Yaxin Li

Professor Gu

Advanced Big Data Analysis

March 1st 2020

Summary on Curriculum Learning in Deep Neural Networks for Financial Forecasting

Document Click Here

Following recent advances in applying deep neural networks (DNNs) in the time series domain, this paper explores applying Long Short-Term Memory (LSTM) and Dilated Convolutional Neural Network (DCNN) models to hierarchical financial time series.

The authors' goal is to forecast the revenue for Microsoft products. They present five neural network models and use the Mean Absolute Percentage Error (MAPE) as error evaluation metric. Those five models are LSTM with Categorical Indicators, LSTM with Seasonality, LSTM with Curriculum Learning, Dilated CNN with Categorical Indicators and Dilated CNN with Curriculum Learning. They find that both LSTM and DCNN models with curriculum learning out- perform the respective models without curriculum learning. By comparing against Microsoft's production baseline accuracy, authors find that their curriculum learning method can be successfully applied to various neural networks on time series data to achieve higher accuracy and positive results in bias and variance.

The authors suggest that seasonal trend decomposition be used only after careful consideration of the durability of financial seasonality. In their application, they only present LSTM results including seasonality since they find it beneficial conjointly with curriculum learning; experimentally, they displayed results fare better than the alternative of curriculum learning sans seasonality.

Summary on Github Link

Github Link Click Here

This github file includes four notebooks. They use data from Kaggle, the Web Traffic Time Series Forecasting to Wikipedia pages. Kaggle Link Here

The first notebook is the introduction to seq2seq for time series forecasting. It walks through constructing an encoder/decoder model with LSTM, with architectures for both training and prediction.

The second notebook is the introduction to the WaveNet CNN model for time series forecasting. It walks through constructing a simple WaveNet-style CNN model and using it to generate predictions.
The third notebook is the full-fledged WaveNet CNN model for time series forecasting. It walks through constructing a WaveNet CNN model including all major components (dilated causal convolutions, gated activations, skip and residual connections) and using it to generate predictions.
The fourth notebook is about the full-fledged WaveNet CNN model for time series forecasting, incorporating exogenous features.