

**ANALISIS SENTIMEN PADA MEDIA SOSIAL X
MENGUNAKAN METODE *NAIVE BAYES*
(STUDI KASUS : *JERSEY* TIM NASIONAL INDONESIA PRODUKSI
APPAREL ERSPO)**

TUGAS AKHIR

Diajukan sebagai salah satu syarat memperoleh gelar Sarjana
Program Studi S-1 Sistem Informasi



Disusun oleh:

SAIFUL AMAR
202103018

**PROGRAM STUDI S-1 SISTEM INFORMASI
FAKULTAS TEKNIK & TEKNOLOGI INFORMASI
UNIVERSITAS JENDERAL ACHMAD YANI YOGYAKARTA
2024**

HALAMAN PENGESAHAN

TUGAS AKHIR

ANALISIS SENTIMEN PADA MEDIA SOSIAL X
MENGUNAKAN METODE *NAIVE BAYES*
(STUDI KASUS : *JERSEY TIM NASIONAL INDONESIA PRODUKSI APPAREL ERSPO*)

Diajukan oleh:

SAIFUL AMAR
202103018

Telah dipertahankan di depan dewan penguji dan dinyatakan sah
sebagai salah satu syarat untuk memperoleh gelar Sarjana
di Fakultas Teknik & Teknologi Informasi
Universitas Jenderal Achmad Yani Yogyakarta

Tanggal: 16 Juli 2024

Mengesahkan:

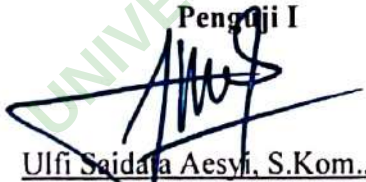
Pembimbing I


Kharisma, S.T., M.Cs.
NIDN: 0502108201

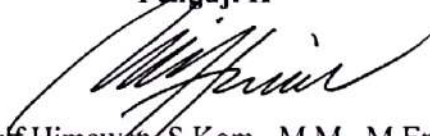
Pembimbing II


Ahmad Hanafi, S.T., M.Eng.
NIDN: 0528088301

Penguji I


Ulfi Saidara Aesvi, S.Kom., M.Cs.
NIDN: 0515129002

Penguji II


Arif Himawan, S.Kom., M.M., M.Eng.
NIDN: 0517127402

Ketua Program Studi S-1 Sistem Informasi
Fakultas Teknik & Teknologi Informasi
Universitas Jenderal Achmad Yani Yogyakarta



PERNYATAAN

Saya yang bertanda tangan di bawah ini, adalah mahasiswa Fakultas Teknik dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta,

Nama : Saiful Amar
NPM : 202103018
Program Studi : S-I Sistem Informasi
Judul Tugas Akhir : Analisis Sentimen Pada Media Sosial X Menggunakan Metode *Naïve Bayes* (Studi kasus : *Jersey Tim Nasional Indonesia Produksi Apparel Erspo*)

Menyatakan bahwa hasil penelitian dengan judul tersebut di atas adalah asli karya saya sendiri dan bukan hasil plagiarisme. Semua referensi dan sumber terkait yang dikutip dalam karya ilmiah ini telah ditulis sesuai kaidah penulisan ilmiah yang berlaku. Dengan ini, saya menyatakan untuk menyerahkan hak cipta penelitian kepada Universitas Jenderal Achmad Yani Yogyakarta guna kepentingan ilmiah.

Demikian surat pernyataan ini dibuat dengan sebenar-benarnya tanpa ada paksaan dari pihak mana pun. Apabila terdapat kekeliruan atau ditemukan adanya pelanggaran akademik di kemudian hari, maka saya bersedia menerima konsekuensi yang berlaku sesuai ketentuan akademik.

Yogyakarta, 16 Juli 2024



Saiful Amar

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat-Nya sehingga penulis dapat menyelesaikan laporan tugas akhir yang berjudul: “Analisis Sentimen Pada Media Sosial *X* Menggunakan Metode Naïve Bayes (Studi Kasus : Jersey Tim Nasional Indonesia Produksi *Apparel Erspo*)”. Penyusunan laporan ini merupakan salah satu persyaratan untuk menyelesaikan studi di Program Studi S-1 Sistem Informasi Fakultas Teknik & Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta. Laporan ini dapat diselesaikan atas bimbingan, arahan, dan bantuan dari berbagai pihak. Pada kesempatan ini penulis dengan rendah hati mengucapkan terima kasih dengan setulus-tulusnya kepada:

1. Bapak Aris Wahyu Murdiyanto, S.Kom., M.Cs. selaku Dekan Fakultas Teknik dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta;
2. Bapak Ahmad Hanafi, S.T., M.Eng. selaku Ketua Program Studi S-1 Sistem Informasi Fakultas Teknik dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta;
3. Bapak Kharisma, S.T., M.Cs. dan Bapak Ahmad Hanafi, S.T., M.Eng. selaku Dosen Pembimbing Tugas Akhir;
4. Para dosen yang telah memberikan banyak bekal ilmu pengetahuan kepada penulis selama menjadi mahasiswa di Fakultas Teknik dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta;
5. Ayah, ibu, dan adik saya, yang telah memberikan dukungan semangat serta doa restu kepada saya, sehingga dapat menyelesaikan studi saya;
6. Rekan-rekan mahasiswa Prodi S-1 Sistem Informasi di Universitas Jenderal Achmad Yani Yogyakarta yang sudah memberi dukungan dan kerja sama selama pembuatan tugas akhir.

Penulis menyadari bahwa laporan tugas akhir ini masih jauh dari kata sempurna. Maka dari itu dengan segala kerendahan hati penulis sangat menghargai adanya kritik dan saran yang membangun dari semua pihak yang bersedia meluangkan waktu untuk membaca laporan tugas akhir ini.

Yogyakarta, 16 Juli 2024



Saiful Amar

PERPUSTAKAAN
UNIVERSITAS JENDERAL ACHMAD YANI
YOGYAKARTA

DAFTAR ISI

Halaman Pengesahan.....	ii
Pernyataan	iii
Kata Pengantar	iv
Daftar Isi	vi
Daftar Tabel.....	ix
Daftar Gambar	x
Daftar Lampiran	xi
Daftar Singkatan	xii
Intisari.....	xiii
<i>Abstract</i>.....	xiv
Bab 1 Pendahuluan	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Pertanyaan Penelitian.....	3
1.5 Tujuan Penelitian	3
1.6 Manfaat Hasil Penelitian.....	3
Bab 2 Tinjauan Pustaka dan Landasan Teori.....	4
2.1 Tinjauan Pustaka	4
2.2 Landasan Teori.....	9
2.2.1 Python.....	9
2.2.2 Media Sosial “X”	9
2.2.3 <i>Text Mining</i>	10
2.2.4 Analisis Sentimen.....	11
2.2.5 Naïve Bayes Classifier	11
2.2.6 Evaluasi Model.....	14
Bab 3 Metode Penelitian.....	16
3.1 Bahan dan Alat Penelitian.....	16
3.2 Jalan Penelitian.....	17

Bab 4 Hasil Penelitian	23
4.1 Ringkasan Hasil Penelitian	23
4.2 Hasil Penelitian	23
4.2.1 <i>Crawling Data</i>	23
4.2.1.1 Registrasi	24
4.2.1.2 <i>Copy Auth Token</i>	24
4.2.1.3 <i>Install Node.js</i>	24
4.2.1.4 <i>Install Tweet Harvest</i>	25
4.2.1.5 Mengambil Data <i>Tweet</i>	25
4.2.2 <i>Preprocessing Data</i>	25
4.2.2.1 <i>Cleaning Text</i>	26
4.2.2.2 <i>Case folding</i>	27
4.2.2.3 Tokenisasi	27
4.2.2.4 Normalisasi	28
4.2.2.5 <i>Stopword Removal</i>	28
4.2.2.6 <i>Stemming</i>	29
4.2.3 Labeling Data	30
4.2.4 Visualisasi	32
4.2.4.1 Wordcloud sentimen positif	32
4.2.4.2 <i>Wordcloud</i> sentimen negatif	33
4.2.5 TF-IDF	33
4.2.5.1 <i>Tokenize</i>	34
4.2.5.2 Pembobotan <i>Term Frequency</i> (TF)	34
4.2.5.3 Menghitung Document Frequency (DF)	35
4.2.5.4 Menghitung TF-IDF	36
4.2.6 Klasifikasi Naïve Bayes	37
4.2.6.1 <i>Data Training</i>	37
4.2.6.2 <i>Data Testing</i>	38
4.2.7 Evaluasi	39
4.2.7.1 Evaluasi Klasifikasi	39
4.2.7.2 Evaluasi Analisis	41

4.2.7.3 Hasil Sentimen.....	43
Bab 5 Kesimpulan dan Saran	44
5.1 Kesimpulan	44
5.2 Saran.....	44
Daftar Pustaka.....	46
Lampiran	49

PERPUSTAKAAN
UNIVERSITAS JENDERAL ACHMAD YANI
YOGYAKARTA

DAFTAR TABEL

Tabel 2.1 Tinjauan Pustaka	5
Tabel 2.2 <i>Confusion matrix</i>	14
Tabel 4.1 Hasil <i>Stemming</i>	30
Tabel 4.2 <i>Labeling</i> Data	31
Tabel 4.3 <i>Confusion Matrix</i> testing data	40
Tabel 4.4 Hasil <i>Cross Validation</i>	40
Tabel 4.5 <i>Classification report</i> data testing	41
Tabel 4.6 Data dengan sentimen positif	42
Tabel 4.7 Data dengan sentimen negatif	42

DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian	17
Gambar 3.2 Alur <i>crawl data</i>	18
Gambar 3.3 Alur <i>preprocessing</i>	18
Gambar 4.1 <i>Auth Token</i>	24
Gambar 4.2 <i>Instance Chromium Browser</i>	25
Gambar 4.3 Modul <i>Preprocessing</i>	26
Gambar 4.4 <i>Cleaning text</i>	26
Gambar 4.5 <i>Case folding</i>	27
Gambar 4.6 Tokenisasi	27
Gambar 4.7 Normalisasi	28
Gambar 4.8 <i>Stopword removal</i>	29
Gambar 4.9 <i>Stemming</i>	30
Gambar 4.10 <i>Wordcloud</i> sentimen positif	32
Gambar 4.11 <i>Wordcloud</i> sentimen negatif	33
Gambar 4.12 <i>Tokenize TF-IDF</i>	34
Gambar 4.13 <i>Term Frequency</i>	34
Gambar 4.14 <i>Document Frequency</i>	35
Gambar 4.15 Perhitungan <i>TF-IDF</i>	36
Gambar 4.16 Penentuan Sumbu X dan Y	37
Gambar 4.17 Membagi dataset	37
Gambar 4.18 Klasifikasi <i>training data</i>	38
Gambar 4.19 Klasifikasi <i>testing data</i>	38
Gambar 4.20 Hasil prediksi <i>testing data</i>	39
Gambar 4.21 Diagram sentimen	41

DAFTAR LAMPIRAN

Lampiran 1 Jadwal Penelitian	50
Lampiran 2 Lembar Bimbingan Dosen	51
Lampiran 3 Hasil Cek Plagiarisme.....	52

PERPUSTAKAAN
UNIVERSITAS JENDERAL ACHMAD YANI
YOGYAKARTA

DAFTAR SINGKATAN

CSV	<i>Comma-Separated Values</i>
DF	<i>Document Frequency</i>
FN	<i>False Negative</i>
FP	<i>False Positive</i>
HTML	<i>Hypertext Markup Language</i>
IDF	<i>Inverse Document Frequency</i>
NLTK	<i>Natural Language Toolkit</i>
PDF	<i>Portable Document Format</i>
PSSI	<i>Persatuan Sepak Bola Seluruh Indonesia</i>
TF	<i>Term Frequency</i>
TN	<i>True Negative</i>
TP	<i>True Positive</i>
TF-IDF	<i>Term Frequency-Inverse Document Frequency</i>