Desereé Frink-Larcart, Michael Wade

IDS Project: Milestone 1

Tuesday, March 7th

**Identified Problem Domain:**

Our identified problem domain is the stock market. On a foundational level, we assumed that consolidating ticker information could be feasible by simply evaluating various questions regarding stock prices, or recent news events about a ticker. Taking it to the next level, we could extract some candlestick patterns, which one of us has some previous experience doing over the summer. This leads to a potentially more ambitious task that we could incorporate, anticipating market trends. While ultimately the dream for every financial analyst, our group would be interested to see how different model approaches could fare against the general unpredictability of the market.

**Manually Created Sample Conversations:**

“What is Amazon’s stock price?”

“Amazon is currently valued at $100 per share”

“How has Amazon’s price changed over the last week?”

“Amazon’s price has risen by 4% over the last week.”

“Are there any upcoming events for Amazon?”

“Amazon’s earnings call will be on Friday the 3rd.”

“What are some recent news items about Amazon?”

“Here are the three most recent articles about Amazon: ...”

“Where did Amazon close on Friday the 3rd?”

“Amazon closed at a share price of $100.”

“What are the top 5 cryptocurrencies, in terms of daily movers?”

“The top 5 cryptocurrencies, in terms of daily movers, are Bitcoin (BTC), Dogecoin (DOGE), Binance Coin (BNB), Ethereum (ETH), Cardano (ADA).”

“What is the yearly dividend of Apple?”

“The yearly dividend % of Apple is 3.28.”

“Display the volume throughout Google’s trading day.”

“The current volume of Google is 5000000.”

“Grab the 6 month candlestick pattern of Apple.”

“The 6 month candlestick pattern of Apple may be seen below:



“Is Apple experiencing a downward trend this week?”

“Yes, Apple is experiencing a downward trend this week.”

**Sample Specifications for Manually Created Sample Conversations:**

**TopFiveMessage**

TopFive( a, b, c )

a: top five what [‘cryptocurrencies’]

b: in terms of [‘daily movers’]

c: answer [‘Bitcoin (BTC), Dogecoin (DOGE), Binance Coin(BNB), Ethereum (ETH), Cardano (ADA)’]

example rendering: “The top 5 cryptocurrencies, in terms of daily movers, are Bitcoin (BTC), Dogecoin (DOGE), Binance Coin (BNB), Ethereum (ETH), Cardano (ADA).”

**YearlyDividendMessage**

YearlyDividend( a, b )

a: company [‘Apple’]

b: percentage [‘3.28’]

example rendering: “The yearly dividend % of Apple is 3.28.”

**VolumeMessage**

Volume( a, b )

a: company [‘Google’]

b: volume [‘5000000’]

example rendering: “The current volume of Google is 5000000.”

**CandlestickMessage**

Candlestick( a, b )

a: company [‘Apple’]

b: pattern [‘img\_file’]

example rendering: “The 6 month candlestick pattern of Apple may be seen below:



**TrendMessage**

Trend( a, b, c, d )

a: company [‘Apple’]

b: trend [‘downward’]

c: affirmative/negative [‘Yes’]

d: duration [‘this week’]

example rendering: “Yes, Apple is experiencing a downward trend this week.”

**Data Source:**

All data will be pulled and generated from the Python finance library. An example of how we plan to interact and generate data from the API may be seen within the *python/data* directory. In the *data.py* file, there is a primitive query of the AAPL ticker, extracting historical and realtime data from NSE’s website.

In regards to the questions.txt file, we will need to manually generate versions of these questions to build a usable dataset. A sample of what this dataset may look like is also provided within the *python/data* directory.

**“Happy Path”:**

The user will be able to interact with this dialogue system via the terminal, by simply asking for any desired information. Upon receiving the question, the application will return the relevant output as deemed appropriate by the model implemented. Once this output is received, the user is able to ask consecutive questions, but should be aware that the application will not keep track of previous dialogues.