An effective technique in the film industry, sentiment analysis, a subfield of natural language processing (NLP), provides insightful data and predictive capabilities that go beyond traditional cinema. For a variety of stakeholders, including filmmakers, studios, marketers, and researchers, the ability to identify and quantify the emotions and feelings contained in film scripts is critical. (Sánchez-Rada and Iglesias 2019). The promise of sentiment analysis in film scripts lies in its impact on the overall success of the film as well as its ability to provide deeper knowledge of a film's content.

To understand the complex relationship between the emotions portrayed in screenplays and the box office performance of the films depicted in them, this literature review examines the fascinating field of sentiment analysis in screenplays. Additionally, related sentiment analysis is examined in other areas related to films and their resulting box office success. From lexicon-based strategies to sophisticated deep learning techniques, it explores a range of sentiment analysis techniques, each offering a unique perspective on the emotional terrain of cinematic narratives.

This literature review's main research question is: How can sentiment analysis applied to film scripts help us understand the elements that affect a movie's box office success? This is an important question in the contemporary film industry, where producers, marketers, and investors place a high value on being able to predict a movie's financial performance. This research question has significance that extends beyond the scientific domain. The practical decision-making processes in the film industry may be significantly impacted by its conclusions.

Examining the emotional core of scripts and their connection to box office revenue can help film studios and filmmakers make better decisions regarding audience engagement, marketing strategies, and content. The results of this research can also be used to decide how to most effectively allocate money and resources in the highly competitive and unstable film business.

The layout of this literature review allows for a thorough investigation of the topic in question. It opens with a look at sentiment analysis methods used in movies and the film business and outlines some possible uses for them. The methodologies utilized in various research are then reviewed, with the benefits and drawbacks of each method underlined.

The following sections discuss various sentiment analysis approaches and their performance and highlight the capabilities and limitations of these techniques. The aim is to highlight the urgent need for further research and exploration in this area as it has the potential to revolutionize the film industry's decision-making processes.

**Sentimental Analysis Three key approaches**

Sentimental analysis has three key approaches the document-level, aspect-based, and sentence-based level.

1. **Document-level sentimental analysis:** can be approached on two main methods; supervised or unsupervised learning (Feldman 2013), this approach assumes that the document contains an opinion on one main object expressed by the author of the document. Entire documents are classified as having an overall positive or negative polarity (Pang et al. 2002, Pang & Lee 2005).
2. **Sentence-based level sentimental analysis:** to understand more opinions about the entities of documents, the level of individual documents is weighed in depth (Feldman 2013). RIllof & Wiebe (2003) have used a sentence-level approach in their works. A document can contain different opinions about the same entities.
3. **Aspect-based level sentimental analysis:** Most often the Sentence-level based sentimental analysis approach fails to detect opinions about an entity, and its aspects. This is where aspect-based sentiment analysis comes in. In order to address this, researchers created the aspect-based approach, in which an opinion is made up of goals and the emotions that go along with them. Cambria et. al (2013) in order to further clarify provided this example "The new iPhone's screen size is amazing, but its battery life is limited. In this case, the screen size and battery life of the same entity would be the two aspects (opinion targets) that would be evaluated. Positive opinions surround the iPhone screen size, while negative opinions surround the battery life. Works like Hu and Liu (2004) and Di Fabbrizio et. al (2011) also use this strategy. Pattern and NLTK both employ a document-level strategy for sentiment analysis.

Sentimental Analysis techniques related to the review are lexicon-based, emotion analysis, opinions summarization, hybrid approaches, Aspect-Based, Rule-Based, Machine Learning-Based, and Deep Learning based, etc.

**Word Embeddings, Opinions, Machine Learning**

Sentiment analysis using emotions and word embeddings (Giatsoglou 2017) pointed out that sentiment analysis is important for companies, marketers, and researchers. The aim was to identify subjective information from text sources. Different approaches could be used, such as lexicon-based and word embedding-based. The proposed sentiment detection framework has a high accuracy and efficiency. The methodology is generic and can be applied to different languages. The hybrid approach combines Word2Vec and a lexicon for sentiment prediction. The results revealed barriers to transitioning to stricter dietary habits. The proposed approach is evaluated based on user reviews in Greek and English the results showed the importance of reviews. Emotion analysis is not discussed in great detail. The integration of emotion analysis and whether it could enable a more insightful analysis of film scripts is at the core of our research.

Weighing the intensity of opinions is another method through which sentimental analysis can be used as a method (Thet et. al. 2010), this is an aspect-based sentimental analysis. The proposed method performed sentiment analysis at the clause level, enabling individual analysis of sentiments for different elements. The algorithm analyzed the grammatical interdependence of words in a phrase, split them into independent sentences, and calculated the contextual emotion value for each sentence based on a specific aspect. The experimental results showed that the proposed method is successful for aspect-based sentiment analysis of short texts such as message board posts. The clause-level sentiment categorization accuracy for the entire movie, director, cast, story, scene, and music components is 75%, 86%, 83%, 80%, 90%, and 81%, respectively. This method can be used to summarize the opinions of multiple evaluation aspects. Unlike previous approaches that focused on extracting feature-opinion pairings, the proposed approach provided the sentiment score of a phrase or sentence and allowed comparison of the sentiment strength of the clause or sentence with others. For example, the sentiment scores obtained by the proposed approach can be used to highlight the most positive and negative clauses or sentences associated with certain characteristics. In the future, we plan to explore sentiment aggregation across multiple genres (e.g. forums, user reviews, critic reviews, and Twitter) to allow readers to search films based on system-computed sentiment ratings, aspects, and user-provided sentiment ratings and browse the preferences of an individual reader. e.g. (a sentimental summary from a movie script or film review). A further study done in our paper would use the sentimental score, to check the association with the film’s box office success.

(Onalaja, et. al. 2021) This study examined classification models for determining sentiment in movie reviews and discovered that assigning higher driving factors to specific characteristics and genres results in higher accuracy in sentiment prediction models. The study focused on using weighted driving factors to identify movie aspects and their impact on sentiment classification. After conducting numerous iterations and cross-validation, the study found that assigning higher driving factors to certain aspects in different genres resulted in higher accuracy of the models. The study suggested that these driving factor-assisted models can provide insights into which aspects drive sentiment in any unseen test dataset. The emergence of movie streaming platforms has allowed users to share their opinions through various formats, and businesses use opinion mining to understand customer preferences. Aspect extraction and Latent Dirichlet Allocation (an unsupervised learning method) are techniques used for sentiment analysis. The research aimed to build classification models to determine viewers' sentiments. In terms of average accuracy and standard deviation across its CV models, SVM performed the best when employing TF-IDF vectorization. The study went on to use CountVectorizer for NB and TF-IDF for LR and SVM models. Evaluating user-generated reviews is crucial for understanding business performance, and the insights from aspect and genre driving factor assisted models can greatly assist the movie industry and other businesses. The study also highlights the importance of granular sentiment analysis and the predictive capability of ABSA models. Despite the limitations of creating a good lexicon list for aspect-related words, the addition of the aspect and genre-driving factor increased the accuracy of the aspect-based models and reasonably improved the predictions of review sentiment.

The practice of extracting subjective information from textual data is known as opinion mining, sometimes known as sentiment analysis. Opinion mining is the process of finding and extracting people's opinions, which might be positive, neutral, or negative. Opinion mining, also known as sentiment analysis, is used to assess people's feelings and attitudes in movie reviews.

Different techniques of sentiment analysis were compared on two datasets and this highlighted the importance of using sentiment analysis to predict a movie’s performance based on the average sentiment of all the reviews. (Danyal et. al. 2023) This was based on the performance of Linear support Vector machine and Multinomial Naïve Bayes. Linear Support Vector Machines achieved a maximum accuracy of 89.48% on the IMDB dataset, while Multinomial Naive Bayes had a maximum accuracy of 71.04% on the Sentiment Polarity Dataset. Logistic Regression also improved its accuracy to 89.96%. Linear Support Vector Machines and Logistic Regression performed better on high-dimensional textual data, while Multinomial Naive Bayes performed better on larger datasets. According to the hyperparameter optimization results of Sentiment Polarity Dataset version 2, Multinomial Naive Bayes dominates with a maximum accuracy of 71.04%, but the passive-aggressive classifier once again outperformed linear SVM and logistic regression. K-Nearest Neighbor and Decision Tree outperformed other models due to their ability to identify a larger proportion of positive or negative sentiment instances.

The study suggested future work on the comprehensive analysis of sentiment analysis techniques and exploring deep learning models for movie reviews. The team proposed that Linear Support Vector Machines and Logistic Regression performed better on high-dimensional textual data, and Multinomial Naïve Bayes performed better on larger datasets. Naïve Bayes, Passive Aggressive Classifiers, Logistic Regression, and LSVM achieved the highest accuracy and can serve as benchmark models for future research in sentiment analysis.

Quantifying the factors in the British film Industry that make action movies a success, the research by (Xu 2023) opined to consider four factors; director, plot, actors, and music. Regarding the impact of social media and film reviews on the film industry, it has been stated that social media has become a communication platform between film companies and audiences. In contrast, film reviews have served as a method of communication through direct input. Two sentimental analysis model was applied, the VADER (Valence Aware Dictionary for Sentiment Reasoning) a lexicon-based model that evaluates by sentiment scoring of each word where the overall ratings and each aspect or opinion can be evaluated (Hutto & Gilbert, 2014), data was collected based on the qualitative research approach, which is the way to get a content analysis project (Leavy 2014), data was scraped using ParseHub from the IMDB website, social media. Methods such as Text Processing, using NLTK (Natural Language Toolkit), tokenization, normalization, lemmatization, feature extraction, and feature selection, the modeling was done using VADER, and SVM; and evaluated by accuracy, f1-score. The limitation of this study is the lexicon-based model cares about a single world and ignores the context, and for SVM, the count of the features is too large. The testing score was 93.75% for the SVM model. The VADER model had a strong robustness and flexibility the average accuracy was 76.2%. The VADER model answered the research question, and SVM is the standard against which the main model of the project is compared.

Another study looked at sentiment analysis of film scripts and reviews. This study examines the relationship between film scripts and their reviews in predicting film ratings. The proposed model combines sentiment analysis, NLP, and machine learning techniques. Combining vector semantics and sentiment lexicons leads to different results. The proposed model achieves high precision but slightly lower accuracy compared to the best-performing model. Emotion and sentiment analysis are important for predicting movie ratings. VADER and NRC are selected for sentiment and emotion analysis. VADER has been widely used in other research (Newman and Joyner 2018, Park and Seo 2018). Experiments are evaluated using accuracy, precision, recall, and F1 metrics. The NRC lexicon outperforms VADER in considering multiple emotions. However, “The proposed model performed well but was limited by the lack of negation handling and the use of only excerpts of reviews. Obtaining the full text and addressing negation handling and polarity shift could improve accuracy. Different preprocessing methods may be needed for these enhancements,” concede the researchers. They suggest that this paper explores the impact of incorporating emotion and sentiment analysis in predicting movie ratings. The proposed model shows impressive performance, but limitations include the lack of negation handling and the use of only excerpts of reviews. Obtaining the full set of reports and a larger corpus for analysis could improve the accuracy of the model. Improvements such as negation handling and polarity shift testing would require different preprocessing methods. (Frangidis et. al. 2020).

**Ensemble Models**

Even though one of the most common, straightforward, and beneficial jobs in natural language processing is sentiment analysis. However, it is unclear whether these provide a significant advantage over straightforward bag-of-words and bag-of-Ngram techniques (Pang & Lee, 2008; Wang & Manning, 2012). Advanced machine learning techniques can be used, such as recurrent neural networks and their variations (Mikolov et al, 2010; Socher et al, 2011). They attained a new state-of-the-art performance of 92.57% while interpolating the results of recurrent neural networks (RNNs), set vectors, and Naive Bayes using the Support Vector Machine (NB-SVM), as opposed to the 91.22% reported by (Wang & Manning, 2012). The research proposed a very simple but powerful ensemble system for sentiment analysis by combining three rather complementary and conceptually different basic models: one based on a generative approach, one based on continuous representations of sentences, and one based on clever reweighting of the tf- idf bag-of-word representation of the document. Each of these models contributes to the success of the overall system and achieves new, state-of-the-art performance on the demanding IMDB movie review dataset.

The impact of various ensemble and machine learning models for sentiment analysis on movie reviews is compared in this work. Additionally, in order to select the most crucial features and reduce the size of the input data, feature extraction techniques are used, improving overall accuracy. Several classification models were used, including Logistic Regression, Support Vector Classifier, and Random Forest, in addition to boosting and stacking ensemble models. The gathered data demonstrated that ensemble models are more accurate than other models for sentimental analysis of movie reviews. The test data were obtained from the NLTK data set for sentiment analysis of movie reviews. (Sangeetha et. al. 2023).

(Mesnil, et. al 2014) A new ensemble system for sentiment analysis combines three baseline models: generative approach, continuous sentence representations, and tf-idf bag-of-word representation. The system achieves a new state-of-the-art performance of 92.57% when interpolating scores from Recurrent Neural Networks, sentence vectors, and Naive Bayes Support Vector Machine. The system's performance is improved by removing the generative model at a time.

**ML Based analysis (Regression)**

Dellarocas et. al (2007) study explored the use of online review metrics in predicting the success of movies. The results showed that online review metrics can improve forecasting accuracy, especially for sleeper movies. The study also demonstrates that online review metrics can be used as a proxy for consumer word-of-mouth and early box office sales. The researchers collected data from Yahoo! Movies, BoxOfficeMojo, and the Hollywood Reporter to construct their models. The study includes regression equations for predicting movie-specific parameters and uses a cross-validation procedure to test the forecasting accuracy of the models. The results suggest that online review metrics have explanatory power and can be used to estimate competitors' sales. The study concluded by suggesting further research in other industries such as video games and music. 34893 individual users were involved in the analysis. The conclusions may support prior research in this area: “The volume, valence, and dispersion of online movie reviews are statistically significant in predicting future sales in the entertainment industry,” Dellarocas said. (Dellarocas et. al 2007)recommend that this text introduce forecasting models and estimation techniques, present the results of fitting them to a data set, compare their forecasting accuracy to older models, discuss managerial implications, and suggest potential avenues for future research.

A study by (Kim et. al. 2021) examined the factors that contribute to film success by analyzing customer comments and film characteristics, finding significant differences in screening days between positive and neutral groups, negative and neutral groups, and different genres and ratings. The researchers used data from films released in South Korea. Film consumption in South Korea is high and the country has the highest number of cinema admissions per capita in the world (Moon et al. 2015). Revenue exceeded 1.2 trillion South Korean won in 2022, indicating that Korea's film market was sufficiently large and robust. The research used the method of survival analysis in four different ways, another method was based on the Cox regression model, which suggests that the probability of an event occurring depends on the characteristics of the subject (Cox 1972). Finally, the time-dependent Cox model was applied. This model considered the percentage of factors in the analysis process that have an impact on survival and the likelihood that they will do so over time. Related to our work is the use of the regression model for film scripts that are to be analyzed based on box office results. As a result of their research, they identified the main factors affecting the number of screening days and put forward five hypotheses. Two of the five hypotheses they evaluated were fully supported, while three others received only limited support. Previous research determined the impact of Internet reviews on screening days and focused primarily on comparing positive and negative comments about the film. To assess the impact of each factor on screening days, they used a method not previously used in such studies: survival analysis. This analysis opens new opportunities for in-depth research to make strategic judgments about film production and funding.

In relation to our study that uses regression analysis, multiple regression analysis and discriminant analysis are used to predict the success of a movie. Factors such as genre, timing of release, and star power influence the success of a movie. Location-based promotions and positive reviews from critics can increase a movie's revenue. Mohan Raj and Aditya (2017) focused the scope of their research on the Indian Film Industry, which showed that movie news references and sentiment measures are correlated with movie grosses. The goal they achieved was to determine the features that influenced the frequency of movies watched, predict the success of a movie with regression and discriminant analysis that specialized in qualitative and quantitative methods, and use sentimental analysis on two movies, through the use of word cloud. According to the research, there is a strong association between user reviews and critic reviews. This critical review can assist in boosting the first week's revenue. The model was created by employing linear regression analysis to find the elements that influenced the 'Frequency' of movie viewing. The dependent variable is "frequency of moviegoing," and the independent variables are "online booking," "movie purchase," "ease of access," and "peer preferences." The key predictive variables, intensity, and timing, are quite simple to capture.

Simonoff and Sparrow (2000) highlight the factors impacting the film's success, such as genres, such as action, children's, comedy, documentary, drama, horror, science fiction, thriller, and so on. Timing of the film's release, such as before a long weekend or during festivals, as well as Academy Award (Oscar®) nominations for the film.

**Relationship Between Sentiment and Box Office Performance**

Kim et. al. (2018) This study developed a domain-specific sentiment analysis lexicon to predict box office performance in the Korean film market. It achieved 81.69% accuracy in sentiment classification and demonstrated a strong positive relationship between consumer sentiment and box office success. Researchers reported in “Text Mining and Sentiment Analysis for Predicting Box Office Success” that this study focuses on the utility of emotion mining for business decisions in the Korean film market. The researchers developed a film domain-specific mood lexicon and tested its performance in predicting box office success. The sentiment analysis with the lexicon showed that the sentiment dictionary is suitable for the film sector in predicting box office success. The preliminary lexicon achieved a classification accuracy of 81.69% and an F1 score of 88.08%. The study also found a strong positive relationship between electronic word-of-mouth consumer sentiment and box office success. This research provides a theoretical reference for dealing with Korean word of mouth and practical guidance for marketers and business users in analyzing and predicting box office success. (Kim et. al. 2018)

There is a link between references to movies in blog posts and their commercial success, according to research on using blogger sentiment to predict a movie's success. Weblog sentiment analysis can strengthen this link. However, there isn't a strong enough correlation between pre-release sentiment and sales to base a forecasting model solely on sentiment. Sentiment can be effectively used in sales prediction models along with other elements like movie genre and season. Weblogs give businesses the chance to examine unsolicited feedback and comprehend customer opinions (Mishane and Glance 2006).

**Deep Learning Based**

Deep learning-based sentiment analysis is now widely used, and this is the current general trend. This study suggests a sentiment analysis method for movie reviews based on deep learning and word mosaic, which is a principle of word mosaic (word vector) and machine learning. The method presented in this article has, according to experiments, correctly classified the emotional content of movie reviews—83.13%. The study also explores different methods of sentiment analysis, including sentiment dictionaries, machine learning, and deep learning. Word embedding is introduced as a feature extraction method based on deep learning algorithms. The article concludes that sentiment analysis based on word embedding has high accuracy and scalability, particularly in shorter sentences. The research involved 6 data sets. The remark note on this study showed that “The article acknowledges some shortcomings in sentiment analysis, particularly in the consideration of special symbols like text symbols or emoticons. The author plans to expand the data set, increase the diversity of sentences, and improve accuracy in the future.” (Zhang et. al. 2021)

(Rani and Singh 2019) It presents a study on the performance of various classification techniques in predicting positive and negative sentiments in movie comments. compare the effectiveness of artificial neural networks, k-nearest neighbor, and hybrid approaches in sentiment analysis. The results show that the hybrid of Artificial Neural Networks and K-nearest neighbors algorithm (ANN and KNN) outperformed other classifiers in terms of precision and accuracy. The study also describes the proposed methodology for data collection, pre-processing, feature extraction, and classification. The findings of this study demonstrate that the most successful method for predicting sentiment in movie reviews is a hybrid strategy using artificial neural networks and k-nearest neighbor classifiers. The performance of different classification techniques, such as hybrid methods, artificial neural networks, and k-nearest neighbor algorithms, was compared by the authors. According to the findings, the hybrid technique performs better in terms of precision and accuracy than other classifiers. This research implies that combining the advantages of various classifiers can improve sentiment analysis outcomes. The study also emphasizes how crucial proper feature extraction is for sentiment analysis. The correct extraction of the collection of attributes used to detect feelings is a key factor in determining whether classification algorithms are successful. The authors propose that feature extraction methods can be enhanced to increase sentiment analysis's precision. Overall, the study's findings offer insightful information about how to use machine learning methods for sentiment analysis of movie reviews. Decision-making procedures in a variety of industries, including marketing and customer service, can benefit from the findings.

(Zhu 2023) study used natural language processing to analyze movie ratings using big data and big data. The model predicts ratings using synopsis, actors, and director information. The research used random forest and neural network algorithms, and SVD with a Bayesian approach. The dataset used is from Kaggle and includes 45,000 movies. The study emphasized conformity modeling and NLP techniques. However, it acknowledged limitations due to technology and expertise. Future research should explore more machine learning algorithms and parameter tuning to improve accuracy.

The study of the literature emphasizes the significance of sentiment analysis in the film industry, stressing its ability to affect decision-making processes such as audience engagement, marketing strategies, and resource allocation. It emphasizes the necessity for additional research in this area in order to change the decision-making methods of the film business.

The literature review reveals a gap in research on emotion analysis in cinema scripts. The proposed research addresses this by collecting sentiment from movie scripts using Natural Language Processing (NLP) approaches. This approach addresses a neglected aspect of sentiment analysis in the film industry. Previous studies have used various sentiment analysis methodologies but did not explicitly use quantitative methods. The research uses Natural Language Processing (NLP) to collect sentiment from movie scripts, providing data-driven insights into the relationship between sentiment and box office performance.

The proposed research would add to the literature by providing empirical evidence on the relationship between sentiment in film scripts and box office success. This empirical component strengthens the credibility and applicability of the research findings and aligns with the film industry's demand for data-driven decision-making. Complementing existing models: The literature research presented numerous sentiment analysis models and their applications. The research applies quantitative sentiment analysis to film scripts and box office returns, using regression models to gain insights into factors determining a film's performance. This research builds on the literature review's importance and focuses on how sentiment analysis can significantly impact the film industry's decision-making processes. It supports the literature's recommendation to use sentiment analysis to improve audience engagement, marketing strategies, and resource allocation. The proposed research is a logical extension of existing material, addressing a specific gap and using a quantitative approach to analyze sentiment in film scripts. It adds important empirical information and practical implications, confirming the need for further research in this critical sector of the film industry.

# **References**

Cox, D.R., 1972. Regression models and life‐tables. Journal of the Royal Statistical Society: Series B (Methodological), 34(2), pp.187-202.

Danyal, M. M., Khan, S. S., Khan, M., Ghaffar, M. B., Khan, B., and Arshad, M,. 2023. Sentiment Analysis Based on Performance of Linear Support Vector Machine and Multinomial Naïve Bayes Using Movie Reviews with Baseline Techniques. Journal on Big Data, 5, 1–18. <https://doi.org/10.32604/jbd.2023.041319>

Dellarocas, C., Zhang, X. and Awad, N.F., 2007. Exploring the value of online product reviews in forecasting sales: The case of motion pictures. Journal of Interactive marketing, 21(4), pp.23-45.[[Google Scholar]](https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Dellarocas%2C+C.%2C+Zhang%2C+X.+%28Michael%29%2C+%26+Awad%2C+N.+F.+%282007%29.+Exploring+the+value+of+online+product+reviews+in+forecasting+sales%3A+The+case+of+motion+pictures.+Journal+of+Interactive+Marketing%2C+21%284%29%2C+23%E2%80%9345.+https%3A%2F%2Fdoi.org%2F10.1002%2Fdir.20087+&btnG=#d=gs_cit&t=1697940324455&u=%2Fscholar%3Fq%3Dinfo%3Axwbo9gerLe8J%3Ascholar.google.com%2F%26output%3Dcite%26scirp%3D0%26hl%3Den)

Di Fabbrizio, G., Aker, A. and Gaizauskas, R., 2011, December. Starlet: multi-document summarization of service and product reviews with balanced rating distributions. In 2011 IEEE 11th International Conference on Data Mining Workshops (pp. 67-74). IEEE. [[Google Scholar]](https://ieeexplore.ieee.org/abstract/document/6137362/)

Feldman, R., 2013. Techniques and applications for sentiment analysis. Communications of the ACM, 56(4), pp.82-89.

Frangidis, P., Georgiou, K., and Papadopoulos, S. (2020). Sentiment Analysis on Movie Scripts and Reviews: Utilizing Sentiment Scores in Rating Prediction. Artificial Intelligence Applications and Innovations, 583, 430–438. doi: 10.1007/978-3-030-49161-1\_36. PMCID: PMC7256373

Giatsoglou, M, Manolis G. Vozalis, Konstantinos D, Athena V, George S, and Konstantinos C. C. 2017. Sentiment analysis leveraging emotions and word embeddings, Expert Systems with Applications, Volume 69,2017 Pages 214-224, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2016.10.043>.

Hu, M. and 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). [[Google Scholar]](https://aclanthology.org/W03-1014.pdf)

Hutto, C. and Gilbert, E., 2014, May. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the international AAAI conference on web and social media (Vol. 8, No. 1, pp. 216-225)[.[Google Scholar]](https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Hutto+%26+Gilbert%2C+2014&btnG=)

Kim, A., Trimi, S., and Lee, S.G. 2021. Exploring the key success factors of films: a survival analysis approach. Service Business, 15(4), 613–638. <https://doi.org/10.1007/s11628-021-00460-x>

Kim, Y., Kang, M., and Jeong, S. R. 2018. Text Mining and Sentiment Analysis for Predicting Box Office Success. KSII Transactions on Internet and Information Systems, vol. 12, no. 8, pp. 4090-4102, 2018. DOI: 10.3837/tiis.2018.08.030.

Leavy, P. ed., 2014. The Oxford handbook of qualitative research. Oxford University Press, USA.

Newman, H. and Joyner, D., 2018. Sentiment analysis of student evaluations of teaching. In Artificial Intelligence in Education: 19th International Conference, AIED 2018, London, UK, June 27–30, 2018, Proceedings, Part II 19 (pp. 246-250). Springer International Publishing.

Mäntylä, M.V., Graziotin, D. and Kuutila, M., 2018. The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. Computer Science Review, 27, pp.16-32.

Mesnil, G., Mikolov, T., Ranzato, M.A. and Bengio, Y., 2014. Ensemble of generative and discriminative techniques for sentiment analysis of movie reviews. arXiv preprint arXiv:1412.5335.

Mikolov, T., Karafiát, M., Burget, L., Cernocký, J. and Khudanpur, S., 2010, September. Recurrent neural network-based language model. In Interspeech (Vol. 2, No. 3, pp. 1045-1048).

Mishne, G. and Glance, N.S., 2006, March. Predicting movie sales from blogger sentiment. In AAAI spring symposium: computational approaches to analyzing weblogs (pp. 155-158).

Mohan Raj, P. and Aditya, S., 2017. Predictive Model for Movie's Success and Sentiment Analysis. Research Journal of Management Sciences, 6(6).

Moon S, Bayus BL, Yi Y, and Kim J. 2015. Local consumers’ reception of imported and domestic movies in the Korean movie market. J Cult Econ. 2015;39(1):99–121. doi:10.1007/s10824-013-9214-x

Onalaja, S., Romero, E. and Yun, B., 2021. Aspect-based Sentiment Analysis of Movie Reviews. SMU Data Science Review, 5(3), p.10.

Pang, B., Lee, L. and Vaithyanathan, S., 2002. Thumbs up? Sentiment classification using machine learning techniques. arXiv preprint cs/0205070.

Pang, B. and Lee, L., 2005. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. arXiv preprint cs/0506075.

Pang, B. and Lee, L., 2008. Opinion mining and sentiment analysis. Foundations and Trends® in information retrieval, 2(1–2), pp.1-135.

Rani, P. and Singh, J., 2019. Performance Modeling of Classification Techniques on Movie Sentiments. International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249-8958 (Online), Volume-9 Issue-2, December, 2019.

Riloff, E. and Wiebe, J., 2003. Learning extraction patterns for subjective expressions. In Proceedings of the 2003 conference on Empirical methods in natural language processing (pp. 105-112).

Sánchez-Rada, J.F. and Iglesias, C.A., 2019. Social context in sentiment analysis: Formal definition, overview of current trends, and framework for comparison. Information Fusion, 52, pp.344-356.

Sangeetha, B., Sangeetha, S., Goutham, D. T., and Vaibhav Ram N. (2023) "Sentiment Analysis on Movie Reviews: A Comparative Analysis," 2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS), Coimbatore, India, 2023, pp. 218-223, doi: 10.1109/ICISCoIS56541.2023.10100367.

Seo, E.J. and Park, J.W., 2018. A study on the effects of social media marketing activities on brand equity and customer response in the airline industry. Journal of Air Transport Management, 66, pp.36-41.

Simonoff, J.S. and Sparrow, I.R., 2000. Predicting movie grosses: Winners and losers, blockbusters and sleepers. Chance, 13(3), pp.15-24.

Socher, R., Pennington, J., Huang, E.H., Ng, A.Y. and Manning, C.D., 2011, July. Semi-supervised recursive autoencoders for predicting sentiment distributions. In Proceedings of the 2011 conference on empirical methods in natural language processing (pp. 151-161).

Thet, T.T., Na, J.C. and Khoo, C.S., 2010. Aspect-based sentiment analysis of movie reviews on discussion boards. Journal of Information Science, 36(6), pp.823-848.

Wang, S.I. and Manning, C.D., 2012, July. Baselines and bigrams: Simple, good sentiment and topic classification. In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers) (pp. 90-94).

Xu, Y. (2023). Evaluation & Analysis of Movie Aspects: Based on Sentiment Analysis. Frontiers in Management Science, 2(3), 64–116. https://doi.org/10.56397/fms.2023.06.08

Zhang, F., Zeng, Q., Lu, L., and Li, Y. (2021). Sentiment analysis of movie reviews based on deep learning. Journal of Physics: Conference Series, 1754(1), 012234. <https://doi.org/10.1088/1742-6596/1754/1/012234>

Zhu, H., 2023. Sentiment Analysis of Movies Based on Natural Language Processing. In 2023 4th International Conference on Education, Knowledge and Information Management (ICEKIM 2023) (pp. 1232-1240). Atlantis Press.