

**SKRIPSI**

**IMPLEMENTASI METODE *HYBRID FILTERING*  
(*COLLABORATIVE FILTERING* DAN *CONTENT-BASED  
FILTERING*) DALAM MEMBERIKAN REKOMENDASI KOST**

***IMPLEMENTATION OF HYBRID FILTERING METHOD  
(COLLABORATIVE FILTERING AND CONTENT-BASED  
FILTERING) IN PROVIDING BOARDING HOUSE  
RECOMMENDATIONS***

Diajukan untuk memenuhi sebagai persyaratan  
memperoleh gelar Sarjana Teknik



**PADILA AULIA**

**D0220353**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS SULAWESI BARAT  
2024**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**IMPLEMENTASI *HYBRID FILTERING (COLLABORATIVE FILTERING DAN CONTENT-BASED FILTERING)* DALAM  
MEMBERIKAN REKOMENDASI KOST**

Telah dipersiapkan dan disusun oleh

**PADILA AULIA  
D0220353**

Telah dipertahankan di depan Tim Penguji  
Pada tanggal 30 November 2023

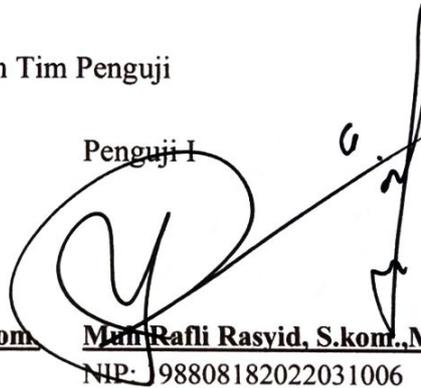
Susunan Tim Penguji

Pembimbing I



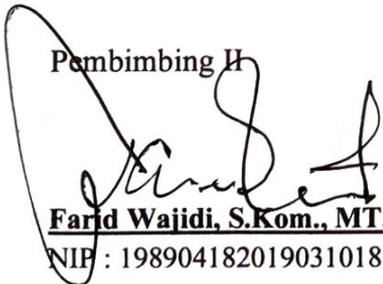
Nuralamsah Zulkarnaim S.Kom., M.Kom.  
NIP: 19891014201031013

Penguji I



Mah Rafli Rasyid, S.kom., M.T  
NIP: 98808182022031006

Pembimbing II



Farid Wajidi, S.Kom., MT.  
NIP : 198904182019031018

Penguji II



Wawan Firgiawan, S.T., M.Kom  
NIDK : 8948080023

Penguji III



A.Amirul Asnan Cirua, S.T., M.Kom  
NIP: 19980402202406100

# LEMBAR PENGESAHAN

## SKRIPSI

### IMPLEMENTASI *HYBRID FILTERING (COLLABORATIVE FILTERING DAN CONTENT-BASED FILTERING)* DALAM MEMBERIKAN REKOMENDASI KOST

Disusun dan diajukan oleh:

**PADILA AULIA**

**NIM. D0220353**

Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam rangka Penyelesaian Studi Program Sarjana Teknik Informatika Fakultas Teknik

Universitas Sulawesi Barat

pada tanggal 19 Desember 2024

dan dinyatakan telah memenuhi syarat kelulusan.

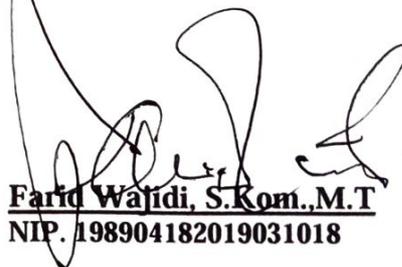
Menyetujui,

Pembimbing I



**Nuralamsah Zulkarnaim, S.Kom., M.Kom**  
NIP.19891014201031013

Pembimbing II



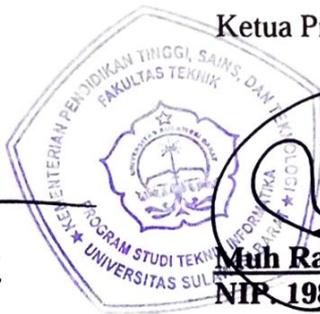
**Farid Walidi, S.Kom., M.T**  
NIP. 198904182019031018

Dekan Fakultas Teknik,  
Universitas Sulawesi Barat



**Dr. Ir. Hafsah Nirwana, M.T**  
NIP.196404051990032002

Ketua Program Studi Informatika



**Muh Rafli Rasyid, S.Kom., M.T**  
NIP. 198808182022031006

## ABSTARK

Dalam era digital, kebutuhan akan sistem informasi di bidang penyewaan tempat tinggal semakin meingkat. Teknologi informasi membantu menyediakan informasi lebih terstruktur bagi pencari kost seperti mahasiswa dan pekerja. Kost sebagai pilihan tempat tinggal populer menawarkan beragam fasilitas, harga informasi keamanan kost, tetapi pencariannya sering terkendala kurangnya informasi akurat dan rekomendasi yang relevan. Penelitian ini mengembangkan sistem rekomendasi kost di Lembang, Kabupaten Majene, menggunakan metode *Collaborative Filtering*, dan *Content-based Filtering* dalam pendekatan *Hybrid*. Kedua metode ini memiliki kekurangan dan kelebihan masing-masing solusi *Hybrid Filtering* menjadi pilihan terbaik. Penelitian ini bertujuan untuk mengimplementasikan metode *Hybrid Filtering* pada sistem rekomendasi kost. Jenis penelitian menggunakan kuantitatif dan jenis data menggunakan data primer yang diperoleh dari wawancara dan kusioner. Jumlah data responden didapatkan sebanyak 111 data dan jumlah data kost sebanyak 60 data. Hasil *Hybrid Filtering* menggunakan *Weight* mendapatkan skor gabungan yang cukup baik dan hasilnya dapat menutupi masing-masing kekurangan metode Hasil MSE 15,888 dan RMSE 3,986. Dapat disimpulkan masih kurang baik karena nilai MSE dan RMSE masih cukup tinggi.

**Kata Kunci** : *Sistem Rekomendasi, Hybrid Filtering, Collaboratiive Filtering, Content-based Filtering*

# **BAB I**

## **PENDAHULUAN**

### **A. Latar Belakang Masalah**

Tuntutan kebutuhan terhadap sistem informasi kian hari kian meningkat hampir di setiap kebutuhan tak terkecuali di bidang penyewaan tempat tinggal. Hal ini didukung oleh perkembangan dunia teknologi informasi yang mampu membuat setiap kegiatan atau informasi harus melakukan peninjauan pekerjaan dengan memanfaatkan teknologi informasi tersebut (Selviana and Dwi Utari 2019). Pemanfaatan sistem dan teknologi informasi secara digital mampu memberikan solusi terhadap permasalahan tersebut. Teknologi dan informasi yang di rancang dapat memberikan gambaran kepada masyarakat secara umum dan kepada pencari tempat tinggal khususnya para pekerja dan mahasiswa yang ingin dengan kampus ataupun tempat kerja (Syam 2018).

Kost-kost adalah unit tempat tinggal yang biasanya disewakan kepada individu atau mahasiswa yang memerlukan tempat tinggal sementara. Unit-unit ini umumnya menawarkan fasilitas dasar seperti kamar tidur, kamar mandi yang bisa bersama atau pribadi, serta dapur untuk memasak. Karena sifatnya yang praktis dan terjangkau, kost-kost sering kali menjadi pilihan populer bagi mereka yang mencari tempat tinggal sementara atau yang lebih terjangkau dibandingkan apartemen atau rumah sewa lainnya. Kost-kost sangat penting bagi masyarakat yang jauh dari kampung halaman mereka, misalnya, karena alasan pendidikan, pekerjaan, atau

kebutuhan lainnya. Setiap daerah, terutama di kota-kota besar, menawarkan beragam kost dengan berbagai variasi fasilitas, harga, dan lokasi. Namun, pencarian kost yang ideal sering menghadapi kendala utama, yaitu kesulitan dalam menemukan kost yang baik dan nyaman. Kendala ini seringkali disebabkan oleh kurangnya informasi yang akurat, yang biasanya diperoleh melalui teman atau postingan di media sosial. Informasi tersebut sering kali tidak akurat, dengan perbedaan harga yang signifikan antara yang diposting online dan kenyataan di lapangan. Masalah lainnya termasuk kurangnya presensi pengguna karena pemahaman yang kurang tentang pentingnya presensi ini. Selain itu, lokasi kost yang kurang aman, dengan risiko pencurian atau insiden lainnya, menjadi perhatian utama bagi para pencari kost. Oleh karena itu, pencari kost cenderung mencari tempat yang sesuai dengan kebutuhan dan preferensi mereka untuk memastikan kenyamanan mereka selama tinggal.

Fasilitas yang ditawarkan oleh penyedia kost juga menjadi faktor utama dalam pemilihan kost. Fasilitas-fasilitas tersebut meliputi koneksi wifi, suplai air, listrik, kamar mandi dalam atau luar, serta dapur untuk memasak. Fasilitas-fasilitas ini sering kali menjadi pertimbangan utama dalam pencarian kost, karena mereka mempengaruhi kenyamanan dan kualitas hidup penghuni. Dengan berbagai macam fasilitas yang ditawarkan, harga kost pun bervariasi. Misalnya, kost dengan fasilitas lengkap seperti wifi dan kamar mandi dalam biasanya memiliki harga yang lebih tinggi dibandingkan dengan kost yang hanya menyediakan fasilitas dasar. Oleh karena itu, para pencari kost cenderung memilih tempat tinggal yang sesuai dengan fasilitas yang mereka inginkan dan butuhkan, karena ini menjadi faktor utama

dalam menentukan harga kost. Fasilitas yang memadai sangat penting karena berkontribusi langsung pada kenyamanan dan kesejahteraan penghuni selama tinggal di kost. Selain itu untuk air bersih dan listrik yang digunakan umumnya biasanya tidak termasuk dalam harga kost. Hal itu bisa menjadi pertimbangan para pencari kost tergantung keinginan mereka masing-masing. Oleh karena itu pentingnya informasi yang relevan dalam memberikan rekomendasi kost.

Lembang, Kecamatan Banggae Timur di Kabupaten Majene adalah salah satu kecamatan yang menawarkan banyak pilihan kost, dari yang murah hingga yang mahal, tergantung pada fasilitas yang disediakan oleh penyedia kost. Lokasi ini menjadi pusat pencarian kost, terutama bagi mahasiswa yang ingin tinggal dekat dengan kampus untuk memudahkan akses ke kegiatan akademis. Lembang menjadi salah satu pilihan utama karena banyaknya kost yang tersedia dengan berbagai variasi harga dan fasilitas, sehingga dapat memenuhi kebutuhan berbagai kalangan. Penelitian ini akan dilakukan di Lembang, Kecamatan Banggae Timur karena daerah ini dikenal sebagai salah satu lokasi di Kabupaten Majene yang banyak menyediakan kost dan menjadi pilihan utama bagi pencari kost, terutama mahasiswa.

Sistem rekomendasi adalah teknologi yang digunakan untuk menyarankan produk, layanan, atau konten kepada pengguna berdasarkan analisis data pengguna sebelumnya dan perilaku. Sistem ini sangat penting dalam era digital karena membantu mengatasi masalah *overload* informasi dengan memberikan saran yang relevan dan personal kepada pengguna. Dalam sistem rekomendasi terdapat tiga jenis sistem rekomendasi yaitu *Collaborative Filtering*, *Content-based filtering* dan

Rekomendasi *Hybrid*. Pada penelitian ini menggunakan *Algoritma Collaborative filtering* dan *Content-based filtering*. *Algoritma Collaborative filtering* adalah suatu metode yang digunakan untuk membandingkan presensi pengguna dengan pengguna lainnya sehingga rekomendasi yang diberikan berdasarkan presensi pengguna yang sama. *Content-based filtering* adalah metode yang digunakan berdasarkan dari kriteria yang sesuai dengan presensi pengguna. Misalnya, sistem dapat merekomendasikan kost dengan fasilitas tertentu kepada pengguna yang memprioritaskan fasilitas tertentu. Ini memungkinkan rekomendasi yang lebih spesifik. Untuk rekomendasi yang lebih baik maka system ini menggunakan Rekomendasi *Hybrid* untuk menggabungkan *Algoritma Collaborative filtering* dan *Content-based filtering* agar rekomendasi kost lebih baik berdasarkan presensi pengguna dan kriteria kost.

Penelitian yang dilakukan oleh (Arifin 2022) sistem rekomendasi dirancang dan diterapkan *Algoritma Collaborative Filtering* dapat memberikan informasi dan merekomendasikan indekos berdasarkan hasil rating pengguna sebelumnya. Penelitian berikutnya dilakukan oleh (Cheryl Ayu Melyani 2022) dalam penelitiannya menerapkan *Content- Based Filtering* berdasarkan kemiripan deskripsi hotel adalah 10 hotel teratas dengan nilai kemipiripan yang terbesar. Penelitian berikutnya yang dilakukan oleh (Larasati and Februariyanti 2021) dalam penelitian sistem rekomenasi emina cosmetic dengan metode *Content- Based Filtering* menghasilkan 10 produk rekomendasi dengan hasil akhir perhitungan cosine tertinggi dengan nilai similaritas 0,7195. Penelitian selanjutnya dilakukan oleh (Diah, Rahmad, and Saleh 2019) dalam penelitiannya menyatakan dengan

Teknik penggabungan secara *Mixed Hybrid* antara metode *Content Based Filtering* dan *Collaborative Filtering* dihasilkan sebuah sistem rekomendasi pemilihan kamera yang mampu menutupi kekurangan dari masing-masing metode yang digunakan. Selanjutnya penelitian yang dilakukan oleh (Handan Siswaningrum, 2022) hasil penelitiannya menyatakan metode *Hybrid* dengan memanfaatkan dua metode rekomendasi (*Content- Based Filtering* dan *Collaborative Filtering* yang digunakan dapat memberikan performa rekomendasi yang baik untuk user yang mampu menutupi kekurangan pada masing-masing metode. Sistem mampu memberikan rekomendasi dari segi kemiripan produk dan penilaian produk terhadap pengguna lain.

Meskipun penelitian terdahulu telah memberikan pandangan yang berharga, masih ada ruang untuk pengembangan lebih lanjut. Penelitian sebelumnya tidak ada yang menyatukan metode *Collaborative Filtering* dan *Content-Based Filtering* dalam pendekatan *Hybrid* dengan rekomendasi kost yang memiliki potensi untuk memberikan rekomendasi yang lebih baik. Oleh karena itu, penelitian ini bertujuan untuk implementasi dalam memberikan rekomendasi kost yang menggunakan *Hybrid* gabungan dari metode *Collaborative Filtering* dan *Content-based Filtering* sambil mengintegrasikan informasi fasilitas, harga, lokasi, jenis kost, keamanan, informasi tambahan dan rating. Penelitian ini akan memperluas pemahaman tentang bagaimana integrasi informasi ini dapat menghasilkan rekomendasi yang lebih baik dan efisien dalam konteks praktis.

## **B. Rumusan Masalah**

Dari latar belakang masalah diatas Adapun rumusan masalah dari penelitian ini adalah:

1. Bagaimana Implementasi rekomendasi *Hybrid Filtering* dalam memberikan rekomendasi kost?

## **C. Tujuan Penelitian**

1. Penelitian ini dapat memberikan rekomendasi kost dengan implementasi rekomendasi *Hybrid Filtering*.

## **D. Batasan Masalah**

Adapun Batasan masalah yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Algoritma yang digunakan dalam penelitian ini adalah rekomendasi *Hybrid (Collaborative Filtering dan Content-Based Filtering)*
2. Kriteria yang digunakan dalam penelitian ini adalah sebagai berikut :
  - a) Fasilitas adalah perlengkapan atau apa saja yang disediakan dalam rumah kost oleh pemilik kost. Fasilitas ini terdiri dari. Fasilitas umum diantaranya dapur, kamar mandi apakah umum atau berada dalam kamar kost. Fasilitas tambahan merupakan fasilitas tambahan seperti wifi, kipas, dan kebutuhan kamar lainnya.
  - b) Harga yang dimaksud disini adalah berapa harga dari tiap kamar kost yang di sediakan harga dapat berupa per kamar, per orang dalam harga ini beberapa

penyedia kost menyediakan kamar kost dengan harga yang berbeda tiap kamar berdasarkan fasilitas yang di sediakan.

- c) Keamanan kost adalah keamanan ini terdiri dari cctv dan parkir
  - d) Jenis kost terdiri dari kost putri, kost putra dan campuran
  - e) Informasi tambahan yaitu informasi tambahan mengenai kost seperti apakah air dan listrik sudah termasuk pembayaran kost atau tidak.
  - f) Rating adalah nilai yang diberikan pada kost dari orang yang tinggal atau pernah tinggal di kost rating ini terdiri dari 5 pilihan yaitu Tidak suka, Kurang suka, Lumayan Suka, Suka dan Sangat Suka.
3. Penelitian ini dilakukan pada Lembang, Kecamatan Banggae Timur, Kabupaten Majene.

### **E. Manfaat Penelitian**

1. Menggabungkan metode *collaborative filtering* dan *content-based filtering*, sistem dapat memberikan rekomendasi kost yang lebih akurat dan relevan bagi pengguna, karena kedua pendekatan ini saling melengkapi dalam mengatasi keterbatasan masing-masing.
2. Implementasi rekomendasi *Hybrid* dapat meningkatkan pengalaman pengguna dengan menyediakan rekomendasi yang lebih tepat sesuai dengan preferensi dan kebutuhan individu, sehingga memudahkan pengguna dalam menemukan kost yang sesuai.

## **BAB II**

### **KAJIAN PUSTAKA**

#### **A. Kost**

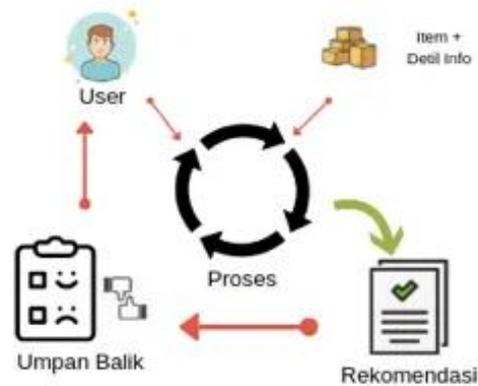
Kost adalah salah satu persewaan kamar yang sangat umum digunakan di Indonesia. Umumnya kost ini biasanya dibangun di wilayah-wilayah yang berdekatan dengan tempat-tempat umum untuk memudahkan jangkauan para pekerja, pelajar, mahasiswa dan lain sebagainya. Kost menjadi pilihan yang paling mudah dalam menentukan tempat tinggal seseorang disamping harga ssyang terjangkau fasilitas yang disediakan juga sudah sesuai kebutuhan pada umumnya. Fasilitas ini terdiri dari kamar mandi, dapur, air, listrik dan fasilitas-fasilitas tambahan lainnya yang disediakan pemilik kost. Selain itu, keamanan kost juga menjadi pertimbangan dalam pemilihan tempat kost yang aman dan nyaman.

#### **B. Sistem Rekomendasi**

Sistem rekomendasi merupakan perangkat lunak atau metode yang memberikan saran berupa tem – item spesifik yang kemungkinan besar menarik bagi pengguna tertentu (Arfisko 2022). Sistem rekomendasi telah di gunakan secara luas di berbagai bidang bisnis di mana konsumen perlu membuat keputusan atau memilih dari berbagai informasi yang tersedia. Sistem rekomendasi dapat memberikan saran – saran item yang bermanfaat bagi pengguna (Larasati and Februariyanti 2021).

Rekomendasi yang diberikan oleh sistem dapat digunakan oleh pengguna untuk membantu mereka dalam membuat keputusan, seperti memilih produk yang akan dibeli, menentukan musik yang akan didengarkan, film yang akan ditonton, atau berita yang akan dibaca. Sistem rekomendasi ini memanfaatkan data dan preferensi pengguna sebelumnya untuk memberikan saran yang lebih personal dan relevan (Larasati and Februariyanti 2021). Dengan demikian, pengguna tidak perlu menghabiskan banyak waktu mencari item yang diinginkan, karena sistem sudah memberikan opsi yang sesuai dengan selera dan kebutuhan mereka. Ini sangat berguna dalam berbagai situasi, seperti saat berbelanja online, menggunakan layanan streaming musik atau video, serta mengakses platform berita digital. Selain itu, sistem rekomendasi juga dapat meningkatkan pengalaman pengguna dengan memperkenalkan mereka pada produk, musik, film, atau berita baru yang mungkin belum pernah mereka ketahui sebelumnya, tetapi sangat sesuai dengan minat mereka.

Umumnya terdapat dua pendekatan dalam membangun sistem rekomendasi, yaitu *Collaborative Filtering* (CF), *Content-based Filtering* (CBF) dan *Hybrid* gabungan kedua metode rekomendasi. Berikut adalah gambaran alur proses dan penjelasan dalam memberikan rekomendasi. *Collaborative filtering* menganalisis preferensi pengguna lain yang serupa untuk memberikan rekomendasi. Sementara *Content-based filtering* fokus pada karakteristik item yang disukai pengguna. *Hybrid filtering* menggabungkan kedua pendekatan ini meningkatkan akurasi rekomendasi. Berikut alur Rekomendasi sistem dan penjelasan dalam memberikan rekomendasi.



**Gambar 2. 1 Alur Rekomendasi Sistem**

**(Sumber : Pandu Wibowo, 2020)**

### 1. Pengguna

Dari pengguna, sistem akan mengumpulkan informasi dan deskripsi yang mencakup riwayat pribadi, riwayat pencarian, aktivitas yang telah dilakukan, dan penilaian yang diberikan secara langsung. Informasi ini memberikan gambaran tentang preferensi dan kebiasaan pengguna, sehingga sistem dapat memahami apa yang disukai dan dibutuhkan oleh pengguna tersebut.

### 2. Item

Sementara itu, dari sisi item mengumpulkan informasi dan deskripsi yang mencakup detail tentang item tersebut serta item – item yang telah dinilai sebelumnya. Item ini lebih bisa berupa produk, barang, komponen kriteria, buku, film, musik dan sejenisnya. Informasi ini membantu sistem dalam mengidentifikasi karakteristik item dan relevansinya dengan preferensi pengguna.

### 3. Proses Rekomendasi

Dalam proses rekomendasi, data yang telah dikumpulkan dari pengguna dan item akan di proses untuk menghasilkan saran yang sesuai. Proses ini dilakukan berdasarkan pendekatan yang dipilih, seperti *Collaborative Filtering* atau *Content-based Filtering*. Pendekatan ini menentukan parameter apa saja yang di perlukan untuk menghasilkan rekomendasi yang akurat dan relevan.

### 4. Hasil Rekomendasi

Hasil dari proses rekomendasi adalah daftar item yang telah diproses menggunakan pendekatan yang dipilih. Item-item ini diharapkan sesuai dengan preferensi dan kebutuhan pengguna berdasarkan data yang telah di analisis.

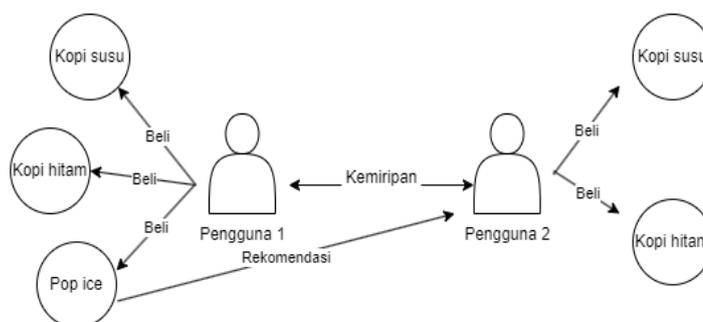
### 5. Umpan Balik

Setelah pengguna menerima rekomendasi, mereka dapat memberikan umpan balik terhadap hasil yang di berikan. Umpan balik ini bisa berupa tindakan seperti mengunjungi item yang di rekomendasikan, memberikan rating atau penilaian. Umpan balik ini sangat penting karena menunjukkan apakah pengguna menyukai atau tidak menyukai rekomendasi yang diberikan. Data umpan balik ini kemudian akan digunakan sebagai sumber informasi baru untuk memperbaiki dan menyempurnakan proses rekomendasi di masa mendatang.

Dalam sebuah artikel dikatakan sistem rekomendasi adalah solusi yang sangat efektif dalam memberikan pengguna rekomendasi yang relevan di tengah melimpahnya informasi di era digital saat ini. Selain itu, sistem ini juga berperan penting dalam meningkatkan interaksi dan keterlibatan pengguna terhadap aplikasi atau tool yang dikembangkan. Dengan adanya sistem rekomendasi

pengguna tidak hanya mendapatkan rekomendasi yang sesuai dengan presensi dan kebutuhan mereka, tetapi juga dapat menghemat waktu dalam mencari informasi yang dibutuhkan, Hal ini membantu meningkatkan efisiensi dan produktivitas pengguna dalam mengakses informasi yang relevan dan bermanfaat bagi mereka (Wibowo, Pandhu. (2020).Medium.com

### C. *Algoritma Collaborative Filtering*



**Gambar 2. 2 Konsep *Collaborative Filtering***

**(Arif Laksito 2020)**

*Algoritma Collaborative Filtering* adalah suatu konsep di mana opini dari pengguna lain yang ada digunakan untuk memprediksi item yang mungkin disukai/diminati oleh seorang pengguna (Yanto, Herwindiati, and Lauro 2021). *Algoritma Collaborative Filtering* adalah metode yang digunakan untuk memberikan rekomendasi yang menghasilkan prediksi dan rekomendasinya berbasis pada nilai dari rating dari pengguna lain dalam sistem (Wulandari and Setiawan 2022). Dengan pemfilteran kolaboratif, sistem membuat rekomendasi berdasarkan karakteristik Anda dan pengguna lain. Oleh karena itu, data historis dan ulasan pengguna merupakan bagian penting dari metode ini. Misalnya,

Pengguna 1 membeli kopi susu, kopi hitam, dan pop ice. Oleh karena itu, pengguna 2 juga kemungkinan besar akan membeli pop ice. Hal ini karena didasarkan pada karakteristik pengguna-ke-pengguna mengenai kompatibilitas. Konsep ini digunakan dalam teknik penyaringan kolaboratif.

*Algoritma Collaborative Filtering* melakukan penyaringan data berdasarkan kemiripan karakteristik konsumen sehingga mampu memberikan informasi yang baru kepada konsumen karena sistem memberikan informasi berdasarkan pola satu kelompok konsumen yang hampir sama (Arifin 2022). *Algoritma Collaborative Filtering* bekerja dengan mengumpulkan umpan balik pengguna dalam bentuk rating item-item dan kemudian memanfaatkan kesamaan perilaku antar pengguna dalam memprediksi rating untuk menentukan bagaimana suatu item di rekomendasikan (Arfisko and Wibowo 2022). Teknik ini bekerja dengan menggunakan data komunitas untuk mencari kesamaan antara pengguna. Pendekatannya adalah mengasumsikan bahwa pengguna yang memiliki preferensi yang mirip di masa lalu kemungkinan besar akan memiliki preferensi yang mirip di pula di masa depan Prinsip ini mendasari penggunaan collaborative filtering dalam menghasilkan rekomendasi item, di mana kita cenderung lebih percaya terhadap rekomendasi dari orang yang memiliki preferensi serupa dengan kita. *Algoritma Collaborative Filtering* adalah teknologi sistem rekomendasi paling sukses saat ini, dikenal dan disukai banyak orang karena sesuai dengan kebutuhan masyarakat pada umumnya (Erlangga et al. 2020).

Metode *Collaborative Filtering* menyediakan rekomendasi berdasarkan pendapat dari pengguna lain yang memiliki kebiasaan yang mirip dengan

pengguna. Dalam menghasilkan rekomendasi sistem perlu mengumpulkan data untuk memahami preferensi pengguna, sehingga dapat memberikan rekomendasi yang sesuai. Ada dua cara utama untuk mengumpulkan data dan juga cara user memberikan pendapat yaitu secara eksplisit atau implicit.

### 1. *Explicit*

*Explicit* adalah metode pengumpulan data yang memerlukan partisipasi aktif pengguna, seperti memberikan rating atau memilih *like/dislike* pada item. Metode ini membutuhkan usaha tambahan dari pengguna, dan sering kali tidak semua pengguna bersedia memberikan penilaian, yang dapat mengurangi efektivitas rekomendasi dalam *Collaborative filtering*. Pengguna secara sadar memberikan rating pada item, mencerminkan apa yang mereka sukai atau tidak sukai.

### 2. *Implicit*

*Implicit* adalah metode pengumpulan data tanpa interaksi langsung dari pengguna. Sistem secara otomatis merekam perilaku pengguna, seperti waktu yang dihabiskan melihat item, genre yang sering dilihat, dan item yang dibeli. Dengan ini, sistem mempelajari preferensi pengguna berdasarkan kebiasaan, membuatnya lebih mudah diterapkan dibandingkan metode *explicit*.

*Collaborative filtering* terbagi menjadi dua jenis, yaitu *Memory-based* dan *Model-based*. *Memory-based* melakukan komputasi secara langsung saat diperlukan secara *factorization* untuk melatih model terlebih dahulu yang dapat dilakukan secara offline. *Memory-based* unggul dalam hal kemudahan implementasi, tetapi mengalami kesulitan ketika berhadapan dengan ukuran

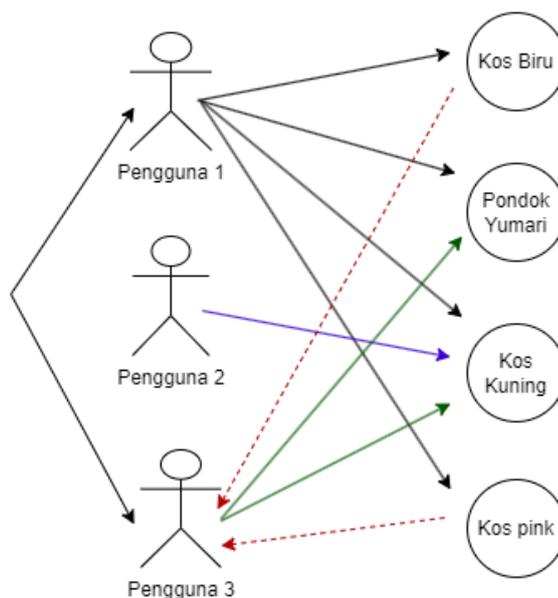
matriks yang besar sehingga mempengaruhi skalabilitas perhitungan. Sementara itu *Model-based* menggunakan beberapa metode seperti: Matrikx factorization, Neural network, Bayesian, dan lainnya. Pada penelitian ini menggunakan *Memory-based* dimana *Memory-based* itu sendiri terbagi menjadi dua kategori, yaitu : *User – based* dan *item – based* (Laksito, Arif. (2021).blog.ariflaksito)

a) *User - based collaborative filtering*

*User-based collaborative filtering* adalah teknik yang digunakan untuk memprediksi item yang mungkin disukai pengguna berdasarkan penilaian yang diberikan pada item tersebut oleh pengguna lain yang memiliki selera yang sama dengan pengguna target *based* (Laksito, Arif. (2021).Blog.ariflaksito. *User-based collaborative filtering* menemukan item menarik bagi pengguna dengan mencari pengguna lain yang memiliki minat atau keinginan serupa. Algoritma ini dimulai dengan menemukan pengguna yang mirip dan kemudian mengidentifikasi pengguna terdekat. Rating yang diberikan oleh pengguna terdekat digunakan sebagai dasar rekomendasi untuk pengguna aktif (Herny Februariyanti, Aryo Dwi Laksono, Jati Sasongko Wibowo 2021). Pendekatan *Collaborative Filtering* menggunakan metode *User-based* didasarkan pada minat pengguna yang memiliki selera atau preferensi serupa dengan memberikan peringkat atau rating pada item tertentu (Handan Siswaningrum, 2022). Tujuan dari pendekatan *User-based collacorative filtering* adalah untuk memberikan rekomendasi item kepada pengguna berdasarkan preferensi dan minat pengguna dengan preferensi, sistem dapat memanfaatkan rating dan *feedback* mereka untuk menyarankan item yang kemungkinan besar akan disukai oleh pengguna yang aktif, sehingga

meningkatkan relevansi dan kepuasan dalam rekomendasi yang diberikan.

Konsep *User-based Filtering* dapat dilihat pada gambar 2.3 berikut.



**Gambar 2. 3 Konsep *User- Based Filtering***

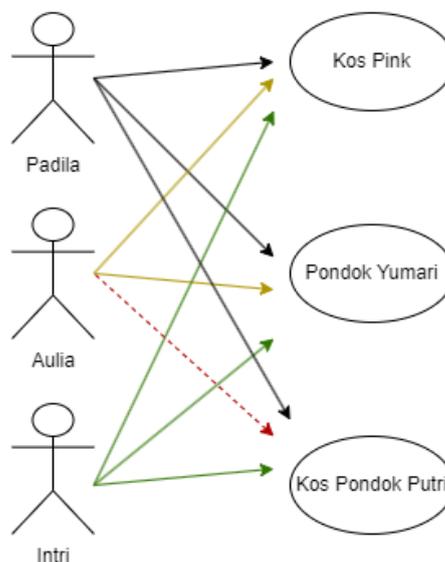
(Sumber : Yanto et al., 2021)

Cara kerja *User-based collaborative filtering* adalah berdasarkan tiap pengguna yang memiliki kesamaan antar pengguna lainnya. Dalam konteks pencarian kost untuk memberikan rekomendasi yang disukai pengguna A maka dari itu harus mencari pengguna lain yang mempunyai kesukaan yang sama dengan pengguna A. Misalkan pengguna 1 menyukai “kost biru”, “pondok yumari”, “kost kuning”, dan “kost pink”, pengguna 2 menyukai “kost kuning”, pengguna 3 menyukai “pondok yumari” dan “kost pink” diimana itu mempunyai kesukaan yang sama terhadap pengguna 1. Kemudian direkomendasikan kost dari pengguna 1 yaitu kost kuning dan kost pink kepada pengguna 3 (Yanto, Herwindiati, and Lauro 2021).

Kelebihan *User – based filtering* adalah mudah di implementasikan dan dipahami. Efektif ketika ada banyak pengguna dengan preferensi yang bisa di bandingkan. Sedangkan kekurangannya adalah skala buruk jika jumlah pengguna sangat besar (perhitungan kemiripan menjadi sangat mahal). Ketergantungan pada jumlah rating yang diberikan pengguna (*cold start*).

b) *Item – based collaborative filtering*

*Item-based collaborative filtering* merupakan teknik yang tidak mencocokkan kemiripan antar pengguna, tetapi melakukan pencocokan setiap item yang dinilai/rating pengguna dengan item serupa, lalu menggabungkan item serupa tersebut ke dalam daftar rekomendasi *based* (Laksito, Arif. (2021).Blog.ariflaksito. *Item- based collaborative filtering* memiliki pola mirip dengan *user-based* jika dalam *user-based* kita mencari hubungan atau kolerasi antar pengguna, maka dalam *item-based collaborative filtering*, kita mencari kolerasi antar item yang disukai atau diminati oleh pengguna. Selanjutnya, item yang berkolerasi tersebut akan direkomendasikan kepada pengguna lain (Herny Februariyanti, Aryo Dwi Laksono, Jati Sasongko Wibowo 2021) . Tujuan dari *item-based* adalah untuk meningkatkan akurasi rekomendasi dengan mengidentifikasi dan menyarankan item yang serupa dengan item yang sudah disukai pengguna, sehingga pengguna dalam menemukan item baru yang relevan dan lebih baik. Konsep *item-based* dapat dilihat pada gambar 2.4 berikut.



**Gambar 2. 4 Konsep *Item-based collaborative filtering***

(Sumber : Yanto et al., 2021)

*Item-based collaborative filtering* bekerja dengan mencari pola rating yang diberikan terhadap sebuah item kemudian memprediksi rating kemiripan yang akan diberikan kepada pengguna terhadap sebuah item. Ilustrasi sederhananya jika aulia menyukai “kost pink” dan “pondok yumari” , namun dia belum menempati “kost pondok putri”. Maka aulia melihat bahwa padila dan intri sudah memberikan peringkat yang sama pada “kost pondok putri” dengan dua kost sebelumnya, maka aulia menyimpulkan bahwa dia juga akan menempati “kost pondok putri”.

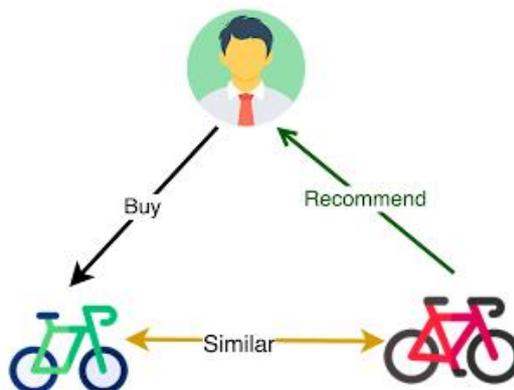
Keuntungan dari *item-based filtering* adalah lebih efisien dalam skala besar karena menghitung kemiripan antara item daripada antar pengguna, serta tidak terlalu bergantung pada jumlah pengguna yang ada (lebih *robust* terhadap *cold start problem*). Namun, kekurangannya adalah membutuhkan banyak data mengenai item yang di rating pengguna, dan bisa jadi kurang efektif jika item –

item tidak memiliki banyak kesamaan atau pengguna memiliki preferensi yang sangat beragam.

Kelebihan dari *Collaborative filtering* adalah dimana pengguna akan merasa surprise mendapatkan rekomendasi produk – produk yang tidak mereka bayangkan sebelumnya. Tetapi di metode *Collaborative filtering* sistem akan menghasilkan rekomendasi yang baik berdasarkan data rating dari pengguna, kondisi jaranganya data – data tersebut menyebabkan sistem rekomendasi berjalan tidak maksimal, atau istilahnya disebut sparse data. Kelemahannya yaitu metode ini tidak dapat memberikan rekomendasi apabila belum ada pemberian nilai pada suatu item. Masalah keterbatasan pada item baru yang belum mendapatkan rating, sehingga ada kemungkinan item tersebut tidak dapat direkomendasikan. Dalam penelitian ini jenis *collaborative filtering* yang digunakan adalah *item-based filtering*.

#### **D. Content – Based Filtering**

*Content- Based Filtering* adalah sebuah rekomendasi sistem dengan cara menganalisis dari deskripsi berdasarkan *content*/atribut yang saling berkaitan dengan item lain yang dibandingkan (Tito Anggoro, 2023). *Content-based filtering* adalah pemfilteran berbasis konten dimana sistem ini memberikan rekomendasi untuk menebak apa yang disukai pengguna berdasarkan aktivitas pengguna tersebut. *Content-based filtering* menggunakan tag untuk produk dengan kata kunci tertentu, memahami preferensi pengguna, mengambil data berdasarkan kata kunci dari database, dan memberikan rekomendasi berdasarkan kesamaan atribut.



**Gambar 2. 5 Konsep *content-based filtering***  
 (Sumber : (Laksito, Arif. (2021).Blog.ariflaksito.)

Sistem rekomendasi dengan metode ini menggunakan data eksplisit dan implisit dari pengguna, membentuk profil pengguna dari data tersebut, dan menggunakannya untuk memberikan rekomendasi produk. Misalnya jika pengguna mencari lagu “Indonesia raya” pada aplikasi youtube sistem akan menyimpan informasi ini dan memberikan rekomendasi berdasarkan yang dicari yang berkaitan dengan “Indonesia raya”. Teknik ini mengandalkan atribut produk/item sebagai referensi kesamaan dan tidak memperhitungkan informasi dari pengguna.

*Content- Based Filtering* merupakan tipe dari sistem rekomendasi dengan konten atau deskripsi sebagai dasar dari algoritma ini. Metode ini memperoleh konten dalam item kemudian membandingkannya dengan konten item yang sebelumnya dicari selanjutnya akan merekomendasikan item yang memiliki konten atau deskripsi serupa dengan apa yang user cari (Cheryl Ayu Melyani 2022). *Content- Based Filtering* tidak melibatkan pengguna lain dalam menentukan rekomendasi, namun hanya pengguna itu sendiri. Berdasarkan apa

yang dicari user, algoritma ini hanya akan memilih item dengan konten yang mirip untuk di rekomendasikan. Hal ini akan membuat keragaman rekomendasi lebih sedikit tetapi akan berfungsi baik walaupun pengguna tidak memberi penilaian (rating) (Larasati and Februariyanti 2021). *Content- Based Filtering* merupakan sistem rekomendasi yang menggunakan profil presensi pengguna dan deskripsi item untuk menghasilkan sebuah rekomendasi (Triana, Adrianti, and Maharani 2019).

*Content-based filtering* memiliki berbagai jenis yang digunakan dalam berbagai konteks dan aplikasi. Berikut adalah beberapa jenis *content-based filtering* yang umum ( Maulid, Reyvan. (2023). DQLab

1. *Content-based filtering* berbasis teks

Menganalisis content teks seperti ulasan produk atau pesan sosial media untuk menentukan sentiment positif atau negatif terhadap suatu topik atau entitas. Mengkategorikan dokumen atau teks berdasarkan topik atau jenis *content*, seperti kategori berita atau klasifikasi spam email.

2. *Content-based filtering* berbasis gambar

Mengidentifikasi objek, wajah, atau katakarakteristik visual dalam gambar. Mengklasifikasikan gambar ke dalam kategori yang sesuai berdasarkan atribut visual seperti warna, tekstur, atau bentuk.

3. *Content-based filtering* berbasis audio

Mengidentifikasi ucapan atau suara dalam file audio. Menilai atau mengategorikan lagu berdasarkan atribut musik seperti tempo, genre, atau lirik.

Namun, sistem rekomendasi berbasis *content-based filtering* memiliki kelebihan dan kekurangan. Kelebihannya adalah sistem melakukan perhitungan berdasarkan kemiripan antara profil pengguna dan profil item. Pada sistem ini, pengguna tidak memerlukan interaksi terhadap pengguna lain sehingga tidak mengalami kekosongan data. Kekurangannya adalah hanya dapat digunakan pada atribut yang sesuai dengan item. Terbatasnya rekomendasi yang mirip saja, sehingga tidak ada kesempatan untuk mendapatkan rekomendasi item yang tidak terduga. (Handan Siswaningrum 2022).

### **E. Rekomendasi *Hybrid***

Rekomendasi *Hybrid* merupakan salah satu metode dari sistem rekomendasi dengan menggabungkan dua atau lebih teknik rekomendasi yang lebih baik. Ide utama metode ini adalah untuk menutupi kekurangan dari metode Tunggal (Handan Siswaningrum 2022). Masing-masing teknik sistem rekomendasi memiliki kelebihan dan kelemahannya tersendiri. Karenanya sistem rekomendasi hybrid yang menggabungkan kekuatan dari model dan target penelitian yang sedang marak dikembangkan (Diah, Rahmad, and Saleh 2019). Dalam penelitian ini, pendekatan *Hybrid CF-CBF (Collaborative Filtering dan Content-Based Filtering)* menunjukkan potensi besar dalam meningkatkan kualitas rekomendasi. Dengan menggabungkan CF, yang berfokus pada pola preferensi pengguna, dan CBF, yang memeriksa kesamaan fitur item, pendekatan ini menghasilkan rekomendasi yang lebih beragam dan personal. Selain itu, pendekatan *Hybrid* ini dapat mengatasi kelemahan masing-masing metode. Misalnya, CF dapat menghadapi masalah *cold-start* ketika data pengguna terbatas,

sementara CBF dapat kurang beragam. Meskipun demikian, pendekatan hybrid ini juga memiliki tantangan, seperti biaya komputasi yang lebih tinggi dan kebutuhan untuk menggunakan metrik evaluasi yang tepat. Namun, penelitian ini menunjukkan bahwa menggabungkan kedua pendekatan ini dapat memperbaiki keakuratan dan personalisasi sistem rekomendasi secara keseluruhan (Widayanti et al. 2023).

Adapun Teknik yang digunakan dalam rekomendasi *Hybrid* adalah sebagai berikut (Handan Siswaningrum , 2022):

1. *Weigth*

Teknik ini melibatkan memberikan bobot atau nilai tertentu pada rekomendasi dari masing-masing metode atau rekomendasi yang terbaik.

2. *Switching*

Teknik ini memilih satu metode rekomendasi secara eksklusif pada suatu waktu.

3. *Mixed*

Pendekatan ini menggabungkan hasil rekomeni dari beberapa metode secara langsung tanpa memberikan bobot pada masing-masing hasil.

4. *Feature combination*

Melibatkan penggabungan fitur atau atribut yang di ambil dari berbagai sumber.

5. *Feature augmentation*

Menambahkan fitur baru atau informasi tambahan ke dalam satu metode rekomendasi berdasarkan informasi dari metode lainnya untuk meningkatkan keakuratan rekomendasi.

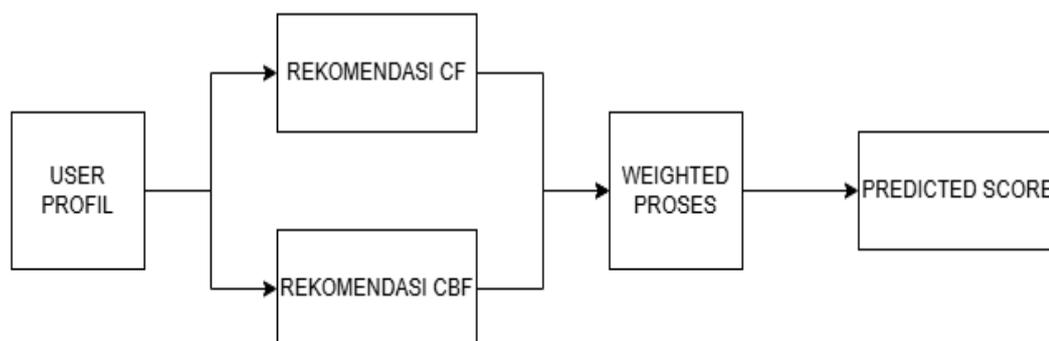
## 6. Cascade

Pendekatan yang mengurutkan metode rekomendasi dimulai dari yang paling sederhana ke yang lebih kompleks,

## 7. Meta level

Teknik ini dari berbagai metode rekomendasi pada satu tingkat yang lebih tinggi mungkin dengan menggunakan model pembelajaran mesin lain untuk menggabungkan atau menilai hasil dari berbagai metode.

Pada penelitian ini jenis *hybrid* rekomendasi yang digunakan adalah *weight*. Hasil dari teknik *hybrid* rekomendasi adalah menggabungkan hasil dari pendekatan *Collaborative Filtering* dan *Content-based Filtering* dengan menggunakan pembobotan tetap untuk masing-masing metode. Bobot ini menentukan kontribusi relative dari setiap pendekatan dalam hasil akhir (Yang, Yang, and Hu 2016).



**Gambar 2. 6 Alur *weighted hybrid* rekomendasi**

(Sumber : Jeffery Chiang)

a) *User* profil

Data *user* profile digunakan pada CF adalah hasil prediksi rating dari tiap *user* berdasarkan hasil kemiripan item atau kost. Data *user* profil untuk CBF adalah membuat *user* profil atribut kost yang diinginkan.

b) Rekomendasi CF (*Collaborative Filtering*)

Hasil rekomendasi dari CF adalah hasil prediksi rating oleh *user* berdasarkan kemiripan item. Hasil prediksi rating diurutkan nilai rating diatas 3 akan menjadi rekomendasi untuk *user*.

c) Rekomendasi CBF (*Content-based Filtering*)

Hasil rekomendasi CBF adalah dari membuat *user* profil kemudian akan dicari *similarity* dengan item yang tersedia berdasarkan atribut yang dibangun dari *user* profil. Selanjutnya hasil rekomendasi berupa nilai kemiripan yang paling mirip akan menjadi rekomendasi untuk *user*.

d) *Weighted* proses

Pada tahap ini prediksi rating dari CF dan skor kemiripan CBF diolah dalam proses penggabungan bobot. Berikut persamaan yang digunakan.

$$rec_w(u, i) = \sum_{k=1}^n \beta_k rec_k(u, i) \quad (2.1)$$

Keterangan :

$rec_w(u, i)$  : Skor rekomendasi gabungan untuk pengguna  $u$  dan item  $i$

$\beta_k$  : Bobot untuk metode rekomendasi ke- $k$

$rec_k(u, i)$  : Skor rekomendasi yang dihasilkan oleh metode ke- $k$  untuk pengguna  $u$  dan item  $i$

e) *Predicted Score*

Hasil akhir adalah *predicted score* yang digunakan untuk menentukan urutan rekomendasi item berprediksi tertinggi menjadi prioritas utama dalam rekomendasi. Hasil akhir ini di urutkan dari nilai terbesar ke terkecil kemudian diurutkan top 10 yang akan di rekomendasikan kepada pengguna.

**F. *Term Frequency-Inverse Document Frequency (TF-IDF)***

TF-IDF digunakan untuk menilai pentingnya suatu kata dalam sebuah dokumen. Metode ini menghitung bobot setiap kata yang sering digunakan dalam information retrieval. TF-IDF dikenal esiseis, mudah, dan menghasilkan hasil yang akurat. TF-IDF merupakan salah satu dari algoritma yang cukup populer dalam melakukan klasifikasi teks dan dapat digunakan untuk menganalisis hubungan antar sebuah kalimat pada sekumpulan dokumen (Tito Anggoro,2023). Algoritma TF-IDF dimanfaatkan sebagai pembentuk profil pengguna untuk mencari kemiripan pada kata kunci. Dalam metode *Content-based filtering*, TF-IDF banyak digunakan untuk menghitung berapa kali sebuah kata muncul dalam sebuah dokumen. Apabila jumlah kemunculan kata tersebut tinggi, dapat ditarik Kesimpulan bahwa kata tersebut penting (Dewa ayu putri diah pramesti 2022). Metode ini menggabungkan dua konsep, yaitu frekuensi kemunculan kata dalam dokumen(TF) dan inverse frekuensi dokumen yang mengandung kata tersebut (IDF).

*Term Frekuensi (TF)* yaitu menghitung seberapa sering suatu kata muncul dalam suatu dokumen. Semakin tinggi frekuensi kemunculan term pada sebuah

dokumen makan akan semakin tinggi juga nilai bobot untuk term itu sendiri (Apriani, Hizbu Zakiyudin 2021). Berikut adalah persamaan yang digunakan dalam menghitung TF-IDF (Rio Trianto Bire 2024).

$$IDF = \log\left(\frac{D}{DF}\right) \quad (2.2)$$

$$W = TF * (IDF + 1) \quad (2.3)$$

Keterangan :

- W : Bobot dokument. Mengukur pentingnya kata dalam dokumen
- TF : Frekuensi kata dalam dokumen
- D : Total jumlah dokumen dalam kospus
- DF : Jumlah dokumen yang mengandung kata tersebut
- IDF : Mengukur seberapa umum atau jarang kata fi seluruh dokumen

### **G. Pearson Colleration**

*Pearson Colleration* adalah salah satu alat ukur yang digunakan untuk menilai kekuatan dan arah hubungan liner antara dua variable. Dua variable dianggap berkolerasi jika perubahan pada salah satu variable diikuti oleh perubahan pada variable lainnya, baik dalam arah yang sama maupun berlawanan. Perlu diingat bahwa nilai koefisien korelasi yang kecil (tidak signifikan) tidak berarti kedua variable tersebut tidak berhubungan. Mungkin saja dua variable memiliki hubungan yang erat nilai kolerasinya mendekati nol. (Setiawan, Ade.(2023).Smartstat.info. Variabel ini terdiri dari variabel dependen dan independent. Kekuatan hubungan antara keduanya verkisar antara 0-1. Jika mendekati angka 1, berarti hubungan kedua variable sangat kuat. Sebaliknya, jika mendekati angka 0, hubungan kedua variable semakin lemah. Sebaliknya jika

mendekati 0, hubungan kedua variable semakin lemah. *Pearson colleration* digunakan untuk menentukan ada atau tidaknya hubungan antara dua variable. Kolerasi ini dapat menghasilkan nilai positif atau negatif. Jika hasilnya positif, maka hubungan antara kedua variable bersifat sejajar. Jika hasilnya negatif,, maka hubungan antara kedua variable tidak sejajar. Sejajar berarti jika nilai variabel independent lebih tinggi, nilai variable dependen juga tinggi. Tidak sejajar berarti jika nilai variabel independek tinggi, nilai variabwl dependen rendah (Yanto, Herwindiati, and Lauro 2021).

Dalam penelitian ini, *Pearson Correlation* digunakan dalam *collaborative filtering*. Jika menggunakan pendekatan *User-based*, *Pearson Correlation* akan menghitung kemiripan antar pengguna berdasarkan rating yang mereka berikan. Sedangkan jika menggunakan pendekatan *Utem-based*, *Pearson Correlation* akan menghitung kemiripan antar item berdasarkan rating yang diberikan oleh pengguna. Dalam konteks data kost, pendekatan *User-based* akan menghitung kemiripan antara pengguna dalam memberikan rating pada berbagai kost. Sementara itu, pendekatan *item-based* akan menghitung kemiripan antar kost berdasarkan rating yang diberikan oleh pengguna. Berikut adalah rumus dari *Pearson Correlation* (Yanto, Herwindiati, and Lauro 2021).

$$sim(k, l) = \frac{\sum_u^m (R_{u,k} - \bar{R}_k)(R_{u,l} - \bar{R}_l)}{\sqrt{\sum_u^m (R_{u,k} - \bar{R}_k)^2} \sqrt{\sum_u^m (R_{u,l} - \bar{R}_l)^2}} \quad (2.4)$$

Keterangan :

$sim(k, l)$  : Nilai *similarity* atau kemiripan kosinus antara item atau objek k dan l

$R_{u,k}$  dan  $R_{u,l}$  : Rating yang diberikaan oleh pengguna u terhadap item k dan l secara berturut – turut

$\bar{R}_k$  dan  $\bar{R}_l$  : Rata – rata rating atau untuk k dan l secara berturut – turut, dihitung semua pengguna u

$m$  : Jumlah pengguna atau observasi yang digunakan dalam perhitungan

## H. Prediksi Rating

Prediksi rating adalah menghitung nilai bobot prediksi yaitu dengan cara membandingkan rating-item yang pernah diberikan pengguna dengan hasil kemiripan antara item tersebut dengan item lainnya (Handan Siswaningrum 2022). Dalam melakukan evaluasi terhadap prediksi rating yang dilakukan menggunakan *Mean Squared Error* (MSE) dan *Root Mean Squared Error* (RMSE). Berikut adalah persamaan yang digunakan dalam prediksi rating (Tommy et al. 2020).

$$P(u, j) = \frac{\sum_{i \in j} (R_{u,i} \times S_{i,j})}{\sum_{i \in j} |S_{i,j}|} \quad (2.5)$$

Keterangan :

$P(u, j)$  : Prediksi untuk responden u pada kost j

$\sum i \in j$  : Himpunan kost yang mirip dengan kost j

$R_{u,i}$  : Rating responden u pada kost i

$R_{u,j}$  : Nilai kemiripan kost i dan kost u

## I. Cosine Similarity

*Cosine similarity* merupakan sebuah metode yang digunakan dalam menghitung tingkat kemiripan (*Similarity*) antara dua dokumen. Sehingga *Cosine*

*similarity* dapat diterapkan pada dua berkas dalam sebuah dokumen teks, dimana terdapat parameter yang digunakan berupa jumlah kata – kata di kedua dokumen untuk di bandingkan. *Cosine similarity* adalah metode untuk mengukur kemiripan antara dua vektor n-dimensi yang sering digunakan dalam pencarian informasi untuk membandingkan dua teks atau dokumen. Kata kunci dalam sebuah dokumen digunakan sebagai ukuran untuk menghitung kemiripan antar dokumen. *Cosine similarity* adalah metode populer untuk mengukur kemiripan dokumen, dengan cara menghitung nilai kosinus dari sudut antara dua vektor. Nilai *cosine similarity* berkisar antara 0 hingga 1. Ketika dua dokumen identik, sudutnya adalah 0 derajat, sehingga nilai kesamaannya adalah 1. Sebaliknya, jika dua dokumen tidak identik, sudutnya adalah 90 derajat dan nilai kesamaannya adalah 0. Semakin mirip dua teks atau dokumen, nilai *cosine similarity* akan mendekati 1, sedangkan jika nilai *cosine similarity* mendekati 0, dua teks atau dokumen tersebut semakin tidak mirip.

Fungsi *similarity* yang paling banyak digunakan adalah *adjust cosine similarity*, *cosine similarity*, dan *pearson coefficient* (Dewa ayu putri diah pramesti 2022). Pada penelitian ini akan menggunakan *cosine similarity* dimana akan digunakan pada *Content-based filtering* untuk menghitung kemiripan antara kost satu dengan lainnya . *Cosine Similarity* dapat dihitung dengan persamaan berikut.

$$similarity(A, B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n A_i^2} \times \sqrt{\sum_{i=1}^n B_i^2}} \quad (2.6)$$

Keterangan :

A : Vektor yang mewakili item A

- $B$  : Vektor yang mewakili item B  
 $A_i$  : Nilai atribut atau fitur ke- $i$  dari item A  
 $B_i$  : Nilai atribut atau fitur ke- $i$  dari item B  
 $n$  : jumlah total atribut atau fitur

#### J. *Mean Squared Error (MSE) dan Root Mean Squared Error (RMSE)*

*Mean Squared Error (MSE)* adalah metrik yang digunakan untuk mengukur kualitas suatu estimator, yaitu rata-rata dari kuadrat selisih antara nilai estimasi dan nilai sebenarnya. MSE yang rendah menandakan bahwa model memiliki tingkat kesalahan prediksi yang rendah. Ini menunjukkan bahwa model memiliki kinerja yang baik dalam memprediksi hasil. (Nugroho et al. 2024). MSE menilai kesalahan prediksi dengan memberikan bobot yang lebih besar, karena perhitungan berdasarkan kuadrat perbedaan.

*Root Mean Squared Error (RMSE)* adalah metrik evaluasi yang digunakan untuk menghitung rata-rata kesalahan kuadrat antara nilai yang di observasi atau nilai sebenarnya dan nilai yang di prediksi oleh model (Nugroho et al. 2024). RMSE sering digunakan dalam evaluasi kinerja model prediksi karena memberikan Gambaran yang intuitif mengenai besarnya kesalahan dalam hasil prediksi. RMSE yang rendah atau mendekati 0 menunjukkan bahwa hasil prediksi sangat mendekati nilai *actual*. (Amansyah et al. 2024). Berdasarkan penjelasan tersebut, MSE dan RMSE digunakan dalam penelitian ini untuk mengukur performa model *Collaborative Filtering* dalam memprediksi rating kost. Berikut adalah persamaan MSE dan RMSE.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.8)$$

$y_i$  : Nilai sebenarnya

$\hat{y}_i$  : Nilai yang di prediksi

n : Jumlah total prediksi

### K. Python

*Python* adalah bahasa pemrograman komputer yang biasa dipakai untuk membangun situs, *software* mengotomatiskan dan melakukan analisis data. Bahasa ini termasuk bahasa yang digunakan dalam tujuan umum. Dalam penelitian ini *python* digunakan untuk menganalisis data untuk *collaborative filtering* dan *content-based filtering*.

### L. Penelitian Terkait

**Tabel 2. 1 Penelitian Terkait**

No	Peneliti dan judul	Hasil Penelitian	Perbedaan	Persamaan
1.	Yanto, Dyah Erny Herwindiati, Manatap Dolok Lauro (2021) Marketplace Chinese Food dengan Sistem	Hasil penelitian menunjukkan pengguna menadapatkan rekomendasi makanan dan restoran yang cocok melalui metode <i>item-based</i> . Pada	Data makanan <i>Chinese food</i> . Tidak melakukan <i>Hybrid Filtering</i>	Persamaan dengan penelitian yang akan dilakukan adalah menggunakan metode <i>Collaborative Filtering Item-based</i> . Menggunakan <i>Pearson Colleration</i> .

	Rekomendasi Menggunakan Item-based Collaborative Filtering	pengujian data menggunakan metode <i>item-based Collaborative Filtering</i> terdapat hasil yang berbeda dengan pengguna lain karena bedanya pemberian rating terhadap suatu item		Menggunakan <i>Weight Sum</i>
2.	Erlangga, Mar Hidayati Anggraini, Fenty Ariani, Yuthsi Aprilinda (2020) Aplikasi E-Marketing Panglong Kayu Menggunakan Metode Collaborative FILTERING	<i>Collaborative Filtering item-based.</i> Hasil penelitian sistem e-marketing panglong kayu dan memudahkan konsumen untuk memesan tanoa meninggalkan rumah. Dari hasil pengujian perangkat lunak menunjukkan bahwa penerapan e-marketing panglong kayu memenuhi kriteria kelayakan. Hasil uji kelayakan sistem dari aspke fungsional nilai 93%	Data yang digunakan adalah produk kayu. Membuat <i>interface</i> yang memudahkan pengguna mencari produk berdasarkan rating.	Menggunakan <i>Collaborative Filtering-item-based.</i> Menggunakan <i>Pearson Colleration.</i> Menggunakan <i>Weight Sum</i>

3.	<p>Irannur Diah, Iwan Fitrianto Rahmad, dan Alfa Saleh pada tahun 2019 melakukan penelitian dengan judul “Implementasi <i>Recommender</i> Pada Pemilihan Kamera Menggunakan <i>Contet Based Filtering</i> dan <i>Collaborative Filtering</i>”.</p>	<p>Hasil penelitian adalah dengan penggabungan secara <i>mixed hybrid</i> dihasilkan sebuah sistem rekomendasi pemilihan kamera yang mampu menutupi kekurangan masing-masing metode. Pada <i>collaborative filtering</i> membutuhkan pengguna untuk memberikan opini / rating terhadap beberapa kamera terlebih dahulu (untuk membangun <i>user profile</i>) sebelum sistem dapat memberikan rekomendasi, sehingga rekomendasi tidak dapat diberikan terhadap pengguna yang baru saja terdaftar dan belum memiliki aktivitas</p>	<p>Data yang digunakan data kamera. Teknik <i>Hybrid</i> yang digunakan adalah <i>Mixed Hybrid</i></p>	<p>Metode <i>Hybrid</i> dalam penggabungan <i>Collaborative Filtering</i> dan <i>Content-Based Filtering</i>.</p>
----	--	--	--	---

		apapun di dalam sistem		
4.	Lukas Tommy, Dian Novianto, dan Yohanes Setiawan Japriadi melakukan penelitian pada tahun 2020 dengan judul penelitian “ Sistem Rekomendasi Hybrid untuk Pemesanan Hidangan Berdasarkan Karakteristik dan Rating Hidangan”	Sistem rekomendasi <i>Hybrid</i> yang diusulkan berhasil memadukan metode <i>content-based</i> dengan <i>collaborative filtering</i> dimana nilai <i>precision</i> dan <i>recall</i> -nya jika diukur dengan <i>confusion matrix</i> berturut-turut adalah 80,73% dan 76,52%. Selain itu, pemanfaatan karakteristik dari hidangan yang pernah dipesan memungkinkan sistem rekomendasi merekomendasikan hidangan baru atau yang belum pernah dipesan selama karakteristiknya mirip dengan hidangan yang	Data yang digunakan data hidangan makanan. Tidak menggunakan TF-IDF	Menggunakan metode <i>Hybrid Filtering</i> , <i>Collaborative Filtering</i> , <i>Content-based Filtering</i> .

		pernah dipesan pelanggan tersebut.		
5.	Cheryl Ayu Melyani pada tahun 2022 melakukan penelitian dengan judul “Sistem Rekomendasi Hotel Dengan Pendekatan <i>Content-Based Filtering</i> ”.	Penelitian ini menggunakan metode <i>Cosine Similarity</i> untuk menghitung kemiripan dokumen, dengan contoh pada dokumen ke-132. Hasilnya, diperoleh 10 dokumen dengan nilai kemiripan tertinggi, yaitu dokumen ke-170, 303, 210, 360, 315, 422, 510, 45, 63, dan 223, dengan nilai <i>Cosine Similarity</i> berturut-turut antara 0.9566 hingga 0.8715. Alur rekomendasi hotel di Yogyakarta melibatkan <i>preprocessing</i> data deskripsi, pembobotan kata menggunakan TF-IDF, perhitungan	Tidak menggunakan <i>Hybrid Filtering</i> . Data yang digunakan adalah data hotel.	Penelitian yang akan dilakukan sama-sama menggunakan <i>Content-based Filtering</i> , <i>Cosine Similarity</i> , TF-IDF

		nilai kemiripan antar dokumen, pengurutan nilai kemiripan dari terbesar, dan pemilihan 10 dokumen dengan nilai tertinggi.		
6.	Handan Siswaningrum melakukan penelitian pada tahun 2022 dengan judul “Sistem Rekomendasi Produk <i>Skin Care</i> Menggunakan Metode <i>Hybrid (Content Based Filtering dan Collaborative Filtering)</i> Berbasis Android.”	Hasil penelitian adalah dengan memanfaatkan metode rekomendasi <i>hybrid</i> dapat memberikan performa rekomendasi yang untuk user yang mampu menutupi kekurangan pada masing-masing metode. Pada metode <i>Content Based Filtering</i> dapat mengatasi masalah <i>cold start problem</i> pada <i>Collaborative Filtering</i> .	Data yang digunakan data produk <i>skin care</i> . Membuat aplikasi berbasis android	<i>Hybrid filtering</i> menggunakan <i>Collaborative Filtering</i> dan <i>Content-based Filtering</i> . Sama-sama menggunakan <i>Pearson Colleration</i> , <i>Weight Sum</i> dan TF-IDF

7.	Muthi Ishlah Rizky, Ibnu Asror, Yusza Reditya Mutri (2020) Sistem Rekomendasi Program Studi untuk Siswa SMA Sederajat Menggunakan Metode <i>Hybrid Recommendation</i> dengan metode <i>Content-based Filtering</i> dan <i>Collaborative filtering</i> mampu diterapkan dalam membangun sistem rekomendasi pemilihan program studi untuk menyeleksi beragam program studi. Dimana pada uji sistem pada <i>naive bayes</i> ditemukan bahwa akurasi yang diberikan <i>Naive Bayes</i> adalah 85.5%, Sedangkan untuk MAE didapatkan sebanyak 0.2194%.	Berdasarkan hasil analisis pengujian, dapat disimpulkan bahwa pendekatan <i>Hybrid Recommendation</i> dengan metode <i>Content-based Filtering</i> dan <i>Collaborative filtering</i> mampu diterapkan dalam membangun sistem rekomendasi pemilihan program studi untuk menyeleksi beragam program studi. Dimana pada uji sistem pada <i>naive bayes</i> ditemukan bahwa akurasi yang diberikan <i>Naive Bayes</i> adalah 85.5%, Sedangkan untuk MAE didapatkan sebanyak 0.2194%.	Menggunakan data program studi, Menggunakan <i>Naive Bayes</i> .	Menggunakan <i>Hybrid</i> rekomendasi, <i>Item-based Collaborative Filtering</i> , <i>Pearson Colleration</i> , <i>Weight Sum</i>
8.	Ardimansyah, Husain,	Berdasarkan hasil penelitian yang	Menggunakan data produk pestisida.	Menggunakan <i>Hybrid Filtering</i>

	<p>Herlinda, Kasmawaru, Nurdiansah, Marsa (2024) Implementasi Metode <i>Hybrid Filtering Technique</i> Pada Penentuan <i>Rating</i> Pestisida</p>	<p>kami lakukan dapat disimpulkan bahwa Kolaborasi Metode <i>Content Based Filtering</i>, <i>Collaborative Filtering</i>, dan <i>Hybrid Filtering Technique</i> pada Menentukan <i>Rating</i> Pestisida dapat dijadikan sebagai rujukan untuk mendapatkan rekomendasi mengenai penentuan pestisida yang cocok digunakan membasmi hama tanaman padi. Dengan menggunakan metode ini dapat menentukan <i>rating</i> pestisida yang paling banyak digunakan petani pada jenis penyakit bercak coklat, penyakit blast,</p>	<p>Tidak menggunakan TF-IDF dan <i>Cosine Similarity</i></p>	<p><i>Collaborative Filtering</i> dan <i>Content Based Filtering</i>. Menggunakan <i>Pearson Correlation</i> dan <i>Weight Sum</i>. Teknik <i>Weight Hybrid</i></p>
--	---	---	--	---

		<p>penyakit busuk batang, penyakit gosong palsu, penyakit hawar daun, dan penyakit tungro. Hasil pengujian sistem diperoleh jenis pestisida yang memiliki nilai rating tertinggi adalah pestisida enquity dengan nilai rating 2,256. Saran untuk penelitian selanjut untuk melakukan kolaborasi metode <i>Hybrid</i> filtering dengan metode K-mens pada penentuan ratin</p>		
9.	<p>Yohana Imelda Lubis, Dohar Josua Napitupulu, Arie Satia Dharma (2020) Impelementasi Metode <i>Hybrid</i></p>	<p>Metode <i>Hybrid filtering</i>, kombinasi <i>Content-based filtering</i> dan <i>Collaborative filtering</i>, diterapkan pada sistem rekomendasi pariwisata untuk mengatasi <i>sparsity</i></p>	<p>Rekomendasi parawisata. Menggunakan data parawisata. Menggunakan MAE dalam pengujian prediksi rating</p>	<p>Persamaan menggunakan <i>Hybrid Filtering</i> gabungan <i>Collaborative</i> dan <i>Content Based Filtering</i>. Menggunakan TF-</p>

	<p><i>Filtering</i> (Collaborative dan Content Based ) untuk Sistem Rekomendasi Pariwisata</p>	<p>data dengan mengisi rating kosong menggunakan Content-based filtering. Selanjutnya, similarity antar item dihitung menggunakan Pearson correlation, memilih k item teratas, dan memprediksi item yang direkomendasikan. Penelitian ini menunjukkan bahwa hybrid filtering lebih akurat, dengan MAE terendah 0,3741 (k 25%), dibandingkan content-based filtering (MAE 1,1742) dan collaborative filtering (MAE 0,3768). Pengisian data sparse</p>		<p>IDF Pearson Colleration.</p>
--	--	--	--	---------------------------------

		<p>menggunakan <i>Hybrid filtering</i> terbukti lebih efektif dibandingkan <i>collaborative Filtering</i> dalam menghasilkan rekomendasi yang lebih baik.</p>		
10.	<p>Ressa Priskila, M.Fajar, Septian Geges, Widiatry (2024) Penerapan Metode <i>Collaborative Filtering</i> dan <i>Content Based Filtering</i> Pada Sistem Rekomendasi Smartphone Android.</p>	<p>Sistem rekomendasi ini dibangun dengan mengimplementasikan metode <i>Collaborative filtering</i> dan <i>Content based filtering</i>. Metodologi pengembangan perangkat lunak menggunakan GRAPPLE (<i>Guidelines for Rapid APPLication Engineering</i>). Hasil Pengujian menggunakan MAPE, sistem yang dibangun termasuk dalam kategori baik,</p>	<p>Menggunakan data <i>Smartphone</i> Android. Tidak menerapkan <i>Hybrid Filtering</i>. Menggunakan metode GRAPPLE.</p>	<p>Menggunakan kedua metode <i>Collaborative Filtering</i> dan <i>Content Based Filtering</i>.</p>

		<p>sehingga dapat memberikan rekomendasi ke pada konsumen. Temuan dari hasil penelitian ini akan memberikan wawasan tentang bagaimana mengimplementasi metode metode <i>Collbarorative</i> dan <i>Content based</i> untuk sistem rekomendasi. Untuk penelitian selanjutnya dapat menggunakan data yang lebih banyak dan metode yang lain.</p>		
--	--	---	--	--

## DAFTAR PUSTAKA

- Amansyah, Ilham, Jamaludin Indra, Euis Nurlaelasari, and Ayu Ratna Juwita. 2024. "Prediksi Penjualan Kendaraan Menggunakan Regresi Linear : Studi Kasus Pada Industri Otomotif Di Indonesia." 4: 1199–1216. <https://j-innovative.org/index.php/Innovative%0APrediksi>.
- Apriani, Hizbu Zakiyudin, Khairan Marzuki. 2021. "Penerapan Algoritma *Cosine Similarity* Dan Pembobotan TF-IDF System Penerimaan Mahasiswa Baru Pada Kampus Swasta." *Jurnal Bumigora Information Technology (BITE)* 1(Vol.3, No.1, Juni 2021, pp. 19–27): 19–27. <https://doi.org/10.30812/bite.v6i1>.
- Arfisko, Hilmi Hidayat. 2022. "Sistem Rekomendasi Film Menggunakan Metode *Hybrid Collaborative Filtering* Dan *Content-Based Filtering*." *e-Proceeding of Engineering* 9(Vol.9, No.3(2022): Juni 2022): 2149. <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/18066>.
- Arfisko, Hilmi Hidayat, and Agung Toto Wibowo. 2022. "Sistem Rekomendasi Film Menggunakan Metode *Hybrid Collaborative Filtering* Dan *Content-Based Filtering*." *e-Proceeding of Engineering* 9(3).
- Arifin, Ali. 2022. "Penerapan Sistem Algoritma Collaborative Filtering Untuk Rekomendasi Pemilihan Indekos Berdasarkan Rating." *Teknologipintar.org* Volume 2 ((6): 1–11.
- Cheryl Ayu Melyani. 2022. "Sistem Rekomendasi Hotel Dengan Pendekatan *Content-Based Filtering*." <https://dspace.uui.ac.id/bitstream/handle/123456789/39737/18611001.pdf?sequence=1&isAllowed=y>.
- Dewa ayu putri diah pramesti, wayan santiyasa. 2022. "Penerapan Metode *Content-Based Filtering* Dalam Sistem Rekomendasi Video Game." *Jurnal Nasional Teknologi iformasi dan Aplikasinya* 1(Volume 1, Nomor 1, November 2022). <https://ojs.unud.ac.id/>.
- Diah, Irfannur, Iwan Fitrianto Rahmad, and Alfa Saleh. 2019. "IMPLEMENTASI RECOMMENDER SYSTEM PADA PEMILIHAN KAMERA MENGGUNAKAN *CONTENT BASED* DAN *COLLABORATIVE FILTERING*." *IT (INFORMATIC TECHNIQUE) JOURNAL* 7(1).
- Erlangga, Erlangga, Mar Hidayati Anggraini, Fenty Ariani, and Yuthsi Aprilinda. 2020. "Aplikasi E-Marketing Panglong Kayu Menggunakan Metode *Colaborative Filtering*." *Explore: Jurnal Sistem Informasi dan Telematika* 11(1): 57.
- Herny Februariyanti, Aryo Dwi Laksono, Jati Sasongko Wibowo, Mardi Siswo Utomo. 2021. "IMPLEMENTASI METODE *COLLABORATIVE FILTERING* UNTUK SISTEM REKOMENDASI PENJUALAN PADA TOKO MEBEL." *JURNAL KHATULISTIWA INFORMATIKA IX(VOL. IX, NO. I JUNI 2021):* 43–50. <https://doi.org/10.31294/jki.v9i1.9859>.
- Larasati, Fatoni Batari Agung, and Herny Februariyanti. 2021. "Sistem

- Rekomendasi Product Emina Cosmetics Dengan Menggunakan Metode *Content - Based Filtering*.” *Jurnal Manajemen Informatika dan Sistem Informasi* 4(1): 45.
- Nugroho, Ariyanto Adi et al. 2024. “MENINGKATKAN KUALITAS DATA PADA PERAMALAN CURAH HUJAN.” 7(2): 301–11.
- Selviana, Renita, and Laili Dwi Utari. 2019. “Sistem Informasi Pencarian Kos Kosan Di Sekitar Kecamatan Beji Berbasis Web.” *Jurnal SPIRIT* 11(1): 73–77.
- Syam, Elgamar. 2018. “Rancang Bangun Sistem Informasi Rumah Kost Dan Kontrakan Teluk Kuantan.” *Jurnal Teknologi Dan Open Source* 1(1): 1–7.
- Tommy, Lukas et al. 2020. “Sistem Rekomendasi *Hybrid* Untuk Pemesanan Hidangan Berdasarkan Karakteristik Dan Rating Hidangan.” 4(2): 137–45.
- Triana, Yaya Sudarya, Fryda Farizha Adrianti, and Fitri Anggri Maharani. 2019. “Implementasi Metode *Content Based Filtering* Pada Aplikasi Pencarian Taman Penitipan Anak.” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)* 3(2): 163–69.
- Widayanti, Riya et al. 2023. “*Improving Recommender Systems Using Hybrid Techniques of Collaborative Filtering and Content-Based Filtering*.” *Journal of Applied Data Sciences* 4(3): 289–302.
- Wulandari, Devi, and Adil Setiawan. 2022. “*UNES Journal of Information System IMPLEMENTATION OF COLLABORATIVE FILTERING AND FP-GROWTH ALGORITHMS IN THE INTRODUCTION OF ECTOURISM IN THE AREA OF*.” 7(2): 64–78.
- Yang, Junrui, Cai Yang, and Xiaowei Hu. 2016. “*A Study of Hybrid Recommendation Algorithm Based on User*.” *Proceedings - 2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2016* 2(2): 261–64. <https://doi.org/10.1109/IHMSC.2016.168>.
- Yanto, Yanto, Dyah Erny Herwindiati, and Manatap Dolok Lauro. 2021. “Marketplace *Chinese Food* Dengan Sistem Rekomendasi Menggunakan *Item-Based Collaborative Filtering*.” *Jurnal Ilmu Komputer dan Sistem Informasi* 9(1): 228. <https://doi.org/10.24912/jiksi.v9i1.11599>.