



Analisis Perbandingan Model *Fully Connected Neural Networks* (FCNN) dan TabNet Untuk Klasifikasi Perawatan Pasien Pada Data Tabular

Edi Ismanto^{*1}, Abdul Fadlil², Anton Yudhana³

Email: ¹edi.ismanto@umri.ac.id, ²fadlil@mti.uad.ac.id, ³eyudhana@ee.uad.ac.id

¹Pendidikan Informatika, Universitas Muhammadiyah Riau

^{1,2,3}Magister Teknik Elektro, Fakultas Teknologi Industri, Universitas Ahmad Dahlan

Diterima: 30 November 2024 | Direvisi: - | Disetujui: 09 Desember 2024

©2020 Program Studi Teknik Informatika Fakultas Ilmu Komputer,
Universitas Muhammadiyah Riau, Indonesia

Abstrak

Electronic Health Records (EHR) menyimpan data tabular yang kaya informasi dan berperan penting dalam mendukung pengambilan keputusan di bidang kesehatan, terutama untuk klasifikasi perawatan pasien. Penelitian ini membandingkan kinerja dua model kecerdasan buatan, *Fully Connected Neural Networks* (FCNN) dan TabNet, dalam menangani data tabular untuk klasifikasi perawatan pasien. Hasil penelitian menunjukkan bahwa kedua model memiliki kinerja yang baik, namun TabNet sedikit lebih unggul. TabNet mencapai akurasi 0,74, lebih tinggi dibandingkan FCNN dengan 0,73. Selain itu, TabNet juga menunjukkan keunggulan pada presisi (0,74 vs. 0,72), *recall* (0,72 vs. 0,71), dan *F1-Score* (0,73 vs. 0,71). Keunggulan ini menegaskan bahwa TabNet lebih andal dalam menghindari false positive dan mendeteksi kasus positif dengan keseimbangan yang lebih baik antara presisi dan recall. Dengan arsitekturnya yang dirancang khusus untuk data tabular dan kemampuan interpretasi langsung, TabNet menawarkan efisiensi dan kemudahan implementasi yang lebih baik dibandingkan FCNN, yang membutuhkan *preprocessing* data lebih kompleks. Untuk penelitian selanjutnya, disarankan untuk menggunakan dataset yang lebih besar dan beragam, mengeksplorasi data dengan fitur yang lebih kompleks, serta melakukan *tuning hyperparameter* yang lebih mendalam guna meningkatkan kinerja kedua model.

Kata kunci: *Electronic Health Records* (EHR), *Data Tabular*, *Fully Connected Neural Networks* (FCNN), *TabNet*, *Klasifikasi Perawatan Pasien*.

Comparative Analysis of *Fully Connected Neural Networks* (FCNN) and TabNet Models for Patient Care Classification on Tabular Data

Abstract

Electronic Health Records (EHR) store tabular data that is rich in information and play a critical role in supporting decision-making within the healthcare field, particularly for patient care classification. This study evaluates the performance of two artificial intelligence models, *Fully Connected Neural Networks* (FCNN) and TabNet, in processing tabular data for patient care classification tasks. The findings reveal that both models demonstrate strong performance, with TabNet showing a slight advantage. TabNet achieves an accuracy of 0.74, marginally surpassing FCNN's 0.73. Furthermore, TabNet excels in precision (0.74 vs. 0.72), recall (0.72 vs. 0.71), and F1-Score (0.73 vs. 0.71), highlighting its greater reliability in minimizing false positives and accurately detecting positive cases with a better balance between precision and recall. With its architecture specifically tailored for tabular data and its capacity for direct interpretability, TabNet offers enhanced efficiency and ease of implementation compared to FCNN, which demands more complex data preprocessing. For future research, it is suggested to employ larger and more diverse datasets, explore data with higher feature complexity, and conduct comprehensive hyperparameter tuning to further improve the performance of both models.

Keywords: *Electronic Health Records* (EHR), *Tabular Data*, *Fully Connected Neural Networks* (FCNN), *TabNet*, *Patient Care Classification*

1. PENDAHULUAN

Seiring dengan kemajuan pesat dalam teknologi informasi dan kecerdasan buatan (AI), sektor kesehatan mulai mengadopsi inovasi-inovasi berbasis data untuk meningkatkan kualitas pelayanan dan keputusan medis. Salah satu inovasi utama yang memiliki potensi luar biasa adalah penerapan data rekam medis elektronik (EHR) dalam mendukung proses diagnosis dan perawatan pasien [1]. EHR menyimpan berbagai informasi penting, seperti riwayat medis, diagnosis, prosedur yang dilakukan, serta hasil pemeriksaan laboratorium, yang dapat memberikan gambaran komprehensif mengenai kondisi kesehatan pasien [2]. Meskipun memiliki potensi besar, data EHR umumnya bersifat kompleks, tidak terstruktur, dan memerlukan pendekatan canggih untuk pengolahan dan analisis yang tepat.

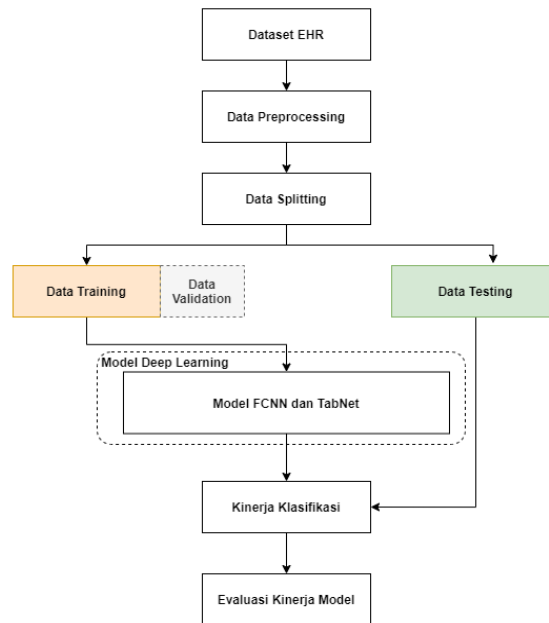
Data tabular, yang terdiri dari data yang terstruktur dalam format tabel dengan kolom dan baris, merupakan salah satu jenis data yang paling umum digunakan dalam berbagai aplikasi, termasuk analisis kesehatan. Dalam konteks perawatan pasien, data tabular sering digunakan untuk menyimpan informasi tentang kondisi pasien, riwayat medis, hasil pemeriksaan, dan rekomendasi perawatan [3]. Dengan meningkatnya volume data yang tersedia, diperlukan metode yang efektif untuk menganalisis dan memanfaatkan data ini guna meningkatkan kualitas pengambilan keputusan klinis. Tradisionalnya, algoritma *machine learning* seperti *Random Forest* atau *Gradient Boosted Trees* sering digunakan untuk menganalisis data tabular [4, 5]. Namun, kemajuan dalam *deep learning* telah membuka peluang baru untuk eksplorasi model seperti *Fully Connected Neural Networks* (FCNN) dan TabNet yang dirancang untuk menangani kompleksitas data tabular dengan lebih baik. Kedua pendekatan ini menunjukkan potensi besar dalam menghasilkan prediksi yang akurat dan dapat diandalkan [6, 7], tetapi perbandingan langsung antara keduanya dalam konteks klasifikasi perawatan pasien masih terbatas.

Fully Connected Neural Networks (FCNN) adalah model *deep learning* yang serbaguna dan telah digunakan secara luas dalam berbagai aplikasi, termasuk analisis data tabular. Meskipun model ini lebih umum digunakan untuk data seperti gambar atau teks, FCNN dapat diadaptasi untuk bekerja dengan data tabular [8]. Sementara itu, TabNet adalah model *deep learning* yang dirancang khusus untuk data tabular dengan fitur seleksi adaptif dan interpretabilitas bawaan. Studi sebelumnya menunjukkan bahwa TabNet sering kali unggul dalam tugas klasifikasi berbasis data tabular, termasuk dalam domain kesehatan [9]. Namun, hanya sedikit penelitian yang secara langsung membandingkan kinerja TabNet dengan FCNN dalam analisis data EHR untuk klasifikasi perawatan pasien.

Analisis data EHR yang akurat dapat memberikan dampak signifikan terhadap kualitas perawatan pasien. Namun, model yang dipilih untuk tugas ini harus mampu menangani kompleksitas data tabular, memberikan hasil yang akurat, dan tetap mudah diinterpretasikan. FCNN dan TabNet memiliki keunggulan masing-masing, tetapi belum ada kesimpulan yang pasti mengenai model mana yang lebih efektif dalam menganalisis data EHR untuk klasifikasi perawatan pasien. Oleh karena itu, penelitian ini diadakan untuk mengevaluasi dan membandingkan kinerja kedua model tersebut, sehingga dapat memberikan panduan yang lebih jelas bagi para peneliti dan praktisi kesehatan. Pertanyaan dalam penelitian ini yaitu (a). Bagaimana performa *Fully Connected Neural Networks* (FCNN) dalam menganalisis data tabular untuk klasifikasi perawatan pasien?. (b). Apakah TabNet memberikan hasil yang lebih baik dibandingkan FCNN dalam hal akurasi, robustitas, dan interpretabilitas pada tugas yang sama?. Sedangkan tujuan dari penelitian ini adalah (a). Membandingkan performa FCNN dan TabNet dalam klasifikasi perawatan pasien berdasarkan data tabular. (b). Menentukan model yang lebih efektif untuk digunakan dalam konteks perawatan pasien. (c). Menyediakan wawasan tentang model *deep learning* untuk mendukung pengambilan keputusan berbasis data di bidang kesehatan.

2. METODE PENELITIAN

Penelitian ini dirancang untuk mengevaluasi dan membandingkan performa model *Fully Connected Neural Networks* (FCNN) dan TabNet dalam klasifikasi perawatan pasien berbasis data *Electronic Health Records* (EHR). Metodologi penelitian melibatkan beberapa tahapan utama yang di tampilkan pada Gambar 1 berikut:



Gambar 1. Tahapan Penelitian

2.1. Dataset EHR

Data yang digunakan dalam studi ini berasal dari salah satu Rumah Sakit Swasta di Indonesia, yang dapat diakses melalui tautan berikut [10]. Dataset berisi 4.412 data hasil pemeriksaan kondisi kesehatan pasien, dengan jumlah fitur sebanyak 12. Tabel 1 menampilkan informasi fitur dari dataset.

Tabel 1. Fitur Dataset

No	Fitur	Tipe Data	Keterangan
1	Hematocrit	Float	Hasil pemeriksaan laboratorium hematokrit pasien
2	Hemoglobin	Float	Hasil pemeriksaan laboratorium hemoglobin pasien
3	Erythrocyte	Float	Hasil pemeriksaan laboratorium eritrosit pasien
4	Leucocyte	Float	Hasil pemeriksaan laboratorium leukosit pasien
5	Thrombocyte	Numeric	Hasil pemeriksaan laboratorium trombosit pasien
6	Mean Corpuscular Hemoglobin (MCH)	Float	Hasil pemeriksaan laboratorium MCH pasien
7	Mean Corpuscular Hemoglobin Concentration (MCHC)	Float	Hasil pemeriksaan laboratorium MCHC pasien
8	Mean Corpuscular Volume (MCV)	Float	Hasil pemeriksaan laboratorium MCV pasien
9	Age	Numeric	Usia pasien
10	Sex	Nominal	Jenis kelamin pasien
11	Severity Level	Nominal	Target label, Berat : Rawat Inap, Ringan : Rawat Jalan

2.2. Data Preprocessing

Data *preprocessing* adalah langkah penting dalam pengolahan dataset *Electronic Health Records* (EHR) untuk memastikan data siap digunakan oleh model *deep learning* [11]. Proses ini dimulai dengan penanganan data yang hilang (*missing values*), akibat kelengkapan yang tidak konsisten pada catatan pasien. Untuk menangani hal ini, dilakukan teknik imputasi, imputasi rata-rata untuk fitur numerik dan imputasi modus untuk fitur kategorikal. Selanjutnya, Teknik *one-hot encoding* data kategorikal dilakukan untuk mengubah data non-numerik, seperti jenis kelamin dan kategori diagnosis, menjadi format yang dapat diproses oleh model. Normalisasi fitur numerik juga diterapkan menggunakan *min-max scaling* untuk memastikan nilai-nilai data berada dalam rentang 0 hingga 1. Hal ini penting untuk menghindari dominasi fitur dengan nilai skala besar terhadap model. Proses berikutnya adalah pembersihan data, untuk mengidentifikasi dan menangani *outlier* yang dapat memengaruhi kinerja model. Agar model dapat divalidasi dengan baik, data dibagi menjadi data pelatihan, data validasi, dan data pengujian dengan stratifikasi berdasarkan label untuk memastikan distribusi kelas yang seimbang di setiap subset. Ini bertujuan agar model dapat diuji secara adil pada data yang tidak terlihat selama pelatihan.

2.3. Model Fully Connected Neural Networks (FCNN)

Fully Connected Neural Networks (FCNN) jaringan saraf tiruan yang terdiri dari beberapa lapisan neuron, di mana setiap neuron dalam satu lapisan terhubung secara langsung ke semua neuron di lapisan berikutnya [8]. Model dirancang untuk menangkap hubungan kompleks antar fitur dalam data melalui transformasi *non-linear*. FCNN terdiri dari tiga komponen utama: lapisan input, lapisan tersembunyi, dan lapisan output. Setiap neuron menerima bobot (w) dan bias (b) sebagai parameter yang disesuaikan selama pelatihan, serta mengaplikasikan fungsi aktivasi (f) untuk menghasilkan keluaran *non-linear*. Proses perhitungan pada setiap neuron dapat dirumuskan sebagai:

$$z = w \cdot x + b \quad (1)$$

$$a = f(z) \quad (2)$$

di mana x adalah masukan, z adalah kombinasi linear dari bobot dan masukan, dan a adalah keluaran setelah fungsi aktivasi diterapkan. Lapisan tersembunyi berfungsi untuk menangkap pola-pola kompleks dalam data, sedangkan lapisan output menghasilkan prediksi berdasarkan tugas spesifik, seperti regresi atau klasifikasi.

2.4. Model TabNet

TabNet model *deep learning* yang dirancang untuk menangani data tabular, yang digunakan dalam tugas klasifikasi dan regresi. TabNet mengintegrasikan mekanisme *attention* yang memungkinkan model untuk secara adaptif memilih fitur paling relevan dalam setiap langkah pengambilan keputusan [12]. Mekanisme ini membantu TabNet dalam mengatasi tantangan data tabular, seperti hubungan antar-fitur yang kompleks dan distribusi data yang tidak seragam, sekaligus menjaga interpretabilitasnya. Arsitektur TabNet terdiri dari tiga komponen utama: *feature transformer*, *sparse attentive mechanism*, dan *decision steps*. Pada setiap langkah keputusan, model menggunakan mekanisme *attention* untuk memilih subset fitur yang akan diteruskan ke komponen berikutnya. Persamaan (3) merupakan formula dari TabNet.

$$M_i = \text{Softmax} \left(\frac{E_i \cdot A_i^T}{\sqrt{d_k}} \right) \quad (3)$$

di mana M_i adalah peta perhatian pada langkah ke- i , E_i adalah representasi fitur yang dihasilkan oleh *feature transformer*, A_i adalah vektor perhatian yang dipelajari, dan d_k adalah ukuran dimensi untuk normalisasi.

2.5. Matriks Evaluasi Kinerja

Model dievaluasi menggunakan *confusion matrix*. Berdasarkan perbandingan antara prediksi model dan label yang sebenarnya, matriks ini menunjukkan seberapa baik model mengklasifikasikan data dan seberapa banyak kesalahan yang dibuat [13]. Persamaan (4), (5), (6), dan (7) metrik evaluasi yang digunakan untuk mengevaluasi kinerja model [14]:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

$$\text{Precision} = \frac{TP}{FP + TP} \quad (6)$$

$$F1 \text{ Score} = 2 \times \frac{\sum \text{Recall} \times \sum \text{Precision}}{\sum \text{Recall} + \sum \text{Precision}} \quad (7)$$

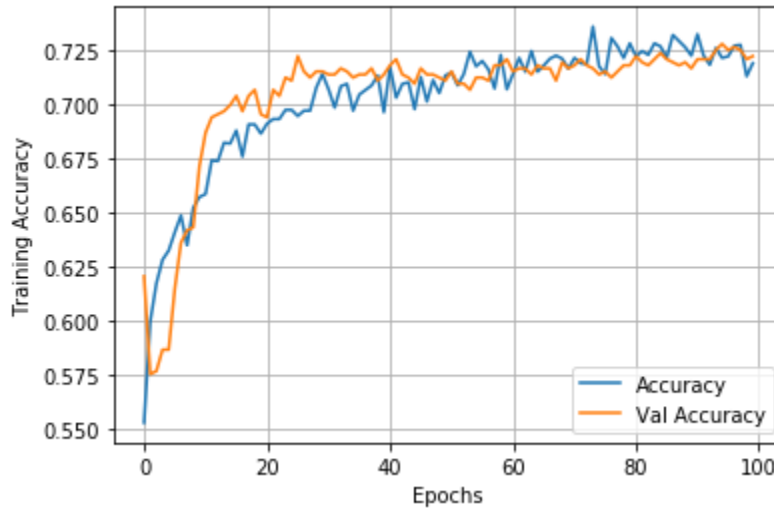
Dalam evaluasi performa model klasifikasi, terdapat empat kategori hasil prediksi: *True Positive* (TP) adalah jumlah kasus ketika model berhasil memprediksi kelas positif dengan benar, sesuai dengan kelas sebenarnya yang juga positif. *False Positive* (FP) terjadi ketika model memprediksi kelas positif, padahal kelas sebenarnya adalah negatif, yang sering disebut *type I error*. Sebaliknya, *True Negative* (TN) adalah jumlah kasus di mana model secara akurat memprediksi kelas negatif sesuai dengan kenyataan bahwa kelas tersebut memang negatif. *False Negative* (FN) terjadi ketika model memprediksi kelas negatif, tetapi kelas sebenarnya adalah positif, yang dikenal sebagai *type II error*. Keempat kategori ini digunakan untuk menghitung berbagai metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score*.

3. HASIL DAN PEMBAHASAN

Eksperimen ini diimplementasikan menggunakan bahasa pemrograman Python, dengan pustaka yang digunakan *TensorFlow* untuk FCNN dan *PyTorch* TabNet untuk model TabNet. Data EHR yang digunakan dalam penelitian ini dibagi menjadi dua bagian: 80% data pelatihan dan 20% data pengujian. Pembagian ini bertujuan untuk melatih model pada sebagian besar data dan mengujinya pada subset yang terpisah untuk mengevaluasi kinerja model yang lebih objektif.

3.1. Hasil Model Fully Connected Neural Networks (FCNN)

FCNN dibangun dengan lapisan input yang menerima 8 fitur, diikuti oleh tiga lapisan tersembunyi yang masing-masing menggunakan fungsi aktivasi ReLU dan Dropout 50% untuk mencegah overfitting. Lapisan output menggunakan *sigmoid* untuk klasifikasi biner. Model ini dikompilasi dengan *binary cross-entropy* sebagai fungsi loss dan *Adam optimizer* dengan *learning rate* 0.001. Pelatihan dilakukan dengan data training dan validasi menggunakan *batch size* 32 selama 100 *epoch*. Gambar 2 menampilkan validasi akurasi model FCNN pada data pelatihan.



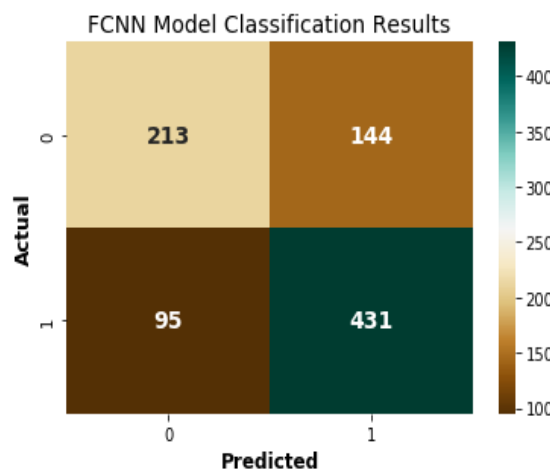
Gambar 2. Validasi Akurasi Model FCNN Pada Data Pelatihan

Model FCNN dievaluasi menggunakan akurasi, presisi, *recall*, dan *F1-Score*. Akurasi mengukur prediksi yang benar, presisi mengukur seberapa tepat prediksi positif, *recall* mengukur kemampuan model dalam mendeteksi contoh positif, dan *F1-Score* menggabungkan presisi dan *recall* untuk memberikan gambaran keseimbangan keduanya. Tabel 2 menampilkan kinerja dari model FCNN.

Tabel 2. Hasil Kinerja Model FCNN

Model	Akurasi	Presisi	Recall	F1-Score
Model FCNN	0.73	0.72	0.71	0.71

Pada Tabel 2 hasil evaluasi model FCNN menunjukkan akurasi 0.73, yang berarti model memprediksi 73% data dengan benar. Presisi 0.72 menunjukkan bahwa 72% prediksi positif model benar-benar positif, sementara *recall* 0.71 berarti model berhasil mendeteksi 71% contoh positif yang sebenarnya. *F1-Score* 0.71 mencerminkan keseimbangan antara presisi dan *recall*, menandakan bahwa model memiliki performa yang cukup baik dalam klasifikasi perawatan pasien. Gambar 3 menampilkan hasil klasifikasi model FCNN.



Gambar 3. Hasil Klasifikasi Model FCNN Pada Data Pengujian

Gambar 3 menunjukkan model FCNN memiliki 213 *true negatives* (TN), 431 *true positives* (TP), 144 *false positives* (FP), dan 95 *false negatives* (FN). Model cukup baik dalam mengklasifikasikan kelas 1 (banyak TP), tetapi masih membuat kesalahan pada beberapa data kelas 0 (FP) dan kelas 1 (FN).

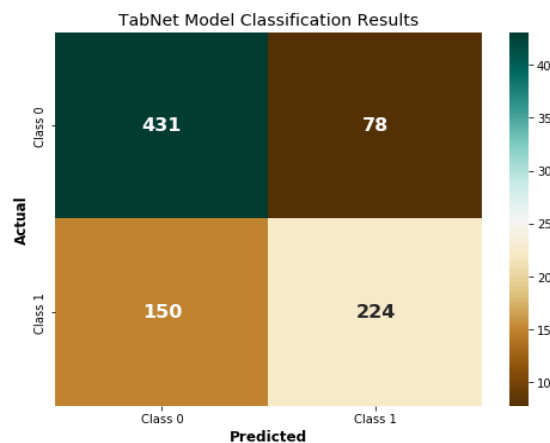
3.2. Hasil Model TabNet

Model TabNet dikembangkan dengan parameter *encoder* dan *decoder* sebesar 8, jumlah langkah model 2, serta penggunaan *gamma* untuk pembelajaran mask dan *sparsemax* sebagai jenis mask untuk *attention*. *Optimizer* yang digunakan adalah *Adam* dengan *learning rate* 0.02. Model dilatih dengan data pelatihan menggunakan *batch size* 1024, *virtual batch size* 128, dan maksimal 100 *epoch*, dengan *patience* 20 untuk menghentikan pelatihan jika tidak ada perbaikan. Setelah pelatihan, model digunakan untuk memprediksi data uji. Tabel 3 menampilkan kinerja dari model TabNet.

Tabel 3. Hasil Kinerja Model TabNet

Model	Akurasi	Presisi	Recall	F1-Score
Model TabNet	0.74	0.74	0.72	0.73

Hasil evaluasi pada Tabel 3 model TabNet menunjukkan kinerja yang cukup baik. Akurasi sebesar 0.74 berarti model memprediksi 74% dari total data dengan benar. Presisi sebesar 0.74 menunjukkan bahwa 74% dari prediksi positif yang dibuat oleh model benar-benar positif, mengindikasikan model menghindari kesalahan *false positive* dengan cukup baik. *Recall* sebesar 0.72 berarti model berhasil mendeteksi 72% dari seluruh contoh positif yang sebenarnya, menunjukkan bahwa masih ada beberapa *false negative*. *F1-Score* sebesar 0.73 mencerminkan keseimbangan yang baik antara presisi dan recall, menunjukkan bahwa model TabNet bekerja dengan performa yang stabil dalam klasifikasi. Gambar 4 menampilkan hasil klasifikasi model TabNet.



Gambar 4. Hasil Klasifikasi Model TabNet Pada Data Pengujian

Gambar 4 menampilkan hasil klasifikasi model TabNet memiliki 431 *true negatives* (TN), 224 *true positives* (TP), 78 *false positives* (FP), dan 150 *false negatives* (FN). Model ini unggul dalam memprediksi kelas 0 (banyak TN) dengan sedikit FP, tetapi kinerjanya lebih lemah dalam memprediksi kelas 1, terlihat dari jumlah FN yang cukup tinggi. Analisis metrik seperti presisi, *recall*, dan *F1-score* diperlukan untuk memahami performa secara lebih detail.

3.3. Perbandingan Hasil Kinerja Antara Model FCNN dan TabNet

Perbandingan kinerja FCNN dan TabNet penting karena keduanya memiliki pendekatan berbeda untuk data tabular. Evaluasi ini membantu menentukan model yang paling efisien dan akurat sesuai kebutuhan dan karakteristik data. Tabel 4 menampilkan perbandingan kinerja model FCNN dan TabNet dalam mengklasifikasikan data tabular perawatan pasien.

Tabel 4. Perbandingan Kinerja Model FCNN dan TabNet

Model	Akurasi	Presisi	Recall	F1-Score
Model FCNN	0.73	0.72	0.71	0.71
Model TabNet	0.74	0.74	0.72	0.73

Perbandingan kinerja antara model FCNN dan TabNet pada Tabel 4 menunjukkan hasil yang relatif dekat, dengan TabNet sedikit lebih unggul dalam beberapa metrik. Akurasi model TabNet (0.74) sedikit lebih tinggi dibandingkan dengan FCNN (0.73), yang menunjukkan bahwa TabNet memiliki sedikit keunggulan dalam memprediksi seluruh data dengan benar. Dalam hal presisi, TabNet juga unggul dengan nilai 0.74, lebih tinggi dari FCNN yang mencapai 0.72, menandakan bahwa TabNet lebih baik dalam menghindari false positive. *Recall* untuk kedua model hampir sama, dengan TabNet sedikit lebih baik (0.72) dibandingkan FCNN (0.71), menunjukkan bahwa TabNet sedikit lebih efektif dalam mendeteksi kasus positif yang sebenarnya. Terakhir, *F1-Score* TabNet (0.73) juga lebih tinggi dari FCNN (0.71), menunjukkan bahwa TabNet memberikan keseimbangan yang lebih baik antara presisi dan recall. Secara keseluruhan, meskipun kedua model menunjukkan performa yang baik, TabNet sedikit lebih unggul dalam hal presisi, *recall*, dan *F1-Score*. FCNN unggul dalam menangkap hubungan non-linear tetapi membutuhkan *preprocessing* data yang signifikan, sementara TabNet dirancang khusus untuk data tabular dengan mekanisme perhatian dan interpretasi langsung.

4. KESIMPULAN

Model *Fully Connected Neural Networks* (FCNN) dan TabNet untuk Data Tabular dalam Klasifikasi Perawatan Pasien menunjukkan bahwa kedua model memiliki kinerja yang baik, tetapi TabNet sedikit lebih unggul. TabNet memiliki akurasi (0.74) yang sedikit lebih tinggi dibandingkan FCNN (0.73), serta keunggulan dalam presisi (0.74 vs. 0.72), *recall* (0.72 vs. 0.71), dan *F1-Score* (0.73 vs. 0.71). Hal ini menunjukkan bahwa TabNet lebih andal dalam menghindari false positive dan mendeteksi kasus positif dengan keseimbangan yang lebih baik antara presisi dan *recall*. Dengan arsitekturnya yang dirancang khusus untuk data tabular dan kemampuan interpretasi langsung, TabNet menawarkan keunggulan dibandingkan FCNN, yang memerlukan *preprocessing* data lebih signifikan. Oleh karena itu, TabNet merupakan pilihan yang lebih baik untuk aplikasi serupa, terutama saat efisiensi dan interpretasi data menjadi prioritas. Untuk penelitian selanjutnya, disarankan menggunakan dataset yang lebih besar dan beragam, menguji model pada data dengan jumlah fitur yang lebih kompleks, serta melakukan tuning *hyperparameter* yang lebih mendalam.

DAFTAR PUSTAKA

- [1] Y. Si, J. Du, Z. Li, X. Jiang, T. Miller, F. Wang, W. J. Zheng dan K. Roberts, "Deep representation learning of patient data from Electronic Health Records (EHR): A systematic review," *Journal of biomedical informatics*, vol. 115, p. 103671, 2021.
- [2] M. Wang, M. S. Pantell, L. M. Gottlieb dan J. Adler-Milstein, "Documentation and review of social determinants of health data in the EHR: measures and associated insights," *Journal of the American Medical Informatics Association*, vol. 28, no. 12, pp. 2608-2616, 2021.
- [3] M. Karatas, L. Eriskin, M. Deveci, D. Pamucar dan H. Garg, "Big Data for Healthcare Industry 4.0: Applications, challenges and future perspectives," *Expert Systems with Applications*, vol. 200, p. 116912, 2022.
- [4] V. Jackins, S. Vimal, M. Kaliappan dan M. Y. Lee, "AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes," *The Journal of Supercomputing*, vol. 77, no. 5, pp. 5198-5219, 2021.
- [5] E. Ismanto, A. Fadlil, A. Yudhana dan K. Kitagawa, "A Comparative Study of Improved Ensemble Learning Algorithms for Patient Severity Condition Classification," *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 6, no. 3, pp. 312-321, 2024.
- [6] D. Jiang, G. Hu, G. Qi dan N. Mazur, "A fully convolutional neural network-based regression approach for effective chemical composition analysis using near-infrared spectroscopy in cloud," *Journal of Artificial Intelligence and Technology*, vol. 1, no. 1, pp. 74-82, 2021.
- [7] L. P. Joseph, E. A. Joseph dan R. Prasad, "Explainable diabetes classification using hybrid Bayesian-optimized TabNet architecture," *Computers in Biology and Medicine*, vol. 151, p. 106178, 2022.
- [8] G. Verma, M. Kumar dan S. Raikwar, "FCNN: fusion-based underwater image enhancement using multilayer convolution neural network," *Journal of Electronic Imaging*, vol. 31, no. 6, pp. 063039-063039, 2022.
- [9] K. McDonnell, F. Murphy, B. Sheehan, L. Masello dan G. Castignani, "Deep learning in insurance: Accuracy and model interpretability using TabNet," *Expert Systems with Applications*, vol. 217, p. 119543, 2023.
- [10] M. Sadikin, "EHR Dataset for Patient Treatment Classification," Mendeley Data, V1, Jakarta, 2020.
- [11] V. Çetin dan O. Yıldız, "A comprehensive review on data preprocessing techniques in data analysis," *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 28, no. 2, pp. 299-312, 2022.
- [12] S. Ö. Arik dan T. Pfister, "TabNet: Attentive Interpretable Tabular Learning," *In Proceedings of the AAAI conference on artificial intelligence*, vol. 35, no. 8, pp. 6679-6687, 2021.
- [13] D. Božić, B. Runje, D. Lisjak dan D. Kolar, "Metrics related to confusion matrix as tools for conformity assessment decisions," *Applied Sciences*, vol. 13, no. 14, p. 8187, 2023.
- [14] M. Heydarian, T. E. Doyle dan R. Samavi, "MLCM: Multi-label confusion matrix," *IEEE Access*, vol. 10, pp. 19083-19095, 2022.