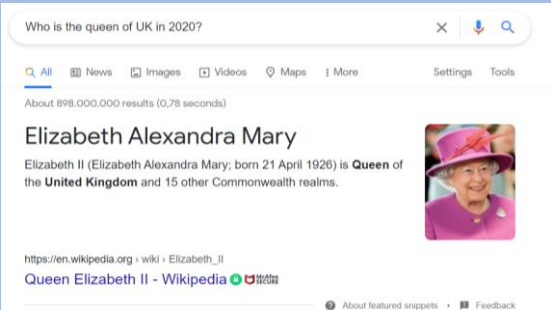


Query Answerability Classifier for Direct Answer Module in Web Search Engines

Yiran Wang: y.wang-71@student.tudelft.nl; Supervisor and Responsible Professor: Claudia Hauff

1. Background

1. Direct Answer Module



2. Machine Reading Comprehension

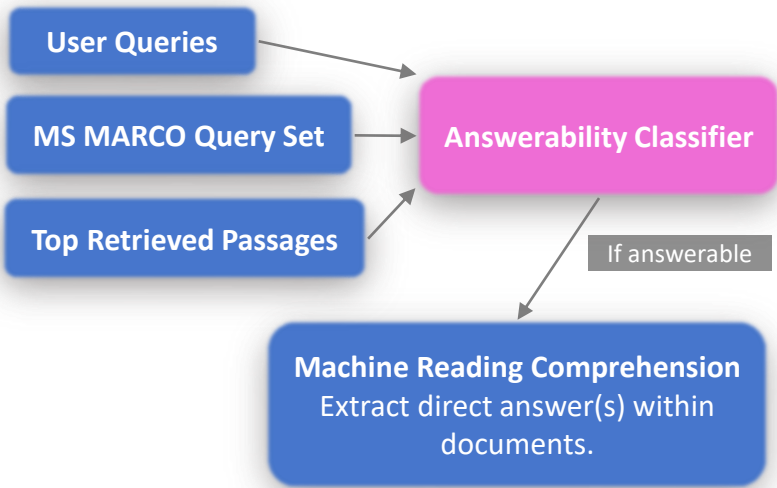
```
Interactive DrQA
>>> process('What is the meaning of life?')
Top Predictions:
+-----+-----+-----+-----+
| Rank | Answer | Doc | Answer Score |
+-----+-----+-----+-----+
| 1 | freedom from suffering | Meaning of life | 3.8928e+05 |
+-----+-----+-----+-----+

Contexts:
[ Doc = Meaning of life ]
Zeno of Citium, a pupil of Crates of Thebes, established the school of Stoicism. According to Zeno, the goal of life is to live in accordance with nature, which means to live in accordance with reason and virtue. The meaning of life is "freedom from suffering" through "apatheia" (indifference to external things) and having "clear judgement", "not" indifference.
```

2. Research Question

How to determine the answerability of queries, which only ask for answers in text?

3. Methods



For developing the answerability classifier:

1. Dataset: Labelled 500 queries from MS MARCO QnA dataset [1].
 1. Answerable queries are human-comprehensible, with sufficient details, generalizable, and text-based.
2. Two Models: Random Forest and XGBoost.
3. Six Features:
 1. **Query dependent features:** GloVe Embedding, number of words, existence of key words.
 2. **Query-corpus dependent features:** Averaged word/ngram-level tf-idf of each query.
 3. **Top-retrieved-passages dependent features:** Fraction of passages containing all the words.

4. Results

Model	The ratio of answerable questions	Prediction Accuracy	F1	The ratio of unanswerable queries identified
Random Forest	0.75	0.78	0.87	0.13
XGBoost	0.75	0.83	0.89	0.35

5. Conclusion

1. XGBoost performs better than Random Forest model in general.
2. High overall prediction accuracy, worse in picking unanswerable queries.

6. Future Work

1. Integrate with MRC to test the whole pipeline.
2. Extend to accept non-text answers.

[1] Payal Bajaj, Daniel Campos, Nick Craswell, Li Deng, Jianfeng Gao, Xiaodong Liu, Rangan Majumder, AndrewMcNamara, Bhaskar Mitra, Tri Nguyen, Mir Rosenberg, Xia Song, Alina Stoica, Saurabh Tiwary, and TongWang. 2016. MS MARCO: A Human Generated Machine Reading Comprehension Dataset. arXiv e-prints, Article arXiv:1611.09268 (Nov. 2016), arXiv:1611.09268 pages. arXiv:1611.09268 [cs.CL]