International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 04 | April - 2022 Impact Factor: 7.185 ISSN: 2582-3930

Covid-19 Prediction using Chest X-Rays

Vinay S Medhekar, Prathamesh D Patil, Aditi A Nagre, Prof. Sonali Pakhmode

Department of Information Technology

Vasantdada Patil Pratishthan's College of Engineering and Visual Arts, Eastern Express Highway, Near Everard Nagar, Sion-Chunabhatti, Mumbai-400 022.

ABSTRACT

Since its inception of 2020, COVID-19 had a devastating impact on the universe. The loss of life has been huge and the global economy has been in total disarray. Even as we approach the end of the year (2022), there appears to be no plausible solution for the deadly virus. The other is to quarantine an infected person, till he/she feels better. By this, we mean to say that if a person shows any of the symptoms, like fever, heavy breathing, etc., they must not be allowed access to public places or gatherings. This web application gives us an idea if the person in the given x-ray is covid positive or negative. We have devised a web application, which uses a Machine Learning model to detect COVID-19, by analyzing the chest X-ray images of the patient. The program is quite accurate in distinguishing these images, and can be used as a simple, cost-effective and user-friendly means to identify a potential COVID-19 infected patient.

I. INTRODUCTION

The Novel Coronavirus, technically known as COVID19, placed its devastating roots, first in China, in the beginning of 2020, and by March 2020, it had spread all over the world. Even as this report is being presently compiled, the loss of life worldwide keeps increasing. More than 3 million people have already been killed by this global pandemic, and another 21 million people are currently infected (as of 1 May 2021), and are currently fighting for their lives. Not to mention the fact that several thousand people are still being infected on a daily basis. From the above stated facts, it is clear that the prevention of this virus is essential. However, it can be hard or forgetful for people to follow this norm at all times. Disease transmission can be suppressed by reducing the likelihood of an uninfected person coming into direct contact with an infected person, resulting in fewer cases and deaths. These symptoms include high fever, chest pain, difficulty in breathing, sore throat, etc. We need a device which can accurately measure few of these symptoms, so that we can do our best to avoid being exposed to COVID-19.

II. LITERATURE SURVEY

The COVID-19 disease is caused by the extreme acute respiratory syndrome coronavirus 2 (SARS CoV-2), which was first laid hands-on Wuhan, China, in late December 2019 and has since spread worldwide (WHO) [4]. This pandemic has put massive strain on the world's health-care systems. The majority of coronavirus patients recover after developing mild flu-like symptoms, but about 20% of patients clinically deteriorate, necessitating hospitalization. This deterioration can happen quickly, requiring intubation and other advanced life support measures before or when the patient arrives at the hospital [5]. In healthy patients, oxygen saturation varies from 95 to 100 percent [6]. When chest x-rays were taken, it was discovered that a significant number of hospital admissions not attributed to COVID-19 were actually suffering from pneumonia caused by COVID-19, with most having severely low oxygen saturation levels of just 50%. COVID-19 pneumonia, unlike normal pneumonia, which causes chest pain and breathing problems, causes oxygen depletion that, is difficult to diagnose at first because the patients do not have any serious breathing difficulties. This is referred to as "silent hypoxia." COVID-19-infected patients' conditions have already worsened into moderate to-severe pneumonia by the time they know they are out of oxygen. COVID-19 patients, on the other hand, should not experience shortness of breath in the early stages of COVID-19 pneumonia. As a result, detecting this silent type of hypoxia in COVID-19 patients until they experience shortness of breath is crucial for preventing pneumonia from progressing to a serious step. The key is to be able to detect this initial drop in oxygen saturation levels so that patients infected with COVID-19 who begins to suffer from pneumonia can be detected very early on and put on a treatment plan to prevent the lungs from deteriorating further [4]

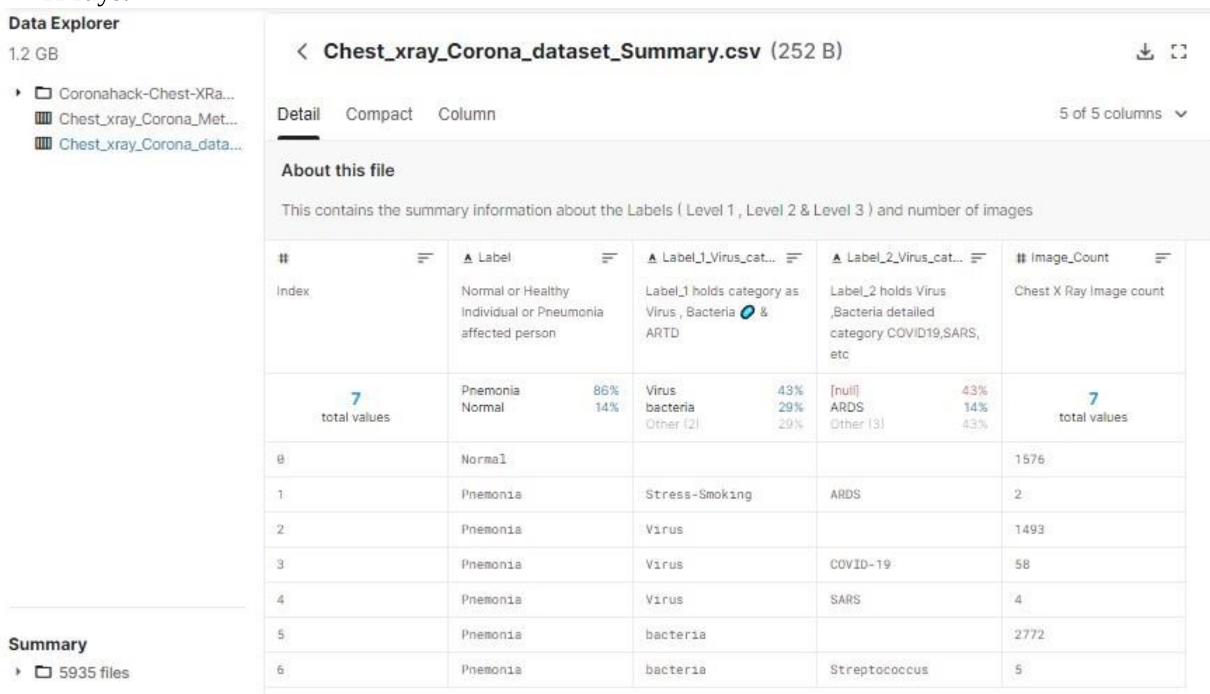
III. METHODOLOGY

A. Loading and importing the dataset.

DATASET 1:

This dataset is called 'Corona Hack -Chest X-Ray-Dataset' and obtained from KAGGLE, which is an online community for data scientists and people working on ML projects.

- This dataset has 2 types of chest x-ray images: i. Chest x-ray images of 'Normal' healthy individuals.
  ii. Chest x-ray images of individuals infected from 'Pneumonia'.
- From this dataset, we extracted all the chest x-ray images of normal healthy individuals and considered them as COVID-19 Negative X-rays.



Data Explorer
1.2 GB
Coronahack-Chest-XRa...
Chest_xray_Corona_Met...
Chest_xray_Corona_data...
Summary
5935 files

Chest_xray_Corona_dataset_Summary.csv (252 B)

Detail Compact Column 5 of 5 columns

About this file

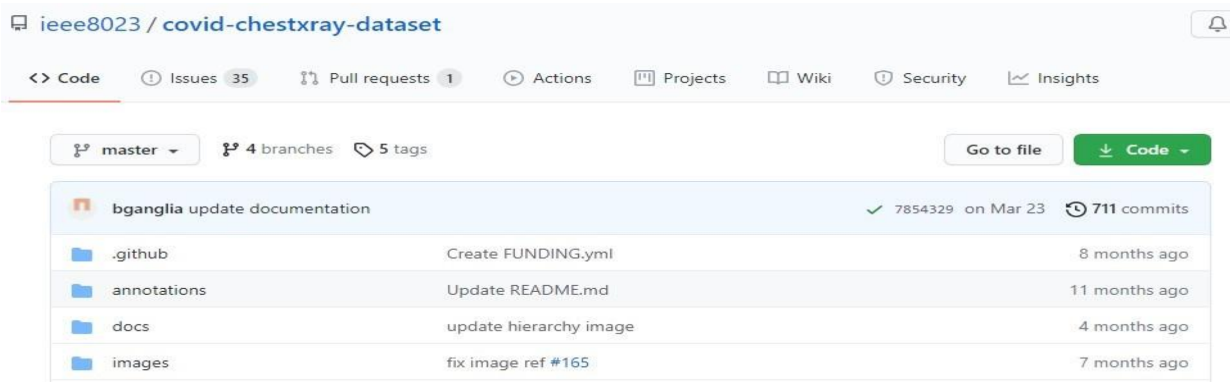
This contains the summary information about the Labels (Level 1 , Level 2 & Level 3) and number of images

#	Label	Label_1_Virus_cat...	Label_2_Virus_cat...	Image_Count
Index	Normal or Healthy Individual or Pneumonia affected person	Label_1 holds category as Virus , Bacteria & ARTD	Label_2 holds Virus ,Bacteria detailed category COVID19,SARS, etc	Chest X Ray Image count
7 total values	Pnemonia 86% Normal 14%	Virus 43% bacteria 29% Other (2) 29%	[null] 43% ARDS 14% Other (3) 43%	7 total values
0	Normal			1576
1	Pnemonia	Stress-Smoking	ARDS	2
2	Pnemonia	Virus		1493
3	Pnemonia	Virus	COVID-19	58
4	Pnemonia	Virus	SARS	4
5	Pnemonia	bacteria		2772
6	Pnemonia	bacteria	Streptococcus	5

DATASET 2:

- This dataset is called 'covid chest x-ray-dataset' and can be found on GITHUB, which is an open-source website software development.
- This dataset consists of x-rays of individuals affected from various diseases. However, we have extracted only the x-rays of individuals affected from 'Pneumonia/Viral/COVID-19' and considered them as COVID-19 POSITIVE.

Hence in total, we have compiled 2155 images, out of which **1576** images are **COVID-19 NEGATIVE** x-rays images and the rest **579** images are **COVID-19 POSITIVE** x-ray images.



B. Preprocessing the data

Before feeding the data into the ML model, it needs to be standardized to a particular format, which can be processed by the model. This process of cleaning and transforming the data to the required standard is known as *'data pre-processing'*. There are two important steps to be taken in data preprocessing:

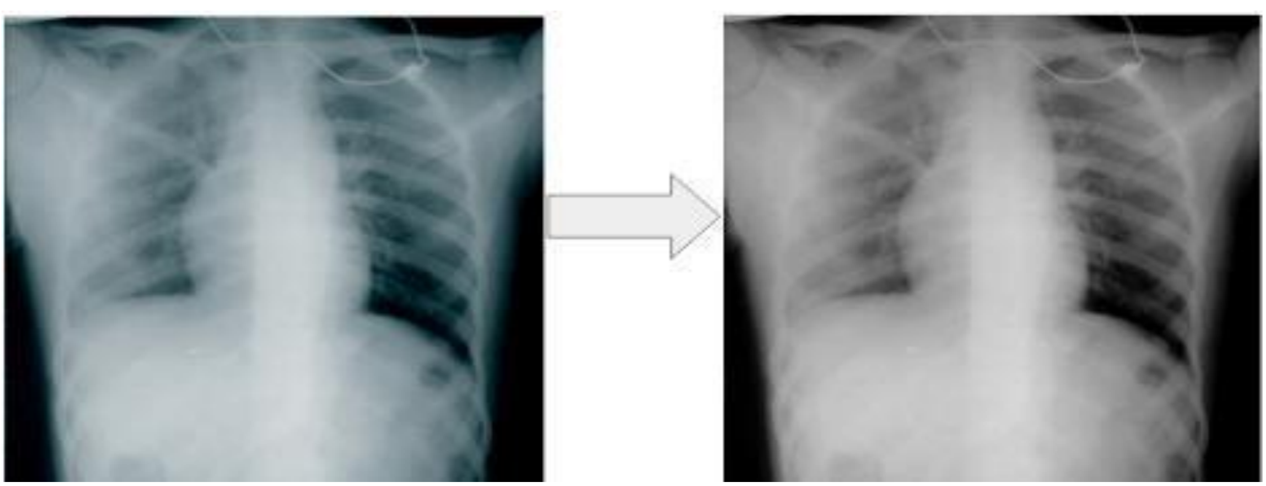
1. **Data cleaning:**

   The process of removing any corrupted, incomplete, or unnecessary data from the dataset is called as data cleaning, data cleansing, or data scrubbing.

2. **Data transformation:**

   The process of transforming or changing the data from one format or structure, to the other, is called data transformation. It is the most essential part of data pre-processing. Different types of data are processed in different ways, according to the need of the model.

The dataset that we used for our model consisted of x-ray images in RGB color format. They were also of different sizes i.e., of different parameters. In data pre-processing, the first image transformation was to convert all the RGB format images in our dataset to standard grayscale format. Through this transformation, we improve upon the space and time complexities of our model.



To further reduce the processing time and standardize the pixel height and width of all the images, as the sizes of all the images were also different, we performed one final image transformation, where once an image is converted from RGB to grayscale, we resized it to 100x100 pixels, which further reduces the time taken by the model to process them.

After the image transformation, we assign each image to a label, where label '0' stands for normal
COVID-19 Negative individual's X-ray and label '1' for a COVID-19 Positive person's X-ray. The images and labels are then stored in 2 separate arrays.

C. Training the CNN model

We have built our CNN model in Keras. The architecture we used for our CNN model is shown below:

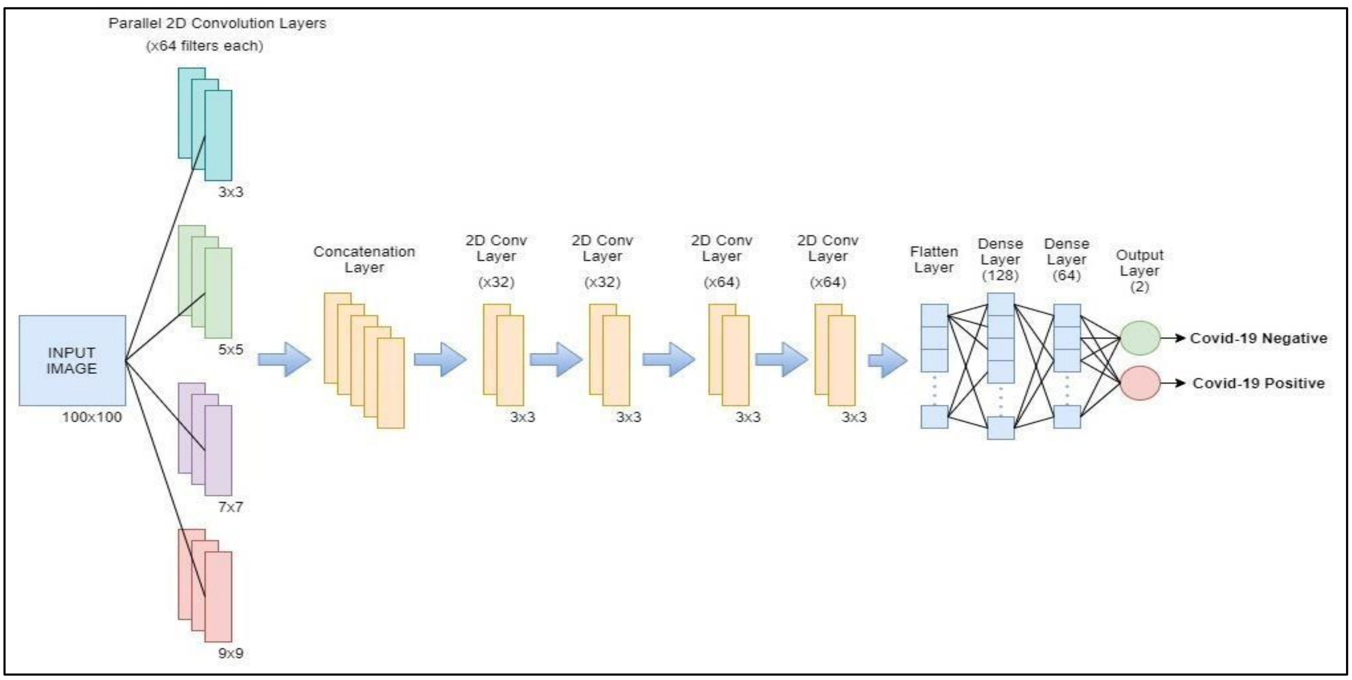


Figure 3.5.4.1 : CNN Model architecture

First, we have taken four filters of varying sizes 3x3, 5x5, 7x7 and 9x9. There are 64 filters for each filter size. Then, the input image is fed parallelly to the four filters (3x3, 5x5, 7x7 and 9x9) simultaneously and then convolution operation is performed between the input image and each of the filters. We pass the result of the convolution operation through ReLU activation function.

Thus, 64 feature maps are produced from convolution operation between input image and 64 3x3 filters. Similarly, 64 feature maps are produced from each of the convolution operation between the input image and 64 5x5, 64 7x7 and 64 9x9 filters.

Now, all the obtained feature maps are concatenated in the Concatenation Layer.

Then, we perform convolution operation using 32 3x3 filters. We pass the result of the convolution operation through ReLU activation function. Max pooling operation is then performed on the obtained feature maps. For max pooling, we have used window size of 2x2 and strides also equal to 2.

Again, we perform convolution operation using 32 3x3 filters. We pass the result of the convolution operation through ReLU activation function. Max pooling operation is then performed on the obtained feature maps. Again, for max pooling, we have used window size of 2x2 and strides also equal to 2.

Now, convolution operation is performed using 64 3x3 filters. We pass the result of the convolution operation through ReLU activation function. Max pooling operation is then performed on the obtained feature maps. For max pooling, we have used window size of 2x2 and strides also equal to 2.

After the convolution and pooling layers, we flatten the obtained layer to get fully connected layer. We have also added a dropout rate of 50%. Then we used 128 neurons in the dense layer and activation function ReLU. Then we have again added a dropout rate of 50%. Then we used 64 neurons in the next dense layer again with the activation function ReLU and again a dropout rate of 50%.

The last dense layer is connected to the output layer containing 2 neurons and the activation function used is SoftMax, since the output is probability of being Covid-19 Positive/Negative.

The loss function used is 'Categorical-cross entropy' since it is a case of multi-class classification.

Our dataset consists of 2155 x-ray images, out of which 1576 x-ray images are of Covid-19 Negative individuals and rest 579 x-ray images are of Covid-19 Positive individuals. For training of our model, we have split the dataset such that 80% of the dataset (i.e., 1724 x-ray images) are used for training and the rest 20% of the dataset (i.e. 431 x-ray images) are used for testing the model.

We performed training of the model on the dataset with 40 epochs.

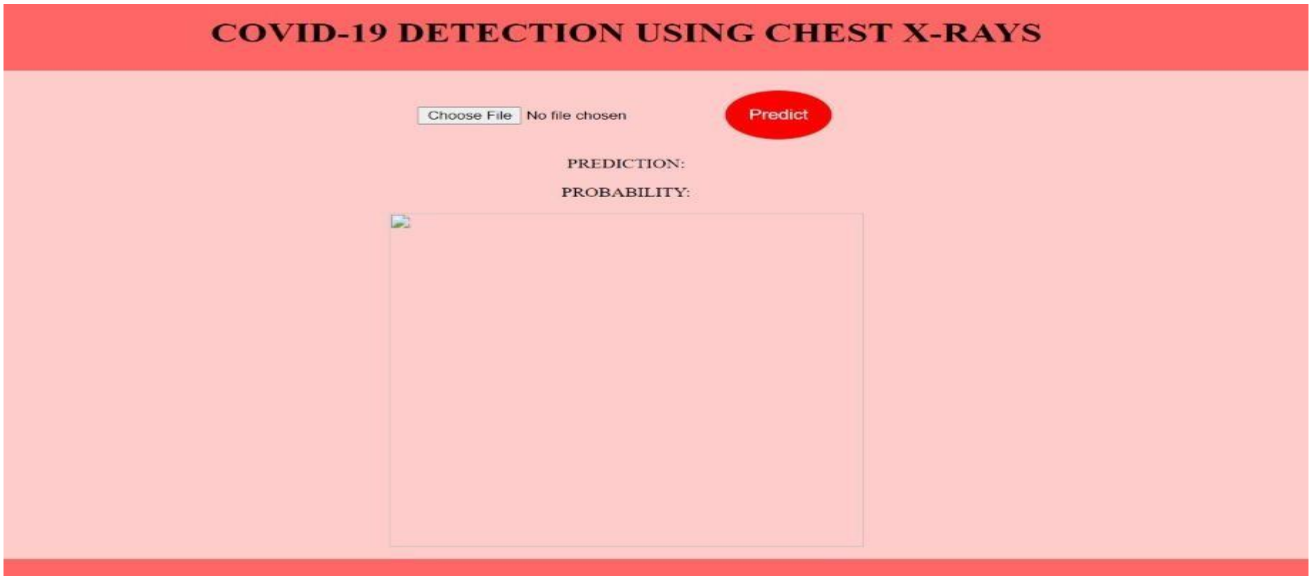
D. Web Application

A web application is also developed in which the CNN model is integrated. The web application provides a simple interface for the user to access the ML model, in order to upload his/her chest Xray, and find out whether they are Covid-19 Positive or Covid-19 Negative.

The web application is developed using HTML, CSS, and JavaScript (frontend), where HTML and CSS have been used to design the actual webpage, and to make it aesthetically appealing and python is used in the back-re end.

Libraries used and purpose of using it:

1) Flask: flask is basically a web application framework and it is used for the same
2) Keras.model: This library is used to load the machine learning model into your backend
3) CV: This library is used to resize and recolor the images
4) NumPy: This library is used to working with arrays that is used to convert image into arrays of pixels
5) PIL: PIL stands for python imagining library and is used because we are working with images.



IV. RESULTS

Here, we have also uploaded an x-ray of an unhealthy person. The following result was obtained:

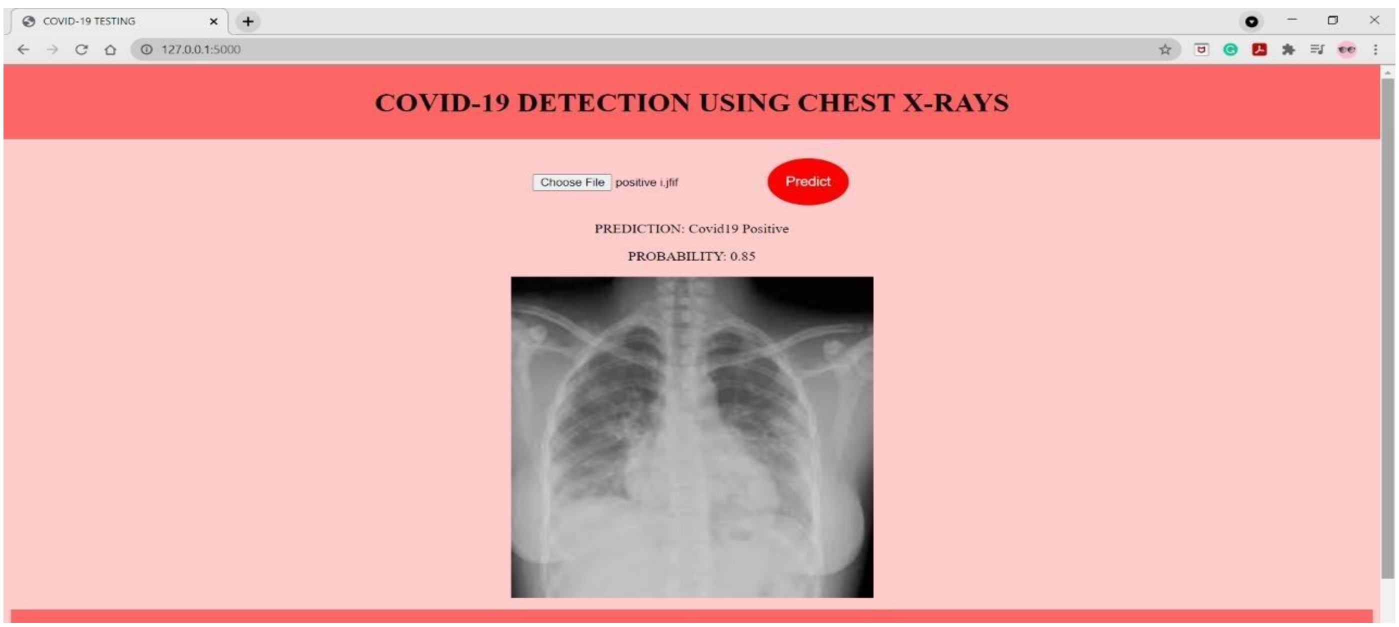


Figure 4.5.5: Prediction and Result

Here, we have also uploaded an x-ray of a healthy person. The following result was obtained:

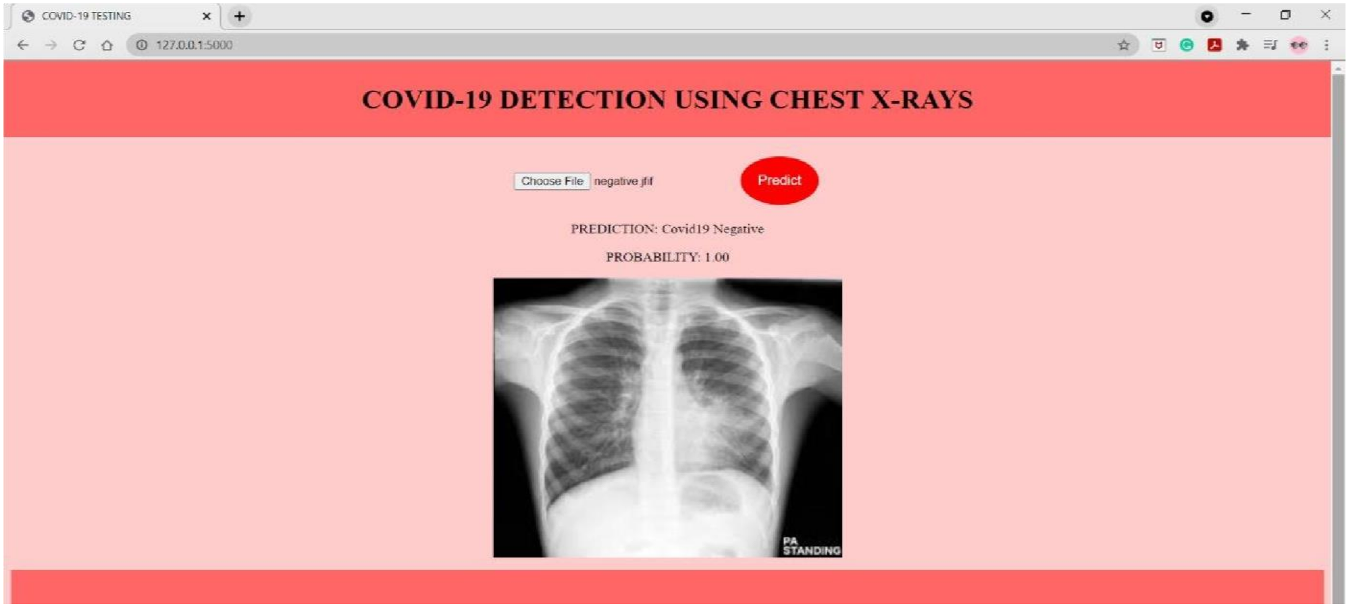


Figure 4.5.6: Result when X-ray of a healthy person is uploaded

V. CONCLUSION

- After doing a lot of research, we learned about the impact of coronavirus on the world, thus we understood the need for the devices and applications developed under this project The Machine Learning model, which predicts the possibility of patients having COVID-19, by analyzing their chest X-rays, has been successfully able to predict COVID-19 in several X-ray cases, and results have been accurate every single time.
- The machine learning model has also been successfully integrated with the web application and is ready for use.

VI. FUTURE SCOPE

The ML model and web application can be transformed into a mobile android application, which shall access the mobile camera to scan and analyze the X-rays of the user. This would get rid of the tedious task of relaying the scanned image between the mobile and the PC.

- The covid19 detection device shall be further worked upon, and turned into a more finished product, along with the casing.
- UI of the COVID-19 web application shall be improved further.
- The X-Ray system can be made more user-friendly.

VII. REFERENCES

[1] B. F. Giraldo etal., "Study of the respiratory pattern variability in patients during weaning trials," The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2004.
[2] J. S. Suri et al., "Systematic Review of Artificial Intelligence in Acute Respiratory Distress Syndrome for COVID-19 Lung Patients: A Biomedical Imaging Perspective," in IEEE Journal of Biomedical and Health Informatics.

[3] B. K. Umri, M. Wafa Akhyari and K. Kusrini, "Detection of Covid-19 in Chest X-ray Image using CLAHE and Convolutional Neural Network," 2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS), 2020.

[4] A. Shankar, Y. Sonar and K. A. Sultanpure, "Detection of COVID-19 using Chest X-Ray Scans," 2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC), 2020.

[5] S. Walvekar and S. Shinde, "Efficient Medical Image Segmentation Of COVID-19 Chest CT Images Based on Deep Learning Techniques," 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), 2021.

[6] A. Ala and Ö. Polat, "Detection of Covid-19 from Computed Tomography Images with DenseNet Based Deep Learning Models," 2021 29th Signal Processing and Communications Applications Conference (SIU), 2021.

[7] T. Padma and C. U. Kumari, "Deep Learning Based Chest X-Ray Image as a Diagnostic Tool for COVID-19," 2020 International Conference on Smart Electronics and Communication (ICOSEC), 2020.