

DAFTAR ISI

ABSTRAK	i
<i>ABSTRACT</i>	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN ORISINALITAS	iv
KATA PENGANTAR	v
DAFTAR ISI	vi
DAFTAR GAMBAR	x
DAFTAR TABEL	xiii
BAB I PENDAHULUAN	1
I.1 Latar Belakang	1
I.2 Perumusan Masalah	6
I.3 Tujuan Penelitian	6
I.4 Batasan Penelitian	6
I.5 Manfaat Penelitian	6
I.6 Sistematika Penulisan	7
BAB II TINJAUAN PUSTAKA	9
II.1 Dasar Teori	9
II.1.1 Twitter	9
II.1.2 Pemilihan Umum Presiden Indonesia 2024	11
II.1.3 <i>Data Mining</i>	12
II.1.4 <i>Text Preprocessing</i>	14
II.1.5 <i>Natural Language Processing (NLP)</i>	15
II.1.6 Pembobotan TF-IDF	17
II.1.7 Analisis Sentimen	18

II.1.8	<i>K-Nearest Neighbor (KNN)</i>	20
II.1.9	<i>K-Fold Cross Validation</i>	22
II.1.10	<i>Particle Swarm Optimization (PSO)</i>	24
II.1.11	<i>Random Undersampling</i>	27
II.1.12	<i>Syntethic Minority Oversampling Technique (SMOTE)</i>	28
II.1.13	Evaluasi	30
II.2	Perbandingan Metode	31
II.3	<i>State of The Art</i>	33
BAB III	METODOLOGI PENELITIAN	37
III.1	Pengembangan Model Konseptual	37
III.2	Sistematika Penyelesaian Masalah	38
III.2.1	Tahap Inisiasi	38
III.2.2	Tahap Implementasi	39
III.2.3	Tahap Evaluasi	40
III.3	Alasan Pemilihan Metode	40
III.4	Rencana Jadwal Kegiatan	41
BAB IV	ANALISIS DAN PERANCANGAN	43
IV.1	Pengumpulan Data	43
IV.2	<i>Data Preprocessing</i>	44
IV.3	<i>Cleaning data</i>	44
IV.4	Pelabelan Data	45
IV.5	<i>Stop Word Removal</i>	46
IV.6	<i>Tokenizing</i>	48
IV.7	Stemming	49
IV.8	<i>Split data</i>	50

IV.9	Implementasi TF-IDF	51
IV.10	Perancangan Model KNN dengan Optimasi PSO	53
IV.10.1	<i>Particle Swarm Optimization</i>	53
IV.10.2	K-Nearest Neighbor	58
IV.11	Metrik Evaluasi	59
BAB V	HASIL DAN EVALUASI	61
V.1	Penjelasan Data	61
V.2	Penjelasan Klasifikasi KNN	62
V.2.1	Menghitung jarak <i>Euclidean</i>	64
V.2.2	Menghitung jarak <i>Manhattan</i>	66
V.3	<i>Skenario 1: Baseline</i>	68
V.3.1	Dataset Anies	68
V.3.2	Dataset Ganjar	73
V.3.3	Dataset Prabowo	79
V.4	<i>Skenario 2: Random Undersampling</i>	84
V.4.1	Dataset Anies	85
V.4.2	Dataset Ganjar	90
V.4.3	Dataset Prabowo	95
V.5	<i>Skenario 3: Oversampling (SMOTE)</i>	100
V.5.1	Dataset Anies	101
V.5.2	Dataset Ganjar	105
V.5.3	Dataset Prabowo	110
V.6	Evaluasi Keseluruhan	115
V.7	Perbandingan performa KNN dengan KNN-PSO	117
V.8	Visualisasi <i>Wordcloud</i>	119

BAB VI	KESIMPULAN DAN SARAN	126
VI.1	Kesimpulan	126
VI.2	Saran	126
DAFTAR PUSTAKA		128