

ABSTRAK

PENGEMBANGAN ALGORITMA RANDOM FOREST MENGGUNAKAN METODE *REGULARIZED IMPURITY REDUCTION (GREEDY RANDOM FOREST)*

(2025: XIII+ 78 halaman + 41 gambar + 11 tabel)

KHAIRUNNISA

062140352374

JURUSAN TEKNIK ELEKTRO

**PROGRAM STUDI SARJANA TERAPAN TEKNIK TELEKOMUNIKASI
POLITEKNIK NEGERI SRIWIJAYA**

Meningkatnya kompleksitas dan besarnya data di bidang kesehatan menuntut pengembangan model *machine learning* yang lebih adaptif dan andal, khususnya untuk tugas klasifikasi berskala besar. *Random Forest* merupakan salah satu metode *ensemble* yang banyak digunakan karena ketangguhannya dalam menghadapi permasalahan seperti ketidakseimbangan kelas dan dimensi fitur yang tinggi. Namun, model *Random Forest* standar masih memiliki kelemahan terhadap *overfitting* dan dominasi fitur, terutama akibat mekanisme pemilihan *split* yang bersifat lokal. Oleh karena itu, penelitian ini mengusulkan *Greedy Random Forest*, sebuah algoritma yang disempurnakan melalui mekanisme pemilihan *split* menggunakan *Regularized Impurity Reduction* yang mempertimbangkan *balance*, *efficiency*, dan *discrimination*. Evaluasi dilakukan pada dua dataset kesehatan, yaitu Alzheimer dan Hypothyroid. Hasil pengujian menunjukkan bahwa *Greedy Random Forest* mampu meningkatkan performa dibandingkan *Random Forest* standar. Pada dataset Alzheimer, *Greedy Random Forest* menghasilkan akurasi sebesar 0,9372 sedikit lebih baik dari *Random Forest* yang memiliki akurasi 0,9349, signifikan terlihat pada dataset Hypothyroid, di mana *Greedy Random Forest* mencapai akurasi 0,9677, mengungguli *Random Forest* yang hanya memperoleh akurasi 0,8991. Dengan hasil tersebut, *Greedy Random Forest* terbukti efektif dalam mengatasi *overfitting*, meningkatkan generalisasi model, serta menghasilkan performa yang lebih konsisten dan interpretatif.

Kata Kunci: *Greedy Algorithm, Health Data, Machine Learning, Random Forest.*

ABSTRACT

DEVELOPMENT OF THE RANDOM FOREST ALGORITHM USING THE REGULARIZED IMPURITY REDUCTION METHOD (GREEDY RANDOM FOREST)

(2025: XIII + 78 halaman + 41 gambar + 11 tabel)

KHAIRUNNISA

062140352374

JURUSAN TEKNIK ELEKTRO

PROGRAM STUDI SARJANA TERAPAN TEKNIK TELEKOMUNIKASI

POLITEKNIK NEGERI SRIWIJAYA

The increasing complexity and scale of healthcare data demand the development of more adaptive and reliable machine learning models, particularly for large-scale classification tasks. Random Forest is one of the most widely used ensemble methods due to its robustness in handling issues such as class imbalance and high-dimensional feature spaces. However, the standard Random Forest model still suffers from overfitting and feature dominance, mainly due to its locally based split selection mechanism. Therefore, this study proposes the Greedy Random Forest, an enhanced algorithm that applies a split selection mechanism based on Regularized Impurity Reduction, which considers balance, efficiency, and discrimination. Evaluation was conducted on two healthcare datasets, namely Alzheimer and Hypothyroid. The experimental results show that Greedy Random Forest outperforms the standard Random Forest. On the Alzheimer dataset, Greedy Random Forest achieved an accuracy of 0.9372, slightly higher than Random Forest with 0.9349. A more significant improvement was observed on the Hypothyroid dataset, where Greedy Random Forest achieved an accuracy of 0.9677, surpassing Random Forest which only reached 0.8991. These results demonstrate that Greedy Random Forest is effective in reducing overfitting, improving model generalization, and providing more consistent and interpretable performance.

Keywords: Greedy Algorithm, Health Data, Machine Learning, Random Forest.