

Komparasi Kinerja Algoritma Logistic Regression, Random Forest, dan Naïve Bayes dalam Klasifikasi Risiko Kredit

Dewi Ayu Nur Wulandari¹, Omar Pahlevi^{2*}, Yopi Handrianto³, Ichsan Ramdhani⁴

¹Sistem Informasi, Universitas Bina Sarana Informatika
Jakarta, Indonesia

^{2,3}Teknologi Informasi, Universitas Bina Sarana Informatika
Jakarta, Indonesia

⁴Teknik Informatika, Univeritas Pamulang
Tangerang Selatan, Indonesia

Correspondence e-mail: omar.opi@bsi.ac.id

ABSTRAK

Klasifikasi risiko kredit merupakan proses penting dalam industri keuangan untuk mengidentifikasi nasabah yang berpotensi mengalami gagal bayar. Penelitian ini bertujuan membandingkan kinerja algoritma supervised learning antara lain Logistic Regression, Random Forest, dan Naïve Bayes. Pada klasifikasi risiko kredit ini menggunakan dataset German Credit yang terdiri dari 1.000 data nasabah. Tahapan penelitian meliputi data preprocessing, normalisasi data, pembagian data latih dan data uji dengan rasio 80:20, pemodelan menggunakan KNIME Analytics Platform 5.9.0, serta evaluasi menggunakan Accuracy, Precision, Recall, F1-Score, Confusion Matrix, dan Cohen's Kappa. Fokus evaluasi penelitian diarahkan pada kemampuan model dalam mendeteksi kelas bad sebagai representasi nasabah berisiko. Hasil pengujian menunjukkan bahwa Random Forest memberikan performa terbaik dengan akurasi 81,50%, precision 0,94, recall 0,82, F1-Score 0,87, dan Cohen's Kappa 0,51. Logistic Regression memperoleh akurasi 76,10% dan Naïve Bayes sebesar 75,50%. Temuan ini menunjukkan bahwa Random Forest lebih efektif dalam mengidentifikasi nasabah berisiko dibandingkan dua algoritma lainnya. Kontribusi penelitian terletak pada pendekatan evaluasi berbasis risiko yang menekankan kemampuan deteksi kelas bad sebagai dasar rekomendasi model klasifikasi risiko kredit.

Kata Kunci: klasifikasi, komparasi, risiko kredit, supervised learning

ABSTRACT

Credit risk classification is a critical process in the financial industry for identifying customers at risk of default. This study aims to compare the performance of supervised learning algorithms, including Logistic Regression, Random Forest, and Naïve Bayes. For this credit risk classification, the German Credit dataset, consisting of 1,000 customer records, was used. The research stages include data preprocessing, data normalization, splitting the data into training and test sets at an 80:20 ratio, modeling using the KNIME Analytics Platform 5.9.0, and evaluation using Accuracy, Precision, Recall, F1-Score, Confusion Matrix, and Cohen's Kappa. The evaluation focused on the models' ability to detect the "bad" class, which represents high-risk customers. The test results showed that Random Forest delivered the best performance with an accuracy of 81.50%, precision of 0.94, recall of 0.82, F1-Score of 0.87, and Cohen's Kappa of 0.51. Logistic Regression achieved an accuracy of 76.10%, and Naïve Bayes achieved 75.50%. These findings indicate that Random Forest is more effective at identifying at-risk customers than the other two algorithms. The contribution of this research lies in its risk-based evaluation approach, which emphasizes the ability to detect the bad class as the basis for credit risk classification model recommendations.

Keywords: classification, comparison, credit risk, supervised learning



1. PENDAHULUAN

Analisis risiko kredit yang dilakukan secara manual seringkali membutuhkan waktu lebih lama dan mungkin kurang akurat. Kondisi perekonomian saat ini menawarkan prospek yang menjanjikan untuk melakukan penilaian risiko kredit secara lebih akurat dan efisien dengan memanfaatkan perkembangan dalam teknologi *machine learning* (Septama et al., 2023). Pendekatan konvensional dalam analisis risiko keuangan umumnya bersifat manual dan kurang optimal dalam menangani data berukuran besar serta pola yang kompleks. Bersamaan dengan tren teknologi maju, penerapan *machine learning* menjadi solusi yang banyak digunakan karena kemampuannya dalam melakukan klasifikasi dan prediksi berbasis data historis secara otomatis (Sakina & Damayanti, 2025).

Pembuatan profil risiko pelanggan merupakan salah satu prosedur dalam mengajukan kredit. Kapabilitas pihak debitur ini menentukan layak atau tidaknya proses cairnya pinjaman kredit. Proses pengajuan kredit bisa diverifikasi melalui kriteria psikologis dan sikap, seperti riwayat kredit. Pendapatan, usia, jumlah tanggungan, status perkawinan, dan pengeluaran merupakan aspek-aspek yang dinilai. Salah satu hambatan dalam analisa profil risiko dikarenakan kinerja proses penilaian kredit yang dilakukan secara tradisional (Bhandary & Ghosh, 2025). Untuk memberikan peluang atau probabilitas statistik dalam mengantisipasi risiko gagal bayar terkait dengan berbagai elemen lainnya, kriteria penilaian kredit menjadi isu penting. Penilaian kredit merupakan indikator penting bagi bank untuk memastikan status keuangan dan kelayakan kredit (Baser et al., 2023).

Evaluasi keputusan pemberian kredit mulai meningkatkan metode penilaian manual dan analisis statistik, yang secara signifikan meningkatkan efisiensi dan kehandalan keputusan pemberian kredit. Faktor-faktor ini disebabkan dengan adanya perluasan *database* lembaga keuangan yang sangat besar. Model algoritma *machine learning* dengan kehandalannya berupa perhitungan statistika yang akurat dalam membantu pengambilan keputusan persetujuan kredit menjadi indikator yang penting (Zhao, 2022). Risiko kredit disebabkan oleh ketidakmampuan debitur untuk membayar kembali atau memenuhi kewajiban utangnya sesuai kesepakatan, atau oleh penurunan ekonomi yang dapat meningkatkan risiko gagal bayar (Khahar & Yamasari, 2024).

Tahapan dalam menganalisa risiko kredit merupakan salah satu metode untuk menangani risiko kredit yang bermasalah. Permohonan pengajuan kredit nasabah selama proses ini diverifikasi oleh analis kredit perbankan (Tamami & Kharisudin, 2023). Tawaran perihal layanan kredit dilakukan oleh pihak perbankan melalui proses pemberian pinjaman uang kepada orang atau bisnis serta meminta debitur memenuhi kewajibannya untuk melakukan pelunasan cicilan berbunga dalam jangka waktu tertentu. Layanan kredit ini pada era modern ini tersedia diakibatkan oleh kebutuhan setiap orang meningkat setiap tahun (Martha et al., 2022). Penerapan *machine learning* dalam klasifikasi risiko kredit mengimplementasikan beberapa model algoritma, diantaranya yakni *Logistic Regression*, *Random Forest*, dan *Naive Bayes* (Chang et al., 2024). Karakteristik dan tingkat kinerja performansi model algoritma tersebut tidak sama satu sama lain, meskipun telah banyak diaplikasikan pada bidang klasifikasi risiko kredit. Studi komparatif secara mendalam pun sangatlah dirasa penting agar dapat diketahui algoritma mana yang paling sesuai dalam hal klasifikasi risiko kredit.

Pada studi ini, peneliti mencoba untuk mencari tahu perihal permasalahan mengenai klasifikasi risiko kredit dengan mengimplementasikan tiga buah model algoritma *supervised learning* dan dikomparasi agar dapat mengetahui model algoritma mana yang terbaik. Variabel *class* atau target dalam penelitian ini dibagi menjadi dua kelas, yaitu *good* dan *bad*. *Class bad* ditetapkan sebagai kelas positif karena merepresentasikan nasabah berisiko yang berpotensi menimbulkan kerugian finansial. Penetapan kelas ini bertujuan agar model lebih menekankan kemampuan dalam mendeteksi risiko dibandingkan sekadar mencapai akurasi tinggi. Model algoritma *supervised learning* yang diterapkan yakni model *Logistic Regression*. Model algoritma diterapkan pada penelitian ini dikarenakan memiliki kemampuan dalam menjelaskan hubungan antara variabel dependen biner dan satu atau lebih variabel independen secara baik (Azam & Qoiriah, 2025). Model algoritma *Random Forest* menjadi model algoritma ke dua yang diimplementasikan pada penelitian ini. Kemampuannya yang telah terbukti dalam menangani data berdimensi tinggi, ketahanan terhadap *overfitting*, dan kapasitas untuk mengidentifikasi pola nonlinier yang rumit menjadi alasan bahwa model ini dipilih. Algoritma ini dapat meningkatkan akurasi prediksi dan stabilitas model dengan membangun beberapa pohon keputusan dan menggabungkan hasil keluarannya (Fadillah & Fauzan, 2025). Selanjutnya model algoritma yang ke tiga yakni *Naive Bayes*, alasan peneliti menerapkan model algoritma ini dikarenakan memiliki kemampuan klasifikasi yang handal dan dapat

digunakan untuk memprediksi probabilitas keanggotaan kelas (Hafid & Khikmah, 2024).

Penelitian yang memiliki relevansi dengan studi ini, antara lain kegiatan penelitian yang dilakukan oleh (Adrian & Verawati, 2025) dengan topik penelitian Analisis Performa *Logistic Regression* dan *Random Forest* dalam Klasifikasi Kelayakan Penerimaan Kredit dimana pada penelitian ini menerapkan model algoritma *Logistic Regression* dan *Random Forest*. Hasil akhir pada penelitian ini menggunakan matrix akurasi, *precision*, *recall*, *F1-Score* untuk klasifikasi kelayakan penerimaan kredit. Model *Random Forest* yang diimplementasikan pada penelitian ini lebih unggul dibandingkan dengan *Logistic Regression* dengan skor akurasi 0.95, presisi 0.93, *recall* 0.98 dan *F1 Score* 0.96. Penelitian terkait yang berikutnya, yakni penelitian yang dilakukan oleh (Dong et al., 2024) dengan judul penelitian *Accuracy Comparison between Five Machine Learning Algorithms for Financial Risk Evaluation*. Dalam studi ini, peneliti mengusulkan lima buah model algoritma klasifikasi *K-Nearest Neighbour* (KNN), *Naïve-Bayes*, *Decision Tree*, *Logistic Regression* dan *Random Forest*. Perbandingan akhir dari kelima algoritma tersebut menunjukkan bahwa akurasi beberapa model akan bergantung pada kuantitas, kompleksitas, dan pola data serta prosedur pra-pemrosesan data.

Penelitian yang dilakukan oleh (Wattanakitrunroj et al., 2024) dengan topik penelitian *Enhancing Supervised Model Performance in Credit Risk Classification Using Sampling Strategies and Feature Ranking*. Tiga teknik pembelajaran mesin *supervised learning* yang populer *Logistic Regression*, *Random Forest*, dan *Gradient Boosting* digunakan dalam eksperimen prediksi risiko kredit untuk menghasilkan hasil yang sangat efektif. Hasil akhir studi ini menunjukkan bahwa teknik *Gradient Boosting* mencapai nilai akurasi, presisi, *recall*, dan *F1 Score* yang hampir sempurna, yaitu lebih dari 99,92%, tetapi nilai MCC-nya lebih besar dari 99,77%.

Meskipun berbagai penelitian telah membandingkan algoritma *machine learning* untuk klasifikasi risiko kredit, sebagian besar studi masih berfokus pada metrik akurasi sebagai indikator utama performa model. Pendekatan tersebut belum sepenuhnya mencerminkan kebutuhan praktis industri keuangan yang lebih menekankan kemampuan model dalam mendeteksi nasabah berisiko tinggi (*bad*). Selain itu, masih terbatas penelitian yang mengombinasikan evaluasi berbasis risiko dengan penggunaan metrik *Cohen's Kappa* untuk menilai konsistensi hasil klasifikasi. Oleh karena itu, diperlukan penelitian yang tidak hanya membandingkan tingkat akurasi model, tetapi juga mengevaluasi kemampuan deteksi kelas berisiko sebagai dasar pengambilan keputusan kredit.

Penelitian yang dilakukan oleh peneliti ini memiliki beberapa kontribusi dan kebaruan yang membahas mengenai pendekatan evaluasi berbasis risiko. Topik pada penelitian ini tidak serupa dengan sebagian penelitian yang hanya menekankan pada nilai *accuracy*, namun lebih memfokuskan evaluasi pada kemampuan model dalam mendeteksi kelas *bad* sebagai kelas berisiko. Pendekatan ini memiliki relevansi dalam konteks risiko kredit karena kesalahan klasifikasi pada kelas *bad* (*false negative*) dapat menimbulkan kerugian finansial yang signifikan (Putra & Rumini, 2025). Kebaruan penelitian ini terletak pada pendekatan evaluasi yang berorientasi pada deteksi kelas *bad* sebagai kelompok risiko utama dalam kredit, serta penggunaan *Cohen's Kappa* sebagai metrik tambahan untuk menilai reliabilitas model. Pendekatan ini memberikan perspektif yang lebih relevan terhadap kebutuhan manajemen risiko dibandingkan evaluasi yang hanya berfokus pada akurasi.

2. METODE PENELITIAN

Penelitian ini dilakukan melalui beberapa tahapan seperti yang tercantum pada gambar 1 kerangka pemikiran penelitian. Komparasi kinerja tiga buah model algoritma *supervised learning* yang memiliki ketepatan menjadi main topic dalam hal klasifikasi risiko kredit. Terdapat tiga buah model algoritma yang dikomparasikan pada penelitian ini, yakni model algoritma *Logistic Regression* (LR), *Random Forest* (RF) dan *Naive Bayes* (NB).

Penelitian ini dilaksanakan melalui beberapa tahapan yang disusun dalam kerangka pemikiran penelitian. Tahap pertama adalah *data collection*, yaitu proses pengumpulan data yang digunakan sebagai dasar analisis. Pada penelitian ini digunakan dataset german credit yang bersifat terbuka dan dapat diakses secara publik melalui Kaggle pada laman <https://www.kaggle.com/datasets/kabure/german-credit-data-with-risk>. Dataset tersebut terdiri dari 1.000 record yang memuat informasi terkait pengajuan kredit, meliputi usia, status perkawinan, saldo rekening giro, pekerjaan pemohon, serta variabel target berupa *risk* yang mengklasifikasikan tingkat risiko kredit. Dataset ini dipilih karena banyak digunakan dalam penelitian klasifikasi risiko kredit dan memiliki karakteristik yang sesuai untuk penerapan algoritma pembelajaran mesin.

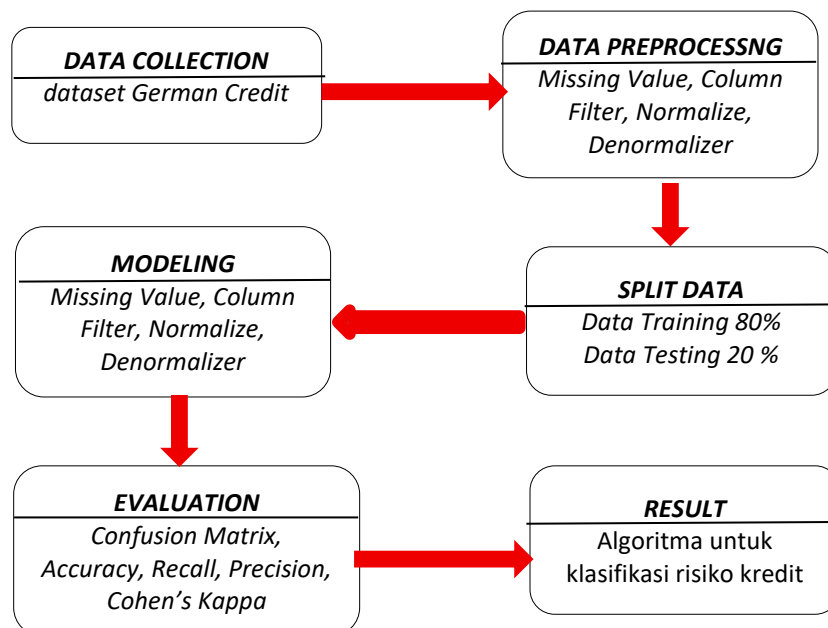
Tahap berikutnya adalah *data preprocessing*, yang bertujuan untuk meningkatkan kualitas data sebelum dilakukan proses pemodelan. Pada tahap ini digunakan node missing value untuk menangani nilai yang hilang pada dataset. Selanjutnya, node *column filter* digunakan untuk memilih atribut yang relevan dengan penelitian dan mengeluarkan atribut yang tidak diperlukan. Proses berikutnya adalah normalisasi data menggunakan node normalizer yang berfungsi melakukan penskalaan terhadap seluruh fitur numerik agar berada pada rentang nilai yang seragam. Setelah proses klasifikasi dilakukan, node *denormalizer* digunakan untuk mengembalikan hasil prediksi ke skala aslinya sehingga interpretasi hasil dapat dilakukan dengan lebih mudah.

Tahap ketiga adalah data splitting, yaitu proses pembagian dataset menjadi data latih dan data uji menggunakan node *table partitioner*. Pada penelitian ini diterapkan rasio pembagian 80:20, di mana 80% data digunakan sebagai data latih dan 20% sisanya digunakan sebagai data uji. Validasi model dilakukan menggunakan metode *hold-out validation* dengan tujuan memperoleh evaluasi performa model yang independen terhadap data pelatihan.

Tahap selanjutnya adalah modeling, yaitu proses pembangunan model klasifikasi menggunakan tiga algoritma *supervised learning*, yaitu *Logistic Regression*, *Random Forest*, dan *Naïve Bayes*. *Logistic Regression* dipilih karena kemampuannya dalam memodelkan hubungan antara variabel dependen biner dan satu atau lebih variabel independen secara efektif (Teary, 2025). Pada penelitian ini, *Logistic Regression* menggunakan parameter bawaan KNIME dengan fungsi sigmoid untuk memodelkan probabilitas kelas. Algoritma *Random Forest* digunakan karena memiliki kemampuan menghasilkan prediksi yang akurat melalui pendekatan *ensemble learning* (Saputra et al., 2024). Model ini dibangun menggunakan 100 *decision tree* dengan mekanisme *bootstrap sampling* dan *majority voting* untuk menentukan hasil klasifikasi akhir. Sementara itu, *Naïve Bayes* dipilih karena mampu menangani data numerik maupun kategorikal dengan baik (Aisyah & Latifa, 2024). Algoritma ini menggunakan pendekatan probabilistik berdasarkan Teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen satu sama lain.

Tahap berikutnya adalah *evaluation*, yaitu proses evaluasi terhadap kinerja model yang telah dibangun. Evaluasi dilakukan menggunakan *Confusion Matrix* sebagai dasar perhitungan beberapa metrik performa, yaitu Accuracy, Recall, dan Precision. Selain itu, penelitian ini juga menggunakan Cohen's Kappa sebagai metrik tambahan untuk mengukur tingkat kesepakatan hasil klasifikasi yang telah dikoreksi terhadap kemungkinan kesepakatan yang terjadi secara acak. Penggunaan Cohen's Kappa memberikan perspektif tambahan dalam menilai reliabilitas model klasifikasi yang dihasilkan.

Tahap terakhir adalah *result*, yaitu penyajian dan analisis hasil komparasi dari ketiga algoritma *supervised learning* yang digunakan. Pada tahap ini dilakukan perbandingan performa *Logistic Regression*, *Random Forest*, dan *Naïve Bayes* berdasarkan metrik evaluasi yang telah ditentukan untuk memperoleh model yang paling efektif dalam melakukan klasifikasi risiko kredit.



Gambar 1. Kerangka Pemikiran Penelitian

3. HASIL DAN PEMBAHASAN

Hasil penelitian berupa komparasi tiga buah algoritma supervised learning yakni model algoritma Logistic Regression, Random Forest dan Naive Bayes dituangkan dalam bentuk implementasi antarmuka dan tabel hasil nilai akhir. Dataset yang diinput berupa data German Credit merupakan langkah pertama dalam siklus kinerja implementasi pengolahan data. Himpunan data ini dibagi menjadi data latih dan data uji. Data latih digunakan untuk pembuatan dan pelatihan model, sedangkan data uji digunakan untuk proses implementasi model klasifikasi yang digunakan.

A. Data Collection

Pada penelitian ini menggunakan himpunan data *German Credit* yang diperoleh pada laman <https://www.kaggle.com/datasets/kabure/german-credit-data-with-risk>. Himpunan data ini terdiri dari sepuluh fitur yang bersifat independent, antara lain *Column1*, *Age*, *Sex*, *Job*, *Housing*, *Saving accounts*, *Checking account*, *Credit amount*, *Duration*, dan *Purpose*. *Feature* yang menjadi *class* atau labelnya yaitu *Risk* dengan tipe nominal *Good* untuk mengindikasikan layak kredit dan *Bad* untuk mengindikasikan berisiko. Tampilan himpunan datanya dapat dilihat pada gambar 2.

#	RowID	Column0	Age	Sex	Job	Housing	Saving ac...	Checking ...	Credit am...	Duration	Purpose	Risk
		Number (L..)	Number (L..)	String	Number (L..)	String	String	String	Number (L..)	Number (L..)	String	String
1	Row0	0	67	male	2	own	NA	little	1169	6	radio/TV	good
2	Row1	1	22	female	2	own	little	moderate	5951	48	radio/TV	bad
3	Row2	2	49	male	1	own	little	NA	2096	12	education	good
4	Row3	3	45	male	2	free	little	little	7882	42	furniture/equipr	good
5	Row4	4	53	male	2	free	little	little	4870	24	car	bad
6	Row5	5	35	male	1	free	NA	NA	9055	36	education	good
7	Row6	6	53	male	2	own	quite rich	NA	2835	24	furniture/equipr	good
8	Row7	7	35	male	3	rent	little	moderate	6948	36	car	good
9	Row8	8	61	male	1	own	rich	NA	3059	12	radio/TV	good
10	Row9	9	28	male	3	own	little	moderate	5234	30	car	bad
11	Row10	10	25	female	2	rent	little	moderate	1295	12	car	bad
12	Row11	11	24	female	2	rent	little	little	4308	48	business	bad
13	Row12	12	22	female	2	own	little	moderate	1567	12	radio/TV	good
14	Row13	13	60	male	1	own	little	little	1199	24	car	bad
15	Row14	14	28	female	2	rent	little	little	1403	15	car	good
16	Row15	15	32	female	1	own	moderate	little	1282	24	radio/TV	bad

Gambar 2. Tampilan Himpunan Data

B. Data Preprocessing

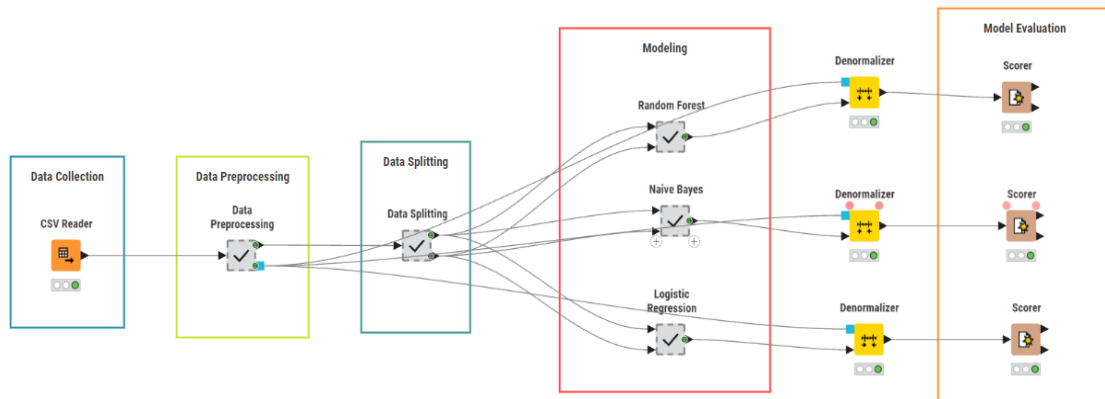
Pada tahapan ini terdapat tiga buah node yang digunakan. Node yang pertama yakni *node Missing Value* digunakan untuk memeriksa atribut yang bertipe *Number* (Integer) dengan memilih kotak dialog *Most Frequent Value* dan atribut yang bertipe *String*/Nominal dengan memilih kotak dialog *Do Nothing*. Node yang kedua yaitu *node Column Filter* digunakan untuk memfilter atribut *Column0* ke kotak dialog *excludes*. Node yang ketiga yaitu *node Normalizer* yang digunakan untuk normalisasi data menggunakan *Min-Max* pada atribut *Age*, *Job*, *Credit amount* dan *Duration*. Node berikutnya yang digunakan yaitu *Denormalizer* di akhir *workflow* setelah proses *modeling*.

C. Split Data

Node Table Partitioner digunakan pada tahapan ini. Pada tahapan ini dilakukan pembagian data latih sebesar 80% dengan jumlah sebanyak 800 *record*, sedangkan data uji sebesar 20% dengan jumlah sebesar 200 *record*. Pembagian ini dilakukan dikarenakan dapat menghasilkan kompromi antara jumlah data pelatihan yang cukup dan jumlah data pengujian yang wajar untuk evaluasi objektif (Joseph, 2022). Rasio 80:20 digunakan sebagai pembagian standar untuk model algoritma *supervised learning*, serta dibahas agar memperoleh evaluasi performa model yang independen dari data pelatihan (Hsieh, 2025).

D. Modeling

Pada tahapan *Modeling* ini menggunakan aplikasi *KNIME Analytics Platform* versi 5.9.0 mengimplementasikan tiga buah model algoritma *supervised learning* yang memiliki *node* atau komponen yang tertuang dalam *platform* tersebut. Model algoritma seperti *Logistic Regression*, *Random Forest* dan *Naive Bayes* memiliki *node* yang berbeda untuk setiap tahap dalam proses pengembangan model. Penjabaran mengenai *workflow* *KNIME 5.9.0* yang menampilkan tiga alur kerja model klasifikasi risiko kredit dituangkan dalam gambar 3.



Gambar 3. Tampilan Tiga Alur Kerja Model Klasifikasi Risiko Kredit menggunakan KNIME

Pada tampilan *workflow* ini, data *German Credit* dibaca menggunakan *node csv reader*. Selanjutnya, himpunan data tersebut difilter dalam penentuan atribut mana yang akan digunakan. Kemudian pada tahapan pengujian algoritma, dimana menerapkan model *Random Forest* menggunakan *node Random Forest Learner* dan *Predictor*, *Logistic Regression* menggunakan *node Logistic Regression Learner* dan *Predictor*, lalu model yang ke tiga yakni *Naive Bayes* menggunakan *Naive Bayes Learner* dan *Predictor*. Hasil akhir dari klasifikasi dapat diperoleh dengan mencantumkan *node scorer* sebagai teknik evaluasi model.

E. Evaluation

Penjabaran model pembelajaran mesin dilakukan dengan variasi parameter untuk menghasilkan model yang terbaik dengan nilai akurasi tertinggi. Dimana hal ini membantu mengoptimalkan model serta mencegah *overfitting*. Kegiatan *evaluation* ini bertujuan memastikan bahwa model yang dihasilkan memiliki performa optimal dan mampu menekan *False Negative* serendah mungkin dengan akurasi yang tetap tinggi. Tabel 1 menampilkan hasil *Confusion Matrix* untuk masing-masing model algoritma yang digunakan pada penelitian. Sedangkan hasil komparasi performa tiga model algoritma pada pembelajaran mesin dituangkan pada tabel 2.

Tabel 1. Hasil *Confusion Matrix* Model Algoritma *Random Forest* (RF), *Logistic Regression* (LR), dan *Naïve Bayes* (NB)

Aktual/Prediksi	Bad (RF)	Good (RF)	Bad (LR)	Good (LR)	Bad (NB)	Good (NB)
Bad	29	132	39	131	38	129
Good	31	8	21	9	22	11

Tabel 2. Hasil Komparasi Performa Tiga Buah Model Algoritma

Algoritma	Accuracy	Recall	Precision	F1-Score	Cohens's Kappa
<i>Random Forest</i>	81,50%	0,82	0,94	0,87	0,51
<i>Logistic Regression</i>	76,10%	0,77	0,93	0,85	0,33
<i>Naïve Bayes</i>	75,50%	0,76	0,92	0,84	0,31

Berdasarkan nilai *Cohen's Kappa*, *Random Forest* memperoleh skor 0,51 yang termasuk kategori *moderate agreement*. *Logistic Regression* memperoleh nilai 0,33 dan *Naïve Bayes* sebesar 0,31 yang termasuk kategori *fair agreement*. Hasil ini menunjukkan bahwa *Random Forest* memiliki konsistensi klasifikasi yang lebih baik dibandingkan dua model lainnya.

Fokus utama penelitian ini adalah kemampuan model dalam mendeteksi kelas *bad* karena kesalahan klasifikasi pada kelas ini berpotensi menyebabkan kerugian finansial. Berdasarkan nilai *recall*, *Random Forest* mencapai skor tertinggi sebesar 0,82, yang menunjukkan bahwa model mampu mengidentifikasi 82% nasabah berisiko secara benar. *Logistic Regression* dan *Naïve Bayes* masing-masing memperoleh *recall* sebesar 0,77 dan 0,76. Hasil ini mengindikasikan bahwa *Random Forest* memiliki kemampuan yang lebih baik dalam meminimalisir kesalahan *false negative*. Keunggulan *Random Forest* diduga berasal dari kemampuannya membangun banyak pohon keputusan dan menggabungkan hasil

prediksi melalui mekanisme *voting*. Pendekatan ini memungkinkan model menangkap hubungan nonlinier antar variabel yang sulit diidentifikasi oleh *Logistic Regression*. Selain itu, *Random Forest* relatif lebih tahan terhadap *noise* dan variasi data sehingga menghasilkan performa klasifikasi yang lebih stabil.

4. KESIMPULAN

Penelitian ini berhasil membandingkan performa *Logistic Regression*, *Random Forest*, dan *Naïve Bayes* dalam klasifikasi risiko kredit menggunakan *dataset German Credit*. Hasil penelitian menunjukkan bahwa *Random Forest* memberikan performa paling baik dalam mendeteksi nasabah berisiko (*bad credit*) dibandingkan dua algoritma lainnya. Temuan ini memperkuat pentingnya penggunaan model *ensemble* dalam mendukung proses penilaian risiko kredit yang lebih akurat. Selain menghasilkan tingkat akurasi yang lebih tinggi, *Random Forest* juga menunjukkan kemampuan deteksi risiko yang lebih baik berdasarkan nilai *recall* dan *Cohen's Kappa*. Meskipun demikian, penelitian ini masih memiliki keterbatasan karena hanya menggunakan satu *dataset* dan metode validasi *hold-out*. Oleh karena itu, penelitian selanjutnya dapat menerapkan teknik *cross-validation*, menambahkan algoritma lain seperti *XGBoost* atau *Support Vector Machine*, serta mengevaluasi pengaruh penanganan data tidak seimbang terhadap performa klasifikasi risiko kredit.

5. REFERENSI

- Adrian, A., & Verawati, I. (2025). Analisis Performa Logistic Regression dan Random Forest dalam Klasifikasi Kelayakan Penerimaan Kredit. *IJCSR: The Indonesian Journal of Computer Science Research*, 4(2), 148–158.
- Aisyah, S., & Latifa, K. (2024). Klasifikasi Naïve Bayes pada Machine Learning untuk Rekomendasi Kredit Customer pada KSP. *IN-FEST 2024 Seminar Nasional Informatika-FTI UPGRIS*, 2(1), 425–432.
- Azam, H. I., & Qoiriah, A. (2025). Prediksi Kelayakan Pinjaman Berdasarkan Profil Risiko Nasabah Menggunakan Logistic Regression dan Random Forest. *JINACS (Journal of Informatics and Computer Science) ISSN*, 07(01), 301–312.
- Baser, F., Koc, O., & Selcuk-Kestel, A. S. (2023). Credit risk evaluation using clustering based fuzzy classification method. *Expert Systems with Applications*, 223, 119882. <https://doi.org/https://doi.org/10.1016/j.eswa.2023.119882>
- Bhandary, R., & Ghosh, B. K. (2025). Credit Card Default Prediction : An Empirical Analysis on Predictive Performance Using Statistical and Machine Learning Methods. *Journal of Risk and Financial Management*, 18(1). <https://doi.org/https://doi.org/10.3390/jrfm18010023>
- Chang, V., Sivakulasingam, S., Wang, H., Wong, S. T., & Ganatra, M. A. (2024). Credit Risk Prediction Using Machine Learning and Deep Learning : A Study on Credit Card Customers. *Risks*, 12(11), 1–33. <https://doi.org/https://doi.org/10.3390/risks12110174>
- Dong, H., Liu, R., & Tham, A. W. (2024). Accuracy Comparison between Five Machine Learning Algorithms for Financial Risk Evaluation. *Journal of Risk and Financial Management*, 17(2), 1–19. <https://doi.org/10.3390/jrfm17020050>
- Fadillah, A. R., & Fauzan, M. N. (2025). Prediction of People's Business Credit Loan Acceptance Using Machine Learning Model : Application of Random Forest Algorithm. *IT Journal Research and Development (ITJRD)*, 10(1), 1–10.
- Hafid, N., & Khikmah, L. (2024). Perbandingan Metode Algoritma C4.5, Naïve Bayes, dan Logistic Regression untuk Penentuan Kelayakan Penerima Kredit. *TEKNOLOGI: Jurnal Ilmiah Sistem Informasi*, 14(2), 85–93. <https://doi.org/https://doi.org/10.26594/teknologi.v14i2.4650>
- Hsieh, P. (2025). ML-based validation of experimental randomization in learning games. *Frontiers in Artificial Intelligence*, 8(1), 1–12. <https://doi.org/10.3389/frai.2025.1541087>
- Joseph, V. R. (2022). Optimal Ratio for Data Splitting. *Statistics and Machine Learning*, 1(1), 1–16.
- Khahar, A., & Yamasari, Y. (2024). Eksplorasi Teknik Pre-Processing Berbasis eXtreme Gradient Boosting (XGBoost) pada Klasifikasi Kredit Default. *JINACS (Journal of Informatics and Computer Science) ISSN*, 06(02), 453–464.
- Martha, S., Andani, W., & Rizki, S. W. (2022). Perbandingan Metode k-Nearest Neighbor , Regresi Logistik Biner , dan Pohon Klasifikasi pada Analisis Kelayakan Pemberian Kredit. *EULER: Jurnal Ilmiah Matematika, Sains Dan Teknologi*, 10(2), 262–273.

- Putra, H., & Rumini. (2025). Comparative Study of Logistic Regression , Random Forest , and XGBoost for Bank Loan Approval Classification. *Journal of Applied Informatics and Computing (JAIC)*, 9(5), 2822–2835.
- Sakina, I. N., & Damayanti, R. F. (2025). Systematic Literatur Review Pendekatan Machine Learning untuk Prediksi Risiko Kredit. *Jurnal PROSISKO*, 12(3), 103–109.
- Saputra, D. B., Atina, V., & Nastiti, F. E. (2024). Penerapan Model CRISP-DM pada Prediksi Nasabah Kredit Menggunakan Algoritma Random Forest. *Idealis: Indonesia Journal Information System*, 7(2), 240–247.
- Septama, H. D., Yulianti, T., Budiyanto, D., Mulyadi, S. M., & Cahyana, A. H. (2023). A Comparative Analysis of Machine Learning Algorithms for Credit Risk Scoring using Chi-Square Feature Selection. *2023 International Conference on Converging Technology in Electrical and Information Engineering (ICCTEIE)*, 32–37. <https://doi.org/10.1109/ICCTEIE60099.2023.10366576>
- Tamami, M. K., & Kharisudin, I. (2023). Komparasi Metode Support Vector Machine dan Naive Bayes Classifier untuk Pemodelan Kualitas Pengajuan Kredit. *Indonesian Journal of Mathematics and Natural Sciences*, 46(1), 38–44.
- Teary, M. G. (2025). Optimalisasi Model Credit Scoring Menggunakan Algoritma Logistic Regression dan Decision Tree. *Jurnal Ilmiah Informatika Global*, 16(03), 352–358.
- Wattanakitrunroj, N., Wijitkajee, P., Jaiyen, S., Sathapornvajana, S., & Tongman, S. (2024). Enhancing Supervised Model Performance in Credit Risk Classification Using Sampling Strategies and Feature Ranking. *Big Data and Cognitive Computing*, 8(3). <https://doi.org/10.3390/bdcc8030028>
- Zhao, Y. (2022). Credit Card Approval Predictions Using Logistic Regression, Linear SVM and Naïve Bayes Classifier. *2022 International Conference on Machine Learning and Knowledge Engineering (MLKE)*, 207–211. <https://doi.org/10.1109/MLKE55170.2022.00047>