

DAFTAR ISI

HALAMAN PERSETUJUAN.....	ii
HALAMAN PERNYATAAN	iv
PRAKATA.....	v
DAFTAR ISI.....	vi
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	x
INTISARI.....	xii
ABSTRACT.....	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
1.6 Metodologi Penelitian	5
1.7 Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA.....	7
BAB III LANDASAN TEORI.....	11
3.1 Analisis Sentimen	11
3.2 Twitter.....	11
3.3 Text Preprocessing	12
3.4 Tokenisasi	13
3.5 DistilBERT Embedding	13
3.6 Self-training.....	16
3.7 LSTM	16
3.8 Confusion Matrix	19
3.9 J&T Express	24
BAB IV METODOLOGI PENELITIAN	25
4.1 Deskripsi Penelitian	25
4.2 Tahapan Penelitian	26

4.2.1	Pengumpulan data	26
4.2.2	<i>Text preprocessing</i>	27
4.2.3	Pemilahan data	28
4.2.4	Pembagian data	29
4.2.5	<i>Tokenisasi</i>	29
4.2.6	<i>Embedding</i>	31
4.2.7	<i>Self-training</i>	32
4.2.8	Pengujian.....	39
4.2.9	Evaluasi	40
BAB V IMPLEMENTASI PROGRAM		43
5.1	Pengumpulan data	43
5.2	Preprocessing data.....	44
5.2.1	Data cleaning dan sorting.....	45
5.2.2	Data <i>stemming</i>	47
5.2.3	Pemilihan data latih.....	47
5.2.4	Pelabelan data dengan BERT.....	48
5.3	Implementasi Prediksi	51
5.3.1	Tokenisasi	51
5.3.2	Embedding	52
5.3.3	<i>LSTM</i>	53
5.3.4	Self-training.....	60
5.3.5	Implementasi pelatihan	64
5.4	Implementasi Testing dan Evaluasi	70
5.4.1	Testing.....	70
5.4.2	Implementasi evaluasi.....	76
BAB VI HASIL DAN PEMBAHASAN		83
6.1	Hasil Preprocessing Data	83
6.1.1	Hasil <i>cleaning</i> dan <i>sorting</i>	83
6.1.2	Hasil <i>stemming</i> data	85
6.1.3	Hasil pelabelan awal	86
6.2	Hasil <i>Self-Training</i>	87
6.2.1	Dataset normalisasi teks	87
6.2.2	Dataset dengan penghilangan stopword.....	89
6.3	Hasil Evaluasi Dataset.....	93

6.3.1	Dataset normalisasi teks	94
6.3.2	Dataset penghilangan stopword	97
6.4	Pembahasan	99
6.4.1	Pengaruh ukuran batch dan tingkat pembelajaran	100
6.4.2	Dampak penghilangan stopword	100
6.4.3	Evaluasi pada dataset uji	101
BAB VII KESIMPULAN		102
7.1	Kesimpulan	102
7.2	Saran	103
DAFTAR PUSTAKA		104