

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai klasifikasi kanker payudara menggunakan algoritma Support Vector Machine (SVM) dan Decision Tree, dapat disimpulkan bahwa SVM memberikan performa yang lebih unggul dalam mendeteksi kanker payudara dibandingkan Decision Tree. Hal ini ditunjukkan oleh nilai akurasi sebesar 98,25%, precision 98,61%, recall 98,61%, F1-score 98,61%, dan AUC sebesar 99,37% yang diperoleh oleh model SVM, menunjukkan tingkat presisi dan kemampuan generalisasi yang sangat baik. Sementara itu, algoritma Decision Tree juga menunjukkan kinerja yang kompetitif dengan akurasi 92,11%, precision 94,37%, recall 93,06%, F1-score 93,71%, dan AUC 92,68%, serta memiliki kelebihan dari segi interpretabilitas melalui feature importance yang menunjukkan bahwa "worst radius" merupakan fitur paling berpengaruh dalam klasifikasi. Hasil hyperparameter tuning menggunakan GridSearchCV menunjukkan bahwa parameter optimal untuk SVM adalah $C=0.1$, kernel='linear', dan gamma='scale' dengan CV score 97,80%, sedangkan untuk Decision Tree adalah max_depth=5, min_samples_split=10, min_samples_leaf=5, dan min_impurity_decrease=0.001 dengan CV score 93,63%. Analisis confusion matrix menunjukkan bahwa SVM hanya melakukan 2 kesalahan prediksi (1 False Positive dan 1 False Negative) dari 114 sampel uji, sedangkan Decision Tree melakukan 9 kesalahan (4 False Positive dan 5 False Negative). Keseluruhan hasil penelitian ini mengindikasikan bahwa SVM merupakan pilihan algoritma yang lebih optimal untuk kasus klasifikasi kanker payudara berbasis data Breast Cancer Wisconsin Diagnostic. Dengan demikian, tujuan dan rumusan masalah dalam penelitian ini telah terjawab melalui analisis dan evaluasi model yang telah disajikan pada bab sebelumnya.

5.2 Saran

Penelitian ini masih dapat dikembangkan lebih lanjut. Disarankan untuk mengeksplorasi algoritma ensemble seperti Random Forest atau XGBoost yang

dapat memberikan performa lebih baik dengan menggabungkan kelebihan dari multiple decision trees. Selain itu, penggunaan dataset yang lebih bervariasi atau data nyata dari instansi medis lokal dapat memperkuat generalisasi model dan validasi eksternal. Penelitian selanjutnya juga dapat mengimplementasikan teknik feature selection atau feature engineering untuk meningkatkan efisiensi model dengan mengurangi dimensi fitur tanpa mengurangi akurasi. Selain itu, penerapan teknik handling imbalanced data seperti SMOTE dapat dipertimbangkan jika menggunakan dataset dengan ketidakseimbangan kelas yang lebih ekstrem. Terakhir, penelitian dapat dilengkapi dengan pengembangan sistem pendukung keputusan berbasis web atau mobile yang mengintegrasikan model SVM terbaik untuk membantu tenaga medis dalam proses deteksi dini kanker payudara secara lebih praktis dan efisien.

