



DAFTAR ISI

HALAMAN JUDUL	ii
HALAMAN PENGESAHAN	iii
PERNYATAAN BEBAS PLAGIASI	iv
PRAKATA	v
INTISARI	xii
ABSTRACT	xiii
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
II TINJAUAN PUSTAKA	5
III LANDASAN TEORI	9
3.1 Genre dan Sinopsis dalam Film	9
3.2 Klasifikasi Teks	9
3.3 <i>Deep Learning</i>	10
3.4 <i>Text Preprocessing</i>	11
3.4.1 <i>Case Folding</i>	11
3.4.2 <i>Remove Punctuation</i>	12
3.4.3 <i>Tokenization</i>	12
3.4.4 <i>Stopwords Removal</i>	12
3.4.5 <i>Stemming dan Lemmatization</i>	12
3.5 Ekstraksi Fitur	12
3.6 Word2Vec	13
3.7 <i>Long Short-Term Memory (LSTM)</i>	14
3.8 BERT	17



3.8.1	<i>Pre-training</i>	18
3.8.2	<i>Fine-tuning</i>	19
3.9	<i>Regularization</i>	19
3.10	Evaluasi Model	20
IV	METODE PENELITIAN	22
4.1	Tahapan Penelitian	22
4.2	Pengumpulan Dataset	23
4.3	<i>Preprocessing</i>	25
4.3.1	<i>Preprocessing</i> untuk Word2Vec	25
4.3.2	<i>Preprocessing</i> untuk BERT	30
4.4	Pemodelan	31
4.4.1	Word2Vec dan LSTM	31
4.4.2	BERT	32
4.5	Evaluasi	33
V	IMPLEMENTASI	35
5.1	Metode 1: Word2Vec + LSTM	35
5.1.1	<i>Import Library</i>	35
5.1.2	<i>Data Splitting</i>	36
5.1.3	<i>Preprocessing</i>	37
5.1.4	Model Word2Vec	37
5.1.5	Vektorisasi dan <i>Padding</i>	38
5.1.6	Model Klasifikasi	39
5.1.7	Evaluasi	41
5.2	Metode 2: BERT	42
5.2.1	<i>Import Library</i>	42
5.2.2	<i>Data Splitting</i>	42
5.2.3	<i>Input Formatting</i>	43
5.2.4	Pemodelan BERT	44
5.2.5	Evaluasi	45
VI	HASIL DAN PEMBAHASAN	47
6.1	Hasil Pengujian dengan Word2Vec + LSTM	47
6.2	Hasil Pengujian dengan BERT	52
6.3	Perbandingan Hasil Pengujian Kedua Metode	57



6.4	Analisis Hasil Klasifikasi	59
VIIKESIMPULAN DAN SARAN		65
7.1	Kesimpulan	65
7.2	Saran	65