Emotion based Customer Behaviour prediction using Machine learning

Priyank Sirohi
Ph.D. Research Scholar, Shobhit Institute of Engineering and Technology (Deemed-to-be University), Meerut, India
and Assistant Professor, Sir Chhotu Ram Institute of Engineering and Technology, C.C.S. University, Meerut, India Email: priyanksirohi01@gmail.com
Niraj Singhal
Director, Sir Chhotu Ram Institute of Engineering and Technology, Chaudhary Charan Singh University, Meerut, India
Email: drnirajsinghal@gmail.com
Pradeep Kumar
Department of Computer Science and Engineering, JSS Academy of Technical Education Noida, Uttar Pradesh, India
Email: pradeep8984@jssaten.ac.in
Mahboob Alam
Department of Computer Science and Engineering, JSS Academy of Technical Education Noida, Uttar Pradesh, India
Email: mahboobalam@jssaten.ac.in

Abstract- The overwhelming volume of knowledge accessible on the Internet makes it impossible to efficiently retrieve information from it. Recommender mechanisms filter through vast quantities of material to provide consumers with specific knowledge. Typically, a recommender framework offering efficient and relevant feedback draws the attention of consumers or clients. There are some holes in the literature that need to be filled. Work has been done on emotion-based recommendation systems, but there are still several accessible areas to explore. This paper investigates the user rating data for recommendation of product using collaborating filtering. The ultimate aim is to explore the recommender method in wider-scaled application. This paper uses the user rating of social media user rating data for recommendation of Emotion-Based RS’s Architecture which is user collaborating filtering techniques. The outcome as product recommendation has been come out in final layout. This investigation proposed the Emotion-Based RS’s Architecture integrated with collaborating filtering techniques.

Keywords: Recommender System, Recommender mechanisms, Emotion-Based RS’s Architecture, User's Emotional State, collaborating filtering
DOI Number: 10.48047/nq.2022.20.19.NQ99117 **NeuroQuantology2022;20(19): 1281-1298**

I. INTRODUCTION

Recommendation mechanisms direct the customer in an ad hoc fashion in the sense of electronic commerce to fascinating or beneficial items in a broad variety of available choices. Recommendation systems must correctly document client expectations and desires in a user profile to produce a credible recommendation. However, consumer sentiment plays a remarkably key position in the decision-making phase for personal and nuanced items such as film, music, and news. Since the standard user profile model does not take the effect of user feelings into consideration, it is difficult for suggested systems to consider or catch shifting user desires. In an effort to boost the effectiveness of recommendation systems, researchers have started to switch to much more user - generated content descriptors in recent years. The management of feelings and personality as drugs that reflect a bigger fraction of the variety of customer demands than formerly used standardised descriptions has become possible thanks to advancements in affective computing, particularly in automated emotion recognition systems (such as gender). By identifying a compact collection of objects from a broad community, suggestion systems help users make decisions. In order to design effective recommendation processes, it is also necessary to consider the human factors involved in decision making. Feelings and personalities are two major influences of human decisions, feelings are an important element (Kahneman 2013). So is attitude, which describes human behavioural variations (McCrae and John 1992). Emotions last two seconds and count as the character is calm and reflects global prejudices over stretches of time. Emotions and personality have been employed to optimise aspects of recommendation systems like cold start, context-sensitive recommendation systems, implicit input, and diversity.

A. Describing Emotions

Emotion can be defined as "a condition that generally results from a major event in a subject. It usually includes (a) a conscious mental state with a distinctive characteristic of feeling and directed toward something, (b) a physical disorder of some kind, (c) characteristic facial expressions, tone of voice and signal. D) Preparation for specific types of work "[Oatley *et al*. nineteen ninety-nine. Emotion may be described, according to Pickard *et al*. 2001], as a sequence of adjustments in state, coherently and concurrently, in reaction to an evaluation of the value of an external stimulation or the learner. In addition, since they may affect interactions, activities, and thought, emotions are essential [Goleman 1995]. Emotion, for example, will often alter the content of a communication and this often suggests that the most significant element is not what was said, but how it was said [Picard *et al.* 2001]. Therefore, it is important for a machine in human and human-computer interaction to understand the various variations between the emotional states of an individual.

There are two key models of defining a user's emotional state:

(1) the model of general emotions and

(2) the model of dimensions.

The general model of emotions suggests that a small number of distinct emotional types are present. However, the categories suggested by Ekman [5] (i.e., pleasure, rage, sorrow, anxiety, disgust, and surprise) tend to be very general. There is no agreement on universal emotions. Dimensional based model, instead, defines every feeling of customer as a point in a continuous multifaceted space in which the quality of the emotion is expressed by each dimension. Equivalence, arousal and domination are the two widely used measurements (hence the abbreviation for VAD), although some scholars refer to these dimensions by other terms (for example, pleasure instead of equivalence in [13] or activation instead of arousal in [7]). The rotation framework suggested by Posner *et al.* [15], Simple emotion maps in the space of VAD (shown in Figure 1)

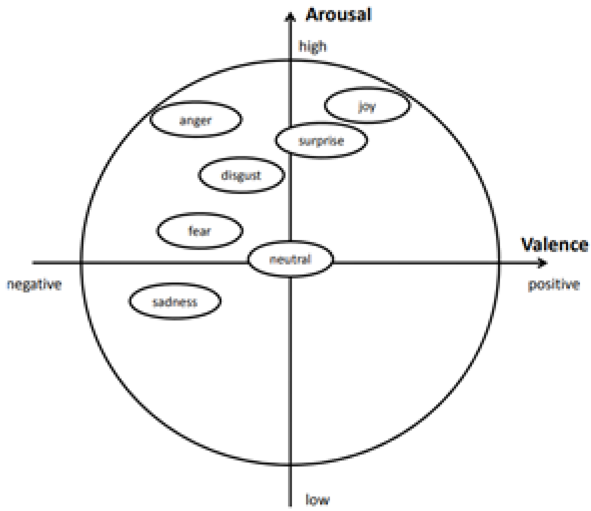

Fig. 1. Fundamental feelings in the multidimensional model's valence-arousal plane

When consumers scan, check, and dig for fascinating material, the recent exponential rise in information (for example, on the World Wide Web(www)) has intensified the require for intellectual support. Every day, with new data created (*e.g.,* on social media such as Twitter, Facebook and E-Newspapers), People are constantly receiving massive amounts of information that they are unable to handle on their own [1,2]. In comparison, the time it takes to view the social messages they produce rises, creating a lack of competitiveness, as more persons and subjects are pursued. Often, human beings are sometimes disturbed when social media changes are immediately disseminated, and This could make it more difficult for them to focus on a challenging work, especially when social media postings don't directly further their current goal. (for example, An effects of discontinuity may be expensive).

II. RELATED WORK

Three key categories may be categorized as preferred schemes (RS): collective filtering (CF), content-based filtering (CB) and a mixed solution incorporating the two preceding approaches. Pure CF approaches were introduced in the first experiments in the RS area, consisting of the measurement of similarities between users. Knowing the neighbours of the consumer (users with common views or tastes), according to the interests of the neighbours, the device suggests components. In reality, for this technique, the higher users rate the goods, the more specific the suggestions are. In this method, two major paradigms exist: model-based technologies and memory-based technologies, with the prevailing paradigm being the latter. Model-based CF technology combines various users of the learning sample into a limited count of groups depend on their grouping habits, unlike memory-based CFs, where a full user item classification data collection is used to produce a forecast.

Fotopoulou *et al.* (2020), developing the abilities of students in social and emotional areas will greatly increase their educational and social behaviour. This includes resources to support teachers in carrying out programs in social and emotional learning and measuring the effect made. To achieve this, it can be seen that it is useful to design intelligent and self-learning instruments with the potential to suggest tasks in accordance with the social and emotional interests of educational classes by integrating recommendation systems with machine learning techniques. They outline a simulation strategy for an integrated recommendation method in the current dissertation that aims to propose instructional opportunities for teachers to strengthen the social and emotional competencies of students and allow use of motivational learning strategies. The enhanced learning model is structured to take into account the creation of the social and emotional characteristics of students and the input provided across a series of experiences. A brief assessment of the detailed solution, with a focus on checking its suitability to satisfy educational needs, is presented.

Polignano *et al.* (2021), Decision-making is the cognitive process of defining and selecting options in accordance with interests, values, and the importance given to certain things or behaviours by the decision-maker. For starters, it is an easy and compact decision-making method to select which movie to see. Recommendation mechanisms help individuals with certain sorts of decisions, typically by estimating a limited list of recommendations that decreases the room for possible solutions. These programs depend highly on understanding of consumer desires, but they must be focused on a holistic interpretation of user behaviour to adequately serve individuals, which often requires how feelings, moods, and character characteristics impact them.

In this work, they explore how relational factors should be used in the recommendation method. They propose that a user's emotional condition, defined by a set of emotions (such as excitement and surprise), is part of the option of a mind-set that must be taken into consideration when designing the desires of the user. A generalized algorithm that considers emotions dependent on manipulating user profiles where every desire, like a movie's five stars ranking, is linked to the emotional condition that the user felt at the moment the preference was obtained, is the key contribution of the paper. The model calculates when the unseen variable blends into the present emotional condition of the user, a measure of emotional cohesiveness that considers both the traits of the influencing user and those of the non-influential component. The technique was applied in the recommendation method for emotion-conscious songs, the effectiveness of which was tested by performing in vitro studies on two standard data sets. The primary result is that, relative to baselines that did not provide any shaping details in the recommendation form, our framework shows increased consistency of recommendations.

Saraswat *et al.* (2020), User-generated material, such as ratings and feedback, includes details and user-confirmed thoughts on a given product. As internet usage increases, there are more types of user-generated data, including reviews. By sharing their comments and views regarding the products they buy online, after watching a video or reading books, etc., individuals express their perspectives, viewpoints, feelings and emotions. An emotional lexicon including pleasure, sorrow, and surprise is included in this user-generated content. A modern aspect of suggesting new products based on the emotional desires may be analysed from these emotions. In this analysis, emotions are extracted from the data produced by this consumer using lexical ontology, WordNet, and psychological knowledge. It is important to use these distilled emotions as suggestions. In contrast with the standard paradigm of resemblance of elements dependent on description, the emotion predictor measurement verifies the efficacy of the proposed model. They often compare this to ambiguities of interpersonal traits.

Almomani *et al.* (2019), the decision-making mechanism influencing option behaviour is discussed in this article. First, by executing a selection task on the web interface, the neuronal and theoretical behaviour of distinct subjects was experimentally registered. Second, the models of preference were designed with rational, emotional, and attentive functionality. For each user, the accuracy of the model predictions was measured and rankings were produced. The findings reveal that (1) the models of treatment are the strongest among all consumers in terms of overall efficiency, but (2) each subject better illustrates a different model.

Sailunaz & Alhajj (2019), online social networking has evolved as a new medium that provides users with a platform to express their viewpoints and thoughts with their friends, families, colleagues, etc. on different subjects and topics. Via text messages, photographs, photos, audio and posts, they will express our feelings, our state of mind, our moments and deal with unique social sites, national and foreign problems. In reality, text remains one of the most common social media contact tools, considering the availability of other means of communication. The work discussed in this article aims to identify, examine, and use the emotions and feelings people convey in the text of their Twitter posts to generate suggestions. Tweets and responses on particular subjects were gathered and a dataset was generated with email, users, feelings, emotional knowledge, etc. They used a dataset to reveal and calculate the effect of user ratings depending on the user of different Tweet-based metrics, the thoughts and sentiments of the Tweets and their replies. Finally, to build general and customized suggestions for users based on their activities on Twitter, they use the latest details. Few fascinating novelties are used in the approach they used in this paper, such as

(1) Including twitter responses in the data set and taking measurements

(2) Providing the degree of consensus, In the degree of impact measurement, the replies'

sentiment and emotion scores. (3) Produce a general and personal recommendation found in a collection of users who decided on the same subject and demonstrated the same thoughts and feelings on that specific matter.

Yang *et al.* (2018), Depression is an important mental disease of global concern. The dynamic etiology and persistent psychiatric features make it impossible for users to perceive their suicidal feelings and severely endanger the patient's wellbeing in existence. The smart suggestion framework has opened up new possibilities with the advancement of e-commerce to track the personal health of emotionally disturbed consumers. Therefore, this essay discusses health, a smart health advisory framework with a mental health predictor of depression. This essay discusses the monitoring and optimization of the psychological and physiological states of consumers by advancing customized recovery options for emotionally disturbed patients. In specific, the architecture of an e-health infrastructure is first proposed in this paper. They then developed smartphone applications devoted to gathering emotional data for people with unpleasant depressed moods and, via a Pearson correlation study, identified the top five outward symptoms of depression. They separated the data from 1,047 volunteers into a training and testing package, utilizing a decision tree to construct a depression prediction model and supporting algorithms for vector machines. As for the different external causes that contribute to depression, they include tailored recommendations and an intelligent decision-making solution, and push specific advice for relational change to guide the behaviour of consumers. Finally, a particular implementation scene is seen where the families of the patient help the patient with psychiatric therapy, to check the system's operability and suitability. This system’s positive results can fulfil the electronic market's desires and can be marketed and popularized.

Bodaghi & Homayounvala (2018), In the recommendation method, collaborative recommendation frameworks include people. And the sort of contact they have with digital recommendation systems, people with varying degrees of ability have different tastes. According to their skill levels, tailoring the form of user experiences with recommended systems seems to be a positive development to enhance the recommendation process and user retention of immersive recommended systems. In this paper, based on their contact with the integrated recommendation framework, expert users are automatically listed. To monitor user experiences, Shopr, an immersive suggestion framework for smartphones, is used. Based on their experiences with Shopr in our User Review, consumers are divided into two groups: experts and beginners. In order to know the ability level of the users, the work completion period is used. The outcome of the study indicates that all users known as expert users are still experts, but this algorithm did not find any expert users who invested more time in the recommendation framework to enhance their shopping range. For consumers who are specialists in digital recommendation systems, the findings of this quest may be used to have customized engagement.

Iliev & Stanchev (2018), The method for extracting details regarding works of art based on textual sources, descriptive of relative works of art, accessible via the metadata itself, is defined in this paper. Using alternate search methodologies and keywords, broad data sets can be conveniently retrieved from the diagrams. A text query is done using the keyboard in the most common search technique. Based on text queries for predefined metadata grouped into two groups, they propose a way to scan, locate, and recommend digital media material, and then map it to verbal emotion signals derived only from the emotion layer of expression. They further clarify the disparity between male and female speakers' emotional expression and further indicate that this differentiation can enhance the efficiency of the method.

Chakraverty & Saraswat (2017), The advantages of cross-domain recommendation systems (CDRs) are being leveraged by several e-commerce platforms to cross-sell goods, target potential customers, and raise sales. Current research enhances consumer item ratings with a range of additional details to provide an appropriate CDR, such as location, personality, geo - tags, and multimedia content, connecting multiple domains. They suggest a new perspective in this paper to produce suggestions in various realms by taking advantage of the feelings that are encapsulated in the user-generated textual material, such as feedback, blogs and comments. These emotions act as clear social and psychological connections between diverse entertainment fields and have the potential to prevent problems with cold start. A rich emotional lexicon is used by our CDR Blueprint to evaluate emotions in online material shared by users in the realms of source and destination and construct emotional profiles for objects and users in both domains. Therefore, to align these profiles to suggest objects in the target area, it applies collective filtering. Demonstrates CDR map function focused on empathy utilizing film and book domains as a case study. Compared to the previously published topic modelling methodology for the entertainment sectors, the experimental findings on the Movie lens and Book crossing data sets produce an F1 metric of 28.9 percent, which is a substantial 71.1 percent increase.

Saraswat & Chakraverty (2017), the emotions of feelings that occur after reading a book or seeing a video are encapsulated by user-generated content, such as reviews. It is important to use these feelings gathered from ratings to suggest elements of related feelings from the entertainment sphere. So far, more study has been conducted utilizing feelings in the recommendation framework as independent characteristics. In this post, in emotion groups, they dig into the usage of blur. In recognizing objects with features similar to the guidelines, the use of emotional characteristics such as affection, excitement, surprise, rage, sorrow, and fear has been found to be effective. There is, moreover, a certain degree of complexity and blurring of the distinctions between the lexicons of these features of categorical emotion that have been mostly neglected thus far. In this paper, by presenting a method for suggesting the cinematic usage of ambiguous emotion characteristics by using each of the feeling groups as a linguistic component, they discuss the issue of complexity implicit in emotional characteristics. To derive ambiguous ranking rules for recommending movies from feelings gleaned from their subsequent ratings, they established the Mamdani model. The findings indicate that the Gaussian Fuzzy model provides an F - measurement of 68.43 percent with 5 linguistic variables, which is a 10.5 percent increase over the clear SVM-based model for the suggestion of movies.

Santos (2016), This chapter discusses how affective computation in adaptive e-learning systems is regarded (in terms of methods of identification and methods of intervention). The goal behind this is to enrich the personal support offered in online educational settings by understanding the effect on the learning process of emotions and personalities. The key material of the chapter consists of a study of 26 articles presenting recent research patterns concerning the exploration of the emotional conditions of students and the availability of sufficient emotional help in different learning environments. The chapter also analyses open-ended issues pertaining to affective computing in school.

Tkalcic *et al*. (2011), in consumer and material modelling, recommendation systems have historically focused on data-driven descriptors. They have seen an increasing range of efforts in recent years to use feelings in multiple forms to increase the efficiency of recommendation systems. They introduce a unifying structure in this article that places the study work carried out thus far in a three-stage model in a distributed manner. They include samples of studies addressing numerous facets of feeling detection and translating them into suggestion structures.

Tkalcic *et al.* (2012), Latest analysis has shown that the accuracy of advisory programs utilizing emotional marks has improved. In this study, they propose a methodology for emotionally sensitive recommender system and conduct a survey of their performance. Within the picture suggestion scheme, they propose a structure focused on the usage chain and evaluate three types of tagging:

(1) Public Tagging,

(2) Overt Emotional Tagging, and

(3) Implied Emotional Tagging.

Mariappan *et al.* (2012), Researchers are developing context-sensitive software that reacts to a user's emotional state by recognising the user's facial gestures. The field of computer vision is actively researching the recognition of face emotion. They implement Face Fetch in this post, a new context-based recommendation framework for multimedia content that understands the current emotional condition of a consumer (happiness, sorrow, anxiety, disgust, surprise, and anger) by identifying facial expressions and recommending the user. Our gadget will collect multimedia content from the cloud, comprising songs, movies, and other films of interest to the user, via the desktop and smartphone user interface, and do so with almost real-time results.

III. APPROACH

The key components suggested for the emotion-based SR as seen in Figure 2. It is possible to view this device design as a interface between customer requirements (implicit or explicit) and knowledge accessible. The main aim is to provide the consumer with knowledge selectively. As we know, users' intentions and behaviour are not necessarily mechanisms that are independent, but appear to be shaped by their social environment [5]. It is not, however, a simple job to identify the user's priorities and ambitions, since it is focused on being acquainted with the dynamic programme. Therefore, a agent (PAA) must also understand social website patterns in addition to considering the consumer model (that is, desires, intentions, and emotions depicted in Figure 2, as human awareness and emotional traits). Trends consist of new data in the process of this work that arouses novelty, surprise or even interest in the culture of the method. The introduction of cooperative agents to share and explore new patterns is crucial for this problem.

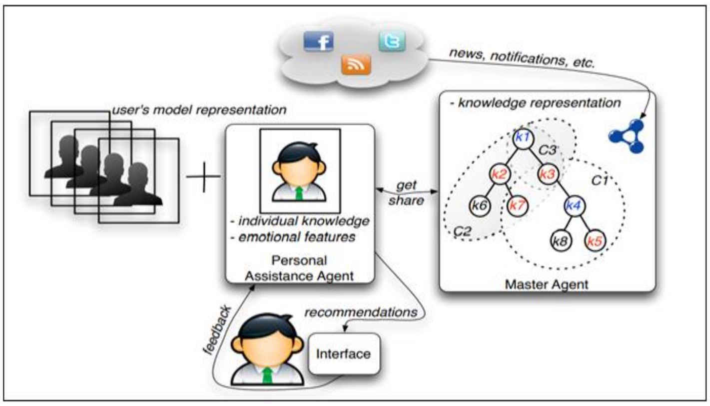

Fig. 2. Emotion-Based RS's Architecture.

An Emotion-Based RS's Architecture is the system that collect the response from the social media, then some agent processes the data and arrange it in some collected manner of understanding presentation with some mechanism. Further this form of information has been categorically recommended. A limited recommendation filter can serve as the user agent. And if it doesn't match the set of tacit user expectations, it can equate the user type with that of other sections of the software and extract new details that may be useful. In brief, in a setting where any of their particular information is exchanged, these agents can work. It would be important to establish a formal representation of information that these agents will use while executing activities in order for the mechanism to work well. The details gathered and the social patterns received from user feedback should be integrated into this system (for example, the number of views of a particular product and its emotional effect). By questioning consumer about the feelings triggered by such details, the emotional effect can be calculated. Another requirement of the method is to organize classes of related elements using main phrases (represented in Fig. 2). A concise overview of the substance of the text includes main phrases. More precisely, a short representation of the records is the main phrases. The importance of this overview knowledge rises when vast sets of records, such as news stories, proliferate and are generated every second. Key phrases are also extremely beneficial since they may be understood separately and independently of one another. Furthermore, key phrases are always manually selected. Although this task is less vulnerable to error, it is hard to repeat, time-consuming, and often arbitrary. Our code requires, however, to retrieve main phrases from heterogeneous network sources and make them fully automatically in a structured representation.

A. Recommender system

A recommender system is a **compelling information filtering system** running on machine learning (ML) algorithms that can predict a customer's ratings or preferences for a product.

Working of Recommendation system

B. Framework based on collaborative filtering

Data filtration and classification techniques are the foundation of recommendation system, which combine two perspective user and content. The heart of the recommender revolves around two categories of consumer data:

- Explicit data is produced when a person uses a specific action to express his or her interests (for example, ranking the product)
- The bulk of the user account is comprised of implicit data acquired, or user behaviour themselves (for example, type of products customer views)

In general, the recommender engine works as follows:

- Collecting customer information
- Identifying customer behaviours and action patterns
- Obtaining insightful information
- Probability computation
- Assessing them to the inventories of available items
- Outlining the most logical pairings

This extends what was already seen in the previous section and now represents one of the most cutting-edge ways. The consumer grid with ratings always serves as the place to start: Equation 1, show rating matrix for user with respect to product.

$$M_{uxl}=\begin{pmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mn} \end{pmatrix} \qquad \text{Eqn-1}$$

C. Memory-based collaborative filtering

Memory-based CF algorithms produce predictions by using all or a portion of a customer-item database. Each user belongs to a community of individuals with comparable characteristics. The goal of the neighbourhood-based CF method, a popular memory-based CF technique, is to identify a user's closest neighbours (so-called active user). Assume we have a rating matrix with users as row, objects as sections, and each cell representing a user's assessment of an object. To put it another way, every row shows a customer vector or ratings vector. An existing user vectors is the current user's rating vector. A rating matrix including one attribute data is shown in Table 1. Note that the question mark (?) indicates a missing value. For instance, user 4 doesn't really rate on things 3 and 4 since r43 and r44 have incomplete data.

Table 1. Rating matrix (Four customer active)

$s.no$	$prd1$	$prod2$	$prod3$	$prod4$
$cust1$	$r_{11} = 1$	$r_{12} = 2$	$r_{13} = 1$	$r_{14} = 5$
$cust2$	$r_{21} = 2$	$r_{22} = 1$	$r_{23} = 2$	$r_{24} = 4$
$cust3$	$r_{31} = 4$	$r_{32} = 1$	$r_{33} = 5$	$r_{34} = 5$
$cust4$	$r_{41} = 1$	$r_{42} = 2$	$r_{43} =?$	$r_{44} =?$

Let us consider ui = ($r_{i1}, r_{i2}, , r_{in}$) and a = ($r_{a1}, r_{a2},\ldots, r_{an}$) be the normal customer vector i and the active customer vector a, respectively where r_{ij} is the rating of customer i to item j. According to table 1, we have u_1 = (1, 2, 1, 5), u_2 = (2, 1, 2, 4), u_3 = (4, 1, 5, 5), and a = u_4 = (1, 2,?).

If any of the cells in the existing user vectors are blank, the ratings matrices are considered to be sparse since the current user did not rate the relevant items. The existing user element's missing data must be predicted, and the items with the highest values must then be suggested to the authorized member. The procedure of anticipating null values involves two steps:

1. Finding out nearest neighbours of active user.
2. calculating prognostic values (or predictive ratings).

Be aware that determining the closest neighbours of an active user is the basis for calculating prediction performance.

Methodology for model based collaborative filtering

In order to generate fresh suggestions, the Collaborative filtering technique for recommendation system only relies on the historical contacts that have observed among users and products. Technologies for making suggestions based on content take into consideration information that users have both directly and indirectly contributed. As an illustration, consider social media information included in this study, which includes user ratings, mean ratings, and recommendations.

D. Result outcome

This analysis has been executed in python with user rating of social data. As presented below the screen shot of the execution of the python for analysis of social data.

```
import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import os
from IPython.core.interactiveshell import InteractiveShell
InteractiveShell.ast_node_interactivity = "all"
import math
import json
import time
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics.pairwise import cosine_similarity
from sklearn.model_selection import train_test_split
from sklearn.neighbors import NearestNeighbors
from sklearn.externals import joblib
import scipy.sparse
from scipy.sparse import csr_matrix
from scipy.sparse.linalg import svds
import warnings; warnings.simplefilter('ignore')
%matplotlib inline

for dirname, _, filenames in os.walk('/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))
```

Fig. 3. Libraries

As the above figure presented that import of various library function has been called for execution of the whole programming.

```
electronics_data=pd.read_csv("input/ratings_Electronics.csv",names=['userId', 'productId','Rating','timestamp'])
```

```
electronics_data.head()
```

	userId	productId	Rating	timestamp
0	AKM1MP6P0OYPR	0132793040	5.0	1365811200
1	A2CX7LUOHB2NDG	0321732944	5.0	1341100800
2	A2NWSAGRHCP8N5	0439886341	1.0	1367193600
3	A2WNBOD3WNDNKT	0439886341	3.0	1374451200
4	A1GI0U4ZRJA8WN	0439886341	1.0	1334707200

Fig. 4. CSV data

As the above figure presented in table presented that the user Id, Product ID, rating and timestamp of the user data.

```
electronics_data.dtypes

userId        object
productId     object
Rating       float64
timestamp      int64
dtype: object
```

```
#Five point summary

electronics_data.describe()['Rating'].T

count    1.048576e+06
mean     3.973380e+00
std      1.399329e+00
min      1.000000e+00
25%      3.000000e+00
50%      5.000000e+00
75%      5.000000e+00
max      5.000000e+00
Name: Rating, dtype: float64
```

Fig. 5. CSV data Statistics

As the above figure presented the data types in one figure and rating of each attribute of the figure.

```
#Find the minimum and maximum ratings
print('Minimum rating is: %d' %(electronics_data.Rating.min()))
print('Maximum rating is: %d' %(electronics_data.Rating.max()))

Minimum rating is: 1
Maximum rating is: 5
```

Fig. 6. Minimum and maximum rating

The above figure presented that the minimum and maximum rating of each attribute as per the rating given by the user through social platform.

```
#Check for missing values
print('Number of missing values across columns: \n',electronics_data.isnull().sum())

Number of missing values across columns:
 userId       0
productId    0
Rating       0
timestamp    0
dtype: int64
```

Fig. 7. Missing Values

The above of the figure is treating the missing values in the collected sample data in CSV. This is very important steps for rehandling of the data.

```
# Check the distribution of the rating
with sns.axes_style('white'):
    g = sns.factorplot("Rating", data=electronics_data, aspect=2.0,kind='count')
    g.set_ylabels("Total number of ratings")
```

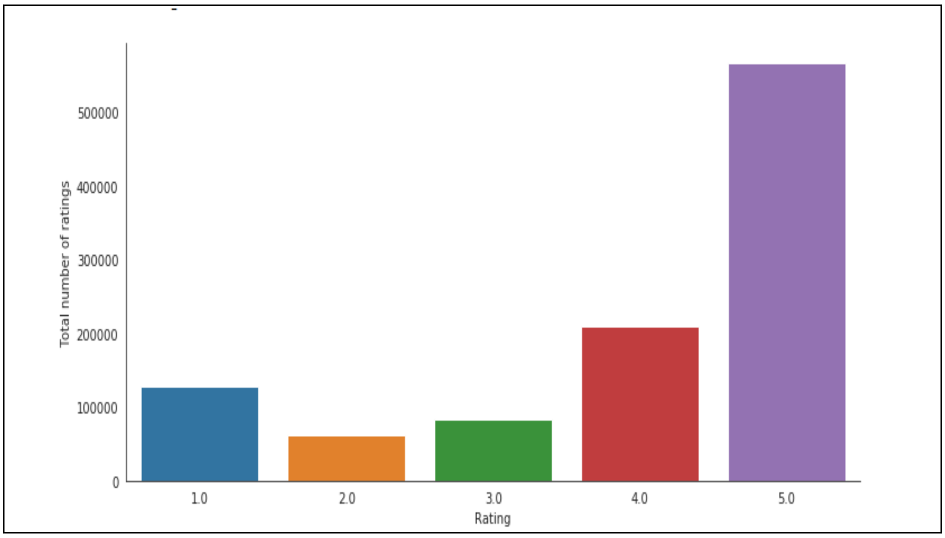

Fig. 8. Distribution of Rating

Now the distribution graph of the rating has been presented in above figure. Total number of rating and start provided in rating has been demonstrated in the figure.

```
print("Total data ")
print("-"*50)
print("\nTotal no of ratings :",electronics_data.shape[0])
print("Total No of Users   :", len(np.unique(electronics_data.userId)))
print("Total No of products  :", len(np.unique(electronics_data.productId)))

Total data
--------------------------------------------------

Total no of ratings : 1048576
Total No of Users   : 786330
Total No of products  : 61894
```

Fig. 9. Total number of Rating

Further the total number of ratings along with users and products has been number out in the above figure with the python very basic print code.

```
#Analysis of rating given by the user

no_of_rated_products_per_user = electronics_data.groupby(by='userId')['Rating'].count().sort_values(ascending=False)

no_of_rated_products_per_user.head()

userId
A5JLAU2ARJ0BO     412
A231WM2Z2JL0U3    249
A25HBO5V8S8SEA    164
A6FIAB28IS79      146
AT6CZDCP4TRGA     128
Name: Rating, dtype: int64
```

Fig. 10. Rating given by user

Above table shows the rating provided by user for each product Id has been presented.

```
no_of_rated_products_per_user.describe()

count    786330.000000
mean          1.333506
std           1.385612
min           1.000000
25%           1.000000
50%           1.000000
75%           1.000000
max         412.000000
Name: Rating, dtype: float64
```

Fig. 11. Rating Statistics

Now the count, mean, standard mean, minimization has been further extracted.

```
no_of_ratings_per_product = new_df.groupby(by='productId')['Rating'].count().sort_values(ascending=False)

fig = plt.figure(figsize=plt.figaspect(.5))
ax = plt.gca()
plt.plot(no_of_ratings_per_product.values)
plt.title('# RATINGS per Product')
plt.xlabel('Product')
plt.ylabel('No of ratings per product')
ax.set_xticklabels([])

plt.show()
```

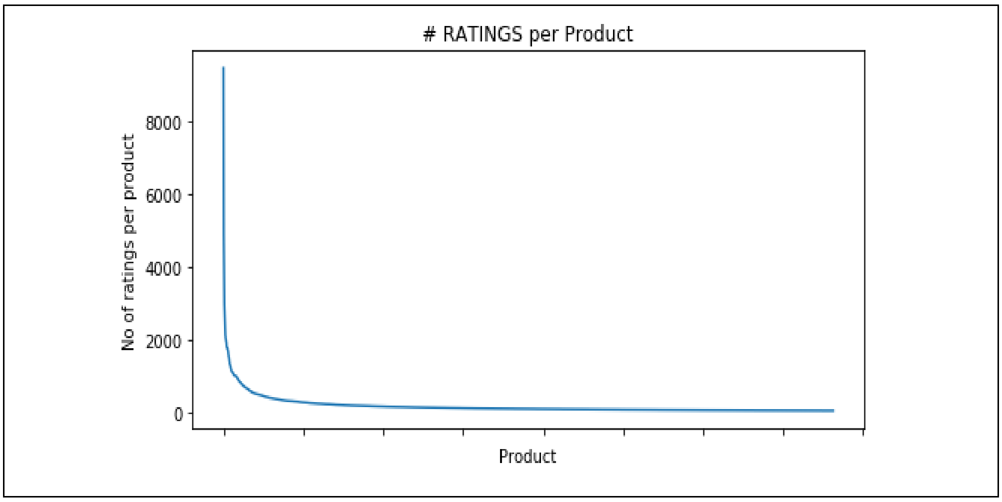

Fig. 12. Rating per product

Now product as per the rating giving has been plotted in above figure.

```
#Average rating of the product

new_df.groupby('productId')['Rating'].mean().head()
```

```
productId
0972683275    4.470980
1400501466    3.560000
1400501520    4.243902
1400501776    3.884892
1400532620    3.684211
Name: Rating, dtype: float64
```

Fig. 14. Average rating of product

Now the average rating has been estimated for the product and presented above in the figure.

```
ratings_mean_count['rating_counts'] = pd.DataFrame(new_df.groupby('productId')['Rating'].count())
```

```
ratings_mean_count.head()
```

productId	Rating	rating_counts
0972683275	4.470980	1051
1400501466	3.560000	250
1400501520	4.243902	82
1400501776	3.884892	139
1400532620	3.684211	171

Now the above figure has both as the average rating and rating count of each product.

```
plt.figure(figsize=(8,6))
plt.rcParams['patch.force_edgecolor'] = True
sns.jointplot(x='Rating', y='rating_counts', data=ratings_mean_count, alpha=0.4)
```

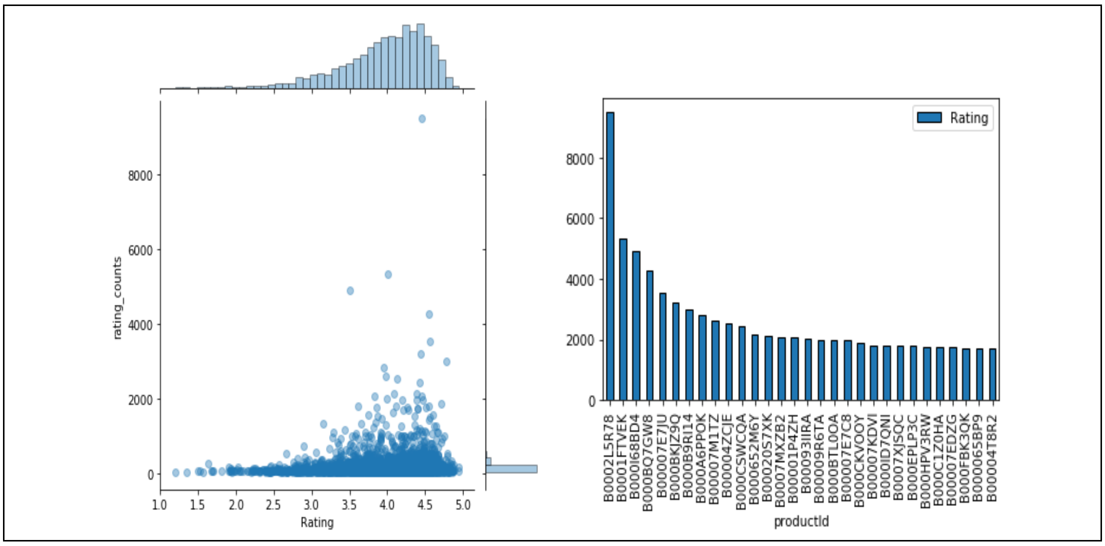

Fig. 16. Rating product id

The above figure has two distinct figures, one represents the rating count vs rating and second figure is the number of count vs product is in visual representation. Over all this above is the part of exploratory research of the data, in our investigation we are exploring the social media data as user rating and the product which has been rated by user. Now Collaborative filtering technique has been used to recommend the product as per user buying behaviour.

Collaborative filtering (Item-Item recommendation)

```
from surprise import KNNWithMeans
from surprise import Dataset
from surprise import accuracy
from surprise import Reader
import os
from surprise.model_selection import train_test_split
```

Now the library has been imported for Collaborative filtering for the same file which has been explored earlier for initiation of recommender system.

```
#Reading the dataset
reader = Reader(rating_scale=(1, 5))
data = Dataset.load_from_df(new_df,reader)

#Splitting the dataset
trainset, testset = train_test_split(data, test_size=0.3,random_state=10)

# Use user_based true/false to switch between user-based or item-based collaborative filtering
algo = KNNWithMeans(k=5, sim_options={'name': 'pearson_baseline', 'user_based': False})
algo.fit(trainset)
```

Now the data set has been training and test has been performed with code as presented above.

```
Estimating biases using als...
Computing the pearson_baseline similarity matrix...
Done computing similarity matrix.
```

```
test_pred = algo.test(testset)

test_pred
```

Now the person base line similarity matrix has been testing for this data.

```
[Prediction(uid='A2CTXDLXKOHCEH', iid='B00003CWDH', r_ui=5.0, est=4.050870065449826, details={'was_impossible': True, 'reason':
'User and/or item is unkown.'}),
 Prediction(uid='ABGCDODB3OLPE', iid='B0002D6QJO', r_ui=3.0, est=4.050870065449826, details={'was_impossible': True, 'reason':
'User and/or item is unkown.'}),
 Prediction(uid='AIL0HKZRRCE3L', iid='B00097CUIE', r_ui=5.0, est=4.050870065449826, details={'was_impossible': True, 'reason':
'User and/or item is unkown.'}),
```

```
# get RMSE
print("Item-based Model : Test Set")
accuracy.rmse(test_pred, verbose=True)
```

```
Item-based Model : Test Set
RMSE: 1.3436

1.343641161111319
```


Fig. 17. Finding RMSE

The above figure found the RMSE value of the model (Item Based Model). This RSME is significant when the whole recommendation has been provided. This value is the kind of expected error which could be accountable for product recommendation.

Model-based collaborative filtering system

Model-based recommender system builds a model from training data using machine learning methods, which is then

used to predict future.

```
new_df1=new_df.head(10000)
ratings_matrix = new_df1.pivot_table(values='Rating', index='userId', columns='productId', fill_value=0)
ratings_matrix.head()
```

productId	0972683275	1400501466	1400501520	1400501776	1400532620	1400532655	140053271X	1400532736	1400599997	1400698987	...
userId											
A01852072Z7B68UHLI5UG	0	0	0	0	0	0	0	0	0	0	...
A0266076X6KPZ6CCHGVS	0	0	0	0	0	0	0	0	0	0	...
A0293130VTX2ZXA70JQS	5	0	0	0	0	0	0	0	0	0	...
A030530627MK66BD8V4LN	4	0	0	0	0	0	0	0	0	0	...
A0571176384K8RBNKGF8O	0	0	0	0	0	0	0	0	0	0	...

The above is the pivot table of the rating, userID, ProductID in the form of matrix.

```
X = ratings_matrix.T
X.head()
```

userId	A01852072Z7B68UHLI5UG	A0266076X6KPZ6CCHGVS	A0293130VTX2ZXA70JQS	A030530627MK66BD8V4LN	A0571176384K8RBNKGF8O	A0590501PZ7I
productId						
0972683275	0	0	5	4	0	
1400501466	0	0	0	0	0	
1400501520	0	0	0	0	0	
1400501776	0	0	0	0	0	
1400532620	0	0	0	0	0	

Now the rating Matrix has been presented above in matrix formatted.

```
X1 = X
```

```
#Decomposing the Matrix
from sklearn.decomposition import TruncatedSVD
SVD = TruncatedSVD(n_components=10)
decomposed_matrix = SVD.fit_transform(X)
decomposed_matrix.shape
```

```
(76, 10)
```

```
#Correlation Matrix

correlation_matrix = np.corrcoef(decomposed_matrix)
correlation_matrix.shape
```

```
(76, 76)
```

```
X.index[75]
```

```
'B00000K135'
```

```
i = "B00000K135"

product_names = list(X.index)
product_ID = product_names.index(i)
product_ID
```

```
75
```

```
correlation_product_ID = correlation_matrix[product_ID]
correlation_product_ID.shape
```

```
(76,)
```

Fig. 18. collaborative filtering system

Now further decomposing the matrix and further the particular product “i” as B00000K135 has been corelation index found as 76.

Findings from above

```
Recommend = list(X.index[correlation_product_ID > 0.65])

# Removes the item already bought by the customer
Recommend.remove(i)
Recommend[0:24]
```

```
['1400501520',
 '1400501776',
 '1400532620',
 '1400698987',
 '6301977173',
 '9573212919',
 '9966694544',
 'B00000J1SC',
 'B00000JCT8',
 'B00000JD34']
```

Fig. 19. Recommended product

Over all after the various step has been performed for Collaborative filtering recommendation, and the result of the suggestive product has been found in above figure.

IV. CONCLUSION AND FUTURE WORK

The growth in the volume of Internet-linked data allows it more complicated to collect knowledge from all possible sources. Recommender programs concentrate on discovering valuable material for consumers through filtering through vast quantities of dynamically produced knowledge. It is impossible to retrieve information from the Internet in an effective manner due to the overwhelming volume of information that is accessible on the internet. Consumers are provided with specific knowledge by recommendation mechanisms after these mechanisms sift through vast quantities of material. The attention of customers or clients is typically garnered by a recommender system that provides feedback that is both effective and pertinent. There are some gaps in the existing body of knowledge that require attention from researchers. Although research has been conducted on feeling-based recommendation systems, there are still a number of uncharted territories that are easily accessible. In this paper, we investigate the user rating data for the purpose of using collaborative filtering to make product recommendations. The ultimate goal is to investigate the applicability of the recommender method on a more extensive scale. For the purpose of providing a recommendation of Emotion-Based RS's Architecture, which is user collaborating filtering techniques, this paper makes use of the user ratings left on social media posts by other users. The conclusion, which included product recommendations, has been presented in its final format. The Emotion-Based RS's Architecture Integrated with Collaborating Filtering Techniques was proposed as the result of this investigation. The ultimate aim of this paper is to analyse the Recommender Method with more adaptive techniques. The future work of this research is to extend the same and provide the substantial proof as simulation for emotion-based recommendations.

references

1. Fotopoulou, E., Zafeiropoulos, A., Feidakis, M., Metafas, D., & Papavassiliou, S. (2020, June). An interactive recommender system based on reinforcement learning for improving emotional competences in educational groups. In *International Conference on Intelligent Tutoring Systems* (pp. 248-258). Springer, Cham.
2. Polignano, M., Narducci, F., de Gemmis, M., & Semeraro, G. (2021). Towards Emotion-aware Recommender Systems: an Affective Coherence Model based on Emotion-driven Behaviours. *Expert Systems with Applications*, *170*, 114382.
3. Saraswat, M., Chakraverty, S., & Kala, A. (2020). Analyzing emotion-based movie recommender system using fuzzy emotion features. *International Journal of Information Technology*, *12*(2), 467-472.
4. Almomani, A., Monreal, C., Sieira, J., Graña, J., & Sánchez, E. (2019, April). Rational, emotional, and attentional choice models for recommender systems. In *World conference on information systems and technologies* (pp. 557-566). Springer, Cham.
5. Sailunaz, K., & Alhajj, R. (2019). Emotion and sentiment analysis from Twitter text. *Journal of Computational Science*, *36*, 101003.
6. Yang, S., Zhou, P., Duan, K., Hossain, M. S., & Alhamid, M. F. (2018). emHealth: towards emotion health through depression prediction and intelligent health recommender system. *Mobile Networks and Applications*, *23*(2), 216-226.
7. Bodaghi, A., & Homayounvala, E. (2018, April). Personalization of interactive recommender systems for expert users. In *2018 4th International Conference on Web Research (ICWR)* (pp. 58-62). IEEE.
8. Iliev, A., & Stanchev, P. (2018, April). Information retrieval and recommendation using emotion from speech signals. In *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)* (pp. 222-225). IEEE.
9. Chakraverty, S., & Saraswat, M. (2017). Review based emotion profiles for cross domain recommendation. *Multimedia Tools and Applications*, *76*(24), 25827-25850.
10. Saraswat, M., & Chakraverty, S. (2017, October). Leveraging movie recommendation using fuzzy emotion features. In *International Conference on Recent Developments in Science, Engineering and Technology* (pp. 475-483). Springer, Singapore.
11. Santos, O. C. (2016). Emotions and personality in adaptive e-learning systems: an affective computing perspective. In *Emotions and personality in personalized services* (pp. 263-285). Springer, Cham.
12. Tkalcic, M., Kosir, A., & Tasic, J. (2011, October). Affective recommender systems: the role of emotions in recommender systems. In *Proc. The RecSys 2011 Workshop on Human Decision Making in Recommender Systems* (pp. 9-13).
13. Tkalcic, M., Burnik, U., Odić, A., Košir, A., & Tasič, J. (2012, September). Emotion-aware recommender systems–a framework and a case study. In *International Conference on ICT Innovations* (pp. 141-150). Springer, Berlin, Heidelberg.
14. Mariappan, M. B., Suk, M., & Prabhakaran, B. (2012, December). Facefetch: A user emotion driven multimedia content recommendation system based on facial expression recognition. In *2012 IEEE International Symposium on Multimedia* (pp. 84-87). IEEE.