

# PREDICTION OF CARDIAC DISABILITIES IN DIABETIC PATIENTS

## Abstract

Cardiac arrest is a condition which is caused by an abnormal heart rhythm and the heart unexpectedly stops pumping blood to various organs, without prompt intervention, it can result in the person's death. Diabetes is associated with an increased risk of cardiac arrest.

Machine learning can be implemented in the medical domain to make better decisions and predictions. In the absence of medical professionals, the prediction model plays a very crucial role in the early detection and diagnosis of diseases.[3]. Various available techniques such as Support Vector Machine (SVM), Linear Regression, Logistic Regression, Decision tree, Random Forest, so on, can be used in the prediction process.

The principal objective of this paper is to predict cardiac arrest in diabetic patients using machine learning techniques. This paper gives the comparison between the different machine learning techniques used and delineates on the most reliable technique and also predicts whether or not a diabetic patient will experience a heart attack.

**Keywords:** Cardiac arrest, Diabetic patients, Machine Learning Techniques, Prediction Model.

## Authors

### Sheela. S

Assistant Professor  
Department of CSE  
Global Academy of Technology  
Visveswaraya Technological University  
Bangalore, Karnataka, India.

### Dr. Shalini. S

Associate Professor  
DSATM  
Bangalore, Karnataka, India.

## I. INTRODUCTION

The heart's electrical failure is the main cause for cardiac arrest. It is the condition where the heart stops beating properly and is caused by irregular heart rhythms called arrhythmias. Diabetes is a condition that is caused due to high levels of blood glucose or blood sugar. There are different health conditions in diabetic patients that could lead to cardiac arrest. They are: high blood pressure, abnormal cholesterol, and high triglycerides, poorly controlled blood sugar levels (too high/out of normal range). High glucose levels in the blood can damage blood vessels and nerves which increases the chances of having a cardiac arrest.

Machine learning when employed in healthcare, aids in processing and analyzing huge imbalanced medical datasets. These systems can be used by medical professionals in diagnosing patients. As diabetes and cardiac arrest are closely related, it is extremely important to predict the occurrence of.

There are many prediction systems developed to predict the presence of any heart disease. There have been several researches to analyze different techniques to be implemented in the prediction models and to increase the accuracy of prediction. Researches are being conducted to find out cardiac arrest in diabetic patients priorly.

Whether diabetes is closely associated with sudden cardiac arrest among different age groups. It is very important for normal people with diabetes to simply feed in their data and know the chances of them having to suffer a cardiac arrest or any other heart-related issue. This helps in spreading the awareness to people with diabetes to take proper precautions and any prior medical help if necessary.

### **The major contributions of our work are:**

- We present a prediction model that is used to predict cardiac arrest in diabetic patients.
- We pre-process the diabetic dataset completely and understand the pattern of the data. The database is cleansed and visualised to find the best features to detect cardiac arrest in diabetic patients.
- To further analyse the data, some of the major features from the selected features are chosen as the dependent variables and categorical and continuous plots are plotted.
- The model is trained and recursive function elimination is performed.
- The patient dataset has been divided into training data (75%) and testing data (25%). These data are fed to four different machine learning models and the accuracy of each obtained. The highest accuracy yielding model is regarded as the best model.
- This model is chosen and used to predict the occurrence of cardiac arrest in an individual diabetic patient using various parameters like age, cholesterol, BP levels, etc.

## II. RELATED WORKS

In the survey [1], several supervised machine learning techniques used in the heart disease prediction models are analyzed based on the accuracy of the results they provide.

In a previously published study[2], the database considered is analyzed and segregated. This segregated database is used in a model to predict heart disease. Machine learning techniques such as Logistic regression, Naïve Bayes are used to increase the accuracy rate. In this proposed system, logistic regression (classification) algorithm and Sklearn library is used to calculate the score. Finally, results are analyzed by comparing Models and using the Confusion Matrix.

In [3], several machine learning techniques like K-means, Naïve Bayes classification, Decision tree algorithm, and ID3 algorithms were used in the prediction of cardiac arrest and to predict the chance of occurrence of cardiac arrest. The preprocessing of the data includes cleaning, transformation, integration, and reduction.

Prediction of sudden cardiac arrest using a machine learning technique two minutes before occurrence using optimal heart rate variability features is the aim of the paper [4]. In this study, the HRV signal is analyzed based on three different feature extraction methods namely, time, frequency, and nonlinear domains.

The logistic regression with backward elimination is used to derive a model for risk score. The machine learning techniques Random forest classifiers, logistic regression classifiers are used for pre-processing the ECG dataset, diabetes dataset, gene expression dataset, blood pressure dataset, and the risk score is calculated.[6]

The research article “Diabetes, glucose tolerance, and the risk of sudden cardiac death” aims to assess the risk of sudden cardiac death and other heart-related diseases associated with diabetes and IGT(impaired glucose tolerance). It also emphasizes on whether the risk of SCD would be pronounced over the risk of non-sudden cardiac death among diabetes patients and the timing of the SCD risk associated with diabetes.[7]

In the paper [8], the proposed system explains how the hospital data is randomly divided into training set and test set. The application of several machine learning algorithms on the training set and also explains the processes of testing the model using the test data.

### **III. PROPOSED SYSTEM**

We present a prediction model that predicts whether or not a diabetic patient suffers a cardiac arrest using the available dataset. Based on the score, the participants can take the necessary precautions and choose to conduct a more thorough check-up with a doctor.

### **IV. METHODOLOGY**

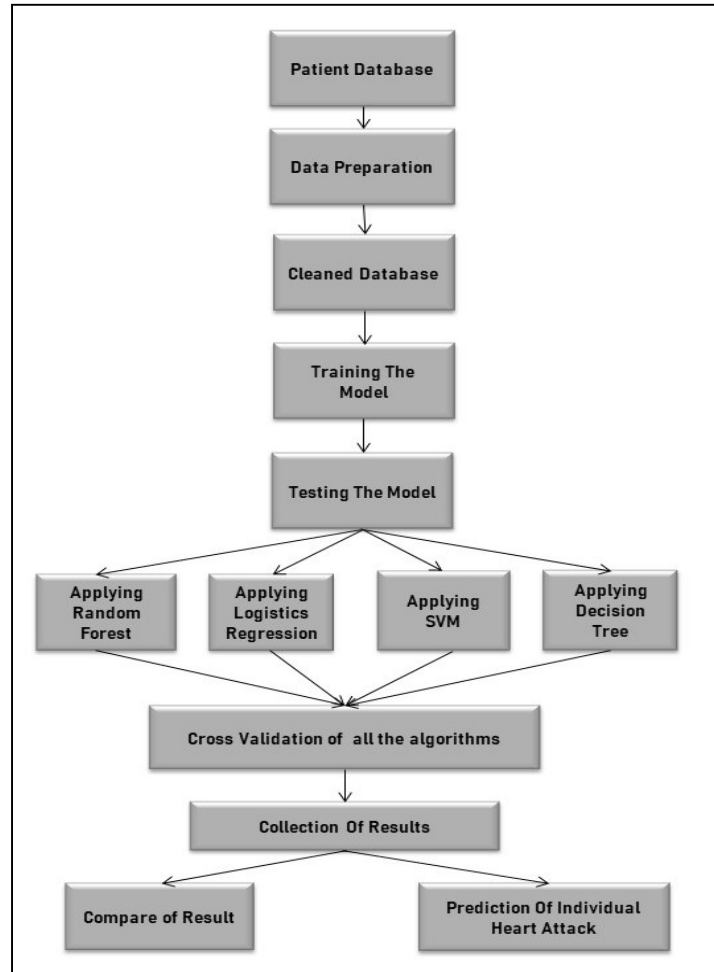
With the continuous growth of technology in the medical field, we have seen an exponential increase in the electronic health records (EHR) which generates huge amount of health records. [4 ]

The majority of the available medical data are extremely disproportionate and to derive a stabilised training dataset, there are many methods used relevant to the parameters chosen in that particular research/ experiment. [5]

Analysis of imbalanced datasets reduces the accuracy significantly. Hence the need for balancing of the dataset is high. [4]

Generally, the imbalanced learning problem concern is dealt with by choosing the best suitable standard algorithm which involves choosing a base class classifier and several other parameters, by keeping these as the reference, the imbalanced dataset is analyzed and grouped.

## V. SYSTEM ARCHITECTURE



**Figure 1:** System Architecture

## VI. TRAINING SET

All the models are applied to get the results. The evaluation metric used is the confusion matrix.

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

**Figure 2: Confusion Matrix**

The confusion matrix displays the correctly predicted as well as incorrectly predicted values by a classifier.

The sum of TP and TN, from the confusion matrix, is the number of correctly classified entries by the classifier.

Confusion Matrix for SVM			
		0	1
0	0	124	13
	1	5	100
Training Set			
		0	1
0	0	32	9
	1	3	17
Test Set			

**Figure 3: Confusion Matrix for SVM**

**Logistic Regression**

Confusion Matrix for Logistic Regression			
		0	1
0	0	118	22
	1	11	91
Training Set			
		0	1
0	0	32	9
	1	3	17
Test Set			

**Figure 4: Confusion Matrix for Logistic Regression**

**Decision Tree**

Confusion Matrix for Decision Tree			
		0	1
0	0	129	0
	1	0	113
Training Set			
		0	1
0	0	29	8
	1	6	18
Test Set			

**Figure 5: Confusion Matrix for Decision Tree**

## Random Forest

Confusion Matrix for Random Forest							
		0	1				
0	0	129	2	0	0	32	10
	1	0	111		1	1	3
Training Set				Test Set			

**Figure 6:** Confusion Matrix for Random Forest

## VII. RESULTS

In this work, we have used four machine learning techniques namely Random Forest, Logistic regression Decision Tree and Support Vector Machine (SVM). The accuracies obtained from these techniques is shown in Fig .2

	Model	Training Accuracy	Test Accuracy
0	Logistics Regression	0.843478	0.844156
1	SVM	0.830435	0.805195
3	Random Forest	1.000000	0.805195
2	Decision Tree	0.843478	0.688312

**Figure 7:** Results obtained

As seen from the Fig 2 Logistic Regression yields the highest accuracy and it is chosen to predict cardiac arrest in an individual.

## VIII. CONCLUSION

Unlike the previous prediction systems, this system predicts whether or not a diabetic patient will suffer a heart attack based on the parameters provided by them. This system helps in the prediction of cardiac arrest at an early stage, which can help patients take the necessary precautions and medical help. This model uses four different machine learning techniques for prediction rather than only one, for better prediction results. The proposed system is designed to compare the different machine learning techniques used based on their accuracy and suggest the best technique.

The proposed model has immense scope for improvement and can be implemented to develop a model that can predict the percentage chances of having a cardiac arrest in diabetic patients. This can help patients and medical professionals understand the severity of the condition.

## REFERENCES

- [1] V.V. Ramalingam, AyantanDandapath, and M Karthik Raja, “Heart disease prediction using machine learning techniques: a survey”, International Journal of Engineering &Technology, 2018
- [2] Reddy Prasad, Pidaparathi Anjali, S.Adil, and N.Deepa, “Heart Disease Prediction using Logistic Regression Algorithm using Machine Learning”, International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8, Issue-3S, February 2019
- [3] Rajesh N, T Maneesha, Shaik Hafeez, and HariKrishna, “Prediction of Heart Disease Using Machine Learning Algorithms”, International Journal of Engineering &Technology, 2018
- [4] L. Murukesan, M. Murugappan, M. Iqbal, and Dr. K. Saravanan “Machine Learning Approach for Sudden Cardiac Arrest Prediction Based on Optimal Heart Rate Variability Features” Journal of Medical Imaging and Health Informatics, Vol. 4, 1–12, 2014
- [5] David S. Siscovick, Nona Sotoodehnia, Thomas D. Rea, Trivellore E. Raghunathan, Xavier Jouven, Rozenn N. Lemaitre “Type 2 diabetes mellitus and the risk of sudden cardiac arrest in the community” Springer Science+Business Media, NIH Public Access, Rev EndocrMetabDisord, 11(1): 53–59. doi:10.1007/s11154-010-9133-5, March 2010
- [6] Nachiket Tapas, Tushar Lone, Damodar Reddy and VenkatanareshKuppili “Prediction of cardiac arrest recurrence using ensemble classifiers” Indian Academy of Sciences, Sadhana Vol. 42, No. 7, pp. 1135–1141 \_DOI 10.1007/s12046-017-0683-z, July 2017
- [7] Antti Eranti1, Tuomas Kerola, Aapo L. Aro, Jani T. Tikkanen, Harri A. Rissanen, Olli Anttonen, M. Juhani Junttila, Paul Knekt and Heikki V. Huikuri "Diabetes, glucose tolerance, and the risk of sudden cardiac death," Bio Med Central, Eranti et al. BMC Cardiovascular Disorders 16:51 DOI 10.1186/s12872-016-0231-5, 2016
- [8] Min Chen, Yixue Hao, Kai Hwang, Lu Wang, and Lin Wang “Disease Prediction by Machine Learning over Big data from Healthcare Communities,” IEEE Transaction, 2018