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

Krzysztof Wołk

Potential competing interests: No potential competing interests to declare.

This critical review examines the paper titled "A Simple Preprocessing Method Enhances Machine Learning Application to EEG Data for Differential Diagnosis of Autism" by Enzo Grossi, Rebecca White, and Ronald J. Swatzyna. The paper presents a novel preprocessing approach for EEG data to enhance machine learning's effectiveness in differentiating Autism Spectrum Disorder (ASD) from other Neuro-Psychiatric Disorders (NPD).

Strengths:

1. **Innovative Preprocessing Method:** The study introduces a unique preprocessing technique using a triangular matrix of Manhattan distances among EEG electrodes, followed by the calculation of a Minimum Spanning Tree (MST). This method simplifies the EEG data into a more manageable form for machine learning analysis, potentially making it a significant contribution to the field.

2. **High Accuracy of Results:** The use of the KNN algorithm achieved a high accuracy rate (93.2%) in differentiating between ASD and NPD subjects, which is notable for its potential clinical implications.

3. **Efficiency and Simplification:** The study successfully reduces complex EEG data to 38 figures, which allows for faster processing and reduced computational time, a significant advantage in practical applications.

Limitations and Areas for Improvement:

1. **Sample Size and Diversity:** The study sample is limited to 50 ASD children and 50 children with other NPDs, matched for age and gender. Expanding the sample size and including a more diverse range of neurodevelopmental disorders could improve the generalizability of the findings.

2. **Lack of External Validation:** The study primarily relies on internal validation. External validation with different datasets and EEG acquisition systems would strengthen the confidence in the model’s generalizability and robustness.

3. **Potential Overfitting and Bias:** While the preprocessing method reduces the data to 38 numbers, there's a risk of overfitting the machine learning model to this specific dataset. More rigorous validation methods, such as leave-one-out cross-validation or testing with independent datasets, could address this concern.

4. **Comparison with Other Preprocessing Methods:** The paper could benefit from a comparative analysis with other

EEG preprocessing methods, providing a clearer context of the advantages and limitations of their proposed method.

5. **Clinical Applicability:** While the high accuracy is promising, the real-world application requires further investigation. The study should consider factors such as the variability in EEG recording conditions in clinical settings and the integration of this method into existing diagnostic protocols.

Conclusion:

The paper by Grossi et al. contributes a novel EEG preprocessing method that shows promising results in the differential diagnosis of ASD. The high accuracy and efficiency of their method are notable strengths. However, the study's limitations, particularly concerning sample size, lack of external validation, and potential overfitting, suggest the need for further research. Future studies should focus on external validation with larger and more diverse samples, comparative analyses with other methods, and exploration of the method's integration into clinical practice.