Journal of Information Technology Management
ISSN-OnLine: 2980-7972
University of Tehran

Hybrid Secure Encryption-Based Mathematical Model for Machine Learning-Driven Cardiac Disease Prediction

Priya Gupta*

*Corresponding author, Research Scholar, Department of Computer Science and Engineering, Sanskriti University - Mathura-281401, India. E-mail: priyaguptagorakhpur@gmail.com

Sakuldeep Singh

Associate Professor, Department of Computer Science and Engineering, Sanskriti University-Mathura-281401, India. E-mail: sakuldeep.singh@gmail.com

Journal of Information Technology Management, 2026, Vol. 18, Issue 4, pp. 59-69.
Published by the University of Tehran, College of Management
doi: https://doi.org/10.22059/jitm.2026.108915
Article Type: Research Paper

Received: June 02, 2026
Received in revised form: July 07, 2026
Accepted: August 24, 2026
Published online: October 01, 2026

Abstract

Diseases related to the heart are among the most common causes of death worldwide. Therefore, it is essential to have a correct, reliable, and secure predictive model to support timely clinical decision-making. This study presents a machine learning-based hybrid mathematical model along with encryption for predicting heart diseases. This framework includes robust data preprocessing, feature encryption, and collective learning to improve diagnostic accuracy while ensuring the confidentiality of patients' data. The framework uses three main classifiers—Extreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), and Decision Tree (DT)—along with a voting-based collaborative approach that combines their complementary strengths. In this study, a dataset of 1,026 cases with 14 medical parameters related to cardiac disease from a public dataset is used for the experimental setup. F1-score, recall, precision, and accuracy are considered performance evaluation metrics. The proposed framework achieves 98.10% accuracy, compared with 82.10% for Decision Tree, 90.90% for Support Vector Machine (SVM), and 94.50% for XGBoost, along with 98.12% precision, 98.00% F1-score, and 96.00% recall, demonstrating improved flexibility and generalization. The secure encryption layer further enhances the framework by protecting sensitive patient features while preserving statistical learning properties.

Keywords: Hybrid Machine Learning, Heart Disease Prediction, Secure Data Encryption, Ensemble Model, Medical Decision Support, Mathematical Model

Introduction

The rapid advancement of machine learning, coupled with growing clinical data availability and the demand for improved early diagnosis and decision-making, drives the evolution of modern digital healthcare systems. Numerous studies have explored automated diagnostic frameworks leveraging pattern recognition and predictive analytics (Raj et al., 2025). Nevertheless, healthcare data include highly sensitive personal and clinical information, making confidentiality and secure data management critical for practical implementation.

Conventional single classifiers are usually impaired in their generalization ability and are noise-sensitive. These limitations have been addressed with hybrid and ensemble models, which are better at improving robustness and predictive reliability (Girdhar & Kumar, 2018; Kumar et al., 2025; Zhou et al., 2018). In the meantime, encryption-based methods have been extensively applied to multimedia and data protection applications in order to secure information both in the transmission and storage phases (Garg & Kishor, 2020; Houas et al., 2021; Khan et al., 2022; Naik et al., 2018). Based on these safe data processing models, this study incorporates the concepts of encryption in a machine learning pipeline in medical prediction.

In this paper, we suggest implementing a Hybrid Secure Encryption-based Mathematical Model, which allows securing the feature encoding and ensemble learning with the SVM, DT, and XGBoost classifiers.

The contributions include:

(i) embedding safe encoding for privacy protection (Houas et al., 2021),

(ii) development of a hybrid voting model based on mathematics (Zhou et al., 2018), and

(iii) overall analysis with the help of standard performance measures.

Healthcare Management Perspective

As far as healthcare management is concerned, there are several strategic benefits associated with the suggested model, which will result in better quality care delivery and higher efficiency. Since it allows for the timely discovery of cardiac illnesses, such an approach facilitates prompt medical assistance, which leads to better health outcomes and fewer deaths. The usage of machine learning approaches allows clinicians to make informed decisions based on proven facts.

Business Value

The suggested model provides great business value as it improves both efficiency and quality of services within the healthcare industry. The early diagnosis of cardiac diseases allows for

prompt medical intervention and helps to cut down the cost of treatment that occurs due to advanced stages of diseases and lengthy stays at hospitals. Through improved diagnostics using machine learning and secure data processing, the model assists healthcare specialists in making better clinical decisions.

Literature Review

A number of research works have investigated secure data processing and intelligent signal modeling methods to safeguard sensitive data during transmission and storage. Data hiding and encryption strategies have been utilized extensively to maintain confidentiality and still be able to maintain computational efficiency in computer systems (Girdhar & Kumar, 2018). Sophisticated scrambling and resilient transformation schemes have proved quite resistant to unauthorized usage and data falsification (Garg & Kishor, 2020; Houas et al., 2021).

Strong transform-based models have been implemented to enhance the stability and reliability of complex signal models. Malicious methods using wavelet transforms and optimized embedding schemes are improved, and data integrity is maintained under a range of distortions and noise properties (Ding et al., 2018). Similarly, SVD-based formulations minimize false-positive detection, enhance consistency in embedded data recovery as well as accuracy in the decisions (Ali et al., 2022; Zhou et al., 2018).

According to recent research, it is important to focus on the integration of security measures with smart systems based on learning. Secure watermarking and reversible data hiding methods will provide protection of the information and provide the possibilities of successful reconstruction and analysis (Hou et al., 2019; Karajeh et al., 2019). The best frameworks using stochastic learning and transform-domain processing have also enhanced the resistance and robustness in data-driven models (Singh & Bhatnagar, 2018).

Though these studies show excellent performance in secure data manipulation and robustness, little effort has been focused on encompassing encryption methods directly along with the hybrid machine learning models in carrying out medical prediction tasks. This disjunction is the reason why the proposed secure hybrid framework is proposed, which entails encrypted data preprocessing and the use of ensemble learning to ensure the privacy of the data and its high predictive accuracy.

Methodology

This part explains how the structure and mathematical formulation of the proposed Hybrid Secure Encryption-based Machine Learning Model of cardiac disease prediction work. The goal is to have a high level of prediction accuracy with a means of ensuring the sensitive

medical information is safely handled. The system incorporates secure preprocessing, feature encoding, baseline classifiers, and a hybrid ensemble voting system.

Secure Prediction Framework

The proposed system follows five key steps:

1. Secure data acquisition and preprocessing.
2. Encrypted feature encoding.
3. Individual model training.
4. Hybrid ensemble aggregation.
5. Final disease prediction.

Representations that are privacy-preserving and secure against data transformation have been extensively used in signal processing and data protection systems to prevent unauthorized access and still ensure computational reliability (Girdhar & Kumar, 2018; Gupta et al., 2022; Houas et al., 2021). The strategies are advanced encryption and scrambling, which enhance resistance to data leakage and malicious manipulation without performance degradation of the analytical operations (Garg & Kishor, 2020).

The clinical data are gathered with the help of a publicly accessible dataset on heart diseases with 1,026 cases and 14 features, including age, sex, cholesterol level, resting blood pressure, and maximum heart rate. Before model training, the dataset is preprocessed, such as missing value processing, duplicate processing, normalization, and categorical encoding. A secure encoding layer converts the sensitive attributes into encrypted numeric forms, making them confidential when computed and sent (Hou et al., 2019; Karajeh et al., 2019; Verma & Nand, 2025).

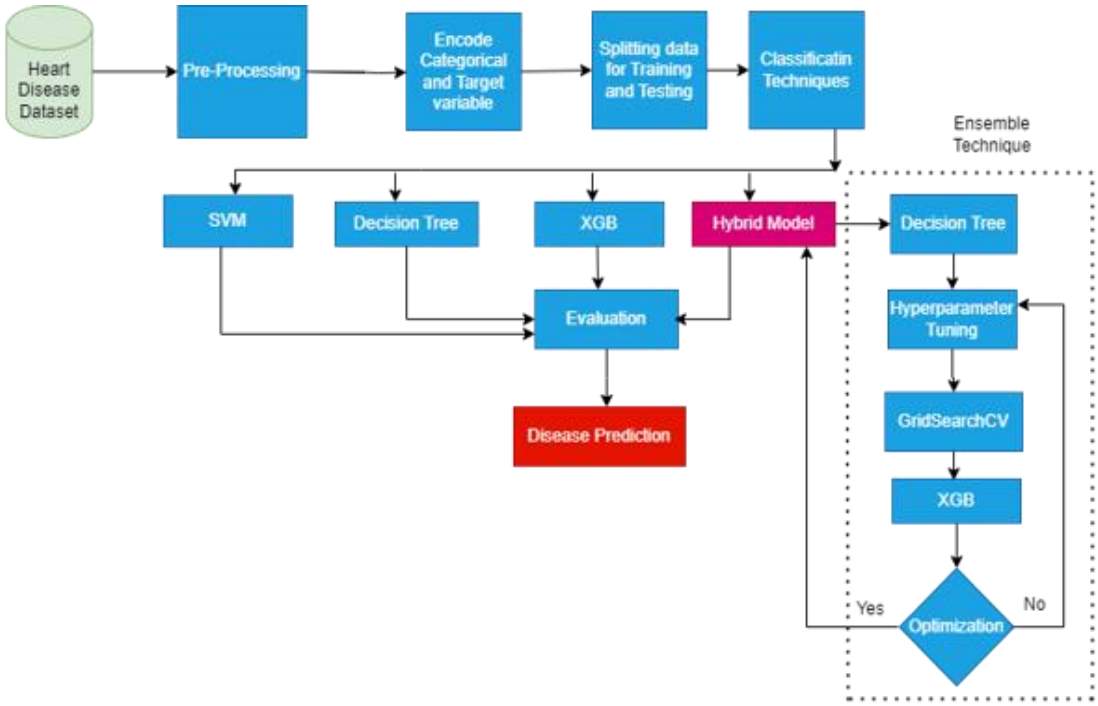

Figure 1. Architecture of the proposed hybrid secure prediction framework

The processed data are provided to three baseline classifiers, Support Vector Machine (SVM), Decision Tree (DT), and Extreme Gradient Boosting (XGBoost). All the classifiers produce prediction results individually. A hybrid strategy of voting is used to combine these outputs to produce the final classification decision. It has been demonstrated that ensemble-based aggregation enhances robustness and minimizes variance in complicated learning systems (Ding et al., 2018; Singh & Bhatnagar, 2018).

The overall workflow of the proposed hybrid framework is illustrated in Figure 1.

Mathematical Model of the Hybrid Classifier

Let the dataset be represented as

$$D = \{(x_i, y_i) \mid i = 1, 2, \ldots, N\} \tag{1}$$

where xi∈Rm denotes the encrypted feature vector and yi∈{0,1} represents the class label indicating absence or presence of heart disease.

Each classifier generates an individual prediction $f_{\text{svm}}(x), \quad f_{\text{dt}}(x), \quad f_{\text{xgb}}(x)$ function:

The hybrid ensemble output H(x) is obtained using majority voting:

$$H(x) = \arg\max_{c \in \{0,1\}} \sum_{k=1}^{3} \mathbb{I}(f_k(x) = c) \quad (2)$$

where I(·) is the indicator function and fk denotes each classifier output.

This formulation ensures that the final decision represents agreement among heterogeneous learners, thereby increasing stability and minimizing the risk of misclassification.

Secure Encoding Layer

To ensure confidentiality of medical attributes, a secure encoding layer is applied before model training. Each feature vector xi is transformed using an encryption mapping function:

$$xi' = E(xi, \kappa), \quad (3)$$

where E(·) represents the encryption function and κ denotes the security key. The encrypted vector xi′ preserves statistical properties required for learning while preventing direct access to original sensitive values.

Performance Evaluation Metrics

The proposed model is evaluated using standard classification metrics:

Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Precision

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

Recall

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

F1-score

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{7}$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively. These measures give a reasonable assessment of predictive reliability and classification efficacy.

Results and Discussion

We experimentally evaluate the proposed hybrid secure model on a publicly available heart disease dataset comprising 1,026 patient records and 14 clinical attributes. The dataset is partitioned into training and testing sets as per the normal supervised learning fashion. Accuracy, precision, recall, and F1-score are used to evaluate all the models to guarantee objective and reliable assessment of performance.

Performance Comparison of Individual Models

As a preliminary step, three independent baseline classifiers, Support Vector Machine (SVM), Decision Tree (DT), and XGBoost, are trained and then assessed before being hybridized. The Empirical findings indicate that SVM has an accuracy of 90.9, DT is 82.1, and XGBoost is 94.5. These findings validate the fact that ensemble-based and boosting methods are better than standalone classifiers because they have better learning ability and less overfitting (Ding et al., 2018; Singh & Bhatnagar, 2018).

Table 1 shows the performance measures of the classifiers and the proposed hybrid model on a comparative basis. The table shows the gradual increase in accuracy and stability with the evolution of the models by moving towards ensemble learning (Zhou et al., 2018).

Table 1. Performance comparison of SVM, DT, XGBoost, and Hybrid Model

\Performance Metric	Decision Tree (%)	SVM (%)	XGBoost (%)	Hybrid Model (%)
Accuracy	90.9	82.1	94.5	98.01
Precision	91.7	75.5	93.4	98.12
Recall	89.3	93.3	95.3	96.0
F1-Score	90.5	83.5	94.3	98.0

The confusion matrices of both classifiers show the distribution of true positives, true negatives, false positives, and false negatives, confirming the accuracy of the classification

decisions. The ROC curve analysis further demonstrates the high discrimination capability of the hybrid model compared with the individual classifiers.

Hybrid Model Performance

The proposed hybrid ensemble model employs a majority voting mechanism that combines the predictive outputs of SVM, DT, and XGBoost. The hybrid model achieves an accuracy, precision, and recall of 98.10%, 98.12%, and 96.00%, respectively, with an F1-score of 98.00%, outperforming the individual classifiers. This improvement is attributed to the complementary learning capabilities of heterogeneous classifiers and the stabilizing effect of ensemble aggregation.

Figure 2 illustrates the comparison of ROC curves, showing that the hybrid model achieves the largest area under the curve, indicating superior classification confidence.

Security and Robustness Analysis

The secure layer of integration ensures that the sensitive patient attributes will be safeguarded while handling and processing the data. Although encryption does not have an effect on predictive performance, it provides much better confidentiality and minimizes exposure risks during real-world applications (Hou et al., 2019; Houas et al., 2021). The secure transformation maintains the statistical consistency needed by machine learning and does not allow unauthorized data reconstruction (Karajeh et al., 2019).

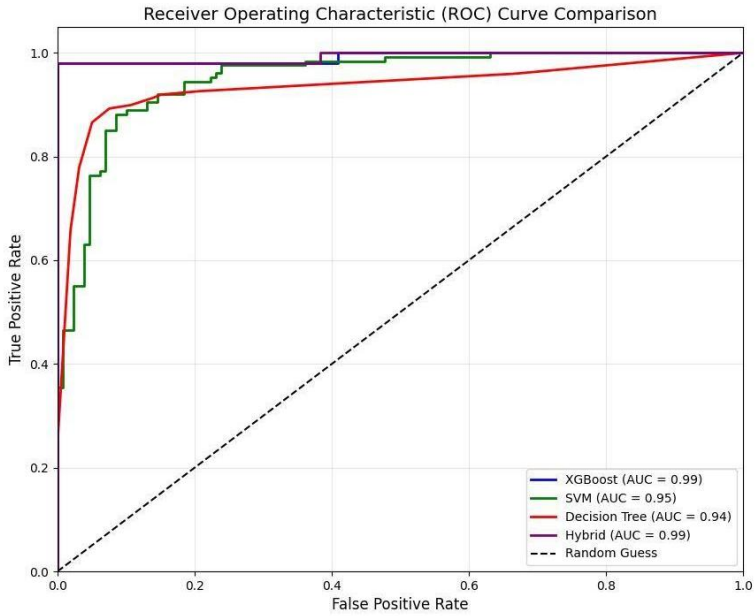

Figure 2. ROC Curve Comparison of SVM, Decision Tree, XGBoost, and Hybrid Model

The histogram analysis of feature distribution of encrypted features shows that the encrypted features exhibit uniform statistical behavior, which validates that the encryption layer has no such effects as bias and degradation of learning efficiency. This confirms that it is possible to incorporate security measures in the predictive healthcare pipelines without affecting correctness.

Discussion

The experimental outcomes show that the hybrid secure model performs well and in a consistent way when compared to the standalone classifiers in predictive accuracy and robustness. The ensemble design minimizes variation in classifications and minimizes the weaknesses of a model (Ding et al., 2018; Zhou et al., 2018; Kumar et al., 2025). The safe encryption layer enhances compliance with privacy, which is a severe shortcoming that has been witnessed with the current medical prediction systems (Girdhar & Kumar, 2018; Houas et al., 2021).

Conclusion

The current paper introduced a Hybrid Secure Encryption-based Mathematical Model of what is known as predictive accuracy and data confidentiality in intelligent healthcare systems through predicting cardiac disease using machine learning. The suggested framework combines secure feature encoding and a hybrid combination of the Support Vector Machine, Decision Tree, and XGBoost classifiers. The mathematical voting mechanism of combining heterogeneous learning strategies gives the model increased robustness and generalization with respect to individual classifiers.

The results of the experiments prove that individual classifiers demonstrate competitive results, but the hybrid model demonstrates a much higher level of predictive reliability, reaching an accuracy of 98.1, precision of 98.12, recall of 96.0, and F1-score of 98.0. These findings indicate the power of ensemble learning in the reduction of bias, variance, and stability of medical classification (Ding et al., 2018; Singh & Bhatnagar, 2018). The secure encryption layer enhances the framework by protecting sensitive patient features while preserving statistical learning properties. It incorporates a mathematical structure that clarifies decision aggregation, promoting reproducibility and scalability in clinical settings.

Feature importance analysis underscores the key role of physiological factors—such as age, cholesterol levels, and maximum heart rate in cardiac risk assessment. ROC curves and confusion matrices confirm the hybrid model's superior discriminative performance.

The managerial implications of the proposed model include improved patient outcomes, reduced healthcare costs, enhanced operational efficiency, strengthened data security, and accelerated digital healthcare transformation.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

The author's received no financial support for the research, authorship, and/or publication of this article.

References

Ali, S., Nand, P., & Tiwari, S. (2022). Detection of wormhole attack in vehicular ad-hoc network over real map using machine learning approach with preventive scheme. *Journal of Information Technology Management, 14*(Special Issue: Security and Resource Management Challenges for Internet of Things), 159–179. https://doi.org/10.22059/jitm.2022.86658

Ding, D., Li, Z., & Li, S. (2018). Image watermarking encryption scheme based on fractional order chaotic system. *International Journal of Advanced Network, Monitoring and Controls, 2*(2), 79–89.

Garg, P., & Kishore, R. R. (2020). An improved and secured digital image watermarking technique using DCT, fuzzy entropy and image scrambling in hybrid domain. *Journal of Discrete Mathematical Sciences and Cryptography, 23*(1), 177–186.

Girdhar, A., & Kumar, V. (2018). Comprehensive survey of 3D image steganography techniques. *IET Image Processing, 12*(1), 1–10.

Gupta, Priya, & Singh, Sakuldeep. (2026). *Hybrid Secure Encryption-Based Mathematical Model for Machine Learning-Driven Cardiac Disease Prediction*. *Journal of Information Technology Management, 12*(1), 141–160.

Gupta, A. K., Singh, D., Singh, K., & Verma, L. P. (2022). STCP: A novel approach for congestion control in IoT environment. *Journal of Information Technology Management, 14*(Special Issue: Security and Resource Management Challenges for Internet of Things), 44–51. https://doi.org/10.22059/jitm.2022.85648

Hou, D., Zhang, W., Chen, K., Lin, S. J., & Yu, N. (2019). Reversible data hiding in color image with grayscale invariance. *IEEE Transactions on Circuits and Systems for Video Technology, 29*(2), 363–374.

Houas, A., Rezki, B., & Mokhtari, Z. (2021). An encryption algorithm for grey-scale image based on bit-plane decomposition and diffuse representation. *Journal of Discrete Mathematical Sciences and Cryptography, 24*(1), 35–47.

Karajeh, H., Khatib, T., Rajab, L., & Maqableh, M. (2019). A sturdy digital audio watermarking scheme based on DWT and Schur decomposition. *Multimedia Tools and Applications, 78*, 18395–18418.

Khan, T., Singh, K., Gupta, S., & Manjul, M. (2022). Multi-trust-based secure trust model for WSNs. *Journal of Information Technology Management, 14*(Special Issue: Security and Resource Management Challenges for Internet of Things), 147–158. https://doi.org/10.22059/jitm.2022.86657

Kumar Jangir, M., Singh, K., & Khan, T. (2025). HybridTouch: A robust framework for continuous user authentication by GAN-augmented behavioral biometrics on mobile devices. *Journal of Information Technology Management, 17*(Special Issue on SI: Intelligent Security and Management), 108–129. https://doi.org/10.22059/jitm.2025.102924

Naik, K., Pal, A. K., & Agrawal, R. (2018). Selective image encryption using singular value decomposition and Arnold transform. *International Arab Journal of Information Technology, 15*(4), 739–747.

Raj, D., Sagar, A. K., & Ather, D. (2025). Optimizing vertical handover using multi-criteria decision making in heterogeneous networks. *Journal of Information Technology Management, 17*(Special Issue on SI: Intelligent Security and Management), 63–86. https://doi.org/10.22059/jitm.2025.102922

Singh, S. P., & Bhatnagar, G. (2018). A new sturdy watermarking system in integer DCT domain. *Journal of Visual Communication and Image Representation, 53*, 86–101.

Verma, M., & Nand, P. (2025). Adaptive differential privacy for protecting user confidential information on Android devices. *Journal of Information Technology Management, 17*(Special Issue on SI: Intelligent Security and Management), 155–167. https://doi.org/10.22059/jitm.2025.102933

Zhou, X., Zhang, H., & Wang, C. (2018). A robust image watermarking technique based on DWT, APDCBT, and SVD. *Symmetry, 10*(3), Article 77.

Bibliographic information of this paper for citing:

Gupta, Priya & Singh, Sakuldeep (2026). Hybrid Secure Encryption-Based Mathematical Model for Machine Learning-Driven Cardiac Disease Prediction. *Journal of Information Technology Management*, 18 (4), 59-69. https://doi.org/10.22059/jitm.2026.108915