

BAB II

TINJAUAN PUSTAKA

2.1. Studi Literatur

Bagian ini membahas literatur yang relevan untuk mengidentifikasi pendekatan dalam prediksi stroke terutama yang menggunakan metode NODE dan optimasi *hyperparameter*. Literatur *diagnose* berdasarkan algoritma yang digunakan, teknik optimasi, dan penerapan pada ranah medis. stroke timbul ketika perfusi serebral terputus secara mendadak, dipicu oleh sumbatan (iskemik dan juga perdarahan akibat pecahnya pembuluh (hemoragik), sehingga pasokan oksigen dan nutrisi terhambat [14]. Penyediaan model yang dapat mendeteksi dan memprediksi resiko stroke akan sangat membantu untuk pencegahan dan perawatan dini [15].

Arsitektur berbasis metode NODE menjadi fokus utama dalam pengembangan model prediksi penyakit stroke. Metode *Neural Oblivious Decision Ensembles* (NODE) pertama kali dikenalkan oleh [7] sebagai pendekatan inovatif untuk menangani data tabular dengan memanfaatkan jaringan neural. Metode ini memanfaatkan arsitektur jaringan neural yang mampu mengolah data terstruktur dengan efisien, mengatasi keterbatasan model tradisional seperti *Gradient Boosting* pohon keputusan [7]. Dalam konteks penelitian data tabular, NODE telah terbukti kompetitif dalam berbagai tugas klasifikasi, menawarkan kemampuan untuk menangani kompleksitas data dengan akurasi tinggi sekaligus mempertahankan interpretabilitas model.

Optimasi *hyperparameter* adalah langkah yang sangat penting untuk meningkatkan kinerja *Neural Oblivious Decision Ensembles* (NODE). Pendekatan seperti *Grid Search* [8], *Random Search* [9], dan *Bayesian Optimization* [10] digunakan untuk menentukan kombinasi *hyperparameter* terbaik. *Grid Search* bekerja dengan mengeksplorasi setiap kemungkinan kombinasi parameter secara sistematis, seperti yang berhasil diterapkan dalam klasifikasi data *mikro array* untuk meningkatkan akurasi model melalui konfigurasi parameter yang optimal

[16]. Sementara itu, *Random Search* menawarkan fleksibilitas dengan memilih kombinasi parameter secara acak, memungkinkan pencarian yang lebih luas dalam ruang *hyperparameter* dengan efisiensi komputasi yang lebih baik [17]. Di sisi lain, *Bayesian Optimization* menggunakan pendekatan berbasis probabilitas untuk memperkirakan parameter terbaik secara iteratif, sehingga dapat meminimalkan waktu tuning tanpa mengorbankan performa model [18]. Dengan menggunakan ketiga pendekatan ini, penelitian ini bertujuan untuk meningkatkan performa NODE dalam tugas-tugas klasifikasi seperti prediksi stroke pada data tabular.

Penggunaan AUC (*Area Under Curve*) sangat penting untuk mengevaluasi performa metode *Neural Oblivious Decision Ensembles* (NODE) yang telah dioptimasi menggunakan *Grid Search* [8], *Random Search* [9], dan *Bayesian Optimization* [10]. Sebagai metrik penilaian, AUC menyediakan pandangan menyeluruh mengenai performa model dalam mengklasifikasikan data ke kelas positif serta negatif pada berbagai ambang keputusan, sehingga cocok untuk menangani dataset yang tidak seimbang [11]. Dalam proses optimasi *hyperparameter*, AUC mempermudah identifikasi konfigurasi parameter terbaik dengan mempertimbangkan keseimbangan antara sensitivitas dan spesifisitas model. Penelitian sebelumnya juga menunjukkan bahwa penggunaan AUC dapat meningkatkan kualitas pemilihan model dan hasil optimasi *hyperparameter*, terutama pada dataset medis yang kompleks seperti prediksi stroke [11]. Mememertimbangkan pertimbangan kelebihan dan kelemahan tersebut, penelitian ini bertujuan untuk mengembangkan model prediksi stroke berbasis metode NODE dengan dioptimalkan menggunakan *hyperparameter Grid Search* [8], *Random Search* [9], dan *Bayesian Optimization* [10].

Selain AUC, evaluasi model dalam penelitian ini juga mencakup metrik tambahan seperti akurasi, *confusion matrix*, dan waktu eksekusi untuk memberikan penilaian yang lebih komprehensif terhadap kinerja model klasifikasi. Akurasi merupakan metrik yang umum digunakan karena memberikan gambaran langsung mengenai proporsi prediksi yang benar dari keseluruhan data. Namun, akurasi dapat menyesatkan pada kasus data tidak seimbang seperti prediksi stroke, di mana model

cenderung bias terhadap kelas mayoritas. Oleh karena itu, *confusion matrix* digunakan sebagai alat analisis yang lebih mendalam untuk melihat distribusi hasil prediksi terhadap kelas sebenarnya. Berdasarkan studi oleh Tharwat (2021) [12], *confusion matrix* memungkinkan evaluasi spesifik terhadap kesalahan klasifikasi melalui komponen *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN), serta memungkinkan perhitungan metrik turunan seperti *True Positive Rate* (TPR) dan *False Positive Rate* (FPR). TPR atau sensitivitas menggambarkan kapabilitas model mendeteksi kelas positif, sedangkan FPR mengukur kesalahan klasifikasi pada kelas negatif. Sementara itu, evaluasi waktu eksekusi dilakukan untuk menilai efisiensi komputasi dari masing-masing metode optimasi hyperparameter yang diterapkan, yaitu *Grid Search*, *Random Search*, dan *Bayesian Optimization*. Dengan menggabungkan AUC, akurasi, *confusion matrix*, dan waktu eksekusi, penelitian ini tidak hanya menilai efektivitas prediksi, tetapi juga efisiensi dan keseimbangan model dalam menangani data tabular medis yang kompleks.

2.2. Penelitian terdahulu

penelitian telah dilakukan dalam bidang prediksi stroke dengan memanfaatkan algoritma machine *learning* dan teknik optimasi *hyperparameter*. Penelitian oleh Setyarini et al. [19], mengevaluasi performa metode *Gradient Boosting* seperti *LightGBM*, *Categorical Boosting*, dan *Histogram-based Boosting* untuk memprediksi risiko stroke. Penelitian tersebut juga mengimplementasikan strategi optimasi hyperparameter menggunakan *Grid Search* dan *Random Search* guna meningkatkan kinerja model. Berdasarkan hasil uji, *LightGBM* dengan *Grid Search* menunjukkan waktu eksekusi tercepat, sedangkan *Categorical Boosting* memberikan hasil prediksi yang paling optimal dari sisi akurasi. Penelitian ini menjadi rujukan penting dalam memahami pendekatan-pendekatan klasikal yang digunakan pada data tabular medis. Tabel 2.1 menampilkan beberapa penelitian terkait yang mencakup nama penulis, tahun, algoritma yang digunakan, serta hasil evaluasi performa yang diperoleh.

Tabel 1. Penelitian Terdahulu

NO	JUDUL PENELITIAN	PENULIS DAN TAHUN	METODE	HASIL PEMBAHASAN
1	Stroke Prediction with Enhanced Gradient Boosting Classifier and Strategic Hyperparameter	Dela Ananda, Agnes Ayu Maharani, Dyah Gayatri, Christian Sri Kusuma Aditya, dan Didih Rizki Chandranegara	Categorical Boosting (CatBoost), Light Gradient Boosting Machine (LightGBM), Histogram Gradient Boosting (HGB), dan Extreme Gradient Boosting (XGBoost)	<p>Penelitian mengimplementasikan empat metode boosting, yaitu CatBoost, LightGBM, HGB, dan XGBoost dengan strategi tuning hyperparameter menggunakan Grid Search dan Random Search. Dari hasil pengujian, CatBoost memperoleh performa terbaik dengan AUC sebesar 88% dan akurasi 89%. LightGBM menyusul dengan AUC sebesar 87% dan akurasi 88%. Sementara HGB dan XGBoost masing-masing memperoleh AUC sebesar 86% dan 85%.</p>

2.3. Metode Neural Oblivious Decision Ensembles (NODE)

Metode *Neural Oblivious Decision Ensembles* (NODE) pertama kali diperkenalkan oleh Sergei Popov et al. pada tahun 2019 sebagai pendekatan baru untuk menangani data tabular [7]. Arsitektur NODE menggabungkan kekuatan pohon keputusan *oblivious* dengan optimasi berbasis *Gradient Descent* [7]. Metode ini terdiri dari lapisan-lapisan pohon keputusan diferensial yang disusun seperti jaringan *neural*, memungkinkan pelatihan secara *end-to-end* [7]. Setiap pohon pada NODE menggunakan *entmax transformation* untuk memilih fitur secara diferensial, sehingga dapat menangani data dengan kompleksitas tinggi tanpa kehilangan efisiensi komputasi [7].

Keunggulan utama NODE terletak pada fleksibilitasnya dalam memproses data tabular dan kemampuannya untuk mengungguli model *Gradient Boosting* tradisional seperti *XGBoost*. Penelitian Popov et al [7]. menunjukkan bahwa metode NODE mampu memberikan hasil akurasi yang lebih tinggi dalam tugas klasifikasi pada berbagai dataset dibandingkan model berbasis pohon lainnya. Selain itu, NODE mendukung arsitektur multilayer yang memungkinkan pemodelan aturan keputusan yang lebih kompleks. Dalam penelitian terbaru, NODE juga dapat diadaptasi menjadi versi interpretable seperti metode NODE-GAM, yang memungkinkan visualisasi dampak fitur individu terhadap prediksi [20].

Metode *Neural Oblivious Decision Ensembles* (NODE) dikenal karena kemampuannya yang unggul dalam menangani tugas klasifikasi, terutama pada data tabular dengan struktur kompleks. Metode NODE mengadopsi pendekatan yang memungkinkan pelatihan secara menyeluruh melalui lapisan-lapisan pohon keputusan diferensial, yang menggunakan *entmax transformation* untuk pemilihan fitur secara fleksibel [7]. Menurut penelitian penelitian [7] dan [20] menunjukkan desain ini tidak hanya meningkatkan efisiensi komputasi, tetapi juga menghasilkan akurasi yang lebih baik dibandingkan model tradisional seperti *XGBoost* dan *CatBoost*.

2.4. Teknik Optimasi Hyperparameter

Optimasi *hyperparameter* adalah langkah yang sangat penting dalam meningkatkan performa metode pembelajaran mesin, termasuk Neural Oblivious Decision Ensembles (NODE). Beberapa metode optimasi yang sering digunakan meliputi *Grid Search* [8], *Random Search* [9], dan *Bayesian Optimization* [10]. Menurut studi oleh [21] *Grid Search* mengeksplorasi seluruh kombinasi parameter dalam ruang pencarian secara sistematis, sementara menurut [22] *Random Search* lebih efisien karena memilih kombinasi *hyperparameter* secara acak, sehingga menghemat waktu komputasi. Di sisi lain penelitian hasil penelitian [23] *Bayesian Optimization*, memanfaatkan pendekatan probabilistik untuk memperkirakan parameter terbaik berdasarkan evaluasi sebelumnya, menjadikannya lebih canggih dan efisien.

Masing-masing metode optimasi memiliki kelebihan tersendiri. *Grid Search* memastikan evaluasi komprehensif karena menguji seluruh kombinasi parameter yang mungkin dalam ruang pencarian, memberikan jaminan hasil optimal pada ruang pencarian kecil [10]. *Random Search*, meskipun berbasis acak, lebih cepat dan fleksibel karena mampu menjelajahi ruang pencarian dengan jangkauan lebih besar dalam durasi yang lebih pendek [22]. *Bayesian Optimization*, yang memanfaatkan model probabilistik, unggul dalam memprediksi kombinasi parameter yang ideal dengan jumlah evaluasi lebih sedikit, menjadikannya metode yang sangat efisien untuk model yang kompleks [23].

Pada metode NODE, optimasi *hyperparameter* menjadi elemen krusial untuk meningkatkan kinerja model. *Grid Search* diterapkan untuk mengevaluasi berbagai kombinasi parameter seperti *learning rate* dan jumlah unit dalam setiap lapisan jaringan [10]. *Random Search* menawarkan fleksibilitas lebih tinggi dalam menjelajahi konfigurasi *hyperparameter* kompleks dengan efisiensi waktu yang lebih baik [22]. Sementara itu, *Bayesian Optimization* memanfaatkan pendekatan berbasis probabilitas untuk menemukan kombinasi parameter optimal secara lebih cepat dan akurat [23], menjadikannya pilihan yang sangat efektif untuk model kompleks seperti metode NODE.

2.5. Metrik Evaluasi

AUC (*Area Under Curve*) adalah metrik evaluasi yang diimplementasikan untuk menilai kemampuan klasifikasi biner model dengan menguji kemampuan metode dalam mengidentifikasi kelas positif dibanding negatif di berbagai ambang pengujian. AUC dihitung dari curva yang merupakan grafik korelasi antar *True Positive Rate* (TPR) dan *False Positive Rate* (FPR) pada rentang nilai *threshold* [24]. Dalam skala AUC 0–1, nilai 1 menunjukkan kapabilitas klasifikasi maksimal, sedangkan 0,5 menandakan performa setara peluang [11]. Selain AUC, penelitian ini juga mengadopsi metrik evaluasi lain seperti akurasi, yang digunakan untuk mengetahui seberapa tepat metode dalam memprediksi seluruh kelas *Matrix*, dipakai guna mengukur jumlah klasifikasi tepat dan salah di setiap kelas serta waktu eksekusi, mengevaluasi efisiensi waktu pemrosesan metode dalam menjalankan prediksi. Pendekatan evaluasi ini sejalan dengan metode yang digunakan oleh Setyarini et al.[19], yang juga menggunakan keempat metrik tersebut guna menilai performa dan efisiensi algoritma klasifikasi secara menyeluruh.

AUC bersifat *threshold-invariant*, sehingga menawarkan evaluasi kinerja yang komprehensif. Selain itu, metrik ini cocok untuk berbagai jenis distribusi data, termasuk dataset dengan keseimbangan kelas yang baik [23]. Namun, AUC memiliki kelemahan ketika diterapkan pada dataset yang sangat tidak seimbang, karena metrik ini dapat memberikan hasil yang terlalu optimistis pada kelas minoritas. Dalam kasus seperti ini, *precision-recall curve* sering kali memberikan wawasan yang lebih relevan [11], [24].

Kelemahan AUC dalam menangani dataset tidak seimbang dapat diatasi dengan pendekatan seperti *weighted AUC*, yang memperhitungkan distribusi kelas [25]. Dalam konteks metode NODE, AUC digunakan untuk mengevaluasi hasil optimasi *hyperparameter*, seperti *Grid Search* [8], *Random Search* [9], dan *Bayesian Optimization* [10]. Dengan mengintegrasikan AUC, analisis metode NODE dapat memastikan performa optimal pada seluruh *threshold* dan memberikan penilaian yang lebih akurat terhadap kemampuan klasifikasinya [11].

Teknik akurasi mengukur tingkat ketepatan keseluruhan, dihitung sebagai proporsi prediksi benar dibanding total data. Indikator ini memberi gambaran umum seberapa sering model mengklasifikasikan dengan tepat, khususnya ketika kelas relatif seimbang. Namun, karena dataset dalam penelitian ini memiliki ketidakseimbangan kelas, akurasi dapat menjadi kurang representatif jika digunakan secara tunggal. Oleh karena itu, *Confusion Matrix* dilibatkan untuk memberikan informasi lebih rinci tentang jumlah klasifikasi yang benar (*True Positive* dan *True Negative*) serta prediksi salah (*False Positive* dan *False Negative*) [12]. Melalui analisis confusion matrix, peneliti bisa mengetahui jenis kesalahan klasifikasi yang paling sering terjadi dan bagaimana model menangani kelas minoritas. Evaluasi ini juga penting untuk mengidentifikasi trade-off antara recall dan precision secara lebih konkret.

Selanjutnya, aspek efisiensi pemrosesan model juga diperhatikan melalui pengukuran waktu eksekusi. Waktu eksekusi mencerminkan seberapa cepat model dapat melakukan pelatihan dan inferensi terhadap data yang diberikan, yang sangat penting dalam aplikasi nyata di bidang medis di mana keputusan perlu dibuat secara cepat. Pengukuran ini dilakukan untuk membandingkan performa temporal antar model dan strategi *hyperparameter* yang diterapkan. Dengan mempertimbangkan waktu eksekusi, penelitian ini tidak hanya menilai akurasi prediktif tetapi juga efisiensi komputasi. Pendekatan evaluasi menyeluruh ini menegaskan bahwa model yang diperoleh tidak hanya unggul dalam prediksi, tetapi dapat layak diterapkan dalam konteks dunia nyata. Melalui kombinasi metrik ini, hasil penelitian menjadi lebih seimbang dan dapat diinterpretasikan dari berbagai perspektif performa. Evaluasi multi-metrik seperti ini juga konsisten dengan praktik dalam penelitian terdahulu yang dijadikan rujukan, sehingga memungkinkan perbandingan yang adil dan terukur.