

BAB I PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi informasi mendorong pemanfaatan data di bidang kesehatan. Selain menyimpan data, teknologi dapat membantu analisis, klasifikasi, prediksi, dan pengambilan keputusan. Pemanfaatan machine learning membuat pengolahan data pasien menjadi lebih cepat, sistematis, dan objektif (1,2).

Infeksi Saluran Pernapasan Akut (ISPA) ialah gangguan pernapasan yang umumnya dialami masyarakat. Penyakit ini dapat menyerang berbagai usia dengan gejala seperti demam, pilek, batuk, sesak napas, sakit tenggorokan, serta gangguan pernapasan lainnya (3,4).

Kemiripan gejala ISPA dengan penyakit pernapasan lain menyebabkan identifikasi awal tidak selalu mudah. Banyaknya pasien dan variasi kondisi klinis juga menuntut proses analisis yang cepat dan tepat, sehingga diperlukan pendekatan berbasis data yang lebih terukur (4,5).

Klasifikasi dalam machine learning dapat dimaksudkan guna mengelompokkan data pasien berdasarkan pola atribut, seperti gejala, tanda vital, dan indikator klinis. Pendekatan ini membantu analisis penyakit secara komputasional tanpa menggantikan peran tenaga medis (1,6)

Decision Tree merupakan algoritma klasifikasi berbentuk pohon keputusan. Alur keputusannya mudah ditelusuri karena terbentuk dari aturan pada setiap atribut, sehingga sesuai untuk penelitian kesehatan yang membutuhkan interpretasi model (7,8).

K-Nearest Neighbor (KNN) menentukan kelas dengan berlandaskan pada kedekatan data pengujian dengan data latih. Algoritma ini sederhana dan fleksibel, namun kinerjanya dipengaruhi oleh skala data, jumlah fitur, pemilihan nilai K, dan metode jarak (9,10).

Perbedaan cara kerja Decision Tree dan KNN menjadi dasar perbandingan performa menggunakan akurasi, presisi, recall, F1-score, confusion matrix, dan validasi model (11,12).

Penelitian ini berjudul “Perbandingan Algoritma Decision Tree dan K-Nearest Neighbor untuk Klasifikasi Penyakit ISPA” dan bertujuan mengetahui algoritma yang lebih optimal untuk klasifikasi ISPA.

1.2 Rumusan Masalah

Merujuk pada penguraian latar belakang, kemudian dirumuskan permasalahan penelitian berikut:

1. Bagaimana penerapan algoritma Decision Tree dan K-Nearest Neighbor dalam klasifikasi penyakit ISPA?
2. Bagaimana kinerja Decision Tree dalam mengklasifikasikan penyakit ISPA pada dataset penelitian?
3. Bagaimana kinerja K-Nearest Neighbor dalam mengklasifikasikan penyakit ISPA pada dataset penelitian?
4. Algoritma manakah yang memiliki performa lebih baik berdasarkan metrik evaluasi yang digunakan?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan Penelitian

Penelitian ini dimaksudkan guna:

1. Mengaplikasikan algoritma Decision Tree dan K-Nearest Neighbor untuk klasifikasi penyakit ISPA.
2. Mengevaluasi kinerja Decision Tree berdasarkan hasil pengujian model.
3. Mengevaluasi kinerja K-Nearest Neighbor berdasarkan hasil pengujian model.
4. Membandingkan performa kedua algoritma untuk menentukan metode yang lebih optimal.

1.3.2 Manfaat Penelitian

Manfaat yang diharapkan dengan pelaksanaan penelitian ini mencakup:

1. Bagi peneliti, penelitian ini memperluas pemahaman tentang pengimplementasian algoritma klasifikasi pada data kesehatan.
2. Bagi akademisi, penelitian ini mampu menjadi rujukan terkait penggunaan machine learning pada kasus penyakit ISPA.
3. Bagi pengembangan informatika kesehatan, penelitian ini memberi gambaran perbandingan algoritma klasifikasi pada data medis.
4. Bagi peneliti selanjutnya, penelitian ini mampu menjadi rujukan bagi pengembangan algoritma, fitur, atau sistem klasifikasi yang lebih lengkap.

1.4 Batasan Masalah

Dalam menghasilkan pembahasan yang lebih terstruktur, kemudian dirumuskan pembatasan masalah penelitiannya, yakni:

1. Penelitian berfokus pada klasifikasi penyakit ISPA menggunakan data terstruktur.
2. Algoritma yang dibandingkan hanya Decision Tree dan K-Nearest Neighbor.
3. Hasil penelitian tidak dimaksudkan menggantikan diagnosis dokter, tetapi sebagai kajian komputasi.
4. Evaluasi dengan akurasi, presisi, recall, F1-score, confusion matrix, dan metrik relevan lainnya.
5. Pembahasan difokuskan pada proses klasifikasi dan perbandingan performa algoritma, bukan kajian klinis mendalam.
6. Hasil penelitian dipengaruhi oleh kualitas, kelengkapan, dan karakteristik dataset yang digunakan.

1.5 Keterbaruan

Penelitian terbaru memperlihatkan bahwa teknologi komputasi semakin banyak dimanfaatkan untuk mendukung deteksi dan klasifikasi penyakit. Caldwell et al. (2025) menekankan pentingnya diagnosis cepat pada infeksi saluran pernapasan akut melalui pendekatan point-of-care dan pengujian molekuler multiplex, terutama bagi kelompok rentan (4).

Pada penerapan machine learning, Yildiz dan Kalayci (2024) membahas model berbasis decision tree pada data medis tabular untuk mendukung diagnosis penyakit. Penelitian tersebut memperlihatkan bahwa data medis terstruktur dapat dianalisis menggunakan model yang memiliki performa baik serta mudah diinterpretasikan (7).

Chadha dan Singh (2025) membandingkan beberapa algoritma klasifikasi, termasuk Decision Tree serta K-Nearest Neighbor, sedangkan Simarmata et al. (2024) mengevaluasi Decision Tree, Naive Bayes, KNN, dan SVM. Kedua penelitian tersebut menegaskan bahwa perbandingan algoritma penting karena setiap metode memiliki karakteristik dan performa yang berbeda (11,12).

Keterbaruan penelitian ini terletak pada fokus perbandingan Decision Tree dan K-Nearest Neighbor dalam mengklasifikasikan penyakit ISPA. Dengan membandingkan dua algoritma pada data yang sama, penelitian ini diharapkan dapat

memperlihatkan algoritma yang lebih sesuai berdasarkan hasil evaluasi model.