



Item-Based Recommender Systems Applying Social-Economic Indicators

Mahdieh Atashkar^{1,2} · Faramarz Safi-Esfahani^{1,2}

Received: 13 June 2019 / Accepted: 5 March 2020 / Published online: 2 April 2020
© Springer Nature Singapore Pte Ltd 2020

Abstract

Despite the development of online items in the market, Item-based Recommender Systems play an essential role to assist consumers to access their targets rapidly. Research activities show that users (consumers) in various societies have different purchasing behaviors. In the latest studies, the similarity of users was calculated based on different indications such as Cultural Indicators. Nonetheless, they neglect social-economic indicators expected to be more factors that are efficient. In this research, a framework was designed for recommender systems called SEIRS that include several algorithms based on social-economic indicators. In fact, the similarity of items was defined in terms of the similarity between users' evaluation based on social-economic indicators. The presented method categorizes items and recommends them to users so that they match with users' favorites more precisely. The results were compared not only with the standard item-based system but also with different mixtures of social-based and economical-based indicators. In addition, the presented idea was used to solve the cold start problem that occurs when a newcomer customer with no information on his previous purchases and evaluations becomes available. Furthermore, it is possible to present suggestions that are more precise by using the SVD algorithm to solve the data sparsity. The present research applied three datasets: first, internet purchase by individuals (isoc_ibuy) dataset, second the related data on social-economic indicators from OECD, and third the actual *BookCrossing dataset*. The experimental results represent that in comparison with the standard recommender system, by applying social indicators the proposed recommender algorithm improved precision by 17.19%, and economic indicators improved precision by 17.60%, respectively. Applying both social and economic indicators provides 20.91% more accurate recommendations.

Keywords Recommender systems · Item-based · Social and economic indicators · E-commerce · Cold start · Sparsity

Introduction

Electronic commerce (e-commerce) provides a variety of products for consumers that mislead them in having the right choice. Recommender systems provide appropriate recommendations according to the consumer's requirements to attract them, and thus the consumers approach their purchasing goals and are enticed to purchase more products from the websites [1]. Recommender systems are categorized in terms of the applied knowledge, recommender algorithm, and the way of offering suggestions to the consumers [2].

Collaborative Filtering (CF) is one of the filtration methods in recommender systems that suggest items based on both consumers' interests, and purchasing comments of other consumers with similar behaviors. Collaborative filtering algorithm suffers from two problems: (1) cold start (new users) and (2) data sparsity [3]. The cold start problem is caused by the lack of pre-assessment of the intended user and the excessive sparsity in the recommender system; thus, it decreases the system precision when suggesting products to new users. In addition, the problem of data sparsity occurs when the number of items increases; therefore the density of each user is reduced, which was respected to all products. In article [4], items are ranked based on the consumers' national culture. When a new consumer from a different country joins the system, new items that correspond to their culture are presented to them. Therefore, a cultural-based recommender system represents a more accurate suggestion and better performance in comparison with the item-based

✉ Faramarz Safi-Esfahani
fsafi@iaun.ac.ir

¹ Faculty of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran

² Big Data Research Center, Najafabad Branch, Islamic Azad University, Najafabad, Iran

recommender system. However, due to the diversity of cultures in different countries, this recommender system might decline to purchase in e-commerce. Considering different factors such as cultural, social, economic and demographic characteristics of the users when a new person from another country comes to the system, the recommender system presents more precise and more efficient items that correspond to these factors.

In the recommender system, whenever there is no evaluated information (cold start problem) or there is less information (sparsity issue), the system can present highly accurate suggestions by considering social and economic indicators (SEI)¹ information. In order to achieve this goal, the item-based framework called Social-Economic Indicator-based recommender system (SEIRS) is presented. Having social information and economic conditions in society, and users' attitude toward items, SEIRS² is able to recommend more accurate suggestions along with considering users' behavior and tact. Then, the obtained results are compared with the basic item-based standard system [4] and some mixtures of social and economic indicators. Then, each social-economic indicator recommender system is compared in terms of performance and improving the accuracy of the proposal. This idea is expected to mitigate (1) the cold start problem in which there is no adequate information on users' previous purchases and evaluations, and (2) the amount of data sparsity. Social-economic indicators recommender system is evaluated using three datasets: first, internet purchase by individual's (isoc_ibuy) datasets, second the related data on social-economic indicators from OECD [5], and finally the actual Book-Crossing Dataset.³

In order to develop recommender system, at first check the effect of social-economic indicators on item-based recommender system, then survey the impact of the proposed framework on the new user problem based on additional information on the social-economic indicators of each user's community in the recommender system as an additional measure to help identify neighbors; in the absence of any information from the user, it will be able to get information from these indicators in order to provide suitable suggestions. Therefore, the new user, based on his country's social and economic indicators, is recommended for items that other users with similar social and economic indicators tended to have. In the third experiment, data sparsity challenge is solved when there are many users and items in the system and the rating coverage of users among items is low. The number of evaluations or purchases that a user performs

is much lower than the total items available on the site. It means the rating matrix is sparse. The proposed system is based on social-economic indicators based on which users with similar interests to each product are known. The experimental results in comparison with the previous research activities represented that the proposed recommender algorithm by applying only social indicators improved precision up to 17.19%, and by applying only economic indicators improved precision up to 17.60%, respectively [4]. Applying both social and economic indicators provided 20.91% more accurate recommendations. According to the results, the percentage of newcomers' challenge improvement among social indicators related to social cohesion indicator was close to 41% and the lowest percentage of improvement was related to social health indicator of nearly 20.38% and in general, the mixture of social and economic indicators was close. It shows a 25% improvement in error compared to the mean-based prediction algorithm [6]. Moreover, applying Singular Value Decomposition (SVD) along with the social-economic indicators improved the accuracy of the recommendation up to 37%.

In the sequel, a survey is done on the background concept, the recommender system methods, and the summary of the related researches. The third section introduces the idea presented. Finally, in section four the results and experiments are analyzed.

Concepts and Related Works

General Problems in Recommender Systems

The item-based recommender systems face serious challenges that need to be addressed in order to improve system performance [2]. Cold start, new user problem is one of the challenges of item-based recommender systems. When a new user logs in, there is no purchase and evaluation information that can identify their neighbors and provide appropriate bids based on their background. This can be seen as a very severe matrix scattering problem. Additional user information regarding the matrix complexity issue will significantly improve the offer to new users [1]. The sparsity problem of data dissemination is also a major challenge for recommender systems, as users deal with less than 1% of the items on a website. The users rate the items that produces a large matrix in which most of the elements are null and looking up the elements in such a large matrix is difficult. The recording density of each user will be decreased with respect to the total items that results in decreasing the accuracy, either [7]. This means that the number of reviews or purchases made by a user is much less than the total volume of items available on the site. In these systems, even

¹ Social-Economic Indicators (SEI).

² Social-Economic Indicator-based recommender system (SEIRS).

³ <http://www.informatik.uni-freiburg.de/~ciegler/BX/> <http://www.grouplens.org/node/73>.

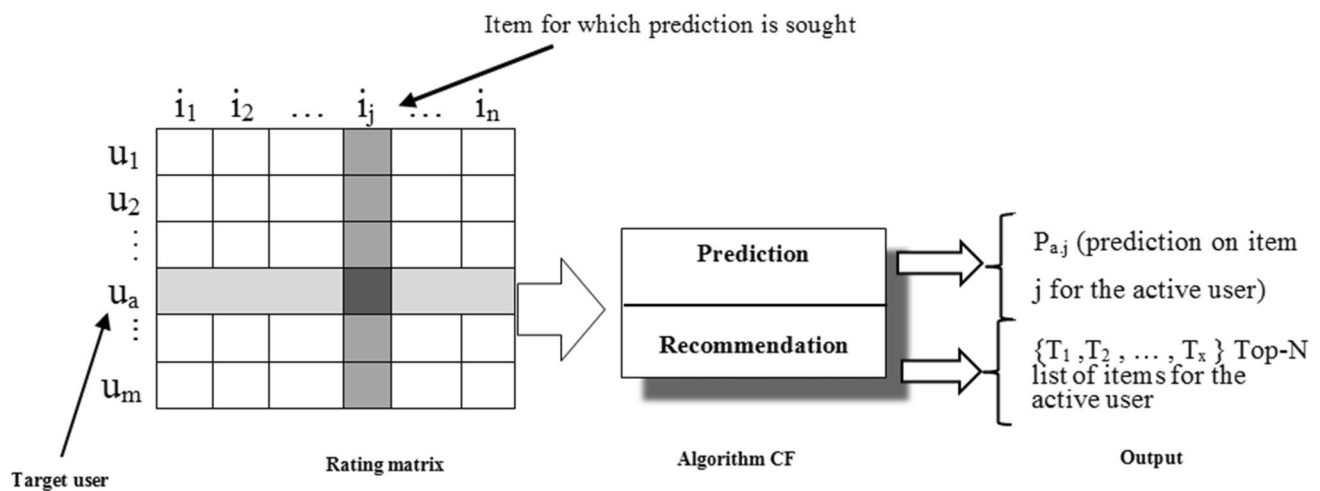


Fig. 1 Collaborative filtering process [10]

active users may have purchased much less than one percent of the items [8].

Concepts of Collaborative Filtering in Recommender Systems

Collaborative filtering (CF) is one of the most successful and widely used technologies that measure the behavior and interest similarity of customers to recommend them suitable items. For example, the recommendation algorithm in Amazon uses the CF technique to collect, rank, and predict similar purchasing of users [9]. An integrated database is created based on which new items are recommended to users.

In general, a collaborative filtering (CF) scenario starts with a list of m users as $U = \{u_1, u_2, \dots, u_m\}$, a list of n items as $I = \{i_1, i_2, \dots, i_n\}$, and a mapping between user–item pairs and a set of weights. The latter mapping can be represented as $m \times n$ matrix M . In the traditional CF domain, the matrix M usually represents user ratings of items, and thus the entry $M_{r,j}$ represents a user U_r 's rating on an item i_j . In this case, the users' judgments or preferences are explicitly given by matrix M . This weight may be binary (representing the existence or nonexistence of the item in the user session), or it may be based on the amount of time spent on the particular item during the session. As shown in Fig. 1, for a given active user (also called the target user) u_a , the task of the CF system is to (1) predict $M_{a,t}$ for a given target item which has not already been visited or rated by u_a ; or (2) recommend a set of items that may be interesting to user u_a [2].

Although there are many collaborative filtering techniques, they can essentially be divided into two major categories [11]:

Model-Based Approaches: Model-based techniques provide recommendations by estimating the parameters of

statistical models for user ratings and do not use all the available information to make a prediction.

Memory-based Approaches: Memory-based CF uses the entire or a sample of the user–item database to generate a prediction. Every user is part of a group of people with similar interests. Such approaches have two main steps, i.e., Calculation of similarities between users and items using rating information and predicting the unknown rating and thus providing either a single value or a list of top N items that the user may like. Memory-based approaches can be classified into two main types: (1) User-based CF: In user-based systems, the similarity between users is calculated by comparing their ratings on the same item, and then computing the predicted rating for item j by user i as a weighted average of the ratings of j by users similar to user i , where the weights are the similarities of these users with i . (2) Item-based CF: In item-based systems, the similarity between two items is determined by comparing the rating made by same user on the items.

Concept of Feature Selection and Dimension Reduction by Singular Value Decomposition (SVD) Algorithm

In some problems, the number of problem features increases because of which records in search space become sparser. Therefore, algorithms are required to decrease the dimension of data. Singular Value Decomposition (SVD) finds the coordinates of users and items to provide predictions [12]. At the first stage, there is a dataset to collect n users and m items along with their priorities. Mathematically, for solving data sparsity problem, matrix $R^{n \times m}$ is used to provide the user–item matrix. In this method, users' matrix is called U , and Items matrix called V . The main goal is to predict the lost scores in the database. Σ contemplates exclusive

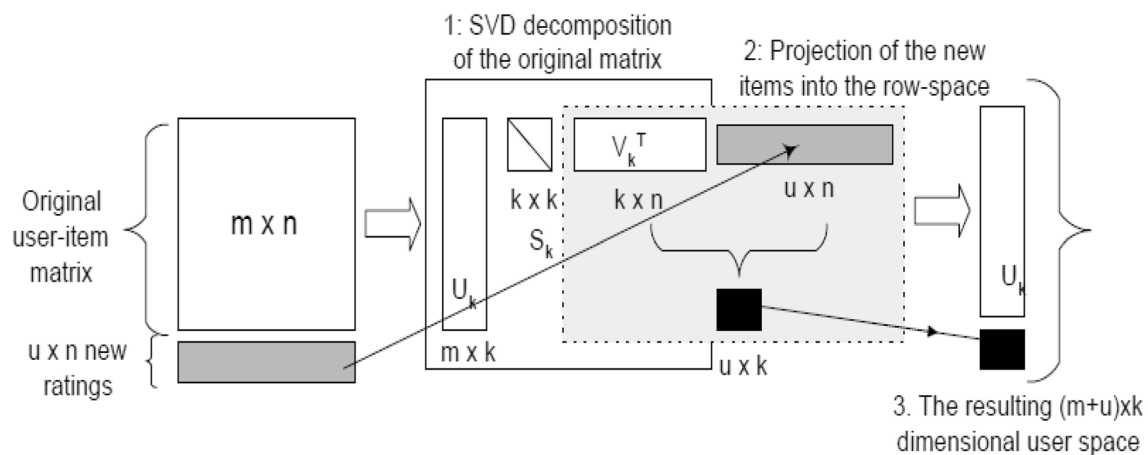


Fig. 2 The diagram of SVD folding-in technique [11]

numbers and the biggest isolated numbers as a scale of rating. Therefore, the perfect matrix will be earned just by the most important approximate properties. In this example, calculating U , V , Σ by protecting just two important properties and just contemplating two first columns in U , V^T pay to decrease data sparsity. So, users and items' score must be equal to the similarity between user vectors and Item vectors [9] (Fig. 2).

Related Research Works on Collaborative Filtering Recommender Systems

By the continued growth of e-commerce, recommender systems are being used to help users in prioritizing [10]. Collaborative filtering is one of the recommender systems to model and/or analyze consumers' prioritizing that suggests suitable recommendations [13]. Figure 3 presents the classification of recommender systems, and the highlights are shown in the mainstream of this research article [14].

Recommender systems that are based on the main idea of collaborative filtering apply the behavior of himself and similar users in suggesting new items to them. Article [4] ranks the similarity between items based on Hofstede's cultural dimension of the buyer [15]. When a new person from a different culture enters the system, new items are recommended according to the user's culture. However, this system does not consider the entire cultural dimensions that affect consumers' behavior. Researchers in [16] consider consumers' behavior and trust relations in social networks in order to solve the cold start problem and present more accurate recommendations. However, people in social networks have no perception of their interests that means applying more effective factors and techniques is required to present more careful recommendations. Therefore, there is need to explore more influential factors among individuals and techniques to find similar friends in social networks in

order to offer suggestions with higher efficiency and accuracy. To alleviate the sparsity and new user problems, this paper [8] presents a new collaborative filtering system in which users are clustered based on their 'personality traits'. The results demonstrate that in facing the data sparsity and new user problems, this method reduces the mean absolute error and improves the precision of the recommendations. However, this system does not consider the social relation in social networks, and demographic features such as age, gender, and occupation. Therefore, this recommender system does not have desirable accuracy. Researchers in [17] present an item-based recommender system called trust CF. In this research, to obviate both cold starts and data sparsity problems, first, appropriate evaluation data is provided by applying both users' rates and trust relations among users. Then, according to the Person Correlation Coefficient (PCC) formula, the trust CF integrates the ratings of trusted friends, similar users, and the ratings of similar items. Since users' data which are used to generate a recommendation model may contain sensitive information, we will study the privacy-preserving recommendation in the future. Also, people's trust relationships should be assessed with the social solidarity and social cohesion indicators. In research [18], the item-based recommender system uses the demographics of users' in the weighted plan in order to personalize web pages. Users' demographic information was used as an auxiliary measure to find similarity between users and resolve cold start users' problems and also present more careful recommendations on whether users' rating was less than Items'; in this type of system, there is a need to resolve the data sparsity problem. Authors in research [19] combined social information and cold start users' problems in a social network (that) depended on group-based and individual-based recommender systems that had been present with more quality. But in group recommendation that depended on behavior and demographics' feature, the recommendations presented

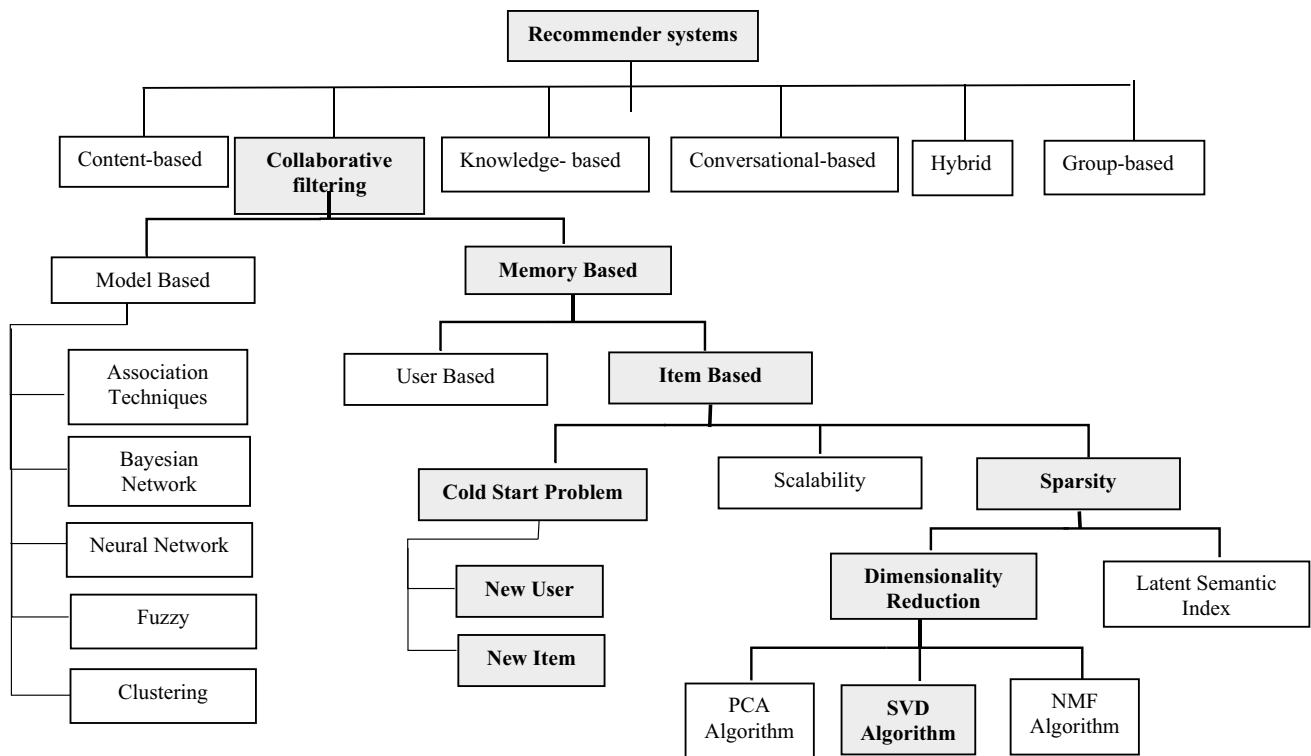


Fig. 3 The sketch of recommender systems [14]

must be considered in all of the perimeters in many different groups. The method in this experiment uses factors affecting e-commerce that involved the mixture of social and economic indicators in order to solve cold start users’ problem and data sparsity, as a source of information besides purchasing information and evaluation that leads to more careful recommendation. The measurement of similarity between users and items according to criterion-affected consumers’ behavior in Table 1 has been analyzed.

The Introduction of Social Indicators

The different behavior of people in confronting various problems and making different decisions in similar conditions comes from social differences and their attitudes toward problems [4]. This is the reason why studying the social attributes of people in predicting their purchasing behavior is important. Figure 4 illustrates social indicators such as Social Solidarity, Social Cohesion Indicators, Equity Indicators, General Context Indicators, Help Indicators, and Self-Indicators, among 24 European countries [5].

Social Solidarity Indicators

Social solidarity is described in terms of 1—**Subjective Wellbeing** (it is commonly known as “happiness” and is

calculated on the response of people to two standard questions: first, how people are satisfied with their own life as a whole? and second, how happy do they feel [20]?) 2—**Group Membership** (it refers to formal/informal participation in social groups such as religious, sport, political, cultural, and associations. It is difficult to say how active the membership is either. It is usually described based on the number of voluntary activities performed in the groups [21]) 3—**Social Isolation** (refers to the lack of contact with society [20]) 4—**Pro/Anti-Social Behavior** (prosocial behavior or voluntary behavior intended to benefit another is a social behavior that benefits other people or society such as helping, sharing, donating, co-operating, and volunteering. Anti-social behaviors also include subtle forms of behavior such as withdrawal and refusal to share with or help others [22]).

Social Cohesion Indicators

Social cohesion indicators with the general goal of identifying indicators: 1—**Life Satisfaction** (life satisfaction is determined not only by economic development but also by the diverse experiences and living conditions.) 2—**Tolerant** (tolerant is the degree of community acceptance of minority groups which is a measurable dimension of social cohesion.) 3—**Confidence** (confidence analyzing the measure of

Table 1 Item-based recommender systems that depend on similarity of users' behaviors

Weakness						
Model preference	Cold start	Sparsity	Dataset	The evaluated indicators	The standard recommendation	
[4]	✓	×	1—Culture Hofstede (dimensions of national culture) 2—Eurostat 3—Bookcrossing	Culture indicator	The studying of the buyer's national culture	
[16]	✓	×	Durban epinions	Social regularization	The studying interprets the differences between social-based recommender systems and trust-aware recommender systems	
[18]	✓		Movie lens	Demographics	The studying demographics information user in a recommender system	
[8]	✓	✓	South tyrol suggests	Personality traits	The studying personality traits user in a recommender system	
[17]	✓	✓	1—Epinions 2—Ciao	Trust CF	The studying trust CF recommender system	
[19]	✓	✓	1—Nielsen 2—movielens	Group membership	The studying enhancing recommender systems using social indicators	
^a The proposed idea in this paper	✓	✓	1—Social Indicators (OECD) 2—Eurostat 3—Bookcrossing	Social-economic indicators	The studying social indicators and economic indicators in the electronic purchasing of individuals	

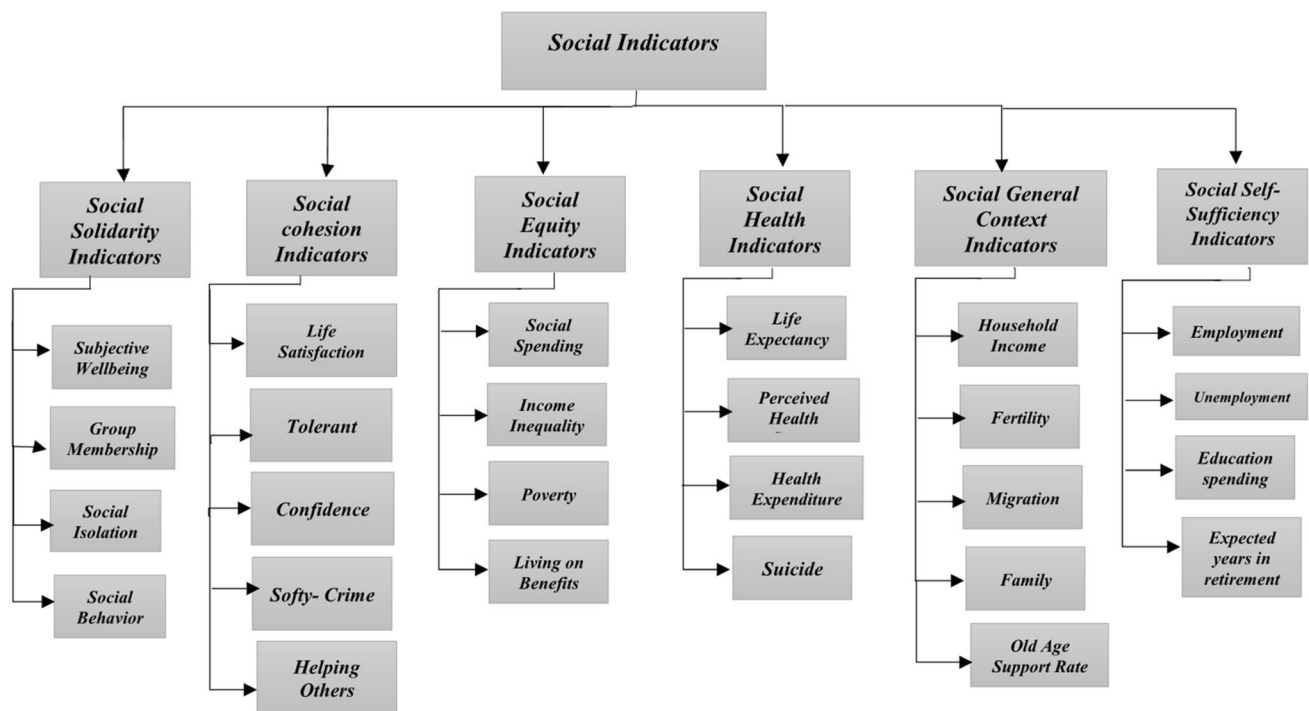


Fig. 4 Classification of social indicators that affect users' behaviors in recommender systems

assurance in one community and the operation depending on the moral made a strong confidence between social relation and economic situation in different societies.) 4—**Softy and Crime** (softy would show the amount of peace in society.) 5—**Helping Others** (helping others indicator either presents the measure of help or shows the correlation [23]).

Equity Indicators

Equity indicators are another purpose of social politics that involved [23]: 1—**Social Spending** (amount of creative expenses, insurance expense, gross production costs.) 2—**Income Inequality** (in every country, there would be a

difference between wealthy and poor income, the main cause of which is unemployment.) 3—**Poverty** (the poverty rate is the ratio of the number of people whose income falls below the poverty line.) 4—**Living on Benefits** (these guaranteed minimum-income benefits (GMI) provide financial support for low-income families and aim to ensure an acceptable standard of living).

Health Indicators

Health status is a fundamental objective of the health-care system, 1—**Life Expectancy** (public health measures, that access to health care and continuing progress in medical technology, have contributed to significant improvements in health status as measured by life expectancy.) 2—**Perceived Health Status** (often the health focus is on objective health status that can be important to assess overall wellbeing.) 3—**Suicide** (suicide gives additional information reasons why some people commit suicide.) 4—**Health Expenditure** (health expenditure is part of the policy response of health-care systems related to health conditions.) [24].

General Context Indicators

The general context indicators in order compare social condition and social indicators in the background of these indicators such as: 1—**Household Income** (real household net disposable income is defined as the sum of household final consumption expenditure and savings.) 2—**Fertility** (the total fertility rate in a specific year is defined as the total number of children that would be born to each woman if she were to live to the end of her childbearing years and give birth to children in alignment with the prevailing age-specific fertility rates [25].) 3—**Migration** (according to the large increasing population in each country and other indicators, such as economic one, had been considered.) 4—**Family** (in view of the strong demand for cross-national indicators on the situation of families and children, the OECD Family Database was developed to provide a cross-national indicator on family outcomes and family policies across the OECD countries, its enhanced engagement partners, and EU member states.) 5—**Old-Age Support Rate** (the old-age support rate is the ratio of the population who are economically active to older people who are more likely to be economically inactive [26].

Self-Sufficiency Indicators

Self-sufficiency is an underlying social policy objective in the background of: 1—**Employment** (access to paid work is crucial to people's ability to support themselves. Employment rates are sensitive to the economic cycle, but in the longer term, they are significantly affected by governments'

higher education and income support policies and by policies that facilitate the employment of women and disadvantaged groups.) 2—**Unemployment** (the unemployment rate is the number of unemployed people as a percentage of the labor force, where the latter consists of the unemployed plus those in paid or self-employment.) 3—**Education Spending** (education spending covers expenditure on schools, universities and other public and private educational institutions. Spending includes instruction and ancillary services for students and families provided through educational institutions.) 4—**Expected years in retirement** (the duration of expected years in retirement illustrates the length of the expected remaining life expectancy from the time of average labor market exit.).

Economic Indicators

In this section, a framework is presented that uses OECD social-economic indicators. The annual report of OECD introduces several indicators as social indicators, including social solidarity, equity, general context, self, health, and social cohesion, as shown in Fig. 4. Here are some of the economic indicators: gross domestic product, gross national income, consumer confidence index, a composite leading indicator, and business confidence index introduced as economic indicators, as shown in Fig. 5. The organization for economic Co-operation and development, OECD by measuring economic indicators, would predict the alteration time in economic cycles, based on consumers' behavior while purchasing, each year. According to Fig. 5, economic indicators that affected consumers' behaviors had been explained: 1—**Gross Domestic Product** (gross domestic product (GDP) at market prices is the expenditure on final items and services minus imports: final consumption expenditures, gross capital formation, and exports fewer imports.) 2—**Gross National Income** (gross national income (GNI) is defined as gross domestic product, plus net receipts from abroad of wages and salaries and of property income, plus net taxes and subsidies receivable from abroad) 3—**Consumer Confidence Index** (the consumer confidence index (CCI) is based on households' plans for major purchases and their economic situation, both currently and their expectations for the immediate future.) 4—**Composite Leading Indicator** (the composite leading indicator (CLI) is designed to provide early signals of turning points in business cycles showing fluctuation of the economic activity around its long-term potential level and development of the population.) 5—**Business Confidence Index** (the business confidence index (BCI) is based on enterprises' assessment of production, orders, and stocks, as well as its current position and expectations for the immediate future.) [27].

Previous investigations in recommender systems did not consider social and economic indicators together. This

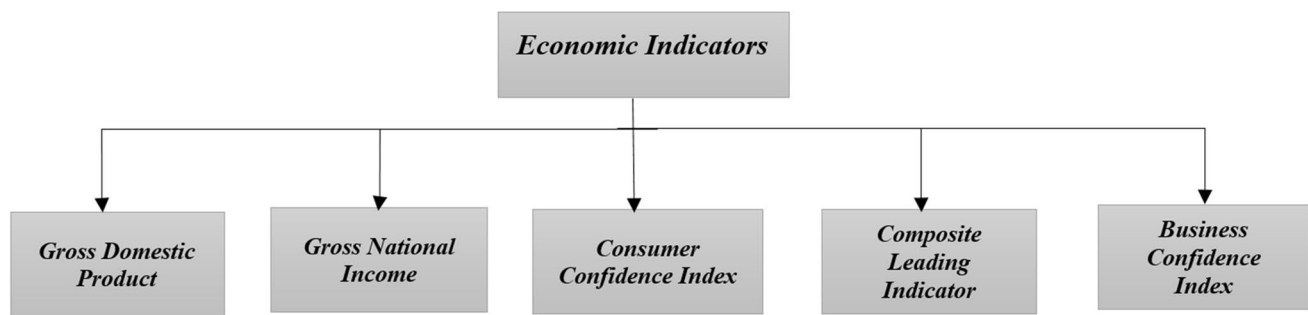


Fig. 5 The classification of economic behavior affected users behavioral in recommender systems

research calculates the correlation between different indicators and their influence on online purchasing to determine the effective indicators, and finally, an item-based recommender framework is presented that is called **Social-Economic Indicator-based Recommender System (SEIRS)**. The framework uses the social-economic indicators to solve cold start and data sparsity problems and presents recommendations.

The Proposed Recommender System

This section encompasses the components of the proposed Social-Economic Indicator-based Recommender System (SEIRS): firstly, the correlation between social and economic indicators and their impact on online purchasing. Based on their relations, the SEIRS framework is introduced that includes algorithms to recommend more recommendations that are accurate in a collaborative filtering recommendation system. The algorithms are presented to satisfy the cold start and data sparsity problems.

Correlation Calculator: The Impact of Social-Economic Indicators on Online Purchasing

In this study, by examining the correlations between these indices and their effect on e-purchasing, we extract effective indicators and present a product-based recommendation system. Correlation is a criterion, to determine the values of two vectors, based on a numerical criterion between $(-1$ and $1)$ where 1 represents the positive correlation of two vectors, i.e., by increasing the values of one vector, the corresponding values in the other vector increase and the number (-1) depicts their negative correlation, which means that by increasing the values of one vector, the corresponding values in the other vector decrease. The zero value represents that the two vectors are not correlated with each other. In addition, the two vectors have a semantic relationship (the significance level should be less than 0.05) [28]. The Pearson's correlation

coefficient can be significant at the 0.05 level (2-tailed) shown by $*$ or it can be significant at the 0.01 level (2-tailed) shown by $**$.

In this section, the correlation between indicators and the amount of their influence on online purchasing are studied. Firstly, the correlation between social indicator vectors and economic indicator vectors is calculated to determine profitable indicators in different societies. Secondly, to find the social-economic indicators that affected online purchasing, the correlation between indicators and the last online purchases is calculated [28].

Studying the Correlation Between Social and Economic Indicators

In order to determine useful indicators in societies, the correlation between social and economic indicators is studied applying the OECD dataset [5]. The research work shows how the entire economic indicators affect social indicators and how they contribute to recommender systems to improve the accuracy of recommendations. Authors in [29] only showed GDP as one of the economic indicators with an affirmative correlation with social indicators, especially with health, and self-income indicators; however, the GDP has a negative correlation with negative indicators such as unemployment, suicide, and criminal. As a matter of fact, indicators such as gross domestic product, gross national income, consumer confidence index, the composite leading indicator, and business confidence index improve the recommendation accuracy in the sequel, as shown in Table 2.

Considering the Effect of Social-Economic Indicators on Online Purchases

In this section, the correlation between online purchasing⁴ and social-economic indicators [5] is calculated in order

⁴ http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database.

Table 2 Correlation between social and economic indicators

	Economic indicators				
	GDP	GNI	CCI	CLI	BCI
<i>Social solidarity indicators</i>					
SBI					
Pearson correlation	.248	.782	.198	-.317	-.168
Sig. (2-tailed)	.762	.218	.412	.654	.627
N	24	24	24	24	24
GMI					
Pearson correlation	.145	.525*	-.378	.298	.319
Sig. (2-tailed)	.553	.021	.110	.265	.183
N	24	24	24	24	24
SWB					
Pearson correlation	.336	.539**	-.198	.100	-.323
Sig. (2-tailed)	.649	.341	.673	.888	.671
N	24	24	24	24	24
SII					
Pearson correlation	-.177	-.194	.148	.260	.100
Sig. (2-tailed)	.468	.427	.553	.283	.683
N	24	24	24	24	24
<i>Health indicators</i>					
Life Expectancy					
Pearson correlation	.558**	.060	.168	.681**	.071
Sig. (2-tailed)	.006	.781	.469	.000	.943
N	24	24	24	24	24
Perceived health status					
Pearson correlation	.548**	.518**	.046	.610	.414
Sig. (2-tailed)	.013	.010	.838	.002	.055
N	24	24	24	24	24
Suicide					
Pearson correlation	-.130	-.290	-.044	-.096	-.067
Sig. (2-tailed)	.552	.160	.817	.655	.343
N	24	24	24	24	24
Health expenditure					
Pearson correlation	.601**	.056	.610	-.109	.002
Sig. (2-tailed)	.005	.801	.002	.540	.511
N	24	24	24	24	24
<i>Social cohesion indicators</i>					
Life Sat					
Pearson correlation	-.098	-.202	-.392	-.063	-.424
Sig. (2-tailed)	.649	.355	.071	.783	.049
N	24	24	24	24	24
Tolerant					
Pearson correlation	.132	.717**	.301	.220	.352
Sig. (2-tailed)	.538	.000	.174	.325	.109
N	24	24	24	24	24
Confidence					
Pearson correlation	.041	.848**	.191	-.071	-.110
Sig. (2-tailed)	.849	.000	.396	.752	.627
N	24	24	24	24	24
Safety					
Pearson correlation	.082	.653**	.126	.133	.308
Sig. (2-tailed)	.702	.002	.577	.556	.163
N	24	24	24	24	24
Help other					

Table 2 (continued)

	Economic indicators				
	GDP	GNI	CCI	CLI	BCI
Pearson correlation					
	.053	.630**	.374	.216	.338
Sig. (2-tailed)					
	.807	.001	.086	.334	.123
N					
	24	24	24	24	24
<i>Equity indicators</i>					
Social spend					
Pearson correlation	.231	.289	-.050	.212	.148
Sig. (2-tailed)	.277	.180	.825	.344	.517
N	24	24	24	24	24
Income inequality					
Pearson correlation	.416*	-.393	-.177	.010	-.033
Sig. (2-tailed)	.043	.064	.430	.964	.883
N	24	24	24	24	24
Poverty					
Pearson correlation	.283	-.399	-.353	-.103	-.282
Sig. (2-tailed)	.181	.059	.107	.648	.182
N	24	24	24	24	24
Living on benefits					
Pearson correlation	-.049	.057	-.256	.225	-.190
Sig. (2-tailed)	.820	.796	.258	.313	.398
N	24	24	24	24	24
Social protection					
Pearson correlation	.175	.383	.276	.136	.539**
Sig. (2-tailed)	.424	.079	.255	.557	.012
N	24	24	24	24	24
<i>General indicators</i>					
Household income					
Pearson correlation	.038	.948**	.079	.259	-.039
Sig. (2-tailed)	.861	.001	.728	.248	.861
N	24	24	24	24	24
Family					
Pearson correlation	.219	-.145	-.472*	.241	-.114
Sig. (2-tailed)	.304	.509	.027	.280	.615
N	24	24	24	24	24
Migration					
Pearson correlation	-.081	.587**	.008	-.024	-.369
Sig. (2-tailed)	.713	.003	.974	.916	.099
N	24	24	24	24	24
Fertility					
Pearson correlation	-.089	.262	.373	-.011	.208
Sig. (2-tailed)	.681	.228	.087	.962	.353
N	24	24	24	24	24
Old-age support rate					
Pearson correlation	.046	-.248	-.160	-.283	-.052
Sig. (2-tailed)	.830	.253	.478	.202	.819
N	24	24	24	24	24
<i>Self-indicators</i>					
Employment					
Pearson correlation	.123	.581**	.478*	.157	.244
Sig. (2-tailed)	.758	.004	.025	.487	.274
N	24	24	24	24	24

Table 2 (continued)

	Economic indicators				
	GDP	GNI	CCI	CLI	BCI
Unemployment					
Pearson Correlation	-.024	-.542	-.340	.530	-.334
Sig. (2-tailed)	.913	.008	.121	.815	.128
N	24	24	24	24	24
Education spending					
Pearson correlation	.135	.908**	-.028	.106	.320
Sig. (2-tailed)	.570	.002	.911	.661	.195
N	24	24	24	24	24
Expected years' men					
Pearson correlation	.324	.392	-.138	.329	-.231
Sig. (2-tailed)	.122	.065	.541	.135	.302
N	24	24	24	24	24
Expected years' women					
Pearson correlation	.255	.152	-.165	.172	-.224
Sig. (2-tailed)	.239	.489	.463	.443	.317
N	24	24	24	24	24

*correlation is significant at the 0.05 level (2-tailed)

**correlation is significant at the 0.01 level (2-tailed)

to find indicators that are more effective in e-commerce. It seems that with the existence of the internet, the national and geographical boundaries should become irrelevant. Consequently, global expansions on the internet could promise more benefits to customers. Although the rate of internet shopping is relatively high in the West, it is still generally unpopular in the East. This research paper mentions that internet shopping is still systematically affected by social-economic differences.

Ability to identify consumers' purchasing behavior and preferences would rise by combining social-economic information with recommender systems. Also, recommender systems could make appropriate offers, especially when there is a lack of customers' purchase information.

The correlations between the internet shopping data and social-economic indicators are calculated to find the indicators that affect e-shopping. The percentage of E-Purchases in different countries was studied, and the correlation of the indicators was calculated to extract the indicators that affect E-Purchasing. Therefore, the social-economic indicators would help recommendation systems to resolve newcomers and cold start problems. To examine the effect of societies on consumers' behaviors, the dataset-commerce by individuals and enterprises [30] is applied. The dataset is released annually by the European Union and includes purchasing information about 24 European countries. We calculate the correlation between different social-economic indicators and the number of online purchases in these countries during 2005–2015. To this point, according to Table 3, the percent of European countries 'online purchases are

extracted, and then their correlations with each indicator are calculated, as well. Table 3 shows that the following indicators have affirmative influence on e-commerce: **Social Solidarity Indicators** (subjective wellbeing, group membership, pro/anti-social behavior) [22], **Social Cohesion Indicators** (tolerant, confidence, safety, help other), **Health Indicators** (life expectancy, perceived health status, health expenditure) [24], **Equity Indicators** (social spend, income inequality, living on benefits, social protection) [23], **General Indicators** (household income, migration, fertility) [24], **Self-Indicators** (employment, education spending, expected years) and **Economic indicators** (gross domestic product, gross national income, consumer confidence index, composite leading indicator, business confidence indicator.) [27].

The Proposed Social-Economic Indicator-Based Recommender System (SEIRS)

Figure 6 shows the presented recommender system that is based on social-economic indicators. It is based on item-based collaborative filtering along with considering social-economic indicators in recommending items. It consists of two main sections: (1) determining neighborhoods and (2) presenting recommendations. At first, the social-economic indicators of each item are determined and then the social and economic organization of people attracted to this item is extracted. As soon as users enter the site, the IP address of his/her resident country is determined. Based on the user's residence country, the recommender system is able to extract his social-economic indicators. Therefore, each user in this system, irrespective of whether he/she has evaluated any item in the system, has some social and economic information based on his/her society. Subsequently, for each item, the social-economic indicators of people who have already bought or evaluated the item would be calculated. Then, a combined similarity measure is designed based on which a list of items is recommended for users that include items that are more similar to the users. Research on the social-economic indicators requires a commodity-based collaborative filtering (CF) framework as a computational advantage over the user-based approaches to examine the similarity of items online prior to giving recommendations.

The Impact of Social-Economic Indicators on the Accuracy of Recommendations

Suppose that there are users in the system with their own social-economic indicators based on their country of residence. Each user has already evaluated one or more items. In order to determine the social-economic indicators of each commodity, a mixture of the social-economic indicators of users that had already evaluated the items is used. Since the computation of items' similarities is independent of the

Table 3 Correlation between different social-economic indicators and the number of online purchases

Social solidarity indicators	SWB	SII	SBI	GMI	
Pearson correlation	.367	-.121	.737**	.631**	
Sig. (2-tailed)	.078	.627	.000	.003	
N	24	24	24	20	
Health indicators	Life expectancy	Perceived health status	Health expenditure	Suicide	
Pearson correlation	.0408*	.493*	.681**	-.055	
Sig. (2-tailed)	.048	.014	.000	.800	
N	24	24	24	24	
Social cohesion indicators	Life Sat	Tolerant	Confidence	Safety	Help other
Pearson correlation	-.0215	.587**	.607**	.588**	.633**
Sig. (2-tailed)	.311	.003	.000	.003	.003
N	24	24	24	24	24
Equity indicators	Social Spend	Income inequality	Poverty	Living on benefits	Social protection
Pearson correlation	.427*	.223	-.392	.287	.455*
Sig. (2-tailed)	.038	.914	.058	.175	.023
N	24	24	24	24	24
General indicators	Household income	Family	Migration	Fertility	Old age support rate
Pearson correlation	.792*	-.023	.474*	.273	-.298
Sig. (2-tailed)	.038	.901	.022	.264	.157
N	24	24	24	24	24
Self-indicators	Employment	Unemployment	Education spending	Expected years men	Expected years women
Pearson correlation	.660*	-.421	.545**	.302	.104
Sig. (2-tailed)	.000	.041	.009	.152	.629
N	24	24	24	24	24
Economic indicators	GDP	GNI	CCI	CLI	BCI
Pearson correlation	.157	.847**	.120	.156	.152
Sig. (2-tailed)	.463	.000	.595	.487	.501
N	24	24	24	24	24

*correlation is significant at the 0.05 level (2-tailed)

**correlation is significant at the 0.01 level (2-tailed)

methods used for generating predictions or recommendations, other evidence (in addition to item ratings or weights) can be used to calculate similarities.

The proposed system that is based on social-economic indicators at first determines the social-economic indicators of each commodity applying the indicators of the users who have evaluated the commodity according to Eq. (1). Then, the similarity of social-economic indicators between commodities has been calculated based on the indicators of users who have evaluated that commodity according to Eq. (2). Then, to achieve the social-economic composite similarity the inter-commodities were calculated using the evaluation matrix applying the cosine similarity (average ratings made by each user) according to Eq. (3). Therefore, the overall similarity scale was determined in terms of the above-mentioned similarities. Finally, the evaluation and prediction of user “u” on the commodity “i” have been calculated according to Eq. (4).

Determining the Similarity Between Items According to The Item Indicators

In this section, social-economic indicators of items had been determined on the basis of a combination of information from social-economic indicators of individuals that have previously purchased or evaluated the items. Equation (1) calculates the similarity of users and items based on social-economic indicators (*d*) for each item (*i*) computed based on the social-economic indicators of users who have rated that item. variable (*k*) is also the number of users who have evaluated item (*i*) [32]:

$$SEIDim_{id} = \frac{\sum_{u=1}^k (R_{u,i} \times Dim_{u,d})}{\sum_{u=1}^k R_{u,i}} \tag{1}$$

In Eq. (1), social indicators for each item *i* (Social Dim (*i*, *d*)) are calculated based on the social indicator of users who

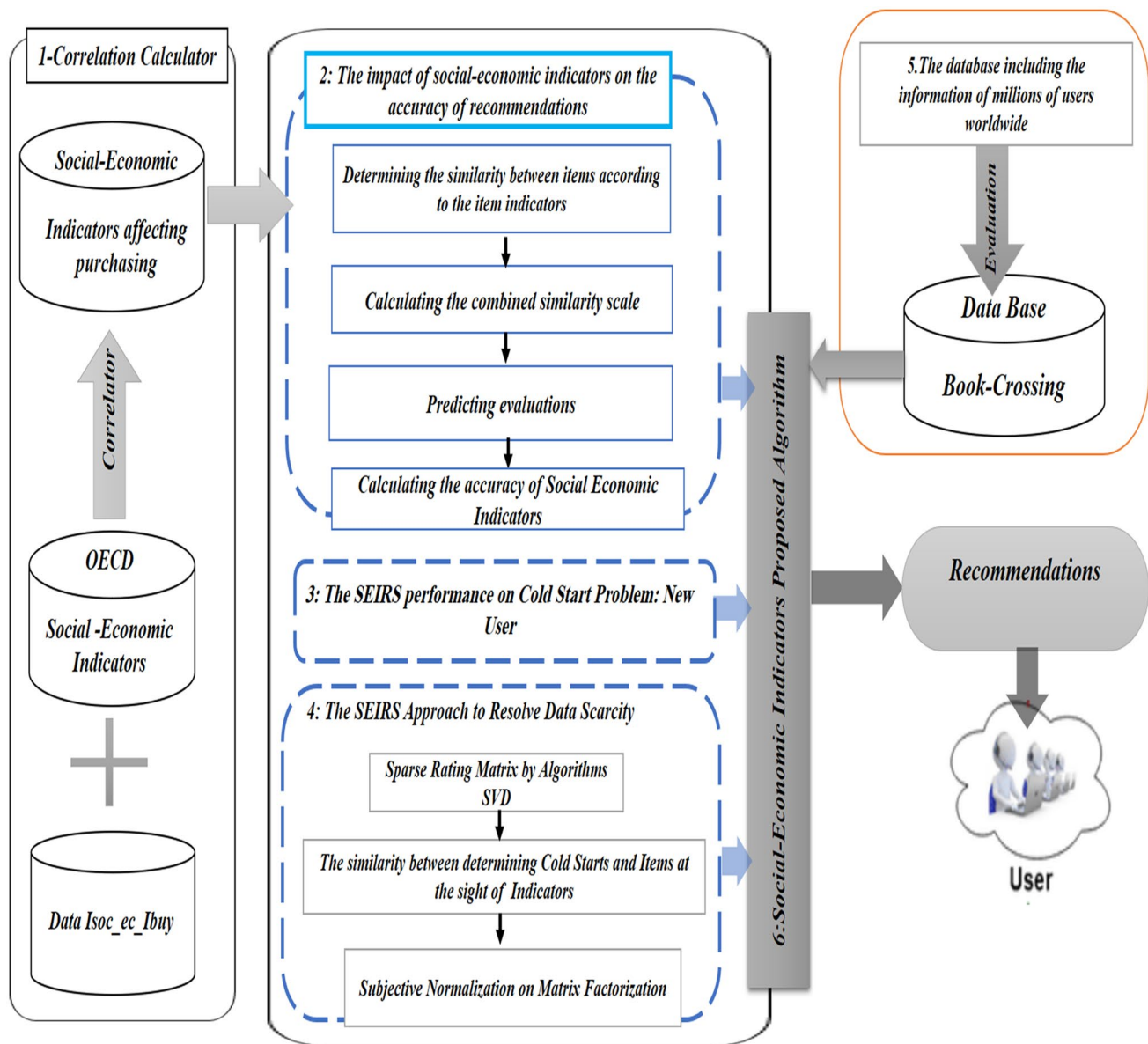


Fig. 6 The framework of social-economic indicator-based recommender system (SEIRS)

have rated that item, for example if “ u_x ,” “ u_y ” and “ u_z ” are rated items and i, j are the social indicators for i, j items computed based on the social indicators of “ u_x ,” “ u_y ” and “ u_z ” that was presented by organization for economic Co-operation and development. Based on the user’s country, the social indicators for each item will be measured according to Eq. (2) [4].

The result of Eq. (1), on n commodities in the system, is $(n \times m)$ matrix in which the $SEID_{i,d}$, d^5 member represents the social-economic indicators value d for the (i) item.

⁵ Social-Economic Indicators (SEI).

Calculating the Combined Similarity Scale

There is a need to design a **combined similarity** measure to determine the distance between social-economic indicators and items. Less distance between indicators is analogous to more similarity of them that is calculated by $SEISim(i, j)$ for a pair of items i and j by Eq. 2 [32]:

$$SEISim(i, j) = \frac{1}{\sqrt{\sum_{d \in D} (\text{SocDim}_{i,d} - \text{SocDim}_{j,d})^2}} \quad (2)$$

In Eq. (2), D is a set including three social indicators. Similarly, we compute item similarities based on the user-item

matrix M . This relation results in an $(n \times n)$ square matrix in which the “ i ” and “ j ” elements of the matrix represent the social-economic similarity between the commodity “ i ” and “ j .” We use the adjusted cosine similarity measure to consider the variation in user ratings. We denote the rating two similarity between items i and j as $\text{RateSim}(i, j)$ [4].

$$\text{RateSim}(i, j) = \frac{\sum_{u=1}^m [(\bar{R}_{u,i} - \bar{R}_u) \times (\bar{R}_{u,j} - \bar{R}_u)]}{\sqrt{\sum_{u=1}^m (R_{u,i} - \bar{R}_u)^2} \times \sqrt{\sum_{u=1}^m (R_{u,j} - \bar{R}_u)^2}} \tag{3}$$

In Eq. (3), $R_{u,i}$ represents the rating of user u on item i , and \bar{R}_u is the average rating value of user u on all items. Finally, for each pair of items i and j , we combine these two similarity measures to get $\text{CombinedSim}(i, j)$ as their linear combination: $\text{CombinedSim}(i, j) = \alpha \cdot \text{SEISim}(i, j) + (1 - \alpha) \cdot \text{RateSim}(i, j)$, where α is a parameter between 0 and 1, specifying the weight of Social-Economic indicators in the combined measure [32]. If $\alpha = 0$, then $\text{CombinedSim}(i, j) = \text{RateSim}(i, j)$, in other words we have the standard item-based filtering. On the other hand, if $\alpha = 1$ then only the Social-Economic indicators are used that essentially, results in a form of SEI-based filtering. Finding the appropriate value for α is not a trivial task, and is usually highly dependent on the characteristics of the data.

Predicting Evaluation

After finding the appropriate similarity scale (combination similarity scale), the set of neighbors of the items, and their similarity value, the next step is the prediction stage, regardless of what method has been used in the previous step to obtain similarity and identify the item’s neighbors. In this step, according to the items set of neighbors, the value of a particular user’s evaluation for a particular product is calculated applying a weighted sum method. We choose the proper value by performing a sensitivity analysis for the datasets in the experimental section below in the sequel. In order to compute predicted ratings, the weighted sum approach is used:

$$M_{u,i} = \frac{\sum_{j=1}^k (R_{u,j} \times \text{CombinedSim}(i \cdot j))}{\sum_{j=1}^k \text{CombinedSim}(i \cdot j)} \tag{4}$$

In Eq. (4) $M_{u,i}$ denotes the prediction value of target user u on target item i .

Calculating the Accuracy of Social-Economic Indicators

In order to measure the accuracy of the recommendations, the standard Mean Absolute Error (MAE) is used

between evaluations and predicted evaluations. Having pairs of $\langle \text{evaluations, evaluated} \rangle$ as $\langle a_i, p_i \rangle$ for the item i , MAE is computed according to Eq. (5). Note that lower MAE values represent higher recommendation accuracy. In this case, the ratings are based on a discrete value of 1 to 10. Thus, the maximum possible value for MAE is 9 (indicating a maximum possible error on all predictions).

$$\text{MAE} = \left(\frac{\sum_{i=1}^n |a_i - p_i|}{n} \right) \tag{5}$$

The SEIRS Approach to Resolve Cold Start Problem

Cold start problem happens when new users enter the system with no prior evaluation on items, no previous evaluation on users. In this case, recommending suitable items by a recommender system is much more difficult. This section aims at using social-economic information of the user from his/her society when there is no adequate information about the user. Suppose that a user from France (IP address) has logged in for the first time. According to the social-economic indicators of this country and applying Eq. (6), the similarity between the indicators of that user with different items will be obtained. Then, by predicting the new user evaluation of the items and considering the average of the evaluations by different users, Eq. (7) predicts the similarity of new users with the items. Equation (8) also calculates items that are less similar to that user.

(1) The criterion for determining the similarity between target user community indicators and social indicators of items is shown in Eq. (6).

$$\text{SEISim}(\text{new user}, i) = \frac{1}{\sqrt{\sum_{d \in D} (\text{SEIDim}_{\text{new user}, d} - \text{SEIDim}_{i,d})^2}} \tag{6}$$

In Eq. (6), $\text{SEIDim}_{i,d}$ is the value of index “ d ” for commodity “ i ” and $\text{SEIDim}_{\text{new user}, d}$, specifies the value of this index for user u . With this formula, the social-economic similarity of the new user to all the items will be obtained and the new user’s interest to the target product can be seen.

The prediction strategy, in this case, is that if the social-economic indicators of the target user are close to the social-economic indicators of the product, it means that the product is closely related to that user’s social-economic indicator and the user will give the product a better rating. Items that have a little social-economic similarity. Therefore, the predicted value is much higher than the average of the estimates made in that commodity, and the prediction of the valuation of these commodities is obtained using Eq. (7). The important point here is that where should we identify the social-economic gap between the new user and the Items?

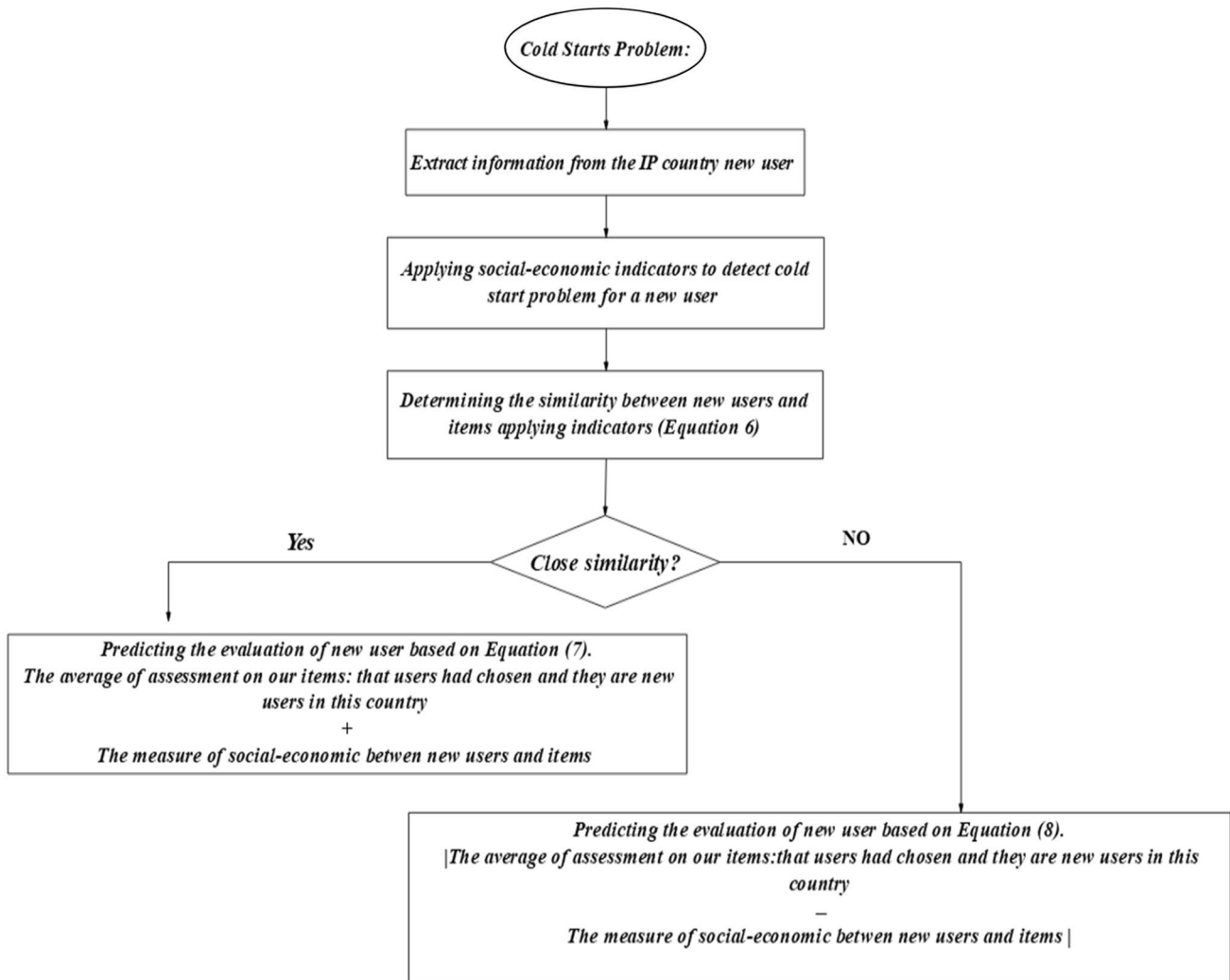


Fig. 7 The algorithm of SEIRS to solve cold start problem

The number of neighbors of the proposed system is 20. In this regard, the order of similarity between the new user and the items is incrementally arranged, the first 20 items being considered as the items most closely related to the user and formula 6 relationship applied to them, and the rest as new ones; they have little resemblance, and he will give them a low score and an Eq. (8), the relationship will apply to them. This formula calculates the social-economic similarity of the new user with all items to find an interest rate of the target user to items. If the target user’s social-economic indicators are more similar to social indicators of an item than others are, it means that the item is close to the target user socially and the user scores better than items with less similarity. The score of the user for a special item is predicted as in Eq. (7).

$$M_{\text{new user},i} = \text{SEISim}_{\text{new-user},i} + \bar{R} \tag{7}$$

In this formula, \bar{R} is the mean ratings performed by users live in the same country as the new user, and $\text{SEISim}_{\text{new-user},i}$ is the social-economic indicators similarity between the new user and item i and $M_{\text{new user},i}$ is the prediction of the new user’s score. On the other hand, if there is a low social-economic similarity between the new user and items, it means that the new user is not close enough to items’ social-economic and does not like them enough and the new user is reluctant to rate the items better than others with more similarity (Fig. 7). The system will predict evaluation according to Eq. (8). In addition, Fig. 8 shows the flowchart of the algorithm, that addresses the cold start problem for new users.

$$M_{\text{new user},i} = \left| \text{SEISim}_{\text{new user},i} - \bar{R} \right| \tag{8}$$

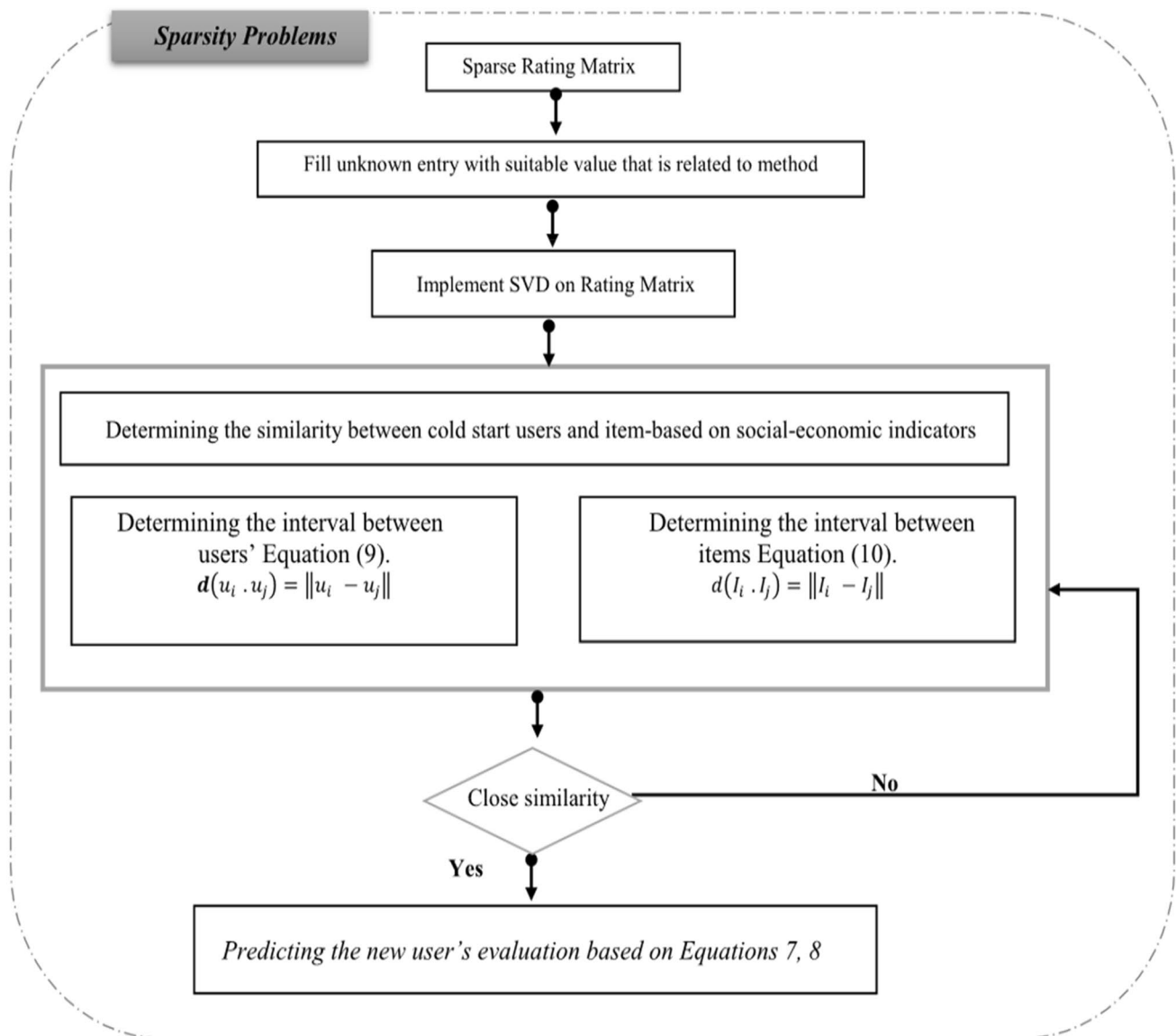


Fig. 8 The algorithm of SEIRS to resolve the sparsity problem

The SEIRS Approach to Resolve Data Scarcity

Data sparsity happens when there are a large number of users and items in the system and the rating coverage of users among items is low. It means the rating matrix is sparse. As a result, finding suitable neighbors for users is difficult. Another problem is users with unusual interests that mislead the recommender system in finding suitable neighbors and finally recommending valid items to users (Fig. 8). Singular Value Decomposition is one of the usable methods in matrix factoring that rating matrix R_n is factored to three matrices $R = USV^T$, and null values in the rating matrix are filled with the average user's score in order to extract secret relations. The completed matrix is standardizing and normalized by

changing values to z-score [33]. According to Fig. 8, firstly, the SVD algorithm is applied to resolve data sparsity problems and secondly by predicting the clustering of items based on social-economic indicators for users' country; the recommender system offers more recommendations that are accurate. In this research, considering the cold start problem and lack of target users' evaluation information, there is a need to determine the interval between users' indicators and item indicators. To eliminate the data dispersion, first, apply the SVD algorithm on the rating matrix, then by proposing the clustering of items based on the social-economic indicators of each user's country, more precise suggestions (applied on cold start) are presented. Then, using both Eq. (9) and Eq. (10), the similarity between users and items is calculated applying both

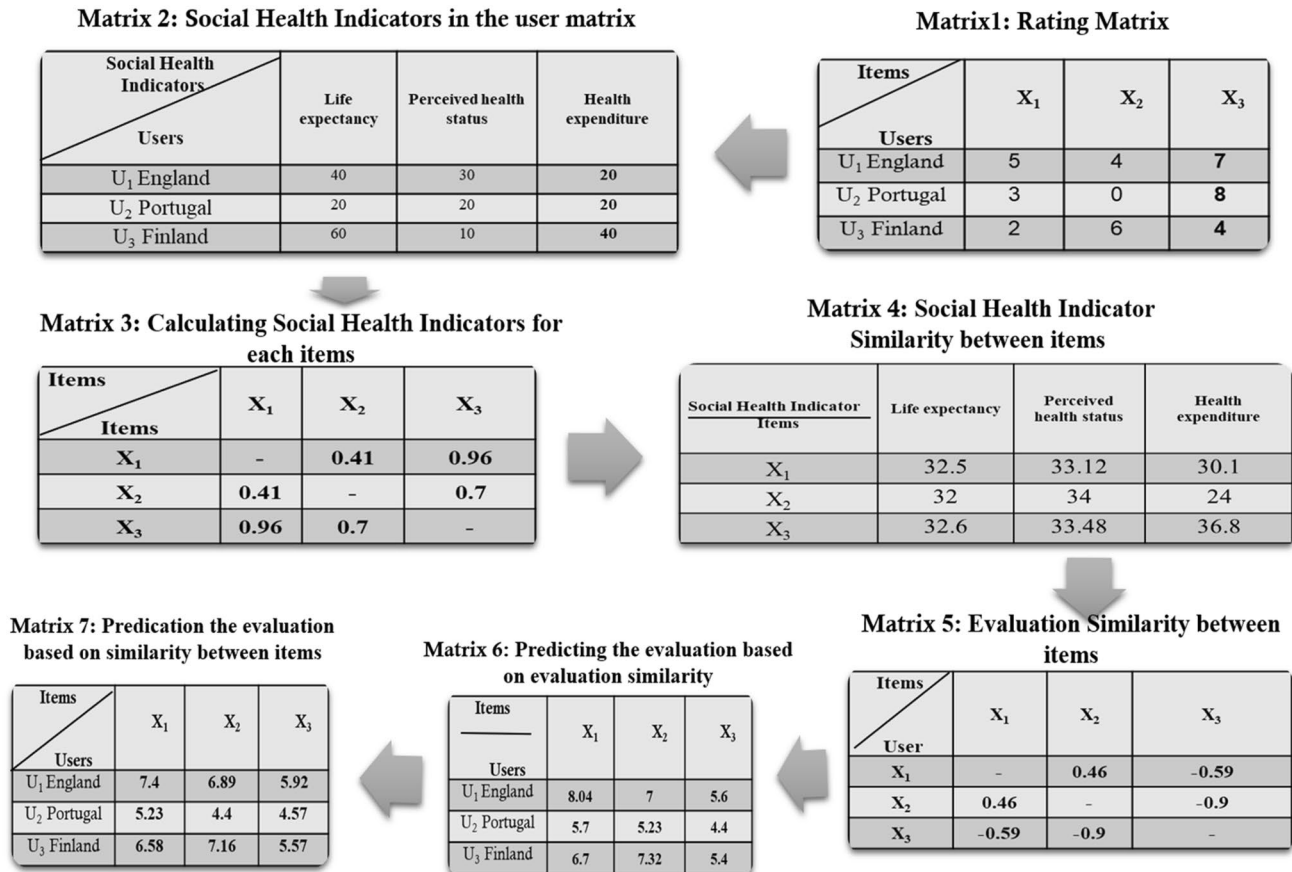


Fig. 9 A case study of social-economic indicators in SEIRS recommender system

matrixes U and M. The SVD algorithm determines whether the items and/or users are similar.

$$d(u_i, u_j) = \| U_i - U_j \| \tag{9}$$

$$d(M_i, M_j) = d(I_i, I_j) = I_i - I_j \tag{10}$$

The Case Study of the Social-Economic Indicator Recommender System (SEIRS)

Figure 9 illustrates a case study of the recommender system that uses social-economic indicators. In Matrix (1), the rating matrix is used in calculating the conformance of users to the social health indicators in Matrix (2). At first, according to Eq. (1) in Matrix (3), social health indicators of the users must be calculated to rate items. Next, the similarity between items according to social health indicators of the users who rated items is calculated according to Eq. (2) in Matrix (4). Then, Matrix (5) considers the similarity of items in terms of assessments performed by items (Eq. (3)). In order to predict the users' rating on different items, Eq. (4) evaluates

the similarity in Matrix (6). Finally, Matrix (7) evaluates and predicts the similarity between items and social health indicators.

Evaluations

Figure 10 illustrates the experimental environment to illustrate how the SEIRS framework is implemented by which tools and/or programming language. Accordingly, statistical calculations and data processing activities are implemented by the SPSS tool; however, the recommender algorithms are implemented in Visual Studio/C#. The choice of the technology stack is justified as follows:

(1) There are several research articles in the field of recommender systems that apply C# language, such as [4, 8, 19]. SEIRS has been developed based on the related research works applying C# as its implementation language.

(2) Different research papers [4, 16, 17] apply a variety of criteria to find similar items. Also, SEIRS presented in this article applies correlation calculations to determine several

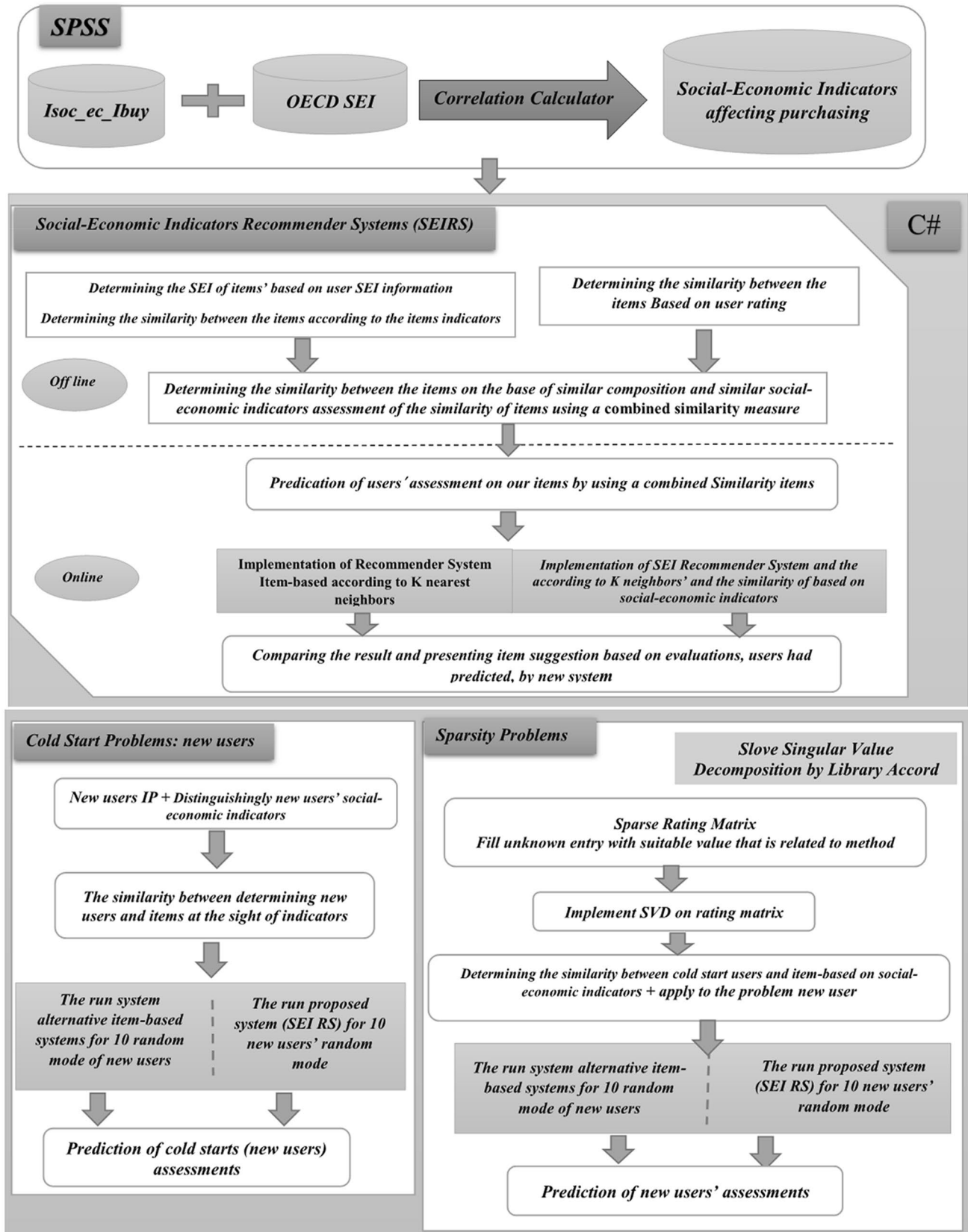


Fig. 10 The process of implementing the proposed algorithm in the recommender system

social-economic indicators to analyze the latest online purchasing of users. The SPSS statistical analytical software provides libraries and tools that make the statistical calculations easy. This is the reason SPSS has been applied in our research work.

In the first phase of this study, a set of correlation calculations with the help of SPSS statistical analysis software is among the internet purchase by individuals (isoc_ibuy) dataset and the data on social-economic indicators from OECD, to extract effective social-economic indicators on electronic shopping.

In the second phase of the study, there are many studies in the field of recommender systems—item based on language $c\#$, each benchmark system has been tested and evaluated on the same dataset and in equal terms, on the same dataset and the original memory of 12 GB and software implemented by programming language $c\#$, tested.

According to Fig. 10, the implementation phase of the SEIRS consists of several steps: (1) deriving social-economic indicators from users via IP address; (2) determining the social-economic indicators of items; (3) the design of the hybrid similarity scale (using the mixture of similarity of social-economic scale similarity with the scale of assessment); and (4) predication of the assessment of the desired user and provide a list of Items.

In this system, the stage of determining social-economic indicators and finding the neighborhood of items could be offline. This system has no effect on the predicted time of system evaluation, such as item-based algorithms, which have a higher speed than the user base, in which all stages have to be done online.

In the third phase, the new user problem, when the new user has entered the system and has not done any evaluation of the items, makes it impossible for the user to provide an appropriate recommendation to the user, thus providing the information related to the users social-economic indicators, from the mean of the evaluations performed on it, and to analyze the algorithm's answers, a basic algorithm is needed to compare the results. This is why a basic algorithm is used in the average assessments.

Finally, for the last phase, how the proposed framework operates on the issue of the data scattering and the new user, assume that a user from the French state (recognition of IP) is applied to the imported system. Given that the ranking matrix has empty values, the ranking matrix needs to be supplemented by means of an average of each user's scores to extract meaningful hidden relationships. The completed matrix is standardized by the conversion of values to the z-score. By applying the single quantity decomposition algorithm on the rate matrix to meet the data dispersion and reduce the recording density of each user with respect to items, the problem of new users can be improved using

the base algorithm in which the average assessments for the goods are used as predictions.

The Dataset and Evaluation Matrices

The datasets applied in this paper are: (1) Book-Crossing Dataset⁶ that contains 278–858 users, 1149,780 ratings (explicit/implicit), and about 271,379 books, (2) The internet purchase by individuals (isoc_ibuy) dataset,⁷ and (3) Social-economic indicators from OECD. The most common criteria for evaluating recommender systems are Precision, Recall and F1 applied in Eqs. (11–13), respectively. Compared to the rating datasets, where users typically rate only a very small number of items, making the dataset extremely sparse, binary selected datasets are dense, as each item was either selected or not by the user α is the number of items that are considered related to the top of the recommended items for the user. ω is the number of items that are not included in the top of the recommended items suggested to the user that they are using. β represents the number of remaining Items that attracted users' traits. The closer the Precision and F1 values of one are, the better the suggestion system is [36].

$$\text{Precision} = \left(\frac{\alpha}{\alpha + \omega} \right) \quad (11)$$

$$\text{Recall} = \left(\frac{\alpha}{\alpha + \beta} \right) \quad (12)$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (13)$$

Experiments

In this section, the prediction accuracy of the SEIRS recommendations framework is evaluated in contrast to those produced by the baseline item-based collaborative filtering. Table 4 shows the experimental design to study different aspects of the SEIRS.

Experiment-1: Studying the Accuracy of Social-Economic Indicators

The MAE is calculated with respect to the number of neighbors (similar items) in the k-nearest-neighbor algorithm. In each experiment, the parameter α is the degree to which the social-economic indicators are used, and the

⁶ <http://www.informatik.uni-freiburg.de/~cziegler/BX/http://www.grouplens.org/node/73>.

⁷ http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database.

Table 4 Experimental design

Experiments	Effective indicators in E-commerce	Indicators						
		1—Social indicators					2—Economic indicators	
		1—1—Social solidarity indicators	1—2—Health indicators	1—3—Social cohesion indicators	1—4—General indicators	1—5—Equity indicators	1—6—Self-indicators	2—Economic indicators
		1—SWB 2—SBI 3—GMI	1—Life Expectancy 2—Perceived Health Status 3—Health expenditure	1—Tolerant 2—Confidence 3—Safety 4—Help other	1—Household Income 2—Migration 3—Fertility	1—Social spend 2—Living on benefits 3—Social protection	1—Employment 2—Education spending 3—Expected years men-women	1—GDP 2—GNI 3—CCI 4—CCL 5—BCI
Experiment 1: Social-economic enhanced accuracy recommender	1—1—Social indicators	✓	-	-	-	-	-	-
	1—2—Health indicators	-	✓	-	-	-	-	-
	1—3—Social cohesion indicators	-	-	✓	-	-	-	-
	1—4—General indicators	-	-	-	✓	-	-	-
	1—5—Equity indicators	-	-	-	-	✓	-	-
	1—6—Self-indicators	-	-	-	-	-	✓	-
	1—7—Economic indicators	-	-	-	-	-	-	✓
	1—8—Social and health indicators	✓	✓	-	-	-	-	-
	1—9—Social cohesion and general and equity and self-indicators	-	-	✓	✓	✓	✓	-
	1—10—Complex of social indicators	✓	✓	✓	✓	✓	✓	-
	1—11—Social and health and economic indicators	✓	✓	-	-	-	-	✓

Table 4 (continued)

Experiments	Effective indicators in E-commerce	Indicators						
		1—Social indicators					2—Economic indicators	
		1—1—Social solidarity indicators	1—2—Health indicators	1—3—Social cohesion indicators	1—4—General indicators	1—5—Equity indicators	1—6—Self-indicators	2—Economic indicators
		1—SWB 2—SBI 3—GMI	1—Life Expectancy 2—Perceived Health Status 3—Health expenditure	1—Tolerant 2—Confidence 3—Safety 4—Help other	1—Household Income 2—Migration 3—Fertility	1—Social spend 2—Living on benefits 3—Social protection	1—Employment 2—Education spending 3—Expected years men-women	1—GDP 2—GNI 3—CCI 4—CCL 5—BCI
	1-12— Social cohesion and general and equity and self- and economic indicators	–	–	✓	✓	✓	✓	✓
	1-13—The mixture of social and economic indicators	✓	✓	✓	✓	✓	✓	✓
Experiment 2: The behavior of the proposed method cold start	2-1—Item based (as the baseline method)	–	–	–	–	–	–	–
	2-2—Social solidarity indicators	✓	–	–	–	–	–	–
	2-3—Health indicators	–	✓	–	–	–	–	–
	2-4—Social cohesion indicators	–	–	✓	–	–	–	–
	2-5—General indicators	–	–	–	✓	–	–	–
	2-6—Equity indicators	–	–	–	–	✓	–	–
	2-7—Self-indicators	–	–	–	–	–	✓	–
	2-8—Economic indicators	–	–	–	–	–	–	✓
	2-9—Complex of social indicators	✓	✓	✓	✓	✓	✓	–

Table 4 (continued)

Experiments	Effective indicators in E-commerce	Indicators						
		1—Social indicators					2—Economic indicators	
		1—1—Social solidarity indicators	1—2—Health indicators	1—3—Social cohesion indicators	1—4—General indicators	1—5—Equity indicators	1—6—Self-indicators	2—Economic indicators
		1—SWB 2—SBI 3—GMI	1—Life Expectancy 2—Perceived Health Status 3—Health expenditure	1—Tolerant 2—Confidence 3—Safety 4—Help other	1—Household Income 2—Migration 3—Fertility	1—Social spend 2—Living on benefits 3—Social protection	1—Employment 2—Education spending 3—Expected years men—women	1—GDP 2—GNI 3—CCI 4—CCL 5—BCI
	2—10—The mixture of social-economic indicators	✓	✓	✓	✓	✓	✓	✓
Experiment 3: The influence of the proposed sparsity method on cold start	3—1—Social solidarity indicators	✓	—	—	—	—	—	—
	3—2—Health indicators	—	✓	—	—	—	—	—
	3—3—Social cohesion indicators	—	—	✓	—	—	—	—
	3—4—General indicators	—	—	—	✓	—	—	—
	3—5—Equity indicators	—	—	—	—	✓	—	—
	3—6—Self-indicators	—	—	—	—	—	✓	—
	3—7—Economic indicators	—	—	—	—	—	—	✓
	3—8—Complex of social indicators	✓	✓	✓	✓	✓	✓	—
	3—9—The mixture of social and economic indicators	✓	✓	✓	✓	✓	✓	✓

parameter k shows the number of neighbors or similar items. According to Table 4, in this experiment, first, each social and economic index is considered separately, and then it is combined with social indices and is analyzed, subsequently.

In Figs. 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, in order to find the similarity between items and neighbors, each indicator had been calculated. In Fig. 11, the results of the recommender system based on social-economic indicators’ are shown in all cases in return for different K values,

Fig. 11 The impact of no. of neighbors on the recommendation with all Social and Economic Indicators (20.91% improvement)

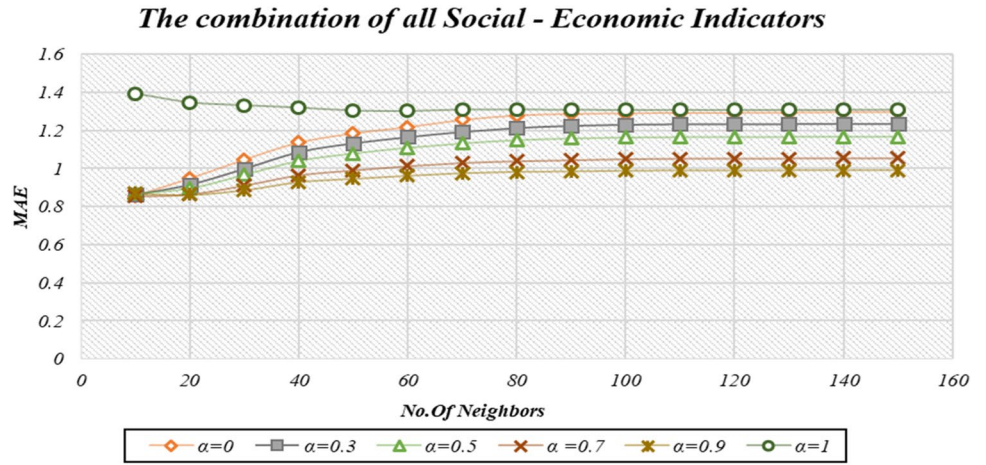


Fig. 12 The impact of no. of neighbors on the recommendation with social indicators (17.19% improvement)

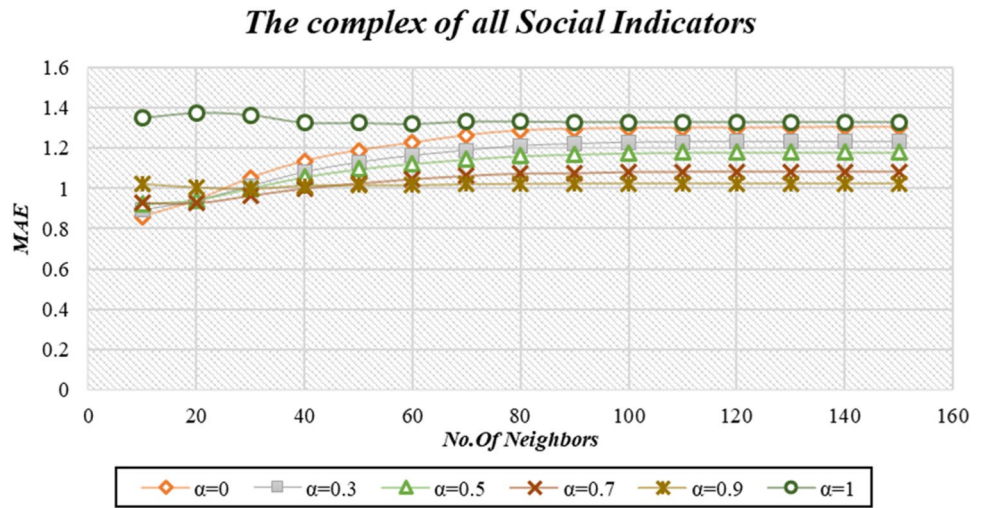
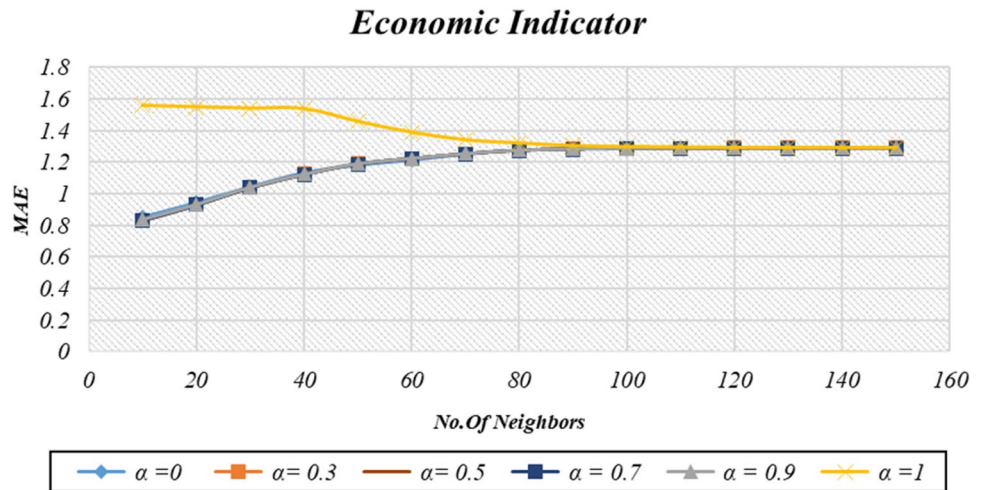


Fig. 13 The impact of no. of neighbors on the recommendation with economic indicators (17.62% improvement)



and these results are much better than the implemented item-based recommender system. The bidding results show (20.91%) improvement, by using social-economic

information related to users as additional information. According to the results in Figs. 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, the first hypothesis of research regarding

Fig. 14 The impact of no. of neighbors on the recommendation with social solidarity indicators (18% improvement)

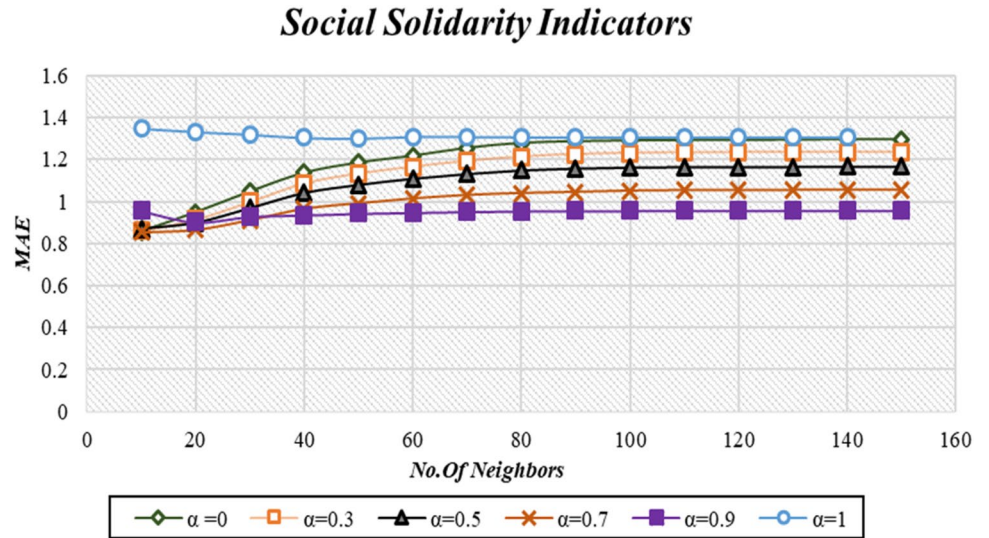


Fig. 15 The impact of no. of neighbors on the recommendation with social health indicators (13.16% improvement)

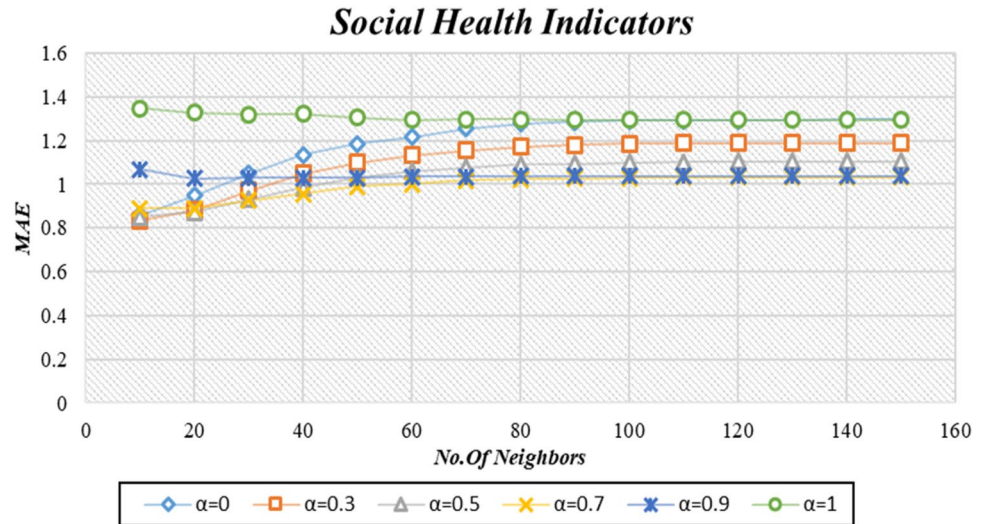


Fig. 16 The impact of no. of neighbors on the recommendation with social self-indicators (12% improvement)

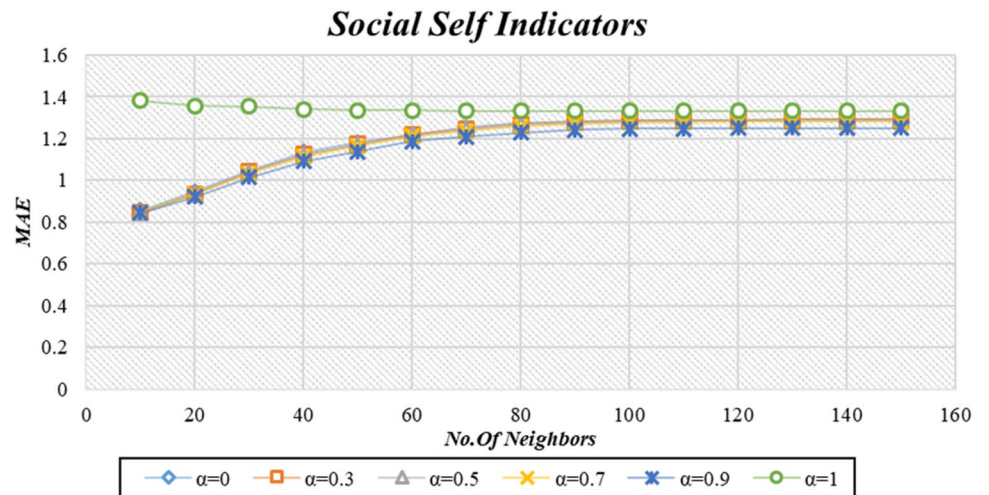


Fig. 17 The impact of no. of neighbors on the recommendation with social cohesion indicators (18.72% improvement)

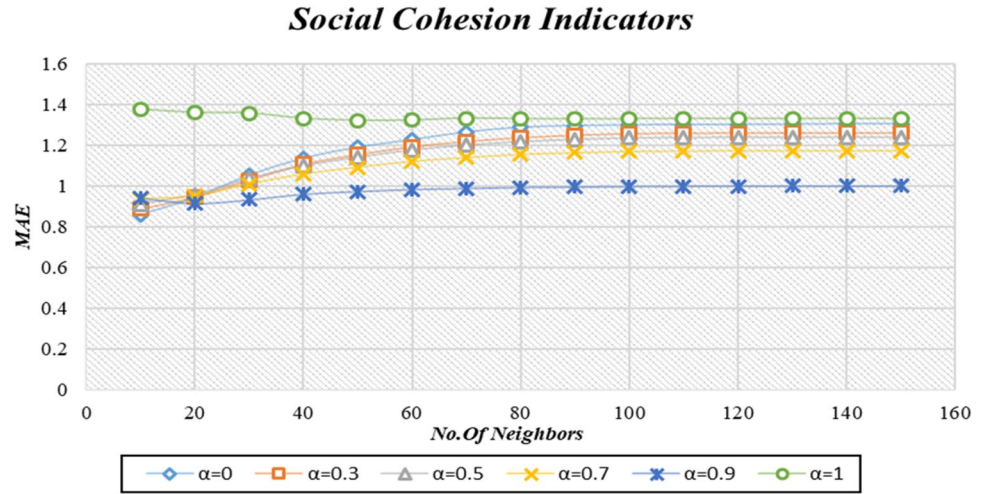


Fig. 18 The impact of no. of neighbors on the recommendation with Social Equity Indicators (19.36% improvement)

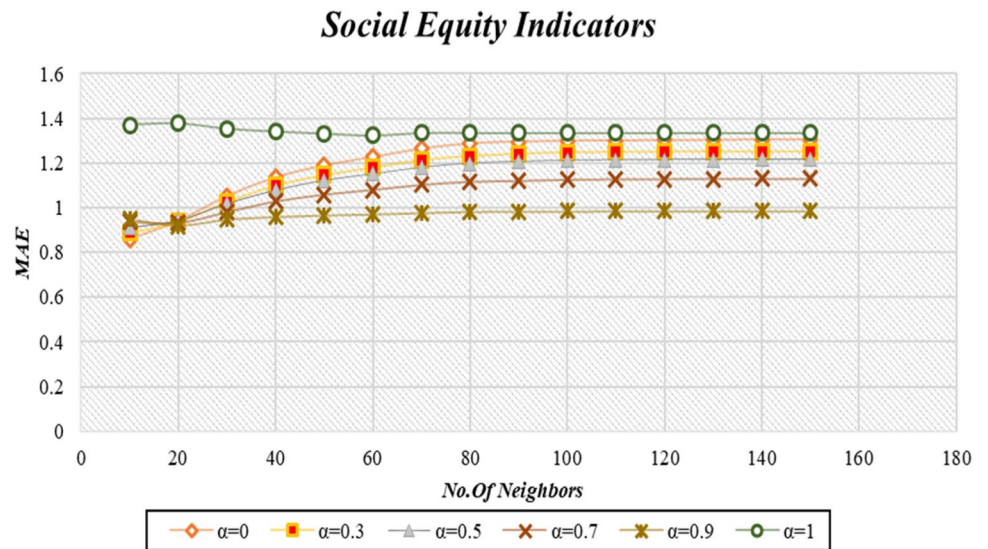


Fig. 19 The impact of no. of neighbors on the recommendation with general social indicators (20.67% improvement)

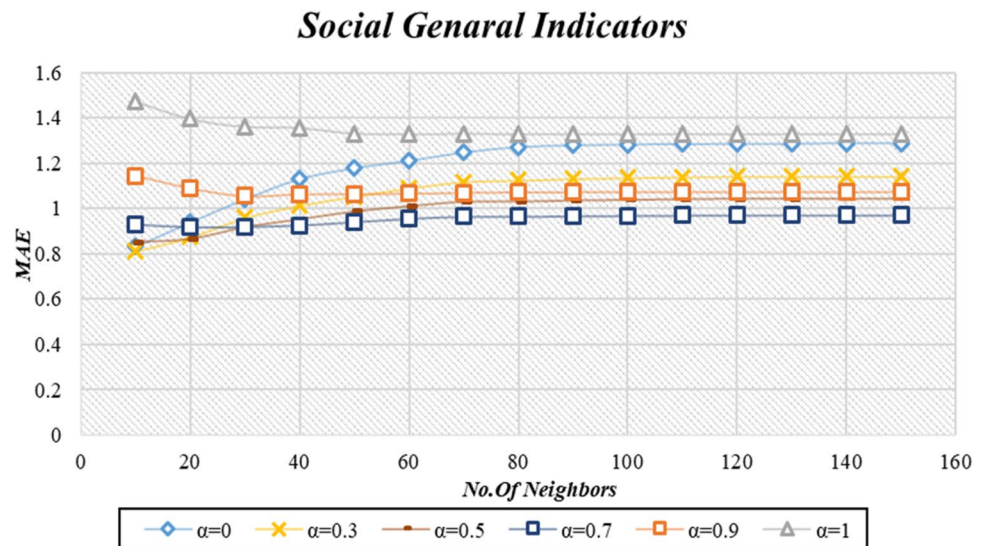


Fig. 20 The impact of no. of neighbors on the recommendation for the complex of social cohesion—self-general-equity indicators (16.09% improvement)

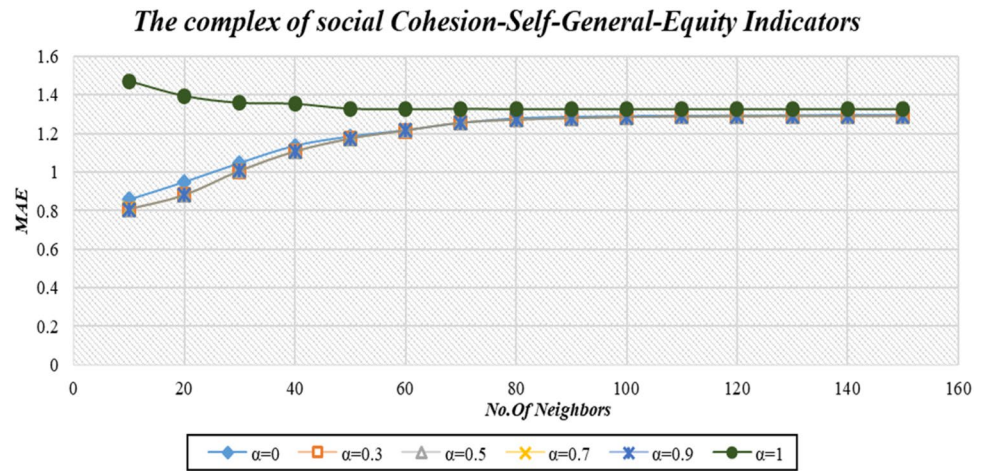


Fig. 21 The impact of no. of neighbors on the recommendation for the mixture of social solidarity indicators, enhanced health indicators (19.82% improvement)

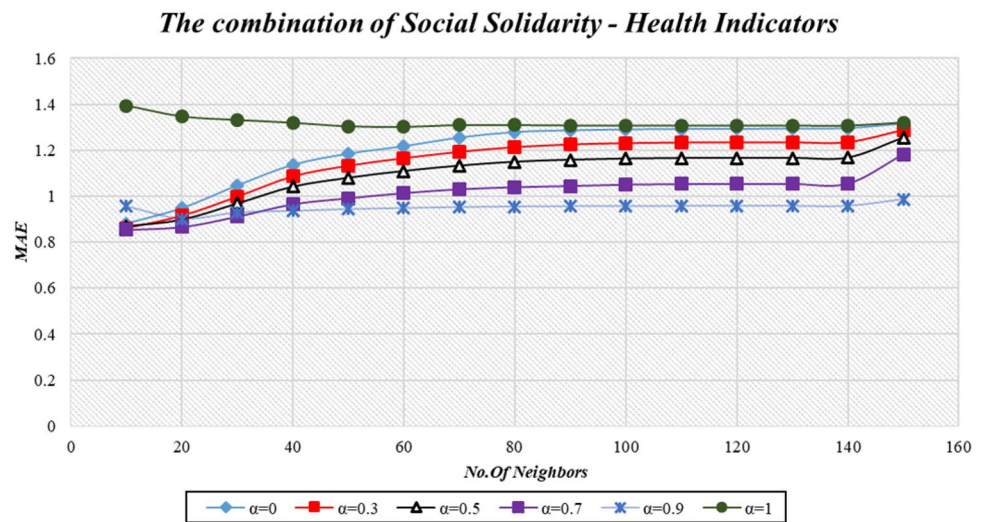


Fig. 22 The impact of no. of neighbors on the recommendation for the complex of social cohesion—self-general-equity indicators, enhanced economic indicators (19.63% improvement)

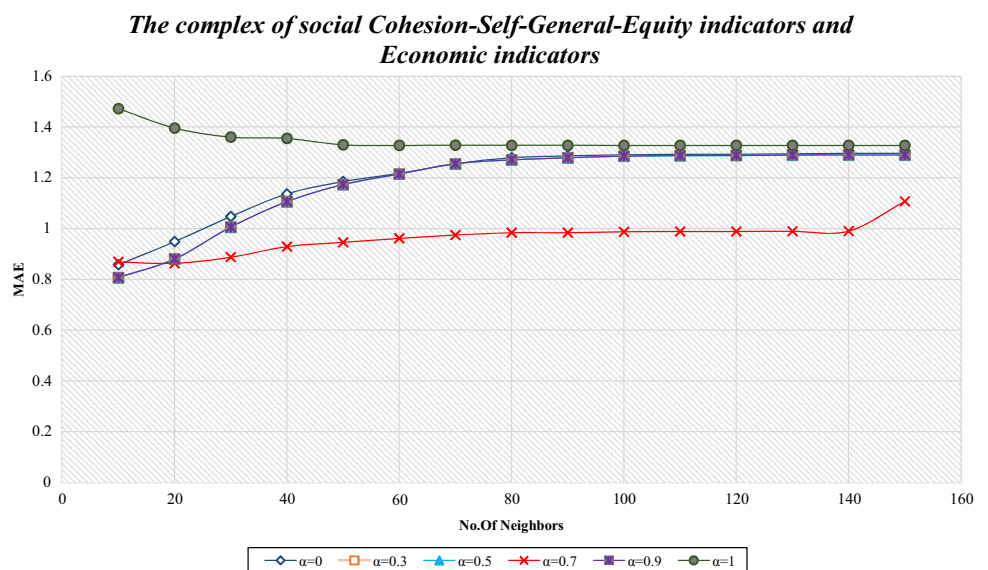


Fig. 23 The impact of no. of neighbors on the recommendation for the mixture of social solidarity-health indicators (15.48% improvement)

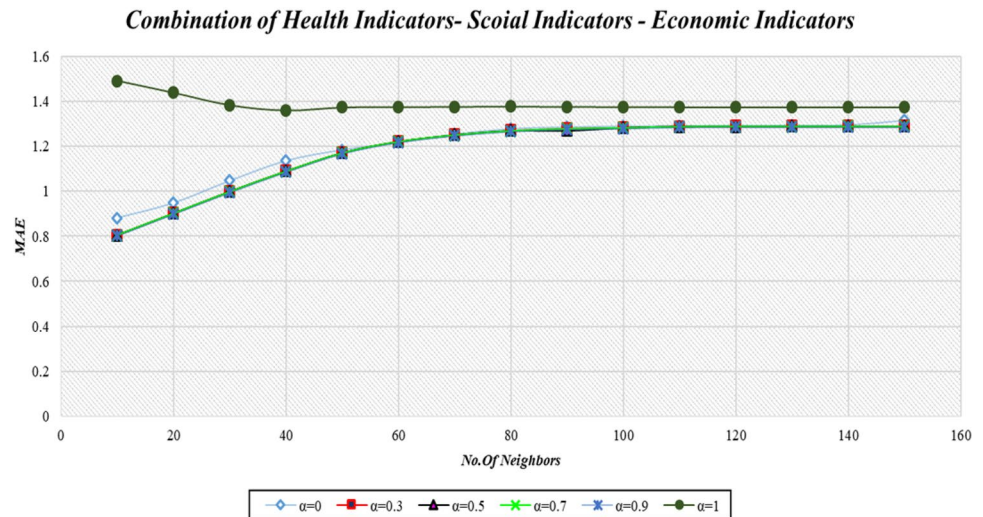
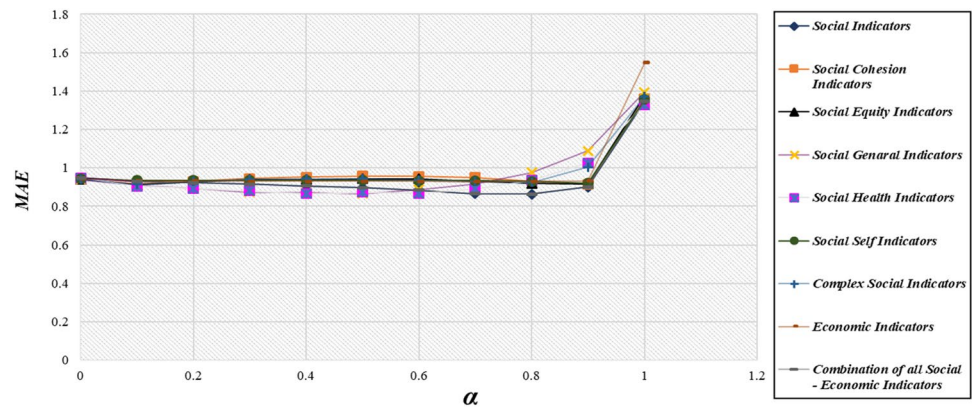


Fig. 24 Impact of α on MAE for social-economic indicators enhanced approach



the accuracy of implementing enterprise filtering systems was realized using the data of social-economic indicators.

Recommender systems are evaluated based on the similarity and closeness of the presented recommendation and users’ opinions. Therefore, the most accuracy between social-economic indicators is related to general indicators that are about 20.67%, and the least accuracy is related to self-indicators that are about 12%. By increasing the number of neighbors in an item-based recommender system, MAE is increased.

In other words, the SEIRS with various values of K shows better results than item-based recommender systems. As illustrated in Fig. 24, for social-economic indicators with $\alpha = 0.8$, and $k = 20$, the algorithm shows the least MAE. Then, the precision results for each individual recommender system based on each social-economic indicator are shown

in Fig. 24, which suggests the better performance of the SEIRS.

Figure 25 evaluates the recommender systems in terms of precision, recall, and f1 applying the social-economic indicators. As much as the F1 value approaches 100%, the output recommendation would be closer to the user’s desire.

In Table 5, the results of the evaluation of accuracy and each of the Precision, Recall and F1 criteria for multiple social-economic indicators in the social-economic indicator recommender system compared to the product-based system are shown.

Recommender systems are evaluated based on the proximity criteria proposal presented to the user. Thus, for example, in systems where users evaluate and rate the items, the distance between the evaluation and the predicted evaluation of the system is decreased and the accuracy of the proposed system will be higher.

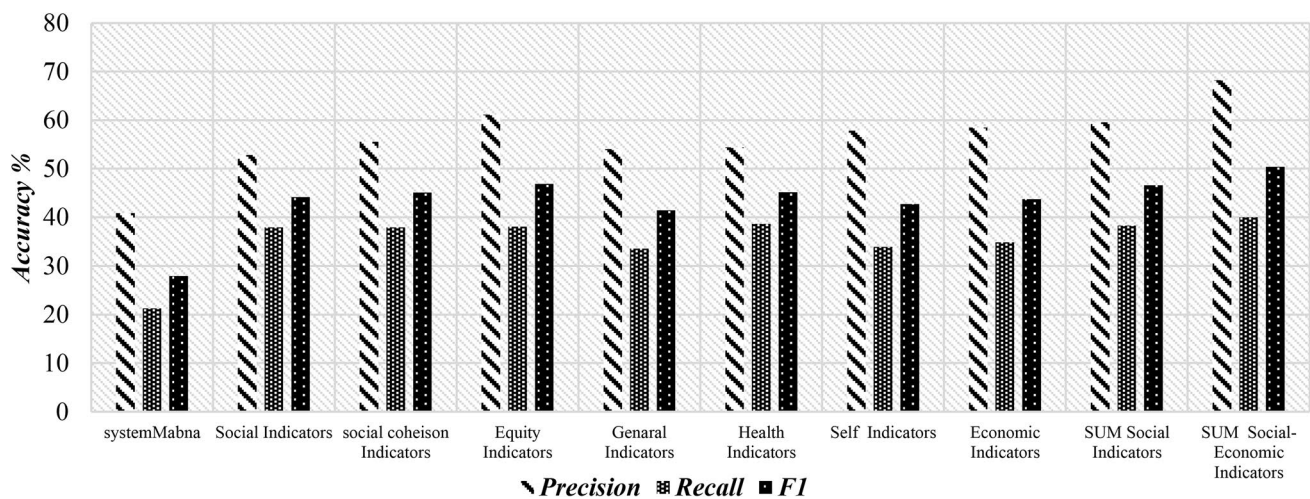


Fig. 25 The precision of the methods with respect to the SEIRS of the user-item matrix

According to Table 5, the recommender system based on social-economic indicators in combination of social-economic indicators was estimated: the MAE criteria (20.91%), Precision criterion (68.18%), Recall (39.97%) and F1 (50.39%), the result verifying the competitive accuracy of the social-economic indicator recommender system over the commodity item-based recommender system.

Experiment-2: Studying the Cold Start Problem

The proposed recommender system that is based on social-economic information is evaluated in this section. Due to entering a new user and not to have any evaluation of items, the item-based collaborating filtering algorithm is not applicable. This is the reason why a basic algorithm is used that uses the average of ratings by other users in order to predict the rating of the new user for an item [6]. The results are illustrated in Table 5. In addition, Fig. 26 shows MAE for ten random executions on both basic and proposed approaches in order to increase the accuracy of the new users.

According to Fig. 26, in all cases, MAE in the social-economic indicator-based system is less than the basic system that is based on the average ratings. According to the results, the percentage of improving cold start users between social indicators related to social solidarity indicators is about 41%, and the least improvement percent related to health indicators is 20.30%, and in general complex of social-economic indicators is near to 25%.

Experiment-3: Studying the Impact of Social-Economic Indicators on Data Sparsity and Cold Start Problem

In this section, the data sparsity problem is resolved by using the SVD algorithm to predict users' priority in collaborative filtering in order to improve prediction accuracy. By processing the SVD algorithm on the rating matrix to solve data sparsity, it also makes us able to improve cold start problems in comparison with the basic algorithm in which the average of item evaluations is considered as a prediction for the new user. According to Table 5, the SVD algorithm and the basic method act similarly in that both apply the average of the evaluations to predict the new user behavior. As a prediction for cold start, this performance will not recur in the process of the SVD algorithm. Figure 27 shows MAE values for random testing of ten newcomers in the proposed social and economic system to increase recommender accuracy and solving data sparsity.

According to Fig. 27, MAE values for ten random newcomers in the SEIRS are less than the baseline research that is based on the average ratings. According to the results, the percentage of improving the SVD algorithm on the rating matrix to solve data sparsity in cold start users' problem between social indicators related to social solidarity indicators is about 40.19%, and the least improvement percent related to Economic indicators is 12.31%, and in general complex of social-economic indicators is near to 16.54%.

Table 5 Improvement comparison of MAE, Precision, Recall and F1 for social-economic indicators

	Baseline approach (item-based recommender system) (%)	Social solidarity indicator (%)	Social cohesion indicator (%)	Social equity indicator (%)	Social general indicator (%)	Social health indicator (%)	Social self-indicator (%)	Economic indicator (%)	The complex of all social indicators (%)	The combination of all social-economic indicators (%)
MAE	10.03	18	18.72	19.36	20.67	13.16	12	17.62	17.19	20.91
Precision	40.82	52.79	55.55	61.15	54.03	54.35	57.84	58.4	59.54	68.18
Recall	21.27	37.9	37.92	38.02	33.55	38.63	33.86	34.85	38.27	39.97
F1	27.96	44.1	45	46.8	41.39	45.16	42.71	43.66	46.59	50.39

Experimental Discussion

In this section, a general study was made on the improvement percent of the proposed method to examine the impact of the social-economic indicators on the accuracy of recommender systems. According to Table 6, the improvement percent of the social-economic indicators is observable in all three experiments. According to the obtained results, the recommender systems were evaluated based on the proximity of recommendations to the user's opinion. For example, in the recommender systems that users evaluate and rate the items, the lower the distance between the evaluation and the predicted evaluation of the system, the higher the accuracy of the proposed system. The highest accuracy among social indicators belonged to general social indicators with 20.67% improvement and social equity indicator with 19.36% improvement. The self-sufficiency indicator showed the lowest accuracy of about 12% and the social health indicator about 13.16%. According to the results, the improvement percent of cold start and data sparsity problems is related to self-indicators about 67%, and the least percent relates to social solidarity indicators about 18%. In general, the percentage of complex social-economic indicators is near to 37% for MAE in comparison with the baseline method.

According to the results, the percentage of improving cold start problems and decrease in data sparsity is the most top percent related to self-indicators about 67%, and the least percent relates to social solidarity indicators about 18%. In general, the percentage of complex social-economic indicators is near to 37% for MAE improvement in comparison with the baseline method.

The MAE improvement percent of the cold start problem without data sparsity in the best condition is with social solidarity indicators that provide about 40%, and the least MAE improvement percent is with the health indicators about 25.18%. The complex of social-economic indicators shows up to 16.54% in terms of MAE in Fig. 28.

Conclusions and Future Works

Social-Economic Indicators in the Recommender system analyzed users' behavior and collecting information, to suggest useful recommendations based on their requirements. Item-based recommender systems suffer from two data sparsity and cold start for new user problems that decrease the accuracy of the recommender systems. In order to resolve the problems, another source of information is required to mitigate the deficiency of the recommender systems whenever enough purchase information is not available. Therefore, considering both social and economic indicators along with users' purchase information, item-based recommender

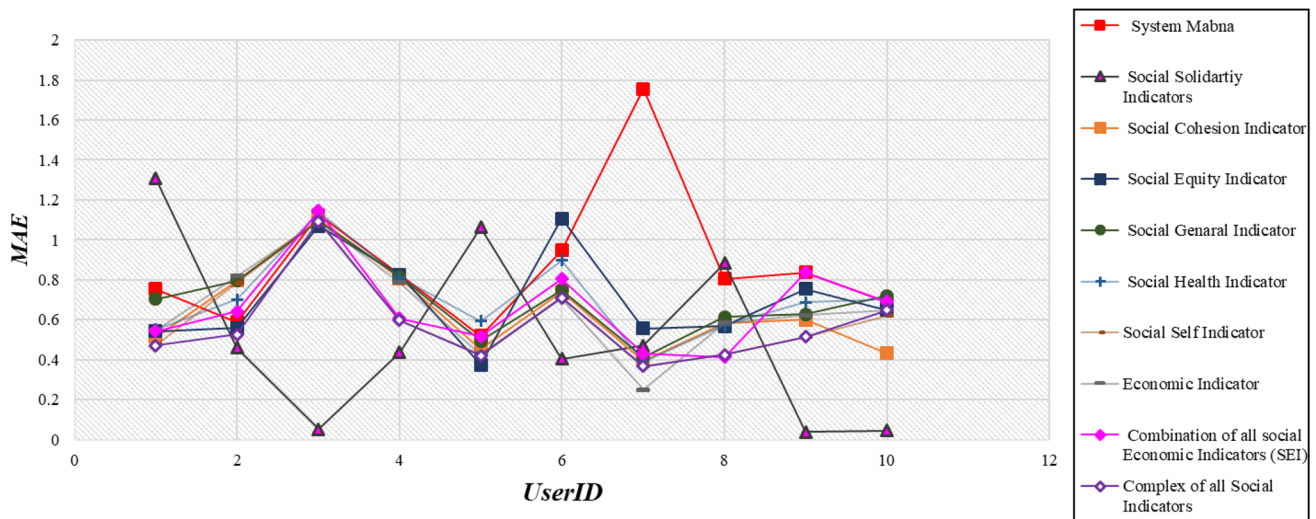


Fig. 26 MAE value for random testing of ten newcomers in the proposed system based on social-economic indicators and average rating-based system

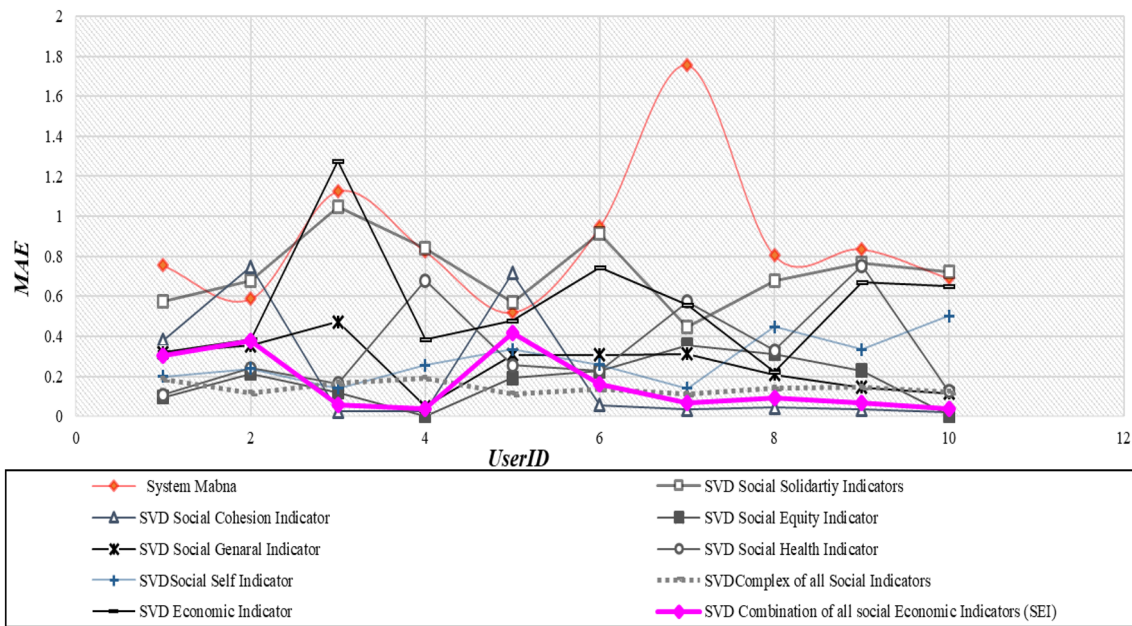


Fig. 27 MAE values for random testing of ten newcomers in the proposed social and economic system to increase recommender accuracy and solving data sparsity

systems could recommend more accurate recommendations of up to 20.91%.

Several ideas may improve the idea presented in this paper. (1) In order to improve social-economic indicators in recommender systems, behaviors of people in different countries to make decisions in similar situations may improve the accuracy of the recommendations. (2) Considering demographics characters of users like age, gender, education,

household income level, etc. may improve the accuracy of recommendations, as well. (3) Users’ purchase information such as rating of items, their history of purchases, date of purchasing, the incidence of purchasing with events, etc. may also provide recommendations that are more accurate. (4) Categorizing friends, family, etc. may give more results that are accurate in social-economic recommender systems. (5) Considering the mixture of social, economic,

Table 6 The summary of results in three experiments

Experimental	Social soli- darity indica- tor (%)	Social cohe- sion indicator (%)	Social equity indicator (%)	Social gen- eral indicator (%)	Social health indicator (%)	Social self- indicator (%)	Economic indicator (%)	The complex of all social indicators (%)	The combina- tion of all social-eco- nomic indica- tors (%)
Studying the accuracy of social-economic indicators	18	18.72	19.36	20.67	13.16	12	17.62	17.19	20.91
The baseline method (basic algorithm) to resolve the cold start problem	41.67	27.38	20.82	20.58	20.30	28.34	26.68	34.69	25.02
The SEIRS improve- ment percent in solv- ing new users' cold start problem against the basic algorithm	18.23	39.90	40.76	28.98	42.59	67.87	35.71	13.31	37.43
The SEIRS improve- ment percent in solving data sparsity problem against the basic algorithm	40.19	36.85	25.18	25.19	27.96	17.72	12.31	17.42	16.54

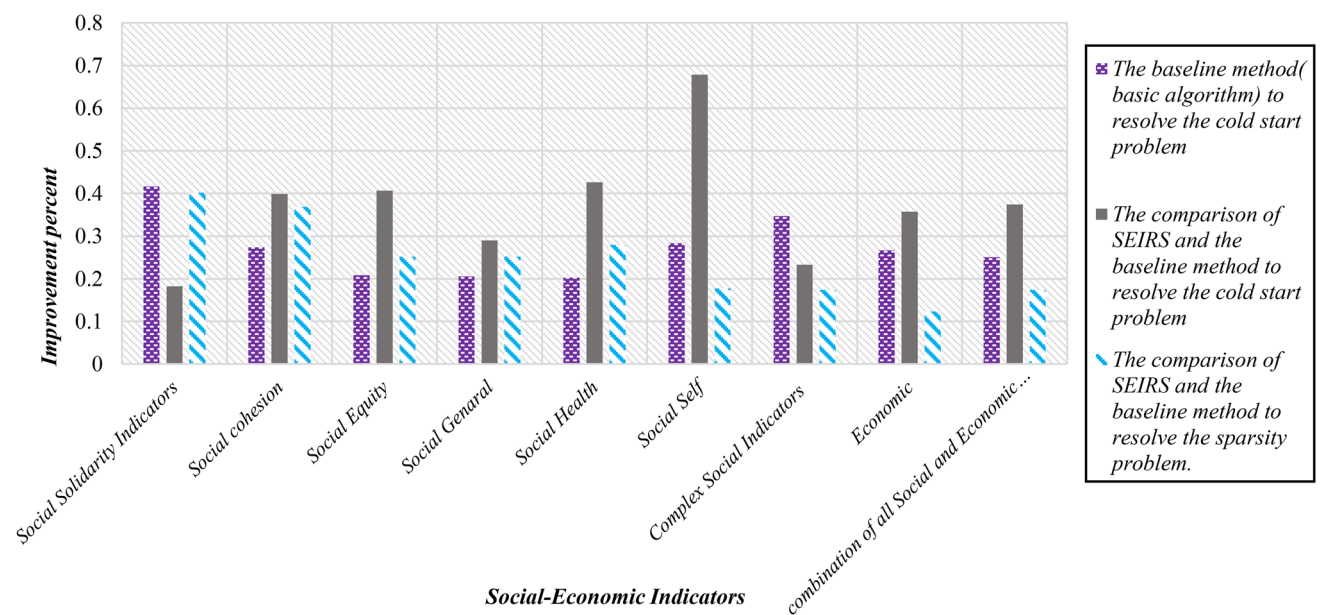


Fig. 28 The comparison of SEIRS and baseline method to resolve the cold start and data sparsity problems

and cultural features of people is expected to present more accurate recommendations.

Funding This study was funded by the authors.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Wei J, He J, Chen K, Zhou Y, Tang Z. Collaborative filtering and deep learning-based recommendation system for cold start items. *Expert Syst Appl.* 2017;69:29–39.
2. Kumar S. Survey on personalized web recommender system. *Int. J. Inf. Eng. Electron. Bus.* 2018;10(4):33.
3. Sarwar B, Karypis G, Konstan J, Riedl J. Item-based collaborative filtering recommendation algorithms. In: *Proceedings of the*

- 10th international conference on world wide web 2001 Vol.1 (pp. 285–295).
4. Zardosht M, Ghasem-Aghaee N. Culturally enhanced collaborative filtering. *World Acad. Sci. Eng. Technol.* 2011;88(59):452–7.
 5. OECD. Social indicators. In: *Society at a glance 2016: OECD social indicators*. <http://dx.doi.org/10.1787/9789264261488-en>, 2016.
 6. Wang Y, Chan SCF, Ngai G. Applicability of demographic recommender system to tourist attractions: a case study on trip advisor. In *Proceedings of The 2012 IEEE/WIC/ACM International joint conferences on web intelligence and intelligent agent technology*, volume 03, 2012, pp. 97–101.
 7. Gong S. A collaborative filtering recommendation algorithm based on user clustering and item clustering. *JSW.* 2010;5(7):745–52.
 8. Hafshejani ZY, Kaedi M, Fatemi A. Improving sparsity and new user problems in collaborative filtering by clustering the personality factors. *Electron Commer Res.* 2018;18(4):1–24.
 9. Sharma M, Mann S. A survey of recommender systems: approaches and limitations. *Int. J. Innov. Eng. Technol.* 2013;2(2):8–14.
 10. Su X, Khoshgoftaar TM. A survey of collaborative filtering techniques. *Adv Artif Intell.* 2009;2009:4.
 11. Zhang S, Wang W, Ford J, Makedon F, Pearlman J. Using singular value decomposition approximation for collaborative filtering. In: *E-commerce technology, 2005. CEC 2005. seventh IEEE international conference on 2005*. p.257–64.
 12. Sarwar B, Karypis G, Konstan J, Riedl J. Incremental singular value decomposition algorithms for highly scalable recommender systems. In: *Fifth international conference on computer and information science, 2002*. p. 27–8.
 13. Mohana H, Suriakala M. A study on ontology-based collaborative filtering recommendation algorithms in E-commerce applications.
 14. De Mooij M. *Consumer behavior, and culture: consequences for global marketing and advertising*. Thousand Oaks: Sage; 2010.
 15. Ma H, Zhou D, Liu C, Lyu MR, King I. Recommender systems with social regularization. In: *Proceedings of the fourth ACM international conference on web search and data mining, 2011*, p. 287–96.
 16. Liu J, Song H, Pei Q, Li Z, Zhan Y, Fan K. TrustCF: a hybrid collaborative filtering recommendation model with trust information. In: *IEEE international conference on communications (ICC). 2018;2018:1–6*.
 17. Gupta J, Gadge J. Performance analysis of recommendation system based on collaborative filtering and demographics. In: *International conference on communication, information & computing technology (ICCICT). 2015;2015:1–6*.
 18. Gartrell CM. Enhancing recommender systems using social indicators. 2014.
 19. OECD. Headline Social indicators. In: *Society at a glance 2009: OECD social indicators*. http://dx.doi.org/10.1787/soc_glance-2008-2-en, 2009.
 20. OECD. Social cohesion indicators. In: *Society at a glance 2005: OECD social indicators*, http://dx.doi.org/10.1787/soc_glance-2005-8-en, 2005.
 21. OECD. Pro-and anti-social behavior. In: *Society at a glance 2011 : OECD social indicators* http://dx.doi.org/10.1787/soc_glance-2011-28-en, 2011.
 22. OECD. Social indicators. In: *Society at a glance 2014: OECD Social indicators*. http://dx.doi.org/10.1787/soc_glance-2014-en, 2014.
 23. OECD. Health at a glance 2013: OECD indicators. OECD publishing. http://dx.doi.org/10.1787/health_glance-2013-en, 2013.
 24. Cleland J, Wilson C. Demand theories of the fertility transition: an iconoclastic view. *Popul. Stud. (NY)*. 1987;41(1):5–30.
 25. Agneessens F, Waeghe H, Lievens J. Diversity in social support by role relations: a typology. *Soc. Networks.* 2006;28(4):427–41.
 26. OECD. Main economic indicators, vol.2. OECD Publishing, Paris <http://dx.doi.org/10.1787/mei-v2018-2-en>, 2018.
 27. Hicks N, Streeten P. Indicator of development: the search for a basic needs yardstick. *Dev. Areas A B. Readings Res.* 1995;27:31–44.
 28. Boarini R, Johansson Å, D'Ercole MM. Alternative measures of well-being. OECD Library, 2006.
 29. Hicks N, Streeten P. Indicators of development: the search for a basic needs yardstick. *World Dev.* 1979;7(6):567–80.
 30. OECD. Pro-and anti-social behavior. In: *Society at a glance 2011 : OECD social indicators*, http://dx.doi.org/10.1787/soc_glance-2011-28-en.
 31. Al Faris A., Soto R, editors. *The economy of Dubai*. Oxford University Press; 2016.
 32. Wen Z. Recommendation system based on collaborative filtering. In: *CS229 lecture notes*. 2008.
 33. Kunaver M, Požrl T, Pogačnik M, Tasič J. Optimisation of combined collaborative recommender systems. *AEU-Int J. Electron. Commun.* 2007;61(7):433–43.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.