MDPI

Article

An Open-Access Data Set of Hospitalized Cardiac Arrest Patients: Machine Learning-Based Predictions Using Clinical Documentation

Lahiru Rajapaksha [1,‡,*], Sugandima Vidanagamachchi [1,‡], Sampath Gunawardena [2,‡] and Vajira Thambawita [3,‡,*]

[1] Department of Computer Science, University of Ruhuna, Matara, Sri Lanka;
[2] Faculty of Medicine, University of Ruhuna, Karapitiya, Sri Lanka.
[3] SimulaMet,Oslo, Norway.
* Correspondence: ltwrajapaksha@gmail.com (L.R); vajira@simula.no (V.T)
‡ These authors contributed equally to this work.

Abstract: Cardiac arrest is a sudden loss of heart function with serious consequences. In developing counties, healthcare professionals use clinical documentation to track patient information. These data can be used to predict developing cardiac arrest. To contribute to the advancement of the research domain, we are publishing the data set through open access. Based on this data set our work revolves around generating and utilizing synthetic data by harnessing the potential of synthetic data vaults. We conducted a series of experiments by employing state-of-the-art machine learning techniques. These experiments were aimed to assess the performance of our developed predictive model in identifying the likelihood of developing a cardiac arrest. This approach is effective in identifying the risk of cardiac arrest in inpatients, even in the absence of electronic medical recording systems. The study evaluated 112 patients who were transferred from the Emergency Treatment Unit to the Cardiac Medical Ward. The developed model achieved 96% accuracy in predicting the risk of developing cardiac arrest. In conclusion, our study showcases the potential of leveraging clinical documentation and synthetic data to create robust predictive models for cardiac arrest. The outcome of this effort will provide valuable insights and tools for healthcare professionals to preemptively address this critical medical condition.

Keywords: Bed Head Ticket; Cardiac Arrest; Clinical Documents; Decision Tree Classification Model; Early Warning System; Deep Learning; Developing Country; Machine Learning; Recurrent Neural Network (RNN)

Citation: Rajapaksha, L.; Vidanagamachchi, S.; Gunawardena, S.; Thambawita, V; Title. *Journal Not Specified* **2023**, *1*, 0. https://doi.org/

Received:
Revised:
Accepted:
Published:

1. Introduction

Cardiac arrest often occurs due to dysfunction in the heart's conduction system, which can cause the heart to stop pumping blood. Ventricular fibrillation is the most common cause (65-80% of cases) [1]. Various heart-related causes include coronary artery disease, cardiomyopathy, inherited conditions, congenital heart disease, heart valve disease, acute myocarditis, and conduction disorders like long QT syndrome. Many patients may not recognize or ignore symptoms, with chest pain being the most common.

The ischemic coronary illness causes over 70% of cardiac arrests and is the leading cause of cardiac dysfunction. Risk factors include hypertension, hyperlipidemia, diabetes, smoking, age, and family history of coronary diseases [2]. In Sri Lanka, Cardiovascular Diseases (CVD) account for a high mortality rate of 534 deaths per 100,000 [3] and 22.64% of total deaths [4]. Cardiac arrest risk is 20-30% within the first 24 hours after a heart attack. Survival rates are around 25%, even with proper medical treatment.

The time to receive help and treatment is crucial for survival in sudden cardiac arrest cases.Early Warning System (EWS) and can help identify deteriorating patients at risk of death from cardiac arrest [5]. Patients who experienced sudden cardiac arrest and were

admitted to the Intensive Care Unit (ICU) showed signs of deterioration hours before the event. Early detection and treatment can reduce mortality rates, and Early Warning System (EWS) is designed based on patients' vital signs to aid in this process [6]. This study makes several important contributions by,

- Publishing an Open Access bedhead ticket dataset.
- Introducing a machine learning model to predict the risk of fatal cardiac arrests, which has shown improved results.
- Analysing the data set with machine learning models to compare the usability of the dataset.

Repository links

- Dataset(Zenodo)
- Model(GitHub)

2. Materials and Methods

2.1. Methodology

The research was designed as a retrospective cohort study, focusing on patients admitted to the cardiac ward between August 13th, 2018, and February 6th, 2020, at Teaching Hospital, Karapitiya (THK), Galle, Sri Lanka. The study population included patients who were transferred to the cardiac ward from the Emergency Treatment Unit (ETU), excluding other types of patients and pediatric patients. A total of 112 patients, aged 15-89 years, were included in the study, with 82 male patients and 30 female patients.

Ethical clearance was obtained from the Ethics Review Committee of the Faculty of Medicine, University of Ruhuna, Galle, Sri Lanka, and the study adhered to the relevant guidelines and regulations established by the committee. Permission to access data was granted by the Director of Teaching Hospital Karapitiya. As this was a retrospective study, obtaining informed consent from all subjects was not applicable.

Data collection involved extracting information from the Bed Head Ticket (BHT) in the hospital's record room. Due to the absence of electronic clinical data, each BHT was manually examined to collect the necessary data. The BHT contained information on the patient's health status, clinical history, management actions, investigations, treatments, progress, and diagnosis.

The extracted BHT data were categorized into five main categories: demographic features, examinations, lab reports, patient history, and outcomes (Fig. 1). The extracted features were further analyzed to determine their availability concerning patients' hospital stays. Features with an availability rate of over 60% (Fig. 2) were selected for model development.

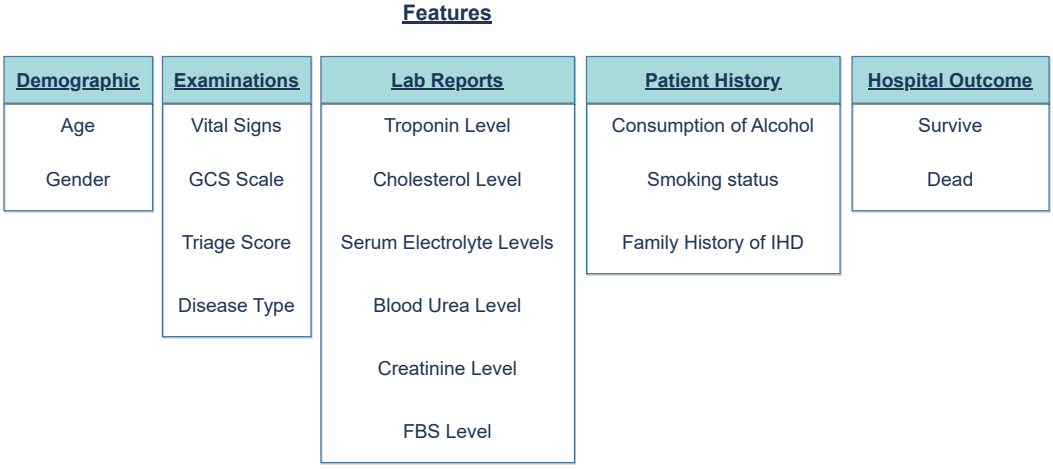


Figure 1. *Categorization of extracted features into groups*

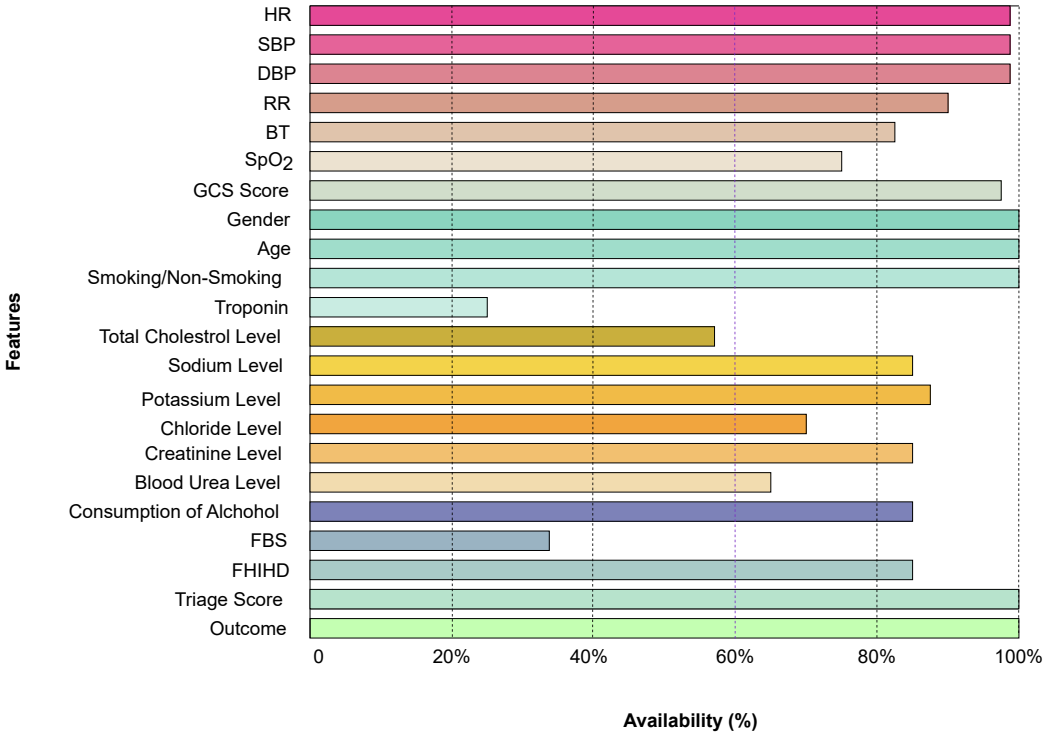


Figure 2. *Feature availability of 112 patients*

2.2. Data preprocessing

As illustrated in Fig. 3, the extracted clinical data were primarily divided into two categories: time-series data and non-time-series data. Time-series data consisted of patient information recorded in relation to time, while non-time-series data encompassed the remaining data. The clinical dataset contained 21 features: age, heart rate (HR), Systolic Blood Pressure (SBP), Diastolic Blood Pressure (DBP), Respiratory Rate (RR), Body Temperature (BT), Oxygen Saturation (SpO2), level of consciousness, troponin level, total cholesterol level, Fasting Blood Sugar level (FBS), serum electrolytes (Sodium, Potassium, Chloride), urea level, creatinine level, triage score (risk score on admission), alcohol consumption, smoking status, Family History of Ischemic Heart Diseases (FHIHD), and hospital outcome.

Out of the 21 features, 19 were selected for inclusion, while three were excluded from the final feature list. Troponin level, total cholesterol level, and fasting blood sugar level, which can be used to detect comorbidities in patients, were excluded due to insufficient records in the BHTs for the majority of patients. Data preprocessing consisted of two steps. First, 19 common features, including the patient's hospital outcome (target variable), were selected. Second, missing value imputation was performed for the selected features. If any feature data was missing, the most recent value was used; if no value was available, the median value was used.

Regarding time-series data (SBP, DBP, HR, RR, BT, SpO2), patients were monitored hourly throughout their hospital stay. In the selected data set, patients were monitored for a minimum of one hour and a maximum of 266 hours (11 days). A suitable time step (time window) for observation was chosen from this range. The time step was determined based on the average observation time of a patient (52 hours). This 52-hour observational time window is considered the prediction window of the model. For non-time-series data, only the results from the patient's lab tests conducted on the hospital admission date were considered.

2.3. Data Division and Preparation

After preprocessing the data set, we partitioned the data set into distinct three subsets to the ratio of 8:1:1, namely the training set, testing set, and validation set. Then the training

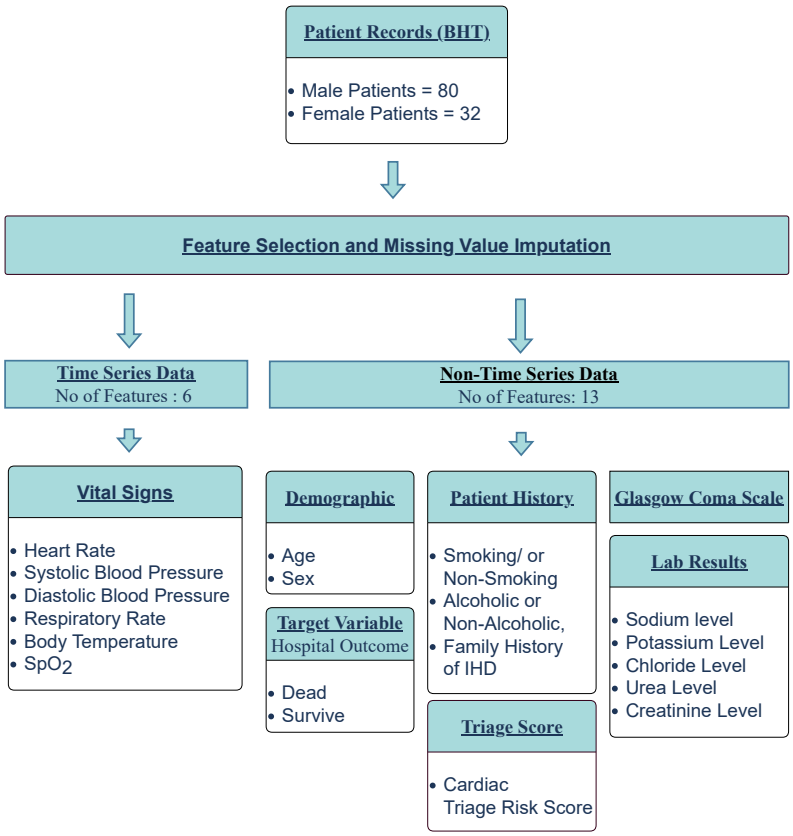


Figure 3. *Data preprocessing workflow*

data set is further enhanced through the infusion of synthetic data generated by Synthetic Data Valult (SDV) (Figure 4). The goal of the SDV is to build generative models of relational databases [14]. By incorporating its capability of creating artificial instances that mimic the statistical properties of the original data set we expanded our training data, fostering greater diversity and aiding the model's ability to capture complex patterns. When generating the

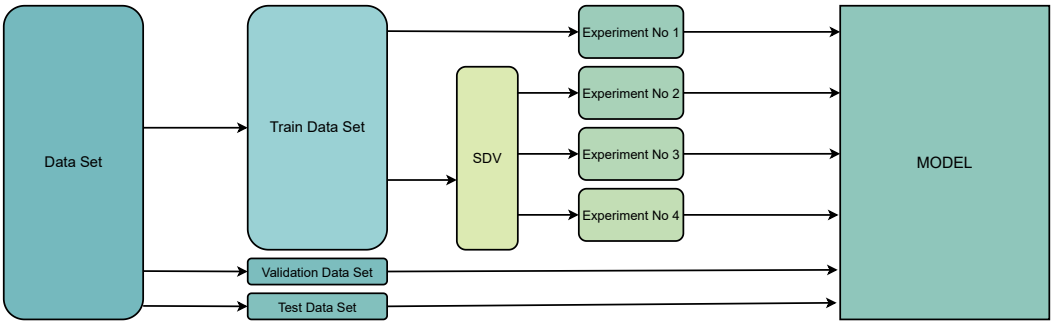


Figure 4. *Data Division and preparation for the model*

synthetic data for experiments no 2-4, our consistent aim is to achieve a balanced ratio of survived patient data to dead patient data at 1:1, ensuring equal representation(Table 1).

2.4. Model architecture

The integration of artificial intelligence into clinical practice has led to numerous studies focused on predicting adverse events, such as cardiac arrests before they occur. Evidence from these studies suggests that deep learning models are more efficient at identifying high-risk patients compared to existing EWSs [15]. Recurrent Neural Network (RNN) models are particularly well-suited for handling temporal sequence data. Among RNN variants, Long Short-Term Memory (LSTM) has demonstrated remarkable performance in various sequence-based tasks [16]. Recent studies have also shown that deep learning

Table 1. Summary of the number of patient records and sequences contains in training data sets. Except for experiment No 01, we maintained the survived: dead ratio as 1:1

	Training Data Set 01		Training Data Set 02		Training Data Set 03		Training Data Set 04	
	No of Sequences	No of Records	No of Sequences	No of Records	No of Sequences	No of Records	No of Sequences	No of Records
Real Survived Patients Data	4287	74	4287	74	0	0	4287	74
Synthetic Survived Patients Data	0	0	0	0	4870	74	4870	74
Real Dead Patients Data	743	15	743	15	0	0	743	15
Synthetic Dead Patients Data	0	0	1768	59	4702	74	5259	133
Total	5030	89	6798	148	9572	148	15159	296

models employing RNN architectures with LSTM outperform clinical prediction models developed using logistic regression [17,18]. Consequently, we chose the LSTM structure to model the temporal relationships within data extracted from the BHT.

2.4.1. LSTM model

The Deep-Learning Cardiac Arrest Prediction Model (DLCAPM) consists of a LSTM designed to handle time-series data, followed by a dense layer. During the model's development, we employed a sigmoid activation function for the dense layer, the Adam optimizer with default parameters, and binary-cross entropy as the loss function. The complete data set was divided into 30% for validation and 70% for training. Inputs to the LSTM model include Systolic Blood Pressure (SBP), Diastolic Blood Pressure (DBP), Heart Rate (HR), Respiratory Rate (RR), Body Temperature (BT), and Oxygen Saturation (SpO2) levels. The input data which was fed into the LSTM comprised a three-dimensional array of 112 × 52 × 6 (5824 patient record instances), as illustrated in Fig 5. The first phase of the model aimed to handle temporal data, while the second phase combined the results of the time-series data with non-time-series data. This combined data was then input into the decision tree model to predict the final outcome based on the model's results.

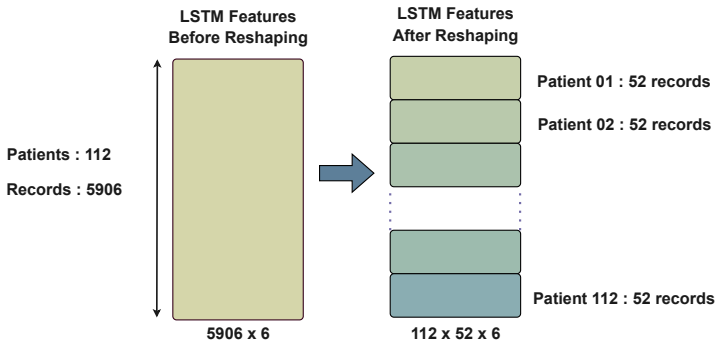


Figure 5. *Data reshaping*

2.4.2. Decision tree model

In the medical domain, the clinical practice involves continuous decision-making, where optimal decision-making strives to maximize effectiveness and minimize loss [19]. Clinical decision analysis (CDA) highlights the significant role of decision trees in this process [20]. Among the five methodologies employed for decision-making, designing a decision tree is considered one [21]. Owing to their reliability, effectiveness, and high accuracy in decision-making, decision trees are widely utilized in various medical decision-making studies [22]. These factors led us to select the decision tree as the model to handle non-time-series data, including the LSTM outcome. The LSTM model is designed to generate a risk score based on time-series data acquired within a time window of up to 52 hours.

2.4.3. Latent vector space

Latent vector space is employed in machine learning to analyze data that can be mapped to a latent space where similar data points are in close proximity. In simpler terms, it can be described as a representation of compressed data. This latent space representation retains all the essential information required to represent the original data points, thereby

facilitating data analysis. Latent space representations are utilized to transform more complex forms of raw data, such as images and videos, into simpler representations, a concept implemented in Representation Learning.

In the context of the LSTM model, the latent vector space serves as an input parameter for the decision tree model. The decision tree inputs are a combination of non-time-series data, including demographic information, lab results, triage score, Glasgow Coma Scale (GCS) values, and patient history, as well as the latent output of the LSTM model (Fig 6). This approach allows for a more comprehensive analysis by integrating both time-series and non-time-series data, enhancing the model's decision-making capabilities.

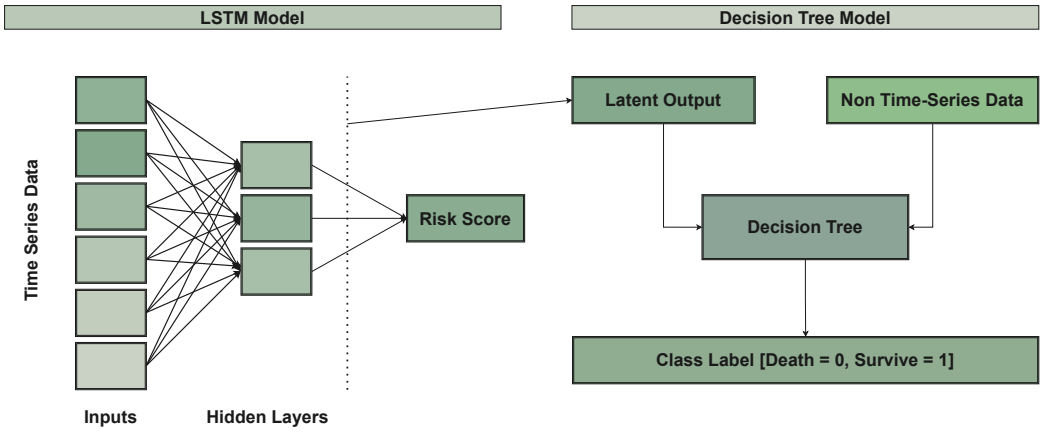


Figure 6. *Model architecture*

2.4.4. Dealing with Data Imbalance Problem

Class imbalance is a common challenge in real-world data, particularly in medical fields, making it difficult to optimize machine learning algorithm performance. In this study, there were 93 majority-class (survived) and 19 minority-class (dead) patients, yielding a 1:5 ratio.

To address the class imbalance, Synthetic Minority Over-sampling Technique (SMOTE) was used first to over-sample the minority class, generating synthetic data and achieving a 1:1 class ratio. But as was recently demonstrated by Van den Goorbergh, SMOTE does not improve discrimination but does lead to models with strong miscalibration [24] we moved forward with the SDV to address the data imbalance problem.

3. Results

In this result section, we present the results of a series of comprehensive experiments conducted to assess the performance of the Deep Learning Cardiac Arrest Prediction Model. To enhance the robustness of our experiments, we incorporated SDV to generate synthetic data, permitting us to thoroughly assess the predictive capabilities of the model. (Table 2)

Table 2. Experiments performed by incorporating synthetic data to evaluate the results of the combined model. RS = Real Survived, RD = Real Dead, SS = Synthetic Survived, SD = Synthetic Dead. The survived: dead ratio of the patient was maintained at 1:1.

Experiment No	RS	RD	SS	SD	Accuracy
01	74	15	-	-	0.954
02	74	15	-	59	0.954
03	-	-	74	74	0.964
04	74	15	74	133	0.967

After executing the designed experiments we were able to achieve peak accuracy through experiment 4 where the experiment encompassed a data set of 296 patient records. Hyperparameters were optimized according to the (Table 3)

Table 3. Hyperparameter values

No of Epochs	Learning Rate	Batch Size	LSTM Nodes	Optimizer
100	0.001	10	2	Adam

Creatinine levels were the most critical predictor, followed by sodium and blood urea levels. Studies suggest that these markers relate to renal function, which is linked to cardiovascular diseases [23]. Additionally, potassium levels, FHIHD, and age played key roles in classifying cardiac patients, supporting the association between cardiovascular diseases and decreased potassium levels [25].

FHIHD is recognized as a well-established risk factor for cardiovascular diseases [26]. In the study's findings, the risk factor FHIHD emerged as one of the most reliable predictive features for cardiac arrests. (Fig 7) displays the probabilistic prediction window for six randomly selected patients (three from each of the two classes). The LSTM model itself was proficient in providing predictions from the time of admission up to 52 hours later. In other words, the model's prediction window spanned 52 hours. The model demonstrated a prediction accuracy of 96% with a confidence interval of 95.01% to 95.85%.

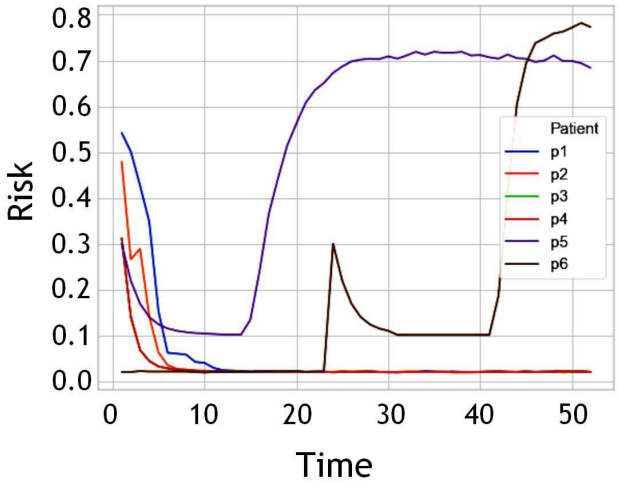


Figure 7. *Prediction of the risk score with 52 hours*

3.1. Comparison with existing models

We assessed the performance of the developed Deep Learning Cardiac Arrest Prediction Model (DLCAPM) by comparing it with well-established machine learning algorithms, such as Logistic Regression, Random Forest, Naïve Bayes, and Support Vector Machine (SVM). The evaluation was carried out using all the features incorporated in the DLCAPM model. Table 4 presents the performance metrics, including Accuracy, Sensitivity, Specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), and F-Score, for each of the compared models. The results demonstrate that the LSTM component of the DLCAPM model exhibits superior performance in comparison to the selected models.

3.2. Correlation analysis

Correlation analysis evaluates the relationships among model input features. Pearson's and Spearman's coefficients are commonly used; the former is for normally distributed variables and the latter for skewed or ordinal variables. A coefficient near ±1 indicates a strong correlation, either positive or negative [27]. In this study, Spearman's rank correlation was used due to Gaussian distribution, and Fig 8 shows the heatmap of correlation coefficients. A high positive correlation exists between Systolic and Diastolic Blood Pressure.

Table 4. Performance comparison with existing machine learning models (LSTM & Decision Tree are the two models of Deep Learning Cardiac Arrest Prediction Model[DLCAPM])

Model	Accuracy	Sensitivity	Specificity	PPV	NPV	F-Score
LSTM	**0.96**	**0.95**	**0.93**	**0.98**	**0.81**	**0.86**
Decision Tree	**0.76**	**0.69**	**0.81**	**0.72**	**0.79**	**0.80**
SVM	0.89	0.84	0.82	0.81	0.82	0.844
Logistic Regression	0.88	0.93	0.81	0.92	0.81	0.87
Random Forest	0.88	0.89	0.90	0.87	0.81	0.91
Naive Bayes	0.85	0.89	0.80	0.82	0.88	0.91

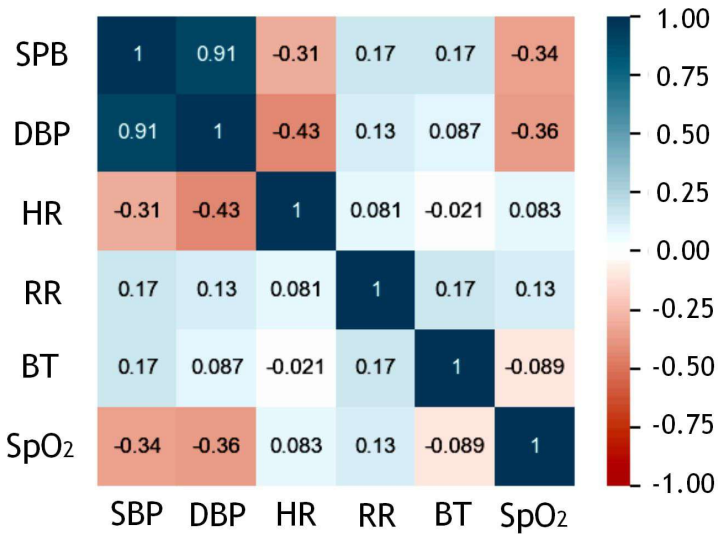


Figure 8. *Correlation heatmap between time-series inputs*

3.3. Characteristics of the study population

In the cohort study, patient data were analyzed to assess the observed characteristics of the study population, as shown in Table 5. From these observations, we can infer that males may have a higher susceptibility to cardiovascular diseases and cardiac arrests. This study also examined the impact of various risk factors that could potentially contribute to the development of CVDs. Among these risk factors, alcohol consumption, smoking, and a family history of ischemic heart disease (FHIHD) were identified as the most significant contributors. In the total population (comprising both males and females), 37% of individuals reported having an FHIHD. Notably, none of the female patients were documented to consume alcohol or smoke. Among the male patients, 73% (81 patients) were found to engage in at least one of these behaviors, with 69% (55 patients) consuming alcohol, 63% (51 patients) smoking, and 26% (21 patients) engaging in both alcohol consumption and smoking.

4. Discussion

The performance of the model was assessed using two optimization algorithms, 'Adam' and 'Admax'. Multiple iterations were performed, altering hyperparameter combinations to identify the four most optimal configurations, which yielded the highest accuracy. Table 6 presents the evaluation metrics for each of these four selected runs. Among these experiments, the best results were achieved in experiment number 04 for training the model.

Table 7 shows the sensitivity, specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), and accuracy measures for the outperformed LSTM model and decision tree model in experiment number 04, respectively.

The ROC curve is a graphical representation that demonstrates the diagnostic ability of binary classifiers by plotting sensitivity against specificity. A better-performing classifier

Table 5. Characteristics of the study population

Characteristic	Data
Study period	13th of August 2018 - 6th of February 2020
Hospital	Teaching Hospital Karapitiya, Galle, Sri Lanka
Total patients, n	112
Input vectors, n	19
Age group	59 – 76years
Male, n (%)	73%
Symptoms before admission	Chest pain on the left side (1/2 hour before the admission), Tightening of the chest, Vomiting, Sweating, Nausea, Cough, Fever
Patients with FHIHD, n (%)	37%
Consume alcohol, Male n (%)	69%
Smoking, Male n (%)	63%
Smoking & use alcohol, Male n (%)	26%

Table 6. Performance of the model for four experimental data sets. Here, we tested the LSTM and the Decision tree model separately

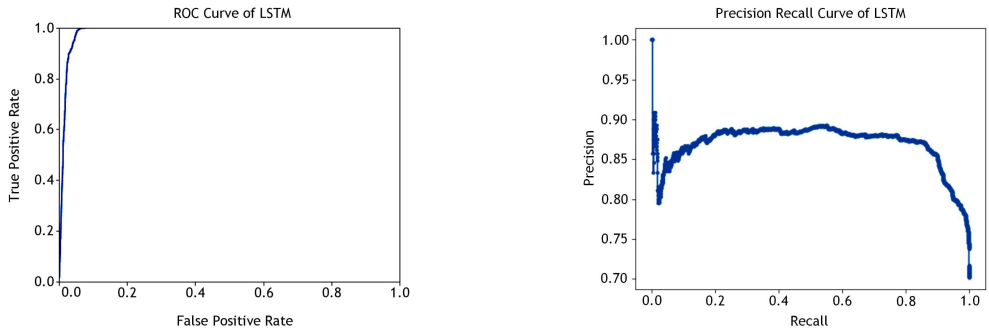
Metric	LSTM				Decision Tree			
	Experiment number				Experiment number			
	01	02	03	**04**	01	02	03	**04**
Accuracy	0.93	0.85	0.94	**0.96**	0.80	0.83	0.80	**0.76**
Precision	1.0	1.0	1.0	**1.0**	1.0	1.0	1.0	**1.0**
Recall	1.0	1.0	1.0	**1.0**	1.0	1.0	1.0	**10**
F-Score	0.81	0.86	0.81	**0.86**	0.82	0.85	0.82	**0.80**
AUC score	0.97	0.95	0.97	**0.98**	0.79	0.83	0.79	**0.75**

Table 7. Evaluation metrics of LSTM and Decision Tree models of the Deep Learning Cardiac Arrest Prediction Model

Statistic	Value		95% CI	
	LSTM Model	Decision Tree Model	LSTM Model	Decision Tree Model
Sensitivity	95.83%	69.57%	95.37% to 96.25%	47.08% to 86.79%
Specificity	93.42%	81.82%	92.07% to 94.61%	64.54% to 93.02%
Positive Predictive Value	98.71%	72.73%	98.44% to 98.93%	55.19% to 85.24%
Negative Predictive Value	81.04%	79.41%	79.37% to 82.60%	67.07% to 87.96%

will have a curve closer to the top-left corner. To compare classifiers' performance, a common approach is to calculate the area under the ROC curve. Figure 9a presents the ROC curve for the LSTM model.

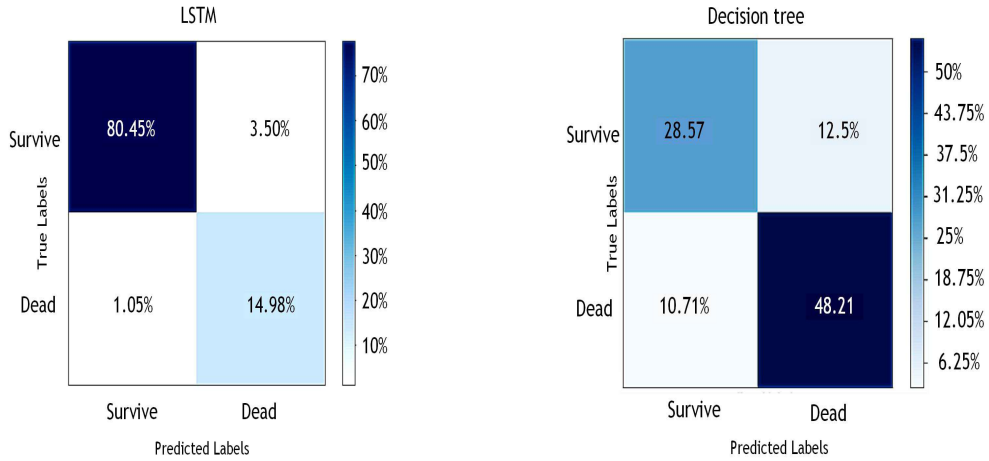
However, visual interpretations and comparisons of ROC curves can be misleading for imbalanced datasets. To address this issue, precision-recall curves are utilized. Figure 9b illustrates the precision-recall curve for the LSTM model, and the supplementary figure shows the curve for the decision tree classifier model.



(a) *ROC curve* **(b)** *Precision-Recall curve*

Figure 9. ROC curve and Precision-Recall curve of LSTM model

The confusion matrix, crucial for statistical classifications in machine learning, is a table describing a classification model's performance and identifying class confusion. Figure 10a and Figure 10b display the confusion matrices for the LSTM and decision tree models, respectively.



(a) *Confusion matrix for LSTM model* **(b)** *Confusion matrix for the decision tree model*

Figure 10. *Confusion matrices for (a) LSTM model and (b) decision tree model*

4.0.1. Comparison with existing EWS

DLCAPM was evaluated against existing cardiac arrest early warning scores, including Modified Early Warning Score (MEWS), Cardiac Arrest Risk Triage Score (CART), and National Early Warning Score (NEWS), based on research by [28,29]. MDCALC calculator [30–32] was used to calculate risk scores for CART, MEWS, and NEWS using collected data (RR, SpO2, BT, SBP, DBP, HR, Age, Triage Score). Table 8 displays the performance of each score.

Table 8. Performance comparison of DLCAPM(Deep Learning Cardiac Arrest Prediction Model: Combination of LSTM and Decision Tree) against CART(Cardiac Arrest Risk Triage Score), MEWS(Modified Early Warning Score) and NEWS(National Early Warning Score) models. Highlighted are the results of the DLCAPM.

Model	Accuracy	Sensitivity	Specificity	PPV	NPV	F-Score
LSTM	**0.96**	**0.95**	**0.93**	**0.98**	**0.81**	**0.86**
Decision Tree	**0.76**	**0.69**	**0.81**	**0.72**	**0.79**	**0.80**
CART	0.60	0.50	0.75	0.75	0.75	0.60
MEWS	0.80	0.93	0.40	0.82	0.4	0.50
NEWS	0.80	0.84	0.66	0.94	0.66	0.66

4.0.2. Limitations

The THK record room utilizes Microsoft Excel to store limited data from bedhead tickets, making it necessary to manually review each patient's record to obtain the required information for this study, which was time-consuming. The small sample size, potential patient heterogeneity, and focus on a specific patient population within a single hospital unit limit the research's generalizability, making it potentially inapplicable nationwide. These limitations stem from the legacy methods of maintaining patient data and difficulties in retrieval.

Due to limited resources, only 112 patient records were extracted, which prevented reaching the desired sample size for the model. To address this limitation, the Synthetic Data Vault Technique (SDV) was employed to increase the minority sample size. Furthermore, the study faced constraints due to the scarcity of local research on the development of cardiac arrest EWSs in Sri Lanka [8,10].

5. Conclusions

An efficient deep-learning cardiac risk prediction model was developed using clinical features from THK cardiac patients' BHTs. This simple model uses accessible patient data, offering bedside support for healthcare workers and assisting decision-making. An open-access dataset was published to encourage further research.

Early and accurate predictions can aid in timely interventions and prevent cardiac events. Despite high accuracy, addressing data limitations could improve the model. Further research is needed to explore applicability in other healthcare settings and integration with existing patient monitoring tools for enhanced prediction.

Author Contributions: L.R. Study concept and design; acquisition of data; analysis and interpretation of data; drafting of the manuscript; statistical analysis. S.V: Study concept and design; Critical revision of the manuscript; Drafting of the manuscript. V.T: Study concept and design; Critical revision of the manuscript; Drafting of the manuscript; Supervision. S.G: Study concept and design; Critical revision of the manuscript; Drafting of the manuscript; Supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Institutional Review Board Statement: The study was conducted in accordance with the Ethics Committee of Faculty of Medicine University of Ruhuna, Sri Lanka (protocol code 2019/P/063 and date of approval 27.11.2019)."

Informed Consent Statement: Since this study was carried out as a retrospective study, obtaining informed consent from all subjects was not applicable.

Data Availability Statement: The dataset used for all experiments is publicly available at https://zenodo.org/record/7603772

Acknowledgments: The authors express their gratitude for the support provided by Dr. Susitha Amarasinghe, Professor Sampath Gunawardena, Mrs. S.G.H. Uluwaduge, Mr. Bhashithe Abeysinghe, Dr. Samadhi Rajapaksha, and the Ethics Review Committee of the Faculty of Medicine, University of Ruhuna.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

LSTM	Long Short-Term Memory
CVD	Cardiovascular Diseases
EWS	Early Warning System
RRT	Rapid Response Teams
DRT	Dedicated Resuscitation Teams
LMIC	Low to Middle-Income Country
HIC	High-Income Countries
BHT	Bed Head Ticket
THK	Teaching Hospital, Karapitiya
ETU	Emergency Treatment Unit
HR	Heart Rate
SBP	Systolic Blood Pressure
DBP	Diastolic Blood Pressure
RR	Respiratory Rate
BT	Body Temperature
SpO2	Oxygen Saturation
FBS	Fasting Blood Sugar
FHIHD	Family History of Ischemic Heart Diseases
DLCAPM	Deep-Learning Cardiac Arrest Prediction Model
RNN	Recurrent Neural Network
CDA	Clinical Decision Analysis
SMOTE	Synthetic Minority Over-sampling Technique
GCS	Glasgow Coma Scale
SVM	Support Vector Machine
PPV	Positive Predictive Value
NPV	Negative Predictive Value
TPR	True Positive Rate
FPR	False Positive Rate
ROC	Receiver Operating Characteristic
MEWS	Modified Early Warning Score
CART	Cardiac Arrest Risk Triage Score
NEWS	National Early Warning Score
MD CALC	Medical Calculator
GRU	Gated Recurrent Unit
CNN	Convolutional Neural Networks
EMR	Electronic Medical Records

References

1. Tang W, Weil MH. Cardiac arrest and cardiopulmonary resuscitation. In: Critical Care Medicine. Elsevier; 2008. p. 3–15.
2. Beane A, Ambepitiyawaduge PDS, Thilakasiri K, Stephens T, Padeniya A, Athapattu P, et al. Practices and perspectives in cardiopulmonary resuscitation attempts and the use of do not attempt resuscitation orders: A cross-sectional survey in Sri Lanka. Indian J Crit Care Med [Internet]. 2017;21(12):865–8. Available from: http://dx.doi.org/10.4103/ijccm.IJCCM_314_17
3. Abeywardena MY. Dietary fats, carbohydrates and vascular disease: Sri Lankan perspectives. Atherosclerosis [Internet]. 2003;171(2):157–61. Available from: http://dx.doi.org/10.1016/s0021-9150(03)00157-6
4. Who.int. [cited 2022 Jun 26]. Available from: https://www.who.int/docs/default-source/gho-documents/world-health-statistic-reports/6-june-18108-world-health-statistics-2018.pdf
5. Ye C, Wang O, Liu M, Zheng L, Xia M, Hao S, et al. A real-time early warning system for monitoring inpatient mortality risk: Prospective study using electronic medical record data. J Med Internet Res [Internet]. 2019;21(7):e13719. Available from: http://dx.doi.org/10.2196/13719
6. Nishijima I, Oyadomari S, Maedomari S, Toma R, Igei C, Kobata S, et al. Use of a modified early warning score system to reduce the rate of in-hospital cardiac arrest. J Intensive Care [Internet]. 2016;4(1). Available from: http://dx.doi.org/10.1186/s40560-016-0134-7
7. Marinkovic O, Sekulic A, Trpkovic S, Malenkovic V, Pavlovic A. The importance of early warning score (EWS) in predicting in-hospital cardiac arrest—Our experience. Resuscitation [Internet]. 2013;84:S85. Available from: http://dx.doi.org/10.1016/j.resuscitation.2013
8. De Silva AP, Sujeewa JA, De Silva N, Rathnayake RMD, Vithanage L, Sigera PC, et al. A retrospective study of physiological observation-reporting practices and the recognition, response, and outcomes following cardiopulmonary arrest in a low-to-middle-income country. Indian J Crit Care Med [Internet]. 2017;21(6):343–5. Available from: http://dx.doi.org/10.4103/ijccm.IJCCM_136_17

9. Ranawaka UK, Wijekoon CN, Pathmeswaran A, Kasturiratne A, Gunasekera D, Chackrewarthy S, et al. Risk estimates of cardiovascular diseases in a Sri Lankan community. Ceylon Med J [Internet]. 2016;61(1):11–7. Available from: http://dx.doi.org/10.4038/cmj.v61i1.8253
10. Beane A, De Silva AP, De Silva N, Sujeewa JA, Rathnayake RMD, Sigera PC, et al. Evaluation of the feasibility and performance of early warning scores to identify patients at risk of adverse outcomes in a low-middle income country setting. BMJ Open [Internet]. 2018;8(4):e019387. Available from: http://dx.doi.org/10.1136/bmjopen-2017-019387
11. Smith GB, Prytherch DR, Schmidt PE, Featherstone PI, Higgins B. A review, and performance evaluation, of single-parameter "track and trigger" systems. Resuscitation [Internet]. 2008;79(1):11–21. Available from: http://dx.doi.org/10.1016/j.resuscitation.2008.05.004
12. Gerry S, Birks J, Bonnici T, Watkinson PJ, Kirtley S, Collins GS. Early warning scores for detecting deterioration in adult hospital patients: a systematic review protocol. BMJ Open [Internet]. 2017;7(12):e019268. Available from: http://dx.doi.org/10.1136/bmjopen-2017-019268
13. Kwon J-M, Lee Y, Lee Y, Lee S, Park J. An algorithm based on deep learning for predicting in-hospital cardiac arrest. J Am Heart Assoc [Internet]. 2018;7(13). Available from: http://dx.doi.org/10.1161/JAHA.118.008678
14. N. Patki, R. Wedge and K. Veeramachaneni, "The Synthetic Data Vault," 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), Montreal, QC, Canada, 2016, pp. 399-410, doi: 10.1109/DSAA.2016.49.
15. Kim J, Chae M, Chang H-J, Kim Y-A, Park E. Predicting cardiac arrest and respiratory failure using Feasible Artificial Intelligence with Simple Trajectories of patient data. J Clin Med [Internet]. 2019;8(9):1336. Available from: http://dx.doi.org/10.3390/jcm8091336
16. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. J Am Med Inform Assoc [Internet]. 2017;24(2):361–70. Available from: http://dx.doi.org/10.1093/jamia/ocw112
17. Ge W, Huh J-W, Park YR, Lee J-H, Kim Y-H, Turchin A. An interpretable ICU mortality prediction model based on logistic regression and Recurrent Neural Networks with LSTM units. AMIA Annu Symp Proc. 2018;2018:460–9.
18. Aczon M, Ledbetter D, Ho L, Gunny A, Flynn A, Williams J, et al. Dynamic mortality risk predictions in pediatric critical care using recurrent neural networks [Internet]. arXiv [stat.ML]. 2017. Available from: http://arxiv.org/abs/1701.06675
19. Aleem IS, Schemitsch EH, Hanson BP. What is a clinical decision analysis study? Indian J Orthop [Internet]. 2008;42(2):137–9. Available from: http://dx.doi.org/10.4103/0019-5413.40248
20. Bae J-M. The clinical decision analysis using decision tree. Epidemiol Health [Internet]. 2014;36:e2014025. Available from: http://dx.doi.org/10.4178/epih/e2014025
21. Myers J, McCabe SJ. Understanding medical decision making in hand surgery. Clin Plast Surg [Internet]. 2005;32(4):453–61, v. Available from: http://dx.doi.org/10.1016/j.cps.2005.05.001
22. Podgorelec V, Kokol P, Stiglic B, Rozman I. J Med Syst [Internet]. 2002;26(5):445–63. Available from: http://dx.doi.org/10.1023/a:1016409317
23. Kurniawan LB, Bahrun U, Mangarengi F, Darmawati ER, Arif M. Blood urea nitrogen as a predictor of mortality in myocardial infarction. Zinc suppl improv heme biosynth rats expo lead [Internet]. 2013 [cited 2022 Jun 26];32(3):172–8. Available from: https://univmed.org/ejurnal/index.php/medicina/article/view/91
24. R. van den Goorbergh, M. van Smeden, D. Timmerman, and B. Van Calster, "The harm of class imbalance corrections for risk prediction models: illustration and simulation using logistic regression," J. Am. Med. Inform. Assoc., vol. 29, no. 9, pp. 1525–1534, 2022, Available from: https://doi.org/10.1093/jamia/ocac093.
25. Kughapriya P, Ponnudhali D, Jones E. Evaluation of serum electrolytes in Ischemic Heart Disease patients [Internet]. Njbms.in. [cited 2022 Jun 26]. Available from: http://njbms.in/uploads/19/1564_pdf.pdf
26. Tan BY, Judge DP. A clinical approach to a family history of sudden death. Circ Cardiovasc Genet [Internet]. 2012;5(6):697–705. Available from: http://dx.doi.org/10.1161/CIRCGENETICS.110.959437
27. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi Med J. 2012;24(3):69–71.
28. Kim J, Park YR, Lee JH, Lee J-H, Kim Y-H, Huh JW. Development of a real-time risk prediction model for in-hospital cardiac arrest in critically ill patients using deep learning: Retrospective study. JMIR Med Inform [Internet]. 2020;8(3):e16349. Available from: http://dx.doi.org/10.2196/16349
29. Medlej K. National Early Warning Score (NEWS) [Internet]. MDCalc. [cited 2022 Jun 26]. Available from: https://www.mdcalc.com/national-early-warning-score-news
30. Medlej K. Modified Early Warning Score (MEWS) for clinical deterioration [Internet]. MDCalc. [cited 2022 Jun 26]. Available from: https://www.mdcalc.com/modified-early-warning-score-mews-clinical-deterioration
31. CART (Cardiac Arrest Risk Triage) Score [Internet]. MDCalc. [cited 2022 Jun 26]. Available from: https://www.mdcalc.com/cart-cardiac-arrest-risk-triage-score
32. van der Ploeg T, Austin PC, Steyerberg EW. Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints. BMC Med Res Methodol [Internet]. 2014;14(1):137. Available from: http://dx.doi.org/10.1186/1471-2288-14-137
33. Tonekaboni S, Laussen P, Eytan D, Greer R, Goodfellow SD, Goodwin A, et al. Prediction of cardiac arrest from physiological signals in the pediatric ICU [Internet]. Squarespace.com. [cited 2022 Jun 26]. Available from: https://static1.squarespace.com/static/59d5ac1780bd5ef9c396eda6/t/5b73738d88251beb6017ede8/1534343143656/Tonekaboni_S
34. Mehmood A, Iqbal M, Mehmood Z, Irtaza A, Nawaz M, Nazir T, et al. Prediction of heart disease using deep convolutional neural networks. Arab J Sci Eng [Internet]. 2021;46(4):3409–22. Available from: http://dx.doi.org/10.1007/s13369-020-05105-1

35. Li H, Wu TT, Yang DL, Guo YS, Liu PC, Chen Y, et al. Decision tree model for predicting in-hospital cardiac arrest among patients admitted with acute coronary syndrome. Clin Cardiol [Internet]. 2019;42(11):1087–93. Available from: http://dx.doi.org/10.1002/clc.23255