

Application of Recommender Systems on Web Portals to Increase User's Visit Statistics

Parnian Zare

Department of Information Technology Engineering, Shiraz Branch,
Islamic Azad University, Shiraz, Iran

Abstract

In today's competitive world, the rapid advancement of technology, Internet and electronic commerce raise demand for existence of a mechanism which can predict user's requirements and requests. These mechanisms can lead us to outsource our competitors. The main issue is that we encounter large amount of information in web portals which are mostly heterogeneous and unrelated, therefore with no proper strategies of categorizing data and information preparation, users get involved in confusion accessing correct content. The most important challenges are reaching most relevant information in order to provide users. As a matter of fact, this problem could get solved by using domain of recommender systems which can help us finding and selecting related information according to user needs. Although, recommender systems help people dealing with massive data, these systems are less employed on Web portals. Certainly, the application of these types of systems on web portals will bring decent improvements to users. The study, uses MovieLens20m dataset that includes ratings and user labels for movies. User ratings and movie rating relationships are used in order to make an appropriate recommendation for users. labels which users considered for the movies are also employed in extraction phase. Finally, a combination of these two categories is offered as a recommendation to user.

Keywords: Recommender Systems, Web Portals, User Visits, MovieLens

DOI: 10.7176/JIEA/9-3-03

Publication date: May 31st 2019

1. Introduction

The electronic world is moving towards information saturation since over the past decade, a large amount of data has been stored on data servers and databases. Nowadays, the amount of available data is doubled every 5 years therefore considering a method employing wide variety of information related to markets, services and business environments in order to gain right data accessibility is essential to decision making process. High data volume on the Internet and its improvements have caused major issues for users to find relative information due to products and services in appropriate time. In other words, finding data and final information tailored to user needs has become a complex and time-consuming process. [1-3]. In such environments, existence of a system capable of recognizing and updating user interests and priorities plus the ability to index and store information in a searchable manner is strongly suggested. Recommender systems can direct users in electronic environments to find qualified information based on different services, they also have ability to identify users and prioritizing them. Furthermore, recommender systems gather user interest's history to maintain acceptable time and energy savings. [3-6]. They also extract user's past behavior, even services and information that he did not notice, in order to predict and suggest interesting results. In fact, these systems are one of the main tools of overcoming information accumulation. They become intellectual segments of data retrieval and screening systems because of their user behavior analyzing power. Today, recommender systems are utilized in various fields of web data in accordance with user requests, e-commerce, film suggestions, music, books and articles. Considering the position of this system in today's e-world and taking into account user's expectations to access correct information in the right place at the right time, researchers have introduced recommender systems a tool for empowering users in order to find qualified services with optimal manner [6-8].

2. Research Background

Following topics are discussed regarding to user interests and preferences in recommender systems [9]. In this research, the role of user modellings and personalized accessibility is vital. When massive amount of information is sent to users, they got confused in choosing needed products or information. Therefore, users need a personalized backup to change provided information volume and type according to the interests. Information filtering systems, such as user behavior monitoring systems, are designed to create personal profiles for users based on their interests and click tracking in order to help information screening process. One of the good features of the recommender systems is that they can extract and subdivide inconsistent information. Although web portals direct huge amount of information to users, displaying selected information will not develop a single problem, so using recommender systems in web-portals is crucial. While, recommender systems are not usually used in web portals, their application in web portal architecture can certainly have a positive impact on overall performance. Because if users find that they are looking at their preferences and interests, they are more likely to visit the portal. [9] Lee et al.

believe that increasing webcam quality can attract users' perceptions [10]. They considered four factors of data adequacy, beautiful appearance, usability and security are critical for attracting viewers. Although it has been acknowledged that only two first factors are enough to satisfy users at first glance, we have to mention that addressing further factors will attract other users to employ the services provided by the portal. For example, we often deal with organizational portals, and sometimes due to the scope of the work, there is no way to offer suggestions, so users must perform a series of routine tasks in accordance with certain procedures. But in commercial, news and educational portals, there should be some possibilities to search and offer items that are close to user's needs or desires [10]. Schuhann et al. attempted to automatically search the web with the help of suggestions provided by the system [11]. In their proposed system, which is now considered trivial, they provide series of web pages plus evaluation process based on page attractiveness for users. The feedbacks were returned back to the system in order to perceive better suggestions for future. In this type of system, heuristic searches for similar user selections were considered, perhaps their contents were not much important. But now the content is the first priority for the user [11]. During the research, the issue of usability and Web portal designs were discussed [10]. Mohammad Ali addressed the issue of increasing patients' application of online medical webcams. Medical web portals include advisories, record reviews and service charging [12]. Herrera et al. examines a qualitative model for assessing the quality of Web portals. This model is based on the ISO standard [13]. Sharma and Agrawal provided a report on several recommender systems used in Web portals. They have looked at different types of recommender systems employed in each Web portals [14].

3. Proposed Method

The vast amount of information on the Web and the Internet has made it difficult for many web users to make decisions and select related information to data or products they need. This was the motivation for researchers to find a solution to this fundamental problem of the new age known as data overflow. To address this issue, a recommender system has been proposed based on previous user requests, product labels which are made by the manufacturer and user's feedbacks. In this research, both previous user requests and requests from all users are considered and prioritizing occurs due to previous users' feedbacks. The proposed method is divided into two main phases: 1) when the user enters the site; 2) when the user enters his request in the form of a phrase. See the process as shown below.

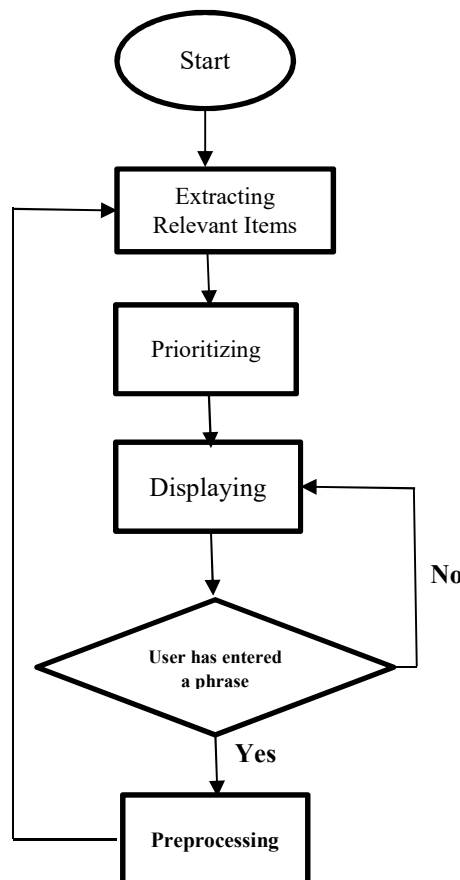


Figure 1. Flowchart of Proposed Method

4. Initial Recommendation

As previously stated, different suggestions are recommended to user as he enters the site. These suggestions are preceded before any input from the user. This phase consists of three different parts which are extracting relevant items, prioritizing and displaying. the second and third sections are related to the phrases that user enters. but the first section which is the most important part of this phase is used to make initial recommendations for user.

4.1 Extracting Relevant Items

In this research it is assumed that the content is labeled, for example, if we consider a movie provider site, genre is pre-defined for each movie. This assumption is logical because manufacturers know that which category their products are classified in and what labels they maintain. According to this default, firstly we pay attention to users previous visited contents, then by gathering the labels which are used for the seen contents we can find relevant new content. For example, if a user has already selected a movie that is a fictional genre, imaginary films are extracted.

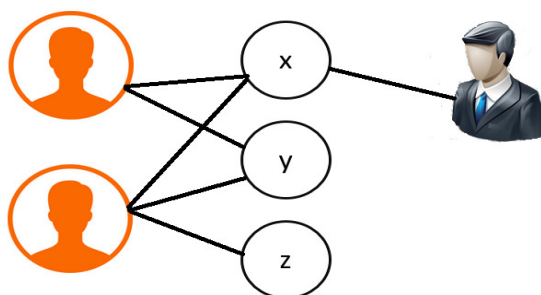


Figure 2. An example of user requests for different content

On the other hand, according to the user-visited content, a search is done based on other users in order to find other users visited contents. For example, if the user has visited the content x, a search will be done based on the users who have visited content x. After finding these users, the items that have been viewed will be listed. We gave all of these retrieved items a single rate. This rating is based on the replies of this content. For example, consider another user who has visited the content x and then visited the content y, according to previous number of users who have visited x and y, the rate is assigned. For each repeat of these two items the rate would get increased. In addition, the visiting time difference between two items x and y is considered. For example, when users view x and y contents with a distance of less than one value, more points would be considered for displaying content y. The pseudo-code which is used for this purpose is shown in the Figure 3. For rating each repetition, a number between zero and one is considered and precise determination of this value is assigned according to the time it takes. In order to get involved in timeframe, it should be noted that what kind of information is recommender system offering to users. For example, the timeframe intended for clothing store site varies with the timeframe for a movie site. this is because users usually change their opinion to buy clothes over short period of time, while changing the views about movies are more in the same timeframe. To this end, the determination of timeframe in this research has been taken as system-user determinable variable.

	Item1	Item2	...	Itemn
User1	R11	R12		R1n
User2	R21	R22		R2n
...				
Userm	Rm1	Rm2		Rmn

Table 1. Users, items and rates allocation

As shown in Table 1, a rate is assigned to each item for different users, which is specified by R. The first index of R specifies the user and the second index of R specifies the item. This table includes zero and one value. Value zero indicates that user did not view the item while value one indicates user visited the item. If user has given a rate to the viewed item, then given rate is considered instead of zero and one. The set of items which are viewed by specific user get extracted and other common users are extracted through these items. After that, by using other users', new items will be extracted.

	Itemn1	Itemn2	...	Itemnj
Userm1	Rm1,n1	Rm1,n2		Rm1,nj
Userm2	Rm2,n1	Rm2,n2		Rm2,nj
...				
Usermi	Rmi,n1	Rmi,n2		Rmi,nj

Table 2. Results of user extraction from common items

You can see the results of other user extraction from common items in the Table 2. All items are viewed by different users. Given these users Userm1, Userm2, ..., Usermi, new items are extracted. If items are rated by users, rating score will be used for items, so for each item that is extracted through a user, one score is obtained, moreover the item which user is interested in may be extracted from one user or other users. Finally, for each item these ratings are collected and considered as the final rate of the item. Item rate calculation is shown in Formula 1.

$$S_{item\ i} = \left(\left(\frac{1}{x_1} \right) \times R_{m_1, n_1} \right) + \left(\left(\frac{1}{x_2} \right) \times R_{m_2, n_1} \right) + R_{m_2, i} \quad x_1, x_2 > 1 \quad \text{Formula 1}$$

In Formula 1, $S_{item\ i}$ means the rate of the item i . In this formula, the user is m_2 , which extracted item i from item n_1 . Item i has been extracted for displaying to m_1 . R also means a rate which user has given to viewed items. As previously stated, for each item extracted, this score is calculated. If an item extracted from several users, the total rate of all items will be considered. After extracting the same items, we sort them based on their ratings, so the content with higher rate is prioritized. The results obtained from this section can Combine with label strategy. For the combination of these scores we should consider common subset of two sets, then the contents related to other users ultimately will be assigned to same label.

4.2 Prioritizing

It is assumed that a rating system is considered for each user after viewing the product, and users should specify a score after viewing each item. It is usually considered as a 5-star rating system presented on different websites, where each star can include three modes; silent, semi-bright and bright. User can fill the stars with the desired score. This score shows users satisfaction in confrontation to the content. The average of users score for a specific item could be considered as one criterion for prioritizing the display. In the proposed method, this scoring is used in two phases. Firstly, to sort the content that user viewed and secondly to prioritize similar content extracted in previous step. In the first phase items that have been assigned to higher scores have higher priority and if rating is not performed, they are considered to be the same priority. In the second phase each content contains average user ratings. These rates which are given by the user and also by other users to extracted contents are used to increase the rate calculated in the previous step. Applying rate increments can be different and in this research is considered as a system-user determinable variable. After rate increment, the extracted content will be reassessed based on the new values of the score.

4.3 Displaying

In this section, the extracted content is displayed to intended user based on the prioritization. Displaying methods and user references to the content can be considered as a feedback in order to examine correctness of user's interest detection. For example, a user who visits the first few pages indicates correct diagnosis of the system, in comparison, when user does not use any of the suggested contents or visits low-priority contents, he indicates a poor diagnosis for recommendation system.

```

Procedure Relevance Items

input:
R : set of contents with tag same User Visited contents
output:
C: set of Relevance Item for User i
Begin
    D ← Visited contents with User Visited contents
    R ← ∅
    C ← ∅
    for each Visited content  $D_j \in D(n - 1)$  do
        Calculate the similarity ( $D_j$ )
    End for
    Sort D According to similarity Feature
    C ← R ∩ D
    C ← C ∪ D
    C ← C ∪ R
end procedure
    
```

Figure 3. Relevant item extraction code

5. Recommendation Based on Users Requests

In this section, recommender system is going to consider user's request. These requests are mainly in the form of words or phrases. It should be noted that many users counted as ordinary so proposed words or phrases may not be exact. By the way, the system's purpose is to help user finding related content faster.

5.1 Preprocessing

After system receives words or phrases from the user, preprocessing task begins. This process extracts important words from the whole phrase. We have to mention that preprocessing task includes different segments which require a study through linguistics. For example, preprocessing can be used to correct spelling mistakes in the system which can cause better identification of user related words. In this research, we examine each word individually by cutting users phrases in to words. After that, we try to remove additional marks that may accompany some of the words. For example, when word containing "@" has been added, "@" will be removed from the word in preprocessing task. This process also deletes relative-words like 'because' and prepositions like "with". Finally, remaining words are given as an input for the next section. It should be noted that these input words may include inappropriate or even meaningless words. This issue is inevitable in the proposed research. Furthermore, removing these words require precise examination of the language with high processing volume. This should be taken into account that since these words have low frequency, they will become worthless in the system over a time with no impact on related item extraction.

5.2 Extracting Relevant Items

Products are going to be extracted and displayed to users by utilizing words obtained from previous step. First, the words are given specific priority which illustrates importance of the word in extracting content. This priority is measured by calculating word frequency in other user requests. Therefore, the word with higher frequency is given higher priority. After determining the priority, related words associated with the user requests are verified and subsequently relevant content would be selected. Contents are searched based on word labels of user requests, then the items that have label of the user's words or the title included user words are displayed as the first results to the user. Of course, it should be noted that in the proposed system the main focus is on the user's inaccurate requests; otherwise, a simple search will take place and the system will lose its true application.

For each word, a monitoring task is considered in previous requests. Requests that contain mentioned words are extracted, then mentioned words will be extracted in the whole requests in order to find frequency. For example, if user inserts x and y in his request, for both x and y monitoring task will be done. If x is supposed to be used in three other user requests, then the words that come with x are extracted. Consider these three requests are (x, z), (x, r) and (x, z). The score of the word z is equal to two, and the score of the word r is equal to one. The pseudo-code of this algorithm is shown in Figure 4. This pseudo-code will be called for each user-requested word. After obtaining the words used in conjunction with mentioned words, we apply a subscribing task. This subscription shows relevance rate of user's words. In subscribing task, prioritization is determined according to the rate of each word and accompanying related words. After completing these steps, words and corresponded rates are sorted. Words with higher rates are ranked first.

```

Procedure Relevance word
input:
W : Word request user
output:
C : set of Relevance words for W
Begin
D ← User Requests
R ← ∅
C ← ∅
for each request  $D_j \in D(n - 1)$  do
    if  $W \in D_j$ 
        words request  $D_j$  add R
    end if
end for
for each word  $R_j \in R(n - 1)$  do
    if  $R_j \in C$ 
         $C_i ++$ 
    else
        words  $R_j$  add C
    end if
End for
Sort C According to Rating
end procedure
    
```

Figure 4. Words relationship in user requests code

Before explaining how to extract related words, it's necessary to consider user's request and corresponded content he visited. User words and visited content should be stored for each mentioned word. For example, when user employs three words in order to access related content, for each three words a separate content ID is stored. In the next phase, the relationship between each word and content ID would be determined. This relationship is shown in form of a number between zero and one. This Relation is obtained from each word based on content ID referring. for example, consider the word x , assuming that the word is used for three contents with identifiers 1, 2, 3. If one assumes that the word x is used for the content with ID 1 seven times, ID 2 ten times, ID 3 three times, word relation for these three contents is respectively 0.35, 0.5 and 0.15.

Content is extracted in order to make recommendation for users by employing stored words which are ranked by rating scores. In the extraction of these contents, words are also considered. At first, these words are placed on recommended content, then extracted content and related words are prioritized based on their relation. This priority is based on words rating score and word-content relation.

After extracting recommended contents, we assign each one a rating score. Finally, initial recommendation results and corresponded ratings are used in final content extraction. in this section, these two categories which are initial recommendation and user request-based recommendation are combined in to one aggregated recommendation with specified rating score ready to display to user. The point to be taken into consideration is content relative value which is perceived as conditional variable.

Two sections of prioritizing and displaying are similar to the initial recommendation described in previous section.

6. Timing of Proposed Method

For each algorithm, timing plays an important role gaining faster executions. As in this research, timing is directly related to user, we consider timeframe of recommender system based on user's request. The first phase which introduced as initial recommendation is less time-consuming, therefore it has major effects in final results. As stated, preprocessing task is performed immediately after user request insertion. Due to the fact that preprocessing time is only depend on the user's request and in most cases the request is not containing more than 10 words, we can skip preprocessing execution time.

In the extracting phase, two main tasks occur which are word and content extractions. In order to consider word extraction a search will be performed among all user requests so if we consider n as the number of requests, search time will be order of $\theta(n)$. After that, for gaining rating score of the words we consider their frequency. Rating score specification in worst case would be order of $\theta(n^2)$. Considering that only those requests with related mentioned words are selected, rating time will be much less than $\theta(n^2)$. Finally, content extraction in relation with user words will be in order of $\theta(n)$.

In the situation of multiple user requests and system low performance, we can perform word-based search in user requests. As each word contains a unique value, by mapping them in to memory blocks we maintain search time in order of $\theta(1)$. Lastly, we can assign each word a content ID. This approach is very effective for high volume user requests.

For prioritizing task if we consider m as the number of contents, timing would be the order of $\theta(m)$. At the beginning of recommender system execution, m is greater than n , but by time passing, user request maintaining and content rating, n will grow greater than m . It can be concluded that total timing of $\theta(m)$ series are lower than $\theta(n)$.

7. Dataset of MovieLens 20M

The MovieLens 20M dataset is provided for movies. This dataset contains three sections of movies, ratings and tags presented in csv files. In the movie file we have movie ID, title and genre. This file contains more than 27,000 different movies. The rating file includes user ID, movie ID, rating score and timestamp. This file contains more than 20 million user rates. The tag file includes labels which are given by users to movies. Each row contains user ID, movie ID, user assigned tag and timestamp. This file covers 460,000 rows.

File	Sample number
Movies	27,278
Rating	20,000,263
Tags	465,564

Table 3. Dataset Information

The dataset has been collected from 138,000 users which are distinguished by identities and no more information is available. This dataset is performed from 1995 to 2015 and updated in 2016.

8. Evaluation

score and accompanying words frequency score, are added in order to maintain final score. The words which are

directly obtained from user requests, are added to extracted words and the rating score is multiplied by factor of 3. On the other hand, movies are extracted from these words, and each movie contains a score for the words. Scores are derived from user's application of the word for specified movie. So, for the final score calculation, total score of the movie is multiplied by factor of 3 plus the word rating score. Also, the user's requested words are used in labels and movie titles.

Second, for the extraction of suggestions, users mentioned movies which are presented in tag file are used. In way that movies which users have viewed in relation with the mentioned movie is extracted. The more users viewed the extracted movie the higher rating score it concludes. To continue evaluation phase scores obtained from first and second parts are used in order to evaluate performance of the proposed method. The tag file has been utilized for evaluation since the beginning of the series, and the results are presented in five categories. The first four categories, contain hundred thousand tags and the last category includes over 65 thousand labels.

The evaluation criterion applied in this research is based on precision which is obtained from the proportion of correct suggestions to all offers of recommender system.

8.1 Tools and Environment

To implement proposed method, Windows operating system version 10 was used on the system with 4GB original hardware and the Intel core i5 CPU. Also, C # programming language has been used.

8.2 Proposed Evaluation Method

First, every single word obtained in the form of a user request is given a score; this score is calculated in the system for each word. According to these words, the words accompanying them in other user requests are extracted, depending on the frequency of the words by other users, another rating score is assigned. As stated, system words added together and results are shown to user based on this measurement. Movies which earn higher rating score has higher priority in displaying to user.

The results of four evaluation models are presented on the 20M MovieLens Dataset. These four models comprise results using words, user rating scores, viewed words and movies, and once again viewed words and movies plus rating score. Results are shown in 5 rows presenting classification of the dataset in to five categories of hundred thousand. The results are also represented in 5 columns. The first category illustrates the percentage of states that user request includes in first 10 suggestions of recommender system, the second category represents suggestions of 11 to 30, the third category represents suggestions of 31 to 60, the fourth category represents suggestions of more than 60, and the last one illustrates the state in which user request does not exist among all system suggestions.

8.3 Evaluation Based on User words

The results of user words evaluation are presented in Table 4. In the first category, results are lower than other categories. This is because there is no information about user's movie applications. Given the fact that dataset is arranged using timeframe, there is no first indication of what word is user going to employ for movie label, therefore many of the movies are seen by the user for the first time which is a reason for not having appropriate recommendations and less qualified results. Another point to note is similar results in other categories which caused by extraction of the words and their relation to the movies. Another consideration is user possibility of first entrance in each timeframe with no regarded information, that can initiate increment in offering inappropriate recommendations.

	1-10	11-30	31-60	60 <	Not recommended
First category	14.17	8.29	13.71	15.51	48.32
Second category	23.11	15.08	14.87	20.27	26.67
Third category	24.13	15.76	14.74	21.16	24.21
Fourth category	23.67	15.39	15.25	21.02	24.67
Fifth category	22.99	16.42	15.31	21.44	23.84

Table 4. Evaluation results based on user words

In Figure 5, you see evaluation results graph. This graph shows the percentage of correct recommendations in each timeframe for different categories.

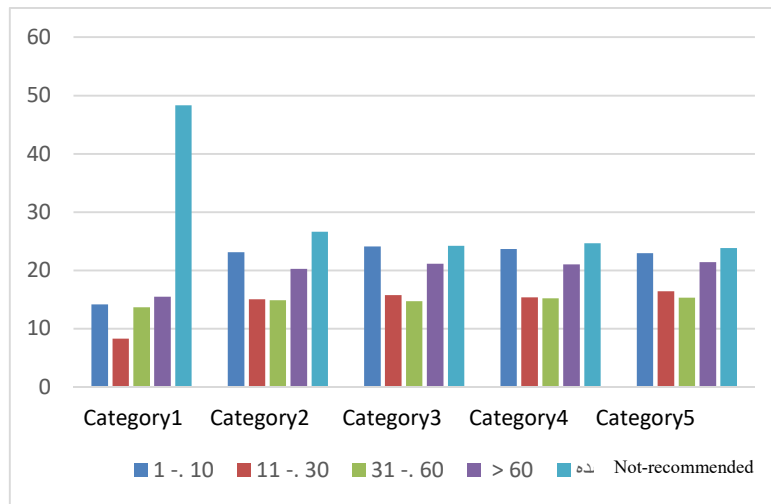


Figure 5. User words evaluation graph

In the graph, accuracy development observed in terms of increasing input data. In horizontal mode, the graph indicates hundred thousand numbers and in vertical mode graph illustrates correct percentage of recommendations. In this graph results are divided into two general categories which are suggested and not suggested. All categories will map to one of these two classes depending on user's request. Another point to note is that each hundred thousand category utilizes previous categories information. For example, words and movies which are employed by users are going to be reused in next category.

As it can be seen in Figure 6, first category correct recommendations are 51.68%. In the second category, there has been a significant improvement to 73.33%. Final recommendations with regard to all categories and division into successful and unsuccessful classes is 70.03%. Above all, we can conclude that 70.03% of the movies recommended by the system are correct regardless of the number of suggestions.

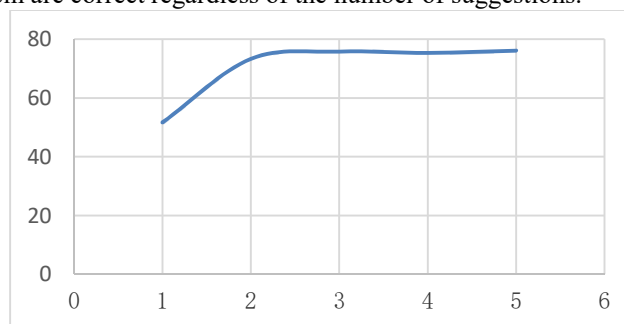


Figure 6. Precision of the proposed method using words

8.4. Evaluation Based on User Ratings and User Words

Utilizing user rating scores of suggested movies, has affected correct recommendation procedures. The effect is on the movies which are being recommended to user, therefore it does not have any influence on whether to recommend a movie or not. The results of this evaluation are shown in Table 5. Obviously, in many cases results have been improved, because of the efficacy of rating score and user words in recommendation displaying. Of course, it should be noted that this process does not always raise system rating scores, and in some cases, it can even reduce it, but mostly this trend is positive and can be beneficial.

	1-10	11-30	31-60	60 <	Not recommended
First category	16.03	9.77	15.42	10.46	48.32
Second category	26.34	17.28	16.79	12.92	26.67
Third category	25.78	16.97	18.87	14.17	24.21
Fourth category	26.80	18.21	17.85	12.47	24.67
Fifth category	26.06	17.63	19.26	13.21	23.84

Table 5. Evaluation results based on user words and user ratings

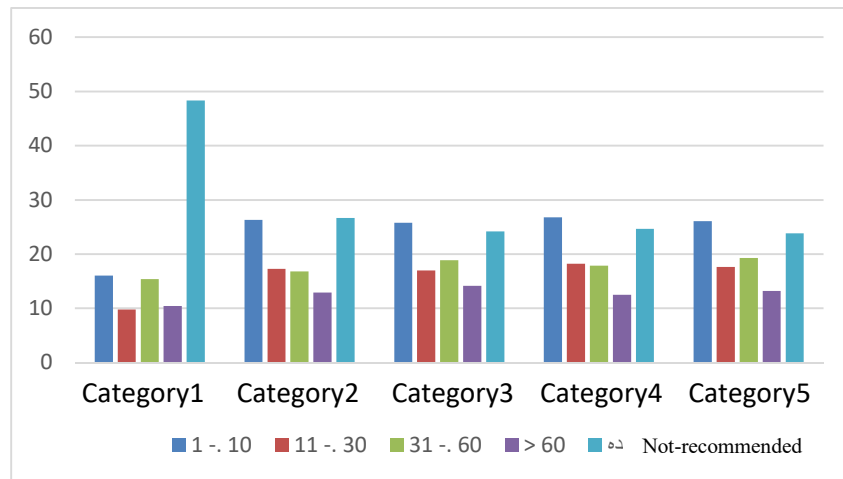


Figure 7. User words and user ratings evaluation graph

8.5. Evaluation Based on User Words and Viewed Movies

In this section, the evaluation results are presented, employing user words and viewed movies. As previously stated, evaluated movies are combined into two sections. Movies which belong to both sections will be added in terms of rating score and consequently final score will be obtained for each movie. The results of this evaluation are shown in Table 6.

	1-10	11-30	31-60	60 <	Not recommended
First category	21.43	17.81	20.02	13.21	27.53
Second category	29.38	23.83	20.09	16.74	10.23
Third category	29.87	22.95	20.78	16.62	9.87
Fourth category	28.91	21.43	22.54	16.51	10.61
Fifth category	29.04	22.11	22.89	16.67	9.29

Table 6. Evaluation results based on user words and viewed movies

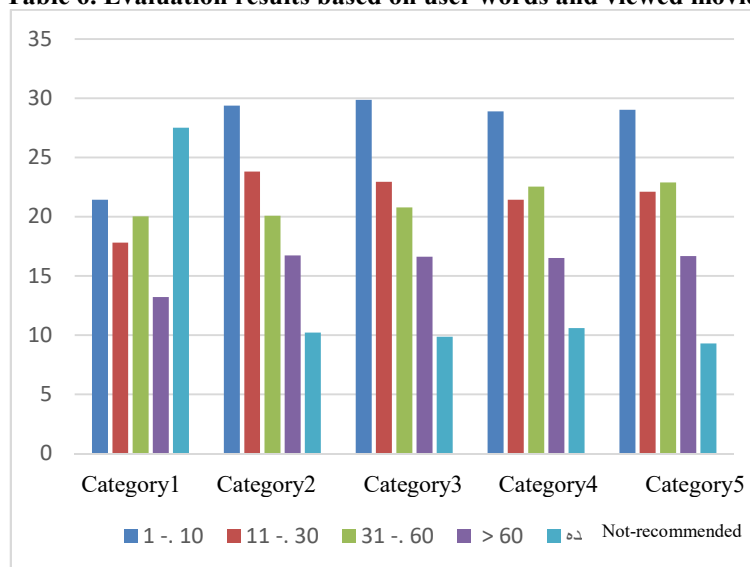


Figure 8. User words and viewed movies evaluation graph

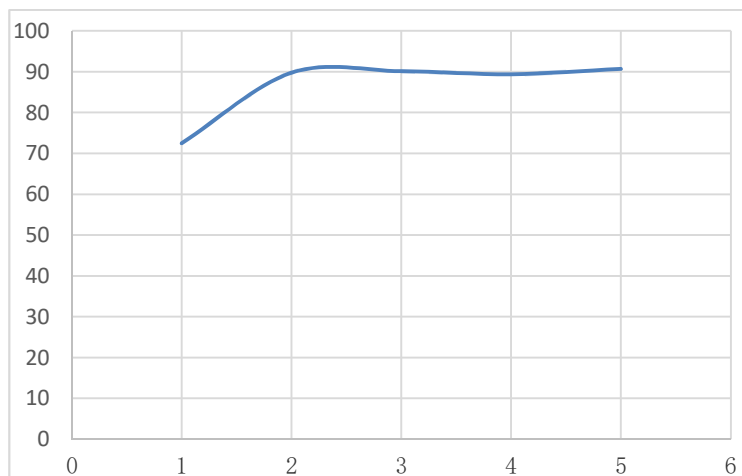


Figure 9. Precision of the proposed method using words and viewed movies

8.6. Evaluation Based on User Words, Viewed Movies and User Ratings

In this section, we assign scores to user-selected movies from the rating file. This task is considered for all extracted movies. As stated, employing user ratings for movies does not affect movie selection process and it only increases the rating score of better movies.

	1-10	11-30	31-60	60 <	Not recommended
First category	22.43	18.67	21.11	10.26	27.53
Second category	30.81	25.12	22.60	11.24	10.23
Third category	31.19	24.89	22.29	11.76	9.87
Fourth category	31.12	24.17	23.03	11.07	10.61
Fifth category	30.92	24.41	24	11.38	9.29

Table 7. Evaluation results based on user words, viewed movies and user ratings

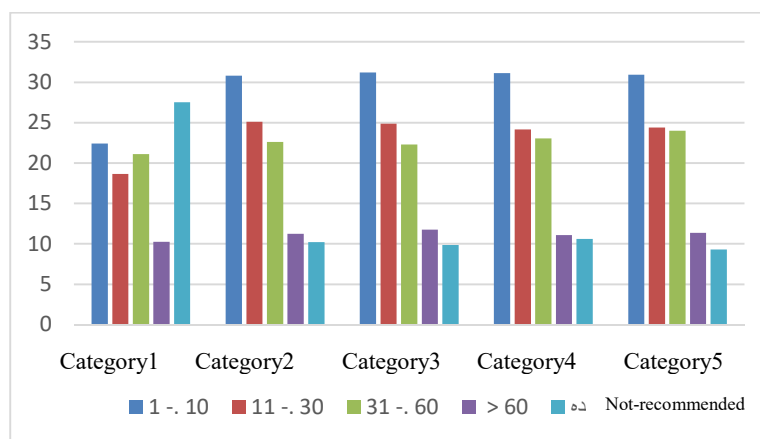


Figure 10. User words, viewed movies and user ratings evaluation graph

As it can be derived from evaluation results, this section contains best performance of the proposed system.

9. Conclusion

Recommender systems, suggest most appropriate services such as data, information, products and items by analyzing user's behavior. Simply, recommender systems attempt to identify most applicable and closest item to user based on their attitude and perception with the help of information obtained from his behavior in relation to other similar users. In this research, a method has been proposed for recommender systems which mainly applies information retrieval techniques on a movie dataset consisting of two sections. The first section is arranged right after preprocessing task on user's request and corresponded sorting score. Requests using mentioned words are declared and new related words get extracted. Due to words frequency in other user requests each word assigned with a new score which is added to words specified score in order to maintain final rate. On the other hand, each word contains specific score which is gained from word frequency of the exact movie that can be used in order to get content extracted, dependent from user requests. The other section is extracting content employing user-visited

contents. In other words, if user had visited a content which other users visited too, then other users' relevant contents are presented as recommendations. In this study, MovieLens-20m dataset is used for evaluation procedure. This dataset contains information about movies, user tags and movie ratings. For evaluation, the results of first section and rating files are presented separately. For the other section evaluation, extracted movies obtained from second section are added to previous extracted films and similar trend of section 1 is considered for displaying results. Finally, we conclude that by employing relationship between user words and corresponded movies as well as user history in viewed films, strong applicable recommendations are offered.

10. References

- [1]. U. Leimstoll and H. Stormer, "Collaborative Recommender Systems for Online Shops," pp. 1–12, 2007.
- [2]. M. F. Iván Cantador Pablo Castells, "A Collaborative Recommendation Framework for Ontology Evaluation and Reuse," *Actas Int. Work. Recomm. Syst. 17th Eur. Conf. Artif. Intell. (ECAI 2006)*, pp. 67–71, 2006.
- [3]. J. Nam, "Understanding the motivation to use web portals," *Comput. Human Behav.*, vol. 36, pp. 263–273, 2014.
- [4]. A. M. Aladwani and P. C. Palvia, "Developing and validating an instrument for measuring user-perceived web quality," *Inf. Manag.*, vol. 39, no. 6, pp. 467–476, 2002.
- [5]. R. P. Bringula, "Influence of faculty- and web portal design-related factors on web portal usability: A hierarchical regression analysis," *Comput. Educ.*, vol. 68, pp. 187–198, 2013.
- [6]. B. Mobasher, "Data mining for web personalization," *Adapt. Web Lect. Notes Comput. Sci.*, vol. 4321, pp. 90 – 135, 2007.
- [7]. R. Burke, "Hybrid Recommender Systems: Survey and," *User Model. UserAdapted Interact.*, vol. 12, no. 4, pp. 331–370, 2002.
- [8]. E. P. Agustini, "Perancangan Ontologi Sebagai Meta Data Aplikasi Berbasis Web Semantik (Studi Kasus : Perpustakaan Universitas Bina Darma Palembang)," pp. 2–5, 2014.
- [9]. L. Sharma and A. K. Agrawal, "Study of Recommendation System for Web Portals," *Int. J. Comput. Appl.*, vol. 84, no. 9, 2013.
- [10]. M. De Gemmis, L. Iaquinta, P. Lops, C. Musto, F. Narducci, and G. Semeraro, "Preference learning in recommender systems," *Prefer. Learn.*, pp. 41–55, 2009.
- [11]. C. T. Liu, T. C. Du, and H. H. Tsai, "A study of the service quality of general portals," *Inf. Manag.*, vol. 46, no. 1, pp. 52–56, 2009.
- [12]. B. S. Mohamedaly, "Increasing Portal System Utilization among Patients of a Community Health Center," 2017
- [13]. M. Herrera, M. Á. Moraga, I. Caballero, and C. Calero, "Quality in Use Model for Web Portals (QiUWeP)," *Proc. 10th Int. Conf. Curr. trends web Eng.*, pp. 91–101, 2010.
- [14]. A. T. Siddiqui and M. Masud, "An E-learning System for Quality Education," vol. 9, no. 4, pp. 375–380, 2012.