Project Final Report

UIUC: Text Information Systems

Group Name: Team Coco

Theme: Free Topics

Specific Topic: Question-and-Answer Chatbot

Captain/Project Leader (and sole team member): Joel Haas, joelah2@illinois.edu

---

Project Task

Configure a question-and-answer "chatbot"

Project Update

I was not able to successfully install the AllenNLP python library. I spent 3 hours researching the AllenNLP capability. I then tried to install the AllenNLP library so I could start learning the library. However, after 2 hours troubleshooting and researching the installation challenges that others have also had trying to install the library, I decided I needed to move on. So, I configured a question-and-answer "chatbot" using native python libraries rather than the AllenNLP python library.

Programming Language:	Python
Tools:	Scikit-Learn, NLTK, Pandas
Systems:	Personal laptop
Datasets:	Deepmind NarrativeQA Reading Comprehension https://github.com/deepmind/narrativeqa
Expected Outcome:	Predict the correct answers to the questions being asked
Evaluation Methodology:	Compared the algorithm's predicted answers against the "right" answers given in the annotated question and answer dataset
Software Code:	Three Jupyter notebooks provided: PrepQuestions, PrepSummaries, TFIDF+CosineSimilarity
Documentation:	Included in README.md and throughout each Jupyter notebook

Completed Tasks by Workload (26 hours):

3 hours - AllenNLP research
2 hours - Troubleshooting installation of allennlp python library
2 hours - Research labeled question / answer datasets
6 hours - Cleaning and preparing datasets

9 hours – Researching, configuring and troubleshooting TF-IDF and cosine similarity algorithms
2 hours – QA TF-IDF and cosine similarity results
2 hours - Documentation

Challenges

Preparing (wrangling) the datasets proved more challenging than expected. This includes parsing and cleaning the provided datasets, and then getting the data into the correct format or the TF-IDF and cosine similarity algorithms.

Have I completed what I planned?

The original plan was to configure a deep learning NLP pipeline using the AllenNLP platorm. Unfortunately, I was not able to successfully install the AllenNLP python library. I spent 3 hours researching the AllenNLP capability. I then tried to install the AllenNLP library so I could start learning the library. However, after 2 hours troubleshooting and researching the installation challenges that others have also had trying to install the library, I decided I needed to pivot. My final project consisted of configuring a (rudimentary) question and answer chatbot using the Google Deepmind NarrativeQA Reading Comprehension project files and native python libraries.

Have you gotten the expected outcome?

Yes and no. Out of the 1040 predicted answers, I manually QA'd 267 predicted answers with each respective annotated "right" answer as provided by Google. 95% of the predicted answers were generated from the correct Wikipedia article. 41% of the predicted answers were correct answers.

Although having an accuracy of 41% in predicting the correct answer to each question, this approach is clearly not sophisticated enough for a RC system. Additionally, the correctly predicted answers returned the entire sentence that included the answer, rather than curating the answer and returning only the portion of the sentence that answered the question most directly.