Predicting Ratings Based On Drug Reviews

Data

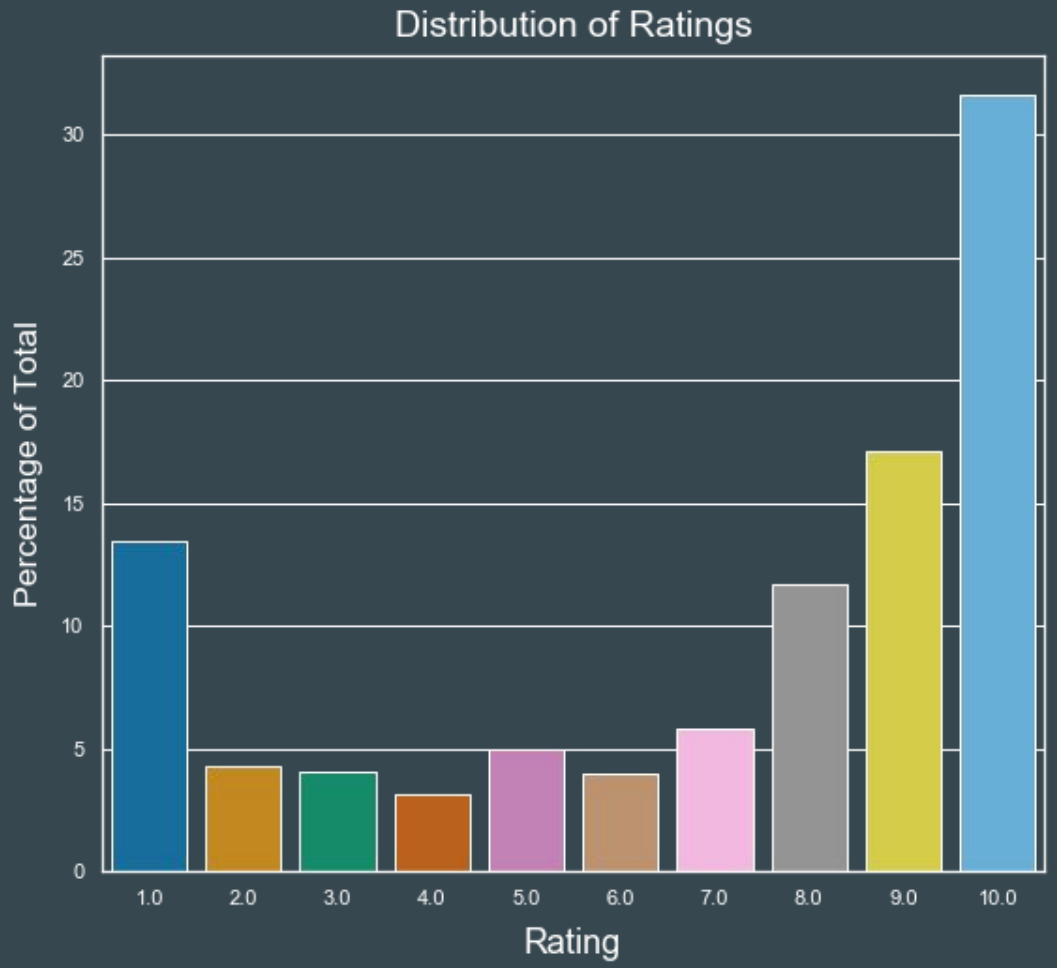
- Data pulled from Kaggle
- ID, drug name, condition, review, rating, date, useful counts

Goal

- Predict rating of drug based on the review

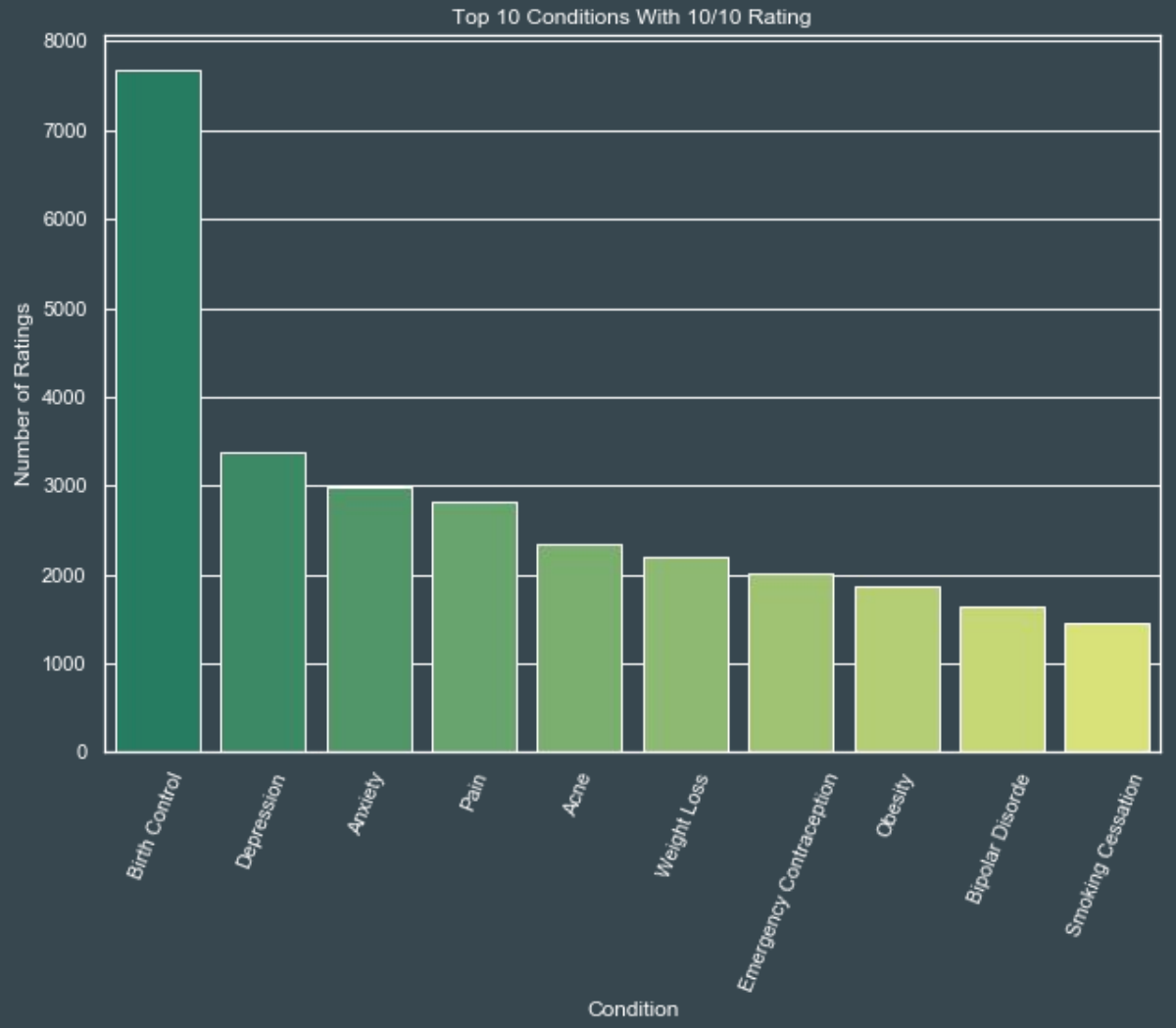
Distribution of Ratings

- Most ratings are 8 or higher
- Denser at the ends



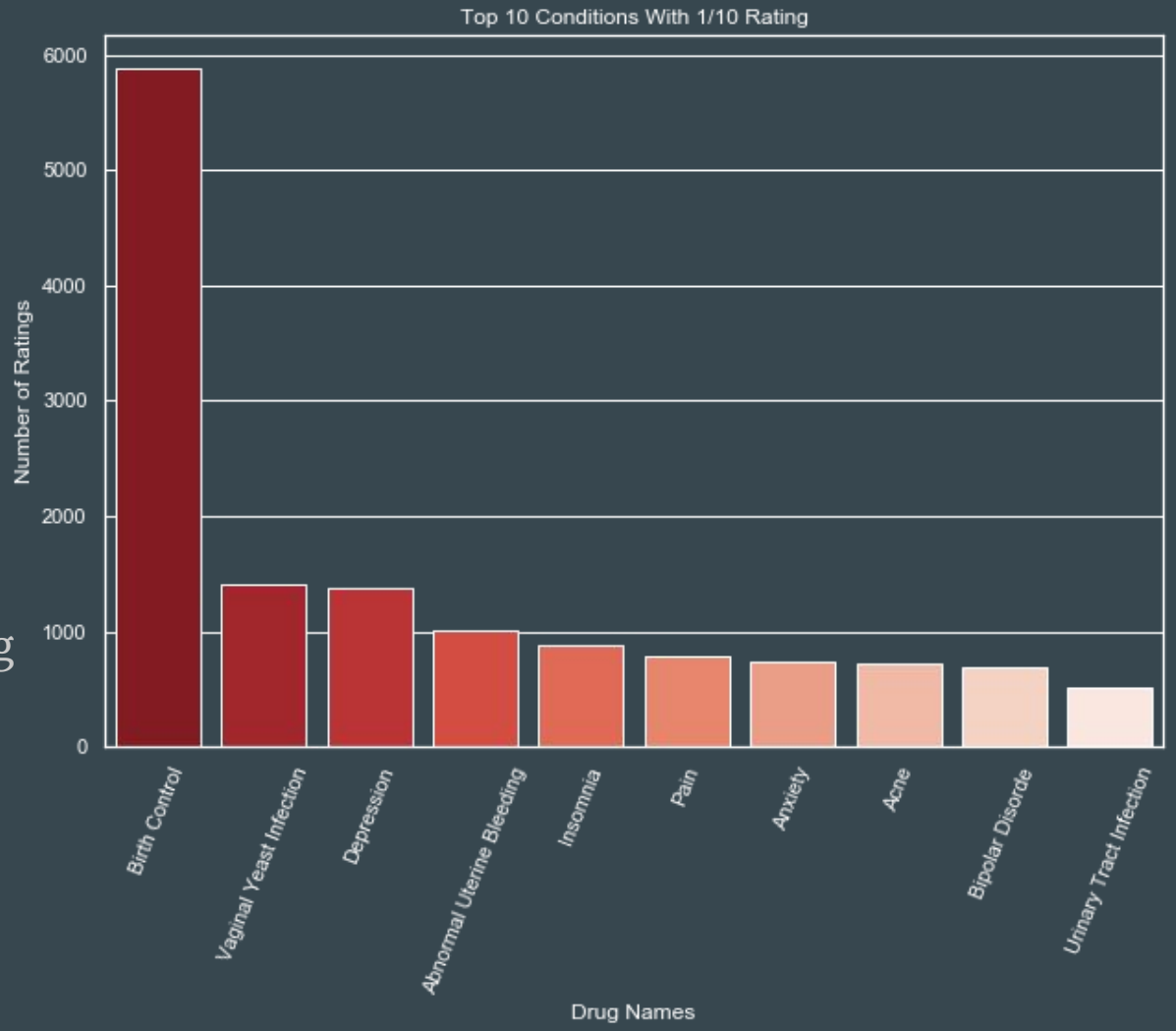
Conditions That Had Best Ratings

- Birth Control
- Depression
- Anxiety
- Pain
- Acne



Conditions That Had Worst Ratings

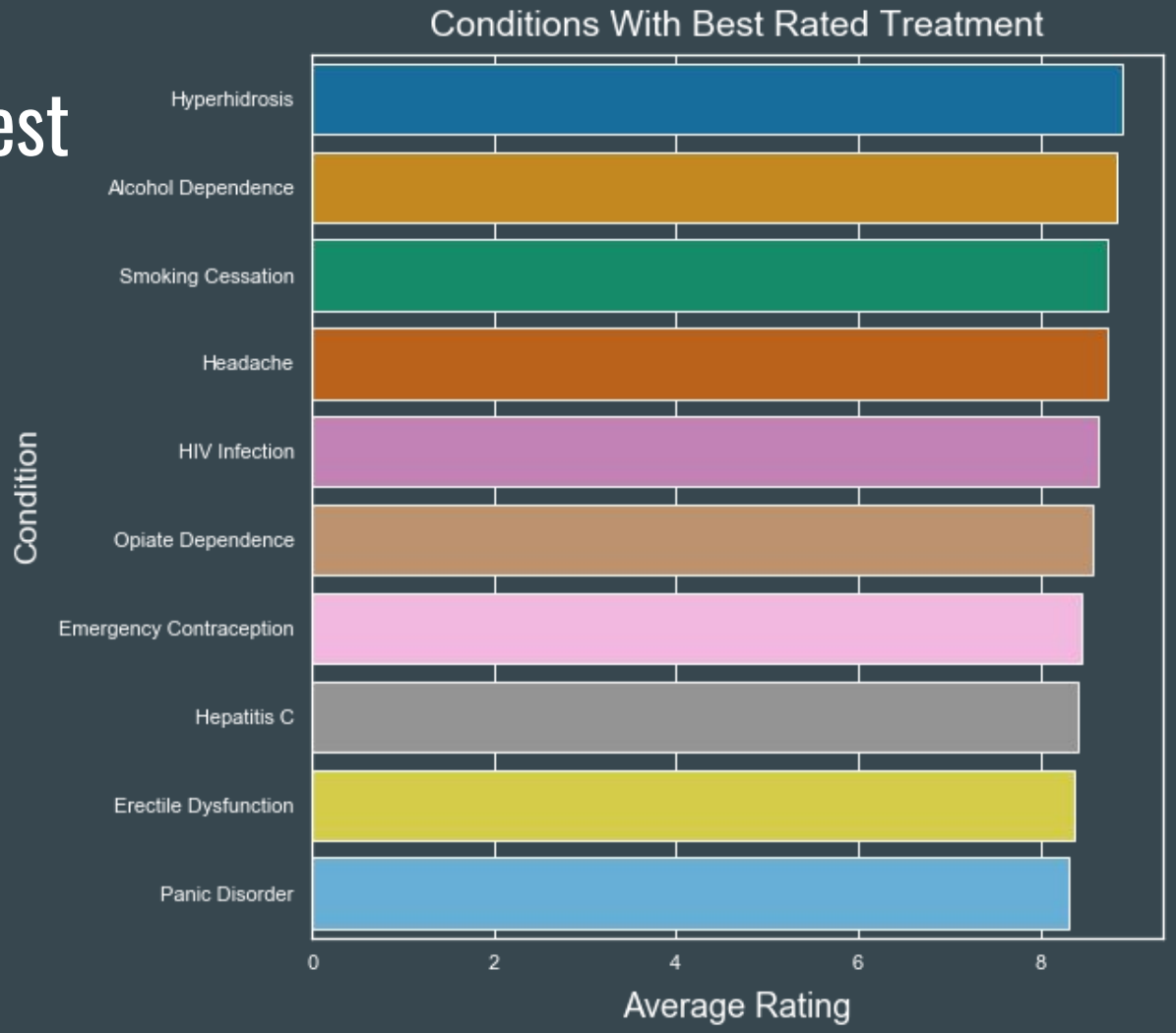
- Birth Control
- Vaginal Yeast Infection
- Depression
- Abnormal Uterine Bleeding



Conditions With Highest Rated Treatments

- Average all ratings of treatments for each condition

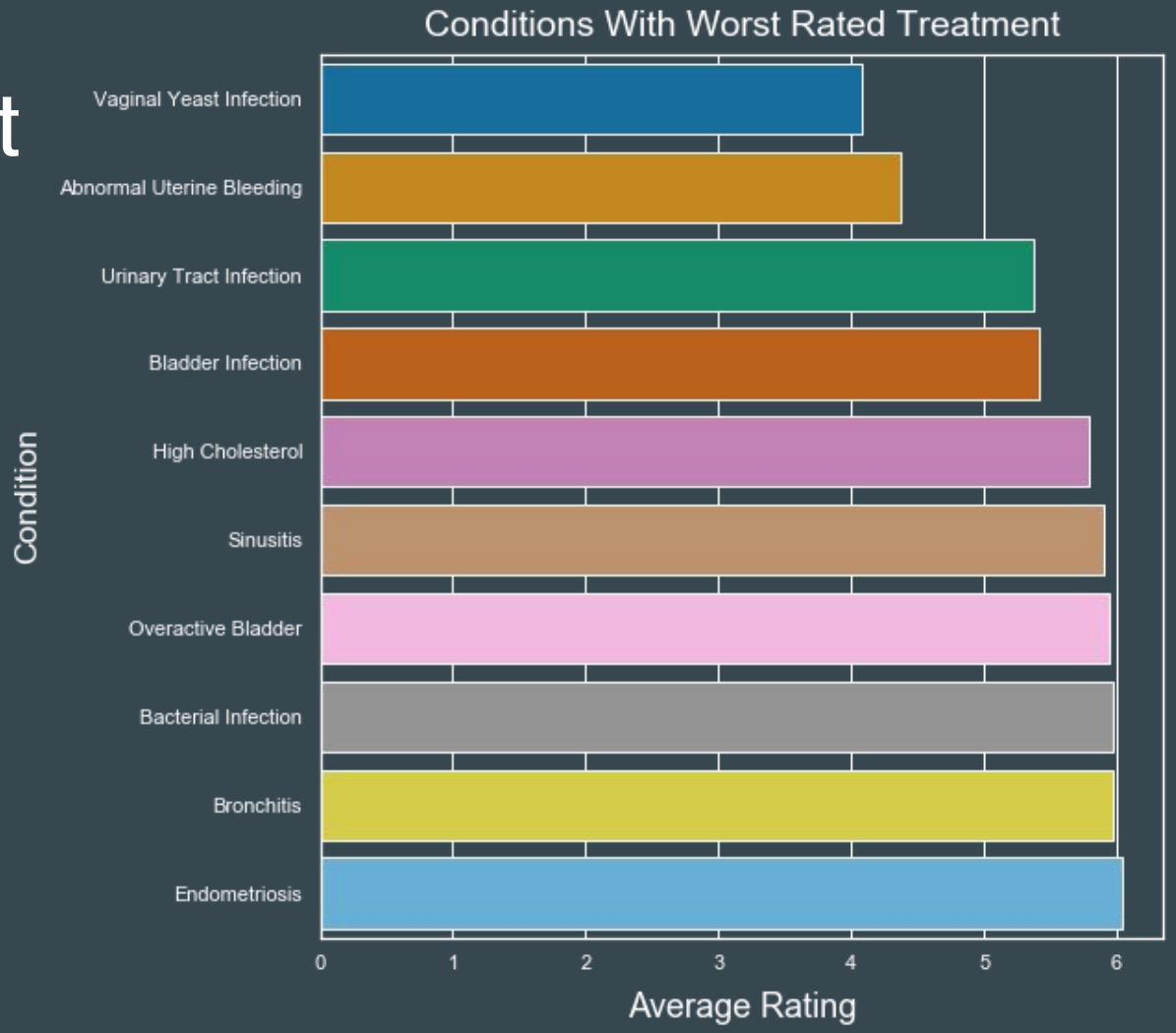
- Hyperhidrosis
- Alcohol Dependence
- Smoking Cessation
- Headache



Conditions With Worst Rated Treatments

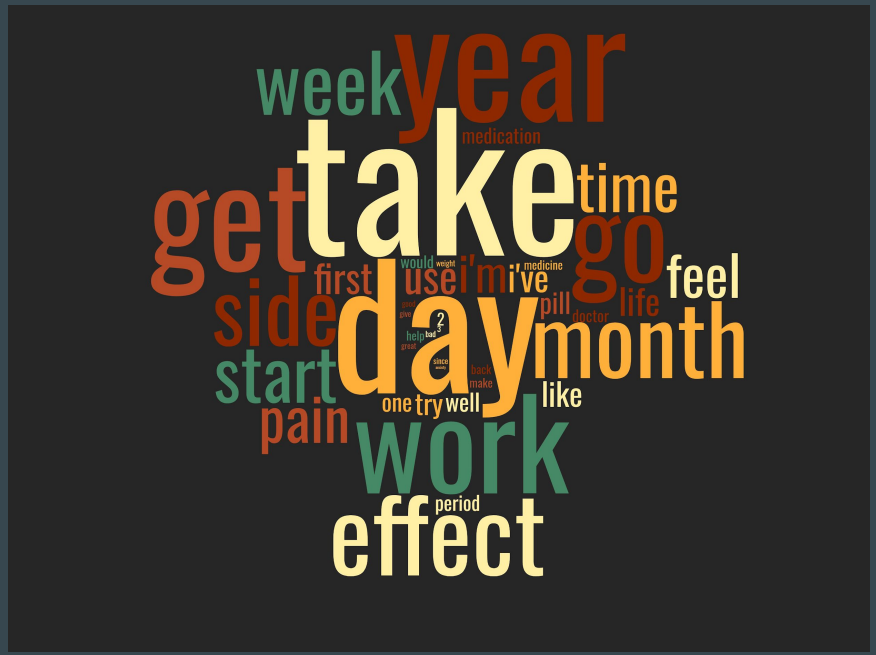
- Average all ratings of treatments for each condition

- Vaginal Yeast Infection
- Abnormal Uterine Bleeding
- Urinary Tract Infection
- Bladder Infection

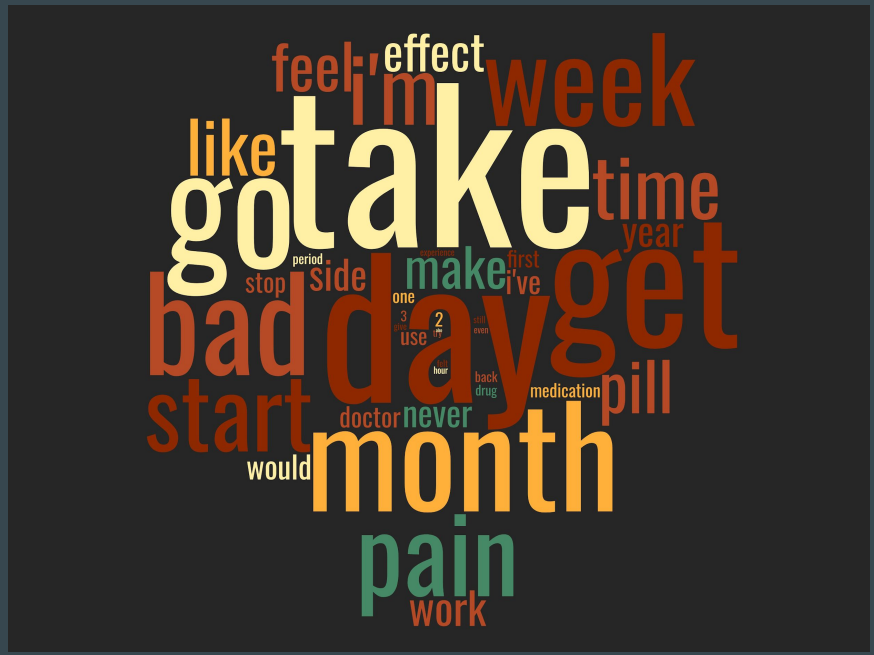


Most Commonly Used Words By Rating

10/10 Rating



1/10 Rating



Model 1

Accuracy	0.49
Loss	1.58

Sequential

Embedding

Bidirectional LSTM

Dense

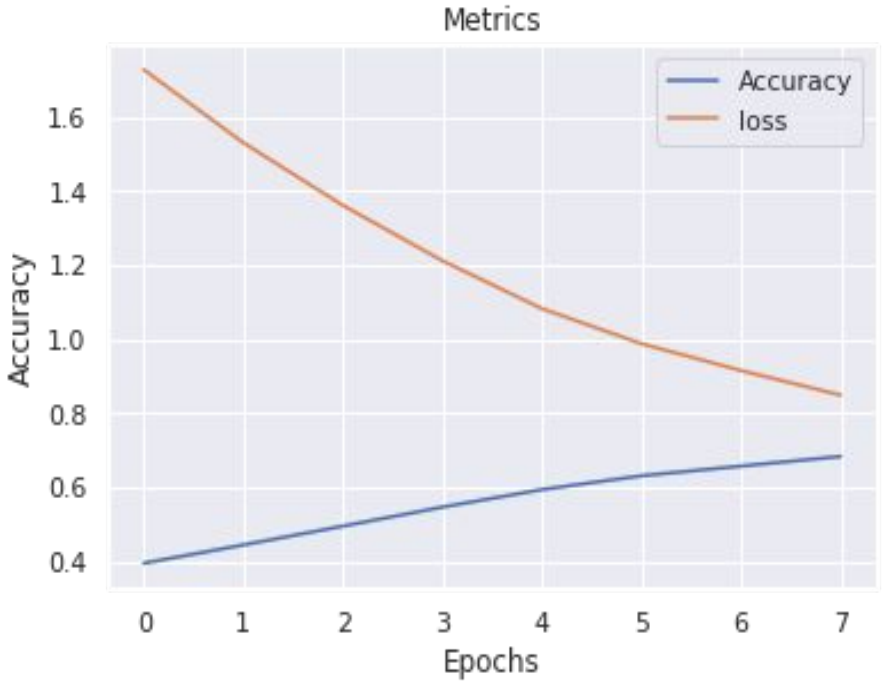
Model 2

- Sequential
- Embedding
- Dropout
- Conv1D
- GlobalMaxPooling1D
- Dense
- Activation

Model 2 results

Accuracy 0.53

Loss 1.64

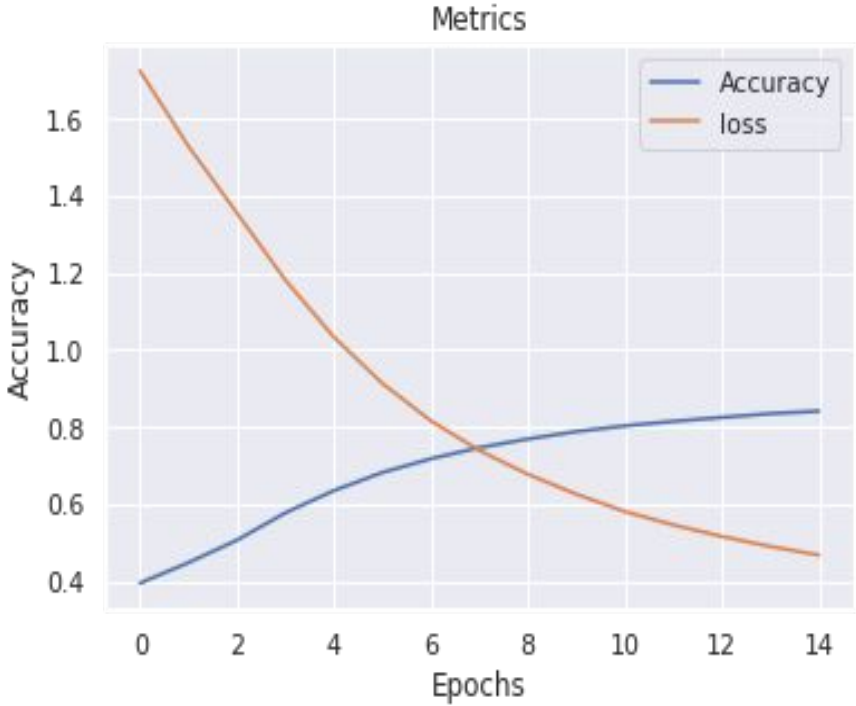


Model 3

- Sequential
- Embedding
- Dropout
- Conv1D
- MaxpoolingID
- LSTM
- Dense
- Activation

Model 3 results

Accuracy	0.62
Loss	1.66



Classification Report

Rating	Precision	Recall	F1-score	Support
1	0.70	0.69	0.69	9534
2	0.53	0.47	0.50	3098
3	0.43	0.52	0.47	2834
4	0.56	0.41	0.47	2184
5	0.52	0.46	0.49	3528
6	0.49	0.45	0.47	2754
7	0.51	0.48	0.49	4133
8	0.54	0.54	0.54	8002
9	0.55	0.64	0.59	11985
10	0.75	0.73	0.74	22525
Accuracy			0.62	70577
Macro Avg	0.56	0.54	0.55	70577
Weighted Avg	0.62	0.62	0.62	70577