

DAFTAR ISI

Abstrak	i
<i>Abstract</i>	ii
Lembar Pengesahan	iii
Lembar Pernyataan Orisinalitas	iv
Kata Pengantar	v
Daftar Isi.....	vi
Daftar Tabel	x
Daftar Gambar.....	xii
Bab I Pendahuluan	1
I.1 Latar Belakang	1
I.2 Perumusan Masalah.....	6
I.3 Tujuan Penelitian.....	6
I.4 Batasan Penelitian	7
I.5 Manfaat Penelitian.....	7
I.6 Sistematika Laporan	7
Bab II Landasan Teori	9
II.1 Info BMKG	9
II.2 <i>Delone & McLean IS Success Model</i>	9
II.3 <i>Knowledge Discovery in Databases</i>	10
II.4 <i>Data Mining</i>	11
II.5 <i>Text Mining</i>	12
II.6 <i>Natural Language Processing</i>	12
II.7 <i>Text Preprocessing</i>	13
II.8 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	14
II.9 <i>Oversampling</i>	15

II.10	<i>Labelling</i>	16
II.11	<i>Cross Validation</i>	16
II.12	<i>Analisis Sentimen</i>	17
II.13	<i>Machine Learning</i>	18
II.14	<i>Ensemble Learning</i>	19
II.15	<i>Gradient Boosted Decision Tree</i>	20
II.16	<i>Decision Tree</i>	22
II.17	<i>CatBoost</i>	23
II.18	<i>Hyperparameter Tuning</i>	25
II.19	<i>Metrik Evaluasi</i>	27
II.20	<i>Topic Modelling</i>	28
II.21	<i>BERTopic</i>	29
II.22	<i>Topic Coherence</i>	30
II.23	<i>Penelitian Terdahulu</i>	32
II.24	<i>Alasan Pemilihan Teori, Kerangka Kerja, atau Mekanisme</i>	34
Bab III	<i>Metode Penyelesaian Masalah</i>	37
III.1	<i>Metode Penelitian</i>	37
III.2	<i>Sistematika Penyelesaian Masalah</i>	38
III.2.1	<i>Initialization</i>	39
III.2.2	<i>Pengumpulan Data (Data Selection)</i>	39
III.2.3	<i>Pengolahan Data (Data Preprocessing)</i>	39
III.2.4	<i>Data Transformation</i>	39
III.2.5	<i>Data Mining</i>	40
III.2.6	<i>Evaluation</i>	40
Bab IV	<i>Penyelesaian Permasalahan</i>	41
IV.1	<i>Initialization</i>	41

IV.2	<i>Data Selection</i>	42
IV.3	<i>Data Preprocessing</i>	44
IV.3.1	<i>Labelling</i>	44
IV.3.2	<i>Case Folding</i>	48
IV.3.3	Penghapusan Tanda Baca (<i>Punctuation</i>)	50
IV.3.4	Penghapusan Emoji dan Simbol	51
IV.3.5	<i>Text Normalization</i>	53
IV.3.6	Penghapusan Angka	56
IV.3.7	<i>Removing Stopword</i>	58
IV.4	<i>Data Transformation</i>	60
IV.4.1	<i>Data Splitting</i>	60
IV.4.2	<i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	62
IV.4.3	<i>Oversampling</i>	65
IV.5	<i>Data Mining</i>	65
IV.5.1	Klasifikasi Topik CatBoost	66
IV.5.2	Klasifikasi Sentimen CatBoost	67
IV.5.3	<i>Topic Modelling BERTopic</i>	69
IV.6	<i>Interpretation and Evaluation</i>	70
IV.6.1	<i>Classification Report</i>	70
IV.6.2	<i>Confusion Matrix</i>	71
IV.6.3	<i>Hierarchical Clustering</i>	72
Bab V	Validasi, Analisis Hasil, dan Implikasi	74
V.1	Validasi dan Analisis Hasil	74
V.1.1	<i>Hyperparameter Tuning Model Klasifikasi Topik</i>	74
V.1.2	Evaluasi Klasifikasi Topik <i>System Quality</i>	75
V.1.3	Evaluasi Klasifikasi Topik <i>Information Quality</i>	79

V.1.4	Evaluasi Klasifikasi Sentimen Topik <i>System Quality</i>	83
V.1.5	Evaluasi Klasifikasi Sentimen Topik <i>Information Quality</i>	88
V.2	Analisis Hasil Subtopik	92
V.2.1	Subtopik <i>System Quality</i> Sentimen Negatif	92
V.2.2	Subtopik <i>System Quality</i> Sentimen Positif	96
V.2.3	Subtopik <i>Information Quality</i> Sentimen Negatif	98
V.2.4	Subtopik <i>Information Quality</i> Sentimen Positif	101
V.3	Implikasi Tugas Akhir	106
Bab VI	Kesimpulan dan Saran	107
VI.1	Kesimpulan	107
VI.2	Saran	108
Daftar Pustaka		109