



SKRIPSI

PENERAPAN METODE *ROCCHIO* UNTUK KLASIFIKASI *BUZZER* DI MEDIA SOSIAL INSTAGRAM (STUDY KASUS DEMO KENAIKAN TUNJANGAN ANGGOTA DPR)

Rizky Rianto
NPM 20081010237

DOSEN PEMBIMBING
Fetty Tri Anggraeny, S.Kom. M.Kom
Andreas Nugroho Sihananto, S.Kom., M.Kom.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2026**

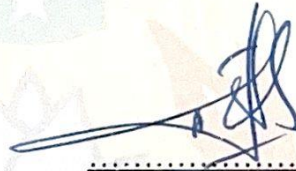
LEMBAR PENGESAHAN

Penerapan Metode *Rocchio* untuk Klasifikasi Buzzer di Media Sosial Instagram (Study Kasus Demo Kenaikan Tunjangan Anggota Dpr)

Oleh :
Rizky Rianto
NPM. 20081010237

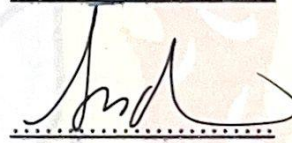
Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi *Prodi* Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 19 Januari 2026

Fetty Tri Anggraeny, S.Kom. M.Kom.
NIP. 19820211 202121 2 005



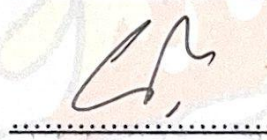
(Pembimbing I)

Andreas Nugroho Sihananto, S.Kom., M.Kom.
NIP. 19900412 202406 1 003



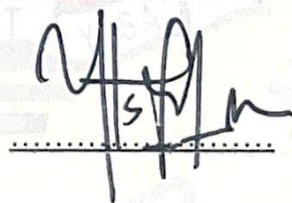
(Pembimbing II)

Eva Yulia Puspaningrum, S.Kom., M.Kom
NIP. 19890705 202121 2 002



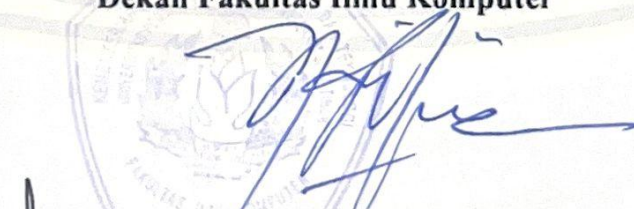
(Ketua Penguji)

Yisti Vita Via, S.ST., M.Kom.
NIP. 19860425 202121 2 001



(Anggota Penguji)

Mengetahui,
Dekan Fakultas Ilmu Komputer



Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

**Penerapan Metode *Rocchio* untuk *Klasifikasi Buzzer* di Media Sosial
Instagram (Study Kasus Demo Kenaikan Tunjangan Anggota Dpr)**

Oleh :

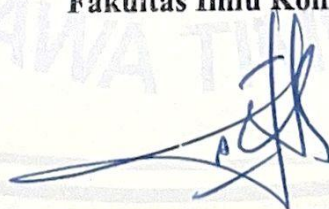
Rizky Rianto

NPM. 20081010237

Menyetujui,

Koordinator *Program Studi Informatika*

Fakultas Ilmu Komputer



Fetty Tri Anggraeny, S.Kom. M.Kom

NIP. 19820211 202121 2 005

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertandatangan di bawah ini:

Nama : Rizky Rianto
NPM : 20081010237
Program : Sarjana (S1)
Program Studi : Informatika
Fakultas : Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.

Surabaya, 19 Januari 2026
Yang Membuat Pernyataan,



Rizky Rianto
NPM. 20081010237

ABSTRAK

Nama Mahasiswa / Rizky Rianto / 20081010237
NPM :
Judul Skripsi : Penerapan Metode *Rocchio* untuk *Klasifikasi Buzzer* di Media Sosial Instagram (Study Kasus Demo Kenaikan Tunjangan Anggota DPR)
Dosen Pembimbing : 1. Fetty Tri Anggraeny, S.Kom. M.Kom
2. Andreas Nugroho Sihananto, S.Kom., M.Kom.

Aktivitas *buzzer* yang terkoordinir di media sosial Instagram berpotensi memanipulasi opini publik, khususnya dalam isu sensitif seperti demonstrasi penolakan kenaikan tunjangan anggota DPR. Deteksi manual terhadap komentar-komentar *buzzer* menjadi tidak efisien mengingat volume data yang besar. Oleh karena itu, penelitian ini bertujuan mengimplementasikan metode *Rocchio* (Nearest Centroid) guna membangun sistem *Klasifikasi* otomatis yang mampu mengkategorikan komentar Instagram ke dalam tiga kelas: *buzzer Pro* (mendukung), *buzzer Kontra* (menentang), dan *Non-buzzer* (netral). Data penelitian diperoleh melalui crawling komentar Instagram menggunakan kata kunci "#tunjangan DPR" dan "#demodpr" pada periode 19–30 Agustus 2025, menghasilkan 3.745 komentar mentah. Sebanyak 3.457 komentar kemudian dilabeli melalui pendekatan analisis sentimen otomatis dengan library VADER. Tahapan *preProcessing* meliputi cleaning, case folding, tokenizing, stopword removal, dan lemmatization. Representasi fitur teks menggunakan pembobotan TF-IDF (Term Frequency-Inverse Document Frequency). Model *Rocchio* dievaluasi dengan lima skenario pembagian data holdout (90:10, 80:20, 70:30, 60:40, 50:50) dan diukur menggunakan confusion matrix serta metrik akurasi, presisi, recall, dan F1-score. Hasil eksperimen menunjukkan bahwa skenario 90:10 menghasilkan kinerja terbaik dengan akurasi 82%, presisi makro 0.82, recall makro 0.82, dan F1-score makro 0.82. Model paling baik dalam mengenali kelas *buzzer Pro* (recall 0.96), diikuti oleh *buzzer Kontra*, sedangkan kelas *Non-buzzer* memiliki recall lebih rendah (0.72) karena variasi bahasa netral yang lebih luas. Selain *Klasifikasi* level komentar, penelitian ini mengembangkan fitur analisis akun yang mengelompokkan komentar berdasarkan username untuk mengidentifikasi kecenderungan utama suatu akun. Untuk menangani kondisi imbang (tie), diterapkan mekanisme tie-breaker berbasis jumlah kata. Kesimpulan penelitian mengkonfirmasi bahwa metode *Rocchio*, meskipun sederhana, efektif dan efisien sebagai model baseline untuk *Klasifikasi* tiga kelas pada data komentar Instagram dengan karakteristik sparse seperti TF-IDF. Analisis akun yang dikembangkan memberikan nilai tambah dengan menawarkan perspektif perilaku pengguna secara agregat. Hasil ini berkontribusi pada ranah computational social science dengan menyediakan metode deteksi *buzzer* yang cepat dan dapat diimplementasikan, sekaligus menjadi fondasi bagi pengembangan sistem pemantauan opini publik yang lebih canggih di platform media sosial.

Kata kunci: *Rocchio*, *Klasifikasi* teks, *buzzer*, Instagram, TF-IDF, komentar media sosial.

ABSTRACT

Student Name / NPM : Rizky Rianto / 20081010237
Thesis Title : Application of the *Rocchio* Method for *Buzzer* Classification on Instagram Social Media (Case Study of the Demonstration of Increase in Allowances for Members of the House of Representatives)
Advisor : 1. Fetty Tri Anggraeny, S.Kom. M.Kom
2. Andreas Nugroho Sihananto, S.Kom., M.Kom.

Coordinated buzzer activities on Instagram social media have the potential to manipulate public opinion, especially on sensitive issues such as demonstrations against the increase in allowances for members of the House of Representatives. Manual detection of buzzer comments becomes inefficient given the large volume of data. Therefore, this study aims to implement the *Rocchio* (Nearest Centroid) method to build an automatic classification system that is able to categorize Instagram comments into three classes: Pro buzzer (support), Contra buzzer (against), and Non-buzzer (neutral). The research data was obtained through Instagram comment crawling using the keywords "#tunjangandpr" and "#demodpr" in the period of August 19–30, 2025, resulting in 3,745 raw comments. A total of 3,457 comments were then labeled through an automated sentiment analysis approach with the VARDER library. The preProcessing stage includes cleaning, case folding, tokenizing, stopword removal, and lemmatization. Representation of text features using TF-IDF (Term Frequency-Inverse Document Frequency) weighting. The *Rocchio* model was evaluated with five holdout data sharing scenarios (90:10, 80:20, 70:30, 60:40, 50:50) and measured using a confusion matrix and accuracy, precision, recall, and F1-score metrics. The results of the experiments showed that the 90:10 scenario produced the best performance with an accuracy of 82%, macro precision of 0.82, macro recall of 0.82, and F1-score of macro 0.82. The model does best at recognizing the Pro buzzer class (0.96 recall), followed by the Contra buzzer, while the Non-buzzer class has a lower recall (0.72) due to the wider variety of neutral language. In addition to comment-level classification, this study developed an account analysis feature that groups comments by username to identify the main tendencies of an account. To handle the tie condition, a word-count tie-breaker mechanism is applied. The conclusions of the study confirm that the *Rocchio* method, although simple, is effective and efficient as a baseline model for the classification of three classes on Instagram comment data with sparse characteristics such as TF-IDF. The developed account analysis provides added value by offering an aggregate perspective of user behavior. These results contribute to the field of computational social science by providing a fast and implementable buzzer detection method, as well as being the foundation for the development of a more sophisticated public opinion monitoring system on social media platforms.

Keywords: *Rocchio*, text classification, *buzzer*, Instagram, TF-IDF, social media comments..

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala rahmat, hidayah, dan karunia-Nya kepada penulis sehingga skripsi dengan judul “Penerapan Metode *Rocchio* untuk *Klasifikasi Buzzer* di Media Sosial Instagram (Study Kasus Demo Kenaikan Tunjangan Anggota Dpr)” dapat terselesaikan dengan baik dan tepat waktu. Penulis menyampaikan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu baik secara langsung maupun tidak langsung selama *Proses* penyusunan skripsi ini. Secara khusus, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, MMT., IPU. selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Tri Anggraeny, S.Kom., M.Kom. selaku Koordinator *Program Studi Informatika* Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Ibu Retno Mumpuni, S.Kom., M.Sc selaku Dosen Wali yang telah banyak membantu penulis dari awal perkuliahan hingga terselesaikannya skripsi.
5. Ibu Fetty Tri Anggraeny, S.Kom., M.Kom, selaku Dosen Pembimbing I yang telah memberikan pengarahan, dukungan, serta masukan berharga selama Proses penelitian berlangsung.
6. Bapak Andreas Nugroho Sihananto, S.Kom., M.Kom., selaku Dosen Pembimbing II yang telah memberikan pengarahan, dukungan, serta masukan berharga selama Proses penelitian berlangsung.
7. Ibu Eva Yulia Puspaningrum, S.Kom., M.Kom selaku Ketua Penguji yang telah memberikan masukan dalam menyusun skripsi ini.
8. Ibu Yisti Vita Via, S.ST., M.Kom. selaku anggota penguji yang telah memberikan masukan dalam menyusun skripsi ini.
9. Seluruh Dosen Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur atas segala ilmu pengetahuan yang diberikan kepada penulis selama masa perkuliahan. Semoga penulis dapat mengamalkan ilmu pengetahuan yang telah diberikan

dan akan menjadi bekal yang berguna bagi penulis di masa yang akan datang.

10. Ayah dan Ibu yang telah memberikan dukungan secara materi dan non – materi sekaligus selalu mendoakan penulis dalam kondisi apapun. Sehingga penulis dapat menyelesaikan perkuliahan dari awal hingga terselesaikannya skripsi ini dengan baik.
11. Fajar Chandra Aditiya, Wahyu Ageng Rahino dan Lugas Fernanda Wirayudha yang telah membantu penulis dengan memberikan semangat untuk proses pengerjaan skripsi ini dari awal hingga skripsi ini terselesaikan serta memberikan bantuan secara moral kepada penulis dari awal bertemu hingga saat ini.
12. Muhammad Alviriza Ramadhan dan Rama Ariya Candra yang telah membantu memberikan arahan kepada penulis dalam proses pengerjaan program dan laporan skripsi dengan baik.
13. Rafi Anggara Nindiaputra, Kevin Austin Lumbanraja dan Calvin Symadhan Santoso yang telah membantu dan menemani menyelesaikan skripsi ini

Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat kekurangan. Oleh karena itu, kritik dan saran yang membangun dari berbagai pihak sangat diharapkan demi kesempurnaan karya ini. Akhirnya, dengan segala keterbatasan yang ada, penulis berharap semoga laporan skripsi ini dapat memberikan manfaat bagi semua pihak, khususnya dalam pengembangan ilmu pengetahuan di bidang pengolahan citra digital dan kecerdasan buatan.

Surabaya, 19 Januari 2026

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN.....	III
LEMBAR PERSETUJUAN.....	V
SURAT PERNYATAAN BEBAS PLAGIASI.....	VII
ABSTRAK.....	IX
ABSTRACT	XI
KATA PENGANTAR.....	XIII
DAFTAR ISI.....	XVI
DAFTAR GAMBAR.....	XIX
DAFTAR TABEL.....	XXII
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan Penelitian.....	3
1.4. Manfaat Penelitian.....	3
1.5. Batasan Masalah.....	5
BAB II TINJAUAN PUSTAKA.....	8
2.1. Penelitian Terdahulu	8
2.2. Media Sosial Instagram.....	12
2.3. <i>Buzzer</i> dan Disinformasi	13
2.4. Pelabelan <i>Buzzer Pro</i> , <i>Buzzer Kontra</i> , dan <i>Non Buzzer</i>	14
2.5. Demo Kenaikan Tunjangan DPR.....	15
2.6. Metode <i>Rocchio</i>	16
2.7. <i>Klasifikasi</i> Teks	18
2.8. Implementasi Metode <i>Rocchio</i> untuk <i>Klasifikasi</i> Teks.....	19
2.9. <i>Exploratory Analysis Data (EDA)</i>	20
2.10. TF-IDF	20
2.11. Confusion Matrix	21
2.12. Tie - Breaker.....	22
2.13. Evaluasi Model <i>Klasifikasi</i>	23

BAB III METODOLOGI	26
3.1. Studi Literatur	27
3.2. Kebutuhan <i>Software</i> dan <i>Hardware</i>	27
3.3. Pengumpulan Data	27
3.4. Pelabelan Data.....	28
3.5. Pembagian Data.....	28
3.6. Perancangan Model.....	28
3.7. Implementasi Model.....	40
3.8. Validasi Model.....	40
3.9. Kesimpulan dan Saran.....	40
3.10. Analisis Akun <i>Buzzer</i>	40
BAB IV HASIL DAN PEMBAHASAN.....	43
4.1. Analisis Data	43
4.2. Persiapan Data.....	43
4.2.1. Pengambilan Data	43
4.3. Pelabelan Data.....	45
4.4. Pembagian Data.....	46
4.5. Perancangan Model.....	47
4.5.1. <i>PreProcessing</i> Data	47
4.5.2. Pembagian Dataset Untuk Training dan Testing Model <i>Rocchio</i> ..	55
4.5.3. TF-IDF	56
4.5.4. <i>Klasifikasi Buzzer</i> Model <i>Rocchio</i>	56
4.5.5. Evaluasi Model.....	63
4.5.6. Penyimpanan Model.....	72
4.6. Penerapan Model.....	73
4.7. Validasi Model.....	77
4.8. Analisis Akun <i>Buzzer</i>	78
BAB V KESIMPULAN DAN SARAN	84
5.1. Kesimpulan.....	84
5.2. Saran.....	85
DAFTAR PUSTAKA	87
BIODATA PENULIS.....	91

DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian	26
Gambar 3. 2 Data Komentar	27
Gambar 3. 3 Data Sudah Terlabel	28
Gambar 3. 4 Perancangan Model	29
Gambar 3. 5 Alur PreProcessing.....	30
Gambar 3. 6 Alur <i>Rocchio</i>	36
Gambar 4. 1 Gambar Jumlah Komentar yang mengandung karakter tidak bermakna	47
Gambar 4. 2 Jumlah Kata Pada tiap Komentar	48
Gambar 4. 3 Jumlah Karakter Pada tiap Komentar.....	48
Gambar 4. 4 Panjang Rata-Rata Kata pada Tiap Komentar	48
Gambar 4. 5 Jumlah Data Tiap Kelas	49
Gambar 4. 6 Wordcloud	49
Gambar 4. 7 Hasil Cleaning	51
Gambar 4. 8 Hasil Casefolding	52
Gambar 4. 9 Hasil Tokenizing	52
Gambar 4. 10 Hasil Stopword Removal.....	53
Gambar 4. 11 Hasil Lemmization	53
Gambar 4. 12 Jumlah Data Sebelum <i>Balancing Data</i>	54
Gambar 4. 13 Jumlah Data setelah <i>Balancing Data</i>	55
Gambar 4. 14 Accuracy Skenario Pertama 90:10	58
Gambar 4. 15 Accuracy Skenario Kedua 80:20	59
Gambar 4. 16 Accuracy Skenario Ketiga 70:30	60
Gambar 4. 17 Accuracy Skenario Keempat 60:40	61
Gambar 4. 18 Skenario Kelima 50:50	62
Gambar 4. 19 Evaluasi Skenario Pertama	63
Gambar 4. 20 Evaluasi Skenario Kedua.....	65
Gambar 4. 21 Evaluasi Skenario Ketiga	67
Gambar 4. 22 Evaluasi Skenario Keempat.....	69
Gambar 4. 23 Evaluasi Skenario Kelima	71

Gambar 4. 24 Tampilan Interface Website	74
Gambar 4. 25 Tampilan Hasil Dari Text Komentar	75
Gambar 4. 26 Tampilan Hasil File Excel	76
Gambar 4. 27 Tampilan Menu Analisis Akun	78
Gambar 4. 28 Analisis Username yang berkomentar	79
Gambar 4. 29 Tampilan Analisis Nama Akun	81
Gambar 4. 30 Kesimpulan Analisis Akun	82

DAFTAR TABEL

Tabel 2. 1 Tabel Confusion Matrix	21
Tabel 3. 1 Tahapan Cleaning.....	31
Tabel 3. 2 Tahapan Case Folding.....	32
Tabel 3. 3 Tahapan Tokenizing	32
Tabel 3. 4 Tahapan Stopword Removal	32
Tabel 3. 5 Tahapan lemmatization	33
Tabel 3. 6 Frekuensi Kata	34
Tabel 4. 1 Hasil Crawling Data Komentar Instagram.....	44
Tabel 4. 2 Data Yang Sudah Dilabeling.....	45