

**ANALISIS OPINI PUBLIK TERHADAP INVESTASI DENGAN
MENGUNAKAN ALGORITMA MULTINOMIAL NAIVE
BAYES CLASSIFIER**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

KELVIN PRATAMA

19.11.2791

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

**ANALISIS OPINI PUBLIK TERHADAP INVESTASI DENGAN
MENGUNAKAN ALGORITMA MULTINOMIAL NAIVE
BAYES CLASSIFIER**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

KELVIN PRATAMA

19.11.2791

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

HALAMAN PERSETUJUAN

SKRIPSI

ANALISIS OPINI PUBLIK TERHADAP INVESTASI DENGAN
MENGUNAKAN ALGORITMA MULTINOMIAL NAIVE BAYES
CLASSIFIER

yang disusun dan diajukan oleh

Kelvin Pratama

19.11.2791

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 24 Juli 2025

Dosen Pembimbing,


Mafid Rahardi, S.Kom., M.Eng
NIK. 190302393

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS OPINI PUBLIK TERHADAP INVESTASI DENGAN
MENGUNAKAN ALGORITMA MULTINOMIAL NAIVE BAYES
CLASSIFIER

yang disusun dan diajukan oleh

Kelvin Pratama

19.11.2791

Telah dipertahankan di depan Dewan Penguji
pada tanggal 24 Juli 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Harvoko, S.Kom., M.Cs.
NIK. 190302286

Andika Agus Slameto, S.Kom., M.Kom.
NIK. 190302109

Majid Rahardl, S.Kom., M.Eng.
NIK. 190302393

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 24 Juli 2025

DEKAN FAKULTAS ILMU KOMPUTER

Prof. Dr. Kusri, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Kelvin Pratama
NIM : 19.11.2791

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Opini Publik terhadap Investasi dengan menggunakan Algoritma Multinomial Naive Bayes Classifier

Dosen Pembimbing : Majid Rahardi, S.Kom, M.Eng

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 24 Juli 2025

Yang Menyatakan,


Kelvin Pratama

HALAMAN PERSEMBAHAN

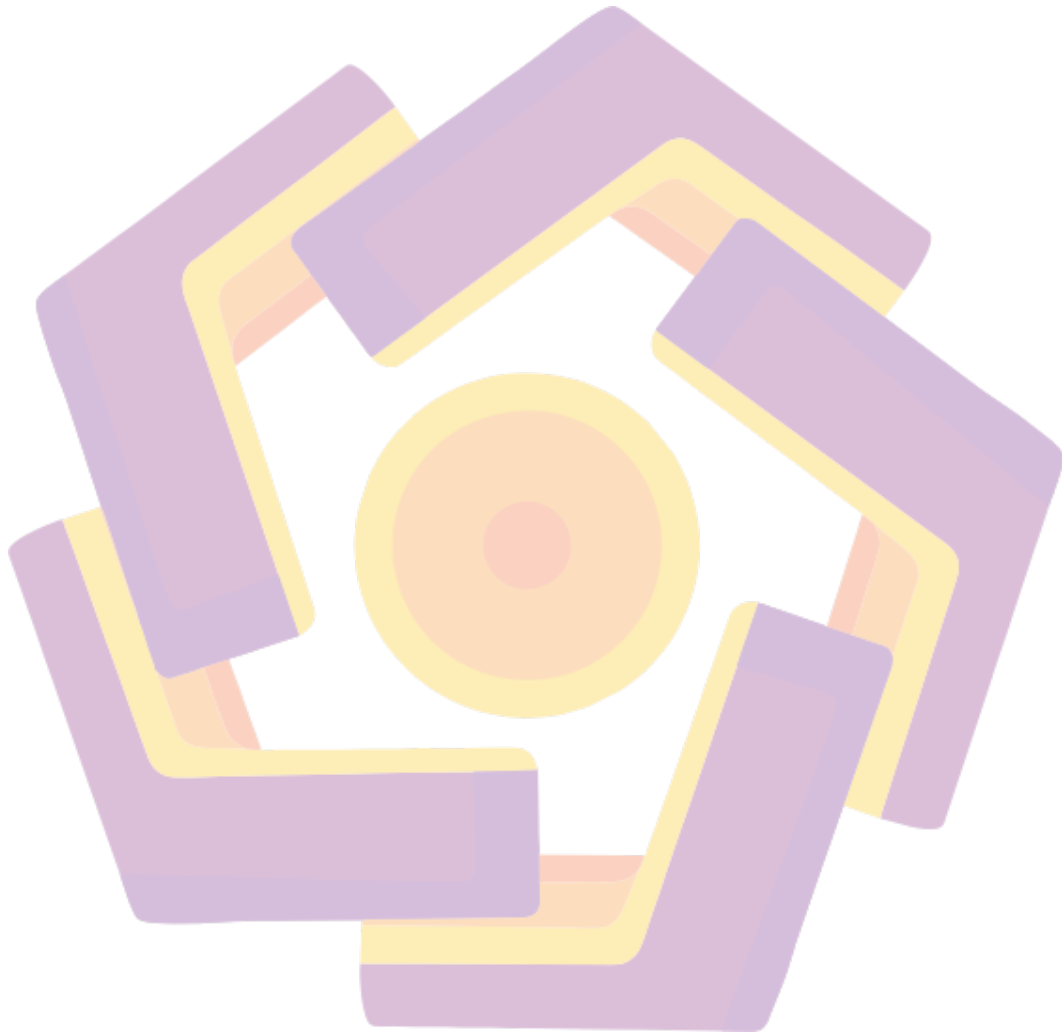
Bismillahirrahmanirrahim

Assalamualaikum Wr. Wb

Puji Syukur selalu tercurahkan kepada Ilahi Rabbi Allah Swt. Yang telah melimpahkan rahmat dan hidayah-Nya sehingga penulis dapat melaksanakan dan menyusun skripsi ini dengan baik. Penulisan karya ilmiah (skripsi) ini saya persembahkan kepada pihak ataupun element yang, memberikan berbagai pengaruh di kehidupan saya di masa-masa akhir jenjang pendidikan yang penuh perjuangan dan cobaan yang mendewasakan pribadi ini. Saya ingin mengucapkan terima kasih kepada:

1. Alhamdulillahirobbil'alamin, segala puji dan syukur hanya milik Allah SWT yang telah memberikan rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini.
2. Penulis juga ingin mengucapkan terima kasih sebesar-besarnya kepada kedua orang tua, yang selalu memberikan dukungan, doa, dan kasih sayang tanpa henti. Tanpa dukungan dan bimbingan mereka, penulis tidak akan mampu menyelesaikan pendidikan dan skripsi ini. Semoga Allah SWT senantiasa memberikan kesehatan, keselamatan, dan kebahagiaan yang berlimpah pada mereka.
3. Penulis juga ingin menyampaikan terima kasih yang tak terhingga kepada dosen pembimbing, Bapak Majid Rahardi, S.Kom, M.Eng. Yang telah memberikan bimbingan, masukan, dan arahan yang sangat berharga selama penulisan skripsi ini. Dosen pembimbing telah membantu penulis untuk mengembangkan ide dan wawasan, serta memberikan pandangan yang kritis dan konstruktif untuk memperbaiki skripsi ini. Penulis sangat berterima kasih atas waktu dan usahanya yang telah diberikan.

Terakhir, penulis juga ingin menyampaikan terima kasih kepada semua pihak yang telah membantu dan mendukung dalam penyelesaian skripsi ini, baik secara langsung maupun tidak langsung. Semoga Allah SWT senantiasa memberikan berkah dan kesuksesan pada kita semua.



KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh. Puji syukur peneliti panjatkan kehadiran Allah SWT karena atas limpahan rahmat, hidayah serta inayah-Nya, peneliti masih diberikan kesempatan dan kemudahan untuk menyelesaikan skripsi ini.

Skripsi ini disusun dalam rangka memenuhi salah satu syarat kelulusan perguruan tinggi program studi Strata 1 Informatika di Universitas Amikom Yogyakarta dan meraih gelar Sarjana Komputer (S. Kom) Pembuat skripsi ini tidak lepas dari berbagai pihak yang telah membantu baik dari segi materi dan spiritual. Penulis juga mengucapkan terimakasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. Suyanto, M.M., selaku rector Universitas Amikom Yogyakarta.
2. Bapak Majid Rahardi, S.Kom., M.Eng., selaku dosen pembimbing yang telah memberikan masukan, saran, bantuan dan bimbingan dalam menyelesaikan naskah skripsi ini.
3. Ibuk Prof. Dr. Kusrini, M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
4. Ibuk Eli Pujastuti, M.Kom., selaku ketua Program Studi Informatika Universitas Amikom Yogyakarta.
5. Dosen Universitas Amikom Yogyakarta yang telah memberikan ilmu dan pengalaman, terimakasih semua jasa Bapak dan Ibu sekalian.
6. Orang tua yang tidak pernah lelah dalam memberikan dukungan restu dan do'anya.
7. Teman-teman dan sahabat yang telah memberikan semangat, motivasi dan bantuan dalam pengerjaan skripsi ini
8. Seluruh staff karyawan Universitas Amikom Yogyakarta yang banyak membantu kelancaran segala aktivitas dan administrasi dalam penyusunan skripsi ini.

9. Semua pihak yang telah membantu sampai selesainya penyusunan skripsi ini yang tentunya sangat berharga dan tidak bisa disebutkan satu persatu.

Penulis menyadari sepenuhnya, bahwa skripsi ini masih jauh dari kesempurnaan, baik dalam hal penyajian skripsi maupun cara penyajian materi. Untuk itu dengan rendah hati peneliti memohon saran dan kritik yang membangun dari pembaca. Semoga skripsi ini dapat bermanfaat bagi penulis pada khususnya dan bagi pembaca pada umumnya serta dapat digunakan sebagai referensi untuk penelitian yang lain. Wassalamu'alaikum Warahmatullahi Wabarakatuh

Yogyakarta, 24 Juli 2025



Kelvin Pratama

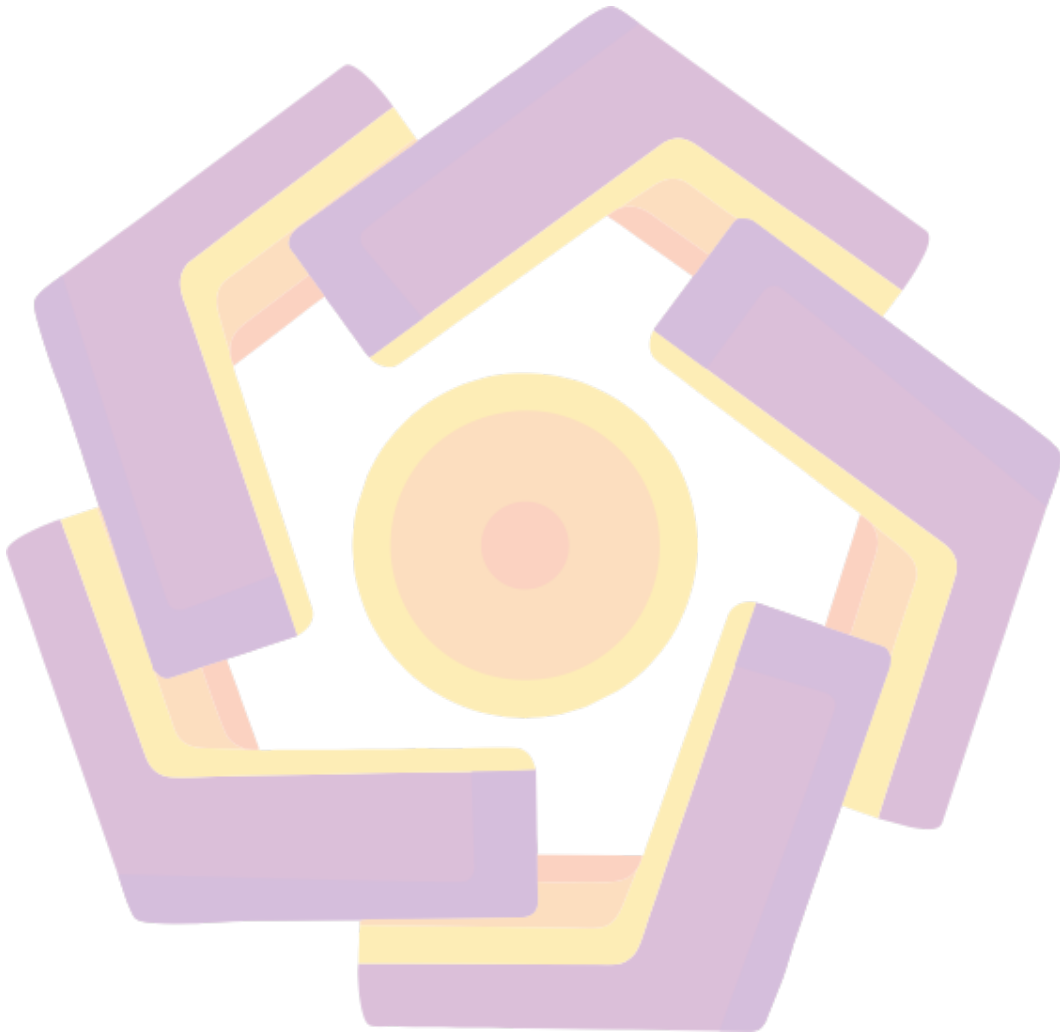
DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN.....	xv
DAFTAR LAMBANG DAN SINGKATAN	xvi
INTISARI	xviii
<i>ABSTRACTs</i>	xix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6

2.1	Studi Literatur	6
2.2	Landasan Teori.....	11
2.2.1	Investasi	11
2.2.2	Analisis Sentimen	11
2.2.3	Preprocessing	12
2.2.3.1	Terjemah	12
2.2.3.2	Normalisasi	12
2.2.3.3	Case Folding	12
2.2.3.4	Clean Punctuation	13
2.2.3.5	Tokenization	13
2.2.3.6	Stopword Removal.....	13
2.2.3.7	Lemmatization	13
2.2.4	Feature Extraction TF-IDF	14
2.2.5	Multinomial Naïve Bayes Classifier	15
2.2.6	Evaluasi Matrix	16
BAB III METODE PENELITIAN		19
3.1	Objek Penelitian.....	19
3.2	Alur Penelitian	19
3.4.	Crawling dan Menentukan Kata Kunci.....	20
3.5.	Preprocessing Data.....	20
3.4.2.1	Terjemah	20
3.4.2.2	Case Folding	21

3.4.2.3	Clean Punctuation	21
3.4.2.4	Tokenizing	22
3.4.2.5	Stopwords Removals	23
3.4.2.6	Lemmatization	23
3.6.	Feature Ekstraksi TF-IDF	24
3.7.	Validation dan Evaluasi	27
3.8.	Alat dan Bahan.....	29
BAB IV HASIL DAN PEMBAHASAN		31
4.1	Crawling.....	31
4.2	Preprocessing	32
4.2.1	Terjemah	32
4.2.2	Case Folding	33
4.2.3	Clean Punctuation	33
4.2.4	Tokenizing	34
4.2.5	Stopword Removals	34
4.2.6	Lemmatization	35
4.3	Labeling	36
4.4	Term Weighting TF-IDF.....	37
4.5	Pseudocode Multinomial Naïve Bayes	37
4.6	Evaluasi Matrix	38
BAB V PENUTUP		42
5.1	Kesimpulan	42

5.2	Saran	42
REFERENSI		43
LAMPIRAN.....		46



DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian	8
Tabel 2. 2 Confusion Matrix.....	17
Tabel 3. 1 Terjemah Bahasa.....	21
Tabel 3. 2 Case Folding	21
Tabel 3. 3 Clean Punctuation	22
Tabel 3. 4 Tokenizing	23
Tabel 3. 5 Stopwords Removals	23
Tabel 3. 6 Lemmatization	24
Tabel 3. 7. Tabel Sample Document.....	25
Tabel 3. 8 Term Frequency	25
Tabel 3. 9 Invers Document Frequency	26
Tabel 3. 10 TF-IDF Weighting	26
Tabel 3. 11 Confusion Matrix.....	28

DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian	19
Gambar 3. 2 K-Fold Cross Validation	27
Gambar 4. 1 Kode Program Binding Crawling Harvest.....	30
Gambar 4. 2 Kode Program Jarak waktu crawling dengan kata kunci.....	31
Gambar 4. 3 Kode Program Terjemah	31
Gambar 4. 4 Kode Case Folding.....	32
Gambar 4. 5 Kode Clean Tag pada Punctuation.....	32
Gambar 4. 6 Kode Clean Punctuation.....	32
Gambar 4. 7 Kode Tokenizing.....	33
Gambar 4. 8 Kode Stopwords Removals	33
Gambar 4. 9 Kode Lemmatization.....	34
Gambar 4. 10 Kode Program Labeling	35
Gambar 4. 11 Chart Komparasi Label (a) 70 (b) 80 dan (c) 85.....	35
Gambar 4. 12 Kode Program TF-IDF	36
Gambar 4. 13 Pseudocode Naive Bayes Classifier	37
Gambar 4. 14 Threshold Classification 0.70	37
Gambar 4. 15 Cross Validation Threshold 70.....	38
Gambar 4. 16 Threshold Classification 0.80	38
Gambar 4. 17 Croos Validation Threshold 0.80	39
Gambar 4. 18 Threshold Classification 0.85	39
Gambar 4. 19 Croos Validation Threshold 0.85	40

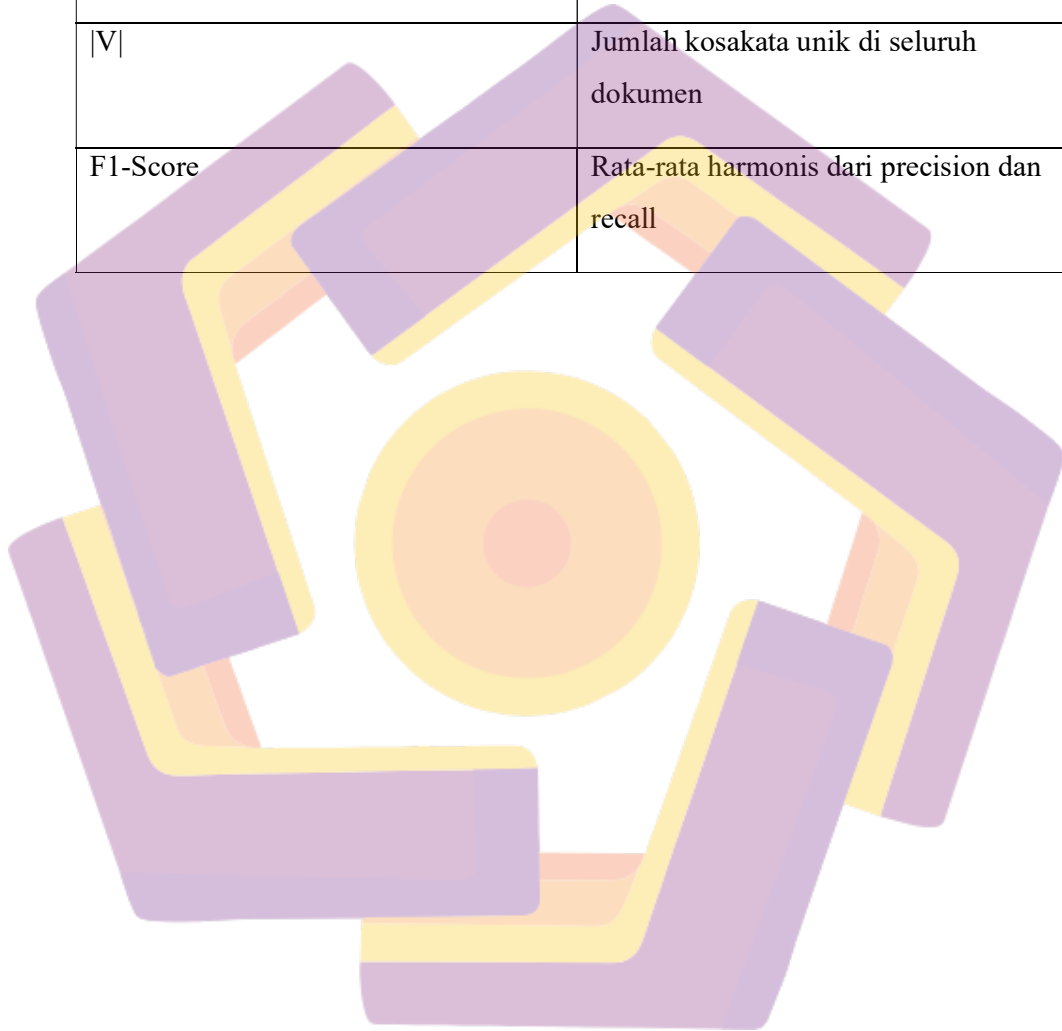
DAFTAR LAMPIRAN

1. Lampiran 1: Dataset Awal (Tweet Mentah).....32
Deskripsi: Kumpulan data mentah hasil crawling dari akun Twitter terkait investasi digital.
2. Lampiran 2: Hasil Pre-processing.....33
Deskripsi: Data setelah melalui tahapan cleansing, case folding, stopword removal, dan stemming.
3. Lampiran 3: Kode Program Pre-processing.....32
Deskripsi: Cuplikan kode Python untuk proses pembersihan data menggunakan NLP dan NLTK.
4. Lampiran 4: Hasil TF-IDF.....37
Deskripsi: Representasi numerik dari data teks setelah dihitung bobot TF-IDF.
5. Lampiran 5: Kode Program TF-IDF.....37
Deskripsi: Implementasi kode Python untuk menghitung nilai TF dan IDF.
6. Lampiran 6: Kode Program Multinomial Naive Bayes.....38
Deskripsi: Script utama klasifikasi menggunakan algoritma Multinomial Naive Bayes.
7. Lampiran 7: Confusion Matrix.....38
Deskripsi: Visualisasi hasil prediksi berupa matrix klasifikasi: TP, TN, FP, FN.
8. Lampiran 8: Nilai Akurasi, Precision, Recall, F1-Score.....38
Deskripsi: Output evaluasi performa model.

DAFTAR LAMBANG DAN SINGKATAN

Lambang/Singkatan	Keterangan
TF	Term Frequency
IDF	Inverse Document Frequency
TF-IDF	Term Frequency – Inverse Document Frequency
NLP	Natural Language Processing
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
NBC	Naïve Bayes Classifier
SVM	Support Vector Machine
α	Alpha (parameter smoothing untuk Naive Bayes)
B	Banyaknya kosakata (Vocabulary size)
N	Jumlah seluruh dokumen
df	Document Frequency (jumlah dokumen yang mengandung term tertentu)
Colab	Google Colaboratory

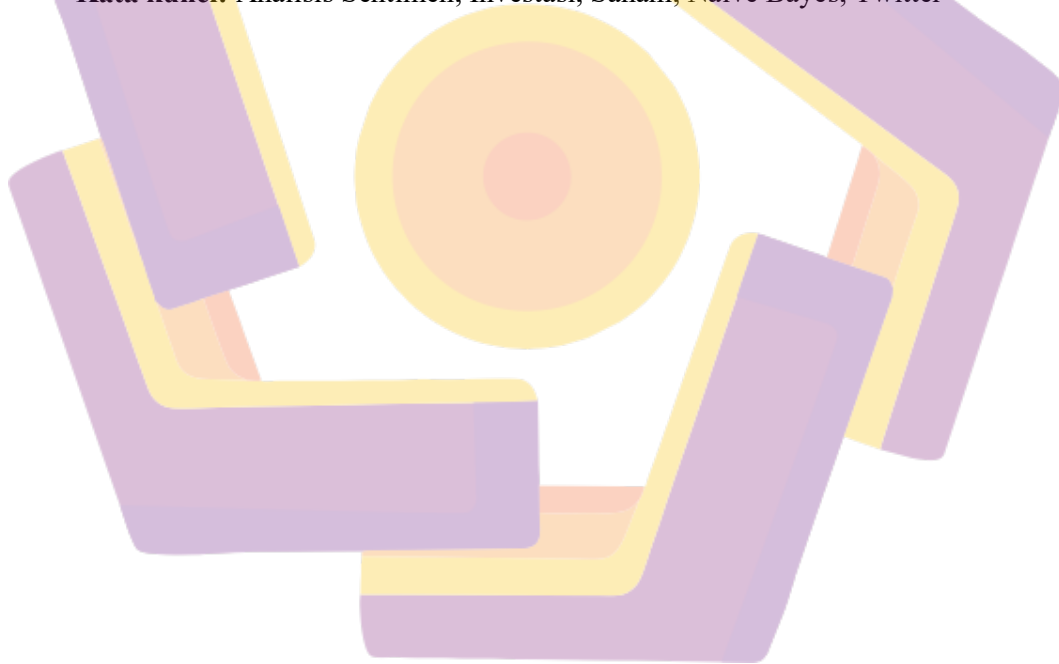
NLTK	Natural Language Toolkit
$TF(t,d)$	Frekuensi kata t dalam dokumen d
$IDF(t)$	$\log(N/df(t))$ – Invers frekuensi dokumen
$ V $	Jumlah kosakata unik di seluruh dokumen
F1-Score	Rata-rata harmonis dari precision dan recall



INTISARI

Perkembangan teknologi informasi di era Industri 4.0 telah memengaruhi berbagai aspek kehidupan, termasuk investasi. Namun, kesadaran masyarakat Indonesia terhadap pentingnya investasi masih rendah. Penelitian ini bertujuan menganalisis sentimen masyarakat di Twitter terkait investasi online menggunakan algoritma Multinomial Naïve Bayes Classifier untuk mengklasifikasikan opini menjadi positif dan negatif. Data tweet dikumpulkan dengan kata kunci investasi dan saham dari Januari 2023 hingga Desember 2024. Tahapan penelitian meliputi preprocessing terjemah, case folding, tokenisasi, stopword removal, lemmatization, ekstraksi fitur TF-IDF, dan klasifikasi. Hasil evaluasi menunjukkan akurasi tertinggi 90% dengan presisi 0,92 dan recall 0,93 untuk sentimen positif, serta akurasi cross validation dengan akurasi 87%. Temuan ini dapat menjadi acuan bagi regulator, pelaku bisnis, dan edukator keuangan dalam meningkatkan literasi investasi masyarakat di Indonesia.

Kata kunci: Analisis Sentimen, Investasi, Saham, Naïve Bayes, Twitter



ABSTRACTs

The development of information technology in the Industry 4.0 era has influenced various aspects of life, including investment. However, Indonesian people's awareness of the importance of investment is still low. This study aims to analyze public sentiment on Twitter regarding online investment using the Multinomial Naïve Bayes Classifier algorithm to classify opinions into positive and negative. Tweet data was collected with the keywords investment and stocks from January 2023 to December 2024. The research stages include translation preprocessing, case folding, tokenization, stopword removal, lemmatization, TF-IDF feature extraction, and classification. The evaluation results showed the highest accuracy of 90% with a precision of 0.92 and a recall of 0.93 for positive sentiment, as well as cross validation accuracy with an accuracy of 87%. These findings can be a reference for regulators, business actors, and financial educators in improving investment literacy in the Indonesian community.

Keyword: Sentiment Analysis, Investment, Stocks, Naïve Bayes, Twitter

