SLACK NAME: @Ini

TEAM: MOBILE KIDSTORIES

TEAM’S TASK:

Build a Kids’ stories mobile app that automatically switches to night mode in the evening and also suggest stories based on previous books read by the user.

MY CONTRIBUTION:

Building a popularity recommender system that recommends popular stories to users based on the age of the user.

HOW IT IMPROVED THEIR PROJECT:

My Machine Learning Model helped improve their project by recommending popular stories based on the age of the user. The users of the app can know the most trending stories recommended for their age.

WORKFLOW:

I created two .csv files based on the database that was given to me by the team lead. The first dataset, Kidstories.csv had the following features [Book\_ID, Title, Category, Body, Author, Likes, Dislikes]. The second dataset, Age\_interaction.csv had the following features [Book\_ID, Age].

Initially I didn’t have the likes and dislikes column, so I had to improvise and build a model that showed stories recommended for a user based on the user’s age. When the team submitted their app for presentation, I was able to download the app and extract the users’ ratings to get the likes and dislikes column. I updated my model to show popular books for a user, based on the user’s age.

HOW IT WORKS

I imported the pandas library and I read the two datasets and stored them in data frames using the imported pandas library. The data frames where stored in two different variables; age\_int and stories. With the age\_int data frame, I printed all the books assigned to the age inputted by the user with the .loc method. After getting a list of stories for that inputted age, I used the .sort\_values() method to sort the stories based on likes and dislikes. Likes were sorted in descending order, while dislikes were sorted in ascending order. So the popular stories are the stories with the highest number of likes and lesser number of dislikes. I then merged the sorted data frame to the stories data frame, to get the names of the stories. The merge was based on the Book\_ID. From the merged data frames, I selected only the ‘Title’ column to be displayed as output using the .loc() method. I also used the input() function to let the user input his age. If the age is greater than 13, then a message will be displayed that the age is over the expected range. I did this because the age range for the stories is between 0 – 13 years.

After creating the model, I used flask framework to create an API for deploying my model. Besides data collection, this part was the most challenging. I created an app.py file and in the file I imported a function in the other ks2.py file.