Journal of Cases on Information Technology
Volume 24 • Issue 3

Analyzing Linguistic Features for Answer Re-Ranking of Why-Questions

Manvi Breja, National Institute of Technology, Kurukshetra, India
Sanjay Kumar Jain, National Institute of Technology, Kurukshetra, India
https://orcid.org/0000-0003-1999-5530

ABSTRACT

Why-type non-factoid questions are ambiguous and involve variations in their answers. A challenge in returning one appropriate answer to users requires the process of appropriate answer extraction, re-ranking, and validation. There are cases where the need is to understand the meaning and context of a document rather than finding exact words involved in question. The paper addresses this problem by exploring lexico-syntactic, semantic, and contextual query-dependent features, some of which are based on deep learning frameworks to depict the probability of answer candidate being relevant for the question. The features are weighted by the score returned by ensemble ExtraTreesClassifier according to features importance. An answer re-ranker model is implemented that finds the highest ranked answer comprising largest value of feature similarity between question-and-answer candidate and thus achieving 0.64 mean reciprocal rank (MRR). Further, the answer is validated by matching the answer type of answer candidate and returning the highest-ranked answer candidate with matched answer type to a user.

KEYWORDS

Answer Candidate Extraction, Answer Re-Ranking, Contextual, ExtraTreesClassifier, Feature Importance, Lexical-Syntactic, Natural Language Processing, Semantic

INTRODUCTION

The advent of IBM's Watson (IBM Watson, 2020) has shown remarkable results in answering open-domain questions. Watson is now no longer treated as only a question answering (QA) system rather it also has ability to sense. Research in question answering domain has achieved high accuracy around 85% in answering factoid-type questions. However, today researchers are motivated to go beyond factoid QA, addressing non-factoid question answering such as 'why' and 'how' type questions. Some of the work from Verberne et al. (2010), Jansen and Surdeanu (2014), Fried and Jansen (2015), Oh et al. (2012, 2013) has been successful in answering open-domain non-factoid questions whereas Tran and Niederee (2018) has investigated deep learning frameworks for answering insurance and financial domain non-factoid questions but still performance is lower than factoid QAS such as IBM Watson.

The question answering system presents an accurate answer satisfying the need of user. Answering why-type questions is complex and the need is to tackle the complexity because of ambiguity and redundancy involved. The paper is contributed towards extracting answer candidates to a question by finding cue phrases reflecting cause-effect relations between terms in retrieved passages. Further

DOI: 10.4018/JCIT.20220701.oa10

an answer re-ranker is developed exploring the set of features based on similarity between question and answer candidates, weighted by feature importance scores. The method is able to achieve 0.64 Mean Reciprocal Rank (MRR) which significantly improves over other previous research works in why-type answer re-ranker.

The remaining paper is structured in the sections where Section 2 explores the previous work in answer re-ranking for why-type questions. Section 3 provides main focus of article describing issues and their solutions. Section 4 describes the system architecture utilized for research. Section 5 highlights the data which is setup for answer re-ranking. Section 6 presents features employed for re-ranking answer candidates and their relevance. Section 7 puts light on the algorithm proposed to weigh each feature set based on the importance of each feature. Section 8 briefs the algorithm used for answer validation process. Section 9 highlights the implementation details in Python. Section 10 compares our proposed work with other previous research works. Finally Section 11 concludes the work with future research directions.

BACKGROUND

A considerable work has already been done in addressing non-factoid type questions and improving answer re-ranker module. This section discusses major contributions in answer re-ranking of English and Japanese non-factoid questions.

Referring to English non-factoid QAS, Verberne et al. (2010) applied various machine learning techniques for ranking answer candidates. The authors explored linguistic features comprising tf-idf, syntactic overlaps, WordNet synsets, cue terms, common words in question & document title and WordNet relatedness. Learning to rank approaches categorized as pairwise, pointwise and listwise (Liu, 2011) were applied with their default hyperparameter settings where Support Vector Regression with its hyperparameter tuning performs best with MRR 0.350. Although the authors provided a good baseline for re-ranking answers but the need is to explore more semantic and contextual features with assigning a weight to each feature. Surdeanu et al. (2011) exploited non-factoid QA pairs from social QA sites, further trained answer re-ranker model by extracting various features comprising similarity (BM25), translation (IBM's Model 1), density/frequency, and web correlation (query-log correlation using PMI and chi-square) features using Perceptron and SVM-rank model to achieve 0.6416 & 0.6381 MRR respectively. Since QA pairs were retrieved from social QA sites, the authors could have considered more features such as number of votes, genres of QA pairs, user comments and answers rating that will help instigate ranking answers. Jansen et al. (2014) integrated lexical semantics with shallow and deep discourse features. The model was trained on open-domain Yahoo! Answers corpus comprising how-type QA pairs and Biology Textbook corpus comprising both how and why-type questions using SVM Rank thus achieved 26.57 P@1 and 49.31 MRR. Molino and Aiello (2014), Fried et al. (2015) have trained answer re-ranker module with lexical-semantic models on dataset of non-factoid how-type questions to achieve 0.7909 MRR and 0.5396 MRR respectively. The authors have significantly addressed the issue of answer re-ranking by learning word representations and finding direct & indirect associations between QA pairs. Tran and Niederee (2018) utilized deep learning frameworks for re-ranking answers of non-factoid questions from insurance & financial domain, achieving 0.616 MRR using SRanker mlp, 0.606 MRR using SRanker bilinear and 0.653 MRR using CARanker.

Considering the research in Japanese non-factoid QAS, Higashinaka and Isozaki (2008) represented answer candidates by causal expressions, content similarity between Q&A and causal relations representing cause & effect. These feature sets are utilized to train Answer Re-Ranker with RankingBoost and SVM rank thus achieving top 5 MRR as 0.305. The authors only considered 'cause' relation representing causality, other relations such as 'purpose', 'condition' need to be further explored. Oh et al. (2012) trained and tested answer re-ranker using TinySVM with features combining morphological and syntactic analysis, semantic word classes based on n-grams and sentiment analysis

Related Content