Log

Week 1:   ()
    2/19/20   First meeting

Week 2:   **(7hr 36 min)**
    2/26/20   Second meeting
    2/24/20   11:00PM to 1:30AM   (2hr 30 min)
        -Downloading and trying to set things up
    2/25/20   4:00PM to 5:45PM   (1hr 45 min)
        -learning how to code with python
        11:15PM to 12:15AM   (1hr)
        -working on code
    2/26/20   4:00PM to 4:30PM
        -learning how to read a excel file in python
    2/28/20   8:37PM to 9:40PM   (1hr 3 min)
        -fixing tensorflow
        -create a Github and inviting Professor
    3/1/20   10:20PM to 11:38PM   (1hr 18 min)
        -watching video on Tensorflow on Colab

Week 3:   **(6hr)**
    3/4/20   Third meeting
    3/8/20   8:40PM to 9:48   (1hr)
        -added some code that actual works in Tensorflow
    3/10/20   8:00PM to 10:00PM   (2hr)
        -try to read a csv file and then try to turn it into a dataset
    3/11/20   1:33PM to 4:30PM   (3hr)
        -trying to write code to create a dataset

Week 4:   **(2hr 38 min)**
    3/11/20   Fourth meeting
    3/14/20   5:37PM to 8:15PM   (2hr 38 min)
        -upgrade Tensorflow to 2.1.0
        -tried out some sample code of KNN
        -started to rewrite project code

Pandemic and now everyone staying home

Week 5:   (8hr 30 min)
    3/18/20   No meeting
    3/18/20   5:45PM to 7:25PM   (1hr 40 min)
                          -putting some sample code on Github
                          -added more code to project
    3/23/20   1:10PM to 2:15PM   (1hr 5 min)
                          -looking up information
    3/24/20   12:00PM to 3:45PM   (4hr 45 min)
                          -
    3/28/20   3:30PM to 4:30PM   (1hr)
                          -testing out other code

Week 6:   (2hr)
    4/1/20   No Meeting
    4/3/20   2:15PM to 4:15PM   (2hr)
                          -working on code

#first meeting after pandemic hit

Week 7:   (9hr 30 min)
    4/6/20   Fifth meeting with professor on Skype
                12:45PM to 1:15PM   (30 min)
                          -getting ready for the meeting
                3:25PM to 6:00PM   (2hr 35 min)
                          -after meeting with professor
                          -download Jupyter notebook
                          -started working on new code
    4/7/20   1:30PM to 3:30PM   (2hr 30 min)
                          -working with new code that actually work
                          -change the data
    4/9/20   4:50PM to 6:05PM   (1hr 15 min)
                          -working on code
    4/10/20   1:33PM to 4:10PM   (2hr 37 min)
                          -working on code
    4/13/20   1:30PM to 3:33PM   (2hr 3 min)
                          -working on code

Week 8:   (6hr 7 min)
    4/15/20   Sixth meeting with Professor
                12:30PM to 1:15PM   (45 min)
                          -looking at code before meeting

2:00PM to 2:32PM   (32 min)

    -meeting with professor

    -discuss on how to edit the dataset

        -removing 'NaN'

        -keep doc_id 1 to 5

        -delete doc_id 6 to 7

2:35PM to 4:05PM   (1hr 30 min)

    -fixing the dataset

4/16/20   2:50PM to 4:40PM   (1hr 50 min)

    -working on the code

    -fixing some stuff on the dataset

4/17/20   4:00PM to 4:30PM   (30 min)

    -writing the interim status report

4/20/20   2:00PM to 3:00PM   (1hr)

    -playing around with current code

    -found another code that also work and might use as backup

Week 9:   **(8hr 20 min)**

   4/22/20   Seventh meeting with Professor

1:00PM to 1:30PM   (30 min)

    -fixing/testing things before meeting starts

2:00PM to 3:05PM   (1hr 5 min)

    -discuss what to do for this week

    Step 1:

        Label = 'User_ID'

        Don't use = 'Doc_ID', 'Phone_ID'

    Step 2:

        -try to find a way to combine 2 row into 1

            -using the original csv file

3:15PM to 4:30PM   (1hr 15 min)

    -fixing dataset and code to finish step 1

4/23/20   1:30PM to 4:00PM   (2hr 30 min)

    -figuring out how to read the results

    -working on step 2

4/25/20   2:00PM to 3:00PM   (1hr)

    -working on code

4/27/20   1:15PM to 3:15PM   (2hr)

    -trying to do step 2 from meeting

    -couldn't find a way to do it

Week 10:   **(8hr 10 min)**

4/29/20   Eighth meeting with Professor

        2:12PM to 3:12PM   (1hr)

            -meeting with professor

            Step 2:

               -create a script to manipulate csv file

            Step 3:

               -print confusion matrix and classification report

               -save them to some variable

            Step 4:

               -replace 'NaN' with the median of the column

        3:15PM to 4:15PM   (1hr)

            -working on the dataset

            -finished everything from step 4 except replacing 'NaN'

5/1/20   4:00PM to 6:25PM   (2hr 25 min)

            -modifying dataset

            -finished **step 2**

5/4/20   12:30PM to 2:45PM   (2hr 15 min)

            -modifying the code

            -fixing up the dataset

5/5/20   12:00PM to 1:30PM   (1hr 30 min)

            -modifying dataset

Week 11: (8hr 14 min)

  5/6/20   Ninth meeting with Professor

        2:00PM to 3:06PM   (1hr 6 min)

            Step 5:

               -replace 'NaN' and 'inf' with the mean of column

            Step 6:

               -script that combine the previous row with the current row

            Step 7:

               -run the code with fixed original csv file

               -run the code with the combined rows file

        3:10PM to 4:20PM (1hr 10 min)

            -working on the stuff we discuss in the meeting

  5/9/20   1:30PM to 2:30PM   (1hr)

            -working on new script

  5/11/20   1:00PM to 5:33PM   (4hr 33 min)

            -working on dataset

               featMatVersion2_6

                  -fill in column 3

featMatVersion2_7
        -fill in the rest of columns

featMatVersion2_8
        -found a 'NaN' at column T row 10970

5/12/20   3:25PM to 4:50PM   (1hr 25 min)
        -fixing up dataset
        -rewriting some code to run new dataset

Week 12:   (8hr 56 min)

5/13/20   Tenth meeting with professor
      1:57PM to 2:05PM   (8 min)
        -short meeting because professor has something to do
        -only discuss what's left to do
      2:10PM to 5:00PM   (2hr 50 min)
        -fixing the problem with database

5/14/20   1:25PM to 5:00PM   (3hr 35 min)
        -still fixing the dataset
        -row 888 column 'AA' and 'AC' are missing numbers
        -fixing code

5/16/20   3:15PM to 4:00PM   (45 min)
        -

5/19/20   3:02PM to 4:40PM   (1hr 38 min)
        -converting result to a csv file

Week 13:   (59 min)

5/20/20   Eleven meeting with professor
      2:00PM to 2:16PM   (16 min)
        -discuss what left to do
      2:17PM to 3:00PM   (43 min)
        -trying to write code to print result into a csv
          -couldn't find a way