

# Problem:

Would a little validation be helpful in telling if you are right or wrong?

Are you becoming too mature for facebook venting? Or are you a venter whose family members don't talk to you at Thanksgiving and you don't know why?

Have you ever had to confront conflicts via email, and things spiraled out of control?

Sometimes, you just need to know...

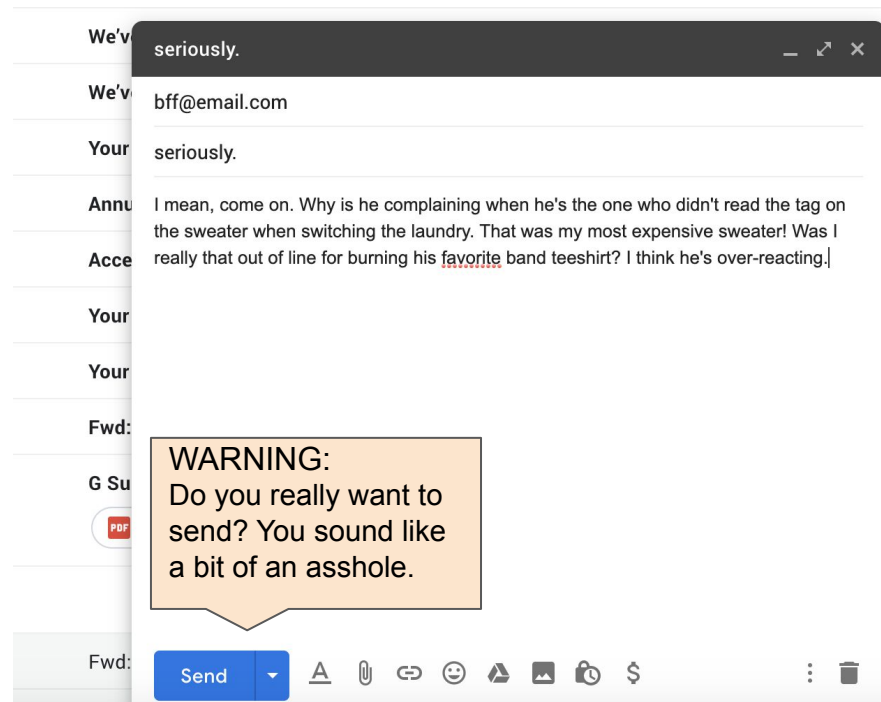
Am I the Asshole?

Meet:

assHo.ly

# Assho.ly will tell you if you need to check yourself before you wreck yourself.

- Stand-alone App
- Browser Extension



To develop the app,  
we need training data  
for our predictor.

# Thank you Reddit: r/AmItheAsshole



r/AmItheAsshole

A catharsis for the frustrated moral philosopher in all of us, and a place to finally find out if you were wrong in an argument that's been bothering you. Tell us about any non-violent conflict you have experienced; give us both sides of the story, and find out if you're right, or you're the asshole. See our ~~\*Best Of\*~~ "Most Controversial" at /r/AITAFiltered!

1.7m

Potential Assholes

25.3k

Judging you right now



Created Jun 8, 2013

JOIN

Example submissions  
include:

AITA because I ate more  
than “my share” of a 6 foot  
party sub last night?



AITA for telling my mom and  
stepdad that my stepdad  
isn't my dad?

AITA for saying 'I told you  
so' when my cat bit  
somebody?

# Methods for Utilizing Reddit to Inform the Predictive Model for Our App



548



r/AmItheAsshole · Posted by u/verrrryanonnn 2 hours ago

## AITA for putting grease on my car to catch the person/people who often hit it when parallel parking on my street?

I moved into an apartment that only has street parking, and in the few months I've been here, my car has gotten more and more dings and scratches. Mostly on the front and back where they appear to be caused by other drivers bumping it when parallel parking. It's a rusty old car but it's like the principle of the thing... It's still my car, my very first car that's like a baby to me, a car I was really proud of saving up for and purchasing as a teenager. I'm 23 now and she's still going strong. And it sucks to see its been backed into and run into regularly

Parking is tight here but imo there's no need to try to squeeze into a spot if it means bumping and scraping other cars... You can just park further away.

I got frustrated and bought a motion triggered dashcam but it interacted badly with my car's power and drained the car battery when left overnight, so I returned it.

Then I had a more old-school idea. My workplace has a machine shop and I

Commenters vote on assholeness with an acronym:

Judgment	Abbreviation
<b>You're the Asshole</b> (& the other party is not)	<b>YTA</b>
You're <b>Not the A-hole</b> (& the other party is)	<b>NTA</b>
<b>Everyone Sucks Here</b>	<b>ESH</b>
<b>No A-holes here</b>	<b>NAH</b>
<b>Not Enough Info</b>	<b>INFO</b>

Process for using the subreddit data:

Gave the submissions a 'vote' based on whether they garnered more NTA's or YTA's in their associated comments.

↑ ahhhtoomuchpressure **Asshole Aficionado [11]** 734 points · 2 hours ago  
↓ NTA. That's a pretty harmless way to figure it out and it's not like you're now keying their car or something. And it solved the problem. Grease comes off with soap, so they should be able to get it off in a car wash. They'll be fine.

↑ Soccermad23 114 points · 1 hour ago  
↓ I agree with this. It also means that if their car or paint gets damaged, its done by their own volition.

↑ Shpate **Partassipant [1]** Score hidden · 40 minutes ago  
↓ His coworkers are dumb. "That might fuck up their paint while they are hitting your car". How stupid can a person be?

+ cwinparr Score hidden · 24 minutes ago (3 children)

↑ serisianstrings 281 points · 2 hours ago  
↓ NTA. Clever! The car is your property, do whatever you want to it. Hitting your car with another is illegal so don't feel bad for those assholes - they should be leaving insurance information every single time it happens, so this is appropriate karma.

+ lonelyygirl69 Score hidden · 51 minutes ago (4 children)

↑ SassyTeacupPrincess **Partassipant [4]** 136 points · 2 hours ago  
↓ NTA Your solution was brilliant. Those two guys probably are bad at parking their own cars and feel like it's a personal attack.

The process for creating a predictive model based on the submission text and post title:

# Dealing with the problem of unbalanced classes:

For the training data, 5,000 submissions were narrowed down to 1,714.

Discarded:

- Ties (equal # YTA and NTA)
- Those without any votes in their comments
- A bunch of NTAs (so that the # of NTAs and YTAs matched)



# Using text data as features in a predictive model:

## Problems:

- A lot of words that don't add to the meaning
- Repeat words (same word in different forms)
- Non-numerical

## Solutions:

- Remove 'stopwords' from features (like 'the' & 'and')
- Streamline text by transforming words into their base forms ('stems' or 'lemmas')
- Transform word features into the frequency of use in each post, or their overall importance with CountVectorizer and TfidfVectorizer tools

# Models used:

- Logistic Regression
- Random Forest

# And how did the models perform?

Very poorly with 50-something percent accuracy!

GASP!

# So why is it so hard to predict who's the asshole based on words?

- It's usually just one or two phrases in a story that color the submitter as Asshole or Not Asshole.
- The same words in a different context in a different story could imply the opposite!
- Hard to detect insensitivity, outrage, selfishness.

# Sometimes it's just a few words that make a difference:

“So my fiancé (37M) and I (30F) are getting married in April. He has a 15 year old daughter ... Her mother (my fiancé's first wife) passed away 2 years ago from cancer. ... Anyways, recently she told my fiancé that she'd like to read a poem at our wedding during the speech portions. .... I read the draft of it and am incredibly unhappy with how it sounds. .... I get that she's still grieving and she's a kid and all that. But I really don't want this read at our wedding ... I've talked to my fiancé and he sees 100% NOTHING wrong with the speech which has frankly upset me beyond what I'd imagined. I feel like he's insisting on prioritizing her during a day that should be about me and him ONLY, not his kid or ex. ...”

[https://www.reddit.com/r/AmItheAsshole/comments/ew2zcr/aita\\_for\\_not\\_wanting\\_my\\_future\\_stepdaughte\\_r\\_to/](https://www.reddit.com/r/AmItheAsshole/comments/ew2zcr/aita_for_not_wanting_my_future_stepdaughte_r_to/)

# Subtly Different versions of the same story can result in different judgments.

- Intentionality
- Self reflection
- Blame and responsibility

# Who are we talking about?

How does the computer know which words are associated with which person in the story?

- Describing the same action performed by the narrator or someone else can result in YTA or NTA votes respectively.

# Considerations for improving the model:

- Divide story into sentences, and do sentiment analysis on sentences containing pronouns like 'I' or 'they' or "my" ("my bf") which have directionality that might be hard for off-the-shelf NLP algorithms to pick up in the whole text
- Analyze strategy for stemming and choosing stopwords because tiny variances can have big effect on meaning. Focus on phrases and ngrams?
- Can we specifically find feelings of remorse or reflection?