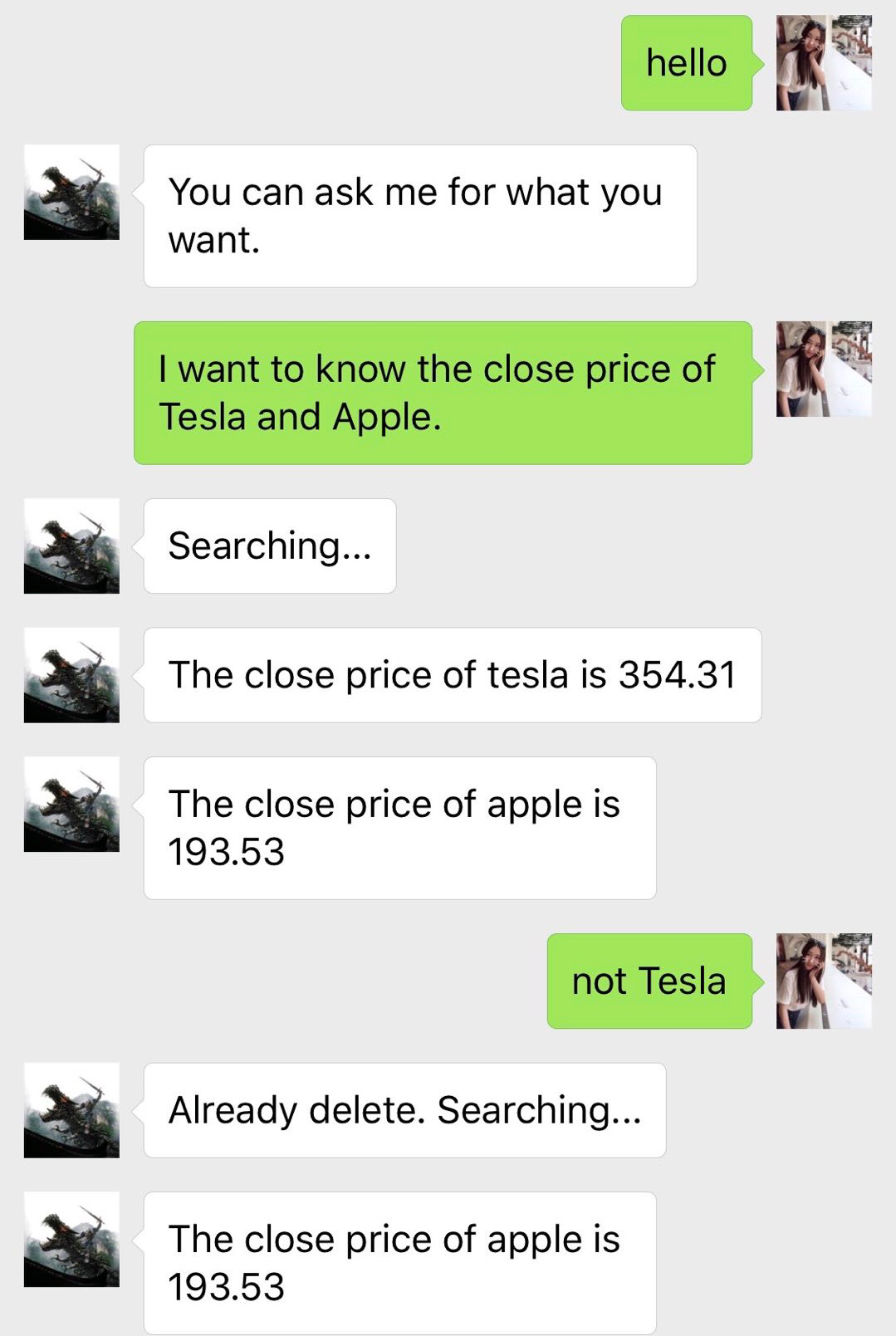
Final Project-

Stock Chatbot on Wechat

Yuming Sun

2018-11-18

Table of Content

1. Introduction：

1.1 Artificial Intelligence …………………………………… 3

1.2 Machine Learning …………………………………… 3

1.3 Chatbot …………………………………… 3

1.4 Financial Chatbot and WXPY ……………………….. 4

2. Theoretical knowledge:

2.1 Regular expression …………………………………. 4

2.2 Support Vector Machine ……………………………….. 4

2.3 Named-entity recognition ……………………………….. 5

2.4 Database Technology …………………………………… 5

2.5 Docker ……………………………………………… 5

3. Coding:

3.1 json ……..…………………………………. 6-8

3.2 Python coding …..…………………………………. 8-14

3.3 WXPY coding …..…………………………………. 14

4. Results:

4.1 Python results …..…………………………………. 15

4.2 WXPY results …..…………………………………. 16-17

5. Conclusion .…………………………………. 17-18

1. Introduction：

1.1 Artificial Intelligence

Artificial Intelligence is definitely one of the most cutting-edged technology nowadays. Artificial Intelligence means simulated intelligence in machines. Human beings create machines that can think of and act like human. The goal of artificial intelligence is to achieve the best performance by learning, reasoning and solving problems. The artificial intelligence faces controversial from the beginning. Some experts treat artificial intelligence as a huge progress of human civilization and believe that artificial intelligence can totally promote the way of living. Some people concern that the highly development of such machines might be dangerous to human’s life and privacy and thus supplant human forces.

1.2 Machine Learning

Machine learning is one of the artificial intelligence which gives computer programs ability to deal with data analysis. There are a few types of machine learning which include Supervised learning, Semi-supervised learning, Active learning, Unsupervised learning and Reinforcement learning. Machine learning focuses on the prediction of data. Machine learning is a type of optimization which helps find the best solution of the problems. The idea of machine learning formulated from statistics, indicating that machine learning has a history of statistics. The theory of machine learning is to generalize from the experience. Generalization refers to a situation when a machine could perform on new tasks precisely after having accepted a learning data set. As a branch of theoretical computer science, Machine learning can be used to different aspects of the life in order to achieve the best solution of problems.

1.3 Chatbot

The chatbot is a computer program that includes a conversation of textual methods. The chatbot is designed in the way similar to a normal conversations performed between human beings. The early chatbots are ELIZA (1996) and PARRY(1972). Chatbot can be used in many aspects including message apps, company internal platforms and toys. There are a few steps in the creation of chatbot. First, the interaction between the user and chatbot is designed. Second ,the context of the user input needs to be fully understand and the correct answer should be built. Then, the usage of chatbot has to be monitored. Sometimes when chatbot faces problems and errors, we need to ensure the effectiveness of chatbots is performed well. The final step needs the help of cloud Platform, which is very useful in building and testing of chatbots.

1.4 Financial Chatbot and WXPY

The main focus of this project is to construct a simple intelligent chat robot which could implement basic conversations about stock issues. Our chatbot is designed to be able to give feedbacks of the following stock issues:

1. The current stock price
2. The closing stock price
3. The value of the stock
4. The stock volume
5. The price of the stock under given condition

Users could query the stock issues of specific companies on a particular date.

Wechat is an application which is similar to Facebook and messenger. It is one of the most widely used chat application in Asia. WXPY is the API provided by Wechat that enable users to build their own chatbot. By using the WXPY, users could communicate with the financial chatbot built on Wechat.

2. Theoretical knowledge:

2.1 Regular expression

Regular expression is a search pattern defined by a sequence of characters and it is a logical formula for string operations. Regular expression is usually used in theoretical computer science programming and formal language theory to find and replace operations on strings by using string searching algorithms. Elements or members in the set of strings required for a particular purpose need to be listed in order to specify the set of strings. The specification is not unique. If at least one regular expression could be found matching with a particular set, there will exist an infinite number of other regular expression which matches with the set.

2.2 Support Vector Machine

Support Vector Machines (SVMs) are supervised learning models. SVMs, as well as the associated learning algorithms, could be used to analyze classification data and regression analysis. The main idea is to map a linearly inseparable set of sample data to a high-dimensional space to make it linearly separable. Given a set of data which has been marked and classified into two categories, SVM would build a model with new examples adding to one category or the other.

2.3 Named-entity recognition

Named-entity recognition (NER), also known as entity identification, is a part of information extraction. NER would extract name entity in the unstructured text and then locate and classify these entities into pre-defined categories. Name entities could be classified into categories like companies, names, locations, moods, quantities and Et cetera. For instance, “Apple”, “Tesla”, “Microsoft” and “Alcoa” in the database and functions we constructed are all identified as the pre-defined category, company.

2.4 Database Technology

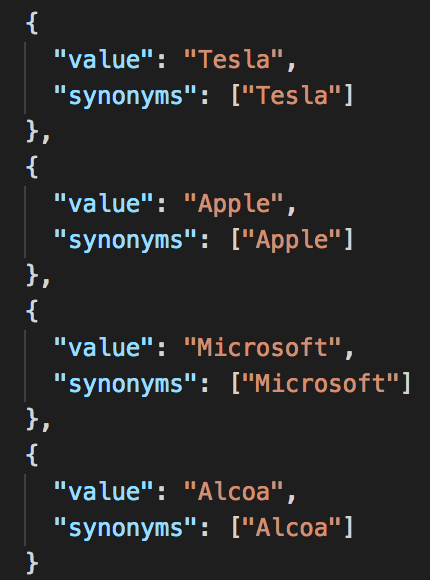
Database technology is a technology that studies, analyzes, and understands the data in a database by studying the basic theory and implementation methods of the structure, storage, design, management, and application of the database. Since data is the object of database technology research and management, all processes involved in database technology are dealing with the date. For example, the database technology establishes corresponding database and data warehouse according to the specified structure by unifying organization and managing data; it uses database management system and data mining system to design a data management and data mining application system that can add, modify, delete, process, analyze, understand, report and print data in the database; it also achieves the effectiveness of processing data by utilizing, analysis and understanding the application management system.

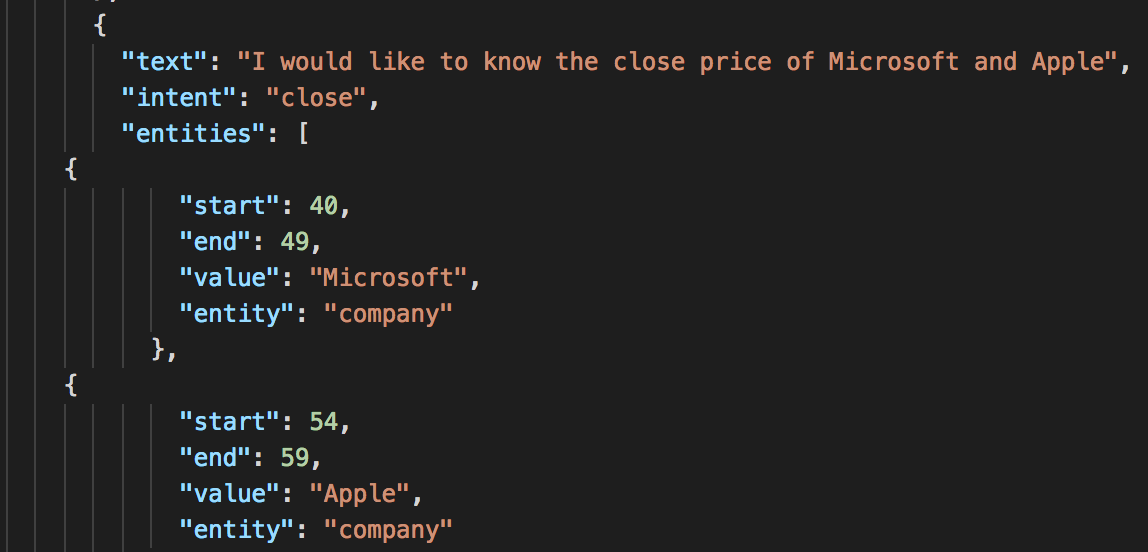
2.5 Docker

Docker is a world-leading container platform that supports open source software. Docker gives the users freedom to design and secure applications with containerization technology. It provides an additional layer of software abstraction on the Linux operating system and an automated management mechanism for operating system layer virtualization. With such advantage, Docker is free of the limitations for the traditional technology which suffers from the technology and infrastructure lock-in, meanwhile Docker can save IT complexity and costs.

3. Coding:

3.1 json

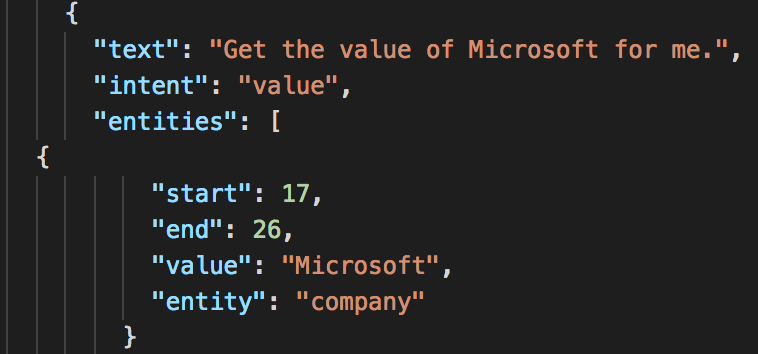
The theme of this project is the construction of an intelligent chat robot based on python library such as rasa\_nlu. The goal is to build a simple financial chat robot which could implement conversations about stock issues. In order to achieve the transaction amount, market value and stock price inquiry at any time (based on the online database), I used techniques such as intent recognition, entity extraction, multiple rounds of multiple queries to build my financial chat robot. The main program structure requires us to first build the json data then import rasa\_nlu to run the model. The structure of the json file is as follows:

***Figure 1: Examples of “entity\_synonyms”***

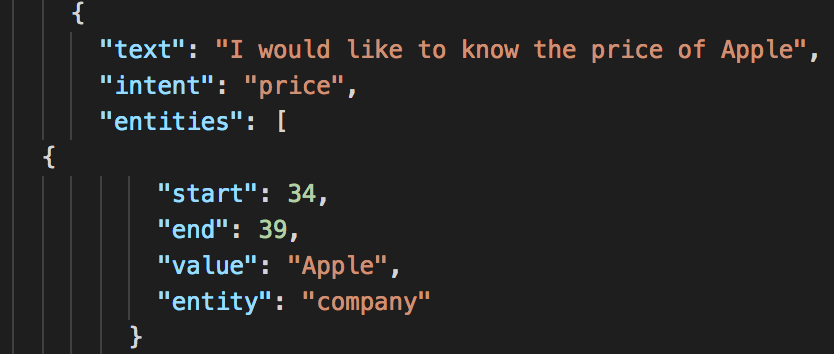
***Figure 2: Example of “intent” : close***

******

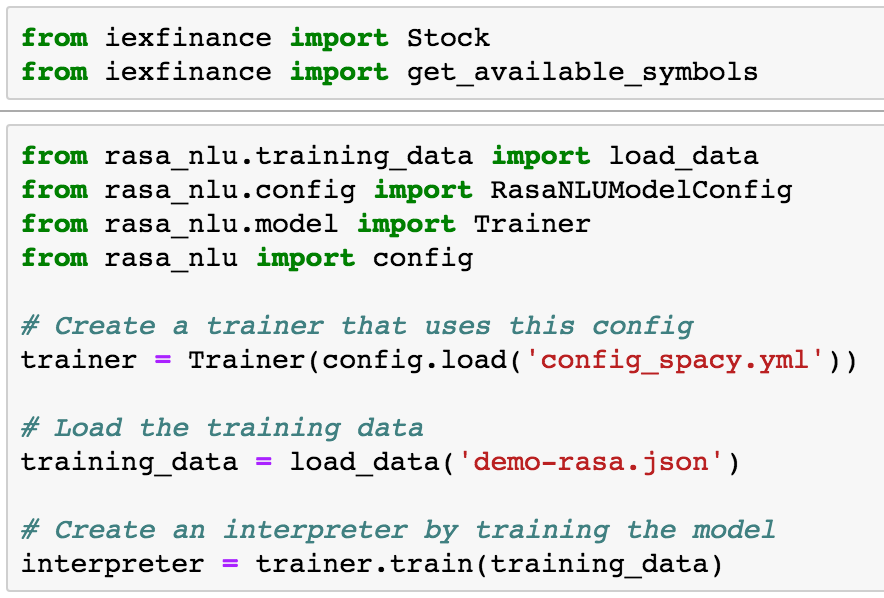
***Figure 3: Example of “intent” : time***

***Figure 4: Example of “intent” : volume***

***Figure 5: Example of “intent” : value***

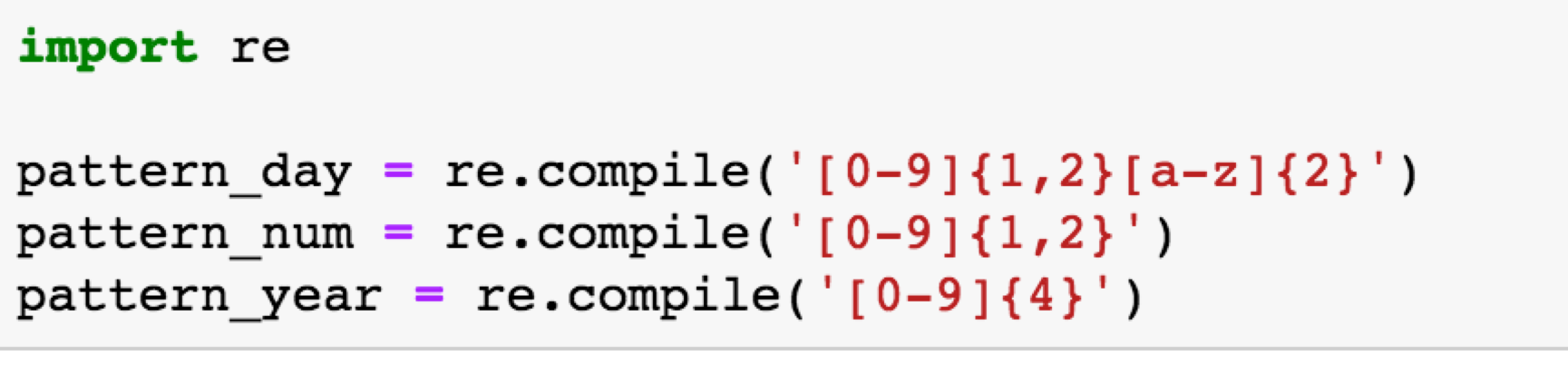
******

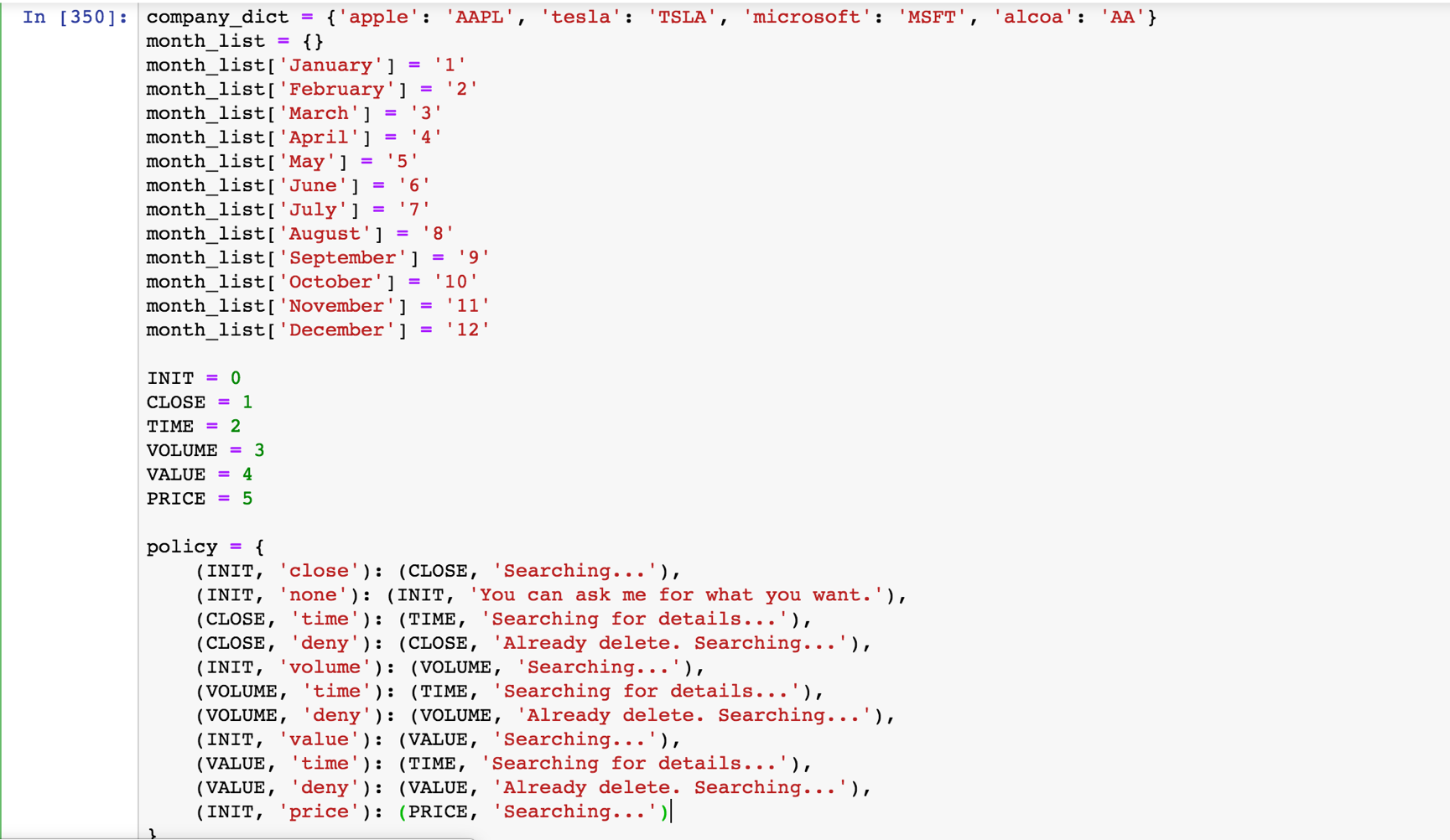
***Figure 6: Example of “intent” : price***

3.2 python coding

***Figure 7***

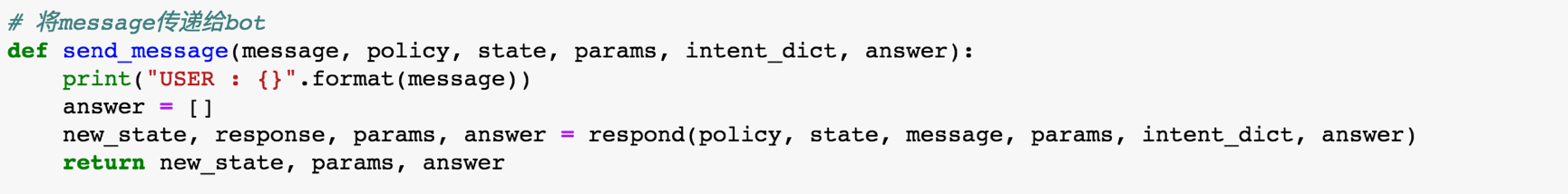
First of all, define constants to be used between functions. For example, the month\_list dictionary is used to convert the month information in the form of English words into a month information represented by a number. The policy defines the initial state, inquiries and the transition rules for the status of the closing price and etc.

***Figure 8*** and ***Figure 9*** show the definitions of the constants:

***Figure 8***

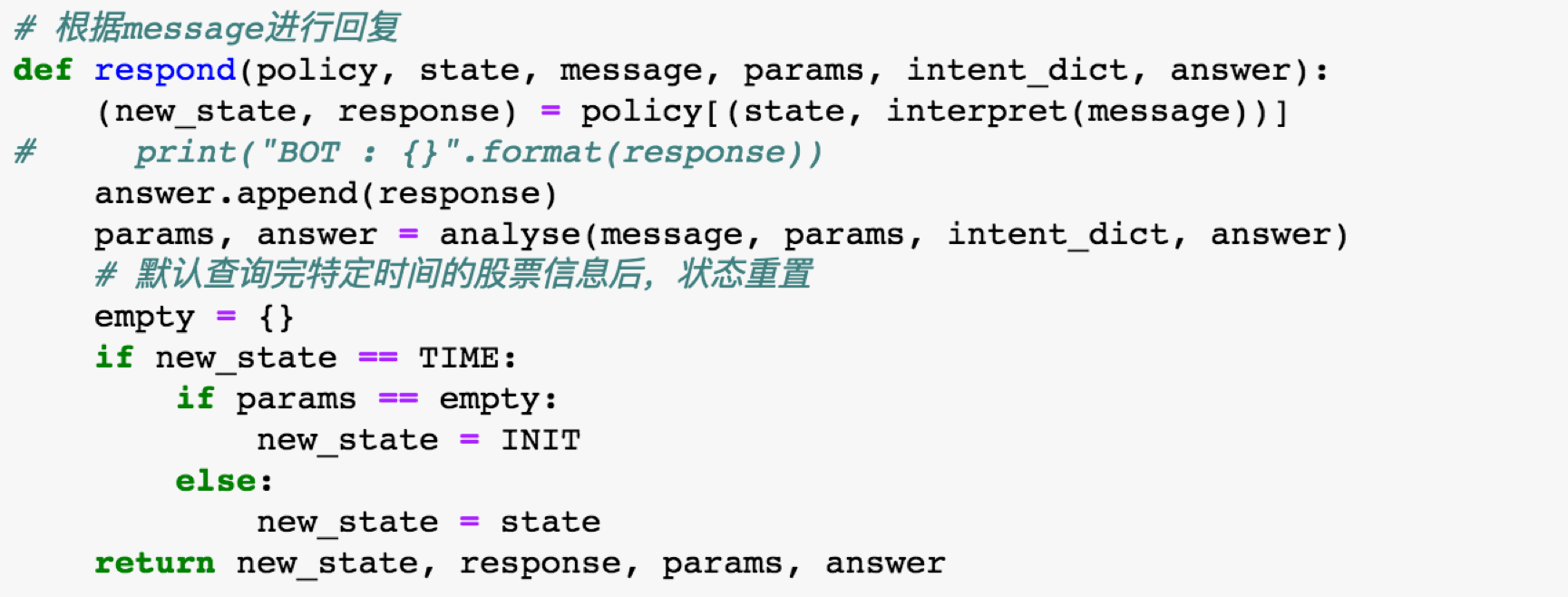
***Figure 9***

The send\_message function defines the basic framework of the conversation. The first step is to enter the user’s request. The request message is processed corresponding to the respond function. The feedback of the processed message is the output of the respond function. The params and the current state of the function need to be returned because they are relevant to the transfer of the following state information and the analysis of request messages.

***Figure 10*** shows functions of the send\_message:

***Figure 10***

In the respond function, the transition of state is first proceeded using the intent recognition of the interpret function. Meanwhile, the required information of state transition, which is the information stored in the policy function, is returned. Then, the request message is processed using analyze function. To be noticed, the parameter params would be emptied because the system defaults to think that users would begin new intent inquiry after time variables are restricted.

***Figure 11*** shows the respond function:

***Figure 11***

The main usage of the interpret function is to process the information based on intent recognition and return the corresponding state information. Each state information here corresponds to the state information in the policy.

***Figure 12*** shows the interpret function:

***Figure 12***

The most important part of the system is the analyze function. The analyze function analyzes the information, extracts the entity, uses API to proceed query based on the intent recognized, and finally returns the result. The analyze function of our chatbot contains the processing method of deny. When users enter two queries and then delete one of them, the deny function would be used to return the final result. Also, the analyze function contains the procession of time. Month and date would be identified from the time variables according to the regular expression. The information of month and date would then be converted into digital form and store in params.

***Figure 13***, ***Figure 14***, ***Figure 15***, ***Figure 16*** and ***Figure 17*** show the analyze function:

***Figure 13***

***Fi******gure 14***

***Figure 15***

***Figure 16***

***Figure 17***

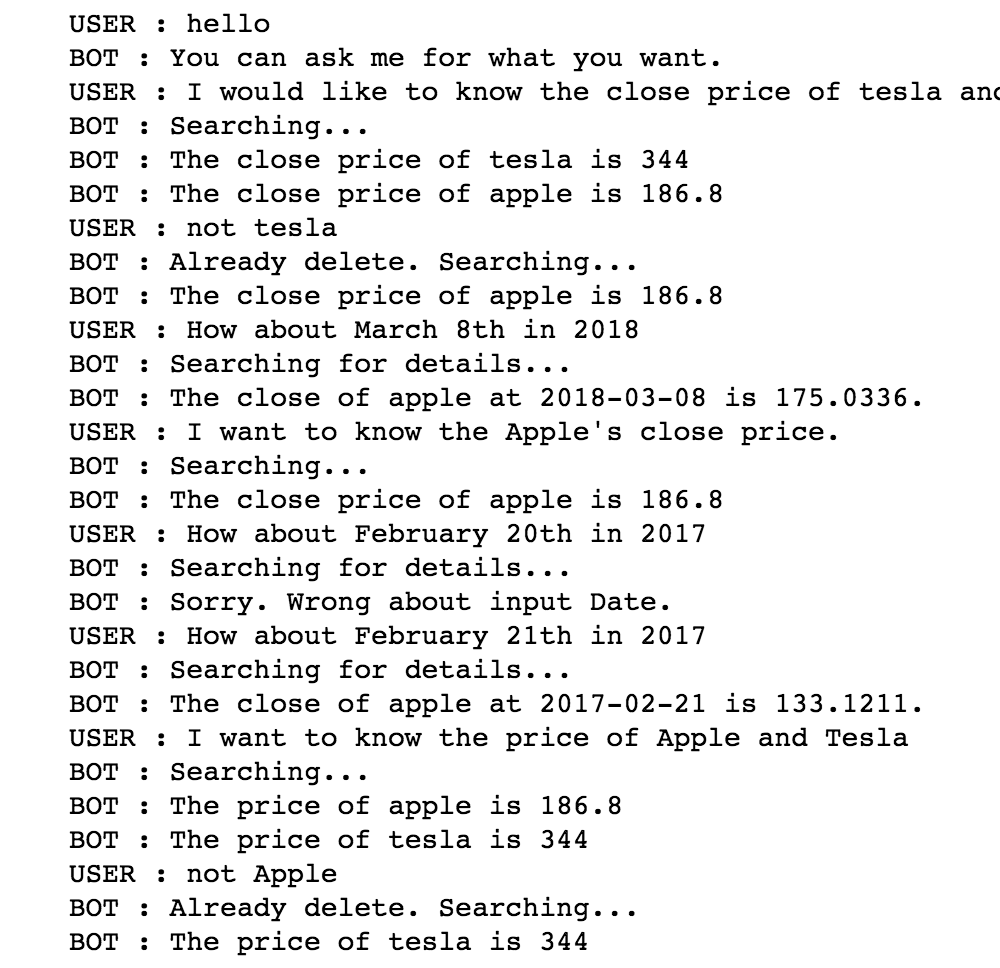
******

3.3 WXPY Coding

***Figure 18: WXPY Codes***

4. Results:

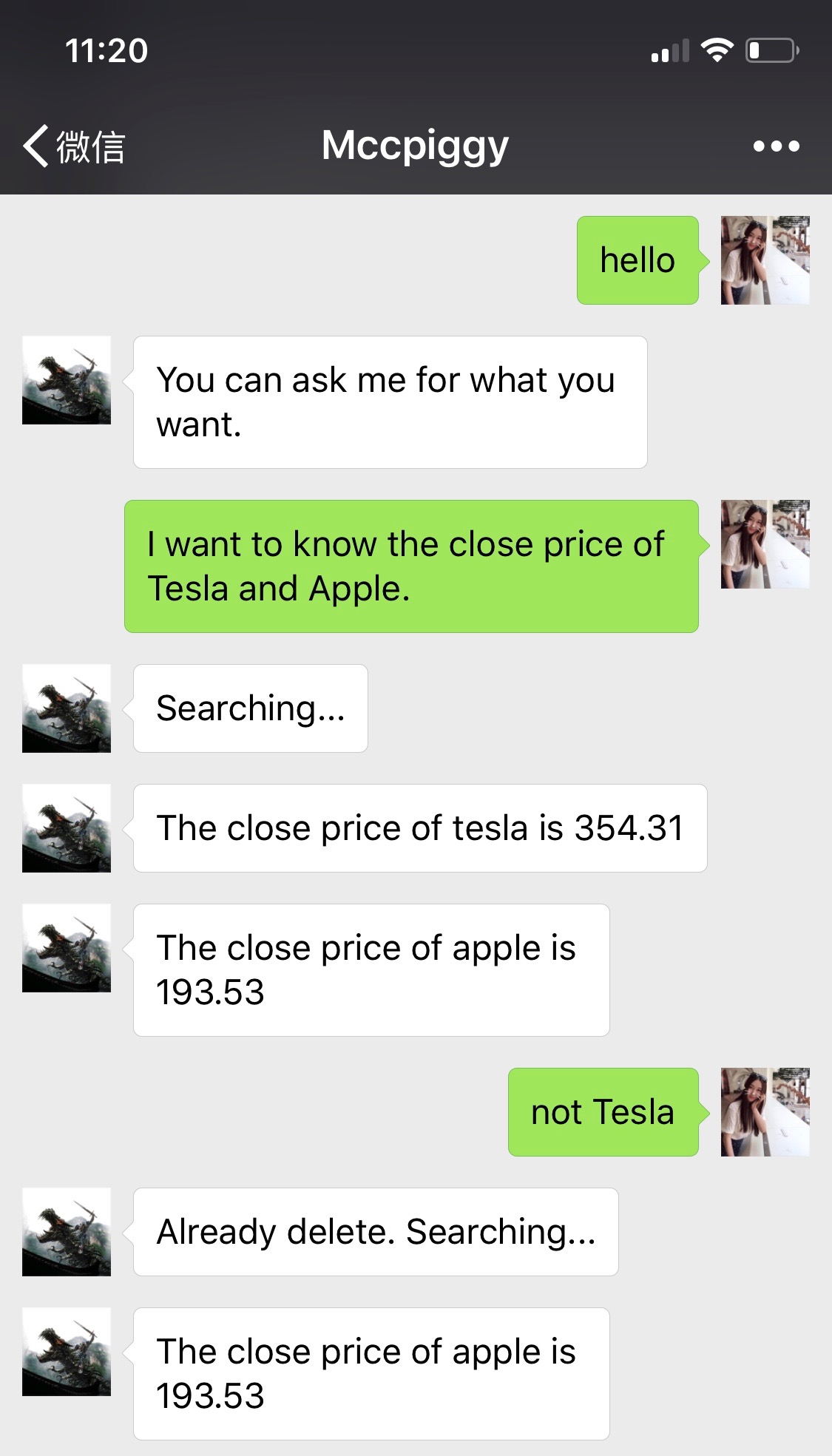
4.1 Python results:

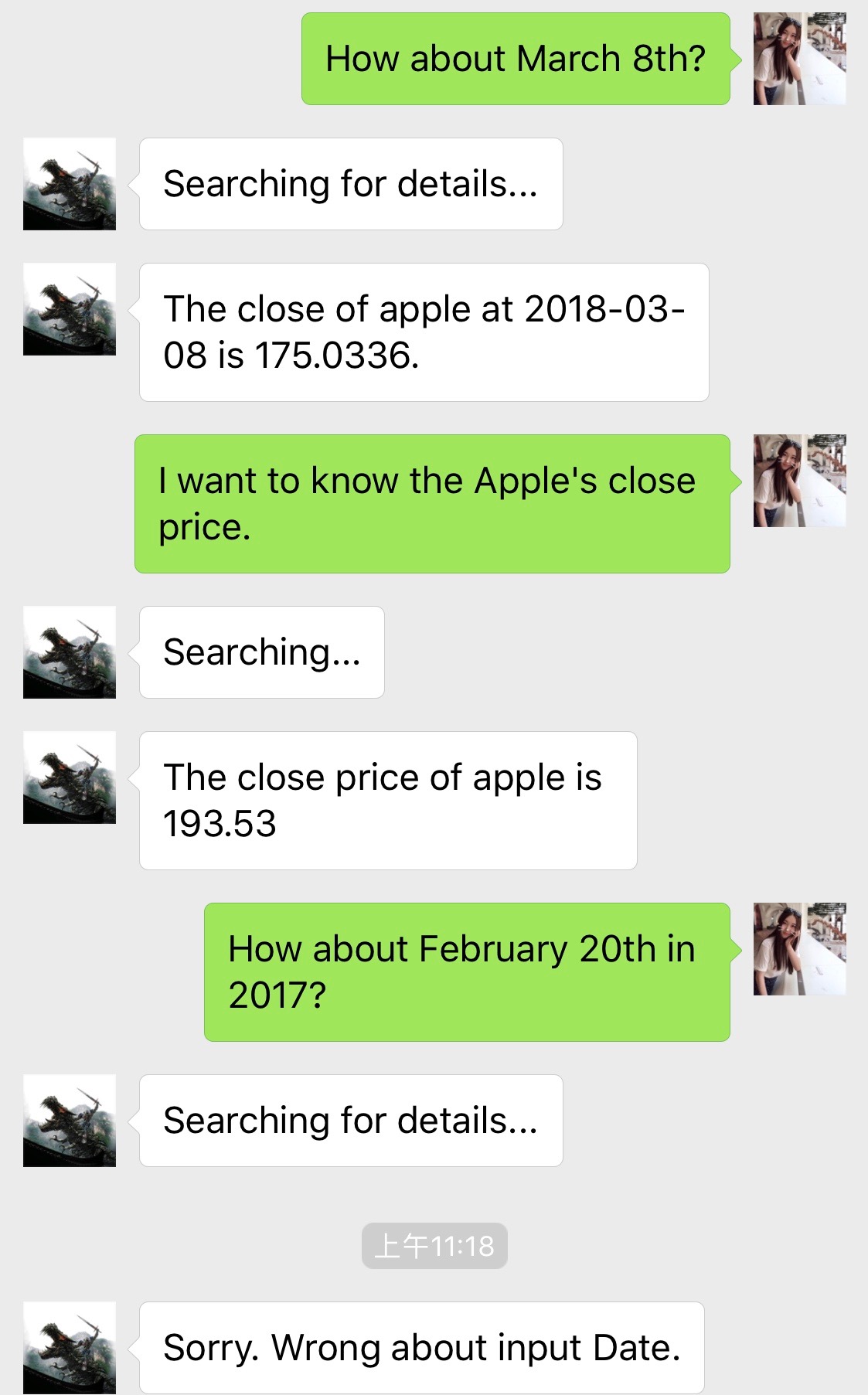


***Figure 19: Python Results***

***Figure 19*** shows the results of the python codes, which is also the conservation between the chabot and user.

February 20th, 2017 is not a weekday, as a result, the stock market did not open. The output of the chatbot is an error message.

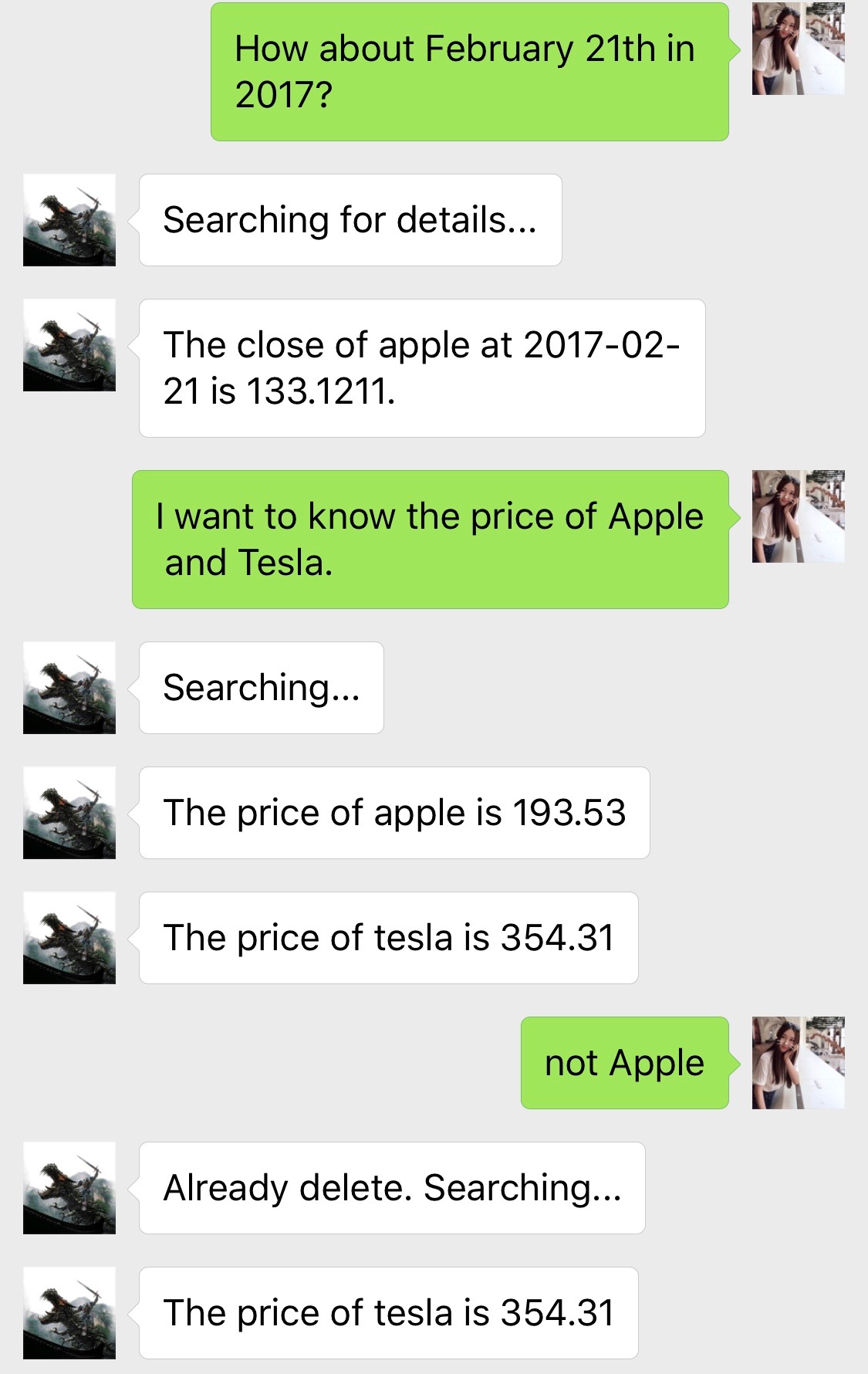
4.2 WXPY results

***Figure 20: Conversation on Wechat (1)***  ***Figure 21:Conversation on Wechat (2)***

***Figure 22 :Conversation on Wechat (3)***

***Figure 20, Figure 21*** and ***Figure 22*** show the conservation between chatbot and user on the application Wechat. Speaker on the left side is the chatbot and speaker on the right side is the user. The output of python codes and the conversation on Wechat are the same.

5. Conclusion

The main focus of this project is to construct a chatbot on the application named Wechat. I worked with another undergraduate student named Hefeng Lin to built our chatbot. Our chatbot is designed to be able to give feedbacks of the following stock issues: current stock price, closing stock price, value of the stock, stock volume and price of the stock based on given conditions. Users could query the stock issues of specific companies on a particular date on Wechat, and the chatbot would give feedbacks. Our chatbot give feedbacks based on the injuries of users accurately and precisely. The chatbot is tested to be effective and successful. There are enormous amount of functions that could be added to the chatbot, for instance, the opening stock price, highest price, lowest price, and even the prediction of future stock market.

This research experience enables me to learn a lot of valuable knowledge. As a student with dual degree in Mechanical Engineering and Statistics, I have plenty of experiences in problem solving, individual and group projects and data analysis. However, it is the first time I use machine learning to proceed data analysis and built a chatbot which analyzes big data and then gives feedback based on the inquiry of users on the application that I use most frequently. Although this research experience is short in time, I comprehende knowledge about big data and machine learning, which are both unfamiliar to me previously, and the utilization of these theoretical knowledge. Also, I have learned about the coding and operations of chatbot during this research experience. I am so glad that my friends and my family members are all really interested in my chatbob. The addition of more functions to the chatbot and the execution of the chatbot on Facebook messenger will be the next step of my project. Building the chatbot on Facebook Messenger and adding more functions are my next project goals.