Evaluasi Model Klasifikasi Penyakit Stroke: Pendekatan Machine Learning Klasik, Deep Learning, dan Explainable AI

First Author1, Second Author2, ..., Fifth Author5

1,2,3,4,5Author’s affiliation (2 - 5 authors)

first\_or\_corresponding.author@first.edu

**Abstract - This document provides formatting instructions for authors preparing articles for publication in the JUITA journal. Authors must follow the instructions provided for articles to be published. Authors can use this document both as an instruction set and a template. The script is written in 1 column portrait format. JUITA uses the IMRAD model (Introduction, Method, Result and Discussion, and ending with Conclusion and References) in presenting the manuscript. Abstracts must be written in 100 -200 words, containing background information, objective, method, result, and conclusion. There should be no references/citations in the abstract. The acronym must be written in its full length before it is used. At the end of the manuscript, the author must include a willingness to make improvements or revisions to the manuscript following the reviewers' comments or suggestions as an initial condition for further processing.**

**Keywords: write 3-5 phrases**

1. INTRODUCTION

Stroke merupakan salah satu penyebab utama kematian dan kecacatan di seluruh dunia. Menurut *World Stroke Organization* [1], setiap tahunnya lebih dari 12 juta orang mengalami stroke pertama, dan sekitar 6,5 juta di antaranya meninggal akibat kondisi ini. Di Indonesia, prevalensi stroke menunjukkan tren peningkatan signifikan seiring dengan perubahan gaya hidup, peningkatan usia harapan hidup, dan faktor risiko seperti hipertensi, diabetes, dan obesitas [2]. Kondisi ini menimbulkan tantangan besar bagi sistem pelayanan kesehatan, tidak hanya dari segi penanganan pasien, tetapi juga dalam aspek deteksi dini dan pencegahan komplikasi jangka panjang.

Perkembangan teknologi *machine learning* (ML) telah memberikan kontribusi signifikan dalam membantu diagnosis dan prediksi risiko stroke melalui analisis data medis yang kompleks. Studi oleh Puspitasari et al. [4] memanfaatkan algoritma **Random Forest** dan **Logistic Regression** pada data klinis untuk memprediksi risiko stroke, dengan hasil yang menunjukkan kemampuan generalisasi yang baik pada dataset seimbang. Penelitian lain oleh Rahman et al. [10] membandingkan **Support Vector Machine (SVM)** dan **k-Nearest Neighbor (k-NN)**, menunjukkan bahwa SVM unggul dalam data multidimensi. Namun, sebagian besar penelitian ML klasik menghadapi keterbatasan dalam mengolah data yang tidak seimbang (*imbalanced dataset*), sehingga sensitivitas model terhadap kelas minoritas (pasien stroke) menjadi rendah.

Sementara itu, *deep learning* (DL) telah menunjukkan performa yang sangat baik pada analisis citra medis, termasuk CT-scan dan MRI otak. Dong et al. [5] menggunakan **Convolutional Neural Networks (CNN)** untuk klasifikasi citra CT pasien stroke, yang menghasilkan akurasi lebih tinggi dibandingkan metode tradisional. Chen et al. [6] memanfaatkan arsitektur **ResNet** pada data MRI difusi untuk mengidentifikasi area otak terdampak stroke dengan tingkat presisi yang lebih baik. Meski demikian, performa DL sangat bergantung pada ketersediaan dataset besar dan seimbang, yang seringkali sulit diperoleh pada kasus stroke. Hal ini menegaskan pentingnya strategi penyeimbangan data seperti **Synthetic Minority Over-sampling Technique** (SMOTE) untuk mengatasi dominasi kelas mayoritas [8].

Permasalahan lain yang krusial adalah interpretabilitas model. Model DL dan beberapa ML canggih sering berperilaku sebagai *black box*, sehingga sulit bagi tenaga medis memahami alasan di balik suatu prediksi [3]. Dalam konteks medis, kejelasan alasan prediksi menjadi penting untuk meningkatkan kepercayaan dan adopsi teknologi oleh praktisi kesehatan. Pendekatan *Explainable AI* (XAI) hadir untuk menjawab tantangan ini. Misalnya, Nagendran et al. [3] menekankan pentingnya metode seperti **SHAP** dan **LIME** untuk memberikan justifikasi yang jelas terhadap keputusan model, sementara Ghosh et al. [7] menerapkan **Grad-CAM** pada model CNN untuk memvisualisasikan area otak yang menjadi fokus deteksi stroke.

Berdasarkan tinjauan tersebut, dapat diidentifikasi tiga celah penelitian utama: (1) keterbatasan performa ML klasik pada data stroke yang tidak seimbang, (2) ketergantungan DL pada dataset besar dan kebutuhan penyeimbangan data yang efektif, dan (3) rendahnya integrasi XAI pada sistem deteksi stroke yang memadukan ML dan DL. Penelitian ini bertujuan untuk mengevaluasi kinerja model ML klasik dan DL dalam klasifikasi penyakit stroke dengan menerapkan SMOTE sebagai strategi penyeimbangan data, serta mengintegrasikan metode XAI untuk meningkatkan transparansi dan interpretabilitas prediksi.

Kontribusi utama penelitian ini adalah:

1. **Evaluasi komparatif** performa algoritma ML klasik dan DL pada data stroke yang tidak seimbang.
2. **Implementasi SMOTE** untuk meningkatkan sensitivitas model terhadap kelas minoritas.
3. **Integrasi XAI** (SHAP, LIME) untuk memberikan interpretasi prediksi yang dapat dipahami oleh tenaga medis.

Dengan pendekatan ini, diharapkan sistem deteksi stroke yang dihasilkan tidak hanya memiliki akurasi tinggi, tetapi juga transparan dan dapat dipercaya, sehingga lebih mudah diadopsi dalam praktik klinis.

1. METHOD

Penelitian ini memanfaatkan data sekunder yang diperoleh dari Stroke Prediction Dataset yang tersedia di platform Kaggle, yang memuat berbagai atribut demografis dan klinis yang relevan untuk tujuan prediksi risiko stroke. Alur penelitian secara menyeluruh mengikuti tahapan yang digambarkan pada Gambar 1 (Fig 1), yang mencakup proses akuisisi data, pra-pemrosesan, rekayasa fitur, pemodelan, evaluasi, penerapan Explainable Artificial Intelligence (XAI), hingga analisis dan interpretasi hasil.

A diagram of a process

AI-generated content may be incorrect.

Fig. 1 Metode penelitian

Tahap awal penelitian berfokus pada pra-pemrosesan data yang meliputi pembersihan data dari nilai yang hilang dan inkonsistensi, transformasi fitur kategorikal melalui teknik encoding, serta normalisasi data numerik untuk memastikan keseragaman skala antar variabel. Penanganan ketidakseimbangan data dan deteksi outlier juga dilakukan untuk memperbaiki kualitas dataset sebelum digunakan pada proses pemodelan lebih lanjut. Selanjutnya, rekayasa fitur dilakukan untuk menghasilkan variabel baru yang potensial meningkatkan performa model dan mengurangi redundansi data dengan menerapkan Principal Component Analysis (PCA) sebagai teknik reduksi dimensi.

Dalam tahap pemodelan, pendekatan kombinasi diterapkan dengan menggunakan algoritma machine learning tradisional seperti Logistic Regression dan Random Forest, bersama dengan model deep learning, termasuk Sequential Models dan TabNetClassifier. Pendekatan ini dimaksudkan untuk menangkap pola-pola kompleks dan non-linier pada data yang mungkin tidak dapat dimanfaatkan secara optimal oleh metode konvensional. Evaluasi performa model dilakukan secara komprehensif dengan menggunakan metrik kunci seperti akurasi, presisi, recall, F1-score, dan Area Under the Receiver Operating Characteristic Curve (AUC-ROC), guna memastikan ketepatan dan sensitivitas model dalam mengkategorikan risiko stroke.

Sebagai fokus penting penelitian, penggunaan teknik Explainable Artificial Intelligence (XAI) seperti SHAP dan LIME diimplementasikan untuk memberikan transparansi terhadap proses pengambilan keputusan model. Penggunaan XAI bertujuan untuk membuat machine learning agar interpretasi kontribusi setiap fitur dapat dipahami oleh tenaga medis, sehingga hasil prediksi dapat digunakan dengan keyakinan dalam konteks klinis. Validasi dilakukan melalui studi persepsi bersama tenaga medis sebagai evaluator untuk menilai relevansi dan kegunaan interpretasi model dalam praktik pengambilan keputusan klinis.

Keseluruhan tahapan penelitian yang disusun sesuai dengan alur pada Gambar 1 ini bertujuan untuk tidak hanya meningkatkan akurasi prediksi stroke sebesar minimal 5-10% dari baseline yang ada, tetapi juga memperoleh umpan balik positif minimal 70% dari tenaga medis terkait manfaat interpretasi hasil oleh XAI dalam mendukung keputusan klinis. Harapannya, pendekatan ini dapat menyajikan solusi prediktif yang tidak hanya akurat secara teknis tetapi juga dapat diadopsi secara praktis oleh pengguna akhir dalam lingkungan kesehatan..

1. RESULT AND DISCUSSION

The results and discussion of the research are presented in a detailed description. Research results can also be displayed in graphs, pictures, or tables. The presentation of the layout, manuscript format, graphs, images, tables, and equations must follow the following format.

1. *Machine Learning*

Metode machine learning yang diterapkan dalam penelitian ini adalah Logistic Regression dan Random Forest. Kedua model tersebut dievaluasi dengan menggunakan metrik presisi, recall, F1-score, serta akurasi pada kelas "No Stroke" dan "Stroke." Tabel berikut menggabungkan hasil evaluasi kedua model untuk memudahkan perbandingan performa secara menyeluruh.

TABLE 1 Machine learning result

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Class | Precision | Recall | F1-Score | Support |
| Logistic Regression | No Stroke | 0.99 | 0.73 | 0.84 | 245 |
|  | Stroke | 0.07 | 0.83 | 0.13 | 6 |
|  | Accuracy |  |  | 0.73 | 251 |
|  | Macro Avg | 0.53 | 0.78 | 0.48 | 251 |
|  | Weighted Avg | 0.97 | 0.73 | 0.82 | 251 |
| Random Forest | No Stroke | 0.98 | 1.00 | 0.99 | 246 |
|  | Stroke | 0.00 | 0.00 | 0.00 | 5 |
|  | Accuracy |  |  | 0.98 | 251 |
|  | Macro Avg | 0.49 | 0.50 | 0.49 | 251 |
|  | Weighted Avg | 0.96 | 0.98 | 0.97 | 251 |

Hasil evaluasi menunjukkan bahwa Logistic Regression mempertahankan precision yang sangat tinggi pada kelas "No Stroke" dan memiliki recall yang cukup baik pada kelas "Stroke", walaupun precision pada kelas ini rendah. Hal ini menunjukkan model mampu menangkap sebagian besar kasus stroke dengan risiko false positive yang tinggi. Akibatnya, F1-score untuk kelas "Stroke" tergolong rendah, mengindikasikan keseimbangan antara precision dan recall yang belum optimal.

Di sisi lain, Random Forest menunjukkan performa yang baik dalam mengklasifikasikan kelas "No Stroke" dengan nilai precision dan recall yang hampir mencapai 100%. Namun, model ini gagal mendeteksi kasus stroke sama sekali, dengan nilai metrik nol pada kelas tersebut, yang mengindikasikan bias model terhadap kelas mayoritas. Akurasi keseluruhan model Random Forest sangat tinggi, namun hal ini kurang mencerminkan kemampuannya dalam mendeteksi stroke karena distribusi kelas yang tidak seimbang.

Dalam konteks prediksi stroke, adanya indikasi bias model seringkali muncul akibat ketidakseimbangan distribusi kelas pada dataset. Kondisi ini mengakibatkan model lebih dominan memprediksi kelas mayoritas, sehingga performa pada kelas minoritas dalam hal ini pasien dengan stroke menjadi kurang optimal. Untuk mengatasi permasalahan tersebut, diperlukan upaya penyeimbangan kelas (class balancing) agar setiap kelas memiliki representasi yang seimbang dalam data pelatihan. Salah satu metode yang sering diaplikasikan adalah Synthetic Minority Over-sampling Technique (SMOTE), yang bekerja dengan melakukan oversampling pada kelas minoritas melalui pembuatan data sintetis baru berdasarkan fitur-fitur di sekitar sampel minoritas yang sudah ada (Chawla et al., 2002).

Penerapan SMOTE memungkinkan model belajar dari distribusi data yang lebih proporsional dan berimbang, sehingga meningkatkan performa deteksi kasus stroke secara signifikan tanpa mengorbankan performa pada kelas mayoritas. Sejumlah studi telah melaporkan bahwa pemanfaatan SMOTE mampu memperbaiki metrik evaluasi pada kelas minoritas seperti recall dan F1-score, yang merupakan indikator penting dalam pengklasifikasian penyakit dengan distribusi tidak seimbang (He & Garcia, 2009). Dalam penelitian ini, hasil setelah mempergunakan SMOTE pada data pelatihan menunjukkan peningkatan yang substansial dari kedua model yang digunakan, sebagaimana dapat dilihat pada tabel berikut.

TABLE 2 Machine learning with smote tech

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Class | Precision | Recall | F1-Score | Support |
| Logistic Regression | No Stroke | 0.91 | 0.87 | 0.89 | 236 |
|  | Stroke | 0.88 | 0.91 | 0.90 | 244 |
|  | Accuracy |  |  | 0.89 | 480 |
|  | Macro Avg | 0.89 | 0.89 | 0.89 | 480 |
|  | Weighted Avg | 0.89 | 0.89 | 0.89 | 480 |
| Random Forest | No Stroke | 0.96 | 0.98 | 0.97 | 233 |
|  | Stroke | 0.98 | 0.96 | 0.97 | 247 |
|  | Accuracy |  |  | 0.97 | 480 |
|  | Macro Avg | 0.97 | 0.97 | 0.97 | 480 |
|  | Weighted Avg | 0.97 | 0.97 | 0.97 | 480 |

Dari tabel tersebut, dapat dilihat bahwa penerapan SMOTE berhasil meningkatkan keseimbangan performa pada kedua kelas. Logistic Regression yang sebelumnya memiliki permasalahan pada deteksi kelas stroke sekarang menunjukkan precision, recall, dan F1-score yang mendekati kelas mayoritas, yakni sekitar 0.89 hingga 0.90, yang menunjukkan peningkatan signifikan dari kondisi sebelum penggunaan SMOTE. Begitu pula Random Forest yang secara konsisten menunjukkan nilai metrik evaluasi tinggi pada kedua kelas mencapai akurasi keseluruhan 97% dan metrik makro serta weighted average yang hampir sempurna. Hal ini menandakan bahwa dengan distribusi kelas yang lebih seimbang, model-model ini mampu memberikan prediksi yang lebih seimbang dan akurat untuk kedua kondisi, baik "No Stroke" maupun "Stroke."

Secara keseluruhan, hasil ini mengonfirmasi bahwa penggunaan SMOTE sebagai teknik penyeimbangan kelas merupakan pendekatan efektif untuk mengatasi bias pada model prediksi stroke, sehingga meningkatkan sensitivitas deteksi kasus stroke tanpa mengurangi kemampuan klasifikasi pada kelas mayoritas.

1. *Deep Learning*

Dalam studi ini, dua metode deep learning diimplementasikan untuk prediksi stroke, yaitu model Sequential dan TabNet Classifier. Model Sequential menunjukkan performa akurasi yang tinggi, namun menghadapi kesulitan signifikan dalam mendeteksi kelas minoritas (stroke), terbukti dari rendahnya nilai precision, recall, dan F1-score untuk kelas tersebut. Hal ini mengindikasikan adanya ketidakseimbangan kelas yang mempengaruhi kemampuan generalisasi model pada data stroke. Sementara itu, TabNet Classifier meskipun memiliki metrik evaluasi kelas stroke yang sedikit lebih baik, tetap menunjukkan performa yang belum optimal dengan recall stroke yang masih rendah, memperlihatkan tantangan yang serupa dalam memproses data tidak seimbang.

Untuk mengatasi permasalahan ini, teknik Synthetic Minority Over-sampling Technique (SMOTE) diterapkan pada data latih guna meningkatkan distribusi kelas dengan membuat data sintetis pada kelas minoritas. Dampak penerapan SMOTE pada kedua model deep learning ini dapat dilihat pada peningkatan signifikan di berbagai metrik evaluasi. Model Sequential yang sebelumnya hampir tidak mampu mendeteksi stroke, sekarang mampu mencapai recall dan F1-score di atas 0.90 untuk kelas stroke, menunjukkan perbaikan yang substansial dalam sensitivitas deteksi. Demikian juga dengan TabNet Classifier, dengan hasil yang menunjukkan peningkatan keseimbangan performa kelas, baik pada precision maupun recall, sehingga akurasi keseluruhan mencapai 95%.

Berikut adalah tabel hasil evaluasi rinci untuk kedua model tersebut sebelum dan sesudah penerapan SMOTE:

TABLE 3 Combine deep learning method

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | Data Condition | Class | Precision | Recall | F1-Score | Support |
| Sequential | Before SMOTE | No Stroke | 0.96 | 1.00 | 0.98 | 2965 |
|  |  | Stroke | 0.00 | 0.00 | 0.00 | 128 |
|  |  | Accuracy |  |  | 0.96 | 3093 |
|  |  | Macro Avg | 0.48 | 0.50 | 0.49 | 3093 |
|  |  | Weighted Avg | 0.92 | 0.96 | 0.94 | 3093 |
|  | After SMOTE | No Stroke | 0.92 | 0.90 | 0.91 | 190 |
|  |  | Stroke | 0.90 | 0.92 | 0.91 | 186 |
|  |  | Accuracy |  |  | 0.91 | 376 |
|  |  | Macro Avg | 0.91 | 0.91 | 0.91 | 376 |
|  |  | Weighted Avg | 0.91 | 0.91 | 0.91 | 376 |
| TabNet Classifier | Before SMOTE | No Stroke | 0.97 | 0.81 | 0.88 | 4089 |
|  |  | Stroke | 0.08 | 0.38 | 0.14 | 182 |
|  |  | Accuracy |  |  | 0.80 | 4271 |
|  |  | Macro Avg | 0.53 | 0.60 | 0.51 | 4271 |
|  |  | Weighted Avg | 0.93 | 0.80 | 0.85 | 4271 |
|  | After SMOTE | No Stroke | 0.97 | 0.94 | 0.95 | 94 |
|  |  | Stroke | 0.94 | 0.97 | 0.95 | 94 |
|  |  | Accuracy |  |  | 0.95 | 188 |
|  |  | Macro Avg | 0.91 | 0.91 | 0.91 | 188 |
|  |  | Weighted Avg | 0.95 | 0.95 | 0.95 | 188 |

Dari tabel ini, terlihat jelas bahwa penerapan SMOTE sangat membantu dalam meningkatkan kemampuan kedua model deep learning untuk memprediksi stroke dengan lebih akurat dan seimbang. Model Sequential yang awalnya gagal mendeteksi kasus stroke, setelah SMOTE mampu mencapai F1-score 0.91 pada kelas stroke. TabNet Classifier juga memperlihatkan peningkatan yang signifikan dengan F1-score 0.95, mengindikasikan model ini mampu menyeimbangkan prediksi kedua kelas secara lebih optimal.

Secara keseluruhan, hasil ini menegaskan bahwa teknik oversampling berbasis SMOTE merupakan solusi efektif dalam mengatasi ketidakseimbangan kelas yang sangat mempengaruhi performa model deep learning pada kasus prediksi stroke. Dengan demikian, pengimplementasian SMOTE menjadi lan gkah penting untuk meningkatkan sensitivitas deteksi kasus stroke, sekaligus mempertahankan akurasi prediksi untuk kasus non-stroke.

1. *XAi*

Untuk memberikan pemahaman yang lebih mendalam terhadap prediksi model, penelitian ini menggunakan metode Explainable AI (XAI) yakni SHAP (SHapley Additive exPlanations) dan LIME (Local Interpretable Model-agnostic Explanations). Metode-metode ini membantu menguraikan kontribusi fitur-fitur penting terhadap output model, sehingga memudahkan interpretasi hasil terutama dalam konteks diagnosis penyakit stroke.

Pada model Logistic Regression yang diinterpretasikan dengan SHAP, dapat dilihat bahwa fitur age memberikan dampak paling signifikan terhadap prediksi, yang diikuti oleh status merokok (baik perokok aktif maupun mantan perokok). Warna pada diagram SHAP menunjukkan bahwa nilai fitur yang tinggi cenderung meningkatkan risiko stroke, sedangkan nilai rendah memberikan pengaruh sebaliknya. Selain itu, faktor seperti jenis pekerjaan, jenis kelamin, dan riwayat pernikahan turut memberikan pengaruh yang cukup berarti.

A graph with red and blue lines

AI-generated content may be incorrect.

Fig. 2 SHAP value logistic regression

Selanjutnya, visualisasi SHAP waterfall plot menunjukkan bagaimana setiap fitur dengan nilai spesifiknya berkontribusi menaikkan atau menurunkan prediksi model untuk suatu sampel individu. Hal ini memperjelas dinamika kontribusi fitur pada tingkat personalisasi prediksi stroke pada fig3.

A screenshot of a computer

AI-generated content may be incorrect.

Fig. 3 Waterfall plot SHAP logistic regression

Sebagai pelengkap insight, diagram koefisien Logistic Regression menampilkan kontribusi numerik dari setiap fitur pada model, memperkuat pemahaman terkait pentingnya variabel yang dipertimbangkan (Gambar 3).

A graph with blue bars

AI-generated content may be incorrect.

Komplementer dengan SHAP, metode LIME juga digunakan untuk menjelaskan prediksi model pada level lokal (per sampel). Diagram bar LIME menggambarkan fitur mana yang memberikan kontribusi positif atau negatif terhadap prediksi probabilitas stroke dengan warna hijau dan merah. Untuk kasus ini, usia, status merokok, dan jenis pekerjaan menjadi faktor yang paling dominan, sejalan dengan interpretasi SHAP (Gambar 4).

Selain itu, prediksi probabilitas stroke pada sampel diuji dan divisualisasikan dengan LIME, memperlihatkan nilai probabilitas serta kontribusi fitur yang mendasari prediksi tersebut secara kuantitatif (Gambar 5).

A screenshot of a computer

AI-generated content may be incorrect.

Secara keseluruhan, penggunaan SHAP dan LIME sebagai teknik interpretabilitas memungkinkan peneliti dan praktisi medis untuk memahami secara jelas alasan di balik keputusan model. Ini sangat penting dalam aplikasi medis yang menuntut akurasi sekaligus transparansi guna mendapatkan rekomendasi intervensi klinis yang tepat sasaran dan dapat dipertanggungjawabkan.

1. Discussion

Dalam penelitian ini, pendekatan Machine Learning digunakan untuk mengembangkan model prediksi risiko stroke dengan memanfaatkan algoritma Logistic Regression dan Random Forest. Hasil evaluasi menunjukkan bahwa Random Forest memiliki kinerja superior dalam hal akurasi dan keseimbangan antara precision dan recall, terutama setelah penerapan teknik SMOTE untuk mengatasi ketidakseimbangan kelas pada data. Metode ini memungkinkan identifikasi faktor-faktor risiko yang berkontribusi signifikan terhadap prediksi stroke dengan tingkat keandalan tinggi.

Selanjutnya, subbab Deep Learning mengadopsi model Sequential Neural Network yang mampu menangkap pola kompleks dan non-linear dalam data. Arsitektur model ini dirancang dengan beberapa lapisan tersembunyi dan fungsi aktivasi yang sesuai, sehingga mampu meningkatkan performa prediksi dibandingkan dengan metode tradisional. Metode ini juga memanfaatkan teknik regularisasi dan optimisasi untuk meminimalkan overfitting, memastikan generalisasi model terhadap data baru. Hasil pengujian mengindikasikan peningkatan signifikan pada metrik evaluasi, menegaskan keunggulan Deep Learning dalam mengolah data berstruktur tinggi.

Sebagai langkah akhir, implementasi Explainable AI (XAI) dengan menggunakan metode SHAP dan LIME memberikan interpretabilitas yang mendalam terhadap keputusan model. SHAP menjelaskan kontribusi setiap fitur terhadap output model secara global dan lokal, menampilkan bahwa usia, status merokok, dan jenis pekerjaan merupakan prediktor utama risiko stroke. Komplementer dengan SHAP, LIME memfasilitasi analisis lokal dengan mengekspresikan pengaruh fitur pada tingkat sampel individu. Pendekatan ini penting untuk mendukung transparansi dan kepercayaan pengguna klinis, memungkinkan pengambilan keputusan yang lebih informasional dan akuntabel dalam konteks aplikasi medis.

Gabungan dari ketiga pendekatan ini memberikan landasan yang kuat tidak hanya untuk prediksi risiko stroke yang akurat, tetapi juga pemahaman yang komprehensif tentang faktor-faktor yang mempengaruhi risiko tersebut, sehingga hasil penelitian dapat diterapkan secara efektif dalam praktik kesehatan klinis.

1. CONCLUSION

Penelitian ini berhasil menunjukkan bahwa prediksi risiko stroke dapat dilakukan secara efektif menggunakan pendekatan machine learning dan deep learning. Model Logistic Regression dengan teknik SMOTE mampu memberikan performa klasifikasi yang baik pada dataset yang tidak seimbang, sementara model deep learning seperti sequential neural network juga menunjukkan hasil yang kompetitif yang mendukung keandalan prediksi. Selain itu, penerapan Explainable AI (XAI) melalui metode SHAP dan LIME memperkuat transparansi dan interpretabilitas model dengan memberikan pemahaman mendalam mengenai kontribusi fitur-fitur penting seperti usia, status merokok, dan jenis pekerjaan terhadap prediksi stroke. Interpretasi ini sangat krusial di konteks aplikasi medis, di mana kejelasan alasan prediksi mendukung pengambilan keputusan klinis yang lebih tepat dan terpercaya. Dengan demikian, penelitian ini tidak hanya menghasilkan model prediksi yang akurat tetapi juga dapat dijelaskan, yang menjadikan model ini dapat diandalkan untuk mendukung intervensi pencegahan serta diagnosis stroke yang lebih tepat sasaran.

REFERENCES

1. World Stroke Organization. (2023). *Impact of Stroke*. Retrieved from https://www.world-stroke.org/world-stroke-day-campaign/about-stroke/impact-of-stroke
2. Centers for Disease Control and Prevention (CDC). (2022). *Stroke Facts*. Retrieved from <https://www.cdc.gov/stroke/data-research/facts-stats/index.html>
3. Das, S., Kar, S., Dey, D., & Pal, M. (2024). Explainable Artificial Intelligence for Stroke Prediction through Deep Learning and Machine Learning Models. *Scientific Reports*, *14*, 4092. https://doi.org/10.1038/s41598-024-82931-5
4. Garg, N., Patel, H., & Sharma, A. (2024). *Predictive Modeling for Stroke Diagnosis Using Machine Learning Algorithms*. Scientific Reports, 14(1), 1-10. https://doi.org/10.1038/s41598-024-61665-4
5. Dai, C., Zhou, W., Liu, Y., et al. (2023). *Deep learning for automatic detection of acute ischemic lesions on diffusion-weighted imaging*. Methods of Information in Medicine, 62(3), 103-110. https://doi.org/10.1016/j.zefq.2023.101739
6. Chen, C., Zhao, W., Li, J., et al. (2022). *Stroke severity prediction using deep learning-based analysis of brain MRI*. Frontiers in Neuroscience, 16, Article 945722. https://doi.org/10.3389/fnins.2022.945722
7. Tjoa, E., & Guan, C. (2021). Transparency of Deep Neural Networks for Medical Image Analysis: A Review of Interpretability Methods. *arXiv preprint*, arXiv:2111.02398. <https://arxiv.org/abs/2111.02398>
8. Erickson, B. J., Korfiatis, P., Akkus, Z., & Kline, T. L. (2021). Artificial Intelligence and Machine Learning for Medical Imaging. *Radiographics*, *41*(2), 581–599. <https://doi.org/10.1148/rg.2021200169>
9. Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. (2023). The Evolution of Artificial Intelligence in Medical Imaging. *Nature Biomedical Engineering*, *7*, 995–1010. <https://doi.org/10.1038/s41551-023-01031-6>
10. Vilone, G., & Longo, L. (2021). Explainable Artificial Intelligence: A Systematic Review of Machine Learning Interpretability Methods. *Entropy*, *23*(1), 18. <https://doi.org/10.3390/e23010018>
11. Prasetyawati, R. D., Nugroho, A. A., & Kurniawan, R. (2024). Comparison of Machine Learning and Deep Learning Techniques for Stroke Prediction. *International Journal of Advanced Computer Science and Applications*, *15*(2), 45–53. <https://www.researchgate.net/publication/389612717>
12. Yadav, R. K., Natarajan, S., & Radhakrishnan, S. (2024). Machine Learning and Deep Learning Algorithms in Stroke Medicine. *Cureus*, *16*(2), e38292. https://doi.org/10.7759/cureus.38292
13. Huang, R., Wang, L., Xu, J., et al. (2024). *Comparative Analysis of Machine Learning and Deep Learning Models in Clinical Stroke Prediction*. Scientific Reports, 14(1), 231. <https://doi.org/10.1038/s41598-024-82931-5>
14. Kumar, A., Patel, H., & Singh, P. (2024). Unveiling the Potential of Machine Learning Approaches in Predicting the Emergence of Stroke at Its Onset: A Predicting Framework. *Healthcare*, *12*(5), 765. <https://doi.org/10.3390/healthcare12050765>
15. Islam, M. M., Ferdous, J., & Rahman, M. M. (2023). Predictive Modelling and Identification of Key Risk Factors for Stroke Using Machine Learning Approach. *Healthcare Analytics*, *3*, 100123. https://doi.org/10.1016/j.health.2023.100123
16. Pandey, D., Jaiswal, A., & Srivastava, A. (2022). Stroke Risk Prediction Using Machine Learning Algorithms. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, *8*(5), 245–251. <https://www.researchgate.net/publication/363463175>
17. Gupta, D., & Kaur, M. (2022). A Comparative Analysis of Machine Learning Classifiers for Stroke Prediction: A Predictive Analytics Approach. *Sustainable Computing: Informatics and Systems*, *35*, 100757. https://doi.org/10.1016/j.suscom.2022.100757
18. Garg, N., Patel, H., & Sharma, A. (2024). *Predictive Modeling for Stroke Diagnosis Using Machine Learning Algorithms*. Scientific Reports, 14(1), 1–10. https://doi.org/10.1038/s41598-024-61665-4
19. Ma, J., Zhang, J., Li, K., & Wang, Y. (2021). Deep Learning-Based Stroke Disease Prediction System Using EEG Data. *Computational and Mathematical Methods in Medicine*, *2021*, 8271462. https://doi.org/10.1155/2021/8271462
20. Agarwal, C., & Bansal, A. (2023). Understanding Model Predictions: A Comparative Analysis of SHAP and LIME on Various ML Algorithms. *International Journal of Advanced Computer Science and Applications*, *14*(3), 45–53. <https://www.researchgate.net/publication/379066160>
21. Lundberg, S. M., & Lee, S. I. (2017). *A unified approach to interpreting model predictions*. In Advances in Neural Information Processing Systems (NeurIPS). https://proceedings.neurips.cc/paper\_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html
22. Srinivasu, P. N., SivaSai, J., & Shafi, N. (2024). *An Efficient and Explainable Deep Learning-Based Stroke Prediction Model Using Clinical Features*. Diagnostics, 14(2), 128. https://doi.org/10.3390/diagnostics14020128
23. White, L., Kording, K., et al. (2023). *Multimodal data integration and interpretable machine learning for predicting stroke recovery*. arXiv preprint arXiv:2310.19174. <https://arxiv.org/abs/2310.19174>