

# Content-Based Filtering pada Sistem Rekomendasi Buku Informatika

*By Dziky Ridhwanullah*

## *Content-Based Filtering* pada Sistem Rekomendasi Buku Informatika

Dziky Ridhwanullah<sup>1</sup>, Yovita Kinanti Kumarahadi<sup>2</sup>, Bayu Dwi Raharja<sup>3</sup>

<sup>1</sup>Sistem Informasi, STMIK Sinar Nusantara

<sup>2</sup>Informatika, STMIK Sinar Nusantara

<sup>3</sup>Sistem Informasi Akuntansi, STMIK Sinar Nusantara

### ABSTRACT

*There are changes taking place in Indonesia's educational system, particularly in universities. The learning approach used in the transformation program is student-centered. A firm foundation in literacy is required for the application of this learning. A library is one of the amenities that students require. The library is an ideal resource for pupils to enhance their critical thinking skills. It's not always simple to find books in the library, though. Students could find it challenging to locate the books they seek because there are so many collections already in existence. One method for doing book searches is through the use of a recommendation system. Using content based filtering is one recommendation system algorithm. This study suggests a content based filtering algorithm-based book recommendation system to facilitate students' search for informatics book titles. TF-IDF and Cosine Similarity are used in a similarity search to find phrases and assign weights to them. The content-based filtering algorithm's research findings might suggest books depending on user-specified parameters. 92,5% accuracy is the average for this method.*

*Keywords: recommendation system, content based filtering, TF-IDF, Cosine Similarity*

### I. PENDAHULUAN

Di Indonesia, pembelajaran *Student-Center Learning* atau SCL sudah mulai diterapkan mulai tahun 2022. Pembelajaran ini sesuai dengan Standar Nasional Pendidikan Tinggi (SN-Dikti), yaitu karakteristik pembelajaran di perguruan tinggi berpusat pada mahasiswa (Sri Gunani Partiwati et al., 2023). Dalam SN-Dikti, disebutkan bahwa karakteristik tersebut ditunjukkan pada Capaian Pembelajaran Lulusan (CPL) yang diraih melalui proses pembelajaran mandiri dalam mencari dan menemukan pengetahuan. Dengan demikian, mahasiswa diharapkan untuk berperan aktif dalam meningkatkan kemampuannya secara mandiri.

Pemberlakuan pembelajaran ini menyebabkan terjadinya perubahan peran pendidik, yaitu menjadi perancang, fasilitator, dan motivator proses belajar. Salah satu fasilitas yang diperlukan mahasiswa adalah perpustakaan. Berisi koleksi buku dan pustaka lain, perpustakaan menjadi sarana yang tepat bagi mahasiswa dalam mengembangkan kemampuan berpikirnya. Namun, melakukan pencarian buku di perpustakaan tidak selalu menjadi hal yang mudah. Banyaknya koleksi yang ada menyebabkan mahasiswa bisa mengalami kesulitan untuk mencari buku yang dikehendaki.

Pencarian buku dapat dilakukan dengan berbagai cara, salah satunya adalah dengan menggunakan sistem rekomendasi. Cara kerja sistem ini yaitu mengolah data dari pengguna, untuk kemudian dijadikan acuan dalam mencari produk yang memiliki kemiripan tinggi. Terdapat beberapa metode pada sistem rekomendasi, antara lain yaitu *collaborative* (Nugroho & Rahayu, 2020) (Herny Februriyanti et al., 2021), *content based* (Mariani Widia Putri et al., 2022), dan *hybrid recommendation system* (Putri et al., 2024), (Hilmi Hidayat Arfisko & Agung Toto Wibowo, 2022).

*Collaborative filtering* memberikan rekomendasi berdasarkan perilaku atau kebiasaan pengguna, *content-based filtering* berdasarkan preferensi pengguna, dan *hybrid*

berdasarkan kombinasi keduanya. Diantara metode tersebut, *content-based filtering* paling cocok digunakan pada penelitian ini karena menggunakan preferensi pengguna untuk memberikan rekomendasi. Dalam hal ini, preferensinya adalah contoh buku yang akan digunakan sebagai bahan ajar.

Penelitian ini dilakukan pada perpustakaan di salah satu perguruan tinggi swasta di Jawa Tengah, batasan pada penelitian ini adalah terkait dengan buku-buku yang ada di program studi Informatika. Agar memudahkan mahasiswa mencari rekomendasi judul buku informatika maka dalam penelitian ini mengusulkan sistem rekomendasi buku menggunakan algoritma *content based filtering*. Metodenya adalah dengan mencari dan memberikan bobot term menggunakan TF-IDF dan kemudian mencari kemiripan kata dengan *Cosine Similarity*. Kemudian dilakukan analisis kualitatif terhadap hasil dari rekomendasi. Sehingga hasil analisis dapat membantu mahasiswa dalam melakukan pencarian judul buku sebagai bahan ajar.

## II. TINJAUAN PUSTAKA

Sistem pembelajaran di Indonesia khususnya perguruan tinggi mengalami perubahan. Perubahan ini sesuai dengan standar nasional yang telah ditetapkan oleh Dikti. Program perubahannya adalah karakteristik pembelajaran berpusat pada mahasiswa. Dengan adanya pemberlakuan pembelajaran ini menyebabkan terjadinya perubahan peran pendidik, yaitu menjadi perancang, fasilitator, dan motivator proses belajar. Salah satu fasilitas yang diperlukan mahasiswa adalah perpustakaan. Berisi koleksi buku dan pustaka lain, perpustakaan menjadi sarana yang tepat bagi mahasiswa dalam mengembangkan kemampuan berpikirnya. Namun, melakukan pencarian buku di perpustakaan tidak selalu menjadi hal yang mudah. Banyaknya koleksi yang ada menyebabkan mahasiswa mengalami kesulitan untuk mencari buku yang dikehendaki. Pencarian buku dapat dilakukan dengan berbagai cara, salah satunya adalah dengan menggunakan sistem rekomendasi.

Sistem rekomendasi adalah sistem yang akan memperkirakan informasi apa saja yang menarik untuk pengguna dan membantu pengguna dalam menentukan pilihan-pilihan terbaik lainnya yang memiliki kemiripan dengan kriteria (Nugroho & Rahayu, 2020). Ada beberapa metode dalam sistem rekomendasi seperti *collaborative* (Nugroho & Rahayu, 2020) (Herny Februariyanti et al., 2021), *content based* (Mariani Widia Putri et al., 2022), dan *hybrid recommendation system* (Putri et al., 2024). Diantara metode tersebut, *content-based filtering* paling cocok digunakan pada penelitian ini karena menggunakan preferensi pengguna untuk memberikan rekomendasi.

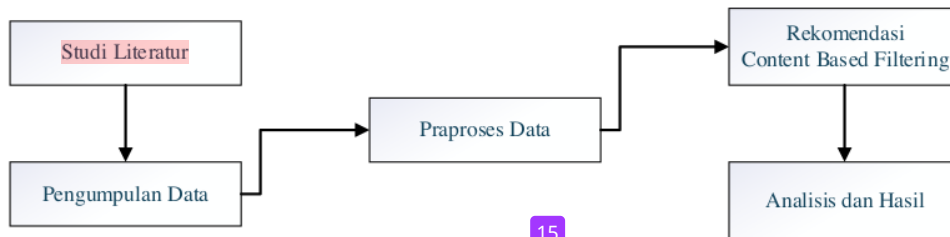
*Content Based Filtering* (CBF) merupakan sistem yang memberikan rekomendasi berdasarkan korelasi antara konten yang diberikan. CBF menggunakan informasi, seperti atribut, untuk menghitung nilai kemiripannya (Son & Kim, 2017). Pada penelitian sebelumnya *Content Based Filtering* (CBF) digunakan untuk merekomendasikan produk minuman kopi. Hal ini dilakukan karena terdapat fenomena *overchoice* di sisi pelanggan. Hasil dari penelitian ini mengatakan algoritma *Content Based Filtering* (CBF) berhasil permasalahan *overchoice* dan pelanggan dapat memilih preferensi berdasarkan kesukaan masing-masing (Kosim & Prihandi, 2024). Kemudian *Content Based Filtering* (CBF) juga dibangun dan diterapkan untuk meningkatkan *customer relation ship* agar dapat meningkatkan *awarness* pelanggan. Hal ini dilakukan agar dapat mencegah gagalnya transaksi. Hasil dari penelitian ini mengatakan algoritma CBF dapat memberikan rekomendasi terbaik dengan nilai bobot tertinggi dengan akurasi hingga 96,6% (Mariani Widia Putri et al., 2022). Penelitian selanjutnya berfokus pada pengembangan sistem

rekomendasi algoritma CBF yang efisien. Dalam penelitian ini metode yang digunakan menggunakan teknik TF-IDF dan kesamaan kosinus. Hasilnya algoritma CBF menunjukkan efektivitas dalam memberikan rekomendasi secara akurat sehingga dapat meningkatkan pengalaman pengguna dan dapat memfasilitasi eksplorasi pasar mobil diecast (Anak Agung Aditya Nugraha & Ngurah Agus Sanjaya ER, 2023).

Perbedaan usulan penelitian dengan penelitian-penelitian sebelumnya adalah pada penelitian ini perpustakaan sebagai sumber data khususnya adalah daftar buku-buku yang ada pada program studi informatika. Sehingga mahasiswa informatika dapat lebih mudah dalam mencari buku-buku informatika sebagai bahan ajar berdasarkan kriteria pengguna.

### III. METODE PENELITIAN

Dalam metode penelitian ini memiliki beberapa tahap, seperti studi literatur, pengumpulan data, praproses data, rekomendasi *Content Based Filtering*, analisis dan hasil. Metode penelitian dapat dilihat pada Gambar 1.



Gambar 1 Metode Penelitian

#### 2.1 Studi literatur

Studi literatur dilakukan guna mempelajari keilmuan yang terkait dengan topik penelitian yakni buku-buku yang ada dalam domain informatika, data science, praproses data, sistem rekomendasi dan algoritma *content based filtering* yang bersumber dari artikel jurnal, internet, buku dan lain sebagainya.

#### 2.2 Pengumpulan Data

Data informasi buku diperoleh dari perpustakaan STMIK Sinar Nusantara. Lingkup program studi yang ada dalam perguruan tinggi ini adalah di domain informatika sehingga buku yang dimiliki sudah mencakup keilmuan-keilmuan yang ada di domain informatika. Jumlah data diperoleh sebanyak 162. Variabel-variabel yang diperoleh dalam dataset seperti tahun terbit, judul buku, pengarang, penerbit, alamat, harga dan dekripsi.

#### 2.3 Praproses Data

Pada tahap ini dilakukan persiapan data yang diperoleh agar siap untuk dianalisis. Dalam prosesnya, pengambilan data mentah atau raw disiapkan agar dapat diserap oleh platform analitik. Praproses data meliputi pembersihan karakter, noise, tokenization, teks noise, teks yang tidak konsisten dan lain sebagainya (Ridhwanullah & Fudholi, 2022). Adapun tahapan ini yaitu seperti:

##### 2.3.1 Proses tokenization

Perolehan data teks akan dirubah menjadi bagian-bagian yang lebih kecil. Proses ini merubah teks kalimat menjadi tiap-tiap kata. Hasil dari proses ini disebut dengan token dan akan dilanjutkan ke tahap selanjutnya.

### 2.3.2 Proses normalization

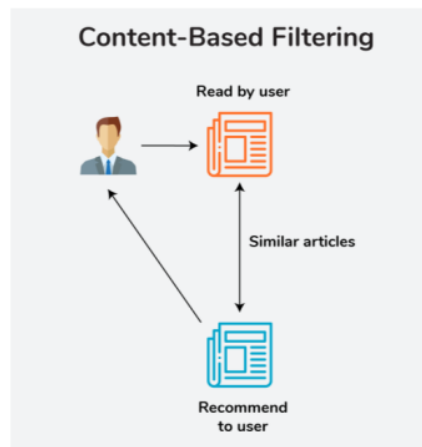
Tahap selanjutnya adalah merubah teks menjadi normal atau memiliki format yang sama, dalam proses ini seperti menghilangkan tanda baca dari segala teks yang mengandungnya kemudian penghapusan imbuhan kata dan merubah huruf besar menjadi huruf kecil.

### 2.3.3 Stopwords

Langkah ini digunakan untuk menghapus kata-kata umum yang biasanya digunakan untuk menyambung kalimat dan hal ini biasanya diabaikan oleh *search engine*. Stopword biasanya dilakukan dalam tahap praproses data, kata-kata yang dinilai tidak mempunyai arti terhadap penelitian dan dapat meningkatkan dimensi data secara tidak perlu (Syahril Dwi Prasetyo et al., 2023). Kata-kata umum dalam bentuk bahasa indonesia telah tersedia pada library sastrawi. Library ini merupakan pengembangan dari PHP sastrawi.

### 2.4 Content based filtering

*Content based recommendation* menggunakan ketersediaan konten yang dimiliki yakni disebut dengan fitur, atribut dan atau karakteristik yang dimana merupakan basis dalam memberikan rekomendasi. *Content based filtering* memanfaatkan minat pengguna dan mencocokkan preferensi pengguna dengan atribut yang dimiliki oleh berbagai item untuk direkomendasikan (Kosim & Prihandi, 2024). Jenis algoritma ini tidak melibatkan pengguna lain. Namun kelibatan pengguna berdasarkan apa yang disukai dan algoritma akan memilih konten yang serupa untuk direkomendasikan.



Gambar 2 Ilustrasi *Content Based Filtering*

### 2.5 Analisis dan evaluasi

Pada tahap ini dilakukan analisis kualitatif dengan memperhatikan hasil dari beberapa rekomendasi buku yang akan dianalisis. Pada tahap evaluasi dilakukan penilaian terhadap akurasi hasil rekomendasi.

## IV. HASIL DAN PEMBAHASAN

Data buku yang telah diperoleh akan melalui tahap praproses untuk mempersiapkan data agar siap untuk di analisis. Hal ini dilakukan karena data yang diperoleh mengandung beberapa karakter, tanda baca, tidak konsisten, dan lain-lain. Pada penelitian ini, tahap praproses diterapkan pada kolom deskripsi. Pada Tabel 1 merupakan potongan data variabel deskripsi buku yang belum dan sudah dilakukan praproses. Hasil kolom deskripsi setelah praproses akan disimpan ke kolom deskripsi\_clean.

Tabel 1 Tahap Praproses Data

2 Sebelum Praproses	11 Hasil Praproses
<p>Selama ini, petunjuk yang dijadikan acuan untuk menentukan keaslian sebuah foto hanyalah berdasarkan informasi dari metadata yang terdapat pada sebuah foto. Buku ini akan mengulas bagaimana metadata ini tidak bisa dipercaya dan bisa diperdaya dengan mudah. Dalam buku ini akan ditunjukkan pula, bagaimana cara mendeteksi foto-foto hasil <b>1</b> kayasa tersebut.</p>	<p>petunjuk dijadikan acuan menentukan keaslian foto berdasarkan informasi metadata foto buku mengulas metadata dipercaya diperdaya mudah buku mendeteksi foto foto hasil rekayasa</p>
<p>Buku ini dimaksudkan untuk digunakan dikalangan pelajar, mahasiswa dan praktisi IT jaringan komputer. Dalam memahami buku ini, diharapkan pembaca sudah pernah dan bisa membuat jaringan under windows maupun under linux, sehingga akan memperlancar dalam mempraktekkan isi buku ini.</p> <p>Sistem operasi yang digunakan untuk menyusun tutorial dalam buku ini adalah MikroTikOS dan software Utility. Adapun untuk software nya bisa didapat dengan cara download di <a href="http://www.mikrotik.com">http://www.mikrotik.com</a>. Konfigurasi jaringannya dengan menggunakan WinBox dan teks command line MikroTikOS.</p> <p>.....</p> <p>.....</p>	<p><b>1</b> buku dikalangan pelajar mahasiswa praktisi jaringan komputer memahami buku diharapkan pembaca jaringannya windows linux memperlancar mempraktekkan isi buku sistem operasi menyusun tutorial buku mikrotikos software utility software nya <b>1</b> download konfigurasi jaringannya winbox teks command line mikrotikos penyusunan buku dibagi topik bahasan komponen topologi jaringan komputer ip address instalasi mikrotikos perintah dasar networking router proxy manajemen bw firewall mrtg implementasi hotspot modem adsl buku dilengkapi tutorial membangun jaringan komputer mikrotikos detail lengkapi contoh implementasi praktis membaca buku diharapkan pembaca langsung mempraktekkan mengundang guru instruktur jaringan</p>

Menurut pada Tabel 1, kolom pertama atau sebelum praproses merupakan data raw terkait dengan deskripsi tiap buku. Teks atau narasi dalam deskripsi tentu saja masih memuat huruf besar dan kecil, terdapat tanda baca, terdapat simbol dan lain-lain. Kondisi ini yang memerlukan perlu adanya langkah praproses yakni seperti tokenization, normalization, noise removal (Ridhwanullah & Fudholi, 2022). Hasil dari praproses dapat dilihat pada kolom ke dua. Pada Gambar 3 merupakan potongan kode program praproses.

Pada penelitian ini dilakukan stopword dua kali yakni bahasa indonesia dan inggris. Stopword merupakan kata umum yang berjumlah besar yang dimana kata-kata ini tidak mempunyai arti atau makna untuk penelitian dan biasanya diabaikan serta disimpan dalam stop list. Penggunaan library stopwords bahasa indonesia menggunakan sastrawi dan bahasa inggris menggunakan bawaan library nltk.

```
#clean_spc1 = re.compile('[/(){}[\]\|@;:]')
clean_spc1 = re.compile(r'(?i)\b(?:https?://|www\d{0,3}[.]|[a-z0-9.-]+[.]|[a-z]{2,4}/)(?:[^\s()<+>]+|\\(?:[^\s()<+>]*\\)+)?(?:[^\s()<+>]+|\\(?:[^\s()<+>]*\\)+)*')
clean_symbol = re.compile('[^0-9a-z #_ ]')
stopwordi = set(stopwords.words('indonesian'))
stopworde = set(stopwords.words('english'))
```

Gambar 3 Potongan kode program praproses

TF-IDF atau Term Frequency – Inverse Document Frequency digunakan untuk menghitung bobot tiap kata yang paling umum digunakan untuk pengambilan informasi. Perhitungan nilai suatu bobot tiap kata tergantung dengan jumlah kemunculan kata tersebut dalam dokumen. Semakin sering kata tersebut muncul maka mengindikasikan pentingnya kata terhadap dokumen (Riza Adrianti Supono <sup>7</sup> Muhammad Azis Suprayogi, 2021). Metode TF-IDF menggabungkan 2 teknik yakni frekuensi kemunculan kata dalam sebuah dokumen dan inverse frekuensi dokumen yang memuat kata tersebut.

Sedangkan IDF digunakan untuk mengukur seberapa umum sebuah kata muncul dalam dokumen. Setelah perhitungan menggunakan metode TF-IDF maka dapat mengkalikan keduanya. Sehingga akan memperoleh nilai tiap kata. Nilai tiap kata inilah yang akan digunakan sebagai bobot. Semakin kata tersebut muncul maka semakin kata tersebut mempunyai makna. Namun tiap kata yang memiliki nilai tinggi belum tentu mempunyai nilai tinggi pula di dokumen lainnya. Dalam penelitian ini, banyaknya kata yang muncul pada dokumen adalah pada kolom deskripsi\_clean.

*Cosine similarity* digunakan untuk mencari kemiripan dari dokumen ke dokumen lainnya. Implementasi dari teknik ini dapat menghitung nilai kemiripan tiap antar kalimat dan dapat mengukur kemiripan teks yang populer. Hasil perhitungan matrik TF-IDF akan dilakukan perhitungan nilai kesamaan antar buku. Pada Gambar 4 merupakan potongan kode program TF-IDF dan cosine similarity dan pada Gambar 5 merupakan hasil perhitungannya.

```
df.set_index('JudulBuku', inplace = True)
tf = TfidfVectorizer(analyzer='word', ngram_range=(1, 3), min_df=0.0, stop_words='english')
tfidf_matrix = tf.fit_transform(df['Deskripsi_clean'])
#cosine_similarities = linear_kernel(tfidf_matrix, tfidf_matrix)
cos_sim = cosine_similarity(tfidf_matrix, tfidf_matrix)
cos_sim
```

Gambar 4 Potongan kode program TF-IDF dan *Cosine Similarity*

```
array([[1.00000000e+00, 5.32852855e-03, 4.97700707e-03, ...,
        5.61933437e-03, 6.97816496e-04, 8.72318604e-03],
       [5.32852855e-03, 1.00000000e+00, 1.37231736e-01, ...,
        1.34753502e-03, 4.47479585e-03, 1.79795208e-02],
       [4.97700707e-03, 1.37231736e-01, 1.00000000e+00, ...,
        5.79627525e-03, 2.12185762e-03, 5.51450804e-03],
       ...,
       [5.61933437e-03, 1.34753502e-03, 5.79627525e-03, ...,
        1.00000000e+00, 1.39695090e-01, 2.65461740e-02],
       [6.97816496e-04, 4.47479585e-03, 2.12185762e-03, ...,
        1.39695090e-01, 1.00000000e+00, 6.54560386e-02],
       [8.72318604e-03, 1.79795208e-02, 5.51450804e-03, ...,
        2.65461740e-02, 6.54560386e-02, 1.00000000e+00]])
```

Gambar 5 Hasil perhitungan TF-IDF dan *Cosine Similarity*

Kemudian percobaan untuk hasil rekomendasi dari persepsi pengguna. Pengguna dapat memasukkan judul-judul buku yang terdapat pada data set. Dalam memasukan judul buku ini harus lengkap. Adapun contoh daftar judul buku seperti pada Gambar 6 dan beberapa hasil rekomendasi judul buku dapat dilihat pada Gambar 7 sampai Gambar 9.

0	Image Forensics : Rahasia Mendeteksi Pemalsuan...
1	Belajar Jaringan Komputer Berbasis Mikrotik OS
2	Jaringan Komputer Dengan TCP/IP : Membahas Kon...
3	Jaringan Komputer Berbasis Mikrotik + CD
4	Pemrograman Arduino & Android Menggunakan App ...
5	Pemrograman Metode Elemen Hingga Berbasis Matl...
6	Pemrograman Web Dinamis menggunakan Java Serve...
7	Belajar Sendiri Adobe InDesign 2019 : Cara mud...
8	Belajar Photoshop CC untuk Fotografi & Desain ...
9	Desain 2D dengan AutoCAD
10	Modul Pemrograman Web (HTML, PHP dan MySQL/Mar...
11	Membuat aplikasi inventory dengan java netbean...
12	Menguasai Adobe Photoshop CC 2021
13	Membangun Sistem Informasi SDM dengan Microsof...
14	Membuat Aplikasi Koperasi Simpan Pinjam dengan...
15	Mudah Belajar CorelDraw
16	Membangun Aplikasi Mobile dengan PWA : Progres...
17	Mudah Membuat dan Berbisnis Aplikasi Android d...
18	Mudah dan Cepat Membuat Program Skripsi dan Tu...
19	Mudah Membuat Website Menggunakan CodeIgniter
20	Mudah Menguasai SPSS : Shortcourse Series
21	Membuat Berbagai Game Android dengan Adobe Ani...
22	Membuat Robot Menggunakan Raspberry Pi + Pemro...
23	Metode Numerik Berbasis Python
24	Network Security Dan Cyber Security : Teori da...
25	Panduan Adobe Illustrator: Cara Mudah Bagi Sia...
26	Panduan Membuat Robot Virtual Menggunakan Ardu...
27	Pengantar Machine Learning: Konsep dan Praktik...
28	Pengenalan Machine Learning Dengan Python
29	Pemrograman Python Untuk Penanganan Big Data
30	Pemrograman Python Untuk Ilmu Komputer dan Teknik
31	Python untuk Membuat Game hingga Face Detector

Gambar 6 Potongan Daftar Judul Buku

```
rekendasi("Belajar Jaringan Komputer Berbasis Mikrotik OS")  
  
['Panduan Lengkap Instalasi Dan Konfigurasi Jaringan LAN - WAN - Wireless - Fiber Optic : Berbasis IoT Industri 4.0',  
'Jaringan Komputer Dengan TCP/IP : Membahas Konsep dan Teknik Implementasi TCP/IP dalam Jaringan Komputer Edisi Revisi ',  
'Panduan lengkap membangun sendiri sistem jaringan komputer',  
'Switch dan Multilayer Switch Cisco Implementasi Jaringan Akses + CD',  
'Router Mikrotik Implementasi Wireless Lan Indoor',  
'Jaringan Komputer Berbasis Mikrotik + CD',  
'Network Security Dan Cyber Security : Teori dan Praktik CISCO CCNA, LINUX, WINDOWS, AMAZON AWS, ANDROID + DVD',  
'Panduan Membuat Robot Virtual Menggunakan Arduino dan Scratch + CD',  
'Sistem Operasi, Konsep Dan Membuat Linux OpenWRT Dan ROM Android',  
'Semua bisa menjadi Programmer VB 2012 Case Study']
```

Gambar 7 Hasil Rekomendasi Judul Buku - 1

```
rekendasi("Pemrograman PHP Dan MySQL Untuk Pemula")  
  
['PHP & MySQL : Langkah Demi Langkah + CD',  
'PHP Komplet',  
'Pemrograman Web dengan PHP7 + DVD',  
'Pemrograman Web Edisi Revisi + CD',  
'Pemrograman Database Komplet',  
'Modul Pemrograman Web (HTML, PHP dan MySQL/MariaDB) : edisi keempat + CD',  
'Pemrograman Database MySQL dengan PHP 7 + CD',  
'PHP MySQL Pemrograman Berorientasi Objek Bagi Programmer + CD',  
'Tuntunan Praktis: Belajar Database Menggunakan MySQL, Edisi Revisi + CD',  
'Rahasia Inti Master PHP & MySQLi (improved)']
```

Gambar 8 Hasil Rekomendasi Judul Buku - 2

Pada Gambar 7 telah dilakukan percobaan untuk memasukkan kriteria judul buku yakni “Belajar Jaringan Komputer Berbasis Mikrotik OS”. Pada percobaan ini hasil rekomendasi judul buku berjumlah sepuluh. Dari sepuluh rekomendasi daftar buku terdapat tiga judul buku yang tidak sesuai dengan isi kriteria judul buku. Kemudian pada Gambar 8, dari hasil rekomendasi yang direkomendasikan menghasilkan hasil yang tepat atau sesuai.



Kemudian penulis melakukan percobaan berikutnya dengan mengatur jumlah keluaran rekomendasi berjumlah lima. Percobaan ini dengan kriteria yang sama dengan percobaan pertama yakni “Belajar Jaringan Komputer Berbasis Mikrotik OS”. Hasil rekomendasi dapat dilihat pada Gambar 9.

```
rekomendasi("Belajar Jaringan Komputer Berbasis Mikrotik OS")  
[ 'Panduan Lengkap Instalasi Dan Konfigurasi Jaringan LAN - WAN - Wireless - Fiber Optic : Berbasis IoT Industri 4.0',  
  'Jaringan Komputer Dengan TCP/IP : Membahas Konsep dan Teknik Implementasi TCP/IP dalam Jaringan Komputer Edisi Revisi ',  
  'Panduan lengkap membangun sendiri sistem jaringan komputer',  
  'Switch dan Multilayer Switch Cisco Implementasi Jaringan Akses + CD',  
  'Router Mikrotik Implementasi Wireless Lan Indoor']
```

Gambar 9 Hasil Rekomendasi Judul Buku - 3

Percobaan ke tiga adalah untuk analisis tahap selanjutnya. Hasil yang direkomendasikan berjumlah lima. Dari hasil tersebut adalah yang mempunyai nilai kemiripan dengan kriteria judul buku yang dimasukkan oleh pengguna dari tertinggi ke terendah. Hasil rekomendasi tersebut adalah tepat semua. Kemudian bagaimana hubungan dengan percobaan pertama sebelumnya. Hal ini disebabkan oleh kurangnya dataset yang dimiliki terhadap kriteria pengguna. Dalam percobaan pertama, mempunyai ketepatan hasil rekomendasi berjumlah tujuh data judul buku yang mempunyai kemiripan. Indikasi kejadian ini adalah kurangnya jumlah data buku yang mirip dengan kriteria judul buku yang dimasukkan oleh pengguna. Jadi ketika jumlah tersebut tidak memenuhi dengan jumlah keluaran rekomendasi yang telah ditetapkan pada sistem maka algoritma akan merekomendasikan dengan nilai kemiripan tertinggi.

```
rekomendasi("Panduan Cepat Belajar HTML, PHP, dan MySQL")  
[ 'PHP & MySQL : Langkah Demi Langkah + CD',  
  'Mudah Membuat Website Menggunakan CodeIgniter',  
  'Pemrograman Web Edisi Revisi + CD',  
  'PHP MySQL Pemrograman Berorientasi Objek Bagi Programmer + CD',  
  'Pemrograman PHP Dan MySQL Untuk Pemula']
```

Gambar 10 Hasil Rekomendasi Judul Buku - 4

Percobaan ke empat pada Gambar 10, penulis memasukan fitur judul buku “Panduan Cepat Belajar HTML, PHP dan MySQL”. Hasil rekomendasi sesuai dengan fitur yang diinginkan oleh pengguna. Hal ini diartikan bahwa sistem rekomendasi mempunyai konten yang cukup untuk direkomendasikan kepada pengguna.

Pada sistem rekomendasi jumlah dataset atau jumlah konten yang dimiliki sangat berpengaruh terhadap hasil rekomendasi. Jika jumlah tersebut kecil maka akan mengakibatkan konten tidak serupa dengan isian fitur pengguna akan muncul sebagai bagian dari hasil rekomendasi. Didalam konten yang dimiliki oleh sistem paling tidak mempunyai jumlah yang cukup banyak tiap judul dalam satu lingkup. Sehingga hal ini akan mempengaruhi ketepatan dari hasil rekomendasi.

Pengujian sistem merupakan hal penting sebagai tolak ukur untuk menilai keberhasilan sistem rekomendasi. Pengujian penelitian ini menggunakan metode kualitatif yaitu mencocokkan hasil rekomendasi. Pengujian dapat dilakukan dengan menghitung nilai akurasi. Adapun rumus menghitung nilai akurasi dapat dilihat pada rumus (1) (Mariani Widia Putri et al., 2022).

$$\text{Nilai akurasi} = \frac{\text{Hasil Sesuai}}{\text{Total Hasil Presisi}} \dots\dots\dots(1)$$

Tingkat presisi dari beberapa percobaan dapat dihitung dengan rumus Presisi = Hasil Sesuai / Total hasil Presisi. Hasil presisi percobaan pertama adalah 70%, percobaan kedua 100%, ketiga 100% dan keempat 100%. Sehingga didapat rata-rata akurasi sebesar 92,5%.

## V. KESIMPULAN DAN SARAN

### 5.1. Kesimpulan

Dalam penelitian ini penulis menerapkan algoritma *Content Based Filtering* pada dataset buku informatika. Algoritma ini dalam beberapa percobaan dapat merekomendasikan daftar buku yang sesuai dengan perspektif pengguna. Algoritma ini tidak bergantung pada rating yang telah diberikan kepada pengguna lain. Sehingga hasil rekomendasi murni dari hasil nilai kemiripan tertinggi. Jumlah dataset dan jumlah tiap judul buku dalam satu lingkup sangat berperan besar dalam algoritma ini. Keluaran rekomendasi adalah berdasarkan dataset yang dimiliki. Semakin banyak dataset maka hasil rekomendasi semakin tepat. *Content based filtering* pada buku informatika memberikan kepuasan yang tinggi karena hasil tingkat akurasi mencapai 100% untuk tiga kali percobaan dan dari semua percobaan algoritma ini mempunyai rata-rata akurasi sebesar 92,5%.

### 5.2. Saran

Setelah melakukan penelitian *content based filtering* buku informatika maka penulis menyarankan untuk penelitian berikutnya agar menggunakan data yang lebih banyak, penggunaan algoritma bobot *term* selain TF-IDF, pengembangan metode *content based filtering* dan penggunaan metode rekomendasi lainnya.

## VI. DAFTAR PUSTAKA

- Anak Agung Aditya Nugraha, & Ngurah Agus Sanjaya ER. (2023). *Penyusunan Sistem Rekomendasi Produk Diecast Mobil Dengan Metode Content-Based Filtering ( CBF )*. 1, 973–976.
- Herny Februariyanti, Aryo Dwi Laksono, Jati Sasongko Wibowo, & Mardi Siswo Utomo. (2021). Implementasi Metode Collaborative Filtering untuk Sistem Rekomendasi Penjualan pada Toko Mebel. *Khatulistiwa Informatika*, IX(I), 43–50.
- Hilmi Hidayat Arfisko, & Agung Toto Wibowo. (2022). Sistem Rekomendasi Film Menggunakan Metode Hybrid Collaborative Filtering Dan Content-Based Filtering. *E-Proceeding of Engineering*, 9(3), 2149–2159.
- Kosim, & Prihandi, R. (2024). Sistem Rekomendasi Menu Minuman Dengan Metode Content – Based Filtering Berbasis Android Pada Mubtada Kopi. *Journal of Computation Science And Artificial Intelligence*, 1(1), 43–69.
- Mariani Widia Putri, Achmad Muchayan, & Made Kamisutara. (2022). *Sistem Rekomendasi Produk Pena Eksklusif Menggunakan Metode Content-Based Filtering dan TF-IDF*. 3(28).
- <sup>4</sup> Nugroho, F., & Rahayu, M. I. (2020). *Sistem Rekomendasi Produk Ukm Di Kota Bandung Menggunakan Algoritma Collaborative Filtering*. 2(3), 23–31.
- Putri, S. N., Zuraiyah, T. A., & Akhmad, D. M. (2024). *Recommender Systems Using Hybrid Demographic and Content-Based Filtering Methods for UMKM Products*. 21(May 2023), 31–44.
- Ridhwanullah, D., & Fudholi, D. H. (2022). Pemodelan Topik pada Cuitan tentang

- Penyakit Tropis di Indonesia dengan Metode Latent Dirichlet Allocation. *Jurnal Ilmiah SINUS*, 20(1), 11. <https://doi.org/10.30646/sinus.v20i1.589>
- Riza Adrianti Supono, & Muhammad Azis Suprayogi. (2021). Perbandingan Metode TF-ABS dan <sup>6</sup>F-IDF Pada Klasifikasi Teks Helpdesk Menggunakan K-Nearest Neighbor. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(5), 911–918. <https://doi.org/10.29207/resti.v5i5.3403>
- Son, J., & Kim, S. B. (2017). Content-based filtering for recommendation systems using <sup>6</sup>multiattribute networks. *Expert Systems with Applications*, 89, 404–412. <https://doi.org/10.1016/j.eswa.2017.08.008>
- <sup>3</sup>Sri Gunani Partiw, Nizam, Dewi Wulandari, Edy Cahoyono, Sri Suning Kusumawardani, Syamsul Arifin, Gatot F Hertono, Wiyanto, Ishaq, Nur Masyitah Syam, Helsa Jumaipa WY, Pradipta Hendrawan Putra, Anisa Rahmawati, Fadhilatul Fajri, Arif Pangaribowo, Ahmad Zuliensyah, Briant Sudwi Julyan, & Yoseph Yulianto. (2023). *Panduan implementasi pembelajaran berpusat pada mahasiswa*. 1–52. [https://repositori.kemdikbud.go.id/29168/1/Panduan Implementasi Pembelajaran Berpusat Pada Mahasiswa.pdf](https://repositori.kemdikbud.go.id/29168/1/Panduan%20Implementasi%20Pembelajaran%20Berpusat%20Pada%20Mahasiswa.pdf)
- Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, & Fitri Nurapriani. (2023). Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naive Bayes dan KNN. *KomtekInfo*, 10. <https://doi.org/10.35134/komtekinfo.v10i1.330>

# Content-Based Filtering pada Sistem Rekomendasi Buku Informatika

---

## ORIGINALITY REPORT

---

# 12%

SIMILARITY INDEX

---

### PRIMARY SOURCES

---

1	<a href="https://openlibrary.telkomuniversity.ac.id">openlibrary.telkomuniversity.ac.id</a> Internet	154 words — 4%
2	<a href="http://www.tokomahal.com">www.tokomahal.com</a> Internet	50 words — 1%
3	<a href="http://dikti.kemdikbud.go.id">dikti.kemdikbud.go.id</a> Internet	46 words — 1%
4	<a href="http://ejurnal.politeknikpratama.ac.id">ejurnal.politeknikpratama.ac.id</a> Internet	39 words — 1%
5	<a href="http://arifindustri.lecture.ub.ac.id">arifindustri.lecture.ub.ac.id</a> Internet	22 words — 1%
6	Wulan Patria Saroinsong, Muhamad Nurul Ashar, Irena Y. Maureen, Lina Purwaning Hartanti et al. "Reimagining Innovation in Education and Social Sciences", Routledge, 2023 Publications	19 words — 1%
7	<a href="http://lib.unnes.ac.id">lib.unnes.ac.id</a> Internet	14 words — < 1%
8	<a href="http://j-ptiik.ub.ac.id">j-ptiik.ub.ac.id</a> Internet	13 words — < 1%

9	123dok.com Internet	10 words — < 1%
10	Erlin Erlin, Bobby Hasbul Fikri, Susanti Susanti, Triyani Arita Fitri. "Deteksi Duplikasi Metadata File pada Media Penyimpanan menggunakan Metode Latent Semantic Analysis", INOVTEK Polbeng - Seri Informatika, 2020 Crossref	9 words — < 1%
11	cendolshare.blogspot.com Internet	9 words — < 1%
12	projects.co.id Internet	9 words — < 1%
13	semnati.uika-bogor.ac.id Internet	9 words — < 1%
14	widyagama.ac.id Internet	9 words — < 1%
15	iptek.its.ac.id Internet	8 words — < 1%
16	Chaira Saidah Yusrie, Ernawati Ernawati, Dadang Suherman, Ujang Cepi Barlian. "Pengembangan Kurikulum dan Proses Pembelajaran Pendidikan Tinggi:", Reslaj : Religion Education Social Laa Roiba Journal, 2021 Crossref	6 words — < 1%
17	Pantea Keikhosrokiani, Goh Man Fye. "A hybrid recommender system for health supplement e-commerce based on customer data implicit ratings", Multimedia Tools and Applications, 2023 Crossref	6 words — < 1%
18	repository.its.ac.id	

---

EXCLUDE QUOTES      OFF

EXCLUDE BIBLIOGRAPHY   OFF

EXCLUDE SOURCES      OFF

EXCLUDE MATCHES      OFF