Short Communication Industrial Engineering & Management 2021
Vol.10 No.1

Surge-Adjusted Forecasting in Temporal Data Containing Extreme Observations - Smaranya Dey, Walmart Labs, Bangalore, India

Smaranya Dey, Subhadip Paul, Uddipto Dutta, Anirban Chatterjee
[1] Walmart Labs, Bangalore, India

Forecasting in time-series data is at the core of various business decision making activities. One key characteristic of many practical time series data of different business metrics such as orders, revenue, is the presence of irregular yet moderately frequent spikes of very high intensity, called extreme observation. Forecasting such spikes accurately is crucial for various business activities such as workforce planning, financial planning, inventory planning. Traditional time series forecasting methods such as ARIMA, BSTS, are not very accurate in forecasting extreme spikes. Deep Learning techniques such as variants of LSTM tend to perform only marginally better than these traditional techniques. The underlying assumption of thin tail of data distribution is one of the primary reasons for such models to falter on forecasting extreme spikes as moderately frequent extreme spikes result in heavy tail of the distribution. On the other hand, literatures, proposing methods to forecast extreme events in time series, focused mostly on extreme events but ignored overall forecasting accuracy. We attempted to address both these problems by proposing a technique where we considered a time series signal with extreme spikes as the superposition of two independent signals - (1) a stationary time series signal without extreme spike (2) a shock signal consisting of near-zero values most of the time along with few spikes of high intensity. We modelled the above two signals independently to forecast values for the original time series signal. Experimental results show that the proposed technique outperforms existing techniques in forecasting both normal and extreme events. A tempest flood hindcast for the west shoreline of Canada was produced for the time frame 1980–2016 utilizing a 2D nonlinear barotropic Princeton Ocean Model constrained by hourly Climate Forecast System Reanalysis wind and ocean level pressing factor. Approval of the displayed storm floods utilizing tide measure records has shown that there are broad zones of the British Columbia coast where the model doesn't catch the cycles that decide the ocean level fluctuation on intraseasonal and interannual time scales. A portion of the inconsistencies are connected to enormous scope variances, for example, those emerging from significant El Niño and La Niña occasions. By applying a change in accordance with the hindcast utilizing a sea reanalysis item that consolidates enormous scope ocean level changeability and steric impacts, the difference of the mistake of the changed floods is fundamentally decreased (by up to half) contrasted with that of floods from the barotropic model. The significance of baroclinic elements and steric impacts to exact tempest flood anticipating in this waterfront area is illustrated, just like the need to consolidate decadal-scale, bowl explicit maritime inconstancy into the assessment of outrageous beach front ocean levels. The outcomes improve long haul extraordinary water level gauges and stipends for the west shoreline of Canada without long haul tide measure records information.

.

This work is partly presented at 7th International Conference on Big Data Analysis and Data Mining, July 17-18, 2020 Vienna, Austria