

© Group of authors, 2021
UDC 616.831-053.31:612.821.7
DOI – <https://doi.org/10.14300/mnnc.2021.16031>
ISSN – 2073-8137

MACHINE LEARNING APPROACH TO CLASSIFICATION OF SLEEP ELECTROENCEPHALOGRAMS FROM NEWBORNS AT RISK OF BRAIN PATHOLOGIES

Selitsky S. ¹, Selitskaya N. ¹, Schult J. ²

¹ University of Bedfordshire, United Kingdom

² Universitat Hamburg, Germany

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ОЦЕНКИ И СИСТЕМАТИЗАЦИИ ЭЛЕКТРОЭНЦЕФАЛОГРАММ СНА У НОВОРОЖДЕННЫХ, ИМЕЮЩИХ РИСК РАЗВИТИЯ ПАТОЛОГИИ ГОЛОВНОГО МОЗГА

С. Селицкий ¹, Н. Селицкая ¹, Д. Шульц ²

¹ Университет Бедфордшир, Великобритания

² Гамбургский университет, Германия

This paper analyzes the Machine Learning approach to classifying sleep electroencephalograms recorded from newborns at risk of brain pathologies. The newborns were in different age groups counted in weeks of post-conceptual age. We consider solutions of the EEG task as a multiclass problem which can be resolved with Decision Tree models, which efficiently predict the weeks of PCA in terms of accuracy. The efficient solution to the multiclass tasks is difficult to find as decision models have to be explored in an ample model parameter space. Moreover, the sleep EEGs have significant overlap between ages because of variations in the newborn maturation patterns and expert evaluations. The experimental results have demonstrated the ability of ML technologies to provide classification accuracy comparable with the expert's opinions.

Keywords: sleep EEG, newborn, postconceptional age, brain maturity, classification, Machine Learning, Artificial Intelligence

Проведен анализ возможности машинного обучения в изучении и систематизации электроэнцефалограмм, записанных в течение сна у новорожденных с риском развития патологии головного мозга. Новорожденные были представлены в разных возрастных группах. Интерпретация ЭЭГ рассматривалась как мультизадачная проблема, решаемая с помощью применения искусственного интеллекта (модель дерева решений), который эффективно и точно оценивал постконцептуальный возраст ребенка. Вместе с тем полной экспертной оценки изменений ряда параметров достигнуть не удалось из-за недостаточной глубины принятия решений/заклучений в предлагаемой модели исследования. Кроме того, ЭЭГ сна имеют значительное перекрытие между возрастными группами из-за различий как в моделях развития мозга новорожденных, так и в оценках экспертов. Результаты исследования продемонстрировали способность предложенной технологии машинного обучения обеспечить оценку и систематизацию результатов ЭЭГ, сопоставимую по точности с заключениями экспертов.

Ключевые слова: ЭЭГ, постконцептуальный возраст, новорожденные, развитие мозга, классификация, машинное обучение, искусственный интеллект

For citation: Selitsky S., Selitskaya N., Schult J. MACHINE LEARNING APPROACH TO CLASSIFICATION OF SLEEP ELECTROENCEPHALOGRAMS FROM NEWBORNS AT RISK OF BRAIN PATHOLOGIES. *Medical News of North Caucasus*. 2021;16(2):140-143. DOI – <https://doi.org/10.14300/mnnc.2021.16031>

Для цитирования: Селицкий С., Селицкая Н., Шульц Д. ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ОЦЕНКИ И СИСТЕМАТИЗАЦИИ ЭЛЕКТРОЭНЦЕФАЛОГРАММ СНА У НОВОРОЖДЕННЫХ, ИМЕЮЩИХ РИСК РАЗВИТИЯ ПАТОЛОГИИ ГОЛОВНОГО МОЗГА. *Медицинский вестник Северного Кавказа*. 2021;16(2):140-143. DOI – <https://doi.org/10.14300/mnnc.2021.16031>

DT – decision tree
EEG – electroencephalograms

ML – Machine Learning
PCA – post-conceptional age

Today there are still many methodological limitations which are expected to overcome in electroencephalograms (EEG)-based newborn age classification. Some attempts have been undertaken to find efficient Machine Learning (ML) solutions in different real-world applications, such as detection of abnormal patterns [1–3], biometric identification [4, 5], trauma severity evaluation [6–8], and survival prediction [9, 10], collision avoidance at Heathrow [11], as well as for early detection of bone pathologies [12, 13].

Following [14, 15], we assume that dichotomisation of the multiclass problem using a set of binary classifiers can improve the accuracy of maturity assessments. We assume that this technique will not achieve better performance on our 10-class EEG problem, because its training will be affected by an imbalance of data. We summarise the performances of the techniques and conclude the paper in Section.

Material and Methods. EEG-based newborn age classification. An essential property of the age classification problem is the natural sequential ordering of class labels that represent the different post-conceptual age (PCA) groups. The conventional multiclass approaches treating the ordered labels as nominal ignore ordering information, which could improve the results. Therefore, it has been shown that such approaches are incapable of providing the best performance for problems with ordered classes. In cases when the class labels are continuous values, regression is typically used. However, the weeks PCA are only coarse estimates in our case, as the brain maturity may change during the week. Therefore the use of regression would be an *ad hoc* approach [16].

Another characteristic of the problem is the large overlap between the samples from the neighboring age groups. This is mainly caused by the normal variation of brain maturity patterns in the range of ± 1 or ± 2 weeks, the uncertainty in estimating the PCA, as well as the noise and artifacts in EEG. It would be desirable to find a classification model that tries to assign most samples to their stated PCA group and minimises the error in the range of ± 1 or ± 2 weeks. For example, given an EEG recording at 36 weeks, it would be unlikely to classify it as a record at 40 weeks then that at 37 weeks, so that the overall results in a range of ± 1 or ± 2 weeks could be improved by taking into account this observation. The overlap between the age groups, uncertainty in PCA estimates, and the noise in EEG data also affect the between-class boundaries making them difficult to learn.

Obviously, the performance of multiclass systems tends to decrease as the number of classes is increased. This is because the larger is the number of classes, the more difficult it is to learn the boundaries separating all the classes.

To handle many classes, a multiclass problem can be transformed into a set of separate two-class problems. This approach is called dichotomisation. The two-class problems are typically much easier to solve as only two classes need to be separated. The most widely used dichotomisation technique is the one-against-all classification transforming a problem with c classes into $c-1$ two-class problems of separating each class. A limitation of his approach is that, for a large number of classes, the portions of training samples in the binary problems may be strongly imbalanced, making training of the classifiers problematic.

Alternatively, the pairwise or one-against-one classification has shown providing better performance in the presence of multiple classes [17]. The idea is to train $c(c-1)/2$ models separating all pairs of classes and then

combine them to make the final decision. The pairwise system has outperformed the multiclass approaches for a large number of classes [18, 19] and has shown promising when samples of different classes overlap because of noise and variations [20, 21]. A drawback of using pairwise classification is that combining classifiers trained on different data may increase the uncertainty in outcomes. To mitigate this problem, instead of using all $c(c-1)/2$ pairs to decide for each sample, the classifiers can be selected and organised into a directed acyclic graph [22].

Problems with ordered labels were attempted to be solved using a modified one-against-all system and the pairwise one [23]. The performances of the approaches have been counted for the exact match between labels and outcomes. However, the performances in a range of ± 1 and ± 2 classes are yet to be explored. In this paper, we assume that EEG maturity assessments can be improved by using dichotomisation. We also assume that performances will be improved in the range of ± 1 and ± 2 . Notably, we expect that pairwise classification will provide better results because of its ability to handle a large number of classes and its robustness to noise and overlap in data.

Next, we show experiments with multiclass Bayesian classification of EEG in 10 PCA groups. The performance of the multiclass approach in the range of 0, ± 1 and ± 2 weeks will be compared to those of the dichotomisation approach.

Multiclass approach to EEG-based newborn age estimation. The settings for running the Bayesian classification were made as follows. The number of DTs sampled in the burn-in phase was 100,000 and in a post burn-in phase 10,000. During the post-burn-in phase, each 10th model was collected to reduce the correlation between decision tree (DT) models. The pruning factor was set to five. Proposal variance was 1.0, and probabilities of making moves of birth, death, change-variable, and change-rule were set to 0.15, 0.15, 0.1, and 0.6, respectively. Under the above settings, the acceptance rate of DT models during the integration was around 0.23 in both phases. The average DT included 66 nodes.

The EEG data included 952 recordings from newborns aged between 36 and 45 weeks of PCA. Each group included approximately 100 patients. All the recordings were additionally processed with an artifact rejection technique removing samples with abnormal amplitude deviation. The EEG recordings were also automatically tested on the presence of the level of detected artifacts.

The EEG were represented by the 36 standard spectral features as well as the new features describing the non-stationarity of an EEG recording. The new features included the segment rate and 10-bin histograms of the pseudo-stationary intervals ranging from 2 to 20 sec. Additionally, the ratio of absolute powers in Alpha and Theta bands was included.

Results and Discussion. The results of age classification within the 10-fold cross-validation are presented in Table 1. The table shows that the performance in terms of the exact match of weeks is $30.1 \pm 12.5\%$, and in the range ± 2 weeks, it is $85.5 \pm 0.8\%$. The ranges are similar to those that can be obtained for expert assessment. We expect that the performances can be further improved within the pairwise classification. Next, we briefly describe the one-against-all classification technique and show the results of experiments on the age classification problem.

One-against-all classification. The most widely used approach to transforming multiclass problems is the

one against-all classification. Within this technique, c binary classifiers are trained to distinguish each c class from the rest. C classifiers evaluate a test sample, and the class whose probability is highest is assigned to the sample. As the binary between-class boundaries are more accessible to learn than the multiclass ones, this technique is expected to improve the classification performance.

Table 1

Performance in intervals and entropy of multiclass discrimination

	0 weeks	1 week	2 weeks	Entropy
Performance in intervals	30.1±12.5	65.5±11.6	85.1±8.2	198.8±10.7

On the other hand, a problem with this technique is that portions of samples in the binary classifiers become imbalanced as the number of classes increases. The imbalance may lead to poor fitting of the classifiers because the errors for the smaller class can become ignored. Consequently, we assume that the one against-all technique will not perform better when the number of classes is large.

To test the assumption, we evaluate the performance of the one-again-all technique on a maturity assessment problem with 10 PCA groups. To integrate the Bayesian classification with the one-against-all technique, a DT ensemble is collected for each of the c binary classifiers. The posterior probabilities of each of the c classes are obtained from the collected ensembles, and the class with the highest probability is assigned to each sample.

The one-again-all technique was run on the same data and with the same settings as the multiclass one. Table 2 shows the average performances of age classification counted in the different ranges of weeks within the 10-fold cross-validation.

We can see that the performance for the exact match is comparable to that of the multiclass approach; however, in the range of ± 1 and ± 2 weeks, the performances are slightly lower by 2 and 1.5 %, respectively.

Table 2

Performance of one-against-all classification in intervals

	0 weeks	± 1 week	± 2 weeks
Indicator	29.9±9.2	63.5±7.8	83.5±8.4

The techniques converting a multiclass problem into a set of binary problems are typically expected to improve the performance. A commonly used technique is the one-against-all classification. However, in our experiments, the one-again-all technique did not improve

the performance of the age classification of 10 PCA groups. One explanation to this may be that the learning of binary classifiers was affected by data size imbalance, has similarly observed that the one against-all approach underperforms the multiclass age discrimination for ten age groups.

In, a one-again-all technique has been explicitly proposed to improve the accuracy for problems with ordered classes. They suggested that the binary classifiers can be trained to distinguish each class from the neighboring ones for such problems. An advantage of this approach is that data imbalance can be avoided. This technique will be explored in further work. In the next section, we briefly describe an alternative approach, the pairwise classification, and show results of experiments on EEG data.

The brain maturity classification task is characterised by a large number of classes, sequential ordering of class labels, and considerable overlap between samples of neighboring age groups. For such applications, the conventional multiclass approaches may perform poorly. We assumed that the performance of brain maturity assessments could be improved by using a classifier system that is more suitable given these characteristics of data.

The dichotomisation techniques, such as the one-against-all and pairwise classification, have been shown to outperform the multiclass approaches by splitting a multiclass problem into a set of binary ones, which are easier to solve. We assumed that pairwise classification, specifically, would provide a better performance, as it has been shown to effectively handling problems with ordered labels and with a large number of classes whose samples overlap. We also assumed that performances in the range of ± 1 and ± 2 weeks could be improved by using an approach suitable for systematic classification.

We have analyzed Machine Learning technologies for the classification of sleep electroencephalograms recorded from newborns at risk of brain pathologies admitted at the universities hospitals in Germany. The recordings were made in different age groups counted in weeks of post-conceptual age.

We have considered ML solutions of the EEG task as a multiclass problem which can be resolved with Decision Tree models which can efficiently predict weeks of PCA. We discussed difficulties in designing the efficient solution to the multiclass tasks, mainly caused by a large model parameter space required to be explored with an ML technology.

Conclusions. The sleep EEGs have significant overlap between ages because of variations in the newborn maturation patterns and expert's evaluations. Nevertheless, our experiments have demonstrated the ability of ML technologies to provide classification accuracy comparable with the expert's opinions.

Disclosures: The authors declare no conflict of interest.

Acknowledgments. The authors are thankful to Dr. L. Jakaite and Dr. V. Schetinin for their research inspiration and support.

References

1. Jakaite L., Schetinin V., Maple C. Bayesian assessment of newborn brain maturity from two-channel sleep electroencephalograms. *Computational and Mathematical Methods in Medicine*. 2012(3):629654. <https://doi.org/10.1155/2012/629654>
2. Nyah N., Jakaite L., Schetinin V., Sant P., Aggoun A. Evolving polynomialneural networks for detecting abnormal patterns. In: 2016 IEEE 8th International Conference on Intelligent Systems. 2016:74-80. <https://doi.org/10.1109/IS.2016.7737403>
3. Miranda-Dominguez O., Mills B. D., Carpenter S. D., Grant K. A., Kroenke C. D. [et al.] Connectotyping: Model based finger-printing of the functional connectome. *PLoS ONE*. 2014;9(11):e111048. <https://doi.org/10.1371/journal.pone.0111048>

4. Schetinin V., Jakaite L., Nyah N., Novakovic D., Krzanowski W. Feature extraction with GMDH-type neural networks for EEG-based person identification. *International Journal of Neural Systems*. 2017;28(6). <https://doi.org/10.1142/S0129065717500642>
5. Selitskaya N., Seliski S., Jakaite L., Schetinin V., Evance F. [et al.] Deep learning for biometric face recognition: Experimental study on benchmark data sets. *Deep Biometrics*. 2020:71-970. https://doi.org/10.1007/978-3-030-32583-1_5
6. Guyer C., Werner H., Wehrle F., Bolsterli B. K., Hagmann C. [et al.] Brain maturation in the first 3 months of life, measured by electroencephalogram: A comparison between preterm and term-born infants. *Clin. Neurophysiol*. 2019;130(10):1859-1868. <https://doi.org/10.1016/j.clinph.2019.06.230>
7. Castro Conde J. R., González Campo C., González González N. L., Reyes Millán B., González Barrios D. [et al.] Assessment of neonatal EEG background and neurodevelopment in full-term small for their gestational age infants. *Pediatr. Res*. 2020;88(1):91-99. <https://doi.org/10.1038/s41390-019-0693-0>
8. Cohen E., Wong F. Y., Wallace E. M., Mockler J. C., Odoi A. [et al.] EEG power spectrum maturation in preterm fetal growth restricted infants. *Brain Res*. 2018;1678:180-186. <https://doi.org/10.1016/j.brainres.2017.10.010>
9. Schetinin V., Jakaite L., Krzanowski W. Bayesian averaging over decision tree models: An application for estimating uncertainty in trauma severity scoring. *International Journal of Medical Informatics*. 2018;112:6-14. <https://doi.org/10.1016/j.ijmedinf.2018.01.009>
10. Schetinin V., Jakaite L., Krzanowski W. Bayesian averaging over decision tree models for trauma severity scoring. *Artificial Intelligence in Medicine*. 2018;84:139-145. <https://doi.org/10.1016/j.artmed.2017.12.003>
11. Schetinin V., Jakaite L., Krzanowski W. Bayesian learning of models for estimating uncertainty in alert systems: Application to air traffic conflict avoidance. *Integrated Computer-Aided Engineering*. 2018;26:1-17. <https://doi.org/10.3233/ica-180567>
12. Akter M., Jakaite L. Extraction of texture features from x-ray images: Case of osteoarthritis detection. Third International Congress on Information and Communication Technology. 2019:143-150. https://doi.org/10.1007/978-981-13-1165-9_13
13. Jakaite L., Schetinin V., Hladuvka J., Minaev S., Ambia A. [et al.] Deep learning for early detection of pathological changes in x-ray bone microstructures: case of osteoarthritis. *Scientific Reports*. 2021;11(1):e2294. <https://doi.org/10.1038/s41598-021-81786-4>
14. Schetinin V., Jakaite L. Extraction of features from sleep EEG for Bayesian assessment of brain development. *PLoS ONE*. 2017;12(3):e0174027. <https://doi.org/10.1371/journal.pone.0174027>
15. Low E., Mathieson S. R., Stevenson N. J., Livingstone V., Ryan C. A. [et al.] Early Postnatal EEG Features of Perinatal Arterial Ischaemic Stroke with Seizures. *PLoS ONE*. 2014;9(7):1-9. <https://doi.org/10.1371/journal.pone.0100973>
16. Frank E., Hall M. A simple approach to ordinal classification. In: Proceedings of the 12th European Conference on Machine Learning. 2001; 2167(2167):145-156. https://doi.org/10.1007/3-540-44795-4_13
17. Furnkranz J. Pairwise classification as an ensemble technique. In: Proceedings of the 13th European Conference on Machine Learning. 2002: 97-110.
18. Jakaite L., Schetinin V. Feature selection for bayesian evaluation of trauma death risk. In: 14th Nordic-Baltic Conference on Biomedical Engineering and Medical Physics: NBC. 2008:123-126. https://doi.org/10.1007/978-3-540-69367-3_33
19. Jakaite L., Schetinin V., Schult J. Feature extraction from electroencephalograms for Bayesian assessment of newborn brain maturity. In: Proceedings of the 24th IEEE International Symposium on Computer-Based Medical Systems. 2011. <https://doi.org/10.1109/CBMS.2011.5999109>
20. Kaminska A., Eisermann M., Plouin P. Child EEG (and maturation). *Handb. Clin. Neurol*. 2019;160:125-142. <https://doi.org/10.1016/B978-0-444-64032-1.00008-4>
21. Navarro X., Porée F., Kuchenbuch M., Chavez M., Beuchée A. [et al.] Multi-feature classifiers for burst detection in single EEG channels from preterm infants. *J. Neural. Eng*. 2017;14(4):046015. <https://doi.org/10.1088/1741-2552/aa714a>
22. Bishop C. M. Pattern Recognition and Machine Learning. *Journal of Electronic Imaging*. 16(4):140-155. <https://doi.org/10.1117/1.2819119>
23. Platt J., Cristianini N., Shawe-Taylor J. Large margin dags for multiclass classification. *Advances in Neural Information Processing Systems*. 2000;12:547-553.

About authors:

Selitsky Stas, PhD student at the School of Computer Science and Technology;
e-mail: selitsky@yahoo.com

Selitskaya Natalya, PhD student at the School of Computer Science and Technology;
e-mail: nselitsk@gmail.com

Schult Joachim, PhD, Lecturer;
e-mail: baj7113@uni-hamburg.de