

# Recommendations of Thesis Supervisor using the Cosine Similarity Method

*By* Hairani Hairani

# Rekomendasi Dosen Pembimbing Skripsi menggunakan Metode *Cosine Similarity*

## *Recommendations of Thesis Supervisor using the Cosine Similarity Method*

Hairani Hairani\*, Mujahid

Ilmu Komputer, Fakultas Teknik, Universitas Bumigora  
BTN Griya Pesona Alam Blok E23, Mataram, Indonesia

\*e-mail: [hairani@universitasbumigora.ac.id](mailto:hairani@universitasbumigora.ac.id)

(received: 6 April 2022, revised: 18 Juli 2022, accepted: 24 Juli 2022)

### Abstrak

Selama proses penulisan skripsi, peran dosen pembimbing sangat dibutuhkan agar penyelesaiannya tepat waktu sehingga kualitas skripsi terjaga. Salah satu permasalahan dalam penentuan dosen pembimbing skripsi S1 ilmu komputer Universitas Bumigora adalah adanya unsur subyektifitas ketua prodi dalam penentuan dosen pembimbing skripsi. Tidak hanya itu, proses pemilihan dosen pembimbing dilakukan secara manual sehingga dapat memakan waktu lama dan dapat memperlambat pengerjaan skripsi mahasiswa. Pengerjaan skripsi mahasiswa menjadi terlambat dan tidak tepat waktu, jika topik skripsinya tidak sesuai dengan kompetensi dosen. Penelitian ini bertujuan implementasi metode *cosine similarity* untuk rekomendasi dosen pembimbing skripsi S1 ilmu komputer Universitas Bumigora. Tahapan penelitian ini terdiri dari pengumpulan dokumen skripsi, pra-pengolahan teks (*Case Folding, Tokenization, Filtering, Stemming*), Pembobotan kata dengan TF-IDF, implementasi metode *cosine similarity*, dan pengujian akurasi. Data yang digunakan sebanyak 113 dokumen skripsi yang terbagi sebagai data *training* sebanyak 90 dokumen dan data *testing* 23 dokumen. Berdasarkan hasil pengujian pada data *testing*, metode *cosine similarity* dapat merekomendasikan dosen pembimbing secara benar sebanyak 21 dari 23 dokumen skripsi dengan akurasi sebesar 91.3%. Dengan demikian, metode *cosine similarity* dapat diterapkan pada kasus pemilihan dosen pembimbing skripsi S1 ilmu komputer Universitas Bumigora karena memiliki akurasi yang sangat baik.

**Kata kunci:** Rekomendasi Dosen Pembimbing, Metode *Cosine Similarity*, Penambangan Teks.

### Abstract

During the thesis writing process, the role of the supervisor is needed so that the completion of the thesis is timely so that the quality of the thesis is maintained. One of the problems in determining the thesis supervisor for undergraduate computer science at Bumigora University is the subjectivity of the head of the study program in determining the thesis supervisor. Not only that, the selection process for supervisors is done manually so that it can take a long time and can slow down student thesis work. Students' thesis work will be late and not on time if the thesis topic is not in accordance with the competence of the lecturer. This study aims to apply the cosine similarity method to the recommendation of a thesis supervisor for undergraduate computer science at Bumigora University. The stages of this research consist of collecting thesis documents, pre-processing text (*Case Folding, Tokenization, Filtering, Stemming*), word weighting with TF-IDF, implementation of the cosine similarity method, and accuracy testing. The data used are 113 thesis documents which are divided as training data as many as 90 documents and testing data 23 documents. Based on the testing data on the test, the cosine similarity method can correctly recommend 21 of 23 thesis documents with an accuracy of 91.3%. Thus, the cosine similarity method can be applied to the case of selecting a thesis supervisor for undergraduate computer science at Bumigora University because it has very good accuracy.

**Keywords:** Thesis Supervisor Recommendations; Cosine Similarity Method; Text Mining.

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

## 1 Pendahuluan

Salah satu persyaratan mahasiswa untuk mendapatkan gelar sarjana adalah pengerjaan skripsi. Skripsi merupakan proses penyelesaian suatu permasalahan dengan kaidah akademik. Dalam penyusunan skripsinya, mahasiswa membutuhkan peran dosen pembimbing. Selama proses penulisan skripsi, peran dosen pembimbing sangat dibutuhkan agar penyelesaiannya tepat waktu dan kualitas skripsi terjaga. Salah satu permasalahan dalam penentuan dosen pembimbing skripsi S1 ilmu komputer Universitas Bumigora adalah adanya unsur subyektifitas ketua prodi dan penentuan dosen pembimbing skripsi. Tidaknya hanya itu, proses penentuan dosen pembimbing dilakukan secara manual sehingga membutuhkan waktu lama dan dapat memperlambat pengerjaan skripsi mahasiswanya. Kesalahan penentuan dosen pembimbing dapat mengakibatkan pengerjaan skripsi menjadi terlambat dan tidak tepat waktu, jika topik skripsi mahasiswa tidak sesuai dengan kompetensi dosen. Solusi yang ditawarkan adalah pembuatan aplikasi rekomendasi dosen pembimbing skripsi dengan pencocokan judul skripsi yang telah dibimbing sebelumnya dengan bidang keahlian dosen menggunakan metode *similarity*. Namun, sebelum dilakukan pembuatan sistem rekomendasi dosen pembimbing skripsi, dibutuhkan sebuah metode *similarity* rekomendasi yang memiliki akurasi yang sangat baik. Oleh karena itu, penelitian ini berfokus untuk menemukan model atau metode yang tepat pada rekomendasi dosen pembimbing.

Penelitian sebelumnya tentang data *mining* ataupun *text mining* bidang pendidikan sudah dilakukan oleh beberapa peneliti seperti metode C4.5 digunakan untuk penentuan beasiswa dengan akurasi sebesar 92% [1] dan akurasi 81% [2], penelitian [3] menggunakan metode C4.5 untuk prediksi kelulusan dengan akurasi 93%, dan penelitian [4] menggunakan metode naïve bayes untuk klasifikasi topik skripsi berdasarkan abstrak dengan akurasi 88%. Adapun penelitian yang berfokus pada permasalahan rekomendasi dosen pembimbing sudah dilakukan oleh peneliti menggunakan berbagai metode, diantaranya adalah metode *Cosine Similarity* oleh [5], [6] dengan masing-masing akurasi sebesar 75 % dan 98% , metode Naïve Bayes [7] mendapatkan presisi dan *recall* masing-masing sebesar 74% dan 100%, metode *Vector Space Model* (VSM) [8] dengan akurasi 93%, SVM dan *Weighted Product* [9] dengan presisi 93%, AHP [10], dan *Dice Coefficient Similarity* [11] dengan akurasi sebesar 88%.

Penelitian [5] menggunakan metode *cosine similarity* untuk pemilihan dosen pembimbing pada jurusan sistem informasi Politeknik Negeri Malang. Penelitian tersebut menghasilkan akurasi sebesar 75% terhadap data uji sebanyak 20 data. Penelitian [6] melakukan klasifikasi dokumen skripsi berdasarkan kategori menggunakan metode *cosine similarity*. Akurasi metode *cosine similarity* pada penelitian tersebut sebesar 98% dengan menggunakan data latih 50 data. Penelitian [7] menggunakan metode *cosine similarity* untuk pencocokkan dosen pembimbing tugas akhir berdasarkan judul, topik, abstrak, dan kompetensi dosen. Hasil penelitiannya memperoleh presisi dan *recall* masing-masing sebesar 74% dan 100% pada data uji yang digunakan sebanyak 10 dokumen.

Berdasarkan penelitian diatas, metode *cosine similarity* memperoleh rata-rata performa yang baik pada kasus pencocokan tingkat kemiripan antar dokumen skripsi. Persamaan penelitian ini dengan penelitian sebelumnya adalah menggunakan metode *cosine similarity* pada kasus rekomendasi dosen pembimbing berdasarkan judul dan kompetensi dosen, sedangkan perbedaannya adalah pada data yang digunakan. Dimana penelitian ini menggunakan data skripsi jurusan ilmu komputer Universitas Bumigora tahun 2020 yang berisi judul skripsi, dosen pembimbing, dan kompetensi dosen. Adapun tujuan penelitian ini adalah implementasi metode *cosine similarity* untuk rekomendasi dosen pembimbing skripsi pada prodi S1 Ilmu Komputer Universitas Bumigora berdasarkan judul dan kompetensi dosen yang diuji berdasarkan akurasi.

## 2 Tinjauan Literatur

Penelitian sebelumnya tentang data *mining* ataupun *text mining* bidang pendidikan sudah dilakukan oleh beberapa peneliti seperti metode C4.5 digunakan untuk penentuan beasiswa dengan akurasi sebesar 92% [1] dan akurasi 81% [2], penelitian [3] menggunakan metode C4.5 untuk prediksi kelulusan dengan akurasi 93%, dan penelitian [4] menggunakan metode naïve bayes untuk klasifikasi topik skripsi berdasarkan abstrak dengan akurasi 88%. Adapun penelitian yang berfokus pada permasalahan rekomendasi dosen pembimbing sudah dilakukan oleh peneliti menggunakan berbagai metode, diantaranya adalah metode *Cosine Similarity* oleh [5],[6] dengan masing-masing akurasi

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

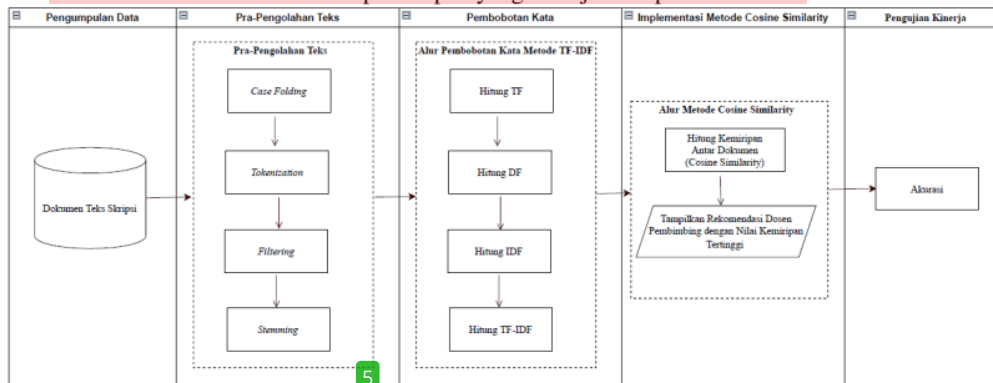
sebesar 75 % dan 98% , metode Naïve Bayes [7] mendapatkan presisi dan *recall* masing-masing sebesar 74% dan 100%, metode *Vector Space Model* (VSM) [8] dengan akurasi 93%, SVM dan *Weighted Product* [9] dengan presisi 93%, AHP [10], dan *Dice Coefficient Similarity* [11] dengan akurasi sebesar 88%.

Penelitian [5] menggunakan metode *cosine similarity* untuk pemilihan dosen pembimbing pada jurusan sistem informasi Politeknik Negeri Malang. Penelitian tersebut menghasilkan akurasi sebesar 75% terhadap data uji sebanyak 20 data. Penelitian [6] melakukan klasifikasi dokumen skripsi berdasarkan kategori menggunakan metode *cosine similarity*. Akurasi metode *cosine similarity* pada penelitian tersebut sebesar 98% dengan menggunakan data latih 50 data. Penelitian [7] menggunakan metode *cosine similarity* untuk pencocokan dosen pembimbing tugas akhir berdasarkan judul, topik, abstrak, dan kompetensi dosen. Hasil penelitiannya memperoleh presisi dan *recall* masing-masing sebesar 74% dan 100% pada data uji yang digunakan sebanyak 10 dokumen.

Berdasarkan penelitian diatas, metode *cosine similarity* memperoleh rata-rata performa yang baik pada kasus pencocokan tingkat kemiripan antar dokumen skripsi. Oleh karena itu, penelitian ini menggunakan metode *cosine similarity* pada kasus rekomendasi dosen pembimbing berdasarkan judul dan kompetensi dosen.

### 13 3 Metode Penelitian

Penelitian ini terdiri dari beberapa tahapan yang ditunjukkan pada Gambar 1.



5  
Gambar 1. Tahap Penelitian

Tahap pertama adalah pengumpulan data. Data yang digunakan penelitian ini berupa data skripsi jurusan ilmu komputer Universitas Bumigora tahun 2020 yang berisi judul skripsi, dosen pembimbing, dan kompetensi dosen. Data diperoleh dari ketua jurusan ilmu komputer Universitas Bumigora sebanyak 113 data.

Tahapan kedua adalah pra-pengolahan teks. Data yang sudah didapatkan pada fase sebelumnya masih bersifat mentah sehingga terdapat *noise* seperti terdapat simbol, angka, tanda baca, dan kata sambung yang dapat menurunkan performa metode yang digunakan. Oleh karena itu pra-pengolahan dibutuhkan untuk mendapatkan data berkualitas sehingga terhindar dari *noise* [13][4]. Pra-pengolahan teks penelitian ini terdiri dari *Case Folding*, *Tokenization*, *Filtering*, dan *Stemming*. *Case Folding* digunakan untuk merubah huruf teks dokumen skripsi menjadi huruf kecil. *Tokenization* digunakan untuk memisahkan teks pada dokumen skripsi menjadi token-token. *Filtering* digunakan untuk menghapus kata tidak penting seperti kata hubung, kata ganti, dll. Sedangkan *stemming* digunakan untuk mengubah kata yang berimbuhan menjadi kata dasar pada teks dokumen skripsi

Tahapan ketiga adalah proses pembobotan setiap *term* atau kata pada dokumen teks skripsi menggunakan metode TF-IDF. Metode TF-IDF menggabungkan dua konsep yaitu TF dan IDF. TF mencari nilai kemunculan term pada dokumen terkait, semakin banyak kemunculan term pada dokumen terkait semakin baik. Sedangkan konsep IDF berbanding terbalik dengan metode TF, semakin jarang kemunculan term pada seluruh dokumen semakin baik. Metode TF-IDF dihitung menggunakan persamaan (1) [14].

2  
<http://sistemasi.ftik.unisi.ac.id>

$$W_{ij} = tf_{ij} \times idf_j = tf_{ij} \times \log \left( \frac{N}{df_j} \right) \quad (1)$$

$W_{ij}$  merupakan bobot term ke- $j$  terhadap dokumen  $i$ .  $tf_{ij}$  merupakan jumlah kemunculan term  $j$  pada dokumen  $i$ .  $N$  adalah jumlah dokumen, dan  $df_j$  merupakan jumlah kemunculan term  $j$  pada seluruh dokumen

Tahapan keempat adalah implementasi metode *Cosine Similarity* untuk rekomendasi dosen pembimbing skripsi S1 ilmu komputer Universitas Bumigora dengan melihat kemiripan antar dokumen data latih dengan data uji pada data yang digunakan. Hasil dari metode *Cosine Similarity* menampilkan list calon dosen pembimbing berdasarkan nilai kemiripan tertinggi. Adapun perhitungan metode *Cosine Similarity* menggunakan persamaan (2) [15][16].

$$sim(x, y) = \frac{x \cdot y}{\|x\| \|y\|} = \frac{\sum_{i=0}^n (x_i * y_i)}{\sqrt{\sum_{i=0}^n x_i^2} * \sqrt{\sum_{i=0}^n y_i^2}} \quad (2)$$

Dimana  $x_i$  dan  $y_i$  merupakan sebuah vektor yang akan dicari tingkat kemiripannya, sedangkan  $\|x_i\|$  dan  $\|y_i\|$  merupakan panjang vektor.

Tahapan terakhir adalah pengujian akurasi metode *Cosine Similarity* menggunakan persamaan (3) [17][18].

$$Akurasi = \frac{\text{Jumlah Klasifikasi Benar}}{\text{Jumlah Data Testing}} \quad (3)$$

## 4 Hasil dan Pembahasan

### 4.1 Pengumpulan Data

Data yang digunakan penelitian ini berupa data skripsi jurusan ilmu komputer Universitas Bumigora tahun 2020 yang berisi judul skripsi, dosen pembimbing, dan kompetensi dosen yang ditunjukkan pada Tabel 1. Data diperoleh dari ketua jurusan ilmu komputer Universitas Bumigora sebanyak 113 data.

**Tabel 1. Data Dokumen Skripsi**

Nomor Dokumen	Judul	Dosen Pembimbing	Kompetensi
D1	Game Education Bahasa Jerman Dasar Berbasis Desktop untuk siswa Kelas X SMA Negeri 3 Mataram	Miftahul Madani, M.Kom	Aplikasi Multimedia dan Augmented Reality
D2	Perancangan Media Pemasaran Rumah Bersubsidi dan Rumah Komersial Berbasis 3 Dimensi dengan Tambahan Audio pada Augmented Reality PT. Varindo Lombok Inti Mataram	Ria Rismayati, M.Kom	Sistem Pakar dan Sistem Pendukung Keputusan
D3	Pengembangan Media Pembelajaran Interaktif IPA pada Materi Mikroorganisme Berbasis Mobile di SMKN 1 Kopang	Suriyati, M.Kom	Aplikasi Multimedia dan Sistem Pendukung Keputusan
D4	Sistem Pakar Diagnosa Penyakit Telinga Tenggorokan (THT) menggunakan Metode Forward Chaining dan Certainty Factor Berbasis Web	Hairani, M.Eng	Data Mining dan Sistem Pendukung Keputusan
D5	Sistem Pendukung Keputusan Penentuan Kelayakan Calon Transmigran dengan Menggunakan Metode Fuzzy Weighted Product	Heroe Santoso, M.Kom	Sistem Pendukung Keputusan dan Augmented Reality
...	.....		
D113	Aplikasi Monitoring Sumber Daya Proxmox VE Berbasis Web	I Putu Hariyadi, M.Kom	Aplikasi Jaringan Komputer dan Cloud Computing

#### 4.2 Pra-Pengolahan Teks

Data yang sudah didapatkan pada fase sebelumnya masih bersifat mentah sehingga terdapat *noise* seperti terdapat simbol, angka, tanda baca, dan kata sambung yang dapat menurunkan performa metode yang digunakan. Oleh karena itu pra-pengolahan dibutuhkan untuk mendapatkan data berkualitas sehingga terhindar dari *noise* [13] [4]. Pra-pengolahan teks penelitian ini terdiri dari *Case Folding*, *Tokenization*, *Filtering*, dan *Stemming*. *Case Folding* digunakan untuk merubah huruf teks dokumen skripsi menjadi huruf kecil. *Tokenization* digunakan untuk memisahkan teks pada dokumen skripsi menjadi token-token. *Filtering* digunakan untuk menghapus kata tidak penting seperti kata hubung, kata ganti, dll. Sedangkan *stemming* digunakan untuk mengubah kata yang berimbuhan menjadi kata dasar pada teks dokumen skripsi. Adapun hasil penggunaan teknik pra-pengolahan teks dokumen skripsi ditunjukkan pada Tabel 2.

**Tabel 2. Contoh Pra-Pengolahan Teks**

Pra-Pengolahan Teks	Hasil
Data Asli	<i>Game Education Bahasa Jerman Dasar Berbasis Desktop</i> pada SMA Negeri 3 Mataram
Case Folding	game education bahasa jerman dasar berbasis desktop pada sma negeri 3 mataram
Tokenization	"game", "education", "bahasa", "jerman", "dasar", "berbasis", "desktop", "pada", "sma", "negeri", "mataram"
Filtering	"game", "education", "bahasa", "jerman", "dasar", "berbasis", "desktop", "sma", "negeri", "mataram"
Stemming	"game", "education", "bahasa", "jerman", "dasar", "basis", "desktop", "sma", "negeri", "mataram"

#### 4.3 Pembobotan Teks

Adapun proses pembobotan menggunakan metode TF-IDF menggunakan 3 dokumen teks skripsi seperti berikut:

**D1** = Game Education Bahasa Jerman Dasar Berbasis Desktop untuk Mahasiswa Kelas X SMA Negeri 3 Mataram

**D2** = Perancangan Media Pemasaran Rumah Bersubsidi dan Rumah Komersial Berbasis 3 Dimensi dengan Tambahan Audio pada Augmented Reality PT. Varindo Lombok Inti Mataram

**D3** = Media Pembelajaran Organ Pencernaan Manusia Berbasis Multimedia Interaktif (Studi Kasus : Kelas V SDN INPRES NTORI Kec. Wawo, Kab. Bima)

Sebelum dilakukan pembobotan teks menggunakan metode TF-IDF terlebih dahulu dilakukan *pre-processing* untuk meningkatkan kualitas teks. Adapun hasil *pre-processing* teks pada ketiga dokumen diatas seperti berikut :

D1 = "game", "education", "bahasa", "jerman", "dasar", "basis", "desktop", "mahasiswa", "kelas", "sma", "negeri", "mataram"

D2 = "rancang", "media", "pasar", "rumah", "subsidi", "rumah", "komersial", "basis", "dimensi", "tambah", "audio", "augmented", "reality", "pt", "varindo", "lombok", "inti", "mataram"

D3 = "media", "ajar", "organ", "cerna", "manusia", "basis", "multimedia", "interaktif", "studi", "kelas", "sdn", "inpres", "ntori", "kec", "wawo", "kab", "bima"

Hasil *pre-processing* pada ketiga dokumen tersebut, selanjutnya dilakukan pembobotan kata menggunakan metode TF-IDF untuk mendapatkan bobot setiap kata atau *term* seperti ditunjukkan pada Tabel 3.

**Tabel 3. Hasil Pembobotan Kata Menggunakan TF-IDF**

Term	TF					W		
	D1	D2	D3	DF	IDF	D1	D2	D3
game	0.08333	0	0	1	0.47712	0.03976	0	0
education	0.08333	0	0	1	0.47712	0.03976	0	0
bahasa	0.08333	0	0	1	0.47712	0.03976	0	0
jerman	0.08333	0	0	1	0.47712	0.03976	0	0

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

dasar	0.08333	0	0	1	0.47712	0.03976	0	0
basis	0.08333	0.05882	0.05882	3	0	0	0	0
dekstop	0.08333	0	0	1	0.47712	0.03976	0	0
.....	.....	.....	.....	.....	.....	.....	.....	.....
bima	0	0	0.05882	1	0.47712	0.03976	0	0.03976
<b>Total</b>					<b>0.38718</b>	<b>0.41364</b>	<b>0.41364</b>	

#### 4.4 Metode Cosine Similarity

Setelah dilakukan proses pen-<sup>5</sup> bobatan kata, selanjutnya dilakukan mencari tingkat kemiripan antar dokumen menggunakan metode *Cosine Similarity* dengan menghitung kemiripan vektor data *testing* dengan data *training* pada dokumen teks skripsi. Misalnya contoh pencarian tingkat kemiripan dokumen D3 dengan D1 dan D2 berdasarkan persamaan (2).

$$sim(x, y) = \frac{\sum_{i=0}^n (x_i * y_i)}{\sqrt{\sum_{i=0}^n x_i^2} * \sqrt{\sum_{i=0}^n y_i^2}}$$

$$sim(D1, D3) = \frac{(0.03976 * 0) + (0.03976 * 0) + ..... + (0 * 0.02806)}{\sqrt{(0.03976)^2 + (0.03976)^2 + ..... + (0.01467)^2} * \sqrt{(0.02806)^2 + (0.02806)^2 + ..... + (0.02806)^2}}$$

$$= \frac{0.00015}{0.01283} = 0.01169$$

$$sim(D2, D3) = \frac{(0 * 0) + (0 * 0.01035) + ..... + (0.028065 * 0)}{\sqrt{(0.01035)^2 + (0.02806)^2 + ..... + (0.02806)^2} * \sqrt{(0.02806)^2 + (0.02806)^2 + ..... + (0.02806)^2}}$$

$$= \frac{0.00011}{0.01283} = 0.00978$$

Berdasarkan hasil perhitungan metode *Cosine Similarity*, dokumen D3 memiliki tingkat kemiripan dengan dokumen D1, karena nilai kemiripannya lebih tinggi dengan dokumen D2.

#### 4.5 Pengujian Hasil

Pada penelitian ini menggunakan data *testing* sebanyak 23 dari 113 *instance* untuk mengukur ketepatan (akurasi) metode *Cosine Similarity* dalam <sup>19</sup> omendasi dosen pembimbing skripsi Universitas Bumigora. Data testing dan hasil pengujiannya ditunjukkan pada Tabel 4.

**Tabel 4. Hasil Pengujian Data Testing**

Data Testing	Judul	Aktual	Prediksi	Keterangan
1	Implementasi <i>Augmented Reality</i> pada Brosur SMK Negeri 2 Mataram Berbasis Android <sup>12</sup>	Ria Rismayati, M.Kom	Ria Rismayati, M.Kom	<b>Valid</b>
2	Penerapan Objek 3 Dimensi pada Buku Kimia Berbasis <i>Augmented Reality</i> SMA <sup>130</sup> eri 1 Plampang	Melati Rosanensi M.Kom	Melati Rosanensi M.Kom	<b>Valid</b>
3	Penerapan <i>Augmented Reality</i> sebagai Media Bantu Promosi Perhiasan Berbasis Android	Heroe Santoso, M.Kom	Heroe Santoso, M.Kom	<b>Valid</b>
4	Penerapan <i>Augmented Reality</i> sebagai Media Pengenalan Tumbuhan Hidroponik	Heroe Santoso, M.Kom	Heroe Santoso, M.Kom	<b>Valid</b>
5	Aplikasi Multimedia Komik Digital Interaktif Cerita Dewi Anjani Berbasis Mobile	Melati Rosanensi M.Kom	Miftahul Madani, M.Kom	<b>Tidak Valid</b>
6	Media Pembelajaran Reprodksi	Melati	Melati	<b>Valid</b>

<sup>2</sup> <http://sistemasi.ftik.unisi.ac.id>

	15 Tubuh Manusia menggunakan Augmented Reality (Studi Kasus : SMAN 2 Selong)	Rosanensi M.Kom	Rosanensi M.Kom	
7	Penerapan Augmented Reality pada Brosur Rumah Sakit Umum Daerah Awet Muda Narmada Berbasis Mobile	Heroe Santoso, M.Kom	Heroe Santoso, M.Kom	<b>Valid</b>
8	Visualisasi Proses Asam Lambung menggunakan Augmented Reality Berbasis Mobile	Miftahul Madani, M.Kom	Miftahul Madani, M.Kom	<b>Valid</b>
9	Implementasi Google Map pada Flutter untuk Aplikasi Kurir Mobile	Pahrul Irfan, M.Kom	Pahrul Irfan, M.Kom	<b>Valid</b>
10	20 Sistem Pakar Diagnosa Penyakit Skizofrenia menggunakan Metode Dempster-Shafer	Hairani, M.Eng	Hairani, M.Eng	<b>Valid</b>
11	33 Sentimen Analisis Terhadap Kenaikan Tarif BPJS Kesehatan menggunakan Algoritma Support Vector Machine	Ni Luh Putu Merawati, M.Cs	Heroe Santoso, M.Kom	<b>Tidak Valid</b>
12	27 Sistem Pendukung Keputusan menggunakan Metode Analytic Hierarchy Process (AHP) dan Topsis untuk Seleksi Penerimaan Mahasiswa Berbasis Web	Hairani, M.Eng	Hairani, M.Eng	<b>Valid</b>
13	Sistem Pendukung Keputusan Penentuan Kelayakan Calon Transmigran dengan Menggunakan Metode Fuzzy Weighted Product	Heroe Santoso, M.Kom	Heroe Santoso, M.Kom	<b>Valid</b>
14	10 Sistem Informasi Penjualan Emas dan Mutiara Berbasis Android (Studi Kasus Sakura Mutiara Mbok)	Ismarmiaty, MMSI	Ismarmiaty, MMSI	<b>Valid</b>
15	Sistem Pakar Diagnosa Penyakit Telinga Hidung Tenggorokan (THT) menggunakan Metode Forward Chaining dan Certainty Factor Berbasis Web	Hairani, M.Eng	Hairani, M.Eng	<b>Valid</b>
16	Analisis Penerapan Unified Threat Management (UTM) untuk Keamanan Layanan Database pada Container Docker	I Putu Hariyadi, M.Kom	I Putu Hariyadi, M.Kom	<b>Valid</b>
17	Analisa dan Implementasi Performa Virtualisasi KVM pada Private Cloud untuk Layanan VOIP	Raisul Azhar, M.T	Raisul Azhar, M.T	<b>Valid</b>
18	Aplikasi Monitoring Sumber Daya Proxmox VE Berbasis Web	I Putu Hariyadi, M.Kom	I Putu Hariyadi, M.Kom	<b>Valid</b>
19	Push Notification untuk Memaksimalkan Promosi Produk menggunakan Google Cloud Messaging (Studi Kasus : CV Jembatan Baru)	Pahrul Irfan, M.Kom	Pahrul Irfan, M.Kom	<b>Valid</b>
20	Analisis Unjuk Kerja Perangkat dan Desain Jaringan Komputer	Raisul Azhar, M.T	Raisul Azhar, M.T	<b>Valid</b>

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



	menggunakan <i>Top Down Network</i> Desain (STAHN Gde Pudja taram)			
21	Implementasi <i>Electronic Nose</i> Berbasis <i>Internet Of Things</i> menggunakan Website sebagai Monitoring Jenis Gas	I Putu Hariyadi, M.Kom	I Putu Hariyadi, M.Kom	<b>Valid</b>
22	Sistem Monitoring Suhu Air untuk Budidaya Ikan Lele Berbasis <i>Internet Of Things</i> (IOT)	Lalu Zazuli Azhar, M.Kom	Lalu Zazuli Azhar, M.Kom	<b>Valid</b>
23	Sistem Monitoring Tumpahan Minyak Berbasis <i>Internet Of Things</i>	Andi Sofyan Anas, M.Kom	Andi Sofyan Anas, M.Kom	<b>Valid</b>

Berdasarkan hasil pengujian data *testing* pada Tabel 1 didapatkan akurasi metode *Cosine Similarity* berdasarkan persamaan (1).

$$\text{Akurasi} = \frac{21}{23} * 100 = 0.9130 * 100 = 91.3\%$$

Implementasi metode *Cosine Similarity* untuk rekomendasi dosen pembimbing skripsi memiliki akurasi sebesar **91,3%** berdasarkan hasil pengujian pada data *testing*. Akurasi metode *Cosine Similarity* dipengaruhi tingkat kemiripan pada judul skripsi dan kompetensi dosen. Selain itu, penggunaan *term* unik sangat berpengaruh pada perhitungan metode *Cosine Similarity* [6].

## 5 Kesimpulan

Penelitian ini bertujuan implementasi metode *cosine similarity* untuk rekomendasi dosen pembimbing skripsi S1 ilmu komputer Universitas Bumigora. Data yang digunakan sebanyak 113 dokumen skripsi yang terbagi sebagai data *training* sebanyak 90 dokumen dan data *testing* sebanyak 23 dokumen. Berdasarkan hasil pengujian pada data *testing*, metode *cosine similarity* dapat merekomendasi dosen pembimbing secara benar sebanyak 21 dari 23 dokumen skripsi dengan akurasi sebesar **91.3%**. Dengan demikian, metode *cosine similarity* dapat diterapkan pada kasus pemilihan dosen pembimbing S1 ilmu komputer Universitas Bumigora karena memiliki akurasi yang **sangat baik**. Penelitian selanjutnya dapat mengkombinasikan metode *cosine similarity* dengan k-nn untuk rekomendasi dosen pembimbing.

## Referensi

- [1] A. S. Suweleh, D. Susilowati, and Hairani Hairani, "Aplikasi Penentuan Penerima Beasiswa menggunakan Algoritma C4.5," *Jurnal BITE*, vol. 2, no. 1, pp. 12–21, 2020.
- [2] H. Hairani, A. S. Suweleh, and D. Susilowaty, "Penanganan Ketidak Seimbangan Kelas menggunakan Pendekatan Level Data," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 20, no. 1, pp. 109–116, 2020, doi: 10.30812/matrik.v20i1.846.
- [3] D. Kumiawan, A. Anggrawan, and H. Hairani, "Graduation Prediction System on Students using C4.5 Algorithm," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 19, no. 2, pp. 358–365, 2020, doi: 10.30812/matrik.v19i2.685.
- [4] H. Hairani, A. Anggrawan, A. I. Wathan, K. A. Latif, K. Marzuki, and M. Zulfikri, "The Abstract of Thesis Classifier by using Naive Bayes Method," in *Proceedings - 2021 International Conference on Software Engineering and Computer Systems and 4th International Conference on Computational Science and Information Management, ICSECS-ICOCSIM 2021*, 2021, no. August, pp. 312–315, doi: 10.1109/ICSECS52883.2021.00063.
- [5] R. Rismanto, A. Rachmad Syulistyo, and B. P. Citra Augusta, "Research Supervisor Recommendation System based on Topic Conformity," *International Journal of Modern Education and Computer Science*, vol. 12, no. 1, pp. 26–34, Feb. 2020, doi: 10.5815/ijmecs.2020.01.04.

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

- [6] R. T. Wahyuni, D. Prastiyanto, and E. Suprpto, "Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF pada Sistem Klasifikasi Dokumen Skripsi," *Jurnal Teknik Elektro*, vol. 9, no. 1, pp. 18–23, 2017, [Online]. Available: <https://journal.unnes.ac.id/nju/index.php/jte/article/view/10955/6659>.
- [7] M. Asfi, N. Fitriani, D. Pembimbing, N. Skripsi, and C. Bayes, "Implementasi Algoritma Naive Bayes Classifier sebagai Sistem Rekomendasi Pembimbing Skripsi," *InfoTekJar : Jurnal Nasional Informatika dan Teknologi Jaringan*, vol. 5, no. 1, pp. 44–50, 2020, [Online]. Available: <https://jurnal.uisu.ac.id/index.php/infotekjar/article/view/2536>.
- [8] R. R. A. Siregar, F. A. Sinaga, and R. Arianto, "Aplikasi Penentuan Dosen Penguji Skripsi menggunakan Metode TF-IDF dan Vector Space Model," *Computatio : Journal of Computer Science and Information Systems*, vol. 1, no. 2, p. 171, 2017, doi: 10.24912/computatio.v1i2.1014.
- [9] Y. R. Pradana, A. A. Supianto, and Y. T. Mursityo, "Prediksi Bidang Penelitian dan Rekomendasi Dosen Pembimbing Skripsi Berdasarkan Konten Latar Belakang pada Naskah Proposal Menggunakan Metode Multi-Class Support Vector Machine dan Weighted Product," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 2, p. 403, 2021, doi: 10.25126/jtiik.2021824511.
- [10] J. S. Simanungkalit, H. T. Sihotang, S. P. Nusantara, J. Iskandar, M. No, and N. Sumatra, "Decision Support System for Selection of Thesis Advisors Supervisors Thesis Advisors according to the Field of Science Using the AHP Method," vol. 3, no. 4, 2020.
- [11] F. T. Anggraeni, I. Y. Purbasari, and E. F. Wulandari, "Undergraduate Thesis Supervisor Recommendation based on Text Similarity," in *The 5th International Conference on Information Technology and Business (ICITB 2019)*, 2019, pp. 248–256.
- [12] L. Yasni, I. M. I. Subroto, and S. F. C. Haviana, "Implementasi Cosine Similarity Matching dalam Penentuan Dosen Pembimbing Tugas Akhir," *Transmisi: Jurnal Ilmiah Teknik Elektro*, vol. 20, no. 1, pp. 22–28, Jan. 2018, doi: 10.14710/transmisi.20.1.22-28.
- [13] A. Rahman and A. Doewes, "Online News Classification using Multinomial Naive Bayes," *ITSMART: Jurnal Ilmiah Teknologi dan Informasi*, vol. 6, no. 1, pp. 32–38, 2017.
- [14] A. Prabhat and V. Khullar, "Sentiment Classification on Big Data using Naive Bayes and Logistic Regression," in *2017 International Conference on Computer Communication and Informatics (ICCCI)*, 2017, pp. 1–5, doi: 10.1109/ICCCI.2017.8117734.
- [15] J. Han, M. Kamber, and J. Pei, "Data Mining Concepts and Techniques," Third Edit., Elsevier, 2012.
- [16] E. L. Amalia, A. J. Jumadi, I. A. Mashudi, and D. W. Wibowo, "Analisis Metode Cosine Similarity pada Aplikasi Ujian Online Otomatis (Studi Kasus JTI POLINEMA)," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 2, pp. 343–348, 2021, doi: 10.25126/jtiik.2021824356.
- [17] Z. Efendi and M. Mustakim, "Text Mining Classification sebagai Rekomendasi Dosen Pembimbing Tugas Akhir Program Studi Sistem Informasi," in *Seminar Nasional Teknologi Informasi Komunikasi dan Industri*, 2017, pp. 235–242, [Online]. Available: <http://ejournal.uin-suska.ac.id/index.php/SNTIKI/article/view/3273>.
- [18] H. Hairani, K. Kurniawan, K. A. Latif, and M. Innuddin, "Metode Dempster-Shafer untuk Diagnosis Dini Jenis Penyakit Gangguan Jiwa Skizofrenia Berbasis Sistem Pakar," *Sistemasi : Jurnal Sistem Informasi*, vol. 10, no. 2, pp. 280–289, 2021, doi: 10.32520/stmsi.v10i2.1195.

# Recommendations of Thesis Supervisor using the Cosine Similarity Method

---

## ORIGINALITY REPORT

---

12%

SIMILARITY INDEX

---

### PRIMARY SOURCES

1	<a href="http://publikasiilmiah.unwahas.ac.id">publikasiilmiah.unwahas.ac.id</a> Internet	32 words — 1%
2	<a href="http://sistemasi.ftik.unisi.ac.id">sistemasi.ftik.unisi.ac.id</a> Internet	30 words — 1%
3	<a href="http://repository.itelkom-pwt.ac.id">repository.itelkom-pwt.ac.id</a> Internet	28 words — 1%
4	<a href="http://ejournal.undip.ac.id">ejournal.undip.ac.id</a> Internet	23 words — 1%
5	<a href="http://id.123dok.com">id.123dok.com</a> Internet	20 words — 1%
6	<a href="http://journal.yrpiiku.com">journal.yrpiiku.com</a> Internet	19 words — 1%
7	<a href="http://eprints.unmer.ac.id">eprints.unmer.ac.id</a> Internet	17 words — < 1%
8	Rahmiati Rahmiati, Dedy Irfan, Agustin Agustin, Siska Hedyati. "APLIKASI PENGUKUR TINGKAT SENTIMEN PELANGGAN BERDASARKAN KOMPLAIN PELANGGAN PLN MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR", INOVTEK Polbeng - Seri Informatika, 2020 Crossref	14 words — < 1%

- 
- 9 [ejurnal.itenas.ac.id](http://ejurnal.itenas.ac.id)  
Internet 14 words — < 1%
- 
- 10 [eprints.unram.ac.id](http://eprints.unram.ac.id)  
Internet 14 words — < 1%
- 
- 11 [rifdt.instifdt.bg.ac.rs](http://rifdt.instifdt.bg.ac.rs)  
Internet 14 words — < 1%
- 
- 12 [www.hipolisis.com](http://www.hipolisis.com)  
Internet 14 words — < 1%
- 
- 13 [Hairani Hairani, Muhammad Ridho Hansyah, Lalu Zazuli Azhar Mardedi. "Integrasi Metode Naive Bayes dengan K-Means dan K-Means-Smote untuk Klasifikasi Jurusan SMAN 3 Mataram", Jurnal Sistem dan Informatika \(JSI\), 2020](#)  
Crossref 13 words — < 1%
- 
- 14 [www.onesearch.id](http://www.onesearch.id)  
Internet 13 words — < 1%
- 
- 15 [Subhan Hadi, Siti Soraya, Melati Rosanensi, Suriyati Suriyati, Baiq Chandra Herawati. "Media Pembelajaran Reproduksi Tubuh Manusia Menggunakan Augmented Reality \(Studi Kasus: SMAN 2 Selong\)", Jurnal Bumigora Information Technology \(BITe\), 2022](#)  
Crossref 10 words — < 1%
- 
- 16 [library.instiki.ac.id](http://library.instiki.ac.id)  
Internet 10 words — < 1%
- 
- 17 [Irwin Supriadi. "Implementasi Algoritma SAW Pada Aplikasi Penentuan Pembimbing Tugas Akhir Mahasiswa \(Studi Kasus: Prodi Informatika Universitas Langlangbuana\)", INTERNAL \(Information System Journal\), 2021](#)  
Crossref 9 words — < 1%

18	<a href="http://docplayer.info">docplayer.info</a> Internet	9 words — < 1%
19	<a href="http://j-ptiik.ub.ac.id">j-ptiik.ub.ac.id</a> Internet	9 words — < 1%
20	<a href="http://jurnal.pcr.ac.id">jurnal.pcr.ac.id</a> Internet	9 words — < 1%
21	<a href="http://openlibrarypublications.telkomuniversity.ac.id">openlibrarypublications.telkomuniversity.ac.id</a> Internet	9 words — < 1%
22	<a href="http://www.coursehero.com">www.coursehero.com</a> Internet	9 words — < 1%
23	Sri Winiarti, Desy Widayanti, Ulaya Ahdiani, Taufiq Ismail. "Klasifikasi Jenis Buku Berdasarkan Cover dan Judul Buku Menggunakan Metode Support Vector Machine dan Cosine Similarity", Sainteks, 2022 Crossref	8 words — < 1%
24	<a href="http://cafeiart.vilagrassa.net">cafeiart.vilagrassa.net</a> Internet	8 words — < 1%
25	<a href="http://ejournals.umn.ac.id">ejournals.umn.ac.id</a> Internet	8 words — < 1%
26	<a href="http://pastel.archives-ouvertes.fr">pastel.archives-ouvertes.fr</a> Internet	8 words — < 1%
27	<a href="http://repository.dinamika.ac.id">repository.dinamika.ac.id</a> Internet	8 words — < 1%
28	<a href="http://unsri.portalgaruda.org">unsri.portalgaruda.org</a> Internet	8 words — < 1%

29	<a href="http://www.journaltoocs.ac.uk">www.journaltoocs.ac.uk</a> Internet	8 words — < 1%
30	<a href="http://bimbinganprogramskripsi.wordpress.com">bimbinganprogramskripsi.wordpress.com</a> Internet	7 words — < 1%
31	<a href="http://jtiik.ub.ac.id">jtiik.ub.ac.id</a> Internet	7 words — < 1%
32	Putri Yuni Ristanti, Aji Prasetya Wibawa, Utomo Pujianto. "Cosine Similarity for Title and Abstract of Economic Journal Classification", 2019 5th International Conference on Science in Information Technology (ICSITech), 2019 Crossref	6 words — < 1%
33	<a href="http://journal.universitasbumigora.ac.id">journal.universitasbumigora.ac.id</a> Internet	6 words — < 1%

EXCLUDE QUOTES ON  
EXCLUDE BIBLIOGRAPHY ON

EXCLUDE SOURCES OFF  
EXCLUDE MATCHES OFF