Review of: "A Simple Preprocessing Method Enhances Machine Learning Application to EEG Data for Differential Diagnosis of Autism"

Fabiola De Marco[1]

1 University of Salerno

Potential competing interests: No potential competing interests to declare.

The authors present work in which EEG data are used to distinguish between children with autism spectrum disorders (ASD) and children with other neuropsychiatric disorders (NPD). The proposed method employs transforming EEG data into Manhattan distance matrices and calculating minimum spanning trees, followed by using a machine learning algorithm (KNN) for classification. The results show an overall accuracy of 93.2% in distinguishing the two groups of subjects. Furthermore, a clustering method is presented that allows an almost perfect separation of subjects based on their diagnosis.

ABSTRACT

It is not clear that the 38 figures are the entire EEG file transformed into just 38 numbers.

INTRODUCTION

RELATED WORKS: “Many different mathematical approaches have been tested in the last few years to disentangle the EEG data complexity and determine if it is possible to distinguish children with ASD from typically developing children or children with other neuropsychiatric disorders”. Previous studies using these methods are not addressed. No parts following the introduction have focused on presenting studies using EEG data for diagnostic purposes. A state-of-the-art analysis is needed that provides information on

- Diagnosis based on EEG

- Conventional diagnostic methods for ASD

- Diagnosis of ASD with EEG.

Understanding the contribution of the presented work in relation to existing scientific knowledge is fundamental and requires a state-of-the-art analysis.

PRE-PROCESSING: Different preprocessing methods are mentioned (CSP, PCA, CAR, SL, ICA), but only the advantages and disadvantages of PCA are explained. For an examination of the pre-processing methods that justifies the methodological choices of the study, it is necessary to illustrate the advantages and disadvantages of the other methods

mentioned.

INTRODUCTION STRUCTURE OF THE PAPER: A brief presentation of the subsequent sessions is absent.

SELECTION OF CONTROL SUBJECTS: Control subjects were selected from a wide variety of neuropsychiatric disorders, including ADHD, mood disorders, anxiety disorders, sleep disorders, and traumatic brain injury. However, this selection does not reflect the diversity of neuropsychiatric disorders present in the general population, and therefore a differential diagnosis between autism and another neuropsychiatric disorder may be more complex. This would affect the value of the results.

EEG DATA COLLECTION: Although the EEG recording was performed under standardized conditions (rest with eyes closed), there may be other variables not considered that could influence the results, such as state of alertness or taking medications. Information on the possible intake of pharmacological therapies is absent in the description of the sample.

MACHINE LEARNING CLASSIFIERS: Although the KNN algorithm has been used as a classifier, there may be other machine learning algorithms that are better suited for this type of data.

RESULTS AND DISCUSSION

LACK OF COMPARISON WITH OTHER METHODOLOGIES: No direct comparison is made with other diagnostic or machine learning methodologies used for the diagnosis of autism. A comparison with other techniques could provide a more complete evaluation of the effectiveness of the proposed model.

The results section lacks details. The discussion part includes information that would ideally have been introduced at the beginning of the work, outlining the state of the art.

The reasons why this work improves the current state of the art are not convincingly presented.