

Skripsi

**ANALISIS SENTIMEN VIDEO PEMBATALAN PIALA DUNIA U-20 DI
INDONESIA OLEH GANJAR PRANOWO MENGGUNAKAN NAIVE
BAYES CLASSIFIER**



ELVIN PRASETYO

Nomor Mahasiswa : 195410209

**PROGRAM STUDI INFORMATIKA
PROGRAM SARJANA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA**

2024

Skripsi
ANALISIS SENTIMEN VIDEO PEMBATALAN PIALA DUNIA U-20 DI
INDONESIA GANJAR PRANOWO MENGGUNAKAN NAIVE BAYES
CLASSIFIER

Diajukan sebagai salah satu syarat untuk menyelesaikan studi

Program Sarjana

Program Studi Informatika

Fakultas Teknologi Informasi

Universitas Teknologi Digital Indonesia Yogyakarta

Disusun Oleh

ELVIN PRASETYO

Nomor Mahasiswa : 195410209

PROGRAM STUDI INFORMATIKA
PROGRAM SARJANA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA

2024

HALAMAN PERSETUJUAN

Judul : Analisis Sentimen Video Pembatalan Piala Dunia U-20
Di Indonesia Oleh Ganjar Pranowo Menggunakan
Naive Bayes Classifier

Nama : Elvin Prasetyo

NIM : 195410209

Program Studi : Informatika

Program : Sarjana

Semester : Genap

Tahun Akademik : 2023/2024



Telah diperiksa dan disetujui untuk diujikan di hadapan Dewan Penguji Skripsi

Yogyakarta, 1 Juli 2024

Dosen pembimbing,

Maria Mediatix Sebatubun.S.Kom., M.Eng.

NIDN:0514089101

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN VIDEO PEMBATALAN PIALA DUNIA U-20 DI
INDONESIA GANJAR PRANOWO MENGGUNAKAN NAIVE BAYES
CLASSIFIER**

Telah dipertahankan di depan Dewan Penguji Skripsi dan dinyatakan
diterima untuk memenuhi sebagian persyaratan guna memperoleh Gelar

Sarjana Komputer

Program Studi Informatika

Fakultas Teknologi Informasi

Universitas Teknologi Digital Indonesia

Yogyakarta

Yogyakarta, 10 Juli 2024

Dewan Penguji

1. Sari Iswanti, S.Si., M.Kom

2. Sri Redjeki, S.Si., M.Kom., Ph.D.

3. Maria Mediatrix Sebatubun, S.Kom., M.Eng

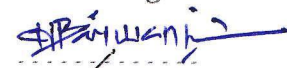
NIDN

0508027202

0521047401

0514089101

Tanda Tangan







Mengetahui
Ketua Program Studi Informatika



Dini Fakta Sari, S.T., M.T.

NIDN: 0507108401

PERNYATAAN KEASLIAN SKRIPSI

Dengan ini saya menyatakan bahwa naskah skripsi ini belum pernah diajukan untuk memperoleh gelar Sarjana Komputer di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara sah diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 5 Agustus 2024



Elvin Prasetyo
NIM: 195410209

KATA PENGANTAR

Dengan penuh kehormatan dan rasa syukur atas rahmat dan kehadiran Allah SWT yang telah memberikan rahmat dan karunia-Nya, penulis menghadirkan skripsi ini sebagai sebuah upaya untuk menggambarkan, menganalisis, dan menyajikan informasi yang relevan terkait dengan Analisis Sentimen Video Pembatalan Piala Dunia U-20 Di Indonesia Oleh Ganjar Pranowo Menggunakan Naive Bayes Classifier yang disusun dan diselesaikan dengan baik.

Skripsi ini disusun sebagai salah satu syarat untuk mendapatkan gelar strata-1 (S1) di program studi Informatika, Fakultas Teknologi Informasi, Universitas Teknologi Digital Indonesia. Dalam masa pengerjaan skripsi ini, penulis mendapatkan banyak dukungan, masukan dan bantuan dari berbagai pihak. Dalam kesempatan ini, penulis ingin menyampaikan penghargaan setinggi-tingginya kepada semua pihak yang telah memberikan dukungan, bimbingan, dan kontribusi. Tanpa kerjasama mereka, dokumen ini tidak akan menjadi kenyataan. Dalam hal ini, penulis mengucapkan terima kasih kepada:

1. Ibu Sri Redjeki, S.Si. ,M.Kom. ,Ph.D., selaku Rektor Universitas Teknologi Digital Indonesia.
2. Bapak Dr. Bambang Purnomosidi DP, S. E. Akt., S. Kom., MMSI, selaku Dekan Fakultas Teknologi Informasi Universitas Teknologi Digital Indonesia.
3. Ibu Dini Fakta Sari, ST, M.T., selaku Ketua Program Studi Informatika Universitas Teknologi Digital Indonesia.

4. Ibu Maria Mediatrix Sebatubun, S.Kom, M.Eng., selaku dosen pembimbing yang telah meluangkan waktunya untuk memberikan bimbingan dan masukan selama penyusunan skripsi.
5. Kedua orang tua tercinta dan seluruh keluarga besar yang selalu mendukung penulis di manapun anda berada.
6. Seluruh dosen di Universitas Teknologi Digital Indonesia yang telah memberikan ilmunya kepada penulis selama masa perkuliahan.
7. Seluruh civitas akademika Universitas Teknologi Digital Indonesia yang banyak memberi bantuan dan dukungan di Universitas Teknologi Digital Indonesia.

Penulis menyadari bahwa penyempurnaan adalah proses yang berkelanjutan. Penulis memohon maaf apabila terdapat kesalahan di dalam skripsi ini dan dengan tulus mengundang setiap masukan, saran, dan tanggapan yang dapat meningkatkan kualitas dokumen ini sarana informasi dan pengetahuan dalam membuat karya tulis di masa yang akan datang. Semoga dokumen ini memberikan nilai tambah dan manfaat bagi pembaca. Terima kasih atas perhatian dan waktu yang telah diberikan.

Yogyakarta, 1 April 2024

Elvin Prasetyo

DAFTAR ISI

HALAMAN JUDUL.....	I
HALAMAN PERSETUJUAN.....	II
HALAMAN PENGESAHAN.....	III
PERNYATAAN KEASLIAN SKRIPSI.....	IV
KATA PENGANTAR.....	V
DAFTAR ISI.....	VII
DAFTAR GAMBAR.....	X
DAFTAR TABEL.....	XII
INTISARI.....	XIII
ABSTRACT.....	XIV
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Ruang Lingkup.....	4
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	5
1.6 Sistematika Penulisan.....	6
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI.....	7
2.1 Tinjauan Pustaka.....	7
2.2 Dasar Teori.....	10
2.2.1 Youtube.....	10
2.2.2 Youtube API.....	10
2.2.3 Python.....	11
2.2.4 Analisis Sentimen.....	11
2.2.5 Naive Bayes Classifier.....	12
2.2.6 Akurasi, Presisi, Recall, dan F1-score.....	13
BAB III METODE PENELITIAN.....	16
3.1 Bahan/Data.....	16

3.2 Peralatan.....	17
3.2.1 Kebutuhan Perangkat Keras (Hardware).....	17
3.2.2 Analisis Kebutuhan Perangkat Lunak (Software).....	18
3.3 Prosedur dan Pengumpulan Data.....	18
3.4 Analisis Dan Rancangan Sistem.....	19
3.4.1 Pengumpulan Data.....	19
3.4.2 Preprocessing Data.....	20
3.4.3 Pelabelan Data Sentimen.....	21
3.4.4 Pembagian Data.....	21
3.4.5 Klasifikasi Naïve Bayes.....	22
3.4.6 Evaluasi Model.....	22
3.5 Perancangan Antar Muka.....	22
BAB IV IMPLEMENTASI DAN PEMBAHASAN.....	26
4.1 Implementasi.....	26
4.1.1 Pengambilan Data.....	26
4.1.2 Preprocessing Data.....	28
4.1.2.1 Case Folding.....	29
4.1.2.2 Data Cleansing.....	29
4.1.2.3 Tokenisasi.....	31
4.1.2.4 Normalisasi.....	32
4.1.2.5 Filtering.....	33
4.1.2.6 Stemming.....	35
4.1.3 Pelabelan Data.....	37
4.1.4 Pembagian Data.....	40
4.1.5 Klasifikasi Naive Bayes.....	41
4.1.6 Evaluasi Model.....	42
4.2 Pembahasan.....	46
4.2.1 Pelabelan Data Secara Manual.....	47
4.2.2 Tampilan Aplikasi Web.....	48
4.3 Hasil Uji Coba.....	52

BAB V PENUTUP.....	55
5.1 Kesimpulan.....	55
5.2 Saran.....	55
Daftar Pustaka.....	57
Lampiran	

DAFTAR GAMBAR

Gambar 3.1.Tahapan Dalam Analisis Sentimen.....	19
Gambar 3.2.Tampilan Halaman Beranda.....	23
Gambar 3.3.Tampilan Halaman Confusion Matrix.....	24
Gambar 3.4.Tampilan Halaman Classification Report.....	25
Gambar 3.5.Tampilan Halaman Hasil Data.....	25
Gambar 4.1.Video Youtube untuk pengambilan data.....	27
Gambar 4.2.Kode untuk pengambilan data.....	29
Gambar 4.3.Kode untuk Case Folding.....	30
Gambar 4.4.Kode untuk Data Cleansing.....	31
Gambar 4.5.Kode untuk Tokenisasi.....	32
Gambar 4.6.Mengubah kata tidak baku di file Normalisasi.....	33
Gambar 4.7.Kode untuk Normalisasi.....	33
Gambar 4.8.Kode untuk Filtering.....	35
Gambar 4.9.Kode untuk Stemming.....	36
Gambar 4.10.Kode Deep Translator.....	37
Gambar 4.11.Kode untuk Pelabelan Data.....	39
Gambar 4.12.Sebaran Dataset.....	40
Gambar 4.13.Kode untuk Pembagian Data.....	42
Gambar 4.14.Kode untuk Klasifikasi Naive Bayes.....	43
Gambar 4.15.Kode untuk Evaluasi Model.....	44
Gambar 4.16.Confusion Matrix.....	44
Gambar 4.17.Classification Report.....	45
Gambar 4.18.Hasil Pelabelan Manual.....	48
Gambar 4.19.Halaman Menu Beranda.....	49
Gambar 4.20.Halaman Menu Confusion Matrix.....	50
Gambar 4.21. Halaman Menu Classification Report.....	51
Gambar 4.22.Halaman Hasil Data.....	51
Gambar 4.23.Tampilan Klasifikasi Sentimen Positif.....	52

Gambar 4.24. Tampilan Klasifikasi Sentimen Negatif.....	53
Gambar 4.25. Tampilan Klasifikasi Sentimen Netral.....	54

DAFTAR TABEL

Tabel 1.1. Tinjauan Pustaka.....	9
Tabel 4.1. Tabel Case Folding.....	30
Tabel 4.2. Tabel Data Cleansing.....	31
Tabel 4.3. Tabel Tokenisasi.....	32
Tabel 4.4. Tabel Normalisasi.....	33
Tabel 4.5. Tabel Filtering.....	35
Tabel 4.6. Tabel Stemming.....	37

INTISARI

Indonesia dijadwalkan menjadi tuan rumah Piala Dunia U-20 pada tahun 2023, tetapi FIFA membatalkan pencalonan Indonesia sebagai tuan rumah Piala Dunia U-20 dan memindahkannya ke negara lain karena beberapa pihak termasuk Ganjar Pranowo di sebuah video *youtube* menolak keikutsertaan Tim Nasional Israel yang lolos kualifikasi dalam turnamen tersebut. Alasan Ganjar Pranowo menolak Israel untuk bermain di Indonesia mencerminkan komitmennya terhadap perjuangan kemerdekaan Palestina yang diperintahkan oleh Presiden pertama Indonesia, Soekarno. Penolakan ini didasarkan pada Undang-Undang Dasar Negara Kesatuan Republik Indonesia (UUD NKRI) Tahun 1945 yang mengatur bahwa negara-negara yang masih melakukan "penjajahan di atas dunia" tidak diperbolehkan masuk ke Indonesia.

Banyak pihak yang kecewa dengan keputusan ini mengkritik Ganjar Pranowo sebagai penyebab pembatalan piala dunia u-20 di Indonesia melalui media sosial, Topik ini pun selalu menjadi perbincangan hangat di *youtube*. Berbagai opini datang dari warganet dengan opini yang baik, yang beropini jelek, dan ada pula yang beropini netral di media sosial dalam menanggapi pernyataan Ganjar Pranowo. Hal tersebut dapat dianalisis menggunakan analisis sentimen.

Analisis sentimen dapat digunakan untuk menganalisis pendapat netizen dan mengelompokkannya menjadi tiga kategori yaitu positif, netral, dan negatif. Penelitian ini memanfaatkan data dari *Youtube* menggunakan *Youtube API* untuk melakukan analisis sentimen yang menggunakan metode *Naïve Bayes Classifier* dengan kata kunci "Bawa Amanat Bung Karno, Ganjar Pranowo Tolak Israel Hadir di Piala Dunia U-20".

Jumlah data yang digunakan sebanyak 2542 data yang diambil selama bulan Maret 2024. Data dibagi dengan rasio 80:20 dimana 2033 data digunakan untuk data latih dan 509 data digunakan sebagai data uji. Penelitian ini menghasilkan akurasi *Naive Bayes* sebesar 69,155% dengan nilai presisi sentimen negatif adalah 66,9%, sentimen netral adalah 79,7%, dan sentimen positif adalah 62,5%. Nilai *recall* untuk sentimen negatif adalah 56%, sentimen netral adalah 68,7%, dan sentimen positif adalah 78%. Terakhir, *f1-score* untuk sentimen negatif adalah 61%%, sentimen netral adalah 74%, dan untuk sentimen positif adalah 69,7%.

Kata Kunci: Analisis Sentimen, Ganjar, Naive Bayes, Piala Dunia, Youtube

ABSTRACT

Indonesia was scheduled to host the U-20 World Cup in 2023, but FIFA canceled Indonesia's candidacy to host the U-20 World Cup and moved it to another country as several parties including Ganjar Pranowo in a youtube video rejected the participation of the Israeli National Team who qualified in the tournament. Ganjar Pranowo's reason for rejecting Israel to play in Indonesia reflects his commitment to the Palestinian independence struggle ordered by Indonesia's first President, Soekarno. This refusal is based on the 1945 Constitution of the Unitary State of the Republic of Indonesia (UUD NKRI) which stipulates that countries that are still "colonizing the world" are not allowed to enter Indonesia.

Many parties who were disappointed with this decision criticized Ganjar Pranowo as the cause of the cancellation of the u-20 world cup in Indonesia through social media, This topic has always been a hot conversation on youtube. various opinions came from netizens with good opinions, bad opinions, and some with neutral opinions on social media in response to Ganjar Pranowo's statement. This can be analyzed using sentiment analysis.

Sentiment analysis can be used to analyze netizen opinions and classify them into three categories: positive, neutral, and negative. This research utilizes data from Youtube using the Youtube API to conduct sentiment analysis using the Naïve Bayes Classifier method with the keyword "Bawa Amanat Bung Karno, Ganjar Pranowo Tolak Israel Hadir di Piala Dunia U-20".

The amount of data used is 2542 data taken during March 2024. The data is divided by a ratio of 80:20 where 2033 data is used for training data and 509 data is used as test data. This research resulted in Naive Bayes accuracy of 69.155% with the precision value of negative sentiment is 66.9%, neutral sentiment is 79.7%, and positive sentiment is 62.5%. The recall value for negative sentiment is 56%, neutral sentiment is 68.7%, and positive sentiment is 78%. Finally, the f1-score for negative sentiment is 61%%, neutral sentiment is 74%, and for positive sentiment is 69.7%.

Keywords: ganjar, naive bayes, sentiment analysis, world cup, youtube