

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN.....	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
DAFTAR RIWAYAT HIDUP	v
HALAMAN PERSEMBAHAN	vi
MOTTO.....	viii
DAFTAR ISI	ix
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xv
DAFTAR PERSAMAAN	xvii
KATA PENGANTAR.....	xviii
INTISARI	xx
ABSTRACT	xxi
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	2
1.3 Batasan Penelitian	2
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Sistem Penulisan.....	4
BAB II TINJAUAN PUSTAKA	6

2.1 <i>Machine Learning</i>	6
2.2 <i>Algoritma Naïve Bayes</i>	8
2.2.1 <i>Multinomial Naïve Bayes</i>	9
2.2.2 <i>Gaussian Naïve Bayes</i>	9
2.2.3 <i>Bernoulli Naïve Bayes</i>	10
2.3 <i>Term Frequency – Inverse Document Frequency (TFIDF)</i>	11
2.4 <i>Confusion Matrix</i>	12
2.5 <i>Kurva ROC</i>	14
2.6 <i>Cross Validation</i>	14
2.7 <i>Software Yang Digunakan untuk Pembuatan Model Klasifikasi</i>	15
2.7.1 <i>Anaconda Navigator</i>	15
2.7.2 <i>Jupyter Notebook</i>	15
2.7.3 <i>Web Scraper</i>	16
2.8 <i>Metode Pengembangan Model Klasifikasi</i>	16
2.8.1 <i>Machine Learning Life Cycle</i>	16
2.9 <i>Penelitian Terkait</i>	19
BAB III <i>METODOLOGI PENELITIAN</i>	37
3.1 <i>Metode Pengembangan Life Cycle</i>	37
3.1.1 <i>Model Requirements</i>	37
3.1.2 <i>Data Collection</i>	38
3.1.3 <i>Data Cleaning</i>	44
3.1.4 <i>Data Labeling</i>	47
3.1.5 <i>Feature Engineering</i>	49
3.1.6 <i>Model Training</i>	51

3.1.7 <i>Model Evaluation</i>	54
3.1.8 <i>Operation</i>	54
BAB IV HASIL DAN PEMBAHASAN.....	56
4.1 <i>Preprocessing Data</i>	56
4.1.1 <i>Hasil Data Collection</i>	56
4.1.2 <i>Hasil Data Cleaning</i>	62
4.1.3 <i>Hasil Data Labeling</i>	117
4.2 <i>Hasil Feature Engineering</i>	125
4.2.1 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	126
4.2.2 <i>Synthetic Minority Over-Sampling (SMOTE)</i>	132
4.3 <i>Hasil Model Training</i>	133
4.3.1 <i>Model Selection</i>	134
4.3.2 <i>Grid Search</i>	136
4.3.3 <i>Model Tuning Hyperparameter</i>	137
4.3.4 <i>Report Model Classification</i>	138
4.4 <i>Hasil Model Evaluation</i>	155
4.4.1 <i>Analisa Pendapat Ahli Bahasa Indonesia</i>	155
4.4.2 <i>Perbandingan Hasil Pendapat Ahli Bahasa Dengan Model Klasifikasi Multinomial</i>	157
4.5 <i>Operation</i>	159
4.5.1 <i>Environment</i>	159
4.5.2 <i>Deployment</i>	160
4.5.3 <i>Monitoring atau updating</i>	166
4.6 <i>Kelemahan Sistem</i>	167

BAB V KESIMPULAN DAN SARAN	169
5.1 Kesimpulan.....	169
5.2 Saran	171
DAFTAR PUSTAKA.....	173
LAMPIRAN	179