

Article

Advanced Machine Learning Techniques for Predicting Inpatient Deterioration in General Medicine

Said Al Jaadi ¹, Laila Al Wahaibi ¹, Mohammed Al-Hinai ², Haneen Hafiz Gaffar ²  and Abdullah M. Al Alawi ^{1,3,4,*} 

¹ Internal Medicine Residency Program, Oman Medical Specialty Board, Muscat 123, Oman; r23070@resident.omsb.org (S.A.J.); r23086@resident.omsb.org (L.A.W.)

² College of Medicine and Health Sciences, Sultan Qaboos University, Muscat 123, Oman; s142659@student.squ.edu.om (M.A.-H.); s136314@student.squ.edu.om (H.H.G.)

³ Medical Research Center, Sultan Qaboos University, Muscat 123, Oman

⁴ Department of Medicine, Sultan Qaboos University Hospital, Muscat 123, Oman

* Correspondence: dr.abdullahalalawi@gmail.com; Tel.: +968-95384990

Abstract

Inpatient deterioration, marked by ICU transfer or mortality, remains a critical challenge in hospital settings. While traditional early warning systems (EWS) have limitations, machine learning (ML) offers a promising approach for the early identification of at-risk patients. This study aimed to develop and validate multiple ML models for predicting inpatient deterioration among general medical patients using electronic health record (EHR) data. A retrospective cohort study was conducted on 524 patients admitted between January 2022 and December 2023. The dataset included demographic, clinical, and laboratory variables, with time-stamped measurements treated as distinct features. After excluding variables with >15% missing data, standard imputation was performed. The training data was balanced using the Synthetic Minority Over-sampling Technique (SMOTE), and feature selection was performed using SelectKBest. A range of models—including Logistic Regression, Random Forest, Gradient Boosting, Support Vector Machines (SVMs), and Neural Networks—were trained and evaluated using AUC, accuracy, precision, recall, and F1-score. During 5-fold cross-validation, the models demonstrated high stability, with the Random Forest achieving a mean AUC of 0.980. On the final independent test set, the optimized Random Forest model yielded the highest performance with an AUC of 0.837 and an accuracy of 85.4%. Functional status, oxygen requirements, and urea levels were identified as key predictors. ML models, particularly Random Forest, can significantly enhance the early detection of inpatient deterioration. The contribution of this work lies in its systematic comparison of multiple algorithms and its robust methodology. Future research should focus on external validation, the integration of temporal data using recurrent neural network architectures, and the application of Explainable AI (XAI) to foster clinical trust and facilitate implementation.



Academic Editor: Hersh Sagreiya Sagreiya

Received: 18 December 2025

Revised: 10 February 2026

Accepted: 10 February 2026

Published: 12 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: machine learning; inpatient deterioration; early warning systems; predictive modeling; ICU transfer; general medicine

1. Introduction

The General Internal Medicine (GIM) service frequently manages a patient population characterized by advanced age and multiple comorbidities, rendering them particularly susceptible to adverse outcomes. Even subtle fluctuations in clinical status may precipitate

significant increases in morbidity and mortality. Consequently, patients admitted to general medical wards demonstrate substantially higher mortality rates compared to those admitted to surgical wards, with reported rates of 31.6% and 10.3%, respectively [1]. Timely escalation of care for patients exhibiting clinical deterioration in inpatient settings remains a persistent challenge for clinicians and healthcare systems. Deterioration following admission to a general ward is associated with a markedly increased risk of mortality. The early phase of clinical decline is particularly critical, as prompt transfer to the intensive care unit (ICU) during this window may significantly improve survival outcomes, especially among patients whose condition worsens after hospital admission [2].

Factors contributing to inpatient deterioration vary based on patient demographics and clinical conditions. These factors can include patient-specific characteristics, such as age, comorbidities, and primary diagnoses, as well as clinical indicators like vital signs, laboratory parameters, and mental status. Health system-related factors, including staffing levels, communication efficiency, and bed availability, also significantly impact outcomes. Furthermore, aspects related to treatment, such as response to therapy and occurrence of postoperative complications, can influence patient deterioration [3]. It has been estimated that each additional year of life increases the likelihood of death by 1.4%, and the average age of deceased patients is approximately nine years older compared to those who survive. Notably, one-quarter of patients who died were aged 85 years or older [4,5]. Other recognized risk factors for in-hospital deaths include nosocomial infections, with the length of stay before the onset of an infection being statistically significant [6,7]. malignancy [8]. male sex, mode of admission (emergency vs. elective) [9]. and need for inter-hospital transfer. Additionally, the number of previous emergency admissions or presentations [10]. functional and cognitive impairment [11]. disease severity, and conditions like cerebrovascular diseases, diabetes, chronic obstructive pulmonary disease (COPD), kidney diseases, dementia, peripheral vascular diseases, liver diseases, and stroke all contribute to increased risk. Moreover, comorbidity scores, polypharmacy, and abnormal vital signs upon admission, as well as hospital admission during non-working days, such as weekends or public holidays [11]. have been linked to increased mortality. Other significant factors include substance abuse [9]. ICU transfer, and certain laboratory markers, such as potassium, sodium, urea, C-reactive protein, haemoglobin, haematocrit, total bicarbonate, white blood cell count, albumin, pH, bilirubin, and creatinine [9,12–15].

Two prominent Early Warning Systems (EWS), namely the Modified Early Warning System (MEWS) and the National Early Warning System (NEWS), are widely used in healthcare settings [7,16]. The MEWS assesses patients based on five physiological components, including temperature, respiratory rate, systolic blood pressure, heart rate, and level of consciousness, using the AVPU (Alert, Verbal response, Pain response, Unresponsive) scale. NEWS includes these five components, as well as oxygen saturation and the use of supplemental oxygen. Data suggest that a MEWS score of 5 or above increases the probability of mortality by 22 times [16]. A systematic review of risk prediction tools for hospital mortality in general wards found that EWS, such as MEWS and NEWS, were effective in predicting imminent mortality risk (within 24 or 48 h). Moreover, EWS performed better at predicting mortality among patients with infections compared to those without. Recent findings from a randomized clinical trial suggested that, compared to EWS, machine learning models reduced the 30-day mortality risk by 16%, although deterioration risk for inpatient patients was not specifically addressed [7].

Numerous studies have developed machine learning models to predict in-hospital mortality, often focusing on specific patient groups, such as those with sepsis [17]. A study analyzed approximately 2.67 million patients with COVID-19 infection across 146 countries to develop a machine learning model aimed at predicting mortality and facilitating timely care [18]. The study utilized several machine learning algorithms, including Support Vector Machine (SVM), Artificial Neural Networks, Random Forest, Decision Tree, Logistic Regression, and K-Nearest Neighbor (KNN), to achieve accurate mortality predictions [18]. Furthermore, machine learning models have been developed for other specific patient populations, such as patients with malignancy [19], patients admitted to ICU with heart failure [20], patients with acute myocardial infarction [21], and those with liver cirrhosis [22].

While the aforementioned studies have demonstrated the potential of machine learning in predicting adverse outcomes in specific populations such as sepsis or COVID-19, there remains a critical gap in the literature regarding the development and validation of such models for the heterogeneous patient population admitted to a General Internal Medicine unit. Furthermore, few studies have conducted a systematic, head-to-head comparison of multiple, diverse machine learning algorithms—from traditional regression to complex ensembles and neural networks—within this specific clinical setting.

Given the limitations of previously developed machine learning models for predicting inpatient deterioration and the lack of validation in local health settings, this study aims to develop model to predict inpatient deterioration for patients admitted under the care of general medicine unit. The model will use patient data from electronic health records and apply machine learning techniques to achieve high accuracy in predicting in-hospital deterioration risk.

The article is structured to provide a comprehensive overview of the research. It starts with an abstract, followed by an introduction that frames the clinical problem and study objectives. The materials and methods section describes the study design, data handling, and modeling approach. Results detail patient characteristics, key predictors, and model performance. The discussion interprets findings, explores clinical implications, and notes limitations. The conclusion summarizes contributions and future directions.

2. Materials and Methods

2.1. Study Setting

This study analysed patients data from individuals admitted to the General Internal Medicine unit between 1 January 2022, and 31 December 2023. Most patients in this unit are admitted through the Emergency Department, presenting with a wide range of underlying aetiologies. The study included patients aged 18 years and older who were admitted during the specified period [23].

2.2. Data Collection

Relevant patient data were extracted from electronic health records (EHRs). The collected data included demographic and clinical factors such as age, sex, primary diagnosis (classified using ICD-10), polypharmacy, length of stay, previous hospitalizations within the last 6–12 months, readmissions within 12 months, and the number of comorbidities. Comorbid conditions, including cerebrovascular disease, COPD, diabetes mellitus, cardiovascular disease, anemia, active malignancy, and cognitive impairment or dementia, were also documented. Vital signs, including heart rate, respiratory rate, blood pressure, oxygen saturation (SPO₂), oxygen supplementation requirements, temperature, were recorded at key intervals (24, 16, and 8 h preceding the outcome). Laboratory data collected within 24–48 h before outcomes included levels of C-reactive protein, haemoglobin, sodium, potas-

sium, urea, creatinine, lactate, blood pH, total bicarbonate, white blood cell count, and liver function tests (bilirubin, albumin, and ALP). Additionally, system-related variables such as day of admission, length of stay, and transfer requirements (including ICU or high-dependency unit transfers) were included.

To incorporate a temporal dimension into the static models, vital signs and laboratory values recorded at different time intervals prior to the outcome (i.e., within 8, 16, and 24 h) were treated as distinct, independent features. For example, a patient's respiratory rate within 8 h and their respiratory rate within 16 h were included as two separate columns in the feature matrix. This method allows the model to learn from the most recent measurements as well as from slightly older values, providing a richer, albeit not truly sequential, representation of the patient's recent clinical history.

2.3. Definition

The primary outcome for this study was inpatient deterioration, a composite endpoint defined as the occurrence of any of the following events: transfer to the intensive care unit (ICU), transfer to a high-dependency unit (HDU), or in-hospital mortality. The non-deterioration endpoint was defined as successful hospital discharge without any of these events.

This definition was chosen for its methodological robustness and clinical relevance. The selected outcomes are significant adverse events that are unambiguously documented within electronic health records, thereby minimizing classification error. Furthermore, they represent a clear escalation of care, which directly aligns with the primary purpose of an early warning system—to identify patients at imminent risk of severe decline. The use of such a binary classification is a standard and well-established approach in the development of prognostic models, as it allows for clear and robust model evaluation [24].

2.4. Data Analysis

Data analysis was conducted using Python 3.12.7, with data manipulation via pandas (v2.0+) and numpy (v1.26+). Machine learning was implemented using scikit-learn (v1.3+), and neural networks were built with Keras (v3.6+) on TensorFlow. Methods included: Preprocessing with SimpleImputer, OneHotEncoder, and StandardScaler; Feature Selection using SelectKBest with ANOVA F-test; Class Imbalance addressed by SMOTE from imblearn; Model Training with LogisticRegression, RandomForestClassifier, GradientBoostingClassifier, and SVC; Hyperparameter Tuning via GridSearchCV; and an Ensemble Method using VotingClassifier.

The dataset was partitioned into a training set, comprising 70% of the data, and a testing set, containing the remaining 30%, to allow for model development and unbiased performance evaluation.

Python was chosen for its robust machine learning ecosystem, integrating libraries like scikit-learn, pandas, and Keras/TensorFlow. This enables efficient data processing and model development. Its flexibility, transparency, and strong community support make it ideal for reproducible, advanced predictive modeling in healthcare research.

The detailed analytical pipeline and data analysis workflow is shown in Figure 1.

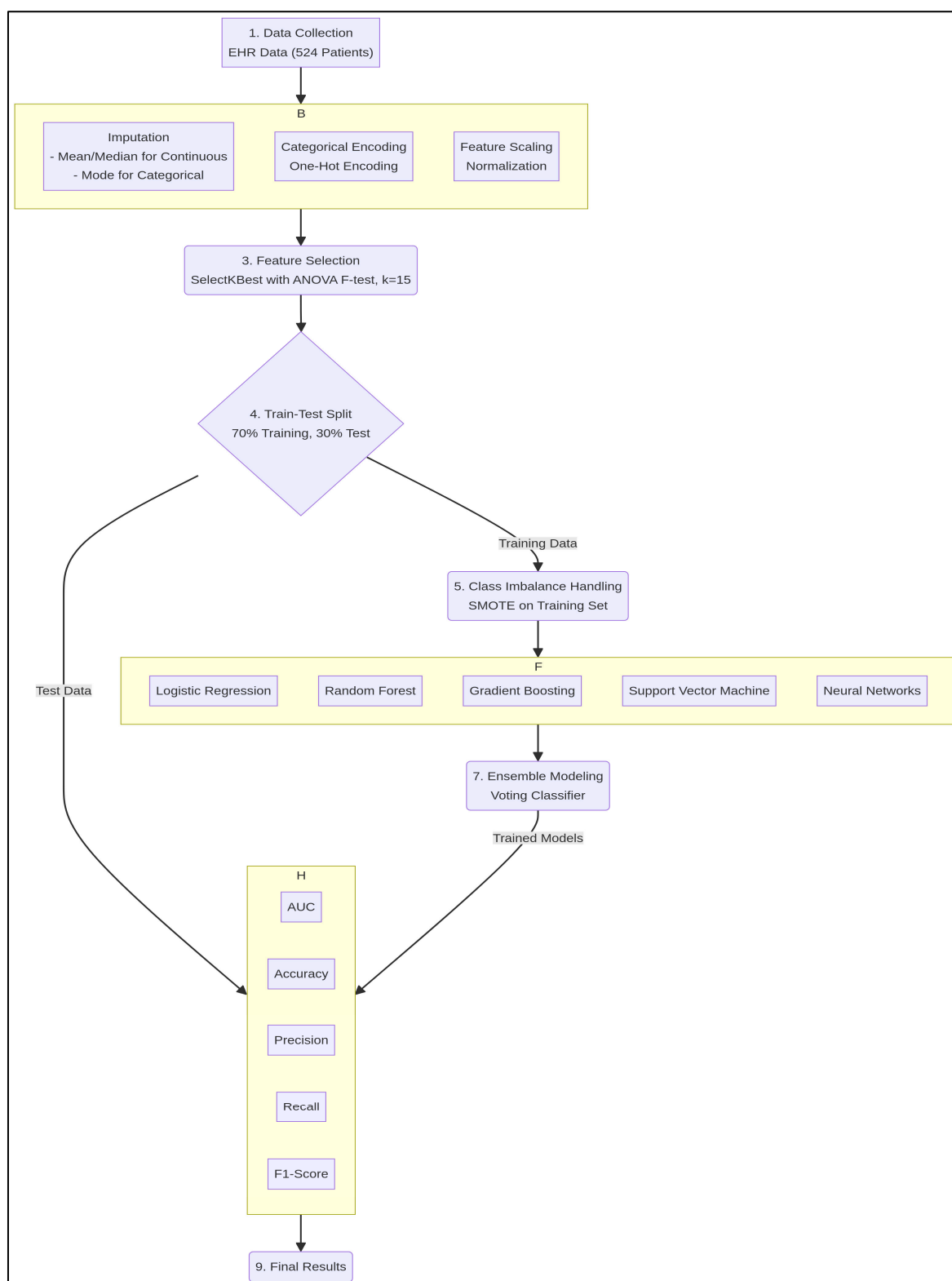


Figure 1. Visual presentation of the data collection and analytical pipeline (total $n = 524$ patients).

2.5. Data Preprocessing

The dataset contained both continuous and categorical variables. Continuous variables, including clinical measurements such as age, BMI, vital signs, and laboratory results, were subjected to imputation for missing data. For non-normally distributed continuous variables, missing values were filled with the median, while for normally distributed variables, the mean was used to replace missing values.

For categorical variables, missing values were imputed with the mode (the most frequent value). The dataset included categorical features such as gender, day of admission, smoking status, and comorbidities, which were also pre-processed using one-hot encoding for categorical variables with multiple levels (such as 'Primary Diagnosis (ICD10)', 'Code Status', 'Day of Admission', and 'Functional Status'). All data was checked to ensure there were no remaining missing values after preprocessing, with the final dataset being complete and ready for analysis.

The handling of missing data was a critical step in our preprocessing pipeline. We first conducted an analysis of data completeness for all candidate variables. To ensure the robustness of our models, any variable with a missing data rate exceeding 15% was excluded from consideration for the final feature set. For all remaining variables, which had a high degree of completeness (i.e., <15% missing values), we employed a standard imputation strategy. Missing values for continuous variables were replaced with the median of that variable, while missing values for categorical variables were replaced with the mode. This two-step approach ensured that our models were trained on a complete and high-quality dataset.

2.6. Feature Selection

To reduce the dimensionality of the dataset and focus on the most important variables, feature selection was performed using the SelectKBest method with ANOVA F-test as the scoring function. This method identified the top 15 most important features that were then used to train the predictive models. These features were selected based on their statistical significance and relevance to the outcome variable.

2.7. Model Development

To predict inpatient deterioration, several machine learning models were trained and assessed. These included Logistic Regression, Random Forest Classifier, Gradient Boosting Classifier, Support Vector Machine (SVM), and both simple and complex Neural Networks. The Synthetic Minority Over-sampling Technique (SMOTE) was applied exclusively to the training set to address class imbalance by creating synthetic samples of the minority class, thereby improving model performance.

The original dataset of 524 patients was imbalanced, containing 156 (29.8%) instances of inpatient deterioration. The SMOTE procedure, applied only to the training set, corrected this by generating synthetic samples of the minority class, resulting in a perfectly balanced training dataset with a 1:1 ratio of deterioration to non-deterioration events.

2.8. Model Evaluation

The performance of each model was assessed using a comprehensive set of evaluation metrics, chosen to provide a robust assessment, particularly given the class imbalance inherent in the dataset. The following metrics were used, where TP, TN, FP, and FN represent True Positives, True Negatives, False Positives, and False Negatives, respectively.

Accuracy: This metric represents the proportion of all predictions that were correct. While commonly reported, it can be misleading in imbalanced datasets and is included for completeness but interpreted with caution.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (1)$$

Precision (Positive Predictive Value): Precision measures the proportion of predicted positive cases that were actually correct. High precision is important to minimize false alarms and prevent alert fatigue.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (2)$$

Recall (Sensitivity): Recall measures the proportion of actual positive cases that were correctly identified. This is a critical metric, as failing to identify a patient who will deteriorate (a false negative) has severe clinical consequences.

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (3)$$

F1-Score: The F1-Score is the harmonic mean of Precision and Recall, providing a single score that balances the two. It is particularly useful for imbalanced datasets.

$$\text{F1-Score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (4)$$

Area Under the Receiver Operating Characteristic Curve (AUC): The AUC represents the model's ability to distinguish between the deteriorating and non-deteriorating classes. An AUC of 1.0 indicates a perfect classifier, while 0.5 indicates performance no better than random chance. It is a robust metric for imbalanced data as it is independent of the chosen classification threshold [25,26].

2.9. Hyperparameter Tuning

To optimize the performance of the machine learning models, hyperparameter tuning was conducted using a 5-fold cross-validated grid search (GridSearchCV) on the training dataset, with the Area Under the Curve (AUC) as the scoring metric. The specific hyperparameters explored and the optimal values identified for each model are detailed in Table 1.

Table 1. Hyperparameter Tuning Details for Machine Learning Models.

Model	Hyperparameter	Range of Values Explored	Optimized Value
Random Forest	n_estimators	[100, 200, 300]	300
	max_depth	[None, 10, 20, 30]	None
	min_samples_split	[2, 5, 10]	2
	min_samples_leaf	[1, 2, 4]	1
Gradient Boosting	n_estimators	[50, 100, 150]	150
	learning_rate	[0.01, 0.05, 0.1]	0.1
	max_depth	[3, 5, 7]	7
	subsample	[0.8, 1.0]	0.8
Support Vector Machine (SVM)	C	[0.1, 1, 10]	10
	kernel	['linear', 'rbf']	'rbf'
	gamma	['scale', 'auto']	'auto'
	class_weight	['balanced', None]	None

2.10. Cross-Validation and Voting Classifier

To assess the stability of the models, 5-fold cross-validation was performed for each model. Additionally, a Voting Classifier was created using a combination of the best-performing models (Logistic Regression, Random Forest, Gradient Boosting, and SVM) to make final predictions. This ensemble approach aggregates the predictions of the individual models, potentially improving the overall performance.

To leverage the strengths of the individual models, we constructed a hard voting ensemble classifier. This classifier aggregates the predictions from the four base models (Logistic Regression, Random Forest, Gradient Boosting, and SVM) and the final output is determined by a majority vote. It is important to note that because a hard voting mechanism outputs a discrete class label rather than a continuous probability score, the Area Under the Curve (AUC) metric could not be calculated for this ensemble model Figure 2.

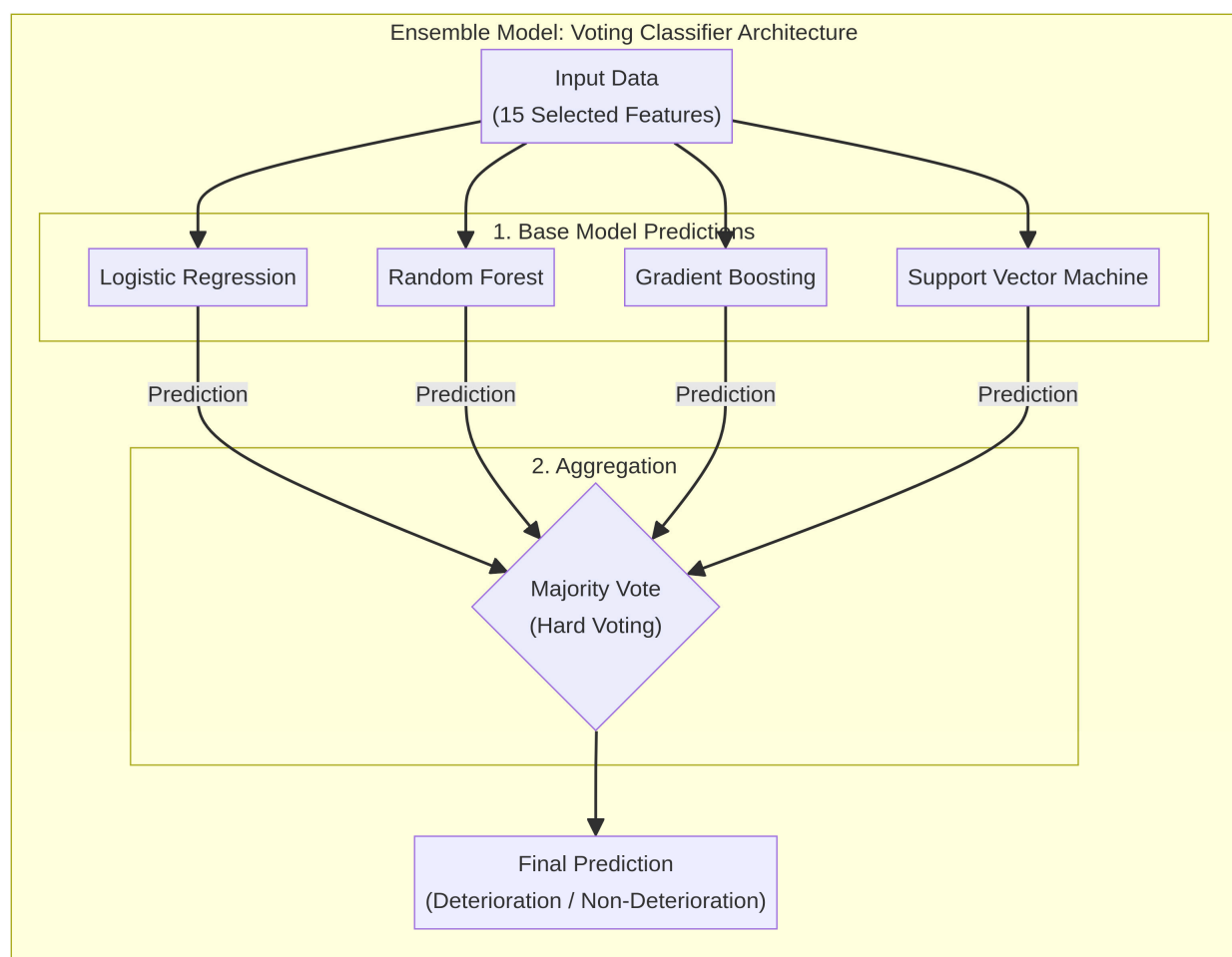


Figure 2. Voting ensemble classifier architecture.

3. Results

The study included 524 patients with a median age of 59.3 years (IQR: 31.7). Most were female (51.5%) with a median stay of 5 days (IQR: 5). Common comorbidities were hypertension (34.4%) and diabetes (38.0%). Most were fully independent (78.2%). All detailed included features were reported in Tables 2 and 3. For continuous variables, the median values varied across the features, with notable findings including a median Body Mass Index (BMI) of 26.0 (IQR: 3.93), a median haemoglobin level of 11.2 g per decilitre (g/dL) (IQR: 3.5 g/dL), and a median C-reactive Protein (CRP) level of 18.0 milligrams per liter (mg/L) (IQR: 35.0 mg/L). The vital signs such as heart rate and blood pressure

also showed variation over different time points, with a median heart rate of 79 beats per minute (bpm) (IQR: 17 bpm) within 8 h of the outcome (Table 2).

Table 2. Categorical variables of all included patients ($n = 524$ patients).

Variable	Details
Women	270 (51.5%)
Functional status	
Functional status Fully dependent	Fully dependent 77 (14.7%); Partially independent 37 (7.1%); Fully independent 410 (78.2%)
Primary Diagnosis according to ICD-10	J15: 44 (8.4%), A90: 32 (6.1%), I50.9: 30 (5.7%), N39.0: 28 (5.3%), I63.51: 18 (3.4%), ...
Code Status (Don't resuscitate)	41 (7.8%)
Need for oxygen therapy (O ₂)	54 (10.3%)
Active Smoking or history of smoking	48 (9.2%)
Alcohol abuse	36 (6.9%)
Substance abuse	15 (2.9%)
Polypharmacy (5 or more medications on admission)	245 (46.8%)
Use of antiepileptics	43 (8.2%)
Use of diuretics	132 (25.2%)
Use of anticoagulant	76 (14.5%)
Hospital Acquired Infection	25 (4.8%)
Comorbidities (≥ 3)	237 (45.2%)
Diabetes Mellitus	199 (38.0%)
Hypertension	180 (34.4%)
Dyslipidaemia	119 (22.7%)
Cardiovascular disease	181 (34.5%)
Heart failure	106 (20.2%)
Chronic obstructive pulmonary disease	132 (25.2%)
Cerebrovascular disease	97 (18.5%)
Active malignancy	13 (2.5%)
Dementia	31 (5.9%)
Acute kidney injury	119 (22.7%)
Chronic kidney disease	0100 (19.1%)
Liver Cirrhosis	15 (2.9%)
Blood transfusion within last 12 months	37 (7.1%)

For categorical variables, 89.7% of patients did not require oxygen therapy, and 93.1% were non-smokers. A significant portion of patients had comorbid conditions, with 45.2% of patients having at least three comorbidities (Table 3).

Table 3. Continuous characteristics of included patients ($n = 524$ patients).

Variable	Median (IQR)
Number of admissions in the preceding 3 months	0.0 (0.0)
Number of admissions in the last 12 months	0.0 (0.0)
Number of emergency presentation in the last 12 months	0.0 (1.0)
Age	59.285 (31.74)
Height (cm)	160.0 (6.0)
Weight (kg)	67.1 (16.3)
BMI	26.0 (3.9)
Temperature within 8 h	36.7 (0.3)
Temperature within 16 h	36.7 (0.3)
Temperature within 24 h	36.7 (0.3)
Heart rate within 8 h	79.0 (17.0)
Heart rate within 16 h	80.0 (16.0)
Heart rate within 24 h	80.0 (18.0)
Respiratory rate within 8 h	18.0 (2.0)
Respiratory rate within 16 h	18.0 (2.0)
Respiratory rate within 24 h	19.0 (2.0)
Systolic Blood pressure within 8 h	125.0 (28.3)
Diastolic Blood pressure within 8 h	67.5 (16.0)
Systolic Blood pressure within 16 h	124.0 (28.0)
Diastolic Blood pressure within 16 h	67.0 (15.0)
Diastolic Blood pressure within 24 h	69.0 (15.0)
Amount of O ₂ in Litres	0.0 (0.0)
Number of regular medications	4.0 (1.0)
Hemoglobin (g/dL)	11.2 (3.5)
Hematocrit (%)	0.36 (0.1)
Platelet ($\times 10^9$ /L)	275.5 (145.5)
White Blood Cell (WBC) Count ($\times 10^9$ /L)	7.15 (4.0)
International Normalized Ratio (INR)	1.03 (0.1)
Calcium (mg/dL)	2.29 (0.1)
Phosphate (PO ₄) (mg/dL)	1.13 (0.3)
C-reactive Protein (CRP) (mg/L)	18.0 (2.0)
Anion Gap (mmol/L)	11.0 (3.0)
Bicarbonate (HCO ₃) (mmol/L)	23.0 (5.0)
Chloride (mmol/L)	102.0 (6.0)
Creatinine (mg/dL)	72.0 (41.25)
Estimated Glomerular Filtration Rate (eGFR) (mL/min/1.73 m ²)	88.0 (31.0)
Potassium (mmol/L)	4.2 (0.6)
Sodium (mmol/L)	137.0 (5.0)

Table 3. *Cont.*

Variable	Median (IQR)
Urea (mg/dL)	5.0 (0.2)
Alanine Aminotransferase (ALT) (U/L)	19.0 (2.0)
Albumin (g/dL)	38.5 (9.0)
Alkaline Phosphatase (ALP) (U/L)	87.0 (42.25)
Aspartate Aminotransferase (AST) (U/L)	22.0 (17.0)
Bilirubin total (mg/dL)	7.0 (6.0)
Magnesium (mg/dL)	0.83 (0.1)
Troponin (ng/L)	27.0 (20.25)
Glucose Glucose (mg/dL)	7.0 (0.0)
Lactate Lactate (mmol/L)	1.4 (0.3)
Lactate Dehydrogenase (LDH) (U/L)	255.0 (0.0)
Brain Natriuretic Peptide (BNP) (pg/mL)	2077.0 (0.0)
Vitamin B12 (pg/mL)	332.0 (0.0)
Ferritin (ng/mL)	247.5 (0.0)
Number of Blood transfusion within last 12 months	0.0 (0.0)

The analysis identified the top 15 factors affecting the risk of deterioration. Factors related to oxygen amount, urea levels, and recent admissions were linked to increased risk. In contrast, functional and code status were associated with decreased risk. Functional status had the strongest influence (Table 4, Figure 3).

Table 4. Key Factors Influencing Risk of Deterioration: Factors with a positive direction indicate an increased risk of inpatient deterioration, while factors with a negative effect indicate a reduced risk of deterioration.

Feature	Coefficient	Effect Direction	Magnitude
Functional status: full independence	−1.008544	Negative	1.008544
Code Status: Full code	−0.80055	Negative	0.80055
Amount of Oxygen (O ₂) in Litres	0.651186	Positive	0.651186
Urea	0.415895	Positive	0.415895
Number of admissions in the last 12 months	0.3915	Positive	0.3915
Anion gap	0.300406	Positive	0.300406
Troponin	0.233988	Positive	0.233988
C-reactive protein (CRP)	0.230803	Positive	0.230803
Alkaline Phosphatase (ALP)	0.229027	Positive	0.229027
International Normalized Ratio (INR)	0.215248	Positive	0.215248
Use of diuretics	0.115222	Positive	0.115222
Albumin	−0.090686	Negative	0.090686
Need for Oxygen (O ₂)	−0.086899	Negative	0.086899
Heart Failure (HF)	−0.047107	Negative	0.047107
Estimated Glomerular Filtration Rate (eGFR)	0.003575	Positive	0.003575

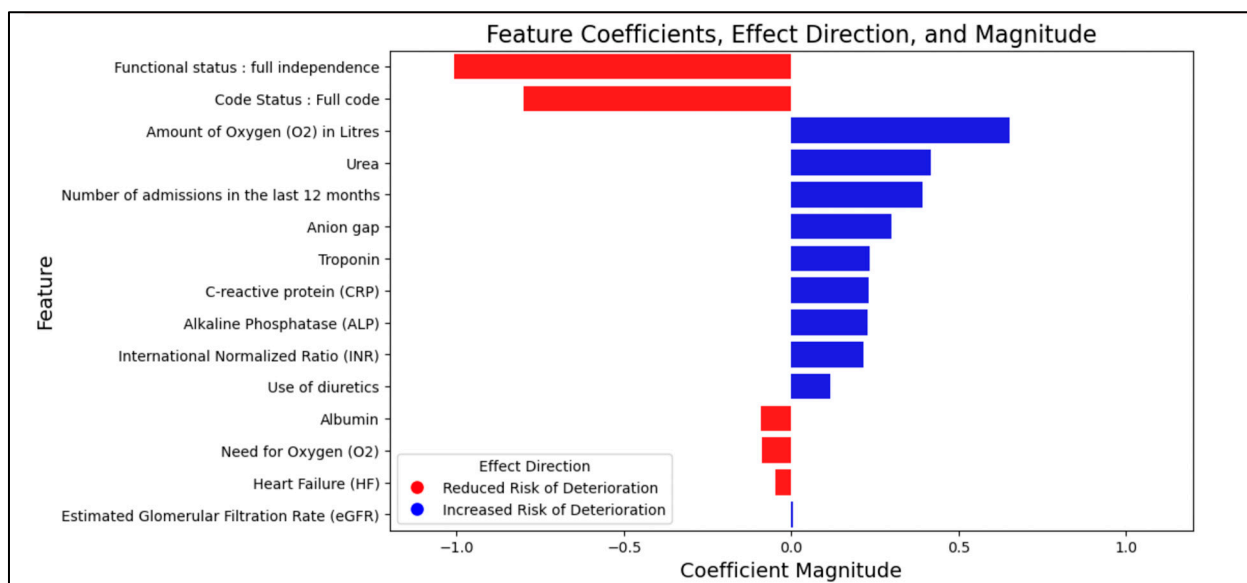


Figure 3. Key Factors Influencing Risk of Deterioration.

The machine learning models trained on the dataset showed varying levels of performance. The Random Forest model emerged as the best-performing model, with an AUC of 0.837, an accuracy of 87.3%, and an F1 score of 47.4% (Table 5).

Table 5. Performance Metrics of Machine Learning Models for Predicting Inpatient Deterioration.

Model	AUC	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Logistic Regression	0.77	81.0	40.6	54.2	46.4
Random Forest	0.837	87.3	64.3	37.5	47.4
Gradient Boosting	0.76	86.7	58.8	41.7	48.8
SVM	0.81	80.4	38.7	50.0	43.6
Neural Network 1	0.69	79.7	38.2	54.2	44.8
Neural Network 2	0.59	79.1	30.4	29.2	29.8
Voting Classifier	N/A	84.8	50.0	29.2	36.4

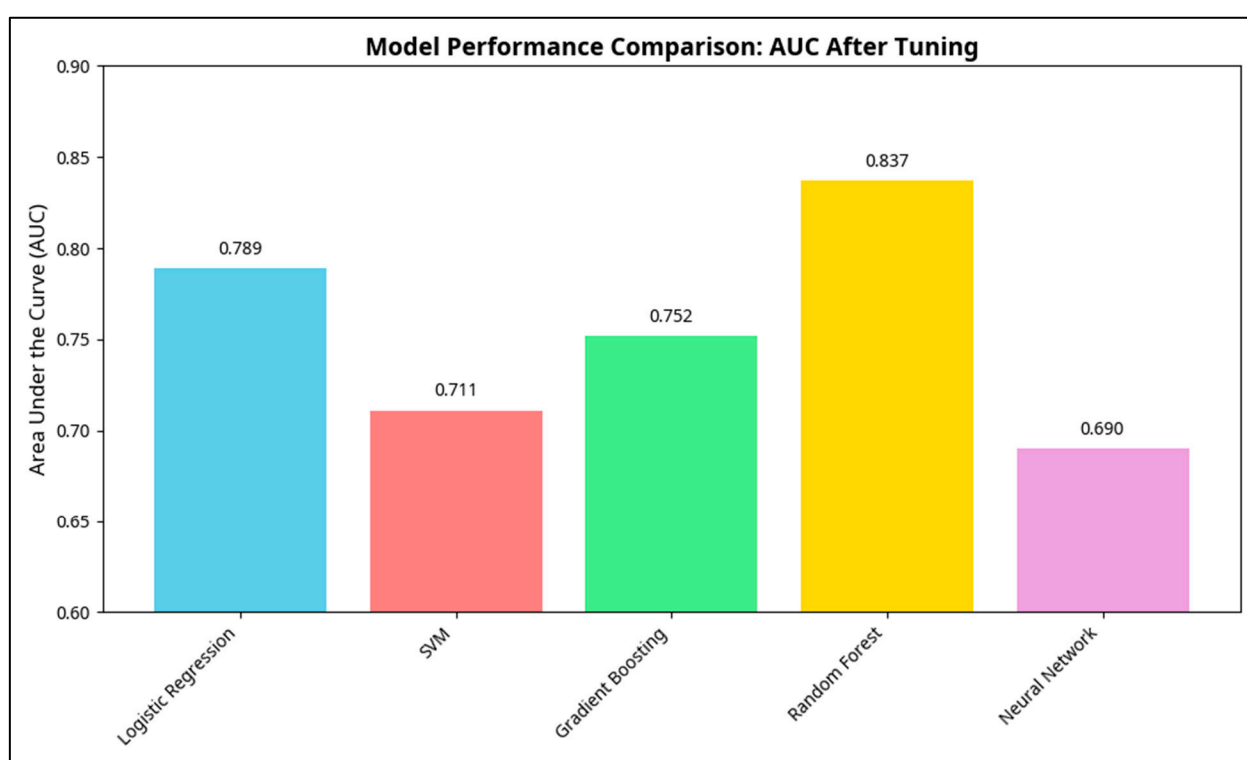
The neural network models performed poorly in comparison, with the first neural network model yielding an AUC of 0.69 and the second model obtaining an AUC of 0.59. The ensemble Voting Classifier, which combined Logistic Regression, Random Forest, Gradient Boosting, and SVM, achieved an accuracy of 84.8%, but the AUC could not be computed due to the lack of probability estimates.

After hyperparameter tuning, the Random Forest model further improved with an AUC of 0.84, accuracy of 85.4%, and F1 score of 43.9%. Gradient Boosting showed a slightly lower performance with an AUC of 0.76 and an F1 score of 48.8%, while the SVM model, despite tuning, achieved a lower AUC of 0.71 and an F1 score of 27.3% (Table 6).

Table 6. Optimized Model Performance After Hyperparameter Tuning.

Model	AUC (Tuned)	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Random Forest	0.84	85.4	52.9	37.5	43.9
Gradient Boosting	0.75	86.1	57.1	33.3	42.1
SVM	0.71	79.7	30.0	25.0	27.3

To visually summarize the comparative performance of the models, Figure 4 presents the Area Under the Curve (AUC) for each of the primary models after hyperparameter tuning. The Random Forest model clearly demonstrates the highest discriminatory ability.

**Figure 4.** Comparative AUC Performance of Machine Learning Models After Hyperparameter Tuning.

To provide context on the computational feasibility of the models for clinical implementation, approximate training and prediction times were recorded. While the hyperparameter tuning phase was computationally intensive, the training times for the final models on the full training set were minimal. Logistic Regression trained in under a second, while Random Forest and Gradient Boosting models trained in approximately 2–7 s. The Support Vector Machine model was the slowest, taking approximately 10–15 s. For all trained models, the prediction latency for a single new patient was negligible (well under one second), confirming their suitability for real-time clinical decision support applications.

During the 5-fold cross-validation performed on the training data, the models demonstrated high stability. The Random Forest model achieved a mean Area Under the Curve (AUC) of 0.980, and the Gradient Boosting model achieved a mean AUC of 0.983 across the validation folds.

4. Discussion

This study introduces a machine learning model to predict inpatient deterioration risk, measured by outcomes such as ICU transfer or mortality versus successful discharge. These models enhance risk stratification, enabling timely interventions and optimized patient care.

The amount of oxygen requirements is a well-recognized marker of hypoperfusion and physiological compromise, making them a critical indicator of clinical decline. In our study, increased oxygen consumption emerged as a significant predictor of inpatient clinical deterioration, corroborating findings from prior research that highlight the role of oxygen needs in identifying patients at risk of adverse outcomes [27–29]. Similarly, elevated urea levels were strongly associated with a higher risk of inpatient deterioration. This association aligns with existing literature, which emphasizes the link between elevated urea and worsening clinical status, likely reflecting underlying renal dysfunction, dehydration, or metabolic stress [30,31].

The anion gap was identified as a predictor of clinical deterioration in our study, with a positive coefficient indicating that higher values were associated with an increased risk of ICU transfer or death. This finding underscores the relevance of metabolic disturbances, such as lactic acidosis, ketoacidosis, or renal dysfunction, in the progression to critical illness. A 2023 study from Boston, USA, concluded that for patients with liver failure, the anion gap serves as a moderate predictor of in-hospital mortality [32]. In contrast, another study showed that in septic patients, the anion gap corrected for albumin and lactate exhibits a strong correlation with the strong ion gap, but neither parameter demonstrated significant predictive ability for mortality. This suggests that while the anion gap reflects underlying metabolic disturbances, its role in predicting outcomes in sepsis may be limited [33].

Prior hospitalization within the past 12 months is a significant predictor of clinical deterioration, linked to higher readmission rates and complications, especially in patients with chronic conditions [34]. Our analysis confirms that frequent admissions correlate with increased deterioration risk. This relationship likely reflects the ongoing vulnerability of patients who frequently require acute care, emphasizing the need for close monitoring and targeted interventions.

This study identified several features associated with a reduced risk of inpatient deterioration, including full code status, fully independent functional status, admission with heart failure, albumin level, and the need for O₂ therapy.

Code status is crucial for determining healthcare needs, reflecting a patient's preferences for life-sustaining measures, from full code to DNR. Our study found full code status linked to reduced inpatient deterioration, likely due to more intensive monitoring and interventions. DNR patients may receive comfort-focused care, aligning with their preferences and potentially avoiding unnecessary ICU transfers [35–37].

Albumin, a marker of nutritional and inflammatory status, is strongly linked to clinical deterioration. Low levels indicate systemic inflammation, chronic disease, or malnutrition, contributing to poor outcomes [33,38–40]. Its predictive value reflects various physiological stressors. Assessing albumin in routine evaluations offers insights into health status and risk of adverse events. Using albumin in predictive models can improve risk stratification and guide interventions like nutritional support and inflammation management [41].

In our analysis, heart failure (HF) showed a negative coefficient, indicating a reduced likelihood of clinical deterioration. This likely reflects the close monitoring and management patients with HF receive, which may prevent acute decompensation. Frequent healthcare encounters for HF patients enables early interventions. While both HF and non-HF hospitalizations increase mortality risk, cardiovascular mortality is higher after HF hospitalizations, especially beyond 30 days [42].

Patients with heart failure (HF), whether with preserved or reduced ejection fraction, have similar re-hospitalization rates. However, those with reduced ejection fraction face higher mortality risk, while resource use is slightly higher in those with preserved ejection fraction. Our model's association of improved outcomes in HF patients should be cautiously interpreted due to potential confounding factors or survivor bias. Future studies should include detailed severity metrics to better understand HF's impact on clinical outcomes [43].

Machine learning models are increasingly utilized to predict clinical deterioration in a specific population like ICU patients with HF [20], sepsis [17], acute myocardial infarction [21]. Such models typically achieve high predictive performance but may lack generalizability. Very few models have been utilized to predict clinical outcomes in disease-independent populations and such models face challenges in optimizing sensitivity and specificity but hold the potential for wider applicability. In our study, we evaluated several machine learning models, including Logistic Regression, Random Forest, Gradient Boosting, SVM, and Neural Networks, to predict clinical deterioration defined as death or ICU transfer in a broader diseases-independent model to ensure generalizability to all admitted patients.

Random Forest models are widely recognized for their superior performance in predicting in-hospital mortality compared to traditional clinical decision rules and other machine learning models, such as logistic regression and classification and regression trees (CART) [44,45]. However, in our study, Logistic Regression emerged as the most balanced model, demonstrating moderate precision and achieving the highest recall among all models. In contrast, Random Forest and Gradient Boosting achieved higher accuracy and precision but were hindered by lower recall, limiting their ability to effectively identify high-risk patients. Neural Network 1 demonstrated recall comparable to Logistic Regression but fell short in precision and F1 score compared to other models. Additionally, the ensemble Voting Classifier, which integrated Logistic Regression, Random Forest, Gradient Boosting, and SVM, exhibited strong accuracy.

It is noteworthy that the tree-based ensemble models, particularly Random Forest, outperformed the neural network models in this study. This finding is likely attributable to several factors. First, neural networks typically require larger datasets to effectively learn complex patterns and generalize well; our dataset of 524 patients may be insufficient for deep learning approaches to reach their full potential. Second, tree-based models are often highly effective and have a strong inductive bias for structured, tabular data of the kind used in this study. Finally, our current model utilizes features from discrete time points rather than as a continuous time-series, a format for which more advanced recurrent neural network architectures (e.g., LSTMs) would be better suited.

To address the trade-offs between precision and recall, we explored ensemble techniques, which aggregate predictions from multiple models to leverage their individual strengths and mitigate weaknesses. Ensembles enhance generalization and reduce overfitting, particularly in datasets with diverse patient characteristics [46]. Our Voting Classifier demonstrated the potential to improve accuracy; however, its recall remained insufficient for high-stakes applications like clinical deterioration prediction.

From a methodological standpoint, the superior performance of the Random Forest model is a key finding. This result aligns with a growing body of evidence suggesting that tree-based ensemble algorithms often exhibit a strong inductive bias for structured, tabular data commonly found in electronic health records. Unlike linear models, they can natively capture complex, non-linear interactions between features without extensive data transformation. In contrast, while the neural network models are theoretically more powerful, their relatively modest performance here highlights their need for larger datasets to effectively learn and generalize. For datasets of this scale and structure,

the trade-off between the high performance of a Random Forest and the inherent interpretability of a simpler model like Logistic Regression becomes a central consideration for clinical implementation.

It is critical to interpret the individual coefficients of the logistic regression model with caution due to potential multicollinearity. For instance, the seemingly counterintuitive positive coefficient for oxygen saturation is likely confounded by the use of supplemental oxygen; patients maintaining high saturation levels while receiving significant oxygen support are inherently at a higher risk. This underscores the importance of evaluating overall model performance and feature importances derive.

Our findings underscore the need for more robust and interpretable models to predict critical outcomes in clinical settings. Accurate and timely prediction of deterioration can facilitate early interventions, such as rapid response team activation or intensive monitoring, ultimately improving patient outcomes and optimizing resource allocation. However, our study also highlights several challenges, including the impact of dataset non-specificity and the difficulty of improving recall without significantly compromising other performance metrics. These challenges are particularly relevant in clinical contexts with diverse patient populations, such as those in rapidly growing healthcare systems facing resource constraints.

Several technical limitations must be acknowledged when interpreting these results. First, as a single-center study, the trained models may have learned patterns specific to our institution's patient population and data-recording practices, which could limit their generalizability when applied to external datasets. Second, the dataset size ($n = 524$) is a significant constraint, particularly for the data-hungry neural network architectures, and increases the risk of overfitting. A third major limitation is the static nature of our feature set; the model was trained on a snapshot of patient data rather than on continuous time-series variables. This approach fails to capture the temporal dynamics of a patient's physiological state, which is a critical component of clinical deterioration.

This work lays the groundwork for several important future research directions from a computational perspective. The most critical next step is the external validation of our best-performing model on a large, multi-center dataset to rigorously assess its generalizability. Furthermore, to address the limitation of static data, future work should focus on employing recurrent neural network (RNN) architectures, such as Long Short-Term Memory (LSTM) or Gated Recurrent Units (GRUs), to model the time-series nature of vital signs and lab results. Finally, to move beyond 'black-box' predictions, the application of Explainable AI (XAI) techniques, such as SHAP (SHapley Additive exPlanations), is essential. Integrating XAI would not only help in debugging and refining the model but would also be crucial for fostering clinical trust and facilitating the adoption of such tools in real-world decision support systems.

5. Conclusions

This study developed and validated machine learning models to predict inpatient deterioration using electronic health record data. Among the models tested, Random Forest achieved the highest performance outperforming logistic regression, support vector machines, and neural networks. Key predictors included functional status, oxygen requirements, and serum urea, highlighting their clinical importance. While logistic regression offered interpretability, Random Forest better captured complex patterns in the data. Neural networks performed modestly, suggesting tree-based methods are preferable for structured datasets of this size. Future work should focus on external validation with larger, multi-center datasets to ensure generalizability. Incorporating temporal data with recurrent neural networks and applying interpretable AI techniques, such as SHAP or LIME, will be

important for understanding model decisions and supporting clinical adoption. These steps will help integrate predictive models into real-time clinical decision support, improving patient care and outcomes.

Author Contributions: Conceptualization, A.M.A.A.; methodology, A.M.A.A.; software, A.M.A.A.; validation, S.A.J. and L.A.W.; formal analysis, A.M.A.A.; investigation, M.A.-H. and H.H.G.; resources, A.M.A.A.; data curation, M.A.-H. and H.H.G.; writing—original draft preparation, S.A.J., L.A.W., M.A.-H. and H.H.G.; writing—review and editing, A.M.A.A.; visualization, A.M.A.A.; supervision, A.M.A.A.; project administration, A.M.A.A.; funding acquisition, A.M.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Ministry of Higher Education, Innovation and Research, Muscat, Oman, grant number BFP/GRG/HSS/24/145.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Medical and Research Ethics Committee of the College of Medicine and Health Sciences, Sultan Qaboos University, Muscat, Oman (Approval Number: MREC# 3293, Ref. No. SQU-EC/103/2024).

Informed Consent Statement: Patient consent was waived due to the retrospective nature of the study using anonymized patient data.

Data Availability Statement: The dataset analyzed in this study is not publicly available due to institutional restrictions but may be provided upon reasonable request to the corresponding author.

Acknowledgments: During the preparation of this manuscript, the authors used generative AI for the purposes of language editing and improving the flow of the text. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Ferrando-Vivas, P.; Jones, A.; Rowan, K.M.; Harrison, D.A. Development and validation of the new ICNARC model for prediction of acute hospital mortality in adult critical care. *J. Crit. Care* **2017**, *38*, 335–339. [[CrossRef](#)]
2. Sankey, C.B.; McAvay, G.; Siner, J.M.; Barsky, C.L.; Chaudhry, S.I. “Deterioration to Door Time”: An Exploratory Analysis of Delays in Escalation of Care for Hospitalized Patients. *J. Gen. Intern. Med.* **2016**, *31*, 895–900. [[CrossRef](#)] [[PubMed](#)]
3. Padilla, R.M.; Mayo, A.M. Clinical deterioration: A concept analysis. *J. Clin. Nurs.* **2018**, *27*, 1360–1368. [[CrossRef](#)] [[PubMed](#)]
4. Hall, M.J.; Levant, S.; DeFrances, C.J. *Trends in Inpatient Hospital Deaths: National Hospital Discharge Survey, 2000–2010*; U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics: Washington, DC, USA, 2013; pp. 1–8.
5. Porcel-Gálvez, A.M.; Barrientos-Trigo, S.; Gil-García, E.; Aguilera-Castillo, O.; Pérez-Fernández, A.J.; Fernández-García, E. Factors Associated with In-Hospital Mortality in Acute Care Hospital Settings: A Prospective Observational Study. *Int. J. Environ. Res. Public Health* **2020**, *17*, 7951. [[CrossRef](#)]
6. Ak, O.; Batirel, A.; Ozer, S.; Çolakoğlu, S. Nosocomial infections and risk factors in the intensive care unit of a teaching and research hospital: A prospective cohort study. *Med. Sci. Monit. Int. Med. J. Exp. Clin. Res.* **2011**, *17*, PH29–PH34. [[CrossRef](#)]
7. Hydoub, Y.M.; Walker, A.P.; Kirchoff, R.W.; Alzu’bi, H.M.; Chipi, P.Y.; Gerberi, D.J.; Burton, M.C.; Murad, M.H.; Dugani, S.B. Risk Prediction Models for Hospital Mortality in General Medical Patients: A Systematic Review. *Am. J. Med. Open* **2023**, *10*, 100044. [[CrossRef](#)]
8. Tang, S.T. When death is imminent: Where terminally ill patients with cancer prefer to die and why. *Cancer Nurs.* **2003**, *26*, 245–251. [[CrossRef](#)]
9. Campbell, S.E.; Seymour, D.G.; Primrose, W.R.; Lynch, J.E.; Dunstan, E.; Espallargues, M.; Lamura, G.; Lawson, P.; Philp, I.; Mestheneos, E.; et al. A multi-centre European study of factors affecting the discharge destination of older people admitted to hospital: Analysis of in-hospital data from the ACMEplus project. *Age Ageing* **2005**, *34*, 467–475. [[CrossRef](#)]

10. Walicka, M.; Chlebus, M.; Śliwczyński, A.; Brzozowska, M.; Rutkowski, D.; Czech, M.; Tuszyńska, A.; Jacyna, A.; Puzianowska-Kuźnicka, M.; Franek, E. Predictors of in-hospital mortality in nonsurgical departments: A multivariable regression analysis of 2,855,029 hospitalizations. *Pol. Arch. Intern. Med.* **2020**, *130*, 268–275. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Silva, T.J.A.; Jerussalmy, C.S.; Farfel, J.M.; Curiati, J.A.E.; Jacob-Filho, W. Predictors of in-hospital mortality among older patients. *Clin. Sao Paulo Braz.* **2009**, *64*, 613–618. [\[CrossRef\]](#)
12. Calvillo-King, L.; Arnold, D.; Eubank, K.J.; Lo, M.; Yunyongying, P.; Stieglitz, H.; Halm, E.A. Impact of social factors on risk of readmission or mortality in pneumonia and heart failure: Systematic review. *J. Gen. Intern. Med.* **2013**, *28*, 269–282. [\[CrossRef\]](#)
13. Goltzman, G.; Perl, S.; Cohen, L.; Avivi, E.; Rapoport, M.J. Single Admission C-reactive protein Levels as a Sole Predictor of Patient Flow and Clinical Course in a General Internal Medicine Department. *Isr. Med. Assoc. J. IMAJ* **2019**, *21*, 686–691.
14. Loekito, E.; Bailey, J.; Bellomo, R.; Hart, G.K.; Hegarty, C.; Davey, P.; Bain, C.; Pilcher, D.; Schneider, H. Common laboratory tests predict imminent death in ward patients. *Resuscitation* **2013**, *84*, 280–285. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Wajner, A.; Zuchinali, P.; Olsen, V.; Polanczyk, C.A.; Rohde, L.E. Causes and Predictors of In-Hospital Mortality in Patients Admitted with or for Heart Failure at a Tertiary Hospital in Brazil. *Arq. Bras. Cardiol.* **2017**, *109*, 321–330. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Jayasundera, R.; Neilly, M.; Smith, T.O.; Myint, P.K. Are Early Warning Scores Useful Predictors for Mortality and Morbidity in Hospitalised Acutely Unwell Older Patients? A Systematic Review. *J. Clin. Med.* **2018**, *7*, 309. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Cheng, C.-Y.; Kung, C.-T.; Chen, F.-C.; Chiu, I.-M.; Lin, C.-H.R.; Chu, C.-C.; Kung, C.F.; Su, C.-M. Machine learning models for predicting in-hospital mortality in patient with sepsis: Analysis of vital sign dynamics. *Front. Med.* **2022**, *9*, 964667. [\[CrossRef\]](#)
18. Pourhomayoun, M.; Shakibi, M. Predicting mortality risk in patients with COVID-19 using machine learning to help medical decision-making. *Smart Health Amst. Neth.* **2021**, *20*, 100178. [\[CrossRef\]](#)
19. Silva, G.; Duarte, L.; Shirassu, M.; Peres, S.; de Moraes, M.; Chiavegatto Filho, A. Machine learning for longitudinal mortality risk prediction in patients with malignant neoplasm in São Paulo, Brazil. *Artif. Intell. Life Sci.* **2023**, *3*, 100061. [\[CrossRef\]](#)
20. Chen, Z.; Li, T.; Guo, S.; Zeng, D.; Wang, K. Machine learning-based in-hospital mortality risk prediction tool for intensive care unit patients with heart failure. *Front. Cardiovasc. Med.* **2023**, *10*, 1119699. [\[CrossRef\]](#)
21. Oliveira, M.; Seringa, J.; Pinto, F.J.; Henriques, R.; Magalhães, T. Machine learning prediction of mortality in Acute Myocardial Infarction. *BMC Med. Inform. Decis. Mak.* **2023**, *23*, 70. [\[CrossRef\]](#)
22. Guo, A.; Mazumder, N.R.; Ladner, D.P.; Foraker, R.E. Predicting mortality among patients with liver cirrhosis in electronic health records with machine learning. *PLoS ONE* **2021**, *16*, e0256428. [\[CrossRef\]](#)
23. Al-Yarabi, A.; Al Balushi, H.; Al Hatmi, K.; Al Yahyaie, R.; Al Alawi, A.M.; Al Zeedy, K.; Al Farhan, H. Inappropriate Hospital Stay of Patients Admitted Under Care of General Medicine Units. *Sultan Qaboos Univ. Med. J.* **2023**, *23*, 174–181. [\[CrossRef\]](#)
24. Efthimiou, O.; Hoogland, J.; Debray, T.P.A.; Aponte Ribero, V.; Knol, W.; Koek, H.L.; Schwenkglenks, M.; Henrard, S.; Egger, M.; Rodondi, N.; et al. Measuring the Performance of Survival Models to Personalize Treatment Choices. *Stat. Med.* **2025**, *44*, e70050. [\[CrossRef\]](#)
25. Raja, V.J.; Dhanamalar, M.; Solaimalai, G.; Rani, D.L.; Deepa, P.; Vidhya, R.G. Machine Learning Revolutionizing Performance Evaluation: Recent Developments and Breakthroughs. In *Proceedings of the 2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS)*; IEEE: Coimbatore, India, 2024; pp. 780–785.
26. Rainio, O.; Teuho, J.; Klén, R. Evaluation metrics and statistical tests for machine learning. *Sci. Rep.* **2024**, *14*, 6086. [\[CrossRef\]](#)
27. Clarke, J.; Gallifant, J.; Grant, D.; Desai, N.; Glover, G. Predictive value of the National Early Warning Score 2 for hospitalised patients with viral respiratory illness is improved by the addition of inspired oxygen fraction as a weighted variable. *BMJ Open Respir. Res.* **2023**, *10*, e001657. [\[CrossRef\]](#)
28. Hoeyer-Nielsen, A.K.; Holmberg, M.J.; Grossestreuer, A.V.; Yankama, T.; Berg, K.M.; Donnino, M.W. Oxygen Consumption as an Early Warning System in the Critically Ill. *Shock* **2023**, *60*, 539–544. [\[CrossRef\]](#)
29. Davis, R.B.; Iezzoni, L.I.; Phillips, R.S.; Reiley, P.; Coffman, G.A.; Safran, C. Predicting in-hospital mortality. The importance of functional status information. *Med. Care* **1995**, *33*, 906–921. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Laville, S.M.; Couturier, A.; Lambert, O.; Metzger, M.; Mansencal, N.; Jacquelinet, C.; Laville, M.; Frimat, L.; Fouque, D.; Combe, C.; et al. Urea levels and cardiovascular disease in patients with chronic kidney disease. *Nephrol. Dial. Transplant. Off. Publ. Eur. Dial. Transpl. Assoc. Eur. Ren. Assoc.* **2022**, *38*, 184–192. [\[CrossRef\]](#)
31. Manu, P.; Asif, M.; Khan, S.; Ashraf, H.; Mani, A.; Guvenek-Cokol, P.; Lee, H.; Kane, J.M.; Correll, C.U. Risk factors for medical deterioration of psychiatric inpatients: Opportunities for early recognition and prevention. *Compr. Psychiatry* **2012**, *53*, 968–974. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Yu, W.; Wen, Y.; Shao, Y.; Hu, T.; Wan, X. Relationship between anion gap and in-hospital mortality in intensive care patients with liver failure: A retrospective propensity score matching analysis. *Am. J. Transl. Res.* **2024**, *16*, 98–108. [\[CrossRef\]](#)
33. Alevrakakis, E.; Papadakis, D.-D.; Vagionas, D.; Koutsoukou, A.; Pontikis, K.; Rovina, N.; Vasileiadis, I. Strong ion gap and anion gap corrected for albumin and lactate in patients with sepsis in the intensive care unit. *Int. J. Physiol. Pathophysiol. Pharmacol.* **2024**, *16*, 10–27. [\[CrossRef\]](#) [\[PubMed\]](#)

34. Breen, T.J.; Padkins, M.; Bennett, C.E.; Anavekar, N.S.; Murphy, J.G.; Bell, M.R.; Barsness, G.W.; Jentzer, J.C. Predicting 1-Year Mortality on Admission Using the Mayo Cardiac Intensive Care Unit Admission Risk Score. *Mayo Clin. Proc.* **2021**, *96*, 2354–2365. [\[CrossRef\]](#)
35. Moore, E.; Munoz-Arroyo, R.; Schofield, L.; Radley, A.; Clark, D.; Isles, C. Death Within 1 Year Among Emergency Medical Admissions to Scottish Hospitals: Incident Cohort Study. *BMJ Open* **2018**, *8*, e021432. [\[CrossRef\]](#)
36. Oliynyk, O.V.; Rorat, M.; Barg, W. Oxygen metabolism markers as predictors of mortality in severe COVID-19. *Int. J. Infect. Dis. IJID Off. Publ. Int. Soc. Infect. Dis.* **2021**, *103*, 452–456. [\[CrossRef\]](#)
37. Lucey, K.; Jones, R.C.; Watson, J.A.; Malakooti, M.; Stephen, R.J. Risk Factors for Deterioration Events in the Pediatric Acute Care Setting. *Hosp. Pediatr.* **2024**, *14*, e260–e266. [\[CrossRef\]](#)
38. Arques, S.; Roux, E.; Sbragia, P.; Gelisse, R.; Pieri, B.; Ambrosi, P. Usefulness of serum albumin concentration for in-hospital risk stratification in frail, elderly patients with acute heart failure. Insights from a prospective, monocenter study. *Int. J. Cardiol.* **2008**, *125*, 265–267. [\[CrossRef\]](#)
39. Tojek, K.; Kowalczyk, G.; Czerniak, B.; Banaś, W.; Szukay, B.; Korzycka-Wilińska, W.; Banaszkiewicz, Z.; Budzyński, J. Blood albumin as a prognostic factor among unselected medically treated inpatients. *Biomark. Med.* **2019**, *13*, 1059–1069. [\[CrossRef\]](#)
40. Turcato, G.; Zaboli, A.; Sibilio, S.; Fanni Canelles, M.; Rella, E.; Giudiceandrea, A.; Pfeifer, N.; Brigo, F. Prognostic Role of Serum Albumin in Predicting 30-Day Mortality in Patients with Infections in Emergency Department: A Prospective Study. *J. Clin. Med.* **2023**, *12*, 3447. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Parthasarathi, A.; Padashetti, V.C.; Padukudru, S.; Chaya, S.K.; Siddaiah, J.B.; Anand, M.P. Association of Serum Albumin and Copeptin with Early Clinical Deterioration and Instability in Community-Acquired Pneumonia. *Adv. Respir. Med.* **2022**, *90*, 323–337. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Pandey, A.; Patel, K.V.; Ayers, C.; Tang, W.H.W.; Fang, J.C.; Drazner, M.H.; Berry, J.; Grodin, J.L. Temporal association between hospitalization event and subsequent risk of mortality among patients with stable chronic heart failure with preserved ejection fraction: Insights from the TOPCAT trial. *Eur. J. Heart Fail.* **2019**, *21*, 693–695. [\[CrossRef\]](#)
43. Nichols, G.A.; Reynolds, K.; Kimes, T.M.; Rosales, A.G.; Chan, W.W. Comparison of Risk of Re-hospitalization, All-Cause Mortality, and Medical Care Resource Utilization in Patients With Heart Failure and Preserved Versus Reduced Ejection Fraction. *Am. J. Cardiol.* **2015**, *116*, 1088–1092. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Huang, T.; Le, D.; Yuan, L.; Xu, S.; Peng, X. Machine Learning for Prediction of in-Hospital Mortality in Lung Cancer Patients Admitted to Intensive Care Unit. *PLoS ONE* **2023**, *18*, e0280606.
45. Taylor, R.A.; Pare, J.R.; Venkatesh, A.K.; Mowafi, H.; Melnick, E.R.; Fleischman, W.; Hall, M.K. Prediction of In-hospital Mortality in Emergency Department Patients With Sepsis: A Local Big Data-Driven, Machine Learning Approach. *Acad. Emerg. Med. Off. J. Soc. Acad. Emerg. Med.* **2016**, *23*, 269–278. [\[CrossRef\]](#)
46. Mahajan, P.; Uddin, S.; Hajati, F.; Moni, M.A. Ensemble Learning for Disease Prediction: A Review. *Healthc. Basel Switz.* **2023**, *11*, 1808. [\[CrossRef\]](#) [\[PubMed\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.