

DAFTAR ISI

ABSTRAK	i
<i>ABSTRACT</i>	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN ORISINALITAS	iv
LEMBAR PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
DAFTAR LAMPIRAN	xiii
DAFTAR SIMBOL.....	xiv
DAFTAR ISTILAH	xv
Bab I PENDAHULUAN.....	1
I.1 Latar Belakang	1
I.2 Perumusan Masalah.....	3
I.3 Tujuan Penelitian.....	3
I.4 Batasan Penelitian	4
I.5 Manfaat Penelitian.....	4
I.6 Sistematika Penulisan Laporan	4
Bab II LANDASAN TEORI	7
II.1 Literatur	7
II.1.1 Natural Language Processing (NLP)	7
II.1.2 <i>Knowledge Discovery in Database</i> (KDD).....	8
II.1.3 <i>Text Mining</i>	10
II.1.4 Analisis Sentimen	11
II.1.5 Algoritma <i>Long Short-Term memory</i> (LSTM)	12
II.1.6 Google Play Store	15

II.1.7	<i>Web Scrapping</i>	15
II.1.8	<i>Labeling data</i>	16
II.1.9	<i>Preprocessing Data</i>	16
II.1.10	<i>Splitting Data</i>	18
II.1.11	<i>Word Embedding FastText</i>	19
II.1.12	<i>Confusion Matrix</i>	19
II.1.13	<i>Pemodelan Topik</i>	21
II.1.14	<i>Term Frequency - Inverse Document Frequency (TF-IDF)</i>	22
II.1.15	<i>Algoritma Non-Negative Matrix FaktORIZATION (NMF)</i>	24
II.1.16	<i>Topic Coherence</i>	25
II.1.17	<i>Word Cloud</i>	25
II.1.18	<i>Streamlit</i>	26
II.2	<i>Penelitian terdahulu</i>	27
II.3	<i>Perbandingan Algoritma</i>	30
Bab III	METODE PENYELESAIAN MASALAH	34
III.1	<i>Metode Penelitian</i>	34
III.2	<i>Sistematika Penyelesaian Masalah</i>	35
III.2.1	<i>Data Selection</i>	36
III.2.2	<i>Preprocessing</i>	36
III.2.3	<i>Transformation</i>	37
III.2.4	<i>Data Mining</i>	38
III.2.5	<i>Evaluation</i>	38
III.2.6	<i>Knowlegde Presentation</i>	39
III.3	<i>Alasan Pemilihan Algoritma</i>	39
Bab IV	PENYELESAIAN MASALAH	41
IV.1	<i>Data Selection</i>	41
IV.1.1	<i>Identifikasi Kebutuhan Data</i>	41

IV.1.2	Pelabelan data.....	42
IV.2	<i>Preprocessing Data</i>	43
IV.2.1	<i>Cleaning Data</i>	43
IV.2.2	<i>Case Folding</i>	44
IV.2.3	<i>Normalize</i>	44
IV.2.4	<i>Tokenization</i>	45
IV.2.5	<i>Stopword Removal</i>	45
IV.2.6	<i>Stemming</i>	46
IV.3	<i>Label Encoding</i>	47
IV.4	<i>Word Embedding FastText</i>	47
IV.5	TF-IDF	48
IV.6	<i>Splitting Data Training dan Data Testing</i>	51
IV.7	Perancangan Model.....	51
IV.7.1	Perancangan Model LSTM	51
IV.7.2	Perancangan Model NMF	54
IV.8	Evaluasi.....	57
IV.8.1	<i>Confusion matrix</i>	57
IV.8.2	<i>Topic Coherence</i>	57
IV.9	<i>Knowledge Presentation</i>	57
Bab V	VALIDASI, ANALISIS HASIL DAN IMPLIKASI	59
V.1	Pengujian	59
V.2	<i>Confusion matrix</i>	60
V.3	<i>Topic Coherence</i>	62
V.4	Analisis Sentimen.....	64
V.5	Analisis topik Non-Negative Matrix Factorization	65
V.6	Penerapan Streamlit untuk Prediksi	75

V.7	Rekomendasi Terhadap Sekolah.mu	75
V.8	Implikasi	76
Bab VI	KESIMPULAN DAN SARAN	77
VI.1	Kesimpulan	77
VI.2	Saran	78
Bab VII	DAFTAR PUSTAKA	79
Lampiran		85