November 2, 2019

To Whom It May Concern:

The manuscript titled “Objective evaluation metrics for automatic classification of EEG events” discusses a very serious issue with respect to performance evaluation of machine learning systems in the biomedical engineering community. The EEG community, specifically, uses lenient scoring metrics which do not correctly characterize the accuracy of EEG event detection in sequential signals.

In this paper, we discuss a few popular traditional performance metrics and introduce two additional metrics which specifically focus on the time alignment and duration of a hypothesis. This is significant because the standardization of the proposed metrics would (1) help more effectively compare research in the field, (2) enable better diagnostics, (3) provide a more realistic view of the machine learning system’s performance, and (4) promote clinical adoption of the technology because these performance metrics better reflect performance in real applications.

We have provided this code as open source software to promote its adoption in the community.

All evaluation metrics discussed in this article and their results are not under consideration or published in another journal.

Best regards,



Dr. Joseph Picone

Professor, Department of Electrical and Computer Engineering