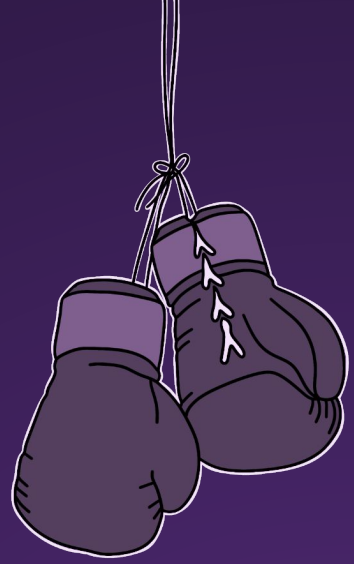
SUBREDDIT MMA vs UFC CLASSIFIER

By Jeff Alexander

FURY

MMA LEAGUE

Is UFC Synonymous with MMA?

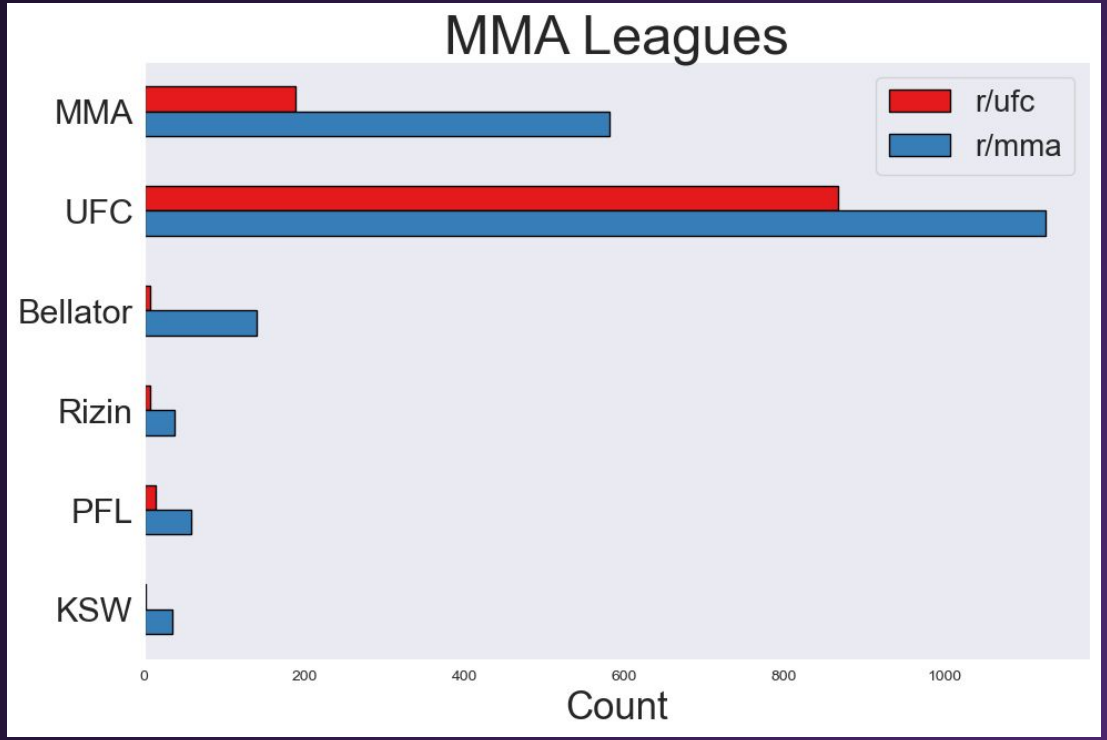


- Can a model accurately classify the subreddits?
- UFC monopoly over MMA fandom?

Over 22%

Of r/MMA Posts include the word ‘UFC’

FIGHTING CHANCE



TAKE OFF
THE
GLOVES.

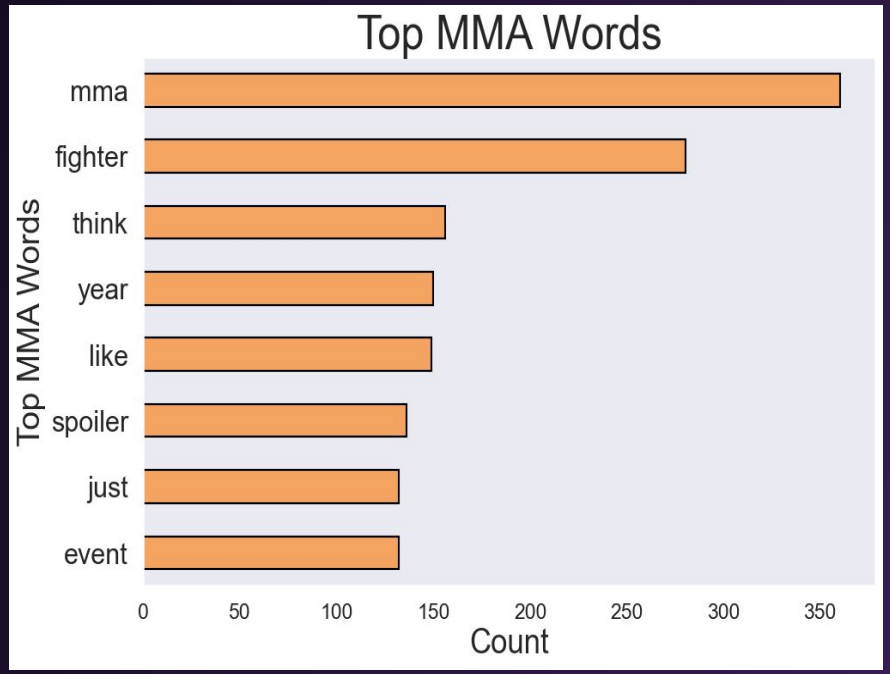
BEST MODEL

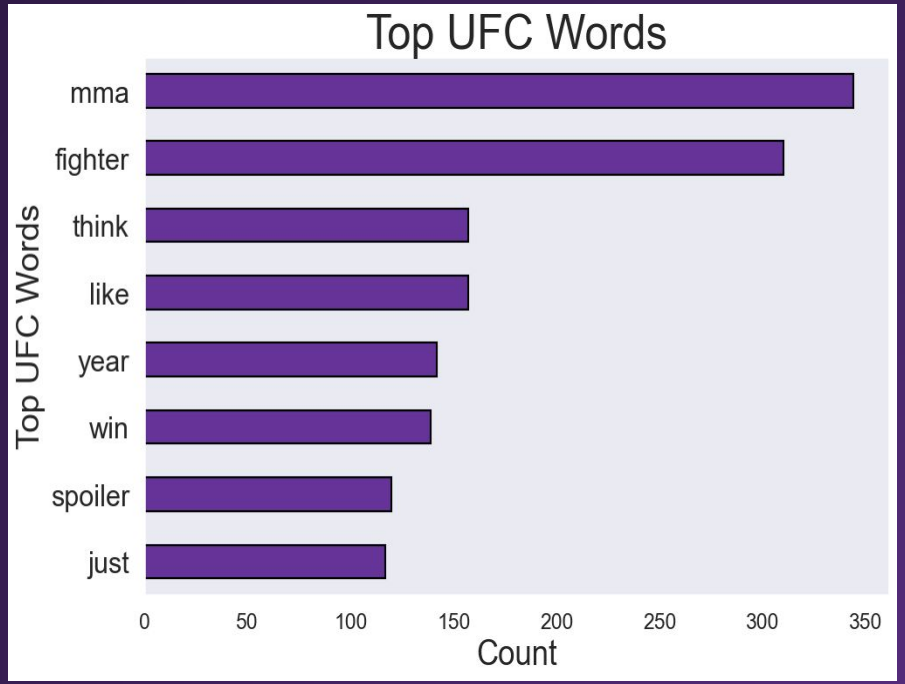
- CountVectorizer as Transformer with Custom Stop Words (‘ufc’, ‘Dana White’, famous fighters)
- Stacked Model with Multinomial NB, Random Forest, and Logistical Regression as Base Estimators
- Logistic Regression as Final Estimator

BEST MODEL

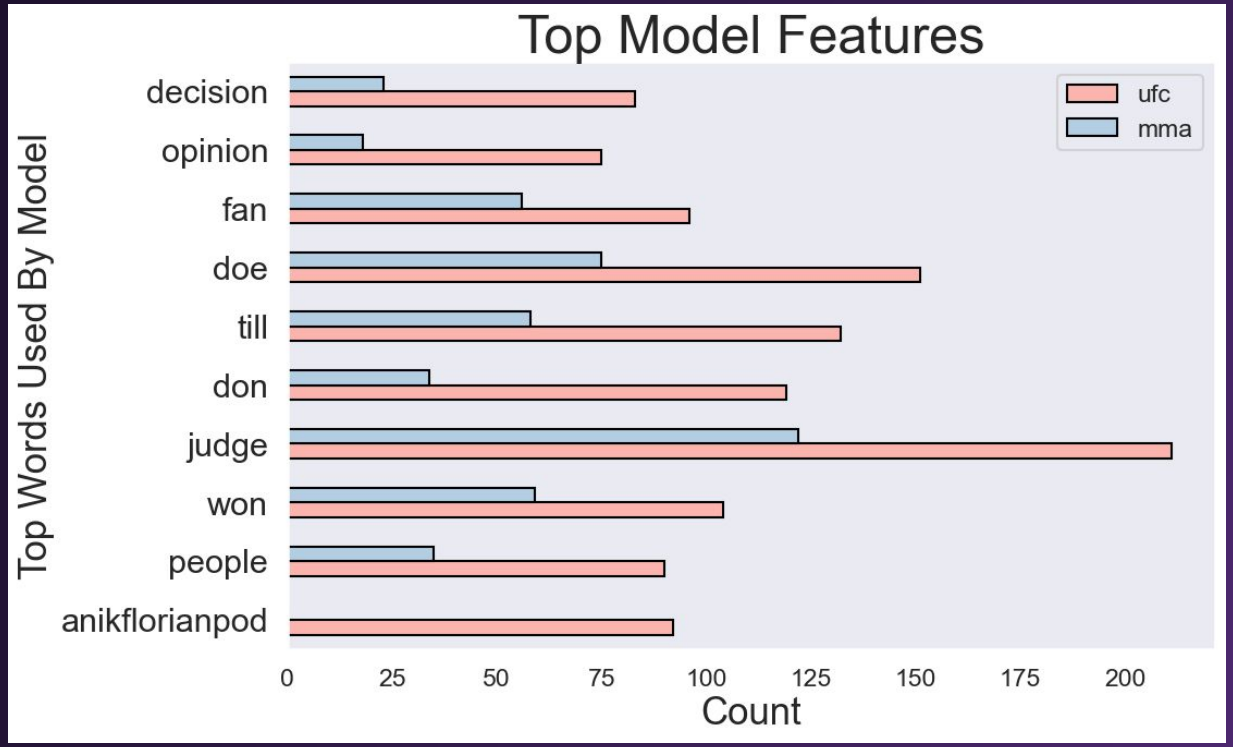
- Final Top Accuracy of 73.19%
- Misclassification of 26.81%

TOP WORDS


Top MMA Words
mma
fighter
think
year
like
spoiler
just
event
Top MMA Words
0
50
100
150
200
250
300
350
Count


Top UFC Words
mma
fighter
think
like
year
win
spoiler
just
Top UFC Words
0
50
100
150
200
250
300
350
Count

TOP FEATURES USED BY THE MODEL



Emerging Interpretation

r/MMA

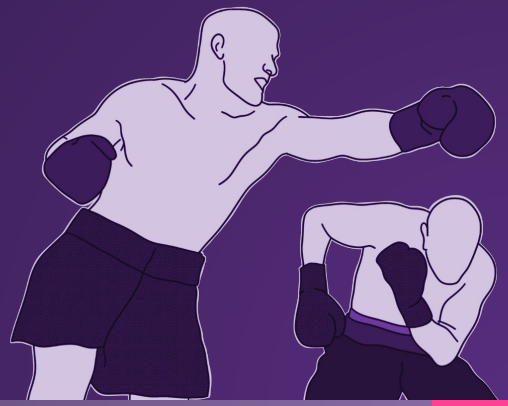
What happened (broader events, people)

r/UFC

How they felt about what happened

MISCLASSIFICATION ANALYSIS

Model Misses



MISCLASSIFIED

False Negative

(r/ufc classified as r/mma)

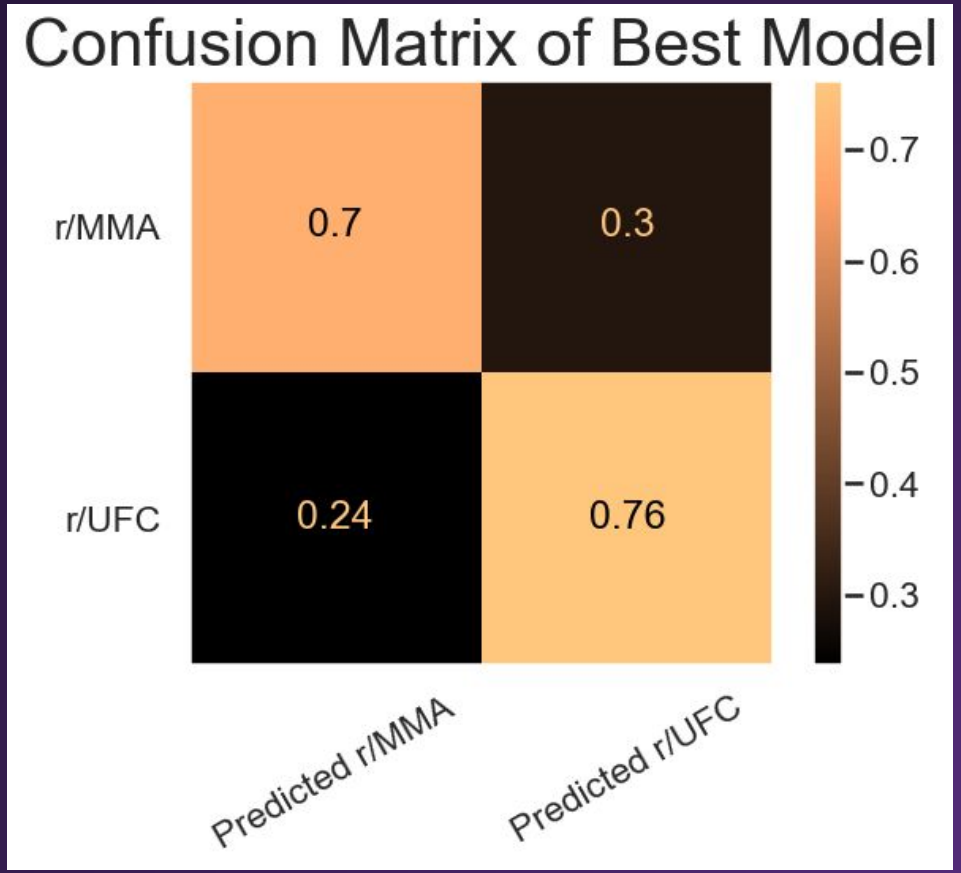
'Confirmed: Khabib Nurmagomedov ha ‘left the MMA industry”

False Positive

(r/mma classified as r/ufc)

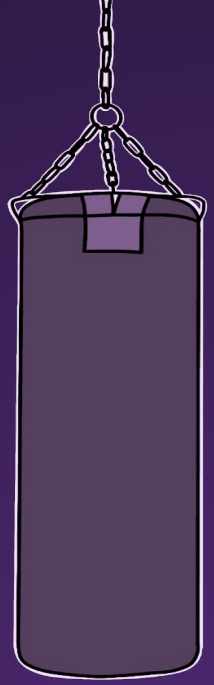
'Every Mythical Fighter Of The UFC | Explained'

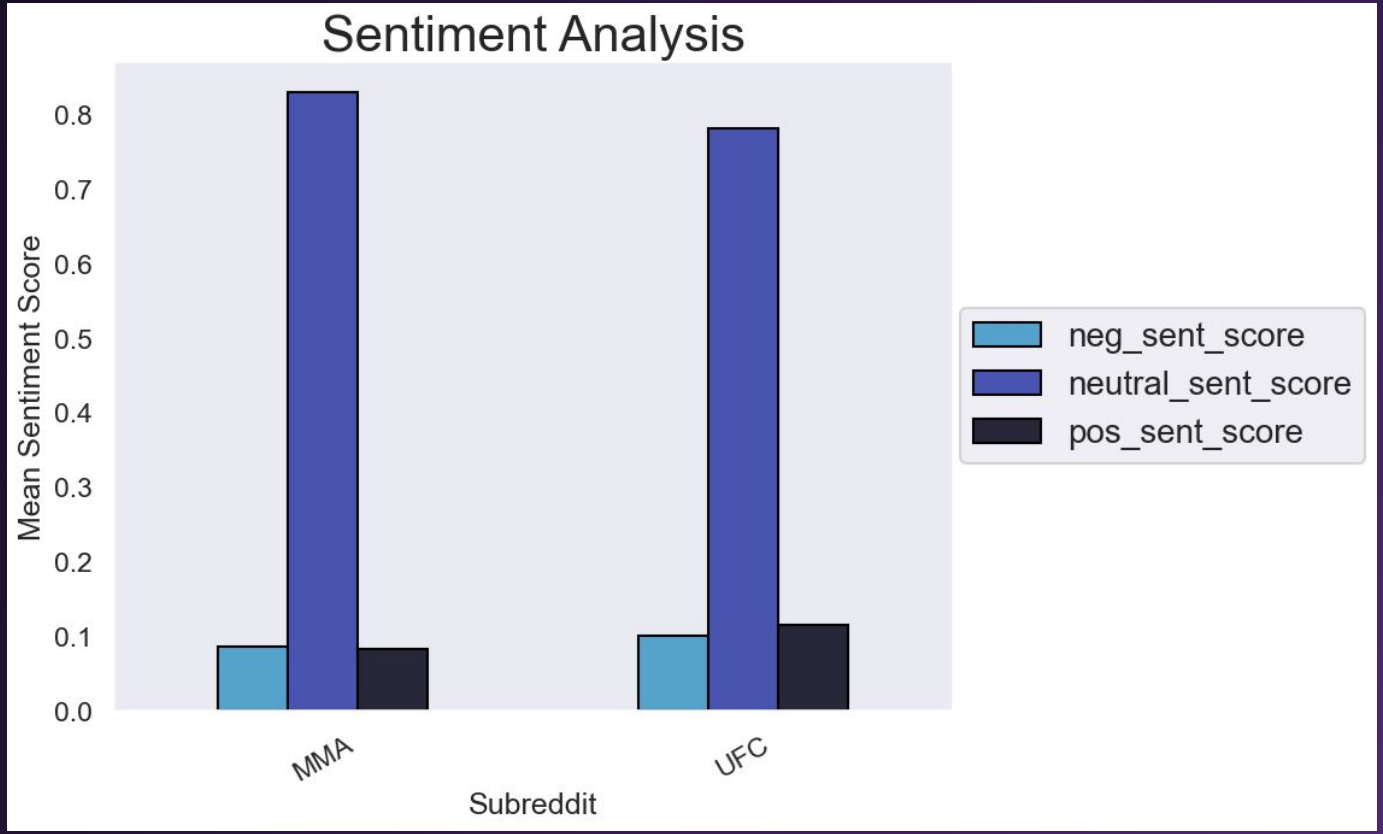
Confusion Matrix

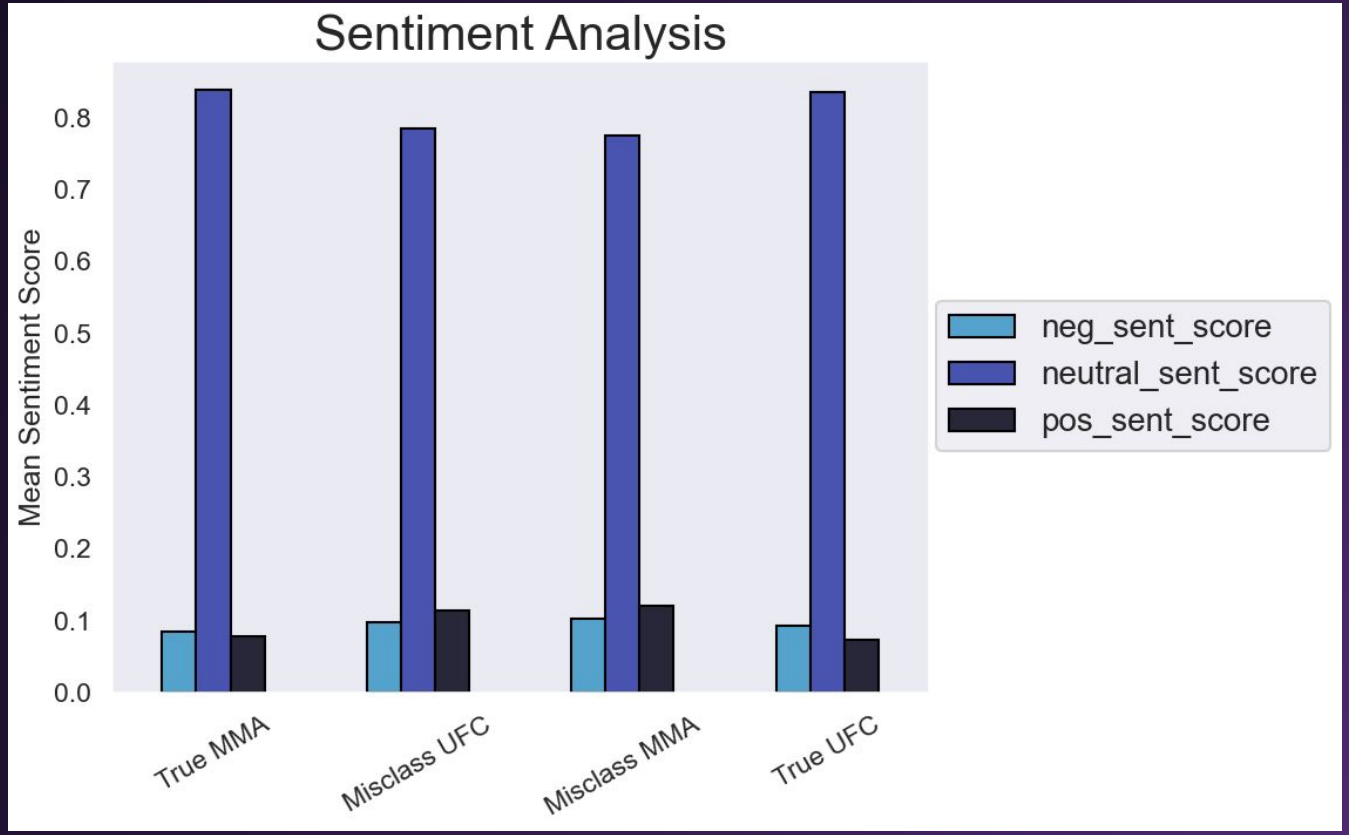


Sentiment Analysis

Misclassification Analysis

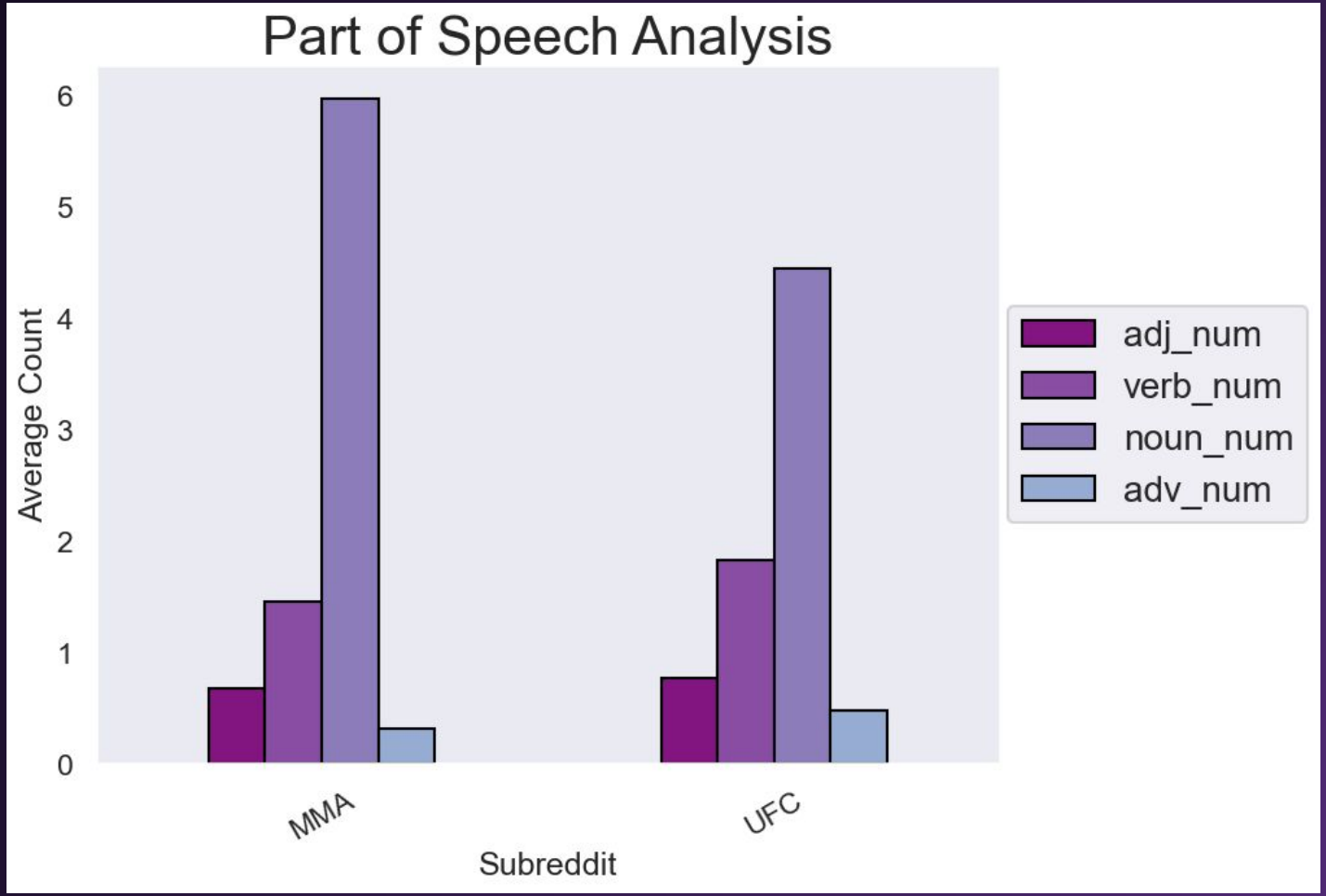


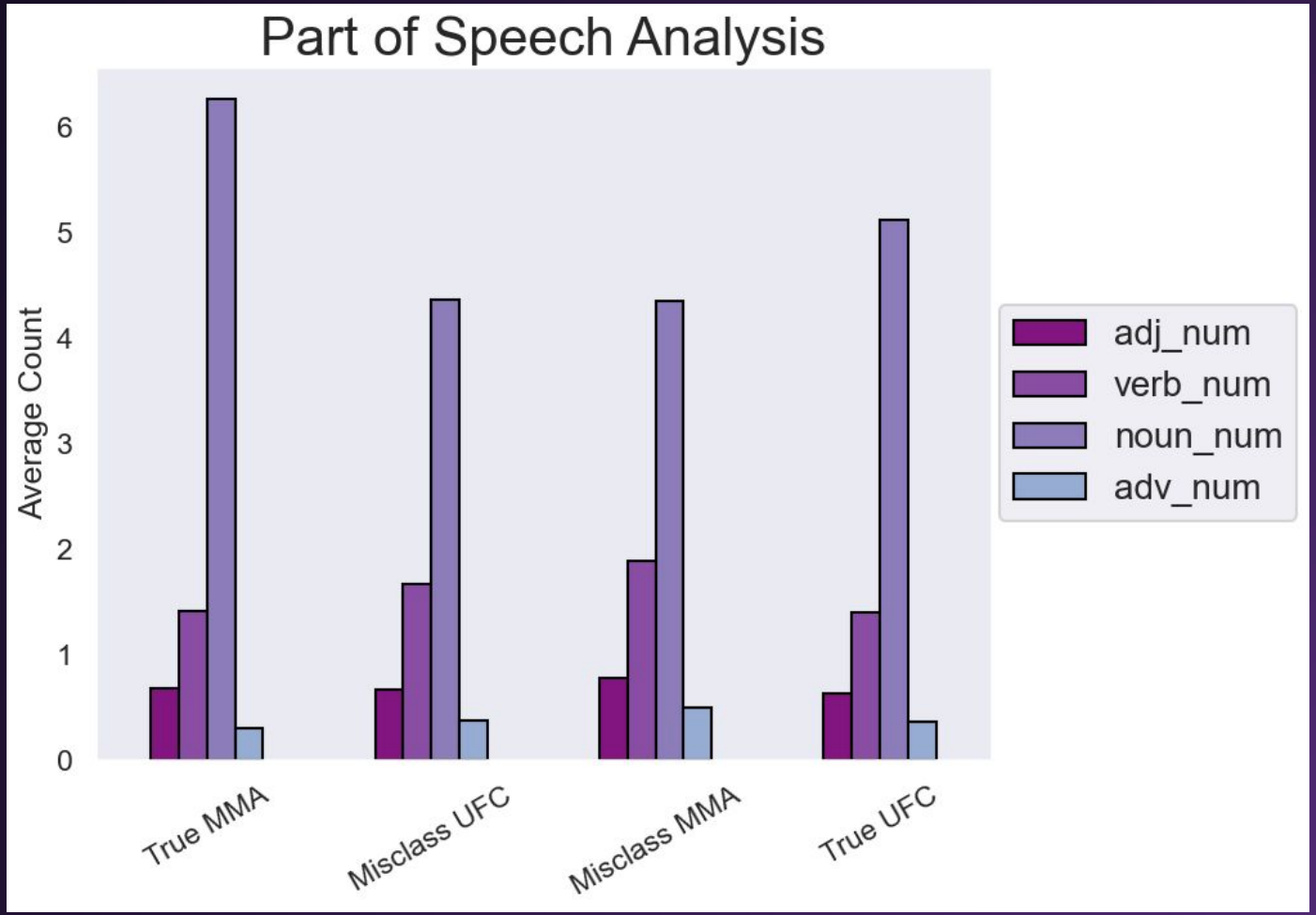

Sentiment Analysis
Mean Sentiment Score
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
MMA
UFC
Subreddit
neg_sent_score
neutral_sent_score
pos_sent_score


Sentiment Analysis
Mean Sentiment Score
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
True MMA
Misclass UFC
Misclass MMA
True UFC
neg_sent_score
neutral_sent_score
pos_sent_score

Parts of Speech

Misclassification


Part of Speech Analysis
Average Count
0
1
2
3
4
5
6
MMA
UFC
Subreddit
adj_num
verb_num
noun_num
adv_num


Part of Speech Analysis
Average Count
0
1
2
3
4
5
6
True MMA
Misclass UFC
Misclass MMA
True UFC
adj_num
verb_num
noun_num
adv_num

FURY

MMA LEAGUE

CONCLUSIONS

- Due to overlapping fandoms and terminologies, the model could not accurately classify to over 75%
- The UFC has a near-monopoly, for now...

FUTURE RESEARCH

- Comment Section
- Upvoting/Downvoting
- More Complex Parts-of-Speech and Sentiment Analysis

THANK YOU