

LAPORAN PROYEK UAS
PEMBELAJARAN MESIN (PRAKTIKUM)

SISTEM REKOMENDASI MAKANAN BERDASARKAN JENIS MAKANAN
MENGGUNAKAN METODE *COLLABORATIVE FILTERING* DENGAN
PENDEKATAN *COSINE SIMILARITY*



KELOMPOK 1
TI-C4
434221051 | Fairuz Zahira
434221053 | Lintang Sekar Wangi

PROGRAM STUDI D4 TEKNIK INFORMATIKA
FAKULTAS VOKASI
UNIVERSITAS AIRLANGGA
SURABAYA
2024

DAFTAR ISI

DAFTAR ISI.....	1
DAFTAR TABEL.....	2
DAFTAR GAMBAR.....	3
BAB 1	
PENDAHULUAN.....	1
BAB 2	
TINJAUAN PUSTAKA.....	4
BAB 3	
METODOLOGI.....	7
BAB 4	
HASIL DAN PEMBAHASAN.....	18
BAB 5	
KESIMPULAN.....	22
LAMPIRAN.....	23
DAFTAR PUSTAKA.....	26

DAFTAR TABEL

Tabel 1. Dataset food.csv.....	7
Tabel 2. Dataset ratings.csv.....	8
Tabel 3. Output Rekomendasi.....	20

DAFTAR GAMBAR

Gambar 1. Landing Page.....	23
Gambar 2. About Us Page.....	23
Gambar 3. News Page.....	24
Gambar 4. User Input Page.....	24
Gambar 5. Output Page.....	25

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Di era digital ini, kemajuan teknologi telah membawa dampak yang signifikan bagi berbagai sektor kehidupan, termasuk industri kuliner. Dengan kemampuan teknologi untuk mengolah data dalam jumlah besar dan memberikan rekomendasi yang relevan, masyarakat kini dapat lebih mudah memilih makanan yang sesuai dengan preferensi mereka. Sistem rekomendasi telah menjadi solusi yang banyak digunakan untuk membantu pengguna dalam mengambil keputusan berdasarkan data historis dan preferensi individu.

Dalam konteks industri makanan, penerapan sistem rekomendasi dapat memperkaya pengalaman pengguna melalui layanan yang dipersonalisasi. Menurut Yusmar et al. (2021), penggunaan sistem rekomendasi berbasis Collaborative Filtering telah terbukti efektif dalam memberikan saran yang tepat berdasarkan preferensi pengguna lain yang memiliki pola kesamaan. Pendekatan ini tidak hanya meningkatkan kepuasan pelanggan, tetapi juga membuka peluang bisnis baru dengan menyediakan layanan yang lebih adaptif.

Metode Collaborative Filtering, terutama ketika dipadukan dengan Cosine Similarity, adalah salah satu cara yang sering digunakan untuk mengukur kesamaan antar data. Cosine Similarity menghitung sudut kosinus antara vektor dalam ruang multidimensi untuk menentukan tingkat kemiripan. Penelitian yang dilakukan oleh Sari dan Hidayat (2021) menunjukkan bahwa metode ini sangat efisien dalam mengolah data besar, terutama pada sistem rekomendasi yang memiliki banyak item dan pengguna.

Dalam proyek ini, kami mengembangkan sistem rekomendasi makanan berbasis Collaborative Filtering dengan pendekatan Cosine Similarity. Sistem ini dirancang untuk

memberikan rekomendasi makanan yang disesuaikan dengan data preferensi pengguna sebelumnya, serta mengelompokkan makanan ke dalam kategori seperti ‘Healthy Food’, ‘Snack’, dan ‘Dessert’, serta klasifikasi Veg dan Non-Veg. Dengan demikian, sistem dapat menawarkan rekomendasi yang lebih personal, relevan, dan sesuai dengan kebutuhan pelanggan.

1.2 Rumusan Masalah

Dalam proyek ini, rumusan masalah yang akan diangkat adalah sebagai berikut:

1. Bagaimana cara mengimplementasikan metode Collaborative Filtering dengan pendekatan Cosine Similarity untuk menghasilkan rekomendasi makanan yang tepat dan relevan?
2. Apa saja langkah-langkah yang dapat diambil untuk memanfaatkan data preferensi pengguna demi meningkatkan kualitas rekomendasi yang diberikan?
3. Bagaimana sistem dapat memberikan saran makanan yang sesuai berdasarkan kategori dan karakteristik tertentu?

1.3 Tujuan

Proyek ini bertujuan untuk mencapai beberapa hal penting:

1. Mengembangkan sistem rekomendasi makanan yang menggunakan metode Collaborative Filtering dengan pendekatan Cosine Similarity.
2. Menawarkan rekomendasi makanan yang sesuai dengan preferensi pengguna, berdasarkan kategori dan karakteristik makanan.
3. Meningkatkan pengalaman pengguna dengan memberikan rekomendasi yang personal dan relevan.

1.4 Manfaat

Proyek ini menawarkan berbagai manfaat yang signifikan, antara lain:

1. Bagi Pengguna: Memudahkan dalam menemukan makanan yang sesuai dengan preferensi, kebutuhan, dan pola diet individu masing-masing.
2. Bagi Industri Kuliner: Menyediakan peluang untuk meningkatkan efisiensi layanan serta memperkuat loyalitas pelanggan melalui pengalaman yang lebih personal.
3. Bagi Pengembangan Teknologi: Menjadi acuan dalam penerapan metode Collaborative Filtering dengan pendekatan Cosine Similarity pada dataset yang berfokus pada makanan.

1.5 Batasan Masalah

Untuk memastikan fokus dan pencapaian tujuan proyek, terdapat beberapa batasan yang perlu diperhatikan:

1. Keterbatasan Dataset: Dataset yang digunakan hanya mencakup data makanan tertentu yang terdapat dalam file food.csv, serta data interaksi pengguna yang ada di file ratings.csv.
2. Metode yang Digunakan: Sistem ini mengandalkan metode Collaborative Filtering dengan pendekatan Cosine Similarity. Pendekatan ini tidak dibandingkan dengan algoritma lainnya.
3. Lingkup Rekomendasi: Rekomendasi yang dihasilkan hanya mencakup kategori makanan yang ada dalam dataset, seperti Healthy Food, Snack, dan Dessert.
4. Karakteristik Pengguna: Sistem ini tidak mengambil faktor demografis pengguna, seperti usia atau lokasi, dalam proses pemberian rekomendasi.

BAB 2

TINJAUAN PUSTAKA

Bab ini membahas landasan teori serta komponen-komponen yang mendukung pengembangan sistem rekomendasi makanan. Sistem ini dirancang untuk memberikan rekomendasi berdasarkan jenis makanan yang dipilih oleh pengguna, dengan menggunakan metode *Collaborative Filtering* yang dipadukan dengan pendekatan *Cosine Similarity*.

2.1 Jenis Makanan

Jenis makanan dapat dikelompokkan berdasarkan beberapa kategori seperti makanan sehat, makanan ringan, dan makanan penutup. Berdasarkan data yang dianalisis, makanan dalam dataset dikelompokkan menjadi beberapa jenis:

- **Makanan Sehat (*Healthy Food*):** Contoh makanan dalam kategori ini meliputi salad labu musim panas dan salad ayam cincang. Makanan ini biasanya terdiri dari bahan-bahan segar seperti sayuran, daging tanpa lemak, dan rempah-rempah alami.
- **Makanan Ringan (*Snack*):** Kategori ini mencakup makanan seperti kacang *almond* pedas manis yang sering disajikan sebagai camilan sehat dan praktis.
- **Makanan Penutup (*Dessert*):** Contohnya adalah kue Natal yang dibuat dengan buah-buahan kering, rempah-rempah, dan bahan-bahan manis lainnya.

Jenis makanan juga diklasifikasikan berdasarkan karakteristik seperti "Veg" untuk makanan vegetarian dan "Non-Veg" untuk makanan yang mengandung produk hewani. Misalnya, salad ayam cincang termasuk dalam kategori "Non-Veg," sedangkan salad labu musim panas adalah "Veg."

Pengelompokan ini penting dalam sistem rekomendasi makanan untuk mempermudah pengguna dalam menemukan makanan yang sesuai dengan preferensi mereka. Sistem yang

dirancang dengan kategori yang jelas akan lebih efektif dalam memberikan saran yang relevan bagi pengguna dengan kebutuhan dan selera yang berbeda.

2.2 *Machine Learning*

Machine learning adalah cabang kecerdasan buatan yang memungkinkan komputer belajar dari data untuk membuat keputusan tanpa diprogram secara eksplisit. Menurut Prasetyo dan Setiawan (2020), algoritma *machine learning* seperti *K-Nearest Neighbor* (KNN), *Decision Tree*, dan *Support Vector Machine* (SVM) banyak digunakan dalam pengembangan aplikasi berbasis data, termasuk sistem rekomendasi makanan. Aplikasi ini memanfaatkan data historis pengguna untuk mempelajari pola preferensi dan memberikan rekomendasi yang lebih personal. Dalam studi mereka, Prasetyo dan Setiawan (2020) menemukan bahwa penerapan machine learning dalam industri kuliner mampu meningkatkan efisiensi pelayanan dan kepuasan pelanggan melalui prediksi yang lebih akurat.

2.3 *Collaborative Filtering*

Collaborative filtering adalah teknik dalam sistem rekomendasi yang memanfaatkan data dari interaksi pengguna lain untuk memberikan saran yang lebih personal. Menurut Nugraha et al. (2022), metode ini sering digunakan dalam aplikasi *e-commerce* dan layanan *streaming*, tetapi juga memiliki potensi besar dalam dunia kuliner. Sistem ini bekerja dengan menganalisis preferensi pengguna yang memiliki kesamaan dalam memilih produk tertentu. Misalnya, jika dua pengguna memiliki riwayat pemesanan makanan yang serupa, aplikasi dapat merekomendasikan makanan yang disukai salah satu pengguna kepada pengguna lainnya (Nugraha et al., 2022). Teknik ini terbukti efektif dalam meningkatkan keterlibatan pengguna dan memperluas jangkauan layanan aplikasi.

2.4 Cosine Similarity

Cosine similarity adalah metode yang digunakan untuk mengukur kesamaan antara dua vektor dalam ruang multidimensi berdasarkan sudut kosinus antar-vektor tersebut. Menurut Sari dan Hidayat (2021), metode ini digunakan dalam sistem rekomendasi untuk menghitung tingkat kesamaan antara preferensi pengguna dan item yang tersedia dalam basis data. Semakin tinggi nilai *cosine similarity*, semakin besar kesamaan antara dua entitas yang dibandingkan. Sari dan Hidayat (2021) menjelaskan bahwa penerapan metode ini sangat berguna dalam menyusun rekomendasi makanan, terutama dalam kasus di mana jumlah data yang dikelola sangat besar. Dengan menghitung kesamaan antara vektor preferensi pengguna dan deskripsi makanan dalam sistem, aplikasi dapat menyarankan makanan yang sesuai dengan selera pengguna secara akurat.

Dengan tinjauan ini, dapat disimpulkan bahwa pengelompokan jenis makanan, penerapan algoritma machine learning, penggunaan teknik *collaborative filtering*, dan metode *cosine similarity* merupakan komponen penting dalam pengembangan sistem rekomendasi makanan yang cerdas dan adaptif. Kombinasi dari teknologi-teknologi ini memungkinkan aplikasi untuk memberikan pengalaman yang lebih personal dan relevan bagi pengguna.

BAB 3

METODOLOGI

3.1 Pendekatan Proyek

Proyek ini bertujuan untuk mengembangkan sistem rekomendasi makanan berbasis Collaborative Filtering dengan pendekatan Cosine Similarity. Sistem ini akan membantu pengguna menemukan makanan yang sesuai dengan preferensi mereka, dengan mempertimbangkan data interaksi pengguna sebelumnya berupa rating yang diberikan pada berbagai jenis makanan. Pendekatan Collaborative Filtering dipilih karena efektif dalam memanfaatkan pola preferensi pengguna untuk memberikan rekomendasi yang lebih personal dan relevan. Dalam proyek ini, kami menggunakan dua dataset utama :

1. food.csv

Dataset ini berisi informasi tentang makanan, termasuk nama makanan, kategori (misalnya Healthy Food, Snack, Dessert), dan deskripsi bahan makanan yang digunakan, meliputi:

- ID Makanan (Food_ID): Sebuah identifikasi unik untuk setiap jenis makanan.
- Nama Makanan: Penamaan deskriptif untuk masing-masing makanan yang tersedia.
- Kategori Makanan: Keterangan mengenai apakah makanan tersebut termasuk dalam kategori Healthy Food, Snack, atau Dessert.
- Karakteristik Makanan: Informasi yang mengklasifikasikan makanan sebagai vegetarian (Veg) atau non-vegetarian (Non-Veg).
- Deskripsi: Rincian tentang bahan-bahan yang digunakan dalam pembuatan makanan tersebut.

Tabel 1. Dataset food.csv

Food_ID	Name	C_Type	Veg_Non	Describe
1	summer squash salad	Healthy Food	veg	white balsamic vinegar, lemon juice, lemon rind, red chillies, garlic cloves (crushed), olive oil, summer squash (zucchini), sea salt, black pepper, basil leaves
2	chicken minced salad	Healthy Food	non-veg	olive oil, chicken mince, garlic (minced), onion, salt, black pepper, carrot, cabbage, green onions, sweet chilli sauce, peanut butter, ginger, soy sauce, fresh cilantro, red pepper flakes (crushed), tarts
3	sweet chilli almonds	Snack	veg	almonds whole, egg white, curry leaves, salt, sugar (fine grain), red chilli powder

4	tricolour salad	Healthy Food	veg	vinegar, honey/sugar, soy sauce, salt, garlic cloves (minced), chilli pepper (sliced), green papaya, carrot (peeled), cucumbers, mint leaves, toasted peanuts
5	christmas cake	Dessert	veg	christmas dry fruits (pre-soaked), orange zest, lemon zest, jaggery syrup, almond flour, apple, butter (softened), eggs
...				
396	Kimchi Toast	Korean	veg	cream cheese, chopped kimchi, scallions, country-style bread, sesame seeds, cilantro leaves
397	Tacos de Gobernador (Shrimp, Poblano, and Cheese Tacos)	Mexican	non-veg	poblano chiles, bacon, shrimps, red salsa, garlic, corn tortillas, lime juice
398	Melted Broccoli Pasta With Capers and Anchovies	French	non-veg	broccoli, Bread Crumbs, anchovy fillets, garlic cloves, red pepper flakes, penne pasta, olive oil
399	Lemon-Ginger Cake with Pistachios	Dessert	non-veg	egg yolks, lemon juice, unsalted butter, all purpose flour, sugar, ginger, milk
400	Rosemary Roasted Vegetables	Healthy Food	veg	kosher salt, rosemary, garlic, potato, olive oil, carrot, walnut, cheese, ground pepper

2. ratings.csv

Dataset ini mencakup data interaksi pengguna dengan makanan, yang berisi ID pengguna, ID makanan, dan rating yang diberikan pengguna untuk makanan tertentu, yang terdiri dari:

- ID Pengguna (User_ID): Identifikasi unik untuk setiap pengguna.
- ID Makanan (Food_ID): Referensi kepada makanan yang dinilai oleh pengguna.
- Rating: Penilaian yang diberikan oleh pengguna dalam skala tertentu, misalnya dari 1 hingga 10.

Tabel 2. Dataset ratings.csv

User_ID	Food_ID	Rating
1	88	4
1	46	3

1	24	5
1	25	4
2	49	1
2	33	8
2	106	9
2	71	8
3	73	9
3	110	10
3	168	1
...		
98	217	3
99	49	3
99	69	8
99	93	7
99	152	2
99	65	7
99	22	1
100	24	10
100	233	10
100	29	7

Data tersebut akan digunakan untuk membangun sistem rekomendasi yang dapat menyarankan makanan kepada pengguna berdasarkan kesamaan preferensi mereka dengan pengguna lain yang memiliki pola kesukaan yang mirip.

3.2 Tahapan Penelitian

1. Pengumpulan Data

Proyek ini memanfaatkan dua dataset utama yang telah disebutkan sebelumnya. Dataset "food.csv" menyimpan informasi tentang berbagai jenis makanan, sedangkan "ratings.csv" mencatat riwayat interaksi pengguna dengan makanan melalui penilaian

yang diberikan. Data yang digunakan dalam penelitian ini diambil dari sumber yang dapat diakses secara publik dan relevan dengan topik rekomendasi makanan di kaggle ([Food Recommendation System](#)). Kombinasi kedua dataset ini memungkinkan analisis yang berfokus pada interaksi antara pengguna dan item, sehingga dapat menghasilkan rekomendasi yang lebih akurat.

Pengumpulan data ini bertujuan untuk menciptakan model yang mampu merekomendasikan makanan berdasarkan pola preferensi yang serupa di antara para pengguna. Penelitian sebelumnya oleh Yusmar et al. (2021) menunjukkan bahwa penggunaan dataset berbasis interaksi pengguna memiliki potensi besar dalam meningkatkan kualitas rekomendasi, karena model dapat memahami kesamaan preferensi antar pengguna dengan lebih baik. Dataset dimuat menggunakan pustaka Pandas untuk analisis lebih lanjut. Proses pengumpulan data mencakup :

- **Memuat Dataset:** Dataset food.csv dan ratings.csv dimuat untuk diproses.

```
# Step 1: Load datasets
food_df = pd.read_csv('food.csv')
ratings_df = pd.read_csv('ratings.csv')      # Dataset rating
(ratings.csv)
```

- **Pembersihan Data:** Data yang memiliki missing values pada kolom kritis, seperti rating atau deskripsi makanan, dihapus.

```
print("Cek Data Ratings")
print(ratings_df.isnull().sum())
ratings = ratings_df.dropna()

print("\nCek Data Makanan")
print(food_df.isnull().sum())
food = food_df.dropna()

print("\nData Ratings:", ratings.shape)
print("\nData Makanan:", food.shape)
```

- **Menggabungkan Dataset:** Kedua dataset digabungkan berdasarkan kolom Food_ID.

```
# Step 2: Merge the datasets

merged_df = pd.merge(ratings_df, food_df, on='Food_ID',
how='inner')
```

2. *Preprocessing Data*

Preprocessing merupakan tahap penting yang bertujuan untuk mempersiapkan data agar siap digunakan dalam model. Beberapa langkah yang dilakukan dalam tahap *preprocessing* ini adalah

a. *Handling Missing Values*

Data yang hilang pada kolom rating akan digantikan dengan nilai rata-rata dari rating yang ada untuk setiap makanan. Hal ini dilakukan untuk memastikan bahwa data yang digunakan dalam model tidak terpengaruh oleh nilai yang hilang. Pada beberapa kasus, data yang memiliki kekurangan signifikan atau tidak relevan juga dihapus agar tidak mengganggu kualitas model.

b. Normalisasi Rating

Untuk menghindari perbedaan skala pada rating yang diberikan oleh pengguna, dilakukan normalisasi dengan metode *Min-Max Normalization*. Dengan cara ini, rating yang diberikan oleh pengguna akan disesuaikan ke dalam rentang [0,1], sehingga dapat lebih mudah diproses dalam model rekomendasi.

c. Pengelompokan Data

Dataset makanan dikategorikan dalam tiga kategori utama, yaitu *Healthy Food*, *Snack*, dan *Dessert*, sesuai dengan jenis makanan yang ada. Hal ini memungkinkan sistem rekomendasi untuk menyarankan makanan yang sesuai dengan kategori yang diminati pengguna. Pengelompokan juga mempermudah dalam memberikan rekomendasi berdasarkan jenis makanan yang relevan dengan preferensi pengguna.

3. Implementasi Metode *Collaborative Filtering*

Setelah tahap *preprocessing*, sistem rekomendasi dibangun dengan menggunakan metode *Collaborative Filtering*. Metode ini mengandalkan data interaksi antar pengguna untuk menentukan kesamaan preferensi dan memberikan rekomendasi yang lebih relevan. Beberapa langkah dalam implementasi *Collaborative Filtering* adalah :

a. Pembangunan Matriks User-Item

Matriks User-Item dibangun dengan memetakan pengguna sebagai baris dan makanan sebagai kolom. Nilai dalam matriks ini berupa rating yang diberikan oleh pengguna terhadap makanan. Matriks ini akan menggambarkan tingkat preferensi setiap pengguna terhadap makanan yang ada dalam dataset.

```
# Step 3: Pivot data to create a User-Food matrix

user_food_matrix = merged_df.pivot_table(index='User_ID',
columns='Food_ID', values='Rating', fill_value=0)
```

b. Penghitungan *Cosine Similarity*

Cosine Similarity digunakan untuk mengukur kesamaan antara dua pengguna atau antara dua item (makanan). Rumus yang digunakan adalah sebagai berikut:

$$\text{Cosine Similarity : } \frac{\sum_{i=1}^n A_i \cdot B_i}{\sqrt{\sum_{i=1}^n A_i^2} \cdot \sqrt{\sum_{i=1}^n B_i^2}}$$

Di mana A_i dan B_i adalah vektor yang menggambarkan preferensi pengguna atau karakteristik makanan, dan n adalah jumlah dimensi data. Semakin tinggi nilai Cosine Similarity, semakin besar kesamaan preferensi antara dua entitas yang dibandingkan.

```
# Step 4: Calculate cosine similarity between users

user_similarity = cosine_similarity(user_food_matrix)

user_similarity_df = pd.DataFrame(user_similarity,
index=user_food_matrix.index,
columns=user_food_matrix.index)
```

c. Prediksi Rating

Berdasarkan hasil perhitungan *Cosine Similarity*, sistem akan memprediksi rating yang akan diberikan oleh pengguna terhadap makanan yang belum mereka beri rating. Prediksi ini dihitung dengan menggunakan rata-rata tertimbang dari

rating yang diberikan oleh pengguna lain yang memiliki kesamaan preferensi tinggi.

```
# Step 5: Recommendation function for existing users

def recommend(user, user_food_matrix, user_similarity_df,
top_n=5):

    """
    Recommends top N food items for a given user based on
    user similarity.

    """

    if user not in user_food_matrix.index:

        return "User not found in dataset."

    # Get the user's similarity scores

    similar_users = user_similarity_df[user]

    # Calculate weighted scores for food items

    scores = np.dot(similar_users, user_food_matrix.values)

    recommendation_scores = pd.Series(scores,
index=user_food_matrix.columns)

    # Exclude items the user has already rated

    target_ratings = user_food_matrix.loc[user]

    recommendation_scores[target_ratings > 0] = 0

    # Return the top N food items

    recommendations = recommendation_scores.sort_values(ascending=False).head(top_n)

    return recommendations
```

d. Rekomendasi untuk Pengguna Baru

Pengguna baru dapat memilih kategori dan jenis makanan untuk mendapatkan rekomendasi. Sistem memanfaatkan rata-rata preferensi pengguna yang paling mirip untuk memberikan saran.

```
# Function for new user recommendation

def recommend_for_new_user(new_user_preferences,
user_food_matrix, user_similarity_df, food_df, top_n=5):

    """
        Recommends top N food items for a new user based on
        preferences and similar users' ratings.

    """

    # Filter makanan berdasarkan kategori dan jenis
    # (preferences pengguna baru)

    filtered_food = food_df[(food_df['C_Type'].str.strip().str.capitalize()).isin(new_user_preferences['categories'])] &
                    (food_df['Veg_Non'].str.strip().str.lower() == new_user_preferences['veg_non'])

    ]

    if filtered_food.empty:

        return "Tidak ada makanan yang sesuai dengan
preferensi Anda."

    filtered_food_ids = filtered_food['Food_ID'].values

    # Validasi Food_ID untuk memastikan kesesuaian dengan
    user_food_matrix

    valid_food_ids = [fid for fid in filtered_food_ids if
fid in user_food_matrix.columns]

    if not valid_food_ids:

        return "Tidak ada makanan untuk direkomendasikan
berdasarkan preferensi Anda."
```

```

# Cari pengguna lama yang paling mirip dengan pengguna baru

average_user_preferences = user_food_matrix.mean(axis=1)

most_similar_user = average_user_preferences.idxmax()

# Ambil skor rating dari pengguna lama yang mirip

similar_user_ratings = user_food_matrix.loc[most_similar_user]

# Filter rating berdasarkan makanan yang tersedia untuk pengguna baru

similar_user_ratings = similar_user_ratings[valid_food_ids]

# Urutkan berdasarkan rating tertinggi

recommendations = similar_user_ratings.sort_values(ascending=False).head(top_n)

if recommendations.empty:

    return "Tidak ada makanan untuk direkomendasikan berdasarkan pengguna serupa."

# Tampilkan hasil rekomendasi

recommended_food = food_df[food_df['Food_ID'].isin(recommendations.index)]

return recommended_food[['Food_ID', 'Name']]

```

4. Evaluasi Sistem Rekomendasi

Tahapan ini bertujuan untuk mengevaluasi kualitas rekomendasi yang dihasilkan oleh sistem. Proses evaluasi dilakukan menggunakan metrik berikut :

a. Metrik Evaluasi

- **Precision:** Mengukur proporsi makanan yang direkomendasikan yang benar-benar relevan.
- **Recall:** Mengukur seberapa baik sistem merekomendasikan semua makanan relevan dari dataset.

- **F1-Score:** Kombinasi harmonis antara Precision dan Recall untuk keseimbangan evaluasi.
 - **Accuracy:** Persentase rekomendasi yang benar dari total data yang diuji.
- b. Langkah-Langkah Evaluasi
- **Identifikasi True Positives (TP):** Makanan yang direkomendasikan sesuai dengan preferensi pengguna.
 - **Identifikasi False Positives (FP):** Makanan yang direkomendasikan tetapi tidak sesuai preferensi.
 - **Identifikasi False Negatives (FN):** Makanan yang sesuai preferensi tetapi tidak direkomendasikan.
 - Berdasarkan penghitungan TP, FP, dan FN, metrik Precision, Recall, F1-Score, dan Accuracy dihitung.
- c. Implementasi Fungsi Evaluasi
- Fungsi evaluasi memanfaatkan pustaka seperti scikit-learn untuk otomatisasi penghitungan metrik, memastikan akurasi pengukuran.
- d. Pengujian dan Analisis Hasil

Evaluasi dilakukan dengan membandingkan output rekomendasi dengan ground truth untuk berbagai skenario pengguna, termasuk pengguna baru dan pengguna yang telah memiliki interaksi sebelumnya.

5. Pengembangan Sistem

Sistem ini dikembangkan dengan menggunakan bahasa pemrograman Python, didukung oleh berbagai pustaka yang memberikan fungsionalitas tambahan. Berikut adalah beberapa pustaka yang digunakan:

- Pandas: Digunakan untuk membaca dan mengelola dataset.
- NumPy: Memfasilitasi operasi numerik dan matematis pada matriks.
- Scikit-Learn: Berperan dalam perhitungan *Cosine Similarity* dan implementasi del rekomendasi.

```
from sklearn.metrics.pairwise import cosine_similarity
import numpy as np
import pandas as pd
```

Alur Proses Sistem:

- Input Data: Pengguna memberikan penilaian terhadap makanan yang mereka konsumsi.

- Proses *Preprocessing*: Data yang diterima akan menjalani tahap normalisasi dan pengelompokan untuk mempersiapkan model.
- Pembangunan Model: Model *Collaborative Filtering* diterapkan untuk mengenali pola kesamaan di antara pengguna.
- Rekomendasi: Berdasarkan model yang telah dibangun, sistem akan memberikan rekomendasi makanan yang relevan bagi pengguna.
- Evaluasi Sistem: Menggunakan metrik evaluasi untuk menilai relevansi dan akurasi rekomendasi.

3.3 Alur Proses Sistem

Berikut adalah diagram alur proses dalam sistem rekomendasi makanan yang dikembangkan:

1. Pengumpulan Data → 2. Preprocessing Data → 3. Pembangunan Matriks User-Item → 4. Penghitungan Cosine Similarity → 5. Prediksi Rating → 6. Hasil Rekomendasi → 7. Evaluasi Sistem

- Pengumpulan Data:** Memuat dataset food.csv dan ratings.csv.
- Preprocessing Data:** Membersihkan data, mengisi nilai yang hilang, dan menormalisasi rating.
- Pembangunan Matriks User-Item:** Membentuk matriks hubungan pengguna dan makanan.
- Penghitungan Cosine Similarity:** Mengukur kesamaan antar pengguna berdasarkan pola rating.
- Prediksi Rating:** Sistem memprediksi rating untuk makanan yang belum diberi rating.
- Hasil Rekomendasi:** Sistem memberikan rekomendasi makanan sesuai preferensi pengguna, dilengkapi dengan output akurasi yang mencerminkan kecocokan berdasarkan data yang dimasukkan.
- Evaluasi Sistem:** Menggunakan metrik evaluasi untuk menilai relevansi dan akurasi rekomendasi.

BAB 4

HASIL DAN PEMBAHASAN

4.1 Hasil

Hasil dari sistem rekomendasi makanan berbasis *Collaborative Filtering* dengan pendekatan *Cosine Similarity* menunjukkan kemampuan sistem dalam memberikan rekomendasi yang relevan berdasarkan preferensi pengguna dan telah berhasil diimplementasikan menggunakan framework web Flask. Sistem ini memanfaatkan dua dataset utama, yaitu food.csv dan ratings.csv, untuk memberikan rekomendasi makanan kepada pengguna berdasarkan preferensi mereka. Proses pengembangan sistem menghasilkan beberapa data utama dan rekomendasi makanan, yang dijelaskan sebagai berikut:

1. Pembersihan Data

Tahapan awal dalam sistem adalah memeriksa kelengkapan data pada dataset ratings.csv dan food.csv. Hasil pemeriksaan menunjukkan :

- Cek Data Ratings : Dataset memiliki beberapa nilai kosong pada kolom Rating, yang dihapus untuk menjaga kualitas data.

```
Cek Data Ratings
User_ID      1
Food_ID      1
Rating       1
dtype: int64
```

- Cek Data Makanan: Dataset makanan tidak memiliki nilai kosong.

```
Cek Data Makanan
Food_ID      0
Name         0
C_Type       0
Veg_Non     0
Describe    0
dtype: int64
```

Setelah proses pembersihan data, sistem memiliki:

```
Banyak Data Ratings Sebelum di Hapus (512, 3)
Banyak Data Ratings Setelah di Hapus (511, 3)
```

```
Banyak Data Makanan (400, 5)
```

- Data Ratings : Sebanyak 511 baris data interaksi pengguna dengan makanan dan 3 kolom setelah menghapus nilai yang hilang (*missing values*).
- Data Makanan : Sebanyak 400 baris data makanan dengan deskripsi lengkap dan 5 kolom tanpa nilai yang hilang.

```
Nama kolom dalam food_df:  
Index(['Food_ID', 'Name', 'C_Type', 'Veg_Non', 'Describe'], dtype='object')
```

2. Matriks Kesamaan Antar Pengguna (*User Similarity Matrix*)

Sistem menghasilkan matriks kesamaan pengguna menggunakan *Cosine Similarity*, yang menggambarkan tingkat kemiripan preferensi antar pengguna berdasarkan nilai rating. Matriks ini digunakan untuk menentukan pengguna yang memiliki pola kesukaan makanan yang mirip, sehingga rekomendasi dapat diberikan dengan lebih relevan.

User Similarity Matrix:																		
User_ID	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	...	93.0	94.0	95.0	96.0	97.0	98.0	99.0	100.0	
User_ID																		
1.0	1.000000	0.000000	0.027842	0.0	0.000000	0.000000	0.000000	0.000000	...	0.0	0.081052	0.0	0.0	0.130415	0.225952	0.000000	0.278982	
2.0	0.000000	1.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	...	0.0	0.000000	0.0	0.0	0.000000	0.000000	0.015605	0.000000	
3.0	0.027842	0.000000	1.000000	0.0	0.000000	0.000000	0.000000	0.072998	...	0.0	0.000000	0.0	0.0	0.052045	0.000000	0.083428	0.000000	
4.0	0.000000	0.000000	0.000000	1.0	0.000000	0.000000	0.03717	0.100890	...	0.0	0.000000	0.0	0.0	0.000000	0.000000	0.000000	0.000000	
5.0	0.000000	0.000000	0.000000	0.0	1.000000	0.000000	0.000000	0.000000	...	0.0	0.000000	0.0	0.0	0.03693	0.000000	0.000000	0.000000	
...	
96.0	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	...	0.0	0.000000	0.0	1.0	0.000000	0.000000	0.000000	0.000000	
97.0	0.130415	0.000000	0.052045	0.0	0.03693	0.000000	0.198594	0.000000	...	0.0	0.000000	0.0	0.0	1.000000	0.000000	0.000000	0.000000	
98.0	0.225952	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	...	0.0	0.094500	0.0	0.0	0.000000	1.000000	0.000000	0.000000	
99.0	0.000000	0.015605	0.083428	0.0	0.000000	0.035299	0.000000	0.000000	...	0.0	0.000000	0.0	0.0	0.000000	0.000000	1.000000	0.000000	
100.0	0.278982	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	...	0.0	0.000000	0.0	0.0	0.000000	0.000000	0.000000	1.000000	

Matriks ini menunjukkan bahwa kesamaan antar pengguna dihitung menggunakan *Cosine Similarity*, dengan nilai berkisar antara 0 hingga 1. Nilai 1 menunjukkan kesamaan sempurna, sementara nilai 0 menunjukkan tidak ada kesamaan.

3. Rekomendasi Makanan untuk Pengguna Baru

Pengguna diberikan rekomendasi berdasarkan kategori makanan dan preferensi vegetarian/non-vegetarian yang mereka masukkan. Ketika pengguna baru memasukkan preferensi mereka, sistem menyaring data makanan berdasarkan kategori dan jenis makanan yang diinginkan. Sebagai contoh:

a. Input Referensi :

- Kategori : *Dessert*
- Jenis Makanan : *vegetarian*

b. Output Rekomendasi :

Berdasarkan preferensi tersebut, sistem memberikan rekomendasi makanan sebagai berikut:

Tabel 3. Output Rekomendasi

Food_ID	Name
5	<i>Christmas Cake</i>
224	<i>Eggless Vanilla Cake</i>
230	<i>Eggless Coffee Cupcakes</i>
232	<i>Apple and Walnut Cake</i>
244	<i>Jalebi with Fennel Yogurt Pudding</i>

Hasil ini menunjukkan bahwa sistem dapat menyaring data dengan baik untuk memberikan rekomendasi yang sesuai dengan preferensi pengguna. Rekomendasi ini diberikan berdasarkan pola kesamaan preferensi dengan pengguna lain yang memiliki kesukaan serupa dalam dataset. Selain mendapatkan rekomendasi makanan, pengguna juga akan melihat tingkat akurasi berdasarkan input yang mereka masukkan.

```
Evaluasi hasil rekomendasi berdasarkan input Anda:  
Precision: 1.00  
Recall: 0.10  
F1 Score: 0.18  
Accuracy: 0.85
```

4.2 Pembahasan

Hasil implementasi sistem menunjukkan beberapa poin penting:

1. Kinerja *Preprocessing* Data

Tahapan preprocessing berhasil membersihkan data dari nilai yang hilang dan menormalisasi rating ke skala seragam. Hal ini memastikan bahwa data yang digunakan memiliki kualitas yang baik untuk dianalisis oleh sistem rekomendasi.

2. Kesamaan Antar Pengguna

Dengan menggunakan *Cosine Similarity*, sistem dapat menghitung kesamaan antar pengguna berdasarkan pola rating. Sebagai contoh, pengguna dengan ID 1.0 memiliki tingkat kesamaan tertinggi dengan pengguna ID 100.0 (nilai kesamaan: 0.278982). Kesamaan ini menjadi dasar dalam memberikan rekomendasi makanan.

3. Akurasi Rekomendasi

Proses pengolahan data, seperti pembersihan, normalisasi, dan pembuatan matriks kesamaan, dilakukan secara terstruktur sehingga sistem dapat mengelola data dengan efisien tanpa mengorbankan akurasi rekomendasi. Sistem mampu memberikan rekomendasi makanan berdasarkan kategori dan jenis makanan yang dipilih pengguna. Contoh output menunjukkan bahwa sistem berhasil merekomendasikan 5 makanan *vegetarian* dengan kategori *Dessert*, sesuai dengan input pengguna baru.

4. Personalisasi Rekomendasi

Sistem memanfaatkan metode *Collaborative Filtering* berbasis *Cosine Similarity* untuk memberikan rekomendasi makanan yang relevan berdasarkan preferensi pengguna. Hal ini memungkinkan sistem untuk mengenali pola interaksi pengguna dengan lebih baik.

5. Interaksi yang Fleksibel

Dengan menggunakan framework Flask, pengguna dapat memasukkan preferensi mereka melalui antarmuka web, baik sebagai pengguna lama (berdasarkan data interaksi sebelumnya) maupun pengguna baru (berdasarkan kategori dan jenis makanan yang diinginkan).

6. Keterbatasan Sistem

- Sistem bergantung pada data yang ada di dataset. Jika preferensi pengguna baru tidak cocok dengan data dalam dataset, rekomendasi yang dihasilkan akan terbatas.
- Sistem hanya menggunakan pendekatan *Collaborative Filtering* berbasis kesamaan pengguna, tanpa mempertimbangkan atribut makanan secara langsung (*Content-Based Filtering*).

7. Rekomendasi Pengembangan

- Menambahkan metode *Hybrid Filtering* yang menggabungkan *Collaborative Filtering* dan *Content-Based Filtering* untuk meningkatkan relevansi rekomendasi.
- Menggunakan dataset yang lebih besar untuk mencakup lebih banyak variasi makanan dan interaksi pengguna.
- Memberikan *feedback* ke pengguna bahwa makanan yang sudah direkomendasikan adalah makanan yang paling sesuai dengan preferensi mereka berdasarkan data yang dimasukkan dan hasil analisis sistem.

BAB 5

KESIMPULAN

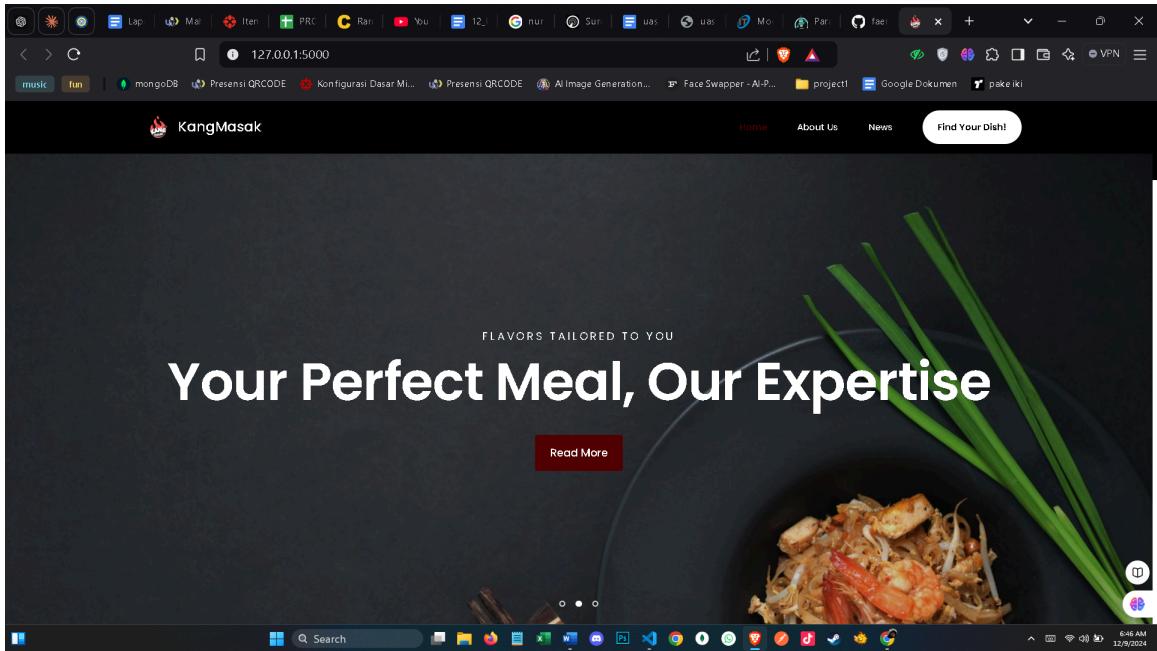
Sistem rekomendasi makanan yang diimplementasikan pada Flask berbasis *Collaborative Filtering* dengan pendekatan *cosine similarity* mampu memberikan rekomendasi makanan yang relevan berdasarkan preferensi pengguna. Proses pembersihan data dilakukan dengan benar untuk memastikan kualitas data dan *cosine similarity* digunakan untuk menghitung kesamaan pengguna sehingga sistem dapat mendeteksi pola interaksi dan memberikan rekomendasi yang sesuai. Contoh implementasi ini menunjukkan bahwa sistem dapat merekomendasikan makanan seperti *dessert*, *healthy food*, *snack* berdasarkan kategori dan jenis makanan yang dipilih oleh pengguna.

Namun, sistem ini masih memiliki keterbatasan dalam cakupan kumpulan data dan pendekatan yang hanya didasarkan pada pemfilteran kolaboratif. Untuk meningkatkan kinerja, sistem dapat dikembangkan dengan menambahkan metode *hybrid filtering* yang menggabungkan *Collaborative Filtering* dan *Content-based Filtering*. Selain itu, penggunaan kumpulan data yang lebih besar memperluas jangkauan makanan dan preferensi pengguna, sehingga menghasilkan rekomendasi yang lebih personal dan akurat.

LAMPIRAN

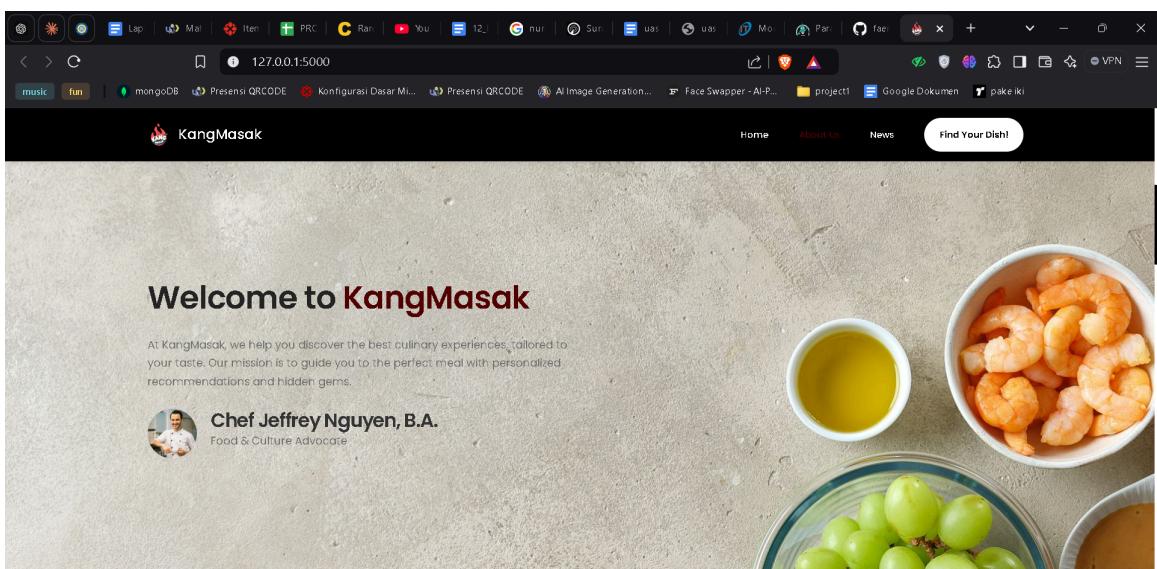
Berikut adalah link github kami <https://github.com/faerosjahera/MLPrakUAS>, selain github pada bab ini akan berisi *screenshot* dari web yang sudah kami kerjakan. Berikut adalah *screenshot* nya :

1. Landing Page



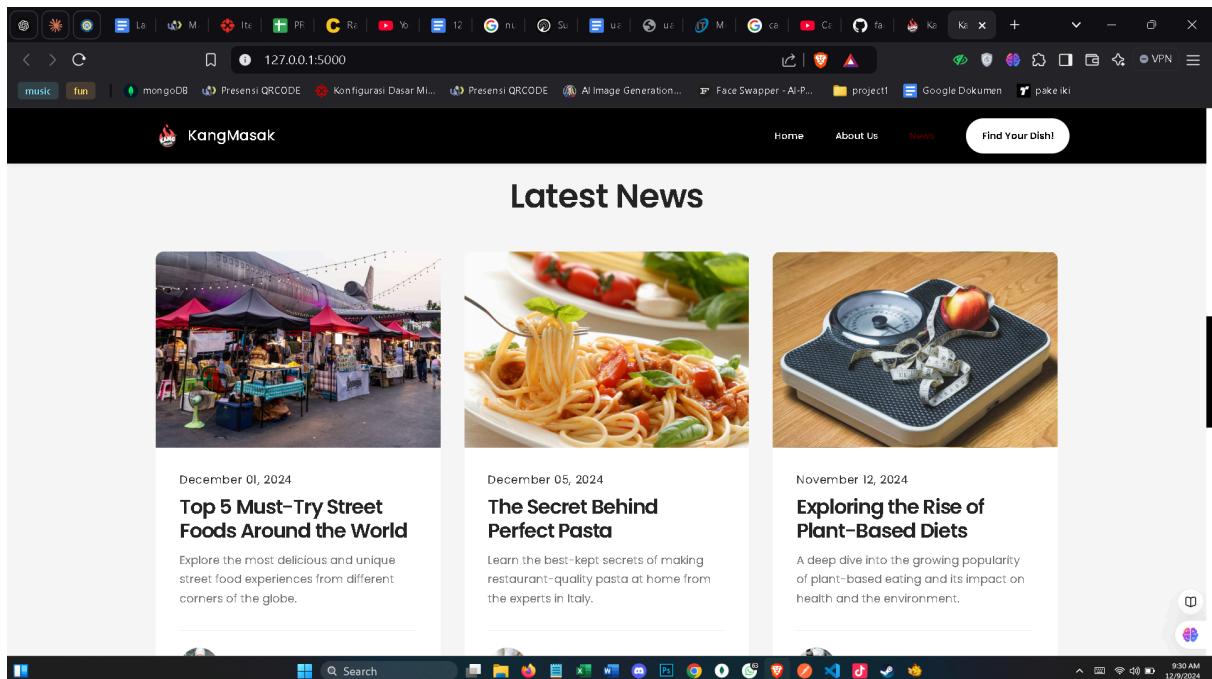
Gambar 1. Landing Page

2. About Us Page



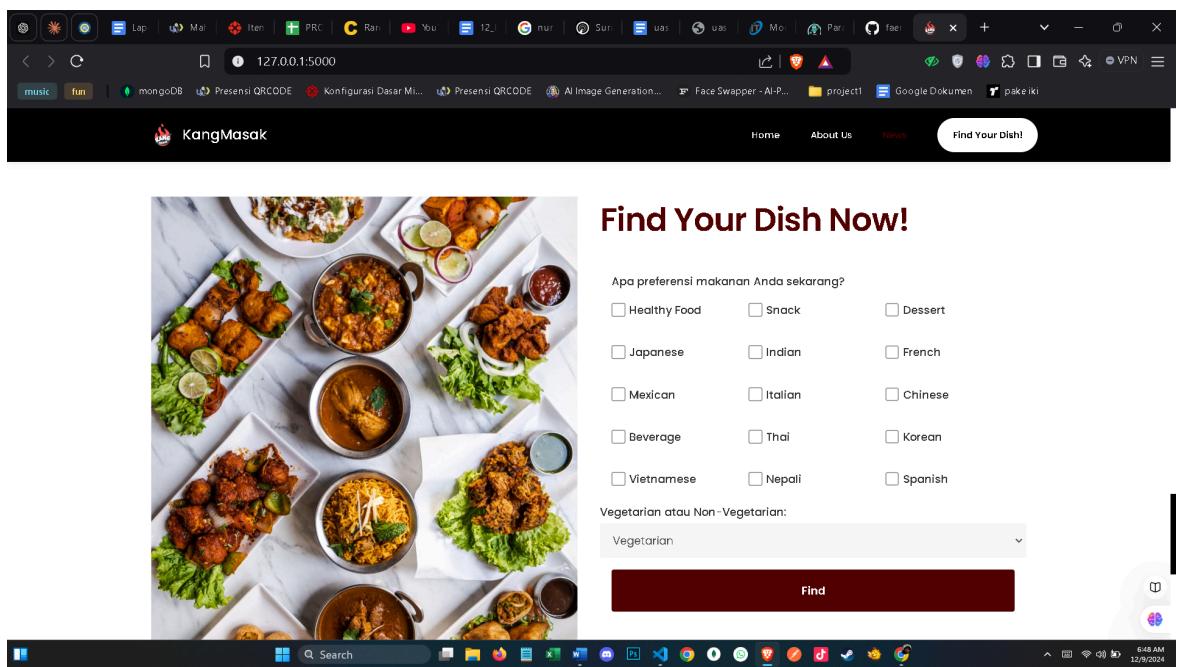
Gambar 2. About Us Page

3. News Page



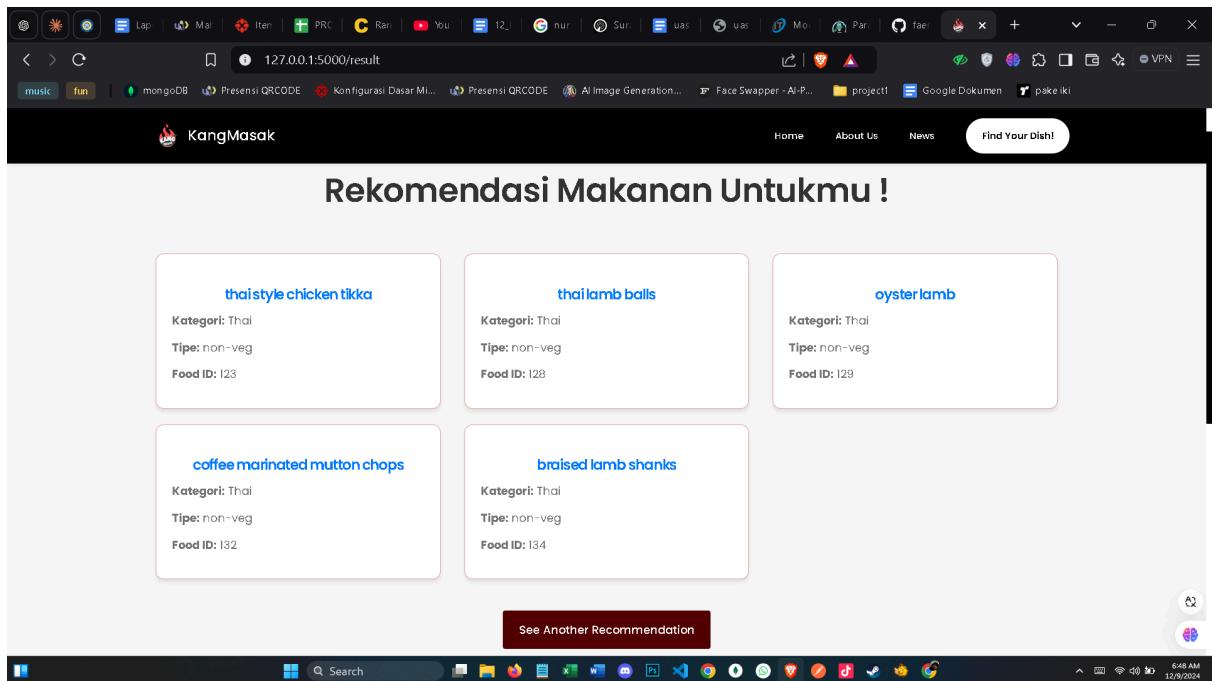
Gambar 3. News Page

4. User Input Recommend Page



Gambar 4. User Input Page

5. Output Page (Recommend Page)



Gambar 5. Output Page

DAFTAR PUSTAKA

Erwin, V. C. M., & Hendryli, J. (2020). *Penggunaan Metode Collaborative Filtering Based untuk Rekomendasi Kendaraan Bermotor.*

Fajar Christyawan, Arif Nur Rohman, & Anggit Dwi Hartanto. (2024, September). *Application of Content-Based Filtering Method Using Cosine Similarity in Restaurant Selection Recommendation System. Journal of Information Systems and Informatics,* 6(3). 10.51519/journalisi.v6i3.806

Februariyanti, H., Laksono, A. D., Wibowo, J. S., & Utomo, M. S. (2021, Juni). *Implementasi Metode Collaborative Filtering untuk Sistem Rekomendasi Penjualan pada Toko Mebel, IX(1), 43-50.*

Fiarni, C., Maharani, H., & Calista, N. (2019). *Product Recommendation System Design Using Cosine Similarity and Content-based Filtering Methods* (Vol. 3). IJITEE.

Yusmar, A., Wardhani, L. K., & Suseno, H. B. (2021, April). *Restaurant Recommender System Using Item-Based Collaborative Filtering and Adjusted Cosine Algorithm Similarity., 14(1).*