IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AI -Driven Prediction Of Cardiovascular Complications In Diabetic Patients

R Kumar[1], S Sujithraj[2], P Vasan[3], J Vijay Pradeep[4],

[1] R Kumar, Associate Professor, CSE, Sri Ramakrishna Institute of Technology, Coimbatore, India.
[2] S Sujithraj, CSE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, India
[3] P Vasan, CSE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, India.
[4] J Vijay Pradeep, CSE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, India

Abstract:
Diabetes are widespread health issues around the world. Both Insulin Dependent (Type 1) Diabetes Mellitus and (Type 2) Diabetes Mellitus are prevalent, with Type 2 diabetes being more common and often associated with lifestyle and diet. The prevalence of heart disease varies by region, but it affects people of all ages, genders, and backgrounds. Neural networks techniques have been applied to diabetes in various ways to assist in diagnosis, management, and treatment. Neural networks can be used to develop predictive models that analyses patient data, including Cholesterol, BMI, to diagnose diabetes or predict the risk of developing diabetes in the future. Neural networks can be trained to predict the risk of diabetes. The existing dataset is pre-processed. Using feature extractions, highly correlated features are extracted. The performance of the model was evaluated through accuracy. A user interface was also developed for clinical use.

[1] INTRODUCTION

Neural Network is a subset of computer science that focuses on improving computers innovation. Automatic learning has several applications in our daily lives, particularly in the domain of medical care. Automatic learning is important in the medical care area due to its powerful data analysis capabilities. Although there is a great advance in the medical treatment options of cardiovascular events (CV) that include antiplatelet and statin agents, ADVERSE CV events remain a significant threat to patients with diabetes.1 cardiovascular disease (ECV) is associated with high mortality among patients with type 2 diabetes (DMADM) where it can explain more of the deaths due to T2DM Typetal Typetal Diabetes. (T1DM). The developed risk models are applied to the general population and, separately, for people with type 3.4 diabetes, but these models are often not well widespread when applied to other populations.5 In fact, ECV risk prediction scores based on traditional risk factors could not identify people who experienced a CV event in 10 years of follow -up among patients with T2DM Evaluated had a comparable and modest discriminative capacity. The severity of diabetes in humans was determined by using various data mining methodologies and neural networks. The main component analysis (PCA) is used to classify the severity of the disease. Because the characteristics of their problems are complex, diseases should be treated with caution.

1.1 General Introduction

Type 2 diabetes mellitus is a generalized chronic condition that continues to raise serious health challenges worldwide. One of its most dangerous and common consequences is cardiovascular disease (ECV), which remains a main cause of death among diabetic individuals. The coexistence of diabetes and cardiovascular complications exerts tremendous pressure on health systems, which underlines the need for more intelligent and most proactive approaches for early diagnosis. Traditional clinical tools often lack the necessary precision and speed to detect cardiovascular risks in diabetic patients at an early stage. These methods can ignore critical indicators, which leads to late interventions when complications are already advanced. However, with the growing availability of digital health records and the rapid evolution of artificial intelligence (AI) and automatic learning (ML), there is now a strong opportunity to build intelligent systems capable of providing more precise and timely predictions. The system gathers advanced data preprocessing, the analysis of main components (PCA) for the selection of characteristics and predictive modeling using neuronal networks of random forests and backs (BPNN). To further improve precision, particle swarm optimization (PSO) is applied to adjust the neuronal network parameters. In addition to building a reliable backend model, the project includes the design of an interactive user interface that allows health professionals to enter patient data and receive instant risk predictions. The results demonstrate how automatic learning can significantly improve medical diagnosis, support more informed decisions and pave the path for personalized attention in real world clinical environments.

1.2 Problem Statement

Diabetes is a widespread and chronic metabolic disorder affecting millions of people worldwide. One of the most concerning aspects of diabetes is its association with an increased risk of cardiovascular diseases (CVD), such as coronary artery disease, stroke, heart failure, and peripheral artery disease. Cardiovascular complications are a leading cause of morbidity and mortality among individuals with diabetes. Therefore, timely identification of patients at risk of these complications is essential for effective prevention and treatment.

This problem statement addresses the need for a comprehensive diabetes prediction model that not only evaluates the risk of developing diabetes but also classifies the risk of specific cardiovascular complications. The goal is to utilize advanced hybrid neural network algorithms to integrate data from various sources and provide a holistic approach to risk assessment and patient care.

Achieving this requires addressing challenges such as data integration, model generalization, handling imbalanced data, and ensuring interpretability. The goal is to develop an accurate and clinically useful tool that can aid in the early detection and personalized management of cardiovascular complications in individuals with diabetes, improving healthcare outcomes and reducing the burden of diabetes-related cardiovascular diseases.

1.3 Algorithm

1. Random Forest Algorithm
2. Particle Swam Optimization
3. Back Propagation Neural Network
4. Random Forest Regressor Algorithm

[2] LITERATURE SURVEY

El_jerjawi and Abu-Naser (2018) explored the application of artificial neural networks (ANN) for diabetes prediction, emphasizing the importance of precise diagnosis given the global prevalence and the high cost of diabetes care. His study focused on minimizing the error function during the ANN training, achieving an average error of 0.01 and a precision rate of 87.3% in the diagnosis of diabetes. This research reinforces the effectiveness of ANN - based models to facilitate early intervention and promote profitable health solutions. It contributes significantly to the growing body of work aimed at improving predictive precision in the management of diabetes.

Mohhamed Khalid Hossen (2022) proposed an automatic learning approach for the prediction of heart disease using patient characteristics such as blood pressure, cholesterol levels, sex and age. The study evaluated several algorithms, including logistics regression, decision trees and neural networks, the incorporation of data preprocessing, characteristics engineering and models selection techniques. The model performance was evaluated by evaluation metrics such as ROC-AUC, precision, and withdrawal. The findings demonstrated the potential of automatic learning to help early detection and risk assessment of heart disease, supporting proactive interventions and personalized medical care. Future recommendations included expanding the range of clinical parameters and improving the interpretability of the model to maximize clinical utility and global impact on cardiovascular health.

Huang et al. (2022) developed a risk prediction model based on machine learning to evaluate coronary artery calcification (CAC) scores for estimating the likelihood of coronary heart disease (CHD). Their approach integrated diverse patient data, including demographic, clinical, and imaging factors, to enhance predictive accuracy. By applying advanced machine learning techniques, the model demonstrated improved early detection and risk assessment capabilities for CHD. This method highlights the potential for machine learning to facilitate targeted preventive interventions and enable personalized risk assessment in clinical practice.

Kwiendacz et al. (2023) investigated the relationship between profiles derived from automatic learning and cardiovascular risk in patients with diabetes mellitus through the Silesia Diabetes-Heart project. Its study aimed to establish comprehensive cardiovascular risk profiles for diabetic patients using advanced automatic learning techniques, including characteristics and predictive modeling. By integrating various patient data, such as biomarkers, clinical history and demographic characteristics, research provided individualized risk assessments. The findings identified new risk factors and underlying mechanisms that contribute to cardiovascular disease in diabetic populations, highlighting the potential for specific interventions and precision medicine approaches to reduce cardiovascular risk.

Yufis Azhara et al. (2023) proposed an approach based on automatic learning to predict heart disease that uses the characteristics of the patient. Several algorithms were evaluated, such as logistics regression, decision trees and neuronal networks, after data preprocessing, characteristics engineering and the selection of models. The model performance was validated using evaluation metrics such as ROC-AUC, precision, precision and withdrawal. The study demonstrated the effectiveness of automatic learning to enable early detection and risk assessment of heart disease, thus supporting proactive therapeutic strategies and personalized medical care. Future addresses highlighted the importance of incorporating additional clinical parameters and improving the interpretability of the model to further improve the clinical impact and global cardiovascular health results.

3. EXISTING METHODOLOGY

The existing system employs machine learning to predict cardiovascular events in patients with diabetes. It integrates patient data, including medical records, clinical measurements, and lifestyle information, to develop predictive models. These models utilize advanced algorithms to identify patterns and risk factors associated with cardiovascular events, such as heart attacks and strokes, in diabetic patients.

The system's algorithms continuously analyze, and update predictions based on the evolving health status of patients. It provides healthcare professionals with valuable insights to better assess and manage the cardiovascular risk in diabetic patients, enabling early intervention and personalized care plans. By leveraging large datasets and artificial intelligence, the Project aims to enhance the overall health outcomes of individuals with diabetes by offering tailored risk assessments and interventions, ultimately reducing the burden of cardiovascular events in this high- risk population.

4. PROPOSED METHODOLOGY

The proposed methodology focuses on developing a predictive system capable of identifying cardiovascular complications in diabetic patients who use automatic learning and deep learning techniques. Initially, diabetes -related data sets and heart disease were collected and preprocessed managing missing values, eliminating atypical values and scale characteristics for consistency. The main component analysis (PCA) was used to reduce dimensionality and extract the most significant attributes of data sets. The refined characteristics were used to train a random forest classifier to establish a strong baseline for prediction. In addition, a neuronal network of posterior propagation (BPNN) was implemented and optimized using particle swarm optimization (PSO) to improve the precision of the prediction and convergence of the model. A comparative analysis was performed between the random forest and the BPNN models, with BPNN combined with PSO that demonstrates a higher performance. In addition, an easy -to -use interface was developed to allow health professionals to enter patient data and obtain real -time risk assessments, thus facilitating early diagnostic strategies and personalized intervention.

4.1 BLOCK DIAGRAM

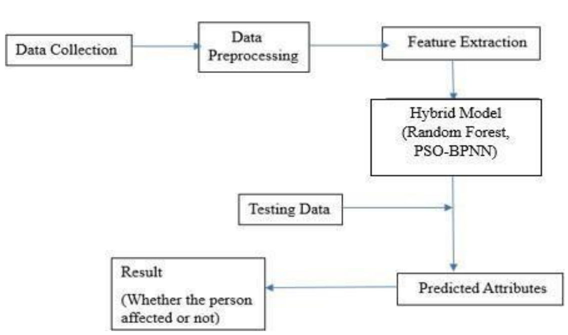


Figure4.1: Architecture of Proposed System

5. RESULTS

5. 1 Data Preprocessing

Initial data sets for diabetes and heart disease contained several missing values in important clinical attributes. To address this problem, the median imputation was carried out based on class labels, ensuring that the filled values reflect the underlying distribution of the data. After preprocessing, the data sets were completely free of missing inputs, providing a robust base for the extraction of characteristics and the training of models.

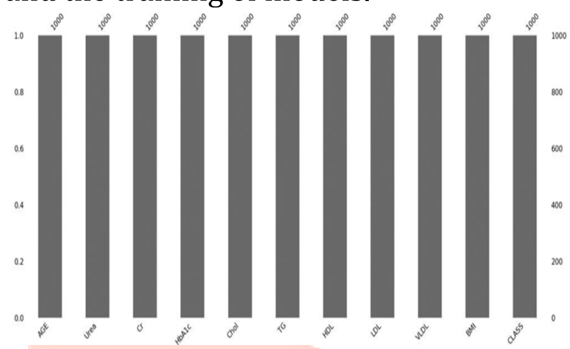
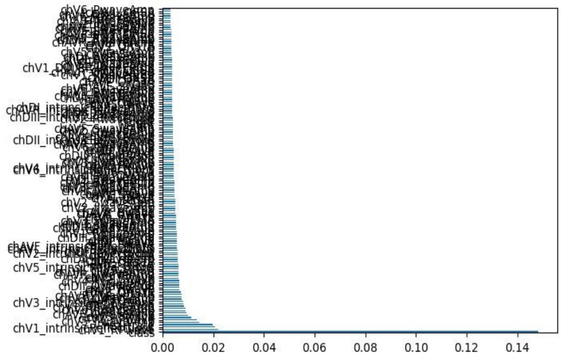


Figure5.1: Correlation matrix of Diabetic Patients

Figure 5.2: Correlation matrix of Heart Disease

5.2 Feature Extraction

The main component analysis (PCA) was used to extract the most significant characteristics of both data sets. This dimensionality reduction technique helped focus on the attributes that contribute to the greatest reducing computational complexity. For the diabetes

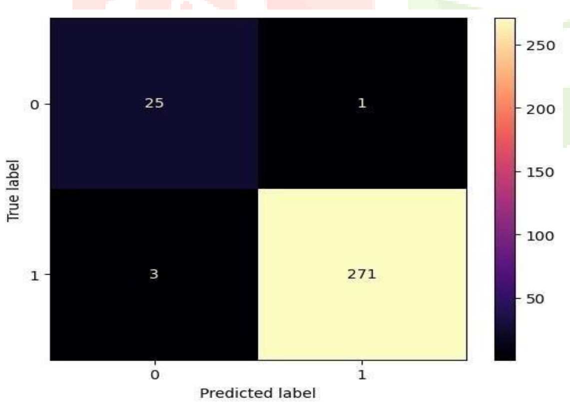


data set, age, HBA1C, BMI and TG identified as the key characteristics. In the data set for heart disease, several ECG parameters were selected according to their high contribution to data variance. variance, improving the performance of the model

Figure5.3: Feature Extraction from Dataset

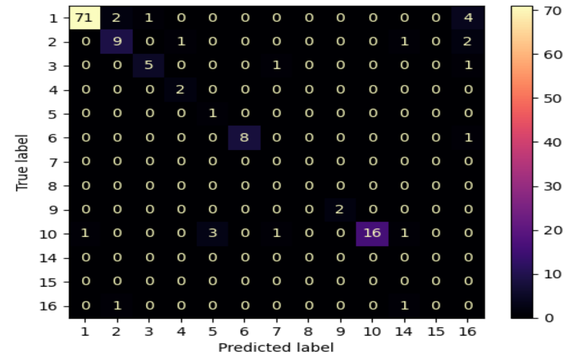
5.3 Classification Model Results

The random forest classifiers were trained in the characteristics extracted to predict the appearance of diseases. The models showed strong predictive capabilities, with the diabetes data set achieving a high classification precision and the data set of heart disease data achieving a general precision of 83%. The



random forest model demonstrated its ability to handle the complex patterns present in the clinical data sets effectively.

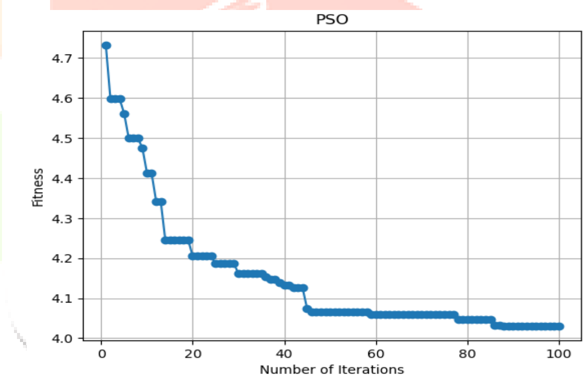
Figure 5.4: Confusion matrix of Diabetic



Patients

Figure 5.5: Confusion matrix of Heart Disease Dataset

To further improve the prediction performance, a neuronal network of posterior propagation (BPNN) was optimized using the optimization of particle swarm (PSO). The optimization process significantly improved the convergence of the model and reduced the possibilities of being caught in local minimums.



The BPNN model, when optimized with PSO, achieved impressive accuracy of 96.5% in the diabetes data set, overcoming traditional BPNN models without optimization.

Figure 5.6: Graphical Representation of PSO

Figure 5.7: Graphical Representation of PSO

5.4 Evaluation Metrics

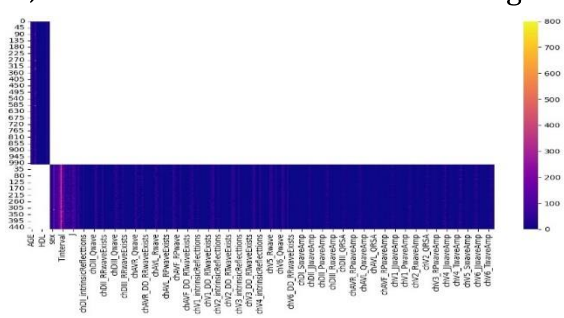
The effectiveness of developed models was evaluated by standard evaluation metrics. It was found that the correlation coefficient between the real and predicted results was 96.60%, indicating a very strong positive relationship. In addition, the average absolute error (MAE) was calculated as 0.0098, confirming the precision and reliability of the predictions made by the models.

Formula:

MAE=n1i=1∑n|yi−y^i|

5.5 Key Influential Features

The analysis of the importance of the characteristic reveals the most critical attributes that influence the prediction results. For diabetes prediction, age levels, HDL, BMI and HBA1C cholesterol emerged as dominant characteristics. For the prediction of heart disease, the characteristics of the ECG signal



such as a Tintval, QRSA and intrinsic. The identification of these characteristics provided valuable clinical information to improve evaluation of the risk of disease.

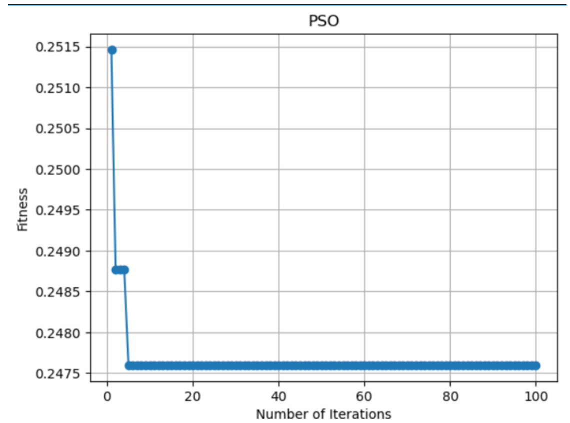


Figure 5.7: Correalated Attributes

5.5 Prediction Results

To translate the models developed into a practical clinical tool, an easy -to -use interface was developed to allow cardiovascular risk prediction in diabetic patients based on clinical and diagnostic parameters. The system allows health professionals to enter important data of patients. After processing inputs through trained predictive models, the interface generates an immediate risk prediction output. Classify the patient into categories of "high risk" or "low risk" to develop cardiovascular complications. This binary classification supports doctors in the identification of patients who require early intervention and specialized monitoring.

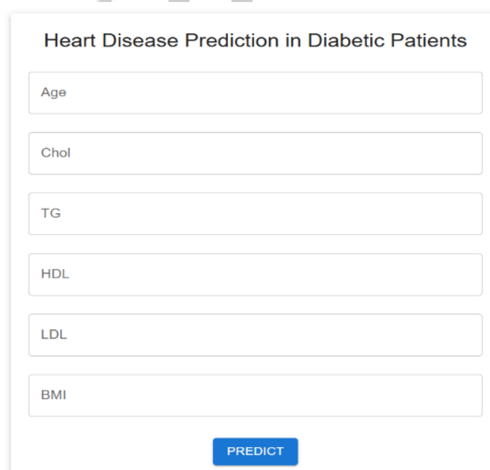


Figure 5.8: UI for Prediction of Heart Disease in Diabetic Patients

6. CONCLUSION

In this research work, a predictive system was developed for the early identification of cardiovascular complications between diabetic patients through the integration of diabetes data and data sets of heart disease. The extraction of characteristics through the correlation analysis and the analysis of main components (PCA) allowed the identification of the most influential attributes that contribute to cardiovascular risks. Automatic learning models were implemented, including the random forest classifier and the back propagation neuronal network (BPNN), to predict cardiovascular complications effectively. The BPNN model, even more optimized using the optimization of the swarm of particles (PSO), demonstrated a significant improvement in predictive precision, which reaches 96.5%, compared to the accuracy of the basal random forest of 83%. The model also achieved a strong correlation coefficient of 96.6% and an absolute medium (MAE) error of 0.0098, confirming the reliability and robustness of the system. The results emphasize the importance of the extraction of combined characteristics of diabetic and heart disease and attribute it to improving prediction capacities. The proposed system not only allows precise risk assessment, but also supports early clinical decision making, thus improving the patient's results. Future work can focus on integrating hospital data in real time, expanding data sets for more diverse patient populations and implementing the model through mobile or cloud -based health applications for broader accessibility.

7. REFERENCES

[1] esreen Samer El_Jerjawi, and Samy S. Abu-Naser, “Diabetes Prediction Using Artificial Neural Network” (2018).

[2] Daniel Sierra-Sosa, Begonya Garcia-Zapirain, Member, IEEE, Cristian Castillo, Ibon Oleagordia, Roberto Nuño-Solinis, Maider Urtaran-Laresgoiti, Adel Elmaghraby, Life Senior Member, IEEE “Scalable Healthcare Assessment for Diabetic Patients using Deep Learning on Multiple GPUs”IEEE Transactions on Industrial Informatics, ID:1551-3203(2019).

[3] Bala Manoj Kumar P, Srinivasa Perumal R, Nadesh R K, Arivuselvan K ,"Type Diabetes mellitus prediction using Deep Neural Networks classifier",International Journal of Computing in Engineering 1 (2020)

[4] MD. Shafiqul Islam, Marwa K. Garaqe (Member IEEE), Samir Brahim Belhaouari (Senior Member,IEEE), and Muhammad A. Abdul-Ghani "Advanced Techniques for Predicting the Future Progression of Type 2 Diabetes", IEEEACCESS.2020.3005540, VOLUME 8, 2020.

[5] Jobeda Jamal Khanam, Simon Y. Foo, “A comparison of machine learning algorithms for diabetes prediction”,ICT Express Volume 7, Issue 4,432-439 (2021).

[6] Boon Feng Wee, Saaveethya Sivakumar1, King Hann Lim1, W. K. Wong, Filbert H. Juwono, “Diabetes detection based on machine learning and deep learning approaches”, s11042-023-16407-5 (2022).