Status Report

Cameron Hatler

Nov. 20, 2024

1 Members

It is still just me working on this project. I have a lot of group projects going on at the moment and am traveling a lot, so I didn't think it would be fair to others if I joined the group and was hard to meet up with to work on the project.

2 Summary

- Finalized Project Plan
- Found dataset to use
- Decided on models to compare
  - Basic implementation from assignment 7
  - Multinomial Naive Baise model from scikit package
  - Classification supervised neural network from scikit package
- Looking at ways to visualize comparisons between the models

3 Results

I don't have any results yet. It took me a little longer than I thought it would finding a good dataset to use and nailing down the final details of the project. But now that I know exactly what my plan is, what models I'm using, what data I'm looking at, and how to compare them I think I'll have a lot more work done.

4 Ethics

Where data comes from is really important to think about. People may not be aware that their data is being used in these models, which is definitely an ethical gray area. It is also important to think about what data you are using and where the findings can be extrapolated to.

5 Problems

Now that I have the data, I've had to do a little bit of a pivot from what I was initially planning. There wasn't any good/clean data that included tweets along with the likes. One solution was to scrape X, but that seemed difficult/morally questionable, so I decided to switch to comparing sentiment models on already created twitter data. One problem I can see for this project is the models being hard to setup, but I think it should be alright.

6 Hours

11/15 an hour looking for data and planning out the project.

11/16 an hour thinking about how to take care of the wide variety of likes there are (figuring multinomial isn't good solution for this so looking into regression models)

11/20 a couple hours realizing the data I was looking at was really bad, but couldn't find other data that had likes/tweets so realizing I would have to switch to a sentiment analysis. Then figuring out a couple different models to test/compare

7 Code

I will slightly adjust the code from assignment 7 to allow a 'neutral' category. I will also have to create some python code using the psykit package to try out a couple already implemented models and compare them to each other and the sentiment model from assignment 7.