

MODEL MACHINE LEARNING UNTUK DETEKSI DINI KETERLAMBATAN BAYAR KREDIT DI PT BANK X

Iqbal Ali Mansyah Wardana

Statistika Bisnis, Fakultas Vokasi, Institut Teknologi Sepuluh Nopember, Indonesia

e-mail: bgiqbalwardana18@gmail.com*

Destri Susilaningrum

Statistika Bisnis, Fakultas Vokasi, Institut Teknologi Sepuluh Nopember, Indonesia

e-mail: destr_s@gmail.com

Abstrak

Risiko kredit merupakan salah satu tantangan utama dalam aktivitas perbankan, terutama dalam menilai kelayakan nasabah kredit yang tergolong *non-performing loan* (NPL) atau memiliki potensi gagal bayar sehingga menjadi indikator awal dari potensi gagal bayar di masa depan. Permasalahan muncul karena sistem klasifikasi tradisional dalam perbankan sering kali tidak mampu mengidentifikasi nasabah yang secara nominal masih lancar membayar, tetapi sebenarnya memiliki risiko tinggi gagal bayar. Akibatnya, bank mungkin salah mengkategorikan kualitas kredit, yang berdampak pada ketidaktepatan penghitungan Cadangan Kerugian Penurunan Nilai (CKPN). Hal ini dapat meningkatkan beban operasional dan risiko finansial bagi bank. Oleh karena itu, penting bagi bank untuk memiliki sistem *early warning system* yang dapat mendeteksi risiko keterlambatan pembayaran sedini mungkin. Penelitian ini menggunakan *machine learning* dengan metode *Support Vector Machine* (SVM) dengan berbagai jenis kernel—linear, RBF, polynomial, dan sigmoid—untuk membangun model klasifikasi risiko keterlambatan pembayaran kredit. Data yang digunakan mencerminkan keberagaman profil nasabah, dengan rata-rata skor kredit 650,53 dan mayoritas berstatus kredit lancar (79,63%). Hasil menunjukkan bahwa kernel RBF memberikan performa terbaik, dengan akurasi 85,45%, sensitivitas 83,83%, dan spesifisitas 90,58%. Model ini diharapkan dapat membantu bank memitigasi risiko kredit secara lebih akurat dan proaktif, sehingga kualitas aset tetap terjaga dan beban CKPN dapat dikendalikan dengan lebih baik.

Kata Kunci: Kredit, *Machine Learning*, *Non-Performing Loan*, Risiko, SVM.

Abstract

Credit risk is one of the main challenges in banking activities, especially in assessing the creditworthiness of customers who are classified as *non-performing loans* (NPL) or have the potential for default, so that they become an early indicator of the potential for default in the future. The problem arises because the traditional classification system in banking is often unable to identify customers who are nominally still paying smoothly, but actually have a high risk of default. As a result, banks may miscategorize credit quality, which has an impact on the inaccuracy of calculating the Allowance for Impairment Losses (CKPN). This can increase the operational burden and financial risk for banks. Therefore, it is important for banks to have an *early warning system* that can detect the risk of late payments as early as possible. This study uses machine learning with the *Support Vector Machine* (SVM) method with various types of kernels—linear, RBF, polynomial, and sigmoid—to build a classification model for the risk of late credit payments. The data used reflects the diversity of customer profiles, with an average credit score of 650.53 and the majority having a current credit status of 79.63%. The results show that the RBF kernel provides the best performance, with an accuracy of 85.45%, sensitivity of 83.83%, and specificity of 90.58%. This model is expected to help banks mitigate credit risk more accurately and proactively, so that asset quality is maintained and CKPN burden can be better controlled.

Keywords: Credit, *Machine Learning*, *Non-Performing Loan*, Risk, SVM.

PENDAHULUAN

Perbankan memiliki peran vital dalam mendukung pertumbuhan ekonomi melalui fungsi intermediasi, yaitu menyalurkan dana dari pihak yang surplus kepada pihak yang defisit. Salah satu bentuk intermediasi tersebut adalah pemberian kredit kepada nasabah. Namun, pemberian kredit tidak

terlepas dari risiko, khususnya risiko kredit (*credit risk*) yang muncul akibat ketidakmampuan nasabah dalam memenuhi kewajiban pembayaran sesuai perjanjian (Basel Committee on Banking Supervision, 2019).

Dalam praktiknya, tidak semua nasabah yang masuk kategori *performing loan* (kredit lancar) memiliki kedisiplinan yang konsisten dalam

membayar cicilan. Sebagian dari mereka mengalami keterlambatan pembayaran (*delinquency*) meskipun belum memenuhi syarat sebagai kredit bermasalah (*non-performing loan*). Fenomena ini sangat penting untuk diperhatikan karena keterlambatan pembayaran berulang sering kali menjadi indikator awal dari potensi gagal bayar (*default*) di masa depan. Hal ini sesuai dengan hasil penelitian oleh (Yuniarti et al., 2022) yang menyatakan bahwa keterlambatan pembayaran pada kredit lancar dapat menjadi indikator dini atas risiko *default*.

PT Bank X sebagai lembaga keuangan nasional yang aktif dalam menyalurkan kredit, menghadapi tantangan yang sama. Sebagian nasabah kredit yang masih berstatus lancar namun memiliki pola keterlambatan berulang, dapat menyebabkan distorsi dalam penilaian kualitas kredit. Ini berdampak langsung terhadap peningkatan beban pencadangan (*impairment loss*) atau CKPN (Cadangan Kerugian Penurunan Nilai), serta peningkatan beban operasional akibat intensifikasi penagihan. Hal ini menuntut bank untuk mengelola risiko kredit secara lebih komprehensif dan berbasis data, sebagaimana diatur dalam Pedoman Manajemen Risiko Kredit oleh Otoritas Jasa Keuangan (Otoritas Jasa Keuangan (OJK), 2021). Kondisi ini menjadi tantangan tersendiri dalam pengelolaan risiko kredit karena segmentasi nasabah berdasarkan status *performing* dan *non-performing* tidak cukup untuk merepresentasikan risiko sesungguhnya. Oleh karena itu, diperlukan pendekatan deteksi dini untuk mengidentifikasi nasabah yang berisiko tinggi mengalami keterlambatan, bahkan sebelum masuk ke kategori NPL. Banyak faktor yang diduga memengaruhi keterlambatan pembayaran, seperti usia, jenis kelamin, lama menjadi nasabah (*tenure*), penghasilan, jenis pekerjaan, skor kredit nasabah, dan sebagainya. Penelitian oleh (Djuarni & Ratnasari, 2022) menunjukkan bahwa karakteristik demografis dan historis kredit berperan signifikan terhadap kemungkinan nasabah mengalami keterlambatan pembayaran.

Di era digital, ketersediaan data historis dalam jumlah besar memberikan peluang untuk mengembangkan sistem prediktif yang lebih akurat dan adaptif. Salah satu pendekatan yang efektif untuk mengatasi tantangan tersebut adalah melalui *machine learning*. Metode ini memungkinkan sistem untuk belajar dari data historis dan mengenali pola kompleks yang tidak mudah terlihat melalui analisis statistik konvensional. Salah satu algoritma *machine learning* yang banyak digunakan untuk klasifikasi

risiko adalah *Support Vector Machine* (SVM). SVM bekerja dengan membentuk *hyperplane* optimal yang memisahkan dua kelas kategori dan sangat efektif untuk menangani data berdimensi tinggi atau data yang tidak linier. Menurut (Vapnik, 1998), SVM dikembangkan dengan dasar teori pembelajaran statistik yang memungkinkan prediksi lebih akurat dalam kasus klasifikasi dua kelas dan minimisasi kesalahan generalisasi. Penelitian oleh (Syafi'i et al., 2022) membuktikan bahwa SVM dapat mencapai tingkat akurasi yang tinggi dalam mendeteksi nasabah kredit yang berisiko telat bayar. Demikian pula, (Pernama et al., 2023) menemukan bahwa SVM lebih stabil dibandingkan metode klasifikasi lain seperti *Artificial Neural Network* atau *Naïve Bayes* dalam konteks prediksi risiko kredit pada data perbankan.

Dengan kemampuan generalisasi yang baik dan minim *overfitting*, SVM menjadi pilihan yang tepat untuk membangun model prediksi risiko keterlambatan pembayaran. Penggunaan algoritma ini diharapkan dapat membantu PT Bank X mengembangkan sistem *early warning* yang dapat memitigasi risiko secara proaktif.

KAJIAN TEORI

Machine Learning

Machine learning (ML) adalah cabang dari kecerdasan buatan (*artificial intelligence*) yang memungkinkan sistem komputer belajar secara otomatis dari data dan meningkatkan kinerjanya tanpa harus diprogram secara eksplisit. Algoritma *machine learning* dirancang untuk mengenali pola dalam data dan kemudian membuat prediksi atau pengambilan keputusan berdasarkan pola tersebut (Mitchell, 1997). Dalam konteks keuangan dan perbankan, *machine learning* banyak dimanfaatkan untuk mendeteksi penipuan, melakukan segmentasi nasabah, serta memprediksi risiko kredit, termasuk keterlambatan pembayaran atau *default*.

Machine learning terbagi ke dalam beberapa pendekatan utama, di antaranya *supervised learning*, *unsupervised learning*, dan *reinforcement learning*. *Supervised learning*, yang digunakan dalam penelitian ini, melibatkan pelatihan model menggunakan dataset yang sudah diberi label, seperti klasifikasi antara nasabah lancar dan bermasalah. Keunggulan *machine learning* terletak pada kemampuannya dalam menangani data besar dan kompleks serta mendeteksi hubungan non-linear yang tidak selalu mudah ditemukan dengan metode statistik tradisional (Kou et al., 2021).

Support Vector Machine

Support Vector Machine (SVM) adalah salah satu algoritma *supervised learning* yang digunakan untuk menyelesaikan masalah klasifikasi dan regresi. SVM bekerja dengan cara mencari hyperplane terbaik yang dapat memisahkan data ke dalam dua kelas berbeda dengan margin maksimum (*maximum margin hyperplane*) (Cristianini & Shawe-Taylor, 2000). Konsep dasar dari SVM adalah memilih garis pemisah (hyperplane) yang tidak hanya memisahkan kelas dengan benar, tetapi juga memaksimalkan jarak antar titik data dari dua kelas terhadap garis tersebut (Cortes & Vapnik, 1995). Di era saat ini, banyak permasalahan yang tidak dapat diselesaikan secara linier karena kompleksitas data yang tinggi dan pola hubungan antar variabel yang bersifat non-linier. Suatu metode untuk mentransformasikan data non-linier ke dalam ruang berdimensi yang lebih tinggi agar dapat dipisahkan secara linier. Proses ini dikenal dengan istilah kernelisasi. Kernel merupakan fungsi yang digunakan untuk mengubah representasi data dari fitur awal (*low-level feature*) menjadi fitur baru (*high-level feature*) tanpa perlu menghitung secara eksplisit koordinat di ruang berdimensi tinggi tersebut (Burges, 1998). variasi kernel yang sering digunakan untuk pemecahan masalah secara non-linier dengan persamaan sebagai berikut.

- Kernel Linear

$$K(x_i, x_j) = x_i^T x_j \quad (1)$$

- Kernel *Radial Basis Function* (RBF)

$$K(x_i, x_j) = \exp\left(\frac{-||x-y||^2}{2\sigma^2}\right) \quad (2)$$

- Kernel *Polynomial*

$$K(x_i, x_j) = (-\gamma x_i^T x_j + r)^d \quad (3)$$

- Kernel *Sigmoid*

$$K(x_i, x_j) = \tanh(-\gamma \cdot x_i^T x_j + r) \quad (4)$$

Dalam algoritma Support Vector Machine (SVM), misalkan x dan y adalah pasangan dua data dari seluruh bagian data pelatihan. Parameter σ (sigma) digunakan dalam kernel *Radial Basis Function* (RBF), C (*complexity*) adalah parameter regularisasi yang mengontrol keseimbangan antara margin maksimum dan kesalahan klasifikasi, serta D (degree) digunakan dalam kernel polinomial. Selain itu, $||x - y||^2$ menunjukkan jarak kuadrat (Euclidean) antara dua vektor fitur (Vapnik, 1995). Untuk membentuk hyperplane pemisah, digunakan bentuk umum persamaan.

$$f(x) = (w, x) + b = 0 \quad (5)$$

di mana:

w adalah vektor bobot yang tegak lurus terhadap hyperplane,

x adalah vektor data, (w, x) adalah hasil dot product antara w dan x , b adalah bias atau offset dari hyperplane terhadap titik asal. Nilai w dan b harus ditentukan melalui proses optimisasi (Schölkopf & Smola, 2002). Untuk itu digunakan konsep hyperplane kanonik, yang mengasumsikan bahwa:

- Titik data dengan label $y=+1$ yang paling dekat ke hyperplane berada pada:

$$(w, x^+) + b = +1 \quad (6)$$

- Titik data dengan label $y=-1$ yang paling dekat ke hyperplane berada pada:

$$(w, x^-) + b = -1 \quad (7)$$

Sehingga jarak margin antar kedua kelas tersebut adalah:

$$\text{Margin} = \frac{2}{||w||} \quad (8)$$

Model SVM berusaha meminimalkan $||w||^2$ untuk mendapatkan margin maksimum, sekaligus memastikan klasifikasi yang tepat terhadap data pelatihan dengan atau tanpa kesalahan.

Kredit

Kredit merupakan salah satu kegiatan utama dalam dunia perbankan yang berfungsi untuk menyalurkan dana kepada pihak yang membutuhkan. Menurut Undang-Undang No. 10 Tahun 1998 tentang Perbankan, kredit adalah penyediaan uang atau tagihan yang dapat disamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam-meminjam antara bank dengan pihak lain yang mewajibkan pihak peminjam untuk melunasi utangnya setelah jangka waktu tertentu dengan imbalan bunga.

Performing Loan

Performing loan adalah kredit yang berada dalam kondisi lancar atau tertib dalam pembayaran, baik angsuran pokok maupun bunga. Dalam konteks perbankan, *performing loan* menunjukkan bahwa debitur memiliki kemampuan dan kemauan yang baik dalam memenuhi kewajibannya sesuai perjanjian kredit. Menurut Otoritas Jasa Keuangan *performing loan* adalah nasabah yang membayar angsuran secara tepat waktu tanpa tunggakan. Penting bagi bank untuk tetap melakukan pemantauan terhadap nasabah *performing loan*.

Non Performing Loan

Non-performing loan (NPL) merupakan kredit yang tergolong bermasalah karena debitur mengalami kesulitan dalam memenuhi kewajiban pembayaran pokok dan/atau bunga dalam jangka waktu tertentu

(Dwihandayani, 2017). Menurut OJK, NPL mencakup tiga kategori kolektibilitas, yaitu kurang lancar, diragukan, dan macet, yang masing-masing menunjukkan tingkat keterlambatan pembayaran dan risiko gagal bayar yang semakin tinggi. Secara umum, kredit dikategorikan sebagai NPL apabila terdapat keterlambatan pembayaran lebih dari 90 hari. NPL menjadi perhatian utama dalam pengelolaan risiko kredit karena dapat berdampak langsung terhadap profitabilitas dan kesehatan keuangan bank. Oleh karena itu, bank diwajibkan untuk membentuk cadangan kerugian penurunan nilai (CKPN) atas kredit bermasalah dan melakukan upaya restrukturisasi jika diperlukan (Otoritas Jasa Keuangan, 2019). Tingkat NPL yang tinggi mencerminkan lemahnya manajemen risiko kredit dan dapat mengganggu stabilitas sistem keuangan secara keseluruhan.

METODE

Penelitian ini menggunakan metode kuantitatif dengan pendekatan deskriptif dan analitis. Proses penelitian diawali dengan identifikasi permasalahan, yaitu keterbatasan sistem klasifikasi konvensional dalam mendeteksi nasabah yang tergolong *performing loan* namun memiliki kecenderungan keterlambatan pembayaran berulang. Tujuan dari penelitian ini adalah mengembangkan model klasifikasi berbasis algoritma *Support Vector Machine* (SVM) untuk mendeteksi nasabah yang berisiko tinggi mengalami keterlambatan pembayaran sebagai bagian dari sistem peringatan dini (*early warning system*).

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari sistem internal PT Bank X. Data mencakup informasi demografis dan historis kredit dari nasabah, seperti skor kredit, usia, jenis kelamin, penghasilan, lama hubungan dengan bank (*tenure*), dan status pembayaran terakhir. Periode data mencakup transaksi dan riwayat pembayaran selama tiga tahun terakhir. Seluruh populasi nasabah yang aktif dalam periode tersebut digunakan sebagai sampel penelitian, sehingga tidak diterapkan teknik sampling. Pengumpulan data dilakukan melalui metode dokumentasi dengan akses resmi dari departemen manajemen risiko PT Bank X.

Tahap awal pengolahan data dilakukan melalui proses *preprocessing*, termasuk penanganan data

hilang, penghapusan duplikasi, dan normalisasi data numerik agar sesuai dengan persyaratan model SVM. Selanjutnya dilakukan pelabelan data menjadi dua kelas: nasabah tepat waktu dan nasabah telat bayar. Pembagian data dilakukan menggunakan rasio pelatihan dan pengujian sebesar 60:40, 70:30, dan 80:20 untuk keperluan validasi model.

Model dikembangkan menggunakan algoritma SVM dengan empat jenis kernel: linear, polynomial, sigmoid, dan radial basis function (RBF). Pemilihan parameter model dilakukan dengan *trial error* untuk mendapatkan kombinasi terbaik dari parameter C dan γ (gamma). Evaluasi performa model dilakukan menggunakan metrik akurasi, sensitivitas (*recall*), spesifisitas, dan *precision*, guna mengukur kemampuan model dalam membedakan nasabah berisiko.

HASIL DAN PEMBAHASAN

KARAKTERISTIK DATA

Pada penelitian ini, data nasabah yang digunakan terdiri dari berbagai variabel yang merepresentasikan karakteristik demografis, keuangan, dan riwayat transaksi nasabah di PT Bank X. Variabel-variabel tersebut dibedakan menjadi dua jenis utama, yaitu variabel numerik dan variabel kategorikal.

Tabel 1. Karakteristik Data Numerik

Variabel	Mean	Maximum	Minimum
Credit Score	650,53	850	350
Usia	38,92	92	18
Tenure	38,98	72	6
Salary	8313849	198208373	5230000

Tabel 1 menunjukkan karakteristik data numerik dari nasabah pada tabel di atas mencerminkan kondisi perekonomian masyarakat yang cukup beragam. Rata-rata *Credit Score* sebesar 650,53 menunjukkan bahwa secara umum nasabah memiliki tingkat kelayakan kredit yang cukup baik, meskipun terdapat individu dengan skor rendah hingga 350 yang dapat mencerminkan potensi risiko finansial atau keterbatasan akses keuangan. Dari segi usia, rata-rata nasabah berada pada 38,92 tahun, yang termasuk usia produktif dan matang secara ekonomi, dengan rentang usia dari 18 hingga 92 tahun yang menunjukkan bahwa layanan keuangan telah menjangkau seluruh lapisan usia. *Tenure* atau lama hubungan nasabah dengan bank yang rata-rata hampir 39 bulan menunjukkan adanya kepercayaan dan kontinuitas hubungan antara nasabah dan lembaga keuangan. Sementara itu, penghasilan

(Salary) menunjukkan angka rata-rata sebesar 8,3 juta rupiah, dengan sebaran dari 5,2 juta hingga lebih dari 190 juta rupiah, yang mencerminkan ketimpangan distribusi pendapatan dalam masyarakat. Hal ini mengindikasikan bahwa meskipun sebagian besar nasabah berada pada tingkat pendapatan menengah, terdapat pula kelompok masyarakat berpendapatan sangat tinggi, yang memperlihatkan kesenjangan ekonomi yang masih terjadi di masyarakat.

Tabel 2. Karakteristik Data Kategorik

Variabel	Kategori	Proporsi
Gender	Pria	54,57%
	Wanita	45,43%
Status	Performing Loan	79,63%
	Non-Performing Loan	20,37%

Tabel 2 menunjukkan berdasarkan karakteristik data kategorikal yang ditampilkan, proporsi nasabah berdasarkan jenis kelamin menunjukkan bahwa mayoritas adalah pria, yakni sebesar 54,57%, sedangkan wanita mencakup 45,43% dari total nasabah. Hal ini mengindikasikan bahwa produk kredit PT Bank X sedikit lebih banyak diakses oleh nasabah pria dibandingkan wanita. Sementara itu, dalam hal status kredit, sebanyak 79,63% nasabah termasuk dalam kategori *performing loan* atau kredit lancar, sedangkan sisanya sebesar 20,37% tergolong *non-performing loan* atau kredit bermasalah. Proporsi tersebut mencerminkan bahwa sebagian besar portofolio kredit berada dalam kondisi yang relatif sehat, meskipun tetap terdapat risiko yang perlu diantisipasi dari sekitar seperlima nasabah yang mengalami keterlambatan atau kegagalan pembayaran.

MODEL SUPPORT VECTORE MACHINE

Pada bagian ini, menjelaskan implementasi performa SVM sangat dipengaruhi oleh jenis kernel yang digunakan dan teknik splitting data yang diterapkan dalam proses pelatihan dan pengujian model.

Tabel 3. SVM Kernel Linear

Pembagian	C	Akurasi
60:40	0,5	65,23%
	1	64,46%
	1,5	62,91%
70:30	0,5	64,32%
	1	67,65%
	1,5	68,91%
80:20	0,5	68,27%
	1	69,77%
	1,5	66,38%

Tabel 3 di atas menunjukkan performa model *Support Vector Machine* (SVM) dengan kernel linear

berdasarkan kombinasi parameter regulasi C dan rasio pembagian data latih-uji. Parameter C berfungsi mengatur keseimbangan antara margin maksimum dan kesalahan klasifikasi, di mana nilai yang terlalu kecil dapat menyebabkan underfitting, sementara nilai yang terlalu besar berisiko overfitting. Dari hasil yang ditampilkan, akurasi terbaik dicapai pada rasio pembagian data 80:20 dengan nilai $C=1$, yang menghasilkan akurasi sebesar 69,77%. Hal ini menunjukkan bahwa model memberikan kinerja optimal ketika menggunakan porsi data latih yang lebih besar (80%) dan parameter regulasi sedang, yang memungkinkan model menangkap pola secara efektif tanpa terlalu sensitif terhadap data pelatihan. Secara umum, terlihat bahwa kenaikan nilai C tidak selalu meningkatkan akurasi, tergantung pada rasio data yang digunakan. Kombinasi 80:20 dengan $C=1$ dapat disimpulkan sebagai parameter terbaik untuk model SVM kernel linear.

Tabel 4. SVM Kernel RBF

Pembagian	C	γ	Akurasi
60:40	0,5	0,01	74,34%
	1	0,1	73,23%
	1,5	1	75,12%
70:30	0,5	0,01	72,54%
	1	0,1	75,66%
	1,5	1	78,81%
80:20	0,5	0,01	79,75%
	1	0,1	82,13%
	1,5	1	85,45%

Tabel 4 menunjukkan hasil pengujian model *Support Vector Machine* (SVM) menggunakan kernel *Radial Basis Function* (RBF) dengan variasi parameter regulasi C, parameter kernel γ , dan rasio pembagian data pelatihan dan pengujian. Parameter C mengontrol penalti terhadap kesalahan klasifikasi, sedangkan γ mengatur sejauh mana pengaruh satu data terhadap lainnya dalam ruang fitur. Berdasarkan hasil, akurasi terbaik diperoleh pada rasio pembagian data 80:20 dengan nilai $C=1,5$ dan $\gamma=1$, yang menghasilkan akurasi sebesar 85,45%. Hal ini menunjukkan bahwa kombinasi nilai regulasi dan parameter kernel yang optimal mampu menangkap kompleksitas pola data dengan baik tanpa overfitting. Secara umum, akurasi cenderung meningkat seiring dengan bertambahnya proporsi data latih (dari 60:40 ke 80:20) dan penggunaan nilai parameter C dan γ yang lebih besar.

Tabel 5. SVM Kernel Polynomial

Pembagian	C	γ	D	Akurasi
60:40	0,5	0,01	2	77,23%
	1	0,1	3	74,21%
	1,5	1	4	75,10%
	0,5	0,01	2	79,63%

70:30	1	0,1	3	74,36%
	1,5	1	4	76,18%
80:20	0,5	0,01	2	83,35%
	1	0,1	3	81,23%
	1,5	1	4	81,49%

Tabel 5 menunjukkan performa algoritma *Support Vector Machine* (SVM) dengan kernel polinomial pada berbagai kombinasi parameter, yaitu nilai C, parameter kernel γ , derajat polinomial D, serta rasio pembagian data pelatihan dan pengujian. Parameter D merepresentasikan derajat dari polinomial yang digunakan dalam kernel, dan mempengaruhi tingkat kompleksitas pemisah data. Hasil pengujian menunjukkan bahwa akurasi model secara umum meningkat seiring dengan proporsi data pelatihan yang lebih besar, yaitu pada pembagian 80:20. Akurasi terbaik tercatat pada kombinasi parameter C=0,5, γ =0,01, dan D=2 dengan rasio pembagian data 80:20, yaitu sebesar 83,35%. Ini menunjukkan bahwa kernel polinomial mampu menangkap pola hubungan non-linear dengan baik ketika derajat polinomial dan parameter lainnya diatur secara optimal. Secara keseluruhan, kernel polinomial memberikan hasil yang cukup kompetitif, terutama pada konfigurasi dengan derajat yang lebih tinggi dan nilai C serta γ yang tepat, menjadikannya salah satu alternatif yang layak dalam pengklasifikasian data nasabah pada penelitian ini.

Tabel 6. SVM Kernel Sigmoid

Pembagian	C	γ	Akurasi
60:40	0,5	0,01	69,76%
	1	0,1	70,11%
	1,5	1	71,23%
70:30	0,5	0,01	72,27%
	1	0,1	69,54%
	1,5	1	69,97%
80:20	0,5	0,01	74,32%
	1	0,1	71,72%
	1,5	1	80,41%

Tabel 6 di atas menunjukkan kernel sigmoid pada algoritma *Support Vector Machine* (SVM), diperoleh variasi akurasi yang berbeda-beda tergantung pada kombinasi parameter C dan γ serta proporsi pembagian data latih dan data uji. Pada skema pembagian data 60:40, akurasi tertinggi dicapai saat C=1,5 dan γ =1 yaitu sebesar 71,23%. Untuk skema 70:30, kombinasi terbaik terjadi saat C=0,5 dan γ =0,01 dengan akurasi 72,27%. Sedangkan pada skema 80:20, akurasi tertinggi diperoleh pada kombinasi parameter C=1,5 dan γ =1 sebesar 80,41%, yang sekaligus menjadi akurasi tertinggi secara keseluruhan. Hal ini menunjukkan bahwa kernel sigmoid bekerja lebih baik ketika model dilatih dengan data yang lebih banyak, serta bahwa

pemilihan parameter yang tepat sangat berpengaruh terhadap kinerja model.

Evaluasi Model

Evaluasi model dari parameter terbaik pada setiap kernel yang digunakan dalam pembangunan model *Support Vector Machine*, evaluasi dilakukan dengan menggunakan beberapa metrik evaluasi yang digunakan, seperti akurasi, sensitivitas, spesifitas.

Tabel 7. Evaluasi Model

Kernel	Akurasi	Sensitivitas	Spesifisitas
Linear	69,77%	76,41%	27,27%
RBF	85,45%	83,83%	90,58%
Polynomial	83,35%	93,69%	52,79%
Sigmoid	80,41%	89,96%	43,75%

Tabel 7 menunjukkan hasil evaluasi model dengan berbagai jenis kernel pada algoritma *Support Vector Machine* (SVM) menunjukkan bahwa kernel RBF (*Radial Basis Function*) memberikan kinerja terbaik dibandingkan kernel lainnya. RBF mencatat akurasi tertinggi sebesar 85,45%, dengan sensitivitas 83,83% dan spesifisitas 90,58%, menunjukkan bahwa model ini sangat baik dalam mengklasifikasikan baik nasabah yang tergolong *performing loan* maupun *non-performing loan*. Kernel polynomial juga menunjukkan performa yang cukup baik, dengan akurasi 83,35% dan sensitivitas tertinggi di antara semua kernel sebesar 93,69%, meskipun spesifisitasnya hanya 52,79%, artinya masih ada kesalahan dalam mendeteksi kategori *non-performing loan*. Sementara itu, kernel sigmoid memberikan akurasi 80,41%, dengan sensitivitas tinggi (89,96%) namun spesifisitasnya lebih rendah (43,75%). Di sisi lain, kernel linear menunjukkan performa terendah, dengan akurasi hanya 69,77% dan spesifisitas sangat rendah sebesar 27,27%, meskipun sensitivitasnya masih tergolong baik (76,41%). Secara keseluruhan, dapat disimpulkan bahwa RBF adalah pilihan kernel paling optimal dalam penelitian ini karena mampu memberikan keseimbangan terbaik antara akurasi, sensitivitas, dan spesifisitas.

PENUTUP

SIMPULAN

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan pada bagian sebelumnya, dapat diperoleh kesimpulan sebagai berikut. Nasabah cukup beragam dan representatif secara sosial ekonomi. Rata-rata skor kredit 650,53 menandakan kelayakan kredit yang baik, meskipun ada sebagian kecil nasabah berisiko tinggi. Usia rata-rata 38,92 tahun dan hubungan rata-rata dengan bank selama 39 bulan mencerminkan stabilitas. Sebaran penghasilan

yang luas menunjukkan ketimpangan pendapatan. Dari sisi kategorikal, mayoritas nasabah adalah pria (54,57%) dan sebagian besar memiliki status kredit lancar (79,63%), yang mencerminkan kondisi portofolio kredit yang relatif sehat. Kernel RBF memberikan performa terbaik dengan akurasi tertinggi 85,45% pada rasio data 80:20 ($C=1,5$; $\gamma=1$), diikuti oleh kernel polynomial (83,35%) dan sigmoid (80,41%). Kernel linear menunjukkan akurasi terendah (69,77%). Kernel RBF direkomendasikan sebagai model terbaik dalam klasifikasi status kredit nasabah karena memberikan hasil paling seimbang dan akurat, dengan akurasi 85,45%, sensitivitas 83,83%, dan spesifisitas 90,58%.

SARAN

Disarankan pada pemilihan parameter C dan γ sangat memengaruhi performa model. Oleh karena itu, tuning parameter secara sistematis seperti grid search atau metode optimasi lainnya sebaiknya dilakukan sebelum implementasi model secara operasional. Model SVM dengan kernel dan parameter optimal dapat diintegrasikan dalam sistem penilaian kelayakan kredit untuk membantu pengambilan keputusan yang lebih objektif dan akurat. Untuk peningkatan kinerja, penelitian selanjutnya dapat mempertimbangkan perbandingan dengan algoritma lain seperti Random Forest, XGBoost, atau model deep learning, terutama pada dataset yang lebih besar dan kompleks.

DAFTAR PUSTAKA

- Basel Committee on Banking Supervision. (2019). *Credit Risk: Review of the Standardised Approach*. <https://www.bis.org/bcbs/publ/d347.htm>
- Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2(2), 121–167.
- Cristianini, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press.
- Djuarni, W., & Ratnasari, R. (2022). IMPLEMENTASI PRINSIP 5C DALAM MENENTUKAN KELAYAKAN PEMBERIAN KREDIT PADA NASABAH. <https://jurnal.unsur.ac.id/ar-rihlah/index>
- Dwihandayani, D. (2017). ANALISIS KINERJA NON PERFORMING LOAN (NPL) PERBANKAN DI INDONESIA DAN FAKTOR-FAKTOR YANG MEMPENGARUHI NPL.
- Kou, G., Lu, C., Sirwongwattana, S., & Huang, Y.-P. (2021). Machine learning applications in credit risk prediction: Recent advances and challenges. *Decision Support Systems*, 154, 113707. <https://doi.org/10.1016/j.dss.2021.113707>
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.
- Otoritas Jasa Keuangan. (2019). *Peraturan Otoritas Jasa Keuangan No. 40/POJK.03/2019 tentang Penilaian Kualitas Aset Bank Umum*.
- Otoritas Jasa Keuangan (OJK). (2021). *Pedoman Manajemen Risiko Kredit pada Bank Umum*.
- Pemerintah Republik Indonesia. (1998). *Undang-Undang Republik Indonesia Nomor 10 Tahun 1998 tentang Perubahan atas Undang-Undang Nomor 7 Tahun 1992 tentang Perbankan*.
- Pernama, B., Dwi Purnomo, H., & Satya Wacana, K. (2023). Analisis Risiko Pinjaman dengan Metode Support Vector Machine, Artificial Neural Network dan Naïve Bayes. *Jurnal Teknologi Informasi Dan Komunikasi*, 7(1), 2023. <https://doi.org/10.35870/jti>
- Schölkopf, B., & Smola, A. J. (2002). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press.
- Syafi'i, Nurdian, O., & Dwilestari, G. (2022). Penerapan Machine Learning untuk Menentukan Kelayakan Kredit Menggunakan Metode Support Vektor Machine. *JURSIMA: Jurnal Sistem Informasi Dan Manajemen STMIK GICI*, 10(2), 108–113. <https://ejournal.stmikgici.ac.id/index.php/jursima/article/view/422>
- Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. Springer.
- Vapnik, V. N. (1998). *Statistical Learning Theory*. Wiley.
- Yuniarti, N., Astuti, B., & Ranidiah, F. (2022). Faktor-Faktor yang Mempengaruhi Non Performing Financing (NPF) pada Bank Umum Syariah yang Terdaftar di Otoritas Jasa Keuangan (OJK) Periode 2015–2019. *Ekombis Review: Jurnal Ilmiah Ekonomi Dan Bisnis*, 10(1), 384–396. <https://doi.org/10.37676/ekombis.v10i1>