NLP Analysis Study Literature Survey and Proposal

Hayden Kwok

May 2024

1 Literature Survey

Question Answering (QA)

Task and Objective

Natural Language Processing (NLP) is a sub-field of Artificial Intelligence (AI) that focuses on developing computer technology to comprehend, interpret, and generate human language, enabling machines to interact with and process natural language data. The branch utilizes a large span of techniques and algorithms that are applied in tasks such as machine translation, sentiment analysis, as well as text summarization. Here, I will be analyzing the scope of the subtopic of Question Answering (QA), noting its functionality, and exploring its limitations.

At a high level, the task of QA involves the development of systems and algorithms that automatically generate responses to user-generated questions in natural language. The objective of these algorithms is to effectively retrieve precise information within prompts to facilitate efficient interactions between humans and computers. In the next section, we will explore a selection of papers that discuss the applications of this task.

QA-GNN: Reasoning with Language Models and Knowledge Graphs for Question Answering

The paper showcases a new NLP model called QA-GNN[1] that is trained from existing language models (LMs) and knowledge graphs (KGs). The model addresses two major challenges: extracting pertinent knowledge from large KGs and jointly reasoning over the QA context and KG. QA-GNN links the QA context and KG to construct a joint graph, updating their representations through graph neural networks (joint reasoning). In addition, it uses LMs to evaluate the importance of KG nodes in the provided QA context (relevance scoring). The authors assess QA-GNN using biological and commonsense QA datasets. The results of the preliminary tests show that QA-GNN performs better than the current LM and LM+KG models and can carry out structured and interpretable reasoning, including answering questions that negate themselves.

GraghVQA: Language-Guided Graph Neural Networks for Graph-based Visual Question Answering

The paper goes over GraphVQA[3], a language-guided graph neural network framework for graph-based visual question answering. It represents images as scene graphs, where objects are encoded as nodes and their pairwise relations as edges. GraphVQA translates and executes natural language questions as multiple iterations of message passing among the graph nodes. Using the model, tests run on the GQA dataset show that GraphVQA outperforms state-of-the-art models by a large margin (88.43% vs. 94.78%). Overall, the paper explores the design space of the GraphVQA framework and discusses the trade-offs of different design choices.

A Survey for Efficient Open Domain Question Answering

This study provides an extensive overview of recent developments in the effectiveness of models for answering open-domain questions (ODQAs)[2]. ODQA is a task in NLP that aims to answer factual questions from a large knowledge corpus without explicit evidence. Although the majority of recent work has been on increasing answering accuracy, the authors point out that there is a trade-off between higher accuracy and a resulting higher memory usage and inference delay (higher is less efficient). As a result, the authors concentrate on examining how accuracy, memory usage, and processing speed are optimized in ODQA models. In addition to providing a quantitative examination of memory cost, query speed, accuracy, as well as overall performance comparison, the paper examines the fundamental approaches for achieving efficiency in ODQA.

2 Project Proposal

After a comprehensive review of the previous papers, I intend to analyze the limitations of the model within the QA-GNN[1] paper, in which the associated codebase can be found at michiyasunaga/qagnn. The repository uses Transformers, PyTorch, Numpy, and Tqdm libraries within a series of Python scripts to run the model. It contains the CommonsenseQA (CSQA) and OpenBookQA (OBQA) datasets and their corresponding knowledge graph datasets. The CSQA dataset involves questions that require commonsense knowledge to answer given multiple-choice answers, where each question is subsequently linked to concepts from a commonsense knowledge graph. OpenBookQA is a dataset designed to test open-book question-answering abilities. The model has access to a set of scientific facts ("open book") and is required to combine the given facts with additional commonsense knowledge to answer multiple-choice questions.

The original paper utilizes Exact Match (EM) and F1 Score to evaluate the model. However, the results of using these metrics identified some potential limitations in the current evaluation approach. I will address this by analyzing the model using Semantic Answer Similarity (SAS) and Mean Reciprocal Rank

(MRR). The EM and F1 scores often underestimate the true performance of the model as it lacks semantic examination between the ground truth and predicted answers. Furthermore, MMR allows us to rank the positions of the correct answer among the model's predictions, which could provide insights into the model's ability to prioritize the most relevant knowledge for answering questions.

Overall, this literature survey has highlighted advancements in Question Answering (QA) systems, including integration with knowledge graphs and graph-based techniques. My project proposal aims to analyze the limitations of the QA-GNN model's abilities to identify semantic relationships and order answers by rank.

References

[1] Antoine Bosselut Percy Liang Jure Leskovec Michihiro Yasunaga, Hongyu Ren. Qa-gnn: Reasoning with language models and knowledge graphs for question answering. 2021.

[2] Dongkuan Xu Qingqing Cao Xiaojun Chen Trevor Cohn Meng Fang Qin Zhang, Shangsi Chen. A survey for efficient open domain question answering. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2023.

[3] Zixuan Liu Weixin Liang, Yanhao Jiang. Graghvqa: Language-guided graph neural networks for graph-based visual question answering. In *Proceedings of the Third Workshop on Multimodal Artificial Intelligence*, 2021.