



**PERBANDINGAN ANALISIS SENTIMEN PADA MEDIA SOSIAL
TWITTER TERHADAP PEMBATALAN PIALA DUNIA U-20 DI
INDONESIA MENGGUNAKAN METODE SUPPORT VECTOR
MACHINE DAN NAIVE BAIYES**

LAPORAN TUGAS AKHIR

**BAGAS DWI HAZTOMO
41520110090**

**UNIVERSITAS
MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**



**PERBANDINGAN ANALISIS SENTIMEN PADA MEDIA SOSIAL TWITTER
TERHADAP PEMBATALAN PIALA DUNIA U-20 DI INDONESIA
MENGUNAKAN METODE SUPPORT VECTOR MACHINE DAN NAIVE
BAIYES**

LAPORAN TUGAS AKHIR

BAGAS DWI HAZTOMO

41520110090

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**UNIVERSITAS
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Bagas Dwi Haztomo

NIM : 41520110090

Program Studi : Teknik Informatika

Judul Proposal Penelitian : Perbandingan Analisis Sentimen Pada Media Sosial Twitter Terhadap Pembatalan Piala Dunia U-20 di Indonesia Menggunakan Metode Support Vector Machine Dan Naive Baiyes

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, 23 November 2024



Bagas Dwi Haztomo

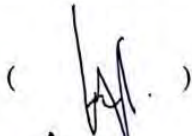
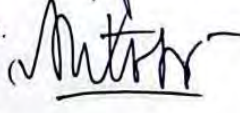
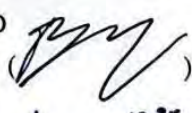
HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Bagas Dwi Haztomo
NIM : 41520110090
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Analisis Sentimen Pada Media Sosial Twitter Terhadap Pembatalan Piala Dunia U-20 di Indonesia Menggunakan Metode Support Vector Machine dan Naive Bayes

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:

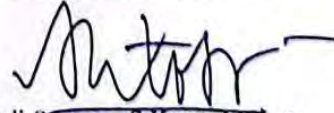
Pembimbing	: Umniy Salamah, S.T.MMSI	()
NIDN	: 0306098104	
Ketua Penguji	: Dr. Hadi Santoso, S.Kom., M.Kom.	()
NIDN	: 0225067701	
Penguji 1	: Bagus Priambodo, ST, M.TI., Ph.D	()
NIDN	: 0313057905	
Penguji 2	: Lukman Hakim, ST, M.Kom	()
NIDN	: 0327107701	

Jakarta, 4 Februari 2025

Dekan


Dr. Bambang Jokonowo, S.Si., MTI
NIDN : 0320037002

Ketua Program Studi


Dr. Hadi Santoso, S.Kom., M.Kom
NIDN : 0225067701

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Proposal Penelitian ini. Penulisan Proposal Penelitian ini dilakukan dalam rangka memenuhi salah satu syarat untuk disidangkan pada seminar proposal. 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. Bapak Prof. Dr. Andi Adriansyah, M.Eng selaku Rektor Universitas MercuBuana Jakarta.
2. Bapak Dr. Bambang Jokonowo selaku Dekan Fakultas Ilmu Komputer Universitas Mercu Buana Jakarta.
3. Bapak Hadi Santoso S.Kom selaku Ketua Program Studi Teknik Informatika Universitas Mercu Buana Jakarta.
4. Ibu Umniy Salamah selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan proposal penelitian ini.
5. Dosen Penguji proposal penelitian atas koreksi dan arahan serta masukannya.
6. Keluarga yang telah memberikan semangat serta dukungan untuk dapat menyelesaikan perkuliahan.

Akhir kata, saya berharap Allah Swt berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Proposal Penelitian ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 04 Februari 2025

Bagas Dwi Haztomo

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

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

Nama	Bagas Dwi Haztomo
NIM	41520110090
Program Studi	Teknik Informatika
Judul Laporan Skripsi	Perbandingan Analisis Sentimen dengan Named Entity Recognition Pada Media Sosial Twitter Terhadap Pembatalan Piala Dunia U-20 di Indonesia Menggunakan Metode Support Vector Machine Dan Naive Baiyes

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, 05 Februari 2025

Yang menyatakan,



Bagas Dwi Haztomo

ABSTRAK

Nama : Bagas Dwi Haztomo
NIM : 41520110090
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Analisis Sentimen dengan Named Entity Recognition Pada Media Sosial Twitter Terhadap Pembatalan Piala Dunia U-20 di Indonesia Menggunakan Metode Support Vector Machine Dan Naive Baiyes
Dosen Pembimbing : Dosen Umniy Salamah, S.T. MMSI

Media sosial twitter merupakan platform media sosial urutan ke 6 yang sering di akses di Indonesia. Terjadinya pembatalan piala dunia u-20 di Indonesia merupakan hal yang sangat ramai di perbincangkan oleh masyarakat Indonesia khususnya para pecinta sepakbola tanah air. Banyaknya isu isu di dalam pembatalan ini menimbulkan pro dan kontra. Tujuan penelitian ini adalah untuk memperoleh pemahaman yang lebih mendalam mengenai sentimen masyarakat. Penelitian bertujuan untuk mengklasifikasikan ulasan masyarakat sebagai positif atau negatif. Penelitian ini merupakan penelitian kuantitatif tentang analisis sentimen yang mengekstrak sentimen pembatalan pialan dunia u-20 di Indonesia pada twitter dengan tujuan menerapkan algoritma Naive Baiyes dan Support Vector Machine (SVM). Alur penelitian ini menggunakan metode SEMMA (Sample, Explore, Modify, Model, dan Assess). Sample berupa pengumpulan data set melalui tweet yang di-crawling selanjutnya menyeleksi atribut yang penting pada tahap explore. Selanjutnya modify berupa text preprocessing sehingga menjadi data terstruktur. Model untuk pelabelan menggunakan masing-masing metode yang akan diuji. Setelah pelabelan akan diklasifikasikan menggunakan kedua metode tersebut. Assess berupa hasil evaluasi menggunakan confusion matrix dan k-fold cross validation. Hasil Penelitian ini akan membandingkan kedua algoritma SVM dan Naive Baiyes

Kata Kunci: Analisis Sentimen, Piala Dunia u-20, Support Vector Machine (SVM), Naive Baiyes, SEMMA

ABSTRACT

Name : Bagas Dwi Haztomo
NIM : 41520110090
Study Program : Informatics Engineering
Title Research Proposal : Sentiment Analysis on Twitter Social Media Regarding the Cancellation of the U-20 World Cup in Indonesia Using Support Vector Machine and Naive Bayes Methods
Dosen Pembimbing : Dosen Umniy Salamah, S.T. MMSI

Twitter social media is the 6th most frequently accessed social media platform in Indonesia. The cancellation of the U-20 World Cup in Indonesia has been widely discussed by the Indonesian community, especially football enthusiasts. The numerous issues surrounding this cancellation have generated both support and opposition. The aim of this research is to gain a deeper understanding of public sentiment. The research seeks to classify public reviews as either positive or negative. This study is a quantitative research on sentiment analysis that extracts sentiments regarding the cancellation of the U-20 World Cup in Indonesia on Twitter, with the goal of applying the Naive Bayes and Support Vector Machine (SVM) algorithms. The research flow uses the SEMMA method (Sample, Explore, Modify, Model, and Assess). The sample involves collecting a dataset through crawling tweets, followed by selecting important attributes in the exploration phase. Subsequently, modification involves text preprocessing to structure the data. The model for labeling uses each method to be tested. After labeling, the data will be classified using both methods. Assessment involves evaluating the results using confusion matrix and k-fold cross-validation. The results of this research will compare the two algorithms, SVM and Naive Bayes.

Keywords: Sentiment Analysis, U-20 World Cup, Support Vector Machine (SVM), Naive Bayes, SEMMA

DAFTAR ISI

HALAMAN PERNYATAAN KARYA SENDIRI	iii
HALAMAN PENGESAHAN	iv
KATA PENGANTAR.....	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS.....	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix-x
DAFTAR GAMBAR	xi
BAB I	1
PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Tujuan	4
1.4 Batasan Masalah.....	4
1.5 Manfaat Penelitian.....	5
1.6 Sistematika Penulisan.....	5
BAB II.....	7
TINJAUAN PUSTAKA.....	7
2.1 Teori Utama	7
2.2 Teori Pendukung	12
BAB III.....	58
METODE PENELITIAN	58
3.1 Pendekatan Penelitian	58
3.2 Desain Penelitian.....	58
3.3 Subjek Penelitian.....	60
3.4 Instrumen Penelitian.....	61
3.5 Teknik Pengumpulan Data	62
3.6 Analisis Data	63
3.7 Prosedur Penelitian.....	63
3.8 Evaluasi Hasil Penelitian.....	64
3.9 Timeline Penelitian	64
BAB IV	66
HASIL DAN PEMBAHASAN	66
4.1 Dataset.....	66
4.2 Explore	66
4.3 Preprocessing Data	67
4.3.1. Cleaning dan Case Folding	67
4.3.2. Normalisasi Teks.....	68
4.3.3. Stopword Removal.....	69
4.3.4. Tokenisasi.....	70
4.3.5. Stemming	71
4.4 Labeling.....	71
4.5 Wordcloud.....	73
4.6 Name Entity Recognition	73
4.7 Naive Bayes	74
4.8 Support Vector Machine	76

4.9 Assess.....	78
BAB V.....	85
PENUTUP.....	85
5.1 Kesimpulan.....	85
5.2 Saran.....	86
DAFTAR PUSTAKA	87
LAMPIRAN.....	90
Lampiran 1 Kartu Asistensi.....	90
Lampiran 2 Curriculum Vitae	91
Lampiran 3 Form Hasil Revisi Dosen Penguji.....	93
Lampiran 4 Hasil Cek Turnitin	95



DAFTAR GAMBAR

Gambar 1. 1 Data Persentase Penggunaan Sosial Media di Indonesia	2
Gambar 1. 2 Data Searches Google di Indonesia.....	3
Gambar 3. 1 Flowchart Naïve Bayes	59
Gambar 3. 2 Flowchart SVM	60
Gambar 3. 3 Pengumpulan Data	62
Gambar 3. 4 Prosedur Penelitian.....	63
Gambar 4. 1 Hasil Crawling Data Twitter	66
Gambar 4. 2 Menampilkan dataset.....	67
Gambar 4. 3.1 Hasil proses pengecekan data duplikat dan kosong	67
Gambar 4. 3.2 Kode cleaning dan case folding.....	68
Gambar 4. 3.3 Hasil cleaning dan case folding.....	68
Gambar 4. 3.4 Normalisasi.....	69
Gambar 4. 3.5 Stopword Removal.....	69
Gambar 4. 3.6 Tokenisasi.....	70
Gambar 4. 3.7 Stemming.....	71
Gambar 4. 2.1 Kode Labeling	72
Gambar 4. 2.2 Hasil Labeling	72
Gambar 4. 5 Wordcloud	73
Gambar 4.6.1 Name Entity Recognition	74
Gambar 4. 7.1 Naïve Bayes.....	75
Gambar 4. 7.2 Hasil Naïve Bayes	76
Gambar 4. 8 Support Vector Machine	77
Gambar 4. 8.2 Hasil Support Vector Machine	78
Gambar 4. 9.1 Cross Validation Naïve Bayes.....	79
Gambar 4. 9.2 Cross Validation Suppoort Vector Machine	80
Gambar 4. 9.3 Tweet Positif	82
Gambar 4. 9.4 Tweet Negatif	83
Gambar 4. 9.5 Tweet Netral	83

UNIVERSITAS
MERCU BUANA