International Journal of Soft Computing 2 (5): 588-591, 2007
ISSN: 1816-9503

Fuzzy Classification for Recommender Systems

[1]K. Satyanarayana and [2]S.P. Rajagopalan
[1]Department of Computer Science, T.S. Narayanaswami College of Arts and Science, Navalur, Chennai 603 103, Tamil Nadu, India
[2]M.G.R. University, Madurayoval, Chennai, Tamil Nadu, India

Abstract: Recommender systems are more and more needed in order to exploit the huge amount of information available on the Internet. Recommeder system makes automatic prediction with regard to the interest expressed by the user. All shop owners do not give same importance to all products since these products do not possess the same value. Any shop owner will prefer products, which will give him/her higher profits and carries high product value. In this study, we have discussed about fuzzy classification of product values and use of linguistic variables, method of associating fuzzy classification with recommender system.

Key words: Recommender systems, collaborative filtering, content-based filtering, linguistic variable, fuzzy logic, standard deviation

INTRODUCTION

In order to draw user's attention and to increase their satisfaction towards online information vendors try to predict user preference based on their behavior. Recommender systems are implemented in commercial web sites to predict user preferences and accurate prediction, results in higher selling rates. These systems allow users to locate the preferable items quickly and avoid the possible information overloads. We know that a number of Recommender systems are used in online shop systems and in the applications of e-commerce these systems plays a major role. Recommender systems are described as the systems that acquire opinion about items from a community of users and that use those opinions to direct with other user within that community to those items which are interested for them (Resnick and Varian 1997).

A typical shop offers a large number of different products. We can assure that it is not necessary that all products are equally valuable as far as shop owner is concerned. We can consider various attributes that influence the value of a product from shop owner's perspective. Profit margin, availability of product, turnover, after sales service, quality and supply are some factors that influence the value of the product. We may use the techniques available in fuzzy logic to analyze these variables. Fuzzy logic was mainly motivated by observing human reasoning, can utilize concepts and knowledge that do not have well defined sharp boundaries, that is, when concepts are vague. Fuzzy logic enables us to have descriptive and qualitative form for vague concepts because they are often described qualitatively by words and hence we will classify the products using this logical system. The advantage of classifying elements using fuzzy classification is that we can assign more number of classes instead of a single class.

Here in this study, we have adopted the classification based on fuzzy logic on products. This classification will influence the output of a recommender system, which will be most valuable from shop owner's perspective. Here we will introduce the fuzzy classification concept applied to products, then we will discuss the recommender systems which will accept this classification.

Products-classification: Customers may get large number of different products from shops. Some attributes like profit margin, quality or turnover may be used to classify the products and we can also use these attributes to define the product range. For customers, it has been shown that a fuzzy classification can be used to calculate an individual customer value (Meier *et al*., 2005).

While estimating the product value we need to consider the factors like current demand, expected margin, quality, stock on hand, future demand and etc. Any shop owner will be interested in strengthening the relationship with high value customers and this can be achieved by assigning personalized discounts or by giving special conditions (Werro *et al*., 2005). Let us now discuss the different attributes for classification.

Corresponding Author: K. Satyanarayana, Plot No: 6, New No: 7, Old No: 9, LIC Colony, I street, Velachery, Chennai-600 042, India

The profit margin is the difference between purchase price and selling price and this value differs from product to product. Selling the products with high profit margin will be the main attraction for any shop owner. Quality is another attribute, which will influence the customer and shop owner. High quality products naturally avoid the service costs and hence shop owners will tend to sell high quality products when compared low quality products. Size could be another important attribute because large products occupy a greater amount of the inventory space and special efforts are needed for packing and logistics. Another attribute, which influences the shop owner, is that the charges offered by the supplies for recommending their products. Similarly availability of product and supplier profile is other important factor in classification and these attributes may vary from shop to shop. Let us now discuss about the linguistic variables and fuzzy sets.

LINGUISTIC VARIABLE AND FUZZY SETS

A set in classical theory always has a sharp boundary because membership in a set for an object either completely belongs to the set or does not belong to the set at all. A Fuzzy set is set with smooth boundary. Fuzzy set theory directly addresses this limitation by allowing membership in a set to be a matter of degree. The degree of membership in a set is expressed by a number between 0 and 1; 0 means entirely not in the set, 1 means completely in the set and a number in between means partially in the set. A fuzzy set is thus defined by a function that maps objects in a domain of concern to their membership value in the set and such function is called membership function, denoted by the symbol μ (John Yen, 2003). The membership function of a fuzzy set A is defined as μ_A and the membership value of x is denoted by $\mu_A(x)$.

Like a conventional set, a fuzzy set can be used to describe the value of a variable. The sentence "The rain is heavy" uses a fuzzy set "Heavy" to describe the quantity of the rain. A linguistic variable enables its value to be described both qualitatively by a linguistic term and quantitatively by a corresponding membership function. The attributes that we have considered for product classification belong to fuzzy sets and linguistic variables, hence fuzzy classification is possible (Chen, 1998).

Suppose the percentage of profit margin in a particular shop ranges from 0 to 40. We have to divide this domain in some equal classes say [0,10], [10,20], [20,30] and [30,40]. With the help of linguistic variables the equivalence classes of the attributes can be defined as shown in the Fig. 1. Here we call the first class as low profit class and final class as high profit class.

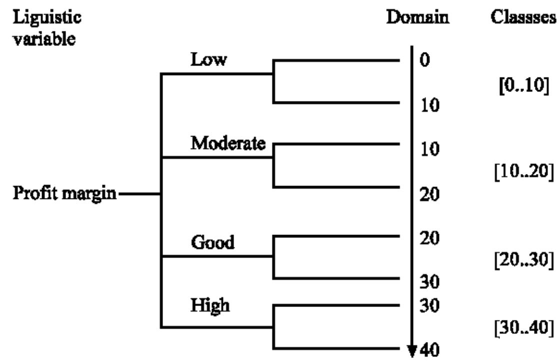


Fig. 1: Linguistic variable-profit margin

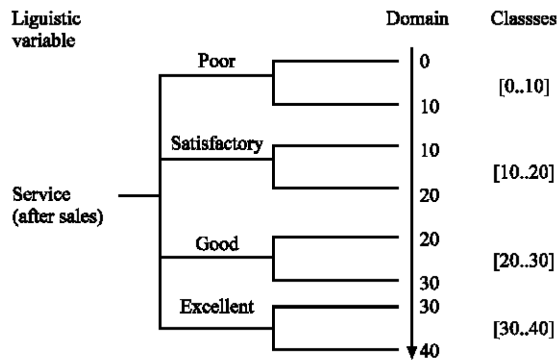


Fig. 2: Linguistic variable-service

Table 1: Classes and their membership degrees

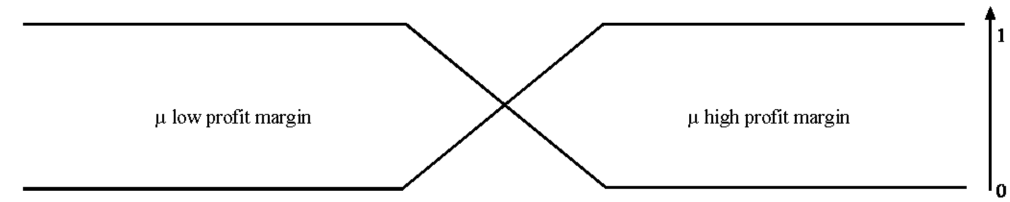
Class	1	2	3	4	5	6	7	8
Membership	1	0.9375	0.875	0.8125	0.75	0.6875	0.625	0.5625
Class	9	10	11	12	13	14	15	16
Membership	0.5	0.4375	0.375	0.3125	0.25	0.1875	0.125	0.0625

Table 2: Percentages that a particular product belongs to various classes

Class	1	2	3	4	5	6	7	8
%	15	14	10	9	8	7.5	7	6
Class	9	10	11	12	13	14	15	16
%	5	4	3.5	3	2.5	2	1.5	1

To derive fuzzy classes from sharp contexts, the qualifying attributes are considered as linguistic variables and verbal terms are assigned to each equivalence class. Similarly we can create equivalence classes for the attribute quality and service. Classes pertaining to the attribute service is explained in Fig. 2. Here each fuzzy set is determined by a membership function μ over the domain of the attribute.

Here $\mu_{\text{satisfactory service}}$, $\mu_{\text{Excellent service}}$ are non numeric attributes and $\mu_{\text{high profit}}$, $\mu_{\text{good profit}}$ are numeric attributes. Attributes profit margin and service determines a two-dimensional classification space as shown in the Fig. 3. The classes are characterized from "excellent product" to "non-ideal product".



← 0		0 10	11 20	21 30	31 40
μ excellent service	Excellent	Less profitable but highly appreciated product Class 6	Profit moderate but highly appreciated product Class 5	Good profit and highly appreciated product Class 3	Excellent product Class 1
	Good	Less profitable but appreciated product Class 8	Profit moderate but appreciated product Class 7	Good profit and appreciated product Class 4	Excellent profit and appreciated product Class 2
μ poor service	Satisfactory	Less profitable product Class 14	Product with moderate product Class 13	Product with good product Class 11	Product with excellent profit Class 9
	Poor	Non profitable and non ideal product Class 16	Moderate profit but non ideal product Class 15	Good profit but not recommended product Class 12	Product with excellent profit but danger to deal with product Class 10

Fig. 3: Two-dimension fuzzy classification space defined by profit and service

The product value can be derived from the fuzzy classification. Membership degrees of the products in the different classes determine the value that the product deserves. Each fuzzy class is associated with some values between 0 and 1 as shown in Table 1. For example

Suppose if a product belongs to all classes at the same time in a fuzzy classification and let us consider the following percentages for various classes as shown in Table 2.
The value of that product is calculated as

(Membership degree * Class percentage / 100) =
* 15/100 + 0.9375 * 14/100 + ……….+ 0.0625 * 1/100)
= 0.601935

RECOMMENDER SYSTEMS-COMMON APPROACHES

Approaches for building a user profile and computing recommendations are Collaborative Filtering (CF), Content Based Filtering (CBF) and Hybrid systems.

- CF systems find similar users to a target user by comparing users' opinions of items. Many CF systems compute similarity between users, by comparing vectors rating using Karl Pearson Coefficient correlation, cosine similarity and other similar techniques. In general CF model facilitates the users to provide ratings for the items they have experience before (Breese *et al.*, 1998). Then the user for whom the recommendations are computed is matched with other users in the system. Finally, predictions for the items that the active user has not yet rated, but the neighbors have rated are computed and these items are presented to the user for decision making.
- Content based information filtering selects the right information for users by comparing representation and searching information to representations, of contents, of user profiles which express interests of users (Baudisch, 1999). In this technique first we gather content data about the items. Secondly we ask the user to provide some rating based on some scale. Next we compile a profile of the user using the content information extracted from the user and rating information provided by the user. Finally we rank the items according to their scores and present them to user in order.
- Another class of recommender systems uses a hybrid approach, which is a combination of the content based and the collaborative filtering (Burke, 2002).

ASSOCIATING PRODUCT CLASSIFICATION WITH RECOMMENDER SYSTEMS

Let us now discuss possible techniques to connect the recommender systems with available Product classification. We can rank the products based on the product values calculated by using the technique discussed above or we may integrate the value of a product directly in a recommnender system.

Let us consider seven different products for the same type of item (say TVs or DVD Players produced by seven different manufacturers). Let us call these products as P1, P2,….,P7. Let us assume the product values for these products (P_v) as 0.3, 0.7, 0.8, 0.4, 0.2, 0.5 and 0.9.

Usually recommender systems list top ranked products in non-increasing order and these values will be adopted by the shop owner for his/her business planning. Let us consider that shop owner want to know the top four products. The following gives an idea of how recommender system associated with fuzzy classification reorders the products for shop owner's use.

Table 3: Ranks given customers (using some technique)

		Products						
		P1	P2	P3	P4	P5	P6	P7
CU	C1	1	0.7	0.6	0.4	0.5	-0.3	-0.1
ST	C2	0.9	0.8	0.5	0.3	-0.1	0.4	0.1
OM	C3	0.8	0.7	0.4	0.5	0.6	0.3	0.2
ER	C4	0.8	0.7	0.5	0.6	-0.1	-0.2	0.1

Table 4: Ranks after fuzzy classification technique

		Products						
		P1	P2	P3	P4	P5	P6	P7
CU	C1	0.9268	0.5292	0.4048	0.3024	0.4512	-0.422	-0.3196
ST	C2	0.8268	0.6292	0.3048	0.2024	0.1488	0.278	-0.1196
OM	C3	0.7268	0.5292	0.2048	0.4024	0.5512	0.178	-0.0196
ER	C4	0.7268	0.5292	0.3048	0.5024	0.1488	0.322	-0.1196

Non-increasing order of products (based on some technique):

$$\underbrace{\text{P2, P4, P3, P6}}_{\text{Top four}}\text{, P1, P7, P5}$$

Non-increasing order of products:

$$\underbrace{\text{P7, P3, P2, P6}}_{\text{Top four}}\text{, P4, P1, P5}$$

(Reordering based on product value)

In the other technique of combining the product classification with recommender systems, we use the two dimensional matrix to represent the relationship between users and products (Table 3). Here the value in each cell is based on implicit or explicit data supplied by the customer and these values will range between -1 and +1. -1 is strong dislike and +1 strong liking.

For simplicity product values discussed in the previous technique is considered. First we will calculate the standard deviation (σ) of product values. Next we need to find the constants for various products with the formula (product values * standard deviation). Then we have subtract these constants from respective columns.

That is Standard deviation (σ) for the product values(0.3, 0.7, 0.8, 0.4, 0.2, 0.5, 0.9) =0.244

$$\text{Constant for product 1} = P_{v1} * \sigma$$
$$\text{Constant for product 2} = P_{v2} * \sigma$$

These constants are subtracted from respective cells and new ranks are be found (Table 4).

CONCLUSION

In this study, we have discussed how fuzzy classification could be used to calculate a value for each product. For a shop owner, the product values have similar meaning like customer values. We have also discussed two methodologies for combining the fuzzy logic concepts with recommender system concepts. We have also discussed how linguistic variables are used in recommender systems. Problematic products may be easily identified. This product value is a good indicator when negotiating with suppliers. It will be possible for shop owner to offer higher discounts when he is able to identify the products with high product values (Blattberg *et al*., 2001).

REFERENCES

Baudisch, P., 1999. Joining collaborative and content-based filtering. In interacting with recommender systems. Online Proceedings of the Chi workshop.

Blattberg, R.C., G. Getz and J.S. Thomas, 2001. Customer Equity-Building and Managing Relationships as Valuable Assets. Harvard Business School Press, Boston.

Breese, J.S., D. Heckerman and C. Kadie, 1998. Empirical analysis of predictive algorithms for collaborative filtering. in Fourteenth Conference on Uncertainty in Artificial Intelligence, (Madison, W.I.), Morgan Kaufmann, pp: 43-52.

Burke Robin, 2002. Hybrid Recommender Systems, survey and Experiments, User Modeling and User Adapted interaction, Vol. 12.

Chen Guoging, 1998. Fuzzy Logic in Data Modeling-Semantics, Constraints and Database Design, Kluwer Academic Publishers, London.

John Yen and Reza Langari, 2003. Fuzzy Logic Intelligence, Control and Information, Pearson Education, Singapore.

Meier Andreas, Werro Nicolas, Albrecht Martin and Sarakinos Miltiadis, 2005. Using a Fuzzy Classification Query Language for Customer Relationship Management, in Proceedings of the 31st VLDB Conference.

Resnick Paul and R. Varian Hal, 1997. Recommender Systems, Commun. ACM., 40: 56-58.

Werro Nicolas, Stormer Henrik and Meier Andreas, 2005. Personalized Discount-A Fuzzy Logic Approach, In: Proceedings of the 5th IFIP International Conference on eBusiness, eCommerce and eGovernment, I3E 2005 Conference.