

Columbia University

IEOR4742 – Deep Learning for OR & FE

Project Summary (Group 14)

Authors contributing to this work:

Abhishek Mukherjee, Xueru Wang, Sie Hendrata

1. Project Objective

We want to use price and fundamentals information to predict future earnings surprise which is the difference between analyst forecast and real EPS. There are two tasks we plan to do:

- 1) Predict an earnings surprise above certain percentage thresholds;
- 2) Predict the level of EPS on a quarterly basis;

Finally, we can use the above information to long/short the specific stocks, thus setting up a portfolio in the financial sector of S&P 500. And, also we can optimize the Portfolio using certain Portfolio optimization metric like Sharpe Ratio or Sortino Ratio.

2. Previous groups' work

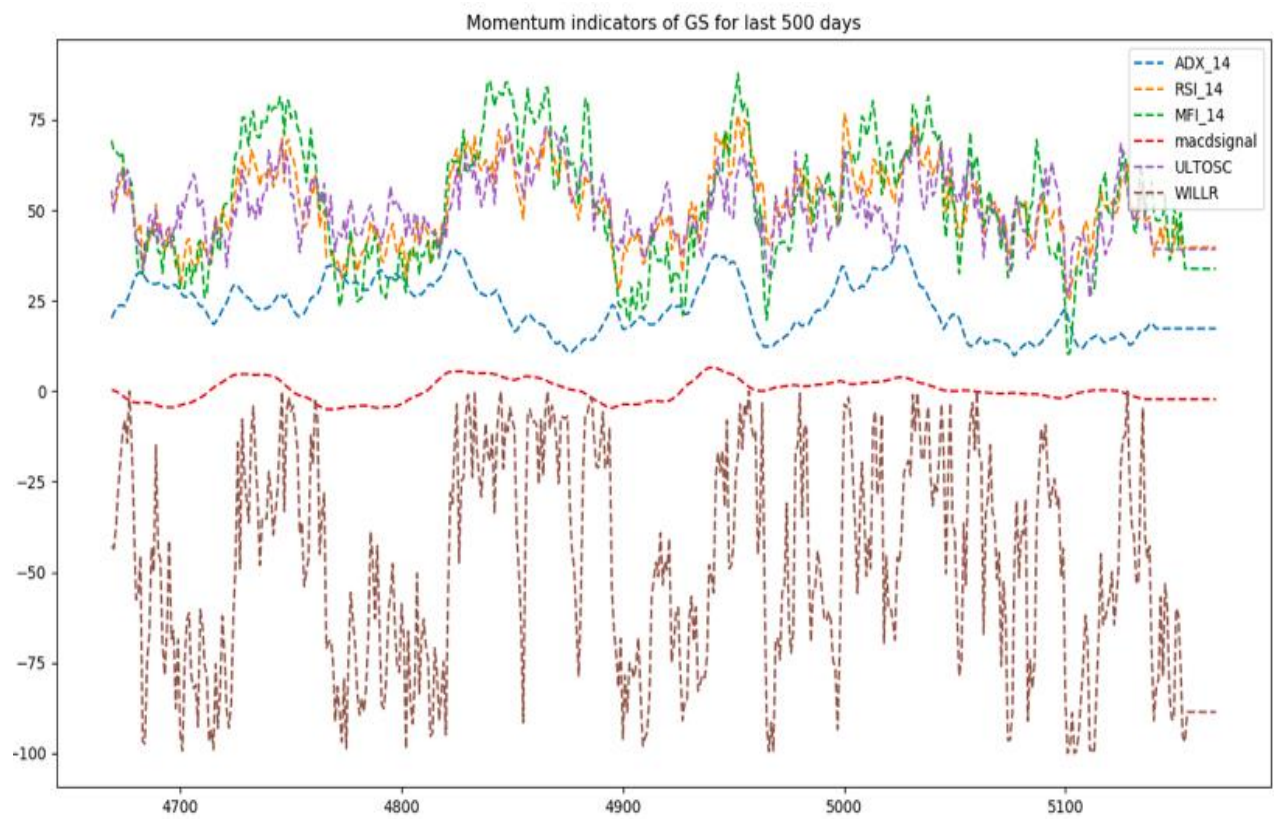
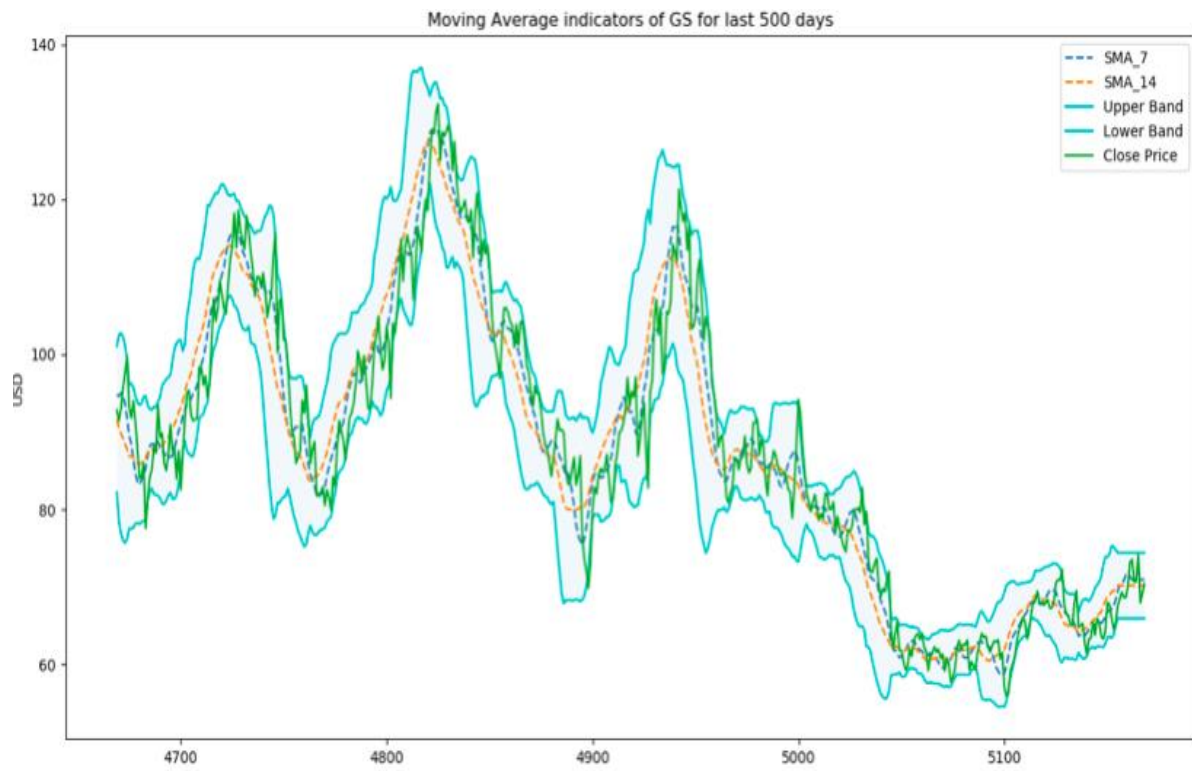
None

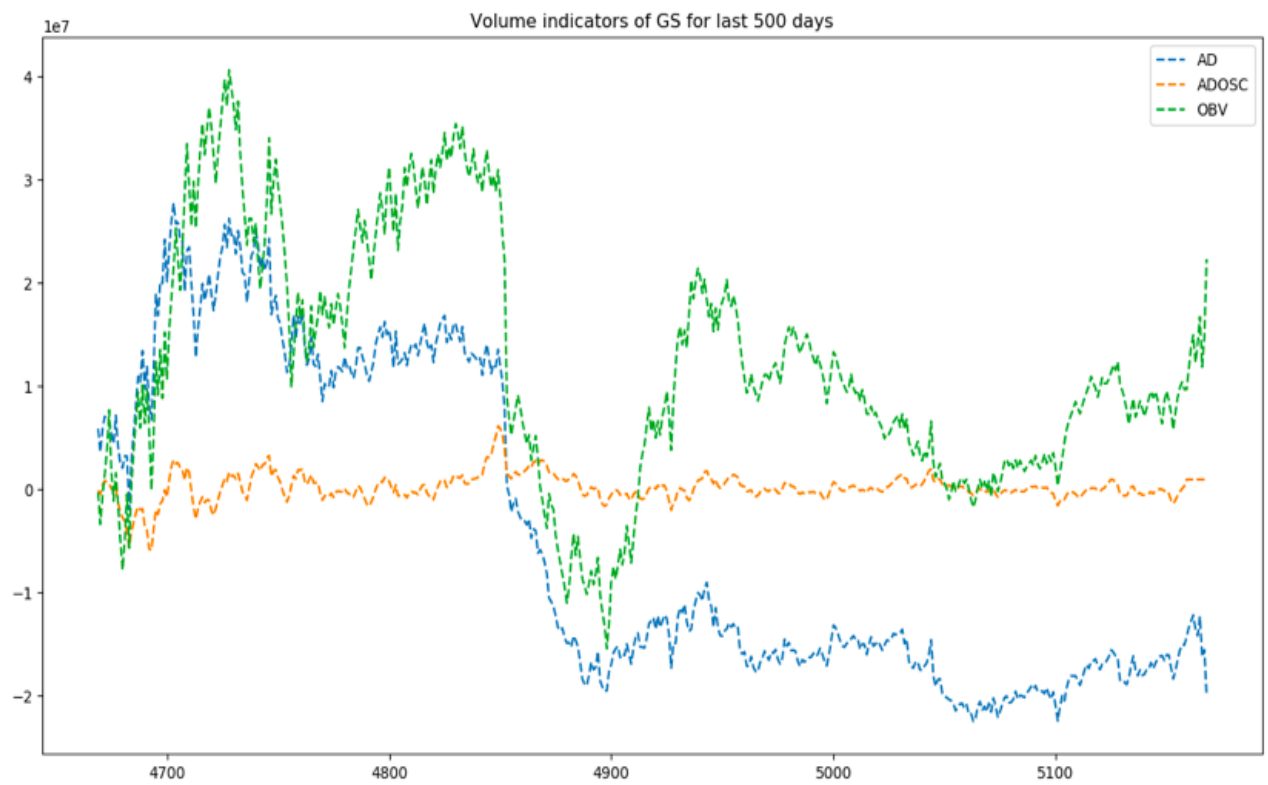
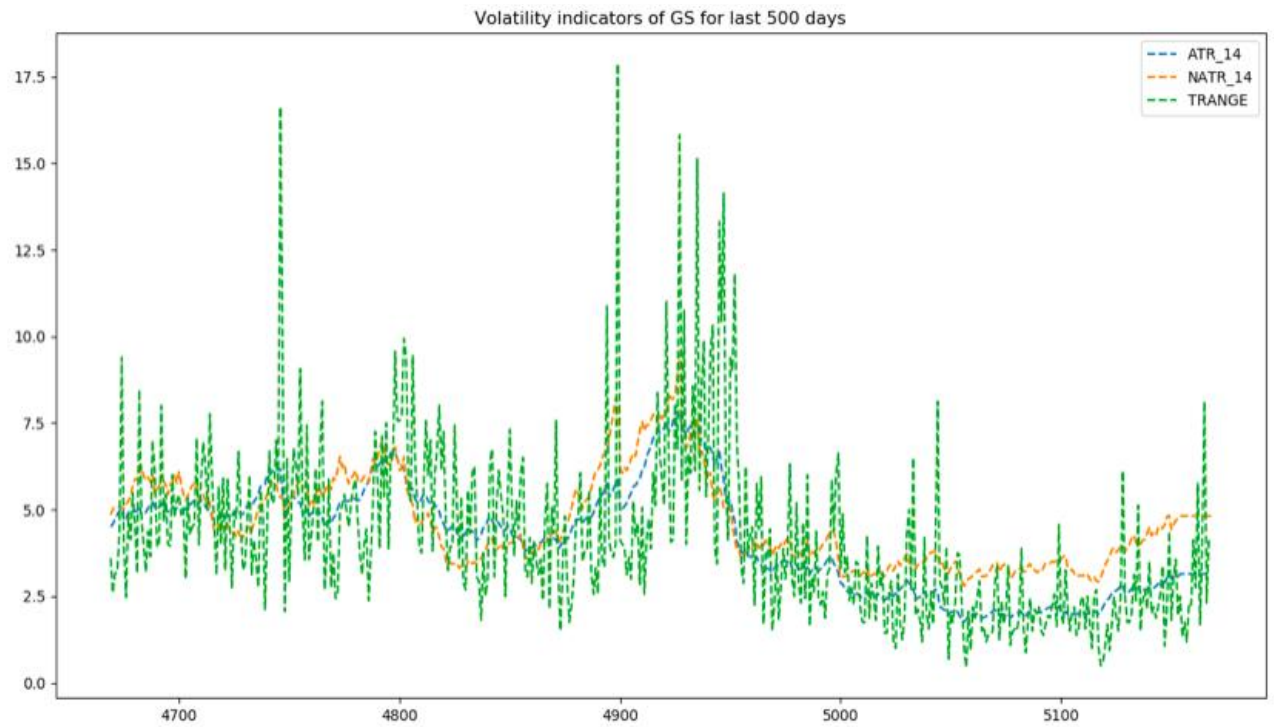
3. What we have done

We have collected daily price data, quarterly fundamental data, and quarterly analyst forecast data of S&P 500 from FactSet. Data collected from 1997 to 2019.

Up to now, we choose GS (Goldman Sachs) as a single stock example and what we have done are as follows:

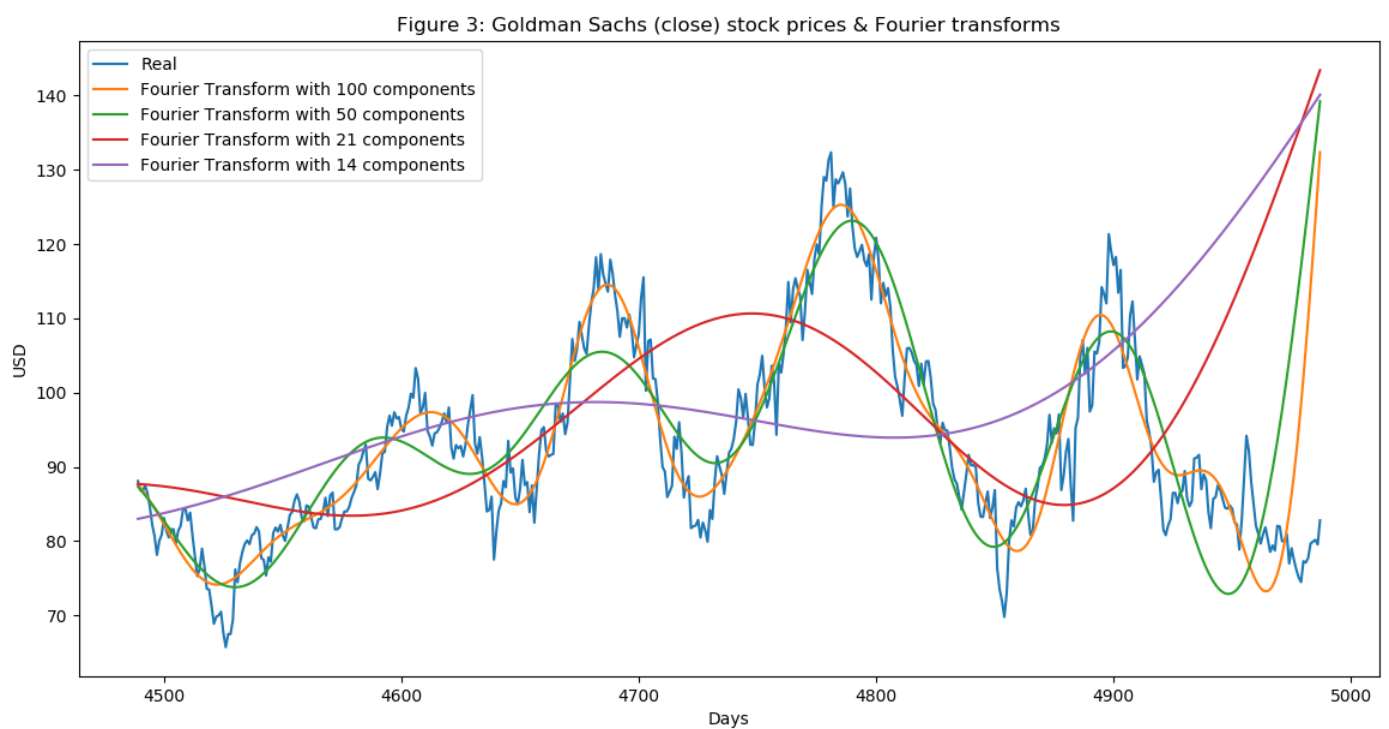
- 1) **Technical Indicators:** we created 178 price Technical Indicators of GS for last 500 days, including Moving Average indicators, Momentum indicators, Volatility indicators, Volume indicators and MACD, the plots are drawn below:



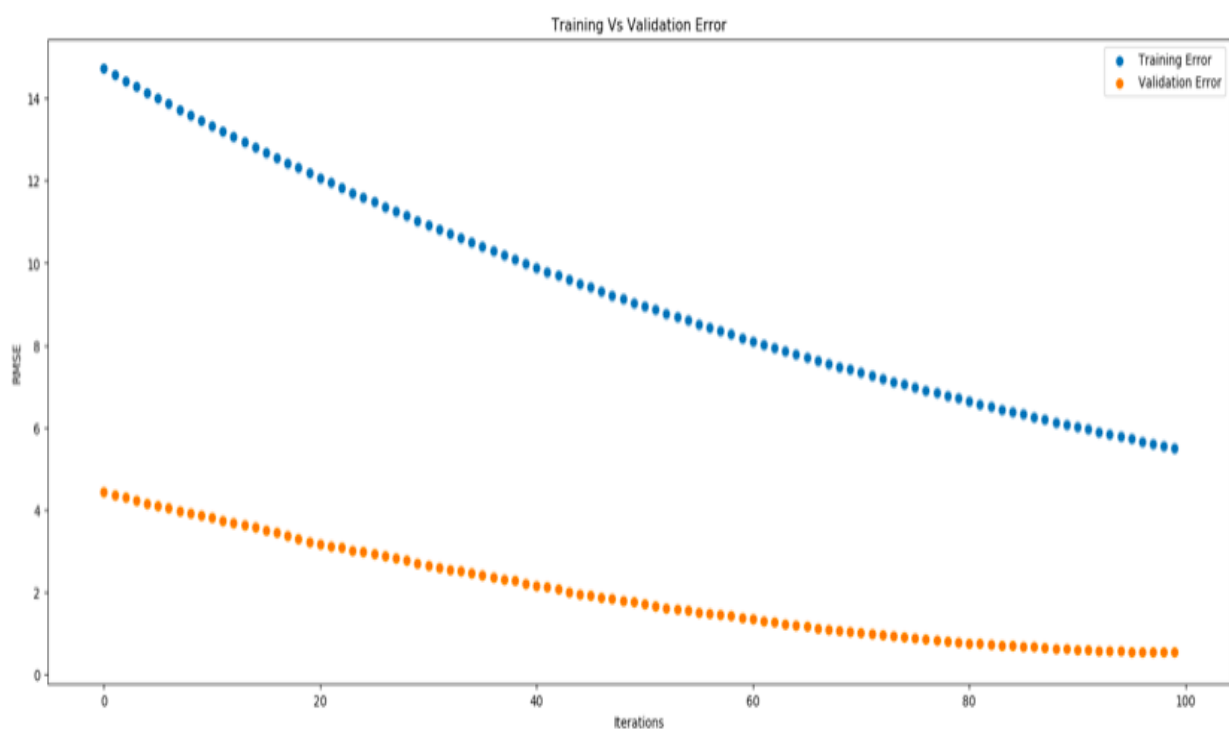
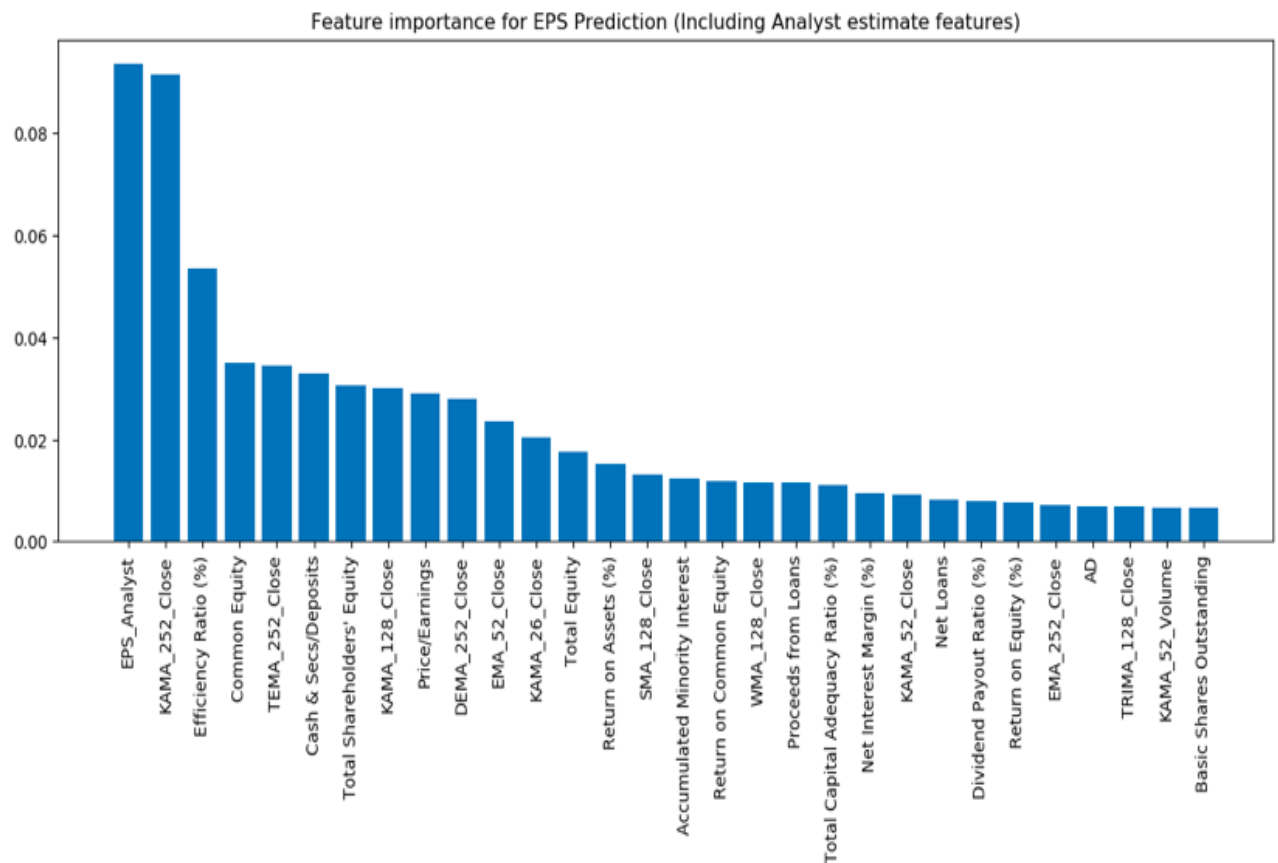


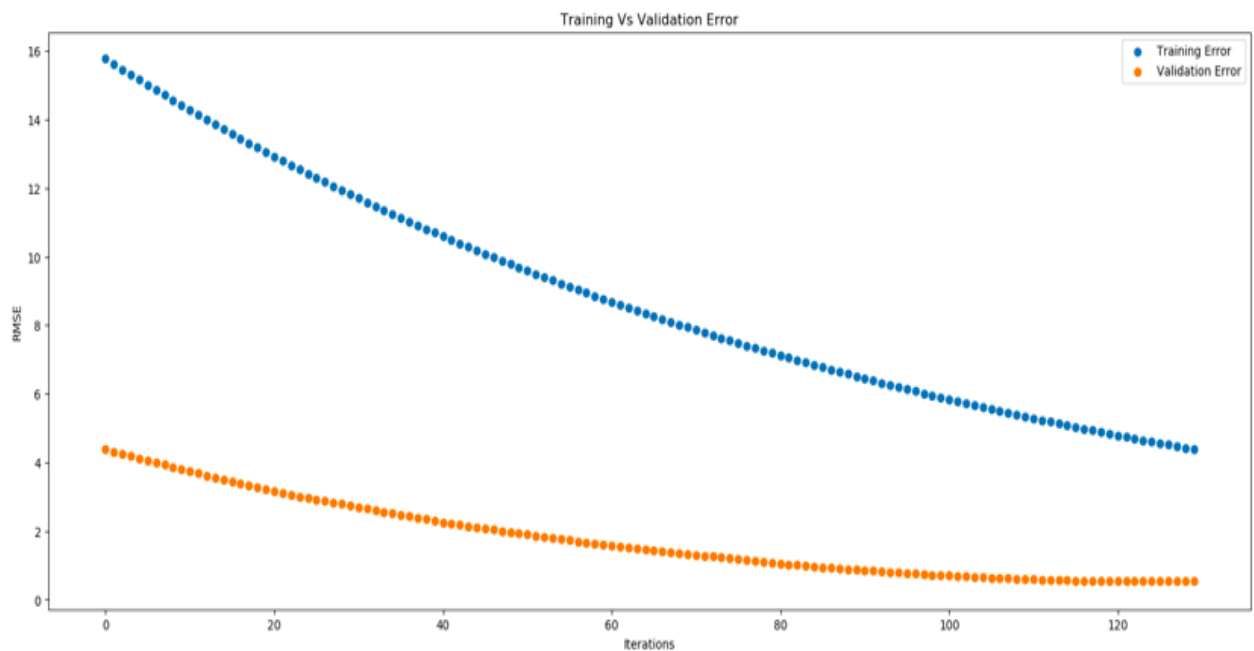
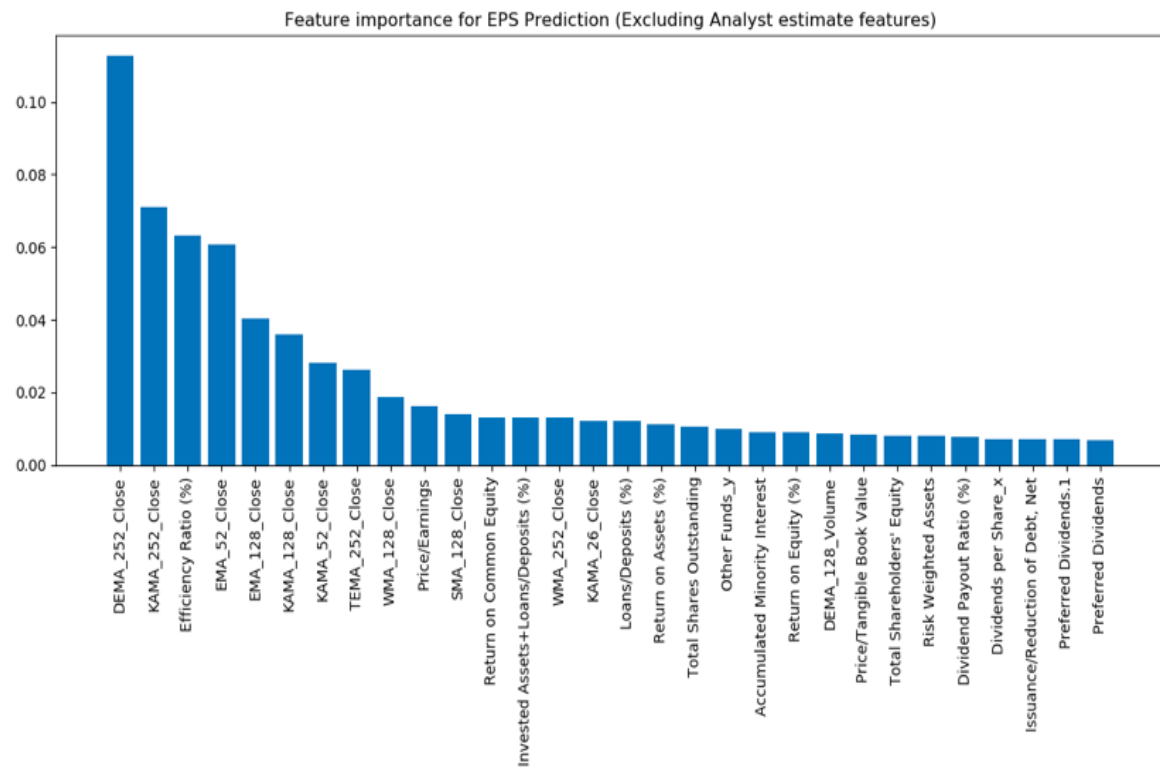


- 2) **Fourier transforms:** along with the daily closing price, we created Fourier transforms in order to generalize several long- and short-term trends. Using these transforms we will eliminate a lot of noise (random walks) and create approximations of the real stock movement, the transforms look like this:



- 3) **Feature importance with XGBoost:** having so many features we have to test feature importance for predicting EPS. Here are our results for including analyst estimate features and excluding analyst features, both of them don't have overfitting problem.





4. Next steps/expected outcome

- 1) Try to set up a classification model that can predict a earning surprise better than a naïve benchmark.
- 2) Use LSTM to predict EPS on quarterly basis.
- 3) Apply the above methods to other stocks.
- 4) Create a portfolio in finaiical section.
- 5) Use GAN to maximize the sharpe of the portfolio.