

PROPOSAL TUGAS BESAR DASAR KECERDASAN ARTIFICIAL

Perancangan Sistem Rekomendasi Buku Menggunakan Logika Fuzzy Berdasarkan Dataset *Goodbooks*



Disusun Oleh:

- | | |
|---------------------------|--------------|
| - Velisia Nihan Rahmawati | 103012300203 |
| - Nabila Putri Aulia | 103012330531 |

**PROGRAM STUDI S1 INFORMATIKA
TELKOM UNIVERSITY
2024/2025**

Latar Belakang

Di era digital, pembaca dihadapkan pada banyak pilihan bacaan yang tersedia secara daring, seperti pada platform Goodreads. Banyaknya pilihan ini membuat proses seleksi buku yang sesuai menjadi semakin kompleks dan membingungkan, terutama ketika preferensi pengguna bersifat subjektif dan tidak terstruktur (Ricci et al., 2011). Sistem rekomendasi menjadi salah satu solusi penting untuk membantu pengguna menemukan buku yang sesuai dengan minat dan kebutuhannya, dengan memanfaatkan data historis dan perilaku pengguna.

Namun, sistem rekomendasi konvensional seperti collaborative filtering atau content-based filtering sering kali tidak cukup efektif dalam menangani ketidakpastian, ambiguitas, dan subjektivitas dari preferensi pengguna (Han et al., 2012). Di sinilah pendekatan logika fuzzy menjadi relevan. Logika fuzzy, yang diperkenalkan oleh Zadeh (1965), memungkinkan sistem untuk menangani informasi yang tidak pasti atau samar dan memodelkan proses pengambilan keputusan dalam bentuk linguistik, seperti “tinggi”, “sedang”, atau “rendah”, yang lebih menyerupai cara berpikir manusia.

Dengan demikian, penerapan logika fuzzy dalam sistem rekomendasi buku diharapkan dapat meningkatkan akurasi dan interpretabilitas rekomendasi, terutama dalam konteks preferensi pengguna yang tidak selalu bisa dijelaskan secara matematis atau eksplisit.

Tujuan

- Merancang dan membangun sistem rekomendasi buku yang dapat memberikan saran bacaan kepada pengguna berdasarkan data numerik seperti *average rating*, jumlah ulasan, dan jumlah ulasan bintang 5 buku dari dataset *Goodbooks*.
- Mengimplementasikan pendekatan logika fuzzy dalam proses evaluasi dan klasifikasi buku, guna menangani ketidakpastian serta subjektivitas dalam data numerik maupun preferensi pengguna.
- Melakukan transformasi data numerik menjadi kategori linguistik, agar hasil rekomendasi lebih mudah dipahami dan dapat digunakan oleh pengguna awam secara intuitif.

- Membandingkan performa pendekatan fuzzy dengan algoritma machine learning seperti Decision Tree dan metode Matrix Factorization dalam konteks sistem rekomendasi.

Permasalahan

Masalah utama yang ingin diselesaikan dalam penelitian ini adalah memprediksi atau merekomendasikan buku yang sesuai untuk pengguna berdasarkan informasi yang tersedia seperti rating, jumlah pembaca, dan jumlah rating bagus. Selain permasalahan utama tersebut, terdapat pula sejumlah isu tambahan yang mendasari penggunaan logika fuzzy dalam proses evaluasi kualitas buku dan penentuan rekomendasi, yaitu:

1. **Kombinasi multi-kriteria yang kompleks**

Sistem rekomendasi tidak hanya bergantung pada satu nilai seperti rating saja, melainkan juga faktor lain seperti *number of reviews*, *publication year*, dan *ratings count*. Fuzzy logic memungkinkan integrasi berbagai kriteria tersebut ke dalam satu kerangka penilaian yang logis dan dapat dikustomisasi sesuai skenario

2. **Ketidakpastian dan ambiguitas dalam preferensi pengguna**

Beberapa indikator seperti average rating, number of ratings, atau rating count tidak memiliki batasan tegas untuk menentukan apakah sebuah buku itu “bagus”, “sedang”, atau “kurang”. Sebagai contoh, rating 4.0 bisa dianggap bagus oleh sebagian pengguna, tetapi biasa saja bagi pengguna lain. Logika fuzzy membantu menangani ketidakpastian ini dengan representasi linguistik seperti “tinggi”, “sedang”, dan “rendah”.

3. **Kebutuhan untuk mentransformasikan data numerik menjadi kategori linguistik yang lebih mudah dipahami**

Agar sistem rekomendasi dapat digunakan oleh pengguna awam atau dikembangkan lebih lanjut ke sistem berbasis antarmuka (misal chatbot atau aplikasi mobile), maka nilai-nilai numerik seperti “average rating = 4.2” perlu dikonversi menjadi istilah linguistik seperti “sangat bagus”.

Solusi

Untuk menyelesaikan permasalahan klasifikasi dan rekomendasi buku berdasarkan data Goodbooks, solusinya adalah penerapan metode yang menggabungkan pendekatan Logika Fuzzy dan Machine Learning. Pendekatan ini ditujukan untuk mengatasi ketidakpastian pada nilai-nilai atribut buku serta meningkatkan akurasi rekomendasi berdasarkan preferensi pengguna. Berikut adalah rincian solusi yang akan diterapkan:

1. Logika Fuzzy untuk Penilaian Kualitas Buku

Digunakan sistem inferensi fuzzy (Mamdani/Sugeno) untuk menilai apakah sebuah buku layak direkomendasikan, berdasarkan 3 parameter utama:

- Average Rating: dinyatakan dalam fuzzy set [*rendah, sedang, tinggi*]
- Number of Ratings: dinyatakan dalam fuzzy set [*Sedikit, Sedang, Banyak*]
- Number of 5 Stars Ratings: dinyatakan dalam fuzzy set [*Sedikit, Sedang, Banyak*]

Contoh aturan fuzzy:

IF Average Rating is Tinggi AND Number of Ratings is Banyak AND Number of 5 Stars Ratings is Sedang THEN:

Output = Sangat Direkomendasikan (Mamdani)

2. Transformasi Data Numerik ke Linguistik

Nilai-nilai seperti:

- Average rating (contoh: 4.2),
- Jumlah rating (contoh: 12.000),
- Jumlah rating dengan bintang 5 (contoh: 500),

akan dikonversi menjadi istilah linguistik berdasarkan fungsi keanggotaan fuzzy. Hal ini memungkinkan sistem untuk menjelaskan rekomendasi dengan kalimat seperti: “Buku ini sangat direkomendasikan karena memiliki rating tinggi dan jumlah ulasan yang banyak.”

3. Machine Learning untuk Pembandingan

- Digunakan algoritma Decision Tree Classifier untuk

mengklasifikasikan apakah buku tersebut akan disukai pengguna, berdasarkan data rating sebelumnya.

4. Matrix Factorization

- Digunakan untuk menganalisis pola interaksi antara pengguna (reader) dan buku berdasarkan rating yang diberikan.
- Teknik ini memecah matriks rating pengguna-buku menjadi dua matriks yang merepresentasikan karakteristik pengguna dan buku.
- Hasilnya membantu memprediksi rating yang belum diberikan pengguna dan mengukur seberapa dekat rating pengguna dengan rating buku (sebagai pembandingan).
- Evaluasi menggunakan Mean Absolute Error (MAE) untuk mengukur akurasi prediksi rating.

Dataset dan Atributnya

Dataset yang kami pilih dari Kaggle berisi informasi lebih dari 20.000 buku, termasuk atribut seperti:

Atribut	Deskripsi
id	ID internal unik untuk setiap entri buku di dataset.
book_id	ID buku dari sistem Goodreads. Ini adalah kunci utama yang sering digunakan untuk penggabungan data.
best_book_id	ID dari versi “terbaik” atau utama dari buku tersebut di Goodreads, sering digunakan untuk menyatukan berbagai edisi.
work_id	ID dari “work” Goodreads, yang mengelompokkan berbagai edisi

	dari sebuah karya buku.
books_count	Jumlah total edisi dari buku tersebut di Goodreads.
isbn	ISBN (International Standard Book Number), identifikasi edisi cetakan buku.
isbn13	Versi ISBN yang terdiri dari 13 digit.
authors	Nama penulis buku. Bisa lebih dari satu penulis, dipisah dengan koma.
original_publication_year	Tahun pertama kali buku diterbitkan.
original_title	Judul asli buku, tanpa tambahan edisi atau subtitle.
title	Judul lengkap buku (bisa termasuk subtitle, edisi, atau keterangan lainnya).
language_code	Kode bahasa dari buku
average_rating	Nilai rata-rata rating dari pengguna Goodreads.
ratings_count	Total jumlah rating yang diberikan oleh pengguna untuk buku tersebut.
work_ratings_count	Jumlah rating untuk seluruh “work” (semua edisi dari buku tersebut).
work_text_reviews_count	Jumlah ulasan berbentuk teks untuk seluruh edisi.
ratings_1, ratings_2, ratings_3, ratings_4, ratings_5	Jumlah rating bintang 1 hingga 5 yang diberikan ke buku tersebut.

Keuntungan Pendekatan Fuzzy

- Mampu menangani ketidakpastian dan subjektivitas data
- Lebih menyerupai cara berpikir manusia
- Cocok untuk sistem berbasis pengguna awam

Alasan Pemilihan Data Set

- Dataset berasal dari dunia nyata dan relevan dalam domain literasi digital.
- Tersedia fitur-fitur numerik dan linguistik yang cocok untuk pendekatan fuzzy.
- Dataset terstruktur dengan baik dan dapat digunakan langsung setelah preprocessing dasar.

Kesimpulan

Fuzzy System merupakan pendekatan yang efektif dan fleksibel untuk mengevaluasi kelayakan rekomendasi buku menggunakan metadata dari dataset Goodreads. Sistem ini dapat menangani nilai rating dan interpretasi panjang buku yang bervariasi. Dalam dunia nyata, proses memilih buku tidak selalu berdasarkan logika statistik yang kaku. Seringkali, pembaca mempertimbangkan aspek-aspek yang bersifat kualitatif, seperti "buku ini cukup panjang tapi sangat bagus" atau "buku ini sangat populer tapi terlalu tebal untuk dibaca sekarang". Fuzzy logic mampu menerjemahkan penilaian subjektif ini ke dalam bentuk sistem yang terstruktur dan dapat dihitung.

Dataset Goodbooks memberikan cukup fitur untuk membangun sistem rekomendasi yang cukup efektif. Metode ini dapat dikembangkan lebih lanjut sebagai filter awal sebelum merekomendasikan buku kepada pengguna secara personal.

\

F. Referensi

Kaggle, "Goodbooks-10k Dataset."

<https://www.kaggle.com/datasets/zygmunt/goodbooks-10k>

Zadeh, L. A. (1965). "Fuzzy sets." *Information and Control*, 8(3), 338–353.

Jang, J.-S. R., Sun, C.-T., & Mizutani, E. (1997). *Neuro-Fuzzy and Soft Computing*. Prentice Hall.

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.

Ricci, F., Rokach, L., & Shapira, B. (2011). *Introduction to Recommender Systems Handbook*. Springer.

Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix Factorization Techniques for Recommender Systems. *Computer*, 42(8), 30-37.
<https://doi.org/10.1109/MC.2009.263>

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer.