**Scalable EEG interpretation using Deep Learning and Schema Descriptors**

Joseph Picone1, PhD; Iyad Obeid1; Sanda M. Harabagiu, PhD2

1 Neural Engineering Data Consortium, College of Engineering, Temple University, Philadelphia, Pennsylvania, U.S.A.

2 Human Language Technology Research Institute, University of Texas at Dallas, Dallas, Texas, U.S.A.

This project addresses a critical market gap in EEG technology – real-time seizure detection for intensive care unit (ICU) and epilepsy monitoring unit (EMU) applications. The ability to auto-scan EEGs and predict seizures in advance will be a transformational clinical technology. Existing products that seek to increase accuracy and productivity via automatic analysis are limited by high rates of false positives, overwhelming healthcare providers with misleading information. Recently, deep learning systems have made tremendous progress in delivering powerful solutions from loosely transcribed data, due to rapid advances in low-cost, highly-parallel computational infrastructure and powerful machine learning algorithms. However, these techniques, which have been very successful in fields such as speech recognition and image understanding because large amounts of training data are available, have yet to be applied to biomedical signals such as EEGs due to a lack of big data resources. Hence, a major goal of this supplement is to complete a pilot study demonstrating the feasibility of applying these big data techniques to biomedical signals.

There are four specific aims: (1) automatic labeling of the TUH EEG Corpus for seizure events; (2) application of deep learning sequential modeling techniques for EEGs; (3) defining Hierarchical epileptiform Activity Descriptors (HAD) for EEGs; and (4) automated Tagging of HADs in medical texts.

The proposed work will produce several novel resources, including (1) a labeled subset of the publicly available TUH EEG Corpus that can support seizure detection research; (2) software tools that apply state of the art sequential modeling systems based on deep learning to biomedical signal processing; (3) The Hierarchical epileptiform Activity Descriptor (HAD) schema, which will provide a scalable solution for representing and annotating epileptiform activities, including those related to seizure prediction, both in biomedical and clinical documents; (4) automatic methods for generating HAD schema tags on EEG reports.

Submitting Author’s career stage: Professor