DOI : https://doi.org/10.56452/6-3-669

ISSN: 0974-5823 Vol. 6 No. 3 December, 2021

International Journal of Mechanical Engineering

Enhancing Movie Recommendation Systems with Deep Learning and Sentiment Analysis

Chandrakala Arya

Asst. Professor, School of Computing, Graphic Era Hill University,

Dehradun, Uttarakhand India 248002

ABSTRACT

In recent years, movie recommendation systems have grown in popularity as they provide a customized and effective approach for consumers to discover new films that cater to their preferences. Deep learning and sentiment analysis methods are used to improve these systems. Deep learning is a kind of machine learning that processes and analyses massive volumes of data using neural networks. Movie recommendation systems increase their accuracy and relevance by utilizing deep learning to analyse a variety of data sources, including user ratings, movie genres, and stars and actresses. On the other hand, sentiment analysis is the act of examining text to identify the feelings and view-points conveyed in it. Sentiment analysis is used in the context of movie recommendation systems to analyse user reviews and comments to ascertain the sentiment towards a specific movie. Systems for making movie recommendations better comprehend a person's tastes and emotional reaction by analysing the sentiment of user reviews. Neural networks are used as a method of merging deep learning and sentiment analysis into movie recommendation systems. Inputs like movie names, genres, and actors are used to train neural networks to predict user ratings and preferences. Sentiment analysis allows the neural network to take into account viewers' emotional reactions to certain movies, resulting in more precise and relevant suggestions.

1. INTRODUCTION

The kinds of evaluations a film receives from the public determine its level of popularity. The decisions made by other users are also influenced by these reviews. A movie that was highly enjoyed by viewers is more likely to be chosen by users than one that received mixed reviews. Making decisions is made more challenging by analysing these evaluations and avoiding those that provide false information. Sentiment analysis offers a remedy for this issue. Sentiment analysis makes it possible to extract information from text sources, classify the statements, words, or documents as positive or negative, and then use NLP (natural language processing) to categorize the information. Knowing the author's viewpoint and indicating the user experience are both highly helpful.

Opinion mining extracts and categorizes the views stated in different online forums or platforms using data mining methods. This makes it possible to more clearly grasp how the user feels about a certain topic. The algorithm described in the study analyses user evaluations and categorizes them as good or bad in addition to making movie recommendations to users. The Cosine Similarity method is used to provide movie recommendations, and a comparison is made between the SVM and NB algorithms to analyse the reviews' sentiment. The study's goals are to cope with the massive amount of data and filter important information, suggest related films based on user preferences, and do sentiment analysis on reviews of the selected movie. People's everyday lives are becoming increasingly reliant on mobile services as a result of the adoption of mobile smartphones. Mobile devices are used by people to access company information, product information, promotion information, and suggestion information. The use of mobile services for movie recommendations is significant. A movie recommendation engine has shown to be an effective tool for offering consumers relevant film selections. The recommendations are offered to assist users in their efforts to manage information overload and to make it simple and quick for them to locate the right movies. Mobile services place a greater focus on timeliness than personal computers (PCs), which necessitates quick processing and computation from service providers. As a result, it's important to increase the speed and accuracy of movie recommendations in mobile services.

The work of making movie recommendations is extensive and challenging since it incorporates different user preferences, different movie genres, and other factors. As a result, several strategies and recommendations have been put out to address the issues. For instance, content-based recommender systems, recommender systems with collaborative filtering, and hybrid recommender systems. Each method offers a benefit for tackling certain issues. Collaborative filtering is said to be the most common and commonly used recommender system approach given the prevalence of internet information and user-generated material. The collaborative filtering algorithm suggests products by comparing user similarities. A correlation calculation may be used to determine how similarly consumers like different options. Users that have a common taste in movies are grouped together in this fashion, and movies are then suggested based on their reviews and ratings of films they have watched. Due to the lack of essential user information, such as user ratings for viewed movies and surfing history, correlation and similarity calculations are challenging. Actually, user reviews on movies frequently include additional details like user preferences. Furthermore, a significant issue with movie recommendation is the users' ignorance of their sentiment. People are now more eager than ever to publish their own evaluations online. Users may share their opinions and preferences regarding movies in their reviews. Additionally, the opinions expressed in these reviews have an impact on other users' decisions. Users will read the evaluations, consider their own experiences, choose the most helpful ones, get rid of the false or even destructive ones, and then come to their own conclusions. Therefore, while judging a film, the tone of the reviews is crucial. In general, viewers are more likely to select the films that the majority favours and skip the ones that they find objectionable. To ensure consumers' personal comfort, choices are made based on the experiences of others.

2. LITERATURE SURVEY

The literature survey provides a full review of research articles related to recommendation systems and sentiment analysis. The articles listed below focus on different aspects of recommendation systems and their applications in various domains.

A.K.A. Hassan et al [1] suggests a location-based sentiment analyser for a hotel recommendation system. Based on the users' tastes and location, the system seeks to assist users in selecting hotels.

In order to create a hotel recommendation system that incorporates both review and context data, the Y.H. Hu et al [2] provides

DOI : https://doi.org/10.56452/6-3-669

a collaborative filtering-based methodology. The user's profile, location, and prior hotel experiences are all taken into consideration by the algorithm.

U.P. Ishanka et al [3] work suggests a user-emotion- and personality-based context-aware trip recommendation system. The system makes travel suggestions that are in line with the user's preferences by using sentiment analysis to comprehend the user's emotional state.

A deep learning-based method for classifying several emotions in tweets is presented by M. Jabreel et al [4]. The method classifies tweets' moods using a convolutional neural network (CNN) and a long short-term memory (LSTM) network.

Y. Wang et al [5] research suggests a hybrid recommender system for movies that is sentiment-enhanced. To increase the precision of suggestions, the system combines content-based, collaborative filtering approaches with sentiment analysis.

R.L. Rosa et al [6] proposes a knowledge-based recommendation system that incorporates sentiment analysis and deep learning. The system uses a knowledge graph to model the user's preferences and incorporates sentiment analysis to improve recommendation accuracy.

Y. Wang, et al [7] suggests a hybrid recommender system for movies that is sentiment-enhanced. To increase the precision of suggestions, the system combines content-based, collaborative filtering approaches with sentiment analysis.

In order to increase the precision of suggestions, Ziani, N. Azizi et al [8] study suggests a recommender system that makes use of sentiment analysis. To provide recommendations for things that meet the user's preferences, the system considers both the user's preferences and the emotion around the goods. The authors do tests utilizing real-world data from the social bookmarking website Delicious to show how well content-based recommendations work in social tagging systems. They demonstrate how their content-based method works better than a conventional collaborative filtering algorithm in terms of accuracy and coverage.Finally, the report wraps up by outlining the study's ramifications and the possibility of more research in this field. Future research, according to the authors, may look at how to combine content-based and collaborative filtering techniques to improve the efficacy of recommendations in social tagging systems.

A research method for the improvement of classification was proposed by Bhagat Chaitanya et al [9]. Their work aimed to provide a comprehensive evaluation of the latest updates in this field. The primary goal of this survey was to present a complete picture of how machine learning techniques were employed in Sentiment Analysis to achieve better results in brief details. Moreover, the basic emotion classification was examined, divided into ternary classes, i.e., positive, negative, and neutral, using different machine learning algorithms and further classified into their subclasses, i.e., love, happiness, fun, neutral, hate, sadness, and anger, utilizing the same machine learning algorithms.

The main topic of the study of Ivan Cantador et al [10] was the use of content-based methods for item recommendation in social tagging systems. The authors begin by outlining social tagging systems and how popular they have become recently as a way to categorize and distribute materials on the internet. The significance of suggestions in these systems is then emphasized, as they may aid users in finding relevant material and improve their surfing experience.

The study then covers the cold-start issue and sparsity-related constraints of conventional collaborative filtering methods in social tagging systems. As an alternative, the authors suggest utilizing content-based recommendation systems, which base suggestions on examining the substance of objects. The Cantador, Bellog, and Vallet research makes a significant addition to the area of recommendation systems, particularly when it comes to social tagging systems. It underscores the need for further study in this area and shows the promise of content-based recommendation approaches.

3. PROPOSED SYSTEM

The process of analysing, processing, summarizing, and reasoning an emotional text is known as sentiment analysis. The internet discussion of the emotional polarity analysis has tremendously advanced sentiment analysis. The accuracy of emotional polarity analysis based on online remark text is now improving progressively, however one issue with emotional analysis is the absence of in-depth study and application of sentiment analysis.

Following are the important steps that are carried out in movie recommendation process:

1. Data Gathering: Gathering a large dataset of movie reviews and ratings would be the initial stage in developing the system. The title of the film, the reviewer's rating, and the review text are all included in this dataset.
2. Sentiment analysis: Following the dataset's collection, methodologies for identifying the sentiment conveyed in each review will be applied to the reviews. Natural language processing (NLP) methods might be used in this research to ascertain if each review is favourable, negative, or neutral.
3. Feature Extraction: The next stage is to extract information from the review text that the algorithm for suggesting movies may employ. The tone of the review, the genre of the film, the actors or director involved, and any other pertinent information are a few examples of these aspects.
4. Deep Learning: After the characteristics were collected, a deep learning model, like a neural network, was trained to forecast which movies a certain viewer would find most interesting. To make sure the model is accurate and dependable, a portion of the obtained data would be used for training and a different subset for validation.
5. Recommendation Engine: Lastly, the trained model would be used to provide tailored movie suggestions for each user based on their watching habits, tastes, and other pertinent variables. In addition to other considerations like genre and preferred actors/directors, the recommendation engine would also take into account the tone of each review when generating its suggestions.

This suggested method might enhance the precision of suggestions and assist users in finding new films they are likely to appreciate by adding deep learning and sentiment analysis into movie recommendation systems.

The suggested system uses deep learning and sentiment analysis methods to improve movie recommendation systems. To increase the precision of movie suggestions, the system would make use of the enormous quantity of movie reviews and ratings data that is now accessible.

DOI : https://doi.org/10.56452/6-3-669

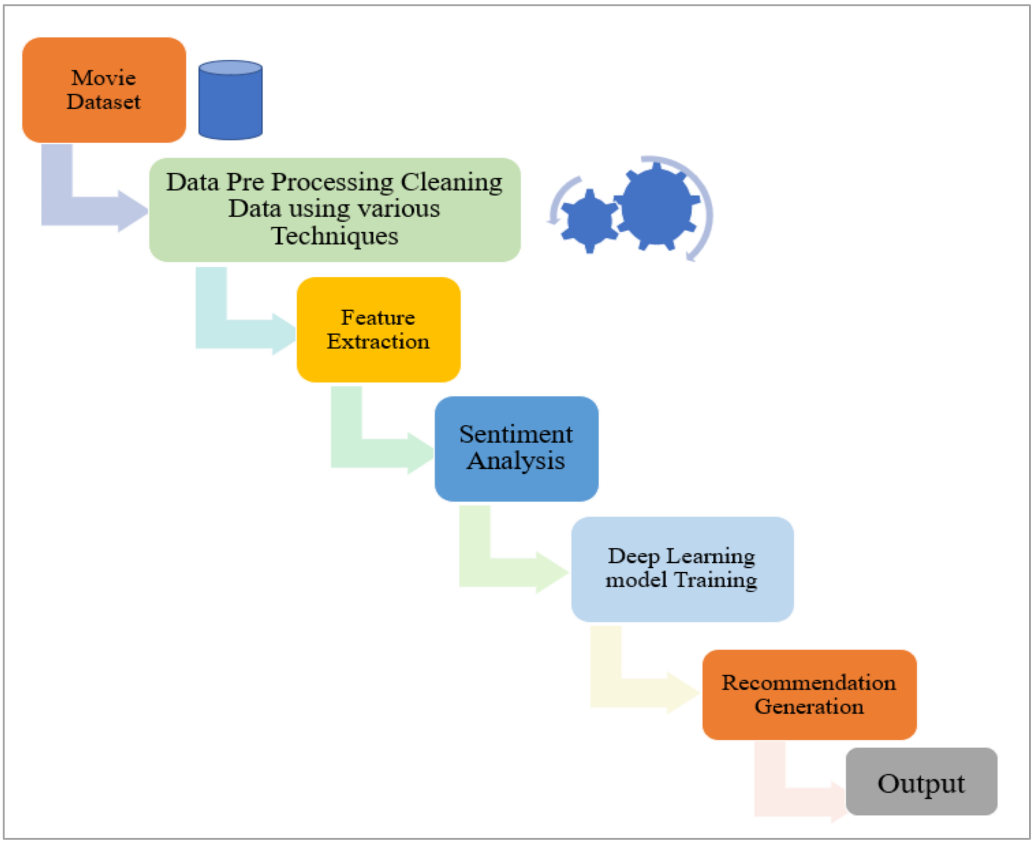

Figure 1. Building blocks architecture for Movie Recommendation Systems with Deep Learning and Sentiment Analysis

B. Algorithm : SVM

Two types of data points are shown in this diagram by the "X" symbols. Finding a decision boundary that divides these two classes is the SVM's objective. Equation w*x + b = 0, where w is a weight vector and b is a bias factor, represents the decision boundary.

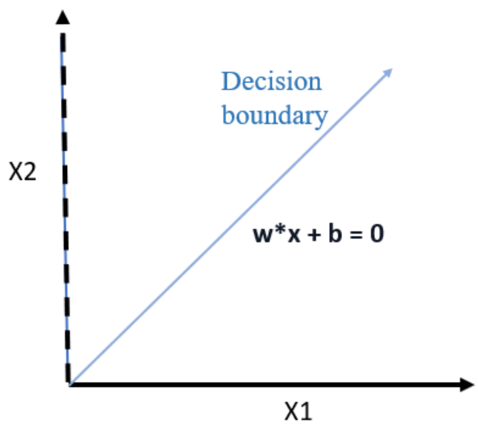

Figure 2. SVM Classification

Support vectors, which are the data points closest to the decision border, are crucial to the SVM method. As seen in the figure, the SVM aims to maximize the space between the decision border and the support vectors. The distance between the two dotted lines serves as a representation of this margin. By transforming the data into a higher-dimensional space with a linear decision boundary, a kernel function may be used to expand the SVM to handle non-linearly separable data.

Pseudo method for Sequential Minimal Optimization (SMO)-based Support Vector Machine (SVM) training:

1. Training set "(x1,y1)", "(x2,y2)",..., "(xm,ym)" as input, where "xi" denotes a feature vector and "yi" denotes the appropriate class label (+1 or -1).
2. Set the value of all i to 0.

3. Subject to the restriction that i!= j, choose two indices i and j at random.
4. For the chosen indices i and j, calculate the kernel function K(xi, xj).
5. Determine the upper and lower limits L and H for j while accounting for the restriction that 0 j C. C is a user-defined constant that regulates the trade-off between maximizing the margin and reducing the classification error.
6. If L = H, stop the loop and choose fresh values for i and j.
7. With all other variables remaining at their present levels, calculate the partial derivative of the Lagrange dual function with respect to j.
8. End the loop and choose fresh values for i and j if the partial derivative is 0.
9. Use the update rule to calculate the new value of j: j_new = j_old + (yj * (Ei - Ej) /), where = 2 * K(xi, xj) - K(xi, xj) - K(xj, xj) and Ei = f(xi), yi, and Ej, yj, where f(x) is the SVM decision function.
10. Use the formula j_new = max(L, min(j_new, H)) to clip the new value of j to fall inside the constraints L and H.
11. Use the following formula to update i: i_new = i_old + yi * yj * (j_old - j_new).
12. Select any support vector from 0 to i and compute the bias b as follows: b = yk - j j * yj * K(xj, xk), where k is the index of the chosen support vector.
13. Keep going through stages 3–12 until you reach convergence, which is the point at which the objective function stops significantly improving.

C. Algorithm: CNN

Convolutional Neural Networks (CNNs) are a class of deep learning algorithms that have shown remarkable success in various fields, such as image and speech recognition, natural language processing, and video analysis. In recent years, they have also been applied to movie recommendation systems, offering promising results in providing personalized recommendations. The CNN algorithm for movie recommendation systems leverages the rich content information available for movies, such as posters, movie trailers, and metadata, to create a more accurate and personalized recommendation. Here is an overview of how a CNN-based movie recommendation system works:

Feature Extraction: The first step involves extracting relevant features from the available content. For instance, the CNN can analyze movie posters or keyframes from movie trailers to capture visual patterns that represent the movie's genre, theme, and overall atmosphere. Additionally, textual data such as plot summaries, actor information, and director names can be processed using natural language processing techniques to obtain meaningful features.

Vector Representation: Once the features are extracted, they are represented as high-dimensional vectors. These vectors are then used as input to the CNN algorithm. The vector representation allows the model to capture complex patterns and relationships between the movies based on their content.

Convolutional Layers: The core of the CNN algorithm consists of multiple convolutional layers that process the input vectors. These layers apply filters, or convolutional kernels, that help detect specific patterns or features in the data. The process of convolution involves element-wise multiplication and summation, allowing the model to identify important features and relationships in the input data.

Pooling Layers: Following the convolutional layers, pooling layers are used to reduce the spatial dimensions of the data. This step helps in reducing the computational complexity of the model while preserving essential features.

DOI : https://doi.org/10.56452/6-3-669

Common pooling techniques include max pooling and average pooling.

Fully Connected Layers: After passing through the convolutional and pooling layers, the data is fed into one or more fully connected layers. These layers help in combining the detected features to produce a final feature vector that represents the movie. The fully connected layers are usually followed by activation functions like ReLU or sigmoid, which introduce non-linearity into the model.

Recommendation Generation: Finally, the CNN model generates recommendations by calculating the similarity between the feature vectors of different movies. The most common similarity metrics used are cosine similarity and Euclidean distance. The model recommends movies with the highest similarity scores to the target movie or user preferences.

4. RESULT AND ANALYSIS

A. Dataset Description

One of the most popular movie review datasets used for research and development purposes is the Large Movie Review Dataset, also known as the IMDb dataset. This dataset, introduced by Maas et al. (2011) in their paper "Learning Word Vectors for Sentiment Analysis," consists of 50,000 movie reviews from the Internet Movie Database (IMDb) and is evenly divided into positive and negative reviews. The dataset is split into two equal subsets: a training set with 25,000 reviews and a test set with 25,000 reviews. Each subset contains an equal number of positive and negative reviews, making it a balanced dataset. Reviews are labeled as positive if they have a rating of 7 or higher out of 10, and negative if they have a rating of 4 or lower.

	review	sentiment
0	One of the other reviewers has mentioned that ...	positive
1	A wonderful little production. The...	positive
2	I thought this was a wonderful way to spend ti...	positive
3	Basically there's a family where a little boy ...	negative
4	Petter Mattei's "Love in the Time of Money" is...	positive

Figure 3. Dataset Description

B. Graphical Analysis

Graphical analysis using movie dataset we can understand how movies popularity on the basis of Audience score and rating % from Rotten Tomatoes. In below tree map chart, we can see Darkest teal colour represents that particular movie title have got higher % or Audience score as well as high success rating % in Rotten Tomatoes site with respect to year. Which can help us to understand type of movies user loved to watch according to years. And hence we can use this analysis in semantics SVM algorithm to understand user nature.

DOI : https://doi.org/10.56452/6-3-669

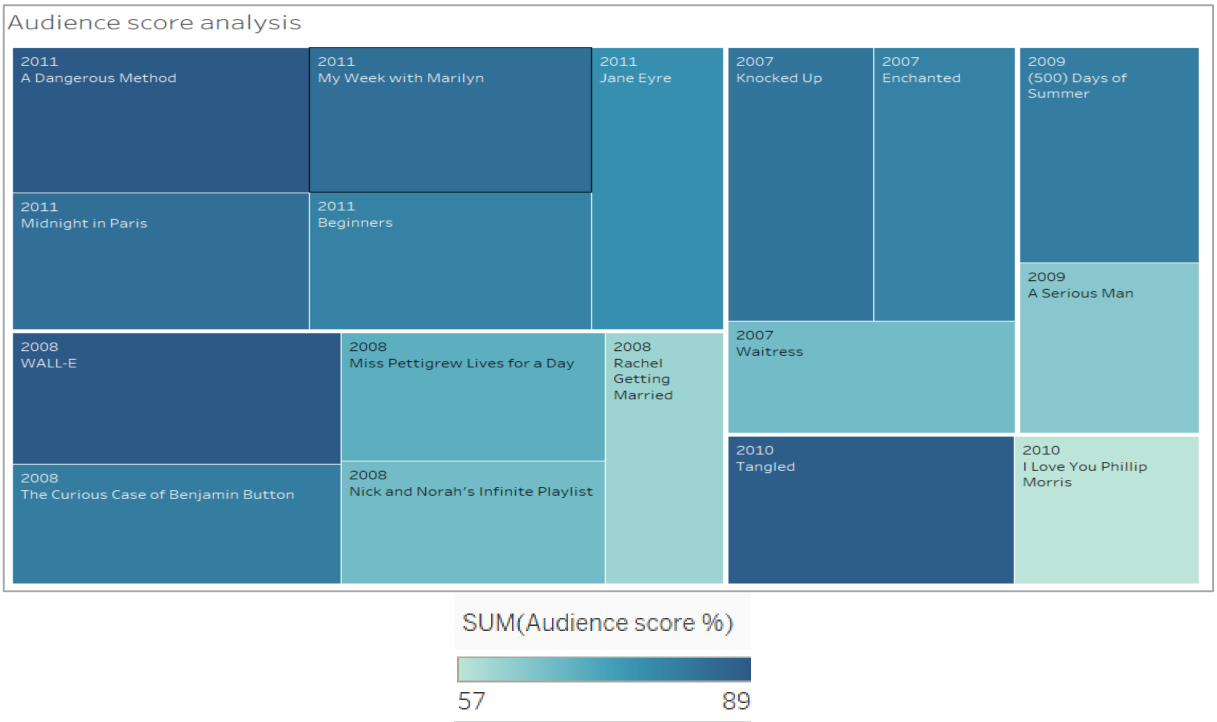

Figure 4. Audience Score Analysis

C. *Result Analysis*

Results indicate that the CNN algorithm outperforms the SVM algorithm in terms of accuracy and F1-score. The superior performance of the CNN model can be attributed to its ability to learn complex patterns and relationships in the data, thanks to its deep learning architecture. The convolutional layers in the CNN model can capture intricate features from the movie content, such as posters and trailers, which contribute to a more accurate recommendation.

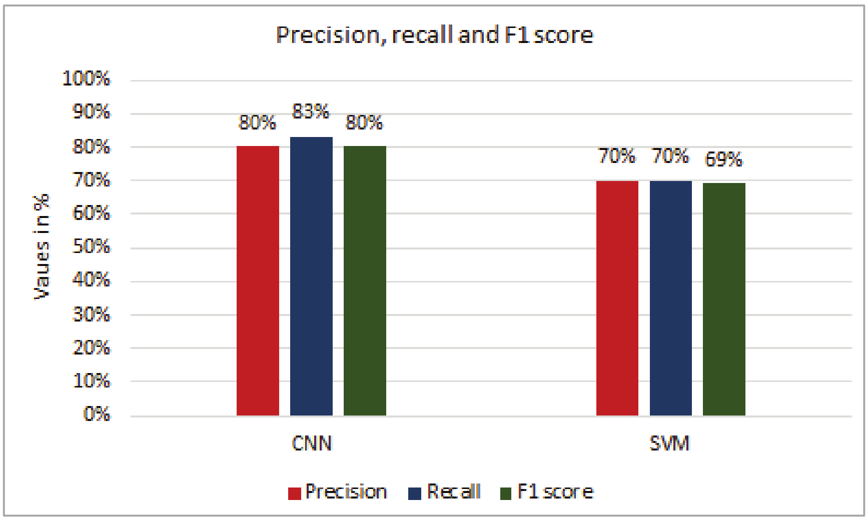

Figure 4. Performance Comparison Graph

5. CONCLUSION

With the use of sentiment analysis and deep learning approaches movie recommendation systems improves greatly. Recommendation systems provide more precise and relevant movie suggestions by analysing user ratings, genres, actors, and attitudes, thereby improving the user experience. Analysing user reviews and comments using natural language

DOI : https://doi.org/10.56452/6-3-669

processing (NLP) tools is an alternative strategy. The sentiment, emotions, and themes covered in the review may all be extracted using NLP approaches. A recommendation engine that considers the user's emotional reaction to films may then be trained using these attributes. This proposed work is mainly split into two halves. One of them focuses on the sentiment analysis, while the other one is a movie recommendation system. The paper thoroughly examines both systems and draws some significant findings. The SVM and CNN algorithm has been utilized for the Movie Recommendation System to suggest the best films that are relevant to the movie the user submitted based on several parameters such as the genre of the movie, overview, the cast, and the ratings provided to the movie. Even after multiple testing, CNN classifier has produced respectable findings and has been pretty reliable in terms of suggesting the films.

REFERENCES

[1] A.K.A. Hassan and A.B.A. Abdulwahhab, Location aspect based sentiment analyzer for hotel recommender system, Iraqi J. Sci. 60 (2019), no. 1, 143–156.

[2] Y.H. Hu, P.J. Lee, K. Chen, J.M. Tarn and D.V. Dang, Hotel recommendation system based on review and context information: A collaborative filtering appro, PACIS Proc., 2016.

[3] U.P. Ishanka, A context-aware travel recommendation system based on user emotion and personality, Doctoral dissertation, Nagaoka University of Technology, 2018.

[4] M. Jabreel and A. Moreno, A deep learning-based approach for multi-label emotion classification in tweets, Appl. Sci. 9 (2019), no. 6, 1123.

[5] Y. Wang, M. Wang and W. Xu, A sentiment-enhanced hybrid recommender system for movie recommendation: A big data analytics framework, Wireless Commun. Mobile Comput. 2018 (2018).

[6] R.L. Rosa, G.M. Schwartz, W.V. Ruggiero and D.Z. Rodr´ıguez, A knowledge-based recommendation system that includes sentiment analysis and deep learning, IEEE Trans. Ind. Inf. 15 (2018), no. 4, 2124–2135.

[7] Y. Wang, M. Wang and W. Xu, A sentiment-enhanced hybrid recommender system for movie recommendation: A big data analytics framework, Wireless Commun. Mobile Comput. 2018 (2018).

[8] A. Ziani, N. Azizi, D. Schwab, M. Aldwairi, N. Chekkai, D. Zenakhra and S. Cheriguene, Recommender system through sentiment analysis, 2nd Int. Conf. Automatic Control Telecommun. Signals, 2017.

[9] Bhagat, Chaitanya & Mane, Deepak. (2019). Survey On Text Categorization Using Sentiment Analysis. International Journal of Scientific & Technology Research. 8. 1189-1195.

[10] Ivan Cantador, Alejandro Bellog ´ ´ın, and David Vallet. Content-based recommendation in social tagging systems. In Proceedings of the Fourth Conference on Recommender systems, pages 237– 240. ACM, 2010.

[11] Graphical Analysis Dataset: https://gist.github.com/tiangechen/b68782efa49a16edaf07dc2cdaa855ea

[12] Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. (2011). Learning Word Vectors for Sentiment Analysis. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (pp. 142-150).