

# Sentiment analysis for IMDB movie reviews.

Intiajul A. Shah,  
Sohrab Hossain

Email: md.intiajul.alam.shah@g.bracu.ac.bd, sohrab.hossain@g.bracu.ac.bd

**Abstract**—In this research we analyze reviews from the Internet Movie Database (IMDb) using sentiment analysis, a branch of natural language processing (NLP). State of the art results in deep learning models have shown remarkable works, despite many difficulties in natural language processing.[1] We use different machine learning techniques to categorize emotional representations of IMDb reviews including logistic regression(LR), Random Forest(RF) and Gradient Boosting(GB). To enhance categorization performance, stop words are eliminated and word normalization is applied to the reviews. Additionally, word embeddings and the TF-IDF are used for a number of datasets in this work.[2] A comparative study is conducted on the experimental results obtained for the different models and input features. We hope these findings will contribute to the growing field of work on sentiment analysis, providing new insights into the application of machine learning(ML) techniques in this field.

**Keywords:** Sentiment, Natural Language Processing (NLP), Emotional View, IMDb Reviews.

## I. INTRODUCTION

As the time and technology is growing, ML has emerged as a key field of research since it makes processing massive volumes of data more effective.[3] This approach has demonstrated particular utility in Natural linguistic Processing (NLP), supporting the classification and prediction of linguistic content. Sentiment analysis is a widely used technique in natural language processing (NLP) that may be used at three different levels: Sentence, Document, and Aspect levels. The objective of this research is to use machine learning to learn and find out the sentimental perspectives of Internet Movie reviews.

Challenges in NLP have made sentiment analysis less accurate and efficient, but new developments in deep learning models have shown remarkable results. To categorize the sentimental representations, we use machine learning techniques such as logistic regression(LR), Random Forest(RF) and Gradient Boosting(GB). To enhance classification performance, stop words are eliminated and the word normalization is applied to the reviews. The features for the classification are then represented by a word matrix created from the reviews.[4]

Additionally, word embedding, the TF-IDF are used for a number of datasets in this work. An analysis is carried out to compare the experimental outcomes for various models and input characteristics. By offering new perspectives, this study adds to the expanding corpus of research in this area.

## II. DATASET

The Internet Movie Database (IMDb) has fifty thousand (50K) reviews that have been classified as either positive or

review	sentiment
One of the other reviewers has mentioned that after watching just 1 Oz episode you'll be hooked. The...	positive
A wonderful little production.   The filming technique is very unassuming-very old-time-B...	positive
I thought this was a wonderful way to spend time on a too hot summer weekend, sitting in the air con...	positive

Fig. 1: Compact view of the dataset.

negative. The binary sentiment analysis dataset utilized in this work is called the IMDb Movie Reviews dataset. There is an equal amount of good and negative evaluations in the dataset, ensuring balance in the data. It only contains reviews that are extremely divisive, with a negative review receiving a score of four or less and a good review receiving a score of seven or higher. Reviews for each movie are limited to thirty. There is more unlabeled data in the dataset.[5]

Because of its range and depth, sentiment analysis research has made extensive use of the IMDb Movie Reviews dataset. It offers a wealth of user-generated content, which can offer insightful information on consumer behavior and public opinion. The English-language dataset may be accessed via a variety of data loaders, including pytorch, tensorflow, and others. Because of its CC-BY-SA license, the dataset is openly accessible for use in research projects with proper acknowledgment to the original author.

The actual review texts from the dataset are preprocessed



```
Row_train, Row_test, Snt_train, Snt_test=train_test_split(Row, Snt, train_size=0.7)
```

Fig. 5: Function for split, train, test.

#### IV. RESULT

From our findings it demonstrates the potential of using several machine learning algorithms for sentiment analysis based on the model's ratings of accuracy and other scores.

We found the output from the Random Forest 0.86 F1 score, 0.86 accuracy, 0.86 precision, and 0.85 recall. The results for Gradient boosting showed 0.86 F1 score, 0.86 precision, 0.86 recall, and 0.86 accuracy. With 0.89 accuracy, 0.89 precision, 0.89 recall, and a 0.89 F1 score, Logistic regression fared better.

It's important to remember that the outcomes might change based on the particular dataset and preprocessing techniques applied. To determine which setup and algorithm is ideal for a particular task, it is therefore essential to experiment with many options.

	precision	recall	f1-score	support
0	0.86	0.85	0.85	7395
1	0.85	0.86	0.86	7480
accuracy			0.86	14875
macro avg	0.86	0.86	0.86	14875
weighted avg	0.86	0.86	0.86	14875

Fig. 6: Accuracy of Random Forest.

	precision	recall	f1-score	support
0	0.87	0.82	0.84	7395
1	0.83	0.88	0.86	7480
accuracy			0.85	14875
macro avg	0.85	0.85	0.85	14875
weighted avg	0.85	0.85	0.85	14875

Fig. 7: Accuracy of Gradient Boosting.

	precision	recall	f1-score	support
0	0.90	0.87	0.89	7395
1	0.88	0.90	0.89	7480
accuracy			0.89	14875
macro avg	0.89	0.89	0.89	14875
weighted avg	0.89	0.89	0.89	14875

Fig. 8: Accuracy of Logistic Regression.

These findings highlight the significance of choosing the right machine learning algorithm for sentiment analysis jobs overall. While deep learning models frequently outperform classic machine learning algorithms, they may nevertheless provide good results, indicating the potential of these cutting-edge methods in natural language processing.

#### V. FUTURE WORK

We believe there are many scopes for further investigation and implementing the models. To expand our knowledge of the advantages and disadvantages of the machine learning techniques applied in sentiment analysis may be the main goal of future research. This can entail researching the fundamental

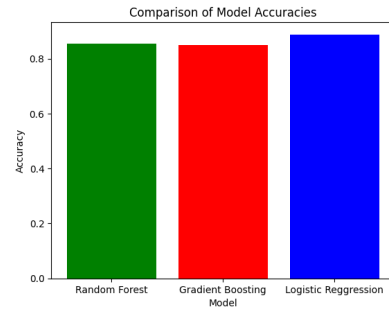


Fig. 9: Bar chart of comparing the results.

principles of existing algorithms, refining their specifications, or creating new iterations. We have also opened the dataset for anyone to use and make their own type of modification.

#### VI. CONCLUSION

We have used this dataset in the study to illustrate sentiment analysis using machine learning techniques such as Random Forest, Gradient boosting and logistic regression. This result highlights the promise of these cutting-edge methods in the field of natural language processing (NLP). Even with the amazing outcomes, the study also identifies a number of areas that need more research. The models' performance might differ based on the particular dataset and preprocessing techniques applied, indicating the need for more thorough testing and validation. Furthermore, compared to standard models, deep learning models performed better but also required more processing power and longer training cycles.

Further study may be done in the future to improve existing models and create more accurate algorithms designed specifically for sentiment analysis jobs. It would also be advantageous to look at how various feature extraction and data preprocessing strategies affect the models' functionality. In order to evaluate these models' effectiveness and practicality, the research might also investigate how these models are applied in actual situations, such as social media monitoring or consumer feedback analysis.[7]

#### REFERENCES

- [1] "Machine Learning based classification for Sentimental analysis of IMDb reviews," *Machine Learning based classification for Sentimental analysis of IMDb reviews*.
- [2] "Papers with Code - Sentiment Analysis Based on Deep Learning: A Comparative Study," 6 2020. [Online]. Available: <https://paperswithcode.com/paper/sentiment-analysis-based-on-deep-learning-a>
- [3] "Papers with Code - Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer," 10 2019. [Online]. Available: <https://paperswithcode.com/paper/exploring-the-limits-of-transfer-learning>
- [4] "Machine Learning based classification for Sentimental analysis of IMDb reviews," *Machine Learning based classification for Sentimental analysis of IMDb reviews*.
- [5] "Papers with Code - IMDb Movie Reviews Dataset." [Online]. Available: <https://paperswithcode.com/dataset/imdb-movie-reviews>
- [6] "IMDB dataset of 50K movie reviews," 3 2019. [Online]. Available: <https://www.kaggle.com/datasets/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews>

[7] "Conclusions." [Online]. Available:  
<https://writingcenter.fas.harvard.edu/conclusions>