[Team 05] ProjF Proposal: Sleepy Time

Matthew Houk
mjhouk@ncsu.edu

Jeremy Park
jipark@ncsu.edu

Tyrone Wu
tkwu@ncsu.edu

1 MOTIVATION

Roughly 50-70 million people in the United States suffer from a chronic sleep disorder, and lack of sleep has been shown to increase the effects of depression, heart attack, and other health complications [6]. The goal of this project is to improve the deep learning methods, specifically convolutional neural networks (CNNs), that can automatically classify sleep stages in polysomnographic (PSG) sleep study data [2, 3, 5, 7]. In our proposed dataset, five different sleep stages were classified manually at 30 second epochs. Automating the sleep stage classification process would enable more efficient sleep pattern analysis, leading to improved public health. This project will support current research at NC State regarding the innovation of artificial intelligence-driven sleep technology [1].

2 DATA DESCRIPTION

For our project, we will use Haaglanden Medisch Centrum sleep staging database [2, 3]. The dataset is collected from 154 patients (88 Male, 66 Female) who were randomly sampled from a population of adults who suffer from varying sleep disorders.

From each patient, the following activities have been recorded for a single night:

- Electroencephalographic (EEG)
- Electrooculographic (EOG)
- Chin Electromyographic (EMG)
- Electrocardiographic (ECG)

In addition, annotations have been marked by domain experts with 30 second epochs to classify the scoring of sleep patterns. In each epoch, the score is categorized into one of the five stages:

- Stage W (Wakefulness)
- Stage N1 (Non-REM 1)
- Stage N2 (Non-REM 2)
- Stage N3 (Non-REM 3)
- Stage R (Rapid Eye Movement)

3 METHODOLOGY

To approach the task of classifying sleep stage from our dataset, we intend to compare against a classical machine learning methodology against a proposed model. Given the various modalities being used in our dataset, we intend to observe how existing approaches for EEG generalize to a hybrid dataset.

3.1 Baseline

For our baseline approach, we chose to use a Support Vector Machine (SVM). These are commonly used in the EEG classification task for comparison, as they perform well at distinguishing features from EEG data.[8].

3.2 Proposed Literature Model

We propose to make use of the EEGNet model, an open-source model from the US Army research laboratory which has been shown to perform well on several EEG paradigms, including P300, ERN, MRCP, and SMR [7]. We hope to demonstrate its ability to expand to non-EEG paradigms. For our model parameters, we intend to perform Bayesian hyperparameter tuning using sci-kit learn [4]. We also use the ADAM optimizer and the Categorical Cross-Entropy loss function to improve the accuracy of the model.

3.3 Proposed Novel Model

We also intend to investigate a Bayesian learning strategy, pending further investigation and an improved understanding of the required processes. We believe that a Bayesian Neural Network (BNN) will offer benefits in indicating when various external influences affect the integrity of the data.

4 EVALUATION

The model will be optimized using sleep stage classification accuracy as the metric. We will be performing a 10-fold cross-validation to ensure the reliability of results, with an 80/20 data split between training and validation data.

For evaluation of the results and comparison, we suggest the use of an Analysis of Variance (ANOVA). This will allow us to better understand the extent of benefit of the EEGNet approach over the classical SVM approach. Other methods of consideration which may be used include precision and recall or a confusion matrix.

REFERENCES

[1] 2021. NSF Award Search: Award # 2037328 - SenSE: AI-Driven, Resilient and Adaptive Monitoring of Sleep (AI-DReAMS). (2021). https://www.nsf.gov/awardsearch/showAward?AWD_ID=2037328&HistoricalAwards=false

[2] Diego Alvarez-Estevez and Roselyne Rijsman. 2021. Haaglanden Medisch Centrum sleep staging database. (2021). https://doi.org/10.13026/GP48-EA60 Version Number: 1.0.0 Type: dataset.

[3] Diego Alvarez-Estevez and Roselyne M. Rijsman. 2021. Inter-database validation of a deep learning approach for automatic sleep scoring. 16, 8 (2021), e0256111. https://doi.org/10.1371/journal.pone.0256111

[4] Ian Dewancker, Michael McCourt, and Scott Clark. 2016. Bayesian Optimization for Machine Learning : A Practical Guidebook. (2016). arXiv:1612.04858 http://arxiv.org/abs/1612.04858

[5] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley. 2000 (June 13). PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* 101, 23 (2000 (June 13)), e215–e220. Circulation Electronic Pages: http://circ.ahajournals.org/content/101/23/e215.full PMID:1085218; doi: 10.1161/01.CIR.101.23.e215.

[6] Institute of Medicine (US) Committee on Sleep Medicine and Research. *Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem.* National Academies Press (US). http://www.ncbi.nlm.nih.gov/books/NBK19960/

[7] Vernon J. Lawhern, Amelia J. Solon, Nicholas R. Waytowich, Stephen M. Gordon, Chou P. Hung, and Brent J. Lance. 2018. EEGNet: A Compact Convolutional Network for EEG-based Brain-Computer Interfaces. 15, 5 (2018), 056013. https://doi.org/10.1088/1741-2552/aace8c arXiv:1611.08024 Number: 5.

[8] Valeria Saccá, Maurizio Campolo, Domenico Mirarchi, Antonio Gambardella, Pierangelo Veltri, and Francesco Carlo Morabito. On the Classification of EEG Signal by Using an SVM Based Algorithm. In *Multidisciplinary Approaches to Neural Computing*, Anna Esposito, Marcos Faudez-Zanuy, Francesco Carlo Morabito, and Eros Pasero (Eds.). Vol. 69. Springer International Publishing, 271–278. https://doi.org/10.1007/978-3-319-56904-8_26 Series Title: Smart Innovation, Systems and Technologies.