

## DAFTAR ISI

HALAMAN SAMPUL	
HALAMAN JUDUL	
PERNYATAAN KEASLIAN DAN PERSETUJUAN UNGGAH TUGAS AKHIR .....	i
PERSETUJUAN DOSEN PEMBIMBING TUGAS AKHIR .....	ii
PERSETUJUAN TIM PENGUJI TUGAS AKHIR.....	iii
ABSTRAK .....	iv
<i>ABSTRACT</i> .....	v
KATA PENGANTAR .....	vi
DAFTAR ISI.....	ix
DAFTAR GAMBAR .....	xii
DAFTAR TABEL.....	xiii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	2
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Batasan Masalah.....	4
BAB II LANDASAN TEORI .....	5
2.1 Landasan Teori .....	5
2.1.1 Penambangan Data ( <i>Data Mining</i> ).....	5
2.1.2 Penambangan Teks ( <i>Text Mining</i> ) .....	7

2.1.3	Algoritma Term Frequency – Inverse Document Frequency (TF-IDF).....	9
2.1.4	<i>Support Vector Machine</i> .....	12
2.1.5	YouTube.....	15
2.1.6	<i>Google Colaboratory</i> .....	16
2.2	Penelitian Terdahulu.....	17
BAB III METODOLOGI.....		19
3.1	Metode Pengumpulan Data .....	20
3.2	<i>Text Preprocessing</i> .....	22
3.2.1	<i>Data Labelling</i> .....	22
3.2.2	<i>Data Cleansing</i> .....	23
3.2.3	Normalisasi .....	23
3.2.4	<i>Case Folding</i> .....	24
3.2.5	<i>Tokenizing</i> .....	24
3.2.6	<i>Stopwords Removing</i> .....	25
3.2.7	<i>Stemming</i> .....	25
3.2.8	Penerapan Metode TF-IDF .....	26
3.2.9	Penerapan Metode SVM ( <i>Support Vector Machine</i> ).....	26
BAB IV HASIL DAN PEMBAHASAN .....		28
4.1	Hasil Pengumpulan Data .....	28
4.2	Hasil Penelitian.....	29
4.2.1	<i>Data Cleansing</i> .....	29
4.2.2	<i>Case Folding</i> .....	29
4.2.3	<i>Tokenizing</i> .....	30
4.2.4	Normalisasi .....	31

4.2.5	<i>Stopwords Removing</i> .....	31
4.2.6	<i>Stemming</i> .....	32
4.2.7	<i>Training Model</i> .....	32
4.3	Pembahasan.....	34
BAB V KESIMPULAN DAN SARAN.....		35
5.1	Kesimpulan.....	35
5.2	Saran.....	36
DAFTAR PUSTAKA.....		37



## DAFTAR GAMBAR

Gambar 2. 1 Optimal Hyperplane .....	12
Gambar 2. 2 Support Vector .....	14
Gambar 3. 1 Metodologi Penelitian .....	19
Gambar 3. 2 Tampilan Website WebScrapers .....	20
Gambar 3. 3 Tampilan WebScrapers .....	21
Gambar 3. 4 Tampilan Hasil <i>Scraping</i> .....	21
Gambar 3. 5 Kumpulan data hasil <i>scraping</i> .....	22
Gambar 3. 6 Potongan kode program data <i>cleansing</i> .....	23
Gambar 3. 7 Potongan kode program normalisasi .....	24
Gambar 3. 8 Potongan kode program <i>case folding</i> .....	24
Gambar 3. 9 Potongan kode program <i>tokenizing</i> .....	24
Gambar 3. 10 Potongan kode program <i>stopwords removing</i> .....	25
Gambar 3. 11 Potongan kode program <i>stemming</i> .....	26
Gambar 3. 12 Potongan kode program TF-IDF .....	26
Gambar 3. 13 Potongan kode program SVM .....	27
Gambar 4. 1 Fungsi Pandas di Google Collab .....	28
Gambar 4. 2 Tampilan visualisasi data mentah .....	28
Gambar 4. 3 Hasil proses data <i>cleansing</i> .....	29
Gambar 4. 4 Hasil proses <i>case folding</i> .....	30
Gambar 4. 5 Hasil proses <i>tokenizing</i> .....	30
Gambar 4. 6 Hasil proses normalisasi .....	31
Gambar 4. 7 Hasil proses <i>stopwords removing</i> .....	32
Gambar 4. 8 Hasil proses <i>stemming</i> .....	32
Gambar 4. 9 Hasil proses <i>training data</i> .....	33

## DAFTAR TABEL

Tabel 2. 1 Perhitungan TF-IDF.....	11
Tabel 4. 1 Tabel <i>Confusion Matrix</i> .....	33

