



Model Rekomendasi Makanan Menggunakan *Content-Based* dan *Collaborative Filtering*

Falahah^{1*}, Rita Komalasari²

¹Program Studi Sistem Informasi, Fakultas Rekayasa Industri, Universitas Telkom

²Program Studi Manajemen Informatika, Politeknik LP3I

¹falahah@telkomuniversity.ac.id, ²ritakomalasari@plb.ac.id

Abstract

The Food Recommendation System is one of the most popular implementations of the recommendation system in everyday life. Food recommendations work similarly to other recommendation systems, but have unique challenges such as the need for information about food ingredients (ingredients), when recommendations are given (time), data availability on the system (cold-starved users and cold-starved food), and the user community. The application of general recommendation techniques such as content-based filtering and collaborative filtering has its own challenges for food recommendation systems. This study aims to find out how to apply content-based filtering and collaborative filtering techniques to provide food recommendations and what obstacles are encountered in the application of these two techniques. The results show that each method requires different dataset characteristics, and the completeness of the dataset is the main obstacle in implementation. Another obstacle is that the design of the food recommendation system is largely determined by the type of recommendation given and the language used to write the content.

Keywords: challenges, dataset, food recommendation, ingredients, language

Abstrak

Sistem Rekomendasi Makanan merupakan salah satu implementasi populer sistem rekomendasi pada kehidupan sehari-hari. Pada dasarnya cara kerja rekomendasi makanan serupa dengan sistem rekomendasi lainnya, namun memiliki tantangan tersendiri seperti kebutuhan informasi terkait bahan pembuat makanan (ingredients), waktu rekomendasi diberikan (time), ketersediaan data pada sistem (cold start user dan cold start food) dan komunitas pengguna. Penerapan teknik rekomendasi yang umum seperti content-based filtering maupun collaborative filtering memiliki kendala tersendiri untuk sistem rekomendasi makanan. Penelitian ini bertujuan untuk mengetahui bagaimana menerapkan teknik content-based filtering dan collaborative filtering pada pemberian rekomendasi makanan dan kendala apa yang dihadapi dalam penerapan kedua teknik tersebut. Hasil penelitian menunjukkan bahwa masing-masing metode memerlukan karakteristik dataset yang berbeda, serta kelengkapan dataset merupakan kendala utama dalam implementasi. Kendala lainnya adalah perancangan sistem rekomendasi makanan sangat ditentukan oleh jenis rekomendasi yang diberikan serta bahasa yang digunakan untuk menuliskan konten.

Kata kunci: Rekomendasi Makanan, kendala, dataset, bahan makanan, bahasa.

1. Pendahuluan

Sistem rekomendasi merupakan salah satu aplikasi populer dari teknik machine-learning pada kehidupan sehari-hari. Saat ini, mesin rekomendasi sudah banyak disertakan pada berbagai aplikasi hiburan, seperti Netflix, spotify, maupun aplikasi marketplace seperti tokopedia, booking.com, agoda, shopee, dan banyak lagi. Lebih spesifik lagi, rekomendasi makanan merupakan salah satu implementasi sistem rekomendasi yang sering kita temui pada aplikasi populer seperti

aplikasi pemesanan makanan (misalnya pada *Traveloka* ataupun *go-food*).

Sistem rekomendasi (*recommender system*) atau RS adalah sistem yang dapat membantu pengguna untuk memilih dari opsi yang tersedia, dengan cara menggunakan rekomendasi dari pengguna lainnya [1].

Sebuah sistem rekomendasi umumnya memuat tiga elemen yaitu: item, user, dan transaksi, yang berupa interaksi potensial antara pengguna dan sistem. Pada umumnya fase rekomendasi dibagi menjadi tiga tahap

yaitu: fase koleksi informasi, fase belajar, dan fase pemberian rekomendasi atau prediksi [2]. RS juga perlu dilengkapi dengan mekanisme *feedback* yang dapat dibagi menjadi tiga bagian yaitu *feedback* implisit, eksplisit dan *hybrid*. *Feedback* eksplisit dilakukan dengan cara pemberian *rating*) [3].

Secara umum, teknik rekomendasi dapat dibagi menjadi empat yaitu teknik *content-based*, *collaborative based*, *knowledge-based*, dan *hybrid*. Pada *content-based filtering*, rekomendasi diberikan berdasarkan data pada item yang direkomendasikan dan perilaku pengguna di masa lampau. [4]. Kendala pada pendekatan ini adalah kelengkapan koleksi data terhadap item tersebut. Teknik rekomendasi pada *content-based filtering* yang umum dilakukan adalah heuristik, yang biasanya menggunakan pendekatan tradisional seperti pengukuran kemiripan cosine, dan penerapan *machine learning* dan metode statistik dengan cara menganalisis dan membuat model dari data pengguna di masa lampau dan membuat prediksi preferensi pengguna.

Sedangkan pada *Collaborative filtering*, rekomendasi diberikan berdasarkan penilaian (*rating*) yang diberikan pengguna. Prinsip *collaborative filtering* adalah mencari kemiripan preferensi antara berbagai kelompok pengguna dan memberikan rekomendasi pada pengguna berdasarkan preferensi kelompok tersebut [5][6]. Teknik rekomendasi *collaborative filtering* dapat dibagi lagi menjadi dua kategori yaitu *memory-based* dan *model-based*. *Memory-based* didasari oleh perilaku pengguna, dan dibagi lagi menjadi *item-based* dan *user-based*. Sedangkan *model-based* merupakan teknik rekomendasi yang menerapkan *machine learning*, misalnya *deep learning* atau matriks faktorisasi.

Sistem Rekomendasi Makanan (*Food Recommendation System*) atau FRS, merupakan salah satu jenis sistem rekomendasi yang umumnya diterapkan untuk memberikan preferensi pilihan makanan kepada pengguna. Jika dibandingkan dengan sistem rekomendasi pada umumnya, FRS memiliki beberapa keunikan sebagai berikut [7]:

FRS melibatkan berbagai konteks dan domain pengetahuan, misalnya konteks kondisi pengguna (kondisi fisik) maupun kondisi lingkungan pengguna.

FRS sangat erat kaitannya dengan kesehatan pengguna, sehingga diharapkan FRS yang ideal dapat menyeimbangkan antara preferensi makanan individu dan kebutuhan Kesehatan nutrisi individu.

FRS memiliki karakteristik khusus dan factor keunikan tersendiri, seperti metode memasak, kandungan bahan makanan (*ingredients*) dan waktu penyajian.

Berdasarkan jenis rekomendasi yang dihasilkan, FRS dapat dibagi menjadi empat jenis yaitu rekomendasi resep makanan, program diet, pembelian makanan, dan penyusunan menu makanan [8]. Berdasarkan tujuan rekomendasi, FRS dapat dibagi menjadi tiga kategori yaitu: Preferensi pengguna (*user preference* FRS);

kebutuhan nutrisi dan kesehatan pengguna (*health and nutrition* FRS); Gabungan antara preferensi dan kebutuhan kesehatan pengguna.

Meskipun FRS sudah banyak dikembangkan dengan berbagai perbaikan pada tingkat akurasi, secara umum FRS masih mengalami kendala-kendala sebagai berikut [9]:

Bahan pembuat makanan (*ingredients*). Pada umumnya FRS mengandalkan pada data penilaian atau *rating* yang diberikan oleh pengguna untuk memberikan rekomendasi berdasarkan pendekatan *collaborative filtering*, tanpa memperhatikan bahan pembuat makanan. Padahal, biasanya individu memilih makanan berdasarkan bahan pembuat makanan. Misalnya ada individu yang memilih makanan yang tidak mengandung ikan laut, karena alergi.

Faktor waktu, FRS pada umumnya dibangun dengan asumsi bahwa preferensi pengguna di masa sekarang akan sama dengan selera di masa lampau. Sedangkan, selera pengguna kemungkinan dapat berubah dari waktu ke waktu karena berbagai sebab, seperti gaya hidup, diet, masalah Kesehatan dan lain-lain.

Cold start user dan *cold start foods*: berdasarkan fakta bahwa pengguna biasanya jarang memberikan *rating* terhadap makanan yang dipilih, yang akan menyulitkan FRS berbasis *collaborative-filtering* untuk mengetahui kemiripan makanan yang dipilih. Hal ini karena pada *collaborative-filtering*, rekomendasi hanya dapat diberikan kepada pengguna yang memberikan cukup banyak *rating* pada makanan, sehingga pengguna yang hanya memberikan *rating* sedikit (*cold-start user*) diabaikan. Demikian halnya dengan makanan baru (*cold start foods*) yang belum mendapatkan banyak *rating* juga diabaikan untuk direkomendasikan.

Komunitas pengguna, aspek lingkungan atau komunitas user juga dapat digunakan untuk memberikan rekomendasi, dengan menggunakan model klustering. Tetapi, model klustering ini juga dapat mengalami kendala seperti berapa klaster yang optimal dan efisiensi dari pengukuran *similarity* yang diterapkan

Penelitian ini bertujuan untuk mengkaji bagaimana proses pemberian rekomendasi makanan menggunakan dua teknik rekomendasi yang paling umum, yaitu *content-based filtering* dan *collaborative filtering*. Pertanyaan-pertanyaan yang akan dijawab adalah: Bagaimana menerapkan teknik *content-based filtering* dan *collaborative filtering* pada pemberian rekomendasi makanan; Kendala apa yang dihadapi dalam penerapan metode tersebut.

Pertanyaan-pertanyaan di atas akan dijawab dengan pembahasan tentang sistem rekomendasi, sistem rekomendasi makanan, teknik *content-based filtering*, *collaborative filtering* dan langkah-langkah penerapannya pada studi kasus rekomendasi makanan.

2. Metode Penelitian

Inti dari sistem rekomendasi adalah teknik yang digunakan untuk menghasilkan rekomendasi. Dua metode atau teknik yang paling umum digunakan adalah content-based filtering dan collaborative filtering. Pada penelitian ini, kedua pendekatan tersebut akan diterapkan pada proses pemberian rekomendasi makanan yang didasarkan dari menu dan bahan-bahan pembuatan makanan (*ingredients*). Tujuannya adalah dapat memberikan rekomendasi berdasarkan kemiripan selera pengguna yang dideteksi melalui kemiripan makanan berdasarkan bahan pembuat makanan (*content-based filtering*) atau berdasarkan pilihan pengguna (*collaborative filtering*).

2.1. Content-based Filtering

Content-based filtering yang akan digunakan pada penelitian ini yaitu dengan cara menghitung kemiripan (*similarity*) pada resep makanan dengan menggunakan metode cosine. Asumsi yang diperlukan adalah: Tersedia dataset menu makanan yang memuat nama menu, bahan dan tahapan pengolahan; Terpasang pustaka untuk *stopword* berbahasa Indonesia. Pada penelitian ini, digunakan pustaka Sastrawi.

Langkah-langkah pemberian rekomendasi menggunakan teknik *content-based filtering* sebagai berikut:

a. Preparation:

Siapkan dataset resep makanan yang terdiri atas nama menu, bahan-bahan (*ingredients*) dan tahapan pengolahan

Instal pustaka (*library*) yang diperlukan untuk pengolahan text (*text processing*), yaitu: Preprocessing: pustaka untuk mengenali *stopwords* dalam bahasa Indonesia; Processing: pustaka untuk mengukur relevansi antar kata dan untuk menghitung *similarity* antar kata

b. Preprocessing:

Baca dataset makanan; Hapus menu/judul menu yang duplikat; Cek cakupan *stopword*, jika terdapat *stopword* yang belum ada pada pustaka eksisting, tambahkan *stopword* secara manual

c. Processing:

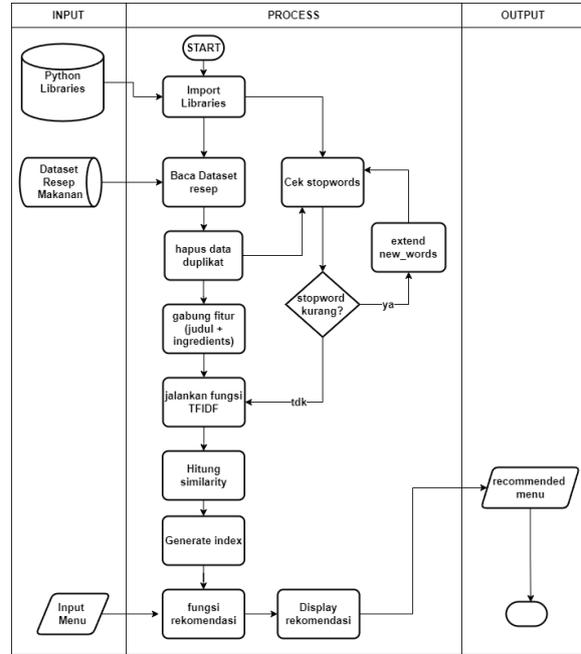
Extract features (judul menu dan bahan-bahan); Cek relevansi antar kata (terhadap gabungan fitur judul menu dan bahan makanan); Hitung *similarity* menu; Buat indeks untuk mengurutkan *similarity*

d. Generate Recommendation:

Buat fungsi untuk rekomendasi, berdasarkan hasil *similarity*; Gunakan fungsi rekomendasi untuk menghasilkan rekomendasi dari masukan menu tertentu.

Tahapan di atas dapat dilihat secara skematik seperti pada Gambar 1.

Rekomendasi *content-based filtering* cocok digunakan ketika pengguna baru pertama kali menggunakan sistem dan belum banyak pengguna yang memberikan rating pada menu makanan yang ada (*cold-start*). Jika pengguna sudah banyak berinteraksi, maka rekomendasi dapat diberikan berdasarkan kemiripan rating pengguna, dengan pendekatan *collaborative-filtering*.



Gambar 1. Langkah Rekomendasi dengan teknik *Content-Based Filtering*

Collaborative filtering yang akan digunakan pada penelitian ini yaitu yang berbasis *item-based*, yaitu mengukur kemiripan item yang mendapatkan rating yang sama. Asumsi yang digunakan adalah tersedianya dataset *rating* dari banyak pengguna terhadap sejumlah menu.

Langkah-langkah pemberian rekomendasi menggunakan teknik *collaborative filtering* sebagai berikut:

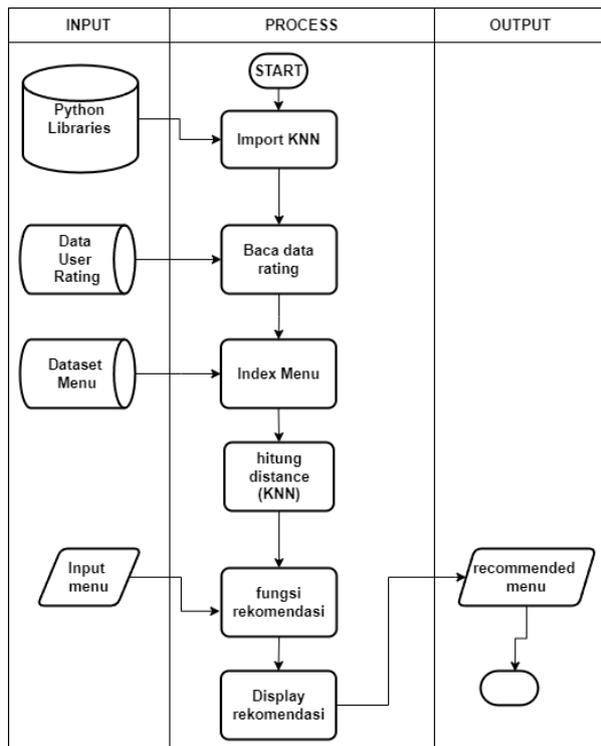
Preparation: Siapkan dataset rating makanan yang terdiri atas pengguna, nomor menu dan *rating* yang diberikan, serta dataset menu yang memuat nomor menu dan nama makanan.

Preprocessing: Baca dataset rating; Baca dataset menu

Processing: Buat index menu; Hitung *similarity* menu dan *rating*

Generate Recommendation: Buat fungsi untuk rekomendasi, berdasarkan hasil *similarity*; Gunakan fungsi rekomendasi untuk menghasilkan rekomendasi dari masukan menu tertentu.

Tahapan di atas dapat dilihat secara skematik seperti pada Gambar 2.



Gambar 2. Langkah Rekomendasi dengan teknik Collaborative Filtering

3. Hasil dan Pembahasan

Pada penelitian ini, langkah-langkah di atas diterapkan menggunakan bahasa pemrograman Python yang dijalankan pada *google-colabs*.

3.1. Content-Based Filtering

Implementasi metode *content-based filtering* dilakukan sebagai berikut

- Dataset: Memilih dataset resep makanan Indonesia yang memuat nama menu dan bahan-bahan pembuat makanan. Dataset diambil dari Kaggle [10]. Sebagai percobaan, akan diambil gabungan dari dua dataset yaitu dataset ikan dan udang. Total dataset berjumlah sebanyak 3926 baris dan 5 kolom
- Pustaka: Install pustaka Sastrawi [11]; Uji pustaka Sastrawi, tambahkan beberapa stopword; Import pustaka *NLTK* dan *regular expression* (re)

- Proses: Baca data menu; Hapus data duplikat `menu1.drop_duplicates(subset=['Title'])`

Ekstrak fitur judul menu dan *ingredients*

```
def create_soup(x):
    return ''.join(x['Title']) + ' + '.join(x['Ingredients'])
menu1['soup'] = menu1.apply(create_soup, axis=1)
```

Buat matriks skor *TF-IDF* untuk setiap kata

```
tfidf = TfidfVectorizer(stop_words=stop_words)
```

```
tfidf_matrix = tfidf.fit_transform(menu1['soup'])
tfidf_matrix.shape
```

Hasilnya adalah matrix *TF-IDF* berukuran (3926, 5279)

Hitung *similarity* antar kata dengan metode cosine

```
cosine_sim = linear_kernel(tfidf_matrix, tfidf_matrix)
```

Buat indeks menu

```
indices = pd.Series(menu1.index, index = menu1['Title']).drop_duplicates()
```

Siapkan fungsi rekomendasi dengan menggunakan metode *cosine*.

```
def get_recommend(title, cosine_sim, indices):
    idx = indices[title]
    sim_scores = list(enumerate(cosine_sim[idx]))
    sim_scores = sorted(sim_scores, key=lambda x: x[1], reverse=True)
    sim_scores = sim_scores[1:14]
    menu_indices = [i[0] for i in sim_scores]
    return menu1['Title'].iloc[menu_indices]
```

Gunakan fungsi rekomendasi untuk mendapatkan rekomendasi

```
get_recommend('Tumis Udang Tempe', cosine_sim, indices)
```

Beberapa contoh hasil rekomendasi:

```
2133 Tumis Udang pete
3146 Tumis udang saus tiram
3874 Udang tempe balado
2639 Udang Tumis Cabe Merah
2481 Udang Tempe Pedas
```

3.2. Collaborative Filtering

Implementasi metode *colabortaive filtering* dilakukan sebagai berikut

- Dataset: Memilih dataset resep makanan Indonesia yang memuat nama menu dan bahan-bahan pembuat makanan. Dataset diambil dari Kaggle.; Membuat data simulasi untuk rating terhadap setiap menu. Pada penelitian ini, rating diberikan hanya untuk 100 menu, dan oleh 100 user. Rating yang diberikan antara 1 sampai 5 dan diberikan secara random.

- Proses: Baca data menu dan *rating* pengguna

Buat indeks untuk menu

```
def get_menu(idx):
    return(df_menu.iloc[idx-1])
```

Hitung *similarity* data *rating* menggunakan pendekatan KNN dengan metrik *cosine* dan algoritma yang digunakan adalah *brute*.

```
knn = NearestNeighbors(metric='cosine', algorithm='brute')
knn.fit(df_rating.values)
```

Hitung *distance* setiap menu, dengan jumlah $n=3$

```
distances, indices = knn.kneighbors(df_
rating.values, n_neighbors=3)
```

Buat fungsi rekomendasi

```
index_user_likes = df_rating.index.to
list().index(menu)
sim_menu = indices[index_user_likes].to
list()
menu_distances = distances[index_user_]
likes].tolist()
id_menu = sim_menu.index(index_user_lik
es)
print('Similar menu to '+str(df_rating.
index[index_user_likes])+': \n')
sim_menu.remove(index_user_likes)
menu_distances.pop(id_menu)
```

```
j = 1
```

```
for i in sim_menu:
print(str(j)+'': '+str(df_rating.index[i
])+', the distance with '+str(menu)+'':
'+str(menu_distances[j-1]))
j = j + 1
```

Gunakan fungsi rekomendasi untuk mendapatkan rekomendasi

```
recommend_menu(17)
Udang sawi saori saus tiram
```

Beberapa contoh hasil rekomendasi:

```
1: 24, the distance with 17:
0.15687365906430617
Tumis Udang tempe kacang panjang
2: 20, the distance with 17:
0.1653572923146065
Udang Saus Asam Manis
```

3.3. Diskusi

Dari beberapa simulasi atas data-data di atas, untuk kasus pemberian rekomendasi makanan menggunakan pendekatan *content-based* dan *collaborative filtering*, ditemui beberapa batasan dan kendala sebagai berikut:

Dataset, yang dapat dibagi menjadi dua kendala, untuk masing-masing pendekatan, yaitu:

Content-based filtering:

Konten, karena dasar pemberian rekomendasi *content-based* adalah konten, maka data makanan harus diuraikan berdasarkan konten. Konten pada makanan ini dapat diuraikan lagi berdasarkan kandungan gizi, cara memasak atau bahan makanan. Pemilihan konten sangat ditentukan oleh tujuan rekomendasi tersebut diberikan. Pada penelitian ini, rekomendasi dibuat dengan tujuan untuk memberikan saran makanan yang mirip dengan selera pengguna dengan masukan berupa nama menu makanan. Jika pemberian rekomendasi bertujuan untuk mengetahui gizi makanan, maka dataset yang tersedia harus memuat informasi kandungan gizi setiap menu.

Bahasa, karena dasar pemberian rekomendasi adalah mencari kemiripan kata-kata pada konten, maka pemilihan pustaka untuk *stopword* yang tepat akan menentukan bagaimana data text difilter dengan membuang *stopword* yang tidak diperlukan. Saat ini, pustaka *stopword* berbahasa Indonesia masih terus dikembangkan.

Collaborative filtering: Kelengkapan *rating*, karena pada *collaborative filtering* dasar pemberian rekomendasi adalah kemiripan *rating* yang diberikan oleh pengguna maka dalam kasus tidak semua menu mendapatkan *rating*, akan menyulitkan sistem menyertakan menu tersebut pada model rekomendasi. Kelengkapan profil pengguna, pada penelitian ini belum disertakan profil pengguna, karena tidak adanya data. Jika profil pengguna tersedia, maka rekomendasi dapat ditingkatkan akurasi dengan memperhatikan kemiripan profil pengguna.

Metode: Pada penelitian ini, hanya diterapkan satu metode mengukur kemiripan menggunakan *cosine*, baik pada *content-based* maupun *collaborative* menggunakan KNN, sehingga belum diuji apakah hasil pengukuran similarity dengan menggunakan metode lain seperti *euclidean distance*, *manhattan distance* dan lain-lain, akan menghasilkan skor dan urutan *similarity* yang serupa.

Selain itu, metode *collaborative filtering* dan *content-based filtering* memiliki karakteristik data yang berbeda sehingga tidak dapat diterapkan pada dataset yang sama. Oleh karena itu, *collaborative filtering* tidak dapat digabungkan dengan *content-based filtering*.

Pengujian Kinerja Model: Pada penelitian ini belum dapat dilakukan pengujian pada kinerja model, dikarenakan belum adanya dataset yang memadai untuk pengujian, khususnya untuk kasus rekomendasi makanan dengan rekomendasi yang diberikan berdasarkan kemiripan bahan pembuat makanan.

4. Kesimpulan

Berdasarkan hasil pembahasan di atas, dapat disimpulkan bahwa: Pada pemberian rekomendasi makanan, perlu ditetapkan tujuan utama pemberian rekomendasi, karena akan mempengaruhi kebutuhan konten yang harus tersedia pada dataset, khususnya untuk pendekatan berbasis *content-based*. Kualitas rekomendasi yang dihasilkan sangat bergantung pada kualitas dataset awal yang tersedia. Tanpa dataset yang lengkap maka hasil rekomendasi yang diberikan sangat rendah akurasi. Pemberian rekomendasi *content-based filtering* tidak dapat digabungkan dengan pendekatan *collaborative filtering* karena keduanya memerlukan karakteristik dataset yang berbeda. Masing-masing pendekatan rekomendasi memiliki kendala tersendiri, yang pada umumnya terletak pada kelengkapan dataset dan bahasa yang digunakan pada dataset.

Daftar Rujukan

- [1] R. Burke, A. Felfernig, and M. H. Göker, "Recommender Systems: An Overview," *AI Mag.*, vol. 32, no. 3, pp. 13–18, Jun. 2011, doi: 10.1609/aimag.v32i3.2361.
- [2] F. O. Isinkaye, Y. O. Folajimi, and B. A. Ojokoh, "Recommendation systems: Principles, methods and evaluation," *Egypt. Informatics J.*, vol. 16, no. 3, pp. 261–273, Nov. 2015, doi: 10.1016/j.eij.2015.06.005.

- [3] W.-T. Kuo, Y.-C. Wang, R. T.-H. Tsai, and J. Y. Hsu, "Contextual restaurant recommendation utilizing implicit feedback," in *2015 24th Wireless and Optical Communication Conference (WOCC)*, Oct. 2015, pp. 170–174, doi: 10.1109/WOCC.2015.7346199. [8]
- [4] S. Wairegi, W. Mwangi, and R. Rimiru, "A Framework for Items Recommendation System Using Hybrid Approach," 2020. [9]
- [5] T. K. Nayak and U. Sharma, "Management Disaster Systems," pp. 542–545. [10]
- [6] P. Sharma and L. Yadav, "MOVIE RECOMMENDATION SYSTEM USING ITEM BASED COLLABORATIVE FILTERING," *International Journal of Innovative Research in Computer Science & Technology*, vol. 8, no. 4. 2020, doi: 10.21276/ijirest.2020.8.4.2. [11]
- [7] W. Min, S. Jiang, and R. Jain, "Food Recommendation: Framework, Existing Solutions, and Challenges," *IEEE Trans. Multimed.*, vol. 22, no. 10, pp. 2659–2671, Oct. 2020, doi: 10.1109/TMM.2019.2958761.
- C. Trattner and D. Elswiler, "Food Recommender Systems: Important Contributions, Challenges and Future Research Directions," Nov. 2017, [Online]. Available: <http://arxiv.org/abs/1711.02760>.
- A. Rostami, "Personal Food Model," *MM 2020 - Proceedings of the 28th ACM International Conference on Multimedia*. pp. 4416–4424, 2020, doi: 10.1145/3394171.3414691.
- D. D. Hermansyah, "Indonesian Food Recipe : Data Visualization," *Kaggle*, 2020. <https://www.kaggle.com/code/dionisiusdh/indonesian-food-recipe-data-visualization/data> (accessed Aug. 10, 2022).
- "Sastrawi 1.0.1," *pypi.org*, 2016. <https://pypi.org/project/Sastrawi/> (accessed Aug. 10, 2022).