

DAFTAR ISI

LEMBAR PENGESAHAN.....	ii
PERNYATAAN BEBAS PLAGIASI.....	iii
KATA PENGANTAR.....	iv
DAFTAR ISI.....	vi
DAFTAR GAMBAR	ix
DAFTAR TABEL.....	xi
DAFTAR LAMPIRAN.....	xii
INTISARI.....	xiii
ABSTRACT.....	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Tinjauan Pustaka	6
2.2 Landasan Teori	12
2.2.1 <i>Chatbot</i>	12
2.2.2 Large Language Model (LLM)	12
2.2.3 <i>Retrieval Augmented Generation</i> (RAG).....	13
2.2.4 Docker	14
2.2.5 n8n.....	15
2.2.6 Komputasi	16
2.2.7 Akurasi	17

2.2.8 Nginx.....	18
2.2.9 Ngrok.....	18
2.3 Hipotesis.....	19
BAB III METODE PROYEK AKHIR	20
3.1 Bahan.....	20
3.2 Alat.....	21
3.3 Tahapan Proyek Akhir	21
3.4 Topologi Sistem	23
3.5 Konfigurasi n8n.....	24
3.6 Integrasi <i>workflow</i> dan <i>website</i>	25
3.6.2 Embedding <i>Workflow</i>	26
3.6.3 NON-RAG <i>Workflow</i>	28
3.6.4 RAG <i>Workflow</i>	32
3.6.5 Rancangan Web Interface	35
3.7 Skenario Pengujian.....	39
3.7.1 Pengujian <i>Baseline</i>	39
3.7.2 Pengujian Integrasi <i>Workflow</i>	40
3.7.3 Pengujian Beban Komputasi	40
3.7.4 Pengujian Akurasi	42
BAB IV HASIL DAN PEMBAHASAN	46
4.1 Hasil Pengujian <i>Baseline</i>	46
4.2 Hasil Integrasi <i>Workflow</i> n8n.....	47
4.2.1 Embedding Model.....	47
4.2.2 NON-RAG <i>Workflow</i>	48
4.2.3 RAG <i>Workflow</i>	50

4.2.4 Website.....	52
4.3 Hasil Pengujian <i>Chatbot</i> Berbasis CPU.....	53
4.4 Hasil Pengujian <i>Chatbot</i> Berbasis GPU.....	56
4.5 Perbandingan Akurasi (NON-RAG vs. RAG).....	57
4.6 Analisis Perbandingan RAG	60
BAB V PENUTUP.....	63
5.1 Kesimpulan.....	63
5.2 Saran.....	64
DAFTAR PUSTAKA	65
LAMPIRAN.....	69

DAFTAR GAMBAR

Gambar 3. 1 Tahapan proyek akhir	21
Gambar 3. 2 Topologi sistem	24
Gambar 3. 3 Pembuatan akun n8n	25
Gambar 3. 4 Proses <i>workflow embedding</i> : pengambilan data.....	26
Gambar 3. 5 Proses <i>workflow embedding</i> : ekstraksi data.....	27
Gambar 3. 6 Proses <i>workflow embedding</i> : simpan vektor	27
Gambar 3. 7 Rancangan Non-Rag.....	28
Gambar 3. 8 <i>Node trigger chat</i>	28
Gambar 3. 9 <i>AI agent NON-RAG</i>	29
Gambar 3. 10 <i>Posgres chat memory</i>	30
Gambar 3. 11 <i>Ollama chat model</i>	30
Gambar 3. 12 <i>log chat</i> ke google doc.....	31
Gambar 3. 13 <i>Chat output</i> ke pengguna.....	31
Gambar 3. 14 Rancangan RAG.....	32
Gambar 3. 15 <i>AI agent RAG</i>	33
Gambar 3. 16 <i>Intruksi AI agent RAG</i>	34
Gambar 3. 17 <i>Qdrant vector node RAG</i>	34
Gambar 3. 18 <i>Embedding node RAG</i>	35
Gambar 3. 19 <i>URL webhook</i>	35
Gambar 3. 20 Berkas <i>index.html website</i>	36
Gambar 3. 21 <i>Akses website dan chatbot</i>	37
Gambar 3. 22 Halaman pendaftaran ngrok	37
Gambar 3. 23 Halaman <i>dashboard</i> ngrok	38
Gambar 3. 24 Status layanan ngrok	39
Gambar 3. 25 <i>Interface Hwinfo64</i>	40
Gambar 4. 1 Hasil pengujian <i>workflow embedding model</i>	48
Gambar 4. 2 Isi dokumen pada qdrant	48
Gambar 4. 3 <i>NON-RAG Workflow</i>	49
Gambar 4. 4 <i>RAG workflow</i>	51
Gambar 4. 5 Hasil pengujian <i>chatbot pada website</i>	53



Gambar 4. 6 Grafik waktu eksekusi NON-RAG dan RAG CPU.....	54
Gambar 4. 7 Grafik waktu eksekusi NON-RAG dan RAG GPU	56
Gambar 4. 8 Grafik waktu eksekusi RAG CPU-GPU	61
Gambar 4. 9 Grafik CPU dan RAM RAG	62

DAFTAR TABEL

Tabel 2. 1 Penelitian-penelitian terkait	9
Tabel 2. 1 Penelitian-penelitian terkait (lanjutan)	10
Tabel 2. 1 Penelitian-penelitian terkait (lanjutan)	11
Tabel 3. 1 Bahan proyek akhir	20
Tabel 3. 2 Spesifikasi Server	21
Tabel 3. 3 Credentials.....	25
Tabel 3. 4 Parameter uji integrasi.....	40
Tabel 3. 5 Tabel pertanyaan komputasi	41
Tabel 3. 6 Tabel pertanyaan akurasi	42
Tabel 3. 7 Tabel pertanyaan akurasi (lanjutan)	43
Tabel 3. 7 Tabel pertanyaan akurasi (lanjutan)	44
Tabel 4. 1 Hasil pengujian sistem sebelum diimplementasikan <i>chatbot</i>	46
Tabel 4. 2 Hasil pengujian integrasi <i>workflow</i> n8n	47
Tabel 4. 3 Data NON-RAG CPU	49
Tabel 4. 4 Data NON-RAG GPU.....	50
Tabel 4. 5 Data RAG CPU	51
Tabel 4. 6 Data RAG GPU	51
Tabel 4. 7 Perbandingan komputasi dengan CPU	54
Tabel 4. 8 Perbandingan komputasi dengan GPU.....	56
Tabel 4. 9 Hasil akurasi NON-RAG	58
Tabel 4. 10 Hasil akurasi RAG	59
Tabel 4. 11 Perbandingan RAG CPU dan GPU.....	61

DAFTAR LAMPIRAN

Lampiran 1 <i>Workflow embedding</i>	69
Lampiran 2 <i>Workflow NON-RAG</i>	69
Lampiran 3 <i>Workflow RAG</i>	69
Lampiran 4 Konfigurasi <i>workflow</i> keseluruhan	70
Lampiran 5 Data hasil <i>query</i> RAG CPU	86
Lampiran 6 Data hasil <i>query</i> NON-RAG CPU	92
Lampiran 7 Data hasil <i>query</i> RAG GPU	101
Lampiran 8 Data hasil <i>query</i> NON-RAG GPU	106