

## DAFTAR ISI

<b>TUGAS AKHIR .....</b>	<b>i</b>
<b>LEMBAR PENGESAHAN .....</b>	<b>ii</b>
<b>LEMBAR ORISINALITAS .....</b>	<b>iii</b>
<b>KATA PENGANTAR.....</b>	<b>iv</b>
<b>DAFTAR ISI.....</b>	<b>v</b>
<b>DAFTAR TABEL .....</b>	<b>viii</b>
<b>DAFTAR GAMBAR.....</b>	<b>x</b>
<b>DAFTAR LAMPIRAN.....</b>	<b>xi</b>
<b>ABSTRAK .....</b>	<b>xiii</b>
<b>ABSTRACT .....</b>	<b>xiv</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah.....	3
1.3 Pertanyaan Penelitian.....	4
1.4 Tujuan Penelitian .....	4
1.5 Batasan dan Asumsi Penelitian.....	4
1.6 Manfaat Penelitian .....	5
<b>BAB II TINJAUAN PUSTAKA .....</b>	<b>7</b>
2.1 Tinjauan Pustaka.....	7
2.2 Landasan Teori.....	13
2.2.1 <i>Machine Learning</i> .....	13
2.2.2 <i>Natural Language Processing (NLP)</i> .....	13
2.2.3 <i>Text Mining</i> .....	14
2.2.4 Mahkamah Agung.....	22

2.2.5	UU ITE.....	22
2.2.6	Jenis Pasal UU ITE pada Penelitian.....	23
2.2.7	Kronologi Pelanggaran UU ITE.....	24
<b>BAB III METODE PENELITIAN .....</b>		<b>30</b>
3.1	Subjek Dan Objek Penelitian .....	30
3.1.1	Subjek Penelitian.....	30
3.1.2	Objek Penelitian .....	30
3.2	Alat Dan Bahan.....	30
3.2.1	Alat.....	30
3.2.2	Bahan.....	31
3.3	Diagram Alir Penelitian .....	31
3.3.1	Studi Literatur .....	32
3.3.2	Pengumpulan Data .....	33
3.3.3	Pra-Pemrosesan Data .....	34
3.3.4	Pembobotan TF-IDF .....	38
3.3.5	Klasifikasi Model K-NN .....	40
3.3.6	Evaluasi Model.....	41
3.4	Jadwal Penelitian .....	42
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>		<b>44</b>
4.1	Pengumpulan Data .....	44
4.2	Pra-Pemrosesan Data .....	45
4.2.1	<i>Data Cleansing</i> .....	45
4.2.2	<i>Data Cleaning</i> .....	46
4.2.3	<i>Stopword Removal</i> .....	49
4.2.4	<i>Stemming</i> .....	56
4.3	Pembobotan TF-IDF .....	59

4.3.1	Menghitung Nilai <i>Term-Frequency</i> (TF).....	59
4.3.2	Menghitung Nilai <i>Inverse Document Frequency</i> (IDF).....	61
4.3.3	Menghitung nilai TF-IDF dan Normalisasi Nilai Tersebut.....	63
4.4	<i>Split Train vs Test – StratifiedKfoldSampling</i> .....	74
4.4.1	<i>Label Encoder</i> .....	74
4.4.2	<i>Split vs Test</i> .....	75
4.5	Analisis Klasifikasi <i>K-Nearest Neighbors</i> (K-NN) .....	75
4.5.1	Perhitungan Jarak ( <i>Euclidean Distance</i> ).....	76
4.5.2	Pemilihan Tetangga Terdekat .....	77
4.5.3	Prediksi Akhir .....	79
4.6	Modeling dan Evaluasi .....	80
4.6.1	<i>K-Nearest Neighbors</i> (K-NN).....	80
4.6.2	<i>SMOTE dan Dimensionality Reduction - PCA</i> .....	85
<b>BAB V KESIMPULAN DAN SARAN</b> .....		<b>93</b>
5.1	Kesimpulan .....	93
5.2	Saran .....	93
<b>DAFTAR PUSTAKA</b> .....		<b>95</b>
<b>LAMPIRAN</b> .....		<b>102</b>