

Analisis Prediksi Stroke Menggunakan Pendekatan Decision Tree dengan Seleksi Fitur dan Neural Network

Indah Werdiningsih
*Sistem Informasi, Fakultas Sains
dan Teknologi*
Universitas Airlangga
Surabaya, Indonesia
indah-w@fst.unair.ac.id

Arum Tiyas Handayani
*Sistem Informasi, Fakultas Sains
dan Teknologi*
Universitas Airlangga
Surabaya, Indonesia
arum.tiyas.handayani-2020@fst.unair.ac.id

Kharristantie Sekarlangit.S
*Sistem Informasi, Fakultas Sains
dan Teknologi*
Universitas Airlangga
Surabaya, Indonesia
kharristantie.sekarlangit.suryade-wi-2020@fst.unair.ac.id

Endang Nurjanah
*Sistem Informasi, Fakultas Sains
dan Teknologi*
Universitas Airlangga
Surabaya, Indonesia
endang.nurjanah-2020@fst.unair.ac.id

Fildzah Akhlaqulkarimah
*Sistem Informasi, Fakultas Sains
dan Teknologi*
Universitas Airlangga
Surabaya, Indonesia
fildzah.akhlaqulkarimah-2020@fst.unair.ac.id

Naurah Hedy Pramiyas
*Sistem Informasi, Fakultas Sains
dan Teknologi*
Universitas Airlangga
Surabaya, Indonesia
naurah.hedy.pramiayas-2020@fst.unair.ac.id

Fakhrana Almas Syah Yahrani
*Sistem Informasi, Fakultas Sains
dan Teknologi*
Universitas Airlangga
Surabaya, Indonesia
fakhrana.almas.syah-2020@fst.unair.ac.id

Iin Mardiyana
*Sistem Informasi, Fakultas Sains
dan Teknologi*
Universitas Airlangga
Surabaya, Indonesia
iin.mardiyana-2020@fst.unair.ac.id

Endah Purwanti
*Sistem Informasi, Fakultas Sains
dan Teknologi*
Universitas Airlangga
Surabaya, Indonesia
endahpurwanti@fst.unair.ac.id

Abstract— Currently, stroke is the second cause of death globally. According to data from the World Health Organization (WHO), 7.9% of deaths in Indonesia are caused by stroke. Based on these data, analysis of the factors influencing the case growth rate is very useful. This paper analyzes various factors in electronic health records for effective stroke prediction with different machine-learning algorithms including Decision Tree and Neural Networks. This research uses a dataset consisting of 12 features, namely ID, gender, age, history of hypertension, history of heart disease, marital status, type of work, type of residence, average glucose level, BMI (Body Mass Index) number, and status. smoking, and prediction of stroke. These features were analyzed using the Neural Network and Decision Tree methods so that selected features were produced for further analysis using the Neural Network method. The feature selection results consist of 5 features: age, history of hypertension, marital status, average glucose level, and BMI (Body Mass Index) number. The highest accuracy results were obtained using the Neural Network method with a feature selection of 88.75, the second highest was obtained with the neural network method of 87.1875, and the lowest accuracy was obtained with the Decision Tree method which had an accuracy result of 81.25. Based on these accuracy results, it can be obtained that the most optimal results are shown by the Neural Network method with feature selection.

Keywords—Stroke, Machine Learning, Decision Tree, Neural Network

I. PENDAHULUAN

Stroke merupakan suatu penyakit cerebrovascular yang terjadi karena terdapat gangguan fungsi otak yang berhubungan dengan penyakit pembuluh darah yang menyuplai darah ke otak[1]. Stroke merupakan penyakit penyebab kematian yang tinggi yaitu menempati posisi kedua secara global. Data *World Health Organization* (WHO) menunjukkan bahwa sebanyak 7.9 % kematian di Indonesia disebabkan oleh

penyakit stroke[2]. Bahkan menjadi penyebab kematian pertama di Amerika menurut *Centers for Disease Control and Prevention* (CDC)[3].

Penelitian mengenai penerapan algoritma pembelajaran mesin *decision tree* dan *neural network* banyak digunakan untuk memprediksi mengenai kesehatan, seperti penelitian-penelitian telah dilakukan oleh sebelumnya yaitu mengenai penerapan *decision tree* dan *neural network* pada penentuan kategori status gizi balita yang dilakukan oleh Elivia Budianto et al[4]. Penelitian tersebut bertujuan untuk mengetahui variabel apa yang mempengaruhi kategori gizi balita berdasarkan *root* dan *leaf* yang terbentuk pada C4.5 dan mengetahui akurasi klasifikasi yang dihasilkan metode C4.5 dan NN. Sehingga diperoleh kesimpulan bahwa hasil penelitian tersebut memiliki nilai akurasi sebesar 62,73% untuk kriteria *gain ratio* dan 84,55% untuk kriteria *information gain* pada metode C4.5. Sedangkan metode NN memiliki hasil akurasi sebesar 97,27%.

Penelitian mengenai *decision tree* menggunakan seleksi fitur telah dilakukan oleh sebelumnya yaitu mengenai penerapan *forward selection* sebagai seleksi fitur untuk menentukan kematian akibat gagal jantung menggunakan algoritma C4.5 yang dilakukan oleh Ulin et al[5]. Penelitian tersebut bertujuan untuk memprediksi kematian akibat penyakit gagal jantung menggunakan metode klasifikasi dengan algoritma C4.5 berbasis *forward selection*. *Forward selection* tersebut digunakan sebagai seleksi fitur untuk meningkatkan hasil klasifikasi yang dihasilkan dan guna menentukan fitur paling berpengaruh yang dalam menyebabkan kematian. Sehingga diperoleh kesimpulan bahwa penelitian tersebut algoritma C4.5 berbasis *forward selection* mempunyai nilai akurasi yang lebih baik dibandingkan dengan pengujian algoritma C4.5 saja yaitu sebesar 84,29% atau nilai akurasi meningkat sebesar 6,4% dari hasil akurasi pengujian algoritma C4.5 sebesar 77,89%. Penelitian mengenai penerapan fitur seleksi juga telah dilakukan oleh penelitian terdahulu pada algoritma *neural network* yang dilakukan oleh Nurlalah & Utami [6]. Penelitian tersebut melakukan perbandingan antara model algoritma *neural network* dan model algoritma *neural network* berbasis *particle swarm optimization* sebagai seleksi fitur untuk diagnosis penyakit liver. Berdasarkan hasil penelitian tersebut dapat disimpulkan bahwa algoritma *neural network* berbasis *particle swarm optimization* mempunyai nilai akurasi yang lebih baik dibandingkan dengan pengujian algoritma *neural network* yaitu sebesar 72,37% atau nilai akurasi meningkat sebesar 5,54% dari hasil akurasi pengujian algoritma *neural network* sebesar 66,83%.

Berdasarkan penelitian-penelitian sebelumnya, maka pada penelitian ini akan dilakukan analisis berbagai faktor dalam catatan kesehatan elektronik untuk memprediksi stroke yang efektif pada dataset “*Stroke Prediction Dataset*” dari situs *Kaggle* dilakukan menggunakan algoritma pembelajaran mesin *Decision Tree* menggunakan fitur seleksi dan *Neural Network*. *Decision tree* merupakan sebuah struktur *flowchart* yang berbentuk menyerupai pohon (*tree*), yang mana setiap simpul internal menandakan suatu tes pada atribut, setiap cabangnya merepresentasikan hasil tes dan setiap simpul daun merepresentasikan kelas atau distribusi kelas[7]. *Decision tree* memiliki metode perhitungan yang bernama *retrograde return* yang melakukan perhitungan pada ujung dan atas pohon (*tree*) kemudian kembali ke awal [8]. *Decision tree* menghasilkan model klasifikasi berbentuk pohon yang dapat digunakan sebagai dasar dalam menangani dan memperbaiki[9]. Kelebihan yang dimiliki oleh algoritma *decision tree* yaitu menghasilkan nilai akurasi dan stabilitas yang tinggi[10]. Algoritma pembelajaran mesin *decision tree* telah digunakan pada penelitian yang dilakukan oleh Randi Estian Pambudi et al menghasilkan nilai akurasi sebesar 96.05 % untuk memprediksi penyakit stroke[11].

Penelitian ini menggunakan metode seleksi fitur. Metode seleksi fitur merupakan sebuah proses pemilihan suatu subset dari fitur asli dengan menghilangkan fitur

yang tidak relevan[12]. Seleksi fitur juga dapat disebut sebagai proses pembatasan jumlah variabel input yang akan digunakan[13]. Seleksi fitur juga digunakan untuk mengurangi fitur data yang berdimensi tinggi, juga dapat mempengaruhi tingkat akurasi[14].

Sedangkan algoritma pembelajaran mesin *neural network* merupakan sebuah algoritma yang memiliki arsitektur yang digunakan untuk membentuk model hubungan yang kompleks antara *neuron input* dan *neuron output* yang bertujuan untuk menemukan pola-pola pada data[7]. *Neural network* juga merupakan node komputasi yang saling terikat dan berhubungan sebagai inti dari algoritma *deep learning*[15]. *Neural network* memiliki kelebihan yaitu dapat memberikan hasil pengolahan data yang lebih cepat dengan nilai akurasi yang lebih tinggi [16]. Penafsiran data yang dilakukan oleh *neural network* menggunakan persepsi mesin dan pelabelan input mentah[17]. Algoritma pembelajaran mesin *neural network* telah digunakan dalam penelitian yang dilakukan oleh Simoen Yuda Prasetyo menghasilkan nilai akurasi sebesar 92.032 % untuk memprediksi gagal jantung biner[15].

Penelitian ini bertujuan untuk memprediksi penyakit stroke menggunakan algoritma pembelajaran mesin *decision tree* sebagai seleksi fitur dan *neural network*. Kontribusi dalam penelitian ini adalah menggunakan *decision tree* sebagai seleksi fitur dan *neural network* sebagai klasifikasi. Hasil penelitian ini diharapkan dapat membantu petugas kesehatan khususnya dokter untuk memprediksi penyakit stroke dan dapat mengetahui performa terbaik antara algoritma pembelajaran mesin *decision tree*, *neural network*, dan *decision tree* dan *neural network*.

II. METHOD/MATERIAL

A. Decision Tree

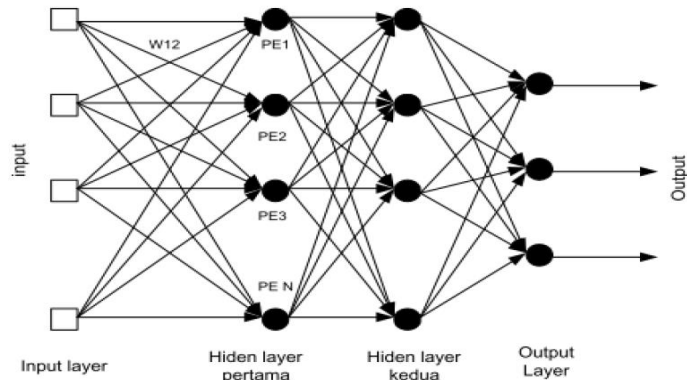
Pohon keputusan merupakan pendekatan “*divide and conquer*” dalam mempelajari masalah dari sekumpulan data independen yang digambarkan dalam bagan pohon[20]. Berikut persamaan data dalam tuple D.

$$Info(D) = \sum_{i=1}^n - p_i \log_2 (p_i) \quad (2.4)$$

p_i merupakan probabilitas tuple dalam D yang menjadi kelas C_i dengan asumsi $|C_i(D)|/|D|$. $Info(D)$ atau disebut juga entropy dari D merupakan rata rata informasi yang diperlukan untuk identifikasi tuple dalam D.

B. Neural Network

Neural network adalah pemrosesan informasi sistem. Secara umum *neural network* dapat dianggap sebagai sistem kotak hitam yang menerima *input* dari lingkungan dan menghasilkan *output*. *Neural network* mengandung elemen pemrosesan dan pembobotan yang saling terhubung. Setiap lapisan dalam jaringan berisi oleh kelompok elemen pemrosesan seperti yang ditunjukkan pada gambar 1 [19].



Gambar 1. Multi Layer Perceptron Neural Networks

Setiap elemen pemrosesan mengumpulkan nilai dari semua *input* yang terhubung ke elemen pemrosesan dan menghasilkan *output* melalui operasi matematika (perkalian operasi). Ada tiga lapisan yang membangun *neural network*, yaitu lapisan *input*, lapisan tersembunyi, dan lapisan keluaran[19].

III. HASIL DAN PEMBAHASAN

Pengolahan data pada penelitian dengan metode klasifikasi prediksi pengidap stroke menggunakan metode *Neural Network* dan *Decision Tree* yang menerapkan seleksi fitur untuk mengidentifikasi akurasi terbaik dari metode *neural network*. Input data yang digunakan diawal adalah seluruh variabel pada data yang terdiri dari jenis kelamin, umur, riwayat hipertensi, riwayat penyakit jantung, status pernikahan, tipe pekerjaan, tipe tempat tinggal, rata-rata level glukosa, angka BMI (*Body Mass Index*), dan status merokok. *Output* yang didapatkan dari pengolahan data ini adalah prediksi penyakit stroke yang mungkin dapat terjadi kepada pasien.

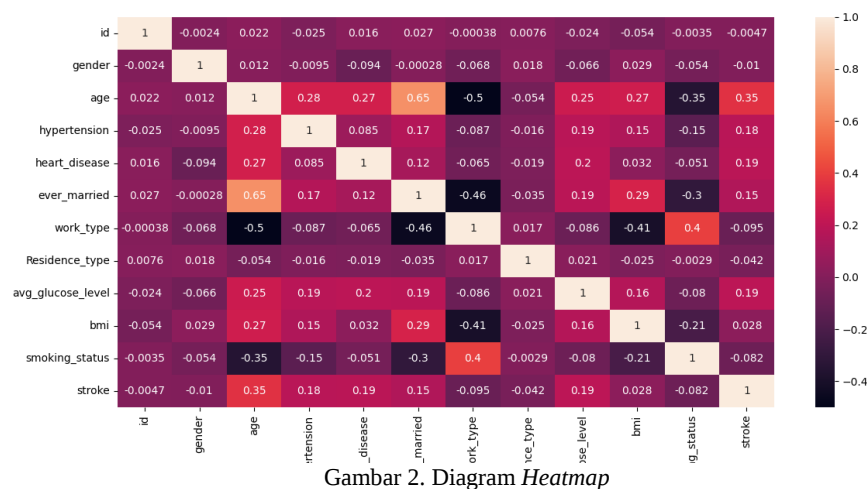
Data yang digunakan adalah dataset yang bersumber dari *website Kaggle*. Dataset ini memiliki 1.599 record data dan 12 atribut.

Tabel 1.
Deskripsi atribut

Atribut	Deskripsi
Id	id pasien
Gender	jenis kelamin pasien
Age	umur pasien
Hypertension	riwayat penyakit hipertensi pasien
heart_disease	riwayat penyakit jantung pasien
ever_married	status pernikahan
work_type	tipe pekerjaan
Residence_type	tipe tempat tinggal
avg_glucose_level	rata-rata level glukosa di dalam darah

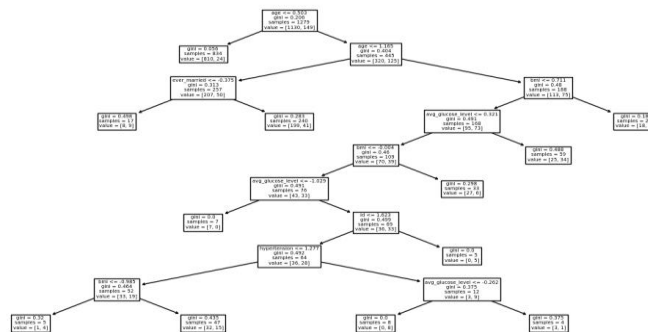
Bmi	index massa tubuh pasien
smoking_status	status merokok
Stroke	prediksi penyakit stroke

Gambar 2 menunjukkan diagram *heatmap* untuk memperlihatkan korelasi dari masing-masing variabel. Variabel yang mempunyai korelasi tertinggi dengan kelas target yaitu variabel stroke adalah variabel *age* (umur). Kemudian diikuti dengan variabel *heart_disease* dan *avg_glucose_level*, dilanjutkan dengan *hypertension*, dan yang terakhir adalah variabel *ever_married*.



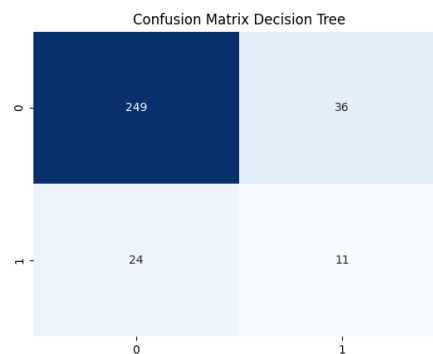
Gambar 2. Diagram *Heatmap*

Hasil dari pengujian menggunakan metode *Decision Tree* dapat dilihat pada gambar 3. Hasil dari pengujian berdasarkan gambar 3 menunjukkan bahwa *root* pada *tree* adalah variabel umur dan terdapat 4 *leaf*, yaitu status pernikahan, BMI, level rata-rata glukosa, dan riwayat hipertensi. Sedangkan variabel jenis kelamin, tipe pekerjaan, tipe tempat tinggal, riwayat penyakit jantung dan status merokok tidak berpengaruh pada prediksi pengidap stroke. Jadi, variabel yang digunakan adalah 5 variabel, yaitu umur, status pernikahan, BMI, level rata-rata glukosa, dan riwayat hipertensi.



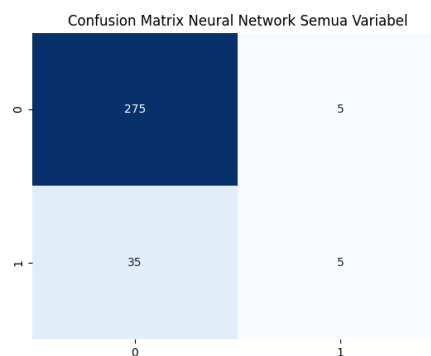
Gambar 3. Hasil *Decision Tree*

Berdasarkan gambar 4, didapatkan hasil bahwa pada kolom True Positif (1,1) yang menunjukkan prediksi positif Stroke dan benar sejumlah 11 orang memiliki arti bahwa terdapat sekitar 11 orang yang didiagnosa bahwa benar telah positif stroke atau terkena stroke. Selanjutnya kolom False Negatif (1,0) yang menunjukkan prediksi tidak stroke dan salah sejumlah 36 orang memberikan arti bahwa terdapat 36 orang yang tidak menderita atau negatif stroke, tetapi sebenarnya orang tersebut terindikasi positif stroke. Kolom False Positif (0,1) dengan prediksi positif Stroke dan salah sejumlah 24 berarti bahwa terdapat 24 orang yang menderita stroke tetapi keadaan sesungguhnya menunjukkan bahwa orang tersebut tidak terindikasi positif stroke. Kolom terakhir yaitu True Negatif (0,0) dengan prediksi tidak stroke dan benar sejumlah 249 memiliki arti bahwa sejumlah 249 orang tidak menderita stroke.



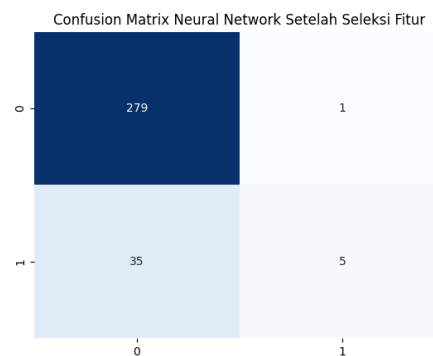
Gambar 4. *Confusion Matrix Decision Tree*

Sementara itu, gambar 5 menunjukkan hasil confusion matriks dari *neural network* semua variabel. Hasilnya menunjukkan bahwa pada kolom True Positif (1,1) yang menunjukkan prediksi positif Stroke dan benar sejumlah 5 orang memiliki arti bahwa terdapat sekitar 5 orang yang didiagnosa bahwa benar telah positif stroke atau terkena stroke. Selanjutnya kolom False Negative (1,0) yang menunjukkan prediksi tidak stroke dan salah sejumlah 5 orang memberikan arti bahwa terdapat 5 orang yang tidak menderita atau negatif stroke, tetapi sebenarnya orang tersebut terindikasi positif stroke. Kolom False Positif (0,1) dengan prediksi positif Stroke dan salah sejumlah 35 berarti bahwa terdapat 35 orang yang menderita stroke tetapi keadaan sesungguhnya menunjukkan bahwa orang tersebut tidak terindikasi positif stroke. Kolom terakhir yaitu True Negatif Negatif (0,0) dengan prediksi tidak stroke dan benar sejumlah 275 memiliki arti bahwa sejumlah 275 orang tidak menderita stroke.



Gambar 5. *Confusion Matrix Neural Network sebelum seleksi fitur*

Gambar 6 menunjukkan bahwa pada kolom True Positif (1,1) yang menunjukkan prediksi positif Stroke dan benar sejumlah 5 orang memiliki arti bahwa terdapat sekitar 5 orang yang didiagnosa bahwa benar telah positif stroke atau terkena stroke. Selanjutnya kolom False Negative (1,0) yang menunjukkan prediksi tidak stroke dan benar sejumlah 1 orang memberikan arti bahwa terdapat 1 orang yang tidak menderita atau negatif stroke, tetapi sebenarnya orang tersebut terindikasi positif stroke. Kolom False Positif (0,1) dengan prediksi positif Stroke dan salah sejumlah 35 berarti bahwa terdapat 35 orang yang menderita stroke tetapi keadaan sesungguhnya menunjukkan bahwa orang tersebut tidak terindikasi positif stroke. Kolom terakhir yaitu False Negatif (0,0) dengan prediksi tidak stroke dan benar sejumlah 279 memiliki arti bahwa sejumlah 279 orang tidak menderita stroke.



Gambar 6. Confusion Matriks Neural Network Setelah Seleksi Fitur

Pada tabel 2 ditunjukkan hasil akurasi dari masing-masing model. Berdasarkan hasil pengolahan data dengan menggunakan model *Decision Tree* dan *Neural Network* menunjukkan hasil bahwa metode dengan akurasi tertinggi untuk memprediksi penyakit stroke adalah metode *neural network* dengan akurasi sebesar 87,1825 dan 88,75 sedangkan untuk *decision tree* hanya sebesar 81,25. Metode *neural network* yang mempunyai akurasi tertinggi dalam memprediksi stroke adalah *neural network* dengan inputan yang telah dilakukan seleksi fitur dari proses *decision tree* yang terdiri dari variabel input age, hypertension, ever_married, avg_glucose_level, dan bmi dengan akurasi hasil *neural network* sebesar 88,75. Hal tersebut menunjukkan bahwa seleksi fitur merupakan proses yang sangat berpengaruh dalam menentukan hasil akhir sebuah pengolahan data karena dengan melakukan seleksi fitur dapat menghilangkan dimensi data dan fitur-fitur yang tidak relevan dalam melakukan proses klasifikasi sehingga diperoleh hasil yang lebih akurat.

Tabel 2.
Hasil pengukuran akurasi tiap model

Performance	Model		
	Decision Tree	Neural Network tanpa Seleksi Fitur	Neural Network dengan Seleksi Fitur
Akurasi	81,25	87,1875	88,75

IV. KESIMPULAN

Berdasarkan dari hasil pengujian yang dilakukan pada penelitian ini, dapat disimpulkan bahwa:

1. Hasil akurasi tertinggi diperoleh menggunakan metode *Neural Network* dengan seleksi fitur yaitu sebesar 88,75. Dan pada metode *Neural Network* didapatkan akurasi sebesar 87,1875. Sedangkan akurasi terendah diperoleh sebesar 81,25 menggunakan metode *Decision Tree*.
2. Metode *Decision Tree* memberikan hasil berupa *root* adalah age dan memiliki 4 *leaf*, di antaranya *ever_married*, BMI, *avg_glucose_level* serta *hypertension*.

REFERENSI

- [1] Widyaswara suwaryo, P. A., Widodo, W. T., & Setianingsih, E. (2019). Faktor Risiko yang Mempengaruhi Kejadian Stroke. *Jurnal Keperawatan* (Vol. 11, No. 4, hal 251-260).
- [2] Mutiarasari, D., Kesehatan, B. I., Komunitas, M.- K., & Kedokteran, F. (2019). ISCHEMIC STROKE: SYMPTOMS, RISK FACTORS, AND PREVENTION. In *Jurnal Ilmiah Kedokteran* (Vol. 6, Issue 1).
- [3] Sailasya, G., & Kumari, G. L. A. (2021). Analyzing the Performance of Stroke Prediction using ML Classification Algorithms. *International Journal of Advanced Computer Science and Applications* (Vol. 12, Issue 6).
- [4] Budianita, E., Oktafalisa., & Ofandi, A. (2020). Penerapan Decision Tree dan Neural Network pada Penentuan Kategori Status Gizi Balita. *Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI)* (Vol. 12).
- [5] Nurlia, Elin., & Enri, Ultach. (2021). Penerapan Fitur Seleksi Forward Selection Untuk Menentukan Kematian Akibat Gagal Jantung Menggunakan Algoritma C4.5. *JUTIM (Jurnal Teknik Informatika Musirawas)* (Vol. 6, No 1).
- [6] Nurlelah, E., Utami, Y, D. (2022). Seleksi Atribut Pada Algoritma Neural Network Menggunakan Particle Swarm Optimization Untuk Diagnosis Penyakit Liver. *Kumpulan jurnaL Ilmu Komputer (KLIK)* (Vol. 9, No. 2).
- [7] Kasih, Patmi. (2019). Pemodelan Data Mining Decision Tree Dengan Classification Error Untuk Seleksi Calon Anggota Tim Paduan Suara. *Innovation In Research Of Informatics* (Vol.1, No. 2, hal. 63-69).
- [8] Mohsena, A., Alsurorib, M., & Buthiena, A. (2019). New Approach to Medical Diagnosis Using Artificial Neural Network and Decision Tree Algorithm: Application to Dental Diseases. *International Journal Information Engineering and Electronic Business* (Vol 4).
- [9] S, Isha., & Borikar, A. (2022). Heart Disease Prediction using Data Mining Techniques. *Journal Of Algebraic Statistics* (Vol. 13, No.3).
- [10] Tigga, P., & Garg, S, (2020). Prediction of Type 2 Diabetes using Machine Learning Classification Methods. *International Conference on Computational Intelligence and Data Science (ICCIDS 2019)*.
- [11] Pambudi, E. R., Sriyanto., & Firmansyah. (2022). Klasifikasi Penyakit Stroke Menggunakan Algoritma Decision Tree C.45. *Jurnal Teknika* (Vol. 16, No. 2, hal 221-226).
- [12] Drajana Rally, C. A., & Bode, A. (2022). Prediksi Status Penderita Stunting Pada Balita Provinsi Gorontalo Menggunakan K-Nearest Neighbor Berbasis Seleksi Fitur Chi Square. *Jurnal Nasional Komputasi dan Teknologi Informasi* (Vol. 5, No. 2).

- [13] Gupta, K., Jiwani, N., Afreen, N., & D, Divyarani. (2022). Liver Disease Prediction using Machine learning Classification Techniques. 11th IEEE International Conference on Communication Systems and Network Technologies.
- [14] Wahyuni, S., Elvira. (2016). Penerapan Metode Seleksi Fitur Untuk Meningkatkan Hasil Diagnosis Kanker Payudara. Jurnal SIMETRIS (Vol. 7, No. 1).
- [15] Prasetyo, Y., Simoen. (2023). Prediksi Gagal Jantung Menggunakan Artificial Neural Network. Jurnal SAINTEKOM (Vol.13, No. 1, hal. 79-88).
- [16] Rane, N., Sunny, J., Kanade, R., & Devi, S. (2020). Breast Cancer Classification and Prediction using Machine Learning. International Journal of Engineering Research & Technology (IJERT).
- [17] Nasser, M., Al-Shawwa., & Abu-Naser, S. (2019). A Proposed Artificial Neural Network for Predicting Movies Rates Category. International Journal of Academic Engineering Research (IJAER) (Vol. 3, Issue. 2).
- [18] Draper, N., Smith, H., 1992. *Analisis Regresi Terapan Edisi Kedua*, PT. Gramedia Pustaka Utama, Jakarta.
- [19] Fadly, M., Uddin, N., & Sutarto, H. Y., 2002. *Flutter Suppression Using Neural Networks : Design and Implementation*.
- [20] I. H. Witten, E. Frank, and M. a. Hall, 2011. *Data Mining Practical Machine Learning Tools and Techniques Third Edition*, vol. 277. Tentang Data Mining.
- [21] Stroke Prediction Dataset, Kaggle, May., 2022. [Online]. Available: <https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset>.
- [22] World Health Organization (WHO). 2019. 10 besar penyebab kematian. Diakses pada 25 April 2023, dari <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>
- [23] World Health Organization (WHO). 2023. Stroke, Cerebrovascular accident. Diakses pada 25 April 2023, dari <https://www.emro.who.int/health-topics/stroke-cerebrovascular-accident/index.html>
- [24] Dev, S., Wang, H., Nwosu, C. S., Jain, N., Veeravalli, B., & John, D. (2022). A predictive analytics approach for stroke prediction using machine learning and neural networks. Healthcare Analytics, 2, 100032.
- [25] Rajora, M., Rathod, M., & Naik, N. S. (2021). Stroke prediction using machine learning in a distributed environment. In Distributed Computing and Internet Technology: 17th International Conference, ICDCIT 2021, Bhubaneswar, India, January 7–10, 2021, Proceedings 17 (pp. 238-252). Springer International Publishing.