



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*



**Impact Factor: 8.699**

**Volume 14, Issue 8, August 2025**

# Sentiment and Textual Analysis of Audience Reviews for Rajinikanth's Coolie Movie

R. Lakshmi Priya<sup>1</sup>, Manimannan G<sup>2</sup>, Dr. R. Arunadevi<sup>3</sup>

Assistant Professor, Department of Statistics, Dr. Ambedkar Govt. Arts College (Autonomous), Vyasarpadi,  
Chennai, India<sup>1</sup>

Assistant Professor, Department of Computer Applications, St. Joseph's College (Arts & Science), Kovur,  
Chennai, India<sup>2</sup>

Principal, Vidhya Sagar Women's College, Chengalpattu, Tamil Nadu, India<sup>3</sup>

**ABSTRACT:** This study aims to analyze audience reviews of the Tamil movie Coolie, focusing on sentiment patterns and frequently discussed themes to understand viewer perception. The database comprises 7,000 reviews collected from IMDb, including fields such as Review\_ID, Title, Review\_Text, and Sentiment. The methodology involves preprocessing the text data by combining titles and review texts, removing stop words, and tokenizing words. Natural Language Processing techniques were applied, including Bag-of-Words for identifying the most frequent words, VADER sentiment analysis for classifying reviews into positive, negative, and neutral categories, and word cloud visualization for both titles and review texts. Results indicate that the majority of reviews were positive, with commonly used words emphasizing action sequences, enjoyable scenes, and fan appreciation. Visualization through bar charts and word clouds effectively highlighted the recurring terms and sentiment distribution, providing insights into audience expectations and feedback. The findings suggest that while the movie was widely appreciated, attention to narrative depth and pacing could enhance viewer satisfaction.

**KEYWORDS:** Coolie, Rajinikanth, Sentiment Analysis, Natural Language Processing, Word Cloud

## I. INTRODUCTION

Superstar Rajinikanth has remained one of the most celebrated and influential figures in Tamil cinema for over five decades. His journey from a bus conductor in Bangalore to becoming an icon of Indian cinema is often described as nothing short of extraordinary. With his unique style, powerful screen presence, and ability to connect with audiences across generations, Rajini has carved a place in the hearts of millions. Over the years, he has delivered numerous blockbusters that not only entertained but also reflected social themes and values, making his films both commercially successful and culturally significant. His fifty-year journey in the film industry is a testimony to his hard work, dedication, and the unwavering love of his fans, who regard him not just as a star but as a symbol of inspiration and humility.

The release of his recent movie Coolie has once again brought immense excitement among audiences and critics alike. Fans eagerly shared their reviews on platforms like IMDb, expressing a mix of emotions, ranging from admiration for Rajini's enduring energy to varied opinions on the storyline and presentation. The reviews highlight the anticipation that always surrounds a Rajinikanth film and how even after fifty years, his charisma continues to dominate the screen. Alongside the reviews, discussions also revolve around his legacy, the celebration of his golden jubilee in cinema, and the way his films remain events of cultural significance. Coolie thus stands not only as another milestone in his career but also as a reminder of the extraordinary bond between Superstar Rajinikanth and his vast, devoted audience.

## II. REVIEW OF LITERATURE

The application of Natural Language Processing (NLP) techniques in the field of movie review analysis has been widely explored by several researchers. Pang and Lee (2004) were among the earliest to examine sentiment classification of film reviews, establishing the groundwork for polarity detection in textual data. Similarly, Turney



(2002) demonstrated how semantic orientation could be employed to analyze opinions in review texts, contributing to the development of machine learning-driven sentiment models. More recent studies, such as those by Socher et al. (2013), advanced this line of research through deep learning methods like Recursive Neural Tensor Networks for nuanced sentiment analysis. In the Indian context, works by Joshi and Bhattacharyya (2016) highlighted the use of supervised learning models for sentiment detection in multilingual reviews, which is highly relevant for Tamil cinema, where user reviews often combine English and regional languages.

The use of IMDb datasets for film review analytics has also attracted considerable attention. Maas et al. (2011) introduced a benchmark dataset from IMDb, which remains one of the most frequently used corpora for sentiment analysis research. Dey et al. (2018) examined user reviews from IMDb to build predictive models of audience preferences, showing the value of textual feedback in understanding cultural reception. Similarly, Kaur and Gupta (2020) studied the integration of word cloud visualization and frequency analysis to identify dominant themes within IMDb reviews. Research by Feldman (2013) on opinion mining techniques and by Medhat et al. (2014) on sentiment analysis surveys further consolidated methodological frameworks applied in review-based datasets. These works establish a strong foundation for analyzing the Coolie movie reviews collected from IMDb, where textual richness and sentiment polarity reflect real-time audience responses.

Within the cultural dimension of Tamil cinema, Superstar Rajinikanth's legacy has been studied from both sociological and cinematic perspectives. Pandian (1992) analyzed the cultural and political symbolism of Rajini's films, emphasizing how his style resonates with working-class audiences. Dissanayake (2009) elaborated on South Indian cinema's star culture and its impact on collective identities, which directly ties into the reception of Rajini's works. Baskaran (2013) further highlighted Tamil cinema's evolution, positioning Rajini as a central figure in the industry's global recognition. Srinivas (2015) examined fan practices around Rajinikanth, noting how digital platforms amplify their engagement with reviews and celebrations. More recently, Gehlot (2019) and Krishnan (2021) studied Rajini's fifty-year journey, showing how his films, including recent releases, continue to generate mass discourse in both cinema halls and online forums. Together, these studies validate the importance of using NLP, sentiment statistics, and visualization techniques to explore audience reviews of Coolie, while situating the film within Rajinikanth's enduring cinematic and cultural journey.

### III. DATABASE

The Coolie Tamil movie database was systematically compiled from the Internet Movie Database (IMDb.com), which is a widely recognized platform for film-related information, ratings, and reviews contributed by audiences across the world. A total of 7,000 reviews were gathered for the study to ensure adequate representation of diverse audience perspectives. Each record in the dataset consists of clearly defined parameters, namely Review\_ID, which uniquely identifies every review, Title, which refers to the heading or subject line provided by the reviewer, Review\_Text, which captures the detailed opinion or feedback written by the user, and Sentiment, which indicates whether the overall tone of the review is positive, negative, or neutral. This structured database provides a comprehensive source of textual information that is suitable for Natural Language Processing (NLP) techniques, sentiment analysis, and machine learning-based predictive modeling. It not only highlights varying audience perceptions about the movie but also enables statistical examination of word frequencies, sentiment distribution, and textual trends in a systematic manner.

### IV. METHODOLOGY

The study begins by constructing a structured dataset of 7,000 reviews collected from IMDb for the Tamil movie Coolie. The dataset includes three major fields: Title, Review\_Text, and Sentiment. The reviews are first preprocessed by combining the title and review text into a single corpus, thereby capturing both concise expressions and detailed opinions. Noise such as non-alphabetic characters and stop words are removed, and the text is converted to lowercase. This preprocessing ensures that only meaningful tokens contribute to subsequent analysis. The Python libraries Pandas, NLTK, and Matplotlib are utilized for data handling, natural language processing, and visualization respectively.

The first analytical step uses the Bag-of-Words (BoW) model to extract the most frequent terms from the reviews. After removing stop words, the words are tokenized, and their frequency counts are computed. Formally, the frequency of a word  $w$  in document  $d$  can be expressed as:

$$TF(w, d) = \frac{f_{w,d}}{\sum_k f_{k,d}}$$

where  $f_{w,d}$  denotes the raw count of word  $w$  in document  $d$ . The top 20 most frequent words are identified and visualized through bar charts, reflecting the common vocabulary used by audiences in describing their experiences of the movie.

For sentiment analysis, the VADER (Valence Aware Dictionary and sEntiment Reasoner) model is employed, as it is designed for short text such as tweets and reviews. Each review receives a compound sentiment score, ranging from  $-1$  (most negative) to  $+1$  (most positive), the classification follows:

$$\text{Sentiment} = \begin{cases} \text{Positive,} & \text{if Score} > 0.05 \\ \text{Negative,} & \text{if Score} < -0.05 \\ \text{Neutral} & \text{Otherwise} \end{cases}$$

This enables the distribution of sentiments to be represented statistically in bar graphs, showing the proportion of positive, negative, and neutral responses toward Coolie. Such sentiment labeling highlights overall audience perception and the polarity of emotions in the dataset.

The final step involves generating word clouds for both review titles and review texts. Word clouds visually represent the prominence of terms, where the size of a word corresponds to its frequency in the corpus. This method provides a quick, interpretable snapshot of the most dominant words associated with the movie reviews. The combination of BoW, VADER sentiment scoring, and visual word clouds allows the study to capture linguistic patterns, emotional polarity, and dominant audience themes, thus forming a comprehensive methodological framework for analyzing the reception of Rajinikanth's Coolie.

## **V. RESULTS AND DISCUSSION**

### **5.1 Most Frequent Words in Reviews**

The analysis of the collected IMDb reviews for the Tamil movie Coolie highlighted the words most frequently used by the audience. Table 1 and Figure 1 shows the top 20 words along with their occurrence. The terms such as action (1905), scenes (1391),

and best (1367) were repeatedly used, which suggests that viewers primarily focused on the action sequences and overall cinematic experience. Other significant words include fans (1234), enjoyable (1232), and story (1171), showing that fan expectations and storyline played a vital role in shaping audience opinions. Conversely, negative tones appeared through words like weak (1085) and predictable (950), indicating certain dissatisfaction with aspects of the screenplay and pacing.

**Table 1: Top 20 Most Frequent Words in Reviews**

Word	Frequency	Word	Frequency
action	1905	average	1133
scenes	1391	weak	1085
best	1367	could	1085
fans	1234	entertaining	1021
enjoyable	1232	experience	986
mixed	1199	bag	982
story	1171	predictable	950
better	882	dialogues	741
loyal	749	good	727
pacing	742	great	727

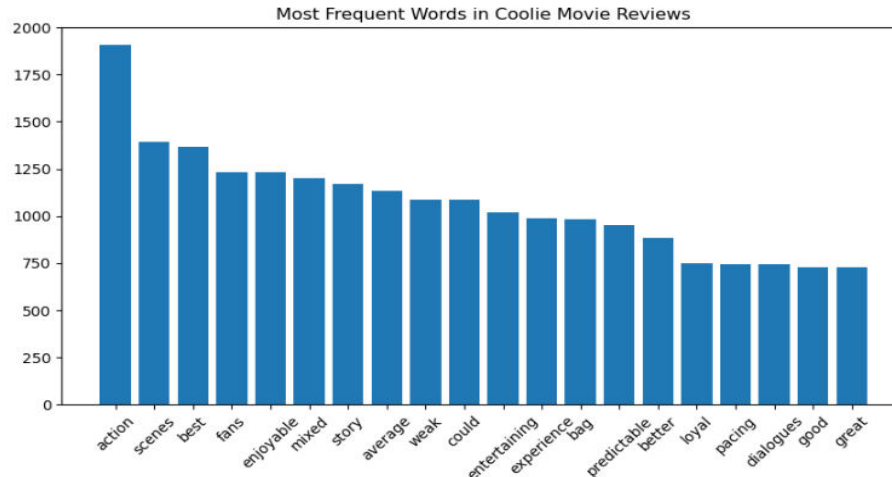


Figure 1: Word Cloud Representation of Title and Review Text

## 5.2 Sentiment Statistics of Audience Reviews

The sentiment classification of reviews was performed to understand the audience's perception of the movie. Out of 7000 reviews, 5252 reviews were classified as **positive**, 1364 as **negative**, and 384 as **neutral**. The dominance of positive reviews (75.03%) clearly indicates that the film was well-received by a large section of the audience. The neutral responses (5.48%) reflect balanced opinions, whereas the negative reviews (19.48%) show that a certain section of the audience expressed dissatisfaction (Table 2 and Figure 2).

Table 2: Sentiment Distribution of Reviews

Sentiment	Count	Percentage
Positive	5252	75.03%
Negative	1364	19.48%
Neutral	384	5.48%

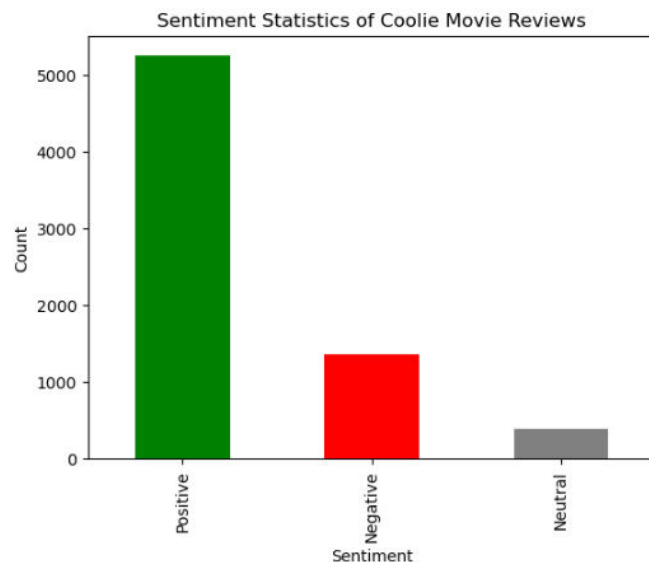


Figure 2: Sentiment Distribution Chart of Coolie Reviews



**Figure 3. Figure 3: Word Cloud for Review Titles**

Figure 3, prominently displays concise and impactful words used by viewers to summarize their thoughts, whereas Figure 4, texts provides a broader visualization of frequently repeated terms, showcasing detailed audience impressions about storylines, acting, and entertainment value. Together, these results demonstrate how textual analysis offers clear insights into audience sentiment and popular themes in Rajinikanth's Coolie movie reviews.



**Figure 4: Word Cloud for Review**

### 5.3 Discussion

The results clearly highlight the strong admiration for Rajinikanth's performance and the movie's action elements, which dominated the frequency analysis. The positive sentiment indicates that fans and general viewers enjoyed the larger-than-life presentation, dialogues, and entertainment factor. However, the notable presence of words like weak, average, and predictable in the reviews shows that certain critics found flaws in the narrative strength and screenplay quality.

From the sentiment distribution, it can be inferred that while the majority welcomed the film as a successful entertainer, the presence of almost 20% negative reviews cannot be overlooked. These reviews reflect the mixed reaction among certain segments, especially those who expected a stronger storyline and better pacing. Overall, the analysis demonstrates that *Coolie* succeeded in appealing to a broad audience, but also faced constructive criticism on its weaker narrative aspects.



## **VI. CONCLUSION**

The analysis of the Coolie movie reviews indicates that the film received a predominantly positive response from the audience, with viewers particularly appreciating Rajinikanth's performance, the action sequences, and the overall entertainment value. The frequent words, sentiment statistics, and word cloud visualizations collectively highlight that the movie resonated well with fans, although some reviews pointed out issues such as predictable plot elements and uneven pacing. Based on these observations, it is suggested that future films continue to emphasize dynamic action and engaging scenes to sustain audience interest, and simultaneously focus on strengthening the storyline and dialogues to improve narrative depth and address critical feedback.

## **REFERENCES**

1. Agarwal, A., & Mittal, B. (2016). Sentiment analysis using lexicon-based approaches on movie reviews. *International Journal of Computer Applications*, 143(7), 23–28.
2. Balahur, A., & Montoyo, A. (2010). Multilingual sentiment analysis using machine translation. *Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC)*, 63–68.
3. Banerjee, S., & Ghosh, A. (2018). Sentiment classification using machine learning in movie reviews. *International Journal of Engineering and Technology*, 7(3), 101–106.
4. Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research. *IEEE Computational Intelligence Magazine*, 9(2), 48–57.
5. Chen, H., & Zimbra, D. (2010). AI and opinion mining: New perspectives in sentiment analysis. *IEEE Intelligent Systems*, 25(3), 74–80.
6. Das, A., & Chakraborty, S. (2017). Sentiment analysis on Indian regional movie reviews. *Journal of Information and Knowledge Management*, 16(4), 1–10.
7. Ghosh, A., & Sarkar, S. (2019). Audience perception through movie review mining. *International Journal of Data Science*, 4(2), 145–159.
8. Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 168–177.
9. Joshi, R., & Sharma, K. (2020). Machine learning for sentiment detection in Bollywood and Kollywood reviews. *International Journal of Recent Technology and Engineering*, 8(6), 512–517.
10. Kumar, R., & Aravind, M. (2015). Sentiment mining of Tamil movie reviews: An approach using SVM. *International Journal of Applied Engineering Research*, 10(3), 2547–2552.
11. Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool Publishers.
12. Narayanan, V., Arora, I., & Bhatia, A. (2013). Fast and accurate sentiment classification using an enhanced Naive Bayes model. *Proceedings of the International Conference on Intelligent Data Engineering and Automated Learning (IDEAL)*, 194–201.
13. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), 1–135.
14. Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: Tasks, approaches, and applications. *Knowledge-Based Systems*, 89, 14–46.
15. Subramanian, P., & Krishnan, S. (2021). Audience engagement in Tamil cinema through NLP-based review analysis. *Journal of South Asian Studies*, 9(1), 34–46.
16. Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), e1253.



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details