

DAFTAR ISI

SKRIPSI	1
SKRIPSI	2
HALAMAN PENGESAHAN	3
PERNYATAAN	1
PRAKATA	2
DAFTAR ISI	3
DAFTAR TABEL	4
DAFTAR GAMBAR	5
ABSTRAK	6
ABSTRACT	7
BAB I	
PENDAHULUAN	8
1.1 Latar Belakang	8
1.2 Rumusan Masalah	10
1.3 Batasan Penelitian	10
1.4 Tujuan Penelitian	11
1.5 Manfaat Penelitian	11
BAB II	
TINJAUAN PUSTAKA	12
BAB III	
LANDASAN TEORI	19
3.1. Multi-Label Classification	19
3.2. Naive Bayes	20
3.3. TF-IDF	28
3.4. Preprocessing	30
3.5. Evaluasi Performa	31
3.5.1. Exact Match Ratio	31
3.5.2. 0/1 Loss	32
3.5.3. Hamming Score/Hamming Loss	32
3.5.4. F1 Score	34
BAB IV	
METODOLOGI PENELITIAN	36
4.1. Deskripsi Penelitian Umum	36
4.2. Desain dan Tahap Penelitian	36
4.2.1. Akuisisi Data	36

4.2.2. Preprocessing Data	37
4.2.2.1. Data Cleaning	37
4.2.2.2. Case folding	38
4.2.2.3. Tokenisasi	38
4.2.2.4. Spell Correction	38
4.2.2.5. Stemming	39
4.2.2.6. Stopwords removal	39
4.2.3. Pelabelan Data	39
4.2.4. Desain Model	44
4.2.4.1. Multinomial Naive Bayes Classifier	44
4.2.4.2. Complement Naive Bayes Classifier	48
4.2.5. TF-IDF	50
4.2.6. Skenario Eksperimen	50
4.2.7. Klasifikasi	51
4.2.8. Evaluasi Performa	51
BAB V	
IMPLEMENTASI	52
5.1. Pembuatan Dataset	52
5.2. Pelabelan Data	53
5.3. Preprocessing Data	53
5.2.1. Data Cleaning	53
5.2.2. Case Folding	55
5.2.3. Spell Correction	55
5.2.4. Stemming	56
5.2.5. Stopwords Removal	56
5.2.6. Tokenisasi	56
5.4. Klasifikasi Multilabel	57
5.5. Skenario Eksperimen	58
5.6. Evaluasi Performa	59
BAB VI	
HASIL DAN PEMBAHASAN	61
6.1. Hasil Pengujian Model	61
6.1.1. Multinomial Naive Bayes	61
6.1.2. Complement Naive Bayes	62
6.2. Analisis Performa	64
BAB VII	
KESIMPULAN DAN SARAN	69

7.1. Kesimpulan	69
7.2. Saran	69
DAFTAR PUSTAKA	71
LAMPIRAN SOURCE CODE	76
1. Mengumpulkan Dataset	76
2. Mencocokkan Data pada Dataframe	76
3. Menyiapkan Modul yang Diperlukan untuk Preprocessing	77
4. Preprocessing	78
5. Proses Training dan Testing MultinomialNB	79
6. Evaluasi MultinomialNB	80
7. Proses Training dan Testing ComplementNB	81
8. Evaluasi ComplementNB	83

DAFTAR TABEL

Tabel 2.1 Riset internasional dalam klasifikasi SDG	11
Tabel 2.2 Perbandingan penelitian klasifikasi multi-label dan SDG	11
Tabel 3.1 Tabel Faktor Kecelakaan dan Kemunculannya	16
Tabel 3.2 <i>Likelihood</i> Cuaca terhadap kecelakaan	16
Tabel 3.3 <i>Likelihood</i> Kondisi Jalan terhadap kecelakaan	16
Tabel 3.4 <i>Likelihood</i> Kemacetan terhadap kecelakaan	16
Tabel 3.5 Contoh Mencari TF-IDF	16
Tabel 3.6 Teknik Preprocessing pada Data Sosial Media	16
Tabel 3.7 Perbandingan hasil prediksi dan <i>ground truth</i>	16
Tabel 3.8 Skema <i>Confusion Matrix</i>	16
Tabel 4.1 Contoh keyword pelabelan	19
Tabel 4.2 Contoh tweet untuk pelabelan data	19
Tabel 4.3 Contoh Data Klasifikasi SDG	19
Tabel 4.4 <i>Likelihood</i> Multinomial Kemiskinan	19
Tabel 4.5 <i>Likelihood</i> Multinomial Pendidikan	19
Tabel 4.6 <i>Likelihood</i> Multinomial Kesenjangan	19

Tabel 4.7 Algoritma MNB	19
Tabel 4.8 <i>Likelihood</i> Complement Kemiskinan	19
Tabel 4.9 <i>Likelihood</i> Complement Pendidikan	19
Tabel 4.10 <i>Likelihood</i> Complement Kesenjangan	19
Tabel 4.11 Algoritma CNB	19
Tabel 4.12 Skenario Eksperimen	19
Tabel 5.1 Struktur Dataset	19
Tabel 5.2 Penjelasan Regex	19
Tabel 5.3 Cuplikan Tabel Kamus Bahasa Gaul	19
Tabel 5.4 Tabel Skenario Eksperimen	19
Tabel 6.1 Evaluasi Performa Multinomial Naive Bayes	19
Tabel 6.2 Evaluasi Performa Complement Naive Bayes	19

DAFTAR GAMBAR

Gambar 4.1 Keseluruhan proses penelitian	12
Gambar 4.2 Proses akuisisi data	12
Gambar 4.3 Diagram <i>Preprocessing</i>	12
Gambar 4.4 Proses <i>Data Cleaning</i>	12
Gambar 4.5 Proses <i>Case Folding</i>	12
Gambar 4.6 Proses <i>Spell Correction</i>	12
Gambar 4.7 Proses <i>Stemming</i>	12
Gambar 4.8 Proses <i>Stopwords Removal</i>	12
Gambar 5.1 Contoh Penggunaan Regex	12
Gambar 5.2 Diagram Proses Klasifikasi	12
Gambar 6.1 Grafik Perbandingan EMR untuk MNB dan CNB	12
Gambar 6.2 Grafik Perbandingan <i>hamming loss</i> untuk MNB dan CNB	12
Gambar 6.3 Grafik Perbandingan <i>F1-Score</i> untuk MNB dan CNB	12
Gambar 6.4 Grafik Perbandingan Nilai-nilai <i>Confusion Matrix</i> untuk MNB dan	

CNB	12
Gambar 6.5 <i>Wordcloud</i> SDG 1-8	12
Gambar 6.6 <i>Wordcloud</i> SDG 9-17	12
Gambar 6.7 Confusion Matrix setiap SDG	12