

4.1.1	Tahap Pengumpulan Data	52
4.1.1.1	Akun Instagram @dkijakarta	52
4.1.1.2	Akun Instagram @jktinfo	53
4.1.1.3	Penggabungan Dataset	55
4.1.2	Preprocessing Data	56
4.1.3	Translate Dataset	57
4.1.4	Pelabelan Dataset	60
4.1.5	Metode BERT	61
4.1.5.1	Mengonversi label kategori	61
4.1.5.2	Split Data	62
4.1.5.3	BERT Tokenizer dan Encoding Data	63
4.1.5.4	BERT Pretrained Model	65
4.1.5.5	Confusion matrix	68
4.1.6	Support Vector Machine (SVM)	69
4.1.6.1	TF-IDF	70
4.1.6.2	Klasifikasi dengan Support Vector Machine	71
4.1.6.3	Confusion Matrix	73
4.2	Pembahasan Hasil Penelitian	74
4.2.1	Evaluasi Metode BERT dan SVM	74
4.2.2	Hasil Analisis Sentimen	74
4.2.3	Perbandingan dengan Peneliti Lain	75
BAB 5 KESIMPULAN DAN SARAN		77
5.1	Kesimpulan	77
5.2	Saran	77
DAFTAR REFERENSI		79
Lampiran 1 Source Code		85

DAFTAR ISI

HALAMAN PERNYATAAN KEASLIAN	iii
HALAMAN PENGESAHAN TUGAS AKHIR	iv
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	v
KATA PENGANTAR	vi
ABSTRAK	vii
DAFTAR ISI	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
BAB 1 PENDAHULUAN	14
1.1 Latar Belakang	14
1.2 Identifikasi Masalah	16
1.3 Tujuan Tugas Akhir	16
1.4 Manfaat Tugas Akhir	16
1.5 Lingkup Tugas Akhir	16
1.6 Kerangka Berpikir	17
1.7 Sistematika Penulisan Tugas Akhir	18
BAB 2 TINJAUAN PUSTAKA	20
2.1 Penelitian Sebelumnya	20
2.2 Analisis Sentimen	27
2.3 Electronic Road Pricing (FRP)	30
2.4 Instagram	31
2.4.1 @dkijakarta	31
2.4.2 @jktinfo	31
2.5 Text Mining	31
2.6 Web Scraping	32
2.7 Preprocessing Data	32
2.7.1 Case Folding	32
2.7.2 Cleansing	33
2.7.3 Stopword Removal	33
2.7.4 Stemming	33
2.7.5 Slangwords	33

2.7.6	Tokenize	33
2.8	Translate	33
2.9	Lexicon Based Features	34
2.10	BERT Tokenizer	34
2.11	TF-IDF	34
2.12	Bidirectional Encoder Representations from Transformers (BERT)	36
2.13	Support Vector Machine (SVM)	37
2.14	Confusion Matrix	38
2.14.1	Akurasi	39
2.14.2	Precision	39
2.14.3	Recall	39
2.14.4	F-Measure	40
BAB 3 METODE PENELITIAN		41
3.1	Rencana Penelitian	41
3.2	Mencentukan Topik	42
3.3	Pengumpulan Data	42
3.3.1	Observasi	42
3.3.2	Studi Pustaka	42
3.3.3	Identifikasi Permasalahan	42
3.3.4	Objek Penelitian	44
3.3.5	Solusi yang diajukan	44
3.3.6	Pengumpulan Dataset	45
3.3.7	Preprocessing	45
3.3.8	Translate Data	47
3.3.9	Pelabelan Data	47
3.3.10	Tokenize	47
3.3.11	Split Data	47
3.3.12	BERT Tokenizer	47
3.3.13	TF-IDF	48
3.3.14	Klasifikasi Dengan Metode BERT dan SVM	48
3.3.15	Evaluasi Menggunakan Confusion Matrix	48
3.4	Bahan dan Peralatan	48
3.5	Jadwal Pelaksanaan Penelitian	49
BAB 4 HASIL DAN PEMBAHASAN		52
4.1	Data Hasil Penelitian	52