

DAFTAR ISI

	halaman
HALAMAN JUDUL	i
SKRIPSI.....	ii
LEMBAR PENGESAHAN SKRIPSI	iii
SURAT PERNYATAAN TENTANG ORISINALITAS	iv
PEDOMAN PENGGUNAN SKRIPSI	v
ABSTRAK.....	vi
ABSTRACT.....	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	2
1.3 Tujuan.....	2
1.4 Manfaat.....	3
1.5 Batasan Masalah.....	3
BAB II TINJAUAN PUSTAKA	4
2.1 Organisasi Pemerintah.....	4
2.2 Data Mining.....	6
2.3 <i>Tool Data Mining</i>	8
2.4 Twitter Crawling	9
2.5 Preprocessing.....	9
2.6 Cluster Analysis.....	19
2.7 Penelitian Terdahulu.....	25
BAB III METODOLOGI PENELITIAN	26
3.1 Objek Penelitian	26
3.2 Tahapan Penelitian	28

BAB IV HASIL DAN PEMBAHASAN	41
4.1 Crawling Tweet	41
4.2 Pre-processing	46
4.3 Cluster Analysis.....	57
4.4 Evaluasi Hasil Analisis.....	62
BAB V KESIMPULAN DAN SARAN	68
5.1 Kesimpulan.....	68
5.2 Saran.....	69
DAFTAR PUSTAKA	70
LAMPIRAN.....	73

DAFTAR TABEL

Nomor	Judul Tabel	Halaman
2.1	Misi Calon Presiden Terpilih	4
2.2	Data Mining Operations dan Association Techiques	6
2.3	Kombinasi Awalan yang Tidak Dibolehkan	16
2.4	Aturan Pemenggalan Kompleks	16
2.5	Metode Cluster Analysis	20
<u>2.6</u>	Tabel Tahapan Inisiasi Centroid	23
2.7	Teknik Validasi	24
3.1	Misi Masing-Masing Pasangan calon	26
3.2	Waktu Debat Calon Presiden dan Wakil Calon Presiden 2019	28
3.3	Tweet Hasil Crawling	29
3.4	Hasil Perhitungan TF-IDF	36
3.5	Hasil Perhitungan Jarak Cosine Similarity	37
3.6	Koleksi A dan B dari Dataset <i>D</i>	38
3.7	Koleksi A dan B inisiasi centroid pertama	39
3.8	Centroid awal untuk $k = 2$	39
3.9	Hasil Perhitungan Silhouette Coefficient	40
4.1	Hasil Operator Get Twitter Statues dari API Twitter	43
4.2	Hasil Removing Uniform Resource Locator (URL)	47
4.3	Hasil Removing Special Symbol on Twitter	48
4.4	Hasil Removing Symbol ASCII, Number, and punctuation	48
4.5	Hasil Tokenization Serta Frekuensi Kata Yang Ditemukan	49
4.6	Jumlah Kata Benar dan Salah	53
4.7	Rangkaian Kalimat Hasil Normalisasi	54
4.8	Rangkaian Kalimat Hasil Stemming	54
4.9	Hasil Removing Stopwords	55
4.10	Besar Dimensi Matriks TF-IDF	55
4.11	P-Value Euclidean	57
4.12	Tweet Centroid Pertama	59
4.13	Hasil Pelabelan Menurut Pakar Ilmu Politik	64

DAFTAR GAMBAR

Nomor	Judul Gambar	Halaman
2.1	Peran Kunci Visi dan Misi	5
2.2	Tokenizing	11
<u>2.3</u>	Arsitektur Skip-Gram	12
2.4	Contoh Hasil Perhitungan Similarity	13
3.1	Diagram Tahapan Penelitian	28
3.2	Tahapan Pre-processing	30
3.3	Input dan Output Tahap Removing URL	30
3.4	Input dan Output tahapan Removing Special Symbols	31
3.5	Input dan Output tahapan Removing Symbol ASCII	32
3.6	Input dan Output Tokenization	32
3.7	Ilustrasi Pencarian Kata Normalisasi	33
3.8	Flowchart normalisasi	34
3.9	Input dan Output Tahap Stemming	35
3.10	Input dan Output Tahap Removing Stopwords	35
3.11	Tahapan Improved K-Means	36
3.12	Visualisasi Density	38
4.1	Token API Twitter	42
4.2	Operator Get Twitter User Statues Masing-Masing Calon	42
4.3	Operator Filter Examples	44
4.4	Operator Select Attributes	44
4.5	Penentuan Attribute Filter Type	45
4.6	Operator Append	45
4.7	Operator Write Excel	46
4.8	Hasil Penyimpanan Dataset Berupa Excel(.xlsx)	46
4.9	Tag Input Token	50
4.10	variabel payload	51
4.11	Variabel Header	51
4.12	Hasil Request HTML	52

Nomor	Judul Gambar	Halaman
4.13	Persentase Hasil Normalisasi Kata	53
4.14	Selisih Euclidean Non-Normalisasi dan Normalisasi	56
4.15	Distance Tweet Normalisasi dan Non-Normalisasi	56
4.16	Density pada Tweet Pasangan Calon 01	58
4.17	Density pada Tweet Pasangan Calon 02	58
4.18	Visualisasi Singular Value Decomposition Tweet	59
4.19	Grafik Silhouette Coefficient Pasangan Calon 01	60
4.20	Grafik Silhouette Coefficient Pasangan Calon 02	61
4.21	Nilai SSE Pasangan Calon 01	61
4.22	Nilai SSE Pasangan Calon 02	61
4.23	Pembagian Tweet Sesuai Kelas	62
4.24	Persentase Tweet berdasarkan Misi Calon Presiden 2019	65
4.25	Hasil Pemetaan Frekuensi Tweet Pasangan Calon 01 Sesuai Waktu Debat	65
4.26	Hasil Pemetaan Frekuensi Tweet Pasangan Calon 02 Sesuai Waktu Debat	67

DAFTAR LAMPIRAN

Nomor	Judul
1	Kata Stopwords
2	Diagram Crawling Tweet