

Deteksi Risiko Kredit Pada Peer to Peer *Lending* Menggunakan CatBoost

Fadhurrahman Akbar Nasution¹, Siti Saadah², Prasti Eko Yunanto³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹fadhurrahmanakbar@students.telkomuniversity.ac.id, ²sitisaadah@telkomuniversity.ac.id,

³gppras@telkomuniversity.ac.id

Abstrak

P2P (*Peer-to-peer*) *lending* sebagai salah satu pilihan yang populer di kalangan peminjam swasta, usaha kecil, dan UMKM karena kemampuannya untuk memberikan akses langsung ke pinjaman tanpa persyaratan ketat yang diberlakukan oleh bank tradisional dan lembaga keuangan. Namun, P2P *lending* juga menghadapi tantangan besar dalam hal risiko kredit, yang menyebabkan tingkat kegagalan pembayaran pinjaman yang tinggi. Penelitian ini bertujuan untuk mengembangkan sistem deteksi risiko kredit dengan menggunakan data pinjaman dari perusahaan Bondora dan mengaplikasikan salah satu algoritma *gradient boosting*, yakni metode CatBoost (*Categorical Boosting*). Kinerja algoritma CatBoost dievaluasi menggunakan kurva ROC (*Receiver Operating Characteristics*) dan AUC (*Area Under Curve*). Lima skenario dilakukan, dan hasilnya menunjukkan bahwa skenario 2, dengan rasio pemisahan data 90:10, mencapai hasil terbaik dengan nilai AUC 0,804600. Skenario 2 ini lebih unggul dibandingkan dengan skenario 1 (rasio pemisahan data 80:20 dengan nilai AUC sekitar 0,791485), skenario 3 (rasio pemisahan data 70:30 dengan nilai AUC sekitar 0,781475), skenario 4 (rasio pemisahan data 80:10 dengan nilai AUC sekitar 0,795319), dan skenario 5 (rasio pemisahan data 70:10 dengan nilai AUC sekitar 0,790287). Pemilihan *hyperparameter* yang sesuai dan pembagian data latihan dengan rasio yang besar dapat meningkatkan kinerja metode CatBoost dalam mengidentifikasi risiko kredit pada P2P *lending*.

Kata kunci : P2P *Lending*, Deteksi Risiko Kredit, CatBoost, AUC, ROC.

Abstract

P2P (*Peer-to-peer*) *lending* has gained popularity among private borrowers, small businesses, and MSMEs due to its ability to provide direct access to loans without the strict requirements imposed by traditional banks and financial institutions. However, P2P *lending* faces a significant challenge in terms of credit risk, resulting in a high rate of loan repayment failures. To address this issue, the study aimed to develop a credit risk detection system using a loan dataset obtained from the Bondora company by implementing one of the gradient boosting algorithm which are called as the CatBoost (*Categorical Boosting*) method. The performance of the CatBoost algorithm was evaluated using ROC (*Receiver Operating Characteristics*) curves and AUC (*Area Under Curve*). In the five scenarios considered, scenario 2, which involved a 90:10 data split, achieved the top performance with an AUC of 0.805. It outperformed scenario 1 (80:20 data split, AUC ~0.791), scenario 3 (70:30 data split, AUC ~0.781), scenario 4 (80:10 data split, AUC ~0.795), and scenario 5 (70:10 data split, AUC ~0.790). Optimizing hyperparameters and dividing training data with a large ratio can improve the performance of the CatBoost method in identifying credit risk in P2P *lending*.

Keywords: P2P *Lending*, Credit Risk Detection, CatBoost, AUC, ROC.

1. Pendahuluan

Latar Belakang

P2P *lending* semakin diminati oleh individu, usaha kecil, dan UMKM karena memberikan persyaratan yang lebih fleksibel dan kriteria yang tidak seketat bank dan lembaga keuangan tradisional [1]. Meskipun begitu, P2P *lending* memiliki tingkat kegagalan yang tinggi dalam pembayaran, yang dikenal sebagai risiko kredit [1]–[3]. Risiko kredit menjadi masalah utama dalam P2P *lending* karena ketidakmampuan pembayar untuk melunasi pinjaman menyebabkan kerugian bagi pemberi pinjaman [2]. Untuk mengurangi risiko kredit, diperlukan sistem yang dapat membantu menilai risiko kredit.

Untuk mengatasi tantangan tersebut, sebuah studi oleh M. Bazarbash [4] merekomendasikan penggunaan pembelajaran mesin dalam menilai risiko kredit dalam P2P *lending*. Pembelajaran mesin melibatkan penerapan algoritma canggih yang dapat dijalankan oleh mesin untuk mengidentifikasi pola dalam data, dengan tujuan utama membuat prediksi. Model pembelajaran mesin dirancang khusus untuk menganalisis informasi lengkap dari berbagai sumber data, sehingga memungkinkan deteksi pola yang mungkin terlewatkan oleh model ekonometrik konvensional. Dengan pemahaman yang lebih baik tentang hubungan kompleks dan non linier antara faktor risiko dan hasil risiko kredit, teknik pembelajaran mesin menawarkan solusi yang menjanjikan untuk mengatasi risiko kredit dalam P2P *lending*. Teknik ini juga memiliki kemampuan untuk mengurangi masalah asimetri informasi, yang merupakan inti tantangan dalam P2P *lending*. Sebagai hasilnya, deteksi risiko kredit dalam P2P *lending* dapat memainkan peran penting dalam mengurangi kegagalan pembayaran pinjaman.