

Analysis of Web Series and Movies Using Sentiment Analysis: A Review

Prachi Avinash Joshi¹, Mahesh Babasaheb Dahiwal²

Department of Information Technology, Government Polytechnic, Aurangabad

Dr. Jayshri Damodar Pagare³

Department of Computer Science and Engineering, Jawaharlal Nehru Engineering college, Aurangabad

Received 22 March 2021; Accepted 06 April 2021

Abstract

In today's world, sentiment analysis is a popular research topic in the Natural Language Processing field (NLP). The primary goal of this research topic is to elicit the emotions and opinions of customers or users through text analysis. Despite the fact that numerous research studies have been conducted in this field, despite the availability of various models, sentiment analysis remains a disadvantage. This paper plans to make a literature review that provides work related to the reviews of web series and movies of different genres using sentiment analysis. Also, Lexicon based on movies review dataset using machine learning sentiment analysis is discussed. Furthermore, the analysis focuses on determining the type of data used. Further, the used environment is discussed, and the paper concludes with appropriate research gaps and challenges, which aids in identifying the application for which sentimental analysis is most needed in future research.

Keywords: Sentiment analysis, web series, natural language processing,

I. INTRODUCTION

Our daily decision-making process has an influence of the opinions of others. Decisions range from buying a property to making investments. All decisions affect various aspects of our daily life. However, in today's internet era, it is much easier to collect opinions from different people. Sources to collect feedbacks are review sites like CNET, Epinions.com e-Commerce sites e.g. Amazon, eBay, Online opinion sites like TripAdvisor, Rotten Tomatoes and also social media like Facebook, twitter. Sentiments can be expressed by words that may be positive, negative, strong or weak. To perform sentiment analysis, one must understand the polarity of words and classify sentiments into categories such as positive, negative or neutral. This task can be accomplished through the use of sentiment lexicons. Sentiment analysis software adds flexibility and insight to the presentation of a brand and its goods, allowing businesses to:

- Monitor consumer perceptions of the brand
- Highlight relevant information about the attitude
- Look for patterns and trends
- Keep a close eye on the influencers' presentations.

Twitter is the most popular social networking site and microblogging platform, with more info. Researchers currently use social media to evaluate the sentiment of users' views on a product, case, or context. Furthermore, sentiment analysis is also known as opinion mining, which is an important NLP activity. The orientation of sentiment is defined by sentiment analysis. The view can be either a user's judgment or assessment, affective state or the deliberated communication of emotion. Sentiment analysis is an area of research that deals with the emotions and sentiments of an individual. The results come in the form of positive data and negative data after the data is analysed through a particular platform. It is basically analyzing the feelings of people all over the world on a particular social networking platform and giving the output of the analysis in short by categorizing the data into positive data and negative data. Sentiment analysis is widely applied to the voice of the client materials like reviews and survey responses. It is basically the concept of NLP which is a field of computer science, linguistics concerned with the interaction between computer and natural languages of humans. Sentiment Analysis is done to fetch and analyse any social media data and then categorizing the fetched data after performing the operations thereafter giving the final output of all the data in short. Social media sentiment analysis is often a wonderful supply of knowledge and may offer insights that may: - 1. Determine to promote

strategy. 2. Improve success during the campaign. 3. Improve electronic messaging 4. Improve the services of the clients A.

Types of Sentiment Analysis

1) Manual Processing Manual processing basically means analyzing data manually. In this process after collection of data from a platform, analysis is performed manually on each and every party of data and the result is also generated manually. This concept is more time consuming and as the data is processed manually so it is difficult to perform analysis on a large amount of data.

2) Keyword Processing Keyword processing means data is analysed through keywords. For example. Data containing words like extraordinary and love will be analysed as positive sentiments and data containing words like awful and detest will be analysed as negative sentiments through keyword processing.

3) NLP (Natural Language Processing) It is basically a concept that acts as a relation between computer language and human language. There is a lexicon dictionary through which the analysis is done. Basically, this dictionary is a collection of data that contains a large amount of information. This dictionary helps in analyzing the data or sentiments into positive and negative ones. Nowadays the concept of NLP and machine learning is used to perform sentiment analysis. This technology has made analyzing data easier. In fact, these concepts give access to a large amount of data in one go due to which results that come through analysis are more accurate.

The following are the key contributions of this paper:

- To be exposed to a critical study of sentiment analysis in film and web series reviews.
- To conduct a thorough examination of different sentiment analysis models based on machine learning algorithms.

The following is the structure of the sentiment analysis classification review:

The following is the structure of the sentiment analysis classification review: The literature review on traditional sentimental analysis is detailed in Section II. Section III addresses sentiment analysis and performance metrics using machine learning algorithms. Section IV specifies Performance Evaluation and Section V Conclusion and challenges in Sentiment Analysis.

II.LITERATURE SURVEY

Many different approaches and work have been performed to analyse data and perform sentiment analysis. Work has been done in many languages. Following are the fields in which sentiment analysis has played an important role:

1. Social media monitoring
2. Brand monitoring
3. Customer feedback
4. Customer support
- A. Sentiment analysis using Machine learning approaches [Lexicon based on movie review dataset] [1]

Sentiment analysis is Lexicon based approach which has a dictionary of positive and negative words which is used to determine sentiment polarity based on inclination of messages from source dataset. Hybrid based approach uses machine learning and lexicon-based approach for classification. Two methods proposed are:

1. Training method
2. Prediction method

In training method model learns to map a specific input(text) to corresponding output (tag) based on test samples for training. Text input into a feature vector is done by feature extractor. Whereas in prediction method, the unseen text inputs are transformed into feature vectors which are further fed into the model generating predicted or expected tags i.e., positive, negative, neutral.

It is possible to perform sentiment analysis by lexicon-based approach, machine learning or hybrid approach. If the size of lexicon is increased, it will increase the approach and becomes more incorrect and time consuming.

Natural language toolkit (NLTK) feature in python can be used for further implementation sample movie review data.

- B. Discussion of sentiment analysis-based application for analyzing reviews of web series and movies and sentiment analysis using machine learning approach i.e. Lexicon based on movie review dataset and application for analysing reviews of movies and web series of different genres has been discussed as follows:

In 2018 Charu Nanda et. At [2] used Hindi Sentiwordnet, Random Forest algorithm to find reviews of Hindi movies. As a result, High accuracy was achieved. But the demerit was only Hindi words were used for analysis and two sentiments were shown positive and negative.

In 2018 Fang et al. [3] used the semantic fuzziness and suggested multi-strategy. The results show that the proposed model has a high degree of efficiency.

In 2019 Tejaswini M. Untwale & Prof. G. Choudhari [4] used Naïve Baye and Random forest algorithms for Feature selection and execution to perform Sentiment analysis of big amount of data. Results were found as web series of comedy and crime categories were more liked by themore liked by the viewers than web series and horror movies and web series are equally liked. The work can be more extended for analytical work. The dataset can be extended along with more categories can be incorporated. Positive and negative review analysis are done.

In 2019 Jayashree Jagdale and Dr. Emmanuel M. [5] used Hybrid corrective critic neural network for Sentiment analysis of movies reviews from community media. Output was developed with the help of technology like metrics. CCNN showed better results than SVM, NB algorithms.

In 2019 Saad and Yang [6] As a first step, tweets were pre-processed, and an effective feature was created using the feature extraction model. For classifying the sentiment analysis, methods such as SVR, RF, Multinomial logistic regression (SoftMax), and DTs were used to give a complete tweet sentiment analysis based on regression with machine learning algorithm. The suggested model has attained the best accuracy.

In 2020 AkshiKumar, KathiravanSrinivasan, ChengWen-Huang, and Albert Y. Zomaya [7] An aggregation model was created to quantify the hybrid polarity. In addition, SVM was used to train the BoVW in order to predict the sentiment of visual content. ConVNetSMany different approaches and work have been performed to analyse data and perform sentiment analysis. Work has been done in many languages. Following are the fields in which sentiment analysis has played an important role:

1. Social media monitoring
2. Brand monitoring
3. Customer feedback
4. Customer support

III. VARIOUS MACHINE LEARNING ALGORITHM FOR SENTIMENT ANALYSIS

Sentiment analysis process on product reviews. As shown in Figure 1, sentiment analysis can be thought of as a classification method. Document-level, sentence-level, and aspect-level sentiment analysis are the three primary classification categories of sentiment analysis. Sentiment analysis at the document level helps to identify an opinion document as reflecting a positive or negative sentiment. It considers the entire document to be a single unit of basic information (talking about one topic). Sentiment analysis at the sentence level helps to describe the sentiment conveyed in each sentence. Determine whether the sentence is subjective or factual as the first step. Sentence-level SA can decide whether the sentence expresses positive or negative thoughts if the sentence is subjective. The Sentiment Classification (SC) techniques, as shown in Figure.2.

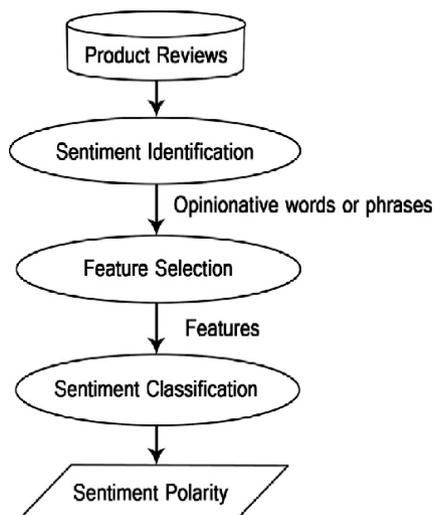


Figure.1. *Sentiment analysis process on reviews.*

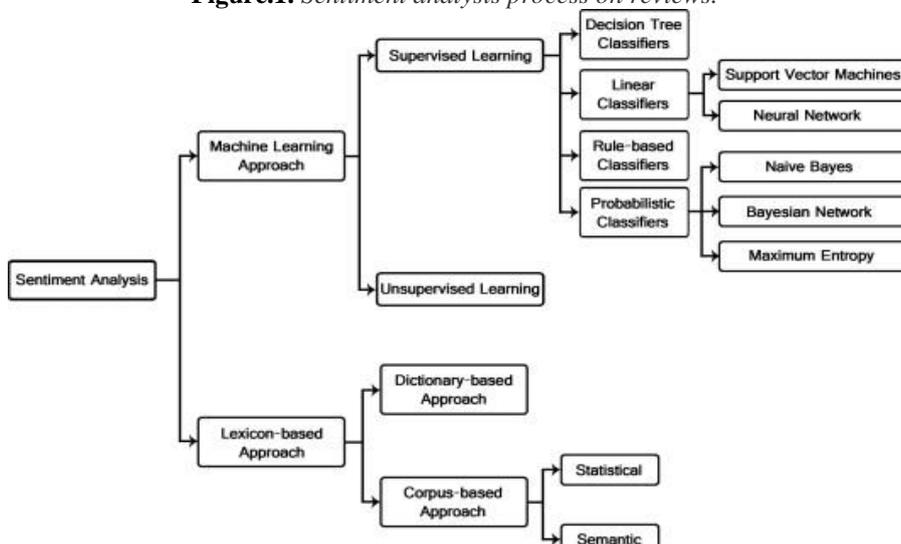


Figure 2. *Sentiment classification techniques.*

IV. PERFORMANCE EVALUATION

The proposed application performs the sentiment analysis for analyzing reviews of web series and movies of different genres from the duration of 2017 to 2019 [8]. As per the analysis of average positive reviews in it is found that the web series of comedy genres is more liked by the audience than the movies of the comedy genre. In the romance genre, it is found that the movies are more liked by the audience than the web series. In the horror genre, it is found that both the movies and web series are equally liked by the audience. In the crime genre, it is found that the web series is slightly liked by the audience than the movies. As per the analysis of average negative reviews in, it is found that the web series of comedy genre is more disliked by the audience than the movies of the comedy genre. In the romance genre, it is found that the movies are less liked by the audience than the web series but the difference is quite less. In the horror genre, it is found that the movies are more disliked by the audience than the web series of the horror genre and the dislike difference is quite high. In the crime genre, it is found that the web series is more disliked by the audience than the movies and the dislike difference is quite high.

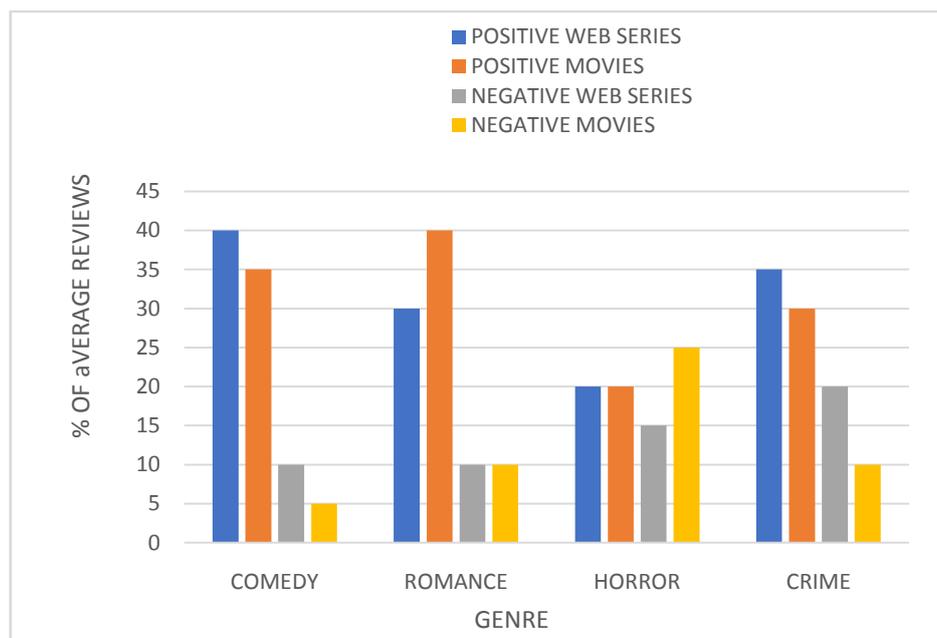


Figure 3. Graphical representation of average positive and negative reviews of web series and movies of all genre

V. CONCLUSION

In today's era of internet people express their views or opinions on the web so for analysing different opinions online content is useful. A lot of research is present in the literature for detecting sentiment from the text still there is huge scope improvement of these existing sentiment analysis models. The strength of the sentiment classification depends on the scale of the lexicon(dictionary) because the size will increase the approach and becomes more incorrect and time consuming. Existing sentiment analysis models can be improved further with more semantic and common-sense knowledge also for the problems with negation, word ambiguity and multipolarity sentiment analysis is a challenging task.

REFERENCES

- [1]. Ayushi Mitra, "Sentiment Analysis using Machine Learning Approaches(Lexicon based on movie review dataset)", Journal of Ubiquitous Computing and Communication Technologies(UCCT)2020,Vol02/No.3 Pages 142-152
- [2]. Charu Nanda, Mohit Dua and Garima Nanda, "Sentiment Analysis on Movies Reviews in Hindi Language using Machine Learning", IEEE, 2018(ICCSP), Chennai, India, 3-5 April 2018.
- [3]. Y. Fang, H. Tan and J. Zhang, "Multi-Strategy Sentiment Analysis of Consumer Reviews Based on Semantic Fuzziness," IEEE Access, vol. 6, pp. 20625-20631, 2018.
- [4]. Tejaswini M. Untawale, G. Choudhari, "Implementation of Sentiment Classification of Movie Reviews by Supervised Machine Learning Approaches",IEEE, 2019 3rd (ICCMC), Erode, India,27-29 march 2019.
- [5]. Jayashree Jagdale, M Emmanuel, "Hybrid Corrective Critic Neural Network for Sentiment Classification in Community Media", 2019 3rd (ICECA),2019.
- [6]. S. E. Saad and J. Yang, "Twitter Sentiment Analysis Based on Ordinal Regression," IEEE Access, vol. 7, pp. 163677-163685, 2019.
- [7]. AkshiKumar, KathiravanSrinivasan, ChengWen-Huang, and Albert Y.Zomaya, "Hybrid context enriched deep learning model for fine-grained sentiment analysis in textual and visual semiotic modality social data", Information Processing & Management, vol. 57, no. 1, January 2020.
- [8]. Aishwarys, ParthWadhwa, Prabhishek Singh, "A New Sentiment Analysis based Application for Analyzing Reviews of Web Series and Movies of Different Genres,10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)