

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP PERDAGANGAN
KARBON DI INDONESIA MENGGUNAKAN
ALGORITMA NAIVE BAYES**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

Kelvin Antonio
00000049042

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2024**

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP PERDAGANGAN
KARBON DI INDONESIA MENGGUNAKAN
ALGORITMA NAIVE BAYES**



SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

Kelvin Antonio
00000049042

UMN

**UNIVERSITAS
MULTIMEDIA
NUSANTARA**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2024

HALAMAN PERNYATAAN TIDAK PLAGIAT

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : kelvin antonio
Nomor Induk Mahasiswa : 00000049042
Program Studi : Informatika

Skripsi dengan judul:

analisis sentimen Pengguna Twitter terhadap perdagangan karbon di indonesia menggunakan algoritma naive bayes

merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/ penyimpangan, baik dalam pelaksanaan Skripsi maupun dalam penulisan laporan Skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas akhir yang telah saya tempuh.

Tangerang, 22 Mei 2024



(kelvin antonio)

HALAMAN PENGESAHAN

Skripsi dengan judul

ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP PERDAGANGAN KARBON DI INDONESIA MENGGUNAKAN ALGORITMA NAIVE BAYES

oleh

Nama : Kelvin Antonio
NIM : 00000049042
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Jumat, 31 Mei 2024

Pukul 10.00 s/s 12.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang



(Sy Yuliani Yakub, S.Kom., M.T. PhD)

NIDN:0411037904

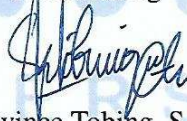
Penguji



(Wirawan Istiono, S.Kom., M.Kom.)

NIDN:0313048304

Pembimbing



(Fenina Adline Twince Tobing, S.Kom., M.Kom.)

NIDN: 0406058802

PJS. Ketua Program Studi Informatika,



(Dr. Eng. Niki Prastomo, S.T., M.Sc.)

NIDN: 0419128203

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Kelvin Antonio
NIM : 00000049042
Program Studi : Informatika
Jenjang : S1
Jenis Karya : Skripsi

Menyatakan dengan sesungguhnya bahwa:

- Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya di repositori Knowledge Center, sehingga dapat diakses oleh Civitas Akademika/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial dan saya juga tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.
- Saya tidak bersedia karena dalam proses pengajuan untuk diterbitkan ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*)**.

Tangerang, 22 Mei 2024

Yang menyatakan



Kelvin Antonio

U M M N
UNIVERSITAS
MULTIMEDIA
NUSANTARA

** Jika tidak bisa membuktikan LoA jurnal/HKI selama enam bulan ke depan, saya bersedia mengizinkan penuh karya ilmiah saya untuk diunggah ke KC UMN dan menjadi hak institusi UMN.

Halaman Persembahan / Motto

”Jika jalannya terlihat terlalu mudah mungkin kamu berada di jalan yang salah.”

Shanks (One Piece)

”Jika jalannya buntu, kita tinggal buat jalan yang baru saja.”

Roronoa Zoro (One Piece)

”Tidak peduli seberapa sulit atau tidak mungkin hal itu, jangan pernah kehilangan tujuanmu.”

Monkey D. Luffy (One Piece)

UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: analisis sentimen Pengguna Twitter terhadap perdagangan karbon di Indonesia menggunakan algoritma naive bayes dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku PJS. Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Ibu Fenina Adline Twince Tobing, S.Kom., M.Kom., sebagai Pembimbing pertama yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya tesis ini. memberikan bimbingan atas terselesainya Skripsi/Tesis ini.
5. Orang Tua dan teman-teman saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tesis ini.

Semoga skripsi ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 22 Mei 2024



Kelvin Antonio

ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP PERDAGANGAN KARBON DI INDONESIA MENGGUNAKAN ALGORITMA NAIVE BAYES

Kelvin Antonio

ABSTRAK

Perdagangan karbon adalah upaya untuk mengurangi emisi karbon melalui mekanisme pasar. Namun, implementasi kebijakan ini di Indonesia menghadapi tantangan, termasuk kurangnya pemahaman dan dukungan dari masyarakat. Oleh karena itu, penting untuk memahami sentimen masyarakat terhadap perdagangan karbon. Penelitian ini bertujuan untuk mengidentifikasi sentimen pengguna Twitter di Indonesia terhadap perdagangan karbon dan mengevaluasi keakuratan algoritma *Naive Bayes* dalam analisis sentimen tersebut. Data yang digunakan adalah *tweet* berbahasa Indonesia dan Inggris dari 20 Januari 2020 hingga 15 April 2024 dengan kata kunci perdagangan karbon. Metode meliputi pengumpulan data, *text preprocessing*, penerapan algoritma *Naive Bayes*, dan evaluasi model menggunakan metrik seperti *Accuracy*, *Recall*, *Precision*, dan *F1-Score*. Hasilnya menunjukkan tingkat akurasi yang memadai dengan akurasi 81% untuk model *Complement* dan *Bernoulli Naive Bayes*. Selain itu, terdapat juga hasil dari *recall* 79%, *precision* 87%, dan *f1-score* 83% pada model *Complement Naive Bayes*. Kemudian, terdapat juga hasil *recall* 71%, *precision* 89%, dan *f1-score* 79% pada model *Bernoulli Naive Bayes*. Akurasi keseluruhan dari penelitian yang dilakukan menggunakan model *Bernoulli Naive Bayes* dan *Complement Naive Bayes* adalah 81%.

Kata kunci: Algoritma *Naive Bayes*, Analisis sentimen, Perdagangan karbon, Twitter.

U M M N
U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

twitter user sentiment analysis on carbon trade in Indonesia using naive bayes algorithm

Kelvin Antonio

ABSTRACT

Carbon trading is an effort to reduce carbon emissions through market mechanisms. However, the implementation of this policy in Indonesia faces challenges, including a lack of understanding and support from the public. Therefore, it is important to understand the public sentiment towards carbon trading. This study aims to identify the sentiment of Twitter users in Indonesia towards carbon trading and evaluate the accuracy of the Naive Bayes algorithm in the sentiment analysis. The data used are Indonesian and English tweets from January 20, 2020 to April 15, 2024 with the keyword carbon trade. Methods include data collection, text preprocessing, application of Naive Bayes algorithm, and model evaluation using metrics such as Accuracy, Recall, Precision, and F1-Score. The results show adequate accuracy with 81% accuracy for Complement and Bernoulli Naive Bayes models. In addition, there are also results from recall 79%, precision 87%, and f1-score 83% in the Complement Naive Bayes model. Then, there are also results of recall 71%, precision 89%, and f1-score 79% in the Bernoulli Naive Bayes model. The overall accuracy of the research conducted using the Bernoulli Naive Bayes and Complement Naive Bayes models is 81%.

Keywords: *Carbon trading, Naive Bayes algorithm, Sentiment analysis, Twitter.*



DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
Listings	xiii
DAFTAR LAMPIRAN	xiv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	4
1.3 Batasan Permasalahan	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	5
1.6 Sistematika Penulisan	5
BAB 2 LANDASAN TEORI	7
2.1 Analisis Sentimen	7
2.2 Twitter	7
2.3 <i>Naive Bayes</i>	7
2.4 <i>Text Mining</i>	9
2.5 <i>Text-Preprocessing</i>	9
2.6 <i>TF-IDF</i>	11
2.7 <i>Confusion Matrix</i>	12
2.8 <i>K-Fold Cross Validation</i>	13
BAB 3 METODOLOGI PENELITIAN	15
3.1 Metodologi Penelitian	15
3.1.1 Studi Literatur	15
3.1.2 Pengumpulan Data	16
3.1.3 Perancangan Sistem	18
3.1.4 Validasi dan Evaluasi	19
3.1.5 Dokumentasi	19
3.2 Rancangan Sistem	20
3.2.1 <i>Text Mining</i>	20
3.2.2 <i>Text Pre-processing</i>	20
3.2.3 <i>TF-IDF</i>	22
3.2.4 <i>Data Splitting</i>	22
3.2.5 <i>Naive Bayes</i>	22
3.2.6 Pengujian Model	23
BAB 4 HASIL DAN DISKUSI	24
4.1 Spesifikasi Sistem	24
4.2 Implementasi Sistem	24
4.2.1 <i>Text Mining</i>	24
4.2.2 <i>Text Pre-processing</i>	27

4.3	<i>TF-IDF</i>	35
4.4	<i>Data Splitting</i>	36
4.5	Penerapan Model, Evaluasi dan Validasi	37
4.5.1	Penerapan Model	37
4.5.2	Evaluasi	47
4.5.3	Validasi	57
BAB 5	SIMPULAN DAN SARAN	63
5.1	Simpulan	63
5.2	Saran	63
DAFTAR PUSTAKA		64



DAFTAR GAMBAR

Gambar 3.1	<i>Flowchart Metodologi Penelitian</i>	15
Gambar 3.2	Pengumpulan Data	17
Gambar 3.3	Perancangan Sistem	18
Gambar 3.4	Evaluasi dan Validasi	19
Gambar 3.5	<i>Flowchart Rancangan Sistem</i>	20
Gambar 3.6	<i>Flowchart Text Pre-processing</i>	21
Gambar 4.1	Dataset	26
Gambar 4.2	<i>Translate Data</i>	30
Gambar 4.3	Hasil <i>Text Pre-processing</i>	35
Gambar 4.4	Hasil <i>Confusion Matrix Complement Naive Bayes</i>	44
Gambar 4.5	Hasil <i>Confusion Matrix Bernoulli Naive Bayes</i>	47



DAFTAR TABEL

Tabel 2.1	Confusion Matrix	12
Tabel 3.1	Tabel Studi Literatur Analisis Sentimen Menggunakan Algoritma Naive Bayes	16
Tabel 4.1	Sampel data	37
Tabel 4.2	Term Frequency	37
Tabel 4.3	Hasil Uji Coba Skenario Complement Naive Bayes	43
Tabel 4.4	Hasil Uji Coba Skenario Bernoulli Naive Bayes	46
Tabel 4.5	Tabel Confusion Matrix Complement Naive Bays	47
Tabel 4.6	Tabel Confusion Matrix Bernoulli Naive Bays	52
Tabel 4.7	validasi silang <i>Complement Naive Bayes</i> dengan <i>data split</i> 80:20	59
Tabel 4.8	validasi silang <i>Bernoulli Naive Bayes</i> dengan <i>data split</i> 80:20	59
Tabel 4.9	validasi silang <i>Bernoulli Naive Bayes</i> dengan <i>data split</i> 70:30	60
Tabel 4.10	validasi silang <i>Bernoulli Naive Bayes</i> dengan <i>data split</i> 70:30	61
Tabel 4.11	validasi silang <i>Bernoulli Naive Bayes</i> dengan <i>data split</i> 60:40	61
Tabel 4.12	validasi silang <i>Bernoulli Naive Bayes</i> dengan <i>data split</i> 70:30	62



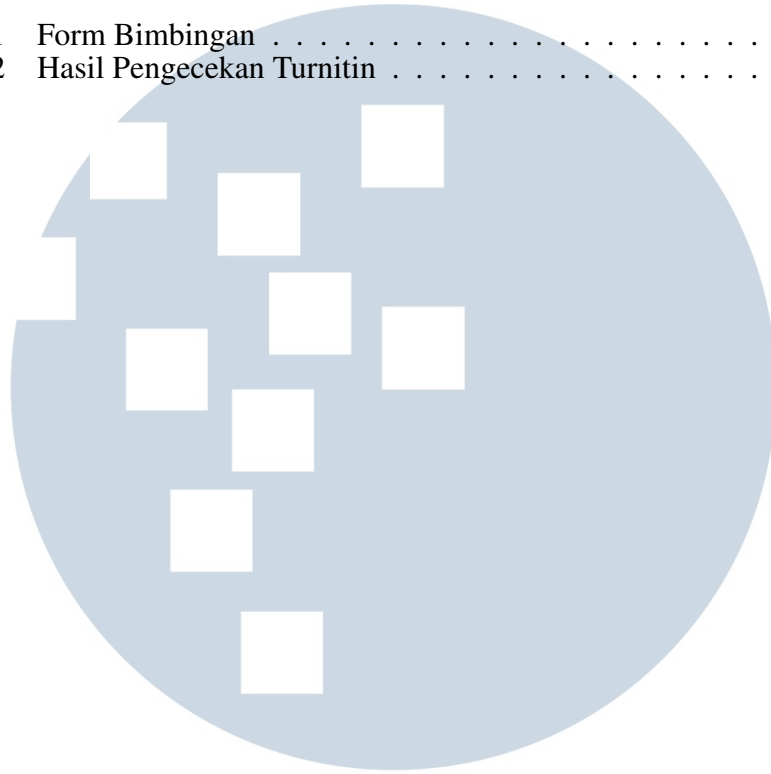
Listings

4.1	Potongan Kode Text Mining	25
4.2	Potongan kode Penggabungan Data	26
4.3	Potongan kode <i>Drop Data Duplicate</i>	28
4.4	Potongan kode untuk <i>check data Nan</i>	28
4.5	Potongan kode <i>slang word</i>	29
4.6	Potongan kode <i>Cleaning Data</i>	30
4.7	Potongan kode <i>Tokenization</i>	31
4.8	Potongan kode <i>Stopwords</i>	32
4.9	Potongan kode <i>Stemming</i>	33
4.10	Potongan kode <i>Labeling Data</i>	34
4.11	Potongan kode <i>TF-IDF Vectorize</i>	36
4.12	Potongan kode <i>Split Data</i>	36
4.13	Potongan kode <i>Complement Naive Bayes</i>	42
4.14	Potongan kode <i>Bernoulli Naive Bayes</i>	45
4.15	Potongan Kode K-Fold Cross Validation Untuk Bernoulli Naive Bayes dan Complement Naive Bayes	58



DAFTAR LAMPIRAN

Lampiran 1	Form Bimbingan	68
Lampiran 2	Hasil Pengecekan Turnitin	69



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA