

**LAPORAN SKRIPSI**  
**PERANCANGAN *CHATBOT* UNTUK LAYANAN INFORMASI**  
**SEKOLAH MENGGUNAKAN *TRANSFORMER-BASED MODEL* (STUDI**  
**KASUS SMP SANTO LEO III)**



**Disusun oleh:**

**Shafa Elysia**

**2020230090**

**PROGRAM STUDI TEKNOLOGI INFORMASI**

**FAKULTAS TEKNIK**

**UNIVERSITAS DARMA PERSADA**

**2024**

## LEMBAR BIMBINGAN



### UNIVERSITAS DARMA PERSADA

Jl. Taman Malaka Selatan, Pondok Kelapa, Jakarta Timur, Indonesia 13450  
Telp. (021) 8649051, 8649053, 8649057 Fax. (021) 8649052  
E-mail : [humas@unsada.ac.id](mailto:humas@unsada.ac.id) Home page : <http://www.unsada.ac.id>

#### Instrumen Bimbingan Skripsi Program Studi Teknologi Informasi Periode 2023/2024 Genap

NIM : 2020230090  
 Nama : Shafa Elysia  
 Judul Skripsi : Perancangan Chatbot untuk Layanan Informasi Sekolah Menggunakan Transformer-Based Model (Studi Kasus: SMP Santo Leo III)  
 Dosen Pembimbing : Herianto, S.Pd., M.T.

No	BAB Utama Skripsi dan BATAS WAKTU Bimbingan	Materi Yang dibahas saat Konsultasi	Tanggal Bimbingan	TTD Dosen
1	BAB I PENDAHULUAN (15 April 2024 s.d 19 April 2024)  Paling lama upload: 19 April 2024	Sistematika penulisan Bab 1	16/04/24	
2		Bimbingan Bab I	17/04/24	
3		Tanggal BAB I di ACC pembimbing =>	18/04/24	
4	BAB II LANDASAN TEORI (22 April 2024 s.d 3 Mei 2024)  Paling lama upload : 3 Mei 2024	Sistematika penulisan Bab 2.	19/04/24	
5		Bimbingan Bab II	30/04/24	
6		Tanggal BAB II di ACC pembimbing =>	2/5/24	
7	BAB III ANALISA DAN PERANCANGAN METODOLOGI (6 Mei 2024 s.d 17 Mei 2024)  Paling lama upload : 17 Mei 2024	Bimbingan Bab III	13/05/24	
8		Bimbingan Bab III	16/05/24	
9		Tanggal BAB III di ACC pembimbing =>	17/05/24	



# UNIVERSITAS DARMA PERSADA

Jl. Taman Malaka Selatan, Pondok Kelapa, Jakarta Timur, Indonesia 13450

Telp. (021) 8649051, 8649053, 8649057 Fax. (021) 8649052

E-mail : [humas@unsada.ac.id](mailto:humas@unsada.ac.id) Home page : <http://www.unsada.ac.id>

10	<b>Percobaan/Demo Aplikasi atau Sistem</b> (20 Mei 2024 s.d 31 Mei 2024)  Paling lama upload : 31 Mei 2024	Bimbingan demo aplikasi	31/05/24	
11				
12				
13				
		Tanggal Aplikasi/Sistem ACC pembimbing =>	31/05/24	
14	<b>BAB IV HASIL DAN PEMBAHASAN</b> (3 Juni 2024 s.d 14 Juni 2024)  Paling lama upload : 14 Juni 2024	Bimbingan Bab 4	19/06/24	
15				
16				
		Tanggal BAB IV di ACC pembimbing =>	19/06	
17	<b>BAB V PENUTUP</b> (17 Juni 2024 s.d 19 Juni 2024)  Paling lama upload : 19 Juni	Bimbingan Bab 5	19/06/24	
18				
			Tanggal BAB V di ACC pembimbing =>	19/6/24

### Catatan :

- Mahasiswa harus konsultasi jauh-jauh hari sebelum batas akhir tanggal per BAB nya.
- Tanggal Bimbingan dan ACC per BAB **HARUS** sebelum batas tanggal maksimum, tetapi boleh sebelum tanggalnya jika bisa lebih cepat
- Dokumen ini WAJIB diupload ke gform yang ditentukan pada range tanggal setiap BAB
- Ujian Seminar ISI akan diadakan pada range tanggal : 24 s.d 28 Juni 2024

Di Acc Untuk Seminar Isi, pada tanggal : 19/06/2024

Oleh Dosen Pembimbing Skripsi

## LEMBAR REVISI

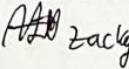
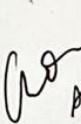


### UNIVERSITAS DARMA PERSADA

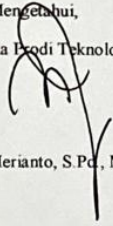
Jl. Taman Malaka Selatan, Pondok Kelapa, Jakarta Timur, Indonesia 13450  
Telp. (021) 8649051, 8649053, 8649057 Fax. (021) 8649052  
E-mail : humas@unsada.ac.id Home page : http://www.unsada.ac.id

#### LEMBAR REVISI - SIDANG SKRIPSI

NIM>Nama : 2020230090 - Shafa Elysia  
Fakultas/Prodi : Teknik / Teknologi Informasi

No.	Keterangan Revisi	Dosen
1.	No revision	 

Mengetahui,  
Ka Prodi Teknologi Informasi

  
Herianto, S.Pd., MT.

MONOLINGUAL • TRILINGUAL • ENERGI TERBARUKAN



## LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Shafa Elysia  
NIM : 2020230090  
Fakultas : Teknik  
Jurusan : Teknologi Informasi  
Judul Penelitian : PERANCANGAN *CHATBOT* UNTUK LAYANAN  
INFORMASI SEKOLAH MENGGUNAKAN  
*TRANSFORMER-BASED MODEL* (STUDI KASUS  
SMP SANTO LEO III)

Dengan ini saya menyatakan bahwa Laporan Skripsi ini disusun secara mandiri berdasarkan hasil survei, observasi, wawancara, serta disesuaikan dengan berbagai referensi lain yang relevan dengan topik Laporan Skripsi ini.

Demikian pernyataan ini saya buat dengan sebenar-benarnya.

Jakarta, 07 Agustus 2024



Shafa Elysia

**LEMBAR PENGESAHAN**  
**PERANCANGAN *CHATBOT* UNTUK LAYANAN INFORMASI**  
**SEKOLAH MENGGUNAKAN *TRANSFORMER-BASED MODEL* (STUDI**  
**KASUS SMP SANTO LEO III)**

Disusun oleh:

Nama: Shafa Elysia

NIM: 2020230090



Herianto, S.Pd., M.T.  
Kajur Teknologi Informasi

Herianto, S.Pd., M.T.  
Pembimbing Laporan

**LEMBAR PENGUJI LAPORAN SKRIPSI**

Laporan Skripsi yang berjudul:

**“PERANCANGAN *CHATBOT* UNTUK LAYANAN INFORMASI  
SEKOLAH MENGGUNAKAN *TRANSFORMER-BASED MODEL* (STUDI  
KASUS SMP SANTO LEO III)”**

Ini telah diujikan pada tanggal

24 Juli 2024





PENYELENGGARA PENDIDIKAN YAYASAN SANTO LEO  
**SMP SANTO LEO III**  
Terakreditasi A  
Jl. Dr. Cipto Mangunkusumo 5B Jababeka, Cikarang Utara  
Telp. (021) 28518283  
Email st.leo3.smp@gmail.com

---

## SURAT KETERANGAN

### Yang bertanda tangan di bawah ini:

Nama : Florentina Dwi Utamingtyas, S.Si., M.M  
Jabatan : Kepala Sekolah  
Unit Kerja : SMP Santo Leo III  
Alamat : Jl. Dr. Cipto Mangunkusumo Blok B-5 Kota Jababeka  
Rt.002/009 Kp. Proponcol, Simpangan, Kec. Cikarang  
Utara, Kab. Bekasi, Prov. Jawa Barat, 17530

### Dengan ini menerangkan bahwa:

Nama : Shafa Elysia  
NIM : 2020230090  
Program Studi : Teknologi Informasi  
Perguruan Tinggi : Universitas Darma Persada

Bahwa yang bersangkutan telah menyelesaikan penelitian di SMP Santo Leo III untuk memperoleh data dalam melaksanakan penyusunan skripsi yang berjudul **“PERANCANGAN *CHATBOT* UNTUK LAYANAN INFORMASI SEKOLAH MENGGUNAKAN *TRANSFORMER-BASED MODEL* (STUDI KASUS SMP SANTO LEO III)”**.

Demikian surat keterangan ini dibuat agar dapat digunakan sebagaimana mestinya.

Bekasi, 7 Agustus 2024

Florentina Dwi Utamingtyas, S.Si., M.M



## KATA PENGANTAR

Puji syukur ke hadirat Tuhan Yang Maha Esa, karena berkat rahmat dan karunia-Nya, penulis dapat menyelesaikan skripsi yang berjudul **“PERANCANGAN *CHATBOT* UNTUK LAYANAN INFORMASI SEKOLAH MENGGUNAKAN *TRANSFORMER-BASED MODEL* (STUDI KASUS SMP SANTO LEO III)”**. Skripsi ini disusun untuk memenuhi salah satu syarat dalam memperoleh gelar Sarjana Teknik pada Fakultas Teknik Universitas Darma Persada.

Dalam penyusunan skripsi ini, penulis banyak mendapatkan bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Bapak Dr. Ade Supriatna, M.T., selaku Dekan Fakultas Teknik Universitas Darma Persada.
2. Bapak Herianto, S.Pd., M.T., selaku ketua jurusan Teknologi Informasi Universitas Darma Persada serta dosen pembimbing yang telah memberikan bimbingan dan arahan selama penyusunan laporan skripsi ini.
3. Seluruh dosen dan staf Fakultas Teknik Universitas Darma Persada yang telah memberikan ilmu dan bantuan selama penulis menempuh pendidikan.
4. Ibu Florentina Dwi Utamingtyas, S.Si., M.M, selaku kepala sekolah SMP Santo Leo III yang menjadi lokasi penelitian ini.
5. Orang tua tercinta dan keluarga besar yang selalu memberikan doa, dukungan, dan motivasi yang tiada henti kepada penulis, khususnya dalam menjalani masa perkuliahan.

6. Teman-teman dan sahabat-sahabat yang selalu memberikan semangat dan bantuan dalam penyelesaian skripsi ini.

Akhir kata semoga laporan skripsi ini bermanfaat bagi kita semua.

Jakarta, 07 Agustus 2024



Shafa Elysia



## ABSTRAKSI

Penelitian ini bertujuan untuk merancang dan mengembangkan sistem *Chatbot* berbasis model *Transformers* dengan metode *Retrieval Augmented Generation* (RAG) untuk layanan informasi di SMP Santo Leo III. *Chatbot* ini dikembangkan untuk meningkatkan aksesibilitas dan efisiensi penyampaian informasi terkait peraturan sekolah, kegiatan ekstrakurikuler, kalender akademik, dan informasi relevan lainnya. Model yang digunakan dalam pengembangan chatbot ini adalah *LLaMA-3-8B-Instruct*, *Mistral-7B-Instruct-v0.3*, dan *Zephyr-7B-β*. Hasil pengujian terhadap dataset testing yang berisi 30 baris pertanyaan terkait SMP Santo Leo III menunjukkan bahwa model *LLaMA-3-8B-Instruct* dan *Mistral-7B-Instruct-v0.3* mencapai akurasi 100%, sementara model *Zephyr-7B-β* mencapai akurasi 70%. Penggunaan *Transformer-Based Model* digabungkan dengan metode RAG terbukti efektif dalam menggabungkan informasi yang relevan dengan konteks percakapan, sehingga meningkatkan kualitas jawaban yang diberikan oleh *Chatbot*. Penelitian ini menunjukkan bahwa implementasi teknologi *Transformers* dan RAG menunjukkan potensi besar dalam meningkatkan interaksi dengan mesin yang lebih intuitif dan responsif pada domain informasi sekolah. Kedua teknologi tersebut dapat mengurangi keterbatasan yang ada pada sistem *Chatbot* berbasis menu, aturan, dan *machine learning* konvensional.

**Kata Kunci:** *Chatbot*, Layanan informasi sekolah, *Retrieval augmented generation*, *Transformer-based model*

## DAFTAR ISI

LEMBAR JUDUL	
LEMBAR BIMBINGAN .....	i
LEMBAR REVISI .....	iii
LEMBAR PERNYATAAN .....	iv
LEMBAR PENGESAHAN .....	v
LEMBAR PENGUJI LAPORAN SKRIPSI .....	vi
SURAT KETERANGAN .....	vii
KATA PENGANTAR .....	viii
ABSTRAKSI .....	x
DAFTAR ISI .....	xi
DAFTAR GAMBAR .....	xv
DAFTAR TABEL .....	xvi
DAFTAR LAMPIRAN .....	xvii
BAB I .....	1
1.1 Latar Belakang .....	1
1.2 Identifikasi Masalah .....	4
1.3 Rumusan Masalah .....	4
1.4 Batasan Masalah .....	4
1.5 Tujuan Penelitian .....	5
1.6 Manfaat Penelitian .....	5
1.7 Metode Pengumpulan Data .....	6
1.7.1 Wawancara .....	6
1.7.2 Dokumen .....	6
1.8 Metode Pengembangan Sistem .....	7
1.9 Sistematika Penulisan .....	8
BAB II .....	11

2.1 Tinjauan Pustaka .....	11
2.1.1 Aplikasi <i>Chatbot</i> sebagai Layanan Informasi Sekolah .....	11
2.1.1.1 Pengertian dan Jenis-Jenis <i>Chatbot</i> .....	11
2.1.1.2 Aplikasi <i>Chatbot</i> sebagai Layanan Informasi Sekolah .....	12
2.1.1.3 Aplikasi <i>Chatbot</i> dengan Pendekatan <i>Neural Networks</i> .....	14
2.1.2 <i>Natural Language Processing</i> dengan Pendekatan <i>Deep Learning</i> .....	15
2.1.2.1 <i>Artificial Intelligence, Machine Learning, dan Deep Learning</i> .....	15
2.1.2.2 <i>Natural Language Processing (NLP)</i> .....	16
2.1.2.2.1 <i>Tokenization</i> .....	17
2.1.2.2.2 <i>Word Embeddings</i> .....	18
2.1.2.3 <i>Language Model (LM)</i> .....	18
2.1.2.4 Parameter LLM untuk Tugas <i>Text Generation</i> .....	19
2.1.3 Metode <i>Transformers</i> .....	20
2.1.3.1 Definisi <i>Transformers</i> .....	20
2.1.3.2 Arsitektur <i>Encoder Model Transformers</i> .....	22
2.1.3.2.1 <i>Self-Attention</i> .....	23
2.1.3.2.2 <i>Feed-Forward Layer</i> .....	24
2.1.3.2.3 <i>Layer Normalization</i> .....	24
2.1.3.2.4 <i>Positional Embeddings</i> .....	25
2.1.3.3 Arsitektur <i>Decoder Model Transformers</i> .....	25
2.1.3.4 <i>Transformer-Based Pretrained Model</i> .....	26
2.1.3.5 Jenis-jenis <i>Transformer-Based Pretrained Model</i> Terkait .....	27
2.1.3.5.1 <i>Sentence-Transformers (SBERT)</i> .....	27
2.1.3.5.2 <i>Large Language Model Meta AI (LlaMA)</i> .....	28
2.1.3.5.3 <i>Mistral-7B</i> .....	29
2.1.3.5.4 <i>Zephyr</i> .....	29
2.1.4 Metode <i>Retrieval Augmented Generation (RAG)</i> .....	30
2.1.4.1 Definisi <i>Retrieval Augmented Generation (RAG)</i> .....	30

2.1.4.2 Metode RAG sebagai Alternatif <i>Fine-Tuning</i> LLM .....	31
2.1.5 CRISP-DM sebagai Tahap Merancang Model <i>Machine Learning</i> .....	34
2.1.6 Pemodelan Sistem UML .....	36
2.1.6.1 <i>Use Case Diagram</i> .....	36
2.1.6.2 <i>Activity Diagram</i> .....	38
2.1.7 <i>Software</i> dan <i>Tools</i> Terkait .....	41
2.1.7.1 Bahasa Pemrograman <i>Python</i> .....	41
2.1.7.2 <i>Streamlit</i> .....	42
2.1.7.3 <i>Hugging Face Ecosystem</i> .....	42
2.1.7.4 <i>LangChain</i> .....	43
2.1.7.5 <i>MongoDB</i> .....	43
2.2 Kajian Literatur .....	44
BAB III .....	49
3.1 Lokasi, Penjadwalan dan Bidang Penelitian .....	49
3.1.1 Lokasi Penelitian .....	49
3.1.2 Jadwal Tahapan Penelitian .....	49
3.1.3 Bidang dan Jenis Penelitian .....	51
3.2 Rancangan Metodologi Penelitian .....	52
3.2.1 Perancangan UML .....	52
3.2.1.1 <i>Use Case Diagram</i> .....	52
3.2.1.2 <i>Activity Diagram</i> .....	55
3.2.2 Perancangan Struktur <i>Database</i> .....	62
3.2.3 Perancangan <i>Interface</i> Aplikasi.....	63
3.2.4 Perancangan <i>Flowchart</i> Algoritma.....	66
3.2.4.1 Algoritma <i>Transformers</i> .....	66
3.2.4.2 Algoritma <i>Retrieval Augmented Generation (RAG)</i> .....	68
3.2.5 Rancangan Tahap CRISP-DM .....	69
3.2.5.1 Analisa Tahap <i>Business Understanding</i> .....	69

3.2.5.2 Analisa Tahap <i>Data Understanding</i> .....	69
3.2.5.3 Rancangan Tahap <i>Data Preparation</i> .....	70
3.2.5.4 Rancangan Tahap <i>Modeling</i> .....	70
3.2.5.5 Rancangan Tahap <i>Evaluation</i> .....	71
3.2.5.6 Rancangan Tahap <i>Deployment</i> .....	71
BAB IV .....	73
4.1 Hasil Penelitian .....	73
4.1.1 Spesifikasi <i>Hardware</i> dan <i>Software</i> yang Digunakan .....	73
4.1.1.1 Spesifikasi <i>Hardware</i> .....	73
4.1.1.2 Spesifikasi <i>Software</i> .....	73
4.1.2 Tampilan <i>Interface</i> Hasil <i>Deploy</i> .....	74
4.1.3 Struktur <i>Database</i> .....	75
4.2 Analisa Hasil .....	77
4.2.1 Percobaan <i>Input – Output</i> .....	77
4.2.2 <i>Testing</i> Hasil .....	78
4.2.3 Modifikasi atau Optimalisasi dari Sistem Terdahulu .....	80
4.2.4 Proses <i>Deploy</i> Sistem Aplikasi .....	81
BAB V .....	83
5.1 Kesimpulan .....	83
5.2 Saran .....	83
DAFTAR PUSTAKA .....	86
LAMPIRAN	

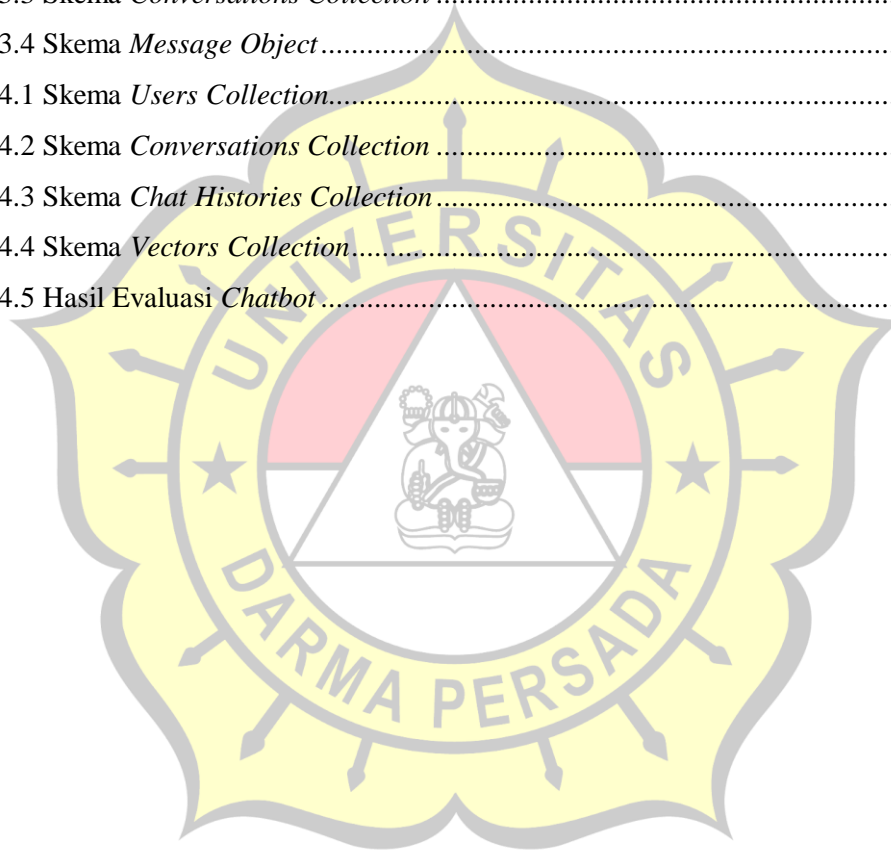
## DAFTAR GAMBAR

Gambar 2.1 Arsitektur <i>end-to-end neural</i> (McTear, 2021:126) .....	14
Gambar 2.2 Alur NLU (Lane et al, 2023) .....	16
Gambar 2.3 Alur NLG (Lane et al, 2024:5) .....	17
Gambar 2.4 Arsitektur <i>Transformers</i> (Lee, 2023:189) .....	21
Gambar 2.5 Gambar arsitektur <i>encoder</i> (Tunstall et al, 2022:60) .....	23
Gambar 2.6 Gambar arsitektur <i>decoder</i> (Tunstall et al, 2022:77) .....	26
Gambar 2.7 Tahap CRISP-DM (Cirillo, 2017:70).....	34
Gambar 2.8 Gambaran umum <i>Hugging Face ecosystem</i> (Tunstall et al, 2022:15) .....	42
Gambar 3.1 <i>Use Case Diagram User</i> .....	53
Gambar 3.2 <i>Use Case Diagram Administrator</i> .....	54
Gambar 3.3 Proses <i>login</i> dan registrasi pengguna .....	55
Gambar 3.4 Proses <i>chat</i> dengan <i>Chatbot</i> .....	56
Gambar 3.5 Proses menghapus riwayat <i>chat</i> .....	57
Gambar 3.6 Proses mengubah data profil pengguna.....	58
Gambar 3.7 Proses mengelola data pengguna .....	59
Gambar 3.8 Proses mengevaluasi kinerja <i>Chatbot</i> .....	60
Gambar 3.9 Proses mengelola <i>knowledge base</i> .....	61
Gambar 3.10 Diagram relasi <i>database</i> .....	63
Gambar 3.11 Tampilan halaman <i>login</i> .....	64
Gambar 3.12 Tampilan halaman registrasi.....	64
Gambar 3.13 Tampilan halaman utama <i>user</i> .....	65
Gambar 3.14 Tampilan profil <i>user</i> .....	65
Gambar 3.15 <i>Flowchart</i> Algoritma <i>Transformers</i> .....	66
Gambar 3.16 <i>Flowchart</i> Algoritma RAG.....	68
Gambar 4.1 Halaman <i>Login</i> .....	74
Gambar 4.2 Halaman Registrasi .....	74
Gambar 4.3 Halaman Utama <i>User</i> .....	75
Gambar 4.4 Halaman <i>Dashboard Administrator</i> .....	75



## DAFTAR TABEL

Tabel 2.1 Tabel Perbandingan Metode RAG dan <i>Fine-Tuning</i> .....	32
Tabel 2.2 Komponen-komponen dari <i>use case diagram</i> .....	37
Tabel 2.3 Komponen-komponen dari <i>activity diagram</i> . ....	39
Tabel 2.4 Paper penelitian terkait. ....	44
Tabel 3.1 Jadwal penelitian .....	49
Tabel 3.2 Skema <i>Users Collection</i> .....	62
Tabel 3.3 Skema <i>Conversations Collection</i> .....	62
Tabel 3.4 Skema <i>Message Object</i> .....	63
Tabel 4.1 Skema <i>Users Collection</i> .....	76
Tabel 4.2 Skema <i>Conversations Collection</i> .....	76
Tabel 4.3 Skema <i>Chat Histories Collection</i> .....	77
Tabel 4.4 Skema <i>Vectors Collection</i> .....	77
Tabel 4.5 Hasil Evaluasi <i>Chatbot</i> .....	79



## DAFTAR LAMPIRAN

Lampiran 1 Surat Keterangan Bebas Plagiat

Lampiran 2 Hasil Cek *Turnitin*

Lampiran 3 *Source Code Function* Koneksi ke *Database*

Lampiran 4 *Source Code Function* Fitur *Chat*

Lampiran 5 *Source Code Function* Fitur RAG

Lampiran 6 Hasil Evaluasi *Chatbot* Menggunakan Model *LLaMA-3-8B-Instruct*

Lampiran 7 Hasil Evaluasi *Chatbot* Menggunakan Model *Mistral-7B-Instruct-v.03*





**TEKNOLOGI INFORMASI**  
**UNIVERSITAS DARMA PERSADA**