

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN	iii
HALAMAN PERSEMBAHAN	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN.....	xii
INTISARI.....	xiv
ABSTRACT.....	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Pembatasan Masalah	4
1.3 Tujuan Penulisan.....	5
1.4 Tinjauan Pustaka.....	5
1.5 Metode Penulisan.....	7
1.6 Sistematika Penulisan	8
BAB II LANDASAN TEORI.....	9
2.1 Matriks	9
2.2 Distribusi Variabel Random Bernoulli	12
2.3 Peluang.....	13
2.3.1 Peluang bersyarat.....	13
2.3.2 Kejadian bebas atau independen	14
2.4 Aturan Bayes.....	15
2.5 Maximum a Posteriori.....	16
2.6 Decision Tree	17
2.6.1 Karakteristik decision tree	18
2.6.2 Proses pembentukan decision tree.....	19

2.7	Machine Learning	20
2.8	Supervised dan Unsupervised Learning.....	21
2.9	Text Mining	22
2.10	Klasifikasi	23
2.11	<i>Imbalanced Data</i>	24
2.12	<i>Website</i>	25
2.13	<i>Web Scraping</i>	26
2.14	Parsehub.....	26
2.15	<i>E-commerce</i>	27
2.16	<i>Preprocessing Data</i>	29
2.17	<i>Word Cloud</i>	30
2.18	Evaluasi Performa Model	31
2.19	Diagram Sebab Akibat.....	33
BAB II IMPLEMENTASI <i>NAÏVE BAYES</i> DAN <i>RANDOM FOREST</i> UNTUK ANALISIS SENTIMEN TERHADAP DATA <i>IMBALANCED REVIEW</i> PRODUK KOSMETIK PADA <i>PLATFORM ONLINE SOCIOLLA</i>		
		35
3.1	Analisis Sentimen	35
3.2	Metode <i>Naïve Bayes</i>	35
3.2.1	Model probabilitas untuk <i>naïve bayes</i>	36
3.2.2	Laplace correction.....	38
3.2.3	Model <i>naïve bayes</i> untuk klasifikasi	38
3.2.4	Nilai <i>likelihood</i> distribusi Bernoulli.....	42
3.2.5	Simulasi klasifikasi dengan metode <i>naïve bayes</i>	46
3.3	<i>Random Forest</i>	53
3.3.1	<i>Classification and Regression Tree (CART)</i>	56
3.3.2	Algoritma <i>random forest</i>	57
3.3.3	Kelebihan dan kekurangan <i>random forest</i>	59
3.4	Teknik Sampling Kelas <i>Imbalanced</i>	60
BAB IV STUDI KASUS DATA <i>IMBALANCED REVIEW</i> PRODUK KOSMETIK PADA <i>PLATFORM ONLINE SOCIOLLA</i>		
		62
4.1	Deskripsi Kasus	62

4.2	Pengumpulan Data dengan <i>Web Scraping</i>	62
4.3	Analisis Statistika Deskriptif	64
4.4	<i>Preprocessing Data</i>	69
4.4.1	<i>Case Folding</i>	70
4.4.2	<i>Tokenizing</i>	71
4.4.3	<i>Filtering</i>	71
4.5	Pelabelan Kelas Sentimen	73
4.6	Pembentukan Data <i>Training</i> dan Data <i>Testing</i>	75
4.7	Metode <i>Naïve Bayes</i>	76
4.8	Metode <i>Random Forest</i>	79
4.9	Evaluasi Performa Model Klasifikasi	81
4.10	<i>Naïve Bayes</i> dan <i>Random Forest</i> dengan Teknik <i>Undersampling</i>	84
4.11	Perbandingan Metode Klasifikasi <i>Naïve Bayes</i> , <i>Random Forest</i> , <i>Naïve Bayes Undersampling</i> dan <i>Random Forest Undersampling</i>	87
4.12	Visualisasi Kata	88
4.13	Diagram Sebab Akibat	91
BAB V PENUTUP		95
5.1	Kesimpulan	95
5.2	Saran	96
DAFTAR PUSTAKA		97
LAMPIRAN		99