

**KLASIFIKASI PERSETUJUAN KREDIT MENGGUNAKAN
LOGISTIC REGRESSION DAN LIGHTGBM DENGAN
PENDEKATAN SOFT VOTING**

SKRIPSI

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



disusun oleh

NI PUTU SUMERTIANI

22.11.4875

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**KLASIFIKASI PERSETUJUAN KREDIT MENGGUNAKAN
LOGISTIC REGRESSION DAN LIGHTGBM DENGAN
PENDEKATAN SOFT VOTING**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

NI PUTU SUMERTIANI

22.11.4875

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

SKRIPSI

**KLASIFIKASI PERSETUJUAN KREDIT MENGGUNAKAN LOGISTIC
REGRESSION DAN LIGHTGBM DENGAN PENDEKATAN SOFT
VOTING**

yang disusun dan diajukan oleh

Ni Putu Sumertiani

22.11.4875

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 25 Februari 2026

Dosen Pembimbing,



I Made Artha Agatsya, S.T., M.Eng., Ph.D.
NIK. 190302352

HALAMAN PENGESAHAN
SKRIPSI
KLASIFIKASI PERSETUJUAN KREDIT MENGGUNAKAN LOGISTIC
REGRESSION DAN LIGHTGBM DENGAN PENDEKATAN SOFT
VOTING

yang disusun dan diajukan oleh

Ni Putu Sumertiani

22.11.4875

Telah dipertahankan di depan Dewan Penguji
pada tanggal 25 Februari 2026

Susunan Dewan Penguji

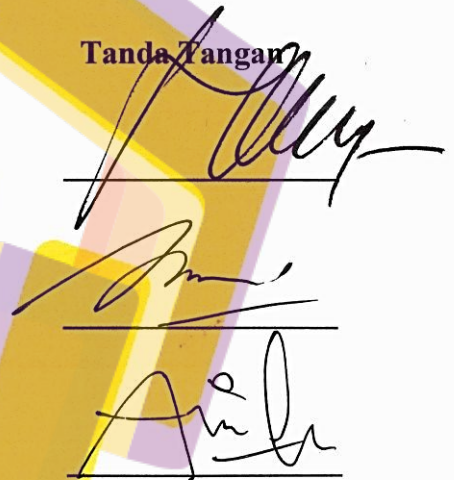
Nama Penguji

Tanda Tangan

Hanif Al Fattah, S.Kom., M.Kom., Ph.D.
NIK. 190302096

Bayu Setiaji, M.Kom.
NIK. 190302216

I Made Artha Agatsya, S.T., M.Eng., Ph.D.
NIK. 190302352



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 25 Februari 2026

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Ni Putu Sumertiani
NIM : 22.11.4875

Menyatakan bahwa Skripsi dengan judul berikut:

KLASIFIKASI PERSETUJUAN KREDIT MENGGUNAKAN LOGISTIC REGRESSION DAN LIGHTGBM DENGAN PENDEKATAN SOFT VOTING

Dosen Pembimbing : I Made Artha Agastya, S.T., M.Eng., Ph.D.

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, 25 Februari 2026

Yang Menyatakan,



Ni Putu Sumertiani

HALAMAN PERSEMBAHAN

Puji syukur penulis haturkan ke hadapan Ida Sang Hyang Widhi Wasa atas asung kerta wara nugraha-Nya, sehingga skripsi ini dapat diselesaikan dengan baik dan tepat waktu. Karya ini penulis persembahkan sebagai bentuk yadnya (korban suci yang tulus ikhlas) sekaligus ungkapan rasa bakti dan terima kasih kepada:

1. Kedua orang tua tercinta, Bapak dan Mama, serta adik tersayang Ketut, yang selalu menjadi sumber doa, kasih sayang, dan dukungan sehingga penulis mampu menyelesaikan tahap ini.
2. Saudara perempuan kembar penulis, Made, yang telah berjuang bersama dalam proses penyusunan skripsi ini, saling menguatkan hingga perjuangan ini secara dapat diselesaikan bersama.
3. Keluarga besar Pakyan, Mekngah, Pakmang, Mekdek, Ajik, Biang Ayu, Ninik dan Dadong, serta Ayuk Savitri, Ayuk Sayu, Kakak, dan Adi Saputra, atas doa, perhatian, dan semangat diberikan.
4. Mbok Saras, yang tulus membantu sejak awal proses penyusunan hingga skripsi ini dapat terselesaikan.
5. Sahabat terdekat, Rifqa Athifa, Ulfatun Badriyah, dan Cornelia, yang selalu berbagi semangat dan menemani perjalanan selama masa perkuliahan.
6. Seluruh teman Angkatan 2022 yang telah menjadi bagian dari proses belajar dan perjuangan hingga mencapai tahap akhir perkuliahan.

Semoga persembahan ini menjadi bagian dari yadnya sebagai wujud rasa syukur dan terima kasih atas doa, dukungan, serta kebaikan yang telah diberikan kepada penulis selama proses penyusunan skripsi ini.

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadapan Ida Sang Hyang Widhi Wasa atas limpahan asung kerta wara anugerah-nya, sehingga skripsi ini yang berjudul “Klasifikasi Persetujuan Kredit Menggunakan Logistic Regression dan LighGBM dengan Pendekatan Soft Voting Ensemble” dapat diselesaikan dengan baik dan tepat waktu. Penyusunan skripsi ini dilakukan sebagai salah satu persyaratan untuk menyelesaikan pendidikan pada Program Studi Informatika, Universitas AMIKOM Yogyakarta.

Selama proses penyusunan skripsi ini tidak terlepas dari bantuan, dukungan, bimbingan, dan arahan dari berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada :

1. Prof. Dr. M. Suyanto, M.M., selaku Universitas AMIKOM Yogyakarta.
2. Prof. Dr. Kusriani, M.Kom., selaku Dekan I Fakultas Ilmu Komputer.
3. Dr. Andi Syafrizal, M.Eng., Ph.D., selaku Wakil Dekan II Fakultas Ilmu Komputer
4. I Made Artha Agastya, S.T., M.Eng., Ph.D., selaku dosen pembimbing yang telah memberikan bimbingan dan arahan selama penyusunan skripsi.
5. Orang-rang terdekat yang tidak dapat disebutkan satu per satu, yang senantiasa memberikan doa, semangat, dan dukungan hingga skripsi ini dapat terselesaikan.

Penulis menyadari bahwa skripsi ini masih memiliki keterbatasan. Oleh karena itu, kritik dan saran yang sangat diharapkan demi perbaikan dimasa mendatang. Semoga skripsi ini dapat memberikan manfaat dan kontribusi bagi pengembangan ilmu pengetahuan.

Yogyakarta, 23 Februari 2026

Ni Putu Sumertiani

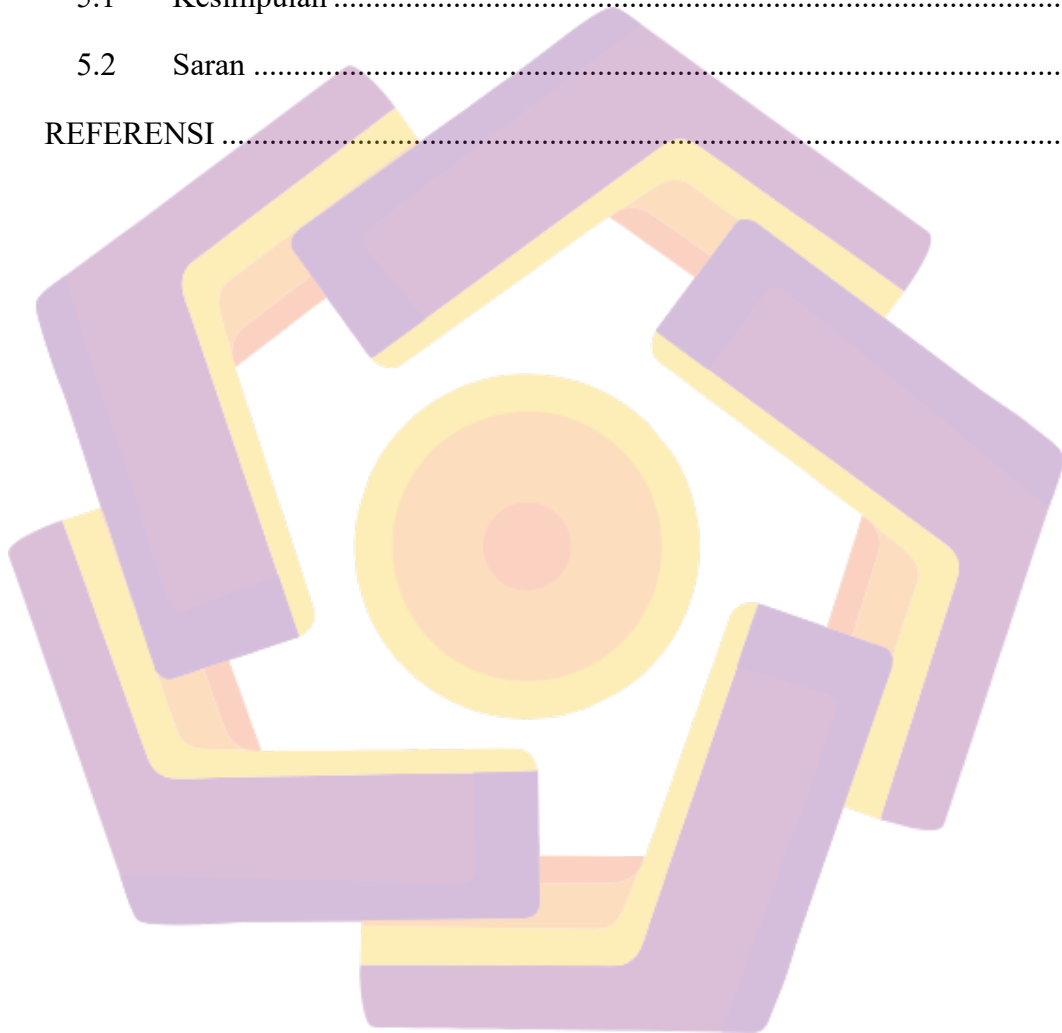
DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xiv
INTISARI	xv
<i>ABSTRACT</i>	xvi
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	6
2.1 Studi Literatur	6
2.2 Dasar Teori.....	12

2.2.1	Pinjaman	12
2.2.2	Risiko Kredit.....	12
2.2.3	Kelayakan Kredit	12
2.2.4	Klasifikasi	13
2.2.5	Machine Learning	13
2.2.6	Feature Selection.....	13
2.2.7	SMOTE (Synthetic Minority Over-sampling Technique)	13
2.2.8	Hyperparameter Tuning (Randomized Search)	14
2.2.9	Stratified K-Fold Cross Validation	14
2.2.10	Logistic Regression (LR).....	16
2.2.11	LightGBM.....	17
2.2.12	Soft Voting Ensemble	18
2.2.13	Confusion Matrix	19
2.2.14	Evaluasi Model	20
BAB III METODOLOGI PENELITIAN		23
3.1	Objek Penelitian.....	23
3.2	Alur Penelitian	23
3.2.1	Pengumpulan data.....	25
3.2.2	Exploratory Data Analysis (EDA).....	25
3.2.3	Preprocessing Data.....	25
3.2.4	Pembagian Data (Split Data)	26
3.2.5	Implementasi Model Klasifikasi	26
3.2.6	Feature Selection.....	26
3.2.7	Penyeimbang data dengan SMOTE	27
3.2.8	Hyperparameter Tuning	27

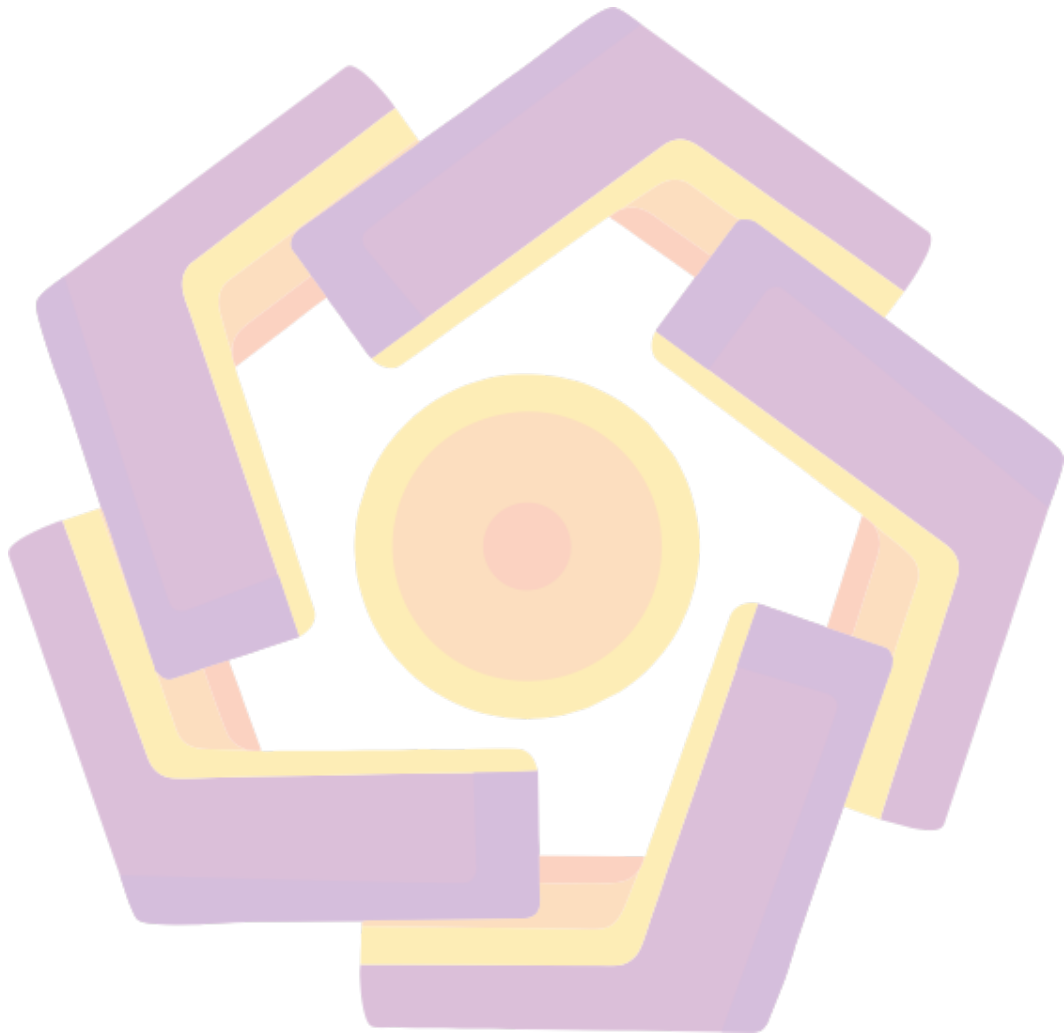
3.2.9	Soft Voting Ensemble	27
3.2.10	Evaluasi Model	28
3.3	Alat dan Bahan.....	28
3.3.1	Data Penelitian	28
3.3.2	Alat/Instrument	29
BAB IV HASIL DAN PEMBAHASAN		31
4.1	Pengumpulan Data	31
4.2	Exploratory Data Analysis (EDA)	31
4.2.1	Analisis Distribusi Target	31
4.2.2	Analisis Distribusi Fitur Numerik (Boxplot).	33
4.2.3	Analisis pada Fitur Numerik Kernel Density Estimation (KDE) ..	35
4.2.4	Analisis pada Kelas Kategorikal	38
4.2.5	Heatmap Korelasi.....	40
4.3	Preprocessing Data.....	41
4.3.1	Cleaning	41
4.3.2	Encoding	43
4.3.3	Pemisahan Fitur dan Label.....	43
4.4	Split Data	44
4.5	Implementasi Model Klasifikasi Awal	44
4.5.1	Logistic Regression.....	45
4.5.2	LightGBM.....	45
4.6	Feature Selection.....	45
4.7	Penyeimbangan data (SMOTE)	47
4.8	Hyperparameter Tuning	48
4.9	Soft Voting Ensemble	48

4.10	Evaluasi model.....	49
4.10.1	Confusion Matrix.....	51
4.10.2	Hasil Evaluasi	52
BAB V		55
5.1	Kesimpulan	55
5.2	Saran	56
REFERENSI		57



DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian	8
Tabel 2. 2 Confusion Matrix	20
Tabel 4. 1 Hasil Penelitian	52

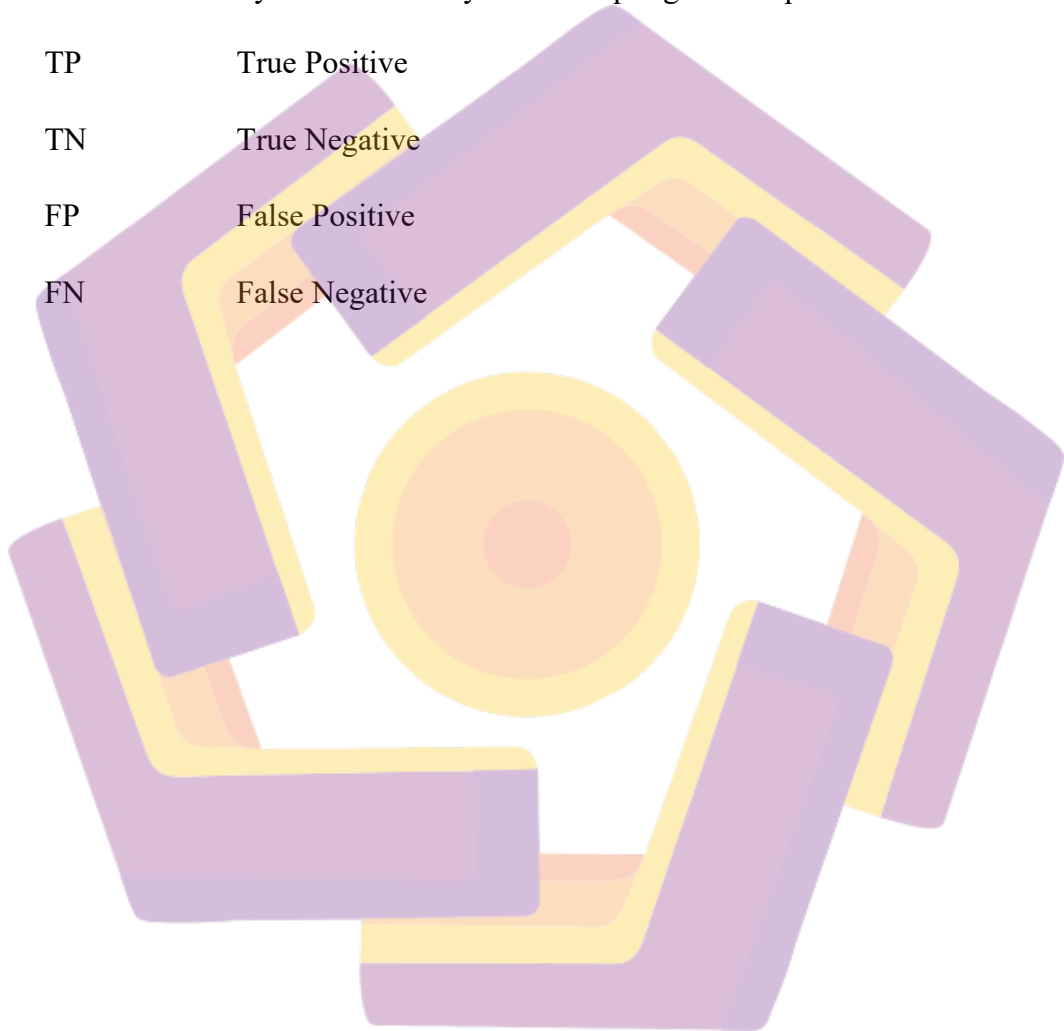


DAFTAR GAMBAR

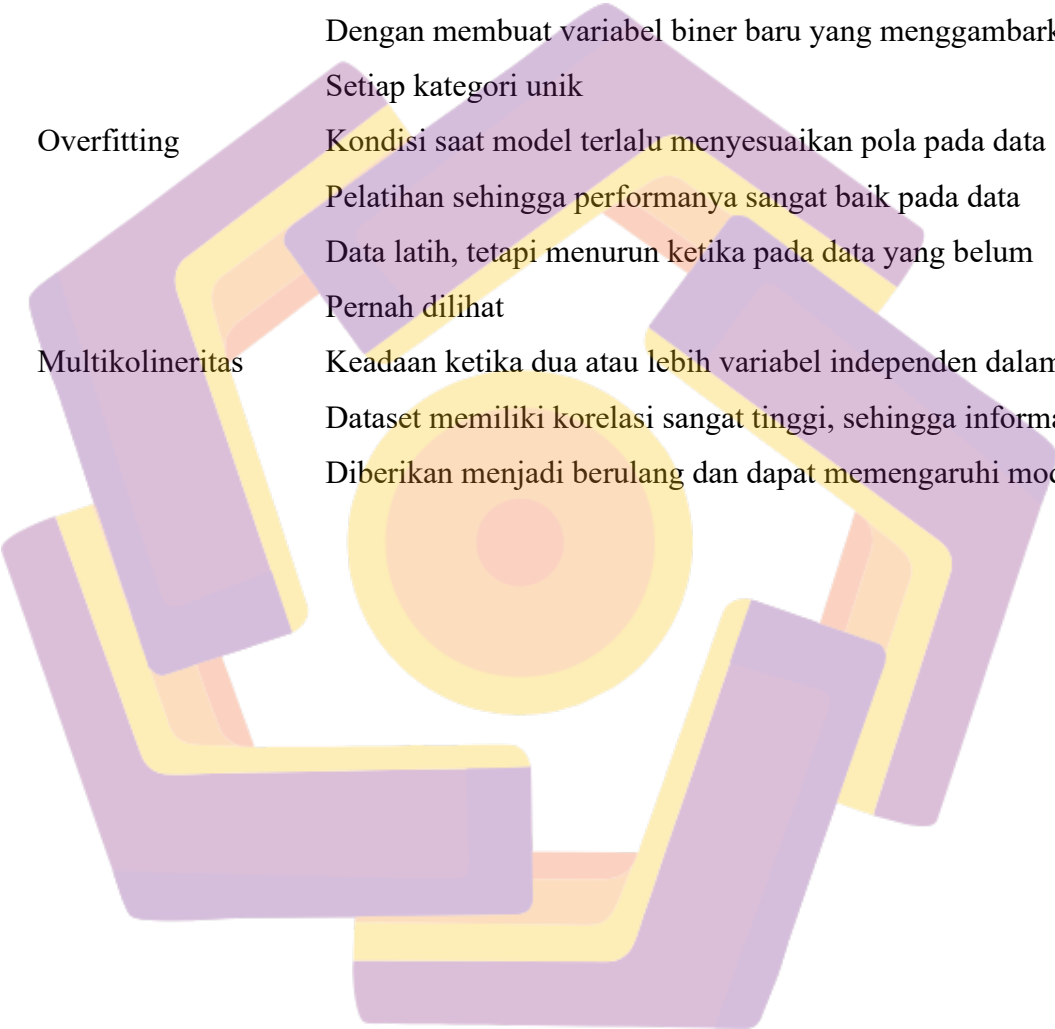
Gambar 2. 1 Ilustrasi Stratified K-Fold Cross Validation	16
Gambar 2. 2 Ilustrasi Logistic Regression	17
Gambar 2. 3 Ilustrasi LightGBM	18
Gambar 2. 4 Ilustrasi Soft Voting Ensemble	19
Gambar 3. 1 Alur Penelitian	24
Gambar 4. 1 Distribusi kelas pada variabel target LoanApproved	32
Gambar 4. 2 Distribusi Fitur Numerik Boxplot	33
Gambar 4. 3 Distribusi Fitur Numerik KDE	35
Gambar 4. 4 Visualisasi countplot fitur kategorikal terhadap target	38
Gambar 4. 5 Heatmap Korelasi	40
Gambar 4. 6 Hasil Seleksi Fitur Menggunakan SelectKBest (ANOVA F-test)	46
Gambar 4. 7 Distribusi kelas target sebelum SMOTE	47
Gambar 4. 8 Distribusi kelas target setelah penerapan SMOTE	48
Gambar 4. 9 Soft Voting Ensemble	49
Gambar 4. 10 Evaluasi Stratified K-Fold	50
Gambar 4. 11 Confusion Matrix	51

DAFTAR LAMBANG DAN SINGKATAN

LR	Logistic Regression
LightGBM	Light Gradient Boosting Machine
SMOTE	Synthetic Minority Over-sampling Technique
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative



DAFTAR ISTILAH



Soft Voting	Metode ensemble yang menggabungkan probabilitas dari beberapa model untuk menentukan kelas akhir berdasarkan nilai probabilitas tertinggi.
One-Hot Encoding	Metode konversi data kategorikal ke bentuk numerik Dengan membuat variabel biner baru yang menggambarkan Setiap kategori unik
Overfitting	Kondisi saat model terlalu menyesuaikan pola pada data Pelatihan sehingga performanya sangat baik pada data Data latih, tetapi menurun ketika pada data yang belum Pernah dilihat
Multikolineritas	Keadaan ketika dua atau lebih variabel independen dalam Dataset memiliki korelasi sangat tinggi, sehingga informasi Diberikan menjadi berulang dan dapat memengaruhi model

INTISARI

Perkembangan teknologi finansial telah mendorong transformasi layanan kredit dari sistem konvensional menuju platform digital. Perubahan ini memungkinkan proses evaluasi kelayakan kredit dilakukan secara lebih efektif dan terstruktur. Namun, meningkatnya penggunaan layanan kredit digital juga menyebabkan lonjakan jumlah pengajuan pinjaman yang harus dianalisis secara akurat dan konsisten. Pertumbuhan penyaluran kredit yang terus meningkat menunjukkan pentingnya sistem evaluasi risiko kredit yang responsif dan mampu mengelola data dalam skala besar. Ketidaktepatan dalam evaluasi dapat meningkatkan risiko kredit bermasalah, seperti gagal bayar, yang berdampak pada kerugian finansial serta menurunkan kepercayaan terhadap sistem keuangan digital.

Penelitian ini bertujuan mengembangkan model klasifikasi persetujuan kredit menggunakan metode *Soft Voting Ensemble* yang mengombinasikan *Logistic Regression* dan *LightGBM*. *Logistic Regression* digunakan untuk menangkap hubungan linier antarvariabel secara interpretatif, sedangkan *LightGBM* mengidentifikasi pola non-linier dan interaksi fitur. Tahapan penelitian meliputi preprocessing data, pelatihan model, serta optimasi melalui hyperparameter tuning dengan pendekatan ensemble untuk meningkatkan performa model.

Hasil penelitian menunjukkan bahwa penerapan *SMOTE* dan hyperparameter tuning meningkatkan kinerja model dibandingkan *baseline*. Model terbaik diperoleh melalui *Soft Voting Ensemble (SMOTE + tuning)* dengan akurasi 97,08% dan *F1-score* 97,11%, sedikit lebih tinggi dibandingkan *LightGBM* tunggal. Temuan ini menunjukkan bahwa penggabungan *Logistic Regression* dan *LightGBM* melalui metode *Soft Voting Ensemble* mampu meningkatkan stabilitas dan performa klasifikasi dalam evaluasi kelayakan kredit. Hasil penelitian ini dapat dimanfaatkan oleh lembaga keuangan serta menjadi referensi bagi pengembangan metode lebih lanjut.

Kata kunci: Kelayakan kredit, Soft Voting Ensemble, Logistic Regression, LightGBM, Klasifikasi.

ABSTRACT

The development of financial technology has driven the transformation of credit services from conventional systems to digital platforms. This transformation enables the creditworthiness evaluation process to be conducted more effectively and systematically. However, the increasing use of digital credit services has also led to a surge in loan applications that must be analyzed accurately and consistently. The continuous growth in credit distribution highlights the importance of a responsive credit risk evaluation system capable of managing large-scale data. Inaccurate evaluations may increase the risk of non-performing loans, such as defaults, which can result in financial losses and reduce trust in the digital financial system.

This study aims to develop a credit approval classification model using the Soft Voting Ensemble method that combines Logistic Regression and LightGBM. Logistic Regression is used to capture linear relationships between variables in an interpretable manner, while LightGBM identifies non-linear patterns and feature interactions. The research stages include data preprocessing, model training, and optimization through hyperparameter tuning with an ensemble approach to improve model performance.

The results show that the application of SMOTE and hyperparameter tuning improves model performance compared to the baseline. The best model is obtained using Soft Voting Ensemble (SMOTE + tuning) with an accuracy of 97.08% and an F1-score of 97.11%, slightly higher than the single LightGBM model. These findings indicate that combining Logistic Regression and LightGBM through the Soft Voting Ensemble method enhances the stability and performance of creditworthiness classification. The results of this study can be utilized by financial institutions and serve as a reference for further methodological development.

Keywords: Creditworthiness, Soft Voting Ensemble, Logistic Regression, LightGBM, Classification.