

SKRIPSI
ANALISIS SENTIMEN DAN DETEKSI *BUZZER* TERHADAP
PASLON GUBERNUR DALAM KAMPANYE PILKADA
JAKARTA 2024 DENGAN NAÏVE BAYES DAN GAUSSIAN
NAÏVE BAYES



Intelligentia - Dignitas

SYALVA SYADILA
1519621001

PROGRAM STUDI
SISTEM DAN TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS NEGERI JAKARTA
2025

**ANALISIS SENTIMEN DAN DETEKSI BUZZER TERHADAP PASLON
GUBERNUR DALAM KAMPANYE PILKADA JAKARTA 2024 DENGAN
NAÏVE BAYES DAN GAUSSIAN NAÏVE BAYES**

Syalva Syadila

Dosen Pembimbing :

Pembimbing I : Lipur Sugiyanta, Ph.D

Pembimbing II: Murien Nugraheni, M.Cs

ABSTRAK

Media sosial X (sebelumnya Twitter) menjadi ruang perbincangan politik selama Pilkada Jakarta 2024, dengan 24,69 juta pengguna aktif di Indonesia. Sifatnya yang *real-time* menjadikan platform ini strategis untuk kampanye, namun juga rawan dimanfaatkan oleh *buzzer* untuk membentuk persepsi publik secara tidak organik, sehingga opini yang tersebar tidak mencerminkan suara masyarakat sebenarnya. Penelitian ini menyajikan pendekatan untuk menganalisis sentimen terhadap tiga pasangan calon gubernur serta mendeteksi *buzzer* untuk mengevaluasi pengaruh *buzzer* dalam membentuk persepsi publik selama Pilkada Jakarta 2024. Pertama, data dikumpulkan melalui crawling selama masa kampanye (25 September - 23 November 2024), dan setelah dilakukan pre-processing menghasilkan 7.781 *tweet*. Kedua, deteksi *buzzer* menggunakan algoritma Gaussian Naïve Bayes berdasarkan fitur seperti *followers*, *following*, *age*, *tweet per day*, *active weeks count*, dan *positive sentiment ratio*. Ketiga, analisis sentimen dilakukan dengan Multinomial Naïve Bayes untuk menentukan polaritas tiap *tweet*. Terakhir, distribusi sentimen dibandingkan sebelum dan sesudah *buzzer* disaring.

Hasil menunjukkan bahwa 55% *tweet* berasal dari 25 akun *buzzer*. Model deteksi *buzzer* mencapai akurasi 90%, dan model sentimen 85,65%. *Buzzer* paling banyak memengaruhi Paslon 1 (58,3%), diikuti Paslon 2 (57,6%) dan Paslon 3 (39%). Aktivitas *buzzer* memuncak pada 14, 16, dan 17 November 2024, bertepatan dengan momen pra dan pasca debat. Rata-rata *tweet buzzer* dalam 3 hari tersebut mencapai 184 *tweet*. Standar deviasi aktivitas *buzzer* sebesar 22,5, jauh lebih tinggi dibanding *non-buzzer* sebesar 0,32. Hal ini menunjukkan bahwa *buzzer* beroperasi secara terorganisir dan masif, serta lebih banyak menyebarkan sentimen positif dalam strategi kampanye bertipe *positive astroturfing*.

Kata Kunci: Analisis Sentimen, *Buzzer*, Pilkada Jakarta 2024

Intelligentia - Dignitas

**SENTIMENT ANALYSIS AND BUZZER DETECTION ON
GUBERNATORIAL CANDIDATES DURING THE 2024 JAKARTA
ELECTION CAMPAIGN USING NAÏVE BAYES AND GAUSSIAN NAÏVE
BAYES**

Syalva Syadila

Dosen Pembimbing :

Pembimbing I : Lipur Sugiyanta, Ph.D

Pembimbing II: Murien Nugraheni, M.Cs

ABSTRACT

Social media platform X (formerly Twitter) served as a major arena for political discourse during the 2024 Jakarta gubernatorial election, with 24.69 million active users in Indonesia. Its real-time and open nature makes it a strategic tool for political campaigns, but also vulnerable to manipulation by buzzers who shape public perception inorganically, thus potentially distorting the actual voice of the people. This study presents an approach to analyze sentiment toward three gubernatorial candidates and detect buzzer activity to evaluate its influence on public opinion throughout the campaign period. First, data were collected through crawling during the campaign period (25 September - 23 November 2024), resulting in 7,782 tweets after pre-processing. Second, buzzer detection was performed using the Gaussian Naïve Bayes algorithm based on features such as followers, following, account age, tweet per day, active weeks count, and positive sentiment ratio. Third, sentiment analysis was carried out using the Multinomial Naïve Bayes algorithm to determine the polarity of each tweet. Lastly, sentiment distribution was compared before and after buzzer filtering.

Results show that 55% of the tweets originated from 25 buzzer accounts. The buzzer detection model achieved 90% accuracy, while the sentiment model reached 85.65%. Buzzers had the strongest influence on Candidate 1 (58.3%), followed by Candidate 2 (57.6%) and Candidate 3 (39%). Their activity peaked on 14, 16, and 17 November 2024—coinciding with pre- and post-debate periods. Buzzers averaged 184 tweets during these three days, with a daily activity standard deviation of 22.5, significantly higher than non-buzzers at 0.32. These findings indicate that buzzers operate in an organized and large-scale manner, predominantly spreading positive sentiment through a positive astroturfing strategy.

Keywords: Sentiment Analysis, *Buzzer*, Jakarta Gubernatorial Election 2024

LEMBAR PENGESAHAN SKRIPSI

Judul : Analisis Sentimen dan Deteksi *Buzzer* terhadap Paslon Gubernur dalam Kampanye Pilkada Jakarta 2024 dengan Naïve Bayes dan Gaussian Naïve Bayes

Penyusun : Syalva Syadila

NIM : 1519621001

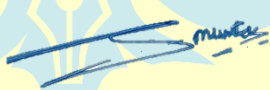
Pembimbing I : Lipur Sugiyanta, Ph.D

Pembimbing II: Murien Nugraheni, M.Cs

Proposal skripsi telah didiskusikan dan diusulkan dari Dosen Pembimbing:

NAMA DOSEN	TANDA TANGAN	TANGGAL
Lipur Sugiyanta, Ph.D (Dosen Pembimbing I)		3 Juli 2025

Murien Nugraheni, M.Cs (Dosen Pembimbing II)		3 Juli 2025
---	---	-------------

NAMA DOSEN	TANDA TANGAN	TANGGAL
Fuad Mumtas, S. Kom., M.T.I. (Dosen Penguji I)		17 Juli 2025

Irma Permata Sari, M.Eng (Dosen Penguji II)		23 Juli 2025
--	--	--------------

Intelligentia - Dignitas

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Skripsi ini merupakan Karya asli dan belum pernah diajukan untuk mendapatkan gelar akademik sarjana, baik di Universitas Negeri Jakarta maupun di Perguruan Tinggi lain.
2. Skripsi ini belum dipublikasikan, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
3. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Universitas Negeri Jakarta.

Jakarta, 23 Juli 2025

Yang membuat pernyataan



Syalva Syadila

1519621001

Intelligentia - Dignitas



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET
DAN TEKNOLOGI UNIVERSITAS NEGERI JAKARTA
UPT PERPUSTAKAAN

Jalan Rawamangun Muka Jakarta 13220

Telepon/Faksimili: 021-4894221

Laman: lib.unj.ac.id

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : Syalva Syadila

NIM : 1519621001

Fakultas/Prodi : Teknik/ Sistem dan Teknologi Informasi

Alamat email : syalvasyadila_1519621001@mhj.unj.ac.id

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada UPT Perpustakaan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

☒ Skripsi ☐ Tesis ☐ Disertasi ☐ Lain-lain (.....)

yang berjudul :

Analisis Sentimen dan Deteksi Buzzer terhadap Paslon Gubernur dalam Kampanye Pilkada Jakarta 2024 dengan Naïve Bayes dan Gaussian Naïve Bayes

Dengan Hak Bebas Royalti Non-Eksklusif ini UPT Perpustakaan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/mempublikasikannya di internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 23 Juli 2025

Penulis

(Syalva Syadila)

KATA PENGANTAR

Alhamdulillah, segala puji bagi Allah SWT atas limpahan rahmat dan karunia-Nya yang senantiasa menyertai setiap langkah, sehingga penulis dapat menyelesaikan skripsi ini sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Sistem dan Teknologi Informasi di Universitas Negeri Jakarta.

Dalam proses penyusunan skripsi ini, penulis banyak menerima bantuan, arahan, dan dukungan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Lipur Sugiyanta, Ph.D selaku Dosen Pembimbing I dan Koordinator Program Studi Sistem dan Teknologi Informasi, atas bimbingan, arahan, dan waktunya yang sangat berharga selama proses penyusunan skripsi ini.
2. Ibu Murien Nugraheni, M.Cs selaku Dosen Pembimbing II, atas dukungan, masukan, dan koreksi yang membangun selama proses pengerjaan skripsi.
3. Seluruh dosen dan staf Program Studi Sistem dan Teknologi Informasi, yang telah memberikan ilmu, serta bantuan selama masa perkuliahan.
4. Mamah, Bapak, Mas, dan seluruh keluarga yang selalu memberikan doa, semangat, dan dukungan moral maupun materiil tanpa henti.
5. Teman-teman seperjuangan Program Studi Sistem dan Teknologi Informasi yang telah menjadi tempat berbagi cerita, belajar bersama, dan saling mendukung dalam menyelesaikan tugas akhir ini.
6. Serta pihak-pihak lain yang telah membantu dalam penyelesaian skripsi ini yang tidak dapat disebutkan satu per satu.

Penulis menyadari masih ada kekurangan dalam skripsi ini. Kritik dan saran yang membangun sangat penulis harapkan. Semoga tulisan ini bermanfaat dan dapat menjadi referensi di masa depan.

Jakarta, 23 Juli 2025



Syalva Syadila

DAFTAR ISI

ABSTRAK	ii
LEMBAR PENGESAHAN SKRIPSI	iv
LEMBAR PERNYATAAN	v
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Pembatasan Masalah	4
1.4 Perumusan Masalah	5
1.5 Tujuan Penelitian	5
1.6 Manfaat Penelitian	5
BAB II KAJIAN PUSTAKA.....	7
2.1 Konsep dan Teori	7
2.1.1. <i>Buzzer</i>	7
2.1.2 Persepsi Publik.....	7
2.1.3. Pasangan Calon Gubernur.....	8
2.1.4. Analisis Sentimen	9
2.1.5. Naïve Bayes	10
2.1.6. Gaussian Naïve Bayes.....	11
2.1.7. <i>Machine Learning</i>	11
2.1.8. Diagram Alir	12
2.1.9. <i>K-fold Cross Validation</i>	13
2.2.10. <i>Confusion Matrix</i>	14
2.2.11. ROC-AUC	15
2.2.12. SMOTE.....	15
2.2 Penelitian Relevan.....	16
2.3 Kontribusi.....	18
2.4 Kerangka Berpikir	20

BAB III METODOLOGI PENELITIAN.....	21
3.1 Tempat dan Waktu Penelitian	21
3.2 Alat dan Bahan Penelitian.....	21
3.3 Langkah-langkah Penelitian.....	22
3.4 Teknik Pengumpulan Data.....	26
3.4.1 Pengumpulan Data <i>Tweet</i>	26
3.4.2 Pengumpulan Data Informasi Akun.....	31
3.4.3. Pre-processing Data <i>Tweet</i>	33
3.5. Metode Evaluasi Data	38
BAB IV HASIL PENELITIAN	39
4.1. Hasil Penelitian	39
4.1.1. Pengambilan Data Akun	39
4.1.2. Pre-Processing Data <i>Tweet</i>	40
4.1.3. Pelatihan Model Klasifikasi Sentimen	41
4.1.4. Pelatihan Model Deteksi <i>Buzzer</i>	43
4.1.5. Penghapusan <i>Tweet Buzzer</i>	46
4.2. Analisis Hasil Penelitian	46
4.2.1. Hasil Pengujian Model Klasifikasi Sentimen	46
4.2.1.1 Hasil Pengujian Model Klasifikasi Sentimen Untuk Pasangan Ridwan Kamil–Suswono	46
4.2.1.2 Hasil Pengujian Model Klasifikasi Sentimen Untuk Pasangan Dharma Pongrekun-Kun Wardana.....	48
4.2.1.3 Hasil Pengujian Model Klasifikasi Sentimen Untuk Pasangan Pramono Anung - Rano Karno	49
4.2.2. Hasil Pengujian Model Deteksi <i>Buzzer</i>	51
4.3. Pembahasan.....	53
BAB V PENUTUP.....	73
5.1 Kesimpulan	73
5.2 Saran.....	73
DAFTAR PUSTAKA	75

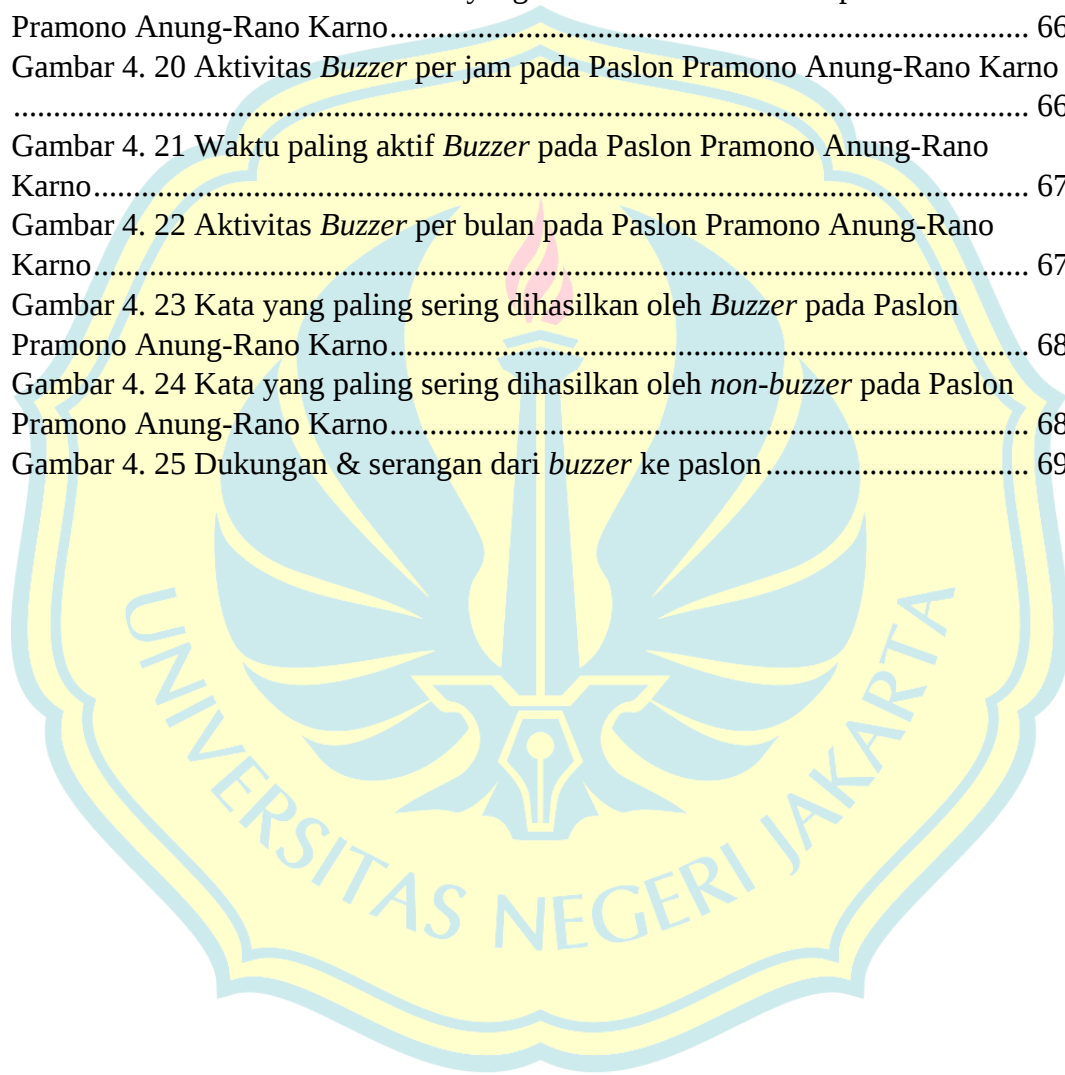
DAFTAR TABEL

Tabel 2. 1 Confusion Matrix	14
Tabel 3. 1 Alat dan Bahan Penelitian.....	21
Tabel 3. 2 Fitur Deteksi <i>Buzzer</i>	24
Tabel 3. 3 Kata Kunci	28
Tabel 3. 4 Daftar Kata Hubung	29
Tabel 3. 5 Contoh <i>Tweet</i> beserta label	30
Tabel 3. 6 Contoh Informasi Akun Beserta Label	32
Tabel 3. 7 Ilustrasi proses Case Folding	33
Tabel 3. 8 Ilustrasi Proses Cleaning	34
Tabel 3. 9 Ilustrasi Proses Tokenizing	35
Tabel 3. 10 Ilustrasi proses stopword removal.....	36
Tabel 3. 11 Ilustrasi proses normalisasi	37
Tabel 4. 1 Atribut data akun yang diambil.....	39
Tabel 4. 2 Hasil pre-processing.....	40
Tabel 4. 3 Contoh Labeling Manual pada Data Latih <i>Tweet</i>	41
Tabel 4. 4 Contoh Labeling Manual pada Data Latih Akun.....	45
Tabel 4. 5 Akurasi model klasifikasi sentimen Ridwan Kamil-Suswono untuk setiap fold.....	47
Tabel 4. 6 Rata-rata Macro Precision, Recall, dan F1-score per-Kelas pada klasifikasi sentimen pasangan Ridwan Kamil-Suswono.....	48
Tabel 4. 7 Akurasi model klasifikasi sentimen Dharma Pongrekun–Kun Wardana untuk setiap fold.....	48
Tabel 4. 8 Rata-rata Macro Precision, Recall, dan F1-score per Kelas pada klasifikasi sentimen pasangan Dharma Pongrekun–Kun Wardana	49
Tabel 4. 9 Akurasi Model Klasifikasi Sentimen Pasangan Pramono Anung- Rano Karno per Fold	50
Tabel 4. 10 Rata-rata Macro Precision, Recall, dan F1-score per-Kelas pada klasifikasi sentimen pasangan Pramono Anung-Rano Karno	50
Tabel 4. 11 Akurasi Model Klasifikasi <i>Buzzer</i> untuk Setiap Fold	51
Tabel 4. 12 Rata-rata precision, recall, F1-score, dan ROC AUC pada Model Klasifikasi <i>Buzzer</i>	51
Tabel 4. 13 Contoh Output Penilaian Model Klasifikasi <i>Buzzer</i>	52
Tabel 4. 14 Akun <i>Buzzer</i> dan Jenis Sentimen terhadap Paslon 1.....	55
Tabel 4. 15 Akun <i>Buzzer</i> dan Jenis Sentimen terhadap Paslon 2.....	60
Tabel 4. 16 Akun ssss dan Jenis Sentimen terhadap Paslon 3	65
Tabel 4. 17 List username <i>buzzer</i> per paslon	70

DAFTAR GAMBAR

Gambar 2. 1 Kerangka Berpikir	20
Gambar 3. 1 Diagram Alir Penelitian	22
Gambar 3. 2 Diagram Alir Pengumpulan Data	27
Gambar 3. 3 Install Library Tweet Harvest	27
Gambar 3. 4 Sampel Pengambilan Data pada Keyword 'paslon 2 pilkada jakarta'	
.....	28
Gambar 3. 5 Diagram Alir pengambilan data informasi akun	31
Gambar 3. 6 Diagram Alir Pre-processing	33
Gambar 3. 7 Diagram Alir Case Folding	34
Gambar 3. 8 Bagian Program Case Folding	34
Gambar 3. 9 Diagram Alir Cleaning	35
Gambar 3. 10 Bagian Program Cleaning	35
Gambar 3. 11 Diagram Alir Tokenizing	36
Gambar 3. 12 Bagian program tokenizing	36
Gambar 3. 13 Diagram Alir Stopword Removal	36
Gambar 3. 14 Bagian Program Normalisasi	37
Gambar 3. 15 Diagram Alir Normalisasi	37
Gambar 4. 1 Bagian program untuk perhitungan umur akun	44
Gambar 4. 2 Distribusi <i>tweet</i> Paslon Ridwan Kamil-Suswono Sebelum & Sesudah	
Penghapusan <i>Buzzer</i>	54
Gambar 4. 3 Persentase Sentimen yang dihasilkan <i>Buzzer</i> pada paslon Ridwan	
Kamil-Suswono	56
Gambar 4. 4 Persentase Sentimen yang dihasilkan <i>non-buzzer</i> pada paslon	
Ridwan Kamil-Suswono	56
Gambar 4. 5 Aktivitas <i>Buzzer</i> per jam pada paslon Ridwan Kamil-Suswono	57
Gambar 4. 6 Waktu Aktif <i>Buzzer</i> pada paslon Ridwan Kamil-Suswono	57
Gambar 4. 7 Aktivitas <i>Buzzer</i> per bulan pada paslon Ridwan Kamil-Suswono	58
Gambar 4. 8 Kata yang paling sering di hasilkan <i>Buzzer</i> pada paslon Ridwan	
Kamil-Suswono	58
Gambar 4. 9 Kata yang paling sering di hasilkan <i>non-buzzer</i> pada paslon Ridwan	
Kamil-Suswono	59
Gambar 4. 10 Distribusi <i>tweet</i> Paslon Dharma Pongrekun-Kun Wardana Sebelum	
& Sesudah Penghapusan <i>Buzzer</i>	60
Gambar 4. 11 Persentase Sentimen yang dihasilkan <i>Buzzer</i> pada paslon Dharma	
Pongrekun-Kun Wardana	61
Gambar 4. 12 Persentase Sentimen yang dihasilkan <i>non-buzzer</i> pada paslon	
Dharma Pongrekun-Kun Wardana	61
Gambar 4. 13 Aktivitas <i>Buzzer</i> per jam pada paslon Dharma Pongrekun-Kun	
Wardana	62
Gambar 4. 14 Aktivitas <i>Buzzer</i> per bulan pada paslon Dharma Pongrekun-Kun	
Wardana	62
Gambar 4. 15 Kata yang paling sering dihasilkan <i>Buzzer</i> pada paslon Dharma	
Pongrekun-Kun Wardana	63

Gambar 4. 16 Kata yang paling sering dihasilkan <i>non-buzzer</i> pada paslon Dharma Pongrekun-Kun Wardana.....	63
Gambar 4. 17 Distribusi <i>Tweet</i> Pasangan Pramono Anung-Rano Karno Sebelum & Sesudah Penghapusan <i>Buzzer</i>	64
Gambar 4. 18 Persentase Sentimen yang Dihasilkan <i>Buzzer</i> pada Paslon Pramono Anung-Rano Karno	65
Gambar 4. 19 Persentase Sentimen yang Dihasilkan <i>Non-buzzer</i> pada Paslon Pramono Anung-Rano Karno.....	66
Gambar 4. 20 Aktivitas <i>Buzzer</i> per jam pada Paslon Pramono Anung-Rano Karno	66
Gambar 4. 21 Waktu paling aktif <i>Buzzer</i> pada Paslon Pramono Anung-Rano Karno.....	67
Gambar 4. 22 Aktivitas <i>Buzzer</i> per bulan pada Paslon Pramono Anung-Rano Karno.....	67
Gambar 4. 23 Kata yang paling sering dihasilkan oleh <i>Buzzer</i> pada Paslon Pramono Anung-Rano Karno.....	68
Gambar 4. 24 Kata yang paling sering dihasilkan oleh <i>non-buzzer</i> pada Paslon Pramono Anung-Rano Karno.....	68
Gambar 4. 25 Dukungan & serangan dari <i>buzzer</i> ke paslon.....	69



Intelligentia - Dignitas

DAFTAR LAMPIRAN

Lampiran 1 Surat Tugas Dosen Pembimbing	78
Lampiran 2 Lembar konsultasi Dosen Pembimbing	79
Lampiran 3 Surat Pernyataan Dosen Pembimbing 1	81
Lampiran 4 Surat Pernyataan Dosen Pembimbing 2	82
Lampiran 5 Lembar persetujuan dosen pembimbing.....	83
Lampiran 6 Data tweet setelah Pre-Processing	84
Lampiran 7 Data username setelah Pre-Processing	85
Lampiran 8. Stopwords	86
Lampiran 9 Surat Permohonan Uji Validitas Analisis Sentimen.....	87
Lampiran 10 Surat Permohonan Uji Validitas Deteksi Buzzer	88
Lampiran 11 Validasi Deteksi Buzzer setelah divalidasi.....	89
Lampiran 12 Validasi analisis sentimen setelah divalidasi.....	91



Intelligentia - Dignitas