



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 5 No. 2 (2026) pp: 18263-18270

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Perbandingan Algoritma *Gradient Boosting* untuk Prediksi Risiko Kredit Pada Sistem Pendukung Keputusan

Ade irmayanti¹, Apriani², Rifqi Hammad³

¹Program Studi Ilmu Komputer, Fakultas Teknik, Universitas Bumigora

²Fakultas Teknik Program Studi Ilmu Komputer

adeirmayaa@gmail.com*, apriani@universitasbumigora.ac.id, rifqi.hammad@universitasbumigora.ac.id

Abstrak

Penelitian ini bertujuan untuk menerapkan algoritma *Gradient Boosting* dalam sistem pendukung keputusan prediksi risiko kredit berbasis *machine learning*. Penerapan *machine learning* memungkinkan analisis data debitur secara otomatis sehingga dapat meningkatkan ketepatan dalam mengidentifikasi risiko gagal bayar dan mendukung proses pengambilan keputusan kredit. Risiko kredit menjadi salah satu permasalahan utama pada lembaga keuangan karena kesalahan dalam penilaian kelayakan kredit dapat meningkatkan potensi kredit macet. Penelitian ini membandingkan performa algoritma *XGBoost*, *LightGBM*, dan *CatBoost* dalam melakukan klasifikasi risiko kredit dengan menerapkan metode *SMOTE-Tomek* untuk menangani ketidakseimbangan data. Dataset yang digunakan terdiri dari 50.000 data nasabah dengan 18 atribut yang diperoleh dari *OpenDataBay* pada dataset *Loan Default Prediction Challenge*. Tahapan penelitian meliputi *Exploratory Data Analysis (EDA)*, *preprocessing*, seleksi fitur, pembagian data, penyeimbangan data, pemodelan, *hyperparameter tuning*, evaluasi model, dan implementasi sistem berbasis web menggunakan *Python* dan *Streamlit*. Evaluasi model dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan *ROC-AUC*. Hasil penelitian menunjukkan bahwa algoritma *CatBoost* memperoleh nilai *ROC-AUC* sebesar 0,737, lebih tinggi dibandingkan *XGBoost* dan *LightGBM*, serta menghasilkan kinerja yang lebih stabil dalam memprediksi risiko kredit dengan data tidak seimbang. Sistem yang dikembangkan mampu membantu proses pengambilan keputusan kredit secara lebih cepat, akurat, dan efisien sehingga dapat mendukung lembaga keuangan dalam meminimalkan risiko pemberian kredit yang tidak tepat.

Kata Kunci: Risiko Kredit, Machine Learning, Xgboost, Lightgbm, Catboost, SMOTE-Tomek, Sistem Pendukung Keputusan

1. Latar Belakang

Di era digitalisasi yang berkembang pesat, sektor keuangan mengalami transformasi signifikan, khususnya dalam layanan pembiayaan dan kredit. Lembaga keuangan konvensional maupun perusahaan *financial technology* (fintech) saat ini menawarkan layanan pinjaman yang lebih cepat, mudah diakses, dan berbasis sistem digital [1]. Kondisi tersebut dapat meningkatkan inklusi keuangan, namun di sisi lain juga meningkatkan potensi risiko kredit apabila proses penilaian kelayakan pinjaman tidak dilakukan secara optimal [2]. Dengan transformasi ini, inklusi keuangan meningkat dan masyarakat lebih mudah mendapatkan dana. Sebaliknya, jika kemudahan mendapatkan kredit tidak diimbangi dengan sistem penilaian risiko yang tepat, hal itu juga dapat meningkatkan risiko kredit.

Risiko kredit merupakan kemungkinan terjadinya kerugian akibat ketidakmampuan debitur dalam memenuhi kewajiban pembayaran pinjaman sesuai perjanjian yang telah ditetapkan [3]. Perkembangan industri keuangan di Indonesia menunjukkan tren pertumbuhan yang positif, dengan pertumbuhan kredit perbankan sebesar 11,35% pada tahun 2022, 10,38% pada tahun 2023, dan 10,39% pada tahun 2024. Pada Januari 2025, penyaluran kredit perbankan kembali meningkat sebesar 10,27% (*year-on-year*) menjadi Rp7.782 triliun, sementara pembiayaan fintech *peer-to-peer lending* mencapai Rp69,39 triliun pada Juli 2024 dengan pertumbuhan tahunan sebesar 23,97% [4]. Meningkatnya volume penyaluran kredit menuntut lembaga keuangan untuk lebih akurat dalam mengidentifikasi calon debitur yang berpotensi mengalami gagal bayar.

Salah satu pendekatan yang banyak digunakan dalam prediksi risiko kredit adalah *machine learning* [5]. Pendekatan ini mampu mempelajari pola kompleks dari data historis sehingga dapat menghasilkan prediksi yang lebih adaptif dibandingkan metode konvensional [6]. Kemampuan *machine learning* dalam mengolah data dalam jumlah besar serta mengenali pola yang sulit dideteksi secara manual menjadikannya semakin banyak diterapkan pada berbagai sektor, termasuk industri keuangan. Dalam implementasinya, *machine learning* sering digunakan

untuk menganalisis karakteristik debitur, seperti pendapatan, riwayat pembayaran, jumlah pinjaman, dan faktor finansial lainnya untuk memprediksi kemungkinan gagal bayar (*default*) [7]. Hasil prediksi tersebut dapat menjadi dasar bagi lembaga keuangan dalam pengambilan keputusan pemberian kredit sehingga risiko kredit bermasalah dapat diminimalkan.

Algoritma *Gradient Boosting* merupakan teknik *ensemble learning* yang membangun model prediktif secara bertahap dengan mengoptimalkan fungsi *loss* melalui *gradient descent*. Seiring berkembangnya, *Gradient Boosting* menjadi dasar bagi berbagai implementasi modern yang lebih efisien dan mampu menangani data berukuran besar, seperti XGBoost, LightGBM, dan CatBoost, yang banyak digunakan untuk klasifikasi dan prediksi *machine learning* [8]. Ketiga metode ini merupakan algoritma berbasis *Gradient Boosting Decision Tree* (GBDT) yang membangun model secara bertahap melalui penambahan pohon keputusan pada setiap iterasi.

Algoritma *Gradient Boosting* seperti XGBoost, LightGBM, dan CatBoost memiliki kemampuan yang baik dalam menangani data kompleks dan menghasilkan performa klasifikasi yang tinggi [9]. XGBoost dikenal unggul dalam optimasi dan regularisasi untuk mengurangi *overfitting* [10], LightGBM memiliki keunggulan dalam efisiensi komputasi [11], sedangkan CatBoost mampu menangani fitur kategorikal secara langsung tanpa proses *encoding* yang kompleks [12]. Perbedaan karakteristik tersebut memungkinkan adanya variasi performa dalam prediksi risiko kredit sehingga diperlukan perbandingan yang komprehensif untuk menentukan model terbaik.

Namun, penerapan *machine learning* pada prediksi risiko kredit umumnya menghadapi permasalahan ketidakseimbangan data (*class imbalance*), yaitu jumlah data debitur gagal bayar lebih sedikit dibandingkan debitur tidak gagal bayar. Kondisi tersebut menyebabkan model cenderung bias terhadap kelas mayoritas sehingga performa klasifikasi pada kelas minoritas menjadi kurang optimal [13]. Untuk mengatasi masalah tersebut, penelitian ini menerapkan metode *Synthetic Minority Over-sampling Technique* (SMOTE) yang mampu menghasilkan data sintesis pada kelas minoritas sehingga distribusi data menjadi lebih seimbang [14].

Beberapa penelitian menunjukkan bahwa algoritma *gradient boosting* memiliki kinerja yang baik dalam prediksi risiko kredit. XGBoost memperoleh nilai AUC sebesar 0,9283 dan presisi 96,4% dalam prediksi gagal bayar pinjaman [15]. Penelitian dengan algoritma XGBoost dan CatBoost menunjukkan bahwa penerapan *SMOTE-Tomek* efektif dalam menangani ketidakseimbangan data [16]. Kombinasi XGBoost dan *SMOTE* bekerja dengan baik untuk menangani data keuangan yang tidak seimbang menunjukkan kemampuan untuk meningkatkan deteksi kelas minoritas secara signifikan, dengan akurasi 96% dan AUC 0,96 [17]. Selain itu, berbagai penelitian menunjukkan bahwa LightGBM unggul dari aspek efisiensi komputasi pada dataset berukuran besar, sedangkan CatBoost mampu menghasilkan performa yang lebih konsisten pada dataset dengan dominasi fitur kategorikal [18]. Namun penelitian yang membandingkan secara langsung kinerja XGBoost, LightGBM, dan CatBoost pada prediksi risiko kredit masih terbatas.

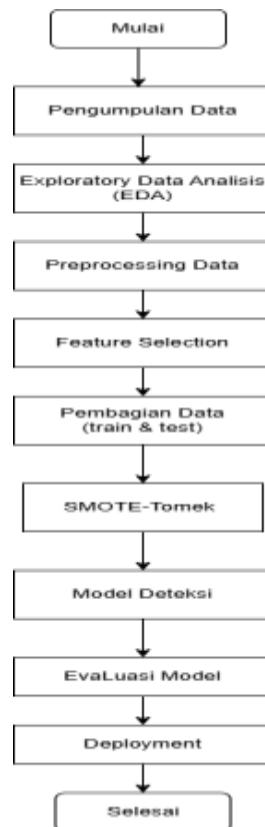
Meskipun demikian, sebagian besar penelitian terdahulu menunjukkan bahwa algoritma *Gradient Boosting* mampu memberikan performa yang baik, sebagian besar penelitian masih berfokus pada satu evaluasi algoritma atau hanya membandingkan dua algoritma, seperti XGBoost dengan Random Forest, XGBoost dengan LightGBM, atau XGBoost dengan CatBoost. Selain itu, sebagian besar penelitian hanya menitikberatkan pada peningkatan performa model klasifikasi tanpa mengimplementasikan model ke dalam sistem pendukung keputusan yang dapat digunakan secara langsung oleh pengguna. Dengan demikian, masih terdapat kesenjangan penelitian berupa belum adanya perbandingan yang komprehensif antara XGBoost, LightGBM, dan CatBoost menggunakan proses *preprocessing* yang sama, penanganan ketidakseimbangan data menggunakan *SMOTE-Tomek*, serta implementasi model terbaik ke dalam sistem pendukung keputusan berbasis web.

Berdasarkan kesenjangan tersebut, penelitian ini membandingkan algoritma XGBoost, LightGBM, dan CatBoost dalam prediksi risiko kredit menggunakan *SMOTE-Tomek* untuk menangani ketidakseimbangan data. Evaluasi dilakukan menggunakan *accuracy*, *precision*, *recall*, *f1-score*, dan *ROC-AUC*, kemudian model terbaik diimplementasikan ke dalam sistem pendukung keputusan berbasis Streamlit untuk mendukung analisis kelayakan kredit secara lebih cepat, akurat, dan berbasis data.

2. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen. Tahapan penelitian dimulai dari pengumpulan data, *Exploratory Data Analysis* (EDA), *preprocessing*, pembagian data, penyeimbangan data

menggunakan *SMOTE-Tomek*, seleksi fitur, pembangunan model, *hyperparameter tuning*, evaluasi model, hingga implementasi sistem berbasis web Streamlit. Berikut Gambar 1. Alur penelitian.



Gambar 1. Alur Penelitian

2.1. Pengumpulan Data

Dataset yang digunakan diperoleh dari platform OpenDataBay melalui *Loan Default Prediction Challenge*. Dataset terdiri atas 50.000 sampel dengan 18 atribut yang mencakup informasi demografis, kondisi finansial, dan karakteristik pinjaman debitur. Variabel target adalah atribut *Default* dengan nilai biner: 1 (gagal bayar) dan 0 (tidak gagal bayar). Tampilan awal dataset yang digunakan dalam penelitian ini ditujukan pada Gambar 2.

LoanID	Age	Income	LoanAmount	CreditScore	HasHousing	HasCreditLines	InterestRate	LoanTerm	OTRatio	Education	EmploymentType	MaritalStatus	HasShortage	HasDependents	LoanPurpose	HasCollateral	Default
ISPPQX3996	56	65894	50637	520	86	4	15.23	36	0.44	Bachelor's	Full-time	Divorced	Yes	Yes	Other	Yes	0
HP9K72W478	69	50432	134446	458	15	1	4.81	60	0.68	Master's	Full-time	Married	No	No	Other	Yes	0
C102BCP487	46	84536	129458	451	26	3	21.17	24	0.21	Master's	Unemployed	Divorced	Yes	Yes	Auto	No	1
V2KXSPM4UN	32	31713	44759	743	0	3	7.07	24	0.23	High School	Full-time	Married	No	No	Business	No	0
EY16LHTZF	60	23437	9139	623	8	4	6.54	46	0.75	Bachelor's	Unemployed	Divorced	No	Yes	Auto	No	0
AB9LHQ2VUS	25	93296	30448	720	16	2	22.72	24	0.10	High School	Unemployed	Single	Yes	No	Business	Yes	1
H8GAPAC871	36	111186	177025	429	86	1	19.11	12	0.16	Bachelor's	Unemployed	Single	Yes	No	Home	Yes	0
DHG2QK3RW	56	136802	155111	531	67	4	8.15	60	0.43	PhD	Full-time	Married	No	No	Home	Yes	0
1R0KLG4NLU	36	42063	82357	637	83	1	25.94	48	0.20	Bachelor's	Self-employed	Divorced	Yes	No	Education	No	1
CMBL1GT7ZF	40	132784	226516	480	114	4	9.09	48	0.33	High School	Self-employed	Married	Yes	No	Other	Yes	0

Gambar 2. Dataset

2.2. Exploratory Data Analysis (EDA)

Pada penelitian ini, *Exploratory Data Analysis* (EDA) dilakukan untuk memahami karakteristik data secara menyeluruh sebelum proses pemodelan. Tahap ini mencakup pemeriksaan distribusi variabel numerik dan kategorikal, deteksi *missing value*, analisis korelasi antar fitur, serta visualisasi distribusi kelas target. Analisis ini

bertujuan untuk mengidentifikasi karakteristik debitur pada masing-masing kelas (*default* dan *non-default*), serta memahami distribusi atribut yang digunakan dalam penelitian ini. Selain itu, EDA digunakan untuk mengidentifikasi potensi permasalahan pada data, seperti ketidakseimbangan kelas, *outlier*, dan hubungan antar variabel yang dapat mempengaruhi kinerja model prediksi. Hasil analisis ini menjadi dasar dalam menentukan tahapan *preprocessing*, seleksi fitur, dan pemilihan metode yang digunakan untuk membangun model prediksi risiko kredit yang lebih akurat.

2.3. Preprocessing

Pada penelitian ini, *preprocessing* data dilakukan dengan pengecekan *missing value* (data kosong), karena dataset yang digunakan dalam penelitian ini bersih jadi tidak perlu mengecek duplikat data. Setelah itu dilakukan *label encoding* pada atribut kategorikal, dan normalisasi data menggunakan *StandarScaler* (z-score).

2.4. Seleksi Fitur

Pada penelitian ini, seleksi fitur dilakukan menggunakan pendekatan *feature importance* gabungan dari XGBoost, LightGBM, dan CatBoost. Tujuan dari proses ini adalah untuk menemukan dan mempertahankan fitur yang berkontribusi besar terhadap proses pembentukan model, sementara fitur yang tidak relevan atau tidak berkontribusi besar terhadap variabel target dihilangkan karna berpotensi menambah *noise* sehingga dapat mempengaruhi kinerja prediksi.

2.5. Penyeimbangan Data

Pada penelitian ini, penanganan ketidakseimbangan kelas dilakukan menggunakan metode *SMOTE-Tomek*, yang menggabungkan metode *oversampling* dan *undersampling*. Sedangkan Tomek Links menghapus pasangan data dari berbagai kelas yang saling berdekatan, SMOTE menambah data sintesis ke kelas minoritas, seperti data *noise* atau data yang ambigu, sehingga batas antar kelas menjadi lebih jelas [14]. *SMOTE* digunakan dalam penelitian ini dengan alasan dataset pada tahap *Exploratory Data Analysis* (EDA) menunjukkan adanya class imbalance yaitu distribusi kelas antara data debitur gagal bayar jauh lebih sedikit dibandingkan debitur tidak gagal bayar. Teknik *SMOTE* nantinya akan bekerja dengan menambahkan sampel sintesis pada kelas minoritas hingga jumlahnya seimbang dengan kelas mayoritas. Penambahan data ini bertujuan mengatasi ketidakseimbangan kelas sehingga model dapat mempelajari pola kedua kelas secara lebih baik dan tidak bias terhadap kelas mayoritas.

2.6. Model Prediksi

Pada penelitian ini, model prediksi merupakan tahapan utama dalam penelitian yang bertujuan untuk membangun model klasifikasi risiko kredit berdasarkan data historis debitur. Algoritma yang digunakan pada penelitian ini adalah *Gradient Boosting* yaitu XGBoost, LightGBM, dan CatBoost yang berfokus membangun model secara bertahap melalui penambahan pohon keputusan pada setiap iterasi untuk memperbaiki kesalahan prediksi yang dihasilkan oleh model sebelumnya [9]. Ketiga algoritma tersebut dipilih karena memiliki kemampuan yang baik dalam menangani data tabular, meningkatkan akurasi prediksi melalui teknik *boosting*, serta telah banyak digunakan pada kasus klasifikasi, termasuk prediksi risiko kredit [8]. Selanjutnya, kinerja masing-masing algoritma akan dibandingkan untuk menentukan model yang paling efektif dalam memprediksi risiko kredit. XGBoost merupakan algoritma berbasis *gradient boosting* yang membangun model secara bertahap untuk mengurangi kesalahan prediksi sebelumnya, dan mampu menangani data besar serta kemampuannya menghasilkan akurasi tinggi dengan *overfitting* yang rendah. Algoritma LightGBM berfokus pada efisiensi komputasi. LightGBM menggunakan pendekatan *leaf-wise tree growth* yang memungkinkan proses pelatihan lebih cepat dan akurat, terutama pada dataset berukuran besar. Algoritma CatBoost menggunakan teknik *ordered boosting* dan *Symmetric Tree Growth*, sehingga mampu menghasilkan model yang lebih stabil dan akurat tanpa memerlukan proses *encoding* yang kompleks.

2.7. Hyperparameter Tuning

Hyperparameter adalah parameter yang ditetapkan sebelum proses pelatihan model dan berfungsi untuk mengatur struktur model. Namun, untuk beberapa kasus model *machine learning hyperparameter* digunakan untuk meningkatkan sebuah akurasi apabila hasil dari sebuah perhitungan mendapat akurasi rendah yang disebut *hyperparameter tuning*.

2.8. Evaluasi Model

Tahap evaluasi dilakukan untuk menilai performa model dengan memanfaatkan *confusion matrix* sebagai dasar perhitungan seperti yang disajikan pada Tabel 1. Berdasarkan matriks tersebut, dihitung evaluasi utama, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*.

Tabel 1. *Confusion Matrix*

Aktual	Prediksi	
	Positif	Negativ
Positive	TP	FN
Negative	FP	TN

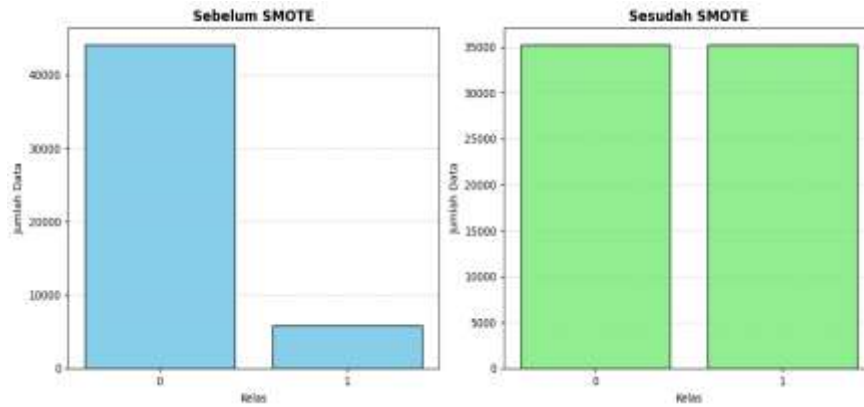
Accuracy mengukur tingkat ketepatan keseluruhan hasil prediksi model terhadap data aktual sebagaimana dirumuskan pada persamaan. *Recall* merupakan metrik evaluasi yang digunakan untuk mengukur kemampuan model dalam menemukan atau mendeteksi seluruh sampel yang benar-benar termasuk dalam suatu kelas tertentu dan dapat dihitung dengan persamaan. *Precision* digunakan untuk mengukur tingkat ketepatan prediksi positif model yang di hitung dengan persamaan, sedangkan *f1-score* digunakan untuk mengukur keseimbangan antara *precision* dan *recall* yang dapat dihitung dengan persamaan. Selain itu keempat metrik tersebut, digunakan pula *ROC-AUC* untuk mengukur kemampuan model dalam membedakan kelas positif dan kelas negatif secara menyeluruh.

3. Hasil dan Diskusi

3.1. Hasil

Dataset yang digunakan dalam penelitian ini berasal dari *Loan Default Prediction Challenge* yang diperoleh melalui OpenDataBay dan terdiri dari 50.000 data nasabah dengan 18 atribut. Pemeriksaan awal terhadap dataset menunjukkan bahwa tidak terdapat missing value maupun data duplikat sehingga data dapat langsung digunakan pada tahap *preprocessing*. Atribut kategorikal kemudian dikonversi ke dalam bentuk numerik menggunakan *Label Encoding*, sedangkan atribut numerik dinormalisasi menggunakan *StandarScaler*. Selanjutnya dilakukan seleksi fitur menggunakan *Feature Importance* gabungan dari algoritma XGBoost, LigtGBM, dan CatBoost sehingga diperoleh atribut-atribut yang memiliki kontribusi paling signifikan terhadap prediksi risiko kredit.

Hasil analisis distribusi kelas sebelum dan sesudah penerapan *SMOTE-Tomek*, menunjukkan bahwa dataset memiliki ketidakseimbangan data, dataset didominasi oleh kelas 0 (*Non-default*) dengan jumlah sekitar 44.200 data, sedangkan kelas 1 (*Default*) hanya sekitar 5.800 data. Kondisi tersebut menunjukkan adanya *class imbalance*, yang berpotensi menyebabkan model lebih cenderung mempelajari pola pada kelas mayoritas sehingga kemampuan dalam mendeteksi debitur yang berpotensi gagal bayar menjadi kurang optimal. Setelah penerapan *SMOTE-Tomek*, jumlah data pada kedua kelas menjadi seimbang, yaitu sekitar 35.200 data untuk masing-masing kelas. *SMOTE* menghasilkan data sintesis pada kelas minoritas (*Default*), sedangkan *Tomek Links* menghapus data yang berada di batas anatar kelas sehingga distribusi data menjadi lebih representatif. Penyeimbangan ini diharapkan dapat meningkatkan kemampuan model dalam mempelajari karakteristik kedua kelas secara lebih proporsional dan menghasilkan performa klasifikasi yang lebih baik. Lebih jelasnya dapat dilihat pada Gambar 3.



Gambar 3. Hasil sebelum dan sesudah SMOTE

Model prediksi dibangun menggunakan algoritma *Gradient Boosting* yakni, XGBoost, LightGBM, dan CatBoost dengan parameter terbaik yang diperoleh melalui *hyperparameter tuning*. Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, *f1-score*, dan *ROC-AUC*. Hasil evaluasi ketiga model algoritma disajikan pada Tabel 2.

Tabel 2. Perbandingan Hasil Evaluasi

Algoritma	Accuracy	Precision	Recall	F1-Score	ROC-AUC
XGBoost	88,82%	44,3%	15%	22%	0,726
LightGBM	88,34%	49,1%	14%	21%	0,732
CatBoost	87,16%	55,4%	20%	26%	0,737

Berdasarkan Tabel 2, XGBoost memperoleh nilai *accuracy* tertinggi sebesar 88,82%, sedangkan CatBoost menghasilkan nilai *precision*, *recall*, *f1-score*, dan *ROC-AUC* tertinggi dibandingkan algoritma lainnya. Hasil tersebut menunjukkan bahwa CatBoost memiliki kemampuan yang lebih baik dalam mengidentifikasi debitur yang berpotensi mengalami gagal bayar. Model CatBoost selanjutnya dipilih untuk diimplementasikan kedalam sistem pendukung keputusan berbasis Streamlit sebagai alat bantu analisis kelayakan kredit.

Berdasarkan hasil confusion matrix pada masing-masing algoritma, dapat dianalisis kemampuan model dalam mengklasifikasikan debitur gagal bayar maupun yang tidak mengalami gagal bayar melalui *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), *False Negative* (FN). Pada penelitian ini XGBoost mencatatkan *True Positive* tertinggi (101), artinya paling banyak berhasil mendeteksi debitur gagal bayar. Namun, XGBoost juga memiliki *False Positive* tertinggi (95), sehingga lebih banyak debitur aman yang keliru diprediksi berisiko. Sebaliknya, CatBoost menunjukkan keseimbangan terbaik dengan *False Positive* terendah (79) dan *True Negative* tertinggi (8.776). Hasil tersebut menunjukkan bahwa CatBoost lebih baik dalam mengurangi kesalahan klasifikasi terhadap debitur yang sebenarnya tidak mengalami gagal bayar. Meskipun nilai *True Positive* CatBoost lebih rendah dibandingkan XGBoost, keseimbangan antara *precision*, *ROC-AUC*, dan jumlah *False Negative* yang rendah menjadikan CatBoost sebagai model yang lebih stabil dan andal untuk prediksi risiko kredit. Perbandingan *confusion matrix* disajikan pada Tabel 3. berikut.

Tabel 3. Perbandingan Hasil *Confusion Matrix*

Algoritma	True Negative	False Positive	False Negative	True Positive
XGBoost	8.750	83	1.066	101
LightGBM	8.764	69	1.072	95
CatBoost	8.776	57	1.088	79

Hasil penelitian ini menunjukkan efektivitas penggunaan algoritma XGBoost, LightGBM, dan CatBoost dalam memprediksi risiko kredit menggunakan data yang telah melalui tahap *preprocessing*, seleksi fitur berbasis feature importance, penyeimbangan kelas menggunakan *SMOTE-Tomek*, serta *hyperparameter tuning*. Berdasarkan hasil pengujian terhadap 50.000 data debitur, algoritma CatBoost menunjukkan kinerja terbaik dengan nilai *ROC-AUC* sebesar 0,737, *precision* sebesar 55,4%, dan jumlah *False Positive* paling rendah, sehingga memberikan performa yang lebih stabil dibandingkan XGBoost dan LightGBM. Meskipun XGBoost memperoleh akurasi tertinggi

sebesar 88,82% dan mampu mendeteksi debitur gagal bayar lebih banyak, CatBoost memiliki kemampuan yang lebih baik dalam mengurangi kesalahan klasifikasi pada debitur yang tidak gagal bayar.

3.2. Diskusi

Hasil penelitian menunjukkan bahwa setiap algoritma *Gradient Boosting* memiliki karakteristik yang berbeda dalam memprediksi risiko kredit. Meskipun XGBoost memperoleh nilai *accuracy* tertinggi sebesar 88,82%, metrik tersebut belum dapat dijadikan indikator utama karena dataset memiliki distribusi kelas yang tidak seimbang. Pada kondisi seperti ini, *ROC-AUC*, *precision*, *recall*, dan *f1-score* lebih representatif dalam mengevaluasi kemampuan model mendeteksi debitur yang berpotensi mengalami gagal bayar. CatBoost memperoleh nilai *ROC-AUC* tertinggi sebesar 0,737 serta menghasilkan nilai *precision* 55,4%, *recall* sebesar 20%, dan *f1-score* sebesar 26%. Hasil ini menunjukkan bahwa CatBoost memiliki kemampuan diskriminasi yang lebih baik dalam membedakan debitur yang berisiko gagal bayar dengan debitur yang tidak berisiko. Kinerja tersebut dipengaruhi oleh mekanisme *ordered boosting* yang mampu mengurangi *prediction shift* selama proses pelatihan serta kemampuannya dalam mengolah fitur kategorikal secara lebih efektif dibandingkan XGBoost dan LightGBM.

Analisis *confusion matrix* juga memperkuat hasil evaluasi tersebut. XGBoost menghasilkan nilai *True Positive* tertinggi, yaitu 101, sehingga lebih banyak mendeteksi debitur yang mengalami gagal bayar. Namun, algoritma ini juga menghasilkan jumlah *False Positive* yang lebih tinggi dibandingkan CatBoost, sehingga lebih banyak debitur yang sebenarnya tidak berisiko diklasifikasikan sebagai berisiko. Sebaliknya, CatBoost menghasilkan nilai *False Positive* terendah dan *True Negative* tertinggi, yang menunjukkan kemampuan lebih baik dalam mengurangi kesalahan klasifikasi terhadap debitur yang sebenarnya tidak mengalami gagal bayar. Dengan demikian, CatBoost memberikan keseimbangan yang lebih baik antara kemampuan mendeteksi risiko kredit dan meminimalkan kesalahan prediksi.

Penerapan *SMOTE-Tomek* juga berkontribusi terhadap peningkatan performa model. Penyeimbangan distribusi kelas melalui pembentukan data sintesis pada kelas minoritas dan penghapusan data pada batas antar kelas membantu model mempelajari karakteristik kedua kelas secara lebih proporsional. Proses ini mengurangi bias terhadap kelas mayoritas sekaligus memperjelas batas pemisah antar kelas, sehingga kesalahan klasifikasi dapat diminimalkan. Dengan demikian, model memiliki kemampuan yang baik dalam mengenali pola debitur yang berpotensi mengalami gagal bayar.

Temuan penelitian ini sejalan dengan penelitian yang dilakukan oleh refrensi [19] yang menyatakan bahwa CatBoost mampu memberikan performa klasifikasi yang lebih stabil pada data kredit. Namun hasil penelitian ini berbeda dengan penelitian [15] yang menyatakan bahwa XGBoost merupakan algoritma dengan performa terbaik. Perbedaan tersebut diduga dipengaruhi oleh karakteristik dataset, teknik *preprocessing*, metode penyeimbangan data menggunakan *SMOTE-Tomek*, seleksi fitur, serta proses *hyperparameter tuning* yang diterapkan pada penelitian ini.

4. Kesimpulan

Penelitian ini menunjukkan, bahwa algoritma CatBoost terbukti menjadi model terbaik dalam prediksi risiko kredit dengan nilai *ROC-AUC* tertinggi sebesar 0,737, diikuti LightGBM (0,732) dan XGBoost (0,726). CatBoost juga menunjukkan keseimbangan prediksi yang lebih stabil pada *confusion matrix* dengan *False Positive* terendah. Penerapan metode *SMOTE-Tomek* efektif dalam menangani ketidakseimbangan kelas pada dataset, sehingga ketiga model mampu mengenali debitur berisiko gagal bayar dengan lebih baik dan menghasilkan prediksi yang lebih optimal. Model CatBoost berhasil diimplementasikan ke dalam Sistem Pendukung Keputusan berbasis web menggunakan *Framework Streamlit*, yang mampu memprediksi risiko kredit secara interaktif dan mendukung pengambilan keputusan kredit secara lebih objektif serta berbasis data.

Referensi

- [1] R. Yudaruddin, B. A. Nugroho, Z. Fitriani, and P. Hapsari, "Liquidity and Credit Risk in Indonesia: The Role of FinTech Development," no. 1, pp. 1–20, 2024, <https://doi.org/10.1177/21582440241245248>
- [2] N. Suhadolnik, J. Ueyama, and S. Da Silva, "Machine Learning for Enhanced Credit Risk Assessment: An Empirical Approach," *J. Risk Financ. Manag.*, 2023, <https://doi.org/10.3390/jrfm16120496>
- [3] T. Novianti, S. A. Mandati, and E. K. Andana, "Peningkatan Evaluasi Risiko Kredit Menggunakan Decision Tree C 4 . 5," vol. 2, no. 2, pp. 1–9, 2023, <https://doi.org/10.30651/mine-tech.v2i2.21749>
- [4] Otoritas Jasa Keuangan, "Konferensi Pers Hasil Rapat Dewan Komisiner Bulanan Agustus 2024," Jakarta, 2024. [Online]. Available: [https://ojk.go.id/id/berita-dan-kegiatan/siaran-pers/Documents/Pages/Sektor-Jasa-Kuangan-Terjaga-Stabil-Dengan-Kinerja-Yang-Baik-di-Tengah-Ketidapastian-yang-Masih-Tinggi/SP RDKB Agustus 2024.pdf](https://ojk.go.id/id/berita-dan-kegiatan/siaran-pers/Documents/Pages/Sektor-Jasa-Kuangan-Terjaga-Stabil-Dengan-Kinerja-Yang-Baik-di-Tengah-Ketidapastian-yang-Masih-Tinggi/SP%20RDKB%20Agustus%2024.pdf)

-
- [5] F. A. Undie, S. A. Hameed, J. A. Ikwen, M. O. Okache, and A. Adebayo, "Artificial Intelligence and Machine Learning : A Review of State-of- the-Art Trends , Global Developments , and Practical Implications," *J. Curr. Trends Comp. Sci. Res.*, vol. 3, no. 1, pp. 1–6, 2024. <https://doi.org/10.33140/JCTCSR.03.01.06>
 - [6] M. E. Ghoujdami, R. Chaabita, and S. & Idamia, "Machine Learning dans l ' évaluation du risque crédit : revue systématique Machine learning in credit risk assessment : A systematic review .," *Int. J. Appl. Manag. Econ.*, vol. 02, pp. 255–274, 2024. <https://doi.org/10.5281/zenodo.12528088>
 - [7] S. S. Islam, S. Haque, M. S. U. Miah, and T. Bin, "Application of machine learning algorithms to predict the thyroid disease risk : an experimental comparative study," pp. 1–35, 2022, <https://doi.org/10.7717/peerj-cs.898>
 - [8] D. Shukla, "A Survey of Machine Learning Algorithms in Credit Risk Assessment," vol. 3, pp. 6290–6297, 2024. <https://doi.org/10.52783/jes.6788>
 - [9] N. Gunasekara, B. Pfahringer, H. Gomes, and A. Bifet, "Gradient boosted trees for evolving data streams," *Mach. Learn.*, vol. 113, no. 5, pp. 3325–3352, 2024, <https://doi.org/10.1007/s10994-024-06517-y>
 - [10] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery (ACM), 2016, pp. 785–794. <https://doi.org/10.1145/2939672.2939785>
 - [11] G. Ke *et al.*, "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2017, pp. 3146–3154. [Online]. Available: https://papers.nips.cc/paper_files/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html
 - [12] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorigush, and A. Gulin, "CatBoost: Unbiased Boosting with Categorical Features," in *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*, Curran Associates, Inc., 2018, pp. 6638–6648. [Online]. Available: <https://papers.neurips.cc/paper/2018/hash/14491b756b3a51daac41c24863285549-Abstract.html>
 - [13] P. S. Rizky *et al.*, "Perbandingan Metode LightGBM dan XGBoost dalam Menangani Data dengan Kelas Tidak Seimbang Universitas Hamzanwadi Selong," vol. 15, no. 2, pp. 228–236, 2022. <https://doi.org/10.36456/jstat.vol15.no2.a5548>
 - [14] N. V. Chawla, K. W. Bowyer, Hall, & K. L. O., and W. P., "SMOTE: Synthetic minority over-sampling technique.," *J. Artif. Intell. Res.*, 2002. <https://doi.org/10.1613/jair.953>
 - [15] Dong, Xia, and W. Xue, "Analisis dan Perbandingan Model Prediksi Gagal Bayar Pinjaman Berdasarkan Algoritma XGBoost dan LightGBM," vol. 6, pp. 32–37, 2023, <https://dx.doi.org/10.25236/AJCIS.2023.060905>
 - [16] A. H. Ginting, R. W. Sembiring, and E. M. Zamzami, *Comparison Analysis : Logistic Regression , Random Forest , XGBoost , and CatBoost in Credit Scoring*, no. Bic 2024. Atlantis Press International BV, 2025. https://doi.org/10.2991/978-94-6463-854-7_16
 - [17] R. Yunanto and U. Budiyo, "Implementasi XGBoost dan SMOTE untuk Meningkatkan Deteksi Transaksi Fraud di Industri Jasa Keuangan Implementing Xgboost Models For Enhanced Detection Of Fraud Transaction In Financial Services Industries," vol. 4, no. 11, pp. 525–535, 2024. <https://doi.org/10.52436/1.jpti.518>
 - [18] A. N. Royana, Y. V. Via, C. A. Putra, J. Raya, R. Madya, and A. Music, "EVALUASI KINERJA LIGHTGBM DAN CATBOOST UNTUK PREDIKSI CHURN BERDASARKAN DATASET PELANGGAN LAYANAN STREAMING MUSIK," vol. 9, no. 4, pp. 6977–6982, 2025. <https://doi.org/10.36040/jati.v9i4.14358>
 - [19] F. A. Nasution, S. Saadah, P. E. Yunanto, D. Informatika, and U. Telkom, "Credit Risk Detection in Peer-to-Peer Lending Using CatBoost," vol. 5, no. 158, 2023. <https://doi.org/10.29207/resti.v7i5.5139>