Motivation

- Post every flight, need to address problems (if any), faced and recorded by pilots.
- The maintenance people need to consult manuals which are voluminous.
- This process consumes a lot of time. Sometimes manuals are overlooked, and the incorrectly detected faulty part is replaced (false positive). A more dangerous situation is a false negative.
- A search-engines-like system would quickly bring up relevant information from aircraft manuals, identify the cause, save time & effort, and reduce maintenance costs.

Introduction

- We contribute an Aviation Knowledge Graph, constructed from accident reports using the domain knowledge and information extraction techniques to the community.
- We propose a Question Answering system that answers the questions asked in natural language from two modules: KG-based QA (KGQA) module, which uses a pipeline to convert Natural Language Queries to SPARQL queries and fetches responses from the Aviation Knowledge Graph. DL-based QA (DLQA) module that extracts answers from the plain text in the documents. The DLQA module has been tested using two different QA models, BERT-QA and GPT3-QA.

We show that a combined Question Answering system, such as ours, outperforms the individual Knowledge Graph and Deep Learning methods on our curated test set.

- The combined system KGQA+BERT-QA attains 7% and 40.3% increase in accuracy over KGQA and BERT-QA modules respectively.
- KGQA+GPT3-QA attains 29.3% and 9.3% increase over KGQA and GPT3-QA modules respectively.

Problem Statement

Given a query, retrieve relevant information from a collection of documents using KG and DL. The question (input) is in natural language, and the output should be the solution to the input. The result can be a direct answer from KG or a fragment of text from the DL model.

Data Repositories

- The NTSB store investigation reports of civil aviation accidents in the US. It captures all the aspects of an accident, namely aircraft specifications, pilot details, environmental state, and a comprehensive description of the suspected cause of the accident.
- The ADREP taxonomy contains a complex multi-level hierarchy of factual descriptions (time, place, aircraft models, engine, component manufacturers, etc.).

Analysis

- GPT3-QA is the best performing model among the individual models owing to its large size and extensive pre-training.
- KGQA+GPT3-QA is the best model for the QA task. It achieves an increase of 9.3% in answer retrieval accuracy over GPT3-QA.

Case Study

- What discrepancy was noted due to which flight landed at La Belle Municipal Airport?
- Which accidents involved aircraft operated by Honeywell?

Table 1: Properties of Aviation Knowledge Graph

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Count (#)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entity Class</td>
<td>239</td>
</tr>
<tr>
<td>Individual</td>
<td>8894</td>
</tr>
<tr>
<td>Object Property</td>
<td>300</td>
</tr>
<tr>
<td>Data Property</td>
<td>71</td>
</tr>
<tr>
<td>Axioms</td>
<td>97829</td>
</tr>
<tr>
<td>Part of Relation among Classes</td>
<td>494</td>
</tr>
<tr>
<td>Property of Relation among Classes</td>
<td>353</td>
</tr>
</tbody>
</table>

A total of 120 SPARQL queries with gold answers of different categories were tested for the KG evaluation, where our KG answered 83 questions.

Results

- We have built a Knowledge Graph guided Deep Learning system.
- We compared the KGQA model with two individual models, KGQA and BERT-QA, as well as a baseline BM25-BERT.
- The KGQA+GPT3-QA model answers both the above questions correctly.
- Table 1 shows the evaluation results of answers and passages predicted by the Question Answering modules. The models are the results of our curated test set created from the NTSB reports.

Figure 1: Aviation KG Construction Process from NTSB reports with assistance from ADREP taxonomy.

Figure 2: Knowledge Graph guided Deep Learning based Question Answering system.

Figure 3: NL2SPARQL pipeline for KGQA system.

Conclusion

- We have successfully created an Aviation Knowledge Graph from NTSB aircraft accident reports to help experts in their study.
- We have developed the KGQA model on our Aviation KG and DLQA model on the NTSB reports.
- We have built a Knowledge Graph guided Deep Learning based Question Answering system, which outperforms the individual KGQA and DLQA systems.
- The dominance of the combined QA system is proved theoretically and empirically by evaluating it on our handcrafted test set.

Future Work

We plan to expand our knowledge scope in aviation safety and make our KGQA + DLQA system robust.

Resource & Code Repository