

by: Kyle Powers



LSTM Stock Prediction Model

Project: Predict Tomorrow's Stock Price

- Created 3 models for 28 stocks using exogenous data sources that I thought could have an impact on broader capital markets
 - 1) LSTM Model Prediction w/ exogenous variables and feature engineering
 - 2) LSTM Model Prediction using only 1,3,5,21,35 day return variables
 - 3) Facebook Prophet Prediction
- Sources: Quandl, Alpha-vantage, University of Michigan Surveys, data.gov, Google.com

Data and Features

Feature Engineering:

- Daily Stock data

20 Variations of each days
volume, high low, close including:

- Percent change
- 3,10,35 day moving average,
standard deviation and percent
change
- High-low difference

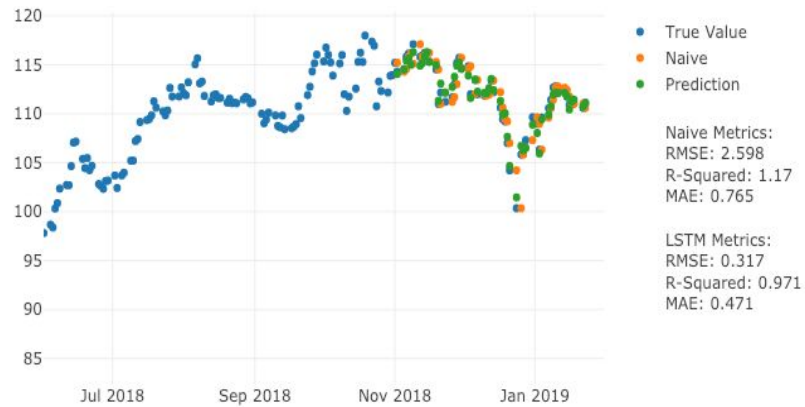
Exogenous Economic Data:

- Consumer Sentiment Report (Monthly)
- Consumer Market Confidence Report (Monthly)
- Change in Home Values During the Past Year (Monthly)
- Gold Prices (Daily)
- Ten-year Treasury yield (Daily)
- Consumer-Price-Index (Inflation) (Monthly)
- S&P 500 Composite Prices (Daily)
- Crude Oil Futures (Daily)
- Current Market Value of Primary Residence (Monthly)*
- Probability of Real Income Gains During the Next 5
Years (Monthly)*

DISNEY (DIS)

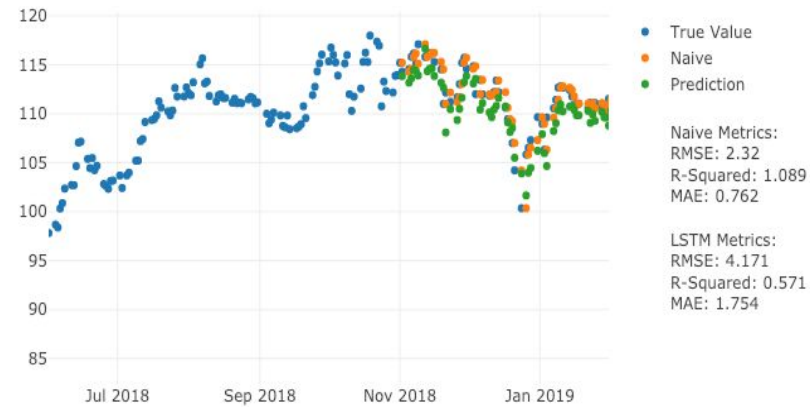
All Features

DIS Prices



Minimal Features

DIS Prices



Best model results:
DIS RMSE reduced
significantly with
Economic variables

Facebook Prophet:

r2_score:

0.95376558183

mean_squared_error:

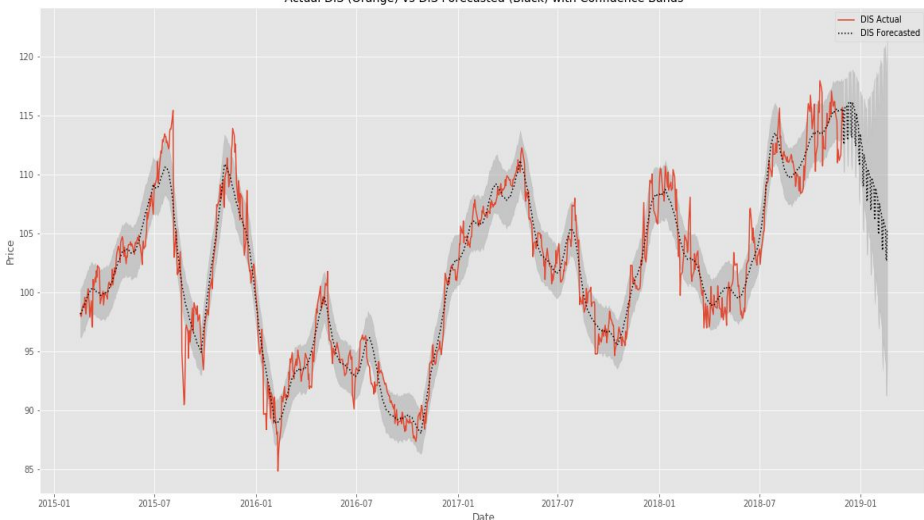
2.508954195

mean_absolute_error:

1.268912

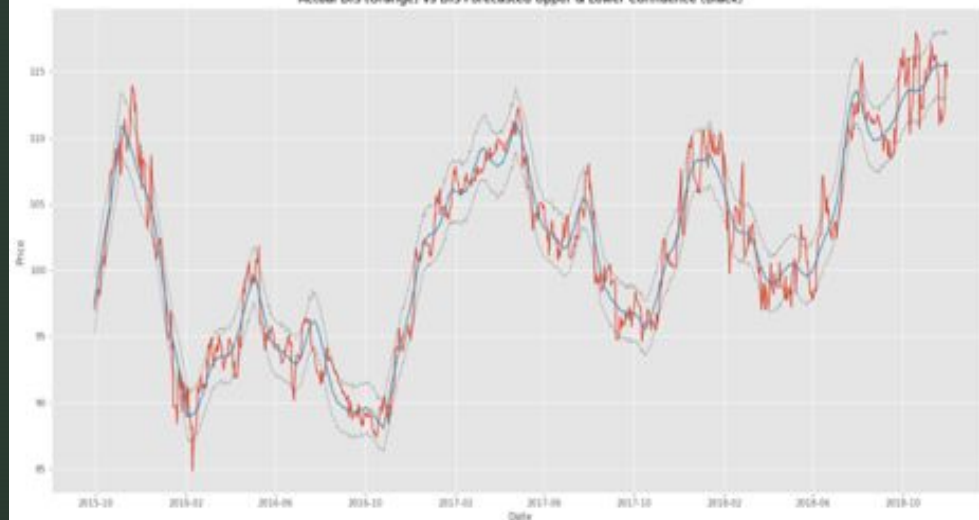
Facebook Prophet

Actual DIS (Orange) vs DIS Forecasted (Black) with Confidence Bands



Facebook Prophet

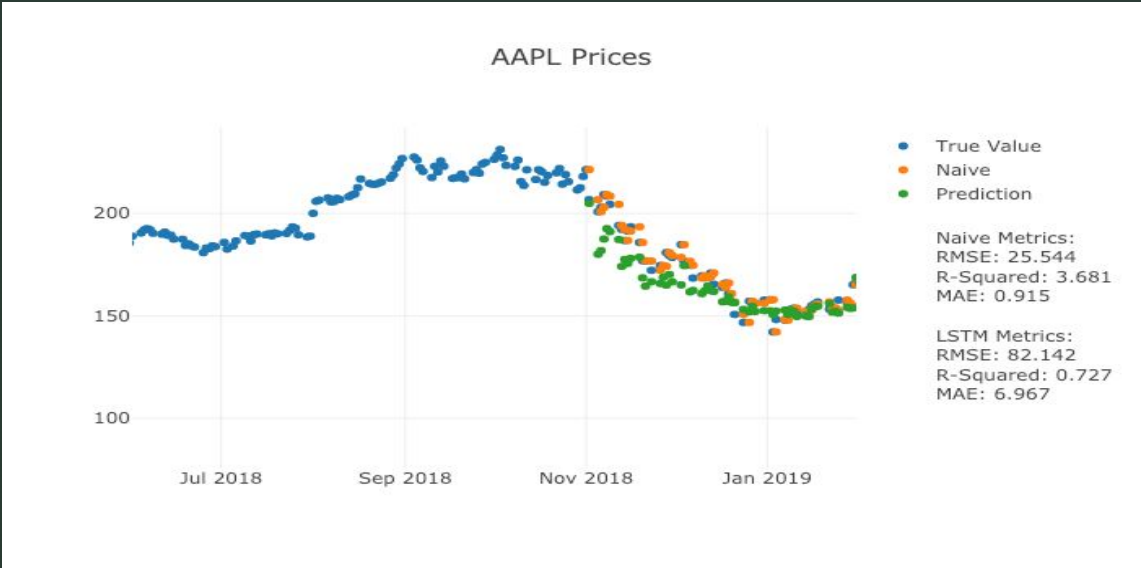
Actual DIS (Orange) vs DIS Forecasted Upper & Lower Confidence (Black)



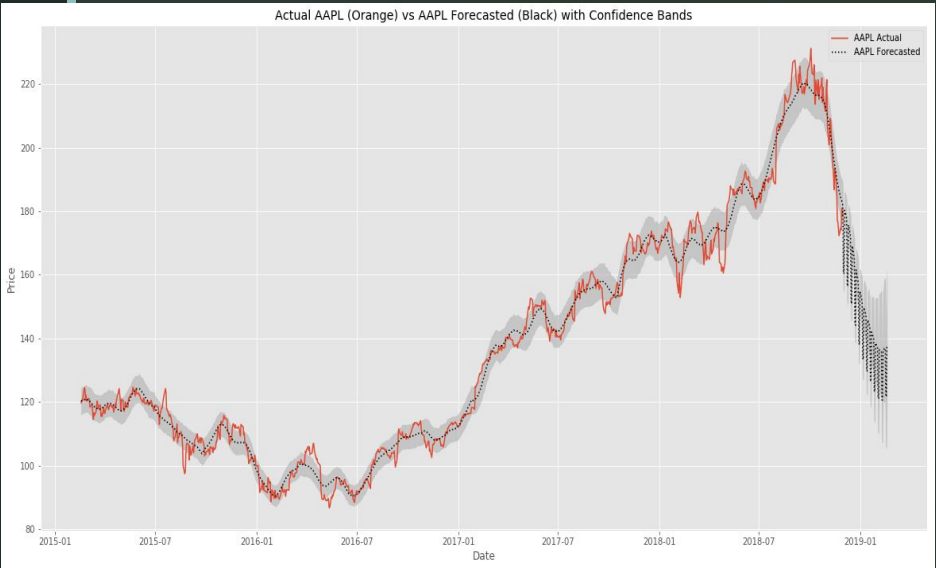
Apple AAPL) All Features



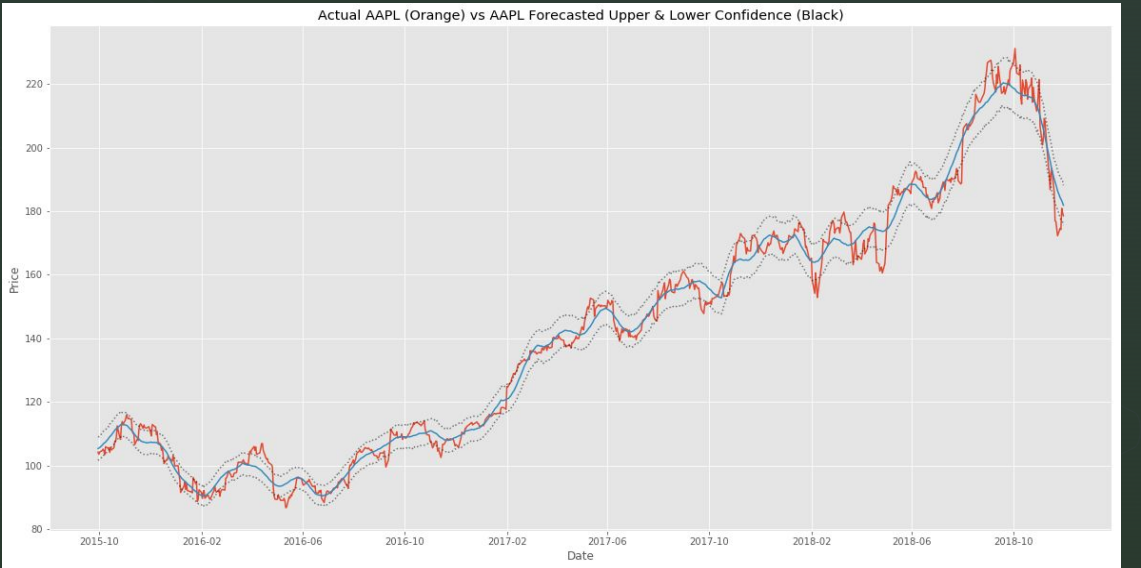
Minimal Features



Facebook Prophet



Facebook Prophet



All Features

WORSE MODEL FLOT:

Interest Rates increased in December and the yield curve inverted. Model expected prices to decline but the increase was already priced in by the market.

Minimal Features

Facebook Prediction

TIP Prices



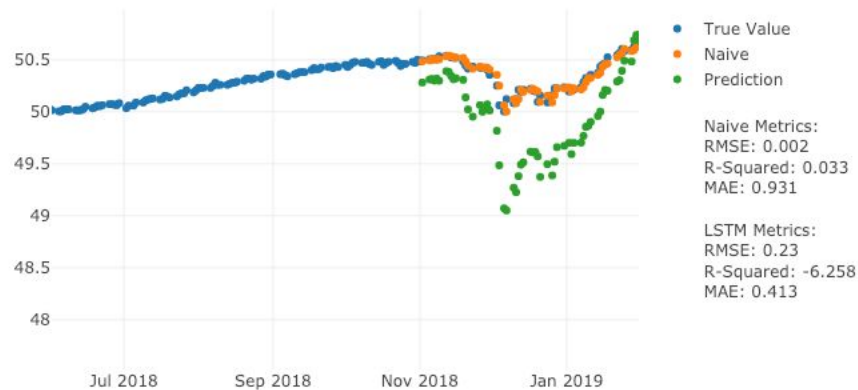
TIP Prices



FLOT Prices



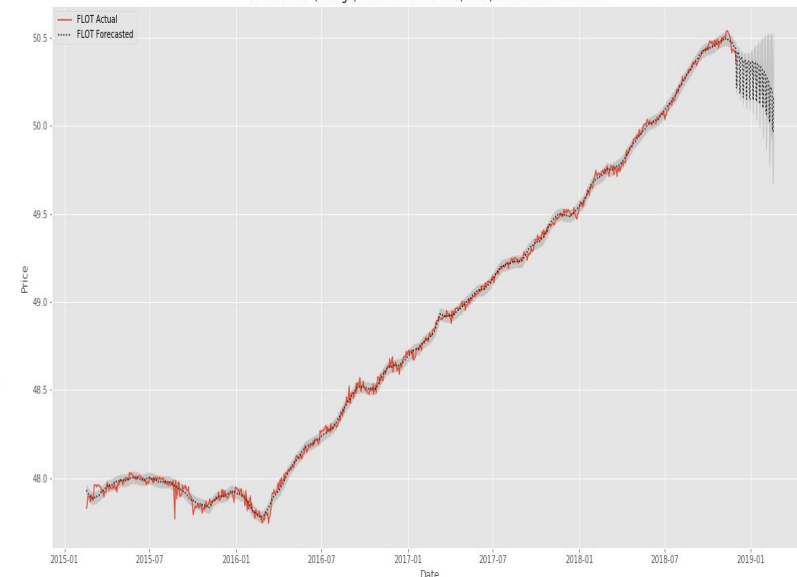
FLOT Prices

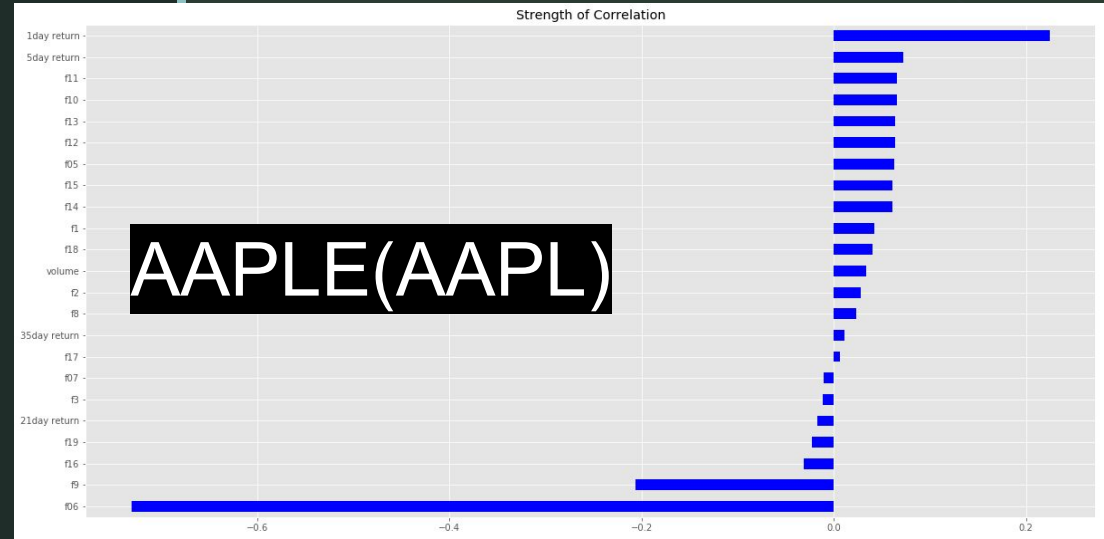
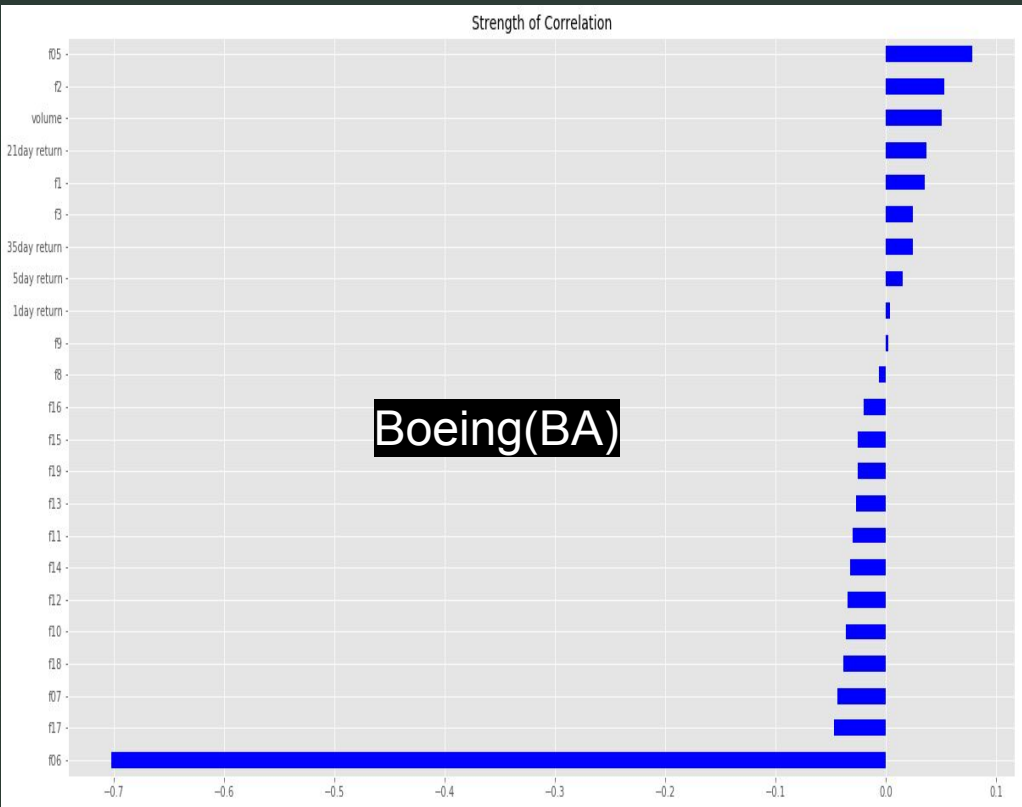
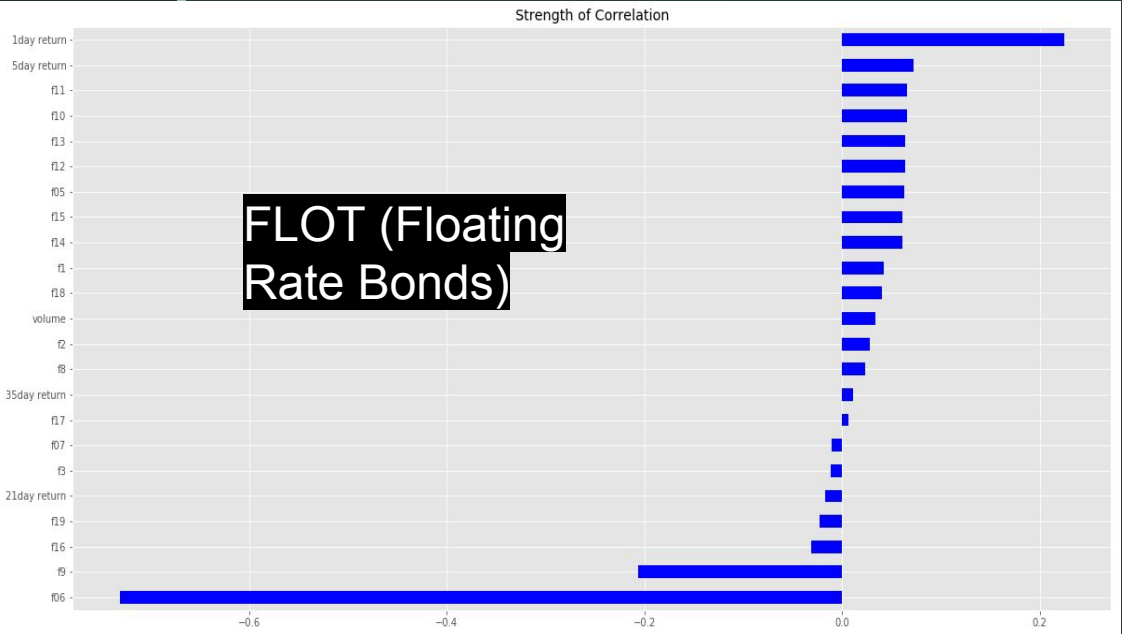


Actual TIP (Orange) vs TIP Forecasted (Black) with Confidence Bands



Actual FLOT (Orange) vs FLOT Forecasted (Black) with Confidence Bands





All Features

Minimal Features

Facebook Prediction

EFA Prices



EFA Prices



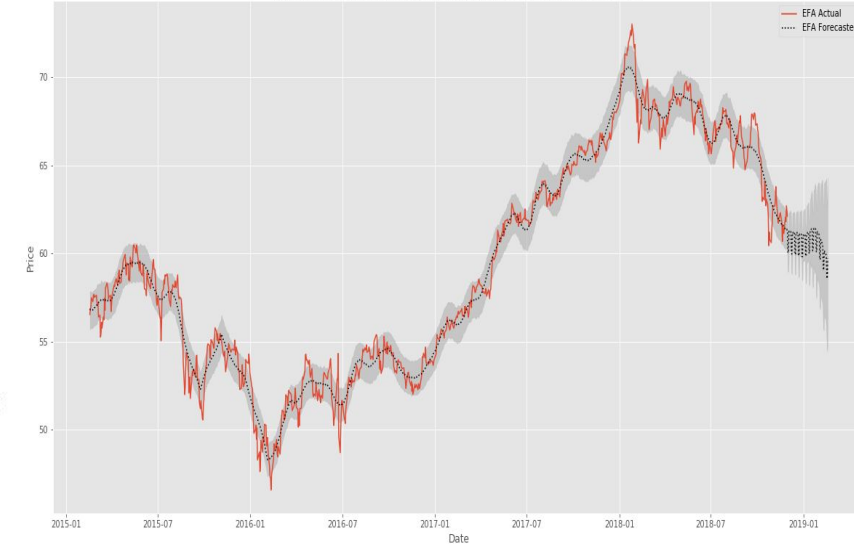
VEA Prices



VEA Prices



Actual EFA (Orange) vs EFA Forecasted (Black) with Confidence Bands

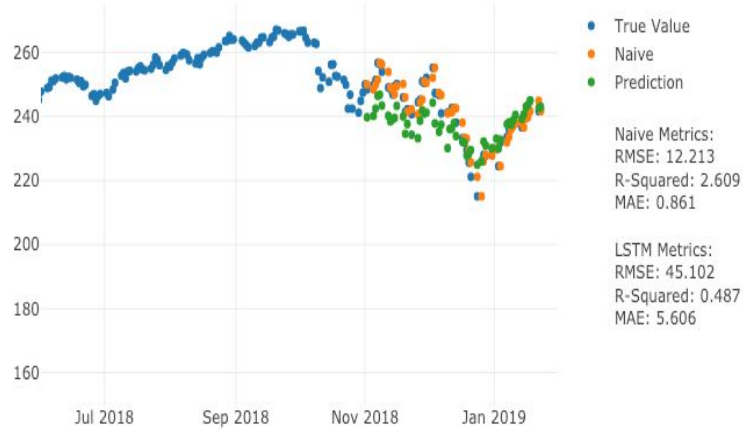


Actual VEA (Orange) vs VEA Forecasted (Black) with Confidence Bands



All Features

VOO Prices



QQQ Prices



Minimal Features

VOO Prices



QQQ Prices



Facebook Prediction

Actual VOO (Orange) vs VOO Forecasted (Black) with Confidence Bands



Actual QQQ (Orange) vs QQQ Forecasted (Black) with Confidence Bands

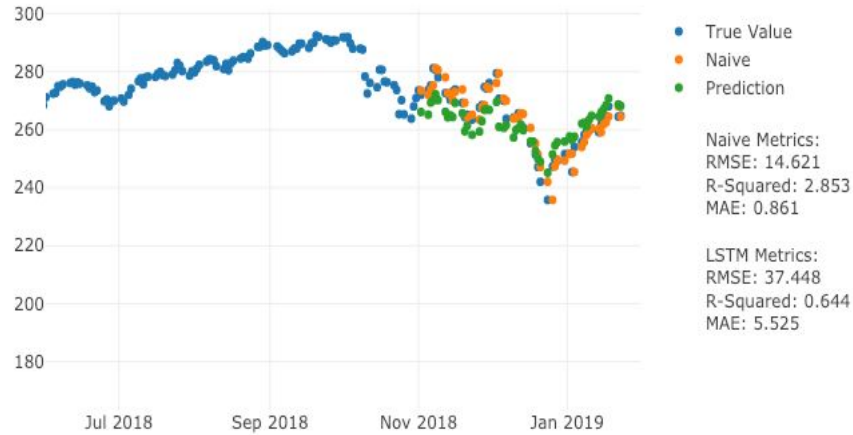


All Features

Minimal Features

Facebook Prediction

IVV Prices



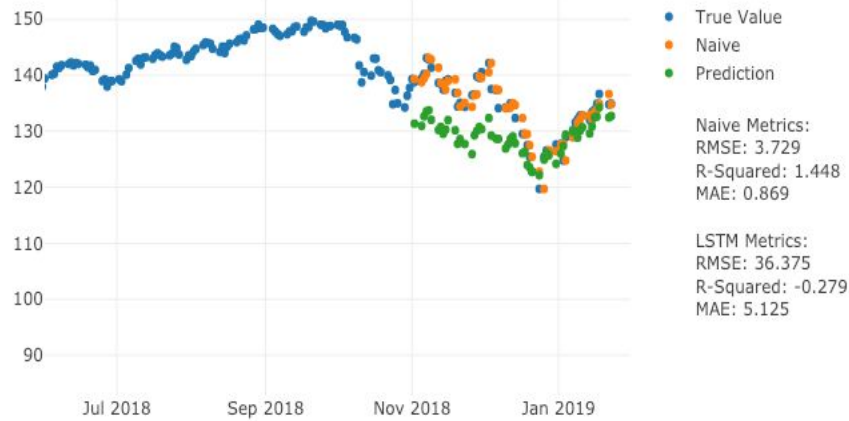
VTI Prices



Actual VTI (Orange) vs VTI Forecasted (Black) with Confidence Bands



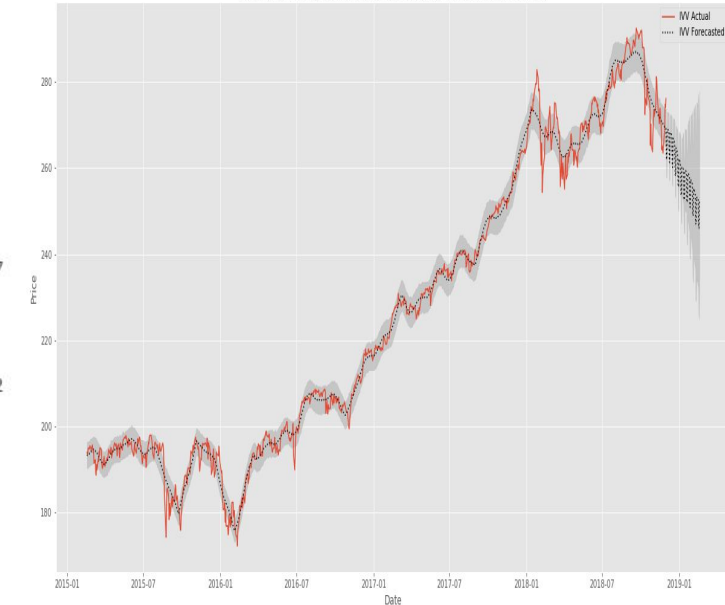
VTI Prices



IVV Prices

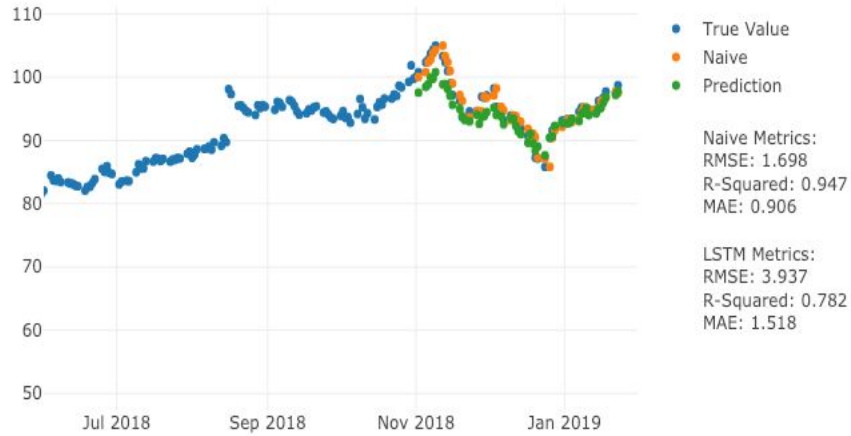


Actual IVV (Orange) vs IVV Forecasted (Black) with Confidence Bands



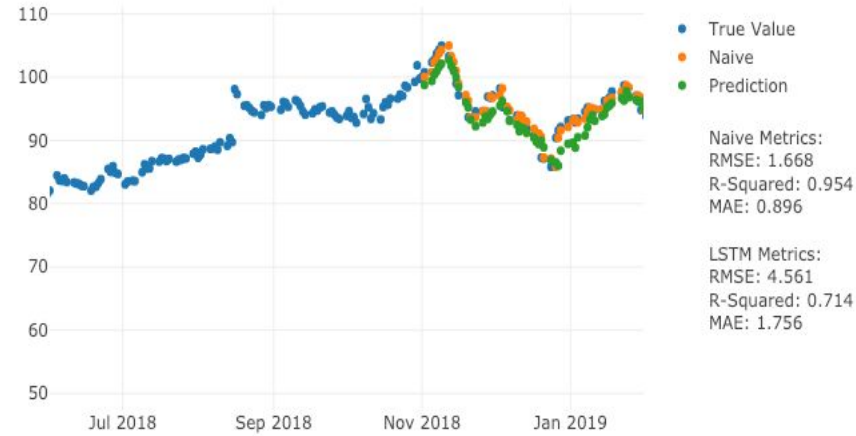
All Features

WMT Prices



Minimal Features

WMT Prices



Facebook Prediction

Actual WMT (Orange) vs WMT Forecasted (Black) with Confidence Bands



RSP Prices



RSP Prices



Actual RSP (Orange) vs RSP Forecasted (Black) with Confidence Bands



All Features

CAT Prices



CVX Prices



Minimal Features

CAT Prices



CVX Prices



Facebook Prediction

Actual CAT (Orange) vs CAT Forecasted (Black) with Confidence Bands



Actual CVX (Orange) vs CVX Forecasted (Black) with Confidence Bands



All Features



Minimal Features

Facebook Prediction

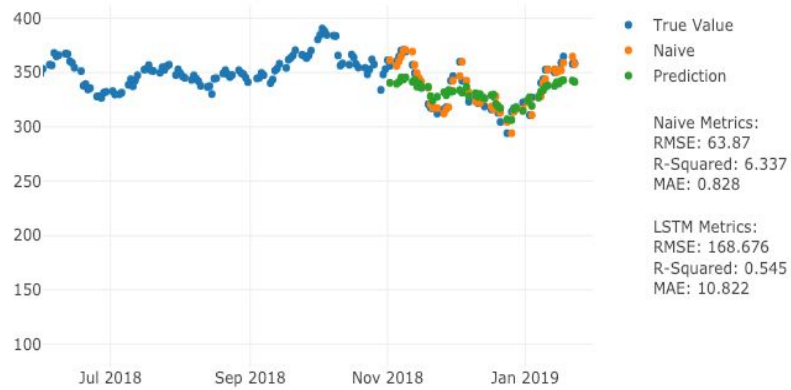
AXP Prices



AXP Prices



BA Prices



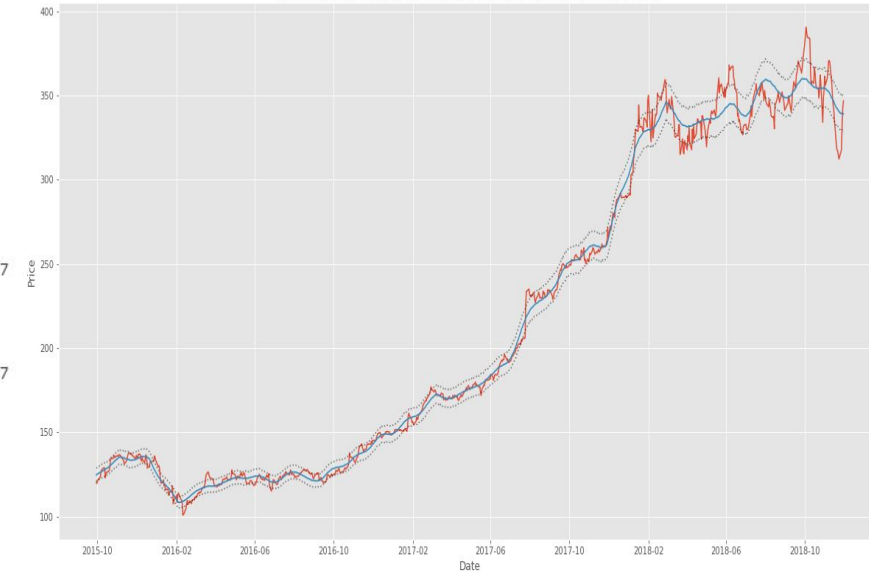
BA Prices



Actual AXP (Orange) vs AXP Forecasted (Black) with Confidence Bands



Actual BA (Orange) vs BA Forecasted Upper & Lower Confidence (Black)

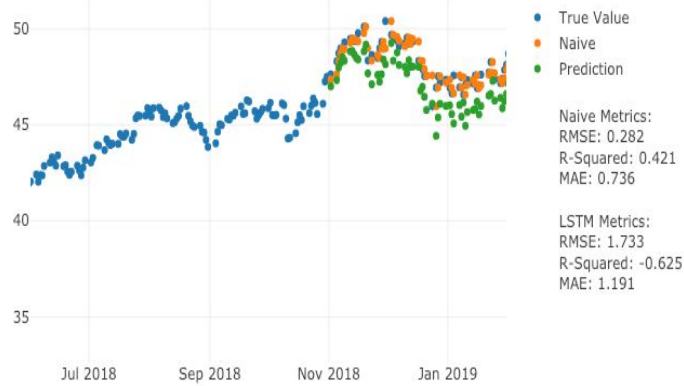


All Features

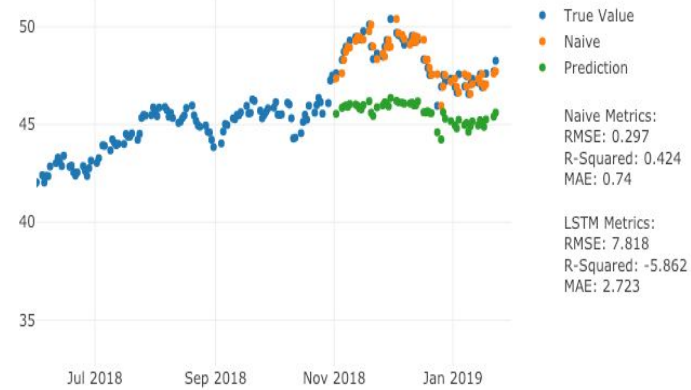
Minimal Features

Facebook Prediction

KO Prices



KO Prices



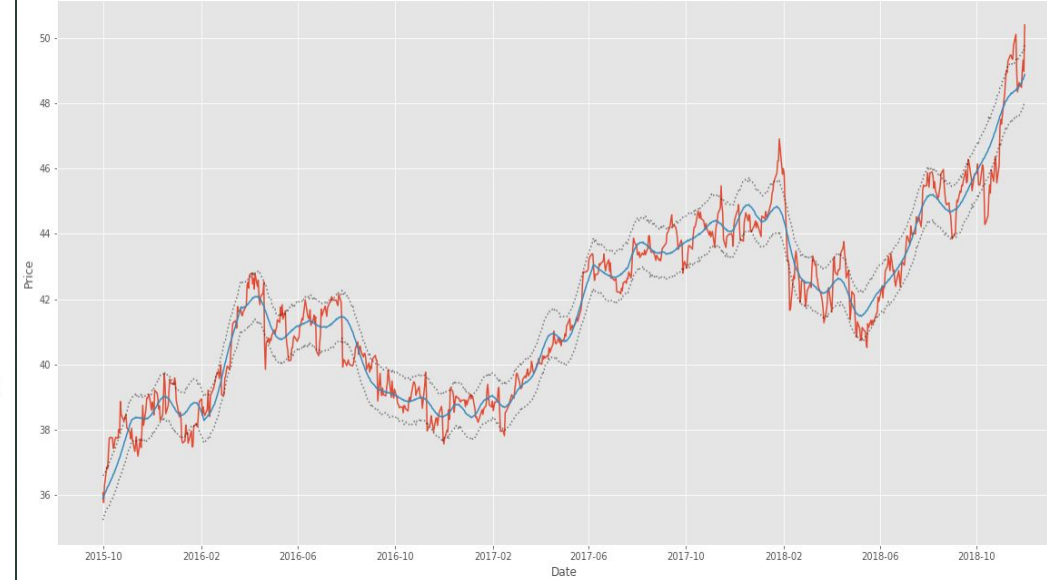
MMM Prices



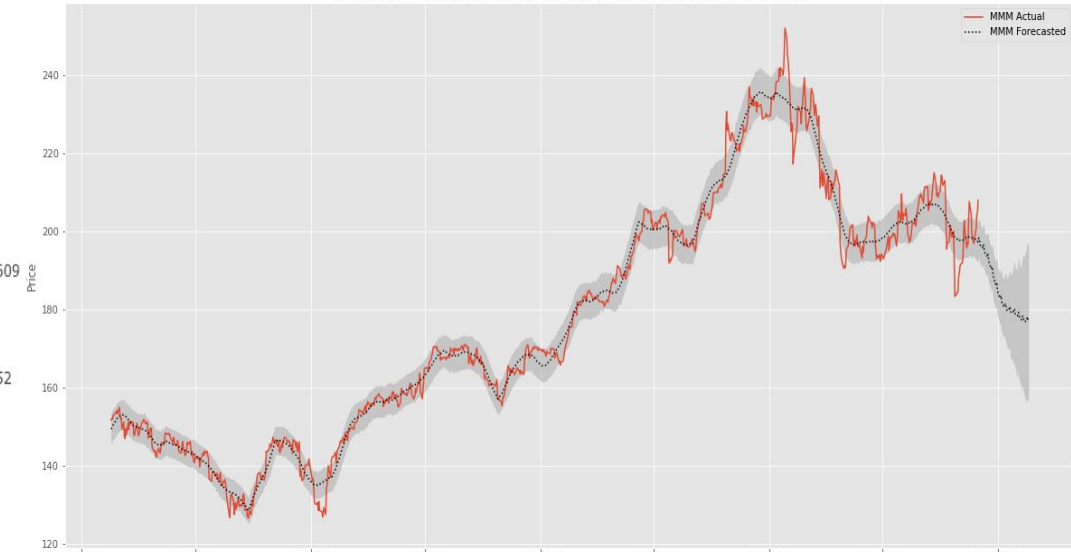
MMM Prices



Actual KO (Orange) vs KO Forecasted Upper & Lower Confidence (Black)



Actual MMM (Orange) vs MMM Forecasted (Black) with Confidence Bands



All Features

Minimal Features

Facebook Prediction

V Prices



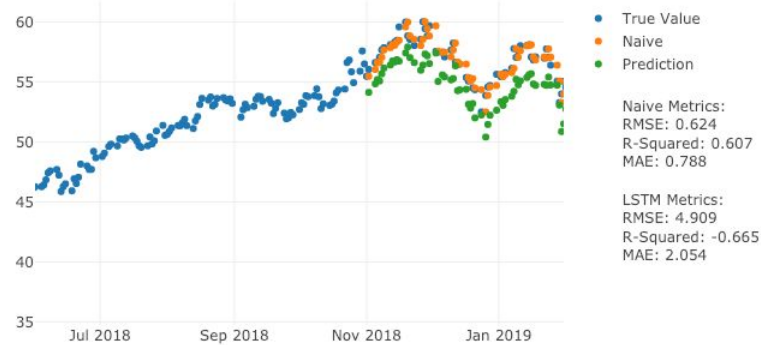
V Prices



VZ Prices



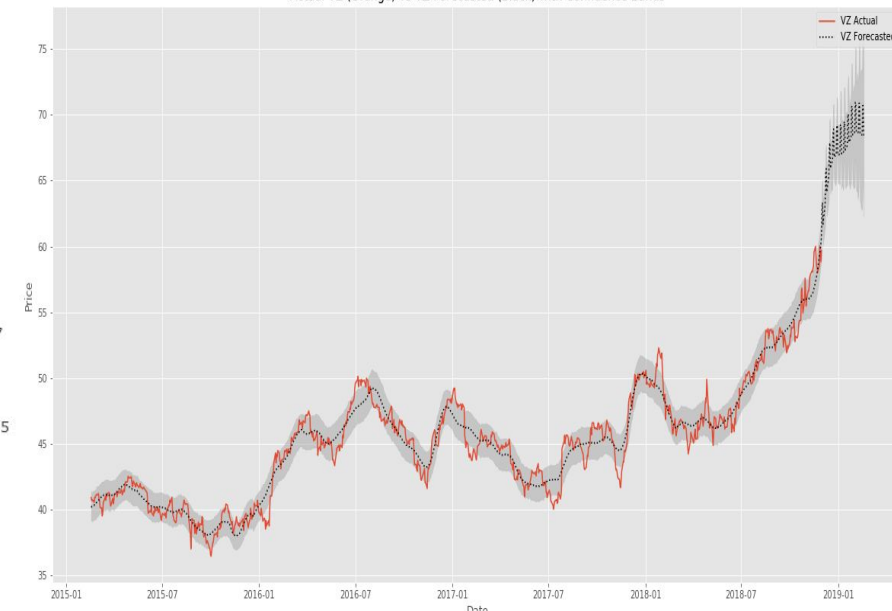
VZ Prices



Actual V (Orange) vs V Forecasted (Black) with Confidence Bands



Actual VZ (Orange) vs VZ Forecasted (Black) with Confidence Bands



All Features

GLD Prices



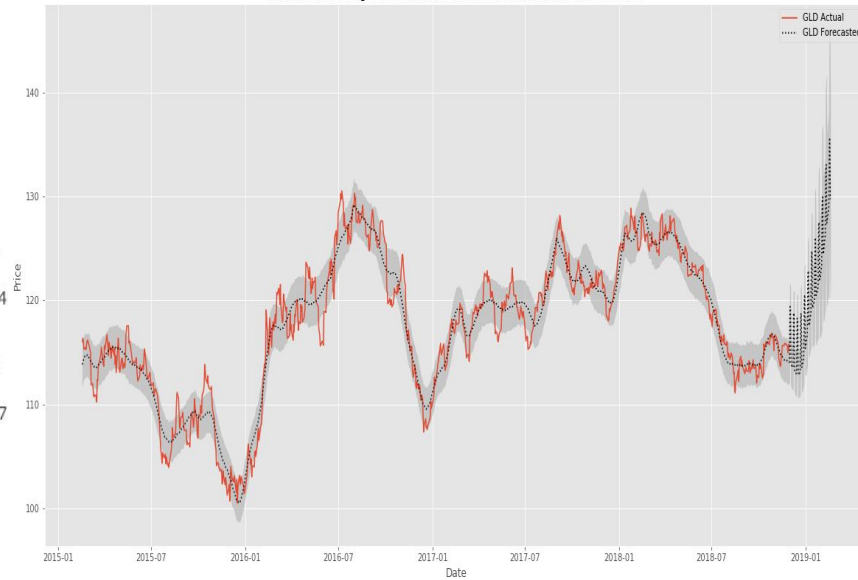
Minimal Features

GLD Prices



Facebook Prediction

Actual GLD (Orange) vs GLD Forecasted (Black) with Confidence Bands



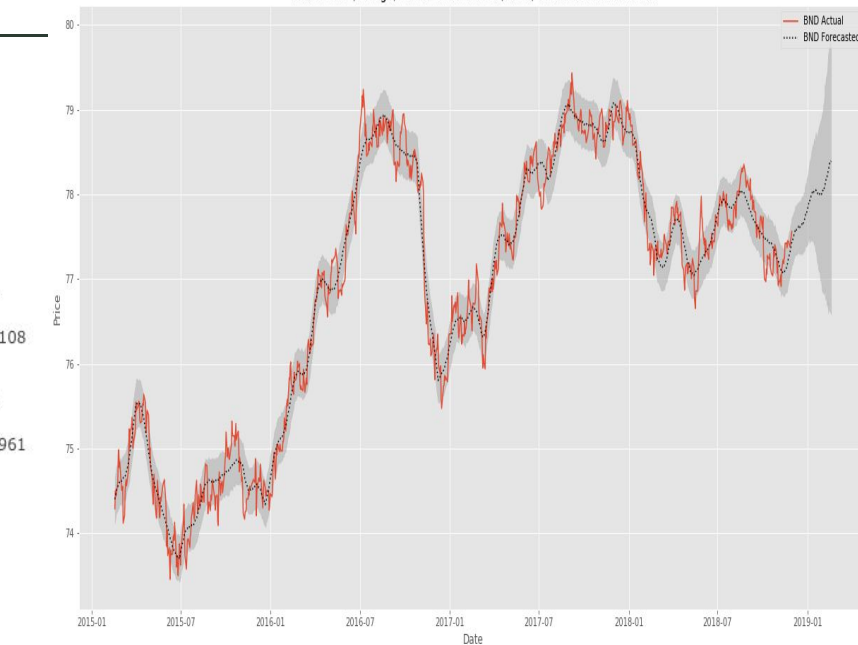
BND Prices



BND Prices



Actual BND (Orange) vs BND Forecasted (Black) with Confidence Bands



All Features

MCD Prices



PFE Prices

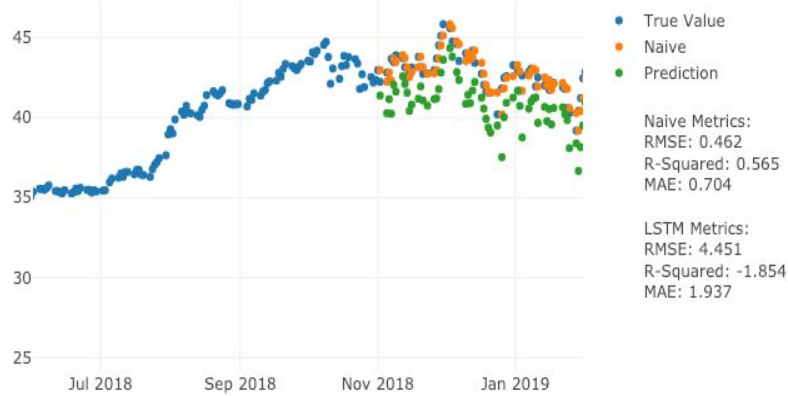


Minimal Features

MCD Prices

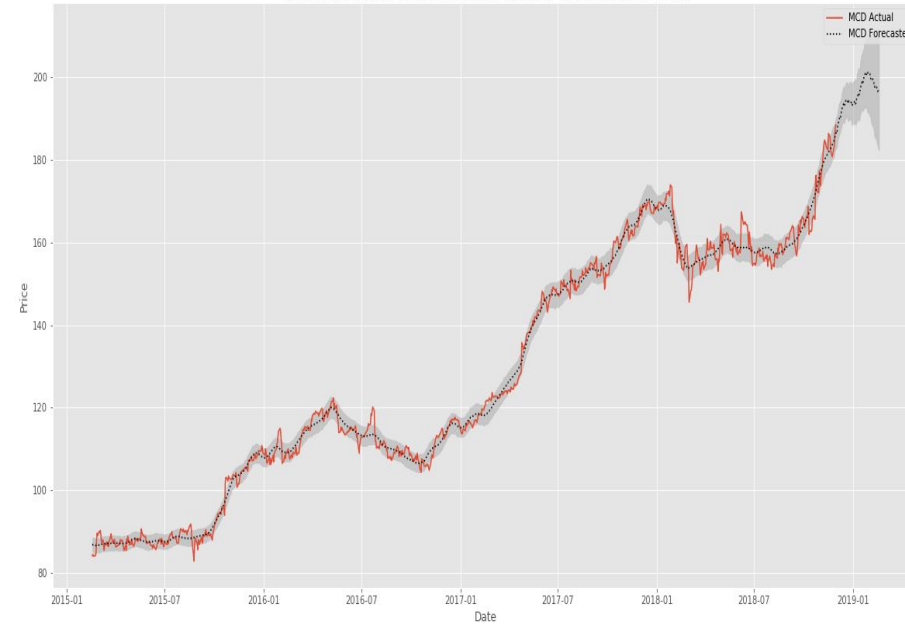


PFE Prices



Facebook Prediction

Actual MCD (Orange) vs MCD Forecasted (Black) with Confidence Bands



Actual PFE (Orange) vs PFE Forecasted (Black) with Confidence Bands



Looking forward

- 1) Look at longer batch windows or predictions further out, because a lot of the economic data is monthly and doesn't have a large meaning on prediction tomorrow's data unless a new report came out. Also a lot of technical indicators are better predictors of medium/longer term trends
 - (Largest correlation to prediction was 1 and 5 day performance for most models)
 - Consistent top correlations: percent change in volume over prior day (log) and difference in volume change (Would like to expand on volume variables)
- 2) Implement NLP on Social Media and 10Q-10K statements to see impacts on stock performance given sentiment regarding stocks
- 3) Add pre-market future data as variables to assess if the stock is likely to close higher or lower before the market opens
- 4) Additional Feature Engineering on Exogenous Data