

**ANALISIS SENTIMEN TENTANG KESEHATAN MENTAL PADA
GENERASI Z MENGGUNAKAN METODE NAIVE BAYES
PADA SOSIAL MEDIA TWITTER**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

SHAFIRA HENTIHU

20.11.3576

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

**ANALISIS SENTIMEN TENTANG KESEHATAN MENTAL PADA
GENERASI Z MENGGUNAKAN METODE NAIVE BAYES
PADA SOSIAL MEDIA TWITTER**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

SHAFIRA HENTIHU

20.11.3576

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN TENTANG KESEHATAN MENTAL PADA GENERASI
Z MENGGUNAKAN METODE NAIVE BAYES
PADA SOSIAL MEDIA TWITTER**

yang disusun dan diajukan oleh

Shafira Hentihu

20.11.3576

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 13 Mei 2024

Dosen Pembimbing,



Windha Mega Pradnya Duhita, M.Kom
NIK. 190302185

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS SENTIMEN TENTANG KESEHATAN MENTAL PADA GENERASI
Z MENGGUNAKAN METODE NAIVE BAYES
PADA SOSIAL MEDIA TWITTER

yang disusun dan diajukan oleh

Shafira Hentihu

20.11.3576

Telah dipertahankan di depan Dewan Penguji
pada tanggal 13 Mei 2024

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Anna Baita, M.Kom

NIK. 190302290

Ike Verawati, M.Kom

NIK. 190302237

Windha Mega Pradnya, M.Kom

NIK. 190302185

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 13 Mei 2024

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.

NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : SHAFIRA HENTIHU
NIM : 20.11.3576

Menyatakan bahwa Skripsi dengan judul berikut:

**ANALISIS SENTIMEN TENTANG KESEHATAN MENTAL PADA GENERASI Z
MENGUNAKAN METODE NAIVE BAYES PADA SOSIAL MEDIA TWITTER**

Dosen Pembimbing : Whindha Mega Pradnya Duhita, M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta,

Yang Menyatakan,


Shafira Hentihu

HALAMAN PERSEMBAHAN

Alhamdulillah segala puji Syukur atas rahmat dan karunia Allah SWT yang telah memberikan petunjuk dan kemudahan bagi penulis dalam menyelesaikan skripsi ini dengan sebaik-baiknya. Dengan ini penulis mempersembahkan karya kecil dan sederhana dalam berbentuk skripsi kepada semua pihak yang telah membantu dengan memberikan dukungan penuh sehingga bisa sampai di tahap ini.

Dengan ini penulis memberikan ucapan terima kasih lewat halaman persembahan ini kepada :

1. Dengan penuh rasa Syukur kepada Allah SWT, bangunku, tidurku, doaku, sujudku, bahagiaku, dan sedihku aku tujukan kepada Allah SWT yang selalu melindungi dan menerangi tiap jalanku.
2. Kepada kedua orang tua penulis Ayah Abd Madjid Hentihu , Umi Gamar Alkatiri, yang selalu berdoa siang malam dan membiayai studiku meskipun semua itu di lakukan dengan penuh pengorbanan dan kesengsaraan dunia tetapi selalu kuat memberikan motivasi dan contoh dalam membentuk karakterku.
3. Kepada tante penulis Iliya Alkatiri (ye) yang bersedia bersama-sama memberikan motivasi dan dorongan serta ucapan doa di setiap sujudnya.
4. Kepada adik tersayang Maulana Rafqi Hentihu, adik semata wayang penulis terima kasih suda selalu menjadi tempat keluh kesah selama ini.
5. Kepada saudara yang tak sedarah yaitu Ano Surniah, Nazifa Alkatiri, Vanesa Rebecca, Sasya Hentihu, dan juga kepada IKPMKB-YK yang senantiasa membantu dan menemani penulis dalam segala kesusahan.
6. Kepada sahabat penulis semasa sekolah yaitu Geng Ladies, Aksha Tukuboya, Fara Almahdaly, Aulia Nurhidayah, Yunisa Bastiman, Nana Hentihu, Dila Tukuboya, Saputri Assagaf, Ira Djogja, yang menemani penulis, memberikan dukungan serta mendengarkan segala curhatan penulis.
7. Kepada seseorang yang memiliki NIM 1320204009 dengan inisial nama (HTN) yang memberikan dukungan serta kasih sayang kepada penulis.

8. Untuk seseorang yang belum bisa kutulis dengan jelas Namanya disini, namun sudah tertulis jelas di *Lauhul Mahfudz* untukku. Terima kasih suda menjadi motivasi penulis dalam menyelesaikan skripsi ini sebagai salah satu bentuk memantaskan diri. Meskipun saat ini penulis tidak tahu keberadaanmu entah di bumi bagian mana dan menggenggam tangan siapa. Seperti kata Bj Habibie “Kalau memang dia dilahirkan untuk saya kamu jungkir balik pun saya yang dapat”.

Last but not least. Terima kasih untuk Shafira Hentihu, diri saya sendiri yang telah bekerja keras dan berjuang sejauh ini. “Orang lain tidak akan paham *struggle* masa sulit kita, yang mereka *ingin* tahu hanya bagian *success storiesnya* saja. Jadi berjuanglah untuk diri sendiri meskipun tidak akan ada yang tepuk tangan. Kelak diri kita di masa depan akan sangat bangga dengan apa yang kita perjuangkan hari ini jadi tetaplah berjuang karena hidupmu masih panjang”. Dan terima kasih pada diri sendiri karena mampu mengendalikan diri dari berbagai tekanan diluar keadaan dan tak pernah memutuskan menyerah sesulit apapun proses penyusunan skripsi ini dengan menyelesaikan sebaik dan semaksimal mungkin, ini merupakan pencapaian yang patut di banggakan untuk diri sendiri.

MOTTO

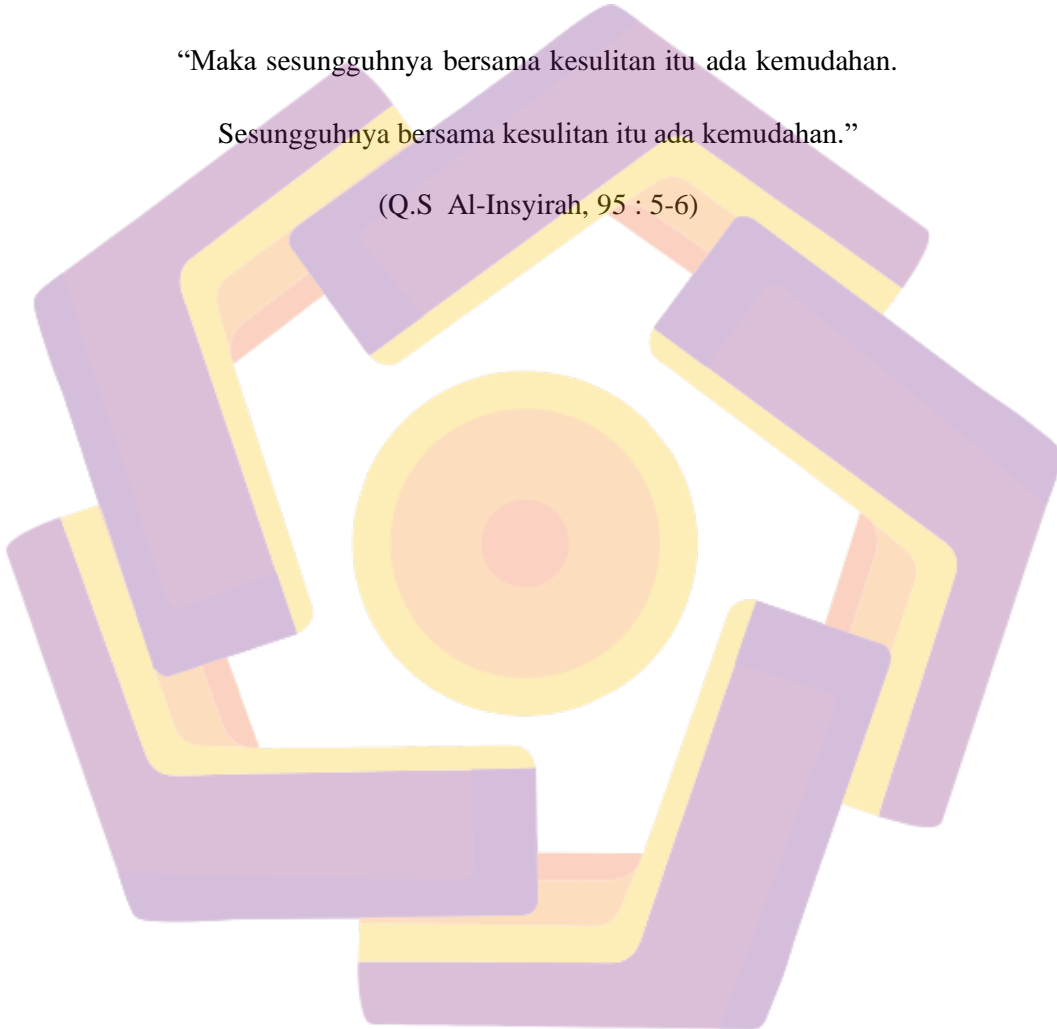
“Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya. Dia mendapat (pahala) dari (kebijakan) yang di kerjakannya dan mendapat (siksa) dari (kejahatan) yang diperbuatnya.”

(Q.S Al-Baqarah : 286)

“Maka sesungguhnya bersama kesulitan itu ada kemudahan.

Sesungguhnya bersama kesulitan itu ada kemudahan.”

(Q.S Al-Insyirah, 95 : 5-6)



KATA PENGANTAR

Alhamdulillah rabbi 'alamin, segala puji dan Syukur penulis panjatkan atas kehadiran Allah subhanahu wa ta 'ala. Dzat yang hanya kepada-Nya memohon pertolongan dengan segala rahmat dan karunia-Nya sehingga penulis mampu menyelesaikan skripsi dengan judul "ANALISIS SENTIMEN TENTANG KESEHATAN MENTAL PADA GENERASI Z MENGGUNAKAN METODE NAIVE BAYES PADA SOSIAL MEDIA TWITTER". Skripsi ini dapat di selesaikan guna memenuhi salah satu persyaratan dan menyelesaikan Pendidikan untuk mencapai Gelar Sarjana pada Fakultas Ilmu Komputer Program Studi Informatika Universitas Amikom Yogyakarta.

Perjalanan panjang telah penulis lalui Dalam rangka menyelesaikan penulisan skripsi ini, banyak hal yang telah dihadapi dalam penyusunan skripsi ini, namun berkat kehendak Allah SWT sehingga penulis berhasil menyelesaikan skripsi ini. Oleh karena itu, dengan penuh kerendahan hati, pada kesempatan ini patutlah kiranya penulis mengucapkan banyak terimakasih kepada:

1. Bapak prof. Dr. M. Suyanto, M.M. selaku rector Universitas Amikom Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom., M.Kom., Ph.D. selaku Dekan fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Windha Mega Pradnya D, M.Kom. selaku ketua program Studi S1 Informatika Universitas Amikom Yogyakarta. Dan juga selaku dosen pembimbing saya yang telah memberikan waktu, saran, kritik motivasi dan bimbingan dalam skripsi ini sehingga dapat menyelesaikan skripsi selesai dengan baik.
4. Bapak Abdul Madjid dan Ibu Gamar Alkatiri selaku orang tua penulis yang memberikan dukungan, doa dan motivasi kepada penulis. Terimakasih telah sabar membesarkan penulis dengan

penyuh kasih sayang dan cinta yang besar, terimakasih sudah berjuang untuk membantu penulis mendapatkan Pendidikan yang sangat baik, terimakasih telah mendukung tanpa syarat.

5. Ibu Iliya Alkatiri (ye) selaku tante penulis, yang selalu ada disaat penulis merasa sedih dan merasa sangat terpuruk, terimakasih atas semua perjuangan, cinta dan kasih sayang kepada penulis, terimakasih sudah sangat sabar membantu dan membesarkan penulis dari kecil hingga dewasa sekarang. Dan terimakasih atas didikan yang selama ini berhasil membuat penulis bisa mampu melewati berbagai cobaan dan bisa membedakan hal yang baik dan buruk selama menjalankan kehidupan. Tidak ada kata-kata yang akan cukup mengatakan betapa bersyukurya penulis memiliki ye.

Penulis menyadari bahwa skripsi ini masih banyak kekurangan karena keterbatasan pengetahuan yang dimiliki oleh penulis. Akhir kata penulis mengharapkan skripsi ini dapat memberikan manfaat bagi penulis khususnya dan pembaca pada umumnya.

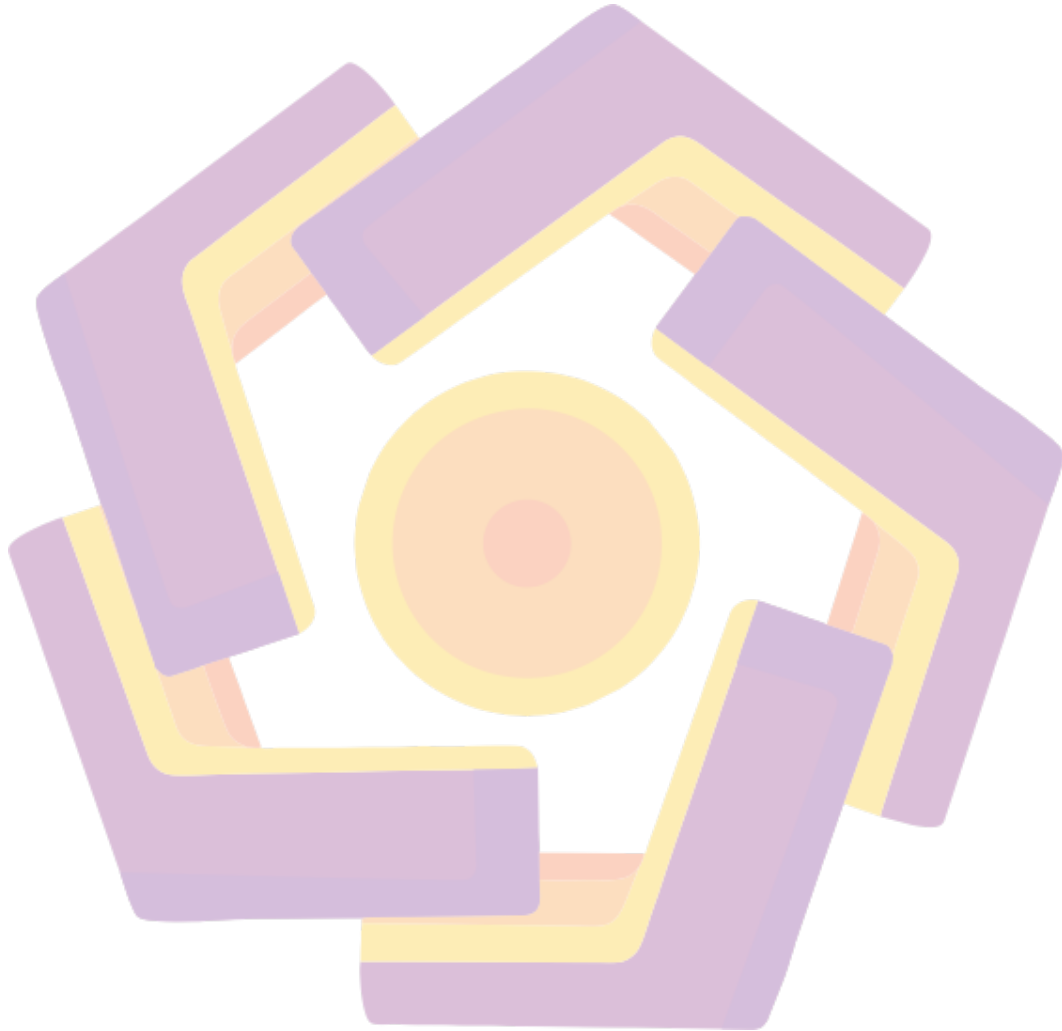
Yogyakarta, 23 April 2024

Penulis

DAFTAR ISI

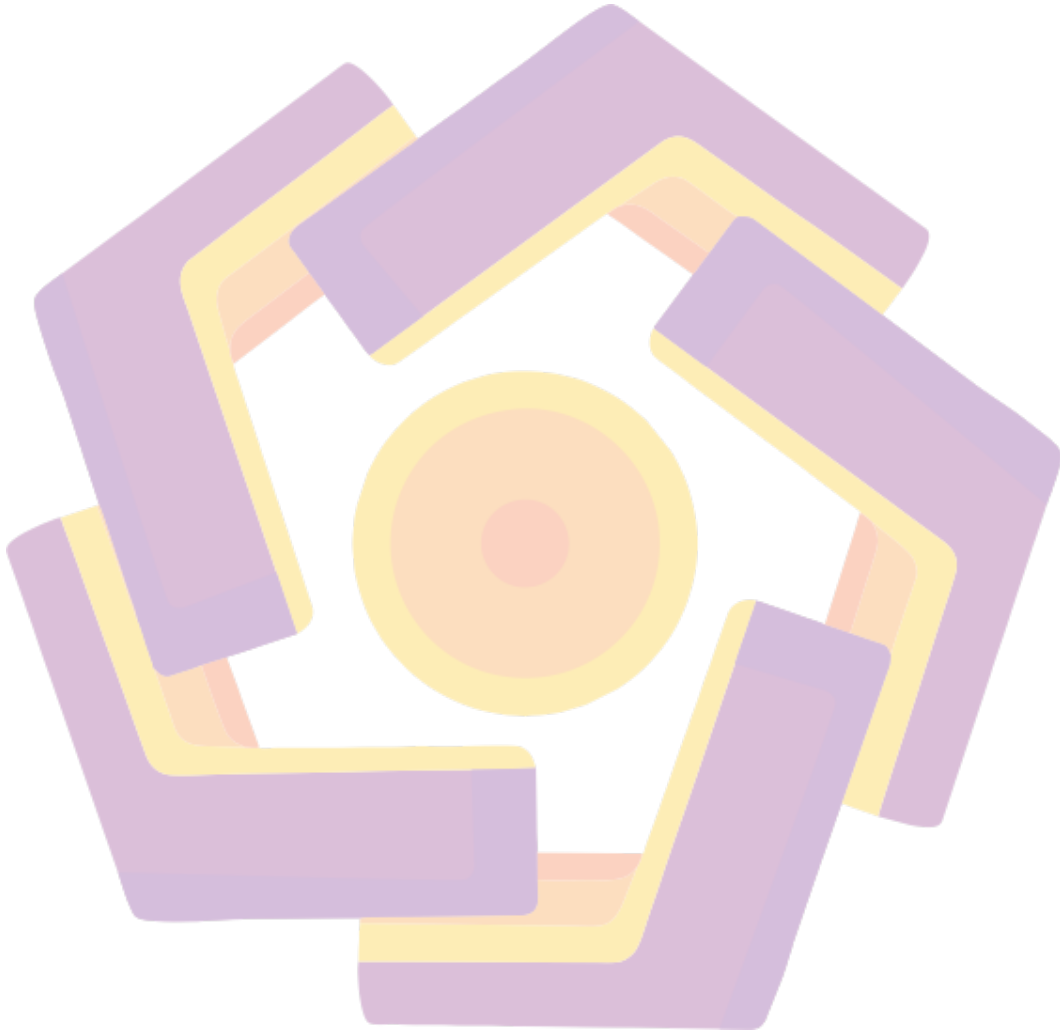
DAFTAR ISI	1
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH.....	xiv
INTISARI	xvi
ABSTRACT	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur.....	5
2.2 Dasar Teori	10
2.2.1 Kesehatan Mental Generasi Z.....	10
2.2.2 Analisis Sentimen pada Data Tweet.....	11
2.2.3 Python	12
2.2.4 Ekstraksi Fitur N-Grams	13
2.2.5 Teknik Synthetic Minority Oversampling Technique (SMOTE).....	14
2.2.6 Algoritma Naïve Bayes	15
2.2.7 Metode Latent Dirichlet Allocation (LDA)	16
2.2.8 Metode Evaluasi.....	17
BAB III METODE PENELITIAN	19
3.1 Objek Penelitian	19
3.2 Alur Penelitian.....	19
3.3 Kebutuhan Sistem.....	26
3.3.1 Perangkat Lunak.....	26
3.3.2 Perangkat Keras	ix

BAB IV HASIL DAN PEMBAHASAN.....	27
4.1 Hasil Penelitian.....	27
4.2 Pembahasan	45
BAB V KESIMPULAN.....	46
5.1 Kesimpulan.....	46
5.2 Saran	46
DAFTAR PUSTAKA.....	48



DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian.....	5
-------------------------------------	---

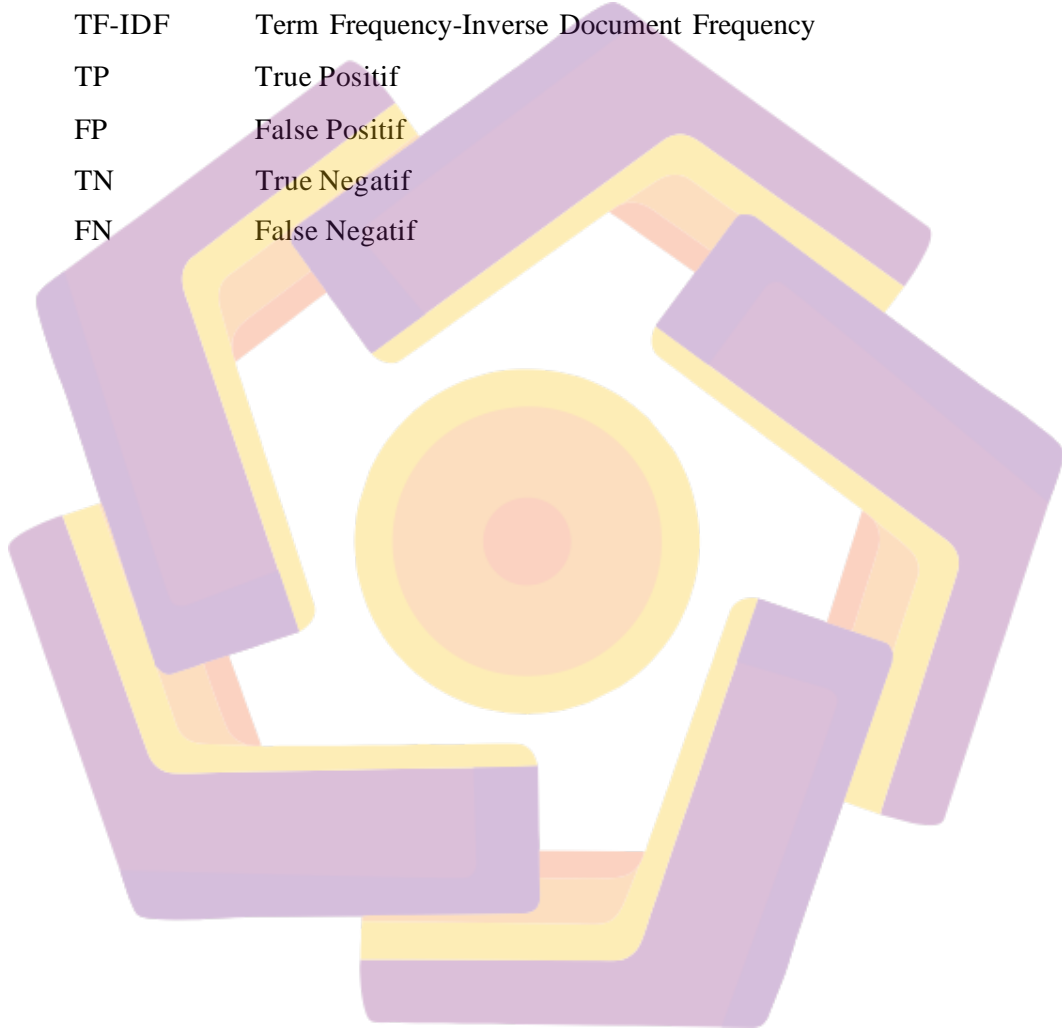


DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian.....	19
Gambar 3. 2 Dataset	20
Gambar 3. 3 Tabel Confusion Matrix.....	25
Gambar 4. 1 Hasil Pelabelan Data	27
Gambar 4. 2 Hasil Pembacaan Dataset.....	29
Gambar 4. 3 Hasil Visualisasi Data Sentiment	30
Gambar 4. 4 Hasil Case Folding dan Cleaning	31
Gambar 4. 5 Hasil Jumlah Data Berdasarkan Label Sentiment Setelah Ekstraksi Fitur dan Penyeimbangan Data	35
Gambar 4. 6 Hasil Classification Report Sebelum Teknik SMOTE	36
Gambar 4. 7 Hasil Classification Report Setelah Teknik SMOTE.....	37
Gambar 4. 8 Hasil Confusion Matrix	38
Gambar 4. 9 Hasil Classification Report dengan Optimasi SMOTE	41
Gambar 4. 10 Hasil Confusion Matrix dengan Optimasi SMOTE.....	42

DAFTAR LAMBANG DAN SINGKATAN

LDA	Latent Dirichlet Allocation
SMOTE	Synthetic Minority Over-sampling Technique
TF-IDF	Term Frequency-Inverse Document Frequency
TP	True Positif
FP	False Positif
TN	True Negatif
FN	False Negatif



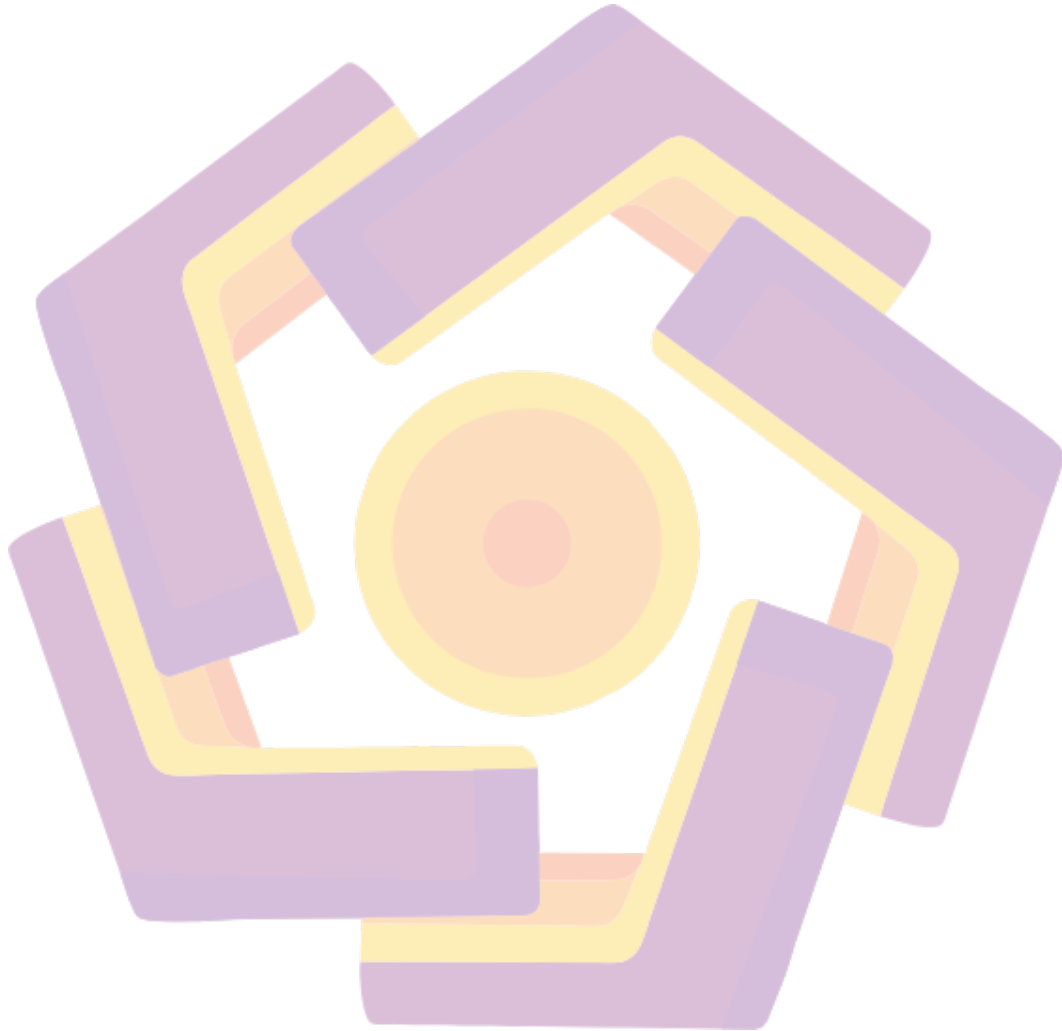
DAFTAR ISTILAH



Analisis Sentimen	Proses analisa terhadap sentimen tertentu terhadap suatu isu
Sentiment	Pendapat suatu orang pada suatu akun di media sosial
Dataset	Data yang akan diolah
Tweet	Cuitan sebuah akun pada twitter
Labelling	Proses pemberian label pada data
Pre-processing	Proses awal sebelum pengolahan data
Akurasi	Ukuran seberapa baik algoritma dapat mengklasifikasikan sentimen dari data dengan benar
Ekstraksi Fitur	Proses mengubah data mentah menjadi representasi yang dapat digunakan oleh model atau algoritma pembelajaran mesin
N-Grams	Sekumpulan n kata yang diambil secara berurutan dari sebuah teks atau dokumen
<i>Precision, recall, f1-score, support</i>	Metrik evaluasi yang umum digunakan dalam pengukuran kinerja model klasifikasi
Case Folding	Proses normalisasi teks dengan mengubah semua karakter huruf menjadi huruf kecil atau huruf besar
Cleaning	Langkah-langkah membersihkan teks dari karakter khusus, tanda baca, atau informasi yang tidak relevan sebelum atau selama pemrosesan
Tokenization	Proses memecah teks menjadi bagian-bagian yang lebih kecil yang disebut token
Stemming	Proses menghilangkan imbuhan dari kata-kata untuk mengembalikan akar kata

Confusion Matrix

Tabel yang digunakan untuk mengevaluasi kinerja model klasifikasi dengan membandingkan hasil prediksi model dengan nilai sebenarnya dari data



INTISARI

Masalah kesehatan mental Generasi Z telah menjadi fokus penelitian karena pengaruh lingkungan digital dan dinamika sosial yang memengaruhi pola pikir dan perilaku mereka. Penggunaan teknologi digital dan media sosial oleh Generasi Z menciptakan tantangan baru dalam memahami dan mengatasi masalah kesehatan mental mereka. Penelitian sebelumnya menunjukkan bahwa algoritma klasifikasi sentimen seperti Naive Bayes masih memiliki kelemahan dalam menghadapi teks yang kompleks dan ambigu terkait kesehatan mental Generasi Z di Twitter. Untuk mengatasi masalah ini, pendekatan yang menggabungkan ekstraksi fitur n-grams dan penyeimbang data Synthetic Minority Over-sampling Technique (SMOTE) diajukan untuk meningkatkan akurasi algoritma Naive Bayes. Penelitian ini bertujuan untuk memberikan kontribusi yang signifikan dalam pemahaman dan deteksi dini masalah kesehatan mental Generasi Z serta menghasilkan rekomendasi intervensi yang lebih efektif.

Metode penelitian yang digunakan mencakup pengumpulan data dari Twitter, pelabelan manual sentimen pada data, dan penerapan teknik SMOTE untuk mengatasi ketidakseimbangan kelas dalam data. Hasil penelitian ini menunjukkan peningkatan signifikan dalam kinerja model klasifikasi Naive Bayes setelah penerapan teknik SMOTE, dengan tingkat akurasi yang meningkat dari 63% menjadi 74%. Setelah optimasi dengan SMOTE menggunakan GridSearchCV, tingkat akurasi model meningkat menjadi 99%, menunjukkan peningkatan yang signifikan dalam prediksi kelas positif. Selain itu, analisis topik menggunakan model LDA menemukan hubungan yang kuat antara kesehatan mental dan Generasi Z dalam dataset Twitter yang diamati.

Kata kunci: Kesehatan Mental, Generasi Z, Naive Bayes, SMOTE, LDA.

ABSTRACT

The mental health issues among Generation Z have become the focus of research due to the influence of digital environments and social dynamics shaping their patterns of thinking and behavior. The utilization of digital technology and social media by Generation Z has posed new challenges in understanding and addressing their mental health issues. Previous studies have indicated that sentiment classification algorithms like Naive Bayes still have weaknesses in dealing with the complex and ambiguous texts related to Generation Z's mental health on Twitter. To address this issue, an approach combining n-grams feature extraction and Synthetic Minority Over-sampling Technique (SMOTE) for data balancing is proposed to enhance the accuracy of the Naive Bayes algorithm. This research aims to provide a significant contribution to understanding and early detection of Generation Z's mental health issues and to generate more effective intervention recommendations.

The research methodology includes data collection from Twitter, manual sentiment labeling of the data, and the application of SMOTE technique to address class imbalance in the data. The results of this study show a significant improvement in the performance of the Naive Bayes classification model after the application of the SMOTE technique, with the accuracy rate increasing from 63% to 74%. After optimization with SMOTE using GridSearchCV, the model accuracy rate further increased to 99%, indicating a significant improvement in predicting positive classes. Furthermore, topic analysis using the LDA model found a strong relationship between mental health and Generation Z in the observed Twitter dataset.

Keyword: Mental Health, Generation Z, Naïve Bayes, SMOTE, LDA