Food & You

Enhancing Restaurant Experience with Neural Networks and Natural Language Processing

Camden Hunt and Na'inoa Loo

Data Collection

- Originally used Yelp API and Tomorrow.io API
  - Rate limits made progress difficult
- Was able to make API calls and get specific data we wanted

Example Data Output

```
Name: Colton's Social House, Rating: 4.2, Temperature: 24°C, Weather Code: 1001
Name: Pismo's Coastal Grill, Rating: 3.9, Temperature: 23.63°C, Weather Code: 1102
Name: Carrillo's Mexican Food, Rating: 4.4, Temperature: 24°C, Weather Code: 1001
Name: The Curry Pizza Company, Rating: 4.6, Temperature: 23.88°C, Weather Code: 1001
Name: The Mogul Restaurant Rating: 4.4, Temperature: 9.13°C, Weather Code: 1000
```

- Eventually we moved to official Yelp Dataset due to rate limits
  - Does weather matter if we are looking at a small geographic area?

Yelp Dataset

- Yelp dataset contains:
  - 6,990,280 reviews
  - 150,346 businesses
  - 200,100 pictures
  - 11 metropolitan areas
- Only looked at Santa Barbara metropolitan area
  - 211,748 reviews
  - 968 businesses

restaurants.json

For each restaurant, contains several variables we looked at, such as:

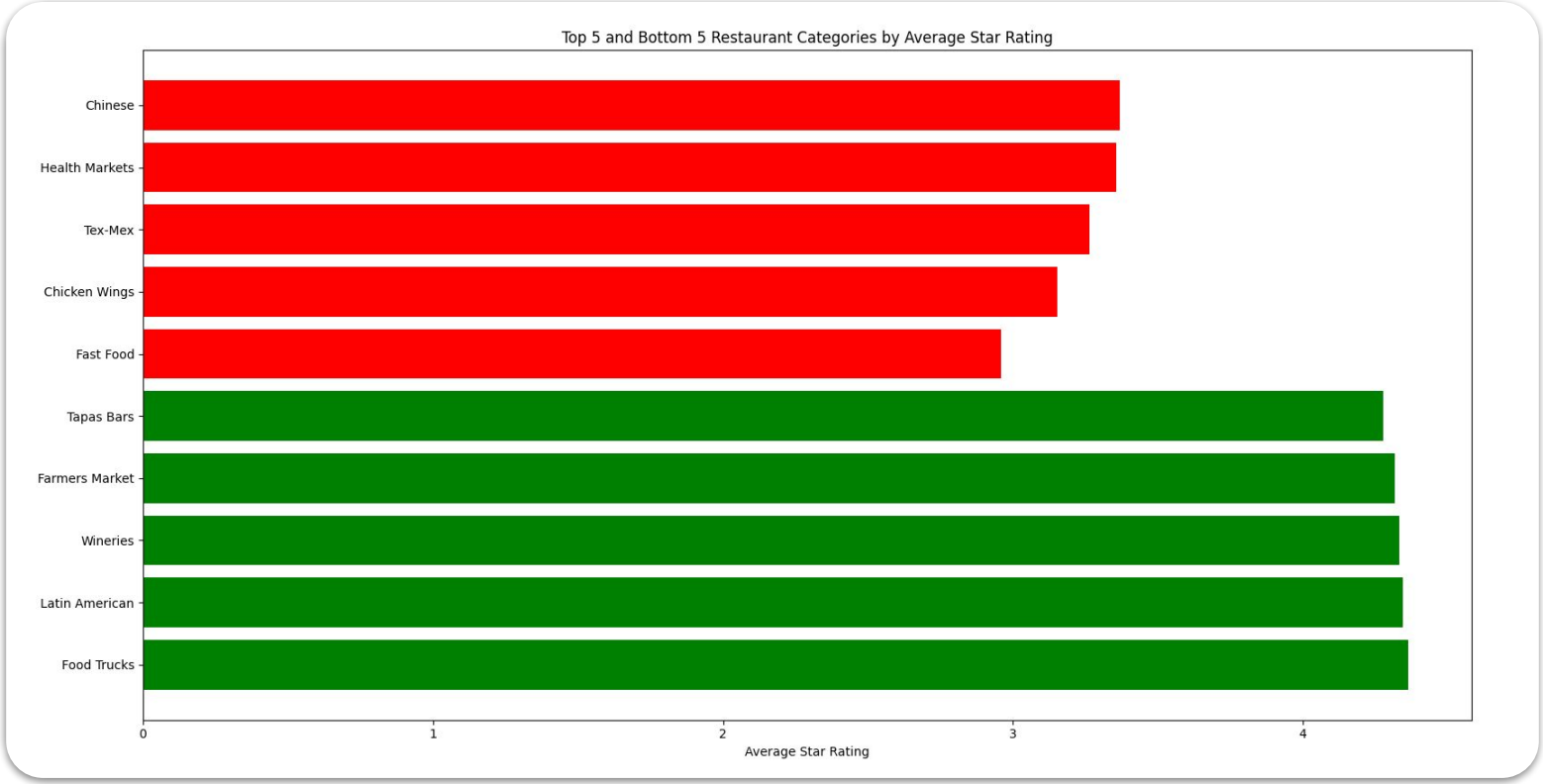
- Serve alcohol?
- BYOB?
- Bike parking?
- Accepts crypto?
- By appointment only?
- Caters?
- Coat check?
- Dogs allowed?
- Drive-thru?
- Dancing?
- Kids?
- Happy hour?
- TVs?
- Music?
- Open 24 hours a day?
- Outdoor seating?
- Delivery?
- Good for groups?
- Reservations?
- Takeout?
- Smoking allowed?
- Wheelchair accessible?

reviews.json

Contains basic data each review, such as:

- Review ID
- User ID
- Business ID
- Star rating
- Review content

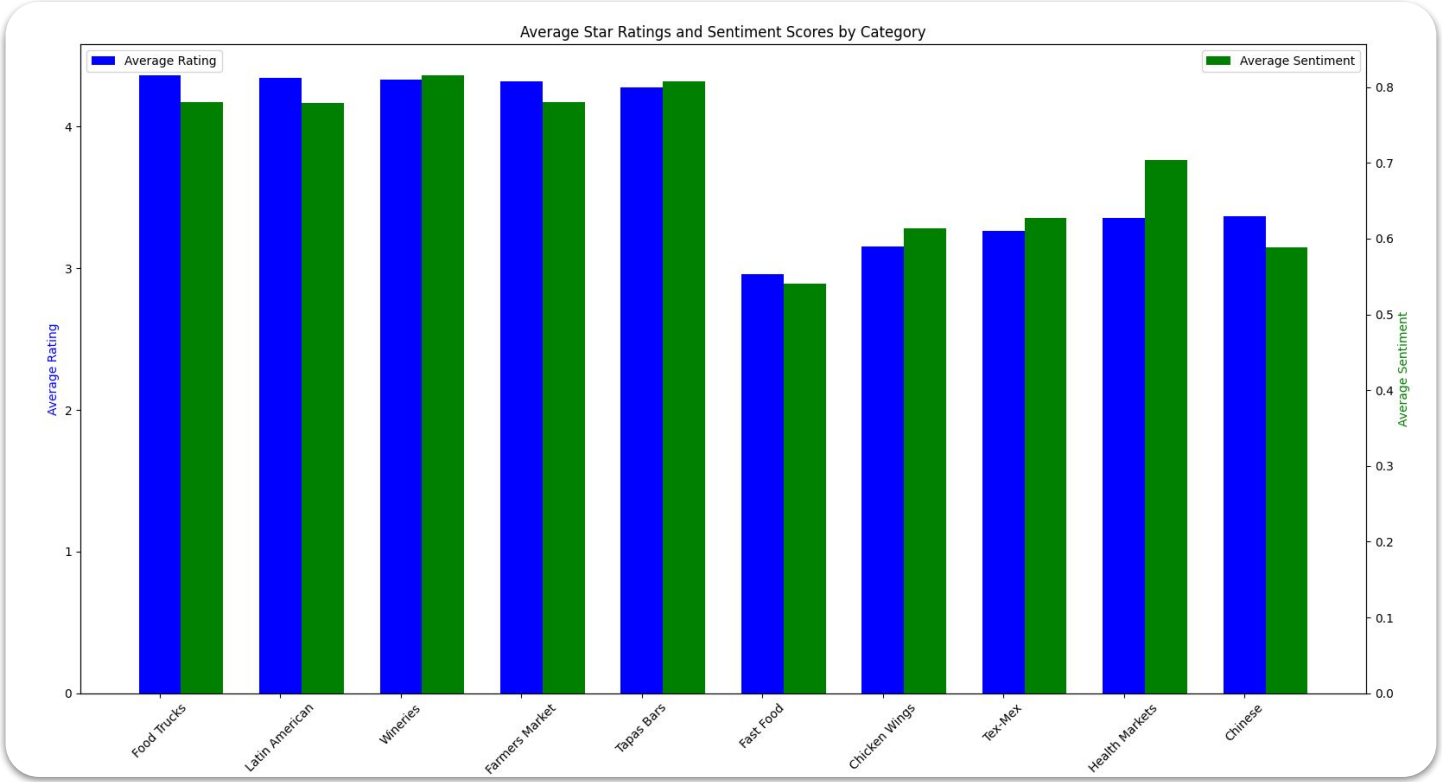
reviews.json



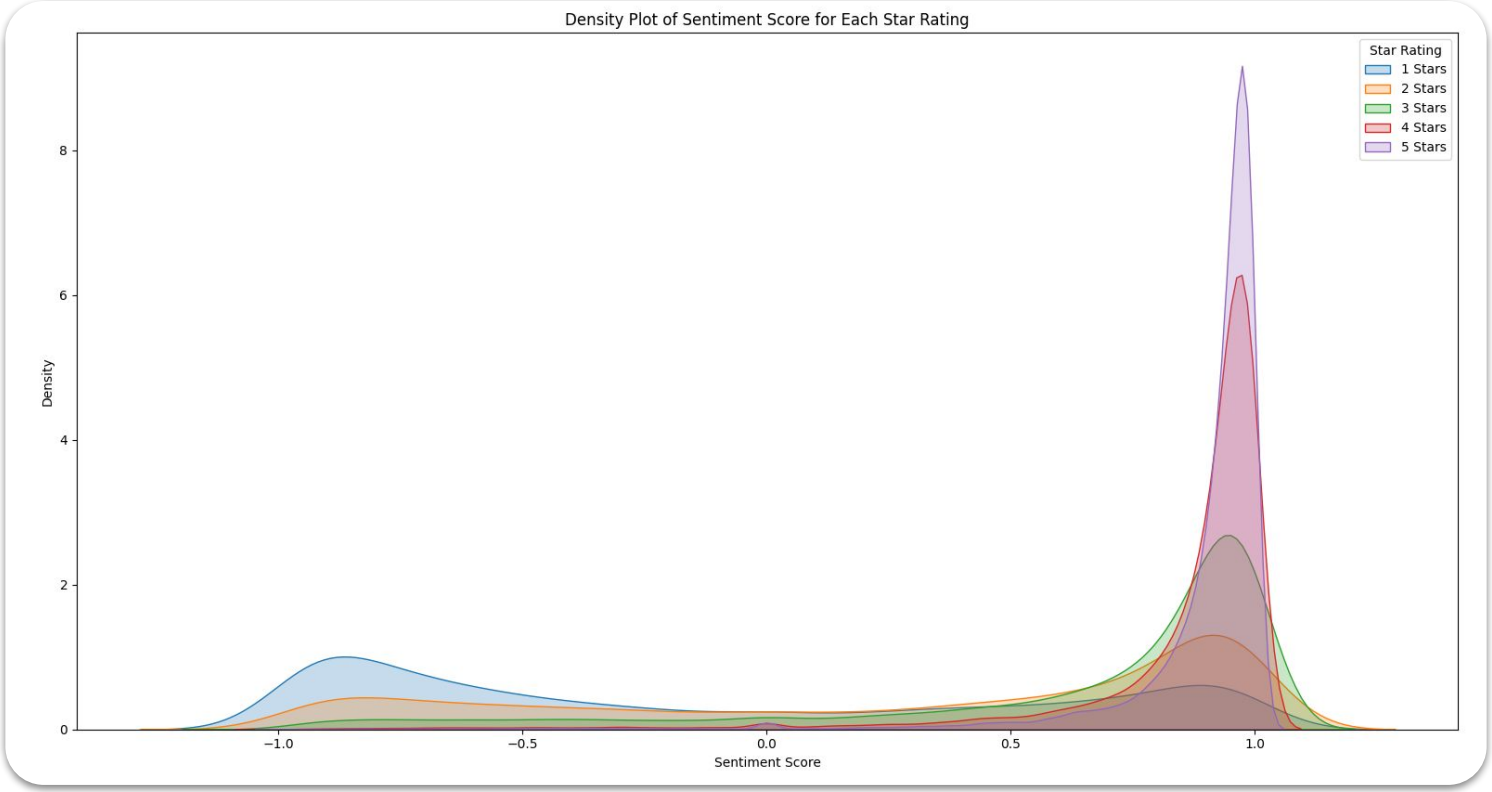
Sentiment Analysis

- Used sentiment analysis (using NLTK) to look at general feelings behind each review
  - Theoretically should correlate strongly with star rating

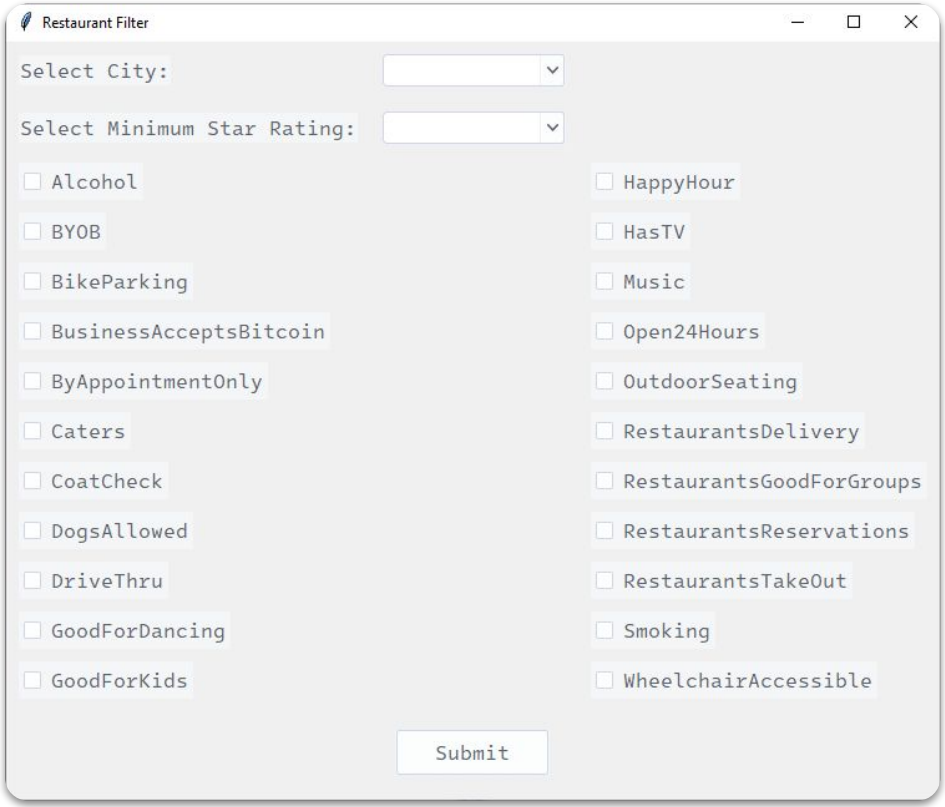
Sentiment Analysis Results



Sentiment Analysis Results

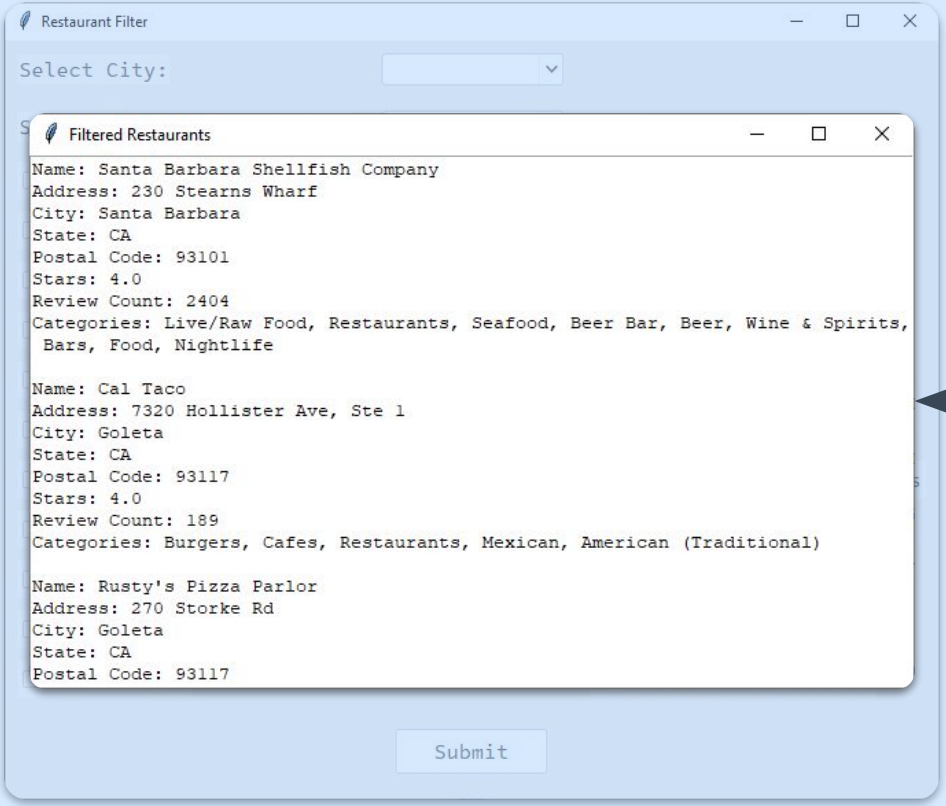


Restaurant Filter GUI



Allows user to search for restaurants to fit their criteria

Restaurant Filter GUI



Opens popup of restaurants that fit criteria

Neural Network

- Trained a neural network on restaurant data using TensorFlow and Keras
- User can search and select one or more favorite restaurants from a list
- Neural network suggests similar restaurants

TensorFlow

Keras

Training the Neural Network

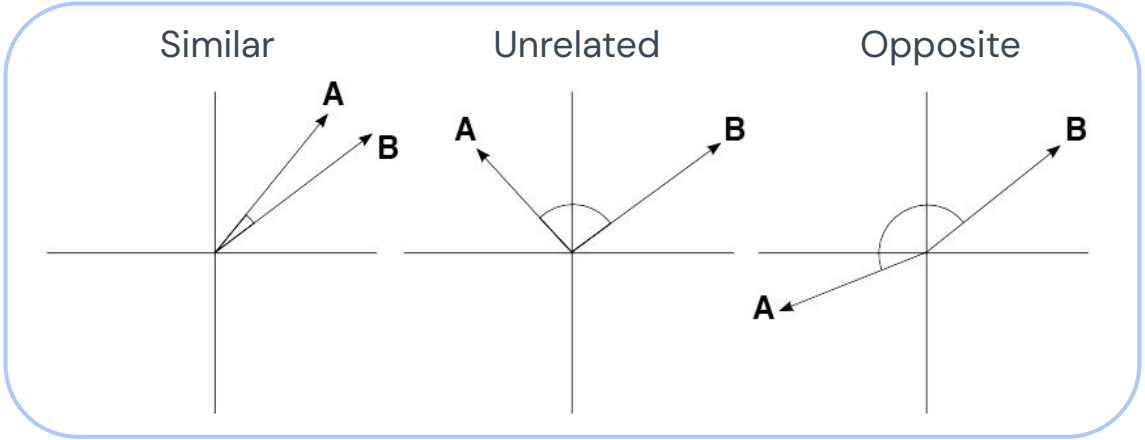
- Dimensionality is determined by the number of unique words in the TF-IDF (Term frequency Inverse document frequency) matrix
    - Creates a matrix of the less important and more important words in a restaurants.json entry
- Network is trained to reconstruct the input data, essentially predicting the TF-IDF vector for each restaurants.json entry

Making Predictions

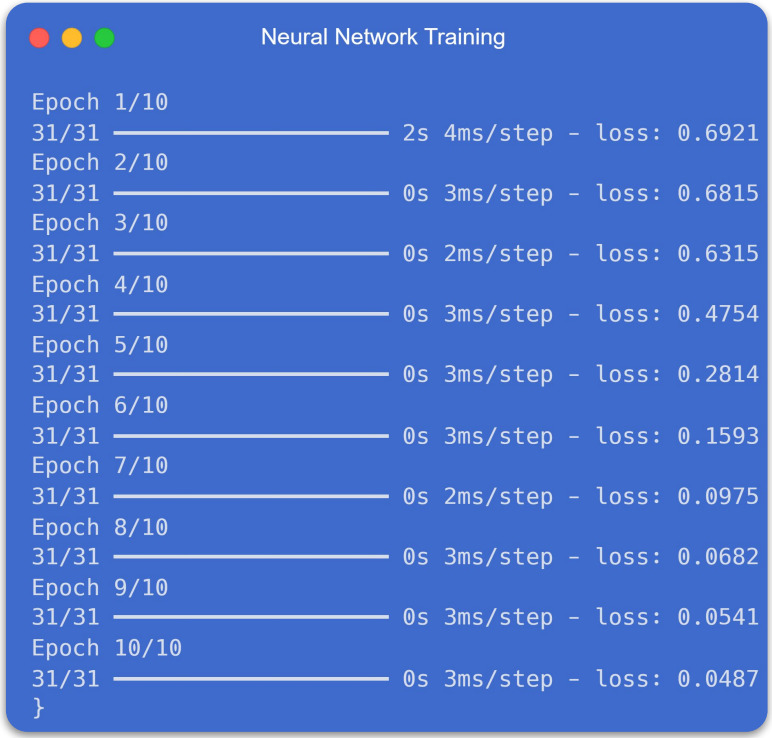
- When the user selects one or more restaurants, the TF-IDF vectors of these restaurants are fed into the trained neural network
  - The network generates new output vectors based on the input vectors
- The output vectors are compared to the vectors of all other restaurants using cosine similarity
  - The top 10 restaurants with the highest similarity scores are returned to the user

Ranking Predictions

- The cosine similarity metric measures the cosine of the angle between two vectors
  - If two TF-IDF vectors have a small angle between them, their cosine similarity score will be closer to 1
- The restaurants are ranked based on this score, and the top 10 are shown to the user

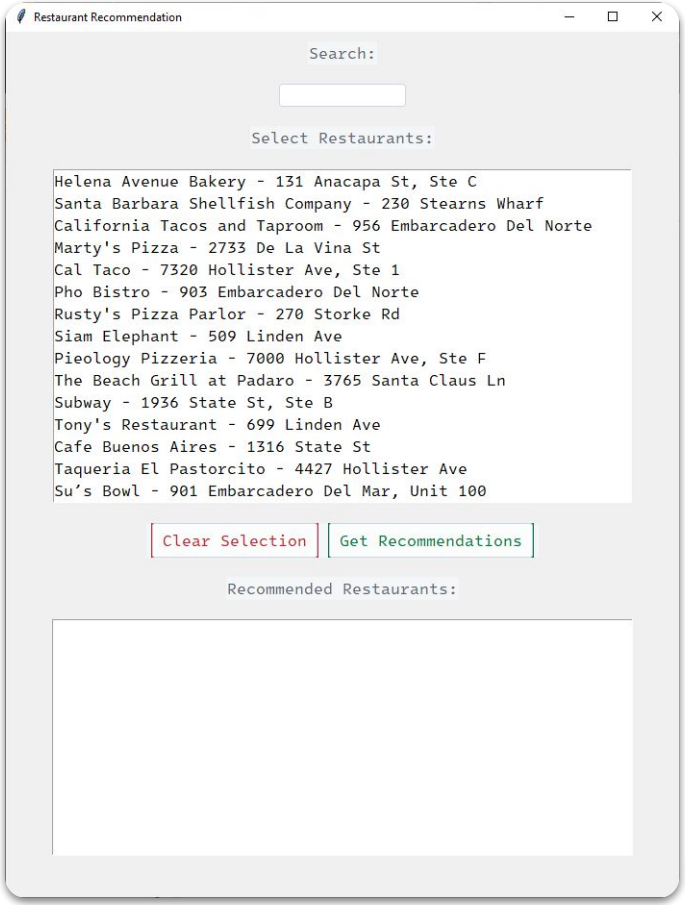


Neural Network



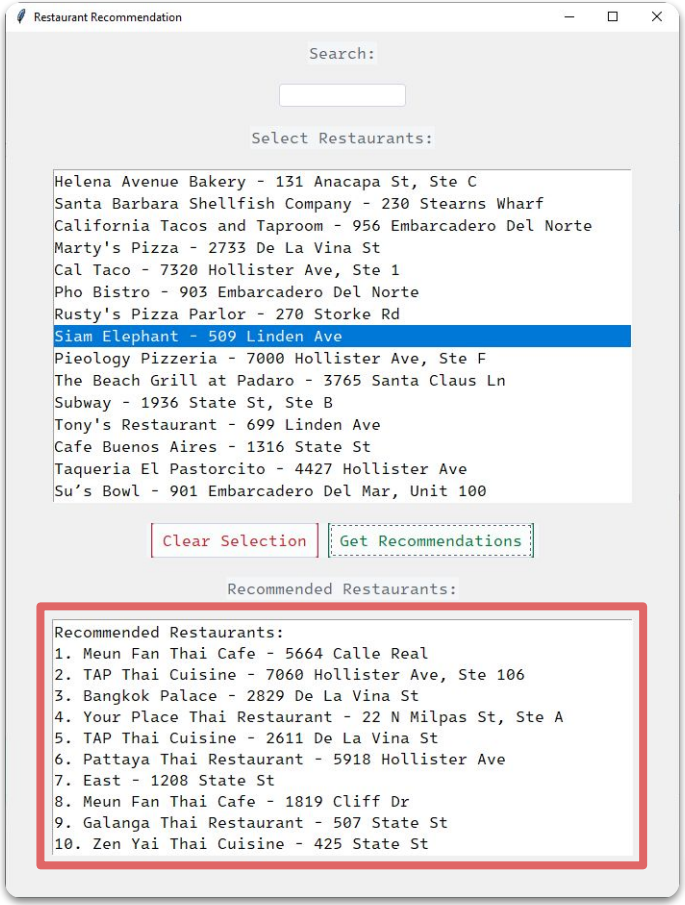
Program trains neural network on runtime

Neural Network (cont.)



To use the neural network, the user can select their favorite restaurants in our application

Neural Network (cont.)



Application gives suggestions for similar restaurants

Evaluating Accuracy?

- Data is generated based on specific user preferences
  - We do not track what restaurants the user goes to
- Integrate Yelp log-in to track if a user goes to a suggested restaurant and the score of their review

Demo