Title: Scalable Machine Learning Architecture for Neonatal Seizure Detection on Ultra-Edge Devices

Paper Id: 64

Authors: Vishal Nagarajan, Ashwini Muralidharan, Deekshitha Sriraman, Pravin Kumar S

Presented by: Vishal Nagarajan

Contents

Overview

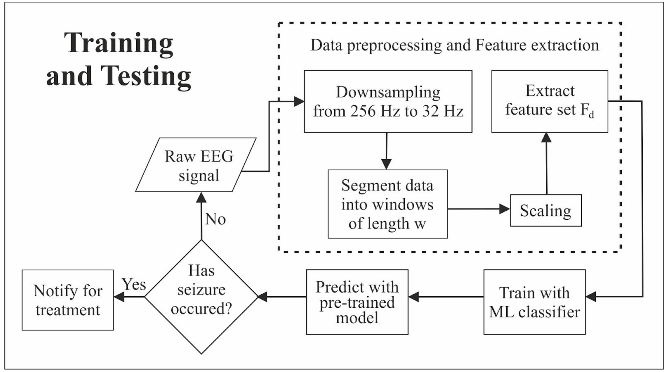
- Neonatal seizures are a commonly encountered neurological condition and is the first clinical sign of a serious neurological disorder. [1]
- Thus, rapid recognition and treatment are necessary to prevent serious fatalities.
- Detecting seizures on time without the need of a medical personnel could potentially prevent the negative effects on the neurocognitive development of the infants.
- Since there is a need for the classification of bio-signals to be computationally inexpensive in the case of seizure detection, this research presents a machine learning (ML) based architecture that operates with comparable predictive performance as previous models but with minimum level configuration.
- Our architecture achieved best sensitivity of 87%, that is more than 6% of that of the standard ML model chosen in this study.
- The model size of the ML classifier is optimized to just 4.84 KB with minimum prediction time of 182.61 milliseconds.

Dataset

- NICU of the Children's Hospital, Helsinki University Central Hospital, Finland [2]
- 18 channel EEG measurements
- 79 neonates participated in the study
- The recorded data was sampled at 256 Hz
- 19 electrodes placed according to the 10-20 international standard
- This dataset is dubbed as the Helsinki dataset

Pipeline Architecture

- This end-to-end architecture receives raw EEG signal, processes it and classifies it as ictal or normal activity.
- After preprocessing, the signal is passed to a feature extraction engine that extracts the necessary feature set F_d.
- It is followed by a scalable machine learning (ML) classifier that performs prediction.



Data Preprocessing and Feature Extraction Engine

- Data Preprocessing and Feature Extraction Engine contains the following sequential stages :
  - Downsampling from 256 Hz to 32 Hz [5]
  - Filtering using a high pass filter with a cut-off frequency of 0.5 Hz
  - Rescaling to values between 0 and 1 by applying minmax scaler.
  - Segmenting into various window lengths of 1, 2, 4, 8 and 16 seconds
  - Transforming output annotations into a single decision value of either 1 (indicating seizure) or 0 (for non-seizure)
  - Extracting features by converting each segment of EEG window to 11 human-engineered features per channel listed in the following slide.

AISP'22
IEEE Conference No. 53593

VIT-AP UNIVERSITY

Feature selection and training

FEATURE SET

TIME DOMAIN FEATURES	DESCRIPTION
Mean	Mean value of the signal in a window.
Standard deviation	Deviation from the mean value of the signal in a window.
Skewness	Measurement of lack of symmetry or asymmetry of an EEG signal.
Kurtosis	Essentially measuring the complexity of EEG signal, it determines if the signal has a peak or is flat at the mean.
Hjorth Activity	Variance of EEG signal in a window.
Hjorth Mobility	Measure of proportion of standard deviation of the power spectral density.
Hjorth Complexity	Compares similarity of the EEG signal to a pure sine wave
ENTROPY FEATURES	**DESCRIPTION**
Permutation Entropy	Measure of the local complexity in a signal.
Shannon Entropy	Measure of uncertainty in random process or quantities.
Approximate Entropy	Measure of the regularity and fluctuation in a time series.
Sample Entropy	Improved form of Approximate entropy, described as index of regularity. It reduces the bias caused by self-matching.

The 11 features were calculated for each channel and this resulted in a large number of feature vectors, which added to the computation overhead owing to its huge dimensionality. [3,4]

Many attributes may be highly correlated and thus, redundant, adding to the high dimensionality of data. High-dimensional data pose a problem of overfitting to predictive based models.

To tackle this problem, Principal Component Analysis (PCA) — a statistical method that reduces dimensionality by projecting most relevant attributes into lower dimensional space was used in this study.

This improves predictive performance of the models by eliminating less significant attributes. In this research, the total features were reduced to subsets of 20, 50, 70 and 100 for experimentation.

SERB INDIA

IEEE GUNTUR SUB SECTION AMARAVATI
IEEE HYDERABAD SECTION

Resource Efficient Classifier

- The feature extractor is connected to a scalable, binary classifier named ProtoNN. [6]
- Based on kNN, this classifier handles the trade-off between prediction accuracy and model size by employing 3 key methods. [7]
    - Projecting the entire data in low-dimension using sparse-projection matrix.
    - Learning prototypes to represent the entire training dataset.
    - Learning the projection matrix jointly with prototypes and labels.
- This allows the classifier to be deployed on devices that have RAM in the order of a few kilobytes.
- In contrast, kNN uses the entire training set for learning and prediction which is dilatory.

Metrics Used

In order to assess the performance of machine learning models, standard classification metrics are employed.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN} \qquad F1 = \frac{2 * TP}{2 * TP + FP + FN}$$

- *Sensitivity* is a measure of *False Negatives* whose value is *Recall-1*.
- *Model size* is the memory footprint of the classifier on the embedded device.
- *Inference time* is termed as the average time for the model to preprocess and predict the class of one segment of data.
- All models were off-loaded to a Raspberry Pi 3 Model B.

Results

WINDOW LENGTH (IN SECONDS)	PCA	TEST ACCURACY	MODEL SIZE (IN BYTES)	PRECISION		RECALL		F1	
				0	1	0	1	0	1
1	20	0.70	1760	0.76	0.70	0.66	0.80	0.71	0.75
	50	0.72	2960	0.8	0.75	0.73	0.82	0.76	0.78
	70	0.75	3760	0.8	0.76	0.73	0.82	0.76	0.79
	100	0.75	4960	0.81	0.8	0.79	0.82	0.80	0.81
2	20	0.71	1760	0.75	0.72	0.67	0.80	0.71	0.76
	50	0.72	2960	0.79	0.78	0.73	0.82	0.76	0.80
	70	0.73	3760	0.79	0.77	0.74	0.81	0.76	0.79
	100	0.76	4960	0.82	0.82	0.80	0.83	0.81	0.82
4	20	0.76	1760	0.79	0.84	0.83	0.80	0.81	0.82
	50	0.77	2960	0.80	0.84	0.83	0.81	0.81	0.82
	70	0.75	3760	0.80	0.85	0.84	0.81	0.82	0.83
	100	0.76	4960	0.83	0.86	0.86	**0.84**	0.84	0.85
8	20	0.75	1760	0.78	0.78	0.84	0.82	0.81	0.80
	50	0.74	2960	0.79	0.84	0.85	0.82	0.82	0.83
	70	0.75	3760	0.8	0.85	0.90	0.84	0.85	0.84
	100	0.77	4960	0.78	0.85	0.86	0.83	0.82	0.84
16	20	0.78	1760	0.79	0.80	0.84	0.85	0.81	0.82
	50	0.79	2960	0.81	0.85	0.88	0.83	0.84	0.84
	70	0.79	3760	0.82	0.84	0.87	0.81	0.84	0.82
	100	0.81	4960	0.84	0.85	0.86	**0.87**	0.85	0.86

ProtoNN Results

Results

WINDOW LENGTH (IN SECONDS)	PCA	K VALUE	TEST ACCURACY	PRECISION		RECALL		F1	
				0	1	0	1	0	1
1	20	37	0.75	0.66	0.71	0.72	0.71	0.72	0.77
	50	13	0.78	0.65	0.71	0.72	0.71	0.72	0.77
	70	37	0.76	0.70	0.73	0.74	0.73	0.74	0.77
	100	21	0.77	0.69	073	0.74	0.73	0.74	0.78
2	20	16	0.69	0.71	0.70	0.70	0.70	0.70	0.69
	50	18	0.75	0.78	0.77	0,76	0.77	0.76	0.875
	70	15	0.76	0.73	0.74	0.75	0.74	0.75	0.77
	100	9	0.77	0.75	0.76	0.76	0.76	0.76	0.77
4	20	19	0.78	0.75	0.76	0.77	0.76	0.77	0.78
	50	8	0.78	0.76	0.77	0.77	0.77	0.77	0.78
	70	11	0.78	0.76	0.77	0.78	0.77	0.78	0.79
	100	23	0.81	0.79	0.80	0.80	**0.80**	0.80	0.81
8	20	15	0.70	0.66	0.68	0.68	0.68	0.68	0.70
	50	11	0.80	0.79	0.79	0.79	0.79	0.79	0.78
	70	22	0.80	0.77	0.78	0.78	0.78	0.78	0.78
	100	9	0.80	0.75	0.76	0.77	0.76	0.77	0.79
16	20	37	0.69	0.68	0.69	0.69	0.69	0.69	0.69
	50	5	0.78	0.75	0.75	0.75	0.75	0.75	0.75
	70	3	0.78	0.72	0.73	0.74	0.73	0.74	0.76
	100	6	0.78	0.75	0.75	0.76	0.75	0.76	0.75

kNN Results

Results

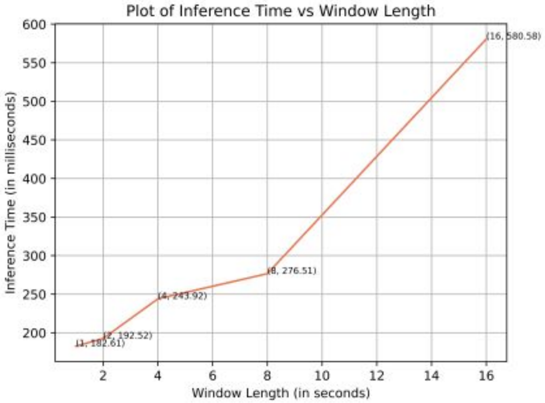


Fig. 4. Inference Time vs Window Length

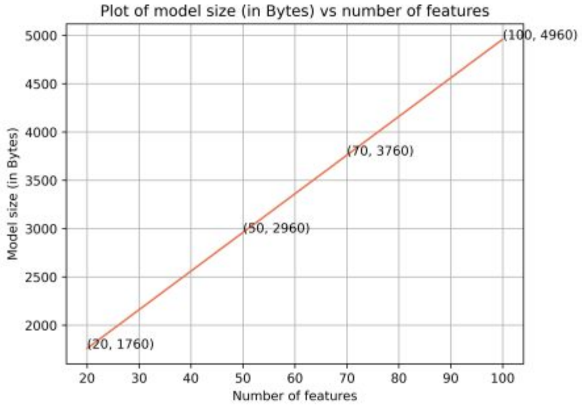


Fig. 5. Model Size vs No. of features

Conclusion

- In this paper, a resource efficient, medically feasible approach for detecting neonatal seizure has been presented and compared with a baseline ML model.
- The proposed architecture makes an optimal trade-off between predictive score and inference time, thereby facilitating accurate and viable seizure detection with practical application.
- The low storage requirement of the model makes it abundantly suitable for deployability on edge-devices.
- In future, the system can be tweaked to perform seizure prediction earlier than its onset by using algorithms such as early stopping and recurrent neural networks.
- Further, this pipeline can be integrated into a wearable device that processes the EEG signals and makes real-time predictions.

References

1. M. Dam, “Children with epilepsy: The effect of seizures, syndromes, and etiological factors on cognitive functioning,” Epilepsia, vol. 31, no. s4, pp. S26–S29, 1990. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1528- 1157.1990.tb05866.x
2. N. Stevenson, K. Tapani, L. Lauronen, and S. Vanhatalo, “A dataset of neonatal eeg recordings with seizures annotations,” 2018. [Online]. Available: https://zenodo.org/record/1280683
3. K. T. Tapani, S. Vanhatalo, and N. J. Stevenson, “Time-varying EEG correlations improve automated neonatal seizure detection,” vol. 29, no. 04, p. 1850030, May 2019. [Online]. Available: https://doi.org/10.1142/s0129065718500302
4. P. Boonyakitanont, A. Lek-uthai, K. Chomtho, and J. Songsiri, “A review of feature extraction and performance evaluation in epileptic seizure detection using eeg,” Biomedical Signal Processing and Control, vol. 57, p. 101702, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1746809419302836
5. D. Y. Isaev, D. Tchapyjnikov, C. M. Cotten, D. Tanaka, N. Martinez, M. Bertran, G. Sapiro, and D. Carlson, “Attention-based network for weak labels in neonatal seizure detection,” *Proceedings of machine learning research*, vol. 126, pp. 479–507, Aug 2020, 32995751[pmid]. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/3299575

References

6. C. Gupta, A. S. Suggala, A. Goyal, H. V. Simhadri, B. Paranjape, A. Kumar, S. Goyal, R. Udupa, M. Varma, and P. Jain, "ProtoNN: Compressed and accurate kNN for resource-scarce devices," in *Proceedings of the 34th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, D. Precup and Y. W. Teh, Eds., vol. 70. PMLR, 06–11 Aug 2017, pp. 1331–1340. [Online]. Available: https://proceedings.mlr.press/v70/gupta17a.htm

7. Z. Lasefr, R. R. Reddy, and K. Elleithy, "Smart phone application development for monitoring epilepsy seizure detection based on eeg signal classification," in 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON), 2017, pp. 83–87.

Thank You!