



Perbandingan Kinerja Algoritma Naive Bayes, SVM, dan Random Forest Terhadap Analisis Sentimen Tragedi Kanjuruhan

LAPORAN SKRIPSI

Ifan Haidar
41519010216

UNIVERSITAS
MERCU BUANA

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2023



Perbandingan Kinerja Algoritma Naive Bayes, SVM, dan Random Forest Terhadap Analisis Sentimen Tragedi Kanjuruhan

LAPORAN SKRIPSI

**Ifan Haidar
41519010216**

**UNIVERSITAS
MERCU BUANA**

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2023**

HALAMAN PERNYATAAN KARYA SENDIRI

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Ifan Haidar

NIM : 41519010216

Program Studi : Teknik Informatika

Judul Laporan Skripsi : Perbandingan Kinerja Algoritma Naive Bayes, SVM, dan
Random Forest Terhadap Analisis Sentimen Tragedi

Kanjuruhan

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, 29 Juli 2023

UNIVERSITAS
MERCU BUANA



Ifan Haidar

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Ifan Haidar
NIM : 41519010216
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Kinerja Algoritma Naive Bayes, SVM,
dan Random Forest Terhadap Analisis Sentimen
Tragedi Kanjuruhan

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 : Vina Ayumi, S.kom, M.kom ()
NIDN : 0311109003
Penguji 1 : Dr. Harwikarya, MT ()
NIDN : 0014075805
Penguji 2 : Harni Kusniyati, ST.M.Kom ()
NIDN : 0324068101

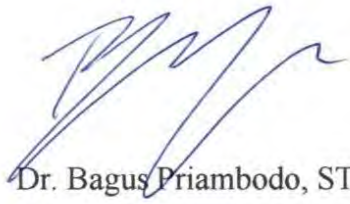
Jakarta, 14 Agustus 2023

Mengetahui,

Dekan

Ketua Program Studi


Dr. Bambang Jokonowo, S.Si., MTI


Dr. Bagus Priambodo, ST., MTI., PhD

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. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer
3. Bapak Bagus Priambodo, ST., MTI., PhD. selaku Ketua Program Studi Teknik Informatika
4. Ibu Vina Ayumi, S.kom, M.kom selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini;
5. (Nama Dosen Penguji) selaku Dosen Penguji Tugas Akhir atas koreksi dan arahan serta masukannya.
6. Bapak Sabar Rudiarto, S.Kom., M.Kom selaku dosen pembimbing akademik saya selama perkuliahan.
6. Seluruh Dosen Program Studi Teknik Informatika yang sudah memberikan ilmu yang bermanfaat selama kuliah berlangsung. Memberi kesempatan untuk belajar, berkarya dan juga berkembang.
7. Orang tua dan keluarga yang banyak memberikan semangat dan motivasi sehingga dapat menyelesaikan penulisan skripsi ini.

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, 29 Juli 2023



Ifan Haidar

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

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

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

Nama : Ifan Haidar
NIM : 41519010216
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Kinerja Algoritma Naive Bayes, SVM, dan Random Forest Terhadap Analisis Sentimen Tragedi Kanjuruhan

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. 29 Juli 2023



Ifan Haidar

ABSTRAK

Nama : Ifan Haidar
NIM : 41519010216
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Kinerja Algoritma Naive Bayes, SVM,
dan Random Forest Terhadap Analisis Sentimen
Tragedi Kanjuruhan
Pembimbing : Vina Ayumi, S.kom, M.kom

Tragedi di Stadion Kanjuruhan Malang terjadi pada 1 Oktober 2022 usai pertandingan sepak bola Arema FC yang kalah 2-3 melawan Persebaya. Tragedi maut Kanjuruhan memakan korban 754 orang, 132 dinyatakan meninggal dunia, 596 luka ringan sedang dan luka berat 26 orang, Minggu (15/10/2022). Penyebab korban tragedi Stadion Kanjuruhan Malang meninggal dunia adalah karena mayoritas mengalami sesak nafas dan terinjak-injak karena panik setelah dilakukan pelepasan gas air mata oleh pihak kepolisian. Situasi panik karena rusuh menyebabkan banyaknya korban dari para supporter yang terluka hingga meninggal dunia. Atas kejadian tragedi Kanjuruhan yang menelan banyak korban menyebabkan banyaknya komentar di sosial media salah satunya adalah Twitter. Dengan banyaknya pengguna Twitter yang menyampaikan opini-opini, maka dapat dimanfaatkan untuk mencari sebuah informasi dan polaritas positif dan negatifnya opini-opini tersebut. Pada penelitian ini dilakukan Perbandingan Kinerja Algoritma Naive Bayes, Support Vector Machine (SVM), dan Random Forest Terhadap Analisis Sentimen Tragedi Kanjuruhan. Hasil eksperimen menunjukkan bahwa kinerja dari Metode SVM menghasilkan nilai akurasi yang lebih baik dibandingkan metode Random Forest dan Naive Bayes, dengan rata-rata akurasi mencapai 84.13%. Untuk metode Random Forest memiliki nilai akurasi rata-rata sebesar 81.12%. Sedangkan untuk metode Naive Bayes dengan rata-rata akurasi mencapai 79.95%.

Kata Kunci : Kanjuruhan, Sentimen, Naive Bayes, Support Vector Machine, Random Forest

ABSTRACT

Name : Ifan Haidar
NIM : 41519010216
Study Program : Informatics
Title Thesis : Performance Comparison of Naive Bayes, Support SVM, and Random Forest Algorithms in Sentiment Analysis of the Kanjuruhan Tragedy
Counsellor : Vina Ayumi, S.kom, M.kom

The tragedy at Kanjuruhan Stadium in Malang occurred on October 1, 2022, after the Arema FC football match, which they lost 2-3 against Persebaya. The Kanjuruhan tragedy claimed 754 victims, with 132 fatalities, 596 suffering moderate injuries, and 26 sustaining severe injuries on Sunday (October 15, 2022). The cause of death for the victims of the Kanjuruhan Stadium tragedy in Malang was primarily due to respiratory distress and trampling caused by panic following the release of tear gas by the police. The chaotic and panicked situation led to numerous casualties among the supporters, resulting in injuries and fatalities. The tragic incident at Kanjuruhan, which claimed many victims, sparked a flurry of comments on social media, particularly Twitter. With numerous Twitter users expressing their opinions, this platform can be utilized to gather information and analyze the positive and negative sentiments expressed. This study compared the performance of the Naive Bayes, Support Vector Machine (SVM), and Random Forest algorithms in analyzing the sentiment surrounding the Kanjuruhan tragedy. The experimental results demonstrated that the SVM method exhibited better accuracy compared to the Random Forest and Naive Bayes methods, with an average accuracy of 84.13%. The Random Forest method achieved an average accuracy of 81.12%. Meanwhile, the Naive Bayes method yielded an average accuracy of 79.95%.

Keywords: Kanjuruhan, Sentiment Analysis, Naïve Bayes, Support Vector Machine, Random Forest

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN	iii
KATA PENGANTAR	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	3
1.5 Batasan Masalah.....	3
BAB II TINJAUAN PUSTAKA	5
2.1 Penelitian Terdahulu	5
2.2 Teori Pendukung	9
BAB III METODE PENELITIAN	11
3.1 Jenis Penelitian.....	11
3.2 Tahap Penelitian.....	11
BAB IV HASIL DAN PEMBAHASAN	17
4.1 Dataset.....	17
4.2 Preprocessing.....	17
4.3 Pembuatan Model.....	19
4.3.1 Pembobotan kata.....	19
4.3.2 Splitting data.....	20
4.3.3 Implementasi Algoritma Klasifikasi.....	20

4.4 Visualisasi data.....	21
4.5 Pengujian Model	22
4.7 Analisa Hasil	41
BAB V KESIMPULAN DAN SARAN	44
5.1 Kesimpulan.....	44
5.2 Saran.....	44
Daftar pustaka.....	45
LAMPIRAN.....	48
LAMPIRAN BIMBINGAN	48
Lampiran Luaran Tugas Akhir.....	49
Halaman Persetujuan.....	50
Lampiran Bukti Submit.....	51
Lampiran Naskah Artikel Jurnal	52
Curriculum Vitae.....	61
Lampiran Surat Pernyataan HKI.....	62
Sertifikat BNSP	63
Plagiarism Check.....	64



DAFTAR TABEL

Table 2.1 Penelitian Terkait.....	5
Tabel 3.1 Contoh Sentimen positif dan negatif	13
Tabel 4.1 Hasil Cleansing.....	18
Tabel 4.2 Hasil Case Folding.....	18
Tabel 4.3 Hasil Tokenizing.....	18
Tabel 4.4 Hasil Stopword Removal.....	19
Tabel 4.5 Hasil Stemming.....	19
Tabel 4.6 Jumlah Persentase Split Data.....	22
Tabel 4.7 TF-IDF dengan Naive Bayes split data 90:10	24
Tabel 4.8 TF-IDF dengan SVM split data 90:10	24
Tabel 4.9 TF-IDF dengan Random Forest split data 90:10	25
Tabel 4.10 TF-IDF dengan Naive Bayes split data 80:20	26
Tabel 4.11 TF-IDF dengan SVM split data 80:20	26
Tabel 4.12 TF-IDF dengan Random Forest split data 80:20	27
Tabel 4.13 TF-IDF dengan Naive Bayes split data 70:30	28
Tabel 4.14 TF-IDF dengan SVM split data 70:30	28
Tabel 4.15 TF-IDF dengan Random Forest split data 70:30	29
Tabel 4.16 TF-IDF dengan Naive Bayes split data 60:40	30
Tabel 4.17 TF-IDF dengan SVM split data 60:40	30
Tabel 4.18 TF-IDF dengan Random Forest split data 60:40	31
Tabel 4.19 TF-IDF dengan Naive Bayes split data 50:50	32
Tabel 4.20 TF-IDF dengan SVM split data 50:50	32
Tabel 4.21 TF-IDF dengan Random Forest split data 50:50	32
Tabel 4.22 TF-IDF dengan Naive Bayes split data 40:60	34
Tabel 4.23 TF-IDF dengan SVM split data 40:60	34
Tabel 4.24 TF-IDF dengan Random Forest split data 40:60	35
Tabel 4.25 TF-IDF dengan Naive Bayes split data 30:70	36
Tabel 4.26 TF-IDF dengan SVM split data 30:70	36
Tabel 4.27 TF-IDF dengan Random Forest split data 30:70	37
Tabel 4.28 TF-IDF dengan Naive Bayes split data 20:80	38
Tabel 4.29 TF-IDF dengan SVM split data 20:80	38
Tabel 4.30 TF-IDF dengan Random Forest split data 20:80	39

Tabel 4.31 TF-IDF dengan Naive Bayes split data 10:90	40
Tabel 4.32 TF-IDF dengan SVM split data 10:90	40
Tabel 4.33 TF-IDF dengan Random Forest split data 10:90	41
Tabel 4.34 Hasil perbandingan Akurasi.....	41



DAFTAR GAMBAR

Gambar 3.1 Tahapan Penelitian.....	11
Gambar 3.2 Proses Crawling Data.....	12
Gambar 3.3 Flowchart Preprocessing.....	13
Gambar 3.4 Flowchart Naive Bayes.....	15
Gambar 4.1 Hasil pengumpulan data.....	17
Gambar 4.2 Pembobotan TF-IDF.....	19
Gambar 4.3 Splitting Data.....	20
Gambar 4. 4 Implementasi Naive Bayes.....	20
Gambar 4.5 Implementasi Support Vector Machine (SVM).....	20
Gambar 4.6 Implementasi Random Forest.....	21
Gambar 4.7 Diagram Bar Data Analisis Sentimen.....	21
Gambar 4.8 Diagram Pie Data Analisis Sentimen.....	22
Gambar 4.9 Confusion Matrix Naive Bayes 90:10.....	23
Gambar 4.10 Confusion Matrix SVM 90:10.....	23
Gambar 4.11 Confusion Matrix Random Forest 90:10.....	24
Gambar 4.12 Confusion Matrix Naive Bayes 80:20.....	25
Gambar 4.13 Confusion Matrix SVM 80:20.....	25
Gambar 4.14 Confusion Matrix Random Forest 80:20.....	26
Gambar 4.15 Confusion Matrix Naive Bayes 70:30.....	27
Gambar 4.16 Confusion Matrix SVM 70:30.....	27
Gambar 4.17 Confusion Matrix Random Forest 70:30.....	28
Gambar 4.18 Confusion Matrix Naive Bayes 60:40.....	29
Gambar 4.19 Confusion Matrix SVM 60:40.....	29
Gambar 4.20 Confusion Matrix Random Forest 60:40.....	30
Gambar 4.21 Confusion Matrix Naive Bayes 50:50.....	31
Gambar 4.22 Confusion Matrix SVM 50:50.....	31
Gambar 4.23 Confusion Matrix Random Forest 50:50.....	32
Gambar 4.24 Confusion Matrix Naive Bayes 40:60.....	33
Gambar 4.25 Confusion Matrix SVM 40:60.....	33
Gambar 4.26 Confusion Matrix Random Forest 40:60.....	34
Gambar 4.27 Confusion Matrix Naive Bayes 30:70.....	35

Gambar 4.28 Confusion Matrix SVM 30:70	35
Gambar 4.29 Confusion Matrix Random Forest 30:70	36
Gambar 4.30 Confusion Matrix Naive Bayes 20:80.....	37
Gambar 4.31 Confusion Matrix SVM 20:80	37
Gambar 4.32 Confusion Matrix Random Forest 20:80	38
Gambar 4.33 Confusion Matrix Naive Bayes 10:90.....	39
Gambar 4.34 Confusion Matrix SVM 10:90	39
Gambar 4.35 Confusion Matrix Random Forest 10:90	40
Gambar 4.36 Visualisasi Perbandingan Akurasi Algoritma	42
Gambar 4.37 Visualisasi Wordcloud Positif.....	43
Gambar 4.38 Visualisasi Wordcloud Negatif	43

