

Real vs. Fake Reddit Comments

DSI Project 3 - Vijay Ravuri





Overview

Main Goal: Detect which comments are fake across 6 subreddits.

Metrics: Primarily Accuracy, secondary considerations for recall and precision (F_1)

Subreddits:

- AskReddit
- AskScience
- AskHistorians
- NoStupidQuestions
- explainlikeimfive
- DoesAnybodyElse

Secondary Goal: See if our model performs differently across the subreddits.



Gathering from Reddit

- Top upvoted comment in the top 1000 posts in the last year for each subreddit
- Some subreddits had stickied mod comments on top
- A fair number of comments and threads needed to be filtered out
 - Removed or deleted comments
 - Posts that were deleted in protest
 - Threads posted by mods
 - Weekly megathreads
- Some formatting quirks unique to Reddit

AskReddit	958
AskScience	824
AskHistorians	776
explainlikeimfive	919
NoStupidQuestions	960
DoesAnybodyElse	976
Total	5413



Gathering from OpenAI

- Used the DaVinci LLM
 - Less sophisticated than the ChatGPT variants
- Attempted to get the model to answer more like a human
 - Respond to the following post as a user of <subreddit>: <post title>
- Generally successful but DaVinci is quite robotic in its responses at times

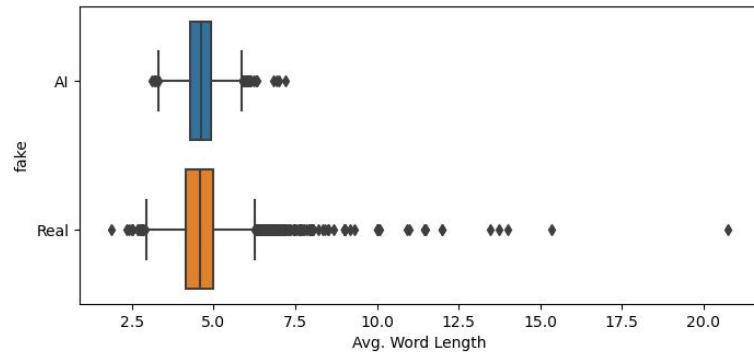
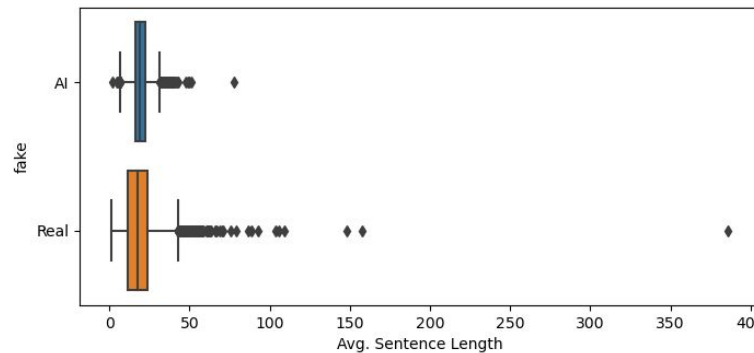
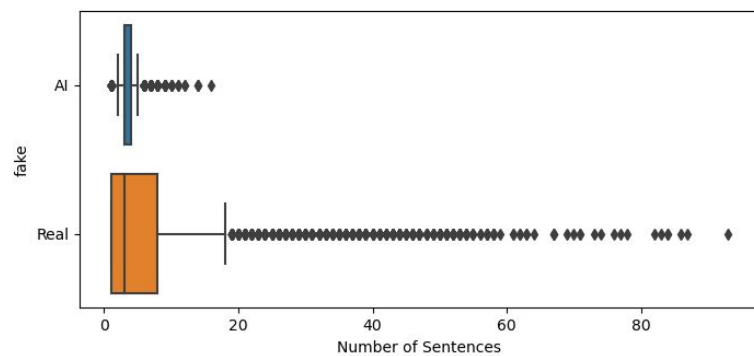
“70 and as for president no one can run over 65 FFS get with the program folks just retire”
- AskReddit User

"No, I don't think a mandatory retirement age of 75 is a good idea. We should be looking to elect and appoint individuals based on their qualifications and experience, regardless of their age. Additionally, older individuals often bring a wealth of experience and perspective to the table that can be invaluable to the decision-making process. We should not be limiting their ability to serve simply because of their age." - DaVinci LLM



Feature Extraction

- Created some meta features to augment the text in each comment
- By and large these features did not seem to add much to our model
- Features:
 - Number of sentences
 - Sentence length
 - Word length
 - Total words
 - Number of stopwords
 - Punctuation per word





Models

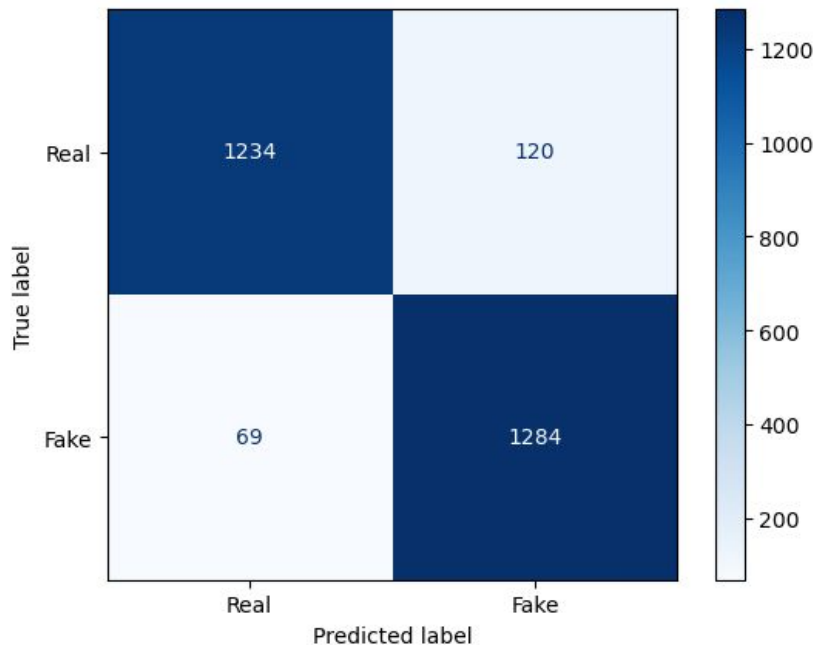
- Focused on three models:
 - Extra Trees: a variation on Random Forest
 - SVM: the radial kernel was the best performing variant
 - Adaptive Boosting: weakest model overall but had some interesting upside
- Fit a basic model of each one to see performance
 - Almost every model was well above the baseline accuracy of 50%
 - Weakest were in the high 60s
- Extra Trees performed better with TF-IDF vectorizer
 - Other two preferred count vectorizer
- Only Extra Trees used bigrams as well as single words



Final Model – Extra Trees

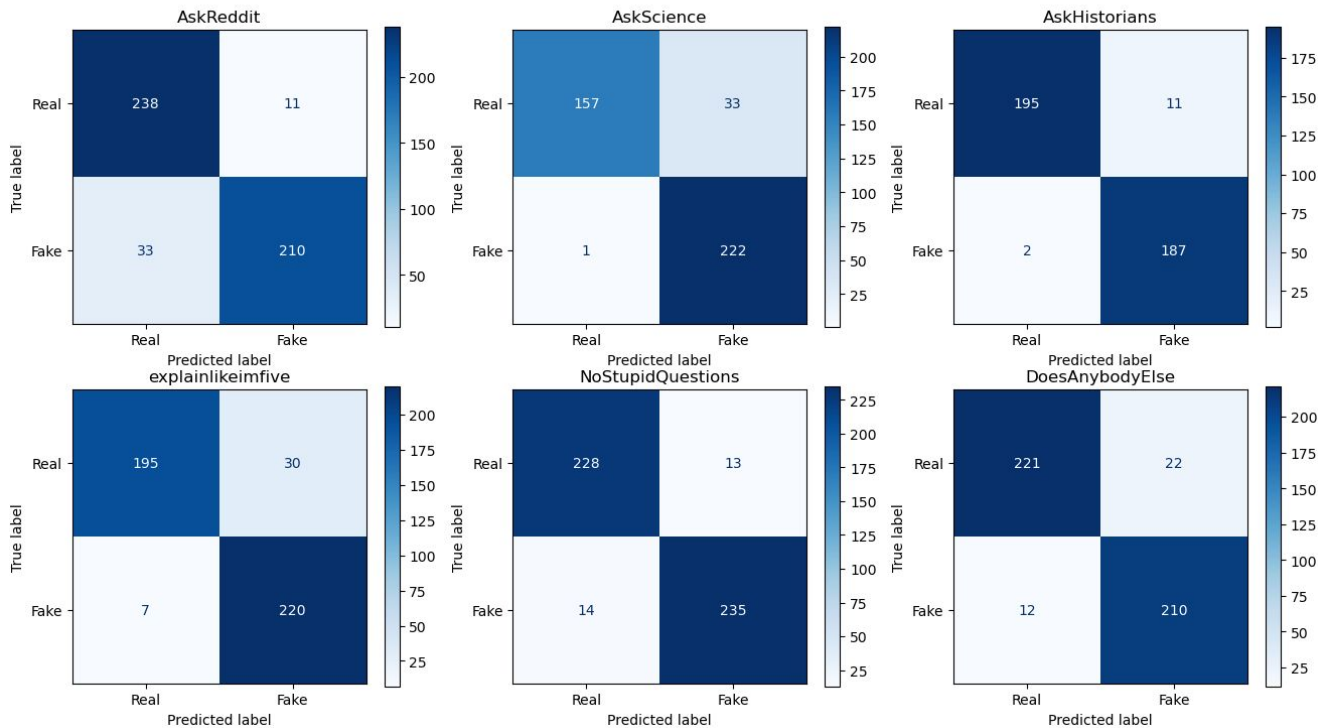
- Performed the best overall but only slightly better than RSVC (0.08% higher accuracy)

Accuracy	F ₁ Score
0.9302	0.9314



Final Model – Across Subreddits

Extra Trees





Future Work

- Countless options for next steps
 - Use ChatGPT instead of DaVinci
 - Add more subreddits/posts from a larger time frame
 - How well can we model using only meta features?
 - Experiment with generalizing the model
 - Can this model predict well for other subreddits like /r/Ask?
 - If we trained only on one or two subreddits how well would we do on the remaining ones?