

PERBANDINGAN ALGORITMA NAIVE BAYES DAN RANDOM FOREST UNTUK KLASIFIKASI KELAYAKAN KREDIT PADA PT. BPR MAGGA JAYA UTAMA

Cornelius Orlando Christiadi¹, Indah Fenriana^{2*}

^{1,2} Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Buddhi Dharma

*Corresponding Author, email: indah.fenriana@ubd.ac.id

ABSTRAK

Kredit menjadi salah satu fasilitas yang disediakan oleh perusahaan khususnya pada bidang keuangan dan cukup diminati oleh masyarakat. Tujuan menggunakan fasilitas tersebut adalah memenuhi keinginan atau kebutuhan baik individu maupun bisnis seperti modal usaha, investasi atau konsumsi pribadi. Penyaluran kredit kepada masyarakat terus meningkat setiap tahunnya dan penyaluran kredit oleh Bank Perkreditan Rakyat (BPR) sejalan dengan meningkatnya rasio *Non-Performing Loan* (NPL) yang berdampak pada kelayakan kredit debitur. Penelitian ini bertujuan untuk membandingkan algoritma Naive Bayes dan Random Forest dalam mengklasifikasikan kelayakan kredit pada PT. BPR Magga Jaya Utama. Sehingga dapat memberikan kemudahan dalam membuat keputusan kelayakan kredit calon debitur dan meminimalisir meningkatnya rasio NPL. Metode yang digunakan meliputi data selection, cleaning, transformation, dan pemodelan menggunakan *RapidMiner* dan implementasi website dengan PHP Code Igniter. Dataset yang belum diproses berjumlah 6983 data, dan setelah melakukan beberapa proses data yang digunakan berjumlah 5553 data nasabah, dibagi menjadi data training dan data testing dengan rasio 70:30. Evaluasi dilakukan menggunakan metrik akurasi, presisi, dan recall. Hasil menunjukkan bahwa algoritma Random Forest memberikan akurasi lebih tinggi (96,13% menggunakan *RapidMiner*, 95,50% menggunakan *website*) dibandingkan Naive Bayes (89,26% di *RapidMiner*, 89,68% menggunakan *website*). Penelitian ini menyimpulkan bahwa Random Forest lebih efektif dalam klasifikasi kelayakan kredit dan dapat dijadikan acuan dalam pengambilan keputusan oleh pihak bank.

Kata kunci: Data Mining, Kelayakan Kredit, Klasifikasi, Naive Bayes, Random Forest

I. PENDAHULUAN

Kredit perbankan merupakan fasilitas keuangan yang penting untuk mendukung kegiatan ekonomi masyarakat (Dasari et al., 2023). Meskipun penyaluran kredit oleh BPR meningkat sebesar 29,63% selama 2019–2023, peningkatan ini diiringi oleh kenaikan rasio Non-Performing Loan (NPL) sebesar 71,98% (“Stat. Perbank. Indones.,” 2023). Rasio NPL mencerminkan kualitas manajemen risiko kredit dan stabilitas keuangan bank (Gayo et al., 2022). Untuk mengatasi risiko tersebut, diperlukan sistem klasifikasi kelayakan kredit yang

akurat dan andal, salah satunya melalui pengembangan model credit scoring berbasis data (Sudirman et al., 2018).

Data mining dan machine learning menjadi pendekatan yang relevan karena mampu menggali pola tersembunyi sebagai nilai tambah dalam dataset besar (Suyanto, 2017; Wanto, M Kom et al., 2020). Salah satu keunggulan machine learning adalah kemampuannya untuk melakukan klasifikasi terhadap data nasabah berdasarkan karakteristik historis yang kompleks. Dalam konteks ini, algoritma Naïve Bayes dan Random Forest banyak digunakan untuk membangun model prediksi kelayakan kredit. Dalam konteks ini, algoritma Naïve Bayes dan Random Forest banyak digunakan. Naïve Bayes didasarkan pada probabilitas antar atribut yang diasumsikan independen (Ramadhan et al., 2022), sedangkan Random Forest menggabungkan banyak pohon keputusan untuk menangani data kompleks dan variabel yang saling terkait (Aziz, 2021).

Beberapa penelitian terdahulu menyatakan bahwa Random Forest memiliki keunggulan dari sisi akurasi. Random Forest menghasilkan akurasi sebesar 75% dibandingkan dengan Naïve Bayes sebesar 71% (Depari, 2022). Lalu Penelitian lain bahkan mencatat selisih akurasi hingga 49,94% lebih tinggi pada Random Forest dibandingkan Naïve Bayes (Indaryono et al., 2024). Oleh karena itu, penelitian ini membandingkan kedua algoritma tersebut menggunakan data nasabah dari PT. BPR Magga Jaya Utama.

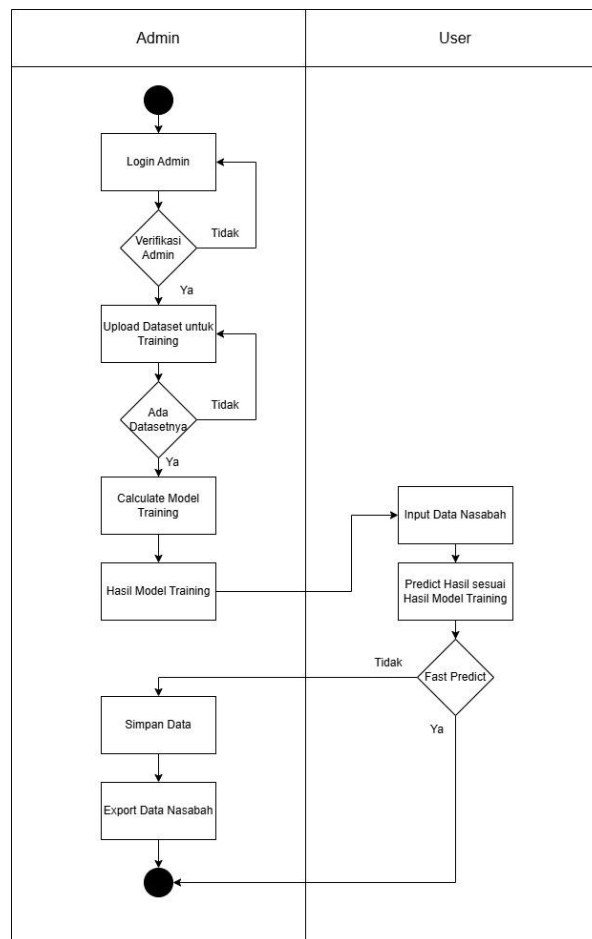
II. METODOLOGI

Penelitian ini menggunakan data nasabah PT. BPR Magga Jaya Utama tahun 2019 hingga 2023 sebanyak 6983 data. Setelah dilakukan data cleaning, data menjadi 5553 baris. Proses penelitian terdiri atas:

1. Data Selection: Pemilihan atribut yang relevan (umur, profesi, pendapatan, plafond, angsuran, jangka waktu, dan kolektibilitas) (Sudirman et al., 2018).
2. Data Cleaning: Menghapus data-data yang tidak memiliki nilai atau null (Sudirman et al., 2018).
3. Data Transformation: Mengubah status kolektibilitas 1 & 2 menjadi “Lancar” dan 3, 4 & 5 menjadi “Macet” (Widia & Asriningtias, 2021).
4. Data Splitting: Pembagian data training (70%) dan testing (30%) (Joseph, 2022).

5. Modeling: Menerapkan algoritma Naïve Bayes dan Random Forest menggunakan RapidMiner dan website PHP CodeIgniter (Herdiansah et al., 2024; Putro, 2021; Rahmat et al., 2017).
6. Evaluasi: Menggunakan metrik akurasi, presisi, dan (Aziz, 2021).

Alur dari implementasi *website* digambarkan dalam gambar 1, *website* dibangun dengan menggunakan bahasa pemrograman PHP CodeIgniter. Untuk menghitung hasil model, pengguna harus *upload* dataset kemudian pengguna dapat melihat hasil dari model tersebut. Pengguna juga dapat menggunakan prediksi cepat, dimana pengguna hanya memasukkan 1 (satu) data nasabah dengan beberapa *field*.



Gambar 1. Alur Penelitian

III. HASIL DAN PEMBAHASAN

Implementasi Algoritma Naïve Bayes dan Random Forest dijalankan pada 2 buah platform: *RapidMiner* dan *Website*. Hasil implementasi menggunakan

RapidMiner ditunjukkan pada Tabel 1, dan Implementasi menggunakan *website* ditunjukkan pada Tabel 2.

Tabel 1. Hasil *RapidMiner*

Algoritma	Akurasi	Presisi	Recall
Naive Bayes	89,26%	92,85%	94,24%
Random Forest	96,13%	95,16%	100%

Tabel 2. Hasil *RapidMiner Website*

Algoritma	Akurasi	Presisi	Recall
Naive Bayes	89,68%	86%	86%
Random Forest	95,50%	97%	91%

Berdasarkan hasil tersebut, Random Forest menunjukkan kinerja lebih baik dalam semua metrik (Aziz, 2021; Depari, 2022). Hal ini menunjukkan bahwa algoritma tersebut lebih andal dalam klasifikasi data kompleks dan cenderung memberikan prediksi lebih akurat (Indaryono et al., 2024).

IV. SIMPULAN

Penelitian ini menyimpulkan bahwa algoritma Random Forest lebih unggul dibandingkan Naïve Bayes dalam klasifikasi kelayakan kredit berdasarkan data nasabah PT. BPR Magga Jaya Utama. Hal ini dibuktikan dengan menggunakan aplikasi *RapidMiner* (Putro, 2021) dan *website*. Dengan akurasi dan presisi tinggi, Random Forest cocok digunakan dalam pengambilan keputusan kredit. Kedepan, penelitian ini dapat dikembangkan dengan menambahkan algoritma lain serta pengujian terhadap data yang lebih besar dan beragam (Indaryono et al., 2024).

DAFTAR PUSTAKA

Aziz, W. A. (2021). Implementasi metode random forest pada klasifikasi data ulasan konsumen perusahaan (studi kasus: aplikasi kai access). In *Repository.Uinjkt.Ac.Id*.

[https://repository.uinjkt.ac.id/dspace/handle/123456789/67842%0Ahttps://repository.uinjkt.ac.id/dspace/bitstream/123456789/67842/1/WILDAN ABDUL AZIZ- FST.pdf](https://repository.uinjkt.ac.id/dspace/handle/123456789/67842%0Ahttps://repository.uinjkt.ac.id/dspace/bitstream/123456789/67842/1/WILDAN%20ABDUL%20AZIZ-FST.pdf)

Dasari, Y., Rishitha, K., & Gandhi, O. (2023). Prediction of Bank Loan Status Using Machine Learning Algorithms. *International Journal of Computing*

- and Digital Systems*, 14(1), 139–146. <https://doi.org/10.12785/ijcds/140113>
- Depari, D. H. (2022). Perbandingan Model Decision Tree, Naive Bayes, dan Random Forest untuk Prediksi Klasifikasi Penyakit Jantung. *UPN VETERAN JAKARTA*.
- Gayo, A. A., Prihatni, R., & Armeliza, D. (2022). Faktor – Faktor yang Mempengaruhi Penyaluran Kredit pada Bank Umum di Indonesia. *Jurnal Akuntansi Dan Keuangan*, 10(1), 25. <https://doi.org/10.29103/jak.v10i1.6099>
- Herdiansah, A., Purnamasari, M., Handayani, T., & Sulaeman, A. (2024). Implementasi Framework Codeigniter Pada Pengembangan Sistem Informasi Dashboard Tracking Data Alumni Sekolah. *JIKA (Jurnal Informatika)*, 8(2), 213. <https://doi.org/10.31000/jika.v8i2.10849>
- Indaryono, N. A. P., Saedudin, R. R., & Hamami, F. (2024). Analisa Perebandingan Algoritma Random Forest dan Naive Bayes untuk Klasifikasi Curah Hujan Berdasarkan Iklim di Indonesia. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 9, 158–167.
- Joseph, V. R. (2022). Optimal ratio for data splitting. *Statistical Analysis and Data Mining*, 15(4), 531–538. <https://doi.org/10.1002/sam.11583>
- Putro, S. S. (2021). *Teori dan Aplikasi Data Mining Dengan RapidMiner*. Literasi Nusantara Abadai.
- Rahmat, B., Agidrama Gafar, A., Fajriani, N., Ramdani, U., Rihin Uyun, F., Purnamasari P., Y., & Ransi, N. (2017). Implementasi k-means clustering pada rapidminer untuk analisis daerah rawan kecelakaan. *Seminar Nasional Riset Kuantitatif Terapan 2017, April*, 58–60. <https://ojs.innov-center.org/index.php/snrkt2017/article/download/10/9>
- Ramadhan, B. Z., Riza, I., & Maulana, I. (2022). Analisis Sentimen Ulasan Pada Aplikasi E-Commerce Dengan Menggunakan Algoritma Naive Bayes Bintang. *Jurnal of Applied Informatics and Computing (JAIC)*, 2, 220–225. <https://doi.org/10.55606/teknik.v3i3.2411>
- Statistik Perbankan Indonesia. (2023). *Statistik Perbankan Indonesia, Vol. 22 No*, 140.
- Sudirman, Windarto, A. P., & Wanto, A. (2018). Data mining tools | rapidminer: K-means method on clustering of rice crops by province as efforts to stabilize



- food crops in Indonesia. *IOP Conference Series: Materials Science and Engineering*, 420(1). <https://doi.org/10.1088/1757-899X/420/1/012089>
- Suyanto. (2017). *Data mining: untuk klasifikasi dan klusterisasi data*. Penerbit Informatika.
- Wanto, M Kom, A., Siregar, M. N. H., Windarto, A. P., Hartama, D., Ginantra, M Kom, N. L. W. S. R., Napitupulu, M Kom, S. D., Negara, M Kom, E. S., Lubis, M. R., Dewi, S. V., & Prianto, C. (2020). *Data Mining: Algoritma dan Implementasi*.
- Widia, D. M., & Asriningtias, S. R. (2021). Cara Cepat dan Praktis Membangun Web Dinamis dengan PHP dan MySQL. *Universitas Brawijaya Press*. https://scholar.google.com/citations?view_op=view_citation&hl=en&user=OGPzBq8AAAAJ&citation_for_view=OGPzBq8AAAAJ:kNdYIx-mwKoC