

NLTK Simple Chatbot Report

Ashley Nnawugo

Introduction

Chatbots are systems built using Artificial Intelligence, they essentially replicate a conversation. My chatbot takes in line by line queries and responds depending on the user's intent, whether it is a question, small talk, name management or to close the program. My chatbot makes use of natural language processing in order to respond with the correct answers. I chose to build it with the following features as this was what had been suggested to us by our lecturers. I believe that these features are the most important features of a chatbot as the main features of a chatbot are to answer questions, match intent and to replicate a human conversation. Human conversation mostly has a lot of small talk, identity management and some question answering.

Background

Systems which do similar things to what my chatbot does include the likes of Apple's Siri, Google's search engine, Amazon's Alexa, and many other chatbots used for customer service developed by large corporations such as banks and retailers. Siri^[1] and Alexa are advanced chatbots which have many features which are attached to it, these include voice control, transactions, small talk and question answering. Google's search engine ^[2] is largest search engine in the world, it works by using information retrieval. Information retrieval is when information needed, is queried, this info needs to be processed and once processed, it can search for the relevant answers. Different techniques used in information retrieval include, string matching, index matching(fast) and machine learning (fastest). Google's search engine makes use of all these techniques. All chatbots make use of large vast datasets, and machine learning in order to process queries and match them to the correct result. For example Siri makes use of large datasets in order to match small talk with small talk responses, and also classifies to know if the user intends to have small talk. Chatbots used by big corporations, such as banks^[3] use intent matching in order to direct customers to the right support, fraud prevention and many other things. I based my chatbot on the features of information retrieval, with string matching and index matching, also small talk using a dataset, and finally intent matching. My chatbot is a much simpler version of Siri's small talk and a much more simpler version of Google's information retrieval system.

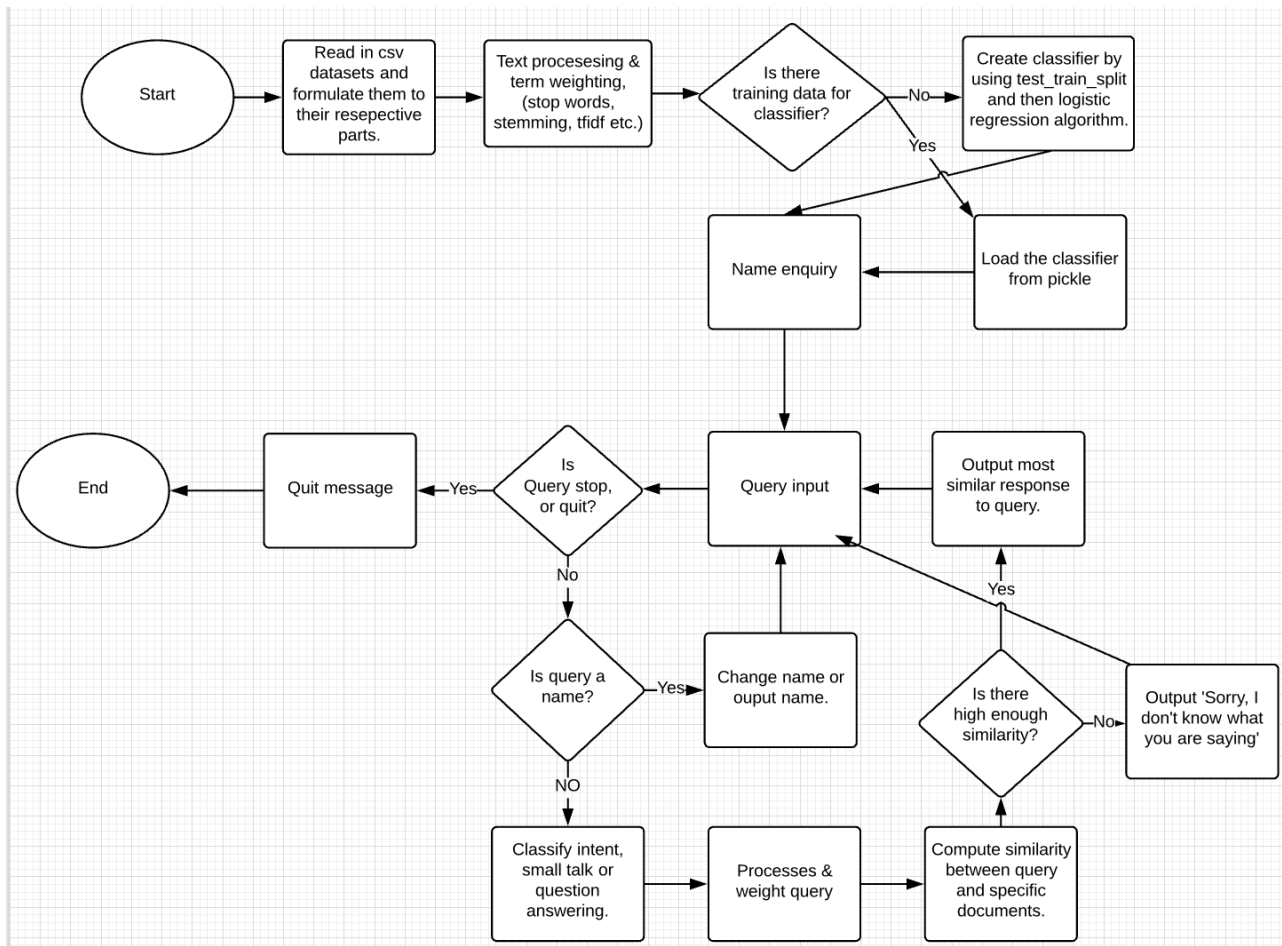
Proposed system

Functionality

The system I built, prompts the user to enter their name at the start, this is then stored. The user is then prompted to ask a query, which is then matched to the correct intent. If the user responds with any stop phrases, the program is ended, if the user responds with any name changing phrases the user is given the chance to change their name. For

everything else, the user's input is processed, then classified to see whether it is small talk or question answering. The system will then respond with an answer related to a question, that the query is most similar to.

This is a flow chart showing the process of how things work in the program



Originality:

This is a very simple chatbot, it performs small talk, name management, question answering and intent matching. It is different from others in the dataset used for small talk, the dataset used was been bootstrapped, and contains some colloquial language. This makes the chat bot more human like and less robotic and monotonous.

Implementation

The first thing I did was to read in the CSV files which were a QnA, small talk and name datasets. The second thing I did was to create a labels list which contains a list of labels to show whether it's question or small talk. After this I pre processed the data, this includes stemming, removing stop-words, and tokenising. This was done through using `CountVectoriser` which is a function which does all of this at once. Following this I then weighted the `count_vect` and transformed it to a TF-IDF. I then set values to equal a `test_train_split` function. Next, I used the `count_vect` and TF-IDF to be applied to the

X_train. I used these functions to prepare a Bag of words model, and then train the logistic regression algorithm. This was then used for classification later, to know whether a query, is a question or small talk. I then implemented blocks of 'if' statements, to check whether the user wants to end the program, ask for their own name or change it. After this an else statement was used to process the query, then I checked using 'if' statements, whether the system predicted, whether the query was small talk or question answering. Once classified, the process for returning answers for small talk and question answering were the same. I created a function which created another TF-IDF for which ever document input into it. The query is then compared to the TF-IDF created from the specific document using a similarity function, which computes the cosine similarity, and creates a list of similarities that are linked to their specific question that are over a certain similarity level. This list is then sorted, and used to match with the question in the documents to output the specific answer from that row. If the similarity level is not met, then an empty list would be returned. This would lead to a response of "Sorry I don't understand".

I didn't use a classifier for name management, this is because the data size for small talk and question answering was much much larger than the name management one. This resulted in the name management being misclassified. Instead I used if statements, whilst telling the user of the chatbot, the specific words to say to change their name or to show their name. I researched this method, which was to use any in an if statement which would check a sentence in query to see if it was in any of the name sentence string and conduct identity management from there.

Evaluation:

I used a set of 12 queries, to test if the query had been correctly classified, and then correctly mapped to the correct answer. There are a mixture of queries to show the breadth of the system. This table shows the queries, the expected answers, the actual answer and the evaluation of it.

Test Query	Expected Answer(s)	Actual answer	Evaluation
how are you (Small talk)	1) Im fine today 2) I'm okay actually 3) Im good 4) Its okay	Its okay	Correct answer
who is isaac newton (Question Answering)	Sir Isaac Newton (25 December 164220 March 1727) was an English physicist and mathematician who is widely regarded as one of the most influential scientists of all time and as a key figure in the scientific revolution .	Sir Isaac Newton (25 December 164220 March 1727) was an English physicist and mathematician who is widely regarded as one of the most influential scientists of all time and as a key figure in the scientific revolution.	Correct answer
hello (Small talk)	Hi	Hi	Correct answer

Test Query	Expected Answer(s)	Actual answer	Evaluation
How African americans were immigrated (Question Answering)	As such, African immigrants are to be distinguished from African American people, the latter of whom are descendants of mostly West and Central Africans who were involuntarily brought to the United States by means of the historic Atlantic slave trade.	As such, African immigrants are to be distinguished from African American people, the latter of whom are descendants of mostly West and Central Africans who were involuntarily brought to the United States by means of the historic Atlantic slave trade .	Correct answer
helo (Small talk)	Hi bro	Hi bro	Correct answer
wassup (Small talk)	I'm cool bro	KeyError: 'Wassup bro'	Error, didn't classify the question correctly, actual small talk pattern in dataset was 'wassup bro'
African americans (Question Answering)	As such, African immigrants are to be distinguished from African American people, the latter of whom are descendants of mostly West and Central Africans who were involuntarily brought to the United States by means of the historic Atlantic slave trade.	As such, African immigrants are to be distinguished from African American people, the latter of whom are descendants of mostly West and Central Africans who were involuntarily brought to the United States by means of the historic Atlantic slave trade.	Correct answer
Universal law of gravity (Question Answering)	Newton's law of universal gravitation states that every point mass in the universe attracts every other point mass with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them.	Newton's law of universal gravitation states that every point mass in the universe attracts every other point mass with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them.	Correct answer

Test Query	Expected Answer(s)	Actual answer	Evaluation
How many muslims live in UK (Question Answering)	The vast majority of Muslims in the United Kingdom live in England and Wales : of 1,591,000 Muslims recorded at the 2001 Census, 1,536,015 were living in England and Wales , where they formed 3% of the population in 2001; 42,557 were living in Scotland , forming 0.84% of the population; and 1,943 were living in Northern Ireland.	The vast majority of Muslims in the United Kingdom live in England and Wales : of 1,591,000 Muslims recorded at the 2001 Census, 1,536,015 were living in England and Wales , where they formed 3% of the population in 2001; 42,557 were living in Scotland , forming 0.84% of the population; and 1,943 were living in Northern Ireland.	Correct answer
How many games on the xbox (Question Answering)	There are currently 952 games (multiplatform: 751; exclusive: 123; console exclusive: 78) on this list as of April 2, 2013.	There are currently 952 games (multiplatform: 751; exclusive: 123; console exclusive: 78) on this list as of April 2, 2013.	Correct answer
Have you seen the news (Small talk)	Yeah interesting developments	Yeah interesting developments	Correct answer
Who am I	{name input}	{name input}	Correct answer

Discussion:

```
[[221  0]
 [ 23 51]]
0.9220338983050848
0.9505376344086022
```

For prediction of 'questions', in the order of confusion matrix, accuracy score and f1score

```
[[221  0]
 [ 14 60]]
0.9525423728813559
0.8955223880597014
```

For prediction of 'smallTalk', in the order of confusion matrix, accuracy score and f1score

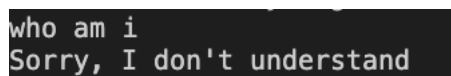
This shows that the classifier is highly accurate, when trying to classify whether a query is a question or whether it is small talk. The accuracy for questions is lower than the accuracy for small talk. However, this doesn't mean that small talk predictions are better classified than questions predictions. The f1score punishes algorithms which favour certain classes over the other, this algorithm shows that predictions for questions is much higher than the one for small talk, hence showing that the predictions for questions is much higher. I noticed whilst testing my chatbot, that questions were answered most of the time, especially when key words were added. This was also true when testing small talk, however, sometimes, my chatbot for certain phrases such as 'wassup bro', would mistakenly classify it as a 'question' and would end up outputting, 'sorry, I don't understand'.

I believe the accuracy is not as high as it should be for small talk, as a result of small talk having a smaller dataset than questions. There is a positive correlation to the size of the dataset and the accuracy and f1 score,

I believe that this chatbot could be deployed into the real world with some small tweaks. These tweaks include using a larger dataset for question and answering, and also a larger dataset for small talk. The chatbot wouldn't do well with avoiding bias, because the chatbot is as biased as the data that is fed into it. For question answering, depending on the dataset fed into it, a biased answer may appear as the query may in itself be related to a biased answer. With small talk, biases are also a thing, Depending on responses for certain small talk, biases can occur. Also, my chat bot has a small minute tendency of misclassifying data, in the real world this could be costly, giving the users of the wrong intent or the wrong response to their query.

However, this chatbot is very simple, it does not have conversation with the user, it cannot carry out a conversation and respond with unrelated new topics. It is chatbot which just responds to small talk of the user. It's question answering functionality is good enough, however it could be better, the chatbots' question answering is as good as the dataset that it is given. The chatbot is based on classification and some similarity matching, it doesn't fair well with spelling errors, it only responds to errors which have been bootstrapped to the dataset to create the illusion that it responds to them correctly, this is shown for 'small talk'.

I manually tested changing name, it would work a majority of the time, however on the rare occasion it would be classified as small talk, as shown below.



```
who am i
Sorry, I don't understand
```

Conclusion

Overall this is a modest chatbot, it can match intent correctly most of the time, answer questions correctly most of the time, when given enough information to find the question, and conducts small talk to a decent level, it handles identity management correctly. It could be better, with more data in its dataset, so it can recognise misspelt words, names and things. Also the chat bot lacks features such as game playing and transactions, which could be implemented to create a more complete and diverse chatbot. I believe the best feature of this chatbot is question answering, as it answers questions correctly a majority of the time. With a larger dataset, I believe it would be suitable for information retrieval for deployment in the real world.

References:

- 1) <https://skywell.software/blog/how-does-siri-work-technology-and-algorithm/>
- 2) <https://www.ionos.co.uk/digitalguide/online-marketing/search-engine-marketing/information-retrieval-how-search-engines-retrieve-data/>
- 3) <https://www.forbes.com/sites/ronshevlm/2021/03/15/every-bank-needs-a-chatbot-or-two-for-its-digital-transformation/?sh=51e15b64275d>