AI & SOCIETY (2020) 35:469–483
https://doi.org/10.1007/s00146-019-00893-z

STUDENT FORUM

Presenting a hybrid model in social networks recommendation system architecture development

Abolfazl Zare[1] · Mohammad Reza Motadel[1] · Aliakbar Jalali[2]

Received: 21 September 2018 / Accepted: 6 May 2019 / Published online: 18 May 2019

Abstract
There are many studies conducted on recommendation systems, most of which are focused on recommending items to users and vice versa. Nowadays, social networks are complicated due to carrying vast arrays of data about individuals and organizations. In today's competitive environment, companies face two significant problems: supplying resources and attracting new customers. Even the concept of supply-chain management in a virtual environment is changed. In this article, we propose a new and innovative combination approach to recommend organizational people in social networks based on organizational communication and SCM. The proposed approach uses a hybrid strategy that combines basic collaborative filtering and demographic recommendation systems, using data mining, artificial neural networks, and fuzzy techniques. The results of experiments and evaluations based on a real dataset collected from the LinkedIn social network showed that the hybrid recommendation system has higher accuracy and speed than other essential methods, even substantially has eliminated the fundamental problems with such systems, such as cold start, scalability, diversity, and serendipity.

Keywords Recommendation systems · Collaborative filtering · Artificial neural network · Fuzzy logic · Supply-chain management · Social networks

Abbreviations

RS	Recommendation systems
SCM	Supply-chain management
ANN	Artificial neural networks
CF-RS	Collaborative filtering RS
D-RS	Demographic RS
CB-RS	Content-based RS
DM	Data mining
ED	Euclidean distance
MF	Membership function
MAE	Mean absolute error
RMSE	Root mean squared error
FPR	False-positive rate
TPR	True-positive rate
OM	Overlap measure

✉ Abolfazl Zare
Zare.Phd@gmail.com

Mohammad Reza Motadel
Dr.Motadel@Gmail.com

Aliakbar Jalali
DrJalali@UMD.Edu

1 Department of Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran

2 Department of Computer Science, University of Maryland, College Park, MD, USA

1 Introduction

Today, in the age of information, the development of web applications and social networks has led to the gathering of vast amounts of data from all around the world. As a result, it is difficult to find information and make decisions, efficiently. Despite the advancement of techniques and methods for managing information, the discovery of useful and appropriate knowledge from among the enormous quantities of data is still a serious and challenging issue. Much research has been done in this area, and more are still underway.

One of the useful information management techniques is recommendation systems. Research into the development of RS focuses solely on the recommendation of items to users and vice versa, in a variety of applications. The challenge is that even though billions of different users exist in various social networks, a few researches have been done individually on recommending people to people in RS, based on organizational communications. Demographic recommendation systems use demographic information to recommend people items. These systems are in significant trouble. Massive data volumes, dispersion of ratings, incomplete and inaccurate information have all caused significant issues such as scalability, cold start, confidence, and trust. These issues have become the subjects of extensive research.

Social networks are complex environments of relationships among individuals, companies, and other entities. Today's companies are faced with two major challenges: supplying resources and attracting new customers. The supply-chain management is based on relationships, and communications in the SCM are shaped by the customer relationship management (CRM) process of the seller and the supplier relationship management (SRM) process of the buyer (Singh and Sohani 2011). Supply-chain managers reduce the overall cost of services and materials. They provide the margin of safety for themselves and suppliers by using close relationships with the leading suppliers (Chopra and Meindl 2007).

The present study is to find how to present a hybrid recommendation system that can recommend relevant organizational people via the use of a social network. In addition to high precision, this approach must be able to solve common problems in recommending organizational people in social networks. In this paper, we propose a new hybrid social recommendation system approach for recommending organizational people in social networks using the basic methods of CF-RS and D-RS and using data mining, artificial neural network, and fuzzy techniques. It has high accuracy and does not have problems in underlying RS such as cold start, scalability, rating dispersion, and so on. The approach of this strategy is similar to D-RS, with the difference that instead of demographic characteristics, exceptional features relevant to the SCM are used to recommend peoples and companies.

For this purpose, a real dataset of the LinkedIn social network was collected. Because of the type of information, it collects from users and companies to conduct the experiments in this article; this website is very suitable. Since, in addition to demographic characteristics, it also includes other specialized information such as industry, function, seniority level of the user, and the related company headcount which are close to the SCM.

The implementation framework of this paper uses a preprocessing step to select and reduce data dimensions using a fully connected neural network layer. In the next step, using several profiling approaches and combining them with different methods of calculating the similarity of features, various combinations are created such that each of these combinations recommends separate k-top lists. Finally, the evaluation metrics are used to validate and select the best generation from this initial population.

The rest of this article is written as follows. The relevant researches are reviewed in Sect. 2. In Sect. 3, we discuss basic concepts. A hybrid approach is presented in Sect. 4. We explain how the proposed hybrid approach works and provides the findings and discussion in Sect. 5. Finally, we close this paper with the conclusion in Sect. 6.

2 Related works

RS became increasingly popular in the mid-1990s with methods on the *GroupLens* dataset (Resnick et al. 1994) which were developed over time. The main idea of RS is the item recommendation based on the rating or preference prediction that a person will give up when faced with that item. This scope has been considered for years as a common interest in industry and academia (Aggarwal 2013). Recommender systems are maybe the most successful tool to support personalized recommendations (Yera et al. 2016).

These systems, which are a kind of automated systems (Martinez-Cruz et al. 2015), are used in e-commerce, multimedia, e-learning, web, document and digital libraries, location-based systems, social networks, tourism, health, and other applications (Zhang et al. 2013). Nowadays, many RS are being used in online sites such as Amazon, Netflix, YouTube, TripAdvisor, Last.fm, and IMDb (Tejeda-Lorente et al. 2014; Zahra et al. 2015), the goal of which is to provide and recommend appropriate information to users. Recommender systems increase product sales by suggesting additional items, building customer loyalty, increasing customers' satisfaction based on their purchasing experience, and increasing the likelihood of repeated visits by satisfied customers. Each of these can in turn lead to increased sales and higher revenue (Beladev et al. 2016).

Therefore, communication is the most crucial operational goal on which the RS structure is based. Of course, RS has other goals, including novelty, coverage, reliability, and trust, serendipity and diversity (Aggarwal 2016). Based on the type of application of the RS, they are divided into three broad categories: collaborative filtering, content-based, and hybrid. Other categories for RS are based on the type of techniques used as follows: demographic, associate rules-based, utility-based systems, and knowledge-based RS. However, in general, two strategies are very prominent in these RS categories: CF-RS and CB-RS. Based on another category of RS, it is divided into two groups of individual and group RS (Kagita et al. 2015).

CF-RS relies solely on past user behavior without the need to create explicit profiles. The main idea of the CF-RS is that those uncertain ratings can also be identified because the ratings through users and items are interrelated. Generally, a CF-RS selects a subset of users who have similar profiles to the target user, and then, by searching for items consumed by this group, recommends items based on the similarity of users to target users (Aggarwal 2016). A review paper divided CF-RS into three categories: memory-based

CF, model-based CF, and hybrid CF combining memory-based and model-based CF algorithms (Su and Khoshgoftaar 2009). Memory-based methods, also called neighborhood-based collaborative filtering algorithms, are used to predict ratings based on neighborhoods and have two types of user-based CF and item-based CF.

CB-RS strategies are provided by automatically matching a customer's interests with items' content. Items that are similar to ones the user preferred in the past are now recommended. In CB-RS, the features used to describe the content are important, and therefore, the more descriptive they are, the more accurate they are (Costa-Montenegro et al. 2012).

Hybrid systems use a combination of the CF-RS and CB-RS approaches. *Burke* has done a complete classification of hybrid systems and listing hybridization methods to combine pairs of recommender algorithms (Burke 2002).

D-RS is based on the similarity between demographic information, such as age, gender, and country of users who rate items and predicate target items ratings. Vozalis and Margaritis (2007) investigate how using SVD with demographic information can enhance understandable CF algorithms. Korfiatis and Poulos (2013) use online consumer reviews as a source for demographic recommendations in online travel reviews. Al-Shamri (2016) explores, discusses, and examines many user-profiling approaches for DRSs. Rahman and Oh (2018) proposed a graph-based algorithm that shows that reducing the number of elements improves the search space. Chen et al. proposed a number of standard terms/term frequency (NCT/TF) collaborate filtering algorithm based on the demographic vector. First, the algorithm generates user demographic vector based on user information (age, occupation, gender) and then calculates two users' similarity, based on the previous result (Chen and He 2009). Davoodi et al. (2013) present a framework to build a hybrid expert recommendation system that integrates the characteristics of content-based recommendation algorithms into a social network-based collaborative filtering system.

We also need evaluation metrics to compare the efficiency and accuracy of the RS. The literature on recommender system evaluation offers a large variety of evaluation metrics the use of which evaluates the performance of recommender systems. Error-based metrics like MAE, RMSE, and ranking-based metrics like precision and recall are discussed in detail (Shinde and Potey 2015). Gunawardana and Shani (2009) suggested many evaluation metrics like MAE, RMSE, and precision, recall for comparing recommendation algorithms.

Today, with the expansion of social networks, RS has become an essential part of social networking.

The LinkedIn social network is an excellent source of specialized and non-specialized information for people and companies. In this article, we suggest a hybrid approach for recommending organizational people to each other based on the specific characteristics and supply-chain communications. This approach has high precision and does not suffer from the above-mentioned common problems. This strategy can be effective and efficient in specialized social networks such as LinkedIn.

3 Basic concept

3.1 Similarity measures

The similarity is a fundamental concept for activities such as recommendation engines, clustering, classification, and anomaly detection. The most crucial step in the RS is to discover the similarity between users and recommend similar items to them. To this end, several similarity measure metrics were used in these systems in the lecture reviews. In the following, the most critical similarity measures metrics are presented.

3.1.1 Euclidean distance

The basis for many similarities and dissimilarity measures are Euclidean distance (Eq. 1):

$$d(x \cdot y) = \sqrt{\sum_{i=1}^{n} \left(x_i - y_i\right)^2} \tag{1}$$

where in this paper x_i and y_i are the ith value of two users, x and y the attribute, and n is the total number of attributes. When data is dense or continuous, ED is the best proximity measure (Sondur et al. 2016).

3.1.2 Pearson correlation coefficient

In statistical discussions, the Pearson correlation coefficient is a linear correlation between two random variables. Contrary to ED, Pearson correlation coefficient determines the degree of dependence of two variables in the interval between 1 and − 1. In this regard, 1 is a total positive linear correlation, 0 is no linear correlation, and − 1 is a total negative linear correlation (Eq. 2) (Sondur et al. 2016).

$$r = r_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - \left(\sum x_i\right)^2}\sqrt{n\sum y_i^2 - \left(\sum y_i\right)^2}}. \tag{2}$$

3.1.3 Cosine similarity function

The Cosine similarity function is one of the functions that is heavily used in information retrieval. Its metric shows the normalized dot product of the two attributes. This similarity measure is the angle between the two objects that is 1 for 0°

angle, and less than 1 for any other value in the interval [0, 0.5π). The less angle between the two vectors is, the more similar and the less difference. In this case, the value of the function tends to zero (Eq. 3) (Sondur et al. 2016).

$$\mathrm{Cos}\left(x_i \cdot y_i\right) = \frac{\sum_{i=1}^{n} x_i \times y_i}{\sqrt{\sum_{i=1}^{n} \left(A_i\right)^2} \times \sqrt{\sum_{i=1}^{n} \left(B_i\right)^2}}. \tag{3}$$

3.1.4 Jaccard coefficient

Jaccard coefficient is one of the measures to compare the similarity and diversity of sample sets. It uses the ratio of the intersecting set to the union set as the measure of similarity. Thus, it equals to zero if there are no intersecting elements and equals to one if all elements intersect (Eq. 4) (Agarwal and Chauhan 2017).

$$\mathrm{Jaccard}_{\mathrm{sim}(u.v)} = \frac{\left|I_u \cap I_v\right|}{\left|I_u \cup I_v\right|}. \tag{4}$$

3.2 Fuzzy logics

Fuzzy logic was first introduced in the fuzzy set theory that was obtained from the generalized classical sets theory (Zadeh 1996). Usually, people do not use accurate information to characterize events and use vague and imprecise linguistic expressions. In other words, this logic transcends one level the computer zero–one level, bringing the responses to a vast and unlimited space in the world between 0 and 1 (Yera et al. 2016). Linguistic terms can be used Instead of numerical evaluations because they are primarily expressed in the form of fuzzy sets (Cheng and Wang 2014). Today, the application of fuzzy theories has expanded in different parts of the RS. Tejeda-Lorente et al. (2015) presented that a new approach called REFORE is a recommender system for researchers based on bibliometric use of a 2-tuple quality-based fuzzy linguistic approach to describe the linguistic information. Another useful hybrid model called HIFCF is presented between picture fuzzy clustering and intuitionistic fuzzy recommender systems for medical diagnosis (Thong 2015). Yera et al. (2016) are devoted to a new fuzzy method for managing in a flexible and adaptable way such uncertainty of natural noise in order to improve recommendation accuracy. Also, Son (2014) presented a precise mathematical definition of FRS including theoretical analyses of algebraic operations and properties and a new hybrid user-based fuzzy collaborative filtering method.

In fuzzy sets, X denotes a universe of discourse, or space of points, with its elements denoted as x. A fuzzy set A is defined as a set of ordered pairs (Eq. 5):

$$A = \left\{x \cdot \mu_A(x) | x \in X\right\}. \tag{5}$$

where $\mu_A(x)$ is the membership function of A (Eq. 6):

$$\mu_A : X \to [0.1] \cdot \mu_A \in \{0.1\} \tag{6}$$

The unique feature of a fuzzy set is its membership function. Since most fuzzy sets have a universe of discourse X consisting of the real line R, it would be impossible to write all the pairs defining a membership function. A more convenient and concise way to define an MF is to express it as a mathematical formula. Several formulas can be used, depending on the nature of the issues as a function of membership. In this article, we use a Gaussian MF that is specified by two parameters (Eq. 7):

$$\mathrm{Gaussian}(x;c \cdot \sigma) = e^{-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2} \tag{7}$$

A Gaussian MF is determined by c and σ; c represents the MFs center, and σ determines the MFs width (Fig. 1).

3.3 Artificial neural networks

An artificial neural network is a signal or information processing system that consists of a large number of simple computational elements that communicate directly with each other and are used in parallel distributed processes to solve the computational task involved (Macukow 2016). In the ANN, the artificial neuron accepts data and tries to provide the description of what happens when such data is processed (Nadin 2018). The first steps of ANN were taken by Warren McCulch, a young neurologist and mathematician, in 1943. Walter Pitts developed the first models of ANN. They published an article titled “A logical calculus of the ideas immanent in nervous activity” that investigated how nerve networks function (McCulloch and Pitts 1943). ANN is derived from the human brain structure and consists of two or more layers interconnected with an input layer and an output layer and possibly a series of hidden layers with elements called neurons. The data stream from the input layer is transmitted to an output layer using an activation function similar to synapse in the human brain (Hassan and Hamada 2017). Paradarami et al. (2017) present a deep

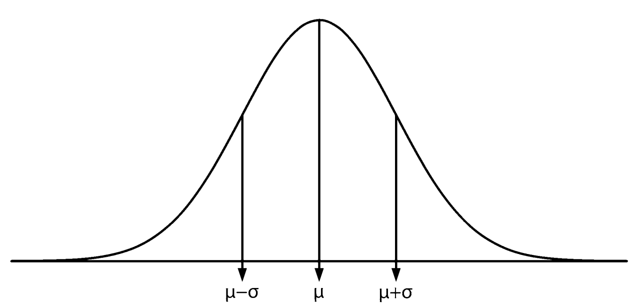


Fig. 1 A Gaussian MF diagram

learning neural network framework that utilizes reviews in addition to CB features to generate model-based predictions for business-user combinations, and Hassan and Hamada (2017) develop an ANN approach for improving the accuracy of multi-criteria recommender systems.

3.4 Common problems in RS

The RS generally suffer some challenges. Some common problems with these systems include data dispersion, cold start, a big data problem, and scalability and over-fitting (Aggarwal 2016). Perhaps one of the main problems with recommending systems is that they need many data to carry out the advice. The RS need many ratings to recommend items to users. In practice, however, it does not happen, and the problem of data dispersion is since most user-item ratings are empty. The problem with the cold start is that when a new user or a new item is added as there is still no information about it, the system cannot provide a good recommendation (Schein et al. 2002). RS also face the problem of having to analyze a massive amount of data. This problem is called big data. The matrix structure of millions of users in millions of items creates a large amount of data that cannot be processed without using special techniques and algorithms, such as data mining. This problem is also referred to as the scalability problem because the system uses large dimensions of data and should use specific techniques such as SVD method, which can produce reliable and efficient recommendations (Isinkaye et al. 2015).

3.5 LinkedIn recommendation system

LinkedIn is the world's largest professional network with more than 562 million users in more than 200 countries and territories worldwide (LinkedIn 2018). The mission of LinkedIn is connecting the world's professionals to make them more productive and successful. LinkedIn uses item-to-item collaborative filtering for people, jobs, companies, groups, and other entity recommendations, and it is a principal component of engagement. That is, for each entity type on the site, there exists a navigational aid that allows members to browse and discover other content. Each of which is called a *Browsemap* by LinkedIn (Wu et al. 2014).

3.6 RS evaluation methods

To compare different methods in the RS, we need evaluation metrics. In the RS for validation presented methods, rated item or items that are precisely followed by the user are usually collected. As a precision indicator, the deviation of predicted recommendations with reality is measured. Afterward, seven commonly used precision measurement indicators are introduced in this paper, called MAE, RMSE, precision, recall, F1, TRP, and FRP.

The two metrics, mean absolute error, and root mean squared error, determine the accuracy of predictions. These two metrics measure how close the predicted results are to actual results (Eqs. 8, 9) (Isinkaye et al. 2015; Lü et al. 2012).

$$\mathrm{MAE} = \frac{1}{|E^P|} \sum_{(i \cdot a) \in E^P} |r_{ia} - \tilde{r}_{ia}|, \tag{8}$$

$$\mathrm{RMSE} = \left(\frac{1}{|E^P|} \sum_{(i \cdot a) \in E^P} |r_{ia} - \tilde{r}_{ia}| \right)^{1/2} \tag{9}$$

The lower value of MAE and RMSE leads to more accurate predictions. The difference between RMSE and MAE is the RMSE before aggregation, and therefore, the effect of significant errors is reduced (Lü et al. 2012).

In addition to the RMSE and MAE measure, precision, recall, and F1-measures also are used to predict recommendation accuracy measures. Precision is the ratio of relevant items selected to the number of items selected (Eq. 10). The recall is the ratio of the selected relevant items or users to the relevant items or users (Eq. 11) (Cheng and Wang 2014).

$$\mathrm{Precision}(t) = 100 \cdot \frac{|S(t) \cap G|}{|S(t)|} \tag{10}$$

$$\mathrm{Recall}(t) = 100 \cdot \frac{|S(t) \cap G|}{|G|} \tag{11}$$

The two precision and recall measures are related to each other, but this relationship is not necessarily monotonic. In other words, the increase in the recall does not always lead to a reduction in precision. In order to combine the precision and recall in a single criterion, a hybrid criterion called the F1 measure is used and is defined in the following way (Eq. 12) (Aggarwal 2016):

$$F_1(t) = \frac{2 \cdot \mathrm{Precision}(t) \cdot \mathrm{Recall}(t)}{\mathrm{Precision}(t) + \mathrm{Recall}(t)} \tag{12}$$

Two other criteria that are used to measure the accuracy of the recommendations system in the literature review are the false-positive rate and true-positive rate.

The TPR, which is the same as the recall, is defined as the percentage of ground-truth positives (Eq. 13) (Aggarwal 2016):

$$\mathrm{TPR}(t) = \mathrm{Recall}(t) = 100 \cdot \frac{|S(t) \cap G|}{|G|} \tag{13}$$

The FPR is the percentage of the falsely reported positives in the recommended list out of the ground-truth negatives (i.e., irrelevant items not consumed by the user). Therefore, if U represents the universe of all items, the ground-truth negative set is given by $(U-G)$, and the falsely reported part in the recommendation list is $(S(t)-G)$. Therefore, the FPR is defined as follows (Eq. 14) (Aggarwal 2016):

$$\mathrm{FPR}(t) = 100 \cdot \frac{|S(t) \cap G|}{|U - G|} \tag{14}$$

The FPR can be viewed as a kind of bad recall, in which the fraction of the ground-truth negatives (i.e., items not consumed), which are incorrectly captured in the recommended list $S(t)$, is reported (Aggarwal 2016).

4 New hybrid approach

In general, there are five phases to build a recommendation system: data collection, user profiling, calculation of similarity, neighborhood selection, and eventually prediction and recommendation (Al-Shamri 2016). The presented study has an approach based on these five phases (Fig. 2).

Several approaches are proposed in the profiling section. In the similarity calculation phase, different strategies are presented. In the sequel, strategies for calculating similarity are combined with profiling strategies, and new methods are presented. After the last phase, all of the proposed methods will be compared in terms of evaluation metrics. Figure 3 illustrates the implementation framework of the model presented in this paper. Approaches and methods are in line with the recommendation of organizational people in the social networks, taking into account the specific features of the communications in the SCM. This approach is a hybrid social recommendation system for recommending organizational people in social networks.

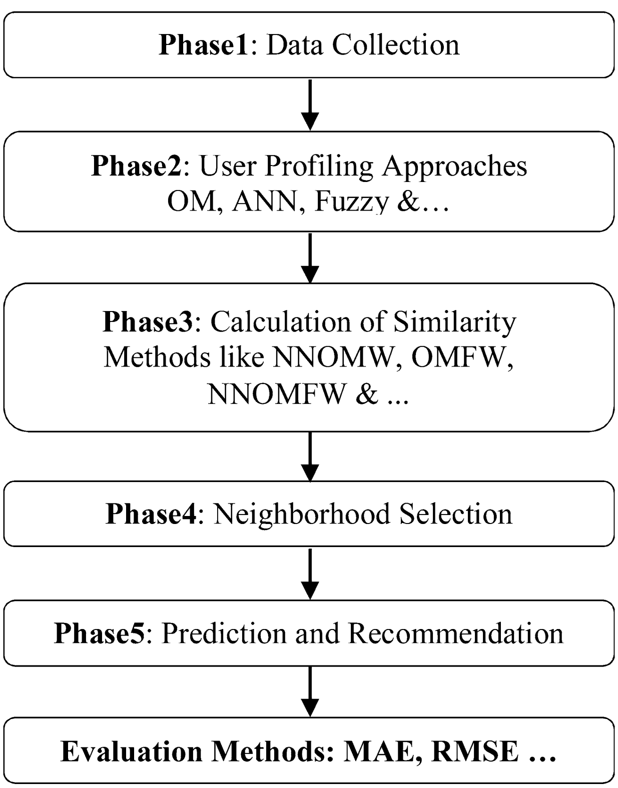


Fig. 2 The new hybrid approach framework

4.1 Data collection and preprocessing

The information gathered from the LinkedIn site is used to conduct experiments. Given the approach of this paper, which focuses on the SCM in the social network to conduct recommendations based on specific characteristics of people, the LinkedIn social network is the most appropriate case study. LinkedIn is a large, specialized community with a wealth of valuable information from professionals in diverse industries and operations, and at various levels of management in companies around the world. In this phase, we use some strategies to collect, preprocess, refine, and reduce data.

4.2 User-profiling strategies

There are different strategies for user profiling. User profiling is done based on information fields. Some strategies are suitable for certain types of data. Here are some innovative strategies along with some basic strategies. These strategies are combined in the similarity computing section and create different methods for experiments of this paper.

4.2.1 ANN based on SCM

An SCM involves all stakeholders involved directly and indirectly in providing the resources of the organization and the customer. Today, companies are challenged to find resources and this challenge affects effectiveness at the level of efficiency and supply-chain response. The SCM involves all functions involved in receiving and completing a customer request. These functions include, not necessarily limited to, the development of new products, marketing, operations, distribution, and financial and customer service (Chopra and Meindl 2007). For example, the LinkedIn social network defines 26 functions for users. Accordingly, two fields of industry and function in specialized networks are seriously related to the SCM. Various industries and functions with different degrees of communication in the SCM are interconnected. Thus far, there has not been a specific study on recommendations in the RS based on SCM. The social network environment is the place where information is interlinked as the interface and the coordinator of the various stages of the

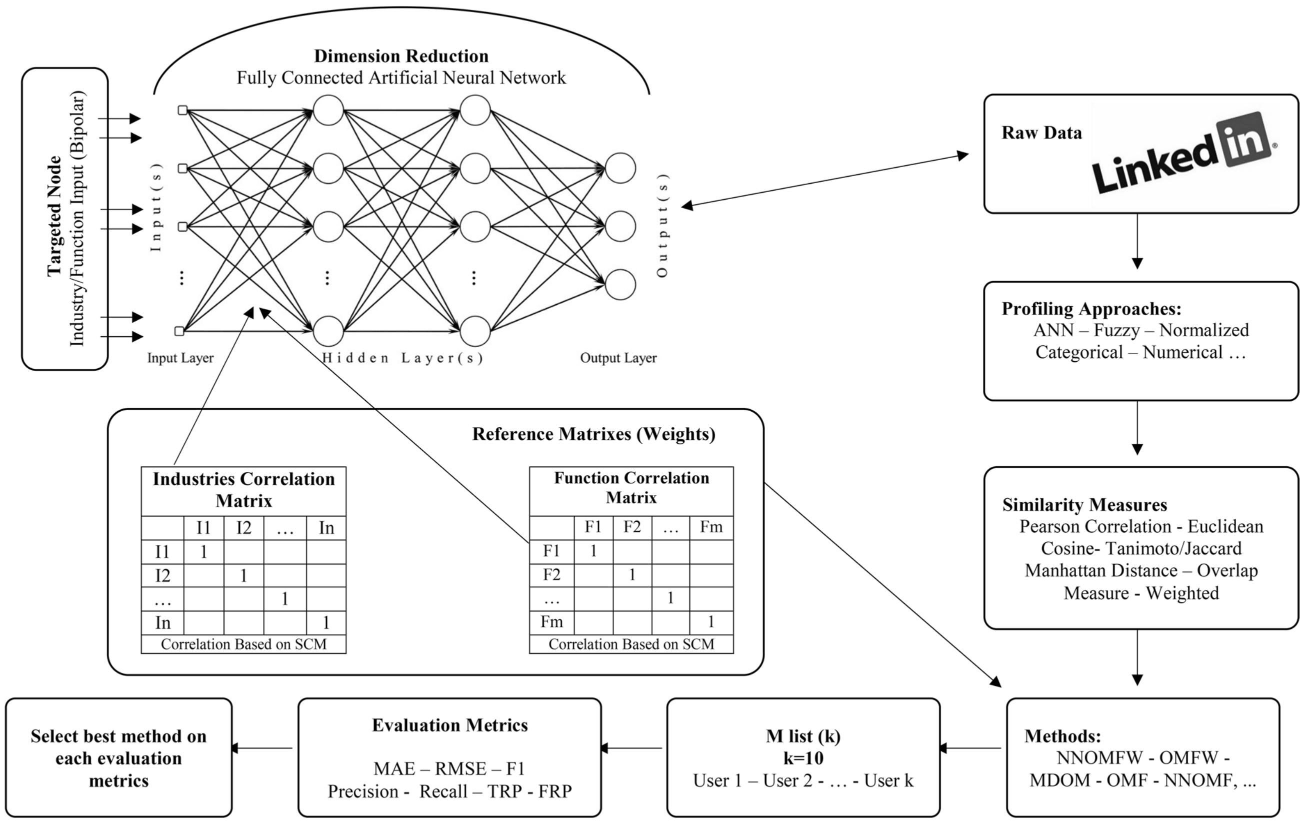


Fig. 3 The implementation framework of the presented hybrid approach

SCM. If the recommendations in these systems are based on the SCM and in order to provide resources, then the overall profitability of the supply chain can be increased (Chopra and Meindl 2007).

It is necessary to draw a graph to model the relationship between different industries. Due to the nature of the industry communications in the SCM, we need to have a weighted directed complete bipartite graph and the related directed adjacency matrix. If the number of industries is n and the number of functions in the network is m, then there are two adjacent matrices n in n and m in m with elements of its primary diameter 1 and other elements are the industries or functions communication degree. Since many different users are in the same industry, each vertex has its loop. The relationship between the vertices in the interval [0–1] is considered. The highest level of communication is one, which is the loops that show the complete communication of industry to itself, and the lowest degree of communication is 0, which results in the lack of communication between the two vertices, and the corresponding directional edge can be eliminated between them. Table 1 shows an adjacency matrix for the communication graph of the five industries selected randomly from LinkedIn. Weights between industries obtained from expert opinions.

To reduce the size of data and overcome the scalability problem, we use a classical ANN layer called fully connected layer, which is a perceptron neural network. In addition to decreasing the size of the data, this layer can be a strategy to increase the richness of the recommendations of the RS and provide more relevant recommendations based on SCM in the second phase. The input of this layer is a vector of numbers. Each input is connected to all outputs. Communications have weights that determine how important each one is (Fig. 4).

To select the similar users in the industry and function attributes that associated with the target user, a fully connected layer is used for each of the two fields. The input layer of each of the two fields is a binary string with lengths of 147 and 26, in which the user industry and function fields are equal to one in the string, and the other values of the input binary string are zero. In the present study, due to the large volume of data in LinkedIn, only five industries and five functions were selected randomly. Thus, the input strings of both layers of the neural network are 5. The outputs are the same number, but their values vary according to the communication weights and the activation function.

The output string for the industry fully connected layer identifies the users related to the target user, and there can

Table 1 The graph adjacency matrix between five selected industrial LinkedIn

5 * 5 LinkedIn Industries Relationship Matrix	Aviation and aerospace (I1)	Banking (I2)	Computer software (I3)	Electrical manufacturing (I4)	Telecommunications (I5)	…
Aviation and aerospace (I1)	1	0	0.65	0.5	0.55	…
Banking (I2)	0	1	0.75	0.5	0.5	…
Computer software (I3)	0.5	0.5	1	0.65	0.7	…
Electrical manufacturing (I4)	0.5	0	0.5	1	0.6	…
Telecommunications (I5)	0.6	0.3	0.55	0.65	1	…
…	…	…	…	…	…	1

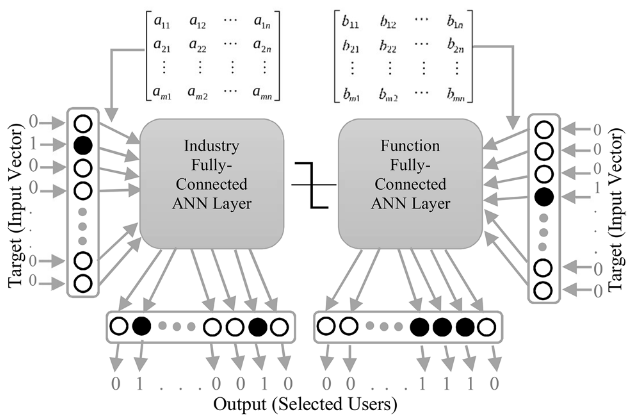


Fig. 4 Two fully connected neural network layer and its adjacency matrixes framework for industry and function fields

be more than one value in the output string. Based on SCM communications, the target user industry is also related to several other industries. The fully connected layer function is the same. The aggregate output of both networks determines which users should be selected.

The weights on the edges in these fully neural networks are assigned using two adjacency matrices associated with industry and function attributes. For each output, the sum of all weighted inputs plus the bias is calculated as output:

$$\text{Output1} = (\text{Input}_1 \times W_{11}) + (\text{Input}_2 \times W_{12}). \cdots . + \left(\text{Input}_5 \times W_{15}\right) + b1$$

This case is very similar to the calculation of a linear relationship in the form $y = ax + b$, where a is a linear coefficient and b is the distance from the y-axis, except that there are distinct slopes (weights) applied to each input element. This linear function is placed inside an activation function:

$$\text{Output1} = \text{Activation Function}\big((\text{Input}_1 \times W_{11}) + (\text{Input}_2 \times W_{12}). \cdots . + (\text{Input}_5 \times W_{15}) + b1\big)$$

The activation function used here is the hard-limit transfer function. This function has a threshold value. The values equal to and above the value of the threshold are activated, and values less than the threshold are inactive and equal to zero.

4.2.2 Fuzzy linguistic approach

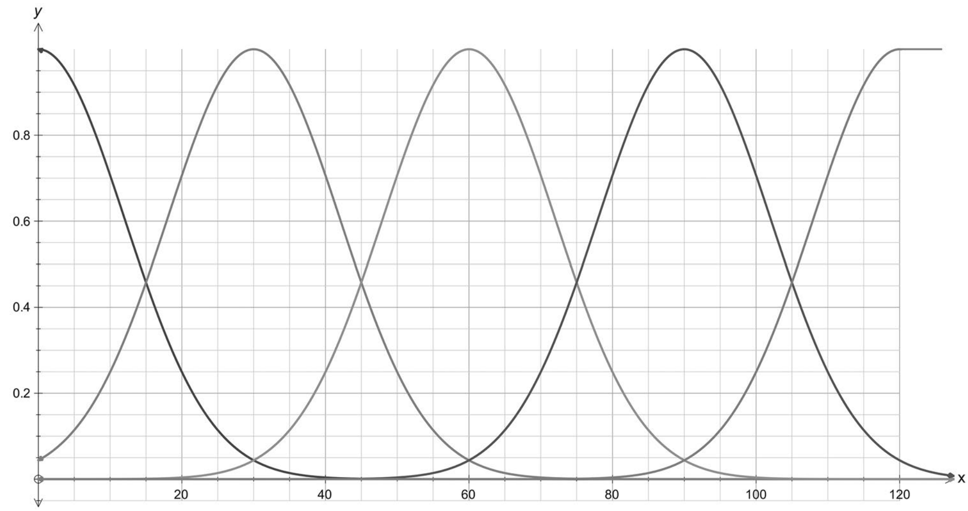
The calculation of similarities between users based on the rating histories is not accurate in many practical applications, and other information such as users' demographic data should be used instead since those data reflect the correlation between users expressed through various attributes of users more strictly than the rating history. An important issue in this observation is that the attributes of users are not only continuous values but also discrete ones such as "Gender" and "Occupation." Thus, in order to calculate the similarity between users based on the demographic data, it is necessary to integrate fuzzy logic with RS.

In this paper, we use a fuzzy approach to fuzzy the specialty of individuals in the social network based on the number of months of activity in the company. Recent studies divide the skills acquisition based on the Dreyfus model into five categories: novice, advanced beginner, competent, proficient, and expert (Ogbuanya and Chukwuedo 2017). A Gaussian function is used to simulate based on the month of work experience field in the fuzzy form (Hameed et al. 2016). Figure 4 plots show a Gaussian MF defined by Gaussian (x; c, 12) and $c = \{0, 30, 60, 90, 120\}$ for experience field based on five level (Fig. 5).

This paper treats age as a fuzzy variable with five fuzzy sets, novice, advanced beginner (*ab*), competent, proficient and expert with the following membership functions (Eqs. 15–19):

$$A_{\text{Novice}}(x) = \left\{ e^{-\frac{1}{2}\left(\frac{x}{12}\right)^2} \quad x \geq 0 \right. \tag{15}$$

Fig. 5 The fuzzy approach diagram with Gaussian MF



$$A_{AB}(x) = \left\{ e^{-\frac{1}{2}\left(\frac{x-30}{12}\right)^2} \quad x \geq 0 \right. \tag{16}$$

$$A_{\text{Competent}}(x) = \left\{ e^{-\frac{1}{2}\left(\frac{x-60}{12}\right)^2} \quad x \geq 0 \right. \tag{17}$$

$$A_{\text{Proficient}}(x) = \left\{ e^{-\frac{1}{2}\left(\frac{x-90}{12}\right)^2} \quad x \geq 0 \right. \tag{18}$$

$$A_{\text{Expert}}(x) = \begin{cases} e^{-\frac{1}{2}\left(\frac{x-120}{12}\right)^2} & 120 \geq x \geq 0 \\ 1 & x \geq 120 \end{cases} \tag{19}$$

To show the degree of a person's experience in a fuzzy form, we present it as a term with five elements according to the membership functions and the following relationship (Eq. 20):

$$P_u \approx \left(A_{\text{Novice}} \cdot A_{AB} \cdot A_{\text{Competent}} \cdot A_{\text{Proficient}} \cdot A_{\text{Expert}}\right) \tag{20}$$

To compare the degree of experience of two users in a fuzzy form, we need similarity measures. There are many similarity measures, but this approach uses a Manhattan distance, to compare the profiles in this paper (Eq. 21).

$$d\left(P_u \cdot P_v\right) = \frac{1}{l} \sum_{l \in S} \left|P_{ul} - P_{vl}\right| \tag{21}$$

As the distance between two users is lower in Manhattan distance, the two users are more similar. We can restrict the fuzzy experience similarity of the two users by using the threshold in the form of the following relationship (Eq. 22):

$$\text{Sim}\left(P_u \cdot P_v\right) = \begin{cases} 1 - d\left(P_u \cdot P_v\right) & d\left(P_u \cdot P_v\right) \leq \theta_1 \\ 0 & d\left(P_u \cdot P_v\right) > \theta_1 \end{cases} \tag{22}$$

In this study, we considered the appropriate value of $\theta_1 = 10^{-3}$. For example, to calculate the similarity of the degree of expertise of two users u_1 and u_2, each of whom has 63 and 8 months of work experience, the target user u_t has 22 months of work experience, the fuzzy similarity of each of the two users with the target user is:

$$\begin{aligned} \text{Sim}\left(u_1, u_t\right) &= \text{Sim}((0.000, 0.023, 0.969, 0.080, 0.000), \\ &\quad (0.186, 0.801, 0.007, 0.000, 0.000)) = 0.599 \\ \text{Sim}\left(u_2, u_t\right) &= \text{Sim}((0.614, 0.614, 0.007, 0.000, 0.000), \\ &\quad (0.186, 0.801, 0.007, 0.000, 0.000)) = 0.753 \end{aligned}$$

The results indicate that the second user is similar to the target user by 0.154 more than the first user.

4.2.3 Weighted/normalized/OM strategies

The similarity calculation of two users is derived from calculating the sum of each element's similarities involved in calculating user similarity. Various attributes are involved in the calculation of the RS similarity. In some studies, the effectiveness of these attributes is not the same. For example, the impact ratio of industry field in RS similarity calculation may be different from the impact ratio of the user expertise degree field. To influence the different effects of fields in the similarity calculation, we need a weighting approach to calculate the weight of similarity. If N relation of similarity calculation would have existed, Eq. 20 assumes the same weight for all similarity relations, while Eq. 21 calculates different weights in the aggregate of all similarity relationships.

The accuracy of this approach in this paper will be considered as a weighted strategy in combination with other methods (Eqs. 23, 24).

$$\mathrm{Sim}(u_x \cdot u_y) = \sum_{i=1}^{N} \mathrm{sim}(x_i \cdot y_i) \tag{23}$$

$$\mathrm{Sim}(u_x.u_y) = \sum_{i=1}^{N} \lambda_i \mathrm{sim}(x_i \cdot y_i), \tag{24}$$

$$\sum_{i=1}^{N} \lambda_i = 1$$

Another issue is that comparing two fields with methods, such as Manhattan distance, leads to the problem of non-uniformity of the scales. The values of the attributes are normalized by the following relationship (Eq. 25). This relationship causes the values to be normalized at intervals of 0 and 1.

$$xn_1 = \frac{x_1 - \min_a}{\max_a - \min_a} \tag{25}$$

One of the most conventional methods used to compare two categorical fields is overlap measure (simple matching measure). In this method, if the two fields are in the same class, the value of the mask is equals to 1, and otherwise, it equals to zero (Eq. 26).

$$\mathrm{sim}(x_i \cdot y_i) = \begin{cases} 1 & x_i = y_i \\ 0 & x_i \neq y_i \end{cases}. \tag{26}$$

5 Experiments and evaluation

5.1 Datasets

It is complicated to gather large volumes of information to conduct experiments on the LinkedIn social network. That is why we focused our data collection on five industries (aviation and aerospace, banking, computer software, electrical manufacturing, and telecommunications) and five functions (accounting, business development, marketing, purchasing, and sales) in a specific geographic area. We collected information about 1404 users in five particular fields (industry, function, seniority level, company headcount, and professional experience in the company) which is close to the SCM. Also, to evaluate methods, 9891 interests in companies that these users followed, were collected. The distinct number of companies followed by the users was 3336.

5.2 Experiments

Due to the presented strategies, experiments are carried out based on five new combination methods called NNOMW, OMFW, NNOMFW, OMF, and NNOMF. These methods are compared to the three classic similarity calculating methods called Cosine, Pearson, and Euclidean. All of these experiments are performed on a dataset to evaluate which method works better and more accurately.

In this paper, the leave-one-out cross-validation method is used to train and test the experiments. In cross-validation, the dataset is divided into q similar sets of data. In the following, one part of q is used for testing, and the rest of the sectors are used for training. This process is repeated for q times to test all of the q sectors. The higher the number of sections (q), the higher the precision and accuracy of the system. A particular instance of this system is a state that q is equal to the total number of records, which is an ideal state. This strategy is known as “leave-one-out cross-validation” that is used in this paper, and the value of q is 1404. The q accuracy average of the different parts is used as the final accuracy of the system.

The number of selected neighbors for each experiment is $k = 10$, and the user’s interests are kept hidden, and the interests of each neighboring user are predicted for q. Then, seven criteria, recall, TRP, precision, F1, FRP, MAE, and RMSE, are calculated, and the average of ten neighboring users counted in these seven evaluation metrics. Finally, the average of each method is calculated as the final value of each metric in the method.

5.2.1 NNOMW method experiment

This method uses the ANN to reduce data, as well as the selection of users related to the target user based on the SCM. Accordingly, the adjacency matrix is selected along with a fully connected neural network layer to select the primary population from the dataset. Similarly, the calculation of the industry and function fields that have a direct relation to the SCM is done using this approach.

In this method, the OM is used to calculate the similarity of users based on the other two fields, the seniority level, and the company headcount, due to the nature of their classification. In order to calculate the individual expertise degree (based on the number of months of works in the company), the values are normalized first, and because of the nature of this field, the MD is used to measure the distance between two users. In the case of this field, because the purpose of the similarity calculation function is maximization, and at MD, fewer values represent more similarity, the values obtained from this criterion are deducted from 1.

Finally, each field of this method is calculated based on the weighted approach. The weights of industry, function, seniority level, company headcount, and degree of expertise fields are, respectively, 0.4, 0.2, 0.1, 0.1, and 0.2 according to experts.

5.2.2 OMFW method experiment

In this method, in addition to the seniority level and company headcount fields, the OM is also used for calculating the similarity of industry and function fields that are categorical.

However, to calculate the similarity of the work experience of two users, we use the fuzzy strategy presented in this paper, which uses a Gaussian function. Finally, the calculation of the overall similarity is based on the weighted strategy and the weights of Sect. 5.2.1-part method.

5.2.3 NNOMFW method experiment

In this method, the neural network strategy is used to calculate the similarity of industry fields and function among users. As with previous methods, for this purpose, two full-connected layers are used for these two fields. Also, since the nature of the seniority level and the company headcount, the OM is used to calculate the similarity between the users in these two fields.

Also, the fuzzy strategy is used to calculate the similarity of work experience between two users. The aggregation of all similarities is done with the weighting strategy and the weights mentioned in the previous two sections.

5.2.4 OMF method experiment

In this method, because of the categorical nature of each field, the OM is used for each of the four fields of industry, function, seniority level, and company headcount. The fuzzy strategy is used to calculate the similarity of users' work experience. Of course, we do not use the weighted strategy in this method.

5.2.5 NNOMF method experiment

In this method, as in Sect. 5.2.3, the ANN is used for industry and function fields, as well as the OM to company headcount filed and seniority levels. The fuzzy strategy is also used for work experience. The only difference with Sect. 5.2.3 method is that the weighting strategy is not done in this way, and the aggregation of the similarity values is considered to be the same.

5.3 Results

In the present study, hybrid methods using the proposed neural network strategy have a higher performance and accuracy than base models, which are well defined in the results of the evaluation metrics. The overall precision and efficiency of the presented ANN approaches are used to the calculation of the similarity of specialized supply-chain characteristics, such as industry, function, seniority level, company headcount, and professional experience in the company and other demographic characteristics, rather than computing and comparing the similarity of user ratings.

This improvement is because the adjacency matrix presented as the weights of the neural network can upgrade the calculation of similarity from the simple binary mode used in the base models to a more advanced state. This strategy, taking into account the supply chain and weights, involves the different value of industry and relationship between different industries in the SCM, and therefore, recommendations are more relevant and more effective than the basic methods. Similarly, the similarity of the function attribute will also be more efficient and more accurate.

Also, combined methods that used fuzzy strategy and Gaussian function in this paper were more efficient and more accurate because instead of simple similarity criterions in the basic models for a feature such as professional experience, a fuzzy feature with Gaussian function was used. This paper attempted to demonstrate the effectiveness and accuracy of these two strategies to organizational peoples' recommendation in social networks, and these strategies can be expanded to each vital feature in the social networks. Some of the methods used in this paper used fuzzy strategy and neural network simultaneously, and the evaluation metrics showed that their performance and accuracy are better than basic models. We use that the binary mode promoted the calculation of similarity in base models to the fuzzy mode by distributing the Gaussian function.

The results show that the proposed approaches, strategies, and methods have a positive impact on improving the accuracy and quality of recommendations. The innovative strategy of the proposed neural network alone can significantly improve the scalability problem and the accuracy of the results.

The goal of this paper was to provide a more accurate recommendation on the LinkedIn network based on SCM. Therefore, the proposed neural network can reduce the amount of data and increase the speed of the recommendation, according to the adjacency matrix and the two critical fields of industry and function provide higher accuracy. For example, the proposed neural network strategy in preparation phase can reduce the volume of users examined the target user from 1404 to 608 for a user with the telecommunication industry and sale function. Also, the adjacency matrix is very useful in calculating the similarity of industries and functions in the various proposed methods using this approach.

Table 2 The results of different methods by the criteria of evaluation

	Recall = TRP	Precision	F1	FRP	MAE	RMSE
NNOMFW	5.385858197	5.407433136	5.396624103	13.77446258	0.932891481	0.953941783
OMFW	5.453646295	5.395497202	5.424415915	13.78398732	0.93301084	0.953987977
MDOM	5.470978724	5.444144323	5.457528538	13.72059735	0.933022944	0.954040631
OMF	5.320366420	5.258952917	5.289481414	13.80002286	0.933023006	0.953096374
Cosine	5.09277984	4.817429951	4.951279671	14.97082498	0.936583535	0.955343977
NNOMF	4.705893786	4.615816181	4.660419764	13.7117094	0.94080765	0.958501289
Euclidean	4.613666281	4.436259012	4.523223774	14.48664302	0.942460772	0.960655966
Pearson	2.402878523	2.850864009	2.607771453	17.20914958	0.959169423	0.970695713

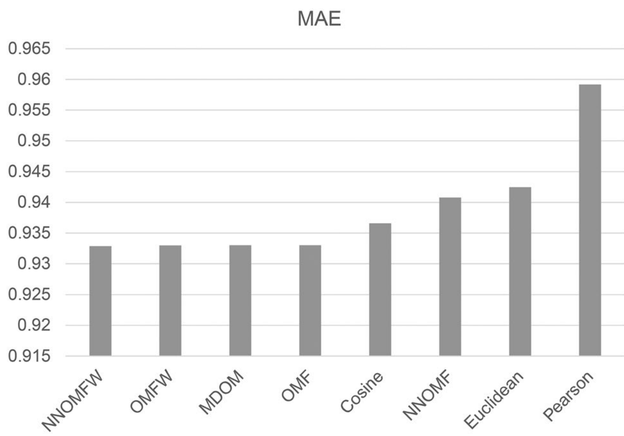


Fig. 6 The MAE metric bar chart for proposed and basic methods

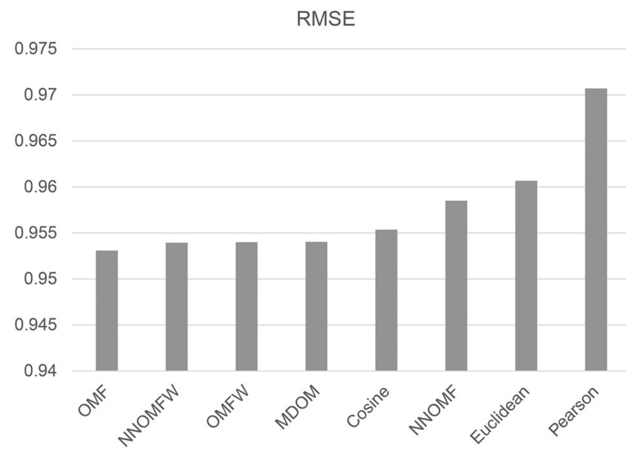


Fig. 7 The RMSE metric bar chart for proposed and basic methods

Also, the fuzzy approach based on the Gaussian function improved in combination with other approaches in the proposed methods. Table 2 illustrates the results of the five combined methods NNOMW, OMFW, NNOMFW, OMF, and NNOMF, and three Cosine, Pearson, and Euclidean basic methods in the seven evaluation metrics (recall, TRP, precision, F1, FRP, MAE, and RMSE).

The least amount of MAE in the methods belongs to the NNOMFW method. The maximum MAE value is also seen in the Pearson method. The NNOMFW MAE is 2.74% less than the Pearson method. Also, the MAE of this method is 1.01%, 0.39%, and 0.012% less than the Euclidean and Cosine methods, and the best method after itself, which is OMFW. Almost all of the proposed methods have less MAE than basic methods. The MAE value of the presented methods is close to each other with very little difference (Fig. 6; Table 2).

The RMSE error rate of the OMF method is 1.81%, 0.78%, and 0.23%, less compared with the Pearson, Euclidean, and Cosine methods, respectively. The RMSE error of this method is only 0.089% less compared to the best method after itself, which is NNOMFW, and the five methods presented in this criterion are also better than basic methods. One of the reasons for the use of this method is to use the fuzzy approach and also neural network approach for the initial selection of users and to calculate the similarity based on communications in the SCM (Fig. 7; Table 2).

Precision is the ratio of relevant items selected to the number of items selected. The best precision value is the MDOM method with a value of 5.4%. The precision of this method compared to the Pearson, Euclidean, and Cosine bases methods is 47.63%, 18.51%, and 11.51% better, respectively. The two methods OMFW and NNOMFW are also superior to the MDOM method with a difference of 0.67% and 0.89%, respectively. The fuzzy approach presented in this article, along with the neural network approach, makes these methods better than the basic methods (Fig. 8; Table 2).

As already mentioned, recall is the ratio of the selected relevant items to the relevant items and equal to the TPR, which is defined as the percentage of ground-truth positives. The highest recall and TPR among methods is MDOM. The recall and TPR of this method is 56.07%, 15.67%, and 6.91% higher compared to the base methods Pearson, Euclidean, and Cosine, respectively. The two OMFW and NNOMFW methods are also superior to the MDOM method with a difference of 0.31% and 1.5%, respectively (Fig. 9; Table 2).

Earlier it was said that to combine two recall and precision criteria, we can use the F1 criterion. As the results of

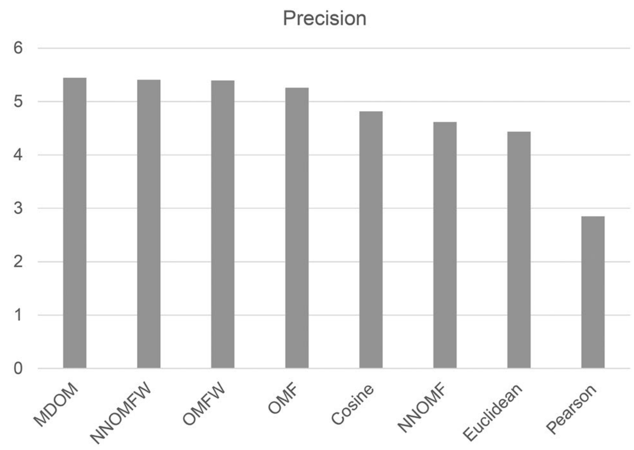


Fig. 8 The precision metric bar chart for proposed and basic methods

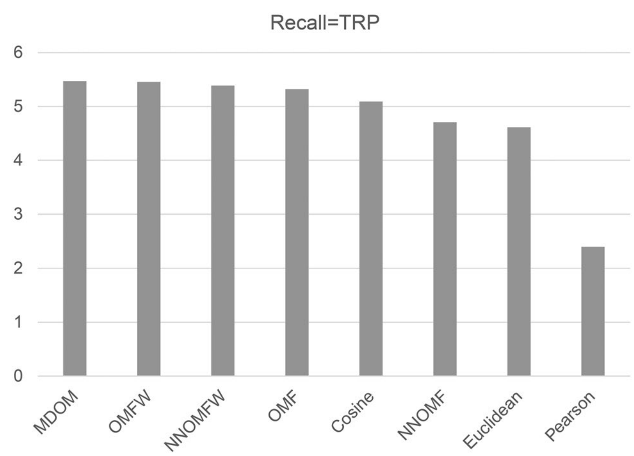


Fig. 9 The recall and TRP metric bar chart for proposed and basic methods

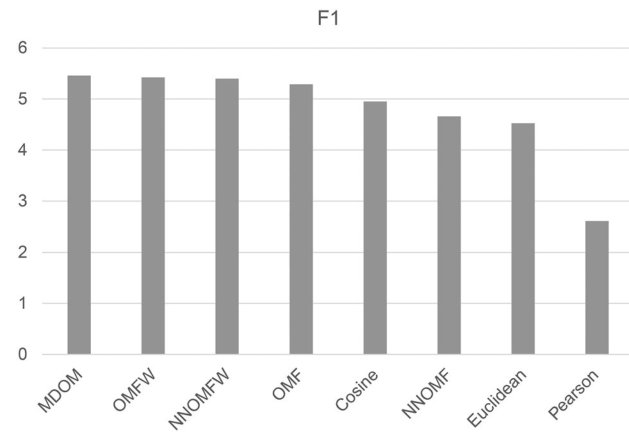


Fig. 10 The F1 metric bar chart for proposed and basic methods

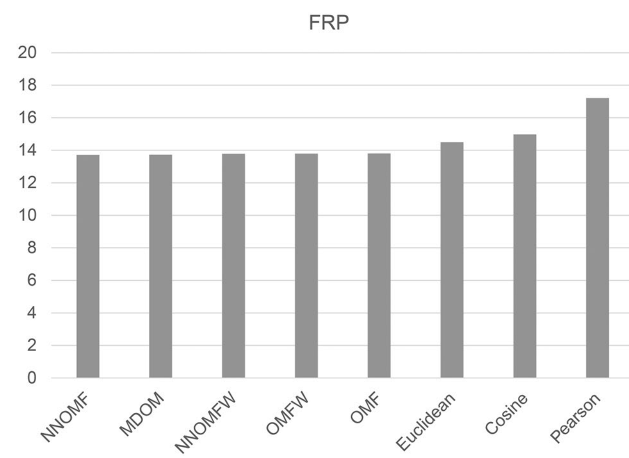


Fig. 11 The FRP metric bar chart for proposed and basic methods

the two preceding sections, the MDOM method is more accurate than other methods. This method has 52.21%, 17.11%, and 9.27% better value than the Pearson, Euclidean, and Cosine bases methods, respectively. This method is only 0.6% better than the best after itself, which is OMFW. All methods presented in this criterion are better than the basic methods (Fig. 10; Table 2). This improvement is due to the use of neural network and fuzzy and weighted approaches in the methods.

The FPR is the percentage of the falsely reported positives in the recommended list out of the ground-truth negatives. In contrast to all previous criteria, the best method in this criterion is NNOMF method. This method is only superior to 0.06% of the method after itself, that is, MDOM. Also, NNOMF method has 20.32%, 8.4%, and 5.3% better value than the Pearson, Cosine, and Euclidean bases methods, respectively (Fig. 11; Table 2).

6 Conclusions and future works

One of the goals of this study was to present a hybrid approach to improve the results of the RS for recommending organizational individuals on social networks. Contrary to the traditional methods in the RS that use user ratings to recommend items to individuals, this paper presents a new hybrid strategy that can, by using user-specific features, recommend people to each other. This system acts like a demographic RS, with the difference that the people's distinctive features in the SCM are taken into account rather than personal characteristics. In social networks, less attention is paid to recommending people to people or, if there is any advice, it is based on demographic characteristics. This fact is one of the papers proposes for specialized social networks such as LinkedIn. In this paper, we have used specific features of users involved in the SCM such as industry, function, seniority level, work experience, and company headcount to recommend people to each other. This will make specific advice based on the supply chain and make the recommendations much more accurate and worthwhile. These methods do not have joint issues and solve data dispersion and cold start, scalability, diversity, and serendipity problems. The hybrid approach combines several approaches, techniques, and methods based on fuzzy logic and ANN. The fuzzy approaches provided by the Gaussian function used to model the work experience of people in the company and made this field are not seen as an absolute value.

The innovative approach of neural network based on a fully connected layer also reduces the amount of data first, which would overcome the problem of scalability, and secondly, sets up the communication between users based on industry and function field in the SCM. This issue has increased the efficiency and accuracy of the proposed methods. In this paper, the proposed methods (NNOMW, OMFW, NNOMFW, OMF, and NNOMF) were compared with the basic methods of the RS with similarity criteria (Cosine, Pearson, and Euclidean) based on the seven criteria of evaluation (recall, TRP, precision, F1, FRP, MAE, and RMSE). In almost all criteria, the presented methods provided better accurate, precision, and accurate. In this research, we had limitations in accessing and collecting data due to the nature of the LinkedIn site. Also, considering that different methods were tested on these data, we also had processing power constraints. In the future, we can work on combining the methods presented using genetic algorithms and gaining more value. The scope of this research can also be expanded to other social networks and tested in these methods.

References

Agarwal A, Chauhan M (2017) Similarity measures used in recommender systems: a study. Int J Eng Technol Sci Res 4:2394–3386

Aggarwal CC (2013) Managing and mining sensor data. Springer, Berlin

Aggarwal CC (2016) Recommender systems. Springer, Berlin

Al-Shamri MYH (2016) User profiling approaches for demographic recommender systems. Knowl-Based Syst 100:175–187

Beladev M, Rokach L, Shapira B (2016) Recommender systems for product bundling. Knowl-Based Syst 111:193–206

Burke R (2002) Hybrid recommender systems: survey and experiments. User Model User-Adapt Interact 12:331–370

Chen T, He L (2009) Collaborative filtering based on demographic attribute vector. In: Future computer and communication, 2009. FCC'09. International conference on. IEEE, pp 225–229

Cheng L-C, Wang H-A (2014) A fuzzy recommender system based on the integration of subjective preferences and objective information. Appl Soft Comput 18:290–301

Chopra S, Meindl P (2007) Supply chain management. Strategy, planning & operation. In: Boersch C, Eschen R (eds) Das summa summarum des management. Springer, Berlin, pp 265–275

Costa-Montenegro E, Barragáns-Martínez AB, Rey-López M (2012) Which App? A recommender system of applications in markets: implementation of the service for monitoring users' interaction. Expert Syst Appl 39:9367–9375

Davoodi E, Kianmehr K, Afsharchi M (2013) A semantic social network-based expert recommender system. Appl Intell 39:1–13

Gunawardana A, Shani G (2009) A survey of accuracy evaluation metrics of recommendation tasks. J Mach Learn Res 10:2935–2962

Hameed IA, Elhoushy M, Osen OL (2016) Interval type-2 fuzzy logic systems for evaluating students' academic performance. In: International conference on computer supported education, 2016. Springer, Berlin, pp 420–441

Hassan M, Hamada M (2017) A neural networks approach for improving the accuracy of multi-criteria recommender systems. Appl Sci 7:868

Isinkaye FO, Folajimi YO, Ojokoh BA (2015) Recommendation systems: principles, methods and evaluation. Egypt Inf J 16:261–273. https://doi.org/10.1016/j.eij.2015.06.005

Kagita VR, Pujari AK, Padmanabhan V (2015) Virtual user approach for group recommender systems using precedence relations. Inf Sci 294:15–30

Korfiatis N, Poulos M (2013) Using online consumer reviews as a source for demographic recommendations: a case study using online travel reviews. Expert Syst Appl 40:5507–5515

LinkedIn (2018) https://about.linkedin.com

Lü L, Medo M, Yeung CH, Zhang Y-C, Zhang Z-K, Zhou T (2012) Recommender systems. Phys Rep 519:1–49

Macukow B (2016) Neural networks—state of art, brief history, basic models and architecture. In: Saeed K, Homenda W (eds) IFIP international conference on computer information systems and industrial management. Springer, Berlin, pp 3–14

Martinez-Cruz C, Porcel C, Bernabé-Moreno J, Herrera-Viedma E (2015) A model to represent users trust in recommender systems using ontologies and fuzzy linguistic modeling. Inf Sci 311:102–118

McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 5:115–133

Nadin M (2018) Machine intelligence: a chimera. AI Soc 33:1–28

Ogbuanya TC, Chukwuedo SO (2017) Career-training mentorship intervention via the Dreyfus model: implication for career behaviors and practical skills acquisition in vocational electronic technology. J Vocat Behav 103:88–105

Paradarami TK, Bastian ND, Wightman JL (2017) A hybrid recommender system using artificial neural networks. Expert Syst Appl 83:300–313

Rahman M, Oh JC (2018) Graph bandit for diverse user coverage in online recommendation. Appl Intell 48:1979–1995

Resnick P, Iacovou N, Suchak M, Bergstrom P, Riedl J (1994) GroupLens: an open architecture for collaborative filtering of netnews. In: Proceedings of the 1994 ACM conference on computer supported cooperative work. ACM, pp 175–186

Schein AI, Popescul A, Ungar LH, Pennock DM (2002) Methods and metrics for cold-start recommendations. In: Proceedings of the 25th annual international ACM SIGIR conference on research and development in information retrieval. ACM, pp 253–260

Shinde S, Potey M (2015) Survey on evaluation of recommender systems. Int J Eng Comput Sci 4:10351–10355

Singh RJ, Sohani N (2011) A proposed model for integration of ERP, CRM, SRM and supply chain management. Paper presented at the International Conference on Industrial Engineering, SVNIT Surat

Son LH (2014) HU-FCF: a hybrid user-based fuzzy collaborative filtering method in recommender systems. Expert Syst Appl Int J 41:6861–6870

Sondur MSD, Chigadani MAP, Nayak S (2016) Similarity measures for recommender systems: a comparative study. J Res 2(3):76–80

Su X, Khoshgoftaar TM (2009) A survey of collaborative filtering techniques. Adv Artif Intell 2009:4

Tejeda-Lorente Á, Porcel C, Peis E, Sanz R, Herrera-Viedma E (2014) A quality based recommender system to disseminate information in a university digital library. Inf Sci 261:52–69

Tejeda-Lorente A, Porcel C, Bernabé-Moreno J, Herrera-Viedma E (2015) REFORE: a recommender system for researchers based on bibliometrics. Appl Soft Comput 30:778–791

Thong NT (2015) HIFCF: an effective hybrid model between picture fuzzy clustering and intuitionistic fuzzy recommender systems for medical diagnosis. Expert Syst Appl 42:3682–3701

Vozalis MG, Margaritis KG (2007) Using SVD and demographic data for the enhancement of generalized collaborative filtering. Inf Sci 177:3017–3037

Wu L, Shah S, Choi S, Tiwari M, Posse C (2014) The Browsemaps: collaborative filtering at LinkedIn. In: RSWeb@ RecSys

Yera R, Castro J, Martínez L (2016) A fuzzy model for managing natural noise in recommender systems. Appl Soft Comput 40:187–198

Zadeh LA (1996) Fuzzy sets. In: Zadeh LA, Klir GJ, Yuan B (eds) Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A Zadeh. World Scientific, Singapore, pp 394–432

Zahra S, Ghazanfar MA, Khalid A, Azam MA, Naeem U, Prugel-Bennett A (2015) Novel centroid selection approaches for KMeans-clustering based recommender systems. Inf Sci 320:156–189

Zhang Z, Lin H, Liu K, Wu D, Zhang G, Lu J (2013) A hybrid fuzzy-based personalized recommender system for telecom products/services. Inf Sci 235:117–129