COVID-19 PREDICTION USING

CLINICAL REPORT TEXT AND X-RAY DATA

Project report submitted in fulfillment of the requirement for degree of

Bachelor of Technology

In

Computer Science and Engineering/Information Technology

By

(HIRDYANSH BHALLA (171362))

Under the supervision of

(DR. HIMANSHU JINDAL)

To

Department of Computer Science & Engineering and Information Technology

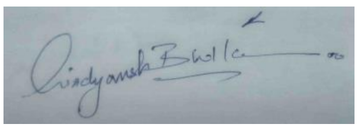
Jaypee University of Information Technology Waknaghat,Solan-173234

Himachal Pradesh

Certificate

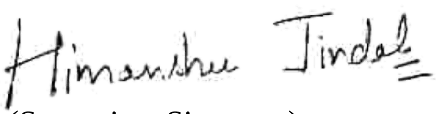
I hereby declare that the work presented in this report entitled "Covid-19 prediction using clinical report text and x-ray data" in the minimum fulfillment of the **Bachelor of Technology** degree in **Computer Science and Engineering / Information Technology** Jaypee University of Information Technology Waknaghat is a true record of my work done in the period from January 2021 to May 2021 under the supervision of **Dr. Himanshu Jindal** (Assistant Professor (SG), Department of Science and Computer Science).

The material included in this report has never been awarded the award of any other degree or diploma.



(Student Signature)
Hirdyansh Bhalla, 171362

This is to certify that the above statement made by the candidate is true to the best of my knowledge.



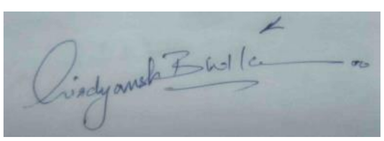
(Supervisor Signature)
Supervisor Name: Dr. Himanshu Jindal
Designation: Assistant Professor (SG)
Department name: Computer Science and Engineering and Information Technology
Dated: 14-05-2021

Acknowledgement

Words are not enough to thank our guide, Dr. Himanshu Jindal, Assistant Professor (SG), Department of Computer Science and Engineering, Jaypee University of Information Technology, Solan, for being so concerned and helping everyone. This is the material needed to prepare the main project report. He has helped us to explore this vast area in an organized manner and has given us all the ideas to work towards a research based organization.

We also thank the professor. (Dr) SP Gera, Head, Department of Computer Science and Engineering and Mr. Ravi Raina, Lab Assistant, for the inspiration and inspiration that inspired us to do this major project work.

Most importantly, we thank our parents, friends and the Almighty for showing us the right direction, for helping us to stay calm at the most uneven times and to continue even at such times, when there is no hope.



Hirdyansh Bhalla

171362

Table of Content

List of Abbreviations

CNN--------------------------------Convolution neural network

BOW________________________Bag of Word

TF__________________________Term Frequency

IDF----------------------------------Inverse Document Freq.

List of Figures

List of Graphs

List of Tables

Abstract

The progress of the discovery will rapidly affect every area of life in the clinical field or any other field. Computerized reasoning has indicated good results in medical services through its mobility by disassembling information. Two different phases are included in this project. We will first use the clinical textual data and after that the image Xray data for the prediction of covid positive patients. The coronavirus has infected more than 100 countries at any one time. People everywhere in the world are helpless against its consequences in the future. Creating a control framework that identifies Covid is fundamental. The disease can be diagnosed with the help of various AI devices to control motion loss. In this paper, we have divided literary clinical reports into four sections using traditional and institutional AI calculations. Highlight designing was done using techniques such as Word Repeat / Contrast Record Repeat accuracy. Also, discoveries made infection. The use of man-made brain power (AI) methods in conjunction with radiological imaging can be used to accurately diagnose the disease, as well as the lack of specific physicians in remote cities to help solve the problem. In this research, another diagnosis of double grouping (Covid vs. no-finding). Model will use two different ML algorithms namely SVM and MNB on the clinical textual data after preprocessing and feature engineering. Finally Deep neural network will be used for the X Ray image data. Our model is designed to be 98.08% accurate for double class. We grasped 17 concentrated layers and introduced a specific transition for each layer. At last both the results would be compared to know the proper efficiency of the model and working of the model.

CHAPTER 1

INTRODUCTION

1.1 Introduction

In the month of December in 2019, first case of covid 19 appeared in Chinese city of Wuhan [1] . WHO came to know about this on 31 Dec 2019. The infection poses a worldwide risk and WHO started to call it by name COVID-19 on Feb 11th, 2020 [1]. COVID-19 is a group of diseases including SARS, ARDS. WHO declared this conflict as a general social problem [2] and identified the corresponding; the infection is communicated by means of the respiratory lot when a solid individual interacts and an unclean person. Infection can communicate between people through a variety of yet vague rootsPeople with various illnesses such as asthma, diabetes, and vascular disease have no immunity to infection and are more likely to be very ill. The person himself analyze dependent on indications and his movement history. Different paramedical organizations have guaranteed of building up an antibody for this infection. Less testing has additionally offered ascend to this infection as we do not have the clinical assets because of pandemic. We have thousands and thousands are tried +ve step by step the world over, it is beyond the realm of imagination to expect to test all the people who show manifestations.

The People with various infections such as asthma, diabetes and coronary heart disease are more vulnerable to infection and more likely to become seriously ill. Person analyze dependent on manifestations and his movement history. Significant signs are clearly recognized that the customer has indicators. No treatment has been received as of April 10, 2020, and patients are openly discussed. Medicines such as hydroxychloriquine, an antipyretic, an antiseptic are used for sexually transmitted infections. From now on, no such vaccine is designed to prevent the spread of this deadly disease, and it can prevent the spread of the disease. Regular hand washing with a 20-second cleanser and avoiding close contact with others by keeping a distance of 1 m can reduce the risk of infection. While breathing, covering the mouth and nose with the help of donated tissues and staying away from contact with the nose, ear and mouth can help avoidance

.

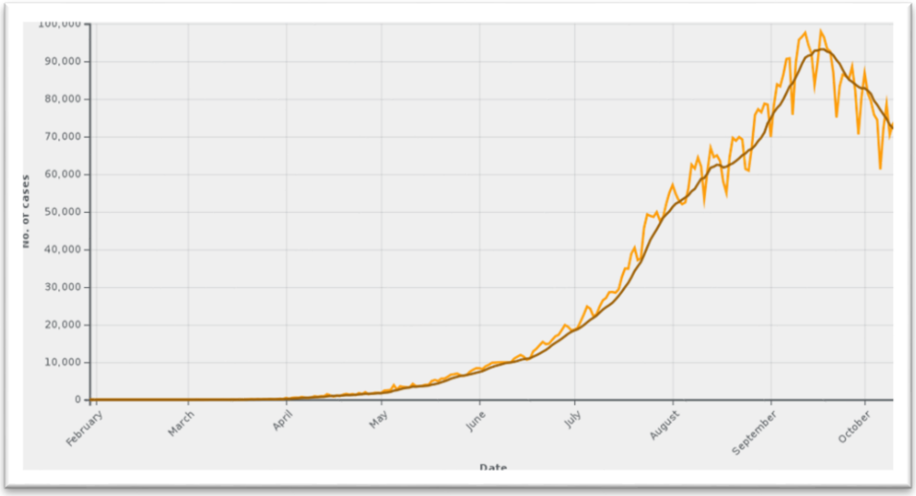


Fig 1.1 India's daily covid-19 cases

1.1.1 Covid-19 detection using clinical Texts

Aside from clinical methodology, AI gives a ton of help in distinguishing the infection with the assistance of picture and printed information. AI can be utilized for the ID of novel Covid. It can likewise estimate the idea of the infection over the globe. In any case, AI requires a gigantic measure of information for arranging or anticipating infections. Administered AI calculations need explained information for ordering the content or picture into various classes. Since the last decade, a major step of progress has been taken here to dispose of some basic firms. Late epidemic pulled in numerous analysts around the world to take care of this issue. Details provided by John Hopkins University as X-beam images and various experts make an AI model that marks the X-beam image on COVID 19 positive or not. So basically we are using the data provided by the Johns Hopkins University for this model and they have provided both X Ray and metadata. We will first use the metadata and extract relevant data from it and do feature engineering and try to classify or divide the results into 4 of our proposed classes which are namely ARDS, SARS, BOTH(COVID AND SARS) and finally the SARS. This we can do by using different machine learning algorithms available specially MNB and SVM.

1.1.2 Covid-19 detection using Radiology (X-Rays)

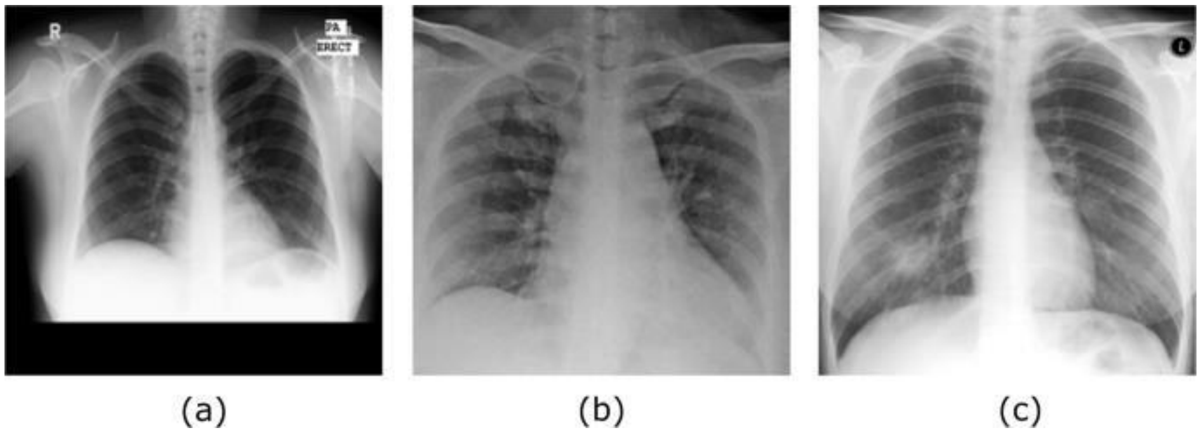


Figure 1.2 **a**) average person, **b**) COVID virus positive, and **c**) sars+ chest X-beam pictures

Toward the start of the in epidemics, the clinical focus of the Chinese was inadequate packaging, which further created a rate which was much more of artificial adverse effects. CT scans is used in the case of COVID-19 in the countries, for example, Turkey, where a low number of test units at the beginning of the epidemic were found. Scientists point out that joining the best clinical images with lab results can help detect COVID-19 early. Radiologic images like Xrays and CT scans prove really helpful. This digital data can be used as a basis to propose a much more efficient model based on this [18]. Sensitive disclosures were allowed by ones who were doing investigations in the image based investigation of COVID-19.

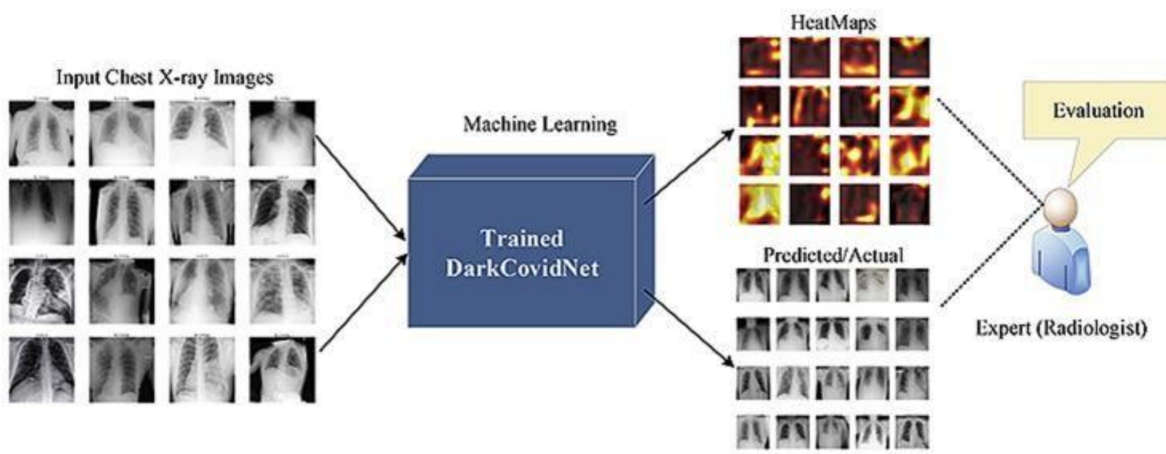


Fig 1.3 Basic Overview

Utilization of AI techniques for programmed analysis in the clinical field have as of late picked up notoriety by turning into an assistant instrument for clinicians. Profound realizing, which is a famous exploration territory of man-made consciousness (AI), empowers the production of start to finish models to accomplish guaranteed results utilizing input information, without the requirement for manual component extraction.

1.2 Problem Statement

1.2.1 Problem Statement

In this project we are using two different approaches to detect whether a person is covid-19 positive or not. In the first method we are taking X-ray data from James Hopkins university and try to analyze the metadata associated with these images. Metadata consists of the clinical reports, We are text mining the relevant features from the clinical report and with help of those features and various machine learning algorithms we are trying to detect whether someone is covid-19 positive or not. At last we are comparing accuracies using different algorithms and judging which one suits best for our model.

In the second part of this project we are trying to use the X-Ray images itself instead of the metadata associated with them. We are using deep learning and CNN to train our model and than using the test set to validate our model and find accuracy for new x-rays images.

At last we will compare both the first way and second way to detect whether someone has covid or not. This will increase the reliability of our project as we are using both the image as well as text data to be sure.

This report on programmed COVID-19 identification is for instructive purposes as it were. It isn't intended to be a solid, profoundly exact COVID-19 conclusion framework, nor has it been expertly or scholastically reviewed.

1.2.2 Problem Statement explained

In this study, an in-depth study model for the planned determination of COVID-19 is proposed. The proposed model has the potential to eliminate engineering without the use of any element extraction techniques, and requires X-beam crude imaging images to restore the findings. This model is remodeled with 125 X-beam chest images, which are not in the standard format and were quickly detected. Symptomatic tests performed after 5–13 days were found to be healthy in patients who recovered [9]. These important findings show us that recovered patients can continue to spread the infection. For these lines, direct access methods are needed. One of the main drawbacks of chest radiography tests is the inability to identify the early stages of COVID-19, as they do not have a sufficient effect on GGO detection [8]. In any case, in-depth learning models revolve around the focus of the unseen on the natural eye, and can work to transform this awareness.

1.3 Objectives

The objective of the report is:

- To develop a model that predicts whether someone has covid-19 or not.
- This model will work with the clinical text reports first to extract the relevant features related to covid and on basis of these features it will predict whether some future clinical report corresponds to someone with covid-19 or not.
- Also it will analyze the X-Ray images using deep learning techniques to predict for some future X-ray image.
- Finally we will compare the results from both these methods and what works best in our case.

1.4 Methodology

The methodology that I have followed for this model is as below:

- First preprocess the metadata associated with the X-Ray images and extract the clinical reports from them. After that use BOW and TF/IDF approach to extract the relevant features and associate each extracted feature with some probabilistic value.
- Now use those probabilistic values to implement various ML algorithms like multinomial Naïve Bayes, SVM and decision tree to decide to which class some future clinical report belongs to. For this case it will classify the data into 3 classes: Covid-19, SARS and ARDS. Finally we will find the accuracy, precision and other judging factors.
- Next we are using the X-Ray images itself and deep learning techniques like CNN, pooling, flattening, transfer learning to first train our model and then predict for some future X-Ray image. For this we are using Keras library.
- At last we are comparing results for both the models and check which one works best for our cause.

1.5 Organization

We will use various machine learning algorithms and deep learning methods to predict whether someone has covid or not. Along with it we are using approaches like BOW, TF/IDF, CNN, Pooling, flattening, transfer learning that helps in our problem statement. Finally we will compare the both proposed methods and check which one works best for us. It is important to mention it is just a predicting model. One can not fully rely on it but still it can help radiologists and doctors with there work.

CHAPTER 2

LITERATURE SURVEY

KEYWORDS

COVID-19(Coronavirus), Deep learning, Convolution, CNN, X-Ray Chest pictures, Radio graphical pictures

2.1 INTRODUCTION

AI and regular language preparation designs use a wide range of information-based models for receipt, clarification and expectation. The NLP has recently gained a lot of interest, usually in the field of text examination, as one of the most important companies in taxonomic material mining and can be done using various calculations [6]. Kumar et al. [6] SWOT research of various controlled and non-aided material agglomeration calculations has been carried out to dig up structured information. Various uses of text characterization are being used extensively for inference research, misrepresentation and spam detection, and subsequent evaluation mining races, advertising, business and more, Verma et al. [7] With the help of vocabulary-based glossary, he examined the attitudes of Indian government agencies. AI has changed the approach to innovation by giving extraordinary results to infections such as diabetes and epilepsy. Emperor et al. [4] Epilepsy is detected using an AI draw, and electroencephalogram (EEG) signals are used to identify specific and epileptic conditions using simulated nerve associations (ANNs). Sarwar et al. [10] Diabetes detection results using AI and dress learning strategies demonstrated that the group method guarantees 94.80% accuracy. These factors can help diagnose and diagnose COVID-19. The firm and precise intent of COVID-19 saves many lives and provides a large amount of information that can build AI (ML) models.

2.2 ML based methods/approaches for Coronavirus detection

ML may give helpful contribution in this case, the analysis is based exclusively on the clinical material in the structure, radiography images and etc. according to Bail et al. [11], machine learning and in-depth practice can replace people by making accurate decisions. Ideal search saves radiologists time and is financially sensitive compared to standard tests of COVID-19. X-beam and registered tomography (CT) sweeps can be used to design AI models. In such a situation some activities are going on. Wong and Wang [12] created the COVID-NET, a deeply convoluted neural network that can detect COVID-19 from the chest.

radiography Pictures. When COVID-19 is detected in a person, the extent to which that person is severely affected is investigated. Not all COVID-19 positive patients need to be fully considered. Having the option of retrieving who is more severely affected will help coordinate assistance and establish clinical asset allocation and utilization.

Yann et al. [13] At Tongji Hospital in Wuhan, China (only) AI was used to construct an estimated prognosis to solve a person's risk of death using data from 29 patients. Jiang et al. Proposed an AI model that could assess a person infected with COVID-19 and cause severe respiratory distress (ARDS). The proposed model brought 80% accuracy. Examples of 53 patients were used to create their sample and these were limited to two Chinese clinics. ML can be used to diagnose COVID-19, which requires a lot of research effort and has not yet worked in general. Because less work is being done on the determination and evaluation of text usage, we have used AI and practice models to classify clinical reports in four classes of infections.

2.3 Automatic detection of COVID positive cases with the help of Deep Neural Networks with X-ray pictures

A few investigations have experienced changes in chest X-beam and CT images before the onset of COVID-19 manifestations [18]. Major disclosures have been accepted by researchers in the imaging investigation of COVID-19. Kong et al. [6] note the right opacities of the infrahilar airspace in COVID-19 patient. Yoon et al. [19] indicated that one in three patients considered had a unique problem of kneeling privately in the lower left lung. Interestingly, the other two had four or five unexpected spots in the two lungs.

Zhao et al. [16] not only did the discovery of opacities of the ground glass (GGO) or GGO included in most patients, but they also noted the union, and the increase in vascular permeability. Li and Xia [17] have identified GGO and combination, interlobular septal thickening and ventricular bronchogram signal, with or without vascular growth, as standard CT patients for COVID patients- 19. Fringe central or multifocal GGO influencing two lungs in half of 75% of patients with one view [9]. In particular, Zu et al. [8] and Chung et al. [6] found that 33% of chest CTs can correct lung opacities.

In Fig. 1, X-beam chest images taken on days 1, 4, 5 and 7 of a 50-year-old COVID-19 patient are diagnosed with pneumonia, and specifications are provided of these images [20].

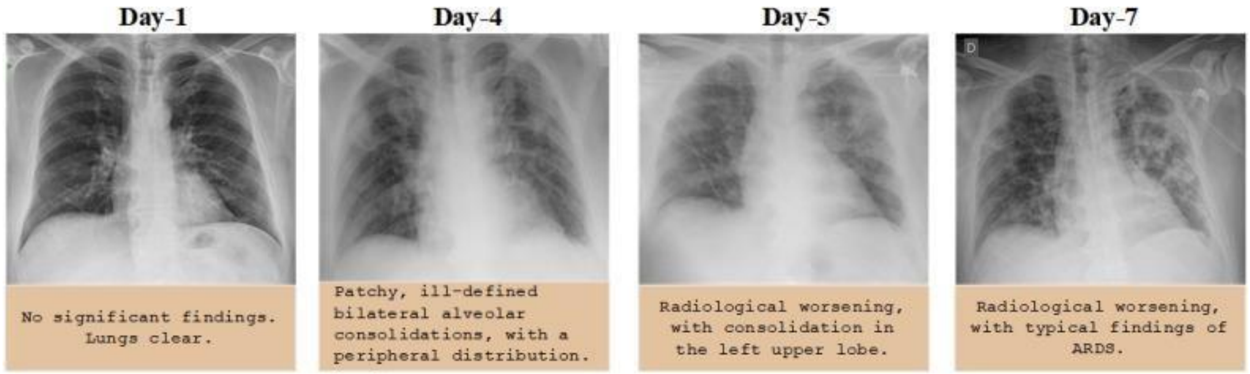


Fig 2.1 Chest X-Ray pics of a patient 50 years old having Pneumonia and Covid-19 over 1 week.

Use of AI strategies for programmed finding in the clinical field have as of late picked up ubiquity by turning into an aide apparatus for clinicians. Profound realizing, which is a famous examination region of man-made consciousness (AI), empowers the making of start to finish models to accomplish guaranteed results utilizing input information, without the need for manual extraction [26,27]. In-depth learning processes have been successfully applied to numerous issues, for example, arrhythmia location [[28], [29], [30]], skin malignancy order [31,32], bosom disease discovery [33,34], mind infection characterization [35], identification of pneumonia from chest X-beam images [36], fundus division division [37], and lung division [38,39]. The rapid rise in COVID-19 disease requires the need for qualification in this field. This has increased the interest in building computerized identification frameworks dependent on AI procedures. It is a moving assignment to give master clinicians to each medical clinic because of the set number of radiologists. Consequently, basic, exact, and quick AI models might be useful to conquer this issue and give convenient help to patients. Despite the fact that radiologists assume a critical part because of their huge involvement with this field, AI expertise in radiology can assist in accurate diagnosis [40]. Also, AI methods can help to overcome barriers, for example, shortages of RT-PCR test units, test costs, and trial time outcomes.

As of late, numerous radiology pictures have been broadly utilized of the COVD-19 area. Hemdan et al. [41] used in-depth study models to analyze COVID-19 in X-beam images and proposed a COVIDX-Net model that included seven CNN models. Wang and Wong

[42] proposed a comprehensive local model of COVID19 (COVID-Net), which found 92.4% clarity in the management of common, non-COVID pneumonia, and COVID-19 classes. Ioannis et al.

[43] developed a comprehensive learning model using 224 certified COVID-19 images. Their model makes up 98.75% and 93.48% of the two- and three-phase development rates, respectively. Narin et al. [44] achieved 98% recognition of COVID-19 using X-beam chest images combined with the ResNet50 model. Sethy and Behera [45] edited highlights from various models of convolutional neural organization (CNN) with the help of a vector-assisted (SVM) machine using X-beam images. Their investigation reveals that the ResNet50 model

with help of Scaler vector machine classifier gave excellent execution. At long last, there are likewise a few ongoing investigations on COVID-19 identification that utilized different profound learning modelswith CT pictures.

Each prepared neural organization picks up information for the specific errand that is thought of. While the fundamental rule of fake neural organizations regenerating human behavior and understanding, alternating learning in non-neural organizations is used to use put away information on a specific errand for another connected undertaking. Profound learning for picture acknowledgment applications is equipped for learning a huge number of pictures, and a few tremendous models are prepared with various techniques.13-17 These pre-prepared models are freely distributed so that all scientists can use the prescribed data. Frames with pre-prepared edges that are publicly accessible, in particular, VGG16,13 VGG19,13 ResNet50,14 Launch V3,15 MobileNet-V2,16 and Densenet121,17 viewed examination.

2.4 Convolution Neural Network

The Convolutional neural Organization (CNN) is one of the best and most effective means of concluding COVD-19 from computer-generated images. A few studies have completed a late commitment to the detection of COVID-19 [8, 15, 24]. For example, in [35], CNN was used relying on the Inception organization to identify COVID-19 disease within CT. In [29], a modified translation of the pre-prepared organization of ResNet-50 was given to visualize CT images in three categories: solid, COVID-19 and bacterial pneumonia. The CXR used in [23] by CNN has been developed based on pre-programmed models of ImageNet to remove high-level points. Those highlights have been taken care of in SVM as an AI detector to detect COVID-19 cases. In addition, in [34], CNN engineering called COVID-Net was used based on interchangeable learning to classify CXR images into four categories: general contamination, viral contamination, non-COVID and COVID-19. In [4], a database of CXR images from patients with pneumonia, confirmed COVID-19 disease, as well as normal exposure, was used to evaluate the cutting edge convolutional neural organization models proposed already for clinical picture grouping. The examination proposed that move learning can separate huge highlights identified with the COVID-19 infection.

2.5 CONCLUSIONS

In this study, a comprehensive learning model for the planned acquisition of COVID-19 is proposed. The proposed model has a start to complete construction without the use of any extraction techniques, and requires X-beam crude imaging images to regain determination. This model is remodeled with 125 X-beam chest images, which are not in the custom building and were quickly detected. Demonstration tests performed after 5–13 days were found to be healthy in recovering patients [52]. This important finding shows us that patients are recovering it can continue to spread the infection. After that, specific approaches to the conclusion are required. One of the most dangerous risks in chest radiography screening is the inability to differentiate the initial stages of COVID-19, as they do not have sufficient implications for GGO detection [8]. In addition, all in-depth learning models are geared towards getting into focus that is not visible to the natural eye, and can work to transform this awareness.

CHAPTER 3

SYSTEM DEVELOPMENT

The development will be divided into two phases or approaches. First we will use the metadata for X-Rays and we will do text mining on that data. This data consists of many different columns like name, age, sex, clinical reports and disease. Basically we are only interested in the clinical reports and the disease to which that report corresponds to. So we will extract these 2 features from the metadata. After that we will do preprocessing on that data and clean that data. By cleaning we mean to remove unnecessary words from the clinical reports. Stop word removal, punctuation removal, lemmatization will be done. Than we will obtain many words which will act as relevant features in the future but we have to choose 10-12 important features from 40 -50 extracted features. To do so we use approaches like BOW (Bag of words) and TF/IDF. These techniques help to tell the relevance of some word in a complete text. Once we get all the important features we will try to attach some probabilistic values with these features as we have to use them as basis for different algorithms. Once everything is done we will finally use different machine learning algorithms like multinomial Naïve Bayes, SVM, Decision Tree to train our model and than validate it against the test set. We will than find out different factors like accuracy, loss, precision etc. At last we will compare all the algorithms and check which one works best for us.

After analyzing all the metadata, to improve the predictable power of this model we will analyze both covid-19 positive and normal X-Rays. We will use Deep Learning to build a model which will be trained on these X-ray images. We will use more than one convolution layer to increase the receptive field (discussed later). CNN will be used to extract different features form the X-ray images. Along with convolution layers we will also use pooling layers so as to increase the receptive field. Finally the resulting image set after convolution will be flattened and fed to a neural network where it will be classified and checked whether a new X-ray image is related to covid-19 positive person or a normal person. Confusion matrix and model summary will be generated to check how efficient our model is working.

Finally when both the above approaches will be checked against new data, we will compare both of them in terms of accuracy and loss. The best approach for our model will be analyzed.

Types of Approach :-

1) Machine learning dependent technique for detection of COVID-19 using medical textual data.

2) Automatic detection for COVID positive cases using deep neural network with X-Ray pics.

3.1. Machine learning dependent techniques for detection of COVID-19

The proposed system comprises of 2.1 to 2.5 developments. Synchronization of information integration 2.1 is done and 2.2 defines information refinement, 2.3 provides a framework for further development, 2.4 provides an integral part of the output. In E there is talk of AI statistics, and 2.5 provides a framework for group AI statistics. Visual representation of the proposed process by appeared in Fig. 2. also, are being examined beneath.

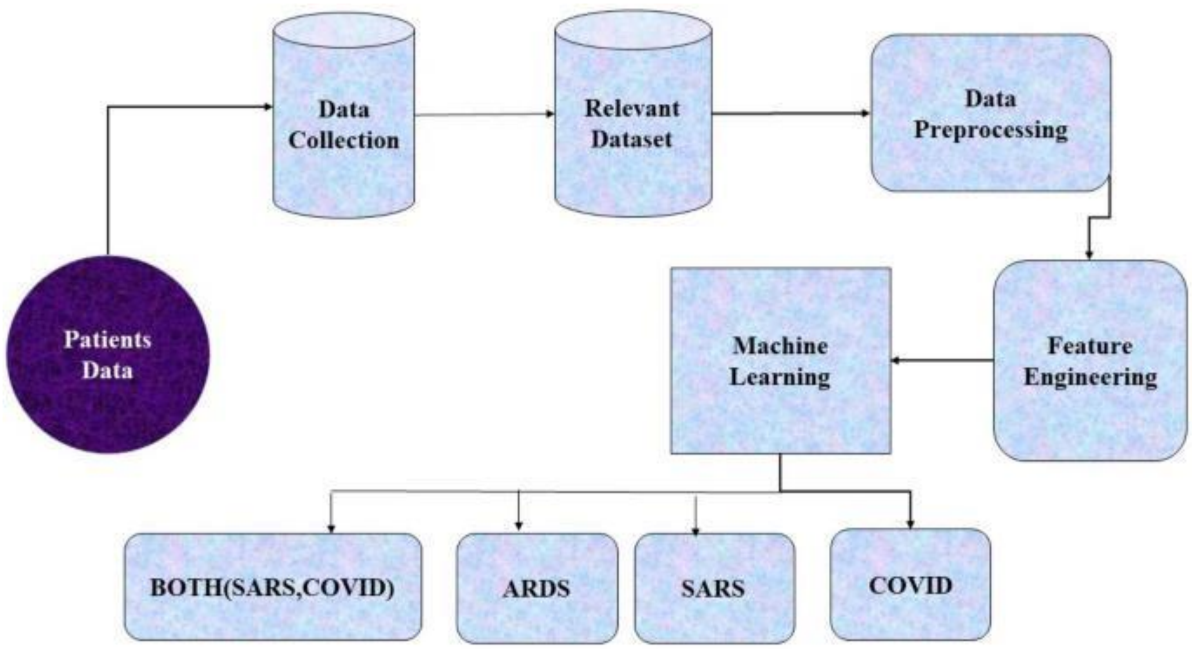


Fig 3.1 Methodology for ML based approach for detecting COVID-19

3.1.1. Data Collection

As World Health Organisation pronounced Coronavirus as a Health Emergency. Scientists and medical clinics offer open access to information about the epidemic. We have compiled GitHub.Footnote1 from which 212 patient information is deleted which shows the side effects of Covid and other diseases. The data contains approximately 24 specifying a straight, moderate id, sex, age, acquisition, tolerance, intubated, requirement of icu, spo2 requirement, extubate period, body temp., pO2's sat., count for leukocyte, check for neutrophil, lymphocyte tally, see, methodology , area, organizer, file name, basic doi's and urls. Permission. Clinical textual report notes provided by doctor and separate notes.

3.1.2 Necessary/Relevant data

As you know our doing is about the extraction of messages so we have separated clinical notes from the findings. Clinical notes contain text while discovering material containing a related text mark. A total of 212 reports were used to determine its length. We look at those reports and try to spereate it out in different categories as shown in the graph below. Here we are trying to detect the report length of all the reports and than trying to classify into four different categories which are proposed in this model namely Ards, sars, both and covid.

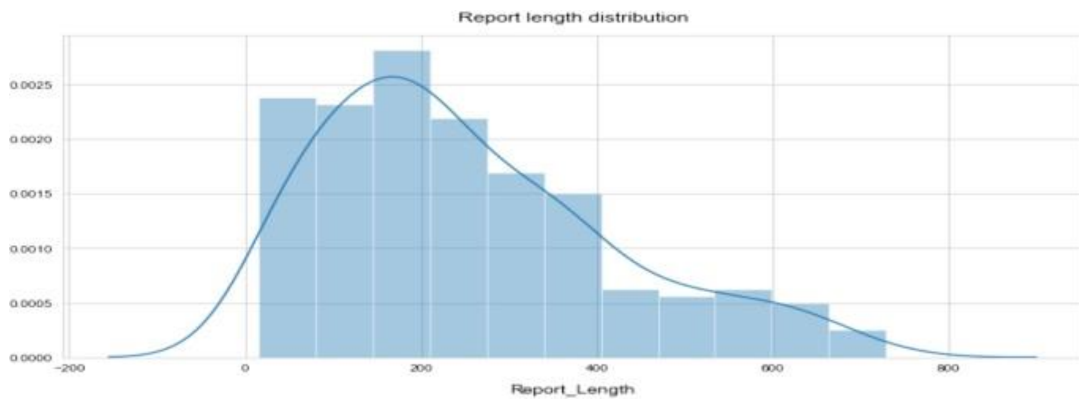


Fig 3.2 Report Length for clinical text report

If we talk about different classes, they will have different report lengths. It is understandable that bigger a report length more the features extracted from it. So it is important to visualize the report lengths for various classes (diseases) that we are trying to predict. Below image shows report length range v/s different classes that we are working on.

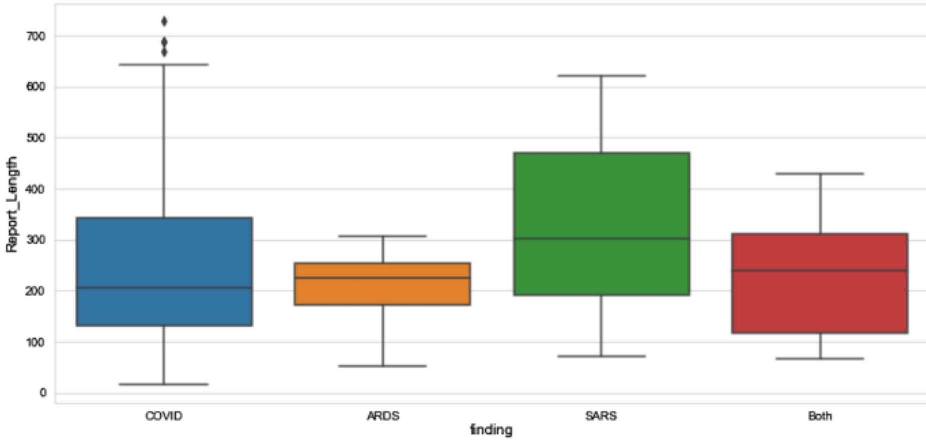


Fig 3.3 Various classes with their respective report length ranges

3.1.3. Preprocessing

The content was unstructured and therefore had to be cleared for the purpose of keeping the AI from happening. Different developments continue at this stage; content is cleared by removing unnecessary content. Emphasis on language is done with the ultimate goal of having the information refined to the highest standard. Stopwords, images, Url's, joins are removed to the point where the collection might be achieved with better precision or accuracy. Fig 3.4 shows the key steps in moving forward.

Clinical_notes	Finding	Report_Length	Punctuation	Lemmatisation	Stop_Word Removal
infiltrate in the upper lobe	COVID	45	infiltrate in the upper l	infiltrate in the upper lobe	infiltrate upper lobe leave lung
progressive infiltrate and	COVID	40	progressive infiltrate a	progressive infiltrate and c	progressive infiltrate consolidation
progressive infiltrate and	COVID	40	progressive infiltrate a	progressive infiltrate and c	progressive infiltrate consolidation
progressive infiltrate and	COVID	40	progressive infiltrate a	progressive infiltrate and c	progressive infiltrate consolidation
diffuse infiltrates in the bil	COVID	48	diffuse infiltrates in th	diffuse infiltrate in the bila	diffuse infiltrate bilateral lower lungs
progressive diffuse inters	COVID	115	progressive diffuse in	progressive diffuse interst	progressive diffuse interstitial opacities consolidation
Severe ARDS. Person is in	ARDS	53	severe ards person is	severe ards person be intu	severe ards person intubate og place
Case 2: chest x-ray obtain	COVID	563	case 2 chest x-ray obt	case 2 chest x-ray obtain o	case 2 chest x-ray obtain jan 6 (2a) brightness lungs
Case 2: chest x-ray obtain	COVID	563	case 2 chest x-ray obt	case 2 chest x-ray obtain o	case 2 chest x-ray obtain jan 6 (2a) brightness lungs
SARS in a 74-year-old mar	SARS	71	sars in a 74-year-old r	sars in a 74-year-old man	sars 74-year-old man develop symptoms 4 days exp
SARS in a 74-year-old mar	SARS	71	sars in a 74-year-old r	sars in a 74-year-old man	sars 74-year-old man develop symptoms 4 days exp
SARS in a 74-year-old mar	SARS	71	sars in a 74-year-old r	sars in a 74-year-old man	sars 74-year-old man develop symptoms 4 days exp
SARS in a 29-year-old wor	SARS	378	sars in a 29-year-old v	sars in a 29-year-old wom	sars 29-year-old woman present 7 days exposure ()
SARS in a 29-year-old wor	SARS	378	sars in a 29-year-old v	sars in a 29-year-old wom	sars 29-year-old woman present 7 days exposure ()
SARS in a 42-year-old wor	SARS	145	sars in a 42-year-old v	sars in a 42-year-old wom	sars 42-year-old woman present 9 days exposure pc

Preprocessed data set

Figure 3.4 Dataset after preprocessing

3.1.4. Feature designing

Now we have got the clinical reports after preprocessing them, distinctly different are taken according to factors chosen and converted numerical values which are based on probabilties. We use Transfer Frequency//Inverse Document Frequency strategy for removing significant highlights. Pack of words was likewise contemplated, unigrams, bigrams were additionally separated. We distinguished 40 applicable highlights by which the order can be executed. These highlights appear in Figure 3.5. By giving weight to compare the object and then giving the same details AI calculations

lungs	chest	patient	multiple	peripheral	bilateral	lower	lung	leave	image	lob	opacities	ct	right	lobe	air	pneum	glass opacities	history
0.379	0	0	0	0	0.34539	0.373	0.379	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0.612	0	0	0	0	0	0	0	0	0.45	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0.31916	0.241498	0.13287	0	0	0.24	0	0	0.119823	0	0	0	0	0.16	0.256674173	0.223
0	0	0	0	0	0.45603	0	0	0	0.573	0	0	0	0	0	0	0	0	0
0.342	0	0	0	0	0	0.336	0.342	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0.50247	0.298721	0	0	0	0	0	0	0	0	0	0	0	0	0.393524911	0
0	0	0	0.26141	0	0	0	0	0	0	0	0	0.3	0.357	0.3	0.3	0	0	0
0	0	0	0	0	0	0	0	0	0	0.41	0	0	0	0	0	0	0	0
0	0	0.2225	0	0.340237	0	0	0	0.2	0	0.17	0	0	0	0	0	0	0	0.314
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.481
0	0	0	0.28544	0	0	0	0	0	0	0	0	0.3	0.39	0.4	0.4	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.317
0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.282	0.3	0.3	0	0	0.238

Features are chosen/selected for classification

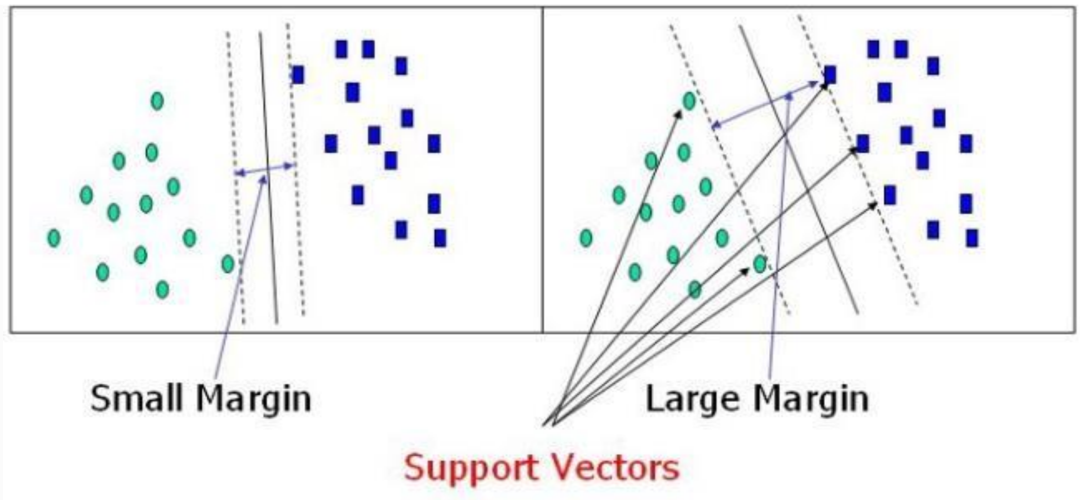
Fig 3.5 Feature Selection

3.1.5. Machine learning based classifications

Finally after getting all the required features we can use different machine learning algorithms to do the classification work. Here we are going to use the multinomial naïve bayes and the scaler vector machine. Apart from that we can use other ensemble machine learning algorithms too. We will compare these algos on some factors like precision, f1 score, accuracy and confusion matrix. Finally we will see which algo works best for us in this particular case.

3.1.5.1. SVM

So here in case of SVM we will try to draw a hyperplane between two different categories. Dimensionality can be increased to achieve that. We will find the margin of the nearest point from both the categories. Based on that a hyperplane will be chosen. There can be many partitions between two classes. We have to choose such a hyperplane so that the margin between both the categories is maximized. Finally that we can divide the training points into 2 separate categories and test for the validation data.



3.1.5.2. Multinomial Naïve Bayes(MNB) algorithm

MNB values are the values of book categories provided through the Bayes Act.Let C means class order in our difficulty we have four categories C = 0, 1, 2 and 3. 40 (40 highlights taken using TF / IDF). At thatpoint MNB relegates ti which is test text to the respective class which has the most noteworthy likelihood P(c|ti) probability by utilizing Bayes theorem demonstrated as follows:-

$$P(c|t_i) = \frac{P(c)\,P(t_i|c)}{P(t_i)}, \quad c \in C$$

3.2. Automated detection with help of Deep Neural Network with X-ray pictures:-

3.2.1.X-ray pictures Data Set

We have taken data from 2 different datasets. One is the James Hopkin university data in which both the X RAY and the metadata was provided. If I particularly talk about the X ray than only

The data for covid positive patients was provided. So for the normal person's chest xray I used the data from keggle. Basically I used a 50-50% data from both these datasets and finally created a new dataset to use in the model. Below are the some images of the X Rays that are present in my dataset.

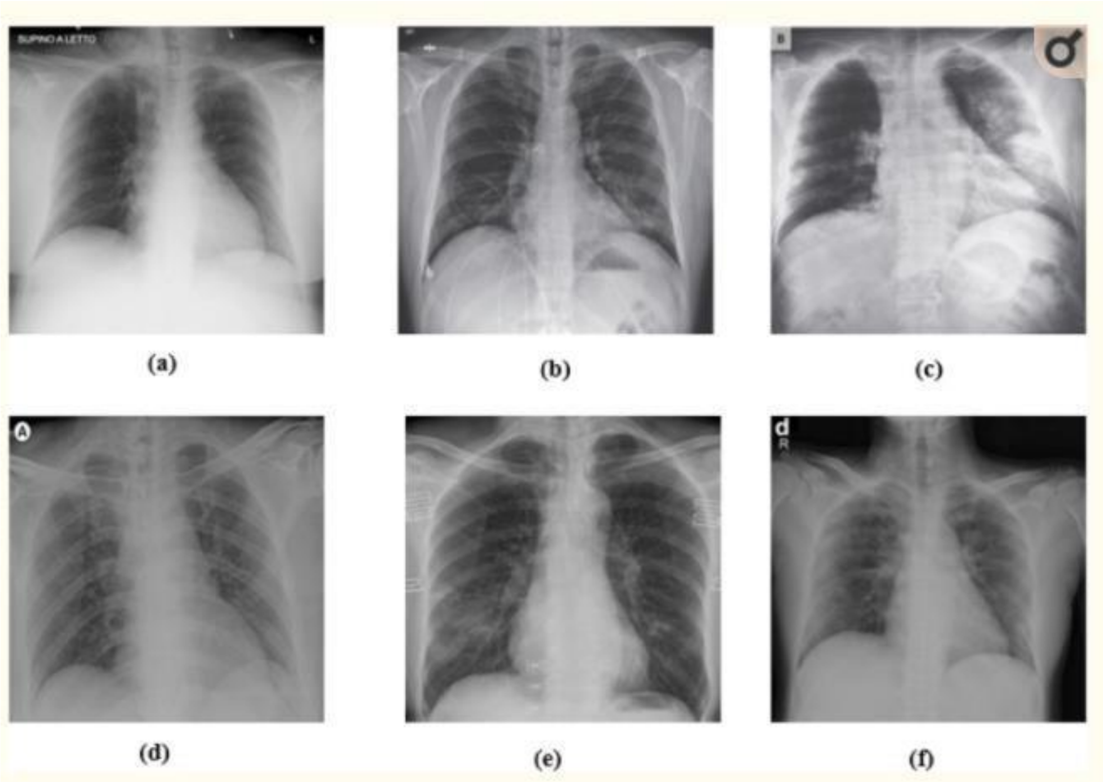


Fig: 3.6

A few cases of COVID-19 and the findings of the database: (a) Cardio-vasal density within the cutting points, (b) Increased basilar left ventricular lesion is obvious, pleasurable concern with pneumonia, (c) Continued infiltration, d) A small combination of the upper right roof and ground glass in both lower projects, (e) The infection indicates the opacities of the right infrahilar airspace, and (f) Continuity of subtle attacks associated with perihilar attacks and abnormal light of the lungs on both sides.

3.2.2. Model designed for this project

The new learning process irritates human thinking. The most profound term refers to the growth in size of this organization through the mass of layers. The structure is named after convolution, the numerical manager. CNN's typical architecture has a convolution layer that focuses on highlights from the provision of channels it uses, the integration layer to reduce the size of computer hacking, and a fully integrated framework, which is a neural organization. By combining at least such layers, a CNN model is created, and its internal parameters are assigned to fulfill a specific assignment, for example, editing or object information. Instead of initiating the most profound development of a model without preparation, a more logical way to build a model using models is shown successfully. Therefore, when designing the in-depth model used in this investigation, the Darknet-19 model is considered as the first phase. Darknet-19 is a differentiating model that lays the foundation for a continuous object recognition framework called YOLO (You only look once).

This framework has a sophisticated engineering designed to identify an object. The DarkNet separator is used based on this functional design. We planned engineering, encouraged by building DarkNet that focused on in-depth learning, rather than building a model without preparation. We used lower layers and channels compared to the original DarkNet architecture. We continued to increase the size of the channels, for example, to 8, 16, 32. To better understand this new model, it is helpful to understand the basics of Darknet-19, which consists of 19 layers of decisions and five layers integration, using Maxpool. These layers are standard CNN layers with various channel numbers, sizes, and step parameters. Let the letter C indicate the convolutional layer, and M mean the Maxpool layer. Since C 1 is considered a layer of information, Darknet-19 has the following layer format.

The basic convolution function in mathematics is given by the below formula. Here we are doing a convolution operation between X and K: -

$$(X * K)(i, j) = \sum_{m} \sum_{n} K(m, n) X(i - m, j - n)$$

here * speaks to the convolution activity(discrete). K structure slides on the information lattice at the phase boundary. The Leaky Modified Straight Unit (Leaky Rail) is used as an activation task.

Darknet design. Defective relu work is calculated as follows:

.
$$f(x) = \begin{cases} 0.01x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$$

The below image shows the representation of the model I have worked on. Different convolution layers and max pooling layers are clearly visible in the model.

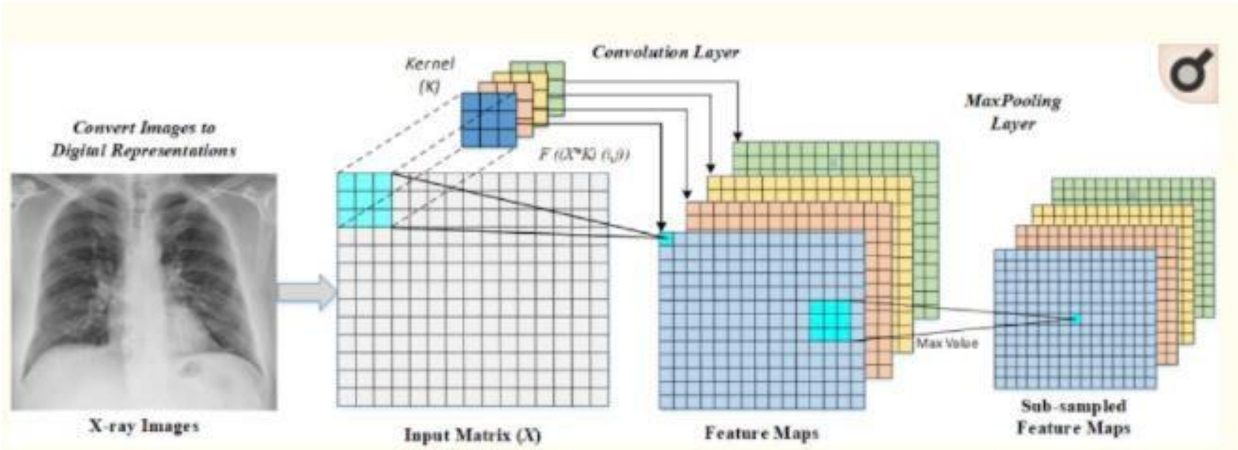


Fig.3.7 : A representative introduction of Max-pooling layer activities and convolution.
The model yields with an average pool and soft max layers. In this research, we encountered the problem of sorting images with obvious nuances. The model that makes such a setting, unlike the Resnets or Resnext [60] models, must have a structure that can capture and learn even minor conflicts that are exceptionally intensive. The proposed model used in this test is illustrated in Figure 3.8.

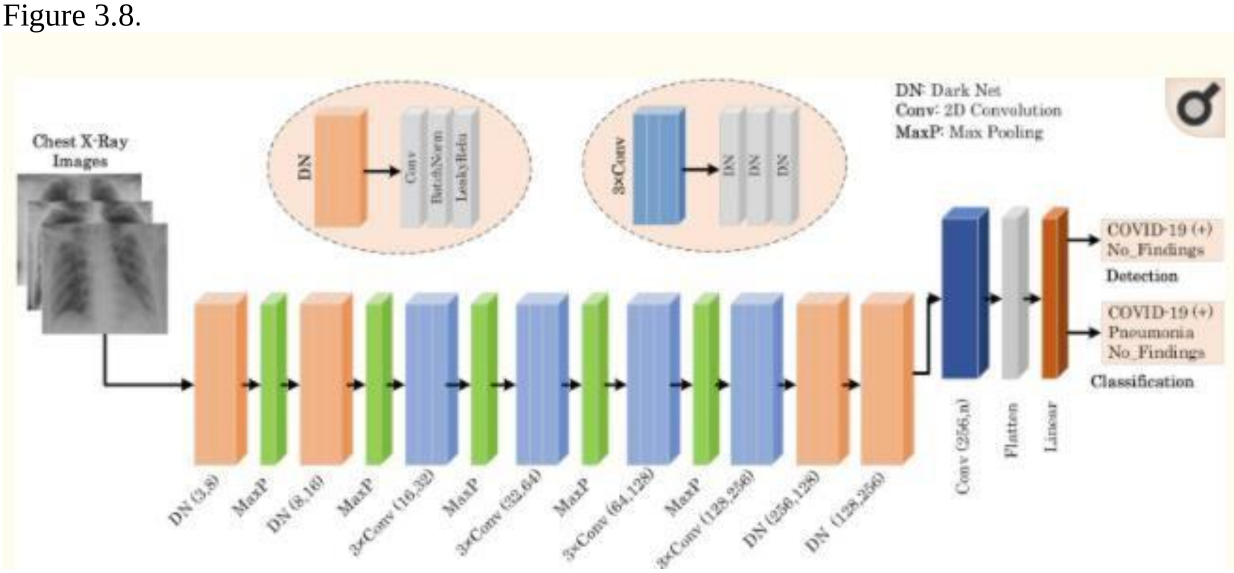


Fig:3.8 The schematic representaion of the devised model ().

The below table clearly shows the different layers used in this model. The output shape, layer type, number of layers and the parameters and the parameters at each layer is clearly visible in the table. One thing here to observe is that as we go deep into the neural network the number of channels as well as the number of parameters increases which is quite understandable. If it is possible to use three different photo classes for information, the same model plays an order assignment to determine the names of X-beam chest images such as COVID 19, Nil, and Pneumonia. We are using the approach of Adam for remodeling agent to regain weight, cross the entropy business function and chose the level of learning as

3e-3

Table 1

The layers and layer parameters of the proposed model (for the binary classification task).

Number of Layer	Layer Type	Output Shape	Number of Trainable Parameters
1	Conv2d	[8, 256, 256]	216
2	Conv2d	[16, 128, 128]	1152
3	Conv2d	[32, 64, 64]	4608
4	Conv2d	[16, 66, 66]	512
5	Conv2d	[32, 66, 66]	4608
6	Conv2d	[64, 33, 33]	18,432
7	Conv2d	[32, 35, 35]	2048
8	Conv2d	[64, 35, 35]	18,432
9	Conv2d	[128, 17, 17]	73,728
10	Conv2d	[64, 19, 19]	8192
11	Conv2d	[128, 19, 19]	73,728
12	Conv2d	[256, 9, 9]	294,912
13	Conv2d	[128, 11, 11]	32,768
14	Conv2d	[256, 11, 11]	294,912
15	Conv2d	[128, 13, 13]	256
16	Conv2d	[256, 13, 13]	294,912
17	Conv2d	[2, 13, 13]	4608
18	Flatten	[338]	0
19	Linear	[2]	678

CHAPTER 4

PERFORMANCE ANALYSIS

Types of Results:-

1) ML dependent techniques for detection of COVID-19 using clinical textual data.

2) Automatic detection of COVID positive cases using neural network with X-ray pictures.

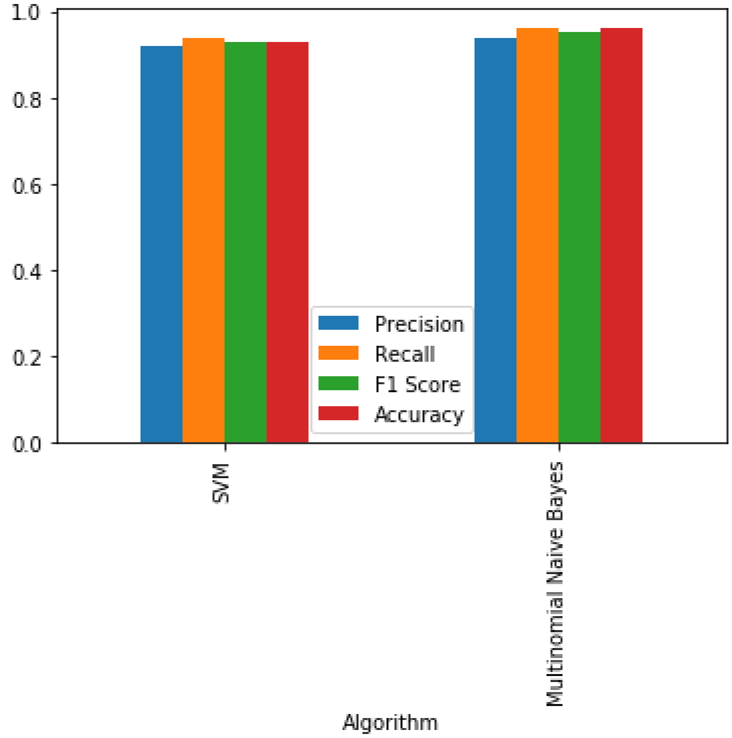
4.1. Machine learning based approaches results:-

We used a window frame with 4 GB Ram and a 2.3 GHz processor to play this function. Skiket learning tools are used to perform AI editing with the help of various libraries such as NLTK, STOPWORDS etc. to improve the accuracy of all AI computation pipelines are used. After playing the measurable number, more information is achieved. Data is part of the 70:30 scale where 70% of the data is used to model the model and 30% is used to test the model. We have content reports from clinics for 212 patients marked in four categories.

Editing is completed using AI statistics by providing them including extracts from the pre-design object. Investigating our model’s speculation from information processing to encrypted information and minimizing the possibility of overcrowding, we disassociate our basic data from the preparation and testing of independent subsets. The ten approved approaches were targeted at all statistics, and this cycle was repeated several times autonomously to maintain a strategic distance from the inspecting inclination presented by arbitrarily dividing the dataset in the cross-validation.

After performing the character, it was found that SVM and Multinomial Naive Bayesian classifier give good results. while executing SVM are 0.92,0.94,0.93,0.93 respectively. category where all the information was used for testing. So we can assume that the more information is provided by these statistics, the more likely it is to improve performance.

Medical data and take important steps. Similarly, it was classified that kovid positive patients report significantly longer doses than the various classes and ranged from 125 lists to 350 characters.



4.2. Automatic detection of COVID + cases with the help of deep neural network with X-ray image results:-

The presentation of the proposed model was tested using a 5-crease strategy to allow for deviations from both similarities and triple descriptions. Most X-beam images are used for preparation and 20% are approved. Investigations are updated as often as they appear in the image. A fragmented fragment is grouped into folders for use in the permission section. We are ready for 100 years. In Fig. 4.2 Preparation and approval of new multi-class chart charts and approval diagrams for approval appeared in Fold-1.

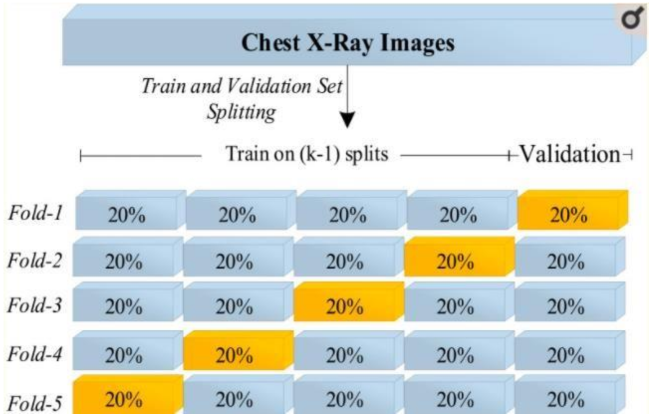


Fig. 4.1 Schematic portrayal of preparing and approval plot utilized in the 5-overlap cross-approval strategy.

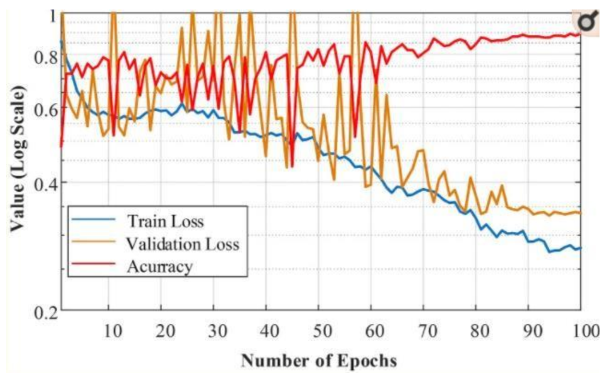


Fig.4.2 Train Loss, Validation Loss and Accuracy bends acquired for Dark Covid-Net model in case of fold-1.

It is usually recognized from Fig. 4.2 that there is a significant increase in unfortunate rates at the beginning of the preparation, which is greatly reduced in the later stage of preparation. The main explanation behind this sharp rise and reduction is given in detail in the class of COVID-19, which is by no throughout the age of preparation, these positive and negative moments gradually diminish in the next phase of preparation.

The use of a set of different components of the model was assessed for each increase, and the performance of standard modeling was determined. Covered as per separate network (CM) appears in Fig. 4.3. Covered CM is made using the number of CMs that have entered all the folders. Hence, it is expected to acquire a thought regarding the overall holes of the model. The model accomplished a normal grouping exactness of 87.02% to order: no discoveries, COVID-19, and Pneumonia classifications. Affectability, explicitness, exactness, F1 scores, as well as accuracy ratings from Table 2 investigate the details of the 3rd class problem model.

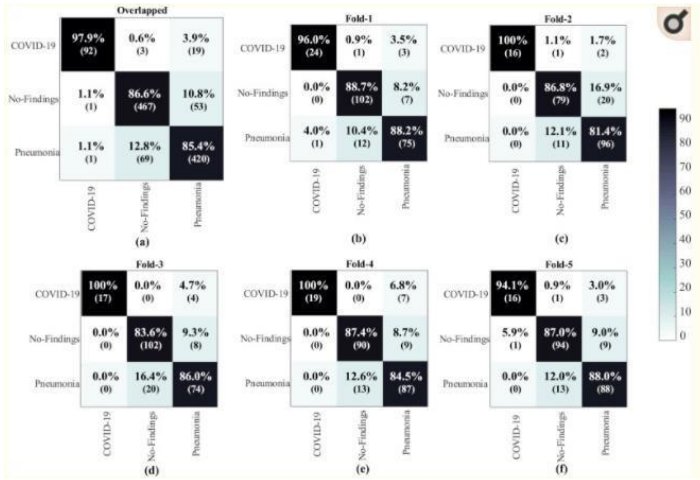


Fig 4.3 The confusing results of the missed and five-fold confusion of multi-class segregation work: (a) scattering matrix, (b) CM Fold-1, (c)CM Fold-2, (d) CM Fold-3, (e)CM Fold-4, and (f)CM Fold- 5.

Table 2

Sensitivity, specificity, precision, F1-score, and accuracy values obtained for each fold of the proposed model.

Folds	Performance Metrics (%)				
	Sensitivity	Specificity	Precision	F1-score	Accuracy
Fold-1	88.17	93.66	90.97	89.44	89.33
Fold-2	84.57	90.61	89.38	86.63	84.89
Fold-3	84.13	91.14	89.88	86.54	85.78
Fold-4	83.66	92.29	90.61	86.42	87.11
Fold-5	85.83	92.75	89.71	87.57	88.00
Average	**85.35**	**92.18**	**89.96**	**87.37**	**87.02**

It is generally observed from a separate network for the task of grouping group collections that the in-depth study model outlined COVID-19 better than pneumonia categories and no results were found. The findings, specifications, and F1-score esteems were 85.35%, 92.18%, and 87.37%, respectively.

Also, the effect of the disruption of the fracture problem of the positive separation of COVID-19 appeared in Fig. 4.4. In addition, the implications, precise specifications, F1 results, and the results of the accuracy of the duplicate function are given in Table 3.

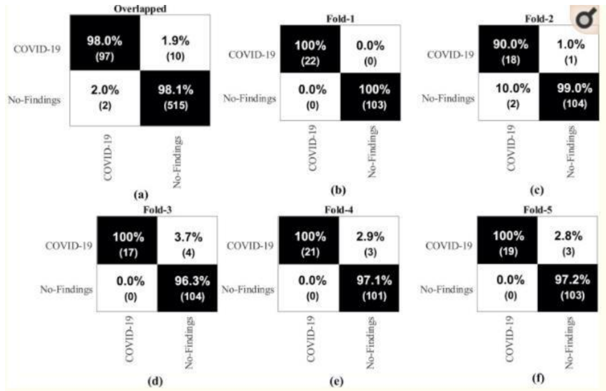


Fig.4.4 The effects of the dispersion and multiplication of five times the function of binary categories. -4 CM, and (f) Fold-5 CM

Table 3

Sensitivity, specificity, precision, F1-score, and accuracy values for No findings and COVID-19 classes of the proposed model.

Folds	Performance Metrics (%)				
	Sensitivity	Specificity	Precision	F1-score	Accuracy
Fold-1	100	100	100	100	100
Fold-2	96.42	96.42	94.52	95.52	97.60
Fold-3	90.47	90.47	98.14	93.79	96.80
Fold-4	93.75	93.75	98.57	95.93	97.60
Fold-5	93.18	93.18	98.58	95.62	97.60
Overlapped					
COVID-19	90.65	99.61	97.97	94.17	98.07
No-Findings	99.61	90.65	98.09	98.84	98.07
Average	**95.13**	**95.3**	**98.03**	**96.51**	**98.08**

We can see from the Table 3 that the supposed model has accomplished a normal precision of 98.08% in identifying COVID-19 and the got normal affectability, particularity, and F1-score estimations of 95.13%, 95.30%, and 96.51%, individually.

Wrong. Therefore, patients diagnosed with COVID-19 are diagnosed with pneumonia (see Figure 4.5 (a)). For this reason, the success rate of the model in the multi-class classification problem

Relatively less than binary class.

.

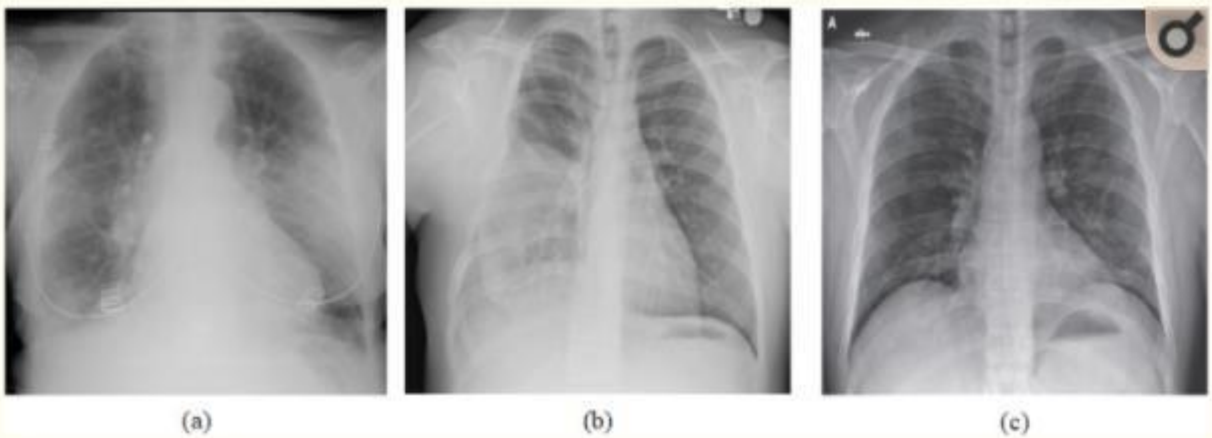


Figure 4.5

Images assessed by radiologists and models: (a) Model pneumonia was diagnosed, although the actual class was COVID-19, (b) Model pneumonia was diagnosed, but the actual class was not detected,

The model is sharp in detecting pneumonia infections. The model can be severely diagnosed with pneumonia and is isolated because it is not found in the dataset, this patient has mass (see Figure 4.5 (b)).

This model generated off-base estimates in patients with low-quality X-beam symbolism and acute respiratory pain disorder (ARDS), in which the lung lung image expands and the lung lungs lose more ventilation (see Figure 4.5 (c)).

This model is valuable for identifying COVID-19 as it includes a preparatory map on general matters. Its effectiveness is reduced in cases of pneumonia and ARDS. The heat map indicated a more important focus area compared to the unobserved area in the X-ray of patients with COVID-19 (see Figure 4.6).

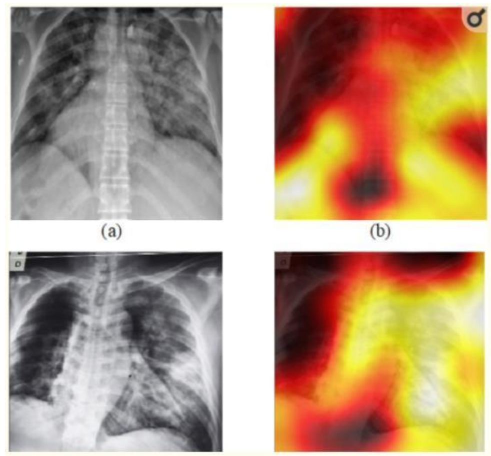


Figure 4.6

X-beam picture and comparative heat map: (a) first X-beam picture, (b) heat guide (a), (c) second X-beam picture, and (d) heat guide (c).

The model further assists in assessing the effectiveness of treatment based on the heat map. It helps specialists in the analysis, development, treatment and isolation of patients.

Fig. 4.7 shows the difference between some COVID and pneumonia case photos. Necessary searches are often found in the chest X-ray of COVID-19 patients.

- Ground Glass-glass opacity (GGOs) (relative, multifocal, subpural, marginal, posterior, medial and basal).
- Ins an insane clearing performance (GGO and middle / intra-lobular septal hardening).
- Air freezing.
- Bronchovascular hardening (in injury).
- Tractional bronchiectasis.

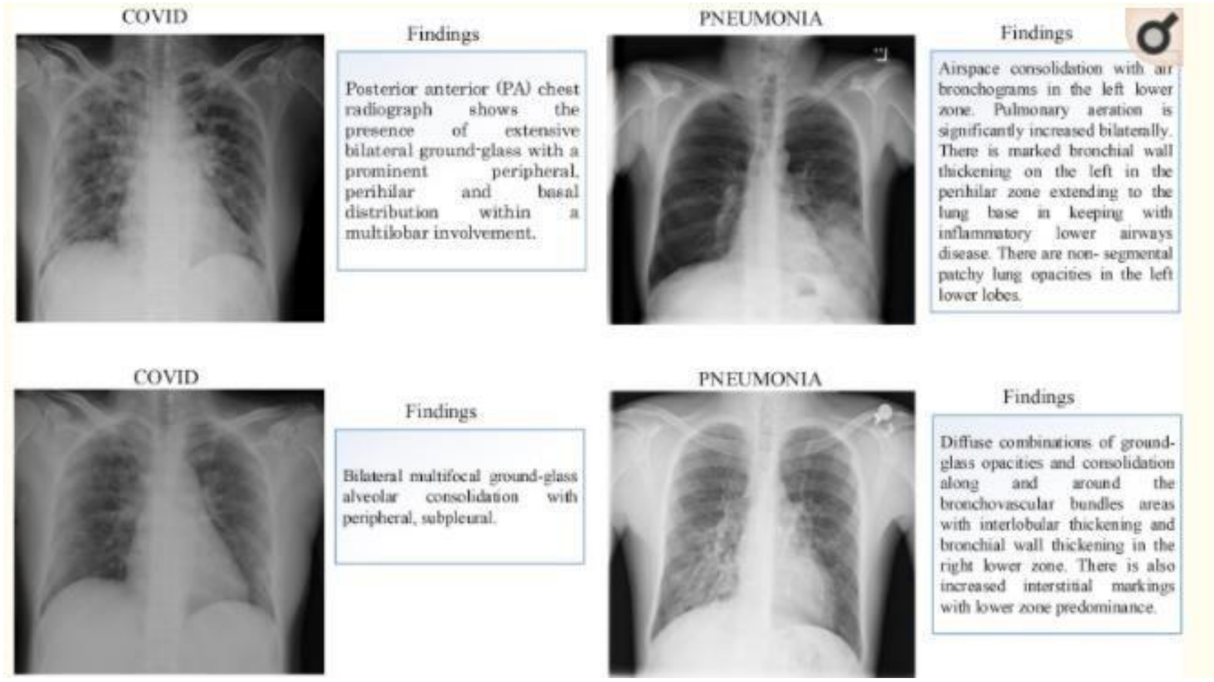


Figure 4.7

Radiologists have noted differences bw some COVID and pneumonia case photographs. In particular, chest X-beam searches of pneumonia patients appear as follows.

- Ground-glass opacity (GGO) Focal diffusion, arbitrary
- Reticular opacity
- Vascular thickening
- Bron is highly distributed in the bronchovascular group
- Tracheal hardening divider

In COVID-19, hilar lymphadenopathy with gland-free solitary lobar or segmental association, various small aspiratorial manifestations are not uncommon. Ken. .

In the COVID-19 crisis, radiological imaging plays an important role despite indicator tests for early determination, treatment and ablation stages of infection. Some trademark searches can be found in the lungs for chest radiography

These models may be promptly utilized for medical services habitats. There are no compelling reason to stand by extended periods of time for the radiologists to screen the pictures. In this manner, medical services laborers and patient family members can zero in on disconnection of dubious cases with the goal that therapy can start. Consequently, the spread of the infection can be essentially diminished. The patients can look for a subsequent assessment on the off chance that they are analyzed as sure by our framework. Consequently, holding up time can be fundamentally decreased, and it will mitigate clinicianremaining burden.

CHAPTER 5

CONCLUSIONS

Coronavirus has stunned the world because of its non-accessibility of immunization or medication. Different analysts are working for vanquishing this savage infection. We utilized different Clinical reports have identified four classes of distinct diseases namely sars,ards,both sars and ards and to be specific the covid. Some headlines such as TFrequency / Inverse Document Frequency and separation of some packs of words. AI computation is used to divide into four distinct classes. After grouping, the SVM and multinomial neo-Bayesian classification was found to give excellent results with 94% accuracy, 96% review, 95% f1 score and an accurate 96.2%. The eff. of the models can be improved by expanding the data measure.

A proper model using both textual and image data was proposed by us and it worked very well on the validation set of data. This model that I made is fully mechanized and has the ability to complete construction without the need for manual element extraction. Our designed framework can individually perform dual and multi-grade work with 90.08% and 7.02% accuracy. The performance of the built model is surveyed by a master radiologist and is suitable for testing with a large database. This framework can be used in remote areas of COVID-19 affected countries to meet less number of experts in radiology. One thing to note here is that i can use these techniques with some other diseases as well. Less number of available data around us related to this subject is one issue here. By taking in account more of the data, we can make this model more efficient and proper. So more data will improve the working of this model.

REFERENCES

1. Wu F., Zhao S., Yu B. A new coronavirus associated with human respiratory disease in China. Nature. 2020;579(7798):265–269. [PMC free article] [PubMed] [Google Scholar]

2. Huang C., Wang Y. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395(10223):497–506. [PMC free article] [PubMed] [Google Scholar]

3. World Health Organization . World Health Organization (WHO); 2020. Pneumonia of Unknown Cause–China. Emergencies Preparedness, Response, Disease Outbreak News. [Google Scholar]

4. Wu Z., McGoogan J.M. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. Jama. 2020;323(13):1239–1242. [PubMed] [Google Scholar]

5. Holshue M.L., DeBolt C. First case of 2019 novel coronavirus in the United States. N. Engl. J. Med. 2020;328:929–936. [PMC free article] [PubMed] [Google Scholar]

6. Kong W., Agarwal P.P. Chest imaging appearance of COVID-19 infection. Radiology: Cardiothoracic Imaging. 2020;2(1) [Google Scholar]

7. Singhal T. A review of coronavirus disease-2019 (COVID-19) Indian J. Pediatr. 2020;87:281–286. [PMC free article] [PubMed] [Google Scholar]

8. Zu Z.Y., Jiang M.D., Xu P.P., Chen W., Ni Q.Q., Lu G.M., Zhang L.J. Coronavirus disease 2019 (COVID-19): a perspective from China. Radiology. 2020 doi: 10.1148/radiol.2020200490. In press. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

9. Kanne J.P., Little B.P., Chung J.H., Elicker B.M., Ketai L.H. Essentials for radiologists on COVID-19: an update—radiology scientific expert panel. Radiology. 2020 doi: 10.1148/radiol.2020200527. In press. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

10. Xie X., Zhong Z., Zhao W., Zheng C., Wang F., Liu J. Chest CT for typical 2019-nCoV pneumonia: relationship to negative RT-PCR testing. Radiology. 2020 doi: 10.1148/radiol.2020200343. In press. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

11. Lee E.Y., Ng M.Y., Khong P.L. COVID-19 pneumonia: what has CT taught us? Lancet Infect. Dis. 2020;20(4):384–385. [PMC free article] [PubMed] [Google Scholar]

12. Bernheim A., Mei X. Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection. Radiology. 2020 doi: 10.1148/radiol.2020200463. In press. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

13. Pan F., Ye T. Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia. Radiology. 2020 doi: 10.1148/radiol.2020200370. In press. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

14. Long C., Xu H. Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT? Eur. J. Radiol. 2020;126:108961. [PMC free article] [PubMed] [Google Scholar]

15. Shi H., Han X. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. Lancet Infect. Dis. 2020;24(4):425–434. [PMC free article] [PubMed] [Google Scholar]

16. Zhao W., Zhong Z., Xie X., Yu Q., Liu J. Relation between chest CT findings and clinical conditions of coronavirus disease (COVID-19) pneumonia: a multicenter study. Am. J. Roentgenol. 2020;214(5):1072–1077. [PubMed] [Google Scholar]

17. Li Y., Xia L. Coronavirus Disease 2019 (COVID-19): role of chest CT in diagnosis and management. Am. J. Roentgenol. 2020:1–7. [PubMed] [Google Scholar]

18. Chan J.F.W., Yuan S. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet. 2020;395(10223):514–523. [PMC free article] [PubMed] [Google Scholar]

19. Yoon S.H., Lee K.H. Chest radiographic and CT findings of the 2019 novel coronavirus disease (COVID-19): analysis of nine patients treated in Korea. Korean J. Radiol. 2020;21(4):494–500. [PMC free article] [PubMed] [Google Scholar]

20. Lorente Edgar. COVID-19 pneumonia - evolution over a week. https://radiopaedia.org/cases/COVID-19-pneumonia-evolution-over-a-week-1?lang=us

21. Litjens G., Kooi T., Bejnordi B.E., Setio A.A.A., Ciompi F., Ghafoorian M.…Sánchez C.I. A survey on deep learning in medical image analysis. Med. Image Anal. 2017;42:60–88. [PubMed] [Google Scholar]

22. Ker J., Wang L., Rao J., Lim T. Deep learning applications in medical image analysis. Ieee Access. 2017;6:9375–9389. [Google Scholar]

23. Shen D., Wu G., Suk H.I. Deep learning in medical image analysis. Annu. Rev. Biomed. Eng. 2017;19:221–248. [PMC free article] [PubMed] [Google Scholar]

24. Faust O., Hagiwara Y., Hong T.J., Lih O.S., Acharya U.R. Deep learning for healthcare applications based on physiological signals: a review. Comput. Methods Progr. Biomed. 2018;161:1–13. [PubMed] [Google Scholar]

25. Murat F., Yildirim O., Talo M., Baloglu U.B., Demir Y., Acharya U.R. Application of deep learning techniques for heartbeats detection using ECG signals-Analysis and Review. Comput. Biol. Med. 2020;120:103726. [PubMed] [Google Scholar]

26. LeCun Y., Bengio Y., Hinton G. Deep learning. Nature. 2015;521(7553):436–444. [PubMed] [Google Scholar]

27. Krizhevsky A., Sutskever I., Hinton G.E. Advances in Neural Information Processing Systems. 2012. Imagenet classification with deep convolutional neural networks; pp. 1097–1105. [Google Scholar]

28. Yıldırım Ö., Pławiak P., Tan R.S., Acharya U.R. Arrhythmia detection using deep convolutional neural network with long duration ECG signals. Comput. Biol. Med. 2018;102:411–420. [PubMed] [Google Scholar]

29. Hannun A.Y., Rajpurkar P., Haghpanahi M., Tison G.H., Bourn C., Turakhia M.P., Ng A.Y. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat. Med. 2019;25(1):65. [PMC free article] [PubMed] [Google Scholar]

30. Acharya U.R., Oh S.L., Hagiwara Y., Tan J.H., Adam M., Gertych A., San Tan R. A deep convolutional neural network model to classify heartbeats. Comput. Biol. Med. 2017;89:389–396. [PubMed] [Google Scholar]

Covid_19_Prediction_Major_Project_171362_2.pdf

ORIGINALITY REPORT

19%	10%	17%	12%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	debuglies.com Internet Source	6%
2	Akib Mohi Ud Din Khanday, Syed Tanzeel Rabani, Qamar Rayees Khan, Nusrat Rouf, Masarat Mohi Ud Din. "Machine learning based approaches for detecting COVID-19 using clinical text data", International Journal of Information Technology, 2020 Publication	3%
3	Tulin Ozturk, Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Baloglu, Ozal Yildirim, U. Rajendra Acharya. "Automated detection of COVID-19 cases using deep neural networks with X-ray images", Computers in Biology and Medicine, 2020 Publication	3%
4	Submitted to Jaypee University of Information Technology Student Paper	2%
5	Asmaa Abbas, Mohammed M. Abdelsamea, Mohamed Medhat Gaber. "Classification of	1%

COVID-19 in chest X-ray images using DeTraC deep convolutional neural network", Applied Intelligence, 2020
Publication

6	Submitted to Coventry University Student Paper	1 %
7	digitalcollections.americanancestors.org Internet Source	1 %
8	Rajit Nair, Santosh Vishwakarma, Mukesh Soni, Tejas Patel, Shubham Joshi. "Detection of COVID-19 cases through X-ray images using hybrid deep neural network", World Journal of Engineering, 2021 Publication	1 %
9	Emtiaz Hussain, Mahmudul Hasan, Md Anisur Rahman, Ickjai Lee, Tasmi Tamanna, Mohammad Zavid Parvez. "CoroDet: A deep learning based classification for COVID-19 detection using chest X-ray images", Chaos, Solitons & Fractals, 2021 Publication	<1 %
10	pubmed.ncbi.nlm.nih.gov Internet Source	<1 %
11	pdfs.semanticscholar.org Internet Source	<1 %
12	Ahmed Hamza Osman, Hani Moetque Aljahdali, Sultan Menwer Altarrazi, Ali Ahmed.	<1 %

"SOM-LWL method for identification of COVID-19 on chest X-rays", PLOS ONE, 2021
Publication

13	Boran Sekeroglu, Ilker Ozsahin. "Detection of COVID-19 from Chest X-Ray Images Using Convolutional Neural Networks", SLAS TECHNOLOGY: Translating Life Sciences Innovation, 2020 Publication	<1 %
14	"Big Data Analytics and Artificial Intelligence Against COVID-19: Innovation Vision and Approach", Springer Science and Business Media LLC, 2020 Publication	<1 %
15	link.springer.com Internet Source	<1 %

Exclude quotes	On	Exclude matches	< 10 words
Exclude bibliography	On		

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

PLAGIARISM VERIFICATION REPORT

Date: **17 MAY 2021**

Type of Document (Tick): PhD Thesis | M.Tech Dissertation/ Report | ✓ B.Tech Project Report | Paper

Name: **Hirdyansh Bhalla** **Department:** **Computer Science** **Enrolment No** **171362**

Contact No. **8837897964** **E-mail.** harshuhirdyansh@gmail.com

Name of the Supervisor: **Dr. Himanshu Jindal**

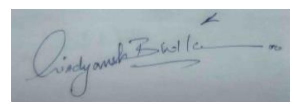
Title of the Thesis/Dissertation/Project Report/Paper (In Capital letters): **COVID-19 PREDICTION USING CLINICAL REPORT TEXT AND X-RAY DATA**

UNDERTAKING

I undertake that I am aware of the plagiarism related norms/ regulations, if I found guilty of any plagiarism and copyright violations in the above thesis/report even after award of degree, the University reserves the rights to withdraw/revoke my degree/report. Kindly allow me to avail Plagiarism verification report for the document mentioned above.

Complete Thesis/Report Pages Detail:

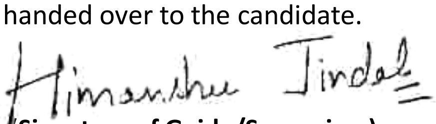
- Total No. of Pages = 45
- Total No. of Preliminary pages = 10
- Total No. of pages accommodate bibliography/references = 3



(Signature of Student)

FOR DEPARTMENT USE

We have checked the thesis/report as per norms and found **Similarity Index** at 19(%). Therefore, we are forwarding the complete thesis/report for final plagiarism check. The plagiarism verification report may be handed over to the candidate.



(Signature of Guide/Supervisor) **Signature of HOD**

FOR LRC USE

The above document was scanned for plagiarism check. The outcome of the same is reported below:

Copy Received on	Excluded	Similarity Index (%)	Generated Plagiarism Report Details (Title, Abstract & Chapters)	
	• All Preliminary Pages • Bibliography/Images/Quotes • 14 Words String	19	Word Counts	8848
Report Generated on			Character Counts	47360
		Submission ID	Total Pages Scanned	45
		1587800514	File Size	2M

Checked by
Name & Signature **Librarian**

……

Please send your complete thesis/report in (PDF) with Title Page, Abstract and Chapters in (Word File) through the supervisor at plagcheck.juit@gmail.com