The Cocktail Experiment: Predicting Drink Preferences

An exploration of factors that influence people's drink preferences and a predictive model to forecast their beverage choices.

What is the Cocktail Experiment?

Problem Statement:

Selecting the perfect cocktail can be overwhelming due to the vast number of options available.

Objective:

To develop a machine learning-based recommendation system to suggest cocktails based on user preferences.

Personalized Recommendations

Using the collected data, the algorithm provides personalized cocktail recommendations tailored to the user's preferences, helping them discover new drinks they are likely to enjoy.

The Cocktail Experiment offers a unique and personalized approach to discovering new cocktails, making it easier for people to find drinks they will love.

The Cocktail Experiment: Dataset Overview

Data Source

The CSV file used in this analysis was sourced from Kaggle, providing comprehensive and structured data essential for our project.

.

Sample Size

The dataset contains 990 cocktails and 653 unique ingredients.

Variables

The dataset features include cocktail name, category, measurements, ingredients, instructions, glass type, and glass size.

Pre Processing Steps

Handling missing values, vectorization.

This dataset provides a comprehensive overview of individual drink preferences and associated demographic factors, offering valuable insights for the Cocktail Experiment project.

Algorithm Explanation

The algorithm starts by gathering data on user preferences, including their favorite cocktail ingredients, flavor profiles, and past drink selections.

Convert the drink ingredients into numerical vectors.

Calculate the similarity between the input drink vector and other drink vectors using cosine similarity..

Recommend the most similar cocktails based on the computed similarity scores.

Model Training and Vectorization

- ## Model Used

  Word2Vec model for vectorizing ingredients. It create dense vector representations of words. These representations, often called word embeddings, capture semantic meanings and relationships between words.

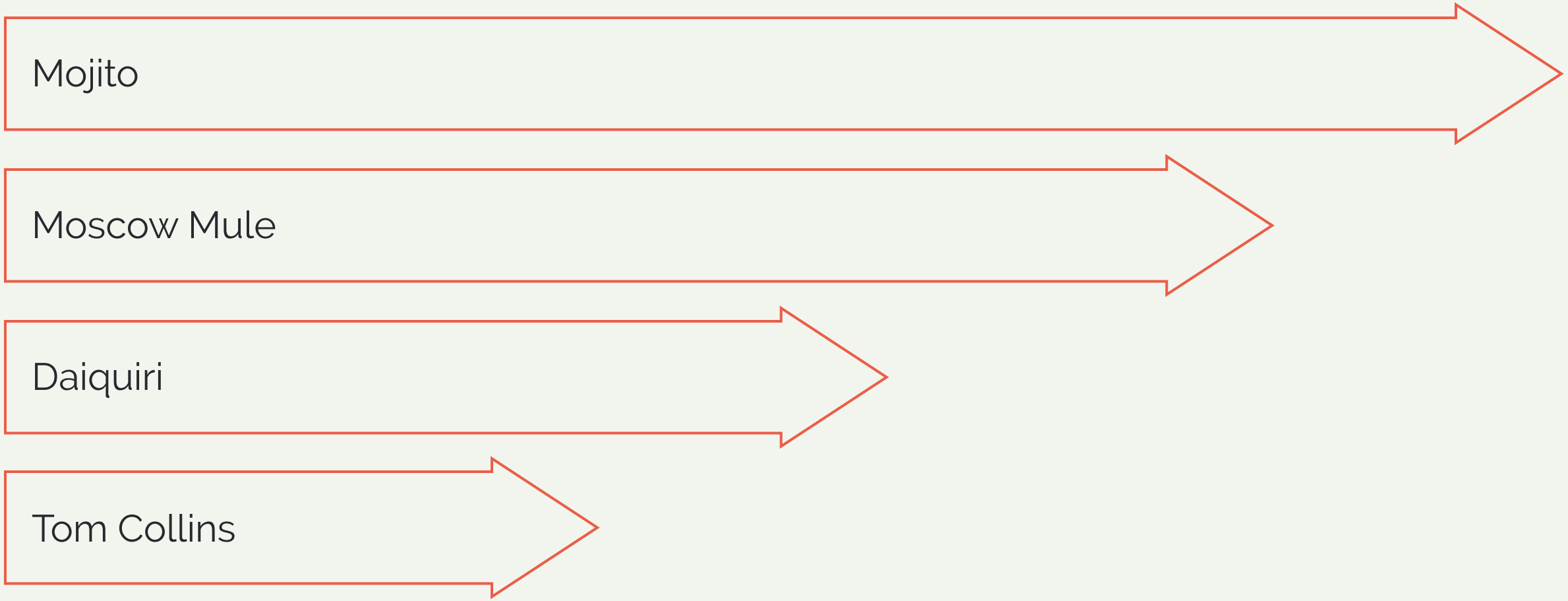
- ## Training Process

  Collected and preprocessed cocktail recipes, tokenized ingredients, trained the Word2Vec model to learn vector representations based on ingredient co-occurrences, and generated embeddings to capture semantic similarities

- ## Vector Size

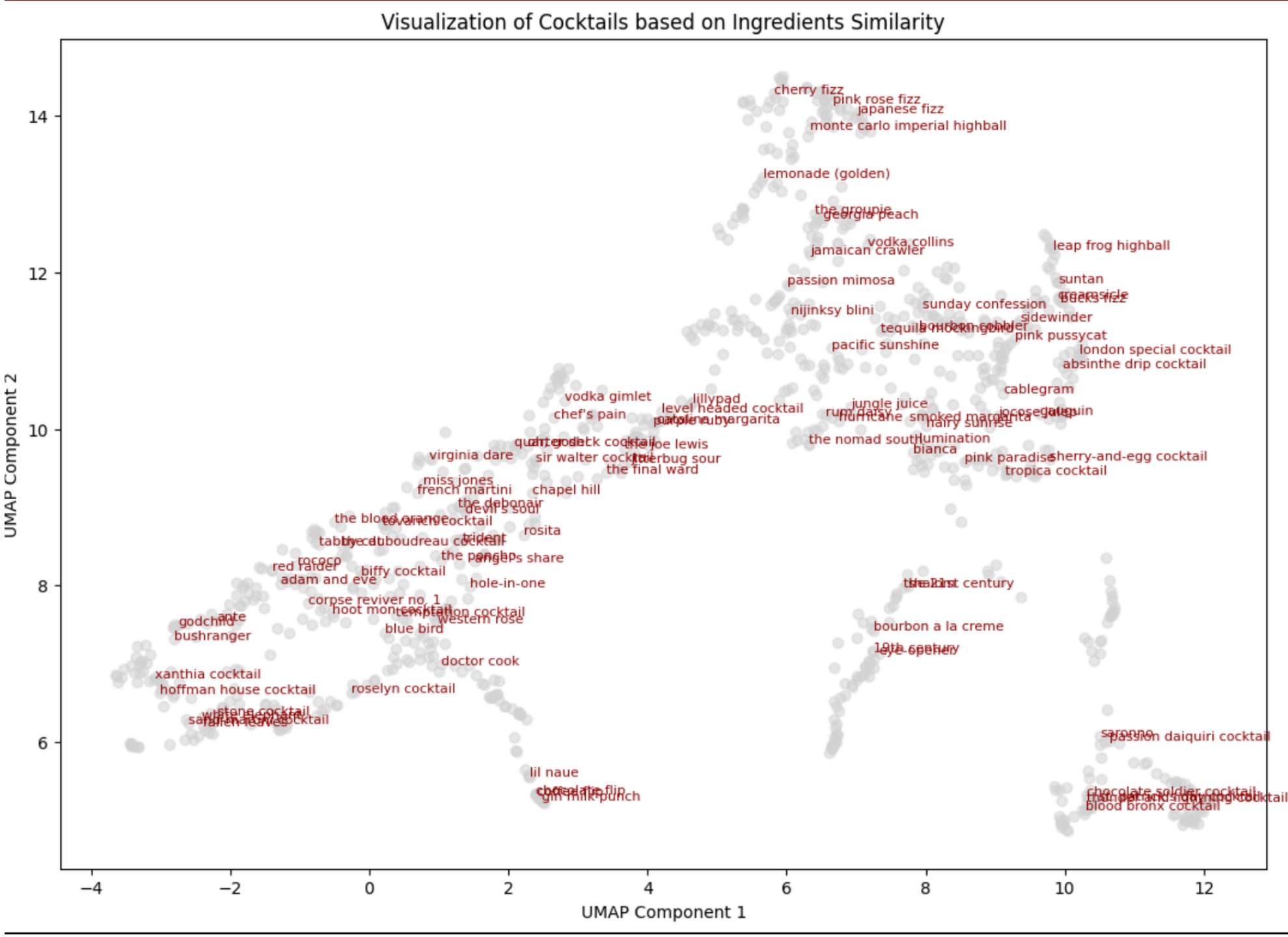
  For the initial vector size, we set it to 100 dimensions. This size provided a balance between capturing enough detail and keeping computations efficient. To ensure consistency, we made sure all vectors were the same size by padding or trimming them as needed. This standardization allowed us to compare the vectors accurately and reliably.

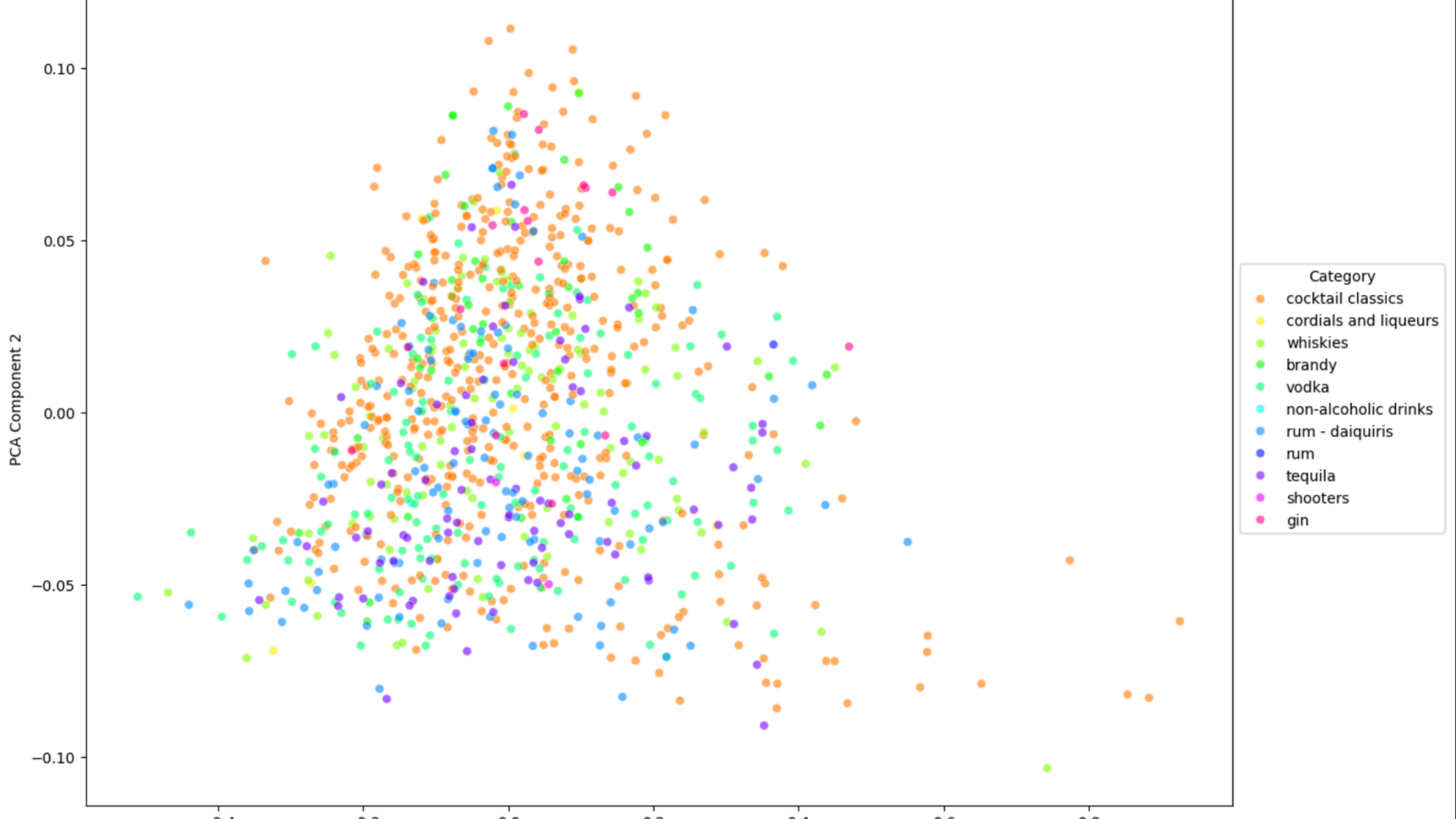
The Cocktail Experiment: Cosine Similarity



Drinks with similar flavors point the same way

Visuals show grouping includes "cherry fizz," "pink rose fizz," and "Japanese fizz," which are likely to share fruity and fizzy components. Another grouping includes "the blood and sand," "french martini," and "corpse reviver no. 1," indicating cocktails that may share more complex, layered ingredients. Each cluster highlights cocktails with similar ingredient profiles, showcasing common flavor characteristics.




0.10
0.05
0.00
−0.05
−0.10
PCA Component 2
Category
cocktail classics
cordials and liqueurs
whiskies
brandy
vodka
non-alcoholic drinks
rum - daiquiris
rum
tequila
shooters
gin

Conclusion and Future Work

Customized Drink Recommendations

The Cocktail Experiment uses machine learning algorithms to analyze customer preferences and provide personalized cocktail recommendations tailored to their individual tastes.

Curious Drinks

Display the similarities of curious drinks based on user likes.

Gain More personalized Data

To improve the recommendations and make them more personalized, gathering more detailed user preferences is key. By asking users not only about the drinks they like but also about specific flavors and ingredients they prefer, we can tailor the algorithm to better match their tastes.

User Feedback Integration

Implementing a feedback loop where users can rate the recommendations. This data can be used to refine the model, improving its predictive accuracy over time.

Flavor Profile Analysis

Conducting detailed analyses of flavor profiles by breaking down ingredients into their flavor components. This would allow the algorithm to make more nuanced recommendations based on users' flavor preferences

The Cocktail Experiment: Predicting Drink Preferences

The deployment page is the final step in the cocktail experimentation process. It is where the predicted drink preferences are presented to the user, allowing them to explore and discover new cocktail options tailored to their personal tastes.