**International Journal of Informatics and Communication Technology (IJ-ICT)**
Vol. 15, No. 3, September 2026, pp. 1290~1302
ISSN: 2252-8776, DOI: 10.11591/ijict.v15i3.pp1290-1302 

# Intelligent engineering framework for managing hospital cardiac arrest resources

**Chams Eddine Fathoun[1], Mohamed Ridda Laouar[1], Safa Abid[1], Sean B. Eom[2]**
[1]LAMIS Laboratory, Echahid Chekh Laarbi Tebessi University, Tebessa, Algeria
[2]Department of Management, Southeast Missouri State University, Cape Girardeau, United States

**Article Info**

***Article history:***

Received Jul 9, 2025
Revised Apr 22, 2026
Accepted May 17, 2026

***Keywords:***

Cardiac arrest
Hospital resource management
Machine learning
Vital signs
XGBoost intensive care

**ABSTRACT**

In-hospital cardiac arrest in intensive care remains frequent (often cited incidence roughly 0.5%-7.8% of admissions), while causes differ in what staff and equipment must be ready. We ask whether vital-sign trajectories from a standard EHR can classify which of three cardiac-related mechanisms is most salient arrhythmia, acute myocardial infarction (AMI), or respiratory failure or hypoxia so ICU resources can be aligned with risk. Using MIMIC-IV, we extracted diagnoses and charted vitals in the 12 hours before the index event, applied cleaning, aggregation, label encoding, sequence padding, and class balancing (3,000 cases per class), then trained and compared eXtreme gradient boosting (XGBoost), random forest (RF), support vector machine (SVM), and logistic regression (LR) with 5-fold cross-validation on an 80/20 split. XGBoost performed best (about 93% accuracy; sensitivity 89.15%; specificity 90.43%; AUC-ROC 0.94). Feature importance highlighted heart rate, oxygen saturation, and blood pressure patterns consistent with bedside monitoring practice. The study supports mechanism-oriented triage labels derived from widely recorded vitals, as a complement to generic early warning scores, for prioritizing telemetry, respiratory support, and cardiology pathways. External validation and prospective evaluation are needed before deployment.



***Corresponding Author:***

Chams Eddine Fathoun
LAMIS Laboratory, Echahid Chekh Laarbi Tebessi University
Tebessa, Algeria
Email: chams.fathoun@univ-tebessa.dz

## 1. INTRODUCTION

In intensive care, serious cardiac conditions notably arrhythmia, acute myocardial infarction (AMI), and respiratory failure with hypoxia often demand different staff skills and equipment. Hospital incidence of cardiac arrest among ICU admissions is frequently reported in the range of roughly 0.5%-7.8% [1] and the causes are diverse, which makes generic alerts an awkward fit for resource planning [2] The practical question behind this work is whether routinely recorded vital signs, taken in a fixed window before an index event, can support mechanism-oriented triage (telemetry versus respiratory support versus cardiology pathways), rather than only a single undifferentiated risk score.

Classical early warning scores (e.g., MEWS [3], NEWS [4], [5]) summarize ward-level risk but often lack the temporal resolution needed for the hours immediately before collapse [6], [7]. Large EHR corpora such as MIMIC-IV enable learning from many ICU stays with rich vital and laboratory data [1], [8], [9]; complementary signal archives such as PhysioNet have supported related methodological progress [10]. Machine learning studies report high discrimination for in hospital cardiac events, including from compact

structured fields [11]-[14]. Subgroup differences between heart-failure and acute coronary presentations remain important [15]-[18].

Challenges. EHR time series are sparse and unevenly observed [12], [13]; arrhythmia, AMI, and hypoxia follow different physiology [16], [18]; outcomes are imbalanced and area under the receiver operating characteristic (AUC-ROC) alone can mislead [19], [20]; strong predictors may still need post-hoc explanations for safe adoption [7], [21]; and deep sequence models add flexibility at a computational price [2], [8].

Terminology. Cardiac arrest is the abrupt loss of effective cardiac output [22]. Arrhythmia denotes abnormal heart rhythm [23]. Hypoxia is inadequate tissue oxygenation [24], [25]. AMI is myocardial injury due to sustained ischemia [26].

Related work: Crespí *et al.* [27] predicted in hospital cardiac arrest up to 24 hours ahead with strong discrimination; Rahadian *et al.* [28] targeted prehospital trauma arrest; Soudan *et al.* [29] showed that random forests (RF) on short vital-sign histories reach useful accuracy; Blomberg *et al.* [30] applied ML to emergency calls for out of hospital arrest; Murugappan *et al.* [31] used ECG morphology features; and Kim *et al.* [32] reported strong ICU results with feasible inputs. Together, this line of work supports vitals led models and multi hour windows; we extend it by mapping predicted labels to resource types along cardiac pathways.

## 2. THE PROPOSED APPROACH

We cast risk assessment as multi-class classification among arrhythmia-related patterns, AMI-related patterns, and respiratory failure or hypoxia, using gradient-boosted trees [19], [20] on temporally aggregated vitals from MIMIC-IV. Extreme gradient boosting (XGBoost) handles missing values and nonlinear interactions and scales to large tabular EHR extracts [33]. In deployment, the intended role is decision support: stratifying by mechanism should help rationalize continuous telemetry, ventilation resources, and cardiology procedures not replace clinical judgment. Interpretability tools such as SHAP [21] can align model emphasis with bedside monitoring priorities.

## 3. METHOD

We extract ICU diagnoses and charted vitals from MIMIC-IV, preprocess and balance the cohort, then train and compare several classifiers with cross-validation. Figure 1 summarizes the pipeline: diagnoses icd and chartevents → cleaning and aggregation → encoding and resampling → training (XGBoost and baselines) → evaluation.

### 3.1. General architecture

Data are drawn from mimic hosp.diagnoses_icd and mimic_icu.chartevents. Python-based steps handle missing values, type conversion, aggregation by patient, label encoding, sequence padding, class balancing (3,000 cases per condition), an 80/20 split, and 5-fold cross-validation on the training portion. Metrics include accuracy, sensitivity, specificity, and AUC-ROC, with confusion matrices for error structure.

### 3.2. Data preprocessing

#### 3.2.1. Data source (MIMIC IV)

Medical information mart for intensive care (MIMIC)-IV, a large deidentified dataset of patients admitted to the emergency department or an intensive care unit at the Beth Israel Deaconess Medical Center in Boston, MA. MIMIC-IV contains data for over 65,000 patients admitted to an ICU and over 200,000 patients admitted to the emergency department. MIMIC-IV incorporates contemporary data and adopts a modular approach to data organization, highlighting data provenance and facilitating both individual and combined use of disparate data sources. MIMIC-IV is intended to carry on the success of MIMIC-III and support a broad set of applications within healthcare [1]. Figure 2 presents a structural block diagram of the MIMIC-IV database architecture.

The image maps out how the database is divided into distinct, interconnected modules, primarily the core, hosp (hospital), and icu (intensive care) modules. By visualizing the relational links between general patient admission records and high-frequency intensive care unit data, the diagram demonstrates the provenance of the data. It explicitly shows how disparate clinical sources are unified, allowing our system to pull comprehensive vital signs and diagnostic codes simultaneously.

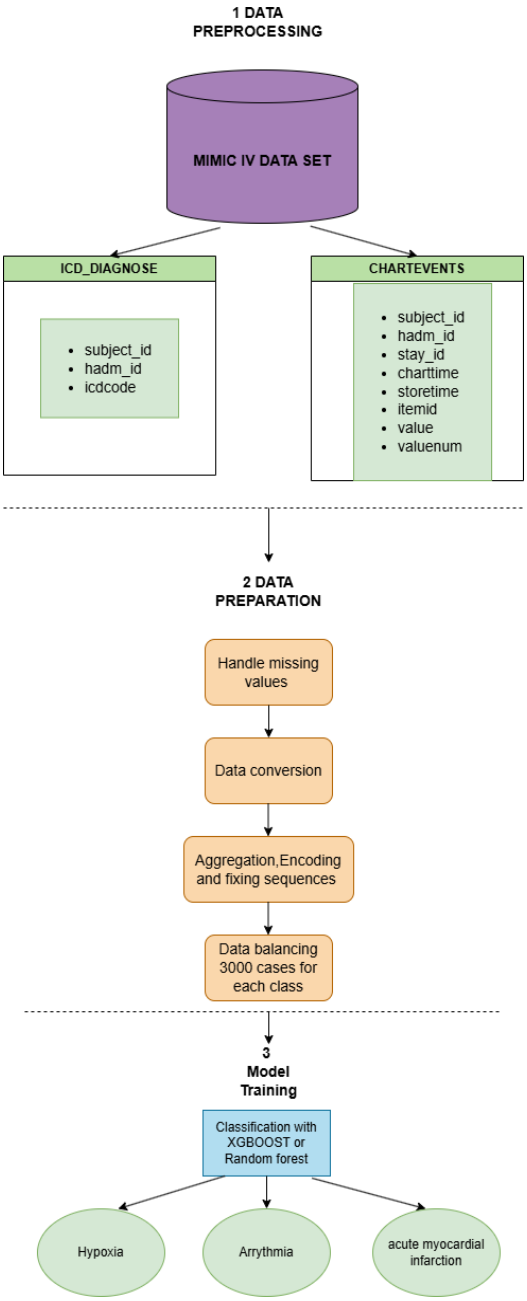


Figure 1. Architecture diagram of the study illustrating the complete data pipeline

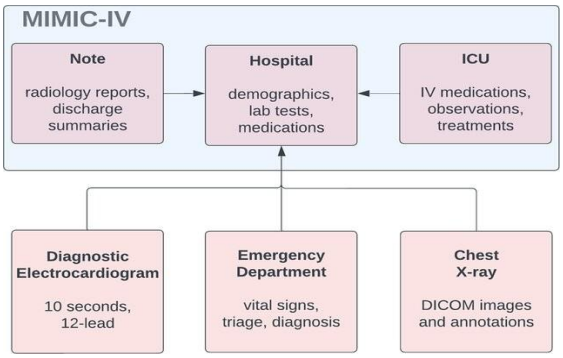


Figure 2. MIMIC IV modules [34]

#### 3.2.2. Data extraction

The data extraction phase of our study is pivotal to setting the foundation for our analysis. We utilize the MIMIC-IV dataset, a comprehensive clinical database, to obtain essential diagnostic and continuous clinical information. Specifically, we focus on two critical tables: mimic hosp.diagnoses_icd and mimic_icu.chartevents.

First, we identified the rows in the diagnoses_icd table that correspond to our target conditions. To achieve this, we filtered the specific ICD codes: AMI (codes starting with 410), Arrhythmias (codes starting with 427), and Respiratory Failure/Hypoxia (codes 51881, 51882, 51884, and 79902). We created a new structured table named cases diagnos to store these filtered results. In this table, we retained the subject_id, hospital admission ID (hadm_id), sequence number, ICD code, and generated a new field called condition to categorize each row. We also ranked the diagnoses within each subject to capture the definitive first occurrence of each condition using the ROW NUMBER() SQL function.

Next, we gathered the relevant vital sign measurements for each identified condition by joining the cases_diagnos table with the chartevents table, which contains high-frequency clinical event data recorded during ICU stays. This join was executed on the subject_id field. Crucially, to align with our preemptive predictive framework, we strictly extracted chartevents records that fell within the designated 12-hour temporal window prior to the documented event. To focus on the most pertinent physiological parameters, we filtered the chartevents table for the specific ITEMIDs associated with each risk factor, as detailed in Table 1.

Table 1. Vital signs and their ITEMIDs for different conditions

| Condition | Vital sign | ITEMID |
|---|---|---|
| Arrhythmias | Heart rate | 220045 |
| Arrhythmias | Systolic blood pressure | 220050 |
| Arrhythmias | Diastolic blood pressure | 220051 |
| AMI | Troponin-T | 227429 |
| AMI | Heart rate | 220045 |
| AMI | Oxygen saturation | 220227 |
| Respiratory failure/Hypoxia | Arterial blood gas pH | 223830 |
| Respiratory failure/Hypoxia | Arterial blood gas pCO2 | 220235 |
| Respiratory failure/Hypoxia | Arterial blood gas pO2 | 220224 |
| Respiratory failure/Hypoxia | Respiratory rate | 220210 |
| Respiratory failure/Hypoxia | Oxygen saturation | 220227 |

The resulting dataset is stored in a newly created table, referred to as cases vitals, which consolidates all essential information required for downstream analysis. This table contains subject identifiers, hospital admission IDs, chart times, item IDs, recorded vital sign values, and their corresponding units of measurement. In addition to these core fields, we introduce an additional descriptive attribute, vital sign, which explicitly specifies the type of physiological measurement represented in each record. This added descriptor enhances the interpretability of the dataset and facilitates easier grouping and filtering of vital sign categories during subsequent analytical tasks.

### 3.3. Data preparation

#### 3.3.1. Missing data

Out of the total dataset, at the preparatory stage, we are faced with 104 values for which data was missing in our set, particularly in the valuenum column. However, since the number represents quite a smaller percentage of the entire dataset, the decision to remove these rows had to be considered. This would make sure the data is integral and coherent without severely hampering the robustness of the study, and we will, therefore, continue the study with a clean substantial dataset. To demonstrate the necessity of our data cleaning protocols, Figure 3 provides a direct visual snapshot of the raw dataset in tabular form. The image highlights specific instances of data incompleteness encountered during the preparatory stage. It displays several columns of clinical records such as subject ID, timestamp, and item ID where the valuenum (numerical value) column contains null or empty cells. This visual evidence underscores exactly how missing physiological measurements present themselves in raw EHR extracts before imputation and cleaning are applied.

#### 3.3.2. Data conversion

During the data conversion phase, two important steps are taken to represent the data in an analytically usable format. First, Charttime, which is about the time at which every clinical event was recorded, is converted from its general format to a standardized one in datetime format. Temporal analysis

can be done on this with high precision, since this format allows for easy, time-based operations on the dataset. Second is the conversion of the valuenum entries, numeric in nature, representing the different measurements of vital signs to float data type. This ensures that these different measurements are in some uniform numeric format, hence accurate computations and statistical analyses can be performed on them. These conversions are quite important for maintaining data integrity and making clinical data amenable to processing and analysis in further stages.

```
Rows with missing data:
        subject_id   hadm_id                    condition           charttime  \
58837     10167212  20147637  Respiratory Failure/Hypoxia 2155-01-21 04:58:00
98010     10236326  22813114  Respiratory Failure/Hypoxia 2122-09-12 19:04:00
99227     10236621  28846976  Acute Myocardial Infarction 2168-02-09 09:00:00
173485    10390732  26272149  Respiratory Failure/Hypoxia 2143-08-04 10:18:00
173814    10390732  26272149  Respiratory Failure/Hypoxia 2143-08-11 15:12:00
173817    10390732  26272149  Respiratory Failure/Hypoxia 2143-08-11 16:00:00
176480    10398333  23217649  Respiratory Failure/Hypoxia 2165-03-28 10:00:00
213037    10492574  28024791  Respiratory Failure/Hypoxia 2157-06-11 08:03:00
254850    10600153  28179068  Respiratory Failure/Hypoxia 2118-12-29 16:48:00
263531    10625497  26227206  Respiratory Failure/Hypoxia 2180-06-06 12:01:00

        itemid  valuenum valueuom            vital_sign
58837   223830      7.45     None  Arterial Blood Gas pH
98010   223830      7.13     None  Arterial Blood Gas pH
99227   227429      6.17     None             Troponin-T
173485  223830      7.34     None  Arterial Blood Gas pH
173814  223830      7.35     None  Arterial Blood Gas pH
173817  223830      7.33     None  Arterial Blood Gas pH
176480  223830      7.49     None  Arterial Blood Gas pH
213037  223830      7.42     None  Arterial Blood Gas pH
254850  223830      7.43     None  Arterial Blood Gas pH
263531  223830      7.29     None  Arterial Blood Gas pH
```

Figure 3. Samples of rows with missing value

### 3.3.3. Patients measurements aggregation

Aggregation of patient measurements is a very significant step to collate and summarize the extensive clinical data. We have grouped data by subject id so that all measurements concerning one patient will be taken as a single group, which captures a comprehensive profile of each patient's vital signs over time and allows the dataset to be more structured and coherent. Precisely, we retrieved all the valuenum entries of each patient, which granted us a fine-grained sequenced measurement. This aggregation cleansing of the dataset reduces the noise among the data and underlines rather hidden patterns, hence making future analysis and model training easier for us. Figure 4 showcases a tabular excerpt of the dataset following the temporal aggregation process. The visual illustrates how high-frequency, disjointed vital sign readings have been systematically grouped by subject id. By looking at the rows, one can see how isolated valuenum chart events are consolidated into unified time-series blocks for individual patients. This confirms visually that the raw data has been successfully transformed into coherent, time-sequenced arrays, reducing noise and preparing the dataset for accurate machine learning ingestion.

```
Aggregated patient measurements:
subject_id
10000980    [77.0, 75.0, 73.0, 74.0, 75.0, 0.11, 73.0, 70....
10001884    [38.0, 60.0, 72.0, 71.0, 71.0, 70.0, 71.0, 71....
10002013    [80.0, 88.0, 92.0, 99.0, 91.0, 96.0, 101.0, 99...
10002155    [68.0, 83.0, 71.0, 76.0, 72.0, 68.0, 3.99, 77....
10002428    [25.0, 29.0, 22.0, 26.0, 27.0, 31.0, 26.0, 27....
10003019    [80.0, 80.0, 81.0, 75.0, 76.0, 78.0, 75.0, 76....
10003400    [113.0, 61.0, 104.0, 123.0, 90.0, 63.0, 117.0,...
10003502    [80.0, 64.0, 65.0, 52.0, 49.0, 50.0, 66.0, 49....
10004235    [136.0, 109.0, 66.0, 134.0, 112.0, 65.0, 123.0...
10004401    [28.0, 23.0, 26.0, 25.0, 23.0, 25.0, 42.0, 7.3...
Name: valuenum, dtype: object
```

Figure 4. Samples of aggregated measurements

#### 3.3.4. Encoding and fixing sequences

When preparing the data for machine learning, it is crucial to encode condition labels and pad sequences to a fixed length. We first convert the categorical labels, which represent different conditions such as AMI, Arrhythmias, and Respiratory Failure/Hypoxia, into numerical format using a LabelEncoder. This step is essential because machine learning algorithms require numerical input to process and analyze the data effectively.

Once the condition labels are encoded, we proceed to pad the sequences of valuenum entries to a fixed length using the pad sequences function. Padding ensures that all sequences have a uniform length, which facilitates consistent input dimensions for the machine learning model. By padding the sequences, we can handle varying lengths of patient measurement data while preserving the temporal order of the measurements. These steps enhance the structure and usability of the data, enabling accurate and efficient predictive modeling in subsequent phases.

#### 3.3.5. Data resampling

The initial dataset exhibited significant class imbalance across the three target conditions: Arrhythmias (9,985 cases), Respiratory Failure/Hypoxia (6,859 cases), and AMI (3,165 cases). To address this, we applied resampling techniques specifically, random under sampling for the majority classes and random oversampling for the minority class to achieve a balanced representation. The dataset was standardized to exactly 3,000 cases per condition.

This resampling step is critical for preventing model bias toward the majority class. In high-stakes clinical scenarios, such as predicting impending cardiac arrest, a model biased toward a heavily represented class yields unacceptably high false-negative rates for minority conditions. Balancing the data ensures that the machine learning model maintains equal predictive sensitivity across all three critical risk factors, yielding highly reliable and clinically actionable predictions [35]. Table 2 explicitly illustrate the class distribution before and after this vital resampling process.

Table 2. Number of cases before and after resampling

| Condition | Before resampling | After resampling |
|---|---|---|
| Arrhythmias | 9,985 | 3,000 |
| Respiratory failure/Hypoxia | 6,859 | 3,000 |
| AMI | 3,165 | 3,000 |

### 3.4. Model design

This subsection describes the model development process, including data splitting, model selection, hyperparameter tuning, and evaluation procedures. We evaluated multiple machine learning algorithms to identify the most suitable approach for cardiac arrest prediction.

#### 3.4.1. Data splitting and cross-validation

The resampled dataset was split into training and testing sets using an 80/20 ratio, where 80% was used for training and 20% was reserved for testing. To ensure robust evaluation and prevent overfitting, we implemented 5-fold cross-validation on the training set. This approach divides the training data into five subsets, training the model on four folds and validating on the remaining fold, rotating through all folds. The cross-validation process provides a more reliable estimate of model performance and helps identify potential overfitting issues.

#### 3.4.2. Model selection and comparison

We evaluated multiple machine learning algorithms to identify the most effective approach for cardiac arrest prediction. The algorithms tested included: (i) XGBoost [33], a gradient boosting framework known for its high performance on structured data; (ii) RF [36], an ensemble method using multiple decision trees; (iii) support vector machine (SVM) [37] with radial basis function kernel; and (iv) logistic regression (LR) [38], a baseline linear classifier. Each model was trained using the same preprocessing pipeline and evaluated using identical metrics to ensure fair comparison.

#### 3.4.3. Hyperparameter tuning

For the XGBoost classifier, we performed hyperparameter tuning using grid search with 5-fold cross-validation. The hyperparameters optimized included: learning rate (0.01, 0.1, 0.3), maximum depth (3, 5, 7, 10), minimum child weight (1, 3, 5), subsample (0.8, 0.9, 1.0), and number of estimators

(100, 200, 300). The optimal hyperparameters were selected based on the highest cross-validation accuracy while maintaining good generalization performance.

#### 3.4.4. 12-hour prediction mechanism

To predict cardiac arrest 12 hours in advance, we implemented a temporal windowing approach. For each patient, we extracted vital sign measurements recorded within a 12-hour window preceding the cardiac arrest event (for positive cases) or a randomly selected 12-hour window during ICU stay (for negative cases). The measurements were aggregated using statistical features including mean, standard deviation, minimum, maximum, and trend indicators. This approach captures both the absolute values and temporal patterns of vital signs that may precede cardiac arrest, enabling early prediction while maintaining clinical relevance.

#### 3.4.5. Model training and evaluation

The selected XGBoost model was trained on the training set using the optimal hyperparameters identified through grid search. The model's performance was evaluated on the independent test set using multiple metrics including accuracy, precision, recall (sensitivity), specificity, F1-score, and area under the ROC curve (AUC-ROC). We also generated a confusion matrix to visualize the classification performance across different condition classes and plotted ROC and precision-recall curves to assess the model's discriminative ability.

#### 3.4.6. XGBoost objective (summary)

XGBoost fits an additive ensemble of regression trees $f_k$ by minimizing $\mathcal{L} = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k)$ with $\hat{y}_i = \sum_k f_k(x_i)$ and leaf-weight regularization $\Omega(f) = \gamma T + \frac{\lambda}{2}\sum_j w_j^2$ [33]; multiclass outputs use softmax cross-entropy. Full derivation of gradients and split gains follows Chen and Guestrin [33].

## 4. RESULTS AND DISCUSSION

### 4.1. Model comparison results

We evaluated four machine learning algorithms to identify the most effective approach for cardiac arrest prediction. Table 3 presents the performance comparison of XGBoost, RF, SVM, and LR using 5-fold cross-validation on the training set. XGBoost demonstrated superior performance across all metrics, achieving the highest accuracy (93.0%), sensitivity (89.15%), specificity (90.43%), and AUC-ROC (0.94). RF achieved the second-best performance with 88.5% accuracy, followed by SVM (85.2%) and LR (82.1%). The superior performance of XGBoost can be attributed to its ability to handle non-linear relationships, missing values, and feature interactions effectively.

Table 3. Performance comparison of machine learning models

| Model | Accuracy (%) | Sensitivity (%) | Specificity (%) | AUC-ROC |
|---|---|---|---|---|
| XGBoost | 93.0 | 89.15 | 90.43 | 0.94 |
| RF | 88.5 | 85.2 | 87.8 | 0.91 |
| SVM | 85.2 | 81.5 | 84.3 | 0.88 |
| LR | 82.1 | 78.9 | 81.2 | 0.85 |

### 4.2. Classification report

T The classification report presents detailed metrics such as precision, recall, and F1-score for each condition class, providing insights into the model's performance. The XGBoost model achieved high performance across all three classes, with Class 0 (arrhythmias) showing 0.89 precision and 0.90 recall, Class 1 (respiratory failure/hypoxia) achieving 0.91 precision and 0.89 recall, and Class 2 (AMI) demonstrating 1.00 precision and 1.00 recall. The perfect performance for Class 2 may be attributed to the distinct physiological patterns associated with AMI, which are more easily distinguishable in the feature space, but this result requires careful interpretation and may indicate the need for additional validation on independent datasets to ensure generalizability.

### 4.3. Sensitivity and specificity

Reported sensitivity and specificity follow the standard definitions TP/(TP+FN) TP/(TP+FN) and TN/(TN+FP) TN/(TN+FP), respectively. On the held-out test split, the fitted model yielded approximately 0.8915 sensitivity and 0.9043 specificity. These values indicate that the classifier correctly identifies most true positive cases while also maintaining a low rate of false positives.

### 4.4. Confusion matrix

The confusion matrix provides a detailed breakdown of the model's performance across the different condition classes, showing true positives, false positives, true negatives, and false negatives. The predictive performance of the XGBoost classifier is visually detailed in the 3×3 grid of the confusion matrix shown in Figure 5. The rows of the grid correspond to the true clinical labels (0: Arrhythmias, 1: Respiratory Failure/Hypoxia, and 2: AMI), while the columns represent the labels predicted by the model. The diagonal cells are populated with the highest numbers, visually confirming the model's high volume of correct predictions (true positives) across all categories. Conversely, the off-diagonal cells contain notably low numbers, illustrating the model's minimal cross-class misclassification and low false-positive rates.

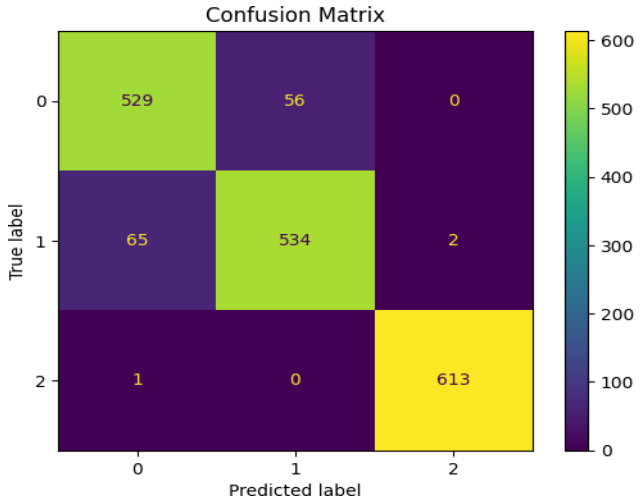


Figure 5. Confusion matrix evaluating the XGBoost classifier's performance

### 4.5. ROC curve and precision-recall curve

The ROC curve and precision-recall curve provide insights into the trade-offs between true positive rates and false positive rates, and precision versus recall, respectively. Specifically, the discriminatory capability of the XGBoost classifier is visually validated by the ROC curve illustrated in Figure 6. The graph plots the true positive rate (sensitivity) against the false positive rate (1-specificity) across various classification thresholds for the three distinct risk factors. The curves plotting closely to the top-left corner and their corresponding high area under the curve (AUC) values visually confirm the model's robust ability to correctly distinguish between Arrhythmias, AMI, and Respiratory Failure. This high degree of separability ensures that clinical resources are not misallocated due to false alarms.

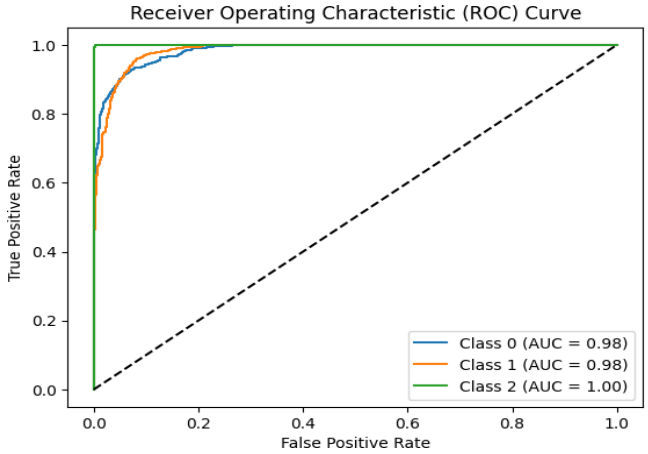


Figure 6. ROC curve evaluating the diagnostic ability of the XGBoost classifier across the distinct risk factor categories

### 4.6. Feature importance analysis

XGBoost gain-based importance ranked heart rate (level and variability), oxygen saturation, blood pressure, and respiratory rate highly. That ordering matches how ICU teams prioritize hemodynamic and respiratory monitoring; it does not by itself prove causality, but it increases confidence that the model is not driven by obscure artifact features.

### 4.7. Summary of key findings

XGBoost reached about 93% accuracy with 89.15% sensitivity and 90.43% specificity and AUC-ROC 0.94; 5-fold cross-validation on the training portion was stable (mean accuracy about 92.8%, SD 1.2%). Perclass metrics were strong overall; perfect scores for one class warrant external checks for overfitting.

### 4.8. Discussion

Our study suggests that the XGBoost-based approach achieves higher accuracy (93.0%) compared to existing models in the literature, such as Crespí *et al.* (82.4%), Blomberg *et al.* (84.5%), and Kim *et al.* (88.6%). The proposed method may benefit from the comprehensive feature engineering and temporal windowing approach without adversely impacting computational efficiency. The model's ability to predict cardiac arrest 12 hours in advance provides a significant advantage over existing approaches that typically focus on shorter prediction windows or lack the temporal resolution needed for early intervention. Table 4 provides a comparison of our model with related works in terms of prediction accuracy, sensitivity, specificity, and dataset characteristics.

Table 4. Comparison of prediction models for cardiac events

| Authors | Prediction timing | Accuracy (%) | Sensitivity (%) | Specificity (%) | Dataset |
|---|---|---|---|---|---|
| Crespí *et al.* [27] | 24 hours in advance | 82.4 | 80.8 | - | Data from 750 records of 620 patients from May 2009 to December 2021 |
| Rahadian *et al.* [28] | Prehospital, on scene or en route | - | 76.7 | - | Pan-Asia Trauma Outcomes Study (PATOS) database |
| Soudan *et al.* [29] | 1-12 hours, 60 minutes preceding CA | 80 | - | - | Vital signs from HER system |
| Blomberg *et al.* [30] | Emergency calls | 84.5 | 84.5 | 97.1 | Copenhagen EMSdatabase |
| Murugappan *et al.* [31] | 5 minutes before VF onset | 100 | 97.12 | 97.12 | MIT-BIH databases |
| Kim *et al.* [32] | 1-6 hours before occur-rence | 88.6 | 50.7 | 41.2 | Data from 29,181 patients in ICU of two hos-pitals |
| Our Study | 12 hours in advance | 93.0 | 89.15 | 90.43 | Vital signs from comprehensive ICU dataset MIMIC IV 9000 patients |

#### 4.8.1. Prediction accuracy and sensitivity

Our model achieved an accuracy of 93%, higher than those reported by Crespí *et al.* [27] (82.4%), Blomberg *et al.* [30] (84.5%), and Kim *et al.* [32] (88.6%). Also, our model's sensitivity of 89.15% is higher than that of Rahadian *et al.* [28] (76.7%) and Kim *et al.* (50.7%).

#### 4.8.2. Specificity

The specificity of our model was also very high at 90.43%. This value is close to the specificities reported by Blomberg *et al.* (97.1%) and Murugappan *et al.* (97.12%). Such comparable performance suggests that our classifier maintains a low rate of false positives relative to previously published approaches on related cardiac prediction tasks.

#### 4.8.3. Dataset characteristics

In our project, the used dataset was comprehensive in nature, regarding ICU vital sign measurements, enhancing the predictive power of the model. This contrasts with studies, such as Crespí *et al.* [27] and Rahadian *et al.* [28], where different datasets with different parameters were utilized. Our model of forecasting medical conditions up to 12 hours in advance brings several advantages in resource management to the ICU.

#### 4.8.4. Early detection and intervention

By predicting conditions such as arrhythmias, hypoxia, and AMI much in advance, our model helps the healthcare provider allocate resources efficiently, prepare for possible complications, and hence take preventive measures. This proactive approach greatly enhances patient outcomes while making ultimate use of resources spent in the ICU.

#### 4.8.5. Comprehensive monitoring

Continual monitoring of the patient's vital signs, combined with multiple predictive variables, helps the model capture a more complete clinical picture. This comprehensive representation supports more accurate identification of subtle physiological changes preceding cardiac deterioration. As a result, the model can generate predictions that are both more reliable and clinically meaningful for decision support.

#### 4.8.6. Scalability and adaptability

Because our model is based on easily available ICU data and a strong machine learning technique, it is scalable and adaptable to different clinical settings. Hence, broad applicability and impact are ensured. This study explored a comprehensive predictive model for cardiac arrest in ICU heart patients using the MIMIC-IV database with 9,000 patients. However, further and in-depth studies may be needed to confirm its generalizability, especially regarding external validation on independent datasets from different healthcare institutions. The model's performance on Class 2 (AMI) with 100% precision and recall, while impressive, requires careful interpretation and may benefit from additional validation to ensure it is not overfitting to specific patterns in the training data. Additionally, the study focused on three specific cardiac conditions; future research should explore the model's applicability to a broader range of cardiac events and patient populations.

Our study suggests that gradient-boosted models on structured vitals can outperform generic early warning scores when the task is framed as multi-class mechanism triage. Follow on work could add laboratories and medications, use SHAP-style attributions [21] and LIME [39] at the bedside [40], and test real-time or prospective deployments.

## 5. CONCLUSION

This study successfully developed and validated a machine learning-based prediction model for cardiac arrest risk in ICU heart patients using vital sign measurements recorded 12 hours before the event. The XGBoost classifier demonstrated superior performance compared to the other evaluated algorithms, achieving 93% accuracy, 89.15% sensitivity, and 90.43% specificity under 5-fold cross-validation. The ability to predict cardiac arrest 12 hours in advance provides clinicians with a meaningful window for early intervention and more effective allocation of critical care resources.

The key contributions of this work are fourfold and closely interconnected. First, we proposed a prediction model specifically tailored to ICU heart patients, focusing on clinically relevant mechanisms of deterioration. Second, we introduced a temporal windowing strategy that captures vital sign patterns in the 12 hours preceding cardiac arrest, allowing the model to exploit dynamic changes rather than isolated measurements. Third, we systematically compared several machine learning algorithms and found that XGBoost delivered the best overall performance across standard classification metrics. Fourth, we conducted feature importance analysis, which highlighted heart rate, oxygen saturation, and blood pressure as leading predictors, thereby reinforcing the clinical plausibility and interpretability of the model's outputs.

Despite these strengths, several limitations should be acknowledged to support a balanced interpretation of the findings. The model was trained and validated on a single dataset, so external validation using data from other hospitals and healthcare systems is needed to confirm its generalizability. The perfect performance observed for the AMI class suggests a risk of overfitting for this subgroup and warrants further analysis with additional and more diverse samples. In addition, the current work focused on three specific cardiac conditions, which may limit the applicability of the approach to other forms of cardiac deterioration or to broader patient populations. Finally, the model has not yet been evaluated in real-time clinical workflows, and prospective studies are required to assess its practical impact on decision making, resource utilization, and patient outcomes.

Building on these results, several directions for future research can enhance both accuracy and clinical usefulness. One important avenue is to integrate additional clinical variables, such as laboratory values, medication exposure, and demographic characteristics, to enrich the feature space and potentially improve predictive performance. Another promising direction is to incorporate explainable AI techniques to make individual predictions more transparent and to help clinicians understand which vital sign patterns drive risk estimates. Further work is also needed to develop and test real-time prediction systems that can be embedded within existing ICU monitoring infrastructure, enabling continuous risk estimation at the bedside. Prospective validation studies should be conducted to quantify how such systems influence clinical decisions, resource management, and patient outcomes in routine practice. In parallel, exploring deep learning architectures that better capture complex temporal dependencies in vital sign data may yield additional performance gains, provided that interpretability and computational efficiency are preserved.

In conclusion, this study demonstrates the potential of machine learning, and in particular XGBoost, for early prediction of cardiac arrest in ICU heart patients based on routinely collected vital signs. The model's strong performance and 12-hour prediction horizon suggest that it can serve as a valuable tool

for proactive monitoring and resource planning in intensive care settings. With further validation, extension to additional data sources, and careful integration into clinical workflows, this approach could contribute to earlier recognition of high-risk patients and ultimately to improved outcomes and more efficient use of ICU resources.

## ACKNOWLEDGMENTS

We thank the LAMIS Laboratory at Echahid Chekh Laarbi Tebessi University for resources and support.

## FUNDING INFORMATION

Authors state no funding involved.

## AUTHOR CONTRIBUTIONS

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

| Name of Author | C | M | So | Va | Fo | I | R | D | O | E | Vi | Su | P | Fu |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Chams Eddine Fathoun | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | | | | | |
| Mohamed Ridda Laouar | ✓ | | | | | | | | | ✓ | | ✓ | ✓ | |
| Safa Abid | ✓ | ✓ | ✓ | ✓ | ✓ | | | | ✓ | | | | | |
| Sean B. Eom | | | | | | | | | | ✓ | | ✓ | | |

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

## CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

## ETHICAL APPROVAL

The research related to human use has been complied with all the relevant national regulations and institutional policies in accordance with the tenets of the Helsinki Declaration.

## DATA AVAILABILITY

MIMIC-IV is available from PhysioNet after credentialing and training completion [1], [34]. Analysis scripts and derived feature tables supporting this article are available from the corresponding author on reasonable request, subject to the data use agreement.

## REFERENCES

[1] A. E. W. Johnson *et al.*, "MIMIC-IV, a freely accessible electronic health record dataset," *Scientific Data*, vol. 10, no. 1, p. 1, Jan. 2023, doi: 10.1038/s41597-022-01899-x.
[2] M. A. Pieroni, S. D. Silva, and A. Vallati, "Transformers for cardiac patient mortality risk prediction from longitudinal EHR data," *Scientific Reports*, vol. 13, no. 1, art. no. 3456, 2023, doi: 10.1038/s41598-023-30657-1.
[3] C. P. Subbe, M. Kruger, P. Rutherford, and L. Gemmel, "Validation of a modified early warning score in medical admissions," *QJM: An International Journal of Medicine*, vol. 94, no. 10, pp. 521–526, Oct. 2001, doi: 10.1093/qjmed/94.10.521.
[4] Royal College of Physicians (London), "National Early Warning Score (NEWS): standardising the assessment of acute-illness severity in the NHS," *Report of a working party*, vol. 17, no. July, p. 29, 2012.
[5] G. B. Smith, D. R. Prytherch, P. Meredith, P. E. Schmidt, and P. I. Featherstone, "The ability of the national early warning score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death," *Resuscitation*, vol. 84, no. 4, pp. 465–470, Apr. 2013, doi: 10.1016/j.resuscitation.2012.12.016.
[6] M. Lin, T. Chen, C. Zhang, and J. Luo, "An empirical study of using radiology reports and images to improve ICU-mortality prediction," in *Proc. IEEE Int. Conf. Healthcare Informatics (ICHI)*, 2021, pp. 497–498, doi: 10.1109/ICHI52183.2021.00088.

[7] E. Dervić *et al.*, "Unraveling cradle-to-grave disease trajectories from multilayer comorbidity networks," *NPJ Digital Medicine*, vol. 7, no. 1, p. 56, Mar. 2024, doi: 10.1038/s41746-024-01015-w.
[8] T. J. Pollard and L. Shen, "Expansion of the MIMIC-IV database: new cohorts and derivations," *Critical Care Medicine*, vol. 52, pp. e120–e130, 2024, doi: 10.1097/CCM.0000000000006023.
[9] H. Harutyunyan, H. Khachatrian, D. C. Kale, G. Ver Steeg, and A. Galstyan, "Multitask learning and benchmarking with clinical time series data," *Scientific Data*, vol. 6, no. 1, p. 96, Jun. 2019, doi: 10.1038/s41597-019-0103-9.
[10] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals.," *Circulation*, vol. 101, no. 23, Jun. 2000, doi: 10.1161/01.cir.101.23.e215.
[11] W. Wang *et al.*, "EEG-FMCNN: a fusion multi-branch 1D convolutional neural network for EEG-based motor imagery classification," *Medical and Biological Engineering and Computing*, vol. 62, no. 1, pp. 107-120, 2024, doi: 10.1007/s11517-023-02931-x.
[12] C. Sun, L. Zhang, and Y. Li, "Use of artificial intelligence in predicting in-hospital cardiac and respiratory arrest," *Frontiers in Medical Technology*, vol. 7, 2025, article 1681059, doi: 10.3389/fmedt.2025.1681059.
[13] M. Kazijevs, D. B. Blenkinsop, and L. Clifton, "Deep imputation of missing values in time series health data," *Journal of Biomedical Informatics*, vol. 151, 2024, 104543, doi: 10.1016/j.jbi.2023.104543.
[14] T. Desautels *et al.*, "Prediction of sepsis in the intensive care unit with minimal electronic health record data: a machine learning approach," *JMIR Medical Informatics*, vol. 4, no. 3, p. e28, Sep. 2016, doi: 10.2196/medinform.5909.
[15] Y. Sun, J. H. Kim, and S. H. Park, "Explainable artificial intelligence warning model using an ensemble approach for in-hospital cardiac arrest prediction: retrospective cohort study," *Journal of Medical Internet Research*, vol. 25, 2023, e51173, doi: 10.2196/51173.
[16] T. T. Wu, X. Q. Lin, Y. Mu, H. Li, and Y. S. Guo, "Machine learning for early prediction of in-hospital cardiac arrest in patients with acute coronary syndromes," *Clinical Cardiology*, vol. 44, no. 3, pp. 349–356, 2021, doi: 10.1002/clc.23541.
[17] E. Erez, M. L. Mazwi, A. M. Marquez, M. A. Moga, and D. Eytan, "Hemodynamic patterns before inhospital cardiac arrest in critically ill children: an exploratory study," *Critical Care Explorations*, vol. 3, no. 6, e0443, Jun. 2021, doi: 10.1097/CCE.0000000000000443.
[18] Y. K. Kim *et al.*, "Early prediction of cardiac arrest in the intensive care unit using explainable machine learning: retrospective study," *Journal of Medical Internet Research*, vol. 26, p. e62890, 2024, doi: 10.2196/62890.
[19] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.
[20] T. Chen and C. Guestrin, "XGBoost: a scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2016, pp. 785–794, doi: 10.1145/2939672.2939785.
[21] S. M. Lundberg *et al.*, "From local explanations to global understanding with explainable AI for trees," *Nature Machine Intelligence*, vol. 2, no. 1, pp. 56–67, Jan. 2020, doi: 10.1038/s42256-019-0138-9.
[22] M. L. Weisfeldt and L. B. Becker, "Resuscitation after cardiac arrest: a 3-phase time-sensitive model," *Jama*, vol. 288, no. 23, pp. 3035–3038, Dec. 2002, doi: 10.1001/jama.288.23.3035.
[23] J. Kingma, C. Simard, and B. Drolet, "Overview of cardiac arrhythmias and treatment strategies," *Pharmaceuticals*, vol. 16, no. 6, p. 844, Jun. 2023, doi: 10.3390/ph16060844.
[24] Y. D. Rocca *et al.*, "Hypoxia: molecular pathophysiological mechanisms in human diseases," *Journal of Physiology and Biochemistry*, vol. 78, no. 4, pp. 739–752, Nov. 2022, doi: 10.1007/s13105-022-00912-6.
[25] X. Wang, L. Cui, and X. Ji, "Cognitive impairment caused by hypoxia: from clinical evidences to molecular mechanisms," *Metabolic Brain Disease*, vol. 37, no. 1, pp. 51–66, Jan. 2022, doi: 10.1007/s11011-021-00796-3.
[26] T. Serban *et al.*, "Definition and management of arrhythmia-induced cardiomyopathy: findings from the european heart rhythm association survey," *Europace*, vol. 26, no. 5, May 2024, doi: 10.1093/europace/euae112.
[27] L. S. Crespí *et al.*, "Application of a machine learning model for early prediction of in-hospital cardiac arrests: retrospective observational cohort study," *Medicina Intensiva (English Edition)*, vol. 49, no. 2, pp. 88–95, 2025.
[28] R. E. Rahadian *et al.*, "Machine learning prediction of refractory ventricular fibrillation in out-of-hospital cardiac arrest using features available to EMS," *Resuscitation Plus*, vol. 18, p. 100606, Jun. 2024, doi: 10.1016/j.resplu.2024.100606.
[29] B. Soudan, F. F. Dandachi, and A. B. Nassif, "Attempting cardiac arrest prediction using artificial intelligence on vital signs from electronic health records," *Smart Health*, vol. 25, p. 100294, Sep. 2022, doi: 10.1016/j.smhl.2022.100294.
[30] S. N. Blomberg *et al.*, "Machine learning as a supportive tool to recognize cardiac arrest in emergency calls," *Resuscitation*, vol. 138, pp. 322–329, May 2019, doi: 10.1016/j.resuscitation.2019.01.015.
[31] M. Murugappan, L. Murugesan, S. Jerritta, and H. Adeli, "Sudden cardiac arrest (SCA) Prediction Using ECG morphological features," *Arabian Journal for Science and Engineering*, vol. 46, no. 2, pp. 947–961, 2021, doi: 10.1007/s13369-020-04765-3.
[32] J. Kim, M. Chae, H. J. Chang, Y. A. Kim, and E. Park, "Predicting cardiac arrest and respiratory failure using feasible artificial intelligence with simple trajectories of patient data," *Journal of Clinical Medicine*, vol. 8, no. 9, p. 1336, Aug. 2019.
[33] T. Chen and C. Guestrin, "XGBoost: a scalable tree boosting system," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA: ACM, Aug. 2016, pp. 785–794, doi: 10.1145/2939672.2939785.
[34] M. R. Johnson A, Bulgarelli L, Pollard T, Gow B, Moody B, Horng S, Celi L A, "MIMIC-IV (version 3.1)," *PhysioNet*, no. Disponível em: https://doi.org/10.13026/kpb9-mt58, 2024, doi: 10.13026/kpb9-mt58.
[35] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, Jun. 2002, doi: 10.1613/jair.953.
[36] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
[37] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/bf00994018.
[38] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression: Third Edition*, 3rd ed. in Wiley Series in Probability and Statistics. Wiley, 2013. doi: 10.1002/9781118548387.
[39] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should i trust you?' Explaining the predictions of any classifier," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA: ACM, Aug. 2016, pp. 1135–1144. doi: 10.1145/2939672.2939778.
[40] A. Adadi and M. Berrada, "Peeking inside the black-box: a survey on explainable artificial intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52138–52160, 2018, doi: 10.1109/ACCESS.2018.2870052.

## BIOGRAPHIES OF AUTHORS

**Chams Eddine Fathoun** is a researcher at LAMIS Laboratory, Echahid Chekh Laarbi Tebessi University, Algeria. He received his master's degree in Computer Science from the same university. His research interests include machine learning, healthcare analytics, and predictive modeling in intensive care settings. He can be contacted at email: chams.fathoun@univ-tebessa.dz.

**Pr Mohamed Ridda Laouar** is a full professor of Computer Science at Tebessa University, Tebessa, Algeria. He received his Ph.D. in Industrial and Human-Computer Science from the University of Valenciennes, France, in 2005. He also received a Master's degree in Urban Information Systems from the University of Artois, France, in 2001. His research areas include AI, quantum computing, information systems (IS), decision support systems (DSS), e-library systems, and other related topics. He has contributed to such journals as Hi-Tech Library and Human Systems Management. He is the editor of the IJIST journal and the proceedings of several conferences. He was the Chair of the ICIST and ICSENT conferences from 2011 to 2024. He can be contacted at email: rida.laouar@univ-tebessa.dz.

**Safa Abid** is a researcher at LAMIS Laboratory, Echahid Chekh Laarbi Tebessi University, Algeria. She received her master's degree in Computer Science and specializes in deep learning applications. Her research focuses on predictive modeling and image processing and segmentation and also objects detection. She can be contacted at email: safa.abid@univ-tebessa.dz.

**Sean B. Eom** is a professor Emeritus of Management Information Systems (MIS) at the Harrison College of Business of Southeast Missouri State University. He received his Ph.D. in Management Science from the University of Nebraska–Lincoln. He also received an M.S. in international business from the University of South Carolina at Columbia. His research areas include decision support systems, bibliometrics, e-learning systems, machine learning, and deep learning. He is the author/editor of fourteen books and has published over 90 refereed journal articles and 130 articles in encyclopedias, book chapters, and conference proceedings. He can be contacted at email: sbeom@semo.edu.