Weekly Mentor Meeting: Every Sunday 2:15-3:15 PM

Our github repo: [acmucsd-projects/fa22-ai-team-3 (github.com)]

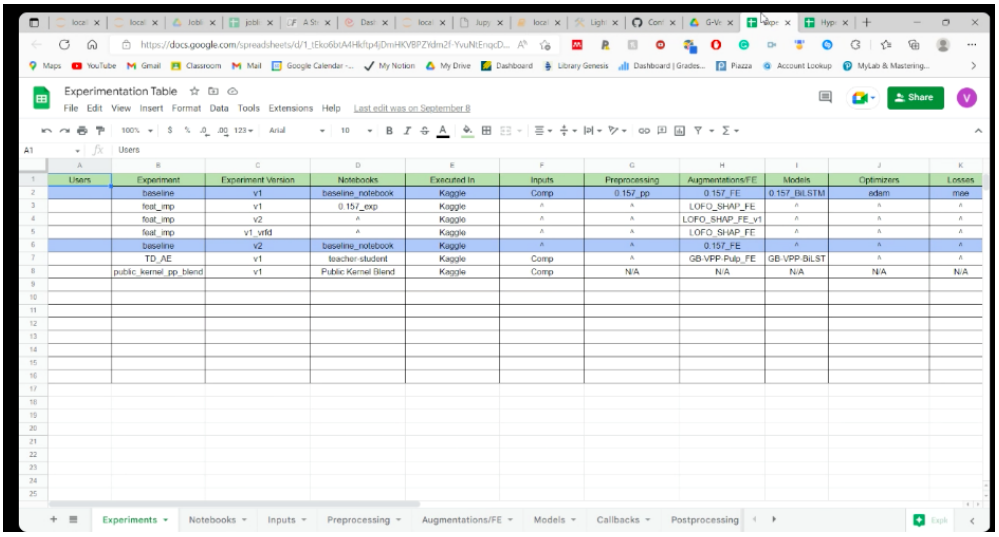
Meeting Time: 10/30/22 2:15-3:15 PM in CSE Basement
Attendees: Max, Chuong, Rebecca, Arvin, Siya, Rohan, Vincent

Action Items

- We will experiment and play with Max's pipeline & start the spreadsheet mentioned below
    - Only things you change are the model and hyperparameters. Do not change the actual pipeline/ pre processing
        - Use Vincent's notebook as a guide (on discord)
    - Use the same cv split: cv=5 splits (train on 4 segments and validate on 1 segment)
    - **Look @ timeline below for more info**
- Save work to folders:
    - We want one folder for every model.
        - Within each model folder, have one folder for every set of hyperparameters you sweep over and save the actual gridsearch psp file and keep it in that folder.
            - If the file size is too big, put it in a dropbox and ensure everyone has access to the dropbox.
    - Save spreadsheet of experiments to repo

Summary of Meeting

- Pipelines
    - Optimizing the pipeline is not a priority; instead, we should stick to one pipeline and use the model
        - Recommended: Standard scaler
    - **Keep one pipeline constant** with various models
- We also need to start a spreadsheet to keep track of and organize the experiments we are running and the different hyperparameters we use

- 

    - Ie keep everything constant and test with/without transformer to check how powerful that component is
- We also decided to use Max's pipeline and fixed it

Timeline

- In this week, we will experiment (set up code to make experimentation very easy)
    - Set up a function that says the best estimator
    - Keep track of hyperparameters, scores, and rank
        - Manually log it into a spreadsheet OR make a function that creates a data frame and logs all the scores to the dataframe
    - Save the entire dictionary and gridsearch object pspp so everyone has access to it
- After this week, deploy and make an app
    - Put reports and model on there
- If we have time, we can make report detailing experiments