Readme:

There are 5 python colab and 4 datasets included in the shared folders.

**Dataset folder:**

**Raw\_Dataset\_For\_Tablet.csv:** Web-scraped raw comments data from Amazon Website

**Ground\_Truth\_For\_Topic\_Training\_Data.csv:** Manual-labeled topics for around 20,000 reviews

**Full\_dataset.csv:** Pre-processed dataset

**Topic\_Weight\_df.xlsx:** Topic-related Predictors obtained after feature engineering

**Google Collab files:**

**Natural Language Processing - Topic Modeling:** Data Cleaning, Word tokenization and lemmatization and topic modeling

**Models - For Sentiment Prediction\_RF\_LR\_NB:** Three classification models for predicting sentiment

**Models - For Sentiment Prediction - NN/SVM/Adaboost:** Three classification models for predicting sentiment

**Models-For Importance Prediction and Statistical Test:** Multiple classification models for predicting importance level

**Models - For Topics Prediction:** Multiple classification models for predicting review topics