



**ANALISIS SENTIMEN KOMENTAR POSTINGAN
INSTAGRAM TERKAIT TILANG UJI EMISI
MENGUNAKAN ALGORITMA K-NEAREST NEIGHBORS
DAN NAÏVE BAYES DENGAN EVALUASI PENGGUNAAN
TEKNIK SMOTE**

LAPORAN TUGAS AKHIR

**UNIVERSITAS
MERCU BUANA**
RYAN ADYTIA 41520010137

**PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS
ILMU KOMPUTER
UNIVERSITAS MERCUBUANA
JAKARTA 2024**



**ANALISIS SENTIMEN KOMENTAR POSTINGAN
INSTAGRAM TERKAIT TILANG UJI EMISI
MENGUNAKAN ALGORITMA K-NEAREST NEIGHBORS
DAN NAÏVE BAYES DENGAN EVALUASI PENGGUNAAN
TEKNIK SMOTE**



LAPORAN TUGAS AKHIR

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

MERCU BUANA

RYAN ADYTIA 41520010137

**PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS
ILMU KOMPUTER
UNIVERSITAS MERCUBUANA
JAKARTA 2024**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Ryan Adytia
NIM : 41520010137
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Sentimen Komentar Postingan Instagram
Terkait Tilang Uji Emisi Menggunakan Algoritma K-Nearest Neighbors Dan Naïve Bayes Dengan
Evaluasi Penggunaan Teknik Smote

Menyatakan bahwa Laporan Tugas Akhir ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 4 Juni 2024



Ryan Adytia

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama Mahasiswa : Ryan Adytia
NIM : 41520010137
Fakultas : Ilmu Komputer
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Sentimen Komentar Postingan Instagram
Terkait Tilang Uji Emisi Menggunakan Algoritma K-Nearest Neighbors Dan
Naive Bayes Dengan Evaluasi Penggunaan Teknik Smote

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing	:	Umniy Salamah, S.T, MMSI	()
NIDN	:	0306098104	
Ketua Sidang	:	Dr. Hadi Santoso, S.Kom., M.Kom	()
NIDN	:	0225067701	
Penguji 1	:	Anis Cherid, SE,MTI	()
NIDN	:	0328127203	
Penguji 2	:	Sabar Rudiarto, S.Kom, M.Kom	()
NIDN	:	0309036902	

Jakarta, 4 Juni 2024
Mengetahui,


Dekan

Ketua Program Studi



Dr. Bambang Jokonowo, S.Si., M.TI.

NIDN: 0320037002



Dr. Hadi Santoso, S.Kom., M.Kom.

NIDN: 0225067701

KATA PENGANTAR

Puji syukur saya panjatkan kepada Allah SWT, Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Laporan Skripsi ini. Penulisan Laporan Skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan Laporan Skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Prof. Dr. Ir. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana;
2. Dr. Bambang Jokonowo, S.Si., M.TI. selaku Dekan Fakultas Ilmu Komputer Universitas Mercu Buana;
3. Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika;
4. Umniy Salamah, S.T, MMSI. selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini;
5. Anis Cherid, SE,MTI. selaku Dosen Penguji Tugas Akhir atas koreksi dan arahan serta masukannya.
6. Sabar Rudiarto, S.Kom, M.Kom. selaku Dosen Penguji Tugas Akhir atas koreksi dan arahan serta masukannya.

Akhir kata, saya berharap Allah SWT, Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Laporan Skripsi ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 4 Juni 2024



Ryan Adytia

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Ryan Adytia
NIM : 41520010137
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Sentimen Komentar Postingan Instagram
Terkait Tilang Uji Emisi Menggunakan Algoritma K-Nearest Neighbors Dan Naïve Bayes Dengan Evaluasi Penggunaan Teknik Smote

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 4 Juni 2024

Yang menyatakan,



Ryan Adytia

ABSTRAK

Nama : Ryan Adytia
NIM : 41520010137
Program Studi : Teknik Informatika
Judul Penelitian : Analisis Sentimen Komentar Postingan Instagram
Terkait Tilang Uji Emisi Menggunakan Algoritma K-Nearest Neighbors Dan Naïve
Bayes Dengan Evaluasi Penggunaan Teknik Smote
Pembimbing : Umniy Salamah, S.T, MMSI

Penelitian ini mengeksplorasi analisis sentimen terhadap komentar postingan Instagram yang berkaitan dengan tilang uji emisi menggunakan algoritma K-Nearest Neighbors (KNN) dan Naïve Bayes, serta mengevaluasi penggunaan teknik Synthetic Minority Over-sampling Technique (SMOTE). Tujuannya adalah untuk memahami persepsi dan respons publik terhadap kebijakan tilang uji emisi dengan menganalisis komentar di Instagram, yang merupakan platform media sosial yang populer di Indonesia. Penelitian ini memberikan kontribusi pada analisis sentimen dengan menerapkan metode pembelajaran mesin yang canggih dan menyediakan wawasan tentang efektivitas algoritma KNN dan Naïve Bayes dalam menganalisis sentimen. Hasil penelitian menunjukkan bahwa kedua algoritma, KNN dan Naïve Bayes, dapat mengklasifikasikan sentimen dengan akurasi tinggi ketika tidak menggunakan SMOTE, namun kinerja pada kelas minoritas (sentimen positif) lebih rendah, yang menunjukkan perlunya penelitian lebih lanjut tentang teknik untuk mengatasi ketidakseimbangan kelas secara efektif. Penggunaan SMOTE tidak signifikan meningkatkan kinerja algoritma KNN tetapi menunjukkan peningkatan kecil dalam recall untuk kelas positif dalam algoritma Naïve Bayes. Kesimpulannya, sementara KNN dan Naïve Bayes efektif dalam klasifikasi sentimen, masih ada kebutuhan untuk penelitian lebih lanjut tentang metode untuk menangani ketidakseimbangan kelas.

Kata Kunci : K-Nearest Neighbor, Naïve Bayes, Machine Learning, Analisis Sentimen, Instagram, Denda Uji Emisi, SMOTE, Ketidakseimbangan Kelas.

ABSTRACT

Name : Ryan Adytia
NIM : 41520010137
Study Program : Informatics Engineering
Title Search : Sentiment Analysis Of Instagram Post Comments Related To Emission Test Files Using The K-Nearest Neighbors And Naïve Bayes Algoritma With Evaluation Of The Use Of The Smote Technique
Preceptor : Umniy Salamah, S.T, MMSI

This research explores sentiment analysis of Instagram post comments related to emissions test tickets using the K-Nearest Neighbors (KNN) and Naïve Bayes algorithms, as well as evaluating the use of the Synthetic Minority Over-sampling Technique (SMOTE) technique. The aim is to understand public perceptions and responses to the emission test ticket policy by analyzing comments on Instagram, which is a popular social media platform in Indonesia. This research contributes to sentiment analysis by applying advanced machine learning methods and providing insights into the effectiveness of KNN and Naïve Bayes algorithms in analyzing sentiment. The results show that both algorithms, KNN and Naïve Bayes, can classify sentiment with high accuracy when not using SMOTE, but the performance on the minority class (positive sentiment) is lower, indicating the need for further research on techniques to effectively overcome class imbalance. The use of SMOTE does not significantly improve the performance of the KNN algorithm but shows a small increase in recall for the positive class in the Naïve Bayes algorithm. In conclusion, while KNN and Naïve Bayes are effective in sentiment classification, there is still a need for further research on methods to deal with class imbalance.

Keywords : K-Nearest Neighbor, Naïve Bayes, Machine Learning, Sentiment Analysis, Instagram, Emission Test Fines, SMOTE, Class Imbalance.

DAFTAR ISI

HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN	iii
KATA PENGANTAR	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI	viii
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
DAFTAR RUMUS	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN	1
1.1. Latar Belakang.....	1
1.2. Perumusan Masalah.....	3
1.3. Tujuan Penelitian.....	4
1.4. Manfaat Penelitian.....	4
1.5. Batasan Penelitian.....	5
BAB II TINJAUAN PUSTAKA	6
2.1 Penelitian Terdahulu.....	6
2.2 Teori Pendukung.....	28
BAB III METODE PENELITIAN	39
3.1 Jenis Penelitian.....	39
3.2 Tahapan Penelitian.....	39
BAB IV PEMODELAN	41
4.1 Visualisasi Data	41
4.2 Dataset	41
4.3 Pre-Processing	46
4.4 Pembuatan Model.....	48
4.5 Evaluasi Algoritma	56
BAB V KESIMPULAN DAN SARAN	62
5.1. Kesimpulan.....	62
5.2. Saran.....	62
DAFTAR PUSTAKA.....	64
LAMPIRAN.....	66

DAFTAR TABEL

Tabel 2.1 Daftar Penelitian Terdahulu.....	6
Tabel 2.2 Data contoh perhitungan jarak Euclidean	33
Tabel 2.3 Data contoh perhitungan Naive Bayes	36
Tabel 2.4 Data contoh perhitungan SMOTE.....	38
Tabel 4.1 Contoh dataset komentar.....	45
Tabel 4.2 Pembersihan data.....	46
Tabel 4.3 Normalisasi data.....	47
Tabel 4.4 Tokenisasi data	47
Tabel 4.5 Stopword data.....	47
Tabel 4.6 Stemming data.....	48
Tabel 4.7 Confusion Matrix	56



DAFTAR GAMBAR

Gambar 2.1 Hasil perhitungan jarak Euclidean	34
Gambar 2.2 Contoh perhitungan Prior	36
Gambar 2.3 Contoh perhitungan Likelihood.....	36
Gambar 2.4 Contoh perhitungan Posterior.....	37
Gambar 2.5 Hasil perhitugan Posterior.....	37
Gambar 2.6 Hasil perhitungan SMOTE.....	38
Gambar 3.1 Flow tahapan penelitian.....	39
Gambar 4.1 Pembagian data positif, netral dan negatif	41
Gambar 4.2 Contoh komentar warga Jakarta.....	43
Gambar 4.3 Tahap Preprocessing.....	51
Gambar 4.4 Pembuatan model K-Nearest Neighbor.....	53
Gambar 4.5 Pembuatan model Naive Bayes.....	55
Gambar 4.6 Hasil Confusion Matrix KNN	58
Gambar 4.7 Hasil Confusion Matrix Naive Bayes.....	60



DAFTAR RUMUS

Rumus 2. 1 Rumus TF.....	31
Rumus 2. 2 Rumus IDF.....	31
Rumus 2. 3 Rumus TF-IDF.....	31
Rumus 2. 4 Rumus Jarak Euclidean.....	33
Rumus 2. 5 Teorema Bayes.....	34
Rumus 2. 6 Representasi sifat independen suatu kelas fitur	35
Rumus 2. 7 Persamaan Prior	35
Rumus 2. 8 Persamaan Likelihood.....	36
Rumus 2. 9 Rumus SMOTE.....	37
Rumus 4. 1 Rumus Precision.....	57
Rumus 4. 2 Rumus Recall.....	57
Rumus 4. 3 Rumus F1 Score.....	57
Rumus 4. 4 Rumus Accuracy.....	58



DAFTAR LAMPIRAN

Lampiran 1 Lampiran Bimbingan.....	66
Lampiran 2 Lampiran CV	67
Lampiran 3 Lampiran HKI.....	68
Lampiran 4 Lampiran BNSP.....	70
Lampiran 5 Halaman Persetujuan	72
Lampiran 6 Form Revisi Dosen Penguji.....	73

