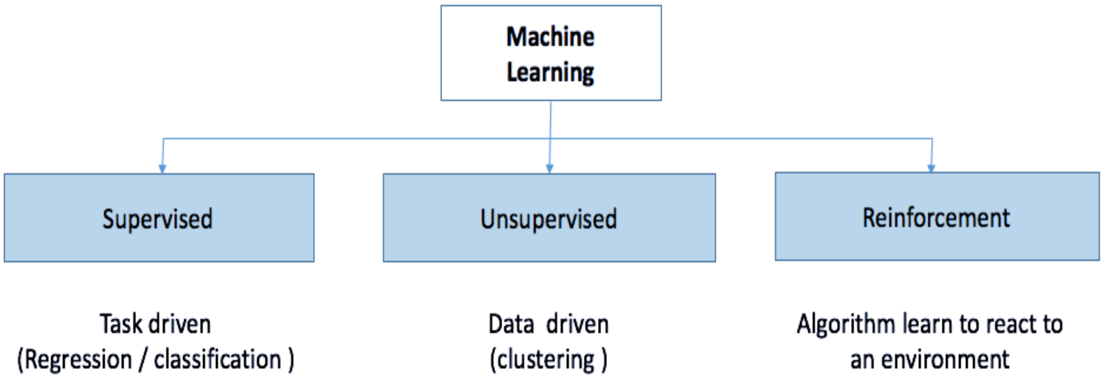


Figure 17:Types of Machine learning

III.4.1.1 Supervised learning

The algorithm is trained using a learning database containing examples of real cases processed and validated. The objective is to find correlations between the input data (explanatory variables) and the output data (variables to be predicted), to then infer the knowledge extracted on inputs with unknown outputs.

In supervised learning, a distinction is made between two types of tasks (regression, classification) .

III.4.1.1.1 Regression

In machine learning, the goal of regression is to predict an output value from the values of a set of input characteristics. For example, predicting the price of oil from its price in recent years, predicting the number of successful candidates in the exam from their scores throughout the semester , etc.

III.4.1.1.2 classification

Supervised classification is a widely used technique with different applications in real life. It makes it possible to generate classification rules (model) from a set of data classified a priori and from an adequate machine learning algorithm. These rules will be used to classify the new instances.

A classification is said to be binary, if the number of classes X is equal to 2. The classifier must predict one of the two classes for the new instances. A classification is said to be multi-class if the number of classes $X > 2$. **[17]**

III.4.1.1.3 Holdout

We use Holdout to split data set into a "train" and "test" set. The training set is what the model is trained on, and the test set is used to see how well that model performs on the data. A common division when using the holdout method is to use 70% of the data for training and the remaining 30% of the data for testing. The principle of this validation method is to partition the data into exactly two subsets of a specified ratio for training and validation. **[18]**

III.4.2 Supervised Classification methods

III.4.2.1 Support Vector Machines (SVM)

Vapnik proposed statistical learning theory based machine learning method which is known as Support vector machine (SVM) . SVM has considered as one of the highest prominent and convenient technique for solving problems related to classification of data and learning and prediction . Support vectors are the data points that lie closest to the decision surface . It executes the classification of data vectors by a hyper plane in immense dimensional space . Maximal margin classifier is the simplest or basic form of SVM that helps to determine the most simple classification problem of linear separable raining data with binary classification . The maximal margin classifier used to find the hyper plane with maximal margin in real world complications **[19]** .

The main advantage of SVM is its capability to deal with wide variety of classification problems includes high dimensional and not linearly separable problems. One of the major drawback of SVM that it requires number of key parameters to set correctly to attain excellent classification results **20]**,

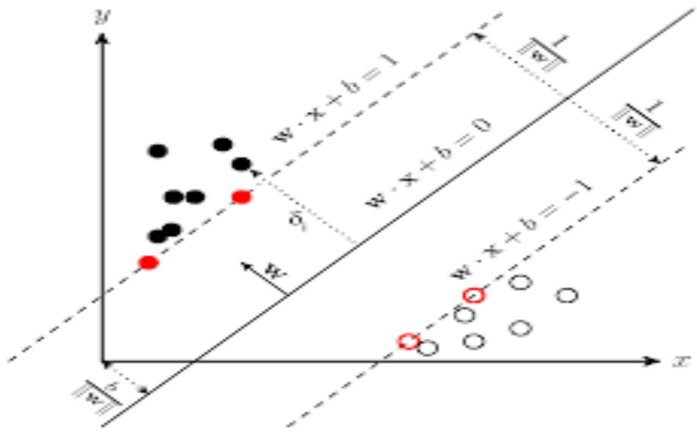


Figure 18:linear support vector machine

III.4.2.2 Least square-support vector machines (LS-SVM)

Support vector machines (SVM) have been widely used in classification and estimation of nonlinear functions. However, the major disadvantage of SVM is its greater computational load for optimization of programming constraints. This drawback has been overcome by least square-support vector machines (LS-SVM), which solves linear equations instead of a programming problem quadratic. **[21]**

III.4.2.3 The main difference between LS-SVM and SVM

the main difference between the two techniques is that the algorithm (LS-SVM) uses the least squares technique to classify the data This is done by calculating and adding the squares and choosing the least square to draw the line separating each class

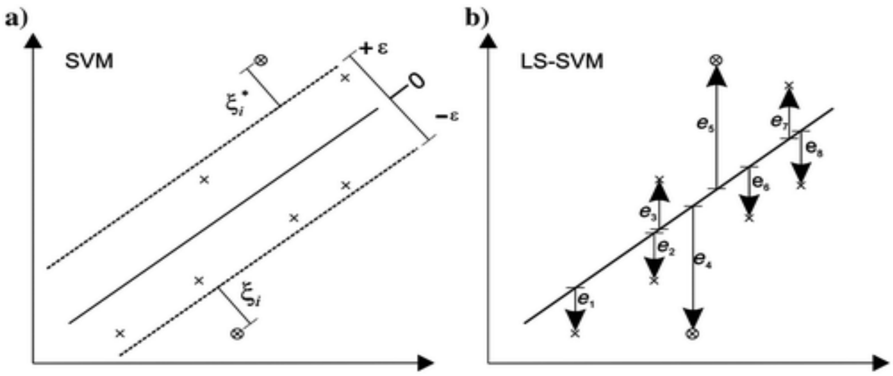


Figure 19:The difference between LS-SVM and SVM

III.4.2.4 Decision tree

Decision trees are one of the best known methods in classification. The principle of decision trees is to achieve the classification of an example by a series of tests on the attributes that describe it. Concretely, in the graphic representation of a tree.

1. An internal node corresponds to a test on the value of an attribute.
2. A branch starts from a node and corresponds to one or more values of this test.
3. A leaf is a node from which no branch starts and corresponds to a class. A decision rule (of the form if . . . then . . .) is created for each path starting from the root of the tree and Traversing the tests (by making conjunctions) to the leaf which is the label

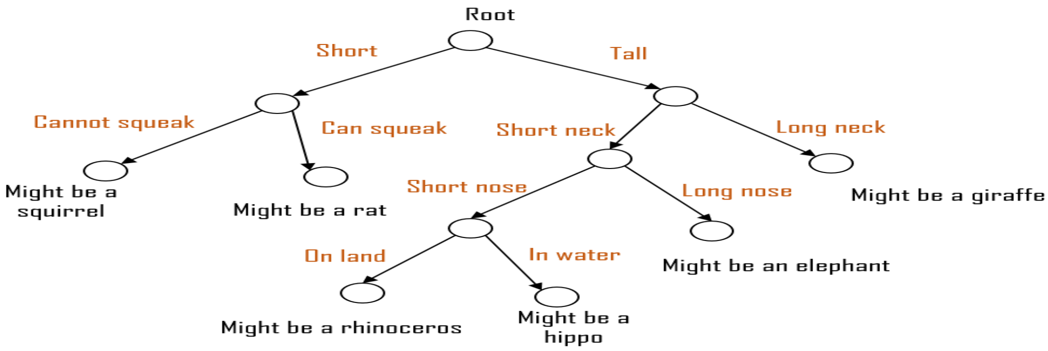


Figure 20:approach of the decision tree working principle

Advantages : Decision Tree is simple to understand and visualize, requires little data preparation, and can handle both numerical and categorical data.

Disadvantages: Decision tree can create complex trees that do not generalize well, and decision trees can be unstable because small variations in the data might result in a completely different tree being generated.

III.4.2.5 Random forests

Random Forests is an extension of Decision Trees with special features. It is an ensemble approach to classification and regression that works by building a multitude of decision trees at training time. The prediction or classification is then made according to a majority voting system within these different trees. The principle of the random forest is then to seek to take advantage of this instability by aggregating them together. **[22]**

Random Forest uses decision trees and statistical theory in combination to reduce classifier variance by averaging a set of decision trees, generating classifiers with very good predictive ability. **[23]**

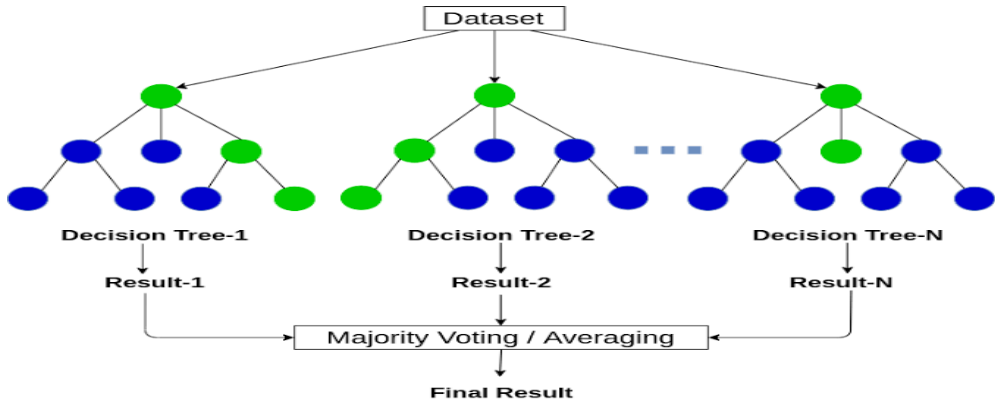


Figure 21:approach of the Random Forest working principle

Advantages: Reduction in over-fitting and random forest classifier is more accurate than decision trees in most cases.

Disadvantages: Slow real time prediction, difficult to implement, and complex algorithm.

III.4.2.6 K- Nearest Neighbor

The k nearest neighbors (kNN) technique is based on the entire data.

Indeed, for an observation, which is not part of the data, that we want to predict, the algorithm will look for the k instances closest to our observation and choose for each observation the majority class among its k closest neighbors.

The k-NN method is a supervised learning technique, and is considered one of the simplest in classification. It allows you to classify a new observation by calculating the distance with the training data, and to take the k nearest neighbors (in terms of distance). Then, observe the class that is predominantly (voted) represented among the k-nearest neighbors and assign this class to the new observation The learning time of the k-NN algorithm is short, the k-NN algorithm calculates the distance between the data points.

Euclidean distance formula is The straight line distance between two points. In a plane with p_1 at (x_1, y_1) and p_2 at (x_2, y_2), it is : $ED = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2 + (Z_2 - Z_1)^2 + \dots}$

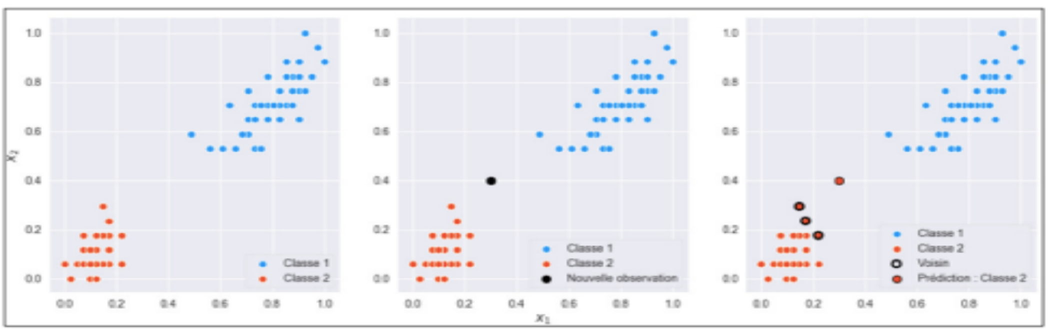


Figure 22:visualization of the k-nearest neighbor

III.4.2.7 Naive Bayes

Bayesian classifiers are statistical classifiers. They can predict class membership probabilities, such as the probability that a given sample belongs to a particular class. Bayesian classifier is based on Bayes' theorem. Naive Bayesian classifiers assume that the effect of an attribute value on a given class is independent of the values of the other attributes. This assumption is called class conditional independence. It is made to simplify the computation involved and, in this sense, is considered ”naive” .Bayes Theorem Bayes Theorem is a principled way of calculating a conditional probability. Conditional probability is the probability that something will happen, given that something else has already occurred. By using conditional probability, we can find out the probability of an event will occur given the knowledge of the previous event.**[24]**

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)}$$

- P(A|B) — is the probability of A given that B has already happened.
- P(B|A) — is the probability of B given that A has already happened.
- P(A) — is the unconditional probability of A occurring.
- P(B) — is the unconditional probability of B occurring.

Naive Bayes is a supervised learning algorithm for classification so the task is to find the class of observation (data point) given the values of features. Naive Bayes classifier calculates the probability of a class given a set of feature values (i.e. p(yi | x1, x2 , … , xn)). Input this into Bayes' theorem: **.[25]**

III.4.2.8 Linear Discriminant Analysis

Linear Discriminant Analysis is one of the commonly used supervised technique for dimensionality reduction. It is also used in classification problems and for data visualizations. The goal is to find a straight line projection of the data that separates the classes as much as possible.. The separation criterion is expressed by maximizing the ratio of inter-class and intra-class variances in this projection. In other words, the distribution of data from the same class should be compact, which would amount to minimizing the dispersion around the mean, as well as maximizing it between the classes.

LDA is a technique that looks for directions that are effective for discrimination
between the data .

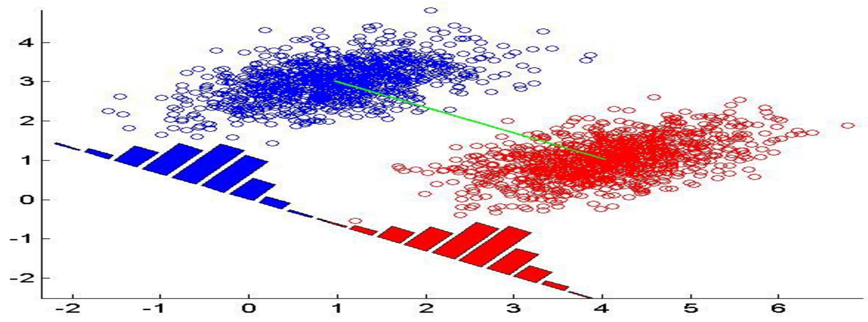


Figure 23:linear discernment analysis (LDA)

III.4.3 Ensemble learning methods

"Strength comes from unity." This old adage reflects the underlying principle that governs the very powerful “ensemble methods” in machine learning. Ensemble learning methods are built on the idea that by mixing models, a far more powerful model can be formed.

So what’s an Ensemble learning?

Ensemble learning is a machine learning paradigm where multiple models (often called “weak learners”) are trained to solve the same problem and combined to get better results. The main hypothesis is that when weak models are correctly combined we can obtain more accurate and/or robust models. **[26].**

So in general it’s a meta approach to machine learning that seeks better predictive performance by combining the predictions from multiple models.

Although there are a seemingly unlimited number of ensembles that you can develop for your predictive modeling problem, there are three methods that dominate the field of ensemble learning. a batter presentation of the methods is shown on the next figure

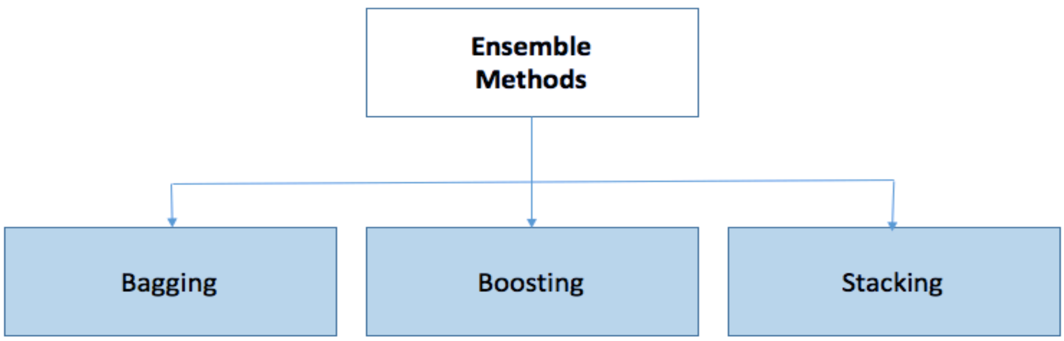


Figure 24:The Main Ensemble Learning Methods

III.4.3.1 Bagging

Bagging, also called bootstrap aggregating, consists of down sampling the data, creating a data set for each model (but similar to the original). Then, during the combination, the predictive analysis is carried out through a majority vote for the classification, or by averaging for the regression. For a batter explanation we present the example in the next figure where the model used is decision tree . At this point we need to mention that random forest is a bagging algorithm because it compose of multiple decision trees as we have mentioned

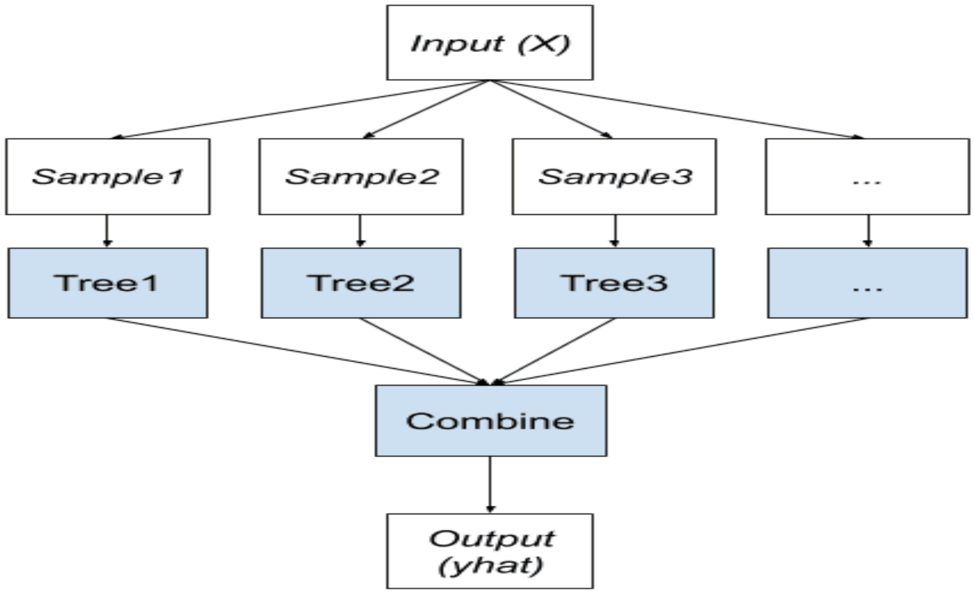


Figure 25:: bagging architecture

III.4.3.2 Random subspace

In machine learning the random subspace method, also called attribute bagging or feature bagging, is an ensemble learning method that attempts to reduce the correlation between estimators in an ensemble by training them on random samples of features instead of the entire feature set.

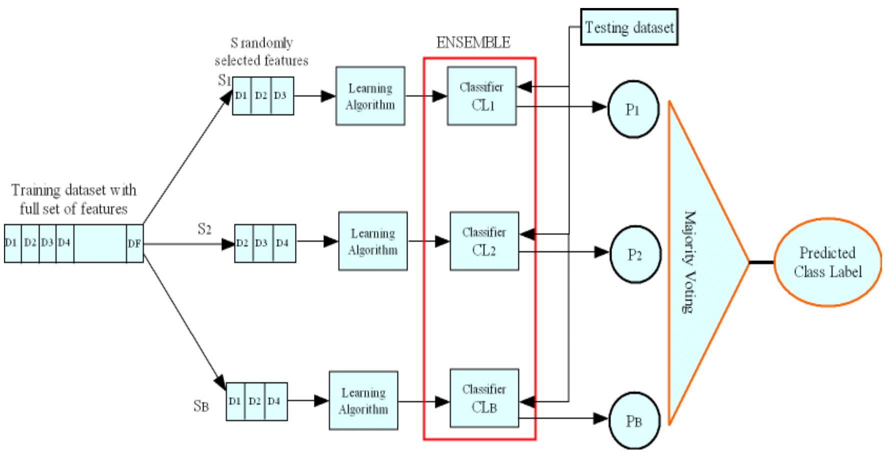


Figure 26:Random subspace architecture

III.4.3.3 Boosting

Boosting will combine the classifier models by weighting them with each new prediction, so that the models having correctly predicted the previous times have a greater weight than the incorrect models. The better a model classifies, the more important it becomes over time.

For a batter explanation we present the example in the next figure where multiple models where used and weighted

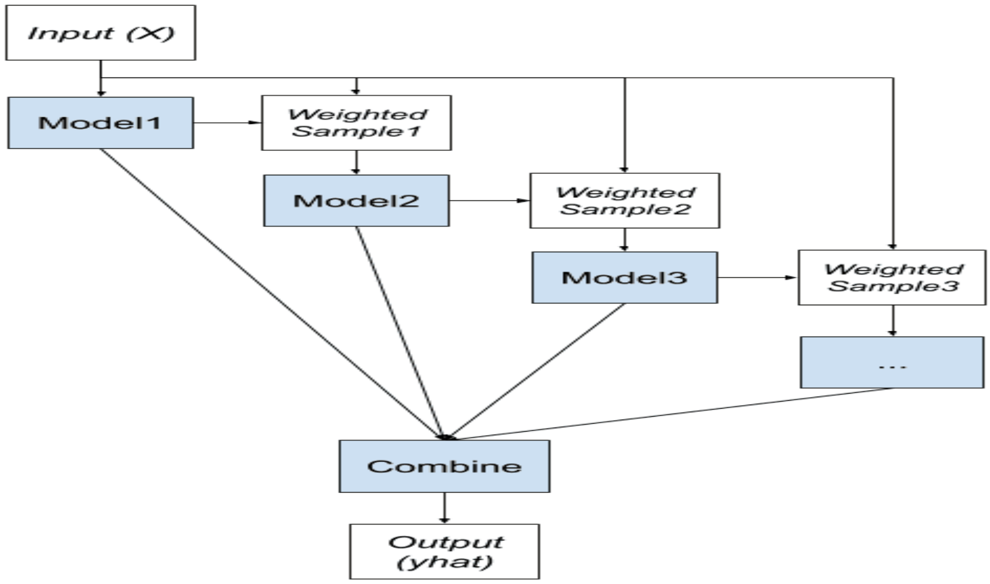


Figure 27:Boosting architecture

III.4.3.4 Stacking

Stacked generalization, or stacking for short, is an ensemble method that seeks a diverse group of members by varying the model types fit on the training data and using a model to combine predictions.

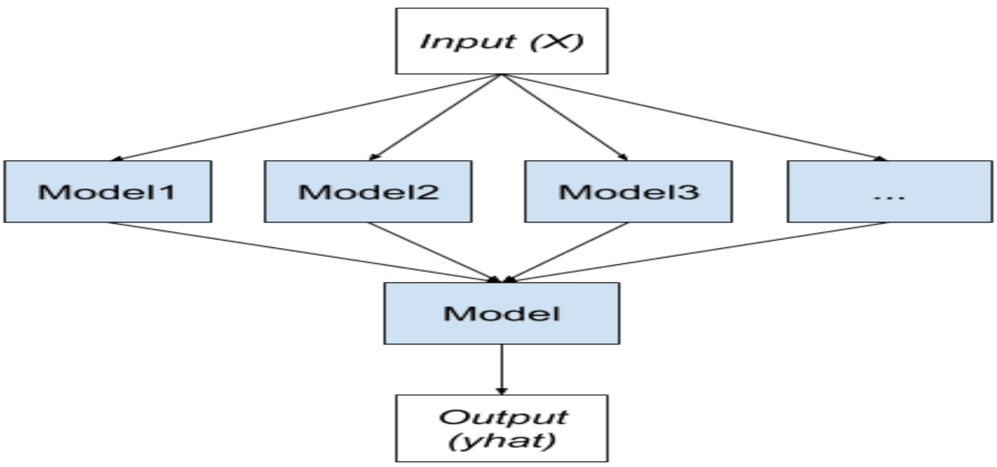


Figure 28:Stacking architecture 1

Stacking has its own nomenclature where ensemble members are referred to as level-0 models and the model that is used to combine the predictions is referred to as a level-1 model

The two-level hierarchy of models is the most common approach, although more layers of models can be used. For example, instead of a single level-0 model we might have 3 or 5 level0 models and a single level-1 model that combines the predictions of level-0 models in order to make a prediction. **[27] .**

A batter presentation is shown the next figure

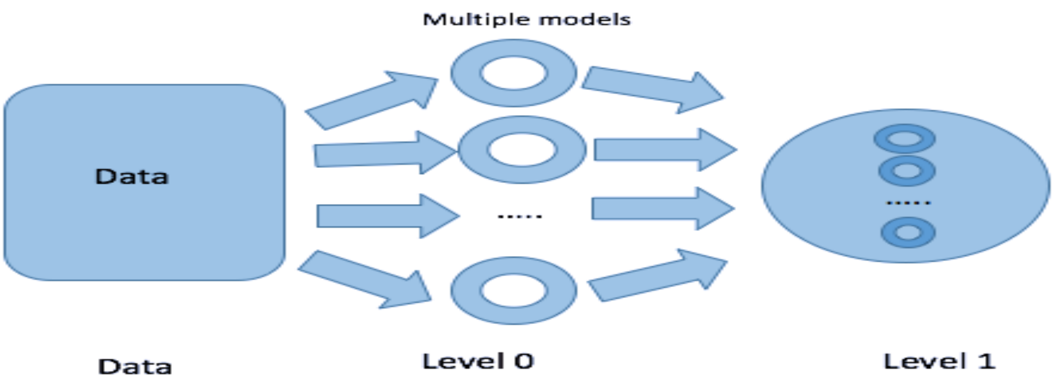


Figure 29:one level Stacking architecture

Stacking is probably the most-popular meta-learning technique. By using a meta-learner, this method tries to induce which classifiers are reliable and which are not.

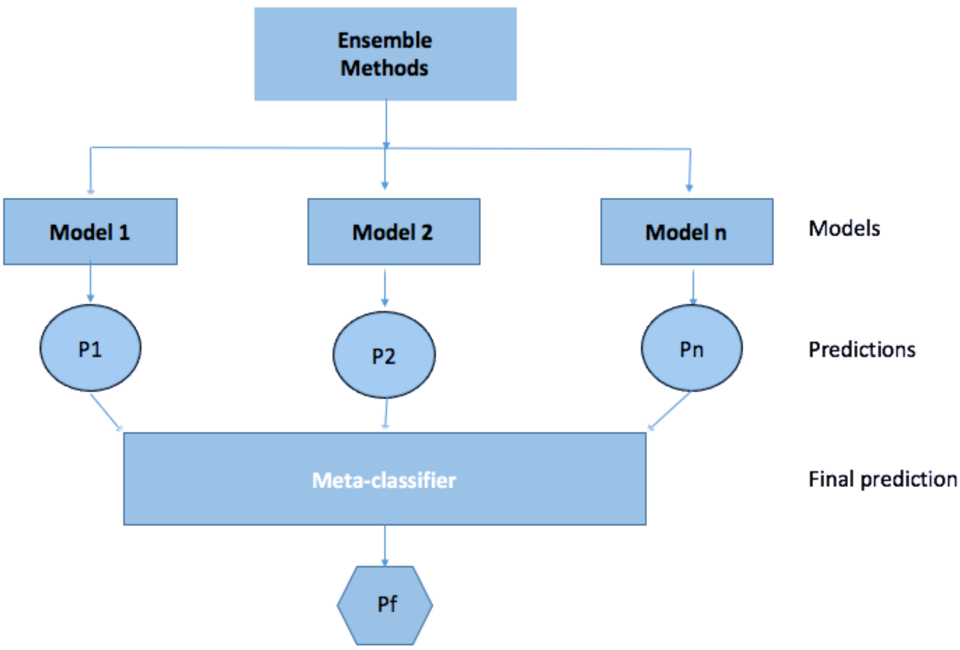


Figure 30:Stacking architecture 2

A possible extension of stacking is multi-level stacking. It consists in doing stacking with multiple layers. As an example, let's consider a 2-levels stacking. In the first level (level 0), we fit the weak learners that have been chosen. Then, in the second level(level 1) , instead of fitting a single meta-model on the weak models predictions (as it was described in the previous subsection) we fit M such meta-models. Finally, in the second level we fit a last meta-model that takes as inputs the predictions returned by the M meta-models of the previous level
From a practical point of view, notice that for each meta-model of the different levels of a multi-levels stacking ensemble model, we have to choose a learning algorithm that can be almost whatever we want (even algorithms already used at lower levels) .

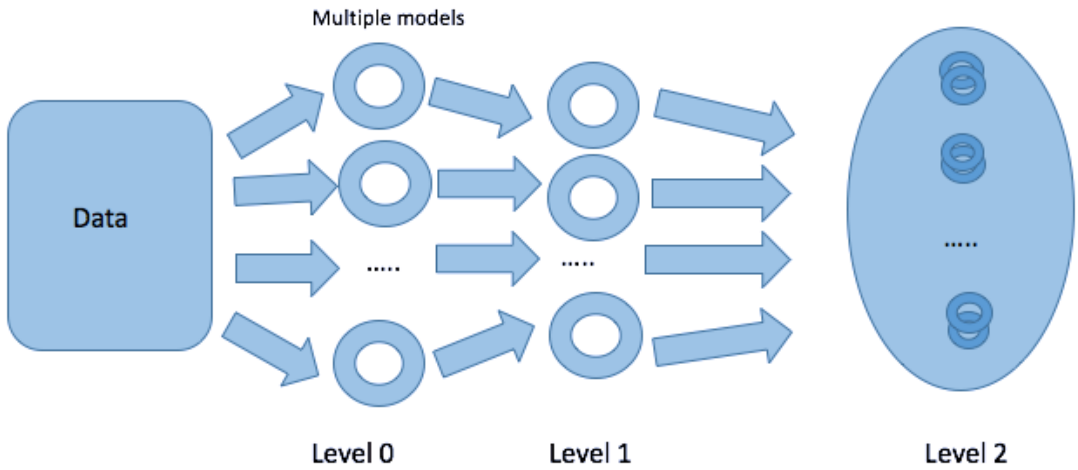


Figure 31:two levels Stacking architecture

III.4.4 Performance evaluation of learning models

III.4.4.1 Confusion Matrix

In the field of machine learning and specifically the problem of statistical classification, a confusion matrix, also known as an error matrix, is a specific table layout that allows visualization of the performance of an algorithm, typically a supervised learning one (in unsupervised learning it is usually called a matching matrix).

Its is an n x n array . Where n is the number of classes in our datasets. In this matrix, the real target classes are crossed with the predicted classes obtained. This gives us the number or percentage of instances that are correctly classified and the number or percentage of misclassified instances.

III.4.4.2 Classification rate (accuracy) and error rate (error accuracy)

We will know the classification rate of the classifier if we can test it on all available examples. Due to the inability to run such tests like most cases, an assessment should be employed instead. The classification rate assessment for a set of Qt test cases is calculated by running the tests on all available examples. The ratio of correctly classified instances (N correct) or incorrectly classified examples (N wrong) must next be calculated (N fault). The percentage of misclassified examples is referred to as error accuracy.

$$accuracy = \frac{NJUST}{QT}$$

$$Erreur\ accuracy = \frac{Nfaut}{QT} = 1 - accuracy$$

III.4.4.3 precision, recall and sensitivity, specificity

Basically, Precision is a way for us to measure how many predictions did our model correctly predict out of all the predictions made.

Actually precision answers about this question: how many correct predictions out of all predictions?

and that’s the formula when we want to calculate it

precision = tp / p = tp / (tp + fp)

Recall and sensitivity on the other hand, is a way for us to measure how many records (or data points, or rows) did we correctly predict out of all records we have .

Recall answers about this question : how many actual record correctly predicted?

and that’s the formula when we want to calculate it

sensitivity = recall = tp / t = tp / (tp + fn)

Actually Precision and Recall only differs on the False Positive and the False Negative

For example purposes, say I want to predict 2 classes which are “Success” and “Fail”. Suppose we want to predict “Success” class ,

False Positive is the number of “Success” class that are not actually “Success”. False Negative is the number of “Success” class we fail to predict as “Success”. So, False Positive and False Negative is, in a way, “the number of predictions that are wrong”.

The Precision and Recall is a metric that we can use to measure model performance when we're doing binary classification or multiclass classification while Sensitivity and Specificity is quite exclusive to binary classification

Specificity talks about the number of negative records correctly predicted.

Actually specificity answers about this question:

how many negative records correctly predicted?

So with Specificity, we can measure how well our model predicts the class that we want to declare as the ***negative*** class. This is, in a way, same as Recall for the ***negative*** class.

We can give this formula as a way to calculate it :

$$\textbf{specificity} = \textbf{tn} / \textbf{n} = \textbf{tn} / (\textbf{tn} + \textbf{fp})$$

III.5 Conclusion

We've covered the basics of Artificial Intelligence and machine learning in this chapter, by giving general definitions, types, and applications. also an explanation of a few major machine learning models.as well as ensemble learning and its most common methods where we focused on stacking and Boosting.

In the next chapter we will present an application in which we will make the diagnostic of

CODLAG system using different models of Machine Learning and Ensemble Learning method with the purpose of improving condition-based maintenance of naval propulsion plants .

Chapter IV: Ensemble Learning (Boosting and Stacking) for classification

IV.1 Introduction

Due to the sensitivity and the necessary precision which should be, and the seriousness of any error in the field where frigates are used we needed to a sufficient method to use for the diagnostic of its naval propulsion plans CODLAG.

To reach this goal we used Simple Learning (classifiers) and Ensemble Learning methods on the data gathered that we have.

In this chapter we will give details of the method used for the diagnostic of our naval propulsion system with interpretation and comparison of the results obtained to demonstrate the effectiveness of the methods used compared to other methods or techniques .

IV.2 Description of system and data base used

Data have been generated from a sophisticated simulator of a Gas Turbines (GT), mounted on a Frigate . characterized by a COmbined Diesel eLectric And Gas (CODLAG) propulsion plant type. The experiments have been carried out by means of a numerical simulator of a naval vessel (Frigate) characterized by a Gas Turbine (GT) propulsion plant.

The different blocks forming the complete simulator (Propeller, Hull, GT, Gear Box and Controller) have been developed and fine tuned over the year on several similar real propulsion plants. In view of these observations the available data are in agreement with a possible real vessel. The propulsion system behavior has been described with this parameters : Compressor degradation coefficient kMc and Turbine degradation coefficient kMt.

so that each possible degradation state can be described by a combination of this triple (kMc,kMt). The range of decay of compressor and turbine has been sampled with an uniform grid of precision 0.001 so to have a good granularity of representation.

In particular for the compressor decay state discretization the kMc coefficient has been investigated in the domain [1; 0.95], and the turbine coefficient in the domain [1; 0.975].

Ship speed has been investigated sampling the range of feasible speed from 3 knots to 27 knots with a granularity of representation equal to tree knots.

A series of measures (16 features) which indirectly represents of the state of the system subject to performance decay has been acquired and stored in the dataset over the parameter's space.**[28]**

- A 16-feature vector containing the GT measures at steady state of the physical asset:

1 - Lever position (lp) [ ]

2 - Ship speed (v) [knots]

3 - Gas Turbine shaft torque (GTT) [kN m]

4 - Gas Turbine rate of revolutions (GTn) [rpm]

5 - Gas Generator rate of revolutions (GGn) [rpm]

6 - Starboard Propeller Torque (Ts) [kN]

7 - Port Propeller Torque (Tp) [kN]

8 - HP Turbine exit temperature (T48) [C]

9 - GT Compressor inlet air temperature (T1) [C]

10 - GT Compressor outlet air temperature (T2) [C]

11 - HP Turbine exit pressure (P48) [bar]

12 - GT Compressor inlet air pressure (P1) [bar]

13 - GT Compressor outlet air pressure (P2) [bar]

14 - Gas Turbine exhaust gas pressure (Pexh) [bar]

15 - Turbine Injection Control (TIC) [%]

16 - Fuel flow (mf) [kg/s] **[28]**

we need to mention before we started using the machine learning algorithms we deleted 3 features that had a fixed value : 12 - GT Compressor inlet air pressure (P1) [bar] , 9 - GT Compressor inlet air temperature (T1) [C] . and the third one had the same values as the previous feature 7 - Port Propeller Torque (Tp) [kN] .

IV.3 Design and methodology of the technique used

In our work we classified the phase of:

GT Compressor decay state coefficient and GT Turbine decay state coefficient.

At this point we need to mention that the classes were defined as :

When GT Compressor decay state coefficient equal to :

0.95 and 0.96 its class 4

0.97 and 0.98 its class 3

0.99 its class 2

1 its class 1

When GT Turbine decay state coefficient equal to :

0.97 its class 4

0.98 its class 3

0.99 its class 2

1 its class 1

IV.3.1 The failures treated in our work

IV.3.1 .1 GT Compressor decay state coefficient.

- Very bad state
- Bad state
- Acceptable state
- Normal state

IV.3.1 .2 GT Turbine decay state coefficient.

- Very bad state
- Bad state
- Acceptable state
- Normal state

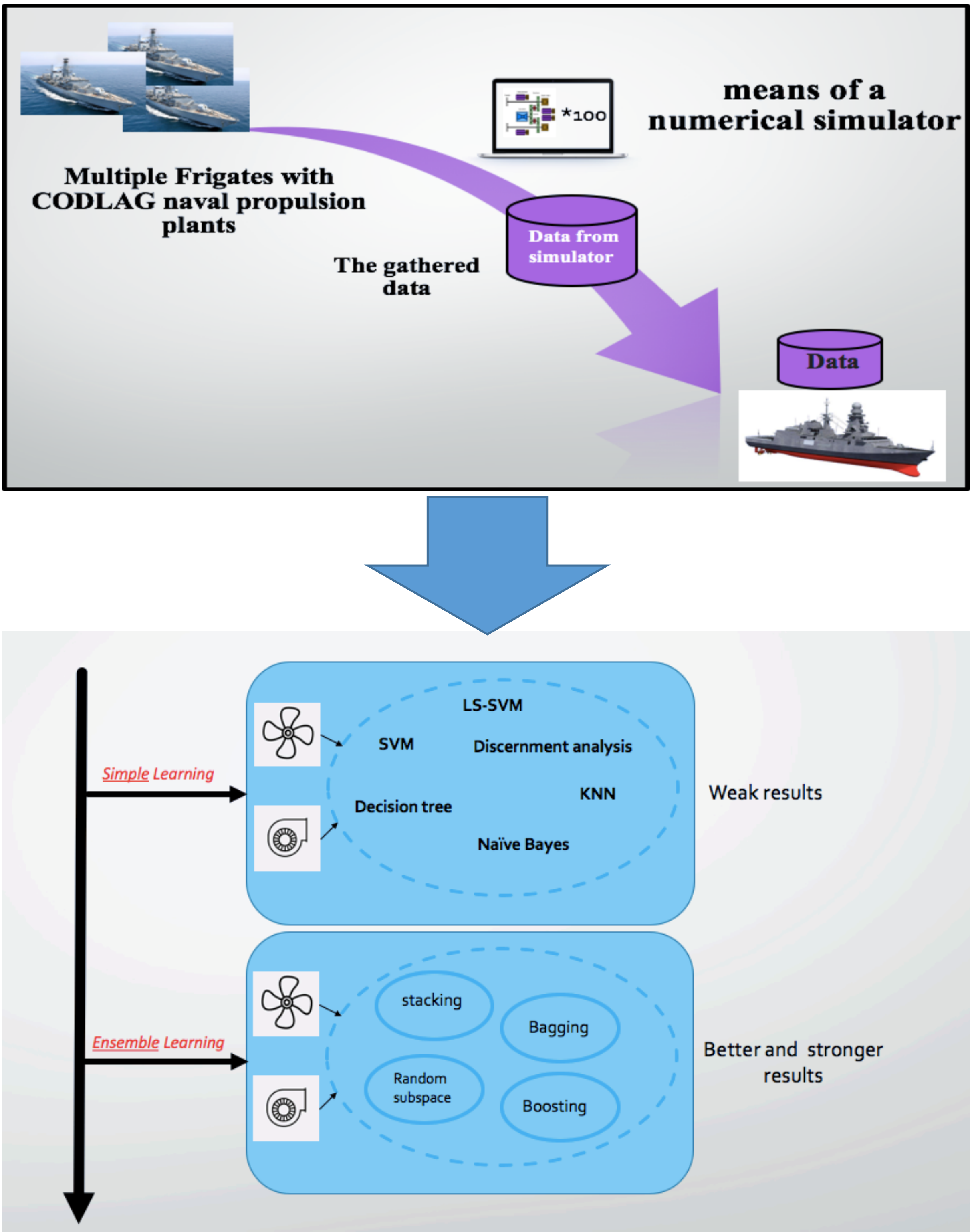


Figure 32:A diagram of the structure and the techniques used

As a first step we thought about using simple learning algorithms which mean using simple classifiers such ass svm to work alone and Run it to Make the best possible prediction , but we wasn't sure about the effectiveness of the results that we can get . That's exactly what happed and we got week accuracy as shown in the diagram above.

The results we got was varying between very week such as svm and Ls-svm and Naïve Bayes but relatively strong such as discernment analysis and KNN by varying the number of the nearest neighbors and above 90 such ass decision tree.

These last classifier results could be considered as good results but in another field maybe , due to the sensitivity of the system that we are working on and where it is used we needed to try more and look for batter and stronger results .

The raison why we thought about Ensemble Learning where multiple learners are set to work together in different methods and structure with optimism for batter results that one learner can get .

That was exactly what we got by using four methods of Ensemble Learning on the both of failures.

where the best result for GT Turbine decay state coefficient obtained by using Boosting methods and the best results for GT Compressor decay state coefficient obtained by using Stacking method.

For a batter visualization of the methods we propose this, diagram below shows the different learners which are classifiers in our case we used for Boosting and Stacking.

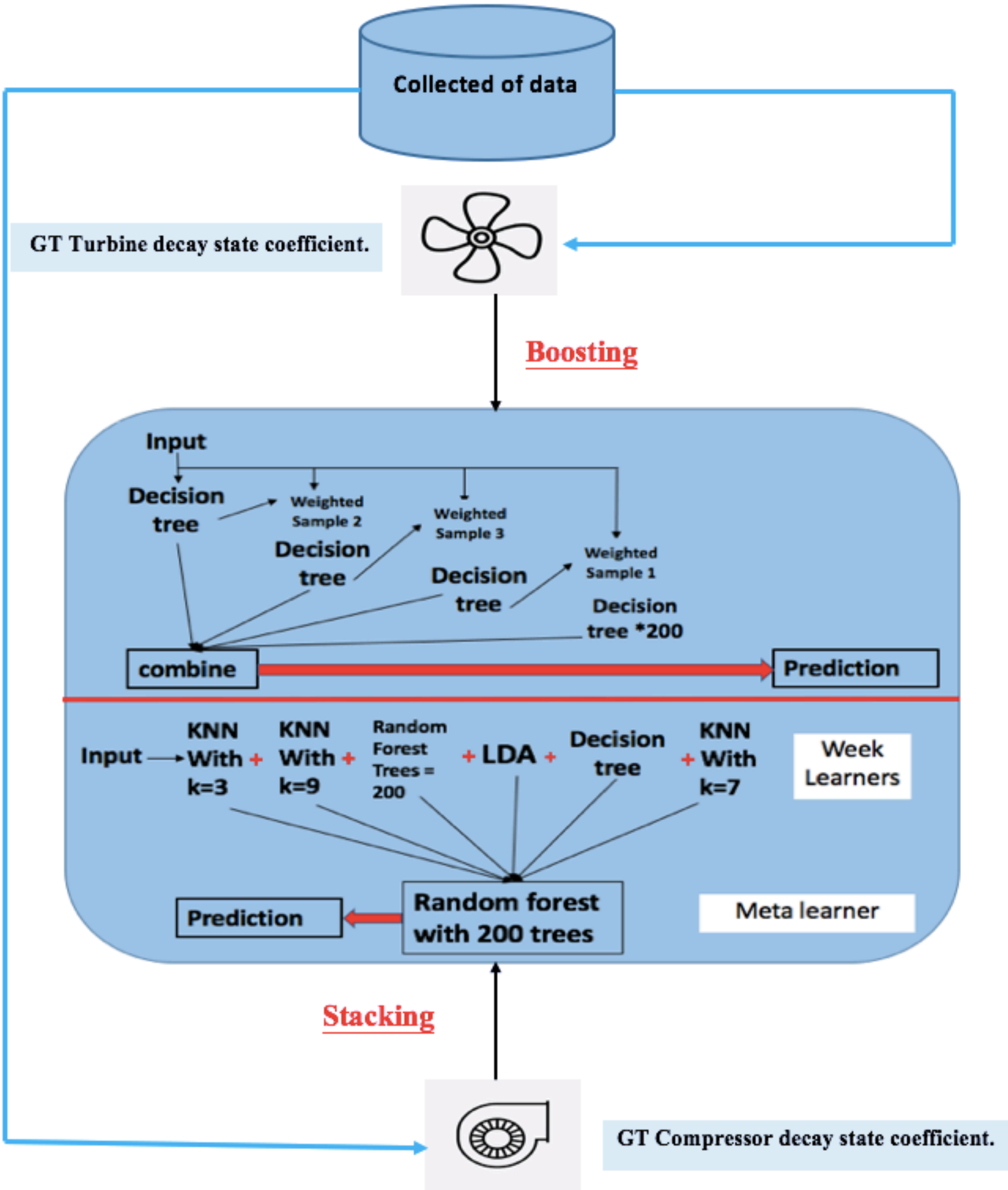


Figure 33:A diagram the of classifiers used in the Boosting and Stacking methods

Chapter IV: Ensemble Learning (Boosting and Stacking) for classification

We varied a lot of learners until we got the best results by this combination of classifiers ,

For GT Turbine decay state coefficient by boosting method where we used 200 decision trees.

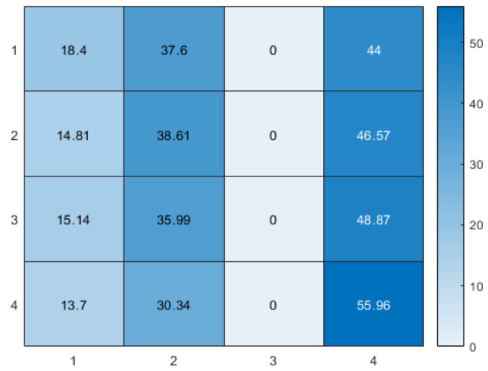
And Stacking for GT Compressor decay state coefficient by using the kNN with different K-nearest neighbor 3,7,9 and random forest with 200 trees and linear discernment analysis and decision tree as week learners and random forest with 200 trees as a meta learner , which was actually the learners that gave batter results when first was used in simple learning means when they were used separated as one classifier .

IV.4 Evaluation and interpretation of the obtained results

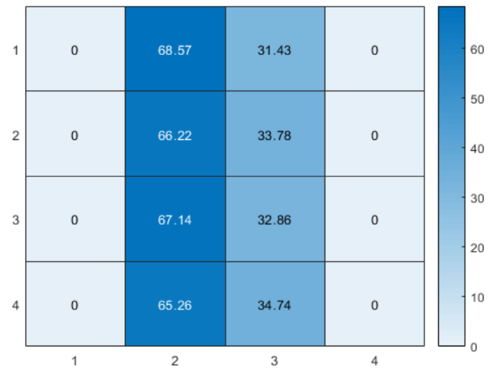
IV.4.1 Results of the simple learning classifiers

	SVM	LS_SVM	KNN	Decision tree	Naïve Bayes	Linear Discernment analyses
	34.4876	31.0541	85.9816	90.7847	35.9955	82.8540
	35.7855	33.4728	87.0427	90.7288	38.7601	82.1000
	34.1801	31.1986	87.1824	90.7568	34.2642	83.2170
	35.6205	31.5752	86.8193	90.0307	38.6205	83.9710
	34.0433	32.1404	86.0933	90.5334	37.0288	83.0494
	34.8857	33.0010	87.5454	90.3937	35.8838	82.9657
	34.8313	32.8670	86.5959	91.0360	37.4756	82.6585
	34.4404	32.3604	87.21033	90.4496	37.5593	83.8313
	34.1971	30.3020	86.8193	90.5054	38.7601	83.4404
	35.8126	31.0220	86.9310	90.8126	37.1963	82.8819
MIN	**34.0433**	**30.3020**	**85.9816**	**90.0307**	**34.2642**	**82.1000**
MEAN	**34.8284**	**31.8993**	**86.8221**	**90.6032**	**37.1544**	**83.0969**
MAX	**35.8126**	**33.4728**	**87.5454**	**91.0360**	**38.7601**	**83.9710**
STD	**0.6845**	**1.0289**	**0.4891**	**0.2815**	**1.4492**	**0.5524**

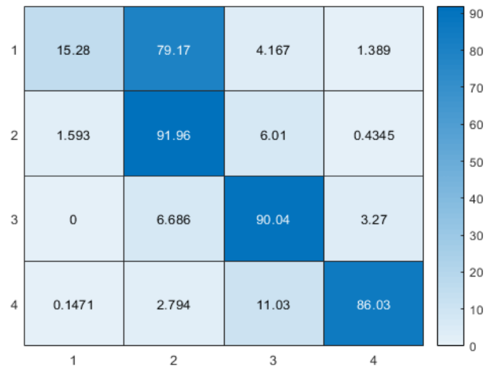
Table 1:Results of Simple Learning classifiers on the GT turbine decay coefficient state



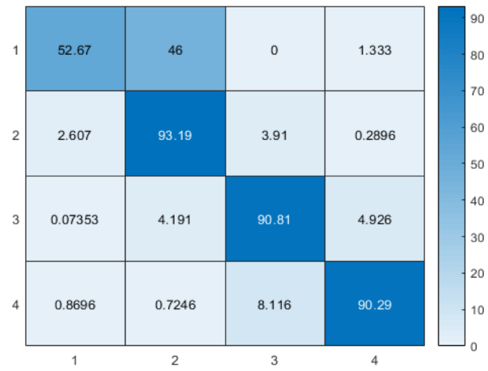
SVM confusion matrix



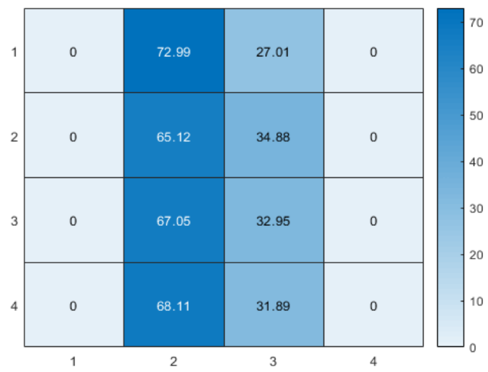
LS-SVM confusion matrix



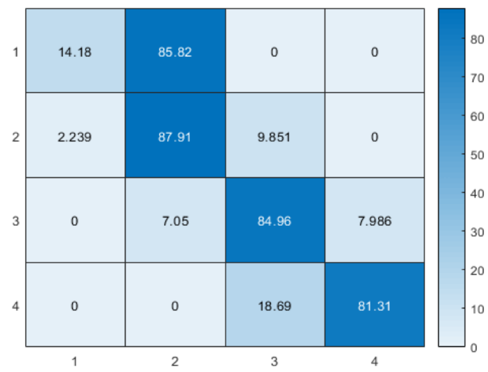
KNN confusion matrix



Decision tree confusion matrix



Naïve Bayes confusion matrix



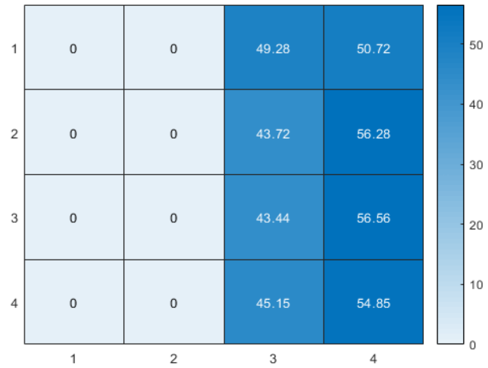
LDA confusion matrix

Figure 34:confusion matrixs of Simple Learning classifiers on the GT turbine decay coefficient state

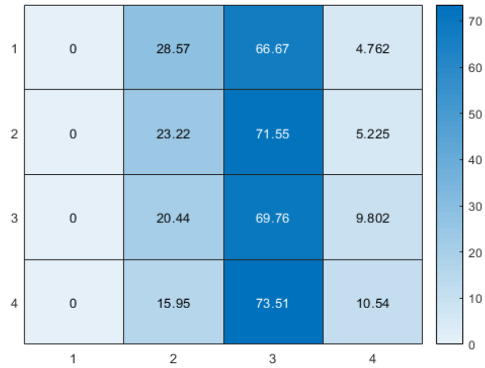
The results of simple classifiers on the turbine represented by the 10 trials and a confusion matrix of each classifier shows that the best prediction until now was predicted by Decision tree with the a mean of 90.6032of accuracy with a low STD 0.2815 .

	SVM	LS_SVM	KNN	Decision tree	Naïve Bayes	Linear Discernment analyses
	41.6643	32.0220	91.5108	92.7395	42.6979	78.8327
	40.6231	34.9110	91.9017	92.1251	44.2893	79.7822
	41.2101	33.1610	90.7568	92.2368	45.2667	79.0841
	41.1079	32.4179	90.1703	92.4540	42.5300	79.5588
	42.0110	31.4210	90.8964	91.8179	41.2455	80.1173
	40.6132	32.4061	90.5064	92.0972	44.1776	79.0841
	41.2193	32.0893	92.1810	92.3764	39.2069	79.8660
	41.0708	31.0178	91.0018	91.8738	39.3466	79.9218
	41.3300	32.2170	91.7342	92.5998	38.6484	78.6931
	40.1718	32.7340	91.3991	92.7395	41.1274	80.3686
MIN	**40.1718**	**31.0178**	**90.1703**	**92.7395**	**38.6484**	**78.6931**
MEAN	**41.1022**	**32.4397**	**91.2058**	**92.3060**	**41.8536**	**79.5309**
MAX	**42.0110**	**34.9110**	**92.1810**	**92.7395**	**45.2667**	**80.3686**
STD	**0.5328**	**1.0605**	**0.6451**	**0.3325**	**2.3241**	**0.5738**

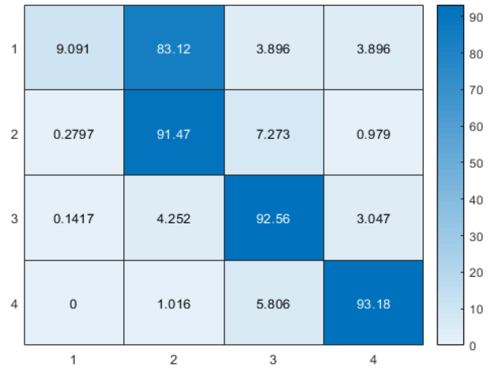
Table 2:Results of Simple Learning classifiers on the GT compressor decay coefficient state



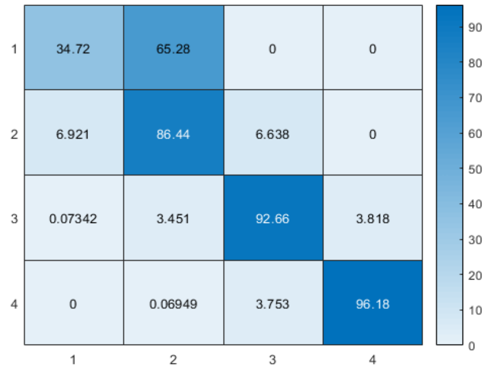
SVM confusion matrix



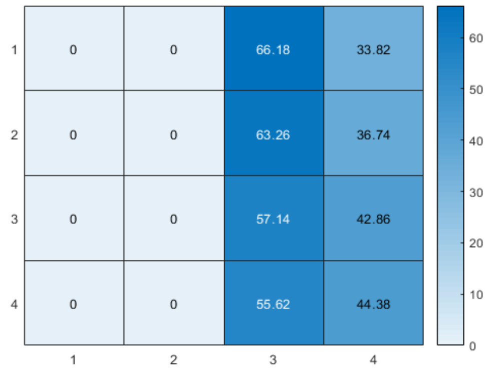
LS-SVM confusion matrix



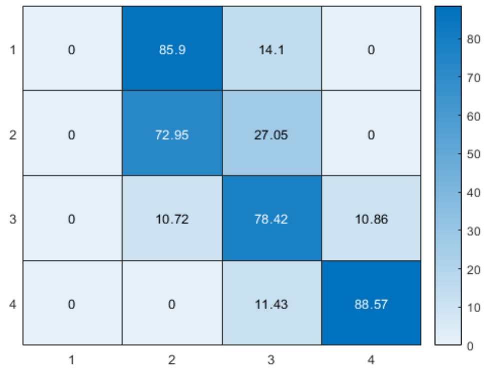
KNN confusion matrix



Decision tree confusion matrix



Naïve Bayes confusion matrix



LDA confusion matrix

Figure 35:confusion matrixes of Simple Learning classifiers on the GT compressor decay coefficient state

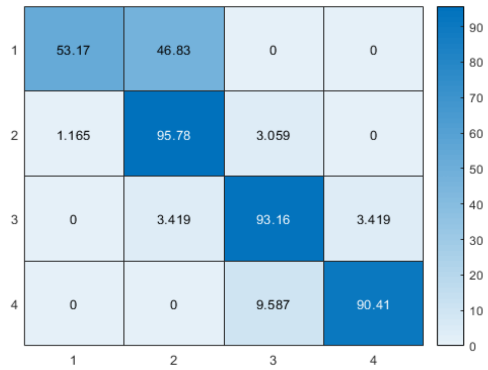
The results of simple classifiers on the compressor represented by the 10 trials and a confusion matrix of each classifier shows that the best prediction until now was predicted by Decision tree with the a mean of 92.3060 of accuracy with a low STD 0.3325 .

But as we have mentioned previously in this chapter, the results of the classifiers alone were not enough that's why we moved to Ensemble Learning.

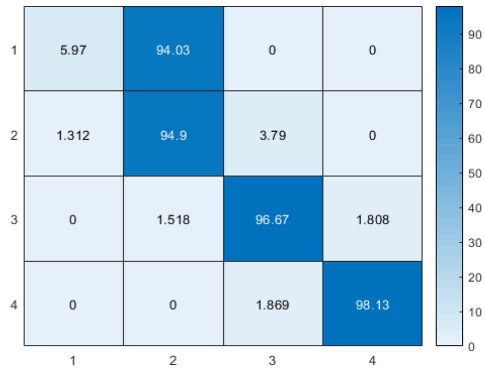
IV.4.2 Results of the ensemble learning classification

	Bagging	Stacking	Random subspace	Boosting
	95.2248	95.3021	94.9735	**96. 3139**
	94.0519	95.9176	95.7833	**96.2863**
	94.3033	95.9509	95.4203	**96.3071**
	93.7448	95.0572	95.6716	**96.3037**
	93.5214	95.3086	93.6942	**96.2940**
	94.2753	95.4203	95.1969	**96.2832**
	94.6104	95.1689	95.2527	**96.2991**
	94.1916	95.7554	94.7501	**96.3021**
	94.3033	95.1969	94.8365	**96.3100**
	92.7957	95.2248	95.6700	**96.2801**
MIN	**95.2248**	**95.0572**	**93.6942**	**96.2801**
MEAN	**94.0022**	**95.4303**	**95.1249**	**96.2979**
MAX	**92.7957**	**95.9509**	**95.7833**	**96.3139**
STD	**0.6921**	**0.3249**	**0.6175**	**0.0117**

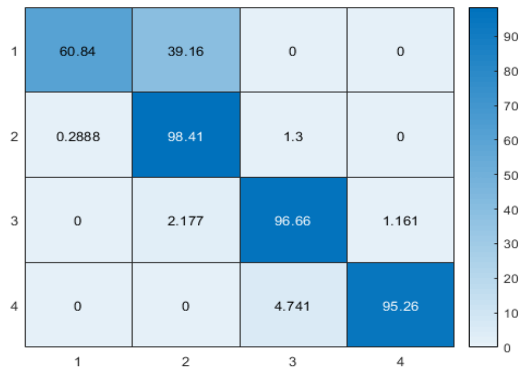
Table 3:Results of Ensemble Learning methods on the GT turbine decay coefficient state



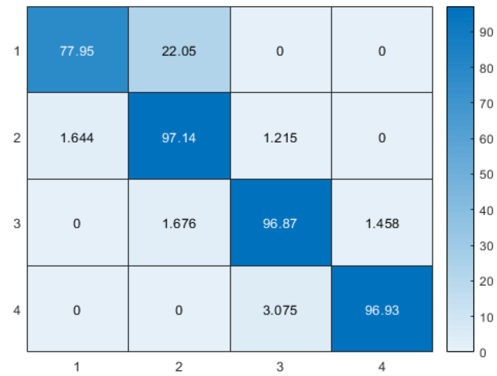
Bagging confusion matrix



stacking confusion matrix



Random subspace confusion matrix



Boosting confusion matrix

Figure 36:confusion matrixes of Ensemble Learning methods on the GT turbine decay coefficient state

The results of Ensemble Learning methods on the GT turbine decay coefficient state represented by the 10 trials and a confusion matrix of each method shows that we got batter and stronger accuracy and results then the simple learning classifiers.

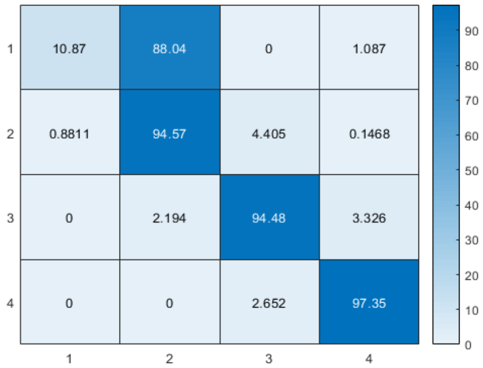
Where the mean accuracy of Bagging was 94.0022 with a relatively high STD , and a batter mean accuracy with Stacking and Random subspace 95.4303 and 95.1249 with a low STD of the Stacking and relatively high one with Random subspace .

But the best prediction means best results considering accuracy and STD was given by Boosting where the mean of accuracy was **96.2979** and the max value that accuracy took was

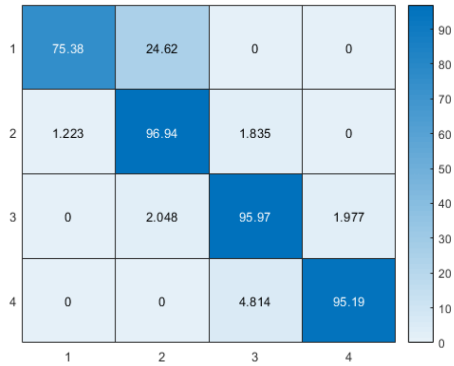
96.3139 we can take it as the best prediction and not take the mean of the teen trials considering the very low STD which was **0.0117.**

	Bagging	Stacking	Random subspace	Boosting
	94.2753	**95.1200**	94.4708	95.5040
	94.2195	**95.1689**	94.5546	94.4150
	94.0799	**95.4203**	94.6663	94.6663
	94.1636	**95.1812**	94.4429	94.5825
	94.0240	**95.1103**	94.6384	94.3591
	94.5267	**95.1902**	94.7221	94.2474
	94.1357	**95.1922**	93.9402	94.8006
	93.8844	**95.1520**	94.7221	94.5267
	93.6610	**95.1401**	94.6942	94.2472
	94.0559	**95.1680**	94.8059	94.2753
MIN	**93.6610**	**95.1103**	**93.9402**	**94.2472**
MEAN	**94.0834**	**95.1843**	**94.5657**	**94.5624**
MAX	**94.5267**	**95.4203**	**94.8059**	**95.5040**
STD	**0.2863**	**0.0875**	**0.2481**	**0.3800**

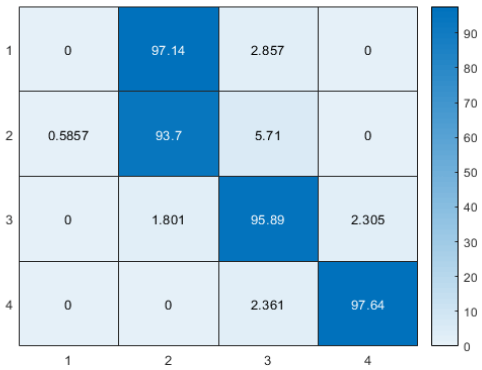
Table 4:Results of Ensemble Learning methods on the GT compressor decay coefficient state



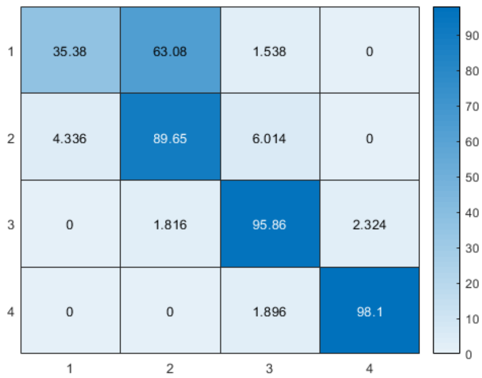
Bagging confusion matrix



stacking confusion matrix



Random subspace confusion matrix



Boosting confusion matrix

Figure 37:confusion matrixes of Ensemble Learning methods on the GT compressor decay coefficient state

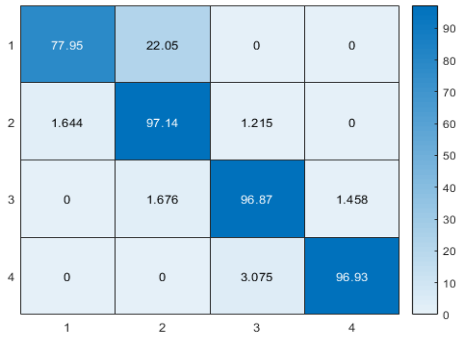
The results of Ensemble Learning methods on the GT compressor decay coefficient state presented by the 10 trials and a confusion matrix of each method shows that we got batter and stronger accuracy and results then the simple learning classifiers.

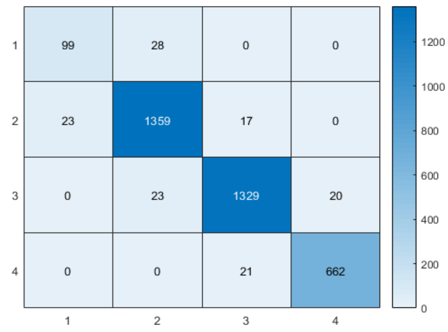
Where the mean accuracy of Bagging and random subspace and boosting was 94.0834 and 94.5657 with and 94.5624 with a relatively low STD between 0.200 and 0.400

But the best prediction means best results considering accuracy and STD was given by Stacking where the mean of accuracy was **95.1843** and the max value that accuracy took was **95.4203** we can take it as the best prediction and not take the mean of the teen trials considering the very low STD which was **0.0875.**

IV.4.3 calculate precision, recall and sensitivity, specificity

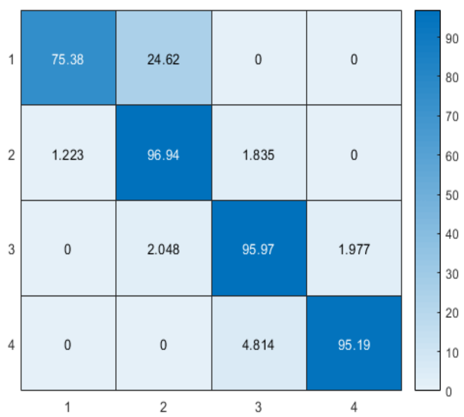
IV.4.3.1 Boosting Matrix

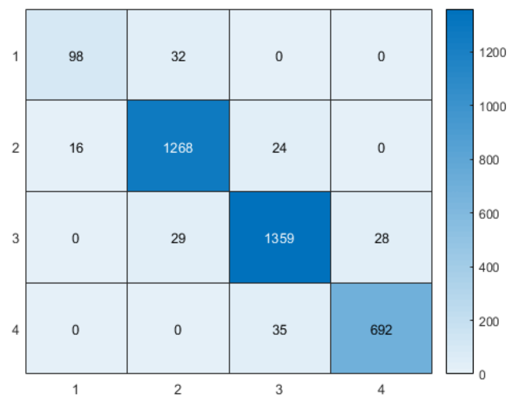




- **class 1 :** precision = 81.14 , sensitivity =77.97 , specificity = 99.32
- **class 2 :** precision = 96.38 , sensitivity =97.14 , specificity = 97.66
- **class 3 :** precision = 97.22 , sensitivity =96.86 , specificity = 98.27
- **class 4 :** precision = 97.06 , sensitivity =96.92 , specificity = 98.82

IV.4.3.2 Stacking Matrix





- **class 1 :** precision = 85.96 , sensitivity = 75.38 , specificity =95.53
- **class 2 :** precision = 95.41 , sensitivity =96.94 , specificity = 97.31
- **class 3 :** precision = 95.83 , sensitivity =95.97 , specificity = 97.27
- **class 4 :** precision = 96.1 , sensitivity =95.18 , specificity = 99.01

Chapter IV: Ensemble Learning (Boosting and Stacking) for classification

IV.5 conclusion

In this chapter we proposed a new method of diagnostic consist of using Machine learning algorithms to improve conditional based maintenance of navel propulsion plants of the frigates by using two ensemble leaning methods which was Boosting and Stacking .

These two methods gave good results that we can trust for the diagnostic of our system.

General conclusion

General conclusion

Any failure in the propulsion plant of these frigates that we are dealing with in this which called (CODLAG) combined diesel electric and gas may inevitably lead to human and natural disasters and an exorbitant maintenance and repairing costs, which leads explicitly to the decline of the yield of the institution and the navy branch of any country.

For these raisons, we proposed and worked on a new method of diagnosis based on artificial intelligence in order to improve conditional based maintenance of the propulsion plant of this important type of ships.

We started by using simple classifiers for the classification of the degradation states of the GT Compressor decay coefficient and GT Turbine decay coefficient, where we tried the SVM, LSSVM, KNN, Decision tree, Naïve Bayes and linear discernment analysis.

Actually We bumped to week results the raison why we thought that "unity is strength "and we moved to ensemble learning where we used Bagging, Random subspace , Stacking and boosting where we got batter and stronger results with Boosting and Stacking as it's shown in chapter four of this work .

The approach and results shown in this work, clearly represent the effectiveness of ensemble learning methods on simple learning methods.

Therefore, This result suggest that stacking and boosting methods could be a useful tool for the diagnosis of the CODLAG naval propulsion plans of a Frigate and it can be relied upon to improve its conditional based maintenance.

References

References

[1] Henderson, James: *Frigates Sloops & Brigs*. Pen & Sword Books, London, 2005.ISBN 1-84415-310-0

[2] *www.military-today.com. Archived from the original on 29 June 2017. Retrieved 13 July 2017*

[3] *Gardiner, Robert, ed. (1980). Conway's All the World's Fighting Ships, 1922–1946. New York: Mayflower Books.ISBN 0-8317-0303-2 .*

[4] Geoffrey Parker, *The Military Revolution: Military Innovation and the Rise of the West 1500–1800*, p. 99

[5] *Lavery, Brian (1989). Nelson's Navy: The Ships, Men and Organisation 1793–1815. Annapolis, MD: Naval Institute Press.ISBN 978-1-59114-611-7 .*

[6] *« Questions et réponses avec BAE Systems sur la mise à jour de la conception des frégates de type 26 à Euronaval 2012 » . Belgique : navyrecognition.com. 10 janvier 2013 . Consulté le 19 mars 2013*

[7] . G. Hill and C. R. Peterson, Mechanics and Thermodynamics of Propulsion. Reading, MA, Addison Wesley Publishing Co., 1992, 764 p., 1992.

[8] J.-C. Han, S. Dutta, and S. Ekkad, Gas Turbine Heat Transfer and Cooling Technology. CRC Press, 2012

[9] (Norme AFNOR X 60-010)

[10] cour de maintenance niversité de batna 2 monsieur mokhtari_messouad introduction a la maintenance

[11] Said Alem ,(maintenance course) univ-boumerdes

[12] **A.Akhenak.** Conception d'observateurs non linaires par approche multi modèle, application au Diagnostic, Thèse de Doctorat, Ecole doctorale IAEM de Lorraine, 2004

[13] Bensana Toufik . (Diagnosis of failures based on the vibration analysis of a steam turbine. UNIVERSITE BADJI MOKHTAR – ANNABA.2008

14] Bouguessa, M. (2015b). Forage de données. notes de cours. Université du Québec à Montréal , 88 pages

[15] Boudheb Tarik. 2018/2019.privacy Preserving Classification of Biomedical Data. Université Djillali Liabes faculté des sciences exactes Sidi Bel Abbes. Page 32

[16]. J. Ah-Pine (Univ-Lyon 2) Apprentissage automatique M2 DM 2019/2020 8 / 359

References

[17] Fatima mahmoude .mars 2018,(approche d'apprentissage automatique pour la détection du spam wep informatique :exploitation de divers caractéristique. Université de Québec à montérial.page 72

[18] WendBenedo Arnaud Bienvenue ZOUNGRANA.2020. (Application des algorithmes d'apprentissage automatique pour la détection de défauts de roulements sur les machines tournantes dans le cadre de l'Industrie 4.0), page 60

[19] Ahmad I, Abdulah AB, Alghamdi AS. Towards the designing of a robust intrusion detection system through an optimized advancement of neural networks, in Advances in Computer Science and Information Technology, ed: Springer, 2010; pp. 597-602.

[20] Han J, Kamber M, Pei J. Data mining: concepts and techniques: Elsevier 2011

[21] HOUAZENE Sabrina et MANSOUR Souad « detection des scrits publiicitaires a base d'apprentissage automatique profond » Master en Informatique 2019.

[22]. J. Ah-Pine (Univ-Lyon 2) Apprentissage automatique M2 DM 2019/2020 8 / 359

[23]. JOSE ALEJANDRO LOPEZ GONZALEZ. Juin 2016.(exploitation des arbre de décision et des support vecteur machine en vue d'application dans l'analyse de texte université du Québec

[24] M. Kantardzic, Data Mining - Concepts, Models, Methods, and Algorithms, IEEE Press, Wiley-Interscience, 2003, ISBN 0-471- 22852-4.

[25] Jiawei Han and Micheline Kamber, Data Mining: Concepts and Techniques, Elsevier 2006, ISBN 1558609016. This part of the lecture notes is derived from chapter 6.4 of this book.

[26]. A book : Patter classification using Ensemble methods , 2010.

[27]. A book : Ensemble Methods in data Mining , 2010.

[28] UCI Machine Learning Repository: Condition Based Maintenance of Naval Propulsion Plants Data Set .