

Pretreatment Spectral Data NIRS Menggunakan Support Vector Regression

Khairunnissa*¹, Arisman Adnan², Syamsudhuha³

^{1,2,3}Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Riau
Email: ¹khairunnissa6890@grad.unri.ac.id, ²arisman.adnan@lecturer.ac.id,
³syamsudhuha@lecturer.ac.id

Abstrak

Kandungan kalium adalah salah satu elemen penting dalam tanah perkebunan kelapa sawit. Teknologi alternatif yang dapat digunakan untuk memberikan informasi terkait pengukuran kandungan pupuk kalium pada tanah kelapa sawit dapat menggunakan teknologi Near Infrared Spectroscopy (NIRS) untuk menggantikan pengujian laboratorium dalam menganalisis. Penelitian ini bertujuan memprediksi kandungan pupuk kalium pada tanah perkebunan kelapa sawit menggunakan metode Support Vector Regression (SVR) pada data NIRS yang merupakan metode cepat dan akurat dalam pengambilan data. Selain itu, penelitian ini bertujuan untuk menentukan metode koreksi spektrum yang paling efektif dan tepat dalam memprediksi kandungan pupuk kalium pada tanah kelapa sawit, dengan menggunakan *pretreatment Standard Normal Variate* (SNV) dan *Multiplicative Spectral Correction* (MSC) serta kombinasi dari keduanya. Penelitian ini menggunakan 100 sampel tanah dengan panjang gelombang 350 nm – 2500 nm. Pengolahan data menggunakan software@R version 4.4.1. Hasil penelitian menunjukkan dengan metode SVR kernel radial basis, menghasilkan metode koreksi terbaik yaitu *Multiplicative Spectral Correction* dengan nilai R^2 sebesar 0.6025 dan RMSE sebesar 0.0201.

Kata kunci: NIRS, Pretreatment, Support Vector Regression

Abstract

Potassium content is one of important element in oil palm plantation soil. Alternative technology that can be used to provide information related to measuring potassium fertilizer content in oil palm soil can use Near Infrared Spectroscopy (NIRS) technology to replace laboratory testing to analyze. The study aims to predict the potassium fertilized content in oil palm plantation soil using the Support Vector Regression (SVR) method on NIRS data which is a fast and accurate method in data collection. In addition, this study aims to determine the most effective and precise spectrum correction method in predicting potassium fertilizer content in oil palm soil, using Standard Normal Variate (SNV) and Multiplicative Spectral Correction (MSC) pretreatment and a combination of both. This study used 100 soil samples with a wavelength of 350 nm – 2500 nm. Data processing using R software@ version 4.4.1. The result showed the prediction of the radial basis kernel SVR method, produced the best correction method in this study, namely Multiplicative Spectral Correction with an R^2 value of 0.6025 and an RMSE of 0.0201.

Keywords: Near Infrared Reflectance Spectroscopy, Pretreatment, Support Vector Regression.

1. Pendahuluan

Perkebunan kelapa sawit merupakan salah satu faktor penggerak perekonomian masyarakat dalam meningkatkan kesejahteraan petani serta pertumbuhan ekonomi daerah dan nasional[1]. Hasil utama dari perkebunan kelapa sawit adalah buah sawit, yang dapat dijadikan sebagai bahan baku untuk mengekstrak *crude palm oil* (CPO), yang selanjutnya diproses menjadi minyak goreng, kosmetik dan produk lainnya yang memiliki nilai ekonomis tinggi[2][3]. Setiap tahun kebutuhan terhadap minyak kelapa sawit mengalami peningkatan, untuk itu pengolahan tanaman dan lahan yang tepat dapat menghasilkan produksi kelapa sawit yang optimal. Salah satu faktor yang berperan dalam meningkatkan produktivitas perkebunan kelapa sawit adalah ketersediaan unsur hara di dalam tanah, terutama Nitrogen (N), Fosfor (P), Kalium (K) dan C-organik. Oleh karena itu, diperlukan informasi mengenai status hara agar pemupukan dapat dilakukan secara tepat sesuai dengan kondisi lahan. Semua unsur hara yang dibutuhkan oleh tanaman kelapa sawit harus tersedia dalam jumlah yang cukup dan seimbang agar mencapai hasil produksi yang optimal[3].

Near Infrared Reflectance Spectroscopy (NIRS) adalah teknologi yang efisien dan cepat untuk mendeteksi kandungan suatu zat. Metode ini adalah teknik non-destruktif yang sangat

potensi dan telah diterapkan dalam berbagai aplikasi, seperti analisis produk farmasi, pemantauan proses industri, serta analisis tanah dalam bidang pertanian[4]. Teknologi NIRS semakin diminati dan banyak digunakan di berbagai disiplin ilmu karena teknik NIRS merupakan pengujian yang cepat dan tidak merusak sampel. Instrumen NIRS terdiri dari sumber cahaya, monokromator untuk memisahkan cahaya, sampel, serta detektor untuk mengatur intensitas cahaya yang melewati sampel pada berbagai panjang gelombang[5].

Data yang diperoleh dari metode NIRS berisi informasi latar belakang dan kebisingan (noise) selain informasi sampel. Untuk mendapatkan kalibrasi yang andal, akurat dan model stabil, sangat diperlukan pra-pemrosesan data spektral sebelum pemodelan. Metode pretreatment yang sering digunakan adalah *Standard Normal Variate* (SNV)[6] dan *Multiplicative Spectral Correction* (MSC)[7].

Data ini selanjutnya akan dianalisis menggunakan metode *Support Vector Regression* (SVR). *Support Vector Machine* (SVM) sering digunakan dalam penelitian karena kemampuannya beradaptasi dengan berbagai jenis data dan memiliki tingkat kesalahan yang rendah. SVM biasanya diterapkan untuk klasifikasi, sedangkan SVR digunakan untuk menangani masalah regresi. Algoritma SVR dikenal dapat menghasilkan prediksi yang akurat karena kemampuannya dalam menangani masalah overfitting. Overfitting terjadi ketika data uji atau pelatihan menghasilkan prediksi dengan akurasi yang hamper sempurna[8]. Metode SVR ini akan diterapkan menggunakan kernel Radial Basis Function (RBF), karena sesuai dengan karakteristik data yang digunakan dan dapat menangani hubungan non-linear pada data [9].

Menurut Kang, et al model SVR memiliki kinerja yang lebih baik dan kapasitas prediksi yang lebih kuat, dan akhirnya dianggap sebagai model analisis kuantitatif yang optimal untuk memprediksi kandungan nitrogen[10]. Guo et al. menjelaskan model terbaik untuk memprediksi bagian C.sinensis adalah SVR yang dikombinasikan dengan SNV sebagai metode *pretreatment*[11]. Sing et al. menjelaskan algoritma SVR pada spektrum NIRS yang dibangun dari hasil klasifikasi sepuluh kali lipat model memiliki akurasi 83% menggunakan pretreatment Standard Normal Variate (SNV) [12].

Oleh karena itu, penelitian ini bertujuan untuk membangun model untuk memprediksi kandungan kalium pada tanah perkebunan kelapa sawit menggunakan metode *support vector regression* yang sebelumnya telah dilakukan *pretreatment* data spectral.

2. Metode Penelitian

Data yang disajikan pada penelitian ini menggunakan data sekunder yang menjelaskan kumpulan data pada spektroskopi inframerah dekat yang digunakan sebagai metode yang cepat dan kuat untuk menentukan atau memprediksi beberapa parameter. Data yang digunakan adalah data NIRS kandungan pupuk kalium pada tanah kelapa sawit untuk variabel terikat (y) dan variabel bebas (x) adalah data spektrum dekat yang diperoleh dan dicatat dalam rentang panjang gelombang 350 nm sampai 2500 nm, dengan tambahan 100 pemindaian tanah per pengukuran sampel yang dihasilkan dari rata-rata setiap sampel tanah dimana pada setiap tanah dilakukan tiga kali perlakuan.

Data tersebut kemudian kita lakukan *pretreatment* data spektral karena terdapat *noise* dalam data. Setelah *pretreatment*, langkah selanjutnya adalah analisis model mengekstrak prediktor guna analisis variabel dependen dan variabel independen. Langkah awal dengan melakukan pembagaim data *training* dan data *testing* dengan proporsi 80% dan 20% masing-masing data. Pemodelan dilakukan dengan metode SVR langsung pada data training hasil dari pretreatment yang sebelumnya kita lakukan dengan kernel radial basis.

2.1 Pretreatment Spectral Data

Data hiperspektral yang dihasilkan mengandung *noise* dan sinyal hamburan yang tidak relevan untuk deteksi, disebabkan oleh berbagai faktor seperti perbedaan bentuk sampel, efek hamburan cahaya, offset garis dasar, perangkat keras system, serta faktor lainnya. Dalam penelitian ini, metode SNV, MSC dan penggabungan kedua metode digunakan untuk *pretreatment* data spektral. MSC adalah metode kemometrik untuk kalibrasi dan pemodelan multipanjang gelombang yang secara efektif dapat menghilangkan pergeseran dan hamburan garis[7], dan mengintensifkan informasi serapan spektral yang terkait dalam kandungan komponen. Dalam pra-pemrosesan MSC, spektrum ideal ditetapkan terlebih dahulu untuk mewujudkan kalibrasi data spektral sampel[10]. Dalam penelitian ini, spectrum rata-rata dari set

kalibrasi diadopsi sebagai spectrum ideal untuk prapemrosesan MSC. Rumus perhitungannya adalah sebagai berikut:

$$x_{ij(MSC)} = \frac{(x_{ij}-a_i)}{b_i} \quad (1)$$

Metode SNV secara efektif menghilangkan interferensi perkalian dari hamburan dan ukuran partikel [6]. Dengan menghitung variasi normal standar spektrum pada setiap panjang gelombang, variasi kemiringan dihilangkan berdasarkan satu sampel. Rumus perhitungannya adalah sebagai berikut:

$$x_{ij(SNV)} = x_{ij} - \frac{\bar{x}_i}{s(x_i)} \quad (2)$$

dimana SNV adalah variasi normal standar individual untuk panjang gelombang, x_{ij} adalah reflektansi spektral di setiap panjang gelombang dari satu sampel, dan \bar{x}_i adalah rata-rata reflektansi spektral dari semua panjang gelombang dari satu sampel.

2.2 Support Vector Regression

Support Vector Regression (SVR) merupakan metode SVM yang digunakan dalam kasus regresi dan data non-linear dengan jumlah input data yang besar dengan output berupa data kontinu. SVR dikembangkan menggunakan pendekatan *machine learning* berbasis statistik dan telah banyak diterapkan untuk menyelesaikan masalah regresi dengan data berdimensi tinggi[13]. Algoritma SVR termasuk dalam kategori algoritma non-linear yang cerdas, dengan hyperparameter seperti ϵ (epsilon), C (nilai penalti), fungsi kernel, serta parameter kernel[14]. Misalkan terdapat l data training, (x_i, y_i) , $i = 1, \dots, l$ dengan data input $x = \{x_1, \dots, x_l\} \subseteq \mathbb{R}^N$ dan l adalah banyak data training. Dengan metode SVR didapat fungsi regresi sebagai berikut:

$$f(x) = w^T \varphi(x) + b \quad (3)$$

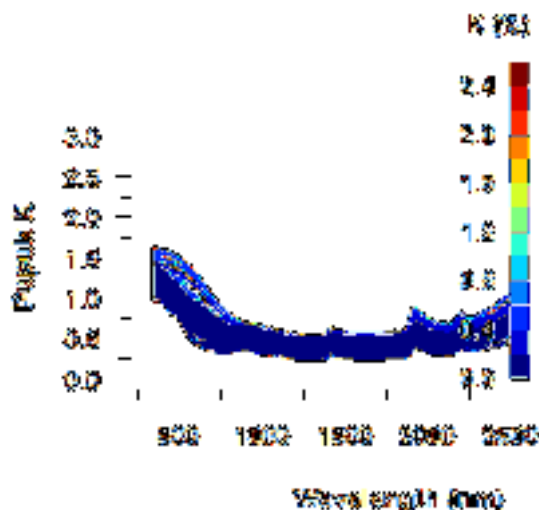
x adalah vektor input yang berupa data, w adalah vektor parameter bobot, dan b adalah bias, $\varphi(x)$ adalah fitur ruang yang mengubah vektor input x melewati ruang dengan dimensi yang lebih tinggi keruang feature φ .

Data yang lebih kompleks biasanya memerlukan fungsi yang lebih ekspresif daripada fungsi linear sederhana. SVR mengubah input non-linear ke dalam ruang fitur berdimensi lebih tinggi, dan proses ini dilakukan oleh kernel. Pada SVR, perhitungan linear dilakukan untuk menemukan hyperplane optimal di ruang fitur tersebut. Kernel memproyeksi data ke ruang fitur berdimensi tinggi guna untuk meningkatkan kapabilitas komputasi dari algoritma pembelajaran linear. Memilih fungsi kernel yang tepat sangat penting karena akan mempengaruhi ruang fitur yang dihasilkan. Dalam penelitian ini, digunakan fungsi kernel radial basis dengan persamaan matematis sebagai berikut:

$$\exp \left(-\gamma \|x_i - x_j\|^2 \right) \quad (4)$$

3.1 Hasil dan Analisis

Penelitian ini menggunakan data spektrum NIRS yang dihasilkan pada panjang gelombang antara 350 nm- 2500 nm. Pada Gambar1, terlihat spektrum yang lebar dengan beberapa puncak dan lembah. Puncak dan lembah dalam spektrum NIRS dipengaruhi oleh kandungan pupuk kalium dalam tanah perkebunan kelapa sawit yang sedang dianalisis. Spektrum mentah dari kandungan kalium tanah kelapa sawit yang ditampilkan menunjukkan puncak-puncak yang terbentuk antara satu spektrum ke spektrum lainnya. Pada spektrum tersebut, terlihat adanya ketebalan yang disebabkan oleh *noise*.

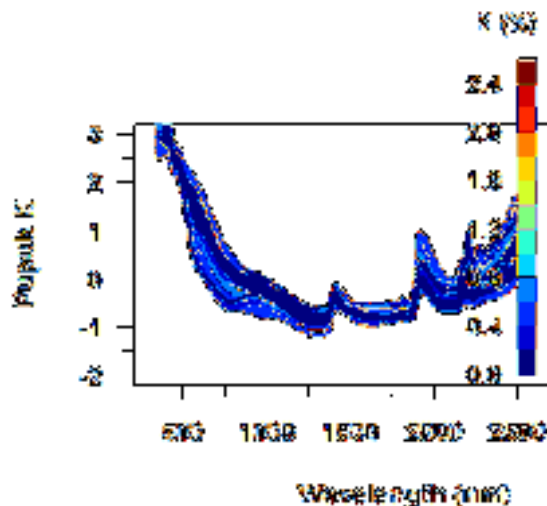


Gambar 1. Panjang Gelombang pada tanah

Data NIRS yang digunakan pada *pretreatment* sebanyak 100 data NIRS yang sudah dirata-ratakan. *Pretreatment* pada spektra NIRS dilakukan untuk memudahkan identifikasi kandungan pupuk kalium tanah perkebunan kelapa sawit pada spectra. Dua jenis *pretreatment* pada spektra NIRS, yaitu *Standard Normal Variate* (SNV) dan *Multiplicative Spectral Correction* (MSC) serta kombinasi keduanya diterapkan dalam penelitian ini. *Pretreatment* pada data spektrum memiliki tujuan yang beragam.

3.1 Pretreatment Standard Normal Variate

Pretreatment SNV, ditunjukkan untuk mengurangi *noise* pada gelombang, menghasilkan spektrum yang lebih halus dan padat[17]. SNV berhasil menghilangkan efek gelombang cahaya dari spektrum, sehingga spektrum kandungan kalium tanah kelapa sawit yang dihasilkan menjadi lebih baik dan lebih jelas dibandingkan dengan spektrum awal yang tidak menggunakan *pretreatment*, seperti yang ditunjukkan pada Gambar 2.

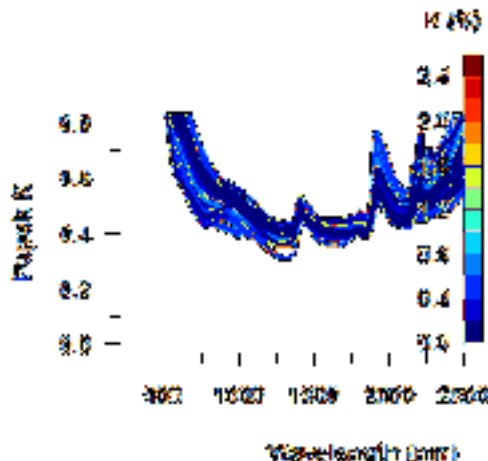


Gambar 2. Pretreatment SNV

3.2 Pretreatment Multiplicative Scatter Correction

MSC adalah salah satu langkah penting dalam pengujian spektrum awal yang berfungsi untuk mengurangi dampak dari hamburan cahaya. MSC menerapkan koreksi spektral perkalian dan penjumlahan pada spektrum asli sesuai dengan rata-rata. Metode ini memisahkan pengaruh hamburan cahaya fisik dari absorbansi cahaya kimia. Setelah diterapkan *pretreatment* MSC,

spektrum asli menunjukkan perbaikan yang signifikan dibandingkan dengan spektrum mentah, karena tingkat kebisingan (*noise*) pada spektrum mentah menurun. Pada Gambar 3 terlihat bahwa spektrum yang telah mendapatkan perlakuan MSC tampak lebih rapat dibandingkan dengan spektrum mentah, dan tingkat kebisingan para *pretreatment* MSC juga terlihat lebih rendah dibandingkan dengan spektrum mentah.



Gambar 3. *Pretreatment* MSC

3.2 Pengembangan Model Kalibrasi SVR

Model Kalibrasi merupakan pengembangan persamaan matematis yang memungkinkan prediksi kandungan pupuk kalium dalam tanah tanpa melakukan pengukuran secara langsung yang dapat merusak sampel karena bersifat destruktif. Proses pengembangan model kalibrasi untuk sampel pupuk kalium pada tanah kelapa sawit memerlukan data acuan dari pengukuran. Data acuan tersebut mencakup jarak antar sampel, rata-rata dan standar deviasi dari pengukuran yang dapat dilihat pada Tabel 1.

Tabel 1. Data Pengembangan Model Kalibrasi

Bagian	Jumlah Sampel	Kandungan Pupuk Kalium	
		Rata-Rata	Standar Deviasi
Kalibrasi	100	0.5687	0.5795

Ketika regresi kalibrasi yang dibangun mampu memberikan prediksi yang akurat maka menunjukkan hasil kalibrasi yang baik. Kualitas prediksi yang baik dapat dilihat dari tingginya nilai koefisien determinasi (R^2) dan rendahnya nilai Root Mean Square Error (RMSE)[16]. Analisis menggunakan NIRS, diperlukan standar error yang rendah daripada standar deviasi dari data awal. Perkiraan kandungan kalium dalam tanah kelapa sawit melalui analisis model kalibrasi SVR dilakukan dengan memanfaatkan panjang gelombang, *pretreatment* SNV, *pretreatment* MSC, serta kombinasi antara *pretreatment* SNV dan MSC. Interpretasi pengembangan model SVR dengan kernel radial basis dalam estimasi kandungan pupuk kalium pada tanah perkebunan kelapa sawit menggunakan NIRS dengan berbagai *pretreatment* dapat dilihat pada Tabel 2.

Tabel 2. Hasil Pengembangan Model Kalibrasi SVR

Perlakuan	Laten	RMSE	R^2
<i>Non pretreatment</i>	80	0.0578	0.6053
SNV	80	0.0186	0.6029
MSC	80	0.0201	0.6025
SNV+MSC	80	0.0152	0.6040

Berdasarkan hasil analisis dari pengembangan model kalibrasi SVR, dapat disimpulkan bahwa semua model prediksi kandungan kalium dalam tanah kelapa sawit menggunakan metode SVR mampu memprediksi kandungan kalium dengan baik. Selain itu, dapat disimpulkan bahwa *pretreatment Multiplicative Scatter Correction* (MSC) merupakan koreksi terbaik, dengan nilai RMSE terendah sebesar 0,0201 dan nilai R^2 tertinggi sebesar 0,6025.

4. Kesimpulan

Penelitian ini membahas penerapan Support Vector Regression (SVR) dengan kernel basis radial untuk memprediksi kandungan pupuk kalium dalam tanah kelapa sawit menggunakan Near Infrared Reflectance Spectroscopy (NIRS). *Pretreatment* dilakukan untuk mengurangi efek *noise* pada data hiperspektral. Hasil evaluasi model menunjukkan nilai yang baik dan mengindikasikan validitas yang tinggi. Metode koreksi terbaik pada penelitian ini adalah *Multiplicative Spectral Correction*, dengan nilai R^2 sebesar 0.6025 dan RMSE sebesar 0.0201. Untuk penelitian selanjutnya, SVR juga dapat dijasikan sebagai opsi bagi peneliti sebagai teknik klasifikasi atau pengelompokan data, selain metode lain seperti Naïve Bayes.

Referensi

- [1] B. Riati and M. Riyadi, "Kesempatan Kerja Dan Kelayakan Ekonomi Usaha Perkebu," *J. Inov. Penelit.*, vol. 27, no. 2, pp. 58–66, 2022, [Online]. Available: <http://117.74.115.107/index.php/jemasi/article/view/537>
- [2] Y. Afrizal, E. Dewi, and Mustain, "Pengolahan crude palm oil (CPO) menjadi minyak sawit merah (MSM) menggunakan filter batuan zeolit, membran keramik dan cartridge filter," *Kinetika*, vol. 13, no. 03, pp. 11–19, 2022, [Online]. Available: <https://jurnal.polsri.ac.id/index.php/kimia/index>
- [3] H. Abdi Hasibuan, L. Afriana, and D. Tamba, "Pengaruh Dosis Bleaching Earth Dan Waktu Pemucatan Crude Palm Oilyang Bervariasi Deterioration of Bleachability Index (Dobi) Terhadap Mutu Produk," *J. Teknol. Ind. Pertan.*, vol. 27, no. 1, pp. 69–75, 2017, doi: 10.24961/j.tek.ind.pert.2017.27.1.69.
- [4] M. Zhu *et al.*, "Fast determination of lipid and protein content in green coffee beans from different origins using NIR spectroscopy and chemometrics," *J. Food Compos. Anal.*, vol. 102, no. May, p. 104055, 2021, doi: 10.1016/j.jfca.2021.104055.
- [5] I. Irfan, I. F. Aji, and D. Yunita, "Application of Near Infrared Spectroscopy (Nirs) in the Measurement of Oil Loss and Its Accuracy Compared To Soxhlet Method," *J. Agroindustri*, vol. 13, no. 2, pp. 107–120, 2023, doi: 10.31186/jagroindustri.13.2.107-120.
- [6] R. J. Barnes, M. S. Dhanoa, and S. J. Lister, "Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra," *Appl. Spectrosc.*, vol. 43, no. 5, pp. 772–777, 1989, doi: 10.1366/0003702894202201.
- [7] D. MacDougall, H. Martens, and P. Geladi, "Linearization and Scatter-Correction for Near-Infrared Reflectance Spectra of Meat," *Appl. Spectrosc.*, vol. 39, no. 3, pp. 491–500, 1985.
- [8] Z. Rais, "Analisis Support Vector Regression (Svr) Dengan Kernel Radial Basis Function (Rbf) Untuk Memprediksi Laju Inflasi Di Indonesia," *VARIANSI J. Stat. Its Appl. Teach. Res.*, vol. 4, no. 1, pp. 30–38, 2022, doi: 10.35580/variansiunm13.
- [9] R. E. Caraka, H. Yasin, and A. W. Basyiruddin, "Peramalan Crude Palm Oil (CPO) Menggunakan Support Vector Regression Kernel Radial Basis," *J. Mat.*, vol. 7, no. 1, p. 43, 2017, doi: 10.24843/jmat.2017.v07.i01.p81.
- [10] S. Kang, H. Deng, L. Chen, X. Zeng, Y. Liu, and K. Chen, "Rapid identification and quality evaluation of medicinal centipedes in China using near-infrared spectroscopy integrated with support vector machine algorithm," *J. Spectrosc.*, vol. 2019, 2019, doi: 10.1155/2019/9636823.
- [11] H. Guo, X. Yang, R. Jing, Y. Li, F. Tan, and M. Li, "Robust multi-class model constructed for rapid quality control of *Cordyceps sinensis*," *Microchem. J.*, vol. 171, no. July, p. 106825, 2021, doi: 10.1016/j.microc.2021.106825.
- [12] D. Sing *et al.*, "Estimation of Andrographolides and Gradation of *Andrographis paniculata* Leaves Using Near Infrared Spectroscopy Together With Support Vector Machine," *Front. Pharmacol.*, vol. 12, no. May, pp. 1–8, 2021, doi: 10.3389/fphar.2021.629833.
- [13] A. Widiarni and M. Mustakim, "Penerapan Algoritma Support Vector Regression dalam Memprediksi Produksi dan Produktivitas Kelapa Sawit," *J. Media Inform. Budidarma*, vol. 7, no. 2, pp. 864–872, 2023, doi: 10.30865/mib.v7i2.6089.
- [14] T. O. Owolabi and M. A. Abd Rahman, "Modeling the optical properties of a polyvinyl alcohol-based composite using a particle swarm optimized support vector regression algorithm," *Polymers (Basel)*, vol. 13, no. 16, 2021, doi: 10.3390/polym13162697.
- [15] F. A. Ngwabebhoh *et al.*, "Preparation and characterization of nonwoven fibrous biocomposites for footwear components," 2020. doi: 10.3390/polym12123016.
- [16] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is

- more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput. Sci.*, vol. 7, pp. 1–24, 2021, doi: 10.7717/PEERJ-CS.623.
- [17] N. Nurhasanah, K. Siregar, and Z. Zulfahrizal, "Prediksi Kadar Air Beras Menggunakan NIRS dengan Metode PLS dan Pretreatment Standard Normal Variate, Derivative-I, Savitzky Golay Smoothing," *J. Ilm. Mhs. Pertan.*, vol. 4, no. 1, pp. 628–637, 2019, doi: 10.17969/jimfp.v4i1.9826.