



**Analisis Sentimen Komentar Pada Saluran Youtube Terhadap
Batalnya Indonesia Menjadi Tuan Rumah Piala Dunia U-20
Menggunakan SVM dan Naive Bayes**

LAPORAN SKRIPSI

Nabiila Salsabiila

41519120026

**UNIVERSITAS
MERCU BUANA**

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

2023



**Analisis Sentimen Komentar Pada Saluran Youtube Terhadap
Batalnya Indonesia Menjadi Tuan Rumah Piala Dunia U-20
Menggunakan SVM dan Naive Bayes**

LAPORAN SKRIPSI

Nabiila Salsabiila

41519120026

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**UNIVERSITAS
MERCU BUANA**

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

2023

HALAMAN PENYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : NABIILA SALSABIILA
NIM : 41519120026
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Sentimen Komentar Pada Saluran
Youtube Terhadap Batalnya Indonesia Menjadi
Tuan Rumah Piala Dunia U-20 Menggunakan
SVM dan Naive Bayes

Menyatakan bahwa Laporan Skripsi 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 Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 18 Januari 2024

UNIVERSITAS
MERCU BUANA



Nabiila Salsabiila

LEMBAR PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Nabiila Salsabiila

NIM : 41519120026

Program Studi : Teknik Informatika

Judul Laporan : Analisis Sentimen Komentar Pada Saluran Youtube Terhadap
Skripsi Batalnya Indonesia Menjadi Tuan Rumah Piala Dunia U-20
Menggunakan SVM 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 : Mohamad Yusuf, S.Kom., M.C.S

NIDN : 0307097606

Ketua Penguji : Yustika Erliani, SE, MMSI

NIDN : 0305057809

Penguji 1 : Sabar Rudiarto, S.Kom, M.Kom

NIDN : 0309036902

Penguji 2 : Muhaimin Hasanudin, ST., M.Kom

NIDN : 0420027508

Jakarta, 19 Februari 2024

Mengetahui,

Dekan

Ketua Program Studi



Dr. Bambang Jokonowo, S.Si., M.T.I



Dr. Hadi Santoso, S.Kom., M.Kom

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Laporan Skripsi ini. Penulisan Laporan Skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada Fakultas Ilmu Komputer Universitas Mercu Buana. 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. Prof. Dr. Ir. Andi Adriansyah, M.Eng selaku Rektor Universitas Mercu Buana
2. Dr. Bambang Jokonowo, S.Si., M.T.I selaku Dekan Fakultas Ilmu Komputer
3. Hadi Santoso, Dr, S.Kom, M.Kom selaku Ketua Program Studi Teknik Informatika
4. Mohamad Yusuf, S.Kom., M.C.S selaku Dosen Pembimbing serta Dosen Pembimbing Akademik yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini
5. Sabar Rudiarto, S.Kom, M.Kom dan Muhaimin Hasanudin, ST., M.Kom selaku Dosen Penguji Tugas Akhir atas koreksi dan arahan serta masukannya.
6. Kedua orang tua saya yang selalu mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercu Buana.
7. Rekan – rekan dan pihak lain yang secara konsisten berbagi informasi serta memberikan dukungan dalam berbagai bentuk.

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Laporan Skripsi ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 18 Januari 2024



Nabiila Salsabiila

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR
UNTUK KEPENTINGAN AKADEMIS**

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

Nama : Nabiila Salsabiila
NIM : 41519120026
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Sentimen Komentar Pada Saluran Youtube Terhadap Batalnya Indonesia Menjadi Tuan Rumah Piala Dunia U-20 Menggunakan SVM dan Naive Bayes

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (Non-exclusive Royalty-Fee Right)** atas karya ilmiah saya yang berjudul diatas 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 Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 18 Januari 2024

Yang menyatakan,


Nabiila Salsabiila

ABSTRAK

Nama : Nabiila Salsabiila
NIM : 41519120026
Program Studi : Teknik Informatika
Judul Laporan : Analisis Sentimen Komentar Pada Saluran Youtube
Skripsi Terhadap Batalnya Indonesia Menjadi Tuan Rumah Piala
Dunia U-20 Menggunakan SVM dan Naïve Bayes
Pembimbing : Mohamad Yusuf, S.Kom., M.C.S

Youtube merupakan aplikasi sosial media terbesar untuk berbagi informasi melalui video. Sebagai media dunia maya, Youtube memberi akses bagi pengguna dalam nonton, unggah, dan komentar pada video. Di Indonesia, Youtube telah mengalami perkembangan yang pesat. Sehingga Youtube dapat menjadi sarana untuk melakukan penelitian analisis sentimen terhadap batalnya Indonesia menjadi tuan rumah Piala Dunia U-20. Data yang diperoleh dari Youtube tersebut diproses dengan text mining kemudian diklasifikasikan menjadi dua kelas yaitu positif dan negatif berdasarkan kata-kata yang terkandung di dalamnya. Bahasa pemrograman yang digunakan yaitu Python. Algoritma yang digunakan yaitu algoritma SVM dan Naive Bayes. Hasil penelitian menunjukkan bahwa pandangan masyarakat terhadap video yang dianalisa banyak mengandung opini negatif. Hasil Klasifikasi pada SVM memiliki hasil akurasi sebesar 86% dan nilai kappa sebesar 68%. Sedangkan, Naive Bayes memiliki hasil akurasi sebesar 80% dan nilai kappa sebesar 50%. Dari Hasil perbandingan disimpulkan bahwa algoritma SVM menggunakan RBF kernel lebih unggul dalam melakukan analisa pada penelitian ini.

Kata Kunci: Youtube, Analisis Sentimen, Piala Dunia U-20, Naive Bayes, SVM

ABSTRACT

Name : *Nabiila Salsabiila*
NIM : *41519120026*
Study Program : *Teknik Informatika*
Title Thesis : *Analisis Sentimen Komentar Pada Saluran Youtube Terhadap Batalnya Indonesia Menjadi Tuan Rumah Piala Dunia U-20 Menggunakan SVM dan Naïve Bayes*
Counsellor : *Mohamad Yusuf, S.Kom., M.C.S*

Youtube is the largest social media application for sharing information via video. As a virtual media, Youtube provides access for users to watch, upload and comment on videos. In Indonesia, Youtube has experienced rapid development. So Youtube can be a means for conducting sentiment analysis research regarding Indonesia's cancellation of hosting the U-20 World Cup. The data obtained from YouTube is processed using text mining and then classified into two classes, namely positive and negative based on the words contained in it. The programming language used is Python. The algorithms used are the SVM and Naive Bayes algorithms. The results of the research show that people's views on the videos analyzed contain many negative opinions. Classification results on SVM have an accuracy of 86% and a kappa value of 68%. Meanwhile, Naive Bayes has an accuracy of 80% and a kappa value of 50%. From the comparison results, it was concluded that the SVM algorithm using the RBF kernel was superior in carrying out the analysis in this research.

Keywords: *Youtube, Sentiment Analysis, Piala Dunia U-20, Naive Bayes, SVM*

DAFTAR ISI

HALAMAN JUDUL	I
HALAMAN PERNYATAAN KARYA SENDIRI.....	II
LEMBAR PENGESAHAN.....	III
KATA PENGANTAR.....	IV
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	V
ABSTRAK	VI
ABSTRACT	VII
DAFTAR ISI.....	VIII
DAFTAR TABEL.....	XI
DAFTAR GAMBAR.....	XII
DAFTAR LAMPIRAN	XIV
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah.....	5
BAB II TINJAUAN PUSTAKA	6
2.1 Teori Penelitian Terkait.....	6
2.1.1 Analisis Sentimen	6
2.1.2 Text Mining	6
2.1.3 Youtube.....	6
2.1.4 Piala Dunia U-20	7
2.1.5 <i>Support Vector Machine</i> (SVM).....	7

2.1.6 <i>Naïve Bayes</i>	8
2.1.7 <i>Confusion Matrix</i>	9
2.2 Penelitian Terkait	10
2.3 <i>Critical Review</i>	18
BAB III METODOLOGI.....	27
3.1 Jenis Penelitian.....	27
3.2 Teknik Pengumpulan Data.....	27
3.3 Diagram Alir Penelitian	27
3.3.1 Pengumpulan Data	28
3.3.2 Tahap Preprocessing.....	28
3.3.3 Pelabelan Data	29
3.3.4 Pembobotan kata <i>Term Frequency-Inverse Document Frequency</i>	29
3.4.5 Pemodelan klasifikasi dengan algoritma SVM dan <i>Naive Bayes</i>	30
3.4.6 Evaluasi Performa Model	30
3.4.7 Visualisasi Data	30
BAB IV HASIL DAN PEMBAHASAN	31
4.1 Dataset.....	31
4.2 Pre-processing.....	32
4.3 Label Dataset.....	38
4.4 Pembobotan kata pada <i>Term Frequency-Inverse Document Frequency</i>	39
4.5 Pemodelan klasifikasi Algoritma <i>Naive Bayes</i> dan SVM.....	40
4.6 Evaluasi Performa Model.....	41
4.7 Visualisasi Data.....	42
4.8 Hasil Analisis	44
BAB V PENUTUP	45
5.1 Kesimpulan	45

5.2 Saran.....	45
DAFTAR PUSTAKA	46
LAMPIRAN.....	48



UNIVERSITAS
MERCU BUANA

DAFTAR TABEL

Tabel 1.1 Data Masalah.....	2
Tabel 2.1 <i>Confusion Matrix</i>	9
Tabel 2.2 Penelitian Terkait	18
Tabel 2.3 <i>Critical Review</i>	25
Tabel 4.4 Hasil pemodelan SVM dan <i>Naïve Bayes</i>	44



DAFTAR GAMBAR

Gambar 3.1 Diagram Alir Penelitian	28
Gambar 4.1 Tampilan Dataset	31
Gambar 4.2 Informasi Dataset	31
Gambar 4.3 Proses <i>Cleaning</i>	32
Gambar 4.4 Hasil Data <i>Cleaning</i>	33
Gambar 4.5 Proses <i>Case Folding</i>	33
Gambar 4.6 Hasil Data <i>Case Folding</i>	33
Gambar 4.7 Implementasi <i>tokenizing</i>	34
Gambar 4.8 Proses Normalisasi Kata	34
Gambar 4.9 Hasil Normalisasi Kata	35
Gambar 4.10 Proses <i>Stopword Removal</i>	35
Gambar 4.11 Hasil Data <i>Stopword Removal</i>	35
Gambar 4.12 Proses implementasi <i>Stemming</i>	36
Gambar 4.13 Implementasi <i>Stemming</i>	37
Gambar 4.14 Implementasi <i>missing value</i>	37
Gambar 4.15 Implementasi menghapus data duplikat	38
Gambar 4.16 Menampilkan dataset hasil <i>pre-processing</i>	38
Gambar 4.17 Menampilkan implementasi pelabelan	39
Gambar 4.18 Implementasi <i>Term Frequency-Inverse Document</i> (TF-IDF)	40
Gambar 4.19 Implementasi pemodelan Algoritma <i>Naive Bayes</i>	40
Gambar 4.20 Implementasi prediksi sentimen Algoritma <i>Naive Bayes</i>	40
Gambar 4.21 Implementasi pemodelan menggunakan Algoritma SVM	41
Gambar 4.22 Hasil Evaluasi Pemodelan <i>Naive Bayes</i>	41
Gambar 4.23 Hasil Evaluasi Pemodelan SVM	42

Gambar 4.24 Hasil visualisasi sentiment analisis	42
Gambar 4.25 Implementasi visualisasi <i>confusion matrix</i>	43
Gambar 4.26 Hasil Visualisasi <i>Confusion Matrix</i> Naive Bayes dan SVM.....	43



DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi	48
Lampiran 2 Bukti Submit / Publish Artikel Ilmah	49
Lampiran 3 Bukti Submit Jurnal	50
Lampiran 4 Naskah Jurnal	51
Lampiran 5 Surat Pernyataan HKI	61
Lampiran 6 Curriculum Vitae (CV)	62
Lampiran 7 Sertifikat BNSP	63

