Copyright © The Author(s) 2027
Jurnal Artificial, Vol 5 No 1, April 2027

Klasifikasi Kelayakan Kredit Berbasis Machine Learning: Evaluasi Komparatif K-Nearest Neighbor dan Naive Bayes

Herwin [1*], Supriadi Sahibu [2], Respaty Namruddin [3]

[1, 2, 3] Universitas Handayani Makassar, Indonesia

* mpsxml332@gmail.com

Abstrak

Penilaian kelayakan kredit memerlukan metode yang mampu mengolah karakteristik pemohon secara konsisten serta mempertimbangkan risiko kesalahan klasifikasi. Penelitian ini bertujuan mengklasifikasikan kelayakan kredit menggunakan algoritma K-Nearest Neighbor (KNN) dan *Gaussian Naive Bayes* serta membandingkan kinerjanya berdasarkan *accuracy, precision, recall*, dan F1-score. Data yang digunakan berupa 14.897 data historis pengajuan kredit dengan fitur hasil rekayasa berupa *Debt Service Ratio* (DSR), *disposable income*, *Loan to Value* (LTV), serta indikator kelayakan usia, status rumah, dan tunggakan. Data dibagi menggunakan train-test split dengan proporsi 80% data latih dan 20% data uji. Model KNN diuji pada beberapa nilai K dan diperoleh K=5 sebagai konfigurasi terbaik. Hasil evaluasi menunjukkan bahwa KNN memperoleh accuracy sebesar 94,66% dan weighted F1-score sebesar 94,84%, lebih tinggi dibandingkan *Gaussian Naive Bayes* yang memperoleh accuracy 91,11% dan *weighted* F1-score 93,39%. Namun, *Gaussian Naive Bayes* menunjukkan kemampuan lebih baik dalam mendeteksi kelas Tidak Layak dengan recall sebesar 100%, sedangkan KNN hanya 30,10%. Temuan ini menunjukkan adanya *trade-off* antara kinerja klasifikasi secara keseluruhan dan kemampuan mendeteksi pemohon berisiko. KNN lebih sesuai untuk klasifikasi agregat, sedangkan *Gaussian Naive Bayes* dapat dipertimbangkan ketika deteksi kelas Tidak Layak menjadi prioritas. Hasil klasifikasi selanjutnya dapat dimanfaatkan sebagai alat bantu dalam proses penilaian kredit.

Kata Kunci: *Credit Scoring, Gaussian Naive Bayes, Kelayakan Kredit, K-Nearest Neighbor, Machine Learning*

Pendahuluan

Kredit merupakan salah satu layanan keuangan yang memiliki peran penting dalam mendukung kebutuhan konsumtif maupun produktif masyarakat. Melalui fasilitas kredit, masyarakat dapat memperoleh pembiayaan untuk berbagai kebutuhan, sedangkan lembaga keuangan memperoleh pendapatan dari aktivitas penyaluran dana kepada debitur. Dalam perspektif manajemen risiko, pemberian kredit merupakan aktivitas yang mengandung risiko karena lembaga keuangan harus menilai kemampuan dan kelayakan calon debitur sebelum dana disalurkan (Shi et al., 2022). Perkembangan layanan keuangan digital dan meningkatnya volume transaksi menyebabkan lembaga keuangan harus mengolah data pemohon dalam jumlah semakin besar sehingga proses penilaian kredit membutuhkan metode yang mampu memanfaatkan informasi secara efisien (Amarnadh & Moparthi, 2023). Sejalan dengan kondisi tersebut, penelitian *credit scoring* menunjukkan adanya pergeseran dari pendekatan statistik konvensional menuju penggunaan metode *machine learning* yang mampu menangani karakteristik data lebih kompleks (Hayashi, 2022). Oleh karena itu, proses penilaian kredit tidak hanya dituntut

https://doi.org/10.54065/artificial.5.1.2027.1974

menghasilkan keputusan secara cepat, tetapi juga perlu memiliki kemampuan prediksi yang baik dan mendukung pengambilan keputusan secara konsisten (Zhang & Yu, 2024).

Risiko kredit pada dasarnya berkaitan dengan kemungkinan debitur tidak mampu memenuhi kewajiban pembayaran sesuai dengan perjanjian kredit. Dalam konteks penilaian kredit, karakteristik peminjam dan informasi kondisi finansial merupakan bagian penting dalam memperkirakan risiko gagal bayar (Anwar et al., 2021). Salah satu pendekatan yang digunakan dalam penilaian calon debitur adalah prinsip *Character, Capacity, Capital, Collateral,* dan *Condition of Economy* atau 5C yang mencakup karakter, kemampuan membayar, modal, jaminan, serta kondisi ekonomi peminjam (Khayatun et al., 2021). Penerapan prinsip 5C pada lembaga perbankan juga dapat menjadi dasar penilaian pemberian kredit untuk membantu meminimalkan risiko kredit (Misral & Finowa, 2021). Selain itu, karakteristik ekonomi dan kepemilikan aset atau jaminan berhubungan dengan akses masyarakat terhadap sumber pembiayaan formal sehingga informasi mengenai kondisi ekonomi pemohon relevan dalam proses penilaian kredit (Santoso et al., 2023).

Dalam praktiknya, penilaian kelayakan kredit dapat melibatkan sejumlah atribut pemohon, seperti usia, pekerjaan, penghasilan, jumlah tanggungan, status kepemilikan rumah, harga kendaraan, jumlah pinjaman, angsuran, serta riwayat pembayaran. Penelitian mengenai *credit scoring* menunjukkan bahwa karakteristik individu, informasi finansial, serta variabel ekonomi dapat digunakan sebagai prediktor dalam membangun model klasifikasi risiko kredit (Xia et al., 2021). Penggunaan sejumlah variabel secara simultan memungkinkan model memperoleh pola hubungan yang lebih kompleks dibandingkan penilaian berdasarkan satu karakteristik saja (Dumitrescu et al., 2022). Pendekatan *machine learning* juga telah digunakan untuk mengolah karakteristik peminjam dalam menghasilkan prediksi kelayakan kredit dengan mempertimbangkan hubungan nonlinier antarvariabel (Liu et al., 2022). Dengan demikian, keberagaman atribut pemohon menghasilkan struktur data yang kompleks sehingga diperlukan metode klasifikasi yang mampu memanfaatkan berbagai karakteristik tersebut secara bersamaan (Liu et al., 2024).

Karakteristik tersebut semakin penting ketika penilaian dilakukan menggunakan data historis dalam jumlah besar. Dataset penelitian ini terdiri atas 14.897 data pemohon kredit, dengan 10.252 pemohon atau 68,82% berstatus *Approve* dan 4.645 pemohon atau 31,18% berstatus *Reject*. Komposisi tersebut menunjukkan adanya ketidakseimbangan kelas antara pemohon yang dinyatakan layak dan tidak layak. Permasalahan *class imbalance* sering ditemukan dalam data *credit scoring* karena jumlah debitur dengan status tertentu, khususnya debitur bermasalah, umumnya lebih sedikit dibandingkan kelas lainnya (Chen et al., 2024). Ketidakseimbangan kelas dapat memengaruhi proses pembelajaran model karena algoritma cenderung memperoleh informasi lebih banyak dari kelas mayoritas sehingga kemampuan mengenali kelas minoritas dapat menurun (Lenka et al., 2022). Teknik *resampling* dan penyesuaian ambang klasifikasi juga dapat digunakan untuk menangani distribusi kelas yang tidak seimbang pada data kredit (Yang et al., 2023). Oleh karena itu, distribusi kelas perlu diperhatikan sejak tahap pembangunan hingga evaluasi model agar performa yang diperoleh tidak hanya mencerminkan kemampuan mengenali kelas mayoritas.

Karakteristik dataset tersebut menunjukkan bahwa pemanfaatan data historis perlu didukung oleh metode klasifikasi yang mampu mempelajari pola dari berbagai atribut

pemohon secara simultan. *Data mining* dan *machine learning* memungkinkan pola dalam data historis dipelajari untuk menghasilkan model klasifikasi terhadap data pemohon baru (Wafda, 2024). Kajian sistematis mengenai prediksi risiko kredit menunjukkan bahwa penggunaan *machine learning* berkembang karena mampu memanfaatkan data dalam jumlah besar dan mengidentifikasi pola yang relevan untuk memperkirakan risiko kredit. Pendekatan tersebut juga dapat meningkatkan kemampuan model *credit scoring* dalam menangani hubungan nonlinier dan karakteristik data kompleks dibandingkan pendekatan regresi konvensional (Dumitrescu et al., 2022). Berbagai algoritma *machine learning* dan *ensemble learning* telah digunakan dalam *credit scoring* untuk memperoleh performa klasifikasi yang lebih baik sekaligus mempertahankan kemampuan model dalam menjelaskan hasil prediksinya (Liu et al., 2024). Perkembangan tersebut menunjukkan bahwa pendekatan berbasis data dapat menjadi instrumen pendukung pengambilan keputusan kredit karena sejumlah atribut pemohon dapat diproses secara simultan berdasarkan pola historis dalam dataset (Nwafor et al., 2024).

Dari berbagai algoritma klasifikasi yang dapat diterapkan pada data kredit, *K-Nearest Neighbor* (KNN) dan Naive Bayes memiliki mekanisme pembelajaran yang berbeda sehingga relevan untuk dibandingkan. KNN menentukan kelas suatu data berdasarkan kedekatannya dengan sejumlah data lain yang memiliki karakteristik serupa, sehingga dapat digunakan untuk mengidentifikasi pola pemohon baru berdasarkan kemiripan dengan data historis (Shi et al., 2022). Sementara itu, Naive Bayes menggunakan pendekatan probabilistik berdasarkan Teorema Bayes dengan asumsi independensi bersyarat antaratribut sehingga proses klasifikasinya relatif sederhana dan efisien (Shi et al., 2022). Kedua algoritma memiliki karakteristik berbeda sehingga performanya dapat berubah ketika diterapkan pada dataset kredit yang berbeda. Berbagai studi komparatif *credit scoring* juga menunjukkan bahwa tidak terdapat satu algoritma yang selalu menghasilkan performa terbaik pada seluruh karakteristik dataset (Lenka et al., 2022; Quan & Sun, 2024). KNN sensitif terhadap representasi dan skala fitur karena proses klasifikasi didasarkan pada ukuran jarak, sedangkan Naive Bayes lebih bergantung pada estimasi probabilitas dan asumsi hubungan antarfitur (Shi et al., 2022).

Penggunaan metode komputasional dalam penilaian kelayakan kredit berkembang melalui berbagai pendekatan, termasuk metode *multi-criteria decision making*. Pendekatan berbasis *machine learning* memungkinkan informasi karakteristik pemohon digunakan untuk membangun model klasifikasi yang dapat membantu pengambilan keputusan kredit (Shi et al., 2022; Zhang & Yu, 2024). Di sisi lain, metode seperti TOPSIS memiliki mekanisme berbeda karena keputusan dibangun melalui pembobotan dan pemeringkatan sejumlah kriteria yang telah ditentukan sebelumnya. Perbedaan tersebut penting diperhatikan karena metode *multi-criteria decision making* lebih menekankan struktur preferensi dan bobot kriteria, sedangkan *machine learning* mempelajari pola hubungan antara atribut input dan kelas target berdasarkan data historis (Jemai & Zarrad, 2023; Zhang & Yu, 2024). Dengan demikian, metode berbasis pembobotan dan pembelajaran dari data dapat menghasilkan karakteristik keputusan yang berbeda meskipun menggunakan atribut pemohon yang serupa.

Sejumlah penelitian terdahulu menunjukkan bahwa KNN dan Naive Bayes dapat digunakan dalam permasalahan *credit scoring* dan *credit risk classification*. Kajian mengenai *machine learning-driven credit risk* menempatkan KNN dan Naive Bayes sebagai metode klasifikasi yang dapat diterapkan pada risiko kredit serta menunjukkan bahwa

performanya dapat berbeda berdasarkan dataset yang digunakan (Shi et al., 2022). Studi mengenai *credit scoring* juga menunjukkan bahwa KNN dapat digunakan dalam evaluasi risiko kredit dan dibandingkan dengan berbagai metode klasifikasi lainnya (Lenka et al., 2022; Rayhan et al., 2026). Penelitian pada dataset risiko kredit menunjukkan bahwa karakteristik algoritma dan struktur data dapat memengaruhi hasil klasifikasi sehingga evaluasi empiris pada dataset tertentu tetap diperlukan (Quan & Sun, 2024). Penelitian mengenai sistem prediksi risiko kredit berbasis *machine learning* juga menunjukkan bahwa KNN dan Naive Bayes dapat digunakan sebagai bagian dari berbagai pendekatan klasifikasi risiko kredit (Emmanuel et al., 2024).

Meskipun kedua algoritma tersebut telah banyak digunakan, perbandingan performanya tetap dipengaruhi oleh karakteristik dataset dan tahapan pengolahan data. Perbedaan jumlah data, atribut, distribusi kelas, teknik *preprocessing*, serta target klasifikasi dapat menghasilkan performa model yang berbeda (Shi et al., 2022). Selain itu, proses pembagian data, reduksi dimensi, dan ketidakseimbangan kelas dapat memengaruhi hasil evaluasi serta interpretasi model pada permasalahan *credit risk assessment* (Chen et al., 2024). Oleh karena itu, hasil penelitian KNN dan Naive Bayes pada domain lain, seperti analisis sentimen terhadap data tekstual, tidak dapat langsung digunakan untuk menyimpulkan performanya pada data tabular finansial. Perbedaan struktur data tersebut menjadi alasan penting untuk menguji KNN dan Naive Bayes secara langsung pada dataset pemohon kredit dengan kombinasi atribut numerik dan kategorikal (Quan & Sun, 2024).

Selain karakteristik distribusi kelas, evaluasi model perlu mempertimbangkan konsekuensi kesalahan klasifikasi pada masing-masing kelas. Pada dataset penelitian, jumlah pemohon berstatus *Approve* lebih besar dibandingkan *Reject*, sehingga kondisi *class imbalance* dapat menyebabkan model menghasilkan *accuracy* tinggi karena mengikuti pola kelas mayoritas, tetapi belum tentu mampu mengenali kelas minoritas dengan baik (Lenka et al., 2022; Chen et al., 2024). Dalam penilaian kredit, kemampuan mengenali kelas berisiko menjadi penting karena kesalahan klasifikasi dapat menimbulkan konsekuensi finansial bagi lembaga pemberi kredit (Zhang et al., 2024). Oleh sebab itu, evaluasi model tidak cukup menggunakan *accuracy*, tetapi perlu dilengkapi dengan *precision, recall, F1-score,* dan *confusion matrix* untuk memperoleh gambaran performa pada masing-masing kelas.

Selain distribusi kelas, karakteristik hubungan antaratribut finansial juga menjadi pertimbangan dalam pembangunan model. Atribut seperti jumlah pinjaman, penghasilan, nilai aset, angsuran, dan kemampuan pembayaran memiliki hubungan yang saling berkaitan sehingga *feature engineering* dapat digunakan untuk membentuk representasi fitur yang lebih informatif dalam mempelajari risiko kredit (Jemai & Zarrad, 2023). Transformasi fitur juga dapat membantu mengungkap karakteristik nonlinier dan meningkatkan performa prediksi (Liu et al., 2024), sedangkan kombinasi *feature engineering*, *feature selection,* dan *machine learning* dapat mendukung pengembangan model prediksi risiko kredit (Muangthanang et al., 2024). Oleh karena itu, penelitian ini menggunakan rekayasa fitur seperti *Debt Service Ratio* (DSR), *disposable income*, dan *Loan to Value* (LTV) untuk memperoleh representasi kemampuan finansial pemohon yang lebih informatif. Fitur tersebut memungkinkan model memanfaatkan atribut mentah sekaligus informasi mengenai hubungan antara kewajiban pembiayaan, nilai aset, dan kemampuan finansial pemohon.

Berdasarkan karakteristik tersebut, penelitian ini memiliki perbedaan dengan penelitian sebelumnya pada fokus penerapan dan karakteristik data yang digunakan. Penelitian terdahulu menunjukkan bahwa KNN dan Naive Bayes dapat digunakan dalam *credit risk assessment*, tetapi performanya bergantung pada karakteristik dataset dan proses pengolahan data (Shi et al., 2022; Lenka et al., 2022). Perubahan fitur, proses seleksi fitur, dan metode transformasi juga dapat menghasilkan perbedaan performa model (Jemai & Zarrad, 2023; Liu et al., 2024). Dengan demikian, penelitian ini tidak hanya membandingkan kedua algoritma, tetapi juga menerapkannya pada data tabular finansial dengan kombinasi atribut numerik dan kategorikal serta fitur finansial hasil rekayasa. Perbedaan tersebut menjadi dasar untuk menguji apakah KNN dan Naive Bayes memberikan performa berbeda pada klasifikasi kelayakan kredit.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk mengklasifikasikan kelayakan kredit berdasarkan data pemohon menggunakan algoritma *K-Nearest Neighbor* dan Naive Bayes serta membandingkan kinerjanya berdasarkan *accuracy, precision, recall,* dan F1-score. Penelitian ini juga mengimplementasikan hasil klasifikasi ke dalam aplikasi berbasis web sebagai bentuk penerapan model pada proses penilaian kredit. Pemilihan KNN dan Naive Bayes didasarkan pada mekanisme pembelajaran keduanya yang berbeda sehingga perbandingannya dapat memberikan informasi empiris mengenai algoritma yang lebih sesuai dengan karakteristik dataset penelitian (Shi et al., 2022; Quan & Sun, 2024). Dengan pendekatan tersebut, penelitian diharapkan memberikan gambaran mengenai algoritma yang lebih sesuai sekaligus menyediakan alternatif sistem pendukung penilaian kredit berbasis data untuk membantu proses pengambilan keputusan.

Metode

Desain Penelitian dan Data

Penelitian ini menggunakan pendekatan kuantitatif dengan metode komparatif eksperimental untuk mengklasifikasikan kelayakan kredit menggunakan algoritma K-Nearest Neighbor (KNN) dan Gaussian Naive Bayes. Perbandingan kedua algoritma dilakukan menggunakan dataset historis nasabah sehingga dapat diketahui algoritma yang memberikan kinerja klasifikasi lebih baik dalam membedakan nasabah yang termasuk kategori Layak dan Tidak Layak.

Data yang digunakan merupakan data historis pengajuan kredit dengan jumlah awal 14.897 record. Data memuat sejumlah informasi mengenai karakteristik pemohon dan kondisi pembiayaan, antara lain usia, penghasilan, jumlah pinjaman, angsuran, harga kendaraan, jumlah tanggungan, kepemilikan rumah, lama tunggakan, dan kolektibilitas. Data tersebut digunakan sebagai dasar untuk membentuk variabel prediktor dan target klasifikasi.

Tahapan penelitian terdiri atas pengolahan data, pembentukan fitur, penentuan label kelayakan, pembagian data, pelatihan model KNN dan Gaussian Naive Bayes, serta evaluasi kinerja model. Alur ini digunakan agar kedua algoritma memperoleh data dan skema pengujian yang sama sehingga hasil perbandingan dapat dilakukan secara objektif.

Preprocessing dan Pembentukan Fitur

Tahap preprocessing dilakukan untuk memastikan data berada dalam kondisi yang sesuai sebelum digunakan dalam pemodelan. Proses meliputi pemeriksaan dan pembersihan data, penanganan data yang tidak sesuai, pengurangan atribut yang tidak relevan terhadap klasifikasi, serta transformasi variabel ke bentuk numerik yang dapat diproses oleh algoritma machine learning. Selanjutnya dilakukan *feature engineering* untuk membentuk indikator yang dapat merepresentasikan kemampuan finansial dan risiko pemohon secara lebih terukur. Fitur yang digunakan dalam penelitian meliputi indikator *Debt Service Ratio* (DSR), *disposable income*, dan *Loan to Value* (LTV). DSR digunakan untuk mengukur proporsi pendapatan yang digunakan untuk membayar angsuran kredit. DSR dihitung dengan persamaan:

$$DSR = \frac{Angsuran\ Bulanan}{Penghasilan\ Bulanan} \times 100\%$$

Disposable income digunakan untuk menunjukkan sisa penghasilan pemohon setelah memperhitungkan estimasi kebutuhan pengeluaran. Nilai *disposable income* yang positif menunjukkan bahwa pemohon masih memiliki sisa kemampuan finansial setelah memenuhi kebutuhan tersebut. Sementara itu, LTV digunakan untuk mengukur perbandingan antara jumlah pinjaman dengan harga kendaraan dan dihitung menggunakan persamaan:

$$LTV = \frac{Jumlah\ Pinjaman}{Harga\ Kendaraan} \times 100\%$$

Berdasarkan hasil feature engineering, dibentuk lima indikator kelayakan, yaitu DSR Layak, Usia Layak, Rumah Layak, Tunggakan Layak, dan LTV Layak. Masing-masing indikator memiliki nilai biner, yaitu 1 untuk memenuhi kriteria kelayakan dan 0 untuk tidak memenuhi kriteria.

Kriteria DSR Layak ditentukan berdasarkan DSR di bawah 30% dan *disposable income* yang bernilai positif. Usia Layak diberikan kepada pemohon berusia 21–60 tahun. Rumah Layak diberikan apabila status tempat tinggal merupakan rumah milik sendiri, keluarga, atau orang tua. Tunggakan Layak diberikan apabila lama tunggakan kurang dari 30 hari. Adapun LTV Layak ditentukan apabila rasio pinjaman terhadap harga kendaraan berada di bawah 85%.

Label target dibentuk berdasarkan status kolektibilitas nasabah. Kolektibilitas 1 dan 2 dikategorikan sebagai Layak, sedangkan kolektibilitas di atas 2 dikategorikan sebagai Tidak Layak. Dengan pembentukan target tersebut, permasalahan penelitian diformulasikan sebagai klasifikasi biner dengan dua kelas keluaran.

Pembagian Data dan Pemodelan

Setelah preprocessing dan feature engineering selesai, data dipisahkan menjadi matriks prediktor (X) dan variabel target (y). Lima indikator kelayakan, yaitu DSR Layak, Usia Layak, Rumah Layak, Tunggakan Layak, dan LTV Layak digunakan sebagai fitur utama dalam proses klasifikasi.

Dataset dibagi menggunakan metode train-test split dengan proporsi 80% data latih dan 20% data uji. Data latih digunakan untuk membangun model, sedangkan data uji digunakan untuk mengetahui kemampuan model dalam melakukan prediksi terhadap

data yang tidak digunakan selama proses pelatihan. Parameter random_state=42 digunakan untuk menjaga konsistensi pembagian data sehingga proses eksperimen dapat direproduksi.

Algoritma pertama yang digunakan adalah K-Nearest Neighbor (KNN). KNN menentukan kelas suatu data berdasarkan kelas dari sejumlah tetangga yang memiliki jarak paling dekat dengan data tersebut. Dalam penelitian ini, nilai (K) diuji pada beberapa nilai, yaitu 3, 5, 7, 9, 11, 13, dan 15. Berdasarkan hasil pengujian, nilai (K=5) digunakan sebagai konfigurasi model KNN. Pengukuran jarak menggunakan Euclidean distance, dengan persamaan:

$$d\ (x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

Algoritma kedua adalah Gaussian Naive Bayes, yang melakukan klasifikasi berdasarkan pendekatan probabilistik dengan asumsi independensi bersyarat antarfitur terhadap kelas. Model ini digunakan untuk menghasilkan probabilitas setiap kelas berdasarkan karakteristik fitur yang dimiliki pemohon. Kedua algoritma dilatih menggunakan data latih yang sama dan diuji menggunakan data uji yang sama. Dengan demikian, perbedaan kinerja yang diperoleh terutama berasal dari karakteristik metode klasifikasi yang digunakan, bukan dari perbedaan pembagian data.

Evaluasi Kinerja

Kinerja kedua model dievaluasi menggunakan *accuracy, precision, recall, F1-score*, dan *confusion matrix*. Penggunaan beberapa metrik dilakukan karena akurasi saja belum cukup untuk menggambarkan kemampuan model dalam menangani masing-masing kelas, terutama ketika distribusi kelas tidak seimbang. Accuracy mengukur proporsi seluruh prediksi yang benar terhadap keseluruhan data:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision menunjukkan tingkat ketepatan prediksi positif:

$$Precision = \frac{TP}{TP + FP}$$

Recall menunjukkan kemampuan model dalam menemukan data yang sebenarnya termasuk dalam suatu kelas:

$$Recall = \frac{TP}{TP + FN}$$

Sedangkan F1-score merupakan rata-rata harmonis antara precision dan recall:

$$F1 = 2\ x\ \frac{Precision\ x\ Recall}{Precision + Recall}$$

Confusion matrix digunakan untuk melihat jumlah prediksi yang benar maupun salah pada masing-masing kelas, yang terdiri atas *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN).

Hasil dari seluruh metrik tersebut kemudian digunakan untuk membandingkan KNN dan Gaussian Naive Bayes. Model dengan kinerja yang lebih baik tidak ditentukan hanya berdasarkan accuracy, tetapi juga mempertimbangkan kemampuan model dalam mengenali kelas Layak dan Tidak Layak melalui *precision, recall, F1-score*, serta pola kesalahan yang ditunjukkan oleh *confusion matrix*.

Hasil

Karakteristik dan Distribusi Data

Dataset yang digunakan dalam penelitian terdiri atas 14.897 data pengajuan kredit. Berdasarkan hasil pembentukan target kelayakan, data terbagi menjadi dua kelas, yaitu Layak dan Tidak Layak. Distribusi masing-masing kelas disajikan pada Tabel 1.

Tabel 1. *Distribusi Data Berdasarkan Kelas Kelayakan Kredit*

Kelas	Jumlah	Persentase
Layak	14.352	96,34%
Tidak Layak	545	3,66%
Total	14.897	100%

Tabel 1 menunjukkan bahwa data didominasi oleh kelas Layak sebesar 96,34%, sedangkan kelas Tidak Layak hanya sebesar 3,66%. Kondisi tersebut menunjukkan adanya ketidakseimbangan kelas yang perlu diperhatikan dalam interpretasi hasil klasifikasi, terutama dalam menilai kemampuan model mendeteksi kelas Tidak Layak.

Hasil Pengujian K-Nearest Neighbor

Model K-Nearest Neighbor (KNN) diuji menggunakan beberapa nilai K, yaitu 3, 5, 7, 9, 11, 13, dan 15. Berdasarkan hasil pengujian, nilai K=5 digunakan sebagai model akhir. Evaluasi dilakukan menggunakan accuracy, precision, recall, dan F1-score.

Tabel 2. *Hasil Classification Report Model K-Nearest Neighbor*

Kelas	Precision	Recall	F1-Score	Support
Layak	0,9748	0,9698	0,9723	2.877
Tidak layak	0,2627	0,3010	0,2805	103
Accuracy			0,9466	2.980
Macro Average	0,6188	0,6354	0,6264	2.980
Weighted Average	0,9502	0,9466	0,9484	2.980

Berdasarkan Tabel 2, model KNN memperoleh *accuracy* sebesar 94,66% dan *weighted* F1-score sebesar 94,84%. Namun, kemampuan model dalam mengenali kelas Tidak Layak masih relatif rendah, dengan recall sebesar 30,10%.

Tabel 3. *Confusion Matrix Model K-Nearest Neighbor*

Aktual/prediksi	Layak	Tidak Layak
Layak	2.790	87
Tidak layak	72	31

Tabel 3 menunjukkan bahwa sebanyak 2.790 data Layak berhasil diklasifikasikan dengan benar, sedangkan 87 data Layak diklasifikasikan sebagai Tidak Layak. Pada kelas Tidak Layak, sebanyak 31 data berhasil dikenali dengan benar dan 72 data justru diklasifikasikan sebagai Layak.

Hasil Klasifikasi Gaussian Naive Bayes

Model berikutnya adalah Gaussian Naive Bayes. Model ini dievaluasi menggunakan metrik yang sama dengan KNN untuk memperoleh perbandingan kinerja yang konsisten.

Tabel 4. *Hasil Classification Report Model Gaussian Naive Bayes*

Kelas	Precision	Recall	F1-Score	Support
Layak	1,0000	0,9079	0,9517	2.877
Tidak layak	0,2799	1,0000	0,4374	103
Accuracy			0,9111	2.980
Macro Average	0,6399	0,9539	0,6945	2.980
Weighted Average	0,9751	0,9111	0,9339	2.980

Tabel 4 menunjukkan bahwa Gaussian Naive Bayes memperoleh accuracy sebesar 91,11%. Meskipun accuracy lebih rendah dibandingkan KNN, model ini menghasilkan recall sebesar 100% pada kelas Tidak Layak. Hal tersebut menunjukkan bahwa seluruh data Tidak Layak pada data pengujian berhasil teridentifikasi oleh model.

Tabel 5. *Confusion Matrix Gaussian Naive Bayes*

Aktual/prediksi	Layak	Tidak layak
Layak	2.612	265
Tidak Layak	0	103

Berdasarkan Tabel 5, seluruh 103 data Tidak Layak berhasil diklasifikasikan dengan benar. Namun, terdapat 265 data Layak yang diklasifikasikan sebagai Tidak Layak. Kondisi tersebut menjelaskan mengapa recall kelas Tidak Layak mencapai 100%, tetapi precision kelas tersebut hanya sebesar 27,99%.

Perbandingan Kinerja Model

Perbandingan kinerja kedua model ditampilkan pada Tabel 6.

Tabel 6. *Perbandingan Kinerja Model K-Nearest Neighbor dan Gaussian Naive Bayes*

Metrik	KNN	*Gaussian Naive Bayes*
Accuracy	94,66%	91,11%
Weighted Precision	95,02%	97,51%
Weighted Recall	94,66%	91,11%
Weighted F1-Score	94,84%	93,39%
Recall Tidak Layak	30,10%	100,00%
F1-Score Tidak Layak	28,05%	43,74%

Tabel 6 menunjukkan bahwa KNN memiliki accuracy dan weighted F1-score yang lebih tinggi dibandingkan Gaussian Naive Bayes. Sebaliknya, Gaussian Naive Bayes memiliki kemampuan yang lebih baik dalam mendeteksi kelas Tidak Layak, yang ditunjukkan oleh recall sebesar 100% dan F1-score sebesar 43,74%. Dengan demikian, KNN lebih unggul dari sisi kinerja agregat, sedangkan Gaussian Naive Bayes lebih unggul dalam mendeteksi kelompok kredit yang berisiko.

Pembahasan

Pengaruh Ketidakseimbangan Kelas terhadap Hasil Klasifikasi

Hasil penelitian menunjukkan bahwa distribusi kelas pada dataset memiliki ketidakseimbangan yang cukup ekstrem. Dari 14.897 data, sebanyak 14.352 data (96,34%) termasuk kategori Layak, sedangkan hanya 545 data (3,66%) termasuk kategori Tidak

Layak. Kondisi tersebut menjadi faktor penting dalam menjelaskan perbedaan kinerja K-Nearest Neighbor (KNN) dan Gaussian Naive Bayes pada penelitian ini. Dominasi kelas Layak menyebabkan sebagian besar pola yang dipelajari model berasal dari karakteristik kelompok tersebut, sehingga akurasi keseluruhan dapat terlihat tinggi meskipun kemampuan mendeteksi kelas minoritas belum tentu optimal. Kondisi serupa juga ditemukan dalam penelitian klasifikasi data tidak seimbang di Indonesia, di mana dominasi kelas mayoritas dapat menyebabkan model cenderung mengikuti karakteristik kelas tersebut dan mengurangi kemampuan dalam mengenali kelas minoritas (Ulfah et al., 2023).

Temuan ini sejalan dengan penelitian terdahulu yang menekankan bahwa evaluasi sistem klasifikasi dengan distribusi kelas yang tidak seimbang tidak cukup hanya menggunakan akurasi. Dalam penelitian ini, ketidakseimbangan kelas terlihat jelas dari confusion matrix. KNN mampu memperoleh akurasi 94,66%, tetapi recall untuk kelas Tidak Layak hanya sebesar 30,10%. Artinya, dari 103 data Tidak Layak pada data pengujian, hanya 31 data yang berhasil dikenali dengan benar, sedangkan 72 data lainnya justru diprediksi sebagai Layak. Kondisi tersebut menunjukkan bahwa tingginya akurasi KNN terutama dipengaruhi oleh keberhasilannya mengklasifikasikan kelas mayoritas. Hasil tersebut memiliki kesesuaian dengan penelitian lain yang membandingkan Naive Bayes, K-Nearest Neighbor, dan *Neural Network* pada data *class-imbalanced*, serta menunjukkan bahwa ketidakseimbangan kelas menjadi persoalan penting dalam menentukan kemampuan model dalam mengenali kelas minoritas (Sahroni et al., 2021).

Rendahnya recall kelas Tidak Layak pada KNN menunjukkan bahwa sebagian besar kesalahan klasifikasi terjadi pada kelas yang justru memiliki proporsi lebih kecil. Dalam konteks penilaian kelayakan kredit, kondisi ini perlu mendapatkan perhatian karena kesalahan dalam mengidentifikasi calon nasabah yang sebenarnya Tidak Layak sebagai Layak berpotensi menimbulkan risiko kredit. Oleh karena itu, model dengan akurasi tinggi belum tentu menjadi model yang paling sesuai apabila tujuan utama penelitian adalah meningkatkan kemampuan dalam mendeteksi kelas minoritas. Penelitian mengenai penanganan ketidakseimbangan data menggunakan Naive Bayes, SVM, dan *Random Forest* juga menunjukkan bahwa evaluasi menggunakan *accuracy, precision, recall*, dan F1-score diperlukan untuk melihat kemampuan model secara lebih menyeluruh pada setiap kelas (Kusuma et al., 2023).

Fenomena tersebut memperkuat temuan penelitian terdahulu bahwa distribusi data yang tidak seimbang dapat membuat model lebih banyak mengikuti pola kelas dominan. Oleh karena itu, penggunaan precision, recall, F1-score, dan confusion matrix menjadi penting untuk memperoleh gambaran kinerja model yang lebih komprehensif. Dalam penelitian ini, recall menjadi indikator yang sangat penting karena menunjukkan kemampuan model dalam menemukan data yang benar-benar termasuk kategori Tidak Layak. Nilai recall yang rendah pada kelas tersebut menunjukkan bahwa masih terdapat sejumlah besar data minoritas yang belum berhasil dikenali oleh model, meskipun nilai akurasi keseluruhan relatif tinggi.

Hasil tersebut juga memberikan gambaran bahwa ketidakseimbangan kelas perlu dipertimbangkan sejak tahap pengolahan data hingga evaluasi model. Beberapa penelitian nasional menunjukkan bahwa pendekatan seperti *oversampling*, *undersampling*, maupun teknik sintetis seperti SMOTE dan MWMOTE dapat digunakan untuk mengurangi dominasi kelas mayoritas dan meningkatkan kemampuan model dalam

mengenali kelas minoritas (Ulfah et al., 2023; Kusuma et al., 2023). Namun, penelitian ini tidak secara langsung menerapkan teknik penyeimbangan kelas, sehingga hasil yang diperoleh dapat digunakan sebagai gambaran kinerja awal KNN dan Gaussian Naive Bayes terhadap kondisi dataset asli.

Dengan demikian, hasil penelitian menunjukkan bahwa akurasi sebesar 94,66% pada KNN perlu diinterpretasikan secara hati-hati karena masih terdapat kelemahan dalam mendeteksi kelas Tidak Layak. Tingginya akurasi tidak secara otomatis menunjukkan bahwa model telah mampu memberikan klasifikasi yang optimal terhadap seluruh kelas. Dalam konteks klasifikasi kelayakan kredit, model dengan kemampuan mendeteksi kelas minoritas yang lebih baik dapat memiliki nilai praktis yang lebih tinggi meskipun akurasinya sedikit lebih rendah. Oleh sebab itu, pemilihan model tidak hanya didasarkan pada nilai akurasi tertinggi, tetapi juga mempertimbangkan precision, recall, F1-score, serta distribusi kesalahan pada confusion matrix.

Kinerja KNN dan Perbedaannya dengan Pendekatan Terdahulu

KNN menghasilkan akurasi sebesar 94,66%, weighted precision 95,02%, weighted recall 94,66%, dan weighted F1-score 94,84%. Kinerja tersebut menunjukkan bahwa KNN mampu melakukan klasifikasi keseluruhan dengan baik pada data kelayakan kredit yang digunakan. Karakteristik KNN yang menentukan kelas berdasarkan kedekatan suatu data dengan sejumlah tetangga terdekat memungkinkan pola karakteristik pemohon yang memiliki kemiripan dengan data historis dapat dimanfaatkan secara langsung dalam proses prediksi. Kemampuan KNN dalam memanfaatkan kemiripan karakteristik antarobservasi juga telah ditunjukkan pada penelitian klasifikasi pemberian pinjaman nasabah di Indonesia (Saputra et al., 2024).

Hasil tersebut memiliki relevansi dengan penelitian lain yang juga membahas persoalan rekomendasi kelayakan penerima kredit (Sholihaningtias, 2023). Penelitian tersebut menggunakan TOPSIS dengan pembobotan ROC dan mempertimbangkan beberapa kriteria seperti harga kendaraan, uang muka, penghasilan, karakter nasabah, dan status tempat tinggal. Hasil penelitian Sholihaningtias menunjukkan bahwa pendekatan berbasis kriteria dapat digunakan untuk menghasilkan rekomendasi kelayakan kredit.

Perbedaannya terletak pada mekanisme pengambilan keputusan. TOPSIS-ROC menentukan kelayakan melalui bobot dan preferensi kriteria yang telah ditentukan, sedangkan penelitian ini menggunakan pembelajaran dari pola data historis melalui KNN dan Gaussian Naive Bayes. Dengan demikian, penelitian ini melengkapi pendekatan berbasis sistem pendukung keputusan tersebut dengan pendekatan *machine learning* yang memungkinkan pola hubungan antardata dipelajari dari dataset historis. Pendekatan KNN juga telah digunakan dalam penelitian nasional sebagai metode klasifikasi yang memanfaatkan kedekatan atau kemiripan karakteristik data untuk menentukan kelas suatu objek (Yustikasari et al., 2022).

Namun, hasil penelitian ini juga menunjukkan keterbatasan KNN ketika digunakan pada data dengan distribusi kelas yang tidak seimbang. Meskipun KNN memperoleh akurasi tertinggi, recall kelas Tidak Layak hanya 30,10%. Dengan kata lain, model masih berpotensi meloloskan sebagian besar pemohon yang sebenarnya termasuk kategori Tidak Layak. Dalam konteks kredit, kesalahan tersebut perlu mendapatkan perhatian karena *false negative* pada kelas Tidak Layak dapat berhubungan dengan risiko pemberian

kredit kepada pemohon yang seharusnya ditolak. Kondisi tersebut menunjukkan bahwa akurasi yang tinggi belum tentu mencerminkan kemampuan model dalam mengenali kelas minoritas secara optimal (Afandi et al., 2024).

Kinerja Gaussian Naive Bayes dan Perspektif Manajemen Risiko

Gaussian Naive Bayes menghasilkan akurasi 91,11%, lebih rendah dibandingkan KNN. Akan tetapi, karakteristik hasil prediksinya berbeda secara signifikan. Model ini menghasilkan recall sebesar 100% untuk kelas Tidak Layak, yang berarti seluruh data Tidak Layak pada data pengujian berhasil teridentifikasi. Sebaliknya, precision kelas Tidak Layak hanya sebesar 27,99%, karena terdapat 265 data Layak yang diklasifikasikan sebagai Tidak Layak. Kondisi tersebut menunjukkan bahwa Gaussian Naive Bayes lebih sensitif dalam mendeteksi kelas Tidak Layak, meskipun konsekuensinya adalah meningkatnya jumlah prediksi positif yang tidak tepat.

Temuan tersebut menunjukkan bahwa Gaussian Naive Bayes memiliki kecenderungan lebih konservatif dalam melakukan klasifikasi risiko. Model berhasil menghindari *false negative* pada kelas Tidak Layak, tetapi konsekuensinya adalah meningkatnya *false positive*. Dalam konteks pengambilan keputusan kredit, karakteristik seperti ini memiliki dua sisi. Dari perspektif mitigasi risiko, kemampuan menemukan seluruh nasabah yang masuk kategori Tidak Layak merupakan keunggulan karena tidak ada kasus Tidak Layak yang terlewat pada data pengujian. Namun, dari perspektif bisnis, terlalu banyak nasabah Layak yang ditolak dapat menyebabkan hilangnya peluang pemberian kredit kepada calon nasabah yang sebenarnya memiliki kemampuan pembayaran. Pendekatan konservatif seperti ini dapat dipertimbangkan ketika konsekuensi kesalahan dalam memberikan kredit kepada pemohon berisiko dinilai lebih besar daripada kehilangan peluang terhadap pemohon yang sebenarnya layak.

Hasil tersebut memperlihatkan bahwa pemilihan model tidak dapat ditentukan hanya berdasarkan nilai akurasi. Jika tujuan utama lembaga keuangan adalah memperoleh klasifikasi umum dengan tingkat kesalahan agregat yang rendah, KNN lebih unggul dalam penelitian ini. Sebaliknya, apabila prioritas utama adalah meminimalkan kemungkinan nasabah berisiko diklasifikasikan sebagai Layak, Gaussian Naive Bayes memberikan karakteristik yang lebih sesuai karena recall kelas Tidak Layak mencapai 100%. Hal ini menunjukkan bahwa pemilihan model perlu mempertimbangkan biaya atau konsekuensi dari masing-masing jenis kesalahan klasifikasi dalam proses pengambilan keputusan kredit.

Temuan ini juga memperlihatkan bahwa performa suatu algoritma sangat bergantung pada karakteristik dataset dan tujuan klasifikasi. Perbandingan Naive Bayes dan KNN pada analisis sentimen aplikasi Mitra Darat menunjukkan bahwa Naive Bayes memperoleh akurasi 99,28%, sedangkan KNN 80% (Wijaya et al., 2024). Perbedaan hasil tersebut dengan penelitian ini menunjukkan bahwa tidak terdapat satu algoritma yang selalu menghasilkan kinerja terbaik untuk seluruh permasalahan klasifikasi. Naive Bayes dapat unggul pada karakteristik data tertentu, sedangkan pada data kelayakan kredit dalam penelitian ini KNN menghasilkan akurasi yang lebih tinggi. Perbedaan kinerja KNN dan Naive Bayes pada berbagai karakteristik data juga ditemukan dalam penelitian klasifikasi pemberian pinjaman nasabah di Indonesia (Saputra et al., 2024).

Dengan demikian, hasil penelitian ini memperkuat pentingnya pemilihan algoritma berdasarkan karakteristik data, distribusi kelas, jenis fitur, dan konsekuensi kesalahan

klasifikasi, bukan hanya berdasarkan reputasi atau hasil penelitian pada domain yang berbeda. Dalam konteks manajemen risiko kredit, model yang dipilih sebaiknya disesuaikan dengan prioritas perusahaan, apakah lebih menekankan ketepatan klasifikasi secara keseluruhan atau kemampuan mendeteksi pemohon yang berisiko.

Perbandingan dengan Penelitian Terdahulu dan Research Gap

Penilaian kelayakan kredit dapat dilakukan menggunakan pendekatan sistem pendukung keputusan berbasis kriteria dengan mempertimbangkan beberapa atribut seperti harga kendaraan, uang muka, penghasilan, karakter pelanggan, dan status tempat tinggal untuk menghasilkan rekomendasi penerima kredit (Sholihaningtias, 2023). Penelitian tersebut memiliki hubungan yang erat dengan penelitian ini karena sama-sama bertujuan membantu proses penilaian kelayakan kredit, tetapi menggunakan pendekatan yang berbeda.

Kontribusi penelitian ini terletak pada penggunaan dua algoritma klasifikasi, yaitu KNN dan Gaussian Naive Bayes, untuk melakukan evaluasi komparatif pada data historis pemohon kredit. Selain itu, penelitian ini tidak hanya menggunakan atribut administratif, tetapi juga membentuk fitur turunan yang berkaitan langsung dengan kemampuan finansial, seperti *Debt Service Ratio* (DSR), *disposable income*, dan *Loan to Value* (LTV). Dengan demikian, informasi dari data pemohon diubah menjadi indikator kuantitatif yang dapat digunakan oleh algoritma untuk mengenali pola kelayakan kredit. Pendekatan *feature engineering* semacam ini penting karena pembentukan fitur yang relevan dapat membantu model menangkap karakteristik risiko dari data kredit secara lebih informatif.

Perbandingan dengan penelitian terdahulu juga memperlihatkan bahwa KNN dan Naive Bayes telah digunakan secara bersamaan untuk permasalahan klasifikasi. Namun, penelitian tersebut berada pada domain analisis sentimen berbasis data tekstual, sedangkan penelitian ini menerapkan kedua algoritma pada data tabular keuangan (Wijaya et al., 2024). Perbedaan domain tersebut penting karena karakteristik data dapat memengaruhi performa algoritma. Oleh karena itu, hasil penelitian ini memberikan bukti empiris bahwa performa KNN dan Naive Bayes perlu dievaluasi secara spesifik pada konteks kelayakan kredit, bukan diasumsikan berdasarkan hasil pada domain lain.

Pengelolaan risiko pada sektor keuangan juga telah dikaji dari perspektif pengawasan terhadap *fraud* dalam bisnis investasi dan peran Otoritas Jasa Keuangan serta hukum ekonomi Islam (Pamungkas & Zulfikar, 2021). Penelitian tersebut berbeda secara tujuan dan metodologi dengan penelitian ini, tetapi memiliki keterkaitan pada aspek pengelolaan risiko di sektor keuangan. Penelitian ini memperluas perspektif tersebut ke ranah komputasional dengan memanfaatkan data pemohon dan algoritma *machine learning* untuk mendukung proses identifikasi risiko kredit secara individual.

Berdasarkan perbandingan tersebut, *research gap* penelitian ini tidak hanya terletak pada penggunaan KNN dan Naive Bayes, tetapi juga pada pengujian kedua algoritma secara langsung pada data kelayakan kredit dengan mempertimbangkan fitur finansial turunan dan evaluasi kelas minoritas. Penelitian ini menunjukkan bahwa evaluasi model kredit harus mempertimbangkan *trade-off* antara performa klasifikasi keseluruhan dan kemampuan mendeteksi nasabah berisiko. Dengan demikian, penelitian ini memberikan kontribusi empiris dalam menentukan algoritma yang lebih sesuai berdasarkan karakteristik data dan konsekuensi kesalahan klasifikasi pada penilaian kelayakan kredit.

Implikasi terhadap Penerapan Machine Learning dalam Kelayakan Kredit

Hasil penelitian memberikan implikasi bahwa sistem klasifikasi kredit berbasis *machine learning* sebaiknya tidak diposisikan sebagai pengganti penuh keputusan analis kredit, tetapi sebagai alat bantu pengambilan keputusan. KNN memberikan performa klasifikasi keseluruhan yang lebih tinggi, sehingga lebih sesuai apabila tujuan utama adalah memperoleh prediksi kelayakan secara umum. Namun, recall kelas Tidak Layak yang hanya 30,10% menunjukkan bahwa penggunaan KNN secara langsung untuk keputusan otomatis memiliki risiko apabila kesalahan *false negative* dianggap penting.

Gaussian Naive Bayes memberikan perspektif yang berbeda. Meskipun akurasinya lebih rendah, kemampuannya mendeteksi seluruh data Tidak Layak menunjukkan bahwa model tersebut dapat digunakan sebagai komponen penyaringan awal untuk mengidentifikasi pemohon yang berpotensi berisiko. Akan tetapi, tingginya jumlah *false positive* juga menunjukkan bahwa hasil model perlu diverifikasi lebih lanjut agar pemohon yang sebenarnya layak tidak langsung ditolak.

Implikasi tersebut sejalan dengan rekomendasi penelitian terdahulu dalam tesis yang menekankan perlunya penanganan ketidakseimbangan kelas, *hyperparameter tuning*, dan validasi yang lebih kuat. Salah satu pengembangan yang relevan adalah penggunaan teknik oversampling seperti SMOTE pada data pelatihan. Pendekatan tersebut berpotensi meningkatkan representasi kelas minoritas sehingga model tidak terlalu dipengaruhi oleh dominasi kelas Layak. Selain itu, *cross-validation* dan optimasi parameter dapat digunakan untuk menguji kestabilan model pada pembagian data yang berbeda.

Secara keseluruhan, penelitian ini menunjukkan bahwa KNN merupakan model yang lebih unggul dari sisi akurasi dan metrik agregat pada dataset yang digunakan, sedangkan Gaussian Naive Bayes memiliki keunggulan dari sisi sensitivitas terhadap kelas Tidak Layak. Oleh karena itu, hasil penelitian tidak semata-mata menghasilkan satu algoritma sebagai pemenang, tetapi menunjukkan adanya trade-off antara akurasi keseluruhan dan kemampuan mendeteksi risiko. Temuan tersebut menjadi dasar penting dalam merancang sistem *credit scoring*, karena model yang dipilih seharusnya disesuaikan dengan konsekuensi bisnis dari masing-masing jenis kesalahan klasifikasi.

Kesimpulan

Penelitian ini menunjukkan bahwa metode *machine learning* dapat digunakan untuk melakukan klasifikasi kelayakan kredit berdasarkan karakteristik data pengajuan dan hasil rekayasa fitur yang meliputi *Debt Service Ratio* (DSR), *disposable income*, *Loan to Value* (LTV), serta indikator usia, status rumah, tunggakan, dan LTV. Berdasarkan pengujian pada 14.897 data, kelas Layak memiliki proporsi sebesar 96,34%, sedangkan kelas Tidak Layak sebesar 3,66%, sehingga terdapat ketidakseimbangan kelas yang perlu menjadi perhatian dalam interpretasi hasil klasifikasi.

Hasil evaluasi menunjukkan bahwa K-Nearest Neighbor (KNN) dengan nilai K=5 memperoleh kinerja agregat terbaik dengan accuracy sebesar 94,66% dan *weighted F1-score* sebesar 94,84%. Namun, kemampuan KNN dalam mendeteksi kelas Tidak Layak masih terbatas, dengan recall sebesar 30,10%. Sementara itu, Gaussian Naive Bayes memperoleh accuracy sebesar 91,11% dan *weighted F1-score* sebesar 93,39%, tetapi mampu mencapai recall 100% pada kelas Tidak Layak. Temuan ini menunjukkan bahwa pemilihan algoritma tidak hanya dapat didasarkan pada accuracy, tetapi juga perlu

mempertimbangkan karakteristik kelas yang menjadi prioritas dalam pengambilan keputusan kredit.

Dengan demikian, KNN lebih sesuai apabila prioritas utama adalah kinerja klasifikasi secara keseluruhan, sedangkan Gaussian Naive Bayes lebih relevan ketika kemampuan mendeteksi seluruh calon debitur dalam kategori Tidak Layak menjadi pertimbangan utama. Penelitian selanjutnya dapat menguji teknik penanganan ketidakseimbangan kelas dan validasi silang untuk meningkatkan kemampuan model dalam mengenali kelas minoritas.

Acknowledgment

-

Daftar Pustaka

Afandi, A., Agtriadi, H. B., Luqman, L., & Susanti, M. (2024). Advanced credit scoring with Naive Bayes algorithm: Improving accuracy and reliability in financial risk assessment. *Jurnal E-Komtek, 8*(2), 399–409. https://doi.org/10.37339/e-komtek.v8i2.2160

Amarnadh, V., & Moparthi, N. R. (2023). Comprehensive review of different artificial intelligence-based methods for credit risk assessment in data science. *Intelligent Decision Technologies, 17*(4). https://doi.org/10.3233/IDT-230190

Anwar, M., Nidar, S. R., Widianto, S., & Layyinaturrobaniyah. (2021). Lending rationale of financial institution and fintech on the perception of borrowers: A qualitative study. *Jurnal Dinamika Manajemen, 12*(2), 291–309. https://doi.org/10.15294/jdm.v12i2.31896

Chen, Y., Calabrese, R., & Martin-Barragan, B. (2024). Interpretable machine learning for imbalanced credit scoring datasets. *European Journal of Operational Research, 312*(1), 357–372. https://doi.org/10.1016/j.ejor.2023.06.036

Dumitrescu, E., Hué, S., Hurlin, C., & Tokpavi, S. (2022). Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. *European Journal of Operational Research, 297*(3), 1178–1192. https://doi.org/10.1016/j.ejor.2021.06.053

Emmanuel, I., Sun, Y., & Wang, Z. (2024). A machine learning-based credit risk prediction engine system using a stacked classifier and a filter-based feature selection method. *Journal of Big Data, 11*, Article 23. https://doi.org/10.1186/s40537-024-00882-0

Hayashi, Y. (2022). Emerging trends in deep learning for credit scoring: A review. *Electronics, 11*(19), 3181. https://doi.org/10.3390/electronics11193181

Jemai, J., & Zarrad, A. (2023). Feature selection engineering for credit risk assessment in retail banking. *Information, 14*(3), 200. https://doi.org/10.3390/info14030200

Khayatun, S., Sutrisno, C. T. S., & Suharti, M. (2021). Pengaruh prinsip 5C terhadap pemahaman kredit pada Kantor Pusat PT BPR BKK Pati (Perseroda) Kabupaten Pati. *Excellent, 8*(2). https://doi.org/10.36587/exc.v8i2.1118

Kusuma, I. G. N. A., Pradipta, I. M., Santosa, I. M. A., & Dharmendra, I. K. (2023). Penanganan ketidakseimbangan data pada klasifikasi pengaduan masyarakat. *Jurnal Teknologi Informasi dan Komputer, 9*(5). https://doi.org/10.36002/jutik.v9i5.2643

Lenka, S. R., Bisoy, S. K., Priyadarshini, R., & Sain, M. (2022). Empirical analysis of ensemble learning for imbalanced credit scoring datasets: A systematic review. *Wireless Communications and Mobile Computing, 2022*, 6584318, 1–18. https://doi.org/10.1155/2022/6584318

Liu, W., Fan, H., & Xia, M. (2022). Credit scoring based on tree-enhanced gradient boosting decision trees. *Expert Systems with Applications, 189*, 116034. https://doi.org/10.1016/j.eswa.2021.116034

Liu, Y., Huang, F., Ma, L., Zeng, Q., & Shi, J. (2024). Credit scoring prediction leveraging interpretable ensemble learning. *Journal of Forecasting, 43*(2), 286–308. https://doi.org/10.1002/for.3033

Misral, M., & Finowa, R. (2021). Penilaian pemberian kredit untuk meminimalisir resiko kredit pada Bank BRI Unit Tampan Kota Pekanbaru. *Jurnal Akuntansi dan Ekonomika, 11*(1), 139–149. https://doi.org/10.37859/jae.v11i1.2558

Muangthanang, C., Mungsing, S., & Chirawichitcha, N. (2024). Credit risk prediction model using feature engineering and machine learning techniques. *International Scientific Journal of Engineering and Technology, 8*(1), 19–26. https://doi.org/10.25126/isjet.202512499

Nwafor, C. N., Nwafor, O., & Brahma, S. (2024). Enhancing transparency and fairness in automated credit decisions: An explainable novel hybrid machine learning approach. *Scientific Reports, 14*, 25174. https://doi.org/10.1038/s41598-024-75026-8

Pamungkas, F. T., & Zulfikar, A. A. (2021). Peran Otoritas Jasa Keuangan (OJK) dalam mengawasi adanya *fraud* dalam bisnis investasi dalam perspektif hukum ekonomi Islam. *Jurnal Penegakan Hukum dan Keadilan, 2*(1), 19–40. https://doi.org/10.18196/jphk.v2i1.9507

Quan, J., & Sun, X. (2024). Credit risk assessment using the factorization machine model with feature interactions. *Humanities and Social Sciences Communications, 11*, 234. https://doi.org/10.1057/s41599-024-02700-7

Rayhan, M. A. (2026). Implementation of Naive Bayes and Support Vector Machine for SMS Spam Classification Using the SMS Spam Collection Dataset. *Journal Artificial: Informatika Dan Sistem Informasi*, *4*(2), 197–212. https://doi.org/10.54065/artificial.1249

Sahroni, M. Y., Setifani, N. A., & Fitriana, D. N. (2021). Analisis perbandingan algoritma Naïve Bayes, k-Nearest Neighbor dan Neural Network untuk permasalahan class-imbalanced data pada kasus credit card fraud dataset. *Teknologi: Jurnal Ilmiah Sistem Informasi, 11*(2). https://doi.org/10.26594/teknologi.v11i2.2393

Santoso, T., Hardiawan, D., Dwiputra, M., Sapulette, M., & Setiawan, M. (2023). Evidence of credit rationing in Indonesia: Income class and collateral. *Economics Development Analysis Journal, 11*(4), 507–523. https://doi.org/10.15294/edaj.v11i4.55287

Saputra, Z., Supahri, H. A., Ismanizan, R., & Rahmaddeni. (2024). Perbandingan algoritma KNN (K-Nearest Neighbors), Naïve Bayes, dan SVM (Support Vector Machine) untuk klasifikasi pemberian pinjaman nasabah. *Jurnal Ilmiah Sistem Informasi dan Teknik Informatika (JISTI), 7*(1), 67–75. https://doi.org/10.57093/jisti.v7i1.182

Shi, S., Tse, R., Luo, W., D'Addona, S., & Pau, G. (2022). Machine learning-driven credit risk: A systemic review. *Neural Computing and Applications, 34*, 14327–14339. https://doi.org/10.1007/s00521-022-07472-2

Sholihaningtias, D. N. (2023). Rekomendasi kelayakan penerima kredit menggunakan metode TOPSIS dengan pembobotan ROC. *Jurnal SAINTEKOM: Sains, Teknologi, Komputer dan Manajemen, 13*(1), 88–99. https://doi.org/10.33020/saintekom.v13i1.376

Ulfah, M., Saragih, T. H., Kartini, D., Mazdadi, M. I., & Abadi, F. (2023). Penerapan MWMOTE untuk mengatasi ketidakseimbangan kelas pada klasifikasi risiko kredit. *Jurnal Informatika Polinema, 9*(4), 477–486. https://doi.org/10.33795/jip.v9i4.1331

Wafda, A. (2024). Integrasi Machine Learning dalam Ritel: Tinjauan Komprehensif tentang Prediksi Harga, Analisis Data Pelanggan, dan Pemanfaatan Media Sosial. *Journal Artificial: Informatika Dan Sistem Informasi*, *2*(2), 90–106. https://doi.org/10.54065/artificial.543

Wijaya, A., Rivaldo, M., & Pribadi, M. R. (2024). Analisis sentimen terhadap aplikasi Mitra Darat menggunakan algoritma Naive Bayes Classifier dan K-Nearest Neighbor. *Informatik: Jurnal Ilmu Komputer, 20*(3), 112–117. https://doi.org/10.52958/iftk.v20i3.7967

Xia, Y., Li, Y., He, L., Xu, Y., & Meng, Y. (2021). Incorporating multilevel macroeconomic variables into credit scoring for online consumer lending. *Electronic Commerce Research and Applications, 49*, 101095. https://doi.org/10.1016/j.elerap.2021.101095

Yang, C., Dong, Y., Lu, J., & Peng, Z. (2023). Solving imbalanced data in credit risk prediction: A comparison of resampling strategies for different machine learning classification algorithms, taking threshold tuning into account. *Proceedings of the 2022 5th International Conference on Machine Learning and Machine Intelligence*, 30–40. https://doi.org/10.1145/3568199.3568204

Yustikasari, Y., Mubarok, H., & Rianto, R. (2022). Comparative analysis performance of K-Nearest Neighbor algorithm and Adaptive Boosting on the prediction of non-cash food aid recipients. *Scientific Journal of Informatics, 9*(2), 179–188. https://doi.org/10.15294/sji.v9i2.32369

Zhang, X., & Yu, L. (2024). Consumer credit risk assessment: A review from the state-of-the-art classification algorithms, data traits, and learning methods. *Expert Systems with Applications, 237*, 121484. https://doi.org/10.1016/j.eswa.2023.121484

.