

DAFTAR ISI

HALAMAN PENGESAHAN TESIS	i
PERNYATAAN.....	ii
PRAKATA.....	iii
DAFTAR ISI.....	v
DAFTAR GAMBAR	vi
DAFTAR TABEL	viii
INTISARI.....	ix
ABSTRACT	x
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah.....	5
1.3 Batasan Masalah	6
1.4 Tujuan	6
1.5 Manfaat	7
BAB II TINJAUAN PUSTAKA.....	8
BAB III LANDASAN TEORI.....	14
3.1 <i>Microtargeting</i>	14
3.2 <i>Political Microtargeting</i>	18
3.3 <i>Machine Learning</i>	19
3.4 <i>Natural Language Processing</i>	22
BAB IV METODOLOGI PENELITIAN	47
4.1 Pengumpulan Data.....	47
4.2 Metode Penelitian	47
BAB V IMPLEMENTASI.....	64
5.1 Alat dan Bahan	64
5.2 Pengumpulan Data.....	65
5.3 <i>Preprocessing Data</i>	67
5.4 <i>Sentiment Analysis</i>	74
5.5 Klasifikasi Pemilih Menggunakan <i>Latent Dirichlet Allocation</i> (LDA)	82
5.6 Klasifikasi <i>Voting Intention</i>	95
BAB VI HASIL DAN PEMBAHASAN	103
6.1 Pengumpulan dan <i>Preprocessing Data</i>	103
6.2 Analisa Sentimen untuk Menentukan Tingkat Dukungan.....	106
6.3 Klasifikasi Pemilih Menggunakan <i>Latent Dirichlet Allocation</i> (LDA)	112
6.4 Klasifikasi <i>Voting Intention</i>	143
6.5 <i>Political Microtargeting</i>	147
BAB VII KESIMPULAN DAN SARAN.....	150
7.1 Kesimpulan.....	150
7.2 Saran	152
DAFTAR PUSTAKA	153

DAFTAR GAMBAR

Gambar 3.1 Klasifikasi algoritma <i>machine learning</i>	20
Gambar 4.1 <i>Flowchart</i> metode penelitian.....	48
Gambar 4.2 Proses analisa sentimen dengan 2 skenario.....	55
Gambar 4.3 Proses analisa sentimen berbasis <i>lexicon</i>	57
Gambar 4.4 Proses klasifikasi tweet	59
Gambar 4.5 Proses <i>supervised machine learning</i>	63
Gambar 5.1 <i>User Interface</i> aplikasi digimind.....	66
Gambar 5.2 Penyusunan <i>query</i> pada aplikasi <i>digimind</i>	66
Gambar 5.3 Fungsi <i>filtering</i> data menggunakan <i>keyword</i> 1	68
Gambar 5.4 Fungsi <i>filtering</i> data menggunakan <i>keyword</i> 2	68
Gambar 5.5 <i>Filtering</i> data dengan kriteria.....	69
Gambar 5.6 Deklarasi fungsi <i>preprocessing</i>	70
Gambar 5.7 Deklarasi fungsi untuk <i>preprocessing</i> lanjutan.....	71
Gambar 5.8 Tokenisasi menggunakan <i>library</i> NLTK	71
Gambar 5.9 <i>Delete stopword</i> dan <i>noise</i>	72
Gambar 5.10 <i>Stemming</i> menggunakan <i>library</i> sastrawi	72
Gambar 5.11 Normalisasi dengan <i>dictionary</i> KBBA	73
Gambar 5.12 Data latih untuk <i>sentiment analysis</i>	75
Gambar 5.13 Deklarasi fungsi <i>import</i> dataset.....	75
Gambar 5.14 Deklarasi fungsi <i>import lexicon</i>	76
Gambar 5.15 <i>Sentiment analysis</i> skenario 1	77
Gambar 5.16 <i>Sentiment analysis</i> skenario 2	78
Gambar 5.17 Deklarasi fungsi <i>found_word</i>	79
Gambar 5.18 Pengujian untuk menentukan skenario terbaik.....	79
Gambar 5.19 Fungsi deklarasi <i>load lexicon</i>	80
Gambar 5.20 Deklarasi fungsi <i>found_word</i>	80
Gambar 5.21 Deklarasi fungsi analisa sentimen.....	81
Gambar 5.22 Fungsi nilai sentimen numerik menjadi label sentimen	81
Gambar 5.23 Analisis sentimen menggunakan skenario 2	82
Gambar 5.24 <i>Import library</i> untuk LDA	83
Gambar 5.25 Fungsi <i>import</i> dataset	83
Gambar 5.26 Fungsi pembentukan <i>bigram</i> dan <i>trigram</i>	84
Gambar 5.27 Model TF-IDF dan transformasi <i>corpus</i>	84
Gambar 5.28 Fungsi iterasi topik dan perhitungan nilai koherensi	85
Gambar 5.29 Model LDA dengan jumlah topik optimal	85
Gambar 5.30 Fungsi penentuan topik dominan dan kontribusi pada dokumen ..	86
Gambar 5.31 Fungsi visualisasi distribusi jumlah kata dalam dokumen	87
Gambar 5.32 Fungsi visualisasi distribusi kata berdasarkan topik dominan	88
Gambar 5.33 Fungsi <i>wordcloud</i>	88
Gambar 5.34 Fungsi distribusi jumlah kata dan bobot kata kunci per topik.....	89
Gambar 5.35 Fungsi visualisasi topik dominan dalam beberapa dokumen	90
Gambar 5.36 Fungsi distribusi dokumen berdasarkan topik dan bobot.....	91
Gambar 5.37 Fungsi <i>clustering</i> menggunakan t-SNE	92
Gambar 5.38 Visualisasi LDA menggunakan pyLDAvis.....	93

Gambar 5.39 Fungsi klasifikasi label.....	94
Gambar 5.40 <i>Import library Scikit-Learn</i>	95
Gambar 5.41 <i>Load dataset dan preprocessing data</i>	95
Gambar 5.42 Fungsi ubah <i>rating</i> klasifikasi menjadi label <i>biner</i>	96
Gambar 5.43 Pembagian data menjadi data latih dan data tes.....	97
Gambar 5.44 Ekstraksi fitur menggunakan TF-IDF	97
Gambar 5.45 <i>Oversampling</i> menggunakan SMOTE	97
Gambar 5.46 Penentuan <i>hyperparameter</i> menggunakan <i>grid search</i>	98
Gambar 5.47 Penentuan model terbaik	99
Gambar 5.48 Penentuan model prediksi terbaik berdasarkan akurasi	99
Gambar 5.49 Evaluasi model prediksi terbaik	99
Gambar 5.50 <i>Import library scikit-learn</i> pada SVM	100
Gambar 5.51 Memuat data dan pembuatan label biner.....	100
Gambar 5.52 Ekstraksi fitur menggunakan TF-IDF	100
Gambar 5.53 Pembagian data menjadi data latih dan data uji	101
Gambar 5.54 Pembentukan <i>pipeline</i> SVM.....	101
Gambar 5.55 Penentuan <i>hyperparameter</i> menggunakan <i>grid search</i>	102
Gambar 5.56 Evaluasi model terbaik	102
Gambar 5.57 Prediksi pada data uji	102
Gambar 6.1 Proses <i>crawling</i> data pada aplikasi digimind	103
Gambar 6.2 Contoh output data hasil <i>crawling</i> menggunakan digimind.....	104
Gambar 6.3 Contoh data hasil <i>filtering</i>	105
Gambar 6.4 Contoh data hasil <i>preprocessing</i>	105
Gambar 6.5 Hasil analisa sentimen pada dataset	110
Gambar 6.6 Nilai koherensi tiap iterasi topik	113
Gambar 6.7 Distribusi frekuensi jumlah kata dan dokumen.....	114
Gambar 6.8 Topik dominan dan persentase kontribusinya pada dokumen	116
Gambar 6.9 Hasil analisa kalimat resrepresentatif tiap topik	117
Gambar 6.10 Word cloud untuk setiap topik	120
Gambar 6.11 Plot jumlah kata dan bobot keyword tiap topik	122
Gambar 6.12 Distribusi dokumen berdasarkan topik dominan dan bobot topik	125
Gambar 6.13 t-SNE <i>clustering chart</i> model LDA	127
Gambar 6.14 Visualisasi LDA menggunakan pyLDAVis	130
Gambar 6.15 Persebaran klasifikasi pengguna twitter pada dataset	142
Gambar 6.16 Niat memilih pengguna twitter pada dataset.....	147
Gambar 6.17 Klasifikasi pengguna twitter.....	148

DAFTAR TABEL

Tabel 2.1 Tinjauan Pustaka	10
Tabel 5.1 Contoh kata dalam <i>dictionary stopwords</i> dan <i>noise</i>	71
Tabel 5.2 Contoh data pada <i>dictionary kbba</i>	73
Tabel 5.3 Contoh <i>keyword</i> untuk analisa topik lda	94
Tabel 6.1 Evaluasi analisa sentimen menggunakan skenario 1	107
Tabel 6.2 Evaluasi analisa sentimen menggunakan skenario 2	108
Tabel 6.3 Nilai koherensi tiap iterasi topik	112
Tabel 6.4 20 kata paling relevan dengan topik 1	131
Tabel 6.5 20 kata paling relevan dengan topik 2	132
Tabel 6.6 20 kata paling relevan dengan topik 3	133
Tabel 6.7 20 kata paling relevan dengan topik 4	133
Tabel 6.8 20 kata paling relevan dengan topik 5	134
Tabel 6.9 20 kata paling relevan dengan topik 6	135
Tabel 6.10 20 kata paling relevan dengan topik 7	135
Tabel 6.11 20 kata paling relevan dengan topik 8	136
Tabel 6.12 20 kata paling relevan dengan topik 8	137
Tabel 6.13 Evaluasi model klasifikasi <i>voting intention</i>	144
Tabel 6.14 Evaluasi model SVM	145