

DAFTAR ISI

Halaman Judul	ii
Halaman Pengesahan	iii
Halaman Pernyataan	iv
Halaman Motto	v
PRAKATA	vi
INTISARI	xii
ABSTRACT	xiii
I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
II TINJAUAN PUSTAKA	5
III LANDASAN TEORI	9
3.1 Natural Language Processing	9
3.2 Arsitektur Transformer	9
3.3 Question Answering	11
3.4 Model BERT	12
3.5 Support Vector Machine	15
3.6 Model T5	15
3.7 Data Preprocessing	17
3.8 Evaluasi Kinerja model	17

3.8.1	F1-Score	18
3.8.2	Accuracy	18
3.8.3	BLEU	19
3.8.4	ROUGE-L	19
IV ANALISIS DAN PERANCANGAN SISTEM		20
4.1	Analisis Sistem	20
4.2	Pengumpulan Data	20
4.3	Perancangan Sistem	21
4.3.1	Pra-Pemrosesan Data	21
4.3.2	Perancangan Model	24
4.4	Pengujian	28
4.4.1	Evaluasi Tugas Klasifikasi Topik	28
4.4.2	Evaluasi Tugas Generasi Jawaban	29
V IMPLEMENTASI SISTEM		30
5.1	Lingkungan Penelitian	30
5.2	Implementasi Pengumpulan Data	30
5.2.1	Implementasi Ekstraksi Buku	30
5.2.2	Implementasi Ekstraksi Menggunakan OCR	32
5.3	Implementasi Prapemrosesan Data	33
5.3.1	Implementasi Menghilangkan Header & Footer	33
5.3.2	Implementasi Menghilangkan daftar isi	34
5.3.3	Implementasi Normalisasi Spasi	34
5.3.4	Implementasi Merge Dataset	35
5.3.5	Implementasi Labeling	35
5.3.6	Implementasi Splitting Data	36
5.4	Implementasi Model BERT	36
5.5	Implementasi Model T5	39
VI HASIL DAN PEMBAHASAN		43
6.1	Analisis Hyperparameter	43
6.1.1	Parameter Learning Rate mBERT	43
6.1.2	Parameter Num Beams T5	44

6.2	Hasil Pengujian	46
6.2.1	Hasil Metrik BERT	46
6.2.2	Hasil Perbandingan TF-IDF+SVM	47
6.2.3	Hasil Metrik T5	47
6.3	Pembahasan	51
6.3.1	Analisis Efektifitas Klasifikasi BERT	51
6.3.2	Analisis Kualitas Jawaban T5	52
6.3.3	Analisis Perbandingan mBERT & TF-IDF+SVM	53
VIIKESIMPULAN DAN SARAN		55
7.1	Kesimpulan	55
7.2	Saran	56