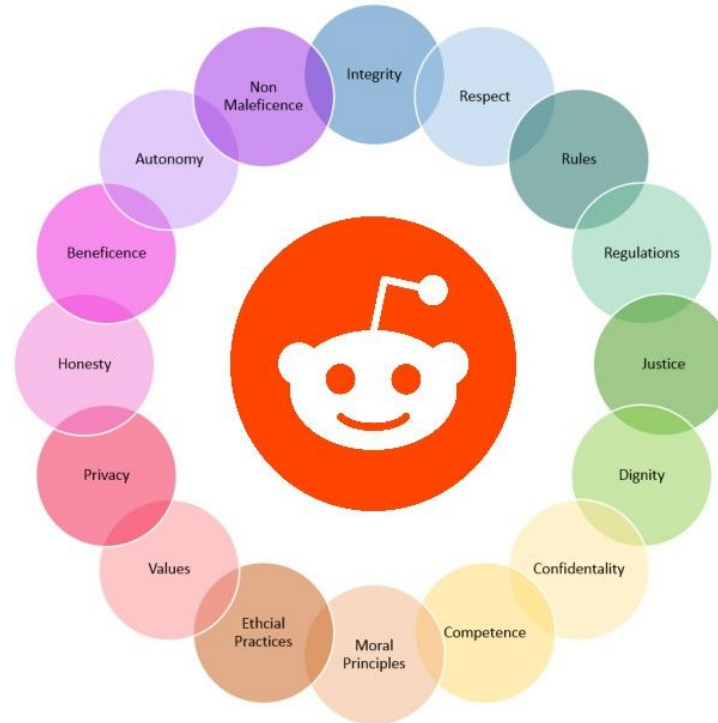


Interpreting Ethics Using NLP and Social Media



Ethics in an Information-Saturated World

With the decline of print and television as media outlets, marketing and politics have moved online. Subsequently, advertising and political propaganda have shifted to the internet and are thriving on social media sites.

As our society is bombarded with media from all outlets, consumers need to be confident the information we are consuming is not only legal and factual, but also ethical.



Examples of Unethical Media

- The wellness company Goop was fined \$145,000 in 2018 for “misleading consumers” by selling “jade eggs” claiming they balanced hormones and regulated menstrual cycles “inner judge flower essence”, claiming it prevented depression.



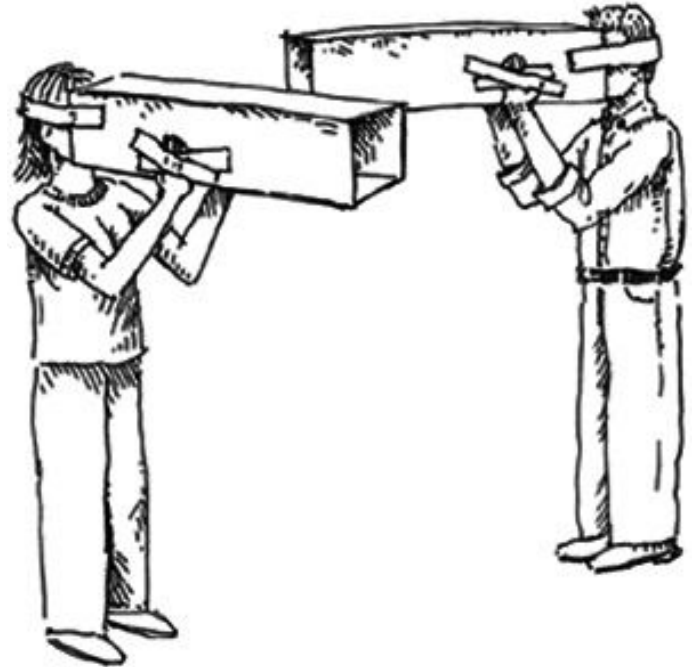
Examples of Unethical Media

- The wellness company Goop was fined \$145,000 in 2018 for “misleading consumers” by selling “jade eggs” claiming they balanced hormones and regulated menstrual cycles “inner judge flower essence”, claiming it prevented depression.
- In 2016, the fake news site Your News Wire first published the conspiracy theory known as “pizza gate” which resulted in a DC restaurant and its staff being assailed by a man with an AR-15-style rifle, an arson attempt, and death threats.



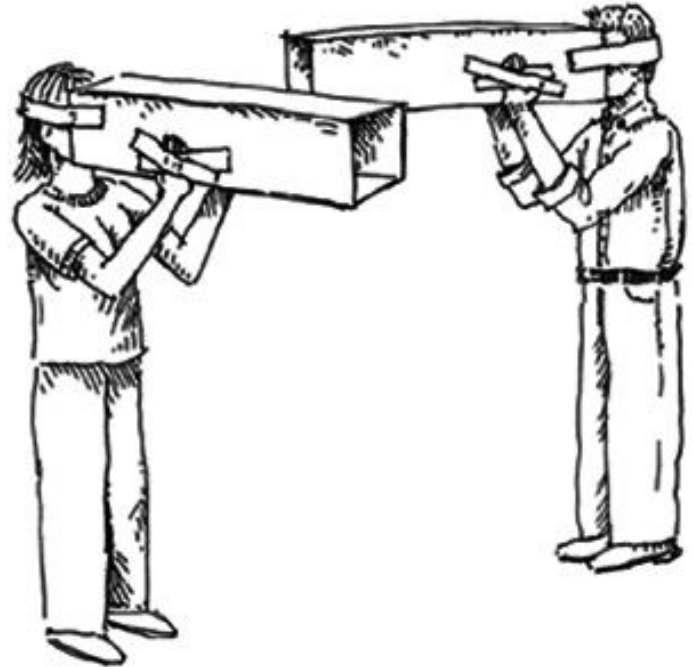
The Problem of Defining Ethics

- How can we accurately categorize ethical vs unethical behavior?
- How can we be sure the person collecting the data isn't using personal bias when defining unethical behavior?
- What if societal norms change?



The Problem of Defining Ethics

- Solution: Collecting unethical behavior that people are self-classifying as unethical.
- Using Reddit and Natural Language Processing to train a regression model to recognize the difference between posts in /r/LifeProTips and /r/UnethicalLifeProTips

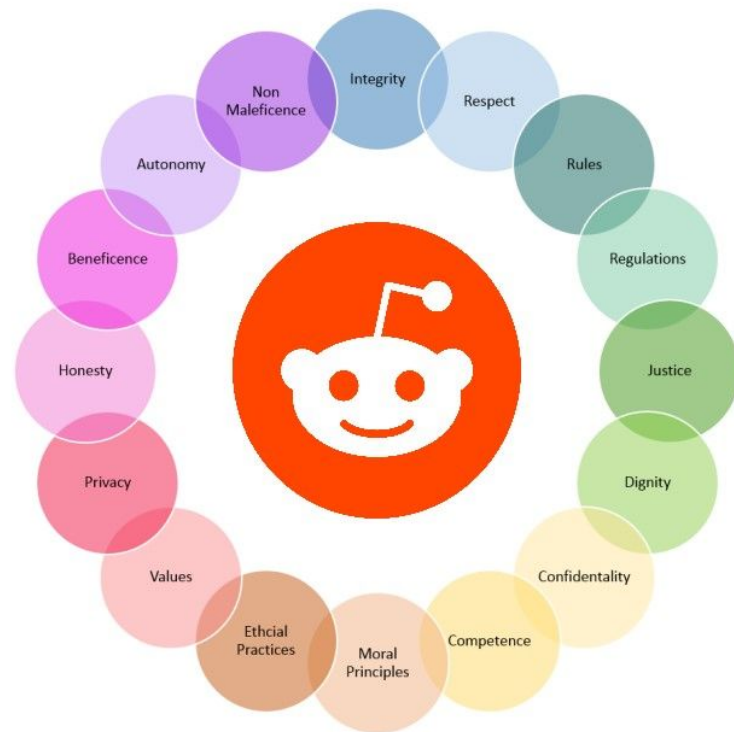


Our Data Source

Reddit is a social news aggregation site and the 19th most popular web site in the world with 330 million active monthly users.

Reddit is composed of “subreddits”, or topic-based communities where members create, discuss, share, and vote on posts.

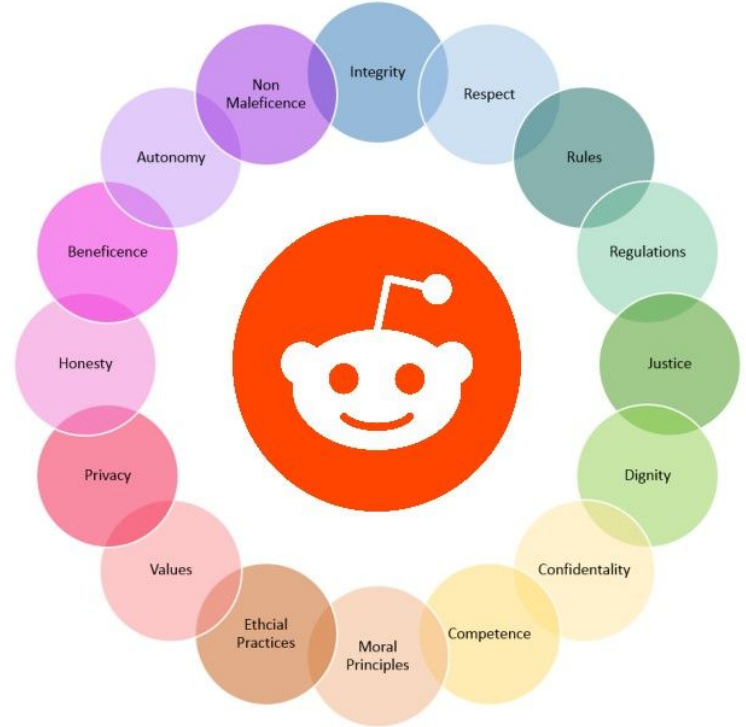
Our model is built on two advice-based subreddits.



Our Model

20,000 raw posts were collected from reddit, equally divided between /r/LifeProTips and /r/UnethicalLifeProTips.

After exploratory data analysis, our dataset contained 19,852 entries.



Our Model

Using a logistic regression model, we are able to correctly identify unethical tips with 81% accuracy.

Exploring Life Pro Tips

- “A Life Pro Tip (or LPT) is a concise and specific tip that improves life for you and those around you in a specific and significant way.”
- 17.7 million members.
- It is against the rules of the forum to post common sense tips, common courtesies, or unethical tips.



Exploring Life Pro Tips

77.1k  Posted by u/cool_much 27 days ago     2 

LPT If a bee colony has set up a hive that you need removed, don't call the exterminator. Beekeepers will relocate the hive often for free. This is an easy way to be nice to the bees and to your wallet :)

Animals & Pets

74.1k  Posted by u/41sa 2 years ago  

LPT: If you are given a prepaid debit card as a gift, save it after you spend the money. You can use it to sign up for free trials online without worry of being scammed.

Money & Finance

57.6k  Posted by u/killgoresalmonman 1 month ago     

LPT: If a credit card of finance offer excludes residents of Iowa, you probably shouldn't apply.

Miscellaneous

If you pay close enough attention, you might notice that a fair amount of finance offers exclude Iowa residents. This is because Iowa has some of the most robust consumer credit laws around that - among other things - put limits on interest rates and late payment penalties, AKA, helping prevent predatory loan practices. This isn't to say that *all* companies who don't do business with Iowa residents are predatory, but I don't think I'd want to take out a loan with a company that won't do business with consumers who have laws that protect them from being abused by creditors.



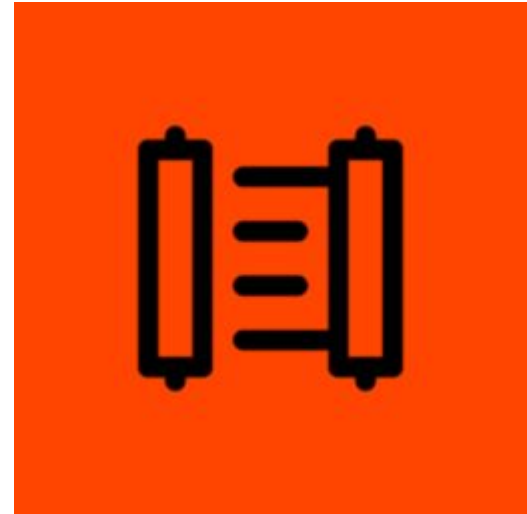
Top 50 “Ethical” Tag Observations

- Language has more of a focus on introspection:
Remember, Life, Important, Consider, Should, Learn, Safe, Helps, Better, Feel
- Language is frequently in the 2nd person tense and has a sense of immediacy:
Don't, You, Your, Always, Especially, Should, Time, Before



Exploring Unethical Life Pro Tips

- “An Unethical Life Pro Tip (or ULPT) is a tip that improves your life in a meaningful way, perhaps at the expense of others and/or with questionable legality. Due to their nature, do not actually follow any of these tips—they're just for fun.”
- 1.1 million members.
- It is against the rules of the forum to post explicitly illegal advice.



Exploring Unethical Life Pro Tips

↑ Posted by u/midwrestless 8 months ago 📄 🔒 🌐 🏆 3 💰

60.6k



ULPT: Look up your buildings washer/dryer model on eBay and order a key for it. I haven't paid for laundry in years and it cost me \$8.00! Sleep like a baby knowing you're not paying for on-site laundry.

Productivity

↑ Posted by u/cannabisqueenkitty 1 year ago 📄

23.2k



ULPT: Drug tests at work? Bring homemade food for the office with slight amounts of thc. They can't fire everyone :)

↑ Posted by u/Sylocule 1 year ago 📄

39.2k

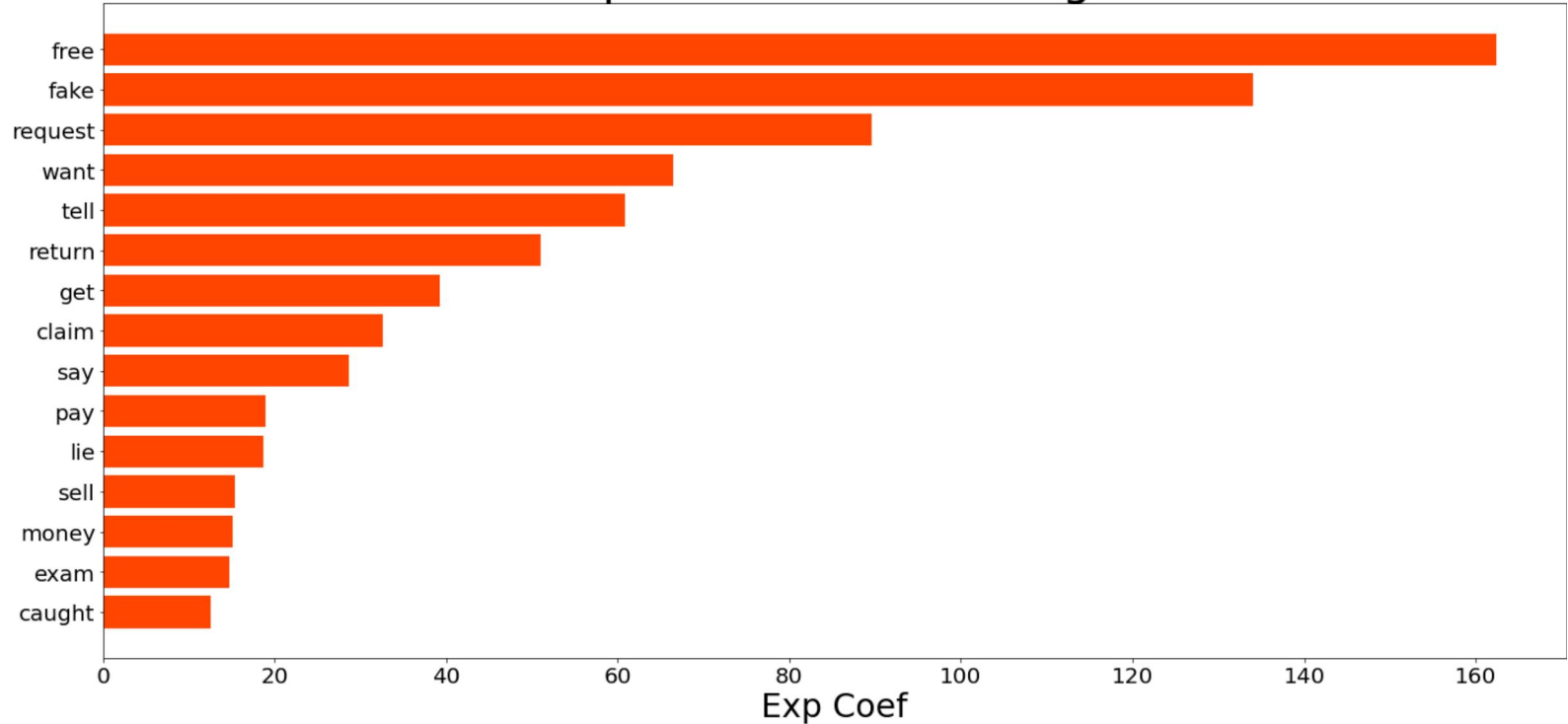


ULPT: If you glue a dead wasp to the palm of your hand, you can hit your boss on the back of the head as hard as you like and act like you saved him.



Exploring Unethical Life Pro Tips

Top 15 "Unethical" Flags



Top 50 “Unethical” Tag Observations

- The obvious:
Free, Fake, Claim, Money, Lie, Illegal, Police, Caught, Cheat, Pretend, Excuse, Drunk
- Language is very direct, instructional:
Get, Tell, Say, Order, How
- Crude and unexpected:
Christmas, Shit, Halloween, Fart, Holiday



Conclusions

Using Natural Language Processing, our model is able to accurately classify 81% of unethical posts. The model is intentionally overfit and on training data, can correctly identify unethical posts as high as 92% of the time.

Although the model does not have exceptionally high accuracy at this time, it can be used an effective tool to influence consumer decisions about the implicit ethics of online content.

