Capstone Project Report

Yuan Wang

Introduction

"Would you recommend a location in Hong Kong to open a new cinema?"
My boss, the stakeholder wants to **open a new cinema as company's new business**.
He explains that watching movie is a part of whole afternoon or night activities. Cinema should has **many restaurants and shopping places nearby**. Transportation is also an important factor. Customer can walk to cinema within **5 minutes** from **public transport facilities** is perfect.
He wants me concentrated on selection of cinema location according to its nearby environment. Cinema facility and rental price is not my concern. He lists out his **top 10 favorite cinemas** in Hong Kong with rating.
I work with my teammates and select **5 possible locations** to build the cinema. Which location should be suggested to the stakeholder?

Data

According to the question, following data are required.
1. Geographic coordinate of Hong Kong cinemas
I need to **compare 5 possible locations with current cinemas** in Hong Kong. Therefore, I need to find a list of Hong Kong cinema and cinemas' geographic coordinates. Luckily, I can find the list and coordinates from the website https://hkmovie6.com/cinema .
By researching the data,there are 68 cinemas in Hong Kong.And I get first five records o f Hong Kong cinemas

	Name	ChiName	Address	Latitude	Longitude
0	Emperor Cinemas - Entertainment Building	英皇戲院 - 娛樂行	3/F, Emperor Cinemas Entertainment Building, 3...	22.281453	114.154230
1	The Coronet @ Emperor Cinemas - Entertainment ...	The Coronet @ 英皇戲院 - 娛樂行	3/F, Emperor Cinemas Entertainment Building, 3...	22.281453	114.154230
2	Emperor Cinemas - Tuen Mun	英皇戲院 - 屯門新都商場	3/F, New Town Commercial Arcade, 2 Tuen Lee St...	22.390776	113.975983
3	Broadway Circuit - CYBERPORT	百老匯戲院 - 數碼港	Shop L1 - 3, Level 1, The Arcade, 100 Cyberpor...	22.261067	114.129825
4	Broadway Circuit - PALACE IFC	百老匯戲院 - PALACE IFC	Podium L1, IFC Mall, 8 Finance Street, Central	22.285545	114.157979

2. Geographic coordinates of 5 possible cinema addresses

Geographic coordinates of 5 possible cinemas are required and I can use Google Map API to find this information

3. Favorite cinema list of stakeholder
The favorite cinema list of stakeholder is an important information that I can **use it as profile to select the best location**.

4. Eating, Shopping and Public transportation facility around cinema
The recommended cinema location needs to have many eating and shopping venues nearby. Convenient public transport is also required.

These data can be found by using FourSquare API to find these venues around the location. The radius of exploration distance is set to 500 meters, which is about 5 minutes walking distance.

Methodology

With above data, I can use content-based recommendation technique to resolve the problem.

Combine with FourSquare API which provides how many venues in different category of Hong Kong cinemas, a matrix which captured characteristic of venues nearby cinema are built. Stakeholder's favorite list is the profile to combine with the matrix to become a weighted matrix of favorite cinema.

The weighted matrix can be applied on 5 target locations with venues information to generate a ranking result. The the top one on the ranking list can be recommended to the stakeholder.

Before building the matrix, I have to prepare the required data and apply some data analysis.

Most of Hong Kong cinemas and stakeholder's favorite cinemas location are built near main road, and centralized in urban area of Hong Kong. The target locations of new cinema are not near to main road.

Then, let's use **Content-Based** or **Item-Item recommendation systems**. In this case, I am going to try to figure out the boss's favorite new cinema location by counting number of nearby venues and ratings given.

Results

With the boss's profile and the complete list of cinemas and their venues count in hand, I am going to take the weighted average of every locoation based on the profile and recommend the top location that most satisfy it.
At last, I found the location: “Tsuen Fung Centre Shopping Arcade, Tsuen Wan”should be recommended.

Conclusion

The stakeholder's problem is resolved. Stakeholder wants to find the best place to build a new cinema in Hong Kong, and the factors of "best location" is based on the number of venues in eating, shopping, transportation category around the location. Stakeholder also provide his favorite list of cinema to further explain what the "best location" is. Content-based filtering machine learning technique is the most suitable method to resolve the problem. It combines stakeholder's preference and cinema profile to make the recommendation result.

The 5 target locations of new cinema may not be a good choices. As the weighting matrix is developed, I can quickly pick other locations and make the recommendation again.