IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# HEALTH MONITORING SYSTEM USING MACHINE LEARNING

**SHARAJ OV, SRABANA PRAMANIK**

Student, B.Tech (Computer Science and Engineering), PU, Bengaluru

Department of Computer Science & Engineering, Presidency University, Bengaluru, Karnataka, INDIA

**Abstract**: In this project, we will focus on predicting and providing early detection of chronic diseases, namely diabetes, cardiovascular disease, and stroke. This is a web-based app that features a user-friendly interface. HealthMate employs predictive analysis and machine learning with a range of algorithms, including random forests, decision trees, logistic regression, support vector machines (SVM), and neural networks. Each algorithm has its strengths and weaknesses, thereby making it suitable for predictive analysis and machine learning. My system allows users to register and access the dashboard, providing users with an overview of their health status. The application ensures data protection by implementing a two-step verification code, thereby securing users' personal data. A robust, predictive health monitoring system that is affordable and user-friendly can help patients and healthcare companies. A health monitoring app helps reduce costs, enhances user engagement through means of goal-setting, and promotes overall health and wellness. Healthcare apps obtain information from the users regarding their weight, age, height, BMI, etc. This can be used for the early detection of chronic diseases, hence ensuring patients take proactive measures to monitor their health and fitness.

**Keywords** - Machine learning, Predictive Analysis, Health Monitoring System, Healthcare, Logistic Regression.

## I. INTRODUCTION

The growing demand for Health applications has become increasingly evident in our society, as a rising number of people are frequently ill. The prevalence of global pandemics and chronic diseases has amplified the need for health monitoring apps among the populace. Health prediction is the process of analyzing and predicting health-related issues through user-system interaction. This is essential for early disease detection, diagnosis, prevention, and treatment. With the availability of large-scale health datasets such as electronic health records, researchers can monitor health trends. This app aims to revolutionize the healthcare industry through enhanced predictive capabilities. For predicting diseases, the HealthMate app employs feature engineering techniques to capture complex patterns in the data, thereby improving model performance. Ensemble methods serve a crucial role in enhancing performance and robustness by combining multiple base models to produce a stronger, more accurate predictive model. HealthMate provides users with an interactive interface, thereby empowering them to seize control of their health and take proactive measures, as the saying goes, “Prevention is better than cure”. The multifaceted approach of HealthMate has significant implications for predictive modelling and health management. By leveraging a diverse array of data sources and techniques, HealthMate aims to enhance the accuracy of predictions that account for a wide range of diseases. The predictive techniques are firmly grounded in an array of supervised machine learning methods. As a result, they offer potential benefits to healthcare industries, as they streamline early detection and aid in the decision-making process. Health prediction strategies are classified into subdivisions, namely symptom-level, disease-level, and multi-disease level. Symptom-level disease analyses individual symptoms, while disease-level analyses specific diseases, and multi-level disease analyses on a much broader scale to predict the risk of multiple health diseases.

Health Care apps foster collaboration and partnerships to achieve their objectives and maximize their impact on public health. HealthMate seeks to collaborate with a range of health companies and organizations. By capitalizing on the expertise and network of these partners, HealthMate aims to advance its mission to improve upon and promote wellness in society. By collaborating with healthcare providers, we can access diverse datasets containing valuable information. By capitalizing on this, we can enhance accuracy and disease prediction. Through rigorous testing and validation, we can prevent data imbalance or outliers. The potential impact of collaboration and partnerships is to address complex health challenges, identify innovative solutions, and strengthen health care systems. Continuous evaluation and iteration are fundamental principles for driving success and the evolution of HealthMate. By employing robust evolution metrics, feedback mechanisms, HealthMate continuously assesses the performance of the system. This ensures that HealthMate identifies areas for improvement, prioritizes features, and adjusts design to better align with the user expectations.

HealthMate relies on SQL for storing user data efficiently and securely. These databases play a crucial role in ensuring the security of sensitive health information. Databases are the backbone of every infrastructure, as they facilitate the storage, retrieval, and manipulation of vast amounts of data.

In this project, we will explore how to create a scalable and cost-effective solution for early disease and preventive healthcare.

## II. LITERATURE REVIEW

The rise of chronic illness and the increasing demand for personalized healthcare have pushed the healthcare industry to develop efficient systems. Traditionally, health care was done by doctors visiting the patients and collecting information, but in the present world and future, the need for consultation would diminish. Patients can monitor their health remotely and adjust their needs without the intervention of a doctor.

[1] Jalluri Harshitha Lakshmi “Multiple Disease Prediction System using Machine Learning and Flask Web Framework Application.”

Machine Learning has emerged as a crucial tool in healthcare, addressing issues such as early detection, predictive analytics, personalized treatment planning, and reduced human error. It plays a crucial role in healthcare, as it extracts relevant information and filters out irrelevant information. Traditional systems can lead to false positives. To mitigate this challenge, Random Forest and Support Vector Machines are used to analyse complex data and detect potential illness with high accuracy. Model performance can also be affected by data imbalance, outliers, missing values, and inconsistent measurements. To address this issue, they use data augmentation strategies to enhance data diversity. Users can receive timely status updates, so they can adopt the recommended dietary plans, exercise routine, and lifestyle modification.

Advanced tools such as Flask and HTML & CSS are used to make the web pages interactive, thereby enhancing user experience and adherence to maintaining or adjusting their health by employing goal-setting. This literature employs object-level health prediction as it provides a broader range of applications.

[2] Aryalekshmi, A., et al. "HealthMentor: A Health Monitoring Mobile App Using Advanced Machine Learning Algorithm."

The incorporation of Machine learning has helped improve the functional efficiency of healthcare. Models, with the help of historical and current patient data, assist in identifying high-risk diseases and ensure pre-emptive measures are implemented. By providing access to health information, users can have access to personalized recommendations that empower users to take control of their health and well-being. Machine learning can revolutionize the healthcare system by predicting life-threatening diseases.

This paper delves into the incorporation of machine learning strategies such as support vector machines, random forests, naïve bayes, and feedforward neural networks.

[3] Yagyanath Rimal, et al. “Comparative analysis of heart disease prediction using logistic regression, SVM, KNN, and random forest with cross-validation for improved accuracy.”

This paper delves into the application of machine learning in healthcare, which would help improve the healthcare system at lower costs. It analyses disease prediction using “Logistic Regression, SVM, KNN, and random forest” to improve the accuracy of the predictions. In cross-validation, the data samples are reshuffled to form random subsets.

Logistic Regression is used to predict the outcome based on 1's and 0's. Example: if the patient is at risk of getting heart disease or not. A support vector machine is used primarily for classification, though it can be used for regression. The main goal is to maximize the gap between the two classes. The larger the difference, the better the model performs. KNN works by finding the closest data points. Random forest is a method of averaging multiple deep decision trees.

[4] Sohrab Effati, et al." Web application using machine learning to predict cardiovascular disease and hypertension in mine workers."

Cardiovascular disease and hypertension are the leading causes of mortality worldwide. This paper evaluates the effectiveness of several machine learning algorithms, such as Logistic Regression, Decision Trees, Support Vector Machines, Extreme Gradient Boosting, and Random Forests. Each of these algorithms has strengths and weaknesses. These algorithms establish their effectiveness in healthcare prediction by their special attributes, such as models that can deal with large datasets, achieve high accuracy, and eliminate biases in algorithms. Extreme Gradient Boosting model is an ensemble method that improves prediction accuracy through gradient boosting. This paper also creates a web application for a clear and intuitive interface. The web application is then integrated with the machine learning models, and when the user submits their data, the application transmits the data to an API, which will then perform the prediction using the model and return the outcome of the model.

[5] Cheng-Sheng Yu, et al. "Development of an Online Health Care Assessment for Preventive Medicine: A Machine Learning Approach"

In this paper, machine learning techniques are applied to construct a database system from EMRs. Due to the digitization of healthcare systems, there has been a continuous expansion of information. With the rapid growth of AI in recent decades, every operation that requires data and forecasting is done using AI. The web-based system provides an online diagnosis of the users who submitted their data to the website. The data is then evaluated, trained, and validated to get the desired outcome. Here, missing values are allowed; however, the users are warned of missing variables. Hierarchical clustering is used as it groups data with similarities. The author describes how they constructed an ML health monitoring system, and their aim is to make sure the system is implemented worldwide.

[6] Rutika Bhagat, Prof. Pragati Patil,"Health Monitoring System Using MachineLearning Techniques Algorithm"

In this paper, the author aims to make a machine learning health app using Flask as the backend. The health monitoring system keeps track of patient data by collecting the patient's BMI, age, gender, temperature, blood pressure, etc. The research aims to move away from traditional health monitoring, as it can be cost-effective, but inefficient, as it is more prone to human error. The app is designed so that the patients are regularly updated about their health status. Decision Tree Classifier is used as it produces higher accuracy, but slower compiling. The app intends to collect data securely, perform data analysis, identify health risks, and predict health risks.

[7] Yogesh Kumar, et al. "A Systematic Review of Artificial Intelligence Techniques in Cancer Prediction and Diagnosis"

In this paper, a federated learning model is used for cancer detection. It evaluates 185 studies, applying PRISMA guidelines to analyse AI techniques. It compares with other metrics by applying metrics such as accuracy, sensitivity, specificity, etc. From this paper, we can conclude that CNNs are most common for image-based tasks. Although AI has made a significant contribution to healthcare systems, there are still challenges faced by researchers, such as limited data size and high dimensionality, and uneven distribution of data.

[8] Shadman Nashif, et al. "Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System"

The paper aims to improve decision-making during heart disease detection. It uses data mining to have the best accuracy and performance among Naïve Bayes, SVM, Logistic Regression, Random Forest & ANN. This paper outlines how machine learning can be used to monitor heart disease hardware, such as Arduino or a Sensor. In traditional health monitoring apps, there is a risk of losing the data. To ensure this doesn't occur, we can use a cloud-based health monitoring app that stores the data from the user. This can protect the data from any loss. From this literature, it is clear that SVM, Random Forest, and Logistic Regression showed higher accuracy compared to other models.

[9] Assefa Senbato Genale, Tsion Ayalew Dessalegn, "Developing a Predictive Model for Stroke Disease Detection Using
A Scalable Machine Learning Approach"

Cardiovascular disease causes roughly 17.9 million deaths globally. As more people from lower-income and higher-income countries get sick from chronic illness, this could hurt the economy and healthcare. This study aims to build an accurate and scalable machine learning model to predict stroke using large healthcare datasets. Early diagnosis can greatly reduce mortality, but traditional methods struggle with the volume of datasets. In order to mitigate this issue, in this paper, they use scalable big data analytics (e.g., Spark), as they are better suited to handle large datasets. Even Hadoop can be used to handle large datasets. After doing a comparative study of all the models, the conclusion is that random forest outperformed all the other models. So, this study proves that by integrating big data analytics like Spark, Hadoop, we can improve the efficiency of stroke prediction, as it can handle large datasets.

[10] Asoke Nath, Tanushree Mondal, "Issues and Challenges in Two Factor
Authentication Algorithms"

Internet Technologies have become a significant part of our lives. It is used all over the world, especially in banks, healthcare, industry, etc. User authentication is needed everywhere; without it, we cannot ensure the safety of our data. To mitigate this, we use two-step verification code, which sends OTP to the user who logged in, and only the user can access the information. This prevents phishing attacks and other attacks, ensuring the safety of the user's data. Authentication is done to verify if the person is genuine or not. There are three types of Factor Authentication:

a) Single Factor Authentication:
It is a security process where users verify their identity using only one form of password.
b) Two-Factor Authentication:
It is a process where the users provide two distinct identifications to access the system.
c) Multi-Factor Authentication:
It is a process where users are required to provide two or more verification methods.

In Conclusion, two-step verification is a powerful tool for securing digital systems. However, the cost is high and is still vulnerable to certain attacks.

## III. EXISTING SYSTEM

The rise of AI and the integration of AI into the healthcare system have significantly transformed the health industry landscape. The most impactful innovation is the development of health monitoring web applications. These applications have altered the way health services are monitored and managed. Health applications like HealthMate aid users in early detection, management, and prevention of diseases. The current health monitoring web applications are designed to collect and analyse various health parameters such as blood pressure, glucose level, age, BMI, height, heart rate, and more.
Examples:

- Apple Health: It synchronizes data with wearable devices.
- HealthMate: It performs disease prediction on diabetes or cardiovascular, or stroke.
- Home Workout: It is a home workout app that encourages user participation through means of goal-setting.

These applications offer basic features such as user authentication, appointment booking, and many more. Some of these systems offer risk analysis for conditions like diabetes, cardiovascular disease, stroke, and heart disease by using models such as logistic regression, decision trees, and random forests.
However, like every system, this application has limitations such as Data privacy concerns, poor rural reach, manual input dependency, and fewer incentives.
Example: Apps like Sweatcoin provide you with Sweatcoin for every walk. Sweatcoin can then be converted to bitcoin to earn money, thereby increasing user engagement.
In conclusion, the current system has laid out the groundwork for health monitoring web applications; however, there is still room for improvement in terms of data privacy, user engagement, and incentives.

## IV. PROPOSED SYSTEM

The growing demand for accessible healthcare options has driven the industry to make groundbreaking discoveries in the healthcare system. Cutting-edge technology provides a comprehensive, efficient, and secure platform for users to not only monitor health parameters but also predict potential risks using advanced technologies like machine learning and deep learning.
The objectives of the Proposed System:

- Address Public Health Challenges.
- Enhance healthcare-decision making.
- Develop accurate disease prediction models.
- Patient-Centred Care.
- Health system efficiency.

The Key features of the proposed system:

I. Health Data Input: Users can input data such as blood pressure, glucose level, age, Body Mass Index, and many more
I. Machine Learning: The models are trained to predict early risks of chronic diseases such as diabetes, cardiovascular disease, and stroke.
II. Seamless UI: The application is designed using front-end web technologies such as HTML, CSS, and JavaScript.
III. Two-step Verification: Secure login and two-factor authentication ensure only authorized users can access sensitive data.

Advantages of the proposed system:

- Improved security.
- Real-time alerts
- Increased accuracy

In conclusion, the proposed system provides better accuracy, user experience, and security.

## V. RESEARCH METHODOLOGY

Methodology is the backbone of any research or project. Each step plays a crucial role in building a Health Application platform. This section analyses each aspect of the methodology, highlighting its strategies and implications for the project.

I. Data Collection:
Data Collection lays the foundation for the development of accurate predictive models. The success of predictive models depends on the quality and diversity of the acquired data.

II. Data Preprocessing:
Data Preprocessing ensures the quality, integrity, and effectiveness of predictive models. This section involves a series of operations aimed at cleaning, transforming, and preparing raw data to make it appropriate for analysis and prediction.

- Missing Data: remove duplicates or irrelevant data.
- Normalization and Standardization: Min-Max scaling transforms numerical features to a common range. By this process, we prevent raw data that exhibit different scales from dominating the modelling process.

- Feature Engineering: Build metrics such as daily/weekly patient inflow or turnover.
- Encoding: This step transforms categorical data to numerical formats so that machine learning models can process it.
- Outlier Detection: Remove outliers by using statistical methods, such as z-score or interquartile range.

III. Model Training: This phase involves the selection of machine learning algorithms to learn patterns and relationships to make accurate predictions. We use machine learning algorithms like support vector machines, logistic regression, linear regression, decision trees, and random forests.

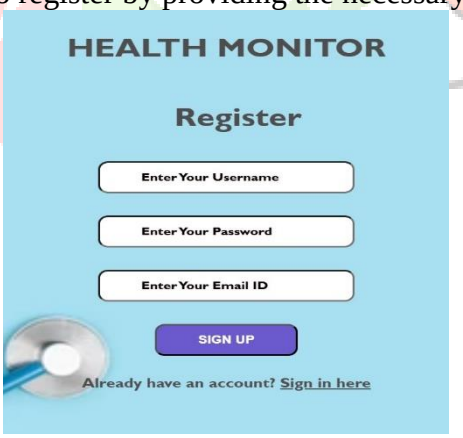
IV. Model Evaluation: This step is essential for assessing the performance of the models on unseen data. By selecting appropriate metrics, such as accuracy, precision, recall, F1 score, AUC-ROC, and AUC-PR, we evaluate the models for better predictions.

V. Visualization:
The dashboard provides predictive insights with respect to the kind of disease and informs the user if they are at risk or not.

This section describes the flow of user interactions:

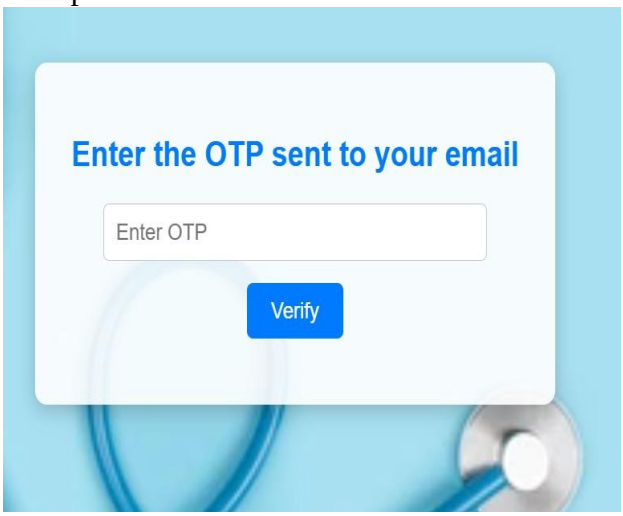
i. Login Page:
Users are required to log in to access their accounts.


ii. Registration Page:
New users are required to register by providing the necessary details.



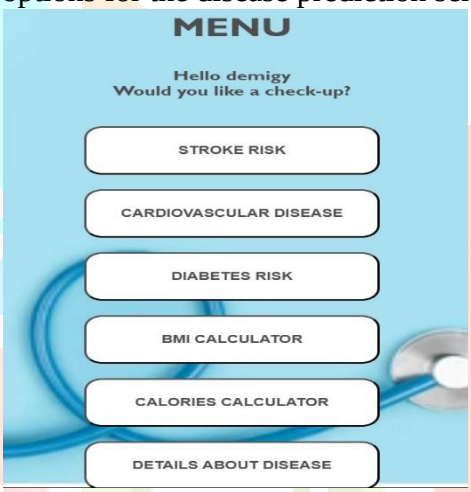
iii. OTP:

After registering, when the users log in to access their accounts, it takes them to another site where they are required to enter the one-time password sent to them via email. This is done to ensure they are genuine.




iv. Home Page:

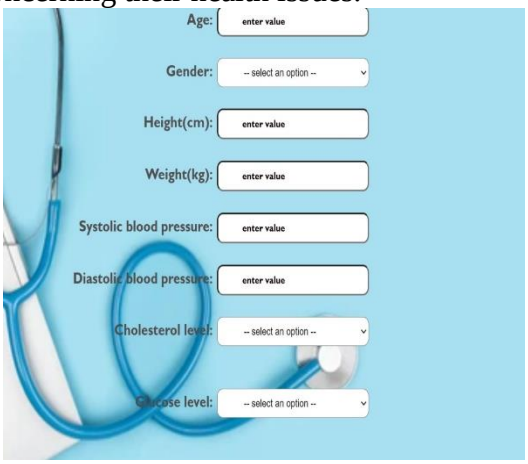
The home page includes options for the disease prediction service.



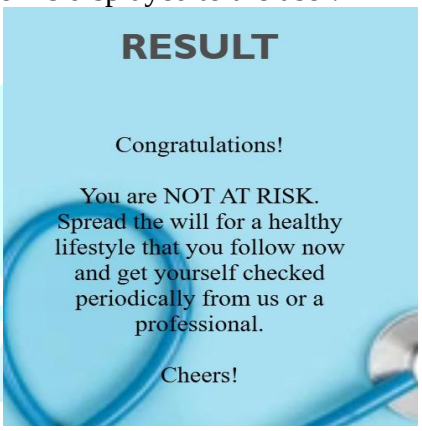

v. Disease Prediction:

Users can select the health service they need, like stroke, diabetes, or cardiovascular disease.

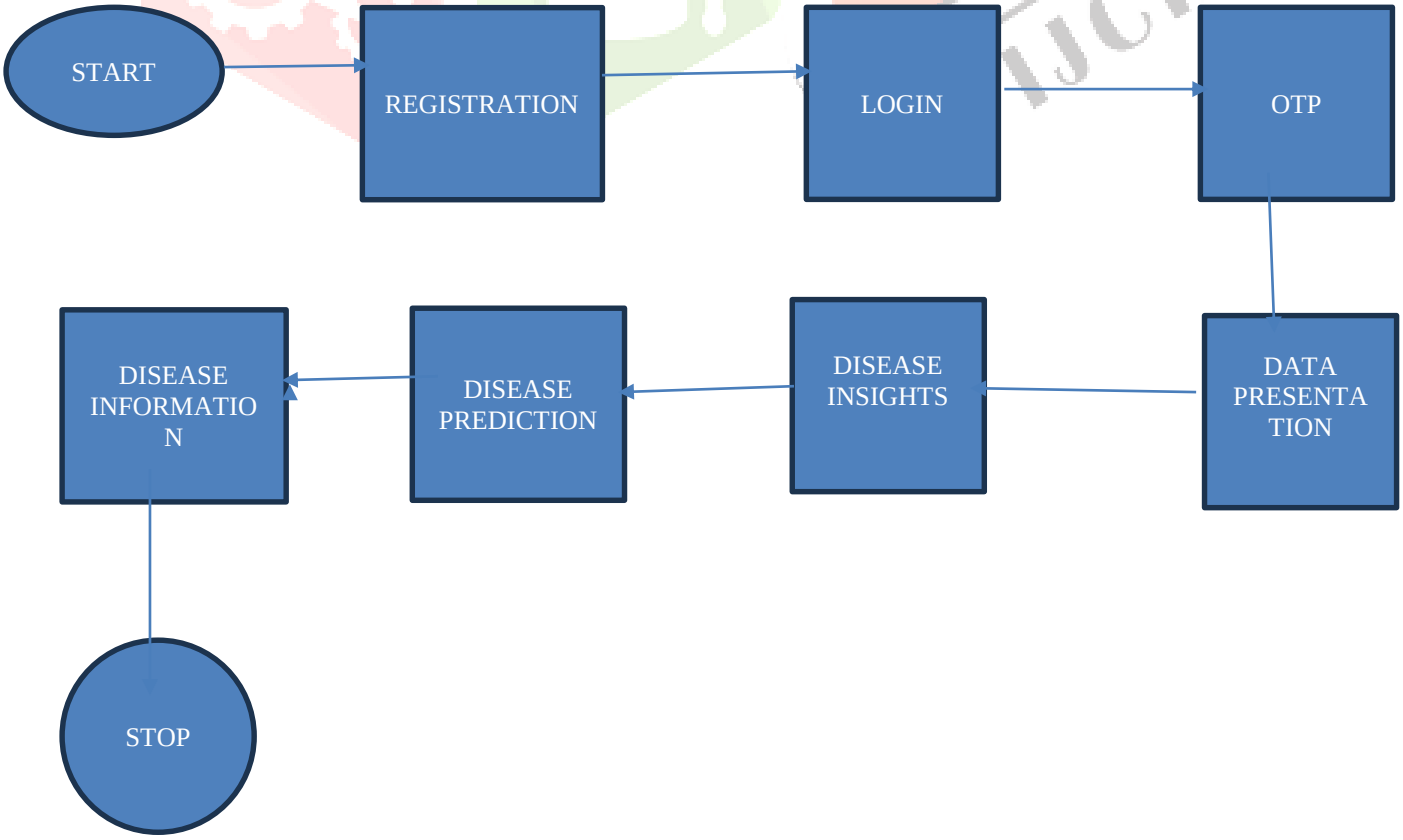
vi. Input Form:
Users fill out the form concerning their health issues.




vii. Output Page:
The result of the prediction is displayed to the user.




viii. Database Update:
The input data is updated in the database and can be used for future reference.



**Technology Stack Used:**

A. Python:
Python is a high-level programming language known for its versatility and wide range of applications.

B. Pandas:
Pandas is a library for Python. It uses data structures such as Data Frame Series to manage structured data.

C. Numpy:
Numpy is employed to manipulate, clean, and analyse health-related datasets.

D. Flask:
Flask is a lightweight and versatile web framework for Python, commonly used for web applications.

E. SQL:
SQL (Structured Query Language) is a standard language used in storing and manipulating relational databases.

F. HTML:
HTML (Hyper Text Markup Language) is the basic building block of all web pages. It provides user interactivity, which can enhance user engagement.

G. CSS:
CSS (Cascading Style Sheets) is utilized for styling and formatting the user interface.

H. Matplotlib:
Matplotlib is a prominent Python tool used for creating static, animated, and interactive visualizations like Line Plots, bar graphs, scatter plots, etc.

## VI. Conclusion

The HealthMate project represents an innovative solution designed to empower individuals in managing their health and wellness. HealthMate offers a diverse approach to promoting health awareness among individuals.

By leveraging machine learning techniques, healthaMate provides a personalized risk assessment and recommendations, and early risk detection. Machine learning tools will identify patterns in patient admission and estimate the likelihood of sickness.

HealthMate provides precise predictions that can assist hospitals in minimizing operational costs and establish an accessible and robust public healthcare system for all. This will not only resolve challenges but also strengthen the trust and efficiency of the public healthcare system.

## VII. Acknowledgment

I would like to express my sincere gratitude to Dr.Srabana Pramanik, Assistant Professor, Presidency University, for their constant guidance, encouragement, and support throughout this research.

I also thank the faculty and staff of the Department of Computer Science & Engineering at Presidency University for providing necessary assistance.

## VIII. References

[1] Jalluri Harshitha Lakshmi, “Multiple Disease Prediction System using Machine Learning and Flask Web Framework Application”, publications at researchgate, 2025

[2] Aryalekshmi, A., et al. "HealthMentor: A Health Monitoring Mobile App Using Advanced Machine Learning Algorithm.", publications at IEEE, 2024

[3] Yagyanath Rimal, et al. “Comparative analysis of heart disease prediction using logistic regression, SVM, KNN, and random forest with cross-validation for improved accuracy.”, publications at Scientific Reports (Nature portfolio), 2025

[4] Sohrab Effati, et al.” Web application using machine learning to predict cardiovascular disease and hypertension in mine workers.”, publications at Scientific Reports (Nature portfolio), 2024

[5] Cheng-Sheng Yu, et al. “Development of an Online Health Care Assessment for Preventive Medicine: A Machine Learning Approach”, publications at jmir, 2020

[6] Rutika Bhagat, Prof. Pragati Patil,” Health Monitoring System Using Machine Learning Techniques Algorithm”, publications at ijraset, 2023

[7] Yogesh Kumar, et al. “A Systematic Review of Artificial Intelligence Techniques in Cancer Prediction and Diagnosis”, publications at Springer, 2021

[8] Shadman Nashif, et al. “Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System”, publications at *ijraset, 2022*

[9] Assefa Senbato Genale, Tsion Ayalew Dessalegn, “Developing a Predictive Model for Stroke Disease Detection Using a Scalable Machine Learning Approach”, publications at Hindawi, 2025

[10] Asoke Nath, Tanushree Mondal, “Issues and Challenges in Two Factor Authentication Algorithms”, published at IJLTET, 2016