**Introduction**

Example complex open-domain question:

“How many people does the EMNLP-IJCNLP 2019 venue hold?”

We present:

- **GoldEn** (Gold Entity) Retriever, an iterative retrieve-and-read model that performs explainable open-domain multi-step reasoning.
- An efficient method for training components that generate natural language queries to retrieve supporting facts.
- Competitive performance without using powerful pretrained neural models like BERT.

<table>
<thead>
<tr>
<th>Efficient</th>
<th>Multi-hop</th>
<th>Explainable</th>
</tr>
</thead>
<tbody>
<tr>
<td>✔️</td>
<td>☒️</td>
<td>✔️</td>
</tr>
</tbody>
</table>

**Previous Work on Open-domain QA**

- One-step retrieve-and-read (e.g., DrQA)

**Language Note**

The HotpotQA dataset we tested on is available in the English language only, thus that is the only natural language our experiments are in. But our principle of semantic overlap is applicable to answering open-domain complex questions in other languages than English if suitably augmented with lemmatization for highly inflected languages.

**Experimental Results**

We test our GoldEn Retriever system on the HotpotQA dataset, which contains ~113k questions requiring two Wikipedia articles to answer. We focus on the fullwiki setting, where the QA system is given a question and ~5 million Wikipedia articles to answer from.

We retrieve 10 paragraphs in total for each question (5 in each retrieval step).

**Acknowledgements**

This research is funded in part by Samsung Electronics Co., Ltd. and in part by the SAIL-JD Research Initiative.