HiLARIS
Journal of Computer Science & Systems Biology
ISSN: 0974-7230
2020
Vol.13 No.6

Surge-Adjusted Forecasting in Temporal Data Containing Extreme Observations

Smaranya Dey, Subhadip Paul, Uddipto Dutta, Anirban Chatterjee

Walmart Labs, Bangalore, India

Abstract

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.

Biography:

Smaranya Dey is a Data Scientist in Walmart Labs, India. She is interested in researching areas of forecasting analysis and natural language processing. Anirban Chatterjee is Staff Data Scientist in Walmart Labs, India. His research interest lies in Time Series Analysis and Modeling with High Dimensional Data.

Speaker Publications:

1. Chelton, D. B., and R. E.Davis, 1982: Monthly mean sea-level variability along the west coast of North America. J. Phys. Oceanogr., 12, 757–784, https://doi.org/10.1175/1520-0485(1982)012<0757:MMSLVA>2.0.CO;2. Google ScholarCrossref

2. Chen, X., X.Zhang, J. A.Church, C. S.Watson, M. A.King, D.Monselesan, B.Legresy, and C.Harig, 2017: The increasing rate of global mean sea-level rise during 1993-2014. Nat. Climate Change, 7, 492–495, https://doi.org/10.1038/nclimate3325. Google ScholarCrossref

3. Church, J. A., and Coauthors, 2013: Sea level change. Climate Change 2013: The Physical Science Basis, T. F. Stocker et al., Eds., Cambridge University Press, 1137–1216.

7th International Conference on Big Data Analysis and Data Mining - July 17-18, 2020 Webinar.

Abstract Citation:

Smaranya Dey and Anirban Chatterjee, Surge-Adjusted Forecasting in Temporal Data Containing Extreme Observations, Data Mining 2020, 7th International Conference on Big Data Analysis and Data Mining – July 17-18, 2020 Webinar.

(https://datamining.expertconferences.org/speaker/2020/smaranya-dey-walmart-labs-bangalore-india)