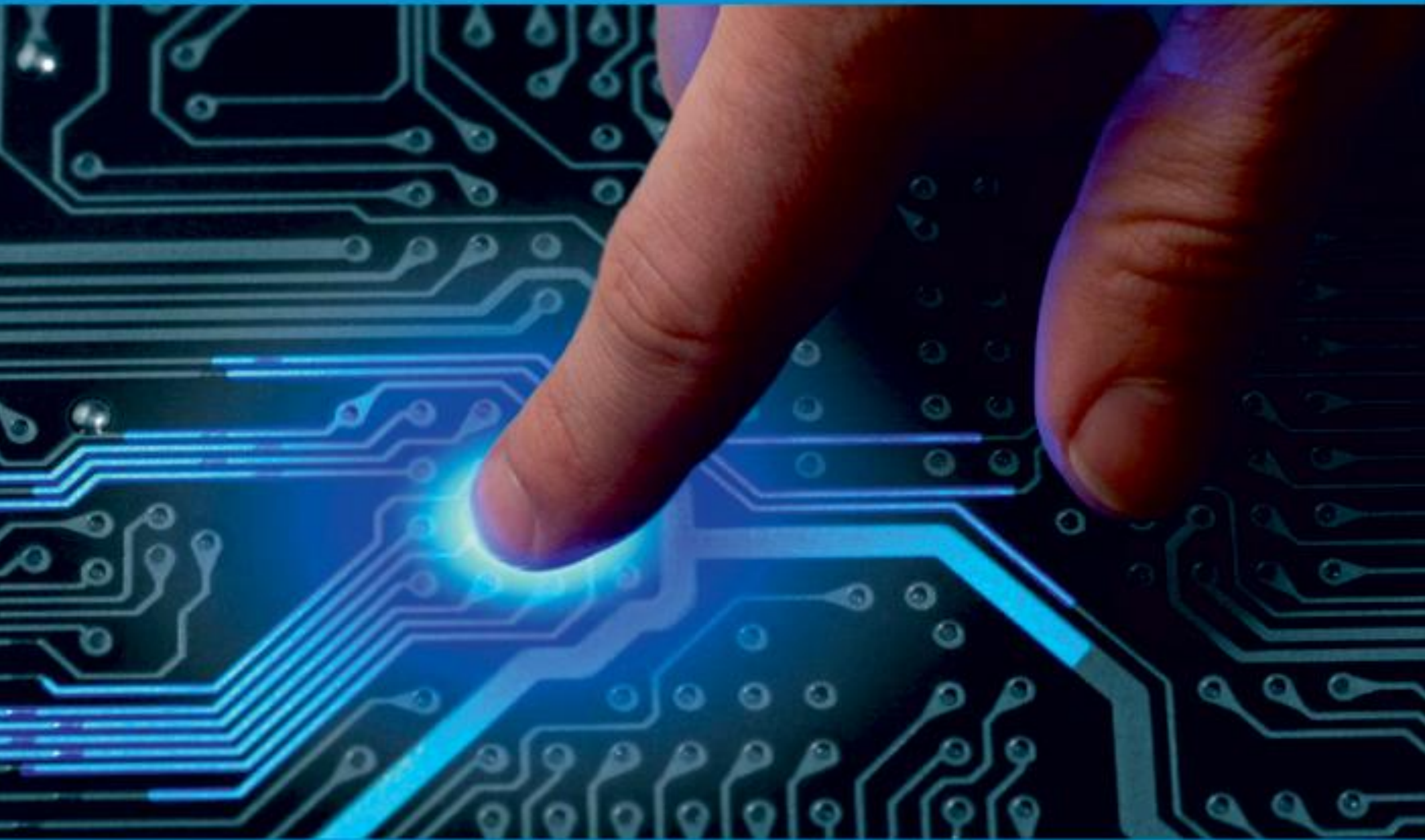
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 4, April 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104049 |

An Extractive Question Answering System

[1]Shahid Afridi Mandal, [2]Rahul Sakhalkar, [3]Pratik Kadlak, [4]Dr Vipul Dalal

[1,2,3] UG Students, Dept of Information Technology, Vidyalankar Institute of Technology, Mumbai, India

[4] Professor, Dept of Information Technology, Vidyalankar Institute of Technology, Mumbai, India

ABSTRACT: Recent advances in the domain of Natural Language Processing (NLP) have led to the development of intelligent model machines that can understand natural language. An extractive Question Answering (QA) system retrieves relevant information from the textual context provided to answer questions asked by users. The objective of this paper is to leverage DeBERTa, a transformer-based deep learning model with self-attention mechanisms and advanced pretraining techniques. The following extractive question answering system accurately captures the context of the text. Our system uses the SQuAD dataset, a popular benchmark dataset for evaluating QA systems. This paper also demonstrates the usefulness of DeBERTa and SQuAD datasets in training extractive QA systems.

KEYWORDS: Extractive question-answering, DeBERTa, SQuAD dataset, natural language processing, deep learning, self-attention mechanism, pre-training.

I. INTRODUCTION

In recent years, the development of Question Answering (QA) systems has become increasingly popular due to their ability to provide accurate responses to user queries. Extractive techniques are used by the systems to retrieve answers from textual contexts or passages, rather than generating responses from scratch.

However, the biggest challenge for these systems is their ability to understand the context of the text and extract relevant information to provide accurate answers to user questions.

Deep learning models such as BERT, RoBERTa, and DeBERTa have shown significant improvements in QA tasks when compared to conventional machine learning models. These models are based on neural networks and can learn the underlying patterns and relationships between words and phrases in language, making them highly suitable for developing extractive question-answering systems that can accurately understand the nuances of natural language and provide precise answers to user queries.

Extractive Question-Answering (QA) systems can sort through extensive amounts of data to select the most relevant information for a given topic, making them capable of handling complex queries. This feature increases the system's accuracy in responding to user queries. These systems find practical applications in customer support, virtual assistants, and search engines.

The performance of the extractive QuestionAnswering system is based on their potential to comprehend the context of the provided text and retrieve useful content to provide accurate answers. While they have a constraint in generating original responses, these systems are invaluable tools for responding to user queries accurately and efficiently.

Based on the literature survey, the development of QA systems has led to significant advancements in natural language processing, and extractive techniques have played a crucial role in this field. Deep learning models such as BERT, RoBERTa, and DeBERTa have enabled these systems to understand the nuances of natural language and provide accurate and efficient responses to user queries. As a result, extractive QA systems are now widely used in various applications, including customer support, virtual assistants, and search engines.

The report is organized into several sections, beginning with an introduction that provides background information on the topic. The literature survey follows, summarizing key findings from previous research, and identifying gaps in previous literatures. The methodology section describes the data collection methods, and data analysis techniques used. The experimental setup section describes what dataset is used, where it is from, and what are the features of the dataset. The result section presents the findings of the study, including statistical analysis. The conclusions section summarizes the key findings, makes recommendations for action, and identifies any limitations and suggestions for future research. The references section has all the sources cited in the report.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104049 |

II. LITERATURE SURVEY

"DeBERTa: A Decoding-enhanced BERT with Disentangled Attention" was developed in 2021. This paper presents DeBERTa, a deep learning model that enhances BERT and RoBERTa's shortcomings by incorporating more sophisticated pre-training methods and self-attention mechanisms. [1]

"SQuAD: 100,000+ Questions for Text Comprehension by Machines." This paper introduces the SQuAD dataset, a popular benchmark dataset for question-answering systems. It has a lot of questions and answers with corresponding answer spans. [2]

In 2019, the study "Improving Question Answering by Commonsense-Based Pre-Training" was published. To improve the efficacy of QA systems, this paper introduces a pre-training strategy that incorporates common sense knowledge. On the SQuAD dataset, the methodology is assessed, and the results are cutting-edge. [3]

"BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," from 2018. To better understand the context of the text, this paper introduces BERT, a deep learning model that makes use of pre-training methods and self-attention mechanisms. BERT has demonstrated appreciable progress in a range of NLP tasks, including question-answering. [4]

"RoBERTa: A Robustly Optimized BERT Pre-Training Approach." 2019. In this paper RoBERTa, a deep learning model that enhances BERT by enhancing the pretraining procedure, is introduced. RoBERTa has demonstrated appreciable progress in a range of NLP tasks, including question answering. [5]

"Reading Wikipedia to Answer Open-Domain Questions" (2008). This paper presents DrQA, an open-domain question-answering system that uses a document retriever and a neural reader to generate answers. [6]

"Neural Machine Reading Comprehension: Methods and Trends". This survey paper provides an overview of neural machine reading comprehension, including model architectures, training techniques, and evaluation metrics. [7]

"Transformer-XL: Attentive Language Models Beyond a Fixed-Length Context". This paper introduces Transformer-XL, a language model that can handle long-term dependencies and generate context-aware representations for question answering. [8]

III. PROBLEM STATEMENT

Systems that automatically react to user inquiries in natural language by giving pertinent information from diverse sources are known as question answering systems. Since the amount of information on the internet exponentially increases, making it challenging for users to quickly and effectively obtain the information they need, these systems are becoming more and more crucial.

In order to analyse and comprehend user inquiries, question answering systems must develop precise and efficient natural language processing algorithms. Understanding the context and subtleties of the language used as well as the question's aim must be understood in order to do this.

Finding and choosing pertinent information from a big volume of unstructured data is a significant difficulty as well. To find the most pertinent and precise responses to user inquiries, question-answering systems must be able to gather and sort through enormous volumes of data.

In order to increase user experience and improve access to information, it is crucial to design efficient questionanswering systems. To overcome these obstacles, creative data gathering and curation methods, cutting-edge natural language processing methods, and ongoing system performance review and improvement are all necessary.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104049 |

IV. PROPOSED METHODOLOGY

Question answering system involves various steps, including data collection and pre-processing, natural language understanding, information retrieval, and answer generation. The following methodology discuss these steps in detail.

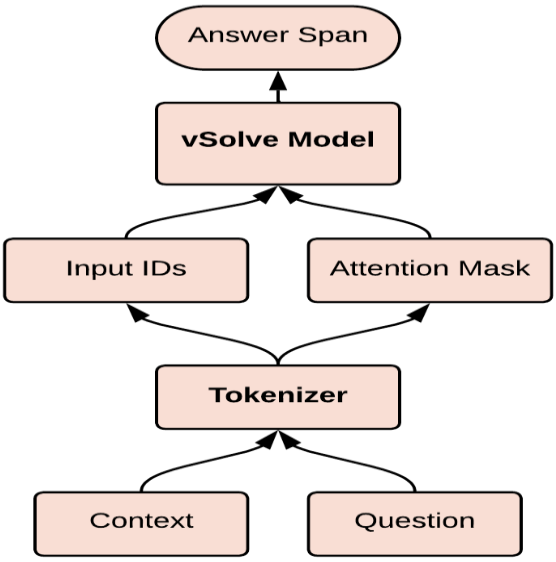


Fig4.1 Proposed System

Step 1: Data Collection and Pre-processing

The first step is to collect and pre-process the data. This involves selecting a high-quality dataset that covers a wide range of topics and domains. The dataset should be pre-processed to remove noise and irrelevant information, such as stop words, special characters, and punctuation. DeBERTa requires a large amount of data for pretraining, so a large dataset should be used if possible.

Step 2: Pre-training

The second step is to pretrain the DeBERTa model on the dataset. This involves feeding the model large amounts of text to learn the underlying patterns and relationships in the language. The pretrained model can then be fine-tuned on a smaller dataset for the specific task of question answering.

Step 3: Fine-tuning

The third step is to fine-tune the DeBERTa model on a smaller dataset for the specific task of question answering. This involves feeding the model pairs of questions and answers and fine-tuning the weights of the model to maximize the probability of generating the correct answer given the question.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104049 |

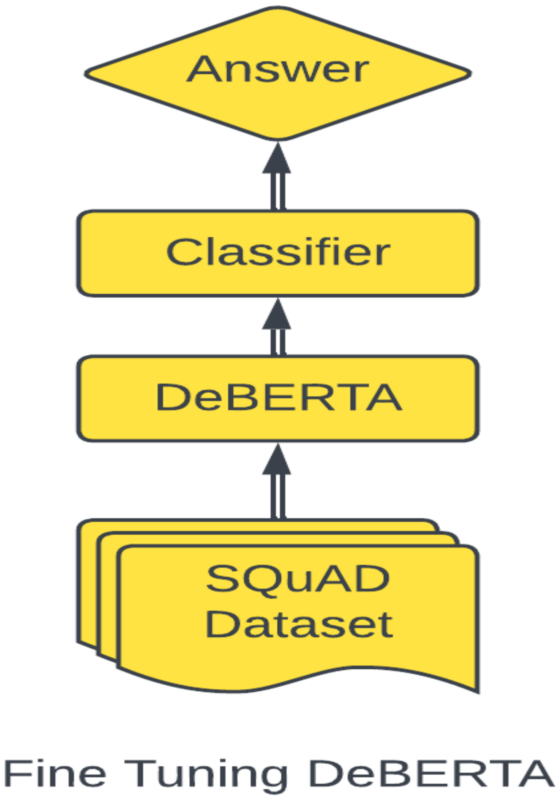


Fig4.2. Fine Tuning DeBERTa

Step 4: Evaluation

The final step is to evaluate the performance of the DeBERTa-based question answering system. This involves testing the system on a held-out test set and calculating metrics such as accuracy, precision, recall, and F1-score. The performance can be improved by fine-tuning the model or using more advanced techniques such as ensemble learning or data augmentation.

Once the DeBERTa model is trained, it can be used to build a question answering system. The system can be test against metrics like precision, recall, F1-score. The performance can be improved by using advanced techniques. The system takes a context and a question as input from the user, process the question, and generate a relevant and accurate answer. The system can be used in search engines, chatbots, virtual assistants, etc.

V. SYSTEM ARCHITECTURE

Input Embeddings: The input sequence of tokens is first processed by a word embedding layer that converts each token into a dense vector representation. DeBERTa uses learned positional encoding to encode the relative position of each token in the sequence.

Disentangled Self-Attention: DeBERTa uses disentangled self-attention to separate the content and position information in the input sequence. This allows the model to address different aspects of the input sequence separately, improving the model's ability to capture complex relationships between words.

Enhanced Masking: DeBERTa uses an enhanced masking strategy that incorporates both random masking and span masking. This improves the model's ability to handle longrange dependencies and reduces overfitting to the masked tokens.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104049 |

Stacked Transformer Layers: The input embeddings are passed through a series of transformer layers, which consist of multi-head self-attention and feed-forward neural network layers. The transformer layers allow the model to capture increasingly complex relationships between the input tokens.

Pooling Layer: After the transformer layers, the final hidden states are passed through a pooling layer that summarizes the input sequence into a fixed-length vector representation.

Modified Output Classifier Layer: Instead of the default output classification layer used in DeBERTa, we have modified it by changing the size of the layer. This involves adding a fully connected layer with a new number of output units that corresponds to the number of output labels required for our task, in this case, question answering.

Fine-tuning: The entire model is fine-tuned on a labelled dataset of question-answer pairs using backpropagation and gradient descent. During fine-tuning, the weights of the transformer layers and the modified output classifier layer are updated to optimize the model's performance on the question answering task.

Overall, this modified DeBERTa architecture with a modified output classifier layer is well-suited for question answering tasks and can be fine-tuned on a labelled dataset to achieve high performance on the task

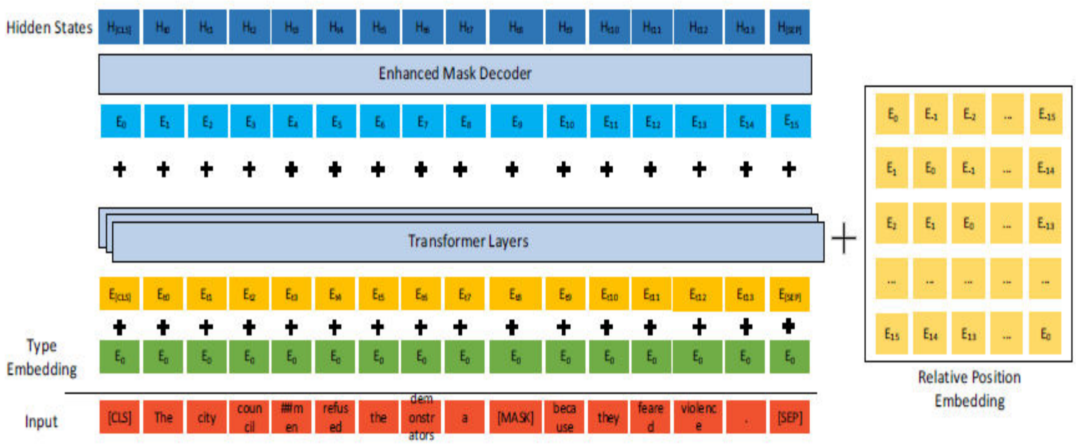


Figure 5.1 System Architecture

VI. RESULT

MODEL	F1 SCORE	EXACT MATCH (EM)
BERT-base	88.5	80.8
DistilBERT	86.9	77.7
Electra-base	90.8	84.5
DeBERTa-v3-base	**92.2**	**84.9**

Table 7.1 Results

BERT-base, DistilBERT, Electra-base, and DeBERTav3-base are four well-known transformer-based language models, and the table 7.1 displays their F1 scores and exact match (EM) performance. The F1 score assesses a model's overall accuracy, while EM gauges how frequently the model produces the precise solution. According to the table, DeBERTa-v3-base has the best overall performance, scoring 92.2 on the F1 scale and 84.9 on the EM scale. With an F1 score of 90.8 and an EM score of 84.5, Electra-base also performs admirably. Bert-base has the greatest F1 score of the two models, whereas DistilBERT and its scores are marginally lower. The DeBERTa-v3-base language model

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104049 |

performs the best overall in this table, demonstrating the high degree of performance that transformer-based language models may attain in question-answering tasks.

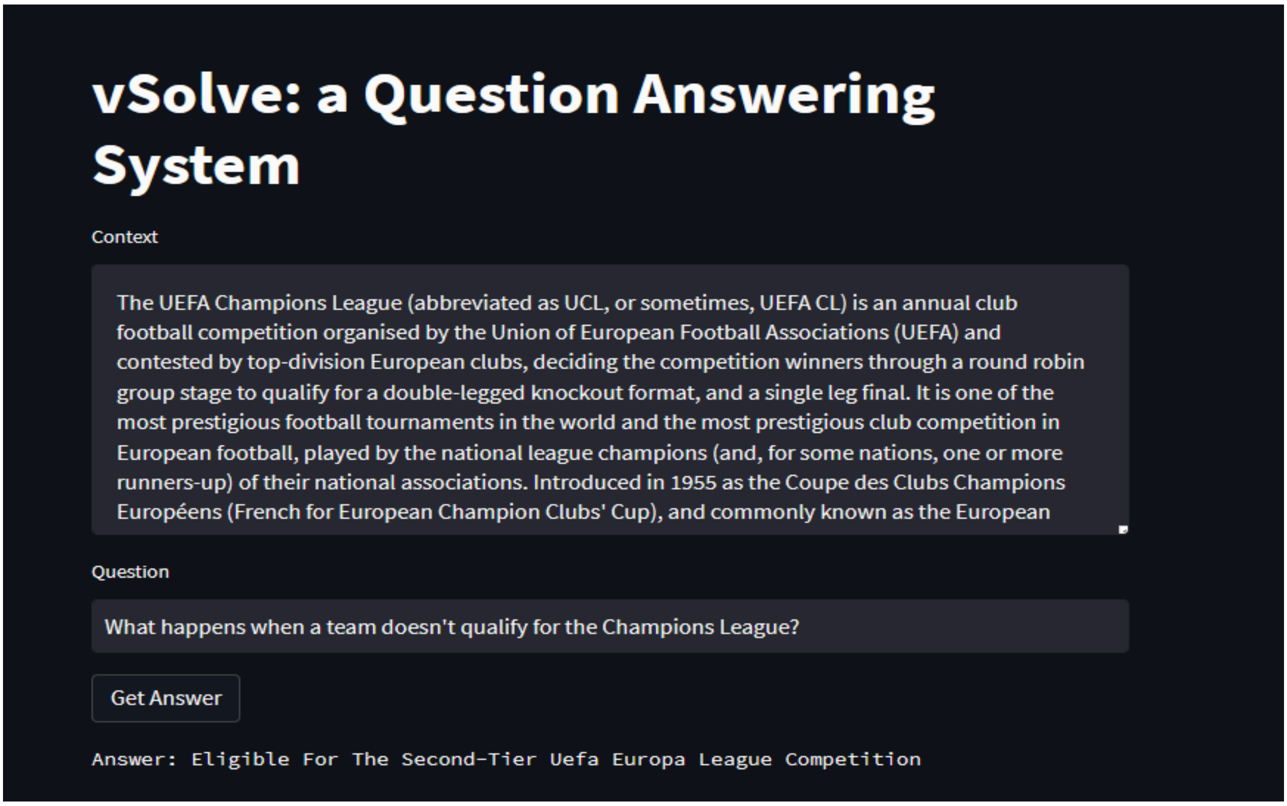


Figure 6.1 Results (Contextual Answer)

Figure 6.2 Results(Out of Context Question)

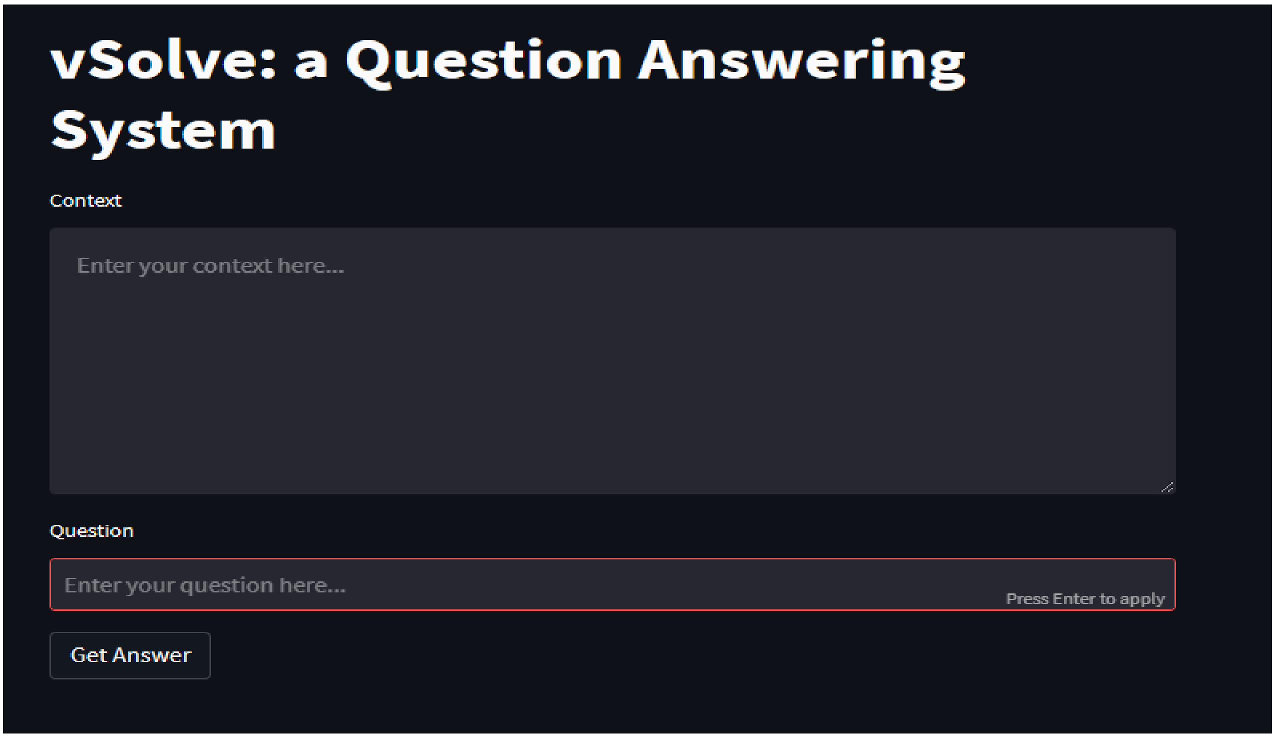


Figure 6.3 Results(Empty field)

VII. CONCLUSION AND FUTURE WORK

In conclusion, both the SQuAD v1.1 and v2.0 datasets have shown remarkable performance for our suggested extractive QA method based on DeBERTa and the SQuAD dataset. The system's impressive performance is proof of the effectiveness of DeBERTa's pre-training techniques and self-attention mechanisms for comprehending the context of texts and extracting pertinent information to offer appropriate replies.

The Question Answering System offers numerous opportunities for future research and development. Some potential avenues for future work include:

- Multilingualism: As the world becomes more and more globalised, there is an increasing requirement for QA systems that can process and respond to enquiries in several languages. Future QA systems are expected to include sophisticated natural language processing techniques that can handle the nuances of various languages and dialects.
- Contextual Understanding: The context of a question might be difficult for current QA systems to grasp, which can result in incorrect or irrelevant replies. Machine learning algorithms that can better grasp the context of a question, including the user's purpose and the larger discourse in which the question is being posed, will probably be incorporated into future QA systems.
- Personalization: QA systems might become increasingly suited to the user's interests, preferences, and past encounters with the system.
- Integration with other technologies: QA systems might be linked up with other tools like chatbots, voice assistants, and augmented reality software. Via a variety of channels, such as text, speech, and visual interfaces, users may be able to pose queries and get responses.

REFERENCES

1. Xiaodong Liu, Jianfeng Gao, Weizhu Chen, Pengcheng He, "DEBERTA: DECODING-ENHANCED BERT WITH DISENTANGLED ATTENTION," 2021.
2. Jian Zhang, Konstantin Lopyrev, Percy Liang, Pranav Rajpurkar, "SQuAD: 100,000+ Questions for Machine

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104049 |

Comprehension of Text," 2016.

3. Duyu Tang, Nan Duan, Ming Zhou, Jiahai Wang, Jian Yin. Wanjun Zhong, "Improving Question Answering by Commonsense-Based Pre-Training," 2019.
4. Ming-Wei Chang, Kenton Lee, Kristina Toutanova, Jacob Devlin, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," 2018.
5. Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, Veselin Stoyanov, Yinhan Liu, "RoBERTa: A Robustly Optimized BERT Pretraining Approach," 2019.
6. Danqi. Chen, "Reading Wikipedia to Answer Open-Domain Questions," 2008.
7. Shanshan Liu, Xin Zhang, Sheng Zhang, Hui Wang, Weiming Zhang, "Neural Machine Reading Comprehension: Methods and Trends," 2021.
8. Zihang Dai, Zhilin Yang, Yiming Yang, Jaime Carbonell, Quoc V. Le, Ruslan Salakhutdinov, "Transformer-XL: Attentive Language Models Beyond a Fixed-Length Context," 2019.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.165

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details