

Content-based Recommendation System for Tourists in Paris

1. Introduction

Personal recommendation system has emerged in recent years. For tourists in a strange city, maybe they would go to somewhere popular or find their friends on first days. Since they have visited some places, we could recommend similar or different places to them. For example, if they like somewhere, they tend to explore similar places on the next day, and vice versa.

2. Data description

Boroughs and respective coordinates can be found in '<http://opendata.paris.fr>', and venues information can be found with Foursquare API. Now we have the venues info of every borough and this is called content of boroughs which could help us make predictions. We can simply visualize how the boroughs shape.

3. Methodology

In recommendation field, Collaborative Filtering is a often-used method to predict what a person would like to choose based on their previous preference. For example, we have preference data like this: $\{A:\{a:5, b:1\}, B:\{a:5, c:5\}\}$, then we could recommend item c to person A because both A and B like a, and B like c. However, maybe c is unknown to A so we can assume that A would also like c.

In this report we first build one-hot matrix for every borough based on what venues it has, and the data has been standardized so we don't need to standardize it. Then we calculate the similarity between different boroughs by multiplying their respective vector. The greater the multiplied outcome is, the more similar the boroughs are.

Then we can list top 5 most similar boroughs for each and recommend similar boroughs for every tourist who rated at least one borough. For instance, if A rates 8th Ardt 5, and 9th Ardt 4, then he or she would probably like 14th Ardt because 14th is the most similar one for 8th and 9th.

4. Results

The similarity table of boroughs:

	10ème Ardt	11ème Ardt	12ème Ardt	13ème Ardt	14ème Ardt	15ème Ardt	16ème Ardt	17ème Ardt	18ème Ardt
10ème Ardt		0.0225	0	0.02	0.055	0.0125	0.01	0.025	
11ème Ardt	0.0225		0.01	0.0175	0.025	0.0125	0.025	0.02	
12ème Ardt	0	0.01		0.01	0	0.01	0.04	0	
13ème Ardt	0.02	0.0175	0.01		0.0675	0.0175	0.02	0.0225	
14ème Ardt	0.055	0.025	0	0.0675		0.0575	0.045	0.09	
15ème Ardt	0.0125	0.0125	0.01	0.0175	0.0575		0.025	0.0525	
16ème Ardt	0.01	0.025	0.04	0.02	0.045	0.025		0.025	
17ème Ardt	0.025	0.02	0	0.0225	0.09	0.0525	0.025		
18ème Ardt	0.0275	0.0325	0	0.04	0.095	0.025	0.02	0.035	
19ème Ardt	0.025	0.02	0.01	0.025	0.0525	0.0225	0.01	0.0325	
1er Ardt	0.0125	0.02	0	0	0.0075	0.0175	0.04	0.0325	
20ème Ardt	0.025	0.0175	0.01	0.025	0.085	0.04	0.03	0.055	
2ème Ardt	0.0275	0.0175	0	0.0225	0.085	0.0375	0.025	0.0575	
3ème Ardt	0.015	0.0175	0.01	0.01	0.0275	0.015	0.01	0.03	
4ème Ardt	0.0325	0.02	0	0.0275	0.0675	0.0175	0.02	0.025	
5ème Ardt	0.035	0.02	0	0.0375	0.0975	0.0325	0.04	0.0475	
6ème Ardt	0.02	0.03	0	0.015	0.0575	0.0175	0.025	0.035	
7ème Ardt	0.03	0.025	0.01	0.04	0.1225	0.045	0.055	0.075	
8ème Ardt	0.0275	0.0175	0	0.0375	0.125	0.04	0.03	0.075	
9ème Ardt	0.0325	0.02	0	0.0325	0.095	0.025	0.03	0.0375	

The recommended result for manually made behavior data:

```
prefer_data = {'Eric':{'9ème Ardt':9,'5ème Ardt':1},'Paul':{'2ème Ardt':9,'16ème Ardt':1}}
```

Recommended and scores:

Eric

```
[('14ème Ardt', 0.9525000000000001), ('7ème Ardt', 0.5900000000000001), ('8ème Ardt', 0.535)]
```

Paul

```
[('14ème Ardt', 0.8100000000000002), ('7ème Ardt', 0.6175), ('17ème Ardt', 0.5425000000000001)]
```

5. Discussion

In order to get better recommendation results, we should have more detailed information to build features of each borough, which is called feature engineering, like crime index, population

information and so one.

Besides, the behavior data is unavailable. The more rating data we have, the more accurate our recommendation system would be. And we could also incorporate metrics to evaluate how well the system works.

6. Conclusion

This system would work better than the way we only recommend most popular boroughs to tourists. We could sharpen the system with more detailed information and data.