

Sistem Rekomendasi Buku Menggunakan Algoritma Rabin-Karp

Arif Nur Rohman¹, Marwan Noor Fauzy², Ahmad Sa'di³

Sistem Informasi

Universitas Amikom Yogyakarta

Yogyakarta, Indonesia

e-mail: ¹arifrahman@amikom.ac.id, ²marwannoorfauzy@amikom.ac.id, ³ahmadsa@amikom.ac.id

Diajukan: 6 Januari 2024; Direvisi: 17 Februari 2024; Diterima: 19 Februari 2024

Abstrak

Toko BukuKita telah menyediakan katalog buku yang dapat diakses atau dibeli melalui website resminya yaitu *bukukita.com*, pada website tersebut pengunjung dapat melakukan pencarian buku dengan memasukkan keyword berupa judul atau pengarang. Namun, rekomendasi buku yang diberikan oleh website ini hanya berdasarkan kesamaan penulis atau juga kategori, padahal buku dengan kategori yang sama belum tentu topiknya sama. Pendekatan yang dilakukan pada penelitian ini adalah menggunakan pendekatan *content based filtering* dengan algoritma Rabin-Karp yang mana deskripsi buku dijadikan sebagai bahan untuk mendapatkan rekomendasi berupa buku lain yang memiliki kedekatan deskripsi. Sebanyak 201 dataset dikumpulkan dari kategori buku agama dan komputer, setelah dilakukan pra-pemrosesan dengan eliminasi buku yang tidak memiliki deskripsi lalu diperoleh dataset akhir 113 baris data yang selanjutnya akan digunakan untuk menguji hasil similarity dengan algoritma Rabin-Karp. Luaran dari algoritma ini adalah skor kemiripan yang mana dalam penelitian ini penulis mengambil yang skornya tertinggi, artinya semakin tinggi skornya maka semakin direkomendasikan. Pada pengujian algoritma di kategori agama buku yang direkomendasikan tetap buku-buku yang ada di kategori agama, begitu juga kategori komputer. Buku dengan kesamaan judul yang tinggi belum tentu menjadi paling direkomendasikan meskipun masuk dalam top 3 rekomendasi.

Kata kunci: Sistem Rekomendasi, Rabin-Karp, Scrapping.

Abstract

BukuKita store has provided a book catalog that can be accessed or purchased via its official website, namely *Bukukita.com*. On this website, visitors can search for books by entering keywords in the form of title or author. However, the book recommendations provided by this website are only based on similarities in authors or categories, even though books in the same category do not necessarily have the same topic. The approach taken in this research is to use a *content-based filtering* approach with the Rabin-Karp algorithm where book descriptions are used as material to get recommendations in the form of other books that have close descriptions. A total of 201 datasets were collected from the religious and computer book categories. After pre-processing was carried out by eliminating books that did not have descriptions, a final dataset of 113 rows of data was obtained which be used to test the similarity results with the Rabin-Karp algorithm. The output of this algorithm is a similarity score, which in this study the author took the highest score, meaning the higher the score, the more recommended it is. In algorithm testing in the religion category, the books recommended are still books in the religion category, as well as in the computer category. Books with high title similarities are not necessarily the most recommended even if they are included in the top 3 recommendations.

Keywords: Recommendation System, Rabin-Karp, Scrapping.

1. Pendahuluan

Toko BukuKita merupakan salah satu toko buku *online* terbesar di Indonesia. Toko buku yang didirikan di Yogyakarta pada 12 juni 2006 ini menjual buku baik secara luring maupun daring melalui *website* resmi dan platform *marketplace*. Toko buku ini menjual buku lebih dari 35 kategori mulai dari agama, anak-anak, komputer hingga sastra dan filsafat. Pada kategori agama saja, per November 2023 toko BukuKita menjual 1.476 buku. Menurut catatan transaksi penjualan, pelanggan toko BukuKita telah mencapai 301.185 dari seluruh Indonesia dan toko ini telah melayani pengiriman mulai dari kabupaten Aceh Barat hingga kabupaten Puncak jaya di Papua.

Toko BukuKita telah menyediakan katalog buku yang dapat diakses atau dibeli melalui *website* resminya yaitu *bukukita.com*, pada *website* tersebut pengunjung dapat melakukan pencarian buku dengan memasukkan *keyword* berupa judul atau pengarang dan *website* akan memberikan hasil sesuai dengan *keyword* yang dimasukkan. Pada halaman detail buku ditampilkan informasi buku yang dipilih oleh pengguna, informasinya antara lain judul, penulis, penerbit, isbn, tahun terbit, jumlah halaman, bahasa, kategori buku, foto sampul, ketersediaan stok dan deskripsi buku. Pada bagian bawah halaman detail buku ini diberikan rekomendasi buku lain untuk pengunjung dan meningkatkan konversi dari pengunjung *website*.

Namun, rekomendasi buku yang diberikan oleh *website* ini hanya berdasarkan kesamaan penulis atau juga kategori, padahal buku dengan kategori yang sama belum tentu topiknya sama. Sebagai contoh pada buku yang berjudul PHP untuk *programmer* pemula pengunjung diberikan rekomendasi buku Adobe Photoshop 8. Kedua buku ini merupakan satu kategori yang sama yaitu komputer namun topiknya sangatlah berbeda, yang satu topiknya pemrograman dan satunya lagi desain grafis.

Penelitian ini bertujuan untuk membuat rekomendasi buku pada toko BukuKita tidak hanya berdasarkan kesamaan kategori dan penulis tetapi juga kemiripan topik yang diperoleh dari deskripsi buku. Pendekatan yang dilakukan pada penelitian ini menggunakan pendekatan content based filtering dengan algoritma Rabin-Karp. Penggunaan algoritma Rabin-Karp sebelumnya sudah pernah digunakan untuk deteksi plagiarisme pada dokumen teks berbasis web[1] yang mana cara kerjanya adalah mencari kemiripan teks antar dokumen. Algoritma Rabin-Karp juga pernah digunakan untuk mendeteksi kemiripan judul tugas akhir dan memperoleh akurasi sebesar 70% [2]. Selain itu, penelitian terkait rekomendasi buku juga pernah dilakukan dengan menggunakan algoritma Weighted Tree Similarity yang memperoleh hasil presisi sebesar 88% [3]. Pada penelitian ini pendekatan yang digunakan adalah content based filtering dengan memanfaatkan sinosis buku sebagai bahan pencarian rekomendasi.

2. Metode Penelitian

Penelitian ini bertujuan menggunakan deskripsi dari suatu buku untuk memberikan rekomendasi buku lain berdasarkan kemiripan antar deskripsi buku, semakin tinggi kemiripannya maka semakin direkomendasikan. Perhitungan kemiripan deskripsi buku menggunakan algoritma Rabin-Karp diharapkan dapat menjadi pembaharuan pada penelitian sistem rekomendasi content based. Hasil dari penelitian ini akan diimplementasikan dalam halaman *website* toko *bukukita.com*.

2.1. Pengumpulan Dataset

Pengumpulan dataset tidak dilakukan dengan cara manual, mengingat banyaknya buku yang ada di toko BukuKita. Pengumpulan dataset dilakukan dengan tool scrapping yang dibuat mandiri oleh penulis menggunakan bahasa pemrograman Python. Penulis menggunakan library BeautifulSoup yang merupakan library khusus untuk membaca data yang terkandung pada antar muka *website*. Penggunaan web scrapping memiliki kemampuan dapat mengumpulkan data berupa teks, tautan, audio, video dan dokumen dengan jumlah yang besar tanpa harus melakukan akses langsung pada basis data [4]. Adapun tool scrapping yang dimaksud disajikan dalam gambar 1 sebagai berikut.

```
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).lstrip().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).lstrip().rstrip()
        sop=BeautifulSoup(requests.get(buku['url']).content, 'html5lib')
        buku['penerbit'] = re.sub('\s+', ' ', t.text).lstrip().rstrip()
        buku['deskripsi'] = re.sub('\s+', ' ', des.text).lstrip().rstrip()
        bukus.append(buku)
    ambil("https://www.bukukita.com/katalog.php?
    masId=10&catId=134&order=")
```

Gambar 1. Kode program scrapping

Kode program ditulis menggunakan Bahasa Python yang memiliki kinerja cepat dalam melakukan scrapping. Teknik scrapping dipilih karena teknik ini dapat mengambil data tanpa perlu akun atau akses basis data *website* toko BukuKita. Prinsip kerja scrapping adalah “*what you see is what you get*” yang artinya apa yang dapat tampil di layar browser maka dapat diambil pula datanya.

Inputan dari tool scrapping ini berupa url dari kategori buku dan luarannya berupa koleksi data yang disimpan pada basis data MySQL untuk memudahkan proses selanjutnya. Pemilihan basis data MySQL dimaksudkan agar mudah melakukan implementasi mengingat toko BukuKita adalah sebuah *website* yang menggunakan basis data MySQL. Scrapping dilakukan pada buku dengan kategori komputer dan agama, pemilihan 2 kategori buku bertujuan sebagai validasi yang mana sistem tidak akan merekomendasikan buku komputer pada kategori agama. Dari dataset tersebut hanya buku yang memiliki deskripsi saja yang akan digunakan oleh penulis.

2.2. Pra-pemrosesan Data

Pra-pemrosesan data dilakukan untuk membersihkan deskripsi buku dari karakter yang tidak diinginkan, karena karakter yang tidak diinginkan tersebut bisa memperburuk hasil perhitungan, antara lain:

- Tanda baca seperti .,#!@& dihilangkan karena dianggap kurang memiliki makna[5]. Jika tanda baca seperti titik dibiarkan, maka tanda titik itu sendiri dapat meningkatkan kemiripan antar dokumen yang sama sama memiliki banyak tanda titik. Selain itu apa bila ada tanda titik yang menempel pada suatu kata, maka kata itu dianggap berbeda meskipun dengan kata yang sama namun tidak memiliki titik.
- Tulisan DESCRIPTION yang secara default sudah ada pada dataset juga dihilangkan.
- Konversi ke huruf kecil agar dataset seragam yang mana kata yang sama dan memiliki makna yang sama akan tertulis sama, yakni sama sama huruf kecil.

Untuk melakukan pra-pemrosesan penulis membuat tool mandiri dengan bahasa PHP sebagai berikut.

```
$buku = [];
$sambil = $db->query("SELECT * FROM buku");
while($tiap = $sambil->fetch_assoc()){
    $deskripsi = preg_replace('/[^A-Za-z]/', " ", $tiap['deskripsi']);
    $deskripsi = str_replace("DESCRIPTION", "", $deskripsi);
    $deskripsi = strtolower($deskripsi);
    if (!empty(trim(trim($deskripsi)))) {
        $id =$tiap['id'];
        $db->query("UPDATE buku SET proses_deskripsi='$deskripsi' WHERE
id='$id'");
    }
}
```

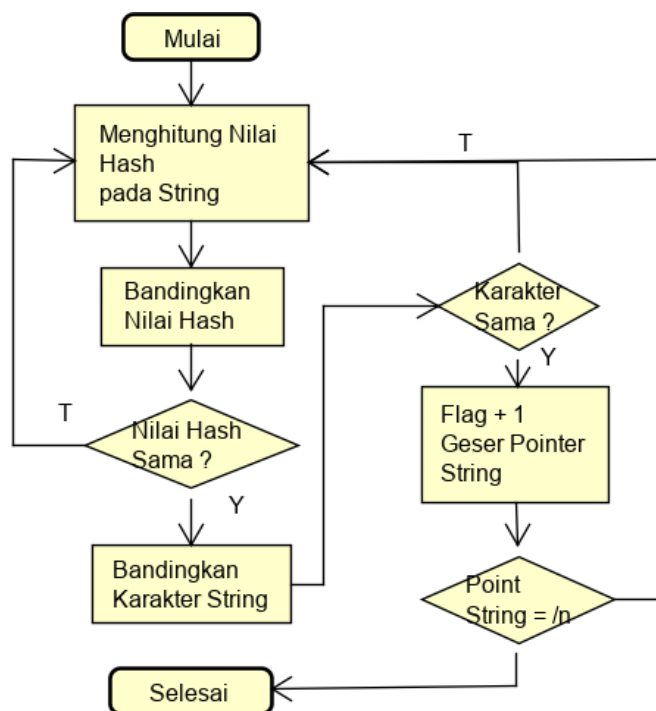
Gambar 2. Tool pra-pemrosesan

Pada proses ini, data-data yang diambil dari kolom deskripsi dan telah diperbaiki disimpan juga pada kolom *proses_deskripsi* di basis data MySQL, tujuannya agar data asli yang diperoleh dari proses *scrapping* tidak hilang dan penulis dapat melihat perbedaan dataset yang belum dan sudah dilakukan pra-pemrosesan. Deskripsi buku yang diperoleh umumnya berisi kalimat yang menggunakan bahasa Indonesia baku, sehingga tidak perlu dilakukan proses *word_normalizer* sebagai perbaikan kata dari istilah slang dan singkatan. Selain itu istilah asing yang ditemukan dalam dataset juga tetap disimpan dan tidak diterjemahkan ke dalam bahasa Indonesia.

2.3. Algoritma Rabin-Karp

Algoritma Rabin-Karp ditemukan oleh Michael O. Rabin dan Richard M. Karp. Teori ini jarang digunakan untuk mencari kata tunggal, namun cukup penting dan sangat efektif bila digunakan untuk pencarian jamak. Algoritma Rabin-Karp merupakan algoritma penggalan data dalam pencarian pola text, Algoritma ini dapat membantu pencarian kemiripan dokumen dengan penerapan metode rolling hash menggunakan kgram dan perhitungan nilai similarity. Algoritma Rabin-Karp adalah algoritma pencocokan

string yang menggunakan fungsi hash sebagai pembanding antara string yang dicari (m) dengan substring pada teks (n)[6]. Cara kerja Rabin-Karp adalah sebagai berikut yang disajikan dalam gambar 3 berikut.



Gambar 3. Alur kerja algoritma Rabin-Karp

Pada dasarnya, algoritma Rabin-Karp akan membandingkan nilai *hash* sebanyak n karakter dari *string* masukan dan *substring* pada teks. Apabila sama, maka akan dilakukan perbandingan sekali lagi terhadap karakter-karakternya. Apabila tidak sama, maka *substring* akan bergeser ke kanan. Kunci utama performa algoritma ini adalah perhitungan yang efisien terhadap nilai *hash substring* pada saat penggeseran dilakukan[7]. Nilai *hash* diperoleh dengan cara konversi huruf A sampai Z menjadi kode ASCII. Rabin Karp dipilih penulis karena merupakan algoritma *multiple pattern search* yang efektif untuk pencarian string dengan pola yang banyak sehingga waktu dan keakuratan pencarian string menjadi lebih baik[8]. Pada penelitian lain sebelumnya, Rabin-Karp memiliki unjuk kerja yang lebih baik dibandingkan algoritma Winnowing dalam kasus mendeteksi kemiripan tugas mahasiswa [9].

3. Hasil dan Pembahasan

3.1. Hasil Dataset

Dataset yang diperoleh dari proses scrapping sebanyak 201 baris data yang tersimpan di MySQL. Data yang diperoleh memiliki atribut id, url, judul, deskripsi, penerbit dan harga. Pengambilan data *url* dirasa penting karena dengan adanya data *url*, *tool scrapping* dapat memasuki *url* atau halaman data tersebut untuk mendapatkan detail deskripsinya. Selain itu, *url* digunakan saat menyajikan hasil rekomendasi pada saat implementasi di antar muka web bukukita.com Adapun dataset yang diperoleh seperti berikut.

Tabel 1. Hasil Dataset

id	url	deskripsi
1	..	DESCRIPTION Sebagai salah satu bahasa pemrograman web yang paling populer di dunia ini...
2	..	DESCRIPTION
3	..	DESCRIPTION Unggul, Handal, Cepat dan Responsif, inilah sistem operasi terbaru dari Microsoft, Windows 8.1!Banyak perubahan yang diusung di wmdows Windows 8.1...
4	..	DESCRIPTION
..
201	..	DESCRIPTION Buku ini merupakan alternatif berbagai kalangan mulai dari pemula maupun,yang telah mahir ...

Pada tabel 1 kolom url dan penerbit tidak penulis tampilkan untuk mempermudah penyajian data, pada kolom deskripsi yang mana kolom ini yang akan digunakan sebagai bahan rekomendasi, terdapat tulisan DESCRIPTION, tanda baca dan huruf besa kecil sehingga perlu dilakukan pra-pemrosesan.

3.2. Hasil Pra-pemrosesan

Pra-pemrosesa pertama yang penulis lakukan adalah elemintasi data yang kosong deskripsinya, sehingga dari 201 baris data tersisa 113 baris data yang dipilih. Data yang dipilih tersebut dilakukan pra pemrosesan selanjutnya dengan dibuang tulisan DESCRIPTION, tanda baca, angka dan konversi ke huruf kecil semua sehingga dataset akhir menjadi sebagai berikut.

Tabel 2. Hasil Pra-pemrosesan

Deskripsi	Pra-pemrosesan deskripsi
DESCRIPTION Sebagai salah satu bahasa pemrograman web yang paling populer di dunia ini...	sebagai salah satu bahasa pemrograman web yang paling populer di dunia..
DESCRIPTION Unggul, Handal, Cepat dan Responsif, inilah sistem operasi terbaru dari Microsoft, Windows 8.1 ...	unggul handal cepat dan responsif inilah sistem operasi terbaru dari microsoft windows ..
DESCRIPTION Amazon Web Services (AWS) adalah layanan cloud computing yang menguasai dunia...	amazon web services aws adalah layanan cloud computing yang menguasai dunia..
DESCRIPTION Buku ini merupakan alternatif berbagai kalangan mulai dari pemula maupun yang telah mahir..	buku ini merupakan alternatif berbagai kalangan mulai dari pemula maupun yang telah mahir..
DESCRIPTION SINOPSISBuku ini mengenalkan bagian dari sebuah pembentukan pemograman web dan menyajikan langkah-langkah perogram yang disusun secara terstruktur. dengan adanya langkah-langkah perogram tersebut,	buku ini mengenalkan bagian dari sebuah pembentukan pemograman web dan menyajikan Langkah langkah program yang disusun secara terstruktur dengan adanya langkah langkah perogram tersebut
DESCRIPTION Wahai orang-orang yang beriman! Makanlah dari benda-benda yang baik (yang halal) yang telah kami berikan kepada kamu dan bersyukurlah kepada Allah, jika betul kamu hanya beribadat kepada-Nya.	wahai orang orang yang beriman makanlah dari benda-benda yang baik yang halal yang telah kami berikan kepada kamu dan bersyukurlah kepada allah jika betul kamu hanya beribadat kepada nya
DESCRIPTION Dalam Islam setiap melakukan apapun pasti ada doanya. Kamu bangun tidur ada doanya, kamu masuk kamar mandi ada doanya, kamu bahkan mau makan, mengenakan pakaian, hingga tidur lagi, ada doanya. Setiap hari dalam kegiatan	dalam islam setiap melakukan apapun pasti ada doanya kamu bangun tidur ada doanya kamu masuk kamar mandi ada doanya kamu bahkan mau makan mengenakan pakaian hingga tidur lagi ada doanya setiap hari dalam kegiatan
DESCRIPTION Ketika media dan sastra Barat cenderung memberikan ilustrasi negative yang berlebihan terhadap orang-orang Arab	ketika media dan sastra barat cenderung memberikan ilustrasi negative yang berlebihan terhadap orang orang Arab

Penghapusan “DESCRIPTION” pada tahap ini tentu membuat dataset menjadi lebih *original*, karena tulisan “DESCRIPTION” merupakan tanda antar muka yang terbawa saat proses pengambilan data. Selain itu dengan menjadikan huruf ke *lowercase* semua, maka kata yang sama menjadi seragam, sama sama huruf kecil sehingga *hash* yang dihasilkan akan sama persis.

3.3. Pengujian Rabin-Karp

Agar setiap kategori buku, baik kategori agama maupun komputer memiliki keterwakilan, maka penulis mengambil sampel pengujian masing-masing kategori 2 buku. Sehingga total ada 4 kali pengujian yang mana tiap pengujian akan menghasilkan 3 rekomendasi sehingga diperoleh hasil akhir 12 rekomendasi. Setiap pengujian yang menghasilkan 3 rekomendasi akan diurutkan berdasarkan skor kemiripan yang paling tinggi, semakin tinggi skornya artinya semakin mirip dan semakin direkomendasikan. Dalam pengujian Rabin-Karp ini, kata yang terdaftar dalam *stopword* tidak dihitung, dikarenakan *stopword* merupakan kosa kata yang kurang bermakna seperti kata hubung dan, atau, yang serta kata ganti orang seperti aku, engkau dan dia[10]. Penggunaan *stopword* dapat mempengaruhi hasil *similarity* Rabin-Karp khususnya dalam kecepatan proses, karena dengan membuang kata yang tidak perlu maka jumlah kata yang harus dibandingkan menjadi berkurang. Untuk memudahkan penyajian data penulis hanya menampilkan judul buku. Adapun hasil pengujian Rabin-Karp disajikan dalam tabel 3 berikut ini.

Tabel 3. Hasil pengujian Rabin-Karp

1	Halal Haram Dalam Islam oleh Muhammad Qasim Kamil	a. Setiap Hari Bisa Menghafal Dzikir&Doa (2015) oleh Ummv Zahirah 32.71 b. Risalah Shalat dan Majmu Syarir Superlengkap oleh IBNU WATINIYAH 31.33
---	---	--

		c. Berdoalah Dalam Mewujudkan Impian-Impian Anda (2013) oleh Khairunnas Abdulhaque Al-Bantanie 30.77
2	Setiap Hari Bisa Menghafal Dzikir&Doa (2015) oleh Ummv Zahirah	a. Risalah Shalat dan Majmu Syarir Superlengkap oleh IBNU WATINIYAH 41.05 b. Halal Haram Dalam Islam oleh Muhammad Qasim Kamil 39.78 c. Sabar Kemudian Syukur Lalu Ikhlas RECOMMENDED oleh Retno D. N 37.06
3	Arab dalam Novel oleh Roger Allen	a. Risalah Shalat dan Majmu Syarir Superlengkap oleh IBNU WATINIYAH 26.97 b. Kumpulan Khotbah Jumat Terbaik Sepanjang Tahun Hijriah - Hard Cover oleh Abdul Bakir, S.Ag 24.31 c. Mendesain Animasi Gif untuk Banner Web oleh Ganesh Progress 22.36
112	HTML, PHP, dan MySQL untuk Pemula oleh Gregorius Agung P	a. Aplikasi Web Database Dengan Dreamweaver MX 2004 oleh Stendy B. Sakur 52.11 b. Buku Pintar Pemrograman Web oleh Adhi Prasetyo 47.11 c. Buku Sakti Webmaster oleh Adhi Prasetyo 46.36
113	Buku Sakti Pemrograman Web HTML, CSS, PHP, MYSQL & JAVASCRIPT oleh Didik Setiawan	a. Membangun Web Store dengan Joomla oleh Adhicipta R. Wirawan 49.03 b. EAZYCOM - Student Exercise Series Pemrograman Web dengan HTML, CSS, dan JavaScript oleh Arief Ramadhan, S.Kom 45.91 c. Otodidak Desain dan Pemrograman <i>Website</i> oleh Jubilee Enterprise 45.50

Sistem rekomendasi ini hanya bekerja dengan cara mencari kemiripan deskripsi buku, sistem rekomendasi ini tidak memeriksa terkait ketersediaan stok buku dan tingkat interaksi pengguna pada suatu buku. Hal ini bertujuan agar pengguna mendapat informasi rekomendasi buku yang utuh karena memang memiliki kesamaan deskripsi buku.

Hasil kerja algoritma Rabin-Karp berupa skor kemiripan beberapa buku terhadap suatu buku yang sedang didetailkan, kemudian beberapa buku tersebut menjadi kandidat rekomendasi yang diurutkan berdasarkan skor tertinggi, setelah diurutkan tidak semua buku diambil sebagai rekomendasi, hanya 3 buku dengan skor tertinggi saja yang digunakan. pada penelitian ini penulis menggunakan skor kemiripan untuk mencari yang tertinggi nilainya dan semakin tinggi semakin baik, sedangkan pada penelitian lain bidang deteksi *plagiarism* skor kemiripan yang dicari adalah yang terendah[1]. Meskipun penelitian ini melakukan penghapusan *stopword* dan penelitian sebelumnya menyebutkan algoritma Rabin-Karp memiliki kecepatan yang baik daripada algoritma lainnya, namun penelitian ini tidak mengukur seberapa cepat proses perhitungan rekomendasi buku meskipun berbagai penelitian sebelumnya menyebutkan Rabin-Karp dapat bekerja maksimal[11].

Pada tabel 3, penulis mengambil sampel 4 hasil test dari total 113 hasil test. Yang dipilih adalah 2 buku untuk kategori agama dan 2 buku untuk kategori komputer. Setiap hasil rekomendasi diurutkan dari skor kemiripan tertinggi, sebagai contoh pada buku halal haram dalam islam oleh Muhammad Qasim Kamil yang memiliki 3 rekomendasi buku antara lain buku setiap hari bisa menghafal dzikir&doa dengan skor 32.71, buku risalah shalat majmu syarir dengan skor 31.11 dan berdoalan dalam segala Impian anda dengan skor 30.77. Dalam hal ini buku setiap hari bisa menghafal dzikir & doa mendapatkan nilai tertinggi karena kesamaan kosa kata yang mirip dengan buku halal haram dalam islam terbanyak yakni 127 kosa kata pada deskripsinya. Pada kategori agama, baik di pengujian 1 maupun 2, skor kemiripan tertinggi diperoleh buku risalah shalat dan masmu syarir dengan skor kemiripan sebesar 41.05 terhadap buku setiap hari bisa menghafal dzikir & doa.

Seperti pada buku agama, penulis juga menguji buku kategori komputer, pada pengujian ke 112 buku HTML, PHP dan MySQL untuk pemula menghasilkan rekomendasi buku yang berjudul aplikasi web database dengan dreamweaver MX 2004. Buku ini menjadi terrekomenasi karena memiliki skor tertinggi yakni 52.11 meskipun pada judul kedua buku ini sama sekali tidak ada kata yang sama. Hal yang sama terjadi pada pengujian ke 113, yang mana pada judul buku student exercise series pemrograman web dengan HTML, CSS, dan JavaScript oleh Arif Ramadhan S.Kom memiliki banyak kesamaan kata dengan judul buku sakti pemrograman web HTML, CSS, PHP, MySQL & JavaScript oleh Didik Setiawan, namun kesamaan judul tidak diperhitungkan dalam penelitian ini sehingga buku tersebut menjadi rekomendasi

kedua dengan skor 45.91 setelah buku membangun web store dengan Joomla oleh Adhicipta R. Wirawan yang memiliki skor 49.03.

Dari 113 hasil testing, ditemukan hasil rekomendasi yang kategorinya berbeda dari suatu buku yang didetailkan. Seperti pada pengujian ke 23 yang mana buku AL-MATSURAT WAZIFAH KUBRA (Jenis 1) pelangi PROMO yang masuk ke dalam kategori agama, menghasilkan 3 rekomendasi yang salah satu rekomendasinya berupa buku kategori komputer dengan judul Hacking & Programming dengan Android SDK untuk Advanced oleh Zainal Arifin. Namun hasil yang tidak akurat ini terjadi pada rekomendasi ke 2 atau 3 dari 3 rekomendasi. Begitu juga pada pengujian ke 68 dengan buku yang berjudul Mudah membangun web profile menghasilkan rekomendasi ke dua yang merupakan buku dengan kategori agama dengan judul Tangan Tuhan di Setiap Fenomena karya Quraish Shihab. Dari 113 pengujian tidak terjadi perbedaan kategori menjadi rekomendasi nomor 1.

3.4. Implementasi Halaman Detail Buku

Sebagai implementasi, penulis memodifikasi halaman detail buku yang sudah ada pada *website* bukukita.com. Halaman ini digunakan untuk melihat detail buku apapun yang diklik oleh pengguna *website* untuk melihat informasi lebih lanjut suatu buku. Pada halaman ini juga terdapat data deskripsi buku yang menjadi dataset penelitian ini. Pada bagian ini, penulis tidak banyak merubah tata letak dan tema antar muka yang ada agar pengguna tetap familiar dan nyaman. Adapun halaman detail buku yang dimaksud disajikan dalam gambar 4 berikut.



Buku Sakti Pemrograman Web HTML, CSS, PHP, MYSQL & JAVASCRIPT

Rp 69.500
Hemat Rp 10.425
Rp 59.075

Judul	Buku Sakti Pemrograman Web HTML, CSS, PHP, MYSQL & JAVASCRIPT
Penulis	Didik Setiawan
No. ISBN	9786026673343
Penerbit	Start Up
Tanggal terbit	Juni - 2020
Jumlah Halaman	-
Berat	250 gr
Jenis Cover	soft cover
Dimensi(L x P)	-
Kategori	Web Programming
Bonus	-
Text Bahasa	Indonesia
Lokasi Stok	Gudang Penerbit (?)

Stok Tidak Tersedia

Beritahu Saya bila akan tersedia

Gambar 4. Halaman detail buku

Halaman detail buku memberikan informasi berupa foto sampul buku pada bagian kiri dan di bagian kanannya berisi meta data buku seperti judul, harga, berat, tanggal terbit, kategori, Bahasa dan ketersediaan stok. Pada bagian bawah halaman ini juga penulis sajikan deskripsi buku dan diikuti dengan buku lain yang direkomendasikan.

3.5. Implementasi Rekomendasi Buku

Implementasi bagian rekomendasi buku masih berada satu halaman dengan detail buku, setiap 1 buku yang didetailkan, maka muncul 3 buku lain yang direkomendasikan. Adapun implementasinya seperti pada gambar 5 berikut.

Rekomendasi Buku Serupa



Membangun Web Store dengan Joomla!
oleh Adhicipta R. Wirawan

Rp. 34,020



EAZYCOM - Student Exercise Series
Pemrograman Web dengan HTML, CSS,
dan JavaScript oleh Arief Ramadhan,
S.Kom

Rp. 17,820



Otodidak Desain dan Pemrograman
Website oleh Jubilee Enterprise

Rp. 38,520

Gambar 5. Rekomendasi buku

Setiap detail buku disajikan 3 buku lain yang menjadi rekomendasi, urutan penyajian berdasarkan skor *similarity* hasil perhitungan algoritma Rabin-Karp. Agar tampilan tetap menyatu dengan tema utama *website*, maka penulis juga menampilkan data gambar dan harga yang tertera pada buku yang mana kedua data tersebut diperoleh bersamaan saat proses *scraping*. Meskipun rekomendasi buku diurutkan berdasar skor *similarity*, namun skor tersebut tidak ditampilkan agar tidak membingungkan pengguna yang mengakses halaman ini.

4. Kesimpulan

Penerapan Sistem rekomendasi buku memberikan manfaat secara langsung bagi pengguna toko BukuKita, yang mana pengguna ketika melihat suatu buku, maka direkomendasikan buku lainnya yang memiliki kesamaan deskripsi yang diurutkan berdasarkan tingkat kemiripannya. *Tool scraping* menjadi alat yang dapat bekerja secara optimal untuk pengambilan data bagi *platform* yang tidak mendapatkan akses database untuk pengambilan data. Pra-pemrosesan menjadi bagian penting agar data lebih seragam dan terhindar dari karakter yang tidak perlu seperti tanda baca dan *stopword*. Algoritma Rabin-Karp dapat digunakan untuk mendapatkan *similarity* atau kemiripan suatu dokumen yang dalam hal ini adalah kemiripan deskripsi buku. Penulis menyarankan untuk penelitian selanjutnya hasil algoritma Rabin-Karp perlu dibandingkan kinerjanya dengan algoritma lain pada dataset yang sama agar rekomendasi buku yang dihasilkan benar-benar optimal. Selain itu, dataset juga perlu ditambahkan agar hasil penelitian lebih komprehensif.

Daftar Pustaka

- [1] A. Aldian and M. Mubarak, "Implementasi Algoritma Rabin-Karp Untuk Pendeteksian Plagiarisme Pada File Dokumen Berupa Text Berbasis Web," *J. Inf. Syst. Res.*, vol. 3, no. 3, pp. 150–154, 2022, doi: 10.47065/josh.v3i3.1404.
- [2] A. Subadri and I. Pratama, "Sistem Deteksi Plagiarism Pada Judul Tugas Akhir Menggunakan Metode Rabin-Karp Berbasis Web," *Technol. J. Ilm.*, vol. 13, no. 4, p. 306, 2022, doi: 10.31602/tji.v13i4.7786.
- [3] M. Alkaff, H. Khatimi, and A. Eriadi, "Sistem Rekomendasi Buku pada Perpustakaan Daerah Provinsi Kalimantan Selatan Menggunakan Metode Content-Based Filtering," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 20, no. 1, pp. 193–202, 2020, doi: 10.30812/matrik.v20i1.617.
- [4] A. Saiful, "Prediksi Harga Rumah Menggunakan Web Scrapping dan Machine Learning Dengan Algoritma Linear Regression," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 1, pp. 41–50, 2021, doi: 10.35957/jatisi.v8i1.701.

-
- [5] A. Agustian, T. Tukiro, and F. Nurapriani, “Penerapan Analisis Sentimen Dan Naive Bayes Terhadap Opini Penggunaan Kendaraan Listrik Di Twitter,” *J. TIKA*, vol. 7, no. 3, pp. 243–249, 2022, doi: 10.51179/tika.v7i3.1550.
- [6] M. O. Pohan, Mesran, and R. D. Sianturi, “Penerapan Algoritma Rabin-Karp Pada Pencarian Sinonim Kata,” *J. Informatics Manag. Inf. Technol.*, vol. 2, no. 1, p. 17, 2022, [Online]. Available: <https://hostjournals.com/>
- [7] S. Yuliyanti and Rizky, “Implementasi Algoritma Rabin Karp Untuk Mendeteksi Kemiripan Dokumen Stmik Bandung,” *J. Bangkit Indones.*, vol. 10, no. 02, p. 1, 2020, doi: 10.52771/bangkitindonesia.v10i02.124.
- [8] S. Fatonah, A. Hadinegoro, A. Hadinegoro, A. D. Hartanto, and A. D. Hartanto, “Deteksi Kemiripan Abstraksi Tugas Akhir Diploma Informatika Universitas AMIKOM Yogyakarta dengan Algoritma Rabin Karp,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 1, p. 1, Feb. 2020, doi: 10.30865/jurikom.v7i1.1927.
- [9] I. Bagus and K. Surya, “KOMPARASI ALGORITMA WINNOWING DAN RABIN KARP MENDETEKSI KEMIRIPAN TUGAS MAHASISWA,” *J. Teknol. Inf. dan Komput.*, vol. 7, pp. 345–351, 2021.
- [10] A. Santosa, I. Purnamasari, and Mayasari Rini, “Pengaruh Stopword Removal dan Stemming Terhadap Performa Klasifikasi Teks Komentar Kebijakan New Normal Menggunakan Algoritma LSTM,” *J. Sains Komput. Inform.*, vol. 6, pp. 81–93, 2022.
- [11] M. F. Asnawi and Zaenal Abidin, “Sistem Pendeteksi Kemiripan Pada Proposal Pengajuan Tugas Akhir Menggunakan Algoritma Rabin-Karp,” *JAMI J. Ahli Muda Indones.*, vol. 2, no. 1, pp. 73–82, 2021, doi: 10.46510/jami.v2i1.61.