Iran J Public Health, Vol. 52, No.1, Jan 2023, pp.175-183

Original Article

A Rule Based Intelligent Software to Predict Length of Stay and the Mortality Rate in Trauma Patients in the Intensive Care Unit

*Mitra Montazeri [1], Mehdi Ahmadinejad [2], *Mahdieh Montazeri [1,3], Kambiz Bahaadinbeigy [1], Mohadeseh Montazeri [4], Leila Ahmadian [1]*

1. *Medical Informatics Research Center, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran*
2. *Department of Anesthesiology, Shahid Bahonar Hospital, Afzalipour Faculty of Medicine, Kerman University of Medical Sciences, Kerman, Iran*
3. *Department of Health Information Sciences, Faculty of Management and Medical Information Sciences, Kerman University of Medical Sciences, Kerman, Iran*
4. *Department of Computer, Technical and Vocational University, Kerman, Iran*

***Corresponding Author:** Email: MahdiehMontazeri@yahoo.com

(Received 19 Apr 2021; accepted 20 Jun 2021)

Abstract

Background: Intensive Care Unit (ICU) has the highest mortality rate in the world. ICU has special equipment that leads to the hospital's most costly parts. The length of stay in the ICU is a special issue, and reducing this time is a practical approach. We aimed to use artificial intelligence to help early and timely diagnosis of the disease to help with health.

Methods: We designed a rule-based intelligent system to predict the length of stay and the mortality rate of trauma patients in ICU. A neuro-Fuzzy and eight machine learning models were used to predict the mortality rate in trauma patients in ICU. The performances of these techniques were evaluated with accuracy, sensitivity, specificity, and area under the ROC curve. Decision-Table was used to predict the length of stay in trauma patients in ICU. For comparison, eight machine learning models were used. The method is compared based on Mean absolute error and relative absolute error (%).

Results: Neuro-Fuzzy expert system and Decision-Table showed better results than other techniques. Accuracy, sensitivity, specificity, and ROC Area of Nero-Fuzzy are 83.6735, 0.9744, 0.3000, 0.8379, and 1, respectively. The mean absolute error and Relative absolute error (%) of the Decision-Table model are 4.5426 and 65.4391, respectively.

Conclusion: Neuro-Fuzzy expert system with the highest level of accuracy and a Decision-Table with the lowest Mean absolute error, which are rule-based models, are the best models. Therefore, these models are recommended as a valuable tool for prediction parameters of ICU as well as medical decision-making.

Keywords: Rule based intelligent software; Neuro-Fuzzy expert system; Decision-table model; Length of stay; Mortality rate

Introduction

The highest hospital mortality rates are in intensive care (1). In the United States, four million patients are admitted daily to the ICU, with an average mortality rate of 8-19%, or about 500,000 deaths a year (2). ICU Hospital is one of the most costly parts of the national health sector (3). These costs are mainly due to the long length of stay in the ICU (4). The most impressive way to reduce the cost of ICU is to reduce the length of stay. There is no straightforward and affordable way to control this issue (5).

Artificial intelligence refers to systems that can behave the same as intelligent human behaviors, including understanding complex situations and reasoning methods of human beings and their successful response, learning, and ability to acquire knowledge and reason for solving problems (6). The application of artificial intelligence in medicine is the main objective of the processing and analyzing medical information and communication between this medical information and the relevant users, based on the knowledge and experience of the operation of various systems in medicine and treatment. One of the most important uses of artificial intelligence in medicine is the recognition and diagnosis of diseases (7). This application includes diagnostic models using various decision-making methods. These models are based on the knowledge and experience of the given system, which in turn provides this information to the computer. After that, the model or system is compared and evaluated with that information or model. As a result, the difference or recognition of the type of variation in the model is shown in the model compared to the natural model. Such as identifying different patterns in medical images (8) and automatic diagnosis of diseases by signal (9-11) are among the other types of this group.

The optimal length of stay can be determined by choosing and optimizing mechanized methods for data collection, even before the patients enter hospitals. From the arrival of patients in different parts of the hospital until they leave, a significant percentage of the time and money is devoted to recording, processing, and storing medical information in handheld systems. While creating a comprehensive intelligent system not only minimizes the cost of information processing and storage, but also it provides quick access to disease records.

In this project, we aimed to design a rule-based intelligent system to predict the length of stay and the mortality rate of trauma patients in ICU. So far, several intelligent systems have been developed to help diagnose illnesses. Still, no rule based system has been reported to predict the length of stay and the mortality rate of trauma patients in ICU. Rule-based system design enables easy understanding, high flexibility, tolerance of inaccurate information, and the ability to model complex nonlinear arbitrary actions. This method is mixed with conditional control techniques and is based on simple conversational language.

Fuzzy system is the well-known intelligent system used in various applications such as medical or medicine applications (12,13). The heart of a fuzzy system is a knowledge base composed of fuzzy IF-THEN rules. A fuzzy IF-THEN rule is an if-then statement that continuous membership functions mark some words. The reasons for using the fuzzy system are simple, easy to understand, high flexibility, tolerance of inaccurate information, and the ability to model non-linear optional complicated applications.

Methods

This study was a cross-sectional study conducted on patients referred to the ICU of Shahid Bahonar Hospital at Kerman University of Medical Sciences, Kerman, Iran from 2010-2012. Sampling was calculated through available samples, and the sample size was 499, including the records of all patients admitted to the ICU at that time.

The selected knowledge sources to create a system for predicting the length of stay and the mortality rate of trauma patients in ICU include specialized articles related to the length of stay of trauma patients in the ICU and specialist and expert physicians. After targeted counseling with two available experts in our hospital, effective parameters in predicting the length of stay and the mortality rate of trauma patients admitted to ICU were identified. Moreover, parameters of decision-making with this topic and their degree of relationship were determined. Finally, software for entering information and predicting the length of stay and the mortality rate of patients were designed and implemented.

The data required for designing the system was collected from patient records of the study population and included demographic and clinical data such as age, sex, etc. The intelligent system was ran using this data, and the diagnosis result of the detection system was compared with the final diagnosis reported in the patient record.

In this research, software designed for this purpose was used to collect the data needed to record the information of traumatic patients admitted to ICU. The software includes patients' demographic information, physical signs, and paraclinical findings. The patient's demographic information included the age and gender of the patient. In the design and development of the system, an interview was conducted with two anesthetist physicians, with Special Care Fellowship, to obtain medical knowledge and identify important and effective parameters related to the identification of essential items in the prediction of the duration of stay and the mortality rate. At the same time, the software was provided for data collection for the design and evaluation phase of the Neuro-fuzzy intelligent system. Specialist doctors approved the software validity and efficiency. Then, the information of 499 patients was recorded in the software.

Data collection

All data used in the model were collected from the ICU patients' records in Shahid Bahonar Hospital. Clinical variables were recorded on a standard form that included patient information: Age, Sex, From Ward, Time of Stay, Urgent Surgery, M.Ventilator, History of Stay, T-°C, MAP, HR, RR, Art-PH, Na, K, Ser/Cr, HCT, WBC, and GCS.

Due to the presence of the Trauma ICU in Shahid Bahonar Hospital and the high number of medical records of patients in the medical records section, the research community included patient files that were admitted to the Hospital.

The data was analyzed by the experts and the software. The necessary pre-processing was done, and Matlab software was used for the analysis of the results. The plan exclusively uses the clinical information recorded in the patient's medical records, and according to the data collection software, it was not necessary to use patient identity data. Therefore, the patients' identities remain confidential at the time of the case examination, and satisfaction with the patients was not required.

Statistical analysis

The proposed method used the cross-validation of k-fold layer for statistical tests. The statistical program in this study was Weka 3.8 (14). It was free software licensed under the GNU General Public License.

In this method, the data were partitioned into *k* subsets. The K-fold cross validation cross-validation holds its advantages, as the method tends to be less biased than other method (15). In each *k* iteration, each subset was used for validation, and *k*-1 others were used for training. This procedure was repeated *k* times, and all data were used exactly *k* times for training and testing. Finally, a mean *k*-time validation result was elected as a final estimate value. In this study, cross-validation of 10-folds was used which is commonly used (16).

With respect to splitting the data into training/test (or train/validation/test) vs K-fold cross validation, it depended on the amount of data that we had, and how well this data represented the distribution of the data where we wanted to apply the model. In an ideal world, we wanted to have an independent test set to verify the performance of our model. Sometimes the dataset was not big enough to be split into training and test set with those characteristics, so people use cross-validation to use as much data as possible for both, training and testing.

Results

Among the patients, 68% (n=343) were male. The Mean age of the patients was 46.5 years old, and about 60% of them (n=301) were over 40 years old. More than half of the patients used a ventilator during the treatment process. The Mean of GCS was 8.6 (SD=4.2). Sixty-eight percent of the patients had no history of hospitalization, and more than half had urgent surgery. The mean length of ICU stay was 7.15 days (SD =11.3).

The data set included 499 patient records in which 398 patients were alive, and 101 patients were dead. Thus, the target variable was divided into two groups: alive or dead.

Performance of the proposed model

Neuro-Fuzzy expert system and eight models of machine learning NB, Naïve Bayes, Logistic regression, SVM, 1NN, AdaBoost, Trees Random Forest, RBF Network, and Multilayer Perceptron were used to predict the mortality rate in trauma patients in ICU.

Neuro-Fuzzy expert system and machine learning models were implemented, and the comparison results are summarized in Table 1. The ROC curve values from 0 to 0.5 represents random classification, and 0.5 to 1 indicates that the model has a general diagnostic ability.

Decision-Table was used to predict the length of stay in trauma patients in ICU. Eight machine learning models were used for comparison: Ensemble Selection, SVM, M5Rules, RBF Network, 1NN, Multilayer Perceptron, Linear Regression, and K-Star. The comparison results based on Mean absolute error and Relative absolute error (%) were reported in Table 2. In each column, the best result was bolded. Decision-Table had the minimum error value and was introduced as the best model in all evaluation merits.

Table 1: The performance of Neuro-Fuzzy expert system and various machine learning models

NO.	*Model*	*ACC*	*Sensitivity*	*Specificity*	*Weighted Avg.*	*ROC Area*
1.	Nero-Fuzzy	**83.6735**	**0.9744**	0.3000	**0.8379**	**1**
2.	Naïve Bayes	81.3627	0.907	0.446	0.814	0.7975
3.	Logistic regression	78.9579	0.874	**0.455**	0.79	0.664
4.	SVM	80.1603	0.915	0.356	0.802	0.636
5.	1NN	75.5511	0.869	0.307	0.756	0.588
6.	AdaBoost	82.1643	0.96	0.277	0.822	0.747
7.	Trees Random Forest	81.1623	0.955	0.248	0.812	0.712
8.	RBF Network	81.7635	0.925	0.396	0.818	0.79
9.	Multilayer Perceptron	77.1543	0.882	0.337	0.772	0.716

Table 2: The performance of Decision Table and various machine learning models

NO.	*Model*	*Mean absolute error*	*Relative absolute error (%)*
1.	Decision-Table	4.5426	65.4391
2.	Ensemble Selection	4.6456	66.9233
3.	SVM	4.7936	69.0553
4.	M5Rules	4.6934	67.6117
5.	RBF Network	6.6233	95.4129
6.	1NN	6.9839	100.6074
7.	Multilayer Perceptron	5.7813	83.2839
8.	Linear Regression	4.9879	71.8545
9.	KStar	6.2485	90.014

Two different feature selection methods (Ranker with relief attribute and Scatter search with CFS subset evaluation) were used to evaluate the role of different features in prediction performance. Table 3 and 4 show the performance of various feature selection methods.

In the Multilayer Perceptron model, we set the learning rate to 0.3, momentum to 0.2, training time to 499, and validation threshold to 20. Multilayer Perceptron used back-propagation to classify instances. The nodes in this network were all sigmoid (except for when the class is numeric, in which case the output nodes become unthreshold linear units). In Naïve Bayes classifier using a default precision of 0.1 for numeric attributes when Classifier is built with zero training instances. In the Trees Random Forest model, the maximum depth of Trees Random Forest was set to 0, and the 10 number trees were to be generated. 1NN model used normalized Euclidean distance to find the training instance closest to the given test instance and predicts the same class as this training instance. If multiple instances had the same (smallest) distance to the test instance, the first one found was used.

Table 3: The performance of Ranker with relief attribute Feature selection method

Classifier model	*Mean absolute error*	*Relative absolute error (%)*
Decision-Table	4.5242	651741
Ensemble Selection	4.6619	671583
SVM	**4.4975**	**64.7893**
M5Rules	4.6761	673628
RBF Network	6.7209	96.8198
1NN	5.6345	81.1684
Multilayer Perceptron	5.4902	79.0896
Linear Regression	4.7142	67.9118
KStar	5.7162	82.3462

Table 4: The performance of Scatter search with CFS subset evaluation Feature selection method

Classifier model	*Mean absolute error*	*Relative absolute error (%)*
Decision-Table	4.5242	65.1741
Ensemble Selection	4.6357	66.7808
SVM	4.5483	65.521
M5Rules	4.5551	65.6197
RBF Network	6.6699	96.0844
1NN	6.0361	86.9548
Multilayer Perceptron	5.1753	74.5539
Linear Regression	4.709	67.8364
KStar	5.7667	83.73

The AdaBoost model's base classifier was Decision Stump. The number of iterations to be performed was set at ten and the weight threshold for weight pruning was set at 100. In the SVM model, the complexity parameter was set to 1, and the epsilon for round-off error (should not be changed) was set to 1.0E-12. The kernel function used in this study was the polynomial kernel. The tolerance parameter (which should not be changed) was set to 0.001.

RBF Network model used the value of 0.1 to minimize the standard deviation for the clusters. The number of clusters for K-Means to generate and the ridge value for the logistic or linear regression was set to 2 and 1.0E-8, respectively.

Discussion

In this study, a rule-based intelligent system to predict the length of stay and the mortality rate of trauma patients in ICU was designed. The proposed Neuro-Fuzzy expert system and Decision-Table method were compared with eight machine learning models.

A rule-based intelligent system, a Neuro-Fuzzy expert system, and Decision-Table showed better results than other techniques. This study demonstrated that a Neuro-Fuzzy expert system with the highest level of accuracy and Decision-Table with the lowest Mean absolute error which were rule-based models were the best models. Therefore, this model was recommended as a useful tool for ICU prediction as well as medical decision-making.

All the resources that somehow point to this innovation can be divided into four categories: mathematical, statistical, data extraction, and multistage. The first group uses mathematical methods to calculate the length of stay in the ICU (17). Conventional mathematical methods are usually computed as mean or median (18). However, this is a straightforward and new way to calculate the length of stay in the ICU. As has been proven in the source (19) how these methods are misleading. The second category includes statistical methods such as logistic regression, linear regression, and covariance (20). In the source (21), we used a regression model to analyze the relationship of several variables determining the length of stay. This article has shown that age, type of admission, and type of hospital effectively correlate with length of stay. Scoring methods are also part of this method. APACHE, for example, is a scoring system introduced in 1981, and in 1985 an edited version of it was published as APACHE II. In addition to this SAPS scoring method, Le Gall (22) used logistic regression to develop SAPS II in 1993.

The third batch method (23) is based on data mining techniques such as sorting (24,25), clustering, etc., to predict residence time and mortality. These techniques seek to fit into a huge database by considering the relationships between variables. Neural networks and logistic regression were used to predict mortality in the ICU (25). The SVM method performs better than APACHE II, while in a study (26), ANN, SVM, and DT performed better than the APACHE III scoring system on ICU patients. Neural network performance showed a logistic regression ratio (27). Better neural network and SVM performance (26). Machine learning techniques were used to predict the length of stay and mortality (28).

The fourth category uses multiple modes and a random procedure for transitions between states (29). The Coxian method falls into this category. In this way, a demonstration of the continuous duration of the patient's stay in the hospital is described as a series of sequential steps that the patient leaves until the hospital leaves (30). A conditional multi-stage distribution was used for the model of patient length of stay (31). Time Slicing Cox regression, an extended form of Cox regression, was used to predict mortality in the ICU (32).

During the evaluation, the Neuro-Fuzzy expert system had the following significant performances:

- An intelligent approach with simplicity.
- Easy to understand and implement.
- Ease of implementation.
- Provide more “user-friendly” and efficient performance.
- capacity to represent inherent uncertainties of the human knowledge with linguistic variables;
- simple interaction of the expert of the domain with the engineering design of the system;
- easy interpretation of the results because of the representation of the natural rule;

- easy extension of the base of knowledge through the addition of new rules;
- robustness about the possible disturbances in the system
- Fast learning; online adaptability; self-adjusting to obtain the small global error possible; slight computational complexity.

Moreover, Decision-Table had the following prominent performances (33):

- Tables were easier to draw up than comparable flow charts. They were easier to change since it was relatively simple to add conditions, rules, and actions to a table.
- Tables forced the programmer to think the problem through. For example, if there were three conditions to be considered, each answered yes or no, then there are 23 or 8 possible paths or rules. Some of these possible paths migth not be pertinent to the problem. However, by knowing the total number of paths, the programmer lessened the danger of forgetting one.
- Several pages of flowcharting migth be condensed into a tiny table and of course, it was easier to follow a particular flow path down one column than it was to follow the same path through several flow chart pages.
- Tables could perform a valuable communication function. An analyst maigth design a new system and present it in the form of a table or tables to other analysts, programmers, managers, and executives. Others easily followed the table format.
- Flow charts, and symbols, on the other hand, were not always standardized, and this factor may hinder their communication value. Tables appeared to be easier for many managers to follow than flow charts. Operating managers could quickly trace and verify those paths in the procedure that ware most interested in them.
- Tables were easier to draw up than comparable flow charts. They were also easier to change since adding conditions, rules, and actions to a table is relatively simple.

Conclusion

The Neuro-Fuzzy expert system had the highest level of accuracy, and Decision-Table had the lowest Mean absolute error. Both methods were rule-based models as well as the best models. Therefore, these models were recommended as a valuable tool for prediction parameters of ICU as well as medical decision-making.

Journalism Ethics considerations

Ethical issues (Including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the authors.

Acknowledgements

Kerman University of Medical Sciences is the financial source of the study.

Conflict of interest

The authors declare that there is no conflict of interests.

References

1. Garland A, Olafson K, Ramsey CD, Yogendran M, Fransoo R (2015). A population-based observational study of intensive care unit–related outcomes. With emphasis on post-hospital outcomes. *Ann Am Thorac Soc*, 12(2): 202-208.
2. Mukhopadhyay A, Tai BC, See KC, et al (2014). Risk factors for hospital and long-term mortality of critically ill elderly patients admitted to an intensive care unit. *BioMed Res Int*, 47: 777-780.

3. Rotta BP, Silva JMd, Fu C, Goulardins JB, Pires-Neto RdC, Tanaka C (2018). Relationship between availability of physiotherapy services and ICU costs. *J Bras Pneumol*, 960575.
4. Rapoport J, Teres D, Lemeshow S, Gehlbach S (1994). A method for assessing the clinical performance and cost-effectiveness of intensive care units: a multicenter inception cohort study. *Crit Care Med,* 22(9):1385-91.
5. Marini JJ, Wheeler AP (2010). *Critical Care Medicine*: The Essentials: Wollters Kluwer/Lippincott Williams & Wilkins.
6. McLeod R, Schell GP (2007). *Management information systems*. Pearson/Prentice Hall USA.
7. Nguyen TP, Ho TB (2012). Detecting disease genes based on semi-supervised learning and protein–protein interaction networks. *Artif Intell Med* , 54(1):63-71.
8. Latacz E, van Dam PJ, Vanhove C, et al (2021). Can medical imaging identify the histopathological growth patterns of liver metastases? S*emin Cancer Biol*, 71: 33-41.
9. Brown C, Chauhan J, Grammenos A, et al (2020). Exploring automatic diagnosis of covid-19 from crowdsourced respiratory sound data. Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.
10. Baghel N, Dutta MK, Burget R (2020). Automatic diagnosis of multiple cardiac diseases from PCG signals using convolutional neural network. *Comput Methods and Programs Biomed*, 197:105750.
11. Salah IB, De la Rosa R, Ouni K, Salah RB (2020). Automatic diagnosis of valvular heart diseases by impedance cardiography signal processing. *Biomed Signal Process Control*, 57:101758.
12. Maldonado S, López J, Jimenez-Molina A, Lira H (2020). Simultaneous feature selection and heterogeneity control for SVM classification: An application to mental workload assessment. *Expert Syst Appl*, 143:112988.
13. Luo M, Zhao R (2018). A distance measure between intuitionistic fuzzy sets and its application in medical diagnosis. *Artif Intell Med*, 89:34-39.
14. Borra S, Di Ciaccio A (2010). Measuring the prediction error. A comparison of cross-validation, bootstrap and covariance penalty methods. *Comput Stat Data Anal*, 54(12):2976-89.
15. Molla MU, Giri BC, Biswas P (2021). Extended PROMETHEE method with Pythagorean fuzzy sets for medical diagnosis problems. *Soft Comput*, *25*(6), 4503-4512.
16. McLachlan, Geoffrey J, Kim-Anh Do, Christophe Ambroise (2005). *Ambroise Analyzing microarray gene expression data.* Wiley Ser Probab Stat.
17. Cheng T-H, Hu PJ-H (2009). A data-driven approach to manage the length of stay for appendectomy patients. *IEEE Trans Syst Man Cybern B Cybern*, 39(6):1339-47.
18. Awad A, Bader–El–Den M, McNicholas J (2017). Patient length of stay and mortality prediction: A survey. *Health Serv Manage Res*, 30(2):105-20.
19. Garg L, McCLEAN S, Meenan BJ, Millard P (2011). Phase-type survival trees and mixed distribution survival trees for clustering patients' hospital length of stay. *Informatica*, 22(1):57-72.
20. Guzman Castillo M. Modelling patient length of stay in public hospitals in Mexico [PhD thesis]. University of Southampton, Malaysia ; 2012.
21. Freitas A, Silva-Costa T, Lopes F, et al (2012). Factors influencing hospital high length of stay outliers. *BMC Health Serv Res*, 12(1):265.
22. Le Gall JR, Lemeshow S, Saulnier F (1993). A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study. JAMA, 270: 2957-2963.
23. Azari A, Janeja VP, Mohseni A (2012). Predicting hospital length of stay (PHLOS): A multi-tiered data mining approach. Data Mining Workshops (ICDMW), 2012 IEEE 12th International Conference on; IEEE.
24. Dou L, Li X, Ding H, Xu L, Xiang H (2020). iRNA-m5C_NB: a novel predictor to identify RNA 5-Methylcytosine sites based on the Naive Bayes classifier. IEEE Access, 8:84906-17.
25. Xu F, Pan Z, Xia R (2020). E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework. *Inf Process Manag*, 57(5):102221.

26. Dumitru D (2009). Prediction of recurrent events in breast cancer using the Naive Bayesian classification. *Annals of the University of Craiova-Mathematics and Computer Science Series*, 36(2):92-6.
27. Breiman L (2001). Random forests. *Machine Learning*, 45(1):5-32.
28. Ferenc G, Lutovac M, Kvrgic V, Stepanic P (2013). A proposed approach to the classification of bearing condition using wavelets and random forests. Embedded Computing (MECO), 2013 2nd Mediterranean Conference on, IEEE.
29. Fathi M, Nemati M, Mohammadi SM, Abbasi-Kesbi R (2020). A machine learning approach based on SVM for classification of liver diseases. *Biomed Eng*, 32(03):2050018.
30. Yasoda K, Ponmagal R, Bhuvaneshwari K, Venkatachalam K (2020). Automatic detection and classification of EEG artifacts using fuzzy kernel SVM and wavelet ICA (WICA). *Soft Computing*, 24(21):16011-9.
31. Sha'abani M, Fuad N, Jamal N, Ismail M (2020). kNN and SVM classification for EEG: a review. *InECCE,* 2019, 555-65.
32. Tabrizi SS, Pashazadeh S, Javani V (2020). Comparative study of table tennis forehand strokes classification using deep learning and SVM. *IEEE Sens J*, 20(22):13552-61.
33. Priyali S. Decision Tables of a Programmer: Meaning, Types and Advantages: yourarticlelibrary. Available from: www.yourarticlelibrary.com