

DAFTAR ISI

LEMBAR PERSETUJUAN.....	i
LEMBAR PENGESAHAN	ii
SURAT PERNYATAAN KARYA SENDIRI.....	iii
LEMBAR PERSEMBAHAN	iv
ABSTRAK.....	v
<i>ABSTRACT</i>	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
DAFTAR PERSAMAAN	xiii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian	3
1.4. Batasan Masalah.....	4
1.5. Kerangka Pemikiran.....	4
1.6. Sistematika Penulisan	5
BAB II KAJIAN LITERATUR	7
2.1. Tinjauan Pustaka	7
2.2. Landasan Teori.....	10
2.2.1. Arsitektur <i>Two-Stages Retrieval (Retrieve & Rerank)</i>	10
2.2.2. Model Representasi Semantik: <i>Bi-Encoder</i> dan <i>Cross-Encoder</i> ..	11
2.2.3. <i>Composite Scoring</i>	13
2.2.4. <i>Intent Classification</i>	13
2.2.5. <i>Keyword Extraction</i>	15
2.2.6. <i>Term Frequency</i>	17
2.2.7. <i>Evaluation Metrics</i>	18
BAB III METODOLOGI PENELITIAN.....	20
3.1. <i>Intent Classification</i>	20
3.2. <i>Keyword Extraction</i>	22

3.3.	<i>Vector Database</i>	25
3.4.	<i>Composite Scoring</i>	25
3.5.	<i>Reranking</i>	28
3.6.	<i>Evaluation</i>	29
BAB IV HASIL DAN PEMBAHASAN		31
4.1.	<i>Composite Scoring Evaluation</i>	31
4.2.	<i>Reranking Evaluation</i>	33
4.3.	Diskusi	35
BAB V PENUTUP		38
5.1.	Kesimpulan	38
5.2.	Saran.....	38
DAFTAR PUSTAKA		40



DAFTAR GAMBAR

Gambar 2. 1 Arsitektur Two-Stages Retrieval	10
Gambar 2. 2 Arsitektur Cross-Encoder	11
Gambar 2. 3 Arsitektur Bi-Encoder	12
Gambar 3. 1 Alur kerja Sequran dimulai dengan memasukkan query pengguna, diikuti dengan Intent Classifier untuk mengidentifikasi intent pencarian. Dokumen yang relevan kemudian diambil menggunakan Sentence-Transformer berdasarkan kesamaan semantik. Proses berlanjut dengan Intent Booster dan analisis Term Frequency untuk meningkatkan relevansi, sebelum Tahap Perangkingan akhir untuk mengoptimalkan hasil.	20
Gambar 3. 2 Ilustrasi keyword extraction dari query pengguna menggunakan dua pendekatan: (1) POS tag dan NER berdasarkan intent query, dan (2) algoritma statistik atau kontekstual seperti Yake! atau KeyBERT. Keyword yang diekstrak digunakan dalam term frequency.	23



DAFTAR TABEL

Tabel 2. 1 State-of-the-art	7
Tabel 3. 1 Kategori intent dengan deskripsi, di mana hanya deskripsi yang divektorkan dan disimpan dalam basis data vektor untuk klasifikasi.	21
Tabel 3. 2 Pemetaan Tag dan Entitas yang diprioritaskan untuk setiap intent.	23
Tabel 4. 1 Evaluasi Composite Scoring menggunakan BM25S, BM25, dan Intent Booster, dengan atau tanpa Keyword Extraction, dinilai berdasarkan P@10, R@10, dan Waktu Eksekusi (s). Angka yang dicetak tebal menyoroti hasil yang lebih unggul.....	31
Tabel 4. 2 Evaluasi penilaian komposit berdasarkan Precision dan Recall pada @30 dan @100, dengan waktu eksekusi yang sesuai dengan Tabel 4. 1. Angka yang dicetak tebal menunjukkan hasil terbaik untuk setiap metrik.	33
Tabel 4. 3 Evaluasi reranking dari berbagai model cross-encoder, yang dinilai berdasarkan P@10, R@10, dan Waktu Eksekusi. Angka yang dicetak tebal menyoroti hasil yang penting untuk didiskusikan.....	33
Tabel 4. 4 Evaluasi reranking berdasarkan Precision dan Recall pada @30 dan @100, dengan waktu eksekusi yang sesuai dengan Tabel 4. 3. Angka yang dicetak tebal menyoroti hasil yang penting untuk didiskusikan.....	35
Tabel 4. 5 Perbandingan hasil antara sistem baseline dan sistem yang diusulkan untuk uji coba query. Teks yang dicetak tebal menunjukkan ayat-ayat yang paling relevan berdasarkan tafsir.....	36



DAFTAR PERSAMAAN

Persamaan 2. 1 Rumus Precision	18
Persamaan 2. 2 Rumus Recall	18
Persamaan 3. 1 Rumus Penilaian Akhir Composite Scoring	27
Persamaan 3. 2 Rumus Normalisasi	27

