

Optimization of referral system for providing medical services to cardiac patients with cardiogenic shock manifestation under uncertainty

Shiva Kanani^a, Iraj Mahdavi^{a,*}, Naghmeh Ziaie^b, Bagher Rahimpour Cami^c

- a. Department of Industrial Engineering, Mazandaran University of Science and Technology, Babol, Iran
- b. Clinical Research Development Unite of Rouhani Hospital, Babol University of Medical Sciences, Babol, Iran
- c. Department of Computer Engineering, Mazandaran University of Science and Technology, Babol, Iran

Abstract

Cardiogenic shock, resulting from cardiac dysfunction, poses a dire threat during cardiac emergencies, necessitating prompt inpatient transfers to intensive care units and aggressive interventions for blood pressure management and adjunctive therapies. Hence, developing an optimal non-invasive decision support system for clinicians is paramount for prognostication and efficient patient transfers to specialized care units. This study aims to enhance the medical referral process for cardiogenic shock patients through Machine Learning (ML) algorithms. Analyzing data from 201 heart patients admitted to emergency wards in 2020, the study employs an Artificial Intelligence (AI)-based model with feature selection and decision phases. The feature selection phase entails analyzing 34 parameters related to the patient's health status, while the decision phase determines treatment outcomes using ensemble-based ML algorithms. Results reveal a mean patient age of 69.44 years, with 57.2% being male, and a concerning 47.7% succumbing within 30 days. Notably, the model's decision phase demonstrates an impressive predictive accuracy of 86% in determining treatment efficacy. Thus, the imperative for an optimal non-invasive decision support system for clinicians is emphasized, enabling proactive prognostication and informed patient transfers to specialized care facilities.

Keywords: Cardiogenic shock; Referral system; Machine learning; Uncertainty; Feature selection

1. Introduction

This paper aims to address two critical components essential for managing the medical referral system catering to cardiac patients specifically focusing on cardiogenic shock and heart failure (HF). The identification component employs machine learning (ML)-based feature selection techniques. Cardiogenic

* Corresponding author. Tel: +98 9111131380

Fax: +(98-11) 3219 0118 E-mail address: irajrash@rediffmail.com

shock denotes a state wherein tissue perfusion diminishes due to cardiac malfunctioning, leading to reduced tissue blood oxygenation and symptoms indicative of lower cerebral blood flow [1]. Heart failure manifests as a clinical syndrome stemming from structural or functional abnormalities in the heart, resulting in reduced blood pumped from the heart or increased intra-cardiac pressure. It presents symptoms such as dyspnea, orthopnea, and lower limb edema; clinical examination often reveals pulmonary congestion and increased jugular vein pressure [2]. In cardiogenic shock, ischemia, decreased myocardial contractility, and cardiac output contribute to an extensive shock necessitating immediate diagnosis and treatment [3]).

Studies indicate that age above 73, a history of cerebrovascular accident, high blood glucose, and serum creatinine at admission are among the risk factors for early and intra-hospital mortality in these patients [4]. Additionally, the therapeutic measures undertaken significantly influence patient mortality rates [5, 6]. Despite recent advancements in science and technology, the management of cardiogenic shock remains a challenge, albeit with a decreasing trend in mortality rates among cardiac patients worldwide [7, 8].

Recent years have witnessed exponential growth in medical device production, data storage capabilities, and data integration, leading to the accumulation of extensive data in repositories. This influx necessitates advanced analytical approaches to extract meaningful insights. While traditional statistical methods such as Generalized Linear Mixed Models (GLMM) focus on analyzing relationships between a limited number of variables, they face limitations in modeling non-linear relationships and handling large-scale data. In contrast, ML techniques excel in modeling complex, non-linear relationships and are adept at handling large-scale data [9, 10]. The healthcare industry generates vast amounts of clinical and administrative data, offering immense potential for personalized care, improved treatment quality, and cost reduction [11]. Data mining techniques play a pivotal role in uncovering patterns and interdependencies within healthcare data, aiding in disease treatment determination, drug efficacy assessment, prediction of medical intervention success rates, and mortality prediction for various diseases. Given that HF is the leading cause of cardiovascular shock, numerous studies have sought to identify predictive factors for cardiovascular shock in HF patients. Identifying these factors can significantly enhance the management of HF patients on cardiovascular medications.

This study aims to identify significant predictive factors for cardiovascular shock in HF patients aged 18 years and above receiving cardiovascular medications, leveraging ML techniques to optimize mortality prediction.

2. Review of the literature

Several studies have explored HF in patients with cardiogenic shock, investigating various aspects such as in-hospital mortality prediction, occurrence of cardiogenic shock, predictors of 30-day mortality in elderly patients with cardiogenic shock, and determinants of disease progression and prehospitalization in HF patients. ML algorithms have emerged as promising tools in predicting, evaluating, and managing diseases while aiding in clinical decision-making.

For instance, Nicolai P. Ostberg et al. (2022) conducted a study focusing on predicting in-hospital mortality among cardiogenic shock patients. Their research found that the LASSO model, with an AUROC of 0.94, outperformed the logistic regression model in predictive performance. Noteworthy predictor factors in the LASSO model included age, Glasgow coma score, D-dimer, lactate, bilirubin, and pH [12].

In another study by Faisal Rahman and colleagues (2022), various ML classification algorithms such as decision trees, random forests, support vector machines, linear discriminant analysis, k-nearest neighbors,

and logistic regression were employed to anticipate cardiogenic shock onset in hospitalized patients with acute HF. Although all algorithms yielded similar results to logistic regression, an algorithm was developed to identify high-risk patients for future cardiogenic shock occurrences. This model facilitated timely adjustments in clinical care [13].

In the 2022 study by Feng Ning Rang and colleagues, entitled 'ML for Predicting Outcomes in Cardiogenic Shock', various models were utilized, including the Cox regression model, the least absolute shrinkage and selection operator (LASSO) regression model, and the Cox regression model. Within the training set, three distinct models were developed. Initially, in the traditional Cox regression model, univariate analysis was employed to select eight predictors with a significance level set at 0.05. Subsequently, predictions underwent two rounds of screening via multiple regression analysis, culminating in the final Cox regression model. Ultimately, six predictors were identified, namely age, heart rate, temperature, white blood cell (WBC) count, anion gap, and blood lactate [14].

Additionally, Xin Li et al. (2022) investigated mortality predictor factors and therapeutic interventions in hospitalized cardiogenic shock patients. Their findings underscored the effectiveness of variables such as age, blood glucose, and heart rate in mortality prediction, guiding appropriate treatment strategies [15].

Furthermore, ML models have proven valuable in predicting in-hospital mortality following medical interventions, as demonstrated by Krittanawong et al. (2022). Their research identified chronic kidney disease as a prominent mortality predictor using neural network models [16].

Moreover, Chang Y et al. demonstrated that utilizing ML models, in conjunction with data from patients' electronic medical records, could predict cardiogenic shock two hours earlier. They also identified associations between cardiogenic shock and several factors, including advanced age, male gender, higher troponin, higher glucose, lower body temperature, lower pulse pressure, an average level of immature granulocytes, higher oxygen saturation, and lower bicarbonate [17].

Awan SE et al. conducted research employing ML techniques, revealing that a small set of variables selected using ML matched the performance of models utilizing the full set of 47 variables in predicting 30-day readmission or death in patients with HF. Furthermore, they found that the predictive model's performance could be significantly enhanced by transforming the original variables using ML methods [18].

Similarly, Ritu et al. introduced a sequential feature selection method aimed at identifying mortality events in patients with heart disease during treatment by isolating the most critical features. They employed various ML methods, including LDA, KNN, RF, SVM, DT, and GBC. Furthermore, they utilized the confusion matrix, receiver operating characteristic curve, precision, recall rate, and F1-score to validate the results of the sequential feature selection (SFS) algorithm. Their experimental findings showcased that the SFS technique achieved an accuracy of 86.67% for the RF classifier [19].

In a separate study, Alotaibi et al. explored the effectiveness of ML techniques in predicting HF. Leveraging the Cleveland Foundation dataset, they employed diverse ML algorithms such as decision tree, logistic regression, RF, naive Bayes, and support vector machine (SVM) for prediction. Notably, the results indicated that the decision tree algorithm exhibited the highest accuracy in predicting heart disease at 93.19%, followed closely by the SVM algorithm at 92.30%. Alotaibi et al. underscored the potential of ML techniques as effective tools for HF prediction, particularly highlighting the decision tree algorithm for future research endeavors [20].

Hasan & Bao investigated an efficient feature selection method for predicting cardiovascular disease, comparing several algorithms. They evaluated models including RF, SVM, Nearest Neighbor, Naive Bayes, and XGBoost to assess their comparative accuracy and identify optimal predictive analysis. Utilizing artificial neural network (ANN) as a standard for comparison, their results revealed the XGBoost classifier, along with the wrapper method, achieved the highest accuracy at 73.74%, followed by SVC at 73.18%, and ANN at 73.20% [21].

In their 2019 study, Rahma Atallah and Amjed Al-Mousa introduced a ML ensemble technique designed to enhance the accuracy and robustness of predicting the likelihood of heart disease. Their approach combined multiple ML techniques to construct an ensemble model, which, through a majority voting mechanism, achieved a notable accuracy rate of 90%. This surpassed the accuracy attained by individual classifiers such as the Stochastic Gradient Descent (SGD) Classifier, K-Nearest Neighbor Classifier, Random Forest Classifier, Logistic Regression Classifier, and Ensemble Classifier [22].

In contrast to prior studies, as demonstrated in Table 1, which predominantly focused on specific age groups or mortality predictors, the current study endeavors to leverage data from patients visiting the hospital's emergency department. The objective is to discern predictive factors for 30-day mortality among cardiogenic shock patients aged over 18 years. Additionally, this study aims to harness ML algorithms to optimize the prediction of 30-day mortality in this patient population. Furthermore, modeling uncertainty is another important feature of the proposed method.

Table (1) literature review table

3. Materials and methods

Given that cardiogenic shock is the most prevalent heart disease, constituting a medical emergency, optimizing proper therapeutic interventions and reducing mortality rates are paramount. To achieve this optimization, ML methods are employed to uncover patterns, hidden interdependencies, and identify a set of key features. By selecting a subset of these key features, machine training time is reduced, and prediction performance is enhanced [23]. Additionally, such predictive models enable physicians and medical staff to focus solely on the minimum necessary variables for outcome prediction. In this study, the feature selection phase involves the utilization of all factors recorded in patients' medical files (state parameters) and employs ML techniques to rank the factors influencing patient status. Subsequently, in the decision-making phase, ensemble techniques are utilized alongside the effective factors identified in the state feature selection phase to predict optimal behavioral therapies and facilitate patient transfer to appropriate treatment departments. The classification framework for patients is illustrated in Figure 1. The core components of the proposed method encompass data preprocessing, the state feature selection phase, and the determination of therapeutic behaviors. The aim is to propose an optimal non-invasive decision support system for cardiogenic shock patients.

Figure 1. The diagram of the research stages

3.1. The data preprocessing stage

Preprocessing of the dataset is crucial to ensure effective representation of data quality. Techniques such as elimination of missing values, standard scaling (SS), and Min-Max Scaling were applied. Missing value management, an essential preprocessing technique, involved determining and handling missing values within the dataset [24, 25]. Various strategies for managing missing values were explored, including ignoring them, substituting with numerical values, or replacing them with the mean feature value. In the HF patient dataset used in this study, two missing values were identified and replaced with the mean feature values.

3.2. Feature selection phase

After calculating the missing values, identifying important features with a strong and positive correlation with the target variable (labels) becomes essential. Extraction of effective features eliminates useless features for prediction and irrelevant ones. The aim of feature selection is to identify the most important features of HF patients with cardiogenic shock manifestations. Additionally, selecting effective features helps develop a more precise model, minimizes learning time, and enhances learning performance. The method of feature selection in the experiments and analyses is presented in Section 3-4.

Feature selection refers to the process of selecting a subset of features from a set of main features based on a special selection criterion [26]. The main advantages of feature selection methods include reducing algorithm computational time, improving prediction performance, identifying relevant features, enhancing data quality, and saving resources in subsequent stages of data collection.

The primary objective of feature selection is to identify a subset of features that significantly enhance decision-making processes.

$$\max_{\text{imize}_{\text{selected_features}}} DS(\text{selected_features}, A), \quad (1)$$

In Equation (1), DS serves as the decision-making function which, through the application of algorithm A on the selected features, yields optimal results. These features are selected using the following model:

$$\text{selected}_{\text{features}} = \arg \text{Max}_{\text{features} \in \text{Features}} FS(\text{features},.), \text{ where} \quad (2)$$

$$\arg \text{Max}_{\text{features} \in \text{Features}} FS(\text{features}) := \{ \text{feature} \in \text{Features} : FS(\text{features}) \geq FS(_ \text{feature}) \text{ for } _ \text{feature} \in \text{Features}$$

In Equation (2), FS stands for the feature selection function in which $_ \text{features}$ denote a subset of features, while Features represents the complete set of selected features.

Feature selection methods can be classified into three groups: wrapper, filter, and embedded (a combination of filter and wrapper methods). The filter method selects statistics-based features independently of the learning algorithm, resulting in shorter computational time. Statistical criteria such as information gain, chi-

square test, Fisher score, correlation coefficient, and variance threshold are used to understand the importance of features [26].

In contrast, the wrapper method's performance heavily depends on the classifier, where the best subset of features is selected based on classification results. In the wrapper feature selection method, a ML model is used to select important features. Initially, all features in the data are input into the model. Then, the ML model runs on all different combinations of available features, and for each combination, the accuracy of the learning algorithms is calculated on the data. Based on resulting accuracy for each combination, important and less important features are identified, and features that increase the accuracy of the learning algorithms are selected as the chosen features for the model [27-30].

The advantage of the wrapper method includes its higher accuracy in feature selection compared to filter methods. Another advantage is its ability to generalize to different algorithms. However, this method may not be efficient for high-dimensional data due to expensive, time-consuming calculations and sensitivity to the choice of the ML model. Examples of wrapper methods include recursive feature elimination, sequential feature selection algorithms, and genetic algorithms [31-33].

3.3. Recursive feature elimination (RFE)

In this study, the Recursive Feature Elimination (RFE) method has been utilized to select the most crucial predictive features. RFE algorithm stands out for its simplicity and effectiveness in feature selection, making it a popular choice in predictive modeling tasks. This method aims to identify the features within the training dataset that are most relevant for predicting the target variables while discarding less informative features. The process involves recursively training the model and iteratively eliminating features with the least importance, based on the weights assigned by the algorithm [32, 33]. The primary steps involved in the RFE method are outlined below:

1. Feature Initialization: All features are initially selected from the main dataset.
2. Initialization of Feature Selection: The process begins by setting the parameter 'n_features_to_select' to an initial value, typically 5.
3. Model Training: The predictor is trained using the selected features, and the importance of each feature is assessed using coefficients or feature importance functions.
4. Feature Elimination: Features with the lowest importance are eliminated from the current set iteratively until the desired number of features is reached.
5. Iterative Process: Steps 3 and 4 are repeated for different values of 'n_features_to_select', ranging from 6 to the total number of features in the dataset.
6. Optimal Subset Identification: Using the results obtained in the previous steps, the optimal subsets of features are determined for each specified number of features.

This iterative approach allows for the identification of the most informative features while gradually reducing the feature space, thus optimizing the model's predictive performance. The effectiveness of the RFE method lies in its ability to systematically evaluate the importance of each feature and select the most relevant ones for prediction, leading to improved model accuracy and generalization across different algorithms. While the RFE method offers notable advantages in feature selection, it may pose challenges for high-dimensional datasets due to its computational complexity and sensitivity to the choice of ML learning models.

3.4. Decision-Making Phase (determining treatment outcomes)

This phase initiates the design of a noninvasive medical decision support system based on classifier models with the effective features identified in the patient's state analysis phase. Classifier methods employed in this stage are among the popular and commonly used methods in heart disease patient data [34, 35]. Classification of the 30-day mortality was trained and evaluated using individual methods of ML learning models. Specifically, 80% of the data (160 patients) were considered for the training dataset, and 20% of the data (41 patients) for the test dataset.

3.4.1 Individual ML Algorithms

Classification is the process of labeling a specific dataset with different classes, applicable to both structured and unstructured data. The goal of classification prediction modeling is to estimate the mapping between input variables and discrete output variables. While various classification algorithms exist, determining the superior algorithm depends on the problem scope and dataset characteristics [33].

Several classic classification methods have been utilized to compare results with the proposed architecture. Given their familiarity, these methods will be briefly introduced.

3.4.1.1 Logistic Regression

Logistic regression, a widely used ML method falling within the category of supervised learning, predicts a categorical dependent variable based on a set of independent variables. For example, it is employed to extract statistically significant cases from the model or predict data trends [34, 35]. The dependent variable in logistic regression is a binary variable, with data coded as 1 (e.g., Yes, Success) or 0 (e.g., No, Failure).

3.4.1.2 Support Vector Machine

Support Vector Machine (SVM) stands as a formidable classifier situated within the supervised ML domain, adeptly employed for both classification and regression tasks. This algorithm proficiently partitions data into distinct categories through the utilization of hyperplanes. SVM showcases prowess in handling multi-dimensional data and effectively managing complex datasets [36].

3.4.1.3 Random Forest

Random Forest stands as one of the methods for classifying datasets, representing an advanced and widely used hybrid method within supervised classification. In this method, a large number of decision trees are constructed for various datasets during the training phase. In the testing phase, each tree within this set assigns a class label to individual data points. Subsequently, following the prediction of class labels by each tree, the final decision for each data point is determined by a majority vote among the members of the tree set. The class label garnering the highest number of votes is deemed the correct label for the test data, thereby enhancing prediction accuracy. This methodology is applied to each data point within the dataset [28].

3.4.2 Decision Component

In the feature selection phase, the identification of factors influencing a patient's condition entails the analysis of 34 parameters pertaining to the patient's health status. Subsequently, in the decision phase, treatment outcomes were determined utilizing ensemble-based ML algorithms. The decision component

implies ML techniques in their hierarchical approaches to create a suitable model for effective referral system.

The following section outlines the main approach to combining ML techniques and then proposes a method for decision-making systems (DS).

3.4.2.1 Hybrid ML Algorithms

Hybrid ML methods, or group models, constitute a subset of ML methods wherein several models, referred to as weak learners or base models, are trained to solve a problem and subsequently combined to yield improved results [28]. When appropriately amalgamated, weak models can generate models boasting higher accuracy and more precise predictions. The hybrid methods employed in this research encompass averaging, weighted averaging, majority voting, and gradient boosting machine (GBM).

For example, GBM Model stands as one of the ML algorithms developed based on decision trees. Leveraging the gradient descent method alongside a set of decision trees, this algorithm comprehensively grasps the intricacies of the training data, thus delivering enhanced performance in predicting new data. In the GBM model, a decision tree is initially constructed, and predictions are made for the training data accordingly. Subsequently, prediction errors are computed for each training sample, and a new tree is generated with the objective of minimizing this error via the gradient descent method. This iterative process is repeated frequently to derive the optimal prediction model. Within the GBM model, each decision tree is bestowed with a weight denoted as the learning rate coefficient. This weight allocates varying degrees of contribution to the final prediction of the model for each decision tree. Additionally, this algorithm employs the regularization method as a means to forestall overfitting and reduce model complexity [37]. The GBM model finds widespread utility across diverse ML applications owing to its commendable performance and stability.

The component of combining for DS to select an appropriate therapeutic behavior is modeled according to Figure 2. Initially selected features enter the combination phase, aiming to construct a fitting ensemble of models. Subsequently, the output results of the model are entered into the behavior selection component in the form of a probability distribution vector, upon which an appropriate outcome is determined. Feature selection is formulated as Equation (1).

Figure 2. General structure of decision-making component

To select appropriate therapeutic behavior for a patient, the parameters of the patient's condition, extracted from the stages, are applied to a hierarchical combined model. This model, with an appropriate combination of well-known ML algorithms such as Logistic Regression, SVM, and RF, creates a suitable hierarchical structure for predicting therapeutic behavior. To create an appropriate combination, based on the selected features, considering Equation (3), function H, which is a combinatorial function of ML algorithms for selecting appropriate combinations and predicting outcomes, generates the best suitable combination. In this function, various combinations of ML algorithms, are performed, and based on the appropriate output, the suitable combination is chosen. MLs in this Equation represents the selected set of algorithms for

creating suitable combinations. and HMLs represents a subset of MLs which are selected for Hybrid Model. Also, Function H, returns the best combination of HMLs.

$$HMLs = \arg \text{Max}_{HMLs \in MLs} H(HMLs), \text{ where} \quad (3)$$

$$\arg \text{Max}_{HMLs \in MLs} H(HMLs) := \{HMLs \in MLs : H(HMLs) \leq H(_ HMLs) \text{ for } _ HMLs \in MLs$$

In this component, based on conducted studies, RF, SVM, and Logistic Regression models have been selected as the candidate ML models for creating an ensemble model. Then, based on the selected appropriate features in the feature selection phase, a combination of these ML models that produces the best result is used as the final model.

Moreover, to model uncertainty in predicting therapeutic behavior, Equation (4) is utilized. In this equation, firstly, according to Equation (5), the outputs of the ML models are filtered, then by creating a probability distribution, the appropriate option for therapeutic behavior is selected.

$$d = \text{choose}(h) := \Pr(d = h_i | h) \quad (4)$$

In this equation, h is the hypothesis vector of different models, which represents the accuracy (probability distribution) of the models. The function Pr also calculates the probability of selecting each hypothesis based on the probability distribution of each one (h_i).

$$h := \{h_i \in H : h_i \geq \alpha\} \quad (5)$$

Accordingly, in the hybrid modeling stage, the output results of each model are expressed as a probability function, and then a continuous probability distribution is created. This distribution function is employed in Equation (4), and the final result is presented for selecting therapeutic behavior. For cases where the probability function is discrete, the categorical distribution (as stated in Equation (6)) is used to create the distribution function and apply it in Equation (4).

$$f(x|P) = \prod_{i=1}^k P_i^{[x=i]} \quad (6)$$

In this Equation, x represents the desired therapeutic behavior to be selected given the probability vector(P). The probability vector p consists of output results from models used in the hybrid stage.

4. Data Analysis and Results

Various assessment criteria were utilized in this study, including precision, F1 score, accuracy, and recall. These metrics were derived from the confusion matrix to assess the effectiveness of 30-day mortality prediction techniques. The confusion matrix enables researchers to evaluate the performance of classification models based on four key factors: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). Enhanced identification of high-risk 30-day mortality cases by the model reflects its precision [38-41].

- **True Positives (TP):** The number of samples correctly predicted as positive.

- **True Negatives (TN):** The number of samples correctly predicted as negative.
- **False Positives (FP):** The number of samples incorrectly predicted as positive.
- **False Negatives (FN):** The number of samples incorrectly predicted as negative.

The specificity measures how precise is attribution to the positive class, which is calculated by the equation (7):

$$specificity = \frac{TN}{TN + FP} \quad (7)$$

Accuracy can be calculated by the Equation (8):

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN \text{ (all data)}} \quad (8)$$

Precision, known as the sum of properly identified cases, refers to the ratio of patients for whom 30-day mortality has been prognosed and actually have the risk of mortality. Precision is calculated by Equation (9):

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

Recall and F1 score are calculated by Equations (10) and (11) respectively in which F1 score represents the harmonic mean of Precision and Recall.

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

$$F - Measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (11)$$

For performance evaluation, a comparison is made between individual models such as SVM, RI, RF, and ensemble models such as AVERAGE, WEIGHTED AVERAGING, MAJORITY VOTING, along with the proposed method using the aforementioned criteria to demonstrate the effectiveness of the proposed model [42-44].

The programming codes utilized in this study were derived from the R programming language version 4.2.2 in the R Studio environment, serving as a robust tool for ML processes. Appropriate packages were employed for the implementation of individual and group models. Additionally, confusion matrices and all Receiver Operating Characteristic (ROC) curves were generated using R software.

4.1 Dataset

The dataset comprised information from 201 patients, each with 34 features (see Appendix), and their respective 30-day mortality rates. These patients, aged above 18 years, were diagnosed with HF and developed cardiogenic shock at the time of referral to Babol Rouhani Hospital. Comprehensive demographic, clinical, and preclinical data were collected via detailed questionnaires, while one-month mortality outcomes were investigated through phone interviews. A census sampling method was employed for data collection. This study, conducted in 2020, constituted a prospective cohort analysis focusing on HF patients exhibiting cardiogenic shock symptoms upon admission.

Patient information, including demographic details, clinical records, and para-clinical data, was meticulously recorded in predefined checklists. Both intrahospital and 30-day mortality rates were ascertained through subsequent phone inquiries. Inclusion criteria encompassed patients above 18 years with HF and cardiogenic shock manifestations at the time of referral or during hospitalization. Table 1 provides a comprehensive overview of the dataset, with output features classified into two categories, indicating 30-day mortality and improvement of HF symptoms with cardiogenic shock manifestations.

The R programming language, executed within the R Studio environment, served as a robust tool for feature selection, model development, and analytical evaluation. Findings are presented in accordance with the research stages. A cohort of 201 HF patients experiencing cardiogenic shock was identified from the cardiac emergency ward of Babol Rouhani Hospital. The mean age of the study population was 69.44 ± 15.71 years, with 47.7% mortality and 52.3% survival rates observed. Gender distribution revealed 57.2% male and 42.8% female patients. Notably, mortality rates among men and women were 45.2% and 51.17%, respectively, while survival rates were 54.8% and 48.83%, as shown in Table 2.

Table 2. The frequency percentage of the prediction variable in terms of gender

4.2 Experiments

Based on the fundamental components of the proposed method, including the selection of appropriate features and the selection of suitable therapeutic behavior, experiments evaluating the proposed method have been conducted in two groups. In the first experiment, the issue of selecting appropriate features has been explored. In the second experiment, the creation of a hierarchical composite model has been examined.

4.2.1 Feature selection

Feature extraction is crucial for identifying relevant variables and eliminating redundant ones in datasets. Our goal in feature selection is to pinpoint key factors among HF patients with cardiogenic shock.

Partial dependency analysis in databases plays a pivotal role in understanding the intricate relationships between variables. This type of dependency unveils how a variable directly influences one or more others within the dataset, disregarding external factors. Such insights are indispensable for meticulous data analysis and elucidating the underlying relationships within databases [45].

Figure 3, illustrates the partial dependence of numerical variables on the 30-day mortality of patients with HF and cardiogenic shock in the cardiac emergency ward of Rouhani Hospital. Notably, increasing age (above 57 years) correlates with higher 30-day mortality rates. This trend is particularly evident in HF patients with cardiogenic shock. Moreover, elevated systolic blood pressure above 75, and diastolic blood pressure above 50, are associated with increased mortality risks compared to their respective lower thresholds. Additionally, the risk of 30-day mortality exceeds 20% for the LVEF feature. Figure 3, further underscores the inverse relationship between hospitalization duration and 30-day mortality risk. Delayed hospital arrival exacerbates mortality risks, while a decrease in pH below 3.7 and elevated lactate levels above 2 are also linked to heightened mortality risks. Notably, a study by Davoodian et al. titled 'The Time and Causes of Mortality in Acute Myocardial Infarction with Cardiogenic Shock' highlights age, time of hospital arrival, high lactate levels, and elevated systolic blood pressure as key factors influencing mortality[46].

Figure 3. Partial dependence of the numerical variables on mortality of patients

The feature selection algorithm RFE has been configured to examine all possible subsets of features. All 34 features have been included in this study. However, while Figure 4 provides an overview of the impact of all 34 features on the 30-day mortality of patients, it is noteworthy that only five features produce results that are reasonably comparable, which can lead to more accurate predictions.

Figure 4. The diagram of accuracy of effective feature selection

Figure 5, indicates the diagram ranking five features selected in the feature selection model via Recursive Feature Elimination (RFE) method. The confusion feature has been identified as the most important one, followed by age of patients, and then CABG, Lactate, and Na.

Figure 5. The bar chart of ranking the features affecting the 30-day mortality prediction

4.2.2 Hierarchical ensemble

In this section, the proposed model is applied to the selected features (with an exception of using main feature in table 3 for performance evaluation) from the previous section, and its results are evaluated against each of the compared models.

Table 3 illustrates that individual models predicting 30-day mortality with the main features, such as SVM and RF, exhibited the best performance, achieving an accuracy score of 0.79. Furthermore, the Precision values for SVM and RF were 0.93 and 0.80, respectively. While Accuracy serves as a fundamental measure of a classifier's quality, it does not distinguish between false negatives (FN) and false positives (FP). In the context of managing cardiogenic shock patients, a higher value of FP is more valuable. According to Formula (17), a lower Precision value for individual models indicates better performance. Therefore, it can

be concluded that the RF model outperformed the SVM model in predicting 30-day mortality in patients with cardiogenic shock using the main features.

Table 3. Comparison of the performance evaluation criteria of 30-day mortality prediction models with principal characteristics

The findings of 30-day mortality prediction of individual models (not combined) with the selected features are summarized in Table 4. It indicates that in the individual models predicting the 30-day mortality, the highest accuracy and precision is associated with the SVM model, with a value of 0.80. The highest F1 score value is observed in the SVM model, with a value of 0.83. Thus, based on the findings presented in Table 4, it can be concluded that the SVM individual model provides the best 30-day mortality prediction for each patient with cardiogenic shock using effective features.

Table 4. Comparison of performance evaluation criteria of the 30-day mortality prediction models with individual models

The evaluation of the proposed methods in Table 5 reveals that in the ensemble models for predicting 30-day mortality, the highest balanced accuracy is associated with the proposed model, measured at 0.87. Similarly, the proposed model exhibits the highest precision, with a value of 0.95. Additionally, the proposed model achieves the largest F1 score value, calculated at 0.86. Consequently, the proposed ensemble model emerges as the most effective predictor of 30-day mortality among patients with HF and cardiogenic shock using the identified features.

Table 5. Comparison of the performance evaluation criteria of the 30-day mortality prediction models of the ensemble models with effective features

The comparison presented in Table 6 evaluates the components of the evaluation criteria for the prediction of 30-day mortality between two models: SVM, an individual ML algorithm, and Proposed Model which is an ensemble-based ML algorithm. The results demonstrate an enhancement in system performance during the determination of patient therapeutic behavior phase when utilizing the proposed ensemble method as the optimal model with highest accuracy (86.0%).

Table 6. Comparison of the performance evaluation criteria of the individual SVM model with Proposed model

Figure 6 depicts comparison of the two confusion matrices reveals that in the Proposed Model, there are more false positives (FP) compared to the SVM method, indicating that patients receive more therapeutic interventions due to more FP in the Proposed Model, indicating increased detectability of the risk of death with Proposed Model. Consequently, enhancing awareness among healthcare providers regarding subsequent therapeutic actions enhances accuracy, thereby leading to improved outcomes for patients with cardiogenic shock.

Figure 6. Comparison of the confusion matrix of SVM, individual model, and the proposed model in decision-making phase

The ROC Curve is a crucial tool for evaluating classification model performance, illustrating the relationship between TP Rate and FP Rate. TP Rate represents correctly identified positive samples, while FP Rate indicates negative samples incorrectly identified as positive. A curve moving upwards and towards the left signifies better model performance. The Area Under the Curve (AUC) reflects model performance, with a larger AUC indicating better performance. This curve serves as a powerful tool for comparing and selecting the best classification model, aiding in the evaluation of classification algorithms.

According to Fig. 7, in ensemble models 30-day mortality prediction with effective features, the proposed model emerges as the optimal model with the highest accuracy (0.86). The ROC plot is a valuable measure for evaluating classifier performance, providing insights into the performance of each algorithm.

Figure 7. Comparison of the ROC curve for individual models and ensemble models predicting the 30-day mortality

Figure 7 illustrates a comparison of the area under the curve (AUC) for individual and ensemble models predicting 30-day mortality based on effective features. The ROC values for the individual models in predicting 30-day mortality for HF patients with cardiogenic shock manifestation using the effective features were 0.88 for logistic regression, 0.85 for RF and 0.85 for SVM, respectively, with the SVM showing the highest accuracy among individual models based on Table 4. Regarding the ensemble models, the Proposed Model achieved the highest ROC value of 0.88, indicating superior performance. These findings suggest that while the use of ensemble models enhances system prediction accuracy during patient state analysis, there is minimal difference in the ROC index between the SVM individual and Proposed ensemble models. The results in Figure 8 compare the ROC curves of two of the most optimal models, SVM and Proposed Model, in the decision phase, showing that with an increase of 3% in Proposed Model, it is accompanied by an improvement in performance.

Figure 8. ROC curve comparison for two SVM and Proposed Model in the decision-making phase

5. Discussion and Analysis

In today's medical research landscape, accurately predicting mortality in heart patients is a crucial endeavor, leading researchers to explore various classification methods. Our study employs meticulous feature selection from the dataset, including confusion, age, CABG, lactate, and Na, are selected from a pool of 34 main features, employing a hybrid approach rooted in recursive feature elimination (RFE) to identify critical features. Evaluation of these models is based on key performance metrics such as accuracy, specificity, F1 score, and precision. In our study, we applied SVM, Logistic Regression (LR), and RF classifiers, achieving accuracies of 0.79, 0.76, and 0.79, respectively. However, recognizing the potential for improvement, we turned to ensemble learning techniques, which involve combining results from different classifiers.

There are several methodologies to model and analyze uncertainty in areas such as logistics, Covid-19 pandemics, and energy so that prediction and robustness were aimed [47, 48, 49, 50]. Ensemble learning has gained attention for its ability to strengthen weaker classifiers, making them more accurate. Unlike using a single algorithm, ensemble methods leverage a mix of classifiers, each contributing its strengths. By carefully selecting these classifiers and how they're combined, ensemble methods can produce more accurate predictions. Our study focused on two key stages: selection of effective features and developing a hierarchical ensemble model for enhancing system performance in selecting an appropriate therapeutic behavior.

There are several ensemble methods available, each with its advantages, such as bagging, stacking, boosting, voting, and averaging. Our study addressed a common issue in medical decision support systems: while many prioritize accuracy, they often ignore uncertainty. To tackle this, we proposed a new ensemble method that not only boosts accuracy but also considers uncertainty.

Our hybrid hierarchical approach involved combining different classifiers, leveraging each one's strengths while compensating for their weaknesses. We also introduced a new way to combine their results, aiming for both high accuracy and acknowledging uncertainty. Comparing our ensemble method to individual classifiers and previous studies showed significant improvements in classification results, leading to selection of suitable therapeutic behavior.

Implementing our improved algorithm demonstrated its effectiveness, achieving an accuracy of 0.86 compared to the existing algorithm's 0.79. Notably, our ensemble method outperformed individual classifiers, highlighting its potential to enhance classification models. This reinforces the value of ensemble learning techniques in improving predictive accuracy in medical research, particularly in predicting mortality in HF patients with cardiogenic shock.

6. Conclusion

In conclusion, optimizing the referral system for providing medical services to cardiac patients with cardiogenic shock manifestation under uncertainty is essential for improving patient care and reducing mortality rates. Through the application of ML techniques, this study has successfully identified key predictive factors and developed accurate prognostic models for 30-day mortality prediction. By leveraging feature selection and ensemble modeling, particularly the proposed model, we have demonstrated the efficacy of ML algorithms in accurately predicting mortality outcomes with high accuracy.

The integration of these predictive models into clinical decision-making processes offers valuable support for healthcare providers, enabling informed treatment strategies, resource allocation, and patient prioritization. By streamlining predictors and optimizing patient management protocols, ML models facilitate efficient and effective healthcare delivery in emergency cardiac settings.

Looking forward, future research endeavors should focus on advancing predictive analytics and personalized medicine approaches in managing cardiogenic shock in HF patients. By refining predictive models and incorporating real-time clinical data streams, healthcare providers can enhance the accuracy and timeliness of mortality predictions, thereby improving patient outcomes. Furthermore, ongoing interdisciplinary collaborations between clinicians, data scientists, and healthcare stakeholders are crucial for driving innovation and translating research findings into clinical practice. By harnessing the power of artificial intelligence and emerging technologies, such as deep learning and precision medicine, we can tailor interventions to individual patient profiles and optimize treatment strategies.

In summary, by embracing technological advancements and fostering collaboration across disciplines, we can pave the way for a future where cardiac patients receive timely and personalized medical services, leading to improved quality of life and better health outcomes.

Since patient mortality in cardiogenic shock is linked to hospitalization days, early discharge increases mortality risk. Future studies should evaluate predictive factors for post-discharge mortality in cardiogenic shock patients to enhance post-discharge patient outcomes. Additionally, predicting mortality risk in other HF groups using ML algorithms is recommended for further research.

References

1. Killip T, 3rd, Kimball JT. "Treatment of myocardial infarction in a coronary care unit. a two year experience with 250 patients", *Am J Cardiol.*, 20(4), pp. 457-464 (1967). doi: 10.1016/0002-9149(67)90023-9.
2. Ponikowski P, Voors AA, Anker SD, et al. "2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) developed with the special contribution of the Heart Failure Association (HFA) of the ESC", *Eur. Heart J.*, 37(27), pp. 2129-2200 (2016). Erratum in: *Eur. Heart J.*, 39(10), p. 860 (2018). doi: 10.1093/eurheartj/ehw383.
3. Jentzer JC, Chonde MD, Dezfulian C. "Myocardial dysfunction and shock after cardiac arrest", *Biomed Res Int.*, 2015, pp. 314796 (2015). doi: 10.1155/2015/314796.
4. Islam MS, Panduranga P, Al-Mukhaini M, et al. "In-hospital outcome of patients with cardiogenic shock complicating acute myocardial infarction: results from royal hospital percutaneous coronary intervention registry, oman ", *Oman Med J.*, 31(1), pp. 46-51 (2016). doi: 10.5001/omj.2016.09.
5. Thiele H, Ohman EM, Desch S, Eitel I, de Waha S. "Management of cardiogenic shock", *Eur Heart J.*, 36(20), pp. 1223-1230 (2015). doi: 10.1093/eurheartj/ehv051
6. Harjola VP, Lassus J, Sionis A, et al. "Clinical picture and risk prediction of short-term mortality in cardiogenic shock", *Eur J Heart Fail.*, 17(5), pp. 501-509 (2015). doi: 10.1002/ejhf.349.
7. Shah AH, Puri R, Kalra A. "Management of cardiogenic shock complicating acute myocardial infarction: A review", *Clin Cardiol.*, 42(4), pp. 484-493 (2019). doi: 10.1002/clc.23168.
8. Alvarez P, Hernandez-Montfort J. "Shock team approaches in managing cardiogenic shock—intersection between critical care and advanced heart failure and transplant cardiology", *Curr. Treat. Options Cardiovasc. Med.*, 22, pp. 20 (2020). doi:10.1007/s11936-020-00820-2.
9. Ngufor C, Van Houten H, Caffo BS, et al. "Mixed effect machine learning: A framework for predicting longitudinal change in hemoglobin A1c", *J Biomed Inform.*, 89, pp. 56-67 (2019). doi: 10.1016/j.jbi.2018.09.001. Epub 2018 Sep 4.
10. Du W, Cheung H, Goldberg I, et al. "A longitudinal support vector regression for prediction of ALS score ", *IEEE Int Conf Bioinform Biomed Workshops*, 2015, pp. 1586-1590 (2015). doi: 10.1109/BIBM.2015.7359912.
11. Veisi H, Ghaedsharaf HR, Ebrahimi M. "Improving the performance of machine learning algorithms for heart disease diagnosis by optimizing data and features" *Soft Comput. J.*, 8(1), pp. 70-85 (2021). doi:10.22052/8.1.70.
12. Ostberg NP, Zafar MA, Mukherjee SK, et al. "A machine learning approach for predicting complications in descending and thoracoabdominal aortic aneurysms", *J Thorac Cardiovasc Surg.*, 166(4), pp. 1011-1020.e3 (2023). doi: 10.1016/j.jtcvs.2021.12.045.
13. Rahman F, Finkelstein N, Alyakin A, et al. "Using machine learning for early prediction of cardiogenic shock in patients with acute heart failure", *J. Soc. Cardiovasc. Angiogr. Interv.*, 1(3), pp. 100308 (2022). doi:10.1016/j.jscai.2022.100308.
14. Rong F, Xiang H, Qian L, et al. "Machine learning for prediction of outcomes in cardiogenic shock", *Front Cardiovasc Med.*, 9, pp. 849688 (2022). doi: 10.3389/fcvm.2022.849688.

15. Li X, Sousa-Casasnovas I, Devesa C, et al. "Predictors of in-hospital mortality among cardiogenic shock patients. Prognostic and therapeutic implications", *Int J Cardiol.*, 224, pp. 114-118 (2016). doi: 10.1016/j.ijcard.2016.09.019
16. Krittanawong C, Hassan Virk H, Hahn J, et al. "Leveraging machine learning models to predict in-hospital mortality after mitralclip", *J. Am. Coll. Cardiol.*, 77, pp. 940 (2021). doi:10.1016/S0735-1097(21)02299-3.
17. Chang Y, Antonescu C, Ravindranath S, et al. " Early prediction of cardiogenic shock using machine learning", *Front Cardiovasc Med.*, 9, pp. 862424 (2022). doi:103389/fcvm.2022.862424.
18. Awan SE, Bennamoun M, Sohel F, et al. "Feature selection and transformation by machine learning reduce variable numbers and improve prediction for heart failure readmission or death", *PLoS One.*, 14(6), pp. e0218760 (2019). doi: 10.1371/journal.pone.0218760.
19. Aggrawal R, Pal S. " Sequential feature selection and machine learning algorithm-based patient's death events prediction and diagnosis in heart disease", *SN Comput. Sci.*, 344 (2020). doi:10.1007/s42979-020-00370-1.
20. Alotaibi F. " Implementation of machine learning model to predict heart failure disease", *Int. J. Adv. Comput. Sci. Appl.*, 10, pp. 637 (2019). doi:10.14569/IJACSA.2019 .0100637
21. Hasan N, Bao Y. "Comparing different feature selection algorithms for cardiovascular disease prediction", *Health Technol.*, 11 (2020). doi:10.1007/s12553-020-00499-2.
22. Atallah R, Al-mousa A. " Heart disease detection using machine learning majority voting ensemble method", In *Proceedings of the 2nd International Conference on New Trends in Computing Sciences (ICTCS 2019)*, pp. 1-6 (2019). doi:10.1109/ICTCS.2019.8923053.
23. Yan K, Zhang D. " Feature selection and analysis on correlated gas sensor data with recursive feature elimination", *Sensors Actuators B Chem.*, 212, pp. 70-75 (2015). doi:10.1016/j.snb.2015.02.025.
24. Ak J, Zongker D. " Feature selection: Evaluation, application, and small sample performance", *Mach. Intell. IEEE Trans.*, 19, pp. 153-158 (1997). doi:10.1109/34.574797.
25. De Stefano C, Fontanella F, Marrocco C, et al. "A GA-based feature selection approach with an application to handwritten character recognition", *Pattern Recognit. Lett.*, 35, pp. 130-141 (2014). doi:10.1016/j.patrec.2013.01.026.
26. G A, Hemalatha S. "Selection of features associated with coronary artery diseases (CAD) using feature selection techniques", *J. Eng. Sci. Technol.*, 13(12), pp. 4034-4051 (2018).
27. Sanchez-Pinto LN, Venable LR, Fahrenbach J, et al. "Comparison of variable selection methods for clinical predictive modeling", *Int J Med Inform.*, 116, pp. 10-17 (2018). doi: 10.1016/j.ijmedinf.2018.05.006.
28. Nazir S, Shahzad S, Mahfooz S, et al. "Fuzzy logic based decision support system for component security evaluation", *Int Arab J Inf Technol.*, (2016).
29. Nazir S, Shahzad S, Riza L. " Birthmark-based software classification using rough sets", *Arab J Sci Eng.*, 42 (2016). doi:10.1007/s13369-016-2371-4.
30. Methaila A, Kansal P, Kumar P. "Early heart disease prediction using data mining techniques", pp. 53-59 (2014). doi:10.5121/csit.2014.4807.
31. Mohd Rosely NFL, Sallehuddin R, Zain A. " Overview feature selection using fish swarm algorithm", *J Phys Conf Ser.*, 1192, pp. 012068 (2019). doi:10.1088/1742-6596/1192/1/012068.
32. Elavarasan D, Vincent P M D, Srinivasan K, et al. " A hybrid CFS filter and RF-RFE wrapper-based feature extraction for enhanced agricultural crop yield prediction modeling", *Agriculture.*, 10, pp. 400 (2020). doi:10.3390/agriculture10090400.
33. Senan EM, Al-Adhaileh MH, Alsaade FW, et al. " Diagnosis of chronic kidney disease using effective classification algorithms and recursive feature elimination techniques", *J Healthc Eng.*, 2021, pp. 1004767 (2021). doi: 10.1155/2021/1004767.
34. Ayon S, Islam M, Hossain R. "Coronary artery heart disease prediction: a comparative study of computational intelligence techniques", *IETE J. Res.*, 68(6), pp. 2488-2507 (2020). doi:10.1080/03772063.2020.1713916.

35. Dissanayake K, Md Johar MG. "Comparative study on heart disease prediction using feature selection techniques on classification algorithms", *Appl. Comput. Intell. Soft Comput.*, 2021, pp. 1-17 (2021). doi:10.1155/2021/5581806.
36. Aleid M, Alyamani K, Rahmouni M, et al. "Modelling the psychological impact of COVID-19 in Saudi Arabia using machine learning", *Comput. Mater. Continua.*, 67, pp. 1-15 (2021). doi:10.32604/cmc.2021.014873.
37. Malik A, Saggi M, Rehman S, et al. "Deep learning versus gradient boosting machine for pan evaporation prediction", *Eng. Appl. Comput. Fluid Mech.*, 16(3), pp. 570-587 (2022). doi:10.1080/19942060.2022.2027273.
38. Fawcett T. "An introduction to ROC analysis", *Pattern Recognit. Lett.*, 27(8), pp. 861-874 (2006). doi:10.1016/j.patrec.2005.10.010.
39. Chawla N. "Data mining for imbalanced datasets: An overview", In: Maimon, O., Rokach, L. (eds) *Data Mining and Knowledge Discovery Handbook*, Springer, Boston, MA, pp. 853-867 (2005). doi:10.1007/0-387-25465-X_40.
40. Witten I, Frank E. "Data mining: Practical machine learning tools and techniques", 1st ed., Elsevier, Amsterdam, Netherlands (2005).
41. Kubat M. "Addressing the curse of imbalanced training sets: One-sided selection", *Proc. 14th Int. Conf. on Machine Learning*, pp. 179-186, San Francisco, CA, USA (2000).
42. Nelay MAI, Nahar N, Hossain MS, et al. "A weighted average ensemble technique to predict heart disease", In: *Trends in Computational and Cognitive Engineering: Proc. 3rd Int. Conf. on TCCE 2021*, Springer, Singapore, pp. 1-12 (2022). doi:10.1007/978-981-16-7597-3_2.
43. Raza K. "Improving the prediction accuracy of heart disease with ensemble learning and majority voting rule", In *Heart Disease Prediction Using Computational Intelligence Techniques*, 1st Ed., Edited by S. S. Kumar and M. Singh, pp. 179-196, Elsevier, Amsterdam, Netherlands (2019). he prediction accuracy of heart disease with ensemble learning and majority voting rule. 2019. p. 179-96. doi:10.1016/B978-0-12-815370-3.00008-6.
44. Brown, G." Ensemble learning", In: *Encyclopedia of Machine Learning*, Sammut, C., and Webb, G.I., Eds., Springer, Boston, MA, pp. 312-320 (2011). doi:10.1007/978-0-387-30164-8_252.
45. Szepannaek G, Lübke K. "How much do we see? On the explainability of partial dependence plots for credit risk scoring", *Argumenta Oeconomica*. 2023(2), pp.137-150 (2023). doi:10.15611/aoe.2023.1.07.
46. Abdulsalam G, Meshoul S, Shaiba H. "Explainable heart disease prediction using ensemble-quantum machine learning approach", *Intell. Autom. Soft Comput.*, 36, pp. 761-779 (2023). doi:10.32604/iasc.2023.032262.
47. Weber, G.-W.; Goli, A.; Tirkolae, E.B. "Logistics and operations modelling and optimization for sustainable supply chain", *Sustainability*, 15(12), pp. 12727 (2023). doi:10.3390/su151712727.
48. Khalilpourazari, S. & Hashemi Doulabi, H., "Robust modelling and prediction of the COVID-19 pandemic in Canada", *Int. J. Prod. Res.*, pp. 1-20 (2021). doi:10.1080/00207543.2021.1936261
49. Kıran, M.S, Özceylan E, Gündüz M, , et al. " A novel hybrid approach based on particle swarm optimization and ant colony algorithm to forecast energy demand of Turkey", *Energy Convers. Manage.*, 53(1), pp. 75-83 (2012). doi:10.1016/j.enconman.2011.08.004
50. Lotfi, R., Kheiri, K., Sadeghi, A. et al. "An extended robust mathematical model to project the course of COVID-19 epidemic in Iran", *Ann Oper Res.*, pp. 1-20 (2022). doi: 10.1007/s10479-021-04490-6.

Appendix

No.	Variable's name	Variable's type	Variable's scale	Variable's role	Measurement unit	Descriptions
1	Mortality	Qualitative	Nominal	Dependent	Yes/No	Intrahospital death + 30-day death
2	Systolic blood pressure	Continuous quantitative	Ratio	Independent	mmHg	Min=55, max=89, mean=81
3	Diastolic blood pressure	Continuous quantitative	Ratio	Independent	mmHg	Min=30, max=71, mean=53.26
4	Pulse rate	Discrete quantitative	Ratio	Independent		Min=71, max=178, median=97
5	Cardiac rhythm	Qualitative	Nominal	Independent	Sinusoidal/non-sinusoidal	1(N=121),2(N=80)
6	Cold extremities	Qualitative	Nominal	Independent	Yes/No	Yes (N=193), No (N=8)
7	Igory	Qualitative	Nominal	Independent	Yes/No	Yes (N=112) – NO (N=89)
8	Confusion	Qualitative	Nominal	Independent	Yes/No	Yes (N=128), No (N=73)
9	LVEF	Continuous quantitative	Ratio	Independent	%	Min=5, max=45, mean=21
10	LVEDD	Continuous quantitative	Ratio	Independent	Mm	Min=38, max=91, mean=58
11	Na	Discrete quantitative	Ratio	Independent	Mmol/L	Min=113, max=160, mean=134
12	K	Discrete quantitative	Ratio	Independent	Mmol/L	Min=2.6, max=7.1, mean=4.4
13	Arterial blood lactate	Discrete quantitative	Ratio	Independent	Mmol/L	Min=0.8, max=15, mean=3.5
14	Arterial blood pH	Continuous quantitative	Ratio	Independent		Min=6.8, max=7.5, mean=7.3
15	Cr	Continuous quantitative	Ratio	Independent	Mmol/L	Min=0.9, max=6.2, mean=2.13
16	Time taken from initiation of clinical symptoms until arriving at hospital	Discrete quantitative	Ratio	Independent	Minute	Hospitalized 0-(64) up to 3(58) – 3 up to 6 (32) – 6 to 12 (21) – 12-24 (17) – more than 24 (6)
17	Duration of hospital stay	Discrete quantitative	Ratio	Independent	Day	Minimum days of hospitalization (1 day) – Maximum days of hospitalization (37 days)
18	Type of received inotrope	Qualitative	Ratio	Independent	Three groups	Group I (N=33), Group II (N=146), Group III (N=23)
19	Duration of received inotrope	Discrete quantitative	Ratio	Independent	Days	Min=1, max=9, mean=7

No.	Variable's name	Variable's type	Variable's scale	Variable's role	Measurement unit	Descriptions
20	MSI	Continuous quantitative	Ratio	Independent		Min=1.73, max=7.45, mean=3.80
21	Shock Index	Continuous quantitative	Ratio	Independent		Min=0.86, max=2.20, mean=1.22
22	Age	Discrete quantitative	Ratio	Other	Years	Min=19, max=96, mean=69
23	Gender	Qualitative	Ratio	Other	Male/female	Male (115) – Female (86)
24	History of coronary artery disease	Qualitative	Nominal	Independent	Yes/No	Yes (113) – No (88)
25	History of MI	Qualitative	Nominal	Independent	Yes/No	Yes (78) – No (123)
26	PCI	Qualitative	Nominal	Independent	Yes/No	Yes (46) – No (155)
27	CABG	Qualitative	Nominal	Independent	Yes/No	Yes (29) – No (172)
28	Hypertension	Qualitative	Nominal	Independent	Yes/No	Yes (105) – No (96)
29	Kidney failure	Qualitative	Nominal	Independent	Yes/No	Yes (25) – No (176)
30	Diabetes	Qualitative	Nominal	Independent	Yes/No	Yes (61) – No (140)
31	Asthma	Qualitative	Nominal	Independent	Yes/No	Yes (16) – No (185)
32	Atrial fibrillation	Qualitative	Nominal	Independent	Yes/No	Yes (77) – No (124)
33	Stroke	Qualitative	Nominal	Independent	Yes/No	Yes (20) – No (181)
34	Ischemic	Qualitative	Nominal	Independent	Yes/No	Yes (113) – No (88)

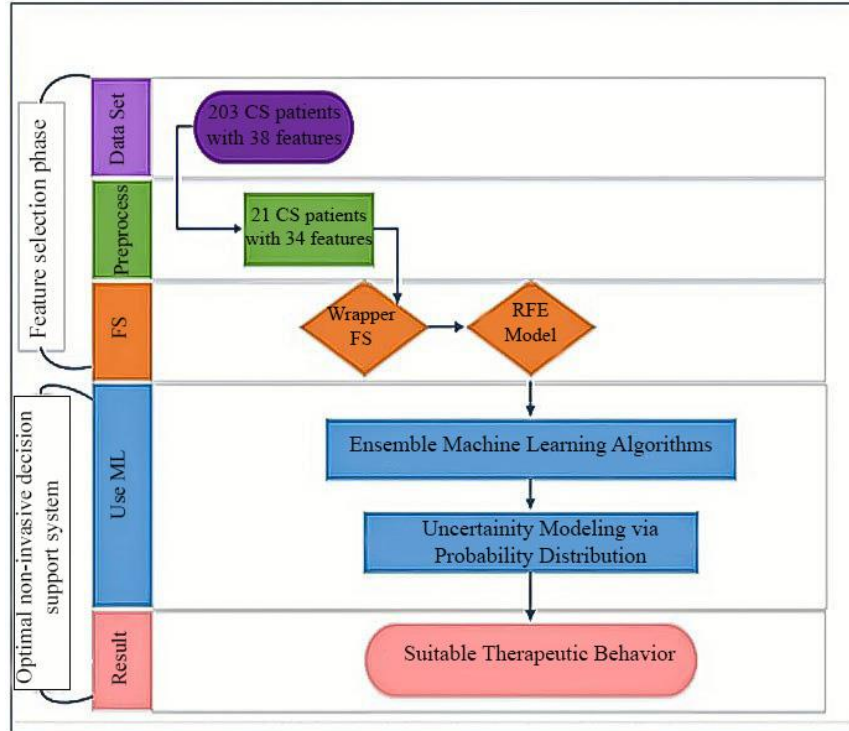


Figure 1. The diagram of the research stages

Entry of Suitable Features

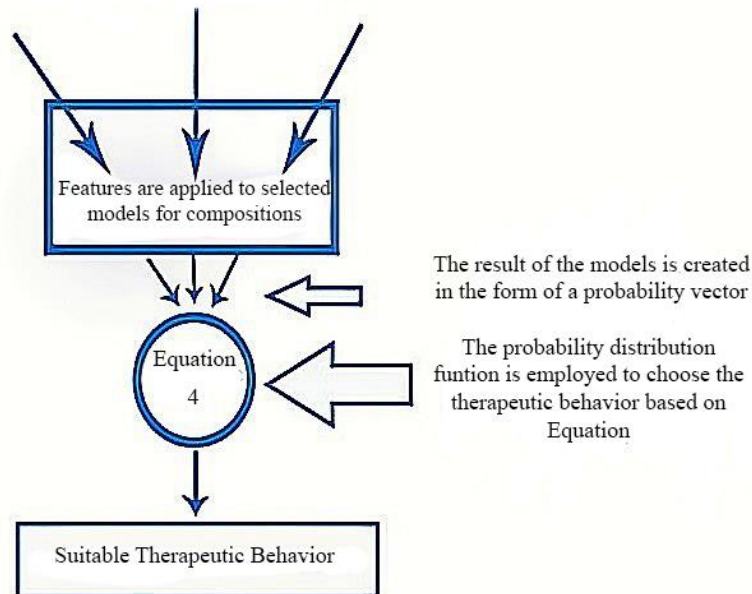


Figure 2. General structure of decision-making component

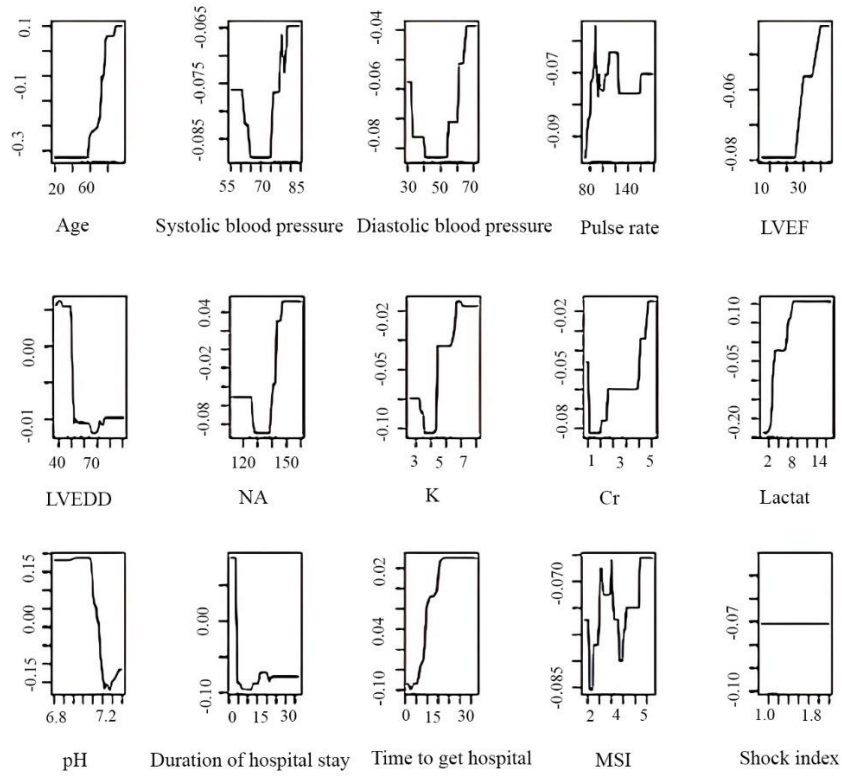


Figure 3. Partial dependence of the numerical variables on mortality of patients

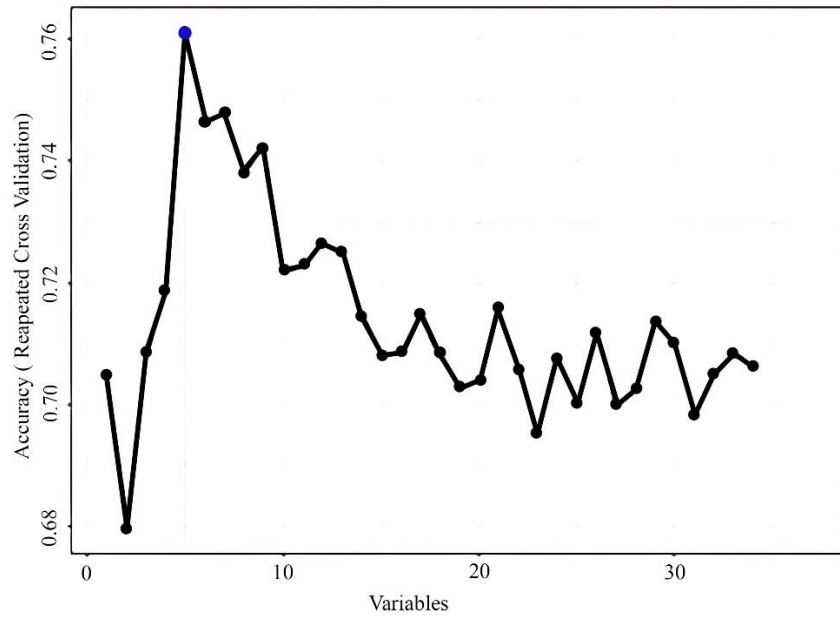


Figure 4. The diagram of accuracy of effective feature selection

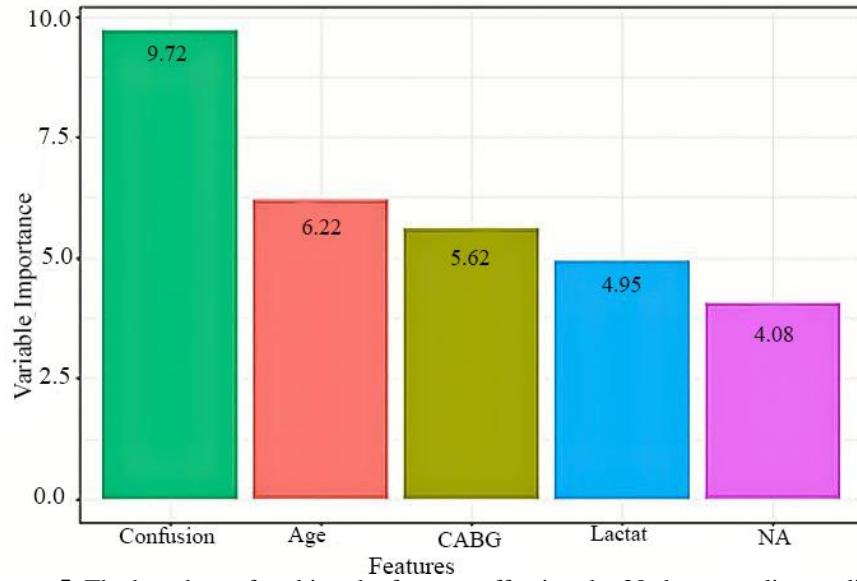
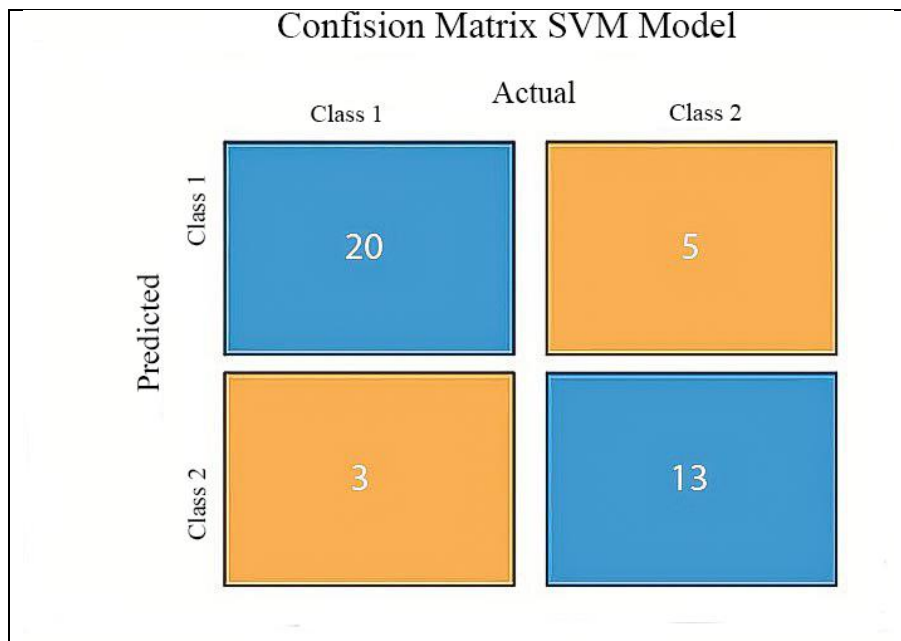


Figure 5. The bar chart of ranking the features affecting the 30-day mortality prediction



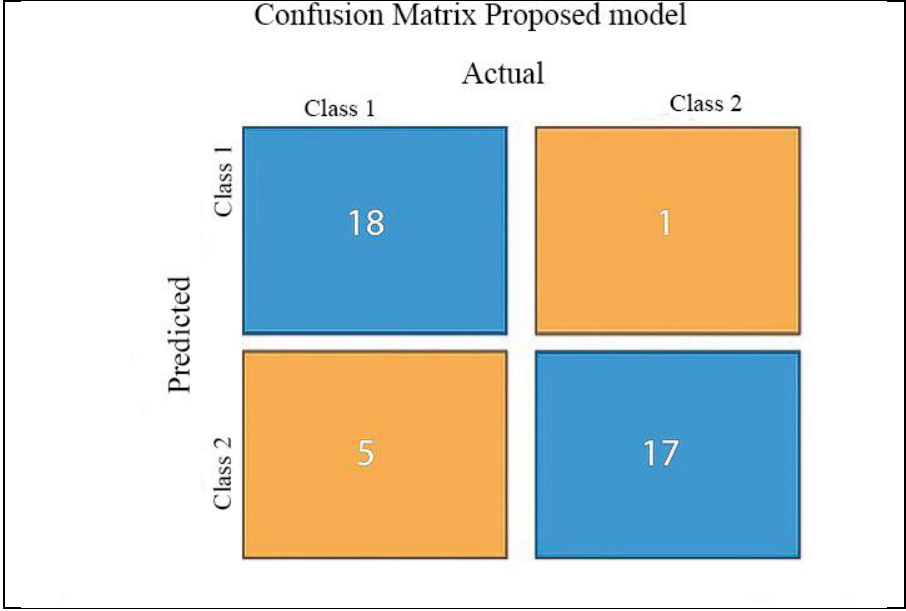
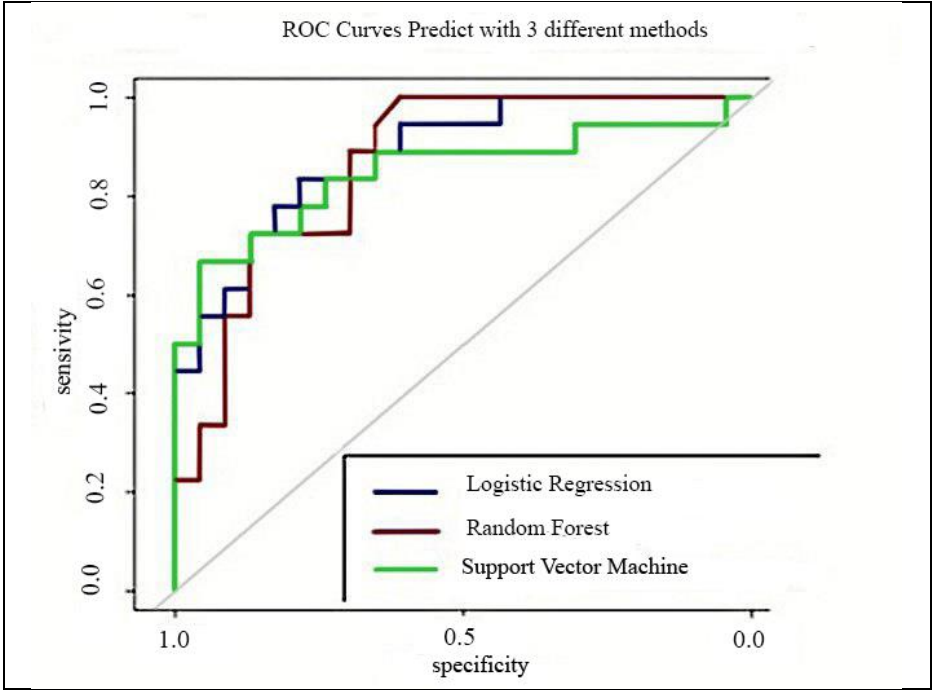


Figure 6. Comparison of the confusion matrix of SVM, individual model, and the proposed model in decision-making phase



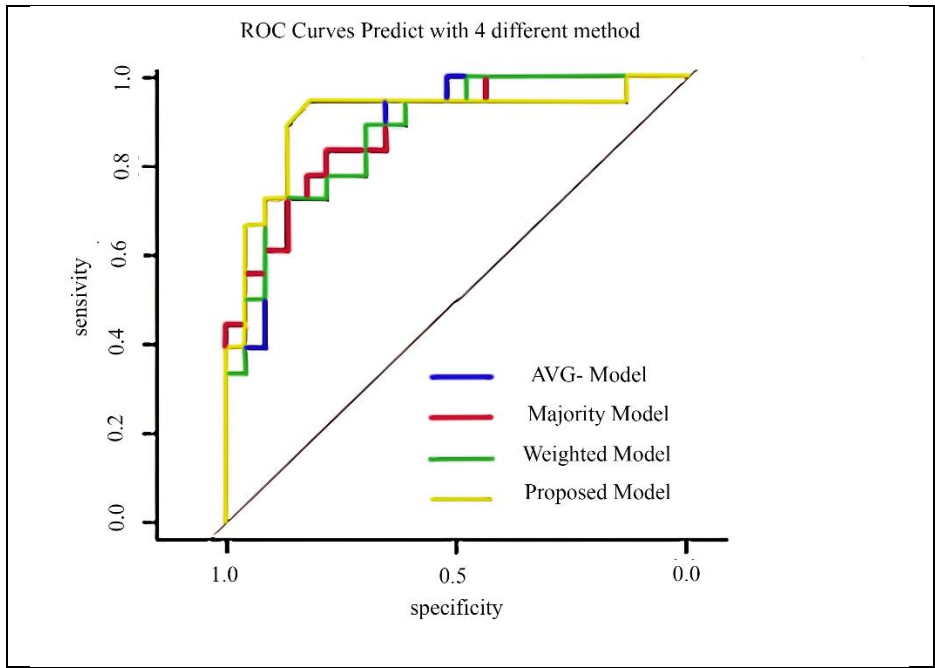


Figure 7. Comparison of the ROC curve for individual models and ensemble models predicting the 30-day mortality

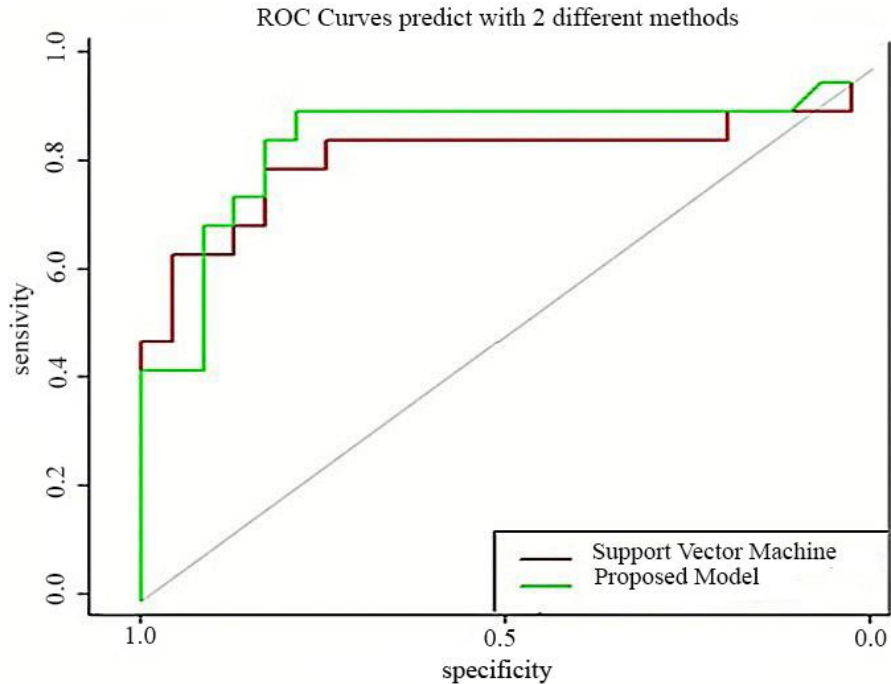


Figure 8. ROC curve comparison for two SVM and Proposed Model in the decision-making phase

Table (1) literature review table

Authors	Novel Approach	Effective Features	Disease & Dataset
Nicolai 2022[12]	LASSO, Random Forest, Gradient Boosted Tree, SVM, Elastic Net	Age, Glasgow coma score, Glasgow coma score, PH, lactate, Bilirubin	Heart Shock - MIMIC-IV
Rahman F 2022[13]	logistic Regression, Random Forest, KNN, Decision tree, SVM, LDA.	Age, systolic blood pressure, heart rate, temperature, blood urea nitrogen, sodium, oxygen saturation, venous pH, hemoglobin, white blood cell count, hydralazine use, trend of respiratory rate, and trend of systolic blood pressure were associated with risk of developing CS	24,461 heart failure patients, 256 patients of cardiogenic shock
Rong F 2022[14]	Least Absolute Shrinkage and Selection Operator (LASSO), Cox Boost model	Age, heart rate, SBP, DBP, respiratory rate, temperature, SpO2, WBC count, RDW INR, PT, anion gap, bicarbonate, blood lactic acid, serum urea nitrogen, and tumor	919 patients from MIMIC-III Affiliated and Hospital Yuying Medical University Children's Hospital of Wenzhou
Xin Li MD 2016[15]	Statistical analysis	Age, Blood Glucose Heart Beat	Cardiogenic Shock, 253 patients
Krittanawong C 2021[16]	Random Forest, SVM, Decision tree, KNN, Boosting	Chronic kidney disease	severe mitral regurgitation
Chang Y 2022[17]	XGB, Multiple Layer Perceptron	Age, Sex, Troponin, Glucose, Low Body Temperature, Lower pulse pressure	Cardiogenic Shock, EHR

Awan SE 2019[18]	Random Forest, SVM, logistic Regression	Age Hospitalization period, Hospitalization period, chronic kidney disease	Heart Failure, HMDC
Alotaibi 2019[20]	Random Forest, Decision tree, logistic Regression, SVM, Naïve Bayes	Predict of heart disease without feature selection	303 patients collected from Kaggle platform
Atallah R 2019[22]	LR, RF, KNN, SGD, Hard voting ensemble model	No mention of feature selection Limited to one dataset	Cleveland Clinic Foundation Heart Disease dataset
Current Study	Feature selection, hybrid model with a hierarchical approach	Age, History of open-heart surgery, Confusion, Lactate level, Na	Cardiogenic Shock, 201 patients

Table 2. The frequency percentage of the prediction variable in terms of gender

Sex	Alive	Mortality	Total
Male	% 54.8	% 45.2	% 57.2
Female	% 48.83	% 51.17	% 42.8

Table 3. Comparison of the performance evaluation criteria of 30-day mortality prediction models with principal characteristics

Index	Logistic regression	RF Model	SVM Model
Specificity	0.91	0.86	0.65
Precision	0.75	0.80	0.93
F-measure	0.92	0.83	0.76
Accuracy	0.76	0.79	0.79
ROC	0.83	0.85	0.89

Table 4. Comparison of performance evaluation criteria of the 30-day mortality prediction models with individual models

Index	Logistic regression	RF Model	SVM Model
Specificity	0.72	0.61	0.72
Precision	0.79	0.74	0.80
F-measure	0.80	0.80	0.83
Accuracy	0.78	0.75	0.80
ROC	0.88	0.85	0.85

Table 5. Comparison of the performance evaluation criteria of the 30-day mortality prediction models of the ensemble models with effective features

Index	Average Model	Majority model	Weighted Average model	Proposed Model
Specificity	0.72	0.72	0.72	0.94
Precision	0.79	0.8	0.79	0.95
F-measure	0.81	0.83	0.81	0.86
Balanced Accuracy	0.77	0.80	0.78	0.87

Table 6. Comparison of the performance evaluation criteria of the individual SVM model with Proposed model

Index	SVM Model	Proposed Model
Specificity	0.72	0.94
Precision	0.8	0.95
F-measure	0.83	0.86
Accuracy	0.8	0.86
ROC	0.85	0.88

Author Biographies

Shiva kanani

A PhD student in Industrial Engineering at Mazandaran University of Science and Technology, this individual is actively involved in teaching at various universities. Their research focuses on data mining, machine learning algorithms, and health management, contributing to advancements in these fields.

Professor Iraj Mahdavi

Professor Iraj Mahdavi is the Director of the Industrial and Computer Engineering College at Mazandaran University of Science and Technology, Babol, Iran. As a professor in the Industrial Engineering Department, he leads research and academic programs focused on integrating advanced engineering methods with practical applications. Professor Mahdavi's work supports the development of innovative solutions in industrial engineering. He can be contacted at PO Box 734, Babol, Iran, or via phone at +(98-911) 113 1380.

Dr. Naghmeh Ziaie

Dr. Naghmeh Ziaie is an Assistant Professor of cardiology at Babol University of Medical Sciences and works in the Clinical Research Development Unit at Rouhani Hospital. Specializing in heart failure and transplantation, Dr. Ziaie focuses on improving clinical practices and patient outcomes in cardiology.

Dr. Bagher Rahimpour Cami

Dr. Bagher Rahimpour Cami is an Assistant Professor in the Department of Computer Engineering at Mazandaran University of Science and Technology, Babol, Iran. He obtained his PhD in Artificial Intelligence from Shahrood University, Iran. His research interests include data mining and recommender systems.