

**HYBRID FOOD RECOMMENDATION SYSTEM USING
TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY
(TF-IDF), K-NEAREST NEIGHBORS (KNN), AND TAG-
BASED SIMILARITY**

LAPORAN NON-REGULER

[SCIENTIST]

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



Disusun oleh :

JUVENTANIA SHEVA MELLANY

22.11.5154

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

**HYBRID FOOD RECOMMENDATION SYSTEM USING TERM-FREQUENCY-
INVERSE DOCUMENT FREQUENCY (TF-IDF), K-NEAREST NEIGHBORS
(KNN), AND TAG-BASED SIMILARITY**

LAPORAN NON-REGULER

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



Disusun oleh :

JUVENTANIA SHEVA MELLANY

22.11.5154

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

JALUR NON-REGULER

**HYBRID FOOD RECOMMENDATION SYSTEM USING TERM FREQUENCY-
INVERSE DOCUMENT FREQUENCY (TF-IDF), K-NEAREST NEIGHBORS
(KNN), AND TAG-BASED SIMILARITY**

yang disusun dan diajukan oleh

Juventania Sheva Mellany

22.11.5154

telah disetujui oleh Dosen Pembimbing

pada tanggal 23 Februari 2026

Dosen Pembimbing,



Kusnawi S.Kom., M.Eng

NIK. 190302112

JALUR NON-REGULER

**HYBRID FOOD RECOMMENDATION SYSTEM USING TERM FREQUENCY-
INVERSE DOCUMENT FREQUENCY (TF-IDF), K-NEAREST NEIGHBORS
(KNN), AND TAG-BASED SIMILARITY**

yang disusun dan diajukan oleh

Juventania Sheva Mellany
22.11.5154

Telah dipertahankan di depan Dewan Penguji
pada tanggal 23 Februari 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Uyock Anggoro Saputro, S.Kom., M.Kom
NIK. 190302419

Majid Rahardi, S.Kom., M.Eng
NIK. 190302393

Kusnawi, S.Kom., M.Eng
NIK. 190302112



Laporan ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer

Tanggal 23 Februari 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrinl, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : Juventania Sheva Mellany

NIM : 22.11.5154

Menyatakan bahwa Laporan dengan judul berikut:

Hybrid Food Recommendation System Using Term-Frequency Inverse Document Frequency (TF-IDF), K-Nearest Neighbors (KNN), and Tag-Based Similarity

Dosen Pembimbing : Kusnawi S.Kom., M.Eng

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan kegiatan SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidak-benaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta,

Yang Menyatakan,



Juventania Sheva Mellany

HALAMAN PERSEMBAHAN

1. Untuk ayahanda tercinta, Ayah Jaka Sidik Purnama. Terima kasih karena sudah berjuang, bertahan, dan menyayangi penulis sampai akhir. Terima kasih karena tidak memilih untuk menyerah. Beliau memang tidak merasakan pendidikan sampai bangku perkuliahan, namun beliau mampu mendidik penulis, memotivasi, memberikan dukungan dan semangat agar penulis mampu menyelesaikan studinya. Berkat beliau lah penulis bisa menjadi pribadi yang mandiri, pantang menyerah dan memiliki keberanian untuk terus melangkah.
2. Untuk ibunda tercinta, Ibu Heny Rahmawati yang menjadi alasan utama penulis bisa bertahan hingga saat ini. Terima kasih atas segala motivasi, semangat, harapan serta bersedia menjadi sandaran terkuat dari kerasnya dunia. Terima kasih karena tidak pernah menuntut akan segala hal dan bahkan senantiasa mendampingi setiap langkah penulis untuk menjadi seseorang yang berpendidikan. Terima kasih atas kasih sayang yang diberikan, mendoakan tanpa henti dan terima kasih atas kesabaran serta pengorbanan yang selalu mengiringi perjalanan hidup penulis. Beliau adalah wanita hebat yang tidak pandai mengeluh juga pandai menyembunyikan segala lukanya sendirian, dan beliau lah sumber kekuatan dan inspirasi penulis. Sekali lagi, terima kasih untuk cucur keringat dan kerja keras yang sudah engkau berikan agar penulis bisa sampai di titik ini.
3. Untuk adik tersayang, Fabyent Janitra Pradadigdy. Terima kasih karena sudah menjadi sumber semangat dan motivasi dalam setiap langkah perjuangan penulis. Meski kami kadang tidak akur, namun dialah salah satu alasan penulis terus berusaha menempuh Pendidikan dengan sungguh-sungguh, agar kelak dia dapat menikmati dan mencontoh tanpa perlu bersusah payah menata hidup dan kariernya dari awal. Semoga apa yang kakak perjuangkan hari ini dapat menjadi pijakan bagi masa depanmu yang lebih baik.

4. Untuk keluarga besar yang tidak bisa penulis sebut satu-persatu namanya. Terima kasih sudah mendukung dan memotivasi penulis hingga penulis bisa sampai di titik ini. Terima kasih untuk segala dukungan baik secara moral maupun finansial.
5. Untuk teman-teman di rumah, Catur, Keisza, Bimbi, Amel, Marshela, Anjeli dan Nanda. Terima kasih karena sudah mendukung dan memotivasi penulis. Terima kasih sudah menjadi sandaran terbaik saat penulis berada dalam keadaan terpuruk.
6. Untuk teman-teman di perantauan, Ika, Fera, Suci, Sekar, Dea, Karen, Sukma, Kak Bella, dan teman-teman lainnya yang tidak bisa penulis sebutkan namanya. Terima kasih karena sudah bersedia menemani penulis. Terima kasih sudah menjadi teman terbaik yang selalu menuntun apabila penulis mengalami kebingungan. Terima kasih sudah menjadi inspirasi dan selalu menegur penulis apabila penulis melakukan kesalahan.
7. Terakhir, untuk saya sendiri, Juventania Sheva Mellany. Anak pertama perempuan yang dipaksa kuat sampai akhir, terima kasih karena sudah mampu bertahan sejauh ini. Terima kasih karena tidak menyerah ketika jalan di depan terasa gelap, ketika keraguan datang silih berganti, dan ketika langkah terasa berat untuk diteruskan. Terima kasih karena memilih untuk melanjutkan, walau seringkali tidak tahu pasti kemana arah ini akan membawa. Terima kasih karena sudah mempercayai proses, meski kadang hasil belum sesuai dengan harapan. Meskipun menghadapi kegagalan, kebingungan, bahkan perasaan menyerah. Terima kasih karena tetap jujur pada rasa takut. Terima kasih karena sudah berani memilih, memilih untuk mencoba, memilih untuk belajar, dan memilih untuk menyelesaikan apa yang kamu mulai. Dan sekali lagi, terima kasih sudah mencintai diri sendiri.

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kehadiran Allah SWT atas rahmat, karunia, dan kekuatan yang senantiasa menyertai setiap proses hingga laporan ini dapat diselesaikan. Penyusunan laporan ini merupakan bagian dari tanggung jawab akademik sekaligus proses pembelajaran yang penulis jalani dengan sungguh-sungguh. Penulis menyampaikan terima kasih kepada:

1. Orang tua dan adik tercinta, yang senantiasa menghadirkan doa, dukungan, dan kekuatan dalam setiap langkah perjuangan penulis.
2. Prof. Dr. Mohammad Suyanto, M.M selaku Rektor Universitas Amikom Yogyakarta.
3. Prof. Dr. Kusrini, M.Kom selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
4. Eli Pujiastuti, M.Kom selaku Ketua Program Studi Informatika Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
5. Kusnawi, S.Kom, M.Eng selaku Dosen Pembimbing yang telah memberikan bimbingan, arahan, serta masukan selama proses penelitian dan penulisan laporan ini.
6. Tim Dosen Penguji yang telah memberikan kritik dan saran konstruktif guna penyempurnaan penelitian ini.

Penulis menyadari bahwa laporan ini masih memiliki keterbatasan. Oleh karena itu, segala bentuk saran dan masukan yang membangun akan sangat berarti untuk perbaikan di masa mendatang. Semoga laporan ini dapat memberikan manfaat dan menjadi bagian dari proses pengembangan keilmuan yang berkelanjutan.

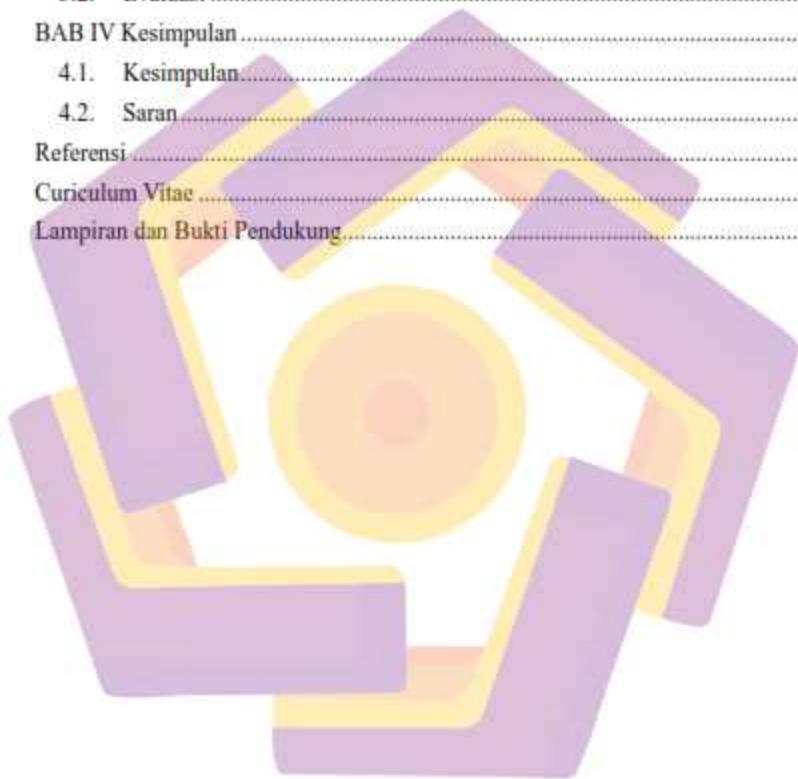
Yogyakarta, 08 Februari 2026

Penulis

DAFTAR ISI

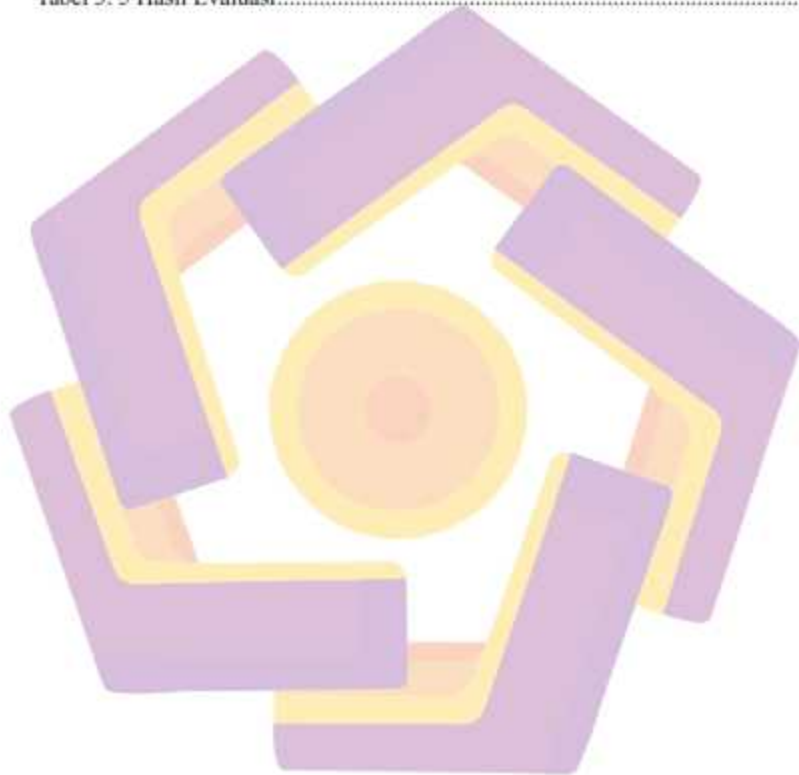
Halaman Persetujuan.....	ii
Halaman Pernyataan Keaslian Karya (*).....	iv
Halaman Persembahan.....	v
Kata Pengantar.....	vii
Daftar Isi (*).....	viii
Daftar Tabel(*).....	x
Daftar Gambar (*).....	xi
Daftar Lampiran (*).....	xii
Daftar Lambang dan Singkatan.....	xiii
Daftar Istilah.....	xiv
Intisari.....	xvi
Abstract.....	xvii
Bab I Pendahuluan.....	1
1.1. Gambaran Umum (*).....	1
1.2. Rumusan Masalah.....	2
1.3. Batasan Masalah.....	2
1.4. Tujuan.....	3
Bab II Teori dan METODE.....	4
2.1. Teori.....	4
2.1.1. Konsep Dasar Sistem Rekomendasi.....	4
2.1.2. Pendekatan Sistem Rekomendasi.....	5
2.1.3. Tahapan Pengolahan Data.....	7
2.1.4. Representasi Fitur dan Pengukuran Kemiripan.....	9
2.1.5. Algoritma K-Nearest Neighbors.....	10
2.1.6. Evaluasi Kinerja Sistem.....	11
2.2. Analisis.....	11
2.2.1. Data Collection.....	12
2.2.2. Data Understanding.....	12
2.2.3. Preprocessing Data.....	13
2.2.4. Hybrid Recommendation Modelling.....	15

2.2.5. System Evaluation.....	18
BAB III Hasil dan Pembahasan	21
3.1. Hasil.....	21
3.1.1. Hasil Preprocessing.....	21
3.1.2 Hybrid Recommendation Modelling.....	22
3.2. Evaluasi	26
BAB IV Kesimpulan	30
4.1. Kesimpulan.....	30
4.2. Saran	31
Referensi	33
Curriculum Vitae	38
Lampiran dan Bukti Pendukung.....	39



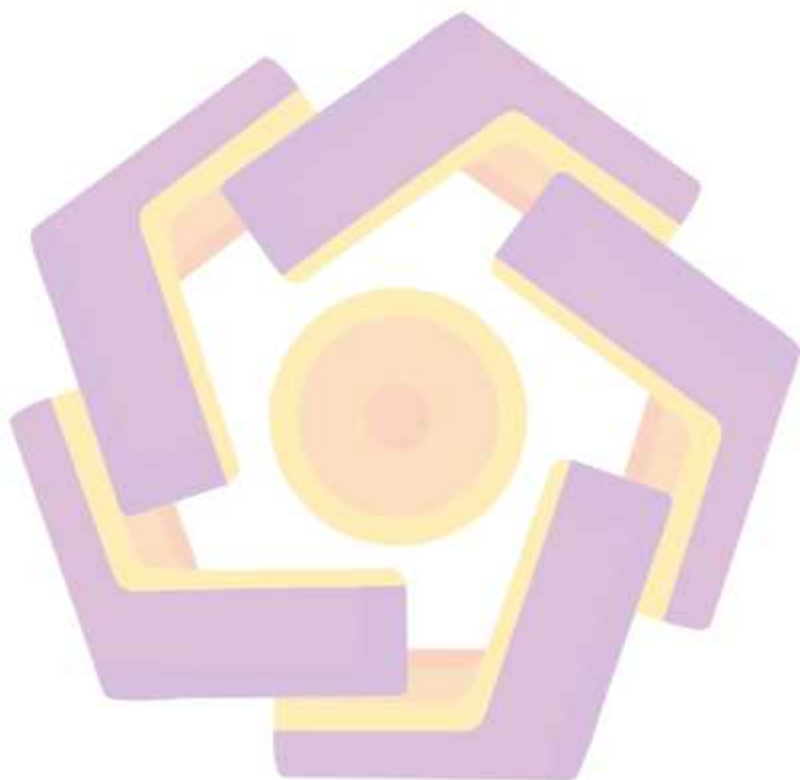
DAFTAR TABEL

Tabel 3. 1 Contoh Hasil <i>One Hot Encoding</i>	22
Tabel 3. 2 Hasil <i>Popularity-Based Ranking</i>	23
Tabel 3. 3 Hasil TF-IDF KNN <i>Item Similarity</i>	23
Tabel 3. 4 Hasil <i>Tag-Based Cosine Matching</i>	25
Tabel 3. 5 Hasil Evaluasi.....	27



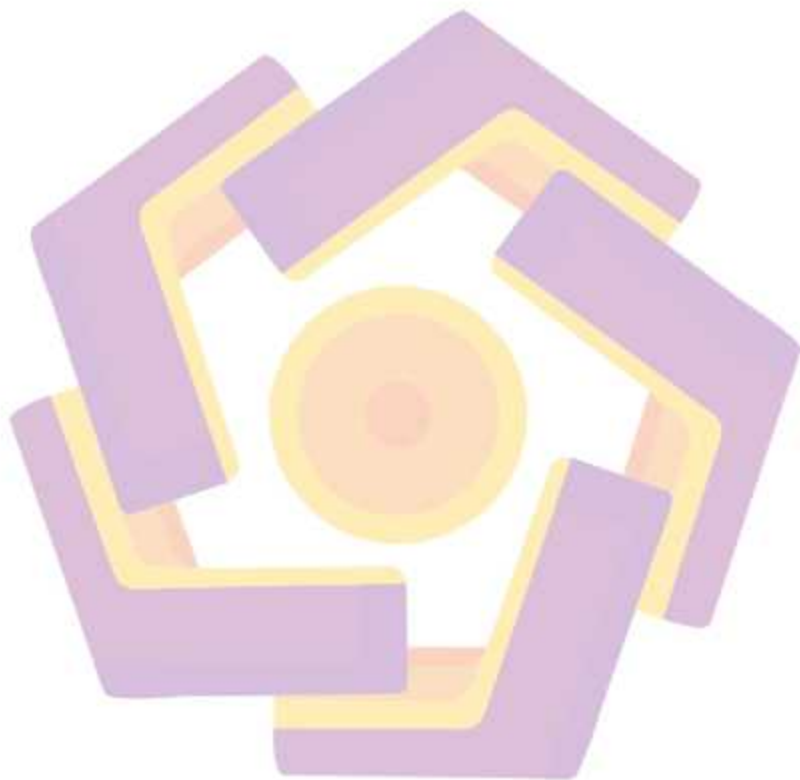
DAFTAR GAMBAR

Gambar 2. 1 Alur Penelitian.....	12
Gambar 3. 1 Chart Evaluasi.....	28



DAFTAR LAMPIRAN

Lampiran 1. Letter Of Acceptance	39
Lampiran 2. Lembar Review.....	40
Lampiran 3. Bukti Pembayaran.....	42
Lampiran 4. Bukti Publikasi Artikel Ilmiah	43



DAFTAR LAMBANG DAN SINGKATAN



TF-IDF	<i>Term Frequency-Inverse Document Frequency</i>
KNN	<i>K-Nearest Neighbors</i>
CBF	<i>Content-Based Filtering</i>
EDA	<i>Exploratory Data Analysis</i>
OHE	<i>One-Hot Encoding</i>
TP	<i>True Positive</i>
TN	<i>True Negative</i>
FP	<i>False Positive</i>
FN	<i>False Negative</i>
μ	Rata-rata (<i>mean</i>)
σ	Simpangan Baku
A	Vektor fitur item pertama
B	Vektor fitur item kedua
$\cos(\theta)$	Nilai <i>cosine similarity</i> antara dua vektor

DAFTAR ISTILAH

<i>Hybrid Recommendation System</i>	Sistem rekomendasi yang menggabungkan lebih dari satu metode rekomendasi dalam satu kerangka
<i>Popularity-Based Ranking</i>	Metode rekomendasi yang menyusun item berdasarkan tingkat popularitas atau frekuensi pemesanan
<i>Item Similarity</i>	Pengukuran tingkat kemiripan antar item berdasarkan fitur atau karakteristiknya
<i>Tag-Based Similarity</i>	Pengukuran kemiripan item berdasarkan kesamaan atribut tag
<i>Cosine Similarity</i>	Pengukuran kemiripan antara dua vektor berdasarkan sudut di antaranya
<i>Term Frequency-Inverse Document Frequency</i>	Metode pembobotan kata untuk merepresentasikan teks dalam bentuk vektor numerik
<i>K-Nearest Neighbors</i>	Algoritma yang menentukan kemiripan berdasarkan sejumlah tetangga terdekat
<i>Content-Based Filtering</i>	Pendekatan rekomendasi berbasis kemiripan karakteristik item
<i>Exploratory Data Analysis</i>	Tahap eksplorasi awal untuk memahami karakteristik dataset
<i>Preprocessing</i>	Tahap pembersihan dan transformasi data sebelum digunakan dalam pemodelan
<i>Data Feature Extraction</i>	Proses pembentukan fitur dari data mentah menjadi representasi numerik.
<i>df_penjualan</i>	Dataset transaksi yang masih mempertahankan kemunculan menu berulang untuk perhitungan popularitas

df_bersih

Dataset hasil penghapusan duplikasi menu yang digunakan untuk pemodelan berbasis kemiripan

One-Hot Encoding

Teknik transformasi data kategorikal menjadi vektor biner

True Positive

Jumlah item relevan yang berhasil direkomendasikan sistem

True Negative

Jumlah item tidak relevan yang tidak direkomendasikan sistem

False Positive

Jumlah item tidak relevan yang direkomendasikan sistem

False Negative

Jumlah item relevan yang tidak direkomendasikan sistem

Accuracy

Proporsi prediksi yang benar terhadap seluruh prediksi

Precision

Proporsi rekomendasi relevan terhadap seluruh item yang direkomendasikan

Recall

Proporsi item relevan yang berhasil ditemukan sistem

F1-Score

Rata-rata harmonik antara *precision* dan *recall*

INTISARI

Pertumbuhan pesat industri kuliner digital meningkatkan kebutuhan akan sistem rekomendasi menu cerdas yang dapat membantu pelanggan dalam membuat pilihan yang akurat dan personal. Studi ini mengembangkan sistem rekomendasi makanan hibrida yang menggabungkan tiga pendekatan komplementer: peringkat berdasarkan popularitas, *Term Frequency-Inverse Document Frequency* (TF-IDF) dengan *K-Nearest Neighbors* (KNN) untuk kesamaan item, dan pencocokan *cosine* berbasis tag. Sistem ini juga dilengkapi dengan komponen *Content-Based Filtering* yang memanfaatkan kesamaan kosinus untuk memperkuat pemodelan kesamaan di antara representasi teks dan berbasis tag. Sebanyak 77.157 catatan transaksi nyata dari Restoran SR Cipali, yang dikumpulkan antara April dan Desember 2024, digunakan sebagai sumber data utama untuk pengembangan dan evaluasi sistem. *Preprocessing* data meliputi pembersihan, penyaringan kategori, transformasi TF-IDF untuk nama produk, *One-Hot Encoding* untuk tag, dan normalisasi harga untuk menghasilkan representasi fitur yang terstruktur dan dapat dibandingkan. Evaluasi dilakukan menggunakan *confusion matrix* dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score* untuk membandingkan kinerja ketiga pendekatan yang diusulkan. Hasil eksperimen menunjukkan bahwa model TF-IDF KNN mencapai kinerja terbaik dengan *accuracy* 0,94, *precision* 0,80, *recall* 1,00, dan *F1-score* 0,89. Model berbasis popularitas memperoleh *accuracy* 0,89 dengan *precision* dan *recall* sebesar 0,80, sedangkan model berbasis tag mencapai *precision* 1,00 namun memiliki *recall* lebih rendah akibat ketidakkonsistenan tag dan selektivitas peringkat. Perbandingan tersebut menegaskan bahwa pendekatan TF-IDF KNN memberikan performa paling optimal dalam konteks data transaksi nyata yang digunakan. Keunikan studi ini terletak pada penggunaan kerangka kerja hibrida ringan yang dievaluasi pada transaksi restoran dunia nyata yang jarang digunakan dalam penelitian sebelumnya yang didominasi oleh dataset *benchmark*, sehingga sistem yang diusulkan menunjukkan praktis yang kuat bagi restoran kecil dan menengah yang kekurangan data penilaian dan masih dapat dikembangkan lebih lanjut melalui peningkatan kualitas tag serta penambahan atribut produk.

Kata kunci: sistem rekomendasi, TF-IDF, KNN, cosine similarity, rekomendasi makanan

ABSTRACT

The rapid growth of the digital culinary industry increases the need for intelligent menu recommendation systems that can assist customers in making accurate and personalized choices. This study develops a hybrid food recommendation system that integrates three complementary approaches: popularity-based ranking, Term Frequency-Inverse Document Frequency (TF-IDF) with K-Nearest Neighbors (KNN) item similarity, and tag-based cosine matching. The system also incorporates a Content-Based Filtering component that leverages cosine similarity to strengthen similarity modeling across textual and tag-based representations. A total of 77,157 real transaction records from SR Cipali Restaurant, collected between April and December 2024, were used as the primary data source for system development and evaluation. Data preprocessing includes cleaning, category filtering, TF-IDF transformation for product names, One-Hot Encoding for tags, and price normalization to generate structured and comparable feature representations. Evaluation was conducted using a confusion matrix with accuracy, precision, recall, and F1-score metrics to compare the performance of the three proposed approaches. Experimental results indicate that the TF-IDF KNN model achieved the best performance with an accuracy of 0.94, precision of 0.80, recall of 1.00, and an F1-score of 0.89. The popularity-based model obtained an accuracy of 0.89 with balanced precision and recall of 0.80, while the tag-based model achieved a precision of 1.00 but demonstrated lower recall due to tag inconsistency and limited tag representation. The comparison of the three models confirms that the TF-IDF KNN approach provides the most optimal performance within the context of real-world restaurant transaction data. The uniqueness of this study lies in the application of a lightweight hybrid framework evaluated on real restaurant transactions, which are rarely utilized in previous studies dominated by benchmark datasets. The proposed system demonstrates strong practical applicability for small and medium-sized restaurants lacking rating data and can be further improved by enhancing tag quality and incorporating additional product attributes.

Keyword: recommender system, TF-IDF, KNN, cosine similarity, food recommendation