Annotation Guidelines for COVID-19 Chatbot

David Oniani

Mayo Clinic

mailto:oniani.david@mayo.edu

Dr. Yanshan Wang

Mayo Clinic

mailto:wang.yanshan@mayo.edu

April 2, 2020

Table of Contents

1 Answer Categorization

Answers will be categorized according to 5 categories:

- Relevant – the answer partially or fully answers the question and/or makes clear attempts to do so and is related to the question. (5 points)

- Well-formed – the answer makes a logical sense and is somewhat related to both the question and COVID-19, yet it does not (partially or fully) answer the question. (4 points)

- Informative – the answer is not related to the question, but provides some information about COVID-19 and makes a logical sense. (3 points)

- Acceptable – the answer makes some logical sense and is weakly related to the question or COVID-19, but is mostly difficult to understand. (2 points)

- Poor – the answer is totally unrelated to the question or COVID-19 and/or does not make a logical sense. (1 point)

2 Example Answers and Evaluation

The answers below are made-up and are not present in the actual evaluation data.

2.1 Relevant (5 Points)

Question: Are there geographic variations in the mortality rate of COVID-19?
Answer: *The mortality rate of COVID-19 depends on the geography a lot and some places in the world might be more vulnerable to it than others. It is also important to consider other causes. I love ice cream!*

Explanation: The answer would be rated 5 since it is clearly related to the question. Although the answer mentions unrelated bits like "I love ice cream", it is mostly relevant to the question.

2.2 Well-formed (4 Points)

Question: Are there geographic variations in the mortality rate of COVID-19?
Answer: *Geography is a great subject! You should learn it! Also, COVID-19 has been tough geographically. It is terrible if you get it.*

Explanation: The answer is somewhat related to the question since it mentions both the geography and COVID-19, however, it does not answer the question.

2.3 Informative (3 Points)

Question: Are there geographic variations in the mortality rate of COVID-19?

Answer: *COVID-19 or Novel coronavirus is a threat to human health. It is bad if you get it so be safe. Some scientists think that it spreads faster than Influenza.*

Explanation: Although the text is not directly related the question, it provides some information about COVID-19 and makes a logical sense.

2.4 Acceptable (2 Point)

Question: Are there geographic variations in the mortality rate of COVID-19?

Answer: *Some disease are viral. There has been viral diseases recently. It can be damaging to people.*

Explanation: The answer is weakly related to COVID-19 since it mentions terms "disease" and "viral." Yet, it fails to deliver the answer related to COVID-19 or the question.

2.5 Poor (1 Point)

Question: Are there geographic variations in the mortality rate of COVID-19?

Answer: *To be, or not to be, that is the question???! What is life? What is Earth? What is human? Soma animals are carnivores and some are herbivores, and some omnivores.*

Explanation: The answer is totally unrelated to the question or COVID-19.