PNEUMONIA PREDICTION USING X-RAY IMAGES

Using Convolution Neural Network

Developed by: Gone Shivakumar , Manchala Nikhil Bharadwaj , Syeda Sara Sarwath , Yada Naveen
Smart Bridge - Remote Summer Internship Program

1 INTRODUCTION:

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 them by human experts in a continuing trial-and-error 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 state-of-the-art pneumonia classification networks.The proposed prediction model is implemented by Convolutional Deep Neural Networks (CNN) using Python programming language and the model is converted into a Website Application. Different machine learning algorithms are trained to measure the performance of CNN with popular and modern classifiers after pre-processing of data. Promising results are achieved, when the results of regular classifiers like SVM, random forest, adaboost, etc. are compared with the the suggested framework using different estimating metrics like accuracy, specificity, area under the curve, and sensitivity, etc.

1.1 OVERVIEW:

Pneumonia is among the top diseases which cause most of the deaths all over the world.

Virus, bacteria and fungi can all cause pneumonia. However, it is difficult to judge the pneumonia just by looking at chest X-rays. The aim of this study is to simplify the pneumonia detection process for experts as well as for novices. We suggest a novel deep learning framework for the detection of pneumonia using the concept of transfer learning. In this approach, features from images are extracted using dierent neural network models pretrained on ImageNet, which then are fed into a classifier for prediction. We prepared five dierent models and analyzed their performance. Thereafter, we proposed an ensemble model that combines outputs from all pretrained models, which outperformed individual models, reaching the state-of-the-art performance in pneumonia recognition. Our ensemble model reached an accuracy of 96.4% with a recall of 99.62% on unseen data from the Guangzhou Women and Children's Medical Center dataset.

1.2 Purpose:

Our aim from the project is to make use of numpy from
python to extract the libraries for machine learning for the pneumonia prediction.
Secondly, to learn how to hyper tune the parameters using grid search cross validation for the
CNN– deep learning algorithm.
And in the end, to predict whether the person is predicted with pneumonia or not by combining the predictions from multiple deep learning algorithms
and withdrawing the conclusions

2 Literature Survey:

Over the last decade, several machine learning based automated methods for identifying different types of pneumonia have been widely studied. Fiszman et al used a natural language processing (NLP) tool to identify acute bacterial pneumonia-related disease in chest X-ray. Performance of this type of resource intensive application is very much comparable to that of the human expert. Chapman et al demonstrated three computerized methods using a rule base, a probabilistic Bayesian network, and a decision tree to diagnose the chest X-ray report associated with acute bacterial pneumonia. In a study of feasibility of an NLP-based monitoring system is done to identify healthcare-associated pneumonia in neonates. However, practical clinical applications of these types of methods are limited due to the dependency on the

information extracted from the narrative reports of the patients. Parveen et al reports an unsupervised fuzzy c-means classification learning algorithm to detect pneumonia infected X-ray images. This approach improves classification accuracy as fuzzy c means allocate weights to all the pixels of the input X-ray images. Rajpurkar et al demonstrated ChexNet, a 121-layer deep convolutional neural network (CNN), that provides the probability of detecting or identifying pneumonia using a heatmap to localize the area of the infection. Kermany et al introduced a transfer learning-based DL framework to diagnose paediatric pneumonia using chest X-ray images. However, none of the methods are exploited to classify X-ray images with pneumonia for the CS framework to meet the need of remote end analysis. Researchers are utilizing the results of machine learning predictions for solving problems of life science. Large volume of information can be extracted and used for future prevention of dangerous diseases. For extracting in
formation from the medical images Python language is used. High dimensional data consists of the medical images which has sizable amount of feature descriptors. To extract the numerous feature descriptors, feature extraction techniques are applied on good quality X-ray images. Deep learning neural networks are trained with the extracted data to create the prediction model. Research is additionally carried for prediction of pneumonia using machine learning classifiers.

2.1 Existing problem:

Description :

In general, a patient suffering from Pneumonia goes to the hospital to take an X-ray image waits for the doctor and then the doctor will check the X-ray then he decides whether the person has pneumonia or not. The results are not only concluded based on just seeing the X-ray images but furthermore, tests were conducted on the patient to verify the results of the doctor. The process is time-consuming and if the patient has severe pneumonia or not he has to wait several days to get the test results. But in recent developments of the artificial intelligence and the computational powers of the computers have increased it helps in predicting pneumonia by just passing the X-ray image as an input to our model.

2.2 Proposed Solution:

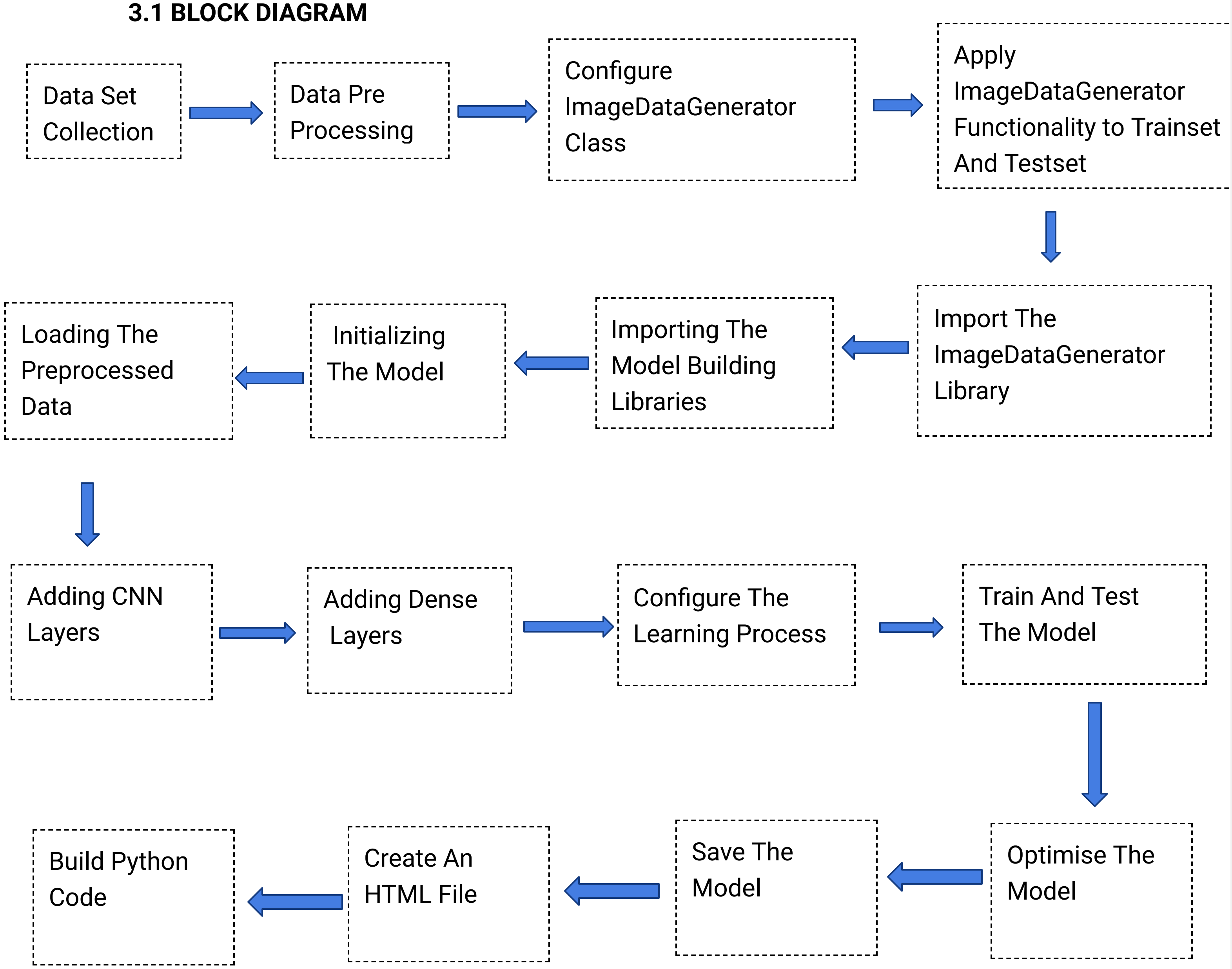
The main objective of this project is to help the doctors to predict the pneumonia disease more accurately using a deep learning model. The objective is not only to help the doctors but also to the patients to verify whether they have pneumonia or not. By using this model we can precisely predict pneumonia. A convolutional neural network model is built from scratch to extract features from a given chest X-ray image and classify it to determine if a person is infected with pneumonia. a web is built where the user can upload the x - ray image and the result is shown on the UI.

3 Theoritical Analysis:

While selecting the algorithm that gives an accurate prediction we gone through lot of algorithms which gives the results abruptly accurate and from them we selected only one algorithm for the prediction problem that is convolutional neural network, it assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. thats how the prediction work great with the convolutional neural network.

At first we got like lot of worst accuracies because we tried lot of algorithms for the best accurate algorithm , finally after all of that we tried the best suitable algorithm which gives the prediction accurately is convolutional neural network. And developed it to use as a real time prediction probelm for the pneumonia prediction using xray images

3.1 BLOCK DIAGRAM



3.2 Software Designing

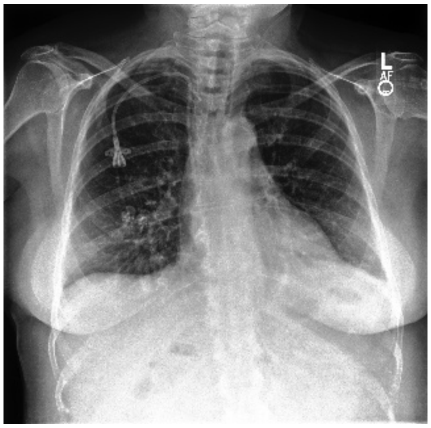
●Google Collab/Jupyter
●Sublime Text/Spyder IDE
● Machine Learning Algorithms (Convolution Neural Networks)
● Python (numpy, keras,tensorflow)
● HTML
● Flask

We developed this Pneumonia Prediction using X-Ray image by using the Python language which is a interpreted and high level programming language and using the Machine Learning algorithms. for coding we used theGoogle Collab of the Anaconda distributions and the Spyder, it is an integrated scientific programming in the python language. For creating an user interface for the prediction we used the Flask. It is a micro web framework written in Python. It is classified as a microframework because it does not require particular tools or libraries. It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions, and a scripting language to create a webpage is HTML by creating the templates to use in th functions of the Flask and HTML

4 EXPERIMENTAL INVESTIGATION:

The dataset comprised a total of 11680 images (Table 1) segmented into two main parts, a training set and a test set. Both bacterial and viral pneumonia were considered as a single category, pneumonia infected. The dataset used in this study did not include any case of viral and bacterial co-infection. All chest X-ray images were taken during the routine clinical care of the patients. Two expert physicians then graded the diagnoses for the images before being cleared for training the AI system. The evaluation set was also checked by a third expert to account for any grading errors. The proportion of data assigned to training and testing was highly imbalanced. Therefore, the dataset was shuffled and arranged into training and test sets only. Finally, there were 10432 images in the training set and 1248 images in the test set. Eleven-point-nine-five percent of the complete dataset was used as the testing dataset. Shows two chest X-ray images, one of a healthy person and the other of a person suffering from pneumonia.

Pneumonia



Normal

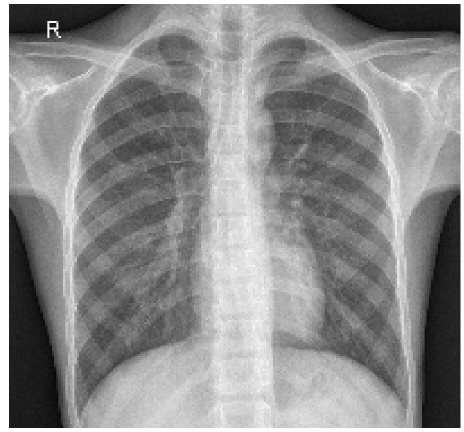
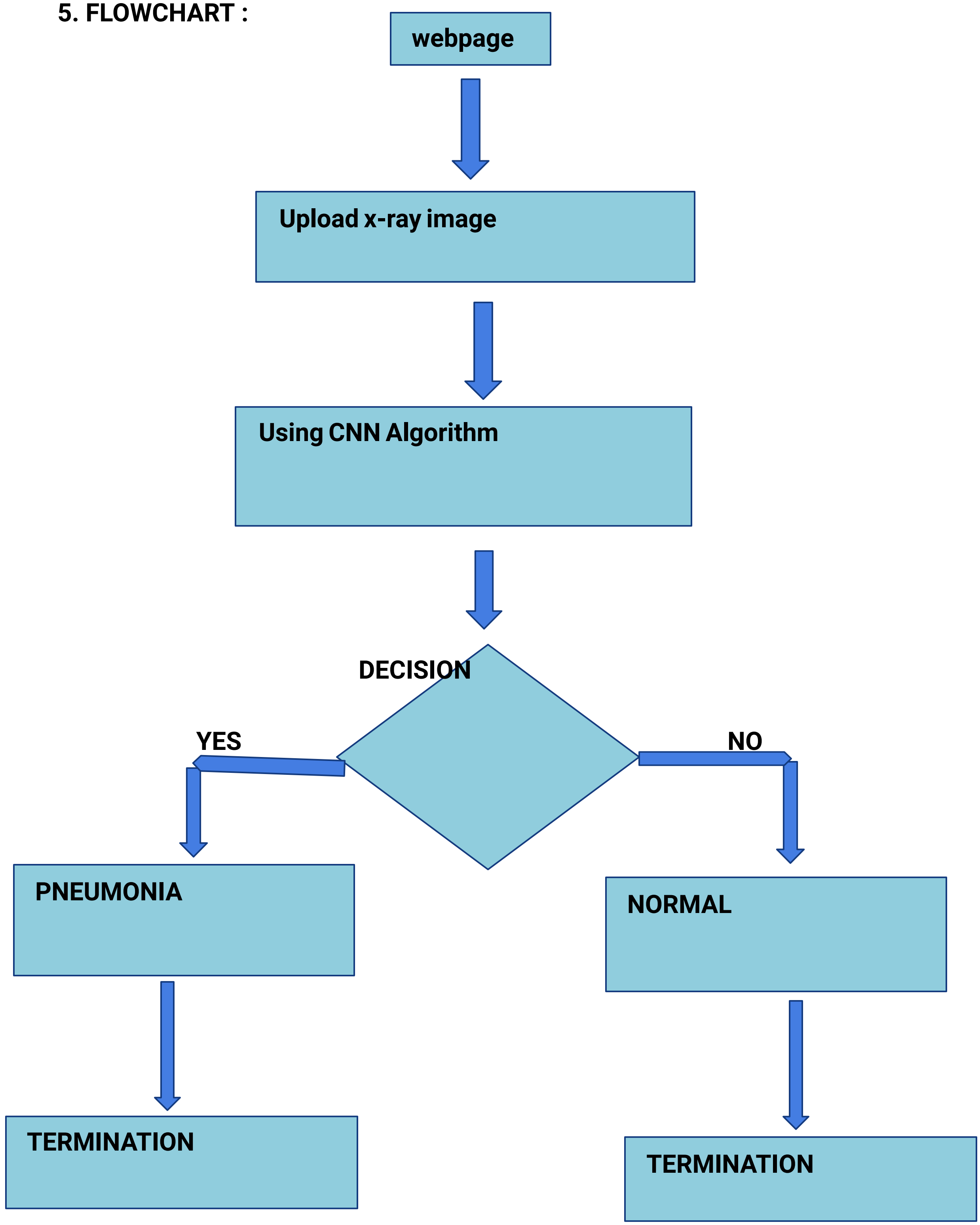


Table 1

Description of the experimental dataset.

Category	Training Set	Test Set
Normal (Healthy)	1283	300
Pneumonia (Viral + Bacteria)	3873	400
Total	5156	700
Percentage	88.05%	11.95%

5. FLOWCHART :


webpage
Upload x-ray image
Using CNN Algorithm
DECISION
YES
NO
PNEUMONIA
NORMAL
TERMINATION
TERMINATION

6 RESULT:

To evaluate and validate the effectiveness of the proposed approach, we conducted the experiments 10 times each for three hours, respectively. Parameter and hyperparameters were heavily turned to increase the performance of the model. Different results were obtained, but this study reports only the most valid.

As explained above, methods such as data augmentation, learning rate variation, and annealing were deployed to assist in fitting the small dataset into deep convolutional neural network architecture. This was in order to obtain substantial results as shown in Figure 4. The final results obtained are training loss = 0.1288, raining accuracy = 0.9531, validation loss: 0.1835, and validation accuracy of 0.8234.

7 ADVANTAGES:

Machine learning algorithms were from the very beginning designed and analyze medical datasets. Today, machine learning provides indispensable tool for data analysis. Especially in the last few years, the digital revolution provided relatively inexpensive and available means to collect and store data. Modern hospitals are well equipped with monitoring and other data collection devices, and data is gathered and shared in large data system. Machine Learning technology is currently compatible for analyzing medical data, and especially there's tons of labor wiped out diagnosis in small specialized diagnostic problems.

8 APPLICATIONS:

As the recall was increased, the precision decreased, and vice versa. In medical applications, all the patients who had the disease needed to be identified, and hence, the recall could be maximized. A low recall could be accepted if the cost of a follow-up medical examination was not high.

it can be seen that the proposed weighted classifier outperformed all the individual models. The generic image features, learned by the deep learning models from ImageNet, served as a good initialization of the weights. The misclassification error for normal (healthy) images as pneumonia images was greater compared to pneumonia images as healthy images. This might be because the number of chest X-ray images of the normal (healthy) case was significantly lower compared to the pneumonia cases.

9 CONCLUSION :
Model from scratch which consists of 5 layers and follows with a fully connected neural network. ThenThroughout the process of developing the CNN model for Pneumonia prediction, we have built a the trained model is evaluated using separate unseen data to avoid bias prediction. As the result, the accuracy of the test dataset reached 87.25% which indicates a decent model. This mini-project allows a beginner to obtain an overview of how to build a model to solve a real-world problem.

FUTURE SCOPE :
It is no doubt that the predictive model can be improved even better by performing data augmentation or implementing a transfer learning concept which facilitates the model a room for improvement. Therefore, this will be added as further enhancement in the upcoming stories.

Based on our obtained results the following are the future directions for continuing the research in Pneumonia Prediction using X-Ray.
1. The work specifically belongs to Pneumonia Prediction with assumption that the images are of good quality and therefore complex normalization, skew removal and slant removal operations are not performed. Accuracy improvement can be observed is these operations are also applied.
2. The developed model can be further extended for the end users .

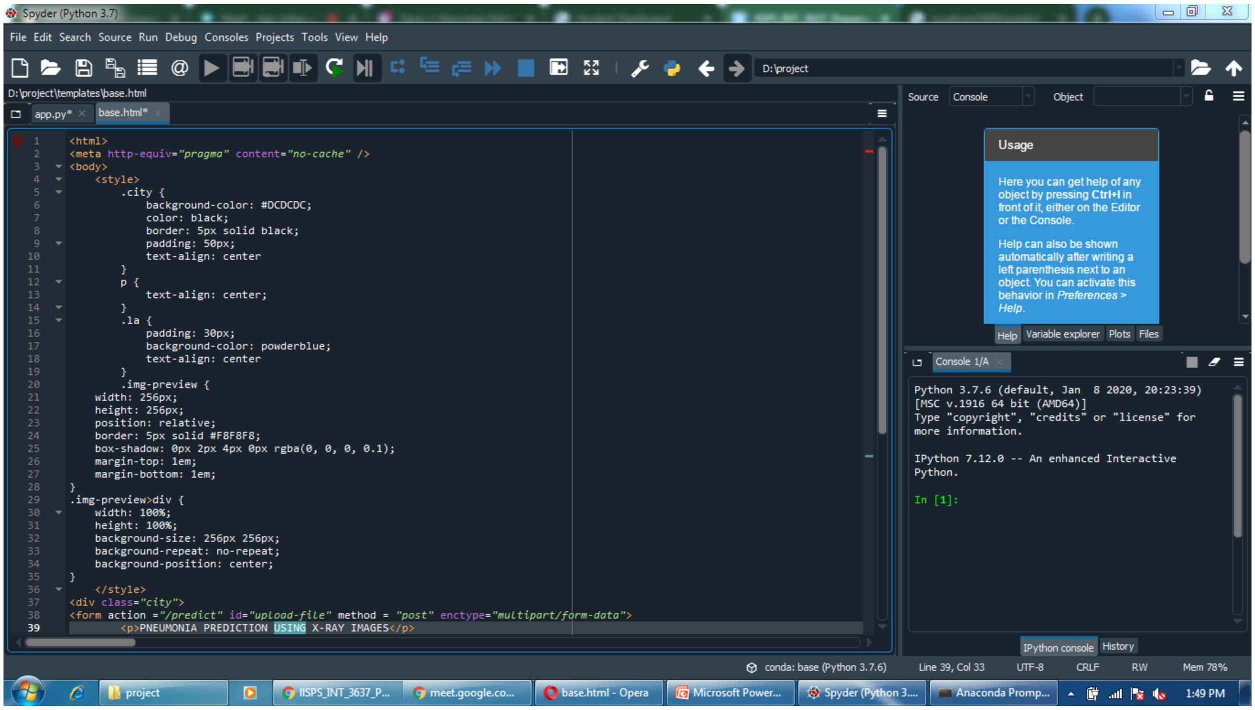
10 Bibliography:
1. Liu, N.; Wan, L.; Zhang, Y.; Zhou, T.; Huo, H.; Fang, T. Exploiting Convolutional Neural Networks With
Deeply Local Description for Remote Sensing Image Classification. IEEE Access 2018, 6, 11215–11228.
[CrossRef]
2. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.W.M.;
Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. Med. Image Anal. 2017, 42,
60–88. [CrossRef] [PubMed]
3. Brunetti, A.; Carnimeo, L.; Trotta, G.F.; Bevilacqua, V. Computer-assisted frameworks for classification of

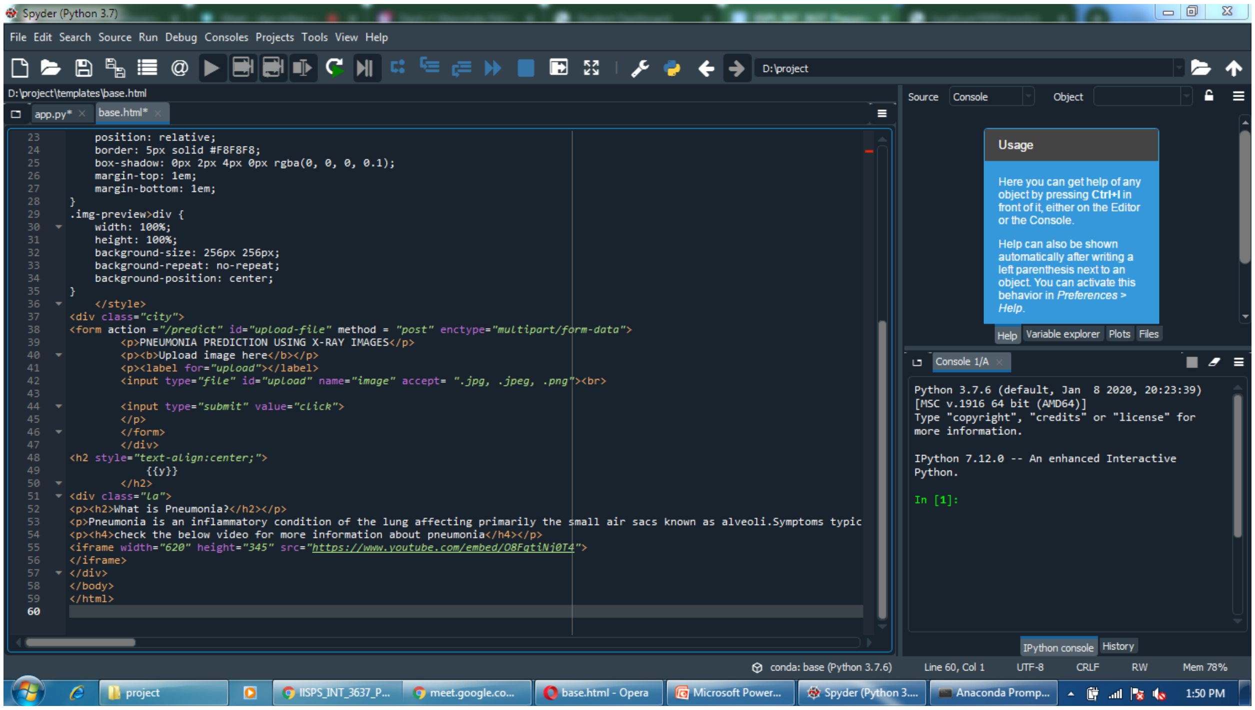
liver, breast and blood neoplasias via neural networks: A survey based on medical images. Neurocomputing
2019, 335, 274–298. [CrossRef]
4. Asiri, N.; Hussain, M.; Al Adel, F.; Alzaidi, N. Deep learning based computer-aided diagnosis systems for
diabetic retinopathy: A survey. Artif. Intell. Med. 2019, 99. [CrossRef] [PubMed]
5. Zhou, T.; Thung, K.; Zhu, X.; Shen, D. Eective feature learning and fusion of multimodality data using
stage-wise deep neural network for dementia diagnosis. Hum. Brain Mapp. 2018, 40, 1001–1016. [CrossRef]
6. Shickel, B.; Tighe, P.J.; Bihorac, A.; Rashidi, P. Deep EHR: A survey of recent advances in deep learning
techniques for electronic health record (EHR) analysis. IEEE J. Biomed. Health Inform. 2018, 22, 1589–1604.
[CrossRef]
7. Meyer, P.; Noblet, V.; Mazzara, C.; Lallement, A. Survey on deep learning for radiotherapy. Comput. Biol.
8. Mal¯ ukas, U.; Maskeli ¯ unas, R.; Damaševi˘cius, R.;Wo´zniak, M. Real time path finding for assisted living using
deep learning. J. Univers. Comput. Sci. 2018, 24, 475–487.
9. Zhang, X.; Yao, L.; Wang, X.; Monaghan, J.; McAlpine, D. A Survey on Deep Learning based Brain Computer
Interface: Recent Advances and New Frontiers. arXiv 2019, arXiv:1905.04149.
10. Bakator, M.; Radosav, D. Deep Learning and Medical Diagnosis: A Review of Literature. Multimodal Technol.
Interact. 2018, 2, 47. [CrossRef]
11. Gilani, Z.; Kwong, Y.D.; Levine, O.S.; Deloria-Knoll, M.; Scott, J.A.G.; O'Brien, K.L.; Feikin, D.R. A literature
review and survey of childhood pneumonia etiology studies: 2000–2010. Clin. Infect. Dis. 2012, 54 (Suppl.
2), S102–S108. [CrossRef] [PubMed]
12. Bouch, C.; Williams, G. Recently published papers: Pneumonia, hypothermia and the elderly. Crit. Care
2006, 10, 167. [CrossRef] [PubMed]
13. Scott, J.A.; Brooks,W.A.; Peiris, J.S.; Holtzman, D.; Mulholland, E.K. Pneumonia research to reduce childhood
mortality in the developing world. J. Clin. Investig. 2008, 118, 1291–1300. [CrossRef]

[PubMed]
14. Wunderink, R.G.; Waterer, G. Advances in the causes and management of community acquired pneumonia
in adults. BMJ 2017, 358, j2471. [CrossRef] [PubMed]
15. National Center for Health Statistics (NCHS); Centers for Disease Control and Prevention (CDC) FastStats:
Pneumonia. Last Updated February 2017. Available online:
http://www.cdc.gov/nchs/fastats/pneumonia.htm
(accessed on 21 November 2019).

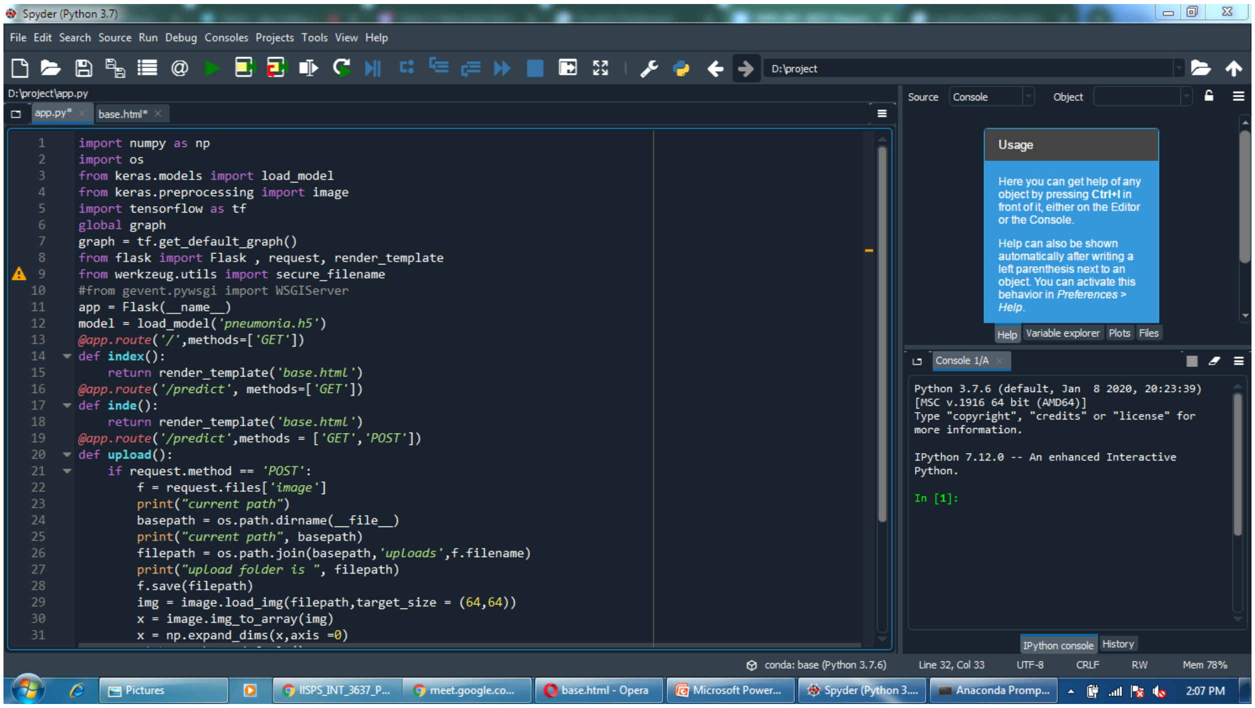
Appendix:

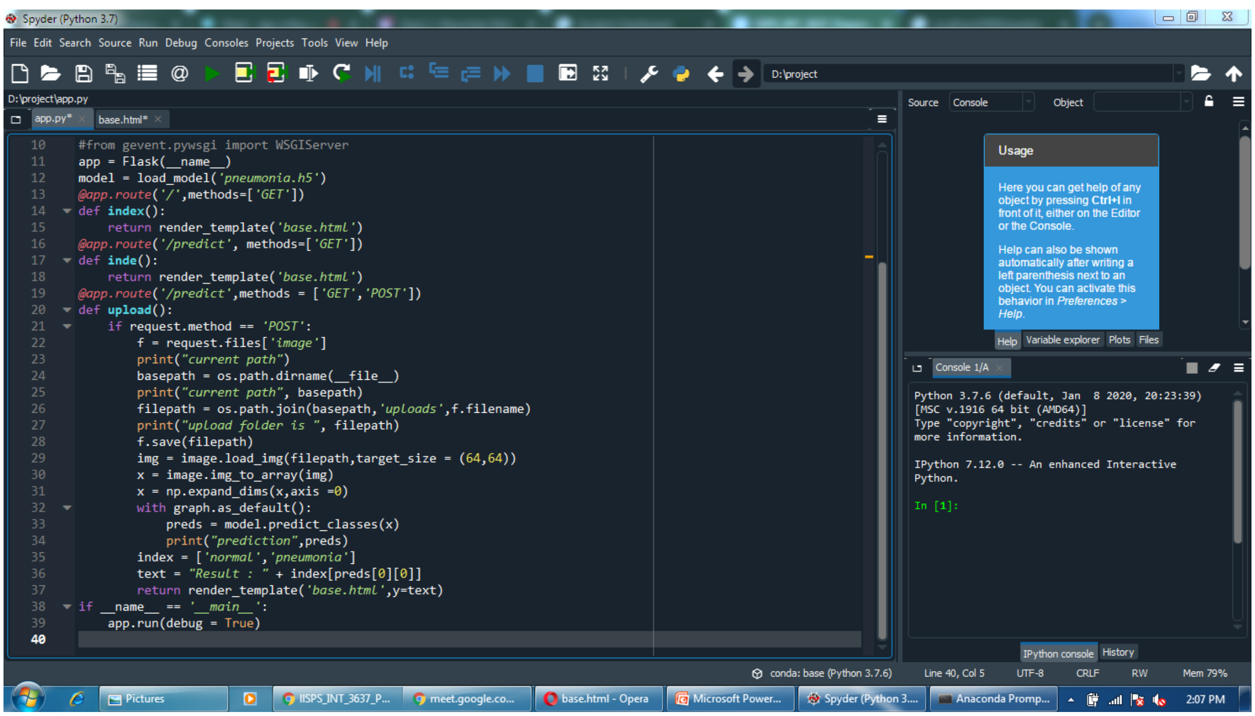
HTML:


Spyder (Python 3.7)
File Edit Search Source Run Debug Consoles Projects Tools View Help
D:\project
D:\project\templates\base.html
app.py*
base.html*
<html>
<meta http-equiv="pragma" content="no-cache" />
<body>
<style>
.city {
background-color: #DCDCDC;
color: black;
border: 5px solid black;
padding: 50px;
text-align: center
}
p {
text-align: center;
}
.la {
padding: 30px;
background-color: powderblue;
text-align: center
}
.img-preview {
width: 256px;
height: 256px;
position: relative;
border: 5px solid #F8F8F8;
box-shadow: 0px 2px 4px 0px rgba(0, 0, 0, 0.1);
margin-top: 1em;
margin-bottom: 1em;
}
.img-preview>div {
width: 100%;
height: 100%;
background-size: 256px 256px;
background-repeat: no-repeat;
background-position: center;
}
</style>
<div class="city">
<form action ="/predict" id="upload-file" method = "post" enctype="multipart/form-data">
<p>PNEUMONIA PREDICTION USING X-RAY IMAGES</p>
Source Console Object
Usage
Here you can get help of any object by pressing Ctrl+I in front of it, either on the Editor or the Console.
Help can also be shown automatically after writing a left parenthesis next to an object. You can activate this behavior in Preferences > Help.
Help Variable explorer Plots Files
Console 1/A
Python 3.7.6 (default, Jan 8 2020, 20:23:39)
[MSC v.1916 64 bit (AMD64)]
Type "copyright", "credits" or "license" for more information.
IPython 7.12.0 -- An enhanced Interactive Python.
In [1]:
IPython console History
conda: base (Python 3.7.6) Line 39, Col 33 UTF-8 CRLF RW Mem 78%
project IISPS_INT_3637_P... meet.google.co... base.html - Opera Microsoft Power... Spyder (Python 3.... Anaconda Promp...
1:49 PM


Spyder (Python 3.7)
File Edit Search Source Run Debug Consoles Projects Tools View Help
D:\project
D:\project\templates\base.html
app.py*
base.html*
position: relative;
border: 5px solid #F8F8F8;
box-shadow: 0px 2px 4px 0px rgba(0, 0, 0, 0.1);
margin-top: 1em;
margin-bottom: 1em;
}
.img-preview>div {
width: 100%;
height: 100%;
background-size: 256px 256px;
background-repeat: no-repeat;
background-position: center;
}
</style>
<div class="city">
<form action ="/predict" id="upload-file" method = "post" enctype="multipart/form-data">
<p>PNEUMONIA PREDICTION USING X-RAY IMAGES</p>
<p>Upload image here</p>
<p><label for="upload"></label>
<input type="file" id="upload" name="image" accept= ".jpg, .jpeg, .png">
<input type="submit" value="click">
</p>
</form>
</div>
<h2 style="text-align:center;">
{{y}}
</h2>
<div class="la">
<p><h2>What is Pneumonia?</h2></p>
<p>Pneumonia is an inflammatory condition of the lung affecting primarily the small air sacs known as alveoli.Symptoms typic
<p><h4>check the below video for more information about pneumonia</h4></p>
<iframe width="620" height="345" src="https://www.youtube.com/embed/O8FatiNj0T4">
</iframe>
</div>
</body>
</html>
Source Console Object
Usage
Here you can get help of any object by pressing Ctrl+I in front of it, either on the Editor or the Console.
Help can also be shown automatically after writing a left parenthesis next to an object. You can activate this behavior in Preferences > Help.
Help Variable explorer Plots Files
Console 1/A
Python 3.7.6 (default, Jan 8 2020, 20:23:39)
[MSC v.1916 64 bit (AMD64)]
Type "copyright", "credits" or "license" for more information.
IPython 7.12.0 -- An enhanced Interactive Python.
In [1]:
IPython console History
conda: base (Python 3.7.6) Line 60, Col 1 UTF-8 CRLF RW Mem 78%
project IISPS_INT_3637_P... meet.google.co... base.html - Opera Microsoft Power... Spyder (Python 3.... Anaconda Promp...
1:50 PM

PYTHON:


Spyder (Python 3.7)
File Edit Search Source Run Debug Consoles Projects Tools View Help
D:\project
D:\project\app.py
app.py*
base.html*
import numpy as np
import os
from keras.models import load_model
from keras.preprocessing import image
import tensorflow as tf
global graph
graph = tf.get_default_graph()
from flask import Flask , request, render_template
from werkzeug.utils import secure_filename
#from gevent.pywsgi import WSGIServer
app = Flask(__name__)
model = load_model('pneumonia.h5')
@app.route('/',methods=['GET'])
def index():
    return render_template('base.html')
@app.route('/predict', methods=['GET'])
def inde():
    return render_template('base.html')
@app.route('/predict',methods = ['GET','POST'])
def upload():
    if request.method == 'POST':
        f = request.files['image']
        print("current path")
        basepath = os.path.dirname(__file__)
        print("current path", basepath)
        filepath = os.path.join(basepath,'uploads',f.filename)
        print("upload folder is ", filepath)
        f.save(filepath)
        img = image.load_img(filepath,target_size = (64,64))
        x = image.img_to_array(img)
        x = np.expand_dims(x,axis =0)
Source Console Object
Usage
Here you can get help of any object by pressing Ctrl+I in front of it, either on the Editor or the Console.
Help can also be shown automatically after writing a left parenthesis next to an object. You can activate this behavior in Preferences > Help.
Help Variable explorer Plots Files
Console 1/A
Python 3.7.6 (default, Jan 8 2020, 20:23:39) [MSC v.1916 64 bit (AMD64)]
Type "copyright", "credits" or "license" for more information.
IPython 7.12.0 -- An enhanced Interactive Python.
In [1]:
IPython console History
conda: base (Python 3.7.6) Line 32, Col 33 UTF-8 CRLF RW Mem 78%
Pictures IISPS_INT_3637_P... meet.google.co... base.html - Opera Microsoft Power... Spyder (Python 3.... Anaconda Promp... 2:07 PM


Spyder (Python 3.7)
File Edit Search Source Run Debug Consoles Projects Tools View Help
D:\project
D:\project\app.py
app.py*
base.html*
#from gevent.pywsgi import WSGIServer
app = Flask(__name__)
model = load_model('pneumonia.h5')
@app.route('/',methods=['GET'])
def index():
    return render_template('base.html')
@app.route('/predict', methods=['GET'])
def inde():
    return render_template('base.html')
@app.route('/predict',methods = ['GET','POST'])
def upload():
    if request.method == 'POST':
        f = request.files['image']
        print("current path")
        basepath = os.path.dirname(__file__)
        print("current path", basepath)
        filepath = os.path.join(basepath,'uploads',f.filename)
        print("upload folder is ", filepath)
        f.save(filepath)
        img = image.load_img(filepath,target_size = (64,64))
        x = image.img_to_array(img)
        x = np.expand_dims(x,axis =0)
        with graph.as_default():
            preds = model.predict_classes(x)
            print("prediction",preds)
        index = ['normal','pneumonia']
        text = "Result : " + index[preds[0][0]]
        return render_template('base.html',y=text)
if __name__ == '__main__':
    app.run(debug = True)
Source Console Object
Usage
Here you can get help of any object by pressing Ctrl+I in front of it, either on the Editor or the Console.
Help can also be shown automatically after writing a left parenthesis next to an object. You can activate this behavior in Preferences > Help.
Help Variable explorer Plots Files
Console 1/A
Python 3.7.6 (default, Jan 8 2020, 20:23:39) [MSC v.1916 64 bit (AMD64)]
Type "copyright", "credits" or "license" for more information.
IPython 7.12.0 -- An enhanced Interactive Python.
In [1]:
IPython console History
conda: base (Python 3.7.6) Line 40, Col 5 UTF-8 CRLF RW Mem 79%
Pictures IISPS_INT_3637_P... meet.google.co... base.html - Opera Microsoft Power... Spyder (Python 3.... Anaconda Promp... 2:07 PM