

ABSTRAK

Stroke merupakan salah satu penyebab utama mortalitas dan disabilitas jangka panjang di dunia, dengan pola konsumsi nutrisi yang tidak seimbang menjadi salah satu faktor risiko utama selain hipertensi, diabetes melitus, dan obesitas. Model machine learning dengan performa tinggi umumnya bersifat black box sehingga sulit dijelaskan, padahal transparansi sangat penting dalam bidang kesehatan. Penelitian ini bertujuan mengimplementasikan algoritma Random Forest untuk mengklasifikasikan tingkat risiko stroke berdasarkan pola konsumsi nutrisi, serta menerapkan metode SHAP (SHapley Additive exPlanations) untuk menginterpretasikan kontribusi setiap variabel terhadap hasil prediksi. Data penelitian diperoleh dari Puskesmas Kanjilo, Kecamatan Barombong, Kabupaten Gowa, sebanyak 4.603 baris data dengan 19 atribut, yang kemudian melalui tahap preprocessing sebelum dibagi menjadi data latih (2.941 data) dan data uji (736 data) dengan proporsi 80:20. Model Random Forest dilatih menggunakan 100 pohon keputusan dan dievaluasi menggunakan confusion matrix, accuracy, precision, recall, dan F1-score. Hasil pengujian menunjukkan model memperoleh akurasi sebesar 99,05%, dengan nilai precision, recall, dan F1-score masing-masing sebesar 99%. Interpretasi SHAP menunjukkan bahwa atribut serat memberikan kontribusi terbesar terhadap prediksi risiko stroke, diikuti oleh BMI, lemak, tekanan darah diastolik, hipertensi, tekanan darah sistolik, berat badan, dan kolesterol. Hasil penelitian ini membuktikan bahwa integrasi Random Forest dan SHAP mampu menghasilkan model klasifikasi risiko stroke yang akurat sekaligus transparan dan dapat diinterpretasikan.

Kata Kunci: *Random Forest; SHAP; risiko stroke; pola konsumsi nutrisi; Explainable Artificial Intelligence (XAI).*