

Predicting New York Times Comments Engagement

...

Carly Tsuda
Capstone Project

Business Question

Proof of concept: determine whether NLP and machine learning can predict which comments will generate further user engagement



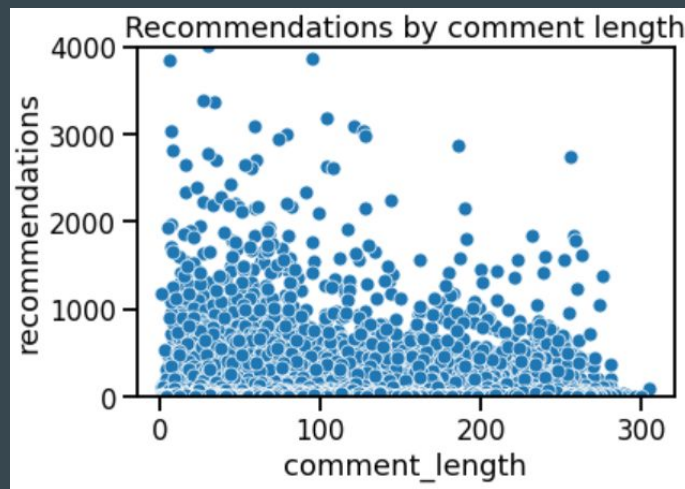
At A Glance

1324 Articles

123,692 Comments

Sections with most engagement:

- News
- Washington
- National
- Foreign
- Opinion
- Business



Detecting Hate Speech

- Used HateSonar to attempt to detect comments with hate speech or offensive language
- Primarily false positives

"I Love It. My Scotch Collie might have finally found her sport!"

"I'm not sure it'll be that easy...."

"At school in sex-ed boys and girls should be taught how to approach each other... Nobody is owed sex."

- Promising potential for a model trained on New York Times comments

Models

- Unable to train models with meaningful amounts of data due to limitations on computational power
- What models did run showed promising potential for predicting whether a comment would receive a high number of recommendations

Recommendations and Future Work

- Offer users Reddit-style 'karma' for positive engagement
- Offer users the option to sort comments based on model's recommendation to prioritize comments that will lead to further engagement
- Expand recommendations to Facebook-style reactions



Thank you!

Questions? Comments?
Contact Carly Tsuda

