



USING NLP & NEURAL NETWORKS TO RECOMMEND BOOKS

BY RONALD
ASSEKO-MESSA
~ DATA SCIENTIST

OBJECTIVES



Background & Preprocessing



Entity Recognition



Modeling & Results



Conclusion & Recommendation





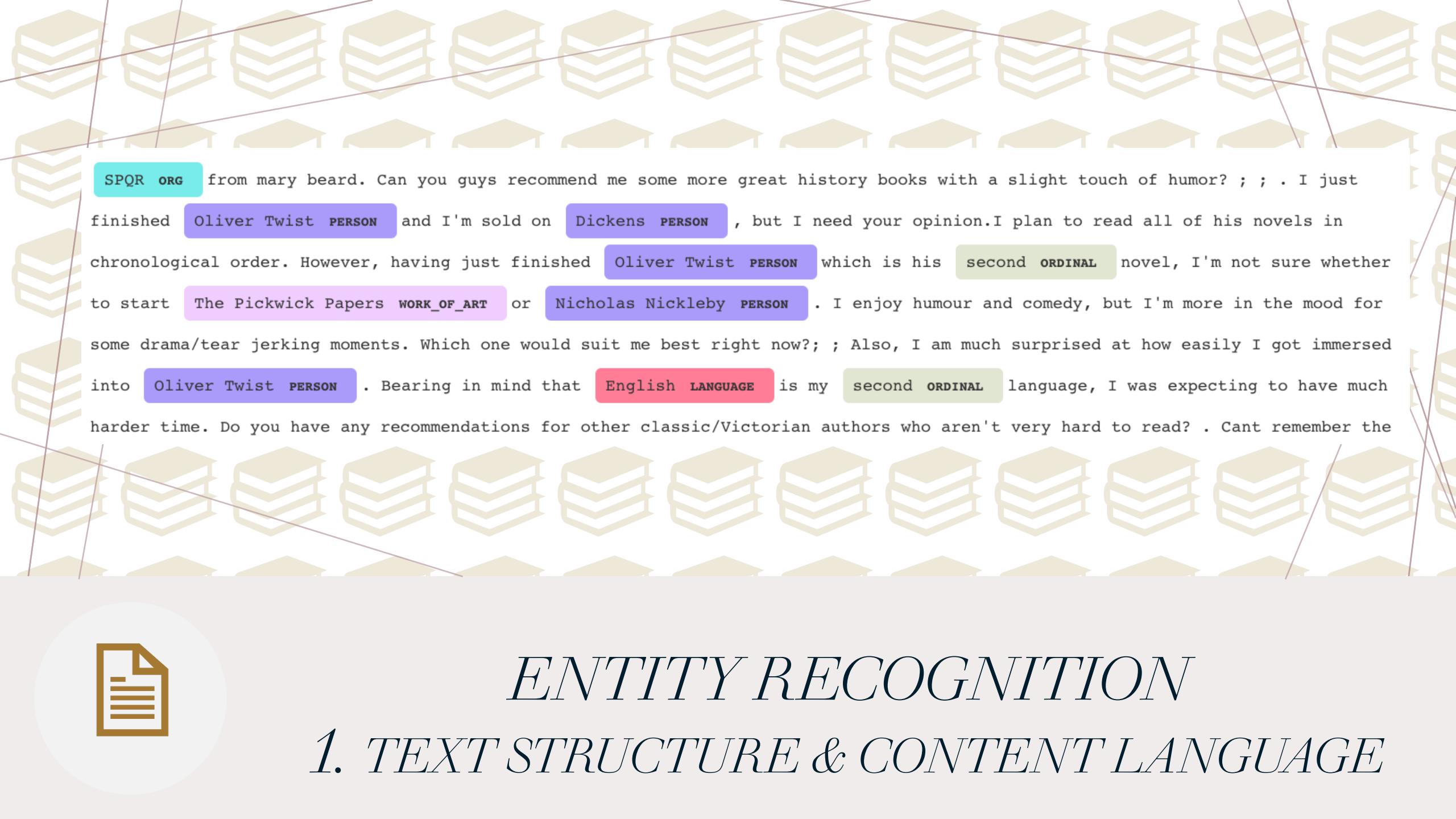
BACKGROUND & PREPROCESSING

Why the need to recommend books?

- Users ask questions about products of their interest (books)
- Receive relevant answers (context)
- Prompt conversations around relevant subjects

How do we go about it?

- Collect data
 - 109,122 posts scrapped from r/booksuggestions (reddit)
 - About 37M Characters
- Preprocessing
 - Raw (No cleaning): 13,159 books
 - Simple cleaning: 14,023 books
 - More advanced cleaning: 10,982 books



SPQR ORG from mary beard. Can you guys recommend me some more great history books with a slight touch of humor? ; ; . I just finished Oliver Twist PERSON and I'm sold on Dickens PERSON , but I need your opinion. I plan to read all of his novels in chronological order. However, having just finished Oliver Twist PERSON which is his second ORDINAL novel, I'm not sure whether to start The Pickwick Papers WORK_OF_ART or Nicholas Nickleby PERSON . I enjoy humour and comedy, but I'm more in the mood for some drama/tear jerking moments. Which one would suit me best right now?; ; Also, I am much surprised at how easily I got immersed into Oliver Twist PERSON . Bearing in mind that English LANGUAGE is my second ORDINAL language, I was expecting to have much harder time. Do you have any recommendations for other classic/Victorian authors who aren't very hard to read? . Cant remember the



ENTITY RECOGNITION

1. TEXT STRUCTURE & CONTENT LANGUAGE



ENTITY RECOGNITION

2. ENTITIES -> BOOKS

1. MODELING & RESULTS



Use NLP to understand the relationship between words and create a vocabulary based on context

Word2Vec (Continuous Skip-gram)
Input words
Output context/similar word



Example of the most similar words

```
[model.wv.most_similar('Bible')]
```

```
[ ('Quran', 0.7612322568893433)
```

2. MODELING & RESULTS



Compute Cosine Similarities to Compare Documents

Measure the angle between documents

- The smaller the angle, the more similar the words/docs, the higher the cosine similarity score



Looking for Horror fiction that explores the uncanny (Lovecraftian?) I love horror films that delve into the outer reaches of human perception, and present monsters or concepts that avoid the usual cliche of the "tormented spirit" or the "psycho killer". I haven't read any Lovecraft, but I believe his brand of horror is the kind of thing I'm looking for. Any recommendations would be much appreciated.



list_of_books	similarity_scores
[the Twilight Zone]	0.989369
[The Ritual]	0.988831
[Teeth]	0.988466

CONCLUSION & RECOMMENDATIONS



The model is indeed able to recommend books based on context



- Give relevant answers
- Prompt conversations around relevant subjects



Test more advanced entity detections methods (Doc2Vec)



Be careful to preprocessing methods



QUESTIONS ■