

**IMPLEMENTASI SISTEM REKOMENDASI BUKU BERBASIS
KONTEN MENGGUNAKAN *VECTOR SPACE MODEL* DAN
SIMILARITY MEASURE UNTUK PENGELOLAAN ANTREAN DAN
SIRKULASI DI PERPUSTAKAAN BERJALAN**

TUGAS AKHIR



**CHIKA HUMAIRA ABIDATILLAH
1202001005**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN ILMU KOMPUTER
UNIVERSITAS BAKRIE
JAKARTA
2024**

**IMPLEMENTASI SISTEM REKOMENDASI BUKU BERBASIS
KONTEN MENGGUNAKAN *VECTOR SPACE MODEL* DAN
SIMILARITY MEASURE UNTUK PENGELOLAAN ANTREAN DAN
SIRKULASI DI PERPUSTAKAAN BERJALAN**

TUGAS AKHIR

**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana pada program studi
Infomatika**



**CHIKA HUMAIRA ABIDATILLAH
1202001005**


**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN ILMU KOMPUTER
UNIVERSITAS BAKRIE
JAKARTA
2024**

HALAMAN PERNYATAAN ORISINALITAS

**Tugas Akhir ini adalah hasil karya saya sendiri,
dan semua sumber baik yang dikutip maupun dirujuk
telah saya nyatakan dengan benar.**

Nama : Chika Humaira A

NIM : 1202001005

Tanda Tangan : 

Tanggal : 15 Agustus 2024

HALAMAN PENGESAHAN


Tugas Akhir ini diajukan oleh:

Nama : Chika Humaira Abidatillah
NIM : 1202001005
Program Studi : Informatika
Fakultas : Teknik dan Ilmu Komputer
Judul : Implementasi Sistem Rekomendasi Buku Berbasis Konten Menggunakan *Vector Space Model* dan *Similarity Measure* untuk Pengelolaan Antrean dan Sirkulasi di Perpustakaan Berjalan

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer (S.Kom) pada Program Studi Informatika Fakultas Teknik dan Ilmu Komputer Universitas Bakrie.

DEWAN PENGUJI

Pembimbing : Guson Prasamuarso Kuntarto, ST, Msc


31.08.2024 19:03WIB
(Guson P. Kuntarto)

Penguji 1 : Irwan Prasetya Gunawan, S.T., M.Eng., PhD




Penguji 2 : Yusuf Lestanto, S.T., M.Sc


31.08.2024
(Yusuf Lestanto...)

Ditetapkan di : Jakarta

Tanggal : 30 Agustus 2024

UNGKAPAN TERIMA KASIH

Segala puji dan syukur dipanjatkan kehadirat Allah Swt. atas kelancaran dalam menyelesaikan tugas akhir ini. Penelitian berjudul “Implementasi Sistem Rekomendasi Buku Berbasis Konten Menggunakan *Vector Space Model* dan *Similarity Measure* untuk Pengelolaan Antrean dan Sirkulasi di Perpustakaan Berjalan” berlangsung selama enam bulan dengan dukungan dari banyak pihak. Oleh karena itu, saya mengucapkan terima kasih kepada pihak-pihak berikut:

1. Guson Prasamuarso Kuntarto, ST., M.Sc selaku dosen pembimbing tugas akhir yang senantiasa memberikan saran dan arahan terkait tugas akhir ini.
2. Kedua orang tua, Papa Sukanta dan Mama Fatmasari, kedua adik, serta seluruh anggota keluarga yang tiada hentinya mendukung, memotivasi, dan mendoakan kelancaran penelitian ini.
3. Teman-teman program studi Informatika 2020 Universitas Bakrie dan para sahabat yang telah menemani dan banyak membantu selama ini.
4. Kak Rachael Yuska, selaku pemilik dari Perpustakaan Berjalan, yang telah mengizinkan saya melakukan penelitian ini; serta Kak Nadia, Kak Farhan, dan Kak Yuli selaku pengurus dan anggota Perpustakaan Berjalan, yang bersedia diwawancara untuk keperluan penelitian.

Semoga dengan tugas akhir ini, penelitian dapat memberikan manfaat dan kontribusi bagi perkembangan ilmu pengetahuan, khususnya di bidang sistem rekomendasi buku.

Jakarta, 16 Agustus 2024



Chika Humaira Abidatillah

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI

Sebagai sivitas akademik Universitas Bakrie, saya yang bertanda tangan di bawah ini:

Nama : Chika Humaira Abidatillah
NIM : 1202001005
Program Studi : Informatika
Fakultas : Teknik dan Ilmu Komputer
Jenis Tugas Akhir : Implementasi Algoritma untuk Sistem Rekomendasi

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bakrie **Hak Bebas Royalti Noneksklusif (Non-exclusive Royalty-Free Right)** atas karya ilmiah saya yang berjudul:

“Implementasi Sistem Rekomendasi Buku Berbasis Konten Menggunakan *Vector Space Model* dan *Similarity Measure* untuk Pengelolaan Antrean dan Sirkulasi di Perpustakaan Berjalan”

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Bakrie berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta untuk kepentingan akademis.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 15 Agustus 2024



(Chika Humaira A)

**IMPLEMENTASI SISTEM REKOMENDASI BUKU BERBASIS
KONTEN MENGGUNAKAN *VECTOR SPACE MODEL* DAN
SIMILARITY MEASURE UNTUK PENGELOLAAN ANTREAN DAN
SIRKULASI DI PERPUSTAKAAN BERJALAN**

Chika Humaira Abidatillah

ABSTRAK

Perpustakaan Berjalan memiliki masalah stagnasi sirkulasi buku, di mana sejumlah buku tidak memiliki peminjam selanjutnya di daftar antrean. Untuk mengatasi ini, penelitian ini mengembangkan sistem rekomendasi buku berbasis konten yang bertujuan meningkatkan sirkulasi dengan merekomendasikan buku-buku relevan sesuai preferensi anggota perpustakaan. Penelitian ini memiliki tiga tujuan utama: (1) membangun *dataset* buku dengan delapan fitur, termasuk kode buku, judul buku, nama penulis, sinopsis, genre, jumlah halaman, riwayat peminjaman, serta *rating* buku yang diekstraksi dari situs Goodreads, menghasilkan 316 data buku dan 1.879 data peminjaman; (2) mengembangkan model sistem rekomendasi menggunakan teknik *Vector Space* dan metode perhitungan kemiripan seperti Term Frequency - Inverse Document Frequency (TF-IDF) dengan Cosine Similarity, Euclidean Similarity, dan Manhattan Similarity, serta Latent Semantic Indexing (LSI) dan Best Matching 25 (BM25). Hasil menunjukkan TF-IDF dengan Cosine Similarity memberikan presisi tertinggi, yaitu 78%; (3) mengukur dampak dari sistem rekomendasi ini terhadap masalah sirkulasi buku yang stagnan. Pengujian kuantitatif dan validasi kualitatif menunjukkan bahwa sistem ini berhasil merekomendasikan buku-buku yang sebelumnya kurang diminati, sehingga memiliki potensi besar dalam mengatasi masalah stagnasi sirkulasi.

Kata Kunci: sistem rekomendasi, sistem antrean, perpustakaan, sirkulasi buku, pemfilteran berbasis konten, *vector space model*, *similarity measure*

**IMPLEMENTATION OF CONTENT-BASED BOOK
RECOMMENDATION SYSTEM USING VECTOR SPACE MODEL
AND SIMILARITY MEASURE FOR QUEUE AND CIRCULATION
MANAGEMENT IN PERPUSTAKAAN BERJALAN**

Chika Humaira Abidatillah

ABSTRACT

Perpustakaan Berjalan faces an issue of stagnant book circulation, where certain books don't have the next borrower on the queue lists. To address this issue, this research developed a content-based book recommendation system aimed at improving book circulation by recommending relevant books based on members' borrowing preferences. This study has three main objectives: (1) to build a comprehensive book dataset with essential features such as book code, book title, author name, synopsis, genre, number of pages, borrowing history, and book ratings extracted from the Goodreads website. The resulting dataset consists of 316 book records and 1,879 borrowing records; (2) to develop a recommendation system model using Vector Space techniques and various similarity measurement methods such as Term Frequency - Inverse Document Frequency (TF-IDF) with Cosine Similarity, Euclidean Similarity, and Manhattan Similarity, as well as Latent Semantic Indexing (LSI) and Best Matching 25 (BM25). The testing results show that the TF-IDF with Cosine Similarity method provided the highest precision score in recommending books that match the preferences of library members, which is 78%; (3) to assess the impact of this recommendation system on the issue of stagnant book circulation. Quantitative testing using precision, recall, f1-score metrics and qualitative validation from the library's administrators indicate that the recommendation system successfully recommended books that were previously less popular, demonstrating significant potential in addressing stagnant circulation.

Keyword: recommender system, queueing system, library, book circulation, content-based filtering, vector space model, similarity measure

DAFTAR ISI

HALAMAN PERNYATAAN ORISINALITAS	i
HALAMAN PENGESAHAN	ii
UNGKAPAN TERIMA KASIH	iii
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	iv
ABSTRAK	v
ABSTRACT	vi
DAFTAR ISI	vii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR RUMUS	xiii
DAFTAR LAMPIRAN	xiv
DAFTAR AKRONIM	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	5
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	6
1.6 Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA	8
2.1 Penelitian Terkait	8
2.2 Perpustakaan Berjalan	12
2.3 Sistem Rekomendasi	14
2.3.1 Content-based Filtering	15
2.3.1.1 Vector Space Model (VSM)	15
2.3.1.2 TF-IDF	16
2.3.1.3 Analisis Semantik Menggunakan Ontologi dan Pengetahuan Ensiklopedis	17
2.3.1.4 Latent Semantic Indexing (LSI)	18
2.3.1.5 Best Matching 25 (BM25)	18

2.3.2	Collaborative Filtering	20
2.3.3.	Pendekatan Hybrid.....	21
2.4	Dataset.....	21
2.4.1	Data Mining.....	21
2.4.2	Data Acquisition	21
2.5	Perhitungan Kemiripan	22
2.5.1.	Cosine Similarity	22
2.5.2.	Pearson Similarity.....	24
2.5.3.	Manhattan Distance Similarity	25
2.6	Evaluasi	25
2.6.1	Precision.....	26
2.6.2	Recall.....	26
2.6.3.	F1-Score	27
BAB III METODOLOGI PENELITIAN.....		28
3.1	Kontribusi Penelitian.....	28
3.2	Kerangka Penelitian	28
3.3	Metode Penelitian.....	29
3.3.1	Tahapan Penyusunan Dataset	29
3.3.2	Implementasi Sistem Rekomendasi Buku dengan <i>Content-based Filtering</i> pada Perpustakaan Berjalan	30
3.4	Evaluasi Efektivitas Sistem Rekomendasi.....	31
3.5	Peralatan Penelitian.....	32
BAB IV HASIL DAN PEMBAHASAN.....		33
4.1	Tahapan Penyusunan Dataset	33
4.1.1	Akuisisi Data	33
4.1.2	Data Preprocessing	36
A.	Case Folding	36
B.	Tokenizing	37
C.	Filtering (Stopword Removal).....	38
D.	Stemming.....	39
4.2.	Implementasi <i>Vector Space Model</i> dan <i>Similarity Measure</i>	40

4.2.1. Ekstraksi Fitur dengan <i>Vector Space Model</i>	40
4.2.2. Perhitungan Kemiripan	41
4.2.3. Sistem Rekomendasi Buku	44
4.2.4. Tampilan Web	45
4.3. Evaluasi Efektivitas Sistem Rekomendasi	47
4.4. Diskusi dan Pembahasan	53
BAB V KESIMPULAN DAN SARAN	57
5.1 Kesimpulan	57
5.2 Saran	58
DAFTAR PUSTAKA	59
LAMPIRAN	64
Lampiran 1. Hasil Survei kepada Anggota Perpustakaan Berjalan	64
Lampiran 2. Tahapan Preprocessing Data	77
Lampiran 3. Data Buku yang Dikumpulkan dari Google Sheets	82
Lampiran 4. Transkrip Wawancara Sebelum Penelitian	82
Lampiran 5. Transkrip Wawancara untuk Evaluasi Sistem Rekomendasi	97

DAFTAR TABEL

Table 2. 1 Penelitian Terkait Implementasi Sistem Rekomendasi Buku Berbasis Konten Menggunakan <i>Vector Space Model</i> dan <i>Similarity Measure</i> untuk Pengelolaan Antrean dan Sirkulasi di Perpustakaan Berjalan	10
Table 2. 2 Syarat dan Ketentuan Peminjaman Buku di Perpustakaan Berjalan	13
Table 2. 3 Perhitungan Cosine Similarity	23
Tabel 4. 1 Precision Dari Setiap Metode	47
Tabel 4. 2 Rangkuman Nilai Perbandingan Skor Precision, Recall, dan F1 setiap metode	48
Tabel 4. 3 Hasil Evaluasi Skor Rekomendasi Buku untuk Buku-Buku dengan Stagnasi	51

DAFTAR GAMBAR

Gambar 2. 1 Flowchart Sistem Peminjaman Buku di Perpustakaan Berjalan	13
Gambar 2. 2 Lanskap Penelitian dalam Pengumpulan Data untuk Machine Learning (Roh et al., 2021)	22
Gambar 3. 1 Kerangka Penelitian Implementasi Sistem Rekomendasi Buku Berbasis Konten Menggunakan Vector Space Model dan Similarity Measure untuk Pengelolaan Antrean dan Sirkulasi di Perpustakaan Berjalan	28
Gambar 3. 2 Data Buku Perpustakaan	29
Gambar 3. 3 Tahapan Data Preprocessing	29
Gambar 3. 4 Flowchart Implementasi Sistem Rekomendasi Buku Berbasis Konten Menggunakan Vector Space Model dan Similarity Measure untuk Pengelolaan Antrean dan Sirkulasi di Perpustakaan Berjalan.....	30
Gambar 4. 1 Dataset Buku Perpustakaan Berjalan	33
Gambar 4. 2 Riwayat Peminjaman Buku Anggota Perpustakaan Berjalan	33
Gambar 4. 4 Dataset Publik Perpustakaan Berjalan	34
Gambar 4. 3 Daftar Buku di Perpustakaan Berjalan dengan Genre Terbanyak	34
Gambar 4. 5 Daftar Buku di Perpustakaan Berjalan dengan Penulis Terbanyak	35
Gambar 4. 6 Daftar Buku di Perpustakaan Berjalan dengan Penulis Terbanyak	35
Gambar 4. 7 Daftar Buku yang Paling Banyak Dipinjam di Perpustakaan Berjalan	35
Gambar 4. 8 Daftar Buku di Perpustakaan Berjalan dengan Antrean Terbanyak	36
Gambar 4. 9 Tahapan Case Folding.....	37
Gambar 4. 10 Hasil dari Proses Case Folding	37
Gambar 4. 11 Proses Penggabungan Beberapa Kolom di Dataset Menjadi Satu Kolom Baru Bernama ‘Merged’	37
Gambar 4. 12 Proses Tokenizing.....	38
Gambar 4. 13 Proses Filtering/Stopwords Removal.....	38
Gambar 4. 14 Hasil Filtering/ Stopwords Removal.....	39
Gambar 4. 15 Tahapan Stemming	39
Gambar 4. 16 Pembobotan data pada kolom “description” menggunakan TF-IDF	40

Gambar 4. 17 Hasil bobot yang diberikan TF-IDF pada setiap istilah pada dataset	41
Gambar 4. 18 Hasil perhitungan Cosine Similarity pada tiap buku dengan keseluruhan buku lainnya	41
Gambar 4. 19 Matrix Cosine Similarity	42
Gambar 4. 20 Hasil perhitungan Euclidean Similarity pada tiap buku dengan keseluruhan buku lainnya.....	42
Gambar 4. 21 Hasil perhitungan Manhattan Similarity pada tiap buku dengan keseluruhan buku lainnya.....	43
Gambar 4. 22 Hasil perhitungan LSI dan Cosine Similarity pada tiap buku dengan keseluruhan buku lainnya.....	43
Gambar 4. 23 Kode untuk Rekomendasi Buku	44
Gambar 4. 24 Rekomendasi Buku yang Dihasilkan	44
Gambar 4. 25 Hasil bobot pada setiap istilah dengan TF-IDF	45
Gambar 4. 26 Hasil perhitungan Cosine Similarity antar buku.....	45
Gambar 4. 27 Hasil Rekomendasi Buku untuk “The Traveling Cat Chronicles”	46
Gambar 4. 28 Hasil Rekomendasi Buku untuk “On Palestine”	46
Gambar 4. 29 Diagram Perbandingan Precision, Recall, dan F1-Score	50
Gambar 4. 30 Wawancara dengan Narasumber 1, admin Perpustakaan Berjalan	51
Gambar 4. 31 Wawancara dengan Narasumber 2, anggota Perpustakaan Berjalan	52
Gambar 4. 32 Perancangan Diagram Relasi Basis Data untuk Pengembangan Sistem Rekomendasi Buku di Perpustakaan Berjalan.....	55

DAFTAR RUMUS

Rumus 2. 1 TF	17
Rumus 2. 2 IDF.....	17
Rumus 2. 3 TF-IDF.....	17
Rumus 2. 4 Latent Semantic Indexing.....	18
Rumus 2. 5 BM25.....	19
Rumus 2. 6 Cosine Similarity.....	23
Rumus 2. 7 Pearson Similarity	24
Rumus 2. 8 Manhattan Distance.....	25
Rumus 2. 9 Euclidean Distance.....	25
Rumus 2. 10 Precision	26
Rumus 2. 11 Recall.....	27
Rumus 2. 12 F1-Score.....	27

DAFTAR LAMPIRAN

Lampiran 1. Hasil Survei kepada Anggota Perpustakaan Berjalan	64
Lampiran 2. Tahapan Preprocessing Data.....	77
Lampiran 3. Data Buku yang Dikumpulkan dari Google Sheets	81
Lampiran 4. Transkrip Wawancara	82
Lampiran 5. Transkrip Wawancara untuk Evaluasi Sistem Rekomendasi.....	109

DAFTAR AKRONIM

BCT	Biblioteca Comunale Terni
BM25	Best Matching 25
CBF	Content-based Filtering
CF	Collaborative Filtering
CSV	Comma-separated Value
CTR	Click-through Rates
FN	False Negative
FP	False Positive
LSI	Latent Semantic Indexing
MIND	Microsoft News Dataset
nDCG	Normalized Discounted Cumulative Gain
NLTK	Natural Language Toolkit
PCC	Pearson Correlation Coefficient
PHP	Hypertext Preprocessor
RAM	Random Access Memory
SQL	Structured Query Language
SVD	Singular Value Decomposition
TF-IDF	Term Frequency – Inverse Document Frequency
TN	True Negative
TP	True Positive
UNESCO	United Nations Educational, Scientific and Cultural Organization
VSM	Vector Space Model