So, Reddit, AITA?

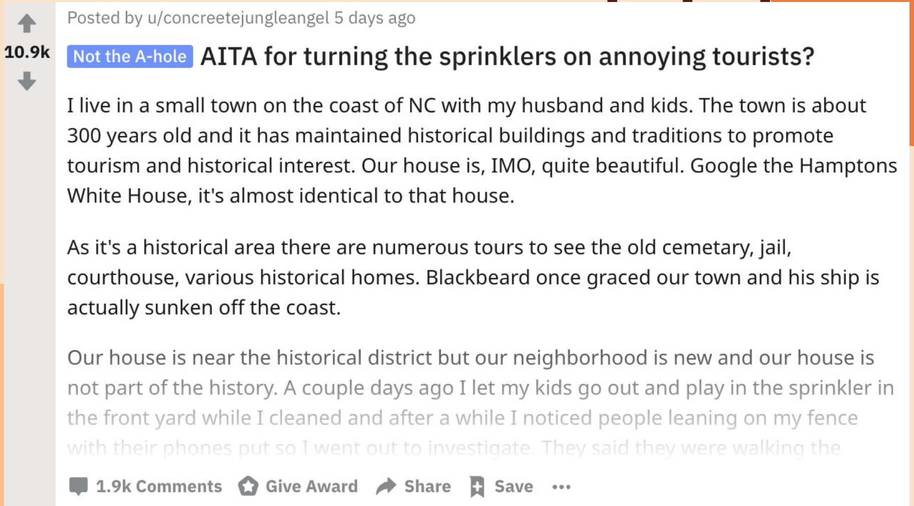
Please excuse the profanity in this presentation

Background

- Social media
- Post, comment, upvote
- Sub-communities on reddit = Subreddits
    - AITA

The Challenge: Can we predict whether reddit users will classify an AITA post as NTA or TA?

The Data



Posted by u/concreetejungleangel 5 days ago

Not the A-hole AITA for turning the sprinklers on annoying tourists?

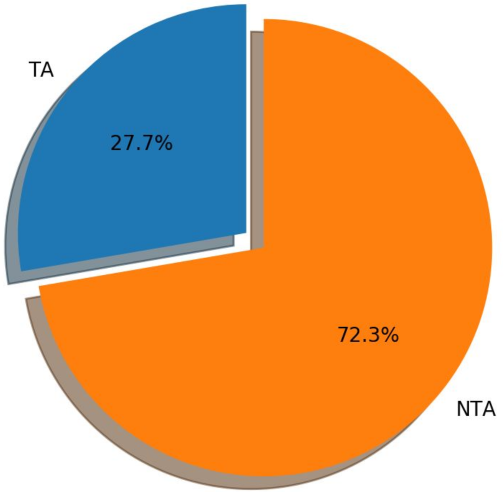
I live in a small town on the coast of NC with my husband and kids. The town is about 300 years old and it has maintained historical buildings and traditions to promote tourism and historical interest. Our house is, IMO, quite beautiful. Google the Hamptons White House, it's almost identical to that house.

As it's a historical area there are numerous tours to see the old cemetary, jail, courthouse, various historical homes. Blackbeard once graced our town and his ship is actually sunken off the coast.

Our house is near the historical district but our neighborhood is new and our house is not part of the history. A couple days ago I let my kids go out and play in the sprinkler in the front yard while I cleaned and after a while I noticed people leaning on my fence with their phones put so I went out to investigate. They said they were walking the

- Pushshift.io API
    - Title
    - Description
    - Flair
        - moderator-tagged short description, unique to the subreddit
        - AITA: NTA, TA, ESH, NAH

Class Distribution



Data Engineering

By Hand

- Description length
- Number of ‘!’, ‘?’, “, ‘...’
- Number of words in all caps
- Number of curse words
- Number of first versus third person pronouns

Using Libraries

- Sentiment analysis of titles & descriptions
- TF-IDF → LDA
  - Latent Dirichlet Allocation AKA unsupervised categorization

NLTK

TextBlob

scikit learn

GENSIM topic modelling for humans

Modeling

First, we SMOTE-d to correct for class imbalance! Then, we tried vanilla versions of…

Logistic Regression

TP: 60%
TN: 38%
FP: 40%

KNN

TP: 60%
TN: 49%
FP: 40%

Random Forest

TP: 73%
TN: 26%
FP: 27%

SVM

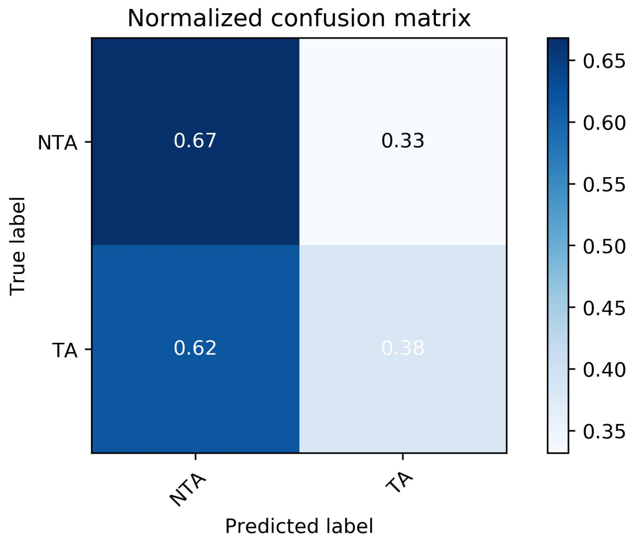
TP: 54%
TN: 46%
FP: 46%

XGBoost

TP: 36%
TN: 66%
FP: 64%

Naive Bayes

TP: 69%
TN: 31%
FP: 31%



Final Model: Random Forest

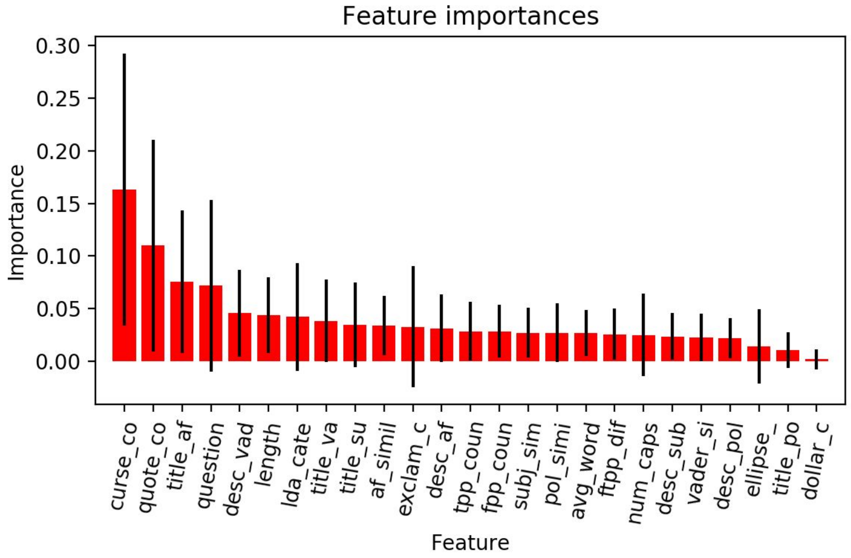
- Highest accuracy ()
- One of the lowest false positives
- The lowest false negative rates

Overall, all models had a 50% or higher false negative rate :(

Feature Importance

Top features:

- Curse count
- Quote count
- Afinn title sentiment
- Question mark count
- VADER description sentiment
- Length
- LDA category


Feature importances
Importance
curse_co
quote_co
title_af
question
desc_vad
length
lda_cate
title_va
title_su
af_simil
exclam_c
desc_af
tpp_coun
fpp_coun
subj_sim
pol_simi
avg_word
ftpp_dif
num_caps
desc_sub
vader_si
desc_pol
ellipse_
title_po
dollar_c
Feature