

Machine Learning Analisis Risiko Kredit pada Lembaga Pembiayaan Menggunakan Algoritma K-Nearest Neighbors (KNN) dan Metode K-Means Clustering

Rizky Ramadhan^{1*)}

¹Jurusan Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur, Surabaya

¹Jl. Rungkut Madya No. 1, Kec. Gn. Anyar, Kota Surabaya, 60294, Indonesia

email: ¹20081010043@student.upnjatim.ac.id

Abstract — Bad credit often becomes a problem when a credit applicant fails to make credit installment payments on time. The liquidity of credit financing institutions can be directly affected if this condition is allowed to continue. Before extending credit, it would be nice for financing institutions to carry out a creditworthiness analysis to reduce the risk of bad credit. The process of evaluating credit worthiness sometimes takes quite a lot of time and is difficult to do. Therefore, to determine which creditors are at risk of default and which are not, a system is needed to analyze creditworthiness quickly and accurately. The development of a system based on machine learning using the k-means clustering method and the k-nearest neighbor algorithm is the goal of this research. In data mining, the k-means clustering method can provide cluster descriptions in a data. Meanwhile, the k-nearest neighbor algorithm uses weight matching on a number of existing attributes to identify the current case by comparing how close the new case is to the old case. The approach used in this study is an experimental approach, which includes literature study, modeling and system development using the k-means clustering method and the k-nearest neighbor algorithm, and system implementation. The results of this study indicate that the application of the k nearest neighbor algorithm for classification, with a k-means coefficient of 30, results in an accuracy of 91.42%.

Abstrak – Kredit macet seringkali menjadi suatu masalah ketika seorang pemohon kredit gagal melakukan pembayaran cicilan kredit tepat waktu. Likuiditas lembaga pembiayaan kredit dapat terkena dampak langsung jika kondisi ini dibiarkan terus berlangsung. Sebelum menyalurkan kredit, alangkah baiknya lembaga pembiayaan perlu melakukan analisis kelayakan kredit untuk mengurangi risiko kredit macet. Proses mengevaluasi kelayakan kredit terkadang memakan waktu yang cukup banyak dan sulit dilakukan. Oleh karena itu, untuk menentukan kreditur mana yang berisiko gagal bayar dan mana yang tidak maka diperlukan sistem untuk menganalisis kelayakan kredit yang cepat dan akurat. Pengembangan sebuah sistem yang berbasis pembelajaran mesin menggunakan metode k-means clustering dan algoritma k-nearest neighbor menjadi tujuan dari penelitian ini. Dalam data mining, metode k-means clustering dapat memberikan deskripsi cluster pada sebuah data. Sedangkan, algoritma k-nearest neighbor menggunakan pencocokan bobot pada sejumlah atribut yang ada untuk mengidentifikasi kasus saat ini dengan cara membandingkan seberapa dekat kasus baru dengan kasus lama. Adapun pendekatan digunakan dalam penelitian ini yaitu pendekatan eksperimental, yang meliputi studi pustaka, pemodelan dan pengembangan sistem yang menggunakan metode k-means

clustering dan algoritma k-nearest neighbor, dan penerapan sistem. Hasil penelitian ini menunjukkan bahwa penerapan algoritma k-nearest neighbor untuk klasifikasi, dengan nilai koefisien k-means sebanyak 30, maka diperoleh hasil akurasi sebesar 91,42%.

Kata Kunci – Pembelajaran Mesin, Sistem Pakar, Analisis Risiko Kredit, K-Nearest Neighbor, Euclidean Distance, K-Means, Clustering

I. PENDAHULUAN

Undang-Undang Perbankan No. 10 tahun 1998 mengatur tentang bank sebagai suatu badan usaha yang mengumpulkan dana dari masyarakat melalui simpanan dan menyalurkannya kembali ke masyarakat melalui pinjaman atau bentuk lainnya untuk meningkatkan kualitas hidup masyarakat. Oleh karena itu, setiap pinjaman harus didasarkan pada perjanjian yang disetujui bersama dan keputusan diambil setelah melakukan analisis. Selain pengecekan oleh BI yang merupakan tahap awal seleksi bagi setiap nasabah yang mengajukan kredit, ketidaktepatan penilaian awal sebelum menjadi nasabah kerap kali menjadi masalah kredit macet. Sebelum memberikan kredit, lembaga pembiayaan harus melakukan analisis kelayakan karena akan mempengaruhi kelancaran dalam pengembalian pinjaman, hal inilah yang nantinya dapat mengurangi risiko kredit macet. Pengambilan keputusan yang baik dalam hal memprediksi apakah pemohon kredit dapat memperoleh kredit atau tidak, menjadi perhatian penting pada penelitian ini [1].

Evaluasi kelayakan kredit sering kali merupakan proses yang rumit dan memakan waktu yang cukup lama. Sehingga, diperlukan sistem yang efisien dan tepat dalam menganalisis kelayakan kredit dan menentukan debitur yang berisiko gagal bayar dan yang tidak. Sebelumnya, banyak peneliti telah meneliti berbagai metode analisis kredit. Azhari dan Wahyu (2006), telah mengembangkan perangkat lunak agen cerdas untuk menganalisis keuangan perusahaan dalam rangka pemberian kredit perbankan [3]. Kurniawan, dkk (2022), mengembangkan aplikasi machine learning untuk strategi peningkatan kualitas perbankan dengan metode predictive analytics [4]. Gultom, dkk (2020), memprediksi tingkat kelancaran pembayaran kredit bank menggunakan algoritma naive bayes dan k-nearest neighbor [5]. Ahsina, dkk (2022), menganalisis segmentasi pelanggan bank berdasarkan

pengambilan kredit menggunakan metode k-means clustering [6].

Dalam penelitian ini, penulis akan menggunakan algoritma k-nearest neighbor dan metode k-means clustering sebagai metode untuk menentukan kelayakan kredit. Kami akan menggunakan atribut yang telah ditentukan dan hasil kedekatan (similarity) dari data kasus sebelumnya untuk memutuskan pemberian kredit. Algoritma k-nearest neighbor sendiri merupakan algoritma yang digunakan untuk menemukan solusi dari kasus saat ini dengan mencocokkan bobot dari sejumlah fitur yang ada dengan menghitung kedekatan antara kasus baru dengan kasus sebelumnya [2].

A. Rumusan Masalah

Rumusan masalah diharapkan akan menentukan instrumen yang digunakan, fokus pembahasan, kesimpulan, dan arah pelaksanaan penelitian. Berikut ini rumusan masalah yang bisa diambil:

1. Bagaimana cara melakukan analisis risiko dalam penyaluran kredit kepada pemohon kredit?
2. Bagaimana cara menentukan apakah seorang pemohon kredit layak mendapatkan pemberian kredit berdasarkan datasets yang ada ?
3. Bagaimana cara membuat sebuah program pembelajaran mesin untuk analisis risiko pemberian kredit?
4. Bagaimana cara menerapkan suatu algoritma dan metode tertentu pada sebuah program machine learning untuk analisis risiko pemberian kredit?

B. Batasan Masalah

Batasan masalah diperlukan untuk menetapkan tujuan yang ingin dicapai dalam penelitian ini. Berikut adalah batasan masalah yang dapat diberikan:

1. Sistem untuk analisis risiko kredit pada penelitian ini hanya berfokus pada penerapan algoritma k-nearest neighbor dan metode k-means clustering.
2. Pembuatan sistem ini tidak mendukung fitur input dan output dari sisi pengguna, melainkan program machine learning yang menggunakan bahasa python.
3. Datasets yang digunakan dalam penelitian ini berasal dari situs web yang menyediakan sumber data (Kaggle), bukan dari lembaga pemberian kredit tradisional.

C. Tujuan Penelitian

Tujuan penelitian adalah untuk memahami atau memprediksi situasi atau memecahkan masalah dalam kondisi yang ada dan membuktikannya dengan tindakan yang diambil. Berikut adalah tujuan penelitian yang dapat diambil:

1. Agar seorang analis kredit mampu mengambil keputusan dalam memberikan kredit kepada calon pemohon kredit dengan cepat, tepat, dan akurat.
2. Agar seorang analis kredit mengetahui, apakah seorang pemohon kredit layak atau tidak untuk diberikan pinjaman kredit berdasarkan datasets yang ada.
3. Agar terciptanya sebuah program pembelajaran mesin untuk analisis pemberian kredit yang efektif dan efisien.
4. Agar lembaga pembiayaan kredit dapat mengurangi kesalahan manusia dengan menerapkan algoritma dan metode yang ada menggunakan pembelajaran mesin.

*) **penulis korespondensi:** Rizky Ramadhan
Email: 20081010043@student.upnjatim.ac.id

II. PENELITIAN YANG TERKAIT

A. Tinjauan Pustaka

Pada penelitian ini, penulis akan mencoba untuk menghubungkan sejumlah karya ilmiah yang telah dikembangkan oleh peneliti lain yang nantinya akan berkaitan dengan karya ilmiah yang sedang penulis kembangkan. Beberapa karya ilmiah yang dimaksud penulis sebagai tinjauan pustaka antara lain sebagai berikut:

1. Jurnal Azhari dan Wahyu, yang berjudul: Pengembangan Perangkat Lunak Agen Cerdas untuk Analisis Kelayakan Keuangan Perusahaan terhadap Pemberian Kredit Perbankan. Pada jurnal ini membahas tentang pembuatan aplikasi smart berbasis agen untuk menganalisis informasi keuangan dan menilai kelayakan manajemen. Studi kelayakan menggunakan kriteria dan bobot perusahaan, dan hasilnya menunjukkan bahwa agen mampu bekerja secara independen dan berinteraksi satu sama lain melalui pengiriman pesan, tanggapan, dan meminta agen lain untuk melakukan sesuatu yang sejalan dengan tujuan utama agen menggunakan sistem cerdas untuk menganalisis kelayakan keuangan dan mengelola perusahaan. Nilai diukur menggunakan kriteria The Five C of Credit, yang terdiri dari (Character, Capacity, Capital, Collateral, dan Condition of Economy), serta kriteria analisis keuangan perusahaan seperti (rasio likuiditas, rasio operasional, rasio struktur modal, rasio kinerja bisnis, dan rasio pertumbuhan), serta analisis manajemen perusahaan [3].
2. Jurnal Kurniawan, dkk, yang berjudul: Pemilihan Metode Predictive Analytics dengan Machine Learning untuk Analisis dan Strategi Peningkatan Kualitas Perbankan. Pada jurnal ini membandingkan berbagai metode klasifikasi machine learning untuk analisis prediktif pada kualitas kredit perbankan dengan menggunakan data yang sangat besar (big data). analitik). Hasil komputasi dari metode-metode tersebut dibandingkan untuk menghasilkan rekomendasi teknik yang paling tepat, yaitu metode Tree Ensemble dengan SMOTE. Perhitungan tersebut dapat menghasilkan rekomendasi bagi bank, dengan contoh yang diberikan dalam jurnal merekomendasikan agar bank fokus pada peningkatan kualitas kredit atau pembiayaan untuk nasabah perempuan atau nasabah wiraswasta, karena berpotensi meningkatkan NPL/NPF [4].
3. Jurnal Gultom, dkk, yang berjudul: Prediksi Tingkat Kelancaran Pembayaran Kredit Bank Dengan Menggunakan Algoritma Naive Bayes dan K-Nearest Neighbor. Pada jurnal ini dilakukan dua proses pengujian. Pengujian pertama menggunakan algoritma Naive Bayes dan K-Nearest Neighbor yang diuji secara terpisah sehingga menghasilkan akurasi yang sangat rendah. Pada pengujian kedua, algoritma Naive Bayes dan K-Nearest Neighbor digabungkan dan diuji