

**ANALISIS KINERJA ALGORITMA SUPPORT VECTOR
MACHINE, RANDOM FOREST, DAN XGBOOST DALAM
KLASIFIKASI RISIKO KREDIT NASABAH**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

HOSSAIN WISMAYA RAYHAN

22.11.4911

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**ANALISIS KINERJA ALGORITMA SUPPORT VECTOR
MACHINE, RANDOM FOREST, DAN XGBOOST DALAM
KLASIFIKASI RISIKO KREDIT NASABAH**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

HOSSAIN WISMAYA RAYHAN

22.11.4911

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS KINERJA ALGORITMA SUPPORT VECTOR MACHINE,
RANDOM FOREST, DAN XGBOOST DALAM KLASIFIKASI RISIKO
KREDIT NASABAH**

yang disusun dan diajukan oleh

Hossain Wismaya Rayhan

22.11.4911

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 26 Februari 2026

Dosen Pembimbing,



Rumini, S.Kom., M.Kom.
NIK. 190302246

HALAMAN PENGESAHAN
SKRIPSI
**ANALISIS KINERJA ALGORITMA SUPPORT VECTOR MACHINE,
RANDOM FOREST, DAN XGBOOST DALAM KLASIFIKASI RISIKO
KREDIT NASABAH**

yang disusun dan diajukan oleh

Hossain Wismaya Rayhan

22.11.4911

Telah dipertahankan di depan Dewan Penguji
pada tanggal 26 Februari 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Dina Maulina, S.Kom., M.Kom.
NIK. 190302250

Toto Indriyatmoko, M.Kom.
NIK. 190302407

Rumini, S.kom., M.Kom.
NIK. 190302246

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 26 Februari 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Hossain Wismaya Rayhan
NIM : 22.11.4911

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Kinerja Algoritma Support Vector Machine, Random Forest, dan XGBoost Dalam Klasifikasi Risiko Kredit Nasabah

Dosen Pembimbing : Rumini, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan **gagasan**, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 26 Februari 2026

Yang Menyatakan,



The block contains a handwritten signature in black ink over a red official stamp. The stamp is rectangular with a double border. Inside, there is a circular emblem of the Indonesian Republic (Garuda Pancasila) on the left and the text 'METRA TEMPEL' in the center. Below the emblem, the number '29AMX361318075' is printed.

Hossain Wismaya Rayhan

HALAMAN PERSEMBAHAN

Puji Syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat, kasih dan karunia-Nya sehingga penulis dapat menyelesaikan penyusunan skripsi ini sebagai salah satu syarat untuk memperoleh gelar Sarjana pada program studi yang penulis tempuh. Penyusunan skripsi ini tidak terlepas dari dukungan, bimbingan, bantuan, dan doa dari berbagai pihak. Oleh karena itu, dalam kesempatan ini penulis ingin menyampaikan rasa terima kasih dan penghargaan setinggi-tingginya kepada:

1. Bapak Suryosumirat Ekopriyanto dan Ibu Eni Ristyani, yang senantiasa menjadi sumber kekuatan utama bagi penulis. Terima kasih atas doa yang tidak pernah putus, kasih sayang, serta dukungan moril maupun materiil yang tak terhingga nilainya.
2. Saudara-sudara tersayang, kakak Ikhsandera Bagus Ramadhan dan Adik Gaddhy Adhyastha Farhan, yang telah menjadi sumber semangat selama proses pengerjaan skripsi ini.
3. Sahabat-sahabat dan Teman-teman seperjuangan yang telah menjadi bagian dari proses panjang ini dengan dukungan dan kebersamaan.
4. Almameter tercinta, sebagai tempat menimba ilmu dan membentuk diri menjadi lebih baik

Penulis juga mengucapkan terima kasih kepada semua pihak yang tidak dapat disebutkan satu per satu yang telah memberikan bantuan secara langsung maupun tidak langsung selama proses penyusunan skripsi ini.

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kepada ke hadirat Tuhan Yang Maha Esa atas kasih, rahmat serta Karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul "ANALISIS KINERJA ALGORITMA SUPPORT VECTOR MACHINE, RANDOM FOREST, DAN XGBOOST DALAM KLASIFIKASI RESIKO KREDIT NASABAH" dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan program studi informatika di Universitas Amikom Yogyakarta.

Proses penyusunan skripsi ini tidak akan berjalan lancar tanpa adanya bimbingan, dukungan, dan doa dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang tulus kepada:

1. Bapak Prof. Dr. M. Suyanto, MM., selaku Rektor Universitas AMIKOM Yogyakarta.
2. Prof. Dr. Kusri, M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Eli Pujastuti, M.Kom., selaku Ketua Program Studi Informatika Universitas Amikom Yogyakarta.
4. Ibu Rumini, S.Kom., M.Kom., sebagai Dosen Pembimbing Skripsi yang telah menyediakan waktu, tenaga, dan pikiran untuk memberikan bimbingan kepada penulis dalam penyusunan skripsi ini. Terima kasih atas ilmu pengetahuan dan pengalaman yang telah diberikan.
5. Seluruh dosen dan tenaga kependidikan di Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta, yang telah memberikan ilmu, bimbingan, serta dukungan selama masa studi penulis menjadi bagian penting dalam perjalanan akademik penulis hingga terselesaikannya skripsi ini

Semoga segala bentuk bantuan, dukungan, dan doa yang telah diberikan mendapat balasan kebaikan dari Tuhan Yang Maha Esa. Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna. Oleh karena itu, segala bentuk kritik dan saran yang membangun akan penulis terima dengan lapang dada demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat dan kontribusi bagi perkembangan ilmu pengetahuan.

Yogyakarta, 26 Februari 2026

Penulis



DAFTAR ISI

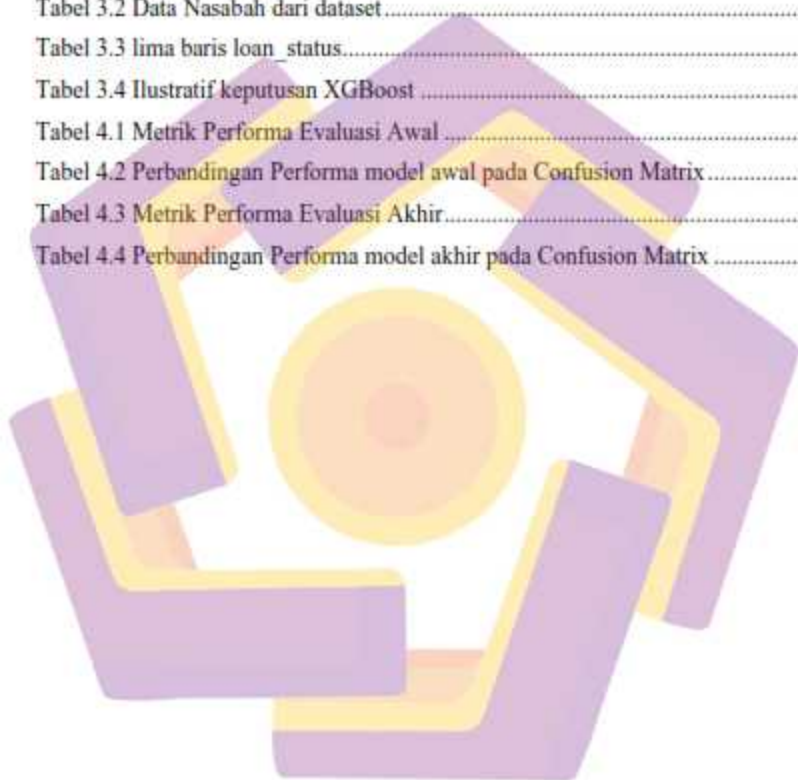
HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN.....	xiv
DAFTAR LAMBANG DAN SINGKATAN	xv
DAFTAR ISTILAH	xvi
INTISARI	xviii
<i>ABSTRACT</i>	xix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5

2.2	Dasar Teori.....	14
2.2.1	Risiko Kredit.....	14
2.2.2	Machine Learning.....	14
2.2.3	Klasifikasi.....	14
2.2.4	<i>Support Vector Machine (SVM)</i>	14
2.2.5	Random Forest.....	16
2.2.6	XGBoost.....	18
2.2.7	Exploratory Data Analysis (EDA).....	19
2.2.8	Feature Selection.....	19
2.2.9	<i>Synthetic Minority Over-Sampling Technique (SMOTE)</i>	20
2.2.10	Evaluasi.....	20
BAB III METODE PENELITIAN.....		24
3.1	Objek Penelitian.....	24
3.2	Alur Penelitian.....	24
3.2.1	Pengambilan Data.....	26
3.2.2	Eksplorasi Data (EDA).....	27
3.2.3	Pre-Processing Data.....	27
3.2.4	Modeling dan Evaluasi awal.....	29
3.2.5	Evaluasi akhir dan Penyimpanan Model.....	33
3.3	Alat dan Bahan.....	34
BAB IV HASIL DAN PEMBAHASAN.....		36
4.1	Pengambilan Data dan persiapan.....	36
4.2	Eksplorasi Data (EDA).....	36
4.3	Pre-Processing Data.....	44
4.3.1	Seleksi Fitur.....	44

4.3.2 Cleaning Data.....	45
4.3.3 Encoding	47
4.4 Splitting Data	48
4.5 features Scaling dan Imbalance Data (SMOTE).....	49
4.6 Evaluasi Awal (Baseline).....	50
4.7 Hypermeter Tuning & Evaluasi Akhir.....	55
4.8 Deployment.....	65
4.8.1 Tampilan dan Pengujian web.....	65
BAB V PENUTUP	71
5.1 Kesimpulan	71
5.2 Saran	72
REFERENSI	74
LAMPIRAN.....	77

DAFTAR TABEL

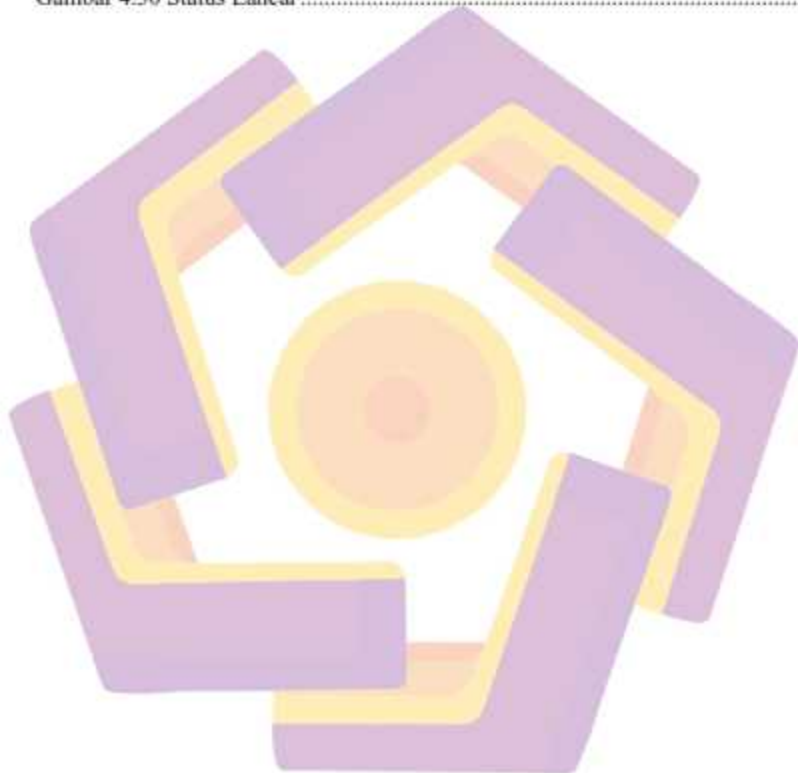
Tabel 2.1 Keaslian Penelitian	7
Tabel 2.2 Confusion Matrix	21
Tabel 3.1 Deskripsi Data Risiko Kredit	26
Tabel 3.2 Data Nasabah dari dataset	29
Tabel 3.3 lima baris loan_status	31
Tabel 3.4 Ilustratif keputusan XGBoost	32
Tabel 4.1 Metrik Performa Evaluasi Awal	51
Tabel 4.2 Perbandingan Performa model awal pada Confusion Matrix	53
Tabel 4.3 Metrik Performa Evaluasi Akhir	57
Tabel 4.4 Perbandingan Performa model akhir pada Confusion Matrix	61



DAFTAR GAMBAR

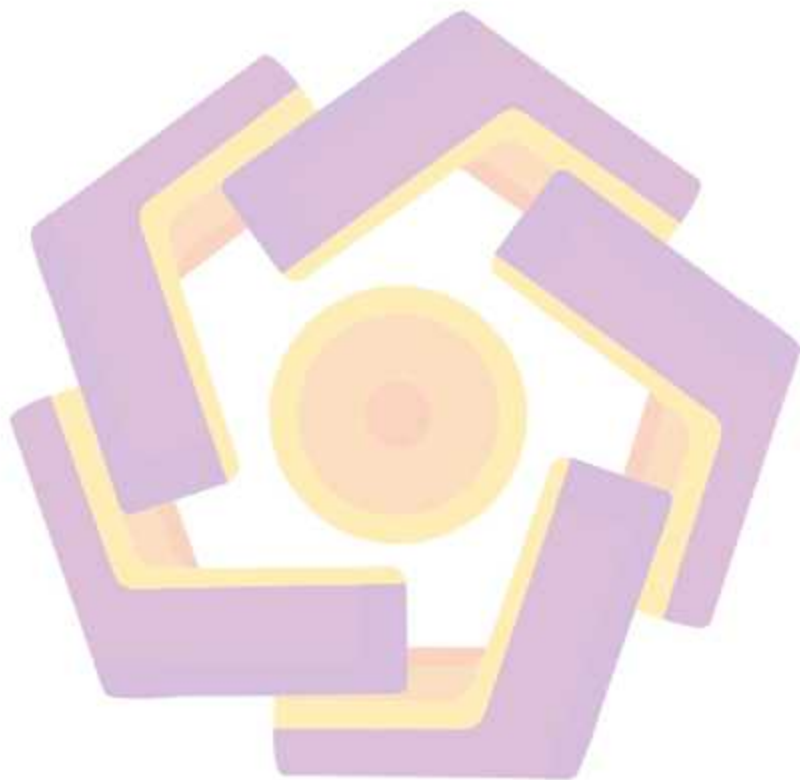
Gambar 2.1 Support Vector Machine	15
Gambar 2.2 Random Forest	17
Gambar 2.3 Ilustrasi Line.....	17
Gambar 2.4 Skema algoritma XGBoost	18
Gambar 3.1 Alur Penelitian	25
Gambar 4.1 Deskripsi dataset	36
Gambar 4.2 Visualisasi Distribusi kolom numerik – Histogram 1	37
Gambar 4.3 Visualisasi Distribusi kolom numerik – Histogram 2	38
Gambar 4.4 Visualisasi Distribusi Kolom Kategorikal 1	39
Gambar 4.5 Visualisasi Distribusi Kolom Kategorikal 2	40
Gambar 4.6 Visualisasi Variabel Target	41
Gambar 4.7 Boxplot Variabel Numerik	42
Gambar 4.8 Heatmap Korelasi Antar Fitur	43
Gambar 4.9 Info Dataset	44
Gambar 4.10 Deskriptif Statistik Dataset	45
Gambar 4.11 Sebelum Handling Missing Values	45
Gambar 4.12 Sesudah Handling Missing Values	46
Gambar 4.13 Penghapusan Outlier	46
Gambar 4.14 Penghapusan Data Duplikat	47
Gambar 4.15 Sebelum Encoding Kategorikal	48
Gambar 4.16 Sesudah Encoding Kategorikal	48
Gambar 4.17 SMOTE Pada Ketidakseimbangan Data	50
Gambar 4.18 Confusion Matrix SVM	52
Gambar 4.19 Confusion Matrix Random Forest.....	52
Gambar 4.20 Confusion Matrix XGBoost	53
Gambar 4.21 Hyperparameter Tuning SVM.....	55
Gambar 4.22 Hyperparameter Tuning RF	56
Gambar 4.23 Hyperparameter Tuning XGBoost	56
Gambar 4.24 Confusion Matrix SVM Tuned	59

Gambar 4.25 Confusion Metrik RF Tuned	60
Gambar 4.26 Confusion Metrik XGBoost Tuned.....	60
Gambar 4.27 Feature Importances RF	63
Gambar 4.28 Feature Importances XGBoost.....	63
Gambar 4.29 Tampilan Deployment.....	65
Gambar 4.30 Status Lancar	66



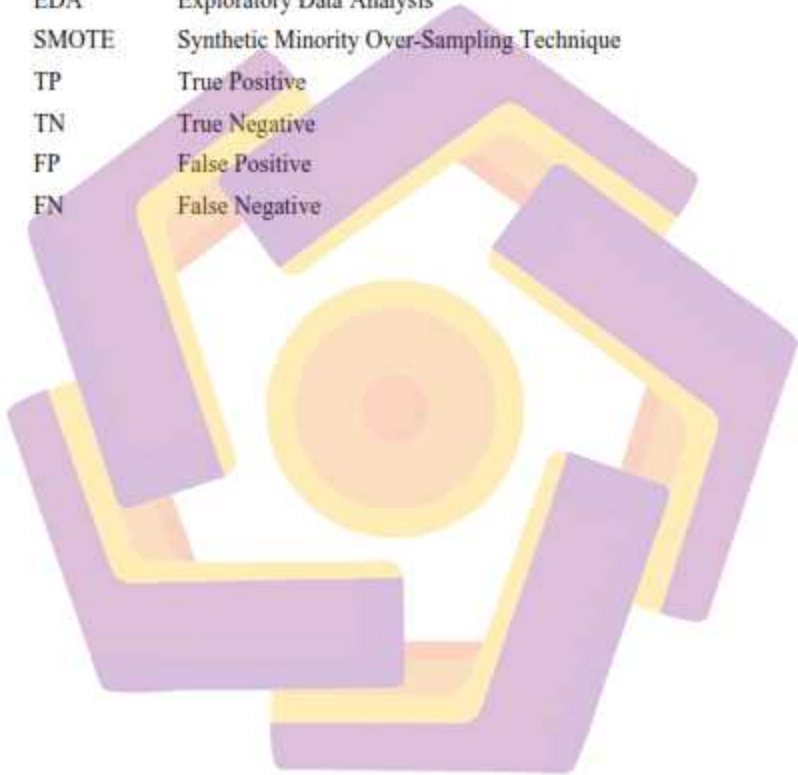
DAFTAR LAMPIRAN

Lampiran 1. Source Code.....	77
------------------------------	----

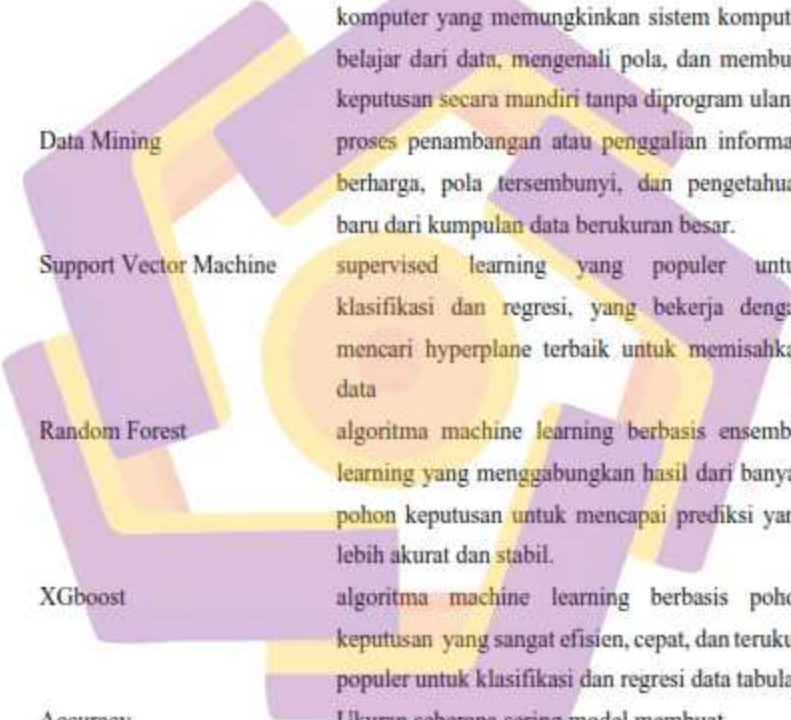


DAFTAR LAMBANG DAN SINGKATAN

RF	Random Forest
SVM	Support Vector Machines
XGBoost	Extreme Gradient Boosting
EDA	Exploratory Data Analysis
SMOTE	Synthetic Minority Over-Sampling Technique
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative



DAFTAR ISTILAH



Resiko Kredit	potensi kerugian finansial yang dihadapi pemberi pinjaman (bank/kreditur) akibat kegagalan debitur memenuhi kewajiban pembayaran pokok atau bunga secara penuh dan tepat waktu.
Machine Learning	cabang kecerdasan buatan (AI) dan ilmu komputer yang memungkinkan sistem komputer belajar dari data, mengenali pola, dan membuat keputusan secara mandiri tanpa diprogram ulang.
Data Mining	proses penambangan atau penggalian informasi berharga, pola tersembunyi, dan pengetahuan baru dari kumpulan data berukuran besar.
Support Vector Machine	supervised learning yang populer untuk klasifikasi dan regresi, yang bekerja dengan mencari hyperplane terbaik untuk memisahkan data
Random Forest	algoritma machine learning berbasis ensemble learning yang menggabungkan hasil dari banyak pohon keputusan untuk mencapai prediksi yang lebih akurat dan stabil.
XGboost	algoritma machine learning berbasis pohon keputusan yang sangat efisien, cepat, dan terukur, populer untuk klasifikasi dan regresi data tabular.
Accuracy	Ukuran seberapa sering model membuat prediksi yang benar dari total keseluruhan prediksi.
Precision	metrik evaluasi yang mengukur rasio antara jumlah data positif yang diprediksi dengan benar dibandingkan dengan total data yang diprediksi positif oleh model.

Recall (Sensitivity)

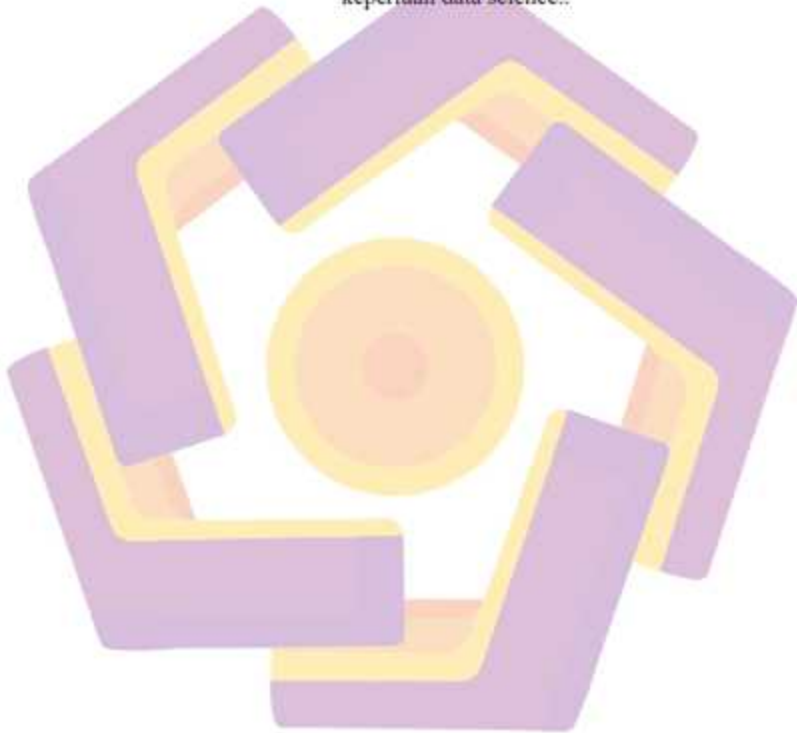
Ukuran kemampuan model dalam menemukan seluruh kasus aktual dari kelas positif.

F1-Score

Rata-rata harmonis dari precision dan recall, menggambarkan keseimbangan antara keduanya.

Google Colaboratory

Platform pemrograman berbasis cloud yang mendukung eksekusi kode Python untuk keperluan data science..



INTISARI

Perkembangan teknologi yang pesat telah mendorong transformasi digital di berbagai bidang, termasuk sektor keuangan. Saat ini, lembaga keuangan semakin bergantung pada data untuk mendukung proses pengambilan keputusan, terutama dalam pemberian pinjaman. Risiko kredit menjadi salah satu aspek penting yang harus diperhatikan, karena kesalahan dalam menilai kemampuan nasabah dapat menyebabkan terjadinya kredit macet, kegagalan pembayaran, suku bunga tinggi dan kerugian finansial yang signifikan. Oleh sebab itu, diperlukan metode yang mampu klasifikasi risiko kredit nasabah secara lebih akurat dan efisien.

Penelitian ini bertujuan untuk membandingkan kinerja algoritma yaitu SVM, Random Forest dan XGBoost dalam memprediksi apakah kredit nasabah mengalami resiko atau tidaknya. Dataset yang digunakan diperoleh dari platform Kaggle, yang memuat atribut-atribut penting seperti pendapatan, usia, masa kerja, kepemilikan rumah, jumlah dan nilai pinjaman, tujuan pinjaman, serta persentase pendapatan terhadap pinjaman. Tahapan penelitian meliputi EDA, pre processing data, pelatihan model, tuning hyperparameter, evaluasi kinerja menggunakan confusion matrix dan metrik akurasi, precision, recall, dan f1-score serta tahap deployment berbasis streamlit.

Hasil penelitian menunjukkan bahwa model XGBoost menghasilkan performa terbaik dengan akurasi sebesar 92,73%, diikuti oleh Random Forest sebesar 92,68% dan SVM sebesar 87,58% setelah tuning. Temuan ini menunjukkan bahwa metode klasifikasi berbasis machine learning efektif dalam memprediksi risiko kredit nasabah. Hasil penelitian ini diharapkan dapat dimanfaatkan oleh lembaga keuangan sebagai sistem pendukung keputusan dalam meningkatkan ketepatan evaluasi risiko kredit secara data-driven, serta menjadi referensi bagi penelitian selanjutnya.

Kata kunci: Resiko Kredit, SVM, Random Forest, XGBoost, Klasifikasi

ABSTRACT

The rapid advancement of technology has driven digital transformation across various sectors, including the financial industry. Today, financial institutions increasingly rely on data to support decision-making processes, particularly in credit lending. Credit risk is a critical aspect that must be carefully considered, as errors in assessing customers' repayment ability may lead to non-performing loans, payment defaults, high interest rates, and significant financial losses. Therefore, effective methods are required to classify customer credit risk accurately and efficiently.

This study aims to compare the performance of three classification algorithms, namely SVM, Random Forest, and XGBoost, in predicting whether a customer's credit is at risk. The dataset used in this research was obtained from the Kaggle platform and includes key attributes such as income, age, employment length, home ownership, loan amount, loan grade, loan purpose, and loan-to-income percentage. The research stages consist of EDA, data preprocessing, model training, hyperparameter tuning, performance evaluation using confusion matrix and metrics including accuracy, precision, recall, and F1-score, as well as deployment using a Streamlit-based application.

The experimental results indicate that the XGBoost model achieves the best performance with an accuracy of 92.73%, followed by Random Forest at 92.68% and SVM at 87.58% after tuning. These findings demonstrate that machine learning-based classification methods are effective for predicting customer credit risk. The results of this study are expected to support financial institutions in developing data-driven decision support systems to improve credit risk evaluation accuracy and to serve as a reference for future research.

Keyword: Credit Risk, SVM, Random Forest, XGBoost, Classification.