Multi-criteria based Item Recommendation Methods

Noor Ifada\textsuperscript{1,}\textsuperscript{2,}\textsuperscript{3}, Syafrurrizal Naridho\textsuperscript{2} and Mochammad Kautsar Sophan\textsuperscript{3}

\textsuperscript{1,2,3)Informatics Department, University of Trunojoyo Madura
\textsuperscript{1)noor.ifada@trunojoyo.ac.id\textsuperscript{2)rizal14115@gmail.com \textsuperscript{3}kautsar@trunojoyo.ac.id

DOI: https://doi.org/10.21107/rekayasa.v12i2.5913

ABSTRACT

This paper comprehensively investigates and compares the performance of various multi-criteria based item recommendation methods. The development of the methods consists of three main phases: predicting rating per criterion; aggregating rating prediction of all criteria; and generating the top-$N$ item recommendations. The multi-criteria based item recommendation methods are varied and labelled based on what approach is implemented to predict the rating per criterion, i.e., Collaborative Filtering (CF), Content-based (CB), and Hybrid. For the experiments, we generate two variations of datasets to represent the normal and cold-start conditions on the multi-criteria item recommendation system. The empirical analysis suggests that Hybrid and CF are best implemented on the normal and cold-start item conditions, respectively. On the other hand, CB should never be (solely) implemented in a multi-criteria based item recommendation system on any conditions.

Keyword: Collaborative Filtering, Content-based, Hybrid, Multi-criteria recommendation method.

INTRODUCTION

Recommendation System is a firmed research topic (G. Adomavicius & Tuzhilin, 2005; Zhang, Zhou, & Zhang, 2011). A multi-criteria recommendation system allows users to specify ratings based on various criteria (Gediminas Adomavicius, Manouselis, & Kwon, 2011; Aggarwal, 2016). For example, a user may rate tourism attraction based on attraction, accessibility, amenities, and ancillary. Such a system implements a recommendation method in which the user’s preference for an item is represented as a vector of ratings corresponding to various criteria (Aggarwal, 2016).

The rating prediction algorithms can be categorized into three approaches: Collaborative Filtering (CF), Content-based (CB), and Hybrid (G. Adomavicius & Tuzhilin, 2005). The CF approach is the most popular rating prediction approach which generates the predictions based on the similarity of users or items (Su & Khoshgoftaar, 2009). The CB approach is popular for generating prediction when the information of the item’s content is available (G. Adomavicius & Tuzhilin, 2005; Lops, Gemmis, & Semeraro, 2011). Meanwhile, the Hybrid approach is comprising both the Collaborative Filtering and Content-based approaches (G. Adomavicius & Tuzhilin, 2005; Burke, 2007). Researchers have proposed methods various methods based on the three rating prediction approaches (Gediminas Adomavicius et al., 2011; Fuchs & Zanker, 2012; Jannach, Karakaya, & Gedikli, 2012; Lakiotaki, Matsatsinis, & Tsoukias, 2011; Manouselis & Costopoulou, 2007). However, to the best of our knowledge, no in-depth work has been done that comprehensively compares the performance of the multi-criteria recommendation method variations.

Cite this as:

© 2019 Noor Ifada, Syafrurrizal Naridho, Mochammad Kautsar Sophan
In this paper, we study a variation of multi-criteria based item recommendation methods. The development of the methods consists of three main phases: predicting rating per criterion; aggregating rating prediction of all criteria; and generating the top-$N$ item recommendations. The multi-criteria based item recommendation methods are varied and labelled based on what approach is implemented to predict the rating per criterion, i.e., Collaborative Filtering (CF), Content-based (CB), and Hybrid. For the experiments, we generate two variations of datasets to represent the normal and cold-start conditions on the multi-criteria item recommendation system.

The contributions of this paper are as follows: (1) Analysis of multi-criteria based recommendation methods and show the different approaches to implementing the rating prediction per criterion phase, and (2) Performance comparison of three multi-criteria based item recommendation methods.

**METHOD**

In this section, we explain how to develop the multi-criteria based methods for generating a top-$N$ list of item recommendation to a target user. The input of the method is the multi-criteria rating histories of users; the item features data; and the significance score of each criterion. At this stage, we can identify and notate the set of users, items, criteria, significance score of each criterion, features, items which have been rated by each user, and features of each item. Table I lists the notations used in this paper to symbolize those sets.

The development of the multi-criteria based item recommendation method consists of three main phases: (a) predicting rating per criterion, (b) aggregating rating prediction of all criteria, and (c) generating the top-$N$ item recommendations. Figure 1 shows the framework of the method.

**Prediction Rating per Criterion**

The calculation of rating prediction can be categorized into three approaches: Collaborative Filtering, Content-based, and Hybrid approach (G. Adomavicius & Tuzhilin, 2005).

**Collaborative Filtering Multi-criteria based Approach**

The Collaborative Filtering (CF) approach is the most popular rating prediction approach (Su & Khoshgoftaar, 2009). A way to implement this

---

Table I. Notations

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U = {u_1, u_2, u_3, ..., u_p}$</td>
<td>set of $p$ users</td>
</tr>
<tr>
<td>$I = {i_1, i_2, i_3, ..., i_q}$</td>
<td>set of $q$ items</td>
</tr>
<tr>
<td>$C = {c_1, c_2, c_3, ..., c_l}$</td>
<td>set of $l$ criteria</td>
</tr>
<tr>
<td>$Y = {y_{1}, y_{2}, y_{3}, ..., y_{l}}$</td>
<td>set of the significance scores of $l$ criteria</td>
</tr>
<tr>
<td>$\Delta = {\delta_1, \delta_2, \delta_3, ..., \delta_f}$</td>
<td>set of $f$ features.</td>
</tr>
<tr>
<td>$r_{u,i}^c$</td>
<td>rating given by user $u$ to item $i$, for criterion $c$</td>
</tr>
<tr>
<td>$\delta_{i,f}$</td>
<td>Designation of item’s feature, i.e., labelled as either 1 or 0</td>
</tr>
<tr>
<td>$I_u$</td>
<td>set of items which have been rated by user $u$</td>
</tr>
<tr>
<td>$F_i$</td>
<td>set of features of item $i$</td>
</tr>
</tbody>
</table>

Figure 1. The Framework of Multi-criteria based Item Recommendation Method
Multi-criteria based Item Recommendation Methods

approach is by calculating the prediction based on the similarities of users rating preferences (Gediminas Adomavicius et al., 2011; Bilge & Yargıç, 2017). Per criterion $c$, we calculate the similarity between user $u$ and $v$ using the cosine similarity function:

$$s_{u,v}^c = \frac{\sum_{i \in I_u \cap I_v} r_{u,i}^c \cdot r_{v,i}^c}{\sqrt{\sum_{i \in I_u} (r_{u,i}^c)^2} \cdot \sqrt{\sum_{i \in I_v} (r_{v,i}^c)^2}}$$ (1)

Once the similarities between target user $u$ and other users are calculated, we can form the top-$k$ nearest neighbours of user $u$, $K_u$. The CF rating prediction of user $u$ to item $i$ for criterion $c$ is formulated as:

$$\alpha_{u,i}^c = \frac{\sum_{v \in K_u} s_{u,v}^c \cdot r_{v,i}^c}{\sum_{v \in K_u} |s_{u,v}^c|}$$ (2)

where $|K_u| \leq k$.

Content-based Multi-criteria based Approach

The Content-based (CB) approach is popular for generating prediction when the information of the item’s content is available (G. Adomavicius & Tuzhilin, 2005; Lops et al., 2011). A way to implement this approach is by calculating the prediction based on the weight of each item feature (Uluyagmur, Cataltepe, & Tayfur, 2012). Per criterion, we calculate the weight of item feature $f$ for user $u$ using the weighting function:

$$w_{u,f}^c = \frac{\sum_{i \in I_u} \delta_{i,f} \cdot r_{u,i}^c}{|I_u|}$$ (3)

The CB rating prediction of user $u$ to item $i$ for criterion $c$ is formulated as:

$$\beta_{u,i}^c = \frac{\sum_{f \in F_i} w_{u,f}^c}{|F_i|}$$ (4)

Hybrid Multi-criteria based Approach

The Hybrid approach is comprising both the Collaborative Filtering and Content-based approaches (G. Adomavicius & Tuzhilin, 2005; Burke, 2007). A way to implement this approach is by calculating the prediction based on the average of CF and CB rating predictions (Aggarwal, 2016). The Hybrid rating prediction of user $u$ to item $i$ for criterion $c$ is formulated as:

$$\gamma_{u,i}^c = \frac{\alpha_{u,i}^c + \beta_{u,i}^c}{2}$$ (5)

Aggregating Rating Prediction of All Criteria

Given the rating prediction of user $u$ to item $i$ for each criterion $c$, we now need to aggregate the prediction of all criteria. The aggregated prediction of all criteria is formulated as (Aggarwal, 2016):

$$\theta_{u,i} = \frac{\sum_{c \in C} \alpha_{u,i}^c \cdot \delta_{i,c}^c \cdot \sum_{c \in C} \beta_{u,i}^c \cdot \delta_{i,c}^c}{\sum_{c \in C} |\delta_{i,c}|}$$ (6)

<table>
<thead>
<tr>
<th>Item</th>
<th>$i_1$</th>
<th>$i_2$</th>
<th>$i_3$</th>
<th>$i_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1$</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>$c_2$</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>$c_3$</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>$c_4$</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Figure 2. A toy example of multi-criteria rating data

| Feature | $\delta_1$ | $\delta_2$ | $\delta_3$ | $\delta_4$ | $\delta_5$
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$i_1$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$i_2$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$i_3$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$i_4$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 3. A toy example of item feature data

| Criteria | $c_1$ | $c_2$ | $c_3$
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Score (y)}$</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 4. A toy example of the criteria significance scores
Generating Top-N Item Recommendation

The list of top-N list item recommendation for target user \( u \), \( \text{Top}_u(N) \), is generated based on the aggregated rating predictions. In this case, item \( i \) is listed in the \( \text{Top}_N \) if \( \theta_{u,i} \) is within \( N \) highest values in \( \theta_u \).

**Toy Example**

This section shows examples of how the multi-criteria based methods for generating a top-N list of item recommendation to a target user are developed based on the CF, CB, and Hybrid approaches. Figure 2, Figure 3, and Figure 4 respectively show the toy examples of multi-criteria rating data, item features data and criteria significance scores. The complete calculation of the three main phases in the multi-criteria based item recommendation method is presented in Figure 5.

Figure 2 shows that there are four users \( U = \{ u_1, u_2, u_3, u_4 \} \), four items \( I = \{ i_1, i_2, i_3, i_4 \} \), and three criteria \( C = \{ c_1, c_2, c_3 \} \). The set of users’ items are formed as \( I_1 = \{ 1, 2, 3, 4 \} \), \( I_2 = \{ 2, 3 \} \), \( I_3 = \{ 1, 2, 3, 4 \} \), and \( I_4 = \{ 1, 2, 3, 4 \} \). On the other hand, Figure 3 shows that there are five item features data \( \Delta = \{ \delta_1, \delta_2, \delta_3, \delta_4, \delta_5 \} \), in which the set of items’ features are formed as \( F_1 = \{ 1, 2, 3, 5 \} \), \( F_1 = \{ 1, 3, 4, 5 \} \), \( F_1 = \{ 1, 5 \} \), and \( F_1 = \{ 1, 4, 5 \} \). Meanwhile, Figure 4 shows that the significance score of each criterion is \( y_1 = 2 \), \( y_2 = 3 \), and \( y_3 = 1 \).

For our target user \( u_2 \), we are generating the top-N item recommendation \( \text{Top}_2(N) \) based on the aggregated rating prediction \( \theta_{2,1} \) and \( \theta_{2,4} \). The rating prediction per criterion based on CF, CB, and Hybrid approaches are respectively calculated using Equation (2), (4), and (5). Whereas the aggregated rating prediction is calculated using Equation (6). Figure 5 shows that the \( \text{Top}_2(N) \) generated based on CF, CB, and Hybrid based methods are respectively \{1,4\}, \{4,1\}, and \{1,4\}.

**RESULTS AND DISCUSSION**

In this section, we present the empirical analysis by conducting experiments to evaluate the performance of the multi-criteria based item recommendation methods built in this paper.

**Dataset and Experiment Procedure**

This paper uses a tourism multi-criteria rating dataset that consists of 77 users, 50 touristic attractions or items, and 900 rating data. Table II lists the details of the dataset. Our experimentations implement the 5-fold cross-validation evaluation approach, in which each fold has a training set \( D_{\text{train}} \) and a test set \( D_{\text{test}} \).

For the experiments, we generate two variations based on the tourism dataset to represent the normal and cold-start conditions:

- **TN**: The dataset is refined such all items and users in \( D_{\text{test}} \) occurred several times in \( D_{\text{train}} \). This dataset represents the normal condition.
- **TCS**: The dataset is refined such that there occurred items in \( D_{\text{test}} \) that have not been rated by users in \( D_{\text{train}} \). This dataset represents the condition in which the cold-start item problem occurs.

**Evaluation Criteria**

Our recommendation methods build the model using \( D_{\text{train}} \) and utilize it afterwards to generate a top-N item recommendations for target users in \( D_{\text{test}} \). In this case, the evaluation is conducted by comparing the top-N list of item recommendations for a user \( u \), \( \text{Top}_u(N) \), to the ground-truth items listed in \( D_{\text{test}} \), \( GT_u \). We use the AP (Average Precision) evaluation metric to measure the performance of recommendations. The AP score of the first \( N \) list of item recommendations for a target user \( u \) is formulated as:

\[
\text{AP}_u(N) := \sum_{n=1}^{N} \text{Precision}_u(n) \cdot \mathbb{1}(\text{Top}_u(n) \in GT_u)
\]

where

\[
\text{Precision}_u(N) := 100 \cdot \frac{|\text{Top}_u(N) \cap GT_u|}{N}
\]
Predicting Rating per Criteria where \( I_1 = \{1, 2, 3, 4\}, I_2 = \{2, 3\}, I_3 = \{1, 2, 3, 4\}, F_1 = \{1, 2, 3, 5\}, \) and \( I_4 = \{1, 4, 5\}\):

<table>
<thead>
<tr>
<th>Criteria</th>
<th>( s_{1,1} )</th>
<th>( s_{1,2} )</th>
<th>( s_{1,3} )</th>
<th>( s_{1,4} )</th>
<th>( a_{1,1}^1 )</th>
<th>( a_{1,4}^1 )</th>
<th>( \theta_{2,1} )</th>
<th>( \theta_{2,4} )</th>
<th>( \text{Top-}N )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \frac{1}{3(2)} + 2x + 2y )</td>
<td>0.87</td>
<td>( \frac{1}{3(2)} + 2x )</td>
<td>0.92</td>
<td>4.48</td>
<td>1.97</td>
<td>4.08</td>
<td>3.00</td>
<td>{1, 4}</td>
</tr>
<tr>
<td>2</td>
<td>0.86</td>
<td>0.89</td>
<td>0.98</td>
<td>( \frac{1}{3(2)} + 2x + 2y )</td>
<td>0.78</td>
<td>0.86</td>
<td>3.48</td>
<td>3.48</td>
<td>{1, 4}</td>
</tr>
</tbody>
</table>

Aggregating Rating Prediction of all Criteria where \( S = \{2, 3, 1\}\):

<table>
<thead>
<tr>
<th>Criteria</th>
<th>( w_{2,1} )</th>
<th>( w_{2,2} )</th>
<th>( w_{2,3} )</th>
<th>( w_{2,4} )</th>
<th>( w_{2,5} )</th>
<th>( \gamma_{2,1} )</th>
<th>( \gamma_{2,4} )</th>
<th>( \text{Top-}N )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \frac{1}{3(3)} + 1(1) + 1(2) )</td>
<td>2.50</td>
<td>1.50</td>
<td>1.50</td>
<td>1.50</td>
<td>3.00</td>
<td>1.88</td>
<td>{1, 4}</td>
</tr>
<tr>
<td>2</td>
<td>( \frac{1}{3(3)} + 1(0) + 1(2) )</td>
<td>1.50</td>
<td>1.50</td>
<td>1.50</td>
<td>1.50</td>
<td>3.52</td>
<td>2.00</td>
<td>{1, 4}</td>
</tr>
<tr>
<td>3</td>
<td>( \frac{1}{3(3)} + 1(1) + 1(2) )</td>
<td>1.50</td>
<td>1.50</td>
<td>1.50</td>
<td>1.50</td>
<td>5.00</td>
<td>1.38</td>
<td>{1, 4}</td>
</tr>
</tbody>
</table>

Hybrid Approach

<table>
<thead>
<tr>
<th>Criteria</th>
<th>( \gamma_{2,1} )</th>
<th>( \gamma_{2,4} )</th>
<th>( \text{Top-}N )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.05</td>
<td>1.82</td>
<td>{1, 4}</td>
</tr>
<tr>
<td>2</td>
<td>2.94</td>
<td>2.76</td>
<td>{1, 4}</td>
</tr>
<tr>
<td>3</td>
<td>2.43</td>
<td>2.57</td>
<td>{1, 4}</td>
</tr>
</tbody>
</table>

Figure 5. Examples of generating the top-\( N \) item recommendation for the target user \( u_2 \) (\( \text{Top-}N \))
It is to be noted that \( I(\cdot) \) results 1 when the condition within the bracket is fulfilled, or 0 otherwise.

**Performance Comparison**

In this sub-section, we compare the performance at \( N = \{1 \ldots 20\} \) of the three developed three multi-criteria based item recommendation methods. For ease of explanation, we label the methods based on the rating prediction per criterion approach implemented, i.e., CF, CB, and Hybrid.

Figure 6 shows the performance comparison of the three methods on the LN dataset. We can observe that CF and Hybrid have comparable performance, while CB performs the worst. Note that Hybrid slightly outperforms CF when \( N \geq 5 \). These results suggest that, on a normal condition, Hybrid is the best method to be implemented in a multi-criteria based item recommendation system.

Figure 7 shows the performance comparison on the LCS dataset. We can notice that CF achieve the best performance, followed by Hybrid and CB. These results advise that a multi-criteria based item recommendation system must implement a CF method on a cold-start item condition.

Additionally, the poor performance of CB confirms that this method should not be implemented solely in a multi-criteria based item recommendation system on any conditions. It is also worth it to note that CF can be more effective than CB at providing recommendations for cold-start items condition since CF performance is better than that of CB on the LCS dataset.
dataset, contradicting the results of previous researches (G. Adomavicius & Tuzhilin, 2005; Aggarwal, 2016).

**CONCLUSION**

In this paper, we develop three multi-criteria based item recommendation methods. The methods are varied and, therefore, labelled based on what approach is implemented to predict the rating per criterion, i.e., Collaborative Filtering (CF), Content-based (CB), and Hybrid. For a deeper analysis, we generate two variations of datasets to represent the normal and cold-start conditions of multi-criteria item recommendation system. The experiment results suggest that Hybrid and CF are best implemented on the normal and cold-start item conditions, respectively. On the other hand, CB should never be (solely) implemented in a multi-criteria based item recommendation system on any conditions.

Nevertheless, we believe that this study needs to be further expanded. Our recommendations are to implement: (1) other functions of CF, CB and, Hybrid approaches for predicting the rating per criterion; or/and (2) other functions for aggregating rating prediction of all criteria. Alternatively, we may build an application to test the user experience on the multi-criteria item recommendation system based on the CF and Hybrid methods developed in this paper.

**REFERENCES**


