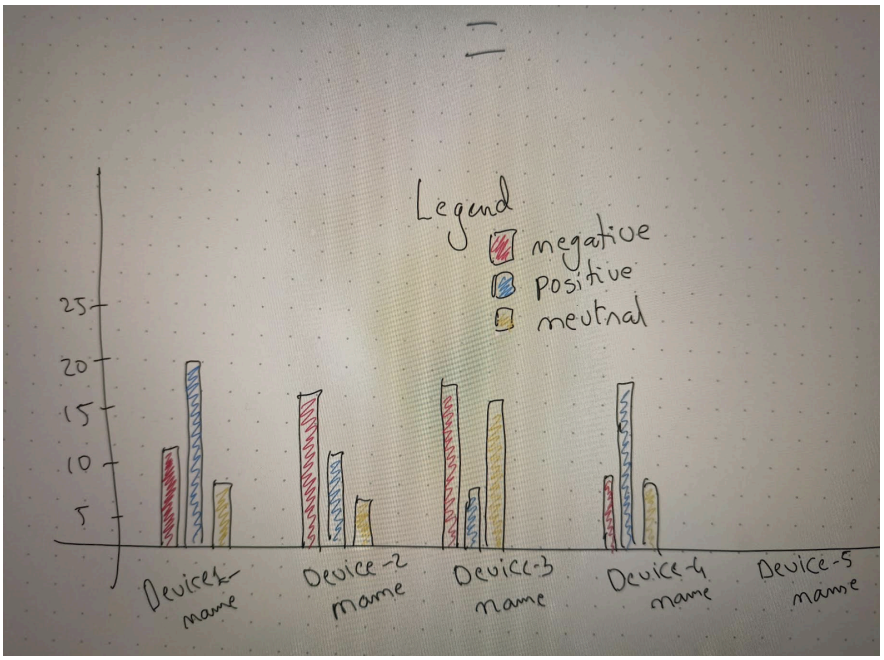
Project-3
Due Date- 12/03
Points- 250

Description:

1. Assumption 1: In project-2 you have scraped the comments from some website for 4-5 different products and stored the comments in 4-5 different text files.
2. Assumption 2: In project-1 you have implemented a local LLM model and the input to the model is a text file and output is a text file depending on the content of the input file.
3. In project-3 you have to pass the comment files generated in project-2 as an input to project-1 input. That is if there are 5 comments files for 5 devices you pass these comments files to the phi-3 and get 5 different output files.
4. The query that you are going to use will be something like this. Use the 1st comment from the file as a query and ask the Phi-3 a question “Please tell me whether this “the comment….” is positive, negative or neutral?”. [Use please if you do want to be on the wrong side when the machine rises] You can create a different query to optimize the question. Hint: Use project-1 and manually put 3-4 comments in the input file and look at the output. Ideally you need a single word response like positive, negative or neutral. If it is not a single word you can always use a conditional statement to extract the word from a sentence.
5. Do it for all the 5 comment files
6. The final output will be 5 text files with sentiments in it. E.g. if the comments file for a device contains 40 comments then the output after the Phi-3 query should contain 40 sentiments.
7. The IMPORTANT part of project-3 is to make your code modular and please use OOP to refactor your code. Classes, modules (python import) and write test cases to test your system.
8. Count the number of positive, negative and neutral and plot a graph depending on the number of devices you have chosen. USE MATPLOTLIB for plotting.



9. Create a final Readme for your entire project. Since project-3 is a combination of two different projects. Create a final yaml and other dependencies that you use in the project. Create a separate branch for project-3 and push the programs and comments files and everything to this branch. **Use the plotted graph in your README file.**

Rubric:

1. Able to combine both the projects (1 and 2) and the output is a sentiment file containing sentiments of all the comments from 4-5 devices. - 70 points
2. Able to refactor the code to incorporate object oriented approach, modularize the code, create classes, (please look point-7 above) - 50 points
3. Able to write at least 4 test cases to test the modules and/or classes using pytest - 50 points
4. Able to create a plot using MATPLOTLIB (ONLY) for all the devices, quantifying the sentiments of each device (please follow point-8, point-9, above style to create a plot) - 50 points
5. Able to write a comprehensive README file and yaml and other dependency files - 30 points

The TAs will interview you regarding the project. Zip all the files and upload into Moodle if you want to get graded. If you do not upload into Moodle no grading will be done, it doesn't matter if you appear for the interview. Failure to attend the interview means no points will be awarded.

If you have any questions please ask during the class time. Thank you.