

BAB I PENDAHULUAN

1.1 Latar Belakang

Dalam era digitalisasi yang semakin pesat, sektor keuangan mengalami perkembangan signifikan, terutama dalam hal kredit dan pinjaman. Layanan pinjaman kini tidak hanya terbatas pada lembaga keuangan tradisional, tetapi juga meluas ke *platform fintech* yang menawarkan kemudahan akses bagi pemohon. Walaupun prosesnya mudah, risiko kredit masih menjadi tantangan yang perlu diantisipasi dengan baik. Lembaga keuangan membutuhkan metode yang efektif untuk memprediksi dengan cara mengklasifikasikan risiko kredit guna mengurangi potensi kerugian.

Saat ini, metode tradisional dalam menilai kelayakan kredit sering tidak mampu menangkap kompleksitas pola pengajuan pinjaman yang dinamis [1]. Pendekatan berbasis data dan machine learning diharapkan dapat memberikan akurasi lebih tinggi dalam analisis risiko kredit. Algoritma seperti XGBoost dan C4.5 menjadi pilihan utama dalam membangun model prediksi yang lebih akurat dan andal dalam mengelompokkan data lama dan baru [2], [3]. XGBoost dikenal efektif dalam menanggulangi kasus machine learning berskala besar dan sangat baik dalam menganalisis data, sementara algoritma C4.5 menawarkan kemudahan interpretasi hasil [2], [3].

Namun, efektivitas kedua algoritma dalam konteks risiko kredit masih menjadi pertanyaan. Perbedaan karakteristik algoritma XGBoost dan C4.5 dalam menangani data pemohon pinjaman memunculkan kebutuhan untuk memahami sejauh mana keunggulan masing-masing dapat diterapkan pada analisis risiko kredit. Beberapa penelitian sebelumnya menggarisbawahi pentingnya memilih metode yang tepat dalam analisis risiko kredit [1].

Berdasarkan uraian latar belakang diatas, bisa disimpulkan pentingnya melakukan penelitian untuk membandingkan kinerja algoritma XGBoost dan C4.5 dalam memprediksi dengan cara mengklasifikasikan risiko kredit lebih baik lagi

dengan faktor-faktor tertentu. Hasil penelitian ini diharapkan dapat memberikan panduan bagi lembaga keuangan dalam memilih metode yang lebih efektif dan efisien, serta menambah referensi dalam analisis risiko kredit berbasis machine learning.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, penelitian ini merumuskan beberapa pertanyaan penelitian sebagai berikut:

1. Bagaimana perbandingan antara algoritma XGBoost dan C4.5 dalam mengklasifikasikan dan memprediksi risiko kredit berdasarkan data pemohon pinjaman?
2. Seberapa efektif algoritma XGBoost dibandingkan dengan algoritma C4.5 dalam mengklasifikasikan risiko kredit berdasarkan data pemohon pinjaman setelah dilakukan optimisasi performa model?

1.3 Batasan Masalah

Dalam penelitian ini, terdapat beberapa batasan masalah yang ditetapkan untuk menjaga fokus dan relevansi penelitian, yaitu sebagai berikut:

1. Penelitian ini hanya menggunakan dua algoritma machine learning, yaitu XGBoost dan C4.5, untuk membandingkan kinerja dalam memprediksi dengan mengklasifikasikan risiko kredit.
2. Data yang digunakan dalam penelitian ini terbatas pada dataset pemohon pinjaman di negara india, meliputi variabel-variabel seperti usia, pengalaman kerja, status perkawinan, kepemilikan rumah, penghasilan tahunan dan lain-lain.
3. Analisis menggunakan bahasa pemrograman Python dan *tool* Google Colab.
4. Kedua algoritma menggunakan split data dengan pembagian nilai *split* rasio sebesar 90:10, 80:20, 70:30, dan 60:40 .
5. Penelitian berfokus pada perbandingan pengujian kinerja model

berdasarkan akurasi, presisi, *recall*, *F1-score*, dan *confusion matrix*.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Membandingkan kinerja algoritma XGBoost dan C4.5 dalam memprediksi dari model klasifikasi risiko kredit berdasarkan data pemohon pinjaman.
2. Menganalisis akurasi, *precision*, *recall*, *F1-score* dan *confusion matrix* dari kedua algoritma dalam mengklasifikasikan risiko kredit.
3. Menentukan algoritma yang lebih efektif dan efisien dalam meningkatkan prediksi kelayakan pemohon pinjaman melalui klasifikasi risiko kredit setelah dilakukan optimisasi performa model.
4. Memberikan rekomendasi bagi lembaga keuangan terkait algoritma terbaik yang dapat digunakan untuk analisis risiko kredit berbasis data pemohon.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

A. Manfaat Teoritis:

1. Penelitian ini memberikan kontribusi pada pengembangan ilmu pengetahuan, khususnya dalam bidang data mining dan machine learning, dengan membandingkan kinerja dua algoritma populer, yaitu XGBoost dan C4.5, dalam konteks prediksi dengan mengklasifikasikan risiko kredit.
2. Hasil penelitian ini dapat digunakan sebagai referensi bagi peneliti selanjutnya yang ingin mengembangkan metode atau model lain untuk meningkatkan akurasi dalam analisis risiko kredit.

B. Manfaat Praktis:

1. Penelitian ini dapat membantu lembaga keuangan, baik perbankan maupun *fintech*, dalam memilih algoritma yang paling efektif dan efisien untuk memprediksi dengan mengklasifikasikan risiko kredit, sehingga dapat mengurangi potensi kerugian akibat debitur yang gagal membayar

pinjaman.

2. Dengan pemilihan algoritma yang tepat, lembaga keuangan dapat meningkatkan ketepatan dalam proses pengambilan keputusan terkait pemberian pinjaman, yang pada akhirnya dapat meningkatkan profitabilitas dan mengurangi risiko operasional.
3. Penelitian ini juga dapat memberikan wawasan bagi pengembang aplikasi atau sistem terkait yang membutuhkan analisis risiko kredit berbasis machine learning, sehingga dapat diimplementasikan dalam sistem penilaian kredit mereka.

1.6 Sistematika Penulisan

Agar laporan penelitian lebih terstruktur dan mudah dipahami, peneliti menyusun sistematika penulisan sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi latar belakang masalah yang mendasari penelitian, rumusan masalah yang hendak dijawab melalui penelitian, batasan masalah untuk memperjelas fokus penelitian, tujuan yang ingin dicapai, serta manfaat penelitian. Selain itu, diuraikan pula sistematika penulisan untuk memberikan gambaran umum mengenai alur penulisan skripsi ini.

BAB II TINJAUAN PUSTAKA

Bab ini mencakup studi literatur yang memuat tinjauan atas penelitian-penelitian sebelumnya terkait topik ini, serta teori-teori dasar yang digunakan sebagai landasan dalam penelitian, termasuk algoritma XGBoost dan C4.5 yang digunakan dalam prediksi dengan mengklasifikasikan risiko kredit.

BAB III METODE PENELITIAN

Bab ini menjelaskan objek penelitian, alur penelitian, serta alat dan bahan yang digunakan dalam penelitian. Proses seperti data preprocessing, pemodelan dengan algoritma XGBoost dan C4.5, dan evaluasi model juga dijelaskan secara rinci.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil penelitian, termasuk hasil prediksi dengan mengklasifikasikan data pemohon untuk risiko kredit menggunakan algoritma XGBoost dan C4.5, serta analisis perbandingan performa kedua algoritma. Pembahasan dilakukan untuk menginterpretasikan hasil yang telah diperoleh dan menilai kelayakan model dalam konteks pemberian pinjaman.

BAB V PENUTUP

Bab ini berisi kesimpulan dari hasil penelitian yang telah dilakukan, serta saran-saran yang dapat diberikan kepada pihak-pihak terkait, baik untuk pengembangan penelitian selanjutnya maupun untuk implementasi di dunia nyata, khususnya dalam analisis risiko kredit.

