Available online at http://ijdea.srbiau.ac.ir

Int. J. Data Envelopment Analysis (ISSN 2345-458X)

Vol. 11, No. 4, Year 2023, Article ID IJDEA-00422, Pages 38-54

Research Article

International Journal of Data Envelopment Analysis

Science and Research Branch (IAU)

Modeling new ranking behavior of users in online social networks to achieve efficient recommending algorithms

Mehdi Safarpour[1], S. Hadi Yaghoubyan[* 1,2], Karamollah BagheriFard[1], Razieh Malekhoseini[1], Samad Nejatian[3]

[1] Department of Computer Engineering, Yasooj Branch, Islamic Azad University, Yasooj, Iran,

[2] Young Researchers and Elite Club, Yasooj Branch, Islamic Azad University, Yasooj, Iran,

[3] Department of Electrical Engineering, Yasooj Branch, Islamic Azad University, Yasooj, Iran

Received 19 March 2023, Accepted 27 June 2023

Abstract

Today, the role of recommender algorithms on online platforms of social networks is considered with the penetration rate increase on the platforms. Moreover, investigating these algorithms' functionality accuracy is important in the appropriate recommendations. These algorithms have been introduced for a long time, and they try to propose appropriate recommendations by the users' behavior modeling. But, nowadays, with an increasing number of people relatives on social networks, the networks' users' behavior is psychological. Moreover, the people's action is related to different events such as publishing a post in addition to the post's contention, the user interest in the post publisher, and the people relationship. This paper shows that user behavior in a social network is predictable and indicates the possibility of incorrect recommendations in participatory filtering algorithms.

Keywords: Social networks, Recommendation systems, Collaborative filtering, Users' behavior prediction, Social relationship

[*] Corresponding author: Email: yaghoobian.h@gmail.com

1. Introduction

Nowadays, online social networks have become a popular platform for information dissemination and people relationships. People can communicate with their family and friends. Governments and commercial organizations can provide new opportunities for business, policy, and services for citizens and consumers. People use social networks for different aims, mainly relationships with their friends and sharing some events [1].

With the advent of Online Social Networks (OSN), using Recommender Systems (RS) that make better recommendations using the users' profile characteristics, activities, and interests enters a more practical phase called Social Recommendation Systems (SRS). OSNs can provide information to identify virtual, physical, and social relationships among users, their profits, common habits, and personal favorites. The RSs propose appropriate recommendations using information extraction and their defined application area. SRS has expanded its recommendation domain from general suggestions to people recommendations for friendship, locations, tags, and other cases.

User's behavior investigation, user's settings and favorites, the subjects and their relationships, identifying the users with the same thinking, and so on can use for information filtering to show only the related subjects, updated subjects, and favorites of the users. Many papers have tried to model one or more aspects of the users' behavior characteristics to extract data from the behavior of social networks' users. But according to the complexity of human behavior, all users' behavior's characteristics are not studied. Content-based and Collaborative Filtering (CF) methods have been used for a long time to predict how people evaluate an item and give points. Filtering the content is based on this assumption that the user wants to observe and select the previously accepted items by the user. In CF, the recommender algorithm shows the items that probably are of user's interest based on the analysis of the user and other users' (their friends) favorites [2].

The main principle of recommender algorithms is the meaningful relationship between the user and the selected items. Hence, the main idea is the possibility of extracting user favorites based on the items' characteristics that the user is rated and accessed previously [3]. But it seems that the users' behavior is not the same for all posts, and an item acceptance by the user in the social networks sometimes depends on the user's psychological factors like emotional relationship with the person who has published the item (social correlation) [4-6], cognitive limitation based on the number of friends [7], brain's limited capacity for social relationships management [8], the location of the item when showing on the page [9], (the issues on top of the users' page on the social network are more significant. Hence, the issues on top of the page may be seen more than the lower issues). Thus, the social and cognitive factors in social media psychology are so effective on the probability of acceptance or not acceptance of an item.

In this research, the reaction of the Instagram virtual network to different posts is studied. Moreover, it is examined whether users accept the published posts by their friends based on their favorites or if they like the posts based on their stable social relationships. Since the proximity criterion of the virtual network considers increasing communication between two people as the base of their similarity in most recommender algorithms, is it possible to have incorrect recommendations using these algorithms?

The rest of this paper is organized as follows. Section 2 includes the investigation. The influential social factors on people's relationships from a psychological point of view are identified in section 3. The user's behavior modeling is explained in section 4. The experiments on the Instagram network are evaluated in section 5. The research findings and the interpretation of the observed behavior are presented in section 6.

2. Investigation

With the advent of social networks, the RSs are considered in most research. Using these systems has gone beyond the recommendation of items and new users. It includes different cases, from job suggestions to detecting the links that are probably converting to stable relationships [10].

The background of each RS is an evaluation of the users' favorites. The users' favorites and preferences in the social networks are determined explicitly or implicitly. The explicit preferences include likes and comments, and the implicit preferences include searching the special information or seeing specific pages or posts. Extracting the user favorites is a base for filtering the information presented to the user. The user's favorites can be modeled based on the characteristics of the items, which are rated or accessed previously by the user [3]. The users' behavior analysis in the social networks is critical for the users' favorites extraction. Kreshna et al. (2016) analyzed user behavior in three steps: user behavior characterization, user behavior recognition, and user behavior prediction. Moreover, they proposed some methods for each step.

Campana and Delmastro (2017) categorized the item suggestion methods and divided the algorithms into two general, memory-based (neighborhood-based) algorithms, which use the relationships and similarities of a user with its neighbors for an item suggestion, and 2 model-based algorithms, which use machine learning techniques to learn the objective function. Both approaches work based on CF algorithms [11].

Many algorithms use memory-based approaches. Agrawal and Chen (2010) combined statistical techniques of subject investigation using CF algorithms to improve the RS power to use the user favorites and the item popularity for the suggestions. They showed that the system performance is improved when the user characteristics and the item's content are considered [12]. Kang and Lerman (2013) used the online users' activities to find users' limited attention to their friends and their published items. The method had two important improvements by proposing the LA-CRT model. It was combined with the item's content. Hence, its explanatory power was added to the suggestions and suggested the new related items. Also, the method could learn the penetration rate and other users' influence on an item acceptance based on the limited attention model [2]. The main problem of the memory-based algorithms is that they cannot predict the scoreless items' scores. The problem is solved using model-based algorithms [11].

Among the model-based item suggestion methods such as association rule-based, Bayesian network, Support Vector Machines (SVM), neural networks, deep learning methods, and Matrix Factorization (MF), the MF is the best method from a scalability and precision point of view [3]. Bokde et al. (2015) believed that the MF method is the most powerful method to find the hidden structure of each data, and they proposed four known models of SVD, PCA, PMF, and NMF [13]. The common idea in the MF-based solutions is that the user's social relationships affect the user's preferences (e. g., scoring). Two friends' selections are more similar than unfamiliar people. Therefore, these relationships determine the weight of the social relationships showing the power of social communications among the users and usually are rated based on similarity [3].

In the CTR-smf method, CTR was integrated with social matrix experience models to consider the social communications among people. The method used the Homophily effect instead of knowing how people attend to their friends [14].

Clustering was used for friends' groups identification in RS based on Social Networks (RSboSN) method. The RSboSN method's precision and calls were better than the SoRec method as the clustering-based approach effectiveness proof [15].

A user's friend information was extracted in the Recommendation with Direct and Indirect Social Influence (RDISI) method, and the users' penetration rate on their neighbors was calculated based on graph theory with a view of the close interests of a user and his(her) trusted friends. The authors used the obtained information for their proposed RS algorithm. They believe that combining the topologic penetration with the recommended algorithm makes better suggestions effect. Moreover, in addition to the penetration among the related users, the users' indirect interaction is investigated, improving the recommendation performance [16].

Based on Fan et al.'s (2019) research, although using the social networks' information for recommendations is possible, there are big challenges. First, finding social interactions in far social relationships is complex, and extracting helpful and correct information for the recommendations is difficult. Second, selecting the neighbors' effective data interacting with different items is complicated. Finally, preferences extraction from the user's interactions and items is challenging. The authors investigated the effect of the users' interactions on their decisions in social networks using deep learning to solve the challenges. They proposed a deep CF model called Deep Social Collaborative Filtering (DSCF). The advantage of this method for an item acceptance prediction by a user was considering direct and far neighbors in the calculations [17].

Popularity-based ItEm Recommender System (PLIERS) is a tag-based item recommendation method that suggests items in order of the user's favorites. It is assumed that the popularity of an item or a tag is related to its meaning. The more general a tag is the more rate of use for it. The more private a tag is the lower its generality. The PLIERS method solves the problem of selection between general and non-general items of the network and guarantees that the recommended item's popularity is compatible with the accepted item by the users. The precision of this method is more than other diffusion-based solutions [18].

In (Bianca, 2018), the number of transferred messages among the users is extracted using a graph. Finally, a tensor including the users, relationships, and the communication duration is created based on the number of messages. Also, the relationships' classification is performed using a hierarchical method [6].

Although the above methods have been successful in providing an RS, they mostly ignore the psychological and cognitive aspects of the user's behavior. Hence, we have studied different papers on social and cognitive factors of social media psychology, and the results are presented in the following sections.

3. Social Factors

Online social media like Facebook, Twitter, Instagram, and so on are created based on social human behavior modeling. Hence, studying the problems of social psychology is important to understand the users' behavior in online social media.

From the social psychology point of view, the proximity principle says that people with more meetings tend to have stronger relationships. The increasing use of technology-based communications makes it possible to have relationships without considering the physical distance. It seems these communications increase people's relationships, but most people use it who already know each other situations [19].

People are interested in other people with the same favorites [20]. In social science and sociology studies, the Homophily rule says that people with similar characteristics tend to have relationships in social networks [21]. Hyon et al. (2020) say that people tend to have relationships with other people with similar characteristics to them as age, gender, and nationality. Moreover, the socially close people may be similar in endogenous distributed changes in their attention distribution way (e. g. in the environment) during the time [22].

People's ability to manage social relationships is limited and determined by brain capacity. Dunbar number explains one of the human's cognitive limitations. The number defined by Robin Dunbar determines the maximum number of people with stable social relationships with a person. The researchers believe that this number is between 150 and 250. However, there is no consensus on the exact value of the number. In stable social relationships, people know others who have a relationship with them and their relations with other people [19]. Although it is believed that online social media develop the size of social networks and the set of related people, Goncalves et al. (2011) and Dunbar (2016) showed that the number of users' friends in online social media is close to the real social networks [23,24]. However, users rarely share their attention. They may have more attention to some of their friends because of their familiarity, trust, social closeness, or influence [2].

Based on the research of Yin et al. (2014), the rating behavior of users essentially doesn’t show their intrinsic interests. They have investigated several social network systems and claim that the rating behaviors of users are generally affected by the intrinsic interest and general attention of users. The first one is relatively stable, but the last one is affected by daily news in specific time durations and is variable. The authors have studied the effect of news and general attention on the viewed items by the user in the time durations [25].

Aivazoglou et al. (2020) performed a study on 38 users (22 to 34 years old, 79% men) to obtain more insights about desired characteristics of social recommendations. The interaction-related feedback presents their expectations and requirements focusing on their friends' recommendations on Facebook. Their goal was better to understand and obtain information about how users spend their time on Facebook. Moreover, the amount of the users' acceptance of entertainment (movie and music) provided by the recommender algorithms is evaluated [10]. Based on their findings, although there is overlap in a people's favorites with their friends, it is better to suggest the content according to the recommendation of the strongly related friends of the person, not based on the number of likes and publication speed on the social networks. They have divided the published content by the user's friends (e. g., posts and links) into different categories (e. g., movies and music) then gathered information to assign that content to a more specific subset (e.g., music subgenres). The system analyzes each text with that content to detect the emotions related to the content and deduce the positive or negative view of the user who published the content. The interest's score to the content is determined based on the characteristics of the user interest and the similarity to the content publisher.

Moreover, it is investigated whether the content should be suggested to the user or not [10].
Social factors investigation shows that at least some users' behavior in social media is affected by social correlation factors.

4. Formal analysis

In OSNs, aggregation of showing platform and the social psychology factors effects make the users' behavior prediction more complicated, as well as individual and environmental factors like users' favorites, the news in the society, and general attention [25]. Other effective factors in addition to the mentioned factors for acceptance or publication of information in social networks are information nature [26], network structure [27,28], impact power, and how does work the social network [29,30], the users' benefits, and the subject's priority and importance [31-33].
The questions in this research include: 1) how much effect do social psychological factors such as social correlations have on the acceptance or not acceptance of an item by the user? 2) Does an item's acceptance due to social relationships make the CF-based algorithms wrong? Hence, we try to model the effect of social correlation or limited attention as an influential factor on the acceptance of an item in the social network using graph theory.
The probability of reception and acceptance of a transmitted item from node u to v can be modeled by a proximity network to calculate the acceptance probability of an item by the user in social networks based on graph theory. We are interested in local approaches, which are only dependent on the neighbors of u and v. It is easy and doesn't need to know any information about the complete graph.
Here, we define the utilized symbols, the common symbols in graph theory at mathematical science based on graph theory book in Table 1 [34].
The acceptance probability of a published post of user u by user v is equal to the probability of going from node u to node v using a random walk algorithm. It is possible by considering parameters in the calculations like the amount of interest in the message content, amount of news about the post, amount of the limited attention of user v to user u, location of the post-presentation, and so on.
The social network graph in the real world is changing and evolving. Therefore, it will become so big and complicated. Hence, finding meaningful proximity criteria and designing fast and low-memory algorithms are critical.

Table 1: The symbols definition

English definition	Mathematical definition	Symbol
a directed graph, or digraph	$G = (V, E)$	G
The node set of a network G	$V\,(G)$	V
The edge set of a network G	$E(G)$	E
Arc from u to v (Edge from node u to node v in digraph called arc (u,v))	$e = (u, v)$	e
the set of all arcs of the form (w,v)		E_v^-
the set of arcs of the form (v,w)		E_v^+
The input neighbourhood of a node v	$N^+(v) = \{x \in V \mid (v, x) \in E_v^+\}$	$N^+(v)$
The output neighbourhood of a node v	$N^-(v) = \{x \in V \mid (x, v) \in E_v^-\}$	$N^-(v)$
The input degree of v	$\deg^+(v) = \lvert E_v^+ \rvert$	$\deg^+(v)$
The output degree of v	$\deg^-(v) = \lvert E_v^- \rvert$	$\deg^-(v)$

A popular method in diagram mining and machine learning, such as scoring in the RSs, is the calculation of proximity between the nodes using a random walk algorithm on the diagrams. The random walk algorithm provides a simple framework for information aggregation from different routes between two nodes [35,36].

In a directed graph, we should move through one of the output edges to another node (like *z*) to go from *u* to *v* and vice versa using a random walk algorithm, which is performed by the probability of $\frac{1}{\deg^{-}(u)}$ according to the number of output edges. Assuming that the current node is *z*, the probability of selecting an appropriate link (edge) to transmit the message to node *v* is equal to $\frac{1}{\deg^{-}(z)}$. Hence, the random walk algorithm can be modeled using equation (1).

$$RandomWalk = \begin{cases} \frac{1}{2}\sum_{z\in\left(N^{-}(u)\cap N^{+}(v)\right)} \frac{1}{\deg^{-}(u)\times\deg^{-}(z)} + \\ \frac{1}{2}\sum_{z\in\left(N^{+}(u)\cap N^{-}(v)\right)} \frac{1}{\deg^{-}(u)\times\deg^{-}(z)} \end{cases} \quad (1)$$

This equation is used to model the applications like the probability of message exchange between users u and v in the social networks using the random walk method.

The attention of the neighbors of u to its published posts can be modeled using equation (2).

$$Limit\ Attention = \sum_{z\in\left(N^{-}(u)\right)} \frac{1}{\deg^{+}(z)} \quad (2)$$

According to the infrastructure of ONS (e.g., Twitter, Facebook, and Instagram) and how to transfer and present the posts in the social network (broadcast for all followers), we can imagine that each user's goal in the network is information presentation for all followers. Moreover, the published posts by node u are published for all followers. In this situation, at first, u transmits the message to all neighbors, including node z, to transmit a message from u to v. Then z rebroadcasts the message. The probability of receiving the message by node v is equal to 1. The common neighbors' proximity method is used to model this proximity (equation 3).

$$Common\ Neighbor = \frac{1}{2}\left|N^{-}(u)\cap N^{+}(v)\right| + \left|N^{+}(u)\cap N^{=}(v)\right| \quad (3)$$

The usual neighbors' proximity method considers simply overlapping and limited attention and has an effective role in the interactions of the epidemic. The equation is written as (3) based on broadcasting and considering limited attention, the equation can be written as (4).

$$Common\ Neighbor\ WithLimit\ Attention = \begin{cases} \frac{1}{2}\sum_{z\in\left(N^{-}(u)\cap N^{+}(v)\right)} \frac{p_u}{\deg^{+}(u)\times\deg^{+}(z)} + \\ \frac{1}{2}\sum_{z\in\left(N^{+}(u)\cap N^{-}(v)\right)} \frac{p_u}{\deg^{+}(u)\times\deg^{+}(z)} \end{cases} \quad (4)$$

5. Investigation

The utilized experimental datasets in this research are described, then the users' limited attention to the items' acceptance is evaluated in section 5.1. The results and diagrams are explained in detail in section 5.2.

5.1. Data Extraction

The aim of this paper is users' limited attention modeling. So, Instagram is the selected social network for complexity reduction of data analysis. In the network, the posts' republishing is so limited, and the users focus on likes and comments on the posts. To this aim, some Instagram users are asked to allow us to follow them for their analysis of the posts. In the request's content, we mentioned that this research does not need their confidential information and the contents of the posts and comments, and only their followers' actions are required. The proposed API by the Facebook company called Instagram Graph API is used for data extraction. In this research, about 3000 accounts (account ID, number of posts, number of followers, bio account, public or private account) have participated. The information includes the ID of more than 28000 published posts by the accounts (each post ID, publication time, number of likes, number of comments, the post location, and the likers' ID) and more than 3000000 followers of the accounts (the account ID) from February 2020 to July 2021. Private information of the users is not used in this research. Moreover, only each post's required information is used without the user ID.

Based on different behaviors of users for various posts, the accounts are divided into public and private accounts, and engagement rate is used instead of the number of likes for data normalization. Figures 1 to 4 show the extracted data dimensions, and tables 2 and 3 present data correlation.

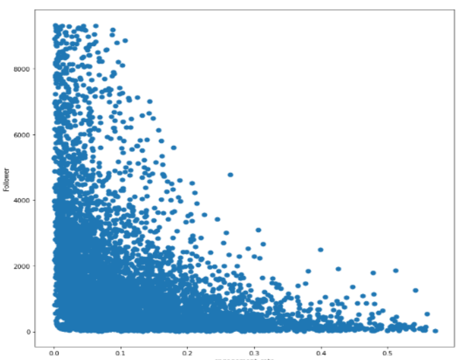


Figure 1: Number of followers vs engagement rate (public account)

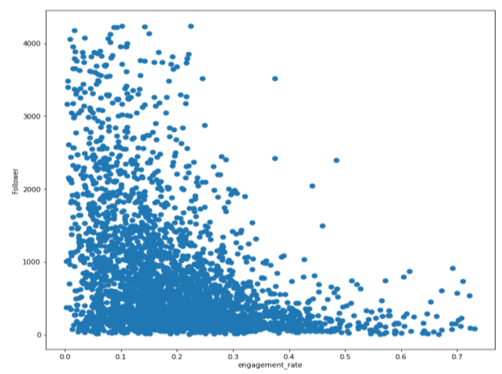


Figure 2: Number of followers vs engagement rate (private account)

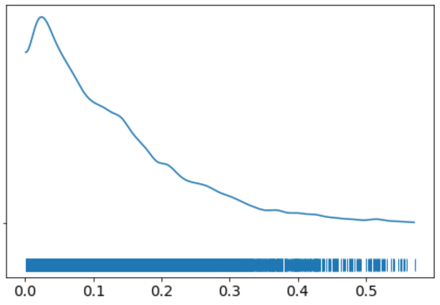


Figure 3: Population distribution based on engagement rate (public account)

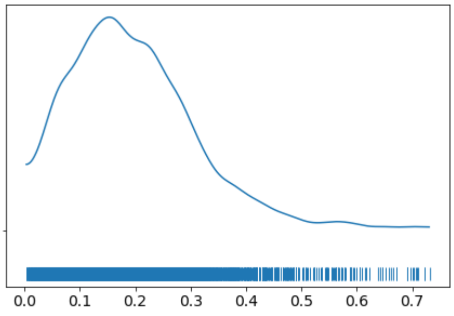


Figure 4: Population distribution based on engagement rate (private account)

Table 2: Parameter's correlation of the private accounts

	Media#	**Follower**	**Following**	**Engagement**
Media #[2]	1.00	0.24	0.22	-0.24
Follower	0.24	1.00	0.51	-0.43
Following	0.22	0.51	1.00	-0.39
Engagement	-0.24	-0.43	-0.39	1.00

Table 3: Parameter's correlation of the public accounts

	Media #	**Follower**	**Following**	**Engagement**
Media #	1.00	0.25	0.02	-0.16
Follower	0.25	1.00	-0.05	-0.06
Following	0.02	-0.05	1.00	-0.39
Engagement	-0.16	-0.06	-0.39	1.00

5.2. Results and Analysis

By determining the extracted data dimension, users' reactions to different posts on Instagram are investigated. An assumed user, X, encounters three types of posts on Instagram. First, recently published posts by X's friends and relatives that are usually for private accounts followed by X. Second, the posts related to the user's favorites that are usually related to the public accounts with a high number of followers followed by X. Third, the posts, which are outputs of the recommender algorithms based on a combination of the X's favorites and X's friends' favorites. Since using the CF-based algorithms in social networks is common, much output of the recommender algorithm is usually the accepted posts by likes or comments of the X's friends. The start of the problem is that the close relationship of X and Y (Y is a following by X) originates from the likes or comments of both users on each other posts. It results in the recommender algorithm suggesting the Y's interested items to X. In this paper, we claim that most of X's likes for the published posts by Y are because of their social relation, not interest in the item's contention. Indeed, a published post on Y's page accepted by X is not because of X's interests but only for their social relationships.

The prediction of whether a user accepts a post by like or comment or not depends on many factors. The most important factor is the user's interest in the published post. Hence, predicting an item acceptance by a user needs much information like the list of the user's interested issues and consistency of the post content with the list. Then we try to show the predictability of the users' reaction to the published posts by their friends. The users' behavior originates from their social relationships, not their interests. Hence, we investigate the reaction of each account's followers to the published posts by the account's user and the predictability of the behavior in the public and private accounts without any knowledge about the message content. To this aim, we extract the likes of about 28000 published posts by about 3000 users on Instagram (about 3000000 records) and investigate them by two recommender algorithms, Baseline and SVD, in two

[2] - Count of media in Instagram account

groups of published posts on public and private pages. We evaluate Baseline and SVD algorithms two times with two different input data from tables 4 and 45.

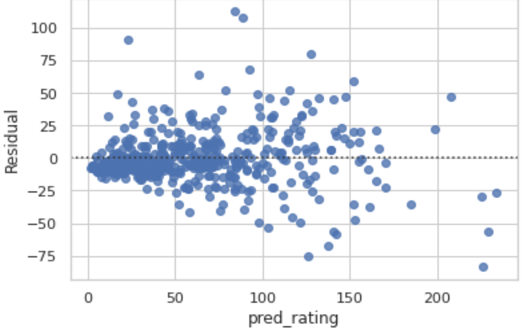
In the first run, input data includes the post ID (Table 4), but the post publisher is not determined. In the second run, the post ID is replaced by the user ID (Table 5). Figures 5 to 12 show the prediction results of the number of likes of the last post of the private and public users using Baseline and SVD algorithms with two different inputs from tables 4 and 5.

Table 4: Data including post ID, follower ID, and posts acceptance or not acceptance

mediaId	FollowerId	Rating
53894	20325	1.0
53892	1249095	0.0
53892	9826734	1.0
53890	32107490	1.0
53890	33959167	0.0

Table 5: Data including user's ID, follower ID, and posts acceptance or not acceptance

userId	FollowerId	Rating
75351	20325	1.0
75351	1249095	0.0
75351	9826734	1.0
75351	32107490	1.0
75351	33959167	0.0



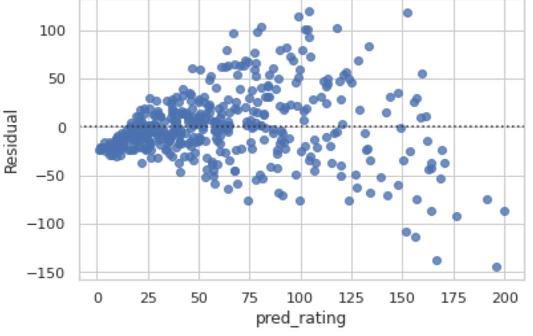


Figure 5: Diagram of residplot or the predicted values of number of likes and residual values or base line algorithm in private account or table 5 entry

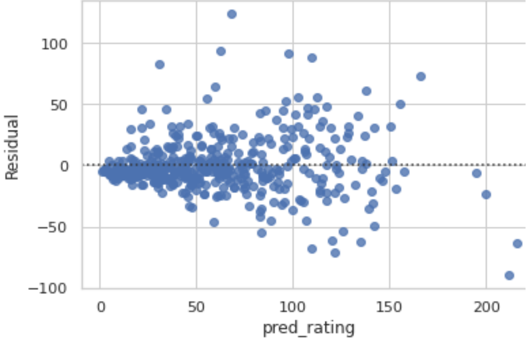


Figure 6: Diagram of residplot[3] or the predicted values of number of likes and residual[4] values or base line algorithm in private account or table 4 entry

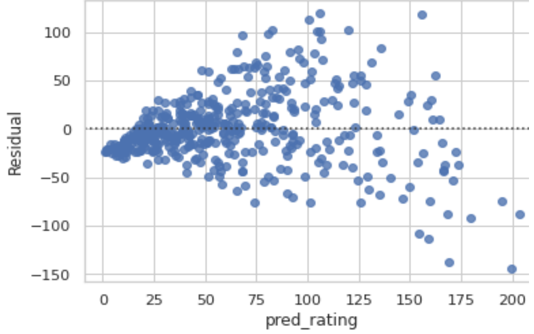


Figure 7: Diagram of residplot or the predicted values of number of likes and residual values or SVD algorithm in private account or table 5 entry

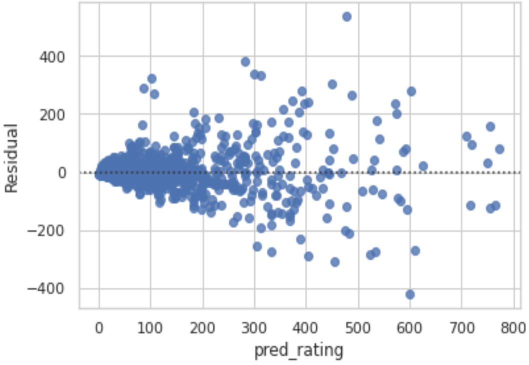


Figure 8: Diagram of residplot or the predicted values of number of likes and residual values or SVD algorithm in private account or table 4 entry

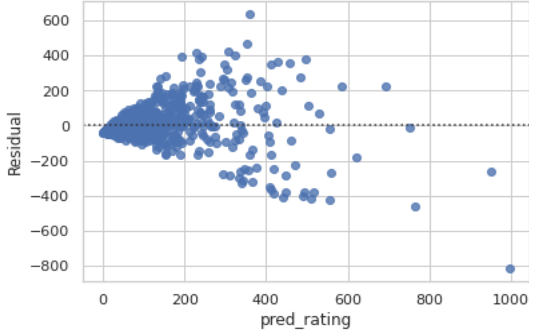


Figure 9: Diagram of residplot or the predicted values of number of likes and residual values or base line algorithm in public account or table 5 entry

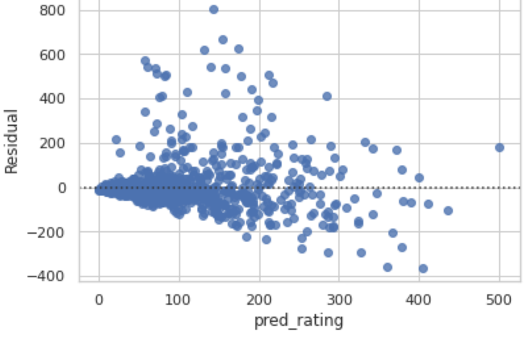
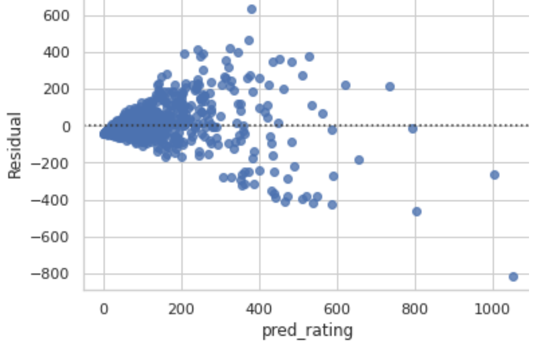


Figure 10: Diagram of residplot or the predicted values of number of likes and residual values or base line algorithm in public account or table 4 entry



[3]Plot the residuals of a linear regression.This function will regress y on x (possibly as a robust or polynomial regression) and then draw a scatterplot of the residuals.

[4]**Residual = Observed – Predicted**

Figure 11: Diagram of residplot or the predicted values of number of likes and residual values or SVD algorithm in public account or table 5 entry

Figure 12: Diagram of residplot or the predicted values of number of likes and residual values or SVD algorithm in public account or table 4 entry

Table 6: Prediction error calculation or two entries of tables 4 and 5 by Baseline and SVD algorithms for private and public accounts

Subject	Input Data	Accunts Type	RMSE Value
BaseLine algorithm	Table 3	private	**39.10**
BaseLine algorithm	Table 4	private	**22.46**
SVD algorithm	Table 3	private	**38.94**
SVD algorithm	Table 4	private	**24.52**
BaseLine algorithm	Table 3	public	**105.13**
BaseLine algorithm	Table 4	public	**108.23**
SVD algorithm	Table 3	public	**104.51**
SVD algorithm	Table 4	public	**112.40**

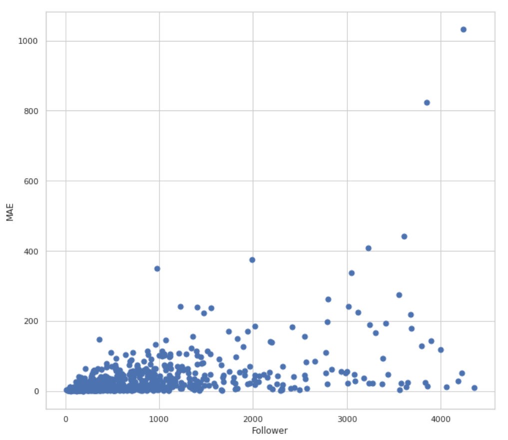


Figure 13. The ratio of MAE to the number of the page's followers shows the prediction precision of the SVD algorithm with data entry of table 5 and it determines that the prediction precision of an item acceptance reduces with the increasing number of followers

Comparing figures 5 and 6 and figures 7 and 8 show that a post's acceptance prediction in private accounts is possible with more precision using the data in table 5, which includes the post publisher information. The comparison is clearer in table 6 based on RMSE calculation (RMSE reduces in the Base Line algorithm from 39.1 to 22.46 and in the SVD algorithm from 38.94 to 24.52). Our inference is that the friends' posts acceptance in virtual networks is performed without considering the intrinsic interest in the post content, and people accept their friends' and relatives' posts (in private accounts) without attention to the posts' content. But in the public accounts, utilizing the same strategy using data in table 5, the predictability error is not better and increased, which is observable by comparing figures 9 and 10 and figures 11 and 12.

The values in table 6 show RMSE prediction errors. It presents that the amount of prediction precision increases in private accounts by data variation using the user ID instead of item ID without any knowledge about the user's interests or the published content. But in the posts of public accounts, the prediction precision decreases by data variation using fixed user ID instead of item ID.

As mentioned previously, the human brain's capacity for stable social relationships is limited, and the increasing number of followers in social networks doesn't show the elimination of this limitation. Based on figure 9, the number of private pages' followers increment results in the reduction of predictability of a post's acceptance by the followers because of a stable reduction of relationships with the increasing number of followers. It is studied in the papers with the subjects of cognitive limitation based on the number of friends and limited brain capacity for social relationships management [7,8].

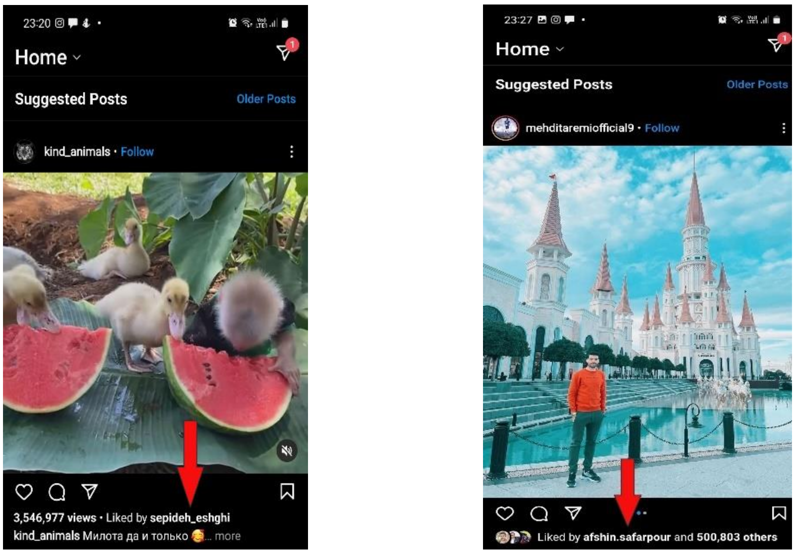
6. Research Findings

Some behaviors with psychological aspects are observed in the social networks because 1) increasing the penetration rate of online social media in different societies makes all relatives of a person in the social network the person's followers, and 2) some limitations such as business or corona pandemic led to low meetings in the real world. Hence, the findings of this research show that the users' behavior toward the published posts by their friends and familiars with private accounts is usually predictable with high precision (more than 75%). It shows that the rating behaviors of the users in their friends' and relatives' private pages essentially don't show the users' intrinsic interests, which are the causes of a post's acceptance predictability without knowing its content. But in public pages, prediction of an item acceptance without seeing its content is impossible. It shows that the users in public accounts accept or rate a post based on their interests. Hence, at least a part of the items' acceptance by a user on social media depends on emotional relationships between the user and the post publisher. It shows that using CF-based algorithms presenting the interests of the user's friends as a recommendation in the suggested list may have some problems. Assuming that user X likes a cat picture published by one of X's relatives, Y, CF recommends cat pictures to X. While X likes the picture only because of X's relationship with Y, not for interest in cat pictures. If the CF-based algorithm considers the X's like for the cat picture as X's interest, it suggests the pictures related to the cat to X that may not be X's interest (picture 1).

We believe that the users accept their close friends' items on online social media due to their emotional and social relations. Hence, the items should not be considered the users' interest. In most existing methods, an item's acceptance by a user is considered the base of a user's interests extraction. Moreover, mostly used CF-based algorithms for recommending an item suggest the items related to the accepted ones due to the social relationships between the post's publisher and other users.

7. Data availability

Data that support the findings of this study have been deposited in google drive by public access on follows link:
https://drive.google.com/drive/folders/1DJIOzD5jEe9IdNczHh3k5ge2r8cGpf9g?usp=sharing



Picture 1: Two examples of a suggested post by Instagram recommending algorithm created due to the social relationship between the user and their followers

8. Conflict of Interest

We undersigned declare that this manuscript is original, has not been published before, and is not currently being considered for publication elsewhere. We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us. We understand that the Corresponding Author is the sole contact for the Editorial process.

References

[1] Yang, M., Zhang, S., Zhang, H., & Xia, J. (2019). A new user behavior evaluation method in online social network. Journal of Information Security and Applications, 47, 217–222. https://doi.org/10.1016/j.jisa.2019.04.008

[2] Kang, J.-H., & Lerman, K. (2013). LA-CTR: A limited attention collaborative topic regression for social media. In arXiv [cs.IR]. http://arxiv.org/abs/1311.1247

[3] Agarwal, C. (2016). Recommender Systems. Springer International Publishing.

[4] Chua, F. C. T., Lauw, H. W., & Lim, E.-P. (2013). Generative models for item adoptions using social correlation. IEEE Transactions on Knowledge and Data Engineering, 25(9), 2036–2048. https://doi.org/10.1109/tkde.2012.137

[5] Rodríguez, N. D., Cuéllar, M. P., Lilius, J., & Calvo-Flores, M. D. (2014). A survey on ontologies for human behavior recognition. ACM Computing Surveys, 46(4), 1–33. https://doi.org/10.1145/2523819

[6] Bianca, B. L. (2018). The user behavior analysis based on text messages using parafac and block term decomposition. International Journal of Advanced Computer Science and Applications: IJACSA, 9(10). https://doi.org/10.14569/ijacsa.2018.091007

[7] Hodas, N. O., & Lerman, K. (2012a). How visibility and divided attention constrain social contagion. 2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing.

[8] Dunbar, R. I. M., & Shultz, S. (2007). Evolution in the social brain. Science (New York, N.Y.), 317(5843), 1344–1347. https://doi.org/10.1126/science.1145463

[9] Buscher, G., Cutrell, E., & Morris, M. R. (2009). What do you see when you're surfing: Using eye tracking to predict salient regions of web pages. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.

[10] Aivazoglou, M., Roussos, A. O., Margaris, D., Vassilakis, C., Ioannidis, S., Polakis, J., & Spiliotopoulos, D. (2020). A fine-grained social network recommender system. Social Network Analysis and Mining, 10(1). https://doi.org/10.1007/s13278-019-0621-7

[11] Campana, M. G., & Delmastro, F. (2017). Recommender systems for online and mobile social networks: A survey. Online Social Networks and Media, 3–4, 75–97. https://doi.org/10.1016/j.osnem.2017.10.005

[12] Agarwal, D., & Chen, B.-C. (2010). fLDA: Matrix factorization through latent dirichlet allocation. Proceedings of the Third ACM International Conference on Web Search and Data Mining - WSDM '10.

[13] Bokde, D., Girase, S., & Mukhopadhyay, D. (2015). Matrix factorization model in collaborative filtering algorithms: A survey. Procedia Computer Science, 49, 136–146. https://doi.org/10.1016/j.procs.2015.04.237

[14] Purushotham, S., Liu, Y., & Kuo, C.-C. J. (2012). Collaborative topic regression with social matrix factorization for recommendation systems. In arXiv [cs.IR]. http://arxiv.org/abs/1206.4684

[15] Sun, Z., Han, L., Huang, W., Wang, X., Zeng, X., Wang, M., & Yan, H. (2015). Recommender systems based on social networks. The Journal of Systems and Software, 99, 109–119. https://doi.org/10.1016/j.jss.2014.09.019

[16] Li, Z., Xiong, F., Wang, X., Chen, H., & Xiong, X. (2019). Topological influence-aware recommendation on social

networks. Complexity, 2019, 1–12. https://doi.org/10.1155/2019/6325654

[17] Fan, W., Ma, Y., Yin, D., Wang, J., Tang, J., & Li, Q. (2019). Deep social collaborative filtering. Proceedings of the 13th ACM Conference on Recommender Systems.

[18] Arnaboldi, V., Campana, M. G., Delmastro, F., & Pagani, E. (2016). PLIERS: A popularity-based recommender system for content dissemination in online social networks. Proceedings of the 31st Annual ACM Symposium on Applied Computing.

[19] Wikipedia contributors. (2020, November 15). Proximity principle. Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Proximity_principle&oldid=988780582

[20] Clore, G. L., & Byrne, D. (1974). A reinforcement-affect model of attraction. In Foundations of Interpersonal Attraction (pp. 143–170). Elsevier.

[21] McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. Annual Review of Sociology, 27(1), 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

[22] Hyon, R., Kleinbaum, A. M., & Parkinson, C. (2020). Social network proximity predicts similar trajectories of psychological states: Evidence from multi-voxel spatiotemporal dynamics. NeuroImage, 216(116492), 116492. https://doi.org/10.1016/j.neuroimage.2019.116492

[23] Gonçalves, B., Perra, N., & Vespignani, A. (2011). Modeling users' activity on twitter networks: validation of Dunbar's number. PloS One, 6(8), e22656. https://doi.org/10.1371/journal.pone.0022656

[24] Dunbar, R. I. M. (2016). Do online social media cut through the constraints that limit the size of offline social networks? Royal Society Open Science, 3(1), 150292. https://doi.org/10.1098/rsos.150292

[25] Yin, H., Cui, B., Chen, L., Hu, Z., & Huang, Z. (2014). A temporal context-aware model for user behavior modeling in social media systems. Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data.

[26] Romero, D. M., Meeder, B., & Kleinberg, J. (2011). Differences in the mechanics of information diffusion across topics: Idioms, political hashtags, and complex contagion on twitter. Proceedings of the 20th International Conference on World Wide Web - WWW '11.

[27] Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012). The role of social networks in information diffusion. Proceedings of the 21st International Conference on World Wide Web - WWW '12.

[28] Weng, L., Menczer, F., & Ahn, Y.-Y. (2013). Virality prediction and community structure in social networks. Scientific Reports, 3(1), 2522. https://doi.org/10.1038/srep02522

[29] Bakshy, E., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). Everyone's an influencer: Quantifying influence on twitter. Proceedings of the Fourth ACM International Conference on Web Search and Data Mining - WSDM '11.

[30] Romero, D. M., Galuba, W., Asur, S., & Huberman, B. A. (2011). Influence and passivity in social media. Proceedings of the 20th International Conference Companion on World Wide Web - WWW '11.

[31] Agarwal, D., & Chen, B.-C. (2009). Regression-based latent factor models. Proceedings of the 15th ACM SIGKDD

International Conference on Knowledge Discovery and Data Mining - KDD ’09.

[32] Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. Computer, 42(8), 30–37. https://doi.org/10.1109/mc.2009.263

[33] Yu, K., Lafferty, J., Zhu, S., & Gong, Y. (2009). Large-scale collaborative prediction using a nonparametric random effects model. Proceedings of the 26th Annual International Conference on Machine Learning - ICML ’09.

[34] Harris, J., Hirst, J. L., & Mossinghoff, M. (2008). Combinatorics and Graph Theory. Springer New York.

[35] Sarkar, P., & Moore, A. W. (2011). Random walks in social networks and their applications: A survey. In Social Network Data Analytics (pp. 43–77). Springer US.

[36] Zhao, J., Wang, P., Lui, J. C. S., Towsley, D., & Guan, X. (2019). Sampling online social networks by random walk with indirect jumps. Data Mining and Knowledge Discovery, 33(1), 24–57. https://doi.org/10.1007/s10618-018-0587-5