

## DAFTAR ISI

PENGESAHAN . . . . .	ii
PERNYATAAN BEBAS PLAGIASI . . . . .	iii
PRAKATA . . . . .	v
ARTI LAMBANG DAN SINGKATAN . . . . .	vi
ABSTRACT . . . . .	vii
INTISARI . . . . .	viii
DAFTAR ISI . . . . .	ix
DAFTAR GAMBAR . . . . .	xi
DAFTAR TABEL . . . . .	xiii
<b>BAB I PENDAHULUAN . . . . .</b>	<b>1</b>
1.1 Latar Belakang . . . . .	1
1.2 Perumusan Masalah . . . . .	4
1.3 Keaslian Penelitian . . . . .	4
1.4 Tujuan Penelitian . . . . .	6
1.5 Batasan Penelitian . . . . .	6
1.6 Manfaat Penelitian . . . . .	7
<b>BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI . . . . .</b>	<b>8</b>
2.1 Tinjauan Pustaka . . . . .	8
2.2 Landasan Teori . . . . .	9
2.2.1 <i>Batch Processing</i> dan <i>Stream Processing</i> . . . . .	9
2.2.2 <i>Machine Learning</i> . . . . .	10
2.2.2.1 Klasifikasi dengan algoritme <i>Naive Bayes</i> . . . . .	14
2.2.3 Sistem Klasifikasi Terdistribusi . . . . .	15
2.2.3.1 <i>Ensemble</i> . . . . .	16
2.2.3.2 Voting . . . . .	18
2.2.4 <i>RabbitMQ Message Broker</i> . . . . .	20
2.2.4.1 Jenis Arsitektur RabbitMQ . . . . .	22
2.2.5 <i>ELK Stack</i> . . . . .	23
2.3 Pertanyaan Penelitian . . . . .	25
<b>BAB III METODOLOGI . . . . .</b>	<b>26</b>
3.1 Alat dan Bahan . . . . .	27
3.1.1 Alat . . . . .	27

3.1.1.1	Perangkat Keras . . . . .	27
3.1.1.2	Perangkat Lunak . . . . .	27
3.1.2	Bahan . . . . .	28
3.2	Jalannya Penelitian . . . . .	29
3.2.1	<i>Data Labelling</i> . . . . .	31
3.2.2	<i>Balancing</i> Data Latih . . . . .	32
3.2.3	Arsitektur Sistem Klasifikasi Terdistribusi . . . . .	34
3.3	Cara Analisis . . . . .	36
3.3.1	Evaluasi Klasifikasi SDGs . . . . .	36
3.3.2	Pengukuran Waktu Proses dan Tingkat Pemakaian CPU . . . . .	39
<b>BAB IV HASIL DAN PEMBAHASAN . . . . .</b>		<b>40</b>
4.1	Pre-processing . . . . .	40
4.2	Klasifikasi . . . . .	42
4.2.1	Model NBC tanpa <i>Pre-processing</i> . . . . .	43
4.2.2	Model NBC dengan <i>Pre-processing Lowercase</i> . . . . .	45
4.2.3	Model NBC dengan <i>Pre-processing Lowercase</i> dan <i>tuning</i> parameter <i>N-Gram</i> . . . . .	47
4.3	Agregasi dan Visualisasi Data . . . . .	50
4.4	Pengujian . . . . .	52
4.4.1	Waktu Pemrosesan . . . . .	52
4.4.2	Pemakaian CPU . . . . .	54
4.5	Pembahasan . . . . .	55
4.5.1	Keunggulan Sistem . . . . .	55
4.5.2	Kelemahan Sistem . . . . .	56
<b>BAB V KESIMPULAN DAN SARAN . . . . .</b>		<b>57</b>
5.1	Kesimpulan . . . . .	57
5.2	Saran . . . . .	57
<b>DAFTAR PUSTAKA . . . . .</b>		<b>59</b>
<b>LAMPIRAN . . . . .</b>		<b>L-1</b>
L.1	Data Latih . . . . .	L-1