**Case Study 5**

Sentiment Decider

Teaching Resources

**Summary**

This case study is focused on teaching the students what sentiment analysis is, as well as how modern online platforms can use sentiment analysis for their benefit. This unit is a great opportunity to teach the students the coding first (by starting with the Sentiment Decider activity) and then doing the class discussion / lecture after the students have already had time to understand what sentiment analysis is through creating their own sentiment analyzer. The goal for this unit is to get students thinking critically about what some of the harms of sentiment analysis can be, and how their personal decisions online when it comes to accepting or rejecting recommendations might help or harm certain individuals.

**CS Topics That Are Covered in This Unit**

1. Functional Decomposition
2. File I/O:
   1. Reading from text files
   2. Reading real Amazon reviews
3. Sentiment Analysis
4. Dictionaries
   1. Defining, Storing, Updating, Accessing
5. Debugging

**Ethics / Social Science Topics That Can Be Covered in This Unit**

1. Ground truth problem (the dangers of assuming that one person's positive / negative is the same as another's)
   1. mechanical turk + the people who drive how AI works through labeling
2. Ranking and producer fairness problem
3. Recommender systems (high level) + Page Ranking
   1. Producer-side fairness + multistakeholder fairness
4. "Hacking" or “Cheating” the algorithm (Amazon + companies that are tricking their algorithm through reviews - assign listening to the reply all episode)
5. Conflation problem (if you are only using a binary classification here (either “positive” or “negative” sentiment) what about all of the reviews that are almost negative or almost positive? (the neutral reviews)
   1. Is there any danger in this?

**Resources (Reading / Watching To Assign the Students)**

1. [The Magic Store](https://gimletmedia.com/shows/reply-all/brhow4) - Reply All Podcast Episode
   1. A fun podcast episode telling stories of how people and organizations have “gamed” Amazon’s recommendation algorithm. This is a great way to get them interested in the “Mamazon” product reviews for the Sentiment Decider coding activity

**Discussion Questions**

1. Do you think that sentiment analysis calculations are accurate?
   1. Yes, I think they are accurate most of the time
   2. No, I think they are wrong most of the time
2. Imagine you were given a list of words and you had to label which of them were positive or negative -- do you think your categorization would match everyone elses?
   1. Yes, I think that it’s fairly objective which words are positive and which words are negative
   2. To a certain degree, I think that most of the words I categorized would match another person’s categorization -- but we might disagree on a few of the words
   3. Not really, I think I might have some words similar with another person’s categorization, but we would probably label a lot of words differently
   4. Not at all, I think that there would be a lot of words that were different between my categorizations and someone else’s
3. How do you feel about people “gaming” algorithms?
   1. I feel like it’s totally fine. It won’t cause harm to anyone and it’s not like you are breaking the law…
   2. I have mixed feelings about it. I can see how it can be harmful to people, but also isn’t it the tech company’s fault for not making the algorithm more robust?
   3. I’m not okay with it. Clearly “gaming” an algorithm can be harmful to people, this shouldn’t be okay for people to do.
   4. I don’t really feel strongly about it. People are gonna do what people are gonna do.
4. Have you ever tried “gaming” an algorithm? (e.g., showing up higher in the search results, trending on social media, etc…)
   1. Yes, I’ve tried and failed
   2. Yes, I’ve tried and I think I figured out some ways to “game” the algorithm!
   3. Nope, never tried to
5. Do you think it’s better for algorithms to be fully transparent to the public (even if it means people can “game” them) -- because then everyone at least knows what the dangers are?
   1. Yes - I think that it would be better if companies were more open about what their algorithms do and how they rank / recommend things
   2. No - I think that if companies become more open about it, it will give too much power to people that want to abuse it
   3. Mixed feelings - I could see this being good and bad
6. Have you ever thought about the impact that recommendations / ranking can have on the producers? (e.g., the movie production company, the book publishers, the amazon sellers, etc..)
   1. Yes - I’ve thought about it a lot and I try to make sure to always look for the recommendations that help out the underdog
   2. Yeah, I’ve thought about it… but usually popular items/sellers/movies,etc.. Are popular for a reason. So, why wouldn’t I click on the highest ranked recommendations?
   3. No I haven’t really thought about it, usually I just click on whatever is recommended to me first
   4. I haven’t really thought about it, usually I don’t click on things that are recommended to me
7. Do you feel like recommendations / ranking are generally pretty good for you?
   1. Yes - it’s creepy how accurate my recommendations can be
   2. No… they aren’t usually useful to me
8. Would you rather have recommendations that are less “personalized” to you if it meant they were helping out a small-scale producer?
   1. Totally!
   2. I’d consider it, if I could have the option to toggle between the two kinds of recommendations
   3. I’d prefer to have my recommendations be more personalized
   4. I’d never consider it
9. Explain how you feel about sentiment analysis in one word.