



PREDICTION OF LONG TERM DISEASES CAUSED AFTER COVID-19 USING CNN AND RANDOM FOREST

Kandregula Jayasri¹, Sai Pavitro Mishra², Gayatri metta³, Lanka swanitha⁴, Bode Prasad⁵

B. Tech Students^{1,2,3,4}, Assistant Professor⁵

Department of Information Technology, Vignan's Institute of Information Technology(Autonomous),
Vishakapatnam.

jayasrikandregula997@gmail.com¹, spmishraoct07@gmail.com², gayatrimetta2001@gmail.com³,
lankaswanitha@gmail.com⁴, arjunprasad.bode@gmail.com⁵

1.ABSTRACT

COVID-19 has created a major impact on healthcare crises across the globe. Many people have affected by this virus with mild to severe symptoms. In most of the cases people with severe symptoms lead to death causing severe problems related to heart , lungs etc. Most of the people lost their lives during the outbreak of covid-19 . Though the recovery rate is good ,most of the people are identified in developing deadly diseases after getting recovered from this virus. On studying the health history records of different types of people affected with covid-19 ,it is derived that in most of the cases people are suffering with diseases related to heart , kidney , lungs etc.The central objective behind this research is to identify the covid-19 patients being affected with such deadly diseases after their recovery so that doctors can easily treat the patients without any complications in future.covid-19 dataset is collected by correcting it with missing and redundant values which is trained with different machined learning algorithms like random forest , decision trees , svm and convolutional neural networks . The proposed model achieves an average accuracy (96%) , precision (95.24%), and recall (92.05%) also prove the utility of the adopted approach in comparison to other techniques for the prediction of diseases.

Keywords

Covid-19 , random forest , convolutional neural networks , Machine learning .

2. INTRODUCTION

Coronavirus disease (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus.

Most people infected with the virus will experience mild to moderate respiratory illness and recover without requiring special treatment. However, some will become seriously ill and require medical attention. Older people and those with underlying medical conditions like cardiovascular disease, diabetes, chronic respiratory disease, or cancer are more likely to develop serious illness. Anyone can get sick with COVID-19 and become seriously ill or die at any age.

The best way to prevent and slow down transmission is to be well informed about the disease and how the virus spreads. Protect yourself and others from infection by staying at least 1 metre apart from others, wearing a properly fitted mask, and washing your hands or using an alcohol-based rub frequently. Get vaccinated when it's your turn and follow local guidance.

The virus can spread from an infected person's mouth or nose in small liquid particles when they cough, sneeze, speak, sing or breathe. These particles range from larger respiratory droplets to smaller aerosols. It is important to practice respiratory etiquette, for example by coughing into a flexed elbow, and to stay home and self-isolate until you recover if you feel unwell.

2.1 Complications of covid-19

Although most people with COVID-19 have mild to moderate symptoms, the disease can cause severe medical complications and lead to death in some people. Older adults or people with existing medical conditions are at greater risk of becoming seriously ill with COVID-19.

Complications can include:

- Pneumonia and trouble breathing
- Organ failure in several organs
- Heart problems
- A severe lung condition that causes a low amount of oxygen to go through your bloodstream to your organs (acute respiratory distress syndrome)
- Blood clots
- Acute kidney injury
- Additional viral and bacterial infections

2.1.1 Major Complications

➤ Heart Disease

Some of the symptoms common in coronavirus “long-haulers,” such as palpitations, dizziness, chest pain and shortness of breath, may be due to heart problems — or, just from having been ill with COVID-19. How do you tell if your symptoms are heart-related, and what can you expect if they are?

Johns Hopkins cardiologists Wendy Post, M.D., and Nisha Gilotra, M.D., clarify which post- coronavirus symptoms may point to a heart issue, when to call your doctor, and other facts all long- term COVID-19 survivors should know.

Temporary or lasting damage to heart tissue can be due to several factors:

- **Lack of oxygen:**

As the virus causes inflammation and fluid to fill up the air sacs in the lungs, less oxygen can reach the bloodstream. The heart has to work harder to pump blood through the body, which can be dangerous in people with preexisting heart disease. The heart can fail from overwork, or insufficient oxygen can cause cell death and tissue damage in the heart and other organs.

- **Myocarditis: inflammation of the heart:**

The coronavirus may infect and damage the heart's muscle tissue directly, as is possible with other viral infections, including some strains of the flu. The heart may also become damaged and inflamed indirectly by the body's own immune system response.

Coronavirus infection also affects the inner surfaces of veins and arteries, which can cause blood vessel inflammation, damage to very small vessels and blood clots, all of which can compromise blood flow to the heart or other parts of the body. "Severe COVID-19 is a disease that affects endothelial cells, which form the lining of the blood vessels," Post says.

- **Stress cardiomyopathy:**

Viral infections can cause cardiomyopathy, a heart muscle disorder that affects the heart's ability to pump blood effectively. When attacked by a virus, the body undergoes stress and releases a surge of chemicals called catecholamines, which can stun the heart. Once the infection resolves, the stressor has ended and the heart can recover.

➤ **Lungs Disease**

COVID-19 can cause lung complications such as pneumonia and, in the most severe cases, acute respiratory distress syndrome, or ARDS. Sepsis, another possible complication of COVID-19, can also cause lasting harm to the lungs and other organs. Newer coronavirus variants may also cause more airway disease, such as bronchitis, that may be severe enough to warrant hospitalization.

- **Pneumonia**

In pneumonia, the lungs become filled with fluid and inflamed, leading to breathing difficulties. For some people, breathing problems can become severe enough to require treatment at the hospital with oxygen or even a ventilator.

The pneumonia that COVID-19 causes tends to take hold in both lungs. Air sacs in the lungs fill with fluid, limiting their ability to take in oxygen and causing shortness of breath, cough and other symptoms.

While most people recover from pneumonia without any lasting lung damage, the pneumonia associated with COVID-19 can be severe. Even after the disease has passed, lung injury may result in breathing difficulties that might take months to improve.

- **Acute Respiratory Distress Syndrome**

If COVID-19 pneumonia progresses, more of the air sacs can become filled with fluid leaking from the tiny

blood vessels in the lungs. Eventually, shortness of breath sets in, and can lead to acute respiratory distress syndrome (ARDS), a form of lung failure. Patients with ARDS are often unable to breath on their own and may require ventilator support to help circulate oxygen in the body.

Whether it occurs at home or at the hospital, ARDS can be fatal. People who survive ARDS and recover from COVID-19 may have lasting pulmonary scarring.

- **Sepsis**

Another possible complication of a severe case of COVID-19 is sepsis. Sepsis occurs when an infection reaches, and spreads through, the bloodstream, causing tissue damage everywhere it goes.

“Lungs, heart and other body systems work together like instruments in an orchestra,” Galiatsatos says. “In sepsis, the cooperation between the organs falls apart. Entire organ systems can start to shut down, one after another, including the lungs and heart.”

Sepsis, even when survived, can leave a patient with lasting damage to the lungs and other organs.

- **Kidney Disease**

Some people suffering with severe cases of COVID-19 will show signs of kidney damage, even those who had no underlying kidney problems before they were infected with the coronavirus. Signs of kidney problems in patients with COVID-19 include high levels of protein or blood in the urine and abnormal blood work.

Studies indicate more than 30% of patients hospitalized with COVID-19 develop kidney injury, and more than 50% of patients in the intensive care unit with kidney injury may require dialysis. Sperati says early in the pandemic, some hospitals were running short on machines and sterile fluids needed to perform dialysis.

“As general treatments for patients with COVID-19 have improved, the rates of dialysis have decreased. This has helped to alleviate shortages, although intermittent supply chain disruptions remain a concern.

“Many patients with severe COVID-19 are those with co-existing, chronic conditions, including high blood pressure and diabetes. Both of these increase the risk of kidney disease,” he says.

But Sperati and other doctors are also seeing kidney damage in people who did not have kidney problems before they got infected with the virus.

3.LITERATURE REVIEW

Prediction of Post Covid complications is one of the most important problem that has to be detected in the early phases of the commencement of the disease so as to reduce the disease progression rate among the individuals .Various researches have been made to find the basic cause of the disease and some have reached to the heights by proposing a system which differentiates the healthy people from those with any affected diseases post affected by covid using various machine learning and deep learning techniques. Lots of pre-processing, feature selection and classification techniques have been implemented and developed in the past decades. Following is the given work done in the prediction of post covid complications.

4.PROBLEM STATEMENT

4.1 EXISTING SYSTEM

As we observe in the present covid situation After being affected with covid-19 many people are suffering with diseases developing their symptoms soon after their recovery and leading to death in the most of the cases.Generally,any diseases can be identified and cured in its initial stage of occurrence only when person

being affected by the disease is health conscious and can identify any drastic change or the health complications that are occurring in their body unless it is impossible for that particular person to know about their health condition and diseases with which they are being suffering with. By this indication of any change, they can consult any doctor and can know their health status and take the necessary precautions to maintain good health. In some situations with the help of some body tests, doctors can easily identify the damage occurring in any part of the patient body and can cure accordingly in the initial stage itself but in most of the cases it becomes a task to identify the actual condition and also may lead to wrong indication of the health status. By our proposed model, doctors can make sure about the health condition of the patients with the diseases related to heart, lungs, kidneys etc. they are being affected with and cure accordingly.

4.2 PROPOSED SYSTEM

As we observe that covid-19 era created a major impact on the healthcare system and its facilities. Generally as we see, many people are being suffered with covid-19 complications and takes a long time for their recovery and also leads to death in the most of the cases. Our proposed model predicts these long term affecting diseases after recovering from covid-19 with best accuracy compared to many different models already proposed as of now. To implement this model covid-19 dataset is taken and corrected with the missing values and the redundant values and trained with the different machine learning algorithms like random forest, decision trees, svm etc and also deep learning techniques like convolutional neural networks. Also, with this model hospital management can be able to store the health status of the patients and can monitor the patients accordingly. Results show that the proposed technique outperforms other baseline techniques and achieves the highest accuracy of 91.34%. Moreover, the results of F1-score (93.16%), precision (92.00%), and recall (94.35%) also prove the utility of the adopted approach in comparison to other techniques for the prediction of heart diseases.

5. RESEARCH METHODOLOGY

This section describes about the architecture, modules, hardware and softwares that are used and implemented in the proposed model of predicting long term diseases caused due to covid-19 using CNN and random forest.

○ Architecture

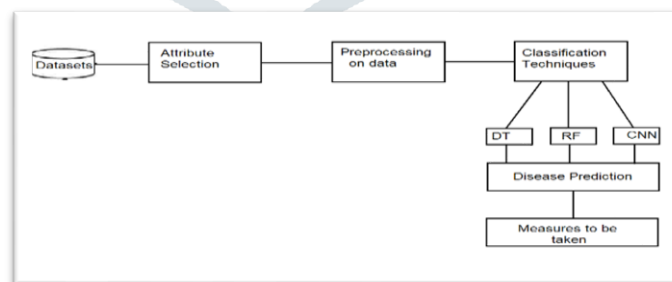


Fig 1: Architecture

➤ Datasets

Dataset is a collection of data that is collected from several patients who are affected with covid-19 and the data represents their health records after their recovery. This data contains different types of attributes and to avoid the wrong prediction of the data, only required data is utilized. Also, the data is corrected with both the missing values and the redundant values.

➤ Attribute Selection

As, the dataset contains different set of attributes in it, using the complete dataset may make the model overfitting that results in less precise output. So, that only the most required attributes are taken into the consideration and trained further to predict the result from the model.

➤ Data preprocessing

Data collected may contain noisy data redundant data, inconsistent data. With such type of data it becomes really challenging to predict the exact output. So, that data must be cleaned by removing all these missing values, redundant values etc and this process is known as data preprocessing. This is most important step in the entire architecture.

➤ Algorithms

Now, after preparing the dataset ready to be trained, it is also necessary to find out the best classification techniques that are to be used and implemented in the model. In our proposed model, we have checked the model with different machine learning algorithms like Decision trees, random forests, svm etc and also deep learning techniques like convolutional neural networks and out of all the best accuracy is achieved with random forest and convolutional neural networks compared to other models.

Hence, the result is predicted and necessary precautions are taken as suggested by the doctor to cure themselves from the deadly diseases.

○ Methodology

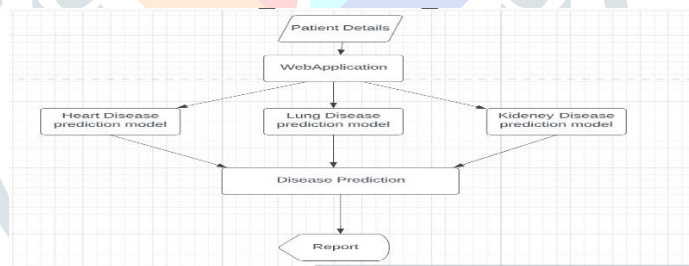


Fig 2: Methodology

○ Modules

➤ Data Gathering

As we are working on the long term diseases that are caused after being affected with the covid-19, data set consisting of health records of covid-19 patients is collected. In our model, as already dataset is available in the online, it is taken from internet and processed further in order to derive the required predictive analysis of the health condition of the covid-19 recovered people. This data not only taken directly and used, but it is corrected with the certain missing values, noisy values, redundant values such that the data will be cleaned and gives better accurate result at the final end. Also, sometimes it is told that the all the attributes should not be considered as it becomes overfitting and accurate result is not obtained. So, that the data is preprocessed and cleaned after the gathering.

➤ Data Splitting

The data collected after processed is trained and tested with the best model in order to get the accurate analysed result. Hence, the data is splitted into train data and test data in the particular ratio and trained with the model with the train data set and tested with the model with the test dataset. This is the most important step in the entire methodology.

➤ Researching the best model

As we already know that there are various machine learning algorithms that are implemented to train and test model . In our proposed model , we have trained and tested the model with different machine learning algorithms like decision trees , random forest , classification techniques , SVM , etc and also with the deep learning techniques like convolutional neural networks etc and out of all the models , random forest and the convolutional neural networks are identified as the best ones with an accuracy of 96% which stands out from all the existing models.

➤ Training and Testing the model

As the model that is choosen in our proposed model is random forest classifier and convolutional neural networks , the dataset that is gathered in the first step is trained with these techniques and tested , implemented further to give the precise output .

➤ Connecting with the interface

As shown in the methodology flow chart , this model is later connected with web UI interface so that , that makes ease to the hospital administration inorder to store the health records of the patients and can treat the patients accordingly . This UI is made with django as backend frameworks and html , css , javascript as front end and also mysql database which store the report generated at end.

6.RESULTS AND DISCUSSIONS

The proposed model produces the following outcomes as shown below

➤ Home page

This is the starting interface where the users are asked to upload their scanned reports and some basic details that are related to their heart , lungs and kidneys . so , that the final report can be generating stating the health condition of the covid-19 recovered patients.

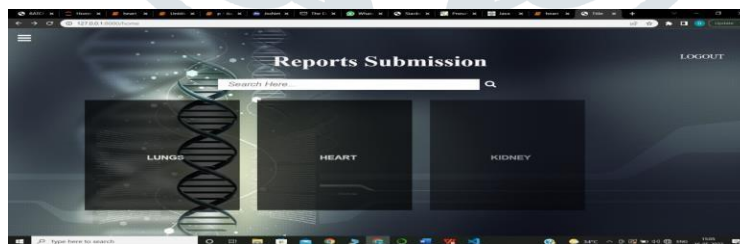


Fig 3.1 : outcome

➤ Uploading details

In this section , starting from the heart patients are asked with all the necessary details according to their health condition and uploaded by the hospital management and processed further to get the necessary finalized report.

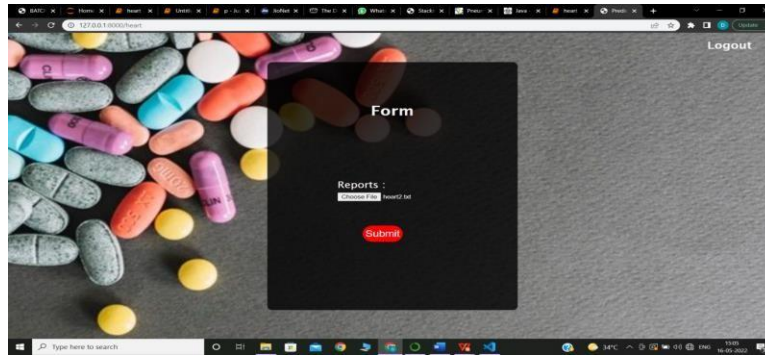


Fig 3.2 : outcome

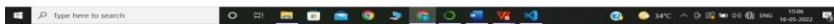
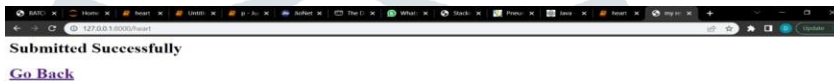


Fig 3.3 : outcome

➤ **Final report**

This is the final report that is generated stating the final condition of the covid-19 recovered patient .



Fig 3.4 : outcome

7.CONCLUSIONS AND FUTURE SCOPE

The proposed system is implemented based on the present situation of the covid-19 outbreak and hence, we can say that this is the best approach to overcome the healthcare crisis and save the lives of many people . Also, with this system , it makes doctors to individually monitor each and every patient to cure them at their max .Also , the proposed system only identifies the diseases that are related to the heart ,

kidneys and lungs and in future we are going to implement this to also predict the health status related to brain . This model can also be added with the most modernized techniques to improve the accuracy of the proposed model in order to get even more precise result and can withstand in the health management.

8. REFERENCES

1. Qian Zhang, Paul Bastard, COVID Human Genetic Effort, Aurelie Cobat, Jean-Laurent Casanova Nature 603, 587-598(2022)
2. Kristin J. Marks, PhD, 1, 2 Michael Whitaker, MPH, 1 Onika Anglin, MPH, 1, 3 Jennifer Milucky, MSPH, 1 Kadam Patel, MPH, 1, 3 Huong Pham, MPH, 1 Shua J. Chai, MD, 4, 5 Pam Daily Kirley, MPH, 4 Isaac Armistead, MD, 6 Sarah McLafferty, MPH, 6 James Meek, MPH, 7 Kimberly Yousey-Hindes, MPH, 7 Evan J. Anderson, MD, 8, 9, 10 Kyle P. Openo, DrPH, 8, 9 Andy Weigel, MSW, 11 Justin Henderson, MPH, 12 Val Tellez Nunez, MPH, 12 Kathryn ComoSabeti, MPH, 13 Ruth Lynfield, MD, 13 Susan L. Ropp, PhD, 14 Chad Smelser, MD, 14 Grant R. Barney, MPH, 15 Alison Muse, MPH, 15 Nancy M. Bennett, MD, 16 Sophrena Bushey, MHS, 16 Laurie M. Billing, MPH, 17 Eli Shiltz, MPH, 17 Nasreen Abdullah, MD, 18 Melissa Sutton, MD, 18 William Schaffner, MD, 19 H. Keipp Talbot, MD, 19 Ryan Chatelain, MPH, 20 Andrea George, MPH, 20 Christopher A. Taylor, PhD, 1 Meredith L. McMorro, MD, 1 Cria G. Perrine, PhD, 1 Fiona P. Havers, MD, 1 and COVID-NET Surveillance Team, Hospitalizations of Children and Adolescents with Laboratory-Confirmed COVID-19 — COVID-NET, 14 States, July 2021– January 2022
3. Senthilkumar Mohan, John A. Aghed, Abugabah, Adimoolam M, Shubham Kumar Singh, Ali kashif Bashir, Louis Sanzogni, An approach to forecast impact of Covid-19 using supervised machine learning model, 01 April 2021, <https://doi.org/10.1002/spe.2969>
4. Matteo Chieragato, Fabio Frangiamore, Mauro Morassi, Claudia Baresi, Stefania Nici, Chiara Bassetti, Claudio Bnà & Marco Galelli, A hybrid machine learning/deep learning COVID-19 severity predictive model from CT images and clinical data, (2022)
5. Alaa M. A. Barhoom & Samy S. Abu-Naser, International Journal of Academic Engineering Research (IJAER) 6 (2):48-68 (2022)
6. Katy Stokes, Rossana Castaldo, Carlo Federici, Silvio Pagliara, Alessia Maccaro, Francesco Cappucci, Giuseppe Ficof, Marco Salvatore, Monica Franzese, Leandro Pecchia, The use of artificial intelligence systems in diagnosis of pneumonia via signs and symptoms: A systematic review, February 2022
7. Ibrahim M. El-Hasony, Omar M. Elzaki, Ali Alsheri, Hanaa Salem, Multi-Label Active Learning- Based Machine Learning Model for Heart Disease Prediction, 4 February 2022
8. Forum Desai, Deepraj Chowdhury, Rupinder Kaur, Marloes Peeters, Rajesh, Chand Arya, Gurpreet Singh, Wandere Sukhpal Singh, Gilla Rajkumar, Buyyaf, HealthCloud: A system for monitoring health status of heart patients using machine learning and cloud computing, March 2022
9. Qiong Bai, Chunyan Su, Wen Tang & Yike Li, Machine learning to predict end stage kidney disease in chronic kidney disease, 19 May 2022
10. Toshiki Kuno, Takahisa Mikami, Yuki Sahashi, Yohei Numasawa, Masahiro Suzuki, Shigetaka Noma, Keiichi Fukuda & Shun Kohsaka, Machine learning prediction model of acute kidney injury after percutaneous coronary intervention, 14 January 2022