

Predicting stocks based on performances through years.

• INTRODUCTION

The aim of this project was to use machine learning to predict the future of a certain stock on the stock market. The predictions are made using previously gotten data about a certain stock, up to 2017.

• FINDINGS

The data that is used is from a Kaggle dataset and it was split into three sections. 60% for training the algorithm, 30% for validation of the training and 10% for the test. Each stock contained data: date, opening price, highest price, lowest price, closing price and volume - how many transactions took place during that day.

• TRAINING THE MODEL

We used two machine learning algorithms: linear regression module and random forest classifier.

• CONCLUSION

RandomForestRegressor seemed to only work fine with the AMD stock. Linear Regression worked with all of the stocks. Although we tried to avoid overfitting using numerous methods, it still looks like the models are overfitting. Our models predicted the AMD stock best, as seen below.

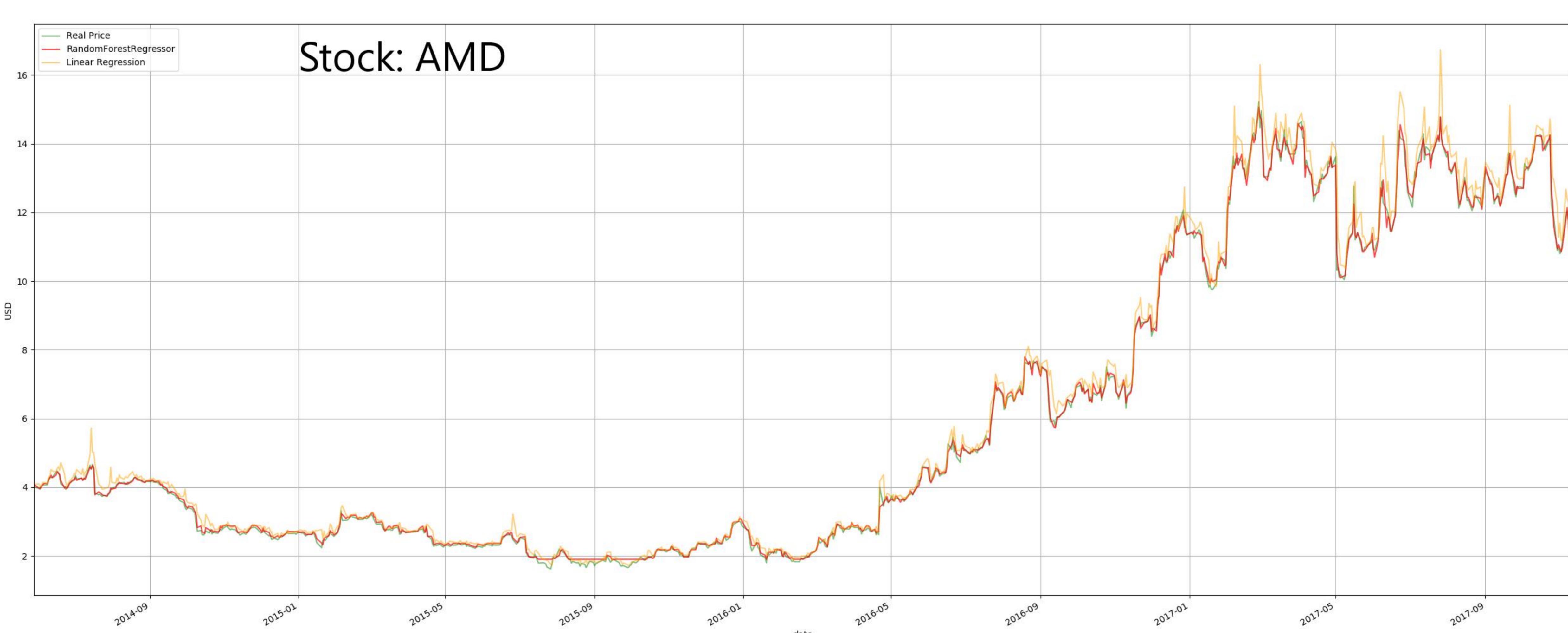
	AMD	MSFT	AAPL
Lin Reg. Score	0.998	0.996	0.999
RandomForest Score	0.999	0.998	-1.25

Data from:

<https://www.kaggle.com/borismarjanovic/price-volume-data-for-all-us-stocks-etfs>

Kernel as a base:

https://github.com/NGYB/Stocks/blob/master/StockPricePrediction/StockPricePrediction_v3_mov_avg.ipynb



QR Code to the repo:

