

RESTAURANT RECOMMENDER SYSTEM

BRAINSTATION CAPSTONE

Progress Standup

Diane Lu





PROBLEM STATEMENT

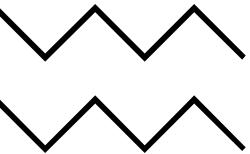
“ Craft a **user-friendly** restaurant recommendation system that understands individual preferences, resulting in **delightful dining experiences** and stronger customer loyalty.



VALUE IN RESTAURANTS

- Personalized **dining suggestions** based on individual preferences
- Potential to **attract new patrons** through targeted recommendations
- Improved **user engagement** and exploration of new dining options
- Valuable **insights** into customer behavior and dining preferences





DATA COLLECTION

Data Source

- Data sourced from Yelp, a renowned platform for business reviews and recommendations
- Our analysis will primarily utilize datasets related to Yelp's businesses, reviews, and users to gain valuable insights

Data Collection

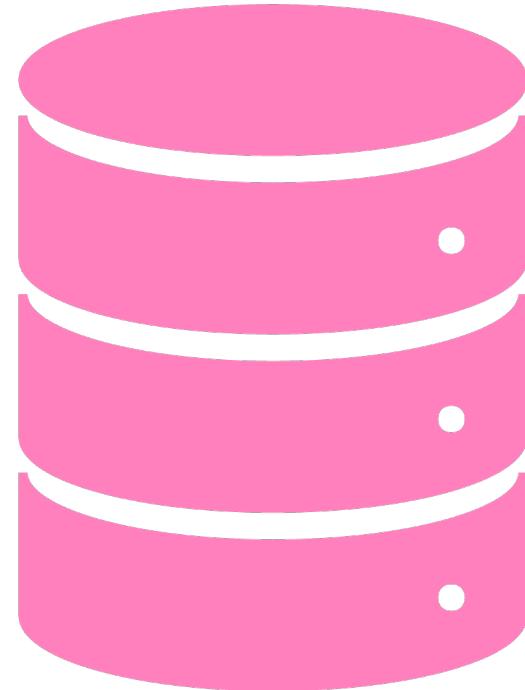
- Data collection involved sourcing datasets from Kaggle, which were provided directly by Yelp.
- This approach ensures reliability and comprehensiveness in deriving valuable insights for the project





DATA DESCRIPTION

- The **Business** dataset:
 - Information on businesses including names, addresses, locations, ratings, categories, and other attributes
- The **User** dataset
 - Information including unique users, user's first names, and average ratings of all reviews
- The **Review** dataset
 - Information on unique reviews and users, star ratings, review text



USER-BASED INFORMATION



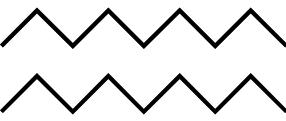
Goal: Provide personalized restaurant suggestions to enhance the user's dining experience



Features: Ratings or interactions given by users to different restaurants



Target: Recommending restaurants that a specific user might like based on their past interactions and preferences



CLEANING THE DATASETS



Data cleaning steps completed for **Business**, **Users**, and **Reviews** datasets

- The **non-restaurant entries** have been removed from all three datasets during the initial data cleaning steps
- Handled **missing values** through appropriate imputation or removal
- Identified and **eliminated duplicate entries** to ensure data integrity

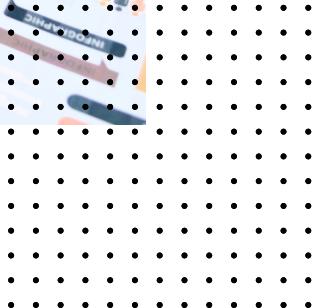
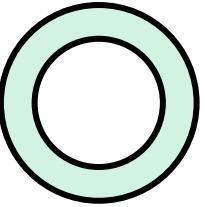
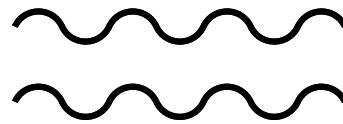
OVERVIEW OF DATASETS

The **Business dataset**: 50,764 entries with 12 columns

The **User dataset**: 2,189,457 entries with 3 columns

The **Review dataset**: 8,635,400 entries with 5 columns

The **Final dataset**: 5,574,714 entries with 18 columns

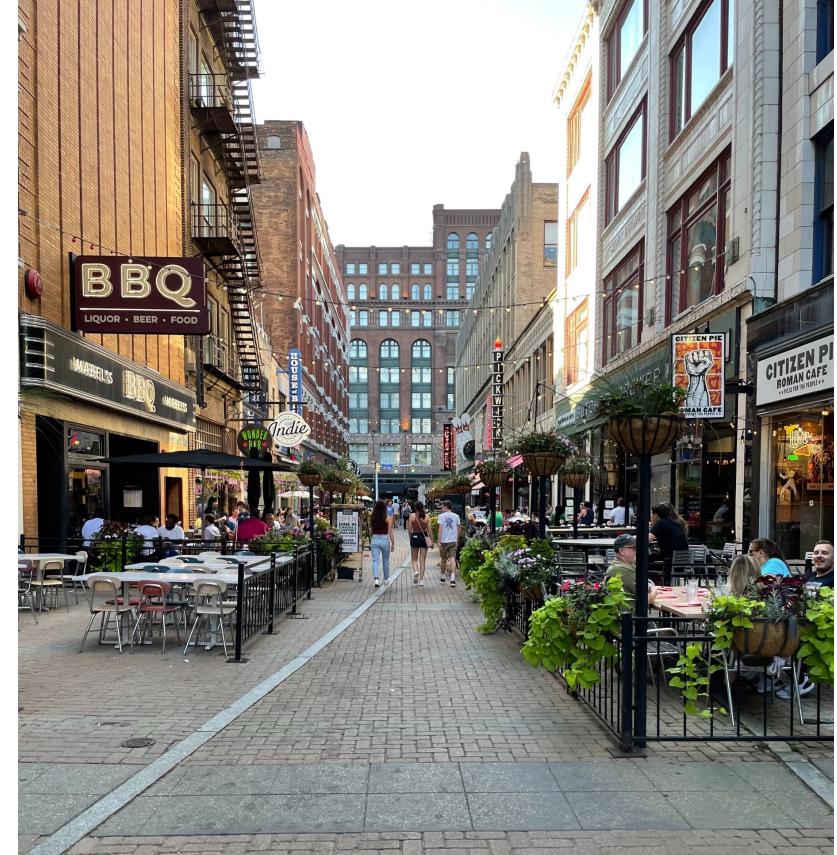




MACHINE LEARNING MODEL

User-based Collaborative Filtering Recommender System

- User-based collaborative filtering model provided **personalized restaurant suggestions** based on user-item interactions
- Initial evaluations showed that the **model recommended restaurants** aligned with users' tastes, enhancing the overall dining experience





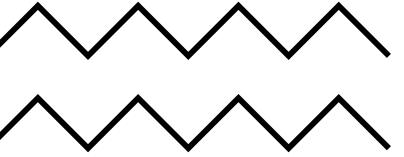
CASE EXAMPLE

Recommendation based on the restaurant **Suika**

- Recommended Restaurants:
 - Kishimoto Japanese Kitchen
 - Hokkaido Ramen Santouka
 - Phnom Penh
 - Kingyo



PLAN OF ACTION



Further analysis and evaluation to enhance model accuracy and performance



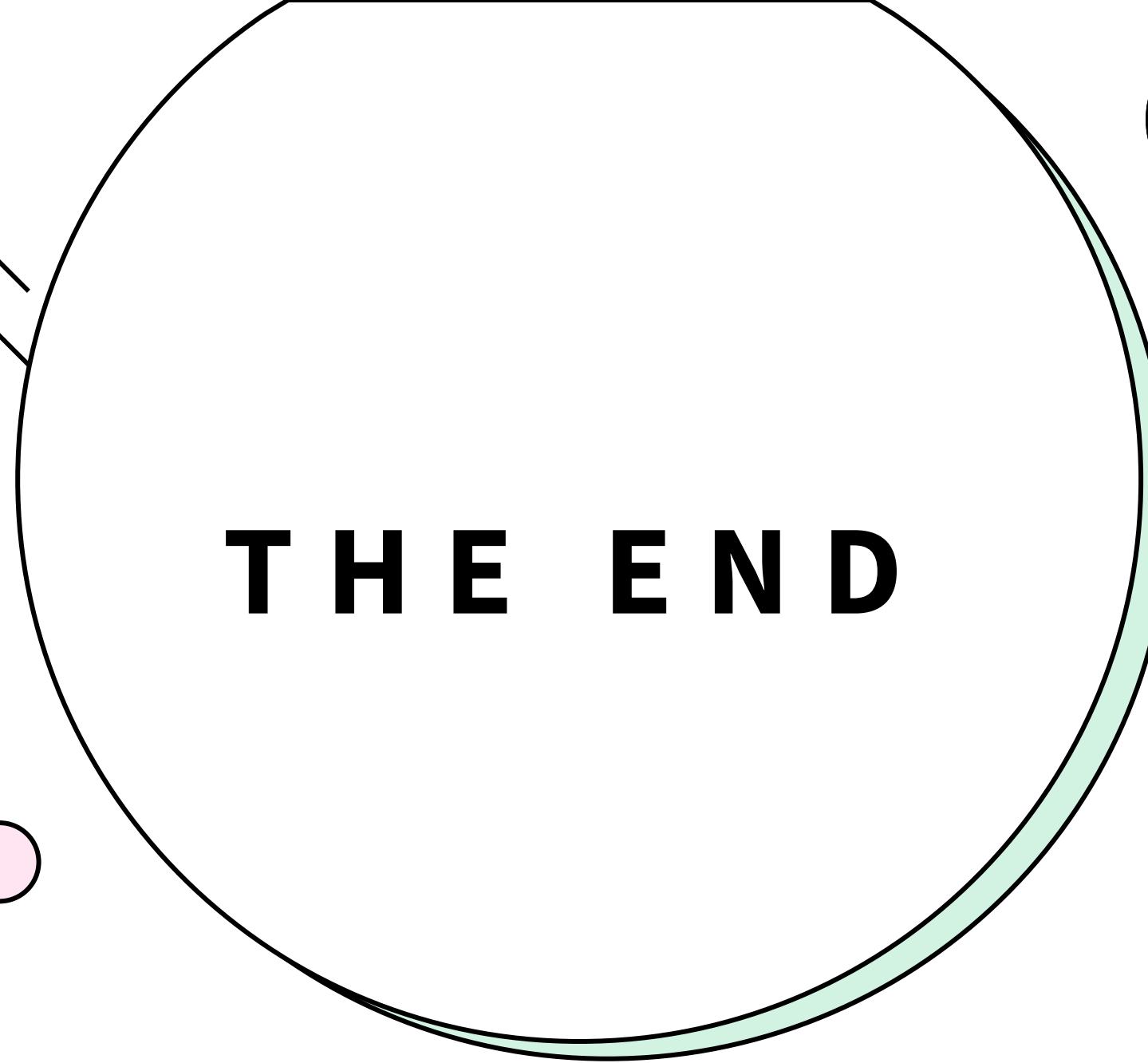
Utilize Singular Value Decomposition (SVD) for handling sparse data and missing values in the user-item matrix



Train the model on the training set and evaluate using metrics based on the User-Item Interaction Matrix



Fine-tune the number of latent features and hyperparameters to optimize the collaborative filtering-based recommendation system



THE END

