

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini telah berhasil menerapkan lima metode *ensemble learning*, yaitu *Random Forest*, *AdaBoost*, *XGBoost*, *LightGBM*, dan *ExtraTrees*, dalam klasifikasi kanker payudara menggunakan dataset *Breast Cancer Wisconsin (Original)*. Sebagai model pembandingan, digunakan algoritma *Decision Tree* yang merupakan model pembelajaran mesin dasar. Evaluasi dilakukan menggunakan lima metrik utama, yaitu *accuracy*, *precision*, *recall*, *F1-score*, dan *AUC-ROC* untuk mengukur performa masing-masing model secara objektif.

Hasil penelitian menunjukkan bahwa semua model *ensemble* mampu melakukan klasifikasi dengan baik, namun *LightGBM* memberikan hasil paling optimal secara keseluruhan. Model ini mencatat *accuracy* sebesar 0,9708, *precision* 0,9821, *recall* 0,9483, *F1-score* 0,9649, serta *AUC-ROC* sebesar 0,9948. Angka-angka ini menjadi bukti bahwa *LightGBM* memiliki keseimbangan terbaik antara ketepatan dan sensitivitas dalam mendeteksi kanker payudara. Model *ExtraTrees* mencatatkan nilai *AUC-ROC* tertinggi, yaitu 0,9961, sedangkan *Random Forest* menunjukkan performa sangat baik pada *F1-score* dan *AUC*. Model *XGBoost* memiliki keseimbangan yang cukup solid antara *precision* dan *recall*, sementara *AdaBoost*, meskipun memiliki *precision* tinggi, masih kurang optimal dalam aspek *recall*, sehingga cenderung melewati beberapa kasus positif. Model *Decision Tree* menempati posisi paling bawah, dengan nilai *recall* dan *AUC-ROC* terendah di antara seluruh model.

Dengan demikian, seluruh rumusan masalah yang diajukan dalam penelitian ini telah terjawab secara menyeluruh. Hasil prediksi menggunakan metode *ensemble learning* menunjukkan performa yang baik, dan model *LightGBM* direkomendasikan sebagai model yang paling akurat dan seimbang untuk klasifikasi kanker payudara berdasarkan dataset yang digunakan.

5.2 Saran

Penelitian ini masih memiliki keterbatasan, khususnya pada jumlah data dan cakupan informasi yang tersedia dalam dataset. Oleh karena itu, disarankan agar penelitian selanjutnya menggunakan dataset yang lebih besar dan lebih kompleks, dengan atribut tambahan seperti usia pasien, riwayat keluarga, dan stadium kanker, agar hasil prediksi menjadi lebih relevan dan aplikatif dalam dunia medis.

Selain itu, proses pelatihan model dalam penelitian ini masih menggunakan pengaturan *hyperparameter* standar. Penelitian lanjutan sebaiknya melakukan optimasi *hyperparameter* secara menyeluruh menggunakan pendekatan seperti *Grid Search*, *Randomized Search*, atau teknik optimasi lainnya guna memperoleh konfigurasi terbaik untuk masing-masing model. Penggunaan teknik *ensemble* lanjutan seperti *stacking* juga sangat potensial untuk diteliti, karena dapat menggabungkan keunggulan dari beberapa model dalam satu kerangka klasifikasi.

Sebagai bentuk implementasi nyata, hasil penelitian ini sangat memungkinkan untuk dikembangkan menjadi sistem pendukung keputusan dalam bentuk aplikasi berbasis web atau mobile yang dapat digunakan oleh tenaga medis. Dengan adanya sistem berbasis *machine learning* ini, proses deteksi dini kanker

payudara diharapkan dapat dilakukan lebih cepat, akurat, dan efisien sehingga berkontribusi dalam peningkatan kualitas layanan kesehatan.