

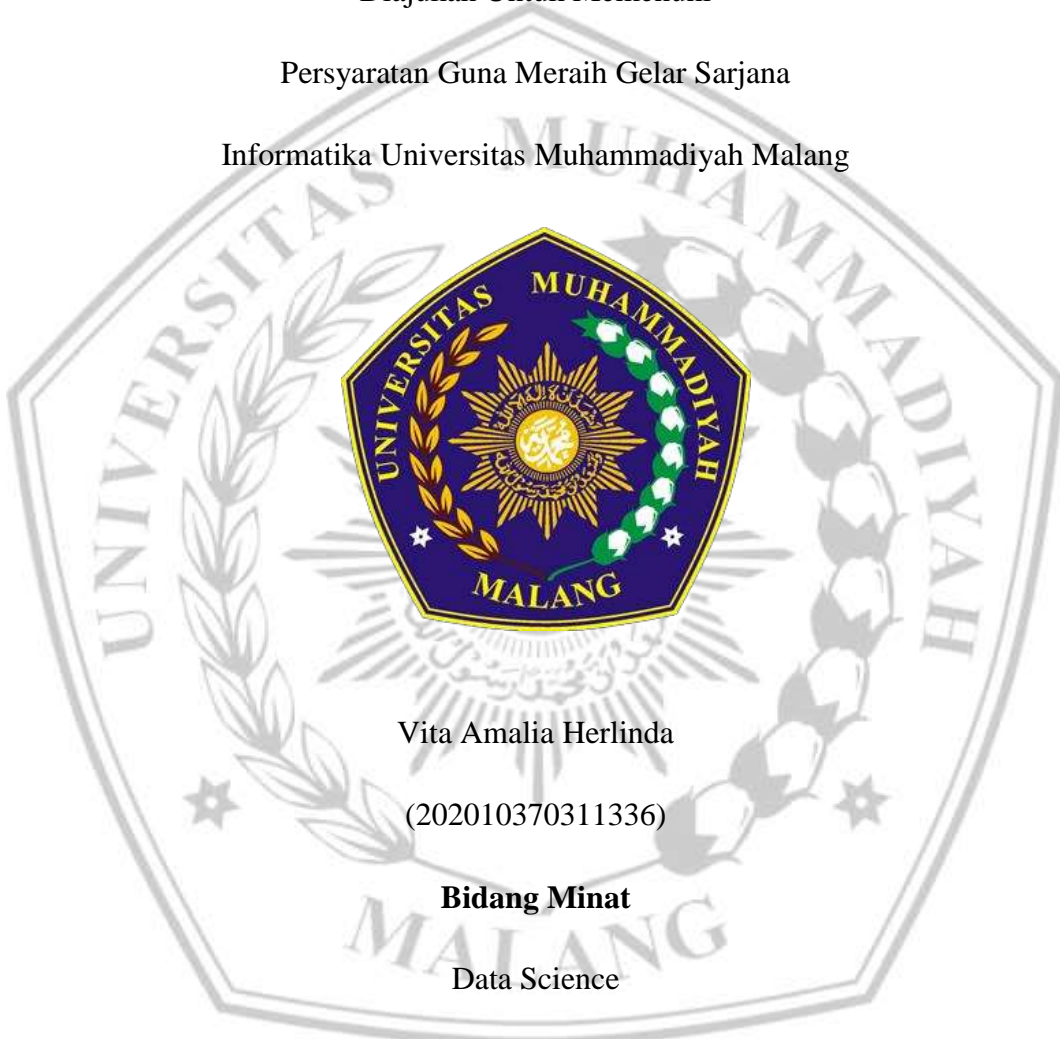
**Analisis Sentimen Masyarakat Terhadap Generasi Z Dalam Dunia Kerja
Pada Media Sosial Twitter Menggunakan Metode Naïve Bayes**

Proposal Tugas Akhir

Diajukan Untuk Memenuhi

Persyaratan Guna Meraih Gelar Sarjana

Informatika Universitas Muhammadiyah Malang



Vita Amalia Herlinda

(202010370311336)

Bidang Minat

Data Science

PROGRAM STUDI INFORMATIKA FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH MALANG

2024

LEMBAR PERSETUJUAN

**Analisis Sentimen Masyarakat Terhadap Generasi Z Dalam
Dunia Kerja Pada Media Sosial Twitter Menggunakan Metode
Naïve Bayes**

TUGAS AKHIR

**Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1
Informatika Universitas Muhammadiyah Malang**

Menyetujui,
Malang, 19 Juli 2024

Dosen 1



Christian Sri Kusuma Aditya, S.Kom.,

M.Kom

NIP. 180327021991PNS.

Dosen 2



Didih Rizki Chandranegara, S.kom.,

M.Kom

NIP. 180302101992PNS.

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

**Analisis Sentimen Masyarakat Terhadap Generasi Z Dalam
Dunia Kerja Pada Media Sosial Twitter Menggunakan Metode
Naive Bayes**

TUGAS AKHIR

Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1
Informatika Universitas Muhammadiyah Malang

Disusun Oleh :

VITA AMALIA HERLINDA

202010370311336

Tugas Akhir ini telah diuji dan dinyatakan lulus melalui sidang majelis penguji
pada tanggal 19 Juli 2024

Menyetujui,

Dosen Penguji 1



Ir. Yufis Azhar S.Kom., M.Kom.
NIP. 10814100544PNS.

Dosen Penguji 2



Yuda Munarko S.Kom., M.Sc
NIP. 10806110443PNS.

Mengetahui,
Ketua Jurusan Informatika



Ir. Galih Wasis Wicaksono S.kom. M.Cs.
NIP. 10814100541PNS.

LEMBAR PERNYATAAN

LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini:

NAMA : Vita Amalia Herlinda
NIM : 202010370311336
FAK./JUR. : TEKNIK INFORMATIKA

Dengan ini saya menyatakan bahwa Tugas Akhir dengan judul “**Analisis Sentimen Masyarakat Terhadap Generasi Z Dalam Dunia Kerja Pada Media Sosial Twitter Menggunakan Metode Naive Bayes**” beserta seluruh isinya adalah karya saya sendiri dan bukan merupakan karya tulis orang lain, baik sebagian maupun seluruhnya, kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya ini, atau ada klaim dari pihak lain terhadap keaslian karya saya ini maka saya siap menanggung segala bentuk resiko/sanksi yang berlaku.

Mengetahui
Dosen pembimbing



Christian Sri Kusuma Aditya, S.Kom.,
M.Kom

Malang, 19 Juli 2024

Yang Membuat Pernyataan



VITA AMALIA HERLINDA
202010370311336

DAFTAR ISI

LEMBAR PERSETUJUAN.....	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN.....	iv
DAFTAR ISI.....	iv
DAFTAR GAMBAR	vii
DAFTAR TABEL.....	viii
LEMBAR PERSEMBAHAN	ix
KATA PENGANTAR	x
ABSTRAK.....	xi
ABSTRACT.....	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Batasan Masalah.....	2
Batasan masalah yakni :	2
BAB II TINJAUAN PUSTAKA.....	4
2.1 Penelitian Terdahulu	4
2.2. Generasi Z Dalam Dunia Kerja.....	6
2.3 Analisis Sentimen	7
2.3.1 Inset Lexicon.....	7
2.4 Twitter.....	9
2.5 Ekstrasi Fitur TFIDF.....	9
2.6. Naïve Bayes	11
2.7 Evaluasi Performa	13
2.7.1 Confusion Matrix	13
2.7.2 Classification Report.....	15
2.8 K-Fold Cross Validation	16

BAB III METODE PENELITIAN.....	18
3.1 Skenario Pengujian.....	18
3.1.1 Pengumpulan Data	19
3.1.2 Preprocessing Dataset	20
3.1.3. Klasifikasi Sentimen Dengan Pendekatan Lexicon	22
3.1.4 Visualisasi Sentimen	22
3.1.5 Metode Naïve Bayes	23
BAB IV HASIL DAN PEMBAHASAN	25
4.1 Hasil Analisis Sentimen	25
4.1.1 Hasil Crawling Data Twitter	25
4.1.2 Hasil Preprocessing	25
4.1.3 Hasil Labeling Dengan Inset Lexicon	27
4.1.4 Hasil Visualisasi WordCloud.....	28
4.1.5 Hasil Akurasi Metode Naïve Bayes	33
4.2. Uji Coba	40
4.2.1 Hasil Pengujian Pengaruh Fitur Stopword.....	40
4.2.2 Hasil Pengujian Pelabelan Manual.....	42
BAB V PENUTUP.....	44
5.1 Kesimpulan	44
5.2 Saran.....	44
DAFTAR PUSTAKA	45

DAFTAR GAMBAR

Gambar 1 Skema Pengujian	18
Gambar 2 Bar Chart Distribusi Label Lexicon	28
Gambar 3 Worcloud Sentimen Positif	29
Gambar 4 Wordcloud Sentimen Negatif.....	30
Gambar 5 Wordcloud Sentimen Netral.....	32
Gambar 6 Confusion Matrix 3x3	34
Gambar 7 Classification Report	37
Gambar 8 Bar Chart Distribusi Label Manual	42



DAFTAR TABEL

Tabel 1. Literatur Review	4
Tabel 2 .Confusion Matrix	13
Tabel 3. Confusion Matrix Multiclass	14
Tabel 4 .Case Folding	25
Tabel 5. Cleaning dan Tokenizing	26
Tabel 6. Normalization.....	26
Tabel 7. Stemming	26
Tabel 8. Filtering Stopword	27
Tabel 9. Tweet Sentimen Positif	30
Tabel 10. Tweet Sentimen Negatif.....	31
Tabel 11. Tweet Sentimen Netral.....	33
Tabel 12. Hasil Precision Multiclass Confusion Matrix	35
Tabel 13. Hasil Recall Multiclass Confusion Matrix.....	35
Tabel 14. Hasil F1-Score.....	36
Tabel 15. Cross Validation K-Fold 5	39
Tabel 16. Classification Report Pengujian Stopword	40
Tabel 17. Perbandingan Classification Report Pelabelan Lexicon dan manual....	42

LEMBAR PERSEMBAHAN

Melalui kesempatan kali ini, penulis ingin mengucapkan terima kasih kepada pihak-pihak yang telah memberikan dukungan luar biasa selama proses penulisan, sehingga penulis dapat menyelesaikan tugas akhir ini dengan baik. Ucapan terima kasih sebesar-besarnya penulis sampaikan kepada:

1. Allah SWT, atas rahmat, hidayah, dan kemudahan yang diberikan.
2. Kedua orang tua, yang selalu memberikan cinta, dukungan, doa, dan semangat yang tiada henti-hentinya. Tanpa kasih sayang dan pengorbanan mereka, penulis tidak akan berada di titik ini.
3. Diri sendiri, yang telah bertanggung jawab, bekerja keras, penuh dedikasi dan ketekunan dalam menyelesaikan tugas akhir ini.
4. Kepada bapak Christian Sri Kusuma Aditya, S.Kom., M.Kom selaku dosen Pembimbing 1 dan bapak Didih Rizki Chandranegara, S.kom., M.Kom selaku dosen pembimbing 2, yang telah membimbing dengan sabar, memberikan arahan, saran, dan kritik konstruktif selama proses penulisan tugas akhir ini.
5. Kepada Ir. Galih Wasis Wicaksono, S.Kom, M.Cs., selaku Ketua Prodi Informatika, dan seluruh dosen informatika atas ilmu, bimbingan akademis, dan inspirasinya selama masa studi.
6. Elfrida, Arolina Rinardi, Ria Wulandari, dan teman-teman lainnya, atas bantuan, semangat, dukungan, dan kebersamaan yang tak ternilai.

Malang, 19 Juli 2024

Penulis



Vita Amalia Herlinda

KATA PENGANTAR

Assalamu'alaikum warahmatullahi wabarakatuh,

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT atas rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan tugas akhir dengan judul "*Analisis Sentimen Masyarakat terhadap Generasi Z dalam Dunia Kerja pada Media Sosial Twitter Menggunakan Metode Naive Bayes*". Shalawat dan salam semoga senantiasa tercurahkan kepada Nabi Muhammad SAW, keluarga, sahabat, dan umatnya hingga akhir zaman. Tugas akhir ini disusun sebagai syarat untuk menyelesaikan pendidikan Strata Satu (S1) pada Program Studi Informatika. Penulis berharap penelitian ini dapat memberikan wawasan tentang pandangan masyarakat terhadap generasi Z dalam dunia kerja dan menjadi referensi bagi penelitian selanjutnya.

Penulis menyadari masih terdapat banyak kekurangan dan keterbatasan dalam penulisan skripsi ini. Oleh karena itu, penulis sangat mengharapkan saran dan kritik yang membangun agar tulisan ini dapat bermanfaat untuk perkembangan ilmu pengetahuan dan teknologi informasi ke depannya. Semoga tugas akhir ini memberikan kontribusi positif bagi pengembangan ilmu, khususnya di bidang analisis sentimen dan data mining.

Wassalamu'alaikum warahmatullahi wabarakatuh.

Malang, 19 Juli 2024

Penulis,



Vita Amalia Herlinda

ABSTRAK

Generasi Z, yang lahir antara tahun 1996 sampai 2010, kini mulai memasuki dunia kerja dan dikenal sebagai "Digital Natives" karena kedekatan mereka dengan teknologi dan media sosial seperti Twitter. Mereka aktif berbagi pemikiran dan pengalaman melalui Twitter, yang telah menjadi platform utama untuk diskusi terbuka dan berbagi perspektif. Penelitian ini menggunakan metode Naïve Bayes untuk menganalisis sentimen tweet terkait Generasi Z dalam dunia kerja, dengan memanfaatkan pelabelan lexicon Inset. Hasil penelitian menunjukkan bahwa pelabelan lexicon mencapai akurasi 76%. Selanjutnya, uji coba dilakukan dengan pendekatan pelabelan manual yang menghasilkan akurasi 87%, menunjukkan perbedaan yang signifikan dibandingkan pelabelan lexicon. Uji coba juga melibatkan pengaruh stopword dalam tahapan preprocessing, di mana hasil menunjukkan pelabelan lexicon tanpa stopword mencapai akurasi 96% dan pelabelan manual tanpa stopword mencapai 90%. Hal ini menunjukkan sensitivitas yang tinggi dari model berbasis lexicon terhadap penggunaan stopword dan potensi untuk meningkatkan variasi term dalam analisis sentimen Generasi Z di dunia kerja melalui Twitter.

Kata kunci : *NLP, Analisis sentimen, Naïve bayes, Inset Lexicon, Twitter*

ABSTRACT

Generation Z, born between 1996 and 2010, is now entering the workforce and is known as "Digital Natives" due to their familiarity with technology and social media such as Twitter. They are actively sharing thoughts and experiences through Twitter, which has become the primary platform for open discussion and sharing perspectives. This research employs the Naïve Bayes method to analyze sentiment in tweets related to Generation Z in the workplace, utilizing the Inset lexicon labeling. The results indicate that the Inset lexicon labeling achieved an accuracy of 76%. Furthermore, experiments with manual labeling yielded an accuracy of 87%, demonstrating a significant difference compared to the lexicon-based approach. The study also examined the influence of stopwords during preprocessing, revealing that the lexicon-based labeling without stopwords achieved 96% accuracy, while manual labeling without stopwords reached 90%. These findings highlight the high sensitivity of lexicon-based models to the use of stopwords and their potential to enhance term variation in sentiment analysis of Generation Z in the workplace via Twitter.

Keywords: *NLP, Sentiment Analysis, Naïve Bayes, Inset Lexicon, Twitter*

DAFTAR PUSTAKA

- [1] S. Gaidhani, L. Arora, B. Kumar Sharma, and A. Professor, "Understanding the Attitude of Generation Z Towards Workplace," *Int. J. Manag. Technol. Eng.*, vol. IX, no. I, pp. 2804–2812, 2019, [Online]. Available: <https://www.researchgate.net/publication/331346456>
- [2] McKinsey Explainers, "What Is Gen Z?," *McKinsey Co.*, no. March 2023, p. 5, 2023.
- [3] D. Rachmawati, "Proceeding Indonesia Career Center Network Summit IV e-Welcoming Gen Z in Job World (Selamat Datang Generasi Z di dunia kerja)," *Proceeding Indones. Career Cent. Netw. Summit IV*, pp. 21–24, 2019.
- [4] A. P. Singh and J. Dangmei, "Understanding the Generation Z: the Future Workforce," *South -Asian J. Multidiscip. Stud.*, no. April, 2016.
- [5] A. P. Dewi and S. Delliana, "Self Disclosure Generasi Z Di Twitter," *Ekspresi Dan Persepsi J. Ilmu Komun.*, vol. 3, no. 1, pp. 62–69, 2020, doi: 10.33822/jep.v3i1.1526.
- [6] Fauziah A, "ANALISIS SENTIMEN MENGGUNAKAN NAÏVE BAYES CLASSIFIER DAN INSET LEXICON PADA TWITTER (Studi Kasus: Mie Gacoan)," 2023.
- [7] A. A. A. Sumanjaya, Indriati, and A. Ridok, "Analisis Sentimen Data Tweets terhadap Penanganan Covid-19 di Indonesia menggunakan Metode Naïve Bayes dan Pemilihan Kata Bersentimen menggunakan Lexicon Based," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 4, pp. 1865–1872, 2022, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [8] D. Musfiroh, U. Khaira, P. E. P. Utomo, and T. Suratno, "Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset

- Menggunakan InSet Lexicon,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 24–33, 2021, doi: 10.57152/malcom.v1i1.20.
- [9] J. T. S. Vol, N. Februari, M. Multinomial, and N. Bayes, “Jurnal ferensi 4,” vol. 2, no. 1, pp. 8–15, 2023.
- [10] M. Hamka and D. Ratna Sari, “Analisis Sentimen Dan Information Extraction Pembelajaran Daring Menggunakan Pendekatan Lexicon,” *Djtechno J. Teknol. Inf.*, vol. 3, no. 1, pp. 21–32, 2022, doi: 10.46576/djtechno.v3i1.2194.
- [11] A. H. Dyo fatra, N. H. Hayatin, and C. S. K. Aditya, “Analisa Sentimen Tweet Berbahasa Indonesia Dengan Menggunakan Metode Lexicon Pada Topik Perpindahan Ibu Kota Indonesia,” *J. Repos.*, vol. 2, no. 7, p. 977, 2020, doi: 10.22219/repositor.v2i7.937.
- [12] I. F. N. Fadhillah, A. Herdiani, and W. Astuti, “Analisis Sentimen Berbasis Leksikon InSet Terhadap Partai Politik Peserta Pemilu 2019 Pada Media Sosial Twitter,” *e-Proceeding Eng.*, vol. 6, no. 3, pp. 1–11, 2019.
- [13] F. Amaliah and I. K. Dwi Nuryana, “Perbandingan Akurasi Metode Lexicon Based Dan Naive Bayes Classifier Pada Analisis Sentimen Pendapat Masyarakat Terhadap Aplikasi Investasi Pada Media Twitter,” *J. Informatics Comput. Sci.*, vol. 3, no. 03, pp. 384–393, 2022, doi: 10.26740/jinacs.v3n03.p384-393.
- [14] T. A. Mutiara, Y. Alkhalifi, N. A. Mayangky, and W. Gata, “Analisis Sentimen Opini Publik Mengenai Larangan Mudik pada Twitter Menggunakan Naïve Bayes,” *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer dan Teknologi Informasi*, vol. 6, no. 2. p. 85, 2020. doi: 10.24014/coreit.v6i2.9727.
- [15] Y. Nooryuda Prasetya, D. Winarso, and Syahril, “Penerapan Lexicon Based Untuk Analisis Sentimen Pada Twiter Terhadap Isu Covid-19,” *J. Fasilkom*,

vol. 11, no. 2, pp. 97–103, 2021.

- [16] T. N. Wijaya, R. Indriati, and M. N. Muzaki, “Analisis Sentimen Opini Publik Tentang Undang-Undang Cipta Kerja Pada Twitter,” *Jambura J. Electr. Electron. Eng.*, vol. 3, no. 2, pp. 78–83, 2021, doi: 10.37905/jjee.v3i2.10885.
- [17] J. D. K. Liza Wikarsa, Apriandy Angdresey, “Implementasi Metode Naïve Bayes dan Lexicon-Based Approach untuk Mengklasifikasi Sentimen Netizen pada Tweet Berbahasa Indonesia,” *J. Ilm. Realt.*, vol. 18, pp. 1–10, 2022, [Online]. Available: <https://ejournal.unikadelasalle.ac.id/index.php/realtech/article/view/5>
- [18] I. G. A. Indrawan and D. Dewi, “Analisis Sentimen Terhadap Presidensi G20 2022 pada Media Sosial Twitter Menggunakan Metode Naïve Bayes,” *KLIK Kaji. Ilm.*, vol. 4, no. 1, pp. 553–561, 2023, doi: 10.30865/klik.v4i1.1104.
- [19] A. F. Hidayatullah, “Pengaruh Stopword Terhadap Performa Klasifikasi Tweet Berbahasa Indonesia,” *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 1, no. 1, pp. 1–4, 2016, doi: 10.14421/jiska.2016.11-01.
- [20] S. Khairunnisa, A. Adiwijaya, and S. Al Faraby, “Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19),” *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 406, 2021, doi: 10.30865/mib.v5i2.2835.



FAKULTAS TEKNIK

INFORMATIKA

informatika.umm.ac.id | informatika@umm.ac.id

FORM CEK PLAGIARISME LAPORAN TUGAS AKHIR

Nama Mahasiswa : Vita Amalia Herlinda

NIM : 202010370311336

Judul TA : Analisis Sentimen Masyarakat Terhadap Generasi Z Dalam Dunia Kerja Pada Media Sosial Twitter Menggunakan Metode Naïve Bayes.

Hasil Cek Plagiarisme dengan Turnitin

No.	Komponen Pengecekan	Nilai Maksimal Plagiarisme (%)	Hasil Cek Plagiarisme (%) *
1.	Bab 1 – Pendahuluan	10 %	10%
2.	Bab 2 – Daftar Pustaka	25 %	8%
3.	Bab 3 – Analisis dan Perancangan	25 %	4%
4.	Bab 4 – Implementasi dan Pengujian	15 %	0%
5.	Bab 5 – Kesimpulan dan Saran	5 %	0%
6.	Makalah Tugas Akhir	20%	4%

*) Hasil cek plagiarisme diisi oleh pemeriksa (staf TU)

*) Maksimal 5 kali (4 Kali sebelum ujian, 1 kali sesudah ujian)

Mengetahui,

Pemeriksa (Staf TU)



Kampus I
Jl. Sekeloa Timur 1 Malang, Jawa Timur
T. +62 341 551 203 (Kampus I)
F. +62 341 480 433

Kampus II
Jl. Perhutangan Selatan No. 100 Malang, Jawa Timur
T. +62 341 551 142 (Kampus II)
F. +62 341 582 180

Kampus III
Jl. Widyadarmas No. 700 Malang, Jawa Timur
T. +62 341 551 319 (Kampus III)
F. +62 341 481 431
E. info@informatika@umm.ac.id