

ABSTRAK

Infeksi Saluran Pernapasan Akut (ISPA) menjadi satu dari sekian penyakit pernapasan yang sering ditemukan di masyarakat serta memiliki gejala yang beragam, sehingga diperlukan pendekatan berbasis data untuk membantu proses klasifikasi penyakit secara lebih sistematis. Penelitian ini dimaksudkan guna membandingkan kinerja algoritma Decision Tree dan K-Nearest Neighbor (KNN) dalam mengklasifikasikan penyakit ISPA. Dataset yang digunakan adalah data Pediatric Respiratory Infections yang terdiri dari 801 data pasien dengan 91 atribut awal. Label klasifikasi dibentuk berdasarkan kolom Main diagnostic dan dikelompokkan menjadi lima kategori, yaitu Asthma/Bronchospasm/Wheezing, Pneumonia/Pneumopathy, Bronchiolitis, Laryngeal/Upper Respiratory, dan Other. Setelah proses seleksi fitur, digunakan 54 fitur yang terdiri dari 24 fitur numerik dan 30 fitur kategorikal.

Tahapan penelitian meliputi preprocessing data, pembentukan label, penanganan nilai kosong, encoding data kategorikal, normalisasi data untuk KNN, pembagian data latih serta data uji dengan rasio 80:20, hingga evaluasi model dengan confusion matrix, accuracy, precision weighted, recall weighted, dan F1-score weighted. Hasil penelitian memperlihatkan bahwa algoritma Decision Tree memperoleh accuracy sebesar 67,08%, precision weighted sebesar 73,72%, recall weighted sebesar 67,08%, dan F1-score weighted sebesar 69,12%. Sementara itu, algoritma KNN memperoleh accuracy sebesar 65,84%, precision weighted sebesar 63,98%, recall weighted sebesar 65,84%, dan F1-score weighted sebesar 64,18%. Berdasarkan hasil tersebut, Decision Tree mempunyai performa yang lebih baik dibanding KNN dalam klasifikasi penyakit ISPA pada dataset yang digunakan. Selain itu, fitur yang paling berpengaruh dalam proses klasifikasi adalah Wheezing, Laryngeal stridor, dan Age (months). Penelitian ini memperlihatkan bahwa algoritma machine learning dapat dipergunakan sebagai pendekatan komputasi dalam membantu analisis data penyakit pernapasan, namun hasil klasifikasi tidak dimaksudkan untuk menggantikan diagnosis dokter atau tenaga medis.

Kata Kunci: ISPA, Decision Tree, K-Nearest Neighbor, Klasifikasi, Machine Learning.