

Penerapan Algoritma Machine Learning Untuk Memprediksi Sentimen Kepuasan Pengguna Weverse

Oleh:

Amalia Rossa Annjani

Uce Indahyanti

Progam Studi Informatika

Universitas Muhammadiyah Sidoarjo

Januari, 2026



Pendahuluan

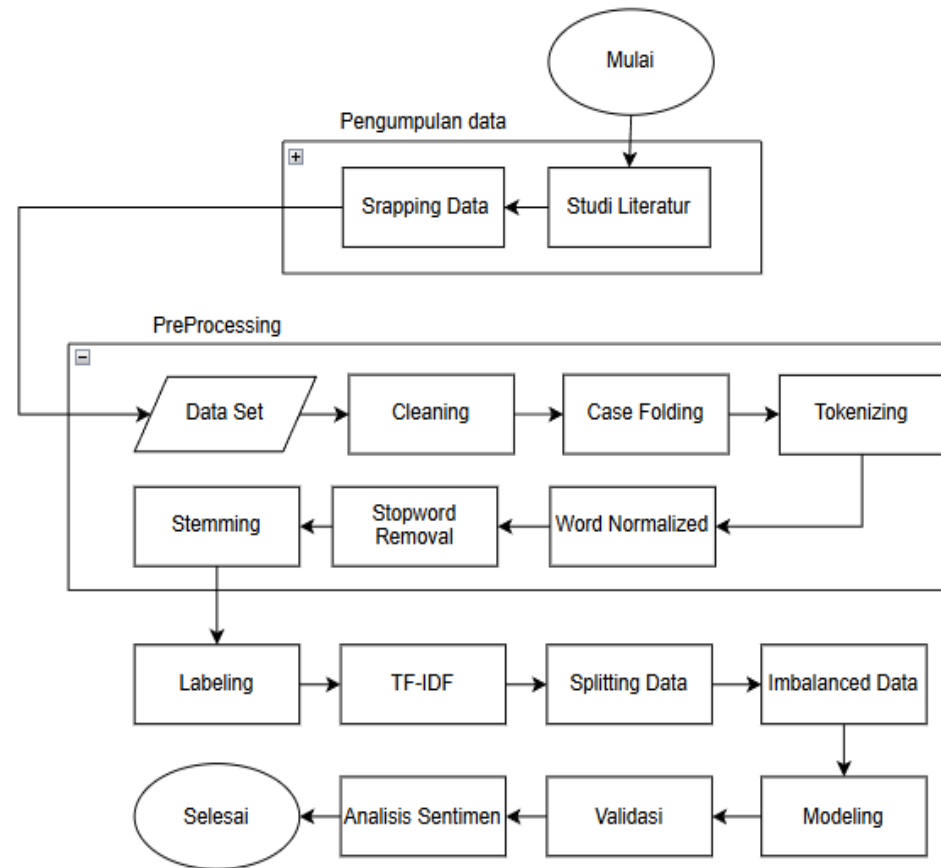
Weverse merupakan platform komunitas penggemar yang dikembangkan oleh HYBE Corporation dan mengalami peningkatan pengguna global sebesar 19% pada tahun 2024 dengan total unduhan lebih dari 150 juta. Tingginya jumlah pengguna menghasilkan volume ulasan yang mencerminkan tingkat kepuasan terhadap berbagai aspek aplikasi, namun ulasan tersebut bersifat tidak terstruktur sehingga memerlukan analisis sentimen. Meskipun analisis sentimen telah banyak diterapkan pada berbagai platform digital, penelitian pada Weverse masih terbatas, khususnya yang membandingkan algoritma Decision Tree dan Naïve Bayes. Oleh karena itu, penelitian ini bertujuan membandingkan kinerja kedua algoritma dalam mengklasifikasikan sentimen ulasan pengguna Weverse menggunakan TF-IDF, Random Oversampling, dan evaluasi confusion matrix.

Pertanyaan Penelitian (Rumusan Masalah)

1. Bagaimana tingkat kepuasan pengguna aplikasi Weverse dapat diklasifikasikan menggunakan algoritma Decision Tree dan Naive Bayes berdasarkan ulasan mereka di Google Play Store?
2. Bagaimana perbandingan antara algoritma Decision Tree dan Naive Bayes dalam hal kinerja saat mengklasifikasikan tingkat kepuasan pengguna untuk aplikasi Weverse, yang dievaluasi melalui confusion matrix dan tingkat akurasi?

Metode

Rancangan Penelitian



Metode

1. Tujuan : Memprediksi sentimen kepuasan pengguna Weverse
2. Data : 10.000 ulasan Google Play Store (Januari 2020 – Januari 2025)
3. Preprocessing : cleaning, Case Folding, Tokenizing. Normalisasi kata, stopword removal, stemming (Sastrawi)
4. Labelling : rating 1–2 (Negatif), 4–5 (Positif), rating 3 (kamus leksikon)
5. Ekstraksi fitur : TF-IDF
6. Algoritma : Decision Tree & Naïve Bayes
7. Split data : 80:20 dan 70:30
8. Imbalanced data : Random Oversampling (ROS)
9. Evaluasi : Akurasi & Confusion Matrix (Precision, Recall, F1-Score)

Hasil

Pengujian model dilakukan terhadap delapan skenario berbeda, mencakup perbandingan dua algoritma klasifikasi (Decision Tree dan Naive Bayes), dua rasio pembagian data (80:20 dan 70:30), serta dua kondisi data latih (Balance dan Imbalance). Tujuan pengujian adalah membandingkan kinerja kedua algoritma sekaligus mengevaluasi pengaruh penyeimbangan data menggunakan Random Oversampling (ROS) terhadap performa model, terutama kemampuan mengenali kelas minoritas

Algoritma	Rasio	Kondisi Train	Akurasi	Precision	Recall	F1-Score
Decision Tree	70:30	Imbalance	8.384%	8.365%	8.384%	8.374%
		Balance (ROS)	8.280%	8.408%	8.280%	8.373%
Naive Bayes		Imbalance	8.651%	8.499%	8.651%	8.295%
		Balance (ROS)	8.043%	8.698%	8.043%	8.247%
Decision Tree	80:20	Imbalance	8.547%	8.536%	8.547%	8.541%
		Balance (ROS)	8.414%	8.536%	8.414%	8.467%
Naive Bayes		8.681%	8.510%	8.681%	8.388%	8.681%
		8.110%	8.803%	8.110%	8.315%	8.110%

Hasil

Pengujian penerapan ROS meningkat pada hasil recall negatif

Algoritma	Split Rasio	Kondisi Train	Recall	F1-Score
Decision Tree	80:20	Imbalance	51.69%	52.20%
		Balance (ROS)	57.49%	52.65%
	70:30	Imbalance	45.98%	46.66%
		Balance (ROS)	53.05%	48.67%
Naïve Bayes	80:20	Imbalance	23.67%	35.51%
		Balance (ROS)	81.64%	57.00%
	70:30	Imbalance	18.97%	30.18%
		Balance (ROS)	76.85%	54.69%

Pembahasan

Hasil penelitian menunjukkan bahwa Naïve Bayes dengan rasio split 80:20 pada kondisi data tidak seimbang menghasilkan akurasi tertinggi, namun kurang optimal dalam mendeteksi sentimen negatif karena bias terhadap kelas mayoritas. Penerapan Random Oversampling (ROS) meningkatkan recall pada kelas negatif secara signifikan, meskipun menurunkan akurasi keseluruhan. Temuan ini menunjukkan adanya trade-off antara akurasi global dan kemampuan deteksi kelas minoritas, sehingga ROS menghasilkan model yang lebih seimbang untuk analisis sentimen ulasan pengguna Weverse.

Temuan Penting Penelitian

1. Naïve Bayes dengan rasio split 80:20 pada kondisi imbalance menghasilkan akurasi tertinggi.
2. Model tanpa penyeimbangan data cenderung bias terhadap kelas positif dan kurang optimal mendeteksi sentimen negatif.
3. Penerapan Random Oversampling (ROS) secara signifikan meningkatkan recall kelas negatif.
4. Terdapat trade-off antara akurasi keseluruhan dan kemampuan deteksi kelas minoritas.
5. ROS menghasilkan model yang lebih seimbang dan relevan untuk evaluasi kepuasan pengguna Weverse.

Manfaat Penelitian

Penelitian ini memberikan gambaran mengenai sentimen kepuasan pengguna terhadap aplikasi Weverse berdasarkan ulasan Google Play Store serta menjadi bahan evaluasi bagi pengembang dalam meningkatkan kualitas layanan aplikasi. Selain itu, penelitian ini menunjukkan perbandingan kinerja algoritma Decision Tree dan Naïve Bayes dalam analisis sentimen, khususnya pada kondisi data tidak seimbang, sehingga dapat menjadi referensi metodologis dan acuan bagi penelitian selanjutnya pada platform komunitas digital.

Referensi

1. A. Z. Tofani, "WWeverse Sebagai Sarana Komunikasi Fans Dengan Idol(Studi Pada Interaksi Seventeen Dan Carat)," vol. 1, pp. 349–357, 2023.
2. E. F. Laili et al., "Komparasi Algoritma Decision Tree Dan Support Vector Machine (Svm) Dalam Klasifikasi Serangan Jantung," J.Sist. Inf. dan Inform., vol. 8, no. 1, pp. 67–76, 2025.
3. D. N. I. Huda, C. Prianto, and R. M. Awangga, "Dellavianti Nishfi Ilmiah Huda Analisis Sentimen Perbandingan Layanan Jasa Pengiriman Kurir Pada Ulasan Play Store Menggunakan Metode Random Forest dan Descision Tree," J. Ilm. Inform., vol. 11, no. 2, pp. 150–158, 2023.
4. Z. Abidin, T. Suratno, and M. F. Putri, "PENERAPAN RANDOM OVERSAMPLING DAN PRINCIPAL COMPONENT ANALYSIS UNTUK MENINGKATKAN AKURASI PREDIKSI KEBANGKRUTAN PERUSAHAAN DI INDONESIA DENGAN MODEL MACHINE LEARNING," vol. 12, no. 5, 2025.
5. M. I. Abas, R. Lamusu, W. E. Pranata, S. Syahrial, and I. Ibrahim, "Penerapan Algoritma Naive Bayes Untuk Sistem Klasifikasi Status Gizi Bayi Balita," vol. 5, no. 5, pp. 1249–1260, 2025
6. B. Franko, N. Wilyanto, H. Irsyad, U. Multi, and D. Palembang, "Analisis Sentimen Terhadap Naturalisasi Pemain pada Youtube Menggunakan Decision Tree dan Naive Bayes Sentiment Analysis of Player Naturalization on Youtube Using Decision Trees and Naive Bayes," vol. 03, no. September, pp. 8–16, 2024.

Referensi

7. A. Syahrul, A. I. Purnamasari, and I. Ali, "Analisis Sentimen Twitter Terhadap Cryptocurrency P (C | X) =," JATI (Jurnal Mhs. Tek.Inform., vol. 8, no. 2, pp. 1–8, 2024.
8. M. I. Prayugah, U. Indahyanti, and N. Ariyanti, "ANALISIS SENTIMEN PUBLIK PADA PEMERINTAH DALAM SERANGAN RANSOMWARE DENGAN PENDEKATAN SMOTE," vol. 8, no. 2, pp. 333–343, 2024.
9. N. H. Nanda Dwi Kurniawan, Praditya Rendi Ferdian, "Analisis Sentimen Algoritma Naïve Bayes, Support Vector Machine, dan Random Forest Pada Ulasan Aplikasi Ajaib," J. Nas. Teknol. dan Sist. Inf., vol. 01, pp. 87–97, 2025.
10. S. Ramadhani, Taufik Hidayat, "ANALISIS SENTIMEN TWEET MASYARAKAT PADA ISU INDONESIA GELAP DI MEDIA SOSIAL X MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN METODE COUNTVECTORIZER," JATI (Jurnal Mhs. Tek.Inform., vol. 9, no. 6, pp. 9664–9670, 2025.
11. G. Kanugrahan, V. Hafizh, C. Putra, and Y. Ramdhani, "Analisis Sentimen Aplikasi Gojek Menggunakan SVM , Random Forest dan Decision Tree," vol. 6, no. 2, 2024.

Referensi

12. S. A. Lubis, Rahma, "Klasifikasi Sentiment Analysis Terhadap Usulan KB Vasektomi Syarat Penerima BANSOS dengan Metode Naive Bayes," JUNRAL KOMPUTER SISTEM INFORMASI KOMPUTER, 2025. [Online]. Available: <https://ejurnal.lkpkaryaprima.id/index.php/juktisi/article/view/631/408>.
13. A. S. Ngabdul Basidt, Eko Supriyadi, "Perbandingan Algoritma Klasifikasi dalam Analisis Sentimen Opini Masyarakat tentang Kenaikan Harga Bbm," Joined J. (Journal Informatics Educ., vol. 6, no. 2, p. 219, 2024.
14. L. Rhomaningtias, A. Khairunisa, S. Shella, M. Wara, and K. M. Hindrayani, "ANALISIS SENTIMEN ULASAN APLIKASI SMILE INDONESIA MENGGUNAKAN METODE NAIVE BAYES DAN SUPPORT VECTOR MACHINE (SVM)," HOAQ J. Teknol. Inf., vol. 16, no. c, pp. 79–91, 2025.
15. F. F. T. Andy Kho, "Analisis Sentimen Pengguna Aplikasi STEAM dengan Algoritma Naive Bayes," J. Multidiscip. Res. Dev., vol. 7, no. 6, pp. 4620–4627, 2025.
16. T. E. Puput Amelia Effendi, "ANALISIS SENTIMEN ULASAN APLIKASI GAMEHAY DAY MENGGUNAKAN ALGORITMA RANDOM FOREST," J. Inform. dan Tek. elektro Terap., vol. 13, no. 3, pp. 1–8, 2025.
17. D. N. Febianty and M. Rahardi, "A Sentiment Analysis of Public Perception Toward Pets in Public Spaces Using Logistic Regression and Word Embedding," J. Appl. Informatics Comput., vol. 9, no. 4, pp. 1846–1851, 2025.

Referensi

18. N. A. V. Andi Nur Halim, Rudiman Rudiman, "Analisis Sentimen Opini Publik Terhadap Peristiwa Bitcoin Halving Pada Data Teks Twitter Menggunakan Metode Naïve Bayes Dan Pembobotan Fitur TF-IDF," vol. 4, no. 3, pp. 2823–2831, 2025.
19. W. M. B. Najib Nurdiansyah, Farhan Sulis Febriyan, Zanuar Gesit Dian Amanta, Dicky Arya Saputra, "Mental Health Analysis to Prevent Mental Disorders in Students Using The K-Nearest Neighbor (K-NN) Algorithm and Random Forest Algorithm Analisis Kesehatan Mental untuk Mencegah Gangguan Mental pada Mahasiswa Menggunakan Algoritma K-Nearest Neighbor (K-NN) dan," Institut Ris. dan Publ. Indones. Indones. J. Mach. Learn. Comput. Sci., vol. 5, no. January, pp. 1–9, 2025.
20. D. S. Wijaya and D. Widyaningrum, "Komparasi metode algoritma klasifikasi pada analisis sentimen komentar cyberbullying di instagram," vol. 7, pp. 466–472, 2024.
21. H. P. Jelita and W. Saad, Muhammad Ibnu, "Penerapan Algoritma Naïve Bayes Dalam Analisis Sentimen Masyarakat Terhadap STMIK Widya Cipta Dharma," Bull. Inf. Technol., vol. 6, no. 2, pp. 148–160, 2025.

