

Chapter 81

Machine Learning Methods as a Test Bed for EEG Analysis in BCI Paradigms

Kusuma Mohanchandra

Dayananda Sagar College of Engineering, India

Snehanshu Saha

PESIT-South, India

ABSTRACT

Machine learning techniques, is a crucial tool to build analytical models in EEG data analysis. These models are an excellent choice for analyzing the high variability in EEG signals. The advancement in EEG-based Brain-Computer Interfaces (BCI) demands advanced processing tools and algorithms for exploration of EEG signals. In the context of the EEG-based BCI for speech communication, few classification and clustering techniques is presented in this book chapter. A broad perspective of the techniques and implementation of the weighted k-Nearest Neighbor (k-NN), Support vector machine (SVM), Decision Tree (DT) and Random Forest (RF) is explained and their usage in EEG signal analysis is mentioned. We suggest that these machine learning techniques provides not only potentially valuable control mechanism for BCI but also a deeper understanding of neuropathological mechanisms underlying the brain in ways that are not possible by conventional linear analysis.

INTRODUCTION

EEG-based brain-computer interface (BCI) has assumed a significant role towards aiding the study and understanding of neuroscience, machine learning, and rehabilitation in the recent years. BCI could be interpreted as a platform for direct communication between a human brain and a computer bypassing the normal neurophysiology pathways. The primary goal of BCI is to restore communication in severely paralyzed population. However, the BCI for speech communication has its applications extended to silent speech communication, cognitive biometrics, and synthetic telepathy (Mohanchandra, Saha, & Lingaraju,

DOI: 10.4018/978-1-7998-2460-2.ch081

2015). Electroencephalography (EEG) is a non-invasive interface, which has high potential due to its superior temporal resolution, ease of handling, portability, and low set-up cost. A general method for designing BCI is to use EEG signals extracted during mental tasks. EEG is the recording of the brain's spontaneous electrical activity from multiple electrodes placed on the scalp. EEG can be altered by motor imagery (Mohanchandra, Saha, & Deshmukh, 2014, pp. 434 - 439) and can be used by patients with severe motor impairments to communicate with their environment and to assist them. Such a direct connection between the brain and the computer is known as an EEG-based BCI.

An extensive exploration of the voluminous literature reveals a gap in the ability to provide speech communication using brain signals to produce meaningful words. Scientific endeavor is directed in the direction of developing a BCI-to-speech communication using the neural activity of the brain through subvocalized speech. Subvocalized speech is talking silently in the mind without moving any articulatory muscle or producing overt activities. The electrical signals generated by the human brain during subvocalized speech are captured, analyzed and interpreted as speech. The book chapter intends to dwell upon the characterization of subvocalized speech, captured via EEG signals. Since EEG signals suffer from poor spatial resolution (Kusuma, & Snehanshu, 2014, pp. 64-71), classification of mental activities using subvocalized speech is a challenging problem and according to the authors, the final frontier of machine learning. EEG signals suffer from the curse of dimensionality due to the intrinsic biological and electromagnetic complexities. Therefore, selecting a representative feature subset which would reduce the size of the dataset without compromising the quality of the information is a research problem worth attention. The problem this chapter would detail on involves efficient classification of subvocalized speech by novel subset selection method minimizing loss of information.

The EEG is acquired through multichannel sensors for a couple of seconds that leads to a large dataset. The deployment of multiple sensors leads to a significant volume of data. Gathering and maintaining the huge amount of data is a challenge and extracting useful information from them is even more demanding. Machine learning is a solution to this problem which helps researchers to evaluate the data in real-time. The machine learning algorithm is used to discover and learn knowledge from the input data. The category and bulk of the data affect the learning and prediction performance of the algorithms. Machine learning algorithms are characterized as supervised and unsupervised methods, also known as predictive and descriptive, respectively.

Supervised methods are built on the training set of the features and corresponding class labels known with confidence. The algorithm is trained on this set of features, and the result is applied to other features of which the target label is not mentioned. In contrast, unsupervised methods do not label data into classes. Unsupervised algorithms require some initial input to one or more of the adjustable parameters, and the solution obtained depend on the input given. It is to be noted that, the success of any predictive or analytic algorithm depends on efficient feature selection paradigms. Feature selection is a dimensionality reduction that efficiently identifies and generates discriminatory features among different classes of data as a compressed feature vector.

This book chapter is arranged as follows. Section 2, illustrates the broad outlook of the data and the methods used for experimentation. Section 2.2, presents the formation and evaluation of the Subset Selection Method (SSM), for feature selection. The SSM algorithm selects a subset of features that has significant variances which have impressive characteristics. The features with low variances are omitted from the feature space as they represent outliers and noise. The algorithm is tested on the EEG signals of subvocalized speech.

19 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/machine-learning-methods-as-a-test-bed-for-eeg-analysis-in-bci-paradigms/252099

Related Content

Managing Seven Dimensions of ICT4D Projects to Address Project Challenges

Devendra Potnis (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 615-637).

www.irma-international.org/chapter/managing-seven-dimensions-of-ict4d-projects-to-address-project-challenges/252048

Password-Less Authentication: Methods for User Verification and Identification to Login Securely Over Remote Sites

Rahul Singh Chowhanand Rohit Tanwar (2019). *Machine Learning and Cognitive Science Applications in Cyber Security* (pp. 190-212).

www.irma-international.org/chapter/password-less-authentication/227582

Explore the Use of Handwriting Information and Machine Learning Techniques in Evaluating Mental Workload

Zhiming Wu, Tao Linand Ningjiu Tang (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 1423-1439).

www.irma-international.org/chapter/explore-the-use-of-handwriting-information-and-machine-learning-techniques-in-evaluating-mental-workload/252090

A Taxonomy of Data Mining Problems

Nayem Rahman (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 512-528).

www.irma-international.org/chapter/a-taxonomy-of-data-mining-problems/252041

Brave New Digital Tools for Action Research in Education: A Beginner's Guide

Reinhard Bauer, Klaus Himpsl-Gutermann, Martin Sankofi, Petra Szucsichand Ruth Petz (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 469-490).

www.irma-international.org/chapter/brave-new-digital-tools-for-action-research-in-education/252039