NLP MAPS Capstone Project

End of Quarter Update

Andrew Simon, Shijie Zhang, Prithviraj Chumble, Meenal Rawlani

Roadmap

- Our Project
- The need for NLP

- Project Architecture
- Summary of progress

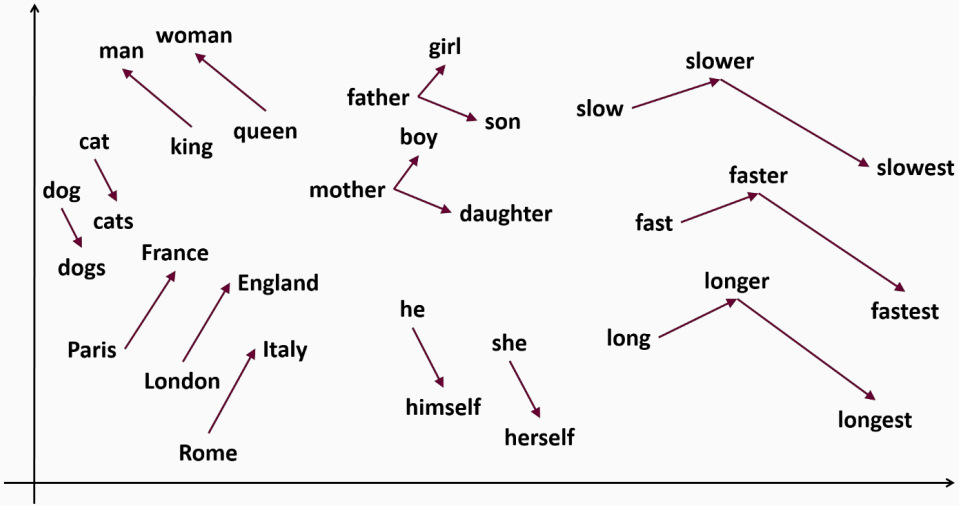
- Models Evaluation Criteria
- Selection Heuristics

- Model Results
- Plans for the future

Our Project

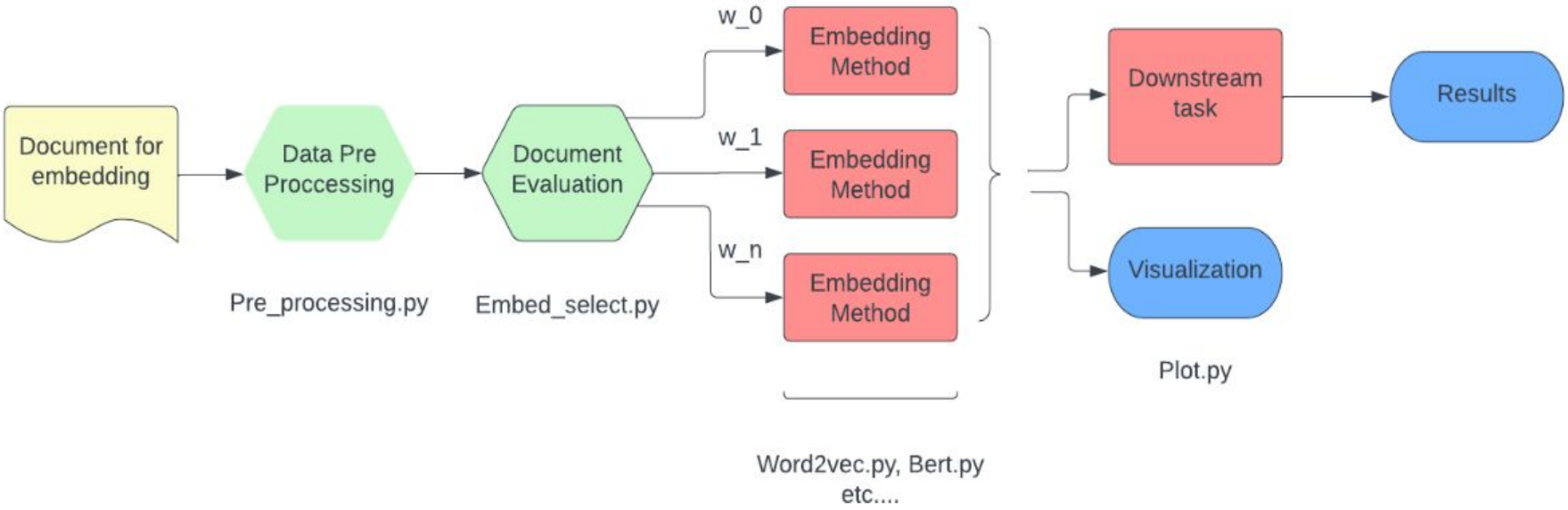
- NLP Maps is a selection algorithm that determines the best NLP algorithm to use for embeddings for downstream tasks
- Factors accuracy, precision, recall, and computational costs to form an ensemble of embedding methods for the given task
- Goal to improve operational efficiency, reduce maintenance costs, and enhance product quality and safety in the chemical engineering plant.

Need for NLP algorithms



- Computers do not inherently understand written text
- NLP algorithms convert written text to vectors that are understandable to downstream tasks
  - I.e Neural Networks, sentiment analysis, name entity recognition
- The accuracy of the task varies with respect to the complexity of the model used

Project Architecture



Summary of Progress

- Four models have been successfully trained on the IMDB movie dataset
- IMDB movie reviews dataset: 25,000 entries for both train and test set, labeled as 0 or 1 for negative or positive reviews, 12,500 for each
- Models show promising accuracy scores from preliminary testing
- Thoroughly tested

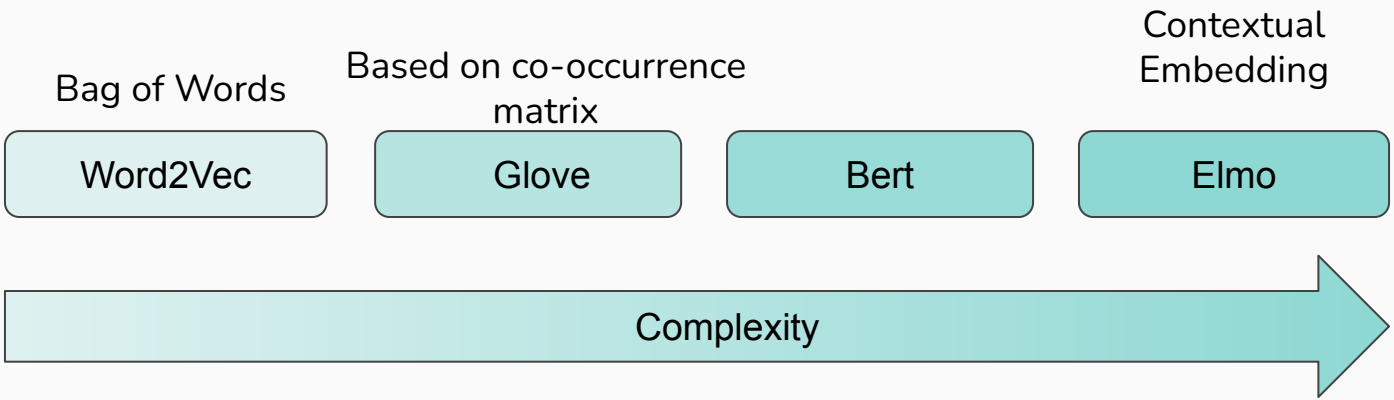
Criteria for evaluating models

- No standardized evaluation method exists, requiring an extensive literature review
- Evaluation techniques were:
  - Highly complex and require more time to fully implement
  - Loosely explained in literature and require further 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	0.587
Bert	0.85
Elmo	NA
GLove	0.754

- Expectedly, the more complex models performed better
- Substantial difficulties with training Elmo model
- Word2Vec underperforms significantly, looking to tune hyperparameters in the future

Plans for the future

- Create a robust selection algorithm
- Incorporate meta embedding
- Add more models
- Include Visualization Packages
- Incorporate models into more downstream tasks

Questions?