**Project Report Format –** **Team 592802**

The provided project report outlines the development of a disease prediction model using machine learning techniques. Here's an elaboration on each section:

1. Introduction

1.1 Project Overview

The project aims to pioneer the development of a disease prediction model using advanced machine learning algorithms such as K Nearest Neighbors, Support Vector Machine (SVM), Decision Tree, and Random Forest. The model is trained on a curated dataset from Kaggle, with the goal of delivering precise predictions through innovative algorithms and a user-friendly interface.

1.2 Purpose

The project addresses challenges in contemporary healthcare, including time constraints, limited online symptom analysis, and financial barriers to traditional doctor visits. The purpose is to empower users with a proactive tool for preventive diagnosis, enabling informed health decisions and overcoming limitations imposed by time and finances.

2. Literature Survey

2.1 Existing Problem

The report highlights the challenges users face in accessing precise health information online and the financial implications of seeking professional medical advice. The proposed solution involves applying machine learning algorithms to enhance online health consultation.

2.2 References

The project draws on references related to Supervised Learning, K Nearest Neighbors, SVM, Decision Tree, Random Forest algorithms, Evaluation Metrics, and Flask Basics, forming the foundation for a comprehensive understanding of the underlying technologies.

2.3 Problem Statement Definition

The central problem lies in the need for accurate disease prediction and preventive diagnosis through machine learning, addressing concerns related to trust, regulatory compliance, and a seamless user interface.

3. Ideation & Proposed Solution

3.1 Empathy Map Canvas

The ideation process involves understanding the user's perspective through an empathy map, guiding the formulation of user-centric solutions in healthcare and disease prediction.

3.2 Ideation & Brainstorming

Intensive brainstorming results in a visionary proposal for a web application with an intuitive interface for symptom input. The core of the solution is the integration of machine learning models for accurate disease prediction.

4. Requirement Analysis

4.1 Functional Requirements

Functional requirements include accurate disease prediction, a user-friendly interface, adherence to healthcare regulations, and measures to instill user trust.

4.2 Non-Functional Requirements

Non-functional requirements specify performance benchmarks, including high model accuracy, secure data handling, and a responsive web design for a seamless user experience.

5. Project Design

5.1 Data Flow Diagrams & User Stories

Data flow diagrams and user stories visually represent information exchange within the system, aiding in crystallizing user interactions and refining project design.

5.2 Solution Architecture

The solution architecture involves a Flask-based web application seamlessly integrated with machine learning models, ensuring a streamlined and efficient system.

6. Project Planning & Scheduling

6.1 Technical Architecture

The technical architecture features a Flask application, HTML templates, CSS for styling, and machine learning models encapsulated as pickle files, ensuring a robust and scalable solution.

6.2 Sprint Planning & Estimation

Sprint planning encompasses tasks spanning data collection, preprocessing, exploratory data analysis, model building, testing, hyperparameter tuning, model deployment, and web application integration for a methodical and incremental approach.

6.3 Sprint Delivery Schedule

Sprints are scheduled incrementally, focusing on specific facets of the project for systematic progression toward the final deliverable.

7. Coding & Solutioning

7.1 Feature 1: Data Collection and Preparation

The coding phase involves implementing Feature 1, which includes collecting data from Kaggle and preprocessing it by eliminating redundant columns and addressing missing values.

7.2 Feature 2: Model Building

Feature 2 entails implementing machine learning models (K Nearest Neighbors, SVM, Decision Tree, and Random Forest) with rigorous evaluation of model performance to identify and preserve the optimal model.

7.3 Database Schema (if Applicable)

No database is incorporated, as the model operates seamlessly on pre-collected and preprocessed data.

8. Performance Testing

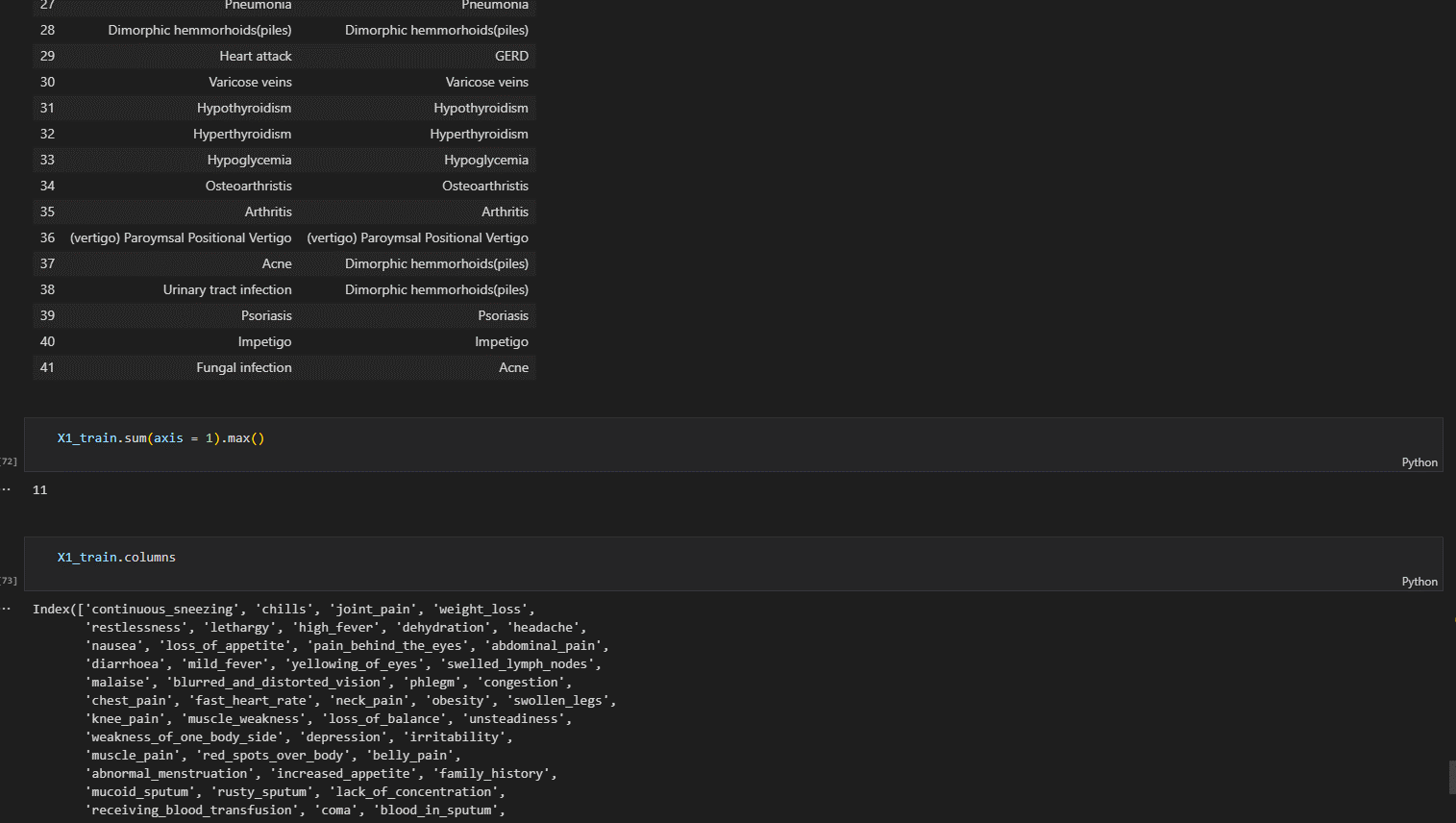
8.1 Performance Metrics

Performance testing involves using metrics such as model accuracy, precision, recall, and F1 score as quantitative benchmarks to assess the robustness and efficacy of the predictive model.

9. Results

9.1 Output Screenshots

Screenshots serve as tangible evidence of the project's success, showcasing the web application interface, symptom input, and the seamlessly predicted disease output.



A screenshot of a computer screen

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

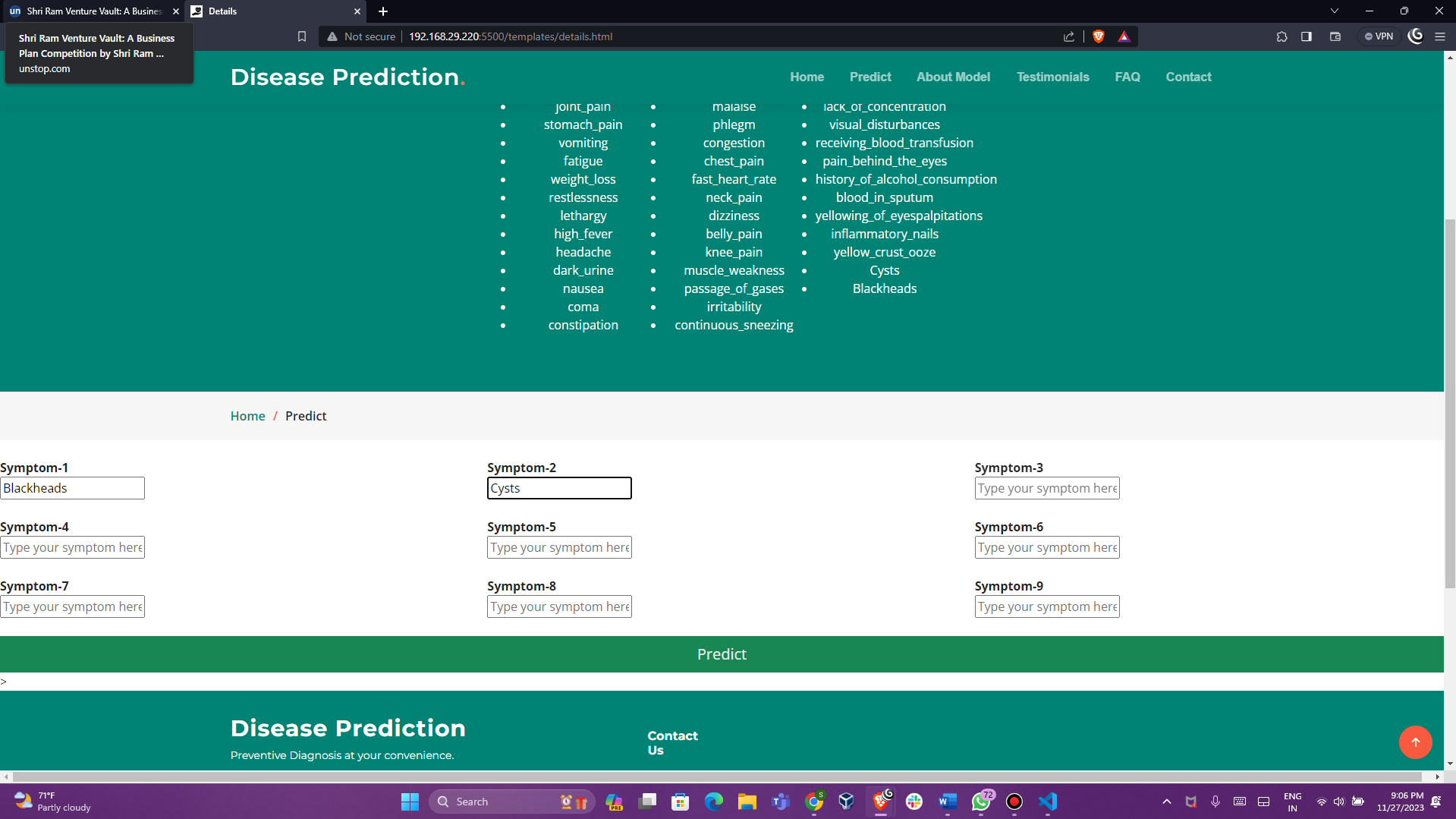
Description automatically generated

A screenshot of a computer

Description automatically generated

A computer screen with text on it

Description automatically generated



A screenshot of a computer

Description automatically generated

10. Advantages & Disadvantages

Advantages

Triumvirate of accurate disease prediction, cost-effective preventive diagnosis, and a user-friendly interface contribute to a transformative healthcare experience for users.

Disadvantages

Acknowledged limitations include a degree of dependency on machine learning models and constraints of being limited to the provided dataset, overshadowed by the project's overarching advantages.

11. Conclusion

The project successfully resolves challenges in online health consultation and preventive diagnosis, delivering a reliable machine learning-based disease prediction system with high accuracy and a tangible impact on healthcare decision-making processes.

12. Future Scope

The future trajectory envisions continuous model training with updated datasets, integration with real-time health monitoring devices, and expansion of the model's scope to cover a broader spectrum of diseases.

13. Appendix

The project repository on GitHub serves as the source code repository, promoting transparency and collaboration within the development community. Links to the GitHub repository and project demo are provided.

In summary, the project report provides a comprehensive overview of the disease prediction model, covering its development, challenges addressed, solution architecture, and future scope. The inclusion of references, performance metrics, and tangible results through screenshots adds depth and credibility to the report.

**GitHub & Project Demo Link:**

- [GitHub Repository]( https://github.com/smartinternz02/SI-GuidedProject-612297-1699706390/tree/main)

- [Project Demo]( https://youtu.be/3PfH29oHUzA)