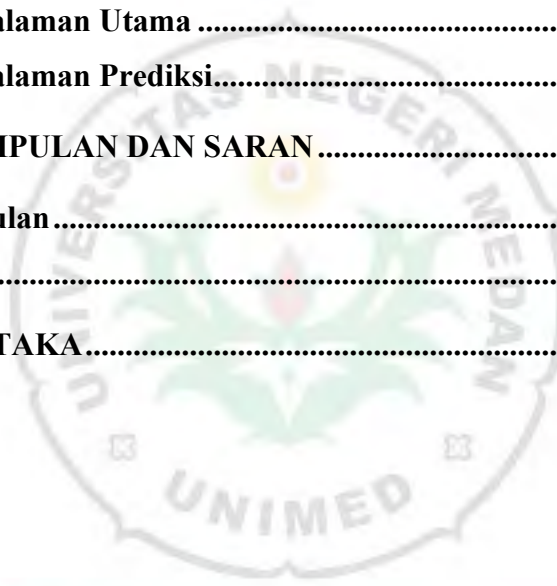


## DAFTAR ISI

<b>Lembar Pengesahan</b> .....	<b>I</b>
<b>Riwayat Hidup</b> .....	<b>II</b>
<b>Halaman Pernyataan Orisinalitas</b> .....	<b>III</b>
<b>Halaman Persetujuan Publikasi</b> .....	<b>IV</b>
<b>Abstrak</b> .....	<b>V</b>
<b>Kata Pengantar</b> .....	<b>VII</b>
<b>Daftar Isi</b> .....	<b>IX</b>
<b>Daftar Gambar</b> .....	<b>XII</b>
<b>Daftar Tabel</b> .....	<b>XIII</b>
<b>Daftar Lampiran</b> .....	<b>XIV</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
<b>1.1 Latar Belakang</b> .....	<b>1</b>
<b>1.2 Rumusan Masalah</b> .....	<b>4</b>
<b>1.3 Ruang Lingkup</b> .....	<b>4</b>
<b>1.4 Tujuan Penelitian</b> .....	<b>4</b>
<b>1.5 Batasan Penelitian</b> .....	<b>5</b>
<b>1.6 Manfaat Penelitian</b> .....	<b>5</b>
<b>BAB II TINJAUAN PUSTAKA</b> .....	<b>6</b>
<b>2.1 Konten Negatif</b> .....	<b>6</b>
<b>2.2 Undang Undang Informasi dan Transaksi Elektronik (UU ITE)</b> .....	<b>6</b>
<b>2.3 Artificial Intelligence</b> .....	<b>7</b>
<b>2.4 Machine Learning</b> .....	<b>9</b>
<b>2.5 Natural Language Processing</b> .....	<b>10</b>
<b>2.6 Klasifikasi</b> .....	<b>12</b>
<b>2.7 Support Vector Machine</b> .....	<b>13</b>
<b>2.8 Random Forest</b> .....	<b>15</b>
<b>2.9 Pembagian Data</b> .....	<b>17</b>

2.9.1	Data Splitting.....	17
2.10	Evaluasi.....	18
2.11	Web Scraping .....	20
2.12	Term Frequency – Inverse Document Frequency (TF-IDF) .....	21
2.13	Penelitian yang Relevan .....	23
<b>BAB III METODE PENELITIAN .....</b>		<b>24</b>
3.1	Tempat dan Waktu Penelitian.....	24
3.2	Perancangan Umum .....	24
3.3	Dataset.....	26
3.4	Preprocessing.....	27
3.4.1	Duplicated Data.....	27
3.4.2	Text Cleansing.....	27
3.4.3	Case Folding .....	27
3.4.4	Stopwords .....	28
3.4.5	Tokenize .....	28
3.4.6	Label Encoding.....	28
3.4.7	Data Splitting.....	28
3.4.8	TF-IDF (Term Frequency — Inverse Document Frequency) ..	28
3.5	Klasifikasi dengan Model Machine Learning .....	28
3.5.1	Support Vector Machine .....	29
3.5.2	Random Forest.....	29
3.6	Evaluasi Model.....	30
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>		<b>31</b>
4.1	Perancangan Sistem .....	31
4.1.1	Perangkat Keras .....	31
4.1.2	Perangkat Lunak .....	31
4.2	Akuisisi Data.....	31
4.3	Preprocessing Data .....	33
4.3.1	Duplicated Data.....	34
4.3.2	Text Cleansing.....	34

4.3.3	Case Folding .....	35
4.3.4	Tokenize .....	35
4.3.5	Stopwords .....	36
4.3.6	Label Encoding.....	36
4.3.7	Data Splitting.....	37
4.3.8	TF-IDF (Term Frequency — Inverse Document Frequency) ..	37
4.5	Perbandingan Hasil Klasifikasi .....	40
4.6	Implementasi Model Machine Learning pada Website .....	47
4.6.1	Halaman Utama .....	47
4.6.2	Halaman Prediksi.....	48
<b>BAB V KESIMPULAN DAN SARAN .....</b>		<b>50</b>
5.1	Kesimpulan.....	50
5.2	Saran .....	50
<b>DAFTAR PUSTAKA.....</b>		<b>52</b>



## DAFTAR GAMBAR

<b>Gambar 2.1.</b> Klasifikasi dengan input $x$ ke dalam label kelasnya.....	13
<b>Gambar 2.2.</b> <i>Support Vector Machine Hyperplane</i> klasifikasi antara dua kelas.	14
<b>Gambar 2.3.</b> Visualisasi dari algoritma <i>decision tree</i> untuk klasifikasi. ....	16
<b>Gambar 2.4.</b> Visualisasi algoritma <i>random forest</i> .....	17
<b>Gambar 2.5.</b> Perbandingan antara data untuk <i>training</i> dan data untuk <i>testing</i> . ..	18
<b>Gambar 2.6.</b> Data <i>training</i> dan <i>testing</i> pada klasifikasi .....	18
<b>Gambar 2.7.</b> Arsitektur web scraping pengambilan data. ....	21
<b>Gambar 3.1.</b> Rancangan umum penelitian.....	26
<b>Gambar 3.2.</b> Ilustrasi Implementasi SVM pada data.....	29
<b>Gambar 3.3.</b> <i>Decision tree</i> pada <i>Random Forest Classifier</i> . ....	30
<b>Gambar 4.1.</b> Hasil scraping website.....	32
<b>Gambar 4.2.</b> Data hasil scraping yang telah disatukan kedalam csv.....	33
<b>Gambar 4.3.</b> Dataset sebelum <i>preprocessing</i> .....	33
<b>Gambar 4.4.</b> Proses <i>cleansing</i> data teks .....	34
<b>Gambar 4.5.</b> Proses <i>case folding</i> .....	35
<b>Gambar 4.6.</b> Proses <i>tokenize</i> .....	35
<b>Gambar 4.7.</b> Proses <i>stopwords removal</i> .....	36
<b>Gambar 4.8.</b> Perbandingan hasil <i>confusion matrix</i> pada <i>machine learning</i> .....	41
<b>Gambar 4.9.</b> Perbandingan <i>Support Vector Machine</i> dan <i>Random Forest</i> . ....	46
<b>Gambar 4.10.</b> Halaman utama aplikasi pendeteksi konten negatif website .....	47
<b>Gambar 4.11.</b> Hasil prediksi kelas <i>whitelist</i> aplikasi pendeteksi konten negatif.	48
<b>Gambar 4.12.</b> Hasil prediksi pendeteksi konten negatif pada website perjudian.	49

## DAFTAR TABEL

<b>Tabel 3.1.</b> Contoh dataset URL .....	27
<b>Tabel 4.1.</b> <i>Label Encoding</i> pada dataset.....	37
<b>Tabel 4.2.</b> Proporsi untuk <i>Data Training</i> dan <i>Data Testing</i> .....	37
<b>Tabel 4.3.</b> Tabel TF-IDF ( <i>Term Frequency - Inverse Document Frequency</i> ) .....	39
<b>Tabel 4.4.</b> <i>True Positive</i> pada kelas Perjudian .....	42
<b>Tabel 4.5.</b> <i>False Positive</i> pada kelas Perjudian .....	42
<b>Tabel 4.6.</b> <i>False Negative</i> pada kelas Perjudian.....	43
<b>Tabel 4.7.</b> <i>True Negative</i> pada kelas Perjudian .....	43
<b>Tabel 4.8.</b> Hasil metrik untuk model <i>Support Vector Machine</i> .....	45
<b>Tabel 4.9.</b> Hasil metrik untuk model <i>Random Forest</i> .....	45
<b>Tabel 4.10.</b> <i>Macro average support vector machine</i> .....	46
<b>Tabel 4.11.</b> <i>Macro average random forest</i> .....	46



## DAFTAR LAMPIRAN

<b>Lampiran 1.</b> Surat Izin Penelitian.....	55
<b>Lampiran 2.</b> Surat Keterangan Telah Melaksanakan Penelitian .....	56
<b>Lampiran 3.</b> Listing Code Program Machine Learning .....	57
<b>Lampiran 4.</b> Listing Code Program Aplikasi Website .....	63

