Kriging-Bootstrapped DNN Hierarchical Model for Real-Time Seizure Detection from EEG Signals

I. L. Olokodana[1], S. P. Mohanty[2], and E. Kougianos[3]

Smart Electronic Systems Laboratory (SESL) [1,2,3]

University of North Texas, TX , USA.[1,2,3]

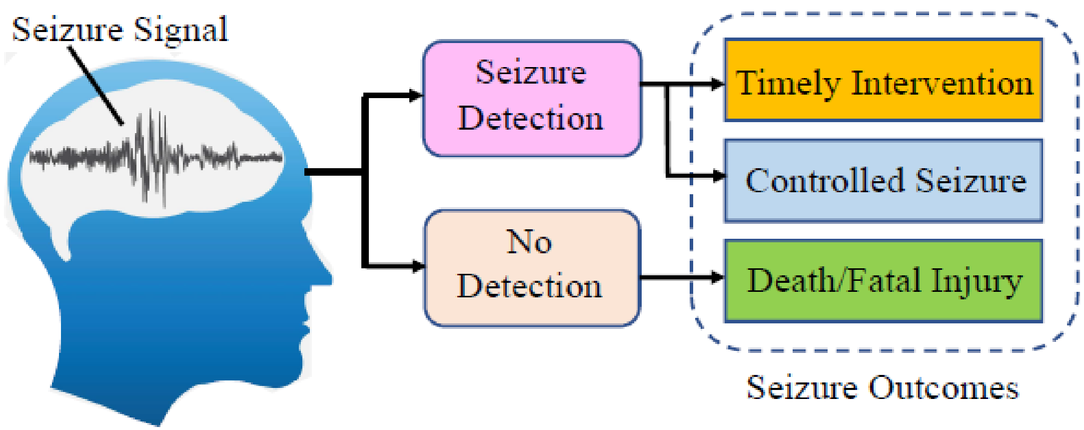
Email: IbrahimOlokodana@my.unt.edu[1], saraju.mohanty@unt.edu[2], elias.kougianos@unt.edu[3]

Outline of the talk

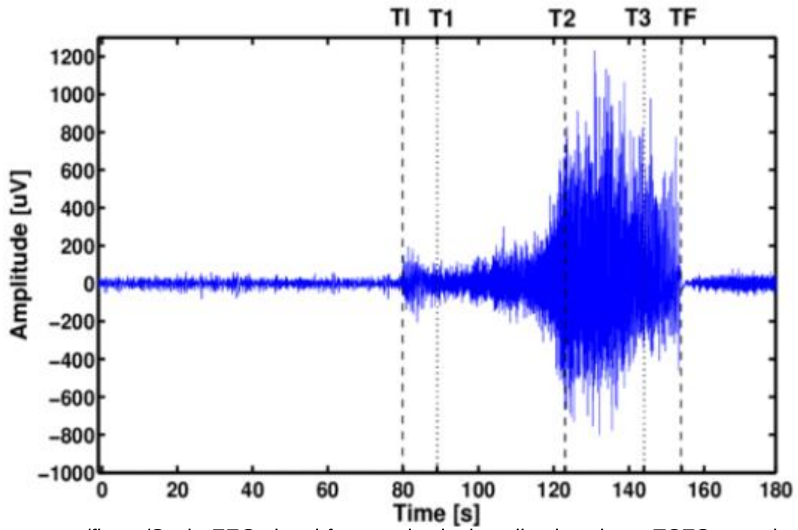
- Epileptic Seizures
- Kriging Methods
- Novel Contributions
- Brain as a Spatial Object
- Proposed Seizure Detection Models
- Experimental Results
- Conclusion & Future Research

Why Seizure Detection?

- Approximately 10% of the world population will have at least one experience of seizure in their lifetime.
- People living with epilepsy experience a higher mortality rate of about 44% compared to the general population with 12.2% mortality rate.
- Timely epileptic seizure detection is an important first step towards effectively managing the seizure disorder.



Characteristics of EEG Signal



Source: https://www.researchgate.net/figure/Scalp-EEG-signal-for-a-tonic-clonic-epileptic-seizure-TCES-recorded-at-the-central_fig2_283645061

- **Complexity**: Highly complex signals depicting the cortical electrical activities in the brain.
- **Intensity**: Low intensity signal measured in µV.
- **Frequency**: Frequency ranges from 0.5–30Hz. Signals are classified as delta, theta, alpha and beta based on frequency.

What are the Research Problems?

- Automatic Seizure Detection.
- Seizure Detection Latency.
- Extended Training Time.
- Patient-Specific Seizure Detection.
- Mobile and Portable Seizure Detection.
- Seizure Crisis Intervention Mechanism.
- Seamless, consistently accurate seizure detection system.
- Low-power seizure detection system.

What are the Challenges?

- Collecting a custom data-set due to the stringent regulations that are involved in collecting data from animals or human subjects.
- The same also applies to testing our models directly on human or animal subjects.
- it is difficult to estimate the level of noise in public data-sets since the conditions of the environment in which they are collected are not known.

Related Research in Seizure Detection – EEG/ML

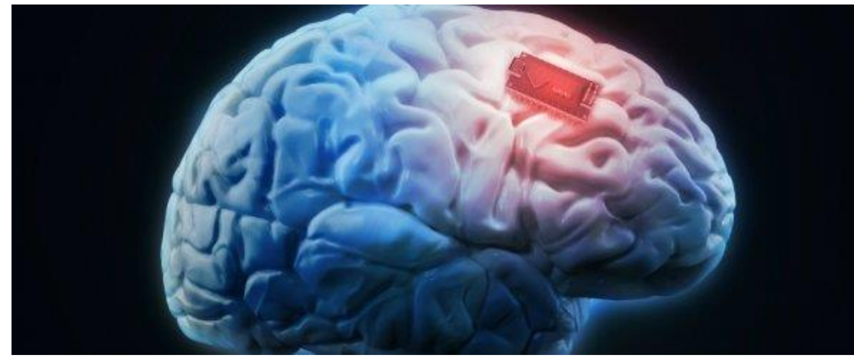
Reference	Dataset	Processing	Classifier Type
Alejandro et al (2017)	CHB-MIT	Time Analysis	Thresholds
Zhou et al. (2012)	Freiburg Database	Wavelet	Bayesian Method
Acharya et al. (2012)	Self-recorded	Freq. Analysis	SVM, KNN
Khan et al. (2012)	Self-recorded	Wavelet	LDA
Runarsson et al (2005)	Self-recorded	Time Analysis	SVM
Rezvan et al. (2017)	Bonn Dataset	Wavelet	MLP
Mursalin et al. (2017)	Bonn Dataset	Time Analysis	Random Forest
Guo et al. (2010)	Bonn Dataset	Wavelet	ANN
Mitra et al. (2009)	Texas' Children	Freq. Analysis	ANN
Zandi et al. (2010)	Vancouver GH	Wavelet	Thresholds

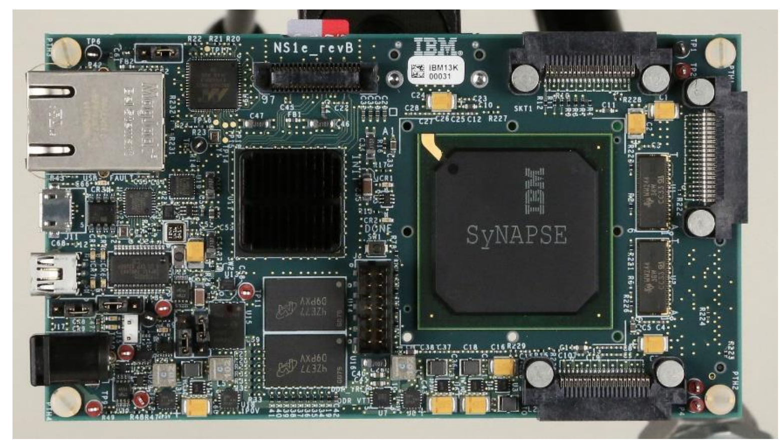
Related Research in Seizure Detection – Non-EEG

- A. Marquez, M. Dunn, J. Ciriaco, and F. Farahmand, “iSeiz: A low-cost real-time seizure detection system utilizing cloud computing,” in Proc. IEEE Glob. Hum. Tech. Conf., 2017, pp. 1–7.
- P. M. Vergara, E. de la Cal, J. R. Villar, V. M. Gonz´alez, and J. Sedano, “An IoT platform for epilepsy monitoring and supervising,” J. Sensors, vol. 2017, July 2017.
- Pavei, J., Heinzen, R. G., Novakova, B., Walz, R., Serra, A. J., Reuber, M., ... & Marques, J. L. (2017). Early seizure detection based on cardiac autonomic regulation dynamics. Frontiers in physiology, 8, 765.
- Jeppesen, J., Fuglsang-Frederiksen, A., Johansen, P., Christensen, J., Wüstenhagen, S., Tankisi, H., ... & Beniczky, S. (2019). Seizure detection based on heart rate variability using a wearable electrocardiography device. Epilepsia, 60(10), 2105-2113.

IBM's Implantable Seizure Detector

- IBM is developing an implantable seizure detector by leveraging on their neurosynaptic computing hardware called TrueNorth.
- The TrueNorth chip is postage stamp-sized and consumes over 1,000 times less power than a conventional processor of similar size.





Source: http://uberveillance.squarespace.com/?category=health_care

Consumer Electronics for Seizure Detection

Source: https://spectrum.ieee.org/the-human-os/biomedical/diagnostics/this-seizuredetecting-smartwatch-could-save-your-life

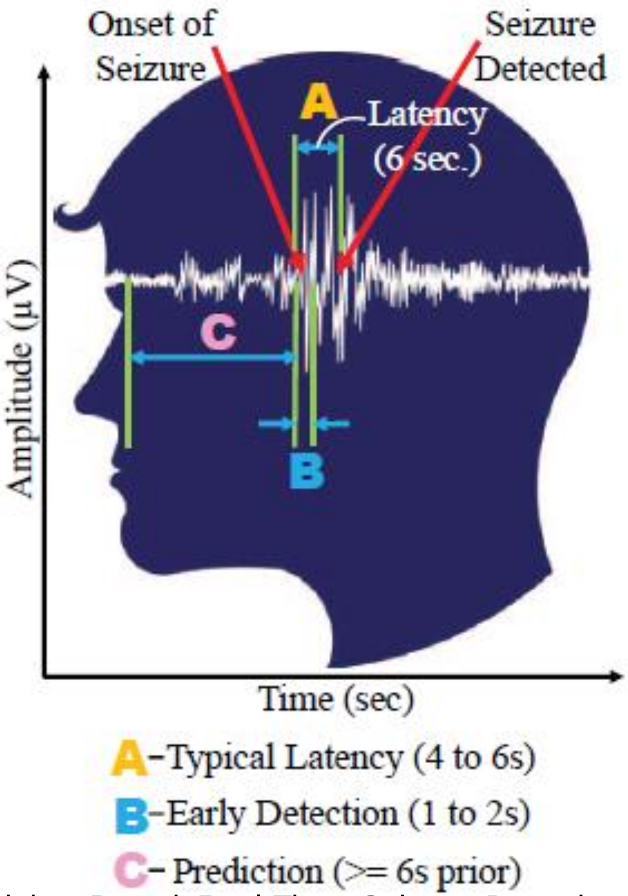
- Embrace2: Smart-band which uses machine learning to detect convulsive Seizures and notifies caregivers.

Source: https://www.empatica.com/embrace2/

- Medical grade smart watch: It detects generalized clonic-tonic Seizures and notifies physicians.

What are the Drawbacks of Existing Works?

- High seizure detection latency.
- Models are complex and unrealistic for real time deployment in the Internet of Medical Things (IoMT).
- Lack of adequate intervention mechanism after detection.



Source: I. L. Olokodana, S. P. Mohanty, and E. Kougianos, "Ordinary-Kriging Based Real-Time Seizure Detection in an Edge Computing Paradigm", in Proceedings of the 38th IEEE International Conference on Consumer Electronics (ICCE), 2020, Accepted

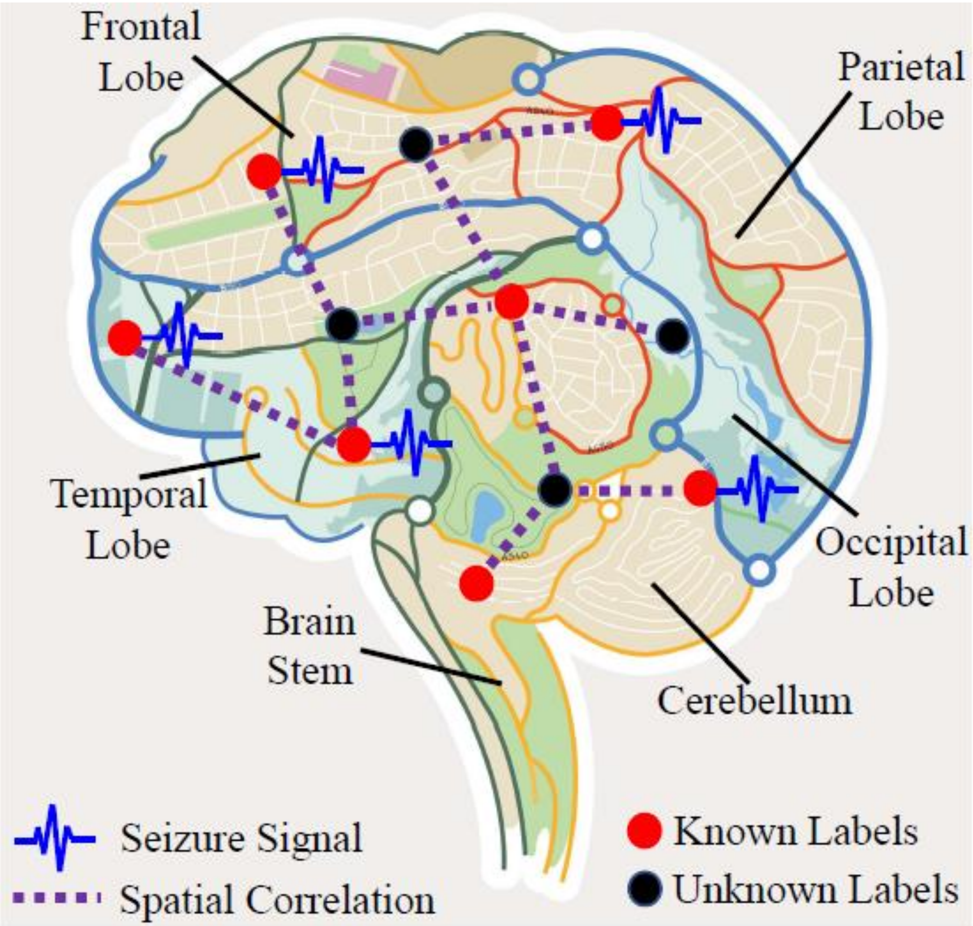
Research Question and Hypothesis

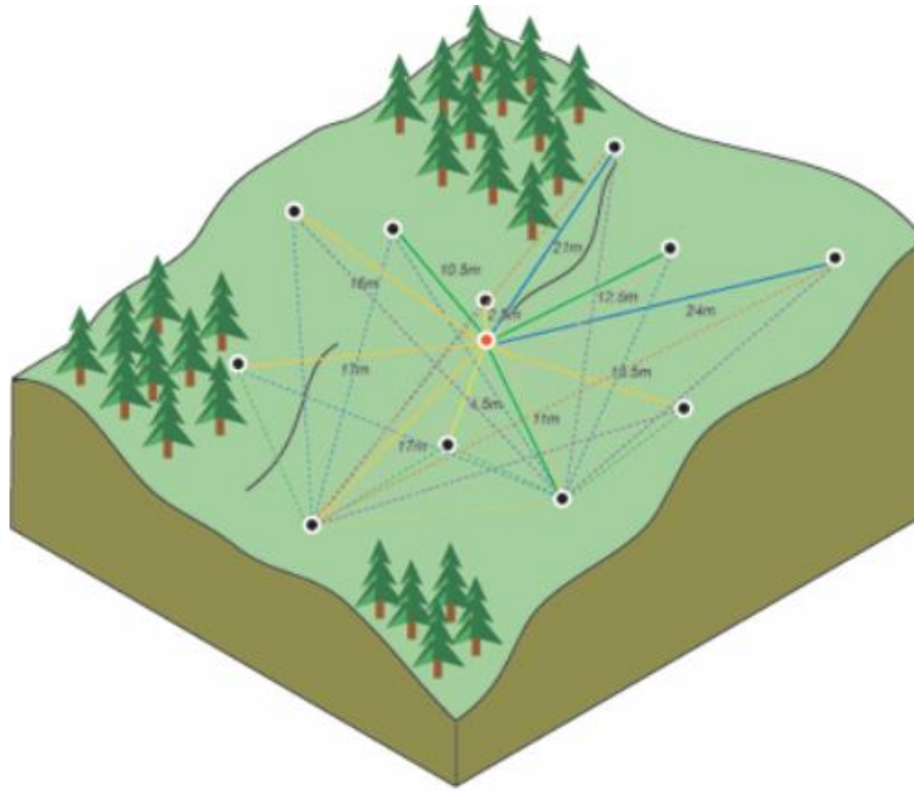
- How can the training time for a DNN be reduced in seizure detection applications without a compromise in performance?
- Since Kriging methods can estimate from the known data highly-correlated values which are previously unknown, will intermediate output from a Kriging model improve DNN training time?

Novel Contributions of the Current Paper

- This is the first Kriging-Bootstrapped hierarchical model for real-time seizure detection to the best of the authors' knowledge.
- Furthermore, the proposed hybrid model achieved a 75% reduction in training time and also improves the performance of the DNN by at least 2.5% after training on the same data size and the same DNN architecture.

Brain as a Spatial Map



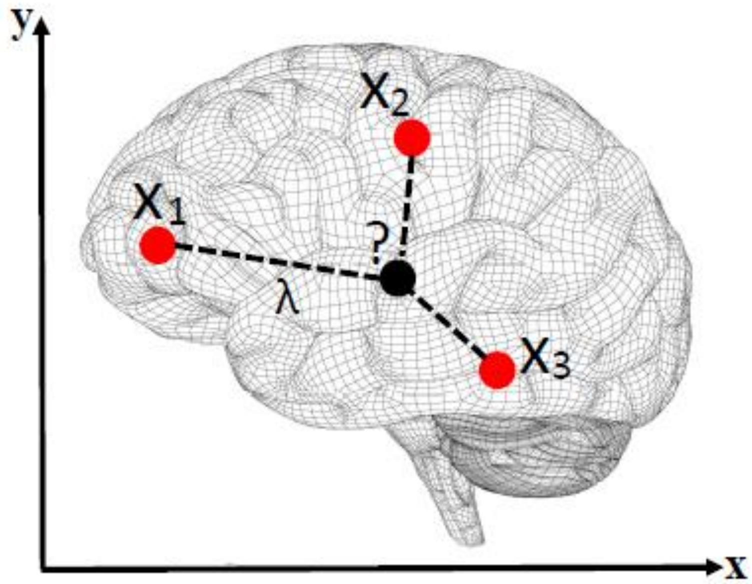


Source:https://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst-toolbox/how-kriging-works.htm

Source: I. L. Olokodana, S. P. Mohanty, and E. Kougianos, “Krig-Detect: Exploring Alternative Kriging Methods for Real-Time Seizure Detection from EEG Signal”, in Proceedings of the 6th IEEE World Forum on Internet of Things (WF-IoT), 2020, pp. Accepted.

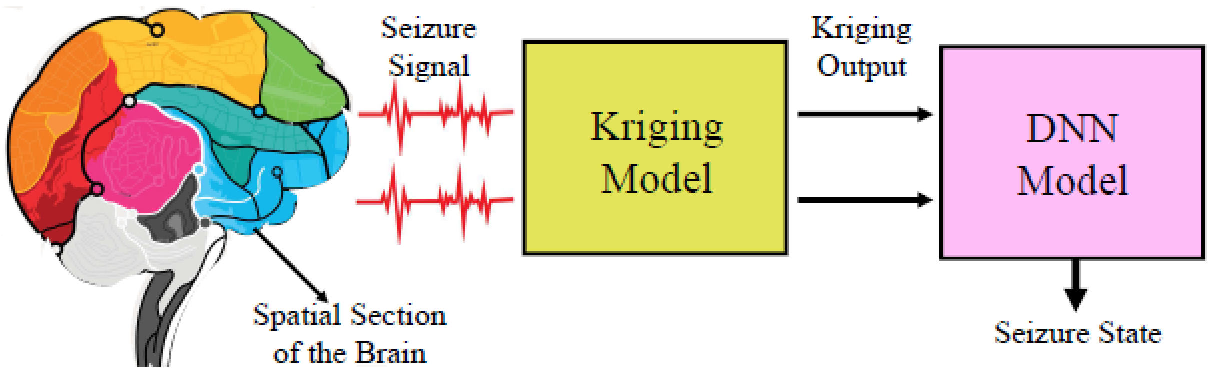
Kriging

- Kriging was originally developed as a geo-statistical model for spatial prediction.
- It is a stochastic process that is governed by a mean value and the relative co-variances of known data points with respect to an unknown.



Motivation

- Although training is not done in real time, a quick turnaround is needed to consistently update the already trained model that is used in a real time seizure detection system.
- Kriging tends to become slower when data size and feature dimension become high.



Bootstrap Sampling

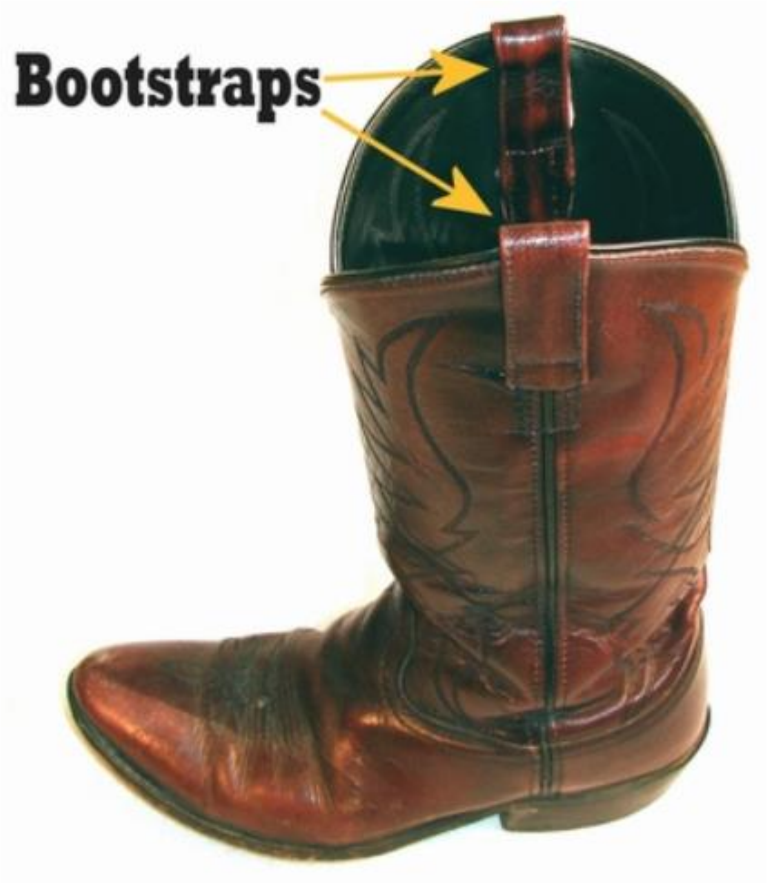
- The term "Bootstrap" as used in this work was borrowed from Bootstrap sampling which refers to a method of estimating the true statistical value of a population from some given samples.

- It involves sampling from the available samples with replacement to create multiple sets of samples.

- The number of sample sets that is created from the original samples is called Bootstrap Size.

The Bootstrap

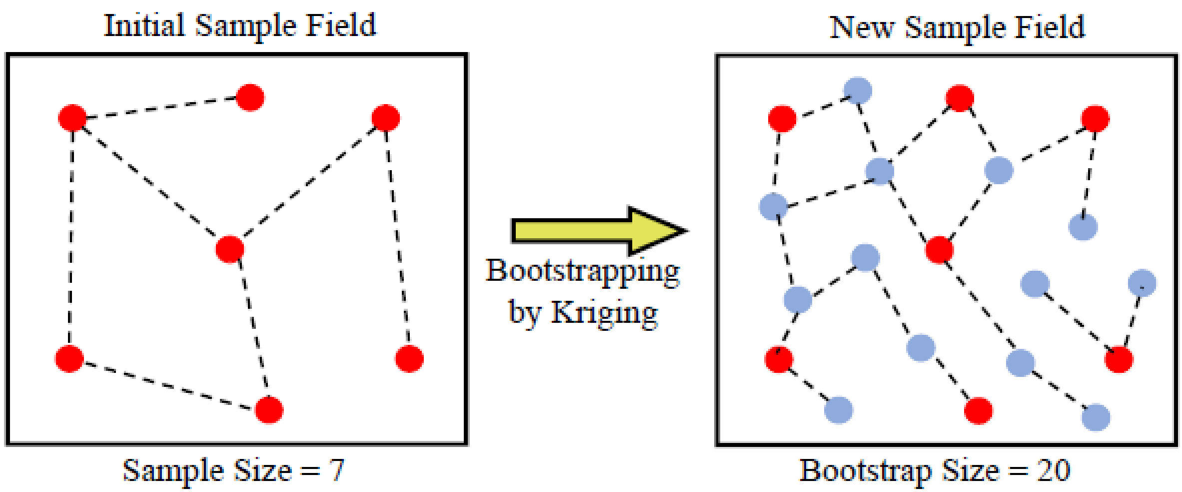
- A Bootstrap is a strap that is looped and sewn to the top of a boot for greater ease in pulling it on.
- It simply means getting oneself through difficulties only with the resources at one's disposal.



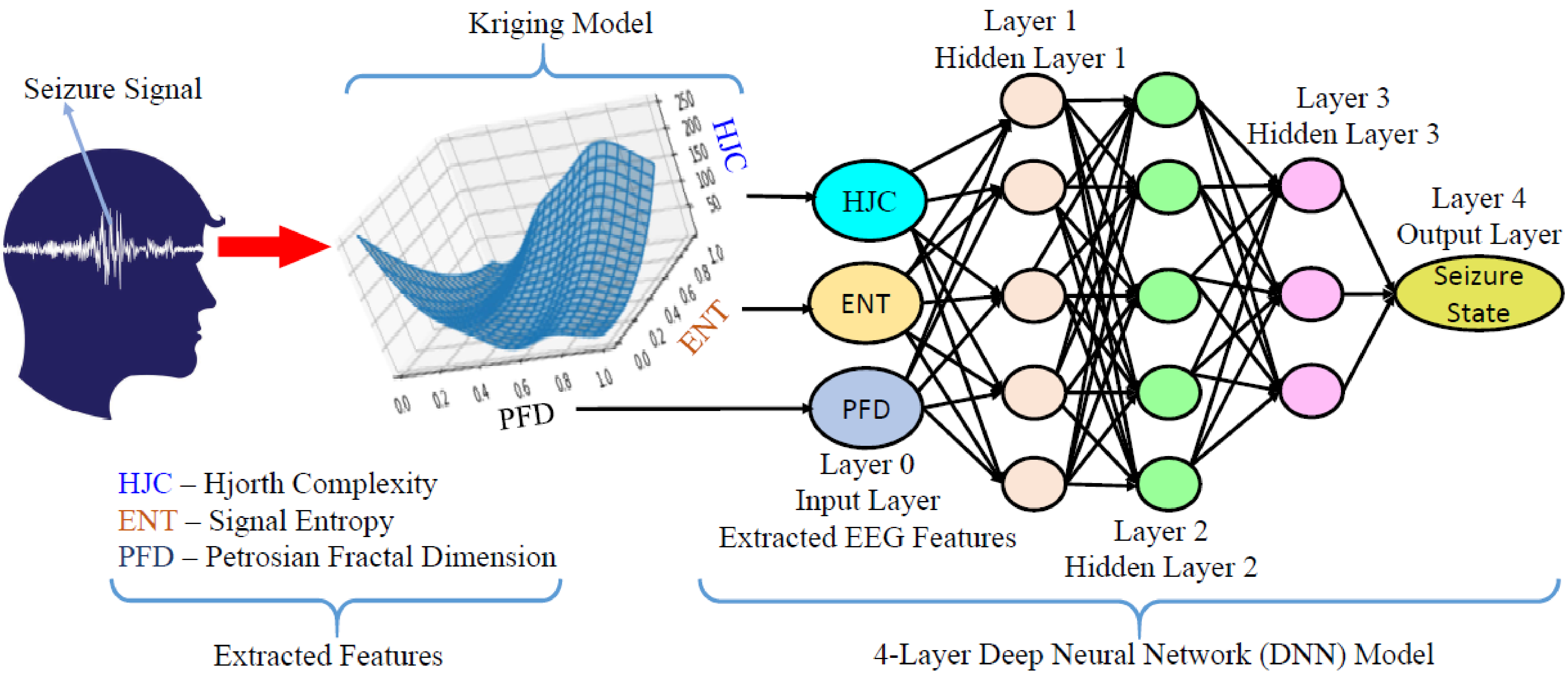
Source: http://www.lemen.com/dictionary-b.html#bootstrap

Bootstrapped Kriging

- In the context of this work, Bootstrapping refers to the process of increasing the correlation between data points in the sample field by increasing the number of data points per unit area via Kriging.

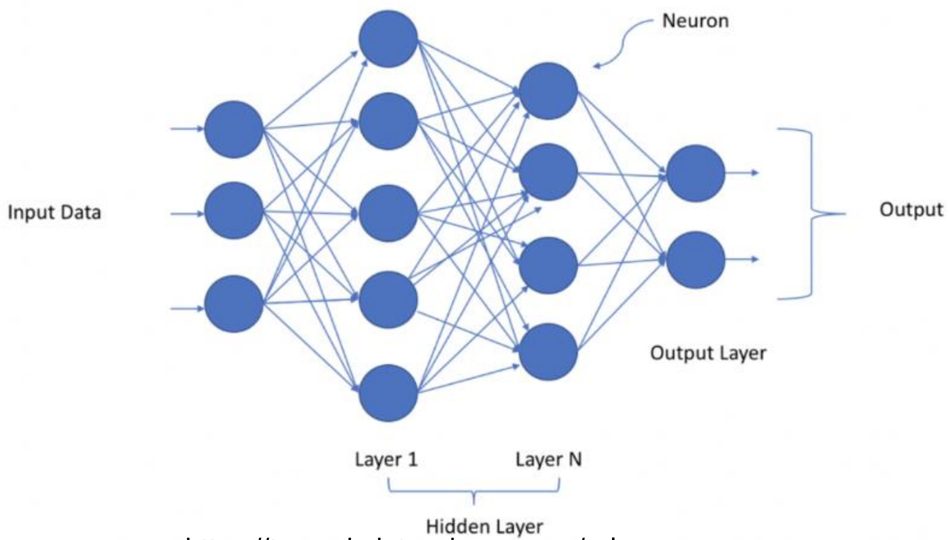


Proposed Krig-DNN Model



Deep Neural Network

- A deep neural network is a neural network with more than one hidden layer.
- The neural network operates by minimizing the cost function.



https://towardsdatascience.com/a-laymans-guide-to-deep-neural-networks-ddcea24847fb

- The cost function is given by the following expression:

$$J(\omega, b) = -\frac{1}{n}\sum_{i=1}^{n}\left[y_i \log \hat{y}_i + (1 - y_i)\log(1 - \hat{y}_i)\right],$$

- Where ω refers to weights, b is bias and n is number of samples. y and y are true are predicted values respectively.

Deep Neural Network

- The gradient descent iterative equations for minimizing the cost are the following:

$$\omega := \omega - \alpha \frac{\partial J(\omega, b)}{\partial \omega}$$

and

$$b := b - \alpha \frac{\partial J(\omega, b)}{\partial \omega},$$

where α is called the learning rate.

- The final prediction ^y for a single neuron are then given by

$$\hat{y} = f(\omega^T \mathbf{X} + b),$$

Kriging-Bootstrapped DNN

- Our proposed hierarchical model generates a highly correlated intermediate data set from a Kriging model.
- Since the intermediate data is more correlated than the original dataset, it is expected that the DNN will spend less time in identifying a coherent pattern within the dataset from which an input-output relationship can be built, hence reducing the training time.

Training or Learning Process

- Given a training data of size n. It is modeled as the number of known locations.
- We set a bootstrap size B and then bootstrap by Kriging. The output size of the bootstrapped Kriging is n+B.
- The "n+B" correlated output is passed into the Deep Neural Network for final training.
- The output of the DNN is binary, with "1" representing a detected seizure state and "0" otherwise.

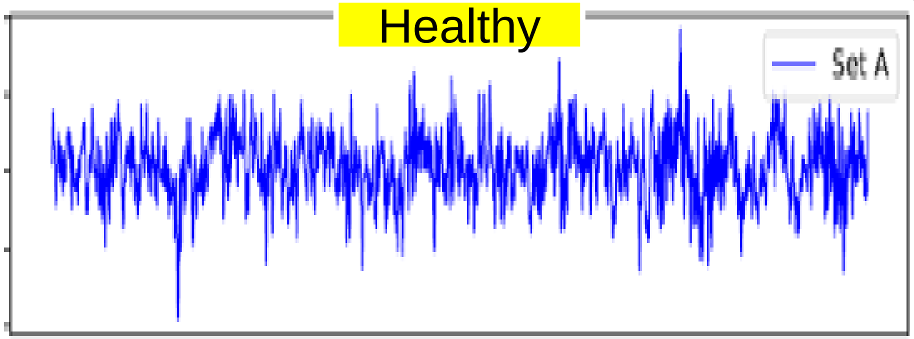
Smart Electronic Systems Laboratory (SESL)

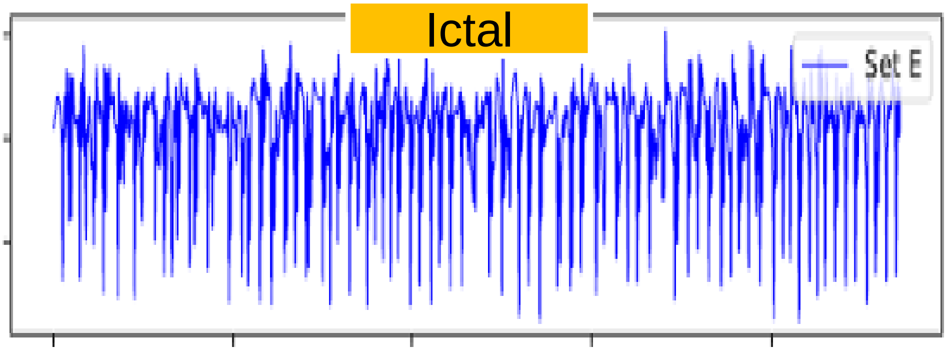
UNT EST. 1890

Experimental Results - EEG Dataset

BONN DATASET

- The datasets were originally collected from 5 healthy volunteers & five epilepsy patients by the University of Bonn. 5 different sets of data were collected as sets A, B, C,D&E.
- Sets A&B are healthy signals, C&D are inter-ictal signals while E is the only set with ictal signals. Each of the sets comprises 100 EEG segments which were collected with a 128-channel EEG system sampled at 173.61 Hz.

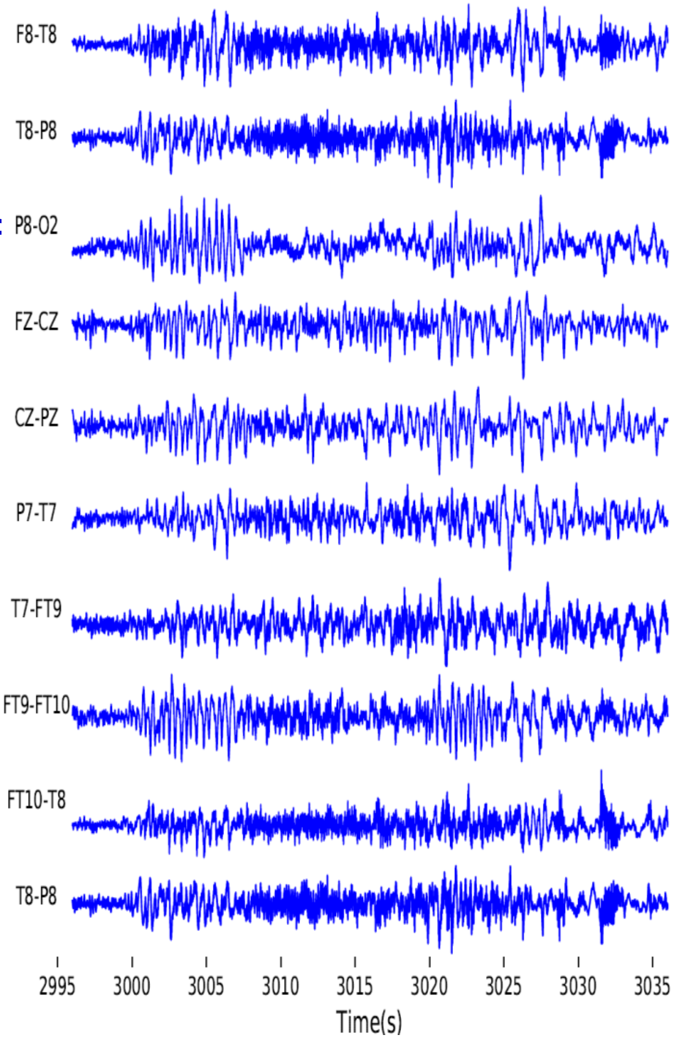




Experimental Results - EEG Dataset

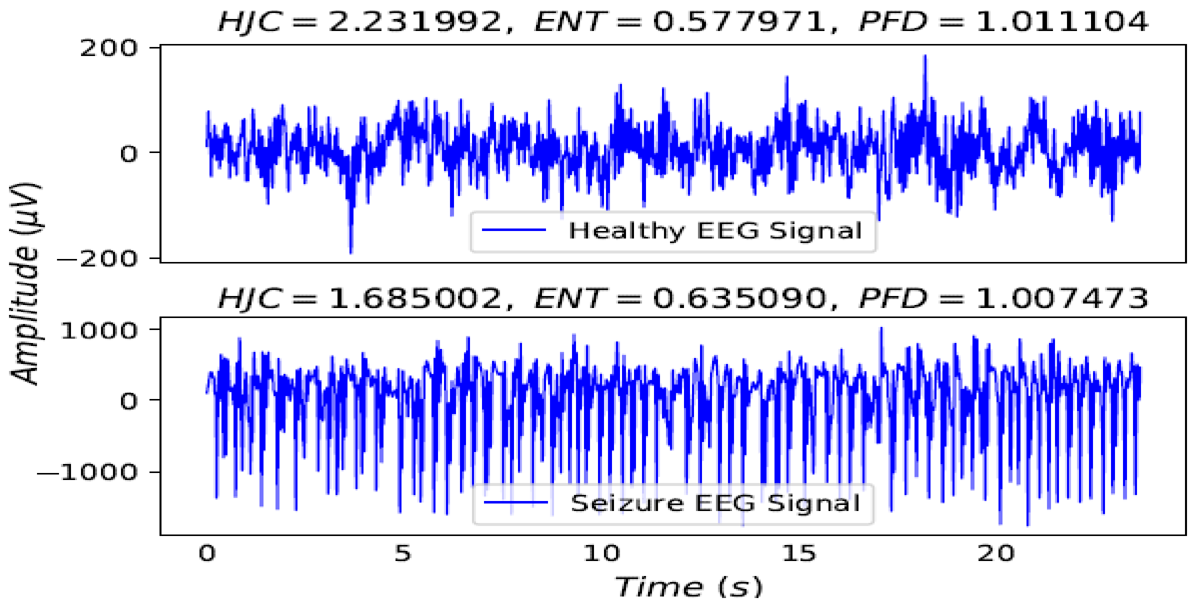
CHB-MIT SCALP EEG DATASET

- This dataset was collected at the Children Hospital Boston (CHB) in conjunction with the Massachusetts Institute Of Technology (MIT). It is therefore referred to as the CHB-MIT Scalp EEG Database.
- The EEG signals were collected from 22 epileptic patients of CHB using a 23-channel EEG, sampled at 256Hz and labeled according to the subjects as chb01 to chb23.
- The dataset consists of a total of 916 hours of continuous EEG recordings across all 22 subjects.



Extracted Features

- Signal Entropy
- **Fractal Dimension**
- Signal Power
- Standard Deviation
- **Singular Value Decomposition Entropy**
- Maximum Fractal Length
- **Hjorth Parameters**
- Hurst Exponent
- Lyapunov Exponent etc.



Experimental Results

TABLE I: Baseline DNN model performance results with 10,000 samples.

Count	Training Accuracy	Testing Accuracy	Training Epochs	Training Time
1	99.82%	80.00%	800	4.29s
2	99.85%	82.50%	1000	5.13s
3	99.95%	92.50%	10000	37.46s
4	99.99%	97.50%	45000	173.57s
5	99.99%	97.50%	50000	199.66s

TABLE II: Kriging-Bootstrapped DNN model performance with 10,000 samples.

Count	Training Accuracy	Testing Accuracy	Training Epochs	Training Time
1	99.14%	97.50%	500	41.07s
2	99.76%	100.00%	800	41.73s
3	99.84%	100.00%	1000	42.02s
4	99.92%	100.00%	1500	43.83s
5	99.92%	100.00%	10000	80.99s

Experimental Results

- DNN Architectural Specification.

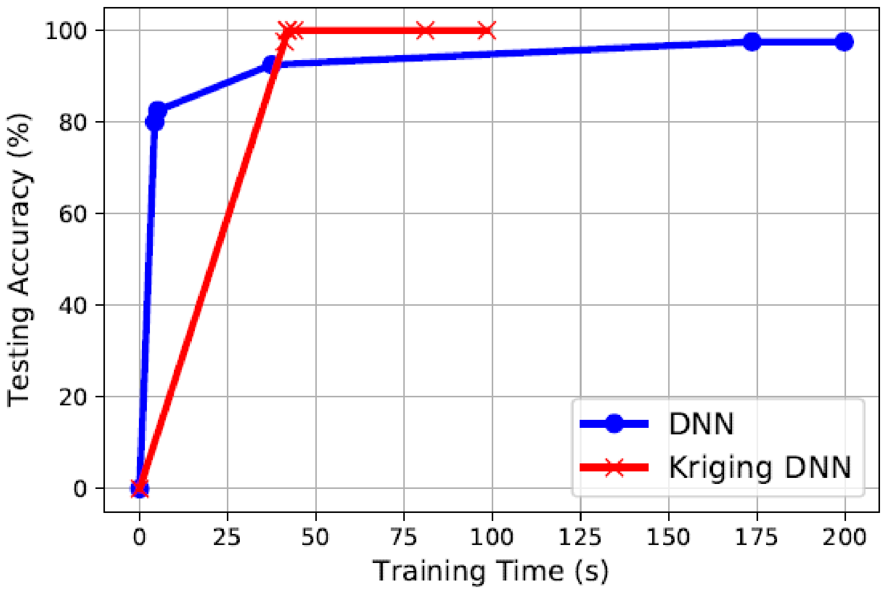
DNN Specifications	Values
No of Layers	4
Hidden Units	5, 5, 3
Hidden L. Activation	Rectified Linear Unit (ReLU)
Output L. Activation	Sigmoid Function
Initialization Method	Xavier Initialization [24]
Optimization Method	Adaptive Momentum [25]

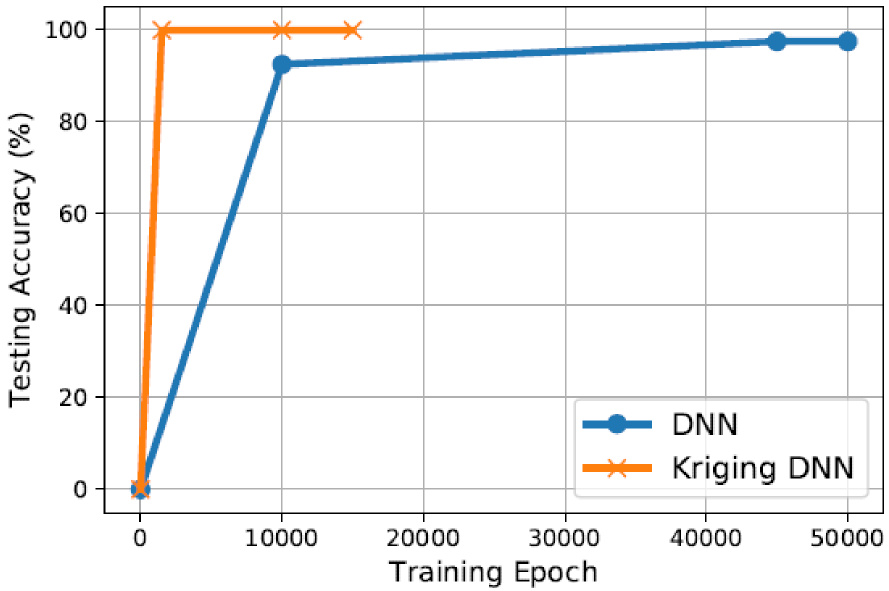
- Comparing best performances for baseline DNN model and the proposed kriging-bootstrapped DNN model.

Models	Training Accuracy	Testing Accuracy	Training Epochs	Training Time
DNN	99.99%	97.50%	45000	173.57s
Kriging DNN	99.92%	100.00%	1500	43.83s

- Proposed model trains in 75% less time and 30 times reduced training epoch size than the ordinary DNN, as well as a 2.5% improvement in testing accuracy.

Experimental Results





Smart Electronic Systems Laboratory (SESL)

Comparison with Related Works

Works	Extracted Features	Classification Algorithm	Accuracy	Sensitivity	Detection Latency
Shoeb, et al. 2010	Spectral, temporal and spatial features.	Support Vector Machine (SVM)	NA	96.00%	4.2 sec.
Zandi, et al. 2012	Regularity, energy & combined seizure indices	Cumulative Sum (CUSUM) thresholding	NA	91.00%	9 sec.
Altaf, et al. 2015	Digital hysteresis	Linear Support Vector Machine (LSVM)	NA	95.70%	1 sec.
Vidyaratne, et al. 2017	Fractal dimension, spatial/temporal features	Relevance Vector Machine (RVM)	99.80%	96.00%	1.89 sec.
Sayeed, et al. 2019	Hyper-synchronous pulses	Signal Rejection Algorithm (SRA)	NA	96.90%	3.6 sec.
Our WF-IoT 2020	Fractal dim., Hjorth complexity& Entropy	Kriging Classifier	97.50%	94.74%	**0.81 sec.**
Current Paper	Fractal dim., Hjorth complexity& Entropy	Kriging-Bootstrapped DNN	100.00%	100.00%	**0.80 sec.**

Conclusions

- This results in this presentation demonstrate the effectiveness of Kriging method for accurate and early seizure detection.
- The detection of seizure onset takes place in near real time with an average detection latency of 0.80 second which is better than previous models in the literature.
- The proposed model achieves a downward spiral in training time up to about 75% reduction compared to a baseline DNN model and improves the performance of the DNN by at least 2.5% after training on the same data size and the same DNN architecture.

Future Research

- In future, we will investigate seizure prediction, which means having prior knowledge that a seizure will occur before it actually does.
- Another future research is to have unified systems that detects seizure before it happens, and then injects drug or performs other control measures right after that.
- We also intend to add security and privacy features to the overall system as it is IoMT-enabled and always connected to Internet.
- We will also use more sophisticated and power-efficient edge devices such as IBM’s neurosynaptic hardware in validating our models.

References

1. I. L. Olokodana, S. P. Mohanty, and Elias Kougianos, "Ordinary-Kriging Based Real-Time Seizure Detection in an Edge Computing Paradigm", *in Proceedings of the 38th IEEE International Conference on Consumer Electronics (ICCE),* 2020, pp. in Press.

2. N. Cressie, "The Origins of Kriging," Math. Geology, vol. 22, no. 3, pp.239–252, Apr 1990.

3. R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," Physical Review E, vol. 64, no. 6, p. 061907, November 2001.

4. M. A. Sayeed, S. P. Mohanty, and E. Kougianos, "cSeiz: An Edge-Device for Accurate Seizure Detection and Control for Smart Healthcare," arXiv Electrical Engineering and Systems Science, vol. 1908.08130, Aug 2019.

5. S. P. Mohanty, V. P. Yanambaka, E. Kougianos, and D. Puthal, "PUFchain: Hardware-Assisted Blockchain for Sustainable Simultaneous Device and Data Security in the Internet of Everything (IoE)," arXiv Computer Science, vol. 1909.06496, Sep 2019.

THANK YOU