HHS Public Access
Author manuscript
Resuscitation. Author manuscript; available in PMC 2024 April 17.

Published in final edited form as:
Resuscitation. 2024 January ; 194: 110049. doi:10.1016/j.resuscitation.2023.110049.

Electroencephalogram-based machine learning models to predict neurologic outcome after cardiac arrest: A systematic review

Chao-Chen Chen[a,b], **Shavonne L. Massey**[c,f], **Matthew P. Kirschen**[c,d,f], **Ian Yuan**[d,f], **Asif Padiyath**[d,f], **Allan F. Simpao**[d,f], **Fuchiang Rich Tsui**[a,d,e,f,*]

[a]Tsui Laboratory, Children's Hospital of Philadelphia, 734 Schuylkill Ave, Philadelphia, PA 19146, United States

[b]Department of Bioengineering, University of Pennsylvania, 240 Skirkanich Hall, 210 S 33rd St, Philadelphia, PA 19104, United States

[c]Department of Neurology and Pediatrics, Children's Hospital of Philadelphia, 3401 Civic Center Blvd, Philadelphia, PA 19104, United States

[d]Department of Anesthesiology and Critical Care Medicine, Children's Hospital of Philadelphia, 3401 Civic Center Blvd, Philadelphia, PA 19104, United States

[e]Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania, 423 Guardian Dr, Philadelphia, PA 19104, United States

[f]Perelman School of Medicine, University of Pennsylvania, 3400 Civic Center Blvd, Philadelphia, PA 19104, United States

Abstract

Aim of the review: The primary aim of this systematic review was to investigate the most common electroencephalogram (EEG)-based machine learning (ML) model with the highest Area Under Receiver Operating Characteristic Curve (AUC) in two ML categories, conventional ML and Deep Neural Network (DNN), to predict the neurologic outcomes after cardiac arrest; the secondary aim was to investigate common EEG features applied to ML models.

Methods: Systematic search of medical literature from PubMed and engineering literature from Compendex up to June 2, 2023. One reviewer screened studies that used EEG-based ML models

---

[*]Corresponding author at: Tsui Laboratory, Children's Hospital of Philadelphia, 734 Schuylkill Ave, Philadelphia, PA 19146, United States. tsuif@chop.edu (F.R. Tsui).

Declaration of competing interest
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement
Chao-Chen Chen: Writing – original draft, Visualization, Methodology, Formal analysis, Conceptualization. **Shavonne L. Massey:** Writing – review & editing, Validation. **Matthew P. Kirschen:** Writing – review & editing, Validation. **Ian Yuan:** Writing – review & editing, Validation, Formal analysis. **Asif Padiyath:** Formal analysis. **Allan F. Simpao:** Writing – review & editing, Validation, Formal analysis. **Fuchiang Rich Tsui:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Appendix A. Supplementary material
Supplementary material to this article can be found online at https://doi.org/10.1016/j.resuscitation.2023.110049.

to predict the neurologic outcomes after cardiac arrest. Four reviewers validated that the studies met selection criteria. Nine variables were manually extracted. The top-five common EEG features were calculated. We evaluated each study's risk of bias using the Quality in Prognosis Studies guideline.

Results: Out of 351 identified studies, 17 studies met the inclusion criteria. Random Forest (RF) ($n = 7$) was the most common ML model in the conventional ML category ($n = 11$), followed by Convolutional Neural Network (CNN) ($n = 4$) in the DNN category ($n = 6$). The AUCs for RF ranged between 0.8 and 0.97, while CNN had AUCs between 0.7 and 0.92. The top-three commonly used EEG features were band power ($n = 12$), Shannon's Entropy ($n = 11$), burst-suppression ratio ($n = 9$).

Conclusions: RF and CNN were the two most common ML models with the highest AUCs for predicting the neurologic outcomes after cardiac arrest. Using a multimodal model that combines EEG features and electronic health record data may further improve prognostic performance.

Keywords

Machine Learning; Deep Learning; Electroencephalography; Heart Arrest; Prognosis

Introduction

Poor neurologic outcome is widely considered the main cause of morbidity and mortality after cardiac arrest.[1] Many survivors suffer from long-term cognitive and motor deficits from resultant hypoxic ischemic brain injury that significantly impact their quality of life.[2] Accurate and timely prediction of neurologic outcome can help inform care management plans and identify patients who might benefit from neuroprotective interventions.

Electroencephalogram (EEG) has been used to predict the neurologic outcome after cardiac arrest.[3,4] However, use of EEG after cardiac arrest is limited in practice because the interpretation of EEGs requires a trained clinician, which is labor-intensive, time-consuming, expensive, and subjective. Hence, an automated process that makes accurate and timely predictions based on EEG data can facilitate the interpretation of EEGs, reduce clinicians' workload, and may improve the quality of patient care after cardiac arrest.

Machine learning (ML) is a subset of artificial intelligence focusing on developing computer algorithms and statistical models to enable computer systems to make predictions from learning data. The capability of ML to handle high-dimensional data and complex waveform has been proven in many fields of studies, e.g., computer vision, natural language processing.[5,6] ML may be applied in an automated process after cardiac arrest to increase predictive accuracy.

The Area Under Receiver Operating Characteristics Curve (AUC) is a common performance metric to evaluate the ability of an ML model to predict an outcome. AUC ranges between 0 and 1, with 0.5 representing a random guess. Generally, a ML model with an AUC between 0.8 and 0.9 is considered to be good, and an AUC > 0.9 is considered to be excellent prediction.[7] Hence, in this systematic review, we aim to answer two questions. First, what is the most common ML model with the highest AUC in conventional ML and DNN. Second, what are the common quantitative EEG (QEEG) features applied to ML models?

Background in machine learning models

Generally, ML models are categorized into two categories, conventional ML and Deep Neural Network models (DNN). The main difference between the two categories is the input to the model. The input of conventional ML is typically handcrafted features, such as demographics (age, sex, etc.) and QEEG features (Shannon's entropy, burst-suppression ratio, etc.). Such input requires domain experts to manually extract clinically relevant features from electronic health record (EHR) and/or raw EEG data. For instance, Shannon's entropy measures the amount of information in raw EEG waveform. Low entropy implies that the brain is less active, which increases the likelihood of a poor neurologic outcome after cardiac arrest. Standard deviation measures the variability of raw EEG waveform. Burst-Suppression measures the continuity of raw EEG waveform.[8] On the other hand, the input of DNN is raw EEG data. DNN automatically extracts features from the raw data.

In this review, we have identified two common conventional ML models, Random Forest (RF) and Logistic Regression (LR). RF model uses multiple (e.g., hundreds of) decision trees to make a prediction. Each decision tree contributes its prediction from the input dataset. The final prediction is the probability of a target outcome (e.g., poor neurological outcome) based on the outcomes of all decision trees. LR model calculates the probability of a target outcome from the input record. The key difference between RF and LR is that RF can better handle collinearity among variables than the LR. Fig. 1 summarized two common conventional ML models.

We have identified two common DNN models, Convolutional Neural Network (CNN) and Long-Short Term Memory Network (LSTM). CNN model has three components, convolutional layers, pooling layers, and fully connected layers. The convolutional layer automatically extracts features from the raw data. Then, features are passed down to a pooling layer to reduce the computational complexity, allowing the model to focus on more relevant and discriminative information. Finally, fully connected layers are used to predict the outcome. LSTM model is a type of recurrent neural network (RNN) capable of addressing temporal/sequential data by capturing temporal dynamics of the temporal data, which is important to process EEG data and its trend information. The key difference between CNN and LSTM is that LSTM can better handle temporal relationships among EEG data compared to the CNN model. Fig. 2 summarized two common DNN models.

Methods

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to conduct the systematic review (Supplemental Table 1).[9]

Eligibility criteria

Included studies met the following inclusion criteria. The cohort dataset should: (1) include use of EEG; (2) use ML models to predict neurologic outcome; (3) include cardiac arrest patients.

Studies that met the following exclusion criteria were excluded: (1) lacking an AUC; (2) a validation dataset was not included, (3) the study cohort consisted of non-human subjects.

Search strategy (study inclusion process)

Two librarians at the University of Pennsylvania provided guidance in our literature search. The search query was based on inclusion criteria, which consisted of three groups of key words, (1) EEG, (2) ML model, and (3) cardiac arrest. Supplemental Table 2 shows the details of the search queries. We then performed the search query on two popular databases: (1) PubMed, a medical literature database and (2) Compendex, an engineering literature database, up to June 2, 2023.

Screening process (study exclusion process)

We used Covidence,[10] to manage the screening process, which consisted of two stages. (1) Screening of the title and abstract and (2) screening of the full text to assess for inclusion and exclusion criteria.

One primary reviewer (CC) performed both stages of the screening process. Four secondary reviewers (SLM, MPK, IY, and AFS) validated the results of the screening process. IY and AFS validated included studies and SLM and MPK validated excluded studies. FT supervised the entire process.

Data extraction and analysis methods

Data extraction was performed by the primary reviewer (CC). Nine variables were extracted to answer the primary question: (1) number of subjects (N), (2) prevalence of poor outcome, (3) recording time of EEG to predict neurologic outcome after cardiac arrest, (4) input format of EEG-based ML model (i.e., raw EEG, QEEG features), (5) poor outcome definition (i.e., cerebral performance category [CPC] 3–5, or pediatric cerebral performance category [PCPC] 4–6), (6) time of outcome assessment in months, (7) ML category (i.e., conventional ML and/or DNN), (8) ML model (e.g., RF and CNN), and (9) the highest AUC in the study.

For the secondary aim, we counted the number of handcrafted QEEG features across the included studies and reported the top-five most common features.

Study risk of bias assessment

Four reviewers (CC, IY, AP, and AFS) rated the risk of bias using the Quality in Prognosis Studies (QUIPS),[11] a bias-assessment tool designed to assess the prognostic study from six domains: (1) study participation, (2) study attrition, (3) prognostic factor measurement, (4) outcome measurement, (5) study confounding, and (6) statistical analysis and reporting. Each domain has three levels: low, moderate, and high risk of bias. A low risk of bias represents a study unlikely to be bias in that assessment domain; a moderate risk of bias represents a study possible to be bias; a high risk of bias represents a study likely to be bias. Disagreement between assessors was resolved by consensus.

Results

Screening process

A total of 351 studies were identified in PubMed and Compendex. Title and abstract screening resulted in 68 studies for full-text screening. Full-text screening resulted in 17 qualified studies ($n = 17$). Fig. 3 summarizes the full screening process.

Study characteristics

The number of patients ranged from 50 to 1,039. The prevalence of patients with poor neurologic outcomes ranged from 25% to 79%. Of the 17 studies, only one study included pediatric patients.[12] The recording time of EEG data to predict neurologic outcomes after cardiac arrest ranged from 20 minutes to 66 hours. The defined poor neurologic outcome across the 17 studies included: Cerebral Performance Category (CPC) 3–5, Pediatric Cerebral Performance Category (PCPC) 4–6, and not awakening from cardiac arrest. The time of neurologic outcome assessment ranged from 0 days (on the discharge day) to six months from the time of hospital discharge. Supplemental Table 3 summarizes the study characteristics of the final 17 studies published between August 2017 and February 2023.

The input format of ML model was raw EEG waveform ($n = 4$), QEEG features ($n = 8$), the integration of raw EEG waveform and EHR data ($n = 1$), the integration of QEEG features and EHR data ($n = 4$). EHR data included age, sex, Pittsburgh Cardiac Arrest Category, presence of a shockable rhythm, time to the return of spontaneous circulation, corneal and pupillary reflex, somatosensory evoked potential (SSEP), anoxic finding using MRI or CT, time from injury to EEG recording, location of arrest.[13–16]

Eleven studies used conventional ML, including Random Forest (RF) ($n = 7$),[12,15,17–21] Logistic Regression (LR) ($n = 3$),[16,18,22] Support Vector Machine (SVM) ($n = 1$),[23] the ensemble model of RF, LR, SVM, K-nearest neighbor (KNN) ($n = 1$)[13] with AUCs ranging between 0.81 and 0.97; Six studies used Deep Neural Network (DNN), including Convolutional Neural Network (CNN) ($n = 4$),[24–27] Long-short Term Memory Model (LSTM) ($n = 1$),[28] the integration of CNN and LSTM ($n = 1$)[14] with AUC ranging between 0.70 and 0.92.[15]

Common QEEG features

The top-five common handcrafted QEEG features applied to ML models are, (1) band power ($n = 12$), (2) Shannon's entropy ($n = 11$) (3) burst-suppression ratio (BSR) ($n = 9$), (4) standard deviation of raw EEG ($n = 7$), and (5) regularity ($n = 5$). We summarized top-five common QEEG features in Table 1.

Study risk of bias assessment

Most studies ($n = 11$) had a low risk of bias across all six assessment domains. The detailed results of study risk of bias can be found in Table 2.

Discussion

After screening 351 abstracts, we performed a full review of 17 studies, and found that RF was the most common ML model with the highest AUC in the conventional ML category, followed by CNN in the DNN category to predict the neurologic outcome after cardiac arrest.

Conventional ML models handle features manually extracted from EEG waveform and/or EHR data, whereas DNN models handle only raw EEG waveform data. Given the advantages of RF and DNN in handling different domains of features, i.e., handcrafted (expert-chosen) features and automatically extracted features,[29,30] we recommend a hybrid model, combining RF and CNN, to improve the prognostic performance. To the best of our knowledge, such a hybrid approach has not yet been developed to predict the neurologic outcome. For the second question, we reported the top-five common handcrafted QEEG features used by RF models to predict the neurologic outcome.

Additionally, we identified two issues that may limit prognostic performance. First, most studies ($n = 14$) only used non-sequential ML models, e.g., RF and CNN; only 3 studies used sequential ML models, e.g., LSTM, to take advantage of temporal relationships in waveform data. For instance, in Tjepkema-Cloostermans's study, they extracted 9 QEEG features from the last 5 minutes of raw EEG, 12 hours after the cardiac arrest and passed them to an RF to predict the outcome.[19] Similarly, in another Tjepkema-Cloostermans's study, they also passed the EEG waveform from the last 5 minutes, 12 hours after cardiac arrest to CNN to predict the outcome.[26] Since such an approach does not consider the entire 12 hours of EEG data, but only the last 5 minutes of EEG data, it may limit the prognostic performance. On the other hand, Zheng et al. applied a sequential DNN model combining CNN and LSTM model to catch the dynamics of long-term EEG data.[14] In general, the study showed that by including longer-term EEG data in the model, AUCs increased from 0.82 (up to 12 hours after cardiac arrest) to 0.91 (up to 66 hours after cardiac arrest), and the sequential models generally outperformed non-sequential models at any time points after cardiac arrest. Similar results are seen in Ghassemi's study. Instead of using a sequential DNN, it applied a sequential LR to predict the outcome.[16] Based on these studies, we recommend applying sequential ML models to improve prognostic performance.

The second issue that may limit prognostic performance is the lack of multimodality. Studies have shown that models using single modality, i.e., only EEG or EHR data (e.g., age, sex, and Pittsburgh Cardiac Arrest Category) had lower performance than models using multimodal data, i.e., the combination of EEG and EHR data. For instance, Aghaeeaval et al. demonstrated that AUCs increased from 0.86 (using only EEG data) to 0.97 when both EEG and EHR data were included.[15] Similarly, Zheng et al. compared model performance between single-modality models using only EHR or EEG data and a multi-modality model using both EHR and EEG data; the AUCs increased from 0.73 (EHR only) or 0.81 (EEG only) to 0.91 (EHR + EEG).[14] However, none of the studies used other physiological signals, such as blood pressure, and blood oxygenation, in multimodality models for improving the performance.

Future works

Based on our results and analyses, we suggest four areas for improving prognostic performance: (1) building a hybrid model that combines RF and CNN to take advantage of features extracted from EEG waveform and EHR data; (2) adding top-five common QEEG features to RF models; (3) using sequential deep learning models such as LSTM to capture the dynamics of temporal trend information of EEG waveform; (4) developing a multimodal model that uses EEG waveform, EHR data, and additional vital-sign waveforms such as blood pressure.

Limitation

This study has a limitation. In the search query, we listed 19 common ML model keywords to extract the relevant studies; some uncommon keywords may not be included. Studies that do not have any of our keywords in their main text are not included.

Conclusions

In this systematic review, we studied 17 papers to answer two questions. For the primary question, RF was the most common ML model with the highest AUC in the conventional ML category, followed by CNN in the DNN category. For the secondary question, the top-five common QEEG features include band power, Shannon Entropy, burst-suppression ratio, standard deviation of raw EEG, and regularity. Future studies may focus on four aspects to improve the prognostic performance: (1) Combining RF and CNN, (2) applying common QEEG features to RF, (3) using a sequential ML model, and (4) developing a multimodal model.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by the Children’s Hospital of Philadelphia. We thank two librarians at the University of Pennsylvania, Maylene Qiu, and Douglas McGee, for providing valuable suggestions on developing comprehensive search queries.

Abbreviations:

EEG	electroencephalogram
ML	machine learning
AUC	the area under receiver operating characteristics curve
DNN	deep neural network models
EHR	electronic health record
RF	random forest
LR	logistic regression

CNN	Convolutional neural network
LSTM	Long-Short term memory network
RNN	Recurrent neural network
CPC	cerebral performance category
PCPC	Pediatric cerebral performance category

REFERENCES

1. Dragancea I, Wise MP, Al-Subaie N, et al. Protocol-driven neurological prognostication and withdrawal of life-sustaining therapy after cardiac arrest and targeted temperature management. Resuscitation 2017;117:50–7. 10.1016/j.resuscitation.2017.05.014. [PubMed: 28506865]
2. Lim C, Alexander MP, LaFleche G, Schnyer DM, Verfaellie M. The neurological and cognitive sequelae of cardiac arrest. Neurology 2004;63:1774–8. 10.1212/01.WNL.0000144189.83077.8E. [PubMed: 15557489]
3. Bauer G, Trinka E, Kaplan PW. EEG patterns in hypoxic encephalopathies (post–cardiac arrest syndrome). J Clin Neurophysiol 2013;30:477–89. 10.1097/WNP.0b013e3182a73e47. [PubMed: 24084181]
4. Sandroni C, D'Arrigo S, Cacciola S, et al. Prediction of poor neurological outcome in comatose survivors of cardiac arrest: a systematic review. Intensive Care Med 2020;46:1803–51. 10.1007/s00134-020-06198-w. [PubMed: 32915254]
5. Khan A, Laghari A, Awan S. Machine learning in computer vision: A review. ICST Trans Scal Inf Syst 2018:169418. 10.4108/eai.21-4-2021.169418. Published online July 13, 2018.
6. Nagarhalli TP, Vaze V, Rana NK. Impact of machine learning in natural language processing: A review. In: 2021 third international conference on intelligent communication technologies and virtual mobile networks (ICICV). IEEE; 2021. p. 1529–34. 10.1109/ICICV50876.2021.9388380.
7. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol 2010;5:1315–6. 10.1097/JTO.0b013e3181ec173d. [PubMed: 20736804]
8. Hirsch LJ, Fong MWK, Leitinger M, et al. American Clinical Neurophysiology Society's standardized critical care EEG terminology: 2021 version. J Clin Neurophysiol 2021;38:1–29. 10.1097/WNP.0000000000000806. [PubMed: 33475321]
9. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021. 10.1136/bmj.n71. Published onlinoe March 29, 2021.
10. Veritas Health Innovation MA. Covidence systematic review software.
11. Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med 2013;158:280. 10.7326/0003-4819-158-4-201302190-00009. [PubMed: 23420236]
12. Lee S, Zhao X, Davis KA, Topjian AA, Litt B, Abend NS. Quantitative EEG predicts outcomes in children after cardiac arrest. Neurology 2019;92:e2329–38. 10.1212/WNL.0000000000007504. [PubMed: 30971485]
13. Yang F, Elmer J, Zadorozhny VI. SmartPrognosis: Automatic ensemble classification for quantitative EEG analysis in patients resuscitated from cardiac arrest. Knowl Based Syst 2021;212. 10.1016/j.knosys.2020.106579 106579.
14. Zheng WL, Amorim E, Jing J, et al. Predicting neurological outcome in comatose patients after cardiac arrest with multiscale deep neural networks. Resuscitation 2021;169:86–94. 10.1016/j.resuscitation.2021.10.034. [PubMed: 34699925]
15. Aghaeeaval M, Bendahan N, Shivji Z, et al. Prediction of patient survival following postanoxic coma using EEG data and clinical features. In: 2021 43rd annual international conference of the IEEE engineering in medicine & biology society (EMBC). IEEE; 2021. p. 997–1000. 10.1109/EMBC46164.2021.9629946.

16. Ghassemi MM, Amorim E, Alhanai T, et al. Quantitative electroencephalogram trends predict recovery in hypoxic-ischemic encephalopathy. Crit Care Med 2019;47:1416–23. 10.1097/CCM.0000000000003840. [PubMed: 31241498]
17. Nagaraj SB, Tjepkema-Cloostermans MC, Ruijter BJ, Hofmeijer J, van Putten MJAM. The revised Cerebral Recovery Index improves predictions of neurological outcome after cardiac arrest. Clin Neurophysiol 2018;129:2557–66. 10.1016/j.clinph.2018.10.004. [PubMed: 30390546]
18. Pham SDT, Keijzer HM, Ruijter BJ, et al. Outcome prediction of postanoxic coma: A comparison of automated electroencephalography analysis methods. Neurocrit Care 2022;37:248–58. 10.1007/s12028-022-01449-8. [PubMed: 35233717]
19. Tjepkema-Cloostermans MC, Hofmeijer J, Beishuizen A, et al. Cerebral recovery index: Reliable help for prediction of neurologic outcome after cardiac arrest. Crit Care Med 2017;45:e789–97. 10.1097/CCM.0000000000002412. [PubMed: 28430695]
20. Admiraal MM, Ramos LA, Delgado Olabarriaga S, Marquering HA, Horn J, van Rootselaar AF. Quantitative analysis of EEG reactivity for neurological prognostication after cardiac arrest. Clin Neurophysiol 2021;132:2240–7. 10.1016/j.clinph.2021.07.004. [PubMed: 34315065]
21. Amorim E, van der Stoel M, Nagaraj SB, et al. Quantitative EEG reactivity and machine learning for prognostication in hypoxic-ischemic brain injury. Clin Neurophysiol 2019;130:1908–16. 10.1016/j.clinph.2019.07.014. [PubMed: 31419742]
22. De-Arteaga M, Chen J, Huggins P, Elmer J, Clermont G, Dubrawski A. Predicting neurological recovery with canonical autocorrelation embeddings. PLoS One 2019;14:e0210966. 10.1371/journal.pone.0210966. [PubMed: 30689648]
23. Carrasco-Gómez M, Keijzer HM, Ruijter BJ, et al. EEG functional connectivity contributes to outcome prediction of postanoxic coma. Clin Neurophysiol 2021;132:1312–20. 10.1016/j.clinph.2021.02.011. [PubMed: 33867260]
24. van Putten MJAM, Hofmeijer J, Ruijter BJ, Tjepkema-Cloostermans MC. Deep Learning for outcome prediction of postanoxic coma. In; 2018:506–9. 10.1007/978-981-10-5122-7_127.
25. Jonas S, Rossetti AO, Oddo M, Jenni S, Favaro P, Zubler F. EEG-based outcome prediction after cardiac arrest with convolutional neural networks: Performance and visualization of discriminative features. Hum Brain Mapp 2019;40:4606–17. 10.1002/hbm.24724. [PubMed: 31322793]
26. Tjepkema-Cloostermans MC, da Silva LC, Ruijter BJ, et al. Outcome prediction in postanoxic coma with deep learning. Crit Care Med 2019;47:1424–32. 10.1097/CCM.0000000000003854. [PubMed: 31162190]
27. Aellen FM, Alnes SL, Loosli F, et al. Auditory stimulation and deep learning predict awakening from coma after cardiac arrest. Brain 2023;146:778–88. 10.1093/brain/awac340. [PubMed: 36637902]
28. Zheng WL, Amorim E, Jing J, et al. Predicting neurological outcome from electroencephalogram dynamics in comatose patients after cardiac arrest with deep learning. IEEE Trans Biomed Eng 2022;69:1813–25. 10.1109/TBME.2021.3139007. [PubMed: 34962860]
29. Cayir A, Yenidogan I, Dag H. Feature extraction based on deep learning for some traditional machine learning methods. In: 2018 3rd international conference on computer science and engineering (UBMK). IEEE; 2018. p. 494–7. 10.1109/UBMK.2018.8566383.
30. Nguyen D, Nguyen H, Ong H, et al. Ensemble learning using traditional machine learning and deep neural network for diagnosis of Alzheimer's disease. IBRO Neurosci Rep 2022;13:255–63. 10.1016/j.ibneur.2022.08.010. [PubMed: 36590098]

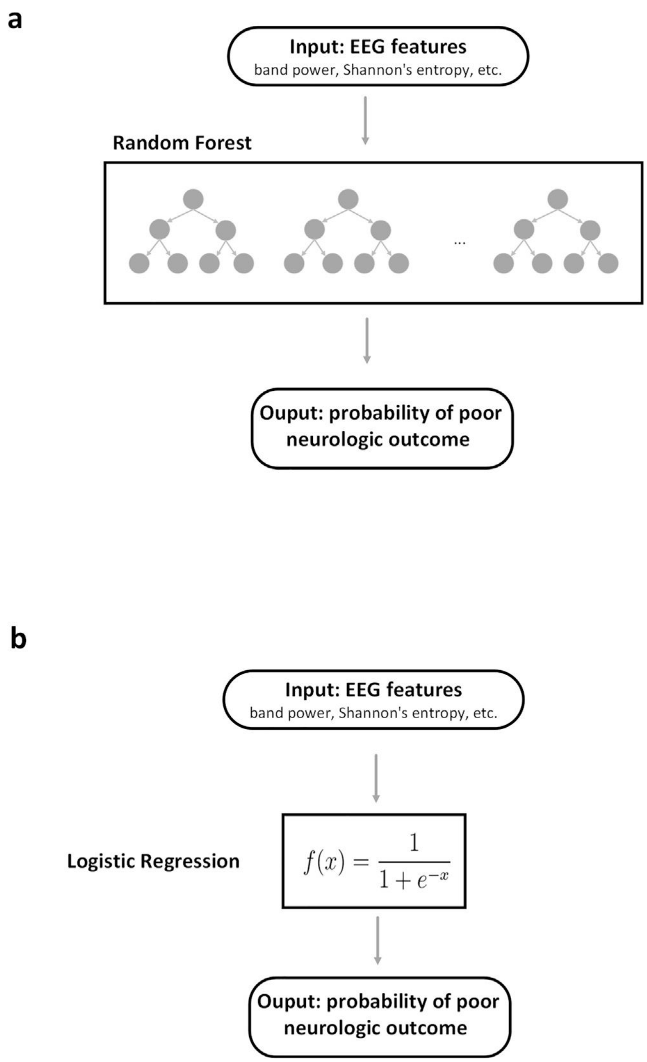

Fig. 1 –.

Two common conventional machine learning models. Plot (a): a Random Forest (RF) model with multiple (e.g., hundreds of) decision trees for outcome prediction. Each decision tree contributes its prediction from the input dataset. The prediction of a RF model is an ensemble predicted probability from all the trees of a target outcome (e.g., a poor neurological outcome). Plot (b): a Logistic Regression model that calculates the probability of a target outcome from the input record.

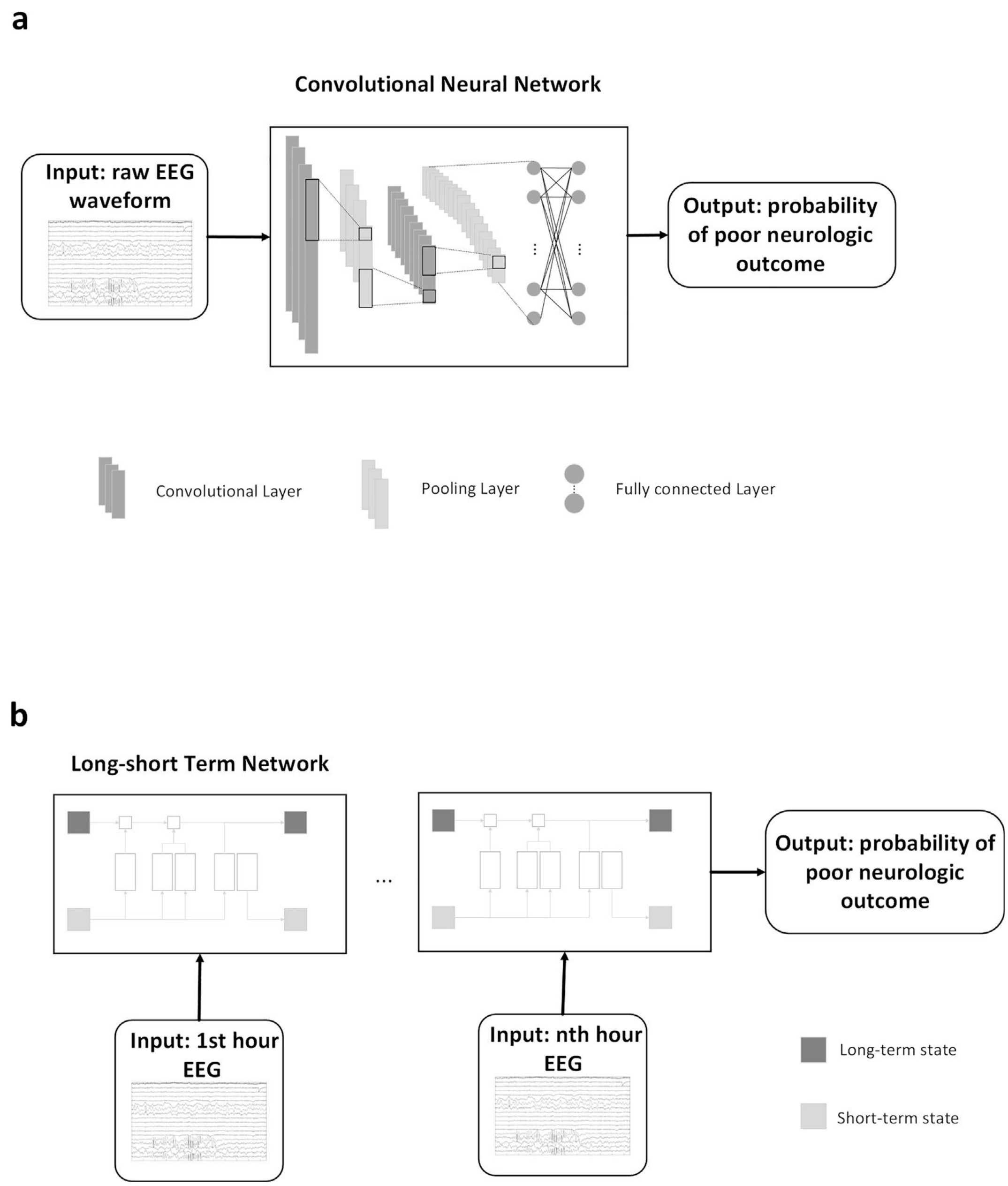

Fig. 2 –.

Two common deep neural networks. Plot (a): a Convolutional Neural Network with three components, convolutional layers, pooling layers, and fully connected layers. The convolutional layer extracts features from the raw waveform data; the pooling layers reduce the computational complexity from the extracted features to focus on relevant and discriminative information; finally, the fully connected layers provide outcome prediction. Plot (b): a Long-Short Term Memory Network (LSTM), a type of recurrent neural network (RNN), capable of addressing temporal/sequential data by capturing temporal dynamics of the temporal data, which is important to process EEG data and its trend information.

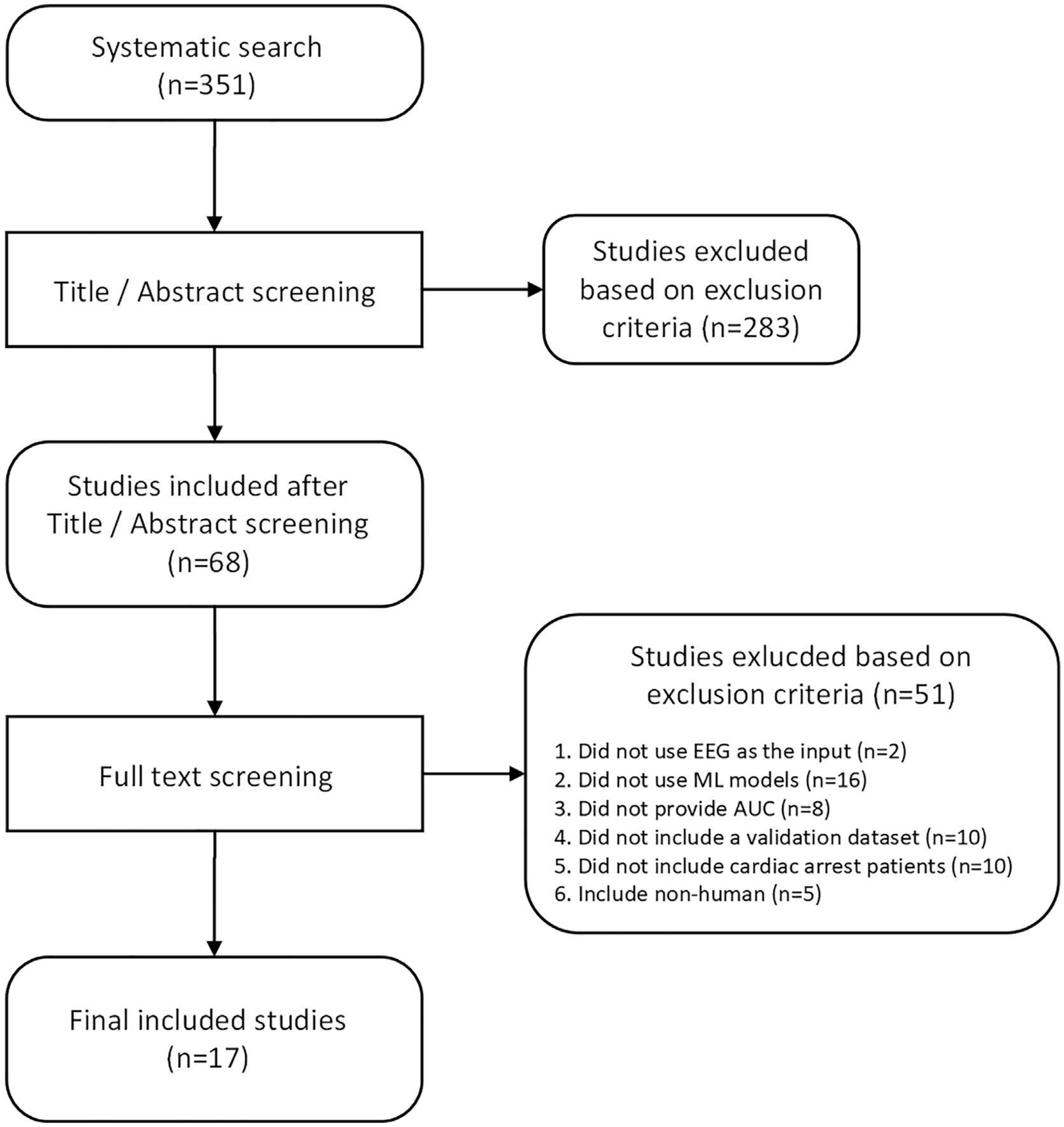

Fig. 3 –.
Screening process of paper review.

Table 1

The top five commonly used quantitative EEG features.

Feature	Number of studies	Brief description
Shannon entropy	11	The expected value of the amount of information
Band power	12	Spectral power in delta (0.5–4 Hz), theta (5–7 Hz), alpha (8–13 Hz), beta (14–20 Hz), or the ratio between different frequency bands.
Burst-suppression ratio	9	The fraction of EEG spent in suppression (<10 μV) per epoch.
Standard deviation	7	The standard deviation of the raw EEG in time-domain.
Regularity	5	An EEG measurement, ranging between 0 and 1, shows the continuity of EEG.

Table 2

Risk of bias among full-text reviewed papers.

Author, year	Study Participation	Study Attrition	Prognostic Factor Measurement	Outcome Measurement	Study Confounding	Statistical Analysis and Reporting
Lee, 2019[12]	Low	Low	Low	Low	Low	Low
Yang, 2021[13]	Moderate	Moderate	Low	Low	Low	Moderate
Zheng, 2021[14]	Low	Low	Low	Low	Low	Low
Aghaeeaval, 2021[15]	Low	Low	Low	Low	Low	Low
Ghassemi, 2019[16]	Low	Low	Low	Low	Low	Low
Admiraal, 2021[17]	Low	Low	Low	Low	Low	Low
Amorim, 2019[18]	Low	Low	Low	Low	Low	Moderate
Aellen, 2023[19]	Low	Low	Low	Low	Low	Low
Nagaraj, 2018[20]	Low	Low	Low	Low	Moderate	Low
Pham, 2022[21]	Low	Low	Low	Low	Low	Low
Tjepkema, 2017[22]	Low	Low	Low	Moderate	Low	Low
De-Arteaga, 2019[23]	Moderate	Low	Low	Low	Moderate	Low
Carrasco-Gómez, 2021[24]	Low	Low	Low	Low	Low	Low
Van Putten, 2018[25]	Low	Low	Low	Low	Low	Low
Jonas, 2019[26]	Low	Low	Low	Low	Low	Low
Tjepkema, 2019[27]	Low	Low	Low	Low	Low	Low
Zheng, 2022[28]	Low	Low	Low	Low	Moderate	Low