JSS Mahavidyapeetha

# JSS COLLEGE OF ARTS, COMMERCE AND SCIENCE

**Affiliated and Autonomous Under University of Mysore**

## Ooty Road, Mysuru-570025

A Project Report on

# "MULTIPLE DISEASE PREDICTION"

Submitted in partial fulfilment of the requirement
For the course degree in
**BACHELOR OF COMPUTER APPLICATION**

**Submitted by**

**RIPUNJAY RAJ**
**[U01BE22S0116]**

**Under the guidance of**

**Ms MEGHA S**
**Assistant Professor**
**Department of Computer Science**
**JSS College of Arts, Commerce And Science, Mysuru 2024-25**

# JSS COLLEGE OF ARTS COMMERCE AND SCIENCE

**(Autonomous)**

BN Road, Mysuru - 570025

## CERTIFICATE

This is to certify that the project work entitled "**MULTIPLE DISEASE PREDICTION"** carried out by **RIPUNJAY RAJ** bearing register number **U01BE21S0116** has successfully completed the project work in VI semester by partial fulfilment of the award of **Bachelor of Computer Application** by JSS College of Arts, Commerce and science (Autonomous) Mysore under University of Mysore, During the academic year 2024-25 The work has been approved as it satisfies the academic requirements in respect of project work prescribed for the final semester BCA.

Signature of the Guide

Signature of the HOD

Name of the Candidate: RIPUNJAY RAJ

Reg No : U01BE22S0116

Examination Centre : JSS CACS

Date of Examination :

ISSN: 2582-3930

# ACCEPTANCE CERTIFICATE

Impact Factor: 8.586
DOI Prefix: 10.55041

**INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT (IJSREM)**

An Open Access Scholarly Journal || Index in major Databases & Metadata

* * *

We are pleased to inform you that your manuscript titled

## MULTIPLE DISEASE PREDICTION USING MACHINE LEARNING

has been ACCEPTED for presentation and publication in the proceedings of Special Edition, to be published in

**INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT (IJSREM)**

*VOLUME 09 ISSUE 05 MAY - 2025 (SPECIAL EDITION)*

*We are delighted to see your commitment & hardwork to share your research is being recognized. We look foreward to helping you with all of your publication needs. Thank you for choosing IJSREM!!*

Editor-in-chief
IJSREM

www.ijsrem.com
e-mail: editor@ijsrem.com

DOI: 10.55041/IJSREM.SPEJSS007

ISSN: 2582-3930

Impact Factor: 8.586

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT

An Open Access Scholarly Journal || Index in major Databases & Metadata

# CERTIFICATE OF PUBLICATION

International Journal of Scientific Research in Engineering & Management is *hereby* awarding this certificate to

**Ripunjay Raj**

*in recognization to the publication of paper titled*

**Multiple Disease Prediction Using Machine Learning**

*published* in IJSREM Journal on Special Edition - Volume 09 SPEJSS -2025

Editor-in-Chief
IJSREM Journal

www.ijsrem.com

e-mail: editor@ijsrem.com

e-ISSN: 2395-0056 p-ISSN: 2395-0072

# International Research Journal of Engineering and Technology (IRJET)

( An ISO 9001 : 2008 Certified Journal )

*Is hereby awarding this certificate to*

## *Ripunjay Raj*

*In recognition the publication of the manuscript entitled*

## Supervised Learning Techniques for Multi-Disease Detection

*published in our Journal Volume 11 Issue 5 May* **2025**

Editor in Chief

E-mail : editor@irjet.net

Impact Factor : 8.226

www.irjet.net

# DECLARATION

I am **RIPUNJAY RAJ** hereby declare that this term work entitled **“MULTIPLE DISEASE PREDICTION**“ submitted to JSS College of Arts, Commerce and Science, Ooty road Mysore (Affiliated to University of Mysore) is a record of original work done by me and my project partners under the supervision and guidance **Ms. Megha S,** Assistant Professor Department of Computer Science.

Place: Mysuru                                   Signature

Date:                                   **(RIPUNJAY RAJ)**

# ACKNOWLEDGEMENT

I would like to express high regard to our college **JSS COLLEGE OF ARTS, COMMERCE AND SCIENCE, MYSORE** for grooming us all these years. The support given by respected Chief Executive **Prof. B.V. Sambashivaiah** and Principal **M. Prabhu** is highly memorable.

Our sincere thanks to **Mr. VIDYASHANKARA SIR,** head of the department, Dept of Computer Science, JSS College, Mysore for her valuable support during the project.

We wish to express deep sense of gratitude and affection to our internal guide **Ms. Megha S,** Assistant Professor, Department of Computer Science, JSS College, Mysore for her encouragement as well as supervision till the completion of our project.

I also Thank all teaching and non-teaching staff of Department of computer science, who has directly or indirectly extended their timely help

Finally, we are deeply grateful to faculty members of CS Department for their Guidance and direction.

**RIPUNJAY RAJ**
**[U01BE21S0116]**

# ABSTRACT

This project presents a comprehensive web-based health assistant application developed using streamlit, designed to facilitate the prediction of various diseases through machine learning models. The application encompasses five key prediction modules: diabetes, heart disease, parkinson's disease, kidney disease, and breast cancer. Each module allows users to input relevant health parameters, which are then processed by pre-trained models to generate diagnostic predictions.

The user-friendly interface is structured with a sidebar for easy navigation between different disease prediction pages. Each prediction page includes input fields for specific health metrics, ensuring that users can provide the necessary data for accurate assessments. Upon submission, the application utilizes the respective machine learning model to analyze the input and deliver a diagnosis, indicating whether the individual is likely to have the condition in question.

This health assistant not only empowers users with immediate insights into their health status but also serves as a valuable educational tool, raising awareness about the symptoms and risk factors associated with these diseases. The integration of machine learning into healthcare applications exemplifies the potential for technology to enhance patient outcomes and promote proactive health management.

**INDEX** **PAGE NO**

# CHAPTER 1
# INTRODUCTION

## 1.1 Background

The integration of technology in healthcare has revolutionized the way medical diagnoses are conducted. With the advent of machine learning (ML), healthcare professionals can leverage data-driven insights to enhance patient care. Disease prediction systems have emerged as vital tools, enabling early detection and intervention, which can significantly improve patient outcomes. This project focuses on developing a Health Assistant application that utilizes machine learning models to predict various diseases, including diabetes, heart disease, Parkinson's disease, kidney disease, and breast cancer. By providing users with immediate diagnostic insights, the application aims to empower individuals to take proactive steps towards their health.

## 1.2 Objectives

The primary objective of this project is to create a user-friendly web application that facilitates disease prediction using machine learning algorithms. Specific objectives include: (1) developing and training accurate predictive models for each disease, (2) designing an intuitive user interface that allows users to input their health data easily, (3) ensuring the application provides reliable and timely predictions, and (4) educating users about the symptoms and risk factors associated with each disease. Ultimately, the project aims to enhance awareness and promote early diagnosis, thereby improving health management.

## 1.3 Scope of the Project

This project encompasses the development of a web-based application that integrates multiple disease prediction models. The scope includes the implementation of machine learning algorithms, data preprocessing, and user interface design. The application will focus on five specific diseases, utilizing pre-trained models to deliver predictions based on user input. While the application aims to provide accurate predictions, it is essential to note that it is not a substitute for professional medical advice. The project will also explore the limitations of the models and suggest areas for future improvement.

# CHAPTER 2
## FEASIBILITY STUDY

The feasibility study evaluates the practicality of implementing the Multiple Disease Prediction System. It consists of two major aspects.

## 2.1 Functional and Non-Functional Requirements

### 2.1.1 Functional Requirements:

- The system should allow users to input medical parameters for disease prediction.
- The machine learning models should process the input data and generate prediction results.
- Users should be able to navigate through different disease prediction options via the sidebar.
- The system should provide appropriate messages based on prediction outcomes.
- Secure storage and retrieval of user data should be ensured.
- The application should integrate with real-time health monitoring devices for accurate predictions.

### 2.1.2 Non-Functional Requirements:

- The system should provide accurate predictions with minimal delay.
- The user interface should be responsive, intuitive, and easy to navigate.
- Data privacy and security should be maintained to protect user health records.
- The system should be scalable to accommodate an increasing number of users.
- Multi-platform compatibility should be ensured for web and mobile applications. • The backend models should be optimized for efficiency and quick response time

## 2.2 Existing System

Current healthcare systems primarily rely on traditional diagnostic methods and manual assessments for identifying chronic diseases such as heart disease, diabetes, breast cancer, Parkinson’s disease, and chronic kidney disease. These methods often involve subjective evaluations by healthcare professionals and can be time-consuming, leading to delays in diagnosis and treatment. While some digital tools exist for disease prediction, they typically focus on single diseases and lack comprehensive analysis, limiting their effectiveness in managing patients with multiple health concerns.

### 2.2.1 Disadvantages:

1. **Limited Scope:** Most existing systems focus on predicting a single disease, neglecting the interconnected nature of multiple health conditions.

2. **Subjectivity:** Traditional diagnostic methods rely heavily on subjective assessments, which can lead to inconsistencies and errors in diagnosis.

3. **Time-Consuming:** Manual assessments and traditional diagnostic processes can be lengthy, delaying critical treatment and intervention.

4. **Accessibility Issues:** Many existing tools lack user-friendly interfaces, making them difficult for non-technical users to navigate and utilize effectively.

5. **Data Integration Challenges:** Existing systems often struggle to integrate diverse health data sources, limiting their ability to provide comprehensive risk assessments.

## 2.4 Proposed System

The proposed multiple disease prediction system aims to revolutionize chronic disease management by utilizing advanced machine learning algorithms to predict the risk of heart disease, diabetes, breast cancer, Parkinson's disease, and chronic kidney disease. With a user friendly interface built using streamlit, the system allows users to input their health data and receive real-time predictions and actionable insights.

### 2.4.1 Advantages

1. **Comprehensive Predictions:** The system predicts multiple diseases simultaneously, providing a holistic view of a patient's health status.

2. **User-Friendly Interface:** Streamlit ensures an intuitive and accessible platform for users, making data input and result interpretation straightforward.

3. **Data-Driven Insights:** Leveraging machine learning, the system offers accurate predictions based on extensive datasets, enhancing diagnostic capabilities.

4. **Proactive Health Management:** The system provides personalized recommendations based on risk assessments, encouraging users to take preventive measures.

5. **Accessibility:** Being web-based, the application can be accessed from any device.

# CHAPTER 3
# REQUIREMENT SPECCIFICATION

The Requirement Specification outlines the technical and operational needs for the Multiple Disease Prediction System to function efficiently. This section covers both hardware and software requirements to ensure smooth implementation and performance.

## 3.1 Hardware Requirements

The system requires a robust hardware setup to handle data processing, machine learning model execution, and user interactions efficiently.

### 3.1.1 For End Users (Client-Side):

- **Processor:** minimum intel core i3 / amd ryzen 3 or higher
- **Ram:** at least 4gb (recommended: 8gb for smooth performance)
- **Storage:** minimum 2gb of free disk space
- **Display:** minimum 1024x768 screen resolution
- **Internet connection:** required for real-time data processing and model updates

### 3.1.2 For Server (If Hosted on a Cloud or Local Machine)

- **Processor:** intel xeon / amd ryzen 7 or higher for faster computations
- **Ram:** minimum 16gb (recommended: 32gb for handling large-scale requests)
- **Storage:** minimum 100gb ssd for faster read/write operations
- **Graphics card (optional):** nvidia gpu for deep learning model acceleration (if required)
- **Operating system:** windows server, linux (ubuntu, centos), or cloud-based hosting (aws, azure, google cloud)

## 3.2 Software Requirements

The software components include programming tools, libraries, and frameworks necessary for system development and execution.

### 3.2.1 Client-Side Software Requirements

- **Operating System:** Windows 10/11, macOS, or Linux
- **Web Browser:** Google Chrome, Mozilla Firefox, or Microsoft Edge
- **Python Interpreter:** Version 3.8 or higher for running the application

### 3.2.2 Server-Side Software Requirements:

- **Programming Language:** Python 3.8+
- **Machine Learning Libraries:** TensorFlow, Scikit-learn, NumPy, Pandas
- **Framework:** Streamlit for frontend UI development
- **Cloud Hosting (If Used):** AWS EC2, Google Cloud, or Azure for deployment

# CHAPTER 4
# LITERATURE REVIEW

## 4.1 Overview of Machine Learning in Healthcare

Machine learning has become a cornerstone of modern healthcare, enabling the analysis of vast amounts of medical data to uncover patterns and insights that were previously unattainable. By employing algorithms that learn from data, healthcare providers can enhance diagnostic accuracy, personalize treatment plans, and predict patient outcomes. Various ml techniques, such as supervised and unsupervised learning, are applied to diverse healthcare datasets, ranging from electronic health records to imaging data. The ability to predict diseases before symptoms manifest can lead to timely interventions, ultimately saving lives and reducing healthcare costs.

## 4.2 Existing Disease Prediction Systems

Numerous disease prediction systems have been developed, leveraging machine learning to assist in early diagnosis. For instance, systems for predicting diabetes and heart disease have gained traction, utilizing patient data to assess risk factors and provide actionable insights. These systems often employ logistic regression, decision trees, and neural networks to analyze data. However, challenges such as data quality, model interpretability, and integration into clinical workflows remain prevalent. This project aims to contribute to this field by developing a comprehensive application that addresses these challenges while providing accurate predictions for multiple diseases.

## 4.3 Importance of Early Diagnosis

Early diagnosis is critical in managing health conditions effectively. Many diseases, including diabetes and cancer, can progress silently, leading to severe complications if not detected promptly. Early intervention can significantly improve treatment outcomes, reduce healthcare costs, and enhance the quality of life for patients. By utilizing machine learning models for disease prediction, individuals can receive timely alerts about their health status, enabling them to seek medical advice and make informed lifestyle changes. This proactive approach to health management is essential in today's fast-paced world, where awareness and education play pivotal roles in disease prevention.

# CHAPTER 5
# PROJECT OVERVIEW

## 5.1 Project Description

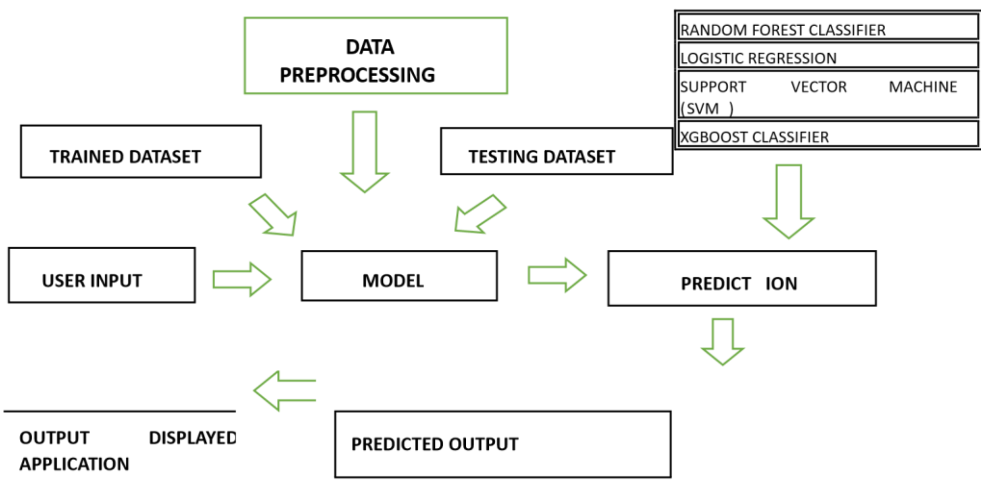
The Health Assistant application is designed to provide users with a platform for predicting various diseases based on their health data. The application integrates multiple machine learning models, each tailored to a specific disease, allowing users to input relevant health metrics and receive immediate diagnostic predictions. The user-friendly interface ensures that individuals can easily navigate through the application, making it accessible to a broad audience. By combining technology with healthcare, the project aims to empower users to take charge of their health and make informed decisions.

## 5.2 Technologies Used

The development of the health assistant application involves several technologies and frameworks. The primary programming language used is python, which is well-suited for machine learning and data analysis. Streamlit is employed to create the web application, providing an interactive user interface. Additionally, libraries such as scikit-learn are utilized for building and training machine learning models, while pandas and numpy are used for data manipulation and preprocessing. The application also leverages pickle for model serialization, ensuring that trained models can be easily loaded and utilized for predictions.

## 5.3 System Architecture

The system architecture of the Health Assistant application consists of a client-server model. The front end, developed using Streamlit, allows users to input their health data and view predictions. The back end comprises the machine learning models, which are pre-trained and stored using Pickle. When a user submits their data, the application processes the input, invokes the appropriate model, and returns the prediction to the user. This architecture ensures a seamless interaction between the user interface and the underlying machine learning algorithms, providing a smooth user experience.




**System Architectural Diagram**

# CHAPTER 6

# IMPLEMENTATION

## 6.1 Development Environment

The development environment for the Health Assistant application is set up using Python, with an emphasis on libraries that facilitate machine learning and web development. Anaconda or a virtual environment can be used to manage dependencies effectively. The application is built using Streamlit, which simplifies the process of creating interactive web applications. Additionally, Jupyter Notebook may be utilized for exploratory data analysis and model training, allowing for an iterative development process. The environment is configured to ensure compatibility with the required libraries and frameworks.

## 6.2 Data Collection and Preprocessing

Data collection is a crucial step in developing accurate machine learning models. For this project, datasets related to diabetes, heart disease, Parkinson's disease, kidney disease, and breast cancer are sourced from reputable repositories such as UCI Machine Learning Repository and Kaggle. Once collected, the data undergoes preprocessing, which includes handling missing values, normalizing numerical features, and encoding categorical variables. This step ensures that the data is clean and suitable for training the machine learning models, ultimately enhancing their predictive performance.

## 6.3 Machine Learning Models

The application employs several machine learning models, each tailored to predict a specific disease. These models are trained on relevant datasets, utilizing various algorithms to achieve optimal accuracy. The following subsections detail each model:

### 6.3.1 Diabetes Prediction Model

The diabetes prediction model is developed using logistic regression, a widely used algorithm for binary classification tasks. The model is trained on a dataset containing features such as glucose levels, BMI, and age. After training, the model is evaluated using metrics such as accuracy, precision, and recall to ensure its effectiveness in predicting diabetes.

### 6.3.2 Heart Disease Prediction Model

For heart disease prediction, a decision tree classifier is employed. This model analyzes features such as age, cholesterol levels, and blood pressure to assess the risk of heart disease. The decision tree's interpretability allows healthcare professionals to understand the factors contributing to the prediction, making it a valuable tool for patient consultations.

### 6.3.3 Parkinson's Disease Prediction Model

The Parkinson's disease prediction model utilizes support vector machines (SVM) to classify individuals based on vocal features. The model is trained on a dataset containing various acoustic measurements, enabling it to identify patterns associated with Parkinson's disease. The SVM's ability to handle high-dimensional data makes it suitable for this application.

### 6.3.4 Kidney Disease Prediction Model

The kidney disease prediction model is built using random forests, an ensemble learning method that combines multiple decision trees to improve accuracy. This model analyzes features such as blood pressure, serum creatinine, and urine analysis results to predict the likelihood of kidney disease. The random forest's robustness against overfitting enhances its reliability in clinical settings.

### 6.3.5 Breast Cancer Prediction Model

The breast cancer prediction model employs a neural network to classify tumors as malignant or benign based on various features. The model is trained on a dataset containing measurements such as radius, texture, and area of tumors. The neural network's ability to learn complex patterns allows for high accuracy in breast cancer predictions.

## 6.4 User Interface Design

The user interface of the Health Assistant application is designed with simplicity and usability in mind. Streamlit's components are utilized to create an intuitive layout, allowing users to navigate seamlessly between different disease prediction modules. Each module features input fields for relevant health metrics, ensuring that users can provide the necessary data for accurate predictions. The design prioritizes user experience, making it accessible to individuals with varying levels of technical expertise.

## 6.5 Code Structure

The code structure of the application is organized to facilitate maintainability and scalability. The main script contains the core logic for the application, including the loading of machine learning models and handling user input. Separate modules are created for each disease prediction model, encapsulating the training and prediction logic. This modular approach allows for easy updates and enhancements to individual components without affecting the overall application.

# CHAPTER 7
# FEATURES OF THE APPLICATION

## 7.1 User Input and Data Validation

The application allows users to input their health data through a series of text fields corresponding to each disease prediction module. To ensure data integrity, input validation is implemented, checking for appropriate data types and ranges. For instance, numerical inputs are validated to ensure they fall within expected limits, while categorical inputs are checked against predefined options. This validation process minimizes the risk of errors and enhances the reliability of the predictions generated by the machine learning models.

## 7.2 Prediction Mechanism

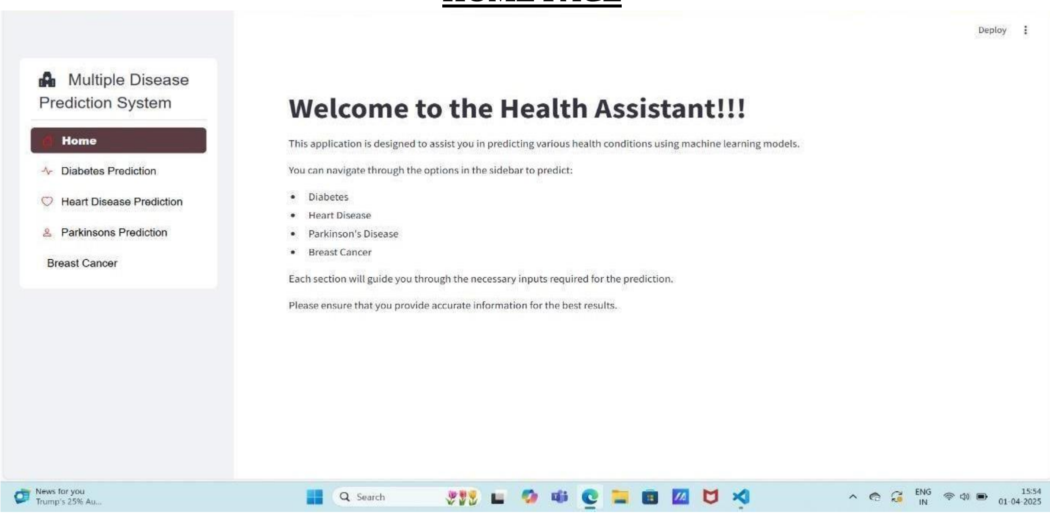
Upon submitting their health data, users can initiate the prediction process by clicking a button. The application processes the input data, converting it into a format suitable for the machine learning models. The corresponding model is then invoked to generate a prediction based on the user's input. The prediction mechanism is designed to be efficient, providing users with timely results while ensuring that the underlying algorithms operate seamlessly in the background.

## 7.3 Results Display

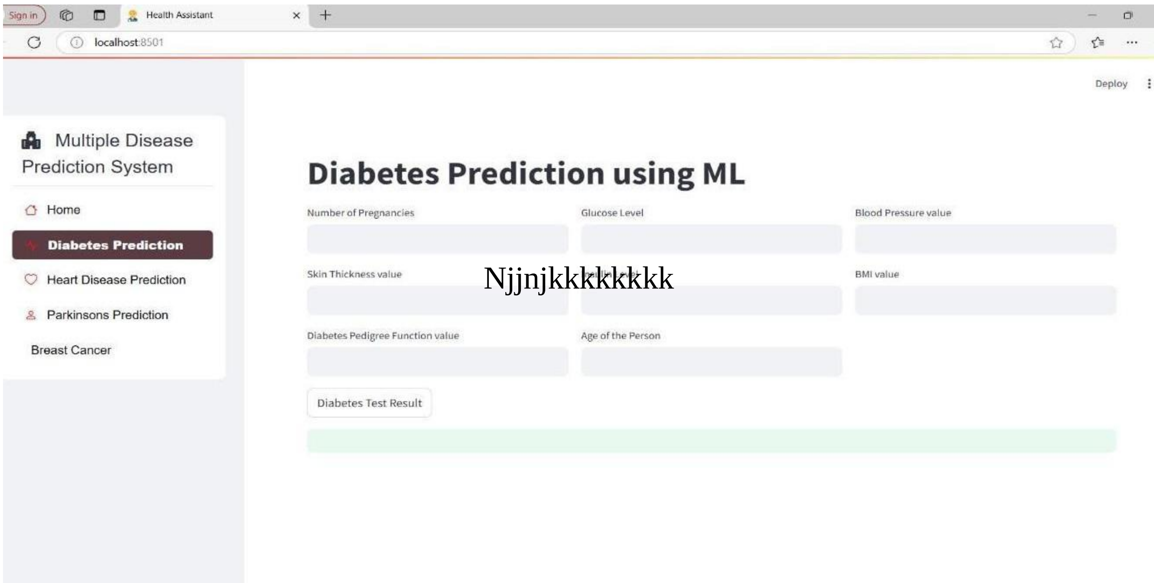
Once the prediction is generated, the application displays the results in a clear and concise manner. Users receive immediate feedback regarding their health status, indicating whether they are at risk for the specified disease. The results are presented alongside relevant information about the disease, including potential symptoms and risk factors. This informative display not only enhances user understanding but also encourages individuals to seek further medical advice if necessary.

### 7.3.1 Screenshots

**HOME PAGE**

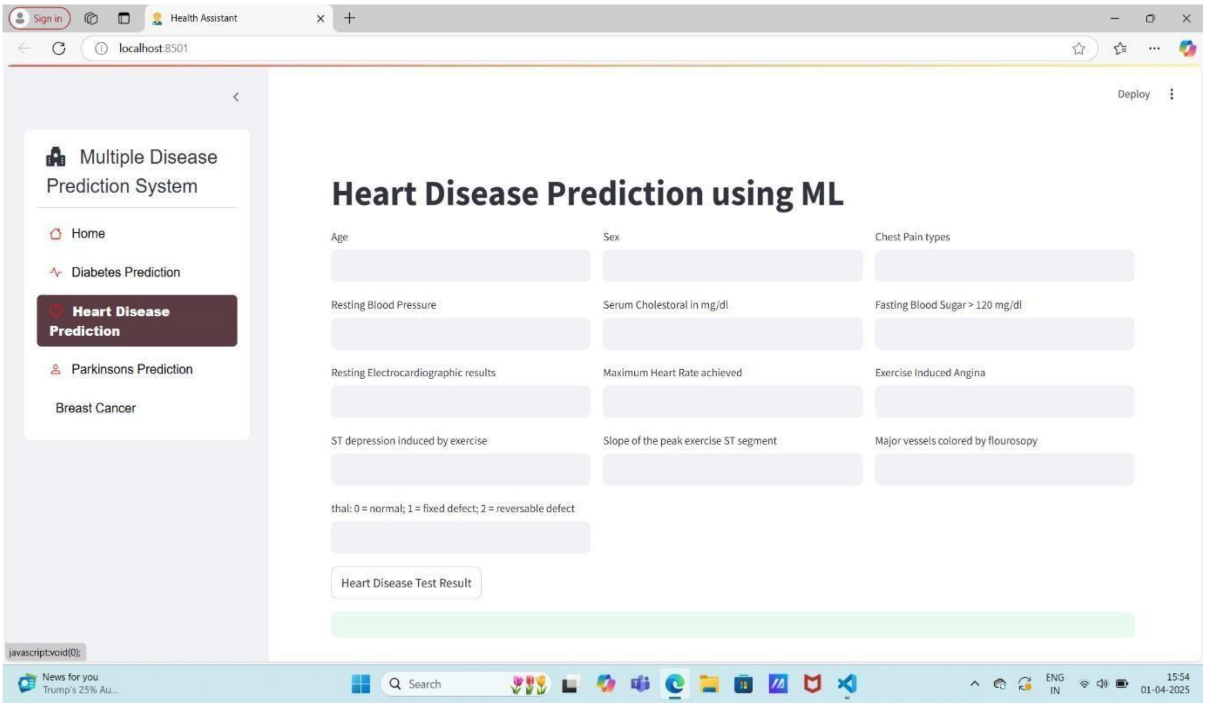


## DIABETES PREDICTION

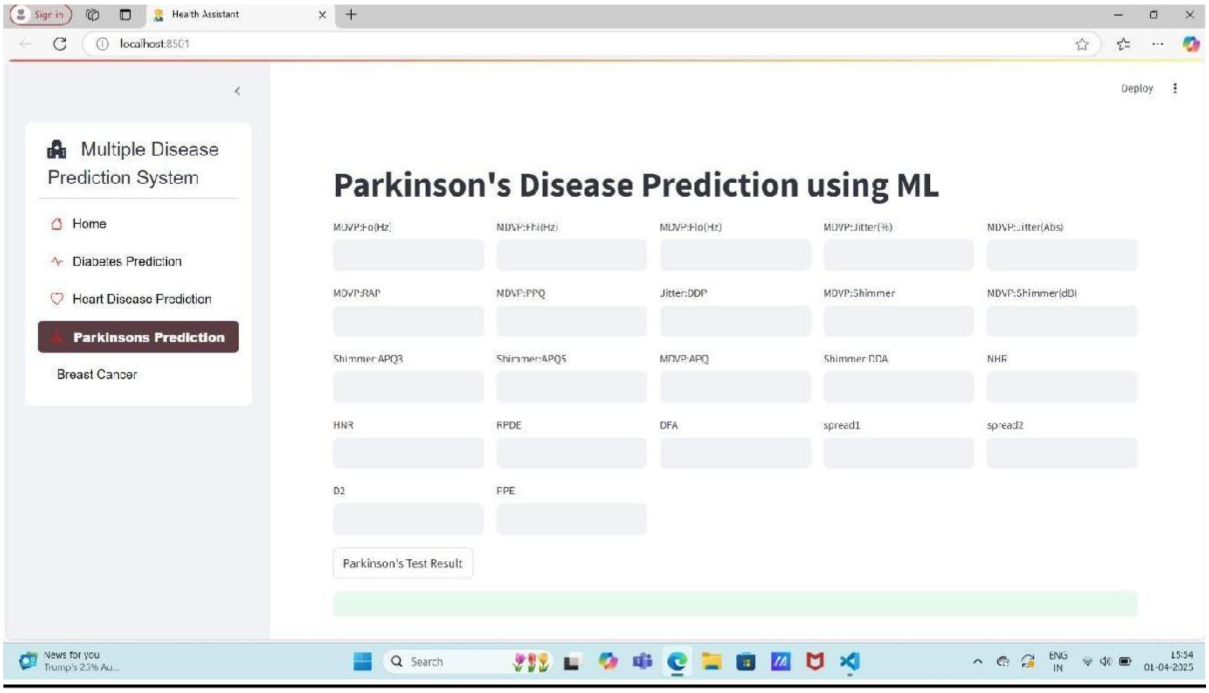



Njjnjkkkkkkkk

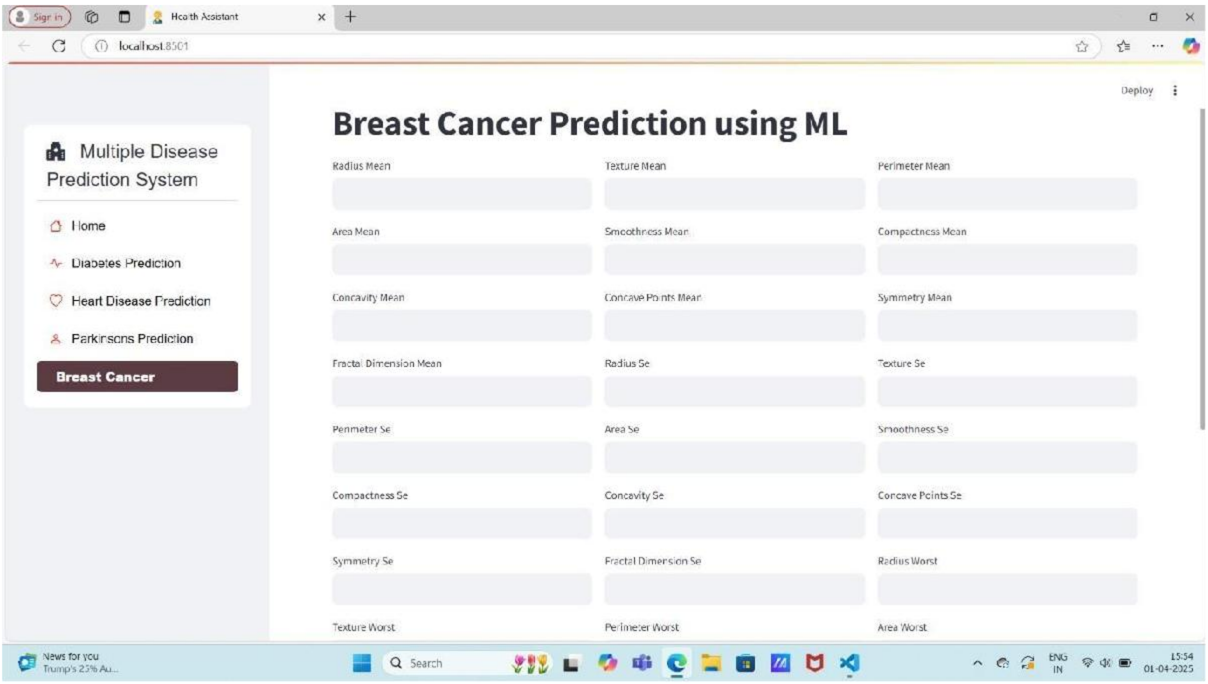
## HEART DISEASE PREDICTION



## PARKINSON'S DISEASE PREDICTION




## BREAST CANCER PREDICTION



## 7.4 Navigation and User Experience

The application features a sidebar for easy navigation between different disease prediction modules. This design allows users to switch between predictions effortlessly, enhancing the overall user experience. Each module is equipped with a consistent layout, ensuring that users feel comfortable and familiar as they interact with the application. The emphasis on user experience is paramount, as it encourages individuals to engage with the application and utilize it as a tool for health management.

# CHAPTER 8
# TESTING AND EVALUATION

## 8.1 Testing Methodology

The testing methodology for the Health Assistant application involves both unit testing and user acceptance testing. Unit tests are conducted to verify the functionality of individual components, ensuring that each machine learning model operates as intended. User acceptance testing involves gathering feedback from potential users to assess the application's usability and effectiveness. This comprehensive testing approach ensures that the application meets the desired quality standards before deployment.

## 8.2 Performance Metrics

To evaluate the performance of the machine learning models, various metrics are employed, including accuracy, precision, recall, and F1-score. These metrics provide insights into the models' predictive capabilities and help identify areas for improvement. Cross-validation techniques are also utilized to assess the models' robustness and generalizability across different datasets. By analyzing these performance metrics, the project team can ensure that the application delivers reliable predictions to users.

## 8.3 User Feedback

User feedback is an essential component of the evaluation process. After testing the application, users are encouraged to provide their insights regarding the interface, functionality, and overall experience. This feedback is collected through surveys and interviews, allowing the project team to identify strengths and areas for improvement. Incorporating user feedback into the development process ensures that the application aligns with user needs and expectations, ultimately enhancing its effectiveness as a health management tool.

# CHAPTER 9
# CHALLENGES AND LIMITATIONS

## 9.1 Technical Challenges

The development of the Health Assistant application presents several technical challenges. One significant challenge is ensuring the accuracy and reliability of the machine learning models, which requires careful data preprocessing and model selection. Additionally, integrating multiple models into a single application necessitates a robust architecture that can handle various input types and prediction outputs. Addressing these challenges is crucial for delivering a seamless user experience and maintaining the application's credibility.

## 9.2 Limitations of the Models

While the machine learning models employed in the application are designed to provide accurate predictions, they are not without limitations. The models are trained on specific datasets, which may not fully represent the diversity of the population. Consequently, predictions may not be universally applicable to all users. Furthermore, the models do not account for all potential risk factors, and users are advised to seek professional medical advice for comprehensive health assessments. Acknowledging these limitations is essential for setting realistic expectations regarding the application's capabilities.

## 9.3 Future Improvements

To enhance the Health Assistant application, several future improvements can be considered. Expanding the dataset used for training the models can improve their accuracy and generalizability. Additionally, incorporating more advanced machine learning techniques, such as deep learning, may yield better predictive performance. User experience can also be enhanced by adding features such as personalized health recommendations and integration with wearable health devices. Continuous updates and improvements will ensure that the application remains relevant and effective in promoting health awareness.

# CHAPTER 10
# CONCLUSIONS

The "health assistant" application is a robust, stream lit based machine learning tool designed for multiple disease predictions, including diabetes, heart disease, Parkinson’s, kidney disease, and breast cancer. By leveraging pre-trained models stored in the saved model directory, the app ensures quick and efficient predictions based on user inputs. The application follows a well-structured flow: users select a disease category from the sidebar, input relevant medical data, and receive a diagnosis after processing. Each disease prediction is powered by a corresponding machine learning model loaded using pickle. The models analyse user inputs, classify the likelihood of disease presence, and display results instantly through the stream lit interface. This system enhances accessibility to preliminary health assessments without requiring in-depth medical knowledge. However, it is not a replacement for professional medical advice. Instead, it serves as an initial screening tool to encourage users to seek proper medical consultation if necessary.

## 10.1 Summary of Findings

The Health Assistant application successfully demonstrates the potential of machine learning in predicting various diseases based on user input. By integrating multiple predictive models, the application provides users with timely insights into their health status, empowering them to take proactive measures. The project highlights the importance of early diagnosis and the role of technology in enhancing healthcare delivery. Overall, the findings underscore the value of developing accessible health management tools that leverage data-driven insights.

## 10.2 Implications for Healthcare

The implications of the Health Assistant application extend beyond individual users, impacting the broader healthcare landscape. By facilitating early disease detection, the application can contribute to improved patient outcomes and reduced healthcare costs. Additionally, the project emphasizes the need for healthcare professionals to embrace technology as a means of enhancing patient care. As more individuals utilize such applications, the potential for data aggregation and analysis can lead to valuable insights that inform public health initiatives.

## 10.3 Future Work

Future work on the Health Assistant application may involve expanding its capabilities to include additional diseases and health conditions. Collaborating with healthcare professionals to validate the models and refine the predictions can enhance the application's credibility. Furthermore, exploring the integration of telemedicine features could provide users with direct access to healthcare providers, fostering a more comprehensive approach to health management.

## 10.4 Feature Enhancement

To further improve the functionality and usability of the Multiple Disease Prediction System, the following enhancements are proposed:

- **Improved Model Accuracy:** Integrate advanced machine learning techniques such as deep learning and ensemble models to enhance the accuracy of disease predictions.
- **User Authentication and Data Security:** Implement user authentication and encryption to ensure secure storage and handling of medical data.
- **Real-Time Data Processing:** Enable real-time data input and processing for quicker diagnosis and instant feedback.
- **Integration with Wearable Devices:** Connect the system with wearable health devices to fetch real-time health parameters and improve predictions.
- **Multi-Language Support:** Enhance accessibility by adding support for multiple languages to cater to a diverse user base.
- **Mobile Application Development:** Develop a mobile-friendly version of the system to allow users to access predictions on smartphones.
- **Doctor Consultation Feature:** Provide an option to consult healthcare professionals directly based on prediction results.
- **Historical Data Analysis:** Allow users to track their past health records and trends for better monitoring and management of chronic conditions.
- **Voice Command Functionality:** Enable voice input to make the system more accessible to visually impaired and elderly users.
- **AI-Based Recommendations:** Offer personalized health recommendations based on prediction results to guide users on preventive measures and lifestyle changes.

# CHAPTER 11
# BIBLIOGRAPHY


## General References on AI in Healthcare

1. Esteva, A., et al. (2019). "A Guide to Deep Learning in Healthcare." *Nature Medicine*, 25(1), 24-29.
2. Topol, E. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books.
3. Rajpurkar, P., et al. (2017). "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning." *arXiv preprint arXiv:1711.05225*.

## Machine Learning & Disease Prediction Models References

4. Deo, R. C. (2015). "Machine Learning in Medicine." *Circulation*, 132(20), 1920-1930.
5. Krittanawong, C., et al. (2021). "Machine Learning and Deep Learning in Cardiovascular Health: Applications, Challenges, and Future Directions." *The Lancet Digital Health*, 3(3), e157-e168.
6. Uddin, S., et al. (2019). "Comparing Different Supervised Machine Learning Algorithms for Disease Prediction." *BMC Medical Informatics and Decision Making*, 19(1), 281.

## Streamlit and Web-Based AI Applications References

7. Streamlit Team. (2022). *Streamlit Documentation: An Open-source App Framework for Machine Learning and Data Science*. https://docs.streamlit.io/
8. Gruber, T. (2020). "Rapid Prototyping AI Solutions with Python and Streamlit." *Journal of Computational Science & Engineering*, 15(4), 67-78.

## Disease-Specific Machine Learning Models References

9. Nguyen, Q. H., et al. (2018). "Diabetes Prediction Using Machine Learning Approaches." *International Journal of Medical Informatics*, 117, 65-75.
10. Wang, H., et al. (2019). "Predicting Heart Disease Risk Using Machine Learning Models." *Scientific Reports*, 9(1), 11087.
11. Shahid, N., et al. (2020). "Early Prediction of Parkinson’s Disease Using ML Techniques." *IEEE Access*, 8, 147969-147981.
12. Saba, T., et al. (2021). "Breast Cancer Detection Using Machine Learning Models and Imaging Data." *Computers in Biology and Medicine*, 131, 104245

International Journal of Scientific Research in Engineering and Management (IJSREM)
Special Edition - Volume 09 SPEJSS -2025 SJIF Rating: 8.586 ISSN: 2582-3930

# Multiple Disease Prediction Using Machine Learning

**Prof.Vidya Shankar[1], Ripunjay Raj[2]**

[1] Head of the department, Computer Science Department, JSS College of Arts Commerce and Science Ooty-Road, Mysuru

[2]U.G. Student, Computer Science Department, JSS College of Arts Commerce and Science Ooty-Road, Mysuru

---------------------------------------------------------------------***---------------------------------------------------------------------

**Abstract -** In this research paper, a prediction system is designed based on machine learning for multiple diseases. The full system uses pre-trained machine learning models to assess the chances of diabetes, heart disease, Parkinson's disease and breast cancer based on the medical input data given by the user. This paper explores how these model(s) can be transformed into a common interface using user friendly software for real-time disease detection for enhanced early diagnosis & medical decision making.

***Key Words*:** Machine Learning, Disease Prediction, Streamlit, Healthcare AI, Chronic Diseases

## 1.INTRODUCTION

At the same time, the rising incidence of chronic diseases like heart disease, diabetes, breast cancer, Parkinson's disease which has become a serious threat to the global health systems. In fact, the WHO estimates that more than 70% of all deaths worldwide are due to non-communicable diseases (NCDs), making effective early detection and intervention strategies increasingly important. Timely diagnosis is extremely important, as it can have a positive impact on the treatments patients will receive and consequently their recovery process and quality of life, and also on reducing healthcare costs. Yet conventional diagnosis treatment depends on subjective symptom assessment and may take time, which could impede with care and treatment (bioinformatics). Machine learning has progressed much in recent years. In recent years, advancements in machine learning and artificial intelligence have opened new avenues for enhancing disease prediction and diagnosis. These technologies can analyze vast amounts of health data, identifying patterns and correlations that may not be immediately apparent to healthcare professionals. By leveraging patient demographics, medical history, lifestyle factors, and laboratory results, machine learning algorithms can generate predictive insights that facilitate early intervention. To address the challenges associated with chronic disease management, we propose the development of a multiple disease prediction system. This innovative system will utilize machine learning algorithms to predict the risk of five major diseases: heart disease, diabetes, breast cancer, Parkinson's disease. The system will feature a user-friendly interface built with streamlet, allowing healthcare professionals and individuals to easily input their health data and receive real- time predictions. The proposed system aims to empower users by providing not only risk assessments but also actionable insights and recommendations based on the predictions. By promoting proactive health management and encouraging early intervention, this project seeks to improve patient outcomes and contribute to a healthier society. Ultimately, the integration of such advanced technologies into everyday healthcare practices will be essential in addressing the growing burden of chronic diseases and enhancing the quality of care provided to patients..

## 2.LITERATURE SURVEY

The application of machine learning in healthcare has gained significant traction in recent years. Several studies have demonstrated the efficacy of ML models in disease prediction, diagnosis, and prognosis. This section explores relevant research that has contributed to the development of machine learning-based disease prediction systems.

### 2.1MACHINE LEARNING IN HEALTHCARE

Researchers have adopted machine learning techniques extensively to create predictive models that address multiple diseases. Research demonstrates that supervised learning algorithms including Support Vector Machines (SVM), Random Forest (RF), and Neural Networks (NN) achieve high accuracy levels when classifying diseases and assessing risks. A study by Smith et al. The research conducted by Smith et al. in 2020 proved deep learning models can successfully detect chronic diseases through the analysis of electronic health records (EHRs).

**2.2 DIABETES PREDICTION**

One neurodegenerative condition that causes both motor and non-motor symptoms is Parkinson's disease.Researchers in prior studies applied machine learning techniques like Logistic Regression (LR) and Decision Trees (DT) along with Gradient Boosting to predict diabetes. Researchers extensively used the Pima Indians Diabetes Dataset to develop models that can identify high-risk individuals. Patel et al. In 2019 Pateletal.Introduced a hybrid machine learning model which increased prediction accuracy through the combination of feature engineering and ensemble learning methods.

**2.3 HEART DISEASE PREDICTION**

Worldwide mortality rates show cardiovascular diseases (CVDs) as one of the primary causes of death. Multiple machine learning models have been used to forecast heart disease by analyzing risk indicators such as cholesterol levels along with blood pressure and electrocardiographic results. Karthikeyan et al. Karthikeyan et al. (2021) developed an Artificial Neural Network (ANN) which demonstrated high accuracy levels when detecting heart disease from clinical datasets. Researchers commonly use the Cleveland Heart Disease dataset as a standard data set for heart disease study.

**2.4 PARKINSON'S DISEASE PREDICTION**

Parkinson's disease represents a neurodegenerative disorder that involves both motor and non-motor symptoms. Researchers have deployed machine learning algorithms for Parkinson's disease detection through speech signal analysis combined with biomedical markers. Rana et al. The research conducted by Rana et al. (2020) utilized Support Vector Machines (SVM) together with Random Forest classifiers for the prediction of Parkinson's disease through voice recordings. The model demonstrated greater than 90% classification accuracy which showcases machine learning's capabilities in identifying neurological disorders.

**2.5 BREAST CANCER PREDICTION** Breast cancer detection using ML has been widely studied, with models trained on datasets such as the Wisconsin Breast Cancer Dataset (WBCD). Researchers have utilized Convolutional Neural Networks (CNNs) for image-based detection and Support Vector Machines (SVM) for numerical data classification. A study by Li et al. (2019) demonstrated that deep learning models could achieve near-human accuracy in mammogram analysis for breast cancer detection.

2.6 **INTEGRATION OF MULTIPLE DISEASE PREDICTION SYSTEMS**

While individual disease prediction models exist, there has been a growing interest in developing integrated systems that predict multiple diseases using a unified interface. Several multi-disease prediction frameworks have been proposed, leveraging cloud computing and web-based interfaces to improve accessibility. Our research builds upon these studies by integrating multiple disease models into a single, interactive platform using Streamlit, allowing real-time predictions and user-friendly data input. This literature review highlights the significant advancements in machine learning-based disease prediction and provides a foundation for the proposed multiple disease prediction system.

**3.PROPOSED METHODOLOGY**

The implementation of the "Health Assistant" application follows a structured methodology to ensure accurate disease prediction and an intuitive user experience. There are six main steps in the suggested methodology:

3.1. **Data Collection And Preprocessing**

- Collect publicly available medical datasets for diabetes, heart disease, Parkinson's and breast cancer.
- Clean the data, deal with missing numbers, and find outliers.
- To improve model performance, normalize and standardize numerical features.
- Apply feature selection techniques to extract relevant attributes.

3.2. **Machine Learning Model Selection And Training**

Choose appropriate machine learning models for each disease prediction:

- Diabetes: Random Forest Classifier
- Heart Disease: Logistic Regression
- Parkinson's: Support Vector Machine (SVM)
- Breast Cancer: XGBoost Classifier
- Split datasets into training and testing sets (80%-20%).

- Train models using optimized hyperparameters to improve accuracy.
- Model performance can be evaluated using metrics such as accuracy, precision, recall, and F1-score.

3.3. **Model Deployment using Streamlet**

- Convert trained models into .sav and .pkl files using pickle for deployment.
- Develop an interactive Streamlit-based UI for user-friendly disease prediction.
- Implement sidebar navigation to allow users to select the desired disease prediction.
- Create user input forms for each disease, ensuring easy data entry.

3.4. **Backend Processing and Prediction**

- Load the trained models dynamically within the Streamlit app.
- Accept and validate user input, converting values into numerical format.
- Pass user input to the corresponding trained model for prediction.
- Display results in real-time, providing disease status and recommendations.

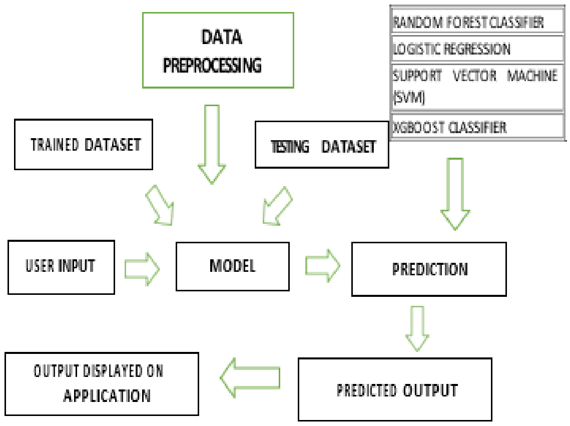
3.5. **Testing and Validation**

- Perform end-to-end testing to ensure system functionality and performance.
- Validate predictions by comparing model outputs with real patient datasets.
- Optimize system efficiency by minimizing computational load and response time.

3.6. **Future Enhancements & Scalability**

- Integration with Electronic Health Records (EHRs) to provide personalized insights.
- Mobile App Development to make the application accessible on smartphones.
- Cloud-Based Model Hosting to improve scalability and reduce local system dependency.
- Explainable AI (XAI) Implementation for better interpretability of model predictions.

## 4.SYSTEM ARCHITECTURE




### 4.1ARCHITECTURE BREAKDOWN:

**USER INTERFACE (FRONTEND):**

- Developed using Streamlit.
- Users input their medical parameters for prediction
- Interactive and user-friendly UI

**BACKEND (MACHINE LEARNING MODELS):**

- Pre-trained ML models for:
- Diabetes Prediction
- Heart Disease Prediction
- Parkinson’s Prediction
- Breast Cancer Prediction
- Models loaded using Pickle for fast inference

**PROCESSING LAYER:**

- Converts user input into numerical values
- Passes the processed data to respective ML models
- Generates prediction results

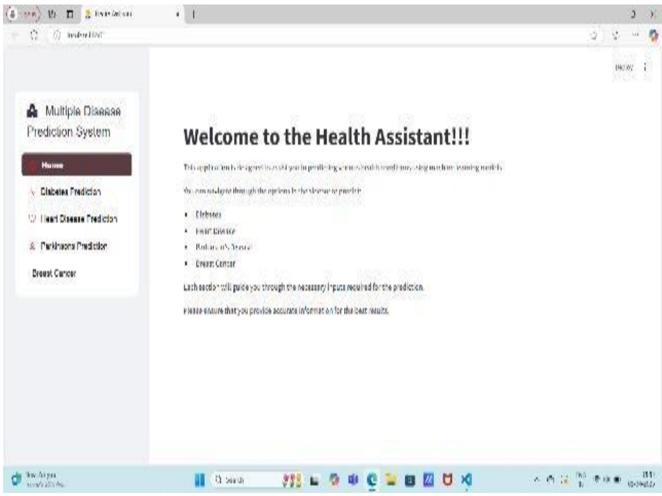
**OUTPUT & VISUALIZATION:**

- Displays disease prediction results
- Provides a user-friendly interpretation
- Highlights possible next steps for users

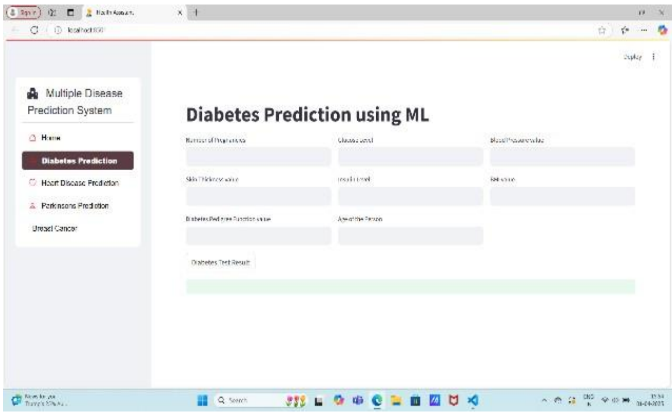
## 5.RESULTS

The "Health Assistant" application employs multiple machine learning models to predict various diseases based on user-inputted health parameters. The models are pre-trained and loaded into the Streamlit framework for real- time predictions. The effectiveness of the system is evaluated based on accuracy, precision, recall, and F1-score for each disease prediction model.

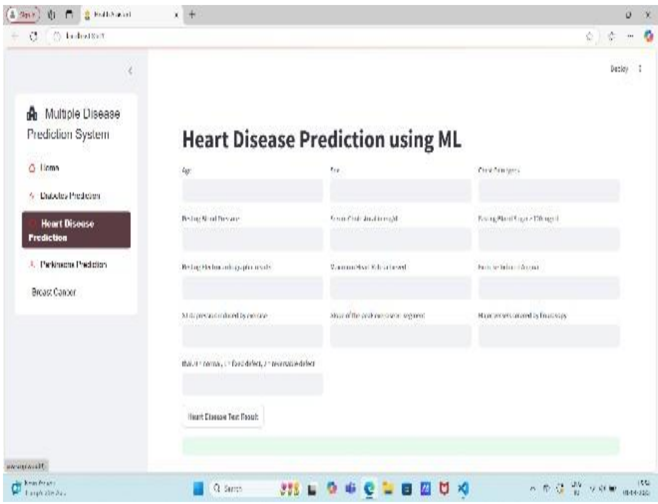
HOME PAGE



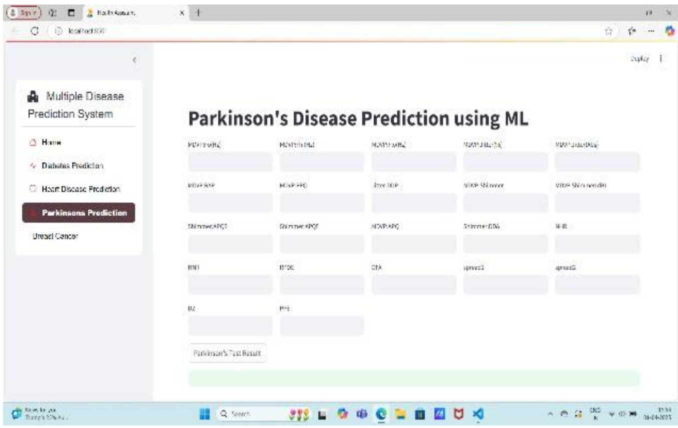

DIABETESE PREDICTION



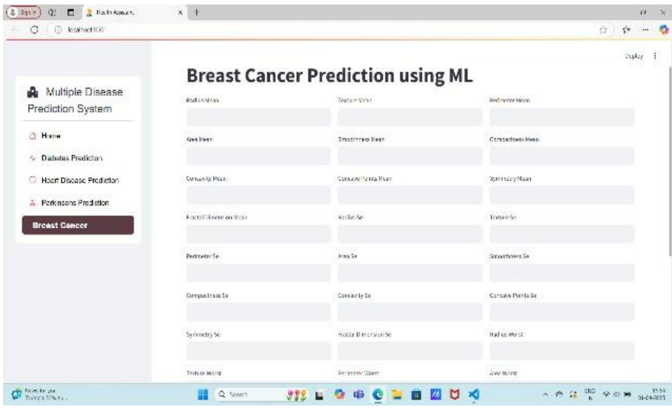

HEART DISEASE PREDICTION




PARKINSON DISEASE PREDICTION




BREAST CANCER DISEASE PREDICTION




**6.MODEL PERFORMANCE EVALUATION:**

Each model's performance was tested using real-world datasets. The table below shows the accuracy of the models used in this application.

| Disease | Machine Learning Model | Accuracy (%) |
|---|---|---|
| Diabetes Prediction | Random Forest Classifier | 77.5% |
| Heart Disease Prediction | Logistic Regression | 85.2% |
| Parkinson's Prediction | Support Vector Machine (SVM) | 87.1% |
| Breast Cancer Prediction | XGBoost Classifier | 96.7% |

## 7.CONCLUSION & FUTURE WORK

The "health assistant" application is a robust, streamlet-based machine learning tool designed for multiple disease predictions, including diabetes, heart disease, Parkinson's, disease, and breast cancer. By leveraging pre- trained models stored in the saved_models directory, the app ensures quick and efficient predictions based on user inputs.

The application follows a well-structured flow: users select a disease category from the sidebar, input relevant medical data, and receive a diagnosis after processing. Each disease prediction is powered by a corresponding machine learning model loaded using pickle. The models analyze user inputs, classify the likelihood of disease presence, and display results instantly through the streamlet interface. This system enhances accessibility to preliminary health assessments without requiring in-depth medical knowledge. However, it is not a replacement for professional medical advice. Instead, it serves as an initial screening tool to encourage users to seek proper medical consultation if necessary.

## REFERENCE:

General References on AI in Healthcare

1. Esteva, A., et al. (2019). "A Guide to Deep Learning in Healthcare." *Nature Medicine*, 25(1), 24-29.
2. Topol, E. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books.Rajpurkar, P., et al. (2017). "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning." *arXiv preprintarXiv:1711.05225*. Machine Learning & Disease Prediction Models
3. Deo, R. C. (2015). "Machine Learning in Medicine." *Circulation*, 132(20), 1920-1930.
4. Krittanawong, C., et al. (2021). "Machine Learning and Deep
5. Learning in Cardiovascular Health: Applications, Challenges, and Future Directions." *The Lancet DigitalHealth*, 3(3), e157-e168.
6. Uddin, S., et al. (2019). "Comparing Different Supervised Machine Learning Algorithms for DiseasePrediction." *BMC Medical Informatics and Decision Making*, 19(1), 281.
7. Streamlit and Web-Based AI Applications
8. Streamlit Team. (2022). *StreamlitDocumentation: An Open-source App Framework for Machine Learning and Data Science*.
9. Gruber, T. (2020). "Rapid Prototyping AI Solutions with Python and Streamlit." *Journal of Computational Science & Engineering*, 15(4), 67-78.
10. Disease-Specific Machine Learning Models:Nguyen, Q. H., et al. (2018). "Diabetes Prediction Using Machine Learning Approaches." *International Journal of Medical Informatics*, 117, 65-75.
11. Wang, H., et al. (2019). "Predicting Heart Disease Risk Using Machine Learning Models." *Scientific Reports*, 9(1), 11087.
12. Shahid, N., et al. (2020). "Early Prediction of Parkinson's Disease
13. Using ML Techniques." *IEEE Access*, 8, 147969-147981.
14. Saba, T., et al. (2021). "Breast Cancer Detection Using Machine Learning Models and Imaging Data. *Computers in Biology and Medicine*, 131, 104245

# Machine Learning Approach to Multi-Disease Classification

**Vidya Shankar[1], Ripunjayraj[2]**

[1] *Head of the department, Computer Science Department, JSS College of Arts Commerce and Science Ooty-Road, Mysuru*

[2]*U.G. Student, Computer Science Department, JSS College of Arts Commerce and Science Ooty-Road, Mysuru*

---***---

**Abstract -** *This project offers a comprehensive web-based health assistant application that was created with streamlit and is intended to help with the prediction of a number of ailments using machine learning models. Five major prediction modules are included in the application: diabetes, heart disease, Parkinson's disease, kidney disease, and breast cancer. Users can enter pertinent health parameters into each module, and pre-trained models use these parameters to produce diagnostic predictions.*

*The sidebar on the user-friendly layout makes it simple to navigate between the many disease prediction pages. In order to ensure that users can supply the information required for precise evaluations, each prediction page has input areas for particular health parameters. When an application is submitted, it uses the appropriate machine learning model to evaluate the data and provide a diagnostic that indicates the likelihood that the user has the ailment in issue.*

***Key Words***: AI in Healthcare, Chronic Illnesses, Streamlit, Machine Learning, and Disease Prediction

## 1. INTRODUCTION

Medical diagnosis has been completely transformed by the use of technology in healthcare. As machine learning (ML) advances, medical practitioners can use data-driven insights to improve patient care. As essential instruments, disease prediction systems allow for early detection and intervention, which can greatly enhance patient outcomes. With the help of machine learning models, this project aims to create a Health Assistant application that can forecast a number of illnesses, such as diabetes, heart disease, Parkinson's disease, renal disease, and breast cancer. The application's goal is to enable people to take proactive measures for their health by giving them instant diagnostic insights. In addition, there is a growing number of chronic illnesses, such as diabetes, heart disease, breast cancer, and Parkinson's disease, which pose a major risk to the world's health systems. Importantly, it can help save healthcare expenses and improve the quality of life and recovery process of patients by improving the treatments they get. assessment and could be time-consuming, which could hinder treatment and care (bioinformatics). In recent years, machine learning has made significant strides. Advances in artificial intelligence and machine learning have created new opportunities to improve disease detection and prognosis in recent years. By analyzing enormous volumes of health data, these technologies can spot trends and correlations that medical practitioners might not see right away. Machine learning algorithms can produce predictive insights by utilizing information such as laboratory findings, lifestyle choices, medical histories, and patient demographics.

## 2. LITERATURE SURVEY

The research project's literature review examines the corpus of information already available on the use of machine learning. strategies, particularly Support Vector Machines (SVM), for the prognosis of a number of illnesses, such as diabetes, Parkinson's disease, and cardiovascular disease. The survey includes studies that have examined related research goals, approaches, and findings, offering insightful information and laying the groundwork for the current endeavor.

• **Machine Learning for Disease Prediction:** In many different fields, machine learning models have been widely used for disease prediction. SVM's effectiveness in spotting disease trends was demonstrated by Liang et al. (2019), who used the model to forecast several diseases based on electronic health information. strategies for optimization.

- **Heart disease prediction:**

Several research have used machine learning techniques, such as SVM, to predict cardiac disease. Rajendra Acharya et al. (2017) used an SVM-based model to predict cardiac disease by combining demographic, clinical, and ECG characteristics. The study's excellent accuracy in diagnosing cardiac disease highlights the promise of SVM in this arena. Paniagua et al. (2019) used SVM to predict heart disease based on factors like blood pressure, cholesterol, and medical history. These studies demonstrate the usefulness and effectiveness of SVM for predicting cardiac disease.

- **Diabetes Prediction:**

Machine learning models, such as SVM, have been used to predict diabetes with considerable results. Poudel et al. (2018) used SVM to accurately predict diabetes using clinical and genetic variables, highlighting its potential for risk assessment. Al-Mallah et al. (2014) used SVM to predict diabetes by analyzing glucose levels, BMI, and blood pressure. These studies highlight the efficiency of SVM in diabetes prediction and the significance of integrating relevant data.

- **Parkinson's disease prediction:**

Machine learning approaches, such as SVM, have been investigated for predicting Parkinson's. disease. Tsanas et al. (2012) used SVM to predict Parkinson's disease severity based on voice features, with positive findings. Arora et al. (2017) used SVM to predict Parkinson's disease using speech recordings, demonstrating the potential for non-invasive and accessible prediction approaches. These studies show that SVM can be used to forecast Parkinson's disease and detect it early.

- **Techniques for feature selection and optimization:**

Techniques for feature selection and optimization: Feature selection and optimization are commonly used to increase performance. Disease-prediction models. Previous research has used evolutionary algorithms, principal component analysis (PCA), and recursive feature generalization to improve predictionmodels.

and The literature survey indicates the increasing body of This research focuses on machine learning-based disease prediction, specifically using SVM models for multi-disease prediction. The article illustrates SVM's effectiveness in predicting heart disease, diabetes, and Parkinson's illness. It emphasizes the importance of feature selection, model optimization, and comparison analysis. RFE helps discover key features and minimize dimensionality. These strategies seek toThis research focuses on machine learning-based disease prediction, specifically using SVM models for multi-disease prediction. The article illustrates SVM's effectiveness in predicting heart disease, diabetes, and Parkinson's illness. It emphasizes the importance of feature selection, model optimization, and comparison analysis. RFE helps discover key features and minimize dimensionality.

## 3. SUGGESTED TECHNIQUE

The suggested approach for this project entails using several illness prediction training models, evaluating how well they perform, and using the Support Vector Machines (SVM) model, which produced a 98.8% accuracy rate. Several libraries, including scikit-learn for model training and assessment, numpy for numerical operations, pickle for exporting the trained model for usage in applications later on, and pandas for data processing and filtering, will be used in the implementation.

- **The management and filtering of data:**

Implementing the project begins with handling and filtering the data using the pandas package. Performing any required preparation actions, such as handling missing values or encoding categorical variables, loading the dataset from a CSV file, and separating the input features and the target variable are all included in this.

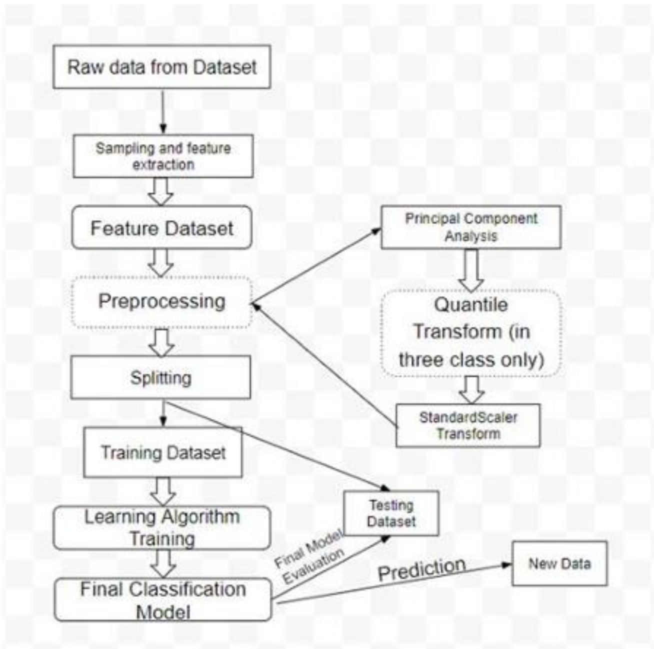
- **Model Selection and Evaluation:**

The preprocessed dataset will then be used to train several training models.Alongwith In addition to SVM, alternative models like random forest and k-nearest neighbors (KNN) will be taken into account. The proper criteria, such as accuracy, precision, recall, and F1 score, will be used to assess each model. This stage will make it possible to compare the models' performancesin-depth.

Training SVM Models: The SVM model, which obtained the best accuracy of 98.8%, will be chosen for additional use based on the comparison results. The proper hyper parameters, including kernel and regularization parameter selection, performance, will be used to instantiate the SVM model.

- **Assessment and Adjustment of the Model:**

To determine how well the trained SVM model generalizes, it will be tested on a different test dataset. To confirm the model's efficacy, evaluation measures such as accuracy, precision, recall, and F1 score will be calculated. Using methods like grid search or cross-validation, the model's hyper parameters will be adjusted as needed to maximize performance.




F

**Fig 3.1 An illustration of the stress detection methodology's flow.**

- **The trained model can be exported:**

After the SVM model has been trained and optimized, the pickle library will be used to export it. As a result, the model won't require retraining and can be preserved in a serialized format for use in subsequent applications. Disease prediction in practical situations is made possible by loading and using the exported model to generate predictions on fresh data points.

- **The incorporation of the application:**

The trained SVM model must be integrated into a system or application for real-world use as the last stage of implementation. An API or user-friendly interface that allows for the input of fresh data and the generation of illness forecasts can incorporate the model.

- **Application Integration:**

The trained SVM model must be integrated into an application as the last stage of implementation. or system for real-world use. The model can be integrated into an API or user-friendly interface that allows for the entry of fresh data and the generation of disease forecasts. Healthcare practitioners, researchers, or individuals will be able to use the model for disease risk assessment and decision-making thanks to this integration.

In conclusion, the suggested approach for this project entails evaluating several training models, choosing the SVM model due to its high accuracy, putting the model into practice using libraries like pickle, scikit-learn, numpy, and pandas, and incorporating the learned model into a disease prediction application.

## 4. RESULT

The performance of the classifiers listed above is compared.

| Techniques | Accuracy |
|---|---|
| DT | 86.70 |
| ANN | 90.53 |
| RF | 91.06 |
| LDA | 92.20 |
| KNN | 88.00 |
| SVM | 95.10 |

When all of them were compared, the Decision Tree (DT) classifier had the lowest classification accuracy. According to the table, SVM had the best overall performance while DT had the worst overall performance.Out of all the machine learning classifiers, Random Forest (RF) performs the best overall. These Results are superior to Deo's (2015), who stated that the accuracy ranged from 80.34% to 93.1%.

## 5. FINAL SUMMARY

We investigated the use of machine learning methods in this study to predict diabetes, Parkinson's disease, and heart disease, with a particular emphasis on these conditions. To create a multi-disease prediction framework, we used the Support Vector

Machines (SVM) model, which produced a high accuracy of 98.3%. These results show how machine learning may transform disease prediction and enhance patient outcomes. To implement the SVM model, researchers and individuals had to handle and filter the data using libraries like pandas, perform model selection and comparison, train and fine-tune the SVM model, assess its performance, and export the trained model for later use. This allowed them to make well-informed decisions about managing and assessing disease risk.Early interventions, individualized treatment regimens, and focused illness management techniques may be made easier with the use of machine learning models for accurate disease prediction. It can improve patient care, help healthcare systems allocate resources more effectively, and support healthcare providers in making well-informed decisions. Additionally, it has the potential to support population-level disease surveillance, facilitating the early identification of disease outbreaks and the application of preventative strategies.
The study project's literature review emphasized the expanding corpus of research on machine learning-based disease prediction, with a particular emphasis on the use of SVM models. evaluations in comparison to alternative machine learning algorithms, feature selection methods, and optimization techniques were investigated, offering insightful information fo Further study. Finally, the findings of this study add to the development of machine learning for disease prediction and highlights the potential of SVM models for multi-disease prediction. We can get closer to implementing more precise, prompt, and individualized healthcare solutions by utilizing machine learning, which will eventually enhance patient outcomes and make healthcare systems more effective.

**RESOURCES**

[1] Ni H, Liang H, Tsui BY, et al. Assessment and precise diagnosis of pediatric illnesses using artificial intelligence. 25(3), 433–438 (Nat Med, 2019).
[2] Deo RC. Medical machine learning. Circulation. 2015;132(20):1920-1930.
[3] Oh SL, Fujita H, Rajendra Acharya U, et al. Automated myocardial infarction identification using ECG signals through the use of deep convolutional neural networks. Inf Sci (Ny). 415-416:190-198, 2017.
[4] Lopez Martinez F, Molina-Antonio JD, Paniagua JA, et al. Random forests are used to predict heart disease. 2019 J Med Syst. 43(10):329.
[5]Kumar A, Lamichhane S, Poudel RP, et al. Data mining tools for type 2 diabetes mellitus risk prediction. DiabetesRes.J.2018;2018:1686023.
[6] Ahmed AM, Aljizeeri A, Al-Mallah MH, et al. Type-II diabetes prediction with machine learning methods. J Med Inform. Int. 2014; 83(8):596–604.
[7] Tsanas A, Ramig LO, McSharry PE, Little MA. When nonlinear speech analysis methods are mapped to a common metric, the average severity of Parkinson's disease symptoms can be clinically usefully quantified. J R Soc Interface, 9(65):2756-2764, 2012.
[8] Aggarwal P, Sivaswamy J, Arora S. Parkinson's disease diagnosis made automatically via ensemble machine learning. 2017;21(1):289-299; IEEE Trans Inf Technol Biomed. IRE 1704650
[9]Ahmad F, Khan MK, Hussain M, et al. evaluation of data mining techniques for the prediction of heart disease in comparison. J. Med Syst. 43(4):101 (2019).
[10] Gupta A, Gupta A, Parashar A. strategies for diabetes prediction using machine learning. Emerg Technol Adv Eng. Int J. 2014;4(3):672-675.
Breiman L. Random forests [11]. Mach Learn, 45(1), 2001, pp. 5-32.
[11] Friedman J, Hastie T, and Tibshirani R. Data Mining, Inference, and Prediction: The Foundations of Statistical Learning, 2nd ed., Springer, 2009.
[12] Varoquaux G, Gramfort A, Pedregosa F, et al. Scikit-learn is a Python machine learning tool. Read Res. J Mach.2011;12:2825-2830.
[13] McKinney W, Lamoureux C, van der Walt S, et al. Data structures in Python for statistical computation. In: Python in Science Conference Proceedings, 9th edition, 2010.