

Sistem Rekomendasi Destinasi Tempat Wisata Provinsi Aceh menggunakan Metode *Content based Filtering*

Aceh Province Tourism Destination Recommendation System using Content based Filtering Method

¹Ariefhan Maulana*, ²Arif Nur Rohman, ³Yoga Pristyanto

^{1,2,3}Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas AMIKOM Yogyakarta
^{1,2,3}Jl. Ring Road Utara, Ngringin, Condongcatur, Kec. Depok, Kabupaten Sleman, Daerah Istimewa
Yogyakarta 55281, Indonesia

*e-mail: ariefhanmaulana@students.amikom.ac.id

(received: 9 October 2025, revised: 5 January 2026, accepted: 6 January 2026)

Abstrak

Wisatawan sering mengalami kesulitan dalam menemukan destinasi wisata di Provinsi Aceh yang sesuai dengan preferensi konten serta memiliki kedekatan geografis dengan lokasi mereka. Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi tempat wisata di Provinsi Aceh menggunakan pendekatan *Content-Based Filtering* dengan algoritma *Cosine Similarity* dan *Haversine Formula*. Data yang digunakan terdiri dari 119 destinasi wisata dengan atribut nama destinasi, deskripsi destinasi, serta koordinat geografis berupa *latitude* dan *longitude*. Tahapan penelitian diawali dengan proses *preprocessing data teks* yang meliputi *case folding*, penghapusan tanda baca, tokenisasi, penghapusan kata duplikasi, penghapusan *stopword*, serta proses *stemming*. Selanjutnya, tingkat kemiripan antar destinasi dihitung menggunakan algoritma *Cosine Similarity* berdasarkan deskripsi konten wisata, sedangkan *Haversine Formula* digunakan untuk menghitung jarak geografis antara lokasi pengguna dan destinasi wisata. Hasil pengujian menunjukkan bahwa sistem yang dikembangkan mampu memberikan rekomendasi destinasi wisata yang relevan dengan mempertimbangkan kesesuaian konten dan kedekatan geografis secara bersamaan, sehingga dapat membantu wisatawan dalam menentukan pilihan destinasi wisata yang sesuai dengan preferensi mereka.

Kata kunci: *content based filtering, cosine similarity, haversine formula, preprocessing, sistem rekomendasi*

Abstract

Tourists often experience difficulties in finding tourist destinations in Aceh Province that match their content preferences and are geographically close to their location. This study aims to develop a tourism destination recommendation system in Aceh Province using a *Content-Based Filtering* approach with the *Cosine Similarity* algorithm and the *Haversine Formula*. The dataset consists of 119 tourist destinations, including attributes such as destination name, destination description, and geographical coordinates (*latitude* and *longitude*). The research process began with text data preprocessing, which included *case folding*, punctuation removal, tokenization, duplicate word removal, *stopword* removal, and *stemming*. Next, the similarity between destinations was calculated using the *Cosine Similarity* algorithm based on tourism content descriptions, while the *Haversine Formula* was applied to measure the geographical distance between the user's location and the tourist destinations. The results indicate that the developed system is able to provide relevant tourism destination recommendations by simultaneously considering content relevance and geographical proximity. Therefore, the system can assist tourists in selecting destinations that best match their preferences.

Keywords: *content based filtering, cosine similarity, haversine formula, preprocessing, recommendation system*

1 Pendahuluan

Sektor pariwisata merupakan salah satu sektor strategis yang berperan penting dalam mendorong pertumbuhan ekonomi daerah. Provinsi Aceh memiliki potensi pariwisata yang besar dengan kekayaan alam, budaya, dan sejarah yang beragam. Berdasarkan data dari Dinas Kebudayaan dan Pariwisata Provinsi Aceh, jumlah kunjungan wisatawan mengalami peningkatan pada tahun 2024 dibandingkan periode yang sama pada tahun 2023, yang menunjukkan adanya peningkatan minat wisatawan untuk berkunjung ke Aceh. Beberapa destinasi wisata yang banyak diminati antara lain Masjid Raya Baiturrahman, Museum Tsunami Aceh, dan Pantai Lampuuk. Meskipun demikian, peningkatan jumlah kunjungan wisatawan tersebut belum sepenuhnya diimbangi dengan kemudahan wisatawan dalam memperoleh informasi dan menentukan destinasi wisata yang sesuai dengan minat dan preferensi mereka[1]. Seiring dengan perkembangan teknologi informasi, berbagai platform digital seperti media sosial dan *Google Maps* telah dimanfaatkan untuk menyediakan informasi mengenai destinasi wisata. Platform-platform tersebut memungkinkan wisatawan untuk mengakses informasi lokasi, ulasan, serta dokumentasi visual suatu destinasi. Namun, ketersediaan informasi yang melimpah belum tentu mempermudah wisatawan dalam menentukan pilihan destinasi yang sesuai, terutama ketika mempertimbangkan kesesuaian karakteristik destinasi dan kedekatan jarak dengan lokasi wisatawan. Kondisi ini menunjukkan perlunya sistem yang tidak hanya menyajikan informasi, tetapi juga mampu memberikan rekomendasi destinasi wisata secara personal dan relevan[1].

Beberapa penelitian sebelumnya telah mengembangkan sistem rekomendasi wisata menggunakan metode Collaborative Filtering[1]. Metode ini terbukti efektif dalam memberikan rekomendasi berdasarkan riwayat dan preferensi pengguna lain. Namun, pendekatan tersebut memiliki keterbatasan, khususnya pada permasalahan cold start, yaitu ketidakmampuan sistem dalam memberikan rekomendasi yang optimal kepada wisatawan baru yang belum memiliki riwayat kunjungan. Keterbatasan ini menyebabkan rekomendasi yang dihasilkan kurang relevan bagi pengguna baru. Selain itu, penelitian lain mengembangkan sistem pendukung keputusan dalam pemilihan destinasi wisata di Provinsi Aceh menggunakan metode Simple Additive Weighting (SAW)[2]. Metode SAW membantu wisatawan dalam menentukan destinasi terbaik berdasarkan sejumlah kriteria yang telah ditentukan. Namun, metode ini bergantung pada bobot kriteria yang bersifat subjektif dan harus didefinisikan secara jelas, seperti biaya, fasilitas, dan jarak, sehingga kurang fleksibel ketika diterapkan pada kondisi data yang tidak memiliki bobot kriteria yang terstruktur. Penelitian lain juga mengembangkan sistem pengenalan tempat wisata di Provinsi Aceh yang lebih berfokus pada aspek perancangan dan pengembangan sistem informasi, seperti metode pengembangan perangkat lunak dan alur kerja sistem[3]. Meskipun penelitian tersebut memberikan kontribusi dalam penyediaan informasi destinasi wisata, sistem yang dikembangkan belum menyediakan fitur rekomendasi destinasi wisata lain yang relevan bagi pengguna berdasarkan karakteristik destinasi dan lokasi geografis.

Berdasarkan keterbatasan dari penelitian-penelitian sebelumnya, diperlukan suatu pendekatan yang mampu memberikan rekomendasi destinasi wisata tanpa bergantung pada riwayat pengguna lain maupun bobot kriteria yang telah ditentukan sebelumnya. Oleh karena itu, penelitian ini mengusulkan pengembangan sistem rekomendasi tempat wisata di Provinsi Aceh menggunakan metode Content-Based Filtering dengan memanfaatkan data nama destinasi, deskripsi destinasi serta informasi geografis berupa latitude dan longitude yang diperoleh dari *TripAdvisor* dan *Google Maps*. Pendekatan ini memungkinkan sistem untuk merekomendasikan destinasi wisata berdasarkan kemiripan karakteristik konten dan kedekatan jarak geografis antara wisatawan dan destinasi wisata, sehingga diharapkan dapat memberikan rekomendasi yang relevan bagi wisatawan baru maupun wisatawan yang telah memiliki pengalaman sebelumnya.

2 Tinjauan Literatur

Penelitian mengenai sistem rekomendasi destinasi wisata telah banyak dikembangkan dengan menggunakan pendekatan Content-Based Filtering yang memanfaatkan informasi deskriptif dari destinasi wisata. Beberapa penelitian menerapkan algoritma TF-IDF untuk pembobotan fitur teks dan algoritma *Cosine Similarity* untuk mengukur tingkat kemiripan antar destinasi. Penelitian oleh Salsabilla dan Utomo[4] menunjukkan bahwa pendekatan tersebut mampu menghasilkan rekomendasi

<http://sistemasi.ftik.unisi.ac.id>

destinasi wisata yang relevan berdasarkan preferensi pengguna dan membantu dalam promosi destinasi wisata di Kabupaten Pekalongan. Hasil serupa juga diperoleh pada pengembangan sistem rekomendasi pariwisata berbasis web di Gunung Kidul, di mana penerapan TF-IDF dan Cosine Similarity mampu memberikan rekomendasi yang sesuai dengan pilihan wisatawan[5]. Temuan ini menunjukkan bahwa pendekatan berbasis konten efektif dalam memodelkan preferensi pengguna tanpa bergantung pada pengalaman pengguna lain.

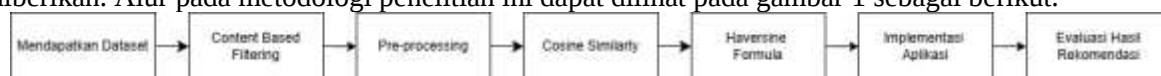
Meskipun demikian, beberapa penelitian juga mengungkapkan keterbatasan dari pendekatan Content-Based Filtering yang hanya mengandalkan kecocokan kata kunci, di mana algoritma TF-IDF belum mampu menangkap hubungan semantik secara mendalam dan sangat bergantung pada kualitas data teks[5]. Keterbatasan serupa juga ditemukan pada sistem rekomendasi non-pariwisata, seperti rekomendasi skincare, yang menunjukkan bahwa kualitas rekomendasi dipengaruhi oleh kelengkapan dan akurasi data[6]. Hal ini menunjukkan bahwa meskipun hasil rekomendasi dinilai relevan, masih terdapat ruang untuk peningkatan kualitas rekomendasi melalui penambahan konteks atau fitur pendukung lainnya.

Selain aspek kemiripan konten, sebagian besar penelitian terdahulu masih berfokus pada karakteristik deskriptif destinasi tanpa mempertimbangkan aspek kedekatan geografis pengguna. Penelitian mengenai rekomendasi wisata kuliner di Madura yang menerapkan TF-IDF dan *Cosine Similarity* berhasil memberikan rekomendasi yang relevan berdasarkan preferensi wisatawan, namun belum mengintegrasikan perhitungan jarak geografis dalam proses rekomendasi[7]. Padahal, dalam konteks pariwisata, jarak antara lokasi pengguna dan destinasi wisata merupakan faktor penting yang dapat memengaruhi keputusan wisatawan dalam memilih destinasi yang akan dikunjungi.

Berdasarkan kajian penelitian sebelumnya, pendekatan *Content-Based Filtering* dengan *Cosine Similarity* terbukti efektif dalam menghasilkan rekomendasi berdasarkan kemiripan konten, namun belum mempertimbangkan aspek geografis. Oleh karena itu, penelitian ini mengintegrasikan *Cosine Similarity* dan Haversine Formula untuk menghitung kedekatan lokasi wisatawan dengan destinasi. Provinsi Aceh dipilih sebagai objek penelitian yang masih jarang dikaji, sehingga diharapkan memberikan kontribusi metodologis dan empiris dalam pengembangan sistem rekomendasi wisata.

3 Metode Penelitian

Alur metodologi penelitian ini, penulis akan melakukan beberapa tahapan - tahapan yang harus dilalui yaitu di mulai dari mendapatkan *dataset*, *content based filtering*, *pre-processing*, *cosine similarity*, *haversine formula*, implementasi aplikasi dan terakhir evaluasi hasil rekomendasi yang diberikan. Alur pada metodologi penelitian ini dapat dilihat pada gambar 1 sebagai berikut.



Gambar 1 Alur metode penelitian

3.1 Mendapatkan dataset

Pada tahap mendapatkan dataset, penulis memperoleh data destinasi wisata melalui teknik web scraping menggunakan ekstensi *Instant Data Scraper* pada peramban *Google Chrome*. Tools ini memungkinkan pengambilan data dalam jumlah besar secara efisien tanpa harus mengakses basis data internal suatu sistem, melainkan dengan mengekstraksi informasi langsung dari halaman situs web publik [8]. Selain melalui scraping, beberapa penelitian lain juga memperoleh dataset sistem rekomendasi dari sumber eksternal seperti *Kaggle* [9].

Namun demikian, pada konteks penelitian ini, dataset destinasi wisata yang secara khusus memuat informasi detail mengenai tempat wisata di Provinsi Aceh belum tersedia secara terbuka pada platform penyedia dataset publik. Kondisi tersebut menyebabkan proses pengumpulan data tidak dapat dilakukan melalui repositori dataset siap pakai, sehingga teknik web scraping dipilih sebagai pendekatan yang paling memungkinkan untuk memperoleh data yang dibutuhkan.

Data destinasi wisata dikumpulkan dari beberapa sumber, yaitu *Google Maps* dan *ChatGPT*. *Google Maps* digunakan untuk memperoleh informasi utama terkait destinasi wisata di Provinsi Aceh, seperti nama destinasi dan koordinat geografis, sedangkan *ChatGPT* dimanfaatkan untuk menghasilkan deskripsi destinasi wisata. Deskripsi yang dihasilkan selanjutnya diverifikasi dengan membandingkannya terhadap informasi yang tersedia pada *Google Maps* dan situs *TripAdvisor* guna memastikan kesesuaian dan keunikan deskripsi, sehingga dataset yang digunakan tetap relevan.

3.2 Content Based Filtering

Content Based Filtering merupakan salah satu metode untuk memberikan rekomendasi berdasarkan kemiripan dari objek satu dengan objek yang lain dimana sistem akan memberikan rekomendasi - rekomendasi yang memiliki kemiripan dengan objek tertentu. Kemudian Content Based Filtering ini dapat diartikan pendukung keputusan untuk menentukan hasil dengan menggunakan item dengan nilai preferensi pengguna[10]. Metode ini sistem yang akan memberikan rekomendasi serta menyarankan item yang sesuai untuk wisatawan. Untuk menghitung kemiripan antar item, Content-Based Filtering menggunakan informasi dari item yang dijadikan atribut[3].

3.3 Pre-processing

Pre-processing merupakan sebuah proses yang bertujuan untuk mengubah data tidak teratur menjadi data yang teratur[11]. Kemudian tahap ini juga salah satu tahap awal ketika pengolahan data dengan tujuan agar menghindari data yang tidak konsisten atau bisa diartikan mengganggu[12]. Ketika tahap ini dapat dilakukan, maka sistem rekomendasi dengan metode *Content Based Filtering* ini akan memberikan informasi lebih akurat dan teratur.

Pada tahap ini, ada beberapa tahap yang harus dilakukan yaitu dapat dilihat di gambar 2 sebagai berikut.



Gambar 2 Tahapan *pre-processing*

Tahapan:

- Case Folding yaitu semua huruf dalam dokumen diubah menjadi huruf kecil.
- Punctual Remove merupakan tahap kedua yang bertujuan untuk menghapus semua tanda baca seperti ,.@!? dikarenakan sistem menganggap tanda baca tersebut tidak bermakna[13].
- Tokenisasi kalimat adalah tahap yang memecahkan dari kalimat menjadi kata dimana bertujuan untuk mengetahui kemiripan kata.
- Eliminasi kata yang sama dalam artian yang memiliki arti yang sama[11]. Eliminasi kata duplikat ini merupakan bagian pendukung dari proses filtering.
- Filtering merupakan tahap kata-kata yang penting diambil dari hasil tahap sebelumnya. Kemudian filtering melakukan penghilangan kata stoplist/stopword seperti ini,dan,jika, dan kata umum lainnya yang tidak memberikan kontribusi signifikan terhadap penentuan kemiripan konten [11].
- Stemming merupakan tahap perubahan bentuk suatu kata yang berimbuhan menjadi kata dasar[11].

Tahapan pre-processing tersebut merupakan tahapan umum yang sering dipakai pada penelitian sistem rekomendasi, tahapan tersebut dipilih karena dataset yang digunakan penulis merupakan dataset yang sudah baik yang memiliki tata bahasa baku, pada penelitian sistem rekomendasi yang menggunakan dataset yang tidak baku, biasanya menggunakan perbaikan kosa kata seperti natural language processing untuk memperbaiki typo[14]

Pada tahap ini, penulis akan membuat sebuah tools tersendiri dengan menggunakan bahasa pemrograman python. Adapun tools tersebut dimulai dari instalasi beberapa pustaka python yang dibutuhkan dan dapat dilihat pada gambar 3 sebagai berikut.

```
import pandas as pd
import re
!pip install spacy
!python -m spacy download xx_ent_wiki_sm
import spacy
from nltk.corpus import stopwords
import nltk
nltk.download('stopwords')
!pip install Sastrawi
# Membaca file yang diunggah dan membaca
df=pd.read_csv('tempatwisataacehdes.csv')
```

Gambar 3 Pustaka python

Gambar 3 terdapat beberapa pustaka yang dibutuhkan seperti pandas, re, spacy, nltk dan sastrawi. Adapun pustaka pandas merupakan salah satu pustaka python yang cukup populer dimana sering

digunakan ketika mengelola data. Pustaka re biasanya digunakan untuk mencari membersihkan teks dari beberapa simbol yang tidak dibutuhkan. Kemudian pustaka spacy juga sering digunakan untuk mengenali kata benda, kata kerja dan lain - lain. pustaka nltk biasanya digunakan untuk mengelola kata - kata umum seperti dan atau dan kata lainnya. Terakhir untuk pustaka Sastrawi akan digunakan ketika ingin mengubah kata imbuhan ke kata dasar seperti terletak menjadi letak. Selain itu, pada gambar 3 juga terdapat `df=pd.read_csv` berfungsi untuk membaca file dataset yang ingin melakukan preprocessing data.

Setelah menginstal beberapa pustaka python, maka masuk awal dari preprocessing data. Berikut tahapan - tahapan tools ketika melakukan preprocessing data.

a. Case Folding

```
# Case Folding
df['after_Case_Folding']=df['markdown'].str.lower()
```

Gambar 4 Alat case folding

Kode program pada gambar ke 4 itu digunakan untuk mengubah semua huruf menjadi dalam bentuk huruf kecil.

b. Punctual Removal

```
# Melakukan Punctual Removal
df['after_Punctual_Removal']=df['after_Case_Folding'].apply(lambda
x: re.sub(r'[\w\s]', '', x))
```

Gambar 5 Alat punctual removal

Gambar 5 bertujuan untuk menghapus semua tanda baca seperti ,?.@! yang terdapat didalam dataset.

c. Tokenisasi

```
#Tokenisasi
nlp = spacy.load("xx_ent_wiki_sm") # Model multibahasa kecil
data = pd.DataFrame(df)
# Tokenisasi menggunakan spaCy
def tokenize_text(text):
    # Proses teks dengan spaCy
    doc = nlp(text)
    # Ambil token dalam bentuk list
    tokens = [token.text for token in doc]
    return tokens
# Terapkan tokenisasi ke setiap baris
data["tokens"]=data["after_Punctual_Removal"].apply(tokenize_text)
```

Gambar 6 Alat tokenisasi

Pada gambar 6 diatas, menunjukkan bahwa setelah melakukan tahapan punctual remove maka kalimat tersebut akan dipisah dalam bentuk tokenisasi.

d. Eleminasi kata

```
# Fungsi menghilangkan duplikat dalam token
def remove_duplicates(tokens):
    return list(dict.fromkeys(tokens))
# eliminasi kata duplikat pada dataset
data["tokens_remove"]=data["tokens"].apply(remove_duplicates)
```

Gambar 7 Alat eliminasi kata

Kode program yang terdapat pada gambar 7 bertujuan untuk menghapus kata yang mengandung arti yang sama sehingga setelah program tersebut berjalan maka akan disimpan kedalam kolom baru dengan nama `tokens_remove`.

e. Filtering

```
# Mendapatkan daftar stopwords untuk Bahasa Indonesia
stop_words=set(stopwords.words('indonesian'))
# Fungsi untuk menghilangkan stopwords
def remove_stopwords(tokens):
    return [word for word in tokens if word.lower() not in stop_words]
# Menerapkan fungsi filter stopwords
data["filtered_tokens"]=data["tokens_remove"].apply(remove_stopwords)
```

Gambar 8 Alat filtering

Gambar 8 diatas menunjukkan bahwa program akan mendapatkan daftar stopwords terlebih dahulu agar dapat dilakukan tahapan filtering.

f. Stemming

```
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
# Inisialisasi Stemmer
factory = StemmerFactory()
stemmer = factory.create_stemmer()
# Fungsi untuk melakukan stemming
def stemming(tokens):
    return [stemmer.stem(word) for word in tokens]
# Terapkan stemming pada dataset
data["stemming_tokens"]=data["filtered_tokens"].apply(stemming)
```

Gambar 9 Alat stemming

Pada gambar 9, penulis melakukan stemming data dengan menggunakan pustaka Sastrawi dimana program akan mengubah kata yang mengandung berimbuhan menjadi kata dasar seperti membaca menjadi baca. Kemudian data yang telah melakukan tahapan stemming ini akan dimasukan kedalam kolom baru dengan nama stemming_tokens. Tahap akhir dilakukan dengan mengubah tokenisasi menjadi kalimat kembali agar mudah dimasukkan ke database MySQL.

3.4 Cosine Similarity

Cosine Similarity merupakan sebuah metode perhitungan untuk melakukan pengukuran tingkat kemiripan antara objek satu dengan object yang lain. Secara fundamental, fungsi dari similarity yaitu fungsi yang menerima dua objek yang berbentuk bilangan riil, yaitu 0, dan 1, dan mengembalikan nilai yang menunjukkan kemiripan, atau kesamaan, antara kedua objek tersebut[11]. Selain itu, Cosine Similarity ini ketika melakukan pengukuran tingkat kecocokan teks dianggap sebagai vector, nilai dari vector X dan Y adalah vektor *term-frequency* dari objek[15]. Kemudian Cosine Similarity juga salah satu metode yang sering digunakan ketika menggunakan sistem rekomendasi content based filtering. Mengapa demikian karena *Cosine Similarity* ini akan menghitung *similarity*/kemiripan berdasarkan object X dan object Y. Perhitungan kemiripan antara objek X dan objek Y dilakukan menggunakan Algoritma *Cosine Similarity* yang ditunjukkan pada Persamaan (1) berikut ini.

$$sim(X,Y) = \frac{n(X \cap Y)}{\sqrt{n(X)n(Y)}} \quad (1)$$

Pada Rumus Persamaan (1) digunakan untuk mencari nilai unsur kemiripan antara objek destinasi tempat wisata X dan objek destinasi Y atau biasanya dikenal dengan persamaan cosine similarty. Dimana hasil dari unsur kemiripan antara objek destinasi tempat wisata X dengan objek destinasi tempat wisata Y atau yang dilambangkan dengan $sim(X,Y)$, sama dengan banyaknya kata yang sama diantara objek destinasi tempat wisata X dengan objek destinasi tempat wisata Y yang dilambangkan dengan $n(X \cap Y)$ akan dibagi dengan akar banyaknya kata yang terdapat didalam object destinasi tempat wisata X dan banyaknya kata yang terdapat didalam object destinasi tempat wisata Y. Kedua akar banyaknya kata yang terdapat didalam objek destinasi tempat wisata X dan Y tersebut dilambangkan dengan $\sqrt{n(X)n(Y)}$.

3.5 Haversine Formula

Haversine Formula adalah sebuah metode yang penting dalam sistem navigasi dimana metode ini bertujuan untuk mengetahui jarak antar lokasi dengan memanfaatkan garis lintang dan garis bujur. Garis lintang tersebut biasanya disebut dengan latitude dan garis bujur disebut dengan longitute. Latitude dan Logitude adalah sistem koordinat geografis yang digunakan untuk menemukan di mana seseorang berada di permukaan bumi. *Haversine Formula* merupakan sebuah metode yang dapat dilakukan untuk menghitung jarak antar dua titik dengan mempertimbangkan bahwa Bumi adalah bidang dengan memiliki derajat kelengkungan[16]. Kemudian Haversine Formula suatu persamaan untuk menentukan posisi dan arah berdasarkan berdasarkan jarak dua titik pada permukaan bumi dalam lingkaran[17]. Pada penelitian ini, *Haversine Formula* akan digunakan untuk menghitung jarak antar wisatawan dengan destinasi yang akan direkomendasikan. Adapun Rumus Persamaan (2), (3), (4), (5), dan (6) dapat dilihat berikut ini.

$$\Delta lat = lat2 - lat1 \quad (2)$$

$$\Delta long = long2 - long1 \quad (3)$$

$$a = \sin^2\left(\frac{\Delta lat}{2}\right) + \cos(lat1) \cdot \cos(lat2) \cdot \sin^2\left(\frac{\Delta long}{2}\right) \quad (4)$$

$$c = 2 \operatorname{atan2}(\sqrt{a}, \sqrt{1-a}) \quad (5)$$

$$d = R \cdot c \quad (6)$$

Pada Persamaan (2) digunakan untuk mencari besaran perubahan latitude keberadaan wisatawan dengan latitude keberadaan suatu destinasi tempat wisata yang hasil besaran perubahannya dilambangkan Δlat sedangkan Persamaan (3) akan mencari besaran perubahan logitute lokasi wisatawan dengan logitute lokasi dari suatu destinasi tempat wisata yang akan dilambangkan $\Delta long$. Kemudian setelah mendapatkan hasil Δlat dan Δlog maka akan dilanjutkan dengan Persamaan (4) agar dapat mengukur jarak angular antara dua titik pada permukaan. Selain itu, Persamaan (5) akan mengukur kalkulasi perpotongan sumbu dilambangkan c . Pada Persamaan (6) akan menghitung jarak sebenarnya d antara dua titik dalam satuan yang sama dengan jari – jari bumi dimana R atau jari – jari bumi 6371 km.

3.6 Implementasi Aplikasi

Pada tahapan ini, penulis membangun aplikasi menggunakan bahasa pemrograman PHP dengan framework CodeIgniter v3.1.13 serta basis data MySQL. Dataset yang telah dikumpulkan diolah untuk menghasilkan rekomendasi destinasi wisata berdasarkan kemiripan nama dan deskripsi serta kedekatan jarak. Perhitungan cosine similarity dan haversine formula diimplementasikan menggunakan PHP untuk menghasilkan rekomendasi yang relevan.

3.7 Evaluasi Hasil Rekomendasi

Evaluasi Hasil Rekomendasi adalah tahap untuk menganalisis kembali hasil yang dihitung menggunakan perhitungan *Cosine Similarity* dan *Haversine Formula* dengan tujuan untuk memastikan kembali apakah rekomendasi yang diberikan oleh sistem sudah sesuai atau belum sesuai dengan preferensi wisatawan serta kedekatan jarak antar wisatawan dengan destinasi tempat wisata Provinsi Aceh. Kemudian dengan melakukan tahapan ini penulis akan menilai kembali akurasi rekomendasi yang diberikan apakah relevan atau tidak.

4 Hasil dan Pembahasan

4.1 Hasil Dataset

Penulis memperoleh dataset sebanyak 119 data yang tersimpan di dalam file *Comma Separated Values (CSV)*. Data yang diperoleh ini, terdapat beberapa kolom yaitu id, url, nama destinasi, rating destinasi, latitude, longitude serta deskripsi destinasi. Dataset yang di peroleh oleh penulis akan dilakukan pre-processing data kemudian akan disimpan kedalam MySql dengan tujuan agar sistem akan mendapatkan dan mengelola data dari MySql. Selain itu, data yang akan digunakan yaitu hanya pada kolom deskripsi destinasi, *latitude* dan *longitude* agar sistem dapat memberikan rekomendasi destinasi tempat wisata yang memiliki unsur kemiripan dan kedekatan jarak antar lokasi wisatawan dengan lokasi destinasi tempat wisata yang direkomendasikan. Adapun data yang didapatkan oleh penulis sebelum melakukan preprocessing seperti pada tabel 1 berikut ini.

Tabel 1 Hasil dataset yang didapatkan

1 Pantai Ujung Karang (4.12571, 96.12789)

Pantai Ujung Karang adalah destinasi wisata alam yang terletak di Aceh, menawarkan pemandangan laut yang indah dengan pasir putih dan hamparan karang di sepanjang pantainya. Pantai ini ...

...

...

119 Ketambe Adventure | Trekking | Rafting aceh (3.69275, 97.65236)

Ketambe Adventure terletak di Kabupaten Aceh Tenggara dan merupakan salah satu destinasi wisata petualangan yang populer, terutama untuk trekking dan rafting. Kawasan ini dikelilingi ...

Pada Tabel 1 untuk mempermudah pemaparan dataset penulis hanya menampilkan beberapa kolom yang terdapat di dataset yaitu id, nama destinasi, latitude, longitude dan deskripsi destinasi. Kemudian untuk kolom url, serta rating destinasi tidak dipaparkan karena pada penelitian ini penulis hanya menggunakan deskripsi destinasi, latitude dan longitude untuk memberikan rekomendasi

<http://sistemasi.ftik.unisi.ac.id>

kepada wisatawan. Kemudian pada tabel diatas data yang dipaparkan belum melakukan preprocessing dimana dataset diatas bisa dilihat masih menggunakan huruf kapital, tanda baca, kata stoplist/stopword serta kata yang mengandung imbuhan. Kata yang berimbuhan diatas dapat dilihat seperti terletak, menawarkan, keindahan, hamparan dan banyak kata berimbuhan lainnya yang akan dibahas oleh penulis pada bagian hasil Preprocessing.

Deskripsi destinasi wisata pada dataset sebagian dihasilkan dengan bantuan Large Language Model (LLM) untuk melengkapi informasi tekstual yang belum konsisten. Untuk menjaga kredibilitas data, setiap deskripsi telah divalidasi secara manual dengan membandingkannya terhadap informasi resmi pada TripAdvisor dan Google Maps, meskipun potensi subjektivitas tetap menjadi keterbatasan penelitian.

4.2 Hasil Pre-processing

Tahap pre-processing dimulai dengan case folding, kemudian dilanjutkan punctuation remove, tokenisasi, eliminasi, filtering, dan stemming kata, dengan hasil penerapan tersebut dapat dilihat pada table 2 berikut ini.

Tabel 2 Hasil Dataset setelah melakukan Preprocessing data

Deskripsi Wisata
Sebelum Pre-processing: Pantai Ujung Karang adalah destinasi wisata alam yang terletak di Aceh, menawarkan pemandangan laut yang indah dengan pasir putih dan hamparan karang di sepanjang pantainya. Pantai ini cocok untuk bersantai, menikmati keindahan matahari terbenam, atau menjelajahi ekosistem laut. ...
Sesudah Pre-processing: pantai ujung karang destinasi wisata alam letak aceh tawar pandang laut indah pasir putih hampar pantai cocok santa nikmat indah matahari benam jajah ekosistem kaya suasana asri jadi tarik ...
Sebelum Pre-processing: Pantai Ujong Batee terletak di Aceh Besar, dikenal dengan pemandangan pantai yang memukau serta jajaran pohon pinus yang rindang di sekitarnya. Pantai ini menawarkan suasana tenang dengan ombak yang relatif tenang, cocok untuk bersantai atau menikmati kuliner khas Aceh yang tersedia ...
Sesudah Pre-processing: pantai ujong batee letak aceh kenal pandang pukau jajar pohon pinus rindang tawar suasana tenang ombak relatif cocok santa nikmat kuliner khas sedia warungwarung unjung pulau weh jauh jadi favorit rekreasi keluarga fotografi

Pada Tabel 2 diatas, dapat dilihat bahwa setelah penulis melakukan tahap Pre-processing, Deskripsi Destinasi Tempat Wisata sebelum melakukan Preprocessing dan sesudah melakukan Preprocessing dapat dilihat memiliki perbedaan dimana sebelum melakukan Preprocessing banyak terdapat kata yang memiliki huruf kapital, tanda baca, kata yang berimbuhan, dan jenis kata yang lain. Deskripsi Destinasi Tempat Wisata Aceh dapat dilihat pada tabel diatas yaitu memiliki kata yang berimbuhan ter “terletak” dimana kata tersebut didapatkan sebelum melakukan tahapan preprocessing. Setelah melakukan preprocessing pada tahapan stemming kata tersebut menjadi “letak” begitu juga dengan kata berimbuhan yang lain “menawarkan” menjadi “tawar”, “hamparan” menjadi “hampar”, “menikmati” menjadi “nikmat” dan banyak kata berimbuhan yang lain yang terdapat didalam tabel 2 yang penulis tidak sebutkan. Setelah melakukan pre-processing dapat dilihat datanya lebih teratur dan konsisten sehingga dengan melakukan tahap pre-processing sistem dapat memberikan informasi rekomendasi destinasi wisata yang lebih akurat dan tepat.

4.3 Pengujian Content Based Filtering

Pada bagian ini, penulis mencoba pengujian pada bagian destinasi tempat wisata provinsi Aceh pada kata kunci pantai dan air terjun dimana kedua kata kunci tersebut sistem akan memberikan rekomendasi sesuai dengan unsur kemiripan destinasi wisata. Penulis mengambil contoh setiap destinasi dengan masing-masing kata kunci 3 pantai dan 2 air terjun. Sehingga jumlah pengujian sebanyak 5 kali. Setiap pengujian akan diberikan rekomendasi 4 destinasi wisata yang memiliki unsur kemiripan berdasarkan perhitungan cosine similarity tertinggi. Kemudian data yang dimasukan kedalam sistem database *Mysql* merupakan hasil dataset yang telah melakukan tahapan *preprocessing* data. Data tersebut yang tidak terdapat tanda baca, huruf kapital, dan lain - lain yang dapat dilihat

<http://sistemasi.ftik.unisi.ac.id>

pada bagian metodologi penelitian. Adapun hasil pengujian sistem rekomendasi content based ini dapat dilihat pada tabel 3 berikut ini.

Tabel 3 Hasil dari pengujian *content based filtering* dengan menggunakan *cosine similarty*

No	Detail Destinasi	Top	Hasil Rekomendasi	Similarity Nama	Similarity Deskripsi	Total
1	pantai ujung karang	1	pantai lancok	0.41	0.59	1.0
		2	pantai leuge	0.41	0.56	0.97
		3	pantai lhok kubu	0.33	0.60	0.93
		4	pantai lambaro	0.52	0.41	0.93
2	pantai ujung batee	1	pantai ujung seukee	0.67	0.43	1.1
		2	pantai ujung blang	0.67	0.37	1.04
		3	pantai ujung kareu ng	0.58	0.42	1.0
		4	pantai lambaro	0.39	0.41	0.80
3	pantai wisata kuala raja	1	pantai lancok	0.35	0.59	0.94
		2	pantai ujung blang	0.29	0.61	0.90
		3	pantai ulee rubek	0.29	0.54	0.83
		4	pantai lampuuk	0.35	0.48	0.83
4	air terjun tingkat tujuh	1	air terjun elnaja	0.58	0.64	1.22
		2	air terjun rujak	0.55	0.58	1.13
		3	air terjun suhom	0.58	0.53	1.11
		4	air terjun sabet	0.58	0.52	1.10
5	air terjun tansaran bidin	1	air terjun kaya	0.58	0.66	1.24
		2	air terjun sabet	0.58	0.63	1.21
		3	air terjun rujak	0.58	0.62	1.20
		4	air terjun elnaja	0.58	0.61	1.19

Dari tabel 3 diatas, penulis dapat menjelaskan dimana sistem rekomendasi bekerja dengan mencari unsur kemiripan berdasarkan kemiripan nama destinasi dan deskripsi destinasi tempat wisata. Kemudian sistem rekomendasi ini, tidak memperhatikan rating yang terdapat pada suatu destinasi tempat wisata melainkan hanya mencari kemiripan antara nama dan deskripsi. Sebagai contoh pada tabel 3 diatas Destinasi Tempat Wisata pantai ujung karang ketika wisatawan ingin mengunjungi destinasi wisata tersebut maka sistem akan merekomendasikan destinasi wisata pantai lancok dengan unsur kemiripan 1.0, pantai *leuge* dengan unsur kemiripan 0.97, pantai lhok kubu dengan unsur kemiripan 0.93 dan pantai lambaro dengan unsur kemiripan 0.93. Kemudian sebagai urutannya sistem akan memberikan rekomendasi pantai lancok terlebih dahulu kemudian diikuti dengan pantai leuge , pantai lhok kubu dan pantai lambaro. Mengapa pantai lancok yang lebih utama direkomendasikan karena hasil penjumlahan similarity nama dan *similarity* deskripsi menunjukkan bahwa pantai ujung karang dengan pantai lancok memiliki unsur kemiripan tertinggi berdasarkan nama dan deskripsi dibandingkan dengan destinasi tempat wisata yang lain.

Algoritma *Cosine Similarity* ini bekerja dengan menghasilkan skor perhitungan kemiripan beberapa destinasi tempat wisata terhadap sebuah destinasi tempat wisata ketika di detailkan. kemudian beberapa destinasi tempat wisata tersebut akan mencari calon kandidat rekomendasi berdasarkan skor total tertinggi yang diperoleh dari hasil perjumlahan similarity nama destinasi dan

similarity deskripsi destinasi. Setelah itu, sistem tidak akan mengambil semua total skor tertinggi melainkan hanya 4 destinasi total skor yang memperoleh sangat tinggi ketika di urutkan. Pada Penelitian sebelumnya, *Content-Based Filtering* bekerja dengan baik dalam menghasilkan rekomendasi wisata[18] begitu juga ketika penulis meneliti dengan menggunakan content-Based Filtering dapat menghasilkan rekomendasi yang tepat berdasarkan kemiripan nama dan deskripsi. Selain itu, penulis mengambil 5 sampel destinasi serta setiap destinasi akan diberikan 4 rekomendasi destinasi yang hampir mirip dengan suatu destinasi tersebut berdasarkan hasil perhitungan total skor tertinggi setelah diurutkan yang diperoleh dari penjumlahan antara hasil similarty nama dan similarty destinasi.

Hasil pengujian *Content-Based Filtering* yang hanya menggunakan *Cosine Similarity* menunjukkan bahwa sistem belum mempertimbangkan kedekatan jarak antara wisatawan dan destinasi. Berdasarkan Tabel 3, rekomendasi yang dihasilkan sudah baik dari sisi kemiripan konten, namun berpotensi merekomendasikan destinasi yang berlokasi jauh dari wisatawan. Oleh karena itu, penelitian ini memadukan *Cosine Similarity* dengan *Haversine Formula* agar sistem mampu mempertimbangkan aspek geografis, sehingga rekomendasi yang dihasilkan menjadi lebih relevan dan tepat sasaran. Adapun hasil perpaduan dapat dilihat pada table 4 berikut ini.

Tabel 4 Hasil uji perpaduan cosine similarty dan haversine formula

No	Detail Destinasi	Rekomendasi wisata ke 1	Rekomendasi wisata ke 2	Rekomendasi wisata ke 3	Rekomendasi wisata ke 4
1	pantai ujung karang	pantai lancok overall score : 0.93 dan jarak 68.49 km	dermage pantai menye overall score : 0.88 dan jarak 28.45 km	pantai ujung blang overall score : 0.87 dan jarak 70.12 KM	ujung nunag overall score: 0.87 dan jarak 20.1 km
2	pantai ujung batee	pantai ujung seukee overall score: 1.02 dan jarak 60.68 km	pantai ujung blang overall score: 0.96 dan jarak 70.12 km	pantai ujung kareu ng overall score: 0.8 dan jarak 216.22 km	dermage pantai menye overall score: 0.74 dan jarak 28.45 km
3	pantai wisata kuala raja	pantai lancok overall score: 0.88 dan jarak 68.49 km	pantai ujung blang overall score: 0.85 dan jarak 70.12 km	dermaga pantai menye overall score: 0.81 dan jarak 28.45 km	pantai ulee rubek overall scroll: 0.77 dan jarak 94.42 km
4	air terjun tingkat tujuh	air terjun tansaran bidin overall score: 1.04 dan jarak 21.64 km	air terjun rujak overall score: 1.02 dan jarak 84.79 km	air terjun elnaja overall score: 1 dan jarak 171.6 km	air terjun sabet overall score: 0.92 dan jarak 155.38 km
5	air terjun tansaran bidin	air terjun rujak overall score: 1.08 dan jarak 84.79 km	air terjun ceuraceu alue teungoh overall scroll: 1.02 dan jarak 111.09 km	air terjun sabet overall score: 1.01 dan jarak 155.38 km	air terjun blang kolam overall score: 1 dan 52.86 km

Pada tabel 4 menunjukkan sistem rekomendasi bekerja lebih maksimal dibandingkan hasil yang diperoleh dari tabel 3 dimana bukan hanya mencari unsur kemiripan berdasarkan nama dan deskripsi melainkan juga menghitung jarak georgrafis antara wisatawan dengan destinasi yang akan direkomendasikan. Sebagai contoh table 4 terdapat pantai ujung batee ketika wisatawan ingin mengunjungi destinasi wisata tersebut maka sistem akan melihat titik kooridinat wisatawan terlebih dahulu dimana titik kooridnat pada sistem harus diatur secara statis yaitu 4.69451 dan 96.76017. Kemudian sistem sudah mendapatkan titik koordinat dari wisatawan untuk mengetahui keberadaan wisatawan maka sistem akan memberikan rekomendasi destinasi wisata pantai ujung seukee dengan memperoleh hasil overall score 1.02 dan kedekatan jarak antara wisatawan dengan destinasi wisata sejauh 60.68 km, pantai ujung blang dengan memperoleh hasil overall score 0.96 dengan kedekatan jarak 70.12 km, pantai ujung kareu ng hasil overall score 0.8 dengan kedekatan jarak 216.22 km dan dermage pantai menye memperoleh hasil overall score 0.74 dengan kedekatan jarak 28.45 km.

Dari contoh tersebut dapat dilihat terdapat nilai overall score. Nilai overall score diperoleh dari hasil kombinasi antara hasil perhitungan cosine similarty dan haversine formula dimana kombinasi dua algoritma tersebut diawali dari menghitung cosine smilarty. Ketika hasil consine similarity sudah diperoleh maka sistem akan mengambil 30 destinasi yang memiliki hasil perhitungan tertinggi. Setelah itu, sistem akan menghitung jarak dan normalisasi skor jarak melalui haversine formula dengan memanfaatkan koordinat lokasi wisatawan dan koordinat lokasi destinasi yang telah dipilih. Kemudian setelah mendapatkan hasil perhitungan haversine formula maka sistem akan menggabungkan skor yang diperoleh dari hasil perhitungan cosine similaty dan haversine formula dengan melakukan pembobotan hasil yang diperoleh. Hasil perhitungan cosine similaty akan diberikan bobot konten 80% dan bobok jarak 20%. Dari pembobotan tersebut kesamaan konten akan lebih diprioritaskan dibandingkan kedekatan lokasi dalam penentuan hasil skor akhir. Maka pada tabel 4 pada percobaan kedua ketika wisatawan ingin mengunjungi pantai ujung batee sistem memberikan dermage pantai menye sebagai urutan keempat dalam hasil rekomendasi padahal dari empat rekomedasi yang diperoleh dermage pantai menye memiliki jarak yang paling dekat dengan wisatawan akan tetapi dermage pantai menye menduduki urutan ke empat dalam hasil rekomendasi. Kemudian terdapat pantai ujung kareu ng yang menduduki urutan ke tiga dimana dari segi jarak dapat dilihat memperoleh jarak yang sangat jauh dengan wisatawan akan tetapi destinasi tersebut menduduki urutan ketiga. Hal tersebut dapat terjadi dikarenakan pada pembobotan hasil akhir perhitungan, bobot konten lebih dipriositaskan dibandingkan dengan bobok jarak. Adapun penulis sudah melakukan percoabaan jika bobot jarak lebih diprioritaskan dibandingkan bobot konten. Hasil dari percobaan dapat dilihat pada tabel 5 sebagai berikut.

Tabel 5 Hasil uji perpaduan cosine – haversine dengan prioritas jarak

No	Detail Destinasi	Rekomendasi wisata ke 1	Rekomendasi wisata ke 2	Rekomendasi wisata ke 3	Rekomendasi wisata ke 4
1	pantai ujung karang	ujung nunang overall score : 0.88 dan jarak 20.1 KM	dermage pantai menye overall score : 0.87 dan jarak 28.45 KM	pantan terong overall score : 0.81 dan jarak 7.42 KM	pantai lancok overall score : 0.79 dan jarak 68.49 KM
2	pantai ujung batee	pantai seukee overall score: 0.86 dan jarak 60.68 km	pantai ujung blang overall score: 0.81 dan jarak 70.12 km	dermaga pantai menye overall score: 0.8 dan jarak 28.45 km	pantai lancok overall score: 0.69 dan jarak 68.49 km
3	pantai wisata kuala raja	dermaga pantai menye overall score: 0.83 dan jarak 28.45 km	pantai lancok overall score: 0.77 dan jarak 68.49 km	pantai ujung blang overall score: 0.75 dan jarak 70.12 km	wisata lhok seulayang overall score: 0.74 dan jarak 56.26 km
4	air terjun tingkat tujuh	air terjun tansaran overall score: 0.97 dan jarak 21.64 km	air terjun blang kolam overall score: 0.83 dan jarak 52.86 km	air terjun rujak overall score: 0.8 dan jarak 84.79 km	air terjun atu kapur overall score: 0.75 dan jarak 78.61 km
5	air terjun tansaran bidin	air terjun blang kolam overall score: 0.87 dan jarak 52.86 km	air terjun rujak overall score: 0.83 dan jarak 84.79 km	pantan terong overall score: 0.78 dan jarak 7.42 km	air terjun atu kapur overall score: 0.75 dan 78.61 km

Berdasarkan Tabel 5, dapat dilihat hasil pengujian algoritma. Nilai overall score diperoleh dari kombinasi bobot kemiripan konten sebesar 40% dan bobot kedekatan jarak sebesar 60%. Kombinasi bobot ini dirancang dengan tujuan untuk lebih memprioritaskan kedekatan geografis, namun tetap mempertimbangkan kesesuaian konten destinasi dengan preferensi pengguna. Hasil pengujian pada

Tabel 5 menunjukkan bahwa sistem cenderung merekomendasikan destinasi wisata yang memiliki jarak lebih dekat dengan lokasi wisatawan. Sebagai contoh, pada skenario pengujian pertama ketika wisatawan memilih destinasi Ujung Karang, sistem memberikan rekomendasi Ujung Nunang pada urutan pertama dengan nilai overall score sebesar 0,88 dan jarak 21,1 km. Rekomendasi selanjutnya juga menunjukkan pola serupa, di mana destinasi dengan jarak relatif lebih dekat memperoleh peringkat yang lebih tinggi dibandingkan destinasi yang memiliki jarak lebih jauh. Meskipun bobot jarak lebih besar, pada beberapa skenario pengujian masih ditemukan kondisi di mana destinasi dengan jarak yang lebih jauh memperoleh peringkat lebih tinggi dibandingkan destinasi yang lebih dekat. Hal ini disebabkan oleh kontribusi bobot kemiripan konten sebesar 40% yang tetap berpengaruh dalam perhitungan overall score. Sebagai contoh, pada skenario pengujian kedua ketika wisatawan memilih Pantai Ujung Batee, terdapat rekomendasi destinasi pada peringkat ketiga yang memiliki jarak lebih dekat dibandingkan peringkat pertama dan kedua. Kondisi ini menunjukkan bahwa kemiripan konten tetap berperan dalam menjaga relevansi rekomendasi agar sesuai dengan karakteristik destinasi yang diminati pengguna.

Pemilihan kombinasi bobot 40% untuk kemiripan konten dan 60% untuk kedekatan jarak didasarkan pada karakteristik alamat pariwisata, di mana faktor lokasi memiliki pengaruh yang signifikan dalam pengambilan keputusan wisatawan. Wisatawan cenderung lebih mempertimbangkan destinasi yang dapat dijangkau dengan jarak yang lebih dekat, namun tetap mengharapkan destinasi yang memiliki jenis atau karakteristik yang serupa dengan preferensi mereka. Oleh karena itu, kombinasi bobot ini dianggap representatif untuk menyeimbangkan relevansi konten dan aspek geografis dalam sistem rekomendasi yang dikembangkan.

Meskipun hasil pengujian secara observasional menunjukkan bahwa sistem mampu memberikan rekomendasi yang relevan, evaluasi tersebut masih bersifat kualitatif. Oleh karena itu, untuk memperkuat analisis, dilakukan evaluasi kuantitatif menggunakan metrik Precision@4 pada tingkat Single Query dan secara agregat melalui perhitungan rata-rata seluruh kueri pengujian dengan menggunakan metrik Precision@4 Multi Query.

Precision@4 Single Query digunakan untuk mengukur tingkat ketepatan rekomendasi dengan menghitung proporsi destinasi yang relevan dari empat rekomendasi teratas yang dihasilkan sistem untuk satu destinasi awal. Penentuan relevansi dilakukan secara manual berdasarkan kesesuaian konteks antara destinasi awal dan destinasi yang direkomendasikan, seperti kesamaan jenis wisata dan karakteristik deskripsi destinasi. Adapun hasil dari *Precision@4 single query* dapat dilihat pada tabel 6 berikut ini.

Tabel 6 Precision@4 Single Query

No	Detail Destinasi	80% Konten – 20% Jarak	40% Konten – 60% Jarak
1	pantai ujung karang	1	0.75
2	pantai ujung batee	1	1
3	pantai wisata kuala raja	1	1
4	air terjun tingkat tujuh	1	1
5	air terjun tansaran bidin	1	0.75

Berdasarkan Tabel 6, skema bobot 80% konten – 20% jarak menghasilkan nilai Precision@4 sebesar 1 pada seluruh query, yang menunjukkan bahwa keempat rekomendasi teratas yang diberikan sistem seluruhnya relevan dengan destinasi awal. Sementara itu, pada skema bobot 40% konten – 60% jarak, terdapat beberapa query yang menghasilkan nilai Precision@4 sebesar 0.75, yang menunjukkan bahwa tiga dari empat rekomendasi teratas dinilai relevan. Hal ini mengindikasikan bahwa peningkatan bobot jarak dapat memengaruhi relevansi konten rekomendasi pada beberapa kasus tertentu. Selain itu, untuk memperoleh gambaran performa sistem secara keseluruhan, dilakukan evaluasi Precision@4 Multi Query dengan menghitung nilai rata-rata Precision@4 dari seluruh query yang diuji. Hasil evaluasi tersebut ditunjukkan pada tabel 7 berikut ini.

Tabel 7 Precision@4 multi query

No	Skema Bobot	Rata – rata Precision@4
1	80% Konten – 20% Jarak	1
2	40% Konten – 60% Jarak	0.9

Berdasarkan Tabel 7, skema bobot 80% konten – 20% jarak memperoleh nilai rata-rata Precision@4 tertinggi, yaitu 1.00, yang menunjukkan bahwa seluruh rekomendasi teratas pada semua query bersifat relevan. Kemudian, skema bobot 40% konten – 60% jarak memperoleh nilai rata-rata Precision@4 sebesar 0.90, yang menandakan bahwa sebagian kecil rekomendasi mengalami penurunan relevansi akibat dominasi faktor jarak. Meskipun demikian, nilai Precision@4 yang tetap tinggi menunjukkan bahwa sistem masih mampu memberikan rekomendasi yang relevan secara keseluruhan.

4.4 Implementasi Halaman Detail Destinasi Wisata

Bagian Implementasi ini, penulis membuat secara manual tampilan user interface pada halaman detail destinasi tempat wisata. Dimana penulis membuat tampilan yang cukup sederhana sehingga wisatawan dapat dengan mudah memahami. Berikut tampilan halaman detail destinasi tempat wisata dapat dilihat pada gambar 10 berikut ini.

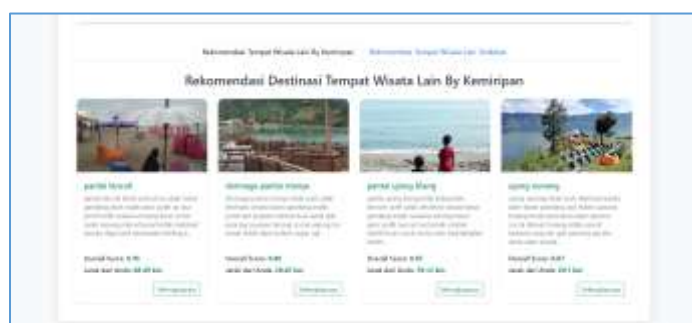


Gambar 10 Halaman detail destinasi tempat wisata

Pada gambar 10 ini dapat dilihat halaman detail destinasi tempat wisata memberikan informasi pada bagian sebelah kiri terdapat gambar destinasi tempat wisata dan pada bagian sebelah kanan terdapat informasi berupa nama destinasi, deskripsi destinasi tempat wisata secara detail, koordinat destinasi tempat wisata, dan koordinat wisatawan.

4.5 Implementasi Rekomendasi Destinasi Tempat Wisata

Implementasi rekomendasi destinasi wisata ini masih berada pada halaman detail destinasi wisata. Setiap wisatawan mengakses halaman suatu detail destinasi, maka sistem memunculkan 8 destinasi wisata lain yang direkomendasikan berdasarkan hasil perpaduan antara algoritma cosine similarty dan haversine formula dengan bobot yang telah ditentukan. Adapun tampilan rekomendasi destinasi wisata agar menarik dan mudah dipahami oleh wisatawan dapat dilihat pada gambar 11 sebagai berikut.



Gambar 11 Rekomendasi destinasi tempat wisata

Penelitian sebelumnya dengan metode *collaborative filtering* mengharuskan pengguna melakukan login atau memiliki *id_user* aktif untuk memperoleh rekomendasi. Meskipun menggunakan Google Colaboratory, sistem tersebut tetap bergantung pada preferensi pengguna dan

belum terimplementasi dalam bentuk aplikasi yang mudah diakses [1]. Berbeda dengan penelitian ini, sistem rekomendasi tidak memerlukan login atau *id_user* karena rekomendasi dihitung berdasarkan kemiripan nama dan deskripsi destinasi. Sistem memberikan delapan rekomendasi destinasi wisata, terdiri dari empat rekomendasi berbasis kemiripan dan empat rekomendasi berbasis kedekatan jarak. Selain itu, sistem diimplementasikan dalam bentuk aplikasi website memudahkan wisatawan.

5 Kesimpulan

Penelitian ini bertujuan untuk membangun sistem rekomendasi destinasi wisata di Provinsi Aceh berbasis *Content-Based Filtering* dengan mengombinasikan algoritma *Cosine Similarity* dan *Haversine Formula* guna mempertimbangkan kemiripan konten dan kedekatan geografis secara bersamaan. Berdasarkan hasil pengujian kualitatif dan kuantitatif menggunakan metrik *Precision@4*, tujuan penelitian telah tercapai, di mana sistem mampu memberikan rekomendasi destinasi wisata yang relevan dan konsisten terhadap preferensi pengguna. Sistem yang dikembangkan berhasil merekomendasikan destinasi dengan tingkat ketepatan yang tinggi, terutama pada skema bobot yang mengutamakan kemiripan konten, tanpa mengabaikan aspek jarak lokasi. Mamfaat utama dari penelitian ini adalah penyediaan model sistem rekomendasi wisata berbasis konten yang mengintegrasikan analisis nama destinasi, deskripsi destinasi dan informasi geografis, sehingga dapat menjadi solusi pendukung bagi wisatawan dalam menemukan destinasi wisata yang sesuai serta menjadi referensi bagi pengembangan penelitian sistem rekomendasi wisata di masa mendatang. Untuk pengembangan penelitian selanjutnya, disarankan agar sistem menggunakan perhitungan jarak berdasarkan kedekatan rute aktual sebagai pengganti jarak garis lurus, sehingga hasil rekomendasi yang diberikan dapat menjadi lebih akurat.

Referensi

- [1] F. Fahira Murzani and D. B. Arianto, "Implementasi Metode *Collaborative Filtering* pada Algoritma Sistem Rekomendasi Destinasi Wisata di Aceh," 2023.
- [2] K. M. Sukiakhy and C. V. R. Jummi, "Sistem Pendukung Keputusan Pemilihan Objek Wisata Ace menggunakan Metode *Simple Additive Weighting (SAW)*," *Jurnal Komputer dan Informatika*, Vol. 9, No. 1, pp. 74–80, Mar. 2021, DOI: 10.35508/jicon.v9i1.3835.
- [3] D. A. Pratiwi and A. Qoiriah, "Sistem Rekomendasi *Wedding Organizer* menggunakan Metode *Content-based Filtering* dengan Algoritma *Random Forest Regression*," *Journal of Informatics and Computer Science*, Vol. 03, 2022, DOI:10.26740/jinacs.v3n03.p231-239.
- [4] C. Salsabilla and D. W. Utomo, "Sistemasi: Jurnal Sistem Informasi Sistem Rekomendasi Wisata Kabupaten Pekalongan dengan *Content based Filtering Pekalongan Regency Tourism Recommendation System with Content based Filtering*," 2025. [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [5] S. O. S. Husna, I. N. Fajri, and Y. Pristyanto, "Sistem Rekomendasi Pariwisata Gunungkidul berbasis Web menggunakan Metode *Content-based*," *JURSISTEKNI*, Aug. 2024.
- [6] K. S. Y. Putri, I Made Agus Dwi Suarjaya, and Wayan Oger Vihikan, "Sistem Rekomendasi *Skincare* menggunakan Metode *Content based Filtering* dan *Collaborative Filtering*," *Decode: Jurnal Pendidikan Teknologi Informasi*, Vol. 4, No. 3, pp. 764–774, Aug. 2024, DOI: 10.51454/decode.v4i3.601.
- [7] L. Cahyani, N. Sephiana, M. Tahir, J. Aisyiah, and S. Artikel, "Jurnal Explore IT|31 Sistem Rekomendasi Wisata Kuliner Madura menggunakan *Content based Filtering Madura Culinary Tourism Recommendation System using Content based Filtering*," 2024, DOI: 10.35891/explorit.
- [8] F. Djiwadikusumah, G. H. Irawan, and R. H. Al-Fadilah, "Web Scraping Situs *E-Commerce* menggunakan Teknik *Parsing DOM*," Vol. 7, No. 2, p. 2021.
- [9] S. Siska, I. N. Fajri, R. Rayhan, A. Pratama, and A. N. Rohman, "Sistem Rekomendasi Wisata Magelang menggunakan Metode *Collaborative Filtering*," *Jurnal Eksplora Informatika*, Vol. 14, No. 1, pp. 63–70, Sep. 2024, DOI: 10.30864/eksplora.v14i1.1084.
- [10] A. Kurniaji and R. C. N. Santi, "Implementasi Metode *Content based Filtering* pada Pemilihan Komik," *Jurnal Informatika*, Vol. 10, No. 2, pp. 109–117, Oct. 2023, DOI: 10.31294/inf.v10i2.16113.

<http://sistemasi.ftik.unisi.ac.id>

- [11] R. H. Mondy and A. W. Wijayanto, “*Recommendation System with Content-based Filtering Method for Culinary Tourism in Mangan Application*,” Dec. 2019, doi.org/10.20961/itsmart.v8i2.35008.
- [12] F. Alghifari and D. Juardi, “Fauzan Alghifari Penerapan Data Mining pada Penerapan Data Mining pada Penjualan Makanan dan Minuman menggunakan Metode *Algoritma Naïve Bayes*,” Sep. 2021, DOI: 10.33884/jif.v9i02.3755.
- [13] A. N. Rohman, M. N. Fauzy, and A. Sa’di, “Sistem Rekomendasi Buku menggunakan *Algoritma Rabin-Karp*,” Jurnal Eksplora Informatika, Vol. 12, No. 1, pp. 86–94, Mar. 2024, DOI: 10.30864/eksplora.v12i1.1074.
- [14] A. J. Hanif, M. N. Farid, and B. Hasanah, “Penerapan *Natural Language Processing* untuk Klasifikasi Bidang Minat berdasarkan Judul Tugas Akhir,” Jurnal Sistim Informasi dan Teknologi, pp. 41–49, Feb. 2023, DOI: 10.37034/jsisfotek.v5i1.196.
- [15] A. Sanjaya, A. B. Setiawan, U. Mahdiyah, I. N. Farida, A. R. Prasetyo, and U. Nusantara PGRI Kediri, “Pengukuran Kemiripan Makna menggunakan *Cosine Similarity* dan *Basis Data Sinonim Kata Measurement of Meaning Similarity using Cosine Similarity and Word Synonyms*,” Vol. 10, No. 4, 2023, DOI: 10.25126/jtiik.2023106864.
- [16] A. H. Kridalaksana, “Penerapan Formula *Haversine* pada Sistem Informasi Geografis Pencarian Jarak Terdekat Lokasi Lapangan Futsal,” 2018, doi.org: 10.30872/jim.v13i1.1027
- [17] A. Muliawan, T. Badriyah, and I. Syarif, “Membangun Sistem Rekomendasi Hotel dengan *Content based Filtering* menggunakan *K-Nearest Neighbor* dan *Haversine Formula*,” Technomedia Journal, Vol. 7, No. 2, pp. 231–247, Sep. 2022, DOI: 10.33050/tmj.v7i2.1893.
- [18] Aprianto, “Penerapan *Algoritma Content-based Filtering* untuk Rekomendasi Destinasi Wisata pada Aplikasi *Picnicker*,” 2022.