

Komparasi sistem regresi metode ensemble: Studi kasus sistem pengukuran kadar karotenoid daun bisbul (*diospyros discolor willd.*) pada citra VNIR = Comparison of the ensemble method of regression system: A case study of the carotenoid level measurement system for bisbul leaf (*diospyros discolor willd.*) in VNIR image.

Aditya Bariq Ikhsan, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20513127&lokasi=lokal>

Abstrak

Kandungan total karotenoid dalam tumbuhan umumnya diukur menggunakan analisis spektrofotometri, dengan sifatnya yang merusak sampel terdapat batasan yang bisa dilakukan untuk penelitian selanjutnya. Digunakan pencitraan hiperspektral menggabungkan analisis spektral dan spasial yang bersifat tidak merusak sampel. Timbul masalah terutama pada bagian algoritma untuk membuat sistem prediksi pada citra hiperspektral karena diperlukan algoritma dengan akurasi yang tepat dan cepat. Penelitian ini membahas tentang komparasi algoritma pembelajaran mesin metode ensemble dengan menambahkan tuning hyperparameter menggunakan random search dan memanfaatkan seleksi fitur yang dimiliki tiap model untuk meningkatkan performa dan mengurangi waktu latih model prediksi kadar karotenoid pada daun Bisbul. Sistem prediksi menghasilkan performa dasar, random forest dengan semua fitur memiliki RMSE sebesar 38,16, serta R2 sebesar 0,95, dan waktu latih 4,27s, xgboost dengan semua fitur memiliki RMSE sebesar 39,82, serta R2 sebesar 0,95, dan waktu latih 0,83s, lightgbm dengan semua fitur memiliki RMSE sebesar 35,59, serta R2 sebesar 0,96, dan waktu latih 1,73s, catboost dengan semua fitur memiliki RMSE sebesar 31,60, serta R2 sebesar 0,97, dan waktu latih 17,34s. Dengan menggunakan fitur hasil seleksi dan I, performa sistem berhasil ditingkatkan, random forest tuning dengan 30 fitur memiliki RMSE sebesar 34,39, serta R2 sebesar 0,96, dan waktu latih 5,85s, xgboost tuning dengan 120 fitur memiliki RMSE sebesar 33,32, serta R2 sebesar 0,96, dan waktu latih 1,73s, lightgbm tuning dengan 50 fitur memiliki RMSE sebesar 32,24, serta R2 sebesar 0,97, dan waktu latih 0,22s, catboost tuning dengan 40 fitur memiliki RMSE sebesar 28,53, serta R2 sebesar 0,97, dan waktu latih 4,92s. Secara umum Catboot memiliki peningkatan RMSE paling tinggi, lightgbm memiliki peningkatan waktu latih paling tinggi.

.....The total carotenoid content in plants is generally measured using spectrophotometric analysis, with its destructive to the sample there are limitations that can be done for further research. Hyperspectral imaging combining spectral and spatial analysis is used that is not destructive to the sample. Problems arise, especially in the algorithm section to create a prediction system on hyperspectral images because an algorithm with precise and fast accuracy is required. This study discusses the comparisons of machine learning algorithm with the ensemble method by adding hyperparameter tuning using random search and utilizing the feature selection of each model to improve performance and reduce training time for predictive models of carotenoid levels in velvet leaves. The prediction system produces basic performance, random forest with all features has RMSE of 38.16, and R2 of 0.95, and training time of 4.27s, xgboost with all features has RMSE of 39.82, and R2 of 0.95, and training time of 0.83s, lightgbm with all features has an RMSE of 35.59, and R2 of 0.96, and training time of 1.73s, catboost with all features has an RMSE of 31.60, and R2 of 0.97, and training time 17.34s. By using the selected features and I, system performance has been successfully improved, random forest tuning with 30 features has an RMSE of 34.39, and R2 of

0.96, and training time of 5.85s, xgboost tuning with 120 features has an RMSE of 33, 32, and R2 of 0.96, and training time of 1.73s, lightgbm tuning with 50 features has RMSE of 32.24, and R2 of 0.97, and training time of 0.22s, catboost tuning with 40 features has an RMSE of 28.53, and R2 is 0.97, and training time is 4.92s. In general Catboot has the highest increase in RMSE, lightgbm has the highest increase in training time.