Drinking Water Engineering and Science Discussions

Drink. Water Eng. Sci. Discuss., referee comment RC1
https://doi.org/10.5194/dwes-2021-7-RC1, 2021

Comment on dwes-2021-7

Anonymous Referee #1

Referee comment on "Implementing and evaluating various machine learning models for pipe burst prediction" by Ahmad Ravanbakhsh et al., Drink. Water Eng. Sci. Discuss., https://doi.org/10.5194/dwes-2021-7-RC1, 2021

This paper presents an application of 5 existing statistical and machine learning methods for pipe burst prediction.

I suggest to reject this paper based of several major issues, such as:

1. There is very little novelty in this paper, which is essentially a poorly done machine learning exercise that would should be rather featured in a blog on Kaggle, rather than in a scientific journal; known techniques, different dataset, no advancing of the state-of-the-art.

2. Most of the paper is about the description of existing techniques, which have been described several times in the literature. The authors employ the RCNN-SVR technique but they fail to cite the correct paper [1] both in the references (Line 375, a paper from 2016) and in the text (Line 148, a paper from 2018). They lack to provide the required details to understand the model, e.g., are those 1D convolutional layers? How many trainable parameters total?

3. The latter information is particularly important because Deep Learning models usually have hundreds of thousand of parameters, and the proposed dataset is made of barely 200 data points. Deep Learning is employed when BIG DATA is available, this is clearly not the case. This raises the issue of whether the models have been trained appropriately or not. In line 52, the authors claim that 85% of data have been selected for training and the rest of them have been used to test the models. Was the training data further split to create a separate validation dataset? Or is the test set used for "validation" and model selection? Is this a fair comparison? Given the data is tabular, not an image or a time-series, I am confident tree-based algorithms such as Random Forest, ExtraTrees or Gradient Boosting trained using a single line of Python would provide at least the same results. Why did the authors use outdated benchmarks for comparison? Most importantly, the authors failed to provide the details needed to understand how they developed and compared the models.

4. The paper lacks organization and structure. Experimental setup (e.g., train/test division) is presented in the Methodology, so are the final trained models. The dataset is cited in the Methodology but introduced in a subsequent section. The results and

discussion section is minimal, and there is no discussion at all.

5. The literature review is outdated and incomplete, see for instance the papers [2-4] reproted at the end of this review.

6. The paper is difficult to read even for an expert in the field.

References:

[1] Zhang, Youshan, and Qi Li. "A regressive convolution neural network and support vector regression model for electricity consumption forecasting." Future of Information and Communication Conference. Springer, Cham, 2019.

[2] Konstantinou, Charalampos, and Ivan Stoianov. "A comparative study of statistical and machine learning methods to infer causes of pipe breaks in water supply networks." Urban Water Journal 17.6 (2020): 534-548.

[3] Snider, Brett, and Edward A. McBean. "Improving urban water security through pipe-break prediction models: Machine learning or survival analysis." Journal of Environmental Engineering 146.3 (2020): 04019129.

[4] Zhou, Xiao, et al. "Deep learning identifies accurate burst locations in water distribution networks." Water research 166 (2019): 115058.