MSc in Computational Cognitive Neuroscience

Goldsmiths UNIVERSITY OF LONDON

Processing EEG data using Density Neural Networks

Student: Vitalie Cervinschi

Supervisor: Nikolay Nikolaev

Abstract

Electroencephalogram (EEG) remains a brain signal processing technique that allows one to gain understanding of the complex inner mechanisms of the brain and abnormal brain waves have shown to be associated with particular brain disorders. These time series data, model the electrical behavioral aspect of the brain's activity and a major objective of data modelling is to learn faithful how to evaluate such models by the use of signal features of the time series.

Introduction

Any model that we are trying to construct in order to simulate the desired process offer a degree uncertainty when running the model over the actual data; the two main reasons for this is data discrepancies and model discrepancies, which can be corrected up to a threshold based on the use of confidence intervals and prediction intervals.

Confidence intervals (CI) evaluate the model's uncertainty with the fluctuation of the standard error (SE). Generating accurate CI and prediction intervals (PI) is extremely important when modelling time series data; assuming that the arriving time series data passed as inputs to the model (in our case it is usually a neural network - NN) are noisy and we would like to evaluate the impact of these corrupted data on the variance of the model's output. Via the nonlinear connectionist function.

Leaning simultaneously the mean and the variance of the target distribution can be accomplished with a two-output NN architecture; having separate output nodes for the mean and the variance. The training of such a network can be performed following the ML method, which involves minimization of the following negative log-likelihood function:

$$C = -log \prod_{t=1}^{T} P(y_t, x_t) = \frac{T}{2} log(2\pi) + \frac{1}{2}\sum_{t=1}^{T} log(\sigma_t^2) + \frac{1}{2}\sum_{t=1}^{T}\left(-\frac{(y_t - f(x_t,w))^2}{\sigma_t^2}\right)$$

The Density NN have been especially developed for learning the input dependent confidence in the predictions when processing time series. The use of DNNs helps us to avoid overfitting both the mean and the time-varying variance. The use of DNNs can also be made for highlighting the novelties in the recorded time series and the use of those can be seen in a wide rage as medical systems and financial ones.

Novelties in time series

We can define novelties as unexpected values or sequence of values in the time series and they can help in the detection of machine failures or frauds in financial application or physiological disorders in patients which have the recorder of their EEG of ECG or EOG being recorded so that the physician can infer a proper treatment of intervention for the subject. The collection set of such examples can be found at http://www.cs.ucr.edu/□eamonn/discords. This set of signals was used to discover *discords* in time series, which can be defined as a sequence within the time series that is least similar to all other sequences (Adriano L, Silvio R. 2006) which can be of great benefit in case of medical research.

Hypothesis

The hypothesis that we want to test is whether we can apply this novelties of the CI to the EEG time series and to fit a DNN to the signal and infer physiological elements out it the outcome of the network.

Methods

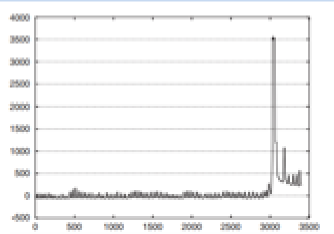
In order to test out hypothesis we need to train out DNN o a set of EEG recording sets that contain different physiological states (sleep, eye-blink and breathing) or different task performing experiments. We process the data sets by using the gradient-descent technique to train the DNN to produce the PIs. We will divide the training of the network in three phases:

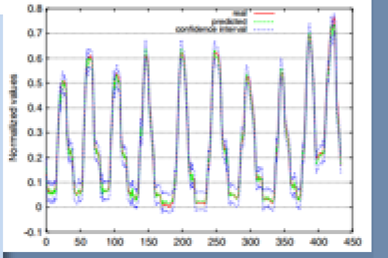
- The weight of the variance-learning part of the network are trained using the BP algorithm.
- The second part of the training in when the weight of the variance-learning part of the network are trained using different subsets from the data, while the mean-learning part is not trained
- The third phase both network parts are trained over all available data to minimize the negative log-likelihood function.

Expected Results

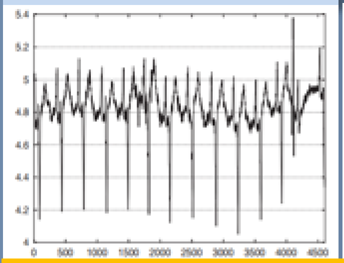
Experiments which contain time series novelties

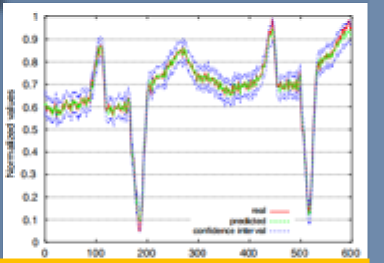
In the figure below, from the research paper of Adriano L, Silvio R 2006 we can see one such example or novelty in a time series which takes place at time 3050 onwards: the novelty belongs to the recording of a patient during sleep and we can see the deep breath and the awakening of the patient. In the next figure from the same study we can depict the case of the simulations of the rob-CI applied on the patient's respiratory time series:





The following two figures show a great example of time series with novelties in the signal of ECG recordings which elicit the presence of the unexpected value in the signal. The first figure only has the time series with novelty and the second with the robust CI for prediction in the rest set of reduced ECG series (from Adriano L, Silvio R 2006)





References

Adriano L.I.Oliveira,Silvio R.L.Meira "*Detecting novelties in time series through neural networks forecasting with robust confidence intervals*", Volume 70, Issues 1–3, December 2006, Pages 79-92

J. Gotman, J. R. Ives, and P. Gloor, "*Automatic recognition of inter-ictal epileptic activity in prolonged EEG recordings*", *Electroencephalogr. Clin. Neurophysiol*. Vol.46 no 5, pp 510-520, 1979

N. Nikolaev and H. Ilba (2006), "*Adaptive Learning of Polynomial Networks: Genetic Programming, Backpropagation and Bayesian Methods*".

D.A.Nix and E. Weigend (1994). "*Estimating the mean and the variance of the target probability distribution*", In: Proc.1994 IEEE Int. Conf.on Neural Networks (ICNN'94), Orlando, Fl, pp 55-60.

D.E.Rumelhart, R.Burbin, R.Golden and Y.Chauin, "Backpropagation: The basic theory".

Edelman B. J., Baxter B., and He B., "EEG source imaging enhances the decoding of complex righthand motor imagery tasks," IEEE Transactions on Biomedical Engineering, vol. 63, pp. 4–14, 2016. https://doi.org/10.1109/TBME.2015.2467312 PMID: 26276986

S. Maiti, R.K. Tiwari: "Neural network modelling and an uncertainty analysis in Bayesian framework: A case study from the KTB borehole site", Journal of Geophysical Research, Vol. 115, 2010, p. 28.

L. Cao, Support vector machines experts for time series forecasting,Neurocomputing 51 (2003) 321–339

E. Keogh, J. Lin, A. Fu, HOT SAX: efficiently finding the most unusual time series subsequence, in: Proceedings of Fifth IEEE International Conference on Data Mining (ICDM'05), 2005, pp. 226–233, longer version of the paper and datasets available at http://www.cs.ucr.edu/eamonn/discordsi.

Time series library hhttp://www-personal.buseco.monash.edu.au/hyndman/TSDLi.

F. Gonzalez, D. Dasgupta, R. Kozma, Combining negative selection and classification techniques for anomaly detection, in: Proceedings of IEEE Congress on Evolutionary Computation, 2002, pp. 705–710.