

Prediksi Kelayakan Pemberian Kredit Menggunakan Metode Random Forest

Oleh:

Noval Firmansah,

Uce Indahyanti

Progam Studi Teknik Informatika

Universitas Muhammadiyah Sidoarjo

Oktober, 2023

Pendahuluan

Saat ini, hampir semua sektor memerlukan kemudahan dalam mengelola informasi yang dimiliki, termasuk dalam bidang pengkreditan. Pembelian dengan sistem kredit sudah menjadi hal umum dalam masyarakat.

Dengan banyaknya jumlah konsumen yang mengajukan kredit akan menimbulkan penumpukan data pengajuan kredit. Hal ini berdampak besar pada bagian Credit Analyst selaku pihak yang menentukan kelayakan pemberian kredit sepeda motor dan memegang data para customer

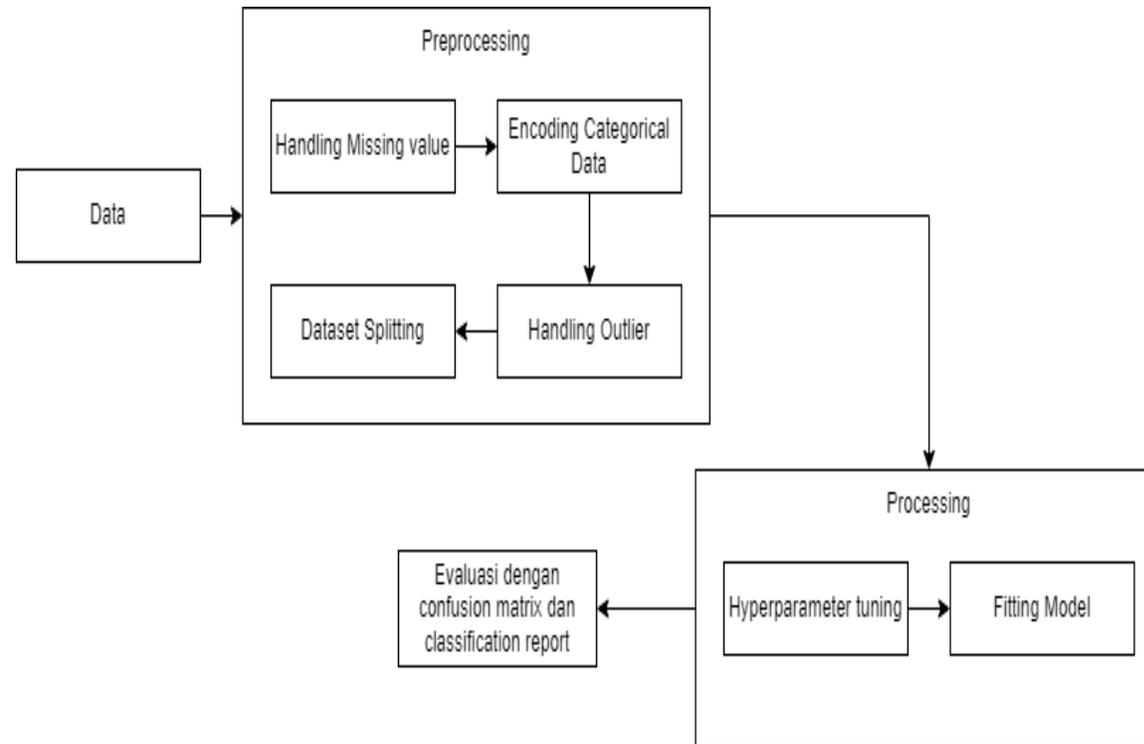
Penggunaan data mining untuk klasifikasi kelayakan pemberian kredit akan sangat membantu pihak perusahaan dimana klasifikasi itu sendiri untuk membedakan kelas data yang layak dan tidak layak untuk melakukan kredit

Tujuan Penelitian

- Membangun sebuah model machine learning yang mampu memprediksi kelayakan kredit dengan baik
- Mengimplementasikan algoritma Random Forest untuk melakukan klasifikasi kelayakan kredit

Metode Penelitian

Flowchart



Metode Penelitian

Pre Processing

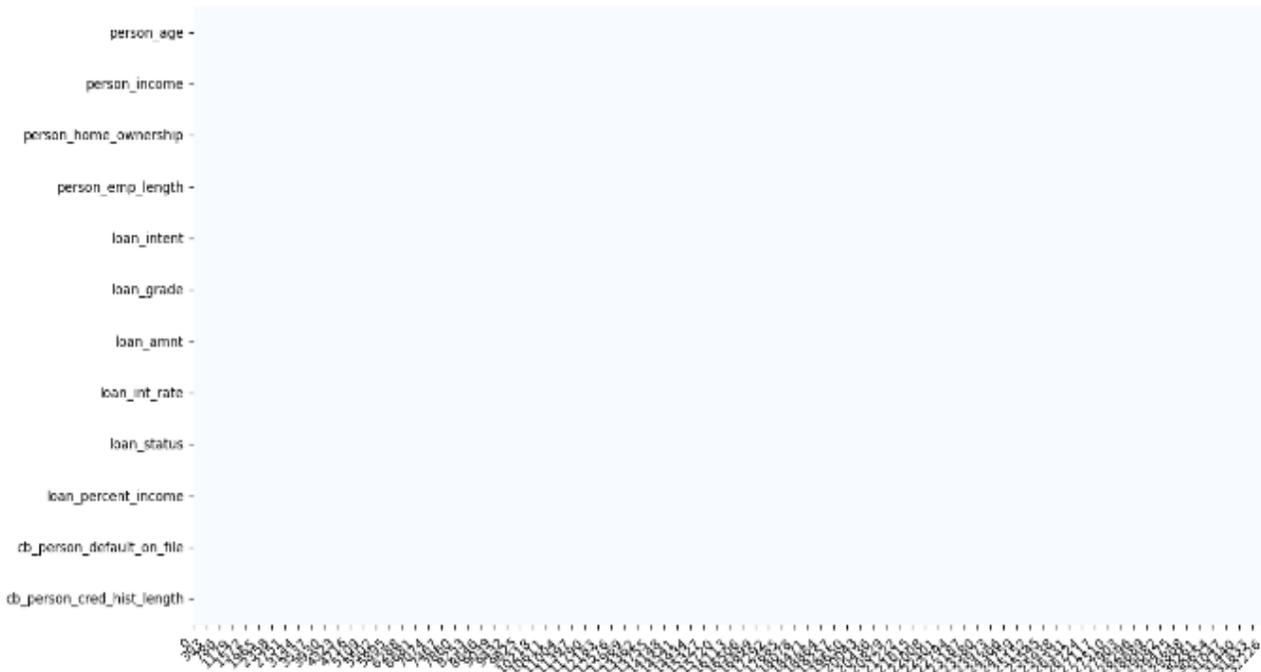
- Handling Missing Values
- Encoding Categorical Data
- Handling Outlier
- Dataset Splitting

Processing

- Hyperparameter Tuning
- Fitting Model

Pre Processing

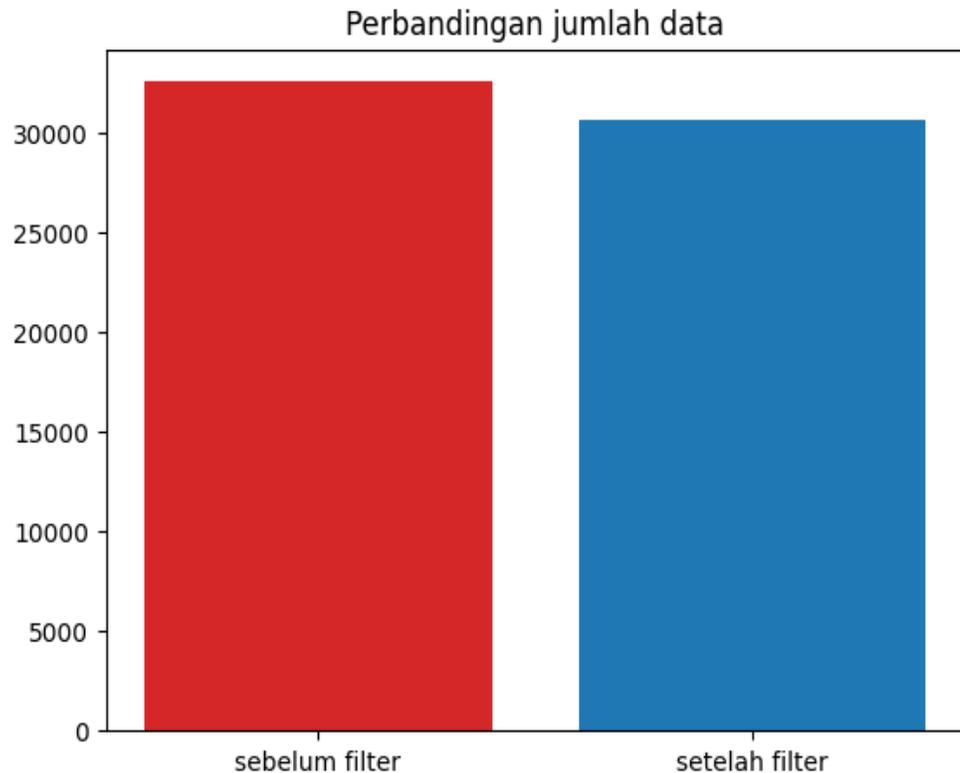
Hasil Handling Missing Value



Metode yang diterapkan adalah imputasi mean, yang memungkinkan peneliti untuk mengisi nilai-nilai yang kosong dengan rerata dari atribut yang bersangkutan.

Hasil Pre Processing

Perbandingan Jumlah Data Setelah Pre Processing



Hasil dari langkah ini adalah pemfilteran data, mengurangi jumlah baris dalam dataset dari 32,581 menjadi 30,649.

Processing

Hyperparameter Tuning

Hyperparameter	Value
n_estimators	10, 50, 100
max_depth	None, 10, 20, 30
min_samples_split	2, 5, 10
min_samples_leaf	1, 2, 4

Proses tuning ini bertujuan untuk menemukan kombinasi hyperparameter terbaik yang akan meningkatkan performa model Random Forest dalam memprediksi kelayakan kredit.

Hasil Processing

Hasil Akurasi Tiap Dataset Split

Rasio Data Tes	Rasio Data Pelatihan	Akurasi
90%	10%	93,28%
80%	20%	93,07%
75%	25%	93,21%

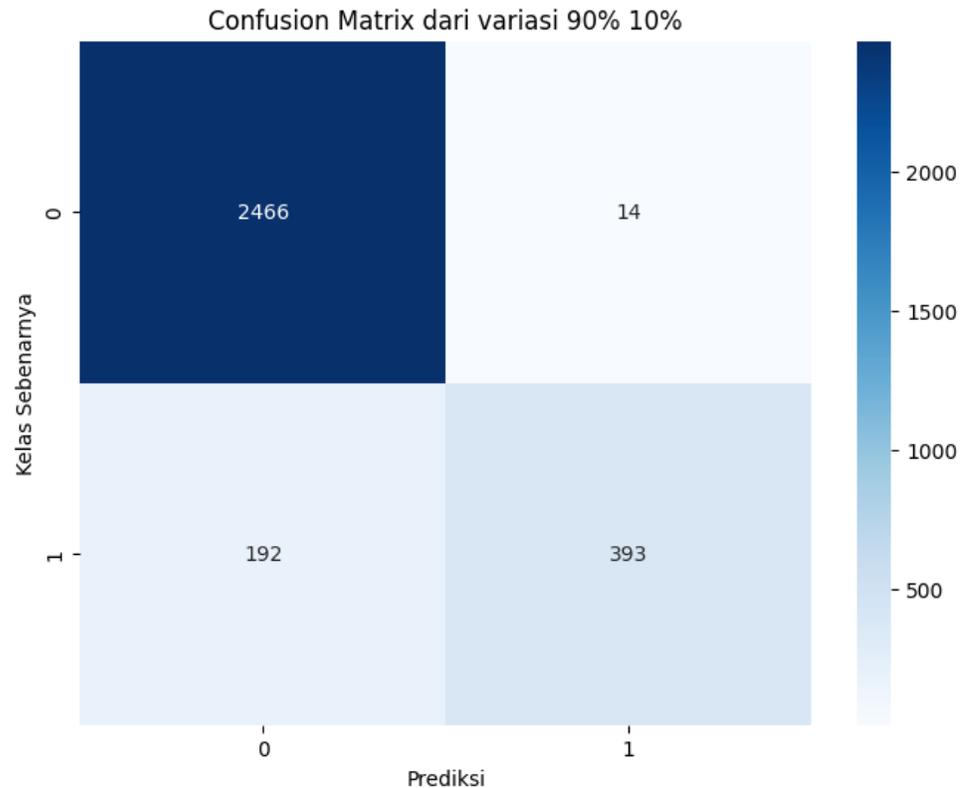
Hasil Processing

Hasil Tuning Tiap Dataset Split

Rasio Data Tes	Rasio Data Pelatihan	Hyperparameter	Value
90%	10%	n_estimators	100
		max_depth	30
		min_samples_split	2
		min_samples_leaf	1
80%	20%	n_estimators	100
		max_depth	None
		min_samples_split	2
		min_samples_leaf	1
75%	25%	n_estimators	100
		max_depth	30
		min_samples_split	2
		min_samples_leaf	1

Evaluasi

Confusion Matrik Hasil Akurasi Terbaik



- True Positive (TP): Sebanyak 2466 kasus berhasil diprediksi sebagai "kredit lancar."
- False Positive (FP): Terdapat 14 kasus yang salah diprediksi sebagai "kredit lancar."
- False Negative (FN): Terdapat 192 kasus yang salah diprediksi sebagai "kredit macet."
- True Negative (TN): Sebanyak 393 kasus berhasil diprediksi sebagai "kredit macet."

Penutup

Dengan menggunakan algoritma Random Forest dan melalui proses pra-pemrosesan data yang cermat, peneliti berhasil mencapai tingkat akurasi sebesar 93,28% dalam memprediksi apakah seorang peminjam layak atau tidak menerima kredit.

Selain itu, parameter terbaik untuk model ini telah diidentifikasi, yaitu 'max_depth': 30, 'min_samples_leaf': 1, 'min_samples_split': 2, dan 'n_estimators': 100, yang menunjukkan konfigurasi optimal untuk analisis kelayakan kredit.

