Final Project
Steve Schmidt
Unsupervised ML

Cuisine Discovery and Recipe Recommendation System

Problem Statement:

As a Data Science Team within a Culinary Discovery Platform, our mission is to enhance the digital culinary experience by navigating the vast landscape of online recipes. We aim to develop a system that intelligently categorizes recipes based on their ingredient profiles and culinary styles, enabling users to explore new dishes that align with their dietary preferences and taste.

Data Description:

We will utilize the Recipe Ingredients Dataset from Kaggle, which features a diverse array of recipes along with detailed ingredient lists and associated cuisines. This dataset is pivotal for analyzing ingredient usage patterns and identifying culinary trends.

Potential Objectives:

Advanced Insights and Analytics:
Undertake a thorough analysis of ingredient co-occurrence within the clusters to pinpoint distinctive combinations that typify particular cuisines or reflect emerging dietary trends.
Delve into user interaction with different cuisine clusters to extract valuable insights regarding dietary preferences and the popularity of emerging food trends.

Recommender System:

Craft a content-based filtering mechanism that suggests recipes to users by drawing parallels in ingredients, cooking methodologies, and individual preferences. This recommendation engine will leverage the clusters formulated through advanced unsupervised learning techniques to deliver tailored suggestions.

Proposed Approach:

PCA (Principal Component Analysis) will be implemented for the initial dimensionality reduction of the ingredient feature matrix, aiding in the clustering process by capturing significant data variance.

t-SNE (t-Distributed Stochastic Neighbor Embedding) will serve as a comparative model for dimensionality reduction alongside PCA, focusing on preserving local data structures.
A comparative analysis of Natural Clustering vs. K-Means Clustering will follow, assessing each method's efficacy in grouping similar recipes. This evaluation will inform the development of a content-based filtering recommender system, leveraging advanced unsupervised learning techniques to offer personalized culinary suggestions.