Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-495-RC1, 2020

Atmospheric Measurement Techniques Discussions

Interactive comment on "Classification of Lidar Measurements Using Supervised and Unsupervised Machine Learning Methods" *by* Ghazal Farhani et al.

Benoît Crouzy (Referee)

benoit.crouzy@meteoswiss.ch

Received and published: 8 September 2020

The authors apply various machine learning techniques (ML) to classify Lidar measurements. They show the potential of different approaches: supervised technique in the presence of known categories and unsupervised techniques in order to detect anomalies in the signal resulting from unusual events (e.g. fires). The paper begins with an extensive introduction on machine learning which could be useful to the Lidar community, and the chosen applications show a good panel of possibilities to apply ML on lidar measurements.

Below I list my comments in the order as they appear in the manuscript.

1) Page 1 line 10: selection of the training and test set. This is a potential major issue. The authors selected randomly scans over the years. In order to have good generalization of the algorithms it however is important to have not too close training and test data. The total number of raw scans out of which events were randomly drawn suggests that this is achieved but I would like the authors to comment on this question, even if only as a caveat to the community. Best practice could be to select a period isolated from the training set to select the test set (e.g. different year).

2) Page 2 lines 3-8: please check for some repetitions (laser fluctuations)

3) Page 2 line 17-18: I suggest to distinguish between supervised and non-supervised from the onset.

4) Page 2 line 26: I would put this line earlier to distinguish between supervised and non-supervised techniques. In addition I would hint for non-specialists that unsupervised techniques are no silver bullet and can be expected to be less powerful due to the absence of training data.

5) Page 2 line 27: "clustering ML" and line 30 "These ML methods" please precise which methods.

6) Page 3 line 23: I find the sigma confusing (summation sign), especially when keeping signed differences.

7) Page 3 line 25: what about "matrix size (m,n)" or m x n ?

8) Page 4 line 13-14: please describe better the Kernel trick that make SVM so powerful. In the current form I do not find the description self-contained (non-uniform level of details). I would describe the parameters to be tuned (e.g. Cost, epsilon insensitive tube, ...)

9) Page 4 line 25: please define all variables

10) Page 5 line 3: maybe I missed it, but if you used LIBSVM or a derived tool please

mention it, as this information could help other users.

11) Page 5 line 25: I found this sentence somewhat disconnected.

12) Equation 3: define the class index.

13)Page 5 line 29: H=0 usually means low entropy which could be seen as a pure prediction. Please clarify.

14) General comment: I would summarize for all methods the hyperparameters to be tuned. This is currently well done only for some of the methods.

15) Page 7 line 1: "a detailed description" this sentence and the introduction to various methods give the impression of ML as a closed list of techniques. As the Lidar community is not very familiar to ML, I would mention that a vast number of other techniques exist. I would also mention ANNs and explain why those were not used in the present paper.

16) I would remove Figure 2 (too much detail in comparison with the rest of the chapter), but this is a matter of personal taste.

17) Page 8 line 15: from and not form

18) Section 3.1: see general comment above, how was it ensured that enough separation between training situations and test situations is achieved (scans taken the same day/hour might not always achieve this). This point needs to be discussed carefully.

19) Section 3.2: In my opinion the discussion on TP/FP/TN/FN does not belong to the results but to the methods.

20) Page 11 line 6: to estimate

21) Page 14 line 11: anomalies and not anomolies

22) General comment: why not making use of thresholds in order to achieve finer objectives, eg. "never tagging good scans as bad"? For example, with SVMs one can

use the distance to the hyperplane to select events with a good likelihood of correct classification.

In all, I recommend the paper for publication after the questions raised are satisfactorily addressed by the authors.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-495, 2020.