

ABSTRAK

Felix Hebert (03082180062)

ANALISIS KOMPARASI METODE *NAÏVE BAYES* DAN *RANDOM FOREST* DALAM KLASIFIKASI KUALITAS *WINE*

(xv + 58; 24; 8; 4)

Penelitian ini bertujuan untuk menganalisis akurasi dua metode klasifikasi, yaitu *Naïve Bayes* (NB) dan *Random Forest* (RF), dalam mengklasifikasikan kualitas *wine*. Data yang digunakan meliputi berbagai faktor yang mempengaruhi kualitas *wine*, seperti kandungan asam, pH, dan atribut lainnya. Penelitian ini membandingkan performa kedua metode dalam hal akurasi, presisi, *recall*, dan *misclassification error* dengan menggunakan berbagai skenario pembagian data (70:30, 80:20, dan 90:10). Hasil analisis menunjukkan bahwa *Random Forest* memberikan kinerja yang lebih baik dibandingkan *Naïve Bayes*, dengan akurasi tertinggi mencapai 91,95% pada skenario 90:10, sedangkan *Naïve Bayes* hanya mencapai 84,55%. Selain itu, *Random Forest* juga menunjukkan presisi dan *recall* yang lebih tinggi serta tingkat kesalahan yang lebih rendah dibandingkan *Naïve Bayes*. Berdasarkan hasil ini, dapat disimpulkan bahwa *Random Forest* lebih efektif dalam menangani kompleksitas data kualitas *wine* dan dapat direkomendasikan sebagai metode yang lebih optimal untuk klasifikasi kualitas *wine*.

Kata Kunci: Klasifikasi, Kualitas *Wine*, *Data Mining*, *Naïve Bayes*, *Random Forest*

Referensi: 17

ABSTRACT

Felix Hebert (03082180062)

COMPARATIVE ANALYSIS OF NAÏVE BAYES AND RANDOM FOREST METHODS IN WINE QUALITY CLASSIFICATION

(xv + 58; 24; 8; 4)

This study aims to analyze the accuracy of two classification methods, Naïve Bayes (NB) and Random Forest (RF), in classifying wine quality. The data used includes various factors that influence wine quality, such as acid content, pH, and other attributes. This research compares the performance of both methods in terms of accuracy, precision, recall, and misclassification error using different data splitting scenarios (70:30, 80:20, and 90:10). The analysis results show that Random Forest performs better than Naïve Bayes, with the highest accuracy reaching 91.95% in the 90:10 scenario, while Naïve Bayes only achieves 84.55%. Additionally, Random Forest also shows higher precision and recall, as well as a lower error rate compared to Naïve Bayes. Based on these results, it can be concluded that Random Forest is more effective in handling the complexity of wine quality data and can be recommended as the more optimal method for classifying wine quality.

Keywords: *Classification, Wine Quality, Data Mining, Naive Bayes, Random Forest*

References: 17