

DAFTAR ISI

HALAMAN ORISINALITAS	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHANAN.....	iii
DAFTAR RIWAYAT HIDUP	iv
HALAMAN PERSEMBAHAN	v
MOTTO	vii
ABSTRAK	viii
ABSTRACT.....	ix
PRAKATA.....	x
DAFTAR ISI	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
BAB I PENDAHULUAN	1
1.1. Latar Belakang.....	1
1.2. Ruang Lingkup	4
1.3. Rumusan Masalah.....	4
1.4. Tujuan Penelitian	4
1.5. Manfaat Penelitian	4
1.6. Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	6
2.1 <i>Cyberbullying</i> (Ujaran Kebencian).....	6
2.2 <i>Python</i>	6
2.2.1 <i>Pandas</i>	6
2.3 <i>Seaborn</i>	7
2.4 <i>Matplotlib</i>	7
2.5 <i>Scikit-Learn</i>	7
2.6 <i>Machine Learning</i>	7
2.7 <i>Natural Language Processing</i>	8
2.8 <i>K-Nearest Neighbor</i>	8
2.9 <i>Feature Engineering</i>	8
2.10 Pembobotan Kata Tf-Idf.....	8
2.11 <i>Confusion Matrix</i>	9
2.12 Kalimat Yang Mengandung <i>Cyberbullying</i> Atau Tidak	11
2.13 Metode Pengembangan Perangkat Lunak <i>Agile Development</i>	11
2.14 Penelitian Sebelumnya	12
BAB III METODOLOGI PENELITIAN	15
3.1. METODE PENELITIAN	15

3.1.1 Model Requirements.....	15
3.1.2 Data Collection	16
3.1.3 Preprocessing Data	18
3.1.4 Data Cleaning.....	19
3.1.6 Data Labeling.....	20
3.1.7 Feature Engineering	20
3.1.8 Model Training.....	20
3.1.9 Knn Confusion Matrix.....	21
3.1.10 Model Evaluation.....	24
3.1.11 Operation.....	24
BAB IV HASIL DAN PEMBAHASAN	25
4.1 Pengumpulan Data.....	25
4.2 Preprocessing Data.....	27
4.2.1 Hasil Import Data	27
4.2.2 Hasil Memisahkan Kolom Tanggal dan Waktu	28
4.2.3 Hasil Penghapusan Data Duplikat	28
4.2.4 Hasil Wordcloud	29
4.3 Hasil Data Cleaning.....	30
4.3.1 Case Folding.....	30
4.3.2 Hasil Tokenization.....	31
4.3.3 Stopword Removal.....	31
4.3.4 Stemming Data.....	32
4.3.5 Hasil Remove Nilai Null.....	33
4.3.6 Hasil Normalization	33
4.4 Hasil Labeling Data	34
4.5 Hasil Feature Engineering.....	36
4.6 Hasil Model Evaluation Algoritma K-Nearest Neighbor	37
4.7 Analisa Pendapat Ahli Bahasa Indonesia (Ibu Hasnawati Nasution).....	38
4.8 Perbandingan Hasil Pendapat Ahli Bahasa Dengan Model Sistem.....	41
4.9 Kelebihan Dan Kekurangan Sistem.....	42
4.9.1 Kelebihan	42
4.9.2 Kekurangan	42
4.10 Kelebihan Dan Kekurangan Ahli Bahasa.....	43
4.10.1 Kelebihan	43
4.10.2 Kekurangan	43
BAB V KESIMPULAN DAN SARAN	44
5.1 KESIMPULAN	44
5.2 SARAN.....	44
DAFTAR PUSTAKA	46

DAFTAR TABEL

Tabel 2.1 Contoh Kalimat yang Mengandung <i>Cyberbullying</i> atau Tidak.....	11
Tabel 2.2 Peneliti Terdahulu	12
Tabel 3.1 Kata Kunci.....	17
Tabel 4.1 Hasil Analisa Pendapat Ahli Bahasa Indonesia	39
Tabel 4.2 Perbandingan Hasil Pendapat Ahli Bahasa Dengan Model Sistem.....	41

DAFTAR GAMBAR

Gambar 1.1 Jumlah pengguna internet di Indonesia	2
Gambar 2.1 Siklus Metodologi <i>Machine Learning by Gartler</i>	8
Gambar 3.1 Alur Model <i>Requirements</i>	15
Gambar 3.2 Alur <i>Crawling</i>	16
Gambar 3.3 Alur <i>Preprocessing Data</i>	18
Gambar 3.4 Alur <i>Cleaning</i>	19
Gambar 3.5 KNN <i>Confusion Matrix</i> Perkalimat	21
Gambar 3.6 KNN <i>Confusion Matrix</i> Perkata.....	23
Gambar 4.1 <i>Registrasi Akun</i>	25
Gambar 4.2 <i>Auth Token</i>	25
Gambar 4.3 <i>Crawling Data</i>	26
Gambar 4.4 Hasil <i>Crawling Data</i>	26
Gambar 4.5 Hasil <i>Import Data</i>	27
Gambar 4.6 Hasil Memisahkan Kolom Tanggal dan Waktu.....	28
Gambar 4.7 Hasil Penghapusan Data Duplikat	29
Gambar 4.8 Hasil <i>WORDCLOUD</i>	29
Gambar 4.9 Hasil <i>Data Cleaning</i>	30
Gambar 4.10 Hasil <i>Floding</i>	30
Gambar 4.11 Hasil <i>Tokenization</i>	31
Gambar 4.12 Hasil <i>Stopword Removal</i>	31
Gambar 4.13 Hasil <i>Steming Data</i>	32
Gambar 4.14 Hasil <i>Remove Nilai Null</i>	33
Gambar 4.15 Hasil <i>Normalisasi</i>	34
Gambar 4.16 Hasil <i>Labeling Data Perkalimat</i>	34
Gambar 4.17 Hasil <i>Labeling Data Perkata</i>	35
Gambar 4.18 Hasil Jumlah Deteksi Perkalimat	35
Gambar 4.19 Hasil Jumlah Deteksi Perkata.....	35
Gambar 4.20 Hasil <i>Feature Engineering</i>	36
Gambar 4.21 Hasil Model Evaluasi KNN Perkalimat	37
Gambar 4.22 Hasil Model Evaluasi KNN Perkata.....	38

