NLP MAPS
Capstone Project

End of Quarter Update

Andrew Simon, Shijie Zhang, Prithviraj Chumble,
Meenal Rawlani

Roadmap

- T
- Updated flowsheet
- Updated Github

- Summary of Literature Review
- Preliminary criteria for evaluating models

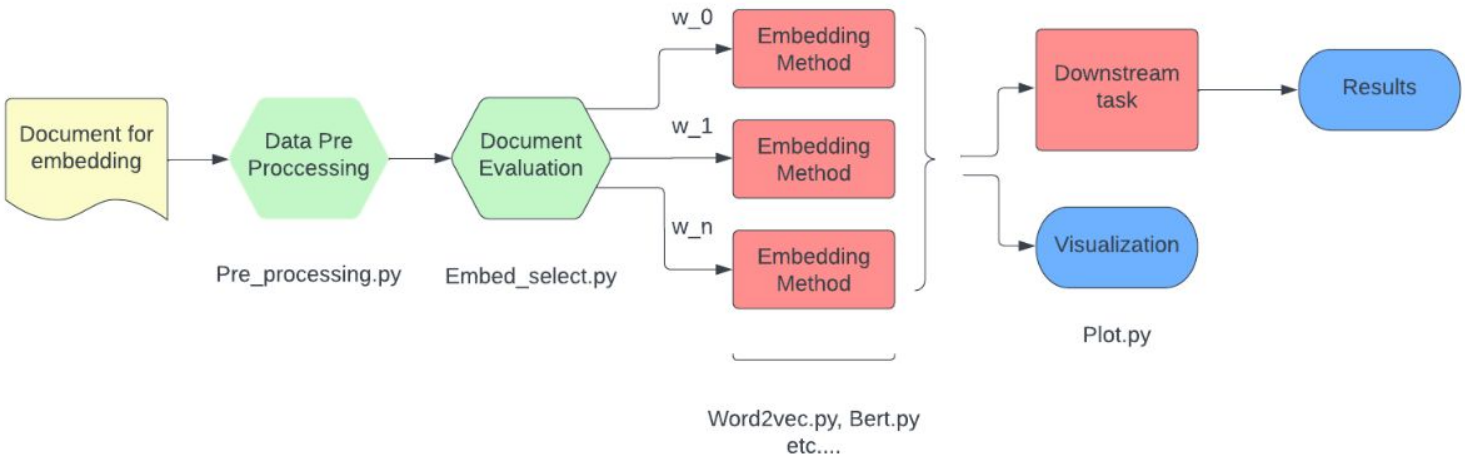
- Models Selection heuristics
- Model Results

- Plans for the future
- Questions

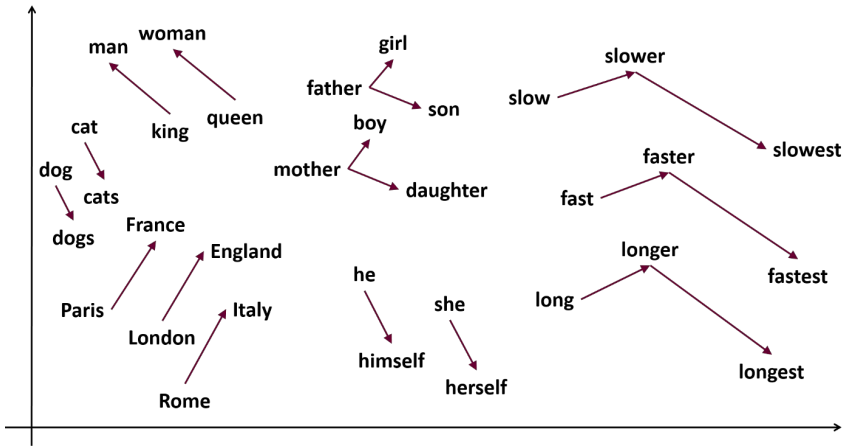
Summary of Progress

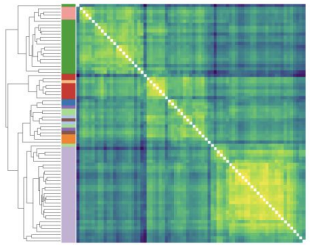
- Updated Github repo to reflect architecture of final version
- Researched criteria for evaluating embedding tasks
- Started training models on the IMDB dataset
- Evaluated models accuracy score

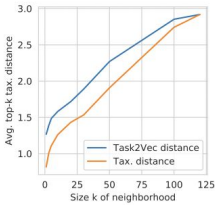
Updated Flow Sheet

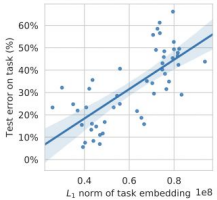


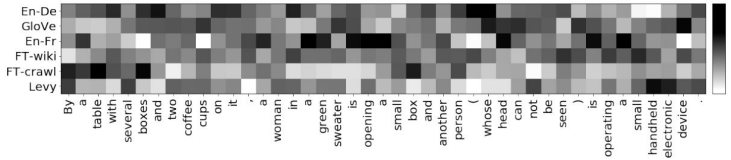
Visualizations



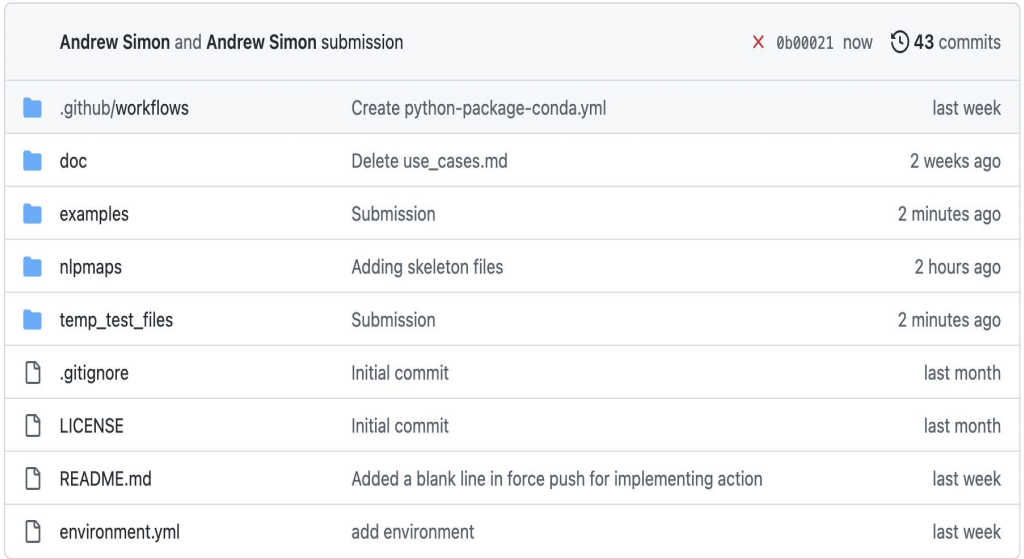






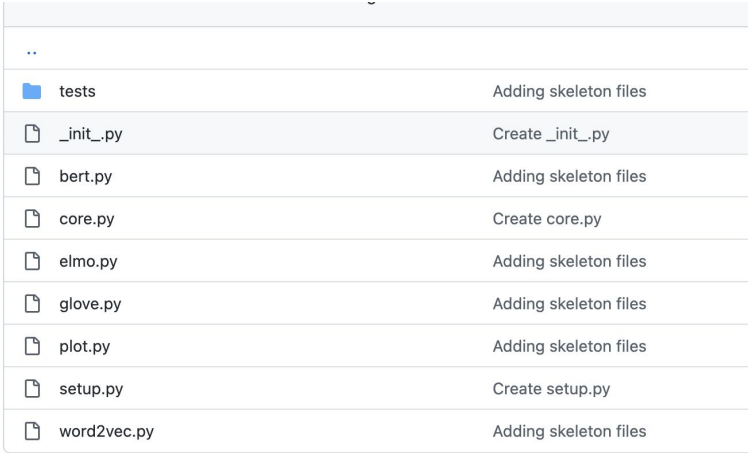


Current Github Repository



Andrew Simon and **Andrew Simon** submission — 0b00021 now — 43 commits

.github/workflows	Create python-package-conda.yml	last week
doc	Delete use_cases.md	2 weeks ago
examples	Submission	2 minutes ago
nlpmaps	Adding skeleton files	2 hours ago
temp_test_files	Submission	2 minutes ago
.gitignore	Initial commit	last month
LICENSE	Initial commit	last month
README.md	Added a blank line in force push for implementing action	last week
environment.yml	add environment	last week



..

tests	Adding skeleton files
init.py	Create _init_.py
bert.py	Adding skeleton files
core.py	Create core.py
elmo.py	Adding skeleton files
glove.py	Adding skeleton files
plot.py	Adding skeleton files
setup.py	Create setup.py
word2vec.py	Adding skeleton files

Literature Review

- Literature review was used for
  - Finding common NLP embedding methods
  - Deciding on evaluation criteria
  - Inspiration for visualization methods
  - Researching meta-embedding methods

Criteria for evaluating models

Task-Optimized Word Embeddings for Text Classification Representations

- AUC for evaluating efficiency

ETNLP: a visual-aided systematic approach to select pre-trained embeddings for a downstream task

- Word Analogy Task
- NER Task

Evaluating Word Embedding Models: Methods and Experimental Results

- Word Similarity
- Word Analogy
- Outlier Detection
- QVEC

Criteria for evaluating models

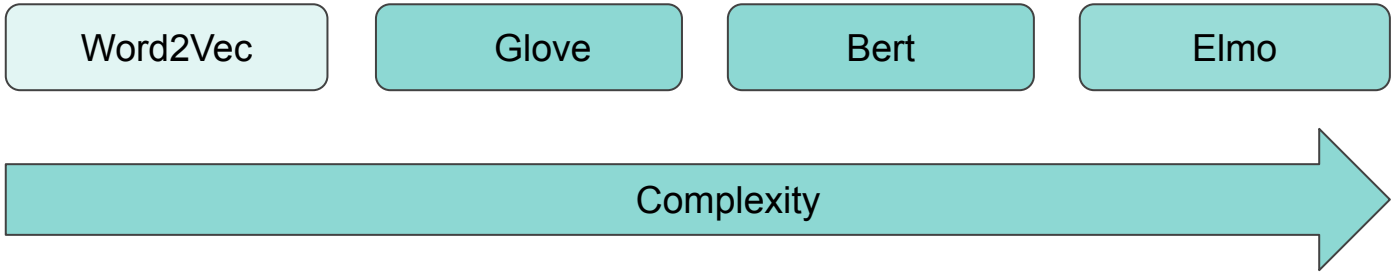
Evaluation techniques were:

- Highly complex and require more time to fully implement
- Loosely explained in the paper and require further literature review
- Significant dependency issues (QVEC)

For the purpose of generating preliminary results, our group decided to move forward with accuracy scores for sentiment analysis

Models Selection Heuristics

- Sought diverse models ranging in complexity
- Well documented training methods
- Easily trainable for Sentiment Analysis



Summary of Results

Method	Accuracy Score
Word2Vec	Na
Bert	0.85
Elmo	Na
GLove	0.76

- Substantial difficulties with training elmo, word2vec resolving the issue this weekend
- Word2Vec underperforms significantly, looking to tune hyperparameters in the future

Plans for the future

- Complete selection algorithm
- Add more models
- Establish a more robust selection criteria
- Incorporate meta embedding
- Add test files
- Train data on a downstream task

Questions

- What will the primary task of the model be for Dow data?
- How to improve current methods
-