Project Report 3

Ethics Engineers

1. Summarize the progress you've made on your project over the last week. Remember, your answers should be BRIEF!
   1. We have cleaned the behavior (user information) and news (article information) dataset by adding correct column names, splitting user history into proper date/time variables, and merging the two datasets so that we have easy access to info about each article each user has clicked on (which is tricky due to how large the files are; we plan to use LazyDesk to solve this). We have also begun some exploratory analysis, like examining NaN values in user history and counting how many times each article is in a user’s history and in a user’s recommended list (by exploding the dataset). We are currently doing some of our work in Google Colab and plan on transferring everything over to Git next week.
2. Have you made any changes in direction? If so, please describe what the change is and why you chose a new path. Remember, your answers should be BRIEF!
   1. We have not – we are committed to exploring this dataset in the same manner we have discussed before.
3. Describe the next steps you plan to take to answer your question. If working in a group, make sure you roughly describe each member's plans. Remember, your answers should be BRIEF!
   1. Next steps consist of finalizing the merge and starting to dive into more complicated analytical methods, such as which types of articles are suggested the most and which kinds of users are most likely to look at certain articles. These two areas are currently the ones we are most interested in discussing in our final product. This process will still include all three of us as we split off into different analytical directions; we have not decided who will look at which issues yet.
4. Github