Real-time flood overflow forecasting in Urban Drainage Systems by using time-series multi-stacking of data mining techniques

Piadeh, Farzad ORCID logoORCID: https://orcid.org/0000-0002-4958-6968, Behzadian, Kourosh ORCID logoORCID: https://orcid.org/0000-0002-1459-8408, Chen, Albert, Campos, Luiza and Rizzuto, Joseph (2023) Real-time flood overflow forecasting in Urban Drainage Systems by using time-series multi-stacking of data mining techniques. In: EGU General Assembly 2023, 23-28 April 2023, Vienna, Austria.

https://doi.org/10.5194/egusphere-egu23-8574

EGU23-8574
EGU General Assembly 2023

Real-time flood overflow forecasting in Urban Drainage Systems by using time-series multi-stacking of data mining techniques

Farzad Piadeh[1], Kourosh Behzadian[1], Albert S. Chen[2], Luiza C. Campos[3], and Joseph P. Rizzuto[1]
[1]University of West London, School of Computing and Engineering, London, United Kingdom of Great Britain – England, Scotland, Wales (farzad.piadeh@uwl.ac.uk)
[2]Centre for Water Systems, Faculty of Environment, Science and Economy, University of Exeter, Exeter, EX4 4QF, UK
[3]Department of Civil, Environmental and Geomatic Engineering, University College London, Gower St, London, WC1E 6BT, UK

Overflow forecasting in early warning systems is acknowledged as an essential task for devastating urban flood risk management. Although many machine learning models have been developed recently to forecast water levels in urban drainage systems (UDS), they usually require big and accurate data resources [1]. Alternatively, ensemble data mining models are becoming more popular, in which time-series numerical data are turned into the categorised features that classify wet weather conditions as two classes of overflow and non-overflow conditions [2]. However, the concept of time-series ensemble modelling i.e., blending different data mining techniques for predictions with different timesteps is still new [3]. Furthermore, the application of more advanced models, particularly multi-blending in these types of ensemble modelling requires more investigation. This study aims to introduce a novel multi-stacking model that integrates different decision tree frameworks by developing various base weak learner data mining techniques and associated base model performance indicators in the process of time-series blending of pre-trained stacked ensemble models. The performance of this new approach is compared by several previously developed ensemble models [2] through confusion matrix performance criteria, including hit rate, overestimation, and underestimation. This method is demonstrated by its application to a real case study of UDS located in the northwest of London for performance assessment up to 5hr ahead (i.e., 20 timesteps with 15-min intervals). In total, 140 base-models and 20 stacked models were developed that are stored in the data warehouse to use as real-time early-warning flood overflow forecasting for this case study. These developed models were used through introduced decision three framework that specified stacking blending methodology. Results show that while base models and stacked models suffer from high miss rate, especially for forecasting more than 3hrs ahead (more than 50%), the proposed multi-stacking model could perfectly maintain the miss rate (i.e., sum of over- and under-estimations) of up to 4hr-ahead predictions less than 10%, but this rate dropped to nearly 30% for 5hr-ahead predictions. However, the rate of overflow forecasting remained acceptably near 80% whereas it is recorded to less than 58% for other benchmark models. Using different decision frameworks for determining importance of each stacked model in blending mode of multi-stacking method shows could reduce errors in forecasting rate and take advantage of each model in real-time early warning urban flood forecasting.

References

[1] Piadeh, F., Behzadian, K., Alani, A. (2022). A critical review of real-time modelling of flood forecasting in urban drainage systems. *Journal of Hydrology*, 607, 127476.

[2] Granata, F., Di Nunno, F., de Marinis, G. (2022). Stacked machine learning algorithms and bidirectional long short-term memory networks for multi-step ahead streamflow forecasting: A comparative study, *Journal of Hydrology*, 613(A), 128431.

[3] Piadeh, F., Behzadian, K., Alani, A.M. (2022). Multi-Step Flood Forecasting in Urban Drainage Systems Using Time-series Data Mining Techniques. *Water Efficiency Conference*, West Indies, Trinidad and Tobago. repository.uwl.ac.uk/id/eprint/9690 [Accessed 31/12/2022].