iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** XI **Month of publication:** November 2021

DOI: https://doi.org/10.22214/ijraset.2021.38844

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue XI Nov 2021- Available at www.ijraset.com

A Comparing Collaborative Filtering and Hybrid Recommender System for E-Commerce

Dr. C. K. Gomathy [1], Thumu Swecha[2], Thondapu Vandana[3]
[1]*Assistant Professor, Dept. Of CSE, SCSVMV (Deemed to be University), Kanchipuram, TamilNadu, India*
[2, 3]*UG Scholar, Student, Dept. Of CSE, SCSVMV (Deemed to be University), Kanchipuram, TamilNadu, India*

Abstract: Here we are building an collaborative filtering matrix factorization based hybrid recommender system to recommend movies to users based on the sentiment generated from twitter tweets and other vectors generated by the user in their previous activities. To calculate sentiment data has been collected from twitter using developer APIs and scrapping techniques later these are cleaned, stemming, lemetized and generated sentiment values. These values are merged with the movie data taken and create the main data frame.The traditional approaches like collaborative filtering and content-based filtering have limitations like it requires previous user activities for performing recommendations. To reduce this dependency hybrid is used which combines both collaborative and content based filtering techniques with the sentiment generated above.
Keywords: machine learning, natural language processing, movie lens data, root mean square equation, matrix factorization, recommenders system, sentiment analysis.

I. INTRODUCTION

In the digital era understanding and creating a digital profile of a customer plays a vital role in increase of sales and attraction of more clients. But, doing it manually for millions of customers is a tedious and error prone task. Here comes machine learning and data science which ease this process and give better results. Recommendation of the best product to customers at amazon or given best movie suggestion to the customer is one of the key successes to their domain. To do this machine learning has already curated models. To do this customer past activities plays an vital role like purchases happened in the past and movies watched are the key data inputs for the model. A vector of the latest events is created for an user to represent his basic profile. Matrix factorization is one the popular techniques used to project the users and items into the shared latent space, using a vector of latent features to represent a user or an item. So the interaction of a product and user is calculated as the inner product of their latent vectors.
A competition led by Netflix has uplifted the matrix factorization and most of the research is happening across the same technique. Despite the effectiveness of matrix factorisation for collaborative filtering it is found that its performance can be enhanced using the choice of interaction function and inner product. While it seems to be a relatively small change but the positive effect it provides of designing a better, dedicated interaction function for modelling the latent feature interactions between users and items

II. LITERARY SURVEY

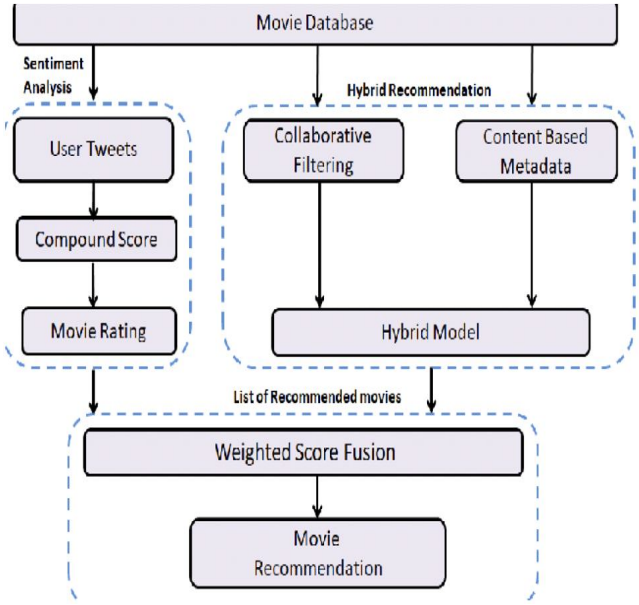
Below is the summary of the papers we had studied and the used references. There are two variations with this technique, user to user collaborative filtering and item to item collaborative filtering. It is a combination of both content based filtering and collaborative filtering and enjoys the advantages of both. In item to item collaborative filtering, the system establishes relationship among items via ratings Issue with User-User collaborative filtering is sparsity, with large item set, small number of ratings too often there are points where no recommendation can be made for a user. .Many recommendation systems suggest items to users by utilizing the techniques of collaborative filtering (CF) based on historical records of items that the users have viewed, purchased, or rated. Thomas Austin Department of Computer Science Recommendation System Using Collaborative Filtering by Yunkyoung Lee Collaborative filtering is one of the well known and most extensive techniques in the recommendation system; its basic idea is to predict which items a user would be interested in based on their preferences.
This project is mainly based to compare the results and metrics between collaborative filtering based recommender systems. The main goal to build this project is to build a collaborative filter recommender system with the use of matrix factorization technique and to build a hybrid recommender using the neural embedding layer from the PyTorch package.

©IJRASET: All Rights are Reserved

III. PROJECT ILLUSTRATION

In this project the main goal is to build collaborative filtering, matrix factorisation and hybrid recommender system using the tweets from twitter based on movie names and recommending them to users. Firstly movie data taken from file is changed to dataframe, and tweets related to movie hashtag has been retrieved using twitter API and scrapping. There after collaborative filter based recommender system has built using the matrix factorisation, Here, we have 2 input file

1) *Input Items:* This is the movie data which has features like movie name, release date, and IMDB Url to check its rating and a one-hot encoded features like unknown, Action, adventure, animation, children, comedy, crime, documentary, drama, fantasy, film-noir, horror, musical, mystery, romance, sci-fi, thriller, war, western. Sentiment extraction from tweets is one of the crucial steps and this is done using the Text-blob package and here are the results.



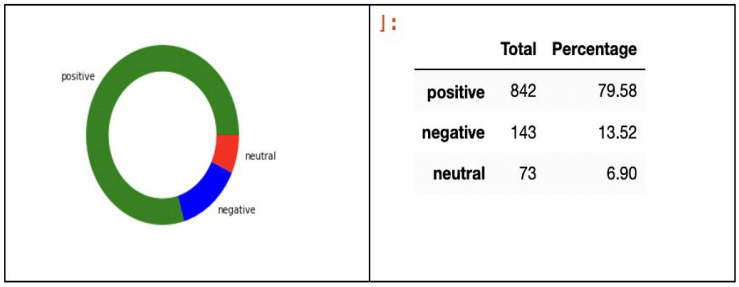
Twitter API and credentials from developer console were used to extract the tweets from twitter using the movie names taken from above datafile as hashtags. Later these tweets are loaded into csv from there to the data frame.

Using Nltk and vader sentiment analysis twitter data has been cleaned by deleting words less than 3 characters, extra spaces, duplicates, punctuation, stemming, lemmatization and calculation of positive, negative and neutral score. In the next step these tweets or comments have been mapped to the movie names retrieved from the above datafile.

Movie data metrics

Movies	1682
Unique users	1642
Ratings	72954

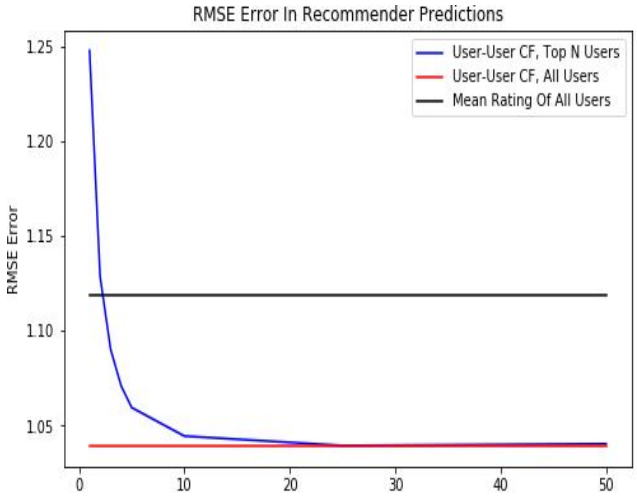
Sentiment distribution chart for the collected tweets



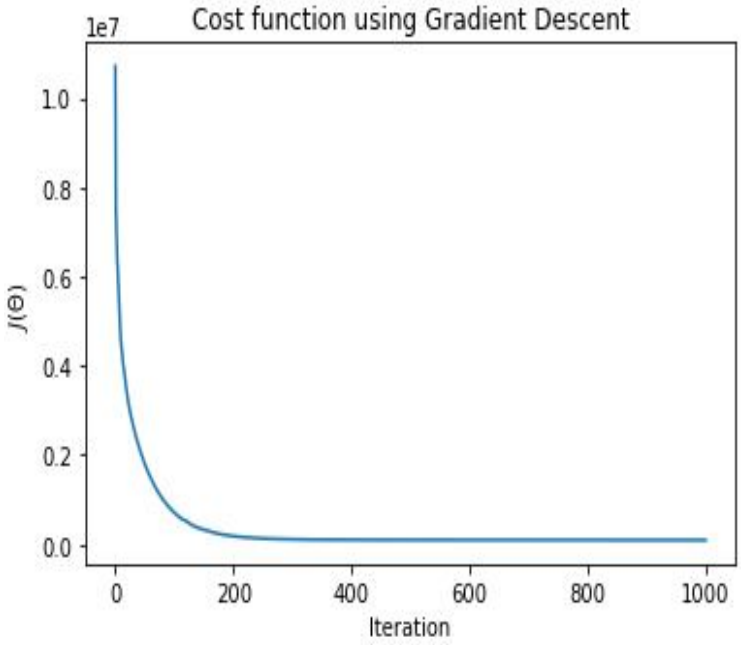
©IJRASET: All Rights are Reserved

IV. RESULTS AND ANALYSIS

We got the following interpretations upon RMSE, precision, Recall, and F1 scores. Our model scored RMSE value of 0.4315 in training and 1.2271 in testing phase.



The below curve informs that the Collaborative filtering is indeed working and with each iteration of Gradient descent the cost function is going down and plateaued out around 200 and from there onwards the change is constant.



V. LIST OF ABBREVIATIONS

ML	Machine learning
NLP	Natural language processing
RMSE	Root mean square equation
PYTHON	python
SA	Sentiment analysis

©IJRASET: All Rights are Reserved

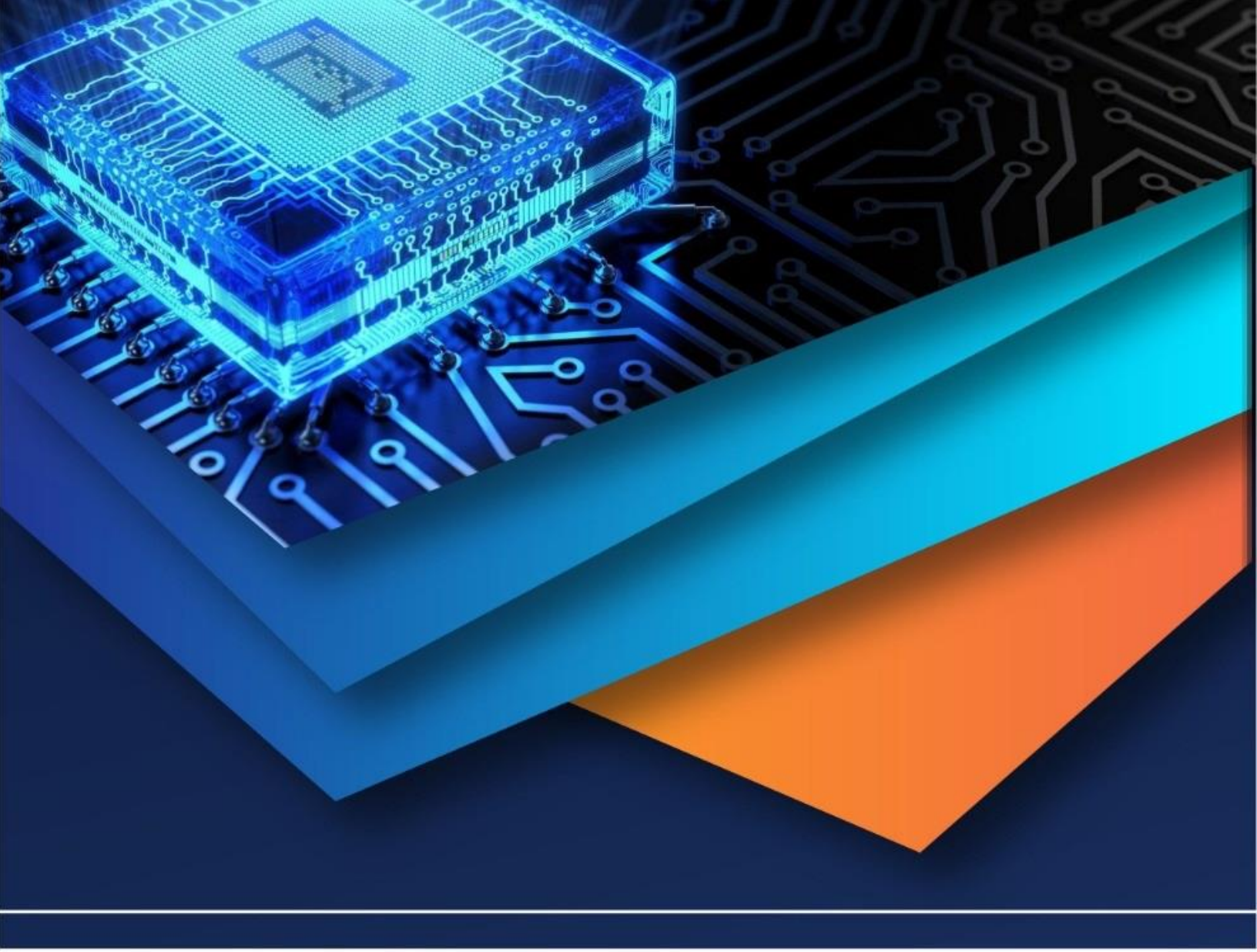
VI. CONCLUSION

As we know recommender systems are playing an vital role in many industries like ecommerce, movie publisher sites etc and used to filter and categories' information and customers to produce higher profit and returns. In this project we used the sentiment data from twitter alongside the user's past activities to recommend movies. Sentiment analysis is to provide data about the user in reacting to a particular movie globally. So this project was built using hybrid methods to provide the accurate result possible even for the user not having past activities. As per our understanding hybrid models work precisely and provide accurate results and this helps increase the domain not limited to movies etc.

REFERENCES

[1] Thomas Tran and Robin Cohen . Hybrid Recommender Systems for Electronic Commerce. AAAI Technical Report WS-00-04. Compilation copyright © 2000, AAAI (www.aaai.org).
[2] 8.Edjalma Queirozda SilvaCelso G.Camilo-JuniorLuiz Mario L.PascoalThierson C.Rosa, An evolutionary approach for combining results of recommender systems techniques based on collaborative filterin
[3] Dr.C K Gomathy, Article: An Effective Innovation Technology In Enhancing Teaching And Learning Of Knowledge Using Ict Methods, International Journal Of Contemporary Research In Computer Science And Technology (Ijcrcst) E-Issn: 2395-5325 Volume3, Issue 4,P.No-10-13, April '2017
[4] Dr.C K Gomathy, Article: A Semantic Quality of Web Service Information Retrieval Techniques Using Bin Rank, International Journal of Scientific Research in Computer Science Engineering and Information Technology (IJSRCSEIT) Volume 3 | Issue 1 | ISSN : 2456-3307, P.No:1563-1578, February-201
[5] Dr.C K Gomathy, Article: A Web Based Platform Comparison by an Exploratory Experiment Searching For Emergent Platform Properties, IAETSD Journal For Advanced Research In Applied Sciences, Volume 5, Issue 3, P.No-213-220, ISSN NO: 2394-8442,Mar/2018
[6] Dr.C K Gomathy, Article: Supply chain-Impact of importance and Technology in Software Release Management, International Journal of Scientific Research in Computer Science Engineering and Information Technology (IJSRCSEIT) Volume 3 | Issue 6 | ISSN : 2456-3307, P.No:1-4, July-2018
[7] Dr.C K Gomathy, Article: A Scheme of ADHOC Communication using Mobile Device Networks, International Journal of Emerging technologies and Innovative Research (JETIR) Volume 5 | Issue 11 | ISSN : 2349-5162, P.No:320-326, Nov-2018
[8] Dr.C K Gomathy, Article: A Study on the recent Advancements in Online Surveying , International Journal of Emerging technologies and Innovative Research (JETIR) Volume 5 | Issue 11 | ISSN : 2349-5162, P.No:327-331, Nov-201.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)