

Article

Patient Mortality Prediction and Analysis of Health Cloud Data Using a Deep Neural Network

Abdullah Alourani ¹, Kinza Tariq ^{2,*}, Muhammad Tahir ² and Muhammad Sardaraz ²

¹ Department of Computer Science and Information, College of Science in Zulfi, Majmaah University, Al-Majmaah 11952, Saudi Arabia

² Department of Computer Science, COMSATS University Islamabad, Attock Campus, Attock 43600, Pakistan

* Correspondence: kinzatariq221@gmail.com

Abstract: Cloud computing plays a vital role in healthcare as it can store a large amount of data known as big data. In the current emerging era of computing technology, big data analysis and prediction is a challenging task in the healthcare industry. Healthcare data are very crucial for the patient as well as for the respective healthcare services provider. Several healthcare industries adopted cloud computing for data storage and analysis. Incredible progress has been achieved in making combined health records available to data scientists and clinicians for healthcare research. However, big data in health cloud informatics demand more robust and scalable solutions to accurately analyze it. The increasing number of patients is putting high pressure on healthcare services worldwide. At this stage, fast, accurate, and early clinical assessment of the disease severity is vital. Predicting mortality among patients with a variety of symptoms and complications is difficult, resulting inaccurate and slow prediction of the disease. This article presents a deep learning based model for the prediction of patient mortality using the Medical Information Mart for Intensive Care III (MIMIC-III) dataset. Different parameters are used to analyze the proposed model, i.e., accuracy, F1 score, recall, precision, and execution time. The results obtained are compared with state-of-the-art models to test and validate the proposed model. Moreover, this research suggests a simple and operable decision rule to quickly predict patients at the highest risk, allowing them to be prioritized and potentially reducing the mortality rate.

Keywords: machine learning; deep neural network; mortality prediction; accuracy; e-health; cloud computing



Citation: Tariq, K.; Alourani, A.; Tahir, M.; Sardaraz, M. Patient Mortality Prediction and Analysis of Health Cloud Data Using a Deep Neural Network. *Appl. Sci.* **2023**, *13*, 2391. <https://doi.org/10.3390/app13042391>

Academic Editors: Vincenzo Conti, Carmelo Militello, Leonardo Rundo, Andrea Tangherloni, Fulvio Zaccagna and Filippo Vella

Received: 4 January 2023

Revised: 8 February 2023

Accepted: 9 February 2023

Published: 13 February 2023



Copyright: © 2023 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) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Healthcare data are increasing at a rapid rate, which leads to the generation of big data. The healthcare data are crucial for patient, as well as for the respective healthcare services provider. The analysis of big data is a challenging issue for the healthcare industry [1]. Several healthcare industries opted for cloud computing to store and analyze healthcare data. Incredible progress has been achieved in making combined health records available to data scientists and clinicians for the activities of healthcare research [2]. An accurate model for the prediction of mortality in Intensive Care Unit (ICU) is an important challenge. Prediction of mortality has been carried out using many machine learning algorithms, including Decision Trees (DT), Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), Discriminant Analysis (DA), Naive Bayes (NB), and Neural Networks (NN). These methods can predict the mortality risk on early basis [3,4].

Many techniques have been developed and used for predicting the risk of ICU mortality in the past decade [5]. In these techniques, the data of different intervals of the patients was used and these theoretical data were documented manually by expert personnel [6]. The approaches of machine learning can acquire the data of the patient for the purpose of mortality prediction. These methods result in better analysis of the data in a cost effective

way [7]. The development of such systems can lead to self-organized networks that leads to service continuity in distributed systems. Realtime data analysis can be used to trigger different events and lead to Service Oriented Networks (SONs) [8].

This research presents an efficient framework for the prediction of mortality rate for cloud health informatics. The study aims to provide crucial information to the physician at an early stage so that the physician can understand the critical condition of the patient and can take timely action. The required data are dynamically accessed and updated from the electronic health care record with short intervals. Machine learning algorithms are used for mortality prediction. For this purpose, different machine learning methods are employed, including Random Forest (RF), Decision Tree (DT), Logistic Regression (LR), Support Vector Machine (SVM), Neural Network (NN), and Gradient Boosting (GB). MIMIC III data are used for training and testing of the proposed model. The models are evaluated using different parameters including accuracy, precision, recall, F-Score, and execution time. The deep neural network model with coefficient correlation matrix yields better results as compared to other models used.

The paper is organized as follows. Section 2 presents literature review with a detailed discussion of state-of-the-art methods. Section 3 presents materials and methods with a detailed discussion of the dataset, features selection, and other steps of the proposed algorithm. Section 4 presents results and discussion, followed by the conclusion section.

2. Literature Review

Machine learning techniques are successfully employed in many domains, including network security [9], resource utilization prediction [10], data compression [11], healthcare [12–14], and many more. There exist review articles covering the use of machine learning in healthcare and other applications [15–17]. Some articles also cover feature selection and feature dimensionality reduction, and weighting [18–21]. This section presents more related works to the proposed model. A comparative analysis of different methods is also presented.

Different techniques have been developed and used for predicting the risk of ICU mortality in the past decade. ICU is a critical care unit that is equipped with advanced technology medical machinery for monitoring patients continuously. The increasing need for these electronic devices creates an opportunity for the approaches of machine learning for the prediction of mortality.

Authors in [22] used machine learning and cloud computing to develop a system for monitoring the health status of heart patients. Different machine learning models are evaluated to obtain the most accurate model while targeting the Quality of Service (QoS) parameters. To validate the selected model, cross-validation methods are used using different parameters, including accuracy, sensitivity, specificity, latency, memory usage, and execution time. The proposed algorithm along with a mobile application is also tested on cloud platform with different input datasets. A lightweight authentication protocol for IoT devices in cloud computing environments is proposed in [23]. The proposed protocol consists of three layers, i.e., IoT devices, trust center at the edge layer, and cloud datacenter. The trust center is used to authenticate and authorize the device to access the services provided by the cloud service provider. An algorithm for the prediction of mortality of heart failure is proposed in [24]. Various machine learning models are applied to heart disease datasets to predict the mortality of patients from heart failure. Different performance measures, including accuracy, precision, and recall, are used to compare different machine learning models. Ensemble learning is used to predict mortality based on the accuracy achieved during the validation. Authors in [25] use a deep reinforcement learning approach for computation offloading and auto-scaling in mobile fog computing. In the proposed approach, first, a greedy approach is used to evaluate the fog device locally. In the next step, deep reinforcement learning is used to find the best destination for the execution of the module. The approach considers the trade-off between power consumption and execution

time during the module selection phase. Comparative experimental results are presented to validate the proposed approach.

Another model for heart disease prediction with the integration of IoT and Fog computing using cascaded deep learning model is presented in [26]. In the initial phase, data are collected from the patients using hardware components. In the next phase, features related to heart diseases are extracted. The selected features are then used to predict the disease using a cascaded convolutional neural network. Galactic swarm optimization is used to optimize the hyperparameters. Comparative experimental results are shown to validate the proposed model. Authors in [27] used different machine learning models to predict patient mortality in the ICU. The MIMIC-III dataset is used to develop machine learning models. The authors employed RM, the Least Absolute Shrinkage and Selection Operator (LASSO), LR, and GB for prediction. The performance of the model is evaluated with different parameters. An efficient CNN-based model for the prediction of coronary heart disease is presented in [28]. The proposed model targets the problem of data imbalance to achieve accurate results. The proposed model is based on two steps, i.e., first LASSO is used for features weights assessment followed by identification of important features with majority voting-based identification. Fully connected layer is used to homogenize the important features. Comparative results are presented to validate the proposed model. Authors in [29] present the concept of e-health cloud with emphasis on many components including the architecture for e-healthcare. Different new opportunities and challenges are discussed. The challenges of privacy and security of e-health cloud are discussed which are the solution of some addressed challenges of e-health cloud. Choi et al. [30] proposed Graph-based Attention Model (GRAM), which uses various leveled data characteristics of clinical ontologies. Experiments on various machine learning algorithms. The proposed model achieved higher accuracy as compared to traditional machine learning models.

A Multilevel Medical Embedding (MiME) system is presented in [31]. The proposed system learns the multilevel embedding of health data and performs predictions. Heart disease prediction and sequential disease prediction are targeted. The method outperforms the baseline methods in different evaluation tests. In ref. [32], the architecture of fog computing and service-oriented computing is proposed which calculates and authenticates the data of healthcare through the Internet of Medical Things (IoMT). In this method, resource-constrained devices are used for analysis and data mining. In recent years, with the emerging field of technology, IoT takes a big part in the lives of people. As IoT has evolved in many fields and provides ease to users including the field of healthcare. Different techniques contain specialized devices that can monitor the condition of the patient continuously and generate big data. Moreover, numerous available IoT devices and portable devices, for example, Electrocardiogram (ECG), Electromyography (EMG) can measure specific parameters, such as the rate of heart, rate of respiration, blood pressure, and many others with a single click. A mortality prediction system in clinical traumatic patients is proposed in [33]. Different machine-learning techniques are used in the proposed system. Supervised learning techniques are constructed with ten-fold cross-validation. Parameters, including accuracy, specificity, precession, recall, F-Score, and AUC, are used to validate the proposed model. Authors in [34] presented a model applying machine learning techniques to the data of the clinic from a large group of patients. The patient suffering from COVID-19 are being treated at the health system of Mount Sinai in New York City, NY, USA, for mortality prediction of the patient. The prediction model is based on the characteristics of the patients and clinical features for mortality prediction with the help of datasets developed. The proposed model is evaluated using AUC scores.

Authors in [35] developed a deep learning structure for mortality prediction for specific patients. Several machine learning models are compared using two sepsis screening tools, including the systemic inflammatory response syndrome (SIRS) and quick sepsis-related organ failure assessment (qSOFA). Different parameters are used

to validate the proposed model. To identify crucial predictive biomarkers of disease mortality a model is presented in [36]. Machine learning tools are employed to implement the model using three biomarkers. Different parameters are used to validate the proposed model. In ref. [37], the authors present the results with benchmarking for prediction of clinical activities for example patients mortality prediction, prediction of time or interval of stay, etc., deep learning models. The MIMI-III (v1.4) dataset is used for prediction. The results show that the proposed model performs well as compared to other methods specifically when the raw data of clinical time series is used as input features to the models. In ref. [38], the authors develop and test a random forest-based prediction system to effectively predict all-cause mortality in patients. Various parameters are used to test the proposed model. The results show that the model outperformed the existing models. Table 1 show the theoretical comparison of different methods used for mortality prediction. The comparison is based on the methods used, targeted parameters, and the datasets used for validation.

Table 1. Theoretical comparison of different methods.

Ref.	Methods	Parameters	Dataset
[22]	SVM, NN, KNN, GB	Accuracy, Sensitivity, Specificity, AUC Score	Cleveland Heart Disease dataset
[24]	SVM, NN, ensemble learning	accuracy, precision, recall	Manually gathered dataset
[26]	OCCNN and Galactic Swarm Optimization	Accuracy, Sensitivity, Specificity	Manually gathered dataset
[27]	LASSO, RF, GB, LR	AUROC, Sensitivity, Specificity	MIMIC-III
[28]	LASSO, CNN	Accuracy, Recall, Specificity, AUC	Data from NHANES
[33]	LR, NN, SMO, JRip, RF	Accuracy, specificity, precision, recall, F-measure, and AUC	RETRAUCI database
[34]	RF, LR, SVM, XGBoost	AUC Score	Patients with COVID-19 treated at the Mount Sinai Health System in New York City
[35]	CNN, KNN, SVM, RF	AUC Score	Manually gathered dataset
[37]	LSTM-based NN	Accuracy, AUC-ROC	MIMIC-III

3. Methods and Material

This section presents the details of the proposed framework. To build the model for mortality prediction machine learning algorithms are used. The MIMIC-III dataset is used for validation. Initially, pre-processing is performed on the data then the dataset is reduced to normalized form. The missing values are replaced with the mean value. Under-sampling is performed on the discharge and deceased record after which the BMI is calculated. Based on these steps and with the required vital features the death rate and the live patients are calculated. After this, the training and testing are performed on the data with the ratio of 20/80. The different algorithms of machine learning are used to find out the evaluation metrics and the results of various algorithms of machine learning are compared to find out the best accuracy and efficient time performance. The steps involved in the proposed prediction model are shown in Figure 1. Figure 2 shows the sequence of activities involved in the proposed model.

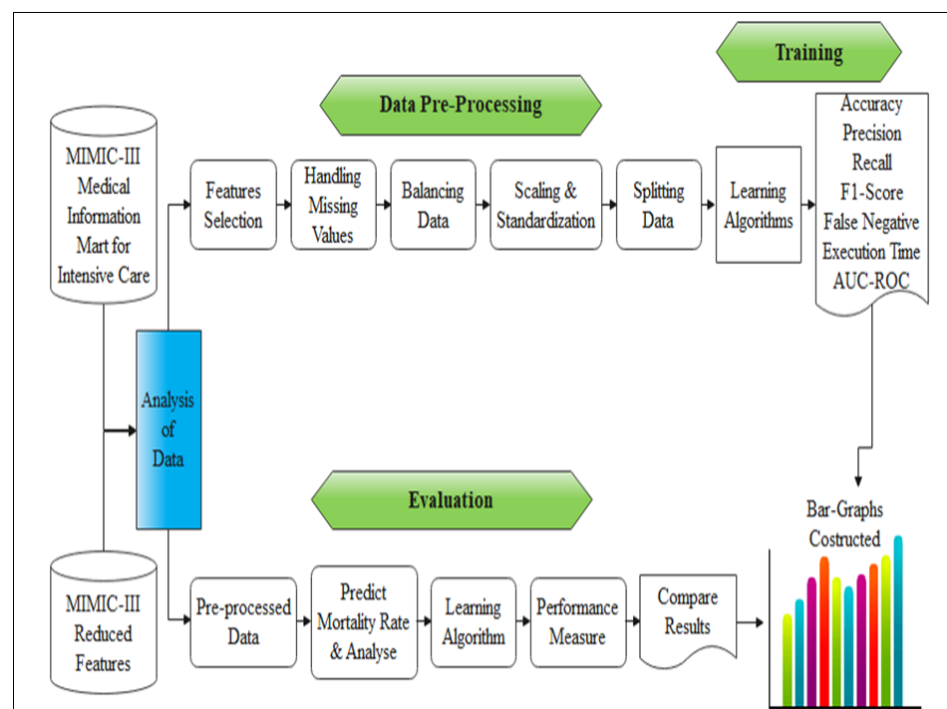


Figure 1. Architecture of the proposed prediction model.

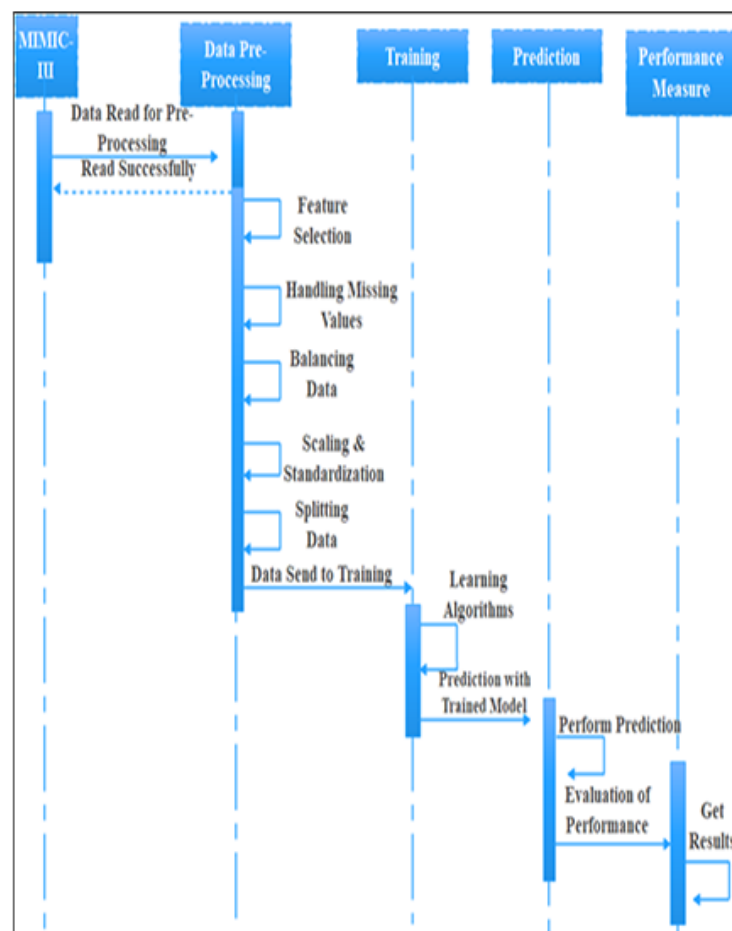


Figure 2. Sequence diagram of the proposed prediction model.

3.1. Dataset

MIMIC-III is a publicly available dataset [39]. Data were gathered from a variety of sources, including critical care information system archives, databases of electronic medical records in hospitals, death master file of the social security administration. This is a huge, centralized database containing patient's related information admitted to the special care center. The dataset has vital signs, medications, laboratory measurements, observations and notes that are recorded by physicians, fluid balance, codes of procedure, codes of diagnosis, reports in the form of images, the interval of staying in hospital, data for survival, etc.

3.2. Features Selection from Data

The selection of features from the dataset is performed on the available dataset. The features in the first 6 h of ICU admission, the average maximum and minimum values of the selected features were calculated and selected as features for prediction. After the stay in ICU, if the patient has died in the hospital, the mortality of the hospital is used as the response variable. Even though the size of the sample is relatively huge, the selection of features is vital for the elimination of extremely associated values. Filter methods pick up the intrinsic properties of the features measured via univariant statistics instead of cross-validation performance. These methods are faster and computationally less expensive than wrapper methods. When dealing with high-dimensional data, it is computationally better to use filter methods.

The flow of the feature selection process is shown in Figure 3. Correlation is a measure of the linear relationship between two or more variables. Through correlation, we can predict one variable from the other. The logic behind using correlation for feature selection is the good variables that are highly related to the target. Furthermore, variables should be correlated with the target but should be uncorrelated among themselves. If two variables are correlated, we can predict one from the other. Therefore, if two features are correlated, the model only needs one of them, as the second one does not add additional information. In this work, we use Pearson Correlation. Absolute value, 0.7 as the threshold for selecting the variables was used. The predictor variables are correlated among themselves. Variables that have a lower correlation coefficient value with the target variable were dropped out. We also compute multiple correlation coefficients to check whether more than two variables are correlated. Even though the size of the sample is relatively huge, the selection of features is significant for reducing the extremely interconnected variables. As an association matrix for the interconnected values is constructed as shown in Figure 4. There are extremely associated values in the dataset. The heat map in Figure 4, uses the intensity of color to show the interconnected variables between the features (light color indicates the negatively related values and darker color indicates the positively related values). This is used to categorize redundantly and even recurrent tests of the laboratory. A total of 138 features were retained for further analysis.

3.3. Missing Values from Data

The percentage of the omitted values of the variables for every feature by the category of response variables i.e., deceased, discharged was computed. Within the first 6 h, all lab tests are not measured for all patients. For every patient who is considered normal, it is expected that there will be missing data. If the record of any patient has multiple misplaced values, it may be expected that they were considered as suffering from minor sickness hence they were not prescribed for the lab test. In the same way, if any patient dies before all the lab tests were performed. By setting up a threshold at two levels raw and features levels, misplaced values were handled. Some features are dropped those which are having more than 60% of missing values and are insignificant clinically. After excluding the features having more missing values, 90 features are remaining in the dataset. Some tests, such as urine analysis and stool test, are not tested in the first six hours of admission, hence, these features are dropped. 11 out of the 90 features are attributes related to admission, such as length of hospital and ICU stay, time of admission, and severity scores. Records with at

most 40% missing values are kept in the dataset which results in a total of 40,246 records. Values that are misplaced in the feature matrix were attributed to the column means.

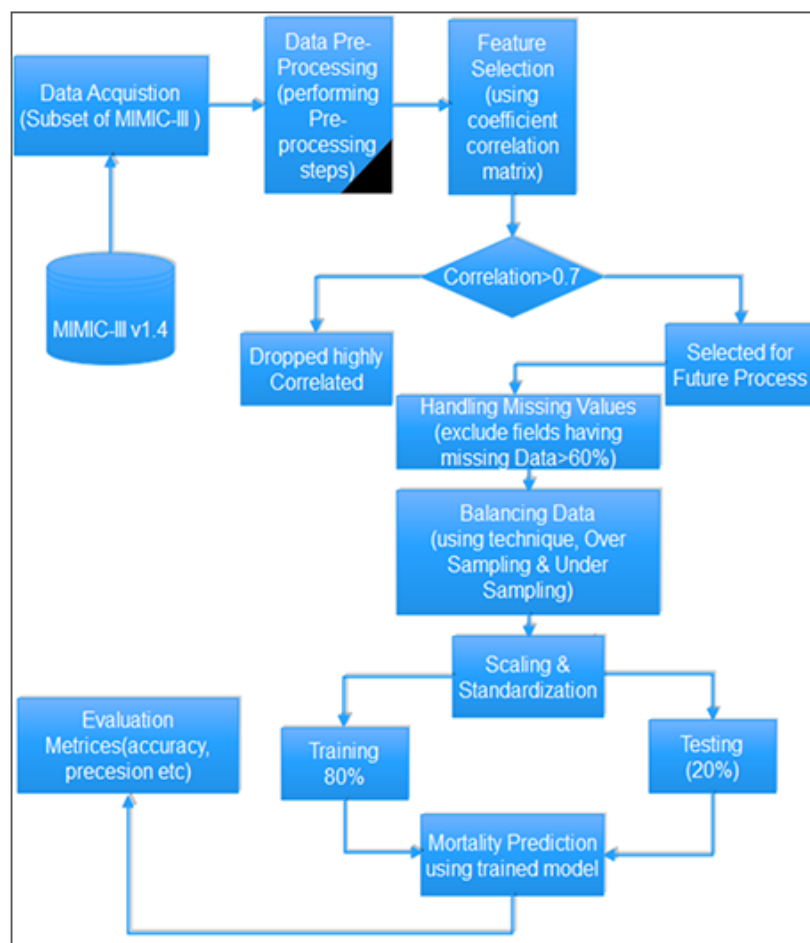


Figure 3. Flow of features selection process.

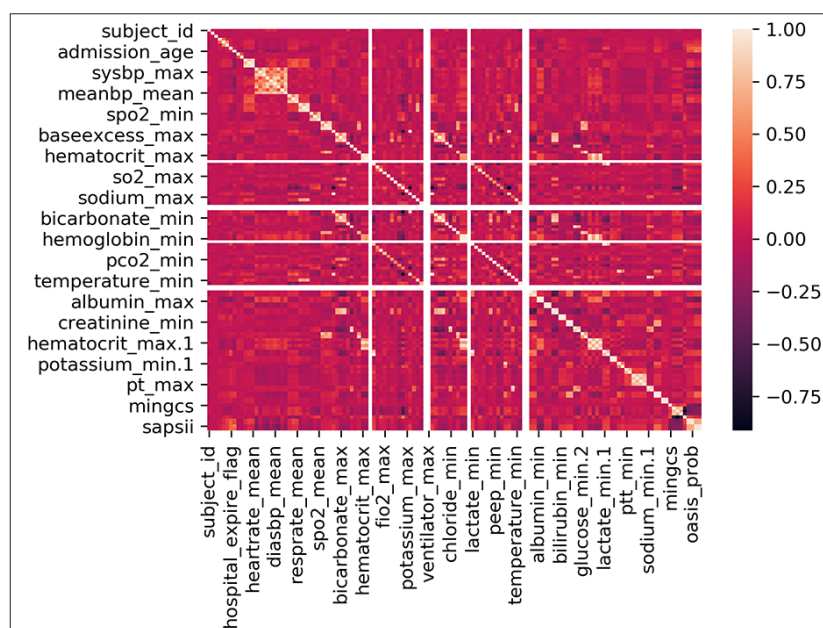


Figure 4. Correlation matrix of all variables selected for experiments.

3.4. Balancing the Data

After addressing the missing values of the variables, the data show that only 18% of the patients died during the hospital stay. This generates classes of imbalance as the ratio between the two classes, i.e., passed away during the hospital stay to discharge from the hospital, is the ratio of 12/88. The use of imbalanced data for training models, generates biased results towards the majority class. The technique of under-sampling is used for data balancing. Even though the broadly used technique is up-sampling, as the data are relatively large and generating an additional class of the alternative (deceased) was not essential. Therefore, model choice algorithms that permit arbitrarily choosing an equivalent number of records from the majority (discharged) class was used for the stability of data. After the under-sampling of data, the aggregate number of records retained was 4861 with a class ratio of 50:50.

3.5. Scaling and Standardization

Continuous values, such as weight, age, and measures of the laboratory were scaled to a range of 0 to 1, as shown in Equation (1).

$$x_s = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (1)$$

where x_s denotes a scaled value, x_i represents the original value, and min and max are the minimum and maximum values of a particular feature. Scaling prevents variables from over-dominating in the supervised learning algorithms.

3.6. Data Splitting

After addressing missing values, balancing the data, and normalization, data splitting is performed. From the whole dataset of 40,246 records, the remaining 80 features were available to use for the development of the model. For the development of the model, the appendix lists are used for the use of features. Finally, the dataset is split into 20/80 splits, i.e., 20% is used for the training purpose and 80% of the data are used for testing. We used sklearn package with model selection and train test split. The splitting is performed to ensure that there is enough data in the training dataset for the model to learn an effective mapping of inputs to outputs. There will also be enough data in the test set to effectively evaluate the model performance in terms of computation and time.

4. Model Training

We select the model parameters using a set of validation with the best model of performance. While the set of testing evaluates the performance of the model and evaluation of results. Training of the model is performed by the training set. For performing these experimentations, the vital predictors from the MIMIC dataset are used to predict the mortality rate with better accuracy. These vital input predictors are *icustay_id*, *icustay_seq*, *los_hospital*, *los_icu*, *oasis*, *oasis_prob*, *sapsii*, *sapsii_prob*, *intime*, *hadm_id*, *specimen*, *gender*, and *intubated* on which the prediction is performed. Different algorithms of machine learning are used for prediction based on the evaluation metrics selected. The experiments have been completed with various arrangements of batch sizes, and the number of epochs to select the best parameters and hyper-parameters for the proposed model. We have applied sigmoid and Relu activation functions in the experiments. Linear activation has been used in the output layer while sigmoid is used in the hidden layer.

Model Configuration

The proposed prediction model was implemented using TensorFlow, a widely used python library for DNN. LR is implemented using scikit-learn python package for comparative analysis. The number of binary features is set equal to the number of possible unique values of categorical features. This approach was utilized for the age variable in the

proposed model. The batch size was set to 1600, init was used as uniform, and the number of epochs was 60. The number of neurons in the hidden layer was selected using hyperparameter optimization. We used the activation function of Relu and Sigmoid. We also used Adam as the algorithm for first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order moments to train the neural networks. The architecture of DNN is shown in Figure 5. Figure 6 shows architecture of the proposed model consisting of different components and activities.

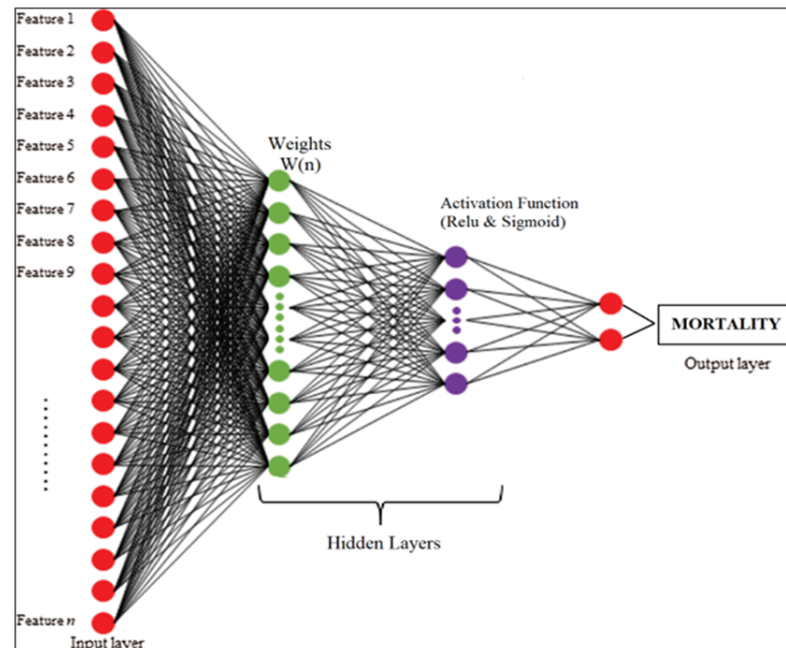


Figure 5. Architecture of DNN.

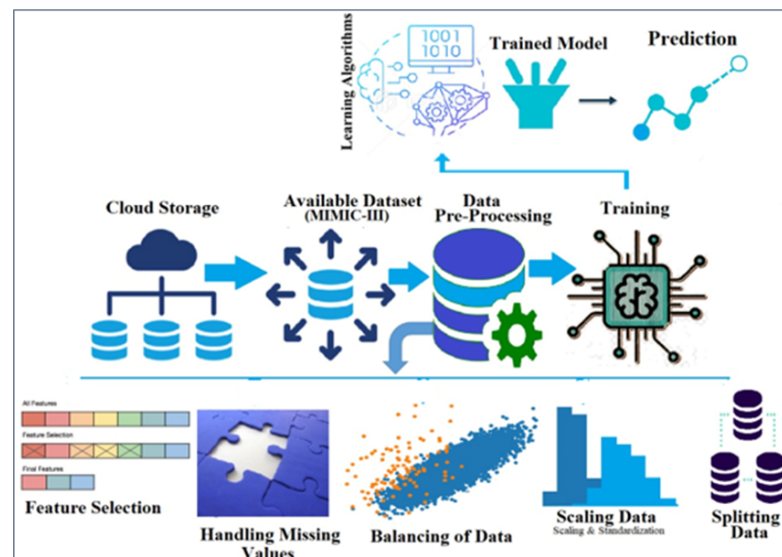


Figure 6. Architecture of the proposed model.

5. Results and Discussion

This section presents the detailed results and comparative analysis of mortality prediction based on various machine learning models. All models were applied to selected features, i.e., vibrant signs, medications, different tests of laboratory, observations, and notes charted by a doctor or a physician, balance of fluid, codes of procedure, codes of diagnostic, imaging-based reports, length of stay, data of survival, etc. A simple layer of multilayer perceptron (MLP) of NN to do a binary classification task with prediction probability is used.

The dataset has an input dimension of 78. The hidden layers are followed by the output layer. Furthermore, the number of neurons is set as uniform. The number of neurons range is set as 78. The activation functions of Sigmoid and Relu are used in layers. For compiling the layers of NN, Adam optimizer is used. As the optimizer is responsible to change the learning rate and weights of neurons in the NN to reach the minimum loss function. Epochs are set to 60. For making the learning process fast, batch size of 500 was used. All the models were applied to the pre-processed dataset and evaluated by evaluation metrics presented in the following subsection.

Evaluation Measures

Evaluation metrics are used to evaluate the performance of the prediction model. Accuracy, precision, recall, and F-score are used to evaluate the performance of different models. The term accuracy refers to how much close the produced results are to the true values. For uniform datasets it is a useful evaluation parameter. Accuracy is calculated using Equation (2).

$$\text{Accuracy} = \frac{TP}{TP + TN + FP + FN} \quad (2)$$

Precision shows how close two or more quantities are to each other. High precision is associated with a low false-positive rate. Precision is calculated with Equation (3).

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

Recall is the division of the related data that are retrieved successfully. It is the number of results that are correct and divided by the number of results that should be returned, as shown in Equation (4).

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

F-score is a measure of the accuracy of a model on a dataset. It is calculated as the normalized average of precision and recall including both false positives and false negatives. F-score is calculated as shown in Equation (5).

$$\text{F-Score} = 2 \times \frac{PR}{(P + R)} \quad (5)$$

Comparative results are presented in Table 2. Results analysis show that DNN outperforms the other frameworks by achieving accuracy of 80% followed by gradient boosting with 78% accuracy. The DT gives the lowest accuracy rate of 70% on the extracted features of the dataset used. In terms of precession, DNN achieved a precision of 86% which is higher as compared to other algorithms. DNN is followed by gradient boosting with 80% precision. While the decision tree gives the lowest precision of 72%. DNN achieved recall score of 79% followed by gradient boosting with a 78% recall score. While decision tree and SVM linear give the lowest recall score of 70%. In terms of F1-score, DNN achieved an F1-score of 80% followed by gradient boosting with a 78% F1-score. While LR and SVM linear give the lowest F1-score of 70% on the extracted features of the dataset used.

According to the results shown in Table 2, DNN obtained 6.67% improvement over RR and SVM-linear, 14.28% improvement over DT, 5.26% improvement gain over SVM-rfb and RF, 3.89% improvement over NN, and 2.56% improvement gain over GB in terms of accuracy. The results show that in terms of precision, DNN obtained 17.18, 19.44, 13.25, 10.25, and 7.5 percent improvement gain in comparison to DT, LR, SVM-linear, SVM-rfb, RF, NN, and GB, respectively. The results show that the percent improvement gain attained by DNN in terms of recall over LR is 2.59 and 12.89 percent over SVM linear. The improvement gain over both SVM-rfb and RF is 5.33 percent. In comparison to GB and NN, DNN attained improvement gain of 3.94 percent. In comparison to SVM-linear and DT, the proposed framework yields 12.85 percent improvement gain. The results

analysis shows that DNN obtained 5.26 percent improvement over LR, SVM linear, and NN. DNN yields 2.56 percent improvement over DT, 3.89 percent improvement over SVM-rfb, and 6.66 percent improvement gain over RF in terms of F1-score. The objective of the proposed study is to predict mortality using machine learning techniques. All these methods are parametric, i.e., based on probability distribution that ensures that the results are statistically significant.

Table 2. Comparative analysis of ML algorithms in terms of accuracy, precision, recall, F1 score, and execution time (ET).

Method	Accuracy	Precession	Recall	F1-Value	ET (Training)	ET (Testing)
LR	0.75	0.73	0.77	0.76	1.89	0.33
DT	0.70	0.72	0.70	0.78	1.42	0.17
SVM-L	0.75	0.76	0.70	0.76	28.08	13.06
SVM-RBF	0.76	0.78	0.75	0.77	33.79	16.67
RF	0.76	0.78	0.75	0.75	1.99	0.82
NN	0.77	0.78	0.76	0.76	8.06	4.33
GB	0.78	0.80	0.76	0.78	1.16	0.84
[28]	0.78	0.81	0.78	0.78	31.3	14.71
[27]	0.78	0.82	0.77	0.78	25.62	13.14
DNN	0.80	0.86	0.79	0.80	44.2	22.31

6. Conclusions

Patient mortality prediction is a challenging task due to the huge amount of data and missing values. To achieve better performance and accuracy, we propose a model based on machine learning. In the layer of prediction, we used DNN. We input the data in the pre-processed form given to the model for building an optimum model with hidden layers, neurons, activation functions optimizers, and epochs. The proposed approach highlights certain significant features that may need to be observed further to estimate their effect on mortality of the patient. Various machine learning models are tested, and evaluation metrics are measured. The best performance is obtained from DNN as compared to the other learning algorithms. Overall, DNN performs best and gives the best performance based on measured parameters, i.e., accuracy, precision, recall, and F1-score.

The execution time of the proposed model is higher as compared to other methods. In future, additional measurement parameters, i.e., time and space complexity need to be considered to justify their significance in the mortality prediction. Moreover, feature selection with other features selection methods can also lead to better results. The model can be tested on other datasets, particularly large-scale datasets to ensure that the model achieves similar results. The proposed model needs to be tested in an IoT-integrated environment to ensure the applicability and validation of the results. The model also needs to be tested in integration with other applications to build a service-oriented architecture for healthcare industry.

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

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The datasets used are available publicly.

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

References

1. AbuKhoussa, E.; Mohamed, N.; Al-Jaroodi, J. e-Health cloud: Opportunities and challenges. *Future Internet* **2012**, *4*, 621–645. [\[CrossRef\]](#)
2. Jiang, F.; Jiang, Y.; Zhi, H.; Dong, Y.; Li, H.; Ma, S.; Wang, Y.; Dong, Q.; Shen, H.; Wang, Y. Artificial intelligence in healthcare: Past, present and future. *Stroke Vasc. Neurol.* **2017**, .. [\[CrossRef\]](#)
3. Awad, A.; Bader-El-Den, M.; McNicholas, J.; Briggs, J.; El-Sonbaty, Y. Predicting hospital mortality for intensive care unit patients: time-series analysis. *Health Inform. J.* **2020**, *26*, 1043–1059. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Metsker, O.; Sikorsky, S.; Yakovlev, A.; Kovalchuk, S. Dynamic mortality prediction using machine learning techniques for acute cardiovascular cases. *Procedia Comput. Sci.* **2018**, *136*, 351–358. [\[CrossRef\]](#)
5. Johnson, A.E.; Pollard, T.J.; Mark, R.G. Reproducibility in critical care: A mortality prediction case study. In *Machine Learning for Healthcare Conference*; PMLR: Boston, MA, USA, 2017; pp. 361–376.
6. Van de Leur, R.; Bleijendaal, H.; Taha, K.; Mast, T.; Ghos, J.; Linschoten, M.; van Rees, B.; Henkens, M.; Heymans, S.; Sturkenboom, N.; et al. Electrocardiogram-based mortality prediction in patients with COVID-19 using machine learning. *Neth. Heart J.* **2022**, 1–7. [\[CrossRef\]](#)
7. 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.* **2016**, *23*, 269–278. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Conti, V.; Militello, C.; Rundo, L.; Vitabile, S. A novel bio-inspired approach for high-performance management in service-oriented networks. *IEEE Trans. Emerg. Top. Comput.* **2020**, *9*, 1709–1722. [\[CrossRef\]](#)
9. Alduailij, M.; Khan, Q.W.; Tahir, M.; Sardaraz, M.; Alduailij, M.; Malik, F. Machine-Learning-Based DDoS Attack Detection Using Mutual Information and Random Forest Feature Importance Method. *Symmetry* **2022**, *14*, 1095. [\[CrossRef\]](#)
10. Malik, S.; Tahir, M.; Sardaraz, M.; Alourani, A. A Resource Utilization Prediction Model for Cloud Data Centers Using Evolutionary Algorithms and Machine Learning Techniques. *Appl. Sci.* **2022**, *12*, 2160. [\[CrossRef\]](#)
11. Alourani, A.; Tahir, M.; Sardaraz, M.; Khan, M.S. Knowledge-Based Framework for Selection of Genomic Data Compression Algorithms. *Appl. Sci.* **2022**, *12*, 11360. [\[CrossRef\]](#)
12. Theis, J.; Galanter, W.L.; Boyd, A.D.; Darabi, H. Improving the in-hospital mortality prediction of diabetes ICU patients using a process mining/deep learning architecture. *IEEE J. Biomed. Health Inform.* **2021**, *26*, 388–399. [\[CrossRef\]](#)
13. Karthikeyan, A.; Garg, A.; Vinod, P.; Priyakumar, U.D. Machine learning based clinical decision support system for early COVID-19 mortality prediction. *Front. Public Health* **2021**, *9*, 626697. [\[CrossRef\]](#)
14. Chui, K.T.; Gupta, B.B.; Alhalabi, W.; Alzahrani, F.S. An MRI scans-based Alzheimer's disease detection via convolutional neural network and transfer learning. *Diagnostics* **2022**, *12*, 1531. [\[CrossRef\]](#)
15. Bottino, F.; Tagliente, E.; Pasquini, L.; Napoli, A.D.; Lucignani, M.; Figà-Talamanca, L.; Napolitano, A. COVID mortality prediction with machine learning methods: A systematic review and critical appraisal. *J. Pers. Med.* **2021**, *11*, 893. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Chen, D.; Liu, S.; Kingsbury, P.; Sohn, S.; Storlie, C.B.; Habermann, E.B.; Naessens, J.M.; Larson, D.W.; Liu, H. Deep learning and alternative learning strategies for retrospective real-world clinical data. *NPJ Digit. Med.* **2019**, *2*, 1–5. [\[CrossRef\]](#)
17. Zemouri, R.; Zerhouni, N.; Racoceanu, D. Deep learning in the biomedical applications: Recent and future status. *Appl. Sci.* **2019**, *9*, 1526. [\[CrossRef\]](#)
18. Jia, W.; Sun, M.; Lian, J.; Hou, S. Feature dimensionality reduction: A review. *Complex Intell. Syst.* **2022**, 1–31. [\[CrossRef\]](#)
19. Leng, L.; Li, M.; Teoh, A.B.J. Conjugate 2DPalmHash code for secure palm-print-vein verification. In Proceedings of the 2013 6th International Congress on Image and Signal Processing (CISP), Hangzhou, China, 16–18 December 2013; Volume 3, pp. 1705–1710.
20. Leng, L.; Zhang, J. Palmhash code vs. palmphasor code. *Neurocomputing* **2013**, *108*, 1–12. [\[CrossRef\]](#)
21. Leng, L.; Li, M.; Kim, C.; Bi, X. Dual-source discrimination power analysis for multi-instance contactless palmprint recognition. *Multimed. Tools Appl.* **2017**, *76*, 333–354. [\[CrossRef\]](#)
22. Desai, F.; Chowdhury, D.; Kaur, R.; Peeters, M.; Arya, R.C.; Wander, G.S.; Gill, S.S.; Buyya, R. HealthCloud: A system for monitoring health status of heart patients using machine learning and cloud computing. *Internet Things* **2022**, *17*, 100485. [\[CrossRef\]](#)
23. Shahidinejad, A.; Ghobaei-Arani, M.; Souri, A.; Shojafar, M.; Kumari, S. Light-edge: A lightweight authentication protocol for IoT devices in an edge-cloud environment. *IEEE Consum. Electron. Mag.* **2021**, *11*, 57–63. [\[CrossRef\]](#)
24. Kedia, S.; Bhushan, M. Prediction of mortality from heart failure using machine learning. In Proceedings of the 2022 2nd International Conference on Emerging Frontiers in Electrical and Electronic Technologies (ICEFEET), Patna, India, 24–25 June 2022; pp. 1–6.
25. Jazayeri, F.; Shahidinejad, A.; Ghobaei-Arani, M. Autonomous computation offloading and auto-scaling the in the mobile fog computing: A deep reinforcement learning-based approach. *J. Ambient. Intell. Humaniz. Comput.* **2021**, *12*, 8265–8284. [\[CrossRef\]](#)
26. Raju, K.B.; Dara, S.; Vidyarthi, A.; Gupta, V.M.; Khan, B. Smart Heart Disease Prediction System with IoT and Fog Computing Sectors Enabled by Cascaded Deep Learning Model. *Comput. Intell. Neurosci.* **2022**, *2022*, 1070697. [\[CrossRef\]](#)
27. Kong, G.; Lin, K.; Hu, Y. Using machine learning methods to predict in-hospital mortality of sepsis patients in the ICU. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 251. [\[CrossRef\]](#)
28. Dutta, A.; Batabyal, T.; Basu, M.; Acton, S.T. An efficient convolutional neural network for coronary heart disease prediction. *Expert Syst. Appl.* **2020**, *159*, 113408. [\[CrossRef\]](#)

29. Gao, F.; Thiebes, S.; Sunyaev, A. Rethinking the meaning of cloud computing for health care: A taxonomic perspective and future research directions. *J. Med. Internet Res.* **2018**, *20*, e10041. [[CrossRef](#)]
30. Choi, E.; Bahadori, M.T.; Song, L.; Stewart, W.F.; Sun, J. GRAM: Graph-based attention model for healthcare representation learning. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 13–17 August 2017; pp. 787–795.
31. Choi, E.; Xiao, C.; Stewart, W.; Sun, J. Mime: Multilevel medical embedding of electronic health records for predictive healthcare. In Proceedings of the 32nd Conference on Neural Information Processing Systems, Montréal, QC, Canada, 3–8 December 2018; pp. 4547–4557.
32. Dimitrov, D.V. Medical internet of things and big data in healthcare. *Healthc. Inform. Res.* **2016**, *22*, 156–163. [[CrossRef](#)]
33. Serviá, L.; Montserrat, N.; Badia, M.; Llompart-Pou, J.A.; Barea-Mendoza, J.A.; Chico-Fernández, M.; Sánchez-Casado, M.; Jiménez, J.M.; Mayor, D.M.; Trujillano, J. Machine learning techniques for mortality prediction in critical traumatic patients: Anatomic and physiologic variables from the RETRAUCI study. *BMC Med. Res. Methodol.* **2020**, *20*, 262. [[CrossRef](#)]
34. Yadaw, A.S.; Li, Y.c.; Bose, S.; Iyengar, R.; Bunyavanich, S.; Pandey, G. Clinical features of COVID-19 mortality: Development and validation of a clinical prediction model. *Lancet Digit. Health* **2020**, *2*, e516–e525. [[CrossRef](#)]
35. Perng, J.W.; Kao, I.H.; Kung, C.T.; Hung, S.C.; Lai, Y.H.; Su, C.M. Mortality prediction of septic patients in the emergency department based on machine learning. *J. Clin. Med.* **2019**, *8*, 1906. [[CrossRef](#)]
36. Yan, L.; Zhang, H.T.; Goncalves, J.; Xiao, Y.; Wang, M.; Guo, Y.; Sun, C.; Tang, X.; Jing, L.; Zhang, M.; et al. An interpretable mortality prediction model for COVID-19 patients. *Nat. Mach. Intell.* **2020**, *2*, 283–288. [[CrossRef](#)]
37. Harutyunyan, H.; Khachatryan, H.; Kale, D.C.; Ver Steeg, G.; Galstyan, A. Multitask learning and benchmarking with clinical time series data. *Sci. Data* **2019**, *6*, 96. [[CrossRef](#)] [[PubMed](#)]
38. Tokodi, M.; Schwertner, W.R.; Kovács, A.; Tóser, Z.; Staub, L.; Sárkány, A.; Lakatos, B.K.; Behon, A.; Boros, A.M.; Perge, P.; et al. Machine learning-based mortality prediction of patients undergoing cardiac resynchronization therapy: The SEMMELWEIS-CRT score. *Eur. Heart J.* **2020**, *41*, 1747–1756. [[CrossRef](#)] [[PubMed](#)]
39. Johnson, A.E.; Pollard, T.J.; Shen, L.; Lehman, L.w.H.; Feng, M.; Ghassemi, M.; Moody, B.; Szolovits, P.; Anthony Celi, L.; Mark, R.G. MIMIC-III, a freely accessible critical care database. *Sci. Data* **2016**, *3*, 160035. [[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.