Chatbot Application on Cryptocurrency

Qitao Xie, Dayuan Tan
Computer Science and Electronic Engineering
University of Maryland, Baltimore County
Baltimore, USA
{qxie1,dayuan1}@umbc.edu

Abstract—Many chatbots have been developed so far and they differ in what they do and how they work. A chatbot is a brand-new conversational agent in the high-speed changing technology world. With the advance of Artificial Intelligence and machine learning, chatbots are getting more and more popular. Essentially chatbot is an extension of other human interface mediums other than phone and other social mediums. In the other hand, Cryptocurrency is a digital or virtual currency designed to work as a medium of exchange. In the current digital exchanging world, investors and interesting parties are eager to know the price and other information of this new type of currency. One of the potential paths to get the info automatically and quickly is through chatbot. We explored the open source python library Chatterbot to apply Itchat API (a WeChat interface) to build a robot chatting application, I&C Chat, on the topic of cryptocurrency. First, we collected question and answer pairs datasets from Quora websites. Furthermore, we also created API call to get the real time quote for the top 25 cryptocurrencies. Then we used the collected data to train our chatbot and implemented a logic adapter to get the price quote of cryptocurrencies based on the incoming question. Itchat API method will return the best matched answer to the asking party automatically. The response time of different questions has been investigated. The results imply that this application is quite useful, feasible and beneficial to the digital currency world.

Index Terms—component, formatting, style, styling, insert

I. INTRODUCTION

There are many words referring to chatbots. Conversational agents, chatbot, or even simply bots have been mentioned previously, while they all mean the same stuff in the point of chatbot view. From wiki [1], a chatbot is a computer program or an artificial intelligence that conducts conversation via auditory or textual methods. People create chatbot to simulate the human-like conversation. Michael Mauldin created the term *ChatterBot* in 1994 [2]. Today Chatbots have been used in various fields, such as marketing, education, entertainment, etc.

Cryptocurrency is new to many people but getting very hot recently. A cryptocurrency is a digital or virtual currency designed to work as a medium of exchange [3]. Different from the conventional currencies, cryptocurrency uses cryptography to secure and verify transactions as well as controls the creation of new units of a particular cryptocurrency. It was introduced in 2009 but it didnt grab mainstream media and peoples eyes until 2012. It may have the potential to compete against traditional payment methods like cash or check, and other online payment method like credit and debit cards and PayPal [4].

Focus on cryptocurrency topic, we designed and implemented our chatbot system to answer users any question on cryptocurrency, form the real time price to definition of different knotty terms used in this filed. The structure of the rest parts of this paper is as follows: In section II, we will introduce why we want to design and implement a chatbot system on cryptocurrency topic. Then in section III the detail of each part of our design will be explained. In section IV our evaluation and results will be illustrated, and we conclude in section V.

II. MOTIVATION AND BACKGROUND STUDY

People pay tons of attention to cryptocurrency for some reasons. One of the most important reasons is that in the past three years their market capitalization and volume have a huge increment. As of May 6, 2018, total market capitalization of cryptocurrencies is bigger than 450 billion USD and record high daily volume is larger than 25 billion USD [5], which is almost 1400% bigger than it was in March 2017. Huge capitalization comes along with huge profile margin. More and more people are attracted into this filed in the hope of cashing in on the trend.

But their prices fluctuate frequently and dramatically. For example, Bitcoin is one of the popular cryptocurrencies that have been in the hot news lately for both growth and volatility. Its highest total market capitalization appeared on Dec 16, 2017 and it was 327 billion USD, rising form 16 billion USD on Mar 29, 2017 with almost 7% daily growth rate. Attracted by the huge growth rate, people are eager to know the real time price of cryptocurrencies.

One more block on the way of getting on board is there are so many new terms and concepts in cryptocurrencies field, as well as detailed operations corresponding to this topic, especially for beginners. Meanwhile those information on the internet is so dispersive that it takes a long time to search and collect all what you want. It would be very convenient if we could get the quick and accurate answers to what we want to know. One of the best approaches is through chatbot, a human-like conversation application.

According to the report [6], chatbot interface has been very popular and keeps a steady growth. It has been used in Instant Messaging system, such as Facebook Messenger, Google Assistant, Amazon Alexa, and so on. Indeed, Instant Messaging has more active users than any other social networks and mail applications [7]. Nearly 4 billion users are accounted for the

top 10 IM platforms. The gradual worldwide acceptance of chat-based interfaces makes the ease of adoption and diffusion of newer technologies built on the top of the pre-existing platforms feasible.

The ease of integration is also a key motivation to develop platforms and frameworks which can synchronize chatbot applications on the field of cryptocurrency. Hence, we decided to implement chatbot application on cryptocurrency.

III. DESIGN

Our chatbot system can provide all those information users may want in two ways. Figure 1 shows the two ways to use chatbot. One is using the web server to host chatbot, which is very convenient and familiar for most web users. The other approach is using WeChat combined with itchat. WeChat is a very popular instant messaging tool which has more than 0.89 billion registered users and it has become a big part of peoples personal social networking [8]. Itchat is an excellent open source WeChat personal number interface [9]. Using itchat to call WeChat has never been easier. We developed a WeChat robot that can handle the information querying. Combining with itchat/WeChat, Chatbot will help you to expand your social connections and facilitate your own working efficiency in the fields you are interested in. Figure 1 demonstrates how chatbot works with WeChat. Just like talking with your friends, the chatbot can give you the answers you want. A simple user case is shown in Figure 1.

Fig. 1. The two ways to use Chatbot.

Our chatbot is working in this way:

- Collect cryptocurrency related popular questions and answers, and build the dataset of Q&As.
- Train the chatterbot using the datasets of the known statements and responses.
- Process input in some logic ways. Using a number of machine learning methods to generate the responses. Search algorithm and classification algorithm are deployed on logic adaptors.
- Return the response. It could be one of the formats such as console, API, speech synthesis, etc.

Logic adapter is very important here. You can use the pre-built data set to select the response via matching logic method,

which is based on pre-built data set. Chatterbot library [10] has a nice way to train the chatbot through the datasets for the better responses. Another way to get the response in logic adapter is to call external API to get the required information. For example, to get the price quote for the cryptocurrencies, we implemented a way to call coin-market API [5] for real-time price information. API calls can be embedded inside logic adapter so that Chatbot can respond the price quote question directly and automatically. We can use multiple logic adaptors in chatbot system. For example, the best match adaptor uses a method to compare the input statement to the known statements. Once the closest match to the input statement is found, another method is implemented to select one of the known responses to that input statement. Another example of logic adaptors is low confidence response adaptor. A specified default response is returned when a response cannot be determined above the threshold of confidence score. Below we discuss each step of building our chatbot in more detail.

A. Data collection

The Knowledge base is built through a series of questions and answers via Quora sites. Chatbots are only as intelligent as the knowledge they can have. Collecting enough training data used to train machine-learning classifiers for information retrieval is crucial to achieve human-like interactions. We wrote a Python crawler program to collect multiple Q&As on the topic of cryptocurrency from Quora.

With Python Requests, Selenium, and Beautiful Soup modules in place, it is easy to scrape the web pages to gather the data for the chats data training that youre working with. The following steps are used for Q&As crawling on Quora:

- The first script is used to get questions given one keyword. In our case, we chose the three words "cryptocurrency", "bitcoin", and "ethereum" as the keywords of Quora search. The questions collected from the script are saved in a text file. As you may know, Quora only shows the 25 questions for a given keywords at a page. Using Python's Selenium module, we are able to simulate the browser's behavior of the JavaScript to move down in the page to get more questions on the Quora querying. This is very useful for web scrawling.
- The second script is used to call Quora to get the title of the question and the answer of each question we collected from the first script. The titles and the answers are stored in a special formatted file. To be used for the data training in our chatbot, the data format has to follow the one shown in the Table 1.

We can also collect more Q&As using Itchat [11], a robot WeChat application. We developed a python application using Itchat and chatterbot API, to automatically collect the questions in some cryptocurrency-topic WeChat groups. We will discuss more on Itchat application below.

For the real time price, we implemented a logic adaptor inside the chatbot application using coin-market API to get the real time price of the top 25 cryptocurrencies. Whenever

```
categories:
- cryptocurrency
conversations:
- - What is cryptocurrency?
  - Cryptocurrency is a medium of exchange, created
    and stored electronically in the blockchain
    ...
```

the input statement has the keyword like "price of bitcoin" or other cryptocurrencies, it will try to call coin-market API to get the real time price of that cryptocurrency. Currently only the top 25 cryptocurrencies are supported for this functionality. We will expand the functionality to support more cryptocurrencies in the future.

B. Chatbot data-training

The open source library Chatterbot comes with a corpus data and utility module that makes it easy to quickly train the bots to communicate. The library has the built-in tools that help simplify the process of training a bot instance. The basic step of data training is loading preset example dialog into chatbot's database. This will create the special data structure that represents the known input statement and response. With the given the data set, the data trainer class will create a knowledge base so that the known input and response are correctly linked. Further, we also specify the collected data set for data training to focus on the topic of cryptocurrency.

TABLE II
ITCHAT SAMPLE PYTHON CODE TO TALK WITH WECHAT.

```
import itchat

@itchat.msg_register(itchat.content.TEXT)
def text_reply(msg):
return msg.text

itchat.auto_login()
itchat.run()

...
```

The data for chatbot's training has to meet the specified format if using the library's data training class. You can develop your own training class for your own data format. For example, if you want to use the data collected from Twitter, you can create a new app using your twitter account. Thus, you can train your chatbot using the data from Twitter.

C. Itchat application

Itchat is a complete and graceful API for WeChat interface [9]. Itchat can be used to respond the text messages posted by the other users in the WeChat. The simple python code is used for this purpose in the Table 2.

We implemented an application to combine Itchat and chatbot, using sqlites3 to save the records including questions and corresponding answers. It can have 2 purposes. First you can enlarge our question pool. Secondly, if the input is known, Itchat can return the preset response right away via chatbots algorithm.

D. Input processing

Based on the different search algorithms, the closely matched response is returned with the highest confidence score. Inputs will be classified into three categories: greetings, price of cryptocurrency, and cryptocurrency related Q&As. It is a technique of artificial intelligence used in the design of a Chatbot. The input is matched with the inputs saved in the database and corresponding response is returned.

To generate the response, a number of search algorithm and classification algorithm are used. For search algorithm, the following have been implemented: the similarity of an input statement to the known statement; the frequency in which similar known responses occur; the most possible category of an input statement. For classification algorithm, the chatbot uses nave Bayesian classification to determine if an input statement meets a particular set of criteria for a response to be generated from that logic adapter.

Typically, there are two types of Chatbot systems. One is rule-based chatbot, and the other is AI-based chatbot. Rule-based chatbot relies on the known rules. It is relatively limited depending on the size of knowledge base and the rules. AI-based chatbot is more intelligent as it understands the natural language. It is not just limited by pre-built commands. It could get "smarter" since artificial intelligent can help bot to understand the unknown things.

An important feature of chatbot is determining the context of the current user expression. Especially when the input has the multiple meanings, it could reply on the history of conversation for the modern chatbot to figure out the best match response.

The chatbot we are building is more like rule-based bot. We use the existing corpus data and collected data set for data-training. We also use the implemented API to get the information we want to know.

E. Database Creating

Two-dimensional string arrays are applied to build a database. Rows in the array are used for requests and responses. All the even rows contain the requests or questions and all the odd rows contain the responses or answers. Columns in the array are applied to save different types of questions that could be asked by the user and responses that a Chatbot can offer. There is one row in the array which contains default responses which is used when the matching question is not found in the array.

F. Chatbot Deploying

We developed the two ways for deploying chatbot. One approach is through WeChat. Another approach is through

web app server. To access it via WeChat, we implemented it via Itchat API application, which can reply to the users automatically as said earlier. To access it via web app, we deploy chatbot via Django, which is a free and open source web framework and can be directly integrated with Chatterbot. Django is used to build the chatbot interaction UI to allow the users to communicate with the bot, as demonstrated in the Figure 4. The deployment is simple and easy to manage.
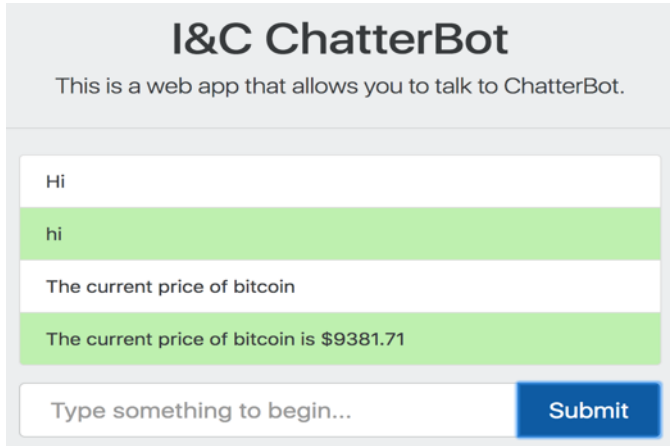

Fig. 2. Chatbot on Web App.

G. Process flow diagram

Chatbot's process flow diagram is shown in the Figure 5 [10]. The chatbot starts with no knowledge if untrained. Each time a statement is input, the chatbot saves the input and the statement that is responsive to into the database. The more the inputs and the correct responses, the more accurate the responses to the input statements are in the future conversations, which is also the purpose of data training for chatbot. When the knowledge base is enriched, and the Q&As are enlarged, the chatbot is able to respond much more correctly. The accuracy of the responses is one of most important factors for evaluating the bot's performance. To talk human-likely, the chatbot should simulate the human's word usage.

In chatbot, the logic adapter determines how to respond to the input statement. After one input is given, the program will compare it with each question statements by calculating a confidence score for each of them. Only one question statement which has highest confidence score will be selected as best matching question statement, and then the corresponding response statement to that best matching question statement will be offered as the response to the input. A threshold of the confidence score will also be preset to make sure the best matching question statement is good enough as we want, in this case, it should have similar meaning of the input. If the response cannot be found in database, or the confidence score is lower than the pre-set threshold, a default response sentence "do not understand" will be offered as the return response to the input.

During the script work, if the designer wants the conversation to be as natural as it can be, open-ended questions have been proved to be a right choice. With the user typing whatever he wants to, the chatbot needs to be diligent in framing questions and processing responses. Moreover, when the chatbot does not come with AI brains, creating relevant responses to user query will become an ordeal. The designer would be better off in framing open-ended questions only when it is extremely essential. Further, open-ended questions can steer the conversation away from end goals. Thus, it is much useful if keeping its communications concise and understandable. One bot cannot cover everything since there are too many topics in the universe. That is why we only implemented logic adaptors to build the bot focusing on one small field or topic.
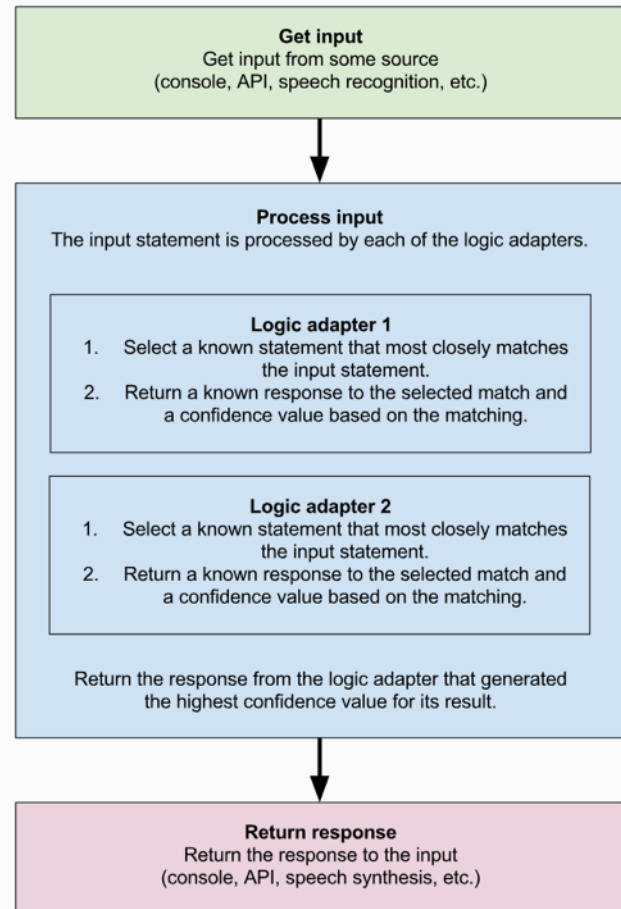

Fig. 3. Chatbots process flow diagram.

IV. RESULTS

The chatbot works by adopting pattern matching. It uses pattern matching to classify the input text and generates a suitable response for the user. Chatbot knows the answer only because the input is in the associated pattern. Similarly, chatbots respond to anything relating to the associated patterns. But it cannot go beyond the associated pattern.

For each kind of question, a unique pattern must be available in the database to provide a suitable response. Nave Bayes is the classic algorithm for text classification. For an instance, a set of statements are given for a particular class. With a new input sentence, each word is counted for its occurrence and is counted for its commonality and each class is assigned a score. The highest scored class is the most likely to be associated with the input statement.

Without data training, the response will always return the input statement back. Table 1 shows the different responses to the input "Most popular cryptocurrencies" based on the selected logic adaptor in the chatbot after data training. Best match logic adaptor returns the response we intend. Since Price adaptor is for the real time price for top 25 cryptocurrencies, the response for this input to this adaptor will be unknown. Multiple logic adaptors (combining those three adaptors) will be ideal.

TABLE III
DIFFERENT RESPONSES BASED ON LOGIC ADAPTORS FOR THE INPUT "MOST POPULAR CRYPTOCURRENCIES".

	Response
Best Match Adaptor	"Bitcoin, Ethereum, Ripple, Litecoin, Monero, Ethereum Classic, Dash, Augur, NEM, and Waves."
Price Adaptor	"Most popular cryptocurrencies."
Low Confidence Response Adaptor	"I am sorry, but I do not understand."
Multiple Adaptors (combining the three adaptors above)	"Bitcoin, Ethereum, Ripple, Litecoin, Monero, Ethereum Classic, Dash, Augur, NEM, and Waves."

The score signifies which intent is most likely to the sentence but does not guarantee it is the perfect match. Highest score only provides the relativity base.

Training time is investigated. The average training time is one second for our data set. The bigger data set, the longer the training time.

TABLE IV
THE RESPONSES OF REAL-TIME PRICE QUOTE FOR THE TOP THREE CRYPTOCURRENCIES.

Input	Response
"Price of bitcoin"	"The current price of bitcoin is $8386.25"
"Price of ethereum"	"The current price of ethereum is $713.123"
"Price of ripple"	"The current price of ripple is $0.710319"

Our chatbot is able to return the real time price for the top 25 cryptocurrencies correctly. Table 4 demonstrates the response of the top 3 currencies respectively, depending on the input question. The testing shows the chatbot responds to the questions consistently. It is averaged 100ms. The performance is acceptable to most users.

Table 5 shows the relationship between the data training time and the number of questions and answers. The more the questions and answers, the more time it takes for data training.

We assume every question should get the answer in the chatbot. But sometimes the question may be unknown to the

TABLE V
THE DATA TRAINING TIME VS THE NUMBER OF Q&AS IN OUR CHATBOT.

Input	Response
20 questions	0.45s
100 questions	2.15s
1984 questions	50.89s

bot. When it happens, the question or the default response may be returned as the answer after searching when low confidence response adaptor is not set. If low confidence response adaptor is set, the default response is returned. The threshold of the low confidence is configurable. In our case, we set the threshold to 0.75. It means that when the input match rate is lower than 75%, it will return the default statement as response, as shown in Figure 4. Even though it is not ideal, our bot is limited just on the topic of cryptocurrency and related knowledge. For the top 25 digital currencies, Chatbot is able to promptly provide the real-time price quotes. Besides the price quote question, Chatbot can answer the questions successfully that are highly matched with the training data set. Certainly, the better Logic adapter on how to return the best responses is for future work.
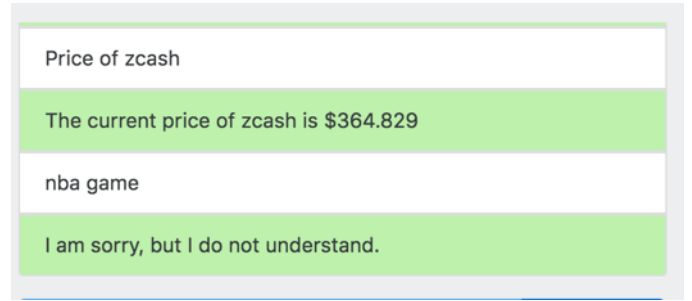

Fig. 4. The response to the unknown input.

It is worth mentioning that some people may argue that Google assistant and Apple Siri can do the same way as what our bot does. As of now, Apple Siri cannot give the real time price but only provide a reference link for the price quote of the cryptocurrencies. Our bot can simply provide the answer to the users. Google assistant could display the answers right away. But for some regions that cannot access Google our bot is an alternative way for the quick user query on the topic of cryptocurrency. Moreover, our chatbot is focusing on the field of the hot topic cryptocurrency.

V. CONCLUSION

In this paper, we use the open source Chatterbot and Itchat to implement the special chatbot for the application on the field of cryptocurrency. We collect the data set for knowledge base of cryptocurrency to data training on our chatbot. The testing shows our chatbot is excellent to respond the certain questions related with cryptocurrency. It is more convenient for users information acquisition especially when the user wants to follow the price trend of the interested cryptocurrency. In the future work, we will continue improving logic adapter

implementing of the response generation and enhancing the knowledge base on the cryptocurrency. In addition, we could extend the simple way of chatbot implementation to build the chatbots on other fields in a practical way.

REFERENCES

[1] Chatbot in Wikipedia. Retrieved from https://en.wikipedia.org/wiki/Chatbot.

[2] Michael L. Mauldin. Chatterbots, tinymuds, and the Turing test entering the Loebner prize competition. AAAI-94 Proceedings, 1994.

[3] Ameer Rosic. What is Cryptocurrency: Everything you need to know. Retrieved from https://blockgeeks.com/guides/what-is-cryptocurrency. 2017.

[4] Gandal, Neil, and Hanna Halaburda. "Competition in the cryptocurrency market." (2014).

[5] Cryptocurrency Market capitalizations. https://coinmarketcap.com/api/. 2017.

[6] Manish Dudharejia. Chatbots are the next big platform. Retrieved from https://www.entrepreneur.com/article/298600. 2017.

[7] W. Mckitterick, "The Messaging App Report: How instant Messaging can be monetized," Business Insider.

[8] 2017 WeChat Users Data Report: There are more than 0.889 billion active users. http://www.sohu.com/a/136382735_184641. 2017.

[9] LittleCoder, Tempdban, Chyroc. ItChat A complete and graceful API for Wechat Retrieved from https://github.com/littlecodersh/ItChat.

[10] Gunther Cox. ChatterBot Machine learning, conversational dialog engine. Retrieved from http://chatterbot.readthedocs.io/en/stable/index.html.

[11] Jia, J. (2003). The Study of the Application of a Keywords-based Chatbot System on the Teaching of Foreign Languages. ArXiv preprint cs/0310018. Retrieved from https://arxiv.org/abs/cs/0310018.