

PREDIKSI PENYAKIT STROKE MENGGUNAKAN ALGORITMA K- NEAREST NEIGHBORS DAN OPTIMASI TEKNIK BAGGING

ABSTRAK

Stroke merupakan salah satu penyebab utama kematian dan kecacatan di seluruh dunia. Prediksi dini risiko stroke menjadi krusial untuk mengurangi dampak penyakit ini. Penelitian ini menerapkan algoritma K-Nearest Neighbors (KNN) yang dikombinasikan dengan teknik Bootstrap Aggregating (Bagging) untuk meningkatkan akurasi prediksi. Hasil penelitian menunjukkan bahwa penggunaan Bagging berhasil meningkatkan akurasi model hingga 95,03%, dengan nilai Area Under Curve (AUC) mencapai 71,7%, yang tergolong dalam kategori Fair Classification. Analisis lebih lanjut mengidentifikasi atribut seperti Age, Heart Disease, dan Ever Married sebagai faktor utama yang memengaruhi prediksi stroke, dengan kontribusi yang lebih stabil berkat teknik Bagging. Tanpa penggunaan Bagging, atribut Age dan Hypertension menjadi lebih dominan, namun model rentan terhadap overfitting. Teknik Bagging membantu mengatasi bias data dan meningkatkan stabilitas model, sehingga menghasilkan prediksi yang lebih akurat dan adil. Temuan ini menunjukkan bahwa pendekatan Bagging pada algoritma KNN dapat menjadi metode yang efektif dalam prediksi risiko stroke berbasis data.

Kata Kunci: Stroke, Prediksi Risiko, K-Nearest Neighbors, Bagging, Bootstrap Aggregating, Area Under Curve (AUC), Fair Classification, Overfitting