

**RANCANG BANGUN CHATBOT AKADEMIK MENGGUNAKAN
ARSITEKTUR *RETRIEVAL-AUGMENTED GENERATION* (RAG)
(STUDI KASUS: PRODI TIF UNIPMA)**

SKRIPSI



OLEH:

ERFIA NADIA SAFARI
NIM. 2205101081

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS PGRI MADIUN

Juli 2026

**RANCANG BANGUN CHATBOT AKADEMIK MENGGUNAKAN
ARSITEKTUR *RETRIEVAL-AUGMENTED GENERATION* (RAG)
(STUDI KASUS: PRODI TIF UNIPMA)**

SKRIPSI

Diajukan kepada Universitas PGRI Madiun untuk Memenuhi Salah Satu
Persyaratan dalam Menyelesaikan Program Sarjana Strata 1
Teknik Informatika

OLEH:

ERFIA NADIA SAFARI
NIM. 2205101081

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS PGRI MADIUN

Juli 2026

LEMBAR PERSETUJUAN DOSEN PEMBIMBING

Skripsi oleh Erfia Nadia Safari diperiksa dan disetujui untuk di uji.

Madiun, 29 Juni 2026
Pembimbing I,



Saifulloh, S.Kom., M.Kom.
NIDN. 0717048906

Madiun, 3 Juli 2026
Pembimbing II,



Pratiwi Susanti, S.Kom., M.MT.
NIDN. 0711089301

LEMBAR PENGESAHAN DOSEN PENGUJI

Skripsi oleh Erfia Nadia Safari ini telah dipertahankan di depan penguji pada hari
Kamis, 9 Juli 2026.

Tim Penguji

Saifulloh, S.Kom., M.Kom.
NIDN. 0711048906

Penguji I

Pratiwi Susanti, S.Kom., M.MT.
NIDN. 0711089301

Penguji II

Yessi Yunitasari, S.Kom., M.Cs.
NIDN. 0703069204

Penguji III

Mengetahui
Dekan FT

Nasrul Rofiah Hidayati, S.T., M.Pd.
NIDN. 0706108202

Menyetujui
Kaprosdi Teknik Informatika

Yoga Prisma Yuda, S.Kom., M.Kom.
NIDN. 0722089002

PERNYATAAN KEASLIAN KARYA ILMIAH SKRIPSI

Yang bertanda tangan di bawah ini :

Nama : Erfia Nadia Safari

NIM : 2205101081

Program Studi : Teknik Informatika

Fakultas : Fakultas Teknik

Menyatakan dengan sebenarnya, bahwa skripsi yang saya tulis dengan judul “Rancang Bangun Chatbot Akademik Menggunakan Arsitektur *Retrieval-Augmented Generation* (RAG) (Studi Kasus: Prodi TIF UNIPMA)” ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilalihan tulisan atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri. Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini plagiat, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Madiun, 9 Juli 2026
Yang membuat pernyataan,



Erfia Nadia Safari
NIM. 2205101081

HALAMAN PERSEMBAHAN

*Skripsi ini penulis persembahkan kepada
orang-orang yang bertanya
“Kapan lulus?”*

MOTTO

“Dan kehidupan dunia tidak lain hanyalah kesenangan yang palsu.”

(QS. Al-Hadid ayat 20)

“Beberapa dari kita dilahirkan dengan potensi sebesar paus, sayangnya mereka
hidup dalam aquarium”

(Emran Bite)

“Penjara tidak selalu berbentuk sempit. Kadang ia berbentuk kemiskinan,
propaganda dan kebodohan.”

(Tan Malaka)

*“The standards you made are more strict for yourself. The thick tree rings in your
life. It's part of you, it's you. Now let's forgive ourselves. Our lives are long, trust
yourself when in a maze. When winter passes, spring always comes.”*

(Answer : Love Myself oleh BTS)

*“A feeling of inferiority, self-loathing compared to others. Aim your gun at these
things starting today.”*

(D-day oleh AgustD)

*“Find euphoria in dystopia. Let our souls become the magnets. No running from
the sadness”*

(Dystopia oleh One Ok Rock)

KATA PENGANTAR

Puji syukur ke hadirat Allah SWT atas segala rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan penulisan Skripsi ini dengan judul “Rancang Bangun Chatbot Akademik Menggunakan Arsitektur *Retrieval-Augmented Generation* (RAG) (Studi Kasus: Prodi TIF UNIPMA)” sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Teknik Informatika Fakultas Teknik Universitas PGRI Madiun. Penulisan Skripsi ini tidak mungkin terwujud tanpa bimbingan, dukungan, dan dorongan dari berbagai pihak yang dengan ikhlas memberikan kontribusi dalam penyelesaian Skripsi ini. Oleh karena itu, penulis ingin menyampaikan penghargaan dan rasa terima kasih kepada:

1. Ibu Nasrul Rofiah H, S.T., M.Pd. selaku Dekan Fakultas Teknik Universitas PGRI Madiun.
2. Bapak Yoga Prisma Yuda, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas PGRI Madiun.
3. Bapak Saifulloh, S.Kom., M.Kom. selaku Dosen Pembimbing I yang telah memberikan pengarahan, bimbingan dan dukungan kepada mahasiswa dalam melakukan penelitian dan penulisan Skripsi
4. Ibu Pratiwi Susanti, S.Kom., M.MT. selaku Dosen Pembimbing II yang telah memberikan pengarahan, bimbingan dan dukungan kepada mahasiswa dalam melakukan penelitian dan penulisan Skripsi
5. Ibu Yessi Yunitasari, S.Kom., M.Cs. selaku Dosen Penguji Seminar Hasil Skripsi Program Studi Teknik Informatika Universitas PGRI Madiun.

6. Kepada seluruh staf dan karyawan Universitas PGRI Madiun yang telah memberikan bantuan kepada peneliti.
7. Seluruh dosen pengampu Mata Kuliah yang telah membimbing, membagikan ilmu, dan membentuk pola pikir saya yang lebih baik sejak semester 1 hingga semester 8.
8. Kedua orang tua serta kakak penulis yang tiada hentinya mengalirkan dukungan penuh, doa, dan kasih sayang serta finansial di setiap langkah perjalanan ini. Semoga sehat selalu dengan anggota keluarga yang selalu utuh bersama hingga akhir hayat.
9. Keci, kucing semata wayang saya yang dengan tingkah lucunya hadir sebagai *support system* mental. Terima kasih sudah hadir dalam hidup penulis, walaupun penulis tahu bahwa umur kucing terlampaui pendek.
10. Evry, sahabat yang selalu setia mendampingi dalam suka maupun duka, serta meyakinkan saya untuk berani melangkah mengambil keputusan terbaik demi masa depan. Terima kasih sudah memberikan dukungan moral dan mental yang sangat berarti hingga detik ini. Semoga takdir baik selalu beriringan dengan lajunya waktu.
11. BTS dengan 7 member yang terdiri dari Kim Namjoon, Kim Seokjin, Min Yoongi, Jung Hoseok, Park Jimin, Kim Taehyung dan Jeon Jungkook yang merupakan boyband favorit penulis yang telah menginspirasi penulis melalui karya musiknya yang hadir dalam hidup penulis selama 9 tahun lamanya untuk selalu berkembang menjadi pribadi yang lebih memaknai hidup.

12. One Ok Rock, Bring Me The Horizon, Green Day, My Chemical Romance dan Paramore, grup musik yang lagu-lagunya selalu menjadi pelampiasan emosi sekaligus pembakar semangat selama masa perkuliahan.
13. Teman-teman seperjuangan, khususnya Kelas D Teknik Informatika Angkatan 2022 yang telah kebersamai, berbagi tawa, dan berjuang bersama melewati dinamika 8 semester ini.
14. Serta kepada seluruh masyarakat yang dengan perhatiannya melalui pertanyaan "kapan lulus?" secara tidak langsung menjadi motivasi ekstra bagi saya untuk segera menyelesaikan studi ini.
15. *Last but not least, i wanna thank me. I wanna thank me for believing in me. I wanna thank me for doing all this hard work. I wanna thank me for having no days off. I wanna thank me for never quitting. I wanna thank me for always being a giver and tryna give more than i receive. I wanna thank me for tryna do more right than wrong. I wanna thank me for just being me at all times.*

Penulis juga ingin menyadari bahwa karya ini masih jauh dari sempurna. Oleh karena itu, penulis dengan rendah hati menerima segala kritik dan saran yang membangun untuk perbaikan di masa yang akan datang. Akhir kata, semoga laporan skripsi ini dapat memberikan manfaat dan kontribusi positif bagi perkembangan ilmu pengetahuan, khususnya dalam bidang Teknik Informatika. Penulis berharap hasil penelitian ini dapat menjadi sumber referensi dan inspirasi bagi peneliti-peneliti yang akan datang.

Wassalamu'alaikum Wr. Wb.

Madiun, 9 Juli 2025

Erfia Nadia Safari

DAFTAR ISI

HALAMAN SAMPUL DEPAN	i
HALAMAN JUDUL.....	ii
HALAMAN PERSETUJUAN DOSEN PEMBIMBING.....	iii
HALAMAN PENGESAHAN DOSEN PENGUJI	iv
HALAMAN PERNYATAAN KEASLIAN KARYA ILMIAH SKRIPSI.....	v
HALAMAN PERSEMBAHAN.....	vi
HALAMAN MOTTO	vii
KATA PENGANTAR.....	viii
DAFTAR ISI	xii
DAFTAR TABEL.....	xv
DAFTAR GAMBAR	xvii
DAFTAR LAMPIRAN	xviii
ABSTRAK	xix
ABSTRACT.....	xx
BAB I PENDAHULUAN	1
A. Latar Belakang Masalah.....	1
B. Pembatasan Masalah	3
C. Perumusan Masalah	4
D. Tujuan Penelitian.....	4
E. Kegunaan Penelitian.....	4
BAB II KAJIAN PUSTAKA	6
A. Kajian Teoritis	6
1. Chatbot	6
2. <i>Artificial Intelligence</i>	7
3. <i>Natural processing Language (NLP)</i>	9
4. <i>Large Language Model (LLM)</i>	10
5. <i>Retrieval-Augmented Generation (RAG)</i>	10
6. <i>Vector Database</i>	11
7. <i>Text Chunking</i>	12
8. <i>Cosine Similarity</i>	13
9. <i>Embedding</i>	15
10. ChromaDB	15

11. Google Gemini	15
12. RAGAS	16
13. Skala Likert	16
14. Flowchart	17
15. UML	18
B. Kajian Empiris	23
C. Kerangka Berpikir	25
BAB III METODE PENELITIAN	28
A. Tempat dan Waktu Penelitian	28
B. Metode Pengembangan Sistem	28
C. Rancangan Penelitian	30
D. Teknik Pengembangan Sistem	34
BAB IV HASIL DAN PEMBAHASAN	37
A. Perencanaan Sistem	37
1. Analisis Permasalahan	37
2. Analisis Sistem yang Diusulkan	39
3. Kebutuhan Fungsional	41
4. Kebutuhan Non-Fungsional	42
5. Identifikasi Dataset	45
6. User Stories	46
B. Perancangan Sistem	46
1. Perancangan Arsitektur RAG	47
2. <i>Use Case Diagram</i>	49
3. <i>Class Diagram</i>	53
4. <i>Activity Diagram</i>	54
5. <i>Sequence Diagram</i>	58
C. Implementasi Sistem	61
1. Implementasi Arsitektur dan Antarmuka Website (<i>Front-end</i>)	61
2. Implementasi Arsitektur dan Organisasi Kode RAG (<i>Back-end</i>) ..	65
3. Implementasi <i>Pipeline Ingestion</i>	65
4. Implementasi Modul <i>Retrieval</i>	76
5. Implementasi Modul <i>Generation</i>	78
D. Pengujian Sistem	84

1. Pengujian <i>Retrieval Augmented Generation Assessment</i> (RAGAS)	
84	
2. Pengujian <i>User Acceptance Testing</i> (UAT).....	102
BAB V PENUTUP.....	114
A. Kesimpulan	114
B. Saran.....	115
DAFTAR PUSTAKA	116
LAMPIRAN.....	120

DAFTAR TABEL

Tabel 2. 1 Notasi <i>Flowchart</i>	17
Tabel 2. 2 Simbol <i>Use Case Diagram</i>	19
Tabel 3. 1 Waktu Penelitian.....	28
Tabel 4. 1 Kebutuhan Fungsional.....	41
Tabel 4. 2 Kebutuhan Non-Fungsional	42
Tabel 4. 3 Spesifikasi Perangkat Keras	43
Tabel 4. 4 Spesifikasi Perangkat Lunak	44
Tabel 4. 5 Daftar Biaya Kebutuhan Sistem	44
Tabel 4. 6 Identifikasi Dataset.....	45
Tabel 4. 7 <i>User Stories</i>	46
Tabel 4. 8 Parameter Konfigurasi Arsitektur RAG	48
Tabel 4. 9 <i>Use Case</i> Mengajukan Pertanyaan	50
Tabel 4. 10 <i>Use Case</i> Panduan Akademik	50
Tabel 4. 11 <i>Use Case</i> Topik Populer	52
Tabel 4. 12 Arsitektur <i>front-end</i>	61
Tabel 4. 13 Arsitektur <i>Back-end</i>	65
Tabel 4. 14 Statistik Hasil Proses <i>Chunking</i>	69
Tabel 4. 15 Contoh Metadata Hasil <i>Chunking</i>	70
Tabel 4. 16 Konfigurasi <i>Embedding</i>	71
Tabel 4. 17 Struktur Metadata <i>Embedding</i>	73
Tabel 4. 18 Penyimpanan ke <i>Vector Store ChromaDB</i>	75
Tabel 4. 19 Integrasi LLM Pada Query.....	82
Tabel 4. 20 Perhitungan Biaya Penggunaan Token.....	82
Tabel 4. 21 Hasil Evaluasi <i>Knowledge Base</i> Pertama.....	85
Tabel 4. 22 Hasil Metriks <i>Knowledge Base</i> Pertama	86
Tabel 4. 23 Hasil Evaluasi <i>Knowledge Base</i> Kedua	86
Tabel 4. 24 Hasil Metriks <i>Knowledge Base</i> Kedua.....	87
Tabel 4. 25 Hasil Evaluasi <i>Knowledge Base</i> Ketiga	88
Tabel 4. 26 Hasil Metriks <i>Knowledge Base</i> Ketiga.....	88
Tabel 4. 27 Hasil Evaluasi <i>Knowledge Base</i> Keempat	89
Tabel 4. 28 Hasil Metriks <i>Knowledge Base</i> Keempat.....	89
Tabel 4. 29 Hasil Evaluasi <i>Knowledge Base</i> Kelima	90
Tabel 4. 30 Hasil Metriks <i>Knowledge Base</i> Kelima	91
Tabel 4. 31 Hasil Evaluasi <i>Knowledge Base</i> Keenam.....	92
Tabel 4. 32 Hasil Metriks <i>Knowledge Base</i> Keenam	92
Tabel 4. 33 Hasil Evaluasi <i>Knowledge Base</i> Ketujuh	93
Tabel 4. 34 Hasil Metriks <i>Knowledge Base</i> Ketujuh	94
Tabel 4. 35 Metrik <i>Semantic Similarity</i>	95
Tabel 4. 36 Metrik <i>Answer Relevancy</i>	95
Tabel 4. 37 <i>Faithfulness</i>	96
Tabel 4. 38 Metrik <i>Context Recall</i>	97
Tabel 4. 39 Metrik <i>Factual Correctness Precision</i>	98
Tabel 4. 40 Metrik <i>Factual Correctness F1</i>	98
Tabel 4. 41 Metrik <i>Factual Precision Recall</i>	99

Tabel 4. 42 Interpretasi nilai metrik RAGAS.....	101
Tabel 4. 43 Rata-Rata Metrik Secara Keseluruhan	101
Tabel 4. 44 Pernyataan UAT	103
Tabel 4. 45 Perhitungan Bobot.....	104
Tabel 4. 46 Hasil Perhitungan Pernyataan	106
Tabel 4. 47 Range Nilai Interval	107

DAFTAR GAMBAR

Gambar 2. 1 Arsitektur Chatbot	6
Gambar 2. 2 Implementasi AI di berbagai bidang	7
Gambar 2. 3 Proses NLP bekerja	10
Gambar 2. 4 Arsitektur RAG	11
Gambar 2. 5 Komponen <i>Vector Database</i>	12
Gambar 2. 6 Contoh penggunaan <i>text chunking</i>	13
Gambar 2. 7 Kerangka Berpikir	27
Gambar 3. 1 Tahapan <i>Extreme Programming</i>	29
Gambar 3. 2 Rancangan Penelitian	31
Gambar 4. 1 <i>Flowchart</i> Analisis Permasalahan	38
Gambar 4. 2 <i>Flowchart</i> Analisis Sistem yang Diusulkan	40
Gambar 4. 3 Rancangan arsitektur RAG.....	47
Gambar 4. 4 <i>Use Case Diagram</i>	49
Gambar 4. 5 <i>Class Diagram</i>	53
Gambar 4. 6 <i>Activity Diagram Ingestion</i>	55
Gambar 4. 7 <i>Activity Diagram Proses Query</i>	57
Gambar 4. 8 <i>Sequence Diagram Proses Ingestion</i>	59
Gambar 4. 9 <i>Sequence Diagram Proses Query</i>	60
Gambar 4. 10 <i>Landing Page</i>	62
Gambar 4. 11 Halaman Chatbot.....	63
Gambar 4. 12 Fitur Interaktif	64
Gambar 4. 13 Fitur Responsif.....	64
Gambar 4. 14 <i>Document Loader</i>	67
Gambar 4. 15 <i>Text Chunking</i>	68
Gambar 4. 16 <i>Knowledge Base</i>	69
Gambar 4. 17 <i>Embedding Service</i>	72
Gambar 4. 18 Visualisasi Persebaran <i>Embedding Knowledge Base</i>	74
Gambar 4. 19 <i>Database ChromaDB</i>	75
Gambar 4. 20 <i>Query Embedder</i>	76
Gambar 4. 21 <i>Retriever</i>	77
Gambar 4. 22 <i>Context Assembler</i>	79
Gambar 4. 23 <i>Prompt Engineering</i>	80
Gambar 4. 24 Intruksi Sistem	81
Gambar 4. 25 <i>Conversation History</i>	83
Gambar 4. 26 <i>Heatmap</i> Evaluasi RAGAS	100

DAFTAR LAMPIRAN

Lampiran 1. Daftar Riwayat Hidup.....	120
Lampiran 2. Kuesioner Pada Google Formulir	121
Lampiran 3. Lembar Bukti Plagiasi	122
Lampiran 4. Validasi Sumber Pustaka	123