

Abstrak

Kanker tiroid, meskipun relatif jarang (0,85–2,5% dari seluruh kasus kanker), menjadi perhatian serius karena prevalensinya yang lebih tinggi pada perempuan dan tantangan dalam diagnosis akibat keterbatasan metode konvensional seperti *fine-needle aspiration biopsy* dan USG. Penelitian ini bertujuan untuk prediksi risiko kekambuhan kanker tiroid dengan menerapkan metode *Random Forest*, *XGBoost*, dan *Logistic Regression*. Mengklasifikasikan kekambuhan penyakit kanker tiroid menggunakan 14 variabel dataset yang di peroleh dari rumah sakit ken saras yang berjumlah 2000 dataset. Data yang digunakan diklasifikasikan menggunakan 3 model metode dan di evaluasi menggunakan *confusion matrix* untuk mencari nilai evaluasi akurasi yang terbaik. Berdasarkan hasil evaluasi *Logistic Regression* mendapatkan nilai akurasi sebesar 83%, serta *Random Forest* dan *XGBoost* mendapatkan akurasi sebesar 82%. Hasil temuan kami membuktikan bahwa pendekatan machine learning dapat berfungsi sebagai sistem pendukung keputusan klinis yang efektif dalam meningkatkan efisiensi diagnosis dan memfasilitasi intervensi medis yang tepat waktu. Implementasi ini dalam praktik klinis tetap memerlukan integrasi dengan pertimbangan medis komprehensif dan supervisi profesional kesehatan untuk menjamin keamanan. Hasil penelitian memberikan kontribusi bagi pengembangan alat diagnostik kanker tiroid yang lebih andal dan efisien.

Kata kunci: kanker tiroid; keputusan klinis; logistic regression; *random forest*; *xgboost* .

Abstract

Thyroid cancer, although relatively rare (0.85-2.5% of all cancer cases), is of serious concern due to its higher prevalence in women and challenges in diagnosis due to limitations of conventional methods such as fine-needle aspiration biopsy and ultrasound. This study aims to predict the risk of thyroid cancer recurrence by applying Random Forest, XGBoost, and Logistic Regression methods. Classifying the recurrence of thyroid cancer using 14 dataset variables obtained from ken saras hospital which amounted to 2000 datasets. The data will be classified using 3 method models and evaluated using confusion matrix to find the best accuracy evaluation value. Based on the evaluation results Logistic Regression gets an accuracy value of 83%, and Random Forest and XGBoost get an accuracy of 82%. Our findings prove that machine learning approaches can serve as an effective clinical decision support system in improving diagnosis efficiency and facilitating timely medical interventions. The implementation of this in clinical practice still requires integration with comprehensive medical considerations and supervision of healthcare professionals to ensure safety. The results contribute to the development of more reliable and efficient thyroid cancer diagnostic tools.

Keywords: clinical decision; logistic regression; *random forest*; thyroid cancer; *xgboost* .

