

Sistem Rekomendasi Buku Fiksi Indonesia Menggunakan Metode *Content Based Filtering* (Studi Kasus : Badan Perpustakaan Dan Kearsipan Provinsi Jawa Timur Surabaya)

Anggi Yhurinda Perdana Putri¹, Zendi Asriel Adrian Jaya², Ruli Utami³

Institut Teknologi Adhi Tama Surabaya¹²³
e-mail: zendiasriel1@gmail.com

ABSTRACT

A book is a set of bound sheets of paper. Books are not bound sheets of paper but can also be from other materials that are made into one. Inside the sheets of paper, some contain writing, pictures, and some are empty. Each sheet in a book is called a page. Prose or fiction means a narrative work that tells something fictional, imaginary not based on reality or can also mean a reality that is born based on fantasy. With the development of information technology, several systems can support decisions to choose books; including recommendation systems. The recommendation system can help book readers determine which fiction books to read. The recommendation system uses content-based filtering techniques and content-based filtering techniques. The system will display 5 users and 102 Indonesian fiction books. By performing the process of user profile similarity and word similarity, the recommendation system obtains the maximum value for Indonesian fiction books. After getting the recommended results, researchers tested the recommendation system with precision and got a score of 54%.

Kata Kunci: *recommendation system; content-based filtering; fiction; precision.*

ABSTRAK

buku merupakan sekumpulan lembaran kertas yang terjilid, dalam lembaran tersebut berisi tulisan maupun kosong. Buku merupakan kumpulan kertas atau bahan lainnya kemudian dijadikan satu dan berupa tulisan maupun gambar. Setiap lembaran dalam buku disebut halaman. Prosa atau fiksi memiliki arti sebuah karya naratif yang menceritakan sesuatu yang bersifat rekaan, khayalan, tidak berdasarkan kenyataan atau dapat juga berarti suatu kenyataan yang yang lahir berdasarkan khayalan. Dengan berkembangnya teknologi informasi ini terdapat beberapa sistem yang dapat mendukung keputusan diantaranya adalah sistem rekomendasi. Sistem rekomendasi dapat membantu pembaca buku untuk menentukan buku fiksi yang akan dibaca. Permasalahan ini dapat diatasi dengan menggunakan sistem rekomendasi dengan teknik *content based filtering*. Sistem rekomendasi ini menggunakan teknik penyaringan berbasis konten dan menampilkan 5 pengguna dan 102 buku fiksi Indonesia. Dengan melakukan proses kesamaan profil pengguna dan kesamaan *similarity* kata, sistem rekomendasi memperoleh nilai maksimal untuk buku fiksi Indonesia. Setelah mendapat hasil yang direkomendasikan, diuji dengan *precision* dan mendapatkan nilai skor 54%.

Kata Kunci: *sistem rekomendasi; content based filtering; fiksi; precision.*

PENDAHULUAN

Berdasarkan permendiknas no. 11 tahun 2005 adalah buku merupakan acuan wajib untuk digunakan di satuan Pendidikan dasar dan menengah atau perguruan tinggi yang memuat materi pembelajaran dalam rangka untuk meningkatkan penguasaan ilmu pengetahuan, peningkatan kepekaan, kemampuan Kesehatan dan teknologi yang disusun berdasarkan standar nasional Pendidikan. Isi buku didapat diperoleh dari banyak cara misalnya dari sebuah hasil penelitian, pengamatan, atau dari imajinasi seorang pengarang yang bisa disebut fiksi. Salah satu tempat terdapat koleksi buku untuk diserap informasinya terdapat di perpustakaan. Pengunjung seringkali mengalami kesulitan dalam mencari dan menemukan buku yang sesuai dengan preferensi yang diinginkan oleh pengunjung karena kurangnya informasi mengenai berbagai jenis buku yang ada pada perpustakaan. [1]

Dengan adanya perpustakaan pengunjung bisa mendapat informasi yang diperlukan, salah satu pelayanan perpustakaan terdapat di Dinas Perpustakaan dan Kearsipan Provinsi Jawa Timur yang berada di jalan menur pumpungan no 32 kecamatan sukolilo kota Surabaya. Dikarenakan banyaknya pengunjung yang meminjam buku dengan kategori fiksi indonesia yang tercatat dalam data desember 2022 sampai dari januari 2023 yang mempunyai persentase 70% dari keseluruhan pengunjung yang ada di Dinas Perpustakaan dan Kearsipan Provinsi Jawa Timur dan tidak ada nya sistem tentang sistem rekomendasi buku fiksi Indonesia disana. Terkadang pengunjung sewaktu melakukan kunjungan ke perpustakaan merasa kebingungan dan tidak mengetahui tujuan untuk memilih buku yang ingin dipinjam di perpustakaan tersebut, maka adanya sistem rekomendasi buku hadir untuk mengatasi permasalahan di Dinas Perpustakaan dan Kearsipan Provinsi Jawa Timur.

TINJAUAN PUSTAKA

Metode Content Based Filtering

Metode Content-Based Filtering merupakan metode dalam sistem rekomendasi yang mampu menghasilkan sebuah rekomendasi berdasarkan *content/atribut*. [2] Keuntungan dari pendekatan ini adalah pengguna mendapatkan wawasan tentang mengapa suatu item dianggap relevan untuk mereka, karena konten di setiap item diketahui dari representasinya[3]. Namun pendekatan ini juga mempunyai kelemahan, misalnya kenyataan bahwa pendekatan ini berfokus pada kemiripan kata kunci. Pendekatan ini tidak mampu menangkap hubungan yang lebih kompleks pada level semantik yang lebih dalam, berdasarkan pada berbagai jenis atribut yang berhubungan dengan objek terstruktur dari teks.[4] Kesamaan antara representasi dari pengguna dan representasi dari item akan didasarkan pada prinsip kedekatan yang menyatakan bahwa jarak dari dua deskripsi item secara langsung berkaitan dengan kesamaan mereka. [5] Dalam metode content based filtering ini terdapat Langkah-langkah agar mencapai sebuah hasil rekomendasinya, berikut merupakan Langkah-langkah nya:

1. Term Frequency–Inverse Document Frequency (TF–IDF) merupakan suatu metode untuk memberikan bobot hubungan suatu kata (term) terhadap dokumen. Metode ini merupakan tata penting cara untuk mengimplementasi Cosine-Similarity ke dalam sistem rekomendasi karena tingkat akurasi sangat tinggi dan efisien.[6]

Rumus perhitungan :

$$TF-IDF: W(t, d) = tf(t, d) * idf(t) \dots \dots \dots (1)$$

2. Cosine similarity merupakan perhitungan yang sering digunakan untuk menghitung kemiripan diantara item – item.[7] Secara umum fungsi dari similarity adalah fungsi yang menerima dua buah objek berupa bilangan riil (0 dan 1) dan mengembalikan nilai kemiripan (similarity) antara kedua objek tersebut berupa bilangan riil.

Fungsi cosine similarity item A dan item B sebagai berikut :

$$sim(A, B) = \frac{n(A \cap B)}{\sqrt{n(A)n(B)}} \dots \dots \dots (2)$$

3. Precision merupakan sebuah pengujian dasar yang biasanya digunakan untuk menguji sistem rekomendasi. TP atau true positive berarti sistem yang menghasilkan informasi relevan, sedangkan FP atau false positive berarti sistem yang menghasilkan seluruh item yang tersedia. Dengan kata lain, precision merupakan pengujian yang membandingkan hasil dari sistem dengan informasi yang diminta oleh pengguna. Precision dapat dihitung menggunakan persamaan. [8] $TP/FP \times 100\% \dots \dots \dots (3)$

METODE

1. TF-IDF

Term Frequency–Inverse Document Frequency (TF–IDF) Term Frequency–Inverse Document Frequency (TF–IDF) merupakan suatu metode untuk memberikan bobot hubungan suatu kata (term) terhadap dokumen. Metode ini merupakan tata penting cara untuk mengimplementasi Cosine-Similarity ke dalam sistem rekomendasi karena tingkat akurasi sangat tinggi dan efisien. TF–IDF dapat mengevaluasi beberapa kalimat yang dianggap penting dalam suatu kumpulan data yang besar.

2. Cosine Similarity

Cosine similarity merupakan perhitungan yang sering digunakan untuk menghitung kemiripan diantara *item – item*. Secara umum fungsi dari *similarity* adalah fungsi yang menerima dua buah objek berupa bilangan riil (0 dan 1) dan mengembalikan nilai kemiripan (*similarity*) antara kedua objek tersebut berupa bilangan riil. Metode ini digunakan untuk menghitung nilai dari kosinus sudut antara dua vektor dan digunakan untuk mengukur kemiripan.

3. Precision

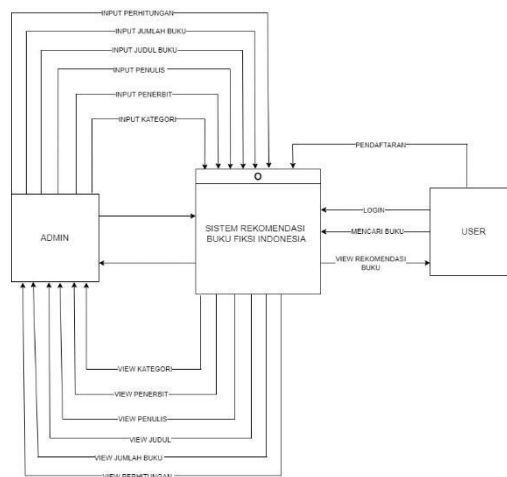
Precision merupakan sebuah pengujian dasar yang biasanya digunakan untuk menguji sistem rekomendasi. TP atau *true positive* berarti sistem yang menghasilkan informasi relevan, sedangkan FP atau *false positive* berarti sistem yang menghasilkan seluruh item yang tersedia. Dengan kata lain, *precision* merupakan pengujian yang membandingkan hasil dari sistem dengan informasi yang diminta oleh pengguna.

4. Recall

Recall adalah perbandingan antara jumlah *records* relevan yang didapatkan dengan jumlah keseluruhan *records* relevan pada *database*. [10]

HASIL DAN PEMBAHASAN

1. Perancangan DFD Sistem



Gambar 1. DFD Sistem rekomendasi Buku

Gambar 1 di atas menjelaskan alur sistem yang akan dibangun nantinya :

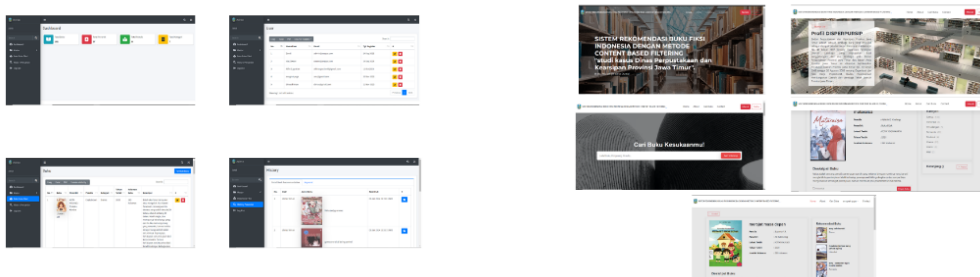
- Hak akses admin di dalam hanya melakukan *input* data buku dan melihat hasil rekomendasi.

Admin bisa mendapatkan informasi tentang *item*, *rating item*, data user tersebut.

- b. *User* dapat memilih *item* dan melihat hasil rekomendasi yang diinginkan dan mendapatkan hasil dari peratingan dan *item* buku yang direkomendasikan dari sistem.
- c. Pengunjung yang tanpa melakukan login hanya bisa melakukan pencarian *item*.

2. Tampilan Antarmuka (user interface)

Pada tampilan antarmuka (*user interface*) penulis disini mencoba untuk membagi dua (2) tampilan atau login diantaranya adalah : tampilan untuk *user* dan tampilan admin sendiri seperti pada gambar 2.



Gambar 2. Interface Sistem Rekomendasi

3. Hasil rekomendasi

Hasil rekomendasi dilakukan berdasarkan dengan banyak nya kata yang sama dengan buku yang telah dipilih oleh 5 user dan sebanyak 102 data buku yang ada didalam database, setelah user mengklik buku tersebut, maka sistem baru memunculkan hasil dari rekomendasi buku tersebut. Dan dibawah ini merupakan contoh dari salah satu hasil rekomendasi dari *user*.

Tabel 1. Hasil rekomendasi dari user

Hasil rekomendasi							
Nilai cosine similarity	no	judul_buku	penulis	penerbit	tahun terbit	kategori	halaman
0.2038098661	1	Rekindled	Dhamala Shobita	Yogyakarta : Laksana	2020	drama	220
0.1874158282	2	Little studygrammer	Mutiara Sya'bani	Bandung : Mizan	2020	drama	84
0.1849000654	3	Merayakan kehilangan	Brian Khrisna	Jakarta : Media kita	2021	romantis	242
0.1849000654	4	Lolita	Vladimir Nabokov	Yogyakarta : Diva Press	2021	Romantis	420
0.1689815924	5	Bising	Kurniawan Gunadi	Bandung : Mizan	2021	motivasi	164

Hasil TF-IDF

Berikut merupakan hasil perhitungan TF-IDF Sebagian contoh dari buku ke 1 dalam rekomendasi tabel diatas.

Tabel 1. Hasil perhitungan TF-IDF

Data buku 1			
no	kata	DQ	D1
1	mataraisa	0.005068313851	0
2	rekindled	0	0.005068313851
3	raisa	0.005068313851	0
4	adalah	0.005068313851	0
5	seorang	0.005068313851	0
6	penulis	0.005068313851	0
7	perempuan	0.005068313851	0
8	cantik	0.005068313851	0

Precision

Berikut ini salah satu merupakan pengujian dari rekomendasi diatas dalam menggunakan *precision* dan mempunyai hasil dari yang sudah diuji :

$$\frac{9}{9+10} \times 100\% = 90\%$$

9 merupakan hasil rasio dari *true positive* atau buku yang relevan dengan kategori buku yang sudah dipilih oleh user dan angka 10 merupakan *false positive* atau hasil dari rekomendasi yang ditampilkan dalam sistem untuk user. Untuk hasil dari precision hasil rekomendasi user 1 mempunyai angka precision yakni 90%.

Hasil Recall

Setelah dilakukan hasil precision diatas lalu menghitung recall dari hasil rekomendasi diatas. berikut merupakan hasil perhitungan dari recall :

User memilih kategori buku petualangan, romantis, dan drama. Hasil rekomendasi yang dihasilkan oleh sistem terdapat pada tabel diatas dalam hasil rekomendasi 7 buku drama, 2 buku romantis dan 1 buku motivasi. Untuk total yang ada dalam kategori buku yang disukai user ada 83 buku yang ada pada dalam database.

Hasil *recall* :

$$\frac{9}{9+74} = 0,108433$$

Untuk hasil yang didapat dari perhitungan recall adalah 0.120481.

KESIMPULAN

Berdasarkan penelitian yang dilakukan dan uraian pembahasan dapat diambil beberapa kesimpulan bahwa dari uji coba sistem yang dilakukan, sistem dapat berjalan sesuai dengan

fungsi-fungsi sesuai kegunaan yang dibuktikan dari pengujian 5 user menggunakan precision maka hasil rekomendasi mempunyai nilai precision sebesar 54%. Serta proses *content based filtering* dapat menghasilkan rekomendasi yang dapat membantu user dalam memilih buku sesuai dengan ketertarikan user.

DAFTAR PUSTAKA

- [1] Ardiansyah, R., Ari Bianto, M., & Saputra, B. D. (2023). Sistem Rekomendasi Buku Perpustakaan Sekolah menggunakan Metode Content-Based Filtering. *Jurnal CoSciTech (Computer Science and Information Technology)*, 4(2), 510–518. <https://doi.org/10.37859/coscitech.v4i2.5131>
- [2] Ayu, D., Safitri, N., Helilintar, R., & Wahyuniar, L. S. (2021). Sistem Rekomendasi Skincare Menggunakan Metode Content-Based Filtering dan Algoritma Apriori. *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, 5(2), 242–248. <https://proceeding.unpkediri.ac.id/index.php/inotek/article/view/1136>
- [3] Huda, A. A., Fajarudin, R., & Hadinegoro, A. (2022). Sistem Rekomendasi Content-based Filtering Menggunakan TF-IDF Vector Similarity Untuk Rekomendasi Artikel Berita. *Building of Informatics, Technology and Science (BITS)*, 4(3), 1679–1686. <https://doi.org/10.47065/bits.v4i3.2511>
- [4] Arfisko, H. H. (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>
- [5] Nastiti, P. (2019). Penerapan Metode Content Based Filtering Dalam Implementasi Sistem Rekomendasi Tanaman Pangan. *Teknika*, 8(1), 1–10. <https://doi.org/10.34148/teknika.v8i1.13>
- [6] Mulyada, A. H., Studi, P., Informatika, T., Informasi, F. T., & April, U. S. (2022). Penerapan metode content-based filtering dalam sistem rekomendasi tempat wisata di sumedang. 1(November), 229–234. [4] M. Nur *et al.*, “Evaluation of Novel Integrated Dielectric Barrier Discharge Plasma as Ozone Generator,” *Bull. Chem. React. Eng. Catal.*, vol. 12, no. 1, p. 24, Apr. 2017.
- [7] Lindang, D. N., Muniar, A. Y., Halid, A., Muhajirin, M., & Amiruddin, A. (2022). Sistem Penentuan Kemiripan Antar Skripsi Menggunakan Metode *Cosine Similarity* Pada Perpustakaan. *Sntei*, 321–324
- [8] Fajriansyah, M., Adikara, P. P., & Widodo, A. W. (2021). Sistem Rekomendasi Film Menggunakan Metode Content Based Filtering. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 5(6), 2188–2199. <http://e-journal.uajy.ac.id/20600/>
- [9] Ginting, E. T. B., & Pratama, I. (2023). Sistem Rekomendasi Jurusan SMK Menggunakan Metode Content-Based Filtering Di Kabupaten Sleman. *Jurnal Sains Dan Teknologi (JSIT)*, 3(2), 291–300. <https://doi.org/10.47233/jsit.v3i2.954>
- [10] Oeyliawan, R. F., & Gunawan, D. (2018). Aplikasi Rekomendasi Buku Pada Katalog Perpustakaan Universitas Multimedia Nusantara Menggunakan Vector Space Model. *Jurnal ULTIMATICS*, 9(2), 97–105. <https://doi.org/10.31937/ti.v9i2.639>