

BAB I PENDAHULUAN

1.1 Latar Belakang

Proses persetujuan pinjaman memegang peran krusial dalam menjaga keberlanjutan bisnis serta aksesibilitas kredit bagi masyarakat. Lembaga keuangan menggunakan berbagai metode untuk menilai kelayakan kredit calon peminjam. Keputusan yang diambil dalam proses ini tidak hanya berdampak pada individu tetapi juga pada stabilitas sistem keuangan secara keseluruhan.

Lembaga keuangan masih mengandalkan metode konvensional dalam proses persetujuan pinjaman. Sistem penilaian kredit tradisional tetap menjadi pendekatan utama dalam industri keuangan. Lebih dari 70% Lembaga keuangan di Amerika Serikat masih menggunakan metode manual dalam menentukan kelayakan pinjaman, meskipun teknologi baru telah tersedia [1].

Metode konvensional dalam persetujuan pinjaman memiliki beberapa kelemahan. Proses manual sering kali memakan waktu dan meningkatkan potensi bias subjektif dalam pengambilan Keputusan. Validasi data yang tidak optimal mengakibatkan rendahnya transparansi dalam sistem ini. Sistem tradisional memiliki tingkat kesalahan prediksi hingga 15-20%, yang menyebabkan ketidakakuratan dalam evaluasi kelayakan kredit [2].

Machine Learning (ML) menawarkan pendekatan yang lebih modern dalam proses evaluasi kredit. Model ML mampu memproses dan menganalisis berbagai variabel seperti jumlah pendapatan, status pernikahan, jumlah tanggungan, serta riwayat kredit untuk memberikan prediksi lebih akurat. Model ML berhasil meningkatkan akurasi penilaian kredit hingga 85% dibanding dengan metode konvensional [3].

Penerapan *Machine Learning* (ML) dalam sistem perbankan menghadapi beberapa tantangan. Model ML dapat mengalami *overfitting*, di mana model terlalu menyesuaikan diri dengan data pelatihan dan kehilangan kemampuan untuk membuat prediksi yang akurat pada data baru. Selain itu, potensi bias dalam data pelatihan dapat menyebabkan ketidakadilan dalam pengambilan keputusan. Bias dalam model *Machine Learning* dapat muncul dari ketidakseimbangan dalam dataset historis, yang beresiko memperkuat ketidakadilan sistemik dalam industri keuangan [4].

Proyek ini bertujuan untuk membangun model prediksi persetujuan pinjaman berbasis *Machine Learning* (ML) serta menganalisis potensi bias dalam model tersebut. Model ML akan ditetapkan dalam sebuah website sederhana menggunakan Flask sebagai *backend*. Sistem ini memungkinkan pengguna untuk memasukkan data dan mendapatkan prediksi secara langsung. Implementasi ini diharapkan dapat

meningkatkan efisiensi serta transparansi dalam proses persetujuan pinjaman, sekaligus mengurangi subjektivitas dalam pengambilan keputusan kredit.

Flask dipilih karena merupakan *micro-framework* Python yang ringan dan fleksibel, sehingga sangat cocok digunakan untuk mengintegrasikan model *Machine Learning* ke dalam aplikasi web. Dibandingkan dengan framework lain, Flask memberikan control yang lebih besar kepada pengembang untuk membangun aplikasi secara modular dan cepat, tanpa terlalu banyak konfigurasi. Selain itu, bahasa pemrograman Python sendiri mendukung banyak *library* yang sangat kaya untuk pengembangan *Machine Learning*, seperti *scikit-learn* dan *Pandas* [5]. Hal ini menjadikan Flask sebagai pilihan ideal untuk proyek pengembangan sistem prediktif berbasis Python [6].

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, rumusan masalah dalam proyek ini adalah sebagai berikut:

1. Bagaimana membangun model yang dapat menilai kelayakan pinjaman nasabah?
2. Bagaimana mengimplementasikan model prediksi tersebut ke dalam website agar dapat digunakan dengan mudah oleh pengguna?

1.3 Tujuan

Proyek ini bertujuan untuk:

1. Membangun model Machine Learning untuk memprediksi kelayakan pinjaman berdasarkan variable tertentu menggunakan metrik evaluasi yang sesuai.
2. Membangun website sederhana menggunakan Flask untuk mengimplementasikan model prediksi agar dapat diakses oleh pengguna secara interaktif.

1.4 Cakupan Pengerjaan

1. Penyediaan Dataset

Dataset yang digunakan adalah “*Loan Eligible Dataset*” yang diperoleh dari platform Kaggle [7]. Dataset ini berisi atribut-atribut penting terkait peminjam, seperti pendapatan, status pernikahan, jumlah tanggungan, riwayat kredit, jumlah pinjaman, dan lama pinjaman. Dataset ini dipilih karena memiliki struktur yang representatif dan sering digunakan dalam studi prediksi pinjaman. Dataset akan melalui proses pembersihan dan transformasi sebelum digunakan dalam pelatihan model.

2. Pembangunan Model *Machine Learning*

Model prediksi akan dibangun menggunakan algoritma klasifikasi Random Forest Classifier, yang dikenal efektif untuk menangani data tabular dan klasifikasi biner seperti kelayakan pinjaman [8]. Proses Pembangunan model melibatkan beberapa tahapan penting, yaitu:

- a. *Pre-processing Data*: mencakup *dataset cleaning*, *label encoding*, *data splitting*, serta *feature scaling*.
- b. *Training Model*: data yang sudah dibagi pada tahap *data splitting* dengan rasio 70:30 untuk data latih dan data uji untuk memastikan evaluasi yang adil menggunakan parameter default.
- c. *Hyperparameter Optimization (HPO)*: Proses tuning akan dilakukan menggunakan metode *Grid Search* untuk mendapatkan kombinasi parameter terbaik yang nantinya akan digunakan untuk *training* model.
- d. *Evaluasi Model*: Model akan diuji menggunakan metrik seperti *score*, *recall*, *f1-Score*, dan *confusion matrix* untuk mengetahui performa prediksi secara menyeluruh.

3. Implementasi ke dalam Website menggunakan Flask

Sistem prediksi akan diintegrasikan ke dalam aplikasi web menggunakan *micro-framework* Flask, yang ringan dan fleksibel untuk keperluan *deployment Machine Learning*. Website akan memiliki antarmuka pengguna yang sederhana, di mana pengguna dapat memasukkan data secara manual melalui form, kemudian menerima hasil prediksi langsung dari model.

1.5 Tahapan Pengerjaan

Proyek ini akan dilakukan dengan mengadopsi metodologi pengembangan *Cross-Industry Standard Process for Data Mining (CRISP-DM)* melalui beberapa tahapan sebagai berikut. Untuk memberikan gambaran yang lebih jelas mengenai alur pengerjaan proyek ini, penjelasan tiap tahapan akan disertai dengan ilustrasi visual pada Gambar 1.1 berikut:



Gambar 1.1 Tahapan pengembangan metode CRISP-DM (Sumber: Dicoding.com)

1. *Business Understanding*

Menganalisis permasalahan dalam proses persetujuan pinjaman secara manual yang umum dilakukan di Lembaga keuangan kecil hingga menengah.

Langkah-langkah yang dilakukan:

- a. Menyusun rumusan masalah dan tujuan proyek.
- b. Mengkaji kebutuhan pengguna terhadap sistem yang cepat dan objektif.
- c. Mengidentifikasi peluang pemanfaatan teknologi *Machine Learning* untuk menyelesaikan masalah tersebut.

2. *Data Understanding*

Memahami struktur dan karakteristik dataset yang bersumber dari Kaggle dengan nama "*Loan Eligible Dataset*" yang memuat informasi pemohon pinjaman seperti status pernikahan, pendidikan, pendapatan, riwayat kredit, dan jumlah pinjaman sebanyak 614 baris. Langkah-langkah yang dilakukan:

- a. Mengimpor dataset ke dalam lingkungan Python (menggunakan Pandas).
- b. Melakukan eksplorasi awal seperti: jumlah data, jenis variabel, distribusi kelas target.
- c. Mengidentifikasi nilai kosong

3. *Data Preparation*

Setelah memahami struktur data, tahapan yang dilakukan berfokus pada *pre-processing* data agar dapat digunakan oleh algoritma *Machine Learning*.

Langkah-langkah yang dilakukan:

- a. Mengisi nilai kosong pada kolom yang kosong
- b. Melakukan encoding pada kolom kategorik
- c. Memisahkan dataset menjadi training set dan testing set dengan proporsi 70:30
- d. Melakukan *feature scaling* untuk menyeragamkan rentang nilai antar kolom

4. *Modeling*

Melakukan tahapan Pembangunan model prediksi menggunakan algoritma *Rainbow Forest Classifier*, yang cocok untuk masalah klasifikasi dengan data tabular.

Langkah-langkah yang dilakukan:

- a. Membangun baseline model menggunakan parameter *default*.
- b. Menggunakan *GridSearchCV* untuk melakukan *Hyperparameter Optimization* (HPO) untuk mendapatkan kombinasi parameter terbaik.
- c. Melatih model lagi menggunakan parameter terbaik hasil HPO

5. *Evaluation*

Model yang telah dilatih dievaluasi menggunakan metrik-metrik klasifikasi untuk mengukur performa secara menyeluruh.

Langkah-langkah yang dilakukan:

- a. Menggunakan metrik *score*, *recall*, *f1-score* pada data uji.
- b. Membuat *confusion matrix* untuk melihat distribusi prediksi.
- c. Menganalisis hasil evaluasi dan memilih model dengan performa terbaik.

6. *Deployment*

Model yang telah dilatih dan disimpan kemudian diintegrasikan ke dalam sebuah aplikasi web berbasis Flask, agar dapat digunakan oleh pengguna secara langsung.

Langkah-langkah yang digunakan:

- a. Menyimpan model terbaik menggunakan *library joblib*
- b. Mendesain *User Interface* (UI) sederhana menggunakan HTML + CSS (Bootstrap).
- c. Mengembangkan *backend* menggunakan Flask.
- d. Menghubungkan *form input* pengguna ke model prediksi.
- e. Menampilkan hasil prediksi secara real-time.
- f. Melakukan pengujian fungsional untuk memastikan sistem berjalan sesuai harapan.

Dengan tahapan-tahapan tersebut, proyek ini tidak hanya menghasilkan model *Machine Learning* yang dapat memprediksi kelayakan pinjaman, tetapi juga menyediakan antarmuka yang mudah diakses dan digunakan secara interaktif.