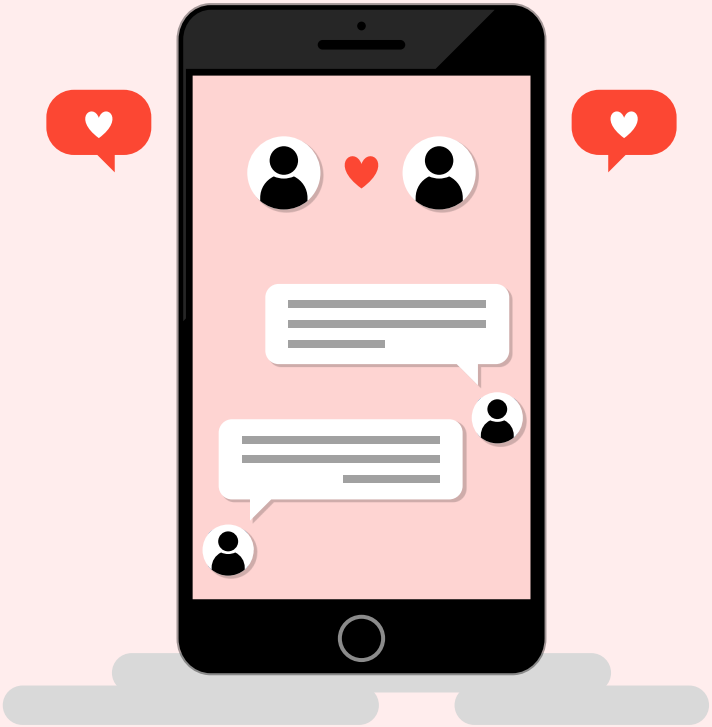


Date-Ah!

Data Driven Restaurant and Activity Planner

Cher Wee Zheng

Table of contents

01

Introduction

Wee Zack

Gender

Man

Age

26

Profession

Unemployed

Traits

Hates sweating	90%
Relationship strength	50%
Knowledge of SG	25%

Strengths

Tech Savvy	90%
Knows what girlfriend likes	90%

Decision Making Process

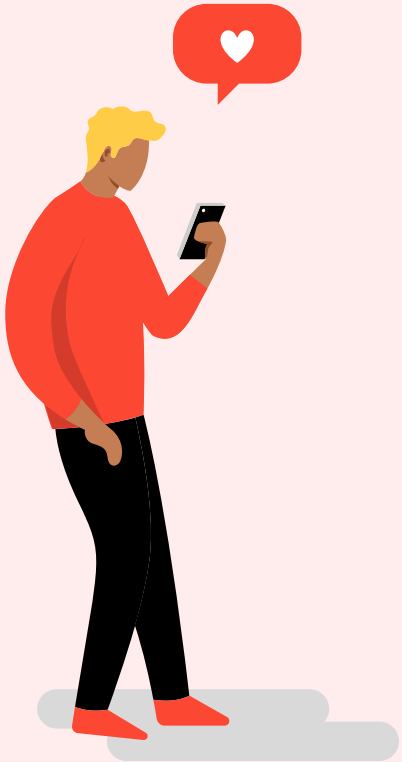
Transport	25%
Please Girlfriend	90%

Wee Zack is a young adult, 26. Recently, he has run out of ideas of where to go and what to do with his girlfriend on dates.

Having had many arguments with his girlfriend recently, he wants to improve their relationship by bringing her out on an enjoyable date, where they spend time doing things she loves.

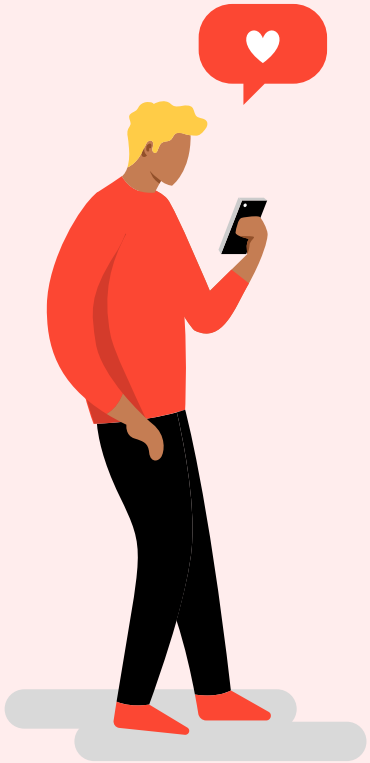
However, he is lazy to do his own research on all the different activities and restaurants and plan the day's itinerary, and he wishes there was an app or website to help him to do this.

What should we eat?



What should we do?

Problem Statement



How might we help individuals to quickly plan their date, ensuring the best experience possible, taking into account their budget, location, food preferences and activity preferences?

Key Objectives

1. Recommend Restaurants and Activities
2. Based on User's inputted preferences
3. Minimise distance

Development Process

Scraping

1. Yelp
2. Google Places
3. TheSmartLocal
4. SecretSingapore

Modelling

1. Creating Dummy Users and interaction data
2. Content-based Filtering
3. Collaborative Filtering
4. Weighted Scoring

Implementation

Working together with dating apps and you to improve our model

Scraping Data

1. Restaurants
    a. Yelp
    b. Google Places
    c. Features:
        i. Cuisine: Chinese, Indian, Malaysian, Western, Indonesian, French, Japanese, Mexican, Middle Eastern, Taiwanese, Vietnamese, Fusion, Italian, Singaporean
        ii. Rating
        iii. Price range $$
        iv. location

2. Activities
    a. TheSmartLocal
    b. Secret Singapore
    c. Features:
        i. Activity Type: Indoor, Outdoor, Animals, Games, Thrill, Wet, Exercise, Nature, Enrichment, Art, Culture, Crafts, Food
        ii. Location

02

Modelling

Matrix Factorisation

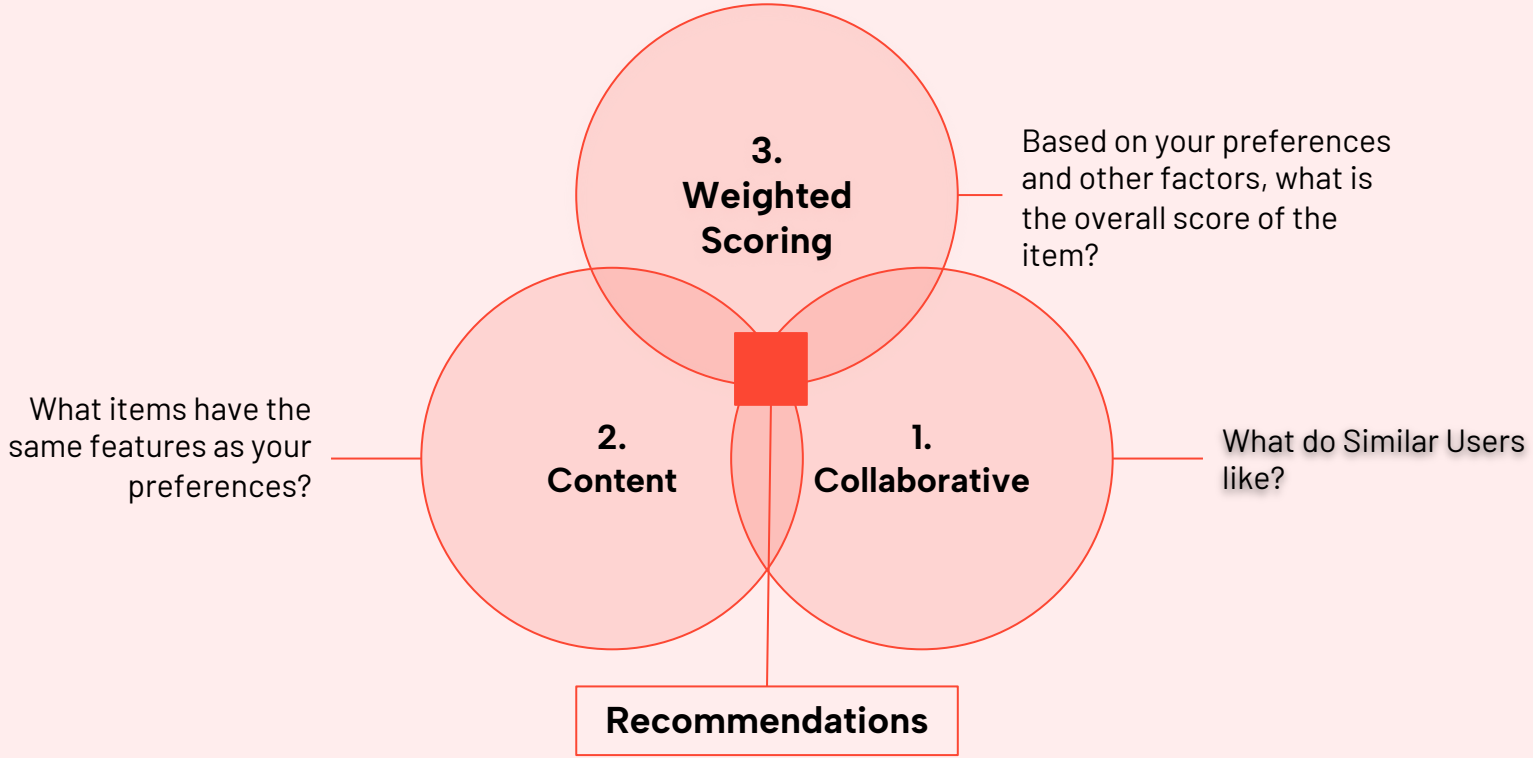
1. Identify latent factors which we cannot tell
2. Decompose a matrix into a product of multiple matrices
3. Often used for tasks such as dimensionality reduction, collaborative filtering, and recommendation systems
4. Algorithms:
    - Alternating Least Squares (ALS)
    - Singular Vector Decomposition (SVD)
    - Singular Vector Decomposition ++ (SVDpp)
    - Non-negative Matrix Factorisation (NMF)
    - SlopeOne

Accuracy Scores

	RMSE	MAE
SVD	1.531	1.305
NMF	1.459	1.26
SlopeOne	1.425	1.228
ALS	1.464	1.27

Rating range: (0 - 4)

Hybrid Model



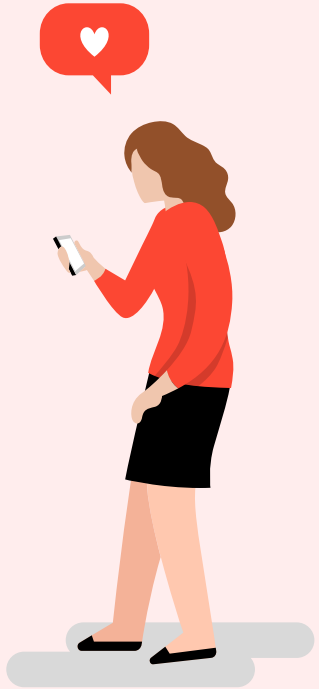
User Inputs

1. How important is Budget, Food, Activity to your date experience?
    a. On a scale of 1 (most important) to 3 (least important)

2. On a scale of 0 to 4, how much do you like (cuisine)?
    a. 0 having no interest and 4 having the most interest

1. How much do you like (activity type)?
    a. 0 to 2, with 0 having no interest, 1 being neutral, 2 having strong interest

1. Postal code



1st Layer: Collaborative Filtering

What do Similar Users like?

1. Generate user profile based on inputs

2. Extract similar users
    a. Cosine similarity

3. Returning the top 50 items for restaurants and activity

2nd Layer: Content-based Filtering

What items have the same features as your preferences?

1. Generate user profile based on inputs

2. Extract similar items
    a. Feature matches

3. Returning the top 50 items for restaurants for activity

3rd Layer: Weighted Scoring

1. Filter the two sets of recommendations, keeping 20 overlaps
    a. If less than 20, populate with highest scoring items from content based
2. Calculate distance to user location for each of these items
3. Restaurant Weights = {3=0.2, 2=0.3, 1=0.4}, Distance Weight = 1 - Budget Weight - Food Weight

    Score = distance weight * 1/distance + food weight*rating + budget weight* price
4. Weights = {1=0.5, 2=0.3, 3 =0.1}, Distance Weight = 1 - Activity Weight

    Score = activity weight*activity score + distance weight*1/distance

    Activity Score = cumulative of inputted activity type score
5. The top 5 items with the best overall scores are then recommended

Wee Zack's Profile

budget	cuisine	activity
1	3	1

Western	Chinese	italian	Vietnamese	Japanese	Singaporean	Mexican	Indian	Malaysian	Middle Eastern	Indonesian	Spanish	Taiwanese	French	Other
0	2	3	1	4	1	3	0	1	2	1	3	0	4	3

indoor	outdoor	animals	wet	enrichment	food	history	thrill	culture	nature	exercise	crafts	games
2	0	2	0	0	0	0	1	0	1	1	2	2

Postal = 098654

App Demo

Wee Zack's Profile

budget	cuisine	activity
1	3	1

Western	Chinese	italian	Vietnamese	Japanese	Singaporean	Mexican	Indian	Malaysian	Middle Eastern	Indonesian	Spanish	Taiwanese	French	Other
0	2	3	1	4	1	3	0	1	2	1	3	0	4	3

indoor	outdoor	animals	wet	enrichment	food	history	thrill	culture	nature	exercise	crafts	games
2	0	2	0	0	0	0	1	0	1	1	2	2

Postal = 098654

Restaurants

Recommended Restaurants:

	name	cuisine category	price	distance	location
251	Hoàng Hâu - Vietnamese Restaurant	Vietnamese	1	2.2662	146 01, 1090 Jalan Bukit Merah, E
181	Jian Bo Shui Kueh	Singaporean	1	2.6028	Tiong Bahru
138	Su Korean BBQ	Other	1	8.9251	Toa Payoh
210	Crystal Jade Kitchen	Other	1	5.0828	Dhoby Ghaut
143	Hup Choon Eating House	Other	1	7.1417	Bukit Timah

Western	Chinese	italian	Vietnamese	Japanese	Singaporean	Mexican	Indian	Malaysian	Middle Eastern	Indonesian	Spanish	Taiwanese	French	Other
0	2	3	1	4	1	3	0	1	2	1	3	0	4	3

Activities

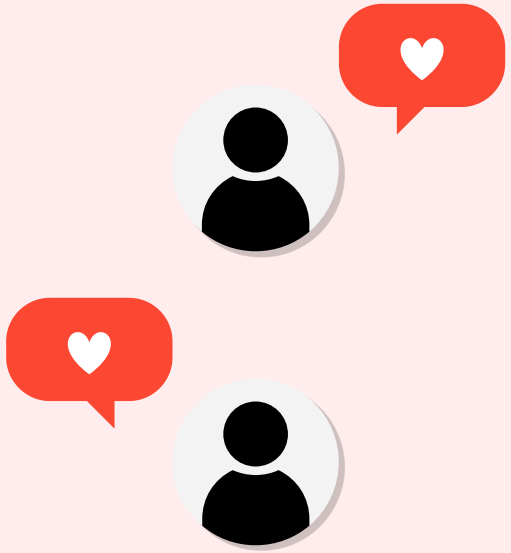
Recommended Activities:

	activity	distance	type	location
145	Dine at Singapore's only underwater restaurant	0.8305	indoor, animals, food	Equarius Hotel, B1M
0	See the S.E.A. Aquarium	1.249	indoor, animals	8 Sentosa Gateway, S
48	Go prawning at FishBugis+	5.8545	indoor, animals	Bugis+, 201 Victoria S
11	Support the Kitten Sanctuary	6.6085	indoor, animals	781A North Bridge Rc
75	Dine at Soul Coffee	10.4908	indoor, animals, food	11 Tanjong Katong Rc

indoor	outdoor	animals	wet	enrichment	food	history	thrill	culture	nature	exercise	crafts	games
2	0	2	0	0	0	0	1	0	1	1	2	2

03

Evaluation



Limitations

1. As of now, the model is able to generate recommendations which meet most of the user's inputted preferences.
2. Limited in database of both items and user interaction.
3. Collection of data is difficult due to formatting of data on different sites and missing information such as location.
4. Over time, the database can grow in size and more items can be added, providing more options and customisation for users.

Recommendations

1. With sufficient interaction data, matrix factorisation techniques can be implemented
    a. latent factors
    b. fill in missing data

1. Collaborative filtering to play a bigger role in the future
    a. Introduction of return users and recommendations based on predicted rating

1. Travel time could be included

2. Collaboration with dating apps to include recommender as a feature into their app

THANK YOU