

## Classifier Directed Prediction of Extreme and Non-Extreme Events in Time Series

Time series consisting of non-extreme and extreme events. Goal predict both accurately.

Issues:

Unbalanced data: large number of non-extreme events (similar segments) and a few extreme events. Both decrease generalization (over-trained and under-trained). Over-trained: too good a fit to the training data which leads to poor performance on test data – over fitting. Under-trained: does not capture the underlying patterns

Single T.S prediction model: cannot expect high prediction accuracies for both events.

Address the classic bias-variance trade-off problem

Hybrid approach: A model for extreme events and a model for non-extreme events. Also a model for extreme vs non-extreme classification. Optimize each model. Each model can be different. For example, CNN model for extreme events, LSTM for non-extreme events, SVM for classifier.

Address issues related to optimizing each model.

Extreme event predictor – overcome underfitting

Non-extreme predictor – overcome training

Classifier: best classifier

Table 1 Expected prediction accuracies of the various models

Model	Entire Time Series	Non-Extreme Events	Extreme Events
Standard Predictor	↑	↑	↓
Non-Extreme Event Predictor	↑	↑	↓
Extreme Event Predictor	↓	↓	↑
Event Classifier*		↑	↑
Classifier-Predictor	↑	↑	↑

\* classification accuracy

Use out flow data only. That is, predicting outflow using only past values of outflow.

Convert hourly outflow to daily outflow by ADDING the 24 hour outflow values (not averaging).

Table 2 MSEs of the various **LSTM** models

Model	Entire Time Series	Non-Extreme Events	Extreme Events
Standard Predictor			
Non-Extreme Event Predictor			
Extreme Event Predictor			
Event Classifier*			
Classifier-Predictor			

\* classification accuracy (%)

Table 3 MSEs of the various **CNN** models

Model	Entire Time Series	Non-Extreme Events	Extreme Events
Standard Predictor			
Non-Extreme Event Predictor			
Extreme Event Predictor			
Event Classifier*			
Classifier-Predictor			

\* classification accuracy (%)

The table below is just an example only. We can select the best sub-models based on the results from the Tables 2 and 3.

Table 4 MSEs of the various sub-models of the **Hybrid Model**.

Model	Entire Time Series	Non-Extreme Events	Extreme Events
Standard Predictor (LSTM)			
Non-Extreme Event Predictor (LSTM)			
Extreme Event Predictor (CNN)			
Event Classifier (CNN)*			

<b>Classifier-Predictor</b>			

\* classification accuracy (%)