Logo

Description automatically generated

**TUGAS AKHIR – QQXXXXXX**

**SISTEM REKOMENDASI PROPERTI BERBASIS HYBRID FILTERING DAN PROFILE MATCHING**

**I GUSTI MADE ARISUDANA**

NRP 5026211188

Dosen Pembimbing I

**Retno Aulia Vinarti, S.Kom., M.Kom., Ph.D.**

NIP 1988201812010

Dosen Pembimbing II

**Izzat Aulia Akbar, S.Kom., M.Eng., Ph.D.**

NIP 1991202011006

**Program Studi Sistem Informasi**

Departemen Sistem Informasi

Fakultas Teknologi Elektro dan Informatika Cerdas

Institut Teknologi Sepuluh Nopember

Surabaya

2025

Icon

Description automatically generated

**TUGAS AKHIR – QQXXXXXX**

**SISTEM REKOMENDASI PROPERTI BERBASIS HYBRID FILTERING DAN PROFILE MATCHING**

**I GUSTI MADE ARISUDANA**

NRP 5026211188

Dosen Pembimbing I

**Retno Aulia Vinarti, S.Kom., M.Kom., Ph.D.**

NIP 1988201812010

Dosen Pembimbing II

**Izzat Aulia Akbar, S.Kom., M.Eng., Ph.D.**

NIP 1991202011006

**Program Studi Sistem Informasi**

Departemen Sistem Informasi

Fakultas Teknologi Elektro dan Informatika Cerdas

Institut Teknologi Sepuluh Nopember

Surabaya

2025

Icon

Description automatically generated

**FINAL PROJECT – QQXXXXXX**

**PROPERTY RECOMMENDATION SYSTEMS BASED ON HYBRID FILTERING AND PROFILE MATCHING**

**I GUSTI MADE ARISUDANA**

NRP 5026211188

Advisor I

**Retno Aulia Vinarti, S.Kom., M.Kom., Ph.D.**

NIP 1988201812010

Advisor II

**Izzat Aulia Akbar, S.Kom., M.Eng., Ph.D.**

NIP 1991202011006

**Study Program Information Systems**

Department of Information Systems

Faculty of Intelligent Electrical and Informatics Technology

Institut Teknologi Sepuluh Nopember

Surabaya

2025

LEMBAR PENGESAHAN

**SISTEM REKOMENDASI PROPERTI BERBASIS HYBRID FILTERING DAN PROFILE MATCHING**

**TUGAS AKHIR**

Diajukan untuk memenuhi salah satu syarat

memperoleh gelar Sarjana pada

Program Studi S-1 Sistem Informasi

Departemen Sistem Informasi

Fakultas Teknologi Elektro dan Informatika Cerdas

Institut Teknologi Sepuluh Nopember

Oleh : **I Gusti Made Arisudana**

NRP. 5026211188

Disetujui oleh Tim Penguji Tugas Akhir :

|  |  |  |
| --- | --- | --- |
| 1. | Retno Aulia Vinarti, S.Kom., M.Kom., Ph.D. | Pembimbing |
| 2. | Izzat Aulia Akbar, S.Kom., M.Eng., Ph.D. | Ko-pembimbing |
| 3. | Amalia Utamima, S.Kom., MBA., Ph.D. | Penguji |
| 4. | Renny Pradina Kusumawardani, S.T., M.T. | Penguji |

**SURABAYA**

**Juni, 2025**

PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini:

|  |  |  |
| --- | --- | --- |
| Nama mahasiswa / NRP | : | I Gusti Made Arisudana / 5026211188 |
| Program studi | : | Sistem Informasi |
| Dosen Pembimbing / NIP | : | Retno Aulia Vinarti, S.Kom., M.Kom., Ph.D. / 1988201812010 |

dengan ini menyatakan bahwa Tugas Akhir dengan judul “Sistem Rekomendasi Properti Berbasis Hybrid Filtering Dan Profile Matching” adalah hasil karya sendiri, bersifat orisinal, dan ditulis dengan mengikuti kaidah penulisan ilmiah.

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia menerima sanksi sesuai dengan ketentuan yang berlaku di Institut Teknologi Sepuluh Nopember.

Surabaya, \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |
| --- | --- |
| Mengetahui  Dosen Pembimbing | Mahasiswa |
|  |  |
| Retno Aulia Vinarti, S.Kom., M.Kom., Ph.D. | I Gusti Made Arisudana |
| NIP. 1988201812010 | NRP. 5026211188 |

ABSTRAK

**SISTEM REKOMENDASI PROPERTI BERBASIS HYBRID FILTERING DAN PROFILE MATCHING**

|  |  |  |
| --- | --- | --- |
| **Nama Mahasiswa / NRP** | **:** | **I Gusti Made Arisudana / 5026211188** |
| **Departemen** | **:** | **Sistem Informasi FTEIC - ITS** |
| **Dosen Pembimbing** | **:** | **Retno Aulia Vinarti, S.Kom., M.Kom., Ph.D.** |

**Abstrak**

Suspensi merupakan komponen penting pada kendaraan bermotor karena berperan penting dalam menjaga kenyamanan dan keamanan saat berkendara. Sebuah ide baru diperkenalkan yaitu, *Series Active Variable Geometry Suspension* (SAVGS), dimana sistem suspensi ini memiliki performa yang lebih baik dari suspensi pasif dan dapat mengatasi kelemahan dari suspensi aktif. Penelitian terus dilakukan guna meningkatkan performa dari SAVGS. Pada penelitian ini akan dipelajari pengaruh panjang *linkage* (*single link)* terhadap performa kendaraan khususnya kenyamanan dan stabilitas. Model seperempat kendaraan digunakan untuk memodelkan dinamika sistem suspensi kendaraan. Pengaruh panjang *single link* dianalisis dalam bentuk koefisien kekakuan dan koefisien peredam. Model linier digunakan untuk merancang *state-feedback control system* (LQR). Kinerja sistem kendali diuji pada model nonlinier yang dibuat dengan menggunakan Simscape Multibody. Hasil simulasi menunjukkan bahwa semakin panjang *single link* yang digunakan maka kenyamanan dan stabilitas kendaraan semakin besar. Namun, semakin panjang *single link* diperlukan input kontrol yang lebih besar.

**Kata kunci: *LQR, Quarter-car, SAVGS, Simscape Multibody, Suspension****.*

ABSTRACT

**PROPERTY RECOMMENDATION SYSTEMS BASED ON HYBRID FILTERING AND PROFILE MATCHING**

|  |  |  |
| --- | --- | --- |
| **Student Name / NRP** | **:** | **I Gusti Made Arisudana / 5026211188** |
| **Department** | **:** | **Sistem Informasi FTEIC - ITS** |
| **Advisor** | **:** | **Retno Aulia Vinarti, S.Kom., M.Kom., Ph.D.** |

**Abstract**

Suspension is an important component in vehicles because it plays an important role in maintaining comfort and safety while driving. A new idea was introduced, namely, Series Active Variable Geometry Suspension (SAVGS), where this suspension system has better performance than passive suspension and can overcome the weaknesses of active suspension. Research continues to improve the performance of SAVGS. The effect of linkage length (single link) on SAVGS performance, especially comfort and stability, is studied. A quarter car is used to model the dynamics of the vehicle suspension system. The effect of single link length is analyzed in the form of stiffness coefficient and damping coefficient. The linear model is used to design the state-feedback control system (LQR). The performance of the control system was tested on a nonlinear model made using Simscape Multibody. The simulation results show that the longer the single link used, the greater the vehicle's comfort and stability. However, the longer the single link required more considerable control input.

**Keywords: *LQR, Quarter-car, SAVGS, Simscape Multibody, Suspension****.*

DAFTAR ISI

[LEMBAR PENGESAHAN i](#_Toc95924952)

[PERNYATAAN ORISINALITAS ii](#_Toc95924953)

[ABSTRAK iii](#_Toc95924954)

[ABSTRACT iv](#_Toc95924955)

[DAFTAR ISI v](#_Toc95924956)

[DAFTAR GAMBAR vi](#_Toc95924957)

[DAFTAR TABEL vii](#_Toc95924958)

[DAFTAR SIMBOL viii](#_Toc95924959)

[BAB 1 PENDAHULUAN 9](#_Toc95924960)

[1.1 Latar Belakang 9](#_Toc95924961)

[1.2 Rumusan Masalah 9](#_Toc95924962)

[1.3 Tujuan 9](#_Toc95924963)

[1.4 Batasan Masalah 9](#_Toc95924964)

[1.5 Manfaat 9](#_Toc95924965)

[BAB 2 TINJAUAN PUSTAKA 10](#_Toc95924966)

[2.1 Hasil Penelitian Terdahulu 10](#_Toc95924967)

[2.2 Dasar Teori 10](#_Toc95924968)

[BAB 3 METODOLOGI 11](#_Toc95924969)

[3.1 Metode yang digunakan 11](#_Toc95924970)

[3.2 Bahan dan peralatan yang digunakan 11](#_Toc95924971)

[3.3 Urutan pelaksanaan penelitian 11](#_Toc95924972)

[BAB 4 Hasil dan Pembahasan 12](#_Toc95924973)

[4.1 Hasil penelitian 12](#_Toc95924974)

[4.2 Pembahasan 12](#_Toc95924975)

[BAB 5 Kesimpulan dan Saran 13](#_Toc95924976)

[5.1 Kesimpulan 13](#_Toc95924977)

[5.2 Saran 13](#_Toc95924978)

[DAFTAR PUSTAKA 14](#_Toc95924979)

[LAMPIRAN 15](#_Toc95924980)

[BIODATA PENULIS 16](#_Toc95924981)

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR SIMBOL

# PENDAHULUAN

## Latar Belakang

Tempat tinggal merupakan salah satu kebutuhan mendasar bagi setiap individu. Tempat tinggal tidak hanya berfungsi sebagai sarana fisik untuk perlindungan, tetapi juga berperan dalam pemenuhan aspek psikologis dan sosial individu. Mengingat pentingnya peran tempat tinggal dalam kehidupan, keputusan untuk memilih properti menjadi proses yang sangat krusial. Sebagian besar individu jarang melakukan transaksi pembelian atau penyewaan properti sepanjang hidup mereka, sehingga setiap keputusan terkait tempat tinggal membutuhkan pertimbangan yang matang dan menyeluruh. Kompleksitas ini tidak hanya muncul dari tingginya nilai ekonomi properti, tetapi juga karena banyaknya faktor yang perlu diperhatikan, seperti lokasi, harga, ukuran, fasilitas, serta preferensi pribadi (Gharahighehi, Pliakos, & Vens, 2021). Oleh karena itu, proses pemilihan tempat tinggal sering kali melibatkan pencarian informasi yang intensif, baik secara konvensional maupun melalui platform digital.

Di era digital saat ini, penggunaan platform daring untuk mencari properti mengalami peningkatan signifikan. Konsumen semakin beralih ke situs web dan aplikasi untuk mengeksplorasi pilihan properti yang tersedia. Namun, meskipun akses informasi menjadi lebih mudah, pengguna sering menghadapi tantangan baru dalam bentuk "*information overload*" atau kelebihan informasi. Banyaknya pilihan properti yang tersedia justru dapat membuat proses pencarian menjadi melelahkan dan membingungkan. Pengguna perlu menyaring berbagai opsi yang tidak relevan sebelum menemukan properti yang sesuai dengan preferensi mereka. Hal ini menyebabkan pengalaman pencarian yang tidak optimal dan berpotensi menurunkan kepuasan pengguna (Yu, Wang, Zhang, Gao, & Wang, 2018). Untuk mengatasi permasalahan ini, sistem rekomendasi (SR) menjadi salah satu solusi yang efektif. SR berfungsi untuk menyederhanakan proses pencarian dengan mengidentifikasi properti yang paling relevan bagi pengguna berdasarkan data yang diperoleh dari interaksi mereka dengan platform.

Berbagai pendekatan telah dikembangkan untuk meningkatkan performa sistem rekomendasi properti. Salah satu penelitian yang signifikan adalah karya Han Jong Jun et al., yang menghasilkan SR berbasis embedding bernama “SeoulHouse2Vec”. SR ini menggunakan pendekatan *Neural Network Collaborative Model* yang memanfaatkan teknik embedding untuk merepresentasikan hubungan antara pengguna dan properti secara lebih baik (Jun, Kim, Rhee, & Chang, 2020). Penelitian lain yang dilakukan oleh Zhang et al. mengusulkan pendekatan *Content-Based Filterin*g yang menggunakan model dua tahap, di mana sistem merekomendasikan properti berdasarkan karakteristik item serta preferensi historis pengguna (Zhang, et al., 2019). Terdapat penelitian lain tentang SR properti yang menggunakan *Content-Based Filtering* dengan mengimplementasikan metode *Term Frequency Inverse Document Frequency* (TF-IDF) untuk memberikan bobot pada judul, deskripsi, dan alamat sebuah iklan properti yang dikunjungi pengguna. Kemudian algoritma Apriori digunakan untuk memeberikan rekomendasi properti yang mirip dengan yang pengguna lihat (Badriyah et al., 2018). Meskipun pendekatan ini menunjukkan hasil yang menjanjikan, masih terdapat beberapa tantangan yang perlu diatasi dalam pengembangan SR yang optimal, terutama dalam menangani berbagai keterbatasan data pengguna.

Salah satu permasalahan utama yang dihadapi oleh SR dalam domain properti adalah fenomena *cold-start problem*, di mana sistem kesulitan memberikan rekomendasi yang akurat bagi pengguna baru atau bagi pengguna yang memiliki sedikit riwayat interaksi dengan platform. Sebagian besar pendekatan saat ini, seperti *Content-Based Filtering*, hanya berfokus pada karakteristik item atau aktivitas historis pengguna, seperti klik atau preferensi visual. Namun, model ini memiliki keterbatasan, karena perilaku pengguna yang terekam tidak selalu merepresentasikan preferensi nyata mereka. Pengguna mungkin melakukan klik secara acak atau terpengaruh oleh faktor eksternal, sehingga hasil rekomendasi yang diberikan bisa saja tidak sesuai dengan profil atau kebutuhan sebenarnya (Knoll, Groß, Schwanke, Rinn, & Schreyer, 2018). Oleh karena itu, diperlukan pendekatan yang lebih personal dan komprehensif dalam menangkap preferensi pengguna.

Untuk mengatasi tantangan tersebut, tugas akhir ini akan mengembangkan sistem rekomendasi berbasis *Hybrid Filtering* yang menggabungkan *Content-Based Filtering* dan *Knowledge-Based Recommender Systems* dengan menggunakan metode *Profile Matching* yang berfokus pada karakteristik demografi dan kondisi sosial pengguna. Pendekatan ini tidak hanya akan meningkatkan kualitas rekomendasi, tetapi juga dapat mengurangi risiko *cold-start* problem yang sering terjadi pada SR konvensional. Tugas akhir ini diharapkan dapat memberikan kontribusi signifikan dalam pengembangan sistem rekomendasi properti yang lebih efektif dan efisien di era digitalisasi ini.

## Rumusan Masalah

Berdasarkan latar belakang permasalahan di atas, maka rumusan permasalahan dalam tugas akhir ini adalah sebagai berikut:

1. Apa saja data demografi dan kondisi sosial yang relevan untuk membentuk profil pembeli?
2. Bagaimana algoritma *Profile Matching* dapat digunakan untuk mencocokkan properti dengan profil pembeli?
3. Bagimana teknik *Hybrid Filtering* dapat diterapkan untuk menentukan properti yang mirip berdasarkan karakteristik item?
4. Bagaimana cara mengukur kepuasan pengguna terhadap rekomendasi yang diberikan oleh sistem?

## Batasan Masalah

Berdasarkan rumusan masalah yang telah dibuat, maka dapat didefinisikan batasan-batasan dalam tugas akhir ini sebagai berikut:

1. Tugas akhir ini akan fokus pada pengembangan sistem rekomendasi properti berbasis *Hybrid Filtering* yang mengintegrasikan Content Based Filtering dan *Knowledge Based Recommender Systems* dengan metode *Profile Matching*.
2. Data yang digunakan dalam tugas akhir ini terbatas pada properti yang berada di daerah JABODETABEK yang diperoleh dari platform digital penjualan properti dengan metode *scraping*.
3. Demografi dan kondisi sosial pengguna yang digunakan untuk membangun profile pengguna diperoleh dari studi literatur penelitian terdahulu.
4. Tugas akhir ini tidak akan mencakup analisis mendalam terhadap faktor eksternal yang mempengaruhi preferensi pengguna, seperti tren pasar properti atau perubahan ekonomi*.*

## Tujuan

Berdasarkan rumusan masalah yang telah diuraikan sebelumnya, tujuan yang ingin dicapai pada tugas akhir ini adalah:

1. Mengidentifikasi dan menentukan data demografi serta kondisi sosial yang relevan untuk membentuk profil pembeli properti.
2. Mengembangkan dan menerapkan algoritma Profile Matching untuk mencocokkan properti dengan profil pembeli berdasarkan karakteristik pengguna.
3. Menerapkan teknik *Hybrid Filtering* untuk menentukan properti yang mirip berdasarkan karakteristik item, seperti lokasi, harga, ukuran, dan fasilitas.
4. Mengembangkan metode untuk mengukur tingkat kepuasan pengguna terhadap rekomendasi yang diberikan oleh sistem.

## Manfaat

Manfaat yang diharapkan dari tugas akhir ini adalah:

1. Bagi Industri Properti, membantu dalam memahami preferensi konsumen dengan lebih baik, sehingga dapat menyesuaikan penawaran properti sesuai dengan kebutuhan pasar
2. Bagi pengguna platform pencarian properti, tugas akhir ini akan membantu mereka menemukan properti yang lebih relevan dan sesuai dengan profil dan kebutuhan spesifik mereka, sehingga proses pencarian menjadi lebih mudah dan efisien.
3. Bagi pengembang sistem rekomendasi, tugas akhir ini dapat menjadi acuan dalam mengembangkan sistem rekomendasi yang lebih personal dan dapat mengatasi masalah *cold-start*, terutama pada aplikasi pencarian properti.
4. Bagi penulis, tugas akhir ini membantu menambah wawasan dalam hal sistem rekomendasi dan profile matching, serta penerapannya dalam dunia nyata.

## Relevansi

Tugas akhir ini memiliki relevansi terhadap salah satu mata kuliah laboratorium Rekayasa Data dan Intelegensi Bisnis (RDIB), yaitu Pemodelan Sistem Kognitif. Pada mata kuliah tersebut, diajarkan teori-teori tentang sistem rekomendasi yang membantu dalam pembuatan tugas akhir ini. Salah satu *roadmap* penelitian laboratorium RDIB, yaitu *Recommender Systems,* memiliki kesesuaian dengan tugas akhir ini.

A close-up of a poster

AI-generated content may be incorrect.

**Gambar 1.1** Roadmap Laboratorium RDIB

# TINJAUAN PUSTAKA

## Hasil Penelitian Terdahulu

Terdapat beberapa penelitian terdahulu yang menjadi landasan dari penelitian yang dilakukan saat ini dan menjadi penunjang dalam penelitian tugas akhir ini.

**Tabel 2.1** Literatur 1

|  |  |
| --- | --- |
| Judul | Recommendation System for Property Search using Content-Based Filtering Method |
| Nama, Tahun | (Badriyah, Azvy, Yuwono, & Syarif, 2018) |
| Gambaran umum penelitian | Penelitian ini mengembangkan sistem rekomendasi properti berbasis web yang menggunakan metode *content-based filtering* untuk membantu calon pembeli menemukan properti yang sesuai dengan preferensi mereka. Sistem ini menggunakan algoritma Apriori untuk menemukan pola dalam data kunjungan pengguna serta metode TF-IDF untuk menganalisis kata kunci pada iklan properti. Sistem ini dirancang untuk meningkatkan efisiensi pencarian properti bagi pengguna berdasarkan preferensi mereka sebelumnya. |
| Keterkaitan penelitian | Penelitian ini relevan dengan tugas akhir ini yang menerapkan metode TF-IDF pada deskripsi properti untuk menemukan informasi pendukung yang relevan untuk direkomendasikan ke pengguna. |

**Tabel 2.2** Literatur 2

|  |  |
| --- | --- |
| Judul | Implementasi Metode Profile Matching Terhadap Sistem Pendukung Keputusan Penerimaan Dana Zakat pada Badan Amil Zakat Pertamina (BAZMA) |
| Nama, Tahun | (Fadilah, 2018) |
| Gambaran umum penelitian | Penelitian ini menggunakan metode *Profile Matching* dalam Sistem Pendukung Keputusan (SPK) untuk proses seleksi penerima dana zakat di Badan Amil Zakat Pertamina (BAZMA). Metode *Profile Matching* dipilih karena keandalannya dalam menghasilkan perangkingan dengan tingkat akurasi yang tinggi. Hasil penelitian ini menunjukkan bahwa penerapan metode *Profile Matching* dalam SPK mampu membantu koordinator dalam pengambilan keputusan secara efisien, menghasilkan keluaran dengan tingkat akurasi yang baik, serta mengurangi subjektivitas dalam evaluasi penerimaan dana zakat. |
| Keterikaitan penelitian | Penelitian ini berkaitan dengan tugas akhir ini, di mana *Profile Matching* diterapkan untuk menghasilkan rekomendasi properti yang lebih terpersonalisasi dengan *user profile* yang telah dibuat. |

**Tabel 2.3** Literatur 3

|  |  |
| --- | --- |
| Judul | *Solving Cold Start Problem for Recommendation System Using Content-Based Filtering* |
| Nama, Tahun | (Chia & Najafabadi, 2022) |
| Gambaran umum penelitian | Penelitian ini mengusulkan metode *content-based filtering* dengan teknik pengukuran kesamaan untuk menyelesaikan permasalahan *cold-start* bagi pengguna baru dalam sistem rekomendasi. Teknik pengukuran kesamaan yang digunakan dalam penelitian ini yaitu *Cosine Similarity* (CS), *Pearson Correlation Coefficient* (PPMCC), dan *Ecludiean Distance* (ED). Hasil dari penelitian menunjukan bahwa CS lebih unggul dalam hal pengukuran kesamaan untuk *content-based filtering* dalam sistem rekomendasi. CS menghasilkan rekomendasi yang lebih presisi dan memuaskan jika dibandingka dengan PPMCC dan ED. |
| Keterikaitan penelitian | Penelitian ini berkaitan dengan tugas akhir ini yang menerapkan pengukuran kesamaan *Cosine Similarity* setelah mengimplementasikan teknik TF-IDF untuk mengukur kesamaan suatu properti dengan properti yang lainnya. |

**Tabel 2.4** Literatur 4

|  |  |
| --- | --- |
| Judul | *Decision-Table Based Testing* |
| Nama, Tahun | (YK, 2015) |
| Gambaran umum penelitian | Penelitian ini mengimplementasikan *Decision-Table Based Testing* dalam pengujian *black-box* untuk memvalidasi fungsionalitas sistem dengan mengevaluasi secara sistematis kondisi input dan tindakan yang terkait guna menghasilkan kasus uji yang komprehensif. Dibandingkan dengan metode *equivalence partitioning* dan analisis nilai batas (*boundary value analysis*), pendekatan ini menghasilkan lebih sedikit kasus uji tetapi lebih efektif, sehingga sangat sesuai untuk memvalidasi sistem berbasis aturan, seperti sistem rekomendasi, yang output-nya sangat bergantung pada kondisi yang digerakkan oleh atribut. |
| Keterikaitan penelitian | Penelitian ini berhubungan dengan tugas akhir ini yang menerapkan *Decision-Table Based Testing* untuk mengevaluasi hasil rekomendasi dari sistem rekomendasi yang dibuat. Dengan menggunakan *Decision-Table Based Testing,* masukan dan tindakan dapat dikendalikan untuk menguji beberapa kemungkinan yang bisa saja terjadi di dunia nyata. |

## Dasar Teori

Sub bab ini berisi mengenai penjelasan beberapa teori yang menjadi bahasan pada tugas akhir. Teori-teori ini digunakan sebagai penunjang pada penelitan tugas akhir.

### **Properti**

Properti dapat didefinisikan sebagai sesuatu yang dimiliki atau dikuasai oleh seseorang, baik berupa aset fisik maupun abstrak (Merriam-Webster, 2024). Dalam konteks hukum, definisi properti mencakup tanah beserta segala sesuatu yang berada di atas atau terhubung langsung dengan tanah tersebut. Menurut NYC BAR, hal ini meliputi benda-benda buatan manusia, seperti bangunan, struktur, jalan, dan pagar, namun tidak mencakup benda-benda yang dapat diambil dari tanah tanpa menyebabkan kerusakan pada tanah itu sendiri (NYC BAR, 2024). Dengan demikian, properti tidak hanya dilihat sebagai hak kepemilikan atas objek, tetapi juga sebagai kontrol terhadap penggunaan dan nilai objek tersebut, termasuk keterkaitannya dengan tanah dan konstruksi di atasnya.

Di Indonesia, kepemilikan tanah dan properti diatur melalui berbagai regulasi hukum, dengan dasar utama adalah Undang-Undang Pokok Agraria (UUPA) Nomor 5 Tahun 1960. UU ini menguraikan berbagai jenis hak atas tanah, seperti Hak Milik, Hak Guna Bangunan (HGB), Hak Guna Usaha (HGU), Hak Pakai, dan Hak Pengelolaan, yang masing-masing memiliki karakteristik serta persyaratan spesifik. Hak Milik, yang merupakan hak tertinggi dan diberikan hanya kepada warga negara Indonesia, memberikan kepemilikan penuh dan diwariskan tanpa batasan waktu. Hak Guna Bangunan (HGB) memberikan izin untuk mendirikan bangunan di atas tanah milik negara atau pihak lain, dan dapat dimiliki oleh warga negara asing melalui mekanisme Penanaman Modal Asing (PMA). Hak Guna Usaha (HGU) diperuntukkan bagi kegiatan usaha agrikultur dan perikanan dengan jangka waktu terbatas, sementara Hak Pakai mengizinkan pemanfaatan tanah oleh warga asing dalam kondisi tertentu, meskipun bernilai lebih rendah daripada Hak Milik atau HGB (Republik Indonesia, 1960). Pada tugas akhir ini lebih difokuskan kepada properti yang memiliki sertifikat Hak Milik dan HGB.

### **Sistem Rekomendasi**

Sistem rekomendasi (SR) adalah sistem penyaringan informasi yang bertujuan untuk mengatasi permasalahan kelebihan informasi dengan menyaring hal yang berkemungkinan cocok dengan preferensi atau selera pengguna. Sistem ini membantu penggua dengan cara menyaring informasi yang tidak relevan saat pengguna mencari informasi yang diinginkan (Philip, Shola, & John, 2014). Sistem ini bekerja dengan menyeleksi potongan informasi penting dari sejumlah besar data yang dihasilkan secara dinamis, berdasarkan preferensi, minat, atau perilaku pengguna terhadap suatu hal (Isinkaye et al., 2015). Pada umumnya SR dapat dikelompokkan ke dalam tiga kategori utama, yaitu *content-based filtering* (CBF), *collaborative filtering* (CF), dan *hybrid filtering*. Selain ketiga kategori utama tersebut, terdapat sistem rekomendasi yang lebih spesifik yaitu *knowledge-based recommender systems* (KBRS).

*Content-Based Filtering* (CBF) adalah pendekatan dalam SR yang memanfaatkan karakteristik atau atribut item untuk memberikan rekomendasi. Pendekatan ini menilai kesesuaian antara profil pengguna dengan konten item menggunakan informasi historis dari item yang telah dinilai atau dikonsumsi oleh pengguna (Isinkaye et al., 2015).

Meskipun CBF cocok digunakan untuk produk yang sering dibeli, seperti buku, berita, dan musik, CBF akan kesulitan saat merekomendasikan produk yang tidak sering dibeli, seperti Komputer, mobil, layanan keuangan, pinjaman, dan apartemen. Hal ini dikarenakan sulitnya mengumpulkan ulasan dari produk yang telah dibeli oleh pengguna lain akibat kuantitas pembelian yang sedikit Untuk mengatasi tantangan tersebut, *Knowledge-Based Recommender Systems* (KBRS) dapat digunakan untuk memanfaatkan pengetahuan mendalam tentang produk dan kebutuhan pengguna (Ameen, 2019). KBRS adalah jenis SR yang didasarkan pada pengetahuan eksplisit tentang hubungan antara preferensi pengguna dan atribut produk tertentu (Rosa et al., 2019).

Salah satu pendekatan dari KBRS adalah *constraint-based recommendation systems* atau *Rule-based recommendation systems*, dimana sistem ini merekomendasikan produk berdasarkan aturan eksplisit yang ditentukan sebagai pembatas dalam basis pengetahuan. Aturan-aturan ini mendefinisikan pemetaan antara kebutuhan pengguna dan fitur produk. *Knowledge based* terdiri dari kebutuhan pengguna, properti produk, dan aturan- aturan yang ada. Produk yang sesuai dengan aturan dan memenuhi kebutuhan pengguna kemudian dihasilkan sebagai rekomendasi (Ameen, 2019).

*Hybrid Filtering* menggabungkan dua atau lebih jenis pendekatan berbeda untuk meningkatkan akurasi dan performa hasil rekomendasi (Pereira et al., 2020). *Hybrid Filtering* memiliki beberapa teknik yang digunakan untuk menggabungkan dua atau lebih pendekatan. Teknik *Weighted* menggabungkan nilai dari beberapa rekomendasi yang diberi bobot sesuai dengan kepentingannya. Teknik *Switching* menggunakan dua atau lebih pendekatan rekomendasi yang digabungkan secara bergantian dengan cara memilih metode yang paling sesuai berdasarkan kondisi tertentu. Teknik *Mixed* menggabungkan rekomendasi dari berberapa pendekatan secara langsung, sehingga pengguna menerima hasil yang mencakup rekomendasi dari semua pendekatan yang digunakan. Teknik *Cascade* menerapkan pendekatan rekomendasi secara berurutan, di mana hasil rekomendasi disaring atau disempurnakan oleh pendekatan berikutnya. Teknik *Feature Augmentation* memanfaatkan *output* dari suatu pendekatan rekomendasi sebagai input atau fitur tambahan untuk pendekatan lainnya. Pada teknik *Meta-level* model yang dihasilkan oleh suatu pendekatan rekomendasi digunakan sebagai input untuk pendekatan lainnya (Rahmawati et al., 2018).

Tugas akhir ini mengimplementasikan pendekatan *hybrid filtering* dengan teknik *feature augmentation*, di mana *output* dari pendekatan CBF sebagai *input* untuk pendekatan KBRS.

### **Term Frequency Inverse Document Frequency (TF-IDF)**

TF-IDF adalah metode statistik yang digunakan untuk mengukur kepentingan sebuah kata dalam suatu dokumen didalam kumpulan badan dokumen. Metode ini menggabungkan dua konsep utama yaitu Frequency Term (TF) dan Inverse Document Frequency (IDF). TF berguna untuk munjukan frekuensi kata dalam dokumen. Dimana, semakin sering sebuah kata muncul, maka sekain relevan kata tersebut dengan isi dokumen. IDF digunakan untuk menunjukan seberapa jarang sebuah kata muncul di seluruh badan dokumen. Semakin jarang sebuah kata muncul, semakin tinggi nilainya, karena kata-kata tersebut memiliki kemampuan yang lebih besar untuk membedakan antar berbagai kategori dokumen (Zhou, 2022).

Rumus dasar dari TF-IDF diberikan sebagai berikut:

1. Term Frequency (TF)

TF digunakan untuk menghitung berapa banyak sebuah kata atau istilah muncul pada sebuah dokumen. Contohnya jika terdapat kata “Dijual” yang muncul dalam sebuah dokumen sebanyak lima kali, maka untuk mencari TF, jumlah sebuah kata yang muncul dibagi dengan jumlah keseluruhan kata yang ada pada dokumen tersebut. Sehingga didapatkan rumus sebagai ini:

|  |  |  |
| --- | --- | --- |
|  |  | (2.1) |

*tfx,y* adalah TF dari kata *ti* dalam dokumen *dj*, *Nx,y* adalah jumlah kemunculan kata *ti* dalam dokumen *tj*, ∑*nz,y* adalah jumlah total semua kata dalam dokumen *dj*, sehingga TF menunjukan proporsi kemunculan *ti* dibandingkan seluruh kata dalam dokumen tersebut (Qaiser & Ali, 2018).

1. Inverse Document Frequency (IDF)

Berbanding terbalik dengan TF, IDF digunakan untuk memberikan bobot yang lebih kepada kata atau istilah yang tidak terlalu bermakna atau tidak penting, namun banyak muncul pada suatu dokumen. Sebagai contoh terdapat kata “Rumah” yang muncul di tiga dari lima dokumen yang dimiliki. Maka didapatkan rumus untuk menghitung bobot IDF sebagai berikut:

|  |  |  |
| --- | --- | --- |
|  |  | (2.2) |

D adalah jumlah total dokumen, dan adalah jumlah dokumen yang mengandung kata *ti* (Qaiser & Ali, 2018).

1. TF-IDF Weight dihtung dengan mengalikan TF dan IDF, sebagai berikut:

TF-IDF menggabungkan kedua perhitungan sebelumnya untuk menghasilkan bobot untuk sebuah kata atau istilah yang memiliki nilai TF yang tinggi dan lebih sedikit muncul pada suatu dokumen sehingga kata tersebut menghasilkan kepentingan yang lebih tinggi (IDF) untuk kata kunci yang dicari dalam sebuah dokumen.

|  |  |  |
| --- | --- | --- |
|  |  | (2.3) |

Nilai dari *tfidf* didapat dari hasil perkalian antara nilai *tf* dan *idf* yang telah dihitung sebelumnya (Qaiser & Ali, 2018).

Agar perhitungan TF-IDF lebih efektif dan akurat, diperlukan langkah-langkah *preprocessing* untuk mempersiapkan data teks. *Preprocessing* ini bertujuan untuk membersihkan teks dari elemen-elemen yang tidak relevan dan menyederhanakan representasi data (Babu Dasari et al., 2024). Dua langkah penting dalam *preprocessing* teks adalah *tokenization* dan *stop words removal*.

*Tokenization* adalah proses membagi teks menjadi unit-unit kecil yang disebut token, seperti kata atau frasa. *Tokenization* memungkinkan teks untuk dianalisis lebih lanjut, seperti perhitungan frekuensi kata (Babu Dasari et al., 2024). Contoh proses *tokenization* pada kalimat “Rumah dijual dengan harga murah” yang berubah menjadi ["Rumah", "dijual", "dengan", "harga", "murah"].

*Stop words* adalah kata-kata umum seperti "dan", "di", "yang", atau "adalah" dalam bahasa Indonesia, yang biasanya tidak memiliki makna signifikan untuk analisis. Menghapus *stop words* mengurangi *noise* dalam data dan meningkatkan fokus pada kata-kata penting (Babu Dasari et al., 2024). Hasil dari *tokenization* ["Rumah", "dijual", "dengan", "harga", "murah"] berubah menjadi ["Rumah", "dijual", "harga", "murah"] setelah dilakukannya proses *stop words removal*.

### **Cosine Similarity**

Cosine similarity adalah metode yang digunakan untuk mengukur derajat kesamaan antara dua vektor, di mana nilai kesamaan tersebut dihitung dengan nilai sudut kosinus yang terbentuk dari kedua vektor yang dibandingkan. Metode ini sering digunakan dalam pemrosesan teks dan klasifikasi data, karena dapat menentukan seberapa mirip suatu teks dengan teks lainnya dalam ruang vektor. Semakin kecil sudut yang terbentuk antara dua vektor, semakin besar nilai kesamaannya, dan nilai kesamaan maksimal (1) diperoleh ketika sudut antara kedua vektor tersebut adalah nol derajat (Lahitani et al., 2016). Adapun rumus untuk cosine similarity antara dua vektor *d1* dan *d2*, seperti berikut:

|  |  |  |
| --- | --- | --- |
|  |  | (2.4) |

Di mana adalah dot product dari vektor *d1* dan *d2*, dan adalah panjang dari dua vektor *d1* dan *d2*, dan adalah hasil perkalian regular dari panjang dua vektor *d1* dan *d2* (Babu Dasari et al., 2024).

### **K-means Clustering**

*Clustering* adalah teknik dalam *machine learning* yang bertujuan untuk mengelompokkan data ke dalam kelompok berdasarkan kesamaan tertentu. Teknik ini termasuk dalam *unsupervised learning*, di mana data tidak memiliki label sebelumnya, sehingga proses pengelompokkan didasarkan pada struktur intrinsik data. Data yang berada dalam kelompok yang sama cenderung memiliki karakteristik yang lebih serupa dibandingkan dengan data di kelompok lain (Wang et al., 2017).

K-means adalah salah satu algoritma *clustering* yang paling sering digunakan karena kesederhanaan dan efisiensinya. Algoritma ini bertujuan untuk meminimalkan jarak antar data dalam satu cluster (itra-cluster) dan memaksimalkan jarak antar *cluster* (inter-cluster) (Punhani et al., 2022). Dalam prosesnya, algoritma ini memulai dengan menentukan sejumlah k centroid awal secara acak. Selanjutnya, data akan dialokasikan ke *cluster* dengan centroid terdekat, lalu centroid diperbarui berdasarkan rata-rata data dalam *cluster* tersebut. Proses ini berulang hingga tidak ada perubahan signifikan dalam posisi centroid (Wang et al., 2017)​.

### **Profile matching**

Metode *profile matching* atau pencocokan profil adalah proses perbandingan antar kompetensi untuk mengetahui perbedaan diantara kompetensi tersebut, atau yang biasa di sebut *gap*. Semakin kecil *gap* maka nilai yang dihasilkan semakin besar dari bobotnya. Metode *profile matching* akan memberikan hasil akhir berupa pemeringkatan untuk setiap kompetensi. Pemeringkatan ini yang akan menjadi dasar dalam mengambil seuatu keputusan (Afijal et al., 2014).

Langkah pertama dalam proses ini adalah identifikasi kriteria, di mana parameter atau kompetensi yang relevan dengan evaluasi didefinisikan. Setiap parameter diklasifikasikan ke dalam Core Factor (essensial) dan Secondary Factor (pendukung), sehingga memberikan dasar yang tersetruktur untuk analisis (Verdian & Wantoro, 2019).

Pada langkah definisi profil target, nilai ideal ditetapkan untuk setiap kriteria sebagai tolok ukur perbandingan. Selanjutnya, pada tahap pengumpulan data, nilai aktual untuk semua subjek atau alternatif yang dievaluasi dikumpulkan dan dinormalisasi guna memastikan konsistensi dalam dataset (Suhartanto et al., 2016).

Tahap berikutnya adalah *gap analysis*, yang menghitung selisih antara nilai aktual dan nilai ideal untuk setiap kriteria dengan rumus

|  |  |  |
| --- | --- | --- |
|  |  | (2.5) |

Kesenjangan yang lebih kecil menunjukkan keselarasan yang lebih dekat dengan profil ideal, sementara kesenjangan yang lebih besar menunjukkan area yang memerlukan perbaikan. Setiap kesenjangan diberi bobot untuk mencerminkan pentingnya, di mana penyimpangan yang lebih kecil diberikan bobot lebih tinggi.

*Core Factor* dan *Secondary Factor* dihitung secara terpisah untuk menangkap atribut kritis dan pendukung. *Core Factor*(NCF) dihitung dengan rumus:

|  |  |  |
| --- | --- | --- |
|  |  | (2.6) |

di mana NC adalah total skor komponen faktor inti, dan IC adalah jumlah item faktor inti. *Secondary Factor*(NSF) dihitung dengan rumus serupa:

|  |  |  |
| --- | --- | --- |
|  |  | (2.7) |

di mana NS adalah total skor komponen faktor sekunder, dan IS adalah jumlah item faktor sekunder. Perhitungan ini memungkinkan pemberian bobot pada kompetensi inti sambil tetap mempertimbangkan faktor pendukung (Verdian & Wantoro, 2019).

Pada fase penentuan bobot, skor total untuk setiap subjek dihitung dengan menggabungkan faktor inti dan faktor sekunder:

|  |  |  |
| --- | --- | --- |
|  |  | (2.8) |

Pendekatan berbobot ini memastikan bahwa atribut kritis memiliki pengaruh lebih besar terhadap evaluasi akhir, sambil tetap memperhitungkan atribut pendukung.

Pada tahap perankingan, subjek diurutkan berdasarkan skor total mereka untuk mengidentifikasi kesesuaian paling dekat dengan profil ideal. Proses ini menjamin pengambilan keputusan yang objektif dan berbasis data.

### **User Profile**

Preferensi generasi milenial dalam memilih hunian dipengaruhi oleh berbagai faktor seperti lokasi, karakteristik unit hunian, biaya, fasilitas, serta gaya hunian. Penelitian ini bertujuan untuk memahami atribut-atribut yang dianggap penting oleh generasi milenial dalam memilih hunian, khususnya di wilayah perkotaan yang memiliki karakteristik unik seperti Kota Bekasi dan DKI Jakarta. Generasi ini cenderung mengutamakan kenyamanan, aksesibilitas, dan efisiensi biaya dalam memilih tempat tinggal.

Penelitian dari kedua paper menyebutkan bahwa:

1. **Kota Bekasi** (Puspitasari et al., 2022) menjadi lokasi penting karena merupakan salah satu wilayah penyangga Jakarta dengan tingkat pertumbuhan generasi milenial yang tinggi. Penelitian ini menggunakan metode kuantitatif dengan data dari 436 responden berusia 20-39 tahun. Data dikumpulkan melalui kuisioner online menggunakan metode snowball sampling. Fokus utama penelitian ini adalah pada kualitas lingkungan seperti keamanan, kebersihan, dan lokasi bebas banjir.
2. **DKI Jakarta** (Yustika et al., 2022) memiliki karakteristik urban yang khas dengan keterbatasan lahan dan tingginya harga properti. Penelitian ini menggunakan metode kualitatif dengan studi literatur pada 369 sampel. Fokusnya mencakup empat faktor utama: lokasi, biaya, atribut fisik, dan legalitas hunian. Di DKI Jakarta, generasi milenial mempertimbangkan kedekatan fasilitas umum seperti transportasi publik, pusat aktivitas kota, dan pendidikan.

Tabel 2.5 Tabel User Profile

|  |  |  |  |
| --- | --- | --- | --- |
| Kebutuhan Pengguna | Pasangan Bekerja dengan Anak | Pasangan Bekerja tanpa Anak | Individu Lajang |
| Lokasi | - Lokasi bebas banjir, aman, dekat dengan fasilitas pendidikan, kesehatan, dan transportasi umum.  - Akses mudah ke jalan utama dan transportasi umum. | - Lokasi strategis di area perkotaan dengan akses ke transportasi umum dan tempat kerja.  - Bebas banjir dan dekat pusat aktivitas kota. | - Dekat pusat aktivitas kota seperti pusat perbelanjaan, kafe, dan hiburan.  - Lokasi strategis dengan akses transportasi umum yang baik. |
| Karakteristik Unit Hunian | - Rumah tapak skala menengah (200-600 m²), 3 kamar tidur, 2 kamar mandi.  - Dilengkapi ruang keluarga 5-10 m² dan taman kecil.  - Lingkungan aman, nyaman, dan bersih. | - Apartemen modern dengan luas 22-70 m², 2 kamar tidur, 2 kamar mandi.  - Desain modern dan fungsional.  - Memiliki balkon dan parkir. | - Apartemen kecil (22-50 m²), 1-2 kamar tidur.  - Desain minimalis dan modern.  - Memiliki ruang minimalis dan area parkir. |
| Fasilitas dan Akses | - Dekat fasilitas pendidikan, kesehatan, taman, dan transportasi publik.  - Bebas banjir dan dekat jalan utama. | - Dekat pusat transportasi umum, tempat kerja, dan fasilitas hiburan. | - Dekat pusat aktivitas seperti pusat belanja, restoran, dan kafe. |

### **Evaluasi**

Pada tugas akhir ini akan dilakukan dua evaluasi pada dua tahap berbeda. Tahap pertama evaluasi dilakukan setelah melakukan perhitungan *cosine similarity* dengan mengimplementasikan evaluasi *silhouette score* untuk mengevaluasi hasil clustering. Kemudian evaluasi selanjutnya dilakukan pada tahap akhir menggunakan *blackbox testing* untuk memastikan sistem bekerja sesuai dengan yang diharapkan.

*Silhouette Score* adalah metrik evaluasi yang digunakan untuk menilai kualitas hasil *clustering* dengan mengukur tingkat kohesi dalam kluster dan pemisahan antar kluster. Skor ini dihitung berdasarkan dua kompenen utama, yaitu jarak rata-rata sebuah data *i* terhadap data lain dalam klusternya sendiri (*a(i)*, disebut kohesi), dan jarak rata-rata data tersebut ke data di kluster terdekat (*b(i)*, disebut separasi) (Wang et al., 2017). Persamaan untuk menghitung *silhouette score* untuk sebuah data adalah sebagai berikut:

|  |  |  |
| --- | --- | --- |
|  |  | (2.9) |

Di mana adalah rata-rata jarak antara data *i* dengan semua data lain dalam kluster yang sama. Sedangkan adala rata-rata jarak antara data *i* dengan semua data dalam kluster terdekat (Punhani et al., 2022).

Nilai Silhouette Score berada pada rentang -1 hingga 1, di mana nilai mendekati 1 menunjukkan data terkelompok dengan baik dalam klusternya sendiri dan jauh dari kluster lain. Nilai mendekati 0 menunjukkan data berada di perbatasan antara dua kluster, sedangkan nilai negatif menunjukkan data lebih dekat ke kluster lain dibandingkan dengan klusternya sendiri​ (Ogbuabor & F. N, 2018).

Metode *Blackbox testing* adalah teknik pengujian yang digunakan untuk mengevaluasi aplikasi tanpa memerlukan pemahaman yang mendalam mengenai detail internal, seperti *source code*. Pengujian ini hanya memeriksa luaran yang dihasilkan berdasarkan masukan yang diberikan. Proses *Blackbox testing* dilakukan dengan cara mencoba program menggunakan berbagai jenis masukan yang bertujuan untuk memastikan apakah program tersebut berfungsi sesuai dengan yang diharapkan (Sasongko, et al., 2021). Terdapat beberapa metode Blackbox testing, namun dalam tugas akhir ini akan difokuskan menggunakan *Decision Table Testing*.

*Decision Table Testing* merupakan metode pengujian *Blackbox* *Testing* yang bertujuan untuk mengevaluasi perilaku sistem berdasarkan berbagai kombinasi input yang berbeda. Metode ini sering disebut juga sebagai tabel *Cause-Effect*, karena mengidentifikasi hubungan sebab-akibat antara input dan hasil yang diharapkan. *Decision Table Testing* diterapkan ketika sistem memiliki sejumlah variasi nilai input yang tidak dapat diuji secara efektif menggunakan metode *Boundary Value Analysis* atau *Equivalent Partitioning*. (Setiawan, 2022)

# METODOLOGI

## Metodologi

Pada subbab ini metodologi tugas akhir dijelaskan dalam bentuk diagram alir yang berisi tahapan-tahapan yang dilakukan dalam pengerjaan tugas akhir.

A black and white diagram

AI-generated content may be incorrect.

**Gambar 3.1** Diagram Alur Tugas Akhir

## Uraian Metodologi

Pada subbab ini, setiap tahap metodologi pada tugas akhir ini dijelaskan dengan lebih rinci.

### **Identifikasi Masalah**

Pada tahap awal tugas akhir, dilakukan identifikasi terhadap kebutuhan pengguna dalam proses pencarian properti yang relevan dan efisien. Selain itu, berbagai kendala yang sering dihadapi oleh sistem rekomendasi, seperti *cold-start* problem, relevansi hasil rekomendasi, dan efisiensi pencarian, juga dianalisis secara mendalam. Hasil dari identifikasi ini akan digunakan sebagai dasar perancangan sistem rekomendasi yang mampu mengatasi permasalahan tersebut dan memenuhi kebutuhan pengguna dengan lebih baik.

### **Studi Literatur**

Pada tahap studi literatur dilakukan kajian terhadap penelitian-penelitian terdahulu yang berkaitan dengan pengembangan sistem rekomendasi. Fokus utama kajian ini adalah metode *Content-Based Filtering* (CBF), *Knowledge-Based Recommender Systems* (KBRS), *Hybrid Filtering*, *Profile Matching*, algoritma Term Frequency-Inverse Document Frequency (TF-IDF). Tugas akhir ini juga mempelajari konsep *Cosine Similarity* sebagai metode penghitungan kesamaan antara properti dalam ruang vektor. Untuk mengevaluasi hasil rekomendasi dilakukan pengkajian terhadap salah satu metode *black box testing* yaitu *decision table-based testing*.Kajian ini bertujuan untuk membangun landasan teoritis yang kuat bagi pengembangan sistem rekomendasi properti yang lebih akurat dan efisien.

### **Pengumpulan Data**

Data yang digunakan dalam tugas akhir ini dikumpulkan melalui proses *web scraping* dari platform digital penjualan properti. Data yang dikumpulkan meliputi atribut properti, seperti lokasi, harga, ukuran properti, dan fasilitas, serta data demografi dan kondisi sosial pengguna untuk pembuatan profil pengguna. Data yang terkumpul kemudian disaring untuk menghilangkan duplikasi dan memastikan kualitas data sesuai untuk digunakan dalam proses analisis lebih lanjut.

### **Pra-proses Data**

Data yang telah terkumpul diolah melalui beberapa tahap untuk memastikan kualitasnya sebelum digunakan dalam implementasi algoritma. Proses ini melibatkan pembersihan data untuk menghapus kesalahan dan nilai yang hilang, serta dilakukan *tokenization* dan *stop words removal* untuk data yang berisi deskripsi properti. Selanjutnya, bobot diberikan pada setiap fitur teks menggunakan algoritma TF-IDF untuk memastikan relevansi antar properti dapat diukur dengan baik.

### **Implementasi Algoritma**

Pada tahap ini mengintegrasikan semua pendekatan utama untuk membangun sistem rekomendasi. Pertama, CBF dengan menggunakan metode TF-IDF untuk menganalisis relevansi properti berdasarkan atributnya, seperti deskripsi, lokasi, atau fasilitas. Kesamaan antar properti dihitung menggunakan *Cosine Similarity* untuk mendapatkan kelompok properti yang relevan. Kedua, K-means *clustering* digunakan untuk mengelompokkan properti yang memiliki kesamaan atribut. Selanjutmya, *rule-base* KBRS diimplementasikan untuk menyaring properti berdasarkan aturan yang telah dirancang sesuai dengan preferensi setiap profil pengguna. Hasil dari *rule-based* KBRS ini adalah daftar properti yang kemudian digunakan untuk langkah selanjutnya. Metode *Profile Matching* diterapkan pada daftar properti yang telah dihasilkan pada langkah sebelumnya dan dilakukan analisis *gap* untuk mencocokkan properti dengan kebutuhan pengguna dengan mempertimbangkan kesesuaian atribut-atribut penting.

### **Uji Coba & Evaluasi Algoritma**

Tahap ini dilakukan evaluasi terhadap hasil kluster dari metode K-Means untuk mengukur kesamaan antar properti dalam satu domain dengan menggunakan pengukuran *Sillhouette Score*. Selain itu, untuk memastikan algoritma berjalan sesuai yang diharapkan, peneliti mengimplementasikan uji Blackbox menggunakan metode Decision Table Testing untuk memverifikasi apakah sistem rekomendasi bekerja sesuai dengan spesifikasi yang telah dirancang. Uji ini memastikan bahwa setiap masukan menghasilkan keluaran yang diharapkan.

### **Penyusunan Buku Tugas Akhir**

Tahapan terakhir adalah penyusunan laporan tugas akhir yang memuat seluruh proses, hasil implementasi, analisis, dan evaluasi algoritma dalam tugas akhir ini.

## Waktu pengerjaan

Sub bab ini berisi lini waktu pengerjaan tugas akhir.

Tabel 3.1 Lini Waktu Pengerjaan

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No | Kegiatan | Bulan | | | | | | | | | | | | | | | |
| 1 | | | | 2 | | | | 3 | | | | 4 | | | |
| 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| 1 | Identifikasi Masalah |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 | Studi Literatur |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 | Pengumpulan Data |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 | Pra-proses Data |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 | Implementasi Algoritma |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 6 | Uji Coba & Evaluasi Algoritma |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 7 | Penyusunan Buku Tugas Akhir |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

# Hasil dan Pembahasan

## Hasil penelitian

## Pembahasan

# Kesimpulan dan Saran

## Kesimpulan

## Saran

DAFTAR PUSTAKA

Afijal, Iqbal, M., Najmuddin, & Iskandar. (2014). Decision Support System Determination for Poor Houses Beneficiary Using Profile Matching Method. Academic Research International, 5(4), 385–394.

Ameen, A. (2019). Knowledge based Recommendation System in Semantic Web - A Survey. International Journal of Computer Applications, 182(43), 20–25. https://doi.org/10.5120/ijca2019918538

Babu Dasari, S., Keerthi, P. V. L. D., & Snehitha, G. (2024). A Pipeline Approach to Research Paper Recommendation System Integrating TF-IDF and KNN. 2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems, ICITEICS 2024, 1–8. https://doi.org/10.1109/ICITEICS61368.2024.10625308

Badriyah, T., Azvy, S., Yuwono, W., & Syarif, I. (2018). Recommendation system for property search using content based filtering method. 2018 International Conference on Information and Communications Technology, ICOIACT 2018, 2018-Janua, 25–29. https://doi.org/10.1109/ICOIACT.2018.8350801

Chia, E. J., & Najafabadi, M. K. (2022). Solving Cold Start Problem for Recommendation System Using Content-Based Filtering. Proceedings - 2022 International Conference on Computer Technologies, ICCTech 2022, 38–42. https://doi.org/10.1109/ICCTech55650.2022.00015

Isinkaye, F. O., Folajimi, Y. O., & Ojokoh, B. A. (2015). Recommendation systems: Principles, methods and evaluation. Egyptian Informatics Journal, 16(3), 261–273. https://doi.org/10.1016/j.eij.2015.06.005

Lahitani, A. R., Permanasari, A. E., & Setiawan, N. A. (2016). Cosine similarity to determine similarity measure: Study case in online essay assessment. Proceedings of 2016 4th International Conference on Cyber and IT Service Management, CITSM 2016, 1–6. https://doi.org/10.1109/CITSM.2016.7577578

Ogbuabor, G., & F. N, U. (2018). Clustering Algorithm for a Healthcare Dataset Using Silhouette Score Value. International Journal of Computer Science and Information Technology, 10(2), 27–37. https://doi.org/10.5121/ijcsit.2018.10203

Pereira, J. G., Tiwari, S., & Ajoy, S. (2020). A survey on filtering techniques for recommendation system. Proceedings - 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security, ISSSC 2020. https://doi.org/10.1109/iSSSC50941.2020.9358819

Punhani, A., Faujdar, N., Mishra, K. K., & Subramanian, M. (2022). Binning-Based Silhouette Approach to Find the Optimal Cluster Using K-Means. IEEE Access, 10(October), 115025–115032. https://doi.org/10.1109/ACCESS.2022.3215568

Puspitasari, D., Adianto, J., & Khoirunurrofik, K. (2022). Preferensi Hunian di Perkotaan pada Generasi Y di Kota Bekasi. Jurnal Pembangunan Wilayah Dan Kota, 18(1), 54–65. https://doi.org/10.14710/pwk.v18i1.37176

Qaiser, S., & Ali, R. (2018). Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents. International Journal of Computer Applications, 181(1), 25–29. https://doi.org/10.5120/ijca2018917395

Rahmawati, S., Nurjanah, D., & Rismala, R. (2018). Analisis dan Implementasi pendekatan Hybrid untuk Sistem Rekomendasi Pekerjaan dengan Metode Knowledge Based dan Collaborative Filtering. Indonesian Journal on Computing (Indo-JC), 3(2), 11. https://doi.org/10.21108/indojc.2018.3.2.210

Rosa, R. L., Schwartz, G. M., Ruggiero, W. V., & Rodriguez, D. Z. (2019). A Knowledge-Based Recommendation System That Includes Sentiment Analysis and Deep Learning. IEEE Transactions on Industrial Informatics, 15(4), 2124–2135. https://doi.org/10.1109/TII.2018.2867174

Suhartanto, A., kusrini, & henderi. (2016). Decision Support System untuk Penilaian Kinerja Guru dengan Metode Profile Matching. Jurnal Komputer Terapan, 2(2), 149–158. http://jurnal.pcr.ac.id

Verdian, A., & Wantoro, A. (2019). Komparasi Metode Profile Matching Dengan Fuzzy Profile Matching Pada Pemilihan Wakil Kepala Sekolah. Jurnal Ilmiah Media Sisfo, 13(2), 97–105. https://doi.org/10.33998/mediasisfo.2019.13.2.652

Wang, F., Franco-Penya, H. H., Kelleher, J. D., Pugh, J., & Ross, R. (2017). An analysis of the application of simplified silhouette to the evaluation of k-means clustering validity. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10358 LNAI, 291–305. https://doi.org/10.1007/978-3-319-62416-7\_21

YK, A. (2015). Decision Table Based Testing. International Journal on Recent and Innovation Trends in Computing and Communication, 3(3), 1298–1301. https://doi.org/10.17762/ijritcc2321-8169.150388

Yustika, F. P., Ridlo, M. A., Widyasamratri, H., Islam, U., & Agung, S. (2022). PREFERENSI GENERASI MILENIAL DALAM MEMILIH HUNIAN STUDI KASUS : DKI JAKARTA , JAKARTA BARAT , DAN Generasi milenial yaitu generasi terbanyak di Indonesia , Menurut data Survey Susenas memilih hunian serta faktor-faktor apa saja yang menjadi acuan bagi gen. 2(1), 72–88.

Zhou, H. (2022). Research of Text Classification Based on TF-IDF and CNN-LSTM. Journal of Physics: Conference Series, 2171(1), 218–222. https://doi.org/10.1088/1742-6596/2171/1/012021

LAMPIRAN

BIODATA PENULIS

Penulis dilahirkan di Madiun, 29 Januari 1985, merupakan anak pertama dari 4 bersaudara. Penulis telah menempuh pendidikan formal yaitu di TK ABA 18 Madiun, SDN Beteng 1 Madiun, SMPN 2 Madiun dan SMAN 2 Madiun. Setelah lulus dari SMAN tahun 2020, Penulis mengikuti SBMPTN dan diterima di Departemen Teknik Mesin FTIRS - ITS pada tahun 2020 dan terdaftar dengan NRP 02112040000130.

Photo closed-up



Di Departemen Teknik Mesin Penulis sempat aktif di beberapa kegiatan Seminar yang diselenggarakan oleh Departemen, Himpunan Mahasiswa Teknik Mesin (HMM) dan aktif sebagai Asisten Praktikum Mesin Konversi Enersi maupun Grader mata kuliah Termodinamika.