UNIVERSITAS
GADJAH MADA
Integration of Low-Cost Multispectral Sensors and Machine Learning for Quantitative Prediction of Physicochemical Properties in Diluted Milk
Khalid Saifullah, Dr.Eng. Ahmad Kusumaatmaja, S.Si., M.Sc.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

INTISARI

INTEGRASI *LOW-COST MULTISPECTRAL SENSOR* DAN *MACHINE LEARNING* UNTUK PREDIKSI KUANTITATIF SIFAT FISIKOKIMIA PADA SUSU YANG DIENCERKAN

Oleh

Khalid Saifullah
24/547324/PPA/06881

Pengenceran air merupakan salah satu bentuk pemalsuan susu yang dapat mengubah karakteristik fisikokimia dan optik susu. Metode konvensional untuk menilai kualitas susu umumnya memerlukan instrumen laboratorium, preparasi sampel, dan tenaga terlatih, sehingga kurang sesuai untuk pengujian cepat dan portabel. Penelitian ini mengembangkan dan mengevaluasi pendekatan penginderaan multispektral berbiaya rendah berbasis sensor AS7265x untuk penilaian kuantitatif pengenceran air terkontrol pada susu komersial. Tiga produk susu UHT komersial disiapkan pada enam tingkat pengenceran, yaitu 0% hingga 25% (v/v). Sensor AS7265x dengan 18 kanal digunakan untuk mengukur respons spektral, sedangkan *total dissolved solids* (TDS), *electrical conductivity* (EC), dan *refractive index* (RI) digunakan sebagai variabel acuan fisikokimia. Tiga transformasi spektral, yaitu REF, ABS, dan KBM, dievaluasi menggunakan model regresi *Random Forest* (RF), *Extra Trees* (ET), dan XGBoost (XGB). *Recursive Feature Elimination* (RFE) digunakan untuk mengurangi prediktor spektral dari 18 menjadi 10 kanal. Kinerja model dievaluasi menggunakan nested cross-validation lima lipatan, dengan RFE dan GridSearchCV dilakukan pada data pelatihan dalam setiap *outer fold*. Validasi dalam produk dan data gabungan menghasilkan kinerja prediksi yang tinggi. ET menunjukkan kinerja tertinggi pada seluruh 16 perbandingan target-validasi. Pada data gabungan, RFE menghasilkan nilai R^2 sebesar 0,9868 untuk pengenceran, 0,9942 untuk TDS, 0,9945 untuk EC, dan 0,9881 untuk RI. RFE secara umum meningkatkan R^2 serta menurunkan RMSE dan MAE dibandingkan model dengan 18 fitur. Namun, validasi lintas produk menunjukkan kinerja eksternal yang lebih rendah pada produk susu yang tidak digunakan dalam pengembangan model, dengan nilai R^2 negatif pada beberapa kombinasi target-produk. Hasil ini menunjukkan bahwa AS7265x dapat memberikan informasi kuantitatif mengenai pengenceran air terkontrol dan perubahan fisikokimia terkait dalam kondisi pengujian, tetapi kemampuan transfer model antarproduk susu komersial masih terbatas.

Kata kunci: Spektroskopi Multikanal, Susu, Pengenceran, *Total Dissolved Solids, Electrical Conductivity, Refractive Index, Recursive Feature Elimination, Machine Learning*

UNIVERSITAS
GADJAH MADA
Integration of Low-Cost Multispectral Sensors and Machine Learning for Quantitative Prediction of Physicochemical Properties in Diluted Milk
Khalid Saifullah, Dr.Eng. Ahmad Kusumaatmaja, S.Si., M.Sc.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

INTEGRATION OF LOW-COST MULTISPECTRAL SENSORS AND MACHINE LEARNING FOR QUANTITATIVE PREDICTION OF PHYSICOCHEMICAL PROPERTIES IN DILUTED MILK

By

Khalid Saifullah
24/547324/PPA/06881

Water dilution is a common form of milk adulteration that alters the physicochemical and optical characteristics of milk. Conventional milk quality assessment generally requires laboratory instruments, sample preparation, and trained personnel, limiting its suitability for rapid and portable testing. This study developed and evaluated a low-cost multispectral sensing approach based on the AS7265x sensor for quantitative assessment of controlled water dilution in commercial milk. Three commercial UHT milk products were prepared at six dilution levels from 0% to 25% (v/v). An 18-channel AS7265x sensor measured the spectral response, while total dissolved solids (TDS), electrical conductivity (EC), and refractive index (RI) were measured as physicochemical reference variables. Three spectral transformations, REF, ABS, and KBM, were evaluated using Random Forest (RF), Extra Trees (ET), and XGBoost (XGB) regression models. Recursive Feature Elimination (RFE) reduced the spectral predictors from 18 to 10 channels. Model performance was evaluated using five-fold nested cross-validation, with RFE and GridSearchCV performed within the outer training data and performance assessed on the held-out outer folds. Within-product and pooled validation produced high predictive performance. ET showed the highest-performing RFE configuration across all 16 target-by-validation comparisons. For the pooled models, RFE achieved R^2 values of 0.9868 for dilution, 0.9942 for TDS, 0.9945 for EC, and 0.9881 for RI. RFE generally increased R^2 and reduced RMSE and MAE compared with the corresponding 18-feature models. However, cross-product validation showed substantially lower external performance on previously unseen milk products, with negative R^2 values for several target-product combinations. These results indicate that AS7265x can provide quantitative information on controlled milk dilution and associated physicochemical changes under the tested conditions, while model transferability across commercial milk products remains limited

Keywords: Multichannel Spectroscopy, Milk, Water Dilution, Total Dissolved Solids, Electrical Conductivity, Refractive Index, Recursive Feature Elimination, Machine Learning