

Background

Politicians frequently use social media share updates, promote their policies, engage with constituents, and shape public opinion.

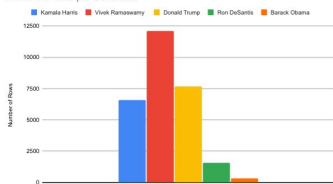
Research question: How does the public engage with this media? Are there key perspectives and issues that resonate with the public?

Introduction

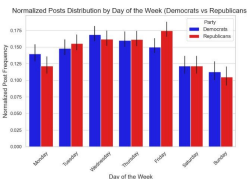
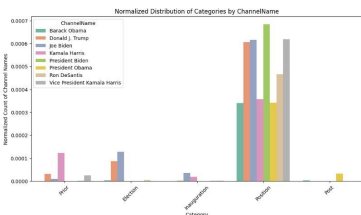
Data: **Junkipedia** scraped posts from 20+ million social media channels and 15+ platforms'

Selected tweets from 5 prominent political figures: Joe Biden, Barack Obama, Donald Trump, Kamala Harris, Ron DeSantis

Amount of Data per Candidate



Total # of data:
23132



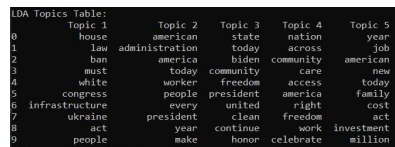
Text Analysis

NLTK Sentiment Intensity Analysis [-1, 1]



Latent Dirichlet Allocation (LDA)

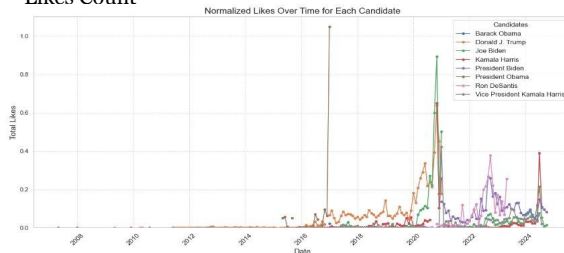
Bayesian topic model to determine topic of each tweet



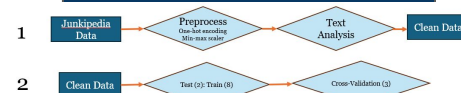
Emphatic language detection: regular expression parse for '!', capital letters

Top 20 words, presence of key words

Target Variable: engagement levels (likes, comments, shares) normalized by follower count, ex. Likes Count



Model Training



Random Forest Regression

{ n_estimators: [10, 50, 100, 200],
max_depth: [None, 10, 20, 30],
min_samples_leaf: [1, 2, 4] }

Ridge linear model (multi-predictors)

{ alpha: [0.1, 1, 10, 100],
solver: ['auto', 'svd', 'cholesky', 'lsqr', 'saga'] }

MLP Regression (Neural Network)

{ hidden_layer_sizes: [(50,), (100,), (100, 50)],
activation: ['relu', 'tanh', 'logistic'],
alpha: [0.0001, 0.1, 0.001, 0.01],
max_iter: [200, 500, 1000] }

Results

Model	MSE	MAE	R2
Random Forest	0.0003	0.0052	0.4539
Ridge	0.0004	0.0070	0.1786
Neural Network	0.0005	0.0068	0.1002

Random Forest Feature Importance:

1. Post length: 0.2030
2. Days into Category: 0.1204
3. ChannelID_10622231 (Obama): 0.1191
4. Follower count: 0.0963
5. Sentiment: 0.0786