

## Membangun Sistem Rekomendasi Hotel dengan Content Based Filtering Menggunakan K-Nearest Neighbor dan Haversine Formula

Agung Muliawan<sup>1</sup>  
Tessy Badriyah<sup>2</sup>  
Iwan Syarif<sup>3</sup>

Program Studi Teknik Informatika dan Komputer  
Politeknik Elektronika Negeri Surabaya  
Kampus Jl. Raya ITS, Keputih, Kec. Sukolilo, Kota SBY, Jawa Timur 60111, Indonesia  
E-mail: [agungmuliawan96@gmail.com](mailto:agungmuliawan96@gmail.com)<sup>1</sup>; [tessy@pens.ac.id](mailto:tessy@pens.ac.id)<sup>2</sup>; [iwanarif@pens.ac.id](mailto:iwanarif@pens.ac.id)<sup>3</sup>



Notifikasi Penulis  
18 Juli 2022  
Akhir Revisi  
16 September 2022  
Terbit  
01 Oktober 2022

Muliawan, A., Badriyah, T., & Syarif, I. (2022). Membangun Sistem Rekomendasi Hotel dengan Content Based Filtering Menggunakan K-Nearest Neighbor dan Haversine Formula. *Technomedia Journal*, 7(2 Oktober), 231–247.

<https://doi.org/10.33050/tmj.v7i2.1893>

### ABSTRAK

Peningkatan pertumbuhan industri hotel pada tiap tahunnya dan preferensi konsumen yang bervariasi dalam kebutuhan layanan hotel mengakibatkan konsumen lebih konsumtif dalam memilih hotel. Kurangnya pilihan kriteria bobot pada penyedia layanan hotel mengakibatkan konsumen mengalami kesulitan dalam memilih hotel yang sesuai dengan preferensinya, sehingga diperlukan sebuah sistem rekomendasi hotel sebagai pilihan alternatif dalam memilih hotel. Dalam penelitian ini digunakan permodelan Case Based Reasoning (CBR) untuk memberikan pembelajaran kepada sistem. Pilihan dari user pada pilihan hotel secara otomatis akan disimpan ke dalam database dan dijadikan sebagai data training sehingga sistem akan mendapatkan informasi secara berkelanjutan. Pada penelitian ini diberikan tiga jenis kebutuhan antara lain Kebutuhan Prioritas (KP), Kebutuhan Umum (KU) dan Kebutuhan Tambahan (KT) dan atribut yang digunakan terdapat enam yaitu: fasilitas, lokasi, harga, tipe kamar, bintang dan skor yang sangat memengaruhi hasil rekomendasi. Untuk setiap nilai bobot yang ada, dilakukan uji validitas bobot kepentingan menggunakan pairwise comparison matrix (PCM) sehingga nilai bobot menjadi valid dengan rentang nilai 0-1. Selain itu penerapan content based filtering menggunakan metode haversine formula dan K-Nearest Neighbor (KNN) dalam menentukan nilai terdekat dengan data training. Dari eksperimen, didapatkan hasil pengukuran performansi yang memuaskan berupa rata-rata kemiripan (similarity) sebesar 84.50%.

Kata kunci : Case Based Reasoning, Content Based Filtering, Haversine Formula, K-Nearest



### ABSTRACT

*The increasing growth of the hotel industry every year and varying consumer preferences in the needs of hotel services have made consumers more consumptive in choosing hotels. The lack of choice of weight criteria for hotel service providers causes consumers to have difficulty choosing a hotel that suits their preferences, so a hotel recommendation system is needed when selecting a hotel. In this study, Case-Based Reasoning (CBR) modeling was used to provide learning to the system. The user's choice of hotel will automatically be saved into the database and used as training data so that the system will receive information on an ongoing basis. In this study, three types of needs are given, including Priority Needs (KP), General Needs (KU), and Additional Needs (KT), and the attributes used are six, namely: facilities, location, price, room type, stars and scores that significantly affect the results. Recommendation. For each weight value, the importance weight validity test was carried out using a pairwise comparison matrix (PCM) so that the weight value became valid with a value range of 0-1. In addition, the application of content-based filtering uses the haversine formula and K-Nearest Neighbor (KNN) methods to determine the closest value to the training data. From the experiment, the performance measurement results were satisfactory in the form of an average similarity of 84.50%.*

*Keywords: Case Based Reasoning, Content Based Filtering, Haversine Formula, K-Nearest Neighbor, Recommendation System*

### PENDAHULUAN

Perkembangan teknologi dan arus informasi yang sangat pesat terutama di website yang terus berkembang dengan menyediakan berbagai kebutuhan dan informasi salah satunya hotel [1]. Hotel merupakan akomandasi dan ditujukan kepada konsumen untuk dapat menginap dan memperoleh layanan atau fasilitas yang telah disediakan sesuai dengan kebutuhan setiap konsumennya [2]. Apabila diperhatikan pesatnya informasi dan jumlah hotel mengakibatkan konsumen lebih selektif dalam memilih hotel yang sesuai dengan preferensinya mengingat setiap konsumen memiliki beragam preferensi yang beragam. Dari kondisi tersebut dibutuhkan layanan dan fasilitas yang berbeda-beda pada setiap hotel yang ada [3].

Atribut yang digunakan pada penelitian ini terdapat 6 (enam) kriteria yaitu fasilitas hotel, tarif hotel, tipe kamar, bintang, review dan lokasi hotel. Kemudian, keenam kriteria tersebut dipetakan ke dalam tiga kategori yaitu Kebutuhan Prioritas (KP), Kebutuhan Umum (KU) dan Kebutuhan Tambahan (KT). Pemberian nilai bobot pada setiap kebutuhan pengguna dilakukan karena setiap pengguna memiliki preferensi yang berbeda-beda. Pemberian nilai bobot dilakukan validasi bobot terlebih dahulu sebelum dilakukan proses algoritma KNN [4]. Validasi bobot pada penelitian ini menggunakan metode *Pairwise Comparison Matrix* (PCM) dengan melihat nilai Consistency Random (CR) dalam menentukan nilai batas diluar konsistensi [5].

Penelitian ini menggunakan terminologi yang sama dengan yang dilakukan oleh

peneliti [6] bahwa data sebelumnya merupakan history pemesanan hotel user lain selama melakukan kunjungan ke dalam situs rekomendasi. Yang selanjutnya dijadikan sebagai data training. Penggunaan Metode Content Based Filtering atau CBF dengan melihat transaksi pemesanan hotel pengguna sebelumnya [7]. Untuk algoritma kemiripan digunakan K-Nearest Neighbor (KNN) yang melakukan klasifikasi dengan melihat jarak terdekat dan haversine formula dalam menentukan jarak terdekat. Sejalan dengan penelitian yang dilakukan oleh Hartatik & Rosyid (2020) [8] Algoritma cosine similarity cocok dipergunakan pada data yang tidak terdapat nilai subjektif seperti similaritas antara teks berdasarkan kemiripan kata dalam teks.

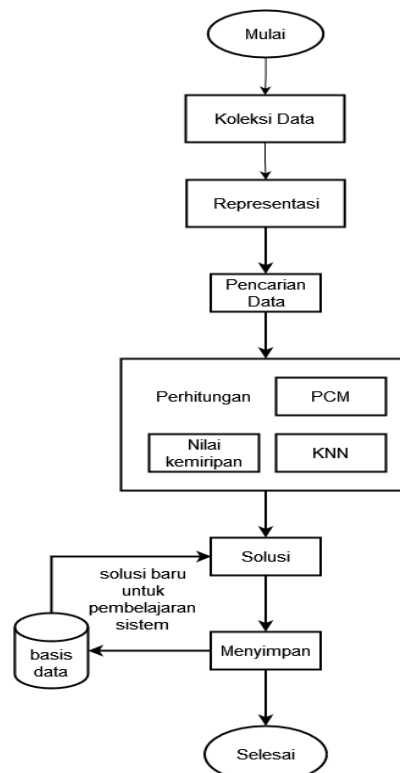
Tujuan dari penelitian atau research ini untuk mempermudah konsumen dalam memberikan alternatif pilihan kepada konsumen sesuai dengan referensinya.

## PERMASALAHAN

Faktanya jumlah dan variasi informasi yang beredar di internet sangatlah banyak dari segi layanan dan fasilitas bahkan harga penginapan sehingga memberikan masalah kepada pengguna dalam menentukan pilihan yang sesuai dengan preferensi [9]. Untuk mengatasi permasalahan tersebut menggunakan sistem rekomendasi, yaitu suatu teknik yang bisa menyediakan saran/pilihan sesuai dengan preferensi pengguna. Sistem rekomendasi adalah metode pemberian solusi kepada pengguna [10]. Sistem rekomendasi juga digunakan dalam menyaring informasi yang melimpah dari website untuk dijadikan informasi yang lebih spesifik kepada penggunanya [11].

## METODOLOGI PENELITIAN

### 1. Tahap Penelitian



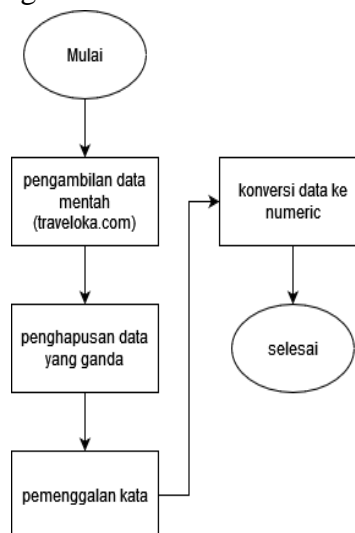
Gambar 1. Tahap Penelitian

Berdasarkan gambar 1 tahapan penelitian dapat dijelaskan sebagai berikut :

- a. Pengumpulan data. Pada tahap ini dilakukan pengambilan data hotel di situs traveloka.com berjumlah 473 data hotel bintang 3-5 di lokasi Surabaya [9]. Selanjutnya dilakukan proses *pre-processing*. Kemudian di *input* ke dalam *database* sebagai informasi hotel.
- b. Representasi. Pemberian informasi kepada user untuk menentukan pemetaan kriteria kebutuhan bobot. Pada penelitian ini digunakan 50 data *dummy* untuk dijadikan data *training*
- c. Proses validitas 3 kebutuhan bobot (Kebutuhan Prioritas, Kebutuhan Umum dan Kebutuhan Tambahan) dan perhitungan nilai kemiripan menggunakan *haversine formula* dan *KNN* pada pemetaan 6 kriteria kedalam kebutuhan (Fasilitas, Harga, Tipe Kamar, Bintang, Skor dan Lokasi)
- d. Proses pemberian rekomendasi kepada user berdasarkan ranking yang memiliki persentase tertinggi yang didapatkan pada tahapan sebelumnya.
- e. Menyimpan respon user dari rekomendasi yang ditampilkan pada tahap 4. Saat user menentukan pilihan hotel yang dibutuhkan data akan yang disimpan secara otomatis untuk dijadikan data *training*

## 2. Koleksi Data

Tahapan koleksi data sebagai berikut :

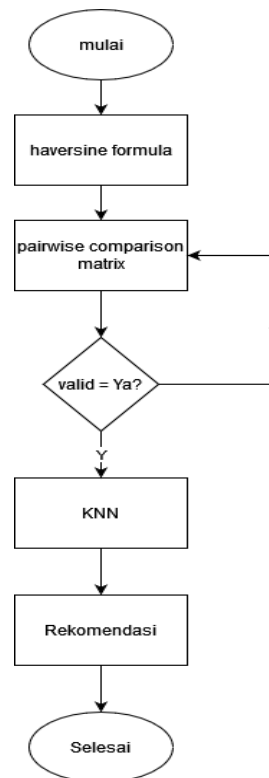


**Gambar 2.** Tahapan koleksi data

Berdasarkan gambar 2 diatas data yang digunakan adalah data hotel kota surabaya [12] dan di situs traveloka.com [13]. Data yang didapatkan dilakukan proses pre-processing terlebih dahulu dalam mendapatkan informasi data yang dapat diproses perhitungan kemiripan.

## 3. Proses Perhitungan

Proses selanjutnya dimana penggunaan informasi dan pengetahuan sehingga menghasilkan penarikan kesimpulan atau solusi rekomendasi hotel. Berikut travmerupakan alur proses perhitungan rekomendasi hotel pada Gambar 3 :



**Gambar 3.** Tahapan proses perhitungan

Penjelasan tahapan perhitungan sebagai berikut :

a. Haversine Formula

Metode *haversine formula* adalah persamaan untuk navigasi berdasarkan jarak lingkaran 2 titik pada permukaan bumi [14]. Berikut merupakan rumus yang digunakan menghitung kedekatan terdekat lokasi pengguna dan lokasi hotel menggunakan metode *haversine formula* [15]:

$$\Delta lat = lat2 - lat1 \quad (1)$$

$$\Delta long = long2 - long \quad (2)$$

$$a = \sin^2\left(\frac{\Delta lat}{2}\right) + \cos(lat1) \cdot \cos(lat2) \times \sin^2\left(\frac{\Delta long}{2}\right) \quad (3)$$

$$c = 2 \operatorname{atan2}(\sqrt{a}, \sqrt{1-a}) \quad (4)$$

$$d = Rxc \quad (5)$$

Keterangan :

Lat1 : Latitude asal

Lat2 : latitude tujuan

Long1 : longitude asal

Long2 : longitude tujuan

R : jari-jari bumi 6371 km

b. Pairwise Comparison Matrix

Pada penelitian ini dilakukan validasi pada setiap kebutuhan kepentingan (KP, KU, KT). Tujuan validasi bobot untuk mendapatkan nilai konsistensi nilai bobot yang

wajar. Proses validasi dimulai dari penentuan nilai kepentingan setiap bobot kemudian dilakukan matriks perbandingan berpasangan dengan cara perkalian tiap bobot Berikut adalah tampilan comparison matrix A yang ditunjukkan pada Gambar 4

$$\begin{pmatrix} 1 & a_{21} & a_{13} & \dots & a_{1n} \\ a_{21} & 1 & a_{23} & \dots & a_{25} \\ a_{31} & a_{32} & 1 & \dots & a_{35} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{51} & a_{52} & a_{53} & a_{54} & 1 \end{pmatrix}$$

**Gambar 4.** Comparison Matrix [13]

Persamaan yang digunakan untuk memperoleh nilai consistency ratio :

$$CI = \frac{n \max(A) - n}{n-1} \quad (6)$$

$$CR = \frac{CI}{RI} \quad (7)$$

Keterangan :

CI = Consistency Index

n = jumlah banyak kriteria

CR = Consistency Random

RI = Random Index

Nilai dapat diterima apabila nilai CR dibawah 10% sehingga tidak perlu dilakukan revisi atau perbaikan pemberian nilai [5].

c. K-Nearest Neighbor (KNN)

Metode KNN melakukan klasifikasi objek dari data training yang memiliki jarak terdekat dengan objek tersebut dan nilai K merupakan banyaknya nilai yang terdekat. Cara kerja metode KNN ialah mencari kecocokan atau kemiripan dari kelompok objek data training dengan objek data baru atau data testing di dalam database [16]. Pada persamaan (3), merupakan rumus metode KNN yang digunakan dalam mencari nilai kemiripan untuk rekomendasi hotel [17].

$$S(P, C) = \frac{S_i \times W_i}{\sum_{i=1}^n W_i} \quad (8)$$

Keterangan :

P = Basis data solusi ke i

C = Masalah attribute ke i

n = jumlah atribut pada basis data

i = attribute individu dari 1 hingga n

S = nilai kemiripan

w = bobot

## HASIL DAN PEMBAHASAN

### A. HASIL

#### 1. Pengambilan Data

Pada penelitian data yang terkumpul sebanyak 437 data dengan 97 hotel bintang 3-5 dari situs traveloka.com. Berikut merupakan contoh data mentah sebelum dilakukan proses *pre-processing*. Berikut merupakan contoh data sample A:

Hotel : Best Western Papilio Hotel  
Fasilitas : AC, TV, wifi gratis, ruang rapat, tanpa sarapan, pusat kebugaran, restoran, parkir, kolam renang, area merokok  
Harga : Rp488.000  
Tipe Kamar : Superior Twin  
Review : 87  
Bintang : 4  
Lokasi : Jalan Ahmad Yani 176-178, Wonokromo, Surabaya

Selanjutnya dilakukan penghapusan data ganda atau double karena pengambilan data pada detail page memiliki informasi yang sama sehingga data yang semula berjumlah 437 menjadi 363. Kemudian dilakukan pemenggalan data pada attribute tipe kamar dan lokasi seperti tipe kamar Superior Queen menjadi superior dan lokasi Jalan Kedungdoro, Sawahan, Surabaya menjadi sawahan untuk mempermudah proses pemberian nilai *numeric*.

Pada tahap ini dilakukan konversi data ke *numeric* supaya dapat diproses pada proses perhitungan nilai kemiripan KNN. Konversi data ke *numeric* ini dilakukan dengan cara memberikan nilai skala pada setiap kriteria (harga, tipe kamar, bintang, skor, tipe kamar, dan lokasi) berikut merupakan skala pada kriteria :

**Tabel 1.** Nilai Skala *Numeric*

Kriteria	Item	Nilai Skala
Fasilitas	AC	Nilai 1 dan 0 (Ya atau Tidak)
	TV	
	Internet	
	Breakfast	
	Ruang Meeting	
	Fitness	
	Restoran	
	Parkiran	
	Kolam Renang	
	Area Rokok	
Harga	> 1.000.000	1
	500.000 – 1.000.000	2
	300.000 – 500.000	3
	100.000 – 300.000	4
	< 100.000	5
Tipe Kamar	Suite	1

	Deluxe	2
	Superior	3
	Standart	4
	Single	5
Review	> 98	1
	96 – 98	2
	94 – 96	3
	92 – 94	4
	< 92	5
Bintang	Bintang 5	1
	Bintang 4	2
	Bintang 3	3
Lokasi	Dukuh pakis	metode haversine formula

Berdasarkan tabel 1 menunjukkan data sampel A yang telah dilakukan proses pre-processing :

Hotel : Best Western Papilio Hotel  
 Fasilitas : 1, 0, 0, 1, 0, 0, 0, 0, 0, 1  
 Harga : 2  
 Tipe Kamar : 3  
 Review : 2  
 Bintang : 2  
 Lokasi : 1

## 2. Data Training

Pada tahap ini dilakukan representasi dari pilihan kriteria kebutuhan user yang akan digunakan sebagai data training. Pada penelitian ini data training yang digunakan berjumlah 50 data dummy sebagai perhitungan kemiripan. Berikut merupakan sampel data data training yang digunakan pada penelitian ini :

**Tabel 2.** Sample Data Training

C	Fasilitas										H	TK	RW	BT	LOK	R
	1	2	3	4	5	6	7	8	9	10						
<b>1</b>	1	0	0	1	0	0	0	0	0	1	2	3	2	2	1	1
<b>8</b>	0	0	1	0	0	1	0	0	1	0	0	3	5	0	12	3
<b>15</b>	0	1	0	0	0	0	0	0	1	1	4	2	0	2	4	1 0
<b>20</b>	0	1	0	0	1	0	0	1	1	0	1	3	0	0	11	4
<b>27</b>	1	1	0	0	1	0	1	0	0	1	3	0	3	3	15	0
<b>30</b>	1	1	1	0	0	1	0	0	1	0	1	3	0	0	11	4



Keterangan :

F1-F10: Pilihan fasilitas 1-10

H : Harga tarif kamar

TK : Tipe Kamar

RW : Skor review

BT : Bintang

LOK : Lokasi hotel

R : Result (hasil pilihan user)

C : Kasus ke-i

### 3. Perhitungan Metode

#### a. Haversine Formula

Penggunaan metode *haversine formula* dalam melihat lokasi terdekat pada tujuan pengguna. Pada penelitian ini kecamatan di Surabaya yang digunakan peneliti terdapat 14 daerah.

Sebagai contoh perhitungan kita akan menghitung jarak terdekat daerah dukuh pakis dengan 13 kecamatan lainnya. Berikut merupakan hasil tabel perbandingan dan perhitungan menggunakan metode *haversine formula* menggunakan perbandingan daerah Dukuh pakis dengan kecamatan di Surabaya.

**Tabel 3.** Perbandingan lokasi

No	Lok 2	Lat 2	Long 2
1	Sukomanunggal	-7,26427	112,69734
2	Sawahan	-7,26302	112,72118
3	Tegalsari	-7,27986	112,73606
4	Wonokromo	-7,302660	112,73310
5	Wiyung	-7,31527	112,68542
6	Gubeng	-7,27928	112,75392
7	Genteng	-7,25913	112,74807
8	Wonocolo	-7,31983	112,74203
9	Gayungan	-7,33054	112,72391
10	Tenggilis Mejoyo	-7,32036	112,75800
11	Mulyorejo	-7,26932	112,79267
12	Sukolilo	-7,28819	112,81651
13	Rungkut	-7,31898	112,80462

**Tabel 4.** Perhitungan haversine formula

No	Alat	$\Delta \log$	a	c	d	n	$\Delta h$
1	0,0002965 56	0,0002211 96	2,19863 E-08	3,68631E- 09	2,04	2,67	0,63
2	0,0003183 61	0,0006370 71	2,53384 E-08	3,06355E- 08	3,01	4,3	1,29
3	2,46E-05	0,000897	1,51E-10	5,92E-08	3,10	5,65	2,55
4	-0,00037	0,000845	3,48E-08	5,07E-08	3,73	5,86	2,13

No	Alat	$\Delta \log$	a	c	d	n	$\Delta h$
5	-0,00059	1,33 E-05	8,79E-08	1,22E-11	3,78	3,78	0,00
6	3,47 E-05	0,001208	3,01E-10	1,08E-07	4,18	7,8	3,62
7	0,000386	0,001106	3,73E-08	9,29E-08	4,60	7,51	2,91
8	-0,00067	0,001001	1,13E-07	6,91E-08	5,44	7,6	2,16
9	-0,00086	0,000685	1,85E-07	3,17E-08	5,93	6,85	0,92
10	-0,00068	0,001279	1,16E-07	1,13E-07	6,10	9,08	2,98
11	0,000208	0,001884	1,09E-08	2,65E-07	6,70	11,91	5,21
12	-0,00012	0,0023	3,64E-09	3,84E-07	7,94	12,85	4,91
13	-0,00066	0,002093	1,08E-07	3,02E-07	8,17	13,83	5,66

Pada perhitungan haversine formula didapatkan perbedaan yang tidak begitu jauh antara nilai perhitungan dan nilai real yang didapatkan dari google maps dengan rata-rata nilai perbandingan sebesar 2,8km.

#### b. Pairwise Comparison Matrix

Selanjutnya dilakukan validasi bobot menggunakan pairwise comparison matrix dimulai dari penentuan nilai bobot pada setiap kebutuhan yang ada (Kebutuhan prioritas, kebutuhan umum dan kebutuhan tambahan).

- Prioritas 3x lebih penting dari umum
- Prioritas 5x lebih penting dari tambahan
- Umum 2x lebih penting dari tambahan

Berikut merupakan perbandingan matrik berpasangan :

**Tabel 5.** Perbandingan matrik berpasangan

F	KP	KU	KT	X	$\sqrt[3]{X}$	B
KP	1	3	5	15	2,47	0,65
KU	0,33	1	2	0,67	0,88	0,23
KT	0,2	0,5	1	0,10	0,46	0,12
$\Sigma F$	1,53	4,5	8		3,81	1

Penjelasan isian tabel perbandingan matrik berpasangan :

- 1) Perbandingan KP dan KU mendapatkan nilai 0,33 karena  $KP = 1$  dan  $KU=3$  sehingga  $1/3 = 0,33$
- 2) Nilai X merupakan perkalian baris KP, KU dan KT mendapatkan nilai 15 kemudian di akar 3 untuk mendapatkan nilai 2,47
- 3) B merupakan nilai bobot dari hasil  $2,47 / 3,81$  mendapatkan 0,65

Apabila telah didapatkan nilai tiap bobot maka akan dilakukan perkalian tiap bobot seperti pada tabel perkalian tiap bobot seperti pada tabel berikut ini :

**Tabel 6.** Perkalian tiap bobot

<b>Ket</b>	<b>KP</b>	<b>KU</b>	<b>KT</b>	$\lambda$ Max
$\Sigma$	1,53	4,5	8	
$\Sigma x B$	0,99	1,03	0,98	<b>3,00</b>

Penjelasan isian tabel perkalian tiap bobot :

- 1) Nilai 0,99 pada kolom KP diperoleh dari  $\Sigma x$  nilai bobot kriteria sehingga  $1,53 \times 0,65$  (bobot nilai KP)
- 2) Hasil  $\lambda$  Max didapatkan dari penjumlahan  $\Sigma x B$  mendapatkan nilai 3
- 3) Setelah didapatkan nilai jumlah bobot maksimum akan dilakukan analisis konsistensi. Kriteria akan dibandingkan dengan kriteria lain yang dapat menjadi tidak konsisten. Pada persamaan (6) merupakan perhitungan nilai *CI* atau *Consistency index* dan persamaan (7) perhitungan nilai *CR* atau *Consistency Random*. Tujuan mencari nilai *CR* untuk mengetahui batas ketidakkonsistenan tidak boleh diatas 0,1 atau 10%

Untuk nilai *Random Index* atau *RI* digunakan tabel tetapan pada Tabel 6, maka nilai *RI* dihasilkan sebesar 0,58

**Tabel 7.** Ketetapan nilai *RI*

<b>N</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
<i>RI</i>	0,00	0,00	<b>0,58</b>	0,9

Dari persamaan tersebut didapatkan nilai *CR* sebesar 0,002. Nilai *CR* tersebut memiliki nilai lebih kecil dari 0,1 atau 10% sehingga pemberian nilai kepentingan dapat diberikan dan tidak perlu dilakukan revisi atau perbaikan. Selanjutnya nilai bobot setiap kriteria akan dijadikan acuan dalam penghitungan metode KNN.

#### c. K-Nearest Neighbor (KNN)

Selanjutnya dilakukan perhitungan kemiripan menggunakan metode KNN. Sebelum menggunakan rumus KNN di persamaan (8) diberikan nilai kemiripan terlebih dahulu kepada kriteria bintang, review, harga dan tipe kamar karena untuk kriteria fasilitas hanya skala 1(ya) dan 0 (tidak). Untuk lokasi juga sudah ditentukan menggunakan haversine formula dengan skala 1-0.

Pemberian nilai tersebut digunakan untuk mempermudah proses pemberian nilai kemiripan kriteria yang akan dibandingkan. Seperti contoh kasus user membutuhkan hotel dengan harga lebih dari Rp.1.000.000, pilihan kriteria tersebut akan dibandingkan dengan harga lainnya. Semakin tinggi kemiripan harga maka semakin mendekati angka 1 dan semakin tidak mirip diberikan angka 0. Berikut merupakan tabel kemiripan harga

**Tabel 8.** kemiripan kriteria harga

<b>Harga yang dipilih</b>	<b>Perbandingan</b>	<b>Kemiripan</b>
---------------------------	---------------------	------------------

> 1.000.000	> 1.000.000	1
> 1.000.000	500.000 – 1.000.000	0.8
> 1.000.000	300.000 – 500.000	0.6
> 1.000.000	100.000 – 300.000	0.4
> 1.000.000	< 100.000	0.2

Pada penelitian kali ini nilai kemiripan algoritma KNN bernilai dari 0 hingga 100% yang berarti semakin tinggi presentase maka kemiripan kebutuhan user semakin tinggi dengan data training yang ada. Berikut merupakan contoh sebuah pemetaan kriteria kebutuhan oleh user yang akan diproses kemiripannya dengan data training :

**Tabel 9.** kebutuhan user

Kriteria	Deskripsi	Kebutuhan
Fasilitas	Breakfast	KP
	Ruang meeting	KP
	Kolam Renang	KU
Harga	300.000 – 500.000	KP
Tipe Kamar	Deluxe	KU
Lokasi	Dukuh pakis	KP

Berikut merupakan sample perhitungan perbandingan kebutuhan user dengan *data training* pada tabel pengujian kemiripan KNN :

**Tabel 10.** perbandingan kemiripan KNN

Deskripsi	bobot	kemiripan	total
Dukuh Pakis	0,65	1	0,65
TV	0	0	0
Ruang meeting	0,65	1	0,65
Area rokok	0,12	1	0,12
Single	0,23	0,4	0,092
review 92-94	0	0	0

$$\text{kemiripan} = \frac{0,65+0,65+0,12+0,092}{0,65+0,65+0,12+0,23} = 0,91$$

Pada simulasi perhitungan KNN menggunakan data *training* menghasilkan nilai kemiripan 0,91 atau 91%. Untuk perhitungan dilakukan untuk seluruh data training yang didapatkan sehingga dengan cara yang sama untuk mendapatkan persentase kemiripan tertinggi sebagai rekomendasi kepada user.

#### 4. Pemberian Rekomendasi

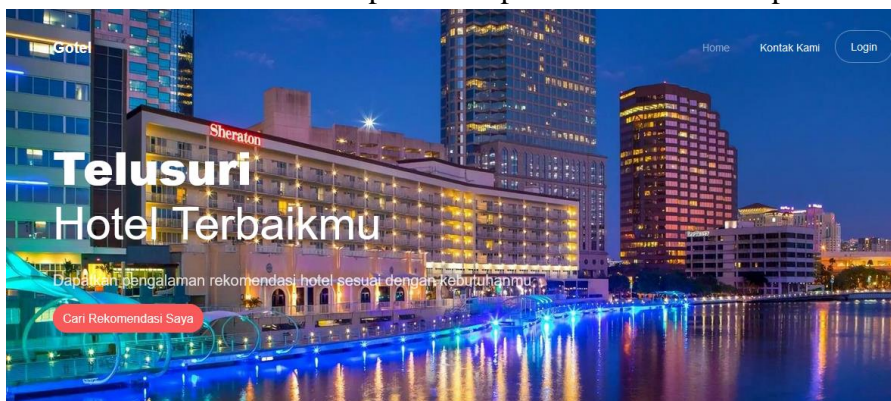
Pada tahap ini nilai kemiripan dari perhitungan KNN akan diambil untuk dijadikan rekomendasi hotel kepada user. Hasil kemiripan akan diurutkan berdasarkan nilai kemiripan tertinggi hingga terendah. Pada tahap ini ditampilkan 5 nilai teratas yang memiliki kemiripan tertinggi yang diberikan kepada user untuk dipilih. Respon dari pilihan user akan disimpan secara otomatis sebagai tambahan data *training* ke dalam sistem rekomendasi.

## B. PEMBAHASAN

Dalam menerapkan aplikasi sistem rekomendasi hotel menggunakan metode haversine formula dan KNN dilakukan tahapan pengkodean menggunakan PHP 7 dan framework CI-3 sedangkan database yang digunakan adalah MySQL.

### 1. Implementasi Antar Muka Aplikasi

Pada bagian ini dijelaskan antarmuka aplikasi sistem rekomendasi hotel. Antar muka pada aplikasi terdiri dari halaman administrator, halaman homepage dan halaman rekomendasi hotel. Berikut merupakan tampilan halaman awal aplikasi.



Gambar 4. Tampilan awal aplikasi

Pada halaman diatas merupakan tampilan awal aplikasi untuk *guest* atau pengunjung sehingga perlu dilakukan registrasi terlebih dahulu untuk menentukan proses rekomendasi. Apabila telah dilakukan registrasi user akan berubah statusnya menjadi *member* atau anggota. Berikut merupakan tampilan halaman pilhan kriteria kebutuhan pengguna.

Tujuan Lokasi  
Dukuh Pakis

**Kebutuhan Prioritas (KP)**

Fasilitas

- Air Conditioner (AC)
- Television (TV)
- Internet
- Breakfast (Sarapan)
- Ruang Meeting / Rapat
- Ruang Kebugaran / Fitness
- Restoran Cepat Saji
- Kolam Renang
- Area Parkiran
- Area Rokok

Harga

- > 1.000.000
- 500.000 – 1.000.000
- 300.000 – 5.000.000
- 100.000 – 300.000

**Kebutuhan Umum (KU)**

Fasilitas

- Air Conditioner (AC)
- Television (TV)
- Internet
- Breakfast (Sarapan)
- Ruang Meeting / Rapat
- Ruang Kebugaran / Fitness
- Restoran Cepat Saji
- Kolam Renang
- Area Parkiran
- Area Rokok

Harga

- > 1.000.000
- 500.000 – 1.000.000
- 300.000 – 5.000.000
- 100.000 – 300.000

**Gambar 5.** Pilihan kriteria kebutuhan pengguna

Hasil dari pilihan kriteria kebutuhan oleh user akan di tampilkan dengan menekan tombol rekomendasi di halaman pilihan kriteria kebutuhan di atas menghasilkan hotel yang memiliki presentase kemiripan tertinggi hingga terendah. Berikut merupakan tampilan halaman rekomendasi hotel.

**Gotel** Home Rekomendasi Hotel Kontak Kami Logout

Hotel	Location	Price	Star Rating	Kemiripan
Garden Palace Hotel	Surabaya	500.000 - 1.000.000	79.38%	79.35%
Bonnet Hotel	Surabaya	100.000 - 300.000	79.13%	79.13%
Singgasana Hotel	Surabaya	500.000 - 1.000.000	79.84%	76.84%
Tab Hotel	Darmo Permai Surabaya	100.000 - 300.000	78.68%	78.08%

**Gambar 6.** Rekomendasi hotel

Untuk melihat perhitungan dari metode *haversine formula* dan KNN dapat dilihat dengan menekan tombol lihat perhitungan pada gambar diatas. Berikut merupakan tampilan perhitungan metode *haversine formula* dan perhitungan metode K-NN.

Perhitungan

Pilihan Kebutuhan Hotel

Tujuan Anda

Dukuh Pakis

Kebutuhan Prioritas

- Ac
- Tv
- Kolam Renang
- Harga : 100.000 – 300.000

Kebutuhan Utama

- Tipe Kamar : Deluxe
- Kelas : Bintang 3

Hitung Lokasi

1			
Jenis	Lokasi	Lat	Long
Lokasi Asal	Dukuh Pakis	-7.28127	112.68466
Lokasi Pemandang	Dukuh Pakis	-7.28127	112.68466

- $\text{delta\_lat} = (3.14 / 180) * (-7.28127 - -7.28127)$
- $\text{delta\_long} = (3.14 / 180) * (112.68466 - 112.68466)$
- $\text{nilai\_a} = \sin(0 / 2)$

Gambar 7. Tampilan perhitungan haversine formula

Hitung Metode

1. Fairfield By Marriott Surabaya [97.89%]

- Harga :500.000 – 1.000.000
- Lokasi :Dukuh Pakis
- Review :96 – 98
- Tipe Kamar :Deluxe

Nama	Bobot	Kemiripan	Total
AC	0	0	0
TV	0.65	1	0.65
INTERNET	0	0	0
BREAKFAST	0	0	0
RUANG MEETING	0	0	0
FITNESS	0	0	0
RESTORAN	0	0	0
PARKIRAN	0	0	0
KOLAM RENANG	0	0	0
AREA ROKOK	0	0	0
HARGA	0.65	1	0.65
TIPE KAMAR	0.23	0.8	0.184

Pengujian

Pada tahap pengujian dilakukan menghitung akurasi kemiripan atau *similarity* hotel terhadap metode KNN. Pada penelitian ini menggunakan 50 data *training* secara dummy dan akan diambil 5 hasil rekomendasi hotel dengan kemiripan tertinggi sebagai rekomendasi user. Berikut pada tabel 9 merupakan nilai kemiripan tertinggi :

Tabel 11. uji kemiripan

Kasus Ke	Hotel	Kemiripan
1	Fairfield By Marriott Surabaya	92,35 %
11	Singgasana Hotel Surabaya	85,49 %
42	Java Paragon Hotel And Residence	83,12 %
18	Tab Hotel Darmo Permai Surabaya	82,30 %
20	The Life Styles Hotel Surabaya	79,21%
	Rata-rata kemiripan	84,50%

Pada perhitungan akurasi rata-rata kemiripan tertinggi mendapatkan nilai sebesar 84,50% maka kemiripan tersebut memuaskan karena mendekati nilai 100%.

## KESIMPULAN

Pada penelitian ini mengerjakan penerapan metode *content based filtering* menggunakan K-NN dan Haversine Formula dalam menentukan rekomendasi hotel. Hasil dari penelitian ini implementasi metode rekomendasi berupa aplikasi *website* yang dapat menghasilkan alternatif rekomendasi hotel berdasarkan preferensi *user*. *User* dapat memilih kriteria dan memetakan sesuai dengan kebutuhannya. Rekomendasi hotel di urutkan berdasarkan kemiripan tertinggi dengan cara membandingkan data *training* dan respon dari *user*. Data *training* sendiri dapat bertambah secara otomatis saat *user* menyimpan respon pilihan hotelnya. Pada penelitian ini validasi bobot kebutuhan mendapatkan nilai 0,002 yang berarti valid karena dibawah 0,1. Berdasarkan pengujian akurasi kemiripan didapatkan nilai rata-rata sebesar 84,50%.

## SARAN

Harapan kedepannya untuk penelitian ini yaitu program ini kedepannya akan terus berkembang dengan pesat dan menyesuaikan peningkatan pertumbuhan dari industri hotel pada setiap tahunnya.

## DAFTAR PUSTAKA

- [1] S. M. Al-Ghuribi and S. A. M. Noah, "Multi-criteria review-based recommender system—the state of the art," *IEEE Access*, vol. 7, pp. 169446–169468, 2019.
- [2] H. Nurrahmi and H. R. Isnanto, "PENGUNAAN K-NEAREST NEIGHBOR DALAM PERANCANGAN APLIKASI PEMESANAN HOTEL DI DEPOK," *Sigma-Mu*, vol. 11, no. 2, pp. 47–55, 2019.
- [3] N. P. A. Mentayani, I. P. Satwika, I. G. A. P. D. Putri, A. A. I. I. Paramitha, and T. Tiawan, "Analisis Dan Perancangan User Interface Sistem Informasi Pembayaran Mahasiswa STMIK Primakara Berbasis Web," *Technomedia Journal*, vol. 7, no. 1 Juni, pp. 78–89, 2022.
- [4] A. Salsabila, R. Yunita, and C. Rozikin, "Identifikasi Citra Jenis Bunga menggunakan Algoritma KNN dengan Ekstraksi Warna HSV dan Tekstur GLCM," *Technomedia Journal*, vol. 6, no. 1 Agustus, pp. 124–137, 2021.
- [5] A. Amanaturohim and S. Wibisono, "Penentuan Parameter Terbobot Menggunakan Pairwise Comparison Untuk CBR Deteksi Dini Penyakit Mata," *J-SAKTI (Jurnal Sains*



- Komputer dan Informatika*), vol. 5, no. 1, pp. 280–294, 2021.
- [6] B. Walek and P. Spackova, “Content-based recommender system for online stores using expert system,” in *2018 IEEE First International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*, 2018, pp. 164–165.
- [7] T. A. Diajukan, M. Salah, S. Persyaratan, and M. Derajat, “SISTEM REKOMENDASI OBJEK PARIWISATA DI PONTIANAK BERBASIS ANDROID MENGGUNAKAN METODE CONTENT-BASED FILTERING,” 2020.
- [8] H. Hartatik and R. Rosyid, “PENGARUH USER PROFILING PADA REKOMENDASI SISTEM MENGGUNAKAN K MEANS DAN KNN,” *Journal of Information System Management (JOISM)*, vol. 2, no. 1, pp. 13–18, 2020.
- [9] K. Wahyudi, J. Latupapua, R. Chandra, and A. S. Girsang, “Hotel content-based recommendation system,” in *Journal of Physics: Conference Series*, 2020, vol. 1485, no. 1, p. 012017.
- [10] T. Badriyah, R. Fernando, and I. Syarif, “Sistem Rekomendasi Content Based Filtering Menggunakan Algoritma Apriori,” *Konferensi Nasional Sistem Informasi (KNSI) 2018*, 2018.
- [11] N. N. Qomariyah and A. N. Fajar, “Recommender system for e-learning based on personal learning style,” in *2019 international seminar on research of information technology and intelligent systems (ISRITI)*, 2019, pp. 563–567.
- [12] “Badan Pusat Statistik.”  
<https://surabayakota.bps.go.id/statictable/2020/06/05/742/jumlah-hotel-menurut-jenis-dan-kecamatan-tahun-2019.html> (accessed Sep. 15, 2022).
- [13] “1.550 Hotel di Surabaya - Booking Hotel Online.” <https://www.traveloka.com/id-id/hotel/indonesia/city/surabaya-103570> (accessed Sep. 15, 2022).
- [14] Y. Yulianto, R. Ramadiani, and A. H. Kridalaksana, “Penerapan Formula Haversine Pada Sistem Informasi Geografis Pencarian Jarak Terdekat Lokasi Lapangan Futsal,” *Inform. Mulawarman J. Ilm. Ilmu Komput*, vol. 13, no. 1, p. 14, 2018.
- [15] R. Hidayati and N. Mutiah, “Penerapan Metode Haversine Formula Pada Pencarian Lokasi Fasilitas Kesehatan Terdekat,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 1, pp. 278–286, 2022.
- [16] H. S. Hopipah and R. Mayasari, “Optimasi Backward Elimination untuk Klasifikasi Kepuasan Pelanggan Menggunakan Algoritme k-nearest neighbor (k-NN) and Naive Bayes,” *Technomedia Journal*, vol. 6, no. 1, pp. 99–110, 2021.
- [17] R. A. Yunmar, “Sistem Rekomendasi Pemilihan Hotel dengan Case Based Reasoning,” *Jurnal Rekayasa Dan Teknologi Elektro*, vol. 11, no. 2, pp. 53–60, 2017.