

Forecasting and Categorization of Cardiac Arrhythmia using Soft Computational Approaches

Dr. Sivaprakash. C

*Department of Artificial intelligence and Machine Learning
Sri Sairam College of Engineering
Anekal, Bangalore
Email:shivadesk@gmail.com*

P.Ramkumar

*Department of Artificial intelligence and Machine Learning
Sri Sairam College of Engineering
Anekal, Bangalore
Email:ramkumarkohila@gamil.com*

Abstract

Health is extremely crucial in the existence of humans, and each part of the human body plays a substantial function in maintaining a well-being lifestyle. Arrhythmia illness is a severe issue in the heart. This disease must be recognized early and preventative measures implemented. Despite remarkable advancements in the medical field, heart arrhythmia remains a difficulty for medical practitioners. As a result, it must be carefully classified in order to diagnose the signs of this condition, which may assist the medical practitioner in treating the patient in an avoidable manner. Despite numerous studies, there are not enough methods available to classify arrhythmias accurately. As a result, this study developed a data mining approach known as K-Nearest Neighbour (KNN) to classify cardiac arrhythmia disorder. This proposed approach yields a beneficial result of 98.1% accuracy and 93.2% precision.

Keywords - *Arrhythmia Disease, Machine Learning, K-Nearest Neighbour, Cardiovascular health, Random Forest*

1. Introduction

Individuals' cardiovascular health is jeopardised when they have arrhythmia, a medical illness characterised by irregular heartbeats. This study digs thoroughly into the complexity of arrhythmia, investigating its causes, kinds, symptoms, diagnosis, and the multiple treatment options currently available. The persons having more awareness about arrhythmia disease are better able to recognise the symptoms, seek immediate medical attention, and take preventative measures to keep their hearts healthy.

An irregular heartbeat is a symptom of arrhythmia, and the presence of aberrant stimulating indications in the heart that disrupt the organ's regular rhythm is categorized as a dysfunction [8]. When this occurs, the electrical system present in the heart, which is in charge of synchronising the muscle contractions of its chambers, suffers, resulting in irregular or abnormally rapid cardiac pulses. Arrhythmias are irregular heartbeats that can affect either the atria (the heart's upper chambers) or the ventricles (the heart's lower chambers), decreasing the heart's ability to adequately pump blood [9][10]. Atrial fibrillation, ventricular tachycardia, and bradycardia are examples of common arrhythmias. Bradycardia is a cardiac condition that is characterized by abnormally slow heart rate, often below 60 beats per minute on average. Bradycardia is cause for concern when it causes symptoms or impairs the heart's ability to pump blood effectively. While certain individuals, like sportsmen, may possess an inherent sluggish heart rhythm that poses no issues, bradycardia can be worrisome when it leads to symptoms. Despite the expansion of various sectors worldwide, this particular ailment continues to pose challenges for early diagnosis by all experts. Consequently, a soft computing method is suggested to identify this condition beforehand.

Tachycardia is one form of arrhythmia distinguished by an abnormally fast heart rate that frequently exceeds 100 beats per minute. Tachycardia that is persistent or repeated can be concerning because it can create symptoms as well as complications. Although transient increases in heart rate in reaction to physical exertion or mental stress are regarded acceptable, chronic or recurring tachycardia should be taken seriously. As a result, this form of sickness must be classified

using cutting-edge technologies. This arrhythmia disease will be classified using a machine learning system. This technology will assist professionals in readily classifying diseases from the healthiest humans.

Artificial intelligence (AI) is employed in the majority of investigative examinations, where it assumes a pivotal position in information extraction. Researchers and academics are prone to mistakes during their research or when trying to make connections between different terrains. This complicates their ability to handle particular challenges [11]. In such instances, machine learning algorithms are trained using labelled datasets to forecast or categorize new occurrences. This method has been demonstrated to be efficient in different fields and has resulted in notable enhancements in system effectiveness and machine development. The characteristics employed by machine learning algorithms to depict data instances can be uninterrupted, grouped, or two-valued. In contrast to unsupervised learning, where instances are not labelled, supervised learning tasks are frequently necessary in numerous machine learning applications. The current literature has focused on the methods needed to achieve this. This work focuses on classification problems where the output of instances accepts only discontinuous and unordered values. It has also incorporated several mentions to groundbreaking investigation that sparked the particular field of study being examined. Machine learning can be observed in data pre-processing and characteristic assortment. There are logic and symbolic methodologies furnished

2. Literature Review

G.K. Prasad et al [1] outlined the use of artificial neural networks (ANN) and wavelets to correctly classify ECG arrhythmias. The suggested approach is resistant to interference and can differentiate between regular sinus rhythm and 12 different forms of arrhythmias. Despite using all 13 arrhythmia types from the MIT-BIH dataset, the average accuracy of identification (96.79%) was similar to that of alternative methods. Alongside the RR interval data, this set of characteristics is a subset of the DWT coefficients produced using the 'Sym6' wavelet. The authors employ a progressive neural network with two hidden layers and employ sigmoid activation parameters for all neurons. The network is trained using the backpropagation algorithm at moderate learning rate and momentum levels.

An efficient and reliable technique for classifying cardiac arrhythmias was presented by Sanjay Tanaji Sanamdikar et al. [2]. This method involved extracting properties from ECG signals using a generic sparsed neural network (GSNN). Different kinds of ECG beats are chosen using the MIT-BIH dataset, and the ECG signal is filtered by calculating the signal-to-noise ratio. Attaining the highest rate of effectiveness, the precision level of arrhythmia diagnosis is 98%. The presented method of identifying and predicting arrhythmias using GSNN would be more successful in identifying arrhythmias than previous methods. Forecast and classification efficacy will be improved by the proposed system.

Kavyshree B[3] and colleagues discussed the utilization of machine learning to forecast cardiac arrhythmia. Arrhythmia refers to a state where the heart beats in an irregular or abnormally fast or slow manner. The study examines the crucial elements that are examined during the usual pattern phase of a heartbeat, known as the ECG signal durations and their associations with each other. Based on these factors, the research recommends utilizing machine learning algorithms to forecast irregular heart rhythms.

Komal S.[4] and colleagues showcased their utilization of machine learning algorithms and methodologies to classify heartbeats using electrocardiogram (ECG) data.. The authors go over numerous approaches for preprocessing ECG data, extracting features, and employing various types of classifiers for automatic heartbeat categorization. The intention of this analysis is to enhance the correctness of detecting heart irregularities, which is critical for diagnosing and treating cardiovascular disorders. Also, this study and analysis mentions a hybrid model that blends genetic algorithms and decision trees to improve arrhythmia classification accuracy. The authors evaluated the functionality of the suggested model using the UCI arrhythmia dataset and reported sensitivity, specificity, and average sen-spec measures.

Karthik.D [5] et al suggested an approach based on machine learning that uses electrocardiography (ECG) data to predict cardiac arrhythmia. Because manual interpretation by doctors is a time-consuming and tiring process, the authors emphasize the importance of automated methods for detecting abnormal heart conditions from ECG data. The suggested approach utilized the Random Forest (RF) algorithm and achieved an accuracy rate of 98%. The Hermite function was also presented as a method for extracting characteristics, and higher-order statistics (HOS) were utilized to improve the characteristics in the support vector machine (SVM) classifier approach. Moreover, a unique classifier technique known as the optimal path forest classifier (OPF) was introduced.

Saira Aziz and her group demonstrated how electrocardiogram (ECG) signals can be employed to identify heart ailments and irregularities. In numerous cases, cardiovascular conditions can be detected by examining the fluctuations in ECG readings, which include P, QRS complex, and T waves. However, variations and abnormalities in ECG interpretations can cause spikes, which must be eradicated for accurate analysis. In order to analyse ECG data, the paper proposes a new technique that makes use of the fractional Fourier transform (FrFT) and two-event related moving averages

(TERMA). Effectiveness-wise, the suggested approach outperforms the state-of-the-art methods. The application of machine learning algorithms for automatically categorising heart diseases based on estimated peak intervals, other ECG signal characteristics, and intervals between various peaks is also being investigated in this research. More over 10,000 patients from the Shaoxing People's Hospital (SPH) database were used to train the machine learning model that was suggested; this made it more useful for classification.

Abhinav Vishwa et al. emphasized the need of using electrocardiogram (ECG) data to diagnose heart problems. The electrocardiogram signals are composed of P waves, QRS complexes, and T waves, denoted by the capital letters P, Q, R, S, and T. The shape, duration, and connection of the P wave, QRS complex, and T wave components, as well as the R-R interval, are the main features examined during the normal heartbeat phase of a cardiac cycle. Changes in these indicators suggest a heart condition that could be caused by anything. Arrhythmia refers to any irregular beat phase, and some arrhythmias are extremely harmful to the patient.

3. Design and Implementation

KNN Classification

Among the supervised learning methods, K-nearest neighbour is a useful classifier. A fresh data set is categorised using this method based on the maximum vote of its nearest neighbour. A distance metric that utilizes Euclidean distance, Manhattan distance, and the Minkowski distance method calculates the distance between closest neighbors. The quantity of neighbors is indicated by the K value.

If the K value is exceedingly low, the results will be less stable. However, increasing the K value allows to upsurge the inaccuracy while still obtaining steady results. Consequently, in the present study, the K value is determined through experimentation to prevent overfitting. Figure 2 illustrates a visual depiction of the KNN algorithm. K-Nearest Neighbor is a direct machine learning model in the realm of Supervised Learning Algorithm.[12][13][15]. The K-NN algorithm prioritises the new case to the group that most closely resembles the current groups based on its assumption of similarity between the new instance and the available data. By classifying new data points according to similarity, the K-NN approach keeps all of the current data. put out a machine learning-based technique that uses electrocardiography (ECG) data to predict cardiac arrhythmia. This means that with the K-NN algorithm, new data can be conveniently classified into an appropriate group [14]. The K-NN method is most commonly used for classification, while it can be applied to both regression and classification. K-NN makes no assumptions regarding the underlying data because of its non-parametric character. Due to its slow learning from the training set, it is sometimes referred to as a lazy learner algorithm. In its place, the dataset is saved and used for classification. During the training phase, the KNN algorithm merely keeps the dataset and adds new data to a group that shares a lot of similarities with the older data.

The separation between two endpoints is indicated by the equation (1) provided in this manner. x_1, x_2 and y_1, y_2 represent the x and y coordinates of the x and y points, respectively. The Random Forest approach surpasses the other methods, particularly the KNN technique, but the SVM method performs satisfactorily.

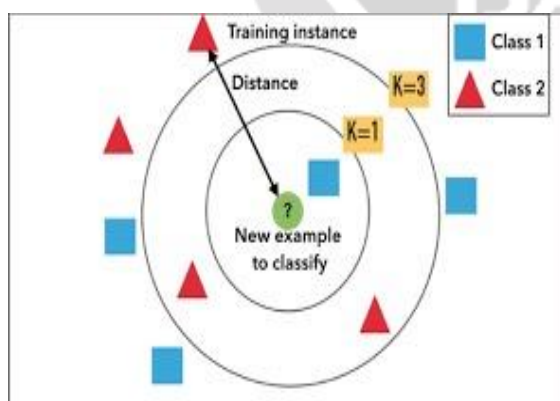


Figure 1. Process of KNN algorithm

The separation between two endpoints is indicated by the equation (1) provided in this manner. x_1, x_2 and y_1, y_2 represent the x and y coordinates of the x and y points, respectively. The Random Forest approach surpasses the other methods, particularly the KNN technique, but the SVM method performs satisfactorily. The evaluation takes into account the correctness, exactness, retrieval, and F1-measure. More summaries can be collected in the coming days utilizing an automatic tool.

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (1)$$

The provided equation computes the distance between two endpoints using recently acquired data. If the distance is insignificant compared to other distance measurements, the freshly obtained data should be included in the cluster.

SVM Classification

The latest supervised machine learning method is Support Vector Machines (SVMs). Hence, alongside giving a concise outline of SVMs, we will mention some newer investigations and notable advancements published prior to the works in this research. It has been demonstrated that by enlarging the gap and establishing the utmost feasible distance between the dividing hyperplane and the examples on both sides, the maximum threshold on the anticipated generalization error can be lessened [16]. When two categories can be linearly distinguished, an ideal separating hyperplane can be achieved by minimizing the squared size of the separating hyperplane. Consequently, the quantity of characteristics existing in the training set does not affect the complexity of the model in an SVM. Therefore, Support Vector Machines (SVMs) are appropriate for learning tasks with a significant number of characteristics in comparison to training instances [17][18]. Notwithstanding the fact that the most notable difference enables the SVM to choose from a wide array of possible hyperplanes, the SVM may encounter difficulties in finding a discerning hyperplane for numerous datasets due to incorrectly classified examples. Nevertheless, the majority of practical scenarios lack data that can be effectively separated by a hyperplane. One strategy to distinguish positive instances from negative ones in the training set is to convert the data into a space with a higher number of dimensions and establish a separating hyperplane there.

This higher-dimensional region is known as the converted feature space [19], which is different from the input domain of the training samples. Any consistent training set can be partitioned using an appropriately chosen converted feature space of appropriate dimensionality. A linear division in the transformed feature space corresponds to a non-linear division in the original input domain. In the feature space, classification points. As the fundamental function determines the modified characteristic area in which the examples of the training set will be classified, selecting the suitable fundamental function is crucial. It is customary to estimate various potential parameters before performing cross-validation on the training set to find the best one. As a result, the slow training speed is one of the disadvantages of SVMs. Choosing fundamental configurations is similar to deciding the number of hidden nodes to use in a neural network.

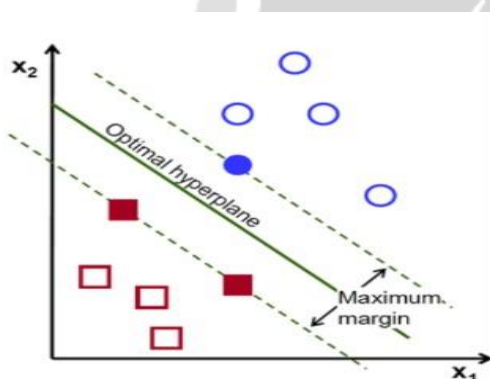


Figure 2. Process of SVM algorithm

Even if the designer is not aware of which training data attributes are being used in the kernel-induced transformed feature space, a Support Vector Machine (SVM) will function as long as the kernel function is legitimate. Without the need for additional matrix storage or the use of any numerical QP optimization steps, Sequential Minimal Optimisation (SMO) is a straightforward method for addressing the SVM QP problem quickly. SMO breaks the general QP problem down into smaller issues. Finally, unlike other search methods like neural networks, the SVM's training optimization problem must achieve a global minimum rather than a local minimum. Because SVM approaches are binary in nature, a multi-class problem must be transformed into a sequence of binary classification problems. Discrete information presents another obstacle, however, by appropriately adjusting the scale, one can achieve satisfactory outcomes.

Random Forest Classification

The random forest classifier is composed of several tree classifiers that produce a random vector that is chosen separately from the input vector. Every tree determines the unit value of the most common class in order to classify an input vector. The Gini Index is a metric for selecting attributes that measures the purity of an attribute based on the classes chosen by the algorithm.

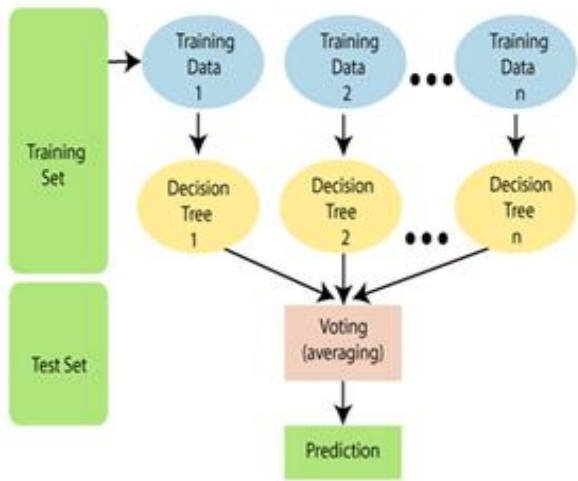


Figure 3. Process of random forest algorithm

4. Results and Discussions

The table beneath provides a summary of the different capabilities of the KNN algorithm using performance indicators. When KNN is utilized for categorization, it achieves a precision of 80.67%. Figure 4 displays the evaluation of the KNN technique in classifying arrhythmia disease. The measurements mentioned in table 1 have been considered for its analysis.

Table 1. Performance of KNN Algorithm

S.No.	Performance metrics	Performance (%)
1	Accuracy	80.67
2	Precision	80.3
3	F-Score	80.01
4	Recall	80.4

Figure 4 demonstrates that by utilizing the KNN technique, the precision has been established to be 80.6%. The accuracy was determined using confusion matrix measurements.

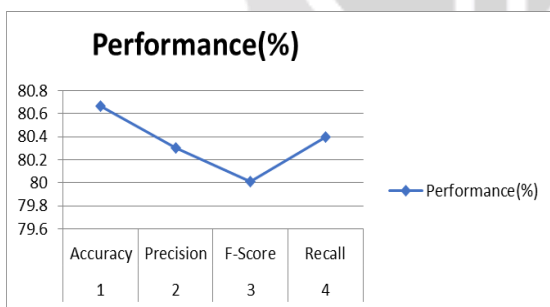


Figure 4. Performance analysis of KNN

Figure 4 demonstrates that by utilizing the KNN technique, the precision has been established to be 80.6%. The accuracy was determined using confusion matrix measurements.

Table 2. Confusion matrix of performance measure

N=Number of Predictions	Real: No	Real: Yes
Expected: No	TN	FP
Expected: Yes	FN	TP

The attached table 2 illustrates a confusion matrix of performance values. In this table, the performance variables TP, TN, FP, and FN refer to true positive, true negative, false positive, and false negative, respectively. The performance assessments and benchmarks for the Support Vector Machine algorithm are displayed below. In equation (2), 'x' signifies the magnitude of the vector, which will be indicated as $\|x\|$. The norm of the value with regard to its vector will be $(x_1, x_2, x_3, \dots, x_n)$ measured as

$$\|x\| = \sqrt{x_1^2 + x_2^2 + x_3^2 + \dots + x_n^2} \quad (2)$$

The vector direction will be measured as

$$\alpha = \frac{x_1}{\|x\|}, \frac{x_2}{\|x\|} \quad (3)$$

The SVM algorithm's efficiency, along with its efficiency metrics, are displayed in table 3 underneath.

Table 3. Performance of SVM Algorithm

S.No	Performance metrics	Performance (%)
1	Accuracy	85.37
2	Precision	84.2
3	F-Score	83.61
4	Recall	82.34

The accuracy of SVM is 85.37%, according to the table above. It stated that when compared to KNN, its performance is slightly superior. Since KNN reported an accuracy of 80.67%. However, SVM has provided more than KNN.

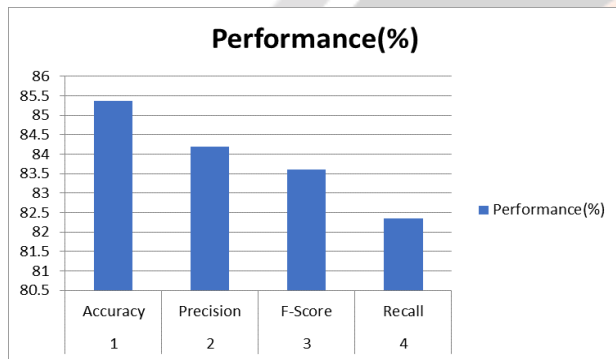


Figure 5. Performance Analysis of SVM algorithm

Figure 3 indicates that the SVM algorithm has an accuracy of 85.37. When compared to KNN, this algorithm produced more credible results in view of its performance analysis.

Analysis of KNN with SVM

In relation to tables 1 and 3, the following table 4 explains the performance comparison of KNN and SVM.

Table 4 Comparison analysis of KNN and SVM

S.No	Performance Metrics	KNN (%)	SVM (%)
1	Accuracy	80.67	85.37
2	Precision	80.3	84.2
3	F-Score	80.01	83.61
4	Recall	80.4	82.34

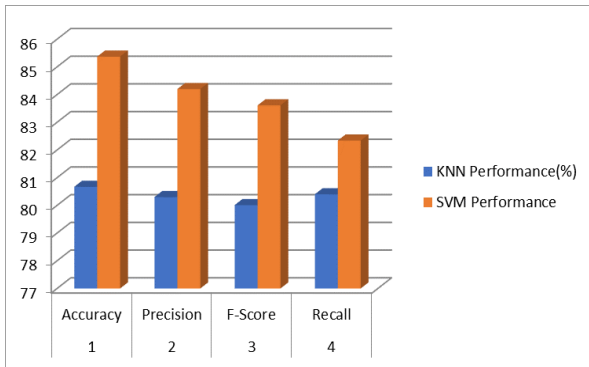


Figure 6. Performance analysis of KNN with SVM

The preceding table 4 addressed the performance analysis of SVM with KNN algorithm to classify the heart arrhythmia disease. The graph below depicts the graphical depiction of the performance analysis.

Figure 6 illustrates a graphical depiction of the evaluation of SVM and KNN. The blue bar and red bar in the image represent the performances of KNN and SVM, respectively. SVM surpassed KNN in all performance measures.

Random Forest Algorithm

This technique incorporates many decision trees that accept input from the dataset and produce output. The most likely outcome will be considered as an output in this scenario. The random forest classification approach is explained below.

$$m_j = n_j C_j - n_{left} C_{left} - n_{right} C_{right} \tag{4}$$

In the aforementioned equation (4), 'm,' C_{left}, and C_{right} denote the node's importance, impurity value, left side sub tree, and right-side sub tree, respectively.

The arbitrary woodland method has been employed to diagnose irregular heartbeat conditions, as demonstrated in table 5.

Table 5. Performance of random forest

S.No	Performance Metrics	Performance (%)
1	Accuracy	99.1
2	Precision	94.2
3	F-Score	93.61
4	Recall	96.34

The preceding table 5 demonstrated the assessment of the random forest algorithm's effectiveness. According to this algorithm, the random forest technique surpassed both KNN and SVM in terms of precision. Figure 7 showcases the examination of the random forest algorithm's effectiveness. Figure 7 depicts the assessment of the random forest method. It demonstrates that the precision of the random forest algorithm is 99.1%. Therefore, this approach outperformed the rest of the algorithms (as stated in table 6).

Table 6 depicted the comparison study of several classifiers. Regarding the numerical values in the table, the RF algorithm produces a greater degree of accuracy compared to other classifiers, as shown in figure 8.

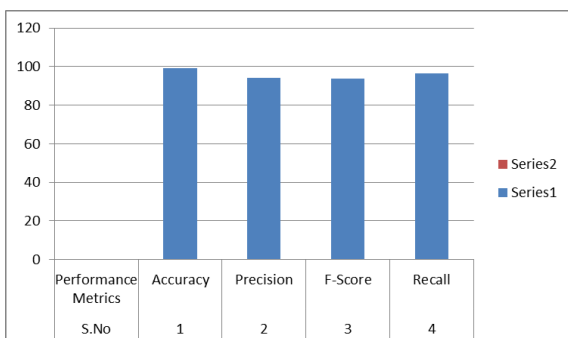


Figure 7. Performance analysis of Random Forest Algorithm

Figure 7 demonstrates that the RF algorithm outperformed other classifiers according to their performance parameters such as accuracy, precision, F- Score, and recall.

Table 6. Comparison of performance of KNN, SVM, RF algorithm

S.No	Performance Metrics	KNN (%)	SVM (%)	RF (%)
1	Accuracy	80.67	85.37	99.1
2	Precision	80.3	84.2	94.2
3	F-Score	80.01	83.61	93.61
4	Recall	80.4	82.34	96.34

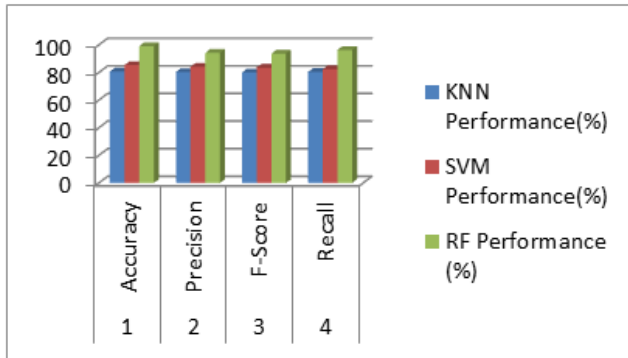


Figure 8 Comparative of performance analysis of KNN, SVM, RF algorithm

The above figure 8 demonstrated the performance of different algorithms. Among the different algorithms RF provides more performance than other algorithms.

5. CONCLUSION

This paper discussed the prediction and classification of arrhythmia disease using a dataset. This paper utilized information from a dataset to categorize by employing various machine learning methods like KNN, SVM, and RF. After exposing the base data to a KNN for the first time and assessing its performance, the data was again exposed to an SVM classifier. SVM was found to outperform KNN based on performance measures that were examined for both models. Last but not least, the RF algorithm was fed the dataset and its performance assessed. KNN, SVM, and RF algorithm performance indicators were compared and analysed. In this analysis, RF provided 99.1% accuracy, which is considerably superior than KNN and SVM. Consequently, the RF technique outperformed other machine learning classifiers in terms of accuracy, as stated in this article.

V. REFERENCES

[1] Prasad, G.K. and Sahambi, J.S. (2020) Classification of ECG arrhythmias using multi-resolution analysis and neural networks. Proceedings of IEEE Conference on Convergent Technologies, 1, 227-231.

[2] Sanamdikar, S., Hamde, S. T., & Asutkar, V. G.(2020). Analysis and classification of cardiac arrhythmia based on general sparsed neural network of ECG signals. SN Applied Sciences, 2(7). <https://doi.org/10.1007/s42452-020-3058-8>

[3] P Vanajakshi , Karthik D R , Somdeb Chanda, 2023, Prediction of Cardiac Arrhythmia using Machine Learning, international journal of engineering research & technology (ijert) Volume 12, Issue 02 (February 2023)

[4] Luz, E., Schwartz, W. R., Cámara-Chávez, G., & Menotti, D. (2016). ECG-based heartbeat classification for arrhythmia detection: A survey. Computer Methods and Programs in Biomedicine, 127, 144–164. <https://doi.org/10.1016/j.cmpb.2015.12.008>

[5] P Vanajakshi , Karthik D R , Somdeb Chanda, (February 2023), Prediction of Cardiac Arrhythmia using Machine Learning, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 12, Issue 02 .

- [6] Aziz, S., Ahmed, S., & Alouini, M. (2021). ECG-based machine-learning algorithms for heartbeat classification. *Scientific Reports*, 11(1). <https://doi.org/10.1038/s41598-021-97118-5>
- [7] Vishwa, A., Lal, M., Dixit, S. G., & Vardwaj, P. (2011). Clasification Of Arrhythmic ECG Data Using Machine Learning Techniques. *International Journal of Interactive Multimedia and Artificial Intelligence*, 1(4),67. <https://doi.org/10.9781/ijimai.20.11.1411>
- [8] Y. Zhou, H. Liu, J. Cao, and S. Li, "Composite learning fuzzy synchronization for incommensurate fractional-order chaotic systems with time-varying delays," *International Journal of Adaptive Control and Signal Processing*, vol. 33, no. 12, pp. 1739–1758,2019.
- [9] C. Usha Kumari, A. Sampath Dakshina Murthy, B. Lakshmi Prasanna, M. Pala Prasad Reddy, and A. Kumar Panigrahy, "An automated detection of heart arrhythmias using machine learning technique:SVM," *Materials Today Proceedings*, vol. 45, pp. 1393–1398, 2021.
- [10] V. Ramalingam, V. Ayantan Dandapath, and M. Karthik Raja, "Heart disease prediction using machine learning techniques : a survey," *International Journal of Engineering & Technology*, vol. 7, p. 684, 2018.
- [11] V. Gupta, M. Mittal, and V. Mittal, "Performance evaluation of various pre-processing techniques for R- peak detection in ECG signal," *IETE Journal of Research*, pp. 1–16, 2020.
- [12] Po-Ya Hsu and C.-K. Cheng, "Arrhythmia classification using deep learning and machine learning with features extracted from waveform-based signal processing," in *Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 292–295, IEEE, Montreal, QC, Canada, July 2020..
- [13] P. Tsinganos and A. Skodras, "On the comparison of wearable sensor data fusion to a single sensor machine learning technique in fall detection," *Sensors*, vol. 18, no. 2, p. 592, 2018.
- [14] R. J. P. Princy, S. Parthasarathy, P. Subha Hency Jose, Raj Lakshminarayanan, and S. Jeganathan, "Prediction of cardiac disease using supervised machine learning algorithms," in *Proceedings of the 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 570–575, IEEE, Madurai, India, May 2020.
- [15] S. Savalia and V. Emamian, "Cardiac arrhythmia classification by multi-layer perceptron and convolution neural networks," *Bioengineering*, vol. 5, no. 2, p. 35, 2018.
- [16] F. I. Alarsan and M. Younes, "Analysis and classification of heart diseases using heartbeat features and machine learning algorithms", *Journal of big data*, vol. 6, no. 1, 2019.
- [17] Soman Swapna and Vinayakumar, "Automated detection of cardiac arrhythmia using deep learning techniques", *Procedia computer science*, pp. 1192- 1201, 2018.
- [18] A. Mustaqeem, S. M. Anwar and M Majid, "Multiclass classification of cardiac arrhythmia using improved feature selection and SVM invariants", *Computational and mathematical methods in medicine*, pp. 1-10, 2018.
- [19] Dilaveris, P. E., Gialafos, E. J., Andrikopoulos, G. K., Richter, D. J., Papanikolaou, V., Poralis, K., et al. (2000). Clinical and electrocardiographic predictors of recurrent atrial fibrillation. *Pacing Clin. Electrophysiol.* 23, 352–358. doi: 10.1111/j.1540- 8159.2000.tb06761.x
- [20] FitzHenry, F., Resnic, F. S., Robbins, S. L., Denton, J., Nookala, L., Meeker, D., et al. (2015). Creating a common data model for comparative effectiveness with the observational medical outcomes partnership. *Appl. Clin. Inform.* 6, 536–547. doi: 10.4338/ACI-2014-12-CR-012