Name: Ankita Yadav

Question Answering System Using COVID-19 Article

Compare the answers you find with the "ground-truth" answers in the json file and analyze the following:

- **What kind of questions can be correctly answered by your system? What kind of questions CANNOT be correctly answered by your system?**

- **Why does your system fail to locate the right answers to these questions?**

- **How should your system be improved so that such questions can be answered?**

Analysis 1:

What kind of questions can be correctly answered by your system? What kind of questions CANNOT be correctly answered by your system?

Answer:

Questions that have been answered correctly with the system are the questions where the answers are direct in the complete sentence. For example, the whole sentence serves as an answer for the question.

As we can see in question 3, "How many states in the U.S. have reported cases of COVID-19?" The answer to the question captured by our system is "All 50 states have reported cases of COVID-19 to CDC". So, if we read the statement, we can make out the answer is 50 which matches the "ground-truth" answer provided in the json file.

However, question which are bit tricky like question 7, "What viruses are similar to the COVID-19 coronavirus?" are not answered correctly. In this case the answer given by our system is "COVID-Emergence COVID-19 is caused by a coronavirus". The correct answer was MERS-CoV and SARS-CoV. The correct answer was not in a form of complete statement however it is a text extracted from the sentence.

Questions	Our System	Status of Answer by Our System
Q.1.	Answered correctly	This question was answered in the form of complete sentence which was able to match the right answer
Q.2.	Answered incorrectly	The answer expected was in for of few words which was not been able to captured correctly by our system
Q.3.	Answered correctly	This question was answered in the form of complete sentence which was able to match the right answer
Q.4.	Answered correctly	This question was answered in the form of complete sentence which was able to match the right answer
Q.5.	Answered correctly	This question was answered in the form of complete sentence which was able to match the right answer
Q.6.	Answered correctly	This question was answered in the form of complete sentence which was able to match the right answer
Q.7.	Answered incorrectly	The answer expected was in form of few words which was not been able to captured correctly by our system
Q.8.	Answered correctly	This question was answered in the form of complete sentence which was able to match the right answer
Q.9.	Answered correctly	This question was answered in the form of complete sentence which was able to match the right answer
Q.10.	Answered semi correctly	This question was answered in the form of complete sentence. There can be more sentence that should have been added to provide complete right answer. For example, the right answer was "People who have serious chronic medical conditions like heart disease, diabetes, lung disease". The answer provided by our system just mentions about the severe chronic conditions but does not mention the type of chronic diseases such as heart, diabetes, and luncg.
Q.11.	Answered correctly	This question was answered in the form of complete sentence which was able to match the right answer
Q.12.	Answered incorrectly	The answer expected was in form of one word which was not been able to captured correctly by our system
Q.13.	Answered semi correctly	This question was answered in the form of complete sentence. There can be more sentence that should have been added to provide complete right answer. Only one statement was provided as an answer however the correct answer compromised of multiple statements.

Analysis 2:

Why does your system fail to locate the right answers to these questions?

Answer:

The system developed fails to locate the right answers to the questions because our system focusses on single sentence which acts as one document. Hence the answer provided by our system is in the form of sentences. For question which have answer in the form of one single statement have been answered correctly however our system has not been able to answer those questions correctly wherein the answer is in the form of few words which need to be extracted from these statements which acts as a multiple single document. There is a failure in locating the correct answer because our system has not been able to extract features, words, entities from the statement.

Analysis 3:

How should your system be improved so that such questions can be answered?

Answer:

Our system can be improved so that the answers are provided correctly through the following ways:

- Name entity recognition shall help us to extract information. It shall provide clear extract, factual information correctly. We can also index the named entities and their reactions can be extracted. For the question answering system, answers are usually entities and thus it shall be able to correctly capture the answer.
- Dependency graph shall help in parsing the sentence and provide relevant answer.

By using the above two techniques the system shall be able to provide the answers to the questions correctly.