Ann. Geophys. Discuss.,
https://doi.org/10.5194/angeo-2019-149-RC1, 2019

Annales Geophysicae Discussions

Interactive comment on "Automatic detection of the Earth Bow Shock and Magnetopause from in-situ data with machine learning" *by* Gautier Nguyen et al.

Anonymous Referee #1

Received and published: 26 November 2019

This paper presents an application of machine learning for the automatic classification of satellite data, specifically to distinguish between magnetosphere, magnetosheath and solar wind regions. The paper is interesting and the machine learning approach is novel. However, there are several flaws that need to be fixed before publication.

1) The authors seem to be aware that a random splitting of the data between training and test set is yields erroneous scores (line 111). Yet, they still present results based on random split. I suggest to completely remove the results obtained in this way and to present only the results obtained with a more correct split in time.

2) The labeling of the data is completely unclear. Reading from line 53 it seems that

they are mixing the 'ground truth' with the result of the classification algorithm. The same argument is repeated on line 118. Obviously you cannot use the same algorithm to label and predict.

3) Similarly, on line 134 they do not explain how do they label such large amount of Cluster data. Similar for other data.

4) In evaluating the model performance, the authors focus exclusively on the AUC. I suggest to compute and show also the True Skill Score and the Heidke Skill Score, that are standard skill scores.

5) Figure 3 is not very informative. It looks like the model can achieve perfect predictions?

6) Line 75: a reference is needed to the original works on boosting algorithms. 7) Line 76: the compute time depends on the CPU (and/or GPU)

8) Line 81: should be 'has been predicted'

9) Line 197: why a decision tree should require less time than an arbitrary decision boundary?

10) Line 199: Decision trees are also threshold-based methods. I do not understand this distinction.

11) Line 203: don't you have cross-calibration issues when you employ the algorithm trained on one satellite to make inference on another satellite?

12) Since decision trees are easily interpretable, it would be interesting to visualize the boundaries and understand their physical implications.

13) Line 226: it is well-known that the probabilities output of boosted ensemble of decision trees are not well-calibrated.

14) Line 275: Do Appendixes B and C really have no text?

15) Finally I suggest to be more specific in the title, by replacing 'Machine Learning' with 'Decision Trees'

Interactive comment on Ann. Geophys. Discuss., https://doi.org/10.5194/angeo-2019-149, 2019.