Forecasting Individual Seizure Risk in Neonatal Hypoxic-ischemic Encephalopathy Using Machine Learning

Jovana Kljajic
Faculty of Technical Sciences
University of Novi Sad
Novi Sad, Serbia
jovanakljajic18@uns.ac.rs

Marija Djermanovic
Faculty of Medicine
Univeristy of Novi Sad
Institute for Child and Youth Health Care of Vojvodina
Novi Sad, Serbia
marija.djermanovic@mf.uns.ac.rs

Slobodan Spasojevic
Faculty of Medicine
Univeristy of Novi Sad
Institute for Child and Youth Health Care of Vojvodina
Novi Sad, Serbia
slobodan.spasojevic@mf.uns.ac.rs

Tamara Skoric
Faculty of Technical Sciences
University of Novi Sad
Novi Sad, Serbia
tamara.ceranic@gmail.com

***Abstract*—This study introduces a novel machine-learning approach for predicting seizure development in neonates with hypoxic-ischemic encephalopathy (HIE) using an Extreme Gradient Boosting algorithm. The model employs a six feature set, consisting of four features extracted from the first 12 hours of single-channel EEG recordings and two clinical parameters available from the newborn's history charts. The proposed ML model was trained/tested on a dataset of 61 neonates with HIE, all born at a gestational age above 35 weeks. The model achieved a Matthews correlation coefficient (MCC) of 0.712 and an area under the receiver operating characteristic curve (AUC) of 0.885, outperforming the current state-of-the-art model, which also incorporates clinical and quantitative EEG features. These findings highlight the potential of the proposed method for early long-term seizure risk prediction in neonates with HIE, identifying those at risk of developing seizures after first 12 hours of recording EEG.**

***Index Terms*—machine learning, single-channel EEG, neonatal seizures, HIE neonates**

I. Introduction

Seizures are a significant concern in neonatology, occurring in approximately 1 to 5 per 1000 live births [1]. Most of them in the neonatal stage occur at the subclinical level, which means that they cannot be detected without monitoring electroencephalography (EEG) [2]. Hypoxic-ischemic encephalopathy (HIE), resulting from oxygen deprivation and reduced blood flow to the brain [3], [4], is a leading cause of seizures in neonates and a major contributor to long-term disabilities such as cerebral palsy and developmental delays [5]. Therapeutic hypothermia, introduced within the first 6 hours of birth, is the current standard treatment for moderate to severe HIE, improving long-term outcomes and reducing seizure burden [6]–[8].

The gold standard for seizure diagnosis is continuous video-EEG monitoring. However, its use is limited by the need for specialized equipment and trained personnel, making it inaccessible in many healthcare institutions [9]. Given these challenges and the importance of early identification and effective management of seizures to improve outcomes for neonates with HIE, the role of artificial intelligence (AI) in neonatal intensive care units (NICU) has increasing trend [10]. The goal is to develop a machine learning model that can identify neonates at high risk of developing seizures, utilizing the existing EEG monitoring equipment in clinical environments.

Early studies investigated correlations between various biochemical parameters and neonatal outcomes like mortality, HIE, and respiratory distress syndrome [11]. In [12], integrated clinical parameters and biochemical measures, combined with EEG background analysis were used. In recent years, machine-learning (ML) has demonstrated potential for improving long-term outcomes in neonates with HIE [13], [14]. These ML models aim to predict which neonates are at a higher risk of developing seizures in the future by leveraging both quantitative and qualitative EEG features, either individually or in combination with clinical parameters [9], [11], [12], [15], [16]. In [9], [17] potential of amplitude-integrated EEG (aEEG) has also been exploder, a simplified trend-monitoring tool that displays one or two channels of processed, time-compressed EEG on a semilogarithmic scale [18].

The aim of this study is to develop a ML model for early prediction of neonates with HIE who later develop seizures, utilizing quantitative-EEG features extracted from the first 12 hours of EEG recordings alongside selected clinical parameters.

II. Dataset

In this study, the dataset included infants born at >= 35 weeks of gestation, requiring continuous EEG monitoring due to a high risk of developing seizures. It comprised single-channel EEG signals recorded from parietal electrode locations P3 and P4 at a sampling rate of 200 Hz, along with clinical and biochemical data. All neonates included were diagnosed with moderate to severe HIE and had EEG recordings of a minimum duration of 12 hours, initiated within the first 6 hours of life. The recordings were obtained using the Olympic Medical CFM 6000 device (Natus Medical Incorporated, 5900 First Avenue South, Seattle, WA 98108, USA) between January 2021 and October 2024 in the Neonatal Intensive Care Unit of the Institute for Child and Youth Health Care of Vojvodina, Novi Sad, Serbia.

The study included 61 neonates, classified into two groups based on the presence and recurrence of neonatal seizures occurring after the first 12 hours of EEG monitoring. The first group included 44 neonates who either did not experience seizures or had only a single seizure episode, while the second group consisted of 17 neonates with multiple seizure episodes. Two medical experts conducted the assessment and classification of the neonates into these groups.

The clinical and biochemical data collected for each neonate included the mode of delivery, the neonate's gender, birth weight (BW), gestational age (GA), Apgar scores at first minute and fifth minute, assisted ventilation, pH, cardiopulmonary resuscitation in the delivery room (CPR), standard bicarbonate ($stHCO_3$), lactate levels, base excess, glycemia, and the age at seizure onset. Demographic characteristics, both overall and stratified by class, are presented in Table 1. Statistical significance ($p < 0.05$) was observed for HIE grades between two observed groups of infants, which is consistent with [9].

A 12 hour epoch of single-channel EEG recording was extracted for each neonate for the analysis. These recordings underwent quantitative analysis to detect patterns associated with seizure activity.

TABLE I
STUDY SAMPLES DEMOGRAPHICS AND CLINICAL CHARACTERISTICS

	All infants (n=61)	Class 1 (n=44)	Class 2 (n=17)	p-value
Gestational age of birth (weeks), mean (SD)	38.69 (1.75)	38.56 (1.78)	39 (1.65)	0.44^a
Birth weight (g), mean (SD)	3183.279 (582.76)	3152.95 (575.45)	3261.76 (611.99)	0.51^a
Male gender, n (%)	36 (59.02)	23 (52.27)	13 (76.47)	0.15^b
Mode of delivery, n (%) *Unassisted vaginal delivery* *Assisted vaginal delivery* *Elective cesarean section* *Emergency cesarean section*	 25 (40.98) 6 (9.84) 4 (6.56) 26 (42.62)	 20 (45.45) 3 (6.82) 3 (6.82) 18 (40.91)	 5 (29.41) 3 (17.65) 1 (5.88) 8 (47.0.6)	0.58^b
HIE grade, n(%) *Mild* *Moderate* *Severe* *Emergency*	 22 (36.07) 28 (45.90) 11 (18.03)	 21 (47.73) 19 (43.18) 4 (9.09)	 1 (5.88) 9 (52.94) 7 (41.18)	$< .002^b$

Note: p-Value < 0.05 was considered statistically significant
[a]p-Value from independent sample t test for parametric data for categorical data
[b]p-Value from chi-square test for categorical data or Fisher's exact test

III. Proposed Method

We propose a model aimed at identifying neonates with HIE who are at risk of experiencing multiple seizures after the first 12 hours of EEG recording. The objective is to detect this vulnerable group by integrating clinical and quantitative-EEG features using ML models. The proposed approach incorporates quantitative EEG features extracted through singular value decomposition (SVD) and spectral analysis of 12-hour EEG recordings. Prior to feature extraction, the raw EEG signal, sampled at 200 Hz, was preprocessed by using bandpass filter with bandpass from 0.5 to 30 Hz [19]. In addition to these features, two clinical parameters, glycemia and the Apgar score at the fifth minute, were selected for inclusion because they contributed to the higher performance of the model. The Apgar score is a quick clinical assessment of a newborn's health based on heart rate, respiratory effort, muscle tone, reflex irritability, and skin color, typically evaluated at 1 and 5 minutes after birth [20]. Glycemia refers to the concentration of glucose in the blood, which can indicate metabolic stability or distress in neonates [21]. Although each parameter had low individual discriminative power, together they contributed to a more effective classifier.

A. Spectral EEG Features Extraction

Spectral flatness and spectral entropy have been shown as useful features. Both features were derived based on the Power Spectral Density (PSD) estimation using Welch's method, on the entire 12 h or EEG signals.

Spectral flatness (S), also known as Wiener entropy, quantifies the flatness of a signal spectrum [19]. It is calculated as the ratio of the geometric mean and the arithmetic mean of the PSD. The spectral flatness was estimated following:

$$S = \frac{\exp\left(\frac{1}{K}\sum_{i=1}^{K}\log P_{\text{Welch}}(f_i)\right)}{\frac{1}{K}\sum_{i=1}^{K}P_{\text{Welch}}(f_i)} \quad (1)$$

where K is the number of frequency bin, and $P_{\text{Welch}}(f_i)$ is the PSD at frequency f_i. PSD was estimated by Welch's method, to minimize spectral leakage we applied Hamming window. The EEG signal segment length was $8 \times fs$ (sampling frequency), corresponding to 8 seconds of signal, and 75% overlap between segments was employed to ensure a smoother spectral estimate.

Spectral entropy quantifies the level of randomness or complexity within a signal [22]. The method involves computing

the Shannon entropy of the PSD, as given by the following equation:

$$H(p_{\text{Welch}}) = -\sum_{i=0}^{N} p_{\text{Welch}}(f_i) \log p_{\text{Welch}}(f_i) \quad (2)$$

where $p_{\text{Welch}}(f_i)$ is the normalized value of the PSD at frequency f_i.

B. EEG Features Extraction Based on Singular Value Decomposition

Singular value decomposition (SVD) has shown as useful for identifying vulnerable groups of infants, applying an aEEG [17], and short-term seizure prediction for pediatric subjects based on multichannel EEG [24]. In this study, we investigate the potential of the singular value of single-channel EEG. The entire 12 hour EEG can be presented as a matrix $X \in \mathbb{R}^{M \times L}$, where M corresponds to the number of non-overlapping EEG epochs, and L represents the number of samples, SVD decomposes X as follows [24]:

$$X = U \cdot \Sigma \cdot V^T = \sum_{i=1}^{r} \sigma_i \cdot U_i \cdot V_i^T, \quad r \leq min(M, L) \quad (3)$$

where T denotes transpose, r is rank and: $U \in \mathbb{R}^{M \times M}$ is an orthogonal matrix with columns U_i - the eigenvectors of the product $X \cdot X^T$; $\Sigma \in \mathbb{R}^{M \times L}$ is a diagonal matrix containing singular values σ_i that indicate the significance of each component; $V \in \mathbb{R}^{L \times L}$ is an orthogonal matrix with columns V_i - the eigenvectors of the product $A^T \cdot A$.

Shannon entropy of singular values was estimated following:

$$H = -\sum_{i=1}^{k} p_i \log p_i \quad (4)$$

where $p_i = \sigma_i / \sum_{j=1}^{5} \sigma_j$ represents the normalized singular values, k represents number of singular values used. Singular values on the diagonal of matrix S are sorted in the descending order and it was empirically proved that the first five represent dominant signal energy carriers.

Spectral decay is an indicator of signal energy concentration. It was calculated as the ratio of the first singular value to the sum of all five singular values:

$$S_{decay} = \frac{\sigma_1}{\sum_{i=1}^{k} \sigma_i} \quad (5)$$

C. Machine Learning (ML) Classifier

A ML model was developed to predict infants with HIE who experienced more than one seizure after the first 12 hours of EEG recording. The model combines quantitative EEG features and clinical features concatenated into a single vector per subject. It was built using the Extreme Gradient Boosting (XGBoost) algorithm, which is based on gradient-boosted decision trees introduced by Chen and Guestrin [25]. Leave-one-out cross-validation approach was used due to the small dataset used (<200). The depth of the decision tree was optimized from the data, with parameters selected using a grid search approach within a nested 10-fold cross-validation.

D. Model Evaluation

Performance of the machine learning models was assessed using the Matthews correlation coefficient (MCC) [26], which has shown to be effective for imbalanced datasets. We also estimated the performance of the ML model using following metrics: area under the receiver operating characteristic curve (AUC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). The binary classification performance was calculated using the following formulas:

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (6)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (7)$$

$$Specificity = \frac{TN}{FP + TN} \quad (8)$$

$$PPV = \frac{TP}{TP + FP} \quad (9)$$

$$NPV = \frac{TN}{FN + TN} \quad (10)$$

where TP represent true positives, TN true negatives, FP false positives and FN false negatives. TP corresponds to HIE infants who later develop seizures.

IV. Results

Table 2 compares the performance of all current state-of-the-art methods to the best author's knowledge. Pavel et al. extracted 41 features (13 clinical and 28 quantitative EEG features), based on golden-standard multichannel EEG. In our previous work [17], clinical data and an aEEG dataset (only 47 HIE infants EEG recordings were available at the time) were used with a total number of 8 features (6 clinical and 2 quantitative aEEG features). The proposed method outperformed the performances of our ML model-based aEEG [17], and most of the performance of the clinical-quantitate ML model for multichannel EEG (described in [9]).

However, a fair comparison of the proposed model is limited by the difference in data set size (all compared datasets are small $<$ 200 samples). The ML model in [9] was a trained/tested on dataset collected from multiple European centers (162 HIE infants), whereas the proposed model was developed using a dataset of 61 HIE infants from one center.

TABLE II
COMPARISON OF PERFORMANCE MEASURES FOR DIFFERENT MODELS

ML model	Feature number	Dataset size	MCC	AUC	Sensitivity	Specificity	PPV	NPV
Quantitative - EEG [a] (CatBoost) [9]	28	162	0.473 (0.337 to 0.612)	0.73 (0.671 to 0.811)	69.8	78.9	61.7	84.3
Clinical and quantitative - EEG [a] (CatBoost) [9]	41	162	0.513 (0.376 to 0.645)	0.746 (0.700 to 0.833)	75.5	78.0	62.5	86.7
Clinical and quantitative - aEEG[b] (AdaBoost) [17]	8	47	0.495	0.758	60.0	87.5	69.2	82.3
Proposed (XGBoost) method (quantitative EEG model) [c]	4	61	0.612 (0.371 to 0.809)	0.852 (0.754 to 0.934)	58.8	95.5	83.3	85.7
Proposed (XGBoost) method (clinical and quantitative EEG model) [c]	6	61	0.712 (0.523 to 0.877)	0.885 (0.803 to 0.951)	59.0	100	100	86.3

[a]EEG=multichannel EEG
[b]aEEG=single channel EEG (P3-P4)
[c]EEG=single channel EEG

V. DISCUSSION

The proposed clinical and quantitative EEG-based model for predicting seizure risk in neonates with HIE, demonstrates improved performance compared to the models which, to the authors knowledge, currently achieve the highest reported performance in the literature [9], [17]. The proposed model achieved higher performance by using six features, out of which two are clinical and four quantitative EEG features. The small database size also constrains the complexity of the machine learning model and increases the risk of overfitting. To mitigate this, cross-validation was employed to provide a more reliable assessment of performance.

Pavel et al. [9] proposed ML models incorporating quantitative and qualitative EEG/aEEG features, either independently or in combination with clinical parameters. The proposed model outperformed it across most of the metrics, reporting an MCC of 0.712 (0.523 to 0.877) and an AUC of 0.885 (0.803 to 0.951). However, a direct comparison with the model proposed in [9] remains challenging. Key distinctions include variations in dataset composition, sample size (both studies analyzed relatively small datasets of $<$200 cases), the duration of EEG epochs, and the type of device and the number of channels for collecting (a)EEG. To address these limitations, we also compared our model with our previous work [17], which utilized aEEG. aEEG primarily captures time-domain features, whereas the proposed model predominantly utilizes frequency-domain features extracted from EEG.

ACKNOWLEDGMENT

This study was supported by EU-funded project CA20124 - "Maximising the impact of multidisciplinary research in early diagnosis of neonatal brain injury" and by the Ministry of Science, Technological Development and Innovation (Contract No. 451-03-65/2024-03/200156) and the Faculty of Technical Sciences, University of Novi Sad through project "Scientific and Artistic Research Work of Researchers in Teaching and Associate Positions at the Faculty of Technical Sciences, University of Novi Sad" (No. 01-3394/1). The authors would like to thank all families for their willingness to participate in studies and clinical teams for database collection.

REFERENCES

[1] H. C. Glass, "Neonatal Seizures," Clinics in Perinatology, vol. 41, no. 1, pp. 177–190, Mar. 2014, doi: https://doi.org/10.1016/j.clp.2013.10.004.
[2] E. Baudou, C. Cances, C. Dimeglio, C. H. Lecamus, "Etiology of neonatal seizures and maintenance therapy use: a 10-year retrospective study at Toulouse Children's hospital," BMC Pediatr., vol. 19, pp.1-9, Apr 2019.
[3] D. Murray, G. Boylan, I. Ali, C. Ryan, B. Murphy, S. Connolly, "Defining the gap between electrographic seizure burden, clinical expression and staff recognition of neonatal seizures," Arch. Dis. Child. Fetal Neonatal Ed., v. 93, pp. 187–191, 2008
[4] H. Tekgul, K. Gauvreau, J. Soul, L. Murphy, R. Robertson, J. Stewart, J. Volpe, B. Bourgeois, A. J du Plessis. "The current etiologic profile and neurodevelopmental outcome of seizures in term newborn infants.", Pediatrics, vol. 117, pp. 1270–80, Apr 2006.
[5] C. Gale, Y. Statnikov, S. Jawad, S. N. Uthaya, and N. Modi, "Neonatal brain injuries in England: population-based incidence derived from routinely recorded clinical data held in the National Neonatal Research Database," Archives of Disease in Childhood - Fetal and Neonatal Edition, vol. 103, no. 4, pp. F301–F306, Oct. 2017, doi: https://doi.org/10.1136/archdischild-2017-313707.
[6] S. Shankaran et al., "Whole-body hypothermia for neonates with hypoxic-ischemic encephalopathy," The New England journal of medicine, vol. 353, no. 15, pp. 1574–84, 2005, doi: https://doi.org/10.1056/NEJMcps050929.
[7] E. Low et al., "Cooling and seizure burden in term neonates: an observational study," Archives of Disease in Childhood - Fetal and Neonatal Edition, vol. 97, no. 4, pp. F267–F272, Jan. 2012, doi: https://doi.org/10.1136/archdischild-2011-300716.
[8] P. Srinivasakumar, J. Zempel, M. Wallendorf, R. Lawrence, T. Inder, and A. Mathur, "Therapeutic Hypothermia in Neonatal Hypoxic Ischemic Encephalopathy: Electrographic Seizures and Magnetic Resonance Imaging Evidence of Injury," The Journal of Pediatrics, vol. 163, no. 2, pp. 465–470, Aug. 2013, doi: https://doi.org/10.1016/j.jpeds.2013.01.041.
[9] A.M. Pavel, J.M. O'Toole, J. Proietti, V. Livingstone, S. Mitra, WP Marnane, M. Finder, E.M. Dempsey, D.M. Murray, G.B. Boylan, "Machine learning for the early prediction of infants with electrographic seizures in neonatal hypoxic-ischemic encephalopathy," Epilepsia, vol. 64, pp. 456-468, Feb 2023.
[10] C. Henry et al., "Application and potential of artificial intelligence in neonatal medicine," Seminars in Fetal and Neonatal Medicine, vol. 27, no. 5, p. 101346, Oct. 2022.
[11] R. L. Andres et al., "Association between umbilical blood gas parameters and neonatal morbidity and death in neonates with pathologic fetal acidemia," American Journal of Obstetrics and Gynecology, vol. 181, no. 4, pp. 867–871, Oct. 1999, doi: https://doi.org/10.1016/s0002-9378(99)70316-9.

[12] D. M. Murray, C. A. Ryan, G. B. Boylan, A. P. Fitzgerald, and S. Connolly, "Prediction of Seizures in Asphyxiated Neonates: Correlation With Continuous Video-Electroencephalographic Monitoring," Pediatrics, vol. 118, no. 1, pp. 41–46, Jul. 2006, doi: https://doi.org/10.1542/peds.2005-1524.

[13] N. Schwalbe and B. Wahl, "Artificial intelligence and the future of global health," The Lancet, vol. 395, no. 10236, pp. 1579–1586, May 2020, doi: https://doi.org/10.1016/s0140-6736(20)30226-9.

[14] B. Abbasi and D. M. Goldenholz, "Machine learning applications in epilepsy," Epilepsia, vol. 60, no. 10, pp. 2037–2047, 2019, doi: https://doi.org/10.1111/epi.16333.

[15] A. O'Shea, G. Lightbody, G. Boylan, A. Temko, "Neonatal seizure detection from raw multi-channel EEG using a fully convolutional architecture", Neural Netw, vol. 123, pp. 12-25, May 2020.

[16] J.M. Perlman, R. Risser, "Can asphyxiated infants at risk for neonatal seizures be rapidly identified by current high-risk markers?,"Pediatrics., vol. 97, pp. 456-462, Apr 1996.

[17] T. Skoric, Marija Djermanovic, J. Kljajic, S. Spasojevic, and J. M. O'Toole, "Early Prediction of Electrographic Seizures in Neonatal Hypoxic-ischemic Encephalopathy Based on Amplitude-integrated EEG and Clinical Data," pp. 1581–1585, Aug. 2024, doi: https://doi.org/10.23919/eusipco63174.2024.10715242.

[18] R. A. Shellhaas, H. C. Glass, T. Chang, 18 - Neonatal Seizures, 6th ed., Swaiman's Pediatric Neurology, Elsevier, 2017, pp. 129-137.

[19] J. M. O'Toole and G. B. Boylan, "NEURAL: quantitative features for newborn EEG using Matlab," arXiv (Cornell University), Jan. 2017, doi: https://doi.org/10.48550/arxiv.1704.05694.

[20] V. Apgar, D. A. Holaday, L. S. James, I. M. Weisbrot, C. Berrien. "Evaluation of the newborn infant - second report," Journal of the American Medical Association, 168(15), pp.1985-1988.

[21] Wasserman, D.H., 2009. "Four grams of glucose," American Journal of Physiology-Endocrinology and Metabolism, 296(1), pp.E11-E21.

[22] C. E. Shannon, "A Mathematical Theory of Communication," Bell System Technical Journal, vol. 27, no. 379-423, p. 312, 1948, doi:https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

[23] A. Shahid, N. Kamel, A. S. Malik and M. A. Jatoi, "Epileptic seizure detection using the singular values of EEG signals," ICME International Conference on Complex Medical Engineering, Beijing, China, pp. 652-655. May 2013.

[24] K. Konstantinides, and K. Yao, "Statisticla analysis of effective singular value in matrix rank determination," IEEE Trans. Acoust., Speech Signal Process, vol. 36, pp.757-763, May 1988.

[25] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pp. 785-794. 2016.

[26] B. W. Matthews, "Comparison of the predicted and observed secondary structure of T4 phage lysozyme," Biochimica et Biophysica Acta (BBA) - Protein Structure, vol. 405, no. 2, pp. 442–451, 1975. DOI: 10.1016/0005-2795(75)90109-9.