

TUGAS AKHIR

**PERBANDINGAN METODE NAÏVE BAYES DAN
SUPPORT VECTOR MACHINE UNTUK ANALISIS
SENTIMEN TERHADAP VAKSIN ASTRAZENECA DI
TWITTER**



**EVA RAHMA INDRIYANI
18102011**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO
2022**

TUGAS AKHIR

**PERBANDINGAN METODE NAÏVE BAYES DAN
SUPPORT VECTOR MACHINE UNTUK ANALISIS
SENTIMEN TERHADAP VAKSIN ASTRAZENECA DI
TWITTER**

***COMPARISON OF NAÏVE BAYES AND SUPPORT
VECTOR MACHINE METHOD FOR SENTIMENT
ANALYSIS OF ASTRAZENECA VACCINE IN
TWITTER***

Disusun Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Komputer



**EVA RAHMA INDRIYANI
18102011**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO
2022**

Lembar Persetujuan Pembimbing

**PERBANDINGAN METODE NAÏVE BAYES DAN
SUPPORT VECTOR MACHINE UNTUK ANALISIS
SENTIMEN TERHADAP VAKSIN ASTRAZENECA DI
TWITTER**

Dipersiapkan dan Disusun Oleh

**EVA RAHMA INDRIYANI
18102011**

Diajukan untuk dapat diselenggarakan Sidang Skripsi sebagai Syarat
Kelulusan Mata Kuliah Tugas Akhir II

Pada
Program Studi S1 Informatika
Fakultas Informatika
Institut Teknologi Telkom Purwokerto

Disetujui oleh
Dosen Pembimbing Tugas Akhir II
Pada tanggal : 9 Juni 2022

Pembimbing I



(Paradise, S.Kom., M.Kom.)

NIDN 0624059501

Pembimbing II



(Merlinda Wibowo, S.T., M.Phil.)

NIDN 0612059203

Lembar Pengesahan Pembimbing

**PERBANDINGAN METODE NAÏVE BAYES DAN
SUPPORT VECTOR MACHINE UNTUK ANALISIS
SENTIMEN TERHADAP VAKSIN ASTRAZENECA DI
TWITTER**

**COMPARISON OF NAÏVE BAYES AND SUPPORT
VECTOR MACHINE METHOD FOR SENTIMENT
ANALYSIS OF ASTRAZENECA VACCINE IN TWITTER**

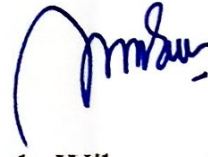
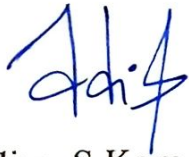
Dipersiapkan dan Disusun Oleh

**EVA RAHMA INDRIYANI
18102011**

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tugas Akhir
Pada hari 28 Juni 2022

Pembimbing I

Pembimbing II



(Paradise, S.Kom., M.Kom.)

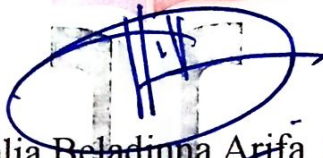
(Merlinda Wibowo, S.T., M.Phil.)

NIDN 0624059501

NIDN 0612059203

Tugas Akhir ini diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 28 Juni 2022

Kaprodi



(Amalia Beladina Arifa, S.Pd., M.Sc.)

NIDN 0606019201

Lembar Penetapan Penguji

**PERBANDINGAN METODE NAÏVE BAYES DAN
SUPPORT VECTOR MACHINE UNTUK ANALISIS
SENTIMEN TERHADAP VAKSIN ASTRAZENECA DI
TWITTER**

**COMPARISON OF NAÏVE BAYES AND SUPPORT
VECTOR MACHINE METHOD FOR SENTIMENT
ANALYSIS OF ASTRAZENECA VACCINE IN TWITTER**

Dipersiapkan dan Disusun Oleh

**EVA RAHMA INDRIYANI
18102011**

**Tugas Akhir Telah Diuji dan Dinilai Panitia Penguji
Program Studi Teknik Informatika
Fakultas Informatika
Institut Teknologi Telkom Purwokerto
Pada Tanggal : 28 Juni 2022**

Penguji I



**(Sena Wijayanto, S.Pd., M.T)
NIDN 0613109201**

Penguji II



**(Atika Ratna Dewi, S.Si, M.Sc)
NIDN 0615128703**

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertandatangan di bawah ini,

Nama mahasiswa : **Eva Rahma Indriyani**
NIM : **18102011**
Program Studi : **Teknik Informatika**

Menyatakan bahwa Tugas Akhir dengan judul berikut:

PERBANDINGAN METODE NAÏVE BAYES DAN *SUPPORT VECTOR MACHINE* UNTUK ANALISIS SENTIMEN TERHADAP VAKSIN ASTRAZENECA DI TWITTER

Dosen Pembimbing Utama : Paradise, S.Kom., M.Kom.
Dosen Pembimbing Pendamping : Merlinda Wibowo, S.T., M.Phil.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Institut Teknologi Telkom Purwokerto maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan, dan penelitian Saya Sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing.
3. Dalam Karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggungjawab Saya, bukan tanggungjawab Institut Teknologi Telkom Purwokerto.
5. Pernyataan ini Saya buat dengan sesungguhnya, apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka Saya bersedia menerima Sanksi Akademik dengan pencabutan gelar yang sudah diperoleh serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Purwokerto, 07 Juni 2022

Yang Menyatakan,



(Eva Rahma Indriyani)

KATA PENGANTAR

Bismillahirrahmanirrahim

Assalamualaikum, Wr. Wb.

Alhamdulillah rabbil ‘alamin segala puji bagi Allah SWT atas rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan Laporan Tugas Akhir yang berjudul “Perbandingan Metode Naïve Bayes dan *Support Vector Machine* Untuk Analisis Sentimen Terhadap Vaksin Astrazeneca di Twitter” dengan lancar dan baik.

Penyusunan Tugas Akhir ini ditujukan sebagai salah satu syarat untuk menyelesaikan pendidikan pada Program Studi Teknik Informatika, Fakultas Informatika, Institut Teknologi Telkom Purwokerto. Penulis mengetahui bahwa penyusunan Tugas Akhir ini tidak luput dari dukungan, bimbingan serta bantuan dari berbagai pihak. Untuk itu penulis ingin mengucapkan terima kasih sebanyak-banyaknya kepada pihak – pihak yang telah banyak membantu dan memberikan dukungan. Dengan kerendahan hati penulis ingin menyampaikan ucapan terima kasih kepada:

1. Allah SWT untuk segala nikmat-Nya sehingga penulis dapat menyelesaikan penyusunan Laporan Tugas Akhir dengan baik dan lancar.
2. Mama dan Ayah dan Adik untuk segala do’a, dukungan, nasihat, motivasi dan kasih sayang yang telah diberikan selama ini sampai sekarang.
3. Kepada keluarga besar saya yang telah memberikan banyak dukungan tanpa syarat.
4. Bapak Dr. Arfianto Fahmi, S.T., M.T., IPM sebagai Rektor Institut Teknologi Telkom Purwokerto
5. Bapak Auliya Burhanuddin, S.Si., M.Kom sebagai Dekan Fakultas Informatika dan Ibu Amalia Beladonna Arifa, S.Pd., M.Cs sebagai Ketua Program Studi S1 Teknik Informatika
6. Kepada Ibu Paradise, S.Kom, M.Kom dan Ibu Merlinda Wibowo, S.T., M.Phil selaku Dosen Pembimbing I dan Dosen Pembimbing II yang telah

memberikan dukungan serta ilmu dan motivasi kepada penulis selama perkuliahan maupun selama penyusunan Tugas Akhir ini.

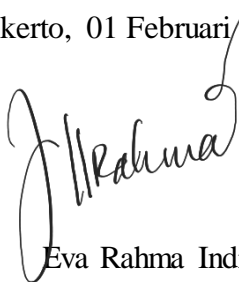
7. Kepada Pak Elfa dan Kak Rani yang telah mendorong dan membimbing saya untuk berkembang selama kuliah.
8. Kepada Riqqah Fadiyah Alya yang telah menjadi tempat saya berkeluh kesah dan telah banyak sekali memotivasi saya untuk tetap semangat dan pantang menyerah semoga kita sukses dan dapat meraih impian.
9. Kepada member iKON dan Hanbin yang telah menjadi alasan saya untuk mengejar impian dan menghibur saya saat berada di titik terendah
10. Kepada seluruh pihak yang telah membantu dalam penyusunan Tugas Akhir ini.

Semoga Laporan Tugas Akhir ini dapat dipergunakan dengan semestinya dan bermanfaat bagi pembaca yang berminat pada topik pembahasan pada Tugas Akhir yang telah penulis susun.

Terima Kasih

Wassalamu'alaikum, Wr. Wb.

Purwokerto, 01 Februari/2022



Eva Rahma Indriyani

DAFTAR ISI

LEMBAR SAMPUL LUAR LAPORAN TUGAS AKHIR	i
LEMBAR SAMPUL DALAM LAPORAN TUGAS AKHIR	i
LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PENGESAHAN PEMBIMBING	iii
LEMBAR PENETAPAN PENGUJI	iv
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
ABSTRAK	xiii
ABSTRACT	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Pertanyaan penelitian	4
1.4 Batasan masalah.....	5
1.5 Tujuan Penelitian	5
1.6 Manfaat Penelitian	6
BAB II TINJAUAN PUSTAKA	7
2.1 Penelitian terdahulu.....	7
2.2 Dasar Teori.....	13
2.2.1 <i>Coronavirus Disease 2019 (COVID-19)</i>	13
2.2.2 Vaksinasi COVID-19.....	13
2.2.3 Vaksin AstraZeneca.....	14
2.2.4 Polaritas Sentimen	14
2.2.5 <i>Natural Language Processing (NLP)</i>	15
2.2.6 Analisis Sentimen	16
2.2.7 Twitter	16
2.2.8 <i>Scraping</i>	17

2.2.9	<i>Text Preprocessing</i>	18
2.2.10	Kamus <i>Lexicon</i> Inset.....	19
2.2.11	<i>Term Frequency Inverse Document Frequency</i> (TF-IDF).....	19
2.2.12	Klasifikasi.....	20
2.2.13	Evaluasi	24
2.2.14	<i>K-Fold Cross Validation</i>	26
2.2.15	Python	26
BAB III METODE PENELITIAN		27
3.1	Objek dan Subjek Penelitian	27
3.2	Alat dan Bahan	27
3.3	Diagram Alir Penelitian	28
3.4	Uraian Tahapan Penelitian	30
3.4.4	<i>Preprocessing Data</i>	31
3.4.5	<i>Labeling Data</i>	33
3.4.6	Ekstraksi Fitur	35
3.4.7	Membangun Model Klasifikasi.....	38
3.4.8	Evaluasi dan Validasi Model Klasifikasi	42
BAB IV HASIL DAN PEMBAHASAN		43
4.1	Hasil Pengumpulan Data	43
4.2	Hasil <i>Preprocessing Data</i>	44
4.3	Pelabelan Data	52
4.4	Ekstraksi Fitur	54
4.5	Implementasi Metode Klasifikasi	54
4.6	Evaluasi Model Klasifikasi	56
4.7	Hasil dan Pembahasan.....	61
BAB V KESIMPULAN DAN SARAN		64
5.1	Kesimpulan	64
5.2	Saran.....	64
DAFTAR PUSTAKA		66
LAMPIRAN		73

DAFTAR TABEL

Tabel 2.1 Penelitian terdahulu terkait analisis sentimen	10
Tabel 2.2 <i>Confusion Matrix</i> dua kelas	24
Tabel 3.1 Spesifikasi <i>Hardware</i>	27
Tabel 3.2 Software yang Digunakan	27
Tabel 3.3 Daftar Kata dalam Kamus Normalisasi	32
Tabel 3. 4 Contoh <i>Stopword</i> Bahasa Indonesia	33
Tabel 3.5 Contoh Daftar Kata Negatif pada InSet Lexicon	34
Tabel 3.6 Contoh Daftar Kata Positif pada InSet Lexicon	34
Tabel 3.7 Kemunculan <i>term</i> pada suatu dokumen	35
Tabel 3.8 Tabel Nilai DF	36
Tabel 3.9 Tabel Skor IDF	37
Tabel 3.10 Skor TF-IDF	37
Tabel 3.11 Contoh Dokumen Untuk Perhitungan Naïve Bayes	39
Tabel 3.12 Contoh Perhitungan Probabilitas Kelas	39
Tabel 3.13 Contoh Perhitungan Probabilitas Kata	39
Tabel 4.1 Contoh Hasil <i>Scraping</i> Data Twitter	44
Tabel 4.2 Contoh Data Setelah Melewati Proses <i>Cleaning</i>	45
Tabel 4.3 Contoh Hasil penerapan <i>Case folding</i> pada dataset	46
Tabel 4.4 Contoh Hasil penerapan <i>Tokenization</i> pada dataset	47
Tabel 4.5 Contoh Hasil penerapan <i>Normalization</i> pada dataset	48
Tabel 4.6 Contoh penerapan <i>Stopword Removal</i> pada dataset	50
Tabel 4.7 Hasil Daftar Kata yang Melewati Tahap <i>Stemming</i>	51
Tabel 4.8 Contoh Hasil penerapan <i>Stemming</i> pada dataset	52
Tabel 4.9 Hasil dataset yang sudah diberi <i>polarity score</i>	52
Tabel 4.10 Hasil akurasi Naïve Bayes dengan <i>10-Cross fold validation</i>	57
Tabel 4.11 Hasil akurasi SVM dengan <i>10-Cross fold validation</i>	60
Tabel 4.12 Perbandingan <i>Accuracy, Precision, Recall dan F1-Score</i>	62

DAFTAR GAMBAR

Gambar 2. 1 <i>Hyperplane</i> yang memisahkan dua kelas positif dan negatif.....	23
Gambar 3.1 Diagram Alir penelitian.....	28
Gambar 4.1 <i>Install library</i> snsrape pada <i>command prompt</i>	43
Gambar 4.2 <i>Scraping Data</i> dari Twitter Menggunakan Snsrape	43
Gambar 4.3 <i>Source code</i> tahap <i>Cleaning</i>	45
Gambar 4.4 <i>Source code</i> tahap <i>Case Folding</i>	46
Gambar 4.5 <i>Source code</i> tahap <i>Tokenization</i>	46
Gambar 4.6 <i>Source code</i> tahap <i>Normalization</i>	47
Gambar 4.7 <i>List Stopword</i> bahasa Indonesia	49
Gambar 4.8 <i>Source code</i> penambahan kata pada <i>list stopwords</i>	49
Gambar 4.9 <i>Source code</i> tahap <i>Stopword Removal</i>	50
Gambar 4.10 <i>Source code</i> tahap <i>Stemming</i>	51
Gambar 4.11 Perbandingan data kelas negatif dan positif	53
Gambar 4.12 <i>Source code</i> deklarasi untuk tahap TF-IDF.....	54
Gambar 4.13 <i>Term unik</i>	54
Gambar 4.14 <i>Source code</i> Implementasi Klasifikasi Naïve Bayes	55
Gambar 4.15 Implementasi Klasifikasi <i>Support Vector Machine</i>	56
Gambar 4.16 Visualisasi <i>Confusion Matrix</i> Naïve Bayes.....	56
Gambar 4.17 Visualisasi <i>Confusion Matrix</i> SVM	59
Gambar 4.18 <i>Wordcloud</i> pada Kelas Label Positif	61
Gambar 4.19 <i>Wordcloud</i> pada Kelas Label Negatif	62
Gambar 4.20 Perbandingan Performa Model Klasifikasi	63

DAFTAR LAMPIRAN

Lampiran 1 Dataset	73
Lampiran 2 <i>Source code Preprocessing</i>	81
Lampiran 3 <i>Source code</i> Klasifikasi Naïve Bayes	86
Lampiran 4 <i>Source code</i> Klasifikasi <i>Support Vector Machine</i>	87