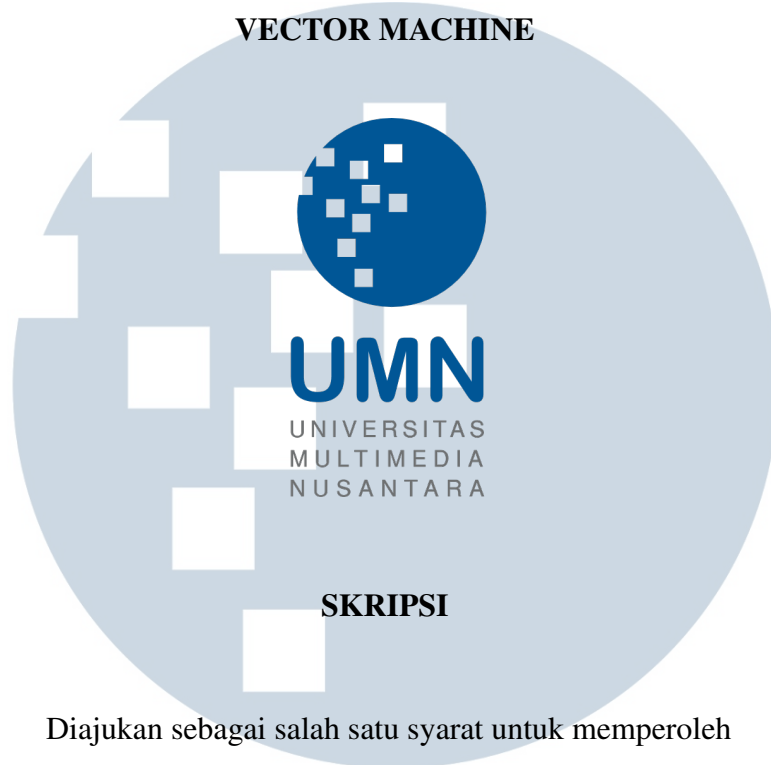


**ANALISIS SENTIMEN TERHADAP PERTANDINGAN PIALA ASIA U-23
2024 INDONESIA VS QATAR MENGGUNAKAN METODE SUPPORT
VECTOR MACHINE**



SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

Ryan Samudra Ginting

00000026250

UMN

UNIVERSITAS

MULTIMEDIA

NUSANTARA

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2024

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Ryan Samudra Ginting

Nomor Induk Mahasiswa : 00000026250

Program Studi : Informatika

Skripsi dengan judul:

Analisis Sentimen Terhadap Pertandingan Piala Asia U-23 2024 Indonesia vs Qatar
Menggunakan Metode Support Vector Machine

merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/ penyimpangan, baik dalam pelaksanaan Skripsi maupun dalam penulisan laporan Skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas akhir yang telah saya tempuh.

Tangerang, 22 Mei 2024



(Ryan Samudra Ginting)

UMM
UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

**ANALISIS SENTIMEN TERHADAP PERTANDINGAN PIALA ASIA U-23
2024 INDONESIA VS QATAR MENGGUNAKAN METODE SUPPORT
VECTOR MACHINE**

oleh

Nama : Ryan Samudra Ginting
NIM : 00000026250
Program Studi : Informatika
Fakultas : Fakultas Teknik dan
Informatika

Telah diujikan pada hari 11 Juni 2024
Pukul 08.00 s/d 10.00 dan dinyatakan
LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang
a/n *N Prastomo* 25/6-24
Niki Prastomo
(Yaman Khaeruzzaman, M.Sc.)
NIDN: 0413057104

Penguji
Eunike Endariahna
(Eunike Endariahna Surbakti, S.Kom., M.T.I.)
NIDN: 0322099401

Pembimbing
Fenina Adline Twince Tobing
(Fenina Adline Twince Tobing, S.Kom., M.Kom)
NIDN: 0406058802

PJS Ketua Program Studi Informatika,
N Prastomo 25/6-24
(Dr. Eng. Niki Prastomo, S.T., M.Sc.)
NIDN: 0419128203

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Multimedia Nusantara, saya yang bertanda tangan di bawah ini:

Nama : Ryan Samudra Ginting
NIM : 00000026250
Program Studi : Informatika
Fakultas : Teknik dan Informatika
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada **Universitas Multimedia Nusantara** hak Bebas Royalti Non-eksklusif (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul:

**ANALISIS SENTIMEN TERHADAP PERTANDINGAN PIALA ASIA U-23
2024 INDONESIA VS QATAR MENGGUNAKAN METODE SUPPORT
VECTOR MACHINE**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non eksklusif ini Universitas Multimedia Nusantara berhak menyimpan, mengalih media / format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis / pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Tangerang, 22 Mei 2024

Yang menyatakan



Ryan Samudra Ginting

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

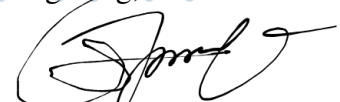
KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: Analisis Sentimen Terhadap Pertandingan Piala Asia U-23 2024 Indonesia vs Qatar Menggunakan Metode Support Vector Machine dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana/Magister Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Yaman Khaeruzzaman, M.Sc., selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Ibu Fenina Adline Twince Tobing, S.Kom., M.Kom, sebagai Pembimbing pertama yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya tesis ini.
5. Teman, kerabat dan rekan kerja yang telah memberikan dukungan dan motivasi untuk penulis.
6. Terimakasih untuk alm. Yustriagus Ginting selaku bapak kandung saya yang telah memotivasi saya untuk menyelesaikan tesis ini untuk mewujudkan cita-cita almarhum.

Semoga skripsi ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 22 Mei 2024



Ryan Samudra Ginting

**ANALISIS SENTIMEN TERHADAP PERTANDINGAN PIALA ASIA U-23
2024 INDONESIA VS QATAR MENGGUNAKAN METODE SUPPORT
VECTOR MACHINE**

Ryan Samudra Ginting

ABSTRAK

Pertandingan Indonesia vs Qatar di piala asia U-23 tahun 2024 menjadi perdebatan yang sedang hangat dibicarakan dikalangan masyarakat Indonesia karena diyakini banyak kecurangan dalam pertandingan tersebut. Banyaknya opini pro dan kontra yang di unggah oleh masyarakat Indonesia di media sosial twitter menjadikan opini tersebut dapat menjadi data untuk menilai pendapat yang dibagi menjadi tiga kategori yaitu positif, negatif dan netral. Dari 2258 data yang telah dikumpulkan, dilakukan labeling data menggunakan vader dan menghasilkan 30.9% sentimen positif, 49.8% sentimen negatif, dan 19.3% sentimen netral. Setelah dibagi menjadi tiga kategori selanjutnya akan dilakukan analisis sentimen menggunakan metode *support vector machine* dengan menggunakan confusion matrix untuk mengukur performa dari klasifikasi maka dapat dinilai accuracy, precision, recall dan f1-score. Didapatkan *kernel rbf* memiliki akurasi tertinggi dengan *Accuracy* 74.89%, *F1-Score* 74.30%, *Precision* 79.46% dan *Recall* 74.89% dengan perbandingan rasio data 60:40. Selain itu dari data yang didapatkan juga berhasil menganalisis teks berbahasa indonesia yang di input secara eksternal dan di kategorikan menjadi positif, netral dan negatif. Dari hasil yang telah didapatkan bahwa metode *support vector machine* berhasil di implementasi dalam analisis sentimen pertandingan Indonesia vs Qatar di piala asia U-23 tahun 2024 pada media sosial twitter.

Kata kunci: *Analisis Sentimen, confusion matrix, Piala Asia U-23, Indonesia vs Qatar, Support Vector Machine, Twitter.*

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

Sentiment Analysis of the 2024 U23 Asian Cup Match Indonesia vs Qatar Using the Support Vector Machine Method

Ryan Samudra Ginting

ABSTRACT

The match between Indonesia and Qatar in the 2024 U-23 Asian Cup has become a hot debate among the Indonesian people because it is believed that there was a lot of cheating in the match. The many pro and con opinions that have been uploaded by the Indonesian people on Twitter social media can be used as data to assess opinions, which can be divided into three categories, namely positive, negative, and neutral. Out of the 2258 data points that were collected, data labeling was performed using Vader, resulting in 30.9% positive sentiment, 49.8% negative sentiment, and 19.3% neutral sentiment. After being divided into three categories, a sentiment analysis is then performed using the *support vector machine* method with the use of a confusion matrix to measure the performance of the classification. Accuracy, precision, recall, and F1-score can then be assessed. The RBF kernel was found to have the highest accuracy with an *Accuracy* 74.89%, *F1-Score* 74.30%, *Precision* 79.46%, and *Recall* 74.89% with a data ratio of 60:40. In addition, the data obtained also succeeded in analyzing Indonesian text that was input externally and categorized it as positive, neutral, and negative. Based on the results obtained, the support vector machine method was successfully implemented in the sentiment analysis of the Indonesia vs Qatar match in the 2024 U-23 Asian Cup on Twitter social media.

Keywords: *confusion matrix, Asian Cup U-23, Indonesia vs Qatar, Sentiment Analysis Support Vector Machine, Twitter*



DAFTAR ISI

HALAMAN PENGESAHAN	ii
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	iii
HALAMAN PERSETUJUAN PUBLIKASI ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	6
2.1 Piala Asia U-23	6
2.2 Analisa Sentimen	6
2.3 VADER	6
2.4 TF-IDF	7
2.5 <i>Support Vector Machine</i>	8
2.5.1 Kernel Linear	8
2.5.2 Kernel Polynomial	9
2.5.3 Kernel RBF	10
2.6 <i>Confusion Matrix</i>	12
2.6.1 <i>Accuracy</i>	13
2.6.2 <i>Precision</i>	14
2.6.3 <i>Recall</i>	14
2.6.4 <i>F1-score</i>	15
BAB 3 METODOLOGI PENELITIAN	16
3.1 Metodologi Penelitian	16
3.2 Perancangan Sistem	17
3.2.1 Gambaran Umum Perancangan	17
3.2.2 <i>Crawling Data</i>	18
3.2.3 <i>Labeling Data</i>	18
3.2.4 <i>Text Preprocessing</i>	20
3.2.5 <i>Train-Test Split Data</i>	21
3.2.6 <i>Penerapan TF-IDF</i>	21
3.2.7 <i>Penerapan SVM</i>	22
3.2.8 <i>Evaluation</i>	23
BAB 4 HASIL DAN DISKUSI	24
4.1 Spesifikasi Sistem	24
4.2 Implementasi Sistem	24
4.2.1 <i>Crawling Data</i>	24
4.2.2 <i>Labelling data</i>	26
4.2.3 <i>Text Pre-processing</i>	28

4.2.4	Training dan Testing Split Data	33
4.2.5	Apply TF-IDF	33
4.2.6	Apply SVM	34
4.2.7	Evaluasi Model	34
4.3	Uji Coba	36
4.3.1	Pengujian dengan Pembagian Data Training dan Data Testing	36
4.3.2	Pengujian dengan Input teks Berbahasa Indonesia	44
4.3.3	Pengujian dengan Input File Csv Berbahasa Indonesia	44
BAB 5	SIMPULAN DAN SARAN	46
5.1	Simpulan	46
5.2	Saran	46
	DAFTAR PUSTAKA	47



DAFTAR GAMBAR

Gambar 2.1	Kernel Linear	9
Gambar 2.2	Kernel Polynomial	10
Gambar 2.3	Kernel RBF	11
Gambar 2.4	<i>Confusion Matrix</i>	13
Gambar 3.1	Flowchart Gambaran Umum Perancangan	17
Gambar 3.2	Flowchart Labeling Data	19
Gambar 3.3	Flowchart Text Preprocessing	20
Gambar 3.4	Flowchart Apply TF-IDF	22
Gambar 3.5	Flowchart Apply SVM	23
Gambar 4.1	kode input file csv	25
Gambar 4.2	Hasil Crawling	25
Gambar 4.3	Potongan kode translate	26
Gambar 4.4	Hasil translate data	26
Gambar 4.5	Kode Vader	27
Gambar 4.6	Hasil Pelabelan Vader	27
Gambar 4.7	Pie Chart Hasil Vader	28
Gambar 4.8	Kode Case Folding	29
Gambar 4.9	Hasil Case Folding	29
Gambar 4.10	Kode Cleaning Data	29
Gambar 4.11	Hasil Cleaning Data	29
Gambar 4.12	Kode Tokenize	30
Gambar 4.13	Hasil Tokenize	30
Gambar 4.14	Kode Normalisasi	31
Gambar 4.15	Hasil Normalisasi	31
Gambar 4.16	Kode Stopwords	32
Gambar 4.17	Hasil Stopwords	32
Gambar 4.18	Kode Stemming	33
Gambar 4.19	Hasil Stemming	33
Gambar 4.20	Kode Input Teks	33
Gambar 4.21	Kode TF-IDF	34
Gambar 4.22	Kode Apply SVM	34
Gambar 4.23	Kode Evaluasi Model	35
Gambar 4.24	Hasil Evaluasi Model	35
Gambar 4.25	Confusion Matrix linear 60:40	37
Gambar 4.26	Confusion Matrix polynomial 60:40	37
Gambar 4.27	Confusion Matrix rbf 60:40	38
Gambar 4.28	Confusion Matrix linear 70:30	39
Gambar 4.29	Confusion Matrix polynomial 70:30	40
Gambar 4.30	Confusion Matrix rbf 70:30	40
Gambar 4.31	Confusion Matrix linear 80:30	41
Gambar 4.32	Confusion Matrix polynomial 80:20	42
Gambar 4.33	Confusion Matrix rbf 80:20	42
Gambar 4.34	Kode uji coba input teks	44
Gambar 4.35	Hasil Uji coba input teks	44
Gambar 4.36	Kode uji coba input csv	45
Gambar 4.37	Hasil uji coba input csv	45

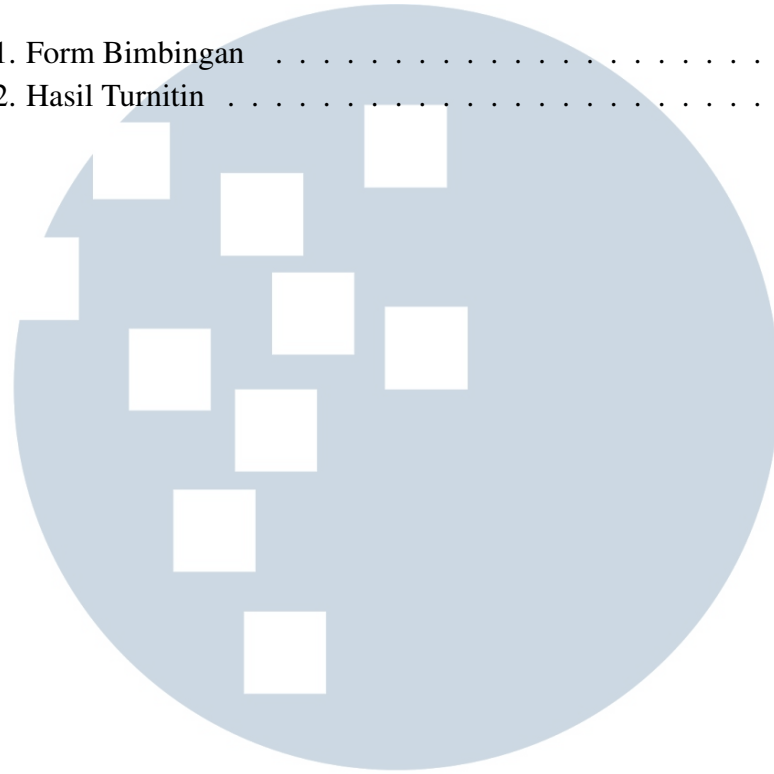
DAFTAR TABEL

Tabel 4.1	Tabel Perbandingan Dataset 60:40	38
Tabel 4.2	Tabel Perbandingan Dataset 70:30	41
Tabel 4.3	Tabel Perbandingan Dataset 80:20	43
Tabel 4.4	Tabel Perbandingan Hasil Pengujian Dataset 60:40, 70:30, dan 80:20	43



DAFTAR LAMPIRAN

Lampiran 1. Form Bimbingan	50
Lampiran 2. Hasil Turnitin	51



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA