Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 430-439
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

Non-Performing Loan Prediction for Credit Application Analysis Using Feature Selection and Ensemble Methods

David Jefri Aruan[1], Rusdah[2*], Ahmad Pudoli[3]

[1,2]Magister Ilmu Komputer, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia
[3]Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia
Email: [1]2411600675@student.budiluhur.ac.id, [2*]rusdah@budiluhur.ac.id, [3]ahmad.pudoli@budiluhur.ac.id
(*: *corresponding author*)

Abstract - *Non-Performing Loans (NPL) are a fundamental indicator of a financial institution's asset health, reflecting loans that fail to meet interest or principal payment obligations as agreed. A high NPL ratio negatively impacts a bank's financial performance, such as decreased profitability as measured by Return on Assets (ROA) and decreased liquidity. Bank Indonesia sets an NPL tolerance limit of 5% of total credit provided by banking financial institutions. Therefore, a predictive model is needed that can detect the possibility of customers experiencing NPLs early. This study aims to identify relevant factors in predicting NPLs and create an NPL prediction model based on these factors. The contribution of this study lies in combining the results of three feature selection techniques: Chi-Square, Mutual Information, and Random Forest feature importance, using the average score eliminated by the Recursive Feature Elimination technique. Several ensemble algorithms, namely Random Forest, XGBoost, Gradient Boosting, and LightGBM, were explored to produce the best-performing model. Then, hyperparameter tuning was performed on the best model. The Random Forest model produced the best performance, with 92.17% accuracy, 78.1% precision, 98.1% recall, and 95.5% AUC. Hyperparameter tuning was shown to improve recall, thus improving the model's ability to measure how much positive data (Current class) was successfully predicted by the model. The results of this study can assist management in making credit decisions. Thus, it is hoped that it can help reduce the number of NPL cases.*

Keywords*: Collectability, Credit Application, Ensemble Method, Feature Selection, Hyperparameter Tuning, Non-Performing Loan.*

Abstrak – *Non-Performing Loan* (NPL) merupakan indikator fundamental kesehatan aset sebuah lembaga keuangan, mencerminkan pinjaman yang gagal memenuhi kewajiban pembayaran bunga atau pokok sesuai kesepakatan. Tingginya rasio NPL berdampak negatif pada kinerja keuangan bank, seperti penurunan profitabilitas yang diukur dengan Return on Assets (ROA) dan penurunan likuiditas. Bank Indonesia menetapkan batas toleransi NPL sebesar 5% dari total kredit yang diberikan oleh lembaga keuangan bank. Sehingga diperlukan model prediktif yang dapat mendeteksi kemungkinan nasabah mengalami NPL sejak awal. Penelitian ini bertujuan untuk menemukan faktor-faktor yang relevan dalam memprediksi NPL serta membuat model prediksi NPL berdasarkan faktor-faktor tersebut. Kontribusi penelitian ini terletak pada penggabungan hasil tiga teknik seleksi fitur Chi-Square, Mutual Information, dan Random Forest *feature importance*, menggunakan rata-rata nilai skor yang dieliminasi dengan teknik *Recursive Feature Elimination*. Beberapa algoritma ensemble, yaitu Random Forest, XGBoost, Gradient Boosting, dan LightGBM, dieksplorasi untuk menghasilkan model dengan performa terbaik. Kemudian dilakukan *hyperparameter tuning* pada model terbaik. Model Random Forest terbukti menghasilkan performa terbaik dengan akurasi 92,17%, presisi 78,1%, *recall* 98,1% dan AUC 95,5%. *Hyperparameter tuning* terbukti meningkatkan *recall*, sehingga *hyperparameter tuning* mampu meningkatkan kemampuan model dalam mengukur seberapa banyak data positif (kelas Lancar) yang berhasil diprediksi oleh model. Hasil penelitian ini dapat membantu manajemen dalam mengambil keputusan pemberian kredit. Dengan demikian, diharapkan dapat membantu menurunkan jumlah kasus NPL.

Kata Kunci: *Ensemble Method, Hyperparameter Tuning*, Kolektabilitas, *Non-Performing Loan*, Pengajuan Kredit, Seleksi Fitur.

1. PENDAHULUAN

Sektor perbankan merupakan salah satu pilar utama dalam sistem perekonomian nasional karena memiliki fungsi strategis sebagai lembaga intermediasi keuangan yang menghimpun dana dari masyarakat dan menyalurkannya kembali dalam bentuk kredit kepada berbagai sektor, baik sektor produktif maupun sektor konsumtif. Penyaluran kredit tidak hanya berperan dalam mendorong pertumbuhan ekonomi, tetapi juga menjadi sumber utama pendapatan bank melalui penerimaan bunga kredit [1]. Kredit yang disalurkan kepada sektor usaha, Usaha Mikro Kecil dan Menengah (UMKM), maupun individu memiliki kontribusi besar terhadap peningkatan investasi, penciptaan lapangan kerja, serta peningkatan daya beli masyarakat. Namun demikian, aktivitas pemberian kredit juga merupakan aktivitas dengan tingkat risiko paling tinggi yang dihadapi industri perbankan, terutama risiko gagal bayar atau kredit bermasalah yang dikenal dengan istilah *Non-Performing Loan* (NPL) [2].

NPL merupakan rasio yang menunjukkan besarnya jumlah kredit bermasalah terhadap total kredit yang diberikan oleh bank. Kredit dikategorikan sebagai NPL apabila debitur tidak mampu memenuhi kewajiban pembayaran pokok maupun bunga sesuai dengan jadwal yang telah disepakati. Secara umum, pinjaman diklasifikasikan sebagai NPL apabila terjadi tunggakan pembayaran selama lebih dari 90 hari untuk pinjaman

Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 430-439
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

komersial dan lebih dari 180 hari untuk kredit konsumtif [3]. Tingginya rasio NPL akan memberikan dampak yang signifikan terhadap kesehatan bank karena menyebabkan meningkatnya biaya pencadangan kerugian kredit, menurunkan profitabilitas yang diukur melalui Return on Assets (ROA), mengurangi likuiditas, serta menurunkan kemampuan bank dalam menyalurkan kredit baru kepada masyarakat [2], [4].

Di Indonesia, pengelolaan risiko kredit menjadi perhatian serius bagi regulator. Melalui Peraturan Bank Indonesia Nomor 15/2/PBI/2013 Pasal 5 ayat (2) huruf a, ditetapkan bahwa batas maksimal rasio NPL yang dapat ditoleransi adalah sebesar 5% dari total kredit yang disalurkan. Rasio NPL yang melebihi batas tersebut menunjukkan adanya penurunan kualitas aset produktif serta mengindikasikan memburuknya kesehatan bank. Sebaliknya, rasio NPL yang berada di bawah batas ketentuan menunjukkan bahwa kualitas kredit masih dalam kategori sehat. Oleh karena itu, pengendalian NPL menjadi salah satu fokus utama dalam penerapan manajemen risiko perbankan agar stabilitas sistem keuangan tetap terjaga.

Fenomena peningkatan kredit bermasalah juga terjadi di Bank Pembangunan Daerah Jakarta (Bank Jakarta). Berdasarkan data NPL selama periode 2020–2024 dari Bank Jakarta, terjadi peningkatan jumlah kasus kredit bermasalah dalam tiga tahun terakhir dengan kenaikan yang cukup signifikan pada tahun 2024. Kondisi tersebut menunjukkan bahwa proses analisis kelayakan kredit yang selama ini diterapkan masih memiliki keterbatasan dalam mengidentifikasi calon debitur yang berpotensi mengalami gagal bayar. Apabila kondisi ini terus berlanjut, risiko kerugian bank akan semakin besar dan dapat memengaruhi stabilitas finansial lembaga dan menghambat kemampuan penyaluran kredit baru [2]. Beberapa penelitian telah dilakukan untuk memprediksi nasabah NPL di perbankan dengan *single classifier* seperti Logistic Regression (LR), Support Vector Machine (SVM), Decision Tree (DT), K-Nearest Neighbors (KNN), Naïve Bayes (NB). Algoritma ini sering digunakan sebagai *baseline* untuk perbandingan atau sebagai komponen dalam metode *ensemble*. Kinerja algoritma tersebut bervariasi; misalnya, LR memiliki akurasi tertinggi dalam studi [5] untuk pinjaman pertanian, sementara Naïve Bayes dapat menjadi yang terbaik dalam studi lain [6]. Namun, LR, Stochastic Gradient Descent Classifier (SGDC), dan NB cenderung mengalami kesulitan pada *dataset* yang tidak seimbang [7].

Kondisi ketika jumlah data nasabah bermasalah jauh lebih sedikit dibandingkan nasabah dengan kredit lancar, ketidakseimbangan data menyebabkan model cenderung bias terhadap kelas mayoritas sehingga akurasi keseluruhan menjadi tinggi, tetapi kemampuan mendeteksi kelas NPL justru rendah. Untuk mengatasi keterbatasan tersebut, berbagai penelitian mulai mengembangkan pendekatan *ensemble learning* yang menggabungkan banyak model klasifikasi sehingga menghasilkan performa yang lebih stabil. *Random Forest,* salah satu algoritma *ensemble*, secara konsisten menunjukkan performa unggul dalam berbagai penelitian prediksi NPL [8], [9], [10], [11], [12], [13]. RF menonjol karena kemampuannya menangani *dataset* kompleks, akurasi prediktif yang superior, dan ketahanan yang kuat terhadap ketidakseimbangan kelas [6].

Selain RF, algoritma berbasis Gradient Boosting seperti *Extreme Gradient Boosting* (XGBoost) dan *Light Gradient Boosting Machine* (LightGBM) juga telah dieksplorasi untuk memprediksi NPL. Model *ensemble learning* ini juga menunjukkan kompetensi prediktif yang sangat kuat, sering kali mengungguli *classifier* berbasis pohon konvensional lainnya [14]. XGBoost dan LightGBM memiliki kemampuan prediktif yang lebih kuat [15]. Sebuah studi pada tahun 2025 secara spesifik menyebutkan bahwa XGBoost mencapai 99,4% akurasi untuk prediksi credit card default dan LightGBM mencapai 90,07% akurasi untuk permintaan pinjaman, meskipun masih menghadapi tantangan berupa rendahnya presisi pada kelas minoritas [15]. Selain pemilihan algoritma, kualitas prediksi NPL juga dipengaruhi oleh variabel prediktor yang digunakan. Berbagai penelitian menunjukkan bahwa variabel finansial seperti pendapatan, jumlah pinjaman, durasi kredit, skor kredit, riwayat pembayaran, jumlah tagihan, serta jumlah pembayaran memiliki kontribusi signifikan dalam memprediksi risiko gagal bayar [16][17].

Di sisi lain, variabel demografi seperti usia, pekerjaan, tingkat pendidikan, status perkawinan, pendapatan keluarga, serta karakteristik sosial ekonomi juga terbukti berpengaruh terhadap kemungkinan terjadinya NPL [18]. Temuan tersebut menunjukkan bahwa faktor mikro yang melekat pada karakteristik debitur memiliki pengaruh yang lebih dominan dibandingkan dengan faktor makroekonomi dalam memprediksi kredit bermasalah. Dengan demikian, pembangunan model prediksi yang memanfaatkan data internal bank secara komprehensif berpotensi menghasilkan tingkat akurasi yang lebih tinggi dibandingkan dengan model yang hanya menggunakan indikator ekonomi makro.

Meskipun berbagai penelitian telah menunjukkan keberhasilan penerapan *Machine Learning* dalam prediksi NPL, masih terdapat beberapa kesenjangan penelitian (*research gaps*). Pertama, sebagian besar penelitian hanya membandingkan algoritma klasifikasi tunggal seperti *Logistic Regression*, SVM, *Decision Tree*, KNN, dan Naïve Bayes, sementara studi yang mengevaluasi secara komprehensif berbagai metode ensemble menggunakan data historis bank di Indonesia masih relatif terbatas. Kedua, banyak penelitian masih menggunakan metrik akurasi sebagai indikator utama evaluasi model, padahal pada kasus NPL yang memiliki karakteristik *imbalanced dataset*, metrik seperti *precision, recall, F1-score*, dan *Area Under Curve* (AUC-ROC) memberikan gambaran yang lebih representatif tentang kemampuan model dalam mendeteksi nasabah berisiko tinggi. Ketiga, belum banyak

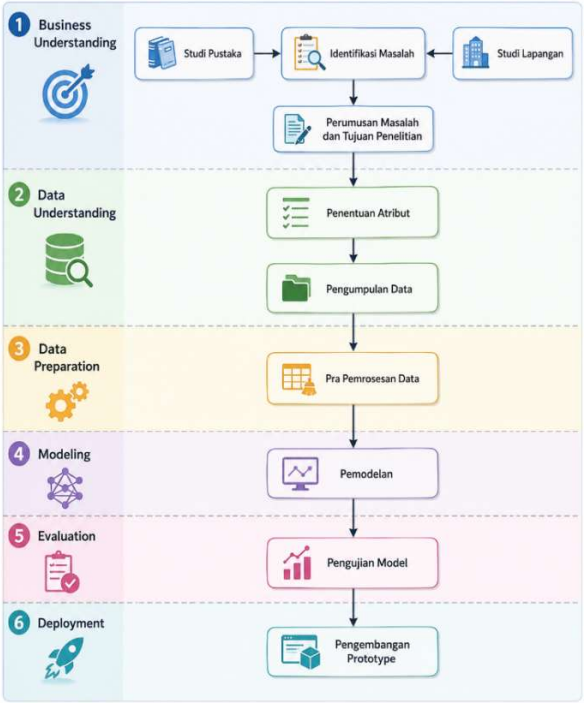
Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 430-439
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

penelitian yang memanfaatkan data operasional Bank Jakarta sebagai studi kasus, sehingga karakteristik spesifik debitur dan kebijakan kredit bank tersebut belum banyak dieksplorasi dalam pengembangan model prediksi.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menemukan faktor-faktor yang relevan dalam memprediksi NPL serta membangun model prediksi NPL berdasarkan faktor-faktor tersebut. Penggabungan tiga teknik seleksi fitur (Chi-Square, *mutual information*, dan Random Forest *feature importance*) menggunakan rata-rata skor, kemudian dieliminasi dengan teknik *Recursive Feature Elimination*, merupakan kontribusi dalam penelitian ini. Penelitian ini menggunakan data historis pengajuan kredit Bank Jakarta pada periode 2020–2024. Penelitian ini membandingkan beberapa algoritma klasifikasi dengan fokus pada metode *ensemble* untuk memperoleh model dengan performa terbaik. Kemudian dilakukan *hyperparameter tuning* pada algoritma terbaik. Evaluasi model dilakukan menggunakan *confusion matrix* dengan parameter *accuracy, precision, recall, F1-score*, dan *Area Under Curve* (AUC-ROC), sehingga diperoleh model yang tidak hanya memiliki akurasi tinggi, tetapi juga mampu mengidentifikasi calon debitur berisiko dengan lebih akurat. Hasil penelitian ini diharapkan dapat menjadi rekomendasi bagi Bank Jakarta dalam meningkatkan kualitas proses analisis kredit, memperkuat manajemen risiko, serta mendukung pengambilan keputusan pemberian kredit yang lebih efektif dan berbasis data.

2. METODE PENELITIAN

Metodologi *Cross-Industry Standard Process for Data Mining* (CRISP-DM) digunakan sebagai kerangka kerja penelitian dalam pengembangan model prediksi *Non-Performing Loan* (NPL). CRISP-DM dipilih karena merupakan metodologi yang sistematis dan banyak digunakan dalam penelitian *data mining* dan *machine learning*. Metodologi ini terdiri atas enam tahapan, yaitu *Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment*, sebagaimana ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.1. Pemahaman Bisnis (*Business Understanding*)

Tahap ini bertujuan untuk mengidentifikasi permasalahan bisnis yang berkaitan dengan meningkatnya jumlah kredit bermasalah di Bank DKI. Kegiatan dilakukan melalui wawancara dengan pihak *Marketing*, *Relationship Manager*, dan Kepala Cabang, serta observasi terhadap proses pengajuan kredit. Selain itu, dilakukan studi pustaka terhadap penelitian terdahulu mengenai prediksi NPL menggunakan *machine learning* untuk mengidentifikasi metode yang telah digunakan serta menemukan peluang pengembangan penelitian. Tahap ini menghasilkan rumusan masalah dan tujuan penelitian.

Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 430-439
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

2.2. Pemahaman Data (*Data Understanding*)

Data sekunder pada penelitian ini bersumber dari transaksi pengajuan kredit di seluruh cabang Bank DKI selama periode 2020–2024. Berdasarkan hasil diskusi dengan pihak bisnis, ditetapkan atribut yang digunakan dalam penelitian, meliputi informasi demografi debitur, karakteristik produk kredit, hasil proses prescreening, serta status pengajuan kredit. Variabel target adalah kolektabilitas kredit yang selanjutnya diklasifikasikan menjadi dua kelas, yaitu Lancar dan Tidak Lancar (NPL). Secara keseluruhan digunakan 23 atribut prediktor dan 1 atribut target. Setelah data terkumpul, dilakukan Exploratory Data Analysis (EDA) untuk memahami distribusi data, mengidentifikasi karakteristik setiap variabel, serta memperoleh insight yang dapat mendukung proses pemodelan.

2.3. Persiapan Data (*Data Preparation*)

Tahap ini bertujuan menghasilkan dataset yang siap digunakan dalam proses pembelajaran model. Kegiatan yang dilakukan meliputi pembersihan data, reduksi atribut, dan transformasi data. Pembersihan data dilakukan untuk mengatasi *missing value*, data duplikat, inkonsistensi data, serta *outlier*. Reduksi atribut dilakukan untuk menghapus atribut yang tidak relevan dalam pengembangan model prediksi NPL. Transformasi data dilakukan dengan membentuk atribut baru, seperti usia berdasarkan tanggal input pengajuan dan tanggal lahir, serta melakukan generalisasi label kolektabilitas menjadi dua kategori, yaitu Lancar dan Tidak Lancar.

Selanjutnya dilakukan *feature selection* menggunakan pendekatan *filter* dan *wrapper* untuk memperoleh atribut yang paling berpengaruh terhadap prediksi NPL. Teknik seleksi fitur Chi-Square, Mutual Information, dan Random Forest feature importance dikombinasikan dengan menggunakan rata-rata nilai skor, kemudian dieliminasi menggunakan teknik *Recursive Feature Elimination*. Setelah seleksi fitur, dilakukan penentuan data training dan data testing. Beberapa skenario pembagian data dengan komposisi 60:40, 70:30, dan 80:20, serta 10-*fold cross-validation*, dieksplorasi untuk memperoleh model yang lebih stabil dan mampu melakukan generalisasi pada data baru.

2.4. Pemodelan (*Modeling*)

Pada tahap pemodelan dilakukan eksplorasi terhadap empat metode ensemble: Random Forest [12], [13], XGBoost [19], LightGBM, dan Gradient Boosting [20]. Eksplorasi keempat algoritma tersebut dilakukan untuk memperoleh performa model terbaik dengan nilai AUC tertinggi menggunakan komposisi dataset yang paling optimal. Model terbaik kemudian dioptimalkan menggunakan *Grid Search Hyperparameter Tuning* dengan 50 kombinasi parameter untuk memperoleh konfigurasi model yang menghasilkan performa prediksi terbaik.

2.5. Evaluasi (*Evaluation*)

Evaluasi model dilakukan menggunakan Confusion Matrix dengan empat komponen utama, yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Berdasarkan nilai tersebut, dihitung beberapa metrik evaluasi yang meliputi akurasi, presisi, *recall*, F1-Score, dan Area Under the Curve (AUC-ROC). Penggunaan beberapa metrik evaluasi diperlukan karena dataset prediksi NPL memiliki karakteristik *imbalanced*, sehingga akurasi saja belum mampu menggambarkan kemampuan model dalam mengidentifikasi debitur yang berpotensi mengalami kredit bermasalah. Oleh karena itu, *precision*, *recall*, F1-*Score*, dan AUC digunakan sebagai indikator tambahan untuk memilih model terbaik.

2.6. *Deployment*

Tahap akhir penelitian adalah Deployment, yaitu melakukan interpretasi terhadap hasil pemodelan untuk menjawab tujuan penelitian serta memberikan rekomendasi kepada Bank DKI dalam meningkatkan proses analisis risiko kredit. Model dengan performa terbaik selanjutnya diimplementasikan dalam bentuk prototipe sistem prediksi *Non-Performing Loan* (NPL). Prototype ini diharapkan mampu membantu bank mengidentifikasi calon debitur berisiko sejak tahap awal sehingga dapat meningkatkan kualitas keputusan kredit dan menekan potensi terjadinya kredit bermasalah.

3. HASIL DAN PEMBAHASAN

3.1 Pemahaman Data (*Data Understanding*)

Penelitian ini menggunakan data sekunder yang berasal dari transaksi pengajuan kredit di seluruh cabang Bank DKI selama periode 2020–2024. Dataset terdiri atas 1.500 data debitur dengan 23 atribut prediktor yang meliputi karakteristik demografi, informasi produk kredit, hasil prescreening, serta status pengajuan kredit. Variabel target adalah kolektabilitas kredit yang digunakan untuk mengidentifikasi status lancar dan tidak lancar (*Non-Performing Loan/NPL*).

Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 430-439
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

Exploratory Data Analysis (EDA) menunjukkan bahwa dataset memiliki karakteristik *imbalanced*, di mana 82% data termasuk kategori lancar dan 18% termasuk kategori kolektabilitas 3–5 (NPL). Setelah dilakukan generalisasi label sesuai ketentuan Direktorat Jenderal Kekayaan Negara, komposisi kelas berubah menjadi 73% Lancar dan 27% Tidak Lancar. Kondisi ini mencerminkan karakteristik umum data perbankan, sehingga diperlukan algoritma yang mampu menangani ketidakseimbangan kelas dengan baik.

Analisis deskriptif juga menunjukkan bahwa pekerjaan merupakan salah satu karakteristik yang paling berkaitan dengan status kolektabilitas. Debitur yang bekerja sebagai Pegawai Pemerintahan/Lembaga Negara memiliki jumlah kasus terbanyak, diikuti oleh Pengajar, Wiraswasta, dan Administrasi Umum. Namun, tingginya jumlah kasus pada kelompok pekerjaan tertentu tidak selalu menunjukkan tingkat risiko yang tinggi karena proporsinya dipengaruhi oleh jumlah debitur dalam kelompok tersebut.

Tabel 1. Penjelasan Atribut dari Data Awal

No	Atribut	Tipe Data	Keterangan	Contoh data
1.	Kode Pekerjaan	Kategorikal	Kode Pekerjaan	37: Pegawai Pemerintahan; 09: Pengajar (Guru/Dosen)
2.	Keterangan Pekerjaan	Kategorikal	Keterangan Pekerjaan	
3.	Tgl_lahir	Kategorikal	Tanggal Lahir Debitur	
4.	Status Pernikahan	Kategorikal	Kode Status Pernikahan	K; B; D
5.	Keterangan Status Pernikahan	Kategorikal	Keterangan Status Pernikahan	K=Kawin; B= Belum Kawin dan D=Cerai
6.	Kelurahan	Kategorikal	Kelurahan domisili Debitur	-
7.	Kecamatan	Kategorikal	Kecamatan domisili Debitur	-
8.	Kota	Kategorikal	Kota domisili Debitur	-
9.	Provinsi	Kategorikal	Propinsi domisili Debitur	-
10.	Produk	Kategorikal	Jenis Produk Kredit	Konsumer; Mikro
11.	Kode sub_produk	Kategorikal	Kode Sub Produk Kredit	KMG; KPR; Mikro KUR; Mikro NON-KUR
12.	Keterangan Sub Produk	Kategorikal	Keterangan Sub Produk Kredit	Kredit Multi Guna (KMG); Kredit Pemilikan Rumah (KPR); Mikro Kredit Usaha Rakyat (Mikro KUR);Mikro Non-Kredit Usaha Rakyat (Mikro NON-KUR)
13.	Tanggal_input	Kategorikal	Tanggal Input Pengajuan	-
14.	Plafond	Numerik	Jumlah pinjaman yang diajukan	-
15.	Jangka_waktu	Numerik	Lama pinjaman dalam bulan	-
16.	Hasil Prescreening SLIK	Kategorikal	Status Hasil Prescreening Sistem Layanan Informasi Keuangan (SLIK)	Low; Medium; High
17.	Hasil prescreening SIKPKUR	Kategorikal	Status Hasil Prescreening Sistem Informasi Kredit Program Kredit Usaha Rakyat (SIKP KUR)	Tanda (-) = Bila sub produk selain Mikro KUR; Terdaftar KUR Tidak Terdaftar KUR Terdaftar KUR di Bank Lain
18.	Hasil prescreening DUKCAPIL	Kategorikal	Status Hasil Prescreening Kependudukan dan Catatan Sipil (Dukcapil)	Sesuai; Tidak Sesuai
19.	Hasil prescreening DHNBI	Kategorikal	Status Hasil Prescreening Daftar Hitam Nasional BI	Tidak; Ya
20.	Hasil prescreening	Kategorikal	Status Hasil Prescreening Keseluruhan	Lolos; Lolos Bersyarat; Tidak Lolos
21.	Status	Kategorikal	Status Pengajuan Debitur	Accept; Reject;Waiting Approval; Lolos Bersyarat
22.	Kolektabilitas	Kategorikal	Kolektabilitas Debitur	1, 2, 3, 4, atau 5
23.	Keterangan Kolektabilitas	Kategorikal	Keterangan Kolektabilitas Debitur	1: Lancar; 2: Dalam Perhatian Khusus; 3: Kurang Lancar 4: Diragukan; 5: Macet

3.2 Persiapan Data (*Data Preparation*)

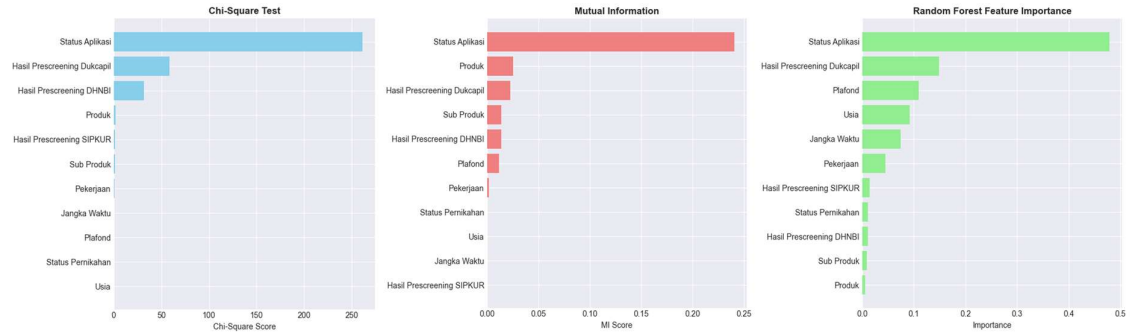
Sebelum masuk ke tahap pemodelan, dilakukan beberapa kegiatan untuk meningkatkan kualitas data. Pertama adalah reduksi data dengan menghapus atribut yang bersifat redundan, seperti atribut keterangan yang

Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 430-439
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

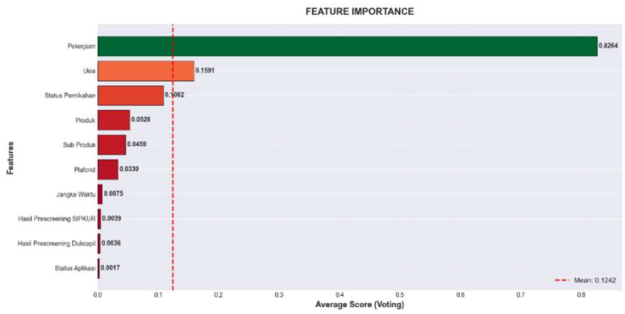
memiliki informasi yang sama dengan atribut kode. Reduksi data bertujuan untuk menyederhanakan data dengan mengurangi jumlah dimensi tanpa kehilangan informasi relevan [21]. Sebagai contoh, hal ini dapat dilihat pada Tabel 1. Atribut Status Pernikahan yang berisi kode K, B, dan D redundan dengan atribut Keterangan Status Pernikahan yang berisi Kawin (K), Belum Kawin (B), dan Cerai (D). Atribut yang dihapus adalah Keterangan Pekerjaan, Keterangan Status Pernikahan, Keterangan Sub Produk, dan Keterangan Kolektabilitas.

Setelah reduksi, jumlah atribut berkurang dari 23 menjadi 19. Selanjutnya dilakukan data *cleaning* yang menunjukkan bahwa seluruh dataset tidak memiliki missing value maupun *outlier*, sehingga tidak diperlukan proses imputasi maupun pembersihan data. Tahap transformasi meliputi pembentukan atribut usia berdasarkan tanggal lahir serta generalisasi label kolektabilitas menjadi dua kelas, yaitu Lancar dan Tidak Lancar. Selain itu, dilakukan transformasi terhadap atribut hasil *prescreening* SIKPKUR agar seluruh nilai memiliki representasi yang konsisten. Untuk memperoleh atribut yang paling relevan untuk prediksi NPL, penelitian ini mengeksplorasi empat teknik seleksi fitur, yaitu Chi-Square Test, *Mutual Information* (MI), *Random Forest Feature Importance* dan penggabungan tiga teknik tersebut menggunakan *Recursive Feature Elimination* (RFE).

Gambar 2 menunjukkan bahwa teknik seleksi fitur Chi-Square Test dan Random Forest *feature importance* menghasilkan sebelas atribut, sedangkan teknik *mutual information* menghasilkan tujuh atribut. Kemudian, semua atribut tersebut dikombinasikan menggunakan rata-rata nilai skor dan dieliminasi dengan teknik RFE (Gambar 3). Hasil kombinasi ketiga teknik tersebut menghasilkan sepuluh atribut, Pekerjaan, Usia, Status Pernikahan, Produk, Sub Produk, Plafond, Jangka Waktu, Hasil *Prescreening* SIPKUR, Hasil *Prescreening* Dukcapil, dan Status Aplikasi. Atribut Pekerjaan menjadi atribut terpenting dengan skor RFE sebesar 0,83. Namun demikian, kesimpulan ini hanya berdasarkan data yang digunakan dalam penelitian ini, sehingga tidak dapat digeneralisasi.



Gambar 2. Hasil Eksplorasi Tiga Teknik Seleksi Fitur



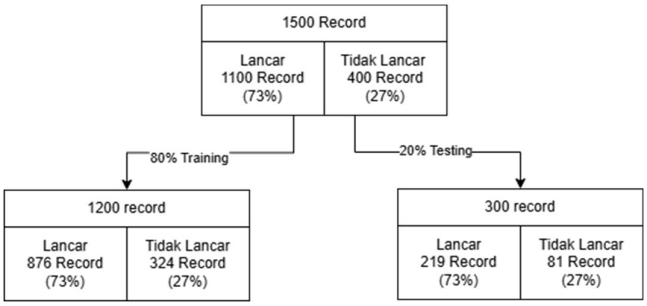
Gambar 3. Hasil Penggabungan 3 Teknik Seleksi Fitur dengan RFE

Tahap akhir *preprocessing* adalah pembentukan data training dan testing menggunakan beberapa skenario, yaitu split 60:40, 70:30, 80:20, serta 10-fold cross-validation. Seluruh skenario menggunakan *stratified random sampling* (Gambar 4) agar proporsi kelas tetap terjaga pada data training maupun data testing. Jumlah data training dan testing pada setiap skenario disajikan pada Tabel 2.

Tabel 2. Komposisi Data pada Setiap Skenario Penentuan Data

Skenario Penentuan Data	Jumlah Data Training	Jumlah Data Testing
Split Data 80-20	1.200	300
Split Data 70-30	1.050	450
Split Data 60-40	900	600
10-fold Cross Validation	1.500	1.500

Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 430-439
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index



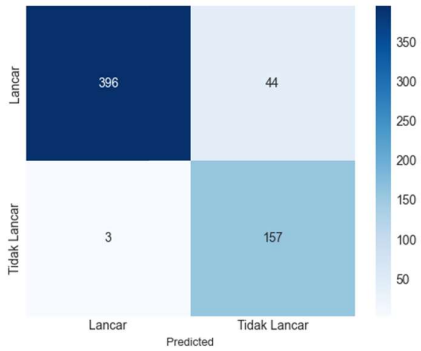
Gambar 4. Ilustrasi dari Teknik Split Data 80:20 dengan *Stratified Random Sampling*

3.3 Hasil Pemodelan

Empat algoritma ensemble, yaitu Random Forest, Gradient Boosting, XGBoost, dan LightGBM, dibandingkan pada empat skenario pembagian data. Evaluasi dilakukan menggunakan metrik akurasi dan Area Under Curve (AUC).

Tabel 3. Hasil Perbandingan Model

Skenario	Model	Akurasi	AUC
80-20 Split	LightGBM	0,920000	0,947386
	Gradient Boosting	0,916667	0,946875
	Random Forest	0,920000	0,945568
	XGBoost	0,910000	0,944830
70-30 Split	LightGBM	0,917778	0,951389
	Gradient Boosting	0,913333	0,953194
	Random Forest	0,920000	0,949192
	XGBoost	0,913333	0,950530
60-40 Split	LightGBM	0,915000	0,953445
	Gradient Boosting	0,915000	0,954524
	Random Forest	0,921667	0,954972
	XGBoost	0.916667	0.953445
10-Fold CV	LightGBM	0.914667	0.951773
	Gradient Boosting	0.912667	0.948068
	Random Forest	0.915333	0.950000
	XGBoost	0.914667	0.951523



Gambar 5. Confusion Matrix Model RF dengan Skenario 60:40

Hasil eksperimen pada Tabel 3 menunjukkan bahwa seluruh algoritma menghasilkan performa klasifikasi yang sangat baik dengan nilai AUC di atas 0,94. Namun demikian, Random Forest memberikan performa terbaik pada skenario split data 60:40, dengan akurasi 92,17% dan AUC 95,50% (Gambar 5). Metrik akurasi dan AUC ini menunjukkan kemampuan Random Forest yang sangat baik dalam membedakan debitur yang lancar dan yang tidak lancar. Berdasarkan klasifikasi Gorunescu [22], nilai AUC yang dihasilkan dalam penelitian ini merupakan *excellent classification*.

Untuk meningkatkan performa model, dilakukan *Grid Search Cross Validation* dengan 50 kombinasi *hyperparameter*. Parameter terbaik yang diperoleh adalah n_estimators = 200, max_depth = 10, min_samples_split

Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 430-439
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

= 2, min_samples_leaf = 2, max_features = sqrt, dan bootstrap = True. Model hasil tuning menghasilkan *accuracy* sebesar 92,17%, *precision* 77,83%, *recall* 98,75%, dan AUC 95,13%.

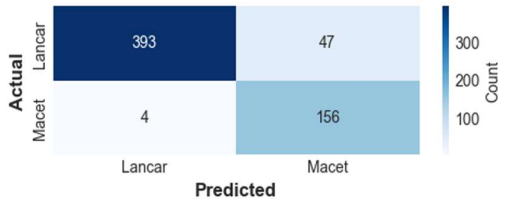
Walaupun nilai AUC mengalami sedikit penurunan dibandingkan dengan model dasar, nilai recall meningkat menjadi hampir 99%, yang menunjukkan bahwa model semakin mampu mengidentifikasi calon debitur yang berpotensi mengalami kredit bermasalah. Dalam konteks mitigasi risiko kredit, peningkatan recall lebih penting dibandingkan dengan peningkatan akurasi karena kesalahan mengklasifikasikan debitur bermasalah sebagai debitur lancar (*false negative*) dapat menimbulkan kerugian yang jauh lebih besar bagi bank.

3.4 Evaluasi dan Pembahasan

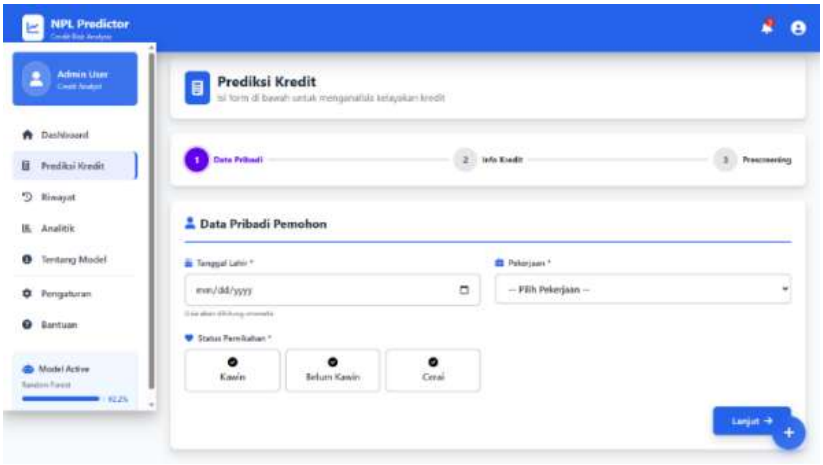
Evaluasi menggunakan Confusion Matrix menunjukkan bahwa model Random Forest mampu mempertahankan performa yang tinggi pada seluruh metrik evaluasi. Perbandingan antara model dasar dan model hasil *hyperparameter tuning* (Tabel 4 dan Gambar 6) menunjukkan bahwa kedua model memiliki tingkat akurasi yang sama, yaitu 97,12%, sedangkan model hasil *tuning* memberikan nilai *recall* yang lebih tinggi (98,75%) dibandingkan model dasar (98,12%). Sebaliknya, nilai *precision* mengalami sedikit penurunan dari 78,11% menjadi 77,83%, sedangkan nilai AUC berubah dari 95,53% menjadi 95,13%.

Tabel 4. Hasil perbandingan *Base Model* dan *Fine-Tuned Model*

Model	Akurasi	Presisi	Recall	AUC
Base Model (Random Forest)	97,12%	78,11	98,12	95,53%
Fine-Tuned Model	97,12%	77,83%	98,75	95,13%



Gambar 6. *Confusion Matrix* setelah dilakukan *Hyper Hyperparameter Tuning*



Gambar 7. Menu Prediksi Kredit Tab Data Pribadi

Hasil penelitian ini menunjukkan bahwa proses *hyperparameter tuning* tidak selalu meningkatkan seluruh metrik evaluasi secara bersamaan. Peningkatan kemampuan model dalam mendeteksi kelas minoritas sering kali diikuti dengan sedikit penurunan *precision* akibat bertambahnya jumlah *false positive*. Dalam kasus prediksi NPL, kondisi tersebut masih dapat diterima karena konsekuensi *false negative* jauh lebih besar dibandingkan *false positive*. Oleh karena itu, model dengan *recall* yang lebih tinggi dianggap lebih sesuai untuk mendukung proses mitigasi risiko kredit di lingkungan perbankan.

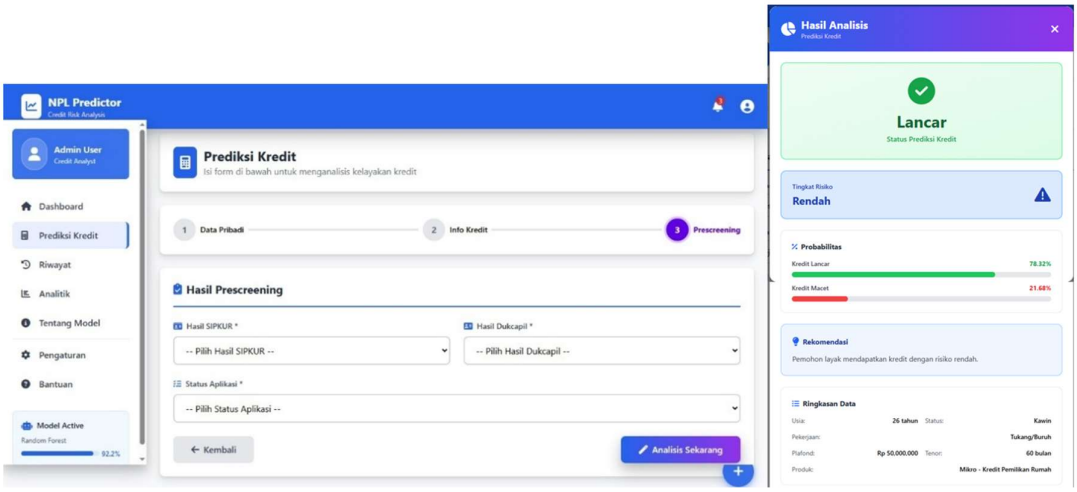
Temuan penelitian juga menunjukkan bahwa algoritma Random Forest memiliki performa yang lebih baik dibandingkan Gradient Boosting, XGBoost, dan LightGBM pada dataset yang digunakan. Demikian halnya

Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 430-439
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

dengan hasil penelitian [8], [14] yang membuktikan kemampuan Random Forest dalam menangani data dengan karakteristik yang kompleks serta ketidakseimbangan kelas. Selain itu, hasil seleksi fitur menunjukkan bahwa faktor-faktor mikro, seperti pekerjaan, usia, plafond pinjaman, serta hasil prescreening, lebih berpengaruh terhadap prediksi NPL dibandingkan dengan indikator makroekonomi, sehingga mendukung temuan [9] bahwa data internal bank memiliki kontribusi yang lebih besar dalam membangun model prediksi risiko kredit.

3.5 Implementasi Model

Sebagai implementasi hasil penelitian, model Random Forest dikembangkan menjadi prototipe sistem prediksi NPL berbasis web menggunakan Python 3.12 dan *framework* Flask. *Prototype* menyediakan modul prediksi kredit yang memungkinkan pengguna memasukkan data pribadi debitur, informasi kredit, serta hasil *prescreening* (Gambar 7). Berdasarkan data tersebut, sistem secara otomatis menghasilkan prediksi status kolektabilitas sebagai dasar pendukung keputusan dalam proses analisis kredit (Gambar 8). Prototipe ini dapat diimplementasikan sebagai *Decision Support System* (DSS) bagi manajemen dalam proses pemberian kredit. Dengan demikian, sistem diharapkan mampu membantu Bank DKI mengidentifikasi calon debitur berisiko lebih awal, meningkatkan kualitas analisis kredit, serta menekan potensi terjadinya *Non-Performing Loan* (NPL).



(a) (b)

Gambar 8. Menu Prediksi Kredit (a) dan Hasil Analisis (b)

4. KESIMPULAN

Kesimpulan penelitian ini adalah (1) hasil seleksi fitur kombinasi tiga teknik seleksi fitur yang dieliminasi dengan RFE (yang menjadi kontribusi dalam penelitian ini) menghasilkan sepuluh atribut yang relevan untuk membangun model prediksi NPL, yaitu Pekerjaan, Usia, Status Pernikahan, Produk, Sub Produk, Plafond, Jangka Waktu, Hasil Prescreening SIPKUR, Hasil Prescreening Dukcapil, dan Status Aplikasi. (2) Dari hasil eksplorasi beberapa model, diperoleh model terbaik, yaitu Random Forest, dengan komposisi data training dan testing 60:40 serta menggunakan *stratified random sampling*. Namun demikian, *hyperparameter tuning* yang dilakukan menggunakan GridSearch dengan 50 iterasi ternyata menurunkan performa model terbaik pada nilai AUC, yaitu turun 0,4 poin, dan presisi turun 0,28 poin. Namun, akurasi tetap dan nilai recall meningkat. Peningkatan *recall* pada model *fine-tuned* dapat meningkatkan kemampuan model dalam mengukur seberapa banyak data positif (kelas Lancar) yang berhasil diprediksi.

Data pada penelitian ini diambil dari seluruh cabang Bank Jakarta di Indonesia. Namun demikian, sampel dari tiap provinsi belum terwakili dengan komposisi yang seragam. Sehingga, pada penelitian berikutnya disarankan untuk menggunakan sampling dari tiap provinsi di Indonesia. Penelitian ini mengeksplorasi metode ensemble dengan teknik *bagging* (Random Forest) dan *boosting* (Gradient Boosting, XGBoost, dan LightGBM). Sehingga perlu dilakukan eksplorasi dengan teknik stacking menggunakan beberapa algoritma klasifikasi yang heterogen.

DAFTAR PUSTAKA

[1] J. C. Chandra, A. E. Putri, A. Z. Ongko, and E. T. Manurung, "The Effect of Non-Performing Loan (NPL) on Net Income at PT Bank Central Asia Tbk," *International Journal for accountancy*, vol. 12, no. 1, pp. 2610–2616, 2025.

[2] D. S. Bhakti, A. Prasetyo, and P. Arsi, "Implementation of Hyperparameter Tuning in Random Forest Algorithm for Loan Approval Prediction," *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 4, pp. 63–69, 2024, doi: 10.52436/1.jutif.2024.5.4.2032.

[3] T. Segal, "Nonperforming Loan (NPL) Definitions, Types, Causes, Consequences," Aug. 02, 2025. Accessed: Aug. 07, 2025. [Online]. Available: https://www.investopedia.com/terms/n/nonperformingloan.asp

[4] S. Masrom, R. A. Rahman, N. H. M. Shukri, N. A. Yahya, M. Z. A. Rashid, and N. B. Zakaria, "The nexus of corruption and non-performing loan: machine learning approach," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 32, no. 2, pp. 838–844, 2023, doi: 10.11591/ijeecs.v32.i2.pp838-844.

[5] M. A. Elnaggar, M. A. EL Azeem, and F. A. Maghraby, "Machine Learning Model for Predicting Non-performing Agricultural Loans," in *Advances in Intelligent Systems and Computing*, Springer, 2020, pp. 395–404. doi: 10.1007/978-3-030-44289-7_37.

[6] A. Soomro, "Loan Default Prediction Using Machine Learning Algorithms: A Systematic Literature Review 2020 - 2023," *Pakistan Journal of Life and Social Sciences (PJLSS)*, vol. 22, no. 2, 2024, doi: 10.57239/PJLSS-2024-22.2.00469.

[7] A. Soomro, H. Zakariyah, S. M. A. Aftab, M. Muflehi, A. Shah, and S. Meraj, "Loan Default Prediction Using Machine Learning Algorithms: A Systematic Literature Review 2020–2023," *Pak. J. Life Soc. Sci.*, vol. 22, no. 2, pp. 6234–6253, 2024, doi: 10.57239/PJLSS-2024-22.2.00469.

[8] M. Abdullah, M. A. F. Chowdhury, A. Uddin, and S. Moudud-Ul-Huq, "Forecasting nonperforming loans using machine learning," *J. Forecast.*, vol. 42, no. 7, pp. 1664–1689, 2023, doi: 10.1002/for.2977.

[9] J. P. Noriega, L. A. Rivera, and J. A. Herrera, "Machine Learning for Credit Risk Prediction: A Systematic Literature Review," *Data (Basel).*, vol. 8, no. 11, 2023, doi: 10.3390/data8110169.

[10] M. Annisa and Rusdah, "Prediction of Non-Performing Loans for Credit Application Analysis of Rural Bank Using Random Forest," in *2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, IEEE, Oct. 2022, pp. 111–114. doi: 10.23919/EECSI56542.2022.9946628.

[11] M. Muhajir and J. Widiastuti, "Random Forest Method to Customer Classification Based on Non-Performing Loan in Micro Business," *Jurnal Online Informatika*, vol. 7, no. 2, pp. 177–183, 2022, doi: 10.15575/join.v7i2.842.

[12] D. B. Saputra, V. Atina, and F. E. Nastiti, "Penerapan Model CRISP-DM pada Prediksi Nasabah Kredit Menggunakan Algoritma Random Forest," *IDEALIS : InDonEsiA journaL Information System*, vol. 7, no. 2, pp. 240–247, 2024, doi: 10.36080/idealis.v7i2.3244.

[13] A. C. Mawarni, R. Rusdah, L. L. Hin, and D. Anubhakti, "Deteksi Dini Gejala Awal Penyakit Diabetes Menggunakan Algoritma Random Forest," *IDEALIS : InDonEsiA journaL Information System*, vol. 6, no. 2, pp. 165–171, 2023, doi: 10.36080/idealis.v6i2.3018.

[14] T. Chaturvedi, S. Halder, U. S. kumar, N. Das, and S. Bittu, "Comparative Performance Analysis of Machine Learning Algorithms for Non-Performing Loan Prediction," in *2025 International Conference on Computational, Communication and Information Technology (ICCCIT)*, IEEE, Feb. 2025, pp. 13–18. doi: 10.1109/ICCCIT62592.2025.10928008.

[15] H. Begum, A. H. M. Z. Haq, and N. A. Shilpa, "machine Learning for Non-Performing Loan Prediction: Enhancing Credit Risk Management," in *The 1st International Online Conference on Risk and Financial Management,* MDPI, Jun. 2025.

[16] W. Andriani, Gunawan, and N. N. P. W. Naja, "Analisis perbandingan machine learning untuk prediksi kelayakan kredit perbankan pada Bank BRI Tegal," *IT-Explore: Jurnal Penerapan Teknologi Informasi dan Komunikasi*, vol. 4, no. 1, pp. 82–92, 2025, doi: 10.24246/itexplore.v4i1.2025.pp82-92.

[17] N. Hazizah, Sharipuddin, and A. Feranika, "Implementasi Algoritma Random Forest Dalam Klasifikasi Risiko Gagal Bayar Kartu Kredit Pada Nasabah Bank," *Jurnal Manajemen Teknologi Dan Sistem Informasi (JMS)*, vol. 5, no. 1, pp. 1050–1059, 2025, doi: 10.33998/jms.2025.5.1.2384.

[18] J. P. Noriega, L. A. Rivera, and J. A. Herrera, "Machine Learning for Credit Risk Prediction: A Systematic Literature Review," *Data (Basel).*, vol. 8, no. 11, p. 169, 2023, doi: 10.3390/data8110169.

[19] D. Pebrianti, H. Kurniawan, L. Bayuaji, and R. Rusdah, "XgBoost Hyper-Parameter Tuning Using Particle Swarm Optimization for Stock Price Forecasting," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 9, no. 4, pp. 1179–1195, 2024, doi: 10.26555/jiteki.v9i4.27712.

[20] Rusdah, B. A. Bregastantyo, and J. E. Riwurohi, "Prognosis Model of The Treatment Period of Tuberculosis Patients with Medication Compliance Parameters using The Gradient Boosting Algorithm," in *2023 6th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, Batam: IEEE, Dec. 2023, pp. 220–225. doi: 10.1109/ISRITI60336.2023.10467254.

[21] F. Ardiansyah, F. Hamdan, S. Sugiyanto, and I. Wahyu Siadi, "Klasifikasi Customer Relationship Management Menggunakan Dataset KDD Cup 2009 dengan Teknik Reduksi Dimensi," *Komputika : Jurnal Sistem Komputer*, vol. 11, no. 2, pp. 193–202, 2022, doi: 10.34010/komputika.v11i2.6498.

[22] F. Gorunescu, "Data Mining Techniques in Computer-Aided Diagnosis : Non-Invasive Cancer Detection," *World Acad. Sci. Eng. Technol.*, vol. 34, pp. 280–283, 2007, doi: doi.org/10.5281/zenodo.1059721.