VISVESVARAYA TECHNOLOGICAL UNIVERSITY

Jnana Sangama, Belagavi-590018

An Internship(21INT68) Report on

"Chronic Disease Prediction using Python"

Submitted in partial fulfillment of the requirements for the award of degree

of

BACHELOR OF ENGINEERING
IN
COMPUTER SCIENCE AND ENGINEERING

By

Rupam Bhattacharyya	**1EP21CS089**

Internal Guide: **Prof. Pavan Kumar R B** **Assistant Professor,** **Dept. of CSE, EPCET**	**Company Guide:** **Shounak Mohanta** **Software Engineer,** **Varcons Technologies**

EAST POINT COLLEGE OF ENGINEERING & TECHNOLOGY

Department of Computer Science & Engineering

Approved by AICTE New Delhi | Affiliated to VTU, Belagavi,
Virgo Nagar, Bengaluru - 560049

2023-2024

EAST POINT COLLEGE OF ENGINEERING & TECHNOLOGY

Department of Computer Science & Engineering

Approved by AICTE New Delhi | Affiliated to VTU, Belagavi,
Virgo Nagar, Bengaluru - 560049

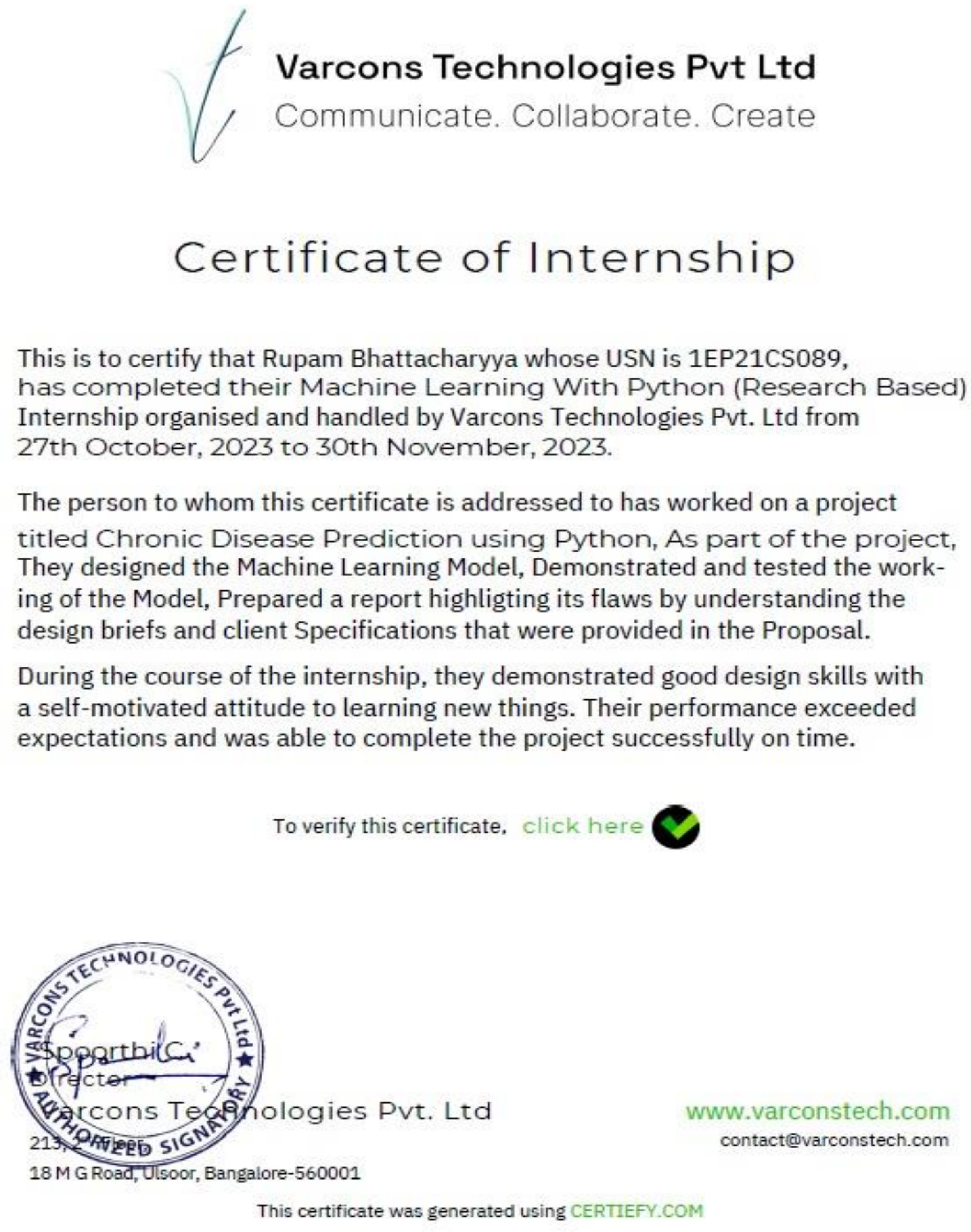
CERTIFICATE

This is to certify that the Internship (21INT68) entitled "Chronic Disease Prediction using Python" is a bonafide work carried out by **Rupam Bhattacharyya[1EP21CS089]**, in partial fulfillment of the requirements of BACHELOR OF ENGINEERING in COMPUTER SCIENCE AND ENGINEERING in VISVESVARAYA TECHNOLOGICAL UNIVERSITY, Belgaum, during the year 2022-2023. It is certified that corrections/suggestions recommended have been incorporated in the Internship report.

Internal Guide	**Company Guide**	**Signature of HOD**	**Signature of Principal**
Prof. Pavan Kumar R B **Assistant Professor,** **Dept. of CSE, EPCET,** **Bangalore.**	**Shounak Mohanta** **Software Engineer,** **Varcons Technologies**	**Dr. I Manimozhi** **HOD, Dept. of CSE,** **EPCET, Bangalore**	**Dr. Mrityunjaya V Latte** **Principal,** **EPCET, Bangalore**

Reviewers:

Reviewer	**Panel Member**
Name: Prof. Shammi L	**Name: Dr. I Manimozhi**
Signature with date:	**Signature with date:**



Varcons Technologies Pvt Ltd
Communicate. Collaborate. Create

Certificate of Internship

This is to certify that Rupam Bhattacharyya whose USN is 1EP21CS089, has completed their Machine Learning With Python (Research Based) Internship organised and handled by Varcons Technologies Pvt. Ltd from 27th October, 2023 to 30th November, 2023.

The person to whom this certificate is addressed to has worked on a project titled Chronic Disease Prediction using Python, As part of the project, They designed the Machine Learning Model, Demonstrated and tested the working of the Model, Prepared a report highligting its flaws by understanding the design briefs and client Specifications that were provided in the Proposal.

During the course of the internship, they demonstrated good design skills with a self-motivated attitude to learning new things. Their performance exceeded expectations and was able to complete the project successfully on time.

To verify this certificate, click here ✓

VARCONS TECHNOLOGIES Pvt Ltd ★ AUTHORIZED SIGNATORY ★

Spoorthi C
Director
Varcons Technologies Pvt. Ltd
213, 2nd Floor
18 M G Road, Ulsoor, Bangalore-560001

www.varconstech.com
contact@varconstech.com

This certificate was generated using CERTIEFY.COM

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere regards and thanks to **Management** of **East Point Group of Institutions, Bengaluru** for providing me an opportunity to work on this Internship.

I am ineffably indebted to **Dr. S Prakash, Senior Vice President**, **East Point Group of Institutions,** for his conscientious guidance and encouragement to accomplish this Internship.

I would like to express my humble and sincere thanks to **Dr. Mrityunjaya V Latte, Principal, East Point College of Engineering and Technology** for his suggestions that motivated me for the successful completion of my Internship.

I would like to express my heartfelt thanks to **Dr. I Manimozhi**, **Professor and Head of Department** of Computer Science and Engineering, EPCET for her valuable advice and suggestions to do my best in this Internship.

I am obliged to guide **Prof Pavan Kumar R B**, Assistant Professor, Dept. of CSE, Internship Coordinators **Prof. Madhushree**, Assistant Professor, Dept. of CSE and **Prof. Nithyananda C R,** Associate Professor, Dept. of CSE who have rendered valuable assistance and guidance for the Internship.

I would like to thank **Mr. Shounak Mohanta**, Software Engineer, Varcons Technologies for providing with technical support and knowledge.

I would like to thank my **Parents** and **Friends** for their support and encouragement duringthe course of my Internship. Finally, I offer my regards to all the **Faculty members of theCSE Department** and all those who supported me in any respect during the Internship.

Rupam Bhattacharyya [1EP21CS089]

CONTENTS

List of Figures

Chapter 1

COMPANY PROFILE

Varcons Technologies, established with a visionary goal of "Empowering Businesses through Innovative Technological Solutions," has quickly positioned itself as a leading player in the ever-evolving tech landscape. The company recognizes that each business is unique, with distinct models and technological requirements, and thus, it places a strong emphasis on tailoring solutions to meet the specific needs of its clients.

Specializing in a wide range of technology domains, Varcons Technologies is a comprehensive Technology Organization offering expertise in web design and development, database management (MYSQL), programming languages such as PYTHON, and front-end technologies like HTML and CSS. The company also excels in ASP.NET and LINQ, showcasing a versatile skill set to address diverse client needs.

Varcons Technologies is at the forefront of addressing the increasing demand for automation, with a keen focus on ERP solutions, connectivity, SEO services, and tailor-made software products. The company is committed to delivering effective web promotion strategies and has a notable presence in Conference Management solutions.

In its pursuit of excellence, Varcons Technologies aims to stand out as an innovative software development company based in Bangalore, India. Leveraging well-researched expertise, the company is dedicated to bringing creativity and innovation to the forefront of software development.

Understanding that effective collaboration and understanding client demands are crucial for success, Varcons Technologies actively engages with clients to define precise solution requirements. The company's collaborative approach often leads to the redefinition of solutions or the emergence of entirely new application requirements during brainstorming sessions. Varcons Technologies positions itself not just as a service provider but as an IT solutions consulting group, comprising high-caliber consultants dedicated to finding optimal solutions for clients.

Chapter 2

ABOUT COMPANY

Varcons Technologies is a dynamic and innovative technology organization dedicated to providing a diverse range of solutions in the ever-evolving landscape of IT services. Specializing in cutting-edge technologies, Varcons Technologies offers expertise in various domains, including but not limited to Web Design and Development, Database Management with a focus on MYSQL, and proficiency in Python Programming.

Expertise of Varcons Technologies:

Web Design and Development:

At Varcons Technologies, we take pride in our comprehensive approach to web design and development. Our skilled professionals are adept at creating seamless client-server applications, ensuring an immersive user experience. Whether it's crafting engaging web interfaces, developing robust backend systems, or optimizing websites for performance, our team is committed to delivering excellence.

Mobile App Development:

Embracing the fast-paced world of mobile technologies, Varcons Technologies excels in Android app development. Leveraging the Android SDK, Java, and Kotlin, our team creates innovative and user-friendly applications tailored to meet the specific needs of our clients. From concept to deployment, we ensure a smooth and efficient development process.

Database Management - MYSQL:

Our proficiency in MYSQL allows us to design and implement robust database solutions. We understand the critical role databases play in applications, and our experts ensure optimal performance, security, and scalability for your data.

Python Programming:

Varcons Technologies is at the forefront of Python programming, utilizing its versatility for a wide range of applications. Whether it's web development, automation, or data analysis, our Python experts deliver solutions that align with industry best practices.

ERP Solutions:

Understanding the growing demand for automation, we specialize in ERP solutions that streamline business processes. Our ERP services encompass a wide range of functionalities, including connectivity, SEO services, and conference management, ensuring a holistic approach to organizational efficiency.

Web Promotion and SEO Services: Varcons Technologies recognizes the significance of effective web promotion and search engine optimization. Our dedicated team works towards enhancing the online visibility of your business, driving organic traffic, and ensuring a strong digital presence.

Tailor-made Software Products:

We take pride in our ability to conceptualize and develop tailor-made software products that align with the unique requirements of our clients. Our focus on customization ensures that the solutions we provide are not only efficient but also perfectly aligned with the business goals of our clients.

Training and Skill Development:

Varcons Technologies places a strong emphasis on mentorship and skill development. Our team of skilled mentors is committed to guiding and nurturing talent, ensuring that each trainee is well-equipped with the latest industry-relevant skills. We organize skill development programs to continuously enhance the capabilities of our mentors, creating a learning ecosystem that benefits both the team and our clients.

Services Offered:

Core Java and Advanced Java

Web Services and Development

Dot Net Framework

Python

Selenium Testing

Conference/Event Management Service

Academic Project Guidance

On The Job Training

Software Training

Chapter 3

INTRODUCTION

3.1 Introduction to ML

Machine Learning is a transformative field within artificial intelligence that empowers computers to learn and adapt without explicit programming. It involves the development of algorithms and models that enable systems to analyze data, recognize patterns, and make intelligent decisions autonomously. From predicting outcomes to optimizing processes, machine learning has diverse applications, revolutionizing industries such as healthcare, finance, and technology. As a dynamic and evolving discipline, machine learning continues to drive innovation, offering unprecedented capabilities for data-driven insights and problem-solving.

The core idea behind machine learning is to enable computers to recognize patterns, correlations, and trends within datasets, ultimately allowing them to make predictions or decisions based on this acquired knowledge. There are three main types of machine learning:

3.1.1. Supervised Learning

In this approach, the algorithm is trained on a labeled dataset, where the input data is paired with corresponding output labels. The model learns to map inputs to outputs, making predictions on new, unseen data. Common applications include image recognition, speech recognition, and predictive modeling.

3.1.2. Unsupervised Learning

Unlike supervised learning, unsupervised learning deals with unlabeled data. The algorithm explores the data's inherent structure and relationships, identifying patterns and grouping similar data points. Clustering and dimensionality reduction are common tasks within unsupervised learning, applicable in fields like customer segmentation and anomaly detection.

3.1.3. Reinforcement Learning

This learning paradigm involves an agent learning to make decisions by interacting with an environment. The agent receives feedback in the form of rewards or penalties based on its actions, allowing it to learn optimal strategies over time. Reinforcement learning is widely used in areas such as game playing, robotics, and autonomous systems.

Machine learning algorithms leverage various techniques, including linear regression, decision trees, neural networks, and support vector machines, among others. Additionally, deep learning, a subset of machine learning, employs neural networks with multiple layers to learn intricate representations of data, achieving remarkable success in tasks like image and speech recognition.

The applications of machine learning are vast and diverse, spanning industries such as healthcare, finance, marketing, and beyond. In healthcare, ML is utilized for disease prediction and personalized treatment plans. In finance, it aids in fraud detection and stock market predictions. In marketing, ML algorithms optimize advertising strategies and customer segmentation.

Continuous advancements in machine learning, driven by research and technological innovation, contribute to its growing impact on society. Ethical considerations, transparency, and accountability in ML algorithms are areas of active research as the technology becomes increasingly integrated into various aspects of our lives.

3.2 Problem Statement

Chronic Disease Prediction using ML Algorithms is a novel application of machine learning that aims to bridge the gap between patient symptoms and accurate disease diagnosis. This project explores the development of a system capable of predicting chronic diseases based on various symptoms and patient data, leveraging advancements in machine learning and data analysis technologies.

3.3 Solution

The following Python script is designed to train and evaluate machine learning models for disease prediction based on a dataset containing 132 symptoms. The script utilizes several machine learning algorithms and includes functionalities for data preprocessing, training, evaluation, and making predictions. Key dependencies include pandas for data manipulation, scikit-learn for machine learning, seaborn and matplotlib for data visualization, and joblib for model serialization.

3.3.1 Code Structure

1. **Import Dependencies**:
   - Various libraries are imported, including yaml for configuration file parsing, pandas for handling data, sklearn for machine learning models and evaluation, and matplotlib and seaborn for visualizations.
2. **Class Definition: DiseasePrediction**:
   - **Initialization (__init__)**:
     - The configuration file (config.yaml) is loaded to get parameters for data paths, model settings, etc.
     - Training and test datasets are loaded using private methods _load_train_dataset and _load_test_dataset.
     - A correlation heatmap of the features in the training data is optionally generated for exploratory data analysis.
   - **Private Methods**:
     - _load_train_dataset: Loads and preprocesses the training dataset, ensuring that it contains the expected number of features and labels.
     - _load_test_dataset: Loads and preprocesses the test dataset, aligning it with the training data features.
     - _feature_correlation: Generates a correlation matrix heatmap to visualize the relationships between features.
   - **Training the Model**:
     - train_model: This method splits the training data into training and validation sets, selects the model based on the specified name (e.g., Naive Bayes, Decision Tree, Random Forest, Gradient Boost), trains the model, and evaluates its performance on the validation set. The trained model is saved using joblib.
   - **Making Predictions**:
     - make_prediction: This method loads a saved model and uses it to make predictions on the test dataset. It evaluates the model’s performance by calculating accuracy and generating a classification report.

3.3.2 Execution

- **Main Script**:
    - The script specifies the model to be trained (e.g., 'decision_tree').
    - An instance of the DiseasePrediction class is created with the specified model name.
    - The model is trained using the train_model method.
    - The trained model's performance is evaluated on the test data using the make_prediction method, and the results are printed.

3.4 Naive Bayes Algorithm

The Naive Bayes algorithm is a probabilistic classifier based on Bayes' Theorem with the "naive" assumption of independence between every pair of features. Despite this assumption being unrealistic in many real-world situations, the algorithm performs remarkably well in various applications, particularly in text classification, spam filtering, sentiment analysis, and disease prediction.

Bayes' Theorem provides a way to update the probability estimate for a hypothesis as more evidence or information becomes available. Mathematically, Bayes' Theorem is expressed as:

$P(H|E) = P(E|H) \cdot P(H) / P(E)$

Where:

- $P(H|E)$ is the posterior probability of hypothesis H given the evidence E.
- $P(E|H)$ is the likelihood of evidence E given that hypothesis H is true.
- $P(H)$ is the prior probability of hypothesis H.
- $P(E)$ is the marginal likelihood of evidence E.

There are three main types of Naive Bayes classifiers:

1. **Gaussian Naive Bayes**: Assumes that the features follow a normal distribution. It is suitable for continuous data.

2. **Multinomial Naive Bayes**: Used for discrete data, particularly in text classification problems where the features represent the frequency of words.
3. **Bernoulli Naive Bayes**: Similar to the Multinomial Naive Bayes but designed for binary/Boolean features. Each feature is considered as either present or absent.

3.4.1 Steps to apply

1. **Training Phase**:
   - Calculate the prior probability for each class.
   - For each class, calculate the likelihood (probability) of each feature given the class. For continuous data, this often involves calculating the mean and variance of the feature within each class.
2. **Prediction Phase**:
   - For a new instance, calculate the posterior probability for each class using Bayes' Theorem.
   - Assign the class with the highest posterior probability to the instance.

3.5 Implementation

Implementation is the stage where the theoretical design is turned into a working system. The most crucial stage in achieving a new successful system and in giving confidence on the new system for the users that it will work efficiently and effectively.

The system can be implemented only after thorough testing is done and if it is found to work according to the specification. It involves careful planning, investigation of the current system and its constraints on implementation, design of methods to achieve the change over and an evaluation of change over methods a part from planning. Two major tasks of preparing the implementation are education and training of the users andtesting of the system. The more complex the system being implemented, the more involved will be the system analysis and design effort required just for implementation.

The implementation phase comprises of several activities. The required hardware and software acquisition is carried out. The system may require some software to be developed.

For this, programs are written and tested. The user then changes over to his new fully tested system and the old system is discontinued.

3.5.1 Testing

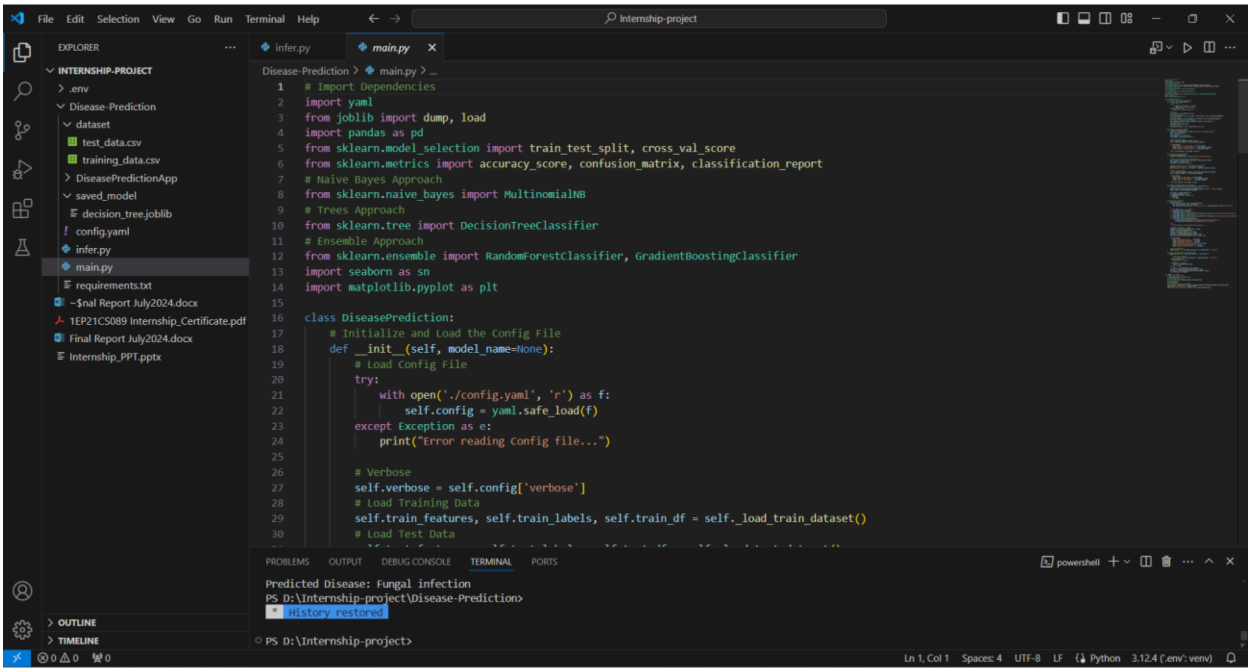
The testing phase is an important part of software development. It is the Information zed system will help in automate process of finding errors and missing operations and also a complete verification to determine whether the objectives are met and the user requirements are satisfied. Software testing is carried out in three steps:

1. The first includes unit testing, where in each module is tested to provide its correctness, validity and also determine any missing operations and to verify whether theobjectives have been met. Errors are noted down and corrected immediately.

   Unit testing is the important and major part of the project. So errors are rectified easily in particular module and program clarity is increased. In this project entire system is divided into several modules and is developed individually. So unit testing is conducted to individual modules.

2. The second step includes Integration testing. It need not be the case, the software whose modules when run individually and showing perfect results, will also show perfect results when run as a whole.

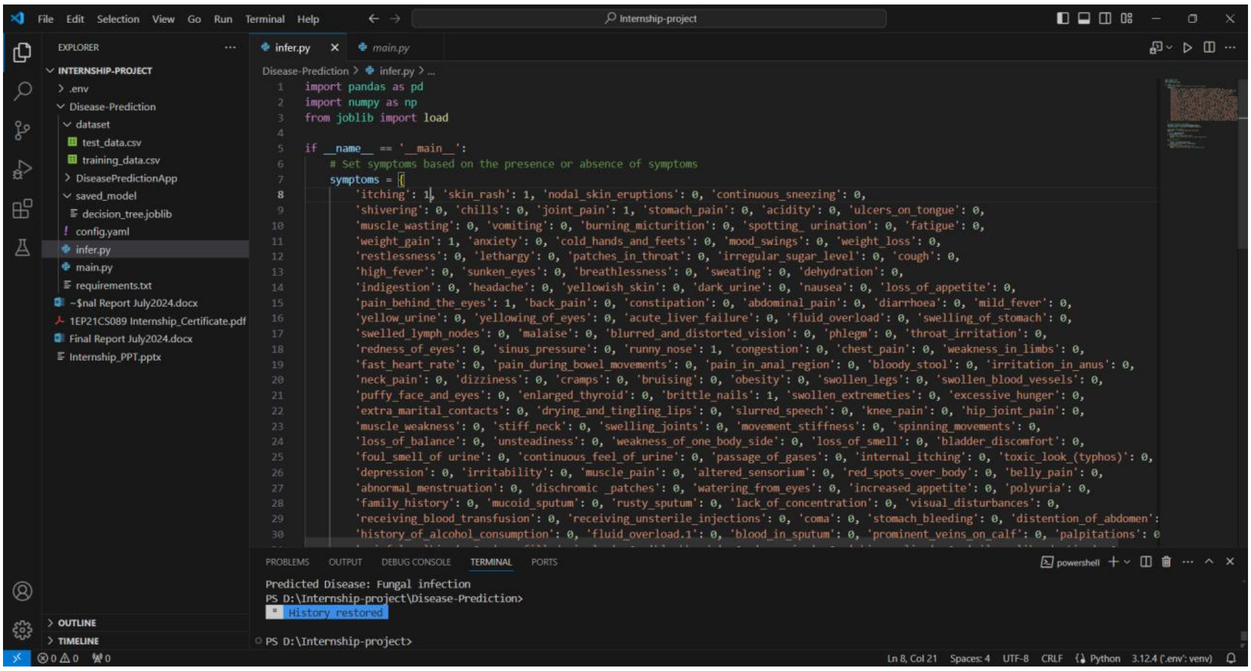
3.5.2 Sample Code


File Edit Selection View Go Run Terminal Help
Internship-project
EXPLORER
INTERNSHIP-PROJECT
.env
Disease-Prediction
dataset
test_data.csv
training_data.csv
DiseasePredictionApp
saved_model
decision_tree.joblib
config.yaml
infer.py
main.py
requirements.txt
~$nal Report July2024.docx
1EP21CS089 Internship_Certificate.pdf
Final Report July2024.docx
Internship_PPT.pptx
OUTLINE
TIMELINE
infer.py
main.py
Disease-Prediction > main.py > ...
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
powershell
Ln 1, Col 1 Spaces: 4 UTF-8 LF Python 3.12.4 ('.env': venv)

```python
# Import Dependencies
import yaml
from joblib import dump, load
import pandas as pd
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.metrics import accuracy_score, confusion_matrix, classification_report
# Naive Bayes Approach
from sklearn.naive_bayes import MultinomialNB
# Trees Approach
from sklearn.tree import DecisionTreeClassifier
# Ensemble Approach
from sklearn.ensemble import RandomForestClassifier, GradientBoostingClassifier
import seaborn as sn
import matplotlib.pyplot as plt

class DiseasePrediction:
    # Initialize and Load the Config File
    def __init__(self, model_name=None):
        # Load Config File
        try:
            with open('./config.yaml', 'r') as f:
                self.config = yaml.safe_load(f)
        except Exception as e:
            print("Error reading Config file...")

        # Verbose
        self.verbose = self.config['verbose']
        # Load Training Data
        self.train_features, self.train_labels, self.train_df = self._load_train_dataset()
        # Load Test Data
```

```
Predicted Disease: Fungal infection
PS D:\Internship-project\Disease-Prediction>
 *  History restored

PS D:\Internship-project>
```

Figure 3.6 Creation of Model


File Edit Selection View Go Run Terminal Help
Internship-project
EXPLORER
INTERNSHIP-PROJECT
.env
Disease-Prediction
dataset
test_data.csv
training_data.csv
DiseasePredictionApp
saved_model
decision_tree.joblib
config.yaml
infer.py
main.py
requirements.txt
~$nal Report July2024.docx
1EP21CS089 Internship_Certificate.pdf
Final Report July2024.docx
Internship_PPT.pptx
OUTLINE
TIMELINE
infer.py
main.py
Disease-Prediction > infer.py > ...
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
powershell
Ln 8, Col 21 Spaces: 4 UTF-8 CRLF Python 3.12.4 ('.env': venv)

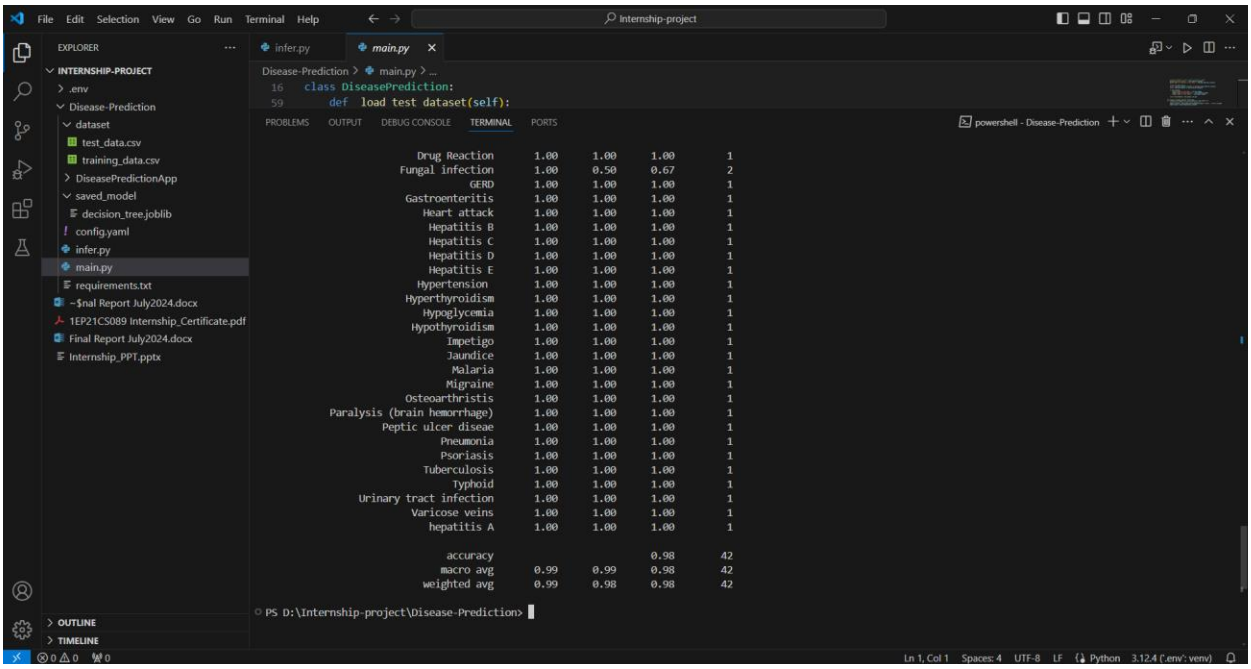
```python
import pandas as pd
import numpy as np
from joblib import load

if __name__ == '__main__':
    # Set symptoms based on the presence or absence of symptoms
    symptoms = {
        'itching': 1, 'skin_rash': 1, 'nodal_skin_eruptions': 0, 'continuous_sneezing': 0,
        'shivering': 0, 'chills': 0, 'joint_pain': 1, 'stomach_pain': 0, 'acidity': 0, 'ulcers_on_tongue': 0,
        'muscle_wasting': 0, 'vomiting': 0, 'burning_micturition': 0, 'spotting_ urination': 0, 'fatigue': 0,
        'weight_gain': 1, 'anxiety': 0, 'cold_hands_and_feets': 0, 'mood_swings': 0, 'weight_loss': 0,
        'restlessness': 0, 'lethargy': 0, 'patches_in_throat': 0, 'irregular_sugar_level': 0, 'cough': 0,
        'high_fever': 0, 'sunken_eyes': 0, 'breathlessness': 0, 'sweating': 0, 'dehydration': 0,
        'indigestion': 0, 'headache': 0, 'yellowish_skin': 0, 'dark_urine': 0, 'nausea': 0, 'loss_of_appetite': 0,
        'pain_behind_the_eyes': 1, 'back_pain': 0, 'constipation': 0, 'abdominal_pain': 0, 'diarrhoea': 0, 'mild_fever': 0,
        'yellow_urine': 0, 'yellowing_of_eyes': 0, 'acute_liver_failure': 0, 'fluid_overload': 0, 'swelling_of_stomach': 0,
        'swelled_lymph_nodes': 0, 'malaise': 0, 'blurred_and_distorted_vision': 0, 'phlegm': 0, 'throat_irritation': 0,
        'redness_of_eyes': 0, 'sinus_pressure': 0, 'runny_nose': 1, 'congestion': 0, 'chest_pain': 0, 'weakness_in_limbs': 0,
        'fast_heart_rate': 0, 'pain_during_bowel_movements': 0, 'pain_in_anal_region': 0, 'bloody_stool': 0, 'irritation_in_anus': 0,
        'neck_pain': 0, 'dizziness': 0, 'cramps': 0, 'bruising': 0, 'obesity': 0, 'swollen_legs': 0, 'swollen_blood_vessels': 0,
        'puffy_face_and_eyes': 0, 'enlarged_thyroid': 0, 'brittle_nails': 1, 'swollen_extremeties': 0, 'excessive_hunger': 0,
        'extra_marital_contacts': 0, 'drying_and_tingling_lips': 0, 'slurred_speech': 0, 'knee_pain': 0, 'hip_joint_pain': 0,
        'muscle_weakness': 0, 'stiff_neck': 0, 'swelling_joints': 0, 'movement_stiffness': 0, 'spinning_movements': 0,
        'loss_of_balance': 0, 'unsteadiness': 0, 'weakness_of_one_body_side': 0, 'loss_of_smell': 0, 'bladder_discomfort': 0,
        'foul_smell_of urine': 0, 'continuous_feel_of_urine': 0, 'passage_of_gases': 0, 'internal_itching': 0, 'toxic_look_(typhos)': 0,
        'depression': 0, 'irritability': 0, 'muscle_pain': 0, 'altered_sensorium': 0, 'red_spots_over_body': 0, 'belly_pain': 0,
        'abnormal_menstruation': 0, 'dischromic _patches': 0, 'watering_from_eyes': 0, 'increased_appetite': 0, 'polyuria': 0,
        'family_history': 0, 'mucoid_sputum': 0, 'rusty_sputum': 0, 'lack_of_concentration': 0, 'visual_disturbances': 0,
        'receiving_blood_transfusion': 0, 'receiving_unsterile_injections': 0, 'coma': 0, 'stomach_bleeding': 0, 'distention_of_abdomen':
        'history_of_alcohol_consumption': 0, 'fluid_overload.1': 0, 'blood_in_sputum': 0, 'prominent_veins_on_calf': 0, 'palpitations': 0
```

```
Predicted Disease: Fungal infection
PS D:\Internship-project\Disease-Prediction>
 *  History restored

PS D:\Internship-project>
```

Figure 3.7 Giving Input

Chapter 4

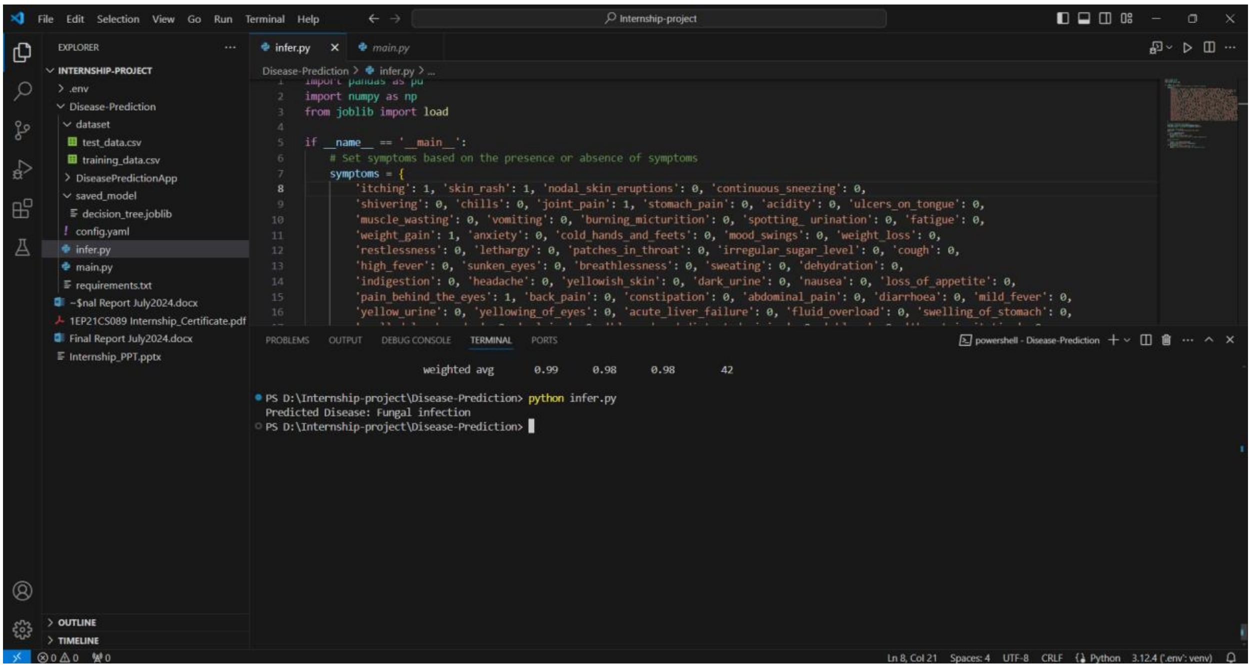
REFLECTION


File Edit Selection View Go Run Terminal Help
Internship-project
EXPLORER
INTERNSHIP-PROJECT
.env
Disease-Prediction
dataset
test_data.csv
training_data.csv
DiseasePredictionApp
saved_model
decision_tree.joblib
config.yaml
infer.py
main.py
requirements.txt
~$nal Report July2024.docx
1EP21CS089 Internship_Certificate.pdf
Final Report July2024.docx
Internship_PPT.pptx
OUTLINE
TIMELINE
infer.py
main.py
Disease-Prediction > main.py > ...
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
powershell - Disease-Prediction
Ln 1, Col 1 Spaces: 4 UTF-8 LF Python 3.12.4 ('.env': venv)

```
class DiseasePrediction:
    def  load test dataset(self):
```

```
                 Drug Reaction       1.00      1.00      1.00         1
               Fungal infection       1.00      0.50      0.67         2
                           GERD       1.00      1.00      1.00         1
                 Gastroenteritis       1.00      1.00      1.00         1
                    Heart attack       1.00      1.00      1.00         1
                     Hepatitis B       1.00      1.00      1.00         1
                     Hepatitis C       1.00      1.00      1.00         1
                     Hepatitis D       1.00      1.00      1.00         1
                     Hepatitis E       1.00      1.00      1.00         1
                   Hypertension        1.00      1.00      1.00         1
                Hyperthyroidism       1.00      1.00      1.00         1
                   Hypoglycemia       1.00      1.00      1.00         1
                 Hypothyroidism       1.00      1.00      1.00         1
                        Impetigo       1.00      1.00      1.00         1
                        Jaundice       1.00      1.00      1.00         1
                         Malaria       1.00      1.00      1.00         1
                        Migraine       1.00      1.00      1.00         1
                Osteoarthristis       1.00      1.00      1.00         1
     Paralysis (brain hemorrhage)       1.00      1.00      1.00         1
             Peptic ulcer diseae       1.00      1.00      1.00         1
                       Pneumonia       1.00      1.00      1.00         1
                       Psoriasis       1.00      1.00      1.00         1
                    Tuberculosis       1.00      1.00      1.00         1
                         Typhoid       1.00      1.00      1.00         1
         Urinary tract infection       1.00      1.00      1.00         1
                  Varicose veins       1.00      1.00      1.00         1
                     hepatitis A       1.00      1.00      1.00         1

                        accuracy                           0.98        42
                       macro avg       0.99      0.99      0.98        42
                    weighted avg       0.99      0.98      0.98        42

PS D:\Internship-project\Disease-Prediction>
```

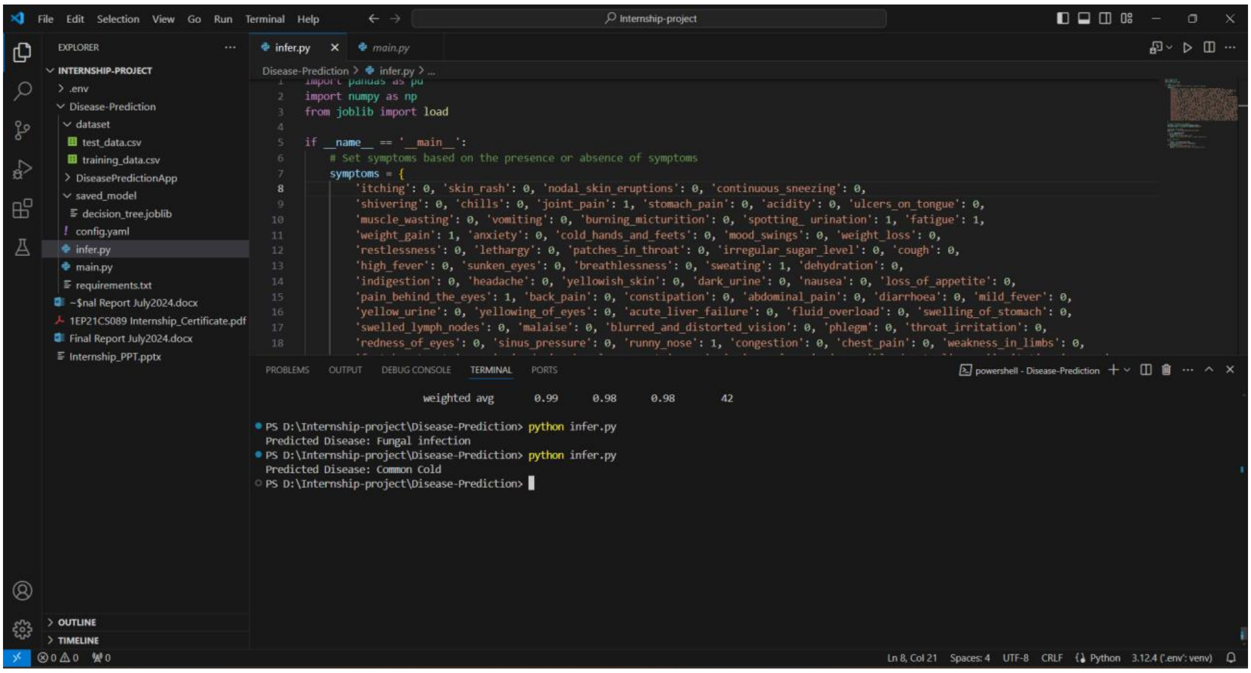
Figure 4.1 Model Training


File Edit Selection View Go Run Terminal Help
Internship-project
EXPLORER
INTERNSHIP-PROJECT
.env
Disease-Prediction
dataset
test_data.csv
training_data.csv
DiseasePredictionApp
saved_model
decision_tree.joblib
config.yaml
infer.py
main.py
requirements.txt
~$nal Report July2024.docx
1EP21CS089 Internship_Certificate.pdf
Final Report July2024.docx
Internship_PPT.pptx
OUTLINE
TIMELINE
infer.py
main.py
Disease-Prediction > infer.py > ...
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
powershell - Disease-Prediction
Ln 8, Col 21 Spaces: 4 UTF-8 CRLF Python 3.12.4 ('.env': venv)

```
import pandas as pd
import numpy as np
from joblib import load

if __name__ == '__main__':
    # Set symptoms based on the presence or absence of symptoms
    symptoms = {
        'itching': 1, 'skin_rash': 1, 'nodal_skin_eruptions': 0, 'continuous_sneezing': 0,
        'shivering': 0, 'chills': 0, 'joint_pain': 1, 'stomach_pain': 0, 'acidity': 0, 'ulcers_on_tongue': 0,
        'muscle_wasting': 0, 'vomiting': 0, 'burning_micturition': 0, 'spotting_ urination': 0, 'fatigue': 0,
        'weight_gain': 1, 'anxiety': 0, 'cold_hands_and_feets': 0, 'mood_swings': 0, 'weight_loss': 0,
        'restlessness': 0, 'lethargy': 0, 'patches_in_throat': 0, 'irregular_sugar_level': 0, 'cough': 0,
        'high_fever': 0, 'sunken_eyes': 0, 'breathlessness': 0, 'sweating': 0, 'dehydration': 0,
        'indigestion': 0, 'headache': 0, 'yellowish_skin': 0, 'dark_urine': 0, 'nausea': 0, 'loss_of_appetite': 0,
        'pain_behind_the_eyes': 1, 'back_pain': 0, 'constipation': 0, 'abdominal_pain': 0, 'diarrhoea': 0, 'mild_fever': 0,
        'yellow_urine': 0, 'yellowing_of_eyes': 0, 'acute_liver_failure': 0, 'fluid_overload': 0, 'swelling_of_stomach': 0,
```

```
                    weighted avg       0.99      0.98      0.98        42

PS D:\Internship-project\Disease-Prediction> python infer.py
Predicted Disease: Fungal infection
PS D:\Internship-project\Disease-Prediction>
```

Figure 4.2 Output Case One


File Edit Selection View Go Run Terminal Help
Internship-project
EXPLORER
INTERNSHIP-PROJECT
.env
Disease-Prediction
dataset
test_data.csv
training_data.csv
DiseasePredictionApp
saved_model
decision_tree.joblib
config.yaml
infer.py
main.py
requirements.txt
~$nal Report July2024.docx
1EP21CS089 Internship_Certificate.pdf
Final Report July2024.docx
Internship_PPT.pptx
OUTLINE
TIMELINE
infer.py
main.py
Disease-Prediction > infer.py > ...
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
powershell - Disease-Prediction
Ln 8, Col 21 Spaces: 4 UTF-8 CRLF Python 3.12.4 ('.env': venv)

```python
import pandas as pd
import numpy as np
from joblib import load

if __name__ == '__main__':
    # Set symptoms based on the presence or absence of symptoms
    symptoms = {
        'itching': 0, 'skin_rash': 0, 'nodal_skin_eruptions': 0, 'continuous_sneezing': 0,
        'shivering': 0, 'chills': 0, 'joint_pain': 1, 'stomach_pain': 0, 'acidity': 0, 'ulcers_on_tongue': 0,
        'muscle_wasting': 0, 'vomiting': 0, 'burning_micturition': 0, 'spotting_ urination': 1, 'fatigue': 1,
        'weight_gain': 1, 'anxiety': 0, 'cold_hands_and_feets': 0, 'mood_swings': 0, 'weight_loss': 0,
        'restlessness': 0, 'lethargy': 0, 'patches_in_throat': 0, 'irregular_sugar_level': 0, 'cough': 0,
        'high_fever': 0, 'sunken_eyes': 0, 'breathlessness': 0, 'sweating': 1, 'dehydration': 0,
        'indigestion': 0, 'headache': 0, 'yellowish_skin': 0, 'dark_urine': 0, 'nausea': 0, 'loss_of_appetite': 0,
        'pain_behind_the_eyes': 1, 'back_pain': 0, 'constipation': 0, 'abdominal_pain': 0, 'diarrhoea': 0, 'mild_fever': 0,
        'yellow_urine': 0, 'yellowing_of_eyes': 0, 'acute_liver_failure': 0, 'fluid_overload': 0, 'swelling_of_stomach': 0,
        'swelled_lymph_nodes': 0, 'malaise': 0, 'blurred_and_distorted_vision': 0, 'phlegm': 0, 'throat_irritation': 0,
        'redness_of_eyes': 0, 'sinus_pressure': 0, 'runny_nose': 1, 'congestion': 0, 'chest_pain': 0, 'weakness_in_limbs': 0,
```

```
                    weighted avg       0.99      0.98      0.98        42

PS D:\Internship-project\Disease-Prediction> python infer.py
Predicted Disease: Fungal infection
PS D:\Internship-project\Disease-Prediction> python infer.py
Predicted Disease: Common Cold
PS D:\Internship-project\Disease-Prediction> 
```

Figure 4.3 Output Case Two

Chapter 5

CONCLUSION

The package was designed in such a way that future modifications can be done easily. From the development of the project, several conclusions can be deduced. Automation of the entire system improves efficiency and provides a friendly graphical user interface, which proves to be better than existing systems. It gives appropriate access to authorized users based on their permissions, effectively overcoming delays in communication. Updating information becomes much easier, and the system's security, data security, and reliability are striking features. Additionally, the system has adequate scope for future modifications if necessary.

The user-friendly design of the system ensures that it is accessible and easy to navigate, which significantly enhances user experience and reduces the learning curve for new users. The robust security measures implemented in the system protect sensitive data, ensuring that only authorized personnel have access to critical information. Furthermore, the system's modular architecture allows for seamless integration of new features and functionalities, providing flexibility for future enhancements. Overall, the system not only meets the current needs but is also well-prepared to adapt to evolving requirements, ensuring its longevity and relevance in the rapidly advancing field of technology.

REFERENCES

1. Python for Data Science Handbook: A comprehensive guide by Jake VanderPlas that covers essential Python libraries and tools for data science, including NumPy, Pandas, Matplotlib, Scikit-Learn, and more. Python for Data Science Handbook
2. Machine Learning Mastery with Python: A practical guide by Jason Brownlee that teaches machine learning concepts and how to implement them using Python, covering a range of algorithms and techniques. Machine Learning Mastery with Python
3. Kaggle - Chronic Disease Prediction Dataset: A dataset available on Kaggle, along with kernels and notebooks shared by the community, demonstrating various approaches to predicting chronic diseases using Python. Kaggle - Chronic Disease Prediction Dataset
4. Scikit-Learn Documentation: The official documentation for Scikit-Learn, a powerful Python library for machine learning, including tutorials and examples on implementing various algorithms. Scikit-Learn Documentation
5. Towards Data Science - Chronic Disease Prediction: Articles and tutorials on Towards Data Science that explore methods and case studies for chronic disease prediction using Python and machine learning techniques. Towards Data Science - Chronic Disease Prediction
6. "Python Machine Learning" by Sebastian Raschka and Vahid Mirjalili: A comprehensive book that covers machine learning algorithms and techniques with practical examples and case studies using Python. Python Machine Learning