

Factoid Question and Answering System

Aditya SRM , S.R Rajeswari, Dinesh Reddy, Varshini

Abstract: “The Natural Language processing (NLP) can be said as a form of artificial intelligence (AI) which is used to process the natural language data such as text, image, video, and audio. It acts as a tool for the computer to understand and analyse the real-time data in human language which the humans speak. Our system answers the factoid questions over texts, images using neural networks along with tensor flow framework. In order to justify the retrieved answer reasoning is also used. Reasoning is the process of analysing data in a logical way to make decisions. In question and answering system reasoning plays an important role for extracting the answers with better accuracy and precision”.

Keywords: NLP, AI, CHATBOT, FUZZY.

I. INTRODUCTION

The Question Answering System is an answering system which is a sub field of computer science, Artificial Intelligence and natural language processing. In this the main focus is on building question Answering systems that automatically answer questions asked by humans in natural language. By this the system is able to understand natural language and is capable of translating sentences and help this system to generate valid answers to questions asked by user. Valid answers are the answers which are appropriate to the questions posed by the user. For forming an appropriate answer, we have to execute the syntax analysis of the given question along with semantic analysis. The “CHAT BOT” Project is an open-source question and answering system for answering Factoid Questions using NLP. It contains classes to represent questions and answers, annotations, and it is able to read and write to a variety of file formats. Separate modules extend capabilities to question alignment, question structure, sequence, and machine learning. Chat Bot is one of a number of projects designed to answer questions. This project proposal is to make a system for answering people questions and their access to a lot of answers by using NLP. Significant part of the “Chat Bot” code base has to be modified to adapt the new question types.

Manuscript published on 30 April 2019.

* Correspondence Author (s)

S.R Rajeswari*, Assistant Professor SRM Institute of Science and Technology

Aditya SRM Institute of Science and Technology

Dinesh Reddy, SRM Institute of Science and Technology

Lakshmi Priya, SRM Institute of Science and Technology

Varshini, SRM Institute of Science and Technology

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

This not only improves the efficiency of software, but also provides easy answers to various Factoid Questions. These are the questions that the regular question and answering systems fails to answer (Factoid Questions).

II. LITERATURE SURVEY

1) International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) 2013 Research and reviews in question answering system. In this paper the author tells about computational intelligence, modelling techniques and the methods needed for retrieving the answer.

2) Zheng Answer Bus question answering system. In: The Second International Conference on the HLT Research, pp.399-405. Morgan Kaufmann Publishers Inc. (2002). It is paper on question and answering system which deals with query processing and answer generation.

3) Department of computer science engineering and neural Sciences at the University of California, San Francisco. This paper deals with neural science and psychology of humans and the methods required to understand the natural language.

4) Department of Computer Science, B.B.A. University Lucknow, Uttar Pradesh, 226025, India. In this the authors deal with computer science domain and the method of using Artificial Intelligence to process a user query and answer it in a most efficient way.

III. IMPLEMENTATION

The implementation includes the NLP, AI. Our project implementation starts with the processing of Natural language processing. AI is used to make the systems better at answering questions.

Keyword Extraction: -

Keywords are the words that are used to describe the type and content of a question. Keywords are words which may be in question or any other words related to the question. These Keywords will be used to retrieve data and paragraphs which may contain answer to the user's question. In other words, wrong keywords result in not retrieving the correct answer. Keywords extraction is done by looking at each word of the question. In the process of extracting the keywords, we also remove the unnecessary words from the question called stop words. Stop word list that is used in this system obtained from the internet and it has been modified by removing and adding some of the words. Stemming is the process of getting the keywords from the given question.



Factoid Question and Answering System

Passage Retriever: -

Corpus is used in Question Answering System to search answer to the given question. The corpus must be documents in human language and should include only a single topic/domain. There are two main alternatives of language corpus that can be used in terms of accessing methods: 1. Offline e.g. articles from electronic media. This corpus is stored in text files in databases. 2. Online corpus e.g. collection of articles from web. We should know how to access the articles and the parts of the article in this corpus. From these alternatives along with offline corpus obtained from Wikipedia we will be able to answer the user questions. However, there are still some weaknesses in this QA system which cannot answer Factoid. Therefore, we made a few changes to the existing system and enhanced the capabilities of the QA system so that the system is able to understand the natural language and will be able to answer Factoid Questions which the previous QA systems failed to do.

Searching Technique: -

The searching technique which we use for answer retrieval is key word analysis. In this the user question broken into many keywords. After that a lexical chain is created based on the created keywords. The system analysis all the keywords and creates a list of related answers. There will be a limit for frequently asked questions if the user asked question meets the FAQ limit, it is added to FAQ else the question will be processed based on the keywords. The lexical chain is used to retrieve the exact answer what the user is looking for. AI, NLP is also used in answering retrieval

IV. ARCHITECTURE DIAGRAM

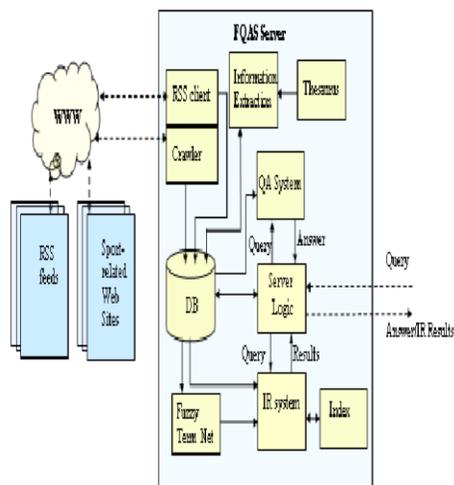


Fig: Architecture of the System.

The system architecture contains data base, query, answer extraction, server logic, question answering system, fuzzy team, IR system. The implementation of the product starts with the user asking a question to the QA system. The QA system analyses the question, process it and retrieve the answer. The answer is given to the user.

V. RESULTS

The results are the answers which the user gets back after asking the query is asked. The user asks QA system a query and the system divides the type of question, process the question and give back the answer. The answer retrieval depends on many factors like psychology of the user, type of the question and many other factors. Thus our system understands the natural language questions and give the answer with high accuracy and precision.

VI. CONCLUSION

Thus by this project, we strongly believe that we have developed a question and answering system which interacts with humans in natural language. It also overcomes the drawbacks of the previous QA systems. Hence our QA system will be able to answer even the factoid questions which the previous QA systems failed to do.

VII. FUTURE ENHANCEMENTS

The future scope of the project is to develop a QA system which adopts to natural language. The code will be modified so that it can answer the Factoid Questions with high precision and accuracy. As the time progress, our system gets better as it analyses all the question types and the context in which they will be asked. It helps in exact answer retrieval instead of some approximate answers. Thus this overcomes all the drawbacks which the previous QA. The performance of QA systems is affected by some factors like Psychology, language of the user who is asking the question etc. In the future QASs should understand natural language and should be able to answer Factoid Questions. The QA systems should be able to satisfy the users with correct answers instead of some related answers. As a result the systems gets better with time and the user gets better answers.

REFERENCES

1. Mengqiu Wang, School of computer Science, Carnegie Mellon University.
2. Chen Qu, Liu Yang, W. Bruce Croft, University of Massachusetts Amherst.
3. Department of Computer Science, B.B.A. University (A Central University) Lucknow, Uttar Pradesh, 226025, India.
4. The Department of Neural Studies of the University of Tubingen.
5. Department of computer science engineering and neural Sciences at the University of California, San Francisco.
6. Department of Computer Science and Engineering, Bilkent University, Ankara, Turkey.
7. Department of Computer Science and Engineering, Jabalpur University, India.
8. Zheng Answer Bus question answering system. In: The Second International Conference on HLT Research, pp.399-404. Morgan Kaufmann Publishers Inc. (2002).
9. Nagehan Pala Er Department of Computer Engineering, Bilkent University, Ankara, Turkey
10. International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) 2013 Research and reviews in question answering system.