

BAB I

PENDAHULUAN

1.1 Latar Belakang

Usaha Mikro, Kecil, Menengah (UMKM) merupakan salah satu penopang stabilitas perekonomian negara (Kementerian Koperasi dan UKM RI, 2022). Untuk meningkatkan perkembangan di sektor ini, pemerintah menerapkan program Kredit Usaha Rakyat (KUR) yang disalurkan melalui berbagai lembaga keuangan, salah satunya adalah Bank XYZ di Kabupaten Mojokerto. Program ini memberikan dampak positif bagi pelaku usaha, namun lembaga penyalur dihadapkan pada ancaman resiko finansial yang sangat tinggi, yaitu ancaman *Non-Performing Loan (NPL)* yaitu kredit bermasalah atau nasabah gagal bayar (Wijaya & Santoso, 2022). Tingkat NPL yang tidak terkendali dapat mengganggu kemudahan mencairkan uang kas bank dan menghambat penyaluran kredit pada periode berikutnya (Otoritas Jasa Keuangan, 2020).

Dalam proses pengajuan pinjaman, bank umumnya memberlakukan penilaian kelayakan atau *Credit Scoring* yaitu sistem pemodelan matematis untuk mengevaluasi kelayakan peminjam berdasarkan kondisi keuangan calon nasabah (Wati & Astuti, 2021). Variabel-variabel seperti batas pinjaman (plafon), estimasi pendapatan (omset), beban operasional, dan pengeluaran rumah tangga menjadi indikator utama dalam mengukur *Repayment Capacity* yaitu kapasitas atau kemampuan nyata peminjam untuk melunasi cicilan (Xia et al., 2020). Namun, proses penilaian yang masih mengandalkan analisis konvensional dari pihak pengamat sering kali tidak mampu menangkap pola tersembunyi dari hubungan rumit antar-variabel keuangan tersebut. Hal ini beresiko memunculkan kesalahan fatal yang berujung pada lolosnya calon peminjam beresiko tinggi.

Kemajuan teknologi yang sangat masif mendorong adopsi *Data Mining* yaitu informasi guna mengungkap pola tersembunyi serta *Machine Learning* yaitu pembelajaran mesin berbasis rekam jejak historis, sehingga

lembaga keuangan mampu mengarahkan kebijakan strategis mereka berdasarkan analisis data nyata atau *data-driven*. (Addo et al., 2018). Algoritma ini mempelajari rekam jejak nasabah masa lalu untuk memprediksi kelancaran kredit calon nasabah baru dengan jauh lebih objektif dan cepat. Salah satu algoritma klasifikasi yang terbukti ampuh menangani resiko finansial adalah *Support Vector Machine (SVM)* yaitu algoritma cerdas yang memisahkan kelas data menggunakan jarak batas yang paling optimal (Bisong, 2019).

Metode SVM sangat tangguh dalam menyelesaikan studi kasus klasifikasi yang kompleks serta sangat kebal terhadap *Overfitting* yaitu kondisi cacat di mana model terlalu kaku menghafal data latihan sehingga gagal menebak data baru (VanderPlas, 2016). Berbagai studi membuktikan bahwa pemodelan *Credit Scoring* menggunakan SVM menghasilkan kestabilan akurasi yang lebih unggul dibanding *Decision Tree* yaitu algoritma klasifikasi pohon keputusan maupun *Logistic Regression* yaitu metode analisis regresi logistik (Hidayat & Purwinarko, 2021). Tujuan matematis SVM adalah menemukan *Hyperplane* atau bidang multidimensi pemisah data yang menyajikan celah jarak paling lebar antara nasabah lancar dan macet.

Meskipun model ini sangat handal, tinjauan terhadap sejumlah penelitian perbankan menemukan adanya *Research gap* atau kesenjangan kajian yang belum terselesaikan. Banyak peneliti yang terburu-buru mengolah klasifikasi kredit menggunakan fungsi algoritma bawaan tanpa menerapkan *Hyperparameter Tuning* yaitu optimasi penyetelan parameter komputasi tingkat lanjut secara mendalam (Setiawan et al., 2020). Selain itu, evaluasi model umumnya hanya diukur melalui persentase Akurasi. Padahal, wujud asli dari data historis perbankan seperti KUR selalu berupa *Imbalanced Data* atau ketidakseimbangan volume data yang drastis, di mana populasi nasabah lancar mendominasi puluhan kali lipat dari nasabah macet (Xuan et al., 2021). Di dalam ekosistem seperti ini, penggunaan metrik Akurasi akan memicu jebakan *Accuracy Paradox* yaitu ilusi statistik

menyesatkan di mana akurasi sistem terlihat tinggi, padahal model sama sekali buta dalam mendeteksi ancaman di kelas minoritas (Trivedi, 2020).

Keberhasilan SVM bergantung pada penggunaan fungsi pemetaan yang tepat (Xia et al., 2020). Penggunaan Kernel *Linear* yang mengandalkan fungsi pemetaan garis lurus sederhana secara paksa pada karakteristik data UMKM yang cenderung kompleks dan heterogen berisiko mendongkrak tingkat kesalahan prediksi, sehingga sistem berpotensi meloloskan calon peminjam berisiko gagal bayar (Pratama & Sukmawati, 2023). Oleh karenanya, *Novelty* atau unsur kebaruan yang ditawarkan penelitian ini adalah uji perbandingan secara mendalam terhadap empat dasar pemetaan SVM yaitu *Linear*, *Radial Basis Function (RBF)*, *Polynomial*, dan *Sigmoid*. Keempatnya dioptimasi melalui algoritma *GridSearchCV* yaitu metode penelusuran kisi-kisi kombinasi parameter secara otomatis. Kinerja model akan dibedah tuntas menggunakan instrumen *Precision* atau rasio ketepatan tebakan positif, *Recall* atau tingkat kepekaan sistem melacak seluruh potensi gagal bayar, dan *F1-Score* atau nilai harmonis penyeimbang *presisi* dan *recall*) (Trivedi, 2020). Luaran dari riset ini diproyeksikan untuk diintegrasikan ke dalam arsitektur *Decision Support System (DSS)*, guna mewujudkan sistem komputasi pendukung keputusan perbankan yang andal, adaptif, serta optimal dalam memitigasi risiko kredit.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, berikut adalah rumusan masalah yang ditemui dalam penelitian ini:

- a. Bagaimanakah penerapan algoritma *Machine Learning* yaitu *Support Vector Machine (SVM)* dan penyetelan *hyperparameter* atau konfigurasi batas mesin dapat membangun model klasifikasi kelayakan kredit nasabah KUR di Bank XYZ berdasarkan variabel keuangannya?

- b. Bagaimanakah hasil perbandingan performa di antara 4 fungsi kernel SVM (*Linear*, *RBF*, *Polynomial*, *Sigmoid*) saat memproses dataset finansial yang memiliki karakteristik *Imbalanced Data* atau data tidak seimbang?
- c. Jenis fungsi kernel beserta kombinasi parameternya manakah yang paling tangguh berdasarkan pengukuran matriks evaluasi *Akurasi*, *Presisi*, *Recall*, dan *F1-Score* untuk menekan resiko kelolosan prediksi nasabah macet?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diusulkan diatas, tujuan penelitian ini adalah:

- a. Menerapkan pengklasifikasi *Support Vector Machine (SVM)* terpadu dengan algoritma penelusuran *GridSearchCV* untuk membangun sistem kelayakan kredit nasabah (KUR) Bank XYZ.
- b. Mengukur, menganalisis, serta membandingkan performa batas pemisah empat kernel SVM secara objektif menggunakan instrumen bedah *Confusion Matrix* atau disebut sebagai matriks kebingungan.
- c. Mendapatkan rancangan arsitektur fungsi dan parameter yang paling responsif menangkal resiko gagal bayar pinjaman atau disebut dengan *Non-Performing Loan (NPL)*.

1.4 Manfaat Penelitian

Penelitian ini diharapkan mampu memberikan kontribusi secara akademis dan praktis, yaitu:

- a. Manfaat Akademis

Menjadi perluasan pustaka keilmuan literatur *Data Science* atau sains data komputasi, spesifik pada komparasi optimasi kernel

algoritma berbasis margin terhadap dataset finansial mikro yang tidak seimbang.

b. Manfaat Praktis

Menjadi pondasi berbasis komputasi cerdas bagi Bank XYZ di dalam mengkaji kelayakan kredit calon peminjam, sekaligus menekan kerugian kas karena model dirancang dengan nilai ukur *sensitivitas* atau kepekaan peredam resiko yang tinggi.

1.5 Batasan Penelitian

Agar pembahasan dalam penelitian ini lebih terarah, fokus, dan tidak menyimpang dari tujuan awal, maka ditetapkan beberapa batasan masalah sebagai berikut:

- a. Ruang Lingkup Data: Dataset yang digunakan bersumber dari data historis nasabah KUR Bank XYZ wilayah Kabupaten Mojokerto pada periode 2022 hingga 2023, dengan total populasi sampel sebanyak 398 data observasi.
- b. Variabel Prediktor (Fitur): Model dilatih murni dari 9 komponen pengeluaran dan pemasukan awal *Capacity* atau kapasitas bayar yang meliputi: plafon, omset, harga pokok penjualan (HPP), tenaga kerja, sewa, listrik, pengeluaran rumah tangga, biaya lain-lain, dan tagihan lain. Variabel riwayat cicilan wajib dihilangkan dari awal untuk mencegah cacat *Data leakage* yaitu kebocoran informasi dari masa depan yang membayakan objektivitas uji coba.
- c. Variabel Target: Klasifikasi bersifat biner yang merepresentasikan status akhir kredit nasabah, yaitu Label 1 untuk nasabah "Lancar" atau sanggup bayar dan Label 0 untuk nasabah "Macet" atau gagal bayar.
- d. Algoritma dan Parameter: Eksperimen dibatasi pada bahasa pemrograman Python dengan pustaka algoritma *Scikit-Learn*. Skenario mencakup transformasi data lewat *Z-score* atau

penyetaraan satuan angka secara statistik, perlindungan lewat argumen *class_weight='balanced'* yaitu bobot penyeimbang porsi nasabah, hingga uji parameter silang C dan gamma.

- e. Evaluasi Kinerja: Penilaian wajib dibedah ke dalam matriks komponen turunan yaitu *Recall*, *Presisi*, *F1-Score* untuk menggagalkan bias *Accuracy Paradox* atau pembacaan akurasi semu.

1.6 Metodologi Penelitian

Penelitian ini menggunakan pendekatan kuantitatif eksperimental karena data yang di analisa berupa angka rekam jejak atau data finansial nasabah yang dapat diukur secara matematis, serta pengujian kinerja pemodelan dilakukan melalui perhitungan metrik statistik secara objektif (Sugiyono, 2019).

Sementara itu, jenis penelitian ini dikategorikan sebagai penelitian terapan atau disebut *Applied Research* dengan desain eksperimental. Penelitian ini membandingkan empat varian fungsi kernel yaitu *Linear*, *RBF*, *Polynomial*, dan *Sigmoid* serta melakukan pencarian konfigurasi parameter terbaik menggunakan metode *GridSearchCV* pada algoritma *Support Vector Machine (SVM)*. Eksperimen ini bertujuan untuk mengamati dan menemukan model paling optimal dalam memprediksi klasifikasi data kredit yang tidak seimbang. Kinerja dari setiap skenario eksperimen tersebut akan di evaluasi menggunakan *Confusion Matrix* melalui turunan metrik *Akurasi*, *Presisi*, *Recall*, dan *F1-Score*.

1.7 Sistematika Penulisan

Agar penjabaran laporan terstruktur dengan rapi, rincian bab disusun sebagai berikut:

BAB I PENDAHULUAN

Menjabarkan latar belakang resiko NPL KUR, ketimpangan kelas nasabah, serta mencantumkan urgensi, rumusan masalah, batasan teknis, dan target penelitian.

BAB II KAJIAN PUSTAKA

Mengkaji teori fundamental terkait kriteria kelayakan finansial, mekanisme algoritma SVM, fungsi turunan kernel, standarisasi data, dan instrumen matriks prediktif.

BAB III METODE PENELITIAN

Mendeskrripsikan rancang bangun operasional sistem, alur dari proses pembagian set data latih-uji, stratifikasi sebaran, perbandingan parameter, hingga teknik validasi.

BAB IV HASIL DAN PEMBAHASAN

Menjalankan penemuan komputasi. Melaporkan pencarian bobot optimal, menganalisis jebakan akurasi melalui tabel metrik sebaran asli, serta pembuktian arsitektur SVM penangkal kredit macet terbaik.

BAB V PENUTUP

Memberikan perumusan simpulan objektif penutup riset yang ditujukan untuk melengkapi rekomendasi ke pihak bank penyelenggara dan para akademisi di masa mendatang.