

Feel free to work with other students, but make sure you write up the homework and code on your own (no copying homework *or* code; no pair programming). Feel free to ask students or instructors for help debugging code or whatever else, though.

1 Find some financial dataset and a paper describing a machine learning technique then apply the technique to the financial data. Explain your methods and paste your output below.

Per Prof Gu's instructions in her email, I am summarizing a paper with machine learning methods.

I found this paper about predicting high-variance events in time series. It is not a financial paper, but I believe that the methods and algorithms discussed are highly applicable to algorithm trading. When we model an equity using a time series, we often will smooth out the data by removing the outliers, which means that we will have trouble predicting events that will cause a large deviation in variance. Therefore, this paper outlines a machine learning method called Long Short Term Memory (LSTM) along with a novel architecture to predict events that would normally fall outside the prediction range of normal time-series analysis. Here is a link to the paper: [paper](#)

The authors work at Uber, which uses time series forecasting to predict rideshare demand at any given time. However, their current forecasting methods are not great at accounting for localized random high variance events that will cause a dramatic increase in the demand for rides, such as protests, Black Friday, and events like those. Therefore, the authors proposed a novel approach to accurately forecast such events, using neural networks embedded in a novel implementation of LSTM.

The authors broke from the traditional approach to training feature extraction models, where normally, the features are derived manually. Instead, the authors used auto feature extraction, which they claim is an essential step to capture the complexity of dynamic events. The authors found that their new, multilayer LSTM model trained with raw data had a 14 percent improvement over industry-standard models.

This novel approach to modeling time series with machine learning is exciting because now, there can be a generic time series that can accurately model both normal events as well as extreme events. In algorithm trading, this can be applied to volatility trading. Because normal financial time series do not account for extreme events as well, such as the beginning of COVID-19 or the release of Fed minutes, these extreme events could render our algorithms unprofitable. Therefore, having a time series model that is adaptable to extreme events is extremely important, especially in times of market unrest like now.

Therefore, I will try to implement the models described in this paper as a new approach to time series data.

■