IRMEX Journal of Artificial Intelligence & Data Science
Vol. 1, No. 1, Mei 2025, pp. 13-18
DOI: -
E-ISSN 3090-8507

Research Article Open Access (CC–BY-SA)

Klasifikasi Customer Churn pada Data Pelanggan Telekomunikasi Menggunakan Metode Voting Classifier

Customer Churn Classification on Telecommunication Customer Data Using the Voting Classifier Method

Rachmat Rakes[1]; Syahrul Usman[2]; Firman Aziz[3,*]; Rahmat Fuadi Syam[4]; Muhammad Nur Arafah[5]
[1,2,3,4,5] *Universitas Pancasakti, Makassar 90121, Indonesia*
[1] *rakesrachmat@gmail.com;* [2] *syahrul.usman@unpacti.ac.id;* [3] *firman.aziz@unpacti.ac.id;* [4] *rahmat@unpacti.ac.id;* [5] *mnurarafah18@gmail.com;*
** Corresponding author*

Abstrak

Customer churn merupakan tantangan besar dalam industri telekomunikasi karena berdampak langsung pada penurunan pendapatan. Penelitian ini bertujuan mengembangkan model prediksi churn pelanggan menggunakan metode Voting Classifier yang menggabungkan algoritma Naive Bayes, Random Forest, dan Logistic Regression. Data yang digunakan adalah dataset pelanggan telekomunikasi publik yang terdiri dari 7043 entri dengan 20 atribut. Proses pelatihan dilakukan dengan pembagian data latih dan uji pada beberapa rasio partisi serta penyeimbangan kelas menggunakan metode SMOTE. Evaluasi kinerja model dilakukan dengan metrik klasifikasi seperti accuracy, precision, recall, dan F1-score. Hasil menunjukkan bahwa pendekatan Voting Classifier menghasilkan performa lebih unggul dibandingkan model individu, dengan akurasi mencapai 88,8%, precision 93,61%, recall 86,27%, dan F1-score 89,79%. Pendekatan ini berpotensi menjadi dasar pengambilan keputusan dalam strategi retensi pelanggan di industri telekomunikasi.

Kata Kunci: Klasifikasi, Customer Churn, Machine Learning, Voting Classifier, Pelanggan Telekomunikasi.

Abstract

Customer churn is a major challenge in the telecommunications industry due to its direct impact on revenue decline. This study aims to develop a customer churn prediction model using the Voting Classifier method, which combines Naive Bayes, Random Forest, and Logistic Regression algorithms. The dataset used is a public telecommunication customer dataset consisting of 7,043 entries with 20 attributes. The training process involved various train-test split ratios and class balancing using the SMOTE method. Model performance was evaluated using classification metrics such as accuracy, precision, recall, and F1-score. The results show that the Voting Classifier approach outperforms individual models, achieving an accuracy of 88.8%, precision of 93.61%, recall of 86.27%, and F1-score of 89.79%. This approach has the potential to support decision-making in customer retention strategies within the telecommunications sector.

Keywords: *Classification, Customer Churn, Machine Learning, Voting Classifier, Telecommunications Customers.*

Pendahuluan

Customer churn atau berhentinya pelanggan dari layanan merupakan masalah utama yang dihadapi oleh perusahaan di sektor industri yang kompetitif, terutama dalam industri telekomunikasi[1]. Dalam lingkungan bisnis yang dinamis, pelanggan memiliki banyak pilihan layanan yang serupa, sehingga perusahaan harus mampu memahami dan mengantisipasi perilaku pelanggan yang berpotensi untuk meninggalkan layanan mereka. merancang sistem prediksi churn pelanggan yang memanfaatkan proses data mining [2]. Sistem yang dihasilkan dapat melakukan integrasi data, pembersihan data, transformasi data, sampling dan pemisahan data, konstruksi model prediksi, memprediksi churn pelanggan dan menampilkan hasil prediksi dalam format laporan tertentu yang diperlukan[3]. Menurut M.Jolfoo [4], churn pelanggan dapat menyebabkan penurunan signifikan dalam pendapatan, karena biaya untuk mempertahankan pelanggan jauh lebih rendah dibandingkan biaya untuk mendapatkan pelanggan baru. Dalam menghadapi permasalahan ini, pendekatan berbasis data seperti machine learning mulai banyak diadopsi untuk membangun sistem prediktif yang mampu mengidentifikasi pelanggan yang berpotensi churn. Algoritma klasifikasi seperti Logistic Regression, Naive Bayes, dan Random Forest telah digunakan secara luas untuk membangun model prediksi churn [5]. Namun, masing-masing algoritma memiliki kelebihan dan kelemahan tersendiri dalam hal akurasi dan stabilitas prediksi. Untuk mengatasi keterbatasan tersebut, pendekatan ensemble learning seperti Voting Classifier menjadi solusi yang menjanjikan[6]. Voting Classifier menggabungkan prediksi dari beberapa model dasar untuk

menghasilkan keputusan akhir yang lebih akurat dan robust [7]. Beberapa studi telah menunjukkan bahwa metode ensemble mampu meningkatkan performa klasifikasi dibandingkan penggunaan model tunggal[8]. Dalam studi yang dilakukan oleh P.Boozary [9], penggunaan soft voting classifier dengan kombinasi Random Forest dan Naive Bayes menghasilkan peningkatan akurasi dalam prediksi churn pada industri telekomunikasi. Sementara itu, penelitian oleh Alfarez et al. (2024) di Indonesia juga membuktikan bahwa integrasi beberapa model klasifikasi dapat meningkatkan keandalan sistem prediktif, terutama dalam mengatasi permasalahan ketidakseimbangan data [10]. Ketersediaan dataset publik seperti yang disediakan oleh Kaggle atau UCI Machine Learning Repository memungkinkan peneliti untuk mengembangkan dan mengevaluasi model prediksi churn secara lebih luas dan replikatif. Melalui pemanfaatan data historis pelanggan yang mencakup perilaku, interaksi, dan profil demografis, model klasifikasi dapat dilatih untuk mengenali pola-pola penting yang menjadi indikator risiko churn [11]. Oleh karena itu, penelitian ini bertujuan untuk membangun model klasifikasi churn pelanggan pada data pelanggan telekomunikasi menggunakan algoritma Voting Classifier. Model ini diharapkan mampu memberikan hasil prediksi yang akurat serta mendukung perusahaan dalam merancang strategi retensi pelanggan yang lebih efektif.

Metode

A. Pengumpulan Data

Pengumpulan data dilakukan dengan memanfaatkan dataset yang tersedia di platform publik seperti Kaggle atau UCI Machine Learning Repository. Dataset yang digunakan biasanya telah terstruktur dan siap diolah. Selain itu, jika memungkinkan, data pelanggan telekomunikasi yang diperoleh dari perusahaan dapat digunakan untuk menambah keberagaman data. Proses pengumpulan data dilakukan melalui teknik crawling atau dengan mengakses data yang disediakan dalam format file CSV atau database yang relevan. Data yang digunakan dalam penelitian ini adalah data Benchmark sebanyak 7043 data yang terdiri dari 19 variabel dengan deskripsi sebagai berikut: Sumber data primer dalam penelitian ini adalah data Qualitative dan Quantitative (numerik) pelanggan telekomunikasi. Pengumpulan data di dapatkan melalui https://www.kaggle.com/datasets/tarekmuhammed/telecom-customers.

B. Teknik Pengumpulan Data

Teknik pengumpulan data yang digunakan dalam penelitian ini sebagai bahan pembuatan sistem, yaitu Studi Pustaka, dan Observasi.

C. Perancangan dan Pemodelan Sistem

Tahap awal dalam penelitian ini dimulai dengan melakukan studi literatur terkait algoritma **Voting Classifier** guna memahami prinsip kerja serta keunggulannya dalam pengklasifikasian data. Selanjutnya, dilakukan perancangan sistem klasifikasi berbasis **data mining**, yang melibatkan pembagian data menjadi data latih dan data uji untuk kebutuhan pemodelan dan evaluasi. Proses pengumpulan data dilakukan dengan memanfaatkan **dataset benchmark pelanggan telekomunikasi** yang tersedia secara publik melalui UCI Repository dan dapat diakses melalui platform Kaggle di tautan: https://www.kaggle.com/datasets/tarekmuhammed/telecom-customers. Dataset ini digunakan sebagai dasar dalam mengembangkan model klasifikasi churn pelanggan. Tahap selanjutnya adalah perancangan sistem klasifikasi menggunakan **algoritma Voting Classifier**, yang mencakup implementasi model ensemble untuk menggabungkan beberapa algoritma pembelajaran mesin. Tujuannya adalah membangun sistem prediksi churn pelanggan telekomunikasi yang lebih akurat dan andal melalui pendekatan kolektif dari beberapa model dasar.

D. Uji Coba Sistem

Tahap uji coba sistem yaitu, melakukan uji coba terhadap algoritma yang diusulkan untuk melihat hasil maupun kinerja yang didapatkan. Tahap analisis hasil menggunakan algoritma Machine learning yaitu, mengevaluasi dan menganalisis hasil dari data yang diteliti berdasarkan algoritma Voting Classifier[12].

E. Analisa Hasil

Hasil uji coba sistem menunjukkan bahwa algoritma Algoritma NaiveBayes berbasis Python terbukti efektif dalam mengklasifikasikan pelanggan telekomunikasi ke dalam kategori "Pelanggan Tetap" atau "Pelanggan yang Akan Berhenti (Churn)"[13]. Model ini mampu menangani ketidakseimbangan data dengan teknik seperti SMOTE[14], sehingga distribusi kelas lebih seimbang dan performa klasifikasi meningkat. Kinerja model sangat bergantung pada kualitas data serta pemilihan fitur yang relevan. Berbagai pembagian data latih dan uji (dari 90:10 hingga 10:90) menunjukkan hasil yang bervariasi. Akurasi model berkisar antara 70% hingga 90%, dengan performa terbaik pada pembagian data 90% latih dan 10% uji (akurasi 88.8%, presisi 93.61%, recall 86.27%, dan F1-score 89.79%). Hasil ini menunjukkan bahwa model lebih stabil ketika data latih lebih dominan. Penerapan SMOTE (teknik oversampling) membantu meningkatkan generalisasi model dengan menyeimbangkan distribusi kelas. Namun, hasil sebelum SMOTE justru lebih tinggi, menunjukkan bahwa ketidakseimbangan alami pada dataset mungkin tidak selalu merugikan performa model [15].

F. Evaluasi Kinerja

Evaluasi kinerja merupakan langkah penting untuk menentukan sejauh mana sebuah model atau sistem mampu menjalankan tugas yang telah ditetapkan secara akurat. Dalam bidang penelitian maupun pengembangan model machine learning, proses ini dilakukan untuk mengukur seberapa efektif model dalam melakukan prediksi atau klasifikasi terhadap data. Penilaian efektivitas tersebut biasanya dilakukan dengan menggunakan sejumlah metrik evaluasi tertentu.

Pada kasus klasifikasi, beberapa metrik evaluasi yang sering digunakan antara lain:

- Akurasi mengukur proporsi prediksi yang benar terhadap seluruh prediksi yang dilakukan oleh model. Metrik ini menunjukkan seberapa sering model menghasilkan klasifikasi yang tepat. Akurasi dapat dihitung menggunakan rumus berikut:

$$\text{Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Jumlah Total Prediksi}} = \frac{TP + TN}{TP + TN + FP + FN}$$

- Precision merupakan metrik yang digunakan untuk menilai ketepatan model dalam melakukan prediksi terhadap kelas positif. Metrik ini menunjukkan proporsi prediksi positif yang memang benar-benar sesuai dengan kondisi sebenarnya. Nilai precision dihitung menggunakan rumus sebagai berikut:

$$\text{Precision} = \frac{TP}{TP + FP}$$

- Recall digunakan untuk menilai sejauh mana model mampu mengenali seluruh data yang termasuk dalam kelas positif. Metrik ini menunjukkan proporsi data positif yang berhasil diprediksi dengan benar dari total data positif yang sebenarnya. Rumus recall dituliskan sebagai berikut:

$$\text{Recall} = \frac{TP}{TP + FN}$$

- F1-Score adalah metrik yang menggabungkan precision dan recall ke dalam satu nilai tunggal dengan mengambil rata-rata harmonis dari keduanya. Metrik ini sangat berguna ketika diperlukan keseimbangan antara ketepatan dan kemampuan model dalam mengenali kelas positif, terutama pada data yang tidak seimbang. Rumus perhitungannya adalah sebagai berikut:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Hasil dan Diskusi

Setelah dilakukan serangkaian eksperimen dalam penelitian ini untuk mengklasifikasikan customer churn menggunakan algoritma Machine Learning, dapat disimpulkan bahwa pendekatan tersebut memberikan hasil klasifikasi yang cukup memuaskan. Algoritma Random Forest, yang diimplementasikan menggunakan bahasa pemrograman Python, terbukti efektif dalam mengelompokkan pelanggan berdasarkan kecenderungannya untuk berhenti berlangganan.

Keberhasilan model ini turut didukung oleh penerapan teknik Synthetic Minority Over-sampling Technique (SMOTE) yang digunakan untuk menangani ketidakseimbangan kelas dalam dataset. Dengan menyamakan jumlah data antara kelas mayoritas dan minoritas, SMOTE membantu meningkatkan kualitas proses pelatihan model dan mengurangi kecenderungan model untuk berat sebelah terhadap kelas tertentu.

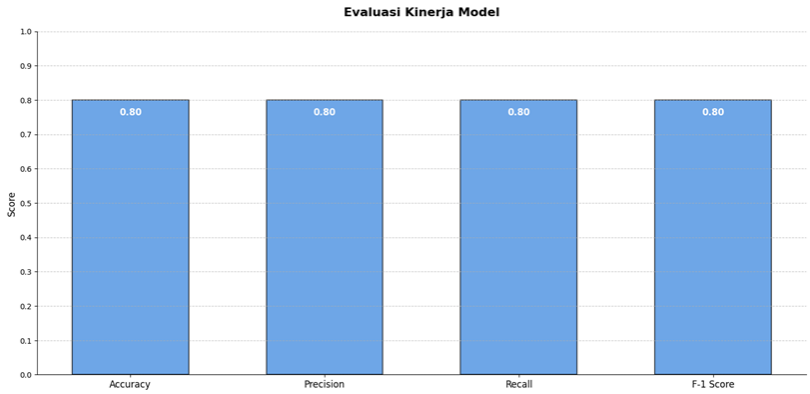
Tabel 1. Confusion Matrix Customer Churn

Random Forest		***Kelas Prediksi***	
Kelas Aktual	Prediksi Pelanggan	Pelanggan Tetap	Pelanggan Yang Akan Berhenti
	Pelanggan Tetap	892	157
	Pelanggan yang akan berhenti	253	768

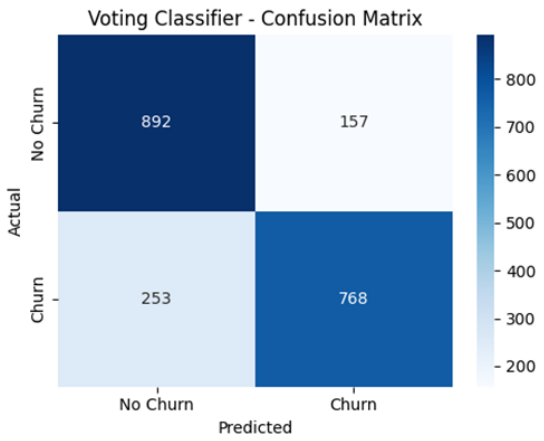
Tabel 2. Hasil Kinerja Model VotingClassifier

Evaluasi Kinerja Random Forest	***Hasil Evaluasi (%)***
Accuracy	80%
Precision	80%
Recall	80%
F1-Score	80%

Berdasarkan hasil pada Tabel 2, metode Voting Classifier menunjukkan performa klasifikasi yang cukup baik, dengan nilai accuracy, precision, recall, dan F1-score masing-masing sebesar 80%. Capaian ini menunjukkan bahwa model memiliki keseimbangan yang baik antara ketepatan dalam mengenali kelas positif dan kemampuannya dalam menangkap seluruh kasus positif. Namun, kinerja model sangat dipengaruhi oleh kualitas data dan pemilihan fitur yang relevan selama proses pelatihan. Salah satu teknik yang terbukti efektif dalam penelitian ini adalah Synthetic Minority Over-sampling Technique (SMOTE). Teknik ini berperan penting dalam menyeimbangkan distribusi data antar kelas, sehingga model dapat belajar pola dari kedua kelas secara lebih optimal. Dengan data yang lebih seimbang, proses pembelajaran menjadi lebih efisien dan model tidak bias terhadap kelas mayoritas. Eksperimen dilakukan dengan berbagai skenario pembagian data latih dan uji, menghasilkan variasi nilai metrik evaluasi, dengan akurasi berkisar antara 60% hingga 84%. Temuan penting lainnya adalah bahwa model cenderung memberikan hasil lebih baik ketika proporsi data uji lebih kecil, karena model mendapatkan lebih banyak data untuk proses pelatihan, yang meningkatkan kemampuan generalisasinya. Penelitian ini juga mengidentifikasi beberapa faktor yang memengaruhi performa model, seperti ukuran dataset, distribusi kelas, dan konfigurasi parameter. Oleh karena itu, pada penelitian selanjutnya disarankan dilakukan proses hyperparameter tuning secara sistematis, misalnya menggunakan metode Grid Search atau Random Search, untuk memperoleh parameter yang paling optimal. Sebagai bentuk evaluasi tambahan, dilakukan perbandingan kinerja Voting Classifier dengan beberapa model individu, yaitu Logistic Regression, Naive Bayes, dan Random Forest. Hasilnya ditunjukkan pada Gambar 1, yang memperlihatkan bahwa Voting Classifier menghasilkan performa yang lebih stabil dan akurat dalam berbagai konfigurasi data. Temuan ini memberikan kontribusi terhadap pengembangan model klasifikasi yang lebih tangguh dan adaptif terhadap dinamika data di dunia nyata.



Gambar 1. Hasil Kinerja Model



Gambar 2. Tabel Confusion Matrix Model VotingClassifier

Kesimpulan

Penelitian ini berhasil mengklasifikasikan perilaku konsumen cokelat melalui penerapan algoritma ensemble Bagging yang dikombinasikan dengan teknik Synthetic Minority Over-Sampling Technique (SMOTE) untuk mengatasi ketidakseimbangan kelas dalam data. Berdasarkan hasil evaluasi, metode ini terbukti mampu meningkatkan nilai akurasi, presisi, recall, dan F1-score secara signifikan jika dibandingkan dengan metode klasifikasi konvensional. Selain itu, struktur pohon keputusan (Decision Tree) yang dihasilkan turut memperjelas pola konsumsi berdasarkan variabel-variabel seperti usia, jenis kelamin, dan tingkat pendidikan. Hasil penelitian juga menunjukkan bahwa penggunaan model Random Forest, sebagai bagian dari algoritma machine learning, memberikan performa klasifikasi yang unggul dalam konteks data pelanggan telekomunikasi. Kinerja model dipengaruhi oleh sejumlah faktor, termasuk kualitas dataset, pemilihan fitur yang relevan, serta proporsi antara data latih dan data uji. Untuk pengembangan lebih lanjut, penelitian mendatang disarankan untuk memfokuskan pada proses optimasi hyperparameter dan penerapan feature engineering guna menciptakan atribut baru yang lebih informatif. Selain itu, pendekatan ini memiliki potensi besar untuk diterapkan dalam klasifikasi perilaku pelanggan lainnya di sektor telekomunikasi, guna mendukung strategi pencegahan churn berbasis machine learning serta meningkatkan interpretabilitas model ensemble dalam pengambilan keputusan.

Daftar Pustaka

[1] F. Pratiwi, M. Barata, and A. Ardianti, “IMPLEMENTASI METODE SMOTE DAN RANDOM OVER-SAMPLING PADA ALGORITMA MACHINE LEARNING UNTUK PREDIKSI CUSTOMER CHURN DI SEKTOR PERBANKAN,” *J. Sist. Inf. dan Inform.*, vol. 8, no. 1 SE-Articles, Jan. 2025, doi: 10.47080/simika.v8i1.3678.

[2] A. Ishaq *et al.*, “Improving the Prediction of Heart Failure Patients’ Survival Using SMOTE and Effective Data Mining Techniques,” *IEEE Access*, vol. 9, pp. 39707–39716, 2021, doi: 10.1109/ACCESS.2021.3064084.

[3] R. Govindaraju, T. Simatupang, and A. Samadhi, “PERANCANGAN SISTEM PREDIKSI CHURN PELANGGAN PT. TELEKOMUNIKASI SELULER DENGAN MEMANFAATKAN PROSES DATA MINING,” *J. Inform.*, vol. 9, Jan. 2009, doi: 10.9744/informatika.9.1.33-42.

[4] M. Joolfoo, R. Jugurnauth, and K. Joolfoo, “Customer Churn Prediction in Telecom Using Machine Learning in Big Data Platform,” *J. Crit. Rev.*, vol. 7, p. 1991, 2020, doi: 10.31838/jcr.07.11.308.

[5] Z. Muteb and M. A. Haq, “Customer Churn Prediction for Telecommunication Companies using Machine Learning and Ensemble Methods,” *Eng. Technol. Appl. Sci. Res.*, vol. 14, pp. 14572–14578, 2024, doi: 10.48084/etasr.7480.

[6] Y. Bharambe, P. Deshmukh, P. Karanjawane, D. Chaudhari, and N. Ranjan, *“Churn Prediction in Telecommunication Industry.”* 2023. doi: 10.1109/ICONAT57137.2023.10080425.

[7] T. G. Dietterich, “Ensemble methods in machine learning,” in *Multiple Classifier Systems: First International Workshop, MCS 2000, Lecture Notes in Computer Science*, 2000, pp. 1–15.

[8] M. Azim Mim, N. Majadi, and P. Mazumder, “A soft voting ensemble learning approach for credit card fraud detection,” *Heliyon*, vol. 10, no. 3, p. e25466, 2024, doi: https://doi.org/10.1016/j.heliyon.2024.e25466.

[9] P. Boozary, S. Sheykhan, H. Tanhaei, and C. Magazzino, “Enhancing customer retention with machine learning: A comparative analysis of ensemble models for accurate churn prediction,” vol. 5, pp. 1–15, 2025, doi: 10.1016/j.jjimei.2025.100331.

[10] R. Alfarez, R. Rianto, and V. Purwayoga, “Penerapan Naïve Bayes untuk Prediksi Customer Churn (Studi Kasus: PT Hutchison 3 Indonesia),” *J. Ris. dan Apl. Mhs. Inform.*, vol. 5, no. 2, pp. 301–307, 2024, doi: 10.30998/jrami.v5i2.8556.

[11] K. Coussement and D. den Poel, “Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques,” *Expert Syst. Appl.*, vol. 34, pp. 313–327, 2006, doi: 10.1016/j.eswa.2006.09.038.

[12] E. Retnoningsih and R. Pramudita, “Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python,” *Bina Insa. Ict J.*, vol. 7, no. 2, p. 156, 2020, doi:

10.51211/biict.v7i2.1422.

[13] A. Manzoor, M. Qureshi, E. Kidney, and L. Longo, "A Review on Machine Learning Methods for Customer Churn Prediction and Recommendations for Business Practitioners," *IEEE Access*, vol. PP, p. 1, Jan. 2024, doi: 10.1109/ACCESS.2024.3402092.

[14] P. R. Undersampling, "Effect of Random Under sampling , Oversampling , and SMOTE on the Performance of Cardiovascular Disease Prediction Models terhadap Kinerja Model Prediksi Penyakit Kardiovaskular," vol. 21, no. 1, pp. 88–102, 2024, doi: 10.20956/j.v21i1.35552.

[15] D. Patil, "Explainable Artificial Intelligence (XAI): Enhancing transparency and trust in machine learning models," Nov. 2024.