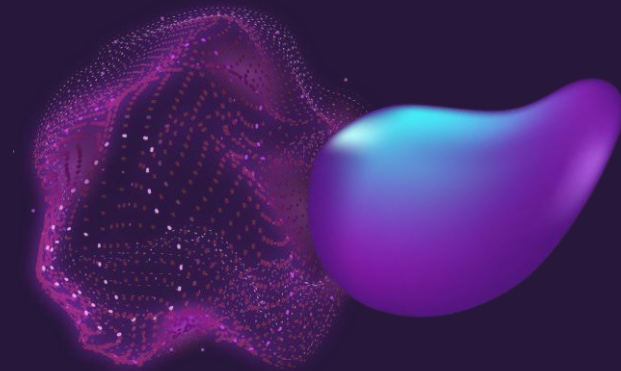




x



Conversational AI for COVID-19^x Empowering Information Exchange through Chatbots

Mariam Tmane - Undergraduate

x

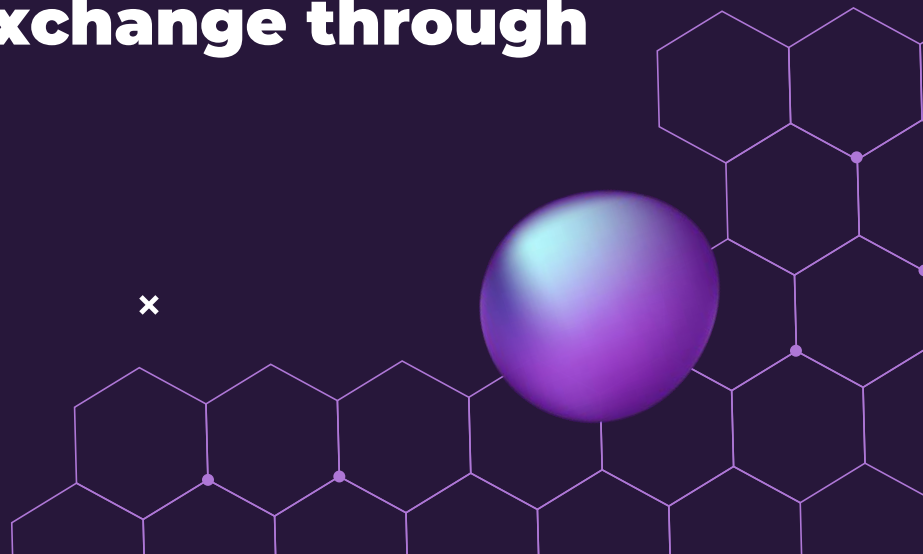
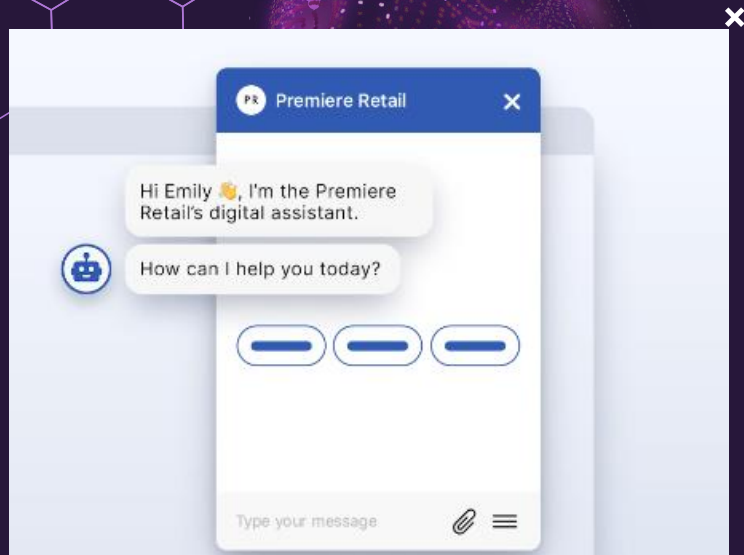


Table of contents

<u>Description</u>	A description of the project, problem formulation and purpose.
AI Agent / methodology	A detailed description of methodologies, and techniques used to develop the chatbot.
<u>Tools / libraries used</u>	Python, Numpy, SpaCy.
<u>Data Source</u>	Information about data used to for information retrieval.
<u>Conclusions / Drawbacks</u>	Conclusions, and results.
Future plans	Future plans and improvement.



COVID-19 Chatbot:

COVID-19 conversational Chatbot, that answers common questions about COVID-19, such as symptoms, information about vaccines, and more.

PROJECT PLAN

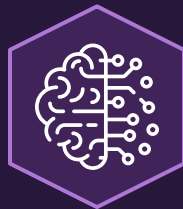
×



Research

Reading artificial intelligence
Research papers.
Papers with code,
kaggle.

+



Design

Learned about existing
models and methodologies,
and designed my
method/solution



×

Implement

Chose libraries, coding
language and start
writing code to
implement the logic of
the proposed solution



Papers With Code

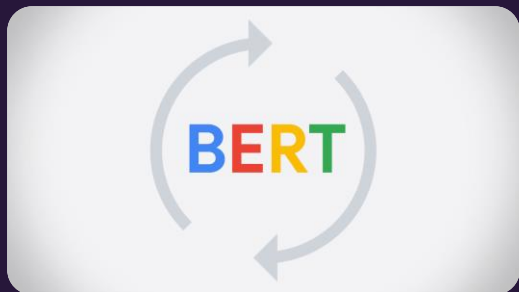
Current Methods

×



Bert Model

Bidirectional Encoder Representations from Transformers. BERT is pre-trained on a large corpus of text data. Can be further fine-tuned



Tf-Idf

Term Frequency-Inverse Document Frequency, It's a statistical measure used in information retrieval and text mining to evaluate the importance of a word in a document relative to a collection of documents.



Technologies Used

×

Python

Language used to implement the program, using Visual Studio Code

spaCy

A powerful and popular open-source library for natural language processing (NLP), provides word embeddings as part of its functionality. Word embeddings are dense vector representations of words, that capture semantic meaning.

```
chatbot.py > ...  
1  import spacy  
2  import json  
3  import numpy as np  
4  import nltk  
5  from nltk.tokenize import word_tokenize  
6  from nltk.corpus import stopwords  
7
```

Implementation

1. ^x Word Embeddings(spacy English model) to vectorize questions and user input.
2. Cosine Similarity(.similarity()), to compare the similarity and find the most similar questions

```
✓ def get_most_similar_question(user_question):  
    #convert user input and questions to vector using spacy  
    questions_to_vector = vectorize_questions(training_questions)  
    input_to_vector = feed_model(user_question)
```

Data Source

×

The data used was carefully gathered from the World Health Organization website, UNICEF, and Centers for Disease Control and Prevention. It was then organized into JSON files that contain popular questions about COVID-19 and their answers.

```
{ } questions.json > [ ] questions
1  {
2    "questions": [
3      {
4        "question": "How does COVID-19 spread?",
5        "answer": "COVID-19 primarily spreads through respiratory droplets when an infected person coughs, sneezes, or talks. It can also spread through contact with contaminated surfaces."
6      },
7      {
8        "question": "What should I do if I think I have COVID-19?",
9        "answer": "If you have symptoms of COVID-19 or have been in close contact with someone who has COVID-19, you should stay home, isolate yourself, and seek medical attention."
10     },
11     {
12       "question": "How can I protect myself and others from COVID-19?",
13       "answer": "You can protect yourself and others by wearing masks, practicing good hand hygiene, maintaining physical distance from others, and avoiding crowded places."
14     },
15     {
16       "question": "What are the symptoms of COVID-19?",
17       "answer": "The most common symptoms of COVID-19 are fever, cough, and shortness of breath. Other symptoms include fatigue, muscle aches, and loss of taste or smell."
18     }
19   ]
20 }
```


Example of chatbot performance

```
chatbot.py > ...
1 import spacy
2 import json
3 import numpy as np
4 import nltk

OUTPUT TERMINAL PORTS PROBLEMS 5 DEBUG CONSOLE
PS C:\Users\masoh\Desktop\AI_covid_chatbot> python chatbot.py
[nltk_data] Downloading package punkt to
[nltk_data] C:\Users\masoh\AppData\Roaming\nltk_data...
[nltk_data] Package punkt is already up-to-date!
[nltk_data] Downloading package stopwords to
[nltk_data] C:\Users\masoh\AppData\Roaming\nltk_data...
[nltk_data] Package stopwords is already up-to-date!
Chatbot: Hi! I'm the chatbot. You can start the conversation by typing your message.
User: hi
Chatbot: Hello! What can I help you with?
User: what are the common symptoms of covid
Chatbot: Symptoms of COVID-19 include Fever or chills, cough, shortness of reath or difficulty breathing, fatigue, muscle or body aches, headache, sore throat, loss of taste or smeel, congestion or runny nose, and nausea or vomiting
User: what does it mean if i have a fever
Chatbot: Fever is a common symptom of COVID-19. It indicates that your body is responding to an infection, such as the coronavirus. It's important to monitor your temperature and seek medical advice if you have a fever.
User: how long is the incubation period
Chatbot: Review of studies conducted from 2020 to March 2022 indicated an average incubation period for COVID-19 at six to seven days, ranging from 1.8 to 18.87 days.
User: what is the difference between the flu and covid
Chatbot: Influenza (Flu) and COVID-19 are both contagious respiratory illnesses, but they are caused by different viruses. You cannot tell the difference between flu and COVID-19 by symptoms alone because some of the symptoms are the same. Some PCR tests can differentiate between flu and COVID-19 at the same time.
User: when can i go back to normal activities
Chatbot: You can go back to normal activities if for more that 24 hours, your syptoms are getting better overall, and you have not had a fever (and are not using fever-reducing medication). When you go back to your normal activities, take added precaution over the next 5 days, Keep in mind that you may still be able to spread the virus that made you sick, even if you are feeling better. You are likely to be less contagious at this time, depending on factors like how long you were sick or how sick you were.
User: bye
Chatbot: Goodbye!
PS C:\Users\masoh\Desktop\AI_covid_chatbot> 
```

Testing and Evaluation

To test the chatbot, I created a validation dataset see how well the chatbot can compare questions and retrieve the correct answer.

Challenges:

- ❑ Making the dataset used by the chatbot contain a wide variety of questions about COVID-19.
- ❑ Making sure the questions are clear and coherent with the answers.

```
{ } validation_data.json > [ ] questions > { } 2
```

```
1  {
2    "questions": [
3      {
4        "question": "What are the different types of COVID-19 tests?",
5        "answer": "There are several types of COVID-19 tests, including PCR tests, antigen tests, and antibody tests."
6      },
7      {
8        "question": "How long does it take to get COVID-19 test results?",
9        "answer": "The time it takes to get COVID-19 test results can vary depending on the type of test and the testing location. Some test
10      },
11      {
12        "question": "Can I get COVID-19 from food?]",
13        "answer": "There is currently no evidence to suggest that COVID-19 can be transmitted through food. However, it's important to pract
14      },
15      {
16        "question": "What should I do if I've been exposed to someone with COVID-19?",
17        "answer": "If you've been exposed to someone with COVID-19, you should quarantine for a period of time as recommended by health auth
18      },
19      {
20        "question": "How effective are face masks at preventing the spread of COVID-19?",
21        "answer": "Face masks are highly effective at reducing the spread of COVID-19, especially when combined with other preventive measur
22      }
23    ]
24  }
```

Example of output after running program on validation data

OUTPUT TERMINAL PORTS PROBLEMS **5** DEBUG CONSOLE

powershell + [icon] [icon] ... ^ x

```
PS C:\Users\masoh\Desktop\AI_covid_chatbot> python testing_chatbot.py
```

```
[nltk data] Downloading package punkt to
```

```
[nltk data] C:\Users\masoh\AppData\Roaming\nltk data...
```

```
[nltk data] Package punkt is already up-to-date!
```

```
[nltk data] Downloading package stopwords to
```

```
[nltk data] C:\Users\masoh\AppData\Roaming\nltk data...
```

```
[nltk data] Package stopwords is already up-to-date!
```

```
[nltk data] Package stopwords is already up-to-date!
```

Chatbot: Hi! I'm the chatbot. You can start the conversation by typing your message.

User Question: --> What about testing

----- Most Similar Question Found: --> what do I need to know about testing?

User Question: --> what are some symptoms of covid-19

----- Most Similar Question Found: --> What are the symptoms of COVID-19?

User Question: --> do you think the vaccine is safe

----- Most Similar Question Found: --> is the vaccine safe and effective?

User Question: --> does wearing masks prevent from

----- Most Similar Question Found: --> How effective are cloth masks in preventing COVID-19?

User Question: --> what if i have a cough

```
----- Most Similar Question Found: --> does it mean i have covid if i have a cough
```

User Question: --> how do i know when to

----- Most Similar Question Found: --> When can I go back to normal activities?

User Question: --> what if i lose my sense of taste

----- Most Similar Question Found: --> What does it mean if I lose my sense of

User Question: --> can covid be transmitted to pets

----- Most Similar Question Found: --> Can COVID-19 be transmitted through pets?

User Question: --> do children get covid-19

----- Most Similar Question Found: --> do children get severe illness if they ge

User Question: --> what if i am vomiting?

```
----- Most Similar Question Found: --> What does it mean if I am vomiting?
```

Overall score on validation data 10 /10

```
PS C:\Users\masoh\Desktop\AI_covid_chatbot>
```




RESOURCES

1. D. Oniani and Y. Wang, "A qualitative Evaluation of Language Models on Automatic Question-Answering for COVID-19," 2020. Accessed: Apr. 21, 2024. [Online]. Available: <https://arxiv.org/pdf/2006.10964v2.pdf>
2. T. Zhang, V. Kishore, F. Wu, K. Q. Weinberger, and Y. Artzi, "BERTScore: Evaluating Text Generation with BERT," arXiv preprint arXiv:2006.10964, 2020. [Online]. Available: <https://arxiv.org/abs/2006.10964>
3. [1]"COVID-19: Frequently asked questions | UNICEF," www.unicef.org.
<https://www.unicef.org/turkiye/en/stories/covid-19-frequently-asked-questions>
4. S. Wang, S. Yu, and C. Zhang, "A Novel Approach to Sentiment Analysis using BERT," IEEE Transactions on Knowledge and Data Engineering, vol. 33, no. 5, pp. 1234-1245, May 2022. [Online]. Available: <https://arxiv.org/pdf/2001.09896v2.pdf>
5. A. Jalilifard, V. Fernandes Caridá, A. Fernandes Mansano, R. Cristo, and F. Carvalho Da Fonseca, "Semantic Sensitive TF-IDF to Determine Word Relevance in Documents," Feb. 2021. Accessed: Apr. 21, 2024. [Online]. Available: <https://arxiv.org/pdf/2001.09896v2.pdf>
6. D. Oniani, "oniani/covid-19-chatbot," *GitHub*, May 08, 2023. <https://github.com/oniani/covid-19-chatbot> (accessed Apr. 21, 2024).