

**PENCARIAN REFERENSI PADA KALIMAT JURNAL  
DENGAN KESAMAAN MAKNA MENGGUNAKAN METODE  
COSINE SIMILARITY**

**SKRIPSI**



**Oleh :  
ROFIATUL ADAWIYAH  
NPM : 1955201000996**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNOLOGI INFORMASI  
UNIVERSITAS MERDEKA PASURUAN**

**2023**

**PENCARIAN REFERENSI PADA KALIMAT JURNAL  
DENGAN KESAMAAN MAKNA MENGGUNAKAN METODE  
COSINE SIMILARITY**

**SKRIPSI**

Diajukan sebagai salah satu persyaratan dalam menyelesaikan  
program Sarjana Komputer (S.Kom)



**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNOLOGI INFORMASI  
UNIVERSITAS MERDEKA PASURUAN**

**2023**

## PERSETUJUAN PEMBIMBING SKRIPSI

### PENCARIAN REFERENSI PADA KALIMAT JURNAL DENGAN KESAMAAN MAKNA MENGGUNAKAN METODE COSINE SIMILARITY

Nama : Rofiatul Adawiyah

NPM : 1955201000996

Program Studi : Informatika

Pasuruan, 25 Juli 2023

Telah diperiksa dan disetujui oleh:

Pembimbing I

Pembimbing II

(Mohammad Zoqi Sarwani, S.Pd., M.Kom.)

(Dian Ahkam Sanj, S.Kom., M.Kom.)

NIDN. 0716049101

NIDN. 0724079202

## PERSETUJUAN DAN PENGESAHAN SKRIPSI

Skripsi oleh Rofiatul Adawiyah ini  
telah dipertahankan di depan dewan penguji  
pada tanggal 25 Juli 2023

Dewan Penguji

Ketua Penguji



(Mohammad Firman Arif S.Kom., M. Kom.)  
NIDN. 0704029201

Anggota Penguji I



(Rudi Hariyanto, S.Kom, M.T.)  
NIDN. 0718018401

Anggota Penguji II



(Anang Aris Widodo, S.Kom, M.T.)  
NIDN. 0702038102

Mengesahkan,

Dekan Fakultas Teknologi Informasi

Mengetahui,

Ketua Program Studi Informatika



(M. Misdrum, S.Kom., M.Kom.)

NIDN. 0717046704



(Dian Ahkam Sani, S.Kom., M.Kom.)

NIDN. 0724079202

## PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan di bawah ini:

Nama : Rofiatul Adawiyah

NPM : 1955201000996

Program Studi : Informatika

Fakultas : Teknologi Informasi

Menyatakan dengan sesungguhnya bahwa **skripsi** yang saya tulis ini benar-benar tulisan saya, dan bukan merupakan plagiasi baik sebagian atau seluruhnya.

Apabila di kemudian hari terbukti atau dapat dibuktikan **bahwa skripsi** ini hasil plagiasi, baik sebagian atau seluruhnya, maka saya bersedia menerima sanksi atas perbuatan tersebut sesuai dengan ketentuan yang berlaku.

Pasuruan, 25 Juli 2023



menyatakan dan membuat pernyataan

Rofiatul Adawiyah

1955201000996

## ABSTRAK

Adawiyah, Rofiatul. 2023. *Pencarian Referensi pada Kalimat Jurnal Dengan Kesamaan Makna Menggunakan Metode Cosine Similarity*, Program Studi Informatika, Fakultas Teknologi Informasi Universitas Merdeka Pasuruan.

Pembimbing : (I) Mohammad Zoqi Sarwani, S.Pd., M.Kom. (II) Dian Ahkam Sani, S.Kom., M.Kom.,

### ***Kata kunci*— Pencarian, Referensi, Jurnal, Fasttext, Cosine Similarity**

Pencarian referensi menjadi hal penting bagi peneliti dalam memperkuat argumennya. Dalam pencarian referensi akan berpacu pada kesamaan kata atau makna pada hal yang akan dicari. Maka dibutuhkan Fasttext yang akan mengubah hasil token dari preprocessing data menjadi vektor. *Fasttext* adalah *library* yang diciptakan oleh Facebook untuk membantu mengubah kata menjadi vektor. Setelah mendapatkan hasil vektor dari kalimat pencarian dan kalimat referensi, kemudian dilanjutkan ke tahapan metode *cosine similarity*. Metode *cosine similarity* digunakan untuk menentukan seberapa dekat vektor satu sama lain. Dengan dataset yang digunakan sebanyak 414 data, maka diperoleh akurasi yang baik sehingga bisa dikatakan sistem ini dapat menemukan pencarian referensi dengan kesamaan makna dengan baik.

## ABSTRACT

Adawiyah, Rofiatul. 2023. *Search for References in Journal Sentences with Similarity of Meaning Using the Cosine Similarity Method*, Informatics Study Program, Faculty of Information Technology, Merdeka University, Pasuruan.

Advisor : (I) Mohammad Zoqi Sarwani, S.Pd., M.Kom. (II) Dian Ahkam Sani, S.Kom., M.Kom.,

Keywords— Search, Reference, Journal, Fasttext, Cosine Similarity

The search for references is important for researchers in strengthening their arguments. In the search for references will race on the similarity of words or meanings in the things to be searched. Then it takes Fasttext that will convert the token result from preprocessing data into vectors. Fasttext is a library created by Facebook to help turn words into vectors. After getting vector results from search sentences and reference calimics, then proceed to the cosine similarity method stage. The cosine similarity method is used to determine how close the vectors are to each other. With the dataset used as many as 414 data, good accuracy was obtained so that it can be said that this system can find reference searches with similar meanings well.

## KATA PENGANTAR

Puji syukur atas kehadiran Allah SWT yang telah memberikan rahmat dan karunia-Nya sehingga penulisan Skripsi dengan judul “**PENCARIAN REFERENSI PADA KALIMAT JURNAL DENGAN KESAMAAN MAKNA MENGGUNAKAN METODE *COSINE SIMILARITY***” dapat terselesaikan dengan baik. Skripsi ini merupakan tugas akhir yang diajukan untuk memenuhi salah satu persyaratan ujian guna memperoleh gelar Sarjana Komputer (S.Kom) pada Program Studi Informatika Fakultas Teknologi Informasi di Universitas Merdeka Pasuruan.

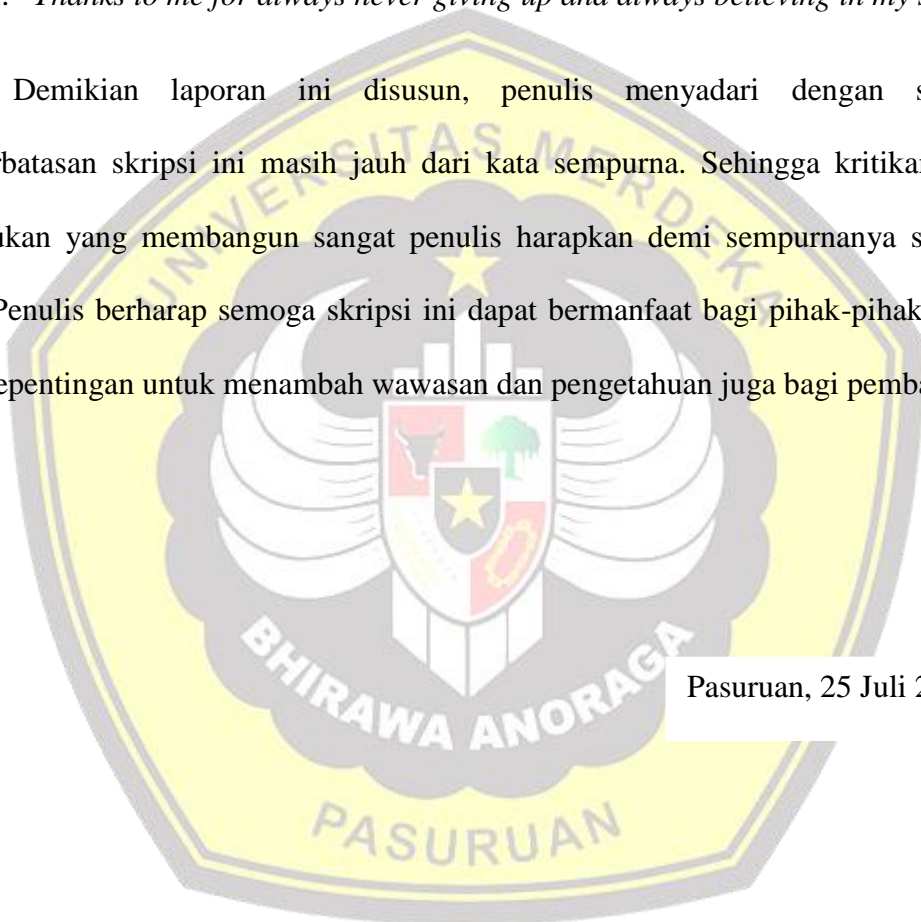
Dalam penyusunan skripsi ini banyak pihak yang telah membantu, oleh karena itu tidak lupa penulis mengucapkan terima kasih banyak kepada :

1. Dr. Ir. Sulistyawati, M.P., selaku Rektor Universitas Merdeka Pasuruan.
2. Muhammad Misdram, S.Kom., M.Kom., selaku Dekan Fakultas Teknologi Informasi.
3. Dian Ahkam Sani, S.Kom., M.Kom., selaku Ketua Program Studi Informatika.
4. Mohammad Zoqi Sarwani, S.Pd., M.Kom., selaku Dosen Pembimbing I dan Dian Ahkam Sani, S.Kom., M.Kom., selaku Dosen Pembimbing II serta segenap Dosen Universitas Merdeka Pasuruan.
5. Keluarga yang selalu mendukung, menanyakan progress skripsi saya dan selalu mendo'akan agar skripsi saya cepat selesai.



6. Sahabat-sahabat yang selalu memberikan dukungan dan motivasi untuk menyelesaikan skripsi ini.
7. Teman-teman FTI Angkatan 2019 yang selalu memberi semangat kepada saya untuk menyelesaikan skripsi ini.
8. Seluruh pihak yang membantu hingga tersusunnya skripsi ini.
9. *Thanks to me for always never giving up and always believing in my self.*

Demikian laporan ini disusun, penulis menyadari dengan segala keterbatasan skripsi ini masih jauh dari kata sempurna. Sehingga kritikan dan masukan yang membangun sangat penulis harapkan demi sempurnanya skripsi ini. Penulis berharap semoga skripsi ini dapat bermanfaat bagi pihak-pihak yang berkepentingan untuk menambah wawasan dan pengetahuan juga bagi pembaca.



Pasuruan, 25 Juli 2023

## DAFTAR ISI

<b>ABSTRAK .....</b>	<b>i</b>
<b>KATA PENGANTAR.....</b>	<b>iii</b>
<b>DAFTAR ISI.....</b>	<b>v</b>
<b>DAFTAR GAMBAR.....</b>	<b>vii</b>
<b>DAFTAR TABEL .....</b>	<b>viii</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>ix</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 LATAR BELAKANG .....	1
1.2 PERUMUSAN MASALAH .....	4
1.3 BATASAN MASALAH .....	4
1.4 TUJUAN PENELITIAN.....	4
1.5 MANFAAT PENELITIAN.....	5
1.6 SISTEMATIKA PENELITIAN.....	5
<b>BAB II TINJAUAN PUSTAKA.....</b>	<b>7</b>
2.1 PENELITIAN TERDAHULU .....	7
2.2 LANDASAN TEORI.....	12
2.2.1 Referensi .....	12
2.2.2 Jurnal.....	12
2.2.3 Kesamaan Makna.....	13
2.2.4 Data .....	13
2.2.5 Flowchart .....	14
2.2.6 Preprocessing Data.....	16
2.2.7 Fasttext.....	16
2.2.8 Cosine Similarity .....	17
2.2.9 Microsoft Excel.....	17
2.2.10 Google Colaboratory.....	18
2.2.11 Pandas .....	18
2.2.12 Numpy.....	19

<b>BAB III ANALISA MASALAH DAN PERANCARAN.....</b>	<b>20</b>
3.1 TAHAP PENELITIAN .....	20
3.2 STUDI LITERATUR.....	21
3.3 ANALISIS KEBUTUHAN SISTEM .....	22
3.3.1 Hardware.....	22
3.3.2 Software .....	23
3.4 DESAIN SISTEM.....	23
3.4.1 Flowchart Sistem .....	23
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>35</b>
4.1 IMPLEMENTASI SISTEM.....	35
4.2 HASIL AKURASI.....	39
<b>BAB V PENUTUP.....</b>	<b>44</b>
5.1 KESIMPULAN.....	44
5.2 SARAN .....	45
<b>DAFTAR PUSTAKA.....</b>	<b>46</b>
<b>LAMPIRAN.....</b>	<b>48</b>



## DAFTAR GAMBAR

Gambar 3.1 Tahapan Penelitian .....	20
Gambar 3.2 Flowchart Sistem.....	24
Gambar 3.3 Cosine Similarity.....	31
Gambar 4.1 Tampilan Library.....	35
Gambar 4.2 Load Model Fasstext .....	36
Gambar 4.3 Tampilan Dataset.....	37
Gambar 4.4 Input Pencarian.....	38
Gambar 4.5 Case Folding dan Tokenizing Dataset.....	38
Gambar 4.6 Membuat Vektor .....	38
Gambar 4.7 Menghitung Cosine Similarity .....	39
Gambar 4.8 Menghitung Akurasi.....	40
Gambar 4.9 Membuat Tabel DataFrame.....	40
Gambar 4.10 Tabel Menampilkan Hasil Pencarian .....	41
Gambar 4.11 Menghitung Akurasi.....	42
Gambar 4.12 Contoh Hasil Pencarian 1 .....	42
Gambar 4.13 Contoh Hasil Pencarian 2.....	43
Gambar 4.14 Contoh Hasil Pencarian 3.....	43

## DAFTAR TABEL

Gambar 2.1 Flowchart.....	14
Gambar 3.1 Hardware .....	22
Gambar 3.2 Case Folding.....	25
Gambar 3.3 Tokenizing.....	26
Gambar 3.4 Tabel Vektor Kata 1 .....	31
Gambar 3.5 Tabel Vektor Kata 2 .....	33



## DAFTAR LAMPIRAN

Lampiran 1 : Contoh Dataset .....	48
Lampiran 2 : Source Code Program.....	60
Lampiran 3 : Daftar Riwayat Hidup .....	64
Lampiran 4 : Surat Keterangan Cek Plagiasi .....	65

