

Trade-off Analysis Between Temporal Window Durations and Performance of Patient Deterioration Predictive Models Based on Machine Learning

Kai Wei Tay

Department of Biomedical Engineering and Health Sciences, Faculty of Electrical Engineering,
Universiti Teknologi Malaysia

<https://doi.org/10.5109/7395508>

出版情報 : Proceedings of International Exchange and Innovation Conference on Engineering & Sciences (IEICES). 11, pp.131-136, 2025-10-30. International Exchange and Innovation Conference on Engineering & Sciences

バージョン :

権利関係 : Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International



Trade-off Analysis Between Temporal Window Durations and Performance of Patient Deterioration Predictive Models Based on Machine Learning

Kai Wei Tay¹, Asnida Abdul Wahab^{1,2}, Yuan Wen Hau^{1,3}, Gan Hong Seng⁴

¹Department of Biomedical Engineering and Health Sciences, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia.

²Medical Device and Technology Centre (MEDITEC), Institute of Human Centered Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia.

³IJN-UTM Cardiovascular Engineering Centre, Institute of Human Centered Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia.

⁴School of AI and Advanced Computing, Xián Jiaotong-Liverpool University, Suzhou, Jiangsu, P.R. China 215400
Corresponding author email: asnida.aw@utm.my

Abstract: *Intensive Care Units (ICUs) patients often have life-threatening diseases and rapidly changing health conditions. Timely response towards patient deterioration was found to increase the survival chances of patients. However, conventional early warning score (EWS) systems rely on predefined thresholds which only indicate the current risk of patients instead of predicting the future trend or risk. Previous studies proved that predictive models based on machine learning outperformed EWS systems. This study developed models based on logistic regression (LogReg) and support vector machine (SVM) to predict short-term mortality of ICU patients based on the data extracted from MIMIC-III database. Results obtained show that the trade-off analysis between observation window and prediction window as well as prediction performance of models were conducted. The models achieved optimum accuracy performance of 0.91 for LogReg and 0.92 for SVM at 4-hours observation window and 1-hour prediction window. The developed models demonstrate high potential in providing early warnings on patient deterioration and hence giving the medical team a longer response time and improving the survival chance of ICU patients.*

Keywords: Patient Deterioration; Mortality Prediction; ICU; MIMIC-III; Machine Learning

1. INTRODUCTION

Intensive Care Units (ICUs) patients often have life-threatening diseases and rapidly changing health conditions. According to the study by Antonio e Silva *et al.* [1], 537 out of 1,596 ICU patients died in the unit, indicating a mortality rate of 33.65%. Timely response towards patient deterioration was found to increase the survival chance of patients. The mortality rate of patients following cardiac arrest was 66% when the medical team responded in less than 3 minutes. Meanwhile, the mortality rate increased to 91% when the medical team responded in more than 3 minutes [2]. Previous studies also identified readmission as a key cause of ICU mortality [1], [3]. The study by Tejerina Álvarez [3] showed that readmitted patients generally had worse prognosis and higher mortality rate (i.e., 2 to 10 times higher) than others. Thus, it is important to identify patients at potential risk of deterioration to provide further medical support.

1.1 Early Warning Score Systems

A conventional approach in patient deterioration prediction is the early warning score (EWS) systems. The systems assigned scores to different clinical parameters based on predefined scoring range. However, EWS systems rely on predefined thresholds which only indicate the current risk of patients instead of predicting the future trend or risk. The systems are also unable to define complex patterns in clinical patient data. In this context, machine learning (ML) algorithms are applicable to identify and learn from the patterns in clinical data [4], [5]. Choi *et al.* [6] developed several predictive models using machine learning to predict in-

hospital mortality for ICU patients. The accuracy of the models ranged from 0.95 to 0.96. The study also included EWS systems as benchmarks and the accuracy of the systems ranged from 0.60 to 0.83 [6]. The study proved that ML-based predictive models outperform traditional EWS system in predicting patient deterioration.

1.2 Machine Learning Predictive Models

Based on the literature review and study by Jahandideh *et al.* [4], logistic regression (LogReg) and support vector machine (SVM) are among the most commonly used machine learning algorithms, especially in patient deterioration prediction. LogReg is a machine learning algorithm used for binary classification. It can be used to predict the probability between two different classes. It utilized a similar concept as linear regression (LogReg) but eventually maps the predicted values to a range between 0 and 1 using logistic function (i.e., sigmoid function) [7], [8]. The strengths of regression model were highlighted on low computation costs and easy interpretation of results [4]. Besides, SVM is a machine learning algorithm that utilizes hyperplanes to separate different classes. Support vectors are the closest point to the hyperplane for each class [8], [9]. Generally, SVM is widely used in classification and regression tasks. SVM possesses several strengths including low generalization error, low computation costs as well as less chance of overfitting [4].

There were previous studies that focused on predicting in-hospital mortality or cardiac arrest without a specific time frame [6], [10]. This resulted in the predicted deterioration events may onset any time within the hospitalization. Such consequences may be troublesome for the medical teams to decide a suitable intervention

manner. Choi *et al.* [6] extracted first 24 hours patient data to predict in-hospital mortality. J. H. Kim *et al.* in their study [10] developed ML models in predicting in-hospital cardiac arrest for emergency department patients. They identified that features such as oxygen saturation, vital signs, mental status and hospital factors were correlated with in-hospital cardiac arrest.

Other than that, other previous studies developed their prediction models with pre-defined prediction window [11], [12], [13], [14]. In this context, the onset of deterioration is within or exactly after the end of prediction window based on the definition in the studies. Alshwaheen *et al.* [11], [12] developed an LSTM-RNN model to predict short-term mortality of ICU patients and sudden transfer of normal ward patients to ICU. The authors explored the combination of observation window and prediction window of predictive models. The optimum model was found to have 578 minutes observation window and 286 minutes prediction window. Provided with a high accuracy of 0.921 and AUROC of 0.933, the model utilized a long observation window whereas 578 minute-by-minute data were needed for each feature. Moreover, the authors also developed LogReg and SVM as baseline models. The baseline models possessed comparable results with shorter observation window required (i.e., 240 minutes) [11], [12].

Another study by Ruiz *et al.* [13] utilized XGBoost, random forest and LogReg models to predict clinical deterioration for pediatric patients aged 6 months or younger. Clinical deterioration was defined as emergent endotracheal intubation, CPR, ECMO and ECPR in the study. The same study highlighted the exploration of prediction windows ranging from 1 hour to 8 hours

before deterioration. Among the developed models, LogReg model performed comparable with other models. The model possessed AUROC of 0.920 with a prediction window of 4 hours. The best model was XGBoost with AUROC of 0.923. However, the models utilized a large number of 1,028 features in order to achieve competitive performance [13].

Moghadam *et al.* [14] explored the predictive power of different ML models in hypotensive events. The study employed 7 ML algorithms, including regression model and SVM model. Based on the results, LogReg and SVM had the highest accuracy among the models (i.e., 0.95 and 0.94 respectively). The models only utilized a 5-minute prior patient data to predict hypotensive event's onset within the next 30 minutes. In contrast to their excellent accuracy performance, the models possessed lower sensitivities of 0.85 and 0.83 respectively. Besides, the onset of deterioration fell within the 30-minutes time frame in this study which was different from some previous studies [11], [12], [13].

2. METHODOLOGY

Fig. 1 outlines the overall study scope, comprising the database and tools employed. This study utilized data from Medical Information Mart for Intensive Care (MIMIC-III). The database is housed at PhysioNet which is a repository of medical research data managed by Massachusetts Institute of Technology (MIT) [15], [16]. The database is comprised of 53,423 distinct hospital admissions of adult patients and includes comprehensive data such as demographics, vital signs, laboratory and microbiology tests, medications, notes and reports [17].

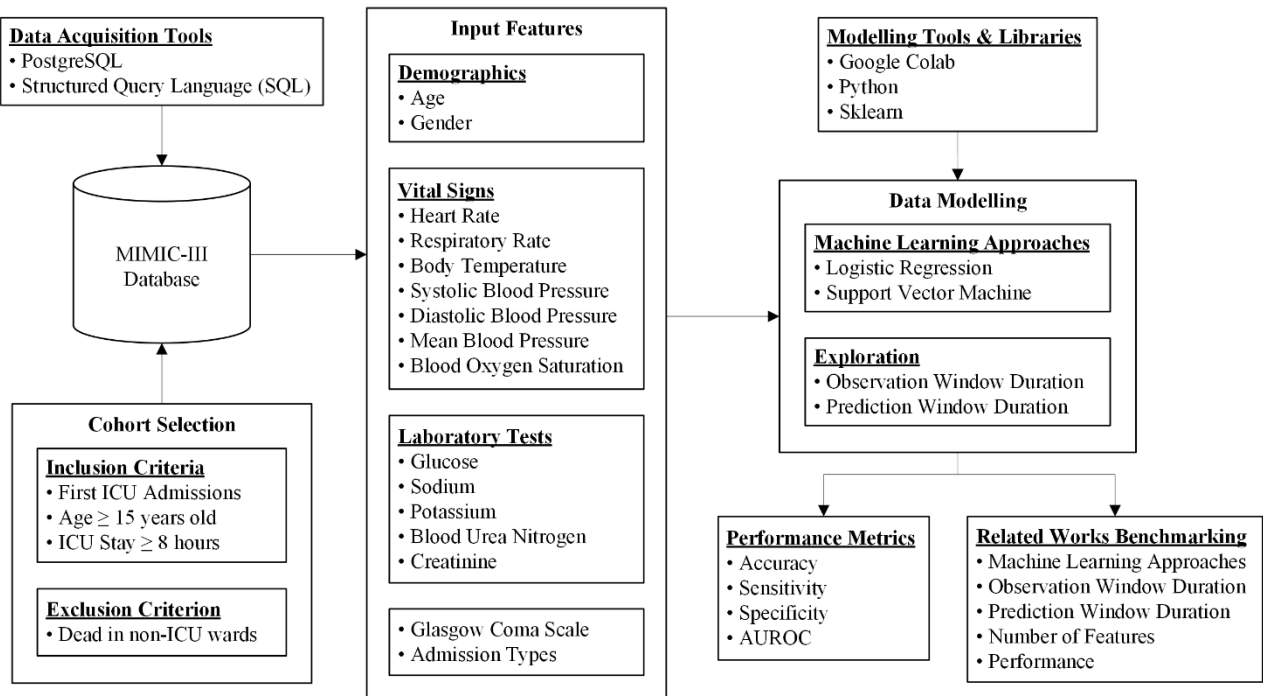


Fig. 1 Overall scope of study

Performance metrics used to evaluate the models developed include accuracy, sensitivity, specificity and AUROC. Accuracy, sensitivity (i.e., recall) and specificity are defined in Equation (1), Equation (2) and Equation (3). TP, TN, FP and FN are defined as True

Positive, True Negative, False Positive and False Negative respectively. Meanwhile, AUROC is the area under a receiver operating characteristics (ROC) curve [8], [18], [19].

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

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

Previous studies set their prediction goals in different aspects including in-hospital mortality, in-hospital cardiac arrest, septic shock, neurological deterioration, hypotensive events, CPR and ECMO [6], [10], [13], [14], [20], [21]. This study defined patient deterioration as short-term mortality whereas the models developed would predict binary classification of survival or mortality among ICU patients. This prediction goal was more general and could be applied to a wider range of patients to predict their survival status. The models would make predictions at the index time after collecting enough data in observation window. The predicted outcomes were based on the prediction window whereas the survival status of patients were based on the onset time. In other words, the models would predict mortality in advance of prediction window duration [5], [11], [12]. The overall timeframe is illustrated in Fig. 2 while Fig. 3 shows the overall workflow of this study from data preparation to model training and evaluation.

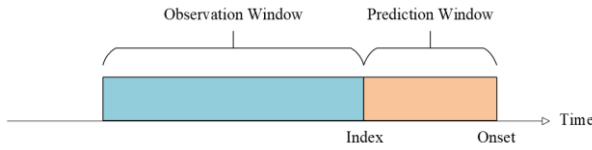


Fig. 2. Definition of windows in predictive models

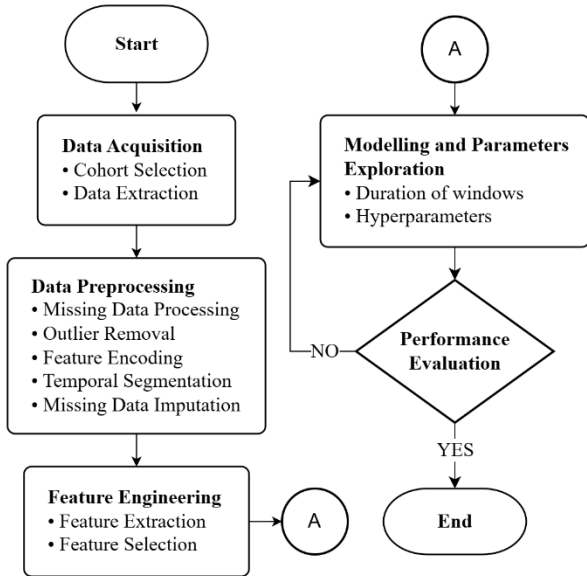


Fig. 3. Modelling flowchart

2.1 Data Acquisition

The first step of data acquisition was cohort selection. This study employed 3 inclusion criteria and 1 exclusion criteria in selecting patient cohort from MIMIC-III database. The inclusion criteria were (a) only first ICU admission of each patient was considered, (b) subjects of aged 15 years and above and (c) the ICU stay had to be

at least 8 hours and above. The exclusion criterion was the removal of any subjects died in non-ICU ward.

In this study, relevant clinical data were systematically extracted from the patient database. This included demographic information (age and gender), admission type, vital signs, laboratory test results, and level of consciousness. The vital signs comprised of heart rate, respiratory rate, body temperature, systolic blood pressure, diastolic blood pressure, mean blood pressure and blood oxygen saturation (SPO₂). The laboratory tests included glucose, sodium, potassium, blood urea nitrogen (BUN) and creatinine. The level of consciousness extracted was the Glasgow Coma Scale (GCS). All raw data was extracted in a 7-hours observation window and 1-hour prediction window setting to accommodate the window explorations of the study. The raw dataset would consist of 7 hours minute-by-minute data with 1 hour prior to the onset time. The onset time referred to mortal time for case group (i.e., label 1) or discharge time for control group (i.e., label 0).

2.2 Data Preprocessing

The missing rate of features (i.e., vital signs, laboratory tests and GCS) existed in each subject were calculated. In this context, subjects with missing rate higher than 20% were excluded from the dataset [6]. Furthermore, the major group identified was the control group (i.e., survive until discharge) while the minor group was the case group (i.e., die within the ICU admission). To address the class imbalance problem, the major group was down sampled to produce a dataset with equal number of classes (i.e., 825 each). In terms of outlier removal, SPO₂ and GCS values were checked for their valid range which are 0% to 100% and scale of 3 to 15 respectively. Meanwhile, the outlier removal of other vital signs and laboratory tests were based on the Equation (4) and Equation (5) [13]. Q1, Q3 and IQR were defined as first quartile, third quartile and interquartile range. Values beyond the lower bound or upper bound were eliminated.

$$Lower\ Bound = Q1 - 1.5 * IQR \quad (4)$$

$$Upper\ Bound = Q3 + 1.5 * IQR \quad (5)$$

Categorical features such as gender and admission type were encoded with numerical values. MIMIC-III database groups patients aged 89 years and above into categories of 300 years old and above for deidentification purposes [17]. In this study, the age of this group of subjects was shifted to 90 years old to keep the feature's ordinal order. Next, the 7-hours dataset was segmented into 11 datasets to accommodate the exploration of observation window and prediction window durations. The explored observation window and prediction window durations ranged from 1 hour to 5 hours and 1 hour to 4 hours respectively. The distinct combinations of window durations were presented along with their prediction performance in result and discussion section. After segmentation, the missing values within the dataset were filled using linear interpolation and followed by nearest fill. Median value of each feature was imputed only if the feature value was missing for the entire observation window.

2.3 Feature Engineering

For each laboratory test, body temperature and GCS, 2 statistical features (i.e., maximum and standard deviation value) were extracted. Meanwhile, 5 statistical features (i.e., minimum, maximum, standard deviation, last value and rate of change) were extracted from the other rapid-changing vital signs. After the feature extraction, these 47 engineered features were standardized and evaluated using Pearson correlation coefficient (i.e., Equation (6)). C and μ are the correlation coefficient and mean value respectively [22]. In this context, only the top 70% (i.e., 33) engineered features with the highest correlation with target (i.e., survival status) were selected.

$$C = \frac{\sum_{i=1}^N [(x_i - \mu_x)(y_i - \mu_y)]}{\sqrt{\sum_{i=1}^N (x_i - \mu_x)^2 \sum_{i=1}^N (y_i - \mu_y)^2}} \quad (6)$$

2.4 Modelling Pipeline

The dataset was split into a training set and testing set with a ratio of 80:20. Afterwards, the dataset was ready to be trained into predictive models using LogReg and SVM [9], [23]. The hyperparameters were tuned using Grid Search with 5-fold cross validation [8], [13], [19]. A trade-off analysis between window durations and prediction performance was presented in the next section.

3. RESULTS AND DISCUSSION

The prediction performance of each model was presented in Table 1 (i.e., LogReg models) and Table 2 (i.e., SVM models). The accuracy performance of each model was illustrated in Fig. 4. Based on the results, LogReg and SVM models shared comparable performance under similar window durations. However, SVM generally showed a slightly better performance than LogReg. Both models generally possessed higher specificity than sensitivity which was in line with previous studies [14]. This indicated a better performance in predicting patients' survival than death.

Moreover, the extension of observation window from 1 hour to 4 hours gradually improved the prediction performance. Longer observation window provided more meaningful data to be interpreted by the models. Besides, shortening of prediction window from 4 hours to 1 hour generally improved the prediction performance. Datasets with shorter prediction windows might contain more early signs of patient deterioration. This condition was found by previous studies [11], [12] as well. In this context, LogReg and SVM models showed optimum performance at a 4-hours observation window and 1-hour prediction window (i.e., 0.91 and 0.92 accuracy respectively).

Table 1. Prediction performance based on LogReg model

Observation Window (Hour)	Prediction Window (Hour)	Accuracy	Sensitivity	Specificity	AUROC
1	1	0.76	0.69	0.84	0.89
2	1	0.85	0.83	0.86	0.93
2	2	0.85	0.84	0.87	0.92
3	1	0.89	0.88	0.91	0.95
3	2	0.88	0.85	0.90	0.93
3	3	0.86	0.82	0.90	0.93
4	1	0.91	0.89	0.93	0.96
4	2	0.88	0.86	0.89	0.95
4	3	0.87	0.84	0.90	0.94
4	4	0.89	0.88	0.90	0.94
5	1	0.91	0.87	0.95	0.97

Table 2. Prediction performance based on SVM model

Observation Window (Hour)	Prediction Window (Hour)	Accuracy	Sensitivity	Specificity	AUROC
1	1	0.76	0.65	0.87	0.89
2	1	0.87	0.85	0.89	0.95
2	2	0.85	0.84	0.87	0.91
3	1	0.91	0.92	0.90	0.97
3	2	0.88	0.87	0.90	0.93
3	3	0.88	0.88	0.89	0.93
4	1	0.92	0.92	0.93	0.97
4	2	0.91	0.92	0.90	0.95
4	3	0.85	0.83	0.87	0.93
4	4	0.90	0.90	0.90	0.93
5	1	0.92	0.92	0.93	0.97

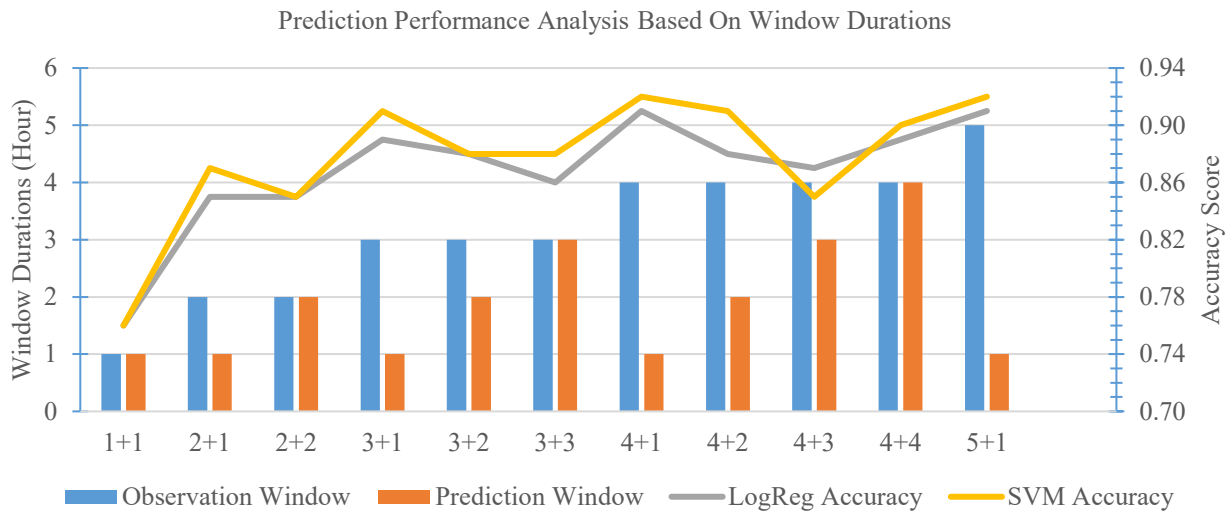


Fig. 4. Prediction performance analysis based on window durations

In addition to the optimum 4-hours observation window and 1-hour prediction window combination, other window combinations were found valuable as well. The predictive model with a 3-hours observation window and 1-hour prediction window showed a competitive accuracy of 0.91. This model was valuable in terms of its shorter observation window. Models with shorter observation windows were applicable to patients who need urgent justification for health conditions. The 4-hours observation window and 4-hours prediction window model also showed a comparable accuracy of 0.90. Although the model's accuracy was slightly lower than the optimum model (i.e., 0.92), its advantage was that it provided a longer prediction time frame. This allowed earlier interventions by the medical teams to save the patients. In this context, septic shock and cardiac arrest are among the critical causes of mortality which require continuous monitoring and timely intervention. Rapid and early predictions of mortality provide higher survival rates by enabling timely interventions, especially for ICU patients suffering from severe illnesses such as sepsis [2], [24].

[6]

4. CONCLUSION AND RECOMMENDATIONS

This study demonstrated the trade-off analysis between observation window and prediction window durations as well as prediction performance. Performances of the developed models were found optimum at 4-hours observation window and 1-hour prediction window. SVM model showed higher accuracy and sensitivity of 0.92, specificity of 0.93 and AUROC of 0.97. The developed models demonstrate high potential in providing early warnings on patient deterioration and hence giving the medical team a longer response time and improving the survival chance of ICU patients. Recommendations on future research include benchmarking with EWS systems as well as employing multi-source datasets and advanced data processing steps.

5. REFERENCES

- [1] L. G. Antonio e Silva, C. M. Dantas de Maio Carrilho, T. B. Talizin, L. T. Q. Cardoso, E. L. Lavado, and C. M. C. Grion, "Risk factors for hospital mortality in intensive care unit survivors: a retrospective cohort study," *Acute and Critical Care*, vol. 38, no. 1, pp. 68–75, Feb. 2023, doi: 10.4266/ACC.2022.01375.
- C. Sandroni *et al.*, "In-hospital cardiac arrest: survival depends mainly on the effectiveness of the emergency response," *Resuscitation*, vol. 62, no. 3, pp. 291–297, Sep. 2004, doi: 10.1016/J.RESUSCITATION.2004.03.020.
- E. E. Tejerina Álvarez, K. A. Gómez Mediavilla, C. Rodríguez Solís, N. Valero González, and J. Á. Lorente Balanza, "Risk factors for readmission to ICU and analysis of intra-hospital mortality," *Medicina Clínica (English Edition)*, vol. 158, no. 2, pp. 58–64, Jan. 2022, doi: 10.1016/J.MEDCLE.2020.11.038.
- S. Jahandideh, G. Ozavci, B. W. Sahle, A. Z. Kouzani, F. Magrabi, and T. Bucknall, "Evaluation of machine learning-based models for prediction of clinical deterioration: A systematic literature review," *Int J Med Inform*, vol. 175, p. 105084, Jul. 2023, doi: 10.1016/J.IJMEDINF.2023.105084.
- S. Muralitharan *et al.*, "Machine Learning-Based Early Warning Systems for Clinical Deterioration: Systematic Scoping Review," *J Med Internet Res*, vol. 23, no. 2, p. e25187, Feb. 2021, doi: 10.2196/25187.
- M. H. Choi *et al.*, "Mortality prediction of patients in intensive care units using machine learning algorithms based on electronic health records," *Sci Rep*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-11226-4.
- X. Zou, Y. Hu, Z. Tian, and K. Shen, "Logistic Regression Model Optimization and Case Analysis," in *2019 IEEE 7th International Conference on Computer Science and Network Technology (ICCSNT)*, IEEE, Oct. 2019, pp. 135–139. doi: 10.1109/ICCSNT47585.2019.8962457.
- R. Krishnamoorthy, K. Tanaka, and R. Thiagarajan, "Analysis of Obese Patients Using Machine Learning to Categorize Hidden Risk Factors in Explorative Assessment," in *Proceedings of International Exchange and Innovation Conference on Engineering & Sciences (IEICES)*, Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Oct. 2024, pp. 685–691. doi: 10.5109/7323336.
- M. Awad and R. Khanna, "Support Vector Machines for Classification," in *Efficient Learning Machines*, Berkeley, CA: Apress, 2015, pp. 39–66. doi: 10.1007/978-1-4302-5990-9_3.

- [10] J. H. Kim, A. Choi, M. J. Kim, H. Hyun, S. Kim, and H. J. Chang, "Development of a machine-learning algorithm to predict in-hospital cardiac arrest for emergency department patients using a nationwide database," *Sci Rep*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022[24] 26167-1.
- [11] T. I. Alshwaheen, "Patient Deterioration Predictive Model using Long Short-Term Memory Recurrent Neural Network with Genetic Algorithm Optimization," Universiti Teknologi Malaysia, 2021.
- [12] T. I. Alshwaheen, Y. W. Hau, N. Ass'Ad, and M. M. Abualsamen, "A Novel and Reliable Framework of Patient Deterioration Prediction in Intensive Care Unit Based on Long Short-Term Memory-Recurrent Neural Network," *IEEE Access*, vol. 9, pp. 3894–3918, 2021, doi: 10.1109/ACCESS.2020.3047186.
- [13] V. M. Ruiz *et al.*, "Early prediction of clinical deterioration using data-driven machine-learning modeling of electronic health records," *J Thorac Cardiovasc Surg*, vol. 164, no. 1, pp. 211–222.e3, Jul. 2022, doi: 10.1016/J.JTCVS.2021.10.060.
- [14] M. C. Moghadam, E. M. K. Abad, N. Bagherzadeh, D. Ramsingh, G. P. Li, and Z. N. Kain, "A machine-learning approach to predicting hypotensive events in ICU settings," *Comput Biol Med*, vol. 118, p. 103626, Mar. 2020, doi: 10.1016/J.COMPBIOMED.2020.103626.
- [15] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. 215–220, Jun. 2000, doi: 10.1161/01.CIR.101.23.e215.
- [16] A. Johnson, T. Pollard, and R. Mark, "MIMIC-III Clinical Database (version 1.4)," *PhysioNet*, 2016, doi: 10.13026/C2XW26.
- [17] A. E. W. Johnson *et al.*, "MIMIC-III, a freely accessible critical care database," *Sci Data*, vol. 3, no. 1, p. 160035, May 2016, doi: 10.1038/sdata.2016.35.
- [18] R. Guido, S. Ferrisi, D. Lofaro, and D. Conforti, "An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review," *Information*, vol. 15, no. 4, p. 235, Apr. 2024, doi: 10.3390/INFO15040235.
- [19] Veianthan Jayaramu, Zed Zulkafli, and Fariq Rahmat, "Spatio-Temporal Prediction of Leptospirosis in Kuantan Using Hydrometeorological Variables and Random Forest Machine Learning," in *Proceedings of International Exchange and Innovation Conference on Engineering & Sciences (IEICES)*, Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Oct. 2024, pp. 916–922. doi: 10.5109/7323369.
- [20] S.-H. Kim *et al.*, "Interpretable machine learning for early neurological deterioration prediction in atrial fibrillation-related stroke," *Sci Rep*, vol. 11, no. 1, p. 20610, 2021, doi: 10.1038/s41598-021-99920-7.
- [21] R. Li, J. K. Agor, and O. Y. Özaltın, "Temporal pattern mining for knowledge discovery in the early prediction of septic shock," *Pattern Recognit*, vol. 151, p. 110436, Jul. 2024, doi: 10.1016/J.PATCOG.2024.110436.
- [22] Y. Kim, T. H. Kim, and T. Ergün, "The instability of the Pearson correlation coefficient in the presence of coincidental outliers," *Financ Res Lett*, vol. 13, pp. 243–257, May 2015, doi: 10.1016/J.FRL.2014.12.005.
- [23] M. Sharma, Dr. S. K. Agarwal, and Dr. M. Bundeale, "DECISIVE ANALYSIS OF MULTIPLE LOGISTIC REGRESSION APROPOS OF HYPER-PARAMETERS," *Indian Journal of Computer Science and Engineering*, vol. 13, no. 1, pp. 188–196, Feb. 2022, doi: 10.21817/indjcse/2022/v13i1/221301190.
- A. Kumar *et al.*, "Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock," *Crit Care Med*, vol. 34, no. 6, pp. 1589–1596, Jun. 2006, doi: 10.1097/01.CCM.0000217961.75225.E9.