

Team - NLP Ninjas  
CS 410 Group Project

# CS 410 Group Project





# Team Members

- Rohit Pandey (Captain) (NetID - rp23)
- Prasad Gole (NetID - gole2)
- Asmita Chihnara (NetID - asmitac2)
- Avinash Pathak (NetID - apathak3)
- Surya Sindhu Mallimanugula (NetID - ssm13)



# Agenda

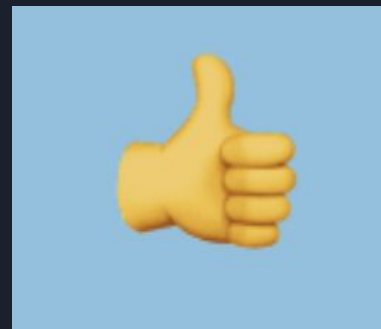
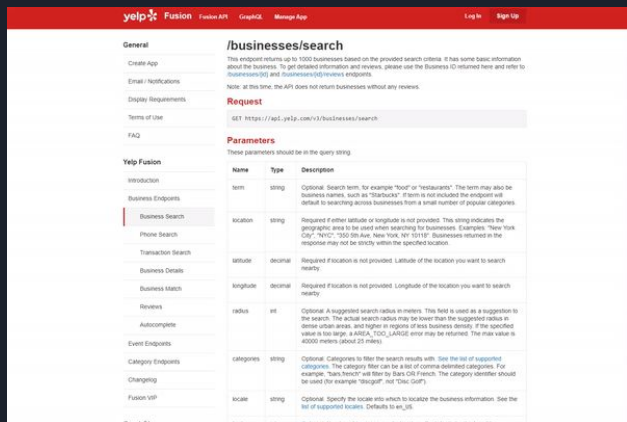
- Why this project?
- Software Implementation
- How to run?
- Demo
- Results
- Future Scope



# Why this project?

1. Addressing a Need: The inception of MapMingle was driven by the growing demand for personalized and dynamic location-based services. In an era where convenience and customization are paramount, there's a clear need for a system that dynamically recommends nearby places based on user preferences and location.
  2. Enhancing User Experience: Traditional location-based services offer static and often irrelevant suggestions. MapMingle is designed to revolutionize this by providing real-time, tailored recommendations, enhancing user engagement and satisfaction.
  3. Leveraging Technology: With advancements in spatial filtering, ranking algorithms, and mobile technology, the opportunity to create a more intelligent and user-friendly recommender system has never been more feasible.
  4. Filling the Gap: Despite the abundance of location-based services, there's a gap in the market for a system that intelligently combines user preferences, location data, and real-time adaptability. MapMingle aims to fill this gap.
  5. Impact: The project has the potential to significantly impact how individuals interact with their surroundings, making the discovery of new places more intuitive, personalized, and enjoyable.
- Innovation and Learning: MapMingle represents an opportunity for our team to innovate and learn, applying cutting-edge technologies in data filtering and ranking algorithms to solve real-world problems.

# Source of Data





# Software Implementation

## Technologies and Frameworks:

- Utilizes Python for backend development, incorporating libraries like `geopy` for spatial calculations and `rank_bm25` for the ranking algorithm.
- Employs modern frontend technologies (React or Angular) to create an interactive and intuitive user interface.

## Data Acquisition and Preprocessing:

- Leverages a comprehensive Yelp dataset, integrating latitude and longitude coordinates for spatial relevance.
- Implements preprocessing steps, including data cleaning, tokenization, and spatial filtering to ensure data quality and usability.

## Ranking Algorithm Development:

- Adopts the BM25Okapi model from the Python library 'rank\_bm25' for ranking places based on user queries and relevance to categories.
- Incorporates a geospatial filtering function using 'geopy' to provide recommendations within a user-specified radius.

## Backend and Frontend Development:

- Backend development focuses on data handling, processing, and response to user queries.
- Frontend development is geared towards creating a user-friendly interface that allows easy interaction and displays recommendations effectively.

## Integration and Testing:

- Engages in ongoing efforts to integrate the frontend UI with backend services for real-time data exchange and interaction.
- Implements comprehensive testing phases to ensure system reliability, efficiency, and optimal user experience.

# Analysis

## 1. DATA

Total Reviews

453,953

Total Restaurants

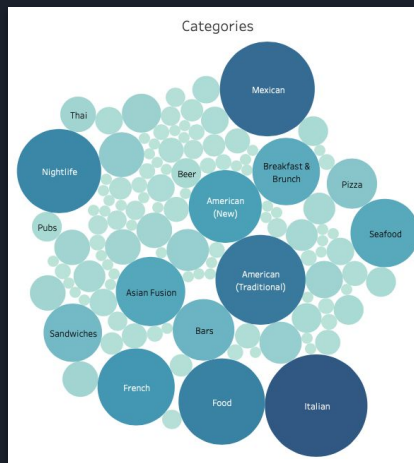
1,000

Total Categories

123

Robust Data Set

## 2. CATEGORIES



7 main categories

## 3. \$ VS RATING



Intermediate price  
with higher reviews

# Product Demo

## Recommendations from MapMingle!

### United States

Follow the instructions:

Let us know what you're looking for!

Restaurants

Where are you now? Select your location in the map and the values will be displayed automatically.

Latitude

39.95421110943039

Longitude

-75.18668522117144

Within Distance (in KM)

10

Minimum rating

2

Submit

Refresh Page

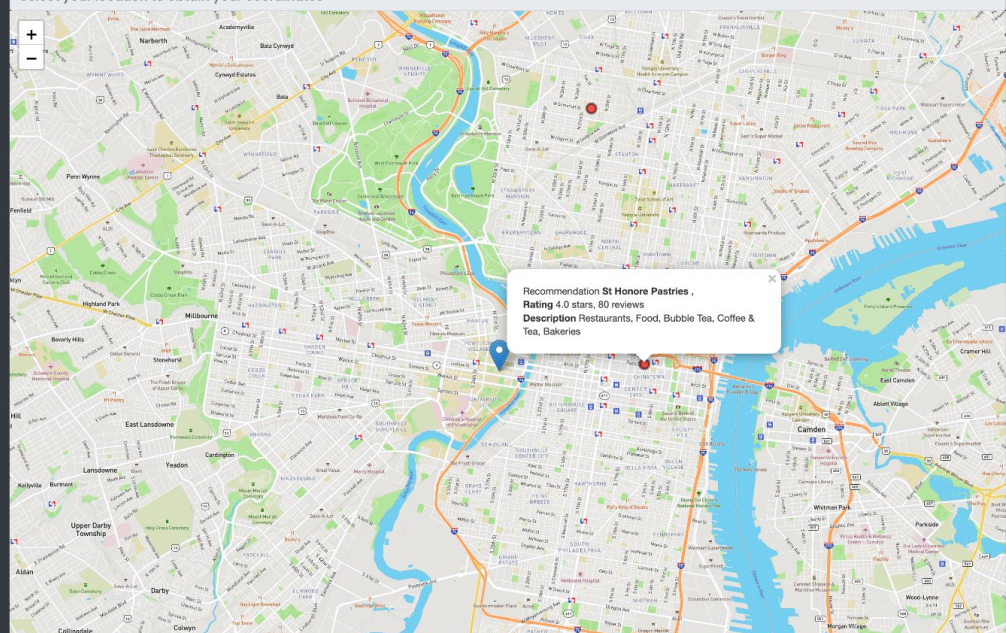
### Gpt Summary

St Honore Pastries is the perfect spot for any food lover. With a 4.0 star rating and 80 reviews within 2.66 km, this bakery offers a variety of delicious pastries, bubble tea, coffee, and more. It's the perfect place to satisfy your sweet tooth and grab a quick bite to eat.

No matter where you are in US and what you are looking for, we can find that for you!

MapMingle is a dynamic nearby places recommender system that provides personalized suggestions based on user-specified latitude, longitude, and search queries. The system sorts and ranks points of interest (POIs) using user reviews and ratings, ensuring high-quality and relevant recommendations.

Select your location to obtain your coordinates





# Demo Video





# Ranking Criteria

Ranking criteria:

1. Find the closest businesses by latitude longitude within 10 (or something) KM radius. and then show top 10 nearest.
2. For top 10, we will also include query term matching with the category of the business in the dataset
3. Stretch - Use review as another parameter



# Challenges

## Filtering Specific Business Types:

- Challenge: Identifying and filtering relevant businesses (like restaurants) from a diverse dataset that includes all kinds of businesses.
- Solution: Implement advanced categorization and tagging algorithms that accurately classify businesses based on their descriptions, services, and user reviews.

## User Location Input Method:

- Challenge: Deciding the most user-friendly and accurate method for users to input their location – via manual entry, map interaction, or automatic geolocation.
- Solution: Offer multiple options for location input, including automatic geolocation with user permission, and manual entry through map selection or text input.

## Handling Dynamic Data:

- Challenge: Ensuring that the data, especially business information and user reviews, are up to date and reflect current statuses.
- Solution: Regularly update the dataset and possibly integrate real-time data feeds where available.

## Scalability and Performance:

- Challenge: Maintaining system performance and speed as the user base and data volume grow.
- Solution: Optimize backend processes, consider cloud-based solutions for scalability, and implement efficient data caching methods.

## Accuracy of Recommendations:

- Challenge: Ensuring that the recommendations are accurate, relevant, and personalized to individual user preferences.
- Solution: Continuously refine the recommendation algorithms using user feedback and machine learning techniques.

## User Privacy and Data Security:

- Challenge: Protecting user privacy, especially regarding location data, and ensuring data security.
- Solution: Implement robust security protocols and transparent privacy policies, giving users control over their data.

## User Interface and Experience:

- Challenge: Creating an intuitive and engaging user interface that caters to diverse user groups.
- Solution: Conduct user experience research, perform usability testing, and iterate design based on user feedback.

## Cross-Platform Compatibility:

- Challenge: Ensuring consistent functionality and experience across different devices and platforms.
- Solution: Use responsive design principles and conduct thorough cross-platform testing.



# How to run?

1. Clone the code and navigate to src folder
2. copy the dataset in the "data" folder. the expected file name is:  
yelp\_academic\_dataset\_business.json
3. run -- pip install pandas numpy nltk geopy geojson rank\_bm25 matplotlib seaborn scikit-learn  
-- to install relevant packages
4. run -- pip install flask. -- to host the site
5. once this is done run -- python app.py
6. you can access the site at --http://127.0.0.1:5000



# Future Scope

## Advanced Personalization:

- Introduce machine learning algorithms for more nuanced user preference analysis, leading to highly personalized recommendations.
- Implement user behavior tracking to refine suggestions based on individual interaction patterns.

## Integration with Social Media:

- Plan for integration with social media platforms to incorporate user-generated content and reviews, enhancing the richness of recommendations.

## Real-Time Traffic and Event Data:

- Incorporate real-time traffic information and local event data to provide recommendations that are not just location-based but also time-sensitive.

## Augmented Reality (AR) Interface:

- Explore the possibility of an AR interface to provide users with an immersive way to explore their surroundings and discover places.

## Expansion to New Markets and Regions:

- Scale the system to cover more geographical locations, adapting to different cultural and regional dynamics in recommendations.

## Voice-Activated and In-Car Navigation Integration:

- Develop voice-activated features for ease of use while on the move, and consider integration with in-car navigation systems for a seamless experience.

## Sustainability and Accessibility Features:

- Integrate features that promote eco-friendly places and accessibility information for users with disabilities.



**Thank You!**