

BAB I

PENDAHULUAN

1.1 Latar Belakang

Teknologi informasi pada era sekarang berkembang dengan sangat cepat. Sistem informasi sendiri terdapat dalam berbagai aspek kehidupan, seperti bidang kesehatan, hukum, pendidikan dan keuangan. Dengan meningkatnya sistem informasi maka dapat dibuatnya sistem sebagai pertimbangan untuk pengambilan sebuah keputusan atau kesepakatan. Penggunaan kartu kredit sangat diminati oleh berbagai kalangan masyarakat di Indonesia. Peningkatan ini diperkirakan karena terdapat berbagai keuntungan dari menggunakan kartu kredit seperti mempermudah transaksi jual beli, jangka penggunaan kartu kredit yang panjang serta limit peminjaman yang besar. Dalam setiap bank atau lembaga keuangan memiliki prinsip kehati-hatian terhadap kemampuan dalam mengelola kelayakan kredit sehingga risiko kredit dapat di monitoring dengan efisien. Risiko kredit merupakan risiko dengan dampak yang berpengaruh terhadap neraca dan laporan laba rugi dari sebuah bank atau lembaga keuangan. Sehingga sebuah lembaga harus mengambil keputusan yang tepat dalam menentukan kelayakan seorang calon debitur, bila tidak maka bank atau lembaga keuangan dapat mengalami kerugian.

Salah satu data historis yang dapat digunakan untuk membantu mempertimbangkan keputusan penggunaan kartu kredit adalah data credit scoring. Data ini biasanya dimiliki oleh bank atau lembaga keuangan lainnya dan berisi data dari debitur yang pernah mengajukan kredit di bank atau lembaga keuangan

tertentu. Melalui proses ini, dapat ditentukan apakah calon debitur yang mengajukan kredit dapat diklasifikasikan sebagai calon debitur yang layak untuk diberikan pinjaman atau tidak. *Credit scoring* atau penilaian kredit merupakan sebuah sistem yang digunakan oleh bank atau lembaga keuangan untuk mengambil keputusan mengenai kelayakan seorang calon debitur untuk diterima menjadi nasabah kredit (Kristina Silalahi et al., 2017).

Metode yang bisa digunakan untuk *credit scoring* adalah *data mining*. *Data mining* merupakan pengungkapan informasi dalam kumpulan data yang banyak. *Data mining* menggunakan interaksi antara data factual, numerik serta strategi kecerdasan buatan untuk mengelola data sehingga bermanfaat dan menghasilkan informasi tertentu terkait data yang akan dikumpulkan (Herrni Yurlianti et al., 2022). Dalam melakukan proses data mining terdapat berbagai proses yaitu salah satunya adalah penggunaan algoritma klasifikasi. Metode yang pernah digunakan untuk mengklasifikasi *credit scoring* antara lain adalah K-Nearest Neighbor, Synthetic Minority Oversampling Technique (SMOTE), Extreme Gradient Boosting (XGBOOST), Random Forest, Adaptive Boosting, Particle Swarm Optimization dan lainnya.

Data *credit scoring* memiliki dimensi data yang sangat besar untuk menentukan kelayakan dari seorang calon debitur. Namun, dimensi data yang lebih sederhana relatif lebih optimal untuk menghasilkan model prediktif yang dapat menghasilkan performa lebih baik (Astoni & Haris, 2022). *Principal Component Analysis* (PCA) yang merupakan salah satu metode *feature selection* untuk digunakan untuk mengurangi dimensi fitur. Metode PCA ini dapat mengurangi

variabel dengan menganalisis korelasi yang besar menjadi data yang memiliki variabel yang lebih sedikit. Penelitian yang menggunakan PCA adalah penelitian dengan judul “Analisis Penerapan *Principal Component Analysis* (PCA) pada Deteksi Kecurangan Kartu Kredit Menggunakan *Random Forest*” oleh Khabib Astoni dan Muhammad Haris pada tahun 2022, yang menghasilkan model dengan penerapan PCA membutuhkan waktu yang lebih cepat untuk melakukan *running* dibandingkan model tanpa penerapan PCA dikarenakan PCA berhasil mereduksi jumlah fitur sehingga pemrosesan data menjadi lebih cepat.

Metode yang digunakan adalah *Bacterial Foraging Optimization* (BFO) yang merupakan algoritma yang bekerja dengan mengoptimasikan dalam mencari solusi dalam sebuah data. Algoritma ini terinspirasi dari cara kehidupan bakteri khususnya *E.coli* dalam perilaku pencarian makanan. Metode *Bacterial Foraging Optimization* ini digunakan untuk mengoptimalkan metode *Principal Component Analysis* (PCA) dengan mencari ciri-ciri terbaik pada PCA yang dapat memisahkan setiap kelas atau atribut satu dan lainnya. (Badriyah et al.,) Penelitian yang menggunakan BFO adalah penelitian dari Tessy Badriyah, Iwan Syarif dan Fitriani Rohmah Hardiyanti dengan judul “*Development of Java Library with Bacterial Foraging Optimization for Feature Selection of High-Dimensional Data*” pada tahun 2024 yang menghasilkan algoritma *Bacterial Foraging Optimization* dapat meningkatkan seleksi fitur dan mendapatkan hasil akurasi sebesar 86%. Hal ini menunjukkan algoritma *Bacterial Foraging Optimization* dapat mengoptimalkan sebuah data dengan baik.

Support vector machine (SVM) merupakan salah satu algoritma *supervised learning* yang dapat digunakan untuk klasifikasi, SVM bekerja dengan cara mencari *hyperplane* terbaik dengan margin terbesar. Margin merupakan jarak antara *hyperplane* dan pola terdekat dari setiap kelas. *Support Vector Machine* (SVM) digunakan untuk mengelola jumlah data yang besar serta dapat menghasilkan fungsi *classifier* yang optimal untuk memisahkan dua kelompok data (Sugiyarto et al., 2019). Penelitian yang menggunakan SVM adalah penelitian oleh Abdul Razak dan Akrim Teguh pada tahun 2023 dengan judul “Penerapan Fitur Seleksi dan *Particle Swarm Optimazation* pada Algoritma *Support Vector Machine* untuk Analisis Credit Scoring” yang menghasilkan akurasi sebesar 87.23%. Oleh karena itu, dapat dinyatakan SVM terbukti dapat meningkatkan hasil akurasi dalam memprediksi kelayakan pemberian kredit (Naufal & Suseno, 2023). Berdasarkan latar belakang yang telah dijelaskan, maka untuk penelitian yang akan dilakukan adalah ***Principal Component Analysis* dan *Bacterial Foraging Optimization* untuk *Credit Scoring*.**

1.2 Identifikasi Masalah

Berdasarkan latar belakang diatas, maka didapatkan rumusan permasalahan pada penelitian adalah sebagai berikut:

1. Bagaimana cara menerapkan fitur *Principal Component Analysis* pada *credit scoring*.
2. Bagaimana cara mengimplementasikan *Support Vector Machine* terhadap hasil dari fitur *Principal Component Analysis*.
3. Bagaimana hasil dari pengaruh *Bacterial Foraging Optimization* terhadap *Support Vector Machine*.

1.3 Batasan Masalah

Batasan masalah pada penelitian ini ditetapkan agar peneliti dapat fokus pada permasalahan yang ada, maka batasan masalah dalam penelitian ini sebagai berikut:

1. *Dataset* yang akan digunakan adalah dataset publik dengan nama *German.data-numeric* dari *Massachusetts Institute of Techonology*.
2. Proses seleksi fitur menggunakan *Principal Component Analysis*.
3. Proses klasifikasi menggunakan metode klasifikasi *Support Vector Machine*.
4. Algoritma untuk optimasi menggunakan *Bacterial Foraging Optimization*.
5. Implementasi akan menggunakan bahasa pemrograman *Python* dengan *Google Colab*.

1.4 Tujuan Penelitian

Penelitian ini memiliki pengaruh pada bidang keuangan atau perbankan, adapun beberapa tujuan penelitian ini adalah sebagai berikut:

1. Menerapkan fitur *Principal Component Analysis* dan optimasi *Bacterial Foraging Optimization* pada *credit scoring*.
2. Menentukan model yang memiliki performa paling baik dalam penerapan fitur *Principal Component Analysis* dan optimasi *Bacterial Foraging Optimization* pada *credit scoring*.

1.5 Manfaat Penelitian

Manfaat dalam penelitian ini adalah sebagai berikut :

1. Memberikan alternatif terhadap pilihan fitur serta metode yang dapat digunakan dalam bidang *credit scoring* dengan *data mining*.
2. Memberikan hasil penelitian dengan dataset publik sehingga dapat menambah pengetahuan serta menjadi bahan pembandingan untuk peneliti lain.
3. Memberikan hasil akurasi, presisi, *recall* dan *f-1 score* dari *credit scoring* setelah penerapan fitur *Principal Component Analysis* dan *Support Vector Machine*, dan akurasi, presisi, *recall* dan *f-1 score* dari *credit scoring* menggunakan *Principal Component Analysis* dan *Support Vector Machine* yang dioptimasi

1.6 Sistematika Penulisan

Tahapan dalam penulisan proposal tesis dibagi menjadi beberapa bab yaitu:

BAB I PENDAHULUAN

Bab ini berisi latar belakang masalah, identifikasi masalah, batasan masalah, rumusan masalah, tujuan penelitian, manfaat penelitian serta sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini berisi mengenai dasar-dasar teori, penelitian terdahulu, serta hipotesis penelitian.

BAB III METODOLOGI PENELITIAN

Bab ini akan menjelaskan tentang mengenai metode pengumpulan data serta metode penelitian yang digunakan, tahapan penelitian yang akan dilakukan oleh peneliti.

BAB IV HASIL DAN PEMBAHASAN

Bab ini akan menjelaskan tentang hasil penelitian yang telah peneliti lakukan serta penjelasan atas hasil yang didapatkan berdasarkan literatur dan dasar teori.

BAB V PENUTUP

Bab ini akan menjelaskan tentang kesimpulan serta saran dari peneliti atas penelitian yang telah dilakukan. Kesimpulan dipaparkan berdasarkan dengan tujuan penelitian serta akan diberikan saran dari peneliti.