Movie Sentiment Analysis

DREAM TEAM

Maddie, Jungho, James

AGENDA

01 DATASET

02 MODELS

03 RESULTS

04 GRADIO

05 CONCLUSION

DATASET

- IMDB Dataset
  - 50K Movie Reviews
  - Detailed & Informative

- Rotten Tomatoes Dataset
  - 10K+ Movie Reviews
  - One sentence reviews

- **Sentiment building models**
  - LinearSVC
    - TFIDFVectorizer - Custom stopwords
  - RNN with Tokenizer
  - Vader

LinearSVC with TFIDF - Stopwords
- Accuracy: **92%**

RNN with Tokenizer
- Accuracy: **89%**

Vader
- Accuracy: **70%**

DARTH VADER
(NLTK MODULE)

R2D2
(RNN WITH TOKENIZER)

LUKE SKYWALKER

(LINEARSVC WITH TFIDF - STOPWORDS)

- LINEARSVC with TDIDF Vectorizer: most accurate on IMDB dataset

- Evaluated on RT database with less accuracy

- 2 trained models > VADER

- SPACY is the better NER Classification Model

Runners up...

```
--------------------------------------------------
Review: I first saw this movie as a younger child. My sister had told me about it and I thought it would be more of a kid's movie. However it remains to a
  Entity: first, Label: ORDINAL
  Entity: first, Label: ORDINAL
  Entity: yesterday, Label: DATE
  Entity: 7 years later, Label: DATE
  Entity: first, Label: ORDINAL
  Entity: a few years, Label: DATE
  Entity: yesterday, Label: DATE
  Entity: Especially Bobby, Label: PERSON
  Entity: first, Label: ORDINAL
```

NER Classification: attempted to draw conclusions of what made the prediction using named entities.
- Spacy
- BERT

Skywalker model optimization

R2D2 model optimization

the gorgeously elaborate contir
lord of the rings " trilogy is
1   column of words cannot adequately describe
co-writer/director peter jackson's expanded
vision of j . r . r . tolkien's middle-earth .

PER', 'score': 0.9222884, 'index': 33, 'word': 'jack',
'start': 161, 'end': 165}, {'entity': 'B-PER', 'score':
0.41237748, 'index': 47, 'word': '##lk', 'start': 204,
'end': 206}]