

Using NLP to train a
classifier on which
subreddit a given post came
from: Binary Classification

Can we get input from other
similar subreddits?





Data wrangling/gathering/acquisition



Push Shift to Csv

- AITA
- Started with: 9999
- After Cleaning: 6081
- RA
- Started with : 9999
- After Cleaning: 6807

Empty rows

- unddit.com
- Dropped
- 1 post per user/author

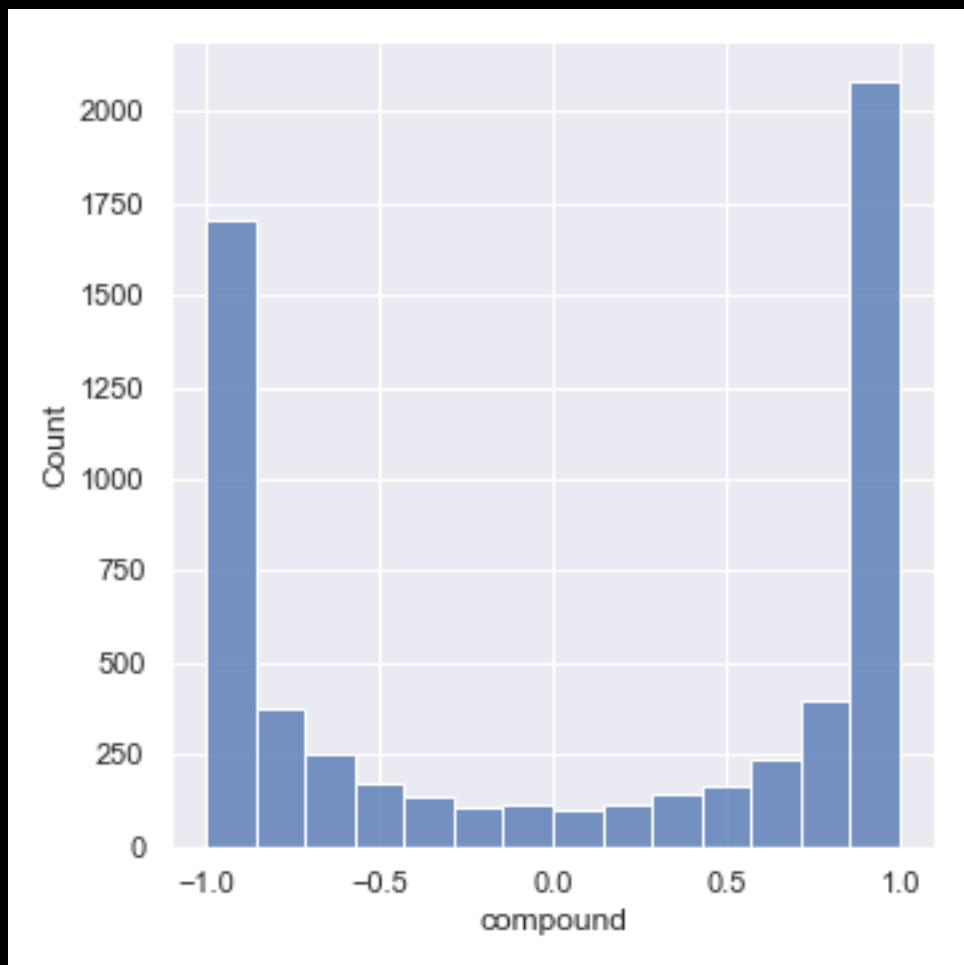
Changing UTC

- AITA
- Started with: 9996
- After Cleaning: 6408
- RA
- Started with : 9999
- After Cleaning: 6834
- 13,242 posts
- 11/19/2022 -11/24/2022

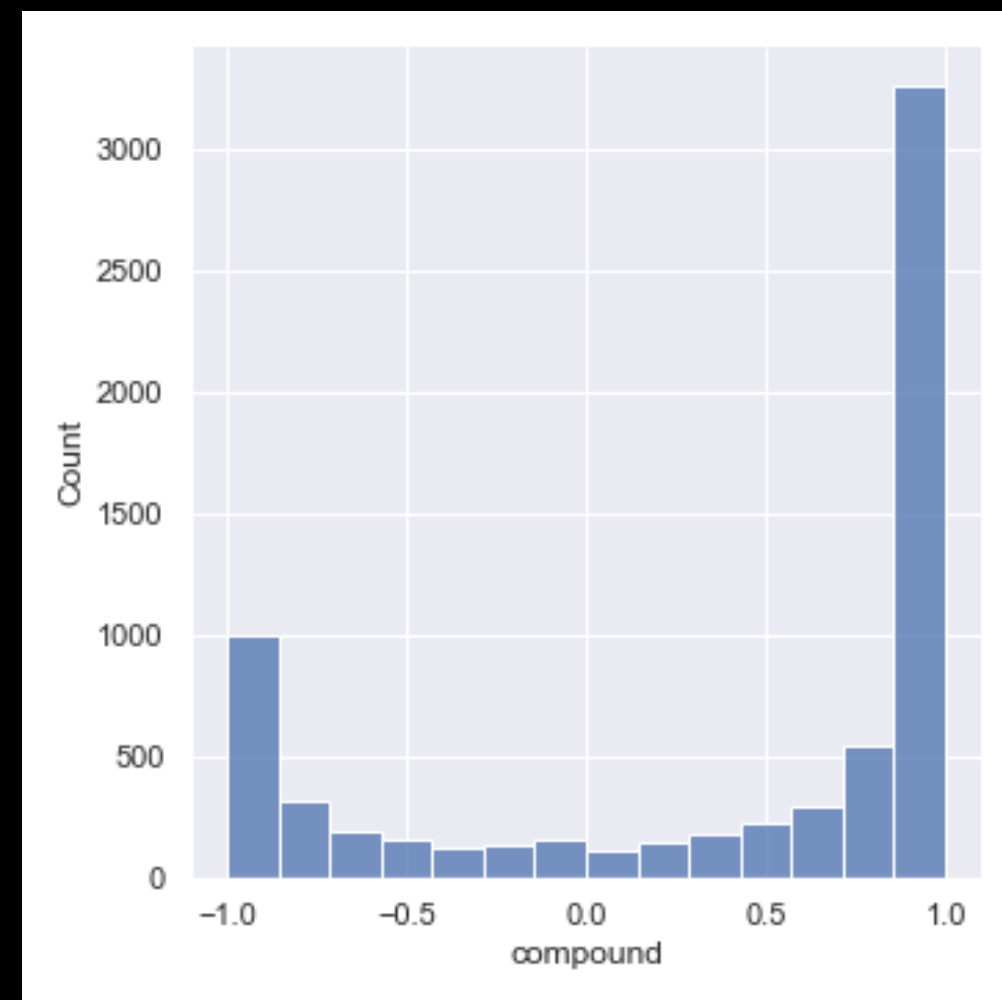


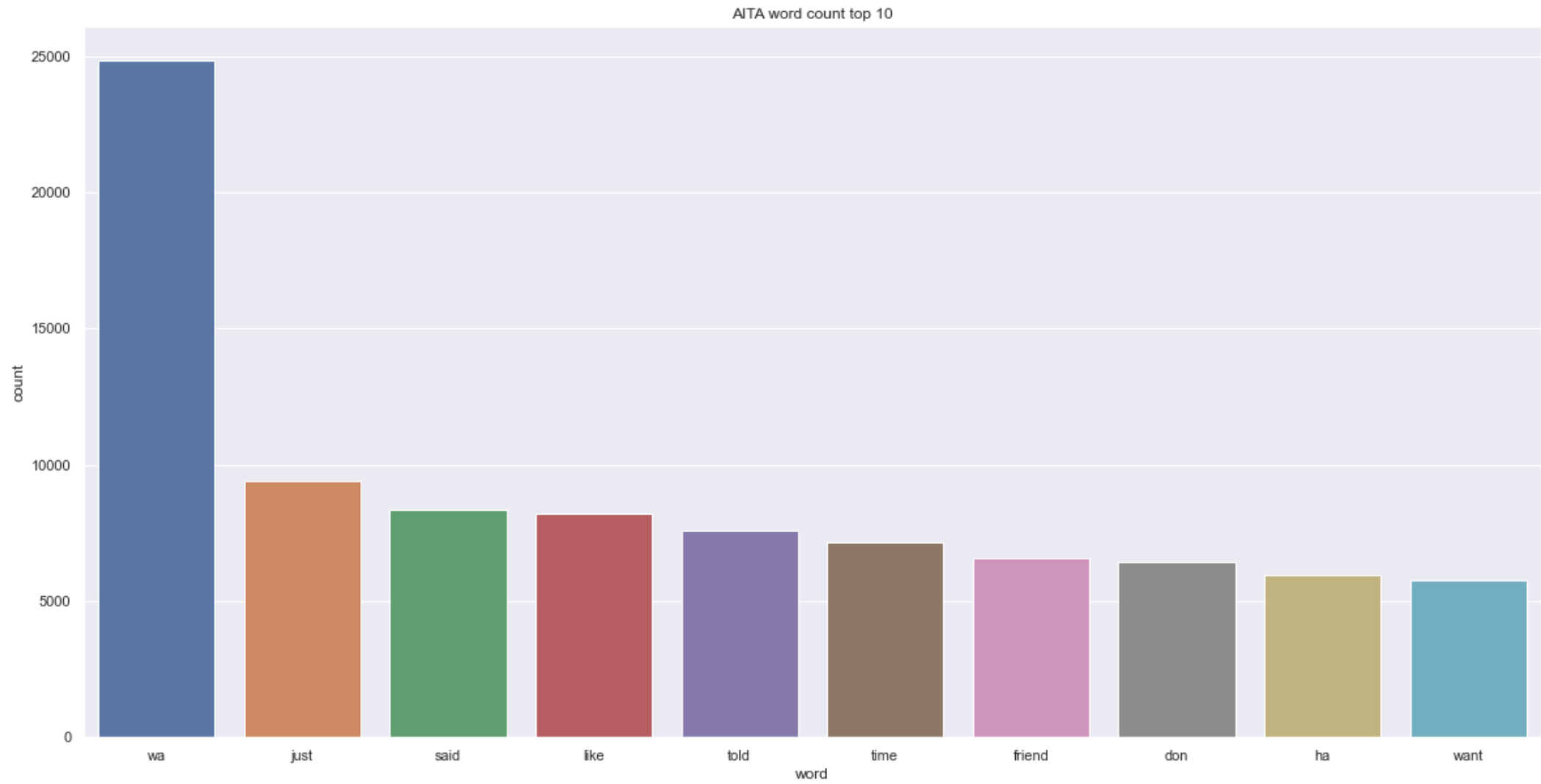
Compound Distribution

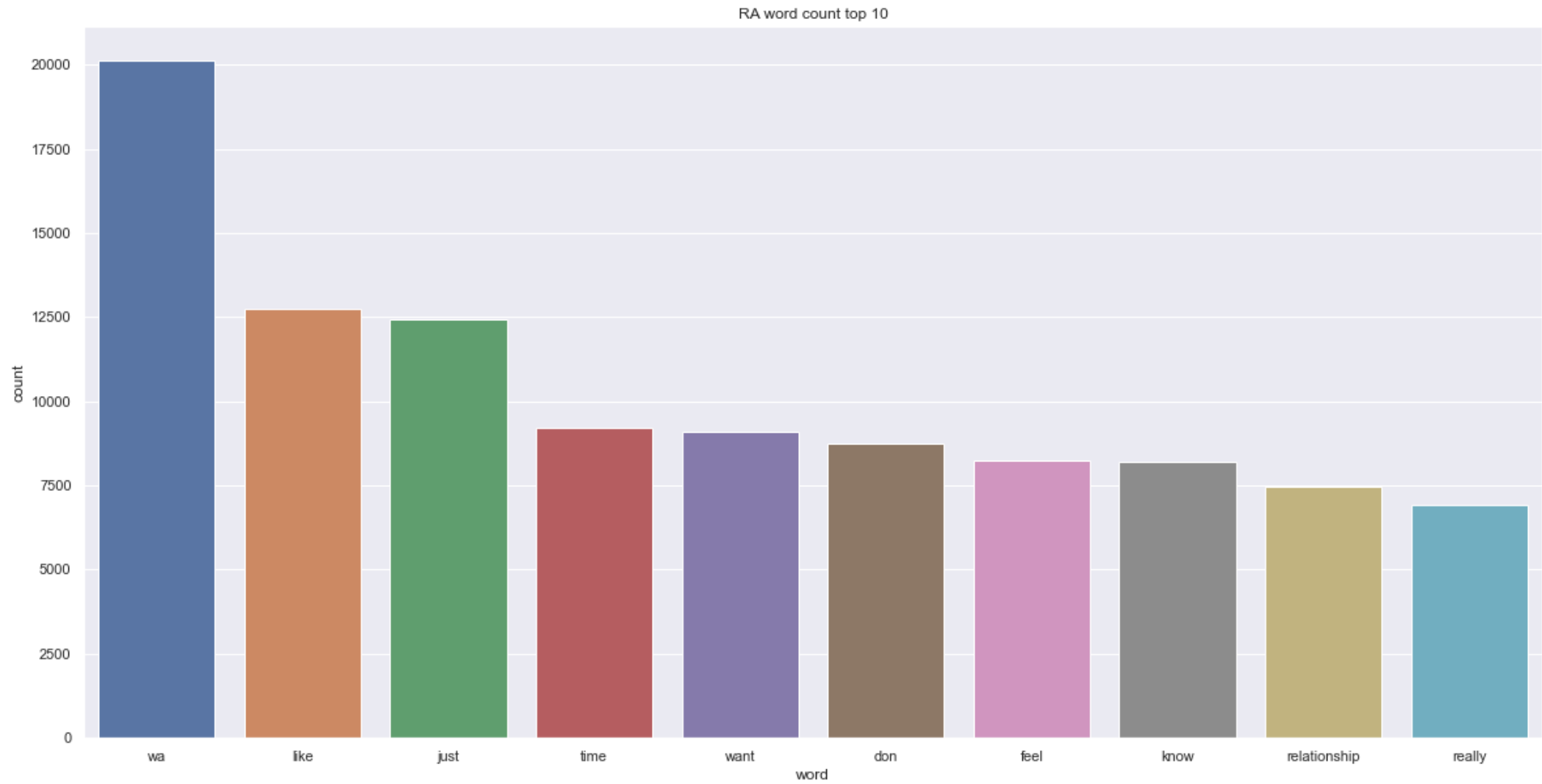
AITA

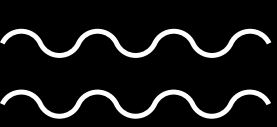


RA

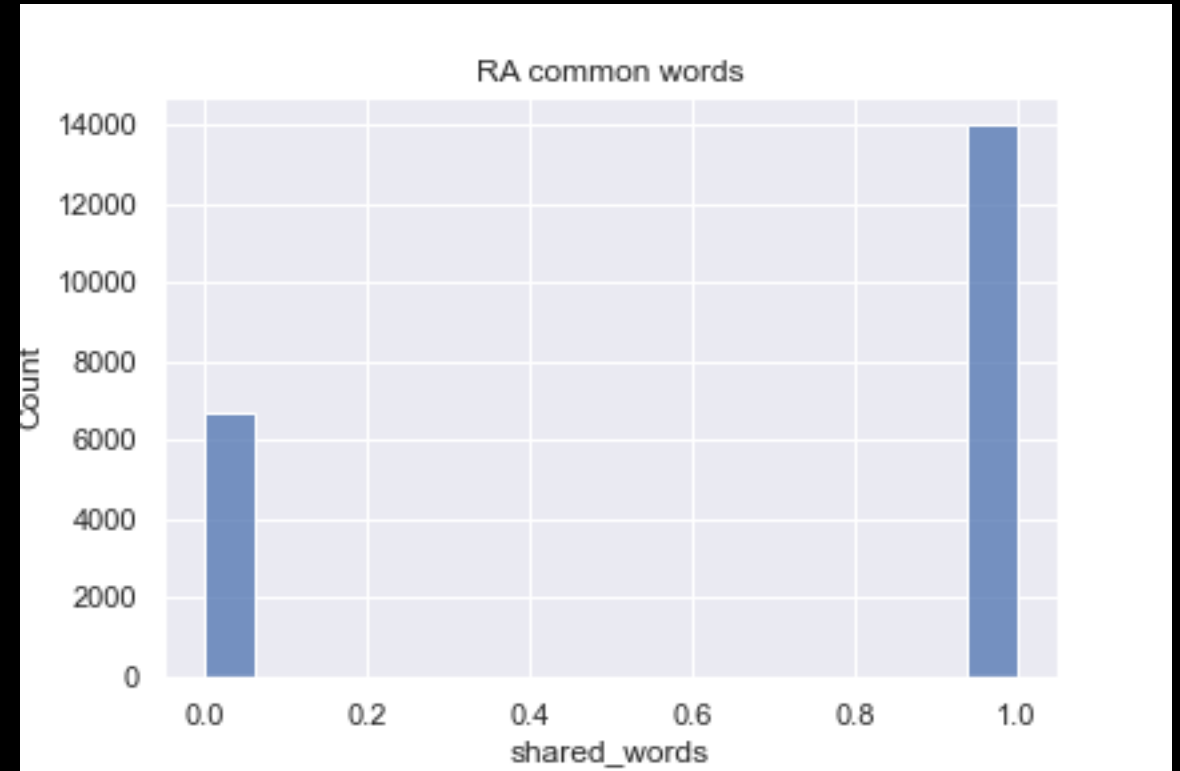
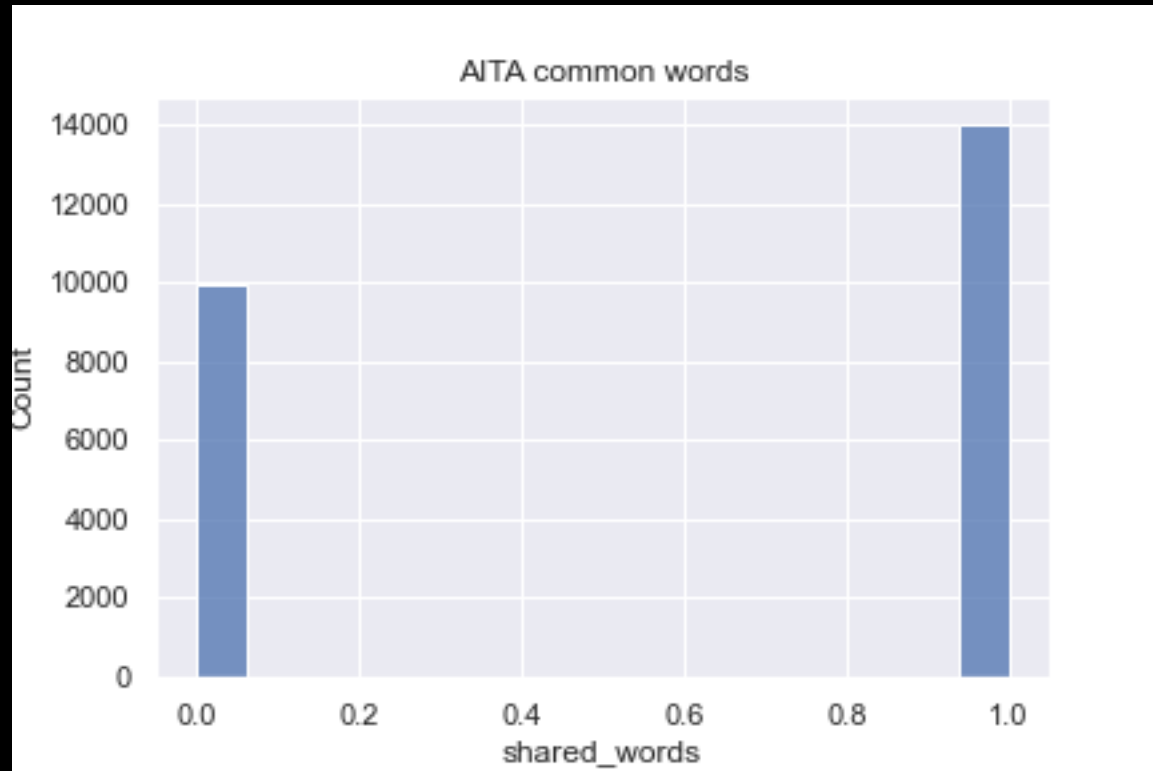








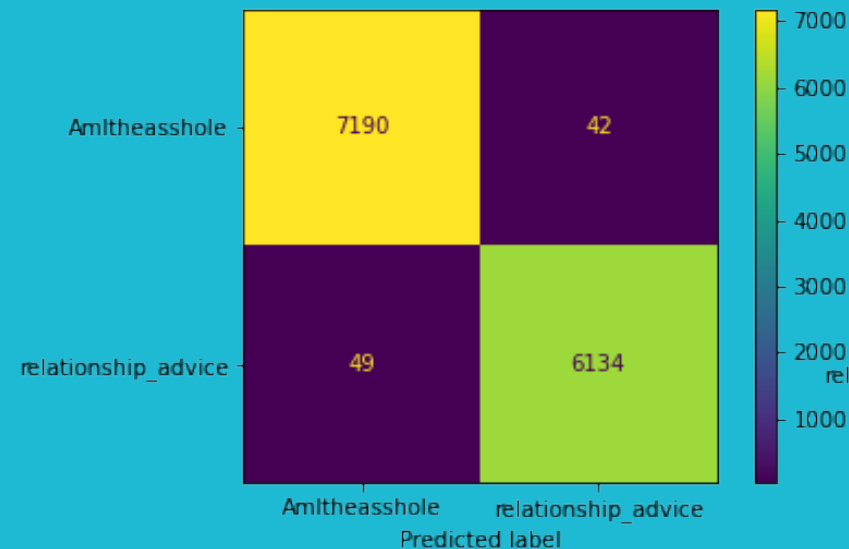
Shared Words



Model Performance:

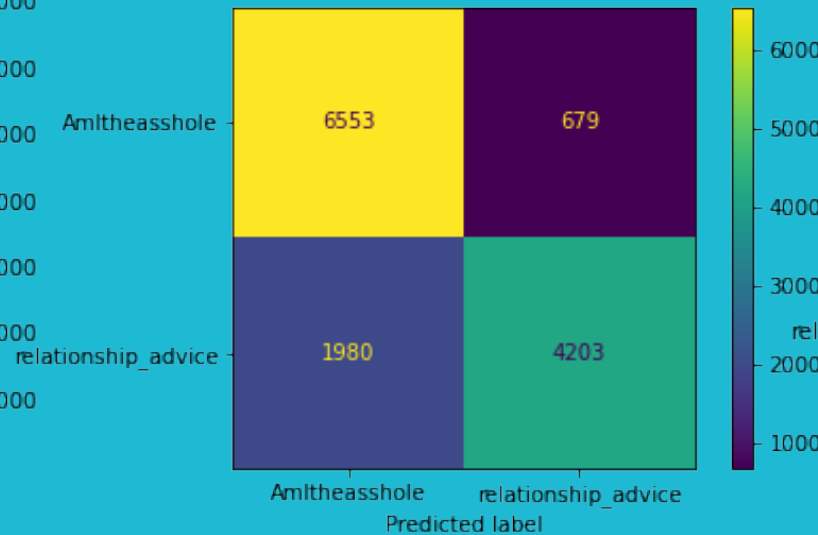


KNN



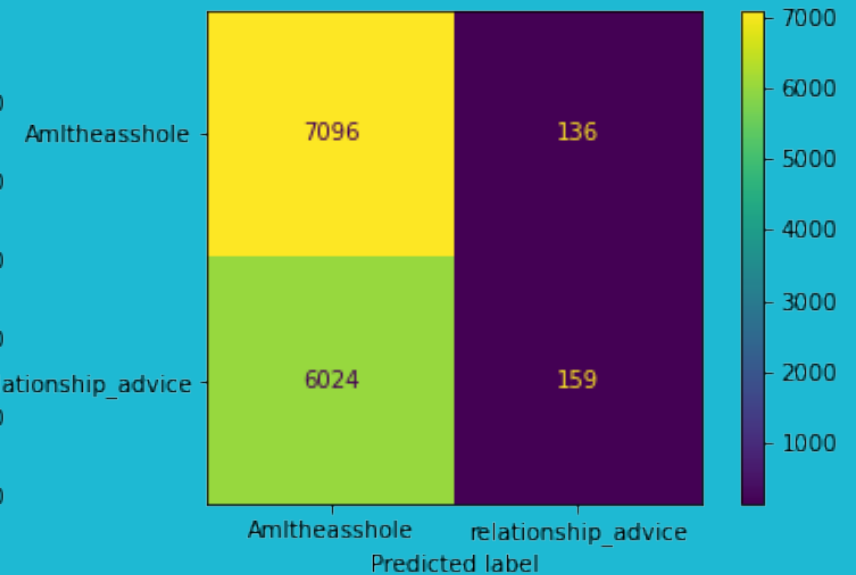
Accuracy: 0.99
Precision: 0.99
Sensitivity recall: 0.99
Specificity: 0.99
F1_score: 0.99

Logistic Regression



Accuracy: 0.80
Precision: 0.86
Sensitivity recall: 0.68
Specificity: 0.68
F1_score: 0.76

Naïve Bayes



Accuracy: 0.54
Precision: 0.53
Sensitivity recall: 0.02
Specificity: 0.02
F1_score: 0.04





AITA or RA?

Very recently, I found out that my husband had installed a tracking device in his sister's car (he's a mechanic) while he was fixing it for her. I confronted him and he refused to tell me what was going on or expand on the situation. His sister and I are close. I threatened him to tell her and her husband but he lashed out saying that this has nothing to do with me and that I was out of line to get myself involved. He told me "it's complicated" that he can't say anything to me now and that I should keep my nose out of it. We had an argument and he started avoiding me saying there's gonna be an issue if I go through with my threat.



AITA for threatening to tell her? Should I back off?

**THANK
YOU**

