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

Vijayakumar Vijayakumar

Potential competing interests: No potential competing interests to declare.

Limited Generalizability: The study only includes EEG data from a specific population of 50 ASD children and 50 children with other NPD, matched for age and gender. The results may not generalize well to other populations, such as adults with ASD or individuals with different neurodevelopmental or psychiatric disorders.

Lack of External Validation: The paper mentions a rigorous validation protocol, but it's unclear whether the model's performance has been validated on an external dataset. Without external validation, the generalizability and reliability of the predictive model may be limited.

Potential Overfitting: The high accuracy of 93.2% obtained by the KNN algorithm raises concerns about potential overfitting, especially given the relatively small sample size. The paper should provide details on how overfitting was addressed, such as through cross-validation or regularization techniques.

Interpretability of Features: While the paper mentions using a new pre-processing approach to extract topological EEG features, it's unclear how these features relate to underlying neurophysiological mechanisms or clinical characteristics. More transparency about the interpretability and clinical relevance of the features could strengthen the paper's contribution.

Lack of Comparison with Existing Methods: The paper does not compare the proposed approach with existing methods for diagnosing ASD or distinguishing between ASD and other psychiatric disorders using EEG data. Including such comparisons would help contextualize the findings and assess the novelty and effectiveness of the proposed approach.

Computational Efficiency: While the paper mentions achieving modest computational time and reducing the information to 38 figures, it's unclear how these figures were selected and whether further optimization is possible. Providing more details about the computational efficiency and scalability of the proposed approach would enhance its practical utility.

Ethical Considerations: The paper does not discuss potential ethical implications of using EEG data for diagnostic purposes, such as privacy concerns, data security, and informed consent. Addressing these ethical considerations is important for the responsible conduct of research involving human subjects.