

**Analisis Sentimen Media Sosial X Terhadap Isu Tapera  
Menggunakan Metode XGBoost Berbasis PSO**

**Laporan Tugas Akhir**

Diajukan untuk memenuhi  
Persyaratan Guna Meraih Gelar Sarjana Informatika  
Universitas Muhammadiyah Malang



Saniyya Ruzzy Marwa  
202110370311177

**Bidang Minat**  
Data Science

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS MUHAMMADIYAH MALANG**

**2025**

## LEMBAR PERSETUJUAN

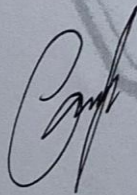
**Analisis Sentimen Media Sosial X Terhadap Isu Tapera  
Menggunakan Metode XGBoost Berbasis PSO**

**TUGAS AKHIR**

Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1  
Teknik Informatika Universitas Muhammadiyah Malang

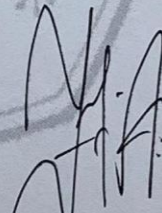
Menyetujui,  
Malang, 5 Mei 2025

Dosen Pembimbing 1



**Christian Sri Kusuma, S.Kom, M.Kom**  
NIP. 180327021991PNS.

Dosen Pembimbing 2



**Yufis Azhar, S.Kom, M.Kom**  
NIP. 10814100544PNS.

## LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini :

NAMA : SANIYYA RUZZY MARWA

NIM : 202110370311177

FAK./JUR. : INFORMATIKA

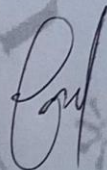
Dengan ini saya menyatakan bahwa Tugas Akhir dengan judul "ANALISIS SENTIMEN MEDIA SOSIAL X TERHADAP ISU TAPERA MENGGUNAKAN METODE XGBOOST BERBASIS PSO" beserta seluruh isinya adalah karya saya sendiri dan bukan merupakan karya tulis orang lain, baik sebagian maupun seluruhnya, kecuali kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya ini, atau ada klaim dari pihak lain terhadap keaslian karya saya ini maka saya siap menanggung segala bentuk resiko/sanksi yang berlaku.

Mengetahui,  
Dosen Pembimbing

Malang, 19 Mei 2025

Yang Membuat Pernyataan



Christian Sri Kusuma, S.Kom, M.Kom



Saniyya Ruzzy Marwa

# LEMBAR PENGESAHAN

## LEMBAR PENGESAHAN

**Analisis Sentimen Media Sosial X Terhadap Isu Tapera  
Menggunakan Metode XGBoost Berbasis PSO**

### TUGAS AKHIR

Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1  
Informatika Universitas Muhammadiyah Malang

Disusun Oleh :

**SANIYYA RUZZY MARWA**


**202110370311177**


Tugas Akhir ini telah diuji dan dinyatakan lulus melalui sidang majelis penguji  
pada tanggal 18 Juni 2025

Menyetujui,


Dosen Penguji 1

Dosen Penguji 2

  
Vinna Rahmayanti S.Si., M.Si  
NIP. 180306071990PNS.

  
Ir. Wildan Suharso S.Kom., M.Kom  
NIP. 10817030596PNS.

Mengetahui,  
Ketua Jurusan Informatika

  
Ir. Gohir Wasis Wicaksono S.kom. M.Cs.  
NIP. 10814100541PNS.

## ABSTRAK

Penelitian ini bertujuan untuk menganalisis sentiment masyarakat terhadap isu Tabungan Rakyat (Tapera) dengan menggunakan data dari media sosial X. Data dikumpulkan dengan cara crawling pada media sosial X menggunakan kata kunci “Isu Tapera” dan mendapatkan sebanyak 2.268 data. Setelah data terkumpul, dilakukan proses pelabelan data secara manual untuk menentukan kategori sentiment yaitu negative, positif, netral dan others. Dilanjutkan dengan tahap pra-pemrosesan data. Selanjutnya dilakukan ekstraksi fitur menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF), kemudian data dibagi menjadi data latih dan data uji dengan perbandingan 80:20. Algoritma yang digunakan dalam penelitian ini adalah Extreme Gradient Boosting (XGBoost) yang kemudian dioptimasi menggunakan Particle Swarm Optimization (PSO) untuk mendapatkan parameter terbaik yang mampu meningkatkan akurasi model. Hasil penelitian menunjukkan bahwa penerapan XGBoost berbasis PSO menghasilkan akurasi sebesar 85%, lebih tinggi dibandingkan dengan akurasi model XGBoost sebelum optimasi. Hasil ini menunjukkan bahwa kombinasi metode XGBoost dan PSO efektif dalam mengklasifikasikan sentimen masyarakat terhadap isu Tapera, serta menunjukkan potensi penerapannya dalam analisis opini publik berbasis media sosial.

**Kata Kunci:** Sentimen, Tapera, Media Sosial X, XGBoost, Particle Swarm Optimization (PSO), TF-IDF

## ABSTRACT

This study aims to analyze public sentiment towards the issue of People's Savings (Tapera) using data from X social media. Data was collected by crawling on social media X using the keyword “Tapera Issue” and obtaining 2,268 data. After the data is collected, a manual data labeling process is carried out to determine the sentiment categories, namely negative, positive, neutral and others. Followed by the data pre-processing stage. Next, feature extraction is carried out using the Term Frequency-Inverse Document Frequency (TF-IDF) method, then the data is divided into training data and test data with a ratio of 80:20. The algorithm used in this research is Extreme Gradient Boosting (XGBoost) which is then optimized using Particle Swarm Optimization (PSO) to get the best parameters that can improve model accuracy. The results showed that the application of PSO-based XGBoost resulted in an accuracy of 85%, higher than the accuracy of the XGBoost model before optimization. This result shows that the combination of XGBoost and PSO methods is effective in classifying public sentiment on the Tapera issue, and shows its potential application in social media-based public opinion analysis.

**Keywords** : Sentiment, Tapera, Social media X, XGBoost, PSO, TF - DF

## LEMBAR PERSEMBAHAN

Puji syukur kepada Allah SWT atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan Tugas Akhir ini. Penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada :

1. Allah SWT yang senantiasa memberikan kehidupan, kesehatan, serta kemudahan dalam setiap langkah yang saya tempuh.
2. Kedua orang tua penulis Bapak Marwakun dan Almh. Umi Nur Hayati yang senantiasa memberikan dorongan, semangat, kasih sayang, doa, dan dukungan sepenuhnya kepada penulis dalam mengerjakan penelitian ini.
3. Kakak saya Silvia Farokhah S.Pd dan putra kecilnya Bartelsi Ibram Santoso yang senantiasa memberikan dukungan dengan memberikan canda dan tawa.
4. Bapak Christian Sri Kusuma Aditya, S.Kom, M.Kom dan Bapak Ir. Yufis Azhar, S.Kom, M.Kom. selaku pembimbing tugas akhir yang telah meluangkan waktu untuk membimbing penulis sehingga dapat menyelesaikan penelitian tugas akhir ini.
5. Bapak/Ibu Dekan Fakultas Teknik Universitas Muhammadiyah Malang.
6. Bapak Ir. Galih Wasis Wicaksono, S.Kom., M.Cs., selaku Ketua Jurusan Teknik Informatika Muhammadiyah Malang.
7. Seluruh Dosen beserta Staff Program Studi Informatika Universitas Muhammadiyah Malang yang telah memberikan ilmu dan pengalaman selama perkuliahan.
8. Kepada teman-teman yang telah memberikan dukungan tanpa henti, semangat yang tiada habisnya, berbagai masukan berharga, serta menjadi tempat berbagi dalam suka maupun duka.
9. Saya pribadi yang telah berhasil menyelesaikan studi.

Malang, 06 Mei 2025

  
Saniyya Ruzzy Marwa

## KATA PENGANTAR

Segala puji bagi Allah SWT atas keberkahan dan hidayah yang diberikan sehingga dapat menyelesaikan penelitian dengan judul “Analisis Sentimen Media Sosial X Terhadap Isu Tapera Menggunakan Metode XGBoost Berbasis PSO”.

Penelitian ini memiliki tujuan untuk melakukan klasifikasi sentimen yang diekspresikan oleh pengguna media sosial X terhadap isu Tapera. Penulis berharap hasil penelitian ini bisa memberikan pengetahuan model klasifikasi yang lebih akurat dan efisien, yang mampu memberikan detail sentimen publik secara lebih mendalam dan memberikan pemahaman yang lebih baik terhadap persepsi masyarakat mengenai isu Tapera yang menjadi topik perdebatan.

Pada penulisan ini masih memiliki keterbatasan. Maka dari itu, penulis menerima saran untuk perbaikan dan penyempurnaan supaya penelitian ini memberikan manfaat khususnya kepada para peneliti.

Malang, 06 Mei 2025

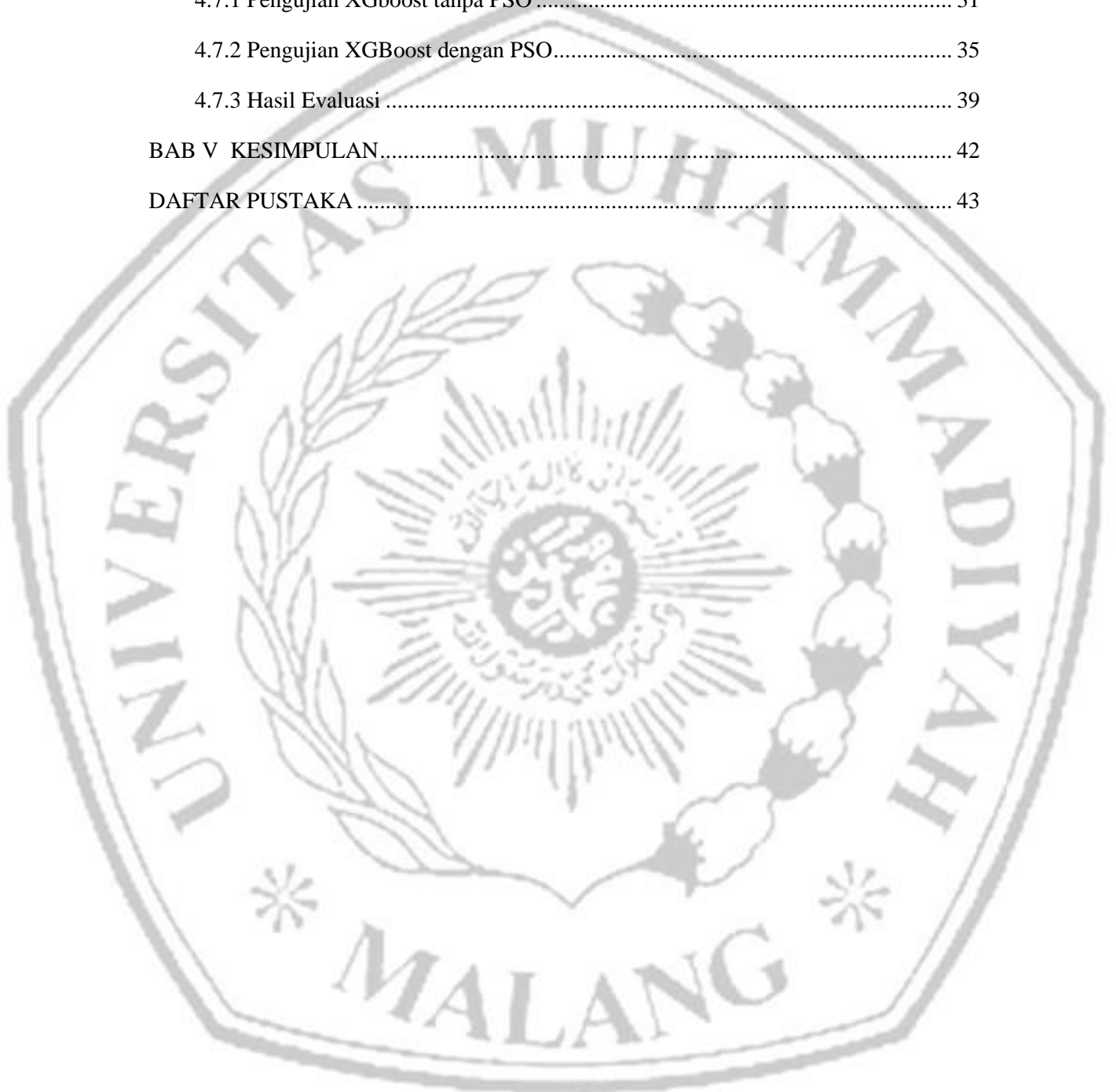
  
Saniyya Ruzzy Marwa

## DAFTAR ISI

LEMBAR PERSETUJUAN .....	ii
LEMBAR PERNYATAAN .....	iii
LEMBAR PENGESAHAN .....	iv
ABSTRAK .....	v
ABSTRACT .....	vi
LEMBAR PERSEMBAHAN .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI .....	ix
DAFTAR GAMBAR .....	xii
DAFTAR TABEL .....	xiii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian .....	4
1.4 Batasan Masalah .....	5
BAB II TINJAUAN PUSTAKA .....	6
2.1 Studi Literatur .....	6
2.2 Implementasi Metode dalam Analisis Sentimen .....	7
2.3 Media Sosial X .....	8
2.4 Natural Language Processing (NLP) .....	8
2.5 TF - IDF .....	9
2.6 Pemodelan XGBoost .....	9
2.7 PSO (Particle Swarm Optimization) .....	10
BAB III METODE PENELITIAN .....	11
3.1 Pengumpulan Data .....	12
3.2 Seleksi Data .....	13

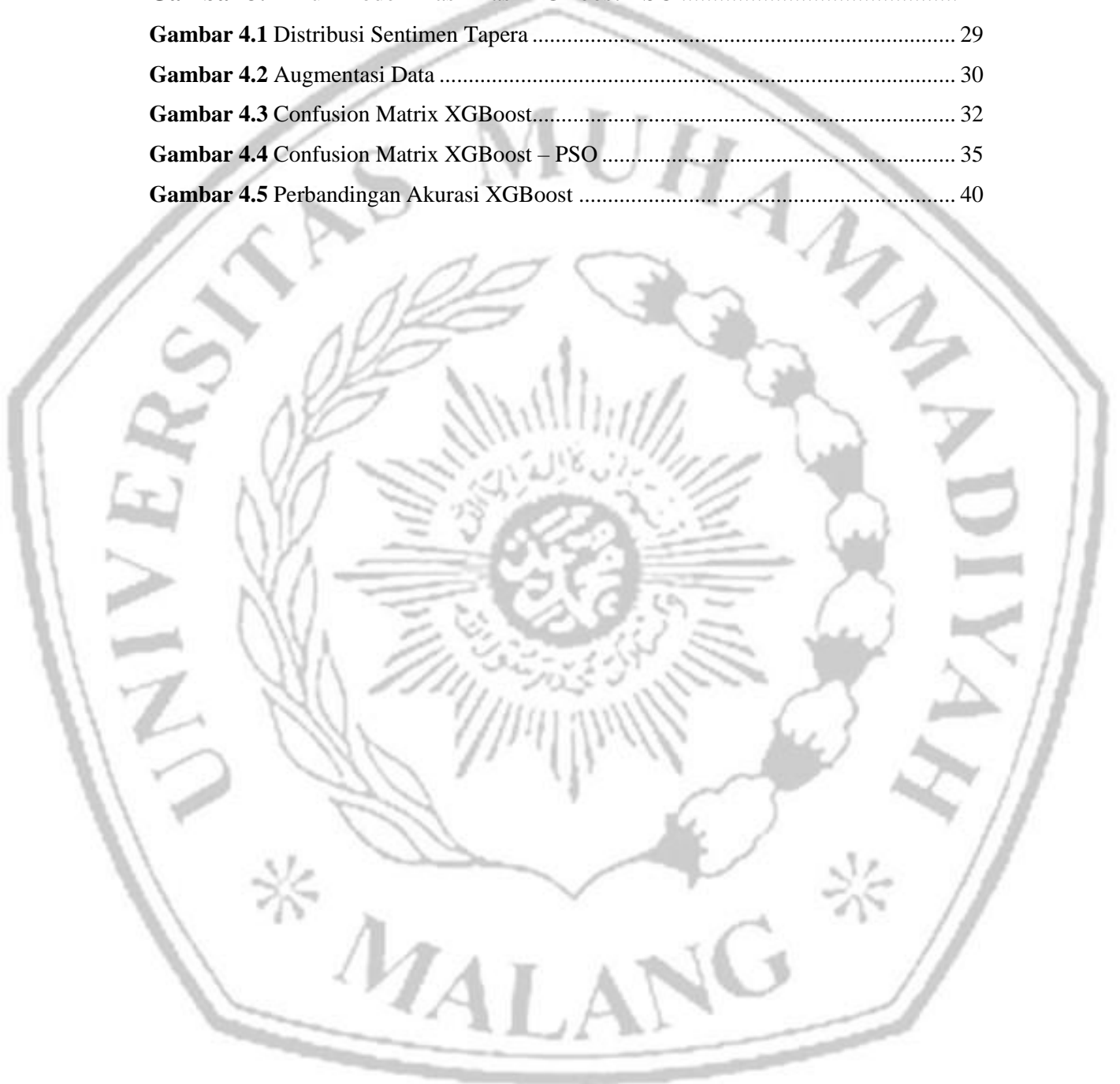
3.3 Pelabelan Data .....	13
3.5 Augmentasi Data.....	14
3.4 Preprocessing Data.....	14
3.4.1 Data Cleaning.....	14
3.4.2 Normalisasi .....	15
3.4.3 Stopword.....	16
3.4.4 Tokenization .....	16
3.4.5 Stemming .....	16
3.5 Term Frequency-Inverse Document Frequency (TF-IDF) .....	18
3.6 Training & Build Model .....	19
3.7 XGBoost berbasis PSO .....	20
3.7.1 XGBoost (Extreme Gradient Boosting).....	20
3.7.2 PSO (Particle Swarm Optimization).....	21
3.8 Evaluation .....	23
3.9 Skenario Pengujian .....	25
3.10 Lini Masa Penelitian .....	26
<b>BAB IV HASIL PENELITIAN .....</b>	<b>27</b>
4.1 Pengumpulan Data.....	27
4.2 Seleksi Data .....	27
4.3 Pelabelan Data .....	27
4.4 Preprocessing Data.....	28
4.4.1 Data Cleaning.....	28
4.4.2 Normalisasi .....	28
4.4.3 Stopword.....	28
4.4.4 Tokenization .....	29
4.4.5 Stemming .....	29
4.5 Analisis Sentimen .....	29

4.6 Pembagian Dataset.....	31
4.7 Pengujian dan Evaluasi Dataset.....	31
4.7.1 Pengujian XGboost tanpa PSO.....	31
4.7.2 Pengujian XGBoost dengan PSO.....	35
4.7.3 Hasil Evaluasi.....	39
BAB V KESIMPULAN.....	42
DAFTAR PUSTAKA.....	43



## DAFTAR GAMBAR

<b>Gambar 3.1</b> Tahap Penelitian .....	11
<b>Gambar 3.2</b> Alur Model Klasifikasi XGBoost-PSO .....	22
<b>Gambar 4.1</b> Distribusi Sentimen Tapera .....	29
<b>Gambar 4.2</b> Augmentasi Data .....	30
<b>Gambar 4.3</b> Confusion Matrix XGBoost.....	32
<b>Gambar 4.4</b> Confusion Matrix XGBoost – PSO .....	35
<b>Gambar 4.5</b> Perbandingan Akurasi XGBoost .....	40



## DAFTAR TABEL

<b>Tabel 3.1</b> Komentar tentang isu Tapera dari sosial media X .....	12
<b>Tabel 3.2</b> Tabel labeling .....	13
<b>Tabel 3.3</b> Tahap Preprocessing .....	17
<b>Tabel 3.4</b> Confusion matrix .....	24
<b>Tabel 3.5</b> Lini masa penelitian .....	26
<b>Tabel 4.1</b> Hasil Pembagian Data Latih dan Data Uji .....	31
<b>Tabel 4.2</b> Classification Report XGBoost tanpa PSO .....	35
<b>Tabel 4.3</b> Classification Report XGBoost dengan PSO .....	39
<b>Tabel 4.4</b> Mis - Klasifikasi .....	39
<b>Tabel 4.5</b> Perbandingan Hasil Pengujian .....	41

## DAFTAR PUSTAKA

- [1] V Rekha, R Raksha, Pradnya Patil, N Swaras, and G L Rajat, "Sentiment Analysis on Indian Government Schemes Using Twitter data," *Int. Conf. Data Sci. Commun. IconDSC*.
- [2] T. S. Sabrila, Y. Azhar, and C. S. K. Aditya, "Analisis Sentimen Tweet Tentang UU Cipta Kerja Menggunakan Algoritma SVM Berbasis PSO," *JISKA J. Inform. Sunan Kalijaga*, vol. 7, no. 1, pp. 10–19, Jan. 2022, doi: 10.14421/jiska.2022.7.1.10-19.
- [3] D. Dewanto, "TOWS matrix as business strategy of BP. Tapera," *Int. J. Res. Bus. Soc. Sci.* 2147- 4478, vol. 11, no. 7, pp. 62–77, Nov. 2022, doi: 10.20525/ijrbs.v11i7.2137.
- [4] R. H. Muhammadi, T. G. Laksana, and A. B. Arifa, "Combination of Support Vector Machine and Lexicon-Based Algorithm in Twitter Sentiment Analysis," *Khazanah Inform. J. Ilmu Komput. Dan Inform.*, vol. 8, no. 1, pp. 59–71, Mar. 2022, doi: 10.23917/khif.v8i1.15213.
- [5] A. Tiwari, R. Gupta, and R. Chandra, "Delhi air quality prediction using LSTM deep learning models with a focus on COVID-19 lockdown," Feb. 21, 2021, *arXiv*: arXiv:2102.10551. Accessed: Jun. 20, 2024. [Online]. Available: <http://arxiv.org/abs/2102.10551>
- [6] Iyad Lahsen Cherif and Abdesselem Kortebi, "On using eXtreme Gradient Boosting (XGBoost) Machine Learning algorithm for Home Network Traffic Classification".
- [7] B. N. S. Aldo Andika Saputra, "Penerapan Algoritma Extreme Gradient Boosting (Xgboost) Untuk Analisis Risiko Kredit," Apr. 2024, doi: 10.5281/ZENODO.10960080.
- [8] R. Rismaya, "Pelanggaran Prinsip Kerja Sama dan Kesantunan Berbahasa dalam Cuitan Twitter Bertema Internalized Sexism 'Internalisasi Seksisme': Suatu Kajian Pragmatik," *Metahumaniora*, vol. 10, no. 3, p. 346, Dec. 2020, doi: 10.24198/metahumaniora.v10i3.31032.
- [9] E. S. Eriana, S. Kom, M. Kom, D. A. Zein, and M. Kom, "ARTIFICIAL INTELLIGENCE (AI)".
- [10] V. W. D. Thomas and F. Rumaisa, "Analisis Sentimen Ulasan Hotel Bahasa Indonesia Menggunakan Support Vector Machine dan TF-IDF," *J. MEDIA Inform. BUDIDARMA*, vol. 6, no. 3, p. 1767, Jul. 2022, doi: 10.30865/mib.v6i3.4218.
- [11] S. Thongsuwan, S. Jaiyen, A. Padcharoen, and P. Agarwal, "ConvXGB: A new deep learning model for classification problems based on CNN and XGBoost," *Nucl. Eng. Technol.*, vol. 53, no. 2, pp. 522–531, Feb. 2021, doi: 10.1016/j.net.2020.04.008.
- [12] D. Agustina, Y. I. Mukti, and S. Muntari, "Integrasi Particle Swam Optimization Menggunakan K-Means untuk Klasterisasi Pengangguran di Kota Pagar Alam," *J.*

- Pustaka AI Pus. Akses Kaji. Teknol. Artif. Intell.*, vol. 3, no. 1, pp. 34–41, Apr. 2023, doi: 10.55382/jurnalpustakaai.v3i1.543.
- [13] E. P. Cynthia and E. Ismanto, “Metode Decision Tree Algoritma C.45 Dalam Mengklasifikasi Data Penjualan Bisnis Gerai Makanan Cepat Saji,” *Jurasik J. Ris. Sist. Inf. Dan Tek. Inform.*, vol. 3, p. 1, Jul. 2018, doi: 10.30645/jurasik.v3i0.60.
- [14] G. Triyono, “PERTIMBANGAN MELAKUKAN DENORMALISASI PADA MODEL BASIS DATA RELASI,” 2011.
- [15] T. W. Putra and A. Triayudi, “Analisis Sentimen Pembelajaran Daring menggunakan Metode Naïve Bayes, KNN, dan Decision Tree,” 2022.
- [16] D. T. Hermanto, A. Setyanto, and E. T. Luthfi, “Algoritma LSTM-CNN untuk Binary Klasifikasi dengan Word2vec pada Media Online,” *Creat. Inf. Technol. J.*, vol. 8, no. 1, p. 64, Mar. 2021, doi: 10.24076/citec.2021v8i1.264.
- [17] I. Prasetyaningrum, K. Fathoni, and T. T. J. Priyantoro, “Application of recommendation system with AHP method and sentiment analysis,” *TELKOMNIKA Telecommun. Comput. Electron. Control*, vol. 18, no. 3, p. 1343, Jun. 2020, doi: 10.12928/telkomnika.v18i3.14778.
- [18] J. Han, K. Shu, and Z. Wang, “Predicting energy use in construction using Extreme Gradient Boosting,” *PeerJ Comput. Sci.*, vol. 9, p. e1500, Aug. 2023, doi: 10.7717/peerj-cs.1500.
- [19] F. Nurrahman, H. Wijayanto, A. H. Wigena, and N. Nurjanah, “PRE-PROCESSING DATA ON MULTICLASS CLASSIFICATION OF ANEMIA AND IRON DEFICIENCY WITH THE XGBOOST METHOD,” *BAREKENG J. Ilmu Mat. Dan Terap.*, vol. 17, no. 2, pp. 0767–0774, Jun. 2023, doi: 10.30598/barekengvol17iss2pp0767-0774.
- [20] N. T. Tran, T. T. G. Tran, T. A. Nguyen, and M. B. Lam, “A new grid search algorithm based on XGBoost model for load forecasting,” *Bull. Electr. Eng. Inform.*, vol. 12, no. 4, pp. 1857–1866, Aug. 2023, doi: 10.11591/eei.v12i4.5016.
- [21] Maher G.M. Abdolrasola\*, , M.A. Hannanb, , S.M. Suhail Hussainc, and , Taha Selim Ustunc, *Optimal PI controller based PSO optimization for PV inverter using SPWM techniques.*
- [22] D. Normawati and S. A. Prayogi, “Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter,” vol. 5, 2021.
- [23] D. P. Javale and S. Desai, “Healthcare Critical Diagnosis Accuracy: A Proposed Machine Learning Evaluation Metric for Critical Healthcare Analysis,” vol. 14, no. 8, 2023.
- [24] M. D. Maulana, A. I. Hadiana, and F. R. Umbara, “ALGORITMA XGBOOST UNTUK KLASIFIKASI KUALITAS AIR MINUM,” vol. 7, no. 5, 2023.

- [25]I. P. A. P. Widiarta, R. Dwiyanaputra, and A. Aranta, “ANALISIS SENTIMEN MASYARAKAT TERHADAP KEBIJAKAN PENERAPAN PPKM DI MEDIA SOSIAL TWITTER DENGAN MENGGUNAKAN METODE XGBOOST,” *J. Teknol. Inf. Komput. Dan Apl. JTika*, vol. 5, no. 2, pp. 154–163, Sep. 2023, doi: 10.29303/jtika.v5i2.342.
- [26]P. Sexton Topper and J. A. Bauermeister, “Relationship Timelines, Dyadic Interviews, and Visual Representations: Implementation of an Adapted Visual Qualitative Technique,” *Int. J. Qual. Methods*, vol. 20, p. 16094069211016708, Jan. 2021, doi: 10.1177/16094069211016708.



## FORM CEK PLAGIARISME LAPORAN TUGAS AKHIR

Nama Mahasiswa : Saniyya Ruzzy Marwa  
 NIM : 202110370311177  
 Judul TA : Analisis Sentimen Media Sosial X Terhadap Isu Tapera  
 Menggunakan Metode XGBoost Berbasis PSO

### Hasil Cek Plagiarisme dengan Turnitin

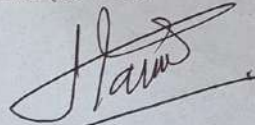
No.	Komponen Pengecekan	Nilai Maksimal Plagiarisme (%)	Hasil Cek Plagiarisme (%) *
1.	Bab 1 – Pendahuluan	10 %	4%
2.	Bab 2 – Daftar Pustaka	25 %	6%
3.	Bab 3 – Analisis dan Perancangan	25 %	2%
4.	Bab 4 – Implementasi dan Pengujian	15 %	15%
5.	Bab 5 – Kesimpulan dan Saran	5 %	3%
6.	Makalah Tugas Akhir	20%	7%

\*) Hasil cek plagiarism diisi oleh pemeriksa (staf TU)

\*) Maksimal 5 kali (4 Kali sebelum ujian, 1 kali sesudah ujian)

Mengetahui,

Pemeriksa (Staff TU)

  
 (..... Agus Muzron .....) )