Perfume Recommendation Model Proposal

Business Understanding

Introduction

The perfume industry faces increasing competition, with many products available in various scents, brands, and price ranges. A personalized perfume recommendation model can enhance customer engagement and satisfaction by helping users discover scents suited to their preferences. This model will leverage perfume data and user preferences to offer tailored recommendations, ultimately increasing purchase likelihood and brand loyalty.

Problem Statement

Customers often find it challenging to select perfumes that match their preferences, when shopping online or at the nearest perfume store. While shopping online, there is minimal opportunity to sample scents directly and clients will only rely on reviews, fragrance descriptions, and recommendations. This project aims to develop a machine learning-based recommendation system that suggests perfumes based on users' preferred fragrance notes, categories (e.g., unisex, men's, women's), and price ranges, improving the customer shopping experience.

Domain

This project lies in the intersection of e-commerce, fragrance retail, and recommendation systems. It uses machine learning to deliver a personalized shopping experience that aligns with the customer's scent profile and budget.

Target Audience

The target audience includes:

- Fragrance Enthusiasts who enjoy discovering new scents.

- E-commerce Retailers looking to enhance their online shopping experience with personalized recommendations.

- Fragrance Brands that want to improve product discoverability and increase sales.

Impact

A successful perfume recommendation model will improve user satisfaction, engagement, and conversion rates. By helping customers discover perfumes they are likely to enjoy with ease, the system can increase the likelihood of repeat purchases and build brand loyalty.

Pre-existing Work

Recommendation systems have been extensively researched in the e-commerce and media sectors (e.g., Amazon, Netflix), with content-based and collaborative filtering approaches being widely adopted. This project builds on existing content-based recommendation techniques but adapts them to the fragrance industry by incorporating perfume-specific attributes such as fragrance notes, categories, and prices

Objective

To develop a personalized Perfume Recommendation System.

Specific Objectives

1. To develop a recommendation engine that can suggest perfumes based on fragrance notes, categories, and price.
2. To implement a user-friendly interface to allow customers to input preferences (e.g., fragrance notes,gender and budget) and receive relevant recommendations.
3. To direct prospective buyers to an e-commerce platform whether they can purchase product of choice.

Success Criteria

The system will be considered successful if it accurately recommends perfumes based on fragrance notes, gender, category, and price range.

Data

The dataset was web-scraped from https://cierraperfumes.com/, an online perfume shop based in Kenya. It contains the following key features:

1. **Category**: The target gender (e.g., women, men, or unisex).
2. **Title**: The name of the perfume.
3. **Price**: The price of the perfume in Ksh.
4. **Link**: The specific URL location on Cierra's website.
5. **Image**: The URL of the perfume packaging image.
6. **Description**: A brief synopsis of the perfume.
7. **Top**: The initial scent, lasting 5–15 minutes, often light, fresh, or citrusy to capture attention.
8. **Middle**: The core fragrance, emerging after the top notes dissipate, lasting 30 minutes to an hour. These are often floral or spicy and represent the perfume's signature.
9. **Base**: The final lingering notes that emerge after the middle notes fade, lasting several hours. Typically rich, including woods, musk, and resins.

Data Preparation

Data cleaning will involve renaming columns, handling missing data, removing duplicates, and combining base, top, and heart notes into a single column. Feature normalization and lemmatization will also be applied to prepare the dataset for analysis.

Modeling

1. Content-Based Filtering:

   - Use cosine similarity on fragrance notes and categories to recommend perfumes similar to those a user likes.

   - Apply TF-IDF to convert perfume notes into feature vectors.

Evaluation

The model will be evaluated based on feature relevance, assessing how well the selected attributes (such as fragrance notes, categories, and price) contribute to accurate recommendations.

Deployment

1. Virtual Server Deployment: Hosting the recommendation system on a virtual server for accessibility and scalability.

2. User Interface: Develop a web interface or integrate the system into an existing e-commerce platform, allowing users to input their preferences (e.g., favorite notes, budget) for tailored recommendations.

Tools/Methodologies

- Data Collection: Scraping tools (BeautifulSoup, Selenium) in Python.

- Data Processing and Analysis: Pandas, NumPy, and scikit-learn.

- Modeling: Scikit-learn for content-based filtering

- Deployment: Flask for API, with possible integration into a web platform.

- Evaluation: Custom metric functions in Python, utilizing metrics from scikit-learn.