

## DAFTAR ISI

<b>LEMBAR PERSETUJUAN</b>	
<b>LEMBAR PENGESAHAN</b>	
<b>LEMBAR PERNYATAAN KARYA SENDIRI</b>	
<b>LEMBAR PERSEMBAHAN</b>	
<b>ABSTRAK</b> .....	i
<b>ABSTRACT</b> .....	ii
<b>KATA PENGANTAR</b> .....	iii
<b>DAFTAR ISI</b> .....	v
<b>DAFTAR TABEL</b> .....	viii
<b>DAFTAR GAMBAR</b> .....	ix
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian .....	3
1.4 Batasan Masalah .....	4
1.5 Kerangka Pemikiran.....	5
1.6 Metodologi Penelitian .....	6
1.7 Sistematika Penulisan .....	6
<b>BAB II KAJIAN LITERATUR</b> .....	8
2.1 Tinjauan Pustaka.....	8
2.2 Landasan Teori.....	11
2.2.1 <i>Text Mining</i> .....	11
2.2.2 Twitter.....	11
2.2.4 Rokok Elektrik ( <i>Vape</i> ).....	12
2.2.5 <i>Naive Bayes Classifier</i> .....	12
2.2.7 <i>Google Colaboratory</i> .....	14
<b>BAB III METODOLOGI PENELITIAN</b> .....	18

3.1	<i>Business Understanding</i> .....	18
3.2	<i>Data Understanding</i> .....	18
3.3	<i>Data Preparation</i> .....	19
3.3.1	<i>Labeling</i> .....	20
3.3.2	<i>Tokenizing</i> .....	20
3.3.3	<i>Casefolding</i> .....	21
3.3.4	<i>Regex Removing</i> .....	21
3.3.5	<i>Stopword</i> .....	22
3.3.6	<i>Stemming</i> .....	22
3.3.7	Pembobotan Kata (TF-IDF).....	22
3.4	<i>Modeling</i> .....	24
3.4.1	Memilih Teknik Pemodelan.....	25
3.4.2	Membangun Model.....	27
3.5	<i>Evaluation</i> .....	29
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....		30
4.1	Hasil <i>Crawling Data</i> .....	30
4.2	Hasil <i>Data Preparation</i> .....	30
4.2.1	Hasil <i>Labeling</i> .....	30
4.2.2	Hasil <i>Tokenizing</i> .....	31
4.2.3	Hasil <i>Casefolding dan Regex Removing</i> .....	32
4.2.4	Hasil <i>Stopword</i> .....	32
4.2.5	Hasil <i>Stemming</i> .....	33
4.2.6	Hasil Pembobotan Kata (TF-IDF) .....	33
4.3	Hasil <i>Modeling</i> .....	34
4.4	Pembahasan.....	34
4.4.1	Pengujian Algoritma Terhadap Data Latih .....	35
4.5	Prediksi Sentimen .....	37
<b>BAB V SIMPULAN DAN SARAN</b> .....		38
5.1	Simpulan .....	38

5.2	Saran .....	38
<b>DAFTAR PUSTAKA</b> .....		<b>39</b>
<b>Lampiran</b> .....		<b>42</b>
<b>DAFTAR RIWAYAT HIDUP</b> .....		<b>54</b>



## DAFTAR TABEL

Tabel 2. 1 <i>The State of the Art</i> .....	8
Tabel 3. 1 Tabel <i>Tokenizing Data</i> .....	20
Tabel 3. 2 Tabel <i>Casefolding</i> .....	21
Tabel 3. 3 <i>Regex Removing</i> .....	21
Tabel 3. 4 <i>Stopword</i> .....	22
Tabel 3. 5 <i>Stemming</i> .....	22
Tabel 3. 6 <i>Term Frequency</i> .....	23
Tabel 3. 7 <i>Term Frequency Normalisasi</i> .....	23
Tabel 3. 8 <i>Document Frequency dan Inverse Document Frequency</i> .....	23
Tabel 3. 9 TF-IDF .....	24
Tabel 3. 10 Contoh Data Latih.....	27
Tabel 3. 11 Perhitungan <i>Naïve Bayes</i> .....	28
Tabel 3. 12 Data Uji.....	28
Tabel 3. 13 Model Probabilistik.....	29



## DAFTAR GAMBAR

Gambar 1. 1 Kerangka Pemikiran.....	5
Gambar 2. 1 CRISP DM. ....	16
Gambar 3. 1 Jumlah Label Positif dan Negatif.....	20
Gambar 3. 2 Arsitektur Sistem.....	25
Gambar 3. 3 <i>Flowchart Naïve Bayes</i> .....	26
Gambar 4. 1 Hasil <i>Crawling</i> .....	30
Gambar 4. 2 Hasil <i>Labelling</i> .....	31
Gambar 4. 3 Hasil <i>Tokenizing</i> .....	31
Gambar 4. 4 <i>Casefolding dan Regex Removing</i> .....	32
Gambar 4. 5 Hasil <i>Stopword</i> .....	33
Gambar 4. 6 Hasil <i>Stemming</i> .....	33
Gambar 4. 7 Hasil Pembobotan Kata (TF-IDF).....	34
Gambar 4. 8 Hasil <i>Modeling</i> .....	34
Gambar 4. 9 <i>Confusion Matrix</i> .....	35
Gambar 4. 10 <i>Classification Report</i> .....	36
Gambar 4. 11 Hasil Analisis Sentimen Terhadap Rokok Elektrik .....	37

