

BAB I

PENDAHULUAN

1.1. LATAR BELAKANG

Kesehatan masyarakat global saat ini menghadapi tantangan serius dengan meningkatnya prevalensi penyakit tidak menular (PTM), khususnya stroke yang memiliki dampak sosial dan ekonomi yang signifikan[1]. Stroke merupakan gangguan pada pembuluh darah otak (serebrovaskular) yang terjadi ketika suplai darah dan oksigen ke otak mengalami hambatan. Kondisi ini dapat menyebabkan kerusakan atau kematian jaringan otak serta mengakibatkan gangguan fungsi otak. Penyempitan, penyumbatan, atau pecahnya pembuluh darah arteri di otak dapat mengurangi aliran darah yang diperlukan otak[2]. Penyakit stroke dipilih sebagai fokus penelitian ini karena merupakan penyebab utama kematian dan kecacatan di seluruh dunia, memiliki angka prevalensi yang tinggi, serta membutuhkan deteksi dini yang cepat dan akurat untuk mengurangi risiko fatalitas dan beban ekonomi[3]. Kesadaran masyarakat yang masih rendah mengenai penyakit stroke menyebabkan mereka sering mengabaikan tanda-tanda awal penyakit ini. Hal tersebut sangat disayangkan karena deteksi dan penanganan dini gejala stroke dapat meningkatkan peluang kesembuhan dan pemulihan bagi penderita. Untuk itu, pemanfaatan teknologi informasi diperlukan sebagai alat bantu dalam memprediksi dan mengenali gejala-gejala awal stroke[4]. Dengan melakukan perbandingan algoritma dapat diketahui bahwa setiap algoritma memiliki karakteristik berbeda

dalam memproses data dan menghasilkan klasifikasi[5]. Dengan membandingkan *K-Nearest Neighbor* (KNN) dan *Artificial Neural Network* (ANN), penelitian ini bertujuan untuk mengetahui algoritma mana yang lebih efektif dalam mendeteksi penyakit stroke. Selain itu, tujuan penelitian ini adalah untuk membantu menyediakan metode klasifikasi yang cepat, akurat, dan efisien dalam mendukung proses diagnosis dini penyakit stroke.

Metode *K-Nearest Neighbor* (KNN) dan *Artificial Neural Network* (ANN) memiliki kelebihan dan kelemahan masing-masing dalam penerapannya pada klasifikasi data medis. KNN dikenal sebagai algoritma yang sederhana dan mudah diimplementasikan. Algoritma ini tidak memerlukan asumsi tentang distribusi data, sehingga cocok untuk dataset yang memiliki karakteristik non-linear. Namun, KNN memiliki kelemahan pada efisiensi waktu komputasi ketika dataset yang digunakan berukuran besar, karena algoritma ini melakukan kalkulasi jarak untuk setiap data baru terhadap seluruh data pelatihan. Selain itu, performa KNN sangat dipengaruhi oleh pemilihan nilai parameter k serta metode pengukuran jarak yang digunakan[6]. Di sisi lain, ANN memiliki kemampuan unggul dalam menangani data dengan pola yang kompleks. ANN mampu mempelajari hubungan non-linear dan fitur-fitur kompleks melalui proses pelatihan yang melibatkan banyak lapisan. Namun, kelemahan ANN terletak pada kebutuhan komputasi yang tinggi serta memerlukan data dalam jumlah besar agar dapat menghasilkan model yang andal[7]. Metode KNN dan ANN dipilih dalam penelitian ini karena keduanya memiliki relevansi kuat dalam konteks klasifikasi data medis. KNN dipilih karena kemudahannya dalam implementasi dan interpretasi hasil, terutama pada dataset berukuran kecil

hingga menengah. Sementara itu, ANN dipilih karena kemampuannya untuk menangani dataset dengan pola kompleks serta menghasilkan akurasi yang tinggi dalam deteksi penyakit. Namun, masing-masing metode memiliki tantangan yang harus diatasi. Pada KNN, tantangannya terletak pada efisiensi komputasi untuk dataset besar dan sensitivitas terhadap outlier. Sedangkan pada ANN, tantangan utama adalah kompleksitas model yang memerlukan waktu pelatihan yang lebih lama serta risiko overfitting pada dataset kecil.

Untuk mengatasi masalah tersebut, penelitian ini akan menerapkan beberapa solusi. Pada metode KNN, akan dilakukan optimasi nilai parameter k dan penggunaan teknik pre-processing data seperti normalisasi untuk meningkatkan efisiensi dan akurasi. Pada ANN, akan diterapkan teknik regularisasi seperti dropout dan penggunaan batch normalization untuk mengurangi risiko overfitting. Selain itu, proses tuning hyperparameter akan dilakukan untuk mendapatkan konfigurasi model ANN yang optimal. Dengan pendekatan ini, diharapkan tantangan yang ada pada kedua metode dapat diminimalkan sehingga performa keduanya dapat dibandingkan secara adil.

Berdasarkan latar belakang tersebut maka peneliti tertarik untuk melakukan penelitian dalam bentuk skripsi yang berjudul **"PERBANDINGAN KLASIFIKASI PENYAKIT STROKE MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR (KNN) DAN ARTIFICIAL NEURAL NETWORK (ANN)."**

1.2. RUMUSAN MASALAH

Berdasarkan latar belakang masalah yang telah dikemukakan di atas, maka penulis merumuskan masalah dalam penelitian ini, yaitu:

1. Bagaimana penerapan Algoritma K-Nearest Neighbor dan Artificial Neural Network dalam klasifikasi penyakit stroke?
2. Seberapa akurat Algoritma K-Nearest Neighbor dan Artificial Neural Network dalam mengklasifikasikan penyakit stroke?

1.3. BATASAN MASALAH

Agar dalam penelitian ini dapat berjalan dengan baik dan terarah, penulis menetapkan ruang lingkup penelitian meliputi:

1. Metode yang digunakan: Penelitian ini hanya menggunakan algoritma K-Nearest Neighbor dan Artificial Neural Network sebagai metode klasifikasi.
2. Sumber data: Dataset yang digunakan adalah penyakit jantung dari Kaggle: (<https://www.kaggle.com/dharshanadhanendran/stroke-disease-prediction>)
3. Fokus analisis: Penelitian ini hanya berfokus pada mengklasifikasi penyakit Stroke.

1.4. TUJUAN PENELITIAN

Berdasarkan pada beberapa permasalahan yang telah disampaikan sebelumnya, penelitian ini memiliki tujuan-tujuan sebagai berikut:

1. Untuk mengklasifikasikan penyakit stroke menggunakan algoritma K-Nearest Neighbor dan Artificial Neural Network.
2. Untuk mengukur tingkat akurasi dalam klasifikasi penyakit stroke menggunakan algoritma K-Nearest Neighbor dan Artificial Neural Network.

1.5. MANFAAT PENELITIAN

Adapun manfaat yang diperoleh dari penelitian ini adalah:

1. Bagi Pengembangan Ilmu Pengetahuan
 - Memberikan kontribusi pada pengembangan ilmu pengetahuan dalam bidang machine learning, khususnya terkait penerapan algoritma KNN dan ANN untuk klasifikasi penyakit stroke.
 - Menjadi referensi akademis untuk penelitian selanjutnya dalam pengembangan sistem klasifikasi penyakit berbasis machine learning.
2. Bagi Institusi dan Tenaga Kesehatan
 - Membantu mempercepat proses screening awal pasien stroke melalui sistem klasifikasi yang akurat.

- Mengoptimalkan penggunaan sumber daya dengan menyediakan sistem klasifikasi yang efisien dan terjangkau.
- Meningkatkan kualitas layanan kesehatan melalui penerapan teknologi machine learning dalam proses diagnosis.

3. Bagi Masyarakat dan Pasien

- Membantu mengurangi risiko kecacatan dan kematian akibat keterlambatan diagnosis stroke.
- Memberikan alternatif diagnosis yang lebih terjangkau dari segi biaya dan waktu.

1.6. SISTEMATIKA PENULISAN

Laporan penelitian ilmiah ini dibuat dalam sistematika yang sesuai dengan kaidah penulisan ilmiah yang benar dan dibagi dalam bab-bab sebagai berikut:

BAB I : PENDAHULUAN

Bab ini merupakan bab pendahuluan yang meliputi latar belakang masalah, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian, serta sistematika penulisan.

BAB II : LANDASAN TEORI

Bab ini memberikan landasan teori yang menjadi dasar untuk membahas laporan tertentu, dan memberikan definisi yang mendasari penelitian

yang dilakukan melalui tinjauan literatur sebagai dasar untuk melakukan analisis.

BAB III : METODOLOGI PENELITIAN

Bab ini merujuk pada kerangka penelitian, metode atau pendekatan yang digunakan, dan alat yang digunakan untuk menjawab pertanyaan penelitian untuk mencapai tujuannya.

BAB IV : ANALISIS DAN HASIL

Pada bab ini, menjelaskan tentang gambaran umum objek penelitian, serta menganalisis data. Hasil analisis disajikan dalam bentuk visualisasi dari tools yang digunakan.

BAB V : IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini berisi implementasi sistem klasifikasi dan hasil pengujian dari kedua algoritma yang digunakan.

BAB VI : PENUTUP

Bab ini merupakan penutup dari penelitian ilmiah ini yang berisi kesimpulan dari pembahasan bab-bab sebelumnya dan saran-saran yang berguna bagi pihak-pihak yang bersangkutan dalam penelitian ilmiah ini.

