

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini menguji dua model klasifikasi risiko kredit, yaitu *Random Forest* dan *Logistic Regression*, dengan skema pemodelan yang konsisten: preprocessing, pembagian data 80/20 secara stratifikasi, penerapan SMOTE hanya pada data latih, serta *hyperparameter tuning* berbasis cross-validation, evaluasi akhir dilakukan pada data uji dengan distribusi kelas asli.

Pada data uji, kedua model menunjukkan kinerja yang sangat baik, namun *Logistic Regression* lebih unggul dalam mendeteksi kontrak *default*=1. *Logistic Regression* menghasilkan TN=1528, FP=15, FN=2, TP=62 dengan *Recall* 0,9688, *Precision* 0,8052, F1 0,8794, sedangkan *Random Forest* menghasilkan TN=1524, FP=19, FN=5, TP=59 dengan *Recall* 0,9219, *Precision* 0,7564, F1 0,8310. Perbedaan yang paling kritis berada pada FN 2 vs 5 karena berkaitan langsung dengan *default* yang terlewat dalam fungsi *early warning*.

Simulasi dampak bisnis berbasis PD–EAD–LGD (LGD 45%) dengan asumsi efektivitas mitigasi 25% dan biaya tindak lanjut Rp100.000 per kontrak menunjukkan manfaat bersih *Logistic Regression* lebih tinggi Rp 832.492.375 dibanding *Random Forest* Rp 791.737.905, dengan beban tindak lanjut yang relatif sama 77 vs 78 kontrak, sekitar 4,8% portofolio uji. Temuan ini menunjukkan bahwa keunggulan model tidak hanya terlihat pada metrik, tetapi juga berdampak pada proyeksi manfaat operasional.

Hasil interpretasi XAI (SHAP) menegaskan bahwa keputusan model banyak dipengaruhi indikator saldo/likuiditas, kecukupan saldo, saldo rekening, sisa saldo, informasi payroll, serta variabel terkait jatuh tempo. *Random Forest* cenderung menggabungkan banyak sinyal kecil secara simultan, sedangkan *Logistic Regression* menampilkan kontribusi yang lebih terstruktur pada beberapa fitur utama, sehingga lebih mudah dijadikan dasar komunikasi dan monitoring risiko.

Berdasarkan kombinasi hasil evaluasi, simulasi dampak bisnis, dan interpretabilitas, *Logistic Regression* setelah tuning merupakan model yang paling tepat untuk prediksi risiko kredit pada dataset penelitian ini, khususnya untuk kebutuhan deteksi dini kontrak berisiko.

5.2 Saran

Optimasi threshold berbasis biaya, bukan *default* 0,5. Penentuan cut-off sebaiknya mengikuti trade-off biaya *False Negative* vs *False Positive* dan kapasitas tindak lanjut. Disarankan membuat risk band (rendah-sedang-tinggi) berbasis rentang PD agar keputusan operasional tidak kaku biner.

Kalibrasi probabilitas (PD) sebelum dipakai untuk keputusan bisnis/ECL. Lakukan kalibrasi misal platt scaling atau isotonic regression dan evaluasi calibration misal brier score/curve agar PD benar-benar mencerminkan probabilitas *default* dan lebih layak digunakan untuk perhitungan berbasis PD–EAD–LGD.

Perkuat simulasi dampak bisnis agar lebih realistis. Skenario saat ini menggunakan LGD 45%, efektivitas mitigasi 25%, dan biaya tindak lanjut Rp100.000/kontrak. Penelitian lanjutan sebaiknya memakai LGD berbasis data *recovery* actual dan menambahkan sensitivity analysis beberapa skenario untuk melihat stabilitas manfaat bersih.

Uji generalisasi dengan validasi lintas waktu dan perluasan periode data. Gunakan time-based split atau train periode awal, test periode berikutnya untuk menguji ketahanan model terhadap perubahan portofolio dan memastikan performa tidak hanya bagus pada snapshot tertentu.

Siapkan kerangka implementasi dan governance interpretabilitas. Pemetaan hasil SHAP ke fitur asli sebelum one-hot perlu dibuat agar mudah dipahami unit bisnis, disertai dokumentasi singkat definisi fitur, batas interpretasi, contoh kasus TP/FN. Jika diimplementasikan sebagai early warning, tetapkan alur tindakan per risk band dan monitoring berkala FN/FP, distribusi fitur kunci, serta jadwal retraining.