# Klasifikasi Sentimen Opini Publik terhadap Program Koperasi Desa/ Kelurahan Merah Putih pada Media Sosial *X* Menggunakan *Multinomial Naïve Bayes*


Disusun Oleh :

**TENI DEINAROSA HERMALIA**

**2C2220005**


**SKRIPSI**

Diajukan sebagai salah satu syarat memperoleh Gelar Sarjana Sains Data

**KONSENTRASI DIGITAL BUSINESS ANALYTIC**


**PROGRAM STUDI SARJANA SAINS DATA**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS KOPERASI INDONESIA**


**2026**

## LEMBAR PENGESAHAN

| | | |
|---|---|---|
| Judul | : | Klasifikasi Sentimen Opini Publik terhadap Program Koperasi Desa/ Kelurahan Merah Putih pada Media Sosial *X* Menggunakan *Multinomial Naïve Bayes* |
| Nama | : | Teni Deinarosa Hermalia |
| NIM | : | 2C2220005 |
| Fakultas | : | Sains dan Teknologi |
| Program Studi | : | Sains Data |
| Konsentrasi | : | *Digital Business Analytic* |

Menyetujui dan mengesahkan

| Pembimbing 1 | Pembimbing 2 |
|---|---|
| Muhammad Haris Fadhillah, S.E.,M.M. | Onoy Rohaeni, Dra., M.Sc. |

| Pjs.Dekan<br>Fakultas Sains dan Teknologi | Pjs.Ketua Program Studi<br>Sains Data |
|---|---|
| Ir. Hj. Hady Nurpady, MP.,CIIQA | Ir. Hj. Hady Nurpady, MP.,CIIQA |

## RIWAYAT HIDUP



Teni Deinarosa Hermalia lahir pada tanggal 14 Desember 2002 di Garut, Jawa Barat. Penulis merupakan anak pertama dari dua bersaudara, putri dari Bapak Almarhum Tito Hermanto, S.Pd., dan Ibu Neneng Meylia Rosidah.

Adapun Riwayat Pendidikan formal yang ditempuh penulis adalah sebagai berikut :

1. Tahun 2009 menyelesaikan Pendidikan di SPS AR-AROHMAH.
2. Tahun 2015 menyelesaikan Pendidikan di SDN Keresek 6.
3. Tahun 2018 menyelesaikan Pendidikan di SMP Negeri 1 Cibatu.
4. Tahun 2021 menyelesaikan Pendidikan di SMK Negeri 4 Garut.

Pada tahun 2022, penulis terdaftar sebagai mahasiswi aktif Program Studi S-1 Sains Data di Universitas Koperasi Indonesia. Selama menempuh Pendidikan, penulis aktif dalam berbagai kegiatan organisasi dan kemahasiswaan, yaitu :

1. Badan Eksekutif Mahasiswa Universitas Koperasi Indonesia sebagai staf magang di Kementerian Media Komunikasi dan Informasi (Medkominfo) pada tahun 2022. Selanjutnya, penulis menjadi anggota pengurus Kementerian Medkominfo periode 2022-2023, kemudian dipercaya menjabat sebagai Menteri Kementerian Medkominfo periode 2024-2025.
2. Unit Kegiatan Mahasiswa Sanggar Seni 12 Juli Ikopin sebagai anggota kepengurusan Divisi Tari selama 2 periode.

3. Asisten Laboratorium Mata Kuliah Praktikum Pemograman Dasar dan Basis Data pada Tahun 2025.
4. Komunitas Penerimaan Beasiswa Bank Indonesia, Generasi Baru Indonesia (GenBI) sebagai Sekretaris Divisi Kewirausahaan perode 2024-2025 dan penulis dipercaya menjabat sebagai Ketua Divisi Kewirausahaan periode 2025-2026.

IKOPIN University

## PERNYATAAN MENGENAI LAPORAN SKRIPSI DAN SUMBER INFORMASI SERTA PEMLIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa laporan skripsi dengan judul **"Klasifikasi Sentimen Opini Publik terhadap Program Koperasi Desa/ Kelurahan Merah Putih pada Media Sosial *X* Menggunakan *Multinomial Naïve Bayes*"** adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apapun kepada perguruan tinggi manapun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam daftar pustaka di bagian akhir laporan akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Universitas Koperasi Indonesia.

Jatinangor, September 2026




Teni Deinarosa Hermalia
2C2220005

## ABSTRACT

**Teni Deinarosa Hermalia, 2026.** *Sentimen Classification of Public Opinion towards the Koperasi Desa/Kelurahan Merah Putih Program on Social Media X using Multinomial Naïve Bayes. Undergraduate Thesis, Undergraduate Data Science Program, Faculty of Science and Technology, IKOPIN University. Supervised by Muhammad Haris Fadhillah and Onoy Rohaeni.*

*The Koperasi Desa/Kelurahan Merah Putih Program is a government program that has generated various public responses on social media X. This study aims to analyze the distribution of sentiments, identify dominant words, compare four classification scenarios, and determine the best-performing scenario. The data were obtained through web scraping from social media X during the period of April 30–May 31, 2026. Of the 8,954 tweets collected, 7,313 tweets remained after selection and preprocessing. The tweets were manually labeled based on their content and context into positive, negative, and neutral categories. The data were divided using stratified random sampling with an 80:20 ratio, resulting in 5,850 training data and 1,463 test data. Classification was performed using TF-IDF and Multinomial Naïve Bayes through four scenarios: unigram without stemming, unigram with stemming, bigram without stemming, and bigram with stemming. Model performance was evaluated using accuracy, precision, recall, Macro F1-score, Weighted F1-score, and Balanced Accuracy, with Macro F1-score as the primary metric.*

*The results show that neutral sentiment was the most dominant category, followed by negative and positive sentiments. The best-performing scenario was TF-IDF unigram without stemming, achieving a Macro F1-score of 0.72 as the primary metric, an accuracy of 72%, a Weighted F1-score of 0.72, and a Balanced Accuracy of 0.71. Classification errors primarily occurred in data containing implicit or ambiguous meanings. These results demonstrate the model's ability to predict manually annotated labels within the research dataset and do not represent the entirety of public opinion. The generalizability of the findings is limited to social media X data collected based on the specified keywords and research period.*



***Keywords: Sentimen Classification, Multinomial Naïve Bayes, Koperasi Desa/Kelurahan Merah Putih Program.***



IKOPIN University

## ABSTRAK

**Teni Deinarosa Hermalia, 2026**. *Klasifikasi Sentimen Opini Publik terhadap Program Koperasi Desa/Kelurahan Merah Putih pada Media Sosial X Menggunakan Multinomial Naïve Bayes. Skripsi Program Studi Sarjana Sains Data, Fakultas Sains dan Teknologi, Universitas Koperasi Indonesia. Di bawah bimbingan Muhammad Haris Fadhillah dan Onoy Rohaeni.*

Program Koperasi Desa/Kelurahan Merah Putih merupakan program pemerintah yang memunculkan beragam tanggapan masyarakat pada media sosial X. Penelitian ini bertujuan menganalisis distribusi sentimen, mengidentifikasi kata dominan, membandingkan empat skenario klasifikasi, dan menentukan skenario terbaik. Data diperoleh melalui *web scraping* pada media sosial *X* periode 30 April–31 Mei 2026. Dari 8.954 tweet yang dikumpulkan, diperoleh 7.313 tweet setelah seleksi dan *preprocessing*. Pelabelan dilakukan secara manual berdasarkan isi dan konteks *tweet* ke dalam kategori positif, negatif, dan netral. Data dibagi menggunakan stratified random sampling dengan rasio 80:20, yaitu 5.850 data latih dan 1.463 data uji. Klasifikasi menggunakan *TF-IDF* dan *Multinomial Naïve Bayes* melalui empat skenario, yaitu *unigram* tanpa *stemming*, *unigram* dengan *stemming*, *bigram* tanpa *stemming*, dan *bigram* dengan *stemming*. Kinerja model dievaluasi menggunakan *accuracy*, *precision*, *recall*, Macro F1-score, *Weighted F1-score*, dan Balanced *Accuracy*, dengan *Macro F1-score* sebagai metrik utama.

Hasil penelitian menunjukkan bahwa sentimen netral merupakan kategori paling dominan, diikuti negatif dan positif. Skenario terbaik adalah *TF-IDF unigram* tanpa *stemming* dengan *Macro F1-score* sebesar 0,72 sebagai metrik utama, *accuracy* sebesar 72%, *Weighted F1-score* sebesar 0,72, dan *Balanced Accuracy* sebesar 0,71. Kesalahan klasifikasi terutama terjadi pada data dengan makna implisit atau ambigu. Hasil tersebut menunjukkan kemampuan model dalam memprediksi label anotasi manual pada data penelitian dan tidak merepresentasikan keseluruhan opini masyarakat. Generalisasi hasil terbatas pada data media sosial *X* yang dikumpulkan berdasarkan kata kunci dan periode penelitian.



Kata Kunci: Klasifikasi Sentimen, *Multinomial Naïve Bayes*, Program Koperasi Desa/Kelurahan Merah Putih



IKOPIN University

## KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah Swt. atas segala rahmat, karunia, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains pada Program Studi Sarjana Sains Data, Fakultas Sains dan Teknologi, Universitas Koperasi Indonesia.

Dalam proses penyusunan skripsi ini, penulis menyadari bahwa penyelesaiannya tidak terlepas dari berbagai tantangan dan keterbatasan. Namun, berkat doa, dukungan, bantuan, serta bimbingan dari berbagai pihak, skripsi ini akhirnya dapat diselesaikan. Oleh karena itu, dengan segala kerendahan hati, penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Orang tua dan keluarga penulis yang selalu memberikan doa, kasih sayang, dukungan, motivasi, serta segala bentuk pengorbanan yang tidak ternilai selama penulis menjalani perkuliahan hingga menyelesaikan skripsi ini.
2. Dekan Fakultas Sains dan Teknologi Universitas Koperasi Indonesia beserta jajaran yang telah memberikan dukungan dan fasilitas kepada penulis selama menjalani pendidikan dan menyelesaikan skripsi ini.
3. Muhammad Haris Fadhillah, S.E., M.M. dan Dra. Onoy Rohaeni, M.Sc. selaku dosen pembimbing yang telah meluangkan waktu membimbing, memberikan arahan, saran, dan masukan kepada penulis dalam menyelesaikan skripsi ini.

4. Muhammad Fahreza, S.E., M.B.A. selaku dosen penelaah konsentrasi yang telah banyak memberikan saran serta masukan-masukan kepada penulis selama proses penyelesaian skripsi ini serta proses pembelajaran selama proses pembelajaran selama perkuliahan.
5. Agus Nugraha, S.T., M.Kom selaku dosen penelaah koperasi yang telah banyak memberikan saran serta masukan-masukan kepada penulis selama proses penyelesaian skripsi ini serta proses pembelajaran selama proses pembelajaran selama perkuliahan.
6. Seluruh dosen Program Studi Sarjana Sains Data, Fakultas Sains dan Teknologi, Universitas Koperasi Indonesia yang telah memberikan banyak ilmu serta pembelajaran selama masa perkuliahan yang akan selalu menjadi bekal bagi penulis untuk kedepanya.
7. Seluruh staf dan tenaga kependidikan Universitas koperasi Indonesia yang telah membantu penulis selama proses perkuliahan.
8. Bank Indonesia yang telah memberikan kesempatan kepada penulis untuk memperoleh Beasiswa Bank Indonesia. Dukungan yang diberikan melalui beasiswa tersebut telah membantu penulis selama menjalani masa perkuliahan dan menjadi salah satu bagian penting dalam perjalanan penulis hingga dapat menyelesaikan pendidikan ini.
9. Mulyati Eka saputri, Indri Nur Sukmawati, Silvi nurinsan, dan Cahya Rahmawati selaku teman seperjuangan selama masa perkuliahan yang telah menjadi bagian dari perjalanan penulis, saling membantu,

memberikan semangat, berbagi cerita, dan menemani berbagai proses hingga penyelesaian skripsi ini.

10. Nadila Delvarisa Utami dan Karima Nur Faidah yang telah menjadi teman sejak masa sekolah hingga saat ini. Terima kasih telah tetap hadir dan menjadi bagian dari perjalanan hidup penulis dalam berbagai fase, serta selalu memberikan dukungan dan semangat.
11. Silvi Nurul fazri dan Rio Aryanto selaku teman Kuliah Kerja Nyata (KKN) yang telah menjadi bagian dari salah satu pengalaman berharga selama masa perkuliahan, berbagi cerita, teman guyon, kebersamaan, dan berbagai pengalaman yang akan selalu dikenang oleh penulis.
12. Terima Kasih penulis sampaikan kepada grup music NCT, khususnya Nakamoto Yuta melalui karya- karya dan dedikasinya secara tidak langsung telah memberikan motivasi, ketenangan pikiran, serta energi positif yang senantiasa menemani setiap Langkah penulisan skripsi ini.
13. Seluruh pihak yang tidak dapat penulis sebutkan satu per satu yang telah memberikan bantuan, dukungan, doa, dan kebaikan hingga penyelesaian skripsi ini.

Jatinangor, 14 Agustus 2026

Penulis

IKOPIN University

## DAFTAR ISI

**LEMBAR PENGESAHAN** ..... **i**
**RIWAYAT HIDUP** ..... **ii**
**ABSTRACT** ..... **v**
**ABSTRAK** ..... **vi**
**KATA PENGANTAR** ..... **vii**
**DAFTAR ISI** ..... **x**
**DAFTAR GAMBAR** ..... **xii**
**DAFTAR TABEL** ..... **xiii**
**DAFTAR LAMPIRAN** ..... **xv**
**BAB I PENDAHULUAN** ..... **1**
1.1. Latar Belakang Penelitian ..... 1
1.2. Identifikasi Masalah ..... 10
1.3. Maksud dan Tujuan Penelitian ..... 11
1.3.1. Maksud Penelitian ..... 11
1.3.2. Tujuan Penelitian ..... 11
1.4. Kegunaan Penelitian ..... 12
1.4.1. Kegunaan Teoritis ..... 12
1.4.2. Kegunaan Praktis ..... 13
**BAB II PENDEKATAN MASALAH DAN METODE PENELITIAN** ..... **15**
2.1. Pendekatan Masalah ..... 15
2.1.1. Pendekatan Koperasi ..... 15
2.1.2. Program Koperasi Desa/Kelurahan Merah Putih ..... 23
2.1.3. Kebijakan Publik dan Opini Publik Digital ..... 29
2.1.4. Media Sosial *X* ..... 32
2.1.5. Sains Data ..... 34
2.1.6. Penelitian Terdahulu ..... 68
2.1.7. Kerangka Berfikir ..... 71
2.2. Metode Penelitian ..... 71
2.2.1. Metode Penelitian yang Digunakan ..... 71
2.2.2. Data yang Diperlukan ..... 72

2.2.3. Sumber Data dan Cara Menentukan Sumber Data ..... 73
2.2.4. Teknik Pengumpulan Data ..... 74
2.2.5. Analisis Data ..... 76
2.2.6. Tempat Penelitian ..... 82
2.2.7. Jadwal Penelitian ..... 82
**BAB III KEADAAN UMUM TEMPAT PENELITIAN ..... 86**
3.1. Gambaran Umum Koperasi Desa/Kelurahan Merah Putih Se-Indonesia ..... 86
3.2. Media Sosial X sebagai Tempat Penelitian ..... 89
3.3. Lingkungan sistem Penelitian ..... 90
3.3.1. Perangkat Keras ..... 90
3.3.2. Perangkat Lunak ..... 91
3.3.3. *Tools* dan *Library* Penelitian ..... 91
**BAB IV HASIL DAN PEMBAHASAN ..... 93**
4.1. Analisis Distribusi Sentimen dan Karakteristik Kata Dominan ..... 93
*4.1.1. Data Understanding* ..... 93
4.1.2. Distribusi Sentimen ..... 96
4.1.3. Karakteristik Kata Dominan ..... 97
4.2. Analisis Penerapan Metode Klasifikasi Sentimen ..... 106
*4.2.1. Data Preparation* ..... 106
*4.2.2. Modelling* ..... 112
4.3. Analisis Performa Klasifikasi Antar Skenario ..... 115
4.3.1. Penerapan *Preprocessing Stemming* ..... 116
4.3.2. Penerapan Representasi Fitur *N-gram* ..... 117
4.3.3. Penentuan Model Terbaik ..... 118
4.4. Analisis Karakteristik Kesalahan Hasil Klasifikasi Model Terbaik ..... 118
*4.4.1. Classification Report* ..... 118
*4.4.2. Confusion Matrix* ..... 119
*4.4.3. Error Analysis* ..... 120
*4.4.4. Deployment* ..... 122
**BAB V SIMPULAN DAN SARAN ..... 124**
5.1. Simpulan ..... 124
5.2. Saran ..... 125
**DAFTAR PUSTAKA ..... 127**
**LAMPIRAN ..... 134**

## DAFTAR GAMBAR

Gambar 1.1 Overview Program Koperasi Desa/Kelurahan Merah Putih ..... 4
Gambar 2.1 Tahapan Metodologi *CRISP-DM* ..... 38
Gambar 2.2 *Confusion Matrix* ..... 63
Gambar 2.3 Bagan Kerangka Berpikir ..... 71
Gambar 3.1 Peta Sebaran Koperasi Desa/Kelurahan Merah Putih di Indonesia .. 88
Gambar 3.2 Logo Media Sosial *X* ..... 89
Gambar 4.1 Distribusi Sentimen *Dataset* ..... 96
Gambar 4.2 *Chart Top* 10 Kata Dominan Keseluruhan ..... 98
Gambar 4.3 *Chart Top* 10 Kata Dominan Positif ..... 100
Gambar 4.4 *Chart Top* 10 Kata Dominan Negatif ..... 102
Gambar 4.5 *Chart Top* 10 Kata Dominan Netral ..... 104
Gambar 4.6 *Classification Report* ..... 119
Gambar 4.7 *Confusion Matrix* Skenario 1 ..... 120

IKOPIN University

## DAFTAR TABEL


Tabel 1. 1 Contoh Percakapan Pengguna Media Sosial *X* ..... 6
Tabel 2.1 Pedoman Pelabelan Sentimen ..... 48
Tabel 2.2 Penelitian Terdahulu ..... 69
Tabel 2.3 Empat Skenario Eksperimen ..... 80
Tabel 2.4 Jadwal Penelitian ..... 83
Tabel 3.1 Jumlah Koperasi Desa/Kelurahan Merah Putih per Provinsi ..... 87
Tabel 3.2 Daftar Perangkat Keras ..... 90
Tabel 3.3 Daftar Perangkat Lunak ..... 91
Tabel 3.4 *Tools* dan *Library* Penelitian ..... 91
Tabel 4.1 Rincian Hasil pengumpulan *Web Scraping* ..... 94
Tabel 4.2 Ringkasan Audit *Dataset* ..... 95
Tabel 4.3 Sampel *Tweet* Keseluruhan *Dataset* ..... 99
Tabel 4.4 Sampel Tweet Sentimen Positif ..... 101
Tabel 4.5 Sampel *Tweet* Sentimen Negatif ..... 102
Tabel 4.6 Sampel *Tweet* Netral ..... 105
Tabel 4.7 Sampel Hasil Tahapan *Case folding* ..... 107
Tabel 4.8 Sampel Hasil Tahapan *Cleaning Text* ..... 108
Tabel 4.9 Sampel Hasil Tahapan *Tokenizing* ..... 108
Tabel 4.10 Sampel Hasil Tahapan Normalisasi ..... 109
Tabel 4.11 Sampel Hasil Tahapan *Stopword removal* ..... 110
Tabel 4.12 Sampel Hasil Tahapan *Stemming* ..... 111
Tabel 4.13 Hasil Akhir Proses *Preprocessing* ..... 111
Tabel 4.14 Pembagian Data Latih dan Data Uji ..... 112
Tabel 4.15 Empat Skenario Eksperimen ..... 112
Tabel 4.16 Sample Hasil Pembobotan *TF-IDF Unigram* (Skenario 1 dan 2) ... 113
Tabel 4. 17 Sample Hasil Pembobotan *TF-IDF Bigram* (Skenario 3 dan 4) ..... 113
Tabel 4.18 Contoh Prediksi *Multinomial Naïve Bayes* ..... 115
Tabel 4.19 Hasil Evaluasi Setiap Skenario ..... 116

Tabel 4.20 Ringkasan Kesalahan Klasifikasi Model Terbaik ..... 121
Tabel 4.21 Contoh *Tweet* yang Mengalami Salah Klasifikasi ..... 121

IKOPIN University

## DAFTAR LAMPIRAN

Lampiran 1. Proses *Crawling* Data ..... 134
Lampiran 2. *Syntak* Evaluasi ..... 135
Lampiran 3. Kamus Normalisasi ..... 135

IKOPIN University