Terna Engineering College

Department of Artificial Intelligence and Data Science

Program : Sem VI

Course: Machine Learning Lab

Experiment No.10

PART A

(PART A: TO BE REFERRED BY STUDENTS)

A.1 Aim: To build Pneumonia Prediction and Detection System using ML.

A.2 Theory:

Improve the accuracy and efficiency of pneumonia diagnosis, enabling timely intervention and treatment planning. Facilitate personalized healthcare by providing tailored risk assessments and diagnostic support. Contribute to the advancement of medical diagnostics through the integration of cutting-edge ML techniques with clinical practice. Ultimately, enhance patient outcomes, reduce healthcare costs, and alleviate the burden of pneumonia on individuals and healthcare systems.

This system aims to enhance diagnostic accuracy and patient outcomes by overcoming obstacles such as subjective interpretation, resource limitations, and insufficient data integration.
The goal is to simplify and streamline the pneumonia diagnosis process, ensuring timely and accurate results.

Algorithm:

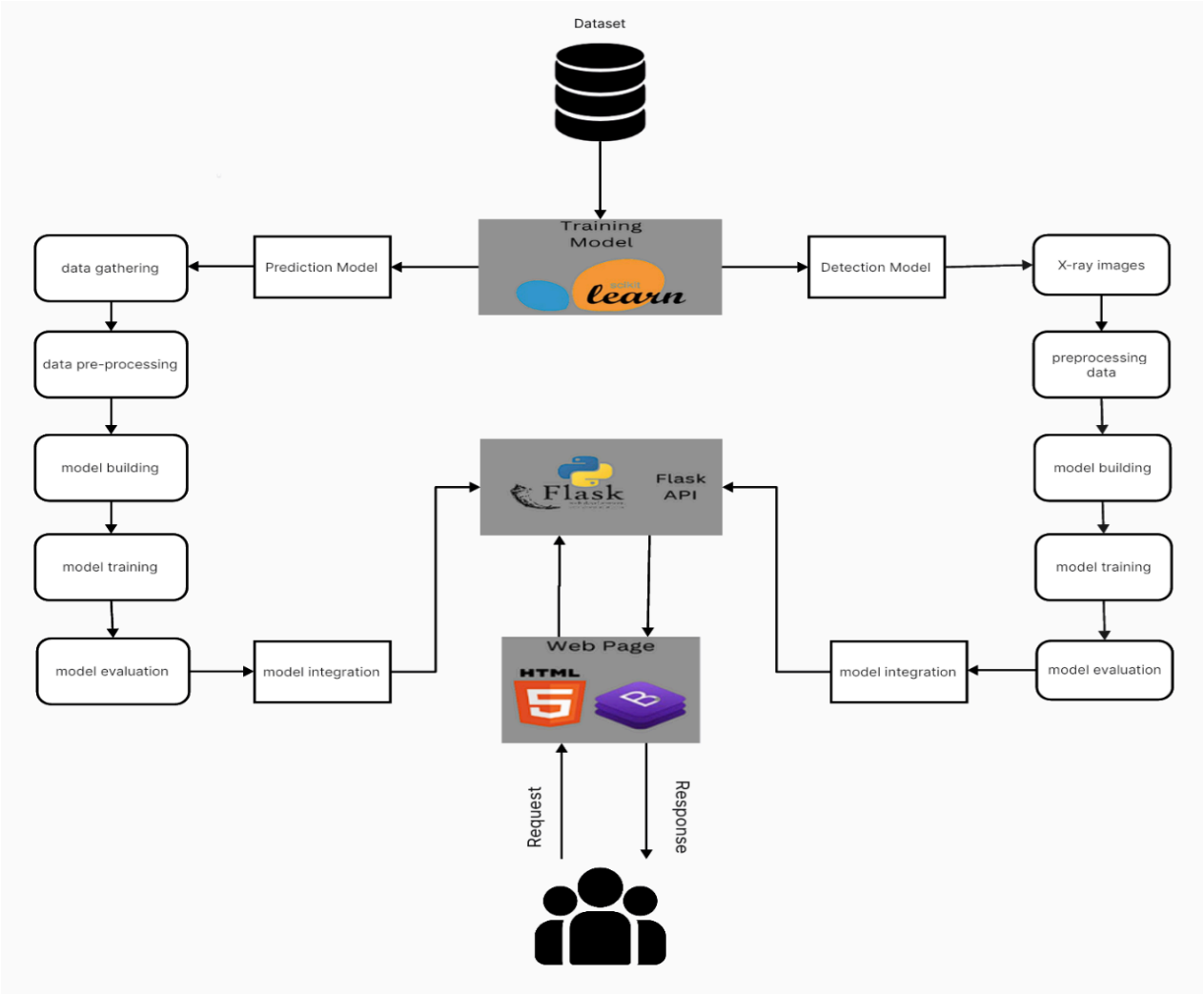
- **Data Collection:** Collect diverse patient data, including clinical results, symptoms, and labeled chest X-ray images.
- **Data Preprocessing:** Clean, preprocess, and augment data for standardization and diversity using Numpy and Pandas Library of the Python.
- **Feature Engineering:** Extract relevant features for predictive models and develop representations for chest X-ray images with the help of LDA, PCA and CNN.

- **Model Selection:** Choose suitable machine learning models for patient data and CNNs for image-based detection.
- **Model Training:** Train models, handling pneumonia case imbalances and optimizing hyperparameters we use LR,SVM, and Neural Networks.
- **Model Interpretability :** Incorporate interpretability techniques for transparency with healthcare professionals.
- **Image-based Detection:** Fine-tune CNNs for accurate pneumonia detection from chest X-ray images.
- **Ethical Considerations:** Ensure data privacy, anonymize information, and obtain necessary ethical approvals.
- **Collaboration with Healthcare Professionals:** Engage healthcare experts for feedback, ensuring clinical relevance.
- **Validation and Testing:** Rigorous validation using cross-validation, testing on separate datasets for generalization.

Literature Survey:

Sr. No	Paper Details	Problem addressed	Methodology	Advantages	Limits
1	Mohammad Farukh Hashmi:Pneumonia detection in chest X-ray images using compound scaled deep learning model| 2021, VOL. 62, NOS. 3–4, 397–406	Difficulty in accurately diagnosing pneumonia, especially in areas lacking trained radiologists.. Urgent requirement for computer-aided diagnosis (CAD) research to improve pneumonia detection. .	• Leveraging deep learning techniques for pneumonia diagnosis. • Creating CAD systems specifically tailored for pneumonia detection. • Analyzing chest X-ray datasets, training deep learning models to enhance predictive capabilities.	• Improved pneumonia diagnosis accuracy compared to traditional methods. • Making pneumonia diagnosis more accessible, especially in regions with limited access to medical professionals. • CAD systems designed to collaborate with clinicians, enhancing the diagnostic process	• Success relies on the quality and diversity of training datasets. • Addressing ethical concerns and regulatory standards in implementing AI-based systems in healthcare. • Real-world clinical validation necessary to ensure reliability and safety of CAD systems in diverse healthcare settings.

Flowchart:


Dataset
Training Model
learn
data gathering
Prediction Model
Detection Model
X-ray images
data pre-processing
preprocessing data
model building
Flask
Flask API
model building
model training
model training
model evaluation
model integration
Web Page
HTML
model integration
model evaluation
Request
Response

PART B

(PART B: TO BE COMPLETED BY STUDENTS)

(Students must submit the soft copy as per following segments within two hours of the practical. Thesoftcopy mustbeuploaded on the Blackboard or emailed to the concerned lab in charge faculties atthe endofthepractical incasethethere is noBlackboard access available)

Roll No.: A12	Name: Sufiyan Khan
Class: TE–AI&DS	Batch: A1
Date of Experiment:	Date of Submission:
Grade:	

B.1 Input and Output:

Code:

In [1]:
```
#import libraries
import pandas as pd
pd.options.mode.chained_assignment = None
import ast
%matplotlib inline
```

In [2]:
```
#import data
data_set=pd.read_csv("pneumonia_blood_test_data.csv")
data_set.head()
```

Out[2]:

	Patient_ID	Age	Gender	Symptoms	White_Blood_Cell_Count	CRP_Level	Procalcitonin_Level	ESR_Level	Diagnosis
0	1	48	Male	['Shortness of Breath', 'Chest Pain']	9261	6.10	0.34	23	No Pneumonia
1	2	21	Male	['Chills', 'Fatigue', 'Productive Cough']	10200	8.17	0.28	17	No Pneumonia
2	3	62	Female	['Fatigue', 'Cough']	14957	17.31	0.81	21	No Pneumonia
3	4	59	Female	['Chest Pain', 'Cough', 'Shortness of Breath',...	12014	7.86	0.36	29	Pneumonia
4	5	43	Female	['Chills']	13237	7.96	0.47	30	No Pneumonia

**Preprocessing data **Cleaning symptoms column

In [3]:
```
#dataset without symptoms and diagnosis
symptoms=data_set['Symptoms']
diagnosis=data_set['Diagnosis']
data_set.drop(labels='Symptoms',axis='columns',inplace=True)
```

In [4]:
```
#Adding cloumns for each symptoms and giving default value as 0
#Cough,Shortness of Breath,Chest Pain,Chills,Fatigue,Productive Cough
data_set['Cough']=0
data_set['Shortness of Breath']=0
data_set['Chest Pain']=0
data_set['Chills']=0
data_set['Fatigue']=0
data_set['Productive Cough']=0
data_set.head()
```

Out[4]:

	Patient_ID	Age	Gender	White_Blood_Cell_Count	CRP_Level	Procalcitonin_Level	ESR_Level	Diagnosis	Cough	Shortness of Breath	Chest Pain	C
0	1	48	Male	9261	6.10	0.34	23	No Pneumonia	0	0	0	
1	2	21	Male	10200	8.17	0.28	17	No Pneumonia	0	0	0	
2	3	62	Female	14957	17.31	0.81	21	No Pneumonia	0	0	0	
3	4	59	Female	12014	7.86	0.36	29	Pneumonia	0	0	0	
4	5	43	Female	13237	7.96	0.47	30	No Pneumonia	0	0	0	

In [5]:
```
#Adding Disgnosis columns having value as 0(No pneumonia) or 1(having pnuemonia)
def pneumonia_check(check):
    if 'Pneumonia' == check:
        return 1
    else:
        return 0

data_set['Diagnosis']=data_set['Diagnosis'].apply(pneumonia_check)
data_set.head()
```

Out[5]:

	Patient_ID	Age	Gender	White_Blood_Cell_Count	CRP_Level	Procalcitonin_Level	ESR_Level	Diagnosis	Cough	Shortness of Breath	Chest Pain	Ch
0	1	48	Male	9261	6.10	0.34	23	0	0	0	0	
1	2	21	Male	10200	8.17	0.28	17	0	0	0	0	
2	3	62	Female	14957	17.31	0.81	21	0	0	0	0	
3	4	59	Female	12014	7.86	0.36	29	1	0	0	0	
4	5	43	Female	13237	7.96	0.47	30	0	0	0	0	

In [6]:
```
new_symptoms=[]
for symptom in symptoms:
    string_representation=symptom
    new_symptoms.append(ast.literal_eval(string_representation))

#Cough,Shortness of Breath,Chest Pain,Chills,Fatigue,Productive Cough
for index in range(len(new_symptoms)):
    if 'Cough' in new_symptoms[index]:
        data_set['Cough'][index]=1
    if 'Shortness of Breath' in new_symptoms[index]:
        data_set['Shortness of Breath'][index]=1
    if 'Chest Pain' in new_symptoms[index]:
        data_set['Chest Pain'][index]=1
    if 'Fatigue' in new_symptoms[index]:
        data_set['Fatigue'][index]=1
    if 'Productive Cough' in new_symptoms[index]:
        data_set['Productive Cough'][index]=1
    if 'Chills' in new_symptoms[index]:
        data_set['Chills'][index]=1
```

In [7]:
```
## Final Cleaned dataset
data_set.head()
```

Out[7]:

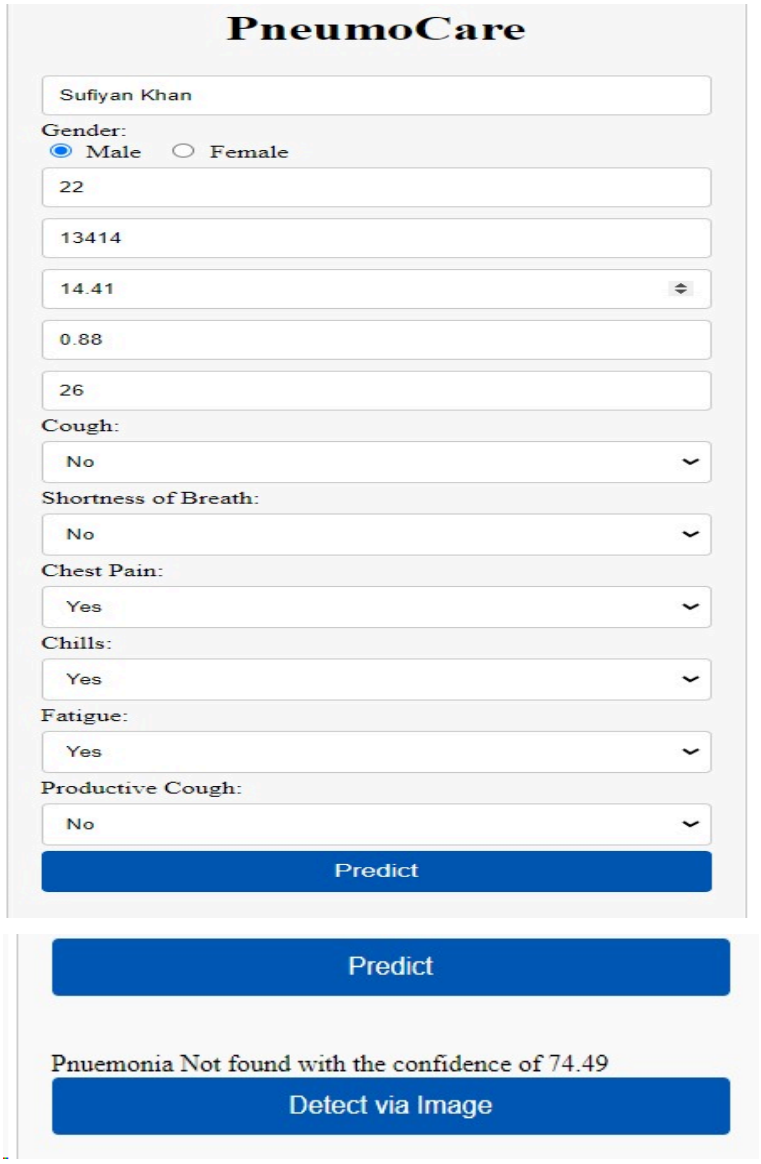
	Patient_ID	Age	Gender	White_Blood_Cell_Count	CRP_Level	Procalcitonin_Level	ESR_Level	Diagnosis	Cough	Shortness of Breath	Chest Pain	Ch
0	1	48	Male	9261	6.10	0.34	23	0	0	1	1	
1	2	21	Male	10200	8.17	0.28	17	0	0	0	0	
2	3	62	Female	14957	17.31	0.81	21	0	1	0	0	
3	4	59	Female	12014	7.86	0.36	29	1	1	1	1	

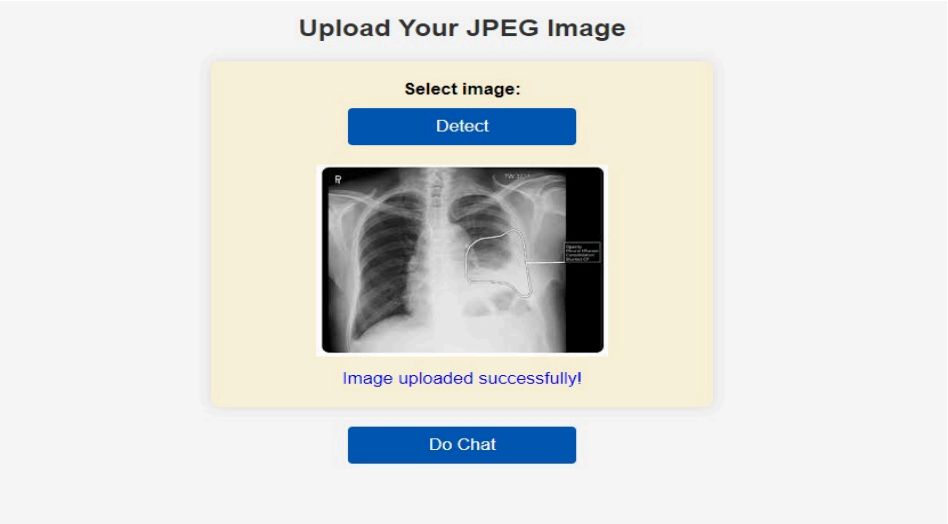
In [8]:
```
data_set.info()
```

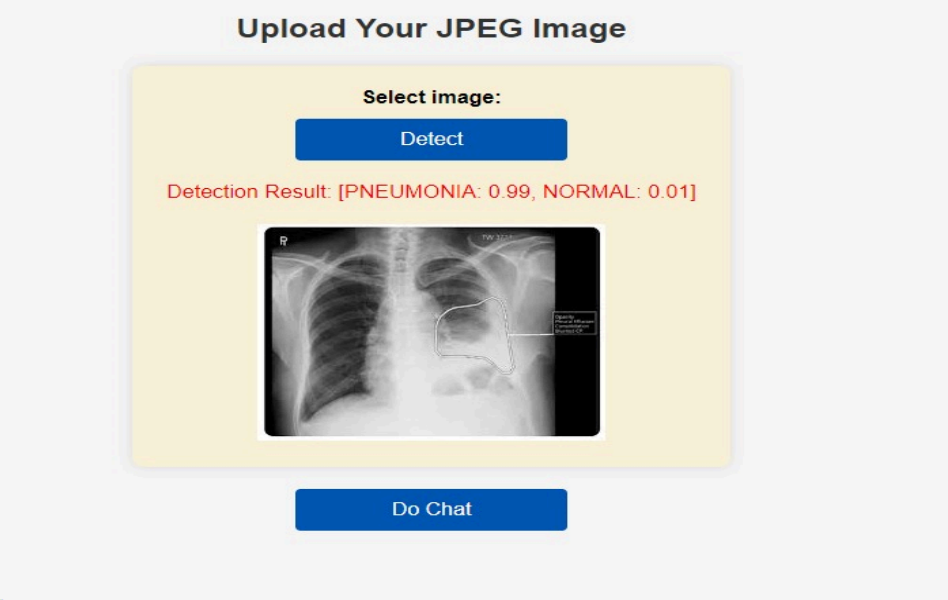
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5467 entries, 0 to 5466
Data columns (total 14 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   Patient_ID              5467 non-null   int64
 1   Age                     5467 non-null   int64
 2   Gender                  5467 non-null   object
 3   White_Blood_Cell_Count  5467 non-null   int64
 4   CRP_Level               5467 non-null   float64
 5   Procalcitonin_Level     5467 non-null   float64
 6   ESR_Level               5467 non-null   int64
 7   Diagnosis               5467 non-null   int64
 8   Cough                   5467 non-null   int64
 9   Shortness of Breath     5467 non-null   int64
 10  Chest Pain              5467 non-null   int64
 11  Chills                  5467 non-null   int64
 12  Fatigue                 5467 non-null   int64
 13  Productive Cough        5467 non-null   int64
dtypes: float64(2), int64(11), object(1)
memory usage: 598.1+ KB
```

In [37]:
```
data_set.to_csv('cleaned_pneumonia_blood_dataset.csv')
```

Output:


PneumoCare
Sufiyan Khan
Gender:
Male
Female
22
13414
14.41
0.88
26
Cough:
No
Shortness of Breath:
No
Chest Pain:
Yes
Chills:
Yes
Fatigue:
Yes
Productive Cough:
No
Predict
Predict
Pnuemonia Not found with the confidence of 74.49
Detect via Image





B.2 Conclusion:

In conclusion, our project represents a significant step forward in the realm of pneumonia prediction and detection using Machine Learning (ML). we have developed a comprehensive system capable of assessing pneumonia risk and detecting the presence of the condition. our project represents a contribution to the field of medical diagnostics, demonstrating the potential of ML-driven approaches to revolutionize pneumonia management and enhance patient care.