IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.-7, Issue-4, April 2019 E-ISSN: 2347-2693

Movie Recommendation System

Kunal Raj[1*], Atulya Abhinav Das[1], Antariksh Guha[1], Parth Sharma[1], Mohana Kumar S[1]

[1]Dept. of Computer Science and Engineering Ramaiah Institute of Technology, Bangalore, India

Corresponding Author: kunalraj1414@gmail.com , Tel.: +91-9071370746

DOI: https://doi.org/10.26438/ijcse/v7i4.10241028 | Available online at: www.ijcseonline.org

Accepted: 18/Apr/2019, Published: 30/Apr/2019

***Abstract*—** In this hustling world, enjoyment is a need for every one of us to refresh our temper and energy. Entertainment regains our self-assurance for work and we can work extra enthusiastically. For revitalizing ourselves, we are able to pay attention to our preferred music or can watch films of our preference. For looking favorable movies online, we will make use of movie recommendation systems, that are extra dependable, when you consider that searching of preferred films would require more and more time which one can ‘t have the funds to waste. In this paper, to improve the quality of a movie recommendation system, a deep learning-based approach is presented to find out what exactly was being talked about in the user's review and the sentiments that people are expressing.

***Keywords*—**Deep Learning , Recommendation System, Review Sentiment Analysis

I. INTRODUCTION

Every time you go about in the web, you most likely go to imdb.com now and again to search for a movie review. Data mining is an interdisciplinary subfield of computer engineering which has been used in this project to accomplice feature wise movie ranking for better recommendation. It is the computational procedure of finding designs in data sets including strategies at the crossing point of man-made reasoning, AI, insights, and database frameworks. The general objective of the data mining process is to extricate data from a data set and change it into a justifiable structure for further use. Beside the crude examination step, it includes database and data the board viewpoints, data pre-preparing, model and induction contemplations, intriguing quality measurements, multifaceted nature contemplations, post-handling of found structures, perception, and web-based refreshing.

This Movie Recommendation System project is created on Java platform and it also uses concepts of Deep Learning such as RNN. It’s a prototype which works on user reviews posted on IMDB and it returns positiveness and negativeness based on certain features of the movie such as direction and production.

II. RELATED WORK

In the paper [2] titled "*Deep Learning for Recommender Systems*," the writers portray that With the consistently developing volume of online information, recommender systems have been an intelligent procedure to beat such information over-burden utility of recommender systems can't be exaggerated, given its across the board appropriation in many web applications, alongside its potential effect to enhance numerous issues identified with over-decision. Lately, deep learning has gathered significant enthusiasm for some exploration fields, for example, PC vision and characteristic language preparing, owing not exclusively to excellent performance deep learning is likewise inescapable, as of late exhibiting its adequacy when connected to information recovery and recommender systems investigate. Clearly, deep learning in recommender framework is feeding. This article intends to give a far reaching audit of ongoing exploration on deep learning based recommender systems. All the more solidly, we give and devise scientific classification of deep learning based suggestion models, alongside giving a complete rundown of the cutting edge.

In the paper [3] titled “*Deep Neural Networks for YouTube Recommendations*” the authors describe that YouTube represents one of the largest scale and most sophisticated industrial recommendation systems in existence. In this work, the authors describe the system at a high level and focus on the dramatic performance improvements brought by deep learning. The calculation is part as indicated by the great two-stage information retrieval dichotomy: first, we detail a deep competitor age model and afterward depict a different deep ranking model.

In the paper [4] titled "*Neural Collaborative Filtering*" the creators portray that profound learning for recommendation,

they primarily used it to model auxiliary information, for example, textual descriptions of things and acoustic highlights of music. With regards to model the key factor in collaborative filtering the interaction among client and things highlights, despite everything they turned to framework factorization and connected an internal item on the idle highlights of clients and things. By supplanting the internal item with a neural engineering that can take in a subjective capacity from information, we present a general system named NCF, short for Neural network based Collaborative Filtering. NCF is conventional and can express and sum up network factorization under its system.

In the paper [5] titled "*Matrix Factorization Techniques for Recommender Systems*" the creators portray that modern consumers are immersed with decisions. Electronic retailers and substance suppliers offer an immense determination of items, with uncommon chances to meet an assortment of unique needs and tastes. Coordinating consumers with the most suitable items is critical to improving client fulfillment and reliability. Therefore, more retailers have turned out to be keen on recommender systems, which examine examples of client enthusiasm for items to give customized proposals that suit a client's taste. Since great customized suggestions can add another measurement to the client experience, web based business pioneers like Amazon.com and Netflix have made recommender systems a striking piece of their sites.

In the paper [6] titled "*Distributed representations of words and phrases and their compositionality*" the creators portray that The as of late presented constant Skip-gram model is an effective technique for adapting high caliber disseminated vector portrayals that catch an extensive number of exact syntactic and semantic word connections. In this paper we present a few expansions that improve both the nature of the vectors and the preparation speed. By subsampling of the continuous words we get huge speedup and furthermore adapt progressively ordinary word portrayals. We additionally depict a basic option in contrast to the various leveled softmax called negative inspecting.

In the paper [7] titled "*Collaborative Deep Learning for Recommender Systems*" the makers depict that Collaborative filtering (CF) is a productive approach commonly used by various recommender systems. Standard CF-based methods use the assessments given to things by customers as the sole wellspring of information for making sense of how to make proposal. In any case, the examinations are every now and again particularly small in various applications, causing CF-based methodologies to degenerate in a general sense in their recommendation execution. To address this sparsity issue, right hand information, for instance, thing content information may be utilized. Network arranged subject backslide (CTR) is a drawing in progressing system embracing this methodology which immovably couples the two sections that gain from two interesting wellsprings of information. Coincidentally, the dormant depiction learned by CTR may not be very convincing when the associate information is lacking. To address this issue, we entirety up continuous advances in significant picking up from i.i.d. commitment to non-i.i.d. (CF-based) input and propose in this paper a different leveled Bayesian model called network arranged significant learning (CDL), which together performs significant depiction learning for the substance information and synergistic isolating for the evaluations (analysis) cross section. Expansive preliminaries on three authentic world datasets from different zones exhibit that CDL would altogether be able to impel the bleeding edge

In the paper [8] titled "*Collaborative Knowledge Base Embedding for recommender Systems*" the creators depict that among various expectation strategies, cooperative separating for the most part experiences restricted execution because of the sparsity of client thing collaborations. To address the issues, helper data is generally used to support the execution.
Because of the fast gathering of data on the web, the learning base gives heterogeneous data including both organized and unstructured information with various semantics, which can be consumed by various applications a heterogeneous network embedding method, termed as Trans, to extract items' structural representations by considering the heterogeneity of both nodes and relationships.

III. METHODOLOGY

This overview of the project gives a concise depiction about the possibility of the project and the execution did. We principally portray System Architecture.

This is an essential advance in any sort of project. It is the initial step that needs to be done, it gives an earlier thought or clear picture about the project will work. Arranging a legitimate Implementation and working as needs be will be key advance for an achievement of project.

A. Review Collection using Web Crawler

This module is responsible for collecting the reviews from the IMDB website by making use of JSoup API and the reviews are stored in the format of {ReviewId, ReviewDesc, MovieName}

The Web Crawler based Online Review Submission will have the input elements like Movie, Web URL and Xpath. The Web URL and Xpath will be validated. If valid then a real HTTP connection is made to the web site and the DOM elements are scanned for text (review) and extracted. The Web Crawler based Online Review Collection can be described in Figure 1

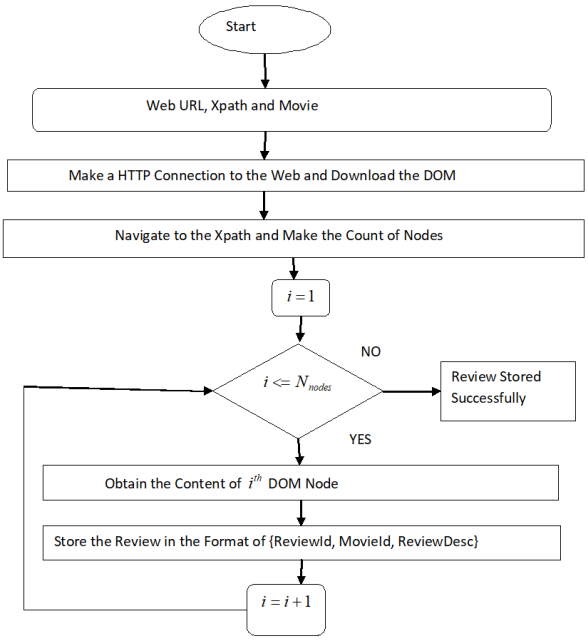

Figure 1 : Web Crawler Based Review Collection

B. Sentiment Analysis per Review and per Feature

The reviews are collected. After that the List of Reviews are scanned and sentiment analysis is performed per feature. The list of features is configurable.

The flowchart for Sentiment Analysis per Review and Per Feature can be described as Figure 2

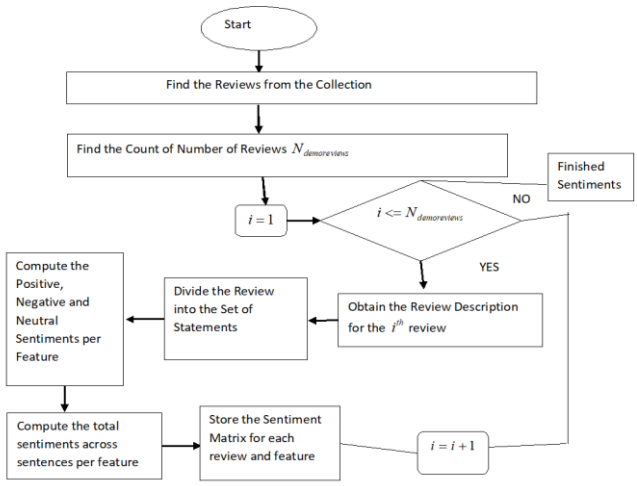

Figure 2 : Review & Feature Based Sentiment Analysis

C. SENTIMENT ANALYSIS PER MOVIE AND PER FEATURE

This Module is responsible for Sentiment Analysis per Movie and Feature. After the review-based sentiments are computed then the following equations are used, and movie-based sentiments are computed and a final matrix is obtained which is movie based sentiments per feature. The process involves finding the unique set of movies for which reviews are collected. The following is computed for each movie. Find the reviews for the movie. For each of the review per feature the following is computed.

The positive sentiment for the movie is computed using the following eq (1):

$$PS_{movie,f_i} = \sum_{i=1}^{N_{reviews}} PS_i \qquad ---Eq\,(1)$$

$$Where,$$

$$N_{reviews} = Number\ of\ reviews$$

$$PS_i = Positive\ sentiment\ of\ i^{th}\ review$$

$$f_j = j^{th}\ feature$$

$$0 \leq f \leq N_f - 1$$

$$N_f = Number\ of\ features$$

The negative sentiment for the movie is computed using the following eq (2):

$$NS_{movie,f_i} = \sum_{i=1}^{N_{reviews}} NS_i \qquad ---Eq\,(2)$$

$$Where,$$

$$N_{reviews} = Number\ of\ reviews$$

$$NS_i = Negative\ sentiment\ of\ i^{th}\ review$$

$$f_j = j^{th}\ feature$$

$$0 \leq f \leq N_f - 1$$

$$N_f = Number\ of\ features$$

The neutral sentiment is computed using the following eq (3):

$$NU = \sum_{i=1}^{N_{reviews}} NU_i \qquad ---Eq\,(3)$$

$$Where,$$

$$N_{reviews} = Number\ of\ reviews$$

$$NU_i = Neutral\ sentiment\ of\ i^{th}\ review$$

$$f_j = j^{th}\ feature$$

$$0 \leq j \leq N_f$$

$$N_f = Number\ of\ features$$

The flowchart for movie-based sentiment analysis can be described as be

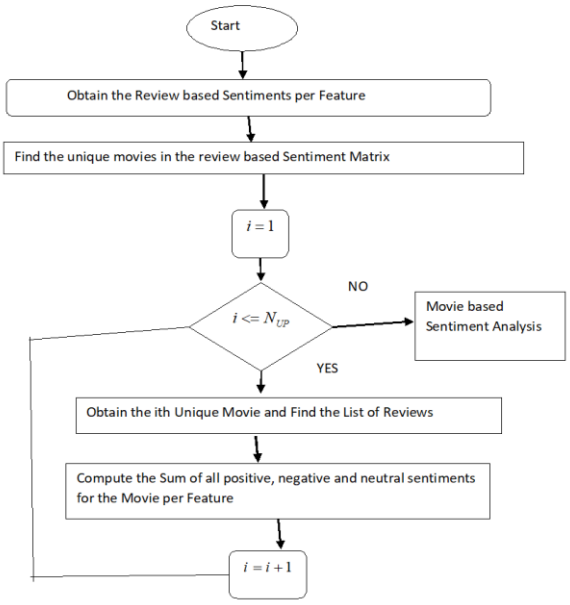

Figure 3 : Movie Based Sentiment Analysis Per Feature

A. Query Based Sentiment Analysis and Recommendations

Query Based Sentiment Analysis is responsible for analysing the searched query for the following cases which have the following use cases

[1] The user searches for a Single Feature. ex- direction. The movies are ranked based on feature that is present in the user query based on the sentiment computation

[2] The user searches for Multiple Feature then a combination feature is computed, and movies are ranked. ex- direction and songs

The flowchart for the Query based Sentiment Analysis can be described in the Figure 4

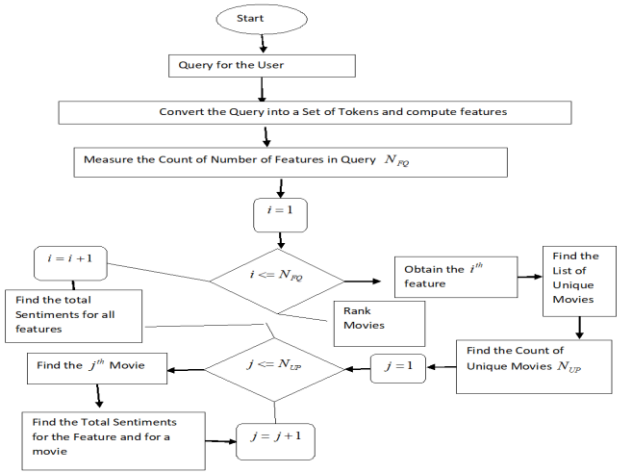

Figure 4 : Recommendations of Movie Based on Feature Sentiments

Figure 4 shows the steps required for ranking of movies based on Features

1. The query is entered by the user
2. The entire query is divided into multiple tokens by using the Delimiter
3. The Number of Unique Features is found out in the query
4. Start from the 1st feature till the number of features
    a. Obtain the Feature
    b. Find the Unique Movies in the algorithm
    c. Obtain the Sentiments of the Feature for each movie
5. Find the sentiments of all Features across Movies
6. Rank the Movies based on Maximum Positive Sentiments, Least Negative Sentiments and Maximum Neutral Sentiments

IV. CONCLUSION

Recommender systems are a powerful new technology for extracting extra value for a business from its user databases. These systems help users discover items they need to purchase from a business. Recommender systems benefit users by enabling them to discover items they like. Recommender systems are quickly becoming a significant instrument in E-commerce on the Web. Movie Recommendation System is used to discover the likeness of the movie based on different features by viewers.

In this paper we presented and experimentally evaluated a algorithm for feature based movie recommender systems. This approach helps us identify the movie according to the features like director, actor, music etc i.e. whether it is good or bad in those features according to the reviews given by viewers and not just ranking the movies according to the rating. So, each feature will have a different movies ranking.

REFERENCES

[1] S. M. Taheri and I. Irajian, "DeepMovRS : A unified framework for deep learning-based movie recommender Systems," 2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS), Kerman, 2018, pp. 200-204. DOI: 10.1109/CFIS.2018.8336633

[2] H.-T. Cheng, L. Koc, J. Harmsen, T. Shaked, T. Chandra, H.Aradhye, G.Anderson, G. Corrado, W. Chai, M. Ispir, R Anil, Z. Haque, L. Hong,V. Jain, X. Liu and H. Shah, Wide & Deep Learning for Recommender Systems, In Proceedings of the 1st Workshop on Deep Learning for Recommender Systems (DLRS 2016), (2016). ACM, New York, NY, USA, 7-10.

[3] P. Covington, J. Adams and E. Sargin, Deep Neural Networks for YouTube Recommendations, In Proceedings of the 10th ACM Conference on Recommender Systems, (2016). ACM, New York, NY, USA, 191-198.

[4] X. He, L. Liao, H. Zhang, L. Nie, X. Hu and T.-S. Chua, Neural Collaborative Filtering, In Proceedings of the 26th International

Conference on World Wide Web (WWW '17), (2017). International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 173-182.

[5] Y. Koren, R. Bell and C. Volinsky, Matrix Factorization Techniques for Recommender Systems, Computer, 42, 8 (August 2009), 30-37.

[6]Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean and G. Hinton, Distributed Representations of Words and Phrases and their Compositionality,NIPS'13 Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2.

[7] H. Wang, N. Wang and D.-Y. Yeung, Collaborative Deep Learning for Recommender Systems, In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '15), (2015). ACM, New York, NY, USA, 1235-1244.

[8] F. Zhang, N. J. Yuan, D. Lian, X. Xie and W.-Y. Ma, Collaborative Knowledge Base Embedding for Recommender Systems, In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), (2016). ACM, New York, NY, USA, 353-362.

Authors Profile

Mr. Kunal Raj is pursuing Bachelor of Engineering in the field of Computer Science from Ramaiah Institute of Technology, Bengaluru, India.

Mr. Atulya Abhinav Das is pursuing Bachelor of Engineering in the field of Computer Science from Ramaiah Institute of Technology, Bengaluru, India.

Mr. Antariksh Guha is pursuing Bachelor of Engineering in the field of Computer Science from Ramaiah Institute of Technology, Bengaluru, India.

Mr. Parth Sharma is pursuing Bachelor of Engineering in the field of Computer Science from Ramaiah Institute of Technology, Bengaluru, India.

Dr. *Mohana Kumar S* is working as an Associate Professor in the department of Computer Science & Engineering, Ramaiah Institute of Technology.