

Capstone Project 1 - Data Storytelling
Springboard Data Science Track
Kamran Ossia
2018-08-29

This project deals with the question of predicting daily stock prices using the Keras deep learning software with a Tensorflow backend and GPU acceleration. The prediction is based on past values of the same stock price, i.e. no other series such as volatility index, cumulative tick, volume, etc. are used. The stock data is downloaded from the IEX. The following questions are addressed in this report.

1. Can you count something interesting?
Counting quantities is not entirely relevant in time series prediction.
2. Can you find trends (e.g. high, low, increasing, decreasing, anomalies)?
Finding trends is an integral part of what is known as technical analysis in trading stocks and commodities based on price action. The most common trend indicators are moving averages, including the simple moving average, i.e. a fixed weighted sum of a past number of values and the exponential moving average where the weighting is an decaying exponential. These values are plotted on top of the price chart to indicate visually the trends at different time frames. Popular trends for daily charts are the 50 and 200 simple moving averages. See TSLA plot below.
3. Can you make a bar plot or a histogram?
A common histogram for stock prices is the volume profile, a histogram plotted horizontally on the price chart to show how often each price level was visited during the time interval of the chart. This histogram is used for technical analysis which is beyond the scope of this report.

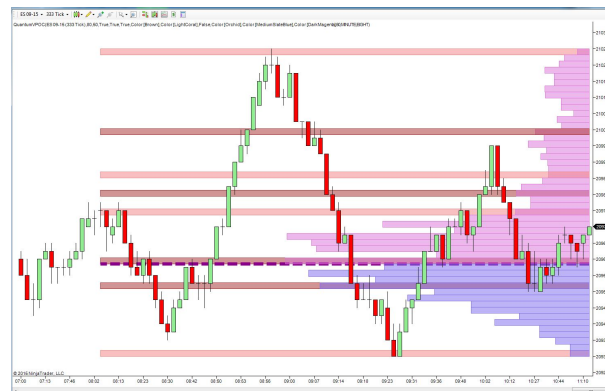
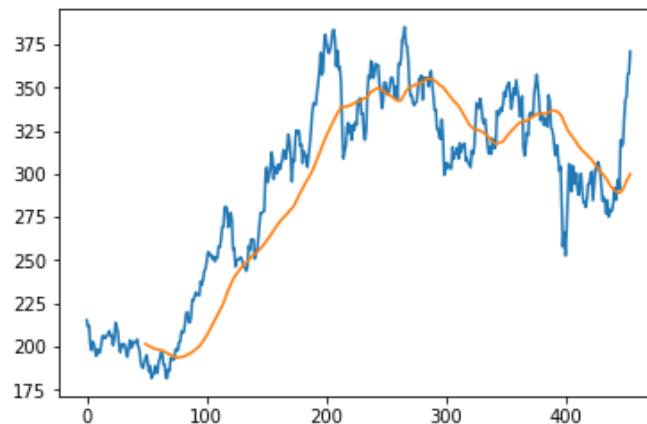


Image from <https://ninjaindicator.quantumtrading.com/>

4. Can you compare two related quantities?
The price action of a stock is influenced by several factors, one of which is the overall stock market trend. Therefore traders often plot the S&P index along with the stock to take advantage of discrepancies or divergences between the two.
5. Can you make a scatterplot?
Scatterplots are not meaningful for our purpose, see next question.

6. Can you make a time-series plot?

The most common for stock prices are the values of open, high, low, close, and volume for the relevant time frame, in our case daily. The following is a daily closing price chart of Tesla (ticker symbol TSLA) for August 2016 to August 2018 overlaid with a 50-day simple moving average.



The following is a daily candlestick chart of Google for February to April 2004. A candlestick chart displays the open and close values as a rectangle, with the high and low shown as 'wicks' above and below. The color of the rectangle indicates whether the close is higher or lower than the open, in our case black and red respectively. The chart below includes empty spots corresponding to weekends when the main exchanges are closed.

The candlestick chart is part of the matplotlib.finance package.



7. Looking at the plots, what are some insights you can make? Do you see any correlations? Is there a hypothesis you'd like to investigate further? What other questions do the insights lead you to ask?

The main conclusion from observing stock charts is that they are nonlinear time series and trying to model them with ordinary difference equations is not going to be successful. The idea of using neural networks to model stock prices is in a way trying to model human reactions to market forces. Fortunately, historical data for stock and commodity prices is freely available for decades, providing plenty of opportunities to implement and test neural network forecasting methods. The question which naturally arises out of this insight is, since trading is a simple game with reward and punishment for human actions, can we implement reinforcement learning methods similar to those used in chess and video games?