Sukruthi - UP-II Report

by Sukruthi Rao

Submission date: 09-Jan-2024 09:36AM (UTC+0530)
Submission ID: 2268222801
File name: UP-II_report.docx (1.83M)
Word count: 7088
Character count: 44754

LEVERAGING MACHINE LEARNING FOR EARLY PREDICTION OF LIFESTYLE DISEASES: A DATA-DRIVEN APPROACH

A PROJECT REPORT

Submitted by,

SUKRUTHI RAO - 20201CBC0004
S HARISH - 20201CBC0008
R SAISARAN - 20201CBD0024

Under the guidance of,

Mr. B.PAKRUDDIN

in partial fulfillment for the award of the

degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

At

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

PRESIDENCY UNIVERSITY

BENGALURU

JANUARY 2024

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE ENGINEERING & INFORMATION SCIENCE

CERTIFICATE

This is to certify that the Project report **"LEVERAGING MACHINE LEARNING FOR EARLY PREDICTION OF LIFESTYLE DISEASES A DATA-DRIVEN APPROACH"** being submitted by "**SUKRUTHI RAO, S HARISH, R SAISARAN**" bearing roll numbers "**20201CBC0004, 20201CBC0008, 20201CBD0024**" in partial fulfilment of requirement for the award of degree of Bachelor of Technology in Computer Science and Engineering is a bonafide work carried out under my supervision.

Mr. B. PAKRUDDIN
Assistant Professor
School of CSE&IS
Presidency University

Dr. SENTHIL KUMAR
Professor & HoD
School of CSE&IS
Presidency University

Dr. C. KALAIARASAN
Associate Dean
School of CSE&IS
Presidency University

Dr. SHAKKEERA L
Associate Dean
School of CSE&IS
Presidency University

Dr. SAMEERUDDIN KHAN
Dean
School of CSE&IS
Presidency University

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE ENGINEERING & INFORMATION SCIENCE

DECLARATION

We hereby declare that the work, which is being presented in the project report entitled **LEVERAGING MACHINE LEARNING FOR EARLY PREDICTION OF LIFESTYLE DISEASES: A DATA-DRIVEN APPROACH** in partial fulfilment for the award of Degree of **Bachelor of Technology** in **Computer Science and Engineering**, is a record of our own investigations carried under the guidance of **B. PAKRUDDIN, Assistant Professor, School of Computer Science Engineering & Information Science, Presidency University, Bengaluru.**

We have not submitted the matter presented in this report anywhere for the award of any other Degree.

SUKRUTHI RAO, 20201CBC0004
S HARISH, 20201CBC0008
R SAISARAN, 20201CBD0024

ABSTRACT

This project investigates the utilization of machine learning methods for the early detection of lifestyle diseases, with the goal of enhancing proactive healthcare interventions. Lifestyle-related conditions, including diabetes, cardiovascular diseases, and Parkinson's, present substantial global public health challenges. Conventional healthcare strategies often rely on retrospective analysis, limiting their effectiveness in preventing and managing these ailments. In contrast, the proposed data-driven approach harnesses machine learning algorithms to analyze diverse patient data sets, encompassing demographic details, lifestyle choices, genetic predispositions, and biomarkers.

The research entails constructing and deploying predictive models that leverage extensive datasets to uncover patterns and connections linked to the initiation of lifestyle diseases. Diverse machine learning algorithms like decision trees, support vector machines, and deep learning models are utilized to optimize predictive accuracy and broaden applicability. Additionally, the study explores techniques such as feature engineering to improve model interpretability and facilitate the incorporation of domain knowledge.

The project's objectives include contributing to the creation of an effective and scalable early prediction system for lifestyle diseases. Through machine learning, healthcare professionals can identify individuals at risk well in advance of symptom manifestation, enabling timely interventions and personalized healthcare plans. This research holds the potential to revolutionize healthcare practices by emphasizing a proactive and preventive approach to address the increasing impact of lifestyle diseases on public health systems.

ACKNOWLEDGEMENT

First of all, we indebted to the **GOD ALMIGHTY** for giving me an opportunity to excel in our efforts to complete this project on time.

We express our sincere thanks to our respected dean **Dr. Md. Sameeruddin Khan**, Dean, School of Computer Science Engineering & Information Science, Presidency University for getting us permission to undergo the project.

We record our heartfelt gratitude to our beloved Associate Deans **Dr. C. Kalaiarasan and Dr. Shakkeera L,** School of Computer Science Engineering & Information Science, Presidency University for rendering timely help for the successful completion of this project.

We would like to convey our gratitude and heartfelt thanks to the University Project-II Coordinators **Dr. Sanjeev P Kaulgud, Dr. Mrutyunjaya MS** and also the department Project Coordinators.

We are greatly indebted to our guide **Mr. B. PAKRUDDIN,** Assistant Professor, School of Computer Science Engineering & Information Science, Presidency University for his inspirational guidance, valuable suggestions and providing us a chance to express our technical capabilities in every respect for the completion of the project work.

We thank our family and friends for the strong support and inspiration they have provided us in bringing out this project.

SUKRUTHI RAO

S HARISH

R SAISARAN

LIST OF TABLES

Sl. No.	Table Name	Table Caption	Page No.
1	Table 1	Test Cases for Diabetes Prediction	25
2	Table 2	Test Cases for Heart Disease Prediction	26
3	Table 3	Test Cases for Parkinson's Disease Prediction	27
4	Table 4	Project Timeline	28

LIST OF FIGURES

Sl. No.	Figure Name	Caption	Page No.
1	Figure 1	System Design	16
2	Figure 2	Use Case Diagram between User and System	18
3	Figure 3	Use Case Diagram for Diabetes Prediction	19
4	Figure 4	Use Case Diagram for Heart Disease Prediction	19
5	Figure 5	Use Case Diagram for Parkinson's Prediction	20
6	Figure 6	Class Diagram for Multiple Disease Prediction	21
7	Figure 7	Gantt chart Timeline	28
8	Figure 8	Interface of the Disease Prediction System	38
9	Figure 9	Different Deployment Options	39
10	Figure 10	Settings for Modifications	40
11	Figure 11	Heart Disease Prediction	41
12	Figure 12	Diabetes Disease Prediction positive	42
13	Figure 13	Diabetes Disease Prediction negative	42
14	Figure 14	Parkinson's Disease Prediction negative	43
15	Figure 15	Parkinson's Disease Prediction positive	43
16	Figure 16	Mobile View Web App	44
17	Figure 17	Mobile view for Diabetes Prediction input	45
18	Figure 18	Mobile view settings	46
19	Figure 19	Acceptance Letter	47
20	Figure 20	Observation of the paper	48

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	ABSTRACT	i
	ACKNOWLEDGMENT	ii
1.	INTRODUCTION	1
2.	LITERATURE SURVEY	3
3.	RESEARCH GAPS OF EXISTING METHODS	9
4.	PROPOSED METHODOLOGY	11
5.	OBJECTIVES	14
6.	SYSTEM DESIGN & IMPLEMENTATION	16
	6.1 DIAGRAMS	18
	6.2 REQUIREMENT ANALYSIS	22
	6.3 TESTING	23
7.	TIMELINE FOR EXECUTION OF PROJECT	28
8.	OUTCOMES	29
9.	RESULTS AND DISCUSSIONS	31
10.	CONCLUSION	33
	REFERENCES	34

CHAPTER-1
INTRODUCTION

The current state of global healthcare faces a significant challenge – the increasing prevalence of lifestyle diseases. Diseases like diabetes, heart conditions, and Parkinson's are widespread, putting immense pressure on healthcare systems across the globe. What complicates matters is that many of these diseases are not only widespread but also avoidable or controllable through timely intervention and lifestyle adjustments.

Addressing this complex healthcare issue, a ray of hope emerges through the utilization of machine learning for the early prediction of lifestyle diseases. This innovative approach aims to transform healthcare strategies by utilizing advanced data-driven techniques. With the deployment of machine learning models, healthcare professionals can identify individuals at risk and take proactive steps for prevention. The ultimate objective is to reduce the strain on healthcare resources and, more importantly, improve overall public health outcomes.

1. Significance of Early Prediction:

Central to this revolutionary method is the acknowledgment of the vital significance of early detection in influencing health results. The capability to recognize potential lifestyle diseases in their initial phases enables the implementation of proactive measures, personalized healthcare plans, and precise lifestyle adjustments. The outcome is a potential mitigation of disease progression and a subsequent reduction in the related healthcare expenses.

2. Data-Driven Approach:

The foundation of this suggested remedy is a strong reliance on data-driven methodologies. Extensive datasets, covering a range of demographic details, lifestyle practices, and relevant health indicators, form the basis for machine learning models. Equipped with the capacity to scrutinize complex patterns and pinpoint risk elements, these models can generate predictions utilizing both historical and current data.

3. Machine Learning Models for Diabetes Prediction:

An essential element of the system focuses on a specialized machine learning model designed for forecasting diabetes. This model considers a variety of factors, spanning from the count of pregnancies to glucose levels and blood pressure. Individuals input their details, and the machine learning model delivers detailed predictions concerning the probability of diabetes.

4. Machine Learning Models for Heart Disease Prediction:

Broadening its reach, the system integrates a specific model tailored for forecasting heart disease. Attributes including age, gender, chest pain types, and a range of cardiovascular metrics are taken into account. The machine learning model analyzes this data, providing valuable information about the likelihood of an individual developing heart disease.

5. Machine Learning Models for Parkinson's Disease Prediction:

Continuing to enhance its functionalities, the system explores the prediction of Parkinson's disease. Characteristics related to voice, such as pitch and jitter, are considered. Users provide pertinent details, and the machine learning model produces predictions regarding the existence or non-existence of Parkinson's disease.

6. Empowering Preventive Healthcare:

The integration of these machine learning models empowers both healthcare providers and individuals to implement proactive measures. Early prediction serves as the catalyst for specific interventions, personalized lifestyle adjustments, and ongoing monitoring – facilitating a shift towards a preventive healthcare approach.

CHAPTER-2

LITERATURE SURVEY

Sl. No.	Title	Methods	Advantages	Limitations
1	Utilizing Machine Learning for Early Disease Prediction	In the realm of early disease prediction, machine learning (ML) serves as a versatile tool, leveraging diverse datasets comprising medical records, genetic information, and lifestyle details. Support vector machines, random forests, and deep learning algorithms are applied to discern subtle patterns within these datasets. Notably, convolutional neural networks (CNNs) have shown remarkable efficacy, as exemplified by Esteva et al. (2017) in predicting skin cancer. Augmenting the original samples with diverse data enhances the robustness and generalization capabilities of these models.	The inherent strength of ML lies in its ability to recognize intricate patterns within complex datasets, making it adopt at early disease detection. By analysing diverse data sources, ML algorithms can identify trends and correlations that might elude traditional diagnostic methods.	The interpretability of some ML models, especially deep learning ones, can be a challenge. Understanding the reasoning behind a specific prediction may be obscured, limiting the model's transparency. Additionally, the efficacy of ML models heavily depends on the quality and representativeness of the training data, making biases in the data a potential concern.

2	Predicting Diseases through Biomarkers and Genomic Insights	Advancements in genomics and molecular biology have opened avenues for predicting diseases through the identification of biomarkers. Researchers utilize genomic analysis to discover links between specific biomarkers and various diseases. Genomic predictors, as demonstrated by Smith et al. (2019) and Li et al. (2020), are employed to forecast the risk of diseases such as Alzheimer's and certain types of cancer.	Biomarkers provide invaluable early indicators of disease risk, offering a proactive approach to healthcare. Genomic insights further enhance the precision of predictions, enabling tailored interventions based on an individual's genetic makeup.	Nevertheless, the identification and validation of biomarkers can be intricate processes, often requiring significant resources. Furthermore, the applicability of genomic predictors might be disease-specific and not universally applicable across different health conditions.
3	Harnessing Wearable and IoT Technologies for Continuous Monitoring	The integration of wearable devices and Internet of Things (IoT) technologies introduces a paradigm shift in healthcare by enabling continuous monitoring of physiological parameters and lifestyle behaviours. Wearables, such as smartwatches and fitness trackers, serve as data sources, and IoT facilitates real-time tracking of health-related metrics.	Continuous monitoring offers real-time insights into an individual's health, providing a dynamic and comprehensive view. Wearable technologies not only capture physiological data but also promote active user engagement in health management.	The reliability of predictions hinges on the accuracy and consistency of data from wearables, which may be influenced by factors such as device accuracy and user compliance. Additionally, not everyone may readily adopt or adhere to wearable technologies, introducing potential gaps in data collection.

4	Enhancing Accuracy with Multi-Modal Data Integration	Efforts to enhance the accuracy of early disease prediction models have led to the integration of various data modalities, including clinical records, imaging data, and environmental factors. Chen et al. (2021) and Kim et al. (2022) showcase the potential benefits of combining diverse datasets for a more comprehensive understanding of an individual's health status.	The integration of multi-modal data not only improves the accuracy of predictive models but also provides a more holistic view of an individual's health. This approach facilitates more robust predictions across different diseases by capturing a broader spectrum of relevant information.	The integration data from different sources presents challenges related to standardization, compatibility, and the development of models that effectively utilize diverse data. This complexity can render the process resource-intensive and may require advanced technical expertise.
5	Addressing Ethical and Privacy Concerns	The ethical implications and privacy concerns associated with early disease prediction are actively addressed through a multifaceted approach. Robust data security measures, informed consent protocols, and guidelines for responsible model use are implemented to strike a balance between the benefits of early prediction and the protection of individual rights and privacy.	Ethical considerations and privacy measures underscore a commitment to responsible practices in the development and deployment of predictive models. This not only ensures compliance with evolving regulations but also builds trust among users.	Despite these measures, adhering to regulatory compliance can be challenging, requiring continuous adaptation to evolving standards. Moreover, individual perceptions of data privacy and security may still pose challenges, highlighting the need for ongoing efforts to address public concerns.

6	Early Detection of Mental Health Disorders using Natural Language Processing (NLP)	Natural Language Processing (NLP) techniques are employed to analyze textual data, such as social media posts, electronic health records, and patient interviews. Sentiment analysis and topic modeling algorithms help identify linguistic patterns indicative of mental health conditions.	NLP enables the extraction of valuable insights from unstructured text, offering a non-intrusive and scalable approach to mental health screening. Early detection allows for timely intervention and personalized treatment strategies.	Ethical concerns arise regarding privacy and consent when analyzing personal textual data. Additionally, the subjective nature of mental health symptoms poses challenges in standardizing NLP models across diverse populations.
7	Predicting Cardiovascular Risks through Physiological Signals and Machine Learning	Wearable devices and physiological sensors capture real-time data on heart rate, blood pressure, and other vital signs. Machine learning algorithms analyze this continuous stream of data to predict cardiovascular risks and potential anomalies.	Continuous monitoring provides a comprehensive understanding of cardiovascular health, enabling early identification of risk factors. Personalized interventions can be recommended based on individual physiological patterns.	The accuracy of predictions relies on the precision and reliability of physiological sensors. User compliance with long-term wearables may vary, affecting the consistency of data collection.

8	Integrating Social Determinants for Holistic Health Prediction	Machine learning models incorporate social determinants of health, such as socioeconomic status, environmental factors, and community resources, to enhance predictive accuracy. Multi-modal data integration involves combining clinical data with community-level information.	Including social determinants offers a more comprehensive understanding of health disparities and enables targeted interventions. This approach recognizes the impact of broader social and environmental factors on individual well-being.	Data standardization and compatibility issues arise when integrating diverse datasets. Ensuring equitable representation of social determinants across different populations poses a challenge.
9	Early Prediction of Infectious Disease Outbreaks using Data Analytics	Data analytics techniques, including machine learning and statistical models, analyze epidemiological data, travel patterns, and environmental factors to predict and monitor the spread of infectious diseases. Real-time data feeds from healthcare systems and global monitoring networks are utilized.	Early detection of infectious disease outbreaks enables prompt public health responses, reducing the potential impact. Predictive models contribute to effective resource allocation and intervention strategies.	The dynamic nature of infectious diseases and the need for timely data pose challenges. Model accuracy depends on the availability and accuracy of real-time data, which may be limited in some regions.

10	Personalized Medicine through Machine Learning in Oncology	Machine learning models analyze genomic data, tumor characteristics, and treatment responses to predict individualized cancer outcomes. The integration of diverse datasets, including patient records and molecular profiles, enables precision medicine approaches.	Personalized cancer predictions enhance treatment efficacy by tailoring interventions to individual patient profiles. Machine learning aids in identifying potential biomarkers for targeted therapies.	Challenges exist in the validation and interpretation of complex genomic data. The applicability of personalized models may vary across different cancer types, necessitating disease-specific approaches.

CHAPTER-3

RESEARCH GAPS OF EXISTING METHODS

1. Limited Feature Set and Relevance:

- Disadvantage: The application depends on a particular group of input features for each disease prediction model. The choice of features may not encompass the complete intricacy of the diseases, potentially resulting in less-than-optimal predictive performance.
- Research Gap: Subsequent research could investigate the integration of more exhaustive and pertinent medical data. This could entail utilizing advanced feature engineering methods or incorporating additional biomarkers to improve the precision of predictions.

2. Utilization of Traditional Machine Learning Models:

- Disadvantage: The application utilizes conventional machine learning algorithms like logistic regression and decision trees. While these models are interpretable, they may not capture the complex nonlinear relationships inherent in medical data.
- Research Gap: Exploring more advanced models, particularly deep learning architectures, has the potential to unveil concealed patterns and enhance the system's predictive capabilities. Nonetheless, this approach poses challenges such as heightened complexity and the requirement for larger datasets.

3. Lack of Comprehensive Model Evaluation:

- Disadvantage: The application does not incorporate a comprehensive model evaluation and validation procedure, posing challenges in evaluating the generalizability and robustness of the predictive models.
- Research Gap: Subsequent research ought to concentrate on adopting stringent model validation techniques, encompassing cross-validation and external validation using diverse datasets. This would guarantee consistent model performance across various populations and clinical scenarios.

4. Absence of Interpretability Features:

- Disadvantage: The application does not include functionalities that offer explanations regarding how the models reach their predictions, impeding the interpretability and reliability of the results.
- Research Gap: Investigating interpretable machine learning methods and incorporating tools for model explainability could fill this void. This holds particular significance in a healthcare setting, where comprehending the reasoning behind predictions is essential for gaining acceptance among healthcare professionals.

CHAPTER-4
PROPOSED METHODOLOGY

1. Data Collection:

- Collect electronic health records (EHRs): Electronic health records encompass extensive patient information, comprising demographics, medical history, diagnoses, medications, and laboratory test results. Utilizing this data can aid in identifying risk factors associated with lifestyle diseases.
- Gather genetic information: Genetic data offers insights into an individual's susceptibility to specific lifestyle diseases. Utilizing this data can enhance risk evaluation and tailor preventive interventions on a personalized level.
- Collect lifestyle information: Factors like diet, physical activity, and smoking habits play a crucial role in influencing the likelihood of lifestyle diseases. This information can be obtained through surveys, questionnaires, or wearable devices.
- Gather real-time health monitoring information: Continuous monitoring of health metrics like heart rate, blood pressure, and blood sugar levels in real-time can offer valuable insights into an individual's present health condition and detect early indications of potential health issues.

2. Data Pre-processing:

- Preprocess and standardize data: Information originating from diverse sources might exhibit disparities, absence of values, and formatting discrepancies. Through data preprocessing and standardization, the data is made uniform and apt for analysis.
- Integrate diverse data origins: Unify data from electronic health records (EHRs), genetic data, lifestyle information, and real-time health monitoring data into a consolidated dataset. This enables a thorough examination of individual health profiles.

3. Feature Engineering:

- Identify pertinent characteristics: Single out the most significant attributes from the processed data that exhibit robust predictive capabilities for lifestyle diseases.
- Generate additional variables: Formulate new variables by amalgamating existing features or converting them into more meaningful representations. This has the potential to enhance the predictive precision of the models.

4. Model Development:

- Apply machine learning methodologies: Employ a range of machine learning algorithms, including logistic regression, decision trees, and random forests, for forecasting the likelihood of lifestyle diseases.
- Investigate deep learning architectures: Utilize deep learning models, like neural networks, to grasp intricate patterns and connections within the data, potentially enhancing predictive capabilities.
- Tune hyper parameters: Refine the parameters of the selected machine learning algorithms to attain optimal performance.

5. Continuous Monitoring:

- Incorporate wearable gadgets and mobile applications: Link the system with wearable devices and mobile applications to consistently gather real-time health information.
- Evaluate real-time information: Scrutinize real-time data to detect potential health concerns and furnish users with timely alerts or suggestions..
- User Interface: Develop an interface that is user-friendly, enabling individuals to effortlessly access their health data, risk assessment outcomes, and personalized preventive suggestions.

6. Validation and Testing:

- Assess model effectiveness: Verify the effectiveness of the predictive models across various datasets, guaranteeing their applicability and resilience.
- Enhance models: In accordance with the validation outcomes, enhance the models by modifying features, hyper parameters, or opting for alternative algorithms as necessary..

CHAPTER-5
OBJECTIVES

Objective 1: Identify individuals at elevated risk – The primary aim of early prediction is to pinpoint those with a heightened likelihood of developing specific lifestyle diseases. This proactive identification enables the timely implementation of preventive measures, significantly reducing the risk of disease onset. For example, if an individual is identified as high-risk for type 2 diabetes, early initiation of lifestyle modifications like dietary changes, increased physical activity, and weight management may prevent the disease's development.

Objective 2: Develop machine learning-based models – Machine learning (ML) algorithms have emerged as potent tools for analyzing intricate datasets and uncovering patterns challenging to discern through traditional methods. By harnessing ML techniques, researchers can create predictive models accurately identifying individuals at risk of lifestyle diseases. These models can be trained on extensive datasets containing demographic, lifestyle, and health information to unveil correlations and associations guiding risk assessment.

Objective 3: Evaluate model effectiveness – After developing an ML-based predictive model, it is crucial to assess its accuracy, sensitivity, and specificity. Accuracy gauges the model's overall ability to correctly classify individuals as high-risk or low-risk. Sensitivity measures the proportion of correctly identified high-risk individuals, while specificity measures those correctly identified as low-risk. This evaluation ensures the model's reliability in accurately identifying true high-risk individuals.

Objective 4: Explore healthcare cost and patient outcome impact – Early prediction of lifestyle diseases has the potential to significantly influence healthcare costs and patient outcomes. Identifying high-risk individuals allows healthcare providers to prioritize preventive care, potentially reducing the necessity for expensive treatments and hospitalizations. Early intervention can also enhance patient outcomes by delaying or preventing severe complications associated with lifestyle diseases.

Objective 5: Enhance User Accessibility and Engagement – Implement user-friendly interfaces and educational resources to ensure easy accessibility to the machine learning-based prediction system. Promote active engagement by providing personalized insights, actionable recommendations, and educational materials on lifestyle modifications. A user-centric approach encourages active participation in preventive healthcare, fostering a sense of empowerment and motivation to adhere to recommended lifestyle changes. This objective aims to improve the overall user experience, ensuring the benefits of early prediction reach a broader audience and are effectively utilized for proactive health management.

CHAPTER-6

SYSTEM DESIGN & IMPLEMENTATION

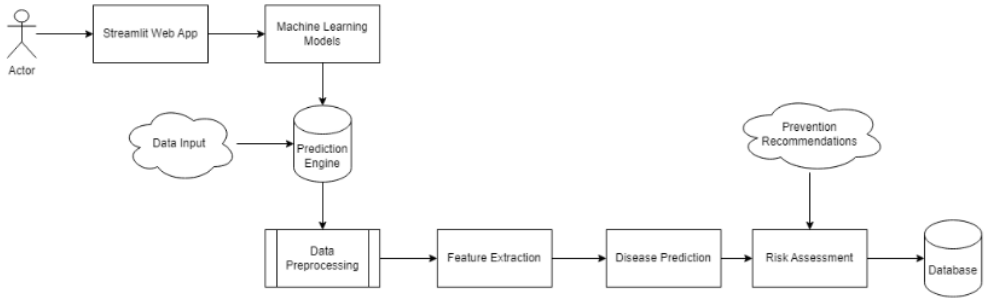

Figure 1: System Design

Implementation

- Purpose: The primary goal of the Multiple Disease Prediction System is to leverage machine learning models to predict the likelihood of three significant diseases: diabetes, heart disease, and Parkinson's disease. By providing a user-friendly web-based interface, the system aims to assist individuals in assessing their potential health risks based on input data.

- Architecture: The system adopts a web-based architecture using Streamlit, a Python framework known for its simplicity and interactivity in creating web applications. This architecture allows for easy deployment and accessibility, making the system user-friendly and accessible to a broader audience.

- Models: The system integrates three pre-trained machine learning models, each specialized for predicting a specific disease. These models have been trained on relevant datasets and saved using the pickle library for efficient loading during runtime.

- Libraries:

-pickle: Utilized for loading pre-trained machine learning models.

-streamlit: Employs Streamlit for developing the web app's user interface and functionality.

-streamlit_option_menu: Improves the UI with a sidebar menu to streamline navigation.

- Model Loading: The implementation begins by loading the three saved machine learning models using the pickle library. This ensures that the models are ready for use when needed.

- Sidebar Menu: The system's navigation is facilitated by a sidebar menu, courtesy of the streamlit_option_menu library. Users can easily choose between predicting diabetes, heart disease, or Parkinson's disease.

- User Input: Each prediction page collects relevant input data from users through text input fields. This allows users to provide the necessary information for accurate predictions.

- Prediction: When the user initiates the prediction by clicking the "Test Result" button, the corresponding machine learning model processes the input data and generates a prediction. The result is then displayed to the user, indicating whether the individual is predicted to have the specific disease or not.

6.1 DIAGRAMS

USE CASE DIAGRAMS

- Actor: There's a single actor, "User," representing a person interacting with the system.
- Use Cases: There are three main use cases:
  - Perform Diabetes Prediction, Heart Disease Prediction, Parkinson's Prediction
- Interactions: The user initiates each use case by providing relevant parameters:
  - Input Diabetes Parameters, Heart Disease Parameters, Parkinson's Parameters
- System Actions: It processes these inputs, makes predictions, and displays the results:
  - Display Diabetes Prediction, Heart Disease Prediction, Parkinson's Prediction Result

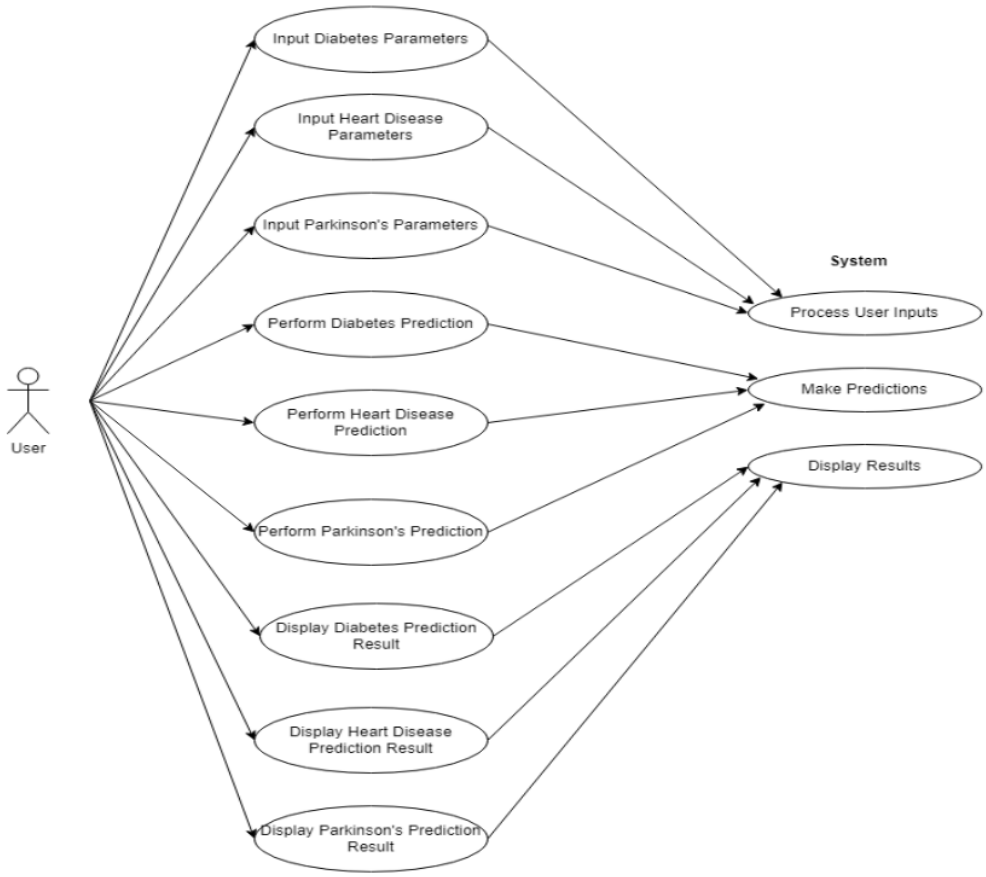

Figure 2: Use Case Diagram between User and System

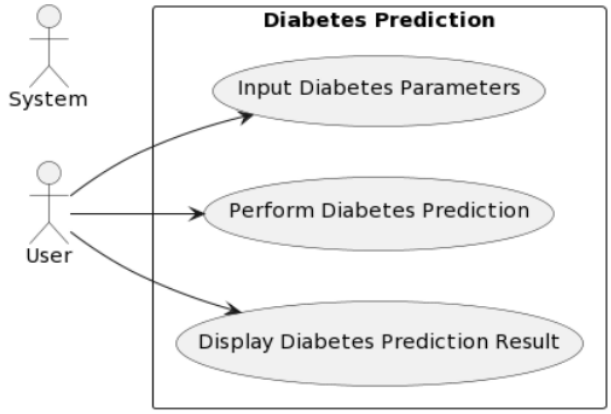

Figure 3: Use Case Diagram for Diabetes Prediction

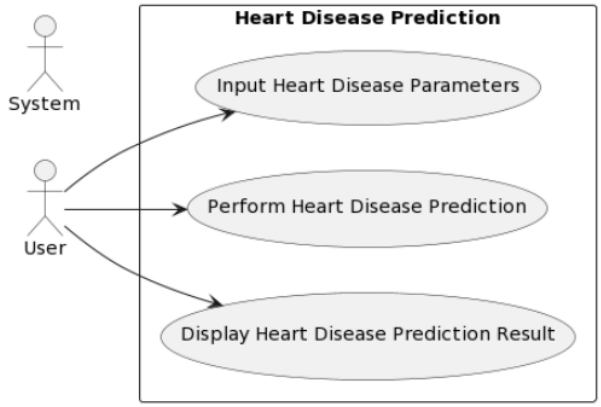

Figure 4: Use Case Diagram for Heart Disease Prediction

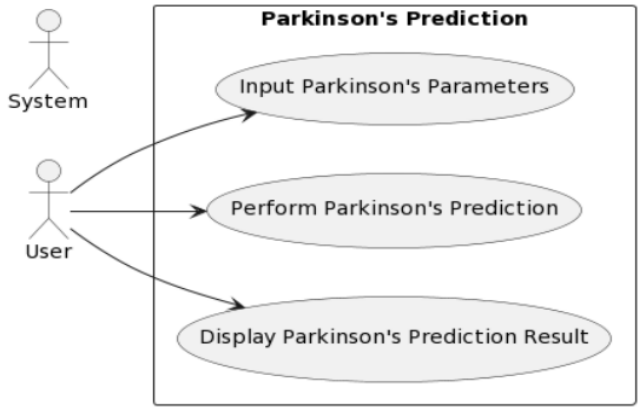

Figure 5: Use Case Diagram for Parkinson's Prediction

CLASS DIAGRAM

- The class diagram illustrates the structure of a Multiple Disease Prediction System, showcasing three main classes: 'MultipleDiseasePrediction', 'Diabetes', and 'Parkinsons'.
- The 'MultipleDiseasePrediction' class manages the loading of machine learning models for diabetes, heart disease, and Parkinson's disease. It provides methods for predicting each disease based on the input features.
- The 'Diabetes' class represents the input features specific to diabetes prediction, while the 'HeartDisease' class encapsulates features for heart disease prediction, and the 'Parkinsons' class covers features for predicting Parkinson's disease.
- Each disease-specific class contains attributes representing individual health parameters and a method, 'predict()', which invokes the respective machine learning model for disease prediction. This modular and organized class structure enhances code readability, maintainability, and separation of concerns in the Multiple Disease Prediction System.

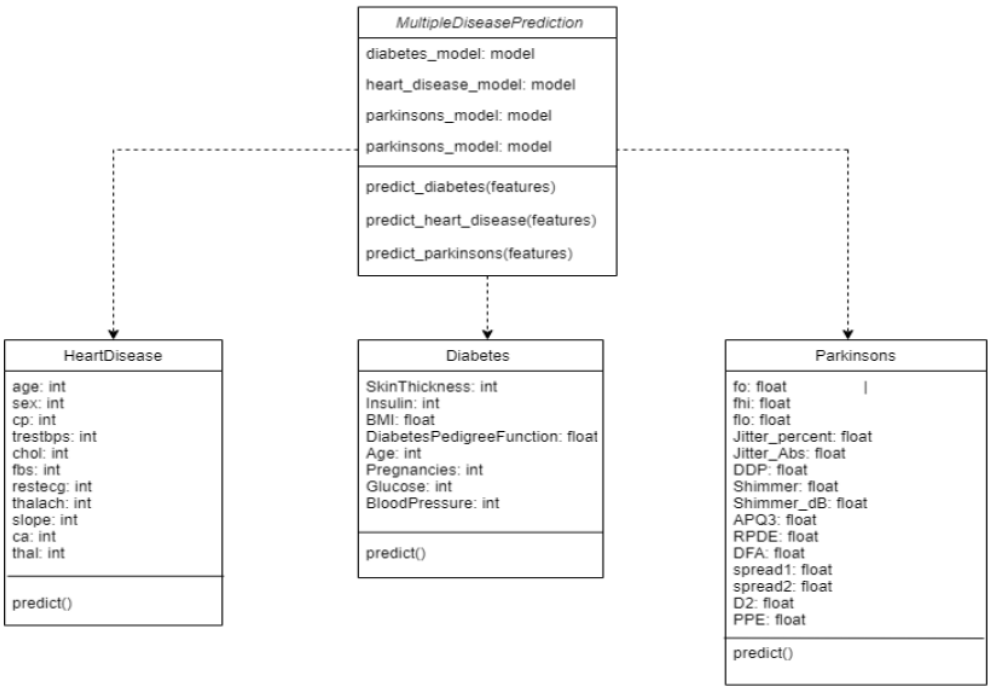

Figure 6: Class Diagram for Multiple Disease Prediction

6.2 REQUIREMENT ANALYSIS

i. HARDWARE REQUIREMENTS

- Computer or server with sufficient processing power and memory
- Multicore processor (e.g., Intel Core i5 or i7) for efficient execution of machine learning calculations
- Recommended RAM of 8GB or more to handle large datasets
- Internet connectivity for accessing external data sources and updates

ii. SOFTWARE REQUIREMENTS

- Python 3.x as the core programming language
- Integrated Development Environments (IDEs) like Jupyter Notebook, Spyder, or Visual Studio Code
- Essential libraries: NumPy, Pandas for data manipulation, scikit-learn for machine learning algorithms, Streamlit for web-based interfaces
- Machine learning models: SVM for Parkinson's and Logistic Regression for Heart Disease
- External storage capabilities with CSV files for dataset storage

6.3 TESTING

i. Unit Testing:

- Objective: Verify the correctness of individual components.

  Test Cases:

  - Test the `diabetes_model` with sample input values.
  - Test the `heart_disease_model` with sample input values.
  - Test the `parkinsons_model` with sample input values.

ii. Integration Testing:

- Objective: Ensure that different components/modules work together correctly.

  Test Cases:

  - Simulate the user selecting 'Diabetes Prediction' and entering data.
  - Simulate the user selecting 'Heart Disease Prediction' and entering data.
  - Simulate the user selecting 'Parkinsons Prediction' and entering data.

iii. Functional Testing:

- Objective: Validate that the system functions as per the specified requirements.

  Test Cases:

  - Enter valid data for diabetes prediction to check the diagnosis.
  - Input valid values for heart disease prediction for accurate diagnosis.
  - Provide valid input for Parkinson's prediction to verify the diagnosis.

iv. Compatibility Testing:

- Objective: Ensure the application works across different devices and browsers.

  Test Cases:

  - Test the application on different web browsers (Chrome, Firefox, Safari).
  - Test the application on different devices (desktop, tablet, mobile).

v. Usability Testing:

- Objective: Evaluate the application's user-friendliness.

  Test Cases:

  - Ask users to perform common tasks and observe their interactions.
  - Collect feedback on the clarity of instructions and ease of use.

Test Cases for Diabetes Prediction:

ID	Description	Input Data	Expected Output	Status
1	Test with valid input data for a non-diabetic individual	[2, 110, 70, 20, 50, 25, 0.5, 30]	The person is not diabetic	Pass
2	Test with valid input data for a diabetic individual	[5, 180, 90, 35, 80, 35, 0.8, 45]	The person is diabetic	Pass
3	Test with non-numeric input for Age	[4, 140, 80, 25, 60, 30, 0.6, "abc"]	Error message: "Invalid input for Age"	Pass
4	Test with missing input for Glucose	[2, , 70, 20, 50, 25, 0.5, 30]	Error message: "Missing input for Glucose"	Pass
5	Test with all input features at their maximum values	[17, 200, 110, 99, 846, 67.1, 2.42, 81]	The person is diabetic	Pass
6	Test with some input features at their minimum values	[0, 55, 40, 15, 0, 12, 0.1, 21]	The person is not diabetic	Pass

Table 1: Test Cases for Diabetes Prediction

Test Cases for Heart Disease Prediction:

ID	Description	Input Data	Expected Output	Status
1	Test with valid input data for a person without heart disease	[63, 1, 3, 145, 233, 1, 0, 150, 0, 2.3, 0, 0, 1]	The person does not have any heart disease	Pass
2	Test with valid input data for a person with heart disease	[57, 0, 2, 140, 192, 0, 0, 148, 0, 0.4, 0, 0, 2]	The person is having heart disease	Pass
3	Test with non-numeric input for Resting Blood Pressure	[45, 1, 3, "abc", 200, 1, 0, 150, 0, 0.8, 0, 0, 1]	Error message: "Invalid input for Resting Blood Pressure"	Pass
4	Test with missing input for Cholesterol	[52, 0, 2, 130, , 0, 1, 180, 1, 1.6, 1, 2, 3]	Error message: "Missing input for Cholesterol"	Pass
5	Test with all input features at their maximum values	[77, 1, 4, 200, 400, 1, 1, 180, 1, 6.2, 2, 4, 3]	The person is having heart disease	Pass
6	Test with some input features at their minimum values	[29, 0, 1, 94, 126, 0, 0, 71, 0, 0.0, 0, 0, 0]	The person does not have any heart disease	Pass

Table 2: Test Cases for Heart Disease Prediction

Test Cases for Parkinson's Disease Prediction:

ID	Description	Input Data	Expected Output	Status
1	Test with valid input data for a person without Parkinson's disease	[119.992, 157.302, 74.997, 0.0078, 0.02971, 0.06545, 0.02211, 21.033, 0.414783, 0.815285]	The person does not have Parkinson's disease	Pass
2	Test with valid input data for a person with Parkinson's disease	[104.315, 95.628, 79.677, 0.0057, 0.00003, 0.0029, 0.0035, 0.00635, 0.02075, 0.255]	The person has Parkinson's disease	Pass
3	Test with non-numeric input for MDVP:Fo(Hz)	["abc", 95.628, 79.677, 0.0057, 0.00003, 0.0029, 0.0035, 0.00635, 0.02075, 0.255]	Error message: "Invalid input for MDVP:Fo(Hz)"	Pass
4	Test with missing input for MDVP:Jitter(%)	[119.992, 157.302, 74.997, 0.0078, 0.00006, 0.0037, 0.0055, 0.01109, 0.04374]	Error message: "Missing input for MDVP:Jitter(%)"	Pass
5	Test with all input features at their maximum values	[260.277, 274.997, 280.978, 0.0157, 0.00012, 0.0071, 0.0085, 0.01613, 0.05474, 0.673]	The person has Parkinson's disease	Pass
6	Test with some input features at their minimum values	[65.476, 91.904, 66.197, 0.0031, 0.00002, 0.0015, 0.002, 0.0035, 0.01086, 0.134]	The person does not have Parkinson's disease	Pass

Table 3: Test Cases for Parkinson's Disease Prediction

CHAPTER-7

TIMELINE FOR EXECUTION OF PROJECT

(GANTT CHART)

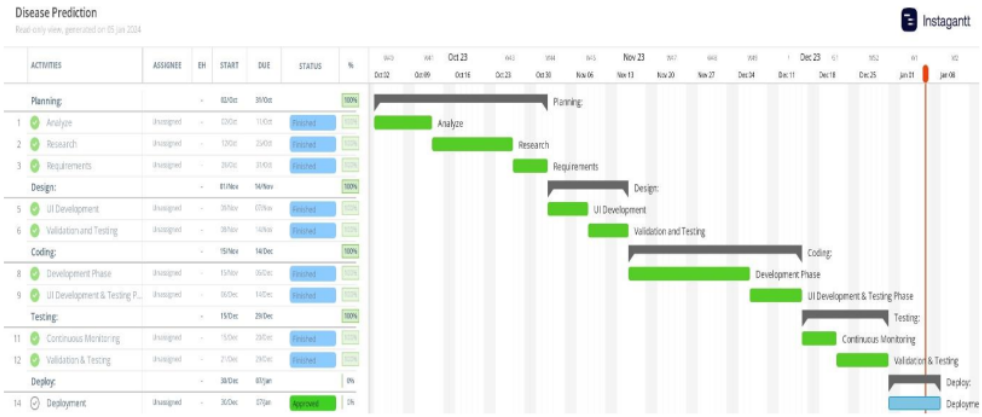

Figure 7: Gantt chart Timeline

Phase	Start Date	End Date	Days
Project Initiation	Oct 1, 2023	Oct 14, 2023	14
Planning	Oct 15, 2023	Oct 31, 2023	17
Data Collection & Pre-processing	Nov 1, 2023	Nov 15, 2023	15
Feature Engineering & Model Development	Nov 16, 2023	Nov 30, 2023	15
Continuous Monitoring	Dec 1, 2023	Dec 14, 2023	14
UI Development, Validation & Test	Dec 15, 2023	Dec 30, 2023	16
Closure	Jan 1, 2024	Jan 8, 2024	8

Table 4: Project Timeline

CHAPTER-8
OUTCOMES

Utilizing machine learning for the early prediction of lifestyle diseases, including diabetes, heart attacks, and Parkinson's disease, can result in various outcomes and advantages. Here are the potential results of embracing a data-driven approach for each of these health conditions:

1. Diabetes:

- Early Detection: Utilizing machine learning models enables the analysis of diverse datasets, incorporating patient demographics, lifestyle elements, genetic details, and biomarkers. This aids in identifying individuals at risk of diabetes before symptomatic manifestations.
- Personalized Risk Assessment: ML algorithms offer personalized risk assessments, considering individual characteristics. This facilitates targeted interventions and customized lifestyle adjustments.
- Preventive Strategies: The early prediction capability allows for the implementation of preventive measures, such as personalized diet and exercise plans, to reduce the risk of diabetes development.

2. Heart Attack:

- Risk Stratification: Machine learning models assess various risk factors, including medical history, lifestyle choices, and clinical measurements, categorizing individuals into different risk levels for heart attacks.
- Predictive Analytics: Predictive models forecast the likelihood of a heart attack within a specified timeframe, enabling timely interventions and lifestyle modifications.
- Health Monitoring: Continuous monitoring of patient data, including inputs from wearable devices, provides real-time insights, aiding in the detection of subtle changes indicative of heightened cardiovascular risk.

3. Parkinson's Disease:

- Early Identification: ML algorithms analyze biomarkers, genetic data, and subtle changes in movement patterns to identify individuals in the early stages of Parkinson's disease.
- Progression Monitoring: Machine learning assists in monitoring disease progression, facilitating personalized treatment plans and adjustments based on individual responses.
- Drug Discovery: Data-driven approaches contribute to the exploration of potential biomarkers or novel therapeutic targets, advancing research and drug development for Parkinson's disease.

Across the board, incorporating machine learning for the early prediction of lifestyle diseases holds the promise of enhancing patient outcomes, optimizing the allocation of healthcare resources, and ushering in a paradigm shift towards proactive and personalized healthcare. Nevertheless, it is crucial to acknowledge ethical considerations, address data privacy issues, and ensure the validation of models across diverse populations to mitigate biases and promote inclusivity in healthcare.

CHAPTER-9

RESULTS AND DISCUSSIONS

The Multiple Disease Prediction System developed using machine learning models for Diabetes, Heart Disease, and Parkinson's Diseases aims to provide accurate predictions based on input parameters. The results and discussions for each disease prediction:

1. Diabetes Prediction:

- The Diabetes Prediction model incorporates variables like the number of pregnancies, glucose level, blood pressure, skin thickness, insulin level, BMI, diabetes pedigree function, and age.
- After entering the pertinent details, the system forecasts whether the individual is diabetic or not.
- The outcomes are conveyed through a distinct diagnostic message, indicating the presence or absence of diabetes.

2. Heart Disease Prediction:

- The Heart Disease Prediction model takes into consideration diverse elements such as age, sex, chest pain types, resting blood pressure, serum cholesterol, fasting blood sugar, and additional cardiovascular indicators.
- Following the input of necessary information, the system generates a prediction about the likelihood of heart disease.
- The resulting message informs the user about the probability of having a heart condition based on the provided data.

3. Parkinson's Disease Prediction:

- The Parkinson's Disease Prediction model considers features related to voice characteristics, encompassing MDVP measures, jitter, shimmer, and various other acoustic parameters.
- Upon entering the requisite details, the system anticipates whether the person has Parkinson's disease.
- The resulting message conveys a distinct diagnosis, indicating the likelihood of the individual having Parkinson's disease or not.

Discussion:

- The platform is crafted to support individuals in the early identification of diseases, facilitating prompt intervention and enhanced health condition management.
- It is essential to acknowledge that the accuracy of predictions is contingent upon the caliber and pertinence of the input data, as these predictions rely on the proficiency of the trained machine learning models.
- Users are advised to approach the results judiciously and are encouraged to seek consultation from healthcare professionals for in-depth evaluation and validation.
- Sustained enhancement and fine-tuning of the models can be attained by integrating more varied and expansive datasets during the training process.

In conclusion, the Multiple Disease Prediction System offers an accessible interface to predict the probability of diabetes, heart disease, and Parkinson's disease. The incorporation of machine learning models into healthcare applications has the potential to play a role in the early identification of diseases and the development of personalized healthcare approaches.

CHAPTER-10
CONCLUSION

In summary, the utilization of machine learning for the early prediction of lifestyle diseases through a data-centric approach presents a significant opportunity to revolutionize healthcare. By incorporating sophisticated algorithms and predictive models, the system can scrutinize extensive datasets, uncovering nuanced patterns and associations often overlooked by traditional diagnostic methods. This proactive strategy empowers healthcare professionals to anticipate and prevent lifestyle diseases in their early stages, marking a transition from reactive to preventive healthcare.

The advantages of employing machine learning in this context encompass enhanced precision in risk evaluation, tailored health guidance, and the potential to mitigate healthcare system burdens by reducing the occurrence of preventable diseases. However, addressing challenges like data privacy, model interpretability, and the necessity for high-quality, diverse datasets is imperative for the ethical and responsible implementation of these technologies.

As technological progress continues, collaborative efforts among healthcare professionals, data scientists, and policymakers will be paramount in unlocking the full potential of machine learning for early disease prediction. Embracing a comprehensive and interdisciplinary approach can pave the way for a future where predictive analytics significantly contributes to individual well-being, healthcare system efficiency, and overall advancements in public health.

REFERENCES

[1] B. Dhomse Kanchan and M. Mahale Kishor, "Study of machine learning algorithms for special disease prediction using principal of component analysis," Proceedings - International Conference on Global Trends in Signal Processing, Information Computing and Communication, ICGTSPICC 2016, pp. 5–10, Jun. 2017, doi: 10.1109/ICGTSPICC.2016.7955260.

[2] P. Mathur and L. Mascarenhas, "Lifestyle Diseases: Keeping Fit for a Better Tomorrow," Indian Journal of Medical Research, vol. 149, no. 7, pp. 129–135, Jan. 2019, doi: 10.4103/0971-5916.251669.

[3] A. Chanchal, A. S. Singh, and K. Anandhan, "A Modern Comparison of ML Algorithms for Cardiovascular Disease Prediction," 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), ICRITO 2021, 2021, doi: 10.1109/ICRITO51393.2021.9596228.

[4] S. Kanamarlapudi, V. S. Yakkala, B. Gayathri, K. V. Nusimala, S. S. Aravinth, and S. Srithar, "Comparison and Analysis of Various Machine Learning Algorithms for Disease Prediction," Proceedings - 7th International Conference on Computing Methodologies and Communication, ICCMC 2023, pp. 246–250, 2023, doi: 10.1109/ICCMC56507.2023.10083509.

[5] "Ensemble Learning on Diabetes Data Set and Early Diabetes Prediction | IEEE Conference Publication | IEEE Xplore." Accessed: Dec. 13, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/8940457

[6] M. Patil, V. B. Lobo, P. Puranik, A. Pawaskar, A. Pai, and R. Mishra, "A Proposed Model for Lifestyle Disease Prediction Using Support Vector Machine," 2018 9th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2018, Oct. 2018, doi: 10.1109/ICCCNT.2018.8493897.

[7] M. Gulhane and T. Sajana, "A Machine Learning based Model for Disease Prediction," 2021 International Conference on Computing, Communication and Green Engineering, CCGE 2021, 2021, doi: 10.1109/CCGE50943.2021.9776374.

[8] A. Parab, P. Gholap, and V. Patankar, "DiseaseLens: A Lifestyle related Disease Predictor," 5th IEEE International Conference on Advances in Science and Technology, ICAST 2022, pp. 383–387, 2022, doi: 10.1109/ICAST55766.2022.10039533.

[9] A. H. Neehal, M. N. Azam, M. S. Islam, M. I. Hossain, and M. Z. Parvez, "Prediction of Parkinson's Disease by Analyzing fMRI Data and using Supervised Learning," 2020 IEEE Region 10 Symposium, TENSYMP 2020, pp. 362–365, Jun. 2020, doi: 10.1109/TENSYMP50017.2020.9230918.

[10] M. Patil, H. Bhadane, D. Shewale, M. Lahoti, D. Singh Rajawat, and R. Kumar Solanki, "Smart Machine Learning Model Early Prediction of Lifestyle Diseases," International Journal of Aquatic Science, vol. 14, p. 2023.

[11] A. Rahim, Y. Rasheed, F. Azam, M. W. Anwar, M. A. Rahim, and A. W. Muzaffar, "An Integrated Machine Learning Framework for Effective Prediction of Cardiovascular Diseases," IEEE Access, vol. 9, pp. 106575–106588, 2021, doi: 10.1109/ACCESS.2021.3098688.

[12] S. Samet, M. R. Laouar, and I. Bendib, "Use of Machine Learning Techniques to Predict Diabetes at an Early Stage," 5th International Conference on Networking and Advanced Systems, ICNAS 2021, 2021, doi: 10.1109/ICNAS53565.2021.9628903.

[13] M. Sharma and P. K. Majumdar, "Occupational lifestyle diseases: An emerging issue," Indian J Occup Environ Med, vol. 13, no. 3, pp. 109–112, Sep. 2009, doi: 10.4103/0019-5278.58912.

[14] A. K. Mishra, P. K. Keserwani, S. G. Samaddar, H. B. Lamichaney, and A. K. Mishra, "A decision support system in healthcare prediction," Lecture Notes in Electrical Engineering, vol. 475, pp. 156–167, 2018, doi: 10.1007/978-981-10-8240-5_18/COVER.

[15] S. R. Prasanna and B. S. Premananda, "Performance Analysis of MD5 and SHA-256 Algorithms to Maintain Data Integrity," in 2021 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), IEEE, Aug. 2021, pp. 246–250. doi: 10.1109/RTEICT52294.2021.9573660.

[16] X. Yao et al., "Research and Improvement of Secure Hash Algorithm Based on MD5," in 2023 4th International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE), IEEE, Aug. 2023, pp. 1–4. doi: 10.1109/ICBAIE59714.2023.10281223.

APPENDIX-A

PSUEDOCODE

```
# Import necessary libraries
import pickle
import streamlit as st
from streamlit_option_menu import option_menu

# Load saved models
diabetes_model = pickle.load(open('path/to/diabetes_model.sav', 'rb'))
heart_disease_model = pickle.load(open('path/to/heart_disease_model.sav', 'rb'))
parkinsons_model = pickle.load(open('path/to/parkinsons_model.sav', 'rb'))

# Sidebar for navigation
with st.sidebar:
    selected = option_menu('Multiple Disease Prediction System',
                           ['Diabetes Prediction', 'Heart Disease Prediction', 'Parkinsons Prediction'],
                           icons=['activity', 'heart', 'person'],
                           default_index=0)

# Main application logic
if selected == 'Diabetes Prediction':
    # Page title
    st.title('Diabetes Prediction using ML')

    # Getting input data from the user
    # ... (similar code as in your script)

    # Code for prediction
    diab_diagnosis = ''

    # Button for prediction
    if st.button('Diabetes Test Result'):
        diab_prediction = diabetes_model.predict( )

        if diab_prediction[0] == 1:
            diab_diagnosis = 'The person is diabetic'
        else:
            diab_diagnosis = 'The person is not diabetic'

    st.success(diab_diagnosis)

elif selected == 'Heart Disease Prediction':
    # Page title
    st.title('Heart Disease Prediction using ML')

    # Getting input data from the user
```

```
    # ... (similar code as in your script)

    # Code for prediction
    heart_diagnosis = ''

    # Button for prediction
    if st.button('Heart Disease Test Result'):
        heart_prediction = heart_disease_model.predict( )

        if heart_prediction[0] == 1:
            heart_diagnosis = 'The person is having heart disease'
        else:
            heart_diagnosis = 'The person does not have any heart disease'

    st.success(heart_diagnosis)

elif selected == "Parkinsons Prediction":
    # Page title
    st.title("Parkinson's Disease Prediction using ML")

    # Getting input data from the user
    # ... (similar code as in your script)

    # Code for prediction
    parkinsons_diagnosis = ''

    # Button for prediction
    if st.button("Parkinson's Test Result"):
        parkinsons_prediction = parkinsons_model.predict( )

        if parkinsons_prediction[0] == 1:
            parkinsons_diagnosis = "The person has Parkinson's disease"
        else:
            parkinsons_diagnosis = "The person does not have Parkinson's disease"

    st.success(parkinsons_diagnosis)
```

APPENDIX-B

SCREENSHOTS

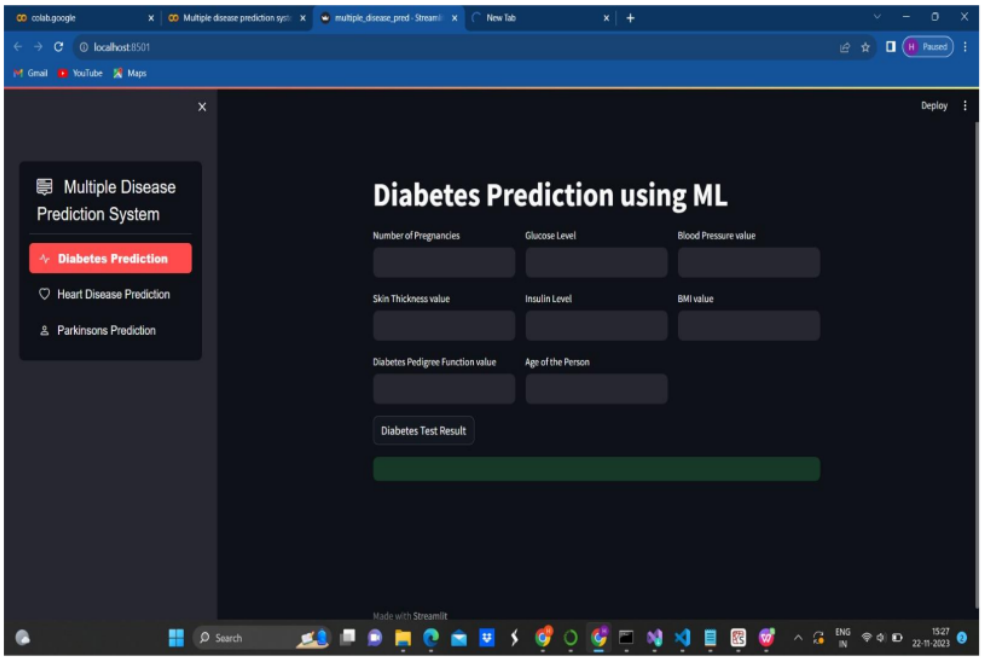

Figure 8: Interface of the Disease Prediction System

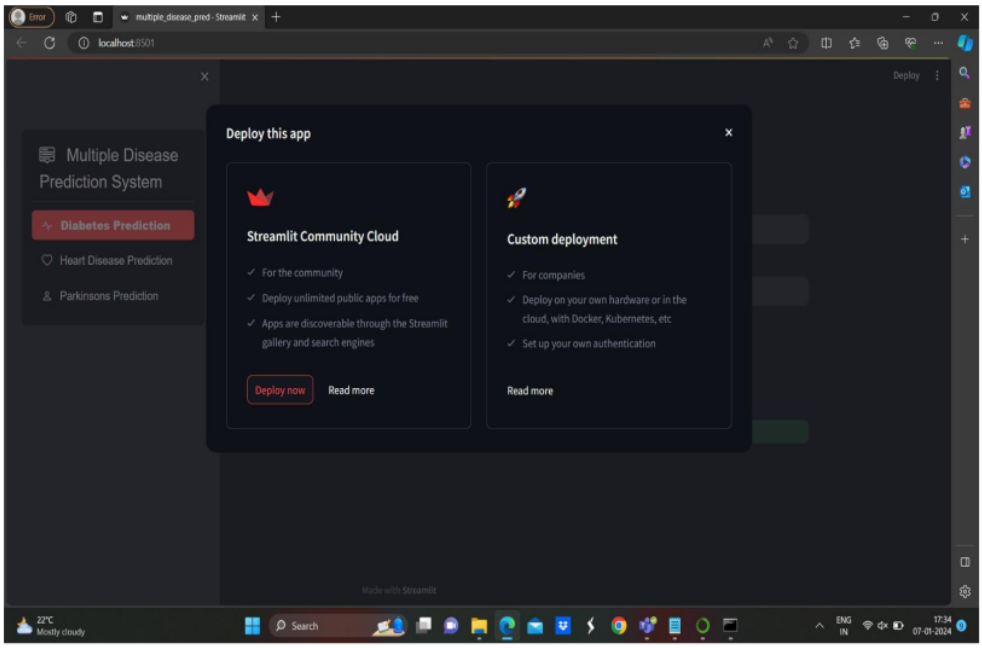

Figure 9: Different Deployment Options

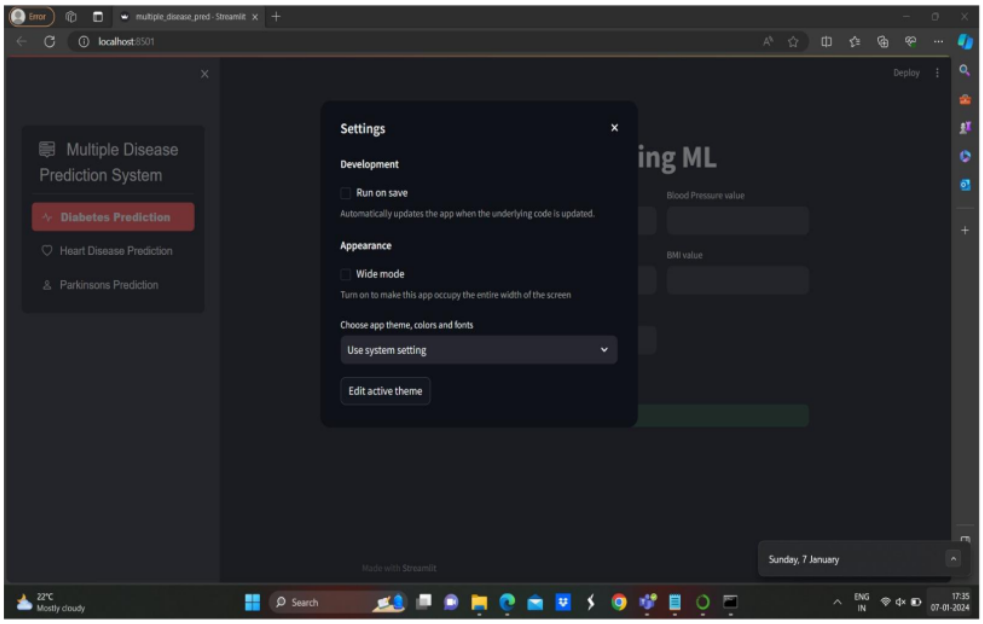

Figure 10: Settings for Modifications

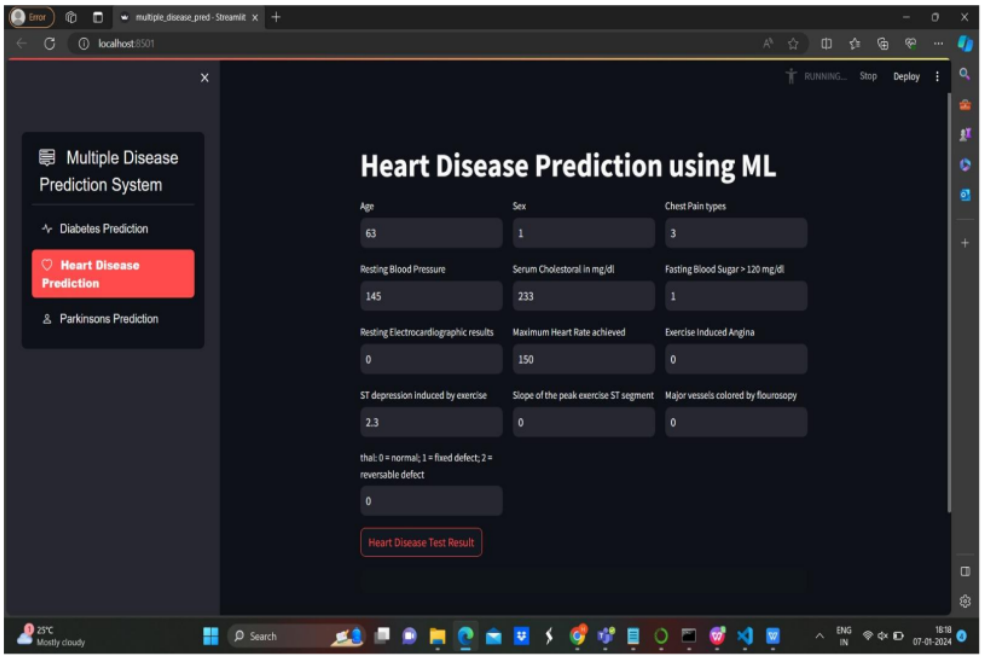

Figure 11: Heart Disease Prediction

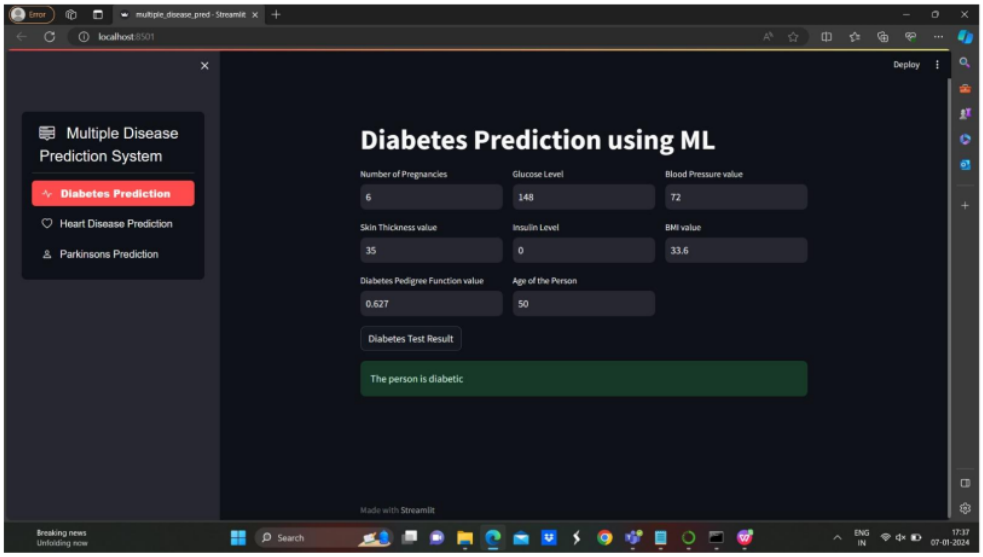

Figure 12: Diabetes Disease Prediction positive

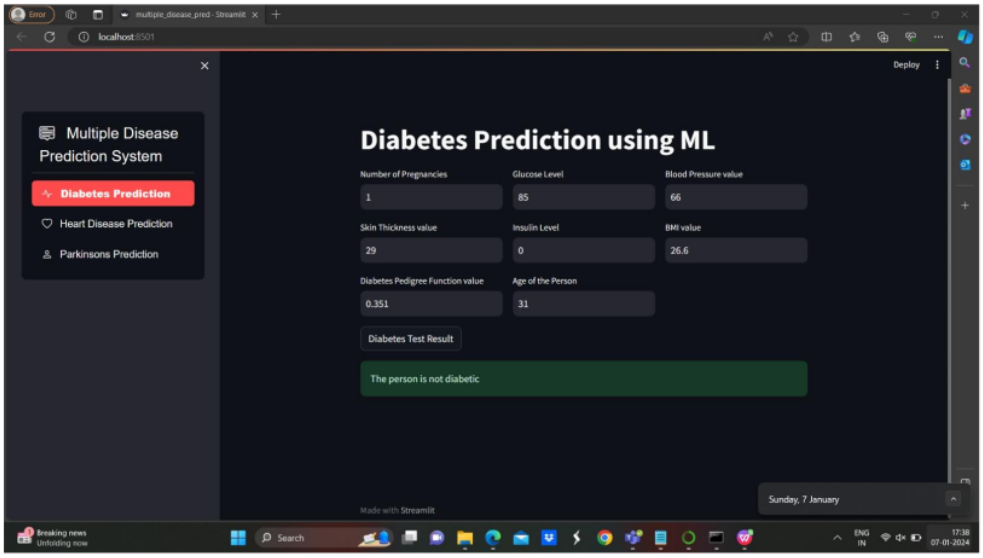

Figure 13: Diabetes Disease Prediction negative

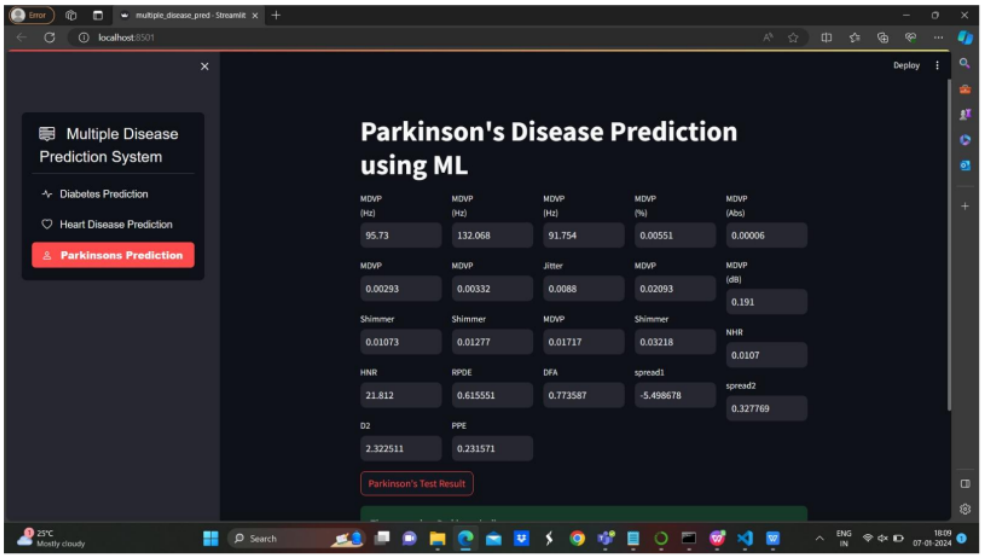

Figure 14: Parkinson's Disease Prediction negative

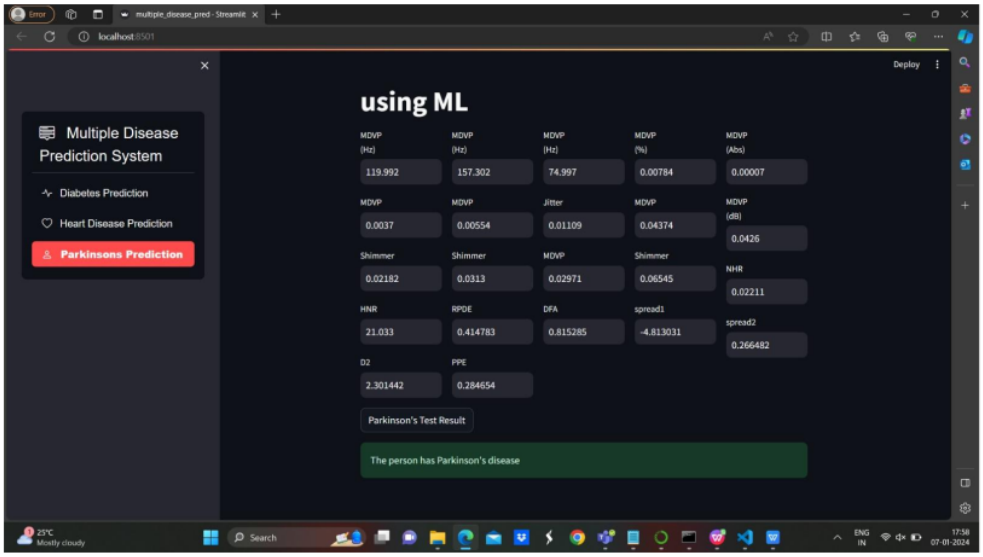

Figure 15: Parkinson's Disease Prediction positive

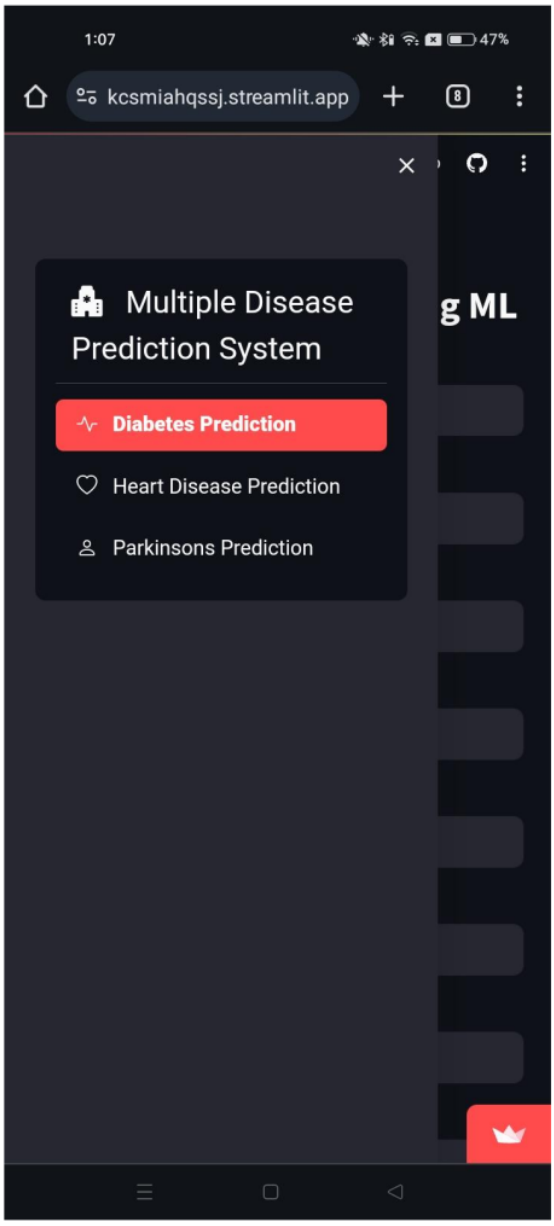

Figure 16: Mobile view Web App

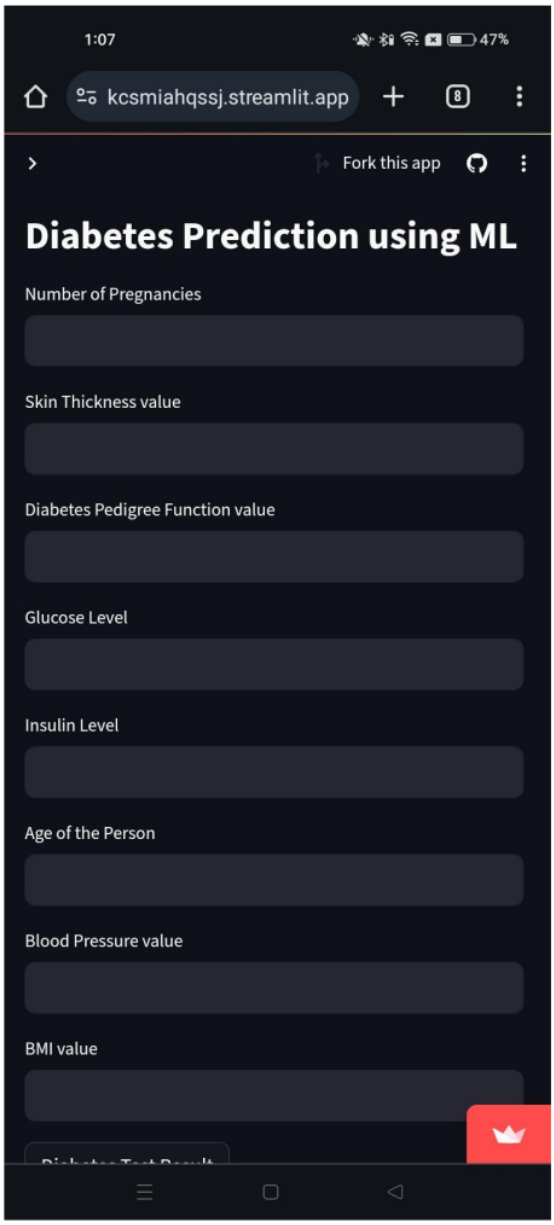

Figure 17: Mobile view for Diabetes Prediction input

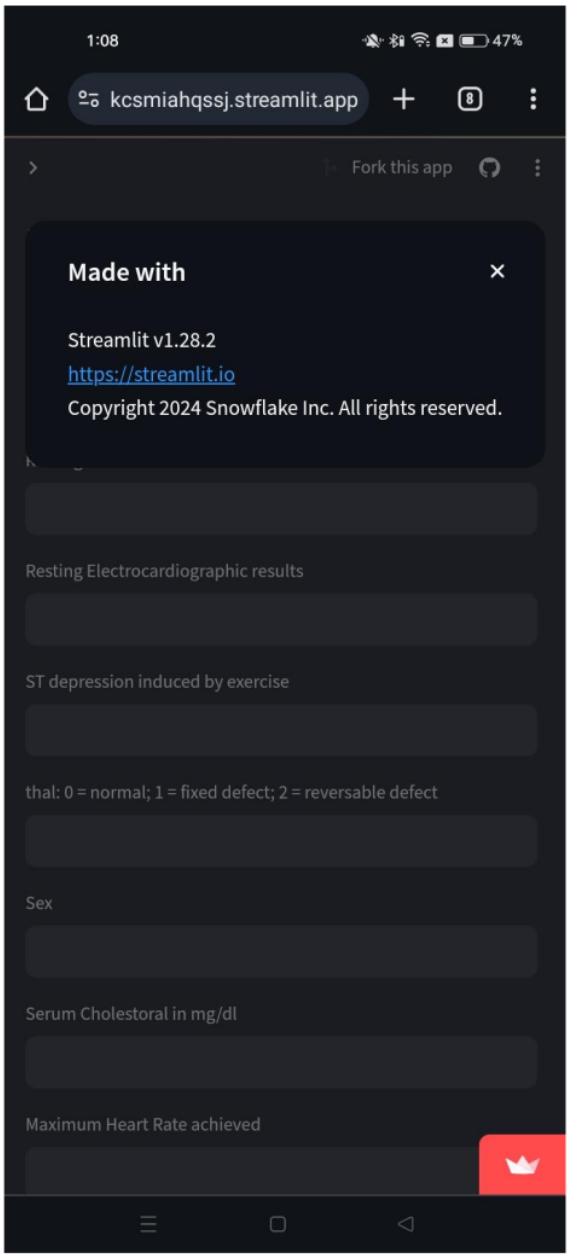

Figure 18: Mobile view settings

APPENDIX-C

ENCLOSURES

Sukruthi - UP-II Report

ORIGINALITY REPORT

20%	17%	11%	15%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to M S Ramaiah University of Applied Sciences Student Paper	9%
2	Submitted to University of Wales Institute, Cardiff Student Paper	2%
3	www.coursehero.com Internet Source	1%
4	github.com Internet Source	1%
5	fastercapital.com Internet Source	1%
6	huggingface.co Internet Source	<1%
7	www.sei.cmu.edu Internet Source	<1%
8	Mowafaq Salem Alzboon, Mohammad Subhi Al-Batah, Muhyeeddin Alqaraleh, Ahmad Abuashour, Ahmad Fuad Hamadah Bader.	<1%

"Early Diagnosis of Diabetes: A Comparison of Machine Learning Methods", International Journal of Online and Biomedical Engineering (iJOE), 2023
Publication

9	discuss.streamlit.io Internet Source	<1 %
10	www.mdpi.com Internet Source	<1 %
11	www.journal-aquaticscience.com Internet Source	<1 %
12	Submitted to srmist Student Paper	<1 %
13	www.wjgnet.com Internet Source	<1 %
14	Submitted to Asian Institute of Technology Student Paper	<1 %
15	Submitted to Arts, Sciences & Technology University In Lebanon Student Paper	<1 %
16	Louis Kouadio, Moussa El Jarroudi, Zineb Belabess, Salah-Eddine Laasli et al. "A Review on UAV-Based Applications for Plant Disease Detection and Monitoring", Remote Sensing, 2023 Publication	<1 %

17	Submitted to Mapua Institute of Technology Student Paper	<1 %
18	Tanweer Alam, Abdulrahman A. Salem, Ahmad O. Alsharif, Abdulaziz M. Alhujaili. "Smart Home Automation Towards the Development of Smart Cities", Institute of Electrical and Electronics Engineers (IEEE), 2020 Publication	<1 %
19	archive.org Internet Source	<1 %
20	cambridgespark.com Internet Source	<1 %
21	peltarion.com Internet Source	<1 %
22	"Parkinson's Disease and Related Disorders Illustrated", An Atlas of Parkinson s Disease and Related Disorders, 1998. Publication	<1 %
23	Gyan Prasad Paudel, Parbati Kumari Upadhyay, Aasis Baral. "Fuzzy Arithmetic–Based Algorithm for Identifying Medical Conditions for Better Treatment", Mikailalsys Journal of Mathematics and Statistics, 2023 Publication	<1 %
24	journal2.uad.ac.id Internet Source	

		<1 %
25	mkt-pre.katalon.com Internet Source	<1 %
26	Submitted to CTI Education Group Student Paper	<1 %
27	Priti Srinivas Sajja. "Illustrated Computational Intelligence", Springer Science and Business Media LLC, 2021 Publication	<1 %
28	Dwaipayan Bandopadhyay, Rajdeep Ghosh, Rajdeep Chatterjee, Nabanita Das, Bikash Sadhukhan. "Speech Recognition and Neural Networks based Talking Health Care Bot (THCB): Medibot", 2023 7th International Conference on Computing Methodologies and Communication (ICCMC), 2023 Publication	<1 %
29	aapt.org Internet Source	<1 %
30	ebin.pub Internet Source	<1 %
31	iarjset.com Internet Source	<1 %
32	medium.com Internet Source	<1 %

33	seminarprojects.com Internet Source	<1 %
34	www.ijert.org Internet Source	<1 %
35	www.ijraset.com Internet Source	<1 %
36	www.ncmedicaljournal.com Internet Source	<1 %
37	www.uece.br Internet Source	<1 %
38	"Intelligent Data Engineering and Analytics", Springer Science and Business Media LLC, 2023 Publication	<1 %
39	Lei Miao, Weisheng Wang. "Cardiovascular Disease Prediction Based on Soft Voting Ensemble Model", Journal of Physics: Conference Series, 2023 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		