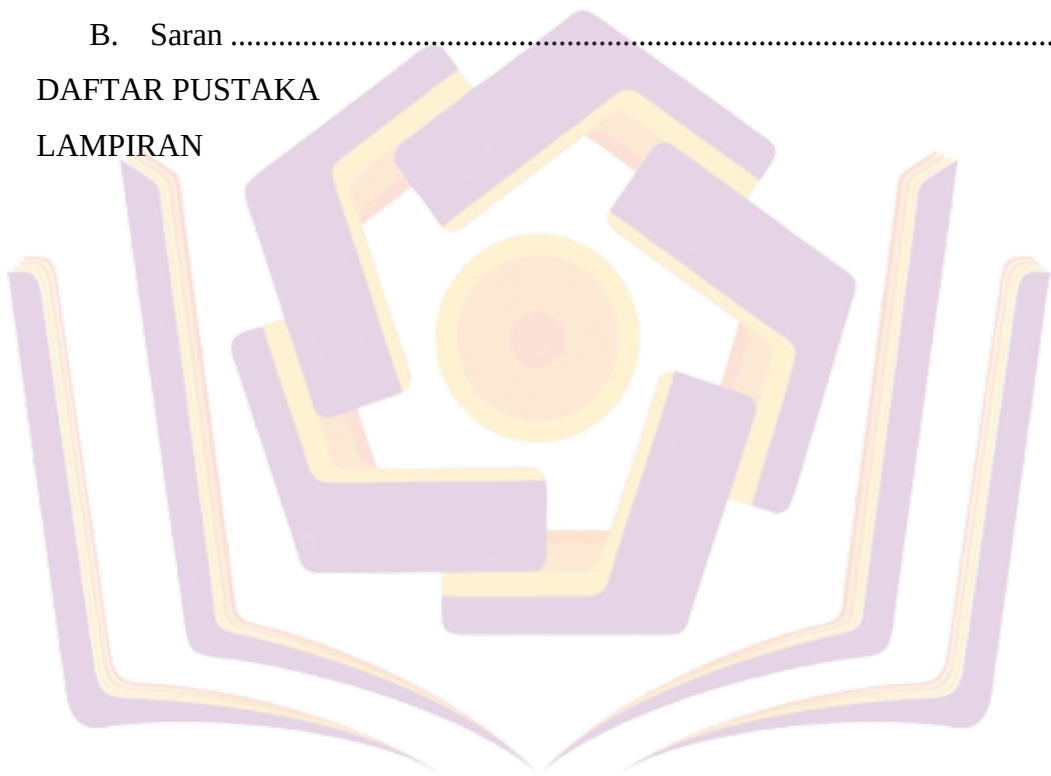


DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL.....	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
HALAMAN PERNYATAAN KEASLIAN	v
HALAMAN PERSEMBAHAN	vi
HALAMAN MOTTO	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiii
DAFTAR ISTILAH	xiv
DAFTAR LAMPIRAN.....	xvii
ABSTRAK.....	xviii
ABSTRACT.....	xix
BAB I PENDAHULUAN.....	1
A. Latar Belakang Masalah	1
B. Rumusan Masalah.....	8
C. Batasan Masalah	8
D. Tujuan Penelitian	10
E. Manfaat Penelitian	11
BAB II TINJAUAN PUSTAKA.....	12
A. Landasan Teori.....	12
B. Penelitian Sebelumnya.....	30
BAB III METODE PENELITIAN.....	35
A. Tempat dan Waktu Penelitian.....	35
B. Metode Pengumpulan Data.....	35
C. Alat dan Bahan Penelitian.....	37
D. Konsep Penelitian	38

BAB IV HASIL DAN PEMBAHASAN	56
A. Hasil Pengumpulan Data.....	56
B. Hasil Persiapan Data	57
C. Skor Metode <i>Retrieval</i>	63
D. Evaluasi Metrik dan Analisis Hasil.....	78
E. Penerapan <i>Retrieval</i> pada <i>Chatbot</i>	91
BAB V PENUTUP.....	94
A. Kesimpulan	94
B. Saran	95
DAFTAR PUSTAKA	
LAMPIRAN	



DAFTAR TABEL

Tabel 2. 1 Penelitian Sebelumnya.....	34
Tabel 3. 1 <i>Software</i> yang Digunakan	37
Tabel 3. 2 Jenis <i>Query-aware Parameter Selection</i>	46
Tabel 3. 3 Rentang Nilai Parameter	48
Tabel 3. 4 Kategori Kueri Pengujian.....	49
Tabel 3. 5 Contoh Kueri Pengujian.....	49
Tabel 3. 6 Skenario Pengujian	50
Tabel 4. 1 Rentang Waktu <i>Dataset</i>	57
Tabel 4. 2 Hasil Scraping Data	57
Tabel 4. 3 Contoh Hasil <i>Preprocessing</i> Dokumen.....	61
Tabel 4. 4 Alpha Terbaik <i>Query-aware Parameter Selection</i>	65
Tabel 4. 5 Nilai Paramter Terbaik <i>Time-Based Weighting</i>	66
Tabel 4. 6 Skor Kategori Leksikal	67
Tabel 4. 7 Skor Kategori Semantik.....	71
Tabel 4. 8 Skor Kategori Temporal.....	75
Tabel 4. 9 Perbandingan Setiap Kategori Metode <i>Retrieval</i>	79
Tabel 4. 10 Evaluasi Metrik Semua Kategori Kueri.....	81
Tabel 4. 11 Hasil Uji <i>Wicoxon Signed-Rank Test</i>	86

DAFTAR GAMBAR

Gambar 2. 1 Alur Kerja RAG	15
Gambar 3. 1 Konsep Penelitian.....	39
Gambar 3. 2 Alur Penelitian.....	40
Gambar 3. 3 Arsitektur Proses RAG <i>Chatbot</i>	44
Gambar 4. 1 Distribusi Sumber Dataset Penelitian.....	56
Gambar 4. 2 Distribusi Spasial Vektor	62
Gambar 4. 3 Frekuensi Kemunculan Kata	64



DAFTAR ISTILAH

<i>BM25 (Best Matching 25)</i>	Algoritma <i>lexical retrieval</i> berbasis <i>Probabilistic Retrieval Framework</i> yang mengukur relevansi dokumen berdasarkan kecocokan kata kunci, frekuensi term, dan panjang dokumen.
<i>Chatbot</i>	Sistem perangkat lunak yang berinteraksi dengan pengguna menggunakan bahasa alami melalui media teks maupun suara.
<i>Cosine Similarity</i>	Metode pengukuran kemiripan antara dua vektor berdasarkan sudut di antara keduanya, digunakan untuk membandingkan kedekatan makna representasi <i>embedding</i> .
<i>Dense Retrieval</i>	Pendekatan pencarian dokumen berbasis representasi vektor (<i>embedding</i>) untuk menangkap kemiripan semantik antara kueri dan dokumen.
<i>Embedding</i>	Representasi teks dalam bentuk vektor numerik berdimensi tinggi yang mempertahankan informasi semantik suatu kalimat atau dokumen.
<i>FAISS (Facebook AI Similarity Search)</i>	Pustaka pencarian kemiripan vektor yang dikembangkan oleh Meta AI, digunakan untuk pencarian <i>semantic retrieval</i> pada dataset berskala besar.
<i>Grid Search</i>	Metode optimasi parameter dengan menguji seluruh kombinasi kandidat nilai secara sistematis untuk memperoleh parameter dengan performa terbaik.

<i>Hallucination</i>	Kondisi ketika model bahasa menghasilkan informasi yang tampak benar namun tidak didukung oleh sumber pengetahuan yang valid.
<i>Hybrid Retrieval</i>	Pendekatan pencarian dokumen yang menggabungkan metode <i>lexical retrieval</i> (BM25) dan <i>semantic retrieval</i> (FAISS) untuk meningkatkan relevansi hasil pencarian.
<i>Large Language Model (LLM)</i>	Model bahasa berbasis arsitektur Transformer yang dilatih pada korpus teks berskala besar untuk memahami dan menghasilkan teks menyerupai bahasa manusia.
<i>Lexical Retrieval</i>	Pendekatan pencarian dokumen berdasarkan kecocokan kata kunci (<i>string matching</i>) antara kueri dan dokumen.
<i>Linear Score Fusion</i>	Mekanisme penggabungan skor pada metode <i>Information Retrieval</i>
<i>Mean Reciprocal Rank (MRR)</i>	Metrik evaluasi yang mengukur seberapa cepat sistem menemukan dokumen relevan pertama dalam hasil pencarian.
<i>Min-Max Normalization</i>	Teknik penyesuaian skala nilai (skor) ke dalam rentang [0,1] agar skor dari metode yang berbeda dapat dibandingkan atau digabungkan.
<i>Natural Language Processing (NLP)</i>	Bidang ilmu yang mempelajari pemrosesan dan pemahaman bahasa alami manusia oleh komputer.
<i>nDCG (normalized Discounted Cumulative Gain)</i>	Metrik evaluasi yang mengukur kualitas urutan peringkat dokumen relevan yang dihasilkan sistem <i>retrieval</i> .
<i>Precision@K</i>	Metrik evaluasi yang mengukur proporsi dokumen relevan di antara K dokumen teratas hasil pencarian.

Recall@K

Metrik evaluasi yang mengukur seberapa banyak dokumen relevan yang berhasil ditemukan dari keseluruhan dokumen relevan yang tersedia.

*Retrieval-Augmented
Generation (RAG)*

Pendekatan yang mengintegrasikan *information retrieval* dan *text generation* untuk meningkatkan kualitas jawaban model bahasa berdasarkan sumber eksternal.

*Rule-Based Query
Classification*

Pendekatan identifikasi karakteristik kueri berdasarkan aturan dan kata kunci tertentu, tanpa memerlukan model klasifikasi yang kompleks.

Sentence Embedding

Teknik representasi kalimat ke dalam bentuk vektor numerik yang mempertahankan makna semantik kalimat tersebut.

Time-Based Weighting

Mekanisme pembobotan skor dokumen berdasarkan keterbaruan tanggal publikasi, agar dokumen yang lebih baru mendapat prioritas lebih tinggi.

Vocabulary Mismatch

Kondisi ketika kueri dan dokumen relevan menggunakan kata yang berbeda meskipun memiliki makna yang sama, sehingga sulit dicocokkan secara leksikal.

Web Scraping

Teknik pengumpulan data secara otomatis dari halaman web menggunakan program komputer.

DAFTAR LAMPIRAN

Lampiran 1. Kartu Bimbingan Dosen Pembimbing 1

Lampiran 2. Kartu Bimbingan Dosen Pembimbing 2

Lampiran 3. Daftar Kueri Pengujian

