PNEUMONIA DETECTION USING X-RAYS USING WATSON STUDIO

A UG PROJECT PHASE-1 REPORT

submitted to

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY, HYDERABAD

In partial fulfillment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

Submitted by

K. HRUTHIC	**20UK5A0513**
P. THIRUMALA	**20UK5A0517**
P. ANURAAG	**18UK1A0545**
K. RUCHITHA	**19UK1A05P5**

Under the guidance of

Mr. V. RANADHEER REDDY

Assistant Professor

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

VAAGDEVI ENGINEERING COLLEGE

Affiliated to JNTUH, HYDERABAD

BOLLIKUNTA, WARANGAL (T.S) – 506005

2019-2023

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

VAAGDEVI ENGINEERING COLLEGE

BOLLIKUNTA, WARANGAL-506005

2019-2023

CERTIFICATE OF COMPLETION

UG PROJECT PHASE-1

This is to certify that the UG Phase-I Report entitled **"PNEUMONIA DETECTION USING X-RAYS USING WATSON STUDIO"** is being submitted by **K.HRUTHIC**(H.NO:20UK5A0513), **P. THIRUMALA**(20UK5A0517), **P. ANURAAG**(H.NO:18UK1A0545), **K. RUCHITHA**(H.NO 19UK1A05P5)in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering** to **Jawaharlal Nehru Technological University Hyderabad** during the academic year **2022-2023**

Project Guide	**Head of the Department**
Mr. V.RANADHEER REDDY	**Mr. Dr. R. Naveen Kumar**
(Assistant Professor)	(Professor)

EXTERNAL

ACKNOWLEDGEMENT

We wish to take this opportunity to express our sincere gratitude and deep sense of respect to our beloved **Dr. P. PRASAD RAO** Principal, Vaagdevi Engineering College for making us available all the required assistance and for his support and inspiration to carry out this UG Project Phase – I in the institute.

We extend our heartfelt thanks to **Dr. R. NAVEEN KUMAR**, Head of the Department of CSE, Vaagdevi Engineering College for providing us necessary infrastructure and thereby giving us freedom to carry out the mini project.

We express heartfelt thanks to Smart Bridge Educational Services Private Limited, for their constant supervision as well as for providing necessary information regarding the mini project anf for their support in completing the mini project.

We express heartfelt thanks to the guide, **Mr. V. RANADHEER REDDY**, Assistant Professor, Department of CSE for his constant support and giving necessary guidance for completion of this mini project.

Finally, we express our sincere thanks and gratitude to our family members, friends for their encouragement and outpouring their knowledge and experiencing throughout thesis.

K. HRUTHIC (20UK5A0513)
P. THIRUMALA (20UK5A0517)
P. ANURAAG (18UK1A0545)
K. RUCHITHA (19UK1A05P5)

ABSTRACT

The risk of pneumonia is immense for many, especially in developing nations where billions face energy poverty and rely on polluting forms of energy. Over 150 million people get infected with pneumonia on an annual basis especially children below 5 years. A patient suffering from Pneumonia takes an X-ray image to the doctor; with them he predicts pneumonia. The results are not just based on seeing the X-ray images, furthermore, tests will be conducted on the patient. The process was time-consuming, but in recent days artificial intelligence helps in predicting pneumonia bypassing the X-ray image. The main objective of this project is to help the doctors to predict the pneumonia disease more accurately using a deep learning model. The objective is not only to help the doctors but also to the patients to precisely predict pneumonia.

This study proposes a Convolutional neural network model trained from scratch to classify and detect the presence of pneumonia from a collection of chest X-ray image samples. Unlike other methods that rely solely on transfer learning approaches or traditional handcrafted techniques to achieve a remarkable classification performance, we constructed a Convolutional neural network model from scratch to extract features from a given chest Xray image and classify it to determine if a person is infected with pneumonia. This model could help mitigate the reliability and interpretability challenges often faced when dealing with medical imagery. Unlike other deep learning classification tasks with sufficient image repository, it is difficult to obtain a large amount of pneumonia dataset for this classification task; therefore, we deployed several data augmentation algorithms to improve the validation and classification accuracy of the CNN model and achieved remarkable validation accuracy.

TABLE OF CONTENTS

1. INTRODUCTION

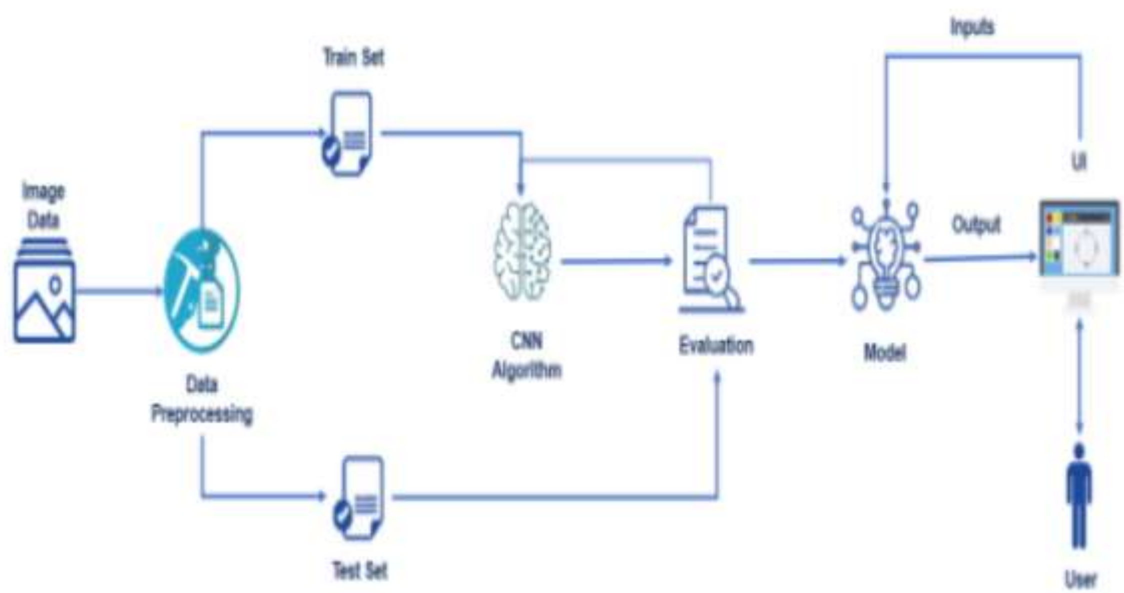
Pneumonia is an intense respiratory disease that influences the lungs. It is a lethal disease wherein the air sacs get loaded up with discharge and other fluid. Pneumonia is ranked eighth in the list of the top 10 causes of death in the United States. Pneumonia causes the passing of around 700,000 youngsters consistently and influences 7% of the worldwide populace. Chest X-rays are fundamentally utilized for the conclusion of this infection. Computer-aided diagnosis using artificial intelligence-based solutions is becoming increasingly popular these days. An issue with this illness is that occasionally, the highlights that depict the actual presence of the sickness regularly get blended in with different infections, and thus, radiologists think that it's difficult to analyse this illness. Deep learning techniques take care of this load of issues, and their exactness in the forecast of the infection is something very similar and, in some cases, considerably more prominent than a normal radiologist. Among the profound learning procedures, convolutional neural networks (CNNs) have shown extraordinary guarantee in picture grouping and division and hence are generally embraced by the exploration local area. Biomedical diagnosis that utilizes deep learning and computer vision has demonstrated to be extremely useful to give a speedy and exact conclusion of the illness that coordinates with the precision of a dependable radiologist.

1.1.OVERVIEW

The risk of pneumonia is immense for many, especially in developing nations where billions face energy poverty and rely on polluting forms of energy. The WHO estimates that over 4 million premature deaths occur annually from household air pollution-related diseases including pneumonia. Over 150 million people get infected with pneumonia on an annual basis especially children under 5 years old . In such regions, the problem can be further aggravated due to the dearth of medical resources and personnel. For example, in Africa's 57 nations, a gap of 2.3 million doctors and nurses exists. For these populations, accurate and fast diagnosis means everything. It can guarantee timely access to treatment and save much needed time and money for those already experiencing poverty. Deep neural network models have conventionally been designed, and experiments were performed upon hem by human experts in a continuing trial-anderror method. This process demands enormous time, know-how, and resources. To overcome this problem, a novel but simple model is introduced to automatically perform optimal classification tasks with deep neural network architecture. The neural network architecture was specifically designed for pneumonia image classification tasks.

The proposed technique is based on the convolutional neural network algorithm, utilizing a set of neurons to convolve on a given image and extract relevant features from them. Demonstration of the efficacy of the proposed method with the minimization of the computational cost as the focal point was conducted and compared with the exiting stateof-the-art pneumonia classification networks. Pneumonia is an inflammatory response in the lung sacs called alveoli. Its often caused by bacteria, viruses, fungi and other microbes. As the germs reach the lung, white blood cells act against the germ and inflammation occurs in the sacs. Thus, alveoli get filled with pneumonia fluid and this fluid causes symptoms like coughing, trouble in breathing and fever. If the infection isnt acted upon during the early periods of the disease, pneumonia infection can spread throughout the body and result in the death of the individual, as a result of the inability to exchange gas in the lungs. In recent times, CNN-motivated deep learning algorithms have become the standard choice for medical image classifications although the state-of-the-art CNN-based classification techniques pose similar fixated network ar- chitectures of the trial-and-error system which have been their designing principle. U-Net, SegNet, and Car- diacNet are some of the prominent architectures for medical image examination.

ARCHITECTURE:



1.2. PURPOSE:

This project is about building a web application that can determine a patient to have pneumonia effectively by breaking down its X-ray picture. A productive model for the recognition of pneumonia prepared on advanced chest X-ray pictures is proposed, which could help the radiologists in their dynamic interaction. An original methodology dependent on a weighted classifier is presented, which consolidates the weighted expectations from the cutting-edge profound learning models like ResNet18, Exception, InceptionV3, DenseNet121, and MobileNetV3 in an ideal way. This methodology is a regulated learning approach in which the organization predicts the outcome dependent on the nature of the dataset utilized. The proposed methodology uses a deep transfer learning algorithm that extracts the featuresfrom the X-ray image that describes the presence of disease automatically and reports whether it is a case of pneumonia.

2. LITERATURE SURVEY

Research Paper 1

Okeke Stephen, Mangal Sain, Uchenna, and Do-Un Jeong proposed a convolutional neural network model trained from scratch to classify and detect the presence of pneumonia from a collection of chest X-ray image samples. Unlike other strategies that depend entirely on transfer learning draws near or conventional handmade strategies to accomplish a wonderful arrangement performance, they built a convolutional neural network model from scratch to extract features from a given chest X-beam picture and group it to decide whether an individual is tainted with pneumonia. This model could help moderate the dependability and interpretability challenges regularly confronted when managing medical imagery.

Research Paper 2

By Daniel S Kermany, Michael Goldbaum, Wenjia paper, here, it established a diagnostic tool based on a deep-learning framework for the screening of patients with common treatable blinding retinal diseases. The exhibition of model relies exceptionally upon the loads of the pre-trained model. Subsequently, the presentation of this model would probably be upgraded when tried on a bigger ImageNet dataset with further developed deep-learning techniques and architecture.

Research Paper 3

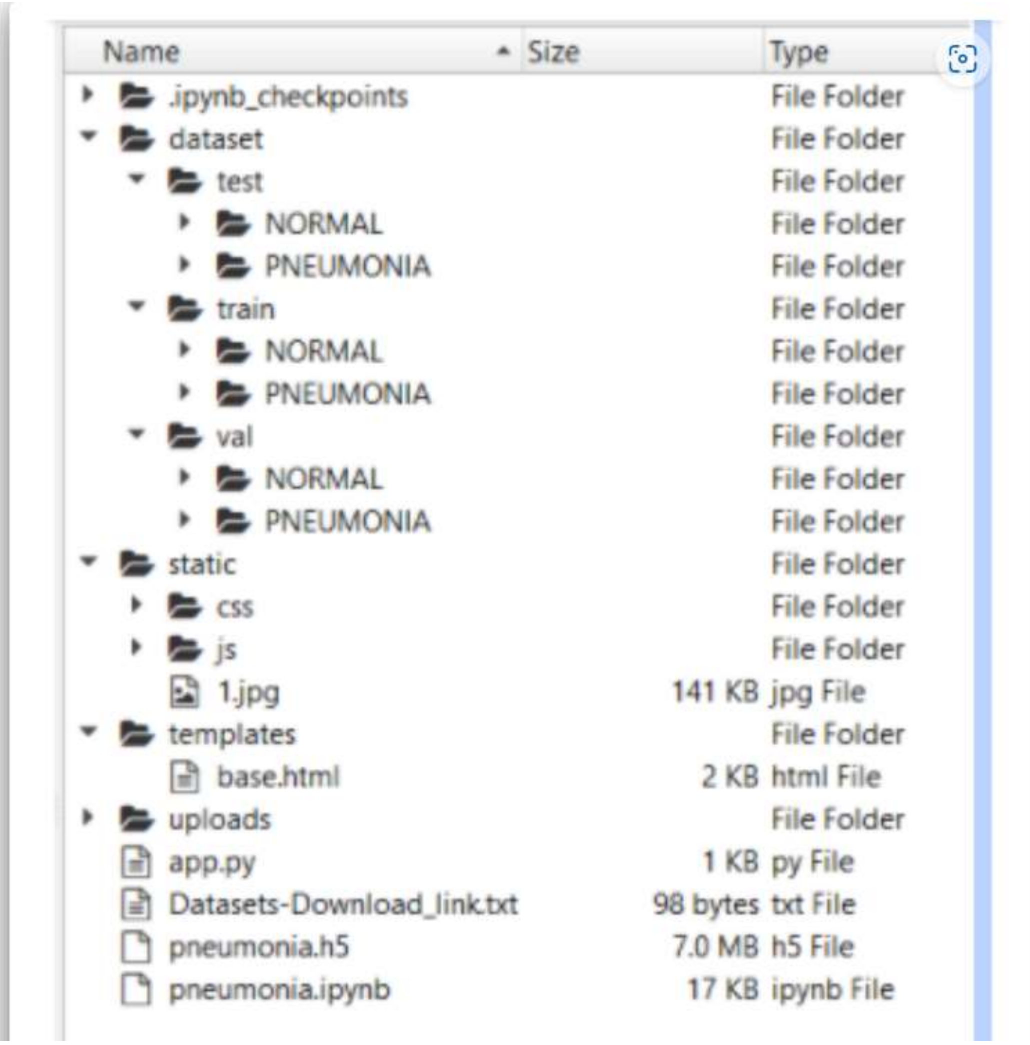
The proposed paper presents a deep neural network based on convolutional neural networks and residual network along with techniques of identifying optimum differential rates using cosine annealing and stochastic gradient with restarts to achieve an efficient and highly accurate network which will help detect and predict the presence of pneumonia using chest x-rays

3.Project Flow

Find below the project flow to be followed while developing the project.

- Download the dataset.
- Classify the dataset into train and test sets.
- Add the neural network layers.
- Load the trained images and fit the model.
- Test the model.
- Save the model and its dependencies.
- Build a Web application using flask that integrates with the model built.

4. PROJECT STRUCTURE



- The dataset folder contains three folders, test, train and validation, each of them having normal and infected x-ray images.
- Flask folder has all the files necessary to build the flask application.
  - static folder has the style sheets, script files and images that are needed in building the web page.
  - templates folder has the HTML page.
  - uploads folder has the uploads made by the user.
  - app.py is the python script for server side computing.
  - .h5 file is the model file which is to be saved after model building.
- pneumonia.ipynb is the training note.

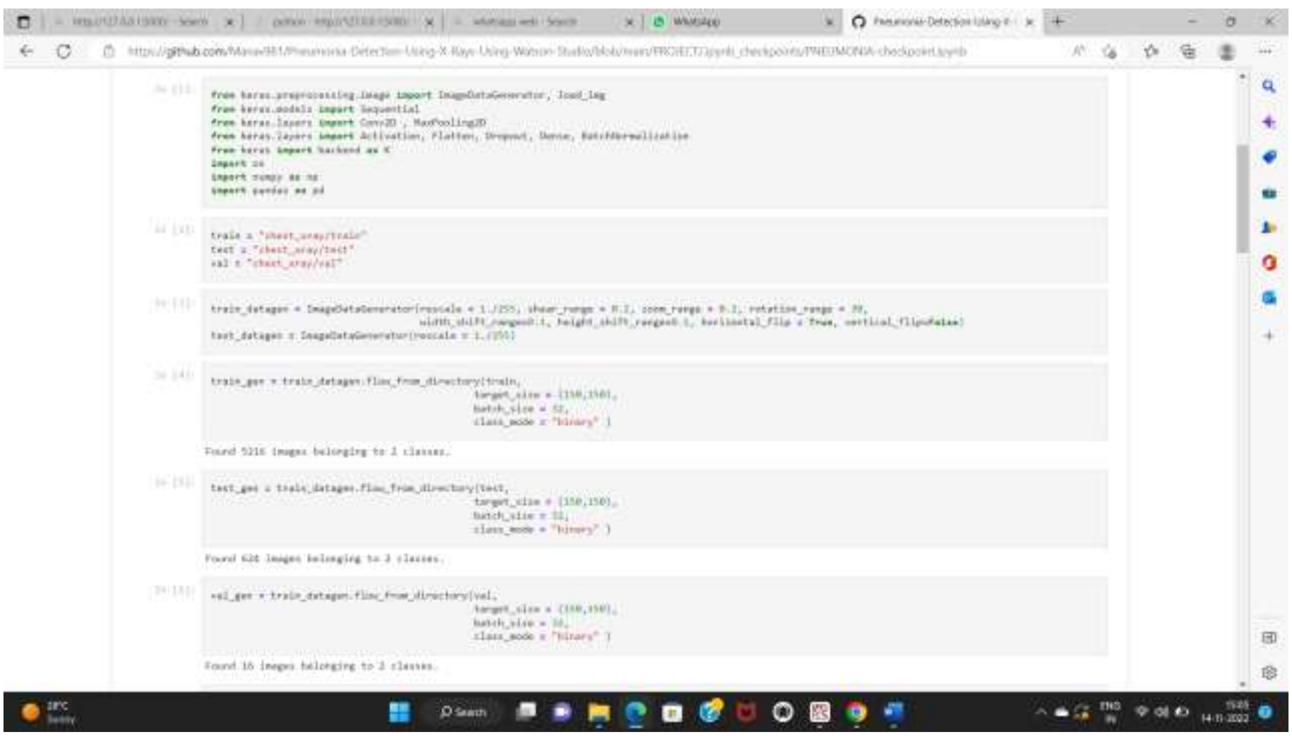
5. Image Preprocessing

In this milestone, we will be pre-processing the Data that is collected.

Image Pre-processing includes the following main tasks

- Import ImageDataGenerator Library.
- Configure ImageDataGenerator Class.
- Applying ImageDataGenerator functionality to the trainset and test set.

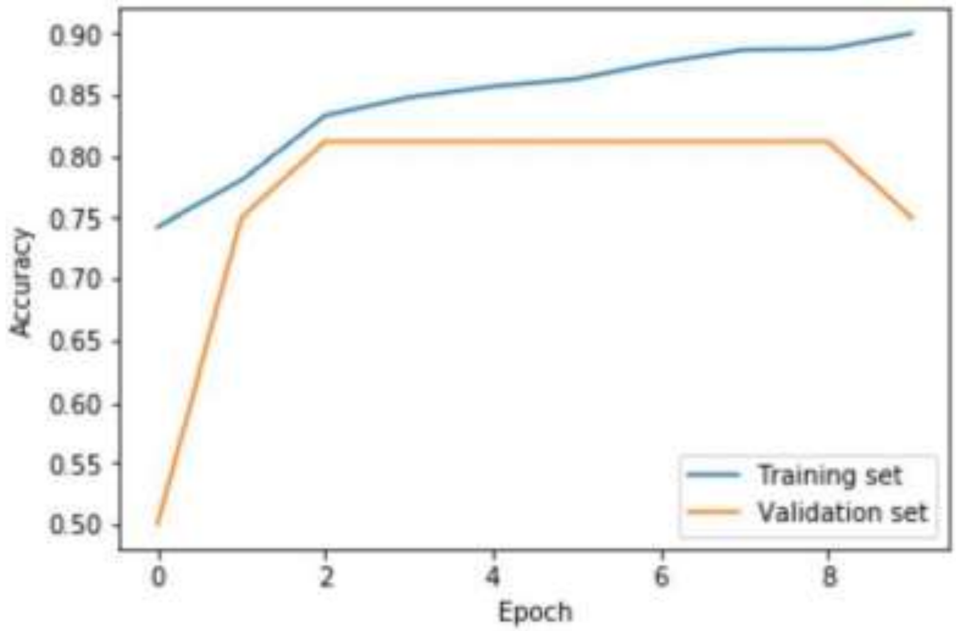
Note: The ImageDataGenerator accepts the original data, randomly transforms it, and returns only the new, transformed data.



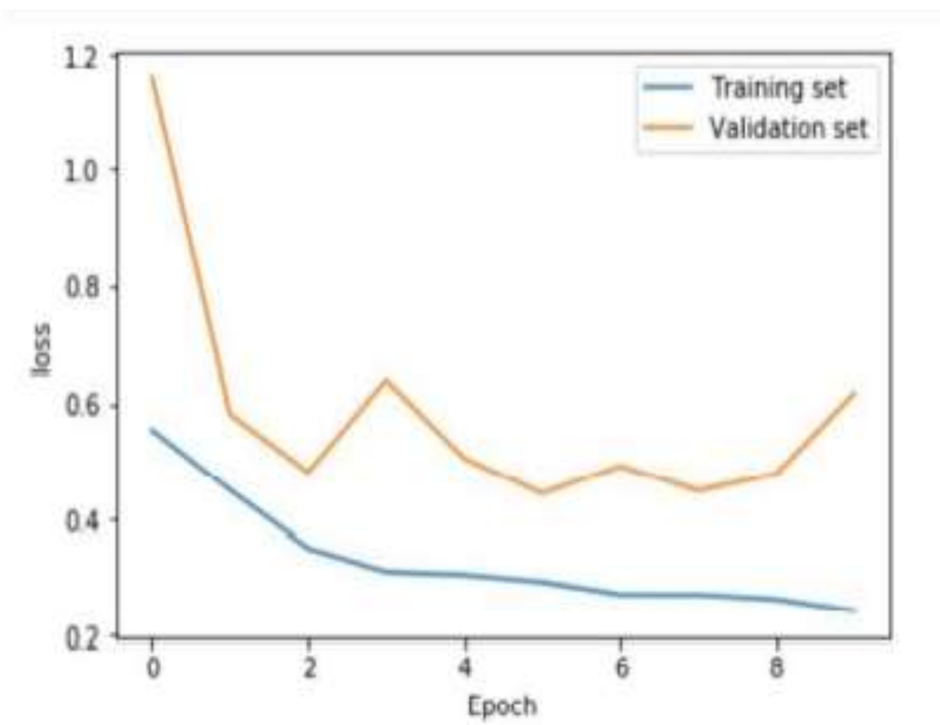
6.MATHEMATICAL DETAILS BEHIND WORKING OF CONVOLUTIONAL NEURAL NETWORKS

CNN is primarily used to look for the patterns in an image. Feature selection in case of images which has hundreds and thousands of dimensions would be a painful process. Feature selection happens all by itself during backpropagation over number of iterations. The input format of CNN is matrix and that of multilayer perceptron is tensor. Convolutional networks were inspired by biological processes in case of connectivity patterns of neurons. Unlike Multilayer Perceptron which uses fully connected layer, CNN uses different layers to detect patterns among images which are feed forwarded. Every image represent some pixels in simple terms. We analyze the influence of nearby pixels on a particular pixel in an image by using a filter or a kernel. Filters are tensors used track of spatial information. These filters learn to extract features like edge detection etc of objects in images in something called a convolutional layer. They help to filter out unnecessary or repetitive information during convolution operation. There are multiple types of filters like high pass, low pass, gaussian etc which are used for different sort of operations. Then the convolution image matrix multiplies with filter matrix which is called Feature Map. Each filter is strided over the image using ot products between filter matrix and corresponding image pixel values. During the convolution operation size of image matrix decreases, if do not intend to do that then there is concept of padding. During padding redundant tensor of zero valued pixels is appended on edges of images, which preserve the size during convolutions. Next, we need to reduce the size of images, if they are too large. Pooling layers section would reduce the number of parameters when the images are too large. Adding pooling layer then decrease the size of the image and hence decrease the complexity and computations. Usually, an activation function ReLu is used in next layer. ReLU stands for Rectified Linear Unit for a non-linear operation. The output is $f(x) = max(0,x)$. The purpose of ReLu is to add non-linearity to the convolutional network. In usual cases, the real-world data want our network to learn non-linear patterns which is the purpose of activation functions. The final step is to flatten our matrix and feed the values to fully connected layer. We need to train the model in the same way, we train other neural networks. Using the certain number of epochs and then backpropagate to update weights and calculate the loss. VII. CONCLUSION In this study, we present a novel method for classifying an X-ray image on its possibility of exhibiting pneumonia in the earl

The accuracy curve is demonstrated in image below:



The Loss to Epoch graph is demonstrated in image as below:



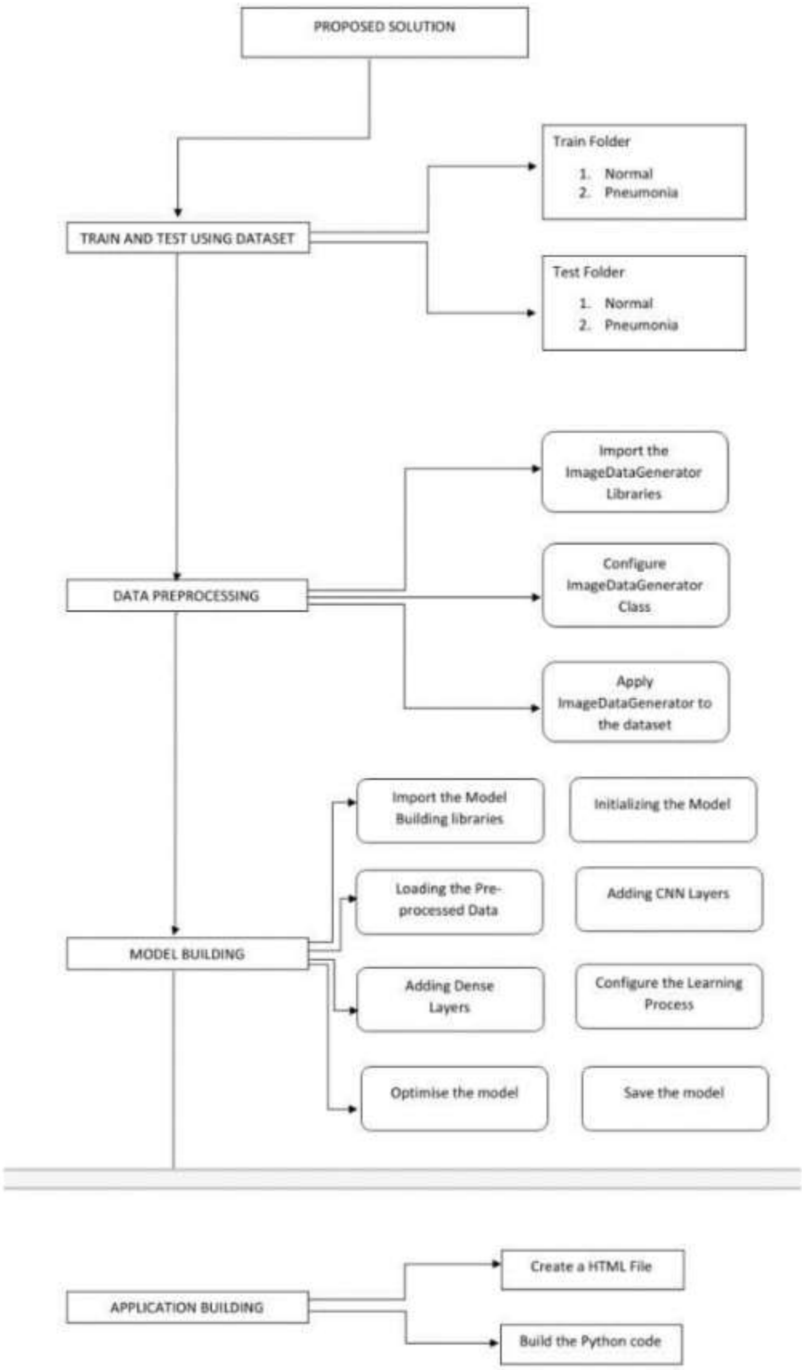
7.RELATED WORK AND COMPARISON

There has also been previous studies done on the early detection of pneumonia. Among the various other methods used by different studies, this paper focuses merely on Pneumonia and its classification. The experimentation was conducted in Koc University Artificial Intelligence Laboratory, Istanbul, Turkey. In their study they presented a novel method for classifying pneumonia existence in an x-ray image. They proposed a two-step image processing before training our deep learning model, in order for making the features of an x-ray image clearer and explicit for easing the classification process. They, then, executed a convolutional neural network followed by a residual neural network for the classification process. Their experimentation was conducted with similar means to ours. They used a 3 layer convolutional network for feature map acquisition for image preprocessing. In our experiment we use 3 convolutional layers, yielding a more efficient and a computationally less costly training process. Our preprocessing methods are similar to real life applications, unlike statistical means that might be ineffective when wide range of data is present. Our proposed architecture yields an accuracy of 85.73%, while their study yielded an accuracy of 78.73%.

8.THEORETICAL SURVEY

BLOCK DIAGRAM

The following is the block diagram of our proposed solution of using the Convolution Neural Networks (CNN) to analyse the images of the X-ray to predict



8.1. HARDWARE/SOFTWARE REQUIREMENTS:

Hardware requirements: Laptop Software requirements: Python-3.6, Keras, TensorFlow, Jupyter Notebook

● **Python**

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. It was created by Guido van Rossum, and first released on February 20, 1991. Its high-level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance.

● **Anaconda Navigator**

Anaconda Navigator is a free and open-source distribution of the Python and R programming languages for data science and machine learning related applications. It can be installed on Windows, Linux, and macOS.Conda is an open-source, cross-platform, package management system. Anaconda comes with so very nice tools like JupyterLab, Jupyter Notebook,QtConsole, Spyder, Glueviz, Orange, Rstudio, Visual Studio Code. For this project, we will be using Jupyter notebook and Spyder.

● **Jupyter Notebook**

The Jupyter Notebook is an open source web application that you can use to create and share documents that contain live code, equations, visualizations, and text. Jupyter Notebook is maintained by the people at Project Jupyter. Jupyter Notebooks are a spin-off project from the IPython project, which used to have an IPython Notebook project itself. The name, Jupyter, comes from the core supported programming languages that it supports: Julia, Python, and R.Jupyter ships with the IPython kernel, which allows you to write your programs in Python, but there are currently over 100 other kernels that you can also use.

● **Spyder**

Spyder, the Scientific Python Development Environment, is a free integrated development environment (IDE) that is included with Anaconda. It includes editing, interactive testing, debugging, and introspection features.

● **Tensor flow**

TensorFlow is an end-to-end open-source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries, and community resources that lets researchers push the state-ofthe-art in ML and developers can easily build and deploy ML-powered applications.

● **Keras**

Keras leverages various optimization techniques to make high-level neural network API easier and more performant.

● **Flask**

Web framework used for building. It is a web application framework written in python which will be running in local browser with a user interface. In this application, whenever the user interacts with UI and selects emoji, it will suggest the best and top movies of that genre to the user.

9. EXPERIMENTAL INVESTIGATIONS

The dataset comprised a total of 5856 images and divided into two main parts, the training dataset testing dataset, validation dataset. Bacterial and viral were considered as one category i.e., pneumonia infected. Finally, there were 5216 images in the training dataset, 624 images in the testing dataset, 16 images in the validation dataset. The below figure shows the chest X-ray images of a healthy person and the other of a person suffering from pneumonia.

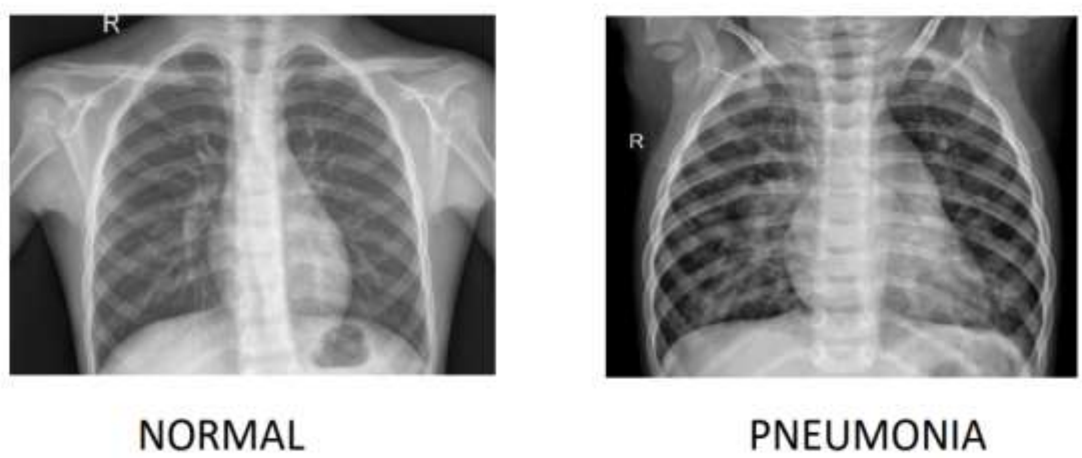
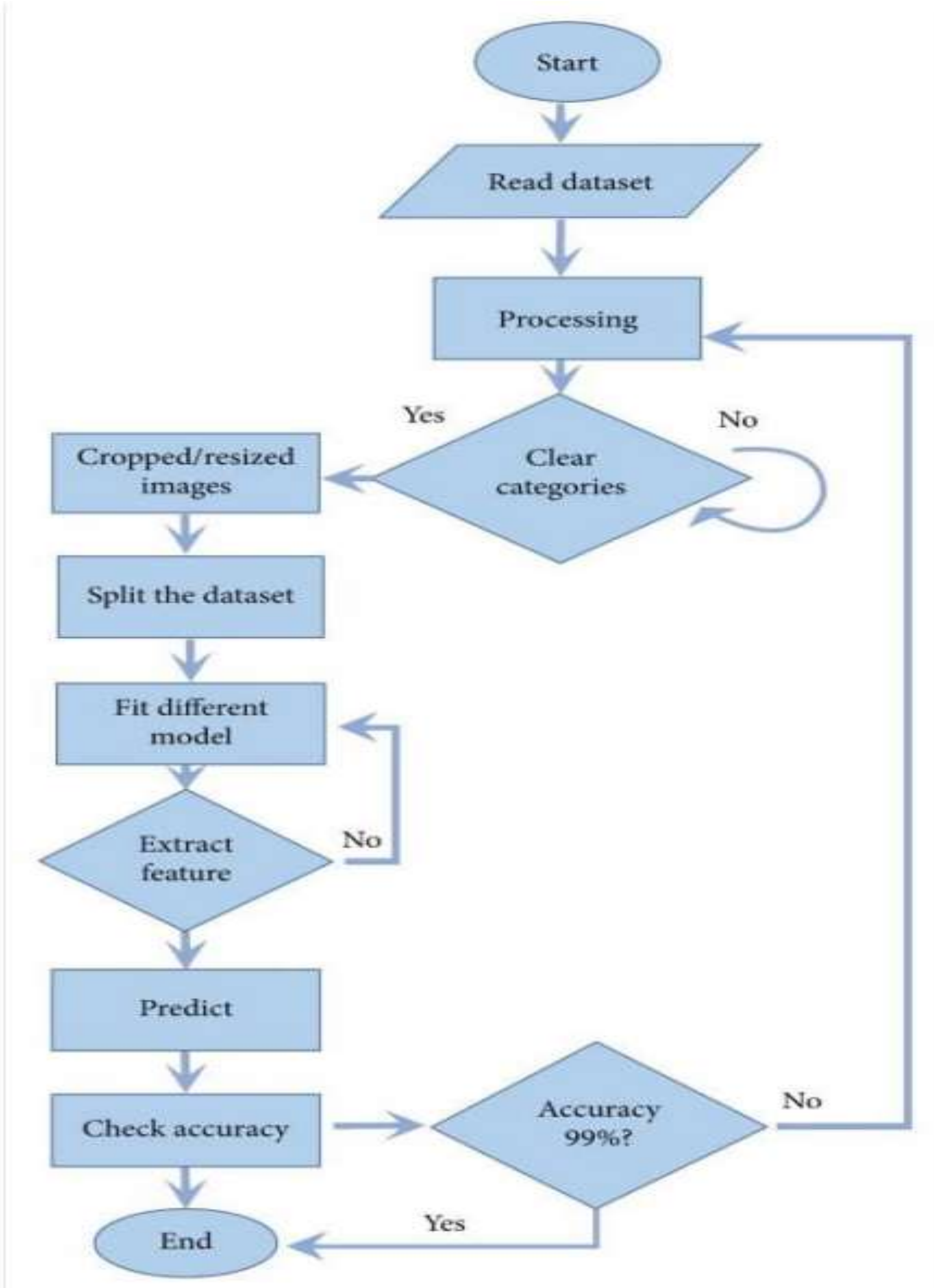


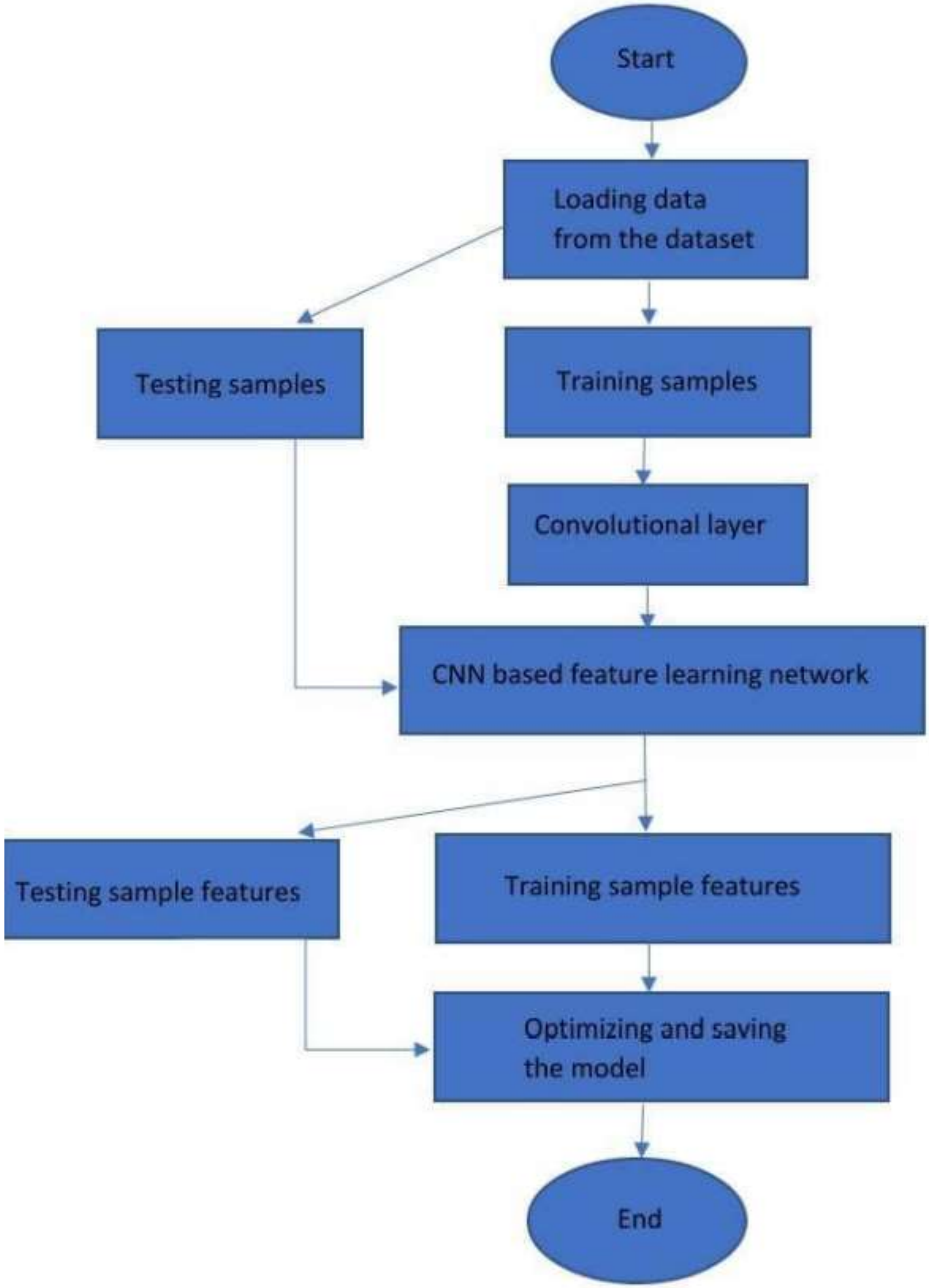
TABLE 1

Description of experimental dataset

CATEGORY	TRAINING SET	TESTING SET	VALIDATION SET
Normal	1341	234	8
Pneumonia	3875	390	8
Total	5216	624	16
Percentage	89.07	10.65	0.27

10.FLOWCHART

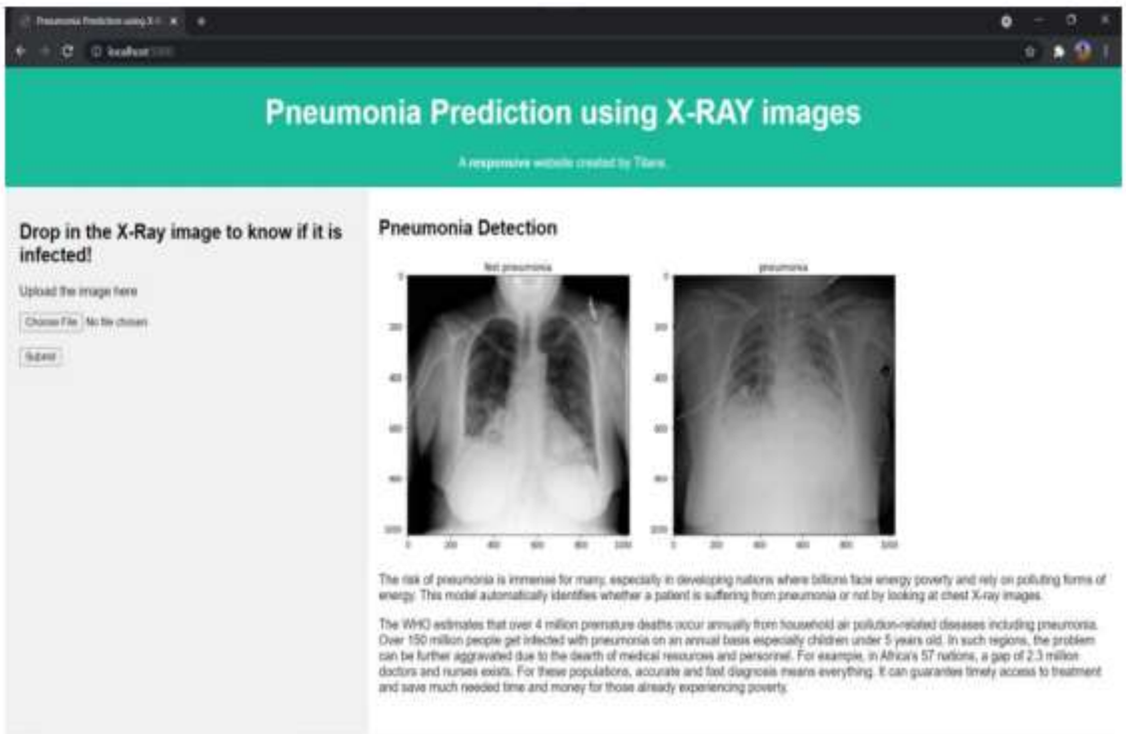



Start
Loading data
from the dataset
Testing samples
Training samples
Convolutional layer
CNN based feature learning network
Testing sample features
Training sample features
Optimizing and saving
the model
End

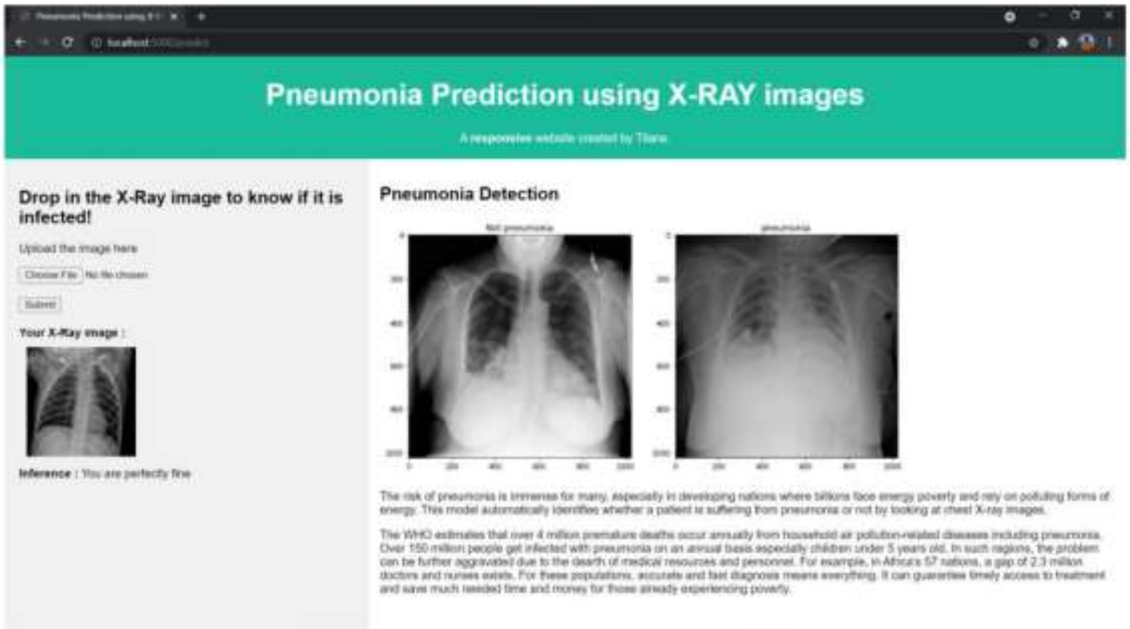
11. RESULTS

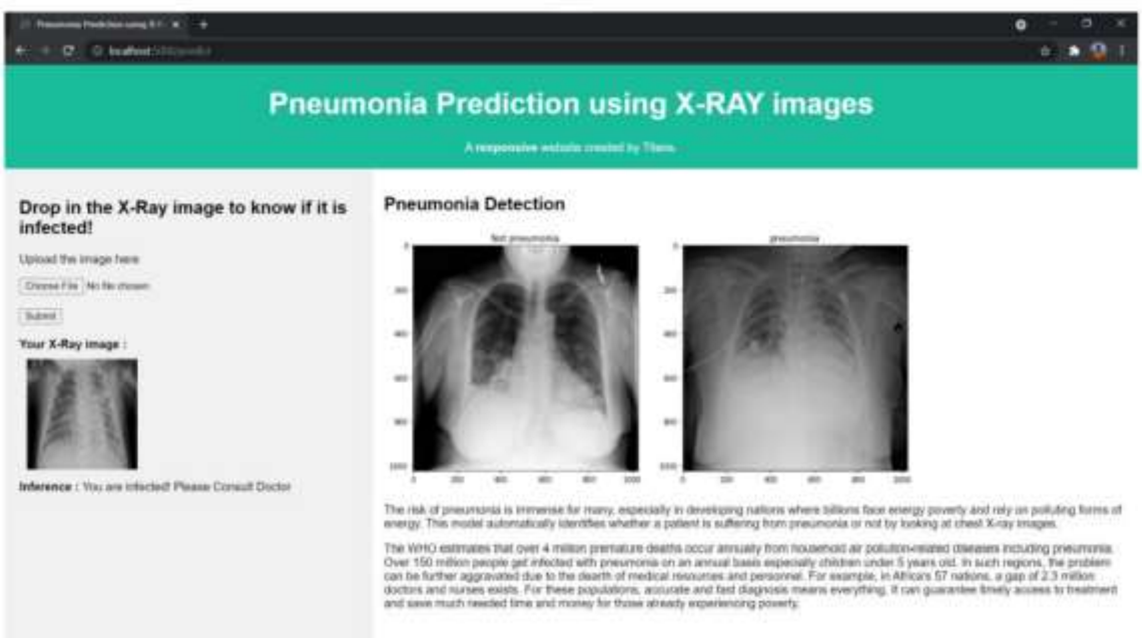
The following show the screenshots of our application of Pneumonia prediction using X-RAY images.

Home Page :



Prediction Page:





12 . ADVANTAGES AND DISADVANTAGES

ADVANTAGES	DISADVANTAGES
Easy model building with less formal statistical knowledge required.	Clinical interpretations of model parameters are difficult.
Capable of capturing interactions between predictors.	Sharing an existing ANN model is difficult.
Capable of capturing non linearities between predictors and outcomes.	Prone to overfitting due to the complexity of model structure.

13.APPLICATIONS

The analysis of X-RAYS using CNN is right now being utilized at different clinical establishments including Singapore’s Changi General hospital (desktop application i.e., installed on radiology workstations) etc... The application of CNN in Pneumonia detection using X-RAYS assists us with using different libraries present to help identify the severity of pneumonia that correlates to the degree of Chest X-RAY (CXR) lung image abnormality

14. CONCLUSIONS

The following is the conclusion of our project:

Unsolved problem that is as of now tormenting the doctors around the world. Our proposed model is designed and advancement to detect and classify pneumonia from chest X-ray pictures. It contains both image processing and convolutional neural network. We developed a model; the algorithm starts by transforming chest X-ray images into sizes less than the original. The next step includes the identification and classification of pictures by the convectional neural network structure, which extracts features from the image and classify them. This work has presented the X-Ray images for Pneumonia discovery based on convolutional neural networks and diverse machine learning. By training a bunch of solid CNNs for an enormous scope dataset, we built a model that can precisely predict Pneumonia. During each epoch data is trained again and again to learn the feature of data. The presentation assessment of the model is estimated by utilizing classification accuracy and cross-validation.

15. FUTURE SCOPE

Pneumonia comprises a significant cause of morbidity and mortality. It represents an impressive number of grown-up emergency clinic affirmations, and countless those patients ultimately die. According to the WHO, pneumonia can be prevented with a simple intervention and early diagnosis and treatment. Nevertheless, most of the worldwide population needs admittance to radiology diagnostics. In any event, when there is the accessibility of imaging hardware, there is a lack of specialists who can analyse X-rays. It is no doubt that the predictive model can be worked on far superior by performing data augmentation or carrying out transfer learning concept which works with the model an opportunity to get better. Along these lines, this will be added as additional upgrade in the forthcoming stories. In the future, it is intriguing to see approaches in which the weights relating to various models can be assessed all the more proficiently and a model that considers the patient's set of experiences while making predictions.

16. BLILOGRAPHY

1. Pneumonia. [(accessed on 31 December 2019)]; Available online: https://www.radiologyinfo.org/en/info.cfm?pg-pneumonia.

2. Johns Hopkins Medicine Pneumonia. [(accessed on 31 December 2019)]; https://www.hopkinsmedicine.org/health/conditions and-diseases/pneumonia.

3. https://www.hindawi.com/journals/mpe/2021/9929274/

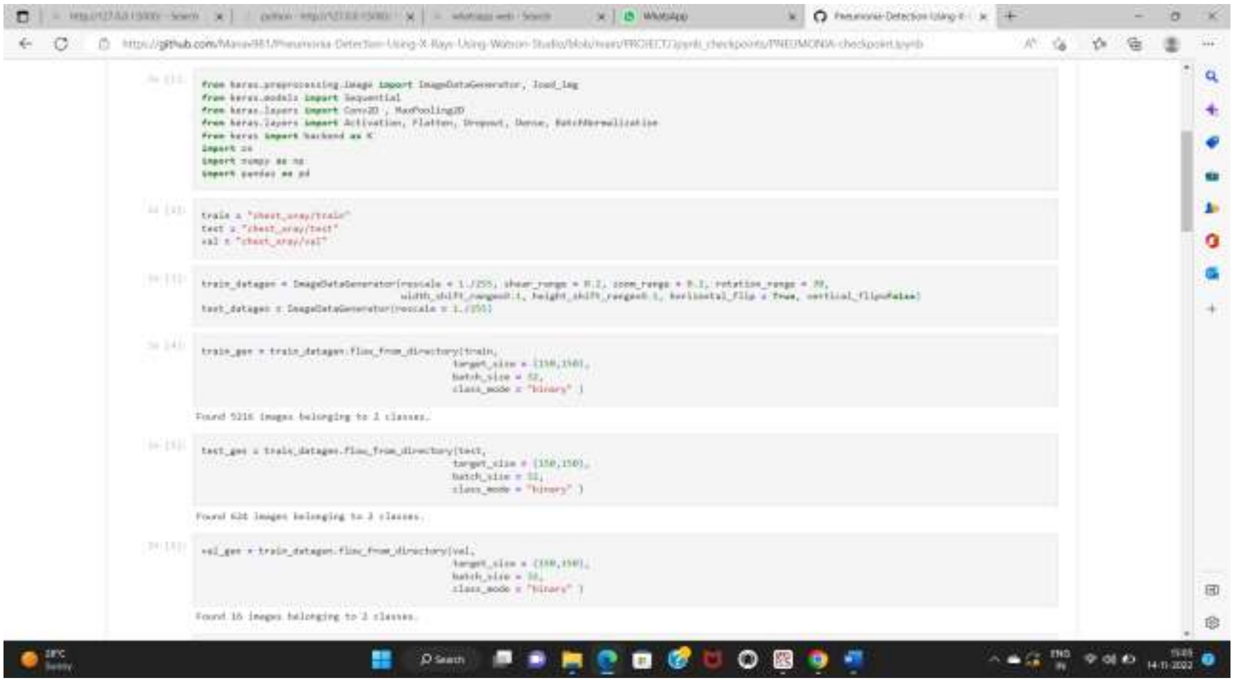
4. Deniz Yagmur Urey, Can Jozef Saul, and Can Doruk Taktakoglu Robert College of Istanbul, Istanbul, Turkey Koc University Artificial Intelligence Laboratory, Istanbul, Turkey Early Diagnosis of Pneumonia with Deep Learning.

5. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770- 778).

6. CheXNet:Radioogist-Level Pnemonia Detection on Chest -Ray with Deep Learning. https://arxiv.org/abs/1711.05225v3

17.APPENDIX

SOURCE CODE:



```
from keras.preprocessing.image import ImageDataGenerator, load_img
from keras.models import Sequential
from keras.layers import Conv2D , MaxPooling2D
from keras.layers import Activation, Flatten, Dropout, Dense, BatchNormalization
from keras import backend as K
import os
import numpy as np
import pandas as pd
```

```
train = "chest_xray/train"
test = "chest_xray/test"
val = "chest_xray/val"
```

```
train_datagen = ImageDataGenerator(rescale = 1./255, shear_range = 0.2, zoom_range = 0.2, rotation_range = 20,
                                   width_shift_range=0.1, height_shift_range=0.1, horizontal_flip = True, vertical_flip=False)
test_datagen = ImageDataGenerator(rescale = 1./255)
```

```
train_gen = train_datagen.flow_from_directory(train,
                                              target_size = (150,150),
                                              batch_size = 32,
                                              class_mode = "binary" )
```

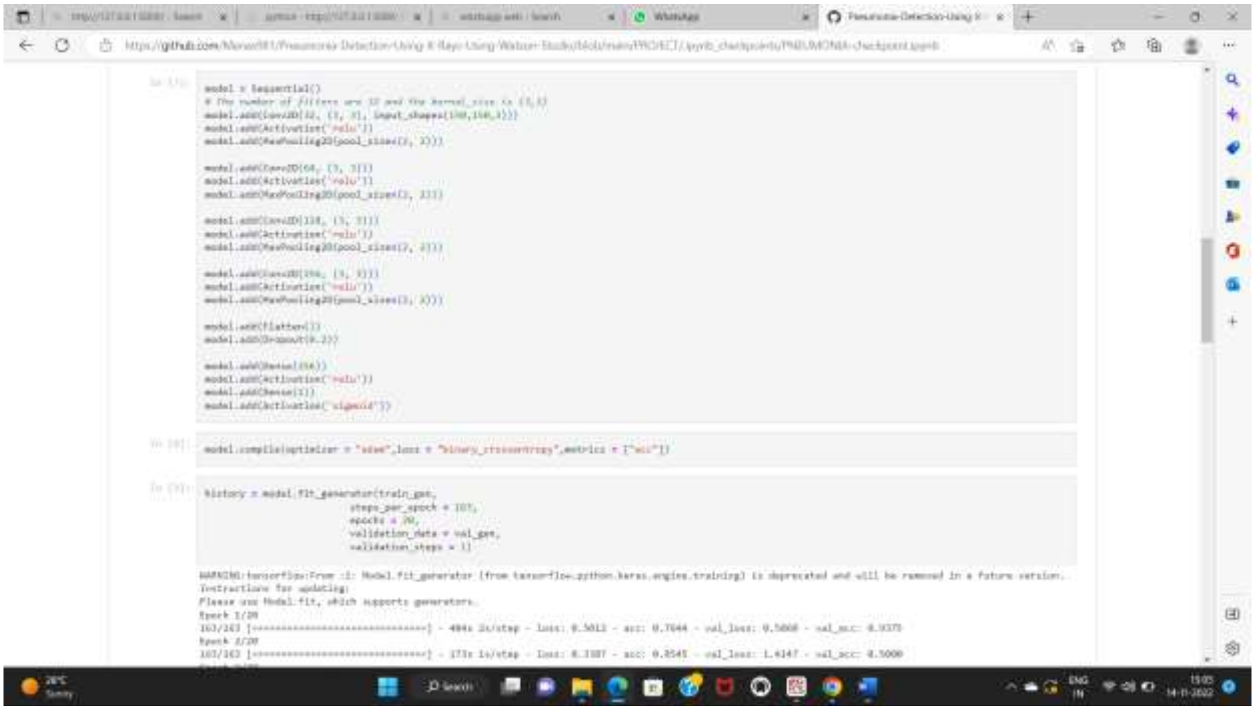
Found 5216 images belonging to 2 classes.

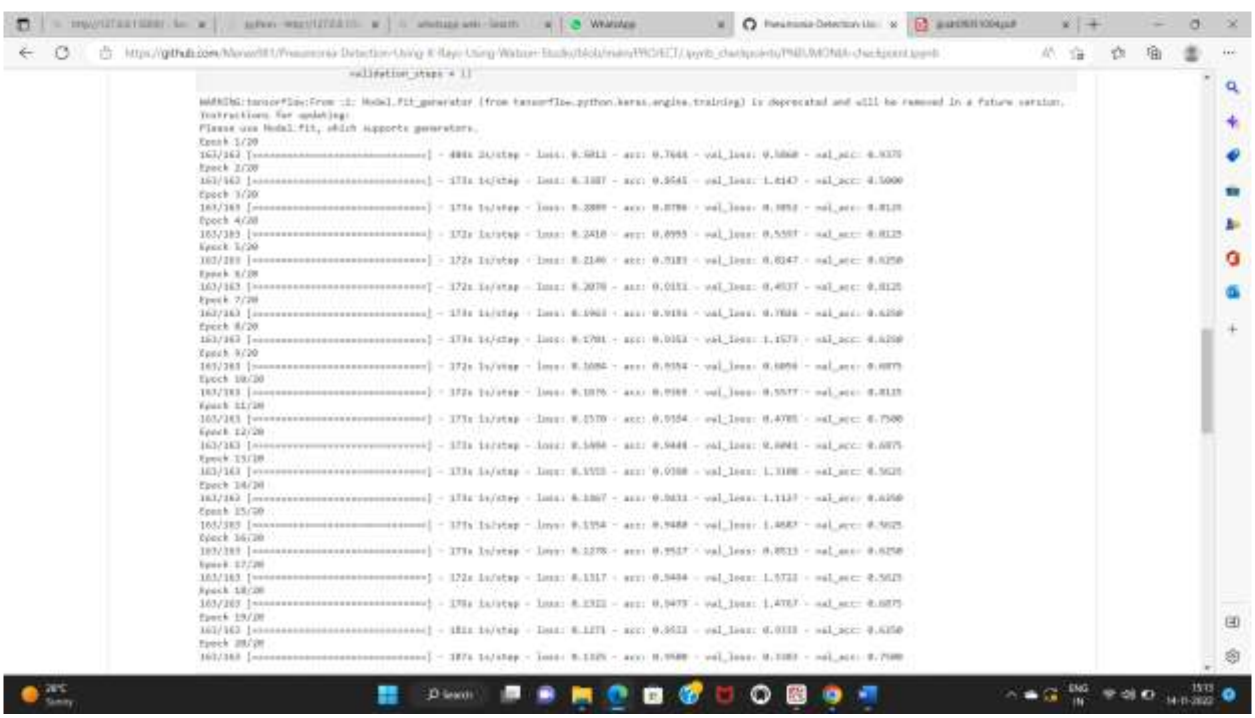
```
test_gen = train_datagen.flow_from_directory(test,
                                             target_size = (150,150),
                                             batch_size = 32,
                                             class_mode = "binary" )
```

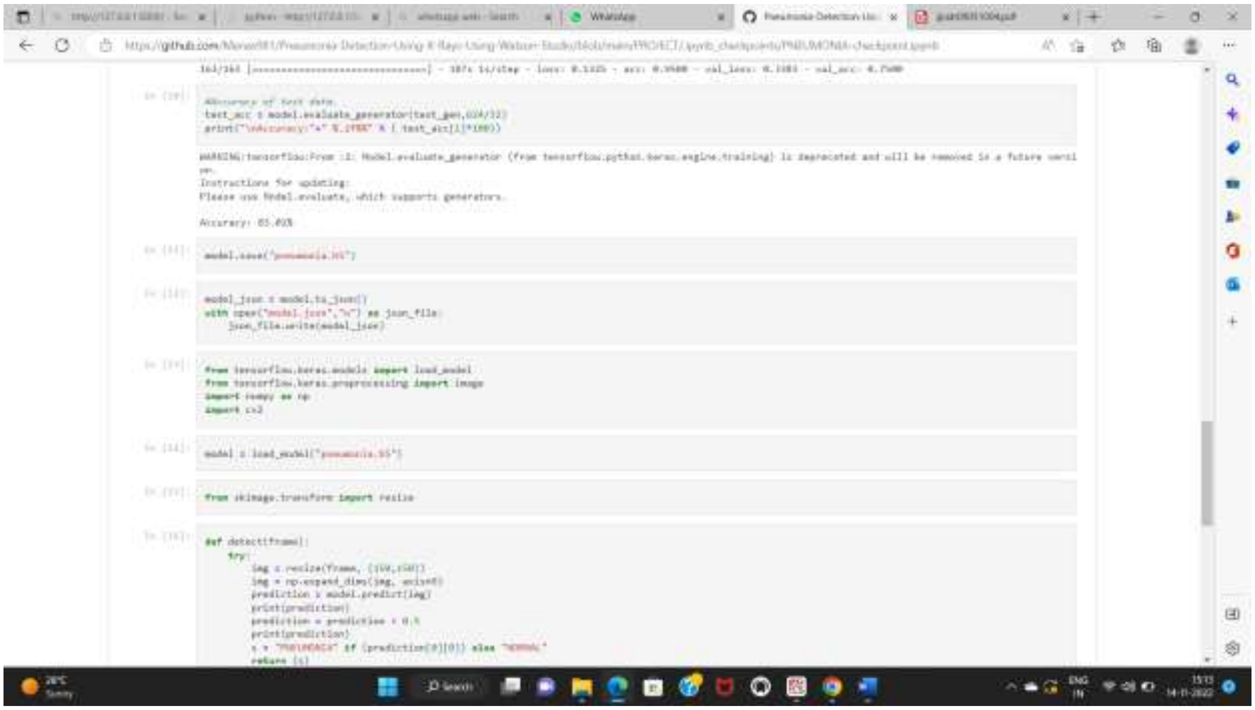
Found 624 images belonging to 2 classes.

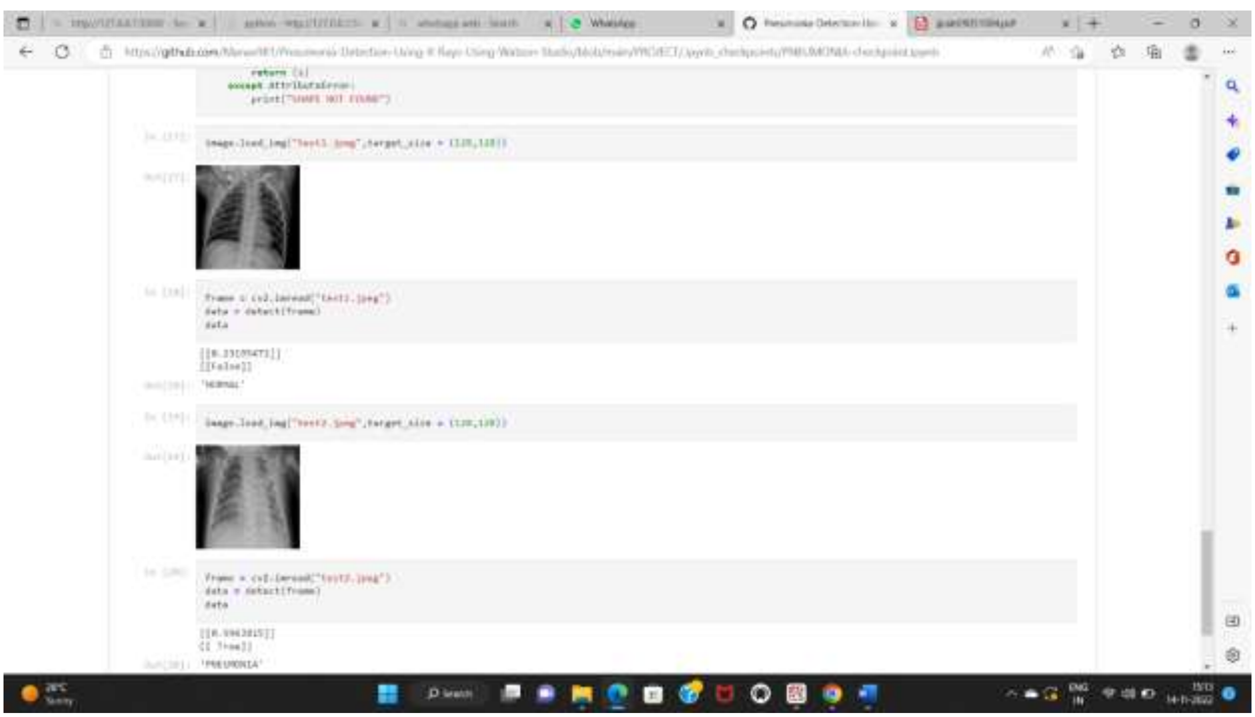
```
val_gen = train_datagen.flow_from_directory(val,
                                            target_size = (150,150),
                                            batch_size = 32,
                                            class_mode = "binary" )
```

Found 16 images belonging to 2 classes.






WhatsApp
Instructions for updating:
Please use Model.evaluate, which supports generators.
from tensorflow.keras.models import load_model
from tensorflow.keras.preprocessing import image
import numpy as np
import cv2
from skimage.transform import resize
def detect(frame):


WhatsApp
'NORMAL'
'PNEUMONIA'