UNIVERSITY OF KASDI MERBAH OUARGLA
Faculty of New Information Technologies and Communication
Department of Electronic and Telecommunication

ACADEMIC MASTER MEMORY

Domain : Science and Technology

Field : Telecommunication

Specialty : System of Telecommunication

Topic

Machine learning to detect covid-19 using cough sounds

Presented by :

- Mr.Brahim Elkhalil GHOUAR
- Mr.Mohammed Larbi SENOUSSI

Publicly defended on : 14 June 2022

before the jury :

Ms.Fella CHARIF	President	MCA	UKM OUARGLA
Mr.Abdelmadjid YOUCEFA	Examiner	MCB	UKM OUARGLA
Mr.Khaled BENSID	Supervisor	MCB	UKM OUARGLA
Mr.Elmoundher HADJAIDJI	Co-Supervisor	Doctorant	UKM OUARGLA

University year : 2021/2022

Abstract

The objective of this project is to design a diagnostic aid system for the early COVID-19 Detection disease from the voice. Mainly, the proposed system is based on two main steps: feature extraction of sound and classification. in this case we have chosen the Mel-Frequency Cepstral Coefficient (MFCC). The classification process is based on three machine learning supervised classifiers: -Support Vector Machine (SVM) -K-nearest neighbors (KNN) - Decision tree (DT). Our proposed system evaluated using TOS. The performance used of our system are the accuracy, sensitivity, and specificity, F1 score, and Receiver Operating Characteristics (ROC).

Keywords — : voice,COVID-19,features extraction, classification.

ملخص

الهدف من هذا المشروع هو تصميم نظام مساعدات تشخيصية للكشف المبكر عن مرض COVID-19 من الصوت. يعتمد النظام المقترح بشكل أساسي على خطوتين رئيسيتين: استخراج السمات للصوت والتصنيف. ميزة الاستخراج التي اخترناها هي معامل Mel-Frequency Cepstral .(MFCC) تعتمد عملية التصنيف على ثلاثة مصنفات خاضعة للإشراف على التعلم الآلي: Machine Vector -Support .(SVM) أقرب-K- الجيران (KNN) - شجرة القرار .(DT) يتم استخراج السمات والتصنيف بواسطة برنامج المحاكاة: MATLAB من MathWorks. يتم استخدام قاعدة بيانات -COVID TOS في تجاربنا. مقاييس الأداء المستخدمة في هذه الدراسة هي الدقة والحساسية والخصوصية ، ودرجة F1 ، وخصائص تشغيل جهاز الاستقبال. (ROC) .

الكلمات المفتاحية--- الكلمات المفتاحية: الصوت ، COVID-19 ، استخراج الميزات ، المصنفات ، .

Apprentissage automatique pour détecter le covid-19 à l'aide sons de toux

Resume

L'objectif de ce projet est de concevoir un système d'aide au diagnostic pour la détection précoce de la maladie COVID-19 à partir de la voix. Le système proposé repose principalement sur deux étapes : l'extraction de caractéristiques du son et la classification. Dans ce cas, nous avons choisi le coefficient cepstral de fréquence Mel (MFCC). Le processus de classification est basé sur trois classificateurs supervisés d'apprentissage automatique : -Machine à vecteur de support (SVM) -Voisins les plus proches (KNN) - Arbre de décision (DT). Notre système proposé a été évalué en utilisant TOS. Les performances utilisées de notre système sont la précision, la sensibilité et la spécificité, le score F1 et les caractéristiques de fonctionnement du récepteur (ROC).

- Mots-clés: voix,COVID-19,extraction de fonctionnalités, classificateurs,Matlab,simulation.

ACKNOWLEDGMENTS

First of all, we thank **ALLAH**

who helped us and provided us with patience and courage during these years of study.

I would like to thank my distinguished Dr. **KHALED BENSID**

who kindly supervises this theme were dedicated and did not spare us his valuable advice and guidance.

my sincere thanks to dr. **ELMOUNDHER HEDJAIDJI**

for his patience, his efforts and his contribution to the development of this research and achieve good results in our study, as well as the success of this great academic year .

We would like to thank the members of the jury MS. CHARIF FELLA and MR. YOUCEFA ABDELMADJID for agreeing to review and evaluate our modest work. We would also like to express our sincere thanks to Dr. HASSANI Mohamed Fouad and Dr.BENLAMOUDI Azzeddine and Dr.MEKIMAH Boualam for all advice on these years of study We take this opportunity to express our gratitude to all our teachers who contributed through their collaboration, availability and sympathy, during our teaching.

DEDICATIONS

We dedicate this modest work as a testimony of affectation, of admiration :

To My Family

My beloved woman who have meant and continued to mean so much to me,who's no longer in this world but her memories continue to regulate my life, who taught me the value of heard work; My grandmother(God rest her and put her in paradise).
to the light of my life my lovely mother who showed me with the indefatigable care, sincere prayers, endless love and unparalleled tenderness who that was pushing further to complete this research. To my father who helped me to chase my dreams and taught me that. to the candle of the house my sister and my brothers .

Dear friends

who helped me in building ideas and editing my many mistakes, who motivate me by sharing words, advice and continuous encouragement with a perfect blend insight and humour through my years of study and through the process of researching and writing this productive and enjoyable thesis.

To The second Family

1001db

Brahim Elkhalil

DEDICATIONS

We dedicate this modest work as a testimony of affectation, of admiration :

To My Familly

With great pleasure, an open heart and immense joy, I dedicate my work to my dearest mother: ”You gave me life, tenderness and courage to succeed. All that I can offer you will not be able to express the love and gratitude I have for you. As a testimony, I offer you this modest work to thank you for your sacrifices and for the affection with which you have always surrounded me”. I also dedicate this work to the memory of my dear late father and grandmother may God have their souls and welcome them in His Vast Paradise. My dedications are also addressed to my dearest sisters Manel and Narimene. I dedicate this work to : my dear grandmother uncles, aunts and My dear cousins

Dear friend

The solid shoulders, the attentive, understanding eye and the persons most worthy of our esteem and respect. No dedication can express our feelings, may God preserve you and give you health and long life with health.

Mohammed Larbi

List of abbreviation

AI:	Artificial Intelligence
ANN:	Artificial Neural Networks
AUC:	Area Under Curve
CNN:	Convolutional neural network
Covid:	Corona virus
CPU:	central processing unit
CWT:	Continuous Wavelet Transform
DBN:	Deep belief network
DCT:	Discrete cosine transform
DNN:	Deep neural network
DSS:	Decision Support Systems
DT:	Decision Tree
F0:	Fundamental Frequency
FFT:	Fast Fourier Transform
FN:	False negative
FP:	False positive
GMM:	Gaussian mixture model
GPU:	Graphical processing unit
GRNN:	eneralized regression neural network
HC:	Healthy Controls

HMM:	Hidden Markov model
HNR:	Harmonic to Noise Ratio
KELM:	Kernel-based Extreme Learning Machine
KNN:	K-Nearest Neighbors
LDA:	Linear didcriminant analysis
LPC:	Linear Predictive Coding
LPCC:	Linear Predictive Cepstral Coefficients
MFCC:	Mel Frequency Cepstral Coefficients
ML:	Machine Learning
PCA:	principal component analysis
PLP:	The Perceptual Linear Prediction
RAM:	Random access memory
RAT:	Rapid Antigen Test
RASTA:	Relative Spectral Transform
RBF:	Radial Basis Function
ROC:	Receiver Operating Characteristic
SVM:	Support Vector Machine
TEO:	Teager energy operator
TN:	True Negative
TP:	True Positive
UBM:	Universal background model
WHO:	World Health Organization
ZCR:	Zero-crossing rate

Contents

List of tables

List of figures

General Introduction

Introduction

Voice communication is an effective element in human life,as it is an essential skill in a person's daily and professional life. People with speech disorders have many difficulties in their daily lives because voice problems are directly related to respiratory diseases, such as COVID-19.

COVID-19, otherwise known as the coronavirus has been declared as a global pandemic by the World Health Organization (WHO) and has rapidly spread over more than 200 countries Worldline [5].

It is caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). The main symptoms include fever, dry cough, sore throat, dyspnea, fatigue, headache, and in severe cases multiple organ failure [[6],[7], [8]].

Implicitly, voice is also affected resulting in a lack of energy to produce sound and loss of voice caused by shortness of breath and upper airway congestion. Recurrent dry coughs can further influence changes in vocal cords affecting voice quality. A recent study has reported changes in the acoustic parameters of voice caused by the insufficient airflow through the vocal tract as a consequence of pulmonary and laryngological involvements in people with COVID-19 [9]. Therefore, all these respiratory conditions caused by COVID-19 can make patients' voices distinctive, creating identifiable voice signatures. With the onset of COVID-19, various efforts were made to develop efficient solutions for automatic diagnosis that would supplement the standard testing methods.

While the Real-time Reverse Transcription-quantitative Polymerase Chain Reaction (RT-qPCR) test serves as the gold standard for COVID-19 detection, it requires an in-person visit to a hospital or laboratory for taking an upper respiratory specimen (nasopharyngeal and oropharyngeal swab). Local clinics may also lack RT-qPCR facilities and skilled staff, requiring transport of specimens and further delaying test results. The Rapid Antigen Test (RAT) is an alternative that does not require laboratory processing and alleviates the time constraint of RT-qPCR, but unfortunately its sensitivity decreases with lower viral loads, thus providing false-negative result [10].

In order to try to get rid of the traditional techniques, to work on the optimal economy of money and time, and on the comfort of the patient, we present a new proposal for the detection of vocal disorders, whereby this approach will rely on signal processing techniques.

Contribution

We contribute to design a diagnostic aid system that allows the early detection of the COVID-19 ,specifically, we are interested in the distinction between COVID-19 subject patients and healthy control voices.

The approach we proposed is based on two main steps: attribute extraction of sound and classification. The feature extraction we have chosen is Mel-Frequency Cepstral Coefficient (MFCC). The classification process is based on three machine learning supervised classifiers: -Support Vector Machine (SVM) -K-nearest neighbors (KNN) - Decision tree (DT). The extraction of the attributes and the classification are performed by simulation software: the MATLAB from MathWorks. The TOS- COVID database is used in our experiments. The performance measures used in this study are the accuracy, sensitivity, and specificity, F1 score, and Receiver Operating Characteristics (ROC).

Memory organization

Our work is organized as follows :

The first chapter: Introduce the voice production anatomy by the voice disorders, their causes, and their different types where we present the COVID-19 disease. Followed by the voice characteristics.

The second chapter: is dedicated to the proposed method.

The third chapter: is dedicated to the experimental results and discussion.

Voice Generalities

I.1 introduction

Voice is one of the most important characteristics of a human, so voice disorders can affect anyone, but are more common in people who use voice extensively in their daily or professional lives. For example teachers or speakers like journalists.
This raises the question of whether these diseases can be detected by artificial intelligence, and the reliance on sound opens up many horizons for the diagnosis of many diseases that directly or indirectly affect the human voice.
In this chapter, starting from the first part, we try to elucidate the mechanisms of sound production and determine their structure. In the second part, we will explore the types of speech disorders through a few examples. finally, we see voice parameters.

I.2 Definition and Anatomy of the Vocal Tract

Our vocal tract, although a relatively small part of the human anatomy, is an amazing part of our body. The vocal tract is the cavity found in humans that is responsible for producing sounds, without which we couldn't speak! Many scientists believe that our ability to communicate in such a sophisticated way sets us far apart from other mammals. So, how does the vocal tract work? .
However, its function and provide a diagram to help you understand how our bodies produce a fundamental process: human speech.
Moreover, identifying the major parts of the vocal tract. In humans, this means the oral cavity, the nasal cavity, the larynx, and the pharynx. Each of these four components is composed of smaller components within, and we need all of these pieces to produce sound.
the easiest part of our anatomy lesson, and also the least complicated. The oral cavity is your mouth, lips, teeth, and cheeks. The nasal cavity is the space behind your nose. The intricate details of these cavities are beyond the scope of this lesson but just remember those basics. Next, let's talk about some of the more complicated pieces of anatomy [11].

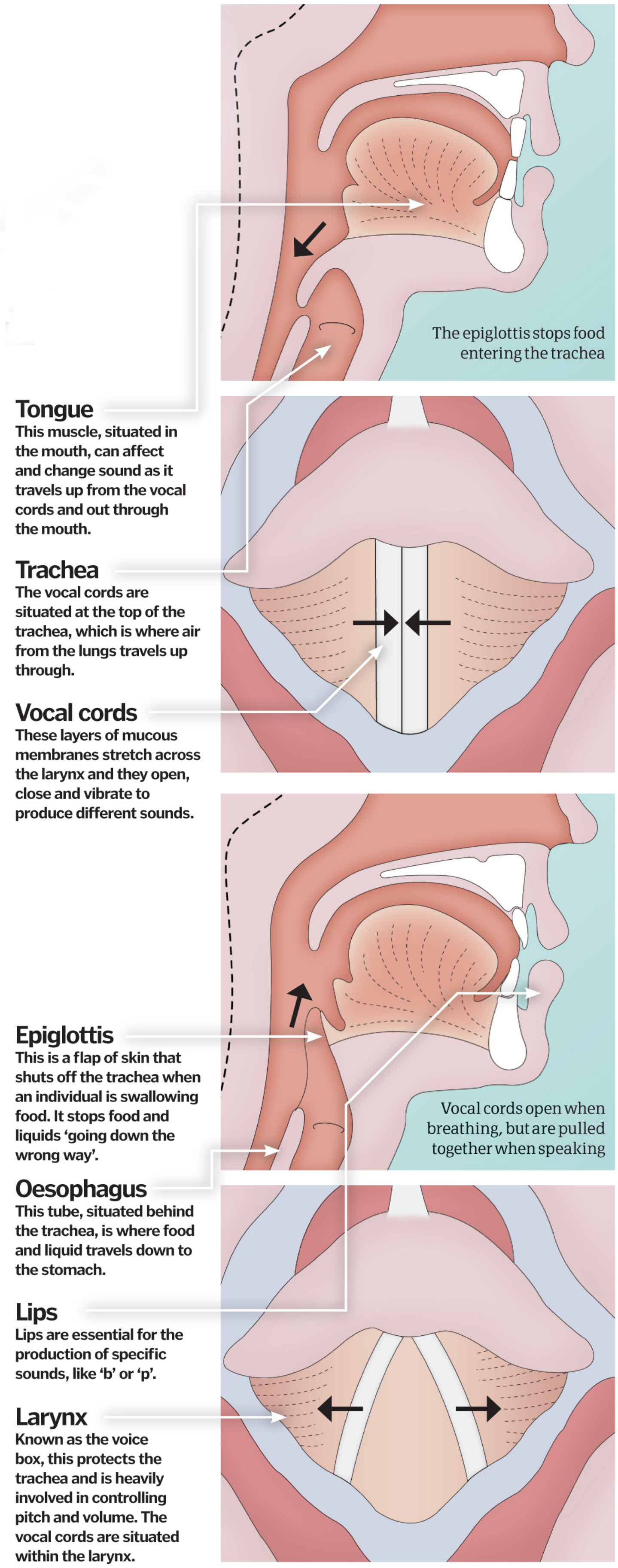

Figure 1: Anatomy of vocal chords

I.2.a The Pharynx:

The pharynx is a flowery manner of pronouncing the top part of your throat. The pharynx is a hollow space that connects the nasal and oral cavities to the larynx, which we‘ll communicate approximately next. As you will see, the larynx is what produces sound, inside the shape of vibration, however, you could think about the pharynx as amplifying or resonating this sound. The form of the throat, oral, and nasal cavities adjustments the vibration sounds produced through the larynx into sounds people understand. The pharynx branches off into two parts:

-The esophagus, which goes down into your stomach.

-The trachea, which goes down into your lungs.

As a result of this, the pharynx is very important in digestion and respiration [12].

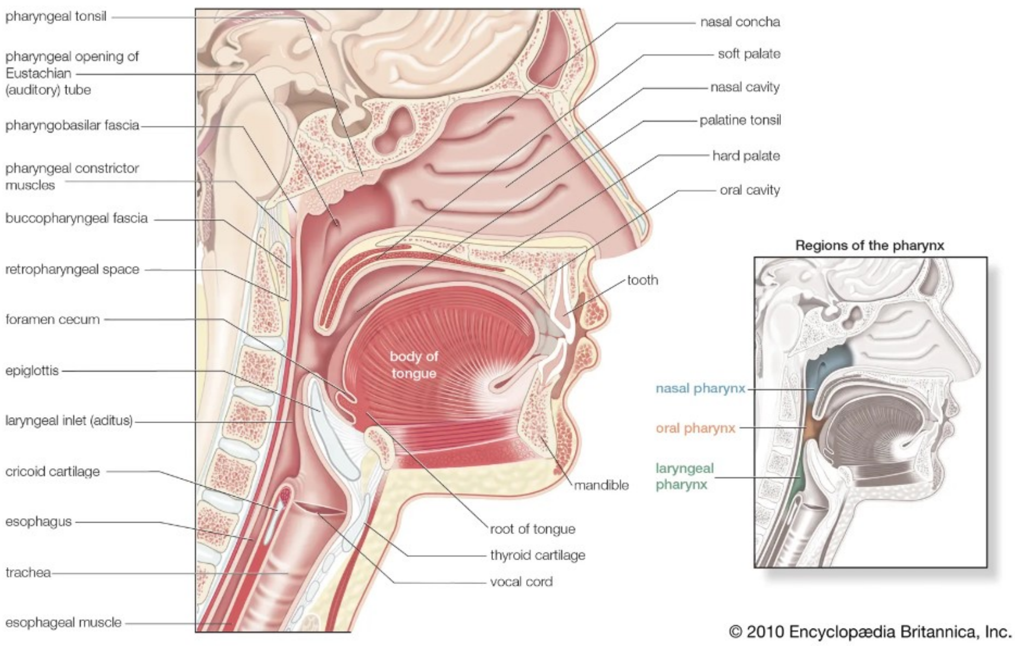

Figure 2: Anatomy of pharynx

I.2.b The Larynx:

The larynx or, as you can have heard of it noted as, your ‘voice box‘, is a muscular organ that holds your vocal cords. It additionally serves to shape an air passage right down to your lungs.

The larynx is made from bone and cartilage and is located on the pinnacle of your trachea (your windpipe), which connects down to your lungs.

Also internal of the larynx are your vocal cords, additionally referred to as vocal folds for the manner they may be shaped. This is what's chargeable for generating your voice. When those flaps of tissue vibrate in opposition to one another, the sound is produced! Ever have a problem speaking due to laryngitis? This is because of an irritation of your vocal cords, which makes speaking hard and painful[13].

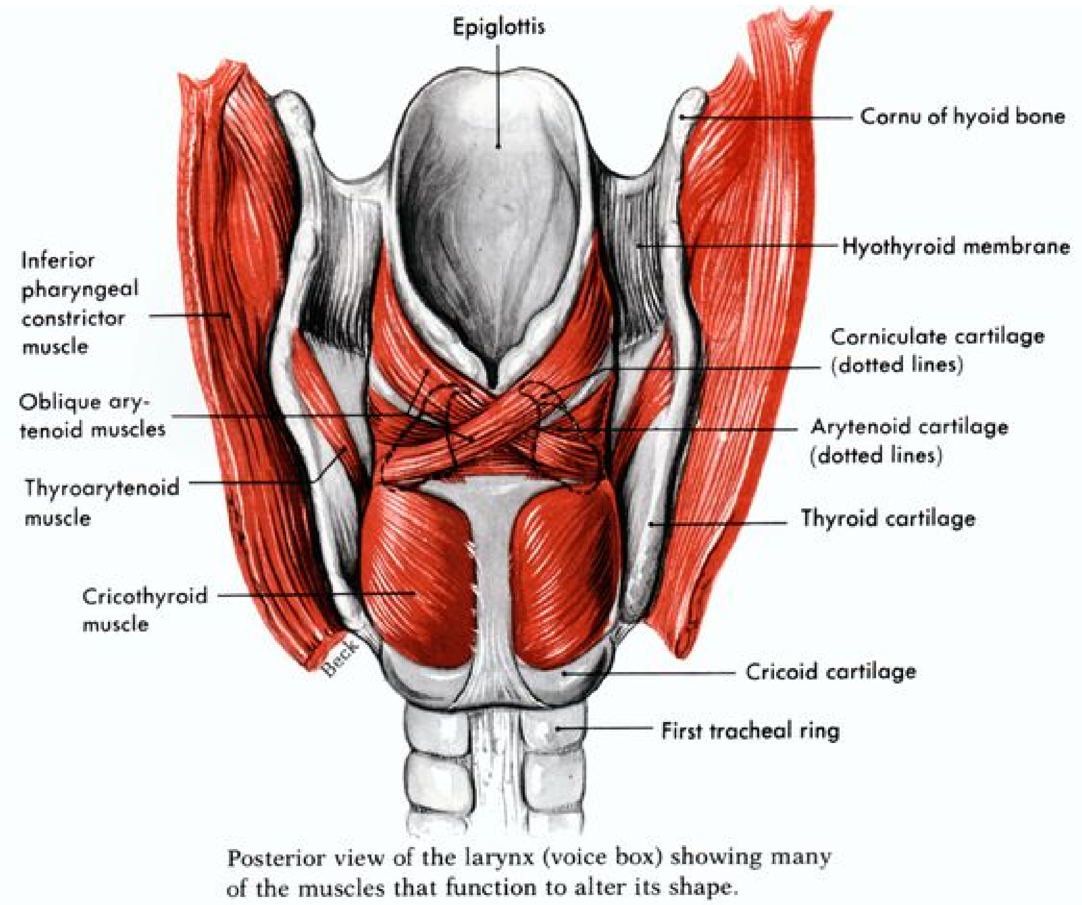

Figure 3: Anatomy of Larynx

I.3 Voice parameters

Acoustic analysis completes the perceptive evaluation of the vocal function and aims at a better approach to diagnosis. In addition, acoustic analysis of speech is a non-invasive method of quantifying speech function. It includes the study of the following acoustic parameters:

I.3.1 Intensity

Intensity is the auditory sensation that allows us to distinguish between a loud and a soft sound. a weak sound. It is expressed in decibels, which also vary in speech. by the air pressure under the glottis [14].

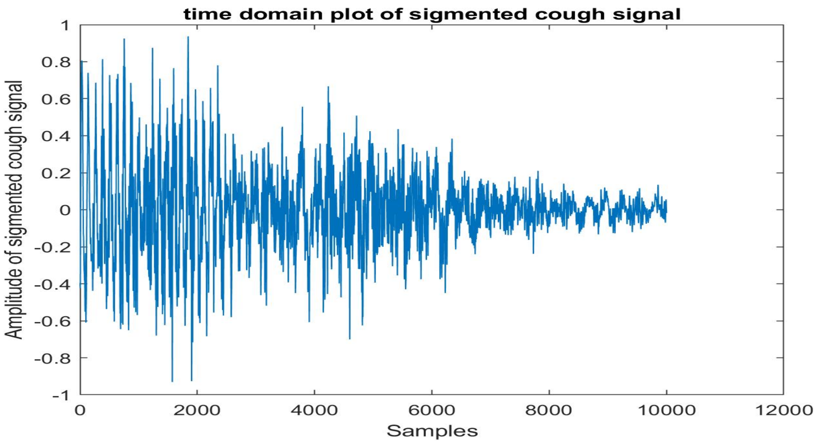

Figure 4: Intensity of voice

I.3.2 Pitch

Pitch or frequency or tonality is the auditory sensation linked to the frequency of vibration of a sound. It is expressed in Hertz. The fundamental frequency is the number of opening-closing cycles of the vocal cords per second, giving a low, medium, or high voice naturally. In the speech, the pitch of the voice varies permanently, it is the intonation, which when intonation, which when it is rich and confers aesthetic and communication qualities to the voice communication [14].

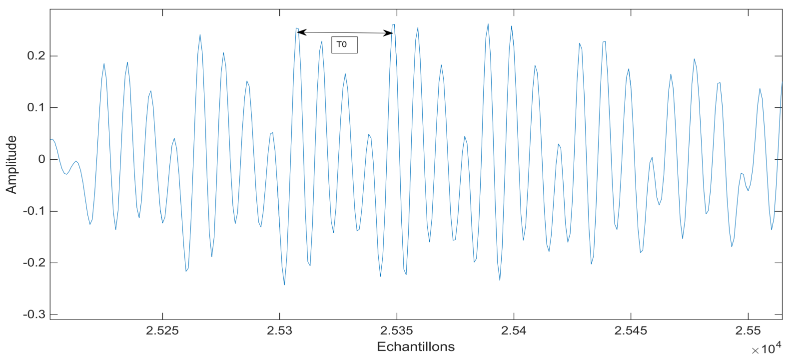

Figure 5: Pitch

I.3.3 Timbre

The timbre is a clue to identifying a voice and recognizing it. The mobility of the palate and tongue, and the tone of the lips and cheeks play an important role in its richness, as well as the quality of the mucous membrane that lines the resonance cavities [14]

I.4 Voice disorder

Language is a complex process. In humans, you push air out of the lungs and through two layers of tissue (vocal cords) in the larynx (voice). The air causes vibrations in the vocal cords that make sounds when touched. If there is a problem with the pitch, tone, or volume of your voice, it means your vocal cords are not vibrating properly - this is called a vocal disturbance. Depending on the cause, it can be temporary or permanent[15] [16].

I.4.1 Organic

I.4.1.a structural

Structural disorders result from some damage to the larynx (physical abnormalities).

- Cysts.
- Granuloma.
- Hemorrhage.
- Hyperkeratosis.
- Laryngitis.
- Leukoplakia.
- Nodules (nodes).
- Papilloma.
- Polyps.
- Trauma.
- Miscellaneous growths.
- **Covid-19.**.

I.4.1.b Neurological

Neurological speech disorders are caused by some problems with the nervous system when interacting with the larynx. Simply put, two nerves flow from the brain to the larynx and control the movement of the larynx. The circulatory laryngeal nerve, the most important of the two nerves, exits the brain, surrounds the aorta, returns upward, and attaches to the left larynx. Because the circulatory laryngeal nerve is located in the neck, it is susceptible to damage during heart, lung, spine, and thyroid surgery. Damage to the nerve causes paralysis (weakness) or paralysis (lack of complete movement) of the vocal cords on the side of the lesion. Other neurological voice disorders are associated with different types of problems in the central nervous system[17].
-Paralysis/Paresis.
-Spasmodic Dysphonia (Laryngeal Dystonia).
-Tremor (Benign Essential Tremor).
-Parkinson.
-Voice problem caused by another neurological disorder (e.g. Parkinson's disease, myasthenia gravis, ALS/Lou Gherig's Disease).

I.4.2 Functional

Functional problems are due to poor muscle function. All helpful problems fall under the category of muscular anxiety dysphonia. The unusual problems listed here indicate unusual types of muscle anxiety. Remember that there may be a disease where there is no choice but to sound, but the use of sound can cause additional exertion, discomfort or fatigue[18].
- Muscle tension dysphonia (general).
- Anterior-posterior constriction.
- Hyperabduction.
- Pharyngeal constriction.
- Ventricular phonation.
- Vocal fold bowing.

I.4.3 Psychogenic

Psychogenic disorders exist because the voice can be impaired for psychological reasons. In this case, there is no structural reason for the speech disorder, and there may or may not be a pattern of muscle tension. A psychological or emotional component is very common in speech disorders, but speech disorders due to psychological disorders are relatively rare. The two most common types of psychogenic disorders are:
- Conversion dysphonia or aphonia.
- Puberphonia (mutational falsetto).

I.4.4 Coronavirus

Coronaviruses are enveloped RNA viruses [19]. These RNA viruses are divided into five branches and coronaviruses are most likely related to branch 2 [20]. Coronaviruses can also be classified as animal, and human coronaviruses in respect to their host targets and the viruses in this subfamily Coronavirinae are divided into four genera, one of which is the beta-coronavirus [21]. There are four different lineages for beta-coronavirus: 2a, 2b, 2c and 2d. SARS-CoV and SARS-CoV-2 belong to 2b lineage and MERS–CoV belongs to 2c lineage [22]. The process of infecting humans through animals by SARS-CoV, MERS-CoV and SARS CoV-2 is still unknown. As 96.2% gene of SARS CoV-2 is analogous to bat's genome so bats are considered as the natural source of virus transformation and it is called bat-to-man SARS-CoV-2 transformation [23].

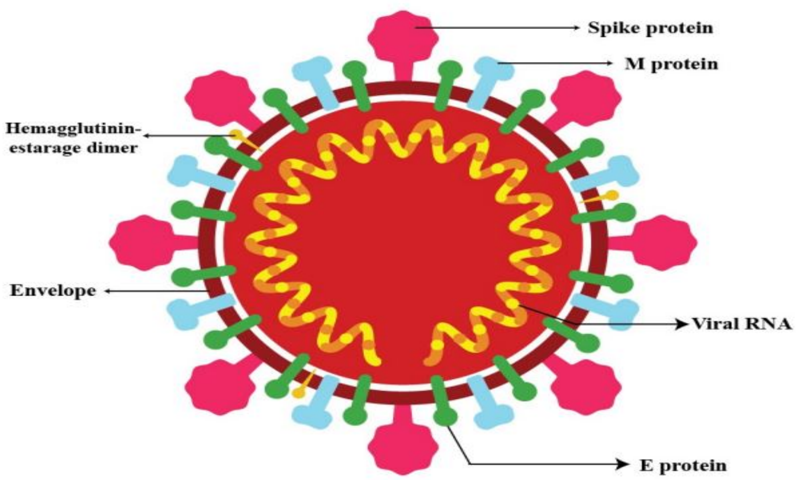

Figure 6: The schematic representation of coronavirus[1].

I.5 Covid-19

Coronavirus is one of the pathologies that has killed humans in the last 3 years, in this work we focus on and try to find a solution to detect it with the lowest possible cost and the fastest time for diagnosis. According to the government of China, the initial case was found on 17th November 2019 with a geriatric patient in the province of Hubei. Besides that, four male and five female infected patients were found in the same month. Coronavirus cases in Hubei steadily increased and crossed two hundred cases, some with fatalities, by the end of the year 2019 [24]. In the last week of December of 2019, an outbreak of strange pneumonia characterized by some unusual features such as dry cough, fatigue, fever and infrequent gastrointestinal symptoms occurred in the Huanan Seafood Wholesale Market in Wuhan city of Hubei province of China. The market was closed on the first day of 2020 after the pronouncement of an epidemiological alert by the neighboring authority [25]. At that time, Wuhan Central Hospital sent a broncho-alveolar lavage fluid (BALF) sample from an unknown clinical case to a sequencing company, and after completing the sequencing work, that institution informed the hospital that a new coronavirus was found in the test.

In the last week of the previous year, a test result addressed to that hospital stated a positive result for SARS, causing some doctors of the hospital to apprise their staffs and relevant hospital authorities of the result. After that, the Wuhan Municipal Health Commission gave notice to some medical institutions on the treatment process of pneumonia of unknown cause. During the initial stages of the outbreak, the number of cases doubled approximately every seven and a half days and the spread of the infection was almost entirely driven by human transmission [26].

Within early and mid-January of the current year, the virus spread to other Chinese provinces besides Hubei. In China, more than one hundred new cases were being detected per day after mid-January, including two people in Beijing and one in Shenzhen. Many people were found with the same signs and symptoms, although the numbers of tests were not adequate to see the real image. Within the same month, the WHO declared the coronavirus an international concern and by this time, the outbreak spread by a factor of one hundred to two hundred [27]. The pathogen of the outbreak was later identified as a novel beta-coronavirus by several studies and named as 2019 novel coronavirus (2019-nCoV) that reminded the scientists about the past pandemic (SARS-2003, caused by another beta-coronavirus) that occurred seventeen years ago [25].

Furthermore, the disease spread to other countries before the beginning of February 2020 and before the ending of the first week of that month, the world saw more than five hundred deaths involving at least twenty-five countries, as documented by the WHO. On the last day of January, Italy had its first confirmed cases who were two tourists from China. Before mid-March, the WHO considered Europe an active center of the pandemic, and on the 19th March, Italy overhauled China as the country with the highest deaths. Before the end of the same month, the United States overtook China and Italy with the most confirmed cases in the world. The majority of COVID-19 cases in the city of New York came from travelers from Europe, rather than directly from China or any other Asian country and in France, retesting of prior samples of the previous year found a person who had the virus [28]. In Bangladesh, transmission is still going on. In some regions, the second wave of this virus attack is being observed. According to the recent updates, as of the 21st of November, 58,283,585 cases were confirmed, with 1,383,651 deaths and 40,364,210 recovered cases [29]. **transmission of covid-19** Corona virus is transmitted in many ways, which greatly facilitated its spread around the world.

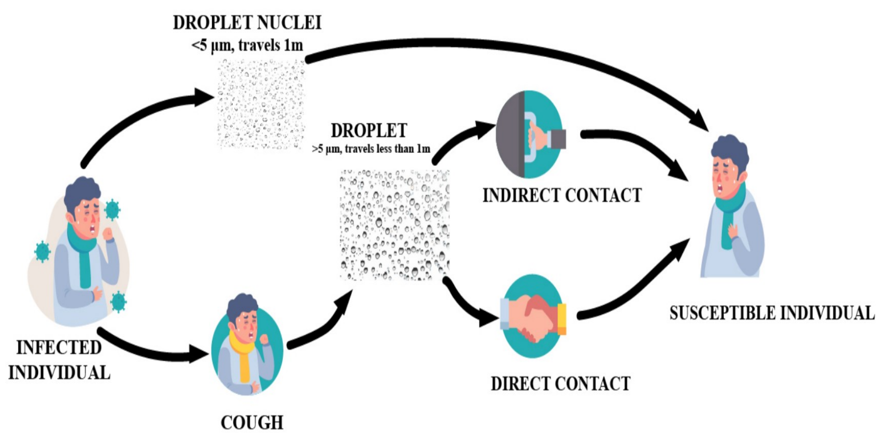

Figure 7: The schematic of Ways of transmission of covid-19[2].

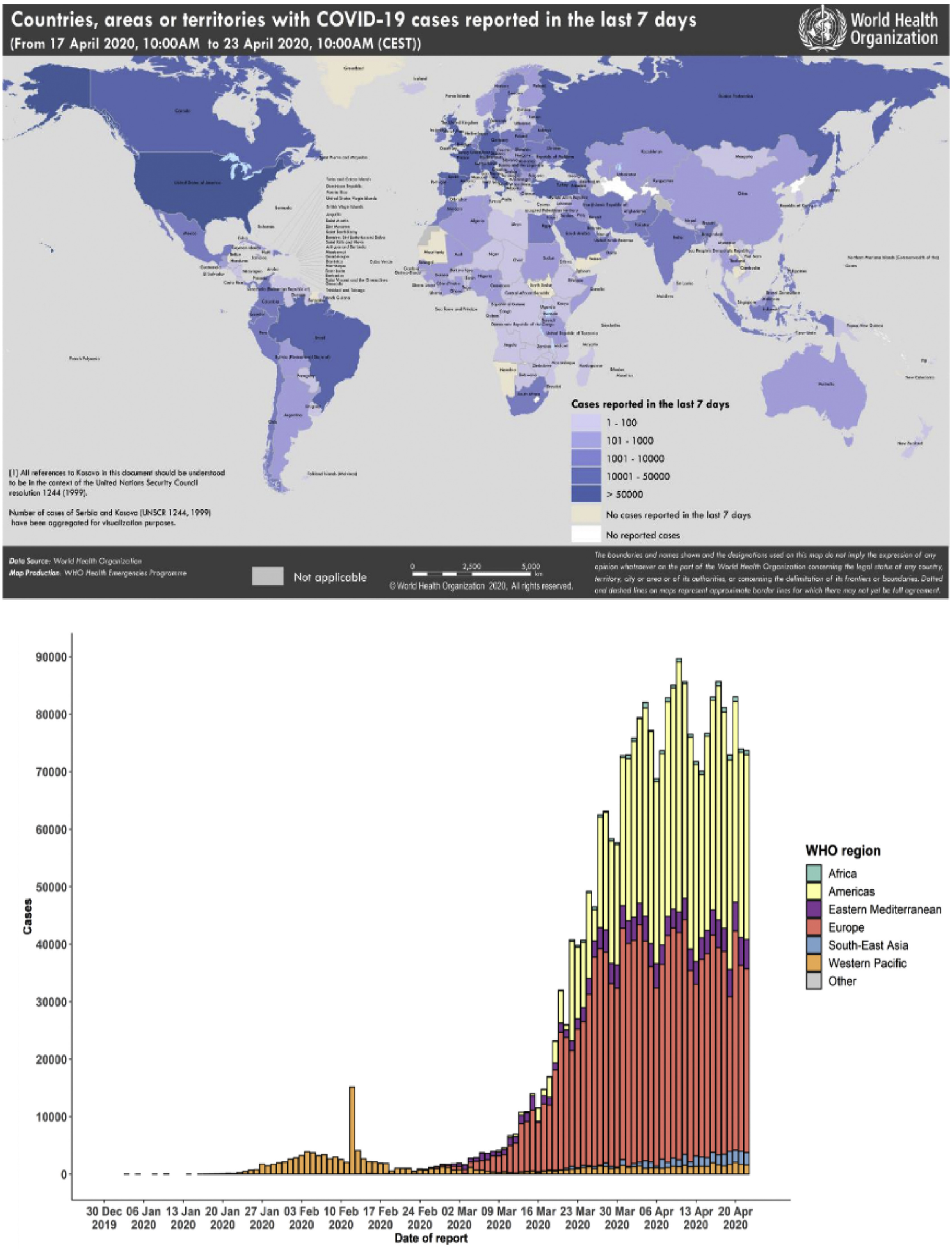

Figure 8: Epidemiology and of covid-19 in the world[3].

I.6 Conclusion

In this chapter, we have mentioned the definition and anatomy of the vocal tract as well as the voice disorders and their types, then we discussed covid-19 disease and his Epidemiology and how he can transmitted between people .Afterward, we established a study of acoustic parameters of the voice.
In the next chapter, we discussed the feature extractions and the classifiers that are used in our study for the earlier detection of covid-19 disease.

Artificial Intelligence, Machine learning, Features Extraction and Classification.

II.1 Introduction

Detection features have been widely used in medical detection system because it is more reliable and fast when compared to conventional methods. voice disorders are common diseases; most people have experienced in their life. Voice disorder sufferers usually do not seek medical consultation attributable of time-consuming and costly medical expenditure. Recently, researchers have proposed various machine learning algorithms for rapid detection of voice disorders based on the analysis of the human voice[30].
The rest of the chapter is organized as follows: Section 2 describes the Artificial intelligence and machine learning . Section 3 describes the feature extraction method and Classifiers. Finally, we conclude this chapter with a conclusion.

II.2 Artificial intelligence

Like machine learning and deep learning, AI isn't "new," but it's definitely undergoing some sort of renaissance. The way people use the word is also changing, much to the chagrin of traditionalists. When Turing first developed his test, the term artificial intelligence mainly referred to a technology that could largely mimic human intelligence. So time travel today is a distant, futuristic thing for us. (It took 60 years, but in 2014, a computer finally passed the Turing test.) Chatbots like Microsoft's short-lived "Tay" are synonymous with many AI technologies.
Today, the term "artificial intelligence" or just "AI" is often used to refer to any kind of machine learning program. In this regard, it is starting to replace "big data" and its followers, "advanced analytics" and "predictive analytics". For those of you who hate the term "big data", this might be a good thing. But some prefer to keep the term AI as a narrowly defined thing that replicates many aspects of human intelligence and becomes a separate entity. We haven't gotten to that stage yet, and we probably never will, although that's probably not the bet you want to bet. A year ago, Facebook CEO Mark Zuckerberg predicted that we were five to 10 years away from developing "artificial intelligence that can truly understand the meaning of content[31].

II.3 Machine learning

machine learning refers to any type of computer program that can “learn” by itself without having to be explicitly programmed by a human.

The phrase (and its underlying idea) originated decades ago - all the way back to Alan Turing’s seminal 1950 paper ”Computers and Intelligence,” which included a section on his famous ”learning machine,” It can fool humans into believing it is real. Today, machine learning is a widely used term that covers many types of programs you will encounter in big data analysis and data mining. Ultimately, the ”brains” that power most predictive programs, including spam filters, product recommendations, and fraud detectors, are machine learning algorithms.

Data scientists should be familiar with the difference between supervised and unsupervised machine learning, as well as ensemble modeling using a combination of method techniques and semi-supervised learning combining supervised and unsupervised methods.[32]

there are three types of machine learning :

II.3.1 Supervised machine learning

In supervised learning, the machine is taught by example. The operator provides the machine learning algorithm with a known dataset that includes desired inputs and outputs, and the algorithm must find a method to determine how to arrive at those inputs and outputs. While the operator knows the correct answers to the problem, the algorithm identifies patterns in data, learns from observations and makes predictions. The algorithm makes predictions and is corrected by the operator – and this process continues until the algorithm achieves a high level of accuracy/performance.

Under the umbrella of supervised learning fall: Classification, Regression and Forecasting.

II.3.1.a Classification: In classification tasks, the machine learning program must draw a conclusion from observed values and determine to what category new observations belong. For example, when filtering emails as ‘spam’ or ‘not spam’, the program must look at existing observational data and filter the emails accordingly.

II.3.1.b Regression:

In regression tasks, the machine learning program must estimate – and understand – the relationships among variables. Regression analysis focuses on one dependent variable and a series of other changing variables – making it particularly useful for prediction and forecasting.

II.3.1.c Forecasting:

Forecasting is the process of making predictions about the future based on the past and present data, and is commonly used to analyse trends[33][34].

II.3.2 Unsupervised machine learning

Here, the machine learning algorithm studies data to identify patterns. There is no answer key or human operator to provide instruction. Instead, the machine determines the correlations and relationships by analysing available data. In an unsupervised learning process, the machine learning algorithm is left to interpret large data sets and address that data accordingly. The algorithm tries to organise that data in some way to describe its structure. This might mean grouping the data into clusters or arranging it in a way that looks more organised. As it assesses more data, its ability to make decisions on that data gradually improves and becomes more refined. Under the umbrella of unsupervised learning, fall:

II.3.2.a Clustering:

Clustering involves grouping sets of similar data (based on defined criteria). It's useful for segmenting data into several groups and performing analysis on each data set to find patterns.

II.3.2.b Dimension reduction:

Dimension reduction reduces the number of variables being considered to find the exact information required[33][34].

II.3.3 Reinforcement learning

Reinforcement learning focuses on regimented learning processes, where a machine learning algorithm is provided with a set of actions, parameters and end values. By defining the rules, the machine learning algorithm then tries to explore different options and possibilities, monitoring and evaluating each result to determine which one is optimal. Reinforcement learning teaches the machine trial and error. It learns from past experiences and begins to adapt its approach in response to the situation to achieve the best possible result [33] [34].

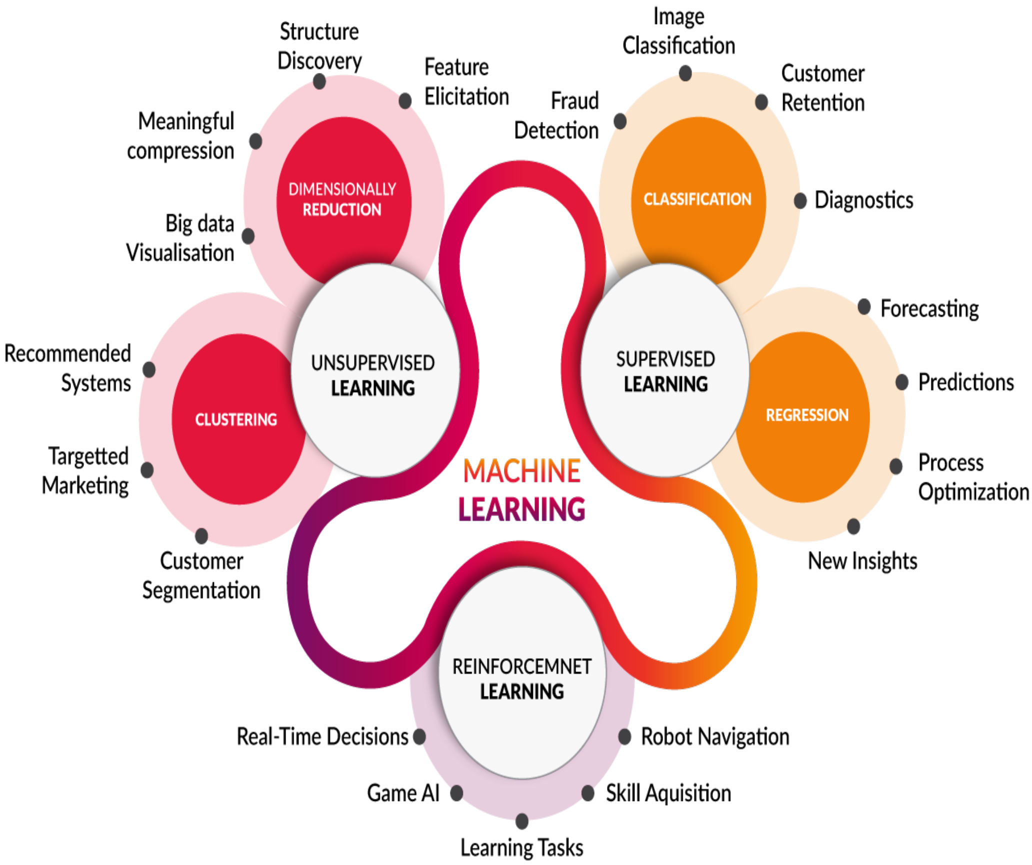

Figure 9: Machin Learning types.

II.4 Feature Extraction:

The Evolution of Audio Features: Simply Put, Feature Extraction highlight the most dominant and discriminatory processes properties of the signal. Appropriate feature imitation properties A very compact signal. Evolution of Audio Signals The function description is shown in Figure below. Development of audio functions can be divided into time domain, frequency domain, joint Time-frequency domain and depth features. oldest and simplest features are extracted from the time domain. Time-domain features were developed until the late 1950s [35],[36],[37].

So far, the time domain features play an important role in audio analysis and classification. Various functions are available for analyzing the frequency spectrum of an audio signal.For example, pitch, formants, etc. are developed from the frequency domain It is used in various applications to date. development of The frequency domain features were around the 1950s to the 1960s [38],[39]. In the late 1960s, common temporal frequencies [40],[41].

Feature extraction is required for better biomedical signal classification performance. The goal of feature extraction is to find the most compact and informative set of features (distinguishable patterns) to improve the efficiency of the classifier. Moreover, feature extraction is employed to extract features from the original signal to accomplish reliable classification. Feature extraction is the most important part of the classification of biomedical signals because the classification performance may suffer if the features are not selected properly.[42]

Figure 10: Evolution of Audio Feature Extraction

We mention here the most famous algorithms developed for features extracting:

- Spectrogram.
- Formant .
- Wavelet analysis.
- The linear predictive coding (LPC).
- Perceptual linear prediction (PLP) .
- Rasta- Perceptual linear prediction.
- Jitter .

- Shimer .
- NNE, GNR, HNR and CHNR.
- The zero-crossing rate (ZCR).
- The linear frequency cepstral coefficient (LFCC).
- The Teager energy operator (TEO).
- **Mel frequency cepstral coefficient (MFCC).**

II.4.1 Mel frequency cepstral coefficient

Extracting optimal parametric representations of acoustic signals is an important task in creating better parametric representations Identify performance. The efficiency of this stage is important for the next stage because it affects its behavior. MFCC is based on human auditory perception and cannot perceive frequencies above 1 kHz. in other words, The MFCC is based on the known variation of the human ear's critical bandwidth with frequency. There are two types of MFCCs The filter is linear at low frequencies below 1000 Hz and logarithmically spaced above 1000 Hz. One Subjective pitch occurs in the Mel frequency range to capture important features of speech in speech.
Mel-Frequency Cepstrum Coefficients (MFCCs) have been very popular in the field of speech processing . MFCCs are a type of cepstral representation of the signal, where the frequency bands are distributed according to the Mel-scale, instead of the linearly spaced approach. In order to extract MFCCs from a frame .Compute Mel-Frequency Cepstral Coefficients (MFCC) from the spectrum of a small window Speech signal obtained by Fast Fourier Transform (FFT) of the signal. get one Approximation of human ear perception to sound frequency and spectrum Subject to a series of triangular filters uniformly distributed over the Mel-frequency range. by discrete cosine, the transform is applied to the output of the filter, and the coefficients are determined in Mel frequencies[43].

II.4.1.a A/D conversion:

In this step, we will convert our audio signal from analog to digital format with a sampling frequency.

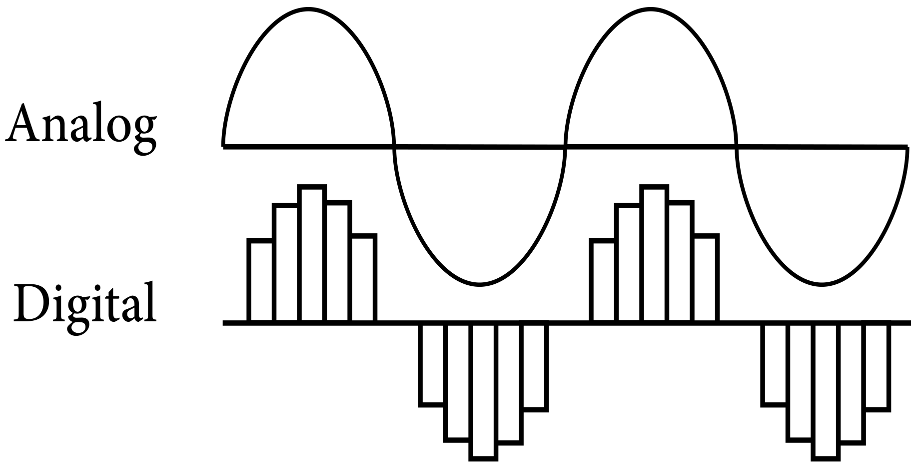

Figure 11: Conversion A/D

II.4.1.b Preemphasis

Pre-emphasis refers to a systematic process designed to increase in-band size The magnitude of some (usually higher) frequencies relative to other (usually lower) frequencies Improve the overall signal-to-noise ratio. So this step handles the process of passing the signal through a filter that Emphasize higher frequencies. This process adds higher frequency signal energy[44].

II.4.1.c Framing

The process of dividing the speech samples received from the ADC into small frames of the length contained in them in the range of 20 to 40 milliseconds. The speech signal is divided into frames of N samples. Adjacent frames will be Separated by M ($M < N$)[44].

II.4.1.d Windowing

MFCC technology is designed to develop features of audio signals that can be used to capture sounds in speech. But there are many calls in a given audio signal, so we divide the audio signal into different segments.Use the Hamming window as the window shape by looking at the next block in the feature extraction processing chain and integrate all the nearest frequency lines. The Hamming window looks like this: $W(n)$, $0 \leq n \leq N-1$ where

N = number of samples in each frame

$Y[n]$ = Output signal

$X(n)$ = input signal

$W(n)$ = Hamming window, then the result of windowing signal is shown below [44]:

$$\mathrm{Y[n] = X(n) * \ W(n)} \tag{1}$$

II.4.1.e DFT(Discrete Fourier Transform):

We will convert the signal from the time domain to the frequency domain by applying the dft transform. For audio signals, analyzing in the frequency domain is easier than in the time domain. FFT is used to convert each frame of N samples from time domain to frequency domain. Fourier Transform is used to transform the convolution of the glottal impulse U[n] and the vocal tract impulse response H[n]. in the time domain. The statement supports the following [44] :

$$\mathrm{Y(w) = FFT\,[h(t)^*X(t)] = H(w)^*X(w)} \tag{2}$$

II.4.1.f Mel-Filter Bank: The frequencies range in FFT spectrum is very wide and voice signal does not follow the linear scale. Each filter's magnitude frequency response is triangular in shape and equal to unity at the Centre frequency and decrease linearly to zero at centre frequency of two adjacent filters.
Then, each filter output is the sum of its filtered spectral components. After that the following equation as shown below" is used to compute the Mel for given frequency f in HZ: [44].

$$\mathrm{mel}(f) = 2595 \ln\left(1 + \frac{f}{700}\right) \tag{3}$$

II.4.1.g Delta energy and delta spectrum

The voice signal and the frames changes, such as the slope of a formant at its transitions. Therefore, there is a need to add features related to the change in cepstral features over time. 13 delta or velocity features (12 cepstral features plus energy), and 39 features a double delta or acceleration feature are added. The energy in a frame for a signal x in a window from time sample t1 to time sample t2, is represented as shown below[44].

$$\text{Energy } = \sum_{t=t_1}^{t_2} X^2[t] \tag{4}$$

II.4.1.h Dynamic Features

In addition to these 13 features, the MFCC technique also considers the first and second derivatives of the features, resulting in another 26 features. Derivatives are calculated by taking the difference of these coefficients between samples of the audio signal and it helps to understand how the transformation happens. Where X[t] = signal , Each of the 13 delta features represents the change between frames corresponding to cepstral or energy feature,while each of the 39 double delta features represents the change between frames in the corresponding delta features[44].

II.4.2 Spectrogram

A speech waveform consists of a series of different events change over time. corresponding to this time-varying property Spectral properties that vary greatly over time. at last, A single Fourier transform cannot capture this speed Time-varying signals and STFTs are used instead [45].

STFT Consists of a single Fourier transform for part of the waveform under a sliding window. The spectrogram of the speech signal is then derived from the STFT.

$$S(\omega) = \left|X\left(m, \omega_k\right)\right|^2 \tag{5}$$

Spectrograms can be the distribution of power density over time and frequency, as shown in Figure 3. The power density distribution of the speech signal varies greatly with time and frequency and can be used to distinguish the normal and sick voice. It can also be seen from the figure that the power distribution of normal speech is uniform in time and frequency. However, it is also not uniform[46].

II.5 Classification methods

Feature extraction is the first step in any voice disorder detection system. In this step, the given voice signals are converted into representative acoustic features using various digital signal processing techniques. We will now discuss the most popularly used techniques for acoustic analysis and feature extraction in the related area and describe.

- **Support vector machine (SVM)** .

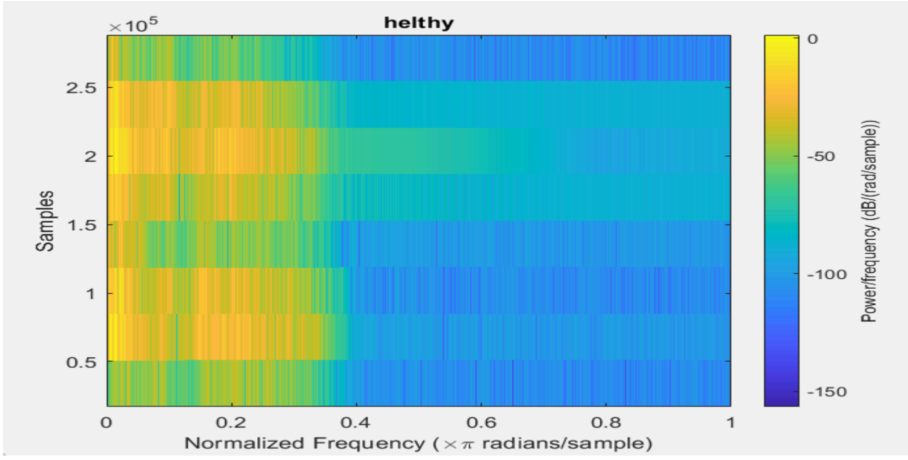

Figure 12: Spectrograms of Cough Covid-19.

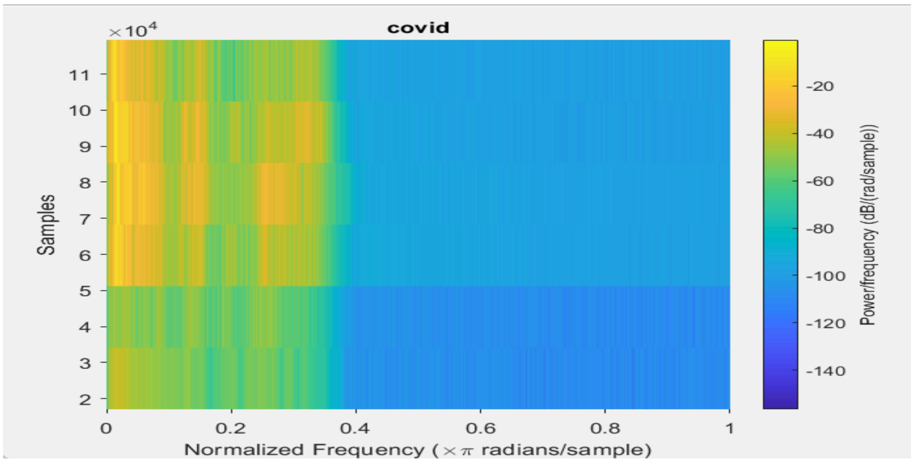

Figure 13: Spectrograms of Cough Covid-19.

- **K-nearst Neighbor (KNN)**.
- **Decision Tree**.
- Gaussian mixture model (GMM) .
- GMM-Universal background model (GMM-UBM).
- Artificial neural network (ANN).
- Hidden Markov model (HMM).
- Deep neural network (DNN).
- Convolutional neural network (CNN).
- Probabilistic neural network (PNN).
- Deep belief network (DBN).
- Generalized regression neural network (GRNN).
- Baysian classifier.
- The k-means clustering.

- The decision tree algorithm.
- Linear didcriminant analysis (LDA).

Recently, several approaches have been used in the field, in our detection system we use four different types (Support Vector Machine (SVM) with (Radial Basis Function (RBF), k-Nearest Neighbor (KNN), Decision Tree (DT)).

II.5.1 Support Vector Machine (SVM)

SVM is a machine learning algorithm that can solve problems such as classification, regression, and detection. A support vector machine is a discriminative technique, a supervised learning method used for classification and regression. It consists of separating two or more sets of points by a hyperplane. Depends on situation and configuration point. The original idea of SVM was based on the use of a kernel function, allowing optimal separation of planned points into different classes. The method uses a set of training data. This allows the hyperplane to separate the best points. We use a multi-class SVM. More formally, SVMs construct a hyperplane or set of hyper-planes in a high- or infinite-dimensional space that can be used for classification, regression, or other tasks such as outlier detection.

intuitively, a good separation is achieved by the hyperplane with the largest distance from the nearest training data point of any class (the so-called functional margin), since in general the larger the margin, the better the generalization error of the classifier smaller [47].

While the original problem can be formulated in a finite-dimensional space, it often happens that the sets to be distinguished are not linearly separable in that space. Therefore, it has been proposed to map the original finite-dimensional space to a higher-dimensional space, which may help the separation in this space. In order to keep the computational load reasonable, the mapping used by the SVM scheme is designed to ensure that the dot product of pairs of input data vectors can be easily computed from the variables in the original space by choosing a kernel function $k(x, y)$ to match the problem[48].

A hyperplane in a higher-dimensional space is defined as a set of points whose dot product with a vector in that space is constant, such a set of vectors is the set of orthogonal (and therefore minimal) vectors that define the hyperplane. The vector defining the hyperplane can be chosen as a linear combination with the parameter α_i of the image of the feature vector x_i appearing in the database. By choosing a hyperplane, the point x in the feature space that maps to the hyperplane is defined by the relation :

$$\sum_i \alpha_i k\left(x_i, x\right) = constant \tag{6}$$

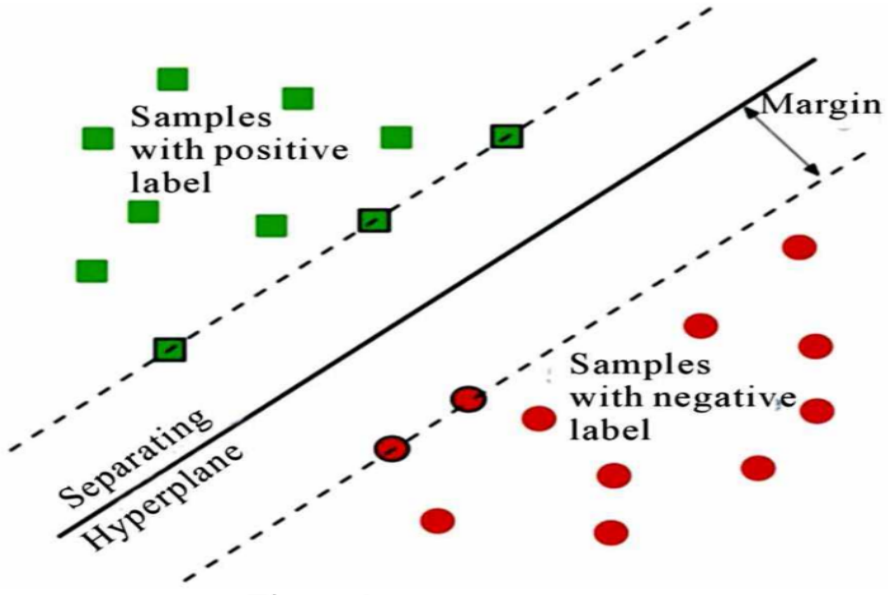

Figure 14: Concept of SVM[4].

We have two types of SVM classifiers:

II.5.1.a Linear SVM

is a classifier that is used for linearly separable data, which implies that if a dataset can be classified into two classes using a single straight line, it is called linearly separable data, and the classifier used is called Linear SVM.

II.5.1.b Non-linear SVM is used for non-linearly separated data, which implies that if a dataset can't be classified using a straight line, it's non-linear data, and the classifier employed is called Non-linear SVM.

II.5.2 K-nearst Neighbor(KNN)

K-Nearest Neighbor is a traditional supervised statistical pattern recognition method that classifies images by comparing the "K" value of the training data with the test data to find the closeness to the test image or data. The "K" value is estimated from feature extraction performed during training.

Use the principle of Euclidean equation in ANN classifier to identify similarity. It is used for classification and regression. In both cases, the input consists of the k closest training samples in the dataset. The output depends on whether KNN is used for classification or regression:

With KNN classification, the output is a class member. Objects are classified by multiple votes from their neighbors, and objects are assigned the most frequent class among their k nearest neighbors (K is a positive integer, usually small). if $k = 1$, then the object is simply assigned to the class of that single nearest neighbor. For KNN regression, the output is the attribute value of the object. This value is the average of the values of the k nearest neighbors. KNN is a type of classification where the function is only approximated locally and all computations are deferred until function evaluation. Because the algorithm relies on distance for classification when features represent different physical entities or are at different scales, normalizing the training data can significantly improve its accuracy.

For classification and regression, it can be a useful technique to assign weights to neighbors' contributions so that closer neighbors contribute more to the mean than farther neighbors. For example, a common weighting scheme is to give each neighbor a weight of 1/d, where d is the distance to the neighbor. Neighbors are taken from a set of objects whose classes (for KNN classification) or object attribute values (for KNN regression) are known. This can be seen as a training set for the algorithm, although no explicit training steps are required. A feature of the KNN algorithm is that it is sensitive to the local structure of the data.

such as fruits and vegetables Cereals can be measured by their crispness and Sweetness (Figure below). for display purpose There are only two features on the 2D plot employed. In fact, there can be any number of predictors, and the example can be extended to include any number of features. In general, fruit is sweeter than a vegetable. The grains are neither crunchy nor sweet. Our job is to determine which category the sweet potato falls into belonging. In this example, we have selected four next types of food is apples, green beans, lettuce, and corn. Because Vegetables got the most votes, sweet potatoes won the prize in the vegetable class. you can see key concepts KNN is easy to understand. In the above example, there are two important concepts. One of them is a way to calculate the distance between them Sweet potatoes and other foods. by default, The KNN function uses Euclidean distance, which can be calculated with the following equation [49],[50].

$$D_m(x_i, x_j) = \left(\sum_{i,j=1}^{n} |x_i - x_j|^m \right)^{1/m} \tag{7}$$

where p and q are the topics to be compared with n characteristics. There are other

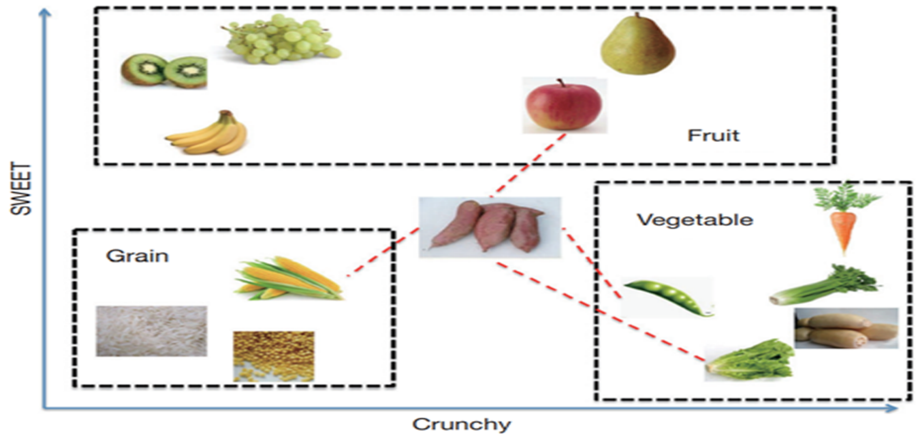

Figure 15: Concept of knn.

calculation methods The distance is the Manhattan distance [51],[52].

Another concept is to determine the parameter K How many neighbors are selected by the KNN algorithm. The appropriate choice of K has a significant impact on the Diagnostic performance of the KNN algorithm. capital k reduces the effect of variance caused by random errors, but at the risk of ignoring small but important patterns. That The key to choosing the right value of K is to find a balance between overfitting and underfitting [53].

some authors It is recommended to set K equal to the square root of the quantity observations in the training dataset [54].

II.5.3 Decision Tree

A decision tree is a classifier represented as a recursive partition of the instance space. A decision tree consists of nodes forming a rooted tree, This means it's a directed tree with a node called "Root" and no entry edge. All other nodes have only one incoming edge. node with outgoing edges is called internal or test nodes. All other nodes are called leaves (also called the end node or decision node). each internal node in a decision tree divides the instance space into two or more subspaces based on given conditions and enters a discrete function of attribute values. In the simplest and most common case, each test looks at one property, the space instance by attribute value. for numeric properties, a condition refers to a range. Each leaf is assigned to a class that represents the most appropriate target value.

Alternatively, a leaf may contain a probability vector that indicates the probability that the target attribute has a particular value. Instances are categorized as follows navigate from root to leaf Test results along the path. (Figure below) depicts a decision tree that reasons for whether or not the prospect responded to direct mail. Internal nodes are represented as circles, while leaves are represented as triangles. Note that this decision tree contains nominal and numeric attributes. The classifier allows analysts to predict potential responses to customers (by sorting them in a tree) and understand the behavioral characteristics of the entire group of potential customers associated with direct mail. Each node is marked with the property it is testing, and its branches are also marked with their corresponding value.[55],[56]

II.6 Conclusion:

In this chapter, we have mentioned the Features extraction and their types and cite the features most used in this domain, then we discussed About the classifiers used in this work and those concept .Afterward, we established a explanation about machine learning and their types and mentioned Artificial intelligence .

In the next chapter, we mentioned others similar works and we discussed the experimental database and our system design and experimental Setup to build our system then we will discuss about result .

Implementation and Evaluation

III.1 Introduction

Cough data for symptomatic and asymptomatic patients. However, as mentioned earlier, the sound of coughing may be reliable in Differentiate COVID-19 patients from healthy groups. It also can be seen that the most powerful scene is to stratify asymptomatic patients with COVID-19 cough sound. Asymptomatic patients are those who unknowingly spread the virus, Our trained network can identify them well by their coughs. That is why This COVID-19 Screening Framework Can Significantly Help Screen Suspected Cases population and reduce the risk of transmission.
In this chapter, we will discuss the results of applying the proposed method for the early detection of Covid-19. The database that was used is TOS COVID. In Covid-19 Detection, the features used for matching are MFCC. Actually, we use it , Because it is simulate of human hearing.
The first part of our work consists in doing an segmentation of the Signal then extract the attributes. The second part of the work uses attribute vectors to build the classifier prediction model. We will evaluate the proposed method by calculating: Accuracy, Sensitivity, specificity F1 score, and ROC curve Finally, we interpret the results.

III.2 Covid-19 detection using voice

The use of voice to detect COVID-19 symptoms is particularly attractive in the current situation as many countries are facing shortages of COVID-19 testing kits. It is therefore imperative to use the available testing kits judiciously. As there is a possibility of symptoms being detected at an early stage using the proposed method, individuals can immediately isolate themselves as a precautionary measure to minimize spreading before they are tested and provided further medical support. Once developed, there will be no need for specialized medical devices. Speech can be collected using a microphone which is available on every mobile device and can be analyzed using an 'app'. If any symptoms are detected, the 'app' can be configured to alert the user as well as health care authorities. The initial screening for suspect COVID-19 symptoms can be done from a remote location, as cough signals can be collected over the telephone.
Anomalies in the biological parameters can be detected much before the symptoms of COVID-19 become conspicuous. The large-scale screening will be possible at low cost

and without complex or expensive medical devices. The proliferation and ubiquity of smartphones allow rolling out the screening app in no time, once the AI models have been trained and incorporated into an 'app' making real-time self-screening for anyone with a smartphone a possibility.

III.3 State of the art

This work presents a novel approach to machine learning for the automatic detection of COVID-19 using cough sounds for symptomatic and asymptomatic patients. As can be seen in below, the proposed innovative approach has provided the best performance compared with A few other studies, instead, have indicated solutions for the detection of Covid-19 disorders based on an analysis of voice.

III.3.1N. Sharma 2020

the first proposed system based on sound to detect covid-19 was by "N.Sharma" in the year 2020 , where he conducted his study based on a database containing (Healthy and COVID-19-positive: 941) Where he studied the following phenomena (Cough, Breathing, Vowel, and Counting (1–20).)The method presented by "N.Sharma" was a system based on a random forest classifier using spectral contrast, MFCC, spectral roll-off, spectral centroid, mean square energy, polynomial fit, zero-crossing rate, spectral bandwidth, and spectral flatness. This system has achieved Accuracy: 76.74(%). [57]

III.3.2 C. Brown et al. 2021

"C.Brown" presented his proposed system for the detection of asthma and Covid together, and his study came in 2021, where he based his study on a database containing (COVID-19-positive: 141,Non-COVID: 298,COVID-19-positive with Cough: 54,Non-COVID-19 with Cough: 32, Non-COVID-19 asthma: 20). this system studied the following phenomena ((Cough, Breathing) and based on CNN-based approach using spectrogram, spectral centroid, MFCC . This system has achieved Accuracy: 80(%). [58]

III.3.3 V. Espotovic 2021

In the same year, ”V.Espotovic” came to present a new study based on the detection of COVID-19 based on coughing and breathing. He used a database containing (COVID-19-Positive: 84,COVID-19-Negative: 419)The method presented by ”V.Espotovic” was a system based on Ensemble-boosted approach using spectrogram and wavelet.This system has achieved Accuracy: 88.52(%) [59]

III.3.4 R.Islam 2022

”R.Islam”came to present a new study based on the detection of Covid based on cough only, where he used a database containing (COVID-19-Positve: 50,Healthy: 50) The method proposed by ”R.Islam” was a system based on CNN-based approach using zero-crossing rate, energy, energy entropy,spectral centroid, spectral entropy, spectral flux, spectral roll-offs, MFCC.This system has achieved Accuracy: 88.52(%) [60]

III.3.5 Rahman.T 2022

”Rahman.T” presented his system for the detection Covid-19 , where he based his study on a database containing (COVID-19-Positve: 237,Healthy: 827). this system studied the following phenomena ((Cough, Breathing) and based on CStacking-based CNN based approach using spectograms . This system has achieved Accuracy: For symptomatic, 96.5% and for asymptomatic 98.85% [61]

III.4 System design

General procedure for the proposed approach to detection of COVID-19 They are described in the flowchart in figure 16 .

In the first part, we restructure and arrange the database and then sample the signal. The analysis method uses an algorithm to estimate the cough signal from the recording signal. I mean the algorithm will extract samples.

In the second part, we go to the next stage, where the special sound parameters are extracted from each sample of the cough signal taken. So that each sound has its own parameters After extracting the audio features , we go to the last step in this part, which is choosing one of the following classifiers (KNN,SVM,DT).

In the third and final part, the selected classifier builds its own classification model and then classifies the cough signals into two classes:

positive negative

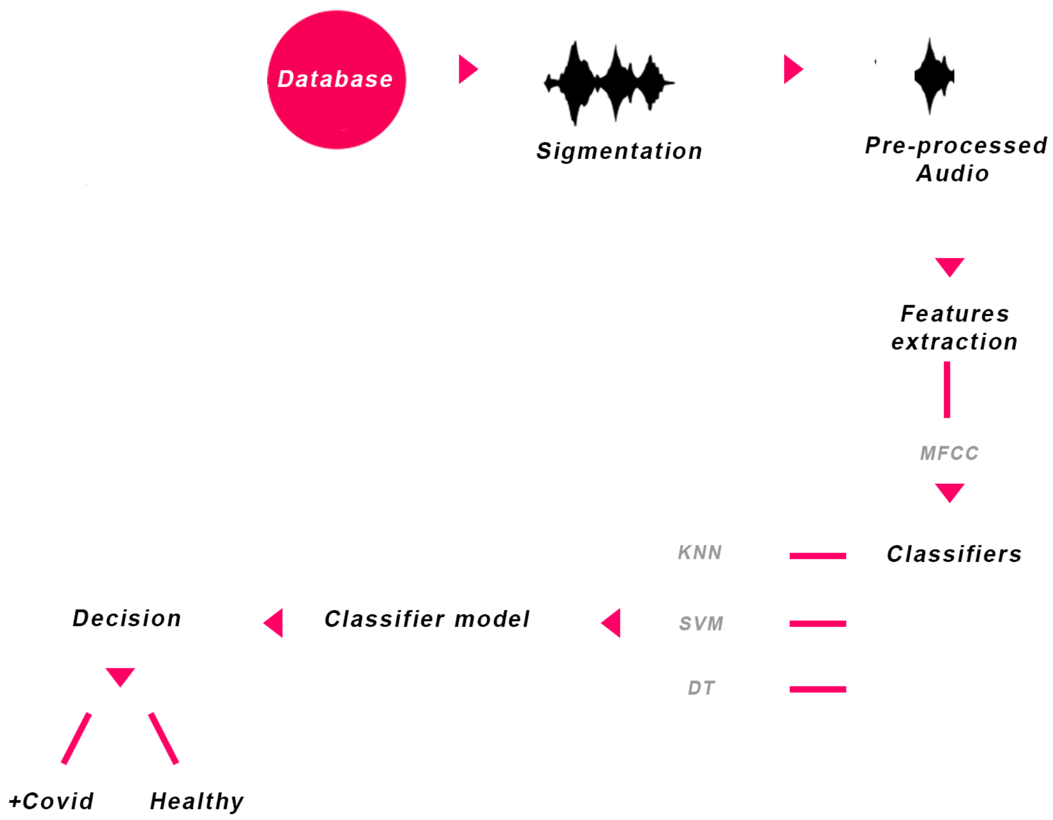

Figure 16: System design

III.5 Experimental database

The audio samples collected for this study belong to individuals Removed in all public and 1 private testing facility in the city of Buenos Aires Institutions conducting RT-PCR studies in patients with suspected COVID, and 14 out-of-hospital isolation wards for patients with confirmed mild COVID cases. That Audio was collected via a WhatsApp chatbot in the city of Buenos Aires Specifically designed to answer citizen inquiries related to the coronavirus pandemic (Coronavirus disease). Participants The first dataset collected corresponds to 2821 people Tested in Buenos Aires from August 11 to December 2, 2020.

People were divided into 1409 people who tested positive for COVID-19 and 1412 people who tested positive for COVID-19 Tested negative.

In this sample group, 52.6% of the subjects were women, 47.4% is male. Based on their PCR swab test results, the audio files for each person were classified as "positive" and "negative." In this work, we use 2771 audio ID files, which have been classified into 1378 detectable and 1393 undetectable.

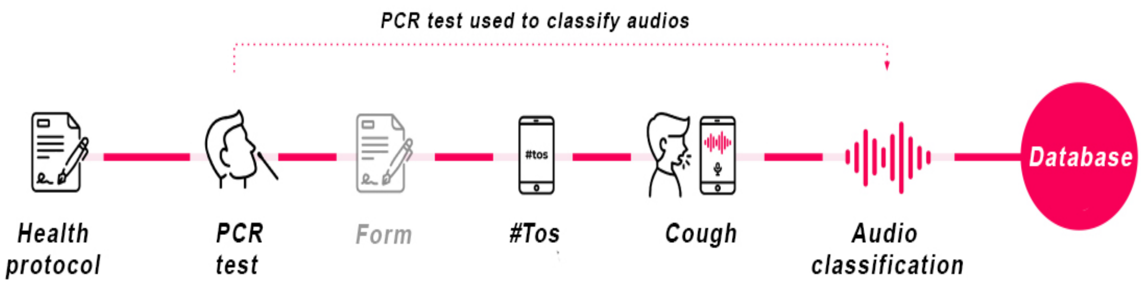

Figure 17: data-base setup design

III.6 Experimental setup

III.6.1 Development Hardware

Our detection system is implemented using MATLAB R2021a in an experimental platform as a workstation (HP Z8 G4), with a 64-bit Microsoft Windows 10 operating system, equipped with :

- *CPU* : Intel Xeon Silver 4108 processor
- *RAM* : 96 GB
- *GPU*: (NVIDIA GeForce RTX 3090 ‘64’,GeForce RTX 2080 Ti)

III.6.2 Development Software

The program used: $MATLABR2021a$

MATLAB is a high-level language for programming and numeric computing platform , He facilitate complex operations.

MatLab is based on formatting data in arrays and the program is similar to most programming languages used by millions of engineers and scientists to analyze data, develop algorithms, and create models.

III.7 Method setup

III.7.1 Data Split

In the first step, we do data cleaning, for new classifying the ID file into six classes ‘Male’, ‘Female’ and ‘detectable’, ’no detectable for each one of them.

III.7.2 Data Segmentation

Two segmented cough sound samples [53] to examine differences between coughs from COVID-negative (ie healthy/controlled) subjects and COVID-positive patients. Cough sound samples from healthy subjects Figure 14 shows a COVID-positive patient. Indicates that a healthy sample is similar to a typical human cough The sound signal is shown in Figure 18. However, the cough sound samples come from COVID-19 patients differ markedly from typical human coughs sound sample. For example, the middle stage and the voiced stage are both COVID-positive patients lived longer than healthy subjects.

Also, during the voiced phase, the signal amplitude is higher For COVID-positive patients, for healthy subjects. amplifier The two coughs also sound differently during the burst phase rehearse,as depicted in Fig. 16. The differences mentioned above indicate that the cough sound can be used as a valuable tool to discriminate the COVID-positive patient from the healthy subject. The recorded cough signal is in digital format. Is pre-processed

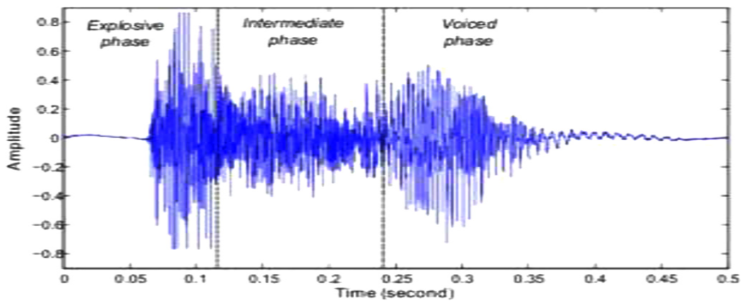

Figure 18: A typical cough sound signal

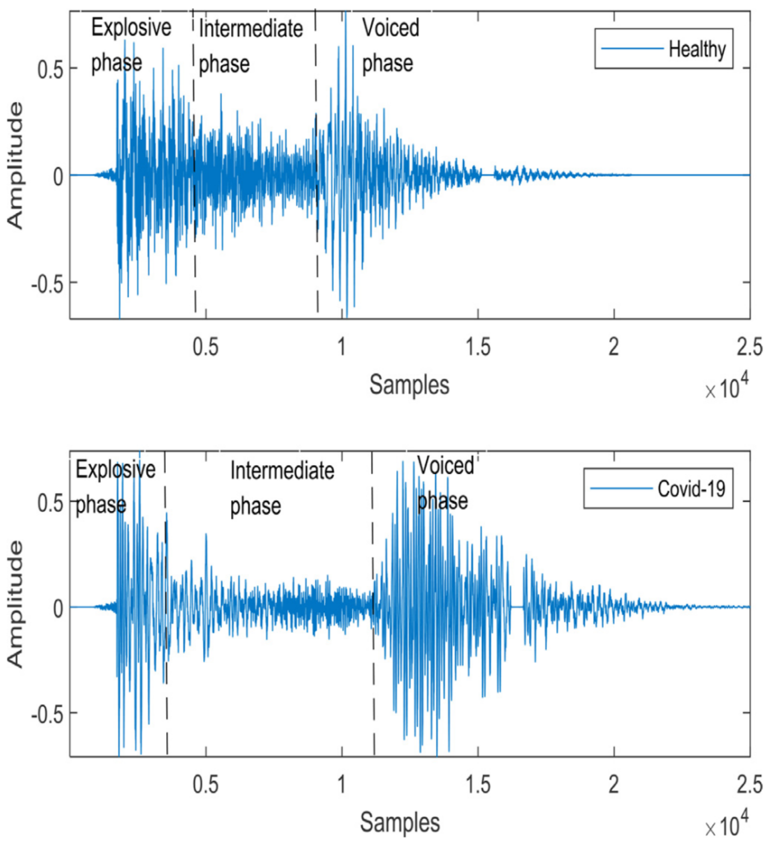

Figure 19: Comparison of the cough sounds for a healthy subject and a COVID-19

to remove unwanted components, background noise, and even silence interval/pause using the signal processing technique.

The cough signal was a random signal we normalized this signal, and after, that we fix the threshold (0.9) amplitude of signal . to localize the cough in the signal recorded. after that we apply Convolutional produce between the signal normalized and the with rect window with size ($2084_m s$) to cut first and the second cough duration.

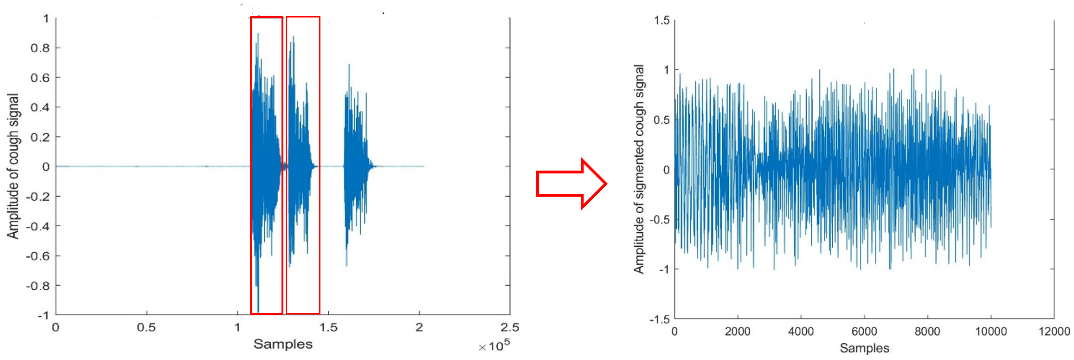

Figure 20: segmented signal

III.7.3 Features extraction

After that, the cough signal is further processed using a features extraction algorithm to extract "features" that characterize the cough signal.
This Algorithm feature was MEL-FREQUENCY CEPSTRAL COEFFICIENT.The MFcc is a simulation of the human ear, which can distinguish between sounds for both sexes, and therefore its settings differ between genders,as well as the sounds between them.

III.7.3.a Mel-Frequency spectral coefficient Algorithm

MFCC are popular features extracted from speech signals for recognition tasks. In the source filter model of speech, the cepstral coefficients are understood to represent the filter (the vocal tract). The frequency response of the vocal tract is relatively smooth, and the source of the voiced sound can be modeled as a pulse train. As a result, the vocal tract can be estimated from the spectral envelope of the speech segment.
The idea behind the mel-frequency cepstral coefficients is to compress information about the vocal tract (smooth spectrum) into a small number of coefficients, based on an understanding of the cochlea. Although there is no fixed standard for calculating coefficients, the figure outlines the basic steps.

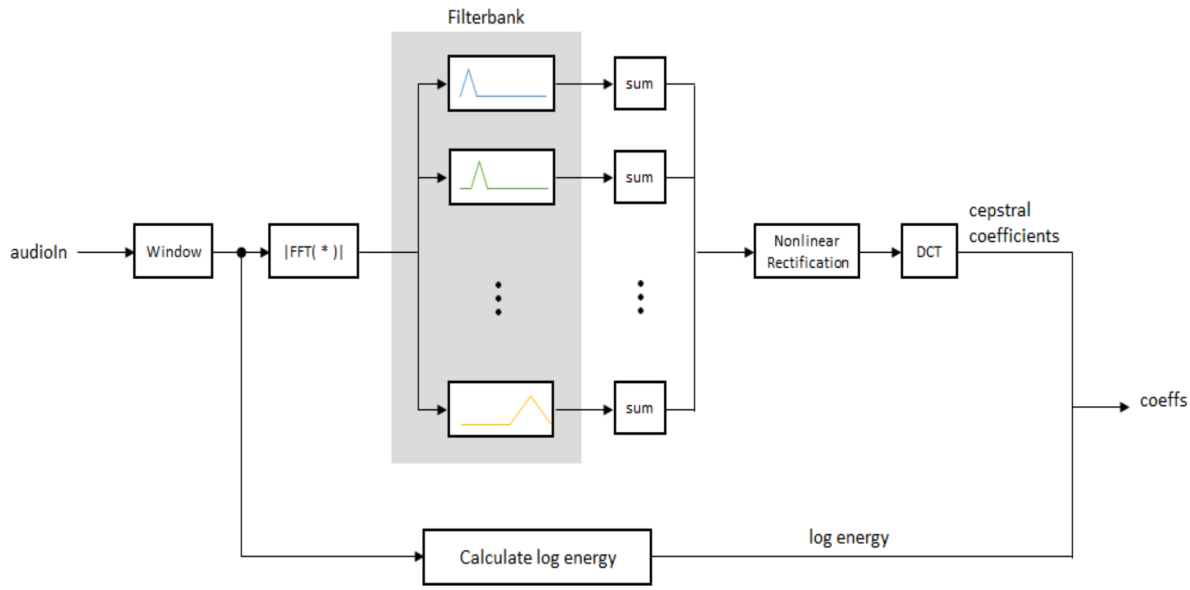

Figure 21: The MFCC Blocks

The default mel filter bank linearly spaces the first 10 triangular filters and logarithmically spaces the remaining filters.

The information contained in the cepstral coefficients of the zeroth Mel frequency is

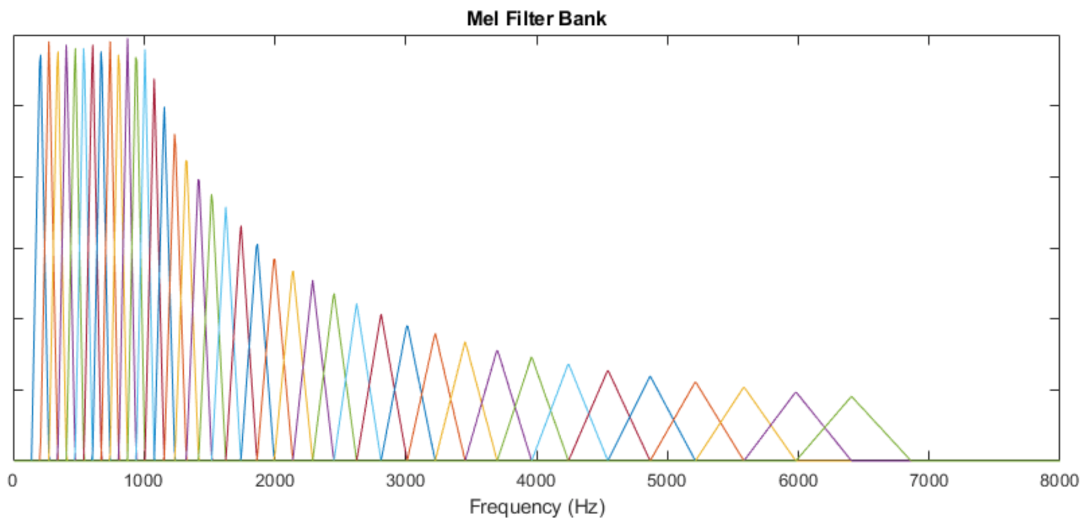

Figure 22: The Mel-filter bank

usually supplemented or replaced with logarithmic energy. The logarithmic energy calculation depends on the input domain.

If the input (audioIn) is a time domain signal, the logarithmic energy is calculated using the following equation:

$$\log E = \log\left(\text{sum}\left(x^2\right)\right) \tag{8}$$

If the input (audioIn) is a frequency domain signal, the logarithmic energy is calculated using the following equation:

$$\log E = \log\left(\operatorname{sum}\left(|x|^2\right) / \text{ FFTLength }\right) \tag{9}$$

III.7.3.b Classification

classification, which is the last stage in our proposed system. the selection of one of the following classifiers with his parameters : SVM, CNN, DT.

KNN: we used "fitcknn" MATLAB function to build our classification model, the parameter we changed is "NumNeighbors" 1 and "Standardize" 1.

SVM: we used "fitcsvm" MATLAB function to build our classification model, the parameter we changed is "KernelFunction" "R B F" and "KernelScale" "auto".

- KernelFunction :

Kernel function used to compute the elements of the Gram matrix, specified as the comma-separated pair consisting of 'KernelFunction' and a kernel function name.

Suppose (X_j, X_k) is element (j,k) of the Gram matrix, where xj and xk are p-dimensional vectors representing observations j and k in X. This table describes supported kernel function names and their functional forms.

"RBF" : Radial Basis Function (RBF) kernel determined by :

$$G\left(x_j, x_k\right) = \exp\left(-\left\|x_j - x_k\right\|^2\right) \tag{10}$$

DT: we used "TreeBagger" MATLAB function to build our classification model, the parameter we changed is "OOBPrediction" On, "Method" classification. In this work, when the classifier SVM was selected, it was based on .

Each classifier, after being identified, Create own model to classify each signal according to which class it belongs to as 'healthy' or 'covid-19'.

III.8 Experimental metrics

In this study, the performances of the voice pathology detection system were assessed by KNN, SVM, DT classifiers, using the MFcc for both of genders. The classification results are provided in terms of accuracy of accuracy, recall, precision, specificity, F1 score, and ROC. These performance indicators are widely used to assess the effectiveness of various medical decision systems.

- **Accuracy:** we can define accuracy as a measure of the capability of classifying the samples correctly, which is expressed as follows:

$$\text{Accuracy}(\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \tag{11}$$

- **Sensitivity:** the probability that abnormal samples will be diagnosed as positive.

$$\text{Recall}_{(}Sensitivity)\ (\%) = \frac{TP}{TP + FN} \times 100 \tag{12}$$

- **Specificity:** the probability of normal samples being incorrectly identified

$$\text{Specificity}(\%) = \frac{TN}{TN + FP} \times 100 \tag{13}$$

- **Precision:** the probability of normal samples being correctly identified

$$\text{Precision}(\%) = \frac{TP}{TP + FP} \times 100 \tag{14}$$

- **F1 Score:** F1-score, is a measure of a model's accuracy on a dataset. It is used to evaluate binary classification systems, which classify examples into 'positive' or 'negative'.

$$\text{F}_1\text{Score}(\%) = \frac{Precision \times Recall}{Precision + Recall} \times 100 \tag{15}$$

- True Positive (TP): the pathogenic nature of the speech sample is recognized by the marker.

- True Negative (TN): the voice sample is healthy, and the marker detects it.

- False Positive (FP): the voice sample is normal, but the marker detects it as pathological.

- False Negative (FN): the diseased voice sample is recognized by the marker as healthy.

III.9 Results and discussions

As we know orientation to medical detection system is considered as one of the best solutions to improve performance.
In this sense, during our experimental study based on a proposed system, we proceeded to all the trains and tests.
In the first step, we tested our system with female then male cough signal .
In the same way. In the second step we measured the performance of our system . The table below summarizes the results of evaluation of Female using the MFCC with this parameters .
for $Female$:
- Window-Length : 1000
- OverlapLength : 305

Table 1: Results of Female with MFCC features

Classifier	Accuracy%	Sensitivity%	specificity%	F1 score%
SVM	89.9	92.08	87.7	88
KNN	99.6	99.3	100	99
DT	99.3	98.6	100	99

The previous table represents the results of the proposed system on how to detect Covid and its effectiveness for females, where we obtained the best accuracy for detection by the Knn determinant algorithm, where the accuracy of its detection of infected individuals reached 99 percent, and the rate of its detection of uninfected persons amounted to one hundred percent. This does not mean that we cannot rely on another classifier in the detection, because our experience with our proposed system proved that the decision tree can detect with high accuracy, as its accuracy also reached ninety-nine percent for infected persons and one hundred percent for uninfected individuals.

The table below summarizes the results of evaluation of Male using the MFCC with this parameters .

for $Male$:

- Window-Length : 1175

- OverlapLength : 395

Table 2: Results of Male with MFCC features

Classifier	Accuracy%	Sensitivity%	specificity%	F1 score%
SVM	88	82.75	95.23	88
KNN	82	73.38	100	84
DT	84	75.75	100	86

The above table presents the results of the proposed Covid detection system and its effectiveness for men, where we achieved the best detection accuracy by the SVM determinant algorithm, detecting infected persons with an accuracy of 88% and a detection rate of The uninfected was 95.23%.

This does not mean that we cannot rely on another classifier for detection, as our experience with our proposed system has yielded similar results, as knn can detect 84% of infected individuals and 100% of uninfected individuals accuracy.

The infected decision tree detection accuracy also reached 82%, and the uninfected Peapole also achieved 100%.

III.10 ROC comapision

The area under the curve (AUC) / Receiver Operating Characteristics (ROC) curve (also known as AUROC (Area under Receiver Operating Characteristic)).

The ROC curve for a binary classification problem plots the true positive rate as a function of the false positive rate. The points of the curve are obtained by sweeping the classification threshold from the most positive classification value to the most negative.

this is the result of plotting ”Roc” for each gender using the ”SVM” classifier :

We notice in the Figure.20 a curve representing a function drawn for TP in terms of FP. The experiment of the proposed system on females presented the following results. The drawing starts ascending, where the TP values are greater than FP until it reaches equality in the end. We notice in the Figure.21 a curve representing a function drawn for

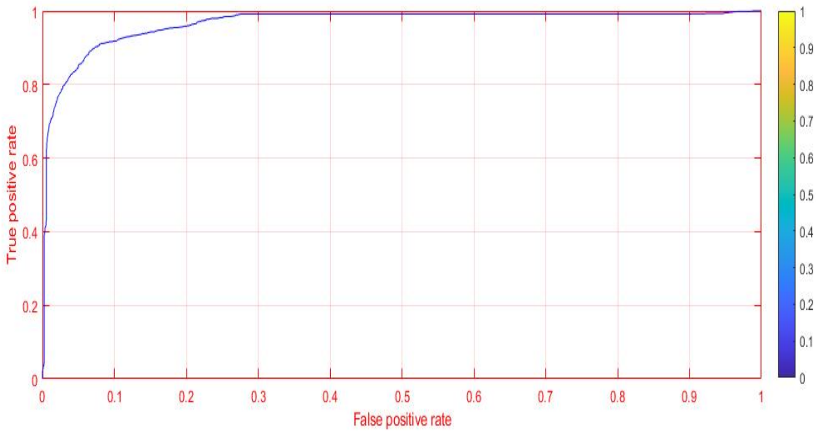

Figure 23: ROC curve for Female classification usin SVM

TP in terms of FP. The proposed system was tested on males, where he presented the following results. The graph starts upward, but there are re-fractions where the values of TP are greater than FP until it reaches equality in the end.

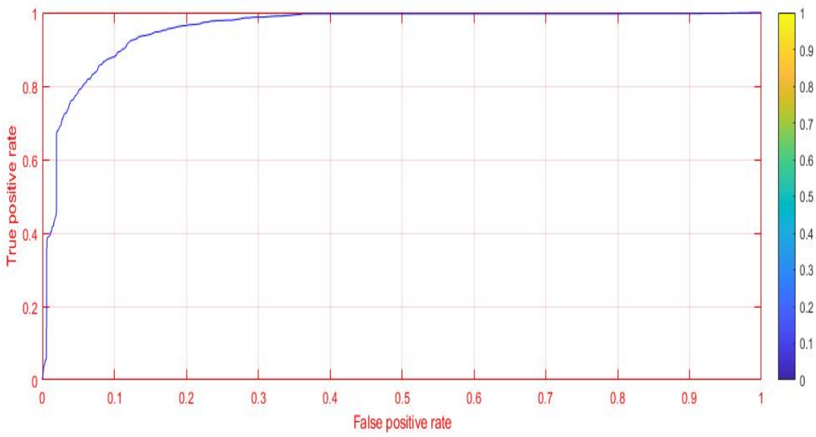

Figure 24: ROC curve for Male classification usin SVM

III.10.1 Comparison

We notice through figure 23 and figure 24 a difference in the roc curve, as the curve for females is greater than the curve for males, as well as the area under the curve, is greater then area under the curve for male , because the accuracy of female is 89.9% and the accuracy of Male is 88% .

so when AUC close to being equal to 1 the results are ideal. We can say that the higher the accuracy of the system, we will see the best roc.

III.11 Conclusion:

In this chapter, we examine and evaluate the state-of-the-art methods and architectures proposed by recent research on the efficient diagnosis of Covid-19 based on medical voice datasets using Machine learning. Most of the architectures examined are able to obtain more or less satisfactory results with high accuracy. However, more research on large and new datasets is always needed to improve the performance of existing methods for more efficient results.

General Conclusion

Medical uses are constantly evolving in technology, it is widely used in many fields for the detection of pathologies. The fundamental goal of Medical detection techniques is to detect covid-19 the person who is being tested. This COVID screening method is a substitute for traditional methods of restorative assays. Everyone is trying to create protect themselves from this broad protection of their lives. So, when people are going to take a selfie in the frame Incredible benefits society. Any innovation is certain for people's well-being and dynamic improvement. A sort of Later mechanical development lies in the field of manufacturing knowledge and its related fields. In this way, it aims to help curb this global spread. With that in mind, these are forward-looking approaches and proved to be effective in terms of accuracy. Our work in this area has limitations number of tests used to build prior knowledge prove. This can of course be used as confirmation provides the concept of machine operability and has well-founded learning methods to prepare diagnosis of COVID. The main goal of our study :

- We tried to propose the realization of detection system based on MFCC used (KNN , SVM ,DT).
- We used a preprocessed database and utilized the machine learning type supervised learning.
- Comparison of classifiers (SVM ,KNN,DT).
- The results obtained in this work show an excellent performance for Pathology and Healthy

We finally presented the experimental results, these results are presented from the metrics and ROC curves of SVM algorithm.

During this work. These results are very encouraging and show the effectiveness of both of genders , in order to develop an efficient Detection system.

Our future work will project to use other modalities, like (Vowels, speech, X-ray image,...etc) with other methods (like unsupervised machine learning, Deep learning).

Bibliography

[1] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang *et al.*, "A pneumonia outbreak associated with a new coronavirus of probable bat origin," *nature*, vol. 579, no. 7798, pp. 270–273, 2020.

[2] B. Chen *et al.*, "Overview of human lethal coronaviruses," *Signal Transduction and Targeted Therapy*, vol. 5, p. 89, 2020.

[3] S. Su, G. Wong, W. Shi, J. Liu, A. C. Lai, J. Zhou, W. Liu, Y. Bi, and G. F. Gao, "Epidemiology, genetic recombination, and pathogenesis of coronaviruses," *Trends in microbiology*, vol. 24, no. 6, pp. 490–502, 2016.

[4] E. Hadjaidji, M. C. A. Korba, and K. Khelil, "Spasmodic dysphonia detection using machine learning classifiers," in *2021 International Conference on Recent Advances in Mathematics and Informatics (ICRAMI).* IEEE, 2021, pp. 1–5.

[5] J. M. Sanders, M. L. Monogue, T. Z. Jodlowski, and J. B. Cutrell, "Pharmacologic treatments for coronavirus disease 2019 (covid-19): a review," *Jama*, vol. 323, no. 18, pp. 1824–1836, 2020.

[6] S. A. Kujawski, K. K. Wong, J. P. Collins, L. Epstein, M. E. Killerby, C. M. Midgley, and G. R. Abedi, "Clinical and virologic characteristics of the first 12 patients with coronavirus disease 2019 (covid-19) in the united states," *Nature medicine*, vol. 26, no. 6, pp. 861–869, 2020.

[7] D. Chang, M. Lin, L. Wei, L. Xie, G. Zhu, C. S. D. Cruz, and L. Sharma, "Epidemiologic and clinical characteristics of novel coronavirus infections involving 13 patients outside wuhan, china," *Jama*, vol. 323, no. 11, pp. 1092–1093, 2020.

[8] P. Sun, X. Lu, C. Xu, W. Sun, and B. Pan, "Understanding of covid-19 based on current evidence," *Journal of medical virology*, vol. 92, no. 6, pp. 548–551, 2020.

[9] M. Asiaee, A. Vahedian-Azimi, S. S. Atashi, A. Keramatfar, and M. Nourbakhsh, "Voice quality evaluation in patients with covid-19: An acoustic analysis," *Journal*

of Voice, 2020.

[10] A. Scohy, A. Anantharajah, M. Bodéus, B. Kabamba-Mukadi, A. Verroken, and H. Rodriguez-Villalobos, "Low performance of rapid antigen detection test as front-line testing for covid-19 diagnosis," *Journal of Clinical Virology*, vol. 129, p. 104455, 2020.

[11] R. D. Kent, "Vocal tract acoustics," *Journal of Voice*, vol. 7, no. 2, pp. 97–117, 1993.

[12] J. Schroeter and M. M. Sondhi, "Techniques for estimating vocal-tract shapes from the speech signal," *IEEE Transactions on Speech and Audio Processing*, vol. 2, no. 1, pp. 133–150, 1994.

[13] J. L. Flanagan, K. Ishizaka, and K. L. Shipley, "Synthesis of speech from a dynamic model of the vocal cords and vocal tract," *Bell System Technical Journal*, vol. 54, no. 3, pp. 485–506, 1975.

[14] L. F. Tracy, R. K. Segina, M. D. Cadiz, and C. E. Stepp, "The impact of communication modality on voice production," *Journal of Speech, Language, and Hearing Research*, vol. 63, no. 9, pp. 2913–2920, 2020.

[15] K. E. Bainbridge, N. Roy, K. G. Losonczy, H. J. Hoffman, and S. M. Cohen, "Voice disorders and associated risk markers among young adults in the united states," *The Laryngoscope*, vol. 127, no. 9, pp. 2093–2099, 2017.

[16] D. R. Boone, S. C. McFarlane, S. L. Von Berg, and R. I. Zraick, "The voice and voice therapy," 2005.

[17] S. G. Adams and A. E. Lang, "Can the lombard effect be used to improve low voice intensity in parkinson's disease?" *European Journal of Disorders of Communication*, vol. 27, no. 2, pp. 121–127, 1992.

[18] J. K. Casper and R. Leonard, *Understanding voice problems: A physiological perspective for diagnosis and treatment.* Lippincott Williams & Wilkins, 2006.

[19] A. Deslandes, V. Berti, Y. Tandjaoui-Lambotte, C. Alloui, E. Carbonnelle, J. Zahar, S. Brichler, and Y. Cohen, "Sars-cov-2 was already spreading in france in late december 2019," *International journal of antimicrobial agents*, vol. 55, no. 6, p. 106006, 2020.

[20] D. S. W. Ting, L. Carin, V. Dzau, and T. Y. Wong, "Digital technology and covid-19," *Nature medicine*, vol. 26, no. 4, pp. 459–461, 2020.

[21] L. K. Tanno, T. Casale, and P. Demoly, "Coronavirus disease (covid)-19:

World health organization definitions and coding to support the allergy community and health professionals," *The Journal of Allergy and Clinical Immunology: In Practice*, vol. 8, no. 7, pp. 2144–2148, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2213219820304700

[22] K. Erles, C. Toomey, H. W. Brooks, and J. Brownlie, "Detection of a group 2 coronavirus in dogs with canine infectious respiratory disease," *Virology*, vol. 310, no. 2, pp. 216–223, 2003.

[23] G. Liya, W. Yuguang, L. Jian, Y. Huaiping, H. Xue, H. Jianwei, M. Jiaju, L. Youran, M. Chen, and J. Yiqing, "Studies on viral pneumonia related to novel coronavirus sars-cov-2, sars-cov, and mers-cov: a literature review," *Apmis*, vol. 128, no. 6, pp. 423–432, 2020.

[24] F. Di Gennaro, D. Pizzol, C. Marotta, M. Antunes, V. Racalbuto, N. Veronese, and L. Smith, "Coronavirus diseases (covid-19) current status and future perspectives: a narrative review," *International journal of environmental research and public health*, vol. 17, no. 8, p. 2690, 2020.

[25] J. E. Kim, J. H. Lee, H. Lee, S. J. Moon, and E. W. Nam, "Covid-19 screening center models in south korea," *Journal of Public Health Policy*, vol. 42, no. 1, pp. 15–26, 2021.

[26] W. Scientific, D. Heymann, N. Shindo *et al.*, "Covid-19: What is next for public health?" *Lancet*, vol. 395, pp. 542–5, 2020.

[27] Y.-C. Wu, C.-S. Chen, and Y.-J. Chan, "The outbreak of covid-19: An overview," *Journal of the Chinese medical association*, vol. 83, no. 3, p. 217, 2020.

[28] I. Futterman, M. Toaff, L. Navi, and C. A. Clare, "Covid-19 and hellp: overlapping clinical pictures in two gravid patients," *American Journal of Perinatology Reports*, vol. 10, no. 02, pp. e179–e182, 2020.

[29] X. Yang, Y. Yu, J. Xu, H. Shu, H. Liu, Y. Wu, L. Zhang, Z. Yu, M. Fang, T. Yu *et al.*, "Clinical course and outcomes of critically ill patients with sars-cov-2 pneumonia in wuhan, china: a single-centered, retrospective, observational study," *The Lancet Respiratory Medicine*, vol. 8, no. 5, pp. 475–481, 2020.

[30] M. Cascella, M. Rajnik, A. Aleem, S. C. Dulebohn, and R. Di Napoli, "Features, evaluation, and treatment of coronavirus (covid-19)," *Statpearls [internet]*, 2022.

[31] H.-D. Wehle, "Machine learning, deep learning, and ai: What's the difference?" 07

2017.

[32] C. A. U. Hassan, M. S. Khan, and M. A. Shah, “Comparison of machine learning algorithms in data classification,” in *2018 24th International Conference on Automation and Computing (ICAC).* IEEE, 2018, pp. 1–6.

[33] K. Wakefield, “A guide to the types of machine learning algorithms and their applications,” *URL: https://www. sas. com/en_gb/insights/articles/analytics/machine-learning-algorithms. html [Accessed on 10 February 2021]*, 2021.

[34] T. M. Mitchell *et al.*, “Machine learning,” 1997.

[35] G. Sharma, K. Umapathy, and S. Krishnan, “Trends in audio signal feature extraction methods,” *Applied Acoustics*, vol. 158, p. 107020, 2020.

[36] F. Goldman-Eisler, “Speech analysis and mental processes,” *Language and speech*, vol. 1, no. 1, pp. 59–75, 1958.

[37] K. Stevens, “Autocorrelation analysis of speech sounds,” *The Journal of the Acoustical Society of America*, vol. 22, no. 6, pp. 769–771, 1950.

[38] C. G. Howard, “Speech analysis-synthesis scheme using continuous parameters,” *The Journal of the Acoustical Society of America*, vol. 28, no. 6, pp. 1091–1098, 1956.

[39] R. K. Potter and J. C. Steinberg, “Toward the specification of speech,” *The Journal of the Acoustical Society of America*, vol. 22, no. 6, pp. 807–820, 1950.

[40] G. Gambardella, “A contribution to the theory of short-time spectral analysis with nonuniform bandwidth filters,” *IEEE Transactions on Circuit Theory*, vol. 18, no. 4, pp. 455–460, 1971.

[41] S. K. Card, T. P. Moran, and A. Newell, *The psychology of human-computer interaction.* Crc Press, 2018.

[42] G. Gambardella, “Time scaling and short-time spectral analysis,” *The Journal of the Acoustical Society of America*, vol. 44, no. 6, pp. 1745–1747, 1968.

[43] S. Shetty, S. Hegde, T. Dodderi *et al.*, “Classification of healthy and pathological voices using mfcc and ann,” in *2018 Second International Conference on Advances in Electronics, Computers and Communications (ICAECC).* IEEE, 2018, pp. 1–5.

[44] A. Bala, A. Kumar, and N. Birla, “Voice command recognition system based on mfcc and dtw,” *International Journal of Engineering Science and Technology*, vol. 2, no. 12, pp. 7335–7342, 2010.

[45] L. Sriratana, S. Poochong, and K. Bisalyaputra, “International journal of basic sci-

ences and applied computing (ijbsac) international journal of basic sciences and applied computing (ijbsac).”

[46] R. Islam, M. Tarique, and E. Abdel-Raheem, “A survey on signal processing based pathological voice detection techniques,” *IEEE Access*, vol. 8, pp. 66 749–66 776, 2020.

[47] Y. Ninja, “Support vector machines in python-svm in python 2019.”

[48] ——, “Support vector machines in python-svm in python 2019.”

[49] A. Subasi, *Practical guide for biomedical signals analysis using machine learning techniques: A MATLAB based approach.* Academic Press, 2019.

[50] R. Short and K. Fukunaga, “The optimal distance measure for nearest neighbor classification,” *IEEE transactions on Information Theory*, vol. 27, no. 5, pp. 622–627, 1981.

[51] K. Q. Weinberger and L. K. Saul, “Distance metric learning for large margin nearest neighbor classification.” *Journal of machine learning research*, vol. 10, no. 2, 2009.

[52] S. Cost and S. Salzberg, “A weighted nearest neighbor algorithm for learning with symbolic features,” *Machine learning*, vol. 10, no. 1, pp. 57–78, 1993.

[53] L. Breiman, “Random forests,” *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[54] Z. Zhang, “Introduction to machine learning: k-nearest neighbors,” *Annals of translational medicine*, vol. 4, no. 11, 2016.

[55] H. Almuallim, “An efficient algorithm for optimal pruning of decision trees,” *Artificial Intelligence*, vol. 83, no. 2, pp. 347–362, 1996.

[56] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees.* Routledge, 2017.

[57] N. Sharma, P. Krishnan, R. Kumar, S. Ramoji, S. R. Chetupalli, P. K. Ghosh, S. Ganapathy *et al.*, “Coswara–a database of breathing, cough, and voice sounds for covid-19 diagnosis,” *arXiv preprint arXiv:2005.10548*, 2020.

[58] C. Brown, J. Chauhan, A. Grammenos, J. Han, A. Hasthanasombat, D. Spathis, T. Xia, P. Cicuta, and C. Mascolo, “Exploring automatic diagnosis of covid-19 from crowdsourced respiratory sound data,” *arXiv preprint arXiv:2006.05919*, 2020.

[59] V. Despotovic, M. Ismael, M. Cornil, R. Mc Call, and G. Fagherazzi, “Detection of covid-19 from voice, cough and breathing patterns: Dataset and preliminary results,” *Computers in Biology and Medicine*, vol. 138, p. 104944, 2021.

[60] R. Islam, E. Abdel-Raheem, and M. Tarique, "A study of using cough sounds and deep neural networks for the early detection of covid-19," *Biomedical Engineering Advances*, p. 100025, 2022.

[61] T. Rahman, N. Ibtehaz, A. Khandakar, M. S. A. Hossain, Y. M. S. Mekki, M. Ezeddin, E. H. Bhuiyan, M. A. Ayari, A. Tahir, Y. Qiblawey *et al.*, "Qucoughscope: An intelligent application to detect covid-19 patients using cough and breath sounds," *Diagnostics*, vol. 12, no. 4, p. 920, 2022.