Towards Generating Personalized Country Recommendation

Ayoub EL Majjodi, Mehdi Elahi, Nabil EL Ioini, and Christoph Trattner
ENSIAS, Mohammed V University, Rabat, Morocco
ayoub.elmajjodi@um5s.net.ma
Free University of Bozen - Bolzano, Bolzano, Italy
nabil.elioini@unibz.it

Mehdi Elahi, University of Bergen, Bergen, Norway
mehdi.elahi@uib.no
Christoph Trattner, University of Bergen, Bergen, Norway
christoph.trattner@uib.no

ABSTRACT
The rise in international migration over the past decades has given more audience to this crucial issue of human life. According to reports by United Nations, more than 243 million people live in a country that is not their place of birth. People decide to immigrate, based on a range of reasons, and choose the country of destination with the hope to begin a new life. However, such a risky decision may not necessarily lead to an improvement of life and in many cases could result in complete dissatisfaction of the emigrating person, and in the extreme cases, cause human catastrophe.

Recommender Systems (RSs) are tools that could mitigate this problem by supporting the people in their decision making process. RSs can interact with the people who are willing to immigrate and acquire certain information about their preferences on potential destinations. Accordingly, RSs can build predictive models based on the acquired data and offer suggestions on where could be a better match for the specific preferences and constraints of people.

This work is an attempt to build a RS that can be used in order to receive personalized recommendation of countries. The system is capable of eliciting preferences of users in the form of ratings, learning from the preferences, and intelligently generating a personalized ranking list of countries for every target user. We have conducted a user study in order to evaluate the quality of the recommendation, measured in terms of accuracy, diversity, novelty, satisfaction, and capability to understand the particular preferences of different users. The results were promising and indicated the potentials of the generating personalized recommendations in this less-explored domain.

1 INTRODUCTION
World immigration accounts for more than 3.3% of the population share. Foreign-born people in developed countries increased from 7% to over 10% in recent years [11]. Today, “much of the developed world is now increasingly composed of nations of immigrants” [4].

Deciding to immigrate is indeed a challenging decision and it could be based on a range of reasons. Immigrants choose the country of destination with the hope to begin a new life. However, such a risky decision may not necessarily lead to a better life. This could be the case even for highly educated immigrants willing to find a better place to live. Improper decisions based on inaccurate data and misleading information could result in a complete dissatisfaction of the immigrating person, and in extreme cases, possibly result in human catastrophe.

The primary challenge for immigrating people is to accurately identify the best possible destination, matching their professional and personal preferences. The belief is that the new place will hopefully meet their expectations and result in a relative life satisfaction. There are a few number of (annually-published) ranking lists, presenting the best destinations to live in the world [35, 38]. However none of these lists offer a proper personalization mechanism that provides a ranking list of countries that match individual needs and constraints of an immigrating individual.

Recommender Systems (RSs) are decision support tools that can assist the users in making better decisions [31]. RSs can mitigate this challenge by making personalized recommendation of countries to immigrating people based on their particular preferences. These systems have been found successful in various application domains, including e-business [32, 37], entertainment [3, 6], tourism [5, 17], restoration [26, 28], health [18, 25], music [33, 34], art [10, 27], education [21, 24], and even e-business [1, 8]. However, immigration domain substantially differs from all of these domains. The cost of irrelevant and incorrect recommendation in many of the noted domains could be negligible. Watching a wrong movie or listening to a boring song would not really cause a serious problem. However, immigrating to an unsuitable place, based on wrong information in social media or wrong suggestions could be extremely costly in any aspect (e.g., relocation cost or difficulty of learning a new language).

This work addresses this challenge by proposing a novel country RS. The system is capable of effectively eliciting the preferences of users, provided in terms of ratings, learn from these preferences...
and builds a recommendation models that can offer personalized

suggestion of countries.

We have developed an evaluation methodology in order to assess

the quality of recommendation made by our system. In a prelim-

inary offline experiment, we have compared a range of common

recommender algorithms in order to identify the top performing

algorithms. Then, we have conducted a real user study where we

requested users to compare the top algorithms in terms of different

dimensions. The results have shown that there is no best recommender
algorithm. While a certain algorithm could offer higher accuracy

of recommendation, however, other algorithms could recommend

a more diverse and novel recommendations that can better suit a

wider range of preferences.

We believe that our work is the first attempt to build a RS in

this novel application domain, and to the knowledge of authors, no

prior work addressed the challenge of recommendation of countries

to people.

2 METHODOLOGY

2.1 Recommender Algorithms

The algorithms adopted for recommendation mainly focus on using

a classical recommendation model called Collaborative Filtering
(CF) [22]. Techniques based on CF exploit ratings provided by a

network of users in order to predict the missing ratings of the items.

The items with the highest predicted ratings are recommended to

users. [13] briefly describes the approaches based on CF.

Neighbourhood-based this kind of algorithm such as k-nearest

neighbors (KNN) computes rating prediction exploiting two sets

of preference data: the ratings of the user for other items and the

ratings of other like-minded users. The item’s rating prediction is

computed based on how the item was rated by the users similar

to the target user. The rating \( \hat{r}_{ui} \) for the user \( u \) and the item \( i \) is

predicted in the following way:

\[
\hat{r}_{ui} = r_u + \frac{\sum_{u' \in N_{Ni(u)}} \text{sim}(u, u') (r_{u'i} - r_{u'})}{\sum_{u' \in N_{Ni(u)}} |\text{sim}(u, u')|}
\]

(1)

where \( \bar{r}_u \) denotes the average ratings of user \( u \), \( \text{sim}(u, u') \) is a similarity

measure between two users \( u \) and \( u' \), and \( N_{Ni(u)} \) is a set of users

similar to user \( u \) (neighbours) who rated item \( i \). We computed

the similarity based on Cosine metric.

Matrix Factorization techniques such as SVD learns for both users

and items factor vectors of the same size. These vectors are inferred

from the user’s rating. Each element of the factor vector, assigned

to an item, reflects how well the item represents a particular latent

aspect. User factor vectors measure the preference of the user for

each factor. The task of the factorization is to split the matrix of

ratings \( R \) into two matrices \( S \) and \( M \):

\[
R \approx SM^T
\]

(2)

where \( S \) is \( |U| \times F \) matrix, and \( M \) is \( |I| \times F \) matrix. \( F \) represents the

number of factors we wish to use. Then, predictions for the ratings

are computed as follows [16]:

\[
\hat{r}_{ui} = \sum_{f=1...F} s_{uf} m_{if}
\]

(3)

where \( s_{uf} \) denotes how much the user \( u \) likes the factor \( f \) and the

value \( m_{if} \) denotes how strong the factor \( f \) is in the item \( i \).

We adopted different types of neighborhood-based recommenders,

i.e., KNN Basic, KNN With Means, and KNN With Baseline. The

former is a simple version while KNN With Means is an extended

version that simply takes it into account the the mean ratings of

each user. KNN With Baseline, on the other hand, takes into ac-

count the baseline rating. This is a factor that is estimated through

a learning process. The number of neighbors for different versions

of KNN algorithms is set to 40.

We also adopted two types of matrix factorization recommenders,

i.e., SVD and SVD++ [23]. The later is an extension of SVD as it is

capable of taking into account implicit ratings [14]. For both we set

the number of factors to 100.

Co-Clustering is a different algorithm that groups similar users

and similar videos clusters [19, 30]. The prediction of \( \hat{r}_{ui} \) is com-

puted by assigning the users and items to some clusters \( \hat{C}_u \), \( \hat{C}_i \) and

co-cluster \( \hat{C}_{ui} \) where

\[
\hat{r}_{ui} = \hat{C}_u + (\mu_u - \hat{C}_u) + (\mu_i - \hat{C}_i)
\]

where \( \hat{C}_ui \) is the average rating of co-cluster \( \hat{C}_{ui} \), \( \hat{C}_u \) is the aver-

age rating of u’s cluster, and \( \hat{C}_i \) is the average ratings of i’s clus-

ter. \( \mu_i \) and \( \mu_i \) represent the average rating for user \( u \) and item

\( i \), respectively. The clusters are assigned using a straightforward

optimization method [19].

2.2 Implementation Details

We implemented a dynamic system, to provide recommendations in

the real time to the active users and collect enough data for further

analysis. We designed a system following a sample architecture

shown in Figure 1.

![Figure 1: The System architecture.](image)

The user interacts with the deployed system, after registration

into the system by providing the basic information and filling up

a short questionnaire. The questionnaire includes questions that

identify the personality characteristics of the user. We collect the

personality characteristics of the users for future studies (e.g., in

[2, 15]). The users are also requested to enter a number of (1-3)

key factors (features) that plays the most important role if making

decision to which country to immigrate. The list of these features

are provided in section 3.1 where we discuss the dataset.

Then, the user enters the preference elicitation phase where she

is requested to choose the countries she has been or well-know and

provide her ratings in the 5-point Likert scale (1-5).
When the preferences are received by the system, the top three recommender algorithms exploit these preferences in order to independently generate personalized recommendation of countries for the user. The recommendations are presented to the user in the form of three different lists, each generated by a different recommender algorithm. Each list consists of 5 recommended countries. The user is requested to proceed and complete a questionnaire, adopted to assess quality of the recommendations in terms of various metrics. We have adopted a validated questionnaire that can measure perceived quality of recommendation [12]. Perceived quality is the degree to which users assess recommendations positively and express their experience with the recommender system. Assessment of perceived recommendation quality could be implemented in terms of different metrics [12, 20, 29]:

- **Accuracy** (also called **Relevance**): how much the recommendations match the users’ interests, preferences and tastes;
- **Diversity**: how much users perceive recommendations as different from each other, e.g., movies from different genres;
- **Perceived personalization ‘Understands Me’**: the user’s perception that the recommender understands their tastes and can effectively adapt to them;
- **User Satisfaction**: the global users’ feeling of the experience with the recommender system.
- **Novelty**: the extent to which users receive new recommended movies;

The questionnaire consists of 14-question, as also used in [12]. The questions address different metrics, i.e., Accuracy (Q1,Q2), Diversity (Q3-Q5), Understand Me (Q6-Q8), Satisfaction (Q9,Q10) and Novelty (Q11-Q14):

- **Q1**: Which list has more selections that you find appealing?
- **Q2**: Which list has more obviously bad suggestions for you?
- **Q3**: Which list has more countries that are similar to each other?
- **Q4**: Which list has a more varied selection of countries?
- **Q5**: Which list has countries that match a wider variety of preferences?
- **Q6**: Which list better reflects your preferences in countries?
- **Q7**: Which list seems more personalized to your preferences?
- **Q8**: Which list represents more mainstream preferences instead of your own?
- **Q9**: Which list would better help you find countries to consider?
- **Q10**: Which list would you be more likely to recommend to your friends?
- **Q11**: Which list has more countries you did not expect?
- **Q12**: Which list has more countries that are familiar to you?
- **Q13**: Which list has more pleasantly surprising suggestions?
- **Q14**: Which list provides fewer new suggestions?

### 3 EXPERIMENTAL SETUP

In this section, we provide the details of the conducted experiments, aimed to investigate the effectiveness of our system.

#### 3.1 Dataset

Initially, we have collected a small dataset of user preferences provided in the form of ratings for countries. We have shared a link and contacted people, through social networks (e.g., LinkedIn), and requested them to participate in a short survey where the participants were asked to enter the most important factors (i.e., features) that they consider if choosing a country for immigration. The participation was on the voluntary basis and participants were not paid. We have performed a literature review and choose twelve factors that are shown to be the most important for the immigration [9, 36, 39]. The factors are **Education quality, Political insecurity, Work opportunities, Health care, Income difference, Wars and dictatorship, Family member abroad, Cultural and linguistic similarities, Working atmosphere, Shorter distance, social conflict, and Crime rate**. Table 1 represents the list of the features that have been chosen by participants, ranked according to the frequency of choices.

The participants were also requested to rate a number of countries that might be familiar to them. The rating would indicate how a country may match the overall expectation of the participant, according to the particular factors participants have chosen. The list of countries have been obtained from yearly migration report of 2018, by the International Organization for Migration (IOM), reported to be the top destination countries for international migrants. During the 13 days of data collection process, we received 3400 ratings from 136 participants. We used this rating dataset and performed a preliminary experiment where we have cross-validated a set of recommender algorithms and selected the top performing algorithms.

We integrated these algorithms in our recommender system where the users can register, provide their particular preferences, and receive lists of recommendation and evaluate them through a questionnaire. The design of the experiment was a within-subjects study. We have contacted the previous participants as well as new ones (through social networks) and requested them to complete the whole process of using our system.

At the end, a total of 281 participants have registered to the system, while only 189 participants provided their preferences and completed the final evaluation questionnaire (i.e., 67 % of all). By average, participants provided 7 ratings (minimum 5 ratings and maximum 41 ratings) resulting in a dataset of 1284 ratings. The participants were from a wide range of countries, namely, USA, Spain, Canada, Germany, Hungary, Netherlands, Italy, Belgium,
Table 2: Offline evaluation: comparison of different recommender algorithms.

<table>
<thead>
<tr>
<th>Recommender</th>
<th>RMSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVD</td>
<td>21.71</td>
<td>16.06</td>
</tr>
<tr>
<td>SVD++</td>
<td>58.21</td>
<td>51.06</td>
</tr>
<tr>
<td>Non-negative MF</td>
<td>48.65</td>
<td>41.43</td>
</tr>
<tr>
<td>Slope One</td>
<td>23.47</td>
<td>18.62</td>
</tr>
<tr>
<td>KNN basic</td>
<td>21.97</td>
<td>16.88</td>
</tr>
<tr>
<td>KNN with Mean</td>
<td>21.59</td>
<td>16.49</td>
</tr>
<tr>
<td>KNN Baseline</td>
<td>21.58</td>
<td>16.47</td>
</tr>
<tr>
<td>Co-clustering</td>
<td>21.39</td>
<td>16.63</td>
</tr>
<tr>
<td>Baseline Only</td>
<td>23.50</td>
<td>19.21</td>
</tr>
<tr>
<td>Random Predictor</td>
<td>37.93</td>
<td>30.91</td>
</tr>
</tbody>
</table>

We have identified three top recommender algorithms: SVD, KNN With Baseline, and Co-Clustering. Accordingly, the users can register, provide their preferences, and receive three lists of recommendations, each generated by one of these algorithms. At the end, the system requests users to compare the quality of recommendation lists in terms of different metrics.

Table 3 presents the results that we have obtained in online evaluation process. As it can be seen, in terms of accuracy, understanding users and user satisfaction, SVD has achieved the best results. In almost all questions, related to these metrics, SVD recommender has overcome the other recommenders. For positively-formulated questions, i.e., Q1, Q6, Q7, Q9, and Q10, majority of users has chosen this recommender as the best among three recommenders. For one of the negatively-formulated questions, i.e., Q2, minority of the users has chosen this recommender which again indicates the excellent performance of this recommender. The only exception is Q8, formulated negatively and user decided that Co-clustering is the best. This question asks about the recommendation list that represents more mainstream preferences than the users preferences. The least percentage of users chose Co-clustering which means that recommender algorithm generates a more specific recommendations than others. This is an unexpected outcome and we will investigate it more in our follow-up studies.

In terms of diversity, the SVD did not perform well in comparison to the other recommenders, i.e., KNN With Baseline and Co-Clustering. For questions Q3 and Q4, both of the KNN and Co-Clustering recommenders similarly performed well. For question Q5, which is positively-formulated, majority of users chose KNN With Baseline as the best recommender. This questions asks about the recommendation list that has countries which match a wider variety of preferences. Accordingly, the KNN With Baseline is capable of generating recommendations matching a divers range of preferences.

In terms of novelty, again SVD did not perform well, while the other recommenders performed well side-by-side. Indeed, each of these algorithms achieved the best results for 2 out of 4 questions. For questions Q13 and Q14, KNN With Baseline outperformed other recommenders whereas for questions Q11 and Q12, Co-Clustering achieved the best results. Hence, the users found suggestions based on KNN With Baseline more pleasantly surprising and containing more new suggestions. On the other hand, users found Co-Clustering recommending countries that users did not expect and hence look more novel.

Overall, these promising results are very interesting; they show that there is no always-winning recommender algorithm in terms of all evaluation metrics. The results also clearly represent a trade-off between (a) accuracy of recommendation, in one side, and (b) novelty and diversity in an other side.

5 CONCLUSIONS

We have developed a novel RS that can be a support in the decision making process of people who might be willing to immigrate to a new country. The system can elicit preferences of such people and learn from the elicited preferences and ultimately build predictive models that can generate personalized recommendations.

We have evaluated the system through an evaluation methodology which consists of offline and online experiments. The results have shown that different recommender algorithms could optimize a different evaluation metric. While certain algorithms (such as SVD) could result in higher recommendation accuracy and
improved perceived satisfaction, however, they may fall short in properly diversifying the recommendations.

It is worth noting that the research field of immigration is highly complex and sensitive. Indeed, this work is an initial step towards exploring the potentials of the recommender systems in this domain. Recommendation approach we have utilized in this work elicits explicit ratings as proposed in classical literature of the recommender systems community. This can be a limitation of our approach. Design and development of more transparent form of preference elicitation can be an interesting direction of research to follow. This can be performed by eliciting a more diverse and novel set of user preferences that can better picture the actual needs and constrains of users.

For future work, we plan to re-design our prototype and implement a new version of the system with a novel user interface offering further functionalities. An example of such functionalities could be critiquing and explanation mechanisms where the users could provide feedback on the recommendation and fine-tune it according to their needs and constrains [7].

We also plan to conduct more experiments while collecting a larger dataset that contains different forms of user preferences. We will compare a wider range of recommender algorithms capable of better learning from different sources of user data, e.g., personality traits and social media profiles.

REFERENCES


