



**ANALISIS SENTIMEN PADA SOSIAL MEDIA *INSTAGRAM*  
MENGUNAKAN ALGORITMA *NAÏVE BAYES*  
(STUDI KASUS : TIMNAS FUTSAL INDONESIA)**

**SKRIPSI**

**DOLI ANANDA EFRAIM**

**1910511104**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA  
2023**



**ANALISIS SENTIMEN PADA SOSIAL MEDIA *INSTAGRAM*  
MENGUNAKAN ALGORITMA *NAÏVE BAYES*  
(STUDI KASUS : TIMNAS FUTSAL INDONESIA)**

**SKRIPSI**

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar  
Sarjana Komputer**

**DOLI ANANDA EFRAIM  
1910511104**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA  
2023**

# LEMBAR PERNYATAAN ORISINALITAS

## LEMBAR PENYATAAN ORISINALITAS

Tugas Akhir ini adalah karya hasil sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Doli Ananda Efrain

NIM : 1910511104

Tanggal : 14 Juli 2023

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia untuk dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 14 Juli 2023

Yang Menyatakan,



(Doli Ananda Efrain)

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

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini :

Nama : Doli Ananda Efraim

NIM : 1910511104

Fakultas : Ilmu Komputer

Program Studi : S1 Informatika

Demi Pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non Ekklusif (*Non-Exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

**ANALISIS SENTIMEN PADA SOSIAL MEDIA *INSTAGRAM*  
MENGUNAKAN ALGORITMA *NAÏVE BAYES*  
(STUDI KASUS : TIMNAS FUTSAL INDONESIA)**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan kata (Basis data), merawat dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada Tanggal : 14 Juli 2023

Yang Menyatakan,



(Doli Ananda Efraim)

# LEMBAR PENGESAHAN

## LEMBAR PENGESAHAN


Dengan ini dinyatakan bahwa Tugas Akhir berikut:

Nama : Doli Ananda Efraim  
NIM : 1910511104  
Program Studi : SI Informatika  
Judul : Analisis Sentimen Pada Sosial Media *Instagram*  
Menggunakan Algoritma *Naive Bayes* (Studi Kasus :  
Timas Futsal Indonesia)


Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian dari persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi SI Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jakarta.

  
**(Dr. Didit Widryanto, S.Kom, M.Si.)**  
Penguji I

  
**(Bayu Hananto, S.Kom., M.Kom.)**  
Penguji II

  
**(Dr. Ermatita, M.Kom.)**  
Dosen Pembimbing

  
**(Dr. Ermatita, M.Kom.)**  
Dekan Fakultas Ilmu Komputer

  
**(Dr. Widya Cholil, M.I.T)**  
Ketua Program Studi

Ditetapkan di : Jakarta  
Tanggal Persetujuan : 11 Juli 2023



## ABSTRAK

Nama : Doli Ananda Efraim

Program Studi : Informatika

Judul : Analisis Sentimen Pada Sosial Media *Instagram* Menggunakan Algoritma *Naive Bayes* (Studi Kasus Timnas Futsal Indonesia)

Sosial media telah menjadi identitas baru bagi masyarakat Indonesia saat ini dan juga merupakan gaya hidup yang sulit dilupakan. Seiring dengan perkembangan zaman, sosial media telah tersebar luas di internet, dan setiap pengguna memiliki kebebasan untuk menggunakan sosial media sesuai keinginan mereka. Bahkan, banyak pengguna yang memiliki lebih dari satu akun sosial media. Salah satu platform sosial media yang populer adalah *Instagram*, yang dapat diakses melalui smartphone Android dan iOS, serta melalui situs web. Tahapan awal dalam penelitian ini adalah pengumpulan data komentar yang diambil dari sosial media *instagram*. Kemudian data komentar tersebut diberi label positif dan negatif yang akan diberikan oleh 2 anotator. Setelah itu dilakukan praproses seperti pembersihan data, *case folding*, *normalization*, *tokenization*, *stopword removal*, dan *stemming* kemudian pembobotan kata dengan *Term Frequency – Inverse Document Frequency*. Hasil dari pelabelan data berjumlah 262 komentar positif dan 142 komentar negatif. Lalu data dibagi menjadi 80% data latih dan 20% data uji. Algoritma yang digunakan untuk melakukan klasifikasi adalah *naive bayes*. Hasil klasifikasi yang didapatkan adalah akurasi sebesar 71%, presisi sebesar 84%, *recall* 69%.

**Kata Kunci:** Analisis Sentimen, Postingan Timnas Futsal Indonesia, *Naive Bayes*, *Instagram*

## ABSTRACT

Name : Doli Ananda Efraim

Study Program : Informatics

Title : Sentiment Analysis on Social Media Instagram Using the Naive Bayes Algorithm (Case Study of the Indonesian Futsal National Team)

Social media has become a new identity for Indonesian people today and is also a lifestyle that is difficult to forget. Along with the times, social media has spread widely on the internet, and every user has the freedom to use social media as they wish. In fact, many users have more than one social media account. One of the popular social media platforms is Instagram, which can be accessed via Android and iOS smartphones, as well as through the website. The initial stage in this research was collecting comment data taken from Instagram social media. Then the comment data is labeled positive and negative which will be given by 2 annotators. After that, pre-processing is carried out such as data cleaning, case folding, normalization, tokenization, stopword removal, and stemming then weighting words with Term Frequency - Inverse Document Frequency. The results of data labeling amounted to 262 positive comments and 142 negative comments. Then the data is divided into 80% training data and 20% test data. The algorithm used to perform the classification is naive Bayes. The classification results obtained are accuracy of 71%, precision of 84%, recall of 69%.

**Keywords:** Sentiment Analysis, Indonesian Futsal National Team Posts, Naïve Bayes, Instagram

## KATA PENGANTAR

Assalamu'alaikum Wr. Wb.

Puji syukur senantiasa penulis panjatkan atas kehadiran Tuhan Yang Maha Esa yang selalu melimpahkan rahmat dan karunia-NYA, sehingga Penulis dapat menyelesaikan penyusunan Skripsi ini yang berjudul “ANALISIS SENTIMEN PADA SOSIAL MEDIA *INSTAGRAM* MENGGUNAKAN ALGORITMA *NAÏVE BAYES* (STUDI KASUS : TIMNAS FUTSAL INDONESIA)”. Skripsi ini disusun oleh penulis sebagai syarat dalam rangka menyelesaikan studi untuk menempuh gelar Sarjana Komputer di Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.

Rasa terima kasih ingin penulis sampaikan dalam kata pengantar ini. Penulis menyadari bahwa terselesainya skripsi ini tidak terlepas dari bantuan dan bimbingan banyak pihak. Oleh sebab penulis ingin mengucapkan banyak terima kasih kepada :

1. Ibu Dr. Ermatita, M.Kom. selaku Dekan Fakultas Ilmu Komputer UPN Veteran Jakarta dan selaku Dosen Pembimbing Skripsi penulis yang telah banyak memberikan arahan, saran, dan masukan yang sangat berguna dalam penyelesaian skripsi ini.
2. Ibu Dr. Widya Cholil, M.I.T. selaku Kepala Program Studi Informatika.
3. Ibu Helena Nurramdhani Irmanda, S.Pd, M.Kom selaku Dosen Pembimbing Akademik di Fakultas Ilmu Komputer UPN Veteran Jakarta.
4. Bapak dan Ibu Dosen Fakultas Ilmu Komputer UPN Veteran Jakarta yang telah memberikan banyak pengetahuan dibidang Hukum dan memberikan bimbingan serta bantuan selama penulis menempuh perkuliahan.



5. Bapak dan Ibu Staf Akademik Fakultas Ilmu Komputer UPN Veteran Jakarta yang telah membantu penulis selama menempuh perkuliahan
6. Orang tua dan keluarga dari Penulis yang selalu memberikan doa dan dukungan baik moral dan materil kepada penulis hingga saat ini.
7. Teman-teman seperjuangan keluarga besar UDAH BUKBER HELLFIRE, Ikhlasul tiktokers Global Edith, Iki leuwiliang, Dipa aneh, Ryan onta, Nopal pucekboy, Ardi jamet tebet dan ada teman - teman para pejuang rupiah dll. yang sebenarnya tidak membantu penulis tetapi menghibur dan selalu memberikan semangat kepada penulis selama proses penyelesaian skripsi.
8. Semua pihak yang tidak dapat disebutkan satu per satu yang telah membantu penulis dalam menyelesaikan skripsi.
  - a. Dalam penulisan skripsi ini, penulis menyadari bahwa masih jauh dari kata sempurna walaupun telah disusun dengan upaya yang maksimal dikarenakan keterbatasan atas kemampuan dan pengetahuan yang penulis miliki. Oleh karena itu, penulis berharap atas saran dan kritik yang bersifat membangun dari berbagai pihak.
2. Besar harapan dari penulis semoga dengan adanya skripsi ini dapat memberikan manfaat dan menambah referensi bagi penulis dan para pembaca secara khusus untuk Mahasiswa Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta (UPNVJ). Akhir kata, semoga tujuan dari pembuatan skripsi ini dapat tercapai sesuai dengan yang diharapkan.

Jakarta, Juni 2023

Doli Ananda Efraim

## DAFTAR ISI

COVER .....	i
JUDUL .....	ii
LEMBAR PERNYATAAN ORISINALITAS .....	iii
PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS .....	iv
LEMBAR PENGESAHAN .....	v
ABSTRAK .....	vi
ABSTRACT .....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI .....	x
DAFTAR TABEL.....	1
DAFTAR GAMBAR .....	2
BAB 1 PENDAHULUAN.....	3
1.1. Latar Belakang Masalah.....	3
1.2. Perumusan Masalah .....	4
1.3. Ruang Lingkup Penelitian.....	4
1.4. Tujuan Penelitian .....	4
1.5. Luaran Penelitian .....	5
1.6. Sistematika Penulisan .....	5
BAB 2 TINJAUAN PUSTAKA .....	7
2.1. <i>Instagram</i> .....	7
2.2. <i>Text Mining</i> .....	7
2.3. Analisis Sentimen .....	8
2.4. Metode Klasifikasi.....	9
2.4.1. Algoritma <i>Naive Bayes</i> .....	10
2.5. <i>Kappa Statistic</i> .....	11
2.6. <i>Term Frequency – Inverse Document Frequency (TF – IDF)</i> .....	13
2.7. Timnas Futsal Indonesia .....	15
2.8. Review Penelitian yang Relevan .....	15
BAB 3 METODOLOGI PENELITIAN .....	18
3.1. Identifikasi Masalah.....	19
3.2. Studi Literatur.....	19
3.3. Pengumpulan Data.....	19
3.4. Pelabelan Data.....	20

3.5.	Praproses .....	20
3.5.1.	Pembersihan data .....	20
3.5.2.	<i>Case Folding</i> .....	20
3.5.3.	<i>Normalization</i> .....	20
3.5.4.	<i>Tokenisasi</i> .....	21
3.5.5.	<i>Stopword Removal</i> .....	21
3.5.6.	<i>Stemming</i> .....	21
3.6.	Pembobotan TF IDF .....	21
3.7.	Klasifikasi .....	21
3.8.	Evaluasi.....	22
3.9.	Hasil Klasifikasi .....	23
3.10.	Alat Bantu Penelitian .....	23
3.10.1.	<i>Hardware</i> .....	23
3.10.2.	<i>Software</i> .....	24
3.11.	Jadwal Penelitian .....	24
BAB 4 HASIL DAN PEMBAHASAN .....		26
4.1	Data.....	26
4.2	Pelabelan Data.....	26
4.3	Praproses .....	29
4.3.1	Pembersihan data .....	29
4.3.2	<i>Case Folding</i> .....	30
4.3.3	<i>Normalization</i> .....	31
4.3.4	<i>Tokenization</i> .....	34
4.3.5	<i>Stopword Removal</i> .....	35
4.3.6	<i>Stemming</i> .....	37
4.4	Pembobotan Kata dengan TF-IDF.....	39
4.5	Klasifikasi <i>Naive Bayes</i> .....	45
4.5.1	Data Latih.....	45
4.5.2	Data Uji .....	46
4.6	Evaluasi.....	48
BAB 5 PENUTUP .....		51
5.1	Kesimpulan.....	51
5.2	Saran .....	51
DAFTAR PUSTAKA .....		53

## DAFTAR TABEL

Tabel 2.1 Variasi Antar-Pengamat/Penilai .....	12
Tabel 2.2 Interpretasi Kappa .....	13
Tabel 3.1 Confusion Matrix .....	22
Tabel 3.2 Jadwal Penelitian.....	24
Tabel 4.1 Hasil Anotasi.....	26
Tabel 4.2 Hasil Sampel Pembersihan Data .....	29
Tabel 4.3 Hasil Sampel Pembersihan Data .....	31
Tabel 4.4 Sampel Kata-Kata Normalization .....	32
Tabel 4.5 Hasil Sampel Normalization .....	33
Tabel 4.6 Hasil Sampel Tokenization.....	34
Tabel 4.7 Hasil Sampel Stopword Removal .....	36
Tabel 4.8 Hasil Sampel Stemming .....	38
Tabel 4.9 Hasil sampel nilai TF, IDF, dan TF-IDF .....	43
Tabel 4.10 Hasil Data Latih .....	46
Tabel 4.11 Probabilitas Sampel Data Uji.....	47
Tabel 4.12 Confusion Matrix .....	49

## DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian.....	18
Gambar 4.1 Hasil Pelabelan Sentimen.....	28
Gambar 4.2 Proses Pembersihan Data .....	29
Gambar 4.3 Proses Case Folding.....	30
Gambar 4.4 Proses Normalization .....	32
Gambar 4.5 Proses Tokenization.....	34
Gambar 4.6 Proses Stopword Removal .....	36
Gambar 4.7 Proses Stemming .....	37