BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil pengujian dan analisis perbandingan performa algoritma *machine learning* dalam kinerja model untuk memprediksi keterlambatan pembayaran pinjaman *online,* penelitian ini menghasilkan beberapa kesimpulan utama yaitu sebagai:

1. Penerapan algoritma Random Forest dan XGBoost dalam penelitian ini dilaksanakan melalui serangkaian tahapan yang sistematis. Dimulai dari pengumpulan *dataset* dari *Lending Club* yang diperoleh dari platform *Kaggle*, dilanjutkan dengan tahap praproses data (*Preprocessing*) yang mencakup pembersihan data, penanganan *missing value, encoding* variabel kategorikal menggunakan Label *Encoding* dan *Ordinal Encoding* serta normalisasi fitur numerik menggunakan *StandarScaller*. Untuk mengatasi ketidakseimbangan kelas, digunakan dua strategi yaitu SMOTE dan *Class Weight*. Selanjutnya, *dataset* dibagi menjadi data latih (80%), data validasi (15%) dan data uji (5%) menggunakan stratified split. Kedua algoritma kemudian dilatih dalam lima skenario berbeda, di mana *Random Forest* diuji dengan skenario SMOTE dan Class Weight sedangkan XGBoost diuji dengan skenario SMOTE, *Class Weight*, serta *GridSearchCV* yang dikombinasikan dengan *Threshold Tuning*. Seluruh model dibangun dalam bentuk *pipeline* yang mengintegrasikan *StandarScaller*, penanganan ketidakseimbangan kelas dan model klasifikasi secara berurutan untuk mencegah terjadinya data *leakage*.
2. Berdasarkan hasil evaluasi menggunakan metriks nilai *precision, recall, F1-Score,* dan *ROC-AUC* pada data uji, algoritma *XGBoost* yang dikombinasikan dengan teknik optimal *GridSearchCV* dan *Threshold Tuning* terbukti menjadi pendekatan paling unggul di bandingkan *Random Forest*. Konfiguarasi ini berhasil mencapai tingkat akurasi sebesar 0,7211. Meski demikian, perlu dicatat bahwa skenario *Random Forest* dengan *Class Weight* menunjukkan kemampuan deteksi kasus *default* yang lebih menyeluruh dengan nilai *recall* dan *F1-Score* yang lebih tinggi sehingga

pemilihan model terbaik tetap bergantung pada prioritas binis yang ingin dicapai perusahaan fintech bersangkutan.

5.2 Saran

Berdasarkan hasil dari analisis dan evaluasi yang telah dilakukan, terdapat beberapa saran strategi yang dapat dijadikan pertimbangan untuk pengembangan penelitian selanjutnya:

1. Penelitian berikutnya di harapkan dapat mengeksplorasi algoritma *gradien boosting* generasi terbaru seperti *CatBoost* atau *LightBM* untuk membandingkan efisiensi prosesnya pada data besar. Selain itu, penerapan metode gabungan seperti SMOTE-Tomek Links patur dipertimbangkan untuk memperjelas batas pemisal antara kelas nasabah lancar dan macet, sehingga kemampuan deteksi *default* dapat ditingkatkan tanpa harus mengorbankan ketepatan prediksi secara keseluruhan.
2. Bagi institusi keuangan maupun penyelenggara pinjaman *online*, model *XGBoost* yang telah dikembangkan dalam penelitian ini layak di gunakan sebagai sistem penyaringan awal pada tahap pengajuan pinjaman. Dengan mengatur ambang batas (*threshold*) secara berkala sesuai dengan kondisi portofolio pinjaman dan toleransi risiko perusahaan, sistem ini dapat membantu memperluas jangkauan persetujuan kredit pada segmen nasabah yang berisiko rendah, sekaligus tetap menekan risiko kesalahan prediksi yang merugikan potensi nasabah.

5.3 Untuk penelitian selanjutnya, disarankan agar menggunakan *dataset* yang lebih beraga, dan tidak hanya terbatas pada *dataset* anonim, sehingga interpretasi fitur dapat dilakukan secara lebih mendalam dan hasil penelitian dapat digeneralisasi pada konteks *fintech lending* yang lebih luas di indonesia. Penggunaan data *real-time* dari *platform fintech* aktif juga dapat meningkatkan relevansi dan kebaruan temuan penelitian.