KNOWLEDGE GRAPH BASED QUESTION ANSWERING SYSTEM

Thisari Patabendi

A dissertation submitted in partial fulfillment for the requirement for
Master of Science Big Data Analytics

Department of Computing
Informatics Institute of Technology, Sri Lanka
in collaboration with
Robert Gordon University

2022

Abstract

Answering natural language questions using a knowledge base is a crucial and difficult task with several applications in natural language processing and information retrieval. The majority of question-answering systems use a method called machine reading comprehension and those are mostly closed domain. However, few contemporary knowledge-based questions answering systems make use of complicated end-to-end neural network techniques that are computationally expensive and time consuming to train. Furthermore, an end-to-end method makes it impossible to examine the query processing process. A non-supervised strategy will be employed in this study, which will decompose the question answering problem into a three-step pipeline of entity-relationship recognition, automatic query formulation, and answer extraction.

For the implementation and testing LC-QuAD 2.0 dataset was used as it's tagged with respective entities, relationships and queries which is helpful to validate the results. Entity relationship identification and automatic query construction are the challenging tasks in the implemented system. However, the project was able to achieve 71% accuracy while giving users the ability to get answers quickly and also allowing interested researches to view the analytics. The final prototype of the system is implemented as a web application.

Key words:

Natural Language Processing, Knowledge Graph, Text processing, SPARQL query, Answer Extraction, Entity relationship recognition