Proceedings of the 5th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/18/20230964

Research on methods and applications of question answering system in the context of ChatGPT

Xiangyu Li

Institute of technology, Wen-Zhou Kean University, Shao Xing, 325060, China

Corresponding author: 1234960@wku.edu.cn

Abstract: In the 21st century, there has been a growing importance placed on the "body" of artificial intelligence, particularly as it relates to language processing. Researchers have developed various machine learning models with a focus on language understanding, including Large Language Model (LLM), Bidirectional Encoder Representation from Transformers (BERT), and Natural Language Processing (NLP). These models have led to the development of numerous applications, such as ChatGPT-3.5, which has recently gained widespread attention. In addition to ChatGPT, other applications have also benefited from these language processing models, including Question Answering Systems (QAS). This paper will examine three QAS that have been enhanced by the context of ChatGPT, discuss the relevant applications, and analyze these different applications in order to predict future trends in this field. One notable QAS is OpenAI's GPT-3-powered AI that can answer questions about any topic. This application leverages the capabilities of GPT-3 to provide accurate and informative responses to a wide range of questions. Another QAS is IBM's Watson, which utilizes natural language processing and machine learning algorithms to understand and respond to user queries. Watson has been used in various industries, including healthcare, finance, and retail. A third QAS is Google's BERT-based system, which uses pre-trained language models to improve its responses to user queries. This system has been integrated into Google Search and other products, allowing users to receive more precise and relevant search results. Overall, the development of these QAS and other language processing applications marks an exciting period of progress in the field of artificial intelligence. As researchers continue to refine these models and explore new applications, we can expect to see even more advanced and sophisticated language processing systems emerge in the future.

Keywords: Q&A system, NLP, deep learning, application, Ai.

1. Introduction

In 1956, John McCarthy introduced the concept of "AI" at the Dartmouth Conference, and the world began to recognize the significance of this field and eagerly dove into research. Despite facing two major setbacks, researchers ultimately overcame these challenges, leading to the development of groundbreaking applications such as ChatGPT-3.5.

On November 30th, 2022, ChatGPT-3.5 was first published and made available for free use. This achievement not only signifies a bright future for AI but also reflects significant improvements in

DOI: 10.54254/2755-2721/18/20230964

machine learning models. Popular models such as NLP and LLM are increasingly used by researchers, leading to advancements in areas such as Q&A systems [1].

This paper will introduce three types of Q&A systems (Visual, Voice, and Text Question Answering Systems) and their current applications. We will then summarize various applications and predict future trends in the use of Q&A systems. Visual Q&A systems utilize computer vision technology to analyze images and answer questions about their content. These systems have a wide range of applications, including automated customer service and content categorization. Voice Q&A systems are becoming increasingly popular as virtual assistants like Siri and Alexa improve their capabilities in answering user queries through natural language processing. Text Q&A systems, on the other hand, rely on machine learning models to analyze text input and provide relevant responses. These systems have been used in various industries, from healthcare to finance. As machine learning models continue to improve, we can expect to see further advancements in Q&A systems in the future. For example, combining multiple input types (such as voice and visual) to provide more comprehensive answers to user questions [2]. Additionally, the integration of AI and Q&A systems in various fields will lead to enhanced efficiency and productivity. Overall, the future looks promising for Q&A systems as they continue to evolve and play an increasingly vital role in various industries.

2. Q&A system

2.1. Visual question answering system

In recent years, VQA models that fuse image visual features and question text features have gained significant attention from researchers. The focus of VQA is on understanding visual content and natural language more comprehensively, accurately representing modal features, and effectively integrating cross-modal information. However, most existing models neglect the autocorrelation information of image regions and question words, instead relying on fine-grained interaction and matching through attention mechanisms and iterative operations based on image regions and question word pairs.

To address this issue, Lu and their group developed a symmetrical attention mechanism-based model. This model utilizes existing methods such as Faster RCNN and Bi-LSTM to build a feature extraction network for images and questions, which are fused together using attention architecture. This approach enables comprehensive bilateral information understanding while improving the correlation between image regional features and problem word pairs. Analyzing semantic associations between images and questions reduces overall semantic deviation, resulting in an improved accuracy of answer prediction. The use of the popular VAQ2.0 dataset with a larger sample size and a greater variety of questions ensures accurate and reliable test results.

This novel model shows great potential for future applications in the field of visuals, particularly for short videos.

2.2. Voice question answering system

In recent years, the rapid development of internet technology and the increasing prevalence of intelligent devices have provided people with unprecedented convenience in their daily life. However, the traditional approach used by Voice Question and Answer Systems to solve problems with deep neural networks (DNN) has limitations. Specifically, it can only be used to solve problems where the input and output are fixed dimension vectors and cannot tackle more complex tasks such as machine translation and speech recognition where the input and output vectors are not fixed dimensions. This disadvantage leads to slower and less accurate identification, incorrect answers, and other issues. To overcome this limitation, Zhang's group introduced RNN to address the problem of mapping variable-length sequences in machine translation using an Encoder-Decoder architecture, as shown in Figure 1.

DOI: 10.54254/2755-2721/18/20230964

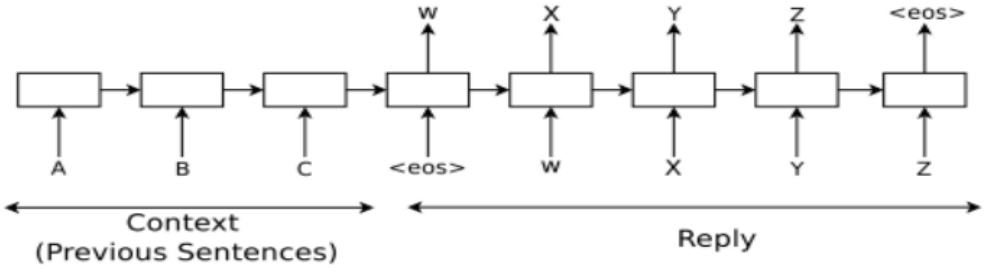

Figure 1. The architecture of Seq2seq.

However, in the aspect of decoding the semantic vector, When the length of the source sequence increases gradually, it is difficult for the vector with a fixed length to represent all the information of the source sequence. Especially, when the length of the input sequence is larger than the length of the source sequence during training, the detection performance of the model will be affected. The longer the input sequence length is, the worse the performance of the model will be. To solve this problem, Zhang's group introduces the Soft Attention Model. Instead of using Encoder to generate fixed semantic vectors, it generates different semantic encodings of source sequences according to different distributions of attention. What's more, Zhang's group also creates a user rating module to fine–tuning the model in real-time.

2.3. Text question answering system

Compared to the Medical Text Question Answering System in English, the Chinese Medical Text Question Answering System faces unique challenges due to the complexity of Chinese medical terminologies. Additionally, the traditional text matching method has limitations, such as the inability to account for word polysemy, contextual information within texts, and effective extraction of high-level semantic features of sentences. Therefore, researchers must delve deeper into sentence semantics when considering text-matching problems. Moreover, medical text-matching encounters the following challenges: complex semantic meanings, extensive medical terms that are difficult to extract, and short question-and-answer texts that make contextual information extraction challenging.

Thankfully, the rapid development of deep learning has opened a new stage for the text-matching task. The method of using pre-training models for fine-tuning domain data is increasingly being accepted. To address text-matching problems with Chinese disease question answering, Wu's group proposed a text-matching approach based on BERT pre-training models and lifting tree models [3]. They used the dataset provided in the disease Question Answering Transfer Learning Evaluation Task held by CHIP2019 for verification purposes. Their approach had higher efficiency and better results compared to traditional methods and BERT models.

However, this study still has its limitations. Feature extraction is not detailed enough, and additional features such as TF-IDF and LDA can be extracted. Furthermore, given the importance of graph features, it is possible to dig deeper and extract graph feature information based on sentence composition, such as degree centrality, proximity centrality, and intermediate centrality. Considering the small size of the data, experiments in data amplification can also be conducted according to the transitivity and symmetry of sentence pairs. These aspects require further research in the future.

3. Related applications

3.1. Manufacturing equipment industry

After the 20th century, with the rise and development of speech recognition and generation technology, natural language understanding technology, semantic analysis technology, mobile Internet technology, and deep learning algorithm, the development of intelligent question answering is also rapid. Because intelligent question answering has 7 X 24 hours online service, fast response speed, a high number of concurrent reception, automatic replies, and other advantages, Internet, finance, medical,

DOI: 10.54254/2755-2721/18/20230964

telecommunications, manufacturing, and other fields have great application prospects. However, through the analysis of the research status of the academic circles and enterprises at home and abroad in the field of intelligent question answering, it can be seen that the typical application of intelligent question answering in various professional fields such as finance, medical treatment, manufacturing is still trying and exploring, and the representative products are few. However, with the continuous deepening of intelligent manufacturing, the demand for intelligent question-answering products in the field of equipment manufacturing has been very strong.

In order to solve the problem of high demand for intelligent question-answering products in the equipment manufacturing field, Wu's group [4] analyzed the application characteristics, interactive logic, and core components of intelligent Q&A in the field of equipment manufacturing, and explored the implementation methods of intelligent Q&A system of equipment manufacturing, considering the different business knowledge used by intelligent Q&A system in different business fields such as product design, process design, production and manufacturing, operation and maintenance support, operation and management. Based on different business knowledge, it is necessary to build different types of professional domain knowledge, and then choose single or combine multiple implementation methods from different implementation frameworks such as FQA, CQA, DQA, and KBQA to solve the complex intelligent question-and-answer application in the equipment manufacturing field.

3.2. Customer service questions and answers

Artificial intelligence technology is developing rapidly and is rapidly applied in all walks of life. The power system is related to the national economy and people's livelihood. The application of artificial intelligence in the power system will certainly effectively improve the efficiency and security of power operation and maintenance. With the design and implementation of an electric intelligent customer service system as the entry point, Lu's team applied natural language processing technology to the automatic operation and maintenance of the electric power system, making the operation and maintenance of the electric power system develop towards automation and intelligence. In addition, with the help of work order big data technology, Lu's group [5] built a dialogue understanding model that can quickly absorb and utilize existing knowledge. The questions of users are analyzed by natural language processing technology, and the question-answering model is constructed and optimized by machine learning methods to build an efficient and accurate intelligent question-answering system. In practical use, the system provides an unlimited number of users with large-scale real-time connectivity and human interaction with customers, making it easy for customers to obtain quality service and a good experience. Through this system, enterprises can efficiently complete automatic operation and maintenance work, so as to save resources and costs, and its economic benefits are remarkable.

3.3. Campus smart services

In recent years, intelligent question-answering technology in the field of artificial intelligence has made significant progress and has been widely applied to various aspects of economic development and social life [6]. Within the field of teaching and education informatization, as smart campus construction rapidly progresses, the campus information service mode is also adapting to the development trend of the intelligent era by gradually transforming from passive to active service, group to personalized service, face-to-face to online service, and manual to intelligent service. In this process, an important task is to consider how intelligent services can be developed as a driving force to solve the many inconveniences that exist in traditional manual-oriented campus consulting services.

Traditional campus consulting services primarily rely on hotlines, offline counseling windows, and e-mail counseling. These types of consulting questions are generally characterized by dispersed sources, concentrated similarity, and fixed answers. For all kinds of issues, multiple levels of departments usually need to conduct repeated manual services, which is time-consuming and labor-intensive. The quality of the replies is often uneven, and the answers are not standardized or unified enough. Particularly during special periods such as holidays or enrollment and orientation, the manual response lags behind or takes

DOI: 10.54254/2755-2721/18/20230964

a long time, making it difficult to guarantee the validity of the service. Additionally, a large amount of manual service data cannot be effectively collected and counted, and user behavior is challenging to quantitatively analyze, making it impossible to determine which aspects of service quality should be improved.

To address these issues, Guo first organized the overall structure of an intelligent question-answering system for the campus, which is divided into a data acquisition layer, knowledge base management layer, intelligent question-answering layer, and application channel layer from the bottom up. Secondly, the knowledge base management is divided into the standard knowledge base, general knowledge base, self-learning knowledge base, and greeting knowledge base to meet the daily knowledge, professional knowledge, and fun-chatting needs of students and teachers. On the intelligent question-and-answer level, Guo utilizes a new word vector technology to convert words or word-after-word segmentation into a word vector matrix representation that can be understood by computers. A classification judgment based on word vector features is made on the question intention to clarify the question intention, put it into a predefined category, and narrow the search scope.

3.4. Network teaching

Due to the outbreak of pneumonia in the new library, network teaching has become a necessary way for college students to continue to receive education at home. High-quality teaching needs to rely on a perfect network teaching system to support students' systematic learning, so the interaction of network teaching is particularly important. However, in the current mainstream network teaching platform, there are often problems such as a lack of synchronous teacher-student interaction, the continuation of interactive information, and individualized interactive answers, which makes the questions generated by students in the network learning can not solved in time, resulting in students' learning is not efficient.

In order to solve this problem, Gu's group [7] first adopted the form of a questionnaire survey to understand the needs of students and then planned to add students/teachers to the online teaching platform of colleges and universities to allow students/teachers to customize chat participants. Interactive data can be used to enter the knowledge map [8]; Set real-time alerts and unanswered message alerts; Relevant hot issue recommendations; Personalized learning map, to achieve an intelligent question-and-answer system and teaching framework cohesion and integration [8]. In view of the low efficiency of teacher-student interaction in the existing network teaching, the teaching system is difficult to achieve high quality and efficient teaching effects of the difficulties of the corresponding system design. It is mainly embodied in the integration of various data sources, knowledge bases, and network information from different sources, and the knowledge graph of the science and technology think tank is used as the search engine [9]. The knowledge graph will use the unit characteristics to complete the storage and retrieval function of data and compare all the relationships and attributes found in the knowledge graph in the BERT model. The NLP model is used for further verification [10].

4. Conclusion

The advances in artificial intelligence methods have enabled Q&A systems to efficiently answer more and more complex questions, which has been an exciting development. These new methods have opened doors for their application in a range of fields, from healthcare and education to manufacturing and finance, where they can automate and streamline various processes. However, despite the remarkable progress made in Q&A systems, they still face challenges when it comes to multi-round dialogues within complex scenarios. Current methods still struggle to provide coherent answers that take into account the context of the conversation, and this remains a critical area for improvement. The ability to engage in multi-round dialogue within complex scenarios is vital for Q&A systems to function effectively in real-world situations, and there is a growing need for research in this area. Further advancements in this field will help to unlock new possibilities for these systems and enable them to offer even more efficient and accurate solutions to complex problems.

DOI: 10.54254/2755-2721/18/20230964

References

[1] Lu Jing, Wu Chunlei, & Wang Leiqaun. (2021). Visual question answering System based on symmetrical attention mechanism. Computer System Application. http://www.c-s-a.org.cn/1003-3254/7925.html

[2] Wu Ziming, Yang Fangyu, Liang Jun, & Lei Jianbo. (2023). Questions and Answers to Chinese diseases based on Model fusion method Research on text matching method. Health Information Management in China.

[3] Zhang Kailiang & Chen Yun. (2022). Intelligent voice question answering system based on deep learning. China Computer & Communication.

[4] Xu Chengwei, Jia Xiaoxia, & Ran Qingyun. (2023). Application of intelligent question answering system in equipment manufacturing industry. Manufacturing Automation.

[5] Lu Zhiliang, Kuang Xianping, & Zeng Zhaojin. (2021). Research on Power Customer Service System based on natural Language processing. Internet and Security.

[6] Guo Li. (2023). Application research of intelligent question answering system in campus Intelligence Service. Light Industry Science and Technology.

[7] Gu Xinyun, Wan Yan, Jiang Runzhe, Zhang Yan, & Hu Cuihua. (2022). The Interactive research of network Teaching based on intelligent Question answering System. Computer Knowledge and Technology.

[8] Omar, R., Mangukiya, O., Kalnis, P., et al. (2023). ChatGPT versus traditional question answering for knowledge graphs: Current status and future directions towards knowledge graph chatbots. arXiv preprint arXiv:2302.06466.

[9] Tan, Y., Min, D., Li, Y., et al. (2023). Evaluation of ChatGPT as a Question Answering System for Answering Complex Questions. arXiv preprint arXiv:2303.07992.

[10] Gilson, A., Safranek, C. W., Huang, T., et al. (2023). How does CHATGPT perform on the United States Medical Licensing Examination? The implications of large language models for medical education and knowledge assessment. JMIR Medical Education, 9(1), e45312.