

# LA Restaurant Health Inspection & Recommendation System

Team HAL9000



# CONTENT

**Overview**

PART ONE



**Models**

PART TWO



**Demo**

PART THREE



**Future Works**

PART FOUR



# Project Objective and Definition

## Pain Points

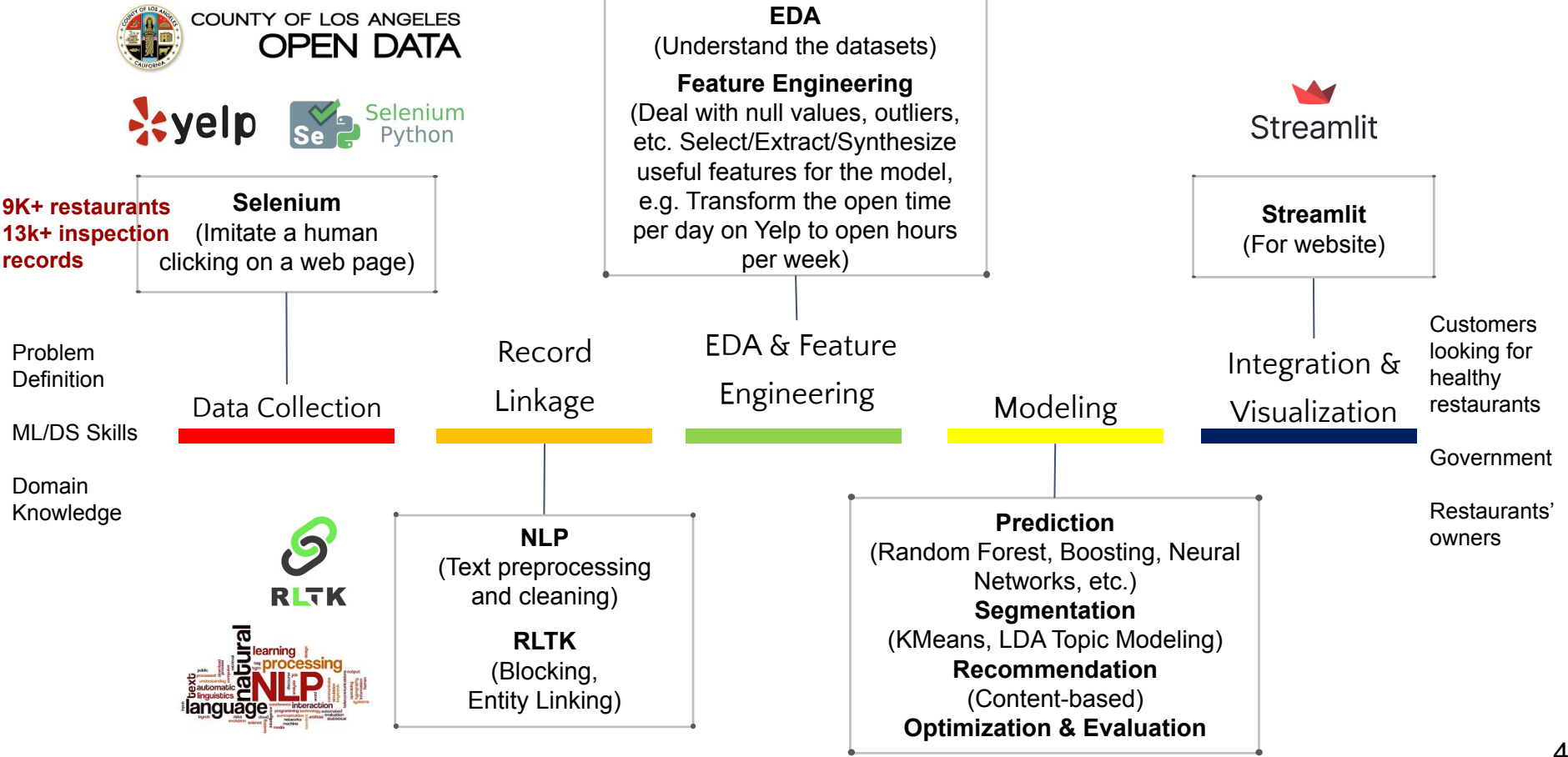
- Public **health issue** has attracted more and more attention. Few smart city projects studied on restaurant inspection.
- For the citizens, health inspection information can be **hard to find** on platforms.
- For the governments, the restaurant inspection is a source of cost and **efficient resource allocation** is needed.

## Our Project

- Predict the restaurant health violation risk
- Segment the restaurants
- Offer personalized restaurant recommendations



# Simplified Process Map



# Prediction

**Goal:** Classify the restaurants into 3 health risk levels (low/medium/high risk)

**Features:**

- Categorical: Location, Price, Size, Photos(bool), Take\_out(bool), Q&A(bool)
- Numerical: Open\_duration, employees, Review\_counts, Review\_sentiment\_polarity

**Machine Learning Models & Evaluation Metrics:**

Model	Accuracy	Roc_auc
Logistic Regression	0.835	0.58
SVM	0.839	0.63
Random Forest	0.842	0.67
XGBoost	0.827	0.71
Neural Networks	0.857	0.51

Baseline

Final Model

# Segmentation

Pipeline: PCA – KMeans – t-SNE & Topic Modeling

Model	Objective	Evaluation
PCA	Dimension Reduction	Explained variance on the first 3 PCs: 0.5
KMeans	Learn the representation of the data	Scree Plot: K=5
t-SNE	2D visualization of the clusters based on PCA	KL divergence = 0.47
Topic Modeling	Extract the keywords from comments	We sampled 10 restaurants and check if the keywords are really included in the descriptions of the restaurants. Included Rate = 70%

Included Rate

70%

Number of Correct Topic/ Detected Topics

Randomly Selected Restaurants

*e.g. 10 Restaurants*



Topic Model



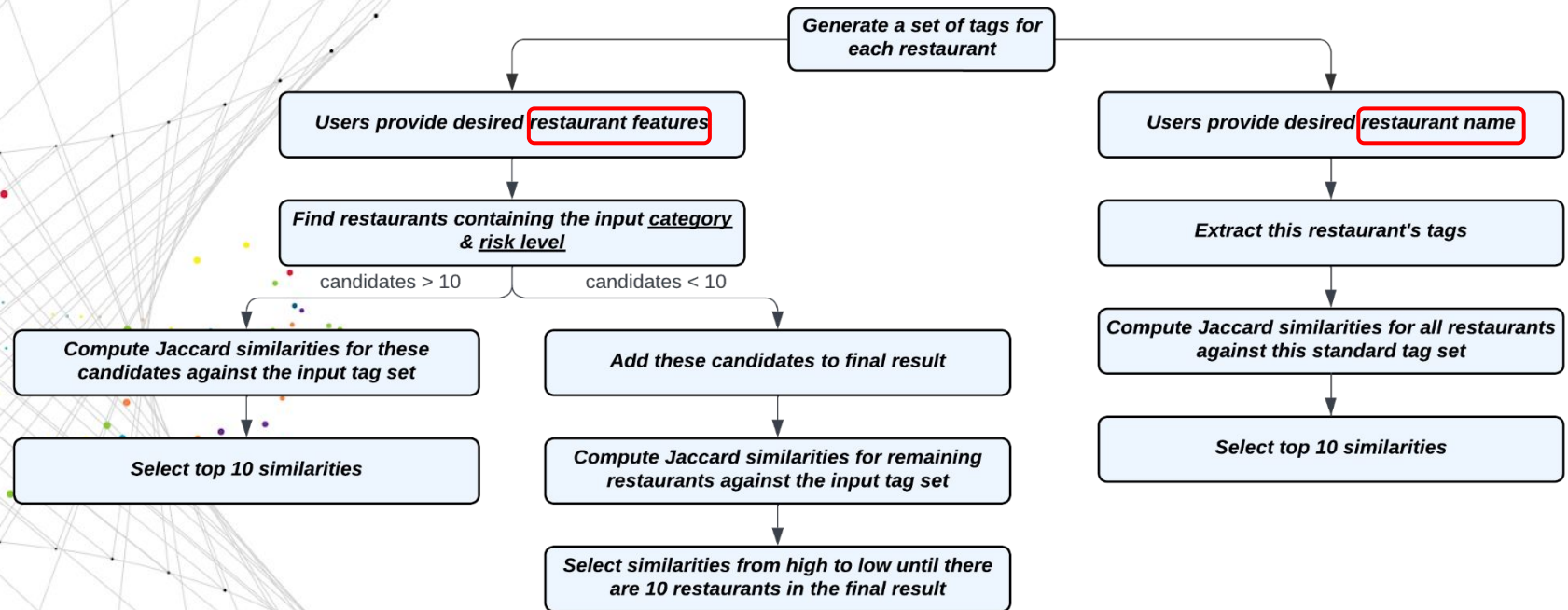
Detected Topic Keywords

Really in the Yelp Description? ?

Results

	restaurant	cluster	is_included
0	Pharaoh Karaoke Lounge	0	False
1	Sanamluang Cafe Hollywood	0	True
2	Eduardo's Border Grill	1	True
3	Mom's Donut and Chinese Food to Go	1	False
4	Ministry Of Coffee	1	True
5	Papa Johns Pizza	2	True
6	Moishe's Restaurant	3	True
7	Hot Thai Restaurant	3	True
8	Kai Ramen West Hollywood	3	True
9	Ginger's Divine Ice Creams	3	False

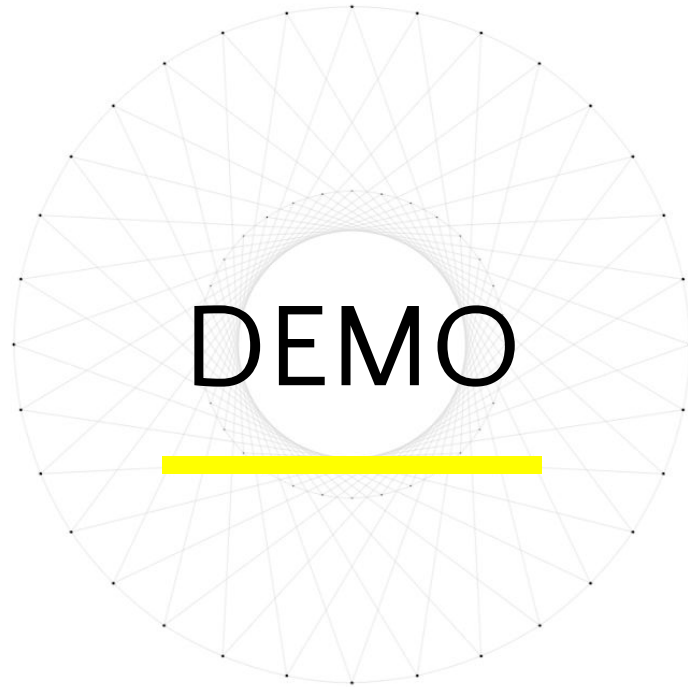
# Recommendation



## Evaluation:

- Among all 10 recommended results, how many are actually relevant?
- Ran some tests - average **74%** relevancy.

Backup Demo



Web application

File

Use Case





## Future Work

1. Optimize UI - adding images, organize layouts, etc.
2. Add an option for users to provide feedbacks on the models
3. Add a rating system for users to provide feedbacks on the recommendation systems
4. Crawl more features to feed into the model to improve the performance
5. Integrate ChatGPT for explaining the final results

1. N. H. M. Shamsuddin, N. A. Ali and R. Alwee, "An overview on crime prediction methods," 2017 6th ICT International Student Project Conference (ICT-ISPC), Johor, Malaysia, 2017, pp. 1-5, doi: 10.1109/ICT-ISPC.2017.8075335.
2. M. Feng et al., "Big Data Analytics and Mining for Effective Visualization and Trends Forecasting of Crime Data," in IEEE Access, vol. 7, pp. 106111-106123, 2019, doi: 10.1109/ACCESS.2019.2930410.
3. Iqbal, Rizwan, et al. "An experimental study of classification algorithms for crime prediction." Indian Journal of Science and Technology 6.3 (2013): 4219-4225.
4. Raschka, Sebastian. "Model evaluation, model selection, and algorithm selection in machine learning." arXiv preprint arXiv:1811.12808 (2018).
5. Llorente, Fernando, et al. "Marginal likelihood computation for model selection and hypothesis testing: an extensive review." SIAM Review 65.1 (2023): 3-58.
6. Barron, Andrew R. "Predicted squared error: a criterion for automatic model selection." Self-organizing methods in modeling. CRC Press, 2020. 87-103.
7. Choi, Jinkyung, Douglas Nelson, and Barbara Almanza. "Food safety risk for restaurant management: Use of restaurant health inspection report to predict consumers' behavioral intention." Journal of Risk Research 22.11 (2019): 1443-1457.
8. Choi, Jinkyung, Douglas Nelson, and Barbara Almanza. "Food safety risk for restaurant management: Use of restaurant health inspection report to predict consumers' behavioral intention." Journal of Risk Research 22.11 (2019): 1443-1457.
9. Siering, Michael. "Leveraging online review platforms to support public policy: Predicting restaurant health violations based on online reviews." Decision Support Systems 143 (2021): 113474.
10. Luna, Julio César, et al. "Food safety assessment and risk for toxoplasmosis in school restaurants in Armenia, Colombia." Parasitology research 118 (2019): 3449-3457.

The background features a complex geometric pattern of thin, light gray lines that form a hyperboloid of two sheets, creating a central void. Scattered throughout this structure are numerous small, multi-colored dots in shades of red, yellow, green, blue, and purple. The text is centered within this abstract composition.

**THANK YOU FOR WATCHING**

**Q & A**