Acute and Critical Care

Acute and Critical Care 2022 February 37(1):45-52
https://doi.org/10.4266/acc.2021.00486

| pISSN 2586-6052 | eISSN 2586-6060

A machine learning model for predicting favorable outcome in severe traumatic brain injury patients after 6 months

Mehdi Nourelahi[1], Fardad Dadboud[2], Hosseinali Khalili[3], Amin Niakan[3], Hossein Parsaei[4,5]

[1]Department of Computer Science, University of Wyoming, Laramie, WY, USA; [2]Department of Electrical Engineering, Sharif University of Technology, Tehran; [3]Trauma Research Center, Shahid Rajaee (Emtiaz) Trauma Hospital, Department of Neurosurgery, Shiraz University of Medical Sciences, Shiraz; [4]Department of Medical Physics and Engineering, School of Medicine, Shiraz University of Medical Sciences, Shiraz; [5]Shiraz Neuroscience Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

Background: Traumatic brain injury (TBI), which occurs commonly worldwide, is among the more costly of health and socioeconomic problems. Accurate prediction of favorable outcomes in severe TBI patients could assist with optimizing treatment procedures, predicting clinical outcomes, and result in substantial economic savings.

Methods: In this study, we examined the capability of a machine learning-based model in predicting "favorable" or "unfavorable" outcomes after 6 months in severe TBI patients using only parameters measured on admission. Three models were developed using logistic regression, random forest, and support vector machines trained on parameters recorded from 2,381 severe TBI patients admitted to the neuro-intensive care unit of Rajaee (Emtiaz) Hospital (Shiraz, Iran) between 2015 and 2017. Model performance was evaluated using three indices: sensitivity, specificity, and accuracy. A ten-fold cross-validation method was used to estimate these indices.

Results: Overall, the developed models showed excellent performance with the area under the curve around 0.81, sensitivity and specificity of around 0.78. The top-three factors important in predicting 6-month post-trauma survival status in TBI patients are "Glasgow coma scale motor response," "pupillary reactivity," and "age."

Conclusions: Machine learning techniques might be used to predict the 6-month outcome in TBI patients using only the parameters measured on admission when the machine learning is trained using a large data set.

Key Words: artificial intelligence; favorable outcome; neuromonitoring; neurotrauma; traumatic brain injury

INTRODUCTION

Traumatic brain injury (TBI) is one of the most common and costly health and socioeconomic problems worldwide. The incidence of TBI is higher than that of complex diseases such as breast cancer, AIDS, multiple sclerosis, and Parkinson disease, such that TBI is considered the leading cause of mortality and disability amongst individuals under 45 years of age. This brain injury is the cause of approximately 50,000 deaths per year in the United States [1–3].

Original Article

Received: June 21, 2021
Revised: August 6, 2021
Accepted: August 24, 2021

Corresponding author
Hossein Parsaei
Department of Medical Physics and Engineering, School of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran
Tel: +98-71-3234-9332
Fax: +98-71-3234-9332
E-mail: hparsaei@sums.ac.ir

It is reported that around 10%–15% of patients with TBI need specialist care leading to high costs for both individuals and society. Accurate assessment and classification of patients with TBI could assist with diagnosing, optimizing treatment procedures, improving clinical outcomes, and may result in substantial economic savings.

A common measure for long-term functional outcomes in TBI patients is the Glasgow outcome scale (GOS) or its extended version extended Glasgow outcome scale (GOSE), which consists of eight ordered categories [4]. These measures are acknowledged to be standard means of describing outcomes in head injury patients due to several advantages such as their reliability, validity, stability, simplicity, availability, and ease of access. Considering the costs of TBI, developing a model capable of predicting favorable or unfavorable GOS/GOSE outcomes in advance might assist the clinicians to optimize the management and treatment of this injury.

Several prediction models using machine learning tools were developed to predict GOS based on medical image modalities [5–8]. Focusing on a specific age group, Hale et al. [5,6] used imaging techniques that display many strengths, such as consistency of computed tomography or ubiquitous usage of magnetic resonance imaging, as well as the multimodal techniques based on this imaging techniques. Nonetheless, imaging data quality might be uneven, have variabilities in data types, atlases, interpretations, and user specifics [9]. Folweiler et al. [10] utilized clustering, which deals with the challenge of picking a confident evaluation method. On the other hand, a combination of machine learning methods and information datasets of patients, including the GOS parameter, was used for providing the prediction models in some of the previous works [11–15]. Employing crucial parameters for outcome prediction is absent in [12,14], while the sample size in [13] is not sufficiently high and hence might not be reliable enough. Moreover, in Eftekhar et al.'s study [15], a specific portion of data was considered, and the time of mortality evaluation was not mentioned clearly. Conveniently, reporting 100% sensitivity for mortality prediction creates an unrealistic picture of the model performance and is a sign of a severe overfitting problem in [11]. In two studies [16,17], authors only focused on feature selection methods, and important features and classification models were not explored in these works. One major problem in [16] is the value of 58.9% for sensitivity that is lower than existing standards; another is the 89.2% specificity that shows unbalanced accuracy for different groups. A reliable model to predict the outcome of TBI could be used as a decision aid system for prognosis, treatment, and neurosurgery with better quality.

KEY MESSAGES

- Machine learning-based models could be reliable in predicting the 6-month outcome in traumatic brain injury (TBI) patients when they are trained using a large data set.
- Logistic regression model is superior to support vector machin and random forest models due to its simplicity and interpretability.
- Three variables of "age," "Glasgow coma scale motor response," and "pupillary reactivity" are important in predicting the 6-month outcome in TBI patients.

In this research, we explored the possibility of developing a machine learning-based model for predicting "favorable" or "unfavorable" outcome after 6 months in severe TBI patients. The parameters measured at admission were used as features (predictors), and three different machine learning models were examined.

MATERIALS AND METHODS

The procedures used in this study were performed in accordance with the ethical standards of the Ethics Committee of Human Experimentation of Shiraz University of Medical Sciences and in accordance with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Written informed consent was obtained from the participants included in the study. The study protocol has been approved by the Ethics Committee of Human Experimentation of Shiraz University of Medical Sciences (SUMS), Shiraz, Iran (Ethic #IR.SUMS.REC.1397.296).

The objective of this work was to develop a model capable of predicting GOSE for TBI patients after 6 months. Precisely, the model was to predict and decide whether GOSE will be "favorable" or "unfavorable" for a patient 6 months after severe TBI. We considered GOSE > 4 as "favorable" and GOSE ≤4 as "unfavorable." The process consisted of three steps: data collection, model development, and evaluation.

Data Collection

Nine features/variables from 2,381 patients admitted to the neuro-intensive care unit of Rajaee (Emtiaz) Hospital (Shiraz, Iran) between 2015 and 2017 were collected and used as pre-

dictors. These features were parameters measured upon admission including age, sex, Rotterdam index, blood sugar (BS) level, prothrombin time-international normalized ratio (PT-INR), GCS motor response, and systolic blood pressure (SBP). The extended Glasgow outcome scale (GOSE) was measured by experts six 6 months after severe TBI trauma.

The GOSE was subsequently categorized into "unfavorable" (GOSE <5 meaning dead, vegetative state, or severe disability) or "favorable" (GOSE > 4 meaning mild disability or full recovery) outcomes. Ultimately, 470 cases (27.9%) were considered as "unfavorable" and 1,212 cases (72.1%) were categorized as "favorable."

Model Development

The model development process consisted of two steps: feature selection and classification. Feature selection was used to identify and remove inappropriate variables from the feature set. The remaining subset of variables was included in the model as the predictors. For this purpose, we employed the sequential forward selection (SFS) technique in which features were serially added to the model one at a time until there were no remaining candidate features that significantly improved the model performance. We used the classification accuracy as the objective function (the wrapper method). The SFS algorithms are discussed in [18]. To construct the predicting model, three well-known algorithms were explored: logistic regression (LR), random forest (RF), and support vector machines.

Logistic regression

LR is a classification algorithm (statistical model) which estimates the probability that a given sample belongs to a certain class. The estimation is conducted using a sigmoid (logistic) function applied to the weighted average of the predictors. The learning procedure is finding the value of these weights [19]. For binary classification, the class label is determined using the thresholding techniques with a threshold value of 0.5 in general. However, the threshold value can be adjusted to balance the sensitivity or specificity when needed.

Random forest

RF is a kind of ensemble classifier based on a different combination of decision trees. Apart from the standard trees, where each node is split based on the most promising splits among all the variables, RF splits each node using the best variable. The best feature is chosen from a subset of randomly assembled variables. This concept has made RF robust against overfitting. The process of developing a RF classifier is well discussed in [20,21].

Support vector machines

Support vector machines emphasize seeking a decision boundary that minimizes the generalization error. This is achieved by maximizing margin. For the problems with nonlinearly separable classes, the data is mapped onto a new space using kernels such that in the new space the data is linearly separable and the SVM optimization problem will fruitfully yield a separating hyperplane [19,22].

Evaluation

The performance of the proposed predicting model was assessed using three performance indices including sensitivity, specificity, and accuracy. Additionally, the area under the curve (AUC) of the receiver operating characteristic (ROC) was used to evaluate the capability of the predictors in discriminating the 6-month post-trauma survival status of the patients. In estimating these indices, we used the k-fold cross-validation procedure to ensure unbiased estimation and ultimately unbiased evaluation of the developed prediction models. The given dataset was randomly divided into k subsets, and then the classifier was trained using $k-1$ subsets and was tested on the remaining subset. This procedure was repeated k times.

RESULTS

The mean age of the patients with TBI was 39.41 years and the majority of patients were male (up to 83%). Approximately 29% of the cases contained at least one missing value in a variable. Out of 2,381 cases analyzed in this study, only 1,682 (up to 71%) cases had full parameters; for the remainder of cases some value for some parameter was missing. After removing cases with missing values, we ended up with 1,682 cases. A summary of the data set is provided in Table 1. The distribution of each variable in both "favorable" and "unfavorable" groups is provided in Figure 1. In this figure, for the six variables of age, BS level, SBP on admission, GCS motor response, PT-INR, and Rotterdam index, within-group boxplots were provided. However, for the categorical variable "pupil reactivity," for each category of anisocoria (A), brisk (B), and fixed (F) the proportion of favorable cases is calculated.

The results of the two employed feature selection methods, step-wise and random forest, are provided in Table 2 along with summaries of the most important features found by each

Table 1. Baseline patient characteristics

Variable	Value
Sex (male:female)	1,414:268
Pupil reactivity	
Anisocoria	121
Brisk	1,301
Fixed	260
N (unable to exam)	0
NA	161
Final GOSE	
1 (Death)	365
2	44
3	20
4	41
5	70
6	107
7	328
SBP	
Mean (range)	128.9 (40–250)
NA	338
BS	
Mean (range)	164.7 (53–681)
NA	100
PT-INR	
Mean (range)	1.3 (0.8–16.2)
NA	190
Age (yr)	
Mean (range)	39.4 (14–96)
NA	14
GCS motor	
Mean (range)	4.7 (1–6)
Rotterdam	
Mean (range)	2.5 (1–6)
NA	131

N: none; NA: not availible; GOSE: extended Glasgow outcome scale; SBP: systolic blood pressure; BS: blood sugar; PT-INR: prothrombin time-international normalized ratio; GCS: Glasgow coma scale.

technique. The performance of the three prediction models estimated using 10-fold cross-validation method is summarized in Table 3. Pair-wise comparison of the mean values was conducted using a multivariate analysis of variance at a 5% significance level and the Tukey-Kramer honestly significant difference test. The results of the ROC and AUC analyses are shown in Figure 2.

Table 2. Variables (features) selected using the feature selection methods examined

Technique	Random forest	Stepwise
Variable	Age, GCS motor response, pupil reactivity, BS level	Age, GCS motor response, pupil reactivity, Rotterdam index, PT-INR

GCS: Glasgow coma scale, BS: blood sugar; PT-INR: prothrombin time-international normalized ratio.

Table 3. Performance of the machine learning-based prediction models developed for predicting unfavorable outcome in severe TBI patients

Model	Accuracy	Sensitivity	Specificity
Logistic regression	0.78±0.01	0.78±0.01	0.78±0.03
Random forest	0.78±0.01	0.79±0.01	0.78±0.03
Support vector machines	0.78±0.01	0.78±0.01	0.78±0.04

Values are presented as mean±standard error.
TBI: traumatic brain injury.

DISCUSSION

This work aimed to use machine learning techniques in developing a model to predict the 6-month clinical outcome using baseline parameters and brain-specific imaging parameters in severe TBI patients. The results shown in Table 3 and Figure 2 indicate that the three examined models performed well for this task. Statistical analyses of the performance of the developed methods showed that there is no significant difference between the estimation ability of the examined methods. However, the LR model can be preferred over the two other methods (i.e., SVM and RF) as it is simple and easily interpretable. Further, the model could be used to estimate the chance of un/favorable outcome for the patients; this could be more beneficial than estimating just the binary class label.

The predicted models were more accurate than other works [11,16] in predicting "unfavorable" outcomes. Our models also provided comparable AUC to several algorithms [23,24]. Moreover, the models presented in this paper compromised between specificity and sensitivity to yield a fair prediction for both groups at the same time. Nevertheless, the performance of a prediction model depends on the data used for training and testing the algorithm and it is hard to have a fair comparison between several models when the data is different and is not available. It should be noted that our goal was to achieve a balanced specificity and sensitivity using simple and easy-to-measure features provided at the early stage of hospitalization, while some previously developed algorithms [25] employed

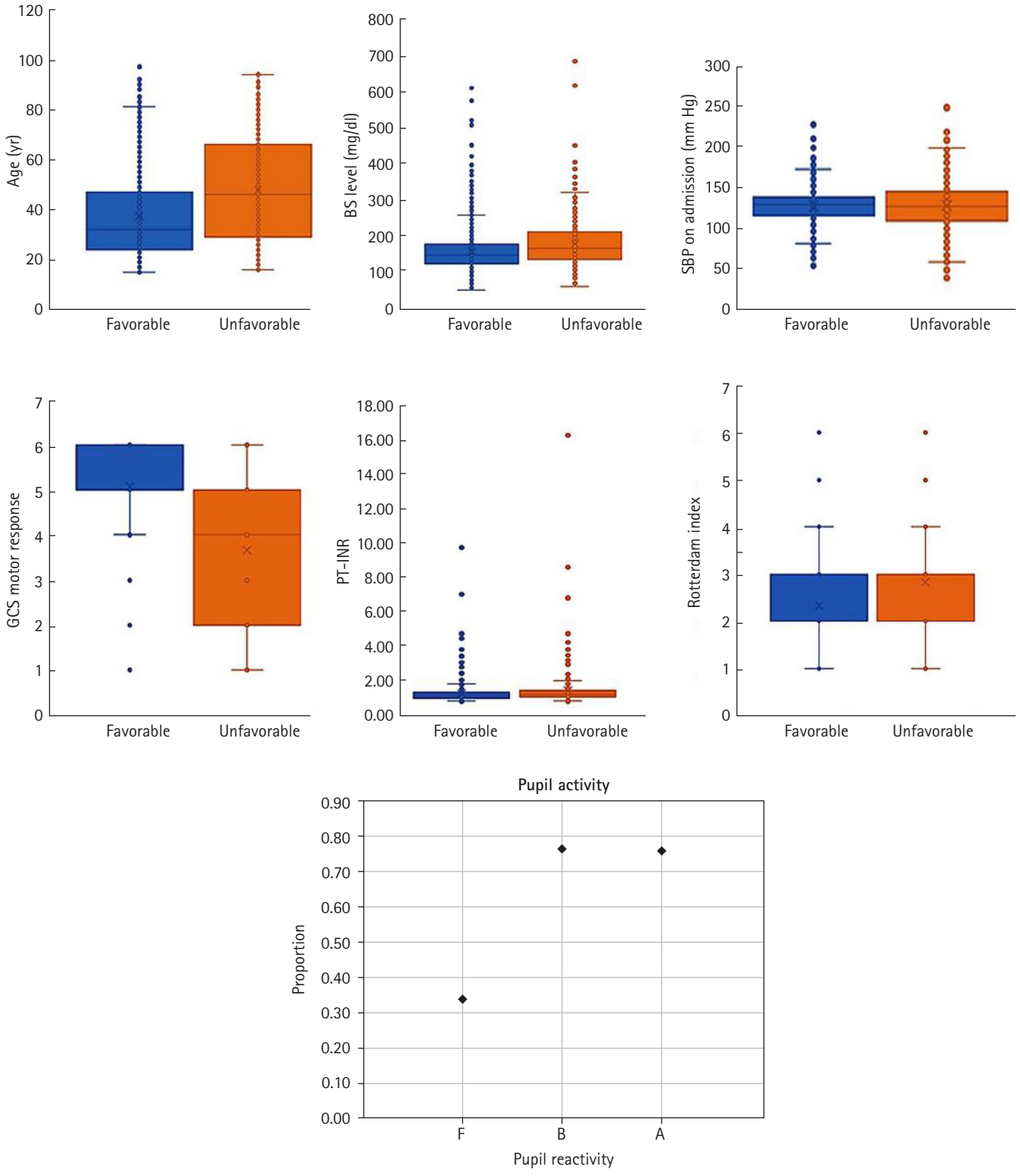

Figure 1. Boxplot for the six variables age, blood sugar (BS) level, systolic blood pressure (SBP) on admission, GCS motor response, coagulation measures prothrombin time-international normalized ratio (PT-INR), and Rotterdam index in "favorable" and "unfavorable" groups. For the categorical variable "pupil activity", for category anisocoria (A), brisk (B), and fixed (F) the proportion of favorable cases is calculated individually.

features from physiological signals (e.g., electroencephalography) which are costly, computationally expensive, and hard to measure. More importantly, we examined interpretable machine learning algorithms (e.g., logistic regression) because for medical applications interpretability of the model is of the essence. Our results showed that by using a set of simple features and interpretable machine learning models trained on large data, it is possible to develop a reliable GOSE predictive model.

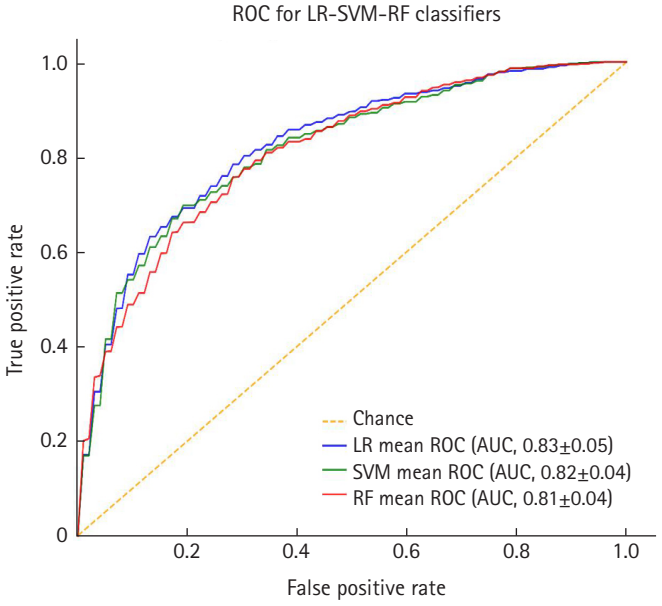

Figure 2. Mean receiver operating characteristic (ROC) curves and area under the curve (AUC) values for the prediction models developed for predicting unfavorable outcome after 6 months in the patients with severe traumatic brain injury. Values are presented as mean±standard error. LR: logistic regression; SVM: support vector machin; RF: random forest.

In terms of the parameters selected by the models, the three variables of "age," "GCS motor response," and "pupil reactivity" were found to be important by both feature selection models and were included in all the models; this finding is in line with previous works [11,16,17]. This result is consistent with the results provided in Figure 1. As shown, there is a shift in the average value between the GOSE categories for the two variables of "age" and "GCS motor response." For the pupil reactivity, the majority (>76%) of cases with category brisk (B) or anisocoria (A) were "favorable" cases. SBP was included in the RF model which is consistent with clinical practice in which blood pressure assessment is part of treatment protocols for severe TBI patients. Besides, it was reported that low SBP before and during ICU intensive care unit admission are linked with high mortality rates in TBI patients. Overall, the parameters found by the RF are more consistent with clinical practice than the ones selected by SFS.

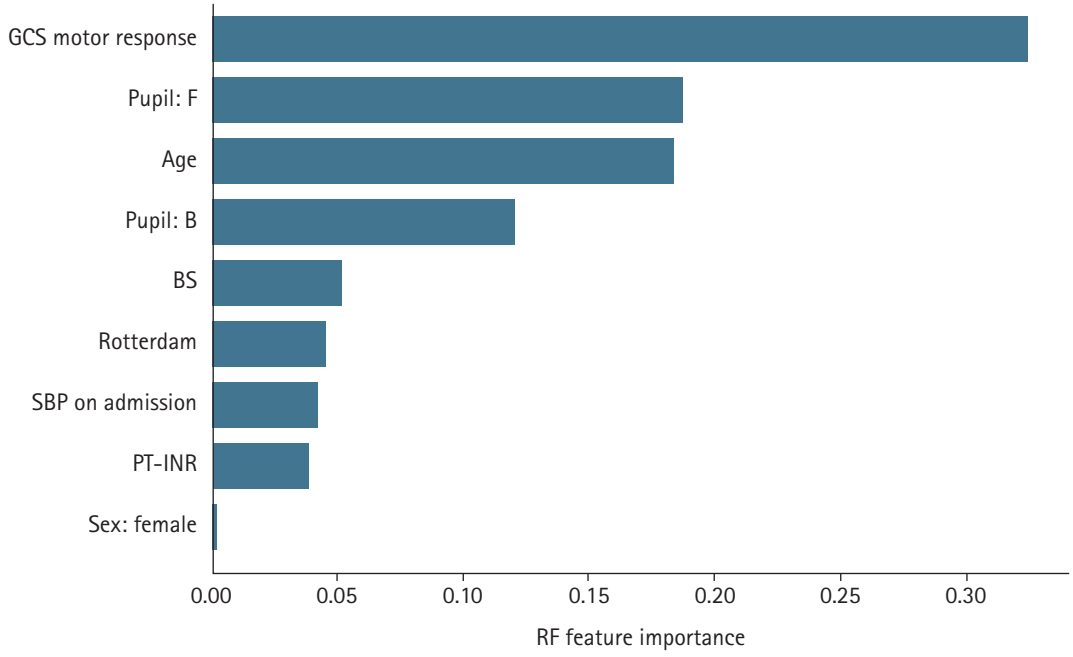

Figure 3. The relative importance of the variables used in random forest (RF)-based prediction model. The higher the value, the more important the feature is to the predicting model. GCS: Glasgow coma scale; F: fixed; B: brisk; BS: blood sugar; SPB: systolic blood pressure; PT-INR: prothrombin time-international normalized ratio.

To conduct an in-depth analysis of the role of features (variables) in predicting GOSE, we measured the importance of the parameters included in the predictor models. The feature importance for RFs was estimated using the Gini importance-based method. The average values of impurity decrease in the RF's structure were computed for all the features as the feature importance. Additionally, we divided the non-numeric parameters, such as pupil and sex, into 0/1 bit parameters according to their unique values. The result is provided in Figure 3. As shown, the "GCS motor response" is ranked as the most important feature. The value of importance of this parameter is > 0.3. The analysis also revealed that "pupil: F" is the second most important feature followed by age and BS. The results are in line with those provided in Figure 1. The box plots of the two variables "age" and "GCS motor response" indicate that there is a shift in the average values of these two variables between the group with "favorable" GOSE and "unfavorable" GOSE. For the categorical variable "pupil reactivity," the majority (>76%) of cases with brisk pupillary response were "favorable" cases. Normally, pupils react to light briskly. However, patients with increased intracranial pressure may show sluggish or slow pupillary response. Fixed (nonreactive) pupils are in general associated with severe brain damage or high intracranial pressure value which is consistent with our results in Figure 1; approximately 65% of cases with fixed pupil activity belonged to the "unfavorable" category. In short, in-depth analysis of the effect of the variable in predicting GOSE revealed that the three parameters of "GCS motor response," "pupil activity," and "age" are the more important factors in predicting GOSE 6-month outcome in TBI patients.

This study showed that machine learning models when trained using a large data set could be used to predict the 6-month outcome in TBI patients using the parameters measured at admission. Our findings showed that the logistic regression model is superior to SVM and RF models because it is simple, more interpretable, and as accurate as the two more complicated models of SVM and RF. These promising results provide early evidence on the capability of machine learning models in predicting outcomes in TBI patients and ultimately assist physicians as possible decision support tools in diagnosing, optimizing treatment procedures, improving clinical outcomes, and offering substantial economic savings.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS

This study was funded by Shiraz University of Medical Sciences under Grant number 97-01-38-16824.

The authors wish to thank Mr. Mohsen Ghofrani for his assistance in model development and editing this manuscript. We wish to thank Dr. Laleh Khojastehat of the Research Consultation Center of Shiraz University of Medical Sciences for her invaluable assistance in editing this manuscript.

ORCID

Mehdi Nourelahi https://orcid.org/0000-0002-6986-3184
Fardad Dadboud https://orcid.org/0000-0001-6157-3046
Hosseinali Khalili https://orcid.org/0000-0002-3390-4212
Amin Niakan https://orcid.org/0000-0003-2644-8543
Hossein Parsaei https://orcid.org/0000-0003-1052-8434

AUTHOR CONTRIBUTIONS

Conceptualization: HK, HP. Data curation: HK, AN. Formal analysis: MN, FD, HP. Funding acquisition: HP. Methodology: all authors. Project administration: HK. Visualization: MN, FD, HP. Writing–original draft: MN, FD, HP. Writing–review & editing: all authors.

REFERENCES

1. Nguyen R, Fiest KM, McChesney J, Kwon CS, Jette N, Frolkis AD, et al. The international incidence of traumatic brain injury: a systematic review and meta-analysis. Can J Neurol Sci 2016;43:774-85.
2. Prins M, Greco T, Alexander D, Giza CC. The pathophysiology of traumatic brain injury at a glance. Dis Model Mech 2013;6:1307-15.
3. Werner C, Engelhard K. Pathophysiology of traumatic brain injury. Br J Anaesth 2007;99:4-9.
4. Weir J, Steyerberg EW, Butcher I, Lu J, Lingsma HF, McHugh GS, et al. Does the extended Glasgow Outcome Scale add value to the conventional Glasgow Outcome Scale? J Neurotrauma 2012;29:53-8.
5. Hale AT, Stonko DP, Lim J, Guillamondegui OD, Shannon CN, Patel MB. Using an artificial neural network to predict traumatic brain injury. J Neurosurg Pediatr 2018;23:219-26.

6. Hale AT, Stonko DP, Brown A, Lim J, Voce DJ, Gannon SR, et al. Machine-learning analysis outperforms conventional statistical models and CT classification systems in predicting 6-month outcomes in pediatric patients sustaining traumatic brain injury. Neurosurg Focus 2018;45:E2.
7. Ledig C, Heckemann RA, Hammers A, Lopez JC, Newcombe VF, Makropoulos A, et al. Robust whole-brain segmentation: application to traumatic brain injury. Med Image Anal 2015;21:40-58.
8. Gong T, Ambastha AK, Tan CL, Su B, Lim TC. Automated prognosis analysis for traumatic brain injury CT images. Proceedings of 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR); 2015 Nov 3-6; Kuala Lumpur, Malaysia. p. 386-90.
9. Agoston DV, Langford D. Big Data in traumatic brain injury: promise and challenges. Concussion 2017;2:CNC45.
10. Folweiler KA, Sandsmark DK, Diaz-Arrastia R, Cohen AS, Masino AJ. Unsupervised machine learning reveals novel traumatic brain injury patient phenotypes with distinct acute injury profiles and long-term outcomes. J Neurotrauma 2020;37:1431-44.
11. Matsuo K, Aihara H, Nakai T, Morishita A, Tohma Y, Kohmura E. Machine learning to predict in-hospital morbidity and mortality after traumatic brain injury. J Neurotrauma 2020;37:202-10.
12. Shi HY, Hwang SL, Lee KT, Lin CL. In-hospital mortality after traumatic brain injury surgery: a nationwide population-based comparison of mortality predictors used in artificial neural network and logistic regression models. J Neurosurg 2013;118:746-52.
13. Klemenc-Ketis Z, Bacovnik-Jansa U, Ogorevc M, Kersnik J. Outcome predictors of Glasgow Outcome Scale score in patients with severe traumatic brain injury. Ulus Travma Acil Cerrahi Derg 2011;17:509-15.
14. Rughani AI, Dumont TM, Lu Z, Bongard J, Horgan MA, Penar PL, et al. Use of an artificial neural network to predict head injury outcome. J Neurosurg 2010;113:585-90.
15. Eftekhar B, Mohammad K, Ardebili HE, Ghodsi M, Ketabchi E. Comparison of artificial neural network and logistic regression models for prediction of mortality in head trauma based on initial clinical data. BMC Med Inform Decis Mak 2005;5:3.
16. Pourahmad S, Rasouli-Emadi S, Moayyedi F, Khalili H. Comparison of four variable selection methods to determine the important variables in predicting the prognosis of traumatic brain injury patients by support vector machine. J Res Med Sci 2019;24:97.
17. Demetriades D, Kuncir E, Velmahos GC, Rhee P, Alo K, Chan LS. Outcome and prognostic factors in head injuries with an admission Glasgow Coma Scale score of 3. Arch Surg 2004;139:1066-8.
18. Guyon I, Elisseeff A. An introduction to variable and feature selection. J Mach Learn Res 2003;3:1157-82.
19. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction. Second edition [Internet]. New York (NY): Springer-Verlag; 2009 [cited 2021 Sep 10]. Available from: http://www.springer.com/gp/book/9780387848570.
20. Probst P, Wright MN, Boulesteix AL. Hyperparameters and tuning strategies for random forest. WIREs Data Min Knowl Discov 2019;9:e1301.
21. Liaw A, Wiener M. Classification and regression by randomforest. R News 2002;2:18-22.
22. Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20:273-97.
23. Bennis FC, Teeuwen B, Zeiler FA, Elting JW, van der Naalt J, Bonizzi P, et al. Improving prediction of favourable outcome after 6 months in patients with severe traumatic brain injury using physiological cerebral parameters in a multivariable logistic regression model. Neurocrit Care 2020;33:542-51.
24. Rubin ML, Yamal JM, Chan W, Robertson CS. Prognosis of six-month Glasgow Outcome Scale in severe traumatic brain injury using hospital admission characteristics, injury severity characteristics, and physiological monitoring during the first day post-injury. J Neurotrauma 2019;36:2417-22.
25. Haveman ME, Van Putten MJ, Hom HW, Eertman-Meyer CJ, Beishuizen A, Tjepkema-Cloostermans MC. Predicting outcome in patients with moderate to severe traumatic brain injury using electroencephalography. Crit Care 2019;23:401.