

**ANALISA SENTIMEN TWITTER TERHADAP PENGGUNAAN ALAT
DETEKSI COVID-19 GENOSE MENGGUNAKAN METODE KNN**

Tugas Akhir disusun untuk memenuhi syarat
mencapai gelar Kesarjanaan Komputer pada
Program Studi Teknik Informatika
Jenjang Program Strata-1



oleh:

MUHAMMAD RIZQI AKBAR

17.01.53.0004

22747

**FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS STIKUBANK (UNISBANK) SEMARANG**

2021

PERNYATAAN KESIAPAN UJIAN AKHIR

Saya MUHAMMAD RIZQI AKBAR, dengan ini menyatakan bahwa laporan Tugas Akhir yang berjudul :


ANALISA SENTIMEN TWITTER TERHADAP PENGGUNAAN ALAT DETEKSI COVID-19 GENOSE MENGGUNAKAN METODE KNN

adalah benar hasil karya saya dan belum pernah diajukan sebagai karya ilmiah sebagian atau seluruhnya atau pihak lain.



MUHAMMAD RIZQI AKBAR
17.01.53.0004

Disetujui Oleh Pembimbing
Kami setuju laporan tersebut diajukan untuk Ujian Tugas Akhir
Semarang, 29 Juni 2021



HERIBERTUS YULIANTON, S.Si., M.Cs.
Pembimbing



Dokumen ini diterbitkan secara elektronik.
Disertai QRCode untuk validasi.

SURAT PERNYATAAN KEASLIAN TUGAS AKHIR/ SKRIPSI

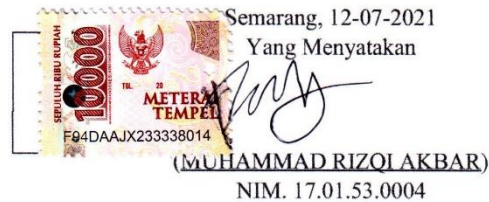
Yang bertanda tangan di bawah ini, saya menyatakan bahwa TUGAS AKHIR / SKRIPSI dengan Judul :

ANALISA SENTIMEN TWITTER TERHADAP PENGGUNAAN ALAT DETEKSI COVID-19 GENOSE MENGGUNAKAN METODE KNN

yang telah diuji di depan tim penguji pada tanggal 12-07-2021, adalah benar hasil karya saya dan dalam TUGAS AKHIR /SKRIPSI ini tidak terdapat keseluruhan atau sebagian tulisan orang lain yang saya ambil dengan cara menyalin, atau meniru dalam bentuk rangkaian kalimat atau simbol yang saya aku seolah-olah sebagai tulisan saya sendiri dan atau tidak terdapat bagian atau keseluruhan tulisan yang saya salin, tiru atau yang saya ambil dari tulisan orang lain tanpa memberikan pengakuan pada penulis aslinya.

Apabila saya melakukan hal tersebut diatas, baik sengaja maupun tidak, dengan ini saya menyatakan menarik TUGAS AKHIR / SKRIPSI yang saya ajukan sebagai hasil tulisan saya sendiri.

Bila kemudian terbukti bahwa saya ternyata melakukan tindakan menyalin atau meniru tulisan orang lain seolah-olah hasil pemikiran saya sendiri, berarti gelar dan ijazah saya yang telah diberikan oleh Universitas Stikubank (UNISBANK) Semarang batal saya terima.



SAKSI 1
Tim Penguji



(HERIBERTUS YULIANTON, S.Si., M.Cs.)

SAKSI 2
Tim Penguji



(JATI SASONGKO WIBOWO, S.Kom., M.Cs.)

SAKSI 3
Tim Penguji



(FATKHUL AMIN, S.T., M.Kom.)



Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan Digisign Unisbank. Untuk memastikan keasliannya, silahkan scan QRCode dan pastikan diarahkan ke alamat <https://digisign.unisbank.ac.id>

HALAMAN PENGESAHAN

Tugas Akhir dengan Judul
**ANALISA SENTIMEN TWITTER TERHADAP PENGGUNAAN ALAT DETEKSI COVID-19
GENOSE MENGGUNAKAN METODE KNN**

Ditulis oleh
NIM : **17.01.53.0004**
Nama : **MUHAMMAD RIZQI AKBAR**

Telah dipertahankan di depan Tim Dosen Penguji Tugas Akhir dan diterima sebagai salah satu syarat
guna menyelesaikan Jenjang Program S1 Program Studi pada Fakultas Universitas Stikubank
(UNISBANK) Semarang.

Semarang, 02-08-2021
Ketua



(HERIBERTUS YULIANTON, S.Si., M.Cs.)
NIDN. 0616077301

Sekretaris



(JATI SASONGKO WIBOWO, S.Kom., M.Cs.)
NIDN. 0621017601

Anggota



(FATKHUL AMIN, S.T., M.Kom.)
NIDN. 0624097401

Mengetahui,
Universitas Stikubank (UNISBANK) Semarang
Fakultas Teknologi Informasi
Dekan



(KRISTOPHORUS HADIONO, Ph.D)
NIDN. 0622027601



HALAMAN MOTTO DAN PERSEMBAHAN

MOTTO :

- “Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya” (QS. Al-Baqarah: 286).
- Kunci kesuksesan adalah fokus pada tujuan, bukan hambatan.
- Apa pun yang sedang kau lakukan, maka lakukanlah dengan sebaik mungkin.

PERSEMBAHAN :

Skripsi ini saya persembahkan kepada:

1. Kedua orang tua saya yang sangat saya sayangi, yang selalu mendukung saya untuk melakukan apapun yang saya inginkan asalkan itu hal baik.
2. Dosen Pembimbing Pak Heribertus Yulianton, S.Si., M.Cs. yang selalu mendukung dan membimbing sampai terselesaikannya skripsi ini.
3. Sahabat – sahabat saya yang selalu mendukung, membantu, dan memberikan semangat saya untuk menyelesaikan skripsi ini.

**FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS STIKUBANK (UNISBANK) SEMARANG**

Program Studi : Teknik Informatika
Tugas Akhir Sarjana Komputer
Semester Genap Tahun 2020 / 2021

**ANALISA SENTIMEN TWITTER TERHADAP PENGGUNAAN ALAT
DETEKSI COVID-19 GENOSE MENGGUNAKAN METODE KNN**

MUHAMMAD RIZQI AKBAR

NIM : 17.01.53.0004

ABSTRAK

GeNose C19 merupakan salah satu alat untuk mendeteksi virus covid-19 yang dikembangkan oleh Universitas Gajah mada (UGM). GeNose bekerja dengan cara mendeteksi Volatile Organic Compound (VOC) yang terbentuk karena adanya infeksi virus covid-19 yang keluar bersamaan dengan nafas yang dimasukkan ke dalam kantong khusus. Saat GeNose mulai digunakan sebagai syarat untuk melakukan perjalanan dengan transportasi umum terutama kereta api dan pesawat terbang, masyarakat memberikan komentar terhadap penggunaan alat tersebut melalui media sosial terutama twitter.

Komentar tersebut memiliki peranan penting karena dapat digunakan untuk melakukan analisis sentimen dalam memprediksi penilaian masyarakat tentang alat tersebut apakah bernilai positif atau negatif. Implementasi analisis sentiment twitter terhadap penggunaan alat deteksi covid -19 GeNose menggunakan metode K-Nearest Neighbor dapat membantu mengklasifikasikan penilaian masyarakat tentang alat tersebut menjadi sentimen negatif atau positif.

Untuk menunjang proses analisis, sebelum data di analisis, ada beberapa tahap yang dilakukan yaitu preprocessing yang didalamnya terdapat proses cleaning, case folding, stemming, stopword removal dan tokenisasi. Proses setelahnya yaitu pembobotan TF-IDF dan yang terakhir klasifikasi metode K-Nearest Neighbor. Setelah proses klasifikasi dan dilakukan pengujian, sistem dapat mengklasifikasikan komentar akurasi terbaik sebesar 87% dengan penggunaan nilai k 1,3, dan 7.

Kata Kunci: analisis sentimen, K-Nearest Neighbor, TF-IDF, Data mining.

ABSTRACT

GeNose C19 is one of the tools to detect the covid-19 virus developed by Gajah Mada University (UGM). GeNose works by detecting Volatile Organic Compounds (VOCs) that are formed due to the Covid-19 virus infection that comes out together with breath that is put in a special bag. When GeNose began to be used as a condition for traveling by public transportation, especially trains and airplanes, people commented on the use of these tools through social media, especially Twitter.

These comments have an important role because they can be used to conduct sentiment analysis in predicting the public's assessment of whether the tool is positive or negative. The implementation of Twitter sentiment analysis on the use of the GeNose COVID-19 detection tool using the K-Nearest Neighbor method can help to classify people's assessments of the tool into negative or positive sentiments.

To support the analysis process, before it is processed into several stages, namely preprocessing which includes the cleaning process, case folding, stemming, stopword removal and tokenization. The next process is TF-IDF weighting and the last is the classification of the K-Nearest Neighbor method. After the classification process and testing, the system can classify the comments with the best accuracy of 87% using k values of 1,3, and 7.

Keywords: *sentiment analysis, K-Nearest Neighbor, TF-IDF, Data mining.*

Pembimbing



(Heribertus Yulianton, S.Si., M.Cs.)

NIDN:0616077301

KATA PENGANTAR

Assalamu'alaikum Wr. Wb.

Dengan memanjatkan puji syukur kehadiran Allah SWT. yang telah melimpahkan segala rahmat, hidayah dan karunia-Nya kepada penulis sehingga penulis dapat menyelesaikan laporan tugas akhir dengan judul “***Analisa Sentimen Twitter Terhadap Penggunaan Alat Deteksi Covid-19 Genose Menggunakan Metode KNN***” guna memperoleh gelar sarjana komputer di Universitas Stikubank Semarang. Dalam penyelesaian laporan banyak pihak yang memberikan bantuan baik secara moril, nasihat dan semangat maupun materil. Oleh karena itu penulis menyampaikan terimakasih kepada:

1. Dr. Safik Faozi, SH, M. Hum, selaku Rektor Universitas Stikubank Semarang.
2. Kristophorus Hadiono, Ph.D, selaku Dekan Fakultas Teknologi Informasi.
3. Eri Zuliarso, M. Kom, selaku Ka. Progdi Teknik Informatika.
4. Heribertus Yulianton, S.Si., M.Cs., selaku Pembimbing yang telah membantu dan memberikan bimbingan serta pengarahan dalam penelitian ini.
5. Dosen-dosen pengampu di Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Stikubank Semarang yang telah memberikan ilmu dan pengalamannya masing-masing, sehingga penulis dapat mengimplementasikan ilmu yang telah disampaikan.
6. Kedua orang tua saya yang sangat saya sayangi dan saya cintai yang telah mencurahkan segenap cinta dan kasih sayang serta perhatian moril maupun materil.

Semoga Allah SWT memberikan balasan yang lebih besar kepada beliau-beliau. Berbagai kekurangan dan kesalahan mungkin pembaca temukan dalam penulisan laporan tugas akhir ini. Semoga apa yang menjadi kekurangan bisa disempurnakan oleh peneliti selanjutnya dan semoga karya tulis ini bisa bermanfaat dan menginspirasi bagi kita semua. Dan pada akhirnya penulis berharap bahwa penulisan laporan tugas akhir ini dapat bermanfaat dan berguna sebagaimana fungsinya.

Wassalamualaikum Wr. Wb.

Semarang, 29 Juni 2021



Penulis

DAFTAR ISI

	halaman
HALAMAN JUDUL.....	i
HALAMAN PERNYATAAN KESIAPAN UJIAN.....	ii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iii
HALAMAN PENGESAHAN.....	iv
HALAMAN MOTTO DAN PERSEMBAHAN.....	v
ABSTRAK.....	vi
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiii
DAFTAR RUMUS	xv
BAB I PENDAHULUAN.....	Error! Bookmark not defined.
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	2
1.3. Batasan Masalah	2
1.4. Tujuan	3
1.5. Metode Penelitian	3
1.6. Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA.....	Error! Bookmark not defined.
2.1. Pustaka yang terkait dengan penelitian.....	Error! Bookmark not defined.
BAB III ANALISA DAN PERANCANGAN SISTEM.....	Error! Bookmark not defined.
3.1. Analisis Data.....	Error! Bookmark not defined.
3.2. Kebutuhan Software dan Hardware..	Error! Bookmark not defined.
3.3. Crawling Data	Error! Bookmark not defined.
3.4. Analisa dan Perancangan	Error! Bookmark not defined.
3.4.1. Preprocessing.....	Error! Bookmark not defined.

3.4.2. Pembobotan TF-IDF.....	Error! Bookmark not defined.
3.4.3. Klasifikasi KNN	Error! Bookmark not defined.
BAB IV HASIL DAN PEMBAHASAN	Error! Bookmark not defined.
4.1. Implementasi sistem menggunakan jupyter notebook.....	Error!
Bookmark not defined.	
4.1.1. Crawling data twitter	Error! Bookmark not defined.
4.1.2. Preprocessing data teks	Error! Bookmark not defined.
4.1.3. Pembobotan Kata	Error! Bookmark not defined.
4.1.4. Membagi Data Menjadi Data Test dan Data Training	Error!
Bookmark not defined.	
4.1.5. Klasifikasi KNN	Error! Bookmark not defined.
4.2. Pembahasan	Error! Bookmark not defined.
BAB V PENUTUP.....	Error! Bookmark not defined.
5.1. Kesimpulan	Error! Bookmark not defined.
5.2. Saran	Error! Bookmark not defined.

DAFTAR PUSTAKA

LAMPIRAN

- Listing Program
- Lembar Bimbingan
- Lembar Berita Acara Revisi
- Lembar Hasil Turnitin

DAFTAR TABEL

Tabel	Halaman
2.1. Tabel Penelitian Terdahulu	9
3.1. contoh data tweet yang telah di crawling	14
3.2. contoh hasil cleaning dan case folding.....	15
3.3. contoh hasil stemming	15
3.4. contoh hasil stopword removal	16
3.5. contoh hasil tokenisasi	16
3.6. contoh hasil akhir preprocessing	17
3.7. contoh penghitungan TF-IDF.....	17
4.1. hasil perbandingan precision, recall, f1-score dan akurasi	33

DAFTAR GAMBAR

Gambar	Halaman
3.1. flowchart implementasi sistem.....	14
3.2. flowchart preprocessing	14
4.1. script import library crawling tweet.....	19
4.2. script proses pemanggilan API twitter	20
4.3. script membuat file csv	20
4.4. script pengambilan tweet	20
4.5. script membaca dan menampilkan file csv	21
4.6. script import library preprocessing	22
4.7. script membaca dan menampilkan file csv	23
4.8. script menghilangkan user dan menampilkan hasilnya.....	23
4.9. script cleaning dan case folding	24
4.10. hasil cleaning dan case folding	24
4.11. script proses stemming.....	24
4.12. hasil proses stemming	25
4.13. script dan hasil proses stopword removal	25
4.14. script proses tokenisasi.....	26
4.15. Hasil proses tokenisasi	26
4.16. script menyimpan hasil preprocessing	26
4.17. script mengubah tipe data.....	26
4.18. script proses pembobotan dan menghitung frekuensi kata	27
4.19. sampel hasil TF – IDF.....	27
4.20. grafik 10 kata yang sering muncul dalam dataset	28
4.21. wordcloud.....	28
4.22. script proses split dataset.....	28
4.23. script proses klasifikasi KNN	29

4.24. hasil sampel 2 jarak terdekat setelah penghitungan jarak euclidean.....	29
4.25. script prediksi dan visualisasi confusion matrix	30
4.26. perbandingan data testing actual dan data testing prediksi sistem.....	31
4.27. visualisasi confusion matrix	31
4.28. script menghitung error rate perbandingan penggunaan nilai K.....	33
4.29. visualisasi hasil error rate nilai K.....	33

DAFTAR RUMUS

Rumus	Halaman
3.1 pembobotan TF-IDF	16
3.2 menghitung IDF	16
3.3 jarak euclidean	18

