

PERBANDINGAN METODE ARTIFICIAL NEURAL NETWORK, DAN RANDOM FOREST PADA KLASIFIKASI TINGKAT OBESITAS

Agung Indra Prasetyo¹⁾, Amak Yunus²⁾, Alexius Endy Budianto³⁾

Universitas PGRI Kanjuruhan Malang

Email: agungindra969@gmail.com

Abstrak

Klasifikasi tingkat obesitas merupakan langkah penting dalam mendukung upaya penanganan obesitas yang semakin meningkat. Penelitian ini bertujuan untuk membandingkan performa metode machine learning, yaitu Artificial Neural Network (ANN) dan Random Forest (RF), dalam melakukan klasifikasi tingkat obesitas berdasarkan dataset yang telah ditentukan. Metode penelitian melibatkan proses preprocessing data, pelatihan model dengan variasi proporsi data training dan data testing (70:30, 80:20, dan 90:10). Hasil penelitian menunjukkan bahwa Random Forest memberikan hasil akurasi yang lebih tinggi dibandingkan dengan Artificial Neural Network. Pada pengujian dengan data training 70% dan data testing 30%, ANN menghasilkan akurasi sebesar 88,20% sementara RF mencapai 97,28%. Dengan proporsi data training 80% dan data testing 20%, akurasi ANN meningkat menjadi 88,76%, sedangkan RF menghasilkan 97,37%. Pada proporsi data training 90% dan data testing 10%, ANN mencapai akurasi tertinggi sebesar 91,39%, namun masih lebih rendah dibandingkan RF yang mencapai 95,69%. Berdasarkan hasil tersebut, dapat disimpulkan bahwa algoritma Random Forest menunjukkan performa yang lebih optimal dibandingkan Artificial Neural Network dalam klasifikasi tingkat obesitas.

Kata Kunci : Artificial Neural Network; Random Forest; Klasifikasi; Obesitas

Abstract

Classification of obesity levels is an important step in supporting efforts to tackle the increasing prevalence of obesity. This study aims to compare the performance of machine learning methods, namely Artificial Neural Network (ANN) and Random Forest (RF), in classifying obesity levels based on a predetermined dataset. The research method involved data preprocessing and model training with varying proportions of training and testing data (70:30, 80:20, and 90:10). The results showed that Random Forest provided higher accuracy than Artificial Neural Network. In testing with 70% training data and 30% testing data, ANN produced an accuracy of 88.20% while RF reached 97.28%. With a training data proportion of 80% and testing data of 20%, the accuracy of ANN increased to 88.76%, while RF produced 97.37%. With a training data proportion of 90% and testing data of 10%, ANN achieved the highest accuracy of 91.39%, but it was still lower than RF, which reached 95.69%. Based on these results, it can be concluded that the Random Forest algorithm shows more optimal performance than Artificial Neural Network in obesity level classification.

Keywords : Artificial Neural Network; Random Forest; Klasifikasi; Obesitas

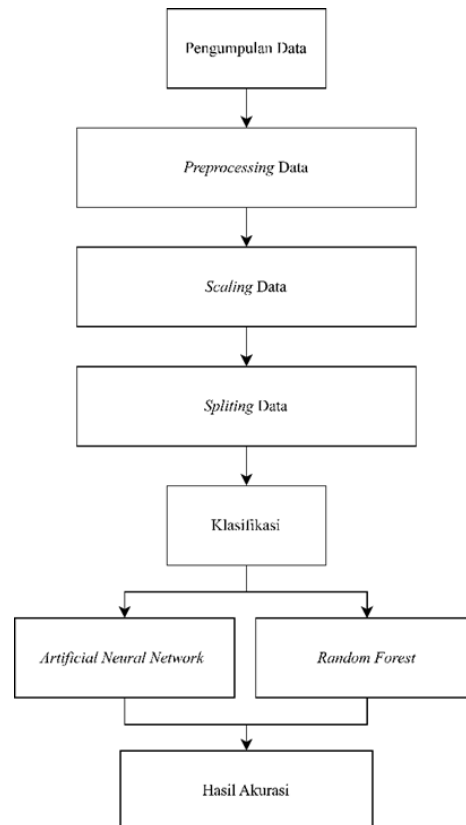
1. PENDAHULUAN

Obesitas adalah kondisi penumpukan lemak berlebih yang terjadi akibat ketidakseimbangan antara energi yang masuk melalui makanan (*energy intake*) dan energi yang digunakan tubuh (*energy expenditure*) dalam jangka waktu panjang. Kondisi ini telah berkembang menjadi permasalahan kesehatan global yang semakin meningkat, dengan jumlah keseluruhan kasus yang terus bertambah setiap tahunnya. Di Indonesia, prevalensi obesitas dewasa meningkat dari 21,8% (2018) menjadi 23,4% (2023), sehingga hampir satu dari empat orang dewasa mengalami obesitas. Kondisi ini tidak hanya berdampak pada kesehatan, tetapi juga

menimbulkan beban ekonomi yang signifikan. Deteksi dini dan klasifikasi tingkat obesitas sangat penting untuk mendukung strategi pencegahan serta penanganan yang tepat. Dalam bidang kesehatan, algoritma pembelajaran mesin seperti *Artificial Neural Network* (ANN) dan *Random Forest* (RF) telah banyak digunakan untuk klasifikasi, dengan penelitian sebelumnya menunjukkan performa yang bervariasi pada berbagai kasus penyakit.

2. METODE / ALGORITMA

Pada penelitian ini, desain penelitian dibuat untuk menjelaskan dari proses awal penelitian hingga hasil penelitian Berikut merupakan sedikit gambaran mengenai metode penelitian :



Gambar 1. Metode Penelitian

Berikut ini merupakan tahapan penelitian:

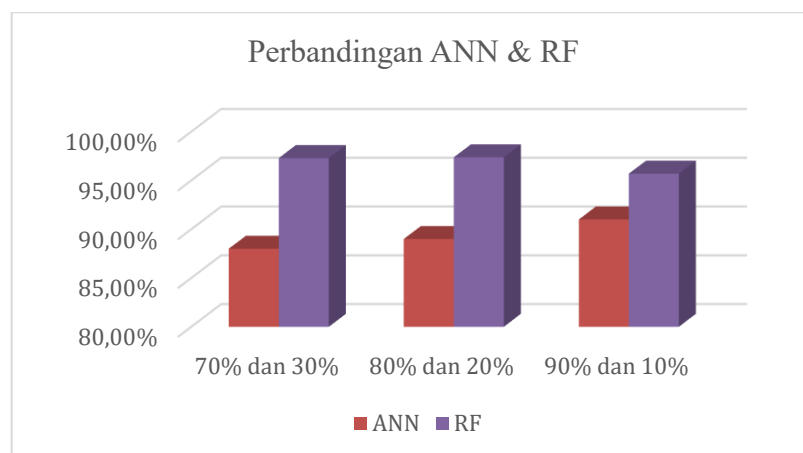
1. Memasukan data yang telah di dapat dari laman *UCI Machine Learning Reppository*
2. Analisis dan eksplorasi data, variabel akan dibagi menjadi kategorikal dan numerik, dan digambarkan secara terpisah untuk memberikan gambaran keseluruhan.
3. Pembersihan data, membersihkan dataset dari record duplikat, pada dataset ini terdapat 24 record duplikat dari 2111 record
4. Data Preprocessing mengimplementasikan Label Encoding untuk mengubah variabel kategorikal menjadi kolom biner, untuk meningkatkan preprocessing data untuk model.
5. Scaling dara Scaling yang tepat dapat meningkatkan performa model, mengurangi waktu pelatihan, dan menghindari bias terhadap fitur yang memiliki skala yang lebih besar.
6. Data Training, memisahkan fitur target dari dataframe lainnya untuk melatih dan menguji model.
7. Membandingkan tingkat akurasi dari algoritma *Artificial Neural Network*, dan *Random Forest*

3. HASIL DAN PEMBAHASAN

Perbandingan *Artificial Neural Network* dan *Random forest* menggunakan dataset obesitas dengan jumlah 2111 data dan 17 atribut, data dibagi menjadi data training dan data testing dengan presentase yang berbeda. Pada tabel dibawah ini merupakan hasil dari perbandingan kedua algoritma tersebut:

Data Training Dan Data Testing	ANN	RF
70% Dan 30%	88,20%	97,29%
80% Dan 20%	88,76%	97,37%
90% Dan 10%	91,39%	95,69%

Pada pengujian algoritma ANN dengan menggunakan data training 70% dan data testing 30% menghasilkan akurasi 88,20% sedangkan untuk algoritma Random Forest menghasilkan akurasi sebesar 97,28%, kemudian untuk data training 80% dan data testing 20% algoritma ANN menghasilkan akurasi 88,76% dan Random forest menghasilkan akurasi 97,37%, dan yang terakhir untuk data training 90% dan data testing 10% algoritma ANN menghasilkan akurasi 91,39% dan Random forest menghasilkan akurasi 95,69%, di bawah ini merupakan grafik dari hasil perbandingan algoritma ANN dan Random Forest :



Gambar 2. Grafik Perbandingan

4. KESIMPULAN

Penelitian ini bertujuan untuk membandingkan hasil akurasi algoritma *Artificial Neural Network* (ANN) dan *Random Forest* (RF), dengan dataset tingkat obesitas. Berdasarkan analisis yang dilakukan, terdapat beberapa temuan penting yang dapat disimpulkan sebagai berikut: Perbandingan Hasil Klasifikasi: Algoritma Random Forest menunjukkan hasil akurasi yang lebih baik dibandingkan dengan Artificial Neural Network. Pada pengujian algoritma ANN dengan menggunakan data training 70% dan data testing 30% menghasilkan akurasi 88,20% sedangkan untuk algoritma Random Forest menghasilkan akurasi sebesar 97,28%, kemudian untuk data training 80% dan data testing 20% algoritma ANN menghasilkan akurasi 88,76% dan Random forest menghasilkan akurasi 97,37%, dan yang terakhir untuk data training 90% dan data testing 10% algoritma ANN menghasilkan akurasi 91,39% dan Random forest menghasilkan akurasi 95,69%.

Tingkat Akurasi: Berdasarkan perbandingan tingkat akurasi, Random Forest memberikan performa yang lebih optimal jika dibandingkan dengan Artificial Neural Network. Meskipun demikian ANN memiliki kemampuan yang cukup baik, hasil akurasi yang dicapai menunjukkan bahwa RF lebih unggul dalam memproses dataset ini, yang disebabkan oleh

kemampuannya dalam menangani fitur yang lebih kompleks dan interaksi antar variabel yang lebih efisien.

Secara keseluruhan, penelitian ini mengindikasikan bahwa dalam konteks klasifikasi tingkat obesitas, Random Forest lebih efektif dibandingkan Artificial Neural Network, namun kedua metode tetap memiliki relevansi tergantung pada kebutuhan spesifik dari penggunaannya.

REFERENSI

- [1] Abbi Rokhman K, Berliana, & Arsi Primadani. (2021). Perbandingan Metode Support Vector Machine Dan Decision Tree Untuk Analisis Sentimen Review Komentar Pada Aplikasi Transportasi Online. *Jurnal Of Information System Management*, Vol. 2, No. 2.
- [2] Almufqi, F. M., & Voutama, A. (2023). Perbandingan Metode Data Mining Untuk Memprediksi Prestasi Akademik Siswa. *Jurnal Teknika*, 15(1), 61–66. <https://doi.org/10.30736/jt.v15i1.929>
- [3] Apriyani, H. (2020). Perbandingan Metode Naïve Bayes Dan Support Vector Machine Dalam Klasifikasi Penyakit Diabetes Melitus. In *Journal of Information Technology Ampera* (Vol. 1, Issue 3). <https://journal-computing.org/index.php/journal-ita/index>
- [4] Bidang SDK (ags). (2023). Permasalahan Obesitas di Indonesia. <https://Dinkes.Jogjaprovo.go.id/Berita/Detail/Permasalahan-Obesitas-Di-Indonesia?Utm>.
- [5] Fattya Ariani, & Andi Taufik. (2020). Perbandingan Metode Klasifikasi Data Mining untuk Prediksi Tingkat Kepuasan Pelanggan Telkomsel Prabayar. *SATIN - Sains Dan Teknologi Informasi*, 6(2), 46–55. <https://doi.org/10.33372/stn.v6i2.666>
- [6] Haganta Depari, D., Widiastiwi, Y., Mega Santoni, M., Ilmu Komputer, F., Pembangunan Nasional Veteran Jakarta, U., Fatmawati Raya, J. R., & Labu, P. (2022). Perbandingan Model Decision Tree, Naïve Bayes dan Random Forest untuk Prediksi Klasifikasi Penyakit Jantung. *JURNAL INFORMATIK Edisi Ke*, 18, 2022.
- [7] Badan Kebijakan Pembangunan Kesehatan. (2023). <https://www.badankebijakan.kemkes.go.id/daftar-frequently-asked-question-seputar-hasil-utama-ski-2023/hasil-utama-ski-2023/?Utm>.
- [8] Jaringan Pangan & Gizi Indonesia. (2023). Pencegahan dan Pengendalian Obesitas di Indonesia. <https://jpg-indonesia.net/2023/11/pencegahan-dan-pengendalian-obesitas-di-indonesia/?Utm>.
- [9] Kusumarini, A. I., Hogantara, P. A., & Chamidah, N. (2021). Perbandingan Algoritma Random Forest, Naïve Bayes, Dan Decision Tree Dengan Oversampling Untuk Klasifikasi Bakteri E. Coli. In *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA) Jakarta-Indonesia*.
- [10] Marito Putry, N., & Nurina Sari, B. (2022). Komparasi Algoritma Knn Dan Naïve Bayes Untuk Klasifikasi Diagnosis Penyakit Diabetes Melitus. *Jurnal Sains Dan Manajemen*, 10(1).
- [11] Mega Santoni, M., Chamidah, N., & Matondang, N. (2020). Prediction of Hypertension using Decision Tree, Naïve Bayes and Artificial Neural Networks in KNIME Analytics Platform (Vol. 19, Issue 4).
- [12] Obesity and overweight. (2024). <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>.
- [13] Subarkah, P., Pambudi, E. P., & Hidayah, S. O. N. (2020). Perbandingan Metode Klasifikasi Data Mining untuk Nasabah Bank Telemarketing. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 20(1), 139–148. <https://doi.org/10.30812/matrik.v20i1.826>

- [14] Wibowo, M., & Ramadhani, R. (2021). Perbandingan Metode Klasifikasi Data Mining Untuk Rekomendasi Tanaman Pangan. *Jurnal Media Informatika Budidarma*, 5(3), 913. <https://doi.org/10.30865/mib.v5i3.3086>
- [15] Yoga Religia, Agung Nugroho, & Wahyu Hadikristanto. (2021). Klasifikasi Analisis Perbandingan Algoritma Optimasi pada Random Forest untuk Klasifikasi Data Bank Marketing. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(1), 187–192. <https://doi.org/10.29207/resti.v5i1.2813>