**Experimental Setup**

**Dataset**

To test the efficiency of different classifiers we have used movie reviews written in plain text by users along with ratings (from 0 to 5) as labels for several different movies.

**Task**

To predict the rating of a review from the provided text.

**Classifiers used**

1. Stochastic Gradient Descent Classifier
2. Gaussian Naïve Bayes Classifier
3. Decision Tree Classifier
4. K-Nearest Neighbor Classifier

**Parameters**

1. Maximum Frequency – 800
2. Minimum Frequency – 10
3. N-gram Range – (1,4)
4. Maximum Features – 200
5. Training Size – 66.7%

**Results**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Accuracy** | **Precision** | **Recall** | **F1** |
| **Linear Classifier** | 0.790909 | 0.435669 | 0.428303 | 0.429211 |
| **Naive Bayes Classifier** | 0.515152 | 0.40923 | 0.477691 | 0.375851 |
| **Decision Tree Classifier** | 0.751515 | 0.436508 | 0.422961 | 0.428635 |
| **K Nearest Neighbor Classifier** | 0.809091 | 0.552222 | 0.430773 | 0.448657 |