

INTISARI

PENGEMBANGAN MODEL UNTUK KLASIFIKASI MUTU CENGIH DENGAN PENDEKATAN SELEKSI DAN PEMBOBOTAN FITUR

Oleh

Firman Tempola

(22/495457/SPA/00843)

Berdasarkan standar nasional Indonesia (SNI) 01-3392-1994, mutu cengkih di Indonesia dibagi menjadi 3 kelas yaitu mutu 1, mutu 2, dan mutu 3. Dalam penentuan mutu cengkih tidak mudah karena setiap kelas memiliki fitur yang hampir mirip, selain itu jumlah data yang masih terbatas sehingga proses ekstraksi fitur dilakukan secara manual. Ada 3 fitur yang diekstraksi yaitu fitur warna, bentuk dan tekstur, dengan total jumlah fitur sebanyak 40 fitur untuk dataset pertama dan 35 fitur untuk dataset kedua. Proses ekstraksi fitur secara manual mengakibatkan ada fitur yang tidak relevan sehingga perlu dilakukan seleksi fitur.

Penelitian ini mengusulkan *ensemble feature selection*. *Ensemble feature selection* sangat penting untuk meningkatkan kinerja model *machine learning* seperti *k-NN*, *Naive Bayes decision tree*, *Support vector machine* dan *multilayer perceptron* dalam klasifikasi mutu cengkih menurut standar SNI 01-3392-1994. *Ensemble feature selection* cara kerjanya dengan memberikan skor pada setiap fitur. Terdapat lima metode *feature selection* yang digunakan untuk memproses *ensemble feature selection*. Fitur-fitur yang dipilih bersifat konsisten dan stabil karena selalu dipilih dalam setiap metode *feature selection* yang diterapkan. Selain diterapkan *ensemble feature selection* untuk meningkatkan *performance machine learning* dalam penelitian ini juga diterapkan proses *feature weighting* dengan menggunakan modifikasi *ant lion optimization*.

Hasil penelitian menunjukkan bahwa kombinasi setiap fitur mempengaruhi kinerja dari model pembelajaran mesin, proses *ensemble feature selection* dapat meningkatkan model pembelajaran mesin seperti *k-nn*, *naive bayes* dan *decision tree*, sedangkan *support vector machine* dan *multilayer perceptron* akurasi mengalami penurunan. Selanjutnya saat diterapkan *feature weighting* akurasi dari model *machine learning* juga mengalami peningkatan bahkan lebih baik dari diterapkan *ensemble feature selection*, dimana akurasi tertinggi yaitu saat diterapkan *k-NN ALO* dengan akurasi yang dihasilkan dalam klasifikasi mutu cengkih mencapai 96,29%.

Kata Kunci: Mutu Cengkih, *Machine learning*, *Feature Selection*, *Feature Weighting*

ABSTRACT

DEVELOPMENT OF A CLOVE QUALITY CLASSIFICATION MODEL USING A FEATURE SELECTION AND WEIGHTING APPROACH

By:

Firman Tempola

(22/495457/SPA/00843)

Based on the Indonesian National Standard (SNI) 01-3392-1994, clove quality in Indonesia is classified into three grades: Grade 1, Grade 2, and Grade 3. Determining clove quality is challenging because each grade exhibits very similar characteristics. Furthermore, due to a limited dataset, the feature extraction process was conducted manually. Three types of features were extracted—color, shape, and texture—resulting in a total of 40 features for the first dataset and 35 for the second. This manual extraction process resulted in the inclusion of irrelevant features, necessitating a subsequent feature selection step.

This study proposes an ensemble feature selection method to enhance the performance of machine learning models, such as k-NN, Naïve Bayes, decision trees, Support Vector Machine, and Multilayer Perceptron, for clove quality classification according to the SNI standard. The ensemble method operates by assigning a score to each feature, utilizing five different feature selection techniques. The features chosen through this ensemble approach are highly consistent and stable, as they are selected across all applied methods. In addition to ensemble feature selection, this research also implements a feature weighting process using a modified Ant Lion Optimization (ALO) algorithm.

The research results indicate that feature combinations significantly influence the performance of machine learning models. The ensemble feature selection process improved the performance of models such as k-NN, Naive Bayes, and Decision Tree; conversely, the accuracy of Support Vector Machine (SVM) and Multilayer Perceptron (MLP) decreased. Furthermore, the application of feature weighting enhanced model accuracy, yielding even better results than ensemble feature selection. Specifically, the highest accuracy was achieved by the k-NN ALO model in classifying clove quality, reaching 96.29%.

Keywords: Clove Quality, *Machine learning*, *Feature Selection*, *Feature Weighting*