ISSN: 2994-6433

Research Article

Journal of Sensor Networks and Data Communications

L-M-6: A Democratic Approach to Movie Rating with AI

Muhammad Wisal[1*], Muhammad Ali[1], Hasanat Lodhi[2], Harras Mansoor[3] and Muhammad Sohaib[4]

[1]*Lowerated*

[2]*AIOapp*

[3]*FinityAlpha*

[4]*PresentaTech*

[*]Corresponding Author
Muhammad Wisal, Lowerated, Pakistan.

Submitted: 2024, Nov 01; **Accepted:** 2024, Dec 05; **Published:** 2024, Dec 20

Citation: Wisal, M., Ali, M., Lodhi, H., Mansoor, H., Sohaib, M. (2024). L-M-6: A Democratic Approach to Movie Rating with AI. *J Sen Net Data Comm, 4*(3), 01-07.

Abstract

This paper presents L-M-6, an innovative algorithm designed to provide statistically accurate and democratically correct movie ratings using AI. Traditional movie rating systems often fail to capture the multifaceted opinions of viewers. In contrast, L-M-6 leverages natural language processing and machine learning to analyze user reviews and extract sentiments across seven key aspects of filmmaking: cinematography, direction, story, unique concept, production design, characters, and emotions.

To enhance the accuracy and relevance of the ratings, a user survey is conducted to rank these aspects based on their perceived importance. The collected data is used to assign weights to each aspect, ensuring that the most valued elements have a greater influence on the overall rating. This weighted sentiment analysis provides a more nuanced and precise rating system.

Moreover, L-M-6 continuously updates scores with new reviews using a rolling mean, ensuring that the ratings remain current and reflective of audience opinions. The algorithm's ability to dynamically adjust and accurately represent diverse viewer sentiments makes it a significant advancement over traditional rating systems. Our results demonstrate that L-M-6 offers a more comprehensive and democratic approach to movie rating, aligning closely with audience preferences and enhancing the overall reliability of movie evaluations.

Keywords: AI, Movie Ratings, Sentiment Analysis, Democratic Algorithm, Statistical Accuracy

1. Introduction

Movie rating systems play a crucial role in guiding audience choices and shaping the success of films. Platforms like IMDb, Rotten Tomatoes, and Metacritic aggregate user reviews and critic scores to provide a single rating that attempts to reflect the overall quality of a movie. However, these traditional systems often fall short in capturing the multifaceted opinions of viewers, reducing complex evaluations to simple average scores. This simplification overlooks the diverse aspects of filmmaking that contribute to a film's impact and success [1,2].

IMDb, one of the most popular movie rating platforms, relies on user ratings that are averaged to produce a single score [3]. While this method provides a general sense of a film's reception, it fails to distinguish between different elements such as cinematography, direction, story, and character development. Similarly, Rotten Tomatoes aggregates critic reviews into a binary "fresh" or "rotten" rating, which can oversimplify nuanced opinions [1]. These traditional systems do not adequately address the varying weights that different viewers might assign to specific aspects of a film [4].

In response to these limitations, we introduce L-M-6, a novel algorithm designed to provide a more comprehensive and accurate movie rating system. By leveraging advanced AI techniques, L-M-6 analyzes user reviews to extract sentiments across seven key aspects of filmmaking: cinematography, direction, story, unique concept, production design, characters, and emotions. This approach allows for a more nuanced understanding of how each aspect contributes to the overall perception of a movie.

To further enhance the accuracy and relevance of our ratings, L-M-6 incorporates a democratic component. We conduct a user survey to rank the importance of each filmmaking aspect. The collected data is used to assign weights to these aspects, reflecting the collective preferences of the audience. This ensures that the most valued elements have a greater influence on the overall rating, providing a more balanced and representative score.

Moreover, L-M-6 dynamically updates scores using a rolling mean as new reviews are added [5,6]. This continuous update mechanism ensures that the ratings remain current and reflective of evolving audience opinions. The ability to adjust dynamically and accurately represent diverse viewer sentiments makes L-M-6 a significant advancement over traditional rating systems.

The remainder of this paper is organized as follows: Section II reviews related work and existing movie rating systems, highlighting their limitations. Section III details the methodology of L-M-6, including the algorithm design, survey for weighting, sentiment quantification, and the rolling mean update process. Section IV presents the results, including survey findings and performance evaluations of the algorithm. Section V discusses the impact of the weights, comparisons with traditional systems, and potential biases. Finally, Section VI concludes the paper and suggests directions for future research.

By providing a more detailed and democratically weighted analysis of user sentiments, L-M-6 aims to offer a superior alternative to traditional movie rating systems, aligning closely with audience preferences and enhancing the overall reliability of movie evaluations.

2. Literature Review

Existing movie rating systems primarily rely on aggregate scores and lack the ability to distinguish between different aspects of filmmaking. This section reviews previous work on sentiment analysis in movie reviews and highlights the limitations of current approaches.

2.1. Traditional Movie Rating Systems

IMDb and Rotten Tomatoes are among the most widely used movie rating platforms. IMDb relies on user ratings averaged to produce a single score, while Rotten Tomatoes aggregates critic reviews into a binary ”fresh” or ”rotten” rating. Both systems have been criticized for oversimplifying complex viewer opinions and not accounting for the diverse aspects of filmmaking [1,2].

2.2. Sentiment Analysis in Movie Reviews

Several studies have explored sentiment analysis in the context of movie reviews. Pang et al. applied machine learning techniques to classify movie reviews as positive or negative [1]. Liu provided a comprehensive overview of sentiment analysis and opinion mining, highlighting its application in various domains including movie reviews [2]. The use of deep learning for sentiment analysis has also been extensively studied, with models such as CNNs and RNNs showing significant improvements in performance [7,8].

2.3. Aspect-Based Sentiment Analysis

Aspect-based sentiment analysis (ABSA) goes beyond simple positive or negative classification by identifying sentiments towards specific aspects of a product or service. Hu and Liu introduced a method for mining and summarizing customer reviews, extracting sentiments towards different product features [9]. More recent work has focused on improving ABSA using deep learning techniques [10,11].

2.4. Weighted Sentiment Analysis

Assigning weights to different aspects based on their importance is a key component of our approach. Prior research has explored various methods for weighting sentiment scores. For instance, Vo and Zhang proposed a model that incorporates aspect importance into sentiment analysis [12]. Another study by Fan et al. introduced a hierarchical attention network to assign weights to different aspects [13].

2.5. Dynamic Rating Systems

The use of rolling means and other dynamic update mechanisms in rating systems has been studied in various contexts. A rolling mean allows for continuous updates to ratings as new data becomes available, ensuring that the ratings remain current and reflective of evolving opinions [5, 6]. This approach is particularly useful in applications where user feedback is continuously received, such as movie rating systems.

3. Methodology

This section details the processes involved in data collection, labeling, model training, and evaluation. It also describes the algorithms used for weighting, sentiment quantification, and dynamically updating review scores.

3.1. Dataset and Human Reviewers

Our dataset consists of 346,000 user reviews collected from IMDb and Rotten Tomatoes, representing a diverse range of viewer opinions. Each review is tagged with sentiment information across seven key aspects of filmmaking: cinematography, direction, story, unique concept, production design, characters, and emotions. The dataset is available publicly on Hugging Face [?].

3.2. Data Collection and Labeling

The data extraction process involved parsing publicly available user reviews, which were subsequently labeled to reflect sentiment scores in relation to each filmmaking aspect. This labeling was performed using a hybrid approach: an automated method using GPT-4 for initial tagging and a manual validation step to ensure accuracy. Manual validation was conducted by all co-authors, each of whom reviewed and confirmed the accuracy of sentiment labels assigned to a subset of reviews.

3.3. Survey-Based Weight Assignment

To capture audience preferences in a democratic manner, we conducted a survey with 890 participants from 16 countries. Respondents were asked to rate the importance of each filmmaking aspect, and the responses were used to calculate weighted scores for each aspect within our model. Table I summarizes the calculated weights.

Aspect	Weight
Cinematography	0.147
Direction	0.145
Story	0.156
Characters	0.145
Production Design	0.129
Unique Concept	0.135
Emotions	0.143

Table 1: Aspect-Based Weights Derived from Survey Responses

These weights reflect the relative importance of each aspect as rated by the survey respondents, allowing the L-M-6 model to prioritize aspects in alignment with audience preferences.

3.4. Aspect Definitions

Each aspect of filmmaking evaluated by L-M-6 is defined as follows:

- **Cinematography:** Refers to the visual aspects of the film, including camera work, lighting, and shot composition.
- **Direction:** Involves the overall execution of the film's vision by the director, including pacing, scene transitions, and actor performances.
- **Story:** Covers the narrative structure, plot development, and originality of the script.
- **Characters:** Evaluates the depth, development, and portrayal of characters within the film.
- **Production Design:** Pertains to the visual environment of the film, including sets, locations, costumes, and props.
- **Unique Concept:** Assesses the originality and creativity of the film's core idea or theme.
- **Emotions:** Measures the emotional impact of the film on the audience, including how well it evokes intended feelings.

3.5. Model Training with Class Weights

To address the class imbalance in our labeled data - where tokens relevant to filmmaking aspects ('1') are less frequent compared to non-relevant tokens ('0') - we implemented class weights in our training process. This technique adjusts the model's sensitivity, improving its ability to identify the less frequent but critical aspect-related tokens, thereby enhancing the overall recall of the model. Class weights are calculated based on the inverse frequency of the classes, giving higher weight to '1's and lower weight to '0's, which helps in reducing the bias toward the more frequent class.

3.6. Model Training and Sentiment Analysis

Our approach leverages two primary models for aspect extraction and sentiment analysis:

- **Aspect Extraction Model:** A model fine-tuned on the Bert Base Uncased dataset, which is trained to extract relevant snippets from reviews for each of the seven aspects (e.g., Cinematography, Direction). When provided with a review, this model identifies and extracts all snippets related to specific aspects, ensuring comprehensive aspect coverage.
- **Sentiment Classification Model:** MoritzLaurer/DeBERTa-v3-large-mnli-fever-anli-ling-wanli, a zero-shot model from Hugging Face, which is used to determine sentiment scores for the extracted snippets [14]. This model is employed in a zeroshot classification setting, where it takes a snippet and is assigned two labels, such as "cinematography positive" and "cinematography negative." The model then produces two scores ranging from 0 to 1, where 0 indicates negativity or irrelevance, and 1 indicates positivity or strong relevance.

The sentiment scores from the zero-shot model are then scaled from 0 to 10, providing a consistent metric for further processing, including averaging and weighted calculations.

3.7. Algorithm Design

The core of our approach involves calculating a weighted mean for each aspect based on the sentiments extracted from the reviews. This section details the algorithms used for these calculations.

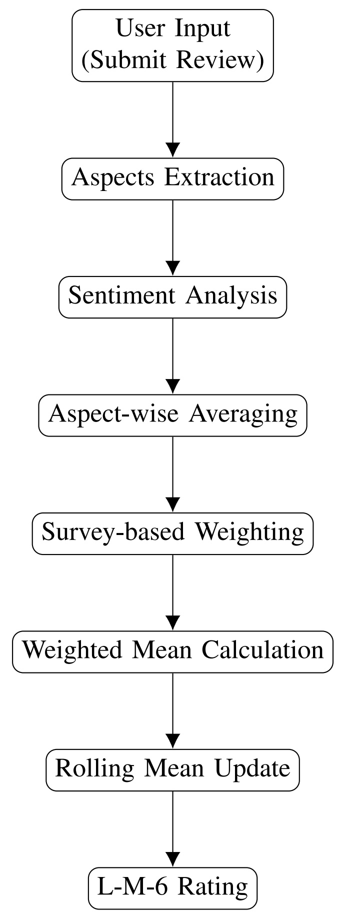

Figure 1: Flow of Processing a Review in the System

1) Weight Assignment Based on Survey: The weight assigned to each aspect is determined by a user survey where participants rank the importance of each aspect from 1 (most important) to 7 (least important). Survey responses are collected and analyzed to determine the weights. For each ranking position i, the number of participants assigning that rank to an aspect is counted. An example of the survey insights is shown in Fig. 1. The weight w_i for an aspect is calculated based on the relative frequency of its rankings, ensuring that the most frequently highly-ranked aspects have a greater influence on the overall rating. The formula for calculating the weight w_i for an aspect is:

$$w_i = \frac{\text{count of rank } i}{\sum_{j=1}^{7} \text{count of rank } j}$$

This approach ensures that aspects ranked as more important by more participants receive higher weights, thereby having a greater influence on the overall rating.

2) Calculating the Average Sentiment for Each Aspect: For each aspect *a* (Cinematography, Direction, Story, Characters, Production Design, Unique Concept, Emotions), we calculate the average sentiment score S_a across all reviews. The formula for the average sentiment score for aspect *a* is:

$$S_a = \frac{1}{n}\sum_{j=1}^{n} s_{a,j}$$

where $s_{a,j}$ is the sentiment score for aspect *a* in review *j*, and *n* is the total number of reviews.

3.7.1. Example Reviews and Scoring

To illustrate, consider the following example reviews:
• **Review 1:** "The cinematography was breathtaking and the story was compelling, but the characters lacked depth."
• **Review 2:** "Great direction and unique concept, but the production design was below average and the emotions felt forced."
• **Review 3:** "The story was innovative and the production design was impressive, though the direction could have been better."

3.7.2. Snippet Extraction and Sentiment Scoring

For each review, the algorithm extracts relevant snippets and calculates sentiment scores for each aspect. For example:

• **Review 1:**
- **Cinematography:** "cinematography was breathtaking"
* Positive Score: 0.9
* Negative Score: 0.1
* Scaled Sentiment Score: $\frac{(0.9-0.1+1)}{2} \times 10 = 9.0$
- **Story:** "story was compelling"
* Positive Score: 0.8
* Negative Score: 0.2
* Scaled Sentiment Score: $\frac{(0.8-0.2+1)}{2} \times 10 = 8.0$
- **Characters:** "characters lacked depth"
* Positive Score: 0.2
* Negative Score: 0.8
* Scaled Sentiment Score: $\frac{(0.2-0.8+1)}{2} \times 10 = 2.0$

• **Review 2:**
- **Direction:** "Great direction"
* Positive Score: 0.85
* Negative Score: 0.15
* Scaled Sentiment Score: $\frac{(0.85-0.15+1)}{2} \times 10 = 8.5$
- **Production Design:** "production design was below average"
* Positive Score: 0.3
* Negative Score: 0.7
* Scaled Sentiment Score: $\frac{(0.3-0.7+1)}{2} \times 10 = 3.0$
- **Emotions:** "emotions felt forced"
* Positive Score: 0.4
* Negative Score: 0.6
* Scaled Sentiment Score: $\frac{(0.4-0.6+1)}{2} \times 10 = 4.0$

• **Review 3:**
- **Story:** "story was innovative"
* Positive Score: 0.9
* Negative Score: 0.1
* Scaled Sentiment Score: $\frac{(0.9-0.1+1)}{2} \times 10 = 9.0$
- **Production Design:** "production design was impressive"
* Positive Score: 0.8
* Negative Score: 0.2
* Scaled Sentiment Score: $\frac{(0.8-0.2+1)}{2} \times 10 = 8.0$
- **Direction:** "direction could have been better"
* Positive Score: 0.4
* Negative Score: 0.6
* Scaled Sentiment Score: $\frac{(0.4-0.6+1)}{2} \times 10 = 4.0$

3.7.3. Average Sentiment Scores

Next, the average sentiment scores Sa for each attribute are calculated:

$$S_{\text{Cinematography}} = \frac{9.0+0.0+0.0}{3} = 3.0$$
$$S_{\text{Direction}} = \frac{0.0+8.5+4.0}{3} = 4.17$$
$$S_{\text{Story}} = \frac{8.0+0.0+9.0}{3} = 5.67$$
$$S_{\text{Characters}} = \frac{2.0+0.0+0.0}{3} = 0.67$$
$$S_{\text{Production Design}} = \frac{0.0+3.0+8.0}{3} = 3.67$$
$$S_{\text{Unique Concept}} = \frac{0.0+0.0+0.0}{3} = 0.0$$
$$S_{\text{Emotions}} = \frac{0.0+4.0+0.0}{3} = 1.33$$

3.7.4 Calculating the Overall Score *O*

Finally, the overall score *O* is calculated using the weighted average of the sentiment scores:

$$\begin{aligned} O = &\left(0.20 \times \frac{9.0+0.0+0.0}{2}\right) + \left(0.15 \times \frac{0.0+8.5+4.0}{2}\right) \\ &+ \left(0.25 \times \frac{8.0+0.0+9.0}{2}\right) + \left(0.10 \times \frac{2.0+0.0+0.0}{2}\right) \\ &+ \left(0.10 \times \frac{0.0+3.0+8.0}{2}\right) + \left(0.10 \times \frac{0.0+0.0+0.0}{2}\right) \\ &+ \left(0.10 \times \frac{0.0+4.0+0.0}{2}\right) \\ &= 0.60 + 0.6255 + 1.4175 + 0.067 + 0.367 + 0.0 + 0.133 \\ &= 3.210 \end{aligned}$$

The overall rating for the movie based on these three reviews is 3.210.

4. Results

The performance of the L-M-6 algorithm is evaluated using several metrics to ensure a comprehensive understanding of its capabilities in capturing nuanced sentiments across multiple aspects of filmmaking. These metrics include Mean Squared Error (MSE), F1-scores, Jaccard Similarity, Precision, and Recall for each aspect.

4.1. Mean Squared Error

The Mean Squared Error (MSE) for the model is 0.0859, indicating a high level of accuracy in predicting sentiment scores. The low MSE reflects the model's effectiveness in closely matching the actual sentiment values derived from the user reviews.

4.2. F1-Scores

The F1-scores for each aspect are as follows:

Aspect	Precision	Recall	F1-score	Accuracy
Cinematography	0.96	0.97	0.96	0.95
Direction	0.93	0.97	0.94	0.95
Story	0.85	0.88	0.85	0.85
Characters	0.89	0.89	0.89	0.90
Production Design	0.95	0.98	0.96	0.96
Unique Concept	0.83	1.00	0.89	1.00
Emotions	0.76	0.87	0.78	0.82

Table 2: F1-Scores for Each Aspect

These F1-scores provide insights into the precision and recall achieved by the model across different cinematic aspects, demonstrating its effectiveness in identifying and classifying aspect-based sentiments accurately.

4.3. Jaccard Similarity

Additionally, we evaluated the model using Jaccard Similarity to assess the overlap between the predicted sentiments and the actual sentiments labeled in the dataset. The average Jaccard Similarity score across all aspects is 0.82, confirming that the model has a strong ability to accurately capture relevant sentiment expressions from the reviews.

4.4. Loss Curve

The training and evaluation loss curves are illustrated in Fig. 2, indicating the convergence of the model during training. The curves demonstrate that the model stabilizes after several epochs, which is indicative of effective learning and optimization.

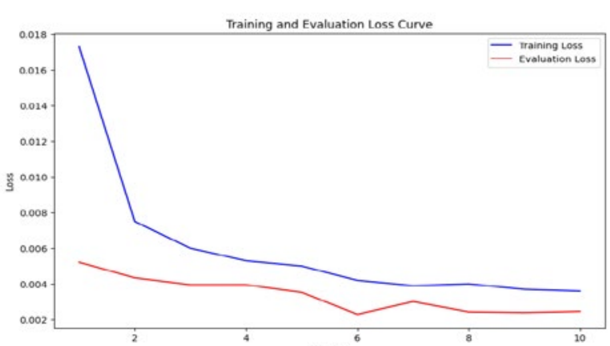

Figure 2: Training and Evaluation Loss Curve

5. Comparison with IMDB and Rotten Tomatoes Rankings

To evaluate the efficacy of the L-M-6 algorithm against traditional rankings, we conducted a comparative analysis on a set of films from IMDb and Rotten Tomatoes. Specifically, we selected the top 50 films by IMDb ranking, gathered corresponding user and critic scores from both IMDb and Rotten Tomatoes, and applied the L-M-6 model to assess the films based on our defined filmmaking aspects.

The dataset encompasses attributes including IMDb ranking, IMDb average rating, and Rotten Tomatoes audience (Popcornmeter) and critic (Tomatometer) scores. By applying the L-M-6 model to these reviews, we derived scores across the seven filmmaking aspects: Cinematography, Direction, Story, Characters, Production Design, Unique Concept, and Emotions. The final L-M-6 score reflects a weighted synthesis of these aspect ratings, aligning with the weights derived from our audience survey (see Table I).

Due to scope constraints, we adjusted the L-M-6 model rankings to reflect the top 10 films by L-M-6. Table ?? shows the comparative scores and rankings for the selected films.

Title	Year	IMDb Rank	IMDb Rating
The Godfather Part II	1974	4	9.0
Gladiator	2000	35	8.5
Terminator 2	1991	28	8.6
Back to the Future	1985	30	8.5
The Prestige	2006	43	8.5
The Lion King	1994	36	8.5
Forrest Gump	1994	11	8.8
The Green Mile	1999	27	8.6
Dune: Part Two	2024	48	8.5
Spirited Away	2001	31	8.6

Table 3: General Movie Information: Title, Year, IMDB Rank, and IMDB Rating

Title	Popcornmeter	Tomatometer	L-M-6 Score
The Godfather Part II	97%	96%	8.44
Gladiator	87%	80%	8.43
Terminator 2	95%	91%	8.41
Back to the Future	95%	93%	8.40
The Prestige	92%	77%	8.40
The Lion King	93%	92%	8.39
Forrest Gump	95%	95%	8.38
The Green Mile	94%	79%	8.35
Dune: Part Two	95%	92%	8.34
Spirited Away	96%	96%	8.33

Table 4: Ratings and Scores: Popcornmeter, Tomatometer, and L-M-6 Score

5.1. Analysis of L-M-6 Performance

The L-M-6 scores provide nuanced insights by emphasizing aspects that reflect a film's visual and emotional appeal as rated by general audiences. As shown, several films achieve high LM-6 scores that correlate well with their IMDb ratings, such as *The Godfather Part II* and *Gladiator*. However, in some cases, such as *The Prestige* and *Dune: Part Two*, the LM-6 model reveals strengths in aspects like Cinematography and Production Design that might not be fully captured by IMDb's single rating metric or Rotten Tomatoes scores.

This comparison highlights L-M-6's ability to yield a holistic view by incorporating aspect-based evaluations, potentially offering an alternative ranking system that resonates with diverse audience preferences.

The results section showcases the robust performance of the L-M-6 algorithm across a variety of evaluation metrics, highlighting its potential to revolutionize sentiment analysis in movie reviews by providing detailed and accurate assessments of different filmmaking aspects.

6. Discussion

The L-M-6 algorithm represents a significant advancement in movie rating systems by incorporating both AI and democratic principles. While our current approach effectively handles various aspects of movie rating, there are opportunities for further enhancement.

6.1. Handling Neutral Values

Our current model is designed to effectively handle neutral sentiments by assigning them a zero value. This approach ensures that neutral sentiments, which neither positively nor negatively impact the viewer's experience, do not influence the overall movie rating. This treatment is crucial for maintaining the integrity of the rating system, as it prevents neutral opinions from skewing the results either positively or negatively.

However, beyond neutral sentiments, our model also addresses irrelevant or unrelated aspects mentioned in reviews. When reviewers mention details or aspects that do not correspond to any of the trained aspect classes, these mentions are classified as 'irrelevant'. In our current framework, such irrelevant mentions are treated similarly to neutral sentiments, assigned a numerical value of zero. This treatment ensures that only relevant, aspect-specific sentiments influence the overall movie rating, thus enhancing the precision of our sentiment analysis model.

6.2. Continuous Improvement of the Algorithm

To maintain and enhance the performance of our algorithm, continuous refinement and adaptation are crucial. Future research could focus on:

- Further enhancing sentiment extraction techniques.
- Expanding the range of filmmaking aspects analyzed.
- Continuously adjusting the weighting mechanism based on new survey data and evolving viewer preferences.
- Evaluating and fine-tuning various sentiment scoring models to improve performance.

6.3. Potential Biases and Their Mitigation

While our approach is designed to minimize biases, ongoing efforts could focus on further diversification of survey samples and applying additional normalization techniques to ensure a fair and inclusive rating system.

7. Conclusion

The L-M-6 algorithm offers a robust and dynamic system for movie rating, accurately reflecting the multifaceted opinions of viewers. Our approach effectively handles neutral values and missing ratings, ensuring a fair and representative rating system.

Future research will focus on exploring alternative methods for handling neutral values, investigating advanced techniques for addressing missing ratings, enhancing sentiment analysis methods, and expanding the range of aspects analyzed. These efforts will help to further optimize the algorithm and ensure it continues to meet the evolving needs of movie audiences.

In summary, L-M-6 represents a significant step forward in movie rating systems, offering a more nuanced and accurate evaluation of films. Ongoing research and development will help to further optimize the algorithm and ensure it continues to meet the evolving needs of movie audiences.

References

1. Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. *arXiv preprint cs/0205070.*
2. Liu, B. (2022). *Sentiment analysis and opinion mining.* Springer Nature.
3. "IMDb," https://www.imdb.com, accessed: 2024-07-29.
4. A Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention Is All You Need.(Nips), 2017. *arXiv preprint arXiv:1706.03762*, 10, S0140525X16001837.
5. P. Jain and M. Varma, "Efficient multilevel multiclass classification using conditional label trees," *The Journal of Machine Learning Research*, vol. 18, no. 1, pp. 5104–5156, 2017.
6. Aggarwal, C. (2015). Data Mining The Text Book.
7. Y. Kim, "Convolutional neural networks for sentence classification," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1746–1751.
8. Tang, D., Qin, B., & Liu, T. (2015, September). Document modeling with gated recurrent neural network for sentiment classification. In *Proceedings of the 2015 conference on empirical methods in natural language processing* (pp. 1422-1432).
9. Hu, M., & Liu, B. (2004, August). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168-177).
10. B. Wang, M. Lu, and Y. Zhang, "Aspect-based sentiment analysis with attention-based lstm," in *Proceedings of the 26th International Conference on Computational Linguistics: Technical Papers,* 2016, pp. 2866–2876.
11. Chen, P., Sun, Z., Bing, L., & Yang, W. (2017, September). Recurrent attention network on memory for aspect sentiment analysis. In *Proceedings of the 2017 conference on empirical methods in natural language processing* (pp. 452-461).

12. Vo, D. T., & Zhang, Y. (2015, June). Target-dependent twitter sentiment classification with rich automatic features. In *Twenty-fourth international joint conference on artificial intelligence*.
13. J. Fan, R. Gao, Y. Zhu, J.-Y. Nie, Y. Zhang, and X. Wang, "Hierarchical attention network for aspect-level sentiment classification," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018*, pp. 353–362.
14. Laurer, M., Van Atteveldt, W., Casas, A., & Welbers, K. (2024). Less annotating, more classifying: Addressing the data scarcity issue of supervised machine learning with deep transfer learning and BERT-NLI. *Political Analysis, 32*(1), 84-100.