Progresif: Jurnal Ilmiah Komputer
https://ojs.stmik-banjarbaru.ac.id/index.php/progresif/index
Jl. Ahmad Yani, K.M. 33,5 - Kampus STMIK Banjarbaru
Loktabat – Banjarbaru (Tlp. 0511 4782881), e-mail: puslit.stmikbjb@gmail.com
e-ISSN: 2685-0877

Prediksi Kelayakan Kredit Nasabah Dengan Penerapan *Cost-Sensitive Random Forest*

DOI: http://dx.doi.org/10.35889/progresif.v22i1.3401

Jolyn Lucretia[1*], Dedy Hermanto[2]
Informatika, Universitas Multi Data Palembang, Palembang, Indonesia
**e-mail Corresponding Author:* jolynlucretia_2226250055@mhs.mdp.ac.id

Abstract

The high risk of credit default and class imbalance in customer data pose major challenges in developing accurate credit scoring systems. This condition causes predictive models to be biased toward the majority class, thereby reducing the ability to detect high-risk borrowers. This study develops a credit scoring model for imbalanced data using the Synthetic Minority Oversampling Technique (SMOTE) and cost-sensitive Random Forest with hyperparameter optimization via GridSearchCV. The dataset consists of 32,581 customer records. Experimental results show that the best configuration with n_estimators = 200 achieves a cross-validation F1-score of 0.813750. On the test data, the model attains an accuracy of 0.927267, precision of 0.911458, recall of 0.738397, and an F1-score of 0.815851, indicating improved and more balanced detection of high-risk borrowers.

Keywords: *Random Forest; SMOTE; Cost-sensitive learning; GridSearchCV*

Abstrak

Tingginya risiko gagal bayar kredit dan ketidakseimbangan kelas pada data nasabah menjadi tantangan utama dalam pengembangan sistem credit scoring yang akurat. Kondisi ini menyebabkan model prediksi cenderung bias terhadap kelas mayoritas sehingga kemampuan deteksi debitur berisiko menjadi kurang optimal. Penelitian ini mengembangkan model *credit scoring* pada data tidak seimbang menggunakan *Synthetic Minority Oversampling Technique* (*SMOTE*) dan *Cost-Sensitive Random Forest* dengan optimasi *hyperparameter GridSearchCV*. Dataset yang digunakan berjumlah 32.581 data nasabah. Hasil pengujian menunjukkan konfigurasi terbaik dengan *n_estimators* = 200 menghasilkan *F1-score* validasi silang sebesar 0,813750. Pada data uji, model mencapai akurasi 0,927267, *precision* 0,911458, *recall* 0,738397, dan *F1-score* 0,815851, yang menunjukkan peningkatan kemampuan deteksi debitur berisiko secara lebih seimbang.

Kata kunci: *Random Forest; SMOTE; Cost-sensitive learning; GridSearchCV.*

1. Pendahuluan

Kredit merupakan salah satu instrumen penting dalam sistem keuangan modern yang berfungsi untuk menyediakan pembiayaan bagi individu maupun badan usaha [1]. Melalui pemberian kredit, lembaga keuangan dapat membantu masyarakat dalam memenuhi kebutuhan konsumtif maupun produktif, sekaligus mendorong pertumbuhan ekonomi [2]. Menurut Andrianto [3], kredit modal kerja merupakan jenis pembiayaan yang disediakan oleh perbankan kepada nasabah untuk memenuhi kebutuhan operasional usaha dalam jangka pendek, sehingga memiliki peran penting dalam menjaga kelangsungan kegiatan bisnis.

Berdasarkan Laporan Surveillance Perbankan Indonesia Triwulan III 2024 yang diterbitkan oleh OJK, penyaluran kredit bank umum tumbuh sebesar 10,85% (yoy), meningkat dari tahun sebelumnya [4]. Namun, meskipun rasio kredit bermasalah (NPL gross) menurun menjadi 2,21%, risiko gagal bayar tetap menjadi perhatian dalam menjaga stabilitas sistem keuangan. Statistik perbankan juga mencatat bahwa pada Februari 2023, rasio kredit bermasalah untuk kredit konsumsi mencapai 5,45%, mengalami kenaikan sebesar 1,54% dibandingkan Desember 2022 dan telah melampaui ambang batas yang ditetapkan oleh Peraturan Bank

Indonesia [5]. Selain itu, sistem penilaian kredit masih banyak dilakukan secara konvensional melalui analisis manual yang memakan waktu, bersifat subjektif, dan kurang adaptif terhadap kompleksitas data nasabah [6]. Tantangan lain yang muncul adalah ketidakseimbangan distribusi kelas, di mana jumlah nasabah lancar jauh lebih besar dibandingkan dengan nasabah gagal bayar, sehingga model cenderung bias terhadap kelas mayoritas [7].

Seiring berkembangnya teknologi informasi, pendekatan berbasis *machine learning* mulai digunakan untuk membangun sistem *credit scoring*, yaitu metode penilaian kelayakan kredit dengan memanfaatkan data historis nasabah dan model prediktif [8]. Dalam konteks data tidak seimbang, metode seperti *SMOTE* dapat membantu menambah representasi kelas minoritas, sementara pendekatan *Cost-Sensitive Learning* memberi bobot kesalahan yang lebih besar pada prediksi gagal bayar untuk mengurangi bias terhadap kelas mayoritas. Algoritma *Random Forest* sendiri dikenal mampu menangani kompleksitas data kredit dan memberikan performa yang stabil pada berbagai kondisi. Untuk meningkatkan performanya, diperlukan *hyperparameter tuning* menggunakan metode seperti *GridSearchCV* agar model bekerja pada konfigurasi yang paling optimal. Kolaborasi metode *machine learning*, *Cost-Sensitive Learning*, *SMOTE*, dan optimasi *hyperparameter* dianggap relevan serta logis untuk menghasilkan sistem *credit scoring* yang lebih objektif dan akurat [9] [10].

Penelitian ini bertujuan mengimplementasikan model *credit scoring* berbasis kombinasi *Cost-Sensitive Learning*, algoritma *Random Forest*, metode *SMOTE*, dan optimasi *hyperparameter* menggunakan *GridSearchCV* untuk meningkatkan kemampuan model dalam mendeteksi nasabah berisiko gagal bayar pada data tidak seimbang. Secara teoritis, penelitian ini diharapkan dapat memperkaya kajian penerapan algoritma klasifikasi dalam bidang keuangan serta pengembangan model *machine learning* yang adaptif terhadap ketidakseimbangan data. Secara praktis, penelitian ini diharapkan dapat membantu lembaga keuangan menilai kelayakan kredit secara lebih objektif, efisien, dan akurat, sehingga mendukung proses pengambilan keputusan yang lebih cepat dan berbasis data dalam pengelolaan risiko keuangan yang berkelanjutan.

2. Tinjauan Pustaka

Putra dan Rumini [7] meneliti masalah ketidakseimbangan kelas pada data kredit, di mana jumlah nasabah yang membayar tepat waktu lebih besar dibandingkan dengan nasabah yang gagal bayar. Penelitian tersebut menjelaskan bahwa kondisi ini dapat menyebabkan model cenderung bias terhadap kelas mayoritas sehingga menurunkan akurasi dalam mendeteksi debitur berisiko gagal bayar.

Selanjutnya, Syah et al. [9] menggunakan algoritma *Random Forest* pada data kredit dengan distribusi kelas yang tidak seimbang. Prosedur penelitian mencakup pembangunan model *Random Forest* serta evaluasi menggunakan akurasi, precision, dan recall. Penelitian ini melaporkan bahwa *Random Forest* mampu mencapai akurasi 96,88% dengan precision dan recall yang seimbang meskipun dataset tidak seimbang.

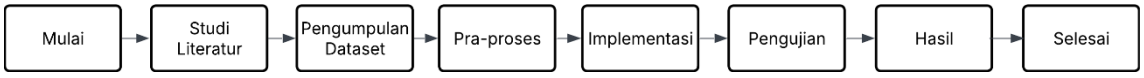
Kemudian, Nanda et al. [10] mengintegrasikan metode *Synthetic Minority Oversampling Technique (SMOTE)* dengan algoritma *Random Forest* untuk menangani ketidakseimbangan data kredit. Tahapan penelitian meliputi proses oversampling menggunakan *SMOTE*, pelatihan *Random Forest*, dan evaluasi melalui akurasi, precision, dan recall. Penelitian tersebut menunjukkan akurasi 94,41%, precision 97,03%, dan recall 91,55%, serta peningkatan akurasi sebesar 2,87%, precision 6,2%, dan recall 31,29% dibandingkan model tanpa *SMOTE*.

Pada penelitian lain, Mienye dan Sun [11] menerapkan *Cost-Sensitive Random Forest (CS-RF)* pada data kredit yang tidak seimbang. Penelitian ini menggunakan pendekatan pembobotan kesalahan pada kelas minoritas dan mengevaluasi performa menggunakan akurasi, recall, F-measure, dan AUC. Hasil penelitian menunjukkan akurasi 98,6% dan AUC 1,0, serta peningkatan recall dan F-measure dibandingkan *Random Forest* standar maupun algoritma *Logistic Regression*, *Decision Tree*, dan *XGBoost*.

Selain itu, Prieto-Egido et al. [12] mengombinasikan *Random Forest*, *SMOTE*, dan *Cost-Sensitive Learning* dalam pemodelan risiko kredit. Prosedur penelitian meliputi oversampling menggunakan *SMOTE*, pelatihan *Random Forest*, dan penerapan pembobotan kesalahan. Evaluasi dilakukan dengan recall, F1-score, dan g-means, dan penelitian tersebut melaporkan peningkatan recall dari 0,580 menjadi 0,615, F1-score dari 0,54 menjadi 0,55, serta g-means dari 0,74 menjadi 0,77 setelah *Cost-Sensitive Learning* diterapkan.

Berdasarkan tinjauan terhadap penelitian-penelitian terdahulu tersebut, dapat disimpulkan bahwa penanganan ketidakseimbangan kelas dalam pemodelan risiko kredit umumnya masih dilakukan secara parsial, baik melalui penerapan *Random Forest*, teknik *oversampling SMOTE*, pendekatan *Cost-Sensitive Learning*, maupun optimasi *hyperparameter* secara terpisah. Belum banyak penelitian yang mengintegrasikan penyeimbangan data, pembobotan kesalahan pada kelas minoritas, serta optimasi *hyperparameter* dalam satu alur pemodelan *credit scoring*. Oleh karena itu, penelitian ini menghadirkan kebaruan dengan mengombinasikan *Cost-Sensitive Random Forest*, *SMOTE*, dan *GridSearchCV* dalam satu rangkaian prosedur pemodelan, sehingga diharapkan mampu meningkatkan sensitivitas model dalam mendeteksi debitur berisiko gagal bayar secara lebih optimal.

3. Metodologi



Gambar 1. Tahapan Metodologi Penelitian

3.1. Studi Literatur

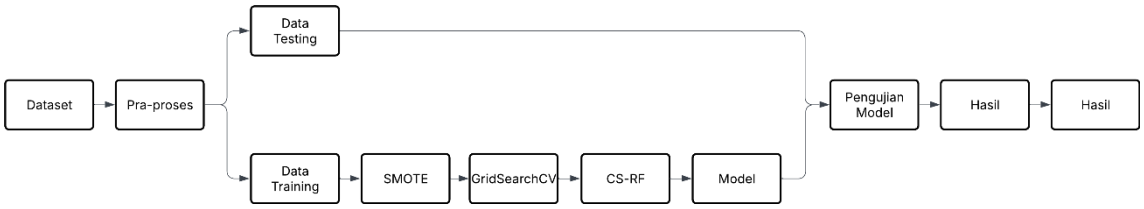
Tahapan ini dimulai dengan melakukan pembelajaran literatur dengan membaca jurnal-jurnal terdahulu yang berfokus pada prediksi kelayakan kredit nasabah yang menggunakan pendekatan *Cost-sensitive Learning Random Forest* dengan Teknik *oversampling Smote* dan *Hyperparameter tuning GridSearchCV*.

3.2. Pengumpulan Dataset

Pada tahap ini, dilakukan proses pengumpulan dataset berupa dataset resiko kredit nasabah yang sifatnya publik dan diperoleh dari *Kaggle*. Dataset ini terdapat 32.581 data yang merepresentasikan profil calon peminjam.

3.3. Pra-proses

Pada tahap ini, sistem *credit scoring* dirancang mengikuti alur mulai dari preprocessing data, pembagian data, penyeimbangan kelas menggunakan *SMOTE*, transformasi fitur, serta penerapan *Cost-Sensitive Random Forest* yang dioptimasi melalui *GridSearchCV* hingga menghasilkan model prediksi yang dievaluasi menggunakan metrik performa.



Gambar 2. Skema Perancangan Model

3.3.1. Pembersihan Data

Berdasarkan hasil pemeriksaan awal terhadap kualitas dataset, diketahui bahwa terdapat 4.011 nilai kosong (*missing values*) dan 165 baris duplikat, sebagaimana ditunjukkan pada Gambar 3. Kondisi ini menunjukkan bahwa dataset memerlukan proses pembersihan sebelum digunakan pada tahap pemodelan, karena keberadaan nilai kosong dan data ganda berpotensi menurunkan akurasi model serta menyebabkan bias dalam proses pelatihan. Oleh karena itu, pada tahap pembersihan data dilakukan penghapusan duplikasi dan penanganan nilai kosong untuk memastikan bahwa dataset berada dalam kondisi yang layak digunakan untuk analisis lebih lanjut.

```
Total missing: 4011
Total duplikat: 165
```

Gambar 3. Pemeriksaan *Missing Value* dan Nilai Duplikat

3.3.2. Transformasi Data

Setelah proses pembersihan selesai, tahapan berikutnya adalah transformasi data, yang meliputi normalisasi fitur numerik dan pengkodean fitur kategorikal menggunakan *One-Hot Encoding*. Hasil transformasi ditunjukkan pada Gambar 4, di mana sebanyak tujuh fitur numerik berhasil dinormalisasi dan 19 fitur kategorikal dikonversi menjadi representasi biner melalui OHE, sehingga menghasilkan total 26 fitur siap pakai untuk proses pelatihan model.

```
=== INFO FITUR HASIL TRANSFORMASI ===
Jumlah fitur numerik (scaled): 7
Jumlah fitur OHE (categorical): 19
Total fitur setelah transformasi: 26
```

Gambar 4. Hasil Penerapan Transformasi *One-Hot Encoding*

3.3.3. Normalisasi Data

Setelah proses transformasi dilakukan, tahap selanjutnya adalah normalisasi data numerik menggunakan metode *Min-Max Scaling*. Contoh hasil normalisasi ditunjukkan pada Tabel 1, yang menampilkan sampel fitur pada baris pertama dari data setelah skala setiap atribut numerik diubah ke rentang 0–1. Tampak bahwa seluruh nilai pada tabel tersebut sudah berada dalam skala yang seragam, menunjukkan bahwa proses normalisasi telah berjalan dengan baik. Normalisasi ini penting untuk memastikan bahwa perbedaan rentang nilai antar atribut tidak menimbulkan bias selama proses pelatihan model. Setelah seluruh proses praproses data selesai, dataset kemudian dibagi menjadi dua bagian, yaitu 80% sebagai data latih dan 20% sebagai data uji, guna memastikan bahwa evaluasi model dilakukan secara objektif dan terpisah dari proses pelatihan.

Tabel 1. Proses *Scaling* Fitur (Sampel)

Sampel Fitur	Sebelum *Scaling*	Sesudah *Scaling*
person_age	37	0.137097
person_income	84000	0.013329
person_emp_length	11	0.089431
loan_amnt	8225	0.223913
loan_int_rate	12.99	0.425281
loan_percent_income	0.10	0.120482
cb_person_cred_hist_length	11.0	0.321429

3.3.4. Pembagian Dataset

Dataset dibagi menjadi dua bagian, yaitu 80% untuk data latih (*training set*) dan 20% untuk data uji (*testing set*) guna memisahkan proses pelatihan dan pengujian model.

3.4. Implementasi

Tahap implementasi dilakukan untuk menerapkan seluruh tahapan penelitian mulai dari pengolahan data, penyeimbangan kelas, pelatihan model, hingga evaluasi hasil. Proses ini mencakup penerapan teknik *SMOTE* untuk menyeimbangkan distribusi data, penggunaan *Cost-Sensitive Random Forest* untuk klasifikasi kelayakan kredit dengan penyesuaian bobot kelas, serta *GridSearchCV* untuk menentukan nilai optimal pada parameter *n_estimators*. Hasil implementasi kemudian dievaluasi menggunakan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score* guna menilai performa model dalam memprediksi risiko gagal bayar pada nasabah.

3.4.1. *Random Forest*

Algoritma Random Forest merupakan salah satu pendekatan *ensemble learning* yang efektif digunakan untuk menyelesaikan permasalahan klasifikasi maupun regresi. Prinsip

kerjanya adalah membangun banyak *Decision Tree* yang masing-masing dilatih menggunakan sampel data dan fitur yang dipilih secara acak (*bagging*). Setiap pohon menghasilkan prediksi tersendiri, kemudian seluruh prediksi tersebut digabungkan untuk memperoleh keluaran akhir yang lebih stabil dan akurat, sehingga mampu menekan risiko *overfitting* dan meningkatkan kemampuan generalisasi [13]. Dalam proses agregasinya, setiap pohon memberikan satu suara terhadap kelas tertentu, dan hasil akhir ditentukan melalui mekanisme *majority voting*. Secara matematis, proses ini dapat diformulasikan sebagai:

$$H(x) = \text{mode}\{h(x, \Theta_k)\}_{k=1}^{K} \quad (1)$$

di mana $h(x, \Theta_k)$ merupakan prediksi dari pohon ke-k yang dibentuk menggunakan subset data dan fitur yang dipilih secara acak. Pendekatan ini menjadikan *Random Forest* mampu menghasilkan prediksi yang lebih stabil dan tahan terhadap variasi data [14]..

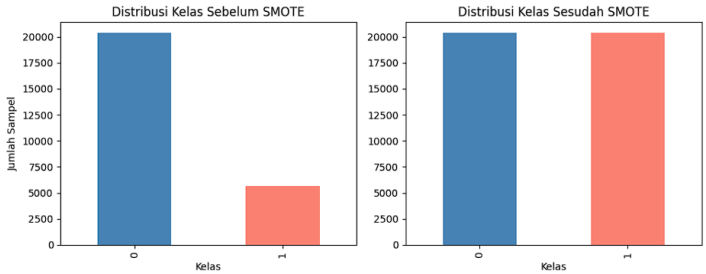
3.4.2. *Cost-Sensitive*

Cost-sensitive learning adalah pendekatan dalam *machine learning* yang mempertimbangkan perbedaan biaya dari setiap jenis kesalahan klasifikasi. Teknik ini sangat efektif diterapkan pada data yang tidak seimbang, karena memungkinkan model untuk memberi bobot lebih besar pada kelas minoritas yang berisiko tinggi jika terjadi kesalahan prediksi. Dengan demikian, *cost-sensitive learning* membantu meningkatkan sensitivitas model terhadap kelas penting tanpa mengorbankan akurasi keseluruhan [15]. metode ini semakin banyak digunakan pada penelitian modern, terutama pada kasus ketidakseimbangan kelas yang ekstrem, karena mampu memberikan performa yang lebih stabil dan relevan secara praktis dibandingkan teknik penyeimbangan data tradisional [16]. Untuk mengimplementasikan konsep tersebut secara matematis, cost-sensitive learning biasanya memanfaatkan *cost matrix* dan bobot kelas yang dihitung berdasarkan tingkat risiko setiap jenis kesalahan, sehingga model dapat meminimalkan *expected misclassification cost* melalui formulasi perhitungan sebagai berikut:

$$w(j) = \frac{C(j) \cdot N}{\sum_i C(i)\ N_i} \quad (2)$$

3.4.3. *Synthetic Minority Oversampling Technique (SMOTE)*

Pada penelitian ini, ketidakseimbangan distribusi kelas pada variabel *loan_status* ditangani menggunakan *Synthetic Minority Oversampling Technique* (*SMOTE*), yaitu metode *oversampling* yang bekerja dengan menghasilkan sampel sintetis baru pada kelas minoritas melalui interpolasi nilai fitur dari beberapa tetangga terdekat [17], [18]. Sebelum *SMOTE* diterapkan, data latih menunjukkan ketimpangan yang cukup besar, di mana kelas 0 memiliki 20.378 sampel dan kelas 1 hanya 5.686 sampel, sebagaimana ditunjukkan pada Gambar X. Ketimpangan ini berpotensi menyebabkan model bias terhadap kelas mayoritas dan mengabaikan pola penting pada kelas minoritas. Setelah *SMOTE* diterapkan pada data latih yang telah melalui proses transformasi numerik, jumlah sampel pada kedua kelas menjadi seimbang, masing-masing 20.378 sampel seperti terlihat pada Gambar 5. Keseimbangan ini memungkinkan model mempelajari karakteristik kelas minoritas dengan lebih baik, sehingga diharapkan meningkatkan performa metrik yang sensitif terhadap kesalahan pada kelas tersebut, terutama *recall* dan *F1-score*.



Gambar 5. Distribusi Kelas Sebelum dan Sesudah *SMOTE*

3.4.4. *GridSearchCV*

Pada tahap optimasi *hyperparameter*, penelitian ini menerapkan *GridSearchCV* untuk menentukan nilai terbaik pada parameter *n_estimators* dari algoritma *Cost-Sensitive Random Forest* yang telah dikombinasikan dengan *SMOTE*. Secara umum, *hyperparameter tuning* bertujuan mencari kombinasi parameter yang menghasilkan performa model paling optimal, dan *Grid Search* merupakan salah satu metode yang sering digunakan karena mampu mengevaluasi seluruh kemungkinan konfigurasi secara menyeluruh, meskipun metode ini memiliki kelemahan berupa waktu komputasi yang tinggi ketika jumlah parameter atau data sangat besar [19]. Dalam penelitian ini, *GridSearchCV* menggunakan lima skema *fold*. Hasil yang disajikan pada Tabel 2 menunjukkan bahwa *n_estimators* = 200 memberikan rata-rata *F1-score* tertinggi, yaitu 0,813750, diikuti konfigurasi 100 dan 50 pohon dengan nilai masing-masing 0,810397 dan 0,810288. Temuan ini mengindikasikan bahwa peningkatan jumlah pohon mampu memberikan stabilitas prediksi yang lebih baik selama proses validasi silang, sehingga *n_estimators* = 200 dipilih sebagai konfigurasi paling optimal untuk pelatihan model akhir.

Tabel 2. Hasil Optimasi Parameter

n_estimators	*F1-Score*
200	0.813750
100	0.810397
50	0.810288

3.5. Pengujian

Pada tahapan ini dilakukan pengujian terhadap data uji menggunakan model yang telah dibangun sebelumnya untuk menilai kemampuan sistem dalam melakukan klasifikasi. Proses pengujian bertujuan untuk mengukur sejauh mana model dapat mengenali data baru yang belum pernah dilatih. Setelah tahap pengujian selesai, hasil prediksi dibandingkan dengan nilai aktual untuk menghasilkan *Confusion Matrix*. Matriks ini digunakan sebagai dasar perhitungan beberapa metrik evaluasi utama, yaitu *precision*, *recall*, dan *accuracy*, dan *f1 score* yang masing-masing dihitung menggunakan Persamaan (3), (4), (5), dan (6). Nilai-nilai tersebut kemudian digunakan untuk menilai tingkat keberhasilan dan keandalan metode yang diterapkan dalam menentukan kelayakan kredit nasabah.

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (4)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (5)$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \times 100\% \quad (6)$$

4. Hasil dan Pembahasan

4.1. Pelatihan Data

Pada tahap pelatihan data, model *Cost-Sensitive Random Forest* yang telah dikombinasikan dengan *SMOTE* dan dioptimalkan menggunakan *GridSearchCV* dilatih menggunakan data hasil praproses yang telah melalui proses *scaling* dan *one-hot encoding*. Selama proses pelatihan, model mempelajari pola antara fitur-fitur kredit dengan status pembayaran untuk membedakan nasabah berisiko dan tidak berisiko secara lebih akurat. Tabel 3 menunjukkan bahwa pada data pelatihan, model mampu menghasilkan prediksi yang konsisten dengan label asli, di mana data dengan label gagal bayar (1) diprediksi secara benar dengan probabilitas yang tinggi, sedangkan data tidak gagal bayar (0) memiliki probabilitas gagal bayar yang rendah. Selanjutnya, pada Tabel 4 yang menyajikan sampel data *testing*, sebagian besar

prediksi model masih sesuai dengan label sebenarnya, meskipun terdapat beberapa kesalahan prediksi pada kelas gagal bayar. Hal ini menunjukkan bahwa model telah mempelajari pola risiko kredit dengan baik pada data training dan tetap mampu melakukan generalisasi secara wajar pada data pengujian.

Tabel 3. Sampel Data Hasil Model *Credit Scoring* (Data *Training*)

Label Asli	Prediksi Model	Probabilitas Gagal Bayar
0	0	0.015
1	1	0.755
0	0	0.095
1	1	0.940
0	0	0.000

Tabel 4. Sampel Data Hasil Model *Credit Scoring* (Data *Testing*)

Label Asli	Prediksi Model	Probabilitas Gagal Bayar
0	0	0.018
1	1	0.821
0	0	0.097
1	0	0.462
1	1	0.905

Berdasarkan hasil yang disajikan pada Tabel 5, dapat dilihat bahwa model menghasilkan performa sempurna pada data *training* dengan nilai *accuracy*, *precision*, *recall*, dan *F1-score* sebesar 1,000. Kondisi ini menunjukkan bahwa model mampu mempelajari pola data latih secara optimal. Sementara itu, pada data testing yang belum pernah dilatih sebelumnya, model tetap menunjukkan performa yang baik dengan nilai accuracy sebesar 0,9273, precision 0,9115, recall 0,7384, dan F1-score 0,8159.

Perbedaan performa antara data training dan data testing menunjukkan adanya trade-off yang wajar pada model dengan kompleksitas tinggi, namun tidak mengindikasikan overfitting yang signifikan. Nilai recall pada data testing yang relatif tinggi mengindikasikan bahwa model mampu mendeteksi debitur berisiko gagal bayar secara konsisten pada data baru, sehingga menunjukkan kemampuan generalisasi yang baik.

Tabel 5. Perbandingan Performa Model pada Data *Training* dan Data *Testing*

Data	Acc	Precision	Recall	F1- Score
Training	1.0000	1.0000	1.0000	1.0000
Testing	0.9273	0.9115	0.7384	0.8159

4.2. Evaluasi Model

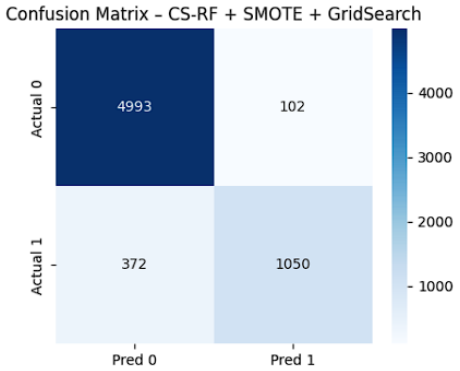
Pada tahap evaluasi model, seluruh skenario dibandingkan menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*. Berdasarkan hasil pada Tabel 5, kombinasi *Cost-Sensitive Random Forest*, *SMOTE*, dan *GridSearchCV* memberikan performa paling seimbang dengan akurasi 0,927267, *precision* 0,911458, *recall* 0,738397, dan *F1-score* 0,815851. Jika

dibandingkan dengan model *RF* standar, model ini mengalami penurunan akurasi sekitar 0,66% serta penurunan *precision* sebesar 6,27%, namun memberikan peningkatan *recall* sebesar 2,20%. Peningkatan *recall* ini sangat penting karena mencerminkan kemampuan model dalam mengenali debitur berisiko yang merupakan kelas minoritas. Dibandingkan dengan *CS-RF* biasa, model gabungan meningkatkan *recall* sebesar 2,40% serta mengalami *trade-off* kecil pada *F1-score* sebesar −1,03%, yang tetap menguntungkan dalam konteks deteksi risiko kredit.

Tabel 5. Ringkasan Performa Model

No	Model	Akurasi	*Precision*	*Recall*	*F1-Score*
1.	CS-RF + SMOTE + GridSearch	0.927267	0.911458	0.738397	0.815851
2.	CS-RF + SMOTE	0.926040	0.910122	0.733474	0.812305
3.	RF + SMOTE	0.926040	0.910122	0.733474	0.812305
4.	RF	0.933712	0.972328	0.716596	0.825101
5.	CS-RF	0.934326	0.978805	0.714487	0.826016

Hasil ini juga selaras dengan *confusion matrix* pada Gambar 6, di mana model mampu mengidentifikasi 1.050 debitur berisiko dengan benar (*True Positive*), sementara masih terdapat 372 data kelas berisiko yang salah diprediksi sebagai aman (*False Negative*). Jumlah *True Positive* yang cukup tinggi menunjukkan bahwa model memiliki kemampuan yang baik dalam menangkap pola pada kelas minoritas, meskipun tetap terdapat *trade-off* berupa kesalahan prediksi pada sebagian data. Secara keseluruhan, integrasi ketiga metode tersebut menghasilkan model yang lebih *robust* serta memberikan performa yang stabil dan seimbang pada tugas klasifikasi *credit scoring* dengan distribusi kelas yang tidak seimbang.



Gambar 6. *Confusion Matriks CS-RF + SMOTE + GridSearch*

4.3. Pembahasan Hasil Temuan

Berdasarkan hasil yang ditunjukkan pada Tabel 4, terlihat bahwa model *Cost-Sensitive Random Forest* yang dikombinasikan dengan *SMOTE* dan *GridSearchCV* menghasilkan performa yang sangat tinggi pada data *training* serta tetap stabil pada data *testing*. Performa sempurna pada data *training* menunjukkan bahwa model mampu mempelajari pola data latih dengan baik, sedangkan penurunan performa yang wajar pada data *testing* mengindikasikan bahwa model masih mampu melakukan *generalisasi* terhadap data yang belum pernah dilatih sebelumnya.

Jika dibandingkan antar skenario pada Tabel 5, kombinasi *Cost-Sensitive Random Forest*, *SMOTE*, dan *GridSearchCV* menghasilkan performa yang paling seimbang, khususnya pada metrik *recall*. Meskipun nilai *accuracy* dan *precision* sedikit lebih rendah dibandingkan *Random Forest* standar, peningkatan nilai *recall* menunjukkan bahwa model lebih efektif dalam mendeteksi debitur berisiko gagal bayar yang merupakan kelas minoritas. Hal ini sejalan dengan

tujuan utama sistem *credit scoring*, yaitu meminimalkan risiko kesalahan dalam mengklasifikasikan debitur berisiko sebagai debitur aman.

Temuan ini memperkuat hasil Putra dan Rumini [7] mengenai dampak ketidakseimbangan kelas yang menyebabkan bias terhadap kelas mayoritas. Hasil penelitian ini juga melengkapi temuan Syah et al. [9], dengan menunjukkan bahwa akurasi tinggi pada *Random Forest* belum tentu mencerminkan kemampuan deteksi risiko yang optimal tanpa penanganan khusus terhadap kelas minoritas.

Selain itu, hasil penelitian ini selaras dengan Nanda et al. [10] yang menunjukkan bahwa penerapan *SMOTE* mampu meningkatkan *recall*, serta mendukung Mienye dan Sun [11] yang membuktikan efektivitas *Cost-Sensitive Random Forest* dalam meningkatkan sensitivitas model terhadap kelas minoritas. Temuan ini juga menguatkan penelitian Prieto-Egido et al. [12], yang mengombinasikan *Random Forest*, *SMOTE*, dan *Cost-Sensitive Learning* untuk meningkatkan performa klasifikasi risiko kredit.

Berbeda dengan penelitian terdahulu yang umumnya menerapkan metode tersebut secara parsial, penelitian ini mengintegrasikan *SMOTE*, *Cost-Sensitive Learning*, dan optimasi *hyperparameter* menggunakan *GridSearchCV* dalam satu alur pemodelan. Integrasi ini terbukti menghasilkan performa yang lebih stabil dan seimbang, sehingga memberikan kontribusi dalam pengembangan pendekatan credit scoring yang lebih efektif dalam menangani permasalahan ketidakseimbangan kelas.

5. Simpulan

Berdasarkan hasil penelitian, dapat disimpulkan bahwa penerapan kombinasi *Cost-Sensitive Random Forest*, *SMOTE*, dan *GridSearchCV* terbukti mampu meningkatkan performa model dalam memprediksi kelayakan kredit pada data yang tidak seimbang. Proses praproses, yang mencakup pembersihan data, transformasi fitur, normalisasi, dan pembagian dataset, berhasil menyiapkan data secara optimal sehingga model dapat mempelajari pola risiko kredit dengan lebih efektif. Penerapan *SMOTE* berhasil menyeimbangkan distribusi kelas, yang semula didominasi oleh nasabah lancar, sehingga model memperoleh representasi yang lebih baik terhadap kelas gagal bayar. Selanjutnya, penggunaan *GridSearchCV* menghasilkan konfigurasi optimal dengan nilai *n_estimators* sebesar 200, yang memberikan rata-rata *F1-score* tertinggi selama validasi silang.

Model gabungan menghasilkan performa yang paling seimbang dengan akurasi 0,927267, precision 0,911458, recall 0,738397, dan *F1-score* 0,815851. Dibandingkan model *Random Forest* standar, model ini memang mengalami penurunan kecil pada akurasi dan *precision*, namun memberikan peningkatan *recall*, sehingga lebih mampu mengenali debitur berisiko yang merupakan kelas minoritas. Hal ini juga tampak pada *confusion matrix*, di mana model berhasil mengidentifikasi 1.050 data gagal bayar dengan benar. Secara keseluruhan, integrasi ketiga metode tersebut menghasilkan model yang lebih kuat, adaptif, dan stabil dalam menangani ketidakseimbangan data, sehingga layak diterapkan sebagai pendekatan alternatif dalam sistem *credit scoring* modern. Ke depan, penelitian dapat dikembangkan melalui eksplorasi *feature selection*, penyesuaian *decision threshold*, atau penggunaan algoritma pembelajaran lain seperti *XGBoost* dan *LightGBM* untuk memperoleh performa yang lebih optimal.

Daftar Referensi

[1] A. Purwatiningsih and A. Suprayitno, “Efektivitas Pemberian Kredit Guna Meminimalkan Kredit Bermasalah Bank Mandiri Cabang Malang,” *Journal of Public and Business Accounting*, vol. 3, no. 2, pp. 108–118, Dec. 2022, doi: 10.31328/jopba.v3i2.281.

[2] N. Dwiastuti, “Pengaruh Kredit Perbankan Terhadap Pertumbuhan Ekonomi dan Hubungannya Dengan Kesejahteraan Masyarakat Kabupaten/Kota di Provinsi Kalimantan Barat,” in *Prosiding Seminar Akademik Tahunan Ilmu Ekonomi dan Studi Pembangunan 2020*, Pontianak, Oct. 2020, pp. 73–91.

[3] H. Wulandari and I. Lubis, “Pengaruh Pertumbuhan Kredit Modal Kerja dan Pertumbuhan Kredit Investasi pada UMKM terhadap Pertumbuhan Ekonomi Indonesia Periode 2013-2023,” *Progressus Humanitatis*, vol. 1, no. 1, pp. 191–204, 2025, doi: 10.70285/s168kc17.

[4] OJK, “Laporan Surveillance Perbankan Indonesia TW III 2024,” 2024.

[5] T. Yulianti, A. H. Cahyana, M. Komarudin, Y. Mulyani, and H. D. Septama, “Penilaian Pembayaran Kredit dengan Logistic Regression dan Random Forest pada Home Credit,” *Pseudocode*, vol. 11, no. 2, pp. 79–88, Sep. 2024, doi: 10.33369/pseudocode.11.2.79-88.

[6] A. M. Jannah, A. Habibi, and B. M. Basuki, “Implementasi Metode Naïve Bayes Classifier Pada Machine Learning Untuk Sistem Alternatif Credit Scoring,” *SCIENCE ELECTRO*, vol. 19, no. 1, pp. 17–4, 2025.
[7] H. Putra and Rumini, “Comparative Study of Logistic Regression, Random Forest, and XGBoost for Bank Loan Approval Classification,” *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 5, p. 2822, 2025, [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC
[8] A. Adrian and I. Verawati, “Analisis Performa Logistic Regression dan Random Forest dalam Klasifikasi Kelayakan Penerimaan Kredit,” *IJCSR: The Indonesian Journal of Computer Science Research*, vol. 4, no. 2, pp. 148–158, Jul. 2025, [Online]. Available: https://subset.id/index.php/IJCSR
[9] I. Syah, S. Sutono, and M. Mulyanto, “Penerapan Algoritma Random Forest Dalam Pencarian Pola Klaim Bpjs Rumah Sakit Umum Kumala Siwi Mijen Kudus,” *JURNAL LENTERA BISNIS*, vol. 14, no. 3, pp. 3986–3995, Sep. 2025, doi: 10.34127/jrlab.v14i3.1770.
[10] N. N. A. Nanda, Y. Farida, and W. D. Utami, “Implementation of SMOTE to Improve the Performance of Random Forest Classification in Credit Risk Assessment in Banking,” *INTENSIF: Jurnal Ilmiah Penelitian dan Penerapan Teknologi Sistem Informasi*, vol. 9, no. 2, pp. 158–177, Jul. 2025, doi: 10.29407/intensif.v9i2.23930.
[11] I. D. Mienye and Y. Sun, “Performance analysis of cost-sensitive learning methods with application to imbalanced medical data,” *Inform Med Unlocked*, vol. 25, pp. 1–10, Jan. 2021, doi: 10.1016/j.imu.2021.100690.
[12] I. Prieto-Egido, A. Guerrero-Curieses, A. Martínez-Fernández, and J. L. Rojo-Álvarez, “Identifying high-risk pregnancies in rural areas with machine-manifold learning,” *Engineer Applications of Artificial Intelligence*, vol. 163, pp. 1–15, Oct. 2025, doi: 10.21950/WDZX9.
[13] K. D. Tzimourta, M. G. Tsipouras, P. Angelidis, D. G. Tsalikakis, and E. Orovou, “Maternal Health Risk Detection: Advancing Midwifery with Artificial Intelligence,” *Healthcare (Switzerland)*, vol. 13, no. 7, pp. 1–21, Apr. 2025, doi: 10.3390/healthcare13070833.
[14] L. Breiman, “Random Forests,” *Mach Learn*, vol. 45, no. 1, pp. 1–32, Jan. 2001.
[15] M. Z. Sarwani, M. Khoiron, and M. Udin, “Optimization of the Naïve Bayes Classifier Algorithm Using Cost-Sensitive Learning to Detect Lung Diseases with an Imbalanced Dataset,” *Journal of Artificial Intelligence and Software Engineering (J-AISE)*, vol. 5, no. 1, p. 332, Mar. 2025, doi: 10.30811/jaise.v5i1.6474.
[16] I. Araf, A. Idri, and I. Chairi, “Cost-sensitive learning for imbalanced medical data: a review,” *Artif Intell Rev*, vol. 57, no. 4, p. 80, Apr. 2024, doi: 10.1007/s10462-023-10652-8.
[17] A. W. Anggraeni, A. R. B. Jamroni, G. Samudra, A. Sarif, and Wiyanto, “Efektivitas Teknik SMOTE Dalam Meningkatkan Performa Naïve Bayes Deteksi Gangguan Kecemasan Mahasiswa,” *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 12, no. 3, pp. 105–15, 2025, [Online]. Available: http://jurnal.mdp.ac.id
[18] B. N. Nuzululnisa and H. Hairani, “Analisis Kinerja Model Random Forest dengan Teknik Manhattan-SMOTE pada Deteksi Fraud Transaksi Kartu Kredit Imbalance,” in *SEMINAR NASIONAL CORISINDO*, Mataram, Sep. 2025, pp. 65–71.
[19] N. A. Suhartono and V. J. L. Engel, “Penerapan Extreme Gradient Boosting (XGBoost) dengan SMOTE untuk Deteksi Penipuan Kartu Kredit,” Institut Teknologi Harapan Bangsa, Bandung, 2022.