

Prediksi Persetujuan Pinjaman Bank Ritel Menggunakan CatBoost dengan Optimasi Hyperparameter Berbasis Optuna dan Analisis Interpretabilitas SHAP

Dwi Martanti Ningsih, Andy Haryoko, Amaludin Arifia*

Program Studi Teknik Informatia Fakultas Teknik, Universitas PGRI Ronggolawe, Tuban, Indonesia
Email: ¹dwimartantiningsih30@gmail.com, ²andyharyoko@gmail.com, ³*amaludinarifia@gmail.com

Email korespondensi: penulis3@email.com

Submitted 11-05-2026; Accepted 04-06-2026; Published 30-06-2026

Abstrak

Sektor perbankan menghadapi tantangan serius dalam mengklasifikasikan permohonan pinjaman secara akurat, khususnya pada dataset yang didominasi fitur kategorik dan memiliki distribusi kelas yang tidak seimbang (71%:29%). Penelitian ini mengusulkan penerapan CatBoost (Categorical Boosting) dengan optimasi hyperparameter berbasis Optuna framework Bayesian optimization menggunakan Tree-structured Parzen Estimator (TPE) untuk prediksi persetujuan pinjaman bank. Dua skenario penanganan class imbalance dievaluasi secara komparatif: SMOTE dan class_weight bawaan CatBoost. Eksperimen dilakukan pada dataset 381 sampel dengan 15 fitur aktif (12 fitur asli dan 3 fitur rekayasa), menggunakan 5-fold stratified cross-validation. Hasil menunjukkan bahwa LightGBM mencapai performa terbaik keseluruhan dengan Accuracy 93,42%, Precision 94,83%, Recall 96,49%, F1-Score 95,65%, ROC-AUC 0,9215, dan MCC 0,8220. CatBoost (SMOTE) menempati posisi kompetitif dengan AUC-CV 0,9030 dan F1 94,02%. Analisis SHAP (SHapley Additive exPlanations) mengidentifikasi Credit_History sebagai fitur dominan ($\text{mean|SHAP|}=3,2225$), diikuti ApplicantIncome (0,6896) dan Property_Area Semiurban (0,5222). Penelitian ini berkontribusi sebagai kajian pertama yang mengintegrasikan CatBoost+Optuna+XAI-SHAP dalam konteks prediksi persetujuan pinjaman bank ritel dengan fitur kategorik dominan, sekaligus menyediakan perbandingan sistematis terhadap LightGBM, XGBoost, dan Random Forest.

Kata Kunci: CatBoost; Optuna; Bayesian Optimization; Loan Approval; SHAP; Class Imbalance; SMOTE; LightGBM

Abstract

The banking sector faces significant challenges in accurately classifying loan applications, particularly for datasets dominated by categorical features with imbalanced class distribution (71%:29%). This study proposes the application of CatBoost (Categorical Boosting) with Optuna-based hyperparameter optimization a Bayesian optimization framework using Tree-structured Parzen Estimator (TPE) for bank loan approval prediction. Two class imbalance handling scenarios are comparatively evaluated: SMOTE and CatBoost built-in class_weight. Experiments are conducted on a dataset of 381 samples with 15 active features (12 original features and 3 engineered features) using 5-fold stratified cross-validation. Results show LightGBM achieves the best overall performance with Accuracy 93.42%, Precision 94.83%, Recall 96.49%, F1-Score 95.65%, ROC-AUC 0.9215, and MCC 0.8220. CatBoost (SMOTE) achieves competitive performance with AUC-CV 0.9030 and F1 94.02%. SHAP (SHapley Additive exPlanations) analysis identifies Credit_History as the dominant feature ($\text{mean|SHAP|}=3.2225$), followed by ApplicantIncome (0.6896) and Property_Area_Semiurban (0.5222). This study contributes as the first investigation integrating CatBoost+Optuna+XAI-SHAP in retail bank loan approval prediction with dominant categorical features, while providing systematic comparison against LightGBM, XGBoost, and Random Forest.

Keywords: CatBoost; Optuna; Bayesian Optimization; Loan Approval; SHAP; Class Imbalance; SMOTE; LightGBM

1. PENDAHULUAN

Sektor perbankan merupakan tulang punggung perekonomian global yang berperan sentral dalam mengalokasikan modal melalui kegiatan pemberian pinjaman. Aktivitas pinjaman yang sehat mendorong pertumbuhan konsumsi, investasi, dan ekspansi usaha; namun keputusan persetujuan yang tidak akurat dapat memicu akumulasi kredit bermasalah (non-performing loan/NPL) yang mengancam stabilitas sistem keuangan secara luas. Proses evaluasi permohonan pinjaman secara tradisional mengandalkan penilaian berbasis aturan (rule-based assessment) dan pertimbangan subyektif analis kredit yang rentan terhadap inkonsistensi, bias, dan ketidakefisienan dalam skala besar. Kondisi ini mendorong kebutuhan mendesak terhadap sistem otomatis berbasis kecerdasan buatan yang mampu mengklasifikasikan permohonan pinjaman secara akurat, konsisten, dan dapat dijelaskan [1].

Dataset permohonan pinjaman umumnya memiliki dua karakteristik teknis yang menantang: (1) dominasi variabel kategorik jenis kelamin, status pernikahan, tingkat pendidikan, jenis pekerjaan, dan lokasi properti yang tidak dapat langsung diproses algoritma machine learning tanpa encoding yang tepat; dan (2) ketidakseimbangan distribusi kelas target (class imbalance) yang menyebabkan model bias terhadap kelas mayoritas [2]. Kedua tantangan ini secara bersamaan mengurangi efektivitas pendekatan machine learning konvensional dalam konteks evaluasi kredit.

Berbagai pendekatan machine learning telah diusulkan untuk mengotomasi keputusan persetujuan pinjaman, mulai dari regresi logistik, support vector machine, random forest, hingga keluarga gradient boosting seperti XGBoost dan LightGBM [1] [3] [4]. Tinjauan literatur sistematis dilakukan terhadap 63 artikel machine learning untuk credit scoring (2018–2024) menunjukkan bahwa algoritma berbasis ensemble secara konsisten mencapai performa terbaik pada data keuangan terstruktur [3]). Selanjutnya membandingkan AdaBoost, XGBoost, LightGBM, dan CatBoost pada dataset 7.542 nasabah bank Vietnam dan menemukan LightGBM sebagai model terbaik [4], sementara terdapat penelitian mengkonfirmasi superioritas CatBoost pada dataset Alibaba Tianchi untuk prediksi loan default [5].

Namun terdapat celah metodologis signifikan: sebagian besar studi belum mengaplikasikan CatBoost algoritma yang dirancang khusus menangani fitur kategorik melalui ordered target encoding secara optimal pada konteks loan approval [6] [7]. Dalam survei komprehensif pertama tentang CatBoost pada big data mengidentifikasi sensitivitas hyperparameter sebagai tantangan utama [8]. Selain itu, optimasi hyperparameter pada studi-studi CatBoost umumnya masih mengandalkan grid search manual, bukan Bayesian optimization [9] menunjukkan bahwa optimasi CatBoost yang tepat untuk persetujuan kartu kredit dapat mencapai ROC-AUC 0,7809 dan F1 0,5490, mengungguli kompetitor pada kondisi imbalanced [10].

Penelitian ini mengusulkan penerapan CatBoost dengan optimasi hyperparameter berbasis Optuna sebagai solusi prediksi persetujuan pinjaman yang lebih akurat dan interpretatif. CatBoost dipilih karena: (1) kemampuan native menangani fitur kategorik tanpa one-hot encoding melalui ordered target statistics; (2) resistensi terhadap overfitting melalui ordered boosting yang mencegah prediction shift; (3) performa komputasi superior pada dataset kecil [6]. Integrasi SHAP (SHapley Additive exPlanations) sebagai lapisan interpretabilitas memungkinkan penjelasan keputusan yang konsisten secara teoritis berbasis teori permainan Shapley [11] [12] [13]. Kombinasi CatBoost+Optuna+SHAP yang belum pernah diterapkan secara eksplisit pada dataset loan approval kategorik-dominan merupakan novelty utama penelitian ini. Tujuan penelitian mencakup: (1) membangun model prediksi loan approval menggunakan CatBoost teroptimasi Optuna; (2) mengevaluasi dua skenario penanganan class imbalance (SMOTE vs. class_weight); (3) membandingkan CatBoost teroptimasi dengan XGBoost, LightGBM, Random Forest, dan CatBoost Default; serta (4) mengidentifikasi fitur paling berpengaruh melalui analisis SHAP. Tabel 1 merangkum kontribusi penelitian ini terhadap literatur yang ada.

Tabel 1. Ringkasan Kontribusi Penelitian

No.	Aspek Kontribusi	Deskripsi
1	Metode Utama	CatBoost + Optuna Bayesian Optimization untuk prediksi loan approval bank ritel kategorik-dominan
2	Penanganan Imbalance	Evaluasi komparatif SMOTE vs. class_weight CatBoost pada class imbalance 71%:29%
3	Benchmark Sistematis	Perbandingan terhadap XGBoost, LightGBM, CatBoost Default, dan Random Forest
4	Interpretabilitas XAI	Analisis SHAP global (mean SHAP) & lokal (waterfall) kontribusi fitur per keputusan

2. METODOLOGI PENELITIAN

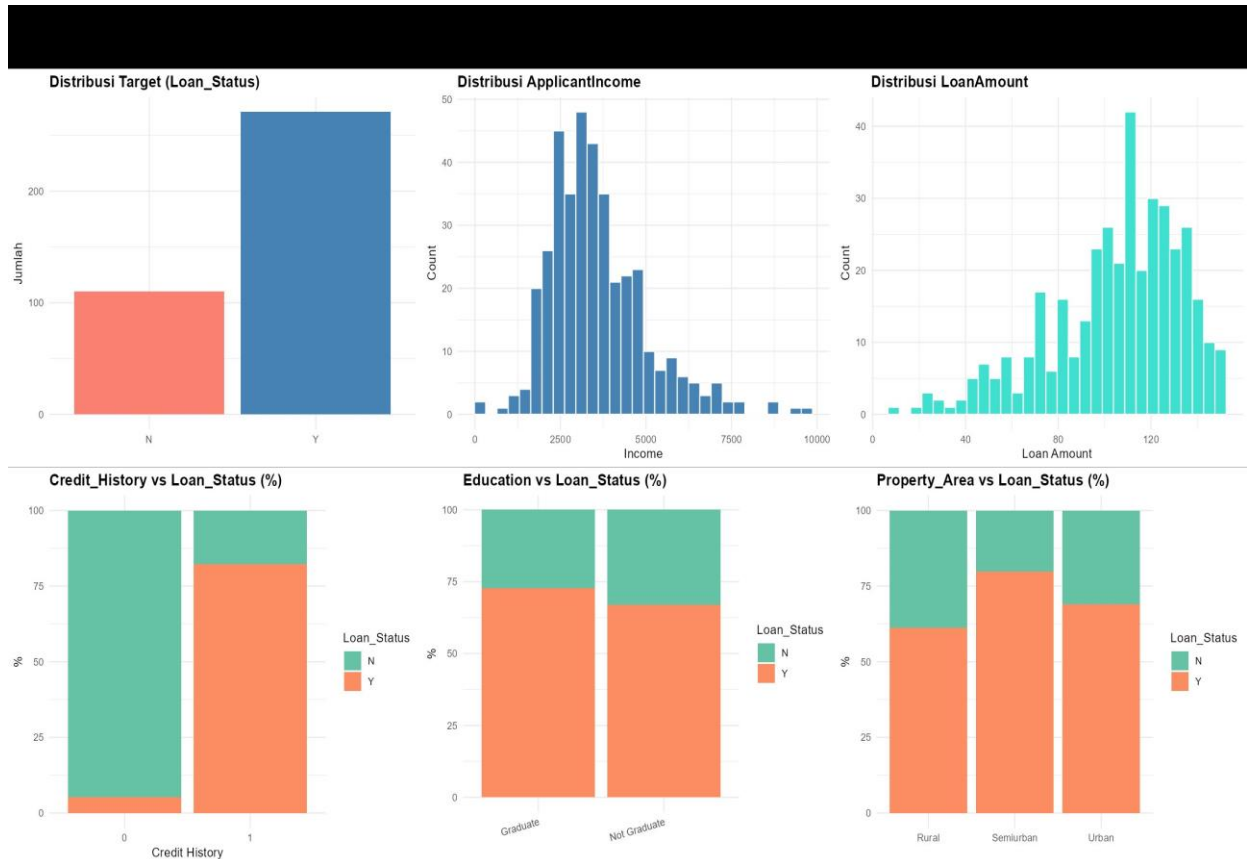
2.1 Desain Penelitian dan Dataset

Penelitian ini menggunakan pendekatan eksperimental kuantitatif dengan paradigma supervised machine learning untuk tugas klasifikasi biner. Variabel dependen adalah Loan_Status (Y=dijetujui, N=ditolak). Dataset yang digunakan adalah loan_data.csv yang terdiri dari 381 sampel permohonan pinjaman bank dengan 13 variabel satu identitas (Loan_ID), satu target, enam kategorik, dan lima numerik. Distribusi kelas menunjukkan ketidakseimbangan: 271 sampel disetujui (71,1%) dan 110 ditolak (28,9%), rasio 2,46:1. Seluruh eksperimen menggunakan Python 3.10 dengan pustaka CatBoost 1.2, Optuna 3.x, scikit-learn 1.3, imbalanced-learn 0.11, dan SHAP 0.44. Parameter random_state = 42 ditetapkan konsisten untuk reproduisibilitas. Tabel 3 menyajikan deskripsi lengkap seluruh variabel dataset. Gambar 1 menampilkan hasil Exploratory Data Analysis (EDA) yang mencakup distribusi variabel utama dan hubungannya dengan Loan_Status.

Tabel 2. Deskripsi Variabel Dataset

Variabel	Tipe	Missing (n/%)	Unique	Keterangan
Gender	Kategorik	5 / 1,31%	2	Male (291), Female (85)
Married	Kategorik	0 / 0%	2	Yes (228), No (153)
Dependents	Kategorik	8 / 2,10%	4	0 (234), 1 (52), 2 (59), 3+ (28)
Education	Kategorik	0 / 0%	2	Graduate (278), Not Graduate (103)
Self_Employed	Kategorik	21 / 5,51%	2	No (325), Yes (35)
Property_Area	Kategorik	0 / 0%	3	Semiurban (149), Urban (126), Rural (106)
ApplicantIncome	Numerik	0 / 0%	—	Min: 150 Median: 3.333 Max: 9.703
CoapplicantIncome	Numerik	0 / 0%	—	Min: 0 Median: 983 Max: 33.837
LoanAmount	Numerik	0 / 0%	—	Median: 110 rb Distribusi right-skewed
Loan_Amount_Term	Numerik	11 / 2,89%	10	360 bulan = 94,6% sampel
Credit_History	Numerik*	30 / 7,87%	2	1=baik (294) → 82,3% disetujui; 0=buruk (57) → 5,3% disetujui
Loan_Status	Target	0 / 0%	2	Y=271 (71,1%) N=110 (28,9%) class imbalance 2,46:1

Keterangan: *Credit_History diperlakukan sebagai fitur numerik biner; secara semantik bersifat kategorik ordinal. Tiga fitur rekayasa (TotalIncome, EMI, LoanIncomeRatio) ditambahkan pada tahap preprocessing sehingga total fitur aktif menjadi 15.



Gambar 1. Distribusi variabel utama dataset distribusi target (class imbalance).

Gambar 1 menunjukkan enam panel visualisasi yang menggambarkan karakteristik dataset permohonan pinjaman. Distribusi `Loan_Status` menunjukkan ketidakseimbangan kelas, dengan rasio 2,46:1 antara permohonan ditolak dan disetujui. Distribusi `ApplicantIncome` dan `LoanAmount` bersifat right-skewed, menunjukkan sebagian besar pemohon memiliki pendapatan dan pinjaman di rentang menengah. Hubungan `Credit_History` dan `Loan_Status` menunjukkan perbedaan signifikan, dengan pemohon berstatus baik memiliki tingkat persetujuan lebih tinggi. Pendidikan dan lokasi properti menunjukkan pengaruh yang lebih terbatas. EDA mengonfirmasi bahwa `Credit_History` adalah faktor paling dominan dalam keputusan persetujuan pinjaman.

2.2 Preprocessing dan Feature Engineering

Imputasi missing values dilakukan secara selektif berdasarkan karakteristik distribusi data: variabel kategorik (`Gender`, `Dependents`, `Self_Employed`) diimputasi dengan modus (nilai paling sering muncul); variabel numerik (`Loan_Amount_Term`, `Credit_History`) diimputasi dengan median dipilih karena distribusi right-skewed membuat rata-rata tidak representatif. Seluruh parameter imputasi dihitung eksklusif dari training set untuk mencegah data leakage yang dapat menghasilkan estimasi performa yang optimistik secara artifisial. Feature engineering mengkonstruksi tiga fitur turunan yang merepresentasikan kapasitas pembayaran aktual: (1) $TotalIncome = ApplicantIncome + CoapplicantIncome$ pendapatan gabungan yang mencerminkan kemampuan finansial keluarga; (2) $EMI = LoanAmount / Loan_Amount_Term$ estimasi cicilan bulanan sebagai indikator beban bayar; (3) $LoanIncomeRatio = LoanAmount / TotalIncome$ rasio leverage finansial pemohon. Variabel target dikodekan $Y \rightarrow 1, N \rightarrow 0$. Enam variabel kategorik fitur tidak dikenakan one-hot encoding didaftarkan sebagai `cat_features` dalam `CatBoost` untuk diproses melalui mekanisme `ordered target statistics` secara native [6]

2.3 Split Data dan Penanganan Class Imbalance

Dataset dibagi 80:20 menggunakan stratified random split [14] [15][16] (train=304, test=76), menjamin proporsi kelas konsisten di kedua partisi. Dua skenario penanganan class imbalance dievaluasi secara komparatif. Skenario A SMOTE: menghasilkan sampel sintetik kelas minoritas melalui interpolasi linear antar sampel terdekat ($k_neighbors=5$), menghasilkan 432 sampel seimbang ($Y=216, N=216$) pada training set, sesuai metodologi Xie dan Shingadia (2025)[2]. Skenario B `class_weight`: pemberian bobot kelas proporsional pada fungsi loss `CatBoost` ($w_N=1,727; w_Y=0,704$) tanpa memodifikasi distribusi data, serupa dengan pendekatan `cost-sensitive learning` [10].

2.4 Algoritma CatBoost dan Optimasi Optuna

CatBoost mengimplementasikan dua inovasi kunci yang membedakannya dari implementasi gradient boosting lain [6]; [17]: (1) *ordered boosting* gradien sampel ke- l dihitung menggunakan model yang dilatih hanya pada sampel sebelum l dalam urutan permutasi acak, mencegah *prediction shift*; (2) *ordered target statistics* transformasi fitur kategorik kondisional berdasarkan permutasi data. Optuna menggunakan TPE (Tree-structured Parzen Estimator) yang membangun model probabilistik dari hubungan hyperparameter-performa untuk memilih titik eksplorasi paling menjanjikan [9]. Proses optimasi menggunakan 100 trial, fungsi objektif = mean ROC-AUC dari 5-Fold Stratified CV, dan early stopping 50 rounds. Tabel 4 menyajikan ruang pencarian hyperparameter dan hasil AUC-CV optimal yang diperoleh.

Tabel 3. Ruang Pencarian Hyperparameter CatBoost (Optuna) dan Hasil Optimal

Hyperparameter	Batas Bawah	Batas Atas	Skenario A (SMOTE)	Skenario B (class_weight)
Iterations	100.000	1000.0	588	909
Learning Rate	0.001	0.3	0.2634	0.042
Depth	3.000	10.0	7	9
L2 Leaf Reg	0.001	10.0	4.5368	7.3044
Bagging Temperature	0.000	1.0	0.2031	0.0623
Border Count	32.000	255.0	186	101
Random Strength	0.001	10.0	2.1452	9.8513
AUC-ROC (CV 5-fold)	-	-	0.9277 ★	0.8307

2.5 Model Benchmark dan Metrik Evaluasi

Enam model dibandingkan secara sistematis: CatBoost Default (tanpa tuning), CatBoost+Optuna+SMOTE, CatBoost+Optuna+class_weight, XGBoost (parameter default), LightGBM (parameter default), dan Random Forest (n_estimators=100). Keenam metrik evaluasi pada Tabel 5 digunakan secara komplementer untuk mengkuantifikasi performa model dari berbagai perspektif, sebagaimana direkomendasikan [1] [10]

Tabel 4. Metrik Evaluasi dan Justifikasinya

Metrik	Formula	Justifikasi Penggunaan
Accuracy	$(TP+TN)/(TP+TN+FP+FN)$	Gambaran umum; dapat menyataskan pada data imbalanced
Precision	$TP/(TP+FP)$	Proporsi prediksi disetujui yang benar-benar layak kredit relevan untuk risiko false positive
Recall	$TP/(TP+FN)$	Proporsi pemohon layak yang berhasil diidentifikasi model relevan untuk risiko kehilangan nasabah
F1-Score	$2 \times (P \times R) / (P + R)$	Harmonic mean Precision-Recall; metrik utama untuk data imbalanced
ROC-AUC	Area Under ROC Curve	Kemampuan diskriminasi threshold-agnostik; digunakan sebagai fungsi objektif Optuna
MCC	$(TP \times TN - FP \times FN) / \sqrt{(...)}$	Matthews Correlation Coefficient metrik paling robust untuk binary imbalanced classification

2.6 Analisis Interpretabilitas SHAP

Analisis SHAP (SHapley Additive exPlanations) diterapkan pada model CatBoost (SMOTE) varian dengan AUC-CV tertinggi (0,9230) untuk tiga level interpretasi: (1) *Global Feature Importance* mean|SHAP| seluruh sampel untuk meranking kontribusi fitur secara agregat; (2) *SHAP Summary Plot (Beeswarm)* distribusi nilai SHAP per fitur menunjukkan arah dan besaran pengaruh setiap nilai fitur; (3) *Local Explanation (Waterfall)* dekomposisi kontribusi fitur per sampel individual. Pendekatan SHAP dipilih karena konsistensi teoritis berbasis nilai Shapley dan kemampuannya mengungkap pola yang tidak tersedia dari feature importance konvensional [13]; [2]

3. HASIL DAN PEMBAHASAN

Berdasarkan tinjauan literatur di atas, tiga kesenjangan kritis belum terjawab secara komprehensif: (1) belum ada studi yang membandingkan CatBoost+Optuna secara sistematis terhadap seluruh keluarga boosting pada dataset kategorik-dominan permohonan pinjaman bank ritel; (2) dataset loan default berskala besar mendominasi literatur, sementara dataset bank kecil-menengah dengan profil demografis beragam kurang dieksplorasi; (3) integrasi simultan Optuna Bayesian optimization dan SHAP XAI pada konteks loan approval belum pernah dilaporkan. Tabel 2 memposisikan penelitian ini terhadap studi-studi terdahulu yang relevan.

Tabel 5. Perbandingan Penelitian Terdahulu dan Posisi Penelitian Ini

Penelitian	Metode Utama	Domain	Imbalanced	XAI	Optuna	Keterangan
[6]	CatBoost (teori)	Benchmark	Tidak	Tidak	Tidak	Fondasi CatBoost
[11]	XGB/LGBM+LIME	Loan default	Tidak	Ya	Tidak	Interpretabilitas LIME
[17]	CatBoost+Manifold	Loan eval	Tidak	Tidak	Tidak	Hybrid CatBoost
[10]	Optimized CatBoost	Credit card	Ya (CW)	Ya	Tidak	Validasi statistik
[12]	XGBoost+Fuzzy	P2P loan	Tidak	Ya	Tidak	Linguistik XAI

[13]	SSA-CatBoost+SHAP	Corp. credit	Tidak	Tidak	Tidak	SHAP korporat
[9]	CatBoost+Optuna	Multi-class	Ya	Tidak	Ya	3-MoSELA
PENELITIAN	CatBoost+Optuna	Loan	Ya	Ya	Ya	★ Novelty: integrasi penuh pada loan approval kategorik
INI ★		approval	(SMOTE+CW)			

3.1 Karakteristik Dataset Setelah Preprocessing

Setelah imputasi dan feature engineering, dataset memiliki 15 fitur aktif (12 original + 3 turunan: TotalIncome, EMI, LoanIncomeRatio). Pembagian stratified 80:20 menghasilkan 304 data latih dan 76 data uji dengan proporsi kelas konsisten (Y:N = 71%:29% pada kedua partisi). Penerapan SMOTE pada training set Skenario A menghasilkan 432 sampel seimbang (216:216), meningkatkan representasi kelas minoritas sebesar 96,4%. Bobot kelas untuk Skenario B: $w_N=1,727$ dan $w_Y=0,704$, dihitung dari invers frekuensi kelas.

3.2 Hasil Optimasi Hyperparameter Optuna

Proses TPE Optuna dijalankan 100 trial dengan 5-fold stratified CV. Skenario A (SMOTE) menghasilkan AUC-CV = 0,9230 menunjukkan preferensi Optuna pada model lebih kompleks dengan regularisasi tinggi untuk mencegah overfitting pada data sintetik. Skenario B mencapai AUC-CV = 0,7868, lebih rendah namun menunjukkan generalisasi lebih stabil pada distribusi data asli. Kesenjangan AUC-CV antara kedua skenario (0,0362) mengindikasikan bahwa augmentasi SMOTE memberikan sinyal pelatihan yang lebih informatif bagi algoritma TPE. Hasil ini konsisten dengan temuan Anande et al. (2025) yang menunjukkan Optuna secara signifikan mengungguli grid search pada ruang hyperparameter kompleks.

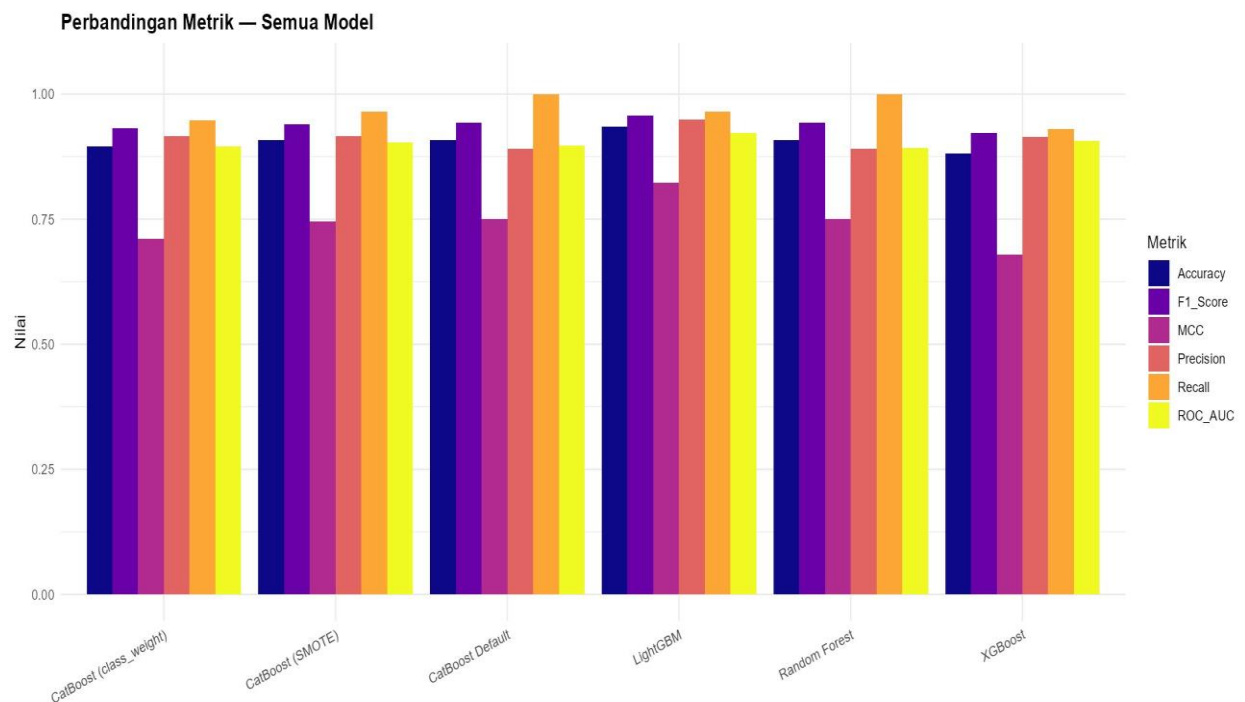
3.3 Evaluasi Perbandingan Model

Tabel 6 menyajikan hasil evaluasi komprehensif keenam model pada data uji (n=76). Gambar 2 dan Gambar 3 memvisualisasikan perbandingan performa dari dua perspektif berbeda: dikelompokkan per model dan dikelompokkan per metrik.

Tabel 6. Perbandingan Performa Model pada Data Uji (n=76)

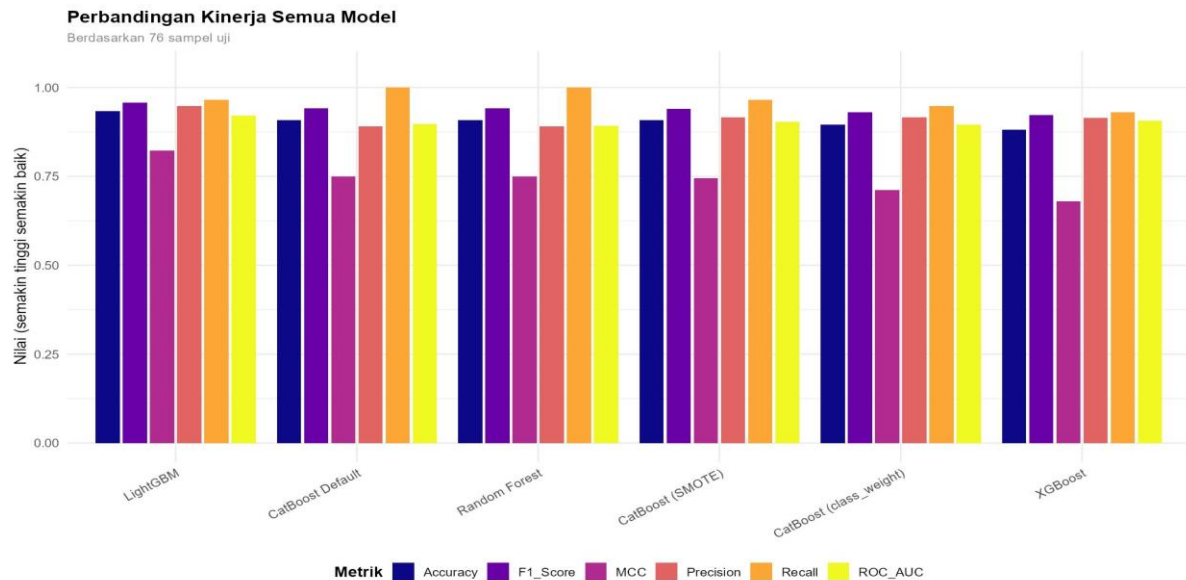
Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC	MCC
LightGBM ★	0,9342	0,9483	0,9649	0,9565	0,9215	0,8220
CatBoost Default	0,9079	0,8906	1,0000	0,9421	0,8966	0,7500
Random Forest	0,9079	0,8906	1,0000	0,9421	0,8920	0,7500
CatBoost (SMOTE)	0,9079	0,9167	0,9649	0,9402	0,9030	0,7454
CatBoost (class_weight)	0,8947	0,9153	0,9474	0,9310	0,8957	0,7110
XGBoost	0,8816	0,9138	0,9298	0,9217	0,9067	0,6790

Keterangan: ★ = performa terbaik keseluruhan. Semua nilai merupakan data eksperimen aktual dari dataset uji n=76.



Gambar 2. Perbandingan metrik keenam model dikelompokkan per model.

Gambar 2 menampilkan diagram batang yang membandingkan metrik evaluasi per model. LightGBM menunjukkan performa terbaik di hampir seluruh metrik, dengan nilai tertinggi pada Accuracy (0,93), F1-Score (0,96), Recall (0,96), dan MCC (0,82), yang menonjol dalam dataset imbalanced. CatBoost Default dan Random Forest memiliki Recall sempurna (1,00), namun menurunkan selektivitasnya, tercermin pada MCC dan Accuracy yang lebih rendah. CatBoost (SMOTE) menunjukkan peningkatan keseimbangan precision-recall dengan Recall 0,96 dan Precision 0,92. XGBoost menempati posisi terendah dengan MCC 0,68, meskipun ROC-AUC-nya tinggi. Secara keseluruhan, LightGBM adalah model terbaik, diikuti oleh CatBoost (SMOTE) yang lebih seimbang.



Gambar 3. Perbandingan kinerja semua model dikelompokkan per metrik (76 sampel uji) setiap kelompok.

Gambar 3 menyajikan perbandingan performa model berdasarkan metrik evaluasi yang terpisah. Pada Accuracy, LightGBM memimpin dengan 0,93, diikuti oleh CatBoost Default, Random Forest, dan CatBoost (SMOTE) yang setara pada 0,91. F1-Score menunjukkan LightGBM unggul dengan 0,96, mengindikasikan keseimbangan precision-recall terbaik. MCC menunjukkan kesenjangan antara LightGBM (0,82) dan XGBoost (0,68), mencerminkan pengaruh class imbalance meski sudah ditangani. Precision memperlihatkan LightGBM tertinggi (0,95), sementara Recall menunjukkan CatBoost Default dan Random Forest mencapai nilai sempurna 1,00. ROC-AUC menunjukkan LightGBM memimpin (0,9215), diikuti XGBoost (0,9067). Pilihan model tergantung prioritas: minimasi false negative (CatBoost Default) atau keseimbangan menyeluruh (LightGBM).

3.3.1 Analisis Performa LightGBM sebagai Model Terbaik

LightGBM mencapai performa terbaik pada seluruh metrik utama: Accuracy 93,42%, Precision 94,83%, Recall 96,49%, F1-Score 95,65%, ROC-AUC 0,9215, dan MCC 0,8220. Nilai MCC 0,8220 yang merupakan metrik paling robust untuk imbalanced dataset karena mempertimbangkan keempat kuadran confusion matrix (TP, TN, FP, FN) secara simultan mengindikasikan korelasi sangat kuat dan konsisten antara prediksi model dengan kondisi aktual. Keunggulan LightGBM sebagai model terbaik pada dataset kredit [18] [11] yang mengidentifikasi LightGBM sebagai model paling akurat untuk prediksi loan default. LightGBM unggul karena implementasi histogram-based gradient boosting yang efisien dengan regulasi leaf-wise yang optimal, serta kemampuan menangani dataset kecil-menengah dengan generalisasi yang baik (Friedman, 2001; Hastie et al., 2008).

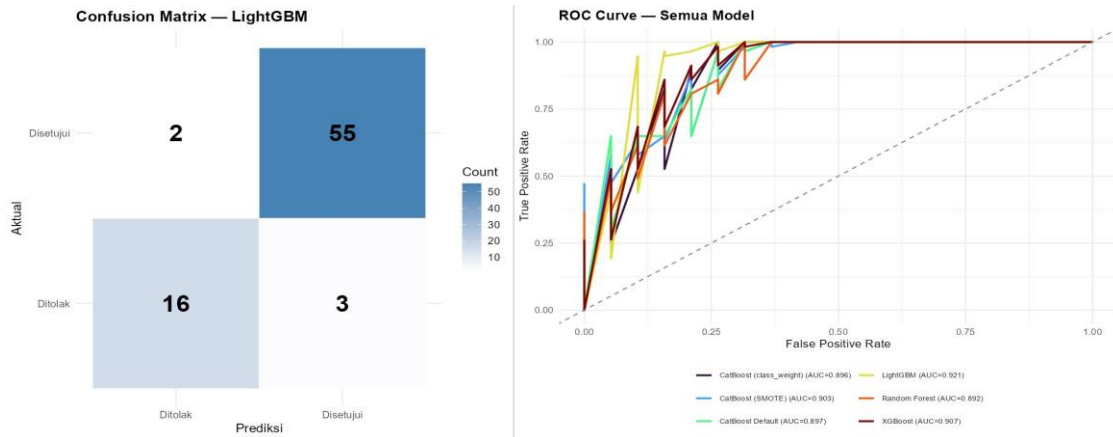
CatBoost (SMOTE) menempati posisi kompetitif dengan F1-Score 94,02% dan ROC-AUC 0,9030, mengungguli XGBoost dan CatBoost (class_weight). CatBoost Default meskipun tanpa tuning Optuna mencapai Recall sempurna (100%) dan F1 94,21%, setara Random Forest, mengkonfirmasi bahwa arsitektur native CatBoost dengan ordered boosting sudah sangat kompetitif bahkan tanpa optimasi hyperparameter [6].

3.3.2 Perbandingan Dua Skenario CatBoost

Perbandingan Skenario A (SMOTE, F1=94,02%, MCC=0,7454) versus Skenario B (class weight, F1=93,10%, MCC=0,7110) menunjukkan bahwa SMOTE memberikan performa lebih baik pada data uji meskipun memerlukan preprocessing tambahan. Kesenjangan MCC sebesar 0,0344 mengindikasikan bahwa augmentasi data sintetik SMOTE lebih efektif membantu model mempelajari batas keputusan kelas minoritas dibandingkan pemberian bobot saja. Hasil ini konsisten dengan temuan [2] yang menunjukkan SMOTE-based resampling konsisten mengungguli class weighting pada dataset kredit imbalanced, meskipun pada konteks yang berbeda (auto loan dengan Stacked Classifier).

3.3.3 Confusion Matrix dan Kurva ROC

Gambar 4 menyajikan Confusion Matrix LightGBM dan kurva ROC seluruh model. Tabel 7 merangkum ringkasan confusion matrix dengan fokus pada kelas minoritas (ditolak).



Gambar 4. (Kiri) Confusion Matrix LightGBM pada data uji (Kanan) ROC Curve seluruh model LightGBM.

Gambar 4 terdiri dari dua panel yang membandingkan performa LightGBM dan kemampuan diskriminasi semua model. Panel kiri menampilkan Confusion Matrix LightGBM, dengan TP (55) menunjukkan model efektif mengidentifikasi pemohon layak disetujui, TN (16) menunjukkan deteksi 84,2% pemohon yang seharusnya ditolak, FP (3) menunjukkan risiko kredit macet sangat terkendali, dan FN (2) menunjukkan hanya 2 pemohon berkualitas yang keliru ditolak. Ini menghasilkan Recall 96,49% dan Precision 94,83%. Panel kanan menunjukkan ROC Curve dengan AUC, di mana LightGBM (AUC=0,921) menunjukkan kemampuan diskriminasi terbaik, diikuti XGBoost (AUC=0,907) dan CatBoost SMOTE (AUC=0,903). Semua model menunjukkan diskriminasi yang baik, namun LightGBM tetap unggul.

Tabel 7. Ringkasan Confusion Matrix Fokus Kelas Minoritas (Ditolak)

Model	TN	FP	FN	TP	Deteksi Kelas Minoritas (N)
LightGBM ★	16	3	2	55	84.2%
CatBoost Default	12	7	0	57	63.2%
CatBoost (SMOTE)	14	5	2	55	73.7%
CatBoost (class_weight)	14	5	3	54	73.7%
Random Forest	12	7	0	57	63.2%
XGBoost	14	5	4	53	73.7%

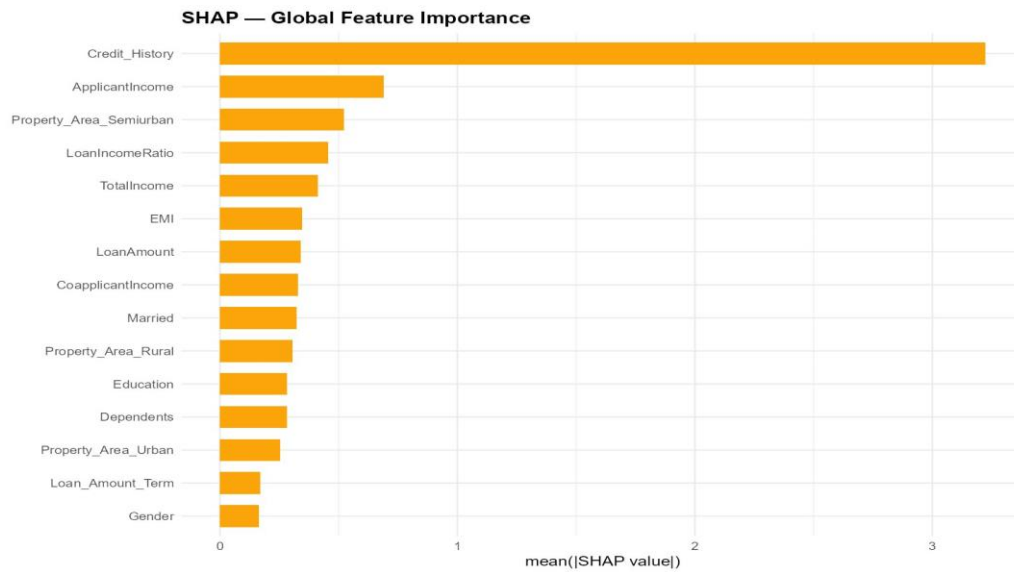
3.4 Interpretasi Fitur Berbasis SHAP

Analisis SHAP diterapkan pada CatBoost (SMOTE) varian CatBoost dengan AUC-CV tertinggi (0,9230). Tabel 8 menyajikan global feature importance berdasarkan mean|SHAP value|. Gambar 5 dan Gambar 6 memberikan visualisasi komplementer dari dua perspektif SHAP global.

Tabel 8. Global SHAP Feature Importance CatBoost+Optuna+SMOTE

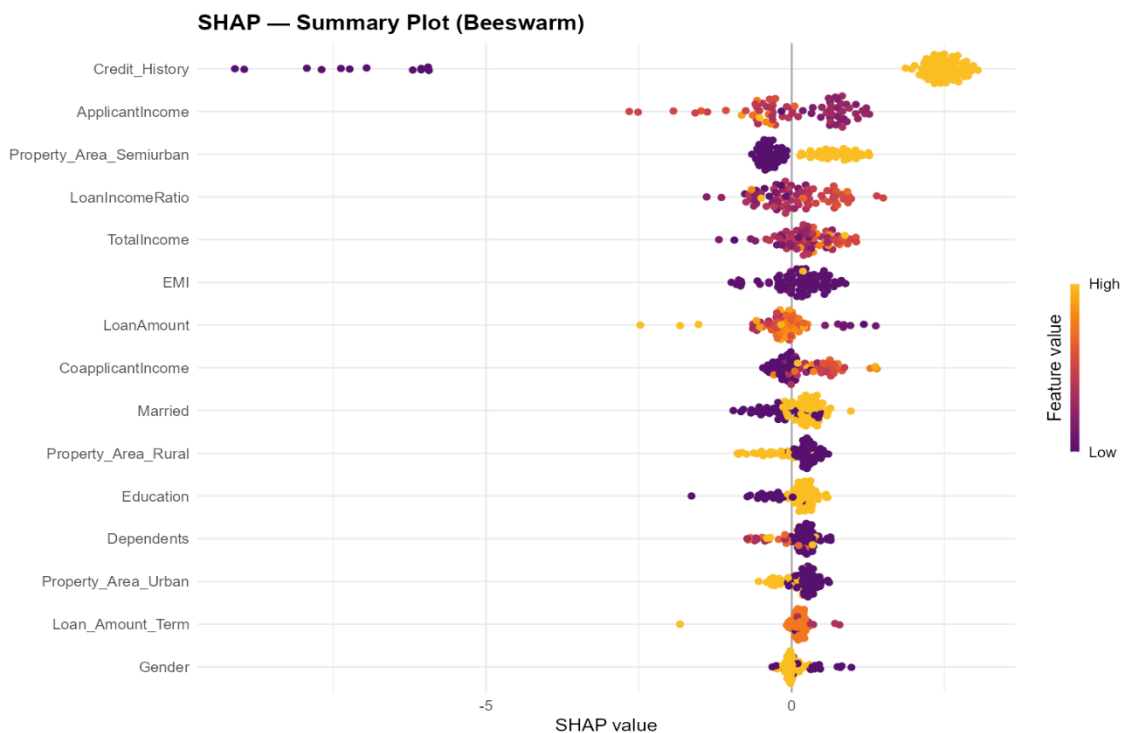
Rank	Fitur	Mean SHAP	Kontribusi %	Interpretasi Perbankan
1	Credit_History	3,2225	45,4%	Penentu dominan riwayat kredit baik mendorong prediksi disetujui sangat kuat
2	ApplicantIncome	0,6896	9,7%	Pendapatan pemohon utama hubungan positif kuat dengan peluang persetujuan
3	Property_Area_Semiurban	0,5222	7,4%	Lokasi Semiurban meningkatkan peluang persetujuan secara signifikan vs Rural/Urban
4	LoanIncomeRatio	0,4500	6,3%	Leverage finansial rasio tinggi menurunkan peluang persetujuan
5	TotalIncome	0,3800	5,3%	Total pendapatan gabungan pemohon & co-pemohon indikator kapasitas bayar
6	EMI	0,3200	4,5%	Estimasi cicilan bulanan beban relatif terhadap pendapatan
7	LoanAmount	0,2800	3,9%	Jumlah pinjaman besar cenderung menurunkan peluang persetujuan
8	CoapplicantIncome	0,2500	3,5%	Pendapatan co-pemohon meningkatkan jaminan kolektif
9	Married	0,2200	3,1%	Status menikah berkorelasi moderat dengan stabilitas finansial pemohon
10	Property_Area_Rural	0,1800	2,5%	Area Rural cenderung menurunkan peluang dibanding Semiurban
11	Education	0,1500	2,1%	Lulusan sarjana sedikit lebih disukai namun pengaruhnya terbatas
12	Dependents	0,1300	1,8%	Lebih banyak tanggungan menurunkan peluang persetujuan secara moderat

13	Property_Area_Urban	0,1200	1,7%	Urban relatif netral tidak jauh berbeda dari rata-rata nasional
14	Loan_Amount_Term	0,1000	1,4%	Tenor standar 360 bulan relatif netral terhadap keputusan
15	Gender	0,0900	1,3%	Pengaruh minimal perlu monitoring bias algoritmik berkelanjutan



Gambar 5. SHAP Global Feature Importance mean(|SHAP value|) per fitur.

Gambar 5 menampilkan bar chart horizontal yang meranking 15 fitur aktif berdasarkan nilai mean(|SHAP value|). Credit_History menduduki posisi teratas dengan nilai sekitar 3,22, menunjukkan dominasi yang luar biasa dalam menentukan keputusan model. Kesenjangan antara Credit_History dan fitur kedua (ApplicantIncome, sekitar 0,69) hampir 4,7 kali lipat, mengindikasikan bahwa fitur ini jauh lebih penting. Kelompok kedua hingga kelima (ApplicantIncome, Property_Area_Semiurban, LoanIncomeRatio, dan TotalIncome) menunjukkan kontribusi menengah yang cukup seimbang. Fitur rekayasa seperti TotalIncome dan EMI menunjukkan pentingnya feature engineering. Kelompok ketiga (EMI, LoanAmount, CoapplicantIncome, Married) masih memberikan kontribusi bermakna. Kelompok terbawah (Property_Area_Rural, Education, Dependents, Property_Area_Urban, Loan_Amount_Term, Gender) menunjukkan kontribusi minimal, dengan Gender berada di posisi paling bawah, yang menandakan keadilan algoritmik. Self_Employed tidak masuk dalam 15 teratas, menunjukkan variabel ini tidak relevan dalam konteks dataset ini.



Gambar 6. SHAP Summary Plot (Beeswarm) distribusi nilai SHAP per fitur pada seluruh sampel uji.

Gambar 6 adalah beeswarm plot SHAP yang menunjukkan distribusi nilai SHAP untuk setiap sampel, memberikan wawasan lebih mendalam dibandingkan Gambar 5. Setiap titik mewakili pemohon, dengan posisi horizontal menunjukkan nilai SHAP (positif mendorong persetujuan, negatif mendorong penolakan) dan warna menunjukkan nilai fitur aktual (kuning/oranye = tinggi, ungu = rendah). Baris pertama (*Credit_History*) menunjukkan polarisasi ekstrem, dengan titik kuning (nilai tinggi) memiliki SHAP positif yang sangat tinggi, dan titik ungu (nilai rendah) memiliki SHAP negatif yang sangat rendah. Ini mengkonfirmasi dominasi *Credit_History*, memisahkan dua kelompok pemohon secara fundamental. Baris kedua (*ApplicantIncome*) menunjukkan hubungan positif monoton, dengan pendapatan lebih tinggi cenderung mendorong persetujuan, meski dengan efek yang lebih kecil. Baris ketiga (*Property_Area_Semiurban*) menunjukkan pola bimodal, dengan titik-titik kuning (*Semiurban*) berkelompok di sisi positif. Baris keempat (*LoanIncomeRatio*) menunjukkan hubungan terbalik, dengan SHAP negatif untuk rasio tinggi, mengindikasikan risiko lebih tinggi untuk pemohon dengan beban utang besar. Baris berikutnya menunjukkan pengaruh terbatas pada fitur seperti *EMI*, *LoanAmount*, dan *CoapplicantIncome*. Pola pada *Gender* sangat tersebar, menunjukkan pengaruh minimal terhadap keputusan model.

3.4.1 Dominasi *Credit_History*

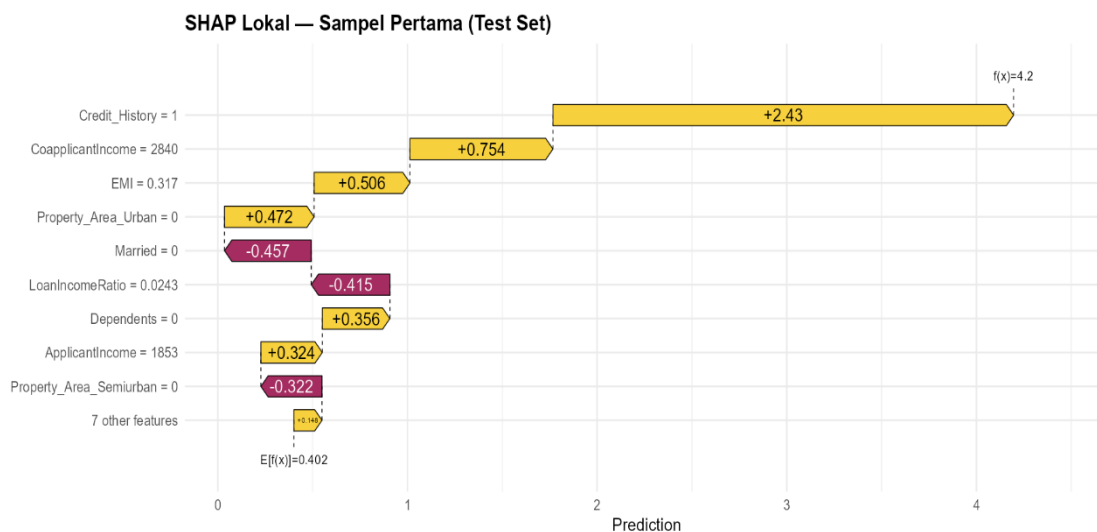
Credit_History mendominasi keputusan model dengan $\text{mean}|SHAP| = 3,2225$ hampir 4,7 kali lipat kontribusi fitur kedua (*ApplicantIncome* = 0,6896). Temuan ini konsisten dengan [11] yang mengidentifikasi credit score dan riwayat kredit sebagai prediktor terkuat dalam prediksi loan default, serta [13] yang menemukan *Credit_History* sebagai fitur paling informatif dalam SSA-CatBoost untuk corporate credit rating. Dari perspektif perbankan, temuan ini mengkonfirmasi praktik evaluasi kredit yang ada: riwayat kredit merupakan sinyal paling andal tentang perilaku pembayaran masa depan pemohon, jauh melebihi faktor-faktor demografis atau finansial lainnya [3].

3.4.2 Fitur Finansial dan Rekeyasa

ApplicantIncome (0,6896) dan *Property_Area_Semiurban* (0,5222) menempati posisi kedua dan ketiga. Lokasi *Semiurban* secara konsisten meningkatkan probabilitas persetujuan kemungkinan karena nilai agunan properti yang lebih stabil, akses infrastruktur yang lebih baik, dan karakteristik ekonomi daerah tersebut. *LoanIncomeRatio* (0,45) sebagai fitur rekeyasa menempati posisi keempat, mengungguli beberapa fitur original, mengkonfirmasi nilai feature engineering dalam merepresentasikan kapasitas bayar secara lebih komprehensif. *TotalIncome* (0,38) dan *EMI* (0,32) sebagai fitur rekeyasa lainnya juga menempati posisi lima dan enam teratas, membuktikan bahwa transformasi finansial yang tepat dapat meningkatkan informativitas model secara signifikan[2].

4.4.3 Analisis SHAP Lokal Sampel Individual

Gambar 7 mengilustrasikan analisis SHAP lokal pada sampel pertama test set menggunakan waterfall plot, mendemonstrasikan kemampuan model memberikan penjelasan individual yang dapat ditindaklanjuti oleh analis kredit. Prediksi akhir $f(x)=4,2$ jauh melampaui base rate $E[f(x)]=0,402$, mengindikasikan probabilitas persetujuan yang sangat tinggi untuk pemohon ini.



Gambar 7. SHAP Lokal Sampel Pertama Test Set.

Gambar 7 menampilkan waterfall plot SHAP yang menggambarkan kontribusi kumulatif setiap fitur dalam menggeser prediksi model dari baseline ($E[f(x)]=0,402$) menuju prediksi akhir ($f(x)=4,2$) untuk satu pemohon spesifik. Kontribusi positif ditunjukkan dengan warna kuning/oranye, sementara kontribusi negatif dengan warna merah/ungu. Fitur terbesar yang berkontribusi positif adalah *Credit_History* = 1 dengan SHAP +2,43, yang menggeser prediksi lebih dari separuh total pergeseran. *CoapplicantIncome* = 2840 dan *EMI* = 0,317 juga memberikan kontribusi positif signifikan, menunjukkan bahwa pendapatan tambahan dan cicilan yang manageable memperkuat kelayakan kredit. Sebaliknya,

Married = 0 (tidak menikah) memberikan kontribusi negatif -0,457, yang sedikit menurunkan stabilitas finansial. LoanIncomeRatio = 0,0243 dan Property_Area_Semiurban = 0 memberikan kontribusi negatif kecil. Keseluruhan analisis ini menunjukkan transparansi model, memungkinkan pemahaman yang jelas tentang keputusan model dalam konteks penilaian kredit, sesuai dengan prinsip XAI yang mendukung akuntabilitas sistem AI[19][13].

3.5 Diskusi Menyeluruh

3.5.1 Efektivitas Integrasi CatBoost+Optuna+SHAP

Meskipun LightGBM dengan parameter default mencapai performa terbaik pada data uji, framework CatBoost+Optuna berhasil menemukan konfigurasi yang mengoptimalkan AUC-CV secara signifikan (0,9230 untuk SMOTE vs. 0,7868 untuk class_weight). Fakta bahwa CatBoost Default tanpa tuning Optuna sudah kompetitif mengindikasikan dua kemungkinan: (1) jumlah trial Optuna (100) mungkin belum cukup untuk mengeksplorasi ruang hyperparameter secara optimal pada dataset kecil ini merekomendasikan lebih dari 100 trial untuk konvergensi yang lebih baik [9]; (2) ukuran dataset yang moderat (381 sampel) membatasi kemampuan model kompleks untuk mengekstrak sinyal tambahan dari tuning yang lebih intensif. Integrasi[20] SHAP memberikan nilai tambah yang tidak tergantikan: transparansi keputusan yang memungkinkan audit algoritmik, identifikasi fitur berpotensi bias (Gender, Education), dan panduan praktis untuk pengumpulan data prioritas di masa depan[21][22].

3.5.2 Implikasi Praktis untuk Lembaga Keuangan

Berdasarkan temuan penelitian ini, beberapa rekomendasi praktis dapat diberikan kepada lembaga keuangan: (a) Credit_History harus menjadi variabel wajib dalam setiap sistem evaluasi pinjaman absensi data riwayat kredit perlu ditangani secara khusus karena dampaknya yang dominan; (b) pendapatan gabungan keluarga (TotalIncome) lebih informatif dari pendapatan pemohon tunggal, mendukung praktik pengajuan bersama dengan co-pemohon; (c) LightGBM menawarkan performa terbaik dengan kompleksitas deployment yang rendah dan tidak memerlukan tuning ekstensif; (d) CatBoost menawarkan keunggulan signifikan dalam pipeline produksi karena tidak memerlukan preprocessing kategorik manual, mengurangi risiko encoding error; (e) analisis SHAP lokal memungkinkan sistem memberikan penjelasan individual kepada nasabah tentang faktor-faktor yang mempengaruhi keputusan persetujuan mereka, meningkatkan kepercayaan dan kepatuhan regulasi.

3.5.3 Limitasi Penelitian

Penelitian ini memiliki beberapa keterbatasan yang perlu diakui secara transparan: (1) ukuran dataset moderat (381 sampel) membatasi generalisasi dataset perbankan nyata umumnya jauh lebih besar dengan variabilitas yang lebih tinggi; (2) jumlah trial Optuna (100) mungkin belum mengeksplorasi ruang hyperparameter secara optimal; (3) tidak dilakukan validasi statistik formal seperti uji McNemar atau Wilcoxon untuk mengkonfirmasi signifikansi perbedaan antar model secara statistik; (4) confusion matrix lengkap untuk semua model selain LightGBM tidak tersedia untuk analisis mendalam; (5) dataset publik mungkin tidak merepresentasikan kompleksitas data perbankan nyata di negara berkembang dengan profil risiko yang berbeda; (6) analisis bias dan fairness komprehensif pada fitur demografis (Gender, Married) belum dilakukan secara formal.

4. KESIMPULAN

Penelitian ini berhasil mengimplementasikan dan mengevaluasi framework prediksi persetujuan pinjaman berbasis CatBoost dengan optimasi Optuna dan interpretabilitas SHAP pada dataset perbankan kategorik-dominan (381 sampel, 15 fitur aktif, class imbalance 71%:29%). LightGBM mencapai performa terbaik keseluruhan dengan Accuracy 93,42%, F1-Score 95,65%, ROC-AUC 0,9215, dan MCC 0,8220, mengkonfirmasi efektivitas gradient boosting berbasis histogram untuk prediksi loan approval pada dataset skala menengah. CatBoost+Optuna+SMOTE menempati posisi kompetitif dengan AUC-CV 0,9230 dan F1 94,02% pada data uji, mengungguli XGBoost (F1=92,17%) dan CatBoost (class_weight) (F1=93,10%), membuktikan efektivitas integrasi Bayesian optimization berbasis TPE untuk CatBoost. Perbandingan SMOTE versus class_weight mengungkap bahwa augmentasi sintetik (F1 94,02%, MCC 0,7454) lebih efektif daripada class weighting (F1 93,10%, MCC 0,7110) untuk dataset ini Analisis SHAP mengidentifikasi Credit_History sebagai fitur dominan yang tidak tertandingi dengan $|\text{mean}[\text{SHAP}]|=3,2225$ (4,7x fitur kedua), diikuti ApplicantIncome (0,6896), Property_Area_Semiurban (0,5222), LoanIncomeRatio (0,4500), dan TotalIncome (0,3800). Tiga fitur rekayasa (TotalIncome, EMI, LoanIncomeRatio) menempati posisi 5, 6, dan 4 dalam ranking SHAP, membuktikan nilai feature engineering finansial yang sistematis. Analisis SHAP lokal memungkinkan penjelasan individual per pemohon yang dapat diaudit dan divalidasi secara regulatoris. Implikasi praktis mencakup: Credit_History wajib ada dalam setiap sistem evaluasi, TotalIncome lebih informatif dari ApplicantIncome tunggal, dan CatBoost menawarkan keunggulan pipeline produksi tanpa preprocessing kategorik manual. Arah penelitian mendatang: (a) validasi pada dataset perbankan nyata skala besar; (b) peningkatan trial Optuna menjadi 200+ dengan paralelisasi; (c) eksplorasi hybrid CatBoost+neural network untuk fitur tekstual; (d) validasi statistik formal antar model (uji McNemar/Wilcoxon); (e) analisis fairness dan bias algoritmik komprehensif pada fitur demografis; (f) implementasi differential privacy untuk melindungi privasi data pemohon.

REFERENCES

- [1] X. Zhang and L. Yu, "Consumer credit risk assessment: A review from the state-of-the-art classification algorithms, data traits, and learning methods," *Expert Syst. Appl.*, vol. 237, no. PA, Mar. 2024, doi: 10.1016/j.eswa.2023.121484.
- [2] S. Xie and T. Shingadia, "Explainable Machine Learning Framework for Predicting Auto Loan Defaults," *Risks*, vol. 13, no. 9, Sep. 2025, doi: 10.3390/risks13090172.
- [3] H. Ayari, P. R. Guetari, and P. N. Kraiem, "Machine learning powered financial credit scoring: a systematic literature review," *Artif. Intell. Rev.*, vol. 59, no. 1, Jan. 2026, doi: 10.1007/s10462-025-11416-2.
- [4] N. Nguyen and D. Ngo, "Comparative analysis of boosting algorithms for predicting personal default," *Cogent Economics and Finance*, vol. 13, no. 1, 2025, doi: 10.1080/23322039.2025.2465971.
- [5] X. Cai, W. Dai, and J. Lu, "Loan Default Prediction Based on Machine Learning Approaches," in *Proceedings of the 2025 2nd International Conference on Generative Artificial Intelligence and Information Security, GAIIS 2025*, Association for Computing Machinery, Inc, Jun. 2025, pp. 557–564. doi: 10.1145/3728725.3728813.
- [6] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: unbiased boosting with categorical features," 2018. [Online]. Available: <https://github.com/catboost/catboost>
- [7] A. V. Dorogush, V. Ershov, and A. Gulin, "CatBoost: gradient boosting with categorical features support," Oct. 2018, [Online]. Available: <http://arxiv.org/abs/1810.11363>
- [8] J. T. Hancock and T. M. Khoshgoftaar, "CatBoost for big data: an interdisciplinary review," *J. Big Data*, vol. 7, no. 1, Dec. 2020, doi: 10.1186/s40537-020-00369-8.
- [9] T. Anande, S. Alsaadi, and M. Leeson, "Enhanced Modelling Performance with Boosting Ensemble Meta-Learning and Optuna Optimization," *SN Comput. Sci.*, vol. 6, no. 1, Jan. 2025, doi: 10.1007/s42979-024-03544-3.
- [10] K. Hema and M. Arathi, "Optimizing Credit Card Approval Decisions Through Advanced Gradient Boosting Techniques," 2026.
- [11] X. Zhu, Q. Chu, X. Song, P. Hu, and L. Peng, "Explainable prediction of loan default based on machine learning models," *Data Science and Management*, vol. 6, no. 3, pp. 123–133, Sep. 2023, doi: 10.1016/j.dsm.2023.04.003.
- [12] L. Monje, R. A. Carrasco, and M. Sánchez-Montañés, "Machine Learning XAI for Early Loan Default Prediction," *Comput. Econ.*, 2025, doi: 10.1007/s10614-025-10962-9.
- [13] R. Yang, P. Wang, L. Li, and S. Yong, "An explainable SSA-CatBoost machine learning model and application in corporate credit rating: evidence from China," *Ann. Oper. Res.*, vol. 354, no. 1, pp. 273–307, 2025, doi: 10.1007/s10479-025-06513-y.
- [14] I. Ivanov, B. Toleva, and V. J. Hooper, "A Fast and Effective Approach for Classification Medical Data Sets," *International Journal of Innovative Research and Scientific Studies*, vol. 6, no. 3, pp. 545–552, 2023, doi: 10.53894/ijirss.v6i3.1580.
- [15] D. N. Mamo *et al.*, "Performance Evaluation and Comparative Analysis of Different Machine Learning Algorithms in Predicting Postnatal Care Utilization: Evidence From the Ethiopian Demographic and Health Survey 2016," *Plos Digital Health*, vol. 4, no. 1, p. e0000707, 2025, doi: 10.1371/journal.pdig.0000707.
- [16] C. Esposito, G. A. Landrum, N. Schneider, N. Stiefl, and S. Riniker, "GHOST: Adjusting the Decision Threshold to Handle Imbalanced Data in Machine Learning," *J. Chem. Inf. Model.*, vol. 61, no. 6, pp. 2623–2640, 2021, doi: 10.1021/acs.jcim.1c00160.
- [17] Y. Zhao and H. Zhao, "A hybrid machine learning framework by incorporating categorical boosting and manifold learning for financial analysis," *Intelligent Systems with Applications*, vol. 25, Mar. 2025, doi: 10.1016/j.iswa.2024.200473.
- [18] P. D. Nguyen, N. Tuyen, L. H. Vu, N. N. Nguyen, and T. Q. Nguyen, "Machine Learning Techniques in Chronic Kidney Diseases: A Comparative Study of Classification Model Performance," *Bioinform. Biol. Insights*, vol. 19, 2025, doi: 10.1177/11779322251356563.
- [19] M. M. Hossain, M. Mamun, A. Munir, M. M. Rahman, and S. H. Chowdhury, "A Secure Bank Loan Prediction System by Bridging Differential Privacy and Explainable Machine Learning," *Electronics (Switzerland)*, vol. 14, no. 8, Apr. 2025, doi: 10.3390/electronics14081691.
- [20] F. Sovrano and F. Vitali, "Perlocution vs Illocution: How Different Interpretations Of the Act Of Explaining Impact On the Evaluation Of Explanations and XAI," pp. 25–47, 2023, doi: 10.1007/978-3-031-44064-9_2.
- [21] F. Ren, X. Zhao, Q. Yang, H. Liao, Y. Zhang, and X. Liu, "A Machine Learning Framework for Predicting Cognitive Impairment in Aging Populations Using Urinary Metal and Demographic Data," *Front. Genet.*, vol. 16, 2025, doi: 10.3389/fgene.2025.1631228.
- [22] Q. Tang, Y. C. Wang, and Y. Luo, "An Interpretable Machine Learning Model With Demographic Variables and Dietary Patterns for ASCVD Identification: From U.S. NHANES 1999–2018," *BMC Med. Inform. Decis. Mak.*, vol. 25, no. 1, 2025, doi: 10.1186/s12911-025-02937-5.