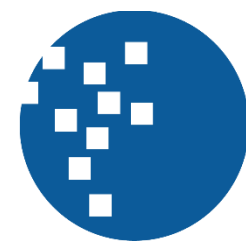


**PENGEMBANGAN CHATBOT BERBASIS LARGE
LANGUAGE MODEL DENGAN METODE RETRIEVAL-
AUGMENTED GENERATION UNTUK LAYANAN
INFORMASI DI UNIVERSITAS MULTIMEDIA NUSANTARA**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

LAPORAN MBKM INDEPENDEN

James Andersen

00000069612

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

**PENGEMBANGAN CHATBOT BERBASIS LARGE
LANGUAGE MODEL DENGAN METODE RETRIEVAL-
AUGMENTED GENERATION UNTUK LAYANAN
INFORMASI DI UNIVERSITAS MULTIMEDIA NUSANTARA**



LAPORAN MBKM PROYEK INDEPENDEN

Diajukan Sebagai Salah Satu Syarat untuk Memperoleh

Gelar Sarjana Komputer

James Andersen

00000069612

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : James Andersen
Nomor Induk Mahasiswa : 00000069612
Program studi : Sistem Informasi

Laporan MBKM Penelitian dengan judul:

PENGEMBANGAN CHATBOT BERBASIS LARGE LANGUAGE MODEL DENGAN METODE RETRIEVAL-AUGMENTED GENERATION UNTUK LAYANAN INFORMASI DI UNIVERSITAS MULTIMEDIA NUSANTARA merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/ penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan MBKM, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk laporan MBKM yang telah saya tempuh.

Tangerang, 17 Juli 2025



(James Andersen)

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Multimedia Nusantara, saya yang bertandatangan di bawah ini:

Nama : James Andersen
NIM : 00000069612
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika
JenisKarya : Laporan MBKM

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Multimedia Nusantara Hak Bebas Royalti Noneksklusif (Non-exclusive Royalty-Free Right) atas karya ilmiah saya yang berjudul:

PENGEMBANGAN CHATBOT BERBASIS LARGE LANGUAGE MODEL DENGAN METODE RETRIEVAL-AUGMENTED GENERATION UNTUK LAYANAN INFORMASI DI UNIVERSITAS MULTIMEDIA NUSANTARA serta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non eksklusif ini Universitas Multimedia Nusantara berhak menyimpan, mengalih media / format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis / pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Tangerang, 22 Juli 2025

Yang menyatakan,



(James Andersen)

KATA PENGANTAR

Puji Syukur atas selesainya penulisan Laporan MBKM ini dengan judul: “PENGEMBANGAN CHATBOT BERBASIS LARGE LANGUAGE MODEL DENGAN METODE RETRIEVAL-AUGMENTED GENERATION UNTUK LAYANAN INFORMASI DI UNIVERSITAS MULTIMEDIA NUSANTARA” dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar <<Strata>> Jurusan Sistem Informaso Pada Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan tugas akhir ini, sangatlah sulit bagi saya untuk menyelesaikan tugas akhir ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Bapak Dr. Erick Fernando, S.Kom., M.S.I., sebagai Pembimbing yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya tesis ini.
5. Dr. Irmawati, S.Kom., M.M.S.I., sebagai Pembimbing Lapangan yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya laporan MBKM Penelitian.
6. Kepada keluarga yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan laporan MBKM ini.

Semoga karya ilmiah ini mendapatkan manfaat positif untuk para pembaca.

Tangerang, 22 Juli 2025



(James Andersen)



PENGEMBANGAN CHATBOT BERBASIS LARGE LANGUAGE MODEL DENGAN METODE RETRIEVAL- AUGMENTED GENERATION UNTUK LAYANAN INFORMASI DI UNIVERSITAS MULTIMEDIA NUSANTARA

(James Andersen)

ABSTRAK

Laporan ini menguraikan pengembangan chatbot berbasis Large Language Model (LLM) yang mengadopsi teknik Retrieval-Augmented Generation (RAG) sebagai solusi inovatif untuk layanan informasi akademik dan administratif di Universitas Multimedia Nusantara (UMN). Proyek ini dipilih karena belum adanya layanan otomatis berbasis AI yang mampu memberikan jawaban cepat dan kontekstual melalui platform yang familiar bagi mahasiswa dan staf.

Pelaksanaan proyek selama satu semester mencakup pengumpulan data internal kampus, pengembangan model LLM dengan metode RAG, integrasi antarmuka pengguna menggunakan Streamlit, serta evaluasi performa chatbot. Berbagai tantangan seperti keterbatasan data, kompleksitas implementasi RAG, dan manajemen waktu berhasil diatasi melalui eksperimen parameter model, peningkatan kerja sama tim, dan konsultasi intensif dengan pembimbing.

Hasilnya, penulis memperoleh pemahaman mendalam tentang pengembangan sistem AI, penguasaan pemrograman Python, serta konsep fine-tuning dan retrieval-based generation. Selain aspek teknis, proyek ini juga memperkuat kemampuan komunikasi, pengambilan keputusan, dan etos kerja dalam konteks kerja tim.

Kata kunci: chatbot, large-language-model, retrieval-augmented generation, universitas multimedia nusantara, layanan informasi

DEVELOPMENT OF CHATBOT BASED ON LARGE LANGUAGE MODEL WITH RETRIEVAL-AUGMENTED GENERATION METHOD FOR INFORMATION SERVICES AT MULTIMEDIA NUSANTARA UNIVERSITY

(James Andersen)

ABSTRACT (English)

This report presents the development of a chatbot leveraging a Large Language Model (LLM) combined with the Retrieval-Augmented Generation (RAG) technique as an innovative approach to enhance academic and administrative information services at Multimedia Nusantara University (UMN). The project was initiated to address the lack of an AI-driven automated system capable of delivering fast and context-aware responses through a user-friendly platform familiar to students and staff.

Conducted over the course of one semester, the project encompassed data collection from internal university sources, development of the LLM model using RAG, integration of the chatbot interface via Streamlit, and performance evaluation. Key challenges faced included limited availability of institutional data, the complexity of implementing RAG, and effective time management. These were mitigated through systematic experimentation with model parameters, strengthened team collaboration, and ongoing guidance from supervisors.

The project provided the author with valuable expertise in AI system development, proficiency in Python programming, and practical experience with fine-tuning and retrieval-based generation techniques. Additionally, it fostered improvements in work ethics, communication skills, and decision-making within a collaborative team setting.

Keywords: *chatbot, large-language-model, retrieval-augmented generation, multimedia nusantara university, information service*

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	2
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS	3
KATA PENGANTAR	4
ABSTRAK	6
<i>ABSTRACT (English)</i>	7
DAFTAR ISI	8
DAFTAR TABEL	10
DAFTAR GAMBAR	11
DAFTAR LAMPIRAN	13
BAB I PENDAHULUAN	14
1.1. Latar Belakang	14
1.2. Rumusan Masalah	16
1.3. Maksud dan Tujuan	16
1.3.1 Maksud Kegiatan	16
1.3.2 Tujuan Kegiatan	16
1.4. Manfaat	17
1.5 Waktu dan Prosedur	18
1.5.1 Waktu Pelaksanaan Kegiatan	18
1.5.2 Prosedur Pelaksanaan Kegiatan	19
BAB II GAMBARAN UMUM MITRA/PERUSAHAAN	21
2.1 Deskripsi Mitra/Perusahaan	21
2.1.1 Visi Misi	23
2.2 Struktur Organisasi Mitra/Perusahaan	24
2.3 Portofolio Mitra/Perusahaan	26
BAB III PELAKSANAAN PROYEK	28
3.1 Kedudukan dan Koordinasi	28
3.2 Tugas dan Uraian Kerja	30
3.3 Uraian Pelaksanaan Kerja	34
3.3.1 Proses Pelaksanaan	34

3.3.1.1	Identifikasi Permasalahan	34
3.3.1.2	Studi Literatur	35
3.3.1.3	Penyusunan Desain Teknis	37
3.3.1.4	Pengembangan Produk/Jasa Layanan	38
3.3.1.5	Pengumpulan Dataset Internal Kampus	38
3.3.1.6	Implementasi dan Pengembangan Sistem	39
3.3.1.7	Pengujian dan Evaluasi Chatbot	52
3.3.1.8	Validasi dan Perbaikan	54
3.3.1.9	Dokumentasi dan Penyusunan Laporan	56
3.4	Kendala yang Ditemukan	56
3.5	Solusi atas Kendala yang Ditemukan	58
BAB IV	SIMPULAN DAN SARAN	59
4.1	Simpulan	59
4.2	Saran	64
DAFTAR PUSTAKA		68
LAMPIRAN		72

DAFTAR TABEL

Tabel 1.1 Timeline Pelaksanaan Kegiatan	17
Tabel 3.1 Timeline Pelaksanaan Proyek	29



DAFTAR GAMBAR

Gambar 2.1 Logo Universitas Multimedia Nusantara	20
Gambar 2.2 Struktur Rektorat UMN	23
Gambar 2.3 Struktur Akademik UMN	24
Gambar 2.4 Struktur MBKM Proyek Independen UMN	25
Gambar 3.1 Website Universitas Multimedia Nusantara	37
Gambar 3.2 Handbook MyUMN	38
Gambar 3.3 Model LLM	38
Gambar 3.4 Import Library	39
Gambar 3.5 Konfigurasi variabel	39
Gambar 3.6 Tokenizer	40
Gambar 3.7 Embeddings	40
Gambar 3.8 Pengolahan dokumen PDF	41
Gambar 3.9 Pemrosesan dokumen PDF	42
Gambar 3.10 Inisialisasi ChromaDB	43
Gambar 3.11 Inisialisasi Model LLM	44
Gambar 3.12 Inisialisasi RAG Query	44
Gambar 3.13 Konfigurasi StreamLit	45
Gambar 3.14 Kostumisasi Kontainer dan Gaya Chatbot	46
Gambar 3.15 Kostumisasi Tampilan Chat	47
Gambar 3.16 Menyesuaikan Tampilan Chat Mode Gelap dan Terang	48
Gambar 3.17 Tampilan Header Aplikasi	49
Gambar 3.18 Inisialisasi Riwayat Chat	49
Gambar 3.19 Tampilan Input Chat User	50
Gambar 3.20 Kondisi Pengiriman Pertanyaan	50
Gambar 3.21 Tampilan Riwayat Chat	51
Gambar 3.22 Hasil Evaluasi	51
Gambar 3.23 Tampilan Website (Mode Gelap)	54



DAFTAR LAMPIRAN

Lampiran A Nama Lampiran	71
Surat Pengantar MBKM - MBKM 01	72
Kartu MBKM - MBKM 02	73
Daily Task MBKM - MBKM 03	74
Lembar Verifikasi Laporan MBKM - MBKM 04	89
Bukti Pendaftaran dan Partisipasi Lomba	90
Lampiran Pengecekan Hasil Turnitin	91