Volume 66, Issue 2, 2022
Journal of Scientific Research
of
The Banaras Hindu University

Entertainment Advisor Using Sentiment Analysis

Abdul Aleem[1][0000-0001-6676-9072], Palash Tewary[1], Shanta Karn[1], and Varun Kumar[1]

[1]Department of Computer Science and Engineering
Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, Odisha
abdulaleem@soa.ac.in, {palastewary, karnshanta, varun2asses}@gmail.com

Abstract. Sentiment or opinion mining analyses include the study of people's written opinions, feelings, attitudes, and emotions. It has been one of the most active areas of research in natural language processing and text mining in recent years. Due to its importance for business and society, research has also moved to management science and social sciences beyond computer science. This article presents an Entertainment Advisor system that uses Sentiment Analysis to calculate the Sentiment Polarity and Subjectivity of Movies or TV shows by reading its reviews. The proposed system is a recommender system, which recommends based on the sentiment polarity. The features provided in the proposed system would let users write their reviews. These reviews would be directly stored in a text file and accessible to all the users for anonymous view. The recommendation is made based on the polarity of all the reviews. The proposed recommender system also allows users to select a particular review and calculate its polarity for demonstration and better understanding.

Keywords: Sentiment Analysis, Text Mining, Natural Language Processing, Recommender System, Entertainment Advisor.

1 Introduction

Sentiment analysis is one of the rigorous growing research areas due to two causes. First, a lot of real-world scenarios exist, where opinion is essential for performing any activity. It is vital to one's behavior and beneficial to take others' opinions while making any decision. Secondly, there are numerous tough challenges with research that have not been tried before 2000. The main reason for the digital text rarely used for opinion mining has been the absence of earlier investigations. It is, therefore, no surprise that the beginnings of the sector and the social media expansion coincides with rapid expansion on the Internet, where sentiment analysis finds a way forward in most real-time applications.

The importance of entertainment in our lives cannot be overstated. It draws people together and is also an excellent method for the entire family to participate in the tie-in. We do not have any refreshments in our lives if we do not have amusement. Almost all the people in our country watch movies and TV series, out of those people, most of them like to watch movies as soon as they are released, whereas some wait to hear how the said movie was and if it is good, then they proceed to watch it [1]. One could say that a thorough review of a particular movie is essential for its success. Others can also say that reviews are also crucial for moviegoers as they are to movie makers. This article proposes a recommender system, termed Entertainment Advisor, to address the requirement of review or opinion for entertainment kinds of stuff. The proposed system recommends a suitable suggestion based on the sentiment polarity of the reviews provided by other users.

The proposed system uses sentiment analysis for drawing an opinion on movies or other entertainment visual series to calculate their success rate. Besides technological enhancement, Entertainment Advisor, in layman terms, also helps the common people select their amusement and interests, which develops better customer relationships [2]. The most interesting feature of the proposed system is that many users can share their reviews in a chat session, which could deliver live understanding also to the seekers. The proposal's main idea is to promote the content of good quality for a user's entertainment. It is known that entertainment will always have a global impact no matter how far the timeline scales up. Hence, this Entertainment Advisor system also needs improvement at a specific stage so that it might be able to succeed with the global recommendation and multiple users.

The rest of the article is organized as follows. The second section discusses the background and related work. The third section describes the proposal of Entertainment Advisor System. The fourth section presents the implementation details along with the outcomes. Finally, the fifth section concludes this article along with some directions for future research.

2 Background and Related Work

Before watching any shows, people always want to have a review. Nowadays, many shows are coming up in the market with no guarantee that all of them will have good reviews. People do not have time and money to waste on the bad shows. So, they prefer to have a second opinion before going to any show. Review is critical at present. People generally use websites such as IMDB, Rotten Tomatoes, and Metacritic to see reviews of movies they desire to see [3]. These websites also allow them to store their reviews in their database on servers. However, users have to access these numerous reviews to conclude individually.

To better understand the requirement of an entertainment advisor system, we conducted an online survey. The survey had a questionnaire that represented a detailed analysis in the form

DOI: 10.37398/JSR.2022.660213

of statistical data, as illustrated in Fig. 1, from which several inferences could be drawn. The survey findings confirm that many people watch or do not watch movies (and video series) based on reviews. Most movie-interested people wait for the reviews to come before going to watch any movies.

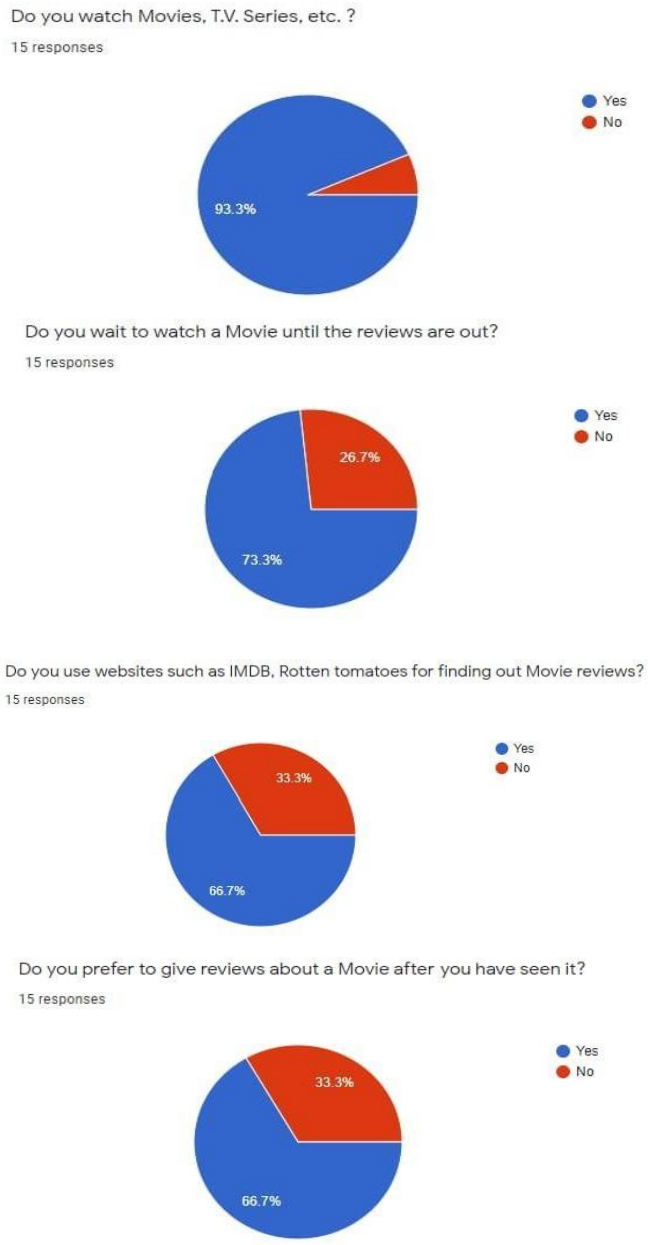

Fig. 1: Findings of online survey regarding recommendation for entertainment

However, websites such as IMDB, Metacritic, etc., are not so efficient nowadays to provide all the required sparse features to integrate the necessary requirements of a general movie enthusiast [4]. In order to gain a better overview of the entire movie review process, each individual review must be taken into account. For this, we need a recommender system that could utilize the sentiments of the reviews. Various researches in the past have addressed this issue.

Gamon [5] proposed an approach for assigning docs sentiment using a four-point scale. The input was taken from the customer feedback. The authors used SVM for the classification of feedback into sentiments. Pang and Lee [6] assigned assign docs sentiment using a three or four-point scale. The sentiments are based on movie reviews. SVM, regression, and metric labeling have been used for the classification of reviews. Choi et al. [7] extracted sources of opinion, emotions, and sentiments using MPQA corpus as the dataset. The authors applied CRF and AutoSlog along with 10-fold cross-validation for sentimental analysis. Wilson et al. [8] assigned expressions +/−/both/neutral to review using MPQA corpus as the dataset. The authors applied BoosTexter and 10-fold cross-validation for sentimental analysis.

König and Brill [9] proposed an approach for assigning docs sentiment. The input was taken from the movie feedback. The authors used Pattern-based, SVM, and hybrid for the classification of feedback into sentiments. Melville et al. [10] gave a framework for sentimental analysis of blogs. The proposed framework includes lexical knowledge in supervised learning for text categorization. This framework is then applied for sentiment classification. Jo and Choi [11] predicted movie rating through sentiment analysis.

There are some other similar works done in refs. [12-18] employed machine learning techniques for sentiment analysis or some other classification techniques. This article does not apply any classification technique or other machine learning techniques for sentiment analysis. The article presents a recommender system that provides suggestions based on all the reviews' collective sentiment polarity.

3 Proposed Work

This article proposed a recommender system, Entertainment Advisor, whose motto is to provide a user-friendly platform to help users submit their reviews and find the reviews and remarks given by other users and authorities. Based on the existing reviews and remarks available in the database, the system also recommends suggestions as per the sentiment analysis of the reviews. The design of the user interface is through a menu-driven approach that asks a user to enter its choices, and based on the user's choice; the functionalities can be further driven. In this proposed recommender system, the reviews get stored in the local system folder where the software itself is located. Hence, this recommender system does not always require an internet connection to work. As soon as the system is connected to the Internet, the database is updated (or synched) as per the reviews and remarks available on the Internet Server. Upon syncing, the proposed recommendation system suggests according to the updated database.

The proposed system aims to provide the review of movies and series based on the user's choice. Entertainment Advisor works without an Internet connection and requires a connection to the Internet only if database synching is required. Multiple users can use the proposed recommender system at the same time. The workflow diagram of the Entertainment Advisor is shown in Fig. 2. The system on start-up will sync with the movie_review.txt file and series_review.txt file kept at the server location of the recommender system (if connected to the Internet). In case the system is not connected to the Internet, it will use these files available locally. The system would then provide a menu with three options to end-users and ask for their choice as per the following details: 1: Movie Review, 2: Series

Review, and 3. Input Your Review. Based on the user's input, a further course of action is decided.

If the input is one, it will further access the movie_review.txt file and display a list of movies to provide sentiment polarity of all the reviews for a selected movie. If the input is two, it will further access the series_review.txt file and display a list of TV series for providing sentiment polarity of all the reviews for a selected TV series. If the input is three, it will give two more options for either entering the user's own review or finding the sentiment of the existing review. In the first case, the Entertainment Advisor will allow the users to write their own reviews in userlist.txt. If the review is about a new item, then it will be added to the list. In the second case, the Entertainment Advisor will find the sentiment polarity of the existing review(s). After finding the sentiment analysis of the review again, it will ask for the continuation whether the user wants to continue(yes/no)? If yes then, it will start again from the beginning, or else it will exit.

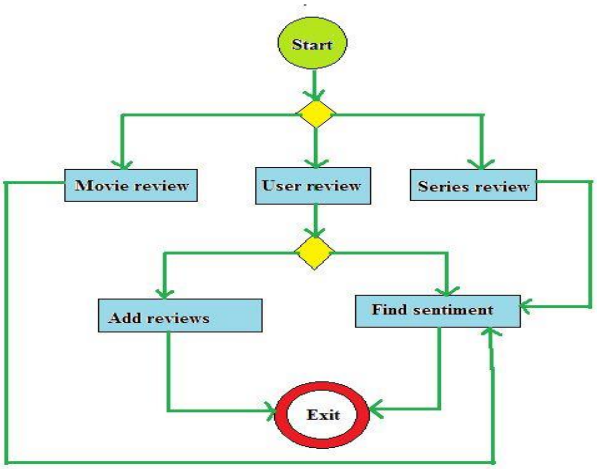

Fig. 2: Workflow diagram of the proposed system

The next significant task of the proposed system is to calculate the sentiment polarity, which is correct up to 9 decimal places. The computed sentiment polarity gives the user a proper value of how exactly the sentiment of other users is for that particular movie/series. Hence, this recommendation is efficient and more accurate than other methods in the existing literature. The proposed Entertainment Advisor is user-friendly and easy to use, which can be operated through a particular menu to execute the commands as per users' choices. The implementation of Entertainment Advisor leads to an easy-to-use recommender system, which even a naïve user can operate efficiently. The implementation details of the proposal have been explained in the next section.

4 Implementation Details

The proposed recommender system – Entertainment Advisor, has been implemented in Python using NLTK and Textblob [16] library that involves the Sentiment Analysis process. The system configuration of the development environment for the implementation of Entertainment Advisor is as follows:

- **Hardware Specification:** The proposed system can be implemented and executed on a machine having Windows 7 or a higher version. A Linux platform, like Ubuntu, could also be used. Entertainment Advisor has the maximum memory requirement of 500 MB storage space only and flawlessly executes with 4 GB or higher RAM. In our experimentation work, four Intel machines have been used with different operating systems. The proposed system is seamlessly executed on both the Windows and Linux machines having minimum specified criteria.
- **Software Specifications:** The complete application, including the Textblob feature that incorporates the Sentiment Analysis process, has been written in Python 3.7. Thus, the software needed is Python 3.7 (or above) with pre-installed Textblob and NLTK libraries.
- **Compiler used:** The user can directly access the software through the Python IDLE platform or by going to the command prompt and typing the file's location which is utilized for storing the database files..

From the user's perspective, Entertainment Advisor is a menu-driven software that initially shows a welcome message and asks the user for the task to be performed, i.e., whether to find out if a movie or a TV show has a good or bad review, or if he/she wants to add their own review for any particular shows of their choice. The software stores the entered review so that the user can assess the review later and find out the sentiment of that particular review by searching for it.

TextBlob utility is used to justify the review. TextBlob is an open-source library for Python 2 and 3[16]. It offers a basic API for standard natural language processing (NLP) activities, including part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, and translation. TextBlob returns the polarity and subjectivity score of a sentence. The polarity score lies between [-1,1], where -1 defines a negative sentiment, and 1 defines a positive sentiment. Negation words reverse the polarity. TextBlob has semantic labels that help with fine-grained analysis. If the Textblob score = 0 then, it is a neutral sentiment. If the Textblob score is < 0 then, it gives a negative sentiment. If the Textblob score is > 0 then, it is a positive sentiment.

Another library utilized is Natural Language Tool Kit (NLTK), which is used for tokenization, parsing, categorization, stemming, tagging, and semantic reasoning text processing packages [16]. NLTK is a Python library function that works on human language data in statistical natural language processing (NLP).

The datasets used here are the text files that contain a selected number of reviews. There are separate text files each for movies and TV series and a separate text file for the user's own personalized set of reviews. The reviews added by the users will not disappear when Entertainment Advisor runs for the next time. The user can directly access the software through Python IDLE platform or the command prompt. In both cases, the software runs smoothly and exit when completes the iteration. The application has been tested and run across different platforms on which it runs smoothly.

5 Results and Analysis

A real-time validation test has been conducted for the Entertainment Advisor. This research product was sent to around 100 people for validation. The people were asked to use the product and give feedback. Around 90% of the people found it very easy to use as compared to any other software. The major feature highlighted by the validators was that, unlike other software, Entertainment Advisor does not force the user to give their own review while looking for the reviews. Test results were overall satisfactory in real-time. The key observations on testing multiple components of the Entertainment Advisor are as follows.

1. The proposed recommender system has handled around 95 client's access at a time without any error.
2. The user interface of this software is straightforward to understand how to use.
3. People with different age groups can access this without any difficulty.
4. This software is platform-independent so that it can be used on machines operated with Windows, Linux, IOS, Android, etc.
5. As this software can run even without an internet connection, it became beneficial for people with no internet connection availability round the clock.

The research product has been pre-loaded with some reviews of few movies. The users have been also requested to upload the reviews for new entries. Upon executing the software, a welcome note is followed by a menu that asks the user for a preferable option to select among movie reviews, series reviews, or user reviews. The first two options are for showing the reviews of the pre-loaded entries. The third option enables user to enter their own review. Based on the user's input, the recommender system lists the items present in the directory whose review can be fetched and also computes the sentiment polarity of the review. Fig. 3 shows the review for "The Salesman" movie. The first option for movie is selected and then the movie "The Salesman" is selected. The review fed before is shown to the user and its sentiment analysis is done. In Fig. 3, the review shown is positive, with the sentiment polarity value of 0.19166666666666.

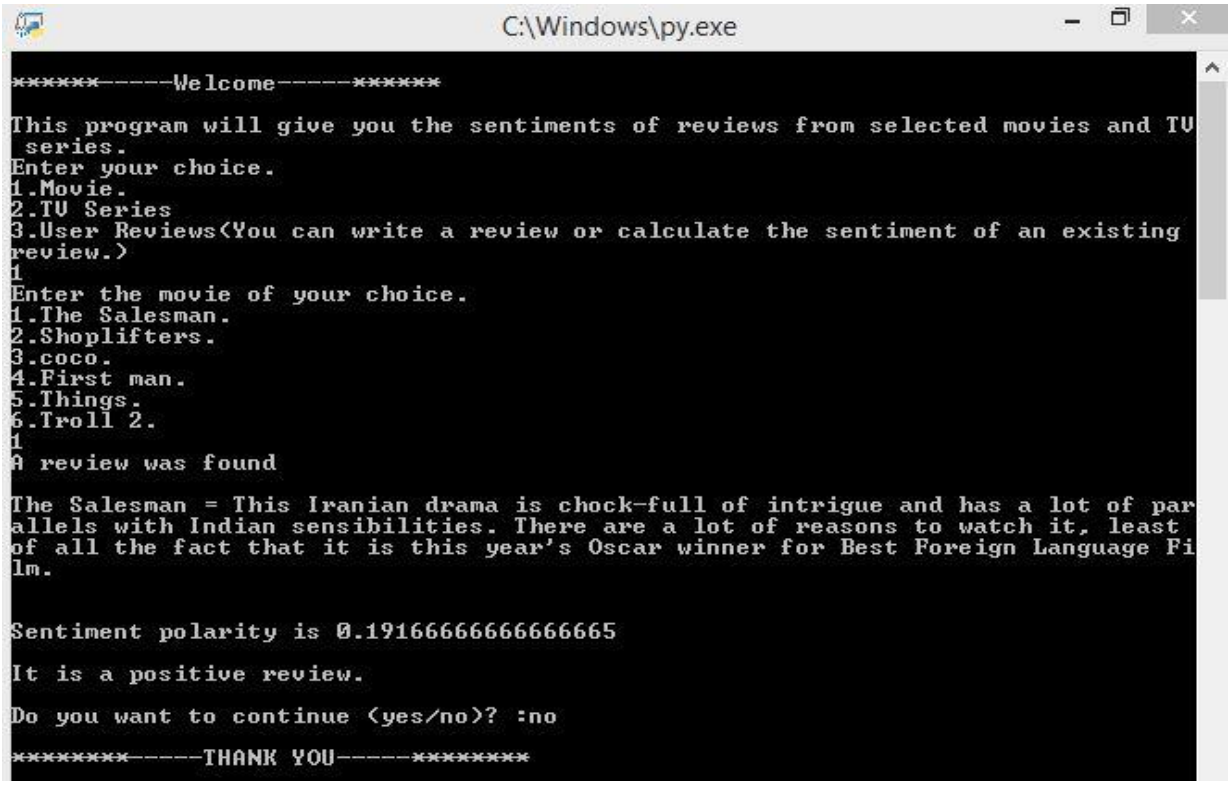

Fig. 3: Review and sentiment polarity of a movie

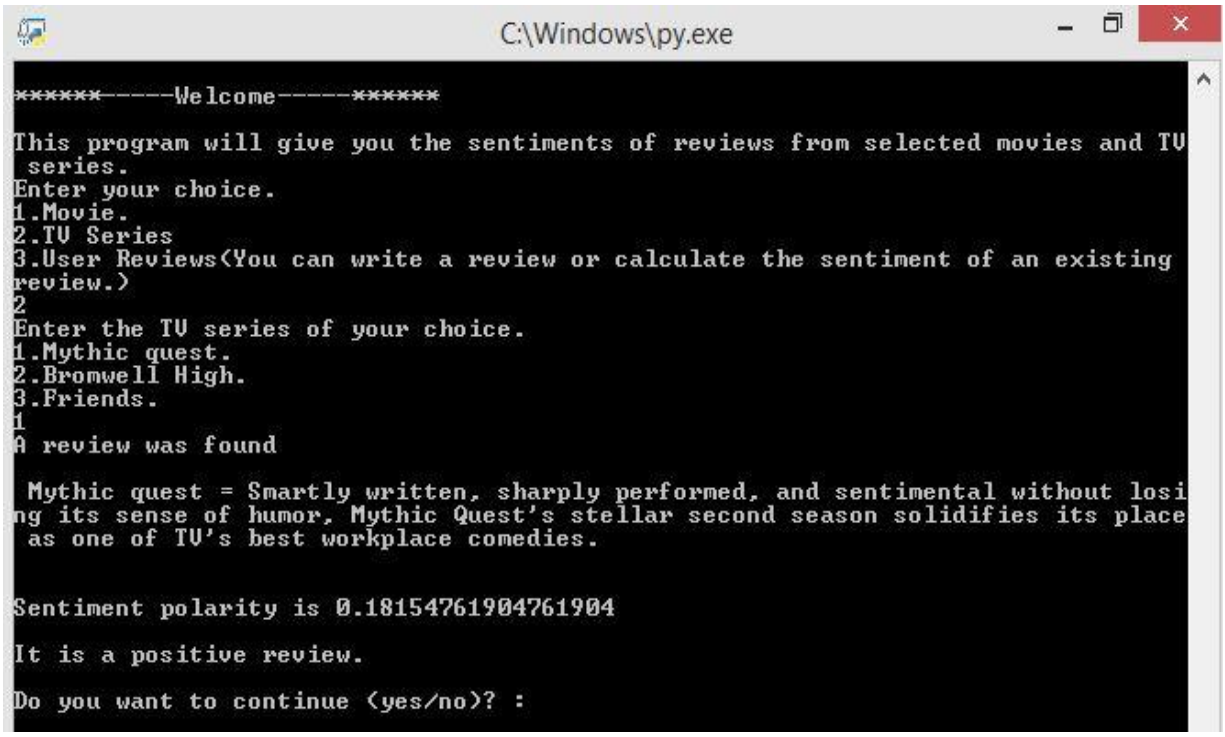

Fig. 4: Review and sentiment polarity of a TV series

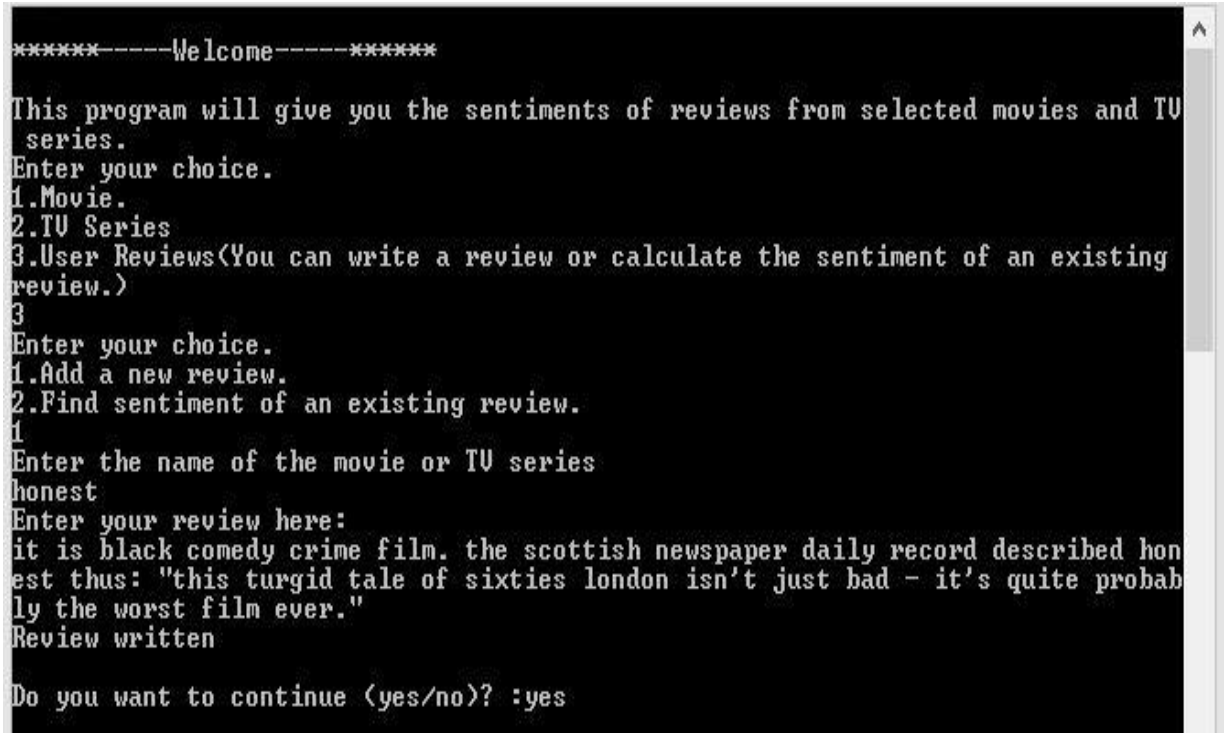

Fig. 5: Adding a customized review by user

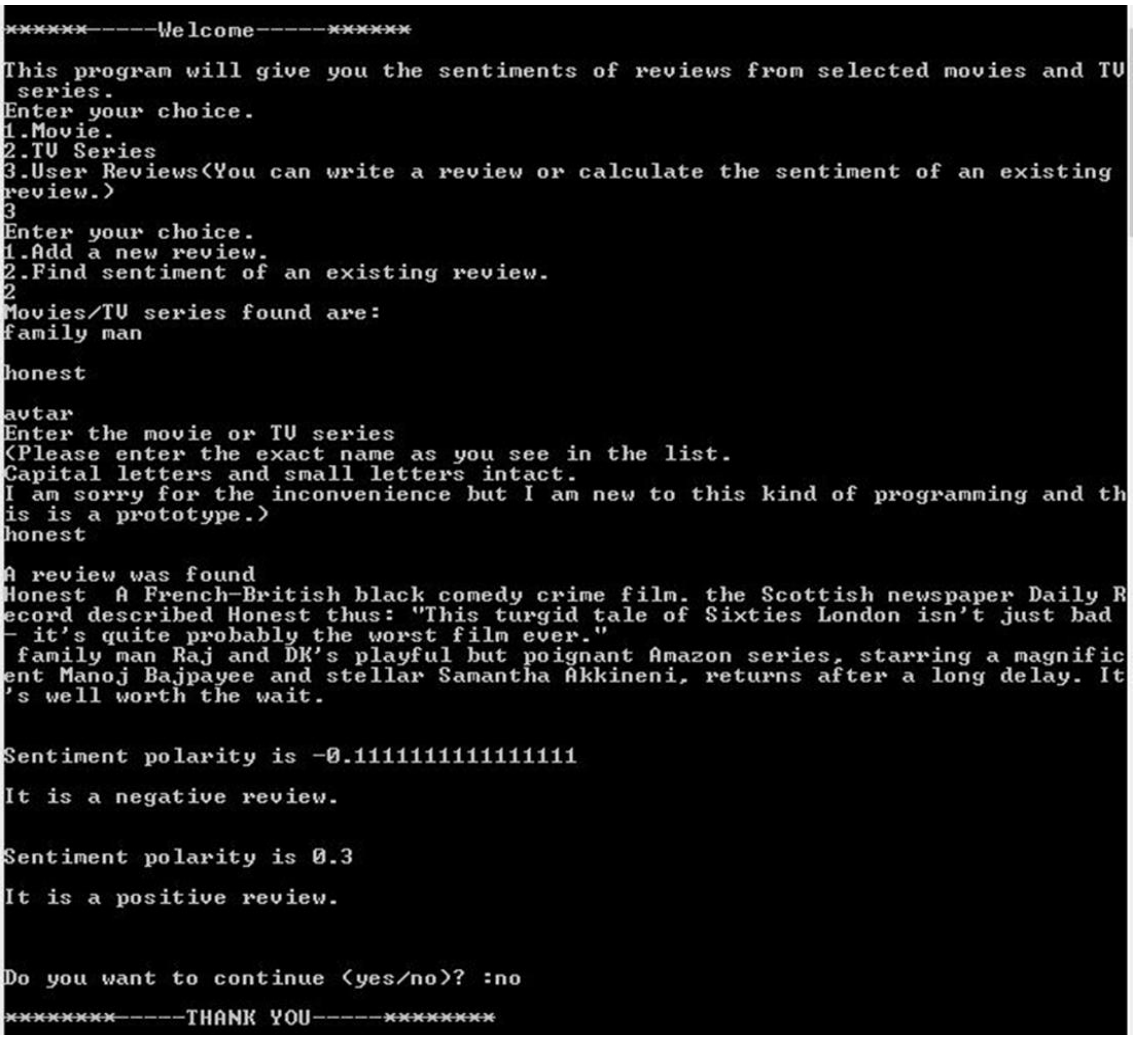

Fig. 6: Sentiment polarity of a customized review

Similarly, the second option leads to the display of TV series whose reviews are already available in the system. Fig. 4 shows the review for the TV series – "Mythic Guest". The testing product had reviews pre-loaded only for three TC series. The second option in the main menu is selected for TV series and then the series "Mythic Guest" is selected. The review fed before is shown to the user and its sentiment analysis is done. In Fig. 4, the review displayed is positive, with the sentiment polarity value of 0.18154761904.

The third option of the main menu is for the customized user review section, as shown in Fig. 5. After the user opts for user reviews, again, two options are provided in sub-menu for either adding a review or finding the sentiment polarity of an existing review. Fig. 5 shows the input of a user review for the movie honest. After the update, the name and review of the movie is added to the list. Again selecting third option would provide user with two options of sub-menu as before. To compute the sentiment polarity, user needs to select the second option in sub-menu. The sentiment polarity of the review entered by the user is computed and displayed through the entered name, as shown in Fig.6.

The results and the validation prove that sentiment analysis is the most appropriate option for deciding the reviews of entertainment shows, whether it is a movie or a TV series. The research product is Entertainment Advisor, which is a recommender system and an easy to use application. It can easily be made open-source and user can provide their reviews and synchronize to the server whenever they have Internet connectivity. At the core, this research is able to deliver a recommender system based on sentiment analysis, which could be easily utilized by end-user.

6 Conclusion and Future Scope

Sentiment Analysis can be used in the entertainment field for developing apps and programs that predict movie success through user reviews. By successfully implementing the proposed recommender system, it can be inferred that Entertainment Advisor is a helpful utility for common people who mostly require a second opinion before choosing something to watch and the people who are willing to help those kinds of people. The proposed system is advantageous over other traditional machine learning techniques as the role of prediction is deduced through exact values of sentiment polarity. Nonetheless, the proposed system does not require Internet connectivity for execution as with the other existing recommender system. It works on the solo machine and the

database files synchronize to the web server periodically whenever there is an Internet connectivity or as per the user's assistance. The lightweight version, easy to use facility and recommendation based on sentiment polarity makes it a perfect tool for all entertainment loving users. The future scope of the proposed system includes the addition of newer features such as regular top 10 movie recommender, movie of the day, trending movies, or series.

References

[1] Reiss S., and Wiltz J.: Why People Watch Reality TV, Media Psychology, vol. 6, no. 4, pp. 363-378, 2004. Available: 10.1207/s1532785xmep0604_3.

[2] Barnes, J.: Establishing meaningful customer relationships: why some companies and brands mean more to their customers, Managing Service Quality: An International Journal, vol. 13, no. 3, pp. 178-186, 2003. Available: 10.1108/09604520310476445.

[3] IMDb: Movies, TV, and Celebrities, Choice Reviews Online, vol. 52, no. 11, pp. 52-5640-52-5640, 2015. Available: 10.5860/choice.190453.

[4] Lindsay E, B.: The Internet Movie Database (IMDb)9952The Internet Movie Database (IMDb). The Internet Movie Database Ltd, http://www.imdb.com, Electronic Resources Review, vol. 3, no. 5, pp. 56-57, 1999. Available: 10.1108/err.1999.3.5.56.52.

[5] Gamon, M.: Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis. COLING 2004: Proceedings of the 20th International Conference on Computational Linguistics. 2004.

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

[7] Choi, Y., Cardie, C., Riloff, E. and Patwardhan, S.: Identifying sources of opinions with conditional random fields and extraction patterns. Proceedings of human language technology conference and conference on empirical methods in natural language processing. 2005.

[8] Wilson, T., Wiebe J., and Hoffmann P.: Recognizing contextual polarity in phrase-level sentiment analysis. Proceedings of human language technology conference and conference on empirical methods in natural language processing. 2005.

[9] König, A. C., and Brill, E.: Reducing the human overhead in text categorization. Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining. 2006.

[10] Melville, P., Gryc, W., and Lawrence, R. D.: Sentiment analysis of blogs by combining lexical knowledge with text classification. Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. 2009.

[11] Jung-Tae, J. and Choi, S.: Sentiment Analysis of movie review for predicting movie rating, Management & Information Systems Review, vol. 34, no. 3, pp. 161-177, 2015. Available: 10.29214/damis.2015.34.3.009.

[12] Zhang, J., Wang, Y. and Xie, T.: Software feature refinement prioritization based on online user review mining, Information and Software Technology, vol. 108, pp. 30-34, 2019. Available: 10.1016/j.infsof.2018.12.002.

[13] Garg S., Aleem A., Gore M.M.: Employing Deep Neural Network for Early Prediction of Students' Performance. In: Udgata S.K., Sethi S., Srirama S.N. (eds) Intelligent Systems. Lecture Notes in Networks and Systems, vol 185. Springer, Singapore. 2021.

[14] Sarvakar, K. and Kuchara, U.: Sentiment Analysis of movie reviews A new feature-based sentiment classification, International Journal of Scientific Research in Computer Science and Engineering, vol. 6, no. 3, pp. 8-12, 2018. Available: 10.26438/ijsrcse/v6i3.812.

[15] Sokol, A.: Movie Recommender System, Southeast Europe Journal of Soft Computing, vol. 8, no. 2, 2019. Available: 10.21533/scjournal.v8i2.180.

[16] Aleem A., Kumar A., and Gore M. M.: A study of manuscripts evolution to perfection. In 2nd International Conference on Adavanced Computing and Software Engneering - ICACSE 2019, pages 278-282, 2019.

[17] Nishadi, A. T.: Text Analysis: Naïve Bayes Algorithm using Python JupyterLab, International Journal of Scientific and Research Publications (IJSRP), vol. 9, no. 11, p. p9515, 2019. Available: 10.29322/ijsrp.9.11.2019.

[18] Yao, J.:Automated Sentiment Analysis of Text Data with NLTK, Journal of Physics: Conference Series, vol. 1187, no. 5, p. 052020, 2019. Available: 10.1088/1742-6596/1187/5/052020.

***