UNIVERSITAS
MUHAMMADIYAH
SIDOARJO

UMSIDA
DARI SINI PENCERAHAN BERSEMI

Prediksi Penyakit Stroke Menggunakan Metode Random Forest

Oleh:

Priyo Wahyu Setiyo Aji

Suprianto

Progam Studi Informatika

Universitas Muhammadiyah Sidoarjo

Agustus, 2023

Pendahuluan

- WHO mendefinisikan strok sebagai kerusakan fungsi otak secara tiba-tiba yang terjadi dalam waktu 24 jam atau lebih [1]. Stroke merupakan penyakit cerebrovascular atau cedera otak dimana pembuluh darah tersumbat dan membatasi suplai darah ke otak.

- Penyakit Stroke merupakan penyakit neurologis terbanyak yang mengakibatkan masalah kesehatan serius dan dampaknya adalah penderita mengalami kecacatan hingga kematian [2].

- Machine Learning adalah satu cabang dari kecerdasan buatan yang dikembangkan agar mesin dapat belajar secara otomatis. Machine learning berfokus pada analisis data untuk menemukan hubungan antar input dan output yang diinginkan

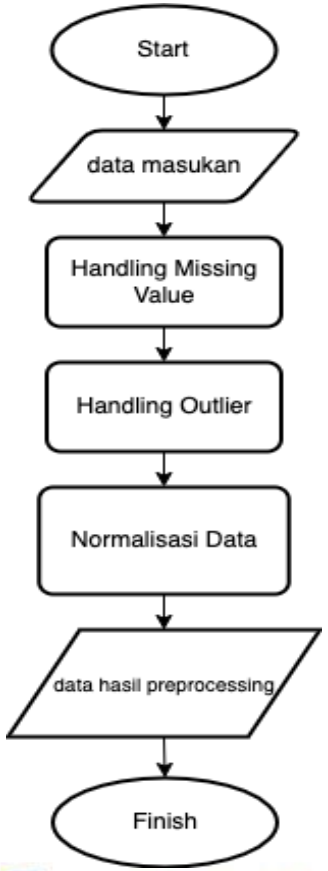
Pertanyaan Penelitian (Rumusan Masalah)

- Permasalahan utama yang diungsung oleh peneliti studi ini adalah bagaimana membangun model machine learning yang dapat digunakan untuk memprediksi penyakit stroke ?
- Bagaimana cara mengimplementasi kinerja perangkat lunak menggunakan Randome Forest ?

Metode

Tahap Preprocessing

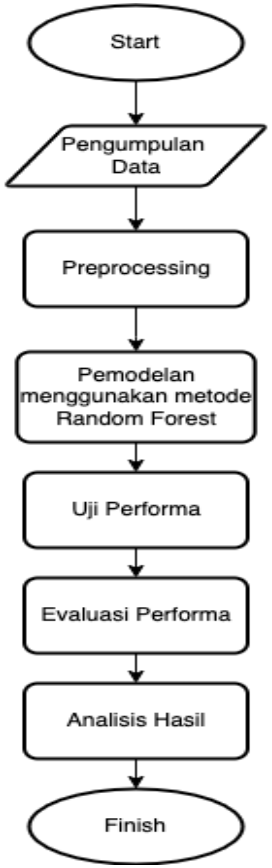
1. Pembersihan Data Setelah dilakukan modeling



Metode

Tahap Preprocessing

2. Pemodelan Menggunakan Metode Random Forest



Metode

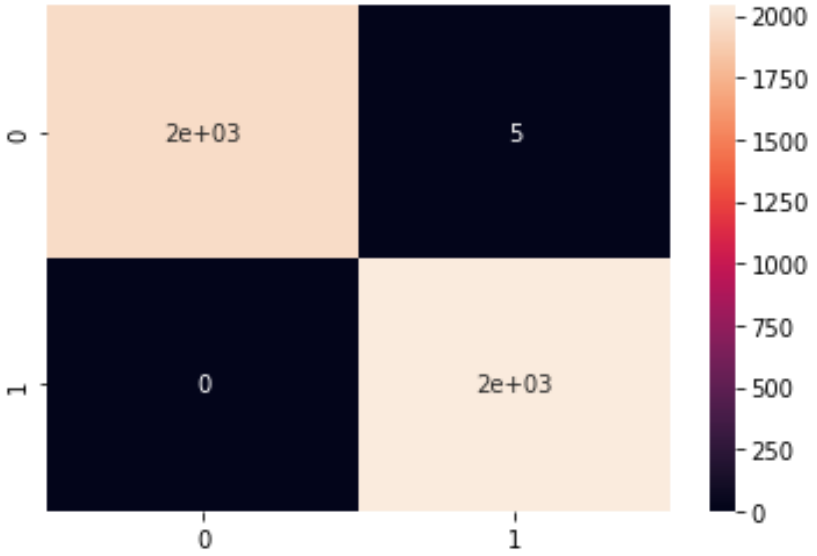
Tahap Preprocessing

3. Klasifikasi dengan Random Forest

Tahap klasifikasi dengan Random Forest adalah membuka data yang telah dilakukan ekstraksi fitur ke dalam jupyter notebook menggunakan library pandas. Ketika data sudah berhasil diload, maka dilakukan pembagian data . Untuk mendapatkan parameter yang paling optimal pada kasus ini, menggunakan Teknik Randomized Search Cross Validation. Teknik ini dapat mencari parameter optimal dari algoritma yang digunakan untuk kasus yang sedang dianalisa. Randomized Search Cross Validation adalah bagian dari modul scikit-learn yang bertujuan secara otomatis dan sistematis melakukan validasi beberapa model dan setiap hyperparameter. Ketika proses running Randomized Search Cross Validation sudah selesai, maka akan didapatkan model beserta score test dan score train.

Evaluasi

Setelah tahap klasifikasi selesai, evaluasi dilakukan untuk mengukur akurasi, presisi, recall, dan skor F1. Pada penelitian ini diperoleh skor akurasi sebesar 99% yang dapat dilihat pada Gambar bawah. Untuk menampilkan confusion matrix yang terdapat pada Gambar atas, peneliti menggunakan pustaka scikit-learn dengan memanggil fungsi metrik. Di sisi lain, untuk menampilkan perhitungan lebih lanjut dari confusion matrix yang terdapat pada Gambar bawah, Peneliti juga memanfaatkan pustaka scikit-learn dengan menggunakan fungsi classification_report.



	precision	recall	f1-score	support
0.0	1.00000	0.99643	0.99821	1961
1.0	0.99658	1.00000	0.99829	2040
accuracy			0.99825	4001
macro avg	0.99829	0.99822	0.99825	4001
weighted avg	0.99826	0.99825	0.99825	4001

Manfaat Penelitian

Ada beberapa manfaat yang bisa diperoleh dari temuan penelitian ini, beberapa di antaranya tercantum di bawah ini:

- Seseorang dapat melakukan pencegahan penyakit stroke lebih dini
- Memberikan konstribusi bagi upaya pengembangan aplikasi yang dapat membantu masyarakat medis terutama dokter dan rumah sakit dalam mendiagnosis penyakit stroke dengan tingkat akurasi yang baik.

Referensi

- **REFERENSI**

- [1] M. A. As Sarofi, I. Irhamah, and A. Mukarromah, “Identifikasi Genre Musik dengan Menggunakan Metode Random Forest,” *Jurnal Sains dan Seni ITS*, vol. 9, no. 1, pp. 79–86, 2020, doi: 10.12962/j23373520.v9i1.51311.
- [2] E. Agustin, A. Eviyanti, and N. Lutvi Azizah, “Deteksi Penyakit Epilepsi Melalui Sinyal EEG Menggunakan Metode DWT dan Extreme Gradient Boosting,” vol. 7, no. 1, pp. 117–127, 2023, doi: 10.30865/mib.v7i1.5412.
- [3] N. Permatasari, “Perbandingan Stroke Non Hemoragik dengan Gangguan Motorik Pasien Memiliki Faktor Resiko Diabetes Melitus dan Hipertensi,” *Jurnal Ilmiah Kesehatan Sandi Husada*, vol. 11, no. 1, 2020, doi: 10.35816/jiskh.v11i1.273.
- [4] J. Antares, “Artificial Neural Network Dalam Mengidentifikasi Penyakit Stroke Menggunakan Metode Backpropagation (Studi Kasus di Klinik Apotik Madya Padang),” *Djtechno: Jurnal Teknologi Informasi*, vol. 1, no. 1, 2021, doi:10.46576/djtechno.v1i1.965.

Referensi

- [5] M. A. As Sarofi, I. Irhamah, and A. Mukarromah, “Identifikasi Genre Musik dengan Menggunakan Metode Random Forest,” *Jurnal Sains dan Seni ITS*, vol. 9, no. 1, 2020, doi: 10.12962/j23373520.v9i1.51311.
- [6] F. Akbar, H. Wira Saputra, A. Karel Maulaya, M. Fikri Hidayat, and Rahmaddeni, “Implementasi Algoritma Decision Tree C4.5 dan Support Vector Regression untuk Prediksi Penyakit Stroke,” vol. 2, no. October, pp. 61–67, 2022.
- [7] U. Amelia *et al.*, “Implementasi Algoritma Support Vector Machine (Svm) Untuk Prediksi Penyakit Stroke Dengan Atribut Berpengaruh,” ... *Student Journal for* ..., vol. III, pp. 254–259, 2022.
- [8] M. N. Maskuri, K. Sukerti, and R. M. Herdian Bhakti, “Penerapan Algoritma K-Nearest Neighbor (KNN) untuk Memprediksi Penyakit Stroke Stroke Desease Predict Using KNN Algorithm,” *Jurnal Ilmiah Intech : Information Technology Journal of UMUS*, vol. 4, no. 1, pp. 130–140, 2022.

Referensi

- [9] A. Byna and M. Basit, “Penerapan Metode Adaboost Untuk Mengoptimasi Prediksi Penyakit Stroke Dengan Algoritma Naïve Bayes,” *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 9, no. 3, pp. 407–411, 2020, doi: 10.32736/sisfokom.v9i3.1023.
- [10] D. E. Cahyani, “Penerapan Machine Learning Untuk Prediksi Penyakit Stroke,” *Jurnal Kajian Matematika dan Aplikasinya*, vol. 3, no. January, pp. 8–14, 2022, doi: 10.17977/um055v3i1p15-22.
- [11] D. Prajarini, S. Tinggi, S. Rupa, D. Desain, and V. Indonesia, “Perbandingan Algoritma Klasifikasi Data Mining Untuk Prediksi Penyakit Kulit,” *Informatics Journal*, vol. 1, no. 3, p. 137, 2016.
- [12] R. S. Rohman, R. A. Saputra, and D. A. Firmansaha, “Komparasi Algoritma C4.5 Berbasis PSO Dan GA Untuk Diagnosa Penyakit Stroke,” *CESS (Journal of Computer Engineering, System and Science)*, vol. 5, no. 1, p. 155, 2020, doi: 10.24114/cess.v5i1.15225.

Referensi

- [13] M. A. As Sarofi, I. Irhamah, and A. Mukarromah, “Identifikasi Genre Musik dengan Menggunakan Metode Random Forest,” *Jurnal Sains dan Seni ITS*, vol. 9, no. 1, pp. 79–86, 2020, doi: 10.12962/j23373520.v9i1.51311.
- [14] J. Xu, Y. Zhang, and D. Miao, “Three-way confusion matrix for classification: A measure driven view,” *Inf Sci (N Y)*, vol. 507, 2020, doi: 10.1016/j.ins.2019.06.064.

UMSIDA
DARI SINI PENCERAHAN BERSEMI