Team name: Outliers
Team members: Yvonne Qiu, Zoey Ye, Jiahuan He, Zenan Wang, Taixi Lu
Date: Mar 3th, 2022

This journal has to be uploaded to 1) a dropbox folder (link) **AND** 2) posted to your project website every two weeks by 11:59 pm on the second Friday. The first due date is Mar. 4th, Mar. 25th and the third due date is April 7th. When uploading to the dropbox folder, change the file name to "Team_name_MM-DD-2022" where MM = month and DD= day of upload).

Team roles for this report (write down name):

Facilitator(s): Zenan Wang Taixi Lu

Recorder(s): Zoey Ye

Deliverer(s): Jiahuan He

Planner(s): Yvonne Qiu

Contact: Taixi Lu

See last page for description of roles. Obviously one person can take more than one role or there can be more than one person per role or make your own roles!

0. **Describe briefly what the main goal of your team is (so the peer reviewer has some context). E.g. we are working on image classification for blah de blah. Our goal is blah de blah etc. In the initial part of the semester before your proposal it is ok to put down "we are still coming up with ideas on team project".**

    We have gathered tweets data pulled form Twitter, and our main goal is to use the existing tweet posts, date, and location information during the COVID pandemic to predict the risk level of that area and the sentiment of the person who sent those tweets. This is important

since we are all under tremendous pressure under COVID, and mental health has become a big problem.

I. **What was done during the report period regarding the project:** If you want to include code include this in the Appendix. Describe what the group did (including contributions of individual team members) with regards to the group project during this report period. Give enough details so I understand what you folks have been doing over the week. Include dates of your meeting(s) and who met on these days.

We have decided on the dataset that we are going to use, and we are trying to clean the data a little bit e.g., filter out invalid data from the dataset. The source of the data is Kaggle, and the link is posted on the [home page](#) of our website.

Taixi will be in charge of the detailed EDA, and other group members will help him constructing the analysis. Yvonne will be doing the biweekly report, concluding what we have discussed during the meeting. Other group members will help for modifications. We have had a meeting on Mar 3rd for two and half hours. We are going to apply logistic regression, recurrent neural network, and more machine learning techniques in our presentation.

II. **What were obstacles faced if any in working on the project?** This could be technical (like not being able to implement or understand particular techniques) or time issues (midterms for other courses etc).

Group cooperation: How to get every people participate into this project, and actively making progress on it.

We are also exploring new ways to do the EDA. As our data is tweets and mainly text message, we will need to find out how to explore the data. It will be somewhat different than what we have seen.

III. **What is the plan for the next reporting period including what each team member is planning to work on. Describe goals and potential timelines (" I plan to finish understanding x to see if it can be implemented for our project by Wednesday etc".)**

We will have finished our EDA by the next reporting period, and we are going to explore the relationship between the variables we have chosen (e.g., strong or weak correlation). We will also start to do some prediction with the techniques we have seen in class.

While in the biweekly document above you will describe what your team did with regards to the team project (with proper attributions of who did what in the week) there are 4 pre-defined roles. I urge you to have different people do these jobs every week so that you gain experience in each of the jobs. There can also be more than one person per job for example 2 people recording the weekly journal.

Facilitator: Manages the group for this week including setting up times for group members to meet, making sure everyone has a say in the meetings etc.

Recorder: Person in charge of recording the meetings as well as the happenings of the past two weeks and describing what was accomplished in the meeting and writing up this report.

Deliverer: Person in charge of checking the entire report and uploading the file to dropbox folder and project website, as well as the representative of the group getting in touch with the instructor.

Planner: Person in charge of what will be happening next two weeks as well as thinking about longer term goals (what more needs to be done for the project).

Team contact: Person I can email if I see any issues in the biweekly report instead of mass spamming everyone in the team.