Interactive Stock Market Time Series Forecasting

(Capstone Project 3)

Pallavi Bothra

Data Science Career Track

Mentor: A J Sanchez

Stock Market Prediction

- Predicting the stock market is very challenging since it has far too many variables to be predicted.
- The market is inherently volatile and unpredictable.
- The traditional algorithms, like algorithmic trading do not fare much better than the average.

Distributed Computing and AutoML for Stock Price Prediction

Capstone Project 3

- Transformed time series dataset to supervised learning dataset.
- Computed additional features, such as previous day's data, last week's mean, median, avg, and std.
- Used AutoML framework to train the data and forecast the stock price for next 30 days.
- Compare the error metrics (MAPE & RMSE) among PMDARIMA, FBPROPHET and supervised learning.

Company	PMDARIMA (MAPE)	PROPHET (MAPE)	AutoML (MAPE)	PMDARIMA (RMSE)	PROPHET (RMSE)	AutoML (RMSE)
Google	0.007	0.03	0.009	11.02	37.16	9.11
Apple	0.008	0.04	0.01	1.95	6.92	1.30
ABM Ind	0.005	0.03	0.01	0.30	1.75	0.40
Alcoa Corp	0.01	0.04	0.01	0.20	0.56	1.12
Future Fintech Grp	0.04	0.32	0.03	0.09	0.66	0.51

https://supervised.mljar.com/

Architectural Diagram

Select company -> select Window -> click forecast

Plotly callback presents the desired output

plotly | Dash

Python Web App

Data Modeling & Forecasting

(PMDARIMA, FBPROPHET, AUTOML)

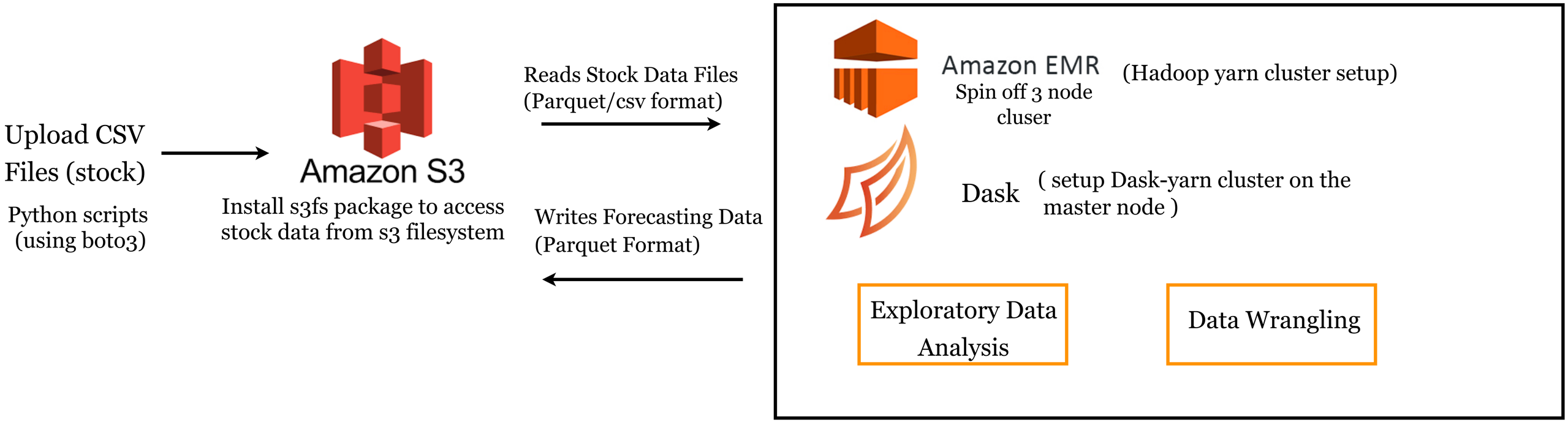
Amazon EMR

Fetches stock data file/model file based on user input (using s3fs)

Amazon S3

- Developed an interactive web application using Plotly Dash where users can provide selection for a company ticker.
- Web application fetches the stock data/model file based on the user input from Amazon S3
- Outputs figures that visualize the stock forecasts for 30 days with error metrics.

Data Preparation WorkFlow



- Uploaded all the stock data files to a bucket in the Amazon S3 object store.
- Spun off 3 node EMR cluster with Hadoop yarn cluster.
- Executed various notebooks that reads all the data files from Amazon S3, performed all the work (wrangling , EDA) and wrote back to Amazon S3 filesystem.

Conclusions & Future Outlook

- Traversed through stock market using various statistical methods, predictive models and compare the accuracy of forecasted resu

- Deployed all the models in the cloud so that it may be accessed on demand by end users or consumers.

- Developed web application so that users can select any company and made informed decision before investment.

- Deep Learning Methods (LSTM)

- Sentimental Analysis

Thank you for your attention