

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

Saniyya Ruzzy Marwa^{*1}, Christian Sri Kusuma Aditya², Yufis Azhar³

Universitas Muhammadiyah Malang

marwasaniyya@webmail.umm.ac.id¹, christiaskaditya@umm.ac.id¹, yufis@webmail.umm.ac.id¹

Abstrak

Penelitian ini bertujuan menganalisis sentimen masyarakat terhadap isu Tabungan Perumahan Rakyat (Tapera) melalui media sosial X. Data dikumpulkan melalui crawling dengan kata kunci "Isu Tapera" dan menghasilkan 2.268 data. Data diberi label secara manual ke dalam kategori negatif, positif, netral, dan others, lalu diproses melalui tahap pra-pemrosesan dan ekstraksi fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF). Data dibagi menjadi data latih dan data uji dengan rasio 80:20. Model klasifikasi menggunakan algoritma Extreme Gradient Boosting (XGBoost) yang dioptimasi dengan Particle Swarm Optimization (PSO) untuk mendapatkan parameter terbaik. Hasil menunjukkan kombinasi XGBoost dan PSO meningkatkan akurasi model hingga 85%. Penelitian ini memberikan gambaran potensi analisis opini publik secara real-time melalui media sosial. Hasilnya dapat digunakan sebagai dasar pertimbangan kebijakan pemerintah terhadap respon masyarakat.

Kata Kunci: Media Sosial, PSO, Sentimen, Tapera, XGBoost

Abstract

This study aims to analyze public sentiment toward the Housing Savings Program (Tapera) issue through the X social media platform. Data was collected using a crawling method with the keyword "Isu Tapera," resulting in 2,268 data points. The data was manually labeled into negative, positive, neutral, and others categories, then processed through preprocessing and feature extraction using Term Frequency-Inverse Document Frequency (TF-IDF). The dataset was split into training and testing sets with an 80:20 ratio. The classification model used Extreme Gradient Boosting (XGBoost), optimized with Particle Swarm Optimization (PSO) to obtain the best parameters. The results showed that the combination of XGBoost and PSO improved the model's accuracy to 85%. This research demonstrates the potential of real-time public opinion analysis via social media. The findings can serve as a basis for government policy considerations regarding public response.

Keywords: Sosial media, PSO, Sentiment, Tapera, XGBoost

1. Pendahuluan

Media sosial telah memberikan dampak signifikan dan menjadi bagian penting dalam kehidupan sehari-hari masyarakat di era digital saat ini. Platform X, sebagai salah satu jejaring sosial yang populer, dikenal karena antarmuka yang ramah pengguna dan jangkauan globalnya. Saat ini, X memiliki sekitar 330 juta pengguna aktif bulanan, dengan pertumbuhan informasi yang sangat cepat [1]. Melalui X, pengguna dapat berkomunikasi, mendokumentasikan aktivitas, mempromosikan produk, berdiskusi, dan menyampaikan opini melalui pembaruan status singkat yang dikenal sebagai tweet, dengan batas maksimal 280 karakter. Tweet ini merepresentasikan opini publik yang dapat dimanfaatkan sebagai sumber informasi untuk analisis dan pengambilan keputusan [2].

Salah satu isu kebijakan publik yang menjadi sorotan dan menimbulkan beragam tanggapan adalah Tabungan Perumahan Rakyat (Tapera). Program ini bertujuan untuk menyediakan pembiayaan murah dalam rangka penyediaan perumahan terjangkau, terutama bagi Masyarakat Berpenghasilan Rendah (MBR). Data menunjukkan adanya penurunan backlog perumahan dari 13 juta unit pada 2015 menjadi 7,6 juta unit pada 2019, sebagai hasil upaya percepatan program perumahan [3]. Tapera bersifat wajib bagi pekerja dengan penghasilan di atas upah minimum, dan menerapkan prinsip gotong royong, di mana mereka yang tidak memenuhi syarat untuk pembiayaan tetap mendapatkan tabungan pensiun. Implementasi

kebijakan ini memicu beragam reaksi publik, mulai dari dukungan hingga penolakan, yang banyak disuarakan melalui media sosial X.

Media sosial kini berfungsi sebagai ruang diskusi publik yang dinamis, termasuk untuk mengungkapkan reaksi terhadap kebijakan pemerintah. Analisis terhadap komentar-komentar tersebut memungkinkan kita untuk memahami persepsi publik yang berkembang terhadap Tapera [4]. Untuk menggali informasi dari data tidak terstruktur seperti tweet, pendekatan berbasis Natural Language Processing (NLP) menjadi penting. NLP banyak digunakan untuk pemrosesan dan ekstraksi data dari platform media sosial, termasuk X [5].

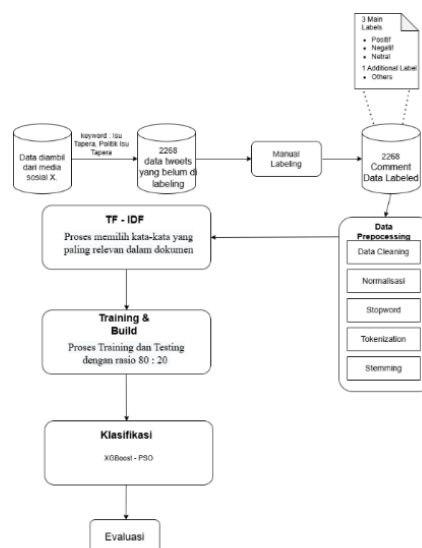
Dalam konteks klasifikasi sentimen, algoritma Extreme Gradient Boosting (XGBoost) banyak digunakan karena keunggulannya dalam menyelesaikan masalah klasifikasi dan regresi dengan pendekatan Gradient Boosting Decision Tree (GBDT) [6]. Untuk meningkatkan performa model, XGBoost dapat dioptimasi menggunakan metode seperti Particle Swarm Optimization (PSO), yang terbukti efektif dalam mengatasi keterbatasan pemilihan parameter model.

Penelitian terdahulu yang menganalisis sentimen terhadap UU Cipta Kerja menggunakan SVM berbasis PSO menunjukkan bahwa PSO mampu meningkatkan performa model secara signifikan, baik dari sisi akurasi, presisi, maupun recall [2]. Selain itu, penelitian terkait XGBoost dalam analisis risiko kredit menunjukkan bahwa algoritma ini efektif menangani ketidakseimbangan data kelas, terutama dengan dukungan teknik oversampling seperti SMOTE, dan mampu menghasilkan AUC hingga 0.918 yang tergolong "very good" [7].

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap isu Tapera melalui data dari media sosial X. Penelitian ini memanfaatkan pendekatan NLP untuk pra-pemrosesan dan ekstraksi fitur, serta menerapkan XGBoost yang dioptimasi dengan PSO guna meningkatkan akurasi klasifikasi. Hasil dari penelitian ini diharapkan dapat memberikan wawasan tentang opini publik serta mendukung pengambilan keputusan kebijakan berbasis data.

2. Metode Penelitian

Bagian ini bertujuan untuk menjelaskan secara rinci tahapan-tahapan yang akan dilaksanakan dalam proses penelitian. Setiap tahapan memiliki peran yang sangat penting dalam memastikan penelitian berjalan sesuai dengan rencana dan menghasilkan data yang valid serta relevan. Secara garis besar, alur penelitian akan dimulai dengan identifikasi masalah, diikuti dengan studi literatur untuk memahami kerangka teori yang ada. Kemudian, proses pengumpulan dan pengolahan data akan dilakukan dengan langkah-langkah yang telah direncanakan secara sistematis. Setelah itu, analisis terhadap data yang terkumpul akan dilakukan untuk menemukan temuan yang sesuai dengan tujuan penelitian. Akhirnya, kesimpulan dan saran akan disusun berdasarkan hasil analisis yang dilakukan. Semua tahapan ini digambarkan dalam Gambar 1, yang memberikan gambaran jelas mengenai urutan dan hubungan antar langkah-langkah tersebut dalam mencapai tujuan akhir penelitian.



Gambar 1. Alur Penelitian

2.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini berasal dari ulasan atau komentar berbahasa Indonesia yang ditemukan di media sosial X. Data ini dikumpulkan selama periode Juni hingga September 2024 dengan menggunakan teknik crawling melalui platform Google Colab, yang secara khusus menargetkan tweet terkait isu Tapera.

Proses crawling ini berhasil mengumpulkan sebanyak 2268 tweet yang relevan dengan topik tersebut. Data yang diperoleh akan dianalisis lebih lanjut untuk mendalami opini dan sentimen publik mengenai Tapera yang tercermin dalam komentar-komentar tersebut.

2.2 Seleksi Data

Seleksi data merupakan tahap penting dalam penelitian ini, yang bertujuan untuk memilih data yang akan digunakan dalam algoritma klasifikasi berdasarkan himpunan data target[8]. Pada penelitian ini, kolom full text dipilih sebagai fokus analisis karena kolom tersebut memuat komentar-komentar yang mencerminkan pendapat masyarakat.

Komentar-komentar yang terkandung dalam kolom tersebut memberikan wawasan penting mengenai pandangan dan opini publik, yang dapat diolah untuk mengidentifikasi pola, tren, atau sentimen yang berkembang di masyarakat. Oleh karena itu, seleksi data yang tepat sangat diperlukan untuk menghasilkan analisis yang akurat dan dapat memberikan informasi yang berguna.

2.3 Pelabelan Data

Setelah proses pengumpulan data selesai, langkah berikutnya adalah memberi label pada setiap ulasan yang telah terkumpul. Proses ini dilakukan secara manual, dengan memberikan label sesuai dengan isi dari ulasan tersebut. Ulasan yang mengandung apresiasi, pujian, atau kesan positif akan diberi label "positif", sementara yang berisi keluhan atau kritik akan diberi label "negatif".

Ulasan yang tidak dapat dikategorikan sebagai positif atau negatif, seperti yang tidak jelas atau ambigu, akan diberi label "netral". Sementara itu, ulasan yang tidak berhubungan dengan topik utama penelitian akan digolongkan ke dalam kategori "others". Proses pemberian label ini penting untuk memastikan analisis sentimen yang tepat dan terstruktur.

2.4 Preprocessing Data

Preprocessing merupakan langkah krusial dalam proses penambangan data, dengan tujuan utama untuk meningkatkan kualitas data sebelum dilakukan analisis atau pengembangan model klasifikasi. Dengan melakukan preprocessing, data yang diperoleh dapat disiapkan dengan lebih baik, memastikan bahwa analisis yang dilakukan lebih tepat dan efektif[7].

Dalam penelitian ini, beberapa langkah preprocessing yang diterapkan meliputi data cleaning, normalisasi, tokenization, serta penerapan teknik stopword removal dan stemming. Langkah-langkah ini bertujuan untuk membersihkan dan menyusun data teks agar lebih siap untuk dianalisis lebih lanjut.

2.4.1 Data Cleaning

- Missing Value*: Penanganan data hilang penting untuk menjaga kualitas analisis. Namun, pada dataset ini tidak ditemukan missing value sehingga tidak diperlukan tindakan seperti imputasi atau penghapusan data.
- Duplikat Data*: Data duplikat dapat menyebabkan bias dalam analisis. Pemeriksaan menunjukkan tidak adanya duplikasi, namun jika ditemukan, harus dihapus untuk menjaga validitas data.
- Remove Punctuation*: Simbol seperti tanda baca dan emotikon dihapus karena tidak relevan dalam analisis sentimen.
- Case Folding*: Seluruh teks diubah menjadi huruf kecil untuk menyamakan kata yang identik namun berbeda penulisan kapitalnya.

2.4.2 Normalisasi

Normalisasi adalah proses yang penting untuk menghindari pengulangan atau duplikasi teks dalam basis data. Selain itu, normalisasi juga berfungsi untuk memperbaiki teks yang mengandung ketidakwajaran atau kesalahan, sehingga menghasilkan bentuk teks yang lebih sederhana, terstruktur, dan mudah dipahami[9].

Proses normalisasi dimulai dengan identifikasi kata-kata yang perlu diperbaiki, seperti kata-kata yang disingkat, salah eja, atau tidak sesuai dengan kaidah bahasa baku, dalam file CSV. Kata-kata yang memerlukan perbaikan kemudian dipindahkan ke file Excel untuk pengelolaan lebih lanjut, di mana setiap kata yang membutuhkan koreksi dihubungkan dengan bentuk baku yang benar. Setelah proses normalisasi selesai, data tersebut dikonversi ke format JSON untuk memudahkan pemrosesan lebih lanjut dan integrasi dengan sistem lainnya.

2.4.3 Stopword

Stopword adalah tahap dalam pengolahan teks yang bertujuan untuk menghapus kata-kata umum yang tidak berkontribusi signifikan dalam analisis data [10]. Kata-kata seperti "dan", "atau", "tapi", dan lainnya biasanya dihilangkan karena dianggap tidak memiliki makna yang relevan dalam konteks analisis.

Sebagai contoh, kalimat "Masyarakat sendiri banyak yg hidup sulit apalagi sekarang sudah ada wacana tapera" akan diproses menjadi "masyarakat hidup sulit wacana tapera" setelah melewati tahap penghapusan stopwords. Proses ini membantu menyederhanakan teks sehingga analisis data dapat lebih fokus pada kata-kata yang memiliki nilai informatif.

2.4.4 Tokenization

Tokenization merupakan langkah dalam pengolahan teks yang bertujuan untuk memisahkan setiap kata dalam suatu dokumen. Pada tahap ini, setiap kata dipisahkan berdasarkan karakter spasi yang ada antar kata [11].

Proses tokenization ini memungkinkan teks yang panjang dapat dipecah menjadi unit-unit kata yang lebih kecil, sehingga mempermudah analisis lebih lanjut. Setiap token yang dihasilkan dapat dianalisis secara individu untuk memperoleh informasi yang lebih terperinci.

2.4.5 Stemming

Stemming adalah metode yang digunakan untuk menghapus imbuhan dari sebuah kata dan mengembalikannya ke bentuk dasarnya [12]. Tujuan utama dari metode ini adalah untuk mengurangi variasi kata yang muncul akibat penggunaan imbuhan, sehingga kata tersebut dapat dianalisis dalam bentuk dasarnya.

Proses stemming dimulai dengan mengidentifikasi kata-kata berimbuhan atau yang merupakan bentuk turunan dalam file CSV. Kata-kata tersebut kemudian dipindahkan ke dalam file Excel untuk pengelompokan dan pencocokan dengan bentuk dasar yang sesuai. Setelah proses stemming selesai, data tersebut diubah menjadi format JSON untuk mempermudah pemrosesan lebih lanjut.

2.5 Term Frequency-Inverse Document Frequency (TF – IDF)

Setelah tahap *preprocessing*, data diproses melalui *term weighting* untuk memberikan bobot pada setiap kata. Penelitian ini menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF), yang umum digunakan dalam *text mining*. TF menunjukkan frekuensi kemunculan kata dalam dokumen, sedangkan IDF mengukur seberapa jarang kata tersebut muncul di seluruh dokumen. Semakin sering kata muncul dalam dokumen tertentu namun jarang di dokumen lain, semakin tinggi bobotnya [13].

$$TF - IDF(t, d, D) = TF(t, d) \times IDF(t, D) \quad (1)$$

TF-IDF digunakan untuk menilai seberapa penting suatu kata (*term*) dalam sebuah dokumen (*d*) terhadap seluruh kumpulan dokumen (*D*). Nilai TF(*t*,*d*) mengukur frekuensi kemunculan kata *t* dalam dokumen *d*, sedangkan IDF(*t*,*D*) mengukur seberapa jarang kata *t* muncul di seluruh dokumen, sehingga memberikan bobot lebih tinggi pada kata-kata yang bersifat khas dan informatif.

2.6 Training & Build Model

Dataset dibagi menjadi data latih dan uji dengan rasio 80:20. Setelah itu, dilakukan *preprocessing* seperti pembersihan dan pemformatan data. Model XGBoost kemudian dilatih menggunakan data latih dengan membangun beberapa pohon keputusan secara bertahap untuk meningkatkan akurasi. Setelah pelatihan, model diuji pada data uji untuk mengevaluasi

kinerjanya menggunakan metrik seperti akurasi dan F1-score. Jika performanya baik, model siap digunakan untuk prediksi data baru.

2.7 Klasifikasi

2.7.1 XGBoost (Extreme Gradient Boosting)

XGBoost adalah algoritma berbasis pohon keputusan yang termasuk dalam metode Gradient Boosting dan dapat digunakan untuk klasifikasi maupun regresi [13]. Pada pohon regresi, node internal mengevaluasi atribut, sedangkan node daun memberikan skor keputusan [14]. Fungsi objektif XGBoost dirumuskan sebagai :

$$obj(\theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \omega(f_k) \quad (2)$$

Di mana fungsi pertama adalah loss function dan yang kedua adalah regulasi kompleksitas model. Sebelum pelatihan, hyperparameter XGBoost perlu dioptimasi, yang terbagi menjadi parameter umum, booster, dan learning task [15].

2.7.2 PSO (Particle Swarm Optimization)

PSO merupakan metode optimasi yang sederhana, tahan terhadap noise, dan memiliki kemampuan konvergensi yang baik [16]. Persamaan kecepatannya:

$$v_i(t+1) = w \cdot v_i(t) + c_1 \cdot r_1 \cdot (pbest_i - x_i(t)) + c_2 \cdot r_2 \cdot (gbest_i - x_i(t)) \quad (3)$$

Menggabungkan pengalaman individu (pbest) dan kolektif (gbest), serta komponen acak untuk eksplorasi. Dalam konteks XGBoost, setiap partikel mewakili kombinasi parameter yang berbeda. Nilai akurasi dari pelatihan digunakan sebagai fungsi fitness. Partikel terus diperbarui hingga iterasi maksimum tercapai, dan parameter terbaik digunakan untuk melatih XGBoost, dengan tujuan meningkatkan akurasi klasifikasi secara optimal.

2.8 Evaluasi

Confusion matrix adalah tabel yang menunjukkan jumlah prediksi benar (True Positive) dan salah (False Positive) dari data uji untuk masing-masing kelas [17]. Contoh struktur confusion matrix untuk tiga kelas disajikan pada Tabel 1.

Tabel 1. Confusion Matrix

	Predicted Positif	Predicted Negatif	Predicted Netral
Actual Positif	TP(Positif)	FP (Positif>Negatif)	FP (Positif>Netral)
Actual Negatif	FP (Negatif > Positif)	TP (Negatif)	FP (Negatif>Netral)
Actual Netral	FP (Netral>Positif)	FP (Netral>Negatif)	TP (Netral)

TP menunjukkan jumlah prediksi yang benar untuk suatu kelas, sedangkan FP adalah jumlah prediksi salah terhadap kelas yang sebenarnya. Evaluasi dilakukan untuk membandingkan model dengan dan tanpa seleksi fitur RFE menggunakan *classification report*, yang mencakup metrik-metrik berikut [18].

$$Akurasi = \frac{TP + TN}{TP + TN + FN} \times 100\% \quad (4)$$

$$Presisi = \frac{TP}{TP + FP} \times 100\% \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (6)$$

$$F1 - Score = 2 \times \frac{TP}{TP + FP + FN} \times 100 \% \quad (7)$$

3. Hasil Penelitian dan Pembahasan

Pada bab ini akan dipaparkan hasil dari penelitian yang telah dilaksanakan. Tahap pengujian penelitian menerapkan bahasa pemrograman *python* untuk proses implementasinya.

3.1 Dataset

Dataset yang digunakan dalam penelitian ini terdiri dari tweet berbahasa Indonesia yang mengangkat isu Tapera. Pengumpulan data dilakukan melalui proses crawling menggunakan Twitter API dan difokuskan pada media sosial X. Total data yang dikumpulkan berjumlah 2267 tweet. Dari hasil pelabelan, sebanyak 1616 data masuk ke dalam kategori negatif, 449 data ke dalam kategori positif, 172 data netral, dan 30 data masuk ke dalam kategori lain. Proses pelabelan dilakukan secara manual oleh tiga orang anotator yang merupakan ahli bahasa Indonesia, guna memastikan interpretasi yang tepat terhadap makna setiap tweet agar dapat diklasifikasikan ke dalam kategori sentimen yang sesuai.

3.2 Hasil Klasifikasi

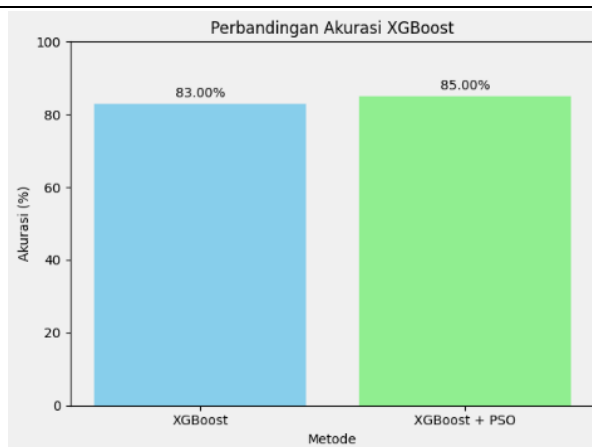
Model klasifikasi awal dibangun menggunakan algoritma XGBoost tanpa proses seleksi fitur, di mana hasil evaluasi menunjukkan performa yang cukup baik namun masih memiliki ruang untuk peningkatan. Metrik evaluasi seperti akurasi, presisi, recall, dan F1-score menjadi acuan untuk menilai kinerja model tersebut. Untuk meningkatkan efektivitas model, dilakukan optimasi dengan metode Particle Swarm Optimization (PSO) yang berfungsi untuk memilih fitur-fitur paling relevan dalam data latih. Strategi ini bertujuan mengurangi dimensi data tanpa kehilangan informasi penting, sehingga model dapat bekerja lebih efisien dan akurat. Setelah penerapan PSO, dilakukan pelatihan ulang model XGBoost dengan subset fitur terpilih, dan hasil perbandingannya dapat dilihat pada Tabel 2.

Tabel 2. Perbandingan Hasil Pengujian

	Pengujian Pertama	Pengujian Kedua	Keterangan	
	XGBoost	XGBoost+PSO	Peningkatan Nilai	Hasil
Acc	83%	85%	2%	Meningkat
Preciss	64%	81%	17%	Meningkat
Recall	56%	63%	7%	Meningkat
F1-Score	59%	66%	7%	Meningkat

Pada Tabel 2 menyajikan perbandingan kinerja model XGBoost sebelum dan sesudah penerapan seleksi fitur menggunakan Particle Swarm Optimization (PSO). Berdasarkan hasil pengujian, terlihat bahwa penerapan PSO memberikan dampak positif terhadap performa model. Akurasi mengalami peningkatan dari 83% menjadi 85%, menunjukkan bahwa model menjadi lebih tepat dalam melakukan klasifikasi. Hal ini juga tercermin pada peningkatan presisi dari 64% ke 81%, yang mengindikasikan bahwa model lebih jarang melakukan kesalahan dalam memprediksi kelas positif.

Selain itu, nilai recall juga mengalami kenaikan sebesar 7%, dari 56% menjadi 63%, yang berarti model mampu mengenali lebih banyak sampel dari kelas yang sebenarnya relevan. Skor F1, sebagai metrik gabungan antara presisi dan recall, juga menunjukkan peningkatan dari 59% ke 66%. Secara keseluruhan, hasil ini memperkuat efektivitas metode PSO dalam membantu proses seleksi fitur, sehingga model yang dihasilkan lebih efisien dan akurat. Hal ini membuktikan bahwa optimalisasi fitur sebelum pelatihan model dapat meningkatkan kualitas hasil klasifikasi secara signifikan.



Gambar 2. Perbandingan hasil akurasi XGBoost dan XGBoost - PSO

Gambar 2 Perbandingan hasil akurasi XGBoost dan XGBoost-PSO menunjukkan perbandingan tingkat akurasi antara model XGBoost tanpa seleksi fitur dan model XGBoost yang telah dioptimasi menggunakan metode Particle Swarm Optimization (PSO). Berdasarkan grafik, tampak bahwa penerapan PSO mampu meningkatkan akurasi model sebesar 2%, dari semula 83% menjadi 85%. Kenaikan ini menunjukkan bahwa proses seleksi fitur berhasil membantu model fokus pada atribut-atribut yang paling berpengaruh dalam proses klasifikasi. Dengan jumlah fitur yang lebih terarah, model dapat menghindari noise dari fitur yang kurang relevan sehingga menghasilkan prediksi yang lebih tepat. Hasil ini memperkuat peran penting optimasi fitur dalam meningkatkan kinerja algoritma pembelajaran mesin.

4. Kesimpulan

Penelitian ini berhasil melakukan analisis sentimen terhadap isu Tabungan Perumahan Rakyat (Tapera) di media sosial Twitter dengan menggunakan model XGBoost yang dioptimasi oleh algoritma Particle Swarm Optimization (PSO). Hasil analisis menunjukkan bahwa sebagian besar sentimen masyarakat terhadap isu Tapera cenderung bersifat negatif. Dengan penerapan PSO, akurasi model XGBoost meningkat dari 83% menjadi 85%, yang menunjukkan bahwa optimasi parameter sangat berpengaruh terhadap performa model klasifikasi. Hal ini membuktikan bahwa kombinasi XGBoost dan PSO efektif digunakan dalam tugas analisis sentimen, khususnya untuk isu-isu sosial yang berkembang di media sosial. Penelitian ini juga dapat menjadi landasan awal bagi pengambilan keputusan oleh pemangku kebijakan dalam merespons opini publik secara cepat dan tepat.

Referensi

- [1] V Rekha, R Raksha, Pradnya Patil, N Swaras, and G L Rajat, "Sentiment Analysis on Indian Government Schemes Using Twitter data," *International Conference on Data Science and Communication (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*, 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," *IJRBS*, 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," *khif*, 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] E. P. Cynthia and E. Ismanto, "Metode Decision Tree Algoritma C.45 Dalam Mengklasifikasi Data Penjualan Bisnis Gerai Makanan Cepat Saji," *jurasic*, vol. 3, p. 1, Jul. 2018, doi: 10.30645/jurasic.v3i0.60.
- [9] G. Triyono, "Pertimbangan Melakukan Denormalisasi Pada Model Basis Data Relasi," 2011.
- [10] T. W. Putra and A. Triayudi, "Analisis Sentimen Pembelajaran Daring menggunakan Metode Naïve Bayes, KNN, dan Decision Tree," 2022.
- [11] D. T. Hermanto, A. Setyanto, and E. T. Luthfi, "Algoritma LSTM-CNN untuk Binary Klasifikasi dengan Word2vec pada Media Online," *citec*, vol. 8, no. 1, p. 64, Mar. 2021, doi: 10.24076/citec.2021v8i1.264.
- [12] I. Prasetyaningrum, K. Fathoni, and T. T. J. Priyantoro, "Application of recommendation system with AHP method and sentiment analysis," *TELKOMNIKA*, vol. 18, no. 3, p. 1343, Jun. 2020, doi: 10.12928/telkomnika.v18i3.14778.
- [13] J. Han, K. Shu, and Z. Wang, "Predicting energy use in construction using Extreme Gradient Boosting," *PeerJ Computer Science*, vol. 9, p. e1500, Aug. 2023, doi: 10.7717/peerj-cs.1500.
- [14] 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. Math. & App.*, vol. 17, no. 2, pp. 0767–0774, Jun. 2023, doi: 10.30598/barekengvol17iss2pp0767-0774.
- [15] 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," *Bulletin EEI*, vol. 12, no. 4, pp. 1857–1866, Aug. 2023, doi: 10.11591/eei.v12i4.5016.
- [16] 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*.
- [17] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," vol. 5, 2021.
- [18] 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.