

IMDB Sentiment Analysis Model Comparison Report

Rohith Bharathithasan

Dataset Summary

The IMDB Movie Review dataset contains 50,000 labeled reviews balanced evenly between positive and negative sentiment. After preprocessing, the dataset was split as follows:

Split	Samples
Training Set	40,000
Test Set	10,000
TF-IDF Vocabulary Size	20,000

All models were trained using TF-IDF features generated from the cleaned reviews.

Models Evaluated

Three classical baseline models were trained and compared:

- **Logistic Regression** (max_iter = 1000)
- **Multinomial Naive Bayes**
- **Linear Support Vector Machine (Linear SVM)**

All models used the same data pipeline to ensure fair comparison.

Accuracy Comparison

Model	Accuracy
Logistic Regression	0.8977
Naive Bayes	0.8642
Linear SVM	0.8970

Detailed Metrics

1. Logistic Regression

Accuracy: 0.8977

Class	Precision	Recall	F1-score	Support
Negative (0)	0.91	0.89	0.90	5000
Positive (1)	0.89	0.91	0.90	5000
Accuracy			0.90	
Macro Avg	0.90	0.90	0.90	10000
Weighted Avg	0.90	0.90	0.90	10000

2. Multinomial Naive Bayes

Accuracy: 0.8642

Class	Precision	Recall	F1-score	Support
Negative (0)	0.86	0.87	0.86	5000
Positive (1)	0.87	0.86	0.86	5000
Accuracy			0.86	
Macro Avg	0.86	0.86	0.86	10000
Weighted Avg	0.86	0.86	0.86	10000

3. Linear Support Vector Machine

Accuracy: 0.8970

Class	Precision	Recall	F1-score	Support
Negative (0)	0.90	0.89	0.90	5000
Positive (1)	0.89	0.90	0.90	5000
Accuracy			0.90	
Macro Avg	0.90	0.90	0.90	10000
Weighted Avg	0.90	0.90	0.90	10000

Observations

- Logistic Regression and Linear SVM achieved nearly identical performance (≈ 0.90 accuracy).

- Multinomial Naive Bayes performed slightly worse, which is expected due to its independence assumptions.
- Both LR and SVM handle high-dimensional sparse TF-IDF features effectively.
- The dataset is perfectly balanced, so macro and weighted metrics are identical.

Conclusion

A comparison of the three classical models shows that Logistic Regression and Linear SVM provide the best performance on TF-IDF features for IMDB sentiment classification. Naive Bayes remains a strong baseline due to its simplicity and speed but underperforms on nuanced text data.