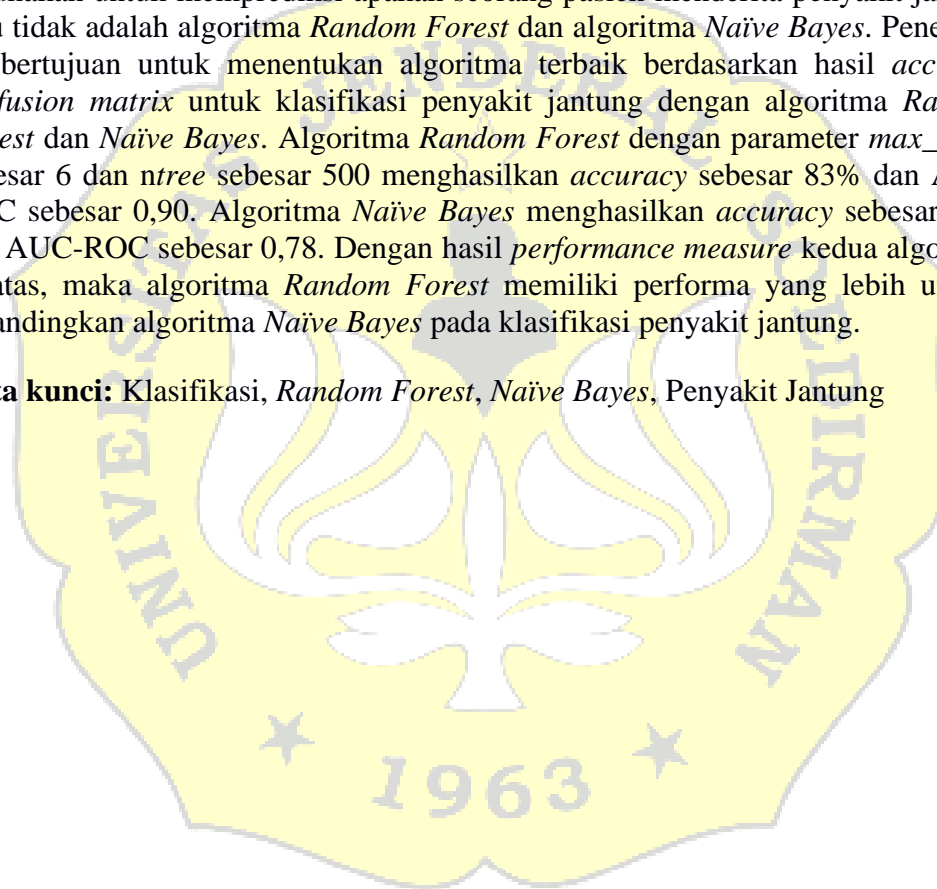


## ABSTRAK

Penyakit jantung adalah masalah kesehatan yang serius dan menjadi penyebab utama kematian di seluruh dunia. Banyak kasus kematian yang terjadi akibat penyakit jantung. Sebagian besar kasus ini terjadi karena kurangnya informasi tentang penyakit jantung yang membuat seseorang sulit mengenali gejalanya, sehingga dibutuhkan suatu cara untuk mendiagnosis penyakit jantung dengan cepat dan mudah. Klasifikasi adalah salah satu teknik dalam *machine learning* yang dapat digunakan untuk memprediksi pendeteksian penyakit jantung berdasarkan gejala atau faktor risiko tertentu. Metode klasifikasi yang dapat digunakan untuk memprediksi apakah seorang pasien menderita penyakit jantung atau tidak adalah algoritma *Random Forest* dan algoritma *Naïve Bayes*. Penelitian ini bertujuan untuk menentukan algoritma terbaik berdasarkan hasil *accuracy* *confusion matrix* untuk klasifikasi penyakit jantung dengan algoritma *Random Forest* dan *Naïve Bayes*. Algoritma *Random Forest* dengan parameter *max\_depth* sebesar 6 dan *n\_tree* sebesar 500 menghasilkan *accuracy* sebesar 83% dan AUC-ROC sebesar 0,90. Algoritma *Naïve Bayes* menghasilkan *accuracy* sebesar 81% dan AUC-ROC sebesar 0,78. Dengan hasil *performance measure* kedua algoritma di atas, maka algoritma *Random Forest* memiliki performa yang lebih unggul dibandingkan algoritma *Naïve Bayes* pada klasifikasi penyakit jantung.

**Kata kunci:** Klasifikasi, *Random Forest*, *Naïve Bayes*, Penyakit Jantung



## **ABSTRACT**

*Heart disease is a serious health problem and the leading cause of death worldwide. Many cases of death have occurred due to heart disease. Most of these cases occur due to a lack of information about heart disease which makes it difficult for someone to recognize the symptoms, so a way is needed to diagnose heart disease quickly and easily. Classification is one of the techniques in machine learning that can be used to predict the detection of heart disease based on certain symptoms or risk factors. Classification methods that can be used to predict whether a patient has heart disease or not are Random Forest algorithm and Naïve Bayes algorithm. This study aims to determine the best algorithm based on the results of accuracy confusion matrix for the classification of heart disease with Random Forest and Naïve Bayes algorithms. Random Forest algorithm with max\_depth parameter of 6 and ntree of 500 produces accuracy of 83% and AUC-ROC of 0.90. The Naïve Bayes algorithm produces an accuracy of 81% and an AUC-ROC of 0.78. With the results of the performance measure of the two algorithms above, the Random Forest algorithm has superior performance compared to the Naïve Bayes algorithm in the classification of heart disease.*

**Keywords:** *Classification, Random Forest, Naïve Bayes, Heart Disease*

