Semiconductor Optoelectronics, *Vol. 41 No. 06 (June 2022), 1089-1098*
https://bdtgd.cn/

Semiconductor Optoelectronics ISSN: 1001-5868

# REVIEW OF RECOMMENDER SYSTEM IN SOCIAL NETWORKS BASED ON CLUSTERING METHOD

**Pooja Rajendra Jayswal**

**Dr. Manav Thakur**
Department of Computer Science, Malwanchal University, Indore

## ABSTRACT

"Review of Recommender System in Social Networks based on Clustering Method" provides a brief overview of the contents of the article. The paper discusses the use of clustering techniques in the development of recommender systems for social networks. The authors provide a review of the different clustering algorithms and techniques used in recommender systems, including K-Means, DBSCAN, and hierarchical clustering. They also discuss the various evaluation metrics used to measure the performance of these systems. The paper emphasizes the importance of clustering algorithms in the development of effective recommender systems. The authors suggest that clustering can help to identify groups of users with similar preferences and interests, which can then be used to make personalized recommendations. The paper concludes with a discussion of the challenges and future directions of recommender systems in social networks. the abstract provides a clear and concise summary of the key points discussed in the paper, highlighting the significance of clustering in the development of recommender systems for social networks.

## INTRODUCTION

The introduction of the paper "Recommender System in Social Networks based on Clustering Method" provides an overview of the increasing popularity of social networks and the need for personalized recommendations in these networks. The authors highlight the challenge of providing relevant recommendations to users in social networks due to the large volume of available data and the diverse interests and preferences of users. The paper proposes the use of clustering methods in the development of recommender systems for social networks. Clustering algorithms can help to identify groups of users with similar interests and preferences, which can then be used to make personalized recommendations. The authors provide a review of different clustering techniques, including K-Means, DBSCAN, and hierarchical clustering, and discuss their strengths and limitations in the context of recommender systems. The authors also discuss various evaluation metrics used to assess the performance of recommender systems, such as precision, recall, and F1 score. They suggest that the choice of evaluation metric should depend on the specific application of the recommender system.

## RECOMMENDER SYSTEM IN SOCIAL NETWORKS

Many computer-mediated technologies are employed for online communication in today's era of rapidly expanding social media. These developments are the highways for public and private message exchanges in online communities and networks; they also serve as conduits for user-generated content, user interaction, the broadcast of news about real-world events, and creation and sharing of user-generated content. Blogs, business networks (e.g. e-commerce), enterprise social networks, CQA forums, microblogs, photo-sharing portals, product review portals, social bookmarking, social gaming, social networks, video-sharing, and virtual worlds are all examples of social media technology. These developments are widely used to facilitate communication between various communities all around the globe.

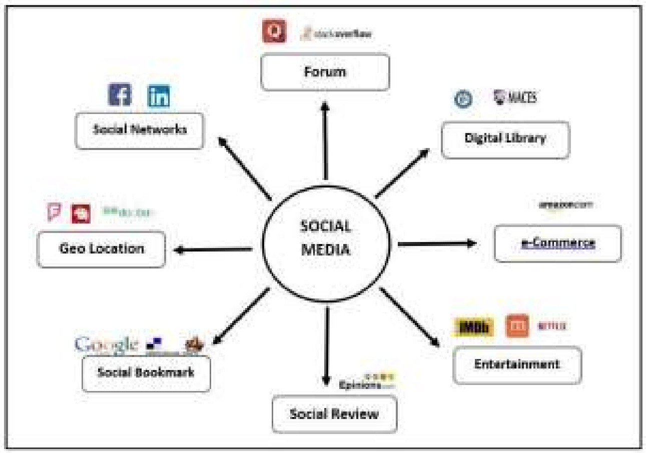


**Figure 1: Categories of social media**

A social media website is one that "facilitates meeting people, finding similar minds, communicating and sharing content, and developing community," and it may or may not promote or allow for a wide range of activities, including commercial ones as well as social ones or a mix of the two. Categories of social media include digital libraries, online stores, shows, discussion boards, maps, locations, bookmarking, reviews, games, and networks. Subset of social media, a social network is an online community of people who share some sort of interest (Figure 1). To communicate with others, people often use social media platforms accessible via various electronic devices, including smartphones and personal computers. These innovations produce highly collaborative environments where people and groups may discuss, rate, comment on, and shape online content created by others. Innovations like this facilitate interaction between companies, groups, communities, and individuals. There has been a significant uptick in the research and development of social media technologies, which alter the ways in which both individuals and huge institutions communicate.

The following are examples of social media categories commonly used to connect people.

Digital Library: A digital library is a collection of electronic documents that can be found on the Internet. Depending on the library, a user can look at magazine articles, books, papers, pictures, sound files, and videos. Many groups have set up digital libraries as an effective way for universities, museums, and some city libraries to share information.

Forums: In forums like Stack Overflow and Yahoo! Answers, users interact with each other by posting (asking questions and getting other people's opinions) and talking to each other. In the context of blog comments, which are comments made by readers at the end of a blog post, the interactions centre on the topic of the post in question.

E-commerce: Sites like Amazon, which allow people to buy and sell goods and services online, encourage social interaction and user contributions.

Entertainment: Live video streaming, movies, video chat chats, and the streaming of music and films are just some of the interactive features and content made possible by today's entertainment websites. Movie Lens, MUADDIB, IMDB, Smart TV, Netflix, Flixster, and the Jester dataset are all examples of this category (jokes).

Blog/Microblog: Twitter is a type of blog or microblog in which users can update their followers on whatever they're up to. These updates are brief and frequently have a word count cap.

Social Network: Social networking services like Facebook and professional networking sites like LinkedIn give users the opportunity to create online profiles and connect with people virtually. In order to communicate with one another and share information, users tap into the network's various functionalities.

Geolocation: Users of a geolocation software like Foursquare can update their social network on their current whereabouts and earn virtual badges by "checking in" to various locations.

Social Games: As contrast to playing games alone, online games encourage or compel players to communicate with one another. Games like pogo, cards, and board games are great for getting people together.

Social Bookmarking: Users can bookmark and organise their favourite websites and webpages using services like Delicious, CiteULike with Google Scholar, and BibSonomy. Such platforms offer a variety of tools for link management and public sharing.

Social Review: Websites in this class include Epinions, which is a popular example of a broad consumer review site where visitors can research a wide range of products by reading both recent and archived evaluations written by other users.

Others (publish, share, and discuss): Here, people can connect with one another through their profiles, messages, and comments on various social media platforms. In addition, this variety lets users broadcast their own material (images and videos). Some websites that fall within this category are YouTube, Flickr, Skype, Gchat, and Wikipedia.

Today, people exchange vast amounts of data, photographs, videos, materials, and documents on social networking platforms. Users now face the new challenge of information overload as a result of the widespread dissemination of data. As a result, several social media platforms employ RSs to circumvent the problem and provide relevant recommendations to specific people. In depth studies of RSs have been conducted since the mid-1990s. Books, movies, music, products, and TV shows can all be suggested with the use of RSs. User feedback in the form of ratings and reviews is analysed by RSs. With the use of effective and accurate recommendation mechanisms built into the systems, the RSs of academic scientific and online libraries help users acquire relevant and trustworthy recommendations, going above and beyond simple catalogue searches. Time and money saved by users, as well as provider procedure, quality, and decision-making strategy, can all be enhanced through RSs.

Recommender systems (RSs) in social media produce helpful article and item recommendations that promote user cooperation. Many people's lives are documented and archived on social media sites like Facebook, Twitter, and blogs. The goal of social networks is to facilitate communication among friends, and they also provide insightful data about users' interests. It is necessary to process the noisy and varied materials shared on these networks before useful information can be extracted from them. Individuals highlight their expertise and share their thoughts in an online forum. The quantity of goods available on the web is expanding rapidly. As a result, it becomes challenging for specialists to identify products within a certain field. RS is widely deployed on social media platforms as a means of solving this issue by pinpointing the exact users who are most likely to be the neighbours of a certain target audience. A user can get recommendations for products they might like based on what their neighbours have liked or shared.

**LITERATURE REVIEW**

Sun, Z., Han, L. et al,(2015) Many academics over the past decade have looked into conventional recommendation systems, particularly collaborative filtering recommendation systems. On the other hand, they don't take into account consumers' interpersonal connections. In fact, these connections can enhance suggestion precision. Research into social-based recommender systems has exploded in recent years. To improve recommender systems, we present a social regularisation strategy that makes use of data collected from social networks. The user-item matrix's missing values (tags) are predicted using information gleaned from users' social networks and rating histories. In particular, we employ a biclustering technique to zero in on the best set of pals for making various last-ditch suggestions. Analyses of real-world datasets prove the suggested method outperforms the state-of-the-art techniques.

Ahmadian, S., Joorabloo, N. et al,(2018) The purpose of a recommender system is to help consumers find what they're looking for by filtering through a big catalogue of options. They've found use in fields as diverse as e-commerce and education and digital health. Yet, recommender systems can benefit from clustering algorithms that assist them place users into relevant groups that are then used as neighbours for the prediction process. Traditional clustering methods ignore the reality that consumers' tastes change over time. In this research, we propose a social recommender system based on a temporal clustering strategy to solve this issue. In particular, the suggested method takes into account both the social information among users and the temporal information of evaluations provided by users on objects. Results from experiments conducted on a benchmark dataset demonstrate that the suggested method yields much higher quality suggestions than the state-of-the-art methods in terms of both accuracy and coverage measures.

Anandhan, A., Shuib, L. et al,(2018) Over the past few years, numerous iterations of the recommender system for conducting reviews have been developed (RS). Recommender Systems (RSs) are created using ratings, reviews, and comparisons provided by actual users. Social media resources (RSs) such blogs, forums, social network sites, social bookmarking sites, video portals, and chat portals facilitate efficient user collaboration. Content, articles, news, e-commerce products, and users are all recommended using RS's access to social media resources. Despite a growing body of research on the role of social media in RSs each year, there is still need for improvement in the way this literature is reviewed and organised. By utilising a methodological decision analysis in six aspects, this study aims to provide a comprehensive review of social media RS on research articles published between 2011 and 2015, including the recommendation approaches, research domains, datasets used within each domain, data mining techniques, recommendation type, and the use of performance measures. From the first 434 RS research articles found in Web of Science and Scopus between 2011 and 2015, 61 are discussed here. In order to achieve the goal of this study, a thorough evaluation and analysis of extracted articles exploring different recommendation approaches utilised in

RS were undertaken.

Amato, F., Moscato, V., et al,(2019) Recent years have seen a meteoric rise in the popularity of online social networks. In particular, people can connect, share, comment on, and view a wide variety of multimedia material via social media networks. There is a tremendous quantity of data generated by this event, and it exhibits characteristics of Big Data, such as a high variation rate, a vast volume, and inherent heterogeneity. For this reason, Recommender Systems have been developed over the past decade to facilitate the exploration of large data repositories, guiding users to "what they actually need" in this sea of data. In this study, we propose and detail a new kind of recommendation engine for big data applications, one that can make suggestions based on the interactions of users and the multimedia material they generate across several social media platforms. A "user-centered" strategy was used to design the system being suggested. To evaluate the suggested method and display how it might achieve extremely promising outcomes, an experimental campaign was conducted using data from numerous social media networks.

Wang, Z., Liao, J.et al,(2014) Nowadays, social networking platforms make friend suggestions to users based on their social graphs, which may not accurately reflect the user's actual tastes when it comes to selecting friends. In this research, we introduce Friendbook, a novel semantic-based friend recommendation system for social networks that makes suggestions about potential friends based on the users' interests and lifestyles rather than their social connections. Using the wealth of information available from the user's smartphone's sensors, Friendbook is able to learn about the user's preferences and habits, calculate the degree to which those habits are comparable to those of other users, and then recommend friends with similar habits. Our approach is inspired by text mining, and we represent a user's life as a set of life documents from which his or her life styles can be derived using the Latent Dirichlet Allocation algorithm. We also suggest a similarity metric to assess the degree to which users' lifestyles are similar, and we use a friend-matching network to determine how much of an impact individual users have in terms of lifestyles. When a user makes a request, Friendbook sends them a list of their most recommended friends. At last, Friendbook has a feedback mechanism to increase the precision of its recommendations. We have deployed Friendbook on Android-powered mobile devices and analysed its performance in lab settings and through extensive modelling. The findings demonstrate that the suggestions accurately match users' tastes when selecting companions.

Davoodi, E., Kianmehr, K. et al,(2013) This study lays out the groundwork for developing a hybrid expert recommendation system by combining the best features of content-based recommendation algorithms with those of a collaborative filtering system based on social networks. The proposed method takes into account the social component of experts' behaviour in an effort to improve the accuracy of recommendation prediction. To achieve this goal, we first build content-based profiles of experts by searching the web for relevant information. Using the foundation of knowledge found in the Wikipedia database, a semantic kernel is constructed.

Zheng, X., Luo, Y., et al,(2018) More and more people are open to sharing their stories and opinions online because to the rise of social media. But as the amount of data stored in networks continues to expand at an exponential rate, the resulting information glut becomes increasingly difficult to manage. To address this issue, a recommender system is useful. Although social network research models and algorithms have been the primary focus of recommender system studies to date, the underlying social network structure of these systems has not been explored in depth, and the so-called cold start problem has yet to be adequately addressed. In this study, we suggest a new type of hybrid recommender system, the Hybrid Matrix Factorization (HMF) model, which employs hypergraph topology to describe and analyse the internal relationship of the social network.

Zhang, D., Hsu, C. H. et al,(2013) In order to provide users with content that they are more likely to find interesting, social recommender systems employ collaborative filtering (CF). There is a diverse range of CF plans that have been proposed. Most, though, aren't equipped to handle the "cold-start dilemma," which occurs when a social media platform's recommendations fail to bring in fresh content, users, or both. Furthermore, they think that all evaluations are equally important to the social media recommendation. This theory contradicts the reality that basic ratings aren't very helpful when it comes to recommending content that readers would enjoy. As such, we suggest a novel approach to the cold-start problem using BiFu approaches in a cloud computing environment. It introduces the notions of popular things and regular raters as a means of locating the rating sources for suggestion. BiFu uses the bi-clustering technique to lower the dimensionality of the rating matrix.

**CLUSTERING ALGORITHMS**

Knowing the total number of clusters is a prerequisite for several clustering techniques. In either instance, as can be seen in Fig. 2, the algorithm attempts to group the data into the specified number of clusters. Such a clustering method is described in Section 3.1 as the K-means and Fuzzy C-means methods. In other cases, the algorithm starts by locating the largest cluster, then moves on to the next largest cluster, and so on, without needing to know the total number of clusters in advance. Types include mountain clustering and subtractive clustering. In both circumstances, it is possible to apply the problem of known cluster numbers; however, K-means and Fuzzy C-means clustering cannot be employed if the number of clusters is unknown.

Furthermore, clustering methods can be used either online or offline. Each input vector is utilised to adjust the cluster centres in real time during online clustering. When fresh data is added to the system, it is able to determine where the cluster hubs are. The system is given a training dataset in offline mode and uses it to locate cluster centres by examining each input vector in the dataset. New input vectors will be classified once the cluster centres have been determined. The procedures detailed here are asynchronous and can be used independently of the internet. These four methods are briefly discussed.

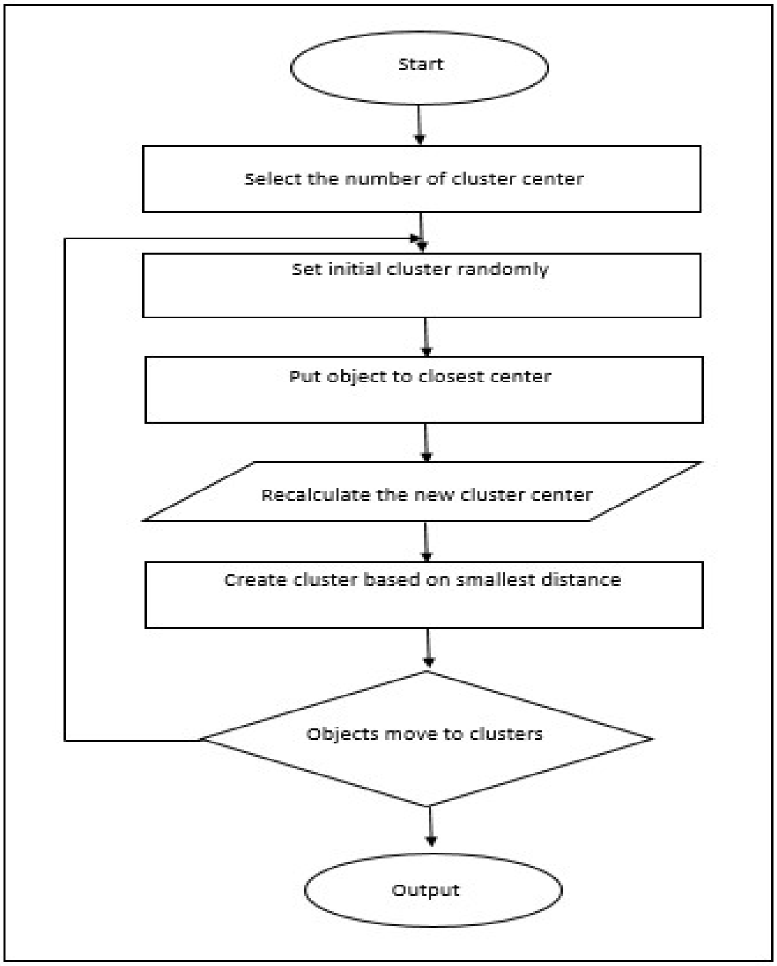


Fig. 2 Flow chart of clustering algorithm

Data points are aggregated using the machine learning technique of clustering. A clustering technique allows us to assign each data point in a series to a distinct category. Features and qualities of data items within the same category may be the same, while data points belonging to distinct categories may have vastly different features and/or characteristics. Clustering is an unregulated research process and is a popular methodology in various fields for the study of statistical results.

**K-Means Clustering**

K-means clustering is a type of unregulated learning used in data without a given group or category. Clustering is not a clustering process. The objective of this algorithm is to find groups in the data that represent the number of groups in variable K. The algorithm works by assigning every data point, based on the features provided, to one of the K groups. Cluster data points dependent on the similarities of features. The results of the algorithm K-means are as follows:

- K cluster centroids that can be used for marking new results.
- Training data labels (one cluster is allocated to each data point).

The clustering allows you to find and analyze the groups that form organically rather than define groups before looking into the data. The following section "Choosing K" describes how to determine the number of groups. A set of function values that identify the resulting classes are a selection of each center of the cluster. The examination of the centroid weights can be used to interpret qualitatively the type of group represented by each group.

**Working principle**

The K-means cluster, or Hard-C-means clustering, is an algorithm based on finding data cluster in a data set, so that dissimilarity (order) measurements have a minimal cost function (or objection function). In most cases, this difference is chosen as the distance from the Euclides. N vectors set, x j = 1, ......, n, must be divided in c groups, ii= 1,2, ..... c. c. According to the Euclidean distance between a vector k x of group j and the cluster center ci, the cost function can be defined by:

$$J = \sum_{i=1}^{c} J_i = \sum_{i=1}^{c} \left( \sum_{k, x_k \epsilon G_i} \| x_k - c_i \|^2 \right)$$

Where $J_i = \sum_{k, x_k \epsilon G_i} \| x_k - c_i \|^2$ is the cost function of group i. The group partitioned is defined by a binary c*n membership matrix U where the element uij is 1 where the jth data point xj is in group I and 0. The minimum uij for equation (3.1) can be defined as follows when the cluster centers are established:

$$u_{ij} = \begin{cases} 1 \; if \;\; \left\| x_j - c_t \right\|^2 \; \leq \; \left\| x_j - c_k \right\|^2, \;\; for\; each\; k \; \neq \; i\,, \\ \qquad 0 \; otherwise. \end{cases}$$

That means that xj is group i if ci is the center nearest to all centres. In the other hand, the ideal center ci that minimizes equation is the mean of all vectors in group i when the membership matrix is fixed, i.e. when uij is fixed,

**Challenges of Clustering-Based Recommender Systems**

Despite the benefits of clustering-based recommender systems, they also face several challenges, including:

Choice of clustering algorithm: The choice of clustering algorithm can be challenging, as it depends on the characteristics of the data and the specific requirements of the system. The performance of the clustering algorithm can have a significant impact on the accuracy of the recommendations generated.

Sensitivity to initial conditions: Clustering algorithms are sensitive to initial conditions, which can affect the final clustering results. This can lead to inconsistency and instability in the generated recommendations, as the clustering results may vary depending on the initialization.

Lack of transparency: Clustering algorithms can be complex and difficult to interpret, which can make it challenging to explain how recommendations are generated to users. This lack of transparency can lead to decreased user trust and engagement with the system.

Computational complexity: Clustering-based recommender systems can be computationally complex, particularly for large datasets, which can make it challenging to generate recommendations in real-time. This can lead to delays in generating recommendations, which can impact the user experience.

Scalability: While clustering-based recommender systems can handle large datasets more effectively than other recommendation approaches, they can still face scalability issues for extremely large datasets. The computational resources required to generate recommendations can increase as the size of the dataset increases.

Cold-start problem: Clustering-based recommender systems can face a cold-start problem when new users or items are introduced to the system. Without sufficient data about the new user or item, the system may struggle to generate accurate recommendations.

clustering-based recommender systems face several challenges, including the choice of clustering algorithm, sensitivity to initial conditions, lack of transparency, computational complexity, scalability issues, and the cold-start problem. Addressing these challenges requires developing more accurate and efficient clustering algorithms, improving transparency and user trust, and optimizing the use of computational resources. Future research can focus on addressing these challenges to enhance the accuracy and effectiveness of clustering-based recommender systems.

## CONCLUSION

Clustering-based recommender systems consist of three main components: data preprocessing, clustering, and recommendation generation. In data preprocessing, the data is collected and cleaned, and it is transformed into a user-item matrix. In clustering, the data is grouped into clusters based on the characteristics of the data using clustering algorithms such as k-means, hierarchical clustering, or density-based clustering. In recommendation generation, recommendations are generated based on the preferences of similar users in the same cluster. Clustering-based recommender systems offer several benefits, including improved accuracy, better scalability, addressing the challenge of data sparsity, and providing more personalized recommendations. However, these systems also face challenges, including the choice of clustering algorithm, sensitivity to initial conditions, lack of transparency, and computational complexity. The future potential of clustering-based recommender systems is significant, particularly as social networks continue to generate enormous amounts of user-generated data. Future research can focus on developing more accurate and efficient clustering algorithms that can address the challenges faced by clustering-based recommender systems, such as sensitivity to initial conditions, computational complexity, and lack of transparency. Additionally, hybrid recommender systems that combine clustering-based recommender systems with other recommendation approaches, such as collaborative filtering and content-based filtering, can lead to improved accuracy and effectiveness in personalized content recommendations.

## REFERENCES


1. Sun, Z., Han, L., Huang, W., Wang, X., Zeng, X., Wang, M., & Yan, H. (2015). Recommender systems based on social networks. Journal of Systems and Software, 99, 109-119.
2. Ahmadian, S., Joorabloo, N., Jalili, M., Meghdadi, M., Afsharchi, M., & Ren, Y. (2018, August). A temporal clustering approach for social recommender systems. In 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) (pp. 1139-1144). IEEE.
3. Anandhan, A., Shuib, L., Ismail, M. A., & Mujtaba, G. (2018). Social media recommender systems: review and open research issues. IEEE Access, 6, 15608-15628.
4. Amato, F., Moscato, V., Picariello, A., & Piccialli, F. (2019). SOS: a multimedia recommender system for online social networks. Future generation computer systems, 93, 914-923.
5. Wang, Z., Liao, J., Cao, Q., Qi, H., & Wang, Z. (2014). Friendbook: a semantic-based friend recommendation system for social networks. IEEE transactions on mobile computing, 14(3), 538-551.
6. Davoodi, E., Kianmehr, K., & Afsharchi, M. (2013). A semantic social network-based expert recommender system. Applied intelligence, 39, 1-13.

7. Zheng, X., Luo, Y., Sun, L., Ding, X., & Zhang, J. (2018). A novel social network hybrid recommender system based on hypergraph topologic structure. World Wide Web, 21, 985-1013.
8. Zhang, D., Hsu, C. H., Chen, M., Chen, Q., Xiong, N., & Lloret, J. (2013). Cold-start recommendation using bi-clustering and fusion for large-scale social recommender systems. IEEE Transactions on Emerging Topics in Computing, 2(2), 239-250.
9. Esslimani, I., Brun, A., & Boyer, A. (2011). Densifying a behavioral recommender system by social networks link prediction methods. Social Network Analysis and Mining, 1, 159-172.
10. Meleshko, Y., Yakymenko, M., & Semenov, S. (2021). A Method of Detecting Bot Networks Based on Graph Clustering in the Recommendation System of Social Network. In COLINS (pp. 1249-1261).
11. Xu, G., Wu, Z., Zhang, Y., & Cao, J. (2015). Social networking meets recommender systems: survey. International Journal of Social Network Mining, 2(1), 64-100.
12. Shokeen, J., & Rana, C. (2020). Social recommender systems: techniques, domains, metrics, datasets and future scope. Journal of Intelligent Information Systems, 54(3), 633-667.
13. Quoc V. Le and Tomas Mikolov., "Distributed Representations of Sentences and Documents", ICML - JMLR Workshop and Conference Proceedings, vol. 32, pp. 1188–1196, 2014.
14. X. Li, M. Wang, and T.-P. Liang, "A multi-theoretical kernel-based approach to social network-based recommendation", Decision Support Systems., vol. 65, pp. 95–104, Sep 2014.