

SISTEM REKOMENDASI PRODUK HUKUM MENGGUNAKAN CONTENT-BASED FILTERING PADA BAPAS SURAKARTA

Stefany Erica Djaksono, Vihi Atina, Fajar Suryani

Program Studi Teknik Informatika, Universitas Duta Bangsa Surakarta
Jl. Bhayangkara No.55 Kota Surakarta, Jawa Tengah, Indonesia
220103168@gmail.com

ABSTRAK

Penelitian ini dilatarbelakangi oleh banyaknya dokumen produk hukum pada portal Sistem Informasi Bapas Surakarta (SIBARATA) di Balai Pemasyarakatan (BAPAS) Surakarta yang menyebabkan proses pencarian dokumen masih dilakukan secara manual dan kurang efisien. Permasalahan tersebut membuat pengguna mengalami kesulitan dalam menemukan dokumen hukum yang relevan secara cepat dan tepat. Penelitian ini bertujuan untuk merancang dan mengimplementasikan sistem rekomendasi dokumen produk hukum menggunakan metode *Content-Based Filtering*. Metode yang digunakan dalam penelitian ini adalah TF-IDF untuk ekstraksi fitur dan *Cosine Similarity* untuk menghitung tingkat kemiripan antar dokumen. Tahapan penelitian meliputi pengumpulan data, preprocessing, ekstraksi fitur, perhitungan *similarity*, dan proses perangkingan dokumen. Hasil penelitian menunjukkan bahwa sistem mampu menghasilkan rekomendasi dokumen hukum berdasarkan tingkat kemiripan konten dengan nilai *similarity* tertinggi sebesar 1.00. Melalui proses perangkingan, sistem dapat menyajikan rekomendasi dokumen secara terurut sehingga membantu pengguna dalam menemukan dokumen hukum secara lebih cepat, tepat, dan efisien.

Kata kunci : sistem rekomendasi, content-based filtering, TF-IDF, cosine similarity, produk hukum

1. PENDAHULUAN

Balai Pemasyarakatan (BAPAS) Kelas I Surakarta merupakan instansi pemerintah yang berperan dalam penyediaan layanan informasi publik, salah satunya melalui portal Sistem Informasi Bapas Surakarta (SIBARATA) yang memuat berbagai dokumen produk hukum seperti Undang-Undang, Peraturan Pemerintah, Peraturan Menteri, serta kebijakan lainnya. Seiring dengan perkembangan digitalisasi layanan publik, kebutuhan akan sistem yang mampu mengelola dan menyajikan informasi secara efektif menjadi semakin penting untuk meningkatkan efisiensi dan aksesibilitas informasi bagi masyarakat [1]. Banyaknya dokumen hukum yang tersedia menyebabkan proses pencarian informasi menjadi lebih kompleks, terutama ketika pengguna harus menemukan dokumen tertentu dalam jumlah data yang besar [2].

Permasalahan yang terjadi pada sistem saat ini adalah proses pencarian dokumen hukum masih dilakukan secara manual melalui daftar atau kategori dokumen yang tersedia. Proses tersebut kurang efisien karena pengguna harus mencari dokumen satu per satu untuk menemukan informasi yang relevan. Selain itu, dokumen hukum memiliki struktur teks formal dan penggunaan istilah khusus yang menyebabkan pencarian dokumen menjadi lebih sulit apabila hanya mengandalkan pencarian biasa [3]. Kondisi ini dapat menghambat pengguna dalam memperoleh informasi hukum secara cepat dan tepat.

Sistem rekomendasi merupakan salah satu solusi yang dapat digunakan untuk membantu pengguna dalam menemukan informasi yang relevan secara otomatis berdasarkan karakteristik data tertentu [4]. Salah satu metode yang banyak digunakan dalam

sistem rekomendasi berbasis teks adalah *Content-Based Filtering*, yaitu metode yang bekerja dengan menganalisis kesamaan karakteristik antar dokumen berdasarkan isi atau konten dokumen tersebut [5]. Metode ini dinilai sesuai digunakan pada dokumen hukum karena mampu menganalisis hubungan antar dokumen berdasarkan kemiripan kata dan konteks isi dokumen.

Dalam implementasinya, metode *Content-Based Filtering* dapat dikombinasikan dengan TF-IDF (*Term Frequency-Inverse Document Frequency*) untuk menentukan bobot kata dalam dokumen serta *Cosine Similarity* untuk menghitung tingkat kemiripan antar dokumen dalam bentuk vektor [6]. Penggunaan kedua metode tersebut memungkinkan sistem untuk melakukan proses ekstraksi fitur, perhitungan *similarity*, serta perangkingan dokumen secara otomatis sehingga dapat menghasilkan rekomendasi dokumen hukum yang relevan.

Berdasarkan latar belakang dan permasalahan yang telah dijelaskan, penelitian ini bertujuan untuk merancang dan mengimplementasikan sistem rekomendasi dokumen produk hukum pada portal SIBARATA menggunakan metode *Content-Based Filtering* dengan pendekatan TF-IDF dan *Cosine Similarity*. Untuk menyelesaikan permasalahan tersebut, penelitian ini dilakukan melalui tahapan pengumpulan data dokumen hukum, *preprocessing* data, ekstraksi fitur menggunakan TF-IDF, perhitungan kemiripan menggunakan *Cosine Similarity*, serta proses perangkingan dokumen untuk menghasilkan rekomendasi yang relevan. Diharapkan sistem yang dikembangkan dapat membantu pengguna dalam menemukan dokumen hukum secara lebih cepat, tepat, dan efisien serta meningkatkan kualitas

layanan informasi publik di lingkungan BAPAS Surakarta.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Penelitian terkait sistem rekomendasi berbasis teks telah banyak dilakukan dengan berbagai pendekatan. Metode TF-IDF dan *Cosine Similarity* terbukti efektif dalam mengukur tingkat kemiripan dokumen berbasis teks dengan tingkat akurasi yang baik [7]. Selain itu, sistem rekomendasi berbasis konten juga mampu meningkatkan relevansi hasil pencarian pada sistem informasi digital [8].

Metode *Content-Based Filtering* memiliki keunggulan karena tidak bergantung pada data historis pengguna dan hanya memanfaatkan karakteristik item yang direkomendasikan [9]. Penggunaan teknik TF-IDF juga mampu meningkatkan kualitas representasi data teks dalam sistem rekomendasi [10]. Selain itu, metode *Cosine Similarity* mampu memberikan hasil yang stabil dan efisien dalam menghitung hubungan antar dokumen teks [11].

Sistem rekomendasi berbasis web juga dinilai mampu meningkatkan efisiensi pencarian informasi secara signifikan [12]. Namun, sebagian besar penelitian tersebut masih berfokus pada data umum seperti *e-commerce* dan belum secara spesifik diterapkan pada dokumen produk hukum yang memiliki struktur bahasa formal dan kompleks. Oleh karena itu, penelitian ini memiliki kebaruan dalam penerapan metode *Content-Based Filtering* pada dokumen hukum di lingkungan layanan publik [13].

2.2. Sistem Rekomendasi

Sistem rekomendasi merupakan sistem yang dirancang untuk memberikan saran kepada pengguna berdasarkan analisis data tertentu guna membantu menemukan informasi yang relevan [14]. Sistem ini banyak digunakan dalam berbagai bidang untuk meningkatkan efisiensi pencarian dan pengalaman pengguna.

Secara umum, sistem rekomendasi terdiri dari beberapa pendekatan seperti *collaborative filtering*, *content-based filtering*, dan *hybrid approach*. *Content-based filtering* lebih fokus pada karakteristik data yang direkomendasikan dibandingkan dengan interaksi pengguna [15].

2.3. Content-Based Filtering

Content-Based Filtering adalah metode yang merekomendasikan item berdasarkan kesamaan karakteristik atau konten yang dimiliki oleh item tersebut [16]. Metode ini bekerja dengan menganalisis atribut seperti teks, kategori, dan deskripsi untuk menentukan tingkat kemiripan antar item.

Keunggulan metode ini adalah tidak memerlukan data historis pengguna sehingga cocok digunakan pada sistem baru atau sistem berbasis dokumen seperti produk hukum [17].

2.4. TF-IDF

TF-IDF merupakan metode pembobotan kata yang digunakan untuk menentukan tingkat kepentingan suatu kata dalam dokumen [18]. Metode ini menggabungkan dua komponen utama yaitu *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF).

Dengan menggunakan TF-IDF, sistem dapat mengidentifikasi kata-kata penting dalam dokumen sehingga meningkatkan kualitas analisis teks dalam sistem rekomendasi [19].

2.5. Cosine Similarity

Cosine Similarity adalah metode yang digunakan untuk mengukur tingkat kemiripan antara dua dokumen dalam bentuk vektor [20]. Metode ini menghitung nilai kemiripan berdasarkan sudut antara dua vektor dalam ruang multidimensi. Secara matematis, *Cosine Similarity* dirumuskan sebagai berikut:

$$\cos(\theta) = \frac{A \cdot B}{\|A\| \times \|B\|}$$

Di mana A dan B merupakan vektor dari dua dokumen yang dibandingkan, $A \cdot B$ adalah hasil perkalian dot product antara kedua vektor, serta $\|A\|$ dan $\|B\|$ adalah panjang (norma) masing-masing vektor. Nilai kemiripan berada pada rentang 0 hingga 1, di mana nilai yang mendekati 1 menunjukkan tingkat kemiripan yang tinggi.

Metode ini banyak digunakan dalam sistem rekomendasi berbasis teks karena mampu memberikan hasil yang stabil dan efisien dalam menghitung hubungan antar dokumen.

3. METODE PENELITIAN

3.1. Jenis dan Sumber Data

Jenis penelitian yang digunakan dalam penelitian ini adalah penelitian kuantitatif dengan pendekatan eksperimen. Penelitian ini bertujuan untuk mengimplementasikan metode *Content-Based Filtering* dalam sistem rekomendasi dokumen produk hukum pada portal Sistem Informasi Bapas Surakarta (SIBARATA) di Balai Pemasyarakatan (BAPAS) Surakarta.

Sumber data dalam penelitian ini diperoleh melalui beberapa teknik pengumpulan data sebagai berikut:

a. Observasi

Observasi dilakukan dengan mengamati secara langsung sistem portal SIBARATA di BAPAS Surakarta. Proses yang diamati meliputi pengelolaan data dokumen hukum oleh admin, struktur data dokumen yang digunakan, serta mekanisme pencarian dokumen dalam sistem yang sedang berjalan. Observasi ini bertujuan untuk memperoleh gambaran nyata mengenai alur sistem serta permasalahan yang dihadapi dalam proses pencarian dokumen hukum.

b. Wawancara

Wawancara dilakukan dengan admin atau pihak yang terlibat dalam pengelolaan sistem SIBARATA. Wawancara bertujuan untuk memperoleh informasi mengenai kebutuhan sistem, kendala dalam pencarian dokumen hukum, serta harapan terhadap fitur rekomendasi yang akan dikembangkan agar dapat meningkatkan efisiensi pencarian informasi.

c. Studi Pustaka

Studi pustaka dilakukan dengan mengkaji berbagai sumber referensi seperti jurnal ilmiah, buku, dan artikel penelitian yang berkaitan dengan sistem rekomendasi, metode *Content-Based Filtering*, TF-IDF, dan *Cosine Similarity*. Studi pustaka ini digunakan sebagai dasar teori dalam perancangan dan implementasi sistem rekomendasi dokumen produk hukum.

3.2. Tahapan Penelitian

Tahapan penelitian dalam pengembangan sistem rekomendasi dokumen produk hukum dilakukan secara sistematis melalui beberapa proses sebagai berikut:

a. Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data dokumen produk hukum dari portal SIBARATA yang meliputi judul, kategori, dan isi dokumen. Data ini digunakan sebagai input utama dalam sistem rekomendasi.

b. Preprocessing Data

Data teks yang telah dikumpulkan kemudian diproses melalui tahap *preprocessing* yang meliputi pembersihan teks, pengubahan huruf menjadi huruf kecil (*case folding*), penghapusan tanda baca, serta penghilangan kata-kata yang tidak memiliki makna penting. Tahap ini bertujuan untuk menghasilkan data yang lebih terstruktur dan siap untuk dianalisis.

c. Ekstraksi Fitur Menggunakan TF-IDF

Pada tahap ini, data teks diubah menjadi representasi numerik menggunakan metode TF-IDF. Proses ini dilakukan dengan menghitung bobot setiap kata berdasarkan frekuensi kemunculan dalam dokumen serta tingkat kepentingannya dalam keseluruhan dokumen, sehingga setiap dokumen dapat direpresentasikan dalam bentuk vektor.

d. Perhitungan Kemiripan Menggunakan *Cosine Similarity*

Setelah diperoleh representasi vektor dari setiap dokumen, dilakukan perhitungan tingkat

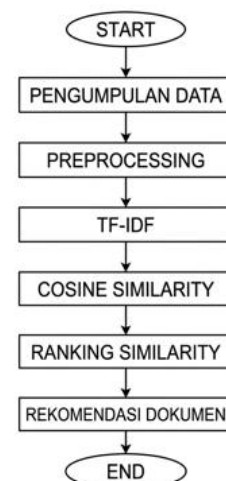
kemiripan antar dokumen menggunakan metode *Cosine Similarity*. Nilai kemiripan yang dihasilkan menunjukkan tingkat kesamaan antar dokumen, di mana nilai yang lebih tinggi menunjukkan kemiripan yang lebih besar.

e. Pembentukan Rekomendasi

Berdasarkan hasil perhitungan kemiripan, sistem akan mengurutkan dokumen berdasarkan nilai similarity dari yang tertinggi hingga terendah. Dokumen dengan nilai kemiripan tertinggi akan dipilih sebagai rekomendasi yang ditampilkan kepada pengguna.

3.3. Tahapan Penelitian

Alur proses analisis yang telah dirancang pada tahap metode penelitian diimplementasikan dalam sistem rekomendasi dokumen hukum. Proses dimulai dari pengumpulan data, preprocessing, ekstraksi fitur menggunakan TF-IDF, perhitungan kemiripan menggunakan *Cosine Similarity*, hingga pembentukan rekomendasi. Hasil implementasi alur proses tersebut ditunjukkan pada Gambar 1.



Gambar 1. Alur Proses

4. HASIL DAN PEMBAHASAN

4.1. Hasil Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data dokumen produk hukum yang berasal dari portal Sistem Informasi Bapas Surakarta (SIBARATA). Data yang digunakan berupa dokumen hukum resmi seperti Undang-Undang Republik Indonesia yang memiliki struktur teks formal. Setiap dokumen memiliki atribut berupa judul dokumen, kategori, dan isi dokumen yang digunakan dalam proses analisis.

Tabel 1. Data Dokumen Dokumen Hukum

No	Judul Dokumen	Kategori	Deskripsi
1	UU RI No 106 Tahun 2024 tentang Kabupaten Kuningan di Provinsi Jawa Barat	Undang-Undang	Pembentukan wilayah Kabupaten Kuningan
2	UU RI No 102 Tahun 2024 tentang Kabupaten Bogor di Provinsi Jawa Barat	Undang-Undang	Pembentukan wilayah Kabupaten Bogor
3	UU RI No 10 Tahun 2024 tentang Kabupaten Aceh Utara di Aceh	Undang-Undang	Pembentukan wilayah Aceh Utara

No	Judul Dokumen	Kategori	Deskripsi
4	UU RI No 15 Tahun 2024 tentang Pemerintahan Daerah	Undang-Undang	Aturan pemerintahan daerah
5	UU RI No 20 Tahun 2024 tentang Administrasi Pemerintahan	Undang-Undang	Sistem administrasi pemerintahan
6	PP No 12 Tahun 2024 tentang Pengelolaan Keuangan Negara	Peraturan Pemerintah	Pengelolaan keuangan negara
7	PP No 25 Tahun 2024 tentang Pelayanan Publik	Peraturan Pemerintah	Standar pelayanan publik
8	PP No 30 Tahun 2024 tentang Sistem Informasi Pemerintah	Peraturan Pemerintah	Sistem informasi pemerintahan
9	Permenkumham No 5 Tahun 2024 tentang Lapas	Peraturan Menteri	Pengelolaan lembaga pemasyarakatan
10	Permenkumham No 7 Tahun 2024 tentang Pembinaan Narapidana	Peraturan Menteri	Pembinaan narapidana
11	Permenkumham No 9 Tahun 2024 tentang Hak Narapidana	Peraturan Menteri	Hak narapidana
12	SOP Pembimbingan Klien Pemasyarakatan	SOP	Prosedur pembimbingan klien
13	SOP Pengawasan Klien Pemasyarakatan	SOP	Pengawasan klien
14	SOP Reintegrasi Sosial Klien	SOP	Reintegrasi sosial
15	SOP Registrasi Klien Pemasyarakatan	SOP	Registrasi klien
16	UU RI No 25 Tahun 2024 tentang Sistem Pendidikan Nasional	Undang-Undang	Sistem pendidikan nasional
17	PP No 40 Tahun 2024 tentang Pengelolaan ASN	Peraturan Pemerintah	Pengelolaan aparatur sipil negara
18	Permenkumham No 12 Tahun 2024 tentang Remisi Narapidana	Peraturan Menteri	Pemberian remisi
19	SOP Pelayanan Administrasi Klien	SOP	Prosedur pelayanan administrasi
20	SOP Monitoring Klien Pemasyarakatan	SOP	Monitoring klien

4.2. Hasil Preprocessing Data

Pada tahap preprocessing dilakukan proses pembersihan dan normalisasi terhadap data teks yang berasal dari isi dokumen hukum. Proses ini meliputi pengubahan seluruh teks menjadi huruf kecil, penghapusan tanda baca, serta penghilangan kata-kata yang tidak memiliki makna penting dalam proses analisis. Selain itu, dilakukan proses tokenisasi untuk memecah teks menjadi kata-kata penyusun sehingga memudahkan dalam proses pengolahan selanjutnya.

Hasil dari preprocessing menghasilkan data yang lebih terstruktur dan hanya berfokus pada kata-kata yang merepresentasikan isi utama dokumen hukum. Dengan demikian, proses ini dapat meningkatkan kualitas data yang digunakan dalam tahap ekstraksi fitur menggunakan metode TF-IDF.

Berdasarkan Tabel 2, proses preprocessing dilakukan dengan mengubah teks menjadi huruf kecil (case folding) serta memecah teks menjadi kata-kata penting melalui proses tokenisasi. Untuk menjaga kejelasan tampilan, hanya sebagian kata kunci utama yang ditampilkan pada tabel.

Table 2. Contoh Hasil Preprocessing Data

No	Data Asli	Case Folding	Tokenisasi
1	UU RI No 106 Tahun 2024 tentang Kabupaten Kuningan	uu ri no 106 tahun 2024 kabupaten kuningan	uu, kabupaten, kuningan
2	PP No 25 Tahun 2024 tentang Pelayanan Publik	pp no 25 tahun 2024 pelayanan publik	pp, pelayanan, publik

4.3. Hasil Ekstraksi Fitur Menggunakan TF-IDF

Pada tahap ini dilakukan ekstraksi fitur menggunakan metode TF-IDF untuk mengidentifikasi kata-kata penting yang merepresentasikan isi dokumen hukum. Kata-kata yang muncul berasal dari istilah umum dalam dokumen hukum seperti “undang-undang”, “kabupaten”, “pemerintahan”, dan “narapidana”.

Tabel 3. Hasil TF-IDF

No	Produk	Term 1	Term 2	Term 3
1	UU RI No 106 Tahun 2024 tentang Kabupaten Kuningan di Provinsi Jawa Barat	undang	kabupaten	jawa
2	UU RI No 102 Tahun 2024 tentang Kabupaten Bogor di Provinsi Jawa Barat	undang	kabupaten	bogor
3	UU RI No 10 Tahun 2024 tentang Kabupaten Aceh Utara di Aceh	undang	kabupaten	aceh
4	UU RI No 15 Tahun 2024 tentang Pemerintahan Daerah	pemerintahan	daerah	aturan
5	UU RI No 20 Tahun 2024 tentang Administrasi Pemerintahan	administrasi	pemerintahan	sistem
6	PP No 12 Tahun 2024 tentang Pengelolaan Keuangan Negara	keuangan	negara	pengelolaan
7	PP No 25 Tahun 2024 tentang Pelayanan Publik	pelayanan	publik	standar

No	Produk	Term 1	Term 2	Term 3
8	PP No 30 Tahun 2024 tentang Sistem Informasi Pemerintah	sistem	informasi	pemerintah
9	Permenkumham No 5 Tahun 2024 tentang Lapas	lapas	pengelolaan	pemasyarakatan
10	Permenkumham No 7 Tahun 2024 tentang Pembinaan Narapidana	pembinaan	narapidana	program
11	Permenkumham No 9 Tahun 2024 tentang Hak Narapidana	hak	narapidana	aturan
12	SOP Pembimbingan Klien Pemasyarakatan	pembimbingan	klien	prosedur
13	SOP Pengawasan Klien Pemasyarakatan	pengawasan	klien	prosedur
14	SOP Reintegrasi Sosial Klien	reintegrasi	sosial	klien
15	SOP Registrasi Klien Pemasyarakatan	registrasi	klien	prosedur
16	UU Sistem Pendidikan	pendidikan	nasional	system
17	PP Pengelolaan ASN	asn	pengelolaan	aparatur
18	Permenkumham Remisi	remisi	narapidana	pemberian
19	SOP Administrasi Klien	administrasi	Klien	pelayanan
20	SOP Monitoring Klien	monitoring	klien	pengawasan

4.4. Hasil Perhitungan Kemiripan Menggunakan Cosine Similarity

Pada tahap ini dilakukan perhitungan tingkat kemiripan antar dokumen produk hukum menggunakan metode *Cosine Similarity*. Perhitungan ini bertujuan untuk mengetahui tingkat kesamaan antar dokumen berdasarkan representasi vektor hasil pembobotan TF-IDF. Nilai kemiripan berada pada rentang 0 hingga 1, di mana nilai yang mendekati 1 menunjukkan tingkat kemiripan yang tinggi.

Sebagai contoh, dilakukan perhitungan kemiripan antara dokumen UU RI No 106 Tahun 2024 tentang Kabupaten Kuningan dengan dokumen lainnya. Misalnya dibandingkan dengan dokumen UU RI No 102 Tahun 2024 tentang Kabupaten Bogor, kedua dokumen memiliki beberapa kata yang sama seperti “kabupaten” dan “provinsi”. Sebagai ilustrasi, kata kunci yang digunakan dalam pencarian adalah “kabupaten jawa barat”. Berdasarkan kata kunci tersebut, sistem akan merekomendasikan dokumen

yang memiliki kesamaan kata dan konteks yang relevan, sehingga menghasilkan nilai kemiripan yang cukup tinggi.

Sebagai contoh, perhitungan nilai similarity dilakukan pada dokumen UU RI No 106 Tahun 2024 tentang Kabupaten Kuningan di Provinsi Jawa Barat terhadap dokumen yang sama. Kata kunci utama yang digunakan adalah “kabupaten”, “kuningan”, dan “jawa”. Karena dokumen dibandingkan dengan dirinya sendiri, maka vektor yang dihasilkan memiliki bobot yang sama. Misalnya vektor dokumen A adalah $A = (1,1,1)$ dan vektor dokumen B adalah $B = (1,1,1)$, maka perhitungan *Cosine Similarity* adalah sebagai berikut:

$$\begin{aligned}
 \text{Cosine Similarity} &= \frac{12 + 12 + 12 \times 12 + 12 + 12(1 \times 1) + (1 \times 1) + (1 \times 1)}{3} \\
 &= \frac{\sqrt{3} \times \sqrt{3}}{3} \\
 &= \frac{3}{3} = 1,00
 \end{aligned}$$

Tabel 4. Hasil *similarity* antar dokumen

No	Produk	Similarity
1	UU RI No 106 Tahun 2024 tentang Kabupaten Kuningan di Provinsi Jawa Barat	1.00
2	UU RI No 102 Tahun 2024 tentang Kabupaten Bogor di Provinsi Jawa Barat	0.88
3	UU RI No 10 Tahun 2024 tentang Kabupaten Aceh Utara di Aceh	0.85
4	UU RI No 15 Tahun 2024 tentang Pemerintahan Daerah	0.6
5	UU RI No 20 Tahun 2024 tentang Administrasi Pemerintahan	0.55
6	PP No 12 Tahun 2024 tentang Pengelolaan Keuangan Negara	0.3
7	PP No 25 Tahun 2024 tentang Pelayanan Publik	0.35
8	PP No 30 Tahun 2024 tentang Sistem Informasi Pemerintah	0.32
9	Permenkumham No 5 Tahun 2024 tentang Lapas	0.4
10	Permenkumham No 7 Tahun 2024 tentang Pembinaan Narapidana	0.42
11	Permenkumham No 9 Tahun 2024 tentang Hak Narapidana	0.38
12	SOP Pembimbingan Klien Pemasyarakatan	0.2
13	SOP Pengawasan Klien Pemasyarakatan	0.18
14	SOP Reintegrasi Sosial Klien	0.15
15	SOP Registrasi Klien Pemasyarakatan	0.17
16	UU RI No 25 Tahun 2024 tentang Sistem Pendidikan Nasional	0.50
17	PP No 40 Tahun 2024 tentang Pengelolaan ASN	0.45
18	Permenkumham No 12 Tahun 2024 tentang Remisi Narapidana	0.43
19	SOP Pelayanan Administrasi Klien	0.22
20	SOP Monitoring Klien Pemasyarakatan	0.21

4.5. Hasil Pembentukan Rekomendasi Dokumen Hukum

Berdasarkan hasil perhitungan kemiripan menggunakan metode *Cosine Similarity* pada tahap sebelumnya, sistem selanjutnya melakukan proses pembentukan rekomendasi dokumen hukum melalui mekanisme perangkingan nilai *similarity*. Proses ini dilakukan dengan mengurutkan seluruh dokumen berdasarkan nilai kemiripan dari yang tertinggi hingga terendah.

Sebagai contoh, digunakan dokumen UU RI No 106 Tahun 2024 tentang Kabupaten Kuningan di Provinsi Jawa Barat sebagai dokumen acuan. Berdasarkan hasil perhitungan *similarity* terhadap 20

dokumen, diperoleh nilai kemiripan yang kemudian diurutkan sehingga menghasilkan ranking rekomendasi.

Nilai tersebut kemudian diurutkan dari yang tertinggi ke terendah sehingga diperoleh ranking dokumen. Berdasarkan hasil perangkingan, sistem merekomendasikan dokumen dengan nilai *similarity* tertinggi sebagai rekomendasi utama, yaitu UU RI No 102 Tahun 2024 dan UU RI No 10 Tahun 2024.

Dokumen dengan nilai *similarity* tertinggi menunjukkan tingkat kemiripan paling tinggi terhadap dokumen acuan dan direkomendasikan sebagai hasil utama sistem.

Tabel 5. Hasil Rekomendasi Dokumen Hukum

No	Rangking	Produk Hukum	Nilai Similarity
1	1	UU RI No 106 Tahun 2024 tentang Kabupaten Kuningan	1
2	2	UU RI No 102 Tahun 2024 tentang Kabupaten Bogor	0.88
3	3	UU RI No 10 Tahun 2024 tentang Kabupaten Aceh Utara	0.85
4	4	UU RI No 15 Tahun 2024 tentang Pemerintahan Daerah	0.6
5	5	UU RI No 20 Tahun 2024 tentang Administrasi Pemerintahan	0.55
6	6	UU RI No 25 Tahun 2024 tentang Sistem Pendidikan Nasional	0.5
7	7	PP No 40 Tahun 2024 tentang Pengelolaan ASN	0.45
8	8	Permenkumham No 12 Tahun 2024 tentang Remisi Narapidana	0.43
9	9	Permenkumham No 7 Tahun 2024 tentang Pembinaan Narapidana	0.42
10	10	Permenkumham No 5 Tahun 2024 tentang Lapas	0.4
11	11	Permenkumham No 9 Tahun 2024 tentang Hak Narapidana	0.38
12	12	PP No 25 Tahun 2024 tentang Pelayanan Publik	0.35
13	13	PP No 30 Tahun 2024 tentang Sistem Informasi Pemerintah	0.32
14	14	PP No 12 Tahun 2024 tentang Pengelolaan Keuangan Negara	0.3
15	15	SOP Pelayanan Administrasi Klien	0.22
16	16	SOP Monitoring Klien Pemasarakatan	0.21
17	17	SOP Pembimbingan Klien Pemasarakatan	0.2
18	18	SOP Pengawasan Klien Pemasarakatan	0.18
19	19	SOP Registrasi Klien Pemasarakatan	0.17
20	20	SOP Reintegrasi Sosial Klien	0.15

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, sistem rekomendasi dokumen produk hukum pada portal Sistem Informasi Bapas Surakarta (SIBARATA) berhasil dibangun menggunakan metode *Content-Based Filtering* dengan pendekatan TF-IDF dan *Cosine Similarity*. Proses pengolahan data yang meliputi *preprocessing*, ekstraksi fitur, dan perhitungan kemiripan mampu merepresentasikan dokumen dalam bentuk vektor sehingga memungkinkan perbandingan antar dokumen secara efektif. Hasil perhitungan *similarity* menunjukkan bahwa dokumen yang dibandingkan dengan dirinya sendiri memiliki nilai kemiripan tertinggi sebesar 1.00. Proses perangkingan berdasarkan nilai *similarity* memungkinkan sistem menyusun rekomendasi secara terurut dari yang paling relevan hingga yang paling rendah, sehingga membantu pengguna dalam menemukan dokumen hukum secara lebih cepat, tepat, dan efisien. Untuk pengembangan selanjutnya, sistem dapat ditingkatkan dengan menambahkan teknik *preprocessing* yang lebih kompleks seperti *stemming* dan *lemmatization*, menggabungkan metode lain

seperti *hybrid recommendation*, serta memperluas jumlah data dan mengintegrasikannya ke dalam sistem berbasis web secara real-time agar menghasilkan rekomendasi yang lebih akurat dan adaptif terhadap kebutuhan pengguna.

DAFTAR PUSTAKA

- [1] H. Zhang, Q. Liu, and S. Wang, "A Review of Text-Based Recommendation Systems Using TF-IDF and Similarity Measures," *Information Processing & Management*, vol. 59, no. 6, pp. 1–15, 2022.
- [2] M. Huda et al., "Sistem Rekomendasi Content-Based Filtering Menggunakan TF-IDF dan Cosine Similarity," *Jurnal Teknologi Informasi*, vol. 14, no. 2, pp. 101–110, 2022.
- [3] M. Ridwansyah, S. Subartini, and Sylviani, "Penerapan Content-Based Filtering dalam Sistem Rekomendasi," *Jurnal Sistem Informasi dan Teknologi*, vol. 6, no. 1, pp. 88–97, 2024.
- [4] S. Ginting and A. Pratama, "Sistem Rekomendasi Produk Menggunakan Content-Based Filtering,"

- Jurnal Informatika*, vol. 10, no. 2, pp. 100–110, 2023.
- [5] T. Wahyudi and A. Nugraha, “Optimalisasi Sistem Rekomendasi Berbasis Web untuk Meningkatkan Akses Informasi UMKM,” *Jurnal Teknologi dan Sistem Informasi*, vol. 8, no. 2, pp. 130–142, 2024.
- [6] L. Chen and H. Wang, “Hybrid Recommendation Systems: A Review,” *Artificial Intelligence Review*, vol. 56, pp. 1–20, 2023.
- [7] V. Kumar and P. Singh, “Content-Based Recommendation Techniques: A Comparative Study,” *International Journal of Computer Applications*, vol. 184, no. 25, pp. 15–22, 2022.
- [8] D. Almazro et al., “Comparative Study of Recommendation Systems,” *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 1, pp. 1–10, 2023.
- [9] Y. Zhao et al., “Feature Extraction using TF-IDF in Text Mining Applications,” *IEEE Access*, vol. 12, pp. 45567–45578, 2024.
- [10] J. Lee and S. Kim, “Document Similarity Measurement Techniques Using Cosine Similarity,” *Applied Sciences*, vol. 13, no. 4, pp. 1–12, 2023.
- [11] R. Pratama and D. Sari, “Web-Based Recommendation System Using Content-Based Filtering,” *Jurnal Teknologi Informasi dan Komunikasi*, vol. 9, no. 1, pp. 50–60, 2024.
- [12] M. Sari and R. Utama, “Recommender Systems: Concepts and Applications,” *Journal of Information Systems*, vol. 18, no. 2, pp. 75–85, 2022.
- [13] A. Setiawan and B. Prakoso, “Text-Based Recommendation System for Digital Documents,” *Jurnal Informatika*, vol. 11, no. 1, pp. 45–55, 2023.
- [14] R. Indrayuni and A. Prasetyo, “Penerapan Sistem Rekomendasi dalam Sistem Informasi Digital,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 9, no. 2, pp. 120–130, 2022.
- [15] D. Suryadi and A. Wibowo, “Analisis Pendekatan Collaborative dan Content-Based Filtering,” *Jurnal Ilmu Komputer*, vol. 10, no. 1, pp. 30–40, 2023.
- [16] F. Rahman and T. Hidayat, “Implementasi Content-Based Filtering pada Sistem Rekomendasi,” *Jurnal Sistem Informasi*, vol. 8, no. 2, pp. 90–100, 2022.
- [17] A. Maulana and D. Saputra, “Pengembangan Sistem Rekomendasi Berbasis Konten pada Platform Digital,” *Jurnal Teknologi Informasi*, vol. 11, no. 1, pp. 60–70, 2023.
- [18] E. Indrayuni et al., “Analisis Pembobotan Kata Menggunakan TF-IDF pada Dokumen Teks,” *Jurnal Informatika*, vol. 9, no. 1, pp. 25–35, 2022.
- [19] R. Wijaya and D. Santoso, “Penerapan TF-IDF untuk Representasi Data Teks,” *Jurnal Sistem Informasi*, vol. 12, no. 2, pp. 80–90, 2023.
- [20] M. Zulkarnain and R. Prasetyo, “Implementasi Cosine Similarity dalam Sistem Rekomendasi Dokumen,” *Jurnal Teknologi Informasi*, vol. 13, no. 1, pp. 40–50, 2024.