

Report: results from five models on CORD-19 dataset (version 2020-03-13) [\[code\]](#)

1st model: multi-qa-mpnet-base-dot-v1

Maximum Sequence Length: 512

Speed (sentences/sec): 2800

Run time: 1 hour on Tesla P100 GPU

QUESTION: What are the coronaviruses?

FROM PAPER WITH TITLE: Natural Bis-Benzylisoquinoline Alkaloids-Tetrandrine, Fangchinoline, and Cepharanthine, Inhibit Human Coronavirus OC43 Infection of MRC-5 Human Lung Cells

PASSAGE ANSWER: Coronaviruses (CoVs) are enveloped, positive-sense, single-stranded RNA viruses that infect a broad range of animal species and cause multiple respiratory outcomes of varying severity, including the common cold, bronchiolitis, and pneumonia [13] .

SCORE: 25.92

QUESTION: What was discovered in Wuhuan in December 2019?

FROM PAPER WITH TITLE: Identification of a Novel Polyomavirus from Patients with Acute Respiratory Tract Infections

PASSAGE ANSWER: This discovery raises many questions for further investigation, such as, Is WU virus a human pathogen?

SCORE: 25.88

QUESTION: What is Coronavirus Disease 2019?

FROM PAPER WITH TITLE: Systematic Comparison of Two Animal-to-Human Transmitted Human Coronaviruses: SARS-CoV-2 and SARS-CoV

PASSAGE ANSWER: This virus causes acute lung symptoms, leading to a condition that has been named as "coronavirus disease 2019" (COVID-19).

SCORE: 27.30

QUESTION: What is COVID-19?

FROM PAPER WITH TITLE: Systematic Comparison of Two Animal-to-Human Transmitted Human Coronaviruses: SARS-CoV-2 and SARS-CoV

PASSAGE ANSWER: This virus causes acute lung symptoms, leading to a condition that has been named as "coronavirus disease 2019" (COVID-19).

SCORE: 29.70

QUESTION: What is caused by SARS-COV2?

FROM PAPER WITH TITLE: Viral Mimicry to Usurp Ubiquitin and SUMO Host Pathways

PASSAGE ANSWER: SARS CoV is the causative agent for severe acute respiratory syndrome, which frequently leads to outbreaks and high mortality.

SCORE: 29.39

QUESTION: How is COVID-19 spread?

FROM PAPER WITH TITLE: Rapid communication

PASSAGE ANSWER: The outbreak has spread rapidly, affecting other parts of China, and cases have been recorded on several continents (Asia, Australia, Europe and North America); further global spread is likely to occur [4] .The spectrum of this disease in humans, now named coronavirus disease 2019 (COVID-19) [5] , is yet to be fully determined.

SCORE: 29.02

QUESTION: Where was COVID-19 discovered?

FROM PAPER WITH TITLE: Clinical Medicine Optimization Method for Forecasting Confirmed Cases of COVID-19 in China

PASSAGE ANSWER: In December 2019, a novel coronavirus, called COVID-19, was discovered in Wuhan, China, and has spread to different cities in China as well as to 24 other countries.

SCORE: 28.08

QUESTION: How many deaths were caused by COVID-19?

FROM PAPER WITH TITLE: Identification of COVID-19 Can be Quicker through Artificial Intelligence framework using a Mobile Phone-Based Survey in the Populations when Cities/Towns Are Under Quarantine

PASSAGE ANSWER: As of February 25 th 2020, the World Health Organization's situational data indicates that there were about 77780 confirmed cases, including 2666 deaths due to COVID-19, including cases in 25 countries [4].

SCORE: 29.09

QUESTION: Does wearing a mask prevent the spread of COVID-19?

FROM PAPER WITH TITLE: Transmission of Influenza A in a Student Office Based on Realistic Person-to-Person Contact and Surface Touch Behaviour

PASSAGE ANSWER: Wearing a mask can control the spread of disease via the long-range airborne, fomite and close contact routes.

SCORE: 27.84

2nd model: all-mpnet-base-v2

Maximum Sequence Length: 384

Speed (sentences/sec): 2800

Run time: 1 hour on Tesla P100 GPU

QUESTION: What are the coronaviruses?

FROM PAPER WITH TITLE: A novel coronavirus capable of lethal human infections: an emerging picture

PASSAGE ANSWER: Coronaviruses are a large family of enveloped, single-stranded RNA viruses that infect a number of different species, including humans.

SCORE: 0.81

QUESTION: What was discovered in Wuhuan in December 2019?

FROM PAPER WITH TITLE: The novel coronavirus outbreak in Wuhan, China

PASSAGE ANSWER: On December 31, 2019, delegates of the Chinese Center for Disease Control and Prevention (CDC) went to Wuhan for field investigations, and the sample of new virus was isolated and further identified as a pathogen of unexplained pneumonia on January 6, 2020.

SCORE: 0.59

QUESTION: What is Coronavirus Disease 2019?

FROM PAPER WITH TITLE: viruses Return of the Coronavirus: 2019-nCoV

PASSAGE ANSWER: The novel CoV (2019-nCoV) was isolated from a single patient and subsequently verified in 16 additional patients [3] .

SCORE: 0.72

QUESTION: What is COVID-19?

FROM PAPER WITH TITLE: Comment

PASSAGE ANSWER: The ongoing coronavirus disease 2019 (COVID-19) outbreak is giving rise to worldwide anxieties, rumours, and online misinformation.

SCORE: 0.73

QUESTION: What is caused by SARS-COV2?

FROM PAPER WITH TITLE: Viral Mimicry to Usurp Ubiquitin and SUMO Host Pathways

PASSAGE ANSWER: SARS CoV is the causative agent for severe acute respiratory syndrome, which frequently leads to outbreaks and high mortality.

SCORE: 0.84

QUESTION: How is COVID-19 spread?

FROM PAPER WITH TITLE: Survey and Visual Detection of Zaire ebolavirus in Clinical Samples Targeting the Nucleoprotein Gene in Sierra Leone

PASSAGE ANSWER: It spread through direct contact with infected individuals, or by touching the blood, organs, bodily secretions and fluids, or the contaminated clothes of such people (MacNeil and Rollin, 2012) .

SCORE: 0.77

QUESTION: Where was COVID-19 discovered?

FROM PAPER WITH TITLE: Clinical Medicine Optimization Method for Forecasting Confirmed Cases of COVID-19 in China

PASSAGE ANSWER: In December 2019, a novel coronavirus, called COVID-19, was discovered in Wuhan, China, and has spread to different cities in China as well as to 24 other countries.

SCORE: 0.74

QUESTION: How many deaths were caused by COVID-19?

FROM PAPER WITH TITLE: Identification of COVID-19 Can be Quicker through Artificial Intelligence framework using a Mobile Phone-Based Survey in the Populations when Cities/Towns Are Under Quarantine

PASSAGE ANSWER: As of February 25 th 2020, the World Health Organization's situational data indicates that there were about 77780 confirmed cases, including 2666 deaths due to COVID-19, including cases in 25 countries [4] .

SCORE: 0.83

QUESTION: Does wearing a mask prevent the spread of COVID-19?

FROM PAPER WITH TITLE: A randomised clinical trial to evaluate the safety, fit, comfort of a novel N95 mask in children

PASSAGE ANSWER: Studies have established the benefits of masks as a non-pharmaceutical modality that serves as a barrier between the virus and the subjects thus breaking the transmission of virus into human host 38 pandemics or disease outbreaks, vaccines may not be available or able to effect its benefits immediately.

SCORE: 0.77

3rd model: all-roberta-large-v1
Maximum Sequence Length: 256
Speed (sentences/sec): 800
Run time: 174 minutes (~3 hours) on Tesla P100 GPU

QUESTION: What are the coronaviruses?

FROM PAPER WITH TITLE: Structure analysis of the receptor binding of 2019-nCoV

PASSAGE ANSWER: They can be classified into four genera: Alpha-, Beta-, Gamma-, and Delta coronavirus [2, 3] .

SCORE: 0.74

QUESTION: What was discovered in Wuhuan in December 2019?

FROM PAPER WITH TITLE: Clinical Medicine Characteristics of and Public Health Responses to the Coronavirus Disease 2019 Outbreak in China

PASSAGE ANSWER: In December 2019In December 2019, a cluster of pneumonia of unknown etiology was detected in Wuhan City, Hubei Province of China.

SCORE: 0.65

QUESTION: What is Coronavirus Disease 2019?

FROM PAPER WITH TITLE: Systematic Comparison of Two Animal-to-Human Transmitted Human Coronaviruses: SARS-CoV-2 and SARS-CoV

PASSAGE ANSWER: This virus causes acute lung symptoms, leading to a condition that has been named as "coronavirus disease 2019" (COVID-19).

SCORE: 0.74

QUESTION: What is COVID-19?

FROM PAPER WITH TITLE: Structure analysis of the receptor binding of 2019-nCoV

PASSAGE ANSWER: The pathogen responsible for this epidemic is a new coronavirus designated 2019-nCoV by the World Health Organization.

SCORE: 0.76

QUESTION: What is caused by SARS-COV2?

FROM PAPER WITH TITLE: SARS-Coronavirus Open Reading Frame-8b triggers intracellular stress pathways and activates NLRP3 inflammasomes

PASSAGE ANSWER: The SARS (severe acute respiratory syndrome) outbreak was caused by a coronavirus (CoV) named the SARS-CoV.

SCORE: 0.79

QUESTION: How is COVID-19 spread?

FROM PAPER WITH TITLE: pathogens Emergence of Novel Coronavirus 2019-nCoV: Need for Rapid Vaccine and Biologics Development

PASSAGE ANSWER: The rapid spread of this infection is frightening and also it causes both mortality and financial loss, which present the global concern of this emerging disease.The transmission of 2019-nCoV is often spread from person to person through the respiratory droplets generated during coughs or sneezes from an infected person.

SCORE: 0.74

QUESTION: Where was COVID-19 discovered?

FROM PAPER WITH TITLE: Clinical Medicine Optimization Method for Forecasting Confirmed Cases of COVID-19 in China

PASSAGE ANSWER: In December 2019, a novel coronavirus, called COVID-19, was discovered in Wuhan, China, and has spread to different cities in China as well as to 24 other countries.

SCORE: 0.79

QUESTION: How many deaths were caused by COVID-19?

FROM PAPER WITH TITLE: Identification of COVID-19 Can be Quicker through Artificial Intelligence framework using a Mobile Phone-Based Survey in the Populations when Cities/Towns Are Under Quarantine

PASSAGE ANSWER: As of February 25 th 2020, the World Health Organization's situational data indicates that there were about 77780 confirmed cases, including 2666 deaths due to COVID-19, including cases in 25 countries [4] .

SCORE: 0.80

QUESTION: Does wearing a mask prevent the spread of COVID-19?

FROM PAPER WITH TITLE: Professional and Home-Made Face Masks Reduce Exposure to Respiratory Infections among the General Population

PASSAGE ANSWER: As part of pandemic preparedness, many are contemplating the contribution wide-spread use of masks could have [8, 9] .

SCORE: 0.79

4th model: msmarco-bert-base-dot-v5

Maximum Sequence Length: 512

Speed (sentences/sec): 2800

Run time: 59 minutes on Tesla P100 GPU

QUESTION: What are the coronaviruses?

FROM PAPER WITH TITLE: Biological Crystallography Structure of the C-terminal domain of nsp4 from feline coronavirus

PASSAGE ANSWER: Coronaviruses are a family of positive-stranded RNA viruses that includes important pathogens of humans and other animals.

SCORE: 173.90

QUESTION: What was discovered in Wuhuan in December 2019?

FROM PAPER WITH TITLE: Clinical Medicine Characteristics of and Public Health Responses to the Coronavirus Disease 2019 Outbreak in China

PASSAGE ANSWER: In December 2019, a cluster of pneumonia of unknown etiology was detected in Wuhan City, Hubei Province of China.

SCORE: 169.46

QUESTION: What is Coronavirus Disease 2019?

FROM PAPER WITH TITLE: Clinical Medicine Characteristics of and Public Health Responses to the Coronavirus Disease 2019 Outbreak in China

PASSAGE ANSWER: Thereafter, this disease was named Coronavirus Disease 2019 (COVID-19) by World Health Organization (WHO), and the causative virus was designated as SARS-CoV-2 by the International Committee on Taxonomy of Viruses [6] .

SCORE: 170.87

QUESTION: What is COVID-19?

FROM PAPER WITH TITLE: Consensus statement The species Severe acute respiratory syndrome- related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2 Coronaviridae Study Group of the International Committee on Taxonomy of Viruses*

PASSAGE ANSWER: The present outbreak of a coronavirus-associated acute respiratory disease called coronavirus disease 19 (COVID-19) is the third documented spillover of an animal coronavirus to humans in only two decades that has resulted in a major epidemic.

SCORE: 167.57

QUESTION: What is caused by SARS-COV2?

FROM PAPER WITH TITLE: Virology Journal G0/G1 arrest and apoptosis induced by SARS-CoV 3b protein in transfected cells

PASSAGE ANSWER: SARS was found to be caused by a novel coronavirus which was designated as SARS coronavirus (SARS-CoV), and the genome of SARS-CoV contains 11 to 14 open reading frames (ORF) and 5 to 8 potential nonstructural proteins [4, 5] .

SCORE: 171.40

QUESTION: How is COVID-19 spread?

FROM PAPER WITH TITLE: Systematic Comparison of Two Animal-to-Human Transmitted Human Coronaviruses: SARS-CoV-2 and SARS-CoV

PASSAGE ANSWER: At present, the COVID-19 outbreak has been spread to all parts of China and around the world, including the United States, Thailand, and Japan.

SCORE: 170.10

QUESTION: Where was COVID-19 discovered?

FROM PAPER WITH TITLE: Clinical Medicine Optimization Method for Forecasting Confirmed Cases of COVID-19 in China

PASSAGE ANSWER: In December 2019, a novel coronavirus, called COVID-19, was discovered in Wuhan, China, and has spread to different cities in China as well as to 24 other countries.

SCORE: 174.16

QUESTION: How many deaths were caused by COVID-19?

FROM PAPER WITH TITLE: Systematic Comparison of Two Animal-to-Human Transmitted Human Coronaviruses: SARS-CoV-2 and SARS-CoV

PASSAGE ANSWER: According to the latest data on Feb. 14th, 2020 [55, 56] , there have been a total of 67,081 clinically diagnosed cases of COVID-19 in worldwide, with 1,526 deaths.

SCORE: 174.26

QUESTION: Does wearing a mask prevent the spread of COVID-19?

FROM PAPER WITH TITLE: Outbreak of Novel Coronavirus (SARS-Cov-2): First Evidences From International Scientific Literature and Pending Questions

PASSAGE ANSWER: Face masks do not seem to be as effective in protecting those who are not infected, and wearing a mask could be useful only when taking care of a person with suspected COVID-19 [20] .There are currently no vaccines against coronaviruses, including SARS-CoV-2.Various vaccine strategies against coronavirus, such as using inactivated viruses, live-attenuated viruses, viral vectorbased vaccines, subunit vaccines, and recombinant proteins are under evaluation.

SCORE: 174.48

5th model: msmarco-distilbert-dot-v5
Maximum Sequence Length: 512
Speed: 4000
Run time: 31 mins on Tesla P100 GPU

QUESTION: What are the coronaviruses?

FROM PAPER WITH TITLE: Alignment-free method for DNA sequence clustering using Fuzzy integral similarity

PASSAGE ANSWER: The coronaviruses are pleomorphic RNA viruses that are widespread among avians, bats, humans and other mammals.

SCORE: 85.51

QUESTION: What was discovered in Wuhuan in December 2019?

FROM PAPER WITH TITLE: Clinical Medicine Characteristics of and Public Health Responses to the Coronavirus Disease 2019 Outbreak in China

PASSAGE ANSWER: In December 2019In December 2019, a cluster of pneumonia of unknown etiology was detected in Wuhan City, Hubei Province of China.

SCORE: 79.24

QUESTION: What is Coronavirus Disease 2019?

FROM PAPER WITH TITLE: Comment

PASSAGE ANSWER: The ongoing coronavirus disease 2019 (COVID-19) outbreak is giving rise to worldwide anxieties, rumours, and online misinformation.

SCORE: 79.98

QUESTION: What is COVID-19?

FROM PAPER WITH TITLE: Consensus statement The species Severe acute respiratory syndrome- related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2 Coronaviridae Study Group of the International Committee on Taxonomy of Viruses*

PASSAGE ANSWER: The present outbreak of a coronavirus-associated acute respiratory disease called coronavirus disease 19 (COVID-19) is the third documented spillover of an animal coronavirus to humans in only two decades that has resulted in a major epidemic.

SCORE: 84.65

QUESTION: What is caused by SARS-COV2?

FROM PAPER WITH TITLE: SARS-Coronavirus Open Reading Frame-3a drives multimodal necrotic cell death

PASSAGE ANSWER: In summary, Rip3-mediated oligomerization of SARS 3a causes necrotic cell death, lysosomal damage, and caspase-1 activation-all likely contributing to the clinical manifestations of SARS-CoV infection. Severe acute respiratory syndrome (SARS) is caused by a coronavirus (SARS-CoV) that at its peak affected more than 8000 people with a 10% mortality rate 1 .

SCORE: 82.65

QUESTION: How is COVID-19 spread?

FROM PAPER WITH TITLE: viruses Perspective Potential Maternal and Infant Outcomes from Coronavirus 2019-nCoV (SARS-CoV-2) Infecting Pregnant Women: Lessons from SARS, MERS, and Other Human Coronavirus Infections

PASSAGE ANSWER: This newly recognized coronavirus, producing a disease that has been termed COVID-19, is rapidly spreading throughout China, has crossed international borders to infect persons in neighboring countries, and humans infected by the virus are travelling via commercial airlines to other continents.

SCORE: 85.82

QUESTION: Where was COVID-19 discovered?

FROM PAPER WITH TITLE: Clinical Medicine Optimization Method for Forecasting Confirmed Cases of COVID-19 in China

PASSAGE ANSWER: In December 2019, a novel coronavirus, called COVID-19, was discovered in Wuhan, China, and has spread to different cities in China as well as to 24 other countries.

SCORE: 85.02

QUESTION: How many deaths were caused by COVID-19?

FROM PAPER WITH TITLE: Systematic Comparison of Two Animal-to-Human Transmitted Human Coronaviruses: SARS-CoV-2 and SARS-CoV

PASSAGE ANSWER: According to the latest data on Feb. 14th, 2020 [55, 56] , there have been a total of 67,081 clinically diagnosed cases of COVID-19 in worldwide, with 1,526 deaths.

SCORE: 87.31

QUESTION: Does wearing a mask prevent the spread of COVID-19?

FROM PAPER WITH TITLE: Estimated effectiveness of symptom and risk screening to prevent the spread of COVID-19

PASSAGE ANSWER: These findings can support evidence-based policy to combat the spread of COVID-19, and prospective planning to mitigate future emerging pathogens.

SCORE: 89.28

Metrics for comparison: 1) Run time (speed) and 2) Accuracy

Model	multi-qa-mpnet-base-dot-v1	all-mpnet-base-v2	all-roberta-large-v1	msmarco-bert-base-dot-v5	msmarco-distilbert-dot-v5
Runtime ¹	~ 1 hour	~ 1 hour	~ 3 hours	~ 1 hour	~ 0.5 hours
Accuracy ²	89%	78%	89%	78%	78%

¹ all run times have been measured on the Tesla P100 GPU offered in Google Colab Pro.

² The accuracy of the model is measured by how many of the 9 questions asked in each model remain unanswered. Note that the answer we sometimes receive may be relevant, but if it does not specifically answer our question by providing the necessary information the question is considered unanswered.

Conclusion: According to the statistics presented in the table above, we can say that the best model overall is the multi-qa-mpnet-base-dot-v1, as it can give us accurate answers to almost all our questions (1 unanswered question) and has a runtime of less than 1 hour (actually 59 minutes). Also, we can say that the msmarco-distilbert-dot-v5 model is "on par", as in half the runtime it leaves only 2 questions unanswered. Therefore, here it is a trade-off between time and accuracy. Note that all these conclusions cannot be precise in choosing the best model and are merely stated as comments on our experiments, since our testing was limited. Finally, we may have high model accuracy rates, but this should not be taken as a benchmark, since again, we only tested each model on 9 questions (limited testing). To be more confident we should experiment with many more questions.