Instructions/flow to execute the code :

1. I ‘ve used google Collab to run my code.
2. The dataset I used is Quora Questions from Kaggle. Below is the link to the dataset.

<https://www.kaggle.com/competitions/quora-question-pairs/data>

1. To begin with please import the required libraries(install packages if prompted) mentioned at the top section in the code to ensure smooth execution.
2. Download the dataset of Image folders from the above link into local machine and upload the folders to google drive as i have mounted drive in the code.
3. I have divided train data into train and test set in 80:20 ratio.
4. Then preprocessing of the question texts was done and the TF-IDF vectorization.
5. SGDClassifiers with logistic regression and SVM classifiers was initialized, and the model is trained with train set .Hyperparameter tuning is done to optimize the model.
6. Then the log loss for train and test data were captured for both classifiers.