

SKRIPSI

**ANALISIS SENTIMEN GLOBAL TERHADAP TAGAR BRICS
INDONESIA MENGGUNAKAN MODEL MULTINOMIAL
NAIVE BAYES DENGAN LAPLACE SMOOTHING DAN
SMOTE**



Intelligentia - Dignitas

ANANDA RIZKY SANDIKA

1512621068

PROGRAM STUDI

PENDIDIKAN TEKNIK INFORMATIKA DAN KOMPUTER

FAKULTAS TEKNIK

UNIVERSITAS NEGERI JAKARTA

2025

ABSTRAK

Ananda Rizky Sandika, Analisis Sentimen Global Terhadap Tagar BRICS Indonesia Menggunakan Model *Multinomial Naïve Bayes* dengan *Laplace Smoothing* dan *SMOTE*. Dosen Pembimbing : Dr. Widodo, M.Kom. dan Murien Nugraheni, M.Cs. Program Studi Pendidikan Teknik Informatika dan Komputer. Fakultas Teknik. Universitas Negeri Jakarta. 2025

Penelitian ini bertujuan untuk mengevaluasi kinerja algoritma *Multinomial Naïve Bayes* dalam analisis sentimen terhadap topik #BRICS Indonesia, dengan membandingkan beberapa pendekatan pengujian. Empat metode diuji, yaitu: (1) *Multinomial Naïve Bayes* tanpa modifikasi, (2) *Multinomial Naïve Bayes* dengan *Laplace smoothing*, (3) *Multinomial Naïve Bayes* dengan *Synthetic Minority Over-sampling Technique* (SMOTE), dan (4) kombinasi antara *Laplace smoothing* dan *SMOTE*. Hasil pengujian menunjukkan bahwa metode *Multinomial Naïve Bayes* dengan *Laplace smoothing* memberikan performa terbaik, dengan *accuracy* 74%, *precision* 86%, *recall* 57%, dan *f1-score* 60%. Sementara itu, *Multinomial Naïve Bayes* tanpa modifikasi menghasilkan *accuracy* 51%, *precision* 56%, *recall* 56%, dan *f1-score* 48%. Metode dengan SMOTE menghasilkan *accuracy* 46%, *precision* 55%, *recall* 55%, dan *f1-score* 44%. Adapun kombinasi *Laplace smoothing* dan *SMOTE* memberikan *accuracy* 71%, *precision* 68%, *recall* 75%, dan *f1-score* 68%. Berdasarkan hasil tersebut, dapat disimpulkan bahwa penggunaan algoritma *Multinomial Naïve Bayes* dengan *Laplace smoothing* mampu memberikan peningkatan kinerja yang paling signifikan dalam klasifikasi sentimen menggunakan *Multinomial Naïve Bayes*. Temuan ini diharapkan dapat menjadi referensi bagi penelitian selanjutnya yang berkaitan dengan analisis sentimen dan penerapan algoritma pembelajaran mesin pada isu-isu global.

Kata Kunci: Analisis sentimen, *Multinomial Naïve Bayes*, *Laplace Smoothing*, *SMOTE*, *BRICS*, Teks klasifikasi, *Machine Learning*.

ABSTRACT

Ananda Rizky Sandika, *Global Sentiment Analysis Towards the BRICS Indonesia Hashtag Using the Multinomial Naïve Bayes Model with Laplace Smoothing and SMOTE. Supervisors: Dr. Widodo, M.Kom. and Murien Nugraheni, M.Cs. Study Program of Technical Education in Informatics and Computer. Faculty of Engineering. State University of Jakarta. 2025.*

This research aims to evaluate the performance of the Multinomial Naïve Bayes algorithm in sentiment analysis on the topic of #BRICS Indonesia, by comparing several testing approaches. Four methods were tested: (1) Multinomial Naïve Bayes without modification, (2) Multinomial Naïve Bayes with Laplace smoothing, (3) Multinomial Naïve Bayes with Synthetic Minority Over-sampling Technique (SMOTE), and (4) a combination of Laplace smoothing and SMOTE. The test results show that the Multinomial Naïve Bayes method with Laplace smoothing provides the best performance, with an accuracy of 74%, precision of 86%, recall of 57%, and f1-score of 60%. Meanwhile, the Multinomial Naïve Bayes without modification yields an accuracy of 51%, precision of 56%, recall of 56%, and f1-score of 48%. The method with SMOTE results in an accuracy of 46%, precision of 55%, recall of 55%, and f1-score of 44%. The combination of Laplace smoothing and SMOTE provides an accuracy of 71%, precision of 68%, recall of 75%, and f1-score of 68%. Based on these results, it can be concluded that the use of the Multinomial Naïve Bayes algorithm with Laplace smoothing significantly improves performance in sentiment classification using Multinomial Naïve Bayes. These findings are expected to serve as a reference for future research related to sentiment analysis and the application of machine learning algorithms on global issues.

Keywords: *Sentiment Analysis, Multinomial Naïve Bayes, Laplace Smoothing, SMOTE, BRICS, Text Classification, Machine Learning.*

LEMBAR PENGESAHAN SKRIPSI

Judul : ANALISIS SENTIMEN GLOBAL TERHADAP TAGAR
BRICS INDONESIA MENGGUNAKAN MODEL
MULTINOMIAL NAÏVE BAYES DENGAN LAPLACE
SMOOTHING DAN SMOTE

Penyusun : Ananda Rizky Sandika

NIM : 1512621068

Disetujui oleh:

NAMA DOSEN	TANDA TANGAN	TANGGAL
Dosen Pembimbing I, Dr. Widodo, M.Kom NIP.1972032520 05011002		30-06-2025
Dosen Pembimbing II, Murien Nugraheni, M.Cs NIP.198710112019032012		30-06-2025

Telah disetujui oleh

NAMA DOSEN	TANDA TANGAN	TANGGAL
Ketua Penguji, Diat Nurhidayat, S.Pd., M.T.I NIP. 198308192018031001		28-07-2025
Dosen Penguji II, Ali Idrus, M.Kom NIP. 198802262019031010		28-07-2025
Dosen Penguji III, Shindy Arty, S.Pd., M.Eng NIP. 199312272024062002		28-07-2025

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Skripsi ini merupakan karya asli dan belum pernah diajukan untuk mendapatkan gelar akademik sarjana, baik di Universitas Negeri Jakarta maupun di Perguruan Tinggi lain.
2. Skripsi ini belum dipublikasikan, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
3. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Universitas Negeri Jakarta.

Jakarta, 30 Juni 2025

Yang membuat pernyataan



Ananda Rizky Sandika

NIM.1512621068



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET DAN TEKNOLOGI
UNIVERSITAS NEGERI JAKARTA
UPT PERPUSTAKAAN

Jalan Rawamangun Muka Jakarta 13220
Telepon/Faksimili: 021-4894221
Laman: lib.unj.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : Ananda Rizky Sandika
NIM : 1512621068
Fakultas/Prodi : Teknik / Pendidikan Teknik Informatika Komputer
Alamat email : rizkysandika19@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada UPT Perpustakaan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

☒ Skripsi ☐ Tesis ☐ Disertasi ☐ Lain-lain (.....)

yang berjudul :

ANALISIS SENTIMEN GLOBAL TERHADAP TAGAR BRICS INDONESIA

MENGUNAKAN MODEL MULTINOMIAL NAIVE BAYES DENGAN LAPLACE

SMOOTHING DAN SMOTE

Dengan Hak Bebas Royalti Non-Eksklusif ini UPT Perpustakaan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/mempublikasikannya di internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 3 Agustus 2025

Penulis

(Ananda Rizky Sandika)
nama dan tanda tangan

KATA PENGANTAR

Assalamu'alaikum Wr.Wb.

Puji syukur penulis panjatkan kepada Allah SWT yang senantiasa memberikan rahmat dan kesehatan, sehingga penulis dapat menyelesaikan penulisan skripsi ini dengan baik. Shalawat dan salam semoga selalu tercurah kepada Nabi Muhammad SAW.

Skripsi yang berjudul “Analisis Sentimen Pandangan Masyarakat Global Terhadap BRICS Menggunakan Algoritma *Multinomial Naïve Bayes*” ini diajukan untuk memenuhi syarat penyelesaian Tugas Akhir pada Program Studi Pendidikan Teknik Informatika dan Komputer, Fakultas Teknik, Universitas Negeri Jakarta.

Penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada semua pihak yang telah terlibat, baik secara langsung maupun tidak langsung, dalam proses penyelesaian skripsi ini. Penulis menyadari bahwa keberhasilan ini bukan hanya hasil usaha keras penulis sendiri, tetapi juga berkat dukungan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Bapak Dr. Widodo, M.Kom., selaku Dosen Pembimbing I yang telah memberikan banyak saran dan nasihat dalam penulisan skripsi ini.
2. Ibu Murien Nugraheni, M.Cs., selaku Dosen Pembimbing II yang telah membimbing, memberi arahan, serta saran dalam penulisan skripsi ini.
3. Bapak Mochammad Ficky Duskarnaen, M.Sc., selaku Koordinator Program Studi Pendidikan Teknik Informatika dan Komputer, Fakultas Teknik, Universitas Negeri Jakarta.
4. Bapak/Ibu dosen dan seluruh staf Program Studi Pendidikan Teknik Informatika dan Komputer yang telah banyak membantu selama perkuliahan serta memberikan motivasi dan inspirasi bagi penulis untuk terus berjuang menyelesaikan proses perkuliahan.
5. Orang tua dan keluarga besar penulis yang selalu mendukung, memberikan banyak masukan, serta motivasi kepada penulis sehingga penulis dapat menyelesaikan skripsi ini dengan baik.

6. Teman-teman mahasiswa Program Studi Pendidikan Teknik Informatika dan Komputer tahun 2021 yang telah berjuang bersama-sama, saling bertukar pikiran, serta memberikan dukungan satu sama lain selama perkuliahan hingga tahap penyelesaian skripsi ini.
7. Semua pihak yang tidak dapat disebutkan satu per satu yang telah terlibat dalam membantu penulisan skripsi ini.

Sebagai penulis menyadari bahwa dalam penulisan skripsi ini masih terdapat berbagai kekurangan dan keterbatasan. Meskipun demikian, penulis berharap bahwa karya ini dapat memberikan manfaat yang signifikan bagi para pembaca. Penulis juga berharap skripsi ini dapat menjadi sumber referensi yang berguna untuk pengembangan penelitian di masa depan.

Dengan penuh kesadaran, penulis memohon maaf jika terdapat kesalahan atau kekurangan dalam penyampaian informasi. Setiap kritik dan saran yang konstruktif sangat penulis hargai, karena hal tersebut akan membantu penulis dan peneliti lainnya untuk terus meningkatkan kualitas penelitian di bidang ini. Semoga skripsi ini tidak hanya menjadi sebuah tugas akademis, tetapi juga dapat memberikan kontribusi positif bagi perkembangan ilmu pengetahuan dan praktik di lapangan.

Jakarta, 13 Mei 2025

Penulis,



Ananda Rizky Sandika

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	i
LEMBAR PERNYATAAN	ii
KATA PENGANTAR.....	iii
ABSTRAK.....	v
ABSTRACT.....	vi
DAFTAR ISI	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
BAB I PENDAHULUAN	12
1.1 Latar Belakang	12
1.2 Identifikasi Masalah	16
1.3 Pembatasan Masalah	16
1.4 Perumusan Masalah	17
1.5 Tujuan Penelitian.....	17
1.6 Manfaat Penelitian	18
BAB II KAJIAN PUSTAKA	19
2.1 Kerangka Teoritik.....	19
2.1.1 Analisis Sentimen.....	19
2.1.2 <i>BRICS</i>	19
2.1.3 Teorema Bayes	20
2.1.4 <i>Naïve Bayes</i>	21
2.1.5 <i>Multinomial Naïve Bayes</i>	22
2.1.6 Crawling Data	23
2.1.7 <i>Text Preprocessing</i>	24
2.1.7.1 <i>Text Cleaning</i>	24
2.1.7.2 <i>Case Folding</i>	25
2.1.7.3 <i>Normalization</i>	26

2.1.7.4 <i>Stopwords</i>	26
2.1.7.5 <i>Tokenization</i>	27
2.1.8 <i>Laplace smoothing</i>	28
2.1.9 <i>SMOTE</i>	28
2.1.10 Confusion Matrix	29
2.2 Penelitian Relevan	31
2.3 Kerangka Berpikir	38
BAB III METODOLOGI PENELITIAN	40
3.1 Tempat dan Waktu Penelitian	40
3.1.1 Tempat Penelitian	40
3.1.2 Waktu Penelitian	40
3.2 Alat dan Bahan Penelitian	41
3.2.1 Perangkat Keras (<i>Hardware</i>)	41
3.2.2 Perangkat Lunak (<i>Software</i>)	41
3.3 Diagram Alir Penelitian	42
3.3.1 Crawling Data	43
3.3.2 Dataset Mentah	44
3.3.3 <i>Translate</i>	44
3.3.4 <i>Filtering dan Text Preprocessing</i>	45
3.3.5 Dataset Terolah	46
3.3.6 <i>Labeling</i>	46
3.3.7 Data Analisis	47
3.3.8 Pembobotan TF-IDF	47
3.3.9 Pengklasifikasian dengan Algoritma <i>Multinomial Naïve Bayes</i> dengan <i>Laplace smoothing</i> dan <i>SMOTE</i>	48
3.3.10 Evaluasi Kinerja	48
3.3.11 Perbandingan Akurasi	49
3.3.12 Penarikan Kesimpulan	50
3.4 Teknik dan Prosedur Pengumpulan Data	50
3.5 Teknik Analisis Data	52
BAB IV HASIL PENELITIAN	53

4.1 Hasil Penelitian	53
4.2 Hasil <i>Crawling</i> Data.....	53
4.3 Translate	54
4.4 Pre-processing	55
4.4.1 Text Cleaning	56
4.4.2 Case Folding	57
4.4.3 Normalization.....	57
4.4.4 Stopwords.....	58
4.4.5 Tokenization	59
4.5 Pelabelan Data.....	60
4.6 Klasifikasi Multinomial Naïve Bayes	61
4.7 Klasifikasi <i>Multinomial Naïve Bayes</i> dan Laplace Smoothing.....	62
4.8 Klasifikasi Multinomial Naïve Bayes dan SMOTE.....	63
4.9 Klasifikasi Multinomial Naïve Bayes, Laplace Smoothing dan SMOTE...	64
4.10 Analisis Hasil Penelitian	65
4.10.1 Evaluasi <i>Multinomial Naïve Bayes</i>	65
4.10.2 Evaluasi <i>Multinomial Naïve Bayes</i> dan <i>Laplace Smoothing</i>	67
4.10.3 Evaluasi <i>Multinomial Naïve Bayes</i> dan SMOTE	69
4.10.4 Evaluasi <i>Multinomial Naïve Bayes, Laplace Smoothing</i> dan SMOTE	71
4.11 Pembahasan.....	73
BAB V PENUTUP.....	79
5.1 Kesimpulan	79
5.2 Saran.....	80
DAFTAR PUSTAKA	81
LAMPIRAN.....	86
RIWAYAT PENULIS	100