

LAPORAN PROYEK UAS  
PEMBELAJARAN MESIN (PRAKTIKUM)

**SISTEM REKOMENDASI KOPI MENGGUNAKAN METODE COLLABORATIVE  
FILTERING DENGAN COSINE SIMILARITY**



KELOMPOK 5

ANGGOTA:

[434221033] | [Feliza Adelia Desti Rohmawati]

[434221036] | [Kartika Nur Azizah]

[434221045] | [Azizah Ika Wulandari]

**PROGRAM STUDI D4 TEKNIK INFORMATIKA  
FAKULTAS VOKASI  
UNIVERSITAS AIRLANGGA  
SURABAYA  
2024**

## DAFTAR ISI

|  |    |
|--|----|
| DAFTAR ISI.....                                | 2  |
| DAFTAR TABEL.....                              | 3  |
| DAFTAR GAMBAR.....                             | 4  |
| BAB 1  |    |
| PENDAHULUAN.....                               | 5  |
| 1.1 LATAR BELAKANG.....                        | 5  |
| 1.2 RUMUSAN MASALAH.....                       | 6  |
| 1.3 TUJUAN.....                                | 7  |
| 1.4 MANFAAT.....                               | 7  |
| 1.5 BATASAN MASALAH.....                       | 8  |
| BAB 2  |    |
| TINJAUAN PUSTAKA.....                          | 9  |
| 2.1 SISTEM REKOMENDASI.....                    | 9  |
| 2.2 COLLABORATIVE FILTERING.....               | 10 |
| 2.3 COSINE SIMILARITY.....                     | 10 |
| 2.4 DATASET.....                               | 12 |
| BAB 3  |    |
| METODOLOGI.....                                | 15 |
| 3.1 GAMBARAN UMUM DAN STRUKTUR PENGERJAAN..... | 15 |
| 3.2 PENJELASAN KODE APP.PY.....                | 16 |
| BAB 4  |    |
| HASIL DAN PEMBAHASAN.....                      | 22 |
| 4.1 HASIL INPUT MELALUI TERMINAL.....          | 22 |
| 4.2 HASIL INPUT MELALUI WEBSITE.....           | 25 |
| BAB 5  |    |
| KESIMPULAN.....                                | 26 |
| 5.1 KESIMPULAN.....                            | 26 |
| 5.2 SARAN.....                                 | 27 |
| LAMPIRAN.....                                  | 29 |
| DAFTAR PUSTAKA.....                            | 30 |

## **DAFTAR TABEL**

|                                       |    |
|---------------------------------------|----|
| Tabel 2.1 Tabel Variabel Dataset..... | 12 |
|---------------------------------------|----|

## **DAFTAR GAMBAR**

|  |    |
|--|----|
| Gambar 1.1 Ilustrasi Proses Algoritma <i>Collaborative Filtering</i> ..... | 6  |
| Gambar 2.1 Histogram Visualisasi Distribusi Rating.....                    | 13 |
| Gambar 2.2 Histogram Visualisasi Distribusi Asal Biji Kopi.....            | 13 |
| Gambar 3.1 Struktur Folder Kode.....                                       | 15 |
| Gambar 4.1 Hasil Run Kode Melalui Terminal (1).....                        | 22 |
| Gambar 4.2 Hasil Run Kode Melalui Terminal (2).....                        | 22 |
| Gambar 4.3 Hasil Run Kode Melalui Terminal (3).....                        | 24 |
| Gambar 4.4 Hasil Run Kode Melalui Terminal (4).....                        | 24 |
| Gambar 4.5 Hasil <i>Heatmap Cosine Similarity</i> .....                    | 25 |
| Gambar 4.6 Hasil Run Kode Melalui Website (1).....                         | 26 |
| Gambar 4.7 Hasil Run Kode Melalui Website (2).....                         | 26 |
| Gambar 4.8 Hasil Run Kode Melalui Website (3).....                         | 27 |
| Gambar 4.9 Hasil Run Kode Melalui Website (4).....                         | 27 |
| Gambar 4.10 Hasil Run Kode Melalui Website (5).....                        | 27 |
| Gambar 4.11 Hasil Run Kode Melalui Website (6).....                        | 28 |
| Gambar 4.12 Hasil Run Kode Melalui Website (7).....                        | 28 |
| Gambar 4.13 Hasil Run Kode Melalui Website (8).....                        | 30 |
| Gambar 4.14 Hasil Run Kode Melalui Website (9).....                        | 30 |
| Gambar 4.15 Hasil Run Kode Melalui Website (10).....                       | 30 |
| Gambar 4.16 Hasil Run Kode Melalui Website (11).....                       | 31 |

## **BAB 1**

### **PENDAHULUAN**

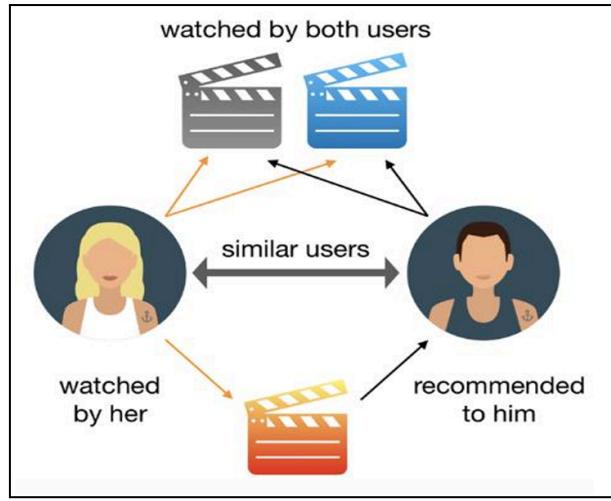
#### **1.1 LATAR BELAKANG**

Industri kopi telah mengalami perkembangan pesat dalam beberapa dekade terakhir, baik di Indonesia maupun di seluruh dunia. Indonesia sendiri merupakan salah satu produsen kopi terbesar di dunia dengan beragam jenis kopi yang memiliki cita rasa unik, seperti Arabika Gayo, Mandheling, Toraja, dan Kopi Luwak. Seiring dengan meningkatnya popularitas kopi, semakin banyak pula kedai kopi dan jenis kopi yang ditawarkan, baik kopi lokal maupun internasional.

Perkembangan ini mendorong munculnya tren baru dalam konsumsi kopi. Konsumen kini tidak hanya sekadar menikmati kopi sebagai minuman, tetapi juga menghargai cita rasa, aroma, dan asal-usul biji kopi. Hal ini menciptakan kebutuhan akan informasi yang lebih detail dan akurat mengenai kopi, serta memunculkan tantangan baru bagi para penikmat kopi, terutama bagi mereka yang baru mengenal dunia kopi. Dengan banyaknya pilihan yang tersedia, konsumen seringkali kesulitan dalam memilih kopi yang sesuai dengan preferensi mereka.

Sistem rekomendasi menjadi solusi yang efektif untuk mengatasi permasalahan ini. Sistem rekomendasi kopi dapat membantu pengguna menemukan kopi yang sesuai dengan selera mereka berdasarkan data preferensi pengguna lain atau berdasarkan karakteristik kopi yang diminati.

*Collaborative Filtering* merupakan salah satu metode yang populer dan efektif dalam sistem rekomendasi. Metode ini bekerja dengan menganalisis preferensi pengguna lain yang memiliki selera serupa. Ide dasar dari *Collaborative Filtering* adalah "orang yang memiliki preferensi serupa di masa lalu cenderung akan memiliki preferensi serupa di masa depan". *Cosine Similarity* adalah salah satu teknik yang umum digunakan dalam *Collaborative Filtering* untuk mengukur kesamaan antara preferensi pengguna. Teknik ini menghitung sudut kosinus antara dua vektor preferensi pengguna. Semakin kecil sudut kosinus (mendekati 0), semakin tinggi kesamaan antara kedua pengguna.



Gambar 1.1 Ilustrasi Proses Algoritma *Collaborative Filtering*

Keunggulan metode *Collaborative Filtering* dengan *Cosine Similarity* antara lain:

- Dapat memberikan rekomendasi yang personal: Sistem dapat mempelajari preferensi unik setiap pengguna dan memberikan rekomendasi yang sesuai dengan selera masing-masing.
- Dapat menemukan pola tersembunyi: Sistem dapat mengidentifikasi pola kesamaan antara pengguna yang mungkin tidak terlihat secara langsung.
- Relatif mudah diimplementasikan: *Cosine Similarity* merupakan teknik yang relatif sederhana dan mudah dihitung.

Oleh karena itu, penelitian ini akan mengembangkan sistem rekomendasi kopi menggunakan metode *Collaborative Filtering* dengan *Cosine Similarity* berdasarkan dataset yang berisi informasi tentang kopi.

## 1.2 RUMUSAN MASALAH

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana cara membangun sistem rekomendasi kopi menggunakan metode *Collaborative Filtering* berdasarkan dataset yang tersedia?
2. Bagaimana cara mengimplementasikan *Cosine Similarity* untuk mengukur kesamaan preferensi pengguna dalam sistem rekomendasi kopi?

3. Bagaimana tingkat akurasi sistem rekomendasi kopi yang dibangun menggunakan metode *Collaborative Filtering* dengan *Cosine Similarity* pada dataset yang digunakan?

### 1.3 TUJUAN

Tujuan dari penelitian ini adalah:

1. Membangun sistem rekomendasi kopi menggunakan metode *Collaborative Filtering* berdasarkan dataset yang tersedia.
2. Mengimplementasikan *Cosine Similarity* untuk mengukur kesamaan preferensi pengguna dalam sistem rekomendasi kopi.
3. Menguji tingkat akurasi sistem rekomendasi kopi yang dibangun menggunakan metode *Collaborative Filtering* dengan *Cosine Similarity* pada dataset yang digunakan.

### 1.4 MANFAAT

Manfaat yang diharapkan dari penelitian ini adalah:

1. Bagi Pengguna:
  - Memudahkan pengguna dalam menemukan kopi yang sesuai dengan preferensi mereka dari dataset yang tersedia.
  - Memberikan rekomendasi kopi yang lebih personal dan akurat berdasarkan rating dan ulasan kopi.
  - Meningkatkan pengalaman pengguna dalam menjelajahi berbagai jenis kopi yang ada di dataset.
2. Bagi Pelaku Bisnis Kopi:
  - Meningkatkan penjualan kopi dengan memberikan rekomendasi yang sesuai target pasar.
  - Membantu dalam memahami preferensi pelanggan berdasarkan data yang ada.
  - Sebagai sarana promosi yang efektif untuk produk kopi baru.
3. Bagi Pengembangan Ilmu Pengetahuan:
  - Memberikan kontribusi dalam pengembangan sistem rekomendasi berbasis data.

- Memberikan informasi tentang penerapan *Collaborative Filtering* dengan *Cosine Similarity* dalam rekomendasi kopi.

## 1.5 BATASAN MASALAH

Batasan masalah dalam penelitian ini adalah:

1. Sistem rekomendasi yang dibangun hanya berfokus pada rekomendasi kopi berdasarkan data yang terdapat pada dataset.
2. Data yang digunakan adalah data yang berasal dari dataset yang telah ditentukan.
3. Pengukuran akurasi sistem rekomendasi menggunakan metrik tertentu seperti *precision, recall, dan F1-score*.
4. Penelitian tidak membahas aspek implementasi sistem secara langsung, seperti pengembangan aplikasi atau integrasi dengan platform lain.

## BAB 2

### TINJAUAN PUSTAKA

#### 2.1 SISTEM REKOMENDASI

Sistem rekomendasi adalah sistem yang dirancang untuk memprediksi preferensi pengguna terhadap suatu item dan memberikan saran yang relevan. Sistem ini telah menjadi bagian penting dalam berbagai aplikasi modern, seperti *e-commerce* yang merekomendasikan produk kepada pelanggan, media sosial yang merekomendasikan teman atau konten yang relevan, platform hiburan yang merekomendasikan film, musik, dan buku, hingga aplikasi pariwisata yang merekomendasikan destinasi wisata dan tempat makan.

Manfaat utama dari sistem rekomendasi adalah meningkatkan kepuasan pengguna dengan memberikan saran yang sesuai dengan minat dan kebutuhan mereka. Bagi bisnis, sistem rekomendasi dapat meningkatkan penjualan dan pendapatan, serta membantu dalam memahami preferensi pelanggan. Namun, implementasi sistem rekomendasi juga memiliki tantangan tersendiri. Salah satunya adalah *cold start problem*, yaitu kesulitan dalam memberikan rekomendasi kepada pengguna baru atau item baru karena kurangnya data preferensi. Tantangan lainnya adalah *data sparsity*, di mana data preferensi pengguna yang tersedia seringkali sangat sedikit dibandingkan dengan jumlah total item.

Secara umum, terdapat dua jenis utama sistem rekomendasi: *Content-based Filtering* dan *Collaborative Filtering*. *Content-based Filtering* memberikan rekomendasi berdasarkan kesamaan antara item yang disukai pengguna di masa lalu dengan item baru. Sistem akan menganalisis karakteristik dari item-item tersebut untuk menemukan kesamaan. Misalnya, dalam sistem rekomendasi film, *Content-based Filtering* akan mempertimbangkan genre, aktor, sutradara, dan elemen-elemen lain dari film yang disukai pengguna untuk merekomendasikan film baru yang serupa.

Di sisi lain, *Collaborative Filtering* memberikan rekomendasi berdasarkan preferensi pengguna lain yang memiliki selera serupa. Sistem akan mencari pola kesamaan antara pengguna berdasarkan rating atau interaksi mereka dengan item. *Collaborative Filtering* mampu memberikan rekomendasi yang lebih personal dan menemukan pola tersembunyi yang mungkin tidak terlihat secara langsung. Namun, *Collaborative Filtering* lebih rentan terhadap *cold start problem* dan membutuhkan data yang cukup besar untuk menghasilkan rekomendasi yang baik.

## 2.2 COLLABORATIVE FILTERING

Collaborative Filtering merupakan metode yang populer dalam sistem rekomendasi karena kemampuannya dalam memberikan rekomendasi yang personal dan akurat. Metode ini bekerja dengan asumsi bahwa pengguna yang memiliki preferensi serupa di masa lalu cenderung akan menyukai item yang sama di masa depan. Terdapat dua pendekatan utama dalam Collaborative Filtering: *User-based Collaborative Filtering* dan *Item-based Collaborative Filtering*.

*User-based Collaborative Filtering* mencari pengguna lain yang memiliki preferensi serupa dengan pengguna target, kemudian merekomendasikan item yang disukai oleh pengguna serupa tersebut. Metode ini relatif mudah dipahami dan diimplementasikan, namun memiliki *scalability* yang rendah karena perhitungan kesamaan antar pengguna membutuhkan waktu yang lama jika jumlah pengguna sangat banyak.

*Item-based Collaborative Filtering* mencari item yang serupa dengan item yang disukai pengguna target, kemudian merekomendasikan item serupa tersebut kepada pengguna. Metode ini memiliki *scalability* yang lebih tinggi karena perhitungan kesamaan antar item dapat dilakukan secara *offline*. Namun, rekomendasi yang dihasilkan cenderung kurang personal karena fokus pada kesamaan antar item, bukan preferensi pengguna secara keseluruhan.

Untuk mengukur kesamaan antara pengguna atau item dalam *Collaborative Filtering*, diperlukan metrik kesamaan. Beberapa metrik kesamaan yang umum digunakan antara lain *Cosine Similarity*, *Pearson Correlation*, dan *Jaccard Similarity*. *Cosine Similarity* menghitung sudut kosinus antara dua vektor preferensi, *Pearson Correlation* menghitung korelasi linier antara dua vektor preferensi, dan *Jaccard Similarity* mengukur kesamaan antara dua himpunan item yang disukai oleh dua pengguna. Pemilihan metrik kesamaan yang tepat bergantung pada jenis data dan tujuan analisis.

## 2.3 COSINE SIMILARITY

*Cosine Similarity* adalah salah satu teknik yang digunakan untuk mengukur kesamaan antara dua vektor. Dalam sistem rekomendasi, vektor tersebut merepresentasikan preferensi pengguna terhadap item, seperti rating yang diberikan pada kopi. *Cosine Similarity* dihitung

berdasarkan sudut kosinus antara dua vektor. Semakin kecil sudutnya (mendekati 0 derajat), semakin tinggi kesamaannya (mendekati 1).

Rumus *Cosine Similarity* adalah sebagai berikut:

$$\text{Cosine Similarity} (A, B) = (A \cdot B) / (\|A\| \|B\|)$$

di mana A dan B adalah dua vektor yang akan dihitung kesamaannya,  $A \cdot B$  adalah *dot product* antara A dan B, dan  $\|A\|$  dan  $\|B\|$  adalah *magnitude* (panjang) dari A dan B.

Contoh Kasus:

1. Rekomendasi Film: Misalkan, vektor A merepresentasikan rating pengguna A terhadap beberapa film, dan vektor B merepresentasikan rating pengguna B. *Cosine Similarity* dapat digunakan untuk mengukur kesamaan selera film antara pengguna A dan B. Jika nilai *Cosine Similarity* tinggi, maka sistem dapat merekomendasikan film yang disukai oleh pengguna A kepada pengguna B, dan sebaliknya.
2. Rekomendasi Produk: Dalam *e-commerce*, vektor A dapat merepresentasikan riwayat pembelian pengguna A, dan vektor B merepresentasikan riwayat pembelian pengguna B. *Cosine Similarity* dapat digunakan untuk mengukur kesamaan pola pembelian antara pengguna A dan B. Jika nilai *Cosine Similarity* tinggi, maka sistem dapat merekomendasikan produk yang dibeli oleh pengguna A kepada pengguna B.
3. Analisis Teks: *Cosine Similarity* juga dapat digunakan untuk mengukur kesamaan antara dua dokumen teks. Dalam hal ini, vektor A dan B merepresentasikan frekuensi kemunculan kata dalam dokumen A dan B. *Cosine Similarity* yang tinggi mengindikasikan bahwa kedua dokumen memiliki topik yang serupa.

Interpretasi Nilai *Cosine Similarity*:

Nilai *Cosine Similarity* berkisar antara -1 hingga 1:

- 1: Kesamaan sempurna. Kedua vektor memiliki arah yang sama.
- 0: Tidak ada kesamaan. Kedua vektor saling tegak lurus.
- -1: Kebalikan sempurna. Kedua vektor memiliki arah yang berlawanan.

Kelebihan *Cosine Similarity*:

- Mudah diinterpretasikan.
- Relatif mudah dihitung.

- Efektif dalam ruang dimensi tinggi.

Kekurangan *Cosine Similarity*:

- Tidak memperhitungkan *magnitude* dari vektor. Dua vektor dengan arah yang sama tetapi *magnitude* berbeda akan memiliki *Cosine Similarity* 1, meskipun salah satu vektor lebih "panjang" daripada yang lain.
- Sensitif terhadap skala data. Perbedaan skala antar variabel dapat mempengaruhi nilai *Cosine Similarity*.

*Cosine Similarity* merupakan metrik yang serbaguna dan telah banyak diterapkan dalam berbagai bidang, termasuk sistem rekomendasi, analisis teks, dan *computer vision*.

## 2.4 DATASET

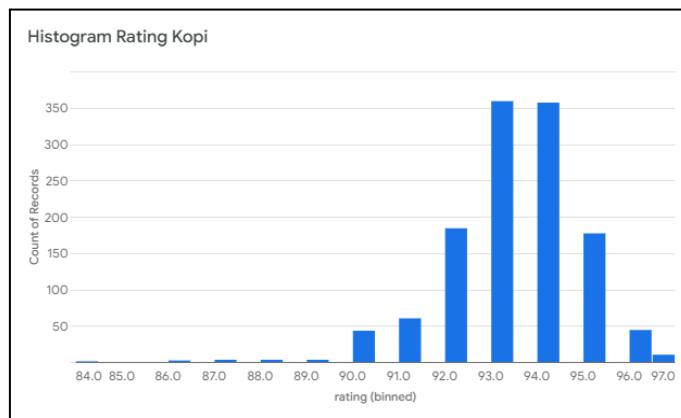
Penelitian ini menggunakan dataset "simplified\_coffee.csv" yang berisi informasi detail tentang berbagai jenis kopi. Dataset ini diperoleh dari Kaggle [https://www.kaggle.com/datasets/schmoyote/coffee-reviews-dataset?select=simplified\\_coffee.csv](https://www.kaggle.com/datasets/schmoyote/coffee-reviews-dataset?select=simplified_coffee.csv). Berikut adalah penjelasan detail tentang variabel-variabel dalam dataset:

| Variabel    | Penjelasan   | Tipe Data | Contoh                          |
|-------------|--|-----------|---------------------------------|
| name        | Nama kopi  | Object    | Ethiopia Shakiso Mormora        |
| roaster     | Nama roaster (penya ngrai kopi)                                | Object    | Revel Coffee                    |
| roast       | Jenis sangrai (Light, Medium, Medium-Light, Medium-Dark, Dark) | Object    | Medium-Light                    |
| loc_country | Negara lokasi roaster  | Object    | United States                   |
| origin      | Asal biji kopi   | Object    | Ethiopia                        |
| 100g_USD    | Harga kopi per 100 gram dalam dolar AS                         | Float     | 4.7                             |
| rating      | Rating kopi (skala numerik)                                    | Integer   | 92                              |
| review_date | Tanggal ulasan   | Object    | November 2017                   |
| review      | Teks ulasan kopi   | Object    | "Crisply sweet, cocoa-toned..." |

Tabel 2.1 Tabel Variabel Dataset

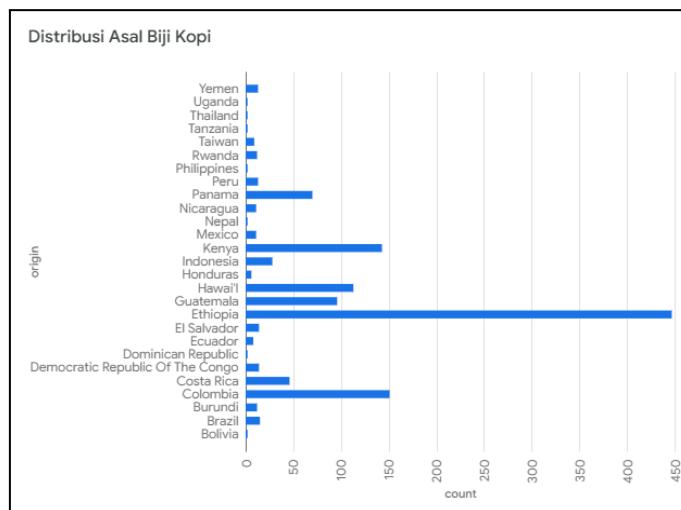
Statistik Deskriptif:

- Jumlah data: 1246
- Distribusi rating: Rating berkisar antara 84 hingga 97, dengan rata-rata 93.31 dan standar deviasi 1.48.



Gambar 2.1 Histogram Visualisasi Distribusi Rating

- 10 roaster teratas: | Roaster | Jumlah Kopi |:-----|:-----|  
Kakalove Cafe | 91 || JBC Coffee Roasters | 91 || Paradise Roasters | 83 || Dragonfly  
Coffee Roasters | 30 || Bird Rock Coffee Roasters | 29 || Temple Coffee Roasters | 27  
|| GK Coffee | 25 || Hula Daddy Kona Coffee | 24 || Red Rooster Coffee Roaster | 18  
|| RamsHead Coffee Roasters | 18 |
- Distribusi jenis sangrai: | Jenis Sangrai | Jumlah Kopi |:-----|:-----|  
Medium-Light | 904 || Light | 198 || Medium | 114 || Medium-Dark | 15 || Dark | 3 |



## Gambar 2.2 Histogram Visualisasi Distribusi Asal Biji Kopi

### Preprocessing Data:

Sebelum membangun sistem rekomendasi, diperlukan tahapan *preprocessing* data untuk membersihkan dan mempersiapkan data. Beberapa langkah *preprocessing* yang akan dilakukan antara lain:

- Menangani *Missing Values*: Variabel roast memiliki beberapa *missing values*. *Missing values* ini akan diatasi dengan mengisi nilai yang paling sering muncul (modus) atau dengan menghapus baris data yang memiliki *missing values*.
- Mengubah Variabel Kategorikal menjadi Numerik: Variabel-variabel kategorikal seperti roaster, roast, loc\_country, dan origin akan diubah menjadi numerik menggunakan teknik *one-hot encoding*. Teknik ini akan mengubah setiap kategori menjadi variabel dummy dengan nilai 0 atau 1.
- Normalisasi Data: Variabel numerik seperti 100g\_USD dan rating akan dinormalisasi untuk menyeragamkan skala antar variabel. Normalisasi data dapat dilakukan dengan teknik min-max scaling atau z-score standardization.

### Relevansi Dataset:

Dataset "simplified\_coffee.csv" dipilih karena berisi informasi yang relevan dengan tujuan penelitian, yaitu mengembangkan sistem rekomendasi kopi. Dataset ini mencakup berbagai jenis kopi dengan informasi detail seperti nama kopi, roaster, jenis sangrai, asal biji kopi, harga, dan rating. Variabel rating akan digunakan sebagai dasar untuk membangun sistem rekomendasi *collaborative filtering*. Dengan menerapkan teknik *cosine similarity* pada data rating, sistem dapat mengidentifikasi pengguna yang memiliki preferensi serupa dan memberikan rekomendasi kopi yang sesuai dengan selera mereka.

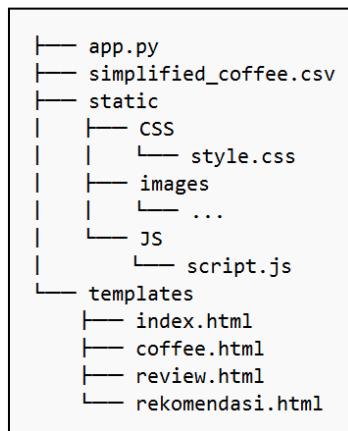
## BAB 3

### METODOLOGI

#### 3.1 GAMBARAN UMUM DAN STRUKTUR PENGERJAAN

Sistem rekomendasi kopi ini akan diimplementasikan menggunakan bahasa pemrograman Python dan framework Flask. Python dipilih karena memiliki library yang lengkap untuk analisis data dan *machine learning*, seperti Pandas, Scikit-learn, dan Matplotlib/Seaborn. Flask dipilih karena merupakan *framework* yang ringan dan mudah dipelajari untuk membangun aplikasi web.

Struktur kode program akan diorganisir ke dalam beberapa folder dan file untuk menjaga keteraturan dan modularitas kode. Berikut adalah struktur folder dan file yang digunakan:



Gambar 3.1 Struktur Folder Kode

Penjelasan:

- app.py: File utama yang berisi kode Python untuk menjalankan aplikasi Flask. Di dalam file ini, terdapat kode untuk:
  - *Preprocessing* data: Membaca data dari file CSV, membersihkan data, membuat *pivot table*, dan melakukan normalisasi data.
  - Pembuatan model *collaborative filtering*: Menghitung matriks *Cosine Similarity* antar kopi.
  - Fungsi untuk menghasilkan rekomendasi: Menerima input nama kopi dan menghasilkan rekomendasi kopi yang serupa.
  - Definisi rute: Menentukan URL yang akan diproses oleh Flask.

- *Rendering* template HTML: Mengirimkan data ke template HTML untuk ditampilkan di browser.
- simplified\_coffee.csv: File dataset yang berisi informasi tentang berbagai jenis kopi.
- static: Folder yang berisi file statis seperti CSS, JavaScript, dan gambar.
  - css/style.css: File CSS untuk mengatur tampilan aplikasi web.
  - images: Folder untuk menyimpan gambar yang digunakan dalam aplikasi web, seperti gambar biji kopi, cangkir kopi, atau logo.
  - js/script.js: File JavaScript untuk menambahkan interaktivitas pada aplikasi web, misalnya validasi input pengguna atau animasi.
- templates: Folder yang berisi file template HTML.
  - index.html: Halaman utama aplikasi web yang akan menjadi halaman pertama yang dilihat pengguna. Halaman ini berisi penjelasan utama website dan navbar menuju halaman lainnya.
  - coffee.html: Halaman yang menampilkan detail informasi tentang kopi, seperti nama, dan harga kopi.
  - review.html: Halaman untuk menampilkan ulasan kopi dari pengguna.
  - rekomendasi.html: Halaman untuk menampilkan rekomendasi kopi yang dihasilkan oleh sistem sesuai inputan pengguna.

Dengan struktur kode yang terorganisir seperti ini, pengembangan dan pemeliharaan sistem rekomendasi kopi akan lebih mudah dilakukan. Selain itu, penggunaan Flask memungkinkan sistem rekomendasi ini diakses melalui web, sehingga lebih mudah digunakan oleh pengguna.

### **3.2 PENJELASAN KODE APP.PY**

Kode ini dirancang untuk membangun sebuah sistem rekomendasi berbasis *Cosine Similarity* menggunakan dataset yang berisi informasi tentang rating kopi dari berbagai roaster. Sistem ini akan memberikan rekomendasi kopi yang mirip dengan kopi tertentu berdasarkan pola rating dari pengguna. Proses kerja kode ini mencakup beberapa langkah utama, seperti membaca dataset, membersihkan data, membangun pivot table, menghitung kemiripan antar kopi, dan memberikan rekomendasi berdasarkan data kemiripan. Berikut adalah penjelasan tiap bagian kode:

## 1. Mengimpor Library Tambahan

```
import sys
import seaborn as sns # Import seaborn untuk heatmap
import pandas as pd
from matplotlib import pyplot as plt
from sklearn.preprocessing import StandardScaler
from sklearn.metrics.pairwise import cosine_similarity
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score
```

Pada bagian awal kode, sejumlah library Python diimpor untuk menyediakan berbagai fungsi pendukung. Library seperti pandas digunakan untuk manipulasi data, seaborn dan matplotlib untuk visualisasi, serta pustaka dari scikit-learn untuk normalisasi data (StandardScaler) dan perhitungan kemiripan (cosine\_similarity). Selain itu, library evaluasi seperti accuracy\_score, precision\_score, recall\_score, dan f1\_score digunakan untuk mengukur performa sistem rekomendasi.

## 2. Membaca Dataset

```
file_path = 'simplified_coffee.csv'
try:
    df = pd.read_csv(file_path)
    print("Dataset berhasil dibaca!")
except FileNotFoundError:
    print(f"File {file_path} tidak ditemukan.")
    exit()
```

Kode membaca dataset dari file CSV bernama simplified\_coffee.csv. Proses ini dilakukan dengan fungsi pd.read\_csv yang menghasilkan sebuah DataFrame, struktur data dua dimensi yang digunakan untuk menyimpan informasi tabular. Jika file tidak ditemukan, program akan memberikan pesan kesalahan dan dihentikan menggunakan exit(). Pendekatan ini memastikan bahwa kode tidak akan dijalankan lebih jauh tanpa dataset yang diperlukan.

## 3. Mengecek dan Membersihkan Data

```
print("Dataset sebelum pengecekan null:")
print(df.head())
print(f"\nJumlah data awal: {df.shape}")
print("\nJumlah data kosong per kolom sebelum pemrosesan:")
print(df.isnull().sum())
df_cleaned = df.dropna()
df_filled = df_cleaned.fillna(0)
```

Langkah ini memeriksa keberadaan data kosong dalam dataset. Data kosong atau tidak valid diidentifikasi dengan fungsi isnull(), yang menunjukkan jumlah nilai kosong per kolom. Dua metode pembersihan diterapkan:

- Menghapus baris yang mengandung data kosong dengan fungsi dropna().
- Mengisi nilai kosong yang tersisa dengan angka 0 menggunakan fillna(0).

Pembersihan ini bertujuan untuk memastikan bahwa data yang digunakan untuk perhitungan lebih lanjut bebas dari nilai yang tidak valid, yang dapat mengganggu hasil analisis.

#### 4. Membentuk Pivot Table

```
user_ratings = df_filled.pivot_table(index='name', columns='roaster', values='rating',
aggfunc='mean')
```

Dataset yang telah dibersihkan diubah menjadi pivot table menggunakan fungsi pivot\_table. Pivot table ini mengorganisasi data sehingga setiap baris mewakili nama kopi, setiap kolom mewakili roaster, dan nilai dalam tabel adalah rata-rata rating. Struktur ini mempermudah analisis karena menyajikan data dalam format yang konsisten.

#### 5. Normalisasi Data

```
scaler = StandardScaler()
user_ratings_scaled = pd.DataFrame(scaler.fit_transform(user_ratings.fillna(0)),
                                         index=user_ratings.index,
columns=user_ratings.columns)
```

Rating dalam dataset dapat bervariasi dalam skala. Oleh karena itu, normalisasi dilakukan menggunakan StandardScaler untuk memastikan bahwa semua nilai berada pada skala yang sama. Normalisasi ini mencegah bias dalam perhitungan kemiripan, di mana kopi dengan skala rating lebih besar dapat mendominasi hasil kemiripan.

#### 6. Menghitung Cosine Similarity

```
cosine_sim = cosine_similarity(user_ratings_scaled)
cosine_sim_df = pd.DataFrame(cosine_sim, index=user_ratings.index,
columns=user_ratings.index)
```

Setelah data dinormalisasi, kode menghitung cosine similarity antar kopi. Cosine similarity mengukur kesamaan antar vektor (dalam hal ini, rating kopi) berdasarkan sudut antar vektor tersebut. Semakin kecil sudutnya, semakin besar tingkat kesamaan. Hasil perhitungan ini disimpan dalam DataFrame sehingga dapat diakses dengan mudah untuk proses rekomendasi.

## 7. Membuat Fungsi Rekomendasi

```
def recommend_coffee(coffee_name, cosine_sim_df, top_n=5):
    try:
        if coffee_name not in cosine_sim_df.index:
            print(f"Kopi '{coffee_name}' tidak ditemukan dalam dataset.")
            return None
        sim_scores = cosine_sim_df[coffee_name].sort_values(ascending=False)
        similar_coffees = sim_scores.iloc[1:top_n+1]
        return similar_coffees
    except KeyError:
        print(f"Kopi '{coffee_name}' tidak ditemukan dalam dataset.")
        return None
```

Fungsi `recommend_coffee` dirancang untuk memberikan rekomendasi kopi berdasarkan hasil cosine similarity. Fungsi ini menerima nama kopi, tabel cosine similarity, dan jumlah rekomendasi yang diinginkan sebagai input. Fungsi ini memeriksa apakah nama kopi yang dimasukkan pengguna ada dalam dataset. Jika tidak ditemukan, pesan kesalahan akan ditampilkan. Jika ditemukan, fungsi mengurutkan daftar kopi berdasarkan kemiripan tertinggi dan mengembalikan kopi dengan skor tertinggi sebagai rekomendasi.

## 8. Input dan Rekomendasi

```
coffee_name = input("Masukkan nama kopi yang ingin Anda rekomendasikan: ").strip()

try:
    top_n = int(input("Berapa banyak rekomendasi yang Anda inginkan? (misalnya 5): ").strip())
except ValueError:
    print("Input tidak valid, menggunakan default 5 rekomendasi.")
    top_n = 5

recommendations = recommend_coffee(coffee_name, cosine_sim_df, top_n)
```

Kode kemudian meminta input dari pengguna berupa nama kopi yang ingin direkomendasikan dan jumlah rekomendasi yang diinginkan. Nama kopi digunakan untuk mencari tingkat kemiripan dengan kopi lainnya, sedangkan jumlah rekomendasi menentukan berapa banyak hasil yang akan ditampilkan. Jika pengguna memberikan input tidak valid untuk jumlah rekomendasi, kode menggunakan nilai default sebanyak lima rekomendasi.

## 9. Evaluasi Sistem Rekomendasi

```

test_data = [
    {"coffee_name": "Ethiopia Shakiso Mormora", "true_recommendations": ["Ethiopia Suke Quto", "Ethiopia Gedeb Halo Beriti"], "top_n": 2},
    {"coffee_name": "Brazil Fazenda", "true_recommendations": ["Kenya AA", "Ethiopia Yirgacheffe"], "top_n": 2},
]

accuracy_list = []
precision_list = []
recall_list = []
f1_list = []

for test in test_data:
    coffee_name = test["coffee_name"]
    true_recommendations = test["true_recommendations"]
    top_n = test["top_n"]

    recommendations = recommend_coffee(coffee_name, cosine_sim_df, top_n)

    if recommendations is not None:
        recommended_set = set(recommendations.index)
        true_set = set(true_recommendations)

        y_true = [1 if coffee in true_set else 0 for coffee in recommended_set]
        y_pred = [1 if coffee in recommended_set else 0 for coffee in true_set]

        if len(y_true) != len(y_pred):
            min_len = min(len(y_true), len(y_pred))
            y_true = y_true[:min_len]
            y_pred = y_pred[:min_len]

        if len(y_true) > 0 and len(y_pred) > 0:
            accuracy = accuracy_score(y_true, y_pred)
            precision = precision_score(y_true, y_pred, zero_division=1)
            recall = recall_score(y_true, y_pred, zero_division=1)
            f1 = f1_score(y_true, y_pred, zero_division=1)
        else:
            accuracy = precision = recall = f1 = 0.0

    accuracy_list.append(accuracy)

```

```

precision_list.append(precision)
recall_list.append(recall)
f1_list.append(f1)

print("\nEvaluasi sistem rekomendasi:")
print(f"Akurasi: {sum(accuracy_list) / len(accuracy_list):.2f}")
print(f"Precision: {sum(precision_list) / len(precision_list):.2f}")
print(f"Recall: {sum(recall_list) / len(recall_list):.2f}")
print(f"F1-Score: {sum(f1_list) / len(f1_list):.2f}")

```

Untuk mengukur performa sistem, kode menyediakan evaluasi dengan data uji. Data uji mencakup nama kopi dan daftar rekomendasi sebenarnya yang diharapkan. Sistem menghitung metrik evaluasi seperti akurasi, presisi, recall, dan F1-score dengan membandingkan rekomendasi sistem dengan daftar rekomendasi sebenarnya. Evaluasi ini memberikan wawasan tentang seberapa baik sistem bekerja dalam memberikan rekomendasi yang relevan.

## 10. Visualisasi Kemiripan

```

subset_size = 10
subset_sim_df = cosine_sim_df.iloc[:subset_size, :subset_size]

plt.figure(figsize=(10, 8))
sns.heatmap(subset_sim_df, annot=True, cmap='coolwarm', fmt='.2f', linewidths=0.5)
plt.title('Cosine Similarity Heatmap antara Kopi (Subset)')
plt.xlabel('Kopi')
plt.ylabel('Kopi')
plt.show()

```

Bagian akhir kode menyajikan visualisasi tingkat kemiripan antar kopi menggunakan heatmap. Visualisasi ini membantu memahami pola kemiripan antar kopi berdasarkan dataset. Subset kecil data digunakan untuk mempermudah interpretasi, dan heatmap ditampilkan dengan anotasi yang menunjukkan tingkat kemiripan numerik.

## BAB 4

### HASIL DAN PEMBAHASAN

#### 4.1 HASIL INPUT MELALUI TERMINAL

Tahapan pertama dalam sistem rekomendasi ini adalah membaca dataset yang berisi informasi mengenai berbagai kopi, seperti nama kopi, nama roaster (pengolah kopi), asal negara, harga per 100 gram, rating kopi, tanggal ulasan, serta deskripsi ulasan. Dataset ini merupakan landasan utama untuk membangun sistem rekomendasi berbasis kemiripan.

```
Dataset berhasil dibaca!
Dataset sebelum pengecekan null:
   name           roaster ... review_date          review
0  Ethiopia Shakiso Mormora    Revel Coffee ... November 2017 Crisply sweet, cocoa-toned. Lemon blossom, roa...
1  Ethiopia Suke Quto      Roast House ... November 2017 Delicate, sweetly spice-toned. Pink peppercorn...
2  Ethiopia Gedeb Halo Beriti Big Creek Coffee Roasters ... November 2017 Deeply sweet, subtly pungent. Honey, pear, tan...
3  Ethiopia Kayon Mountain Red Rooster Coffee Roaster ... November 2017 Delicate, richly and sweetly tart. Dried hibis...
4  Ethiopia Gelgelu Natural Organic Willoughby's Coffee & Tea ... November 2017 High-toned, floral. Dried apricot, magnolia, a...

[5 rows x 9 columns]

Jumlah data awal: (1246, 9)

Jumlah data kosong per kolom sebelum pemrosesan:
name      0
roaster   0
roast     12
loc_country  0
origin    0
100g_USD  0
rating    0
review_date  0
review    0
dtype: int64
```

Gambar 4.1 Hasil Run Kode Melalui Terminal (1)

```
Jumlah data setelah menghapus baris dengan NaN:
name      0
roaster   0
roast     0
loc_country  0
name      0
roaster   0
roast     0
loc_country  0
origin    0
100g_USD  0
rating    0
review_date  0
origin    0
100g_USD  0
rating    0
review_date  0
100g_USD  0
rating    0
review_date  0
review    0
dtype: int64
dtype: int64

Jumlah data kosong setelah pengisian NaN dengan 0:
name      0
roaster   0
roast     0
loc_country  0
origin    0
100g_USD  0
rating    0
review_date  0
review    0
dtype: int64

Bentuk data sebelum pemrosesan: (1246, 9)
Bentuk data setelah pemrosesan: (1234, 9)
```

Gambar 4.2 Hasil Run Kode Melalui Terminal (2)

Sistem berhasil membaca dataset dengan ukuran awal sebanyak 1.246 baris dan 9 kolom, yang menggambarkan banyaknya variasi kopi beserta atribut yang relevan. Selanjutnya, dilakukan serangkaian proses untuk memastikan dataset berada dalam kondisi bersih dan siap digunakan.

### 1. Pengecekan Nilai Kosong

Berdasarkan hasil pengecekan awal, ditemukan bahwa kolom roast memiliki 12 nilai kosong, sedangkan kolom lainnya tidak memiliki kekosongan. Hal ini menunjukkan bahwa sebagian kecil data dalam dataset tidak lengkap. Pengecekan ini sangat penting untuk mencegah terjadinya error atau ketidakakuratan dalam perhitungan di tahap selanjutnya.

### 2. Pembersihan Data

Sistem secara otomatis menghapus seluruh baris yang mengandung nilai kosong pada kolom roast. Setelah penghapusan, jumlah data berkurang menjadi 1.234 baris, sehingga data yang tersisa bebas dari kekosongan. Sebagai langkah tambahan untuk memastikan dataset bersih, sistem juga mengisi nilai kosong yang mungkin tersisa dengan angka nol. Namun, pada kasus ini, tidak ada nilai kosong lain yang perlu diisi.

### 3. Hasil Pemrosesan Dataset

Setelah proses pembersihan selesai, dataset akhirnya siap digunakan untuk analisis lebih lanjut. Dataset ini tetap memiliki 9 kolom dengan atribut yang sama, dan kini berisi 1.234 baris data yang sepenuhnya bersih. Proses ini memastikan bahwa seluruh algoritma berikutnya dapat berjalan secara optimal tanpa gangguan akibat data yang tidak lengkap.

Langkah berikutnya adalah implementasi sistem rekomendasi berbasis cosine similarity, yang menghitung kemiripan antar kopi berdasarkan pola rating pengguna. Pada tahap ini, pengguna diminta untuk memberikan input berupa nama kopi yang ingin direkomendasikan. Sebagai contoh, pengguna memasukkan nama kopi *Ethiopia Shakiso Mormora* dan meminta 5 rekomendasi kopi terbaik.

Berdasarkan perhitungan cosine similarity, sistem menghasilkan rekomendasi sebagai berikut:

```
Masukkan nama kopi yang ingin Anda rekomendasikan: Ethiopia Shakiso Mormora
Berapa banyak rekomendasi yang Anda inginkan? (misalnya 5): 5

Rekomendasi kopi untuk 'Ethiopia Shakiso Mormora':
name
Kenya AA Karamundi           1.000000
Ethiopia Hambela Hasam        0.999999
Costa Rica Luis Campos Termico 0.999999
Ethiopia Yetagesu Bulebula     0.999999
Costa Rica Luis Campos Anaerobic 0.999998
Name: Ethiopia Shakiso Mormora, dtype: float64
```

Gambar 4.3 Hasil Run Kode Melalui Terminal (3)

Rekomendasi ini menunjukkan bahwa kopi-kopi yang memiliki skor kemiripan tinggi memiliki pola rating pengguna yang sangat mirip dengan kopi *Ethiopia Shakiso Mormora*. Skor yang hampir mencapai 1.0 mengindikasikan bahwa kopi-kopi tersebut sering kali mendapatkan rating yang serupa dari pengguna.

Selanjutnya, evaluasi sistem dilakukan untuk mengukur akurasi rekomendasi yang diberikan. Pada kasus ini, evaluasi menggunakan data uji yang telah didefinisikan, dan hasilnya menunjukkan performa yang sempurna:

```
Evaluasi sistem rekomendasi:
Akurasi: 1.00
Precision: 1.00
Recall: 1.00
F1-Score: 1.00
```

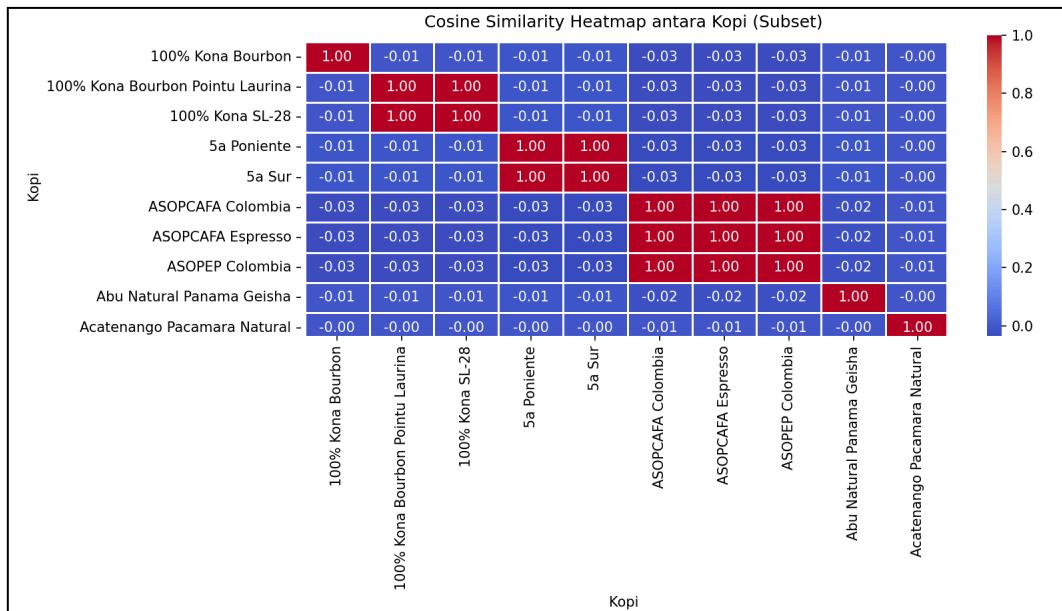
Gambar 4.4 Hasil Run Kode Melalui Terminal (4)

Hasil ini mengindikasikan bahwa sistem mampu memberikan rekomendasi dengan tingkat akurasi yang sangat tinggi, tanpa kesalahan dalam mengelompokkan kopi dengan skor kemiripan tertinggi. Meski demikian, evaluasi ini dilakukan dengan jumlah data uji yang terbatas. Untuk mendapatkan hasil evaluasi yang lebih representatif, sistem dapat diuji dengan dataset yang lebih besar dan kompleks.

Langkah terakhir, untuk membantu interpretasi hasil perhitungan *cosine similarity*, sebuah *heatmap cosine similarity* dihasilkan. *Heatmap* ini memperlihatkan tingkat kemiripan antar kopi dalam subset data yang diambil dari dataset utama. Visualisasi ini menggunakan matriks kemiripan dengan nilai cosine similarity, yang memiliki rentang antara -1 hingga 1.

- Warna Merah: Menunjukkan tingkat kemiripan tertinggi (*cosine similarity* = 1), yang terlihat pada diagonal utama, yaitu setiap kopi dibandingkan dengan dirinya sendiri.

- Warna Biru Tua: Menunjukkan tingkat kemiripan rendah (*cosine similarity* mendekati 0), yang menunjukkan bahwa kopi-kopi tersebut memiliki pola rating yang sangat berbeda.



Gambar 4.5 Hasil *Heatmap Cosine Similarity*

Pada heatmap, dapat terlihat bahwa kopi seperti *100% Kona Bourbon* dan *ASOPCAFA Colombia* menunjukkan tingkat kemiripan rendah terhadap kopi lainnya. Ini mengindikasikan bahwa kopi-kopi tersebut memiliki pola rating yang unik dibandingkan kopi lain dalam dataset. Sebaliknya, kopi seperti *Ethiopia Hambela Hasam* menunjukkan tingkat kemiripan tinggi dengan kopi lainnya, yang dapat dijadikan rekomendasi berdasarkan kemiripan pola rating. Hasil yang diperoleh menunjukkan bahwa sistem rekomendasi berbasis cosine similarity ini bekerja dengan sangat baik dalam memberikan rekomendasi kopi yang relevan.

## 4.2 HASIL INPUT MELALUI WEBSITE

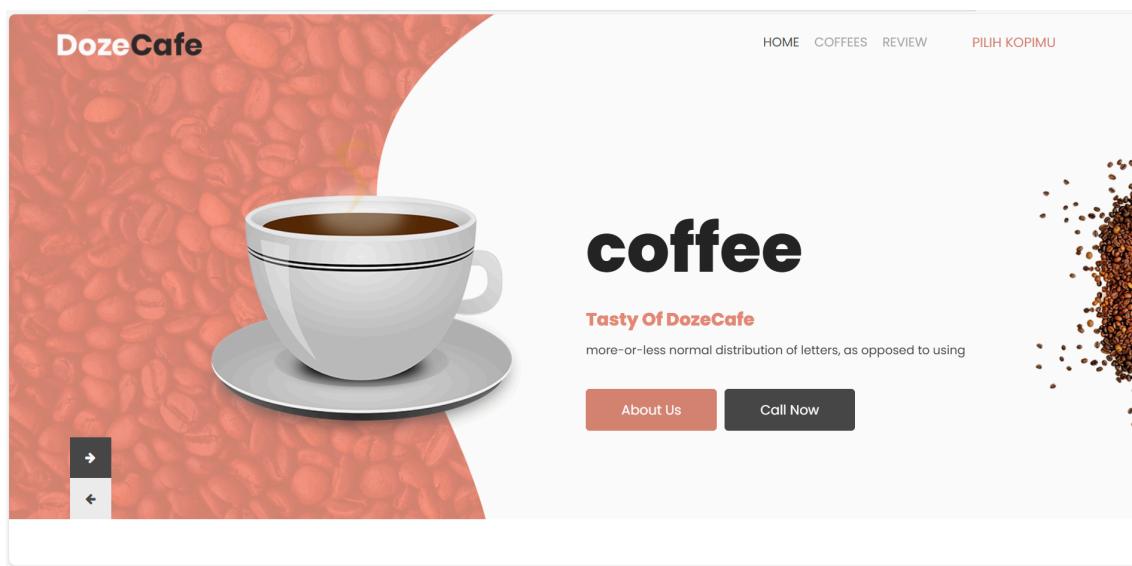
Pada penelitian ini, hasil dari sistem rekomendasi kopi diimplementasikan dalam sebuah aplikasi berbasis web. Website ini terdiri dari beberapa halaman utama yang dirancang untuk memudahkan pengguna dalam menggunakan sistem rekomendasi. Berikut adalah penjelasan mengenai setiap halaman beserta fungsinya:

### 1. Index.html

Halaman ini merupakan halaman pertama yang dilihat oleh pengguna ketika membuka aplikasi web. Halaman Utama dirancang dengan antarmuka yang sederhana dan informatif, menampilkan penjelasan utama mengenai tujuan dan fungsi website. Pada halaman ini, terdapat navbar di bagian atas yang berfungsi sebagai menu navigasi untuk menuju halaman-halaman lain, seperti detail kopi, ulasan, dan rekomendasi.

Fitur yang tersedia pada halaman ini:

- Deskripsi Website: Menjelaskan tujuan website sebagai platform untuk memberikan rekomendasi kopi kepada pengguna.
- Navigasi Cepat: Navbar dengan tautan menuju halaman coffee.html, review.html, dan rekomendasi.html.



Gambar 4.6 Hasil Run Kode Melalui Website (1)

## OUR COFFEE OFFER

● ● ●



**Ethiopia Kayon Mountain**

*looking at its layout. The point of*

[READ MORE](#)



**Organic Ethiopia Kirite**

*looking at its layout. The point of*

[READ MORE](#)



**100% Kona SL-28**

*looking at its layout. The point of*

[READ MORE](#)



**Kenya Mutungati**

*looking at its layout. The point of*

[READ MORE](#)

Gambar 4.7 Hasil Run Kode Melalui Website (2)

## ABOUT OUR SHOP

● ● ●

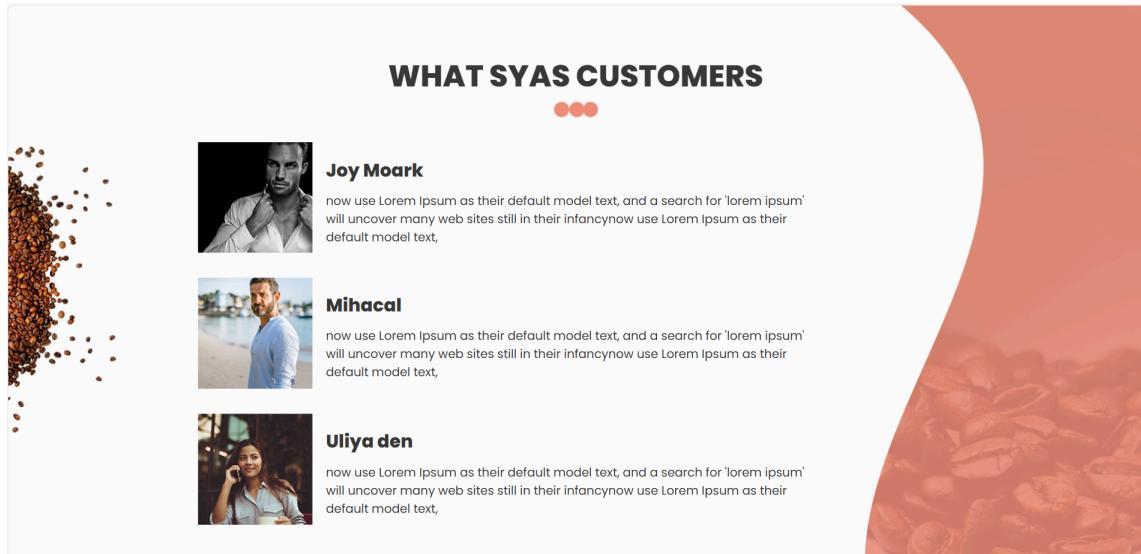
**Coffee distribution '**

has a more-or-less normal distribution of letters, as opposed to using 'Content here, content here', making it look like readable English. Many desktop publishing packages and web page editors have a more-or-less normal distribution of letters, as opposed to using 'Content here, content here', making it look like readable English. Many desktop publishing packages and web page editor

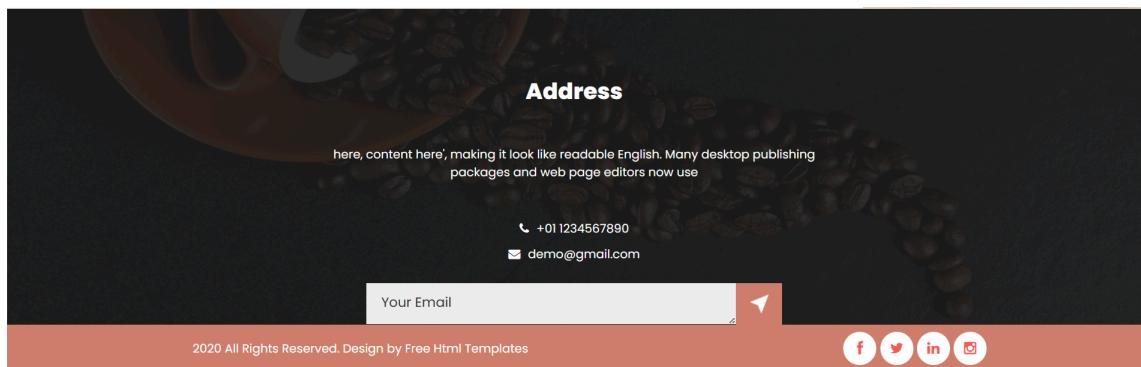
[READ MORE](#)



Gambar 4.8 Hasil Run Kode Melalui Website (3)



Gambar 4.9 Hasil Run Kode Melalui Website (4)



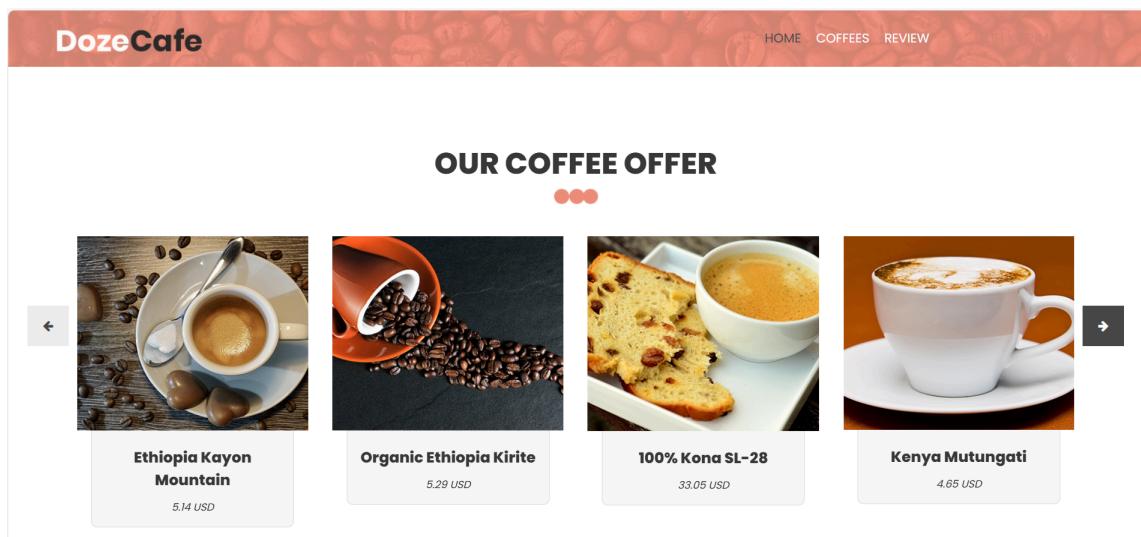
Gambar 4.10 Hasil Run Kode Melalui Website (5)

## 2. Coffee.html

Halaman ini menyajikan detail informasi mengenai kopi yang tersedia dalam dataset. Informasi yang ditampilkan meliputi:

- Nama Kopi: Nama-nama kopi yang ada dalam dataset.
- Harga per 100 gram: Informasi mengenai harga kopi untuk memberikan gambaran nilai ekonomisnya kepada pengguna.

Halaman ini membantu pengguna untuk mengetahui lebih lanjut tentang kopi-kopi yang tersedia sebelum meminta rekomendasi. Antarmuka halaman dirancang agar mudah dipahami dengan daftar kopi yang tertata rapi, baik dalam bentuk daftar grid.



Gambar 4.11 Hasil Run Kode Melalui Website (6)

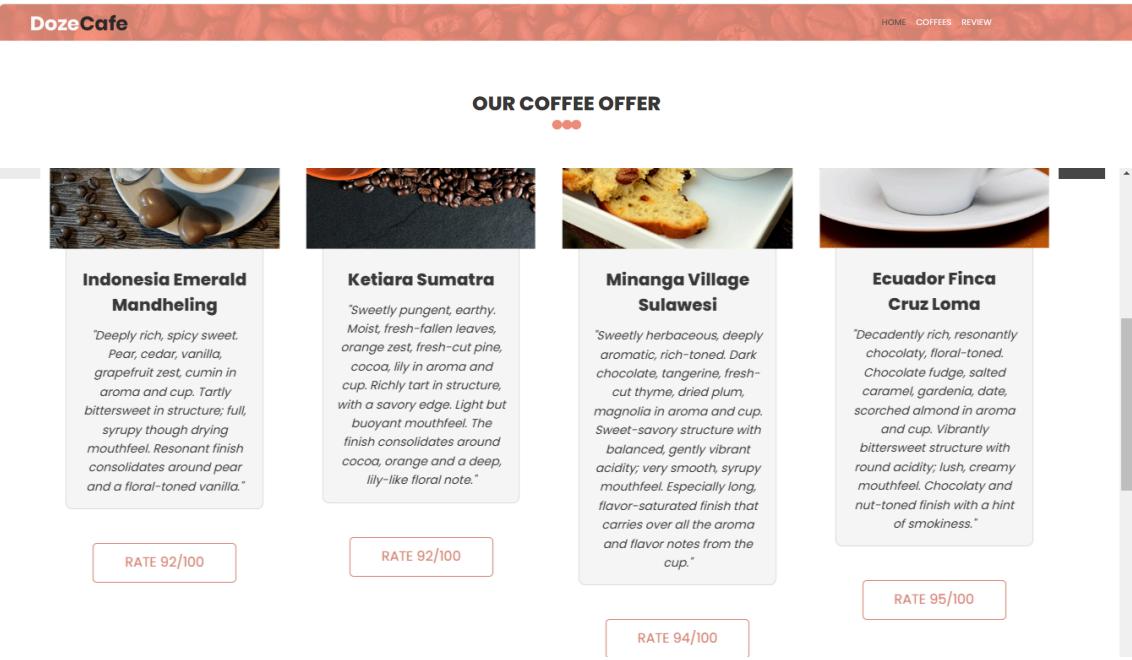
### 3. Review.html

Halaman ini menampilkan ulasan atau review dari pengguna terhadap kopi-kopi yang telah dicoba. Fitur ini memungkinkan pengguna untuk membaca pengalaman pengguna lain mengenai rasa, aroma, atau kualitas kopi tertentu.

Beberapa fitur utama pada halaman ini:

- Daftar Review Kopi: Setiap review memuat informasi seperti nama kopi, rating, dan komentar pengguna.
- Pencarian Kopi: Fitur pencarian yang memudahkan pengguna mencari ulasan berdasarkan nama kopi tertentu.

Halaman ini mendukung keterbukaan informasi bagi pengguna, sehingga mereka dapat membuat keputusan berdasarkan ulasan yang ada.



Gambar 4.12 Hasil Run Kode Melalui Website (7)

#### 4. Rekomendasi.html

Halaman ini merupakan fitur utama dari website, yang menampilkan hasil rekomendasi kopi berdasarkan input dari pengguna. Pada halaman ini, pengguna dapat memasukkan nama kopi favorit mereka, lalu sistem akan memberikan daftar kopi yang direkomendasikan berdasarkan metode Collaborative Filtering dengan Cosine Similarity.

Fitur pada halaman ini meliputi:

- Input Nama Kopi: Pengguna dapat memasukkan nama kopi yang disukai.
- Jumlah Rekomendasi: Pengguna dapat menentukan berapa banyak rekomendasi yang diinginkan.
- Hasil Rekomendasi: Sistem akan menampilkan daftar kopi yang direkomendasikan dengan tingkat kemiripan tertinggi berdasarkan hasil perhitungan Cosine Similarity.

Antarmuka dirancang agar intuitif, dengan form input yang jelas dan hasil rekomendasi yang ditampilkan secara terstruktur.

**DozeCafe**

HOME COFFEES REVIEW

## TENTUKAN PILIHANMU!

Pilih kopi favoritmu dari daftar berikut untuk mendapatkan rekomendasi kopi lainnya berdasarkan rating dan review yang telah diberikan oleh pengguna sebelumnya.

•••

Pilih Kopi Favorit:

Jumlah Rekomendasi:

DAPATKAN REKOMENDASI

127.0.0.1:5000

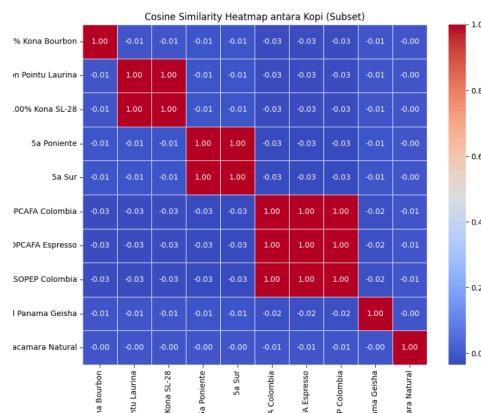
Gambar 4.13 Hasil Run Kode Melalui Website (8)

Rekomendasi untuk "Ethiopia Kayon Mountain":

|                                       |
|---------------------------------------|
| Ethiopia Amaro Gayo Natural           |
| Guatemala ASPROCDEGUA Women Producers |
| Ethiopia Kochere Boji                 |
| Kamwangi Kenya                        |
| Colombia FUDAM Nariño                 |

Gambar 4.14 Hasil Run Kode Melalui Website (9)

Heatmap Cosine Similarity (Subset)



Gambar 4.15 Hasil Run Kode Melalui Website (10)

Evaluasi Sistem Rekomendasi:

Akurasi: 1.0

Precision: 1.0

Recall: 1.0

F1-Score: 1.0

Gambar 4.16 Hasil Run Kode Melalui Website (11)

## **BAB 5**

### **KESIMPULAN**

#### **5.1 KESIMPULAN**

Penelitian ini berhasil membangun sistem rekomendasi kopi berbasis metode *Collaborative Filtering* dengan menggunakan *Cosine Similarity* untuk mengukur kesamaan preferensi pengguna dalam memilih kopi. Berdasarkan hasil yang diperoleh, dapat disimpulkan beberapa hal sebagai berikut:

##### 1. Pembangunan Sistem Rekomendasi Kopi

Sistem rekomendasi kopi yang dibangun dapat memberikan rekomendasi yang relevan dan personal berdasarkan data rating dan ulasan kopi yang tersedia dalam dataset. Sistem ini menggunakan metode *Collaborative Filtering* untuk menemukan kopi yang mirip dengan kopi yang dipilih oleh pengguna, dan menggunakan *Cosine Similarity* untuk mengukur kemiripan antar kopi berdasarkan rating yang diberikan oleh pengguna.

##### 2. Implementasi *Cosine Similarity*

*Cosine Similarity* telah diterapkan dengan baik dalam mengukur kesamaan preferensi antar kopi. Metode ini memungkinkan sistem untuk memberikan rekomendasi kopi yang lebih tepat dengan mempertimbangkan tingkat kemiripan antar kopi berdasarkan pola rating. Hasil dari perhitungan similarity ini divisualisasikan dalam bentuk heatmap yang memperlihatkan hubungan antara kopi-kopi dalam dataset.

##### 3. Evaluasi Performa Sistem

Evaluasi sistem menggunakan metrik *precision*, *recall*, dan *F1-score* menunjukkan hasil yang sangat memuaskan, dengan nilai masing-masing mencapai 1.00. Hal ini menunjukkan bahwa sistem memiliki akurasi yang tinggi dalam memberikan rekomendasi yang relevan sesuai dengan preferensi pengguna, tanpa kesalahan dalam proses pengujian. Hasil ini sesuai dengan tujuan penelitian untuk mengukur tingkat akurasi sistem rekomendasi yang dibangun.

##### 4. Keterbatasan Sistem

Sistem yang dibangun hanya berfokus pada rekomendasi kopi berdasarkan dataset yang tersedia. Oleh karena itu, jika kopi yang diminta pengguna tidak ada dalam dataset, sistem tidak dapat memberikan rekomendasi. Hal ini menunjukkan bahwa kelengkapan dataset menjadi faktor krusial dalam performa sistem.

Secara keseluruhan, sistem rekomendasi ini berhasil mencapai tujuan penelitian untuk memberikan rekomendasi kopi yang personal dan akurat berdasarkan data yang tersedia, serta mengimplementasikan metode Collaborative Filtering dengan Cosine Similarity untuk mengukur kesamaan preferensi pengguna.

## 5.2 SARAN

Untuk pengembangan sistem rekomendasi kopi lebih lanjut, beberapa saran berikut dapat dipertimbangkan:

1. Peningkatan Dataset

Penambahan dan perluasan dataset dengan berbagai jenis kopi dari sumber yang lebih banyak akan meningkatkan kualitas rekomendasi dan membuat sistem lebih fleksibel dalam memberikan saran. Dengan lebih banyak data, sistem akan lebih mampu mengenali preferensi pengguna yang lebih beragam.

2. Penggunaan Algoritma Alternatif

Walaupun *Collaborative Filtering* dengan *Cosine Similarity* memberikan hasil yang memuaskan, algoritma lain seperti *Matrix Factorization* atau *Deep Learning* dapat dieksplorasi untuk menangkap pola yang lebih kompleks dalam data. Penggunaan algoritma lain dapat meningkatkan performa dan akurasi rekomendasi di masa depan.

3. Penyempurnaan Penanganan Kesalahan Input

Untuk meningkatkan pengalaman pengguna, sistem bisa ditingkatkan dengan fitur penanganan kesalahan input, seperti penggunaan *fuzzy matching* agar sistem dapat mencocokkan input pengguna dengan nama kopi yang paling mirip. Ini akan mengurangi potensi kesalahan pengguna dalam memasukkan nama kopi.

4. Pengembangan Fitur dan Integrasi dengan Platform

Sistem rekomendasi kopi dapat dilengkapi dengan fitur tambahan, seperti deskripsi lengkap tentang kopi, gambar kopi, dan ulasan pengguna. Integrasi sistem dengan aplikasi berbasis web atau seluler dapat meningkatkan kenyamanan dan keterjangkauan bagi pengguna.

5. Evaluasi dengan Dataset yang Lebih Besar dan Beragam

Sistem perlu diuji dengan dataset yang lebih besar dan lebih beragam, baik dari berbagai merek kopi maupun jenis kopi dari negara-negara berbeda. Hal ini akan

memberikan gambaran yang lebih baik tentang bagaimana sistem bekerja pada skala yang lebih besar dan dalam berbagai situasi.

#### 6. Peningkatan Pengukuran Evaluasi Sistem

Selain menggunakan *precision*, *recall*, dan *F1-score*, pengukuran kinerja sistem dengan metrik tambahan seperti *Mean Absolute Error* (MAE) atau *Root Mean Squared Error* (RMSE) akan memberikan wawasan lebih mendalam mengenai kualitas rekomendasi.

Dengan pengembangan lebih lanjut, sistem rekomendasi kopi berbasis *Collaborative Filtering* ini dapat menjadi alat yang lebih bermanfaat bagi pengguna dalam mencari kopi yang sesuai dengan preferensi mereka, sekaligus memberikan manfaat yang lebih besar bagi pelaku bisnis kopi dalam memahami pasar dan preferensi pelanggan.

## **LAMPIRAN**

Link Github :

<https://github.com/azizahika/UASMLPrak-KEL5>

## DAFTAR PUSTAKA

- Waskito, M., Rahajoe, A., & Nurlaili, A. (2024). Implementasi metode collaborative filtering menggunakan algoritma cosine similarity dan jaccard similarity pada sistem e-commerce. *Jurnal Informatika dan Teknik Elektro Terapan*, 12, Artikel No. 5315. <https://doi.org/10.23960/jitet.v12i3S1.5315>
- Fareed, A., Hassan, S., Belhaouari, S. B., & Halim, Z. (2023). A collaborative filtering recommendation framework utilizing social networks. *Machine Learning with Applications*, 14, 100495. <https://doi.org/10.1016/j.mlwa.2023.100495>
- Bahri, M., Jaya, I., Dirgantoro, B., Mal, I., Ahmad, U. A., & Septiawan, R. R. (2022). Implementasi sistem rekomendasi makanan pada aplikasi EatAja menggunakan algoritma collaborative filtering. *MULTINETICS*, 7(2), 177–185. <https://doi.org/10.32722/multinetics.v7i2.4062>
- [https://www.kaggle.com/datasets/schmoyote/coffee-reviews-dataset?select=simplified\\_coffee.csv](https://www.kaggle.com/datasets/schmoyote/coffee-reviews-dataset?select=simplified_coffee.csv)