

BAB I

PENDAHULUAN

1.1. Latar Belakang

Secara internasional, stroke berkontribusi besar terhadap beban disabilitas. Defisit neurologis yang muncul dapat berupa gangguan penglihatan, gangguan bicara (disartria/afasia), penurunan mobilitas, dan kelumpuhan. Penegakan diagnosis memanfaatkan temuan klinis beserta modalitas pencitraan otak dalam membedakan stroke iskemik dan ICH [1]. Pada penelitian [2] dijelaskan bahwa stroke menyumbang beban disabilitas yang besar di dunia, dengan gejala neurologis seperti penglihatan kabur atau hilang sebagian, bicara tidak jelas (disartria), kelemahan anggota gerak, gangguan keseimbangan, dan kelumpuhan. Menurut penelitian oleh [3] menjelaskan bahwa stroke berbahaya karena terjadi pada otak, pusat pengendali tubuh. Konsekuensinya termasuk defisit motorik (misalnya hemiparesis atau kelumpuhan) pada anggota gerak [4].

Berbagai algoritma *Mechine Learning* maupun *deep learning* telah diterapkan untuk memprediksi penyakit stroke, seperti *Gradient Boosting Classifier* [4], *Decision Tree* [5], dan *Artificial Neural Network* [6]. Meskipun algoritma tersebut menunjukkan hasil akurasi yang beragam, efektivitasnya masih dipengaruhi oleh karakteristik data dan parameter yang digunakan. Salah satu model algoritma *deep learning* yang inovatif dalam menangani data tabular adalah *Neural Oblivious Decision Ensembles* (NODE), yang diperkenalkan oleh Popov et al. [7]. Metode ini memanfaatkan jaringan *neural* yang dirancang khusus untuk data terstruktur, memberikan performa yang kompetitif dibandingkan model tradisional lainnya. Seperti yang dijelaskan pada penelitian oleh [7], menunjukkan bahwa metode NODE mampu mengolah data dengan kompleksitas tinggi dan memberikan akurasi yang signifikan dalam berbagai tugas klasifikasi, termasuk prediksi medis.

Keberhasilan metode NODE juga sangat dipengaruhi oleh pemilihan parameter yang optimal. Oleh karena itu, teknik optimasi *hyperparameter* seperti

Grid Search [8], *Random Search* [9], dan *Bayesian Optimization* [10] memainkan peran penting dalam meningkatkan performa model. Studi sebelumnya menunjukkan bahwa teknik - teknik ini tidak hanya mempercepat proses *tuning* model, tetapi juga dapat meningkatkan akurasi prediksi secara signifikan. Selain optimasi *hyperparameter*, untuk memastikan dan memvalidasi performa metode NODE, analisis validasi menggunakan AUC (*Area Under the Curve*) menjadi sangat penting. Menurut penelitian [11] dijelaskan AUC memberikan gambaran umum performa model pada berbagai ambang keputusan, metrik ini juga memiliki keterbatasan karena mencakup ambang keputusan yang tidak realistis. Namun, AUC tetap menjadi standar dalam seleksi dan evaluasi model klasifikasi dua kelas, memberikan informasi yang mendalam mengenai kapabilitas model dalam memisahkan sampel positif dari negatif [11]. Selain AUC, evaluasi model juga mencakup metrik lain seperti akurasi, *confusion matrix*, dan waktu eksekusi guna memberikan penilaian yang lebih komprehensif. Akurasi menggambarkan bagian prediksi yang sesuai kebenaran dari semua data, dan meskipun sering digunakan, akurasi bisa memberikan gambaran keliru ketika dataset memiliki ketidakseimbangan kelas yang signifikan. *Confusion matrix* memberikan analisis lebih rinci terhadap hasil prediksi model dengan memetakan total prediksi benar dan salah pada setiap kategori. Melalui *confusion matrix*, dapat dihitung metrik turunan seperti *True Positive Rate* dan *False Positive Rate* yang memberikan sensitivitas dan kesalahan klasifikasi model [12]. Di sisi lain, waktu eksekusi juga menjadi indikator penting dalam mengevaluasi efisiensi komputasi suatu model, terutama ketika model tersebut akan digunakan dalam sistem nyata. Dengan menggabungkan seluruh metrik ini, penelitian menegaskan bahwa kinerja model melampaui akurasi statistik dengan tetap efisien dan dapat diandalkan dalam konteks praktis, termasuk aplikasi prediksi medis.

Berdasarkan penelusuran literatur, hingga saat ini belum ditemukan penelitian yang memusatkan perhatian pada penggabungan metode NODE dengan teknik optimasi *hyperparameter* untuk kasus prediksi stroke. Studi ini ditujukan untuk mengatasi gap yang terdapat dalam penelitian sebelumnya dengan mengintegrasikan kedua pendekatan ini. Dengan memanfaatkan keunggulan

NODE dan optimasi *Grid Search* [8], *Random Search* [9], dan *Bayesian Optimization* [10], diharapkan dapat diperoleh model yang lebih akurat dan efisien dalam memprediksi stroke.

Penelitian ini tidak hanya berfokus pada prediksi stroke, tetapi juga berupaya memberikan kontribusi signifikan dalam pengembangan teknologi prediksi berbasis *deep learning* dengan menerapkan metode *Neural Oblivious Decision Ensembles* (NODE), yang dirancang secara khusus untuk mengolah data tabular secara lebih efisien dan akurat. NODE menggabungkan keunggulan model pohon keputusan dan jaringan saraf dalam satu arsitektur yang adaptif dan mampu menyesuaikan terhadap struktur data *non-linier* yang kompleks. Riset ini melanjutkan dan mengembangkan karya dari Setyarini et al. [13], yang menggunakan model Gradient Boosting dengan strategi optimasi *hyperparameter*. Meskipun penelitian tersebut menghasilkan kinerja yang baik, pendekatan ini belum mengeksplorasi potensi dari metode *deep learning* seperti NODE untuk data tabular, serta belum mengintegrasikan pendekatan pencarian *hyperparameter* yang lebih adaptif seperti *Bayesian Optimization*. Oleh karena itu, penelitian ini menghadirkan pendekatan alternatif yang tidak hanya berbeda dari segi arsitektur model, tetapi juga menawarkan perbaikan dari sisi strategi optimasi dan proses pelatihan untuk meningkatkan akurasi serta sensitivitas model terhadap kelas minoritas. Dengan demikian, penelitian ini diharapkan dapat memperluas cakupan penerapan *deep learning* dalam bidang prediksi medis, khususnya dalam mengidentifikasi risiko stroke secara lebih efisien dan adaptif di masa mendatang.

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah penelitian ini yaitu:

- a. Bagaimana mengembangkan model prediksi stroke yang akurat dengan menggunakan metode *Neural Oblivious Decision Ensembles* (NODE)?

- b. Bagaimana pengaruh Teknik optimasi *hyperparameter Grid Search, Random Search, dan Bayesian Optimization* terhadap performa model NODE dalam memprediksi stroke?
- c. Bagaimana validasi metode *Neural Oblivious Decision Ensembles* (NODE) menggunakan metrik AUC, akurasi, *confusion matrix*, dan waktu eksekusi untuk menilai efektivitas prediksi?
- d. Bagaimana perbandingan performa terbaik metode *Neural Oblivious Decision Ensembles* (NODE) dengan hasil optimal dari penelitian sebelumnya yang menggunakan metode *Gradient Boosting*?

1.3. Tujuan Penelitian

Penelitian ini bertujuan untuk:

- a. Mengintegrasikan metode NODE dengan teknik *hyperparameter* untuk meningkatkan akurasi prediksi stroke.
- b. Bagaimana pengaruh teknik optimasi *hyperparameter (Grid Search, Random Search, dan Bayesian Optimization)* dalam meningkatkan performa metode *Neural Oblivious Decision Ensembles* (NODE).
- c. Memvalidasi performa metode *Neural Oblivious Decision Ensembles* (NODE) menggunakan metrik AUC, akurasi, *confusion matrix*, dan waktu eksekusi.
- d. Membandingkan performa terbaik metode *Neural Oblivious Decision Ensembles* (NODE) dengan hasil optimal dari penelitian sebelumnya yang dilakukan oleh Setyarini et al. [13] yang menggunakan Gradient Boosting.

1.4. Batasan Masalah

Untuk menghindari perluasan topik yang terlalu luas, maka peneliiian ini akan dibatasi pada beberapa aspek berikut:

- a. Penelitian ini menggunakan metode *Neural Oblivious Decision Ensembles* (NODE), yang merupakan metode baru sehingga literatur pendukung dan referensi terkait masih terbatas.
- b. Dataset yang digunakan dalam penelitian ini diperoleh dari Kaggle, dan penelitian hanya difokuskan pada dataset ini tanpa melibatkan dataset lain untuk generalisasi hasil.
- c. Evaluasi model dilakukan menggunakan metrik AUC, akurasi, *confusion matrix*, dan waktu eksekusi tanpa melibatkan metrik lain.
- d. Perbandingan dilakukan hanya terhadap hasil terbaik dari penelitian terdahulu yang menggunakan metode *Gradient Boosting*.

