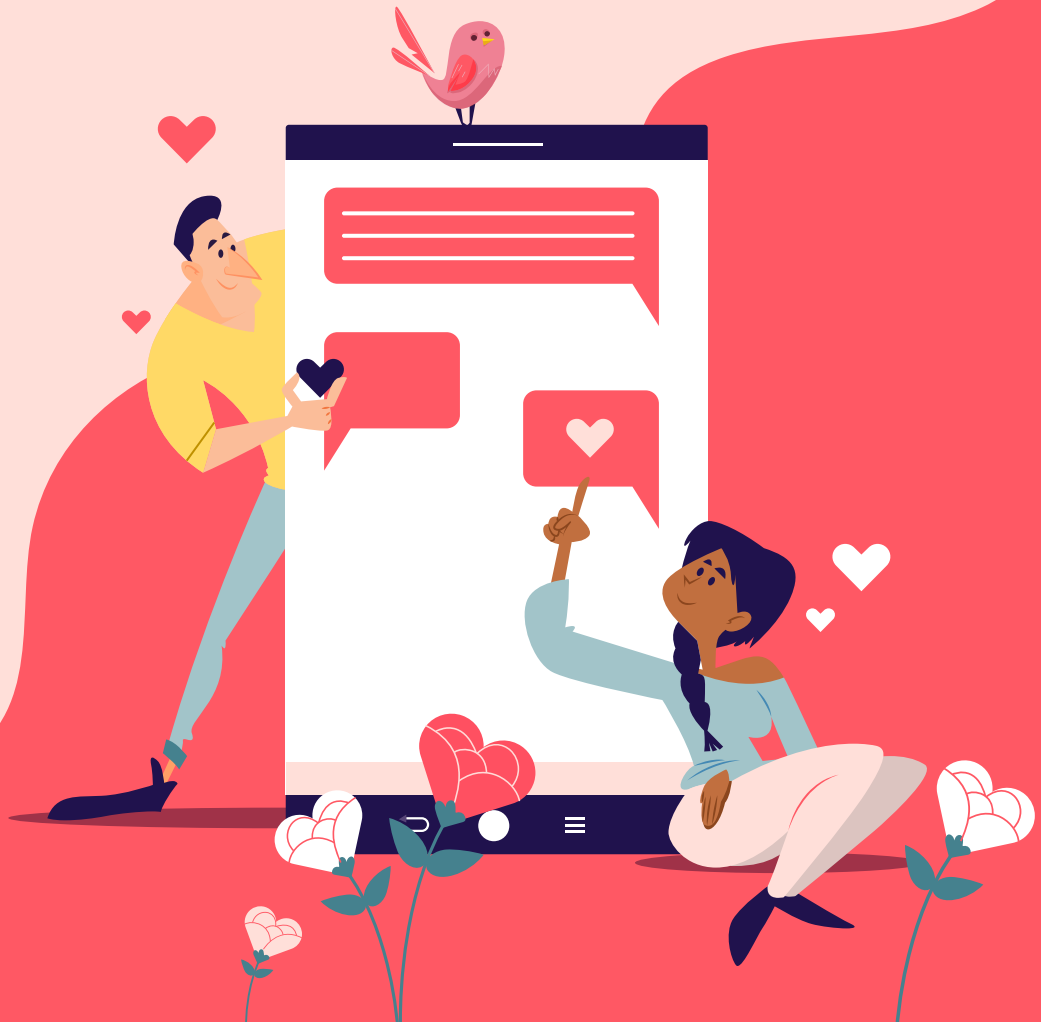
DATING APP FOR FOOD LOVERS

Quan Nguyen




Foodstruck

OBJECTIVES

"MATCH" USERS

Using Content-based Recommender System and Natural Language Processing

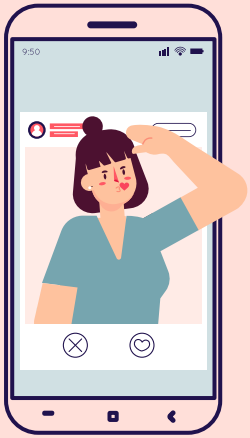
RECOMMEND RESTAURANTS

Using Collaborative Filtering Recommendation System

IMPROVE ACCURACY

Using Neural Network Embeddings

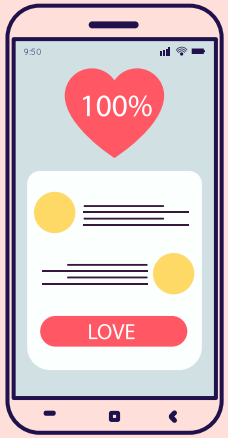
YELP OPEN DATASETS



REVIEWS

6,685,900 reviews

Features: *text, ratings, user_id, business_id*



BUSINESSES

192,609 businesses

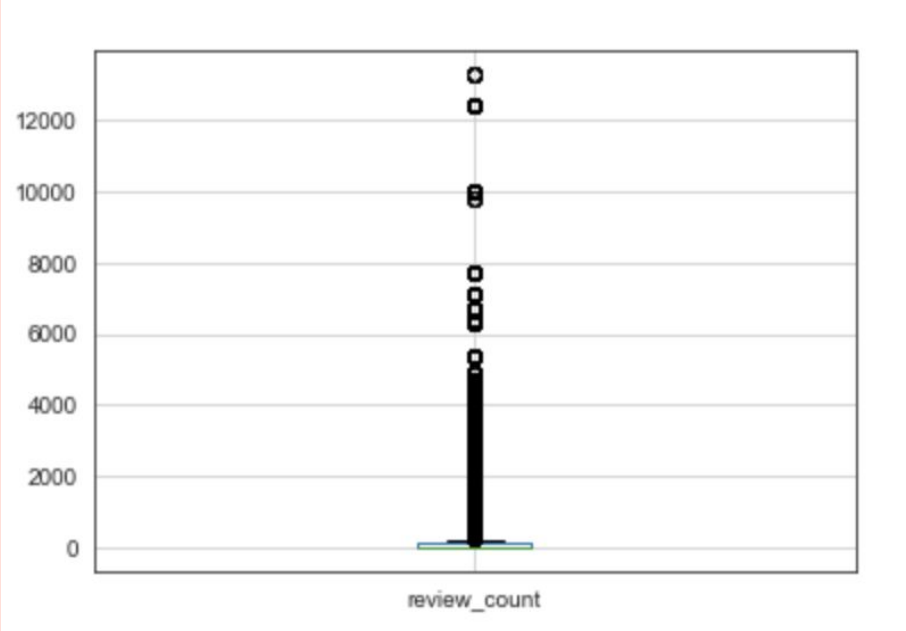
Features: *review_count, user_id*

USERS

1,637,138 users

Features: *city, business_id*

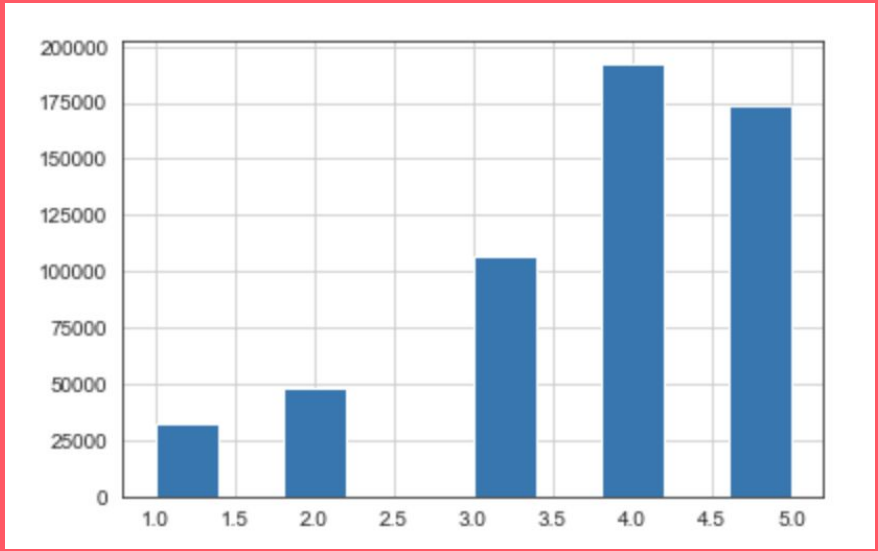
EDA AND DATA PROCESSING



- Max value: 13,000 reviews
- Min value: 1 reviews
- Mean ~ 100 reviews

- Get only businesses in Las Vegas
- Users who have more than 60 and less than 120 reviews

EDA AND DATA PROCESSING



Users Rating after getting only restaurants in Las Vegas and users who have more than 60 reviews and less than 120 reviews

After pre processing:

Users: 31,000

Restaurants: 18,000

Reviews: 700,000

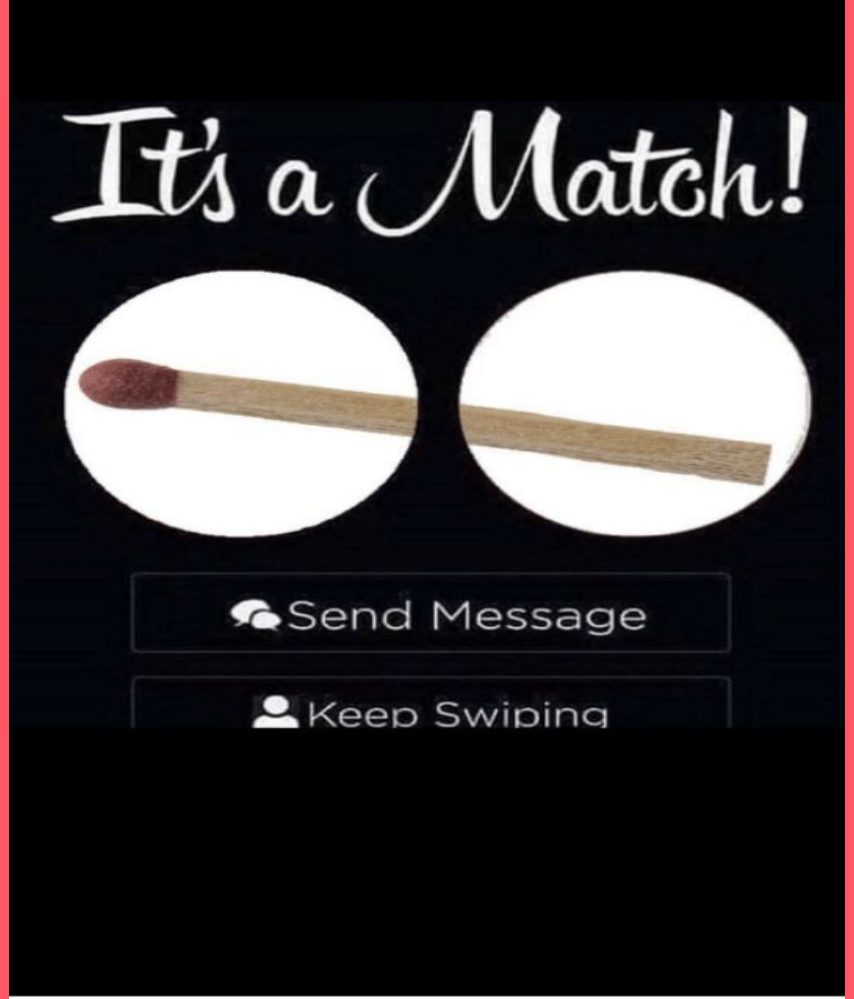
01. "MATCH" USER

NATURAL LANGUAGE PROCESSING

- Create Bag of Words using user's reviews
- TF-IDF Vectorize the BoWs
- Compute Cosine Similarity

CONTENT-BASED RECOMMENDER SYSTEM

Using Cosine Similarity to group users together



Perfect match

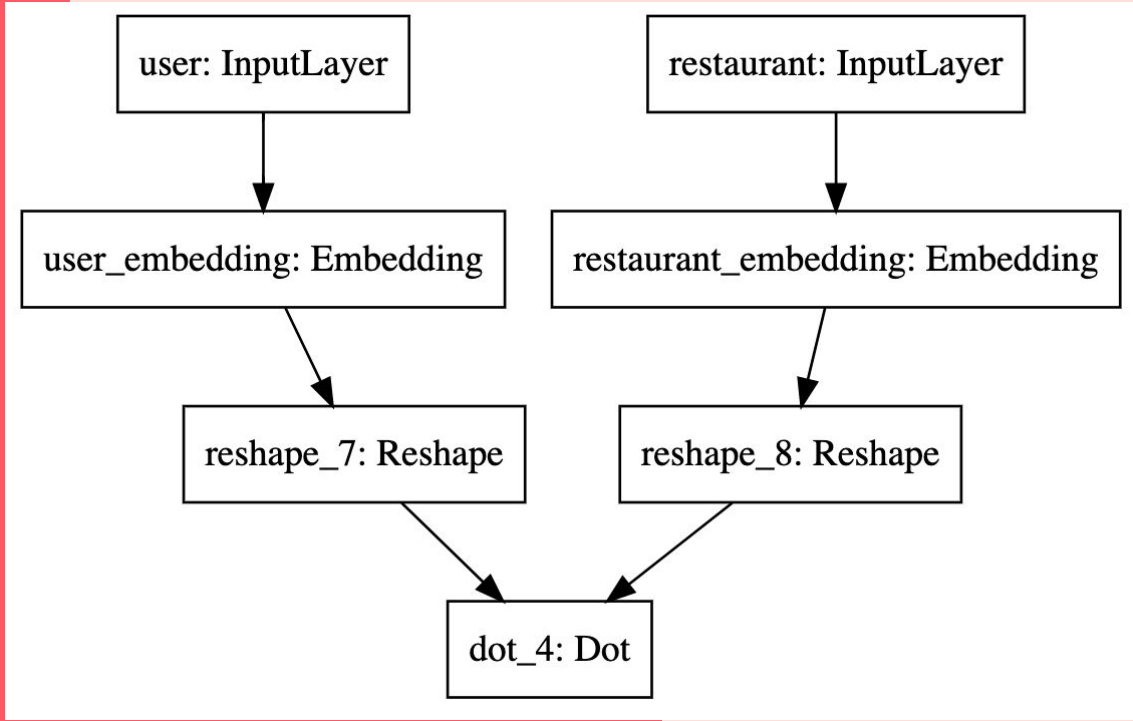
2. RECOMMENDING RESTAURANTS TO THE USERS

COLLABORATIVE FILTERING

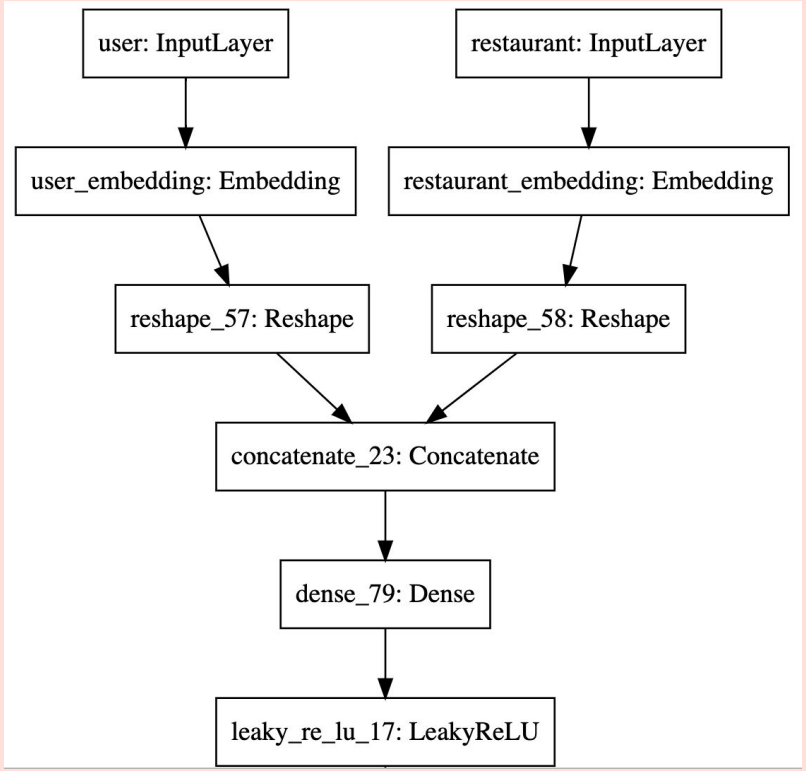
1. KNN Basic:
    - Cosine similarity: 1.72
    - Pearson Correlation: 1.71
2. KNN Baseline:
    - Cosine similarity: 1.61
    - Pearson Correlation: 1.63
3. Singular Value Decomposition : 1.55

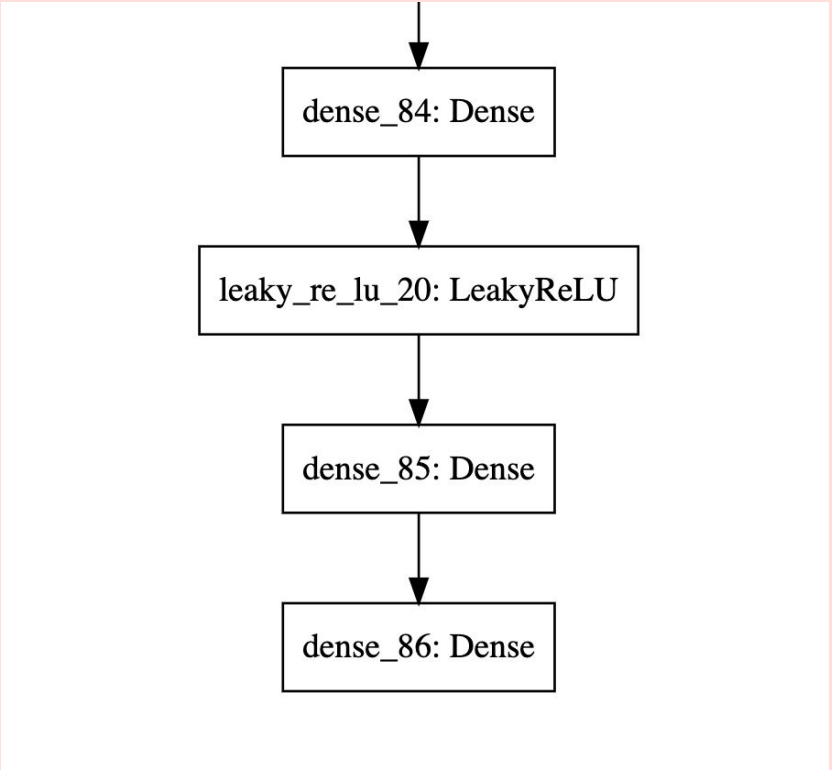
3. NEURAL NETWORKS EMBEDDINGS

MATRIX FACTORIZATION

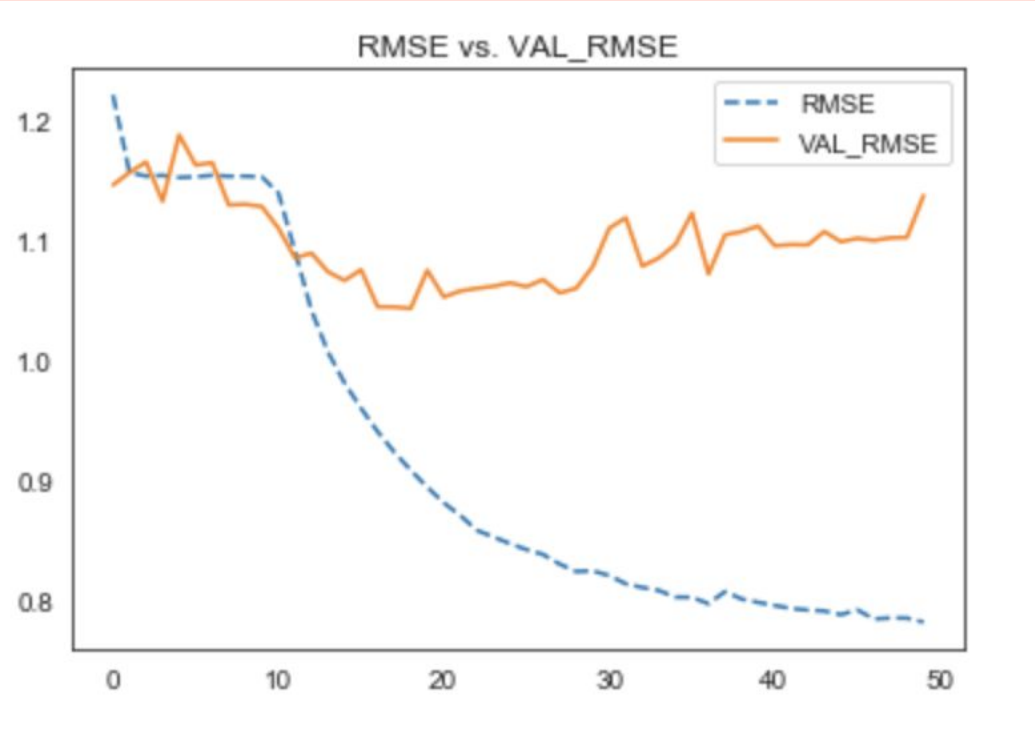


ADDING LAYERS





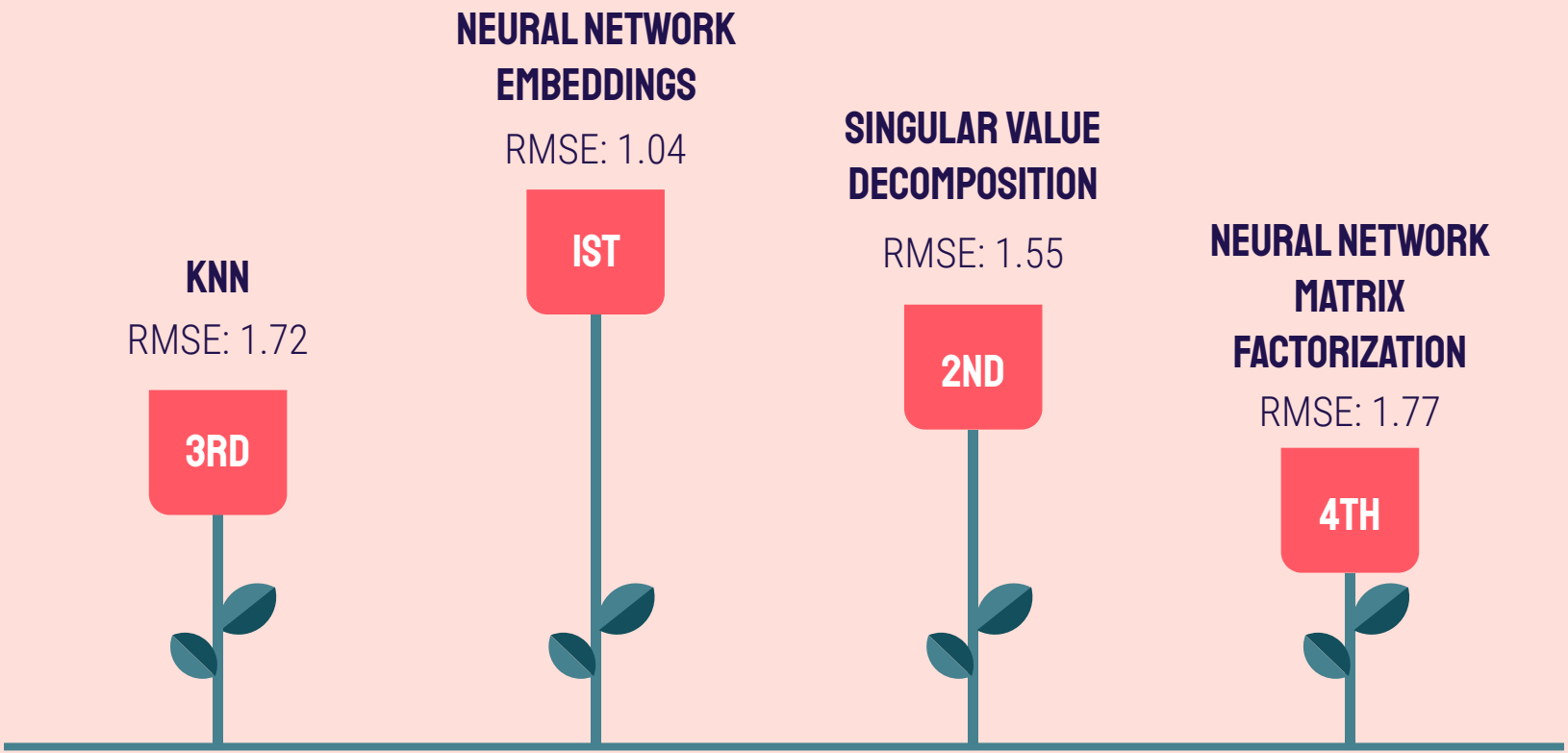
NEURAL NETWORK EMBEDDINGS



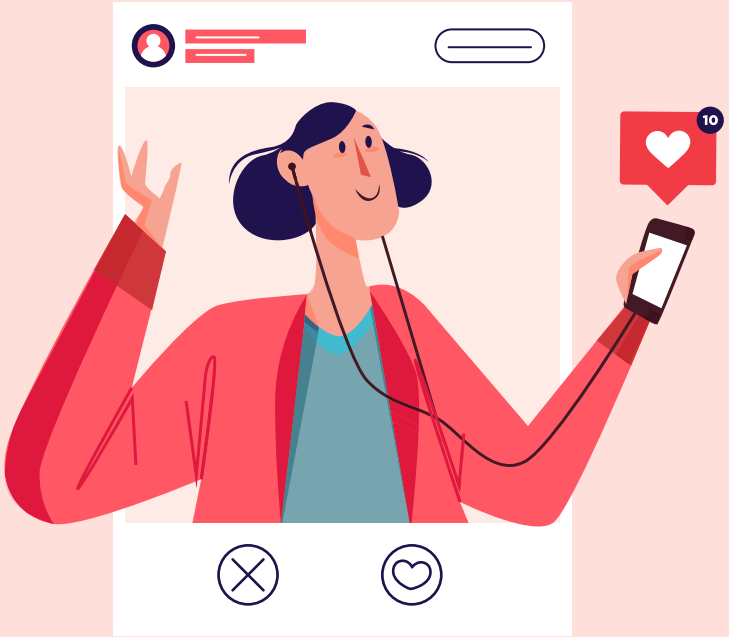
- Best model:
  - Optimizer: Adamax
  - With LeakyReLU Layers
  - Embedding size: 10
  - Lowest RMSE: 1.04

RESULTS

- By using Neural Networks, accuracy improves by 0.7 (RMSE from 1.72 to 1.04)



EXAMPLES



'GC-sitwrqrMkqYj-oTMRkA' and '7NyPj6trSAv0OJKyzdTuGQ' :

- Similarity Score: 0.127
- Recommendations:
    1. Kung Fu Thai & Chinese Restaurant
    2. Joyful House Chinese Cuisine
    3. Shang Artisan Noodles

'l6BmjZMeQD3rDxWUbiAiow' and 'LKd_-BT2GMWp8_1ho18kOg':

- Similarity Score: 0.145
- Recommendations:
    1. The Buffet at Bellagio
    2. Le Village
    3. Paradise Buffet and Cafe

INSIGHTS

- Users with higher number of reviews and longer reviews seem to have more matches.
- A few sentences about a restaurant on Yelp are not enough to show if 2 people are actually similar.

WHAT'S NEXT?

- Add Metadata
- More Data using AWS and PySpark
- Look into 5-star ratings bias

THANKS!