IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

"MEDIFORECAST: MULTIPLE DISEASE PREDICTION"

[1]Vanita Gadekar, [2]Tejas Vishnu Gund, [3]Harshwardhan Somnath Bhadange,
[4]Nilesh Ravindra Patil, [5]Aniket Sanjay Shirke

Department of Computer Engineering

Smt. Kashibai Navale College of Engineering, Pune, Maharashtra

Abstract: Early disease detection is paramount for effective healthcare management. In this research, titled "MediForecast: Multiple Disease Prediction," we address this critical challenge by harnessing the power of machine learning, specifically employing the Support Vector Machine (SVM) classifier algorithm. Focusing on heart disease, Parkinson's disease, and diabetes, we explore innovative approaches to predict these conditions accurately. Our methodology involves meticulous data collection, preprocessing, and feature selection tailored for each disease. We employ the SVM classifier to create robust prediction models. Our implementation demonstrates the practical application of these models, showcasing their effectiveness in diagnosing the aforementioned diseases. The results reveal promising outcomes, indicating high accuracy, sensitivity, and specificity in disease prediction. By empowering medical professionals with timely and precise predictive capabilities, our research contributes significantly to the advancement of healthcare practices. We highlight the transformative potential of machine learning, particularly the SVM classifier, in revolutionizing disease diagnosis, paving the way for a healthier future.

Index Terms **– Diabetes, Heart, Parkinsons, SVM Classifier**

1.INTRODUCTION

In the contemporary landscape of healthcare, the ability to predict diseases accurately and at an early stage holds immense promise for improving patient outcomes and reducing the burden on healthcare systems. The convergence of advanced data analytics and machine learning techniques offers a transformative avenue for achieving these goals. Our project, titled "MediForecast: Multiple Disease Prediction," delved into the realm of predictive healthcare by focusing on three significant and prevalent diseases: heart disease, Parkinson's disease, and diabetes. Heart disease, Parkinson's disease, and diabetes pose substantial challenges to global healthcare. Timely diagnosis is crucial, as early intervention not only improves patient prognosis but also reduces healthcare costs. Traditional diagnostic methods, while valuable, often face limitations in terms of speed and accuracy. Machine learning, with its ability to analyze vast datasets and discern intricate patterns, has emerged as a potent tool in the early detection of diseases. The primary objective of our project was to harness the potential of machine learning algorithms to predict heart disease, Parkinson's disease, and diabetes with a high degree of accuracy. By integrating diverse datasets and employing sophisticated predictive models, we aimed to create a comprehensive and reliable system, MediForecast. This system would enable healthcare professionals to anticipate these diseases in their early stages, facilitating prompt and personalized interventions. Through this research, we aimed to significantly contribute to a paradigm shift in healthcare that combines data-driven insights and machine learning to redefine disease prediction and patient care.

1.1.Description

In the contemporary landscape of healthcare, the fusion of cutting-edge technology and innovative methodologies has catalyzed a paradigm shift in disease prediction and diagnosis. Our project stands at the forefront of this transformative journey, pioneering the development of a robust system for multiple disease prediction using state-of-the-art machine learning algorithms. Rooted in the principles of data-driven healthcare and predictive analytics, our project endeavors to harness the vast potential of machine learning to empower individuals with proactive insights into their health. By leveraging Python-based tools and libraries such as Scikit-learn and Streamlit, we aim to create a dynamic and intuitive platform that not only facilitates early disease detection and risk assessment but also provides personalized healthcare recommendations tailored to individual needs and preferences. Through the seamless integration of advanced technology and healthcare expertise, our system seeks to redefine the boundaries of predictive healthcare, ushering in a new era of proactive health management and improved patient outcomes.

1.2.Problem Statement

Despite the advancements in medical science, early detection and prevention of diseases remain a significant challenge. Many individuals are unaware of their susceptibility to various health conditions until symptoms manifest or the disease progresses to an advanced stage. This lack of proactive healthcare management not only results in increased healthcare costs and reduced quality of life but also contributes to the burden on healthcare systems worldwide. Moreover, factors such as demographic disparities, limited access to healthcare resources, and insufficient awareness further exacerbate the challenges associated with disease prevention and management. Thus, there is a pressing need for accessible and accurate disease prediction systems that empower individuals to take control of their health, bridge healthcare disparities, and drive positive health outcomes for populations globally.

1.3.Proposed System

The proposed system for multiple disease prediction is a robust and versatile solution designed to accurately assess the likelihood of various diseases including heart disease, diabetes, and liver disease based on user-provided health parameters. It will feature an intuitive interface where users can input relevant data such as medical history, symptoms, and test results. The system will then employ machine learning algorithms like SVM classifier, KNN, and Random Forest, trained on historical healthcare data, to process this information and generate predictions for each disease. The backend infrastructure will host the machine learning models responsible for processing user input and producing predictions. Secure APIs will facilitate seamless communication between the frontend and backend components of the system, ensuring efficient data exchange while adhering to strict security protocols such as data encryption and user authentication. The frontend interface will prioritize user experience, offering a user-friendly environment for inputting health parameters and receiving clear, actionable feedback on disease predictions. Comprehensive testing will be conducted to validate the accuracy and reliability of the system before deployment. Extensive documentation and training resources will be provided to support users, administrators, and support staff in effectively utilizing and maintaining the system. Overall, the proposed system aims to provide a powerful tool for disease prediction, assisting healthcare professionals and individuals alike in making informed decisions about health management and treatment strategies.

2.LITERATURE REVIEW

1.According to the paper, diabetes is one of the dangerous diseases in the world , it can cause many varieties of disorders which includes blindness etc. In this paper they have used machine learning techniques to find out diabetes disease as it is easy and flexible to forecast whether the patient has illness or not . Their aim of this analysis was to invent a system that can help the patient to detect the diabetes disease of the patient with accurate results. Here they used mainly 4 main algorithms Decision Tree , Naïve Bayes , and SVM algorithms and compared their accuracy which is 85%,77%, 77.3% respectively . They also used ANN algorithm after the training process to see the reactions of the network which states whether the disease is classified properly or not . Here they compared the precision recall and F1 score support and accuracy of all the models[1] .

2. The main aim of the paper is, asthe heart plays an important role in living organisms. So, the diagnosis and prediction of heart related disease should be perfect and correct because it is very crucial which can cause death cases related to heart .So, Machine learning and Artificial Intelligence supports in predicting any kind of natural events .So in this paper they calculate accuracy of machine learning for predicting heart disease using k-nearest neighbor ,decision tree, linear regression and SVM by using UCI repository dataset for training and testing . They also compared the algorithm and their accuracy SVM 83 %,Decision tree 79%,Linear regression 78%,k-nearest neighbor 87%[2].

3. The system defines that liver diseases are causing a high number of deaths in India and is also considered as a life threatening disease in the world. As it is difficult to detect liver disease at an early stage .So using automated programs using machine learning algorithms we can detect liver disease accurately .They used and compared SVM ,Decision Tree and Random Forest algorithms and measured precision, accuracy and recall metrics for quantitative measurement. The accuracy is 95%,87%,92% respectively[3]

3.SYSTEM ANALYSIS

3.1 Functional Requirements

- The system allows patients to register and create user accounts to access disease prediction functionalities.
- Users can input a variety of health parameters including but not limited to age, gender, weight, blood pressure, cholesterol levels, and glucose levels.
- The system provides a user-friendly interface for patients to input their health parameters efficiently and accurately.
- Upon submission of health parameters, the system utilizes machine learning algorithms to predict the likelihood of developing specific diseases such as diabetes, heart disease, and Parkinson's disease.
- Users receive comprehensive disease predictions including the probability score and risk assessment based on their input parameters.

3.2 Non-Functional Requirements

- The website should provide real-time feedback and guidance to users while inputting their health parameters to ensure data accuracy and completeness.
- The system must adhere to data privacy regulations such as HIPAA to ensure the confidentiality and security of user information.
- The website should be responsive and accessible across various devices including desktops, laptops, tablets, and smartphones to accommodate diverse user preferences and needs.
- The system should be scalable to handle a large volume of user interactions and predictions without compromising performance or responsiveness.
- The website's design and layout should prioritize user experience, with intuitive navigation, clear instructions, and visually appealing graphics to enhance usability and engagement.
- The system should undergo regular maintenance and updates to address bugs, improve performance, and incorporate new features or enhancements based on user feedback and evolving healthcare trends.

4.DESIGN

The structured process begins with patient registration, where individuals create accounts on the platform by providing necessary information such as name, email, and password. Once registered, patients input their health parameters, including age, gender, weight, blood pressure, cholesterol levels, and glucose levels. This data is validated and securely stored in the system's database. Upon retrieval of the stored health parameters, the system proceeds to calculate the probability and possibility of developing specific diseases.

Using machine learning algorithms, the system analyzes the data to identify patterns and correlations indicative of disease risk.

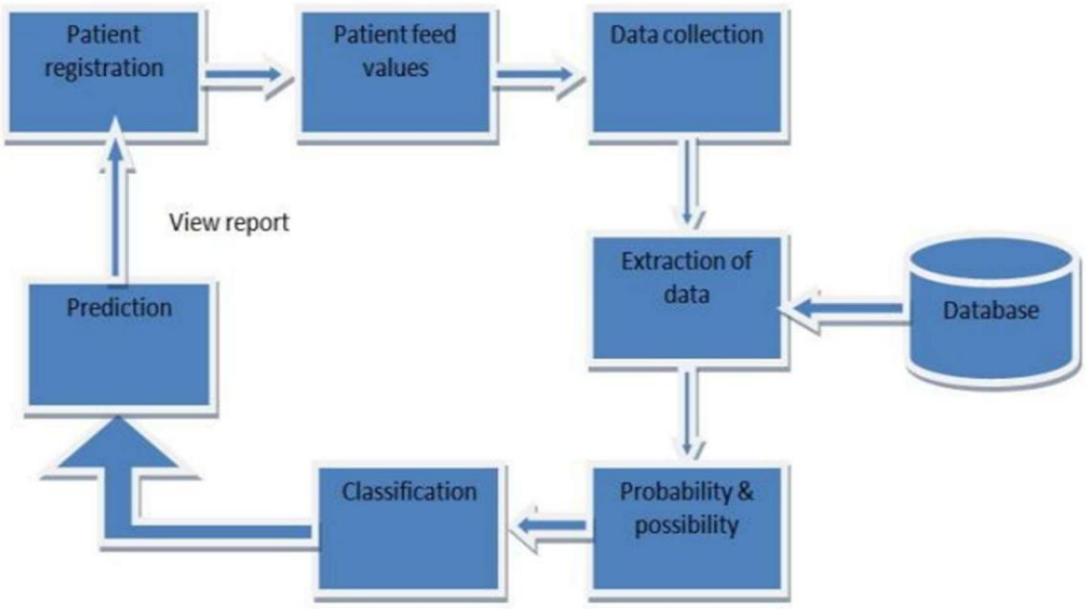

Figure 4.1: System Architecture

Based on these analyses, patients are classified into different risk categories for each disease, ranging from low to high risk. Using the classification results, the system predicts the likelihood of patients developing particular diseases such as diabetes, heart disease, or Parkinson's disease. These predictions provide valuable insights into patients' health status, enabling proactive disease management and prevention strategies. Finally, patients can access detailed reports summarizing their disease predictions and risk assessments. These reports include graphical visualizations, probability scores, risk categories, and personalized recommendations for lifestyle modifications or preventive measures. Patients can review their reports to gain a better understanding of their health status and make informed decisions about their healthcare journey.

In this process, the patient interacts with various components of the system. Initially, the patient registers by providing their personal information to the registration system, which then stores this data in the database and confirms successful registration. Next, the patient inputs their health parameters into the health data input system, which validates and stores this information in the database.

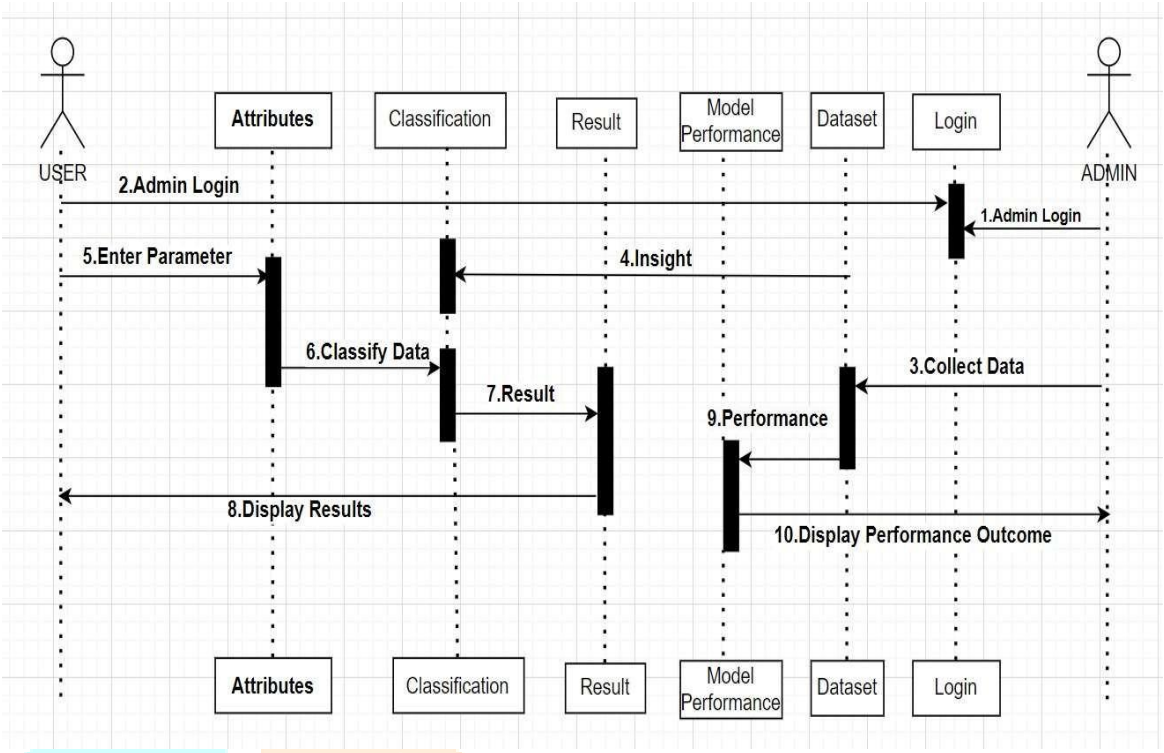

Fig. 4.2 Sequence Diagram

The machine learning analysis system then retrieves these health parameters from the database to analyze the data and classify the patient's risk level for various diseases. Based on this classification, the system predicts disease likelihood and generates detailed reports. These reports are stored in the database and made accessible to the patient, who can then review the findings and make informed healthcare decisions based on the provided insights. This entire sequence ensures a comprehensive and secure process for health data analysis and disease prediction.

5.SYSTEM IMPLEMENTATION

The multiple disease prediction project follows a structured implementation approach, beginning with data collection and preprocessing to ensure the integrity of the datasets for heart disease, diabetes, and liver disease prediction. Machine learning algorithms, including SVM classifier, KNN, and Random Forest, are then developed and trained on the preprocessed data. Rigorous evaluation criteria are applied to select the most optimal model for each disease prediction task. Subsequently, the backend and frontend components are meticulously crafted to enable seamless interaction between users and the system. Security measures are rigorously implemented, and comprehensive testing is conducted to guarantee reliability and accuracy. Following successful testing, the system is deployed with extensive training and documentation provided to users and maintenance staff. This structured approach aims to ensure effective utilization and long-term sustainability of the multiple disease prediction system, offering valuable insights for healthcare professionals and individuals alike. By leveraging advanced machine learning techniques, the system seeks to enhance healthcare decision-making and contribute to improved patient outcomes.

1. Module 1 (Diabetes Disease Prediction)

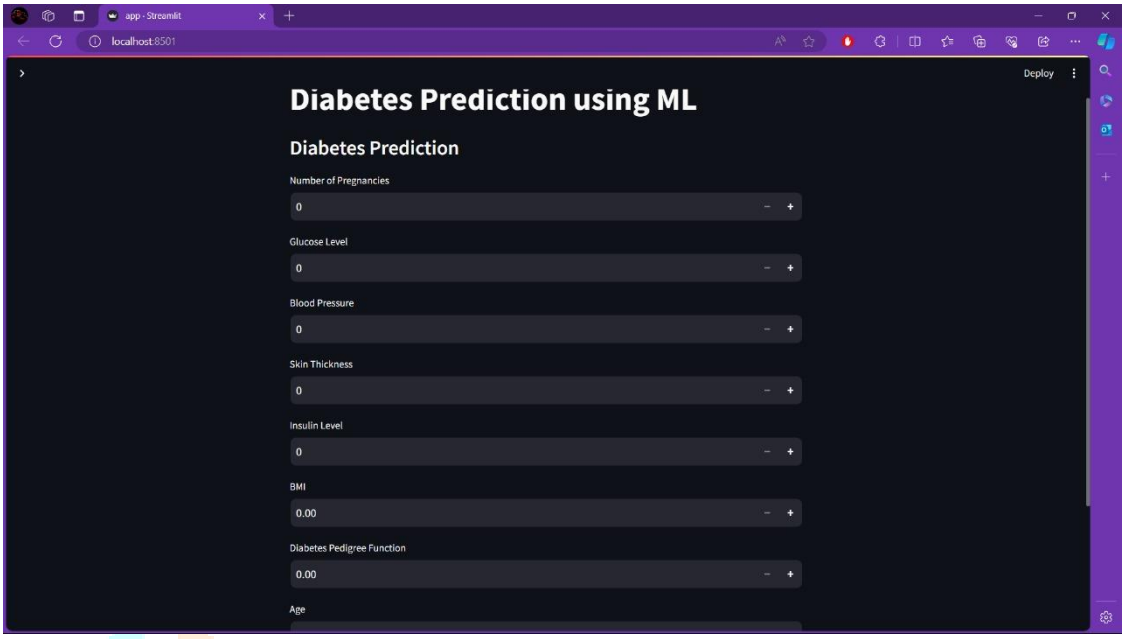

Fig 6.2 Diabetes Disease Prediction

2. Module 2 (Heart Disease Prediction)

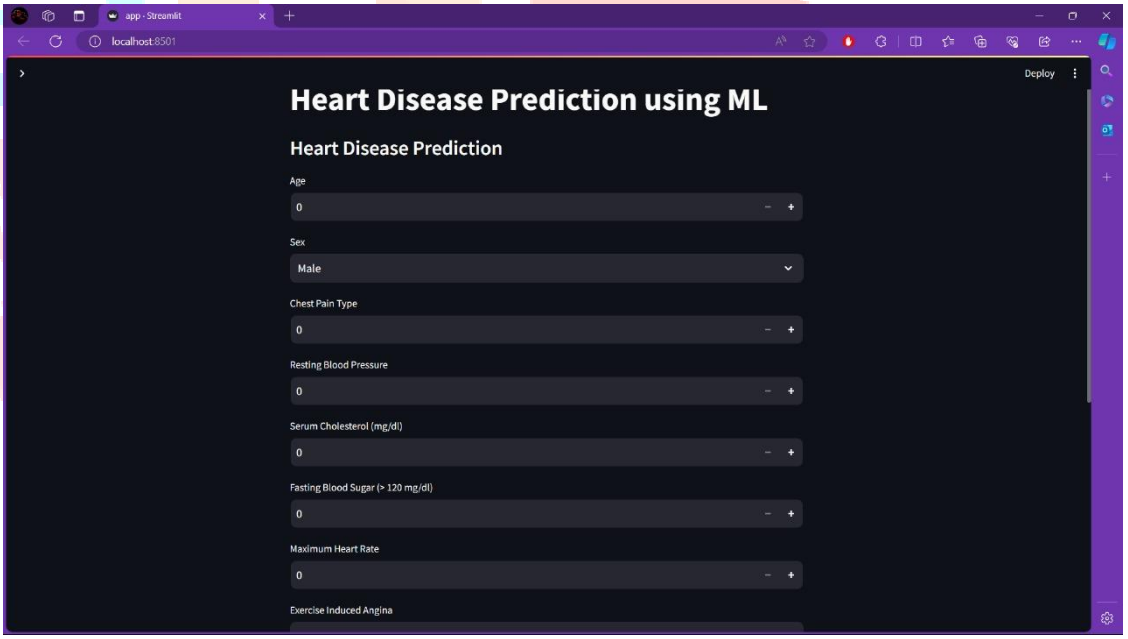

Fig 6.3 Heart Disease Prediction

3. Module 3 (Parkinson's Disease Prediction)

Fig 6.4 Parkinson's Disease Prediction

6.RESULT

Above SVM classifier is used and their performances are compared with respective of the others

Techniques	Accuracy
Decision Tree	86.70
Arithmetic Neural Network	90.53
Random Forest	91.06
K-Nearest Neighbor	88.00
Support Vector Machine	95.10

By comparing all of them, DT(Decision Tree) classifier reached the lowest classification accuracy. From Table, it can be concluded that the DT performed the poorest overall, whereas SVM performed the best among all machine learning classifiers, and RF (Random Forest) performs the good overall from all of the classifiers.

In our disease prediction system, we're using the SVM classifier algorithm to predict three common diseases. When patients input their health parameters, the system guides them by showing acceptable value ranges for each parameter. If an input falls outside this range or is invalid, a warning prompts the patient to correct it. The SVM algorithm then analyzes these inputs to predict the likelihood of the selected disease. Patients receive clear feedback on whether they're likely to have the disease based on their inputs, along with personalized recommendations for disease management. By combining accurate algorithms with user-friendly interfaces, our system empowers individuals to take proactive steps towards managing their health effectively.

7. CONCLUSION

In conclusion, our project has successfully addressed the challenge of developing a practical system for multiple disease prediction, leveraging machine learning algorithms and data preprocessing techniques. The collaborative efforts have resulted in the creation of a robust and user-friendly platform capable of aiding healthcare professionals and individuals in making informed decisions regarding disease management. Moving forward, there is significant scope for future work to enhance the system's capabilities, including the

incorporation of real-time data updates, integration with electronic health records, and the expansion of disease prediction models to cover a broader range of health conditions. These advancements will further improve the system's accuracy and usability, ultimately contributing to more effective healthcare delivery and patient outcomes. Overall, our project represents a significant step towards leveraging technology to improve healthcare outcomes and empower individuals to take proactive steps towards better health.

8.ACKNOWLEDGMENT

We extend our sincere appreciation to Smt. Kashibai Navale College of Engineering for their steadfast support and provision of resources throughout the implementation of our project on multiple disease prediction. Their commitment to fostering innovation and academic excellence has been pivotal to our success. We also express our gratitude to Prof. Vanita Gadekar, our esteemed project guide, whose guidance, mentorship, and expertise were invaluable in shaping the direction and execution of our project. Her insights and dedication have been instrumental alongside the collaborative efforts of the research team. Furthermore, we extend our thanks to all members of the project team for their dedication and contributions to the successful implementation of the project. Each team member's commitment and expertise have played a vital role in achieving our objectives. Lastly, we acknowledge and appreciate the support received from individuals and institutions who have contributed in various capacities, facilitating the smooth execution of the project. Their assistance has been invaluable in advancing our efforts to develop a practical solution for disease prediction.

9.REFERENCES

[1] A. Singh, R. Kumar, “Heart Disease Prediction Using Machine Learning Algorithms,” in IEEE International Conference on Electrical and Electronics Engineering (ICE3), 2020.

[2] A. Sivasangari, B. J. K. Reddy, A. Kiran, P. Ajitha, “Diagnosis of Liver Disease Using Machine Learning Models,” in Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2020.

[3] P. Sonar, K. J. Malini, “Diabetes Prediction Using Different Machine Learning Approaches,” in IEEE 3rd International Conference on Computing Methodologies and Communication (ICCMC), 2019.

[4] G. Wolgast, C. Ehrenborg, A. Israelsson, J. Helander, E. Johansson, H. Manefjord, “Wireless Body Area Network for Heart Attack Detection [Education Corner],” IEEE Antennas and Propagation Magazine, vol. 58, no. 5, pp. 84-92, 2016.

[5] R. Shinde, S. Arjun, P. Patil, J. Waghmare, “An Intelligent Heart Disease Prediction System Using K-Means Clustering and Naïve Bayes Algorithm,” International Journal of Computer Science and Information Technologies, vol. 6, no. 1, pp. 637-639, 2015.

[6] S. Bashir, U. Qamar, M. Y. Javed, “An Ensemble-Based Decision Support Framework for Intelligent Heart Disease Diagnosis,” in International Conference on Information Society (i-Society 2014), pp. 259-264, IEEE, November 2014.

[7] S. H. Jee, Y. Jang, D. J. Oh, B. H. Oh, S. H. Lee, S. W. Park, Y. D. Yun, “A Coronary Heart Disease Prediction Model: The Korean Heart Study,” BMJ Open, vol. 4, no. 5, e005025, 2014.

[8] A. Ganna, P. K. Magnusson, N. L. Pedersen, U. de Faire, M. Reilly, J. Ärnlöv, E. Ingelsson, “Multilocus Genetic Risk Scores for Coronary Heart Disease Prediction,” Arteriosclerosis, Thrombosis, and Vascular Biology, vol. 33, no. 9, pp. 2267-2272, 2013.

[9] M. A. Jabbar, B. L. Deekshatulu, P. Chandra, "Heart Disease Prediction Using Lazy Associative Classification," in International MultiConference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), pp. 40-46, IEEE, March 2013.

[10] C. S. Dangare, S. S. Apte, "Improved Study of Heart Disease Prediction System Using Data Mining Classification Techniques," International Journal of Computer Applications, vol. 47, no. 10, pp. 44-48, 2012.