



**KLASIFIKASI DAN ANALISIS SENTIMEN PADA DATA  
TWITTER MENGGUNAKAN ALGORITMA *NAÏVE BAYES*  
(STUDI KASUS: MATA UANG KRIPTO TKO)**

**SKRIPSI**

**ARYA RANGGA PRAMASTYA  
1710511054**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA  
FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI S1 INFORMATIKA  
2024**



**KLASIFIKASI DAN ANALISIS SENTIMEN PADA DATA  
TWITTER MENGGUNAKAN ALGORITMA *NAÏVE BAYES*  
(STUDI KASUS: MATA UANG KRIPTO TKO)**

**SKRIPSI**

**Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Komputer Pada  
Fakultas Ilmu Komputer**

**ARYA RANGGA PRAMASTYA**

**1710511054**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA**

**FAKULTAS ILMU KOMPUTER**

**PROGRAM STUDI S1 INFORMATIKA**

**2024**

# PERNYATAAN ORISINALITAS

## PERNYATAAN ORISINALITAS

Skripsi adalah karya sendiri dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Arya Rangga Pramastya

NM : 1710511054

Tanggal : 7 Oktober 2024

Bilamana dikemudian hari ditemukan tidak sesuai dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai ketentuan yang berlaku

Jakarta 7 Oktober 2024

Yang Menyatakan



(Arya Rangga Pramastya)

## **PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

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

Nama : Arya Rangga Pramatsya

NIM : 1710511054

Fakultas : Ilmu Komputer

Program Studi : Informatika

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

### **Klasifikasi dan analisis sentimen pada data twitter menggunakan algoritma *Naives Bayes* (Studi kasus mata uang kripto TKO)**

Dengan hak bebas royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan 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 : 31 Juli 2024



(Arya Rangga Pramatsya)

# LEMBAR PENGESAHAN

## LEMBAR PENGESAHAN SKRIPSI

Dengan ini dinyatakan bahwa Tugas Akhir/Skripsi berikut :

Nama : Arya Rangga Pramastya  
NIM : 1710511054  
Program Studi : S1 Informatika  
Judul Tugas Akhir : KLASIFIKASI DAN ANALISIS SENTIMEN PADA DATA TWITTER MENGGUNAKAN ALGORITMA *NAÏVE BAYES* (STUDI KASUS: MATA UANG KRIPTO TKO)

Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian dari persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jakarta

(Dr. Widya Cholil, M.I.T)

Penguji I

(Muhammad Adrezo, S.Kom., M.Sc)

Penguji II

(Rio Wirawan, S.Kom., MMSI.)

Pembimbing



(Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM)

Dekan Fakultas Ilmu Komputer

(Dr. Widya Cholil, M.I.T)

Ketua Program Studi

Disahkan di : Jakarta  
Tanggal : 22 Juli 2024



## KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kehadirat Tuhan Yang Mahakuasa atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan Proposal Tugas Akhir. Proposal tugas akhir ini disusun sebagai syarat Tugas Akhir Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.

Dalam penyelesaian Skripsi ini tidak lepas dari bantuan banyak pihak yang telah memberikan masukan kepada penulis. Untuk itu penulis mengucapkan terima kasih kepada:

1. Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM, selaku dekan Fakultas Ilmu Komputer
2. Dr. Widya Cholil, M.I.T, selaku Ketua Program Studi Sarjana Jurusan S1 Informatika.
3. Rio Wirawan, S.Kom., MMSI, selaku dosen pembimbing .
4. Orang tua yang telah memberikan dukungan baik secara moril maupun materil.
5. Seluruh pihak yang terlibat dalam kelancaran pembuatan makalah karya ilmiah ini dan yang belum disebutkan di atas, penulis ucapkan terima kasih.
6. Xena Gabriella, S.Kom Selaku pacar saya yang memberikan dukungan penuh untuk Skripsi saya.
7. Penulis menyadari bahwa masih banyak kekurangan dari laporan ini, baik dari materi maupun teknik penyajiannya. Oleh karena itu, kritik dan saran yang membangun penulis harapkan.

Jakarta, 31 Juli 2024



Arya Rangga Pramatsya

**KLASIFIKASI DAN ANALISIS SENTIMEN PADA DATA  
TWITTER MENGGUNAKAN ALGORITMA NAIVE BAYES (Studi  
kasus mata uang kripto TKO)**

**Arya Rangga Pramastya**

**ABSTRAK**

Twitter Merupakan sosial media yang saat ini masih digemari masyarakat umum. Twitter merupakan alat komunikasi antar pengguna berbasis internet dan wadah untuk mengekspresikan pendapat pribadi atau opini. Banyak aspek yang bisa di utarakan di media sosial ini. Beberapa topik pun dapat ditemukan dalam twitter termasuk mata uang kripto (cryptocurrency). Saat ini sedang ramai pembahasan tentang mata uang kripto di dalam twitter. Banyak orang mengutarakan opini masing masing melalui sebuah tweet yang ada dalam fitur twitter. Salah satu mata uang kripto yang ramai adalah TKO. TKO merupakan mata uang kripto pertama yang diluncurkan oleh Indonesia. Berdasarkan dengan latar belakang tersebut, maka akan diperlukan penelitian tentang mata uang kripto (cryptocurrency). Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap opini atau pendapat masyarakat tentang mata uang kripto menggunakan algoritma Naives Bayes.

**Kata Kunci:** *cryptocurrency*, Mata uang kripto, Analisis Sentimen, *Naïve Bayes*

***CLASSIFICATION AND SENTIMENT ANALYSIS ON TWITTER  
DATA USING THE NAÏVE BAYES ALGORITHM (CASE STUDY:  
TKO CRYPTO CURRENCY)***

*Arya Rangga Pramatsya*

***ABSTRACT***

*Twitter is a social media that is currently still popular with the general public. Twitter is an internet-based communication tool between users and a forum for expressing personal views or opinions. There are many aspects that can be expressed on this social media. Several topics can be found on Twitter including cryptocurrencies. Currently there is a lot of discussion about cryptocurrencies on Twitter. Many people express their opinions through tweets in the Twitter feature. One of the most popular cryptocurrencies is TKO. TKO is the first cryptocurrency launched by Indonesia. Based on this background, research on cryptocurrencies will be needed. This research aims to conduct sentiment analysis of public opinions about cryptocurrencies using the Naives Bayes algorithm.*

*Keywords: cryptocurrency, Cryptocurrency, Sentiment Analysis, Naïve Bayes*



## DAFTAR ISI

COVER.....	I
PERNYATAAN ORISINALITAS.....	II
PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS .....	III
LEMBAR PENGESAHAN .....	IV
KATA PENGANTAR .....	V
ABSTRAK.....	VI
<i>ABSTRACT</i> .....	VII
DAFTAR ISI.....	VIII
DAFTAR GAMBAR.....	X
DAFTAR TABEL.....	XI
BAB I PENDAHULUAN .....	1
1.1. Latar belakang.....	1
1.2. Rumusan Masalah .....	2
1.3. Batasan Masalah.....	2
1.4. Tujuan Penelitian .....	2
1.5. Manfaat Penelitian .....	2
1.6. Luaran yang Diharapkan .....	2
1.7. Sistematika Penulisan.....	3
BAB II LANDASAN TEORI.....	4
2.1. Twitter.....	4
2.2. Analisis Sentimen.....	4
2.3. Mata Uang Kripto .....	5
2.4. Toko Token(TKO) .....	5
2.5. <i>Text Mining</i> .....	6
2.5.1. Pengumpulan Data.....	6
2.5.2. <i>Text Preprocessing</i> .....	7
2.6. TF-IDF .....	8
2.7. Kappa Statistic .....	9
2.8. Algoritma <i>Naive Bayes</i> .....	9
2.9. Evaluasi Model.....	10
2.10. Penelitian Terdahulu.....	12

BAB III METODOLOGI PENELITIAN .....	13
3.1. Tahapan Penelitian .....	13
3.1.1. Identifikasi Masalah.....	13
3.1.2. Pengambilan Data .....	14
3.1.3. <i>Preprocessing</i> Data.....	14
3.1.4. TF-IDF .....	15
3.1.5. Klasifikasi Algoritma <i>Naive Bayes</i> .....	15
3.1.6. Evaluasi dan Hasil Model .....	15
3.2. Alat Bantuan Penelitian.....	16
3.3. Jadwal Penelitian.....	16
BAB IV HASIL DAN PEMBAHASAN .....	18
4.1. Data .....	18
4.2. Pelabelan Data.....	19
4.3. Pra Proses Data .....	20
4.3.1. <i>Transform Cases</i> .....	21
4.3.2. <i>Cleaning</i> .....	22
4.3.3. <i>Tokenization</i> .....	22
5.1.1. <i>Stopword Removal</i> .....	23
5.1.2. Normalisasi .....	24
4.2. Pembobotan TF IDF.....	25
4.3. Pembobotan TF-IDF .....	25
4.4. Proses Klasifikasi .....	28
4.3.1. Data Latih .....	29
4.4.2. Data Uji.....	31
4.5. Evaluasi Dan Hasil Analisis.....	34
BAB V KESIMPULAN DAN SARAN .....	38
5.1. Kesimpulan .....	38
5.2. Saran.....	38
DAFTAR PUSTAKA .....	39

## DAFTAR GAMBAR

Gambar 3. 1. Alur Penelitian .....	13
Gambar 4. 1. Program untuk pengambilan data .....	18
Gambar 4. 2. Program Transform Cases.....	21
Gambar 4. 3. Program proses cleaning data .....	22
Gambar 4. 4. Program untuk proses tokenisasi .....	23
Gambar 4. 5. Program untuk proses Stopword Removal .....	24
Gambar 4. 6. Program Normalisasi data.....	25
Gambar 4. 7. Program TF-IDF .....	26
Gambar 4. 8. Program untuk proses klasifikasi .....	28

## DAFTAR TABEL

Tabel 2. 1. Kappa Value .....	9
Tabel 2. 2. Confusion Matrix untuk dua kelas .....	11
Tabel 4. 1. Pelabelan Data Tweet .....	19
Tabel 4. 2. Kappa Value .....	20
Tabel 4. 3. Hasil Transform cases .....	21
Tabel 4. 4. Hasil Cleaning .....	22
Tabel 4. 5. Hasil Tokenization .....	23
Tabel 4. 6. Hasil Stopword Removal .....	24
Tabel 4. 7. Hasil Normalisasi .....	25
Tabel 4. 8. Hasil Pembobotan TF-IDF .....	26
Tabel 4. 9. Probabilitas sampel data latih .....	30
Tabel 4. 10. Sampel data uji .....	31
Tabel 4. 11. Hasil probabilitas data uji kelas positif dan negatif .....	32
Tabel 4. 12. Pembagian Data .....	34
Tabel 4. 13. Confusion Matrix evaluasi hasil data uji .....	35
Tabel 4. 14. Hasil accuracy, recall, precision dan F-1 score .....	36