

DAFTAR ISI

PENGESAHAN	ii
PERNYATAAN BEBAS PLAGIASI.....	iii
PRAKATA.....	iv
ARTI LAMBANG DAN SINGKATAN	vi
<i>ABSTRACT</i>	vii
INTISARI.....	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan masalah.....	6
1.3 Keaslian penelitian.....	6
1.4 Tujuan Penelitian	10
1.5 Manfaat Penelitian	10
1.6 Batasan Penelitian.....	10
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	11
2.1 Tinjauan Pustaka	11
2.2 Landasan Teori	13
2.2.1 <i>Natural Language Processing</i>	13
2.2.2 Synthetic Minority Over-sampling Technique (SMOTE).....	15
2.2.3 Term Frequency-Inverse Document Frequency	16
2.2.4 Klasifikasi	19
2.2.5 Pra-proses.....	21
2.2.6 Metode Klasifikasi <i>Naïve Bayes</i>	21
2.2.8 Metode Klasifikasi <i>Logistic Regression</i> (LR).....	23
2.2.9 <i>Split Validation</i>	24
2.3 Pertanyaan Penelitian.....	25
BAB III METODOLOGI	26
3.1 Alat dan Bahan.....	26
3.1.1 Alat.....	26
3.1.2 Bahan.....	27
3.2 Jalannya Penelitian.....	28
3.2.1 Pengumpulan Data	29
3.2.2 Pra-proses Teks	29
3.2.2.1 Stopword Removal.....	31
3.2.2.2 Stemming	31
3.2.2.3 Tokenisasi.....	32

3.2.3	Term Frequency-Inverse Document Frequency	33
3.2.4	SMOTE	34
3.2.5	Klasifikasi	35
3.2.6	Evaluasi	35
3.3	Cara Analisis	37
BAB IV HASIL DAN PEMBAHASAN.....		38
4.1	Hasil Pengumpulan Data	38
4.2	Konversi Format <i>File</i> Text ke CSV	38
4.3	Hasil Pra-Proses Teks.....	39
4.4	Hasil Tokenisasi Kata Unik dengan TF-IDF.....	41
4.5	Hasil Evaluasi Prediksi Pada Model	42
4.5.1	Akurasi	43
4.5.2	<i>F1-Score</i>	44
4.5.3	Waktu Komputasi.....	45
4.5.4	Perbandingan Hasil Evaluasi <i>Hybrid Classifier</i> dengan Penelitian Sebelumnya.....	48
4.5.5	Perbandingan Hasil SMOTE dengan sebelum SMOTE	49
4.5.6	Hasil Klasifikasi Error pada Prediksi	49
BAB V KESIMPULAN DAN SARAN.....		51
5.1	Kesimpulan	51
5.2	Saran	51
DAFTAR PUSTAKA		52
LAMPIRAN.....		L1
Lampiran 1	Dataset	L1