Proximal: Jurnal Penelitian Matematika dan Pendidikan Matematika
ISSN 26158132 (cetak)
ISSN 26157667 (online)

Perbandingan Algoritma *Support Vector Machine* Dan *Random Forest* Dalam Klasifikasi Kelayakan Pemberian Pinjaman Pada Bpr Nusumma Jawa Barat

Tami Salsa Sabila[1], Fithri Sri Mulyani[2], Pramesti Melyna Mustofa[3]
Prodi Sains Aktuaria, Fakultas Sains dan Teknologi, Universitas Cipasung Tasikmalaya[1,2,3]
Email: adetamisalsa.2112@gmail.com[1], fithri.sm@uncip.ac.id[2],
pramesti_melyna@uncip.ac.id[3]

Coressponding Author: Tami Salsa Sabila **email:** adetamisalsa.2112@gmail.com

Abstrak. Pesatnya peningkatan pengajuan kredit pada lembaga keuangan menuntut adanya sistem evaluasi kelayakan pinjaman yang mampu bekerja secara cepat, akurat, dan objektif guna meminimalkan risiko kredit bermasalah. Oleh karena itu, penelitian ini bertujuan untuk membandingkan kinerja dua algoritma machine learning, yaitu *Support Vector Machine* (*SVM*) dan *Random Forest* (*RF*), dalam mengklasifikasikan kelayakan pemberian pinjaman pada BPR Nusumma Jawa Barat. Penelitian ini menggunakan data primer sebanyak 753 observasi dengan enam variabel prediktor dan satu variabel target berupa status pinjaman. Metode penelitian yang digunakan adalah eksperimen kuantitatif melalui beberapa tahapan, meliputi pembersihan data, penyeimbangan kelas menggunakan metode *ROSE*, transformasi data, pembangunan model *SVM* dan *Random Forest*, serta evaluasi kinerja model menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, *F1-score*, dan *Area Under Curve* (*AUC*). Hasil penelitian menunjukkan bahwa algoritma *SVM* menghasilkan nilai *AUC* sebesar 0,9748, sedangkan algoritma *Random Forest* memperoleh nilai *AUC* sebesar 0,9953. Berdasarkan hasil tersebut, *Random Forest* menunjukkan performa klasifikasi yang lebih unggul dibandingkan *SVM*. Temuan ini mengindikasikan bahwa *Random Forest* berpotensi menjadi algoritma yang lebih optimal untuk diterapkan sebagai sistem pendukung keputusan dalam evaluasi kelayakan kredit, sehingga dapat membantu lembaga keuangan dalam meningkatkan efektivitas dan kualitas pengambilan keputusan kredit.

Kata Kunci: **Kelayakan kredit, klasifikasi, *Support Vector Machine*, *Random Forest*, *machine learning*, evaluasi model.**

Abstract. *The rapid increase in credit applications in financial institutions necessitates the development of loan eligibility evaluation systems that are capable of operating quickly, accurately, and objectively in order to minimize the risk of non-performing loans. Therefore, this study aims to compare the performance of two machine learning algorithms, namely Support Vector Machine (SVM) and Random Forest (RF), in classifying loan eligibility at BPR Nusumma Jawa Barat. This study uses primary data consisting of 753 observations with six predictor variables and one target variable representing loan status. The research method employed is a quantitative experimental approach conducted through several stages, including data cleaning, class imbalance handling using the ROSE method, data transformation, development of SVM and Random Forest models, and model performance evaluation using confusion matrix, accuracy, precision, recall, F1-score, and Area Under the Curve (AUC). The results show that the SVM algorithm achieved an AUC value of 0.9748, while the Random Forest algorithm obtained a higher AUC value of 0.9953. Based on these findings, Random Forest demonstrates superior classification performance compared to SVM. This result indicates that Random Forest has strong potential to be implemented as a decision support system for loan eligibility evaluation, thereby assisting financial institutions in improving the effectiveness and quality of credit decision-making.*

Keywords: ***Loan eligibility, classification, Support Vector Machine, Random Forest, machine learning, model evaluation.***

A. Pendahuluan

Pesatnya pertumbuhan sektor keuangan, khususnya dalam bidang perbankan dan *fintech* mengakibatkan jumlah pengajuan kredit atau pinjaman juga semakin meningkat (Nur et al., 2019). Semakin meningkatnya jumlah pengajuan pinjaman atau kredit telah menciptakan tantangan baru bagi pihak penyalur dana seperti bank atau lembaga keuangan lainnya. Salah satu tantangan utamanya adalah tingginya tingkat risiko kredit macet atau kredit bermasalah, yang dapat memengaruhi kestabilan keuangan lembaga pemberi pinjaman (Anand et al., 2022). Oleh karena itu, penilaian kelayakan pemberian pinjaman merupakan tahapan krusial dalam menjaga stabilitas keuangan lembaga pemberi pinjaman. Apabila proses penilaian ini tidak dilakukan secara tepat dan akurat, maka lembaga keuangan berpotensi mengalami peningkatan risiko kerugian yang berdampak pada kinerja dan keberlanjutan operasional (Pahlevi et al., 2023).

Dalam praktiknya, proses penilaian kelayakan kredit sering kali memakan waktu yang lama jika dilakukan secara manual karena melibatkan analisis data yang kompleks, seperti data demografis, riwayat keuangan, dan pola pembayaran (Asana & Yanti, 2023). Kondisi ini mendorong perlunya sistem klasifikasi kelayakan kredit yang mampu bekerja secara cepat, akurat, dan efisien guna mendukung pengambilan keputusan kredit yang lebih objektif.

Perkembangan teknologi informasi memungkinkan penerapan metode *machine learning* dalam proses klasifikasi kelayakan kredit. Metode ini memiliki kemampuan untuk mempelajari pola dari data historis dan menghasilkan prediksi secara otomatis dengan kecepatan dan tingkat akurasi yang tinggi (Elwirehardja et al., 2023). Beberapa algoritma yang sering digunakan dalam klasifikasi kredit adalah *Support Vector Machine* (*SVM*) dan *Random Forest*.

Beberapa penelitian terdahulu menunjukkan bahwa algoritma *Support Vector Machine* (*SVM*) memiliki performa yang baik dalam klasifikasi kredit. Penelitian yang dilakukan oleh Putri et al (2021) menunjukkan bahwa *SVM* mampu mengklasifikasikan data kredit dengan akurasi sebesar 95,08% dan nilai *Area Under the Curve* (*AUC*) sebesar 0,9419, yang mengindikasikan kemampuan klasifikasi yang sangat baik. Hasil serupa juga dilaporkan oleh Asana & Yanti (2023), yang memperoleh nilai akurasi sebesar 85%, presisi 91%, dan *recall* 79%, sehingga menunjukkan bahwa *SVM* efektif dalam mengklasifikasikan data kredit secara umum.

Sementara itu, algoritma *Random Forest* (*RF*) juga banyak digunakan dalam klasifikasi kelayakan pemberian pinjaman dan menunjukkan kinerja yang kompetitif. Penelitian oleh Pahlevi et al. (2023) melaporkan bahwa *RF* menghasilkan nilai akurasi sebesar 78,60% dan nilai *AUC* sebesar 0,907, yang termasuk dalam kategori klasifikasi sangat baik. Selain itu, Reddy et al (2022) menunjukkan bahwa *RF* mampu meningkatkan akurasi prediksi kelayakan pemberian pinjaman secara signifikan dibandingkan metode konvensional, dengan tingkat akurasi sebesar 77,29%, sehingga dinilai efektif dalam mendukung pengambilan keputusan kredit di sektor perbankan.

Perbandingan kinerja antara algoritma *SVM* dan *Random Forest* telah dilakukan pada beberapa penelitian dengan hasil yang bervariasi. Penelitian oleh nqs et al (2020) menunjukkan bahwa *SVM* menghasilkan akurasi klasifikasi yang lebih tinggi dibandingkan *Random Forest*, meskipun memerlukan waktu komputasi yang lebih lama dalam proses pelatihan model. Sebaliknya, penelitian oleh Saheed et al (2020) melaporkan bahwa *Random Forest* secara keseluruhan lebih unggul dibandingkan *SVM*, khususnya pada data dengan tingkat kompleksitas yang tinggi.

Perbedaan karakteristik dan hasil penelitian terdahulu menunjukkan adanya research gap terkait kinerja algoritma *SVM* dan *Random Forest* yang sangat dipengaruhi oleh karakteristik data yang digunakan. Berbeda dengan penelitian sebelumnya, penelitian ini menggunakan data riil pengajuan pembiayaan pada BPR Nusumma Jawa Barat serta mengintegrasikan tahapan pembersihan data, penanganan ketidakseimbangan kelas menggunakan metode *ROSE*, dan

evaluasi kinerja model berdasarkan berbagai metrik klasifikasi. Oleh karena itu, penelitian ini bertujuan untuk membandingkan kinerja algoritma *Support Vector Machine* dan *Random Forest* dalam mengklasifikasikan kelayakan pemberian pinjaman pada data pembiayaan BPR Nusumma Jawa Barat. Hasil penelitian ini diharapkan dapat memberikan rekomendasi dalam pemilihan algoritma yang tepat untuk membangun sistem pendukung keputusan kredit yang lebih objektif, efisien, dan berbasis data.

B. Metodologi Penelitian

Metode penelitian yang digunakan dalam penelitian ini adalah eksperimen kuantitatif, yaitu metode penelitian yang menekankan pada pengujian secara objektif menggunakan data numerik untuk memperoleh hasil yang terukur (Rasyid, 2022). Dalam penelitian ini, eksperimen dilakukan dengan menerapkan algoritma *Support Vector Machine* (*SVM*) dan *Random Forest* (*RF*) sebagai perlakuan pada dataset kelayakan pemberian pinjaman.

Data yang digunakan pada penelitian ini merupakan data primer yang diperoleh dari BPR Nusumma Jawa Barat, terdiri dari 866 observasi dengan variabel target berupa status kelayakan pinjaman (disetujui atau ditolak). Variabel independen yang digunakan meliputi usia, jenis kelamin, kolektibilitas kredit, plafon pinjaman, tujuan penggunaan kredit, tingkat pendidikan, pekerjaan, dan suku bunga kredit. Untuk lebih jelasnya dapat dilihat pada tabel berikut.

Tabel 1 Deskripsi Dataset

No	Variabel	Keterangan	Jenis	Definisi Operasional Variabel
1	Usia	Usia debitur	Numerik	Variabel Independen (Bebas)
2	Gender	Jenis kelamin debitur	Kategorikal	
3	Kolektibilitas	Kualitas kredit	Kategorikal	
4	Plafon	Jumlah kredit yang diajukan	Numerik	
5	Penggunaan	Tujuan penggunaan kredit	Kategorikal	
6	Pendidikan	Tingkat pendidikan terakhir	Kategorikal	
7	Pekerjaan	Jenis pekerjaan atau bidang usaha	Kategorikal	
8	Suku Bunga Kredit	Tingkat suku bunga yang dikenakan pada pinjaman	Numerik	
9	Status	Label target (Disetujui atau Ditolak)	Kategorikal	Variabel Dependen (Terikat)

Proses pengolahan data dilakukan secara digital menggunakan perangkat lunak *RStudio.* Adapun langkah-langkah analisis yang akan dilakukan adalah sebagai berikut.

1. Pembersihan Data

Tahap awal penelitian adalah pembersihan data (*data cleaning*) yang bertujuan untuk memastikan kualitas data sebelum proses pemodelan dilakukan. Pembersihan data meliputi penanganan nilai hilang (*missing values*) dengan menggunakan metode *median imputation*, mengingat proporsi data hilang relatif kecil dan variabel bersifat numerik (Mohammed et al., 2021). Selain itu, dilakukan identifikasi terhadap *outlier* menggunakan pendekatan statistik, namun data pencilan tetap dipertahankan karena nilai yang dihasilkan masih logis dan algoritma yang digunakan relatif *robust* terhadap *outlier*.

Proses pembersihan data juga mencakup penghapusan data duplikat untuk menjaga integritas dataset. Selanjutnya, dilakukan transformasi data dengan mengonversi variabel kategorikal ke bentuk numerik menggunakan *one-hot encoding*, serta standarisasi data untuk menyesuaikan skala antar fitur, khususnya pada algoritma *SVM* yang sensitif terhadap

perbedaan skala (Géron, 2019). Tahap ini diakhiri dengan seleksi fitur, di mana fitur yang tidak relevan atau berpotensi menimbulkan bias dihilangkan berdasarkan karakteristik data dan pertimbangan domain knowledge (Ramasubramanian & Singh, 2017).

2. Pembuatan Model

Tahap pembuatan model diawali dengan pembagian dataset menjadi data pelatihan (70%) dan data pengujian (30%). Selanjutnya, dilakukan penanganan ketidakseimbangan kelas menggunakan metode *Random Oversampling* (*ROSE*) untuk meningkatkan representasi kelas minoritas sebelum proses pelatihan model (Khan et al., 2024).

Pemodelan dilakukan menggunakan dua algoritma *machine learning*, yaitu *Support Vector Machine* (*SVM*) dan *Random Forest* (*RF*). Model *SVM* dibangun menggunakan kernel linear dengan skema 5-fold cross-validation serta proses tuning parameter otomatis. Sementara itu, model *Random Forest* dikembangkan melalui teknik *bootstrap* sampling dan *ensemble decision tree* dengan mekanisme voting mayoritas sebagai hasil prediksi akhir (Ayyadevara, 2018).

3. Evaluasi Kinerja Model

Setelah model dibangun, dilakukan evaluasi kinerja untuk menilai kemampuan masing-masing algoritma dalam mengklasifikasikan data. Evaluasi dilakukan menggunakan *Confusion matrix*, serta metrik *accuracy*, *precision*, *recall*, *F1-score*, dan *ROC-AUC* (Zheng, 2015). Penggunaan berbagai metrik ini bertujuan untuk memberikan gambaran kinerja model secara menyeluruh, terutama dalam kondisi data yang berpotensi tidak seimbang.

1. *Confusion matrix*. Digunakan untuk menggambarkan hasil klasifikasi model berdasarkan empat komponen utama, yaitu *True Positive* (*TP*), *True Negative* (*TN*), *False Positive* (*FP*), dan *False Negative* (*FN*). *Confusion matrix* menjadi dasar dalam perhitungan metrik evaluasi lainnya.
2. *Accuracy*. Mengukur proporsi prediksi yang benar terhadap keseluruhan data uji. Metrik ini memberikan gambaran umum mengenai tingkat ketepatan model dalam melakukan klasifikasi.
3. *Precision*. Menunjukkan tingkat ketepatan model dalam memprediksi kelas positif (pinjaman disetujui), yaitu perbandingan antara jumlah prediksi positif yang benar dengan seluruh prediksi positif yang dihasilkan model.
4. *Recall*. Mengukur kemampuan model dalam mendeteksi seluruh data yang seharusnya diklasifikasikan sebagai kelas positif. *Recall* mencerminkan sensitivitas model terhadap kelas disetujui.
5. *F1-score*. Merupakan rata-rata harmonis antara *precision* dan *recall*, yang digunakan untuk menyeimbangkan kedua metrik tersebut, terutama ketika terdapat ketidakseimbangan kelas pada data.
6. *ROC* dan *AUC*. *Receiver Operating Characteristic* (*ROC*) *Curve* menggambarkan hubungan antara *True Positive Rate* dan *False Positive Rate* pada berbagai nilai ambang batas. Nilai *Area Under the Curve* (*AUC*) digunakan untuk mengukur kemampuan model dalam membedakan kelas secara keseluruhan, di mana nilai *AUC* yang lebih tinggi menunjukkan performa klasifikasi yang lebih baik.

C. Hasil Penelitian dan Pembahasan

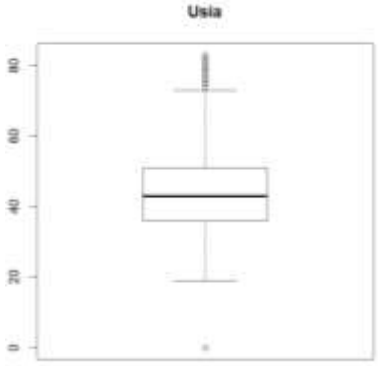
1. Pembersihan Data

Pada tahap pembersihan data, ditemukan nilai hilang pada variabel usia sebesar 1,27% dari total data. Nilai tersebut ditangani menggunakan metode *median imputation*, sehingga tidak terdapat lagi *missing values* pada dataset. Hal tersebut dapat dilihat pada gambar berikut:

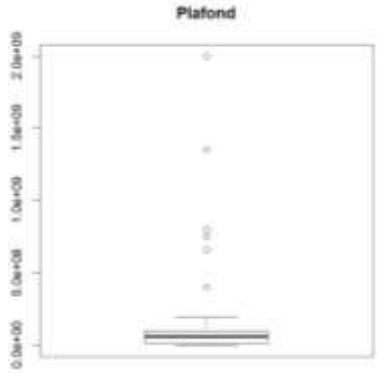
Gambar 1 *Missing Values*

Gambar 2 *Missing Values* Setelah Ditangani

Identifikasi *outlier* pada variabel numerik (usia dan plafon pinjaman) menunjukkan adanya beberapa nilai ekstrem, namun nilai tersebut tetap dipertahankan karena masih logis dan relevan secara kontekstual.



Gambar 3 *Outlier* Fitur Usia



Gambar 4 *Outlier* Fitur Plafon

Selain itu, dilakukan penghapusan data duplikat sehingga jumlah data akhir yang digunakan dalam pemodelan menjadi 753 observasi. Data kategorikal ditransformasikan menggunakan *one-hot encoding*, sedangkan data numerik distandarisasi menggunakan metode *z-score*. Berdasarkan hasil seleksi fitur, dua variabel yaitu kolektibilitas kredit dan suku bunga kredit dieliminasi karena memiliki nilai konstan atau kosong pada kondisi tertentu yang berpotensi menimbulkan bias pada model. Sehingga variabel yang digunakan pada penelitian ini dapat dilihat pada table berikut.

Tabel 2 Deskripsi Variabel Setelah Pemilihan Fitur

No	Variabel	Keterangan	Jenis	Definisi Operasional Variabel
1	Usia	Usia debitur	Numerik	Variabel
2	Gender	Jenis kelamin debitur	Kategorikal	Independen
3	Plafon	Jumlah kredit yang diajukan	Numerik	(Bebas)
4	Penggunaan	Tujuan penggunaan kredit	Kategorikal	
5	Pendidikan	Tingkat pendidikan terakhir	Kategorikal	
6	Pekerjaan	Jenis pekerjaan atau bidang usaha	Kategorikal	
7	Status	Label target (Disetujui atau Ditolak)	Kategorikal	Variabel Dependen (Terikat)

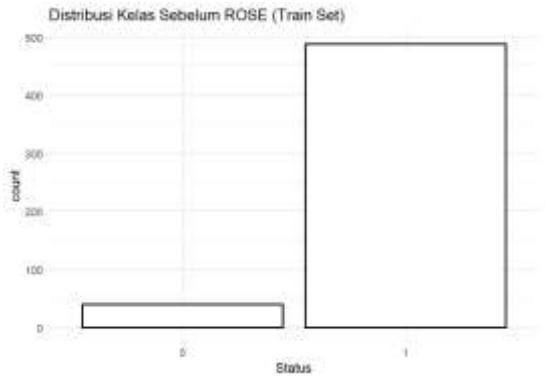
2. Pembuatan Model

Perbandingan yang digunakan pada tahap pembagian data *training* dan *testing* adalah 70:30, 70% untuk data *training* dan 30% untuk data *testing*. Adapun uraiannya dapat dilihat pada tabel berikut.

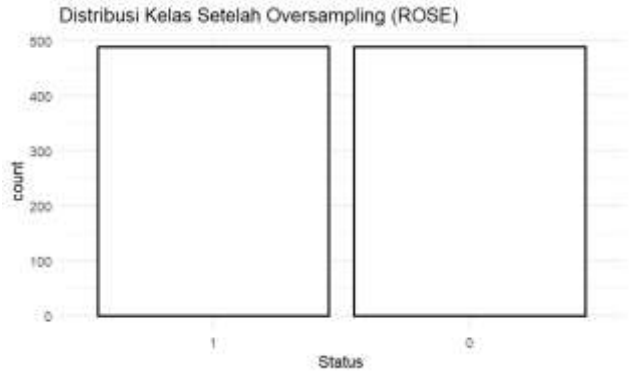
Tabel 3 Pembagian Data *Training* dan *Testing*

Data	Diterima	Ditolak	Jumlah
Training	488	40	528
Testing	208	17	225
Jumlah	696	57	753

Untuk mengatasi ketidakseimbangan kelas pada data pelatihan, diterapkan metode *Random Oversampling* (*ROSE*) sehingga distribusi kelas menjadi seimbang sebelum proses pemodelan dilakukan. Hal tersebut dapat dilihat pada gambar berikut.



Gambar 5 Sebelum *ROSE*

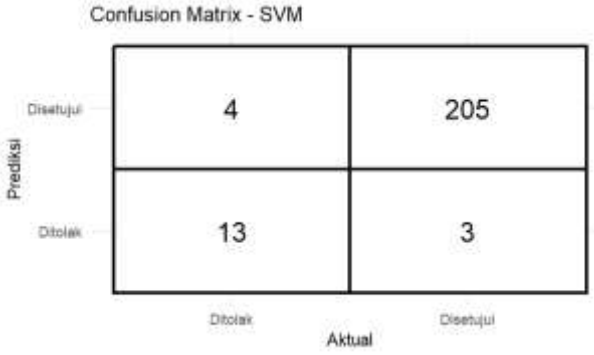


Gambar 6 Setelah *ROSE*

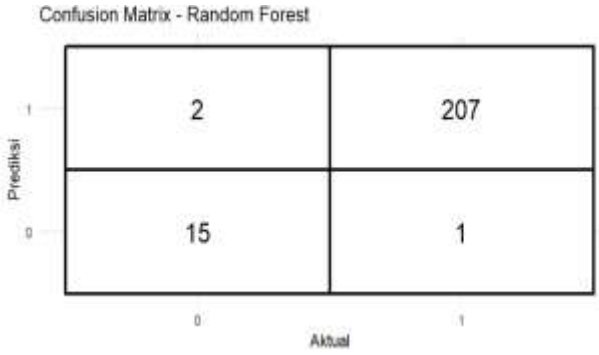
Model *SVM* dibangun menggunakan *kernel* linear dengan skema *5-fold cross-validation*, sedangkan model *RF* dibangun menggunakan 100 pohon keputusan. Evaluasi kinerja model dilakukan menggunakan *confusion matrix* dan metrik evaluasi akurasi, presisi, *recall*, *F1-score*, serta *ROC-AUC*.

3. Evaluasi Kinerja Model

Pada tahap evaluasi dimulai dengan *confusion matrix* yang memberikan gambaran jumlah prediksi benar dan salah yang dilakukan oleh model terhadap masing-masing kelas. Hasilnya dapat dilihat pada gambar berikut.



Gambar 7 *Confusion matrix SVM*



Gambar 8 *Confusion matrix RF*

a. *Support Vector Machine* (*SVM*)

Bedasarkan matriks tersebut dapat dihitung nilai *accuracy*, *precision*, *recall* dan *F1-score*. Untuk algoritma *SVM* nilai akurasinya dapat dihitung sebagai berikut:

$$Accuracy = \frac{205 + 13}{205 + 13 + 4 + 3} = 0{,}9689$$

Untuk menggambarkan seberapa akurat model dalam memprediksi kelas positif (disetujui) dapat dilihat daari nilai presisi yaitu:

$$Precision = \frac{205}{205 + 4} = 0{,}9809$$

Nilai *recall* yang menunjukkan kemampuan model dalam menangkap seluruh data disetujui secara benar, dihitung sebagai berikut:

$$Recall = \frac{205}{205 + 3} = 0{,}9856$$

Sedangkan *F1-score* yang memberikan gambaran menyeluruh tentang kinerja model terhadap kelas positif, nilainya dihitung:

$$F1 - Score = 2 \times \frac{0{,}9809 \times 0{,}9856}{0{,}9809 + 0{,}9856} = 0{,}9832$$

b. *Random Forest*

Nilai akurasi *Random Forest* dapat dihitung sebagai berikut:

$$Accuracy = \frac{207 + 15}{207 + 15 + 2 + 1} = 0{,}9867$$

Hasil perhitungan presisi adalah sebagai berikut:

$$Precision = \frac{205}{205 + 2} = 0{,}9904$$

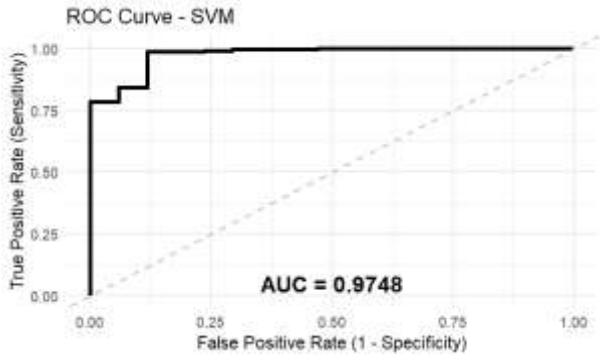
Nilai *recall* dapat dihitung:

$$Recall = \frac{207}{207 + 1} = 0{,}9952$$

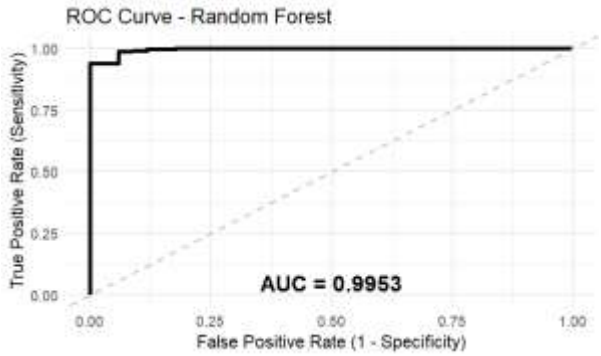
Hasil perhitungan *F1-score* adalah sebagai berikut:

$$F1 - Score = 2 \times \frac{0{,}9904 \times 0{,}9952}{0{,}9904 + 0{,}9952} = 0{,}9928$$

Selain itu, proses evaluasi dilanjukan dengan menggunakan kurva *ROC* (*Receiver Operating Characteristic*) dan perhitungan *AUC* (*Area Under the Curve*). Hasilnya dapat dilihat pada gambar berikut.



Gambar 9 Kurva *ROC SVM*



Gambar 10 Kurva *ROC RF*

Berdasarkan kurva tersebut nilai *AUC* untuk algoritma *SVM* yang dihasilkan adalah 0.9748, mendekati nilai maksimum 1. Ini berarti bahwa secara statistik, ada peluang sebesar 97,48% bahwa model dapat membedakan secara acak antara satu observasi positif dan satu observasi negatif dengan benar. Pada algoritma *Random Forest*, nilai *AUC* yang diperoleh adalah 0,9953, sangat mendekati nilai maksimum 1 dan bahkan lebih tinggi daripada *SVM*. Artinya, terdapat peluang sebesar 99,53% bahwa model dapat membedakan secara benar antara satu observasi positif dan satu observasi negatif secara acak.

Untuk lebih jelasnya, berikut ini adalah perbandingan evaluasi kinerja kedua model yang digunakan dalam penelitian ini.

Tabel 4 Perbandingan Metrik Evaluasi SVM dan *Random Forest*

Metrik	*Support Vector Machine* (*SVM*)	*Random Forest* (*RF*)
Accuracy	0,9689	0,9867
Precision	0,9809	0,9904
Recall	0,9856	0,9952
F1-score	0,9832	0,9928
AUC	0,9748	0,9953

Untuk memperjelas perbandingan kinerja kedua model, Tabel 4 menyajikan ringkasan metrik evaluasi yang digunakan dalam penelitian ini. Berdasarkan tabel tersebut, *Random Forest* menunjukkan nilai yang lebih tinggi pada seluruh metrik evaluasi dibandingkan *Support Vector Machine*, dengan akurasi sebesar 0,9867, presisi 0,9904, *recall* 0,9952, *F1-score* 0,9928, dan nilai *AUC* sebesar 0,9953. Sementara itu, *SVM* juga menunjukkan performa yang sangat baik dengan nilai akurasi 0,9689 dan *AUC* 0,9748.

Hasil penelitian ini menunjukkan bahwa baik algoritma *Support Vector Machine* (*SVM*) maupun *Random Forest* (*RF*) mampu memberikan performa klasifikasi yang sangat baik dalam penilaian kelayakan pemberian pinjaman. Kinerja *SVM* yang tinggi mengindikasikan bahwa algoritma ini efektif dalam menangani data dengan kombinasi fitur numerik dan kategorikal, khususnya setelah melalui tahapan normalisasi dan penyeimbangan kelas. Menurut Sheykhmousa et al (2020), *SVM* memiliki kemampuan yang kuat dalam memisahkan data pada ruang berdimensi tinggi, terutama ketika data telah diproses dengan baik dan memiliki struktur pemisahan yang jelas. Temuan ini sejalan dengan penelitian Teles et al. (2020) serta Asana dan Yanti (2023), yang menyatakan bahwa *SVM* mampu menghasilkan tingkat akurasi yang tinggi dalam klasifikasi kelayakan kredit.

Namun demikian, hasil penelitian ini juga menunjukkan bahwa *Random Forest* memiliki keunggulan dibandingkan *SVM* pada seluruh metrik evaluasi. Keunggulan ini dapat dijelaskan melalui karakteristik *Random Forest* sebagai algoritma ensemble yang menggabungkan banyak pohon keputusan, sehingga mampu menangkap pola data yang kompleks dan mengurangi risiko

overfitting. Kinerja yang sangat baik ini sejalan dengan hasil pada penelitian Pahlevi et al. (2023) yang membuktikan bahwa *Random Forest* mampu memberikan performa yang andal dalam menilai kelayakan kredit. Temuan penelitian ini juga konsisten dengan Saheed et al. (2020) yang menyatakan bahwa *Random Forest* mampu menghasilkan tingkat akurasi yang sangat tinggi terlebih jika digunakan pada data dengan kompleksitas tinggi dan variasi fitur yang beragam.

Implikasi dari hasil penelitian ini menunjukkan bahwa *Random Forest* memiliki potensi yang lebih besar untuk diterapkan sebagai sistem pendukung keputusan (*Decision Support System*) dalam evaluasi kelayakan pemberian pinjaman, khususnya pada lembaga keuangan seperti BPR Nusumma Jawa Barat. Dengan tingkat akurasi dan kemampuan generalisasi yang lebih tinggi, *Random Forest* dapat membantu analis kredit dalam mengambil keputusan yang lebih objektif, konsisten, dan berbasis data historis. Selain itu, penerapan model klasifikasi berbasis machine learning ini diharapkan dapat meningkatkan efisiensi proses penilaian kredit serta meminimalkan risiko kredit bermasalah, sehingga mendukung stabilitas dan kinerja keuangan lembaga perbankan.

D. Kesimpulan

Berdasarkan hasil penelitian mengenai perbandingan algoritma *Support Vector Machine* (*SVM*) dan *Random Forest* (*RF*) dalam klasifikasi kelayakan pemberian pinjaman pada BPR Nusumma Jawa Barat, dapat disimpulkan bahwa kedua algoritma mampu memberikan performa klasifikasi yang sangat baik. Hal ini ditunjukkan oleh nilai *AUC* yang tinggi serta metrik evaluasi lainnya yang konsisten. Model *SVM* menunjukkan kemampuan klasifikasi yang sangat baik dengan nilai *AUC* sebesar 0,9822, sedangkan model *Random Forest* menghasilkan performa yang lebih unggul dengan nilai *AUC* sebesar 0,9953. Perbandingan hasil evaluasi menunjukkan bahwa *Random Forest* memiliki kinerja yang lebih optimal dibandingkan *SVM* pada seluruh metrik evaluasi yang digunakan.

Dengan demikian, *Random Forest* dinilai lebih sesuai untuk diterapkan pada kasus klasifikasi kelayakan pemberian pinjaman pada BPR Nusumma Jawa Barat. Kedua algoritma tetap memiliki potensi untuk dikembangkan lebih lanjut melalui penyesuaian parameter yang tepat sehingga dapat dimanfaatkan sebagai sistem pendukung keputusan yang akurat dan efisien. Untuk penelitian selanjutnya, disarankan menggunakan dataset yang lebih besar dan beragam dengan menambahkan fitur-fitur relevan lainnya agar model yang dihasilkan lebih representatif. Selain itu, disarankan juga untuk menerapkan analisis *feature importance* guna mengidentifikasi variabel yang paling berpengaruh terhadap penentuan kelayakan kredit serta pengujian algoritma lain atau metode *ensemble* yang lebih kompleks dapat dilakukan guna memperoleh performa klasifikasi yang lebih optimal.

DAFTAR PUSTAKA

Anand, M., Velu, A., & Whig, P. (2022). Prediction of Loan Behaviour with Machine Learning Models for Secure Banking. *Journal of Computer Science and Engineering (JCSE)*, *3*(1), 1–13. https://doi.org/10.36596/jcse.v3i1.237

Asana, I. M. D. P., & Yanti, N. P. D. T. (2023). Sistem Klasifikasi Pengajuan Kredit Dengan Metode Support Vector Machine (SVM). *Jurnal Sistem Cerdas*, *2*, 123–133. https://apic.id/jurnal/index.php/jsc/article/view/287

Ayyadevara, V. K. (2018). Pro Machine Learning Algorithms: A Hands-On Approach to Implementing Algorithms in Python and R. In *Pro Machine Learning Algorithms: A Hands-On Approach to Implementing Algorithms in Python and R*. Apress Media LLC. https://doi.org/10.1007/978-1-4842-3564-5

Elwirehardja, G. N., Suparyanto, T., & Pardamean, B. (2023). *Pengenalan Konsep Machine Learning untuk Pemula* (F. A. Munawwar, Ed.). INSTIPER PRESS (IKAPI & APPTI). https://research.binus.ac.id/airdc/2023/01/pengenalan-konsep-machine-learning-untuk-pemula/

Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* (R. Roumeliotis & N. Tache, Eds.; 2nd ed.). O'Reilly Media, Inc. https://www.oreilly.com/library/view/hands-on-machine-learning/9781492032632/

Khan, A. A., Chaudhari, O., & Chandra, R. (2024). A Review of Ensemble Learning and Data Augmentation Models for Class Imbalanced Problems: Combination, Implementation and Evaluation. *Expert Systems with Applications*, *244*, 1–29. https://doi.org/10.1016/j.eswa.2023.122778

Mohammed, M. B., Zulkafli, H. S., Adam, M. B., Ali, N., & Baba, I. A. (2021). Comparison of Five Imputation Methods in Handling Missing Data in a Continuous Frequency Table. *AIP Conference Proceedings*, *2355*. https://doi.org/10.1063/5.0053286

Nur, I. T. A., Setiawan, N. Y., & Bachtiar, F. A. (2019). Perbandingan Performa Metode Klasifikasi SVM, Neural Network, Naive Bayes untuk Mendeteksi Kualitas Pengajuan Kredit di Koperasi Simpan Pinjam. *Jurnal Teknologi Informasi dan Ilmu Komputasi*, *4*. https://doi.org/http://dx.doi.org/10.25126/jtiik.2019641352

Pahlevi, O., Amrin, & Handrianto, Y. (2023). Implementasi Algoritma Klasifikasi Random Forest Untuk Penilaian Kelayakan Kredit. *Jurnal*, *5*(1), 71–76. https://doi.org/http://dx.doi.org/10.31294/infortech.v5i1.15829

Putri, N. H., Fatekurohman, M., & Tirta, I. M. (2021). Credit Risk Analysis using Support Vector Machines Algorithm. *Journal of Physics: Conference Series*, *1836*(1), 1–10. https://doi.org/10.1088/1742-6596/1836/1/012039

Ramasubramanian, K., & Singh, A. (2017). Machine Learning Using R. In *Machine Learning Using R* (1st ed.). Apress Media. https://doi.org/10.1007/978-1-4842-2334-5

Rasyid, F. (2022). *Metode Penelitian Kualitatif dan Kuantitatif*. IAIN Kediri Press. https://repository.iainkediri.ac.id/859/1/buku%20metode%20penelitian%20FATHOR%20RASYID.pdf

Reddy, C. S., Siddiq, A. S., & Jayapandian, N. (2022). Machine Learning based Loan Eligibility Prediction using Random Forest Model. *Proceedings of the 7th International Conference on Communication and Electronics Systems, ICCES 2022*, 1073–1079. https://doi.org/10.1109/ICCES54183.2022.9835875

Saheed, Y. K., Hambali, M. A., Arowolo, M. O., & Olasupo, Y. A. (2020). Application of GA Feature Selection on Naive Bayes, Random Forest and SVM for Credit Card Fraud

Detection. *International Conference on Decision Aid Sciences and Application (DASA)*, 1091–1097. https://doi.org/10.1109/DASA51403.2020.9317228

Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., & Homayouni, S. (2020). Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 6308–6325. https://doi.org/10.1109/JSTARS.2020.3026724

Teles, G., Rodrigues, J. J. P. C., Rabêlo, R. A. L., & Kozlov, S. A. (2020). Comparative Study of Support Vector Machines and Random Forests Machine Learning Algorithms on Credit Operation. *Software - Practice and Experience*, *51*(12), 1–9. https://doi.org/https://doi.org/10.1002/spe.2842

Zheng, A. (2015). *Evaluating Machine Learning Models* (S. Cutt, Ed.; 1st ed.). O'Reilly Media, Inc. https://www.oreilly.com/library/view/evaluating-machine-learning/9781492048756/