

INTISARI

PENERAPAN METODE *ENSEMBLE BOOTSTRAP AGGREGATING* (*BAGGING*) PADA ALGORITMA POHON KEPUTUSAN C5.0

Oleh

Lala Puspita Sari

19/445720/PA/19544

Klasifikasi merupakan salah satu teknik dalam *machine learning* yang digunakan untuk mengelompokkan data ke dalam kelas-kelas. Metode pohon keputusan atau *decision tree* merupakan salah satu metode klasifikasi dalam *machine learning* yang populer karena konstruksinya yang mudah dipahami oleh manusia. Metode pohon keputusan termasuk ke dalam metode yang tidak stabil dibandingkan dengan metode klasifikasi lain dikarenakan perubahan kecil dalam data latih dapat menyebabkan perubahan yang signifikan pada model klasifikasi. Oleh karena itu, metode *bagging* diterapkan pada metode pohon keputusan untuk meningkatkan stabilitas dan meningkatkan performa klasifikasi. Penerapan metode *bagging* memungkinkan pembangunan beberapa model sekaligus melalui sampling dengan pengembalian secara berulang dari data latih dan menggabungkan hasilnya. Pada penelitian ini, metode *bagging* diterapkan pada algoritma pohon keputusan C5.0. Dari hasil analisis dengan menggunakan data *Heart Disease*, diperoleh kesimpulan bahwa penerapan metode *bagging* pada algoritma pohon keputusan C5.0 dapat meningkatkan performa klasifikasi dibandingkan hanya menggunakan algoritma pohon keputusan C5.0 tanpa *bagging* berdasarkan ukuran akurasi, *error rate*, sensitifitas, spesifisitas, presisi, dan *F1-score*.

Kata Kunci: Klasifikasi, Pohon Keputusan C5.0, *Bootstrap Aggregating* (*Bagging*)

ABSTRACT

APPLICATION OF THE ENSEMBLE BOOTSTRAP AGGREGATING (BAGGING) METHOD IN THE DECISION TREE C5.0 ALGORITHM

By

Lala Puspita Sari

19/445720/PA/19544

Classification is one of the machine learning techniques that is used to group data into classes. The decision tree method is a popular classification method in machine learning because its construction is easy for humans to understand. The decision tree method is unstable compared to other classification methods because small changes in the training data can cause significant changes to the classification model. Therefore, the bagging method is applied to the decision tree method to increase stability and improve classification performance. The application of the bagging method allows the development of several models at once through sampling with replacement from training data and combining the results. In this study, the bagging method is applied to the C5.0 decision tree algorithm. From the results of the analysis using Heart Disease data, it is concluded that the application of the bagging method to the C5.0 decision tree algorithm can improve classification performance compared to only using the C5.0 decision tree algorithm without bagging based on accuracy, error rate, sensitivity, specificity, precision, and F1-scores.

Keywords: Classification, Decision Tree C5.0, Bootstrap Aggregating (Bagging)