

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi dan komunikasi yang pesat telah mendorong munculnya berbagai platform pinjaman online. Pinjaman online saat ini disediakan oleh berbagai lembaga, baik yang beroperasi secara legal maupun ilegal. Keabsahan pinjaman online telah diatur dengan baik oleh Otoritas Jasa Keuangan (OJK). Layanan ini memungkinkan pengguna untuk meminjam uang melalui platform keuangan yang terhubung dengan teknologi informasi, di mana seluruh proses, mulai dari pengajuan, persetujuan, hingga pencairan dana, dilakukan secara daring atau melalui konfirmasi via SMS dan telepon. Meski ilegal, pinjaman online tetap diminati masyarakat Indonesia karena persyaratannya yang sederhana, pencairan dana yang cepat, serta limit pinjaman yang besar. Pinjaman online juga memungkinkan peminjam memilih tenor sesuai kebutuhan mereka, dan kebanyakan tidak memerlukan jaminan. Berdasarkan data OJK pada tahun 2023, penyaluran dana melalui fintech lending di Indonesia mencapai Rp20,53 triliun pada Agustus 2023, mengalami peningkatan tipis sebesar 0,78% dibandingkan bulan sebelumnya yang tercatat sebesar Rp20,37 triliun.[1]

Meningkatnya jumlah peminjam juga membawa risiko yang tinggi, baik bagi peminjam yang mungkin tidak mampu membayar kembali maupun bagi penyedia layanan yang berpotensi mengalami kerugian. Oleh karena itu, penting untuk mengembangkan sistem yang dapat memprediksi kelayakan peminjam secara akurat.

Dalam konteks ini, teknik machine learning menawarkan solusi yang efektif untuk analisis data dan prediksi. Tiga algoritma yang umum digunakan dalam klasifikasi adalah Logistic Regression digunakan untuk mengklasifikasikan variabel respons kategorikal dengan memanfaatkan fungsi logistik. Algoritma ini bertujuan untuk mempelajari pola hubungan antara sejumlah variabel independen dan variabel respons. [2]. K-Nearest Neighbors (KNN) Algoritma K-Nearest Neighbor (K-NN) merupakan model klasifikasi yang menetapkan kategori kelas

untuk suatu objek baru berdasarkan kelas mayoritas dari k tetangga terdekatnya dalam kelompok data yang telah dilatih (Wijaya & Ridwan, 2019)[3] dan Random Forest algoritma (Random Forest (RF) dapat diterapkan untuk klasifikasi data tidak seimbang dalam skala besar, dengan menghasilkan performa yang baik dan waktu pemrosesan yang efisien) [4] Masing-masing algoritma ini memiliki karakteristik yang unik, seperti cara pengolahan data, kecepatan dalam pengambilan keputusan, serta akurasi prediksi. Perbandingan efektivitas dari ketiga algoritma ini menjadi penting untuk menentukan metode terbaik dalam memprediksi kelayakan pinjaman online.

Metode CRISP-DM (Cross-Industry Standard Process for Data Mining) menjadi kerangka kerja yang tepat dalam penelitian ini. CRISP-DM telah dikembangkan sejak tahun 1996 oleh tim dari NCR, SPSS, dan Daimler Chrysler. Metodologi ini dipilih karena fleksibilitasnya dalam berbagai bidang bisnis, baik untuk menyelesaikan masalah maupun mendukung kegiatan penelitian [5] CRISP-DM menyediakan pendekatan sistematis yang mencakup tahap pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penerapan. Dengan mengikuti langkah-langkah dalam CRISP-DM, peneliti dapat memastikan bahwa setiap tahap proses analisis dilakukan dengan baik dan hasil yang diperoleh dapat diandalkan.

Melalui penelitian ini, diharapkan dapat memberikan wawasan yang lebih mendalam mengenai perbandingan ketiga algoritma dalam konteks prediksi pinjaman online. Hasil dari analisis ini tidak hanya akan memberikan kontribusi signifikan dalam pengembangan model prediksi yang lebih efektif, tetapi juga menawarkan solusi yang bermanfaat bagi penyedia layanan pinjaman dalam mengelola risiko kredit. Dengan demikian, penelitian ini diharapkan dapat membantu dalam menciptakan sistem pinjaman online yang lebih aman dan efisien.

1.2 Rumusan Masalah

1. Sejauh mana efektifitas lima algoritma *machine learning* yaitu *logistic Regression*, *K-Nearest Neighbors (KNN)*, *Random Forest*, *SVM*, dan *XGBoost* dalam memprediksi kelayakan pinjam online?

2. Bagaimana cara menentukan kelayakan peminjam dalam sistem pinjaman online dengan menggunakan algoritma *machine learning*?

1.3 Batasan Masalah

1. Penelitian ini hanya membandingkan lima algoritma *mechine learning*, yaitu *logistic Reggession*, *K-Nearest Neighbors (KNN)*, *Random Forest*, *SVM*, dan *XGBoost*.
2. Peneliti ini tidak mencakup pengaruh faktor eksternal, seperti kebijakan perusahaan, peraturan pemerintah, atau perubahan ekonomi yang dapat memengaruhi kelayakan peminjam.

1.4 Tujuan Penelitian

1. Menentukan Metode Penilaian Kelayakan Peminjam dengan mengembangkan cara untuk menilai kelayakan peminjam dalam sistem pinjaman online menggunakan algoritma machine learning.
2. Menganalisis efektifitas algoritma Mechine Learning dengan mengevaluasi tingkat efektifitas tiga algoritma *mechine learning* yaitu *logistic Reggession*, *K-Nearest Neighbors (KNN)*, *Random Forest*, *SVM*, dan *XGBoost* dalam memprediksi kelayakan peminjam pada prediksi pinjaman online.

1.5 Manfaat Penelitian

1. Memberikan kontribusi ilmiah dalam pengembangan metode prediksi kelayakan peminjam menggunakan algoritma machine learning dan pendekatan CRISP-DM.
2. Penelitian ini dapat mendukung pengembangan sistem keuangan yang lebih baik dan inklusif yang memberikan pendekatan yang lebih komprehensif dan adil dalam menilai masyarakat yang kesulitan mengakses keuangan.
3. Penelitian ini dapat menjadi referensi bagi peneliti lain yang ingin melakukan penelitian serupa, baik dalam konteks analisis algoritma pembelajaran mesin maupun dalam penerapan teknik CRISP-DM pada studi kasus lainnya.

1.6 Sistematika Penulisan

BAB I: Pendahuluan

Bab ini berisi penjelasan tentang latar belakang masalah yang menjadi dasar penelitian, diikuti oleh rumusan masalah, tujuan penelitian, manfaat penelitian, serta ruang lingkupnya. Latar belakang menjelaskan pentingnya kemampuan memprediksi kelayakan pinjaman online dan bagaimana metode CRISP-DM serta algoritma machine learning yang digunakan berperan. Selain itu, bab ini menguraikan alasan penggunaan wawancara sebagai salah satu teknik pengumpulan data.

BAB II: Tinjauan Pustaka

Bab ini menyajikan teori-teori yang menjadi dasar penelitian. Pembahasannya meliputi konsep pinjaman online, penjelasan mengenai metode CRISP-DM, teori tentang *logistic Reggession*, *K-Nearest Neighbors (KNN)*, *Random Forest*, *SVM*, dan *XGBoost*, serta studi sebelumnya yang relevan. Bab ini juga memuat kajian literatur untuk memperkuat konteks penelitian, dilengkapi dengan kerangka berpikir yang menggambarkan alur logis penelitian secara keseluruhan.

BAB III: Metodologi Penelitian

Bab ini menjelaskan pendekatan yang digunakan dalam penelitian, mencakup:

1. Pendekatan Penelitian: Penjelasan tentang penggunaan metode CRISP-DM sebagai kerangka analisis utama.
2. Data dan Sumber Data: Uraian tentang jenis data yang digunakan, proses pengumpulannya serta sumber data yang relevan.
3. Tahapan Penelitian: Penjabaran langkah-langkah dalam metode CRISP - DM, seperti memahami bisnis, memahami data, persiapan data, pembuatan model, evaluasi model, hingga implementasi.
4. Teknik Analisis Data: Penjelasan proses perbandingan algoritma *logistic Reggession*, *K-Nearest Neighbors (KNN)*, *Random Forest*, *SVM*, dan *XGBoost* berdasarkan metrik evaluasi seperti akurasi, presisi, recall, dan f1-score.

BAB IV: Hasil dan Pembahasan

Bab ini menyajikan hasil penelitian beserta analisisnya. Isi bab ini meliputi:

1. Temuan dari penelitian terkait kebutuhan dan konteks bisnis dalam prediksi kelayakan pinjaman.
2. Perbandingan kinerja algoritma logistic Reggession, K-Nearest Neighbors (KNN), Random Forest, SVM, dan XGBoost menggunakan data penelitian.
3. Diskusi mengenai keunggulan dan kelemahan dari setiap algoritma dalam memprediksi kelayakan pinjaman online.
4. Analisis lebih lanjut tentang peran metode CRISP-DM dalam mempermudah dan menyusun proses penelitian ini.

BAB V: Penutup

Bab ini berisi kesimpulan dan rekomendasi. Kesimpulan menyajikan rangkuman hasil penelitian, termasuk algoritma yang paling optimal dalam konteks penelitian ini. Rekomendasi diberikan untuk peneliti masa depan dan pihak terkait, seperti perusahaan fintech, dalam mengimplementasikan hasil penelitian. Selain itu, disampaikan pula saran untuk penelitian lanjutan agar dapat dilakukan dengan cakupan dan kedalaman yang lebih baik.