

Topic Modelling Based Recommender System

Keerthana Jayaprakash, Siddharth Kanojiya, Sneha Bezawada

Motivation

"I am really very hungry, want to visit a nice fast food place but I am not in the mood to read tons of reviews. Dear Zomato, as I've told you what I like and dislike through my reviews, could you please help?"

Goals

Build a system to extract user's preference from his reviews and suggest him similar restaurants.

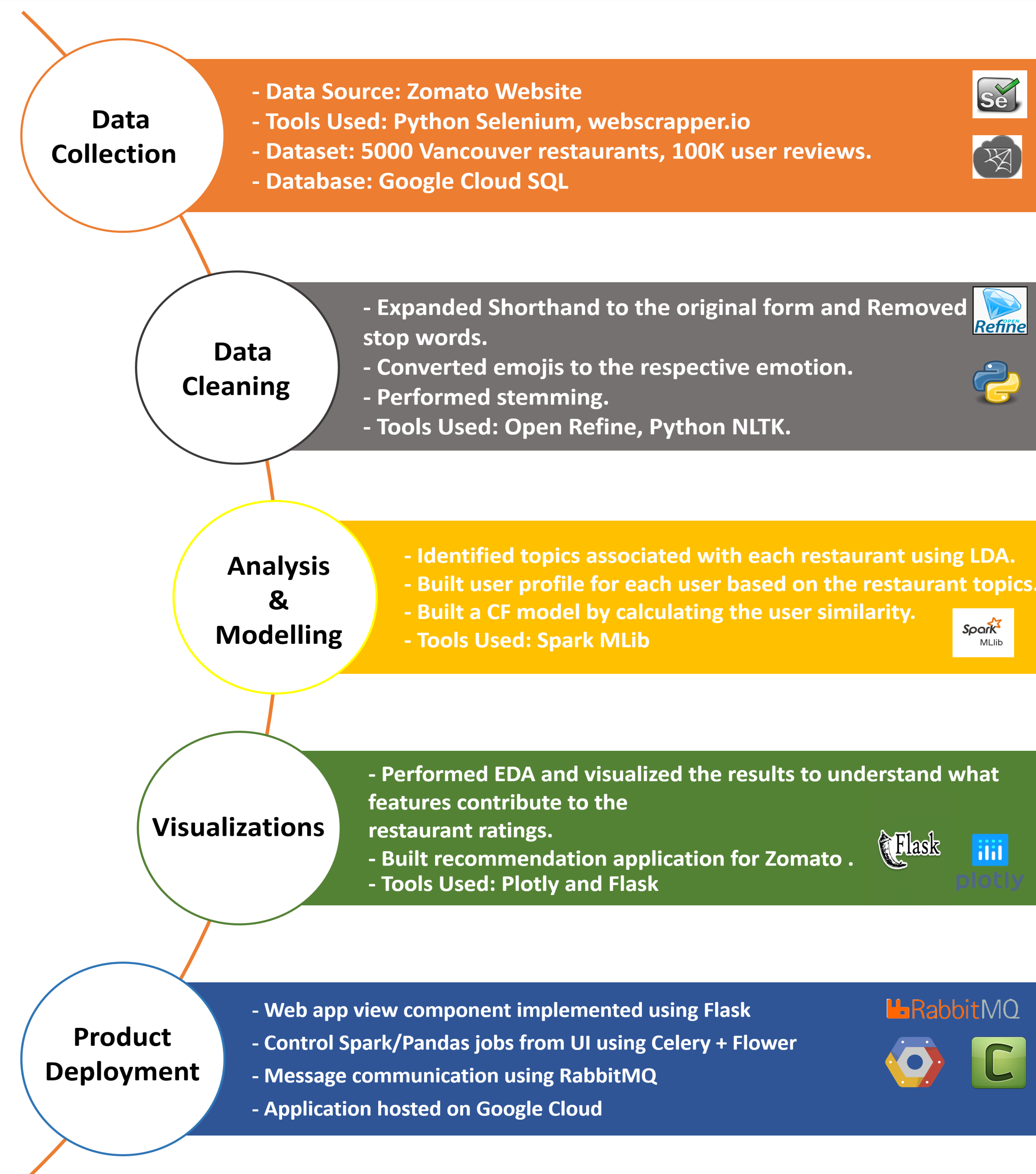
And, find out answers to the following questions:

1. What are the things people really like and dislike in Greater Vancouver area when it comes to food and restaurants?
2. What is the approximate budget to have a good meal from a highly rated restaurant in Vancouver?
3. How busy are the restaurants in Vancouver?
4. Are only international cuisines ruling the restaurant market.
5. If someone wants to start a new restaurant in Vancouver, which cuisines will attract more customers?

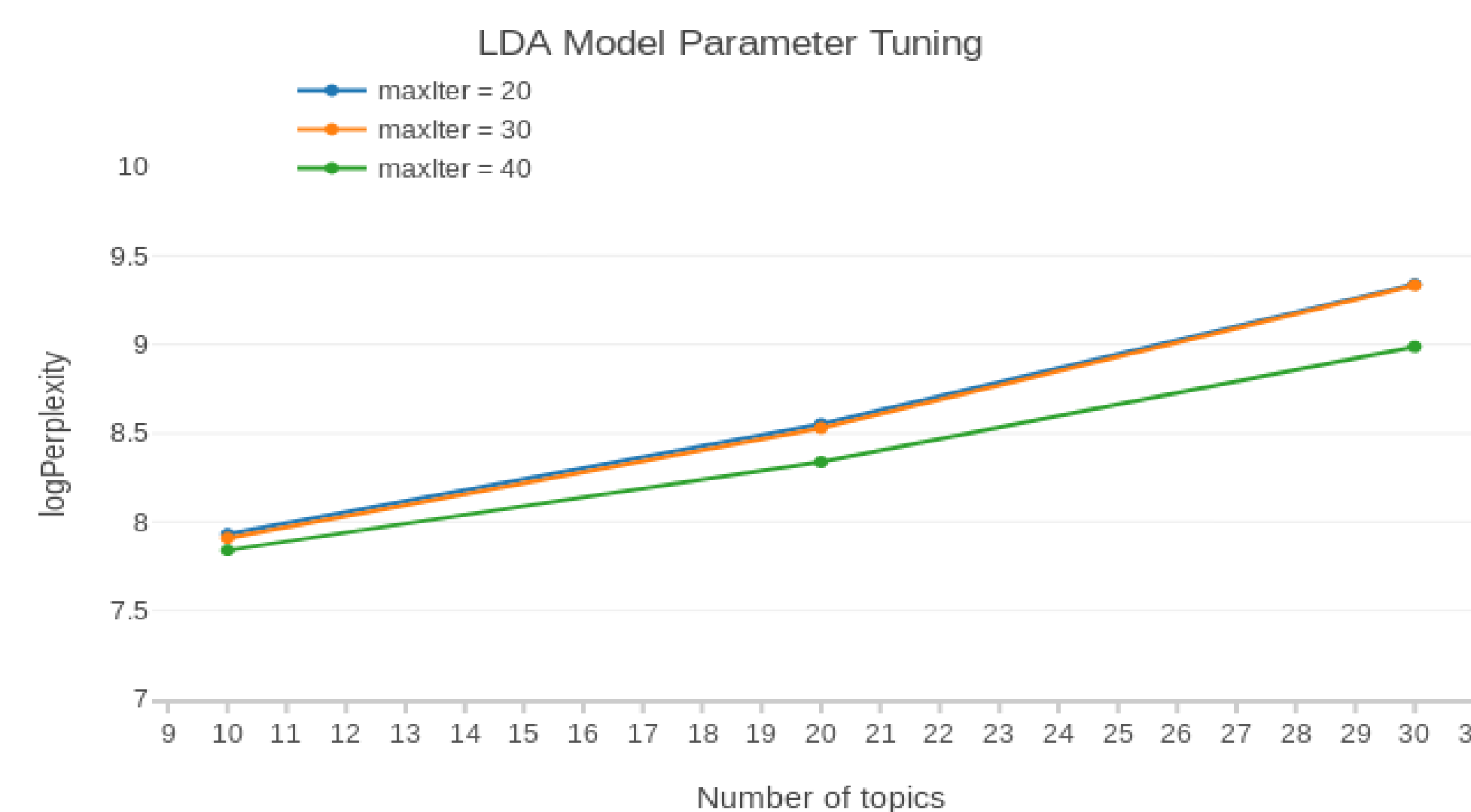
Approach

1. Scrape Vancouver restaurants data from the Zomato Website
2. Perform sentiment analysis on the reviews
3. Identify the most relevant topics for each restaurant using topic modelling technique.
4. Build a collaborative filtering recommender based on topic model and predict recommendations
5. Develop a web interface to emulate a restaurant search portal and show recommendations

Pipeline



Evaluation

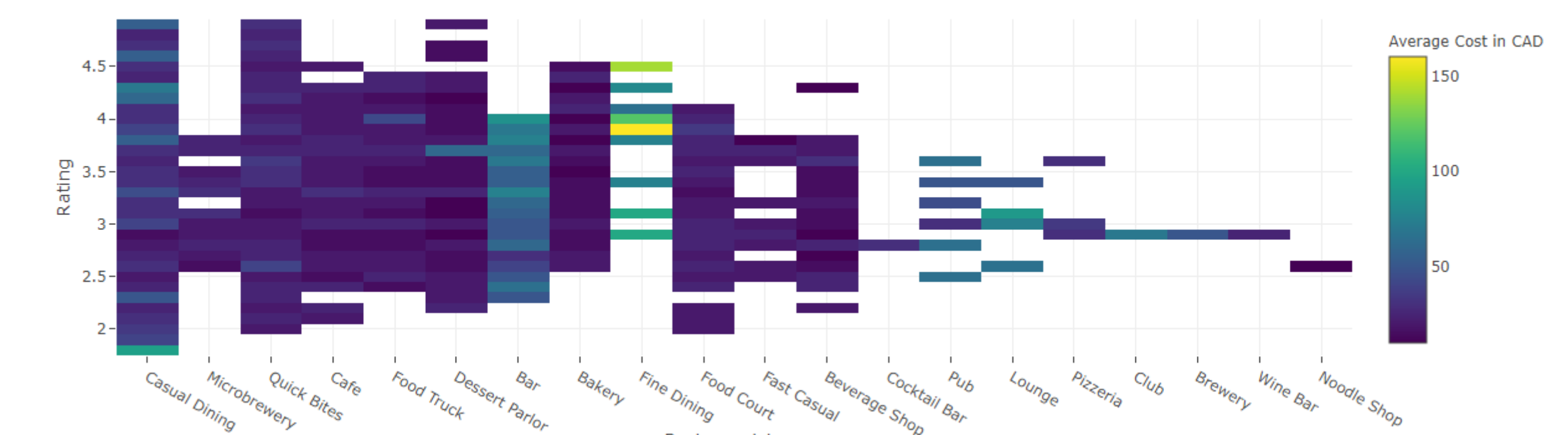


Model Comparison

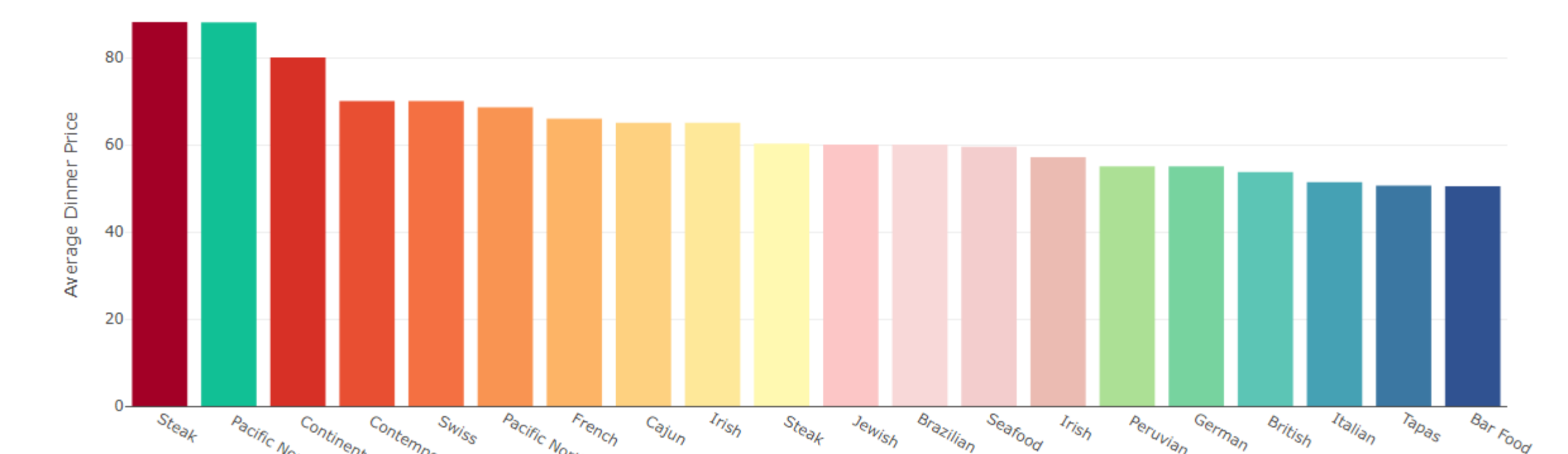
Model	RMSE Value
LDA Based CF	3.636
User Based CF	3.647
Item Based CF	3.650

Visualizations

Restaurant Type Vs Rating & Cost



Restaurant Cuisines Vs Average Dinner Cost



Positive Sentiments



Correlation Between Restaurant Features and Rating



Learnings

- Zomato API allows only 1000 requests per day, too restrictive to collect 5k restaurants and 100k reviews data. Solution - Use standalone Google web scraper and Selenium scripts.
- Recommendations model can't be retrained on the fly as soon as new data arrives. Solution - periodically retrain the model offline as asynchronous jobs

Future Work

Leverage the images posted in reviews and other datasets for example restaurant inspections to improve recommendations.