

## DAFTAR ISI

<b>HALAMAN PENGESAHAN.....</b>	<b>i</b>
<b>PERNYATAAN.....</b>	<b>ii</b>
<b>MOTTO .....</b>	<b>iii</b>
<b>PRAKATA .....</b>	<b>iv</b>
<b>DAFTAR ISI.....</b>	<b>v</b>
<b>DAFTAR TABEL.....</b>	<b>viii</b>
<b>DAFTAR GAMBAR.....</b>	<b>ix</b>
<b>INTISARI.....</b>	<b>xi</b>
<b>ABSTRAK .....</b>	<b>xii</b>
<b>I. PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian.....	5
1.4 Batasan Masalah.....	5
1.5 Manfaat penelitian .....	5
<b>II. TINJAUAN PUSTAKA.....</b>	<b>6</b>
<b>III. LANDASAN TEORI.....</b>	<b>11</b>
3.1 <i>Authorship Attribution</i> .....	11
3.1.1 Text preprocessing .....	11
3.1.2 <i>Case folding</i> .....	11
3.1.3 <i>Tokenization</i> .....	12
3.1.4 <i>Stopwords removal</i> .....	12
3.1.5 <i>Stemming</i> .....	12
3.1.6 Count-Vectorizer (CV).....	13
3.1.7 Support Vector Machine (SVM) .....	13
3.1.8 <i>Random forest</i> .....	15

3.2	<i>Authorship Obfuscation</i> .....	16
3.2.1	Sistem berbasis aturan .....	16
3.2.2	Part of Speech (POS) <i>tag</i> .....	17
3.2.3	<i>Word embedding</i> .....	18
3.2.4	<i>Spelling correction</i> .....	19
3.2.5	Evaluasi .....	20
3.2.6	Akurasi .....	20
3.2.7	Presisi .....	21
3.2.8	<i>Recall</i> .....	21
3.2.9	<i>F1-score</i> .....	21
3.2.10	<i>Cosine similarity</i> .....	22
<b>IV.</b>	<b>METODE PENELITIAN.....</b>	<b>23</b>
4.1	Deskripsi Umum Penelitian.....	23
4.2	Tahapan Penelitian .....	25
4.3	Dataset .....	26
4.4	Rancangan Penelitian .....	27
4.4.1	<i>Text preprocessing</i> .....	27
4.4.2	Ekstraksi fitur .....	29
4.4.3	Rancangan model <i>authorship attribution</i> .....	29
4.4.4	Analisis fitur.....	30
4.4.5	Rancangan model <i>authorship obfuscation</i> .....	30
4.4.6	Evaluasi sistem .....	40
<b>V.</b>	<b>IMPLEMENTASI .....</b>	<b>42</b>
5.1	Spesifikasi <i>Hardware</i> dan Software .....	42
5.2	Hirarki Modul Implementasi Kode .....	42
5.3	Modul <i>Modelling.py</i> .....	42
5.3.1	Mengimpor pustaka yang digunakan .....	42
5.3.2	Memuat dan membagi dataset.....	43
5.3.3	Membersihkan teks (data twitter).....	44
5.3.4	Ekstraksi fitur .....	44
5.3.5	Pelatihan model dan evaluasi .....	45
5.4	Modul <i>Stylometri.py</i> .....	46

5.4.1	Fungsi <code>count_words_per_sentence</code> .....	46
5.4.2	Fungsi <code>fraction_stopwords</code> .....	46
5.4.3	Fungsi <code>fraction_punctuations</code> .....	47
5.4.4	Impor <i>pre-trained pos-tagger</i> .....	47
5.4.5	Menghitung perbandingan kata benda, kerja, dan sifat.....	48
5.5	Modul <code>Obfuscator.py</code> .....	49
5.5.1	Impor pustaka <code>Word2Vec</code> .....	49
5.5.2	Fungsi memecah atau menggabungkan kalimat.....	49
5.5.3	Fungsi penggantian <i>stopwords</i> .....	51
5.5.4	Fungsi <i>spelling correction</i> .....	52
5.5.5	Fungsi menambah atau mengurangi tanda baca.....	55
5.5.6	Fungsi penggantian angka dan simbol ke bentuk <i>string</i> .....	56
5.5.7	Fungsi <i>lowercase word</i> .....	59
5.5.8	Fungsi substitusi kata.....	59
5.5.9	Fungsi menghitung cosine <i>similarity</i> antara dua teks .....	60
<b>VI.</b>	<b>HASIL DAN PEMBAHASAN.....</b>	<b>62</b>
6.1	Pengujian <i>Authorship attribution</i> . .....	62
6.2	Pengujian <i>Author Obfuscation</i> .....	63
6.2.1	Rata-rata metrik fitur stylometri.....	63
6.2.3	Hasil pengujian <i>safety</i> .....	64
6.2.3	Hasil pengujian <i>soundness</i> .....	71
6.2.4	Hasil pengujian sensibleness .....	76
<b>VII.</b>	<b>KESIMPULAN DAN SARAN .....</b>	<b>79</b>
7.1	Kesimpulan.....	79
7.2	Saran .....	80
<b>DAFTAR PUSTAKA.....</b>		<b>81</b>