

**PERBANDINGAN AKURASI ALGORITMA *K-NEAREST
NEIGHBOR* DAN ROBERTA DALAM ANALISIS SENTIMEN
PADA KOMENTAR *CHANNEL YOUTUBE MIAWAUG***

SKRIPSI SKEMA SCIENTIST JURNAL

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



Disusun oleh :

FACHRUDIN OKTA RAHMAWAN

21.11.4294

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

**PERBANDINGAN AKURASI ALGORITMA *K-NEAREST NEIGHBOR* DAN
ROBERTA DALAM ANALISIS SENTIMEN PADA KOMENTAR *CHANNEL*
*YOUTUBE MIAWAUG***

SKRIPSI SKEMA SCIENTIST JURNAL

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



Disusun oleh :

FACHRUDIN OKTA RAHMAWAN

21.11.4294

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

HALAMAN PERSETUJUAN

JALUR NON-REGULER

**PERBANDINGAN AKURASI ALGORITMA K-NEAREST NEIGHBOR DAN ROBERTA DALAM
ANALISIS SENTIMEN PADA KOMENTAR CHANNEL YOUTUBE MIAWAUG**

yang disusun dan diajukan oleh


Fachrudin Okta Rahmawan

21.11.4294

telah disetujui oleh Dosen Pembimbing

pada tanggal 16 Desember 2024

Dosen Pembimbing,



Hanafi, S.Kom, M.Eng, Ph.D
NIK. 190302024

HALAMAN PENGESAHAN

JALUR NON-REGULER

**PERBANDINGAN AKURASI ALGORITMA K-NEAREST NEIGHBOR DAN ROBERTA DALAM
ANALISIS SENTIMEN PADA KOMENTAR CHANNEL YOUTUBE MIAWAUG**

yang disusun dan diajukan oleh

Fachrudin Okta Rahmawan
21.11.4294

Telah dipertahankan di depan Dewan Penguji
pada tanggal 16 Desember 2024

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Nuri Cahyono., M.Kom
NIK. 190302278

Arifiyanto Hadinegoro, S.Kom., M.T.
NIK. 190302289

Hanafi, S.Kom., M.Eng., Ph.D.
NIK. 190302024

Laporan ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 16 Desember 2024

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : Fachrudin Okta Rahmawan

NIM : 21.11.4294

Menyatakan bahwa Laporan dengan judul berikut:

PERBANDINGAN AKURASI ALGORITMA *K-NEAREST NEIGHBOR* DAN ROBERTA DALAM ANALISIS SENTIMEN PADA KOMENTAR *CHANNEL YOUTUBE MIWAUG*

Dosen Pembimbing : Hanafi, S.Kom, M.Eng, Ph.D

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 kegiatan **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 ketidak-benaran 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, 16 Desember 2024

Yang Menyatakan,



Fachrudin Okta Rahmawan

KATA PENGANTAR

Puji syukur ke hadirat Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya, sehingga laporan ini dapat diselesaikan dengan baik dan tepat waktu. Laporan ini disusun sebagai salah satu syarat untuk memenuhi tugas akhir pada program studi yang sedang saya jalani.

Dalam proses penyusunan laporan ini, banyak pihak yang telah memberikan bantuan, dukungan, dan arahan. Oleh karena itu, dengan segala kerendahan hati, saya ingin mengucapkan terima kasih kepada:

1. **Bapak Hanafi, S.Kom, M.Eng, Ph.D** selaku Dosen Pembimbing, atas segala bimbingan, nasihat, dan waktu yang telah diberikan selama proses penyusunan laporan ini.
2. **Bapak Nuri Cahyono., M.Kom, Bapak Ariflyanto Hadinegoro, S.Kom., M.T. dan Bapak Hanafi, S.Kom, M.Eng, Ph.D** selaku Tim Dosen Penguji, atas saran, kritik, dan masukan yang membangun untuk penyempurnaan laporan ini.
3. Orang tua dan keluarga saya, yang selalu memberikan doa, dukungan, dan motivasi tanpa henti selama saya menempuh pendidikan dan menyelesaikan laporan ini.
4. Seluruh teman dan rekan yang telah membantu secara langsung maupun tidak langsung dalam proses penyusunan laporan ini.

Semoga semua bantuan, dukungan, dan doa yang telah diberikan mendapatkan balasan yang setimpal dari Allah SWT.

Saya menyadari bahwa laporan ini masih memiliki kekurangan. Oleh karena itu, saya terbuka terhadap kritik dan saran yang membangun untuk perbaikan di masa depan.

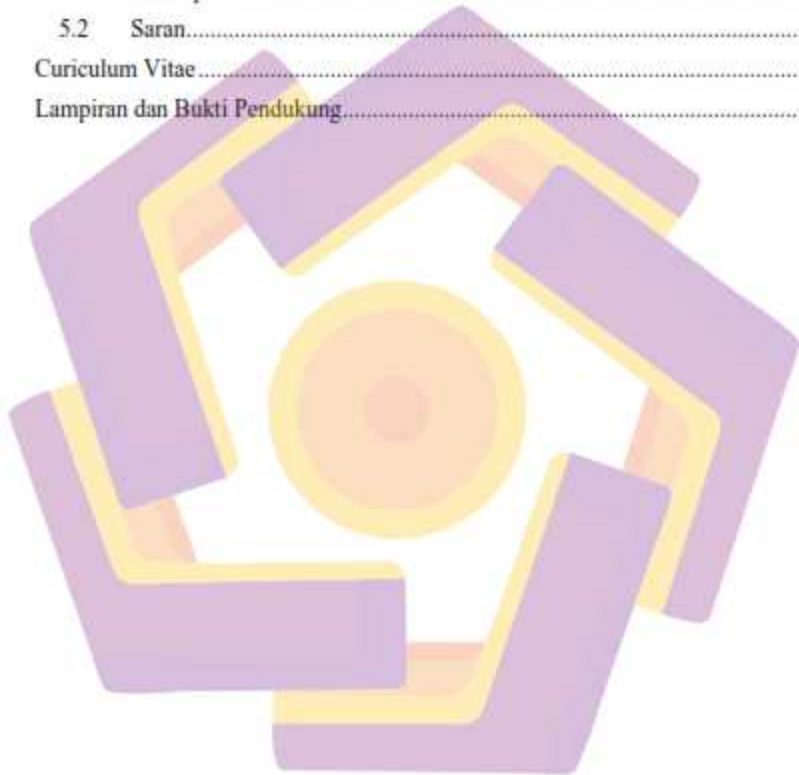
Yogyakarta, 16 Desember 2024

Penulis

DAFTAR ISI

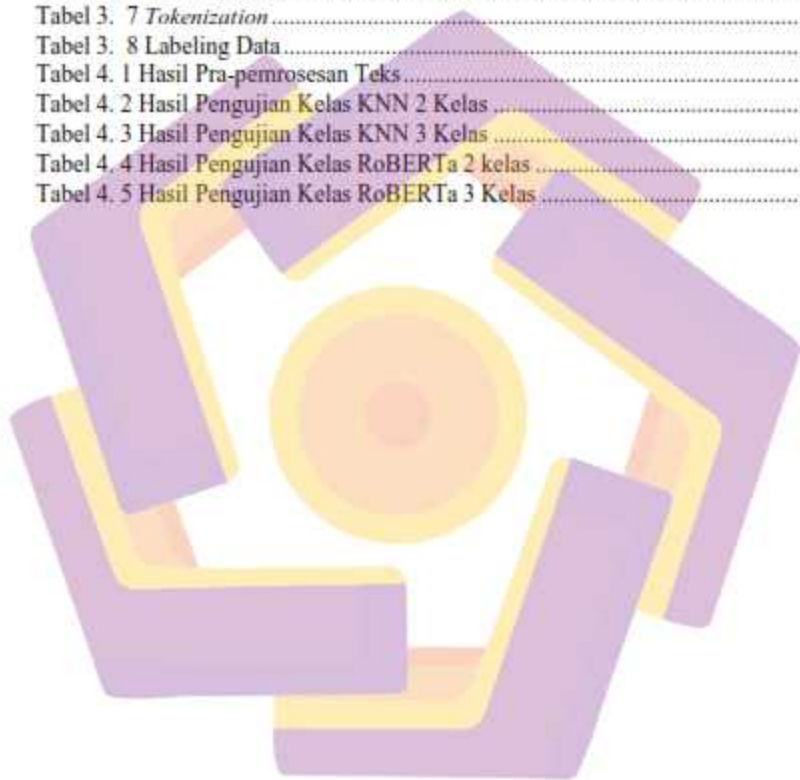
Halaman Judul.....	I
Halaman Persetujuan.....	II
Halaman Pengesahan	III
Halaman Pernyataan Keaslian Karya	IV
Kata Pengantar	V
Daftar Isi	VI
Daftar Tabel	VIII
Daftar Gambar.....	IX
Daftar Lampiran.....	X
Daftar Lambang dan Singkatan.....	XI
Daftar Istilah.....	XII
Intisari.....	XIII
<i>Abstract</i>	XIV
Bab I Pendahuluan	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah.....	3
1.3. Batasan Masalah.....	3
1.4. Tujuan	3
Bab II	5
Tinjauan Pustaka	5
2.1. Studi Literatur	5
2.2. Landasan Teori.....	5
Bab III	6
METODOLOGI PENELITIAN	6
3.1 Metode.....	6
BAB IV	16
HASIL DAN PEMBAHASAN	16
4.1 Hasil Data <i>Crawling</i>	16
4.2 Hasil Pra-pemrosesan Teks.....	16
4.3 <i>Wordcloud</i>	17

4.4	Hasil Labeling	19
4.5	Hasil Resampling	20
4.6	Hasil Perbandingan Sentimen dari Algoritma KNN dan RoBERTa	21
BAB V	27
KESIMPULAN	27
5.1	Kesimpulan	27
5.2	Saran.....	27
Curriculum Vitae	30
Lampiran dan Bukti Pendukung	31



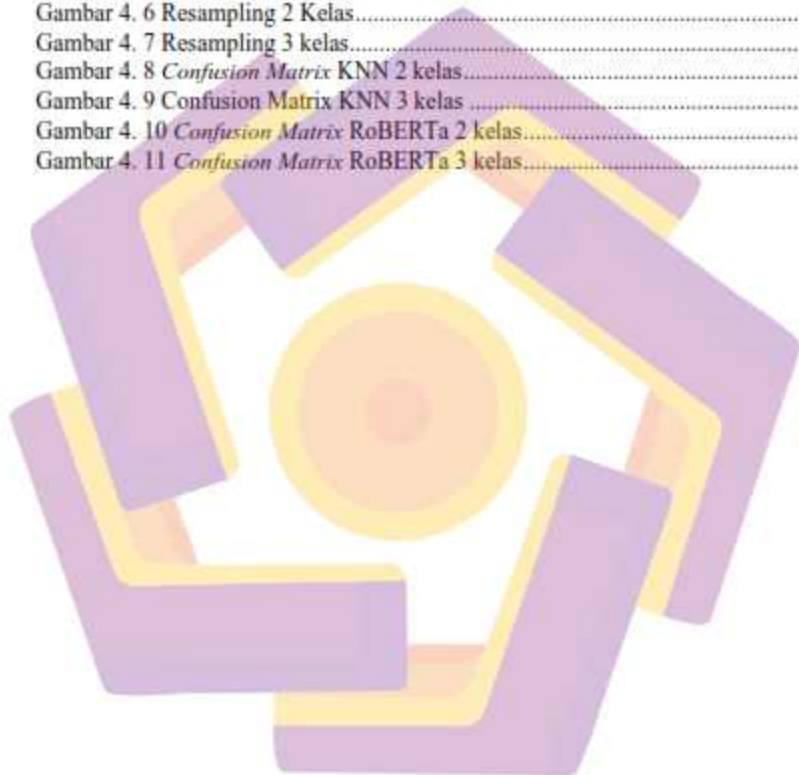
DAFTAR TABEL

Tabel 3. 1 Sumber Data.....	7
Tabel 3. 2 <i>Data Cleaning</i>	8
Tabel 3. 3 <i>Case Folding</i>	8
Tabel 3. 4 <i>Normalization</i>	9
Tabel 3. 5 <i>Steaming</i>	9
Tabel 3. 6 <i>Stopword Removal</i>	9
Tabel 3. 7 <i>Tokenization</i>	10
Tabel 3. 8 Labeling Data.....	13
Tabel 4. 1 Hasil Pra-pemrosesan Teks.....	17
Tabel 4. 2 Hasil Pengujian Kelas KNN 2 Kelas.....	21
Tabel 4. 3 Hasil Pengujian Kelas KNN 3 Kelas.....	22
Tabel 4. 4 Hasil Pengujian Kelas RoBERTa 2 kelas.....	23
Tabel 4. 5 Hasil Pengujian Kelas RoBERTa 3 Kelas.....	25



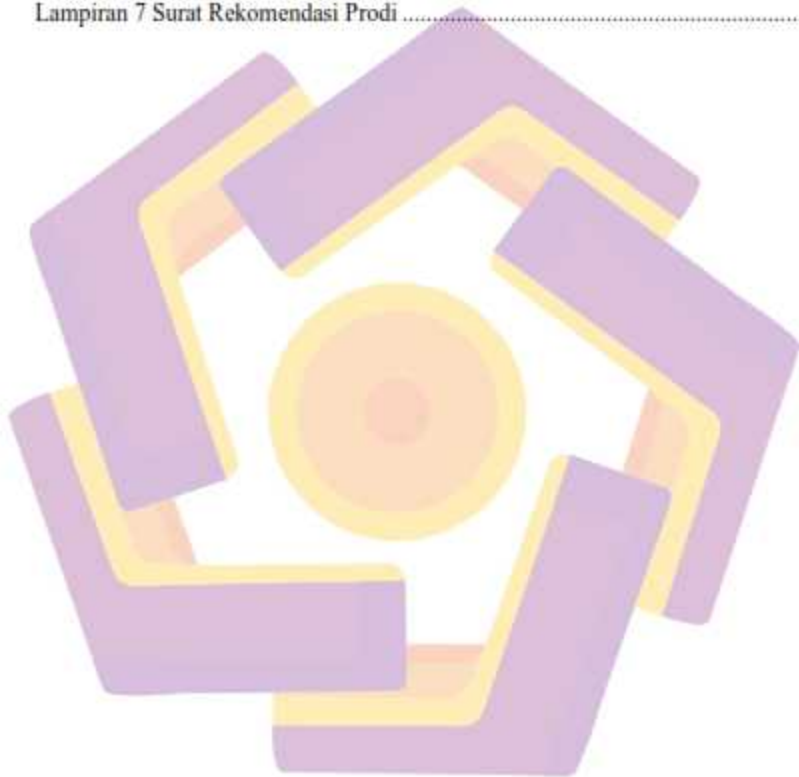
DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian.....	6
Gambar 4. 1 Hasil Data <i>Crawling</i>	16
Gambar 4. 2 <i>Wordcloud</i> Netral.....	18
Gambar 4. 3 <i>Wordcloud</i> Positif.....	18
Gambar 4. 4 <i>Wordcloud</i> Negatif.....	19
Gambar 4. 5 Hasil Labeling.....	19
Gambar 4. 6 Resampling 2 Kelas.....	20
Gambar 4. 7 Resampling 3 kelas.....	20
Gambar 4. 8 <i>Confusion Matrix</i> KNN 2 kelas.....	21
Gambar 4. 9 <i>Confusion Matrix</i> KNN 3 kelas.....	23
Gambar 4. 10 <i>Confusion Matrix</i> RoBERTa 2 kelas.....	24
Gambar 4. 11 <i>Confusion Matrix</i> RoBERTa 3 kelas.....	25



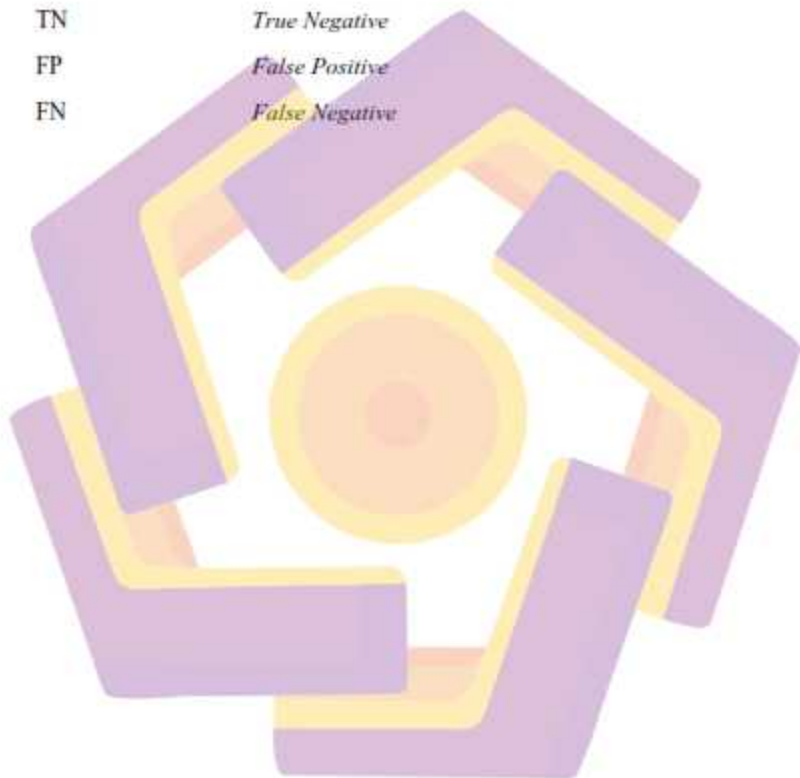
DAFTAR LAMPIRAN

Lampiran 1 LOA	31
Lampiran 2 Lembar Review	33
Lampiran 3 Bukti Terbit.....	34
Lampiran 4 Sertifikat	34
Lampiran 5 Bukti Index	35
Lampiran 6 Bukti Pembayaran.....	36
Lampiran 7 Surat Rekomendasi Prodi	37



DAFTAR LAMBANG DAN SINGKATAN

KNN	<i>K-NEAREST NEIGHBOR</i>
RoBERTa	<i>Robustly Optimized BERT Pretraining Approach</i>
TP	<i>True Positive</i>
TN	<i>True Negative</i>
FP	<i>False Positive</i>
FN	<i>False Negative</i>



DAFTAR ISTILAH

Vektor	Besaran yang mempunyai arah
Akurasi	Merupakan rasio prediksi Benar
<i>Precision</i>	Tingkat ketepatan antara hasil prediksi model dengan data yang diminta
<i>Recall</i>	Rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif
<i>F1-Score</i>	Merupakan perbandingan rata-rata presisi dan recall yang dibobotkan
<i>Confusion Matrix</i>	Pengukuran performa untuk masalah klasifikasi machine learning dimana keluaran dapat berupa dua kelas atau lebih

INTISARI

Penelitian ini bertujuan untuk mengevaluasi keakuratan dua algoritma, yaitu *K-NEAREST NEIGHBOR* (knn) dan *robustly optimized bert approach* (roberta), dalam menganalisis sentimen pada komentar di kanal *YOUTUBE* miawaug. analisis sentimen dilakukan pada dua kategori sentimen: klasifikasi biner (positif dan negatif) dan klasifikasi multi-kelas (positif, netral, dan negatif). dengan menggunakan knn, klasifikasi biner menghasilkan akurasi sebesar 86,12%, f1-score sebesar 87,44%, recall sebesar 96,64%, dan precision sebesar 79,89%. sebaliknya, klasifikasi multi-kelas mencapai akurasi 98,21%, f1-score, dan recall dengan presisi 98,23%. namun, model roberta mengungguli knn dengan akurasi 93,89%, f1-score 93,88%, recall 94,59%, dan presisi 93,22% pada klasifikasi biner. untuk klasifikasi multi-kelas, roberta lebih unggul lagi, mencapai 99,21% untuk akurasi, f1-score, recall, dan presisi. temuan ini menunjukkan bahwa roberta mengungguli knn dalam analisis sentimen, terutama dalam konteks multi-kelas, yang mengindikasikan ketangguhannya yang lebih besar untuk aplikasi ini.

Kata kunci: Analisis Sentimen 1, *K-NEAREST NEIGHBOR* (KNN) 2, RoBERTa 3, Komentar *YOUTUBE* 4.



ABSTRACT

This study aims to evaluate the accuracy of two algorithms, K-NEAREST NEIGHBOR (knn) and robustly optimized bert approach (roberta), in analyzing sentiment within comments on miawaug's YOUTUBE CHANNEL. sentiment analysis was conducted on two sentiment categories: binary classification (positive and negative) and multi-class classification (positive, neutral, and negative). using knn, the binary classification yielded an accuracy of 86.12%, f1-score of 87.44%, recall of 96.64%, and precision of 79.89%. in contrast, the multi-class classification achieved 98.21% accuracy, f1-score, and recall with a precision of 98.23%. however, the roberta model outperformed knn, achieving 93.89% accuracy, 93.88% f1-score, 94.59% recall, and 93.22% precision in binary classification. for multi-class classification, roberta further excelled, attaining 99.21% across accuracy, f1-score, recall, and precision. these findings demonstrate that roberta surpasses knn in sentiment analysis, especially in multi-class contexts, indicating its greater robustness for this application.

Keyword: *Sentiment Analysis 1, K-NEAREST NEIGHBOR (KNN) 2, RoBERTa 3, YOUTUBE Comments 4.*

