Int. J. of Comp. & Info. Tech., (2017) 5(1): 33-42.

IJOCIT
ISBN: 2345-3877
www.ijocit.org
Volume 5, Issue 1

Review Paper

A Taxonomy on Products Recommendation Systems

Mohsen Dindar [1,*],
Shahram Jamali [2]

Received: 23 July 2016
Accepted: 18 Dec 2016

Abstract
With the promotion of commerce and communication networks, e-commerce and specially electronic stores have gained great popularity among users, but the information overload has become a major challenge that is users need to search and find information and services for use the features and applications of communication networks. At this time achieving to the required information through massive amounts of similar information would be a waste a lot of time of users and finally, users may fail to reach their desired information. The widespread electronic stores and diversity of products available in stores may also have the same problem. Therefore, A guidance and in fact a useful suggestion according to user's interests and criteria, keeps them out of confusion and wasting time. For this purpose, Product Recommender Systems (PRS) try to resolve this problem by providing recommendations to purchase products to customers. In this paper we will present taxonomy on product recommendation systems that is divided into three approaches: review-based, feature-based and hybrid PRSs.

Keywords: Product recommendation system (PRS), taxonomy, review-based, feature-based.

Citation: Dindar, M., Jamali, S., (2017). A Taxonomy on Products Recommendation Systems, *Int. J. of Comp. & Info. Tech. (IJOCIT)*, 5(1): 33-42.

1 Computer engineering Department, Germi Branch, Islamic Azad University, Germi, Iran
2 Department of Computer Engineering, University of Mohaghegh Ardabili, Ardabil, Iran
* Corresponding author: Vbcars@gmail.com, Tel: +98-9143555589

1. Introduction

E-commerce like Internet has grown immensely, and today this type of business uses more than other ones. Ease of purchase and availability of boarding and lack of transportation problems and other benefits cause this type of business has gained great popularity among users. In other hand, electronic store Due to the ease of purchase, payment easy to, independence of time (7/24) and find almost all the products needed is quickly spread. So, the rapid development in this field makes electronic store's users and customers have faced with the problem of information overload.

Finding online the desired items that match the user's criteria requires a lot of time and searching. In many cases, the users may not be able to find the products according to their interests. Therefore, a useful suggestion as guidance according to user's interests and criteria, keeps them out of confusion and wasting time. For this purpose, product recommender systems try to resolve this problem by providing recommendations to purchase products to customers [1].

Product recommendation system plays the guidance's rule by taking user's idea about products as input and searches among the available ones in the database in accordance with the user's criteria and finally suggests the list of products to the user that match with user's idea as the output. Users by this recommendation can find required products without wasting time and confusion and make more accurate decisions about purchase product. Also users can recognize products based on the recommendations that associated with their priorities, and decide to buy. Hence, the recommendation system can help to users to identify suitable products for their needs and preferences in an effective way and solve the problem of information overload in e-commerce and help to growth sales [2].

Extremely fast growth of the e-commerce platforms has made marketers to make online recommender to guide the customers in their purchase process and persuade them to make decisions. In the other hand, customers have demanded more personalized information delivery services. Hence, all PRSs try to provide more personalized services and many PRSs could found that work as online guide for customers. So, we can divide online PRSs into three approaches: review-based, feature-based and hybrid methods.

Review-based approach focuses on other customers' opinion about a special product. This approach recommends items to a consumer based on other customers' purchase decisions that have similar preferences. This approach is introduced as content based technique [4] that recommends items similar to those had been bought by other customers. These kinds of PRSs have low quality recommendations, which is due to three major issues. First, these cannot provide recommendations unless multiple-item purchasing profiles for a number of consumers, or at least for the consumer currently using the system, are available. Second, preference estimates based on purchasing profiles are inaccurate when, as is often the case, these profiles contain products purchased as gifts for or on behalf of other consumers. Third, purchasing profiles are historical data, revealing past but not necessarily current preferences.

Feature-based approach focuses on the features of products or customers. This approach gets customers' comments explicitly about products and predicts matched products with customers' criteria. These kinds of PRSs have high quality recommendations because in these approaches are done structured query and very similar products to the customers' requirements are found.

Hybrid methods combine advantages both of review-based and feature-based to develop more flexible PRS and provide more accurate recommendations to encourage customers to make purchase decision.

2. Review-based PRSs

Review-based PRS is widely used for recommendation systems design, which utilizes the content of items to create features and attributes to match user profiles. Items are compared with items previous liked by the users and the best matched items are then recommended. One major issue of CB filtering approach is that RS needs to learn user preferences for some types of items and apply these for other types of items.

Collaborative Filtering (CF) approach is the most popular approach for recommendation systems design. It utilizes a large amount of data collected from user behavior in the past and predicts which items users will like. It does not need to analyze the content of the items. Instead, it relies on the relationship between users and items, which are typically encoded in a rating feedback matrix with each element representing a specific user rating on a specific item [5].

As mentioned in Review-based recommender systems user's interests are expressed as ratings for items, and each additional rating extends the knowledge of the system and affects the system's recommendation accuracy [10]. For example In a PRS the cold start (CS) recommendation problem of providing prediction on the popularity of given CS items to the general users is investigated, and is presented a solution to predict the popularity of complete cold start (CCS) items and incomplete cold start (ICS) items. There are two main motivations for this work:

- CS items need to be recommended to achieve rank for motion recommendation and they should be carefully suggested to give users' better preferences with the recommendation systems. Else the CS items may go to an undesirable cycle of receiving no ratings.

- The estimated ratings for CS can help to make right decisions on product planning and sale strategies. The accuracy of such estimation is critically important for this type of purposes.

This work designs two integrated recommendation models, in which item features are learned from a deep learning architecture SDAE, using the descriptions of items retrieved online, then these features are exploited and integrated into the timeSVD++ CF model. timeSVD++ is one of the best performing CF models which tracks time changing behavior in the data and takes the temporal dynamics into account [5].

This work includes four phases as follow:

- It proposes a general framework of combination the CF approach and machine learning algorithms to improve recommendation performance for CS items. This framework uses deep learning neural networks as the key item factor vectors in the recommendation model for CCS items and approximated by the item factor vectors in the model for ICS items

- It integrates the CF approach and machine learning algorithms for CS items to extract the key content features
and embedding the item features into the CF
recommendation models.

- Based on the general framework specific system design and models are presented, in which the state of the art CF model, timeSVD++ and an advanced deep learning neural network model, SADE, are used for CS items recommendation.

- In addition to the design and evaluation of recommendation models for CCS items and ICS items separately, we also compared the performance of IRCD-CCS model and IRCDICS model on rating prediction for ICS items.

In practice, recommendation systems keep introducing new items into the systems over time. If a newly introduced item is a CCS item, the CF model can't provide rating prediction for it. If the item is an ICS item, the CF models may not give good recommendation. It may be beneficial to apply a CCS recommendation model for ICS item rating prediction. We propose a scheme of switching recommendation models for ICS items and retraining the models to deal with the practical issues of transition of item status from CS to non-CS. To the best of our knowledge, this practical issue has not been studied before in the literature

In other work a social recommendations system for e-commerce has combined similarity, trust and relationship is proposed that detects priority of the members through close friends and social network [4]. The basic idea of trust and reputation systems is to obtain a score for users. According to these results, other users can decide whether they are traded by a trusted user or not. In fact this mark establishes the reliance on the recommendations of any person that is obtained by close friends in social network [4].

Four analytic modules have been developed to analyze the obtained information from social network. The preference similarity analysis module measures the preference similarity between two customers based on the product rating records of each customer. A group of users with the same similarity level can be identified. The criteria of a goal customer towards a specific product can be predicted by a group of other customers with the same preference similarity. The preference similarity degree of two customers can be estimated according to their product purchases or rating records.

The recommendation trust analysis module computes the reputation quality (success rate) of the product recommendations of a customer according to his/her product rating records. The recommendation predication accuracy is positively associated with the recommendation trust (expertise reputation) of recommenders. The recommendation trust of a recommender is evaluated by his/her success rate of product recommendations.

The social relation analysis module analyzes the relation closeness degree between two customers according to implicit interaction records or explicit closeness ratings between them in a social network. The closeness value of a relation path between two users is measured by the weakest tie strength (closeness) at the edge of the path. When there are multiple relation paths between two users, the path with the strongest closeness value is used to represent the social relation strength between the two users.

The personalized product recommendation module computes the personalized factor weights for product evaluation and recommendation based on individual factor ratings with respect to different product categories. These criteria are significantly affected by the impact of personality traits, such as gender, age, and economic status. In order to achieve the personalized recommendation criteria, users are invited to evaluate the relative importance of preference similarity, recommendation trust, and relationship closeness [4].

In This recommendation system, the customer confidence is more considered and lead to rating of recommenders according to the success of the previous recommendations. The important thing is that if the accuracy of the recommendations is greater, its impact on the customer's personal decision to be higher. The strength of this system is emphasizing on the theoretical perspectives of the confidence to recommend. In fact, the customer can select recommender with high confidence and accuracy of recommendations.

The weakness of the social recommendation system is that it anticipates customer’s needs based on the comments of customer's friends and associates on social network. If does not available accurate knowledge of customers on social networks, identifying their needs will not be accurate.

Another weak point is lack of attention to product features. Recommended Product characteristics (include price, quality, brands on the market, year of manufacture, etc.) should be clear to the customer to choose the product carefully.

The new customer is another weakness of this recommendation system because gathering information about a new customer in social networks can be time consuming and difficult.

Another PRS has presented on personalized products recommendation based on user-contributed photos from social media sites [6]. The input of this approach is user shared photos of the same webpage and their corresponding textual descriptions. In this study, a new hierarchical user interests mining (Huim) approach for personalized PRS is proposed, which includes four stages:

1) Use of visual data and UGC for user interest mining.
2) Hierarchical user interest representation. User's information (UGC and enriched tags of the photos) to a hierarchical public topic space is mapped and represented user interest by a high dimensional topic vector and proposed to Huim promote discussions of interest.
3) Products representation. Products are mapped to the same public topic space to get their topic distribution vectors. Each product also corresponds to a point. The public topic space acts as a bridge between user and products.
4) Product ranking. In the public topic space, both user's interest and a product are represented by a high dimensional topic vector. Thus the relevance of user and product can be measured by the correlation of their topic vectors. In this paper used the cosine of them to measure their relevance and then determined the ranks of the products [6].

The advantage of this PRS uses the customer's picture at User's Profile. Although identification customer is more, it gives more accurate personalized recommendations to users. The disadvantage of this PRS is that if sufficient knowledge of user is unavailable or if the user visits from system for first time, this system won't provide accurate and effective advice to users. In fact, the history of customer buying behavior is the system requirements.

3. Feature-based PRS

Feature-based approach concentrates on the products or customers thorough extracting preferences about desired products. This approach provides more accurate advices but because of requirements extracting expression the domain of product is limited.

For example of this approach a PRS based on Association classification for personalization in B2C e-commerce applications is presented [8]. This system consists of four stages which are:

- requires the processing module
- Production Association classification Module
- Pruning classification module
- Validation System Performance Module

At the first, requirement data are collected and transformed into common phrase datasets. Then data mining procedure starts to search for a set of associated, frequently occurring phrase patterns (classifiers). In addition, stemming algorithm and stop word list in English are adopted to decrease the dimensions of the phrase dataset and improve the efficiency of the classifier extraction [8].

In this research, four phrase set collections are used to match the requirements. For every collection, there exist several sub collections, each involve a set of synonyms. A lot of semantic rules are represented as IF–THEN rule formats and stored in the semantic rule database to indicate the inference relations between requirements and a set of predefined phrases. After stop words removal and stemming, one customer requirement is transformed into a word set and the semantic meaning of such a requirement is represented as IF–THEN rule formats.

The next phase is classifying the mined association rules. Also to solve the recommendations problem needed to association rules by considering the class label data. So in this paper, two sets of items defined which one includes customer requirements and other includes items classes' label.

After generating classifiers using association rule learning, one important problem originates from the need for determining appropriate thresholds for the support and confidence levels. If the support and confidence thresholds are sat with low or high values, respectively, the number of the classifiers could be very huge or useful classifiers may be ignored. Besides, noisy and redundant information threaten the classification quality. The generated classifiers are pruned by which only those classifiers with good quality are kept for recommendations.

Finally, to evaluate how accurately the proposed recommendation system assigns class labels based on future customer requirements, this research applies the accuracy measurement to validate the system performance [8].

In this PRS, customer's criteria is collected in set of words and then Using data mining association rules, to find the relationship between these words and classify in the class with the specify label. The product related with this class is provided as a product recommendation to the customer. The strength of this system is collecting the customer's measures carefully and eliminating "new customer's" problem by collecting the customer's buying criteria through explicit way.

Existing noisy and insufficient information in the information presented by the customer can reduce the accuracy of system and threat performance of PRS.

Another PRS is presented [9] that introduced a prototype of e-commerce portal, called e-Zoco, of which main features are: (i) a catalogue service intended to arrange product categories hierarchically and describe them through sets of attributes, (ii) a product selection service able to deal with imprecise and vague search preferences which returns a set of results clustered in accordance with their potential relevance to the user, and (iii) a rule-based knowledge learning service to provide the users with knowledge about the existing relationships among the attributes that describe a given product category [9].

Portal embedded in this study compared the following components:

Products catalogues systems, customer management systems, messaging *systems, products evaluation systems, management issues system, lexicographic based search engine, sales management system, bid management system, secret data management system and reporting system.*

Today features that available in the e-commerce portals in following:

- Catalogue browsing
- Lexicographic search
- Advanced Search
- Mixed search
- Results arrangement

Product selection systems are distinguished from traditional search engines in that they provide more advanced capabilities, such as the interpretation of vague or imprecise search criteria or the results clustering and classification according to their relevance. Product selection systems can be broadly classified according to the kind of products for which recommendation is offered as:

- Product selection systems for low involvement products (LIP), such as books, music albums, or films.
- Product selection systems for high involvement products (HIP), such as appliances, video or photo cameras, musical instruments or vehicles.

In the case of LIP, the click-to-buy rates are usually higher compared to those of HIP. Consequently, recommendations for LIP are usually offered with the help of the customer's past purchases or past searches, demographic details, or explicitly specified interests. Collaborative Filtering (CF) is one of the most widely used techniques to offer recommendations for LIP products. CF techniques try to match the customer's tastes and preferences with that of all other customers to identify those like-minded and then offer the products bought by them as recommendations [9].

In the case of HIP, product selection systems are usually developed to take as input a set of product features or attributes to match against the set of products available in the database. As result, the system generates a ranking of products most likely to be of interest to the customer. Product selection in the case of high involvement products results more complex compared to low involvement products [9].

This system first is a hierarchical catalog contains general and specific classes. Each class is a combination of products that are hierarchically organized. Members of a specific class, inherit general class to properties**.** All classes in the catalog associated with the transaction. Any product that has sold at least is one member of classes of catalogue [9].

As mentioned**,** e-commerce customers want to have accurate and useful metrics that they just have not made aware of certain models and brands**,** but also they will explore models available on the market. Then, this system performs product recommendation in three phases:

- Identify the needs of users (search criteria).
- Search among the items in the catalogue and find the product that is the closest match.
- Organize and present the results to the customer [7].

The advantages of this system can be noted to differentiate products with high involvement (high risk selection) and low involvement products (choose low risk). The accuracy of the products can aware customers of product features and help to customer to select requirements carefully**.**

The disadvantages of this system is assigned it to the C2C stores (auction stores and auction sites) that customers are in contrast with each other and reliability and security of customers in these sites is challenging.

The other disadvantage of this system is high computational complexity for products with high involvement and exist probability of customers confusion during the long process of determine research criteria.

4. Hybrid PRS Method

Hybrid methods combine advantages both of review-based and feature-based to develop more flexible PRS and provide more accurate recommendations to encourage customers to make purchase decision.

For example of this approach a PRS is developed [7] that called HOPE, which integrates CF-based recommendation using implicit rating and SPA-based recommendation. This section presents the overview of the system, followed by the detailed description of each step of the framework.

Overall framework of HOPE system consists of two main processes: CF process and SPA process. The CF process is a way in which the customer experience and preferences are adapted with entire customer and an arbitrary subjective ID is assigned to each customer and purchase product is suggested to customer based on specified criteria of that ID. in this study implicit rating for each user about an item (The ratio of purchase desired item to the total number of items is purchased by the user) is obtained using purchase information. Thus, it calculates the similarity between a target user and other users using the implicit rating and selects the top k users based on the similarity score as neighbors of a target user. Finally, the predicted preferences of a target user on purchased items by the top k neighbors (CFPP) are calculated based on the ratings of the neighbors. The SPA process derives sequential patterns from transaction data of other users, and predicted preferences on items (SPAPP) are calculated by matching all subsequences of a target user's purchase sequence data with each derived sequential pattern.

In this method, transaction data sorted for each person by the date of purchase and then purchase sequential data is generated. Sequential data is series of items which they are arranged according to the time of purchasing; then the sequential patterns of users sequence data except the goal user can be achieved by using SPA. Actually, repeated items that appear in other users' purchase transactions causes sequential patterns are predictable. Finally, the weighted sum of normalized CFPP and SPAPP is calculated as a final predicted preference (FPP) on each candidate item to recommend, and then the top n items with the highest FPP are recommended [7].

Combination of cooperation filter and sequential patterns and choose customers who have a lot of near shopping behavior and investigation products Sequential patterns is the advantage of this PRS.

The problem of noise in the data that predict explicit rating collaborative filtering using the user's mental ID and determine customer needs from this ID can be noted as Disadvantages.

Other PRS proposes a novel hybrid recommendation method that combines the segmentation-based sequential rule method with the segmentation-based KNN-CF method [11]. The proposed method uses customers' RFM (Regency, Frequency, and Monetary) values to cluster customers into groups with similar RFM values. For each group of customers, sequential rules are extracted from the purchase sequences of that group to make recommendations. Meanwhile, the segmentation-based KNN-CF method provides recommendations based on the target customer's purchase data for the current period. Then, the results of the two methods are combined to make final recommendations.

To take advantage of the merits of the CF and sequential rule-based (SR) methods, this hybrid model combines the SSR method with the segmentation-based KNN-CF (SKCF) approach to improve the quality of recommendations.

The SSR method improves the quality of sequential rule-based recommendations by making recommendations based on customer groups. It uses customers' RFM values to cluster customers with similar values into groups. The SSR method then extracts sequential rules from each group of customers and provides recommendations based on the group that the target customer belongs to. The SKCF method is also used to provide recommendations based on customer groups. For a target customer u in a specific customer group G, the purchase data of the customers (including u) in G and in period T is used to derive the K-nearest neighbors of u and make recommendations. We combine the results of the two methods linearly to predict which products the target customer will buy in period T. In other words, to enhance the quality of recommendations, the hybrid method considers customers' purchase sequences in the periods prior to period T derived by the SSR method as well as customers' purchase data for period T derived by the SKCF method.

The rationale for the proposed hybrid approach is as follows. The SSR method does not utilize information about the target customer's purchases in period T, while the SKCF method does not consider the purchase sequences of customers over time. The SSR method assumes that if customers' purchases over time are similar prior to period T, then their purchases will also be similar in the current period T. Thus, the advantage of the SSR method is that it makes recommendations based on customers with similar purchase behavior over time. However, over time, some customers' purchases may be very different from those of other customers. In addition, customers with similar purchases prior to period T may not have similar purchases in period T. Thus, in such scenarios, the SSR method may not perform well. The SKCF method, which utilizes the target customer's purchases in period T to make recommendations, can resolve the drawbacks of the SSR method. Even so, for customers who make very few purchases in period T, the SKCF method does not perform well due to the sparsity of neighbors. In such a scenario, the SSR method can complement the SKCF method because it can utilize customers' purchase data prior to period T to find customers with similar purchase patterns. Basically, the hybrid approach seeks to combine the advantages of the two methods by utilizing customer purchases over time and the target customer's purchases in period T to improve the prediction power [11].

5. Comparison of PRSs

Generally, each PRS try to improve part of the buy recommendations process. This improvement is to increase the accuracy of personalization product recommendations and to help to customers to choose the desired product. Also PRSs want to solve challenges will be faced (which discussed earlier). Hence, for comparison about Efficiency of PRSs, we must evaluate these systems in terms of overcoming challenges.

In this paper, several recommendation systems investigated and the advantages and disadvantages of each mentioned**.** In this section, we get an aggregation of advantages and disadvantages of these PRSs in Table 1. As can be seen Table 1 shows the performances of PRSs to overcoming some of the challenges.

6. Conclusion

PRSs are one of a method that adoption to deal with the problem of overload information in e-commerce. The purpose of PRS is providing buy recommendations to help to electronic store's customers to avoid wasting time and confusion among many products available in store. Providing accurate recommendations will cause to select product carefully by the customer.

In this paper, a taxonomy of PRSs were presented that divided them to three major groups and for each group several recommendation systems investigated and the advantages and disadvantages of each mentioned**.** Can be said that each PRS try to address one or more aspects of current challenges in this area and any of this have been not overcome to the existing problems completely.

Table 2: Solve the challenges by PRSs

PRSs	Noisy data	Attention on the product features	System development	Customer needs Change with changing times	Recommender Ranking	New Customer	Customer Criteria Collection
Collaborative Filtering and Deep Learning Based Recommendation System For Cold Start Items	×	√	√	×	√	√	×
Social Recommender Mechanism	×	×	√	×	√	×	×
Associative classification-based recommendation system	×	√	√	√	×	√	√
Preference-based Clustering Reviews	×	×		×		√	×
Product Recommendation with temporal dynamics	×	√		√		×	×
Mining user-Contributed photos for personalized product recommendation (Huim)	×	√		×		×	×
Combining implicit rating-based collaborative filtering and sequential pattern analysis (HOPE)	×	√		×		×	×
highly adaptive recommender system for C2C	×	√		√		√	√
A hybrid of sequential rules and collaborative filtering for product recommendation	×	√		√		×	√

References

[1] Yung-Ming Li, Chun-Te Wu, Cheng-Yang Lai, (2013). "A social recommender mechanism for e-commerce: Combining similarity, trust, and relationship", *Elsevier Decision Support System*s, Vol 5, pp. 740–752.

[2] Yiyang Zhang, Jianxin (Roger) Jiao, (2007). "An associative classification-based recommendation system for personalization in B2C e-commerce applications", *Elsevier Expert Systems with Applications*, Vol 33, pp. 357–367.

[3] Li Chen, Feng Wan, (2013), "Preference-based clustering reviews for augmenting e-commerce recommendation", *Elsevier Knowledge-Based Systems*, Vol 50, pp. 44–59.

[4] Wenxing Honga, Lei Li, Tao Li, (2012). "Product recommendation with temporal dynamics", *Elsevier Expert Systems with Applications*, Vol 39, pp. 12398–12406.

[5] Wei, J., He, J., Chen, K., Zhou, Y., & Tang, Z. (2017). Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Systems with Applications*, 69, 29-39.

[6] He Feng, Xueming Qian, (2014). "Mining user-contributed photos for personalized product recommendation", *Elsevier Neurocomputing*, Vol 129, pp. 409-420.

[7] Keunho Choi, Donghee Yoo, Gunwoo Kim, Yongmoo Suh, (2010). "A hybrid online-product recommendation system: Combining implicit rating-based collaborative filtering and sequential pattern analysis", *Elsevier Electronic Commerce Research and Applications*, Vol 11, pp. 309–317

[8] Jose Jesus Castro- Schez, Raul Miguel, David Vallejo, Lorenzo Manuel López – López, (2011). "A highly adaptive recommender system based on fuzzy logic for B2C e-commerce portals", *Elsevier Expert Systems with Applications*, Vol 38, pp. 2441–2454.

[9] Mehdi Elah, Francesco Ricci, Neil Rubens, (2016). "A survey of active learning in collaborative filtering recommender systems", *Computer Science Review*, Vol.20, pp.29–50.

[10] Duen-Ren Liu, Chin-Hui Lai, Wang-Jung Lee, (2009). "A hybrid of sequential rules and collaborative filtering for product recommendation", Information Sciences, vol.179, pp.3505–3519.