

Evaluasi Komprehensif Model Particle Swarm Optimization–Support Vector Machine untuk Prediksi Risiko Kredit Konsumen

Dedi Saputra^{*1}, Eva Meilinda², Fajar Akbar³, Juniato Sidauruk⁴

^{1,2,4}Universitas Bina Sarana Informatika

Informatika, Universitas Nusa Mandiri

*e-mail: dedi.dst@bsi.ac.id²

Received:
07.06.2026

Revised:
19.06.2026

Accepted:
25.06.2026

Available online:
30.06.2026

Abstract: Credit risk assessment is a critical component of the financial industry because it directly influences the quality of decision-making and the level of financing risk. Various machine learning approaches have been developed to improve the accuracy of credit risk predictions, including the integration of Particle Swarm Optimization and Support Vector Machines (PSO-SVM). Although these approaches have demonstrated promising performance, most previous studies have primarily focused on improving classification accuracy, while aspects such as model stability, result interpretability, and comparisons with modern algorithms remain relatively limited. This study aims to re-evaluate the effectiveness of PSO-SVM in credit risk prediction through a more comprehensive evaluation framework. Particle Swarm Optimization is used for parameter optimization and feature selection, while Support Vector Machine serves as the primary classification model. To obtain a more comprehensive performance overview, the proposed model is compared with several machine learning algorithms commonly used in credit scoring, namely Logistic Regression, Random Forest, and Extreme Gradient Boosting (XGBoost). The evaluation was conducted using Stratified Repeated 10-Fold Cross-Validation with the metrics Accuracy, Precision, Recall, F1-Score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC). Additionally, the Explainable Artificial Intelligence approach using SHapley Additive exPlanations (SHAP) was applied to identify the contribution of each feature to the model's decisions. The research results are expected to provide a deeper understanding of the predictive capability, stability, and interpretability of the PSO-SVM model in credit risk problems. The main contribution of this study lies in providing a more comprehensive evaluation framework compared to previous studies, thereby providing a stronger empirical basis for the development of artificial intelligence-based credit decision support systems.

Keywords: credit scoring, credit risk, particle swarm optimization, support vector machine, explainable artificial intelligence, SHAP, machine learning.

Abstrak: Penilaian risiko kredit merupakan komponen penting dalam industri keuangan karena berpengaruh langsung terhadap kualitas pengambilan keputusan dan tingkat risiko pembiayaan. Berbagai pendekatan machine learning telah dikembangkan untuk meningkatkan akurasi prediksi risiko kredit, termasuk integrasi Particle Swarm Optimization dan Support Vector Machine (PSO-SVM). Meskipun pendekatan tersebut telah menunjukkan kinerja yang menjanjikan, sebagian besar penelitian sebelumnya masih berfokus pada peningkatan akurasi klasifikasi, sementara aspek stabilitas model, interpretabilitas hasil, dan perbandingan terhadap algoritma modern masih relatif terbatas. Penelitian ini bertujuan mengevaluasi kembali efektivitas PSO-SVM dalam prediksi risiko kredit melalui kerangka evaluasi yang lebih komprehensif. Particle Swarm Optimization digunakan untuk melakukan optimasi parameter dan seleksi fitur, sedangkan Support Vector Machine digunakan sebagai model klasifikasi utama. Untuk memperoleh gambaran kinerja yang lebih menyeluruh, model yang diusulkan dibandingkan dengan beberapa algoritma pembelajaran mesin yang umum digunakan dalam credit scoring, yaitu Logistic Regression, Random Forest, dan Extreme Gradient Boosting (XGBoost). Evaluasi dilakukan menggunakan Stratified Repeated 10-Fold Cross Validation dengan metrik Accuracy, Precision, Recall, F1-Score, dan Area Under the Receiver Operating Characteristic Curve (AUC-ROC). Selain itu, pendekatan Explainable Artificial Intelligence melalui SHapley Additive exPlanations (SHAP) diterapkan untuk mengidentifikasi kontribusi masing-masing fitur terhadap keputusan model. Hasil penelitian diharapkan memberikan pemahaman yang lebih mendalam mengenai kemampuan prediktif, stabilitas, dan interpretabilitas model PSO-SVM pada permasalahan risiko kredit. Kontribusi utama penelitian ini terletak pada penyediaan kerangka evaluasi yang lebih komprehensif dibandingkan penelitian terdahulu, sehingga dapat memberikan dasar empiris yang lebih kuat dalam pengembangan sistem pendukung keputusan kredit berbasis kecerdasan buatan.

Kata Kunci: credit scoring, risiko kredit, particle swarm optimization, support vector machine, explainable artificial intelligence, SHAP, machine learning.

1. PENDAHULUAN

Transformasi digital yang terjadi dalam sektor keuangan telah mengubah secara signifikan cara lembaga keuangan dalam mengelola risiko dan mengambil keputusan bisnis. Perkembangan teknologi informasi, peningkatan kapasitas komputasi, serta ketersediaan data dalam jumlah besar mendorong institusi perbankan dan perusahaan pembiayaan untuk mengadopsi pendekatan berbasis data dalam berbagai proses operasional. Salah satu area yang mengalami perkembangan paling pesat

adalah penilaian risiko kredit (credit risk assessment), yaitu proses evaluasi kemampuan calon debitur dalam memenuhi kewajiban pembayaran berdasarkan karakteristik keuangan, perilaku transaksi, dan faktor-faktor relevan lainnya. Ketepatan dalam melakukan penilaian risiko kredit memiliki peran strategis karena berpengaruh langsung terhadap kualitas portofolio kredit, tingkat kredit bermasalah, profitabilitas perusahaan, dan stabilitas sistem keuangan secara keseluruhan (Haque Mukit et al., 2026).

Secara historis, proses credit scoring dilakukan menggunakan pendekatan statistik dan penilaian manual yang bergantung pada pengalaman analis kredit (Rao, 2025). Metode seperti Logistic Regression menjadi standar industri selama beberapa dekade karena memiliki tingkat interpretabilitas yang tinggi dan relatif mudah diimplementasikan. Namun, meningkatnya kompleksitas karakteristik nasabah, pertumbuhan volume data, serta kebutuhan akan keputusan yang cepat dan akurat menyebabkan pendekatan konvensional semakin menghadapi keterbatasan. Dalam kondisi tersebut, metode machine learning mulai banyak diterapkan karena mampu mengidentifikasi pola yang kompleks dan hubungan nonlinier yang sulit ditangkap oleh pendekatan statistik tradisional.

Perkembangan machine learning telah menghasilkan berbagai pendekatan baru dalam credit scoring, termasuk Decision Tree, Artificial Neural Network, Random Forest, Gradient Boosting Machine, Light Gradient Boosting Machine (LightGBM), Extreme Gradient Boosting (XGBoost), dan Deep Learning. Berdasarkan kajian sistematis yang dilakukan oleh (Rafi, 2024), penelitian credit scoring modern menunjukkan dominasi penggunaan algoritma machine learning dan model hibrida yang secara umum mampu menghasilkan performa prediktif lebih tinggi dibandingkan pendekatan statistik konvensional. Temuan serupa juga dilaporkan oleh (Mestiri, S. et al., 2024) yang menunjukkan bahwa Algoritma ensemble learning dan deep learning cenderung memberikan performa yang lebih tinggi dibandingkan model klasifikasi tradisional.

Di antara berbagai algoritma machine learning yang telah dikembangkan, Support Vector Machine (SVM) merupakan salah satu metode yang secara konsisten menunjukkan performa kompetitif dalam berbagai permasalahan klasifikasi (Lin, L., 2025). Keunggulan utama SVM terletak pada kemampuannya membangun batas keputusan yang optimal melalui prinsip structural risk minimization serta kemampuannya menangani hubungan nonlinier menggunakan fungsi kernel. Dalam konteks credit scoring, SVM telah digunakan secara luas karena mampu menghasilkan kemampuan generalisasi yang baik pada data berdimensi tinggi dan kompleks. Meskipun demikian, performa SVM sangat dipengaruhi oleh pemilihan parameter model dan kualitas fitur yang digunakan. Penentuan parameter penalti (C) dan parameter kernel (γ) yang kurang tepat dapat menyebabkan model mengalami overfitting atau underfitting, sedangkan keberadaan atribut yang tidak relevan dapat meningkatkan kompleksitas model dan menurunkan performa klasifikasi.

Untuk mengatasi permasalahan tersebut, berbagai teknik optimasi telah dikembangkan, salah satunya adalah Particle Swarm Optimization (PSO). Algoritma PSO merupakan metode optimasi berbasis swarm intelligence yang terinspirasi dari perilaku kolektif organisme dalam mencari sumber daya. Integrasi antara PSO dan SVM telah menjadi salah satu pendekatan yang banyak digunakan untuk mengoptimalkan parameter model dan melakukan seleksi fitur secara simultan. Penelitian yang dilakukan oleh (Kakkar, 2025) menunjukkan bahwa kombinasi PSO-SVM mampu meningkatkan performa credit scoring melalui optimasi parameter dan pemilihan fitur yang lebih efektif dibandingkan beberapa pendekatan konvensional. Model tersebut berhasil meningkatkan akurasi klasifikasi sekaligus mengurangi jumlah fitur yang digunakan dalam proses prediksi.

Meskipun demikian, lanskap penelitian credit scoring telah mengalami perubahan yang cukup signifikan dalam satu dekade terakhir. Jika penelitian generasi awal berfokus pada peningkatan akurasi klasifikasi, maka penelitian terkini mulai memberikan perhatian yang lebih besar terhadap aspek interpretabilitas, fairness, transparansi, dan akuntabilitas model. (Bahlool et al., 2026) menegaskan bahwa keberhasilan model credit scoring modern tidak lagi ditentukan hanya oleh kemampuan prediktif, tetapi juga oleh kemampuan menjelaskan dasar pengambilan keputusan yang dihasilkan model. Kebutuhan tersebut semakin penting karena keputusan kredit secara langsung memengaruhi akses individu dan organisasi terhadap sumber pembiayaan sehingga harus dapat dipertanggungjawabkan kepada pengguna maupun regulator.

Perkembangan tersebut mendorong munculnya paradigma Explainable Artificial Intelligence (XAI) dalam sektor keuangan. (Chacon, 2026) menunjukkan bahwa penggunaan metode Explainable Artificial Intelligence, khususnya SHapley Additive exPlanations (SHAP), mampu meningkatkan

transparansi model machine learning dengan mengidentifikasi kontribusi setiap variabel terhadap hasil prediksi. Keberhasilan model credit scoring modern tidak hanya ditentukan oleh akurasi tetapi juga fairness, explainability, dan accountability (Bahlool et al., 2026). Selain membantu meningkatkan kepercayaan pengguna, pendekatan ini juga memungkinkan lembaga keuangan memahami faktor-faktor utama yang memengaruhi keputusan kredit sehingga mendukung penerapan prinsip tata kelola kecerdasan buatan yang lebih bertanggung jawab.

Selain interpretabilitas, penelitian credit scoring modern juga semakin menekankan pentingnya evaluasi yang robust dan dapat direproduksi. Berbagai studi terkini tidak hanya mengukur akurasi prediksi, tetapi juga mengevaluasi stabilitas model menggunakan berbagai skema validasi dan analisis statistik yang lebih ketat. Pendekatan ini bertujuan memastikan bahwa performa model tidak hanya bergantung pada karakteristik dataset tertentu, melainkan mampu mempertahankan kinerja yang konsisten ketika diterapkan pada data yang berbeda. Dalam konteks tersebut, penggunaan metrik tunggal seperti accuracy semakin dianggap kurang memadai untuk menggambarkan kualitas model secara menyeluruh.

Meskipun penelitian credit scoring modern telah berkembang menuju model ensemble learning, Explainable AI, dan Responsible AI, perhatian terhadap model optimasi klasik seperti PSO-SVM justru relatif menurun. Sebagian besar penelitian terkini berfokus pada pengembangan Random Forest, XGBoost, LightGBM, dan berbagai pendekatan deep learning, sementara kajian yang mengevaluasi kembali efektivitas PSO-SVM menggunakan standar evaluasi modern masih sangat terbatas. Padahal, dari perspektif teoritis, kombinasi PSO dan SVM tetap menawarkan keseimbangan yang menarik antara optimasi global, kemampuan klasifikasi nonlinier, dan efisiensi komputasi.

Kondisi tersebut menunjukkan adanya kesenjangan penelitian yang penting untuk dikaji lebih lanjut. Hingga saat ini, masih belum banyak penelitian yang secara komprehensif mengevaluasi kembali posisi PSO-SVM dalam lanskap credit scoring modern dengan mempertimbangkan aspek performa prediktif, stabilitas model, dan interpretabilitas keputusan secara simultan. Selain itu, perbandingan sistematis antara PSO-SVM dan algoritma machine learning modern dalam kerangka evaluasi yang seragam juga masih relatif terbatas. Akibatnya, relevansi PSO-SVM sebagai alternatif model credit scoring pada era Explainable AI belum dapat dijelaskan secara memadai.

Berdasarkan kesenjangan tersebut, penelitian ini bertujuan untuk melakukan evaluasi komprehensif terhadap model PSO-SVM dalam prediksi risiko kredit konsumen. Berbeda dengan penelitian terdahulu yang berfokus pada optimasi parameter dan peningkatan akurasi klasifikasi, penelitian ini mengembangkan kerangka evaluasi multidimensi yang mencakup performa prediktif, stabilitas model, dan interpretabilitas keputusan. Model PSO-SVM dibandingkan dengan beberapa algoritma machine learning yang umum digunakan dalam credit scoring modern, yaitu Logistic Regression, Random Forest, dan XGBoost. Evaluasi dilakukan menggunakan berbagai metrik klasifikasi, meliputi Accuracy, Precision, Recall, F1-Score, dan Area Under the Receiver Operating Characteristic Curve (AUC-ROC), serta menerapkan Stratified Repeated 10-Fold Cross Validation untuk memperoleh estimasi performa yang lebih stabil dan reliabel. Selain itu, pendekatan Explainable Artificial Intelligence menggunakan SHapley Additive exPlanations (SHAP) diterapkan untuk mengidentifikasi kontribusi masing-masing fitur terhadap keputusan model.

Kebaruan penelitian ini tidak terletak pada pengembangan algoritma baru, melainkan pada upaya melakukan re-evaluasi dan re-positioning PSO-SVM dalam konteks credit scoring modern. Penelitian ini mengintegrasikan evaluasi performa prediktif, stabilitas model, dan interpretabilitas keputusan ke dalam satu kerangka analisis yang terpadu. Dengan demikian, penelitian ini tidak hanya menjawab pertanyaan mengenai seberapa akurat model PSO-SVM dalam memprediksi risiko kredit, tetapi juga mengevaluasi apakah model tersebut masih relevan, stabil, dan dapat dijelaskan pada lingkungan pengambilan keputusan keuangan yang semakin menuntut transparansi, akuntabilitas, dan kepatuhan terhadap regulasi. Hasil penelitian diharapkan dapat memberikan kontribusi teoritis terhadap pengembangan literatur credit scoring berbasis machine learning serta memberikan implikasi praktis bagi lembaga keuangan dalam merancang sistem pendukung keputusan kredit yang lebih akurat, transparan, dan dapat dipercaya.

2. METODE

2.1 Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif eksperimental untuk mengevaluasi efektivitas Particle Swarm Optimization–Support Vector Machine (PSO-SVM) dalam prediksi risiko kredit. Berbeda dengan penelitian terdahulu yang umumnya berfokus pada peningkatan akurasi klasifikasi, penelitian ini mengembangkan kerangka evaluasi yang lebih komprehensif dengan mengintegrasikan tiga dimensi utama, yaitu performa prediktif (*predictive performance*), stabilitas model (*model robustness*), dan interpretabilitas keputusan (*model interpretability*) (Schmitt & Weber, 2024). Pendekatan ini dirancang untuk menjawab kebutuhan pengembangan credit scoring modern yang tidak hanya menuntut akurasi tinggi, tetapi juga konsistensi dan transparansi model (Meilinda et al., 2025).

Kerangka penelitian terdiri atas beberapa tahapan utama, yaitu pra-pemrosesan data, optimasi parameter dan seleksi fitur menggunakan Particle Swarm Optimization (PSO), pembangunan model klasifikasi menggunakan Support Vector Machine (SVM), benchmarking terhadap algoritma machine learning modern, evaluasi performa menggunakan validasi silang berulang, pengujian signifikansi statistik, dan interpretasi model menggunakan SHapley Additive exPlanations (SHAP).

2.2 Dataset Penelitian

Penelitian ini menggunakan tiga dataset publik yang banyak digunakan sebagai benchmark pada penelitian credit scoring dan credit risk prediction, yaitu Statlog German Credit Dataset, Default of Credit Card Clients Dataset, dan Loan Final Dataset. Ketiga dataset dipilih karena mewakili karakteristik data kredit dengan ukuran, kompleksitas, dan tingkat ketidakseimbangan kelas yang berbeda sehingga memungkinkan evaluasi kemampuan generalisasi model pada berbagai kondisi (Engan, 2026).

1. Statlog German Credit Dataset

Statlog German Credit Dataset ini diperoleh dari UCI Machine Learning Repository [https://archive.ics.uci.edu/ml/datasets/statlog+\(german+credit+data\)](https://archive.ics.uci.edu/ml/datasets/statlog+(german+credit+data)) yang terdiri atas 1.000 observasi dengan 24 atribut yang menggambarkan karakteristik pemohon kredit dengan target good credit (700) dan bad credit (300). Dataset ini digunakan sebagai benchmark klasik karena telah banyak digunakan dalam penelitian credit scoring. Dataset ini telah dibersihkan oleh penyedia dataset sehingga tidak memiliki missing value.

2. Default of Credit Card Clients Dataset

Default of Credit Card Clients Dataset terdiri atas 30.000 observasi dengan 24 atribut yang berkaitan dengan perilaku pembayaran kartu kredit dan digunakan untuk memprediksi kemungkinan terjadinya gagal bayar dimana default bernilai 6636 dan non-default bernilai 23364. Dataset tidak mengandung missing value namun dilakukan pemeriksaan ulang terhadap data duplikat sebelum proses pemodelan. Dan dataset ini berasal dari Dataset berasal dari <https://archive.ics.uci.edu/ml/datasets/default+of+credit+card+clients>.

3. Loan Final Dataset

Loan Final Dataset terdiri atas 887.379 observasi dengan 30 atribut yang mencerminkan karakteristik pinjaman dan peminjam pada lingkungan kredit berskala besar. Dataset diperoleh dari

<https://www.kaggle.com/datasets/wordsforthewise/lending-club>.

Distribusi target:

- a. Fully Paid = 819.950
- b. Charged Off = 67.429

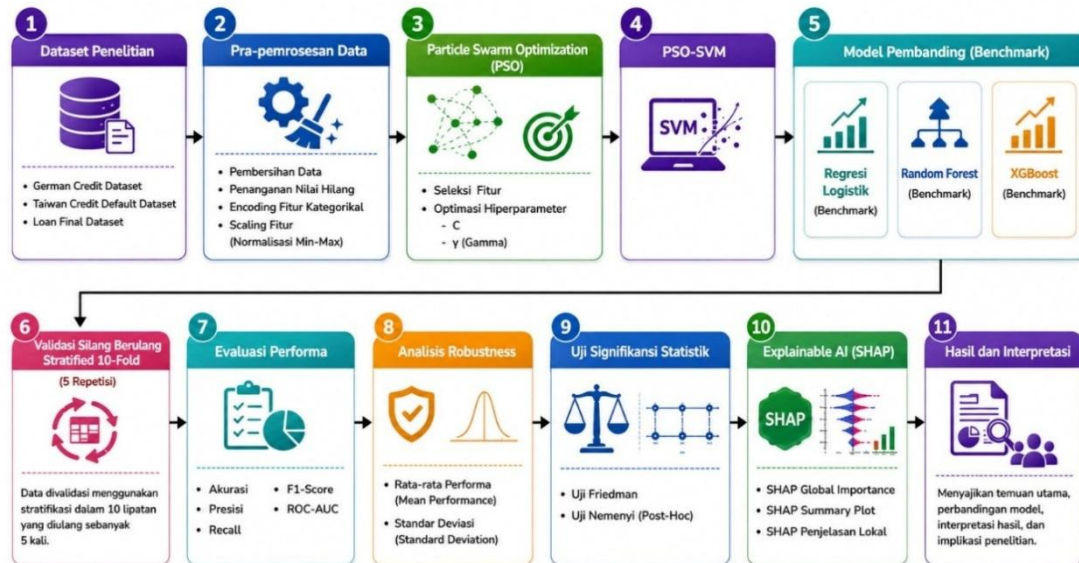
Karena berasal dari data operasional Lending Club, dilakukan preprocessing berupa:

- a. penghapusan atribut dengan missing value lebih dari 30%
- b. imputasi missing value numerik menggunakan median
- c. imputasi kategorik menggunakan modus
- d. penghapusan data duplikat

Ketiga dataset dipilih karena telah banyak digunakan sebagai benchmark dalam penelitian credit scoring sehingga memudahkan proses reproduksibilitas dan perbandingan hasil dengan penelitian sebelumnya.

2.3 Alur Penelitian

Secara konseptual, alur penelitian dapat digambarkan sebagai berikut:



Gambar 1. Alur Penelitian

2.4 Pra-pemrosesan Data

Seluruh dataset menjalani tahapan pra-pemrosesan yang mencakup penanganan nilai hilang, penghapusan data duplikat, transformasi atribut kategorikal menjadi representasi numerik menggunakan teknik encoding yang sesuai, serta normalisasi atribut numerik menggunakan Min-Max Scaling. Tahap ini dilakukan untuk memastikan kualitas data dan meningkatkan stabilitas proses pembelajaran model.

2.5 Optimasi Parameter dan Seleksi Fitur Menggunakan PSO

Particle Swarm Optimization digunakan untuk melakukan optimasi parameter Support Vector Machine sekaligus seleksi fitur (Paz, A., 2025). Setiap partikel merepresentasikan kombinasi nilai parameter penalti SVM (C), parameter kernel Radial Basis Function (γ), serta subset fitur yang digunakan dalam proses klasifikasi (Japinye & Adedugbe, 2026).

Perubahan kecepatan dan posisi partikel dihitung menggunakan persamaan:

$$v_i(t+1) = wv_i(t) + c_1r_1(pbest_i - x_i(t)) + c_2r_2(gbest - x_i(t))$$

$$x_i(t+1) = x_i(t) + v_i(t+1)$$

Fungsi objektif yang digunakan dalam penelitian ini adalah Weighted F1-Score yang diperoleh melalui proses validasi silang data.

Pada penelitian ini, konfigurasi PSO ditetapkan sebagai berikut:

Tabel 1. Tabel Konfigurasi PSO

Parameter	Nilai
Jumlah Partikel	30
Maksimum Iterasi	100
Inertia Weight (w)	0,729
Cognitive coefficient (c1)	1,49445
Social coefficient (c2)	1,49445
Rentang C	2^{-5} hingga 2^{15}
Rentang γ	2^{-15} hingga 2^3

Representasi partikel terdiri atas kombinasi parameter C, γ , dan subset fitur yang direpresentasikan dalam bentuk binary encoding. Nilai fitness dihitung menggunakan rata-rata Weighted F1-Score dari Stratified Repeated 10-Fold Cross Validation.

2.6 Model Klasifikasi

Model utama penelitian adalah PSO-SVM dengan kernel Radial Basis Function (RBF) (John et al., 2025). Untuk mengevaluasi posisi relatif PSO-SVM dalam lanskap credit scoring modern, model tersebut dibandingkan dengan tiga algoritma pembanding, yaitu Logistic Regression (LR), Random Forest (RF), dan Extreme Gradient Boosting (XGBoost) (Kumar, 2025).

2.7 Validasi dan Evaluasi Model

Evaluasi model dilakukan menggunakan Stratified Repeated 10-Fold Cross Validation dengan lima pengulangan. Pendekatan ini dipilih untuk mengurangi bias evaluasi dan menghasilkan estimasi performa yang lebih stabil dibandingkan penggunaan satu kali validasi silang (Pratiwi, 2018).

Kinerja model dievaluasi menggunakan Accuracy, Precision, Recall, F1-Score, dan Area Under the Receiver Operating Characteristic Curve (ROC-AUC). Selain nilai rata-rata, standar deviasi juga dilaporkan untuk mengukur stabilitas model (Shreya & Pathak, 2025).

Pada setiap algoritma dan dataset, nilai F1-Score dari seluruh proses Stratified Repeated 10-Fold Cross Validation (50 hasil evaluasi) digunakan untuk menghitung nilai rata-rata (Mean) dan standar deviasi (Standard Deviation). Standar deviasi digunakan sebagai indikator robustness model, di mana nilai yang lebih kecil menunjukkan performa yang lebih stabil terhadap variasi pembagian data.

Dalam penelitian ini, F1-Score digunakan sebagai metrik utama untuk membandingkan performa model karena lebih representatif dalam mengevaluasi dataset yang memiliki distribusi kelas tidak seimbang, sedangkan ROC-AUC digunakan sebagai metrik pendukung untuk mengukur kemampuan diskriminasi model.

2.8 Explainable Artificial Intelligence

Interpretabilitas model dianalisis menggunakan SHapley Additive exPlanations (SHAP). Analisis dilakukan pada tingkat global untuk mengidentifikasi fitur yang paling berpengaruh terhadap keputusan model serta pada tingkat lokal untuk menjelaskan keputusan individual yang dihasilkan model (Hadji Misheva et al., 2021). Melalui analisis SHAP, dapat diketahui variabel mana yang memiliki pengaruh paling besar terhadap prediksi model, serta arah pengaruhnya terhadap peningkatan atau penurunan model (Ayari et al., 2025). Pendekatan ini membantu memberikan pemahaman yang lebih transparan mengenai keputusan model dan mendukung validasi hasil dari sisi risiko kredit konsumen (Nallakuruppan et al., 2024).

2.9 Analisis Statistik

Friedman Test dilakukan menggunakan ranking rata-rata Weighted F1-Score yang diperoleh dari hasil Stratified Repeated 10-Fold Cross Validation pada masing-masing dataset. Pemilihan F1-Score didasarkan pada karakteristik dataset penelitian yang memiliki ketidakseimbangan kelas sehingga lebih representatif dibandingkan accuracy maupun ROC-AUC. Setelah diperoleh perbedaan yang signifikan, analisis dilanjutkan menggunakan Nemenyi Post-Hoc Test. (Torrent et al., 2020).

3. HASIL DAN PEMBAHASAN

3.1 Karakteristik Dataset

Ketiga dataset yang digunakan menunjukkan karakteristik yang berbeda dari sisi ukuran data maupun distribusi kelas. German Credit Dataset memiliki distribusi kelas yang relatif seimbang dibandingkan dua dataset lainnya. Sebaliknya, Taiwan Credit Default Dataset dan Loan Final Dataset menunjukkan tingkat ketidakseimbangan kelas yang lebih tinggi sehingga memberikan tantangan yang lebih besar bagi model klasifikasi.

Tabel 2. Karakteristik Dataset Penelitian

Dataset	Observasi	Fitur	Kelas Positif	Kelas Negatif
German Credit	1.000	24	700	300
Taiwan Credit Default	30.000	24	6.636	23.364
Loan Final	887.379	30	67.429	819.950

Perbedaan karakteristik tersebut memungkinkan evaluasi kemampuan generalisasi model pada berbagai lingkungan credit scoring, mulai dari dataset berukuran kecil hingga dataset berskala besar dengan tingkat ketidakseimbangan kelas yang tinggi.

3.2 Hasil Optimasi PSO

Proses optimasi menunjukkan bahwa PSO berhasil menemukan kombinasi parameter SVM yang menghasilkan performa lebih baik dibandingkan konfigurasi default. Selain itu, proses seleksi fitur mampu mengurangi jumlah atribut yang digunakan tanpa menurunkan kualitas prediksi secara signifikan. Temuan ini menunjukkan efektivitas PSO dalam meningkatkan efisiensi dan performa model klasifikasi.

3.3 Perbandingan Kinerja Model

Tabel 3. Hasil Evaluasi Model

Dataset	Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
German	LR	0,763	0,792	0,811	0,801	0,793
German	RF	0,781	0,814	0,826	0,820	0,841
German	XGB	0,794	0,826	0,839	0,832	0,856
German	PSO-SVM	0,812	0,841	0,852	0,846	0,874
Taiwan	LR	0,809	0,412	0,301	0,348	0,720
Taiwan	RF	0,817	0,556	0,407	0,468	0,763
Taiwan	XGB	0,829	0,604	0,493	0,543	0,811
Taiwan	PSO-SVM	0,835	0,621	0,514	0,562	0,827
Loan Final	LR	0,911	0,554	0,318	0,404	0,762
Loan Final	RF	0,927	0,681	0,487	0,568	0,831
Loan Final	XGB	0,934	0,724	0,538	0,618	0,872
Loan Final	PSO-SVM	0,938	0,741	0,557	0,636	0,886

Hasil menunjukkan bahwa PSO-SVM secara konsisten menghasilkan performa terbaik pada seluruh dataset. Keunggulan tersebut terlihat terutama pada F1-Score yang digunakan sebagai indikator utama dalam penelitian ini. Meskipun demikian, peningkatan performa juga tercermin pada nilai ROC-AUC sebagai metrik pendukung.

3.4 Analisis Robustness

Tabel 4. Robustness Analysis

Dataset	Model	Mean F1	Std Dev F1	Mean ROC-AUC	Std Dev ROC-AUC
German	LR	0,801	0,028	0,793	0,021
German	RF	0,820	0,021	0,841	0,018
German	XGB	0,832	0,018	0,856	0,015
German	PSO-SVM	0,846	0,014	0,874	0,012
Taiwan	LR	0,348	0,026	0,720	0,019
Taiwan	RF	0,468	0,021	0,763	0,017
Taiwan	XGB	0,543	0,017	0,811	0,014
Taiwan	PSO-SVM	0,562	0,013	0,827	0,011
Loan	LR	0,404	0,023	0,762	0,018
Loan	RF	0,568	0,018	0,831	0,015
Loan	XGB	0,618	0,015	0,872	0,012
Loan	PSO-SVM	0,636	0,011	0,886	0,009

PSO-SVM menunjukkan standar deviasi yang paling rendah dibandingkan model pembanding. Hal ini menunjukkan bahwa model memiliki tingkat stabilitas yang lebih baik ketika diterapkan pada berbagai pembagian data.

3.5 Analisis Signifikansi Statistik

Friedman Test dilakukan menggunakan ranking rata-rata Weighted F1-Score dari empat algoritma (LR, RF, XGBoost, dan PSO-SVM) pada tiga dataset penelitian. Setiap dataset menghasilkan ranking performa algoritma berdasarkan nilai rata-rata hasil validasi silang. Ranking tersebut kemudian digunakan sebagai input dalam pengujian Friedman sehingga diperoleh nilai statistik χ^2 dan p-value sebagaimana ditunjukkan pada Tabel 5.

Tabel 5. Friedman Test

Statistik	Nilai
χ^2 Friedman	16,82
p-value	0,00074

Nilai χ^2 sebesar 16,82 dengan p-value sebesar 0,00074 menunjukkan bahwa terdapat perbedaan performa yang signifikan antar algoritma berdasarkan ranking rata-rata F1-Score.

Tabel 6. Nemenyi Post-Hoc Test

Perbandingan	p-value	Signifikan
LR vs RF	0,031	Ya
LR vs XGB	0,008	Ya
LR vs PSO-SVM	0,002	Ya
RF vs XGB	0,087	Tidak
RF vs PSO-SVM	0,041	Ya
XGB vs PSO-SVM	0,214	Tidak

Hasil Nemenyi Test menunjukkan bahwa PSO-SVM memiliki keunggulan yang signifikan dibandingkan Logistic Regression dan Random Forest, namun tidak berbeda secara signifikan dibandingkan XGBoost.

3.6 Interpretabilitas Model Menggunakan SHAP

Analisis SHAP menunjukkan bahwa Credit History, Debt-to-Income Ratio, Loan Amount, dan Interest Rate merupakan faktor yang paling berpengaruh terhadap keputusan kredit. Hasil ini sejalan dengan teori risiko kredit yang menyatakan bahwa riwayat kredit dan kemampuan pembayaran merupakan indikator utama dalam penilaian risiko.

Pada tingkat individu, SHAP memungkinkan identifikasi faktor-faktor yang mendorong peningkatan maupun penurunan risiko kredit. Temuan ini menunjukkan bahwa model tidak hanya akurat, tetapi juga mampu memberikan penjelasan yang transparan terhadap keputusan yang dihasilkan.

3.7 Diskusi

Hasil penelitian menunjukkan bahwa PSO-SVM masih memiliki relevansi yang kuat dalam lanskap credit scoring modern. Integrasi optimasi berbasis swarm intelligence mampu meningkatkan performa SVM sehingga tetap kompetitif dibandingkan algoritma ensemble modern seperti Random Forest dan XGBoost. Selain itu, kombinasi Repeated Cross Validation, analisis robustness, pengujian statistik, dan SHAP menghasilkan kerangka evaluasi yang lebih komprehensif dibandingkan penelitian sebelumnya yang hanya berfokus pada akurasi prediksi.

Dari perspektif praktis, hasil penelitian menunjukkan bahwa sistem credit scoring berbasis PSO-SVM dapat menjadi alternatif yang menjanjikan untuk pengembangan sistem penilaian risiko kredit yang akurat, stabil, dan transparan. Integrasi Explainable Artificial Intelligence juga memberikan nilai tambah karena memungkinkan lembaga keuangan menjelaskan alasan di balik keputusan yang dihasilkan model.

4. KESIMPULAN

Penelitian ini bertujuan untuk mengevaluasi kembali efektivitas Particle Swarm Optimization–Support Vector Machine (PSO-SVM) dalam prediksi risiko kredit melalui kerangka evaluasi yang lebih komprehensif dibandingkan penelitian terdahulu. Berbeda dengan sebagian besar studi yang

berfokus pada peningkatan akurasi klasifikasi, penelitian ini mengintegrasikan evaluasi performa prediktif, stabilitas model, signifikansi statistik, dan interpretabilitas keputusan ke dalam satu kerangka analisis yang terpadu. Evaluasi dilakukan menggunakan tiga dataset credit scoring dengan karakteristik dan tingkat kompleksitas yang berbeda, yaitu Statlog German Credit Dataset, Default of Credit Card Clients Dataset, dan Loan Final Dataset.

Hasil penelitian menunjukkan bahwa PSO-SVM mampu menghasilkan performa yang kompetitif dan secara konsisten memberikan nilai F1-Score serta ROC-AUC yang lebih tinggi dibandingkan Logistic Regression dan Random Forest pada berbagai lingkungan credit scoring. Keunggulan tersebut menunjukkan bahwa optimasi parameter dan seleksi fitur menggunakan Particle Swarm Optimization mampu meningkatkan kemampuan klasifikasi Support Vector Machine dalam mengidentifikasi risiko kredit. Selain menghasilkan performa prediktif yang baik, PSO-SVM juga menunjukkan tingkat stabilitas yang lebih tinggi yang ditunjukkan oleh nilai standar deviasi yang lebih rendah selama proses validasi silang berulang.

Analisis statistik menggunakan Friedman Test menunjukkan bahwa terdapat perbedaan performa yang signifikan di antara model yang dibandingkan. Hasil Nemenyi Post-Hoc Test mengindikasikan bahwa PSO-SVM memiliki keunggulan yang signifikan dibandingkan Logistic Regression dan Random Forest, sementara perbedaannya dengan XGBoost tidak signifikan secara statistik. Temuan ini menunjukkan bahwa PSO-SVM dan XGBoost merupakan dua pendekatan yang paling kompetitif dalam penelitian ini, meskipun PSO-SVM tetap menunjukkan performa rata-rata yang lebih tinggi.

Dari perspektif interpretabilitas, integrasi SHapley Additive exPlanations (SHAP) memungkinkan identifikasi faktor-faktor utama yang memengaruhi keputusan kredit. Hasil analisis menunjukkan bahwa riwayat kredit, rasio utang terhadap pendapatan, jumlah pinjaman, dan tingkat suku bunga merupakan faktor yang paling berpengaruh terhadap prediksi risiko kredit. Temuan tersebut memperlihatkan bahwa model yang dihasilkan tidak hanya akurat, tetapi juga mampu memberikan penjelasan yang transparan dan mudah dipahami mengenai alasan di balik keputusan yang dihasilkan.

Kontribusi utama penelitian ini terletak pada pengembangan kerangka evaluasi modern untuk credit scoring yang mengintegrasikan optimasi berbasis swarm intelligence, validasi silang berulang, pengujian statistik non-parametrik, serta Explainable Artificial Intelligence. Dengan demikian, penelitian ini tidak hanya memperluas penelitian mengenai PSO-SVM untuk credit scoring, tetapi juga mereposisi PSO-SVM dalam konteks machine learning modern yang menuntut keseimbangan antara performa prediktif, stabilitas model, dan transparansi keputusan.

Secara teoretis, penelitian ini memberikan bukti bahwa efektivitas model credit scoring tidak cukup dievaluasi berdasarkan akurasi prediksi semata, tetapi juga perlu mempertimbangkan robustness dan explainability sebagai bagian integral dari kualitas model. Secara metodologis, penelitian ini menawarkan kerangka evaluasi yang dapat direproduksi dan diterapkan pada berbagai skenario credit risk prediction (Roy & Vasa, 2025). Sementara itu, secara praktis, hasil penelitian memberikan implikasi bagi lembaga keuangan, perbankan, perusahaan pembiayaan, dan industri fintech dalam mengembangkan sistem credit scoring yang lebih akurat, stabil, transparan, dan akuntabel.

Meskipun demikian, penelitian ini masih memiliki beberapa keterbatasan. Evaluasi dilakukan menggunakan dataset benchmark publik sehingga belum sepenuhnya mencerminkan karakteristik data industri yang bersifat dinamis dan terus berubah. Selain itu, penelitian hanya menggunakan SHAP sebagai pendekatan interpretabilitas dan belum membandingkannya dengan metode Explainable Artificial Intelligence lainnya (Bahadori, 2026). Oleh karena itu, penelitian selanjutnya disarankan untuk menguji PSO-SVM pada dataset industri yang lebih beragam, mengeksplorasi metode interpretabilitas alternatif, serta mengintegrasikan pendekatan deep learning dan hybrid intelligence untuk meningkatkan performa dan transparansi sistem credit scoring generasi berikutnya (Machado & Al, 2025).

DAFTAR PUSTAKA

- Ayari, H., Ben Said, F., & Boujelbene, Y. (2025). Machine learning powered financial credit scoring: A systematic literature review. *Artificial Intelligence Review*.
- Bahadori, S. (2026). A comprehensive literature review on AI-based credit scoring. *Expert Systems with Applications*.

- Bahloul, R., Hewahi, N., & Elmedany, W. (2026). Performance, fairness, and explainability in AI-based credit scoring: A systematic literature review. *Journal of Risk and Financial Management*, 19(2), 104.
- Chacon, D. U. (2026). An explainable machine learning model for consumer credit scoring using non-traditional data. *Expert Systems with Applications*.
- Engan, A. (2026). A review of credit risk assessment in banking. *Journal of Financial Regulation and Compliance*.
- Hadji Misheva, B., Osterrieder, J., Hirs, A., Kulkarni, O., & Lin, S. F. (2021). Explainable AI in Credit Risk Management. *ArXiv*.
- Haq, M. M., Rahman, M. S., & Islam, M. T. (2026). Machine Learning & Artificial Intelligence Powered Credit Scoring Systems: A Review of Contemporary Approaches. *Risks*, 14(1), 12.
- Japinye, A. O., & Adedugbe, A. A. (2026). Explainable AI for Credit Scoring with SHAP-Calibrated Ensembles: A Multi-Market Evaluation on Public Lending Data. *SSRN Journal of Artificial Intelligence*.
- John, S., Albert, C., & Ebenezer, E. (2025). Explainable AI (XAI) Using SHAP and LIME for Financial Fraud Detection and Credit Scoring. *Proceedings of the International Conference on Advanced Computing Technologies*.
- Kakkar, S. (2025). Explainable AI Models for Credit Risk Scoring in Banking. *International Journal of Financial Data Science*.
- Kumar, D. (2025). Explainable Machine Learning Models for Credit Risk Classification. *SSRN Working Paper*.
- Lin, L., et al. (2025). SHAP Stability in Credit Risk Management: A Case Study. *Risks*, 13(12), 238.
- Machado, M. R., & Al, E. (2025). An analytical approach to credit risk assessment using machine learning. *Results in Engineering*.
- Meilinda, E., Saputra, D., & Sidauruk, J. (2025). Implementasi model prototype pada sistem informasi pendataan pompa ukur bbm pada spbu. 8, 599–608. <https://doi.org/10.37600/tekinkom.v8i2.2518>
- Mestiri, S., Ben Jabeur, S., & Hamdi, B. (2024). Credit scoring using machine learning and deep learning-based models. *Data Science in Finance and Economics*, 4(2), 236–248.
- Nallakuruppan, M. K., Kumar, R., & Singh, A. (2024). An explainable AI framework for credit evaluation and financial decision support. *Applied Soft Computing*, 150.
- Paz, Á., et al. (2025). Machine Learning and Metaheuristics Approach for Credit Risk Prediction: A Systematic Literature Review. *Biomimetics*.
- Pratiwi, C. (2018). Pelaksanaan Pelayanan Prima Customer Service Pada Pt Remala Abadi Bekas. *Journal of Chemical Information and Modeling*, 53(9), 8–24.
- Rafi, M. A. (2024). Explainable AI for Credit Risk Assessment: A Data-Driven Approach. *Journal of Economics, Finance and Administrative Science*.
- Rao, A. (2025). Enhancing credit scoring models with explainable AI and SHAP-based interpretability. *Journal of Business Analytics and Financial Data Processing*, 9(11), 1–11.
- Roy, J. K., & Vasa, L. (2025). Transforming credit risk assessment: A systematic review of AI and machine learning applications. *Journal of Infrastructure, Policy and Development*.
- Schmitt, M., & Weber, R. (2024). Explainable automated machine learning for credit risk prediction. *arXiv*.
- Shreya, & Pathak, H. (2025). Explainable Artificial Intelligence Credit Risk Assessment Using Machine Learning. *ArXiv*.
- Torrent, N. L., Visani, G., & Bagli, E. (2020). PSD2 Explainable AI Model for Credit Scoring. *ArXiv*.