

Chatbot Report

System Description: The program reads in a text file containing data on Michael Jordan and first processes the text. I first lowercase all the text and then tokenized the text using word tokenize and sent tokenize. These techniques turn the text into a list of words and sentences, respectively. I then removed the stop words using nltk.stopwords. This helped me to find the important terms in the user questions and match them to the text. I then lemmatized the text using the built in Wordnet lemmatizer. This was useful in finding the roots of all the words and help me get more accurate matchings of the user questions to the text. I then used the tf-idf vectorizer and the cosine similarity which I imported from sklearn to find the sentence that was the most similar to the question that the user asked. The tf-idf vectorizer is a measure of how many times a word appears in a document times the inverse document frequency of the term. After the tf-idf assigned an idx to each word, the cosine similarity was used to find the vector that was closest to the vector which represented the question that was asked. This allowed me to find the sentence in the text was closest text-wise to the question that was being asked. The user model that I built is a dict which has the key as the name of the user and the value as a list of question and answer tuples that the user asked. It is being passed through each call of the chatbot function so that the program can “remember” a previous user.

If you do not want to chat, type exit. Else, type anything else.

no

I am Michael Jordan's number one fan. Ask questions about him. Type exit to exit this conversation

What is your name?

Vedant

Ask a question!

six championships

[140] the bulls were strengthened by the addition of rebound specialist dennis rodman, and the team dominated the league, starting the season at 41â€“3.

Ask a question!

third overall pick

[52] trail blazers general manager stu inman contended that it was not a matter of drafting a center but more a matter of taking bowie over jordan, in part because portland already had clyde drexler, who was a guard with similar skills to jordan.

Ask a question!

1997 championship

jordan abruptly retired from basketball before the 1993â€“94 nba season to play minor league baseball but returned to the bulls in march 1995 and led them to three more championships in 1996, 1997, and 1998, as well as a then-record 72 regular season wins in the 1995â€“96 nba season.

Ask a question!

Did Jordan have 20 all star appearances

jordan's individual accolades and accomplishments include six nba finals most valuable player (mvp) awards, ten nba scoring titles (both all-time records), five nba mvp awards, ten all-nba

first team designations, nine all-defensive first team honors, fourteen nba all-star game selections, three nba all-star game mvp awards, three nba steals titles, and the 1988 nba defensive player of the year award.

Ask a question!

exit

Thank you for chatting with me!

If you do not want to chat, type exit. Else, type anything else.

no

I am Michael Jordan's number one fan. Ask questions about him. Type exit to exit this conversation

What is your name?

Vedant

Welcome back!

Ask a question!

Does Jordan have 20 all star appearances

jordan's individual accolades and accomplishments include six nba finals most valuable player (mvp) awards, ten nba scoring titles (both all-time records), five nba mvp awards, ten all-nba first team designations, nine all-defensive first team honors, fourteen nba all-star game selections, three nba all-star game mvp awards, three nba steals titles, and the 1988 nba defensive player of the year award.

Ask a question!

six championships

[140] the bulls were strengthened by the addition of rebound specialist dennis rodman, and the team dominated the league, starting the season at 41â€“3.

Ask a question!

six championships

You asked this already!

Ask a question!

72-9

jordan abruptly retired from basketball before the 1993â€“94 nba season to play minor league baseball but returned to the bulls in march 1995 and led them to three more championships in 1996, 1997, and 1998, as well as a then-record 72 regular season wins in the 1995â€“96 nba season.

Ask a question!

If you do not want to chat, type exit. Else, type anything else.

no

I am Michael Jordan's number one fan. Ask questions about him. Type exit to exit this conversation

What is your name?

Vedant

Ask a question!

six championships

n the 1992â€“93 season, despite a 32.6 ppg, 6.7 rpg, and 5.5 apg campaign, including a second-place finish in defensive player of the year voting,[77][97] jordan's streak of consecutive mvp seasons ended, as he lost the award to his friend charles barkley,[76] which upset him.

Ask a question!

finals john stockton

[91] he scored more than 30 points in every game of the series, including 40 or more points in four consecutive games.

Ask a question!

Did jordan have six championships with Scottie Pippen?

they advanced to the eastern conference finals where their rival, the detroit pistons, awaited them;[87] this time, the bulls beat the pistons in a four-game sweep.

Ask a question!

championships

Be more specific please.

Ask a question!

champion

Be more specific please.

Ask a question!

winning

"[10] he played fifteen seasons in the national basketball association (nba), winning six nba championships with the chicago bulls.

Ask a question!

championships

Be more specific please.

Ask a question!

You asked this already!

Ask a question!

exit

Thank you for chatting with me!

If you do not want to chat, type exit. Else, type anything else.

no

I am Michael Jordan's number one fan. Ask questions about him. Type exit to exit this conversation

What is your name?

Vedant

Welcome back!

Ask a question!

exit

Thank you for chatting with me!

If you do not want to chat, type exit. Else, type anything else.

exit

Bye.

Process finished with exit code 0

My knowledge base is a txt file which has information from the wikipedia page of michael Jordan. It is structured like:

Overview:
Early Career:
Mid-Career
Late-Career:

The overview is a summary of his career. The early career is 1990-1993. The mid career focuses on the championships and his involvement in baseball. The late career focuses on the other three championships and retirement.

My user model is a dict which has a key as the name of the user and the value as a list of question and answer tuples that the user asked.

{Vedant: [{six championships,
[140] the bulls were strengthened by the addition of rebound specialist dennis rodman, and the team dominated the league, starting the season at 41â€“3}, {third overall pick,
[52] trail blazers general manager stu inman contended that it was not a matter of drafting a center but more a matter of taking bowie over jordan, in part because portland already had clyde drexler, who was a guard with similar skills to jordan}, {1997 championship, jordan abruptly retired from basketball before the 1993â€“94 nba season to play minor league baseball but returned to the bulls in march 1995 and led them to three more championships in 1996, 1997, and 1998, as well as a then-record 72 regular season wins in the 1995â€“96 nba season}, {Did Jordan have 20 all star appearances, jordan's individual accolades and accomplishments include six nba finals most valuable player (mvp) awards, ten nba scoring titles (both all-time records), five nba mvp awards, ten all-nba first team designations, nine all-defensive first team honors, fourteen nba all-star game selections, three nba all-star game mvp awards, three nba steals titles, and the 1988 nba defensive player of the year award}}]}

The chatbot's strengths are that it is easily runnable and requires low computing power. It is rules based and does not contain large neural networks. Its weaknesses are that it is inaccurate and produce very inaccurate answers. It is not up to the standard of a modern chatbot. It produced some inaccurate answers and some which did not make any sense, which was surprising since I did use some machine learning techniques to help it find the best sentence in the text that related to the user's answers.

