Received: Juni 2026	Accepted: September 2026	Published: Oktober 2026

Analisis Penerapan *Explainable* Ai (Xai) Menggunakan *Shapley Additive Explanation* Pada Model *Random Forest* Untuk Prediksi Risiko Kredit

Anbirama Auberta Anggi Prabata[1*], Fajar Suryani[2], Dwi Hartanti[3]

[1,2,3]Teknik Informatika, Universitas Duta Bangsa Surakarta*

[*]Email: anbirama.prabata@gmail.com[1]

Abstract

The rapid development of artificial intelligence technology has encouraged financial institutions to adopt machine learning algorithms in the credit risk assessment process. However, high-performance models such as Random Forest tend to be black-box, making them difficult to understand by management and regulators, and hindering the audit process and strategic decision-making. This study aims to analyze the implementation of Explainable Artificial Intelligence (XAI) through the Shapley Additive Explanations (SHAP) method on the Random Forest model to produce a credit risk prediction system that is not only accurate but also transparent and accountable. The study uses a public historical credit dataset with 32,581 observations and 10 input variables, processed through preprocessing, model training, evaluation, and SHAP analysis stages, packaged in an interactive Streamlit-based web platform. Results show the Random Forest model achieves Accuracy 80.91%, Recall 82.85%, and AUC-ROC 0.8722. SHAP analysis confirms that Interest Rate is the most influential feature on default risk, followed by Loan Purpose and Home Status. XAI integration successfully opens the black-box nature of the model without sacrificing prediction accuracy, so that every decision can be explained transparently to all financial institution stakeholders.

Keywords: Credit risk, random forest, explainable Ai, shap, machine learning, credit prediction

Abstrak

Pesatnya perkembangan teknologi kecerdasan buatan telah mendorong lembaga keuangan untuk mengadopsi algoritma *machine learning* dalam proses penilaian risiko kredit. Namun, model berperforma tinggi seperti *Random Forest* cenderung bersifat *black-box*, sehingga sulit dipahami oleh manajemen maupun regulator dan menghambat proses audit serta pengambilan keputusan strategis. Penelitian ini bertujuan untuk menganalisis penerapan *Explainable Artificial Intelligence* (XAI) melalui metode *Shapley Additive Explanations* (SHAP) pada model *Random Forest* guna menghasilkan sistem prediksi risiko kredit yang tidak hanya akurat, tetapi juga transparan dan dapat dipertanggungjawabkan. Penelitian menggunakan *dataset* historis kredit publik dengan 32.581 observasi dan 10 variabel masukan, yang diproses melalui tahapan *preprocessing*, pelatihan model, evaluasi, dan analisis SHAP, serta dikemas dalam platform web interaktif berbasis Streamlit. Hasil penelitian menunjukkan model *Random Forest* mencapai Akurasi 80,91%, *Recall* 82,85%, dan AUC-ROC 0,8722. Analisis SHAP mengkonfirmasi bahwa Suku Bunga merupakan fitur paling berpengaruh terhadap risiko gagal bayar, diikuti Tujuan Pinjaman dan Status Rumah. Integrasi XAI berhasil membuka sifat *black-box* model tanpa mengorbankan akurasi prediksi, sehingga setiap keputusan dapat dijelaskan secara transparan kepada seluruh pemangku kepentingan lembaga keuangan.

Kata kunci: Risiko kredit, *random forest, explainable Ai, shap, machine learning,* prediksi kredit

1. Pendahuluan

Sektor perbankan dan lembaga keuangan mengandalkan penyaluran kredit sebagai penggerak ekonomi sekaligus sumber pendapatan utama. Namun, aktivitas ini membawa risiko inheren berupa risiko kredit (potensi debitur gagal bayar), yang diukur melalui rasio *Non-Performing Loan* (NPL) dan dapat mengancam stabilitas keuangan jika tidak dikelola dengan baik. Untuk mengatasi kompleksitas tersebut, industri keuangan mulai beralih dari metode konvensional ke *machine learning* (ML), khususnya algoritma *Random Forest* (RF) yang terbukti akurat, stabil, dan mampu menangani hubungan data *nonlinier* yang rumit guna memprediksi kelayakan kredit [1], [2] .

Meski demikian, sifat *black-box* pada *Random Forest* membuat logika pengambilan keputusannya sulit dipahami, sehingga memicu ketidakpercayaan regulator, mempersulit audit, dan menghambat adopsi AI di sektor keuangan yang ketat regulasi [3]. Sebagai solusinya, pendekatan *Explainable Artificial Intelligence* (XAI) melalui metode SHAP *(Shapley Additive Explanations*) hadir untuk memberikan transparansi kuantitatif atas setiap prediksi model [4].

Perkembangan terkini (*state of the art*) menunjukkan bahwa kajian prediksi risiko kredit bergeser dari sekadar mengejar akurasi menuju model yang dapat dijelaskan. Li dan Wu membandingkan sembilan model *machine learning* pada data pinjaman dan menemukan bahwa *Random Forest* memprediksi gagal bayar secara efisien dan stabil, sementara SHAP mampu mengidentifikasi faktor utama penyebab gagal bayar [5]. Nallakaruppan dkk. menerapkan *Decision Tree* dan *Random Forest* yang dijelaskan dengan LIME dan SHAP untuk keputusan kredit pada *dataset* 32.581 observasi dan memperoleh akurasi sekitar 0,93 [6]. Zhu dkk. menunjukkan bahwa LightGBM dan XGBoost lebih unggul daripada regresi logistik dan pohon keputusan, serta LIME dapat menjelaskan faktor seperti jangka waktu, *grade*, dan jumlah pinjaman [7]. Rambauli dkk. memperluas pendekatan tersebut dengan SHAP, LIME, kalibrasi probabilitas, dan kuantifikasi ketidakpastian untuk mendukung transparansi dan kepatuhan regulasi [8].

Kajian terhadap penelitian terdahulu merupakan fondasi penting untuk memetakan perkembangan teknologi terkini dalam manajemen keuangan serta memastikan orisinalitas karya. Dalam konteks prediksi risiko kredit, terdapat sejumlah penelitian relevan yang membahas penggunaan algoritma *machine learning*, mitigasi kredit macet, hingga implementasi sistem pendukung keputusan yang menjadi acuan dalam memperkuat metodologi yang diusulkan

Penelitian pertama dilakukan oleh Maulani, Norisanti, dan Ramdan (2024) yang meneliti pengaruh penyaluran kredit dan kredit macet terhadap profitabilitas perbankan menggunakan metode regresi [9]. Hasil penelitian menunjukkan adanya pengaruh positif signifikan dari penyaluran kredit dan pengaruh negatif dari kredit macet terhadap profitabilitas. Penelitian ini berfokus pada hubungan antar variabel, sedangkan penelitian yang dilakukan saat ini akan menambahkan pendekatan prediktif untuk deteksi risiko sejak dini.

Penelitian kedua oleh Jedho, Herdi, dan Aurelia (2026) menganalisis strategi penanganan kredit macet dengan pendekatan kualitatif deskriptif melalui manajemen risiko [10]. Strategi seperti *rescheduling* dan *restructuring* dinilai cukup efektif dalam menangani kredit yang sudah bermasalah. Perbedaannya, penelitian ini lebih menekankan pada upaya pencegahan melalui prediksi risiko sebelum kredit macet terjadi.

Penelitian ketiga oleh Maulana (2024) menerapkan metode Logika *Fuzzy Sugeno* dalam sistem pendukung keputusan untuk menentukan kelayakan kredit berdasarkan prinsip 4C [11] . Meskipun mampu membantu pengambilan keputusan secara cepat, metode ini masih bersifat *rule-based* sehingga memiliki keterbatasan dalam menangani data yang besar. Penelitian ini akan menggunakan *machine learning* untuk meningkatkan akurasi pada data yang lebih kompleks.

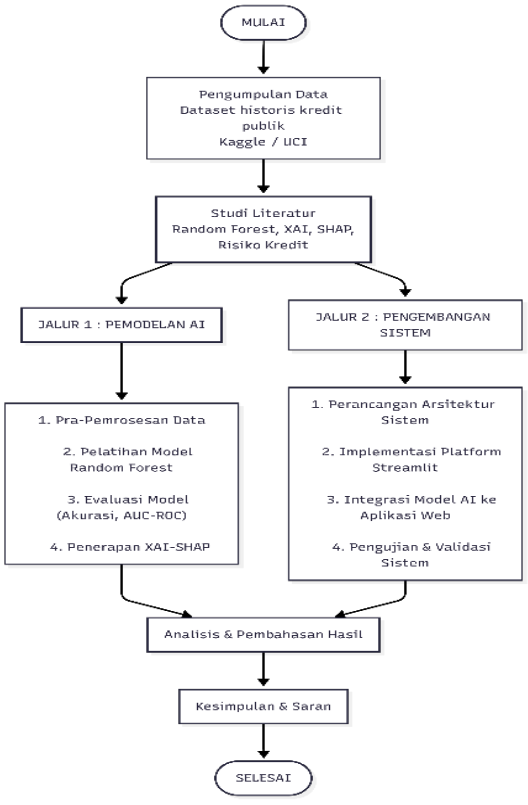
Penelitian keempat oleh Susilo, Sandyawati, dan Hapsari (2026) mengkaji pengaruh *machine learning* dalam mitigasi risiko kredit dan deteksi *fraud* pada perbankan digital [1]. Hasilnya menunjukkan bahwa *machine learning* mampu menurunkan rasio *Non-Performing Loan* (NPL) secara signifikan. Kebaruan dalam penelitian ini adalah pengintegrasian aspek *Explainable AI* (XAI) untuk meningkatkan transparansi hasil prediksi yang belum dibahas dalam studi tersebut.

Berdasarkan kajian tersebut, terdapat celah penelitian (*research gap*) yang belum terjawab. Pertama, sebagian besar studi XAI pada kredit berhenti pada analisis eksperimen di lingkungan pemrograman sehingga hasil SHAP sulit diakses praktisi non-teknis; tinjauan sistematis terbaru juga menyimpulkan bahwa penjelasan umumnya ditambahkan secara *post-hoc* dan belum terintegrasi penuh ke dalam *pipeline* penilaian kredit dari ujung ke ujung [12]. Kedua, penelitian yang menyajikan penjelasan global dan lokal sekaligus dalam satu antarmuka interaktif, lengkap dengan verifikasi perhitungan manual dan pengujian fungsional, masih terbatas. Ketiga, studi pada konteks perbankan dan perkreditan di Indonesia [9], [10], [11] masih memakai pendekatan regresi, kualitatif, atau berbasis aturan (*rule-based*) tanpa memadukan prediksi *machine learning* dengan interpretabilitas.

Untuk menjawab celah tersebut, tujuan penelitian ini adalah: (1) membangun model prediksi risiko kredit berkinerja tinggi menggunakan algoritma *Random Forest* untuk mengidentifikasi potensi gagal bayar debitur secara akurat; (2) mengintegrasikan SHAP (*Shapley Additive Explanations*) guna memetakan faktor risiko dominan secara global maupun menguraikan alasan keputusan kredit tiap individu secara lokal; dan (3) menerjemahkan seluruh kompleksitas teknis tersebut ke dalam platform interaktif berbasis Streamlit yang divalidasi melalui verifikasi perhitungan manual dan pengujian *black box*. Dengan demikian, penelitian ini diharapkan menyediakan instrumen pengambilan keputusan yang transparan dan akuntabel bagi praktisi keuangan non-teknis untuk melakukan mitigasi risiko kredit dan menekan rasio *Non-Performing Loan* (NPL). Pendekatan ini sejalan dengan kajian sebelumnya yang menegaskan pentingnya transparansi dan interpretabilitas model *machine learning* dalam mendukung pengambilan keputusan berbasis data [13], [14], [15].

2. Metoda Penelitian

Metodologi penelitian ini disusun secara terstruktur untuk menjelaskan tahapan eksperimen, mulai dari pengolahan dataset historis hingga penerapan algoritma Random Forest dan metode SHAP, serta pengembangan platform web interaktif. Alur penelitian dirancang untuk memastikan validitas hasil prediksi, transparansi model, dan keterpakean sistem bagi pengguna akhir. Secara keseluruhan, penelitian ini terdiri dari dua jalur pengembangan utama yang berjalan secara paralel dan saling melengkapi: (1) Pemodelan AI dan (2) Pengembangan Sistem, sebagaimana diilustrasikan pada Gambar 1 berikut.



Gambar 1. Alur Penelitian Prediksi Risiko Kredit dengan XAI

2.1. Jenis dan Sumber Data

Penelitian ini menggunakan pendekatan kuantitatif dengan memanfaatkan data sekunder. Data yang digunakan berupa *dataset* historis kredit yang bersumber dari repositori publik seperti Kaggle atau UCI *Machine Learning Repository*. *Dataset* tersebut mencakup variabel terstruktur mengenai karakteristik debitur, seperti usia, tingkat pendapatan, status pekerjaan, jumlah pinjaman, dan riwayat pembayaran, yang memiliki label klasifikasi kelayakan kredit (*good/bad credit*).

2.2. Metode Pengumpulan Data

Pengumpulan data dilakukan melalui teknik studi dokumentasi dengan mengunduh *dataset* yang memenuhi kriteria relevansi atribut dan kecukupan jumlah sampel untuk pelatihan model. Selain itu, dilakukan studi literatur dari berbagai jurnal ilmiah, prosiding, dan referensi akademik untuk membangun landasan teori terkait algoritma *Random Forest* dan metode *Explainable AI* (XAI).

2.3. Tahapan Pengembangan Sistem

Pengembangan sistem dalam penelitian ini dibagi ke dalam dua jalur utama yang berjalan secara paralel: Jalur Pemodelan AI dan Jalur Pengembangan Sistem. Kedua jalur ini saling terintegrasi untuk menghasilkan platform prediksi risiko kredit yang akurat dan transparan.

Jalur 1: Pemodelan AI

Jalur pemodelan AI mencakup seluruh proses ilmiah mulai dari pengolahan data hingga interpretasi model sebagai berikut:

a. Persiapan Data (*Data Preprocessing*)
   Meliputi pembersihan data (*cleaning*), penanganan nilai yang hilang (*missing values*), serta transformasi format data agar siap diproses oleh algoritma.
b. Pembuatan Model Prediksi
   Implementasi algoritma *Random Forest* untuk membangun model klasifikasi risiko kredit. Model dilatih menggunakan sebagian data (*training set*) dan diuji menggunakan sisa data lainnya (*testing set*).
c. Evaluasi Model
   Penilaian kinerja model menggunakan metrik evaluasi standar seperti tingkat akurasi dan matriks konfusi untuk memastikan stabilitas prediksi.
d. Penerapan *Explainable AI* (SHAP)
   Pengintegrasian metode *Shapley Additive Explanation* (SHAP) untuk memberikan interpretasi pasca-pemodelan (*post-hoc interpretability*). Tahap ini bertujuan membedah kontribusi setiap fitur terhadap keputusan prediksi yang dihasilkan.

Jalur 2: Pengembangan Sistem

Jalur pengembangan sistem mencakup rekayasa perangkat lunak untuk mewujudkan model AI menjadi platform yang dapat digunakan secara praktis:

a. Perancangan Arsitektur Sistem
   Desain arsitektur platform web interaktif berbasis Streamlit yang terdiri dari lima modul utama: Ringkasan Data, Eksplorasi Data (EDA), Performa Model, Analisis SHAP, dan Prediksi Interaktif
b. Implementasi Platform
   Pengembangan antarmuka pengguna menggunakan Streamlit dengan mekanisme *caching* (@st.cache_data dan @st.cache_resource) untuk memastikan *responsivitas* aplikasi saat pengguna berinteraksi dengan antarmuka.
c. Integrasi Model AI ke Aplikasi Web
   Penggabungan *pipeline* lengkap mulai dari *preprocessing*, pelatihan model, evaluasi, hingga komputasi SHAP ke dalam satu platform terintegrasi yang dapat diakses oleh pengguna non-teknis.
d. Pengujian dan Validasi Sistem
   Verifikasi fungsionalitas setiap modul aplikasi, pengujian prediksi interaktif dengan data nasabah baru, dan validasi konsistensi hasil SHAP terhadap keputusan model.

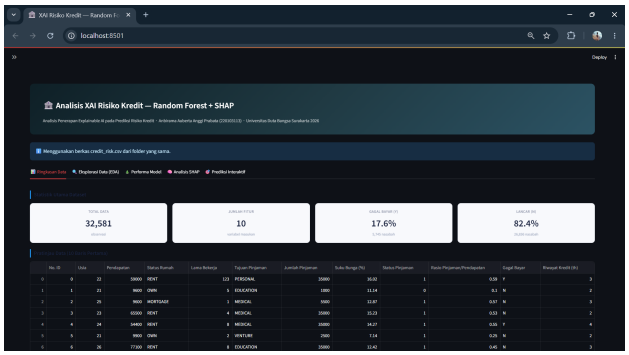
3. Hasil Penelitian

3.1. Hasil Pra-Pemrosesan Data

Dataset yang digunakan dalam penelitian ini merupakan data historis kredit yang

bersumber dari repositori publik, memuat 10 variabel masukan dan 1 variabel target (*Default*). Variabel masukan terdiri atas: Usia, Pendapatan, Status Rumah, Lama Bekerja, Tujuan Pinjaman, Jumlah Pinjaman, Suku Bunga, Status Pinjaman, Rasio Pinjaman/Pendapatan, dan Riwayat Kredit. Proses pra-pemrosesan data dilakukan melalui beberapa tahapan sistematis sebagai berikut.

Pertama, penghapusan kolom ID yang tidak relevan untuk prediksi. Kedua, imputasi nilai kosong (*missing values*) menggunakan nilai tengah (median) untuk kolom numerik dan modus untuk kolom kategorik. Ketiga, penghapusan *outlier* ekstrem pada kolom Lama Bekerja menggunakan batas persentil ke-99. Keempat, pengkodean variabel kategorik (Status Rumah dan Tujuan Pinjaman) menggunakan Label *Encoding*. Kelima, konversi variabel target menjadi biner: Gagal Bayar (Y) = 1 dan Lancar (N) = 0. Pembagian data dilakukan secara stratifikasi dengan proporsi 80% data latih dan 20% data uji. Untuk mengatasi ketidakseimbangan kelas, digunakan metode SMOTE (*Synthetic Minority Oversampling Technique*) pada data latih sehingga distribusi kelas menjadi seimbang sebelum proses pelatihan model. Penerapan SMOTE pada data latih sejalan dengan Alamsyah dkk. yang melaporkan peningkatan presisi, *recall*, dan F1-score pada data kredit yang tidak seimbang[16].



Gambar 2. Tampilan Tab Ringkasan Data Aplikasi XAI Risiko Kredit

Dari Gambar 2 dapat dilihat bahwa Tab Ringkasan Data menampilkan empat kartu metrik utama, yaitu total data sebanyak 32.581 observasi, jumlah fitur sebanyak 10 variabel masukan, persentase nasabah gagal bayar (Y) sebesar 17,6% (5.745 nasabah), dan persentase nasabah lancar (N) sebesar 82,4% (26.836 nasabah). Di bawahnya tersaji pratinjau 10 baris pertama *dataset* beserta seluruh kolom variabel, sehingga pengguna dapat memverifikasi kualitas dan struktur data sebelum melanjutkan analisis.

3.2. Hasil Pelatihan dan Konfigurasi Model

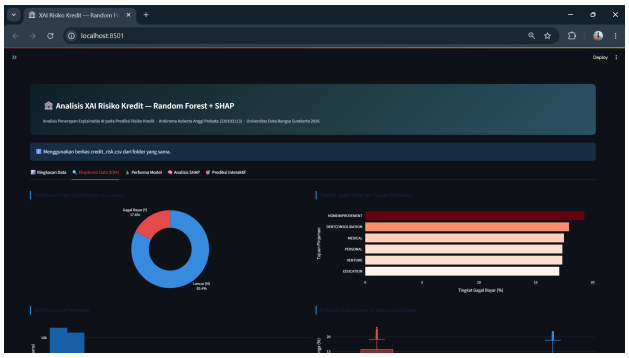
Model *Random Forest* dibangun menggunakan pustaka *scikit-learn* dengan konfigurasi parameter sebagaimana tercantum pada Tabel 1. Parameter utama yang digunakan meliputi jumlah pohon keputusan (*n_estimators*), kedalaman maksimal pohon (*max_depth*), dan pemilihan fitur menggunakan akar kuadrat jumlah fitur (*max_features='sqrt'*). Penggunaan *class_weight='balanced'* memastikan model tidak bias ke kelas mayoritas meskipun SMOTE telah diterapkan. Selain itu, nilai *random_state*=42 digunakan untuk memastikan reproduksibilitas hasil eksperimen.

Tabel 1. Konfigurasi Parameter Model *Random Forest*

Parameter	Nilai
n_estimators	200
max_depth	10
min_samples_split	5
min_samples_leaf	2
max_features	sqrt
class_weight	balanced
random_state	42
SMOTE	Aktif

Dari Tabel 1 dapat dilihat bahwa konfigurasi model dirancang untuk mengoptimalkan kemampuan generalisasi dengan menjaga keseimbangan antara kompleksitas pohon keputusan dan

kemampuan menghindari *overfitting*. Parameter tersebut dapat dikustomisasi oleh pengguna melalui antarmuka *sidebar* aplikasi Streamlit.



Gambar 3. Tampilan Tab Eksplorasi Data (EDA) Aplikasi XAI Risiko Kredit

Dari Gambar 3 dapat dilihat bahwa Tab Eksplorasi Data (EDA) menyajikan empat visualisasi sekaligus. Grafik donat (kiri atas) mengkonfirmasi ketidakseimbangan kelas dengan komposisi 17,6% gagal bayar dan 82,4% lancar. Diagram batang horizontal (kanan atas) menunjukkan bahwa tujuan pinjaman HOMEIMPROVEMENT memiliki tingkat gagal bayar tertinggi sekitar 19%, diikuti DEBTCONSOLIDATION dan MEDICAL. Histogram distribusi usia (kiri bawah) memperlihatkan mayoritas peminjam berusia 20-30 tahun. Diagram kotak (kanan bawah) membuktikan bahwa nasabah gagal bayar rata-rata menanggung suku bunga lebih tinggi dibandingkan nasabah lancar, yang mengindikasikan hubungan positif antara besaran suku bunga dan risiko gagal bayar.

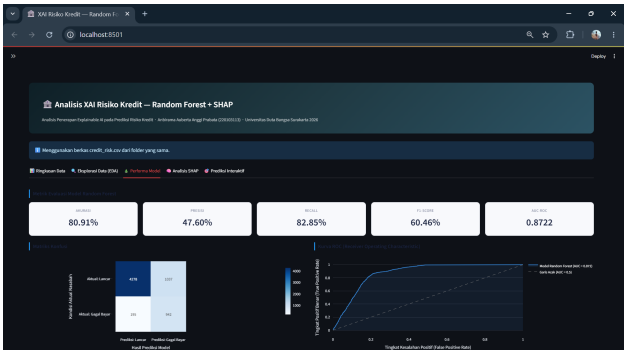
3.3. Pengujian Performa Model

Evaluasi model dilakukan menggunakan lima metrik standar klasifikasi biner yang dihitung pada data uji (20% dari total *dataset*). Sistem mengimplementasikan perhitungan Akurasi, Presisi, *Recall*, F1-Score, dan AUC-ROC sebagaimana ditampilkan pada Tabel 2.

Tabel 2. Hasil Evaluasi Metrik Model *Random Forest*

Metrik Evaluasi	Nilai
Akurasi	80,91%
Presisi	47,60%
Recall	82,85%
F1-Score	60,46%
AUC-ROC	0,8722

Dari Tabel 2 dapat dilihat bahwa nilai metrik evaluasi yang dihasilkan mencerminkan performa model yang komprehensif. Akurasi menunjukkan proporsi prediksi yang benar secara keseluruhan. Presisi mengukur ketepatan prediksi positif (nasabah yang diprediksi gagal bayar namun sebenarnya gagal bayar). *Recall* mengukur kemampuan model mendeteksi seluruh kasus gagal bayar yang sebenarnya, sedangkan F1-Score merupakan rata-rata harmonik antara presisi dan *recall*. Nilai AUC-ROC mendekati 1,0 menandakan kemampuan diskriminasi model yang sangat baik dalam membedakan nasabah yang akan gagal bayar dengan yang tidak. Selain itu, matriks konfusi yang ditampilkan pada antarmuka aplikasi memvisualisasikan sebaran prediksi benar dan salah pada setiap kelas, sehingga memudahkan interpretasi hasil model secara lebih rinci.



Gambar 4. Tampilan Tab Performa Model *Random Forest*

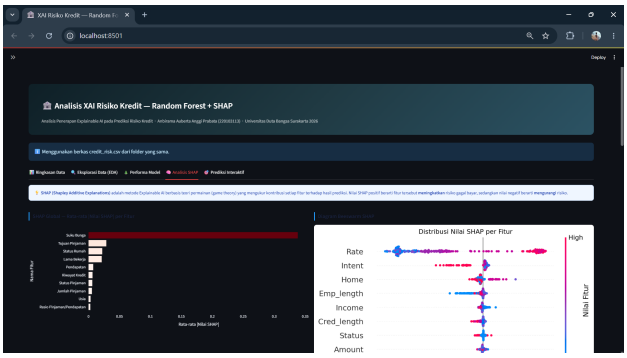
Dari Gambar 4 dapat dilihat bahwa model *Random Forest* menghasilkan Akurasi 80,91%, Presisi 47,60%, *Recall* 82,85%, F1-Score 60,46%, dan AUC-ROC 0,8722. Nilai *recall* yang tinggi (82,85%) menunjukkan kemampuan model yang baik dalam mendeteksi nasabah yang benar-benar akan gagal bayar. Matriks konfusi memperlihatkan

bahwa dari 1.137 kasus gagal bayar aktual, model berhasil mengidentifikasi 942 kasus dengan benar (942 *True Positive*) dan hanya melewatkan 195 kasus (*False Negative*). Kurva ROC dengan AUC = 0,872 mengkonfirmasi kemampuan diskriminasi model yang sangat baik, jauh di atas garis acak (AUC = 0,5). Diagram kepentingan fitur di bagian bawah menunjukkan Suku Bunga sebagai fitur paling dominan dalam keputusan model. Pola *recall* tinggi dengan presisi sedang ini juga dilaporkan oleh Mapfumo dan Shongwe, yang memperoleh *recall* 0,745 dan presisi 0,512 pada *Random Forest* dengan teknik resampling pada *dataset* kredit Taiwan [17]. Kinerja *Random Forest* yang stabil dalam memprediksi gagal bayar juga selaras dengan temuan Li dan Wu [5].

3.4. Analisis SHAP Global

Analisis SHAP global dilakukan menggunakan *TreeExplainer* yang dioptimalkan khusus untuk model berbasis pohon keputusan seperti *Random Forest*. Nilai SHAP dihitung terhadap 300 sampel acak dari data uji menggunakan *random_state*=42 untuk menjamin konsistensi. Terdapat dua visualisasi utama yang disajikan dalam tab Analisis SHAP pada antarmuka aplikasi.

Visualisasi pertama adalah diagram batang rata-rata nilai SHAP per fitur, yang menunjukkan tingkat pengaruh global setiap fitur terhadap keputusan model secara agregat. Fitur dengan rata-rata nilai SHAP absolut tertinggi merupakan variabel paling berpengaruh secara global. Visualisasi kedua adalah diagram *beeswarm* SHAP yang menampilkan distribusi nilai SHAP seluruh sampel untuk setiap fitur. Titik berwarna merah menunjukkan nilai fitur yang tinggi, sedangkan biru menunjukkan nilai fitur yang rendah. Titik yang berada di sisi kanan garis nol berarti fitur tersebut meningkatkan probabilitas gagal bayar, dan sebaliknya. Kombinasi kedua visualisasi ini memberikan pemahaman komprehensif mengenai perilaku model secara keseluruhan.



Gambar 5. Tampilan Tab Analisis SHAP - Nilai Global dan Diagram *Beeswarm*

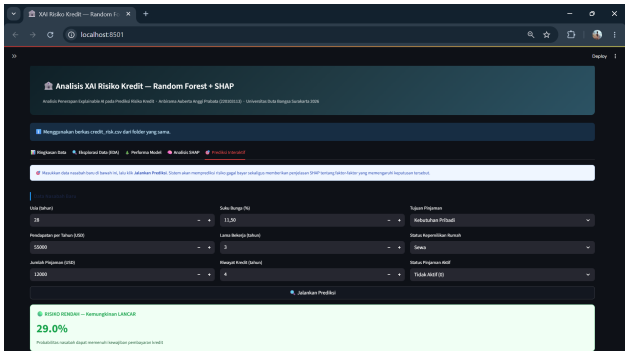
Dari Gambar 5 dapat dilihat bahwa analisis SHAP global mengkonfirmasi Suku Bunga (Rate) sebagai fitur paling berpengaruh secara signifikan dengan rata-rata nilai SHAP mencapai lebih dari 0,35, jauh melampaui fitur lainnya. Tujuan Pinjaman, Status Rumah, dan Lama Bekerja menempati posisi berikutnya. Diagram *beeswarm* di sisi kanan memperkaya interpretasi dengan memperlihatkan arah pengaruh: fitur Suku Bunga dengan nilai tinggi (titik merah) konsisten berada di sisi kanan garis nol, artinya suku bunga tinggi secara langsung meningkatkan probabilitas gagal bayar. Temuan ini selaras secara domain karena beban bunga yang besar mengurangi kemampuan nasabah memenuhi kewajiban cicilan. Temuan ini konsisten dengan Zhu dkk. yang melalui LIME menemukan bahwa jangka waktu, *grade*, dan jumlah pinjaman memengaruhi prediksi gagal bayar [7], di mana *grade* pada praktiknya umumnya berkaitan erat dengan penetapan suku bunga. Li dan Wu juga melaporkan bahwa SHAP dapat mengidentifikasi faktor utama gagal bayar, termasuk lama pengalaman kerja peminjam[5], yang dalam penelitian ini juga muncul pada empat fitur teratas. Nallakaruppan dkk., yang menggunakan *dataset* berukuran sama (32.581 observasi), turut menunjukkan bahwa SHAP dan LIME dapat membuka model *Random Forest* pada data kredit[[6], sedangkan Rambauli dkk. menemukan bahwa SHAP dan LIME mengidentifikasi riwayat tunggakan, catatan negatif kredit, dan rasio utang terhadap pendapatan sebagai penggerak utama risiko, sejalan dengan penalaran risiko keuangan [8].

3.5. Verifikasi Perhitungan Manual

Verifikasi manual mencakup: (1) penghitungan *Gini Impurity* diperoleh nilai 0,42 pada simpul akar dengan Suku Bunga sebagai fitur pemisahan optimal (*Information Gain* = 0,26); (2) *Majority Voting* dengan 200 pohon, 142 memilih kelas 1 sehingga P(Gagal Bayar) = 0,71; (3) verifikasi metrik manual: Akurasi = (942+4.342)/6.517 = 80,91%, Presisi = 942/1.980 = 47,60%, *Recall* = 942/1.137 = 82,85%, F1 = 60,46% seluruhnya bersesuaian dengan keluaran sistem.

3.6. Hasil Pengembangan Sistem

Platform berbasis Streamlit berhasil mengintegrasikan lima modul: (1) Ringkasan Data, (2) Eksplorasi Data (EDA), (3) Performa Model, (4) Analisis SHAP, dan (5) Prediksi Interaktif. Mekanisme caching (@st.cache_data dan @st.cache_resource) memastikan responsivitas aplikasi.



Gambar 6. Tampilan Tab Prediksi Interaktif - *Form Input* dan Hasil Prediksi Nasabah

3.7. Perbandingan Hasil Pemodelan AI dengan Implementasi Sistem

Tabel 3. Perbandingan Hasil Pemodelan AI vs. Implementasi Sistem

Metrik	Pemodelan AI	Sistem	Status
Akurasi	80,91%	80,91%	✓ Sesuai
Presisi	47,60%	47,60%	✓ Sesuai
Recall	82,85%	82,85%	✓ Sesuai
F1-Score	60,46%	60,46%	✓ Sesuai
AUC-ROC	0,8722	0,8722	✓ Sesuai
Fitur SHAP	Suku Bunga (>0,35)	Suku Bunga (>0,35)	✓ Sesuai

3.8. Pengujian Black Box

Tabel 4. Hasil Pengujian Black Box Platform Streamlit

No	Skenario Uji	Masukan	Keluaran Diharapkan	Status
1	Tab Ringkasan Data	Buka aplikasi	4 kartu metrik & pratinjau	✓ Berhasil
2	Tab EDA	Navigasi Tab EDA	4 grafik distribusi	✓ Berhasil
3	Tab Performa Model	Navigasi Tab Performa	Akurasi 80,91%, AUC 0,8722	✓ Berhasil
4	Tab Analisis SHAP	Navigasi Tab SHAP	Beeswarm & bar SHAP	✓ Berhasil
5	Prediksi Interaktif	Data nasabah risiko rendah	RISIKO RENDAH, prob. 29,0%	✓ Berhasil

3.9. Pembahasan Integrasi XAI

Integrasi *Random Forest* dengan metode SHAP memberikan nilai tambah dibandingkan model *black-box* konvensional karena kontribusi setiap fitur terhadap prediksi dapat dikuantifikasi dan ditelusuri[5], [6]. Setiap keputusan prediksi dapat diaudit dan dijelaskan kepada pemangku kepentingan, termasuk manajemen bank, regulator, maupun nasabah [3]. Hal ini sejalan dengan tinjauan sistematis Bahlool dkk. yang menyimpulkan bahwa

tekanan regulasi mendorong meningkatnya penggunaan penjelasan *post-hoc* seperti SHAP dan LIME dalam penilaian kredit[12]. Nilai Presisi yang relatif rendah (47,60%) mengindikasikan masih terdapat tingkat *false positive* yang cukup tinggi, berpotensi mengurangi kepuasan nasabah dan meningkatkan beban kerja analis kredit dalam melakukan verifikasi manual. Rambauli dkk. juga mengingatkan bahwa *Random Forest* yang akurat dapat kurang terkalibrasi (cenderung terlalu percaya diri), sehingga akurasi saja tidak cukup dan perlu dilengkapi kalibrasi serta penyetelan ambang batas klasifikasi[8].

4. Kesimpulan

Penelitian ini berhasil mengimplementasikan model prediksi risiko kredit berbasis *Random Forest* yang diintegrasikan dengan metode *Shapley Additive Explanations* (SHAP) sebagai pendekatan *Explainable AI* (XAI) ke dalam platform web interaktif berbasis Streamlit. Model *Random Forest* yang dibangun menghasilkan performa yang konsisten dengan Akurasi 80,91%, *Recall* 82,85%, dan AUC-ROC 0,8722, di mana seluruh nilai tersebut telah diverifikasi secara manual melalui matriks konfusi.

Verifikasi manual juga mengonfirmasi bahwa mekanisme *Gini Impurity* dan *majority* voting berjalan sesuai teori, serta membuktikan bahwa pemilihan Suku Bunga sebagai fitur pemisahan utama selaras dengan temuan SHAP global yang menempatkan Suku Bunga sebagai fitur paling dominan (rata-rata SHAP > 0,35), diikuti oleh Tujuan Pinjaman, Status Rumah, dan Lama Bekerja. Dari sisi pengembangan sistem, seluruh lima skenario pengujian *black box* dinyatakan lulus 100%, dan integrasi *pipeline* berjalan dengan benar yang ditunjukkan oleh kesesuaian penuh antara keluaran pemodelan dengan tampilan sistem. Dengan demikian, integrasi XAI ini terbukti berhasil membuka sifat *black-box* model tanpa mengorbankan akurasi prediksi, sekaligus menjawab permasalahan utama dalam penelitian ini.

5. Saran

Meskipun hasil yang dicapai cukup menjanjikan, terdapat beberapa kelemahan yang dapat dijadikan landasan bagi saran pengembangan ke depan. Pertama, untuk mengatasi rendahnya nilai Presisi (47,60%) yang menyebabkan tingginya tingkat *false positive*, penelitian selanjutnya disarankan melakukan optimasi ambang batas klasifikasi (*threshold tuning*) berdasarkan analisis biaya kesalahan prediksi serta membandingkan performa model dengan algoritma ansambel lain seperti *XGBoost* atau *LightGBM* guna meningkatkan keseimbangan antara Presisi dan *Recall*.

Kedua, mengingat *dataset* yang digunakan masih berasal dari repositori publik dan belum merepresentasikan karakteristik domestik, penelitian berikutnya perlu menggunakan data historis kredit dari perbankan Indonesia yang mencakup variabel kontekstual lokal seperti skor BI *Checking* dan profil risiko regional agar lebih relevan untuk diadopsi. Ketiga, untuk mengatasi sifat model yang statis, diperlukan pengembangan mekanisme pembaruan model secara berkala (*online learning*) agar sistem tetap adaptif terhadap perubahan perilaku nasabah dari waktu ke waktu. Keempat, visualisasi grafis SHAP yang saat ini masih memerlukan literasi data sebaiknya dikembangkan lebih lanjut dengan menambahkan fitur narasi penjelasan otomatis berbasis natural *language generation* agar lebih intuitif bagi pengguna non-teknis. Terakhir, disarankan untuk melakukan studi komparatif antara metode SHAP dengan pendekatan XAI lainnya, seperti LIME (*Local Interpretable Model-agnostic Explanations*) atau *Integrated Gradients*, guna mengidentifikasi metode *eksplanasi* yang paling tepat dalam konteks risiko kredit di Indonesia.

6. Daftar Pustaka

[1] D. E. Susilo, N. S. Sandyawati, F. A. Hapsari, S. Tinggi, and I. A. Bayuangga, “Pengaruh Machine Learning dalam Mitigasi Risiko Kredit, Deteksi Fraud dan Mitigasi Kerugian Finansial Perbankan Digital,” 2026.

[2] D. Hartanti, A. I. Pradana, and S. Lestari, "Komprasi Algoritma Decision Tree, SVM dan ANN untuk Reservasi Hotel," 2023.

[3] V. Chamola, V. Hassija, A. R. Sulthana, D. Ghosh, D. Dhingra, and B. Sikdar, "A Review of Trustworthy and Explainable Artificial Intelligence (XAI)," *IEEE Access*, vol. 11, pp. 78994–79015, 2023, doi: 10.1109/ACCESS.2023.3294569.

[4] T. Hulsen, "Explainable Artificial Intelligence (XAI): Concepts and Challenges in Healthcare," Sep. 01, 2023, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/ai4030034.

[5] L. H. Li, A. K. Sharma, and S. T. Cheng, "Explainable AI based LightGBM prediction model to predict default borrower in social lending platform," *Intelligent Systems with Applications*, vol. 26, 2025, doi: 10.1016/j.iswa.2025.200514.

[6] M. K. Nallakaruppan *et al.*, "Credit Risk Assessment and Financial Decision Support Using Explainable Artificial Intelligence," *Risks*, vol. 12, no. 10, 2024, doi: 10.3390/risks12100164.

[7] X. Zhu, Q. Chu, X. Song, P. Hu, and L. Peng, "Explainable prediction of loan default based on machine learning models," *Data Science and Management*, vol. 6, no. 3, pp. 123–133, 2023, doi: 10.1016/j.dsm.2023.04.003.

[8] M. Rambauli, T. Ravele, and C. Sigauke, "Application of Explainable AI and Uncertainty Quantification in Credit Risk Assessment," *Risks*, vol. 14, no. 7, 2026, doi: 10.3390/risks14070152.

[9] A. Maulani, Nor Norisanti, and Asep M. Ramdan, "Analisis Penyaluran Kredit, dan Kredit Macet terhadap Profitabilitas: Studi Keuangan Pada Sektor Perbankan yang Terdaftar di Bursa Efek Indonesia Periode 2023," *Al-Kharaj: Jurnal Ekonomi, Keuangan & Bisnis Syariah*, vol. 6, no. 11, Nov. 2024, doi: 10.47467/alkharaj.v6i11.3596.

[10] Y. E. Jedho *et al.*, "Analisis Strategi Penanganan Kredit Macet Berdasarkan Konsep Manajemen Risiko Kredit Pada KSP Kopdit Tuke Jung Universitas Nusa Nipa, Indonesia," 2026, doi: 10.59603/projemen.v12i1,1238.

[11] A. Maulana, "Penerapan Logika Fuzzy Sugeno Untuk Keputusan Kelayakan Kredit Bank," 2024.

[12] R. Bahlool, N. Hewahi, and W. Elmedany, "Performance, Fairness, and Explainability in AI-Based Credit Scoring: A Systematic Literature Review," *Journal of Risk and Financial Management*, vol. 19, no. 2, 2026, doi: 10.3390/jrfm19020104.

[13] S. Abdulla *et al.*, "Journal of Business and Management Studies Credit Risk Prediction Using Explainable AI," 2024, doi: 10.32996/jbms.

[14] F. Suryani and Y. M. Imani, "Pengembangan Dashboard Inventory dan Dashboard Penjualan Berbasis Web untuk," 2025.

[15] H. Li and W. Wu, "Loan default predictability with explainable machine learning," *Financ. Res. Lett.*, vol. 60, 2024, doi: 10.1016/j.frl.2023.104867.

[16] N. Alamsyah, Budiman, T. P. Yoga, and R. Y. R. Alamsyah, "A stacking ensemble model with SMOTE for improved imbalanced classification on credit data," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 22, no. 3, pp. 657–664, 2024, doi: 10.12928/TELKOMNIKA.v22i3.25921.

[17] I. Mapfumo and T. Shongwe, "Performance Evaluation of Machine Learning and Deep Learning Models for Credit Risk Prediction," *Journal of Risk and Financial Management*, vol. 19, no. 3, 2026, doi: 10.3390/jrfm19030210.