

# A New Aware-Context Collaborative Filtering Approach by Applying Multivariate Logistic Regression Model into General User Pattern

**Loc Nguyen**

Sunflower Soft Company, Ho Chi Minh City, Vietnam  
Email: [ng\\_phloc@yahoo.com](mailto:ng_phloc@yahoo.com)

Received 12 September 2016; accepted 15 August 2016; published 18 August 2016

Copyright © 2016 by author and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

Traditional collaborative filtering (CF) does not take into account contextual factors such as time, place, companion, environment, etc. which are useful information around users or relevant to recommender application. So, recent aware-context CF takes advantages of such information in order to improve the quality of recommendation. There are three main aware-context approaches: contextual pre-filtering, contextual post-filtering and contextual modeling. Each approach has individual strong points and drawbacks but there is a requirement of steady and fast inference model which supports the aware-context recommendation process. This paper proposes a new approach which discovers multivariate logistic regression model by mining both traditional rating data and contextual data. Logistic model is optimal inference model in response to the binary question “whether or not a user prefers a list of recommendations with regard to contextual condition”. Consequently, such regression model is used as a filter to remove irrelevant items from recommendations. The final list is the best recommendations to be given to users under contextual information. Moreover the searching items space of logistic model is reduced to smaller set of items so-called general user pattern (GUP). GUP supports logistic model to be faster in real-time response.

## Keywords

Aware-Context Collaborative Filtering, Logistic Regression Model

---

## 1. Introduction

Recent researches on collaborative filtering (CF) focus on inherent information about users and items and how to recommend such relevant items to such users. Database used to build up CF algorithms is in form of rating matrix composed of ratings that users give to items. Additional contextual factors such as time, place, condition and situation existing in real world are not considered in CF algorithms. For instance, if a user prefers to watch news program in morning and movies in evening then contextual information, namely temporary information, should be aware in recommendation tasks and so, it is inappropriate to provide movies to her/him in the morning even though such movies are the most relevant to her/him.

Given a training set which is a rating matrix and a user who requires recommendations, CF algorithm tries to predict rating value on items which are not rated by this user. After that CF algorithm makes a list of such items arranged in order of predictive rating values and recommends this user such list. In other words, CF algorithm constructs a predictive function  $R2$  ([1], pp. 217-250) whose cross domain is the Cartesian product of a set of users  $U$  and a set of items  $I$ . The domain  $U \times I$  is also called rating matrix. Co-domain of function  $R2$  is a set of predictive ratings denoted  $R$ .

$$R2_{U \times I} : U \times I \rightarrow R$$

Function  $R2$  called traditional 2-dimension (2D) mapping doesn't consider contextual factors now and it can lack information necessary to highly accurate prediction. Suppose contextual information including location, time and companion are added to prediction process, the 2D function  $R2$  becomes the 3-dimension (3D) mapping denoted as below:

$$R3_{U \times I \times C} : U \times I \times C \rightarrow R$$

Note that  $C$ ,  $U \times I$  and  $U \times I \times C$  represent context domain, 2D (cross) domain and 3D (cross) domain, respectively. In other words, function  $R3$  gives recommendations to user under circumstances specified as contextual information.

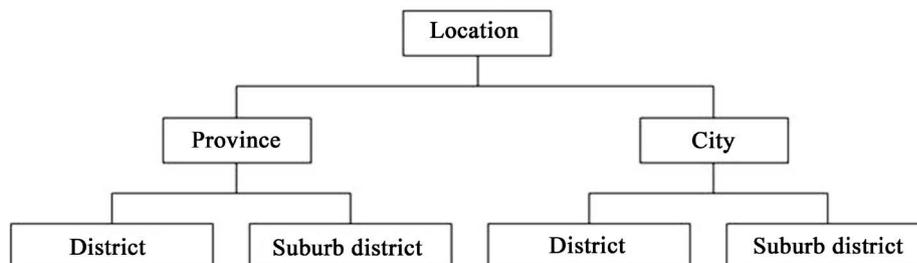
Although context has many different types, we can reduce these types into three main types in order to answer three question forms: when, where and who ([1], pp. 224-225).

- Time type indicates the time when user requires recommendation, for example: date, day of week, month, and year.
- Location type indicates the place where user requires recommendation, for example: theater, coffee house.
- Companion type indicates the persons with whom user goes or stays when recommendation task is required, such as: alone, friends, girlfriend/boyfriend, family, co-workers.

Contextual information is organized in two forms: hierarchical structure ([2], p. 1537) and multi-dimensional data (MD) model.

According to hierarchical form, the context domain  $C$  is defined as a set of contextual dimension  $\mathbf{K} = (K^1, K^2, K^3, \dots, K^n)$ .  $\mathbf{K}$  is represented in hierarchy structure, whose attributes  $K^i(s)$  are associated with the ascending order of fine level. For example, given attributes  $K^i$  and  $K^j$  where  $i < j$ ,  $K^j$  is finer than  $K^i$  and so  $K^i$  contains  $K^j$ . It is easy to recognize that  $K^1$  is the coarsest attribute which contains all remaining attributes  $K^2, K^3, \dots, K^n$ . Each  $K^i$  contains values at the same level  $i$  and it can be split into finer levels. An example of contextual dimension is shown in **Figure 1**.

In above example,  $K^2 = \{\text{City, Province}\}$ ,  $K^3 = (\text{City} \rightarrow \text{District}, \text{City} \rightarrow \text{Suburb district}, \text{Province} \rightarrow \text{District}, \text{Province} \rightarrow \text{Suburb district})$ .



**Figure 1.** Contextual dimension example.

According to MD form, context domain  $C$  is defined as the Cartesian product of  $n$  dimension,  $C = D_1 \times D_2 \times \dots \times D_n$ . Each dimension  $D_i$ , in turn, is a set of attributes,  $D_i = (a_{i1}, a_{i2}, \dots, a_{ik})$ . For example, suppose  $C$  has only one dimension of time denoted  $D_1 = Time$  (day of week). So, the cross domain  $U \times I \times C$  of predictive function  $R3$  constitutes a 3D cube: *User* (name), *Item* (book name) and *Time* (day of week). Each block in this cube is assigned a rating which is the predictive outcome of function  $R3$ . **Figure 2** depicts a MD cube ([1], p. 227).

**Figure 2** indicates that given user “John” (101), item “novel book” (7) and time “Sunday” (1), the predictive outcome of  $R3$  gains the highest value 5.

There are three approaches ([1], pp. 232-233) to apply context into recommendation process:

- Contextual pre-filtering: Firstly, given context  $c \in C$  is used to select user-item pairs  $(u, i)$  which are more relevant to this context, leading to obtain the aware-context cross domain  $U \times I$ . After that traditional 2D function  $R2$  is taken on such cross domain.
- Contextual post-filtering: Firstly, traditional 2D function  $R2$  is used to produce the list of recommended item. After that context  $C$  is used to fine-tune this list in order to remove irrelevant items according to concrete context.
- Contextual modeling: The 3D function  $R3$  is used directly on context-aware cross domain  $U \times I \times C$ .

The basic idea of contextual pre-filtering is to project the 3D domain  $U \times I \times C$  on 2D plane, based a concrete context  $c \in C$  so that such 3D domain is reduced to 2D domain  $U \times I$ . Let  $\prod_c$  be projection operation based on condition context  $c$ , we have:

$$\prod_c(U \times I \times C) = U \times I$$

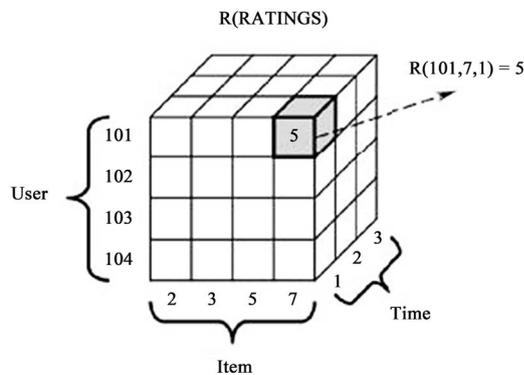
$$R3_{U \times I \times C}(u, i, c) = R2_{\prod_c(U \times I \times C)}(u, i)$$

The concrete context  $c \in C$  which is strict projection condition can make the cross domain  $U \times I$  small or sparse, causing low predictive accuracy. So generalization technique is used to make projection condition loose, namely the exact condition  $c$  is replaced by the more general condition  $c'$ . For example, the context  $c = \text{“Saturday”}$  which indicates that “Mary prefers to go shopping on Saturday” is replaced by more flexible context  $c = \text{“Weekend”}$  because she can like going shopping on Sunday if she often goes on Saturday. The general context not only expands the space of potential recommendation solutions but also improve predictive accuracy.

The essence of contextual post-filtering is to fine-tune the raw recommendation results taken from predictive function  $R2$  which didn't consider contextual factors before. Consequently, this method tries to figure out user's context-aware interests, preferences or attributes by using some artificial intelligent and mining techniques and apply such attributes into raw results so as to remove out irrelevant items or change their ranks in final recommendation list. For example, given context  $c = \text{“evening”}$  which indicates Mary's interest “watching movies in the evening”, contextual post-filtering approach will remove all of news or sport events from her recommendation list.

The strong point of contextual pre-filtering and post-filtering approaches is the ability to take advantage of legacy recommendation algorithms not taking account into contextual factors.

The essence of contextual modeling is to incorporate directly contextual information into predictive function  $R3$  and so,  $R3$  is constructed as inference model such as data mining, machine learning, heuristic model or statistical model, etc. The 2D predictive function isn't used in contextual modeling. It implies that the strong



**Figure 2.** MD cube ([1], p. 227).

point of this method is pure and powerful inference mechanism. It opens a new trend in context-aware recommendation research, giving a lot of prospects although a few related techniques are extensions of 2D algorithms.

The approach in this paper is the hybrid of contextual post-filtering and contextual modeling. Hence a logistic inference model ([3], pp. 372-411) is used as the post filter to make recommendations more relevant to users according to contextual factors. Section 2, which is the main section, describes the proposed approach in detail. The concept of general user pattern (*GUP*) is firstly introduced and then the collaborative filtering algorithm based on *GUP* and logistic model is mentioned. The equation to solve logistic model, which is the most important feature of the proposed algorithm, is constructed with support of mathematical tools. An example is given at the end of section 2 for illustrating the proposed approach. Section 3 is the conclusion.

## 2. Basic Idea and Details of the New Approach

The new approach is suggested by two comments:

- Although contextual information is necessary to improve the quality of recommendation process but it cannot replace essential rating information in collaborative filtering research. Inference model taking into account contextual factor should be used as the filter to adjust recommendations returned from predictive ratings so as to give users more appropriate items in concrete circumstances.
- When additional contextual information is considered, the speed of recommendation process is decreased. So inference mechanism should be fast in real-time response.

The basic idea is to apply a fast inference model, namely logistic regression function into a list of recommended items so as to achieve a better recommendation result under contextual information. The logistic regression model responses immediately the binary request “whether or not a list of items is relevant to concrete context or preferred by users”. Because there are various items and each item is associated with an individual regression function, the domain of regression function becomes huge, which decreases the speed of algorithm. In order to solve this problem, the items space is reduced to “general user pattern”.

General user pattern (*GUP*) is known as a set of items to which “many” users give ratings. Since rating values on *GUP* are not always high, it reflects solely user’s access or rating frequency. Given the threshold  $\theta$ , let  $n$  and  $n_j$  be the number of total users and the number of users who rate on item  $x_j$ , we have:

$$GUP = \{x_j \mid n_j \mid n\} \text{ where } x_j I$$

So *GUP* is defined as a set of items where the ratio of the number of total users to the number of users who rate on such each item is great than or equal to a given threshold  $\theta$ .

Given a *GUP*, context  $c$  and a logistic regression function  $f$ , the regressive (or independent) variables of  $f$  are taken from *GUP*. The response of  $f$  is binary variable taking values between 0 and 1 where value 1 indicates that it is likely that user prefers *GUP* under context  $c$  and otherwise. Suppose  $GUP = \{x_1, x_2, \dots, x_n\}$ , logistic regression function  $f(x_1, x_2, \dots, x_n)$  can be considered to be dependent on *GUP*.

Given the raw recommended list  $R$ , an instance of *GUP* is initialized on  $R$ . This instance denoted *INS* is a set of predictive rating values taken from  $R$  with condition that respective items co-exist in both  $R$  and *GUP*. For example, if  $R = \{x_1 = 5, x_2 = 3, x_3 = 4\}$  and  $GUP = \{x_1, x_3\}$  then  $INS = \{x_1 = 5, x_3 = 4\}$  because item  $x_1$  and item  $x_3$  exist in both  $R$  and *GUP* and their respective values are 5 and 4.

Consequently, regression model  $f$  is evaluated on *INS* with regard to context  $c \in C$ . If the outcome of  $f$  with *INS* is near to 0 then all items in *GUP* are removed from  $R$ . Finally, the list  $R$  is recommended to user after it was fine-tuned by pruning irrelevant items from it.

In general the algorithm has four steps:

1) A 2D predictive function is applied into rating matrix  $U \times I$  so as to produce a raw list  $R$  of recommended items without existence of contextual factors.

2) *GUP* is discovered over contextual 3D cross domain  $U \times I \times C$ . Items in *GUP* are frequent items.

3) Multivariate logistic function  $f$  is learned from cross domain  $U \times I \times C$  by statistical technique.

4) Function  $f$  is used to remove irrelevant items from the list  $R$ . In other words, only aware-context items are kept in  $R$ . So the final filtered list is the best result which is recommended to users. This step includes two sub-steps:

a) The instance of *GUP* so-called *INS* is constructed by matching *GUP* with  $R$ .

b) Function  $f$  is evaluated on *INS* to perform the removal of redundant items from  $R$ .

Because step 3 is the most important, the method to construct multivariate logistic function is discussed in detailed now. The probability that  $GUP = \{x_1, x_2, \dots, x_n\}$  belongs to context  $c \in C$  is  $p$ ; in other words, the probability that user prefers  $GUP$  under context  $c$  is  $p$ . The concept *odd* ([3], p. 11.8) is defined as the ratio of  $p$  to  $1 - p$ . This ratio represents how much user likes  $GUP$  vice versa how much user dislikes item  $GUP$ .

$$odd = \frac{p}{1-p}$$

Note that the *odd* also expresses how likely  $GUP$  belongs to context  $c \in C$  vice versa how likely  $GUP$  does not belongs to context  $c$ . The natural logarithm of *odd* is linear regression function of  $n$  variables being  $GUP$ .

$$\log \log(odd) = \log \log\left(\frac{p}{1-p}\right) = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n \quad (1)$$

Note that  $x_1, x_2, \dots,$  and  $x_n$  are regressive or independent variables whose values are obtained from recommended list  $R$  later and coefficients  $\alpha_i(s)$  are called parameters of logistic model. If *odd* is considered as response or dependent variable, Equation (1) is re-written as following:

$$odd = \frac{p}{1-p} = \exp(\alpha_0 + \alpha_1 i_1 + \alpha_2 x_2 + \dots + \alpha_n x_n) \quad (2)$$

where  $\exp(\cdot)$  denotes exponent function.

The probability  $p$  that user likes items in  $GUP$  is computed according to following function derived from Equation (2) with attention that such probability is logistic regression function  $f(x_1, x_2, \dots, x_n)$ .

$$f(i_1, i_2, \dots, i_n) = p = \frac{1}{1 + \exp(-(\alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n))} \quad (3)$$

Equation (3) represents the multivariate logistic model  $f$  with regard to a concrete context  $c$ . The approach in this paper uses this equation to estimate whether or not a user prefers a list of recommended items based on a concrete context. For example, given user being “John”,  $GUP$  being {“Gladiator”, “Golden Eye”}, recommended movie list being {“Gladiator” with predictive rating 5, “Golden Eye” with predictive rating 4, “Four Rooms” with predictive rating 4} and, if  $f$  produces a value greater than or equal to 0.5 when  $f$  is evaluated on  $GUP$  with regard to time context “evening”, then it asserts that John likes such list and there is no film to be removed. This logistic model is built up in offline mode so as not to affect response time.

The problem needs solved now is to determine parameters  $\alpha_i(s)$  in logistic function  $f$ . We will use method of maximum likelihood estimate (MLE) [4] to construct them. Given training data  $D$  is rating cube whose dimensions are user, item and context. Each volume in rating cube is quantified by a value that a user rate on an item in concrete context. If rating cube is projected onto a context  $c$ , we get a rating matrix  $D_c$  with two dimensions such as user and item. Suppose  $D_c$  has  $m$  rows and  $n$  columns and let  $y_i = (y_{i1}, y_{i2}, \dots, y_{in})$  be the  $i^{\text{th}}$  row of  $D_c$  and it is easy to infer that  $y_{ij}$  is the  $i^{\text{th}}$  instance (rating value) of item  $x_j$ . Suppose rating values  $y_{ij}(s)$  range from 1 to  $v$  where  $v$ -most favorite and 1-most dislike and the value 0 indicates that user does not rate on item. Suppose  $GUP$  also has  $n$  items which are the same to ones in  $D_c$ , let  $a_i = (a_{i1}, a_{i2}, \dots, a_{in})$  be a possible instance of such  $n$  items. Hence, there are  $(v+1)^n$  such instances because  $a_{ij}$  ranges from 0 to  $v$ . For example, if  $n = 2$  and  $v = 2$  we have 9 instances (0, 0), (0, 1), (0, 2), (1, 0), (1, 1), (1, 2), (2, 0), (2, 1), (2, 2). Let  $z_i$  be the number of  $y_i$  in  $D_c$  so that  $y_i = a_i$  and so there are  $N = (v+1)^n$  such  $z_i(s)$ . Let  $p_i$  be an instance of logistic function  $f$  evaluated on  $x_1 = y_{i1}, x_2 = y_{i2}, \dots, x_n = y_{in}$ , we have:

$$p_i = \frac{1}{1 + \exp(-(\alpha_0 + \alpha_1 y_{i1} + \alpha_2 y_{i2} + \dots + \alpha_n y_{in}))}$$

It is easy to infer that  $p_i$  is the probability that the instance  $a_i$  occurs in  $D_c$ . The likelihood function ([4], p. 4) of  $f$  given training data  $D_c$  is:

$$L(\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_n) = \prod_{i=1}^N p_i^{z_i} (1-p_i)^{N-z_i} = \prod_{i=1}^N \left(\frac{p_i}{1-p_i}\right)^{z_i} (1-p_i)^N \quad (4)$$

The logarithm likelihood function of  $f$  is:

$$\begin{aligned}
\text{Log}L(\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_n) &= \log \left( \prod_{i=1}^N \left( \frac{p_i}{1-p_i} \right)^{z_i} (1-p_i)^N \right) \\
&= \sum_{i=1}^N \left( z_i \left( \alpha_0 + \sum_{j=1}^n \alpha_j y_{ij} \right) - N \log \left( 1 + \exp \left( \alpha_0 + \sum_{j=1}^n \alpha_j y_{ij} \right) \right) \right)
\end{aligned} \tag{5}$$

The first order partial derivatives of logarithm likelihood function with regard to  $\alpha_k(s)$  are:

$$\begin{cases} \frac{\partial \text{Log}L}{\partial \alpha_0} = \sum_{i=1}^N \left( z_i - \frac{N \exp \left( \alpha_0 + \sum_{j=1}^n \alpha_j y_{ij} \right)}{1 + \exp \left( \alpha_0 + \sum_{j=1}^n \alpha_j y_{ij} \right)} \right) \\ \frac{\partial \text{Log}L}{\partial \alpha_k} = \sum_{i=1}^N y_{ik} \left( z_i - \frac{N \exp \left( \alpha_0 + \sum_{j=1}^n \alpha_j y_{ij} \right)}{1 + \exp \left( \alpha_0 + \sum_{j=1}^n \alpha_j y_{ij} \right)} \right) \end{cases} \text{ where } 1 \leq k \leq n$$

If there is convention that  $y_{i0} = 1$ , we have:

$$\frac{\partial \text{Log}L}{\partial \alpha_k} = \sum_{i=1}^N y_{ik} \left( z_i - \frac{N \exp \left( \alpha_0 + \sum_{j=1}^n \alpha_j y_{ij} \right)}{1 + \exp \left( \alpha_0 + \sum_{j=1}^n \alpha_j y_{ij} \right)} \right) \text{ where } 0 \leq k \leq n \tag{6}$$

Let  $\alpha_0^*, \alpha_1^*, \dots, \alpha_n^*$  be estimates of  $\alpha_0, \alpha_1, \dots, \alpha_n$ , respectively. These  $\alpha_k^*(s)$  are maximum points of logarithm likelihood function and hence, we set the first order partial derivatives of logarithm likelihood function with regard to  $\alpha_k(s)$  so as to find out  $\alpha_k^*(s)$ .

$$\begin{cases} \frac{\partial \text{Log}L}{\partial \alpha_0} = 0 \\ \frac{\partial \text{Log}L}{\partial \alpha_1} = 0 \\ \vdots \\ \frac{\partial \text{Log}L}{\partial \alpha_n} = 0 \end{cases} \Leftrightarrow \begin{cases} \sum_{i=1}^N \frac{y_{i0} \exp \left( \alpha_0 + \sum_{j=1}^n \alpha_j y_{ij} \right)}{1 + \exp \left( \alpha_0 + \sum_{j=1}^n \alpha_j y_{ij} \right)} = \frac{1}{N} \sum_{i=1}^N y_{i0} z_i \\ \sum_{i=1}^N \frac{y_{i1} \exp \left( \alpha_0 + \sum_{j=1}^n \alpha_j y_{ij} \right)}{1 + \exp \left( \alpha_0 + \sum_{j=1}^n \alpha_j y_{ij} \right)} = \frac{1}{N} \sum_{i=1}^N y_{i1} z_i \\ \vdots \\ \sum_{i=1}^N \frac{y_{in} \exp \left( \alpha_0 + \sum_{j=1}^n \alpha_j y_{ij} \right)}{1 + \exp \left( \alpha_0 + \sum_{j=1}^n \alpha_j y_{ij} \right)} = \frac{1}{N} \sum_{i=1}^N y_{in} z_i \end{cases} \tag{7}$$

Because Equation (7) is the set of  $n + 1$  non-linear equations with  $n + 1$  variables, it is easy to find out its solution  $(\alpha_0^*, \alpha_1^*, \dots, \alpha_n^*)$  by applying some numerical analysis methods such as Newton-Raphson ([5], pp. 67-79) method. Substituting estimates  $\alpha_0^*, \alpha_1^*, \dots, \alpha_n^*$  into Equation (3), it is totally to determine the logistic function  $f$ .

For example, suppose there are 2 items,  $GUP = \{x_1, x_2\}$  receiving binary values such as 0-dislike and 1-like. So, we have  $n = 2, v = 1$  and 4 possible instances  $y_1 = (y_{11} = 0, y_{12} = 0), y_2 = (y_{21} = 0, y_{22} = 1), y_3 = (y_{31} = 1, y_{32} = 0), y_4 = (y_{41} = 1, y_{42} = 1)$ . According to Equation (2), instances of logistic function  $f$  evaluated on  $y_1, y_2, y_3$ , and  $y_4$  are:

$$\begin{aligned}
p_1 &= f(y_{11} = 0, y_{12} = 0) = f(x_1 = 0, x_2 = 0) = \frac{1}{1 + \exp(-\alpha_0)} \\
p_2 &= f(y_{21} = 0, y_{22} = 1) = f(x_1 = 0, x_2 = 1) = \frac{1}{1 + \exp(-\alpha_0 - \alpha_2)} \\
p_3 &= f(y_{31} = 1, y_{32} = 0) = f(x_1 = 1, x_2 = 0) = \frac{1}{1 + \exp(-\alpha_0 - \alpha_1)} \\
p_4 &= f(y_{41} = 1, y_{42} = 1) = f(x_1 = 1, x_2 = 1) = \frac{1}{1 + \exp(-\alpha_0 - \alpha_1 - \alpha_2)}
\end{aligned}$$

Suppose only instance  $y_1 = (y_{11} = 0, y_{12} = 0)$  is observed and so we have  $z_1 = 1, z_2 = z_3 = z_4 = 0$ . The Equation (7) becomes:

$$\begin{cases} 1 - \frac{4 \exp(\alpha_0)}{1 + \exp(\alpha_0)} - \frac{4 \exp(\alpha_0 + \alpha_1)}{1 + \exp(\alpha_0 + \alpha_1)} - \frac{4 \exp(\alpha_0 + \alpha_2)}{1 + \exp(\alpha_0 + \alpha_2)} - \frac{4 \exp(\alpha_0 + \alpha_1 + \alpha_2)}{1 + \exp(\alpha_0 + \alpha_1 + \alpha_2)} = 0 \\ -\frac{4 \exp(\alpha_0 + \alpha_1)}{1 + \exp(\alpha_0 + \alpha_1)} - \frac{4 \exp(\alpha_0 + \alpha_1 + \alpha_2)}{1 + \exp(\alpha_0 + \alpha_1 + \alpha_2)} = 0 \\ -\frac{4 \exp(\alpha_0 + \alpha_2)}{1 + \exp(\alpha_0 + \alpha_2)} - \frac{4 \exp(\alpha_0 + \alpha_1 + \alpha_2)}{1 + \exp(\alpha_0 + \alpha_1 + \alpha_2)} = 0 \end{cases}$$

It is necessary to solve Equation (7) with regard to coefficients  $\alpha_0, \alpha_1,$  and  $\alpha_2$ . Suppose  $\alpha_0 = 0$  and  $\alpha_2 = \alpha_1$ , we have:

$$\begin{cases} 1 - \frac{8 \exp(\alpha_1)}{1 + \exp(\alpha_1)} - \frac{4 \exp(2\alpha_1)}{1 + \exp(2\alpha_1)} = 0 \\ -\frac{4 \exp(\alpha_1)}{1 + \exp(\alpha_1)} - \frac{4 \exp(2\alpha_1)}{1 + \exp(2\alpha_1)} = 0 \end{cases}$$

By using software Mathematica [6], it is easy to find out  $\alpha_0^*, \alpha_1^*$  and  $\alpha_2^*$  as follows:

$$\begin{cases} \alpha_0^* = 0 \\ \alpha_1^* = \alpha_2^* = -2.15109 + 3.14159i \end{cases}$$

where  $i = \sqrt{-1}$  is imaginary unit.

Finally, probabilities  $p_1, p_2,$  and  $p_3$  are determined by substituting complex solutions  $\alpha_0^*, \alpha_1^*$  and  $\alpha_2^*$  into Equation (2) as follows:

$$\begin{aligned} p_1 &= f(x_1 = 0, x_2 = 0) = 0.5 \\ p_2 &= f(x_1 = 0, x_2 = 1) = -0.131679 + 1.82495 * 10^{-17} i \\ p_3 &= f(x_1 = 0, x_2 = 1) = -0.131679 + 1.82495 * 10^{-17} i \\ p_4 &= f(x_1 = 0, x_2 = 1) = 0.0133582 - 3.2281 * 10^{-18} i \end{aligned}$$

where  $i = \sqrt{-1}$  is imaginary unit.

If *GUP* gets the instance  $y_{11} = (x_1 = 0, x_2 = 0)$ , the logistic probability  $p_1 = f(x_1 = 0, x_2 = 0) = 0.5$  which leads to conclude that user does not like such two items.

### 3. Conclusions

The approach in this paper is the hybrid of contextual post-filtering and contextual modeling where logistic model is applied in the post stage of recommendation process. The thinking behind this approach is that rating values obtained explicitly by questionnaires or implicitly by inferring users' behaviors are the most important and contextual factor around users or related to application is additional information which is useful but not essential. Comparing to contextual pre-filtering, this approach restricts the loss of rating information in rating matrix by ignoring data pre-filtering. At the post stage, this approach removes only items which are asserted that users do not like them under contextual condition. Such assertion is the outcome of steady inference model, namely logistic model.

The removal restriction increases recall metric due to reserving solutions space but doesn't lessen precision metric in comparison of pre-filtering method. Comparing to traditional or post-filtering method, this approach is more accurate because of the steady inference mechanism of logistic function when logistic model is appropriate to binary request such as following yes/no question "whether or not Mary prefers to browse commercial websites in the evening". Moreover this approach exploits the relationship among items in general user pattern, which is necessary to recommendation process but is not considered in contextual pre-filtering or post-filtering.

---

## Acknowledgements

I express my deep gratitude to Prof. Dr. Ho, Hang T. T., Vinh Long General Hospital, Vietnam Ministry of Health who funded me to complete and publish this research.

## References

- [1] Ricci, F., Rokach, L., Shapira, B. and Kantor, P.B. (2011) Recommender Systems Handbook (Vol. I). Ricci, F., Rokach, L., Shapira, B. and Kantor, P.B., Eds., Springer, New York, Dordrecht, Heidelberg, London. <http://dx.doi.org/10.1007/978-0-387-85820-3>
- [2] Palmisano, C., Tuzhilin, A. and Gorgoglione, M. (2008) Using Context to Improve Predictive Modeling of Customers in Personalization Applications. *IEEE Transactions on Knowledge and Data Engineering*, **20**, 1535-1549. <http://dx.doi.org/10.1109/TKDE.2008.110>
- [3] Montgomery, D.C. and Runger, G.C. (2003) Applied Statistics and Probability for Engineers. 3rd Edition, John Wiley & Sons, Inc., New York.
- [4] Czepiel, S.A. (2002) Maximum Likelihood Estimation of Logistic Regression Models: Theory and Implementation. <http://czep.net>
- [5] Burden, R.L. and Faires, D.J. (2011) Numerical Analysis. Julet, M., Ed., 9th Edition, Brooks/Cole Cengage Learning, 872.
- [6] Wolfram. (n.d.) Mathematica. *Wolfram Mathematica*, **10**. <http://www.wolfram.com/mathematica>



Scientific Research Publishing

### Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>