

Perbandingan Algoritma Machine Learning untuk Prediksi Customer Churn Perbankan Berbasis Credit Score

Comparison of Machine Learning Algorithms for Banking Customer Churn Prediction Based on Credit Score

¹Suryadillah Hendrawinata*, ²Jasmir Jasmir, ³Gunardi Gunardi

^{1,2,3}Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dinamika Bangsa
Indonesia

^{1,2,3}Jl. Jendral Sudirman Thehok - Jambi 36138 Indonesia

*e-mail: sabitbulan799@gmail.com, gun4rdi.sj@gmail.com, ijay_jasmir@yahoo.com

(received: 8 February 2026, revised: 16 March 2026, accepted: 18 March 2026)

Abstrak

Tingkat pergantian nasabah yang tinggi merupakan tantangan signifikan bagi industri perbankan, yang mengakibatkan kesulitan keuangan yang besar dan biaya yang lebih tinggi untuk nasabah baru. Mengidentifikasi nasabah yang kemungkinan besar akan *churn* secara proaktif sangat penting untuk menerapkan strategi retensi yang efektif. Penelitian ini bertujuan untuk mengatasi masalah ini dengan mengimplementasikan dan membandingkan secara komprehensif tiga algoritma klasifikasi *machine learning* yang berbeda, yaitu Regresi Logistik, *Random Forest*, dan XGBoost. Penelitian ini menggunakan dataset sekunder, yaitu profil nasabah bank yang terdiri dari 10.000 nasabah dengan karakteristik tertentu termasuk skor kredit, saldo, dan aktivitas transaksi. Metodologi penelitian menggunakan Cross-Industry Standard Process for Data Mining (CRISP-DM). Model-model dievaluasi menggunakan metrik seperti Akurasi, *Precision*, *Recall*, F1-Score, dan ROC-AUC. Temuan menunjukkan bahwa model *ensemble* secara signifikan mengungguli model linier (Regresi Logistik) yang hanya mencapai F1-Score 0,286. *Random Forest* muncul sebagai model superior dalam penelitian ini, mencapai Akurasi tertinggi (0,864), F1-Score (0,590), dan ROC-AUC (0,852). Sebagai perbandingan, XGBoost menunjukkan performa yang kompetitif dengan F1-Score (0,579) dan ROC-AUC (0,832). Kesimpulannya adalah *Random Forest* memberikan kinerja paling optimal secara keseluruhan, menawarkan kemampuan terbaik untuk mengidentifikasi nasabah yang berisiko dalam dataset ini.

Kata kunci: klasifikasi, *customer churn*, CRISP-DM, random forest, XGBoost

Abstract

High customer turnover is a significant challenge for the banking industry, resulting in significant financial difficulties and higher fees for new customers. Proactively identifying customers likely to churn is crucial for implementing effective retention strategies. This study aims to address this issue by implementing and comprehensively comparing three different machine learning classification algorithms: Logistic Regression, Random Forest, and XGBoost. This study used a secondary dataset, namely bank customer profiles consisting of 10,000 customers with specific characteristics, including credit scores, balances, and transaction activity. The research methodology used the Cross-Industry Standard Process for Data Mining (CRISP-DM). Models were evaluated using metrics including Accuracy, Precision, Recall, F1-Score, and ROC-AUC. Findings indicate that ensemble models significantly outperform the linear model (Logistic Regression), which achieved an F1-Score of only 0.286. Random Forest emerged as the superior model, achieving the highest Accuracy (0.864), F1-Score (0.590), and ROC-AUC (0.852). In comparison, XGBoost showed competitive performance with an F1-Score of 0.579 and an ROC-AUC of 0.832. In conclusion, Random Forest provides the most optimal performance, offering the best capability to identify at-risk customers within this dataset.

Keywords: classification, *customer churn*, CRISP-DM, , random forest, XGBoost

1 Pendahuluan

Persaingan ketat di industri perbankan menuntut institusi untuk tidak hanya fokus pada akuisisi pelanggan baru tetapi juga secara intensif mempertahankan pelanggan yang sudah ada. Kehilangan pelanggan, atau yang biasa dikenal dengan *customer churn*, merupakan tantangan besar karena mendeskripsikan kondisi di mana pelanggan meninggalkan penyedia layanan saat ini untuk beralih ke kompetitor demi mencari penawaran layanan yang lebih baik. Melindungi loyalitas pelanggan jangka panjang dianggap jauh lebih menguntungkan bagi penyedia layanan dibandingkan mendapatkan pelanggan baru. Biaya untuk mempertahankan pelanggan yang ada jauh lebih rendah daripada biaya untuk mendapatkan pelanggan baru karena nasabah lama cenderung memiliki ketergantungan untuk membeli produk tambahan. Oleh karena itu, kemampuan untuk mengidentifikasi pelanggan yang berisiko *churn* secara dini menjadi aset strategis yang memungkinkan bank untuk menerapkan strategi retensi yang ditargetkan dan efektif, sehingga mengurangi risiko bahaya bisnis dan kerugian finansial[1].

Pengelolaan nasabah dalam dunia perbankan sangat berkaitan erat dengan profil risiko dan perilaku keuangan mereka, yang tercermin melalui skor kredit (*credit score*). Seiring dengan pesatnya perkembangan penelitian mengenai penggunaan data kredit, analisis skor ini kini tidak hanya berfungsi sebagai alat penilai kelayakan pemberian pinjaman, tetapi juga menjadi indikator krusial dalam memprediksi loyalitas nasabah. Melakukan klasifikasi berdasarkan skor kredit menjadi sangat penting karena skor tersebut menggambarkan tingkat kepercayaan dan keterikatan nasabah terhadap layanan bank. Nasabah dengan skor kredit tertentu sering kali menunjukkan pola perilaku yang berbeda dalam hal kecenderungan untuk berhenti berlangganan. Dengan mengintegrasikan analisis skor kredit ke dalam model prediksi *churn*, institusi perbankan dapat melakukan segmentasi nasabah secara lebih akurat dan mengambil keputusan strategis yang lebih tepat sasaran untuk menjaga stabilitas aset Perusahaan.

Seiring dengan perkembangan teknologi digital, bank mengumpulkan data transaksi dan interaksi pelanggan dalam jumlah besar. Data ini dapat dimanfaatkan menggunakan teknik *data mining* dan *machine learning* untuk membangun model prediktif. Model klasifikasi, khususnya, telah banyak digunakan untuk memprediksi *customer churn* dengan menganalisis pola perilaku pelanggan di masa lalu[2]. Berbagai algoritma seperti Regresi Logistik, *Naive Bayes*, *Support Vector Machines*, dan metode *ensemble* telah diterapkan, akan tetapi perbandingan performa antar model seringkali spesifik terhadap *dataset* dan konteks industri.

Namun, implementasi model prediktif ini menghadapi dua permasalahan utama. Permasalahan pertama adalah karakteristik *dataset churn* yang hampir selalu tidak seimbang, dimana jumlah pelanggan yang *churn* jauh lebih sedikit daripada pelanggan yang setia. Banyak penelitian terdahulu terjebak pada metrik akurasi yang tinggi namun menyesatkan, padahal model tersebut gagal mengidentifikasi kelas minoritas yang krusial. Hal ini sejalan dengan studi yang menekankan bahwa penanganan ketidakseimbangan data melalui teknik *resampling* sangat menentukan reliabilitas model prediksi *churn* agar tidak hanya unggul pada nilai akurasi semata[3]. Permasalahan kedua adalah kurangnya interpretabilitas; banyak studi yang membandingkan model-model canggih seperti *ensemble* hanya berhenti pada pelaporan metrik dan memperlakukannya sebagai *black box*. Akibatnya, model tersebut gagal memberikan wawasan yang dapat di tindak lanjuti bagi manajemen mengenai faktor-faktor apa yang sebenarnya paling signifikan mendorong pelanggan untuk *churn*.

Tujuan penelitian ini adalah untuk menerapkan dan membandingkan tiga algoritma klasifikasi populer untuk memprediksi tingkat kehilangan nasabah bank menggunakan *dataset publik 'Churn_Modeling'*. Algoritma yang digunakan meliputi Regresi Logistik sebagai model dasar dan dua metode *ensemble* tingkat lanjut: *Random Forest* dan *XGBoost*. Fokus penelitian ini adalah membandingkan kinerja model menggunakan metrik evaluasi komprehensif, seperti Akurasi, Presisi, Recall, F1-Score, dan ROC-AUC, untuk memberikan rekomendasi model terbaik dalam mengatasi masalah kehilangan nasabah di industri perbankan.

2 Tinjauan Literatur

Beberapa studi terbaru telah memberikan bukti empiris mengenai efektivitas algoritma *eXtreme Gradient Boosting (XGBoost)* dalam mengurangi kompleksitas klasifikasi data di beberapa sektor. Wijaya dkk.[4] serta Wakhidah dkk.[5] mengawali temuan ini dengan mendemonstrasikan bahwa

<http://sistemasi.ftik.unisi.ac.id>

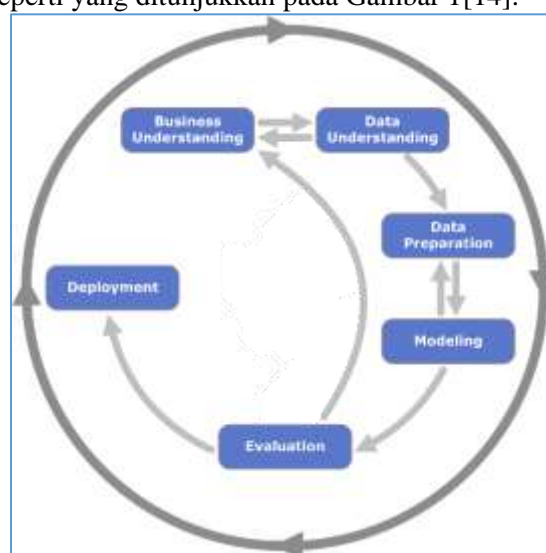
XGBoost mampu memberikan performa yang sangat unggul dalam memprediksi *customer churn* pada industri telekomunikasi, dengan tingkat akurasi mencapai 94,74%. Keunggulan komparatif algoritma ini semakin dipertegas oleh Dachi dan Sitompul[6] yang menemukan bahwa dalam konteks klasifikasi keputusan kredit, *XGBoost* secara konsisten melampaui kemampuan *Random Forest* dengan selisih akurasi yang signifikan. Fenomena dominasi *XGBoost* ini juga divalidasi oleh Givari dkk.[7] dalam kasus persetujuan kredit, di mana model ini mampu mencapai akurasi 96,30%, jauh mengungguli algoritma konvensional lainnya seperti *Random Forest* dan Support Vector Machine (SVM) yang hanya menyentuh angka 92,59%. Temuan-temuan tersebut secara kolektif mengonfirmasi bahwa arsitektur gradient boosting pada *XGBoost* memiliki ketangguhan dalam menangani data yang tidak seimbang (*imbalanced data*) serta mampu menangkap pola non-linear dengan lebih presisi dibandingkan model ensemble lainnya.

Meskipun efektivitas *XGBoost* telah teruji secara luas, sebagian besar literatur yang ada masih memiliki keterbatasan dalam hal generalisasi model pada sektor perbankan penelitian oleh Basri dkk.[8] telah mulai mengeksplorasi penggunaan Regresi Logistik dan *Random Forest* dalam perbankan, namun belum sepenuhnya mengintegrasikan optimasi parameter teknis *XGBoost* dengan analisis fitur perilaku transaksi yang spesifik. Literatur saat ini cenderung berfokus pada perbandingan akurasi antar-algoritma tanpa mengeksplorasi bagaimana penyesuaian *hyperparameter* dapat mengatasi *noise* pada data transaksi perbankan yang sering kali memiliki *outlier* tinggi[9]. Selain itu, aspek interpretabilitas model dalam menjelaskan variabel kritis yang memicu perpindahan nasabah bank masih jarang dikaji secara integratif dalam satu kerangka penelitian.

Oleh karena itu, artikel ini mengisi celah (*novelty*) tersebut dengan memfokuskan analisis pada implementasi *XGBoost* yang telah dioptimasi khusus untuk sektor perbankan menggunakan dataset 'Churn_Modelling'. Kebaruan penelitian ini terletak pada evaluasi komprehensif menggunakan metrik F1-Score dan ROC-AUC untuk mengatasi bias data tidak seimbang, serta membandingkannya secara langsung dengan Regresi Logistik dan *Random Forest* dalam satu kerangka kerja CRISP-DM yang terstruktur [10][11]. Berdasarkan analisis literatur tersebut, dirumuskan hipotesis bahwa penggunaan model ensemble canggih akan secara signifikan mengungguli model linier dalam mengidentifikasi nasabah yang berisiko, terutama jika dikaitkan dengan variabel kritis seperti *credit score* dan aktivitas finansial nasabah [12].

3 Metode Penelitian

Studi ini menerapkan *Cross-Industry Standard Process for Data Mining* (CRISP-DM) sebagai metodologi standar untuk proyek-proyek ilmu data[13]. Kerangka kerja ini menyediakan proses kerja yang terstruktur, mulai dari memahami bisnis hingga mengevaluasi hasilnya. Studi ini didasarkan pada pendekatan CRISP-DM, yang mencakup pemahaman bisnis, pemahaman data, persiapan data, pemodelan, dan evaluasi, seperti yang ditunjukkan pada Gambar 1[14].



Gambar 1 kerangka kerja CRISP-DM [14]

3.1 Pemahaman Bisnis (Business Understanding)

Tahap awal penelitian ini adalah mendefinisikan tujuan dari sudut pandang perbankan. Masalah utama yang dihadapi adalah risiko kerugian akibat nasabah yang berhenti berlangganan. Fokus utama penelitian ini adalah memanfaatkan variabel skor kredit (*credit score*) untuk memprediksi perilaku nasabah. Dengan klasifikasi yang tepat, bank dapat melakukan langkah preventif bagi nasabah berisiko tinggi guna menjaga stabilitas aset dan menekan biaya operasional akuisisi nasabah baru[11].

3.2 Pemahaman Data (Data Understanding)

Data yang digunakan dalam penelitian ini bersumber dari dataset publik '*Churn_Modelling*'. Dataset ini mencakup informasi dari 10.000 nasabah bank yang berasal dari tiga negara di Eropa, yaitu Prancis, Jerman, dan Spanyol. Atribut yang digunakan meliputi data demografis dan finansial seperti skor kredit, saldo rekening, usia, dan keaktifan nasabah. Melalui tahap ini, diidentifikasi adanya ketidakseimbangan kelas pada variabel target (*churn*), di mana jumlah nasabah yang keluar jauh lebih sedikit dibandingkan nasabah yang tetap bertahan.

3.3 Persiapan Data (Data Preparation)

Tahap ini bertujuan untuk mengolah data mentah menjadi format yang siap diproses oleh algoritma. Langkah-langkah yang dilakukan meliputi[15].

1. Pembersihan Data: Langkah awal adalah menghapus kolom-kolom yang tidak relevan untuk proses pemodelan, yaitu '*RowNumber*', '*CustomerId*', dan '*Surname*', karena tidak memiliki nilai prediktif.
2. Transformasi Data: Melakukan *encoding* pada variabel kategorikal (Geografi dan Gender) serta penskalaan fitur (*feature scaling*) pada variabel numerik agar skor kredit dan saldo memiliki rentang yang seragam.
3. Pembagian Data: Setelah dataset dibersihkan dan ditransformasi, dataset tersebut dipisahkan menjadi set pelatihan (*training set*) dan set pengujian (*test set*). Pembagian dilakukan menggunakan fungsi *train_test_split* dengan 80% data pelatihan dan 20% data uji. Parameter *random_state = 42* digunakan untuk memastikan bahwa hasil dapat dihasilkan, dan *stratify=y* diterapkan untuk mengidentifikasi sifat kelas *churn* yang sama pada data pelatihan dan uji, yang menunjukkan bahwa dataset tidak seimbang.

3.4 Pemodelan (Modeling)

Penelitian ini membandingkan tiga algoritma utama dengan landasan teori sebagai berikut:

1. Regresi Logistik

Regresi logistik digunakan untuk memodelkan probabilitas variabel target. Fungsi sigmoid digunakan untuk menghitung hasil linier dalam rentang probabilitas 0 hingga 1, yang didefinisikan sebagai berikut::

$$P(Y = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n)}} \quad (1)$$

Berdasarkan persamaan (1), terdapat parameter β yang merupakan koefisien model. Nilai probabilitas yang dihasilkan akan dikonversi menjadi kelas prediksi (0 atau 1) berdasarkan ambang batas tertentu.

2. Random Forest

Random Forest bekerja dengan membangun banyak pohon keputusan (*decision trees*). Kualitas pemisahan pada setiap simpul ditentukan oleh *Gini Impurity* yang dirumuskan pada persamaan (2):

$$Gini = 1 - \sum_{i=1}^c (P_i)^2 \quad (2)$$

Pada persamaan (2), \mathcal{P}_i merepresentasikan proporsi kelas pada suatu simpul. Nilai *Gini* yang lebih rendah menunjukkan simpul yang lebih murni, yang membantu model dalam melakukan klasifikasi nasabah berdasarkan fitur seperti skor kredit secara lebih akurat.

3. XGBoost

XGBoost mengoptimalkan model secara bertahap dengan meminimalkan fungsi objektif yang menggabungkan *loss function* dan regularisasi, sebagaimana terlihat pada persamaan (3):

$$\text{Obj}(\theta) = \sum_i L(y_i, \hat{y}_i) + \sum_k \Omega(\mathcal{F}k) \quad (3)$$

Dalam persamaan (3), komponen $\Omega(\mathcal{F}k)$ berfungsi sebagai penalti untuk mengontrol kompleksitas model. Hal ini memungkinkan XGBoost untuk menangani data yang tidak seimbang (*imbalanced*) pada kasus *churn* perbankan dengan risiko *overfitting* yang minimal.

3.5 Evaluasi (Evaluation)

Tahap evaluasi bertujuan untuk menilai seberapa baik model yang telah dilatih dapat digeneralisasi pada data baru (data uji)[1]. Untuk mengukur kinerja model secara ilmiah, digunakan beberapa metrik evaluasi sebagai berikut:

- L 1). Akurasi: Persentase prediksi yang benar secara keseluruhan.
- L 2). Presisi: Proporsi prediksi positif (*churn*) yang benar-benar positif.
- L 3). *Recall*: Proporsi kasus positif (*churn*) aktual yang berhasil diidentifikasi.
- L 4). F1-Score: Rata-rata harmonik dari Presisi dan *Recall*.
- L 5). ROC-AUC: Mengukur kemampuan model secara keseluruhan untuk membedakan antara dua kelas.

3.6 Penerapan (Deployment)

Fase terakhir dari pekerjaan CRISP-DM adalah implementasi model dalam kondisi kerja bank. Model dengan kinerja terbaik (XGBoost) akan diintegrasikan ke dalam sistem manajemen hubungan pelanggan bank[16]. Penerapan ini difokuskan pada pemanfaatan skor kredit sebagai parameter otomatis untuk memberikan peringatan dini (*early warning*) kepada pihak manajer hubungan nasabah. Dengan demikian, bank dapat segera memberikan penawaran khusus atau solusi finansial bagi nasabah dengan profil kredit tertentu yang terdeteksi berisiko *churn*, sehingga strategi retensi menjadi lebih tepat sasaran dan terukur[17].

4 Hasil dan Pembahasan

4.1 Deskripsi Dataset dan Persiapan Data

Sesuai dengan alur penelitian, tahap pertama dimulai dengan penjabaran dataset. Penelitian ini menggunakan dataset '*Churn_Modelling*' yang terdiri dari 10.000 catatan nasabah perbankan yang mencakup tiga negara di Eropa yaitu Prancis (50,1%), Jerman (25,1%), dan Spanyol (24,8%). Distribusi data berdasarkan wilayah geografis nasabah tersebut dijabarkan secara rinci pada Tabel 1 berikut:

Table 1 Distribusi geografis nasabah dalam dataset

Negara	Jumlah Nasabah	Presentase(%)
Prancis (<i>France</i>)	5.014	50,1%
Jerman (<i>Germany</i>)	2.509	25,1%
Spanyol (<i>Spain</i>)	2.477	24,8%
Total	10.000	100%

Setelah pendeskripsian data, dilakukan proses persiapan data (*data preparation*) secara runtun sebagai berikut:

1. Pembersihan Data: Menghapus fitur non-prediktif seperti *RowNumber* dan *CustomerId*.
2. Transformasi Kategorikal: Mengubah variabel Geografi dan Jenis Kelamin menjadi numerik melalui teknik *encoding*.
3. Penskalaan Fitur (*Feature Scaling*): Melakukan standarisasi pada variabel skor kredit (*credit score*) dan saldo menggunakan *StandardScaler*.
4. Penanganan Ketidakseimbangan Data: Menerapkan teknik penyesuaian kelas target agar model dapat mengenali pola nasabah yang melakukan *churn* secara lebih akurat.

4.2 Hasil Analisis Berdasarkan Rumusan Masalah

Hasil penelitian ini disajikan secara ringkas untuk mengilustrasikan efektivitas algoritma pembelajaran mesin dalam memprediksi pelanggan yang berhenti berlangganan berdasarkan skor kredit. Perbandingan kinerja ketiga model tersebut ditunjukkan pada Tabel 2 di bawah ini:

Table 2 Perbandingan performa model klasifikasi

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
Logistic Regression	0.806	0.569	0.191	0.286	0.770
Random Forest	0.864	0.766	0.480	0.590	0.852
XGBoost	0.850	0.678	0.505	0.579	0.832

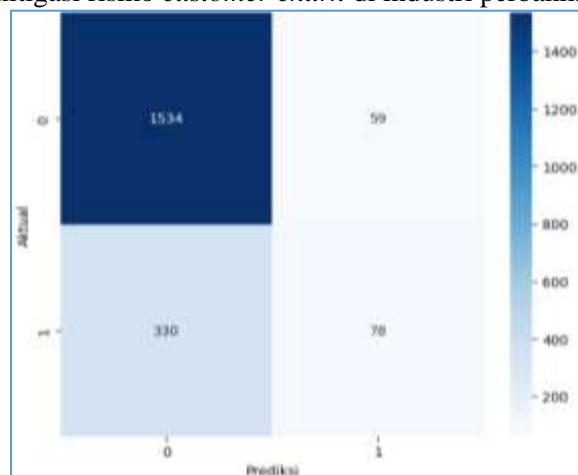
1. Performa Algoritma

Tabel 2 menunjukkan perbandingan kinerja antara XGBoost, Random Forest, dan Regresi Logistik. Sebagai model dasar, Model Regresi Logistik menghasilkan akurasi sebesar 80,6%, tetapi recall (0,191) dan F1-Score (0,286) relatif rendah. Hal ini menunjukkan bahwa model dasar cukup membantu dalam mengidentifikasi pelanggan yang akan berhenti berlangganan, meskipun akurasinya agak rendah karena data yang tidak seimbang.

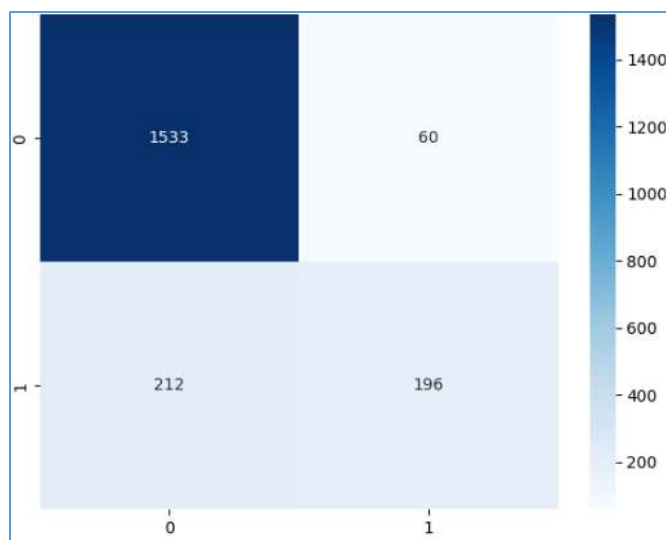
Berdasarkan kelemahan tersebut, ditemukan bahwa model Random Forest menunjukkan performa paling unggul dalam penelitian ini. Model ini mencapai tingkat akurasi tinggi sebesar 0,864, bersama dengan nilai F1-score (0,590) dan ROC-AUC (0,852) yang mengungguli dua algoritma lainnya. Hal ini menunjukkan bahwa metode *ensemble bagging* lebih efektif dalam menangani pola data nasabah pada eksperimen ini dibandingkan model linear maupun *boosting*.

2. Signifikansi Skor Kredit

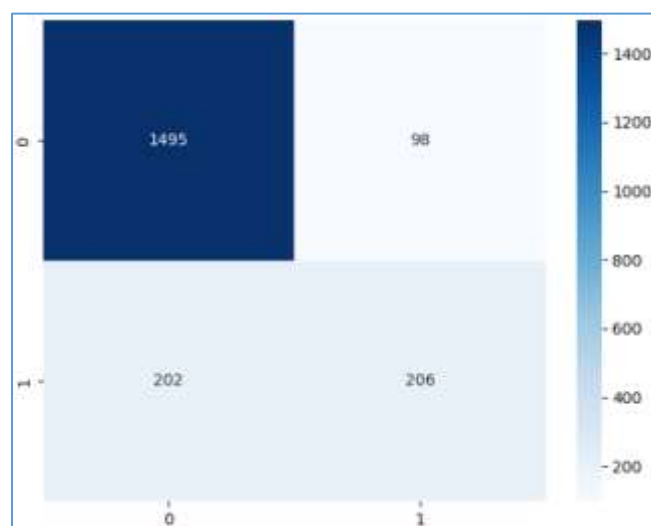
Melalui analisis tingkat kepentingan fitur, ditemukan bahwa skor kredit (*credit score*) tetap menjadi variabel kunci yang memengaruhi keputusan nasabah untuk keluar. Kombinasi antara profil kredit yang rendah dengan variabel ekonomi lainnya terbukti menjadi prediktor kuat dalam model klasifikasi. Hal ini memvalidasi urgensi penggunaan skor kredit sebagai basis utama dalam memitigasi risiko *customer churn* di industri perbankan.



Gambar 2 Logistic regression



Gambar 3 RandomForest

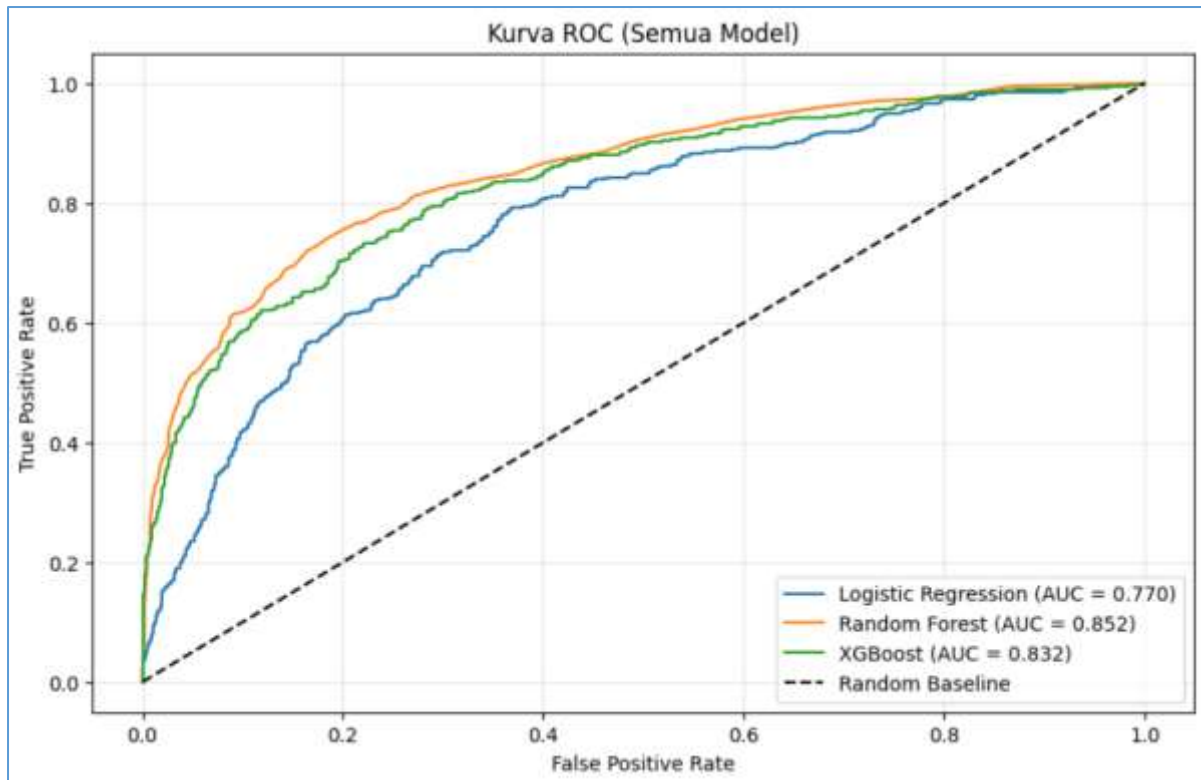


Gambar 4 XGBoost

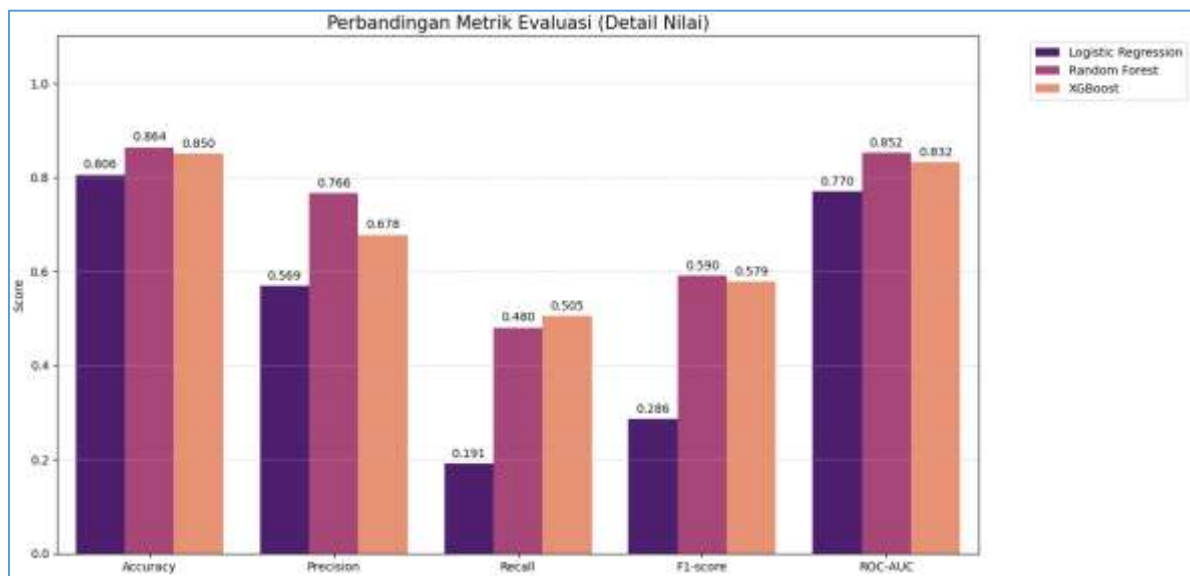
Model gabungan (Random Forest dan XGBoost) menunjukkan peningkatan kinerja yang signifikan jika dibandingkan dengan model linier. Random Forest (Gambar 3) mencapai akurasi 86,4% dengan F1-Score 0,590 dan Recall 0,480. Sementara itu, XGBoost (Gambar 4) mencatatkan akurasi 85,0% dengan F1-Score 0,579, namun memiliki nilai Recall tertinggi yaitu 0,505.

Peningkatan signifikan pada *Recall* dan *F1-Score* menunjukkan bahwa kedua model *ensemble* ini jauh lebih baik dalam menemukan pelanggan yang berisiko *churn* dibandingkan Regresi Logistik. Nilai ROC-AUC, yang mengukur daya pembeda model, juga mengonfirmasi superioritas ini: Regresi Logistik (0,770), Random Forest (0,852), dan XGBoost (0,832). Berbeda dengan draf awal, data terbaru menunjukkan bahwa Random Forest adalah model paling optimal dalam penelitian ini karena memenangkan mayoritas metrik evaluasi, termasuk Akurasi, Presisi, F1-Score, dan ROC-AUC.

Analisis *Confusion Matrix* pada Gambar memperjelas temuan dari Tabel 2. *Confusion Matrix* Regresi Logistik (Gambar 2) menunjukkan jumlah *False Negative* yang sangat tinggi, yang berarti banyak nasabah *churn* yang tidak terdeteksi. Sebaliknya, *Confusion Matrix* untuk Random Forest (Gambar 3) dan XGBoost (Gambar 4) menunjukkan jumlah *True Positive* yang lebih tinggi secara signifikan. Meskipun XGBoost sedikit lebih unggul dalam menangkap jumlah nasabah *churn* (Recall), Random Forest memberikan tingkat kesalahan prediksi (*False Positive*) yang lebih rendah, menjadikannya model yang lebih stabil secara keseluruhan untuk tujuan bisnis.



Gambar 5 Kurva ROC



Gambar 6 Bar Chart

Gambar 5 (Kurva ROC) dan Gambar 6 (Bar Chart) memvisualisasikan perbandingan performa antar model secara komprehensif. Pada Gambar 5, kurva untuk Random Forest dan XGBoost terlihat jelas berada lebih dekat ke sudut kiri atas dibandingkan kurva Regresi Logistik. Hal ini menunjukkan bahwa untuk setiap tingkat *False Positive Rate*, model *ensemble* mencapai *True Positive Rate* yang jauh lebih tinggi.

Seara spesifik, Random Forest menunjukkan area di bawah kurva (AUC) yang paling luas (0,852), mengungguli XGBoost (0,832) dan Regresi Logistik (0,770). Visualisasi *Bar Chart* pada Gambar 6 semakin mempertegas dominasi Random Forest dalam metrik Akurasi dan F1-Score. Hal ini mengonfirmasi bahwa model berbasis *decision tree* dengan teknik *bagging* (Random Forest)

memberikan pemisahan kelas yang lebih stabil dan optimal dalam mendeteksi nasabah yang berisiko *churn* pada karakteristik dataset ini.

5 Kesimpulan

Penelitian ini menyimpulkan bahwa implementasi algoritma *machine learning* berbasis *ensemble* sangat efektif dalam memprediksi *customer churn* di industri perbankan dengan memanfaatkan variabel skor kredit (*credit score*) sebagai prediktor kunci. Berdasarkan hasil pengujian menggunakan kerangka kerja CRISP-DM pada 10.000 data nasabah, model Random Forest muncul sebagai model yang paling optimal dengan tingkat akurasi sebesar 86,4% dan F1-Score sebesar 0,590, mengungguli XGBoost (F1-Score: 0,579) dan Regresi Logistik (F1-Score: 0,286). Analisis ini membuktikan bahwa integrasi skor kredit ke dalam model klasifikasi tidak hanya meningkatkan presisi deteksi nasabah berisiko, tetapi juga memberikan wawasan strategis bagi bank untuk merancang program retensi yang lebih efisien guna menjaga stabilitas aset dan meminimalisir kerugian finansial.

Referensi

- [1] P. P. Singh, F. I. Anik, R. Senapati, A. Sinha, N. Sakib, and E. Hossain, "Investigating customer churn in banking: A machine learning approach and visualization app for data science and management," *Data Science and Management*, vol. 7, no. 1, pp. 7–16, Mar. 2024, doi: 10.1016/j.dsm.2023.09.002.
- [2] B. A. Maulana and N. Hidayati, "Churn Prediction in Credit Customers Using Random Forest and XGBoost Methods," *Indonesian Journal of Data and Science*, vol. 6, no. 1, pp. 82–90, Mar. 2025, doi: 10.56705/ijodas.v6i1.215.
- [3] R. E. Ako *et al.*, "Effects of Data Resampling on Predicting Customer Churn via a Comparative Tree-based Random Forest and XGBoost," *Journal of Computing Theories and Applications*, vol. 2, no. 1, pp. 86–101, Jun. 2024, doi: 10.62411/jcta.10562.
- [4] D. A. Kusuma, A. R. Dewi, and A. R. Wijaya, "Perbandingan Random Forest dan Convolutional Neural Network dalam Memprediksi Peralihan Pelanggan," MEI, 2025.
- [5] L. N. Wakhidah, A. K. Zyen, and B. B. Wahono, "Evaluation of Telecommunication Customer Churn Classification with SMOTE Using Random Forest and XGBoost Algorithms," 2025. [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [6] J. Melvin Ayu Soraya Dachi and Pardomuan Sitompul, "Analisis Perbandingan Algoritma XGBoost dan Algoritma Random Forest Ensemble Learning pada Klasifikasi Keputusan Kredit," *Jurnal Riset Rumpun Matematika dan Ilmu Pengetahuan Alam*, vol. 2, no. 2, pp. 65–71, Jul. 2023, doi: 10.55606/jurrimipa.v2i2.1336.
- [7] M. R. Givari, R. Mochamad, and Y. U. Sulaeman², "Perbandingan Algoritma SVM, Random Forest Dan XGBoost Untuk Penentuan Persetujuan Pengajuan Kredit," vol. 16, no. 1, 2022, [Online]. Available: <https://journal.uniku.ac.id/index.php/ilkom>
- [8] M. Basri, M. Iqbal Pradipta, and K. Aditya, "A Comparative Study : Predicting Customer Churn in Banking Using Logistic Regression & Random Forest," *Ultimatics : Jurnal Teknik Informatika*, vol. 17, no. 1, 2025.
- [9] Y. Xu, C. Rao, X. Xiao, and F. Hu, "Novel early-warning model for customer churn of credit card based on gsaibas-catboost," *CMES - Computer Modeling in Engineering and Sciences*, vol. 137, no. 3, pp. 2715–2742, 2023, doi: 10.32604/cmcs.2023.029023.
- [10] E. J. Casabianca, M. Catalano, L. Forni, E. Giarda, and S. Passeri, "A machine learning approach to rank the determinants of banking crises over time and across countries," *J. Int. Money Finance*, vol. 129, Dec. 2022, doi: 10.1016/j.jimonfin.2022.102739.
- [11] G. Gunawan, "Data Mining Using Crisp-Dm Process Framework On Official Statistics: A Case Study Of East Java Province," *Jurnal Ekonomi dan Pembangunan*, vol. 29, no. 2, pp. 183–198, Dec. 2021, doi: 10.14203/jep.29.2.2021.183-198.
- [12] P. P. Singh, F. I. Anik, R. Senapati, A. Sinha, N. Sakib, and E. Hossain, "Investigating customer churn in banking: A machine learning approach and visualization app for data science and management," *Data Science and Management*, vol. 7, no. 1, pp. 7–16, Mar. 2024, doi: 10.1016/j.dsm.2023.09.002.

- [13] C. Schröer, F. Kruse, and J. M. Gómez, “A systematic literature review on applying CRISP-DM process model,” in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 526–534. doi: 10.1016/j.procs.2021.01.199.
- [14] K. Anam and A. R. Rinaldi, “Komparasi Algoritma Machine Learning dalam Klasifikasi Loyalitas Nasabah Bank Berbasis Particle Swarm Optimization,” 2024. [Online]. Available: <https://www.kaggle.com/>
- [15] R. A. Casonatto, T. D. P. G. Souza, and A. M. Mariano, “Quality and Risk Management in Data Mining: A CRISP-DM Perspective,” in *Procedia Computer Science*, Elsevier B.V., 2024, pp. 161–168. doi: 10.1016/j.procs.2024.08.257.
- [16] C. Mariscal, Y. Yustiawan, F. C. Rochim, and E. Tanuar, “Implementing and analyzing fairness in banking credit scoring,” in *Procedia Computer Science*, Elsevier B.V., 2024, pp. 1492–1499. doi: 10.1016/j.procs.2024.03.150.
- [17] M. Alonso Dos Santos, C. Zarco-Fernández, and F. Liébana-Cabanillas, “The new frontier of customer understanding: Financial satisfaction and AutoML in banking,” *Journal of Retailing and Consumer Services*, vol. 90, Mar. 2026, doi: 10.1016/j.jretconser.2025.104601.