



Perbandingan Akurasi dalam Analisis Sentimen berbasis RoBERTa dan Naive Bayes pada Topik Vaksin di Masa Pandemi

LAPORAN SKRIPSI

**Muhammad Kevin Octorico
41518010184**

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



Perbandingan Akurasi dalam Analisis Sentimen berbasis RoBERTa dan Naive Bayes pada Topik Vaksin di Masa Pandemi

LAPORAN SKRIPSI

**Muhammad Kevin Octorico
41518010184**

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

MERCU BUANA

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

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Muhammad Kevin Octorico
NIM : 41518010184
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Akurasi dalam Analisis Sentimen berbasis RoBERTa dan Naïve Bayes pada Topik Vaksin di Masa Pandemi

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.

UNIVERSITAS
MERCU BUANA

Jakarta, 2 Desember 2023



Muhammad Kevin Octorico

HALAMAN PENGESAHAN

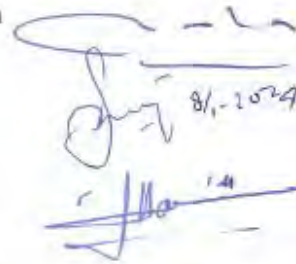
Laporan skripsi ini diajukan oleh

Nama : Muhammad Kevin Octorio
NIM : 41518010184
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Akurasi dalam Analisis Sentimen berbasis RoBERTa dan Naive Bayes pada Topik Vaksin di Masa Pandemi

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata I pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Disahkan Oleh

Pembimbing : Drs. Achmad Kodar, MT., M.Kom
NIDN : 0323085801
Ketua Penguji : Dr. Afyati, S.Si, MT
NIDN : 0316106908
Penguji 1 : Lukman Hakim, ST, M.Kom
NIDN : 0327107704
Penguji 2 : Dr. Hadi Santoso, S.Kom., M.Kom
NIDN : 0225067701



8/1-2024

UNIVERSITAS
MERCU BUANA

Jakarta, 23 Desember 2023

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. Dr. Bambang Jokonowo, S.Si, MTI. selaku Dekan Fakultas Ilmu Komputer
2. Afiyati, S.Si, MT selaku Wakil Dekan Fakultas Ilmu Komputer
3. Dr. Hadi Santoso, S.Kom., M.Kom selaku Ketua Program Studi Teknik Informatika
4. Drs. Achmad Kodar, MT., M.Kom. selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini;
5. Lukman Hakim, ST, M.Kom selaku Dosen Penguji Tugas Akhir atas koreksi dan arahan serta masukannya.
6. Dr. Hadi Santoso, S.Kom., M.Kom selaku Dosen Penguji Tugas Akhir atas koreksi dan arahan serta masukannya.
7. Kedua Orang Tua saya yang selalu memotivasi dan menyemangati saya dalam mengerjakan 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, 2 Desember 2023



Muhammad Kevin Octorico

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

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

Nama	: Muhammad Kevin Octorico
NIM	: 41618010184
Program Studi	: Teknik Informatika
Judul Laporan Skripsi	: Perbandingan Akurasi dalam Analisis Sentimen berbasis RoBERTa dan Naïve Bayes pada Topik Vaksin di Masa Pandemi

Demikian 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, 23 Desember 2023

Yang menyatakan,




METERAI
TEMPEL
5000
BAQALX030792829

Muhammad Kevin Octorico

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 GAMBAR	x
DAFTAR TABEL	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan dan Manfaat	2
1.3.1 Tujuan.....	2
1.3.2 Manfaat	3
1.4 Batasan Masalah.....	3
BAB II TINJAUAN PUSTAKA.....	4
2.1 Penelitian Terdahulu.....	4
2.2 Teori Pendukung	10
2.2.1 Analisis Sentimen	10
2.2.2 Transformers	10
2.2.3 Python	11
2.2.4 RoBERTa.....	11
2.2.5 Naive Bayes	12
2.2.6 AFINN	12
2.2.7 Confusion Matrix.....	13
BAB III METODE PENELITIAN	15
3.1 Jenis Penelitian.....	15
3.2 Tahap Penelitian	16
3.2.1 Pengumpulan Data	16
3.2.2 <i>Pre-Processing</i>	17
3.2.3 Pelatihan Model	19
3.2.4 Evaluasi dan Komparasi Model	19
BAB IV HASIL DAN PEMBAHASAN.....	20
4.1 Dataset.....	20
4.2 Pre-Processing Dataset.....	20
4.3 Pelatihan Model	23
4.3.1 Model RoBERTa	25
4.3.2 Model Naive Bayes	29
4.4 Evaluasi dan Komparasi Model	30
4.4.1 Model RoBERTa	30

4.4.2	Model Naive Bayes	32
4.4.3	Komparasi Model pada Dataset Testing.....	33
4.4.4	Pembahasan Hasil	40
BAB V KESIMPULAN DAN SARAN		43
5.1	Kesimpulan	43
5.2	Saran.....	43
DAFTAR PUSTAKA.....		45
LAMPIRAN.....		48
Lampiran Bimbingan.....		48
Lampiran Luaran Tugas Akhir.....		49
Lampiran Bukti Submit Jurnal		50
Lampiran Jurnal		51
Lampiran Plagiarism		58
Lampiran Surat Pernyataan HKI (jika belum published HKI).....		59
Lampiran Curriulum Vitae		60
Lampiran Sertikat BNSP (SKL)		61
Lampiran Persetujuan Dosen Pembimbing.....		63



DAFTAR GAMBAR

Gambar 1 Tahap penelitian	16
Gambar 2 Alur Pre-Processing.....	18
Gambar 3 Menampilkan bentuk awal dataset	21
Gambar 4 Dataframe yang telah diberikan sentimen otomatis	21
Gambar 5 Total label sentimen otomatis pada dataframe	22
Gambar 6 Test percobaan penerjemahan emoji dan abreviasi untuk diaplikasikan ke dataframe	22
Gambar 7 Hasil dataset yang telah di tokenisasi dan siap di latih pada model naive baiyes.....	23
Gambar 8 Dataset testing RoBERTa.....	24
Gambar 9 Dataset Training AFINN 1	25
Gambar 10 Dataset Training AFINN 2	25
Gambar 11 Pemisahan data untuk RoBERTa.....	26
Gambar 12 Tokenisasi dan pemisahan data pelatihan dan data testing RoBERTa	27
Gambar 13 Fungsi Training pada RoBERTa.....	28
Gambar 14 Penyimpanan dan klasifikasi evaluasi model RoBERTa.....	29
Gambar 15 Proses dataset untuk pelatihan model Naive Bayes	29
Gambar 16 Pelatihan model Naive Bayes.....	30
Gambar 17 Hasil klasifikasi prediksi label RoBERTa 9000 data	30
Gambar 18 Hasil klasifikasi prediksi label AFINN 9000 data pada RoBERTa....	31
Gambar 19 Hasil klasifikasi prediksi label AFINN 31000 data pada RoBERTa... 31	
Gambar 20 Hasil klasifikasi prediksi label RoBERTa 9000 data pada naive baiyes	32
Gambar 21 Hasil klasifikasi prediksi label AFINN 9000 data pada naive baiyes	32
Gambar 22 Hasil klasifikasi prediksi label AFINN 31000 data pada naive baiyes	33
Gambar 23 Perbandingan performa RoBERTa 9000	34
Gambar 24 Confusion Matrix RoBERTa 9000	35
Gambar 25 Perbandingan performa AFINN 9000	36
Gambar 26 Confusion Matrix AFINN 9000	37
Gambar 27 Perbandingan performa AFINN 31000	38
Gambar 28 Confusion Matrix AFINN 31000	39

DAFTAR TABEL

Tabel 1 Penelitian Terdahulu.....	4
Tabel 2 Confusion Matrix	13
Tabel 3 Perbandingan Performa Model.....	40



DAFTAR LAMPIRAN

<u>Lampiran Bimbingan</u>	48
<u>Lampiran Luaran Tugas Akhir</u>	49
<u>Lampiran Bukti Submit Jurnal</u>	50
<u>Lampiran Jurnal</u>	51
<u>Lampiran Plagiarism</u>	58
<u>Lampiran Surat Pernyataan HKI (jika belum published HKI)</u>	59
<u>Lampiran Curriulum Vitae</u>	60
<u>Lampiran Sertikat BNSP (SKL)</u>	61
<u>Lampiran Persetujuan Dosen Pembimbing</u>	63



UNIVERSITAS
MERCU BUANA