Dedicated Question Answering System for Pregnant Women and Neonates Using Corona Virus Literature

Revathi S Nambiar
Department of Computer Science and Engineering, Amrita School of Engineering, Bengaluru,Amrita Vishwa Vidyapeetham,India
revathisujayan@gmail.com

Deepa Gupta
Department of Computer Science and Engineering, Amrita School of Engineering, Bengaluru,Amrita Vishwa Vidyapeetham,India
g_deepa@blr.amrita.edu

***Abstract*— The global pandemic, COVID-19 has made it more important to quickly and precisely retrieve critical information for effective use by specialists in a wide range of fields. The scope of literary exploration and investigation has extended. Domain question answering system will work or produce good to certain extent but still favour more to the pretrained dataset rather than new dataset. In this work we target developing a customized question answering framework that can assist the medical network with retrieval of answers to important logical questions like risk factors, effective modes of communication, various treatment options for target high-risk populaces like pregnant women and neonates.**

The methodology proposed mainly focuses on introducing a novel pipeline architecture using latent dirichlet allocation and bidirectional encoder representation from transformers for embedding the information related to pregnant women and neonates, which can be used for another target population as well. The system is modeled to produce the best and reliable answers for the delicate population, which requires more efficient answers rather than generic population, which can be answered using pretrained systems. Compared to existing work, our work is dedicatedly designed for task informed questions, therefore does not use the existing question answering systems. In this context, the system has showed the accurate and compact answers for different inquiries.

***Abbreviations* – COVID-19, Corona Virus Disease of 2019; CORD-19, COVID-19 Open Research Dataset; LDA, Latent Dirichlet Allocation; BERT, Bidirectional Encoder Representations from Transformers; QnA, Question and Answering.**

***Keywords*—Text Mining, CORD-19, Question and Answering, Haystack, LDA and BERT, Autoencoder, Pregnant Women, Neonates.**

I. INTRODUCTION

The COVID-19 pandemic has made rapid advancements in research into the disease a global necessity . During this time, the measure of information created incredibly surpassed human analysts' capacity to keep up to date. When managing amazingly enormous informational collections, it is hard for people alone to viably dissect and order the entirety of the data [1]. In almost one year, few research work has been initiated in developing question answering system for medical domain, but most of the work focused on modelling the answers using the existing techniques or existing question answering system used to answer the broader domain [8], which failed to produce the accurate or concise answers to the queries. The existing systems used for answering the queries used strategies like voting or ranking for producing the results, which was an overhead for the system and time consuming. We designed a question answering system which did not rely on external rating or ranking system for producing answers, rather developed an innovative strategy for the display of top answers. The information retrieved will not just discuss or provide information of one topic alone, so a technique to incorporate and differentiate among other related topics also is needed. In this research we propose a novel way to capture the relationship of entities from the questions asked, so we needed a technique or a strategy that embeds the full content of the sentence, which can then be bundled with similar subjects . This was achieved by using the combination of latent dirichlet allocation of assignment vectors and bidirectional encoder representations from transformers embedded vectors based on the concept of pipelining.

Thus, retained the semantic knowledge and construct contextual topic identification by integrating latent dirichlet allocation, bidirectional encoder representations from transformers and clustering. Our approach aims to be generalizable to new problems, efficient within the confines. We chose generalizability over precision because we believed that although designing one particular fit-for-purpose approach for each issue could produce accurate results.

The following is our definition of the problem. A background space of documents related to neonates and pregnant women from which extraction of answers takes place. This will be broken down into three steps for any given query: retrieving related documents / sections from a large number of documents using these important records to look for possible answers. Providing enough detail to a researcher in order to present the findings. To evaluate the system, comparing the answers obtained from human assessors and the answers from dedicated system was adopted. We believe that by providing a response that addressed these three objectives, we could significantly improve the approach that researchers could use when looking for answers to their

questions. The presented work aims at the following objectives:

- Novel pipeline architecture using latent dirichlet allocation and bidirectional encoding from transformers for the identification of topics.
- Customization of farm-haystack question answering system by providing the vectors from pipelined architecture.
- Providing more reliable and accurate results when compared to existing question answering systems.

This paper is organised as follows: In section 2, we provide an overview of related work, in section 3, we present our methodology, and in section 4, we present and discuss the results. Finally, in section 4 we present the conclusion of the work.

II. RELATED WORK

A trustworthy question answering system can contribute a lot to the researchers working on corona virus. They can easily locate the articles and ensure that they do not merely redo the previous research. Several works were reported on question answering system which widely focused on a broader domain. Work related to information retrieval or question answering system that addresses how to proficiently respond to the inquiries identified with coronavirus from literary works were few in number. Existing strategies used key word matching, pretrained model, term frequency – inverse document matrix [1]. A topic Modeling based approach has a lot of potential for unsupervised, automated document retrieval for a set of well-defined tasks [2].

The amount of information prepared is insufficient to accomplish equivalent exhibitions of DistilBERT, which is pre-prepared on a larger general-space corpus [3]. Text mining of these reports will assist with fortifying ends and our normal examination of models [4]. Clustering articles would guarantee that researchers do not just modify past investigations, yet rather develop the past work of others. Computerization of such methods utilizing state of the art profound learning order models can incredibly encourage biomedical exploration [5] . Biomedical domains have lots of applications and gets good accuracy using pre-trained models like BioBERT [6]. Most of the papers published under CORD-19 dataset [7] used text mining and analytics concepts like named entity recognition [8] .

Bi-directional transformers were prominent with language-based question answering systems [9]. COVID-19 based datasets had witnessed a few of releasing in the previous years [10]. Exploration of a new domain would need the launching of new applications in order to search the different literatures or publications available with respect to the domain [11]. Pretrained information retrievals have worked good for medical related question answering which gives better accuracy, for instance BioBERT [12].

Research conducted for an innovative research requires lots of domain knowledge and non-domain related knowledge so there was lots of abstracts available for this study [13]. Certain research was beneficial in analysing the zero-shot or transfer capabilities of existing models on COVID-19-related problems [14]. For the objective of retrieving similar questions from community Question Answering forums [15], long Short-Term Memory networks with an attention mechanism that can choose important parts of text are used [16].

Most of the work carried out focused on developing Question Answering systems and Summarisers [17] for different tasks under CORD-19 dataset. Conceptual representation of documents using BERT, Scientific BERT where prominent among the representations. Extractive and Query based summarisers [18], clustering and classifying tools where developed. Question answering systems and summarisers so far developed was not able to convey or capture the effective modes of communication to reduce the spread of corona virus. Open domain question answering based on learning indexable phrase representations [19] that are independent of the query and capture lexical, semantic, and syntactic information using both dense and sparse vectors [20]. When dealing with corona virus-related material, there was a need for natural questions in addition to technical inquiries [21]. The question answering system benefited greatly from transfer learning and transformers, which improved sentence entity relationship [22]. In order to analyse the processing stage, literature research necessitates a variety of visualisation tools [23]. Because the data is so large, better display methods or visualisation strategies are also required [24].

Most of the reported work included general information extraction techniques. Research involved using existing topic identification techniques like latent dirichlet allocation or using clustering methods alone, which failed to prevent the overlapping of clusters, because the information did not involve only a single topic but involved a collection of other topics which was highly context based. Developments so far focused on the broader domain rather than one specific section of population, example - high risk population.

III. DATASET

CORD-19 is a database containing over 400,000 scientific publications about COVID-19, SARS-CoV-2, and associated coronaviruses, including over 150,000 with full text [4].

Papers in CORD-19 are sourced from PubMed Central (PMC), PubMed, the World Health Organization's COVID-19 Database and preprint servers bioRxiv, medRxiv, and arXiv [7]. Some of the metadata entries retrieved from sources do not refer to texts, but rather to materials such as tables of contents, indices, or explanatory records.

These records are found and deleted from the final dataset minimum of one author is required for all conference articles. We screened out full text papers about coronaviruses other than COVID-19 because we are only interested in papers about COVID-19 and no other coronaviruses. After the first screening then we filtered out papers related to neonates and pregnant women which consisted of around 80,000 papers. CORD-19 dataset had many columns and had to pre-process and work only on few columns like year of publishing, abstract and full text of articles. Incorporates two kinds of information to be specific metadata and full content articles, metadata comprises of 19 segments which incorporate data like writers, title, and so on.

IV. METHODOLOGY

The proposed architecture consists of three different methods, that are data acquisition, pipelining of latent dirichlet allocation vectors and bidirectional encoded vectors, and question answering system, as shown in Fig. 1, which is discussed in detail in the following subsections.

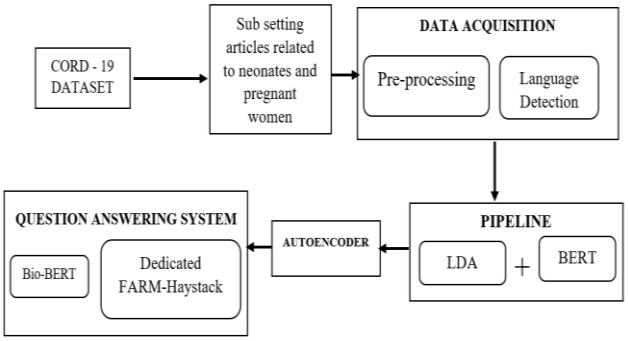

Fig. 1. Flow Diagram of proposed architecture.

A. Data Acquisition

Since the informational dataset index was so enormous and fluctuated in the zones, pre-processing and exploratory information examination is fundamental . The PDF parsing pipeline is used to extract full text and bibliographies from each PDF . In this work, only the articles with full text was used. Utilized a bunch of catchphrases, for example, 'sars-cov-2',' Coronavirus',' 2019-ncov',' ncov2019', 'beta COVID',' pandemic' since the edited compositions had an amazingly wide corpus. After discovering this underlying subset, the papers were separated on whether they contained catchphrases. Furthermore, filtering of articles related to neonates and pregnant women was extracted using keywords like "neo-natal", "natal", "pregnancy", "pregnant"," unborn". Language detection packages like spacy-lang detect was used incorporation with spacy pipeline. In the matching point, transformed our target terms into stemmed versions for compatibility.

B. Novel Pipelined Architecture

The examination into the edited articles indicated that some type of sifting or bunching of subjects would help understand the topics precise. Along these lines, to choose what subject groups were accessible for study, then changed to unstructured bunching utilizing latent dirichlet allocation. The utilization of hierarchical bayesian models, like latent dirichlet allocation has assisted with assembling subjects applicable to another group of papers that seem to focus on airplane and other themes. Even though using Bayesian models alone did not provide a clear differentiation between the topics. Articles involved topics of various kinds which incorporated contextual information. Therefore, a novel pipeline architecture that merged topics identification using contextual information was required.

The articles normally do not intelligibly talk about a solitary theme making it difficult for bayesian model to distinguish the principle topics of the archives. The genuine importance of full content of articles is to a great extent setting based, so word co-event-based strategies like latent dirichlet allocation may fall flat, with these impediments, visualization with this is difficult as the clusters [23] are overlapping. So, emergence of a strategy that inserts the full context of the sentence, which would then be able to bunch with comparable points was much needed. Bidirectional encoded vectors were not able to distinguish the topics perfectly because they focused only on the embedding of the vectors. Sentence embedding procedures like BERT looks at whole sentences and their semantic data as vectors. Extra contextual information is required when texts are incoherent in order to fully reflect the concept of the topics. Fig. 2 shows the pipelining of the embedding vectors using LDA and BERT.

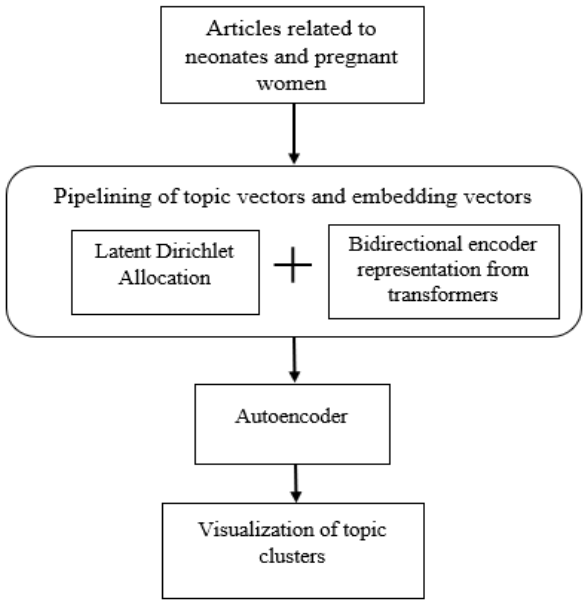

Fig. 2. Flow Diagram of Embedding Pipeline

Sentence transformers was used which helps to produce vectors easily after embedding. In this technique the vectors from bayesian model like LDA and bidirectional encoded vectors are used. For training hierarchical LDA model the number of topics chosen was 33 because after training for various number of topics, the coherence score was best when the chosen number of topics was 33. For Transformer models such as BERT, the input sequence length is 512-word pieces, which correspond to approximately 300-400 words. Longer texts are truncated to the first x word chunks. The provided methods, by default, use a limit of 128-word pieces, longer inputs will be truncated. Loading and storing of BERT embeddings is to pickle and store pre-computed embeddings on disc and load, which can be done using hugging face interface. To balance the relative value of information from LDA and BERT weight hyperparameter called gamma was used, which was set to 15 . The vectors from LDA is multiplied using the weight parameter and then further concatenated with vectors from BERT. The concatenated vector from LDA and BERT were high in dimension so required conversion of vectors into a lower-dimensional latent space using an autoencoder by preserving the contextual information of the vectors,

This was done by using a latent dimension of 32 and activation layer which used ReLU function with a batch size of 128. However, in the event that it begins changing the size of the layers, the process intrinsically learns another approach to capture the information. Thus, this cycle of taking a point in one vector space (Word embedding dependent on LDA and BERT) and placing it into an alternate vector space (the autoencoder-based) is a vector space implanting. Furthermore, implanting the purposes of one vector space into another. Visualization methods like clustering was applied to latent space representations, and contextual topics were extracted from the clusters.

C. *Proposed Question Answering System*

The proposed question answering framework is designed to adopt the embedded vectors. The pipelined vectors from the autoencoder are fed into the dedicated question answering system. The queries were sentence tokenized before sending it to the system, which enabled faster retrieval of answers [4]. Haystack was introduced in recent years and involves techniques like natural language processing, information retrieval and backend engineering technologies. It searches the documents in a DocumentStore for answers to queries. Elasticsearch, SQL, and InMemory are some of the current DocumentStore implementations. By using the Document store, customized embedded vectors from autoencoders related to neonates and pregnant women can be fed into the system. Existing haystack question answering system is pretrained on public dataset, but fine tuning on domain specific data or vectors will bo ost the efficiency by 15-20%. We trained the haystack system with the embedding vectors from autoencoder. The method named write_doc_to_db, will help to fetch the articles from local storage haystack storage. Then using reader as a base model and fine-tune it with auto encoded vectors. Haystack's Readers are trained using the most recent transformer-based language models. The retriever helps the reader by acting as a lightweight filter, reducing the number of documents the reader must process. To answer our actual questions, the Finder connects reader and retriever in a pipeline. Instead of using Elasticsearch's standard BM25, we will use vector similarity of the questions to retrieve them from user question vs. frequently asked ones. For this, we can use the Embedding Retriever and specify a model for the embeddings.

V. RESULTS AND DISCUSSION

A. *Performance Evaluation and Analysis*

The task of evaluating a Question Answering system is difficult. Based on the utility of the system, evaluation can be done in different ways. In this research focus was on the real-world questions like risk factors, effective modes of communication, treatment options related to neonates and pregnant women. Furthermore, classified questions as various types like “OR’, “WHAT’, etc. When evaluating the performance of question answering systems, decision makers must consider not just the situation at hand, but also their own local practices and experiences. Therefore, learnt that human assessor’s assessment will be the right evaluation measure for the proposed question answering system. We selected 15 different questions and seeked the help of 3 medical practitioners from Manipal hospital for answering the questions, which were also routed to the proposed question answering system. Then compared the answers given by them with the answers provided by the question answering system, and discovered that the answers had a very high matching score when compared to pretrained BioBERT which delivers SOTA performance.

B. *Experimental Setup and Results obtained*

Topic modelling results obtained using three different approaches like latent Dirichlet allocation, bidirectional encoder representations and combining both was compared and found that the differentiation of topics was much clearer when combined. From the Fig. 3 it is clear that LDA+BERT gives a better visualization of topics compared to LDA or BERT because the clusters are not overlapping, they are balanced and also provides a clear boundary between the clusters, hence providing and capturing better visualization of topics.

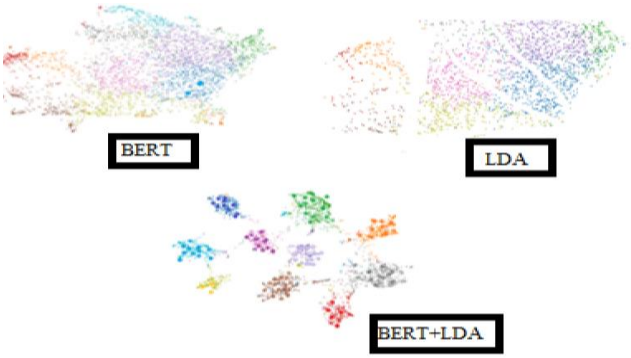

Fig. 3. Visualization of topics using LDA, BERT, LDA+BERT

Contextual topic embedding (BERT + LDA) performed the best overall. The silhouette score for LDA+BERT was 0.27, which was higher than the individual silhouette scores for LDA and BERT when utilized alone.

BioBERT, pretrained question answering system on medical domain was used for comparison, failed to capture the inner semantic relationships of sentences, which motivated us to develop proposed question answering system for pregnant women and neonates using embedded vectors.

Comparison of performance of proposed farm-haystack QnA with Bio-BERT QnA is as follows: Question 1- Do pregnant women with suspected or confirmed COVID-19 need to give birth by caesarean section?

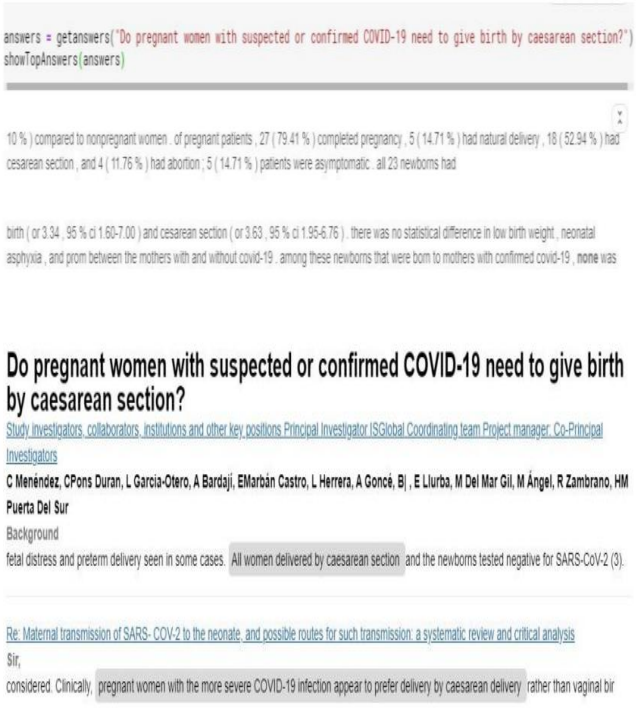

Fig. 4. Snapshot of using Pretrained BioBERT and proposed Farm-Haystack

Questions= getanswers ("Do pregnant women with suspected or confirmed covid-19 need to give birth by caesarean section?) showTopAnswers(answers)
a) Exisitng Biobert Question Answering system
10%) compared to nonpregnant women of pregnant patients , 27 (79.41%) completed pregnancy , 5 (14.71%) had natural delivery , 18 (52.94%) had cesarean section, and 4 (11.76%) had abortion , patients were asymptomatic……..
Birth (or 3.34% to 160-70) and cesarean section (or 3.63, 95%) there was no statistical difference in low birth weight, neonatal , asphysixa…….
a) Proposed Question Answering System
Authors- C Menendez, C Pons Duran, L Garcia-ontero. **Article**- Hydroxychloroquine efficacy and safety in preventing SARS-CoV-2 infection and COVID-19 disease severity during pregnancy (COVID-Preg) **ANSWER**- Fetal distress and preterm delivery seen in some cases. All women delivered caesarean section and the new-born tested negative…….....
Authors- K F Walker , K O'Donoghue , N Grace. **Article**- Maternal transmission of SARS-COV-2 to the neonate, and possible routes for such transmission: a systematic review and critical analysis **ANSWER** -Clinically, pregnant women with the more severe covid-19 infection appear to prefer delivery by casesarean delivery…….

Fig. 5a. Results obtained for "Do pregnant women with suspected or confirmed COVID-19 need to give birth by caesarean section?"

Question 2- Who is at most risk for COVID-19 pregnant women or non-pregnant women?

Questions= getanswers ("who is most at risk for covid-19 pregnant women or non-pregnant women?) showTopAnswers(answers)
b) Exisitng Biobert Question Answering system
Self-reported sars-cov2 testing and difference between pregnant women and non-pregnant women who were hospitalized and those who recovered in the community multivariable regression was used to investigate disease severity and comorbidity…….
Difference between pregnant women and non-pregnant women who were hospitalized and those who recovered in the community multivariable regression was used to investigate disease severity and comorbidity effects results…….
b) Proposed Question Answering System
Authors- KA Pastick, MR Nicol, E Smyth…. **Articles**- A Systematic Review of Treatment and Outcomes of Pregnant Women With COVID-19-A Call for Clinical Trials……. **ANSWER**- Of 70% of the total healthcare workforce. Pregnant women and women of reproductive potential are therefore at a high risk of contracting covid-19……
Authors- H Dong, M Zhang… **Article**- Investigation on the mental health status of pregnant women in china during the pandemic covid-19…… **ANSWER**- Pregnant women were more likely to have an adverse pregnancy outcome when compared to women who were not pregnant ……..

Fig. 5b. Results obtained for "Who is at most risk for COVID-19 pregnant women or non-pregnant women?"

Question type3- What are the treatment options for pregnant women infected with COVID-19?

Questions= getanswers ("what are the treatment options for pregnant women infected with covid-19?) showTopAnswers(answers)
a) Exisitng Biobert Question Answering system
Results thirty-three pregnant women with covid-19 and 28 newborns were identified. one (3%) pregnant woman needed the use of mechanical ventilation. There were no moralities among the pregnant women or newborns…….
5.8% of non pregnant women , after adjusting for age, presence of underlying medical conditions , and pregnant women were significantly more likely to be admitted to the intensive care unit……
a) Proposed Question Answering System
Authors- N Tug, M Yassa, E Kole…….. **Article**- Pregnancy worsens the morbidity of COVID-19 and this effect becomes more prominent as pregnancy advance…… **ANSWER**- Either by hospital admission due to the COVID-19 disease symptoms or by universal testing for all obstetrical admissions……..
Authors- G Eysenbach, M Schoen, YHeng Yaw, L Wei…….. **Article**- Pregnancy worsens the morbidity of COVID-19 and this effect becomes more prominent as pregnancy advance…… **ANSWER**- Hence, early diagnosis, isolation and treatment are crucial for the pregnant women and neonates……

Fig. 5c. Results obtained for "What are the treatment options for pregnant women infected with COVID-19?"

As we find in the outcome, answers utilizing the existing question answering system like BioBERT did not perform well when contrasted with proposed Farm-Haystack, on the grounds that the BioBERT utilizes similar structure to address all the inquiries independent of the area of the inquiry posed. BioBERT was simply ready to catch the words referenced in the inquiry though, proposed question answering framework had the option to capture the semantic likeness between the words in the posed inquiry because the input provided was the embedded vectors from LDA and BERT.

It is clear from Fig. 5b that answers produced using BioBERT did not provide the actual intact answers to the question, rather presented a keyword matching answer. On the other hand, the proposed question answering system clearly distinguished the question and precisely answered the question and subsequently the yield was pregnant women. The answers presented using proposed framework consisted of title of the literature, the author's name, also highlighted the key points within the answer

VI. CONCLUSION AND FUTURE WORK

In this research we were able to build a dedicated question answering system for neonates and pregnant women and also delivered superior results when compared to existing question answering system in medical domain. In this undertaking we have utilized two Question answering frameworks to be specific BioBERT QnA and Farm-Haystack QnA.

QnA was fed with pipelined vectors from LDA+BERT as information. Embedding only the articles related to neonates and pregnant women using the pipelined architecture helped to improve the closeness of the inquiry and the result.

The current inquiry answering frameworks where not trained to address the inquiries from clinical space, that was because question noting framework pretrained on biomedical area was presented (BioBERT). Farm-Haystack question answering frameworks searches away at archive semantic likeness search so this structure with pipelined vectors had the option to give better outcomes. Thus, we found that the dedicated QnA have better results when compared to pretrained models in medical domain. The proposed QnA was evaluated with the help of human assessors and found to be a useful question answering system for dedicated questions.

In the future, this dedicated question answering system has the potential to be generalized for all other domain specific questions. Future work involves addressing the other risk factors related to COVID-19, Question Expansion utilizing Word Embeddings, Utilizing Extractive Summarizer rather than Abstractive Summarizer.

REFERENCES

[1] C. Herlihy and Y. Liu, "Automated task informed document retrieval on the COVID-19 open research dataset using topic modeling," Acl Conference , pp 3-4, 2020.

[2] M. Muffo, A. Cocco, M. Messina and E. BERTino, "Question answering tool for COVID-19,", Acl Conference, pp-1-2, 2020.

[3] A. Poliak et al., "Collecting verified COVID-19 question answer pairs," Acl Conference,2020.

[4] V. Nguyen, M. Rybinski1, S. Karimi and Z. Xing, "Searching scientific literature for answers on COVID-19 questions", Acl Conference, pp 1-2, 2020.

[5] D. Das et al. , "Information retrieval and extraction on COVID-19 clinical articles using graph community detection and bio-BERT embeddings," Acl Conference, 2020.

[6] W. Yoon, J. Lee, D. Kim, M. Jeong, and J. Kang, "Pre-trained language model for biomedical question answering," Acl Conference, 2019.

[7] B. J. Gutiérrez, J. Zeng, D. Zhang, P. Zhang and Y. Su, "Document classification for COVID-19 literature," Acl Conference,2020.

[8] J. W. Park, "Continual BERT: Continual learning for adaptive extractive summarization of COVID-19 literature,",Asian Conference on Information Retrieval, 2020.

[9] J. Devlin, M. Chang, K. Lee and K. Toutanova, "BERT: pre-training of deep bidirectional transformers for language understanding," arXiv:1810.04805, In Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies,pp 2-3,2019.

[10] M. Abdul-Mageed, A. Elmadany, E. M. B. Nagoudi D. Pabbi, K. Verma, and R. Lin, "Mega-cov: A billion-scale dataset of 65 languages for COVID-19," Acl Conference, 2020.

[11] O. W. Lithgow-Serrano, A. Lopez-Fuentes, Y. Balderas-Martinez, F. Rinaldi and J. Collado-Vides, "A smart literature exploration environment for COVID-19 literature," In European conference on information retrieval, 2020.

[12] J. Lee et al., "BioBERT: a pre-trained biomedical language representation model for biomedical text mining," Bioinformatics, vol. 36, 2020, pp. 1234–1240.

[13] T. H. Huang, C. Huang, C. C. Ding, Y. Hsu, and C. L. Giles, "Coda-19: Reliably annotating research aspects on 10, 000+ CORD-19 abstracts using non-expert crowd", Acl Conference, 2020.

[14] R. Tang et al., "Rapidly bootstrapping a question answering dataset for COVID-19," Acl Conference,pp 2-3,2020.

[15] S. Romeo et al., "Neural attention for learning to rank questions in community question answering," In Proceedings of COLING, the 26th International Conference on Computational Linguistics: Technical Papers, pp. 1734–1745,2016.

[16] N. Reimers and I. Gurevych, "SentenceBERT: Sentence embeddings using siamese BERT networks," Acl Conference, 2019.

[17] Y. Xu and M. Lapata, "Query focused multi-document summarization with distant supervision," Acl Conference, 2020.

[18] E. Perez, P. Lewis, W. Yih, K. Cho and D. Kiela, "Unsupervised question decomposition for question answering," Acl Conference, 2020

[19] M. Seo, J. Lee, T. Kwiatkowski, A. Parikh, A. Farhadi and H. Hajishirzi, "Real-time open-domain question answering with dense-sparse phrase index," Acl Conference,2019.

[20] M. T. R. Laskar, E. Hoque and J. Huang, "Query focused abstractive summarization via incorporating query relevance and transfer learning with transformer models," In: Goutte C., Zhu X. (eds) Advances in Artificial Intelligence, vol. 12109,2020.

[21] T. Kwiatkowski et al., "Natural questions: a benchmark for question answering research," Transactions of the Association of Computational Linguistics, vol. 7, pp. 452–466, 2019.

[22] C. Yanover et al., "What factors increase the risk of complications in sars-cov-2 positive patients? a cohort study in a nationwide israeli health organization," medRxiv 2020.05.07.20091652, 2020.

[23] D. Grujicic, G. Radevski, T. Tuytelaars and M. B. Blaschko, "Self-supervised context-aware COVID-19 document exploration through atlas grounding," Acl Conference , 2020.

[24] W. F. Oyewusi, "Interactive Visualization and Simplified Pattern Discovery in the COVID-19 Open Research Dataset (CORD-19)," Acl Conference, 2020.