BAB I
PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi finansial (*finansial technology* atau *fintech*) di Indonesia telah mendorong transformasi besar dalam layanan keuangan digital. Salah satu inovasi yang berkembang pesat adalah pinjaman *online* atau *online lending*, yang memberikan akses pembiayaan cepat, fleksibel, dan tanpa persyaratan kompleks seperti perbankan tradisional. Kemudahan proses pengajuan dan pencairan dana menjadikan layanan ini semakin banyak digunakan oleh Masyarakat. Menurut (Fitria *et al*., 2025) dalam jumlah analisa sosiologi, pertumbuhan *fintech lending* di Indonesia tidak hanya dipicu oleh perkembangan teknologi digital, tetapi juga oleh meningkatnya ketergantungan masyarakat terhadap pinjaman berbasis aplikasi yang menawarkan proses pengajuan tanpa jaminan dan pencairan dana yang instan. Layanan pinjaman daring kini telah menjadi bagian penting dari perilaku ekonomi masyarakat modern, terutama pada kelompok usia produktif yang membutuhkan alternatif pembiayaan non-bank.

Namun, di balik pertumbuhan yang pesat tersebut muncul masalah serius yang sering dihadapi oleh penyedia layanan, yaitu risiko keterlambatan pembayaran (*late payment*) dan bahkan mengalami gagal bayar (*default*). Risiko kredit ini tidak hanya mengancam probabilitas dan kelangsungan operasional penyediaan layanan, tetapi juga berpotensi mengganggu stabilitas sistem keuangan dalam skala yang lebih luas. Tingginya tingkat keterlambatan ini menyebabkan Perusahaan *fintech lending* mengalami kesulitan dalam menjaga arus kas dan stabilitas keuangan. Oleh karena itu, sangat penting untuk memiliki sistem prediksi yang akurat untuk mengidentifikasi potensi keterlambatan pembayaran sejak dini menjadi sebuah kunci bagi perusahaan untuk melakukan langkah yang efektif.

Dalam konteks inilah, pendekatan *machine learning* menawarkan solusi yang tepat. Kemampuannya untuk belajar dari data historis dan mengidentifikasi pola-pola kompleks menjadikannya unggul dibandingkan metode *statistik* tradisional. Algoritma yang sering menjadi andalan untuk tugas klasifikasi pada berbagai kasus, termasuk data pinjaman, adalah algoritma *Random Forest* dan *XGBoost*. *Random forest*, sebagai metode *ensemble* berbasis *bagging*, dikenal dengan stabilitasnya dan

ketahanan terhadap *overfitting*. Sementara itu, *XGBoost* sebagai implementasi canggih dari *gradient boosting*, terkenal dengan kecepatan, efisiensi dan kemampuannya menangani hubungan non-linier yang kompleks antar fitur (Chen *et al*., 2022)

Menurut penelitian (Kristanti *et al*., 2024) menunjukkan bahwa algoritma *Random Forest* yang dikombinasikan dengan teknik *Synthetic Minority Oversampling Technique* (SMOTE) mampu mencapai nilai akurasi hingga 96% dan f1-score 0,95 pada prediksi risiko keuangan di Indonesia. Hasil ini tentu mengidentifikasi bahwa *Random Forest* memiliki stabilitas tinggi saat berhadapan dengan data yang tidak seimbang, seperti kasus keterlambatan pembayaran. Di sisi lain, algoritma *XGBoost* sering kali memberikan hasil yang lebih unggul dibandingkan model lainnya, dengan tingkat akurasi yang mencapai 89% pada kasus prediksi gagal bayar pinjaman (Gupta *et al*., 2024). Hasil ini tentu memperkuat anggapan bahwa *XGBoost* lebih efisien dalam mengelola data berukuran besar dengan korelasi variabel yang kompleks.

Menurut (Alamsyah *et al*., 2025) mempublikasikan hasil riset yang menunjukkan bahwa kombinasi model *ensemble Random Forest* dengan *XGBoost* mampu memberikan akurasi hampir sempurna dalam *credit scoring* berbasis data sosial di sektor *fintech* Indonesia. Maka hasil temuan di atas memperkuat bahwa pendekatan berbasis *machine learning* dapat meningkatkan keakuratan sistem penilaian risiko pinjaman digital di Indonesia.

Penelitian ini menggunakan *dataset* anonim berupa variabel numerik dan kategorikal bertanda N1-N316 dan C1-C79, sesuai dengan standar privasi data yang diterapkan pada industri keuangan digital. Hal ini menyebabkan identitas asli fitur tidak ditampilkan, namun pola data masih dapat dianalisis secara statistik dan digunakan untuk model prediksi. Dengan memanfaatkan algoritma *Random Forest* dan *XGBoost*, penelitian ini bertujuan untuk menentukan model yang paling optimal dalam memprediksi keterlambatan pembayaran pinjaman *online* di Indonesia.

Berdasarkan permasalahan tersebut, diperlukan analisis komparatif antar algoritma *machine learning* untuk menentukan metode paling efektif dalam memprediksi keterlambatan pembayaran. Oleh karena itu, penelitian ini secara

khusus difokuskan pada perbandingan performa algoritma *Random Forest* dan *XGBoost* menggunakan *dataset* anonim *fintech* Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana proses penerapan algoritma *Random Forest* dan *XGBoost* dalam membangun model klasifikasi untuk memprediksi keterlambatan pembayaran pinjaman *online*?
2. Bagaimana hasil evaluasi kedua model berdasarkan matriks *accuracy*, *precision*, *recall*, dan *F1-score* untuk menentukan model yang lebih optimal?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk:

1. Menerapkan algoritma *Random Forest* dan *XGBoost* untuk membangun model klasifikasi dalam memprediksi keterlambatan pembayaran pinjaman *online* pada *dataset fintech* yang bersifat anonim.
2. Mengevaluasi dan membandingkan performa kedua model menggunakan matriks *accuracy*, *precision*, *recall dan F1-Score* untuk mengetahui algoritma yang lebih optimal dalam memprediksi keterlambatan pembayaran.

1.4 Manfaat

Adapun manfaat dari penelitian ini antara lain:

1. Manfaat akademis:

   Memberikan kontribusi pada bidang ilmu data dan dengan memberikan studi komparasi empiris antara *Random Forest* dan *XGBoost* dalam konteks prediksi risiko kredit di Indonesia, yang dapat dijadikan referensi bagi penelitian selanjutnya.
2. Manfaat praktis:

   Menyediakan *insight* yang *actionable* bagi Perusahaan *fintech lending* dalam memilih algoritma *machine learning* yang paling efektif untuk

meningkatkan akurasi sistem deteksi dini keterlambatan pembayaran, sehingga dapat menekan angkat *default* dan mempertahankan Kesehatan portofolio pinjaman.

3. Manfaat sosial:

   Penelitian ini di harapkan dapat berkontribusi terhadap penciptaan ekosistem pinjaman digital yang lebih sehat dan berkelanjutan, dengan cara mengurangi praktik pinjaman bermasalah serta membantu pengguna untuk lebih bijak dalam mengelola pinjaman.

UBP KARAWANG