

Ensemble Machine Learning Framework for Heart Abnormality Classification with Effective Feature Selection

¹Pragash K and ²Jayabharathy J

^{1,2}Department of Computer Science Engineering, Puducherry Technological University, Puducherry, India.
¹pragashkaliyan@gmail.com, ²jayabharathy@ptuniv.edu.in

Correspondence should be addressed to Pragash K : pragashkaliyan@gmail.com

Article Info

Journal of Machine and Computing (<https://anapub.co.ke/journals/jmc/jmc.html>)

Doi: <https://doi.org/10.53759/7669/jmc202505057>

Received 26 May 2024; Revised from 18 October 2024; Accepted 20 January 2025.

Available online 05 April 2025.

©2025 The Authors. Published by AnaPub Publications.

This is an open access article under the CC BY-NC-ND license. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Abstract – Coronary Artery Disease (CAD) is the most common cardiovascular disease. Risk factors impact CAD progression. Diagnostic and therapy methods for this illness include several costly side effects. Consequently, researchers are seeking economical and precise techniques for diagnosing this condition. Machine learning algorithms may assist doctors in the early diagnosis of the condition. Hence this work presents an efficient approach for feature selection and classification of abnormal heart rate patterns by combining Joint Mutual Information (JMI), Quantum Annealing, and a Bayesian ensemble model using CatBoost and XGBoost classifiers. The method starts with Joint Mutual Information to rank features based on their dependency with the target variable, identifying the most informative features for classification. Quantum Annealing, specifically simulated annealing in this case, is then used to optimize the subset of features by exploring the feature space and selecting the most relevant combinations, thus improving the model's performance by avoiding suboptimal solutions. The selected characteristics are then input into a Bayesian ensemble of CatBoost and XGBoost classifiers, which are trained to forecast heart rate irregularities. This ensemble method integrates the advantages of gradient boosting models to improve forecast accuracy and mitigate overfitting. The proposed technique is termed Probability-based Bayesian Statistics with Ensemble Boost (PBSEnsBoost), achieving an accuracy of 98.2%, specificity of 93.6%, sensitivity of 92.6%, and an F1-score of 95.7%.

Keywords – Coronary Artery Disease, Abnormality, Feature Selection, Ensemble, Bayesian, Mutual Information, QR Wave.

I. INTRODUCTION

Cardiovascular Diseases (CVD) are widespread globally, accounting for one-third of annual mortality, with 7.5 million fatalities related to Coronary Heart Diseases (CHD). About 1.8 million of these deaths are sudden and caused by ACS [1]. Heart failure patients are misdiagnosed 16.1% of the time [2]. Coronary calcium scans utilizing X-rays to study arteries can diagnose CHD. CTCA detects CHD with 89% sensitivity and 96% specificity [3,4]. Delivering affordable, high-quality medical treatment is a primary challenge for healthcare organizations. For optimal care, a patient's issues must be accurately diagnosed and suitable treatments must be administered. To address these issues, it is essential to include advanced technology [5]. This entails examining techniques, such as using data analysis to enhance the decision-making processes of healthcare professionals. By using these technologies, healthcare organizations may improve their treatments, therefore establishing a basis for more accurate and efficient patient care [6]. Patient information and treatments are stored in a comprehensive medical data repository. Medical judgments often depend on the healthcare professional's experience, training, and intuition rather than being guided by data-driven insights [7]. Data is crucial in today's changing world. Effective problem-solving relies on it to extract insights, identify patterns, and build crucial links. These innovations improve forecast accuracy and help grasp complex, non-linear relationships across large datasets [8]. This, in turn, enables medical practitioners to make informed judgments and choose the most effective therapies from the outset of a disease. Machine learning, with its capacity for autonomous adaptation and model enhancement, is leading this transformation, enabling medical practitioners to make educated decisions and choose appropriate treatments from the onset of an illness. This evolving domain persistently expands the limits of possibility, ensuring further progress in healthcare and other areas [9]. Machine learning, using data to identify complex patterns and relationships that may elude

human observation, is very beneficial in predicting cardiac disease. It helps doctors diagnose and prognose heart issues quickly and accurately. Beyond diagnosis, machine learning is being used in cardiac care to create personalized therapy procedures. These systems use patient-specific data to prescribe treatment and interventions that improve outcomes and decrease risks [10]. It is essential to assess the likelihood of heart disease manifesting in specific populations using effective data mining techniques; these techniques analyze data to uncover significant patterns and actionable insights within a dataset, which comprises various data relevant to the etiology and resultant characteristics of each monitored patient. Furthermore, to mitigate the influence of trivial attributes in the dataset, feature selection methods are used, wherein characteristics are selected based on their significance relative to the resultant feature [11]. These estimate assessments facilitate the development of a ML model to predict cardiac disease in the testing data based on the patterns or insights acquired from the training data. Hence, the contributions are:

- A hybrid feature selection method that combines Joint Mutual Information (JMI) and Quantum Annealing to optimize feature subset. Quantum Annealing further improves this process by considering the global optimization of feature subsets, rather than relying solely on heuristic methods like forward or backward selection.
- This approach ensures that the selected features are both statistically significant and non-redundant, leading to better generalization and reduced overfitting in machine learning models.
- Hybrid Joint Mutual Information (JMI) Quantum Annealing identifies a promising subset of features by adopting mutual information, dramatically reducing the search space for quantum annealing. It significantly reduces the dimensionality of the data, which decreases the computational load in subsequent processing stages.
- probability Bayesian statistics-based ensemble boosting classifier combines ensemble learning techniques, such as boosting (e.g., XGBoost, CatBoost), with Bayesian inference to integrate model uncertainty and improve prediction accuracy
- Bayesian statistics offers a framework for probabilistic reasoning, whereby predictions are represented as distributions across potential outcomes derived from chosen attributes rather than just point estimates.

The subsequent sections of this article are delineated below. Section 2 provides the description of the literature survey. The proposed solution is outlined in Section 3. Section 4 encompasses the investigation of performance and outcomes through comparative research. Section 5 concludes the work with future approach.

II. RELATED WORKS

Different ML algorithms were created to diagnose cardiac disorders. The publications differed because researchers used association rules, clustering, and classification algorithms to extract the most important heart disease prediction characteristics with high accuracy. In [12] offered PSO and SFFS for efficient feature selection. NAO classifies normal and abnormal cardiac sounds. It has 98.03% accuracy, 97.64% sensitivity, and 98.43% specificity. In [13], an ensemble model including the top three classifiers—RF, XG boost, Gradient Boost—is created using seven highly significant characteristics and compared utilizing algorithms and ensemble learning approaches. Its accuracy is 96.17%. In [14], a neural network may leverage the inherent linkages between enhanced quaternion domain dynamic weight variables and quaternion-valued input features. Accuracy is 97.2%. To reduce overfitting, [15] uses deep convolutional neural networks (DCNN) with transfer learning. The accuracy is above 0.9200 with sensitivity of 0.8775 and specificity of 0.9637. In [16], closest neighbor/naive Bayes and adaptive neuro-fuzzy inference system (ANFIS) detect seventeen CVD risk variables. Prediction accuracy is 91.95%. In [17] used a KNN, RF, and SVM. KNN has 86% f-measure mean. In combined LBP HOG with Bag of Words (BOW) model and feature fusion. Finally, it uses SVM, which has the best accuracy at 87.35%. Ensemble Cost-sensitive SVM was used to construct linear features from the CTG signal's time-domain representation. Its sensitivity is 85.2%, specificity 66.1%, and quality index 75.0%.

Approaches like particle swarm optimization, sequential forward feature selection, and ensemble models combining classifiers such as Random Forest, XGBoost, and Gradient Boosting Machine have demonstrated high accuracy and efficiency. Neural networks, including quaternion dynamic representation and deep convolutional networks with transfer learning, show promise in mitigating overfitting and capturing complex relationships in data. Adaptive neuro-fuzzy inference systems, combined classifiers, and feature fusion methods like Local Binary Patterns and Histograms of Oriented Gradients further enhance predictive capabilities. In summary, while current models show high accuracy, challenges related to computational efficiency, interpretability, and generalization remain, limiting their real-world application in clinical settings. Hence, proposed Efficient Feature Selection using Joint Mutual Information (JMI), Quantum Annealing, and a Probability-Based Ensemble Machine Learning Network helps to overcome the above-mentioned limitation.

III. PROPOSED METHODOLOGY

In our work we have done an accurate heart disease prediction with binary classification machine learning model. It begins with band-pass filtering the ECG signal to remove noise from sources such as power-line interference and baseline drift, ensuring that only the frequency range relevant to cardiac activity (typically 0.5-100 Hz) remains. Next, Empirical Mode Decomposition (EMD) is applied to further decompose the ECG signal into intrinsic mode functions (IMFs), which helps isolate and remove non-cardiac artifacts such as motion or muscle interference. By discarding the high-frequency IMFs associated with noise and reconstructing the signal from the remaining relevant components, the

signal is denoised and prepared for feature extraction. The feature extraction stage employs several Entropies for quantifying the complexity and irregularity of the signal. From the XL dataset some of the features are selected using Joint Mutual Information (JMI) with quantum annealing follows by Bayesian inference and ensemble classifiers for abnormality detection as in Fig 1.

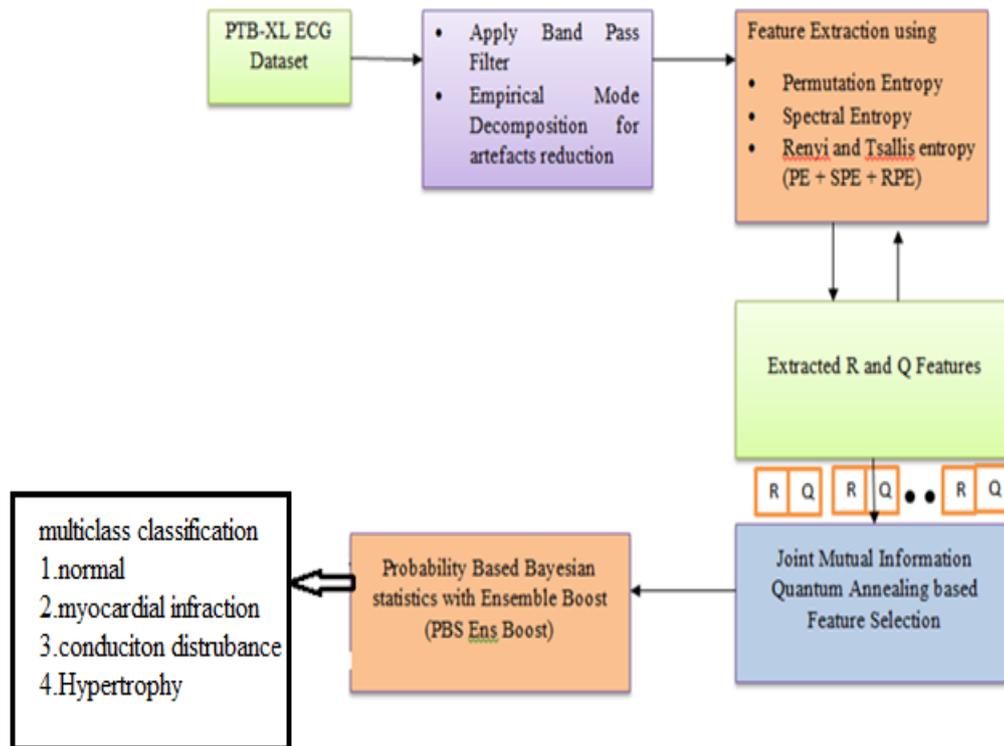


Fig 1. Overall Block Diagram for Coronary Heart Disease Abnormality Detection.

Dataset Description

All ECG data used in this investigation are sourced from the PTB-XL dataset. The PTB-XL database comprises a substantial collection of 21,837 clinical 12-lead ECG recordings. The data has a sample rate of 500 Hz and 100 Hz, with a resolution of 16 bits. Each ECG signal has a duration of 10 seconds and is assessed by cardiologists. The PTB-XL dataset comprises data from 18,885 individuals, with a balanced gender distribution of 52% male and 48% female participants. The dataset has five primary categories: NORM—normal electrocardiogram, CD—myocardial infarction, MI—conduction disturbance, HYP—hypertrophy.

Preprocessing of Data

The fundamental purpose of ECG signal processing is to detect signal components versus noise. This procedure is eq (1). Then, Filter the signal $x(t)$ sampled at $t = t_1, \dots, t_n$ to isolate the information-carrying signal $S(t)$ and remove the distorting interference $n(t)$:

$$x(t) = F[S(t), n(t)] \tag{1}$$

Where, F is a functional, $S(t)$ is the ECG signal, and $n(t)$ is the distorting interference. Usually, it is assumed to be additive:

$$x(t) = S(t) + n(t) \tag{2}$$

Quantile features should be utilized to assess how the bandpass filter cutoff frequency affects ECG signal parameter measurement inaccuracy. Bandpass filters with equations are usually Butterworth bandpass filters used to preprocess ECG signals. This involves filtering out frequencies outside the desired range (e.g., 0.5 to 40 Hz) to retain the essential ECG features while reducing noise. Butterworth filters with infinite impulse response provide flat passband frequency responses. An IIR filter transfer function is

$$H(s) = \frac{B(s)}{A(s)} \tag{3}$$

Where, $B(s)$ and $A(s)$ are polynomials in s and s is the complex frequency variable. Once the filter coefficients a_i and b_i are calculated, we can apply the filter to the ECG signal $x[n]$ by the difference equation

$$y[n] = \sum_{i=0}^N b_i \cdot x[n - i] - \sum_{j=1}^N a_j \cdot y[n - j] \tag{4}$$

Where, $x[n]$ is input signal, $y[n]$ is filtered ECG signal, b_i and a_j Filter coefficients computed for the Butterworth bandpass filter. Moreover, the artefacts refer to specific, non-random distortions caused by external factors or physiological movements unrelated to the target signal. It may be muscle contractions, electrode movement on the skin, breathing-induced baseline wander, and movements during recording. The EMD divides the signal into a few IMFs and a residual value,

$$x(t)_{EMD} = \sum_{i=1}^k IMF_i(t) + r_k(t) \tag{5}$$

The ultimate residual value is r_k , where k is the number of IMFs IMF_i is the i th IMF. As a result noise reduced and artifacts reduced signal is given as follows

$$y_{preprocess} = y[n] + x(t)_{EMD} \tag{6}$$

Feature Extraction Based on Entropy

After preprocessing feature extraction from an ECG signal using Permutation Entropy, Spectral Entropy, and renyi-Tsallis Entropy is a way to capture complex, nonlinear patterns in the signal, which can be useful for classifying and analyzing ECG features

Permutation Entropy

The encoding shows sample rank order in n -length sequences. To define permutation entropy (PE),

$$H_n = - \sum_{j=1}^{n!} p'_j \log_2(p'_j) \tag{7}$$

Where the frequency of symbol sequence patterns, known as permutations, is represented by p'_j . Permutation entropy per symbol is

$$h_n = - \frac{1}{n-1} \sum_{j=1}^{n!} p'_j \log_2(p'_j) \tag{8}$$

Spectral Entropy

The SpE considers the signal's frequency-domain normalized power distribution as a probability distribution,

- Calculation of the Discrete Fourier Transform of the preprocessed time domain data, $x(n)$

$$X(K) = \sum_{n=0}^{N-1} x(n)W_N^{kn}, 0 \leq K \leq N - 1 \tag{9}$$

Where, $W_N = e^{-\frac{j2\pi}{N}}$

The spectral entropy is derived from the Shannon entropy through the formula

$$SpE = - \sum p_k \log p_k \tag{10}$$

Rényi Entropy

Rényi generalizes the information measure to retain event additivity. Diversity indices are represented by ReEn in preprocessed ECG signals. In aberrant preprocessed ECG readings, spikes or high peaks are inferred. Define the ReEn as

$$ReEn(\alpha) = \frac{1}{1-\alpha} \log(\sum p_i^\alpha) \tag{11}$$

Where, p_i is the probability of each amplitude level in the ECG signal, and $\alpha=2$ (here) determines the sensitivity. As a results, the *perm_ent* extract local order and complexity in the ECG signal. SpE helps to measure randomness in frequency domain. ReEn Adds flexibility in quantifying diversity and rare events in the ECG signal distribution. Permutation Entropy (*perm_ent*), Spectral Entropy (SpE), and Rényi-Tsallis Entropy (ReEn) are powerful tools for analyzing R-wave and Q-wave deviations in ECG signals. *perm_ent* quantifies the complexity of signal

patterns, effectively detecting irregularities or abrupt changes in the QRS complex. SpE analyzes the power distribution across frequencies, highlighting shifts in spectral content caused by variations in R-wave amplitude or Q-wave depth. ReEn, which assesses nonlinear and multifractal properties, is sensitive to rare or extreme events and long-range correlations, capturing subtle or significant waveform deviations. Together, these entropy measures identify deviations in amplitude, timing, and morphology of R- and Q-waves, enabling accurate characterization of ECG abnormalities. To identify the R and Q wave Regions, adopt R-peak detection algorithm to locate R peak. Define the regions around R peaks that may contain Q waves or deviations.

These entropy features can be combined to create a feature vector for further analysis. Three entropy measures into a single vector as follows

$$Extract_{feature} = H_{perm_ent}, H_{SpE}, H_{ReEn} \tag{12}$$

Joint Mutual Information Quantum Annealing Based Selection

From the extracted features required features are selected using (JMIQA) Assuming that the joint probability distribution of two features $feat_1$ and $feat_2$ is $p(feat_1, feat_2)$ and their marginal probability distribution is $p(feat_1)$ and $p(feat_2)$ respectively, shared on info The Kullback-Leibler divergence of the joint probability distribution $p(feat_1)$ and $p(feat_2)$ and marginal probability distributions $P(x)$ and $p(y)$ is described as $MI(feat_1, feat_2)$,

$$MI(X, Y) = - \int x \int y f_{(x,y)} \log \frac{f_{x,y}(x,y)}{f_x(x)f_y(y)} dx dy \tag{14}$$

Where, $feat_1 = X$ and $feat_2 = Y$. Thus, feature f_i is more relevant to class label C than to subset S when $I(f_i, S; C) > I(f_i, S; C)$. Combining selected features presupposes $f_i \in F - S$ and $f_s \in S$. When paired with each feature in subset S individually, candidate feature f_i has the lowest joint mutual information with class label C . **Fig 2** shows Flow of Joint Mutual Information Quantum Annealing Based Selection.

After identifying the minimal joint point, quantum simulated annealing is used to represent each feature of a feature subset as a binary decision: 0 means the feature is not selected, 1 means it is selected. We incorporate two coefficients, denoted as α and β to regulate the linear and quadratic terms, respectively, in order to manage both terms effectively. The JMIQA equation can be represented as follow.

$$minimize f(X) = -\alpha \sum_i JMI_{ii}x_i + \beta \sum_{i<j} JMI_{ij}x_ix_j \tag{15}$$

Where, in linear terms, JMI_{ii} refers the joint mutual information among the target and a feature, while JMI_{ij} denotes the MI between two features. The parameters α and β serve as a relative weight, such that $\alpha + \beta = 1$.

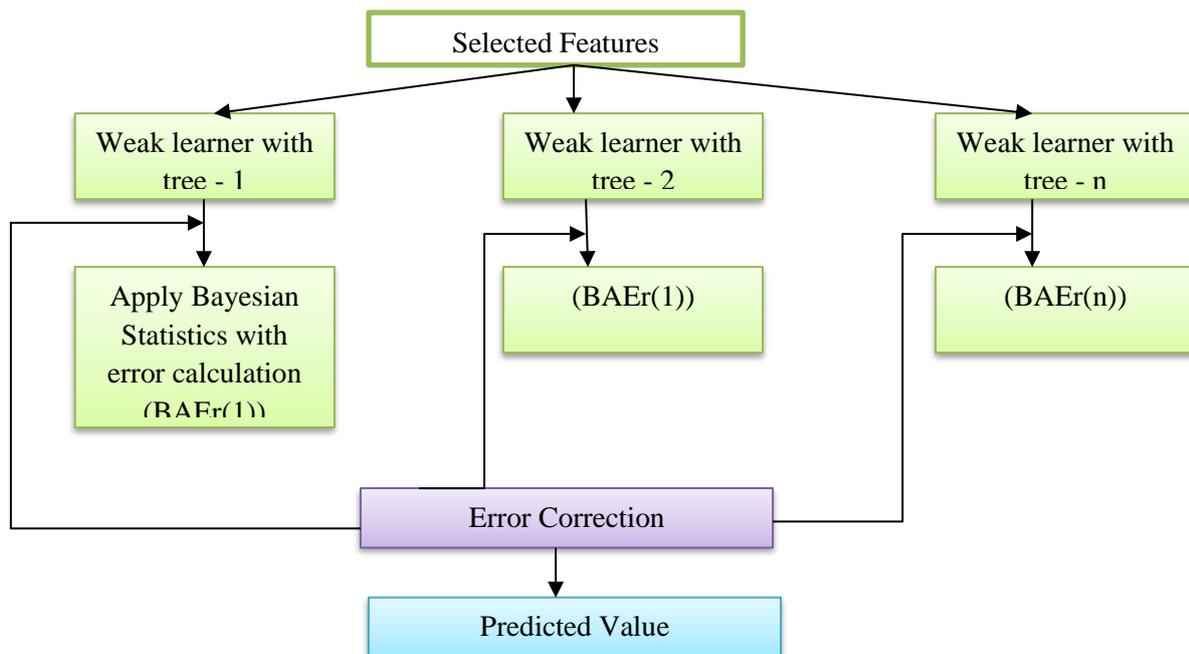


Fig 2. Flow of Joint Mutual Information Quantum Annealing Based Selection.

Algorithm:

Input: Feature set F , class label C , coefficients α and β (where $\alpha + \beta = 1$)
 Output: Selected feature subset S
 (patient_id,report,infraction_stadium1,baseline_drift,extr_beats, pacemaker, filename_lr, R wave and Q wave)
 Step-1 Initialize subset $S = \emptyset$
 Step-2(a) Compute the marginal mutual information
 Step-2(b) Initialize Joint Mutual Information (JMI) score for feature fea_i
 Step-3for each feature compute joint mutual information and update \min_JMI
 Step-4 Define the objective function for quantum simulated annealing
 Step-5 Initialize binary decision variables x_i where,

$$x_i = 1 \text{ if feature } f_i \text{ is selected, otherwise } x_i = 0$$

 Step-6 Apply Quantum Simulated Annealing
 Step-6(a) - Set initial temperature $T > 0$
 Step-6(b) Repeat until convergence
 Step-6(c) Gradually reduce temperature T
 Step-7 Select features f_i with $x_i = 1$ and add them to subset S
 Step-8 Return the final subset S containing the selected features

Probability Based Ensemble Boost Classifier

The selected feature are given for classification. Here, we used two machine learning algorithms such as CatBoost and XGboost with probability-based Bayesian statistics. The CatBoost algorithm adeptly manages categorical characteristics, and the simple method is to substitute them with the mean value of the associated labels. The average value of the label will serve as the criteria for node division in the decision tree. This procedure entails:

$$F_m(x) = F_{m-1}(x) + \vartheta \cdot h_m(x) \tag{16}$$

Where, $F_m(x)$ represents the model's forecast after the addition of m trees. $F_{m-1}(x)$ represents the model's forecast after the incorporation of $m - 1$ trees. ϑ represents the learning rate, and $h_m(x)$ denotes the m -th tree. In conjunction with CatBoost, we use XGBoost to minimise the subsequent objective function, which comprises the loss function and regularisation terms.

$$Loss^{(t)} = \sum_{i=1}^n l(y_i, y_i'^{(t-1)} + f_t(x_i)) + \varphi(f_t) \tag{17}$$

In this context, l denotes the loss function that quantifies the discrepancy between the observed data y_i and the predicted data y_i' , while f_t represents the model at the t -th iteration, with t serving as the iteration index throughout the optimisation process. Bayesian statistics assumes that people evaluate the probability of an event based on its probability and the probability of a dependent event, Where g_i is from catboost and h_i is from XGboost, we may state this mathematically as “Bayes' theorem” or “Bayes' rule”,

$$p(g_i|h_i) = \frac{p(h_i|g_i) \times p(g_i)}{p(h_i)} \tag{18}$$

The objective is to examine the probability of a certain genotype being better, in terms of performance or stability, compared to its counterparts. The pairwise probability of enhanced performance and the pairwise probabilities of enhanced stability were provided, respectively.

$$p(g_i > h_i|y) = \frac{1}{S} \sum_{s=1}^S I(g_i > h_i|y) \tag{19}$$

$$p(var(g_i h_i) < var(g_i h_i|y)) = \frac{1}{S} \sum_{s=1}^S I(var[g_i h_i]^s < var([g_i h_i]^s|y)) \tag{20}$$

Where, $I(g_i^s > g_i^s|y)$ indicates success if g_i^s S i has a higher value than g_i^s , and failure otherwise $(var[g_i h_i]^s < var([g_i h_i]^s|y))$ indicates success if $g_i h_i$ has lower variance. As a result, this will help to classify such as normal, arrhythmia, myocardial infraction and cardiomyopathies, while also taking advantage of Bayesian reasoning for uncertainty estimation and combining multiple classifiers to improve overall performance.

IV. PERFORMANCE ANALYSIS

- Experimental setup - This study employed Google Colab on a Ryzen 7 PC equipped with a 4800-H CPU and 16 GB of RAM. The efficacy of the proposed feature selection and classification techniques is evaluated against existing approaches based on accuracy, specificity, and sensitivity, with the proposed methods demonstrating superior performance. The findings were compared for binary classification between the proposed and existing approaches. The performance of our work on multiclass classification is examined by comparing the suggested PBSEnsBoost with the existing novel automated approach (NAO)[12], deep convolutional neural networks (DCNN) [15] and Ensemble Cost-sensitive Support Vector Machine (ECSVM).
- Accuracy is the ratio of positive predictions to total predictions, including both positive and negative results. The Equation (21) shows,

$$acc = \frac{TP+TN}{TN+FN+TP+FP} \tag{21}$$

Table 1. Comparison Between Existing and Proposed for Accuracy

Number of sample data	NAO	DCNN	ECSVM	PBSEnsBoost
200	98.2	92.4	75	99.2
400	98.1	91.7	74.9	98.5
600	98.2	92.1	74.7	98.2
800	98.5	91.5	75.2	99.4
1000	98.4	91.9	74.7	99.5

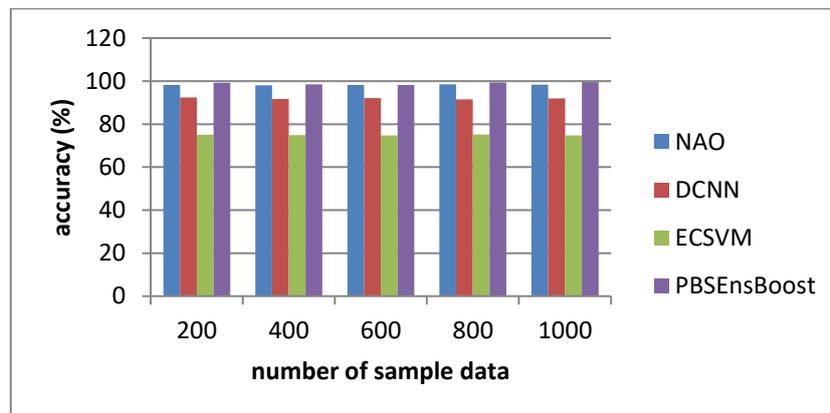


Fig 3. Analysis of Accuracy.

Fig 3 indicates the overall comparison of accuracy. For the increasing order from 200 to 1000 numbers of sample data, the proposed method PBSEnsBoost achieves best accuracy of 99.5%, which is 1.53% better than NAO, 7.5% better than DCNN and 24.5% better than ECSVM. **Table 1** shows Comparison Between Existing and Proposed for Accuracy.

- Specificity pertains to the proportion of correct negative predictions generated by the model relative to the total predictions, both positive and negative. Specificity may be determined using Equation (22),

$$spec = \frac{TN}{TN+FP} \tag{22}$$

Table 2. Comparison Between Existing and Proposed for Specificity

Number of sample data	NAO	DCNN	ECSVM	PBSEnsBoost
200	98.4	96.54	66.1	98.77
400	98.2	96.38	66.9	98.79
600	98.6	96.23	65.9	97.89
800	98.4	96.58	65.3	97.37
1000	98.2	96.36	66.4	98.85

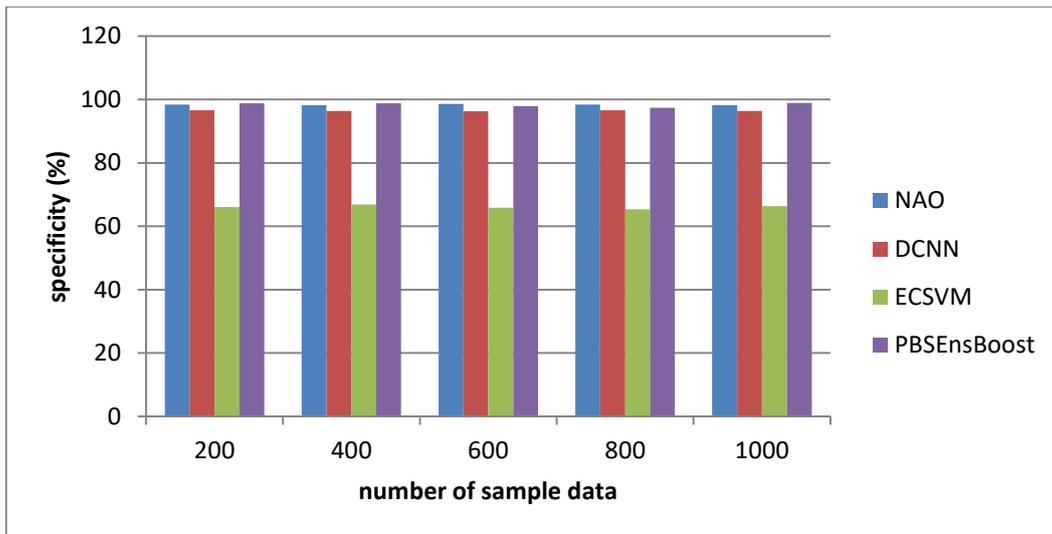


Fig 4. Analysis of Specificity.

Fig 4 indicates the overall comparison of specificity. For the increasing order from 200 to 1000 numbers of sample data, the proposed method PBSEnsBoost achieves best specificity of 98.78%, which is 0.35% better than NAO, 2.41% better than DCNN and 32.68% better than ECSVM. Table 2 shows Comparison Between Existing and Proposed for Specificity.

Sensitivity pertains to the accurate identification of positive predictions made by the developed machine learning model. The specificity may be determined using Equation (23).

$$sen = \frac{TP}{TP+FN} \tag{23}$$

Table 3. Comparison Between Existing and Proposed for Sensitivity

Number of sample data	NAO	DCNN	ECSVM	PBSEnsBoost
200	97.56	87.75	85.56	98.45
400	96.98	87.45	85.31	97.34
600	97.31	86.98	84.85	98.35
800	97.64	86.35	85.84	98.24
1000	97.35	87.68	85.52	98.63

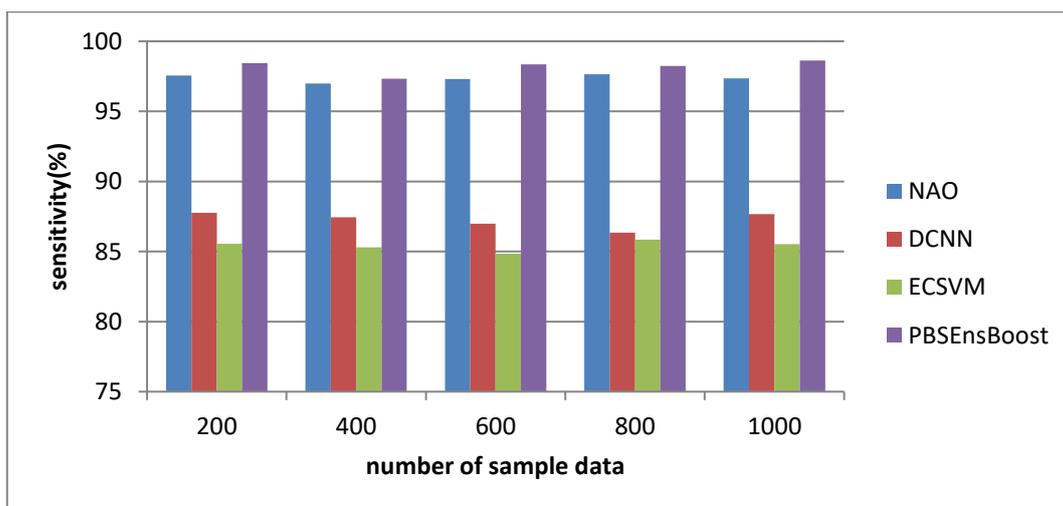


Fig 5. Analysis of Sensitivity.

Fig 5 indicates the overall comparison of sensitivity. For the increasing order from 200 to 1000 numbers of sample data, the proposed method PBSEnsBoost achieves best sensitivity of 98.1%, which is 1.54% better than NAO, 11.65% better than DCNN and 13.1% better than ECSVM. Table 3 shows Comparison Between Existing and Proposed for Sensitivity. Table 4 shows Comparison Between Existing and Proposed for F1-Score.

The F1-score quantifies the harmonic mean of model performance.

$$F1 - score = 2 \times \frac{spec \times sen}{spec + sen} \tag{24}$$

Table 4. Comparison Between Existing and Proposed for F1-Score

Number of sample data	PBSEnsBoost
200	98.4
400	97.8
600	97.2
800	98.2
1000	97.4

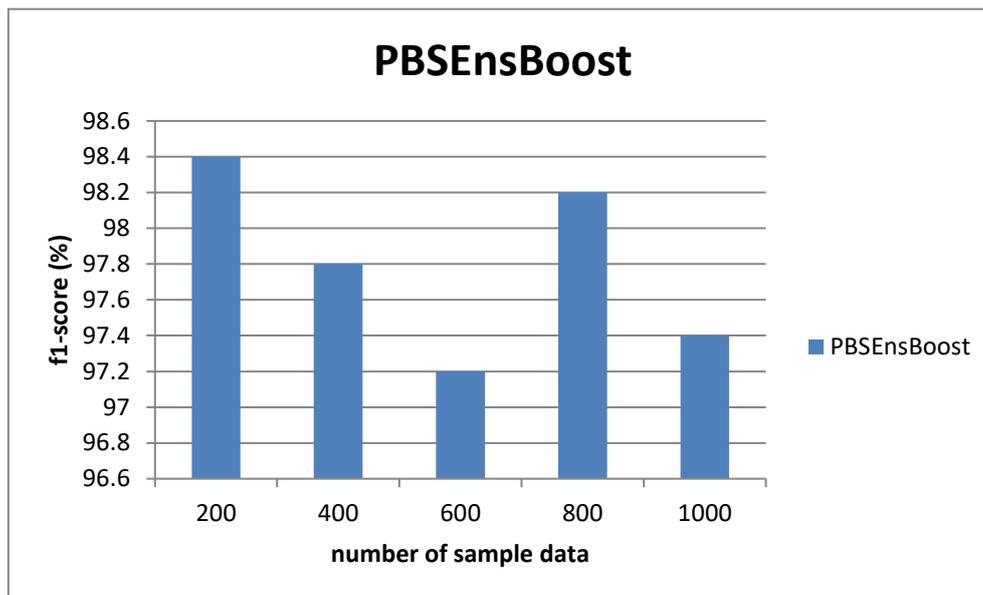


Fig 6. Analysis of F1-Score.

Fig 6 illustrates the comprehensive comparison of the f1-score. As the sample size grows to 200 and 1000, the F1-score exhibits a decrease, reaching its lowest point at around 97.2% for 600 samples. Performance improves again at 800 samples, with the F1-score rising close to 98.3%. However, at 1000 samples, the score drops slightly below 98.0%. These results suggest that the algorithm performs best with smaller or intermediate sample sizes and experiences variability in performance with larger datasets, potentially due to computational or data distribution factors. **Table 5** shows Overall Comparative Analysis Between Existing and Proposed Methods for Binary Classification.

Table 5. Overall Comparative Analysis Between Existing and Proposed Methods for Binary Classification

Parameters	Accuracy (%)	Specificity (%)	Sensitivity (%)	F1-score (%)
NAO [12]	98.03	98.43	97.64	-
DCNN [15]	92	96.37	87.75	-
ECSVM [19]	75	66.1	85.2	-
PBSEnsBoost (proposed)	98.2	93.6	92.6	95.7

V. CONCLUSION

This research examined the efficacy of Joint Mutual Information Quantum Annealing for choosing coronary heart disease variables to forecast the disease's existence using the PBSEnsBoost model. The findings demonstrated that JMIQA is crucial for selecting relevant characteristics and minimising training duration. The JMIQA proficiently discerns the attributes in a training dataset that are most pertinent for forecasting CHD. JMIQA maintains the feature significance of feature-selected data and closely resembles the original data based on the aforementioned finding. Nonetheless, a disadvantage of this work is that the suggested PBSEnsBoost model has not been evaluated on diverse datasets. It is advised that the JMIQA be tested on other datasets in further research to confirm the results of this investigation.

CRedit Author Statement

The authors confirm contribution to the paper as follows:

Conceptualization: Pragash K and Jayabharathy J; **Methodology:** Pragash K; **Software:** Jayabharathy J; **Data Curation:** Pragash K and Jayabharathy J; **Writing- Original Draft Preparation:** Pragash K and Jayabharathy J; **Visualization:** Pragash K; **Investigation:** Pragash K and Jayabharathy J; **Supervision:** Jayabharathy J; **Validation:** Pragash K; **Writing- Reviewing and Editing:** Pragash K and Jayabharathy J; All authors reviewed the results and approved the final version of the manuscript.

Data Availability

No data was used to support this study.

Conflicts of Interests

The author(s) declare(s) that they have no conflicts of interest.

Funding

No funding agency is associated with this research.

Competing Interests

There are no competing interests.

References

- [1]. Cardiovascular diseases (CVDs). (2022). [Online]. Available: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
- [2]. C. W. Wong et al., “Misdiagnosis of Heart Failure: A Systematic Review of the Literature,” *Journal of Cardiac Failure*, vol. 27, no. 9, pp. 925–933, Sep. 2021, doi: 10.1016/j.cardfail.2021.05.014.
- [3]. A. Zriqat, A. M. Altamimi, and M. Azzeh, “A comparative study for predicting heart diseases using data mining classification methods,” *International Journal of Computer Science and Information Security (IJCSIS)*, vol. 14, no. 12, Dec 2016.
- [4]. V. Pham, Q. de Hemptinne, J.-M. Grinda, D. Duboc, O. Varenne, and F. Picard, “Giant coronary aneurysms, from diagnosis to treatment: A literature review,” *Archives of Cardiovascular Diseases*, vol. 113, no. 1, pp. 59–69, Jan. 2020, doi: 10.1016/j.acvd.2019.10.008.
- [5]. N. S. Chandra Reddy, S. Shue Nee, L. Zhi Min, and C. Xin Ying, “Classification and Feature Selection Approaches by Machine Learning Techniques: Heart Disease Prediction,” *International Journal of Innovative Computing*, vol. 9, no. 1, May 2019, doi: 10.11113/ijic.v9n1.210.
- [6]. N. Kannan et al., “Future perspectives on new innovative technologies comparison against hybrid renewable energy systems,” *Computers and Electrical Engineering*, vol. 111, p. 108910, Oct. 2023, doi: 10.1016/j.compeleceng.2023.108910.
- [7]. K. Raza, “Improving the prediction accuracy of heart disease with ensemble learning and majority voting rule,” *U-Healthcare Monitoring Systems*, pp. 179–196, 2019, doi: 10.1016/b978-0-12-815370-3.00008-6.
- [8]. S. Nalluri, R. Vijaya Saraswathi, S. Ramasubbarreddy, K. Govinda, and E. Swetha, “Chronic Heart Disease Prediction Using Data Mining Techniques,” *Data Engineering and Communication Technology*, pp. 903–912, 2020, doi: 10.1007/978-981-15-1097-7_76.
- [9]. B. Dhyanesh, S. G. Ammal, K. Saranya, and K. E. Narayana, “Advanced Cloud-Based Prediction Models for Cardiovascular Disease: Integrating Machine Learning and Feature Selection Techniques,” *SN Computer Science*, vol. 5, no. 5, May 2024, doi: 10.1007/s42979-024-02927-w.
- [10]. S. Asif, Y. Wenhui, Q. ul Ain, Y. Yueyang, and S. Jinhai, “Improving the accuracy of diagnosing and predicting coronary heart disease using ensemble method and feature selection techniques,” *Cluster Computing*, vol. 27, no. 2, pp. 1927–1946, Jun. 2023, doi: 10.1007/s10586-023-04062-2.
- [11]. B. Kolukisa and B. Bakir-Gungor, “Ensemble feature selection and classification methods for machine learning-based coronary artery disease diagnosis,” *Computer Standards & Interfaces*, vol. 84, p. 103706, Mar. 2023, doi: 10.1016/j.csi.2022.103706.
- [12]. B. P. Doppala, D. Bhattacharyya, M. Chakkravarthy, and T. Kim, “A hybrid machine learning approach to identify coronary diseases using feature selection mechanism on heart disease dataset,” *Distributed and Parallel Databases*, Mar. 2021, doi: 10.1007/s10619-021-07329-y.
- [13]. O. Taylan, A. Alkabaa, H. Alqabaa, E. Pamukçu, and V. Leiva, “Early Prediction in Classification of Cardiovascular Diseases with Machine Learning, Neuro-Fuzzy and Statistical Methods,” *Biology*, vol. 12, no. 1, p. 117, Jan. 2023, doi: 10.3390/biology12010117.
- [14]. D. Dudeja et al., “Energy Efficient and Secure Information Dissemination in Heterogeneous Wireless Sensor Networks Using Machine Learning Techniques,” *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–14, Jun. 2022, doi: 10.1155/2022/2206530.
- [15]. B. Wu, P. Liu, H. Wu, S. Liu, S. He, and G. Lv, “An Effective Machine-Learning Based Feature Extraction/Recognition Model for Fetal Heart Defect Detection from 2D Ultrasonic Imageries,” *Computer Modeling in Engineering & Sciences*, vol. 134, no. 2, pp. 1069–1089, 2023, doi: 10.32604/cmescs.2022.020870.
- [16]. R. Zeng, Y. Lu, S. Long, C. Wang, and J. Bai, “Cardiotocography signal abnormality classification using time-frequency features and Ensemble Cost-sensitive SVM classifier,” *Computers in Biology and Medicine*, vol. 130, p. 104218, Mar. 2021, doi: 10.1016/j.combiomed.2021.104218.
- [17]. A. L. Goldberger et al., “PhysioBank, PhysioToolkit, and PhysioNet,” *Circulation*, vol. 101, no. 23, Jun. 2000, doi: 10.1161/01.cir.101.23.e215.