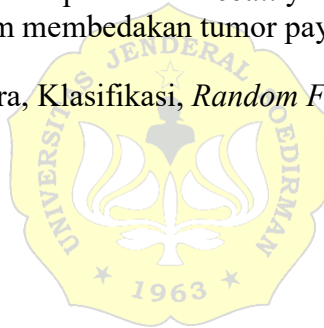


ABSTRAK

Tumor payudara merupakan salah satu masalah kesehatan yang banyak dialami perempuan dan berpotensi berkembang menjadi kanker akibat rendahnya deteksi dini, sehingga diperlukan suatu metode yang mampu melakukan klasifikasi tumor payudara secara akurat berdasarkan data klinis pasien. Penelitian ini bertujuan untuk mengimplementasikan algoritma *Random Forest* dalam mengklasifikasikan tumor payudara menjadi dua kelas, yaitu tumor jinak dan tumor ganas, serta mengevaluasi tingkat akurasi model yang dihasilkan. Data yang digunakan merupakan data rekan medis pasien tumor payudara di RSUD Prof. Dr. Margono Soekarjo. Tahapan penelitian meliputi analisis deskriptif, *preprocessing* data, pembagian data latih dan data uji dengan beberapa rasio, penerapan *hyperparameter tuning*, serta proses klasifikasi menggunakan algoritma *Random Forest*. Evaluasi performa model dilakukan menggunakan komponen *confusion matrix*. Hasil penelitian menunjukkan bahwa algoritma *Random Forest* mampu memberikan performa klasifikasi yang sangat baik. Model menghasilkan akurasi tertinggi sebesar 97,2% pada rasio data 80:20, serta akurasi sebesar 97% dan 96% pada rasio 75:25 dan 70:30. Nilai presisi dan *recall* yang tinggi menunjukkan bahwa *Random Forest* efektif dalam membedakan tumor payudara jinak dan ganas.

Kata kunci: Tumor Payudara, Klasifikasi, *Random Forest*, *Machine Learning*



ABSTRACT

Breast tumors are one of the health problems commonly experienced by women and have the potential to develop into cancer due to low early detection rates. Therefore, a method capable of accurately classifying breast tumors based on patients' clinical data is required. This study aims to implement the Random Forest algorithm to classify breast tumors into two classes, namely benign and malignant tumors, and to evaluate the accuracy of the resulting model. The data used in this study consist of medical record data of breast tumor patients at RSUD Prof. Dr. Margono Soekarjo. The research stages include descriptive analysis, data preprocessing, splitting the dataset into training and testing sets using several ratios, hyperparameter tuning, and classification using the Random Forest algorithm. Model performance evaluation is conducted using confusion matrix components. The results show that the Random Forest algorithm achieves very good classification performance. The highest accuracy of 97.2% is obtained using an 80:20 data split ratio, while accuracies of 97.0% and 96.0% are achieved using 75:25 and 70:30 ratios, respectively. The high precision and recall values indicate that the Random Forest algorithm is effective in distinguishing between benign and malignant breast tumors.

.Keywords: *Breast Tumor, Classification, Random Forest, Machine Learning*

