

The train folder contains 12.5K positive and negative reviews each from IMDB. The train and test dataset contain a disjoint set of movies.

In the reviews no more than 30 reviews are allowed for any movie.

In the labeled (positive and negative) dataset, a negative review has a score ≤ 4 out of 10, and a positive review has a score ≥ 7 out of 10.

Reviews are stored in text files named following the convention `[[id]_[rating].txt]` where `[id]` is a unique id and `[rating]` is the star rating for that review on a 1-10 scale. For example, the file `[train/pos/205_8.txt]` is the text for a positive-labeled test set example with unique id 205 and star rating 8/10 from IMDB.

The unsp directory has 0 for all ratings because the ratings are omitted for this portion of the dataset.

Using the data provided create a Sentiment Analysis model and mention the accuracy while sharing the code back.

If needed, below should be used and not any alternative of these:

- 1) Tensorflow
- 2) Python