Analisis Komparatif Random Forest dan XGBoost Berbasis Class Weighting pada Kasus Imbalanced Dataset Kredit

Daffa Tusi Santoso[1*], Edwin Pavel Ekapagliuca[2], Iqbal Syabudin[3], Linda Nuzul Kharismasari[4]

[1*]Teknik Informatika/Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta
[1*]230103095@mhs.udb.ac.id

[2]Teknik Informatika/Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta
[2]230103098@mhs.udb.ac.id

[3]Teknik Informatika/Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta
[3]230103104@mhs.udb.ac.id

[4]Teknik Informatika/Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta
[4]230103106@mhs.udb.ac.id

***Abstrak*— Penilaian kelayakan kredit merupakan aspek penting dalam manajemen risiko perbankan, namun sering menghadapi permasalahan ketidakseimbangan kelas (imbalanced data) yang menyebabkan model klasifikasi cenderung bias terhadap kelas mayoritas. Kondisi ini meningkatkan risiko False Negatives, yaitu nasabah berisiko gagal bayar yang diprediksi sebagai nasabah lancar. Penelitian ini bertujuan membandingkan kinerja algoritma Random Forest sebagai model baseline dengan XGBoost yang dioptimasi menggunakan class weight balancing dan hyperparameter tuning untuk meningkatkan kemampuan deteksi nasabah berisiko. Penelitian menggunakan pendekatan kuantitatif eksperimental dengan kerangka kerja CRISP-DM yang meliputi business understanding, data understanding, data preparation, modeling, evaluation, dan deployment. Dataset dibagi menjadi data latih dan data uji dengan proporsi 80:20, sedangkan optimasi XGBoost dilakukan menggunakan RandomizedSearch. Hasil penelitian menunjukkan bahwa model XGBoost menghasilkan Recall sebesar 62,47% dan F1-Score sebesar 53,87%, lebih tinggi dibandingkan Random Forest yang memperoleh Recall 35,80% dan F1-Score 45,72%. Meskipun akurasi menurun dari 81,20% menjadi 76,33%, peningkatan sensitivitas berhasil mengurangi False Negatives sehingga lebih sesuai untuk kebutuhan manajemen risiko kredit. Analisis feature importance menunjukkan bahwa riwayat status pembayaran, khususnya PAY_0, menjadi faktor paling berpengaruh dalam prediksi. Model terbaik selanjutnya diimplementasikan ke dalam prototipe Decision Support System berbasis Streamlit yang mampu memberikan prediksi risiko kredit secara real-time. Hasil penelitian menunjukkan bahwa optimasi XGBoost efektif meningkatkan kemampuan klasifikasi pada dataset kredit yang tidak seimbang serta mendukung pengambilan keputusan kredit yang lebih akurat.**
Kata kunci— Risiko Kredit, XGBoost, Random Forest, Class Weight Balancing, CRISP-DM.

***Abstract*— Creditworthiness assessment is a critical component of banking risk management but is often challenged by class imbalance, causing classification models to favor the majority class. This condition increases the risk of False Negatives, where high-risk borrowers are incorrectly classified as low-risk customers. This study aims to compare the performance of a baseline Random Forest model with an optimized Extreme Gradient Boosting (XGBoost) model using class weight balancing and hyperparameter tuning to improve the detection of high-risk borrowers. A quantitative experimental approach based on the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework was employed, consisting of business understanding, data understanding, data preparation, modeling, evaluation, and deployment. The dataset was divided into training and testing sets using an 80:20 ratio, while XGBoost hyperparameters were optimized through RandomizedSearch. The experimental results show that the optimized XGBoost achieved a Recall of 62.47% and an F1-Score of 53.87%, outperforming the baseline Random Forest with a Recall of 35.80% and an F1-Score of 45.72%. Although overall accuracy decreased from 81.20% to 76.33%, the substantial improvement in sensitivity significantly reduced False Negatives, making the model more suitable for credit risk management. Feature importance analysis identified recent payment history, particularly PAY_0, as the most influential predictor of credit default. The best-performing model was deployed as a Streamlit-based Decision Support System capable of providing real-time credit risk predictions. The findings demonstrate that the optimized XGBoost model effectively improves classification performance on imbalanced credit datasets and supports more reliable credit decision-making.**
Keywords— Credit Risk, XGBoost, Random Forest, Class Weight Balancing, CRISP-DM.

I. PENDAHULUAN

Manajemen risiko finansial, khususnya dalam penilaian kelayakan kredit, merupakan elemen krusial bagi stabilitas dan profitabilitas institusi perbankan. Salah satu tantangan komputasional terbesar dalam mengembangkan model prediksi

risiko kredit adalah penanganan fenomena ketidakseimbangan kelas (imbalanced data), di mana representasi nasabah dengan riwayat pembayaran lancar (kelas 0) sangat mendominasi dibandingkan nasabah yang mengalami gagal bayar (kelas 1)[1] Kondisi distribusi yang timpang ini memicu bias pada algoritma klasifikasi prediktif, yang cenderung mengutamakan pengenalan pola pada kelas mayoritas. Akibatnya, model sering kali menghasilkan tingkat False Negatives yang tinggi yakni nasabah gagal bayar yang terprediksi lancar yang mana dalam domain finansial, hal ini membawa dampak kerugian langsung yang jauh lebih signifikan dibandingkan rasio (False Positives)[2]

Penelitian sebelumnya menunjukkan bahwa ketidakseimbangan distribusi data pada kasus prediksi risiko kredit dapat menurunkan kemampuan model machine learning dalam mengenali nasabah yang berpotensi mengalami gagal bayar. Kondisi tersebut menyebabkan model lebih cenderung memprediksi kelas mayoritas sehingga performa klasifikasi terhadap kelas minoritas menjadi kurang optimal[3] Oleh karena itu, berbagai penelitian mulai mengembangkan pendekatan berbasis *ensemble learning* yang dipadukan dengan teknik penanganan data tidak seimbang dan seleksi fitur untuk meningkatkan kemampuan model dalam mengidentifikasi risiko kredit secara lebih akurat. Hasil penelitian menunjukkan bahwa kombinasi metode tersebut mampu meningkatkan performa prediksi serta memberikan hasil klasifikasi yang lebih baik dibandingkan penggunaan model konvensional[4]

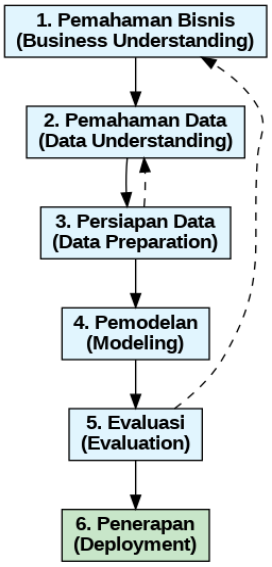
Dalam literatur machine learning, berbagai algoritma ensemble telah diaplikasikan untuk pemodelan prediktif. Algoritma Random Forest secara luas diimplementasikan sebagai model baseline dasar karena ketahanannya terhadap outlier, namun model ini sering kali kesulitan menangani ketimpangan kelas yang ekstrem tanpa intervensi data tambahan. Di sisi lain, algoritma Extreme Gradient Boosting (XGBoost) menawarkan pendekatan yang lebih mutakhir dalam meminimalkan galat prediksi. Melalui optimasi yang melibatkan teknik penyesuaian bobot kelas (class weight balancing) dan pencarian parameter yang tepat (hyperparameter tuning seperti RandomizedSearch), arsitektur model dapat distabilkan untuk mencapai tingkat generalisasi yang tinggi serta mencegah terjadinya overfitting. Pendekatan kombinasi ini diklaim mampu meningkatkan sensitivitas algoritma dalam mendeteksi populasi minoritas[5]Selain itu, penelitian tersebut juga menerapkan teknik penanganan ketidakseimbangan data serta proses seleksi fitur guna mengurangi pengaruh atribut yang kurang relevan. Hasil penelitian menunjukkan bahwa kombinasi berbagai metode dalam model ensemble mampu memberikan performa yang lebih baik dibandingkan penggunaan model tunggal pada kasus credit scoring dengan distribusi data yang tidak seimbang[6]. Penelitian lain juga telah membuktikan bahwa meskipun algoritma Random Forest dan XGBoost memiliki kemampuan klasifikasi yang baik, XGBoost yang dikombinasikan dengan penanganan imbalanced data terbukti menghasilkan evaluasi metrik yang lebih sensitif terhadap kelas minoritas [7]. Temuan tersebut menunjukkan bahwa pemilihan algoritma dan strategi penanganan imbalanced dataset merupakan faktor penting dalam meningkatkan kualitas sistem credit scoring [8].

Berdasarkan latar belakang permasalahan tersebut, maksud dari penelitian ini adalah untuk melakukan analisis komparatif antara kinerja model baseline Random Forest dengan model XGBoost yang telah dioptimasi menggunakan pembobotan kelas. Tujuan utama riset ini dilakukan adalah untuk mengatasi isu ketidakseimbangan kelas pada dataset nasabah, sehingga model dapat mengidentifikasi nasabah berisiko dengan tingkat sensitivitas (Recall) dan rasio keseimbangan (F1-Score) yang jauh lebih akurat dibandingkan sekadar mengandalkan metrik Akurasi global. Lebih jauh lagi, riset ini ditujukan untuk mengekstraksi interpretasi logika keputusan model (feature importance), serta memvalidasi fungsionalitas algoritma secara praktis melalui implementasi prototipe Decision Support System berbasis web interaktif. Selain itu, penelitian ini diharapkan dapat memberikan informasi mengenai efektivitas penerapan *class weighting* pada algoritma Random Forest dan XGBoost sehingga dapat digunakan sebagai bahan pertimbangan dalam pengembangan sistem *credit scoring*. Bagi lembaga

keuangan, hasil penelitian ini diharapkan dapat membantu meningkatkan ketepatan penilaian kelayakan kredit dan mendukung proses pengambilan keputusan yang lebih akurat.

II. METODOLOGI PENELITIAN

Penelitian ini dijalankan menggunakan pendekatan kuantitatif eksperimental dengan mengadopsi kerangka kerja *Cross-Industry Standard Process for Data Mining* (CRISP-DM)[9]. Kerangka kerja ini dipilih karena menyediakan siklus hidup analisis data yang adaptif dan terstruktur, mulai dari pemahaman domain masalah finansial hingga tahap penerapan model ke dalam lingkungan produksi berupa sistem informasi pendukung keputusan. Alur metodologi dalam penelitian ini dibagi menjadi enam tahapan utama yang saling beriterasi, yaitu pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penerapan sistem.



Gambar 1. Alur CRISP-DM

A. *Pemahaman Bisnis (Business Understanding)*

Tahap awal penelitian difokuskan pada identifikasi masalah operasional dan risiko manajemen finansial di industri perbankan[10]Tantangan utama yang dihadapi adalah tingginya potensi kerugian finansial langsung apabila model klasifikasi mengalami kegagalan deteksi atau menghasilkan tingkat False Negatives yang tinggi, yakni kondisi di mana nasabah yang sebenarnya berisiko gagal bayar justru diprediksi sebagai nasabah lancar. Oleh karena itu, tujuan teknis penggalian data dalam riset ini ditetapkan untuk meminimalkan bias algoritma prediktif, serta meningkatkan sensitivitas atau Recall dan rasio keseimbangan metrik (F1-Score) pada klasifikasi populasi nasabah yang berisiko tinggi[11]

B. *Pemahaman Data (Data Understanding)*

Fase ini melibatkan analisis terhadap karakteristik dan struktur dataset kredit yang digunakan dalam eksperimen. Karakteristik utama dari dataset ini adalah adanya ketidakseimbangan kelas (imbalanced data) yang sangat signifikan, di mana distribusi sampel didominasi oleh kelas mayoritas yaitu nasabah dengan status pembayaran lancar (kelas 0) dibandingkan nasabah dengan status gagal bayar (kelas 1). Analisis terhadap atribut dan riwayat nasabah telah terbukti menjadi komponen krusial dalam membangun model prediksi risiko gagal bayar kredit yang akurat [12]. Dalam penelitian ini, evaluasi dilakukan menggunakan dataset (Default of Credit Card Clients) dari (UCI Machine Learning Repository), yang memuat 23 parameter input terstruktur mencakup riwayat aktivitas nasabah. Atribut-atribut tersebut dikelompokkan ke dalam tiga kategori utama, yaitu profil demografis statis (batas limit kredit, jenis kelamin, tingkat pendidikan, status pernikahan, dan usia), riwayat status pembayaran selama enam bulan terakhir, serta perilaku finansial kuantitatif yang mencakup matriks nominal tagihan dan pembayaran riil bulanan.

C. *Persiapan Data (Data Preparation)*

Pada tahapan persiapan data, dilakukan serangkaian manipulasi dataset agar siap digunakan pada proses pelatihan algoritma. Langkah pertama adalah pembagian dataset (*data splitting*), di mana data dialokasikan secara acak menjadi data latih dan data uji dengan proporsi pengujian sebesar 20 persen dari total populasi alokasi ini menghasilkan volume pengujian sebesar 6.000 instans data dengan representasi sampel minoritas sebanyak 1.327 instans di dalamnya. Selanjutnya, untuk mengatasi bias akibat ketimpangan distribusi kelas minoritas, diterapkan teknik penyesuaian bobot kelas (*class weight balancing*) secara langsung ke dalam fungsi objektif algoritma pada saat fase pembelajaran berlangsung.

D. Pemodelan (Modeling)

Tahap pemodelan dilakukan dengan menerapkan pendekatan analisis komparatif antara dua algoritma berbasis *ensemble learning*. Penelitian menggunakan algoritma *Random Forest* standar tanpa optimasi parameter sebagai model *baseline* untuk tolok ukur performa awal, yang kemudian dibandingkan dengan model utama menggunakan algoritma *Extreme Gradient Boosting* (XGBoost) yang diintegrasikan dengan pembobotan kelas. Guna meminimalkan galat prediktif, menyederhanakan arsitektur pohon, dan mencegah model dari risiko *overfitting*, dilakukan proses pencarian parameter optimal (*hyperparameter tuning*) memanfaatkan metode *RandomizedSearch*[13]. Konfigurasi parameter arsitektur pohon terbaik yang diujikan meliputi batas kedalaman pohon maksimal sebesar 3, tingkat pembelajaran sebesar 0.05, jumlah estimasi pohon sebanyak 200, serta rasio pemotongan sampel dan fitur masing-masing pada nilai 0.6.

E. Evaluasi Model (Evaluation)

Evaluasi kapabilitas model dilakukan secara komprehensif menggunakan pendekatan multi-metrik statistik. Metrik yang dianalisis meliputi nilai Akurasi global, Presisi, *Recall*, dan tingkat *F1-Score* khusus pada kelas minoritas[14]. Pergeseran performa klasifikasi dipantau secara detail melalui visualisasi Matriks Kebingungan (*Confusion Matrix*). Ketahanan model terhadap variasi ambang batas diuji melalui analisis Kurva *Receiver Operating Characteristic* (ROC) dengan mengukur luasan Area di Bawah Kurva (AUC), serta Kurva *Precision-Recall* (PR) untuk mengukur capaian tingkat *Average Precision* di atas garis rasio *baseline* populasi minoritas. Tahap ini diakhiri dengan proses Ekstraksi Kepentingan Fitur (*Feature Importance*) untuk memberikan interpretasi logis mengenai variabel yang paling dominan memengaruhi keputusan algoritma[15].

F. Penerapan (Deployment)

Fase akhir dari metodologi ini adalah mentransformasikan model *machine learning* analitis terbaik ke dalam aplikasi fungsional. Arsitektur model XGBoost yang telah dioptimasi diekspor dan diimplementasikan ke dalam prototipe sistem informasi pendukung keputusan berbasis web interaktif. Antarmuka pengguna sistem ini dikembangkan menggunakan *framework* berbasis Python yaitu Streamlit. Guna mencegah beban kognitif berlebih, sistem dirancang dengan membagi 23 parameter nasabah ke dalam tiga segmentasi panel formulir visual. Sistem akan memproses matriks data input secara waktu nyata menggunakan fungsi logistik dari model XGBoost untuk mengekstraksi nilai probabilitas, kemudian menampilkan hasil berupa kelas prediksi final beserta persentase tingkat keyakinannya.

III. HASIL DAN PEMBAHASAN

Bagian ini menguraikan hasil eksekusi dari tahapan inti kerangka kerja CRISP-DM yang telah dirancang pada metodologi penelitian, yang mencakup fase pemodelan (*modeling*), evaluasi model (*evaluation*), dan penerapan sistem (*deployment*). Diskusi berfokus pada efektivitas penanganan data tidak seimbang dan pengujian komparatif performa algoritma.

A. *Hasil Pemahaman Bisnis (Business Understanding)*

Hasil dari fase ini adalah ditetapkannya metrik keberhasilan bisnis untuk model yang dibangun. Mengingat tingginya potensi kerugian finansial langsung akibat kegagalan deteksi nasabah gagal bayar, parameter keberhasilan utama dalam riset ini tidak difokuskan pada Akurasi global, melainkan pada kemampuan model untuk memaksimalkan metrik Recall (sensitivitas) dan F1-Score.

B. *Hasil Pemahaman Data (Data Understanding)*

Berdasarkan eksplorasi awal terhadap dataset *Default of Credit Card Clients* yang bersumber dari repositori publik *UCI Machine Learning*, ditemukan bahwa karakteristik utama data mengalami ketidakseimbangan kelas (*imbalanced data*) yang sangat signifikan. Distribusi populasi sangat didominasi oleh kelas mayoritas (nasabah dengan status pembayaran lancar atau kelas 0), dibandingkan dengan kelas minoritas (nasabah gagal bayar atau kelas 1). Selain itu, terekstraksi 23 parameter input terstruktur yang akan menjadi prediktor, meliputi profil demografis statis, riwayat status pembayaran, dan perilaku finansial kuantitatif.

C. *Hasil Persiapan Data (Data Preparation)*

Melalui proses pembagian data (*data splitting*), dataset berhasil dipisahkan dengan proporsi pengujian sebesar 20 persen. Proses ini menghasilkan volume data uji sebanyak 6.000 instans, di mana representasi sampel minoritas (nasabah gagal bayar) hanya berjumlah 1.327 instans. Untuk mengatasi ketimpangan distribusi ini pada saat fase pelatihan, teknik penyesuaian bobot kelas (*class weight balancing*) diterapkan secara langsung ke dalam algoritma.

D. *Hasil Fase Pemodelan (Modeling)*

Tahap pemodelan mengimplementasikan algoritma *baseline Random Forest* dan model *Extreme Gradient Boosting* (XGBoost). Untuk mengoptimalkan kinerja XGBoost dalam mengatasi masalah ketidakseimbangan kelas (*imbalanced data*) yang didominasi oleh kelas nasabah lancar (0), diterapkan teknik penyesuaian bobot kelas (*class weight balancing*). Selain itu, pencarian parameter optimal (*hyperparameter tuning*) dilakukan menggunakan metode *RandomizedSearch* guna menghasilkan konfigurasi arsitektur pohon yang lebih sederhana namun stabil, memiliki kemampuan generalisasi tinggi, serta mencegah terjadinya *overfitting* terhadap data latih. Hasil pengaturan parameter terbaik yang diperoleh dari proses optimasi tersebut disajikan pada Tabel 1.

Tabel 1. Nilai Random Search Hyperparameter

Hyperparameter	Nilai Hyperparameter
max_depth	3
learning_rate	0.05
n_estimators	200
subsample	0.6
colsample_bytree	0.6

E. *Hasil Evaluasi Model (Evaluation)*

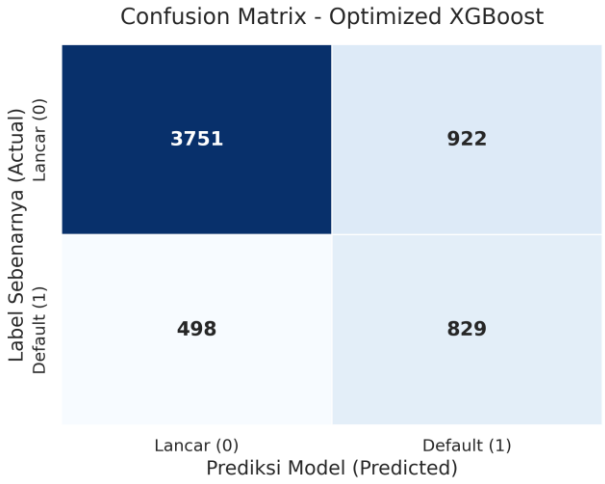
Evaluasi performa klasifikasi dilakukan berdasarkan pengujian terhadap 20 persen data uji yang setara dengan 6.000 instans. Hasil komparasi metrik evaluasi antara model *baseline* dan model XGBoost yang telah dioptimasi dapat dilihat pada Tabel 2.

Table 2. Komparasi Metrik Evaluasi

Metrik Evaluasi	Random Forest(Base)	XGBoost(Tuned0
Akurasi	81.20%	76.33%
Precision(Kelas 1)	63.25%	47.34%
Recall(Kelas 1)	35.80%	62.47%
F1-Score(Kelas 1)	45.72%	53.87%

Berdasarkan Tabel 2, penerapan XGBoost yang dioptimasi berhasil mereduksi bias algoritma terhadap kelas mayoritas. Hal ini dibuktikan oleh lonjakan drastis pada metrik *Recall* untuk kelas 1 (gagal bayar), yang meningkat dari 35.80% menjadi 62.47%. Walaupun terjadi kompromi (*trade-off*) berupa penurunan Akurasi keseluruhan (menjadi 76.33%) dan Presisi (menjadi 47.34%), peningkatan *F1-Score* sebesar kurang lebih 8% menegaskan bahwa XGBoost memberikan rasio keseimbangan klasifikasi yang lebih baik. Dalam domain manajemen risiko finansial, model yang sangat sensitif (*Recall* tinggi) jauh lebih krusial untuk mencegah kebocoran deteksi nasabah berisiko tinggi.

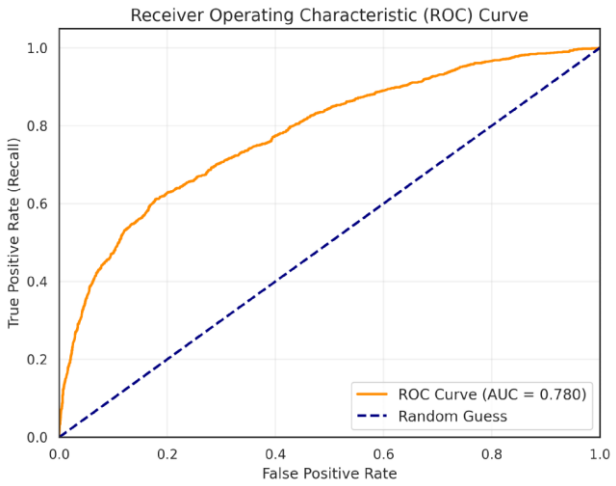
Pergeseran kapabilitas model dalam mengklasifikasikan instans dapat diamati lebih detail melalui Matriks Kebingungan pada Gambar 1.



Gambar 2. Confusion Matrix

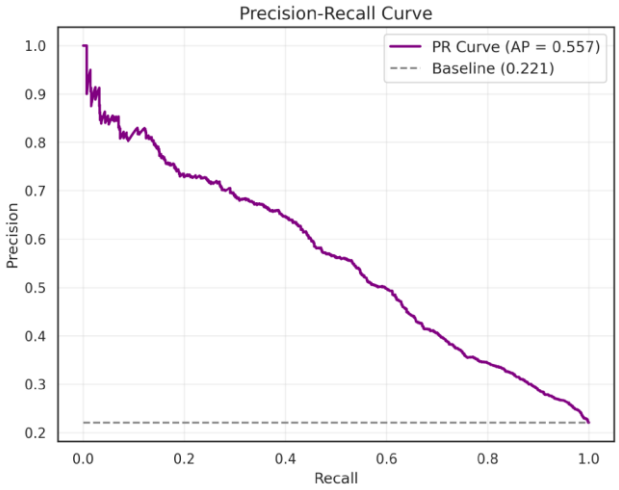
Pada hasil pengujian kelas positif (nasabah default berjumlah 1.327), model XGBoost yang telah dioptimasi berhasil mengidentifikasi dengan benar sejumlah besar nasabah bermasalah (True Positives). Tingkat False Negatives (nasabah gagal bayar yang lolos karena diprediksi lancar) berhasil ditekan secara signifikan dibandingkan model baseline. Peningkatan volume False Positives (nasabah lancar yang diprediksi gagal bayar) dinilai sebagai rasio risiko yang dapat diterima (acceptable

risk) oleh institusi finansial dibandingkan kerugian finansial langsung akibat tingginya tingkat False Negatives.



Gambar 3. Kurva Receiver Operating Characteristic (ROC)

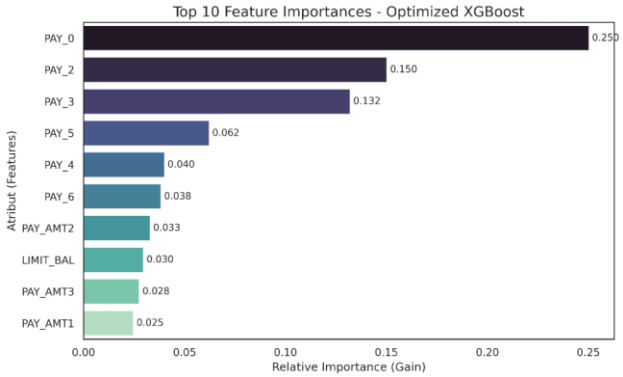
Kurva ROC pada Gambar 2 memperlihatkan nilai AUC sebesar 0.780. Nilai yang merepresentasikan luasan kurva di atas garis acak (0.5) ini mengindikasikan bahwa model memiliki kemampuan diskriminatif yang memadai dalam membedakan nasabah berisiko dan tidak berisiko.



Gambar 4. Kurva Precision-Recall (PR)

Sebagai metrik yang lebih objektif untuk dataset tidak seimbang, Kurva PR pada Gambar 3 memperkuat argumen penelitian ini. Kurva tersebut menunjukkan bahwa model mempertahankan tingkat Average Precision (AP) sebesar 0.557 yang berada di atas rasio baseline (0.221) populasi minoritas, membuktikan bahwa kinerja F1-Score yang dicapai bersifat stabil.

Tahap evaluasi diakhiri dengan interpretasi terhadap logika pengambilan keputusan algoritma (*explainability*) yang divisualisasikan melalui grafik *Feature Importance* untuk mengetahui determinan utama risiko kredit, seperti yang ditunjukkan pada Gambar 4.



Gambar 5. Sepuluh Atribut Berpengaruh Terhadap Prediksi Model

Merujuk pada Gambar 4, analisis menyimpulkan bahwa riwayat status pembayaran pada bulan terakhir (PAY_0) mendominasi perolehan informasi (Gain) algoritma. Prediktor sekunder yang mengikuti pola ini adalah (PAY_2) dan (PAY_3). Temuan algoritmik ini memvalidasi asumsi empiris perbankan bahwa kedisiplinan transaksi historis jangka pendek dan alokasi batas kredit merupakan variabel paling reliabel untuk memproyeksikan status pembayaran di masa depan, mengesampingkan variabel demografis statis seperti usia atau tingkat pendidikan.

F. *Hasil Fase Penerapan Sistem (Deployment)*

Guna memvalidasi fungsionalitas model *machine learning* secara praktis di fase penerapan, arsitektur XGBoost terbaik yang dihasilkan dari proses *hyperparameter tuning* diimplementasikan ke dalam sebuah prototipe aplikasi web interaktif sederhana. Sistem antarmuka pengguna (*User Interface*) ini dikembangkan menggunakan *framework* berbasis Python (Streamlit), yang dirancang untuk menerima input 23 parameter nasabah secara terstruktur. Untuk mengoptimalkan pengalaman pengguna (*User Experience*) dan mencegah beban kognitif akibat

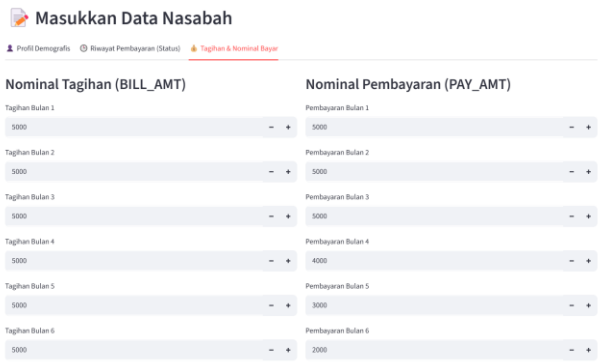
banyaknya kolom input, antarmuka sistem dibagi ke dalam tiga segmentasi formulir utama.

Gambar 6. Antarmuka Input Profil Demografis

Seperti yang ditunjukkan pada Gambar 5, segmentasi pertama merekam variabel-variabel demografis statis dan batas kredit awal. Parameter yang diinputkan meliputi Limit Kredit (LIMIT_BAL), Jenis Kelamin (SEX), Tingkat Pendidikan (EDUCATION), Status Pernikahan (MARRIAGE), dan Usia (AGE). Sistem menggunakan elemen antarmuka *dropdown selectbox* dan *slider* untuk memastikan data yang dimasukkan memiliki format yang valid dan sesuai dengan *encoding* pada saat pelatihan model.

Gambar 7. Antarmuka Input Riwayat Pembayaran

Berdasarkan Gambar 6, pengguna diharuskan memasukkan status pembayaran nasabah selama enam bulan terakhir (PAY_0 hingga PAY_6). Sistem memberikan keterangan indikator numerik secara eksplisit pada layar (misalnya, angka -1 untuk pembayaran tepat waktu, dan angka positif untuk merepresentasikan jumlah bulan keterlambatan). Hal ini meminimalisir kesalahan interpretasi input oleh pengguna akhir.



Gambar 8. Antarmuka Input Tagihan & Nominal Bayar

Gambar 7 menampilkan matriks input ganda yang terbagi menjadi kolom Nominal Tagihan (BILL_AMT1 s.d. BILL_AMT6) dan Nominal Pembayaran (PAY_AMT1 s.d. PAY_AMT6). Pemisahan visual ini memudahkan pengguna untuk melakukan komparasi langsung secara visual antara jumlah yang ditagihkan oleh bank dengan jumlah riil yang dibayarkan oleh nasabah pada bulan yang sama.

Gambar 9. Antarmuka Output Prediksi

Keberhasilan implementasi prototipe interaktif ini secara efektif menjembatani celah antara riset data mining analitis dan pengaplikasian sistem informasi pendukung keputusan (Decision Support System) di tahap produksi industri perbankan.

IV. KESIMPULAN

Penelitian ini berhasil mencapai tujuannya dalam mengatasi masalah ketidakseimbangan kelas pada dataset risiko kredit melalui kerangka kerja CRISP-DM dengan mengekstrak 23 parameter input terstruktur. Melalui pembagian data dengan proporsi 20% untuk pengujian (setara dengan 6.000 instans, di mana 1.327 instans merupakan sampel minoritas nasabah gagal bayar), optimasi model XGBoost menggunakan teknik class weight balancing dan hyperparameter tuning terbukti lebih unggul dalam

manajemen risiko dibandingkan model baseline Random Forest. Model XGBoost yang dioptimasi ini menghasilkan peningkatan signifikan pada metrik sensitivitas, di mana Recall kelas 1 melonjak drastis dari 35.80% menjadi 62.47% dan F1-Score meningkat dari 45.72% menjadi 53.87%. Meskipun terjadi trade-off berupa penurunan Akurasi keseluruhan menjadi 76.33% dan Presisi menjadi 47.34%, model ini terbukti efektif menekan angka False Negatives. Keandalan model juga diperkuat oleh capaian nilai AUC sebesar 0.780 pada Kurva ROC serta nilai Average Precision (AP) sebesar 0.557 (jauh di atas baseline populasi minoritas sebesar 0.221). Selain itu, analisis Feature Importance memvalidasi secara kuantitatif bahwa riwayat status pembayaran jangka pendek, terutama PAY_0, PAY_2, dan PAY_3, merupakan prediktor paling penentu dalam kelayakan kredit dibandingkan variabel demografis statis.

Secara praktis, model arsitektur XGBoost terbaik ini telah diimplementasikan secara utuh ke dalam prototipe Decision Support System berbasis web menggunakan Streamlit. Aplikasi ini mengoptimalkan pengalaman pengguna dengan membagi 23 parameter nasabah ke dalam 3 segmentasi formulir utama untuk mengeksekusi prediksi risiko secara real-time. Sebagai langkah pengembangan di masa mendatang, disarankan untuk menguji ketahanan model pada dataset lintas institusi yang lebih masif dan mengombinasikannya dengan teknik hybrid resampling guna memaksimalkan kemampuan generalisasi sistem.

REFERENSI

[1] A. Manurung and A. Rahman, "Universitas Multi Data Palembang | 1337 5 th MDP STUDENT CONFERENCE (MSC) Implementasi Algoritma Random Forest untuk Memprediksi Penjualan Produk Pada Toko Ritel Elektronik di Kota Palembang."

[2] N. Hazizah and A. Feranika, "Implementasi Algoritma Random Forest Dalam Klasifikasi Risiko Gagal Bayar Kartu Kredit Pada Nasabah Bank," *Jurnal Manajemen Teknologi dan Sistem Informasi (JMS)*, vol. 5, no. 1, 2025, doi: 10.33998/jms.v5i1.

[3] "Prediksi Risiko Gagal Bayar Pinjaman".

[4] Y. Jin, W. Zhang, X. Wu, Y. Liu, and Z. Hu, "A Novel Multi-Stage Ensemble Model with a Hybrid Genetic Algorithm for Credit Scoring on Imbalanced Data," *IEEE Access*, vol. 9, pp. 143593–143607, 2021, doi: 10.1109/ACCESS.2021.3120086.

[5] R. Irsyadul Anam and M. Hamdani, "Januari-Maret," *Jurnal Ilmu Komputer Dan Informatika*, vol. 2, no. 3, p. 2026, 2026, [Online]. Available: https://jurnal.globalscients.com/index.php/jiki.1075

[6] H. Chen, C. Yang, M. Du, and Y. Zhang, "Research on Credit Risk Prediction under Unbalanced Dataset Based on Ensemble Learning," *Math. Probl. Eng.*, vol. 2023, no. 1, Jan. 2023, doi: 10.1155/2023/2927393.

[7] A. Fauziah, "Optimizing Credit Scoring Performance Using Ensemble Feature Selection with Random Forest," *Jurnal Matematika, Statistika dan Komputasi*, vol. 21, no. 2, pp. 560–572, Jan. 2025, doi: 10.20956/j.v21i2.42032.

[8] N. Goodarzi, "Title: Comparison Between XGBoost and Random Forest in Credit Risk Assessment Using the German Credit Dataset."

[9] M. Mendes and T. Farinha, "MIDA—Method for Industrial Data Analysis Based on CRISP-DM," *Mach. Learn. Knowl. Extr.*, vol. 8, no. 4, Apr. 2026, doi: 10.3390/make8040108.

[10] A. Abdullah, D. U. Khairah, and M. W. Pangestika, "COMPARISON OF RANDOM FOREST AND XGBOOST ALGORITHMS IN CREDIT CARD FRAUD CLASSIFICATION," *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 6, no. 3, pp. 392–398, Dec. 2025, doi: 10.37859/coscitech.v6i3.10470.

[11] D. A. A. Pertiwi, K. Ahmad, J. Unjung, and M. A. Muslim, "A Performance Comparison of Data Balancing Model to Improve Credit Risk Prediction in P2P Lending," *Scientific Journal of Informatics*, vol. 11, no. 4, pp. 881–890, Nov. 2024, doi: 10.15294/sji.v11i4.14018.

[12] S. Yang, Z. Huang, W. Xiao, and X. Shen, "Interpretable Credit Default Prediction with Ensemble Learning and SHAP," May 2025, doi: 10.1109/ICAHN67688.2025.00027.

[13] N. P. Nur Fauzi, S. Khomsah, and A. D. Putra Wicaksono, "Penerapan Feature Engineering dan Hyperparameter Tuning untuk Meningkatkan Akurasi Model Random Forest pada Klasifikasi Risiko Kredit," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 12, no. 2, pp. 251–262, Apr. 2025, doi: 10.25126/jtiik.2025128472.

[14] S. Danil, N. Rahaningsih, R. D. Dana, and . M., "PENINGKATAN KLASIFIKASI KEMISKINAN INDONESIA MENGGUNAKAN METODE DECISION TREE," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 2, Apr. 2025, doi: 10.23960/jitet.v13i2.6336.

[15] L. A. Martini, Indrianto, and I. K. Seneng, "Analisis Performa, Explainability, dan Fairness pada Model Klasifikasi Multi-Dataset Medis Menggunakan SVM dan Random Forest," *Jurnal Algoritma*, vol. 23, no. 1, May 2026, doi: 10.33364/algoritma/v.23-1.3339.