J’s Finance Explorer

Student: Zih-Cin Jian (zxj161530)

Professor: Jason Parker

BUAN6340

J’s Finance Explorer

- A web application based on an additive model for forecasting stock prices

- The goal of this project:

1. Understand and analyze time series data
2. Write code and read documents and articles - python developer’s documentation, stock market terms (bear and bull) etc.
3. Communicate with people who have different backgrounds (see slide #8)

J’s Finance Explorer

Distinct differences between J’s Finance Explorer and other time-series examples used in textbooks and on the Internet listed below

Online Examples	J’s Finance Explorer
• Scikit-learn Linear Regression model • Statsmodels Module • TensorFlow (neural network computation framework) • Static Data • Quandl Online Database • Basic Pandas and Numpy • Basic Matplotlib	• Additive Model (AM) • Prophet (open source software designed by Facebook) • Real-time Data • Web Scraping • Dash (a framework for building web apps) • Interactive Visualization • Cleaner Code(Continuous Code Refactoring) • My Own Modules

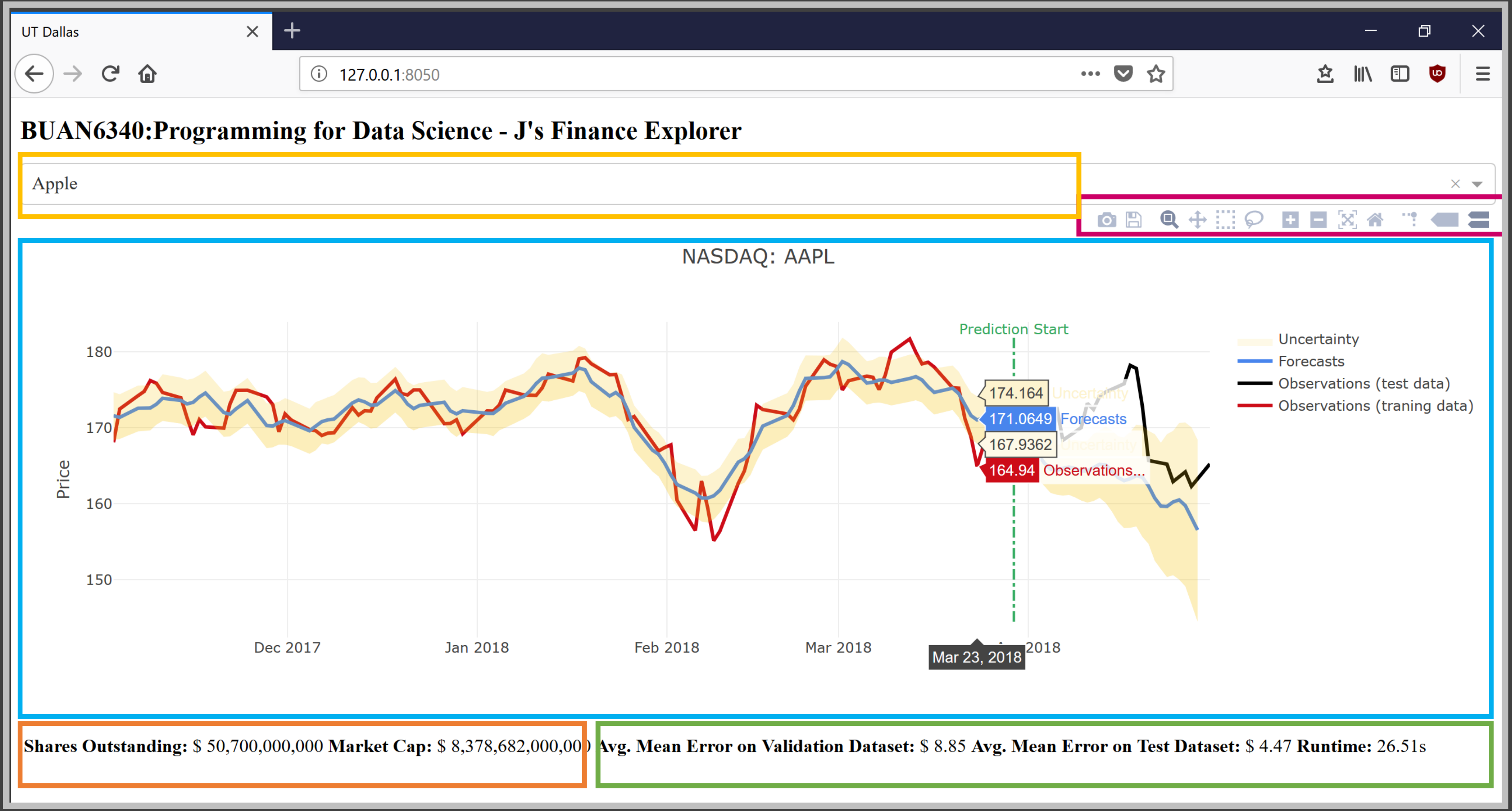
J’s Finance Explorer - Functions

Webpage

Dropdown Menu

Data Visualization

Stock Information



Tool Bar

Model Information

J's Finance Explorer
History

Mar 15

Analyzed Stock Market Raw Data, Researched Trends and Seasonality in Time Series Data

Mar 23

Set up Development Environment – Atom, Auto-Complete-Python, Hydrogen, linterpycodestyle, Github

Mar. 27

Started building Finance Explorer 1st version, using Statesmodels, Matplotlib and Quandl online database

Apr. 6

Did Some Research about How to Perform Cross-Validation and Tune Parameters

Apr. 15

Employed Prophet to Improve Model & Added Web Scraping Feature and Perform Code Refactoring

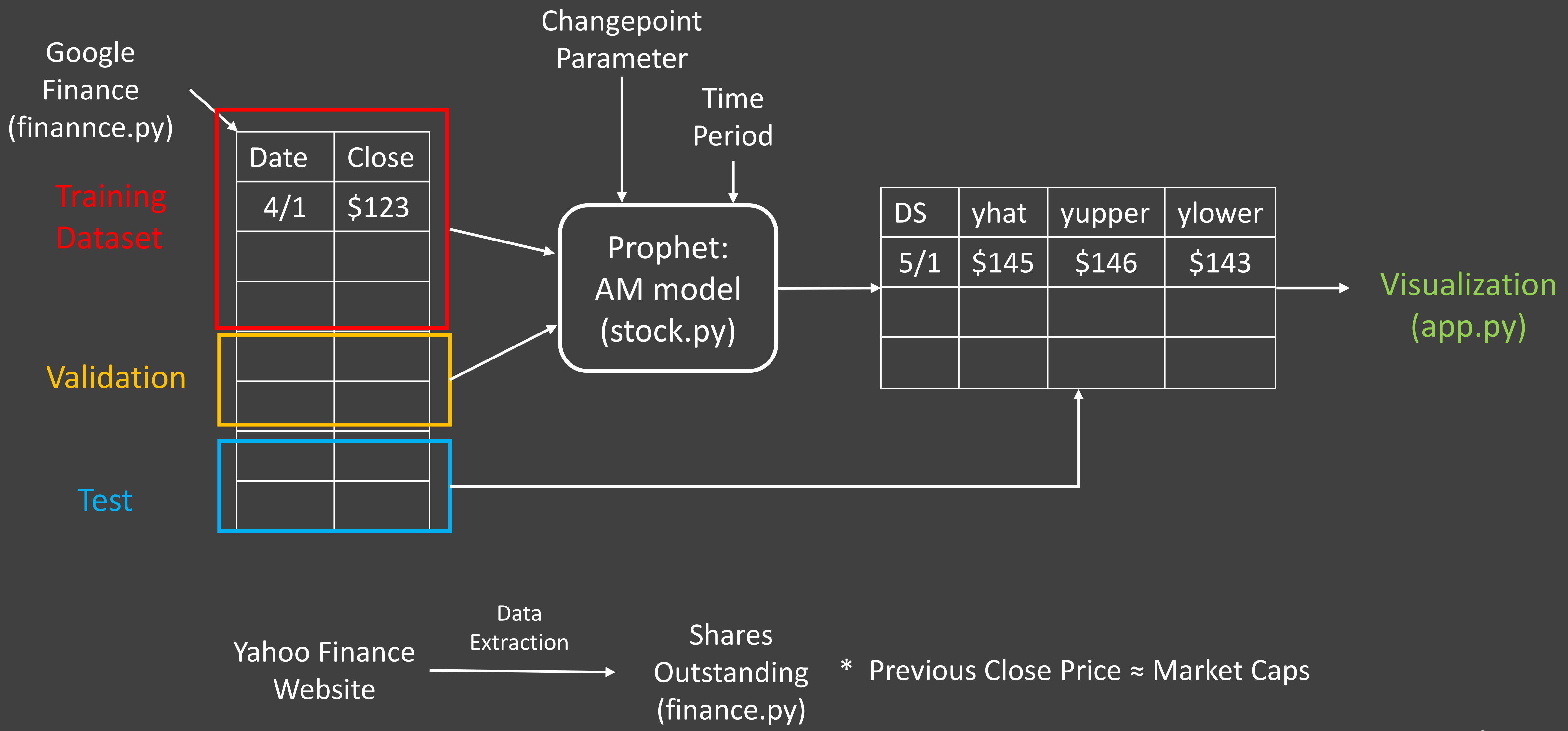
Apr. 20

Added Real-time Data Extraction Feature & Visualized Outputs on Interactive Webpage

Apr. 30

J' Finance Explorer★

J’s Finance Explorer – Framework



J’s Finance Explorer – Interesting Things You May Want to Know

- I was once unable to access google finance data for a few hours right after first data access. Symptoms: requested 1576800 data entries [success] -> requested again in like 15s [failed] -> requested again [failed] -> 24 hours later -> requested 1000 data [success]

- The attribute “data-*” is dynamic in yahoo finance html file. The number (e.g. data-140, data-149) assigned to “Shares Outstanding” changes every time when the page is reloaded.

- The first stock forecast prediction model only had 50% accuracy on trend prediction. (It’s like flipping a coin, literally.)

J’s Finance Explorer – Things I have been asked

Q: Stock price prediction is not something new, and besides you cannot even compete against those apps designed by Google and Yahoo.

A: I think this topic is really suitable for python beginners and those who want to learn how to do analysis on time-series data. In addition, imputing missing data would take a lot time and efforts because It is not reasonable to just insert mean, median or zero into those N/A cells.

Q: Have you consider adding NLP to analyze news, which would help improve the model accuracy?

A: Brilliant idea

Q: Kaggle is a great platform for data science, which allows you share you work with others.

A: Yes Kaggle is a nice platform and GitHub is a good place too. GitHub