

# Implementasi Metode Content Based Filtering Pada Pemilihan Komik

Arba'i Kurniaji<sup>1\*</sup>, Rina Candra Noor Santi<sup>2</sup>

<sup>1,2</sup> Teknik Informatika, Fakultas Teknologi Informasi dan Industri, Universitas Stikubank Semarang  
Jl. Tri Lomba Juang, Mugassari, Kota Semarang, Indonesia

e-mail: <sup>1</sup>akaaji02@gmail.com, <sup>2</sup>r\_candra\_ns@edu.unisbank.ac.id

---

Informasi Artikel      Diterima: 13-06-2023      Direvisi: 06-07-2023      Disetujui: 10-07-2023

---

## Abstrak

Komik adalah rangkaian gambar dan kata yang disusun dengan tujuan untuk menyampaikan informasi kepada orang yang membacanya. Selain dalam bentuk buku saat ini komik memiliki bentuk digital yang dapat dibaca melalui smartphone atau komputer. Komik memiliki beragam alur cerita dan genre, mulai dari aksi, petualangan dan lain-lain. Selain itu setiap negara memiliki komik dengan gaya yang berbeda. Banyaknya komik yang ada saat ini serta memiliki jenis, alur dan genre yang berbeda menjadi sebuah masalah bagi orang yang ingin membaca komik. Permasalahan ini dapat diselesaikan dengan membuat suatu sistem rekomendasi. Pada penelitian ini akan menggunakan metode *Content Based Filtering* untuk mendapatkan hasil rekomendasi. Dalam metode ini menggunakan metode TF-IDF untuk melakukan pembobotan dan *Cosine Similarity* untuk mencari kemiripan komik. Metode ini dipilih karena melihat kebiasaan pembaca komik yang sering membaca komik sesuai dengan yang pernah mereka baca sebelumnya. Kriteria yang akan digunakan pada penelitian ini adalah sinopsis, genre, dan komikus. Jumlah data yang digunakan pada penelitian ini sebanyak 200 komik. Hasil dari penelitian ini mengambil 10 komik untuk dijadikan rekomendasi. Pada pengujian yang telah dilakukan dengan menggunakan komik Ao Ashi sebagai komik preferensi mendapatkan nilai kemiripan sebesar 0.13901. Hasil rekomendasi pada peringkat pertama memiliki nilai kemiripan sebesar 0.03901 dengan persentase kemiripan 100% dan pada peringkat terakhir sebesar 0.09046 dengan persentase kemiripan 65.08%. Rata-rata nilai kemiripan yang dihasilkan adalah 0.10604 dengan persentase kemiripan sebesar 76.38%. Dengan nilai dan persentase yang didapatkan menunjukkan bahwa sistem berjalan dengan baik dan menghasilkan rekomendasi yang sesuai preferensi pengguna.

**Kata Kunci:** Komik, *Content Based Filtering*, Sistem Pendukung Keputusan

## Abstract

Comics are a series of images and words arranged with the aim of conveying information to people who read them. Apart from being in book form, comics now have a digital form that can be read via smartphone or computer. Comics have a variety of storylines and genres, ranging from action, adventure and others. In addition, each country has comics with different styles. Many comics that exist today and have different types, flows and genres become a problem for people who want to read comics. This problem can be solved by creating a recommendation system. This research will use the *Content Based Filtering* method to get recommendation results. This method uses the TF-IDF method to do weighting and *Cosine Similarity* to find comic similarities. This method was chosen because it looks at the habits of comic readers who often read comics according to what they have read before. Criteria that will be used in this research are synopsis, genre, and comic artist. Total data used in this study were 200 comics. The results of this study took 10 comics to be used as recommendations. In tests that have been carried out using Ao Ashi comics as preference comics get a similarity value of 0.13901. The recommendation results in the first rank have a similarity value of 0.03901 with a similarity percentage of 100% and in the last rank of 0.09046 with a similarity percentage of 65.08%. Average similarity value generated is 0.10604 with a similarity percentage of 76.38%. With the values and percentages obtained, it shows that the system runs well and produces recommendations that match user preferences.

**Keywords:** Comic, *Content Based Filtering*, Decision Support System



## 1. Pendahuluan

Komik adalah kumpulan gambar-gambar yang disusun secara teratur sehingga memiliki alur cerita. Secara garis besar komik adalah sebuah buku cerita bergambar (Soedarso, 2015). Komik juga bisa dikatakan sebagai sastra gambar serta menjadi alat komunikasi visual yang dapat digunakan untuk menyampaikan sebuah informasi agar dapat dengan mudah dimengerti. Komik sendiri juga bisa disebut dengan karya sastra bergambar (Bonnett, 1998). Komik biasanya dicetak pada kertas yang biasa kita lihat pada koran dalam bentuk strip atau dicetak dalam bentuk buku (Maharsi, 2011). Semakin berkembangnya zaman apalagi di era teknologi informasi sekarang ini komik sekarang sudah berbentuk digital. Bentuk digital ini sering kita temui pada *website* dan aplikasi *android*. Tidak hanya bentuk mediana yang mengikuti zaman, sekarang genre, cerita, gambar dan bentuk penyajian dari komik sendiri juga mengikuti perkembangan zaman.

Komik memiliki alur cerita dan genre yang beragama seperti action, historikal, drama, horor, romansa dan masih banyak lagi. Alur cerita dan genre yang beragam ini menjadikan komik memiliki banyak pembaca serta menjadikan komik memiliki batasan usia pembacanya. Teknik penggambaran komik sekarang juga semakin lebih bagus lagi dan memiliki gaya-gaya tersendiri di setiap komikus. Setiap negara memiliki komik dengan gaya penggambaran, cerita dan keunikan tersendiri. Banyak sekali orang yang menjadikan membaca komik sebagai hobi.

Dengan banyaknya komik yang beredar serta memiliki alur cerita dan genre yang beragam menjadi masalah bagi orang yang ingin membaca komik. Hal ini membuat orang yang ingin membaca komik menjadi kesulitan dalam memilih komik yang ingin mereka baca. Kesulitan tersebut membuat pembaca komik akhirnya mencari rekomendasi atau saran bacaan komik di *website* forum yang membahas mengenai komik atau perlu melihat review mengenai komik di sosial media. Dengan begitu akan memakan waktu orang yang ingin mencari saran atau rekomendasi dan membuat orang menjadi bingung untuk menentukan komik yang ingin mereka baca.

Salah satu cara untuk mengatasi masalah ini adalah dengan membuat sistem pendukung keputusan yang dapat merekomendasikan komik berdasarkan preferensi pembaca. Dalam era perkembangan teknologi yang semakin maju ini, hampir semua tugas dan masalah manusia dapat diselesaikan dengan cepat. Dengan membuat sebuah sistem rekomendasi akan membantu orang

yang ingin membaca komik dapat menemukan rekomendasi atau saran komik yang sesuai dengan kesukaan atau preferensi mereka.

Sistem Pendukung Keputusan adalah sebuah sistem yang dapat membantu kita dalam melakukan pengambilan keputusan yang akan kita ambil dalam berbagai kegiatan (Turban, 2003). Sistem Pendukung Keputusan memiliki tujuan untuk memberikan informasi dan membantu pengguna dalam memberikan suatu prediksi sehingga pengguna bisa mengambil keputusan dengan baik (Fathoroni dkk, 2020). Dengan membuat Sistem Pendukung Keputusan orang-orang yang gemar membaca, ingin membaca dan membeli komik dapat dengan mudah menentukan serta menghemat waktu dalam memilih komik. Melihat kebiasaan orang-orang yang membaca komik, mereka biasa membaca komik sesuai dengan preferensi dari komik atau bacaan yang pernah mereka baca sebelumnya. Penelitian ini bertujuan untuk membantu pengguna dalam memilih komik berdasarkan preferensi pengguna. Oleh karena dengan membuat sistem rekomendasi dapat menjadi solusi dalam mengatasi permasalahan ini.

Pada penelitian yang pernah dilakukan dengan judul Implementasi *Machine Learning* dalam Penentuan Rekomendasi Musik dengan Metode *Content-Based Filtering* membahas mengenai berkembangnya musik yang semakin signifikan dan banyaknya *platform* penyedia layanan musik secara *online*. Hasil dari penelitian ini adalah sebuah sistem rekomendasi berbasis android dengan memanfaatkan metode *Content-Based Filtering*. Adapun kriteria yang digunakan isi dari keseluruhan musik untuk mendapatkan hasil rekomendasi (Putra & Santika, 2020).

Penelitian lain yang berjudul Sistem Rekomendasi Film Menggunakan *Content Based Filtering* yang membahas banyaknya produksi film yang ada baik di bioskop maupun di situs streaming serta genre yang beragam. Hasil dari penelitian ini adalah sebuah sistem rekomendasi yang dapat membantu orang-orang dalam memilih film. Hasil rekomendasi yang didapatkan adalah film yang memiliki bobot yang mirip berdasarkan sinopsis dan judul film dari preferensi serta diakhiri dengan melakukan filtering berdasarkan genre (Fajriansyah dkk., 2021).

Penelitian lain yang berjudul Sistem Rekomendasi Buku Menggunakan *Weighted Tree Similarity* dan *Content Based Filtering* membahas mengenai pengunjung yang kesulitan menemukan buku yang sesuai dengan buku yang dulu pernah dibaca. Pada penelitian tersebut menghasilkan sebuah sistem rekomendasi berbasis *website*. Penggunaan metode *Content-Based Filtering* akan

memberikan rekomendasi berdasarkan judul buku, sinopsis dan pengarang yang kemudian akan dianalisis dengan Metode *Weighted Tree Similarity* sehingga mendapatkan hasil rekomendasi yang baik (Alkaff dkk., 2020).

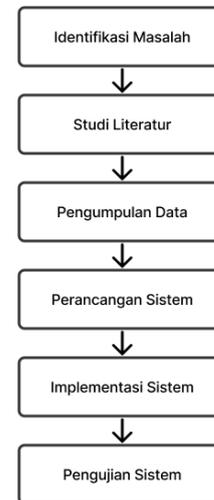
Penelitian lain yang berjudul *Hybrid Recommendation System* untuk Peminjaman Buku Perpustakaan dengan *Collaborative* dan *Content-Based Filtering* membahas mengenai Perpustakaan Universitas Kristen Petra yang memiliki 160.000 dari tahun 2014 dan terus bertambah sehingga menyebabkan pengunjung yang ingin membaca buku lain yang sesuai yang pernah dibaca menjadi memakan banyak waktu. Penggunaan dua metode yaitu *Collaborative Filtering* dan *Content-Based Filtering* menghasilkan hasil rekomendasi yang sesuai dengan kemiripan dan preferensi buku (Widjaja & Palit, 2022).

Penelitian lain yang berjudul Sistem Rekomendasi Film Menggunakan Metode *Hybrid Collaborative Filtering* dan *Content-Based Filtering* membahas mengenai pada saat ini menonton film dan acara TV sangat dimudahkan oleh adanya platform streaming yang memiliki jumlah katalog yang sangat banyak. Dalam penelitian tersebut menggunakan pendekatan metode *hybrid* untuk mendapatkan hasil rekomendasi terbaik. Hasil dari penelitian ini adalah sebuah sistem rekomendasi yang dapat menggunakan metode *Collaborative Filtering*, *Content-Based Filtering* dan *Hybrid* keduanya. Dalam pengujiannya dengan melakukan perbandingan metode *Hybrid Collaborative Filtering* dan *Content-Based Filtering* menghasilkan rekomendasi paling baik (Arfisko & Wibowo, 2022).

Berdasarkan penelitian sebelumnya metode Content Based Filtering dipilih karena melihat kebiasaan dari para pembaca sering membaca komik berdasarkan dengan komik yang mereka sukai dan komik yang mirip dengan yang pernah dibaca sebelumnya. Metode yang akan digunakan adalah Content Based Filtering yaitu sebuah metode yang digunakan dalam proses pemfilteran berdasarkan kemiripan atau hal yang disukai pengguna sesuai aktivitas yang pengguna lakukan. Dengan metode ini sistem akan menebak sesuatu hal yang sesuai dengan kemiripan atau kesukaan user dari aktivitas yang pernah dilakukan (Putri dkk., 2022). Dalam penggunaan metode *Content-Based Filtering* sendiri akan menggunakan metode TF-IDF dalam melakukan pembobotan kata serta metode *Cosine Similarity* dalam mencari kemiripan komik. Adapun kriteria yang akan digunakan adalah sinopsis, komikus dan juga genre dari komik.

## 2. Metode Penelitian

Alur penelitian yang akan dilakukan pada penelitian ini meliputi identifikasi masalah, studi literatur, pengumpulan data, perancangan data, implementasi sistem dan pengujian sistem. Alur penelitian yang akan dilakukan dapat ditunjukkan pada gambar 1 dibawah ini.



Gambar 1. Alur Penelitian

Berikut adalah penjelasan alur penelitian yang akan dilakukan.

### 1. Identifikasi Masalah

Dalam sebuah penelitian tahapan pertama yang dilakukan adalah dengan memahai permasalahan yang akan diteliti. Masalah yang akan diteliti adalah mengenai pemilihan komik yang memiliki berbagai macam jenis.

### 2. Studi Literatur

Tahapan ini adalah mencari informasi dari penelitian-penelitian sebelumnya mengenai masalah yang diteliti. Informasi yang dicari adalah untuk mendapatkan teori dan metode yang akan digunakan.

### 3. Pengumpulan Data

Data yang akan dikumpulkan adalah data komik yang berisi judul, sinopsis, genre, dan komikus. Data tersebut diambil dari website penyedia dataset dan website penyedia daftar komik.

### 4. Perancangan Sistem

Rancangan sistem dibuat agar sistem yang dibangun dapat efektif dan efisien. Sistem digambarkan secara rinci dan dirancang dengan menggunakan pemodelan *Unified Modelling Language*.

### 5. Implementasi Sistem

Sistem yang dirancang akan diimplementasikan dengan melakukan pengkodean. Pada tahapan ini juga metode *Content-Based Filtering* akan diimplementasikan pada sistem. Bahasa pemrograman yang akan digunakan dalam

mengimplemtasikan sistem ini adalah *Python* dan *HTML*.

## 6. Pengujian Sistem

Dalam tahapan ini pengujian akan menguji metode *Content-Based Filtering* yang sudah diimplementasikan serta melakukan analisa terhadap.

### 2.1. Content Based Filtering

*Content Based Filtering* adalah salah satu metode sistem pendukung keputusan yang menggunakan item yang memiliki nilai dari pengguna sebagai preferensi untuk menentukan hasil. Item pada metode ini biasa berupa teks. Salah satu metode yang digunakan pada *Content Based Filtering* adalah TF-IDF (*Term Frequency Inverse Document Frequency*) dan *Cosine Similarity* (Alkaff dkk., 2020).

### 2.2. Term Frequency Inverse Document Frequency

*Term Frequency Inverse Document Frequency* (TF-IDF) adalah nilai statistik yang mencerminkan atau menunjukkan kemiripan suatu term dengan beberapa data dalam dokumen yang membuat term tersebut bisa menjadi kata kunci dari dokumen tersebut. Nilai TF-IDF diambil dari perhitungan *Term Frequency* (TF) yang menggambarkan frekuensi kemunculan istilah atau kata yang muncul dari dokumen yang ada, dan *Inverse Document Frequency* (IDF) yang menggambarkan frekuensi dokumen yang memiliki istilah dalam koleksi (Arfisko & Wibowo, 2022). Berikut adalah rumus dari *Term Frequency*.

$$tf(d, t) = 1 + \log(f_{dt}) \quad (1)$$

Keterangan :

tf (d,t) = Nilai TF untuk kata ke-t dalam dokumen ke-d

f dt = Jumlah kata ke-t dalam dokumen ke-d

Berikut adalah rumus perhitungan dari *Inverse Document Frequency* :

$$IDF(t) = \log \frac{D}{df} \quad (2)$$

Keterangan :

IDF(t) = Nilai IDF untuk kata ke-t pada item

Df = Jumlah dokumen yang mengandung kata ke-t

D = Total jumlah dokumen pada item

Berikut adalah rumus yang digunakan dalam proses TF-IDF :

$$Wd, t = TF(d, t) \times IDF(t) \quad (3)$$

Keterangan :

d = dokumen ke-d

t = kata ke-t dari kata kunci

W = bobot dokumen d terhadap kata ke-t

### 2.3. Cosine Similarity

*Cosine similarity* adalah metode pengukuran kesamaan yang biasa digunakan pada data atau dokumen berupa teks yang mengukur dua vektor dan ruang multidimensional yang mendekati arah yang sama. *Cosine Similarity* akan menghitung kemiripan antara dua data dalam bentuk vektor berdasarkan sudut antara vektor-vektor tersebut. Semakin kecil sudut antara vektor-vektor, maka data-data tersebut memiliki tingkat kemiripan yang tinggi (Lindang dkk., 2023). Berikut adalah rumus dari *Cosine Similarity* :

$$\cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (4)$$

Keterangan :

A = vektor

B = vektor

Ai = bobot term i dalam blok Ai

Bi = bobot term i dalam blok Bi

I = jumlah term dalam kalimat

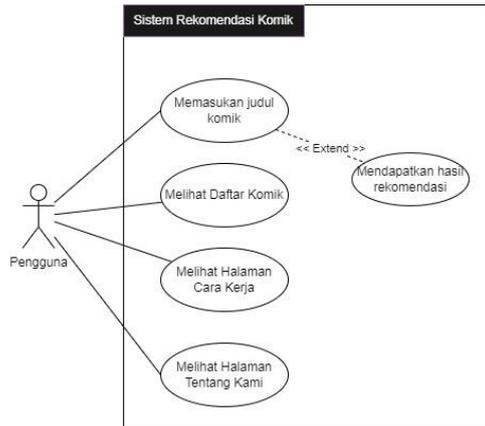
N = jumlah vektor

## 3. Hasil dan Pembahasan

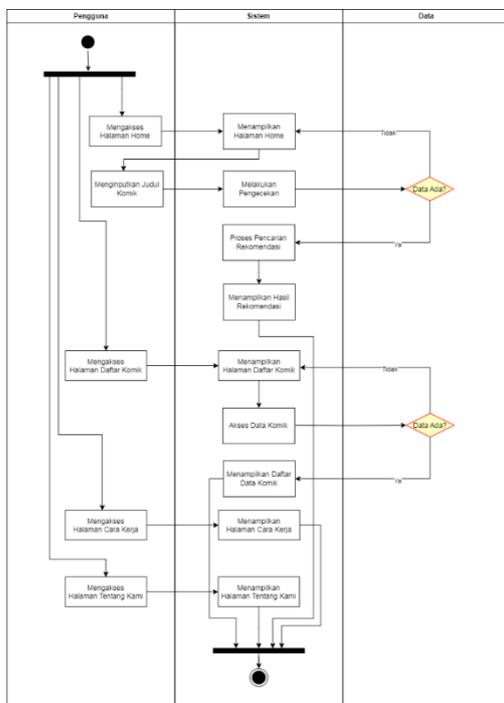
Pada penelitian ini data yang telah dikumpulkan akan disimpan pada sebuah file dengan format *Comma Separated Value* (csv). Jumlah data yang digunakan pada penelitian ini berjumlah 200 data komik. Kolom data yang digunakan sebagai kriteria adalah sinopsis, genre, dan komikus.

Sistem yang akan dibuat pada penelitian ini sebelum diimplementasikan akan dirancang dengan menggunakan *Unified Modeling Language* (UML). *Unified Modeling Language* (UML) adalah sebuah bahasa pemodelan grafis yang digunakan untuk merancang, memodelkan, dan mendokumentasikan sistem perangkat lunak (Christian & Kelvin, 2022).

Gambar 2 menjelaskan use case diagram dari sistem. *Use Case Diagram* adalah sebuah gambaran interaksi antara pengguna dengan sistem. Sistem yang akan dibuat memiliki satu aktor yaitu pengguna. Pengguna memiliki 4 interaksi pada sistem yaitu memasukan judul komik, melihat daftar komik, melihat halaman cara kerja dan melihat halaman tentang kami.



Gambar 2. Use Case Diagram



Gambar 3. Activity Diagram

Pada gambar 3 menjelaskan *activity diagram* keseluruhan sistem rekomendasi komik. Pengguna pertama akan mengakses halaman home kemudian pengguna akan menginputkan judul komik yang ingin dijadikan preferensi untuk mendapatkan hasil rekomendasi. Setelah itu sistem akan melakukan proses untuk mencari hasil rekomendasi yang sesuai dengan preferensi pengguna dengan nilai kemiripan yang sama atau mendekati. Pengguna bisa mengakses halaman daftar komik untuk melihat daftar komik. Selain itu pengguna dapat mengakses halaman cara kerja dan tentang kami yang kemudian akan ditampilkan oleh sistem halaman tersebut.

### 3.1. Proses Metode Content-Based Filtering

Metode *Content-Based Filtering* memiliki beberapa tahapan yang harus dilalui agar dapat mendapatkan hasil rekomendasi. Pada pembuatan suatu sistem rekomendasi tahapan yang harus dilalui yaitu *pre-processing*, pembobotan kata dan normalisasi. Berikut adalah penjelasan pada tahapan *Content-Based Filtering*.

#### 1. Pre-Processing

Tahapan ini bertujuan untuk membersihkan dan mempersiapkan data yang berupa teks untuk dianalisis (Ula dkk, 2021). Data yang akan melalui tahapan ini adalah bagian yang dijadikan kriteria yaitu sinopsis, genre dan komikus. Tahapan *pre-processing* meliputi *clean data*, tokenisasi dan *filtering*. Pada penelitian ini data sinopsis akan terlebih dahulu melalui proses *pre-processing* secara lengkap, hal ini dikarenakan data pada sinopsi sangat banyak dan penting pada tahapan ini. Kemudian pada data genre dan komikus akan melalui tahapan *pre-processing* yang hanya meliputi *clean data* dan tokenisasi pada saat proses penggabungan data. Berikut adalah penjelasan mengenai *clean data*, tokenisasi dan *filtering*.

##### a) Clean Data

Tahapan ini memiliki tujuan untuk membersihkan data dari *noise* seperti mengubah data menjadi huruf kecil, menghilangkan tanda baca, dan menghilangkan angka pada data (Ula dkk, 2021). Tahapan *clean data* pada pembuatan penelitian ini dapat dilihat pada tabel 1.

Sebelum	Sesudah
<i>Guts, a former mercenary now known as the "Black Swordsman," is out for revenge. After a tumultuous childhood, he finally finds someone he respects and .....</i>	<i>guts a former mercenary now known as the black swordsman is out for revenge after a tumultuous childhood he finally finds someone he respects and .....</i>
<i>Gol D. Roger, a man referred to as the "Pirate King," is set to be executed by the World Government. But just before his demise, he confirms the existence .....</i>	<i>gol d roger a man referred to as the pirate king is set to be executed by the world government but just before his demise he confirms the existence .....</i>

##### b) Tokenisasi

Tokenisasi bertujuan untuk mengubah data teks menjadi kata-kata yang terpisah atau memisahkan kata-kata. Kata yang akan dibuat terpisah menjadi kata per kata. Proses tokenisasi ini juga memiliki tujuan agar mempermudah data

teks untuk dihitung term dalam data. Tahapan tokenisasi pada penelitian ini dapat dilihat pada tabel 2.

**Tabel 2. Hasil Proses Tokenisasi**

Sebelum	Sesudah
<i>guts a former mercenary now known as the black swordsman is out for revenge after a tumultuous childhood he finally finds someone he respects and .....</i>	<i>guts, a, former, mercenary, now, known, as, the, black, swordsman, is, out, for, revenge, after, a, tumultuous, childhood, he, finally, finds, someone, he, respects, and ....</i>
<i>gol d roger a man referred to as the pirate king is set to be executed by the world government but just before his demise he confirms the existence .....</i>	<i>gol, d, roger, a, man, referred, to, as, the, pirate, king, is, set, to, be, executed, by, the, world, government, but, just, before, his, demise, he, confirms, the, existence ....</i>

**c) Filtering**

*Filtering* merupakan tahapan yang berguna untuk memfilter atau menghilangkan kata-kata yang tidak dibutuhkan dalam proses analisis dokumen. Dalam penelitian ini akan menggunakan *stopword removal* bahasa inggris dalam proses *filtering* ini. Pada tahapan ini akan menghapus kata-kata yang tidak dibutuhkan pada data yang akan diproses (Alkaff dkk., 2020). Tahapan *filtering* pada penelitian ini dapat dilihat pada tabel 3.

**Tabel 3. Hasil Proses Filtering**

Sebelum	Sesudah
<i>guts, a, former, mercenary, now, known, as, the, black, swordsman, is, out, for, revenge, after, a, tumultuous, childhood, he, finally, finds, someone, he, respects, and ....</i>	<i>guts, former, mercenary, now, known, the, black, swordsman, out, for, revenge, after, tumultuous, childhood, finally, finds, someone, respects, and .....</i>
<i>gol, d, roger, a, man, referred, to, as, the, pirate, king, is, set, to, be, executed, by, the, world, government, but, just, before, his, demise, he, confirms, the, existence .....</i>	<i>gol, roger, man, referred, the, pirate, king, set, executed, the, world, government, but, just, before, his, demise, confirms, the, existence .....</i>

Setelah data sinopsis telah melalui semua tahapan, selanjutnya data genre dan komikus akan melalui tahapan *clean data* dan tokenisasi. Data tersebut tidak melalui tahapan

*filtering* dikarenakan data tersebut tidak perlu ada kata yang dihilangkan. Kemudian data sinopsi, genre dan komikus akan dijadikan menjadi satu untuk dapat dihitung bobot katanya

**2. Pembobotan Kata**

Setelah data telah melalui tahapan *pre-processing* selanjutnya data tersebut akan dihitung bobot kata pada data dengan menggunakan metode TF-IDF. Perhitungan ini mempertimbangkan frekuensi kata pada sebuah data dan frekuensi kata semua data. Proses pertama pada perhitungan TF-IDF adalah menghitung *Term Frequency* (TF) kata pada data. Hasil dari perhitungan *Term Frequency* pada penelitian ini dapat dilihat pada tabel 4.

**Tabel 4. Perhitungan Term Frequency**

Kata	D1	D2	D200	TF
<i>guts</i>	4	0	0	0.0384
<i>former</i>	1	0	0	0.0096
<i>mercena</i>	1	0	0	0.0096
<i>ry</i>				
<i>now</i>	2	0	1	0.0192
...	...	..	...	...
<i>black</i>	1	0	0	0.0096

Setelah mendapatkan nilai *Term Frequency* selanjutnya adalah menghitung *Inverse Document Frequency* (IDF) yaitu seberapa sering kata tersebut muncul pada data. Hasil dari perhitungan *Inverse Document Frequency* pada penelitian ini dapat dilihat pada tabel 5.

**Tabel 5. Perhitungan Inverse Document Frequency**

Kata	IDF
<i>guts</i>	4.6051
<i>former</i>	3.1010
<i>mercenary</i>	3.9120
<i>now</i>	1.4916
...	....
<i>black</i>	3.2188

Setelah mendapatkan nilai dari TF dan IDF, maka selanjutnya melakukan perhitungan untuk TF-IDF untuk setiap kata dalam setiap dokumen. Hasil dari perhitungan TF-IDF pada penelitian ini dapat dilihat pada tabel 6.

**Tabel 6. Perhitungan TF-IDF**

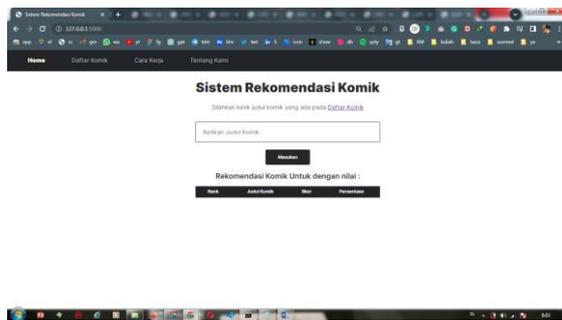
Kata	TF-IDF
<i>guts</i>	0.1771
<i>former</i>	0.0298
<i>mercenary</i>	0.0376
<i>now</i>	0.0286
...	....
<i>black</i>	0.0309

### 3. Normalisasi

Tahapan ini dilakukan untuk menormalisasikan nilai yang didapat pada perhitungan TF-IDF agar memiliki rentan nilai yang sama. Tahapan normalisasi ini akan dihitung menggunakan *Cosine Similarity*. Proses *Cosine Similarity* ini akan dilakukan setelah pengguna menginputkan judul komik yang dijadikan preferensi. Kemudian sistem akan mencari nilai kemiripan yang sama dengan preferensi. Hasil dari proses ini dapat dilihat pada bagaian pengujian.

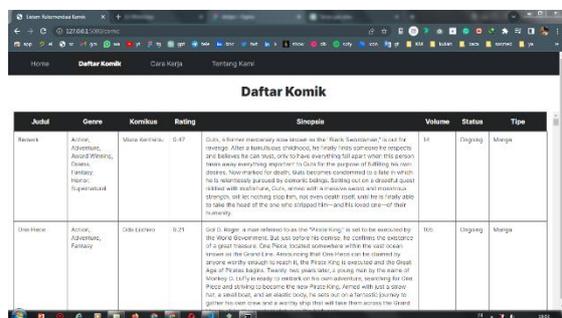
### 3.2. Implementasi Sistem

Setelah mendapatkan nilai dari proses *Content-Based Filtering*, selanjutnya akan diimplementasikan pada *website*. Sistem akan diimplementasikan menggunakan bahasa pemrograman *Python* dan *HTML*. Agar dapat digunakan pada *website* dibantu dengan menggunakan library *Python* yaitu *Flask*



Gambar 4. Halaman Home

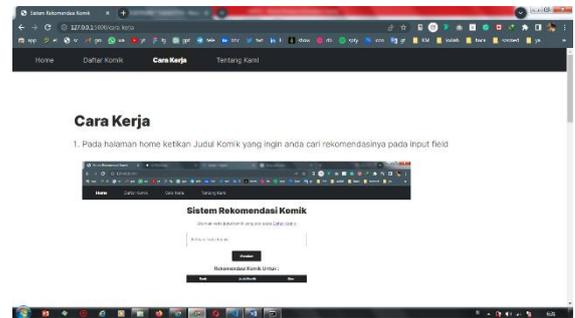
Pada gambar 4 merupakan halaman home atau halaman utama dimana pengguna dapat mencari rekomendasi komik. Setelah memasukkan judul komik pengguna akan mendapatkan hasil rekomendasi mulai dari peringkat, judul komik, skor dan persentase kemiripan.



Gambar 5. Halaman Daftar Komik

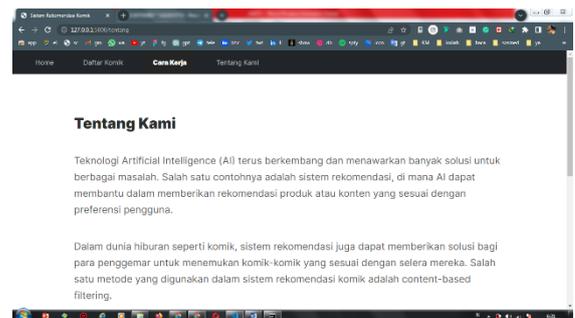
Pada gambar 5 merupakan halaman daftar komik yang dimana berisi list daftar komik yang ada pada sistem. Sehingga sebelum mencari rekomendasi komik pengguna dapat

melihat terlebih dahulu daftar komik yang ada pada sistem. Daftar komik ditampilkan dengan menggunakan tabel.



Gambar 6. Halaman Cara Kerja

Pada gambar 6 merupakan halaman cara kerja yang berisi bagaimana cara menggunakan sistem rekomendasi yang telah dibuat. Halaman ini dilengkapi dengan gambar dan cara-cara penggunaan sistem.



Gambar 7. Halaman Tentang Kami

Pada gambar 7 merupakan halaman tentang kami yang menjelaskan isi atau kegunaan dari sistem rekomendasi yang telah dibuat. Halaman ini juga berisi penjelasan mengenai komik dan metode *Content-Based Filtering*.

### 3.3. Pengujian dan Analisa

Pada tahapan ini akan menguji sistem dengan melakukan percobaan untuk mendapatkan rekomendasi komik serta menganalisis nilai kemiripan. Pada penelitian ini pengujian akan menggunakan komik Ao Ashi sebagai komik preferensi. Langkah yang perlu dilakukan adalah mengetikkan judul komik pada input yang tersedia. Sistem akan mencari nilai kemiripan yang paling mendekati dari komik Ao Ashi. Setelah melakukan perhitungan komik Ao Ashi memiliki nilai *Cosine Similarity* sebesar 0.13901. Kemudian sistem akan mencari 10 komik dengan nilai kemiripan yang sama atau mendekati dari nilai komik Ao Ashi. Hasil rekomendasi dan dari nilai *Cosine Similarity* serta persentase kemiripan pada pengujian dengan

menggunakan komik Ao Ashi sebagai komik preferensi dapat dilihat pada tabel 7.

Tabel 7. Peringkat Hasil Rekomendasi Komik Ao Ashi

Rank	Judul Komik	Cosine Similarity.	Persentase Kemiripan
1	Giant Killing	0.13901	100%
2	Blue Lock	0.12123	87.21%
3	Ariadne in the Blue Sky	0.11194	80.53%
4	Ace of the Diamond	0.10865	78.16%
5	Kuroko's Basketball	0.10375	74.64%
6	Be Blues! Ao ni Nare	0.09921	71.37%
7	Tokyo Ghoul:re	0.09859	70.92%
8	Eyeshield 21	0.09680	69.64%
9	Days	0.09085	65.36%
10	That Time I Got Reincarnated as a Slime	0.09046	65.08%

Berdasarkan tabel 7 dapat dilihat bahwa nilai pada peringkat satu yaitu komik Giant Killing memiliki nilai *Cosine Similarity* yang sama dengan komik preferensi yaitu 0.13901 dengan persentase kemiripan 100% dan pada peringkat sepuluh dengan komik That Time I Got Reincarnated as a Slime memiliki nilai *Cosine Similarity* 0.09046 dengan persentase kemiripan 65.08%. Semua hasil rekomendasi yang didapatkan mendapatkan nilai *Cosine Similarity* yang mendekati atau mirip dengan komik Ao Ashi dan persentase kemiripan yang didapatkan semua diatas 50%. Jika menghitung rata-rata dari hasil rekomendasi mendapatkan nilai 0.10604 sehingga memiliki tingkat persentase kemiripan 76.38%. Hal ini menandakan bahwa hasil rekomendasi yang diberikan memiliki kecocokan dan kemiripan dengan komik preferensi atau Ao Ashi. Dengan persentase kemiripan sebesar 76.38% juga menunjukkan bahwa sistem dapat berjalan dan menghasilkan hasil rekomendasi dengan baik.

#### 4. Kesimpulan

Dari penelitian yang telah dilakukan adalah membuat sebuah sistem rekomendasi komik dengan menggunakan metode Content-Based Filtering. Banyaknya data yang digunakan pada penelitian ini adalah 200 data komik. Kriteria yang digunakan adalah sinopsis, genre dan komikus. Kriteria ini digunakan untuk mencari bobot dan nilai kemiripan antar data. Hasil rekomendasi akan mencari 10 rekomendasi komik berdasarkan komik

preferensi. Berdasarkan percobaan yang telah dilakukan dengan menggunakan komik Ao Ashi sebagai komik preferensi menghasilkan nilai *Cosine Similarity* yaitu 0.13901. Hasil rekomendasi yang didapat komik dengan peringkat pertama memiliki nilai *Cosine Similarity* sebesar 0.13901 yaitu pada komik Giant Killing dengan persentase kemiripan 100% dan pada peringkat sepuluh memiliki nilai sebesar 0.09046 yaitu komik That Time I Got Reincarnated as a Slime dengan persentase kemiripan 65.08%. Pada percobaan yang dilakukan memiliki rata-rata nilai *Cosine Similarity* sebesar 0.10604 dengan persentase kemiripan 76.38%. Dengan persentase kemiripan sebesar 76.38% menunjukkan bahwa hasil rekomendasi cocok dengan komik preferensi yaitu Ao Ashi. Hasil ini juga menunjukkan bahwa sistem yang dibuat berjalan dengan baik dan benar sehingga dapat memberikan rekomendasi yang sesuai dengan preferensi pengguna.

#### Referensi

- Putra, A. I., & Santika, R. R. (2020). Implementasi Machine Learning dalam Penentuan Rekomendasi Musik dengan Metode Content-Based Filtering. *Edumatic: Jurnal Pendidikan Informatika*, 4(1), 121-130. <https://doi.org/10.29408/edumatic.v4i1.2162>
- Fajriansyah, M., Adikara, P. P., & Widodo, A. W. (2021). Sistem Rekomendasi Film Menggunakan Content Based Filtering. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN, 2548, 964X*. <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/9163>
- Alkaff, M., Khatimi, H., & Eriadi, A. (2020). Sistem Rekomendasi Buku Menggunakan Weighted Tree Similarity dan Content Based Filtering. *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput*, 20(1), 193-202. <https://doi.org/10.30812/matrik.v20i1.617>
- Arfisko, H. H., & Wibowo, A. T. (2022). Sistem Rekomendasi Film Menggunakan Metode Hybrid Collaborative Filtering Dan Content-Based Filtering. *eProceedings of Engineering*, 9(3). <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/18066>
- Widjaja, A. A., & Palit, H. N. (2022). Hybrid Recommendation System untuk Peminjaman Buku Perpustakaan dengan Collaborative dan Content-Based Filtering. *Jurnal Infra*, 10(2), 1-6. <https://publication.petra.ac.id/index.php/teknik-informatika/article/view/12512>

- Christian, Y., & Kelvin, K. (2022). RANCANG BANGUN APLIKASI KURSUS ONLINE BERBASIS WEB DENGAN SISTEM REKOMENDASI METODE CONTENT-BASED FILTERING. *Rabit: Jurnal Teknologi dan Sistem Informasi Univrab*, 7(1), 23-36.  
<https://doi.org/10.36341/rabit.v7i1.2181>
- Nastiti, P. (2019). Penerapan Metode Content Based Filtering dalam Implementasi Sistem Rekomendasi Tanaman Pangan. *Teknika*, 8(1), 1-10.  
<https://doi.org/10.34148/teknika.v8i1.139>
- Putri, I. H., Nurakhmadyavi, S. M. K. H., & Wahyudi, E. E. (2022, April). Literature Review: Sistem Rekomendasi untuk Buku dan Film. In *Prosiding Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika (SNESTIK)* (Vol. 1, No. 1, pp. 75-81).  
<https://doi.org/10.31284/p.snestik.2022.2635>
- Ula, N., Setianingsih, C., & Nugrahaeni, R. A. (2021). Sistem Rekomendasi Lagu Dengan Metode Content Based Filtering Berbasis Website. *eProceedings of Engineering*, 8(6).  
<https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/17229>
- Lindang, D. N., Muniar, A. Y., Halid, A., Muhajirin, M., & Amiruddin, A. (2023, February). Sistem Penentuan Kemiripan Antar Skripsi Menggunakan Metode Cosine Similarity Pada Perpustakaan. In *Seminar Nasional Teknik Elektro dan Informatika (SNTEI)* (Vol. 8, No. 1, pp. 321-324).  
<http://118.98.121.208/index.php/sntei/article/view/3640>
- Fathoroni, A., Fatonah, R. N. S., Andarsyah, R., & Riza, N. (2020). Buku tutorial sistem pendukung keputusan penilaian kinerja dosen menggunakan metode 360 degree feedback. CV. Kreatif Industri Nusantara.
- Turban. 2001. *Decision Support System and intelligent system (Sistem Pendukung Keputusan dan Sistem Cerdas)*. Yogyakarta, Andi
- Limbong, T., Muttaqin, M., Iskandar, A., Windarto, A. P., Simarmata, J., Mesran, M., ... & Wanto, A. (2020). *Sistem Pendukung Keputusan: Metode & Implementasi*. Yayasan Kita Menulis.
- Maharsi, I. (2011). *Komik. Dwi-Quantum*. (kuning)
- Soedarso, N. (2015). *Komik: karya sastra bergambar. Humaniora*, 6(4), 496-506.
- Bonnet, M. (1998). *Komik Indonesia*. Jakarta: Perpustakaan Populer Gramedia