

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi telah mengubah cara masyarakat dalam memperoleh referensi resep makanan. Resep tidak lagi terbatas pada media cetak, melainkan tersedia secara luas melalui berbagai platform dan aplikasi berbasis *online*. Hal ini memberikan kemudahan akses terhadap resep makanan [1], [2]. Sebagai contoh, situs web Allrecipes.com merupakan salah satu *platform* berbagi resep terbesar di dunia dengan 40 juta pengguna dan lebih dari 3 miliar akses resep per tahun di 24 negara [3]. Namun sebuah penelitian menemukan bahwa kelimpahan informasi tersebut memunculkan permasalahan yang dikenal sebagai *choice overload*, yaitu kondisi ketika individu dihadapkan pada terlalu banyak alternatif pilihan sehingga mengalami kesulitan dalam menentukan keputusan yang optimal. Dampak dari kondisi ini menyebabkan pendorongan terhadap masyarakat untuk beralih pada konsumsi makanan cepat saji yang lebih praktis dan sering kali rendah nutrisi dan tinggi kalori [3]. Tidak hanya itu, kondisi tersebut dapat menyebabkan pembusukan bahan makanan (seperti sayur dan daging) karena bahan tersebut dibiarkan tidak diolah [4], [5], [6]. Kesulitan menemukan informasi yang relevan dari banyaknya pilihan umumnya diatasi melalui sistem rekomendasi. Hal tersebut diatasi sistem rekomendasi dengan cara merekomendasikan item yang sesuai dengan preferensi pengguna [7]. Oleh karena itu diperlukan sebuah sistem rekomendasi resep makanan yang dapat membantu pengguna dalam menentukan pilihan efektif dan optimal.

Penelitian terdahulu telah mengembangkan sistem rekomendasi resep makanan menggunakan pendekatan *Content-Based Filtering* (CBF). Metode ini memanfaatkan representasi fitur berbasis teks seperti *Term Frequency-Inverse Document Frequency* (TF-IDF) yang dikombinasikan dengan *Cosine Similarity* untuk mengukur tingkat kemiripan antar resep berdasarkan bahan atau kata kunci tertentu. Evaluasi pada penelitian sebelumnya menunjukkan tingkat kesesuaian rekomendasi yang cukup baik berdasarkan persepsi responden, yang mencapai 71% [6]. Penelitian lain membangun hal serupa menggunakan evaluasi metrik *Root Mean Square Error* (RMSE) dengan nilai 0.3, yang menunjukkan *error* relatif tidak terlalu besar [8]. Walaupun demikian, pendekatan CBF memiliki keterbatasan karena hanya mengandalkan karakteristik konten tanpa mempertimbangkan pola interaksi

pengguna. Akibatnya, sistem kurang mampu dalam menghasilkan rekomendasi yang *personal* [9].

Di sisi lain, pendekatan *Collaborative Filtering* (CF) memungkinkan untuk memberikan rekomendasi yang terpersonalisasi dengan cara mempelajari pola preferensi berdasarkan riwayat interaksi antar pengguna dan *item* [10]. Sebuah penelitian telah mengembangkan model hibrida yang menggabungkan CBF berbasis TF-IDF dan CF berbasis *Neural Collaborative Filtering* (NCF). Berdasarkan penelitian tersebut, Model CF yang digunakan dapat memberikan sistem rekomendasi yang menangkap pola preferensi pengguna. Meskipun demikian, model *deep learning* seperti NCF memiliki keterbatasan, terutama dari segi kompleksitas arsitektur dan ketergantungan pada berbagai *hyperparameter* yang menyebabkan waktu pelatihan lebih lama serta kebutuhan komputasi yang tinggi. Hal ini menjadi tantangan dalam penerapan sistem rekomendasi berbasis web [3].

Dalam rangka mengatasi keterbatasan tersebut, kami memilih untuk menggunakan pendekatan CF dengan model yang memiliki waktu pelatihan lebih singkat, dan kebutuhan komputasi yang efisien untuk memungkinkan penerapan sistem rekomendasi makanan yang *personal* dan relevan berbasis web secara efektif. Terdapat sebuah penelitian komparatif yang dilakukan pada tahun 2025 terhadap berbagai model sistem rekomendasi collaborative filtering. Di antara model yang dibandingkan, *Embarrassingly Shallow Autoencoder for Sparse Data* (EASE-R) menawarkan keseimbangan yang optimal antara akurasi, efisiensi komputasi, dan kecepatan pemrosesan. Model ini mampu menangani *dataset* berskala besar dengan baik serta memberikan respons yang cepat, sehingga dipilih sebagai pendekatan utama dalam penelitian ini [11]. Hal ini juga selaras dengan penelitian dimana EASE-R pertama kali ditemukan. EASE-R mencapai efisiensi dan kecepatan tinggi melalui penggunaan solusi bentuk tertutup (*closed-form solution*) pada model linear sederhana tanpa lapisan tersembunyi yang memungkinkan pelatihan dilakukan dalam satu langkah komputasi saja. Sementara, akurasi pada skala besar didukung oleh penerapan batasan diagonal nol dan pemodelan matriks *item-to-item*, yang kompleksitasnya tidak bergantung pada jumlah pengguna [12].

Berdasarkan uraian tersebut, penelitian ini mengusulkan implementasi sistem rekomendasi resep makanan berbasis web menggunakan metode CF dengan model EASE-R. Sistem yang dikembangkan diharapkan mampu menghasilkan rekomendasi yang *personal*, relevan, dan efisien dalam biaya komputasi dan waktu pelatihan, sehingga dapat meningkatkan efektivitas pengguna dalam menemukan resep sesuai preferensinya.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang, permasalahan yang diangkat dalam penelitian ini adalah banyaknya individu yang mengalami kesulitan dalam menemukan resep makanan yang sesuai dengan preferensinya di tengah melimpahnya informasi digital. Maka, diperlukan suatu sistem rekomendasi yang mampu memberikan rekomendasi resep yang *personal* (sesuai dengan preferensi pengguna) dengan cara mempelajari pola interaksi pengguna lain.

1.3 Tujuan

Berdasarkan rumusan masalah tersebut, tujuan penelitian ini adalah untuk merancang dan mengimplementasikan sistem rekomendasi resep makanan berbasis web menggunakan metode CF dengan model EASE-R untuk memberikan rekomendasi resep yang *personal* (sesuai dengan preferensi pengguna) dengan cara mempelajari pola interaksi pengguna lain.

1.4 Manfaat

Manfaat yang didapatkan dari penelitian tugas akhir ini adalah:

1. Mempermudah pengguna untuk menemukan ide resep makanan yang *personal* (sesuai dengan preferensi pengguna) berdasarkan pola interaksi pengguna lain.
2. Menunjukkan potensi model *Embarrassingly Shallow Autoencoder for Sparse Data* (EASE-R) sebagai metode *collaborative filtering* yang menawarkan keseimbangan yang optimal antara akurasi, efisiensi komputasi, dan kemampuan menangani *dataset* berskala besar dalam aplikasi web
3. Memberikan referensi implementasi model EASE-R dalam sistem rekomendasi berbasis web yang dapat dijadikan acuan untuk pengembangan penelitian sejenis.

1.5 Ruang Lingkup

Ruang lingkup penelitian ini disusun untuk membatasi pembahasan agar penelitian tetap terfokus pada tujuan yang telah ditetapkan serta dapat dilaksanakan secara sistematis tanpa mengurangi manfaat yang diharapkan. Adapun ruang lingkup dalam penelitian ini adalah sebagai berikut:

1. Dataset penelitian diperoleh dari platform kaggle.com, bernama Nutrigo_Dataset. Dataset berisi *recipe_file*, *interaction_file*, dan *images*. Data yang digunakan dalam

pembangunan model sistem rekomendasi ini berupa data interaksi pengguna dan item yang direpresentasikan dalam bentuk *user-item interaction matrix*, tanpa mempertimbangkan atribut konten dari item seperti deskripsi, kategori, maupun metadata lainnya.

2. Penelitian ini berfokus pada pengembangan sistem rekomendasi berbasis *collaborative filtering*, dengan menggunakan model EASE-R, tanpa melibatkan *content-based filtering* maupun *hybrid recommendation*.
3. Proses evaluasi sistem rekomendasi dilakukan menggunakan metrik evaluasi rekomendasi yang umum digunakan, seperti Recall@K dan NDCG@K dan membandingkannya dengan *Popularity Baseline* sebagai tolak ukur, untuk mengukur kualitas hasil rekomendasi yang dihasilkan oleh model.
4. *Output* sistem berupa daftar rekomendasi resep makanan yang diurutkan berdasarkan tingkat kemiripan tertinggi. *Output* tersebut dihasilkan berdasarkan *Input* sistem, yaitu interaksi pengguna dengan resep makanan yang tersedia,
5. Pengujian sistem aplikasi web secara fungsional dilakukan menggunakan metode *Black Box Testing*.
6. Sistem berfokus pada rekomendasi resep makanan internasional, dan berbahasa inggris.
7. Model *Collaborative Filtering* EASE-R dikembangkan menggunakan bahasa pemrograman *Python*. Sementara pengembangan web menggunakan bahasa pemrograman *Javascript*.
8. Skema pembagian dataset yang digunakan pada penelitian ini adalah 80/10/10, dimana 80% merupakan data pelatihan, 10% merupakan data uji, dan sisanya menjadi data validasi. Data validasi digunakan untuk mencari nilai parameter regularisasi yang tepat.
9. Terdapat beberapa tahap *pre-processing* yang dilakukan pada dataset, tahap-tahap tersebut dapat dibagi sebagai berikut:
 - a. Split dataset 80/10/10
 - b. Menghapus user yang memiliki kurang dari 5 interaksi
 - c. Menghapus *item* resep yang memiliki kurang dari 5 interaksi
 - d. Menghapus user yang memiliki 500 atau lebih interaksi karena dianggap sebagai *outlier*
 - e. Konversi data interaksi dari eksplisit ke implisit
 - f. Pengaplikasian label *encoding*
 - g. Konversi data akhir ke *Sparse Matrix*