# AI Platform For Predicting The Risk of Having NCDs

Ariane Shimirwa, Claudine Mahoro, Tyson Ntwari Muvunyi
Carnegie Mellon University | Master's of Science in Information Technology

## Abstract

Non-communicable diseases (NCDs) are chronic medical conditions not caused by infectious agents and primarily associated with lifestyle factors. This paper highlights the significant impact of non-communicable diseases (NCDs) on global public health, particularly in low- and middle-income countries where they account for most premature deaths. A privacy-preserving AI platform is proposed as a solution to predict the risks of having NCDs, such as cardiovascular diseases and type 2 diabetes, using clinical data and advanced machine learning algorithms to reduce the number of deaths that occur yearly. The platform alerts individuals with high risks and provides them with further medical consultation options.

## Introduction

Non-communicable diseases (NCDs), also known as chronic diseases, are long-lasting conditions that develop slowly and are not caused by infectious agents. They include cardiovascular diseases, cancer, chronic respiratory diseases, diabetes, and mental health disorders, among others. NCDs require ongoing treatment and care, affecting long-term health. Lifestyle factors such as unhealthy diets, physical inactivity, tobacco use, excessive alcohol consumption, and environmental factors contribute to the development of these diseases. NCDs have become a major public health issue, particularly in low and middle-income countries. They are responsible for a significant number of deaths worldwide, with 17 million individuals dying prematurely each year before the age of 70. Around 86% of these deaths occur in low and middle-income countries [1]. To address this growing health concern, it is crucial to prioritize early detection, prevention, and management strategies for NCDs.

This paper presents a privacy-preserving AI platform for predicting the risk of non-communicable diseases (NCDs) like cardiovascular and type 2 diabetes. The platform utilizes supervised learning algorithms and a Machine Learning framework to uncover hidden patterns and relationships in the data. By employing techniques such as Random Forest, Support Vector Machine, Naïve Bayes, and Artificial Neural Networks, it analyzes clinical data to predict the risks of diabetes or heart disease. The proposed platform not only enhances NCD risk prediction but also ensures privacy protection. It notifies individuals at high risk and offers the option to schedule appointments with medical professionals for further consultation.

## Methods and Materials

The methodology of this study includes several steps: data collection, pre-processing, exploratory analysis, model building, mobile application development, and security implementation. The data was gathered from online sources and underwent pre-processing to ensure its quality. Exploratory analysis was conducted to gain insights from the data. **Chart1** . illustrates that individuals around the age of 60, especially males, were more prone to heart disease risks, and **Chart 2**. illustrates more females having diabetes disease between the age of 40 and 50. This observation aligns with the research from Medicine in Novel Technology and Devices, which indicates that men have a higher risk of heart disease at an early age compared to women, whose risks tend to increase after menopause [2]. To build predictive models, forward stepwise regression was employed to select relevant features. Four supervised models, namely Random Forest, Support Vector Machine, Naïve Bayes, and Artificial Neural Network were developed for each dataset.

The models developed in this study were evaluated using classification metrics such as accuracy, precision, recall, and F1 score to assess their performance. Furthermore, a mobile application was created using Flutter for the app development, a Postgres database for storing patient records, and Fast API for backend services. The mobile application enables users to input their health metrics, which are then utilized by the machine learning algorithm to predict their risk levels. To ensure security and privacy, various measures were implemented, including encryption and hashing of user credentials, authentication verification, and password protection. These measures were put in place to safeguard users' personal information and maintain data privacy.

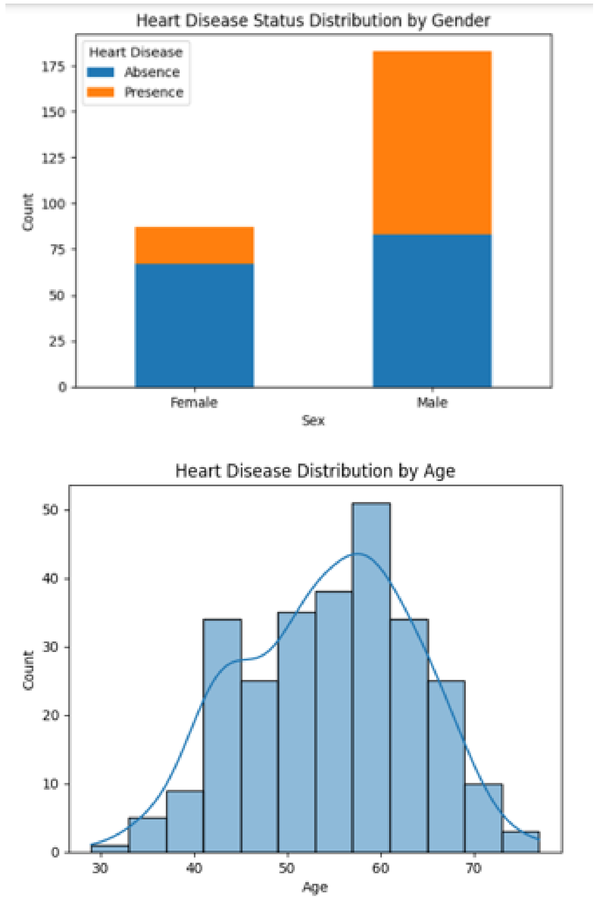


**Chart 1.** Bar graphs showing heart disease distribution by gender and age.

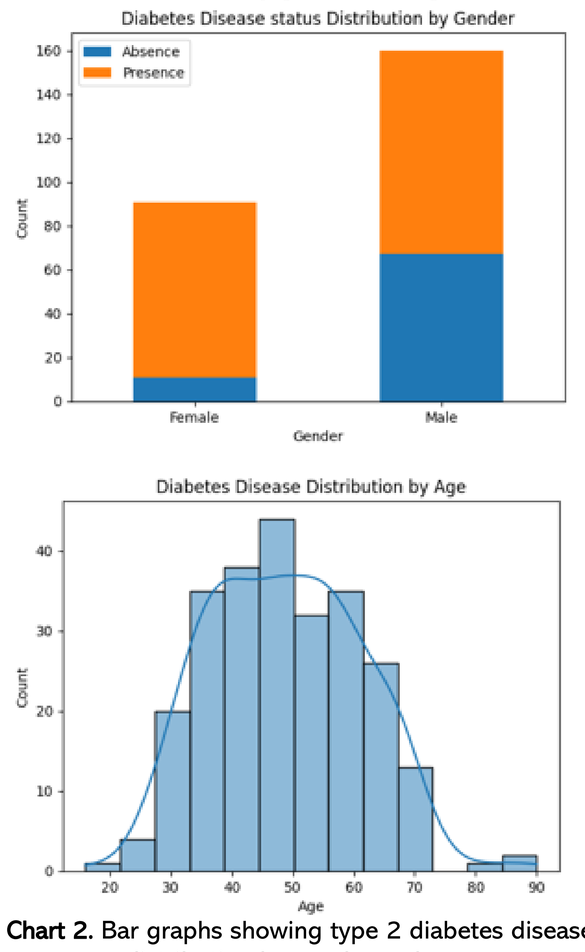


**Chart 2.** Bar graphs showing type 2 diabetes disease distribution by gender and age.

## Results

In the modeling phase, the support vector machine model demonstrated superior performance in predicting cardiovascular disease with an accuracy of 92.59% and the Artificial Neural Network model outperformed other models in predicting type 2 diabetes, achieving an accuracy of 92.15%. As a result, these two models, SVM and ANN, were saved and integrated into the mobile application. The mobile application was successfully developed, allowing users to input their diagnostic data through the user interface and utilize the models for risk prediction (**see Fig 1**). The application displays the results back to the user, providing an evaluation of their risk for the corresponding non-communicable disease (NCD). See

Additionally, the mobile application enables users to promptly schedule an appointment with a specialist for further diagnosis (**see Fig. 2**). The booking system within the application facilitates the process of reserving a medical appointment at a hospital. If the NCD prediction indicates high risk, the user can choose to immediately make an appointment with their preferred hospital or reserve it for a later date. This functionality enhances the user experience by providing seamless access to specialized medical care based on their predicted NCD risk.

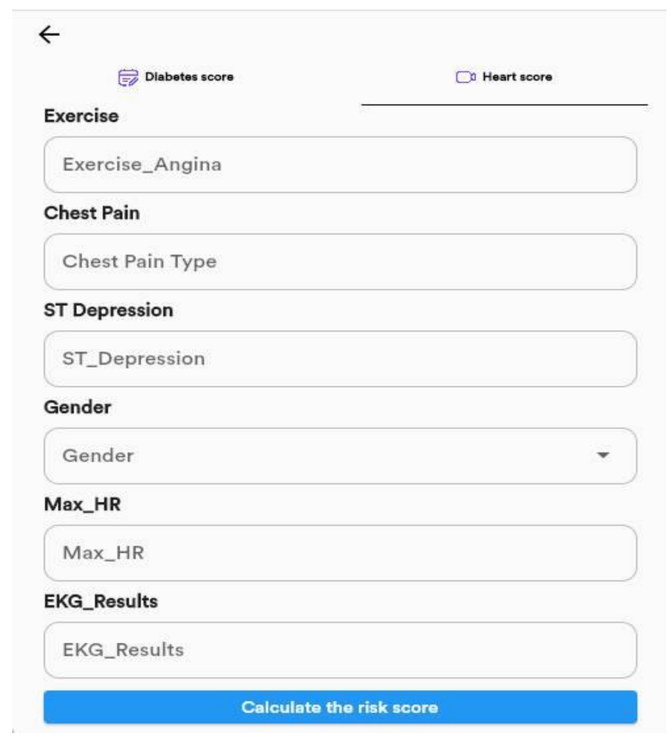


**Fig1.** Interface for heart disease risks prediction

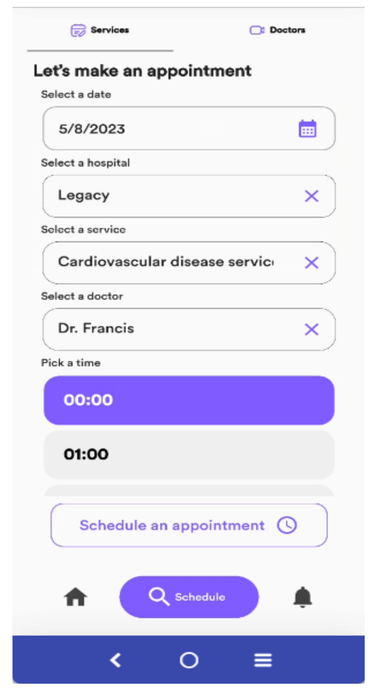


**Fig2.** making an appointment by a user

## Conclusions

In this study, a smartphone application was developed to address the challenge of early detection of diabetes and cardiovascular diseases. Clinical data from an online source was collected, pre-processed, and explored. Feature selection techniques were applied to select relevant variables for prediction, and various machine learning algorithms, including Random Forest, Support Vector Machine, Naive Bayes, and Artificial Neural Network, were trained on each dataset. The models' performance was evaluated using classification metrics, and Support Vector Machine and Artificial Neural Network emerged as the top performers, achieving high accuracies of 92.59% and 92.15% respectively for heart disease and diabetes prediction. These models were integrated into the smartphone application, built using Flutter, allowing users to input their information and receive predictions for diabetes or heart disease. When positive results are obtained, the application prompts the user to book an appointment at a desired hospital for further medical consultation, while negative results encourage the user to maintain a healthy lifestyle to prevent the development of non-communicable diseases in the future.

## Recommendation

Based on the findings presented in the paper, the paper recommends healthcare organizations to adopt AI-powered platforms for predicting non-communicable diseases, with a focus on cardiovascular diseases and diabetes. To improve prediction accuracy, the use of real clinical datasets and wearable sensors is suggested. Specialized sensors, such as Continuous Glucose Monitoring devices and smartwatches, can provide real-time patient data for more efficient screening. Exploring different machine learning algorithms and techniques is also advised. Future research should expand the application of AI platforms and wearable sensors to include other non-communicable diseases, like cancer and respiratory diseases, to enhance disease detection and prevention and ultimately improve patient outcomes.

## Contact

Ahmed Biyabani
Associate Teaching Professor
Carnegie Mellon University Africa
Email: ab3x@andrew.cmu.edu
Phone: + 250 791 569 905

## References

1. "Non-communicable diseases," World Health Organization, September 2022.
2. Z. C. A. S. X. D Zujie Gao, "Gender differences in cardiovascular disease," Medicine in Novel Technology and Devices, vol. 4, no. ISSN 2590-0935, 2019.
3. Health Care Quality Indicators - cardiovascular disease and diabetes - OECD, https://www.oecd.org/els/health-systems/hcqi-cardiovasculardisease-and-diabetes.htm (accessed May 8, 2023).