

# Universal prediction model berbasis random forest untuk sistem pengukuran kadar gula: studi kasus 3 varietas pisang musa acuminata = Universal prediction model for sugar content measurement system based on random forest: a case study of 3 varieties of banana musa acuminata

Fitria Indah Astari, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20514091&lokasi=lokal>

---

## Abstrak

Kadar gula merupakan salah satu karakter kualitas buah pisang yang memengaruhi rasa. Berdasarkan kesamaan karakteristik spektral pada varietas pisang ambon, pisang mas, dan pisang cavendish, yang didapat dari hasil akuisisi citra menggunakan kamera hiperspektral VNIR (400 - 1000 nm), telah berhasil dirancang sebuah universal prediction model yang tidak memerlukan perlakuan destruktif untuk memprediksi nilai kadar gula pada 3 varietas pisang *Musa acuminata*: pisang ambon, pisang mas, dan pisang cavendish. Filter Savitzky - Golay mengurangi noise pada hasil spektral dan juga memperkuat sinyal yang berisi informasi penting. Pada model regresi Random Forest dengan fitur panjang gelombang optimal berjumlah 20 menghasilkan nilai RMSE pelatihan (train) dan RMSE pengujian (test) yang kecil dengan masing-masing sebesar 0,67 dan 1,08 serta nilai koefisien korelasi pelatihan ( $R^2$  Train) dan koefisien korelasi pelatihan ( $R^2$  Test) yang besar masing-masing sebesar 0,98, dan 0,94. Hal menunjukkan penggunaan algoritma Random Forest untuk pembuatan universal prediction model memiliki hasil yang baik.

.....Sugar content is one of the quality attribute that affects the taste of bananas. A non-destructive universal prediction model has been successfully designed based on the similarity of the spectral characteristics of the Ambon bananas, Mas bananas and Cavendish bananas, which were obtained from the results of image acquisition using a VNIR hyperspectral camera (400 - 1000 nm) to predict the value of sugar content on 3 banana *Musa acuminata* varieties: Ambon bananas, Mas bananas and Cavendish bananas. The Savitzky - Golay filter reduces noises in the spectral results and also amplifies the signal that contains some important information. Random Forest regression model with 20 features give a small error value with RMSE Train value is 0.67 and RMSE Test value is 1.08, also give a large correlation coefficient value with  $R^2$  Train value is 0.98 and  $R^2$  Test value is 0.94, it shows that the use of the Random Forest algorithm for designing the universal prediction model has good results.