



Pipelined Ensemble Architecture for Mortality Prediction on MIMIC III

D. Malarvizhi¹, A. Prakash²

¹Regional Assistant Professor, Department of Computer Science, Dr. N.G.P. Arts and Science College Coimbatore

²Professor, Department of Computer Science, Hindusthan College of Arts & Science Coimbatore

*Corresponding author's: D. Malarvizhi

Article History	Abstract
Received: 06 June 2023 Revised: 05 Sept 2023 Accepted: 06 Dec 2023	<p>Automated healthcare decision support has seen a huge rise with the improved data collection models in hospitals and also improved machine learning based techniques that exhibit high possibilities for automation. Automating healthcare systems as an aide for clinical practitioners can ensure fast and more accurate results for the patients and can also aid in hospital administration. Mortality prediction has been one of the major and critical factors that determines the type of treatment and the level of resources that has to be allocated for a patient. This work presents a pipelined ensemble architecture that can be used for effective prediction of mortality levels of a patient. The pipeline model has been designed in multiple levels to ensure improvement of quality of the medical data and effective prediction. The pipelined architecture model has been compared with existing state-of-the-art model, and the results indicate high performance with 92% accuracy levels, ensuring the model is suitable for use in real time mortality prediction.</p>
CC License CC-BY-NC-SA 4.0	Keywords: Mortality Prediction; Machine Learning; Pipelining; Ensemble Learning; Healthcare Decision Support

1. Introduction

Intensive Care Unit (ICU) is one of the vital locations that has to be provided with continuous monitoring facilities, appropriate amount of manpower and other critical resources. Early and appropriate treatments are vital to critical patients admitted in ICUs. It is necessary to provide appropriate manpower and resources, however, it is also mandatory to avoid overcrowding of technicians [1]. It becomes mandatory to identify the state of the patient and also to identify the probable progress level of the patient to ensure they are provided with appropriate facilities. These factors are usually of extreme importance to patient in ICU. Several factors and their complicated associations are to be considered to identify the state of a patient, in specific to determine the probability of mortality. During the initial days scoring systems were used to determine the mortality levels [2, 3]. The major challenges that exist in using clinical data for predictions are, data heterogeneity introduced because of the large number of departments and clinical experts involved for treating a single patient [4], high complexity of data, large size of the data [5] and data sparsity. Further, imbalance also plays a critical role, as most of the patients are healthy during discharge. Improvement in electronic healthcare systems has facilitated collection of medical data, which can be effectively used for analysis.

Electronic health records can be used for automated prediction of mortality levels [6]. Current advancement in machine learning techniques have made automated predictions more appropriate. Machine learning is the process of deriving complex correlations among factors from the data to provide predictions [7, 8]. Clinical data can very highly benefit from the improvement in machine learning techniques [9]. The data is huge and relatively high complexity levels occurs between the dependent and independent variables; hence manual deciphering is impossible. The recent years have witnessed the rising need for clinical decision support systems [10]. Increased use of machine learning based techniques for decision support in clinical systems has been observed in the recent years. These decisions are usually used as reference points for the medical staff to proceed with further treatments [11].

This work presents a pipelined ensemble architecture that is used for mortality prediction. The architectural pipeline includes a data preprocessing model, a feature selection model, multiple machine

learning ensemble models and a combiner model to provide the final results. The model has been facilitated to handle highly complex large sized data. The obtained results exhibit high performance indicating that the model is suitable for deployment in real time.

Related works

Mortality prediction has been one of the major requirements of the current automated healthcare systems. The process is however highly complicated because of the large number of records used for the prediction process. Current and the most significant researches in this domain are discussed in this section.

A machine learning based model that uses gradient boosted machines (GBM) for mortality prediction has been presented by Li et al. [12]. This work uses light GBM for mortality prediction. The work uses features derived from physiological signs of the patient and laboratory tests to build the training data. An artificial neural network based on mortality prediction has been presented by Ding et al. [13]. Data extraction phase identifies and extracts the significant and relevant features. This ensures that high quality data is passed to the neural network model. The neural network model automatically performs feature analysis in its architecture further fine tuning the process to ensure high quality results. An ensemble technique that uses prognostic scores for mortality prediction has been presented by Selcuk et al. [14]. This work includes Box-Cox and Min-Max transformations to improve the efficiency of the ensemble-based prediction technique.

A neural network-based mortality prediction model that deals with sampling methods and feature selection process has been presented by Steinmeyer et al. [15]. The model mainly concentrates on improving the reproducibility and also ensures generalizability of the model to bigger and more generic clinical datasets for predicting the mortality levels. This work presents an extensive processing pipeline and also multiple evaluating techniques for mortality prediction. Feature based analysis has been identified to be one of the major techniques followed in the mortality prediction process. A heart data-based mortality prediction model that focuses on minimal number of clinical variables has been presented by Sadeghi et al. [16]. A temporal based model that predicts mortality during admission has been presented by Veith et al. [17]. A neural network-based model that applies self-normalizing for mortality prediction has been presented by Zahid et al. [18].

A deep learning model to forecast mortality has been presented by Harerimana et al. [19]. This model has been designed to predict both Length of Stay (los) and also mortality rates. It is a two-level prediction model, that uses a multi-model architecture for the prediction process. Several key features that are available during admission are used for the prediction process. A length of stay prediction model to identify the stay levels of new-born during admission has been presented by Thompson et al. [20]. Administrative data plays a major role in deriving the training data for this model. The length of stay prediction model using Hidden Markov Models (HMM) was presented by Sotoodeh et al. [21]. This work considers vital signs of patients monitored during the initial 48 hours after admission to formulate the training data. An ICU based mortality prediction model that uses hierarchical logistic regression has been presented by Moser et al. [22]. The model performs feature selection and feature augmentation by generating physiology scores to improve the prediction process.

A dynamic ensemble to perform mortality prediction based on ICU data has been presented by Guo et al. [23]. It is an ensemble selection model that uses multiple ensemble models and additional logic to determine the best ensemble model to be used for the data. An extreme learning machine-based mortality prediction model has been presented by Krishnan et al. [24]. This work is focused on handling large scale data for the prediction process. An analysis of healthcare process and its impact on in-hospital mortality has been presented by Mandalapu et al. [25]. The model concentrates on identifying the factors that contribute to in-hospital mortality during weekends.

Pipelined Ensemble Architecture for Mortality Prediction (PEAMP)

Mortality prediction generally requires analysis of huge amount of data obtained from multiple departments in a hospital. The huge nature of data and the high complexity results in the need for a complex architecture that can handle the data effectively. This work presents a pipelined architecture that encompasses ensemble-based machine learning techniques to ensure high quality predictions. The proposed architecture has been designed in two major phases; the data preprocessing phase and the pipelined ensemble architecture.

Data Preprocessing

Data preprocessing is the initial phase that is used to clean the data for the pipelined ensemble architecture. The input clinical data is obtained from MIMIC III, which is composed of multiple tables of data obtained from various departments in the hospital. Each table corresponds to a single department and contains details about patients pertaining to the department. In order to obtain a comprehensive view of the patient, the tables have to be analyzed and integrated according to the requirements of the prediction domain. Manual feature analysis is performed to identify the features that are required for mortality prediction. This process is followed by integration of multiple tables containing the necessary features. Standard feature argumentation-based techniques are applied to the data to ensure elimination of missing data and noisy values. Date based features are a common occurrence in hospital-based data. As these features are not directly usable in machine learning models, they are converted to numerical data using date-based operations. Further, hospital records have been identified to contain huge number of categorical features. The categorical entities are converted to numerical entities by applying one hot encoding techniques. However, applying one hot encoding results in a huge increase in the number of features, resulting in a data that is huge and highly complex.

Pipelined Ensemble for Mortality Prediction

The pipelined ensemble model has been designed to handle the highly complex nature of the data in a systematic manner to build an effective prediction model. The preprocessed training data is passed to the data pipeline for feature selection, data training and final prediction integration.

Model based Feature Selection

The initial process in the pipelined model is feature selection. Feature selection is the process of analyzing and identifying features that exhibit a significant impact during the prediction process. Several features in the data might not exhibit any impact on the final predictions, while some features exhibit negative impacts on the final prediction. Such features need to be detected and eliminated to improve the quality of the prediction process. Further, eliminating unnecessary features will also result in the reduction of data size. Huge data size is one of the major issues faced while handling medical data. Large number of features also result in the curse of dimensionality, which reduces prediction quality to a large extent.

This work follows model-based feature selection, which performs feature selection based on the significance of features identified using a machine learning model. The machine learning model is used as a meta transformer. The data is applied to the model, based on the predictions, weights for features are identified. Features exhibiting weights that fall below a certain threshold are eliminated. This work follows mean based thresholding, which uses the mean value of obtained weight values as the threshold value. Logistic regression is used for feature selection.

Multi Model Integrated Training

Model training is the next process in the proposed pipeline. This process is divided into two stages, which can be performed in parallel or sequential manner. Artificial Neural Networks (ANN) is used in the initial stage of the pipeline. The network model has been constructed using 100 neurons in the hidden layer. Relu is used as the activation function. Further, ANN models perform intrinsic feature selections, hence resulting in highly effective predictions. Random forest is used in the next stage of pipeline to improve the prediction efficiency. Random forest is a tree-based algorithm that uses decision tree as it's base algorithm of choice. Multiple instances of decision trees are generated to create a single random forest model. A distinct subset of the training data is passed to each instance of the decision tree. After the completion of the training process, the multiple decision trees are integrated into a single model. Tree pruning is applied to ensure optimized tree construction and prediction. The ANN model uses entire data for prediction; however, it performs internal feature selection. The random forest model uses a subset of the data. These factors ensure that the proposed model can effectively handle data imbalance and noisy entities.

Multiple Prediction Integration

The previous sections of the pipeline use two machine learning models for the prediction process. The test data is passed through the same pipeline and is predicted using both the trained machine learning models. This results in generation of two predictions rather than a single prediction. The two predictions are combined using mean based combiner to produce the final prediction set. The mean-based combiner forms the last phase of the pipeline. The resultant data is passed to the user as the final prediction.

3. Results and Discussion

Performance of the pipelined mortality prediction model has been analyzed using the MIMIC III data set. A classification report presenting a class-based comparison of precision, recall and F1-Score is shown in table 1. Class 0 corresponds to patient who was discharged and Class 1 corresponds to the patient whose mortality was recorded. It could be observed from the table that the proposed algorithm exhibits high precision levels in predicting mortality, and moderate recall levels. The F1-Score which is an aggregated performance metric shows 92% performance for class 0 and 94% performance for Class 1. The overall accuracy level has been observed to be 93%. These factors depict the efficiency of PEAMP model in effectively identifying mortality levels.

Table 1: Classification Report

	Precision	Recall	F1-Score
0	0.85	1	0.92
1	1	0.88	0.94
Accuracy			0.93

The ROC curve representing sensitivity and specificity levels is shown in figure 1. Sensitivity levels represent efficiency of the model in identifying positive classes, which represents mortality and specificity levels represent efficiency of the model in identifying negative classes. High performance in both the metrics exhibit and overall effective classification model. The curve indicates high sensitivity and specificity levels of the PEAMP model. Further, the curve representing PEAMP model occupies a larger area compared to the model presented by Ding et al. [13] these factors indicate efficient performance by the PEAMP model.

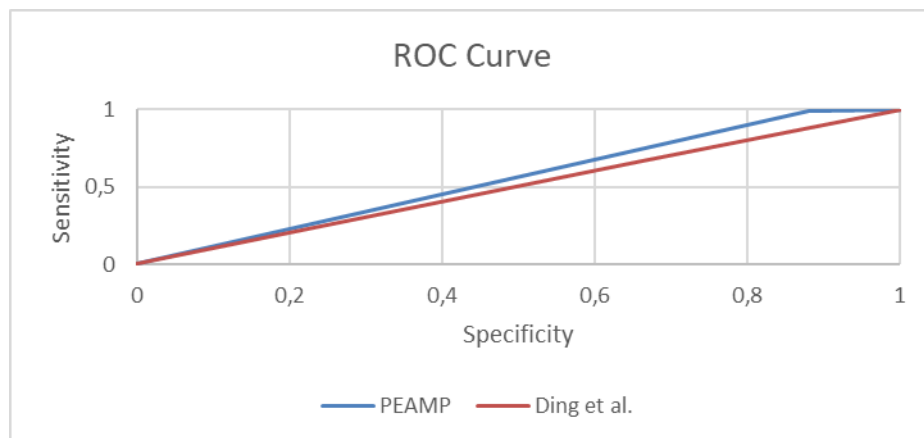


Figure 1: ROC Comparison of PEAMP Model

Comparison of the aggregate metrics accuracy, sensitivity and specificity has been performed and is shown in figure 2. It could be observed that the PEAMP model exhibits better performance in all the metrics when compared with the model presented by Ding et al.

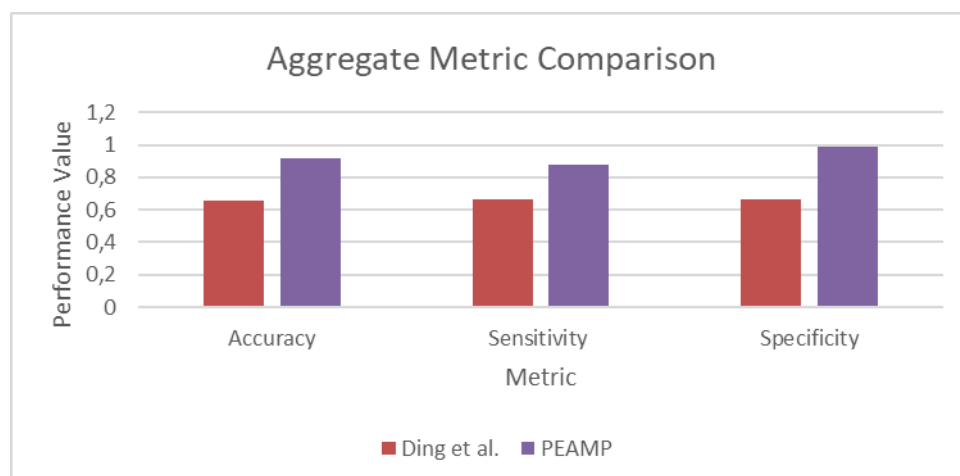


Figure 2: Aggregate Metric Comparison of PEAMP

A comparison of the Positive Predictive Value (PPV) and the Negative Predictive Value (NPV) is shown in figure 3. Positive predictive value indicates the efficiency of prediction of classes representing mortality. It could be observed that the PEAMP model exhibits 99% efficiency in predicting mortality classes, which is 30% higher than the model presented by Ding et al. Negative predictive value indicates the efficiency of prediction of classes representing healthy patients. The proposed PEAMP model exhibits slightly reduced predictions in predicting NPV.

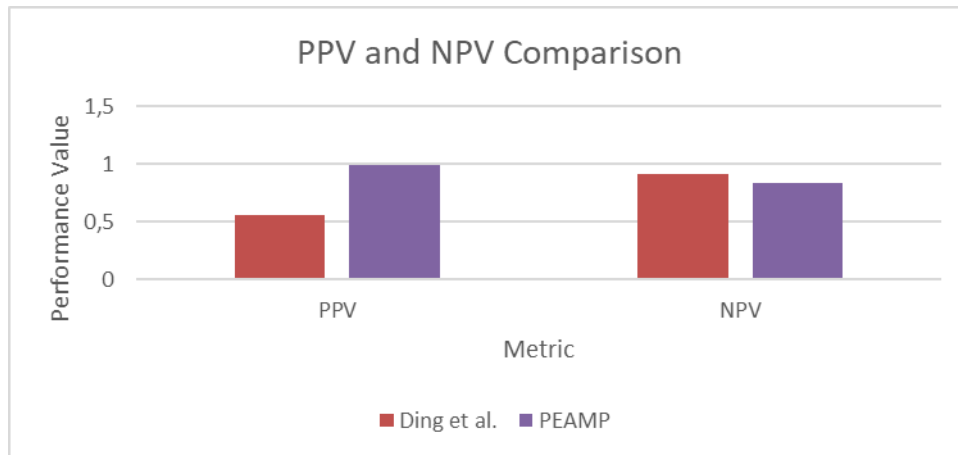


Figure 3: PPV and NPV Prediction Comparison of PEAMP

A tabulated performance comparison of PEAMP with the mortality prediction model presented by Ding et al. is presented in Table 2. The proposed PEAMP model exhibits a slight reduction in NPV at 7%. However, the PEAMP model exhibits a 43% increase in PPV, 33% increase in specificity, 22% increase in sensitivity and 26% increase in overall accuracy. These overall performances indicate that the PEAMP model is highly effective in identifying mortality, ensuring that it can be effectively deployed in real time.

Table 2: Performance Comparison of PEAMP

	Ding et al.	PEAMP
Sensitivity	0.666	0.88
Specificity	0.661	0.99
Accuracy	0.66	0.92
PPV	0.563	0.99
NPV	0.916	0.84

4. Conclusion

Intensive care units are considered to be some of the critical locations and also one of the resource consuming areas in a hospital. Identifying mortality levels accurately can ensure appropriate treatments and appropriate resource allocations. This work presents a pipelined ensemble architecture that uses multiple machine learning models to provide effective handling of the highly complex clinical data. The initial stages are designed for data preprocessing and feature selection, the intermediate stages for multiple model based prediction and the final stage for combining the predictions to provide the final result. The major advantage of this architecture is that it is composed of all the necessary components required for analysis. The proposed model is also flexible and can be parallelized effectively when deployed in a parallelizable architecture. Experimental results indicate high performance with 92% accuracy levels. Comparisons indicate 26% improvement in overall accuracy showing that the model is highly capable of being deployed in real time. However, the model still exhibits scope for improvement, as the sensitivity levels show moderate performance. Future enhancements of the model would be directed towards improving the model efficiency.

References:

- [1] S. Nemati *et al.*, "An interpretable machine learning model for accurate prediction of sepsis in the ICU," *Critical Care Medicine*, vol. 46, no. 4, pp. 547–553, 2018. doi:10.1097/ccm.0000000000002936.
- [2] A. Donabedian, "The apache II severity of disease classification system," *An Introduction to Quality Assurance in Health Care*, pp. 159–162, 2002. doi:10.1093/oso/9780195158090.005.0005.

- [3] J. R. Le Gall, "A new simplified acute physiology score (SAPS II) based on a European/North American Multi center Study," *JAMA: The Journal of the American Medical Association*, vol. 270, no. 24, pp. 2957–2963, 1993. doi:10.1001/jama.270.24.2957
- [4] A. E. W. Johnson *et al.*, "Mimic-III, a freely accessible Critical Care Database," *Scientific Data*, vol. 3, no. 1, 2016. doi:10.1038/sdata.2016.35
- [5] A. E. Johnson, D. J. Stone, L. A. Celi, and T. J. Pollard, "The mimic code repository: Enabling reproducibility in Critical Care Research," *Journal of the American Medical Informatics Association*, vol. 25, no. 1, pp. 32–39, 2017. doi:10.1093/jamia/ocx084.
- [6] Y. Chen, H. Yang, Heterogeneous postsurgical data analytics for predictive modeling of mortality risks in intensive care units, in: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vol. 2014, IEEE, 2014, pp. 4310–4314. <http://dx.doi.org/10.1109/EMBC.2014.6944578>.
- [7] A. E. Johnson *et al.*, "Machine learning and decision support in critical care," *Proceedings of the IEEE*, vol. 104, no. 2, pp. 444–466, 2016. doi:10.1109/jproc.2015.2501978 M. Bhandari, M. Reddiboina, Building artificial intelligence-based personalized predictive models, *BJU Int.* (2019) <http://dx.doi.org/10.1111/bju.14746>.
- [8] M. Ghassemi, L. A. Celi, and D. J. Stone, "State of the art review: The data revolution in critical care," *Critical Care*, vol. 19, no. 1, 2015. doi:10.1186/s13054-015-0801-4
- [9] L. Celi *et al.*, "A clinical database-driven approach to decision support: Predicting mortality among patients with acute kidney injury," *Journal of Healthcare Engineering*, vol. 2, no. 1, pp. 97–110, 2011. doi:10.1260/2040-2295.2.1.97.
- [10] L. Ryan *et al.*, "Mortality prediction model for covid-19, pneumonia, and mechanically ventilated ICU patients: A retrospective study," *SSRN Electronic Journal*, 2020. doi:10.2139/ssrn.3605352
- [11] S. Moosavi Kashani and S. Zargar, "Mortality prediction in emergency department using Machine Learning Models," *Journal of Archives in Military Medicine*, vol. 11, no. 3, 2023. doi:10.5812/jamm-140442
- [12] N. Ding, C. Guo, C. Li, Y. Zhou, and X. Chai, "An artificial neural networks model for early predicting in-hospital mortality in acute pancreatitis in mimic-III," *BioMed Research International*, vol. 2021, pp. 1–8, 2021. doi:10.1155/2021/6638919
- [13] M. Selcuk, O. Koc, and A. S. Kestel, "The prediction power of machine learning on estimating the sepsis mortality in the Intensive Care Unit," *Informatics in Medicine Unlocked*, vol. 28, p. 100861, 2022. doi:10.1016/j.imu.2022.100861
- [14] M. Selcuk, O. Koc, and A. S. Kestel, "The prediction power of machine learning on estimating the sepsis mortality in the Intensive Care Unit," *Informatics in Medicine Unlocked*, vol. 28, p. 100861, 2022. doi:10.1016/j.imu.2022.100861
- [15] R. Sadeghi, T. Banerjee, and W. Romine, "Early hospital mortality prediction using vital signals," *Smart Health*, vol. 9–10, pp. 265–274, 2018. doi:10.1016/j.smhl.2018.07.001.
- [16] N. Veith and R. Steele, "Machine learning-based prediction of ICU patient mortality at TIME OF ADMISSION," *Proceedings of the 2nd International Conference on Information System and Data Mining*, 2018. doi:10.1145/3206098.3206116.
- [17] M. A. H. Zahid and J. Lee, "Mortality prediction with self normalizing neural networks in intensive care unit patients," *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, Las Vegas, NV, USA, 2018, pp. 226-229, doi: 10.1109/BHI.2018.833341.
- [18] N. Veith and R. Steele, "Machine learning-based prediction of ICU patient mortality at TIME OF ADMISSION," *Proceedings of the 2nd International Conference on Information System and Data Mining*, 2018. doi:10.1145/3206098.3206116
- [19] B. Thompson, K. O. Elish and R. Steele, "Machine Learning-Based Prediction of Prolonged Length of Stay in Newborns," *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Orlando, FL, USA, 2018, pp. 1454-1459, doi: 10.1109/ICMLA.2018.00236.
- [20] S. Nemati, *et al.*, "An interpretable machine learning model for accurate prediction of sepsis in the ICU," *Critical Care Medicine*, vol. 46, no. 4, pp. 547-553, Apr. 2018. DOI: 10.1097/CCM.0000000000002936.
- [21] A. Moser *et al.*, *Mortality prediction in intensive care units including premorbid functional status improved performance and internal validity*, 2021. doi:10.31219/osf.io/rb62v
- [22] C. Guo, M. Liu, and M. Lu, "A Dynamic Ensemble Learning algorithm based on K-means for ICU mortality prediction," *Applied Soft Computing*, vol. 103, p. 107166, 2021. doi:10.1016/j.asoc.2021.107166
- [23] G. S. Krishnan and S. K. S., "A novel ga-elm model for patient-specific mortality prediction over large-scale lab event data," *Applied Soft Computing*, vol. 80, pp. 525–533, 2019. doi:10.1016/j.asoc.2019.04.019
- [24] V. Mandalapu, B. Ghaemmaghami, R. Mitchell, and J. Gong, "Understanding the relationship between healthcare processes and in-hospital weekend mortality using MIMIC III," *Smart Health*, vol. 14, p. 100084, 2019. doi:10.1016/j.smhl.2019.100084