CARL VON OSSIETZKY UNIVERSITÄT OLDENBURG · 26111 OLDENBURG

Master Thesis

Investigate and evaluate the performance of various Machine Learning approaches to generate synthetic time series data from a reference sample of Experimental testdata.

Keywords: Data augmentation, Synthetic data generation, Time series data (EEG testdata)

Description: This topic is formulated in accordance to augment the available data for further data analysis in the SmartHelm project. Data is the primary element which plays a major role in Machine learning process pipeline. The amount of data indirectly defines the performance a model, during the process of training a conventional algorithm or complex Machine learning model. Actually there are many constraints which result in poor data acquisition or data accumulation during the progress of a project such as expensive experimental costs in case of acquiring data from Brain computer interaction (BCI) devices, data privacy constraints of the enterprises not permitting to transfer the complete data which may contain sensible information and, finally a case in which contact restrictions between people may not permit to continue with the further experimental data acquisition. Thereby all the above-mentioned examples may end up in insufficient amount about of data. This problem is addressed through or partially solved by using of Generative models as well as current state of the art approaches.

AIM: The main aim of this Topic is to find out the various standalone methods from the literature study, which are capable in generating synthetic data from the given sample of 16 channel time series EEG data. After aggregating different methods, the most preferable methods should be short listed in order to carry out the implementation on in-hand EEG data acquired from SmartHelm test experiments. The synthetically generated data after implementation of state-of-art approaches should be evaluated. Evaluation plays a vital role to benchmark the generated synthetic data as well as to avail this new set of datasets for the further data analysis.

- Study on Generative models
- Data Augmentation
- Synthetic data generation
- Benchmarking the obtained datasets

Depending upon the feasibility and the time constraint the further extension of above work would carrying out Data analysis with appropriate analysis algorithms perform basic classification or such relevant methods.

References:

1. Aznan, Nik Khadijah Nik, et al. "Simulating brain signals: Creating synthetic eeg data via neural-based generative models for improved ssvep classification." *2019 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2019.
2. Kiebel, Stefan J., et al. "Dynamic causal modelling for EEG and MEG." *Cognitive neurodynamics* 2.2 (2008): 121.

Falls Sie Interesse an einer Bearbeitung des Themas haben, melden Sie sich bitte bei:

M. Sc. Harish Moturu (harish.moturu@uni-oldenburg.de)

DEPARTMENT FÜR INFORMATIK

ABTEILUNG
WIRTSCHAFTSINFORMATIK I
VERY LARGE BUSINESS APPLICATIONS

PROF. DR. JORGE MARX GÓMEZ

MITARBEITER
M.SC. HARISH MOTURU

TELEFONDURCHWAHL
017657782274

FAX
(0441) 7 98 – 4472

EMAIL
harish.moturu@uni-oldenburg.de

GEBÄUDE A4
Uhlhornsweg 84 – Raum A4 3-313

OLDENBURG
11.11.2020

VERY LARGE BUSINESS APPLICATIONS
Carl von Ossietzky Universität Oldenburg

POSTANSCHRIFT
D-26111 Oldenburg

PAKETANSCHRIFT
Ammerländer Heerstraße 114 - 118
D-26129 Oldenburg

TELEFONZENTRALE
(0441) 7 98 – 0

BANKVERBINDUNG
Landessparkasse zu Oldenburg
Kto. Nr.: 1 988 112
BLZ: 280 501 00
BIC: BRLADE21LZO