

# Doctor Covid: A Natural Language Processing based Artificial Intelligence Chatbot to Support Accurate Public Health Education in the Pandemic Era



Lily WY Yang; Wei Yan Ng; Xiaofeng Lei; Shaun Tan; Daniel SW Ting

## Introduction

Covid-19 has impacted life on an unprecedented global scale

- Healthcare systems are severely disrupted, due to overwhelmed resources from disease morbidity and containment measures [1]
- Misinformation is rampant on the internet, which may well translate to vaccine hesitancy, and noncompliance to public health measures

With this relentless surge in demand for healthcare and limited physical contact, there is a need to re-explore traditional healthcare models

- Systems utilising artificial intelligence (AI), natural language processing (NLP) and telemedicine have emerged as supporting technology for information delivery, triaging and diagnosis [2]

Existing chatbots for Covid-19 are limited by poor accuracy and language, amongst others [3]. Our study aims to develop a chatbot which:

- Answers Covid-19 related questions that patients may encounter
- Provides multilingualism with 15 languages in addition to English
- Achieves high accuracy using transformer-based high dimensional semantic space representation

## Methodology

We developed a multi-lingual conversational chatbot, Dr. Covid

- The training dataset was developed by creating main unique questions paired with the respective answer (MQA), then expanding each MQA to 5-10 sub-questions to improve coverage and variety
- The testing dataset was created separately based on the training data
- Dr. Covid was assessed based on accuracy and speed, and its performance compared with existing Covid-19 chatbots

The pre-trained NLP models were developed using:

- spaCy transformer
- Bidirectional Encoder Representations from Transformers (BERT)
- Ensemble model (Figure 1), which is created by applying various weightages of the first two models, forming a combined model

Few-shot learning and Easy Data Augmentation (EDA) were utilized in a low-resource setting (Figure 2).

Figure 1. The Ensemble model architecture.

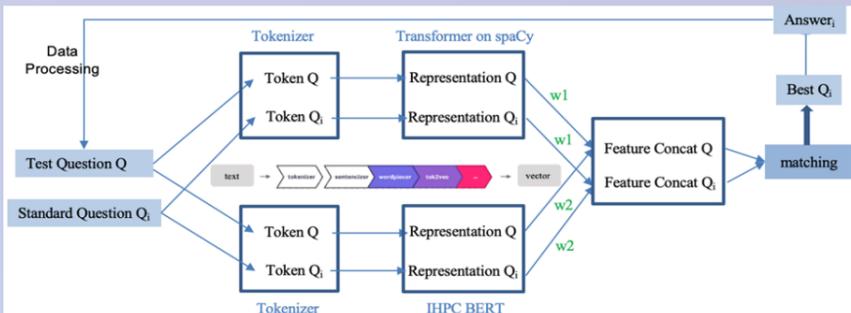
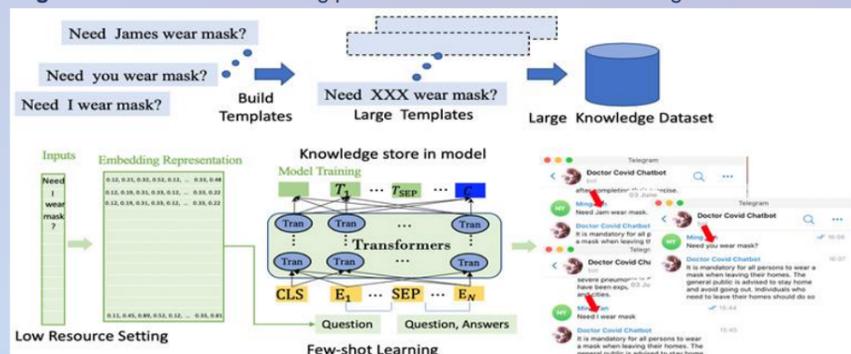


Figure 2. The few-shot learning process under low-resource settings.



## Results

A total of 2703 questions, comprising 1290 local and 1413 global questions, were developed for training. 500 questions were used for testing, comprising 100 local and 400 global questions (Table 1)

- The Ensemble model achieved the highest accuracy, with an overall accuracy of 86.2%, and top 3 accuracy of 94.4%.
- The use of EDA did not value add to existing models, however, with the overall accuracy improving marginally from 84.5% to 84.8%, and the top 3 accuracy worsening from 92.8% to 92.3%.

Table 1. Performance assessment and parameters for Ensemble, spaCy, and BERT.

Model	Parameters	Accuracy (%)	
		Best answer (local/global)	Top 3 answers* (local/global)
Ensemble model	0.7 spaCy vector: 0.3 BERT vector	86.2 (92.0/84.8)	94.4 (97.0/93.8)
BERT	Summing corresponding to all tokens_768 last layer	85.2 (88.0/84.5)	93.2 (95.0/92.8)
	With use of easy data augmentation (EDA)	84.8 (-/84.8)	92.3 (-/92.3)
spaCy transformer	Synonym replacement (SR) and random insertion (RI)	71.0 (69.5/77.0)	80.4 (85.0/79.0)

\* The top 3 closest matching MQA were retrieved if the similarity score fell below 0.90

- Compared to other chatbots (Table 2),
- Dr. Covid (Ensemble) was the most accurate (84.8% vs 22.0-32.8%)
  - With the Graphics Processing Unit (GPU), Dr. Covid generated an answer within 1.0-4.0s on a laptop, and within 1.0s in the model API
  - Dr. Covid could converse in 15 widely-spoken non-English languages, including Chinese, Malay, Tamil, and French

Table 2. Accuracy and user interface assessment for various NLP chatbots and devices, in comparison to Dr. Covid.

NLP chatbot	Accuracy (%)	Average time taken per question (s)		
	Best answer	Laptop	Tablet	Smartphone
Dr. Covid (GPU)	84.8	1.0-4.0	-	-
MOH Ask Jamie	31.8	0.9	0.9	1.1
WHO	32.8	5.0	5.3	5.5
NHS Inform	22.0	5.0	3.9	3.8

## Conclusions

Our study demonstrated the tremendous potential of NLP-based AI chatbots, such as Dr. Covid, in improving healthcare delivery in the pandemic era. Our best model, the Ensemble, showed promising accuracy, multilingualism, and speed, achieving an overall and top 3 accuracy of 86.2% and 94.4% respectively.

Future directions include:

- Expansion to more languages and integration to social media platforms to reach greater audiences
- Exploration of different but complementary domains such as text to speech, and speech to speech, may be of help in specific populations like the visually impaired, or to provide more options for convenience
- NLP-based risk stratification, contact tracing, and patient monitoring

## References

- Mehta S, Machado F, Kwizera A, et al. COVID-19: a heavy toll on health-care workers. *Lancet Respir Med.* 2021;9(3):226-8.
- Gunasekaran DV, Tham YC, Ting DSW, et al. Digital health during COVID-19: lessons from operationalising new models of care in ophthalmology. *Lancet Digit Health.* 2021;3(2):e124-e34.
- Miner AS, Laranjo L, Kocballi AB. Chatbots in the fight against the COVID-19 pandemic. *NPJ Digit Med.* 2020;3(1):65.