NOVEL FEATURE REPRESENTATION STRATEGIES FOR PERIODIC TIME SERIES FORECASTING WITH PREDICTED FUTURE COVARIATES

Anonymous authors
Paper under double-blind review

ABSTRACT

Accurate time series forecasting is a fundamental challenge in data science. Unlike traditional statistical methods, conventional machine learning models, such as RNNs and CNNs, use historical data consisting of previously measured variables including the forecast variable and all its covariates. However, in many applications, some of the covariates can also be predicted with reasonable accuracy for the immediate future. We refer to such covariates as *future covariates*. We consider the problem of predicting water levels at a given location in a river or canal system using historical data and future covariates. For example, in some applications, a covariate of interest could be "precipitation", for which it is possible to use historical data as well as reasonably accurate predictions for the near future, which may be obtained from the weather service. Traditional methods to incorporate future covariates have major limitations. The strategy of simply concatenating the future covariates to the input vector is highly likely to miss the past-future connection. Another strategy that iteratively predicts one step at a time can end up with prediction error accumulation. We propose two novel feature representation strategies to solve those limitations – *shifting* and *padding*, which create a framework for contextually linking the past with the predicted future, while avoiding any accumulation of prediction errors. Extensive experiments on three well-known datasets revealed that our strategies when applied to RNN and CNN backbones, outperform existing traditional methods. Our experiments also suggest a relationship between the amount of shifting and padding and the periodicity of the time series.