

PENERAPAN METODE RANDOM FOREST DALAM SISTEM PENDUKUNG KEPUTUSAN KELAYAKAN KREDIT PADA PEGADAIAN DI KOTA BEKASI

Kristina Natalia bebe Mbupu^{1*}, Sri Wahyuni ², M Andres³
^{1,2,3}Teknik Informatika, Fakultas Sains dan Teknologi
Universitas Panca Sakti Bekasi
Kristinanataliabebembupu@gmail.com

ABSTRACT

This study aims to design and develop a web-based Decision Support System using the Random Forest algorithm to classify credit feasibility in the context of Pegadaian in Bekasi City into Feasible and Not Feasible categories. A quantitative approach with an applied experiment was used on a simulated dataset of 1,501 records. The dataset consisted of seven features: income, occupation, age, collateral type, collateral value, loan amount, and loan term. The data were split into 1,200 training records and 301 testing records using an 80:20 ratio. The model was developed using Python and scikit-learn with 100 decision trees, then integrated into a web application using PHP, MySQL, and a Flask service. Evaluation used a confusion matrix, accuracy, precision, recall, F1-score, AUC-ROC, and feature importance. The results obtained 84.4% accuracy, 85.7% precision, 92.8% recall, 89.1% F1-score, and 88.0% AUC-ROC. Income had the highest relative contribution to the model predictions. The results indicate that Random Forest can be implemented as a data-driven tool for credit feasibility classification. However, because the study used a simulated dataset, final decisions still require verification and officer consideration.

Keywords: Random Forest, decision support system, credit feasibility, classification, Pegadaian.

ABSTRAK

Penelitian ini bertujuan merancang dan membangun Sistem Pendukung Keputusan (SPK) berbasis web dengan menerapkan algoritma Random Forest untuk mengklasifikasikan kelayakan kredit pada konteks Pegadaian di Kota Bekasi ke dalam kategori Layak dan Tidak Layak. Penelitian menggunakan pendekatan kuantitatif dengan eksperimen terapan pada dataset simulasi sebanyak 1.501 data. Dataset terdiri atas tujuh fitur, yaitu pendapatan, pekerjaan, umur, jenis jaminan, nilai jaminan, jumlah pinjaman, dan jangka waktu. Data dibagi menjadi 1.200 data latih dan 301 data uji dengan rasio 80:20. Model dibangun menggunakan Python dan scikit-learn dengan 100 pohon keputusan, kemudian diintegrasikan ke aplikasi web berbasis PHP, MySQL, dan layanan Flask. Evaluasi menggunakan confusion matrix, accuracy, precision, recall, F1-score, AUC-ROC, dan feature importance. Hasil pengujian menunjukkan accuracy 84,4%, precision 85,7%, recall 92,8%, F1-score 89,1%, dan AUC-ROC 88,0%. Variabel pendapatan memiliki kontribusi relatif paling tinggi dalam pembentukan prediksi model. Hasil penelitian menunjukkan bahwa Random Forest dapat diterapkan sebagai alat bantu klasifikasi kelayakan

kredit berbasis data. Namun, karena penelitian menggunakan dataset simulasi, keputusan akhir tetap memerlukan verifikasi dan pertimbangan petugas.

Kata kunci: Random Forest, sistem pendukung keputusan, kelayakan kredit, klasifikasi, Pegadaian.

A. Pendahuluan

Kredit atau pembiayaan memiliki peran penting dalam mendukung kegiatan ekonomi masyarakat. Pegadaian menjadi salah satu alternatif pembiayaan karena menyediakan layanan yang relatif cepat dengan dukungan jaminan. Namun, penentuan kelayakan kredit tetap memerlukan analisis terhadap berbagai karakteristik calon nasabah. Penilaian yang terlalu bergantung pada pengalaman petugas dapat menimbulkan perbedaan dalam proses analisis dan mengurangi konsistensi keputusan.

Perkembangan machine learning membuka peluang untuk mengolah data menjadi dasar rekomendasi keputusan. Salah satu algoritma yang relevan adalah Random Forest, yaitu metode ensemble yang membentuk sejumlah pohon keputusan dan menggabungkan hasil prediksinya. Random Forest telah digunakan pada berbagai penelitian klasifikasi kredit

karena mampu menangani pola data yang kompleks. Penelitian ini menerapkan Random Forest dalam SPK berbasis web untuk mengklasifikasikan kelayakan kredit menjadi Layak dan Tidak Layak.

Penelitian berfokus pada konteks Pegadaian di Kota Bekasi dengan menggunakan dataset simulasi. Tujuan penelitian adalah merancang SPK berbasis web, menerapkan Random Forest pada proses klasifikasi, serta mengukur kinerja model menggunakan accuracy, precision, recall, F1-score, dan AUC-ROC.

B. Metode Penelitian

Penelitian menggunakan pendekatan kuantitatif dengan eksperimen terapan. Dataset penelitian berupa 1.501 data simulasi yang disusun untuk merepresentasikan karakteristik pengajuan kredit. Data bukan arsip internal atau identitas nasabah PT Pegadaian. Setiap data memiliki tujuh

fitur, yaitu pendapatan, pekerjaan, umur, jenis jaminan, nilai jaminan, jumlah pinjaman, dan jangka waktu. Target klasifikasi terdiri atas dua kelas, yaitu Layak dan Tidak Layak.

Tahapan penelitian meliputi pemeriksaan data, data cleaning, transformasi fitur kategorikal, pembagian data, pelatihan model, evaluasi, analisis feature importance, penyimpanan model, dan integrasi ke aplikasi web. Dataset dibagi menggunakan rasio 80:20 dengan `random_state` 42, sehingga menghasilkan 1.200 data latih dan 301 data uji.

Model dibangun menggunakan `RandomForestClassifier` dari `scikit-learn` dengan parameter awal `n_estimators=100`, `criterion='gini'`, `bootstrap=True`, dan `random_state=42`. Arsitektur sistem menggunakan PHP sebagai antarmuka, MySQL sebagai basis data, Python untuk pemodelan, dan Flask sebagai layanan prediksi. Evaluasi dilakukan menggunakan `confusion matrix`, `accuracy`, `precision`, `recall`, `F1-score`, dan

`AUC-ROC`.

C. Hasil Penelitian dan Pembahasan

Hasil pengembangan menunjukkan bahwa sistem dapat mengelola data nasabah, dataset training, proses pelatihan model, prediksi, riwayat prediksi, dan pelaporan. Model `Random Forest` yang telah dilatih diintegrasikan ke dalam layanan `Flask` sehingga aplikasi web dapat mengirim data masukan dan menerima hasil klasifikasi secara otomatis.

Pengujian dilakukan terhadap 301 data uji. `Confusion matrix` menghasilkan 192 `True Positive`, 62 `True Negative`, 32 `False Positive`, dan 15 `False Negative`. Dengan demikian, terdapat 254 prediksi benar dan 47 prediksi salah. Hasil evaluasi menunjukkan `accuracy` sebesar 84,4%, `precision` sebesar 85,7%, `recall` sebesar 92,8%, `F1-score` sebesar 89,1%, dan `AUC-ROC` sebesar 88,0%.

Nilai `recall` yang tinggi menunjukkan bahwa model mampu mengenali sebagian besar data yang termasuk kelas Layak. `Precision` menunjukkan bahwa sebagian besar prediksi Layak sesuai dengan kondisi aktual pada dataset pengujian. `F1-score` menunjukkan keseimbangan

yang baik antara precision dan recall, sedangkan AUC-ROC menunjukkan kemampuan model dalam membedakan dua kelas.

Analisis feature importance menunjukkan bahwa pendapatan memiliki kontribusi relatif paling tinggi dalam pembentukan prediksi model. Setelah itu, kontribusi relatif berasal dari umur, nilai jaminan, jumlah pinjaman, pekerjaan, jangka waktu, dan jenis jaminan. Hasil ini hanya menunjukkan tingkat kontribusi relatif pada model dan tidak membuktikan hubungan sebab-akibat.

Meskipun hasil evaluasi menunjukkan kinerja yang baik, masih terdapat kesalahan klasifikasi. False Positive dapat memberikan rekomendasi Layak pada data yang aktualnya Tidak Layak, sedangkan False Negative dapat memberikan rekomendasi Tidak Layak pada data yang aktualnya Layak. Oleh karena itu, sistem digunakan sebagai alat bantu pengambilan keputusan dan tidak menggantikan kewenangan petugas. Selain itu, penggunaan dataset simulasi menjadi keterbatasan utama penelitian sehingga hasil belum dapat digeneralisasi langsung ke seluruh

nasabah atau kondisi operasional Pegadaian.

D. Kesimpulan

Penelitian berhasil merancang dan membangun Sistem Pendukung Keputusan berbasis web yang menerapkan algoritma Random Forest untuk klasifikasi kelayakan kredit pada konteks Pegadaian di Kota Bekasi. Sistem mengintegrasikan pengelolaan data, pelatihan model, prediksi, riwayat prediksi, dan pelaporan.

Model Random Forest mampu mengklasifikasikan kelayakan kredit berdasarkan tujuh fitur, yaitu pendapatan, pekerjaan, umur, jenis jaminan, nilai jaminan, jumlah pinjaman, dan jangka waktu. Pada 301 data uji, model menghasilkan accuracy 84,4%, precision 85,7%, recall 92,8%, F1-score 89,1%, dan AUC-ROC 88,0%. Pendapatan memiliki kontribusi relatif paling tinggi dalam pembentukan prediksi model.

Penelitian selanjutnya dapat menggunakan data operasional yang sah dan telah dianonimkan, menambah jumlah serta keragaman data, membandingkan Random Forest dengan algoritma lain,

melakukan hyperparameter tuning, dan menerapkan validasi silang untuk meningkatkan kemampuan generalisasi model.

DAFTAR PUSTAKA

- Adrian, A., & Verawati, I. (2025). Analisis Performa Logistic Regression dan Random Forest dalam Klasifikasi Kelayakan Penerimaan Kredit. *The Indonesian Journal of Computer Science Research*, 4(2), 148–158. <https://doi.org/10.59095/ijcsr.v4i2.205>
- Amrin, A., Pahlevi, O., & Rianto, H. (2025). Analisa Komparasi Model Data Mining Algoritma C4.5, CHAID, dan Random Forest untuk Penilaian Kelayakan Kredit. *Computer Science (CO-SCIENCE)*, 5(1), 49–57. <https://doi.org/10.31294/coscience.v5i1.6208>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Christiadi, C. O., & Fenriana, I. (2025). Perbandingan Algoritma Naive Bayes dan Random Forest untuk Klasifikasi Kelayakan Kredit pada PT. BPR Magga Jaya Utama. *POTERS (Proceedings of Technology, Engineering and Computers)*.
- Firmansah, N., Indahyanti, U., & Eviyanti, A. (2023). Prediksi Kelayakan Pemberian Kredit Menggunakan Metode Random Forest. *Jurnal Ilmiah Komputasi*, 22(4), 605–610. <https://doi.org/10.32409/jikstik.22.4.3515>
- Kurniawan, R. (2024). Application of Random Forest Algorithm on Credit Risk Analysis. *Procedia Computer Science*, 245, 740–749. <https://doi.org/10.1016/j.procs.2024.10.300>
- PT Pegadaian. (2026). Kinerja Pegadaian Menguat, Aset Tembus Rp183,8 Triliun.
- Rahmawati, T. I., & Handayani, I. (2025). Implementasi Random Forest untuk Klasifikasi Kelayakan Kredit Nasabah dalam Pengelolaan Risiko Keuangan di PT BPR BKK Wonogiri. *INTECOMS: Journal of Information Technology and Computer Science*, 8(6), 1937–1946.
- Tang, L., Cai, F., & Ouyang, Y. (2019). Applying a nonparametric random forest algorithm to assess the credit risk of the energy industry in China. *Technological Forecasting and Social Change*, 144, 563–572. <https://doi.org/10.1016/j.techfore.2018.03.007>