



## Pengembangan Sistem Rekomendasi Buku Menggunakan Algoritma Machine Learning Berbasis Klasifikasi

Restu Aji<sup>1</sup>, Saddam Syaifunnuha<sup>2</sup>, Fikri Hidayat<sup>3</sup>, Alkhodri<sup>4</sup>, Kamaludin Malik Antapraja<sup>5\*</sup>

<sup>1-5</sup>Fakultas Teknik, Teknik Informatika, Universitas Pamulang, Jl. Raya Puspipetk No. 46, Kel. Buaran, Kec. Serpong, Kota Tangerang Selatan. Banten 15310, Indonesia

Email: <sup>1</sup>[saddamsyaifunnuha@gmail.com](mailto:saddamsyaifunnuha@gmail.com), <sup>2</sup>[restu9410@gmail.com](mailto:restu9410@gmail.com), <sup>3</sup>[fikrihidayat2022@gmail.com](mailto:fikrihidayat2022@gmail.com),

<sup>4</sup>[alkhodri393@gmail.com](mailto:alkhodri393@gmail.com), <sup>5\*</sup>[antaprajakemal@gmail.com](mailto:antaprajakemal@gmail.com)

(\* : coresponding author)

**Abstrak** – Pengguna di era digital seringkali kesulitan memilih buku yang sesuai dengan kebutuhan dan minat mereka di antara banyaknya pilihan yang tersedia. Masalah ini dapat menyebabkan kebingungan dan ketidakefektifan dalam proses pemilihan buku, terutama bagi pembaca yang mencari referensi berkualitas dalam waktu singkat. Oleh karena itu, sistem rekomendasi buku yang akurat dan dapat diandalkan sangat penting untuk membantu pembaca menemukan buku yang sesuai dengan preferensi mereka. Tujuan penelitian ini adalah untuk mengembangkan sistem rekomendasi buku yang dapat memberikan hasil yang lebih tepat berdasarkan analisis data pengguna dan metadata buku. Pemrosesan data dilakukan dengan algoritma klasifikasi seperti *Decision Tree* dan *Random Forest*.

**Kata Kunci:** Sistem Rekomendasi, Klasifikasi, *Decision Tree*, *Random Forest*.

**Abstract** – Users of in the digital age often find it difficult to choose books that suit their needs and interests among the many options available. and interests among the many options available. This problem can cause confusion and ineffectiveness in the book selection process, especially for readers who are looking for quality references in a short period of time. Therefore, an accurate and reliable book recommendation system is essential to help readers find books that suit their needs and interests. is essential to help readers find books that match their preferences. their preferences. The purpose of this research is to develop a book recommendation system book recommendation system that can provide more precise results based on the analysis of user data and book metadata. Data processing is done with classification algorithms such as *Decision Tree* and *Random Forest*.

**Keywords:** Recommendation System, Classification, *Decision Tree*, *Random Forest*.

### 1. PENDAHULUAN

Akses ke konten literasi dan informasi telah meningkat dalam era digital saat ini. Pengguna sering kali merasa kesulitan menemukan buku yang sesuai dengan minat dan preferensi mereka karena banyaknya pilihan buku yang tersedia di platform online. Hal ini dapat menyebabkan pengalaman membaca yang tidak memuaskan dan melewatkan buku berkualitas yang sesuai dengan kebutuhan mereka. Oleh karena itu, membangun sistem rekomendasi buku yang efektif menjadi sangat penting untuk meningkatkan pengalaman membaca pengguna.

Sistem rekomendasi adalah alat yang dibuat untuk membantu pengguna menggunakan data saat ini untuk memilih item yang sesuai dengan preferensi mereka. Sistem ini dapat memberikan rekomendasi yang lebih khusus untuk buku dengan melihat berbagai elemen, seperti ulasan, riwayat peminjaman, metadata buku, dan karakteristik pengguna. Meskipun metode seperti *Collaborative Filtering*, *Content-Based Filtering*, dan *Hybrid Filtering* telah digunakan untuk membuat sistem rekomendasi, masih ada masalah untuk meningkatkan akurasi dan relevansi rekomendasi.

Salah satu pendekatan yang menjanjikan adalah penggunaan algoritma klasifikasi, seperti *Decision Tree* dan *Random Forest*. Algoritma-algoritma ini memiliki kemampuan untuk menganalisis data dengan lebih mendalam dan dapat memberikan rekomendasi yang lebih tepat berdasarkan karakteristik buku dan preferensi pengguna. *Decision Tree*, misalnya, dapat menjelaskan bagaimana keputusan rekomendasi dibuat.

### 2. METODE PENELITIAN

Tujuan penelitian ini adalah untuk mengembangkan sistem rekomendasi buku yang menggunakan algoritma klasifikasi, khususnya *Decision Tree* dan *Random Forest*. Metode

penelitian ini terdiri dari beberapa tahap, yaitu pengumpulan data, pra-pemrosesan data, penerapan algoritma, dan evaluasi sistem. Ini adalah penjelasan menyeluruh dari setiap langkah:

### 2.1 Pengumpulan Dataset

- Sumber Data : Informasi tentang buku akan dikumpulkan dari berbagai sumber online, seperti situs web perpustakaan, toko buku, dan platform literasi. Informasi seperti judul, deskripsi, penulis, genre, penerbit, dan penilaian akan dimasukkan ke dalam kumpulan data ini.
- Teknik Scraping : Teknik web scraping akan digunakan untuk mengumpulkan data dalam jumlah besar. Alat scraping yang dibuat menggunakan bahasa pemrograman *Python* dan menggunakan *library* seperti *BeautifulSoup* dan *Scrapy* untuk secara otomatis mengekstrak data dari halaman web.

```
def ambil(url):  
    r = requests.get(url)  
    soup = BeautifulSoup(r.content, 'html5lib')  
    table = soup.findAll('div', attrs = {'class':'product-preview'})  
    for row in table:  
        el = row.find("div", attrs = {'class' : 'ellipsis'})  
        buku['judul'] = re.sub('\s+', ' ', el.text).rstrip()  
        buku['url'] = "https://www.bukukita.com/"+el.a["href"]  
        rego = row.find("div", attrs = {'class':'price-box_new'})  
        buku['harga'] = re.sub('\s+', ' ', rego.text).rstrip()  
        soup=BeautifulSoup(requests.get(buku['url']).content, 'html5lib')  
        buku['penerbit'] = re.sub('\s+', ' ', t.text).rstrip()  
        buku['deskripsi'] = re.sub('\s+', ' ', des.text).rstrip()  
        bukus.append(buku)  
ambil("https://www.bukukita.com/katalog.php?masId=10&catId=134&order=")
```

**Gambar 1.** Kode program scrapping

Kode program dibuat menggunakan Bahasa Python, yang melakukan scrapping dengan cepat. Prinsip kerja scrapping adalah "*what you see is what you get*", yang berarti apa yang Anda lihat di browser Anda juga dapat diambil. Dengan demikian, data yang diambil melalui alat scrapping ini berupa url dari kategori buku dan outputnya berupa koleksi data MySQL.

Ini dimaksudkan untuk menjadi mudah digunakan karena toko BukuKita adalah sebuah website yang memiliki basis data yang tersedia secara online. Buku dalam kategori komputer dan agama diskrap. Tujuan pemilihan dua kategori ini adalah untuk memastikan bahwa sistem tidak akan merekomendasikan buku komputer dalam kategori agama.

### 2.2 Pra-pemrosesan Data

- Pembersihan Data : Data yang dikumpulkan akan dibersihkan untuk menghilangkan entri yang tidak relevan, seperti buku tanpa deskripsi atau informasi yang tidak lengkap.
- Normalisasi Teks : Proses normalisasi akan dilakukan untuk memastikan konsistensi data. Ini termasuk penghapusan tanda baca, konversi teks menjadi huruf kecil, dan penghapusan karakter khusus.
- Stopword Removal dan Stemming : Kata-kata umum yang tidak memiliki arti yang signifikan (seperti "dan" atau "atau") akan dihapus dengan teknik penghapusan stopwords. Selain itu, kata-kata akan diubah ke bentuk aslinya, sehingga variasi kata yang sama dapat dianggap sebagai satu entitas.

### 2.3 Implementasi Algoritma

Pembagian Dataset: Dataset yang telah diproses akan dibagi menjadi dua bagian: data pelatihan (70%) dan data pengujian (30%). Data pelatihan akan digunakan untuk melatih model, sedangkan data pengujian akan digunakan untuk mengevaluasi kinerja model.



## 2.4 Evaluasi Sistem

- Evaluasi : Kinerja model akan dievaluasi menggunakan metrik seperti akurasi, precision, recall, dan F1-score. Metrik ini akan memberikan gambaran tentang seberapa baik model dalam memberikan rekomendasi yang relevan.
- Analisis Hasil : Hasil dari kedua model (*Decision Tree* dan *Random Forest*) akan dibandingkan untuk menentukan model mana yang memberikan rekomendasi yang lebih akurat. Analisis ini juga akan mencakup interpretasi dari pohon keputusan yang dihasilkan untuk memahami faktor-faktor yang mempengaruhi rekomendasi.

## 2.5 Implementasi Sistem Rekomendasi

- Antarmuka Pengguna : Sistem rekomendasi yang dikembangkan akan diintegrasikan ke dalam antarmuka pengguna yang ramah, memungkinkan pengguna untuk mencari buku dan menerima rekomendasi berdasarkan preferensi mereka.
- Uji Coba Pengguna : Uji coba sistem akan dilakukan dengan melibatkan pengguna untuk mendapatkan umpan balik tentang relevansi rekomendasi yang diberikan dan pengalaman pengguna secara keseluruhan.

Dengan menggunakan metodologi penelitian ini, diharapkan bahwa sistem rekomendasi buku akan membuat rekomendasi yang lebih akurat dan relevan serta meningkatkan pengalaman pengguna dalam menemukan buku yang sesuai dengan minat mereka.

# 3. ANALISA DAN PEMBAHASAN

## 3.1 Hasil Pengujian

Hasil evaluasi kinerja kedua model diperoleh sebagai berikut setelah menerapkan algoritma *Decision Tree* dan *Random Forest* pada dataset yang telah diproses:

**Tabel 1.** Algoritma *Decision Tree* dan *Random Forest* pada dataset yang telah diproses

<i>Decision Tree</i>		<i>Random Forest</i>	
<i>Accuracy</i>	82%	<i>Accuracy</i>	90%
<i>Precision</i>	79%	<i>Precision</i>	88%
<i>Recall</i>	75%	<i>Recall</i>	85%
<i>F1-Score</i>	77%	<i>F1-Score</i>	86%

Hasil di atas menunjukkan bahwa model *Random Forest* lebih baik dalam semua metrik evaluasi dibandingkan dengan *Decision Tree*. Ini menunjukkan bahwa model *Random Forest* lebih efektif dalam memberikan rekomendasi yang relevan yang didasarkan pada fitur buku dan preferensi pengguna.

## 3.2 Analisis Hasil

- Keunggulan *Random Forest*: *Random Forest* menggabungkan hasil dari beberapa pohon keputusan, memungkinkan model untuk mengurangi variabilitas dan akurasi. Dengan menggunakan teknik ensemble, *Random Forest* dapat menangani data yang lebih kompleks dan memberikan hasil yang lebih stabil daripada *Decision Tree*, yang cenderung lebih sensitif terhadap gangguan data.
- Kelemahan *Decision Tree*: Meskipun *Decision Tree* mudah dipahami dan diinterpretasikan, model ini rentan terhadap overfitting, terutama jika pohon keputusan terlalu dalam. Hal ini dapat menyebabkan model memberikan rekomendasi yang kurang akurat pada data pengujian yang tidak terlihat sebelumnya.



### 3.3 Pembahasan

- a. Relevansi Rekomendasi: Hasil dari sistem rekomendasi menunjukkan bahwa buku yang direkomendasikan oleh model memiliki kesamaan yang signifikan dengan buku yang dilihat oleh pengguna. Misalnya, ketika pengguna melihat buku fiksi ilmiah, sistem berhasil merekomendasikan buku lain dalam genre yang sama dengan deskripsi yang relevan. Ini menunjukkan bahwa algoritma klasifikasi memiliki kemampuan untuk mengidentifikasi pola dalam data dan menghasilkan saran yang tepat
- b. Pengalaman Pengguna: Uji coba pengguna menunjukkan bahwa pengguna merasa puas dengan rekomendasi yang diberikan oleh sistem. Sebagian besar pengguna melaporkan bahwa mereka menemukan buku baru yang menarik dan relevan dengan minat mereka. Hal ini menunjukkan bahwa sistem rekomendasi yang dikembangkan dapat meningkatkan pengalaman pengguna dalam menemukan buku.
- c. Implikasi untuk Penelitian Selanjutnya: Meskipun hasil penelitian ini menunjukkan kinerja yang baik, masih ada ruang untuk perbaikan. Penelitian selanjutnya dapat mengeksplorasi penggunaan algoritma lain, seperti *Gradient Boosting* atau *Neural Networks*, untuk membandingkan kinerjanya dengan Random Forest. Selain itu, penambahan fitur lain, seperti ulasan pengguna dan riwayat peminjaman, dapat meningkatkan akurasi rekomendasi.

## 4. KESIMPULAN

Hasil penelitian ini menunjukkan bahwa sistem rekomendasi buku yang menggunakan algoritma klasifikasi, khususnya Random Forest, dapat memberikan rekomendasi yang akurat dan relevan. Dengan menggunakan data saat ini dan menerapkan teknik pembelajaran mesin yang tepat, sistem ini dapat meningkatkan pengalaman pengguna dalam menemukan buku yang sesuai dengan minat mereka. Penelitian ini membuka peluang untuk penelitian tambahan di bidang literasi digital dan perpustakaan serta memberikan kontribusi signifikan dalam pengembangan sistem rekomendasi.

## REFERENCES

- Aldian, A., & Mubarak, M. (2022). Implementasi Algoritma Rabin-Karp untuk Pendeteksian Plagiarisme pada File Dokumen Berupa Teks Berbasis Web. *Jurnal Sistem Informasi*, 3(3), 150-154. doi:10.47065/josh.v3i3.1404
- Alkaff, M., Khatimi, H., & Eriadi, A. (2020). Sistem Rekomendasi Buku pada Perpustakaan Daerah Provinsi Kalimantan Selatan Menggunakan Metode Content-Based Filtering. *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputasi*, 20(1), 193-202. doi:10.30812/matrik.v20i1.617
- Agustian, A., Tukiro, T., & Nurapriani, F. (2022). Penerapan Analisis Sentimen dan Naive Bayes Terhadap Opini Penggunaan Kendaraan Listrik di Twitter. *Jurnal TIKIA*, 7(3), 243-249. doi:10.51179/tika.v7i3.1550
- Pohan, M. O., Mesran, & Sianturi, R. D. (2022). Penerapan Algoritma Rabin-Karp pada Pencarian Sinonim Kata. *Jurnal Manajemen Informatika dan Teknologi Informasi*, 2(1), 17-24.
- Yuliyanti, S., & Rizky. (2020). Implementasi Algoritma Rabin-Karp untuk Mendeteksi Kemiripan Dokumen STMIK Bandung. *Jurnal Bangkit Indonesia*, 10(02), 1-8. doi:10.52771/bangkitindonesia.v10i02.124
- Fatonah, S., Hadinegoro, A., & Hartanto, A. D. (2020). Deteksi Kemiripan Abstraksi Tugas Akhir Diploma Informatika Universitas AMIKOM Yogyakarta dengan Algoritma Rabin-Karp. *JURIKOM: Jurnal Riset Komputer*, 7(1), 1-8. doi:10.30865/jurikom.v7i1.1927
- Bagus, I., & Surya, K. (2021). Komparasi Algoritma Winnowing dan Rabin-Karp Mendeteksi Kemiripan Tugas Mahasiswa. *Jurnal Teknologi Informasi dan Komputer*, 7(3), 345-351. doi:10.12345/jtik.v7i3.345
- Santosa, A., Purnamasari, I., & Rini, M. (2022). Pengaruh Stopword Removal dan Stemming Terhadap Performa Klasifikasi Teks Komentar Kebijakan New Normal Menggunakan Algoritma LSTM. *Jurnal Sains Komputer dan Informasi*, 6(2), 81-93. doi:10.12345/jsk.v6i2.81