



**OPTIMASI PROSES PENYANGRAIAN KAKAO BERBASIS
TEKNOLOGI MINDSPHERE DI UGM COCOA TEACHING AND
LEARNING INDUSTRY, BATANG, JAWA TENGAH**

Enas Duhri Kusuma¹, Addin Suwastono¹, Sujoko Sumaryono¹, Novan Wega Wardhana²

¹Departemen Teknik Elektro dan Teknologi Informasi, Fakultas Teknik, UGM

²Cocoa Teaching and Learning Industry UGM

¹enas@ugm.ac.id*

ABSTRAK

Proses penyangraian adalah salah satu tahap penting dalam pengolahan biji kakao yang menentukan kualitas hasil akhir produk. Penelitian ini bertujuan untuk mengoptimalkan suhu dan durasi penyangraian berbasis data untuk menghasilkan kualitas biji kakao terbaik. Pada proses ini, kandungan air pada biji kakao dikeluarkan dan biji dikembangkan agar didapatkan aroma dan warna yang khas serta memenuhi standar. Data diperoleh dari perekaman secara real-time menggunakan platform IoT Mindsphere yang ada di CTLI dan bagian quality control, mencakup kapasitas roasting, kadar air, pH, jenis biji, dan variabel lainnya. Metode data mining dengan algoritma Support Vector Regression (SVR) digunakan untuk memprediksi suhu dan durasi optimal, yang dievaluasi menggunakan Mean Absolute Percent Error (MAPE) dan Root Mean Squared Error (RMSE). Hasil penelitian menunjukkan bahwa algoritma SVR mampu merekomendasikan kombinasi suhu dan durasi penyangraian dengan akurasi tinggi, dengan nilai rata-rata MAPE sebesar 4,76% dan RMSE sebesar 9,17. Penelitian ini bermanfaat bagi industri kakao sebagai panduan berbasis data untuk meningkatkan efisiensi dan kualitas produk olahan kakao.

Kata Kunci : Kakao, optimasi, IoT, CTLI, Mindsphere. *data mining*

ABSTRACT

The roasting process is a crucial step in cocoa bean processing that determines the final product quality. This study aims to optimize the roasting temperature and duration using data-driven approaches to achieve the best cocoa bean quality. During this process, the water content in cocoa beans is removed, and the beans are developed to achieve distinctive aroma and color that meet standards. Data is collected through real-time recording using the IoT Mindsphere platform at CTLI and quality control, including roasting capacity, water content, pH, bean type, and other variables. The data mining method with the Support Vector Regression (SVR) algorithm is employed to predict optimal temperature and duration, evaluated by Mean Absolute Percent Error (MAPE) and Root Mean Squared Error (RMSE). The results show that the SVR algorithm can recommend optimal roasting conditions with high accuracy, achieving an average MAPE of 4.76% and RMSE of 9.17. This study benefits the cocoa industry by providing a data-driven guide to enhance efficiency and product quality.



Lisensi

Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.

Keywords: Cocoa, optimization, IoT, CTLL, Mindsphere, data mining

1. Pendahuluan

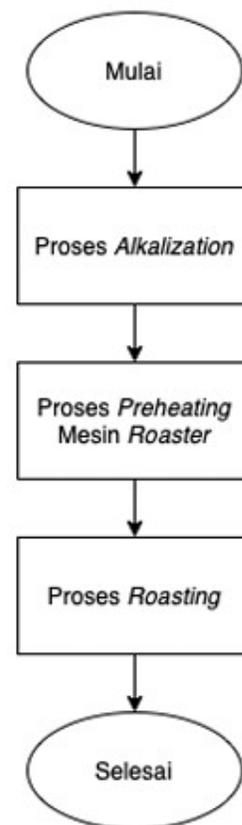
Indonesia merupakan salah satu negara penghasil kakao terbesar di dunia. Tercatat pada tahun 2020, Indonesia mampu menghasilkan 659,7 ribu ton kakao [1] dengan nilai ekspor mencapai US\$1,21 miliar [2]. Hal ini tentunya menjadikan kakao sebagai salah satu potensi komoditas perkebunan unggulan di Indonesia. Menurut data yang diambil dari situs web Kementerian Pertanian Republik Indonesia, pada tahun 2021, luas perkebunan kakao mencapai 1.497.467 Ha [3]. Selain menjadi komoditas perkebunan unggulan di Indonesia, berdasarkan hal tersebut kakao juga dapat menjadi sumber lapangan kerja bagi masyarakat sekitar perkebunan kakao.

Universitas Gadjah Mada melalui UGM Cocoa Teaching and Learning Industry ikut andil dalam menghasilkan produk kakao di Indonesia. UGM Cocoa Teaching and Learning Industry bertujuan untuk mendorong dan mempercepat program hilirisasi industri pengolahan kakao dan sebagai wahana produktif berbasis riset dan inovasi untuk mendukung proses pembelajaran yang bersinergi dengan industri. Industri coklat yang terletak di Kabupaten Batang tersebut merupakan hasil kerjasama dari berbagai pihak yaitu dari Universitas Gadjah Mada, Kementerian Perindustrian, Dikti, dan Pemerintah Kabupaten Batang. Industri ini sangat unik karena berlokasi di tengah-tengah perkebunan kakao warga [4].



Gambar 1. Nibs Biji Kakao

Terdapat dua produk akhir yang dihasilkan pada UGM Cocoa Teaching and Learning Industry yaitu butter dan coklat bubuk. Sebelum menjadi butter dan coklat bubuk, tentunya biji kakao akan melalui beberapa proses produksi. Salah satunya adalah proses roasting atau penyangraian. Proses roasting adalah proses mengubah biji kakao menjadi *nibs* (daging biji) seperti ditunjukkan pada Gambar 1, mengeluarkan kandungan air dari *nibs* biji kakao, mengeringkannya, serta mengembangkan biji tersebut agar mendapatkan aroma dan warna yang khas. Visualisasi proses sangrai kakao ditampilkan pada diagram alir Gambar 2.



Gambar 2. Diagram alir proses sangrai

Variabel yang berpengaruh dalam proses roasting adalah waktu dan suhu yang diatur dalam proses roasting [5]. Agar hasil roasting bagus, maka suhu dan waktu roasting harus

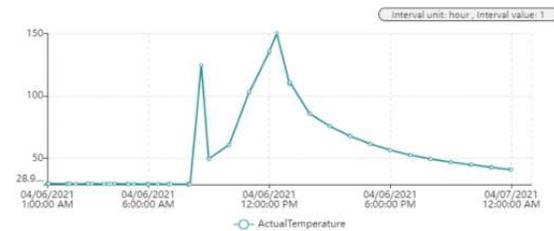
optimal. Pada penelitian ini, dengan menggunakan algoritma *support vector regression* (SVR), dibuat model yang dapat digunakan untuk merekomendasikan suhu dan durasi roasting berdasarkan data-data pada proses roasting seperti kapasitas roasting, kadar air, pH, jenis biji dan variabel lain. Penulis membandingkan tiga algoritma data mining, yaitu Support Vector Regression (SVR), Multiple Linear Regression (MLR), Extreme Learning Machine (ELM), dan Particle Swarm Optimization-Extreme Learning Machine (PSO-ELM) pada penelitian ini. Algoritma data mining tersebut dievaluasi menggunakan ukuran standar regresi seperti MAPE dan RMSE untuk nantinya akan dibandingkan performanya. Dengan demikian, diharapkan operator dapat menentukan suhu dan durasi roasting yang optimal agar hasil roasting yang dihasilkan baik, sehingga kualitas produksi dari pengolahan biji kakao memiliki kualitas yang tinggi.

2. Metode Penelitian

A. Dataset

Dataset yang dipakai pada artikel ini dikumpulkan melalui platform IoT *Mindsphere* yang dilanggan oleh CTLI. *Mindsphere* adalah platform *Internet of Things* (IoT) berbasis *cloud* dari Siemens yang dirancang untuk mengumpulkan, menganalisis, dan memanfaatkan data dari berbagai perangkat dan sistem industri. Platform ini memungkinkan perusahaan untuk menghubungkan aset fisik—seperti mesin dan sensor—ke dunia digital untuk mengoptimalkan operasional dan mempercepat transformasi digital. *Mindsphere* dari Siemens adalah solusi IoT industri yang kuat untuk menghubungkan aset fisik dengan teknologi digital. Dengan memanfaatkan analitik dan data real-time, platform ini membantu perusahaan meningkatkan efisiensi, mengurangi biaya,

dan berinovasi lebih cepat di era Industri 4.0. Pada *Mindsphere*, dapat diperoleh profil suhu yang merupakan fungsi waktu selama menyangrai, sebagaimana ditampilkan pada Gambar 3.



Gambar 3. Profil suhu penyangraian terhadap waktu, direkam menggunakan *Mindsphere*

Dataset yang digunakan adalah dataset pemantauan produksi dari tim laboratorium UGM Cocoa Teaching and Learning Industry serta data suhu roaster dari *MindSphere* mulai bulan April 2021 sampai bulan Juli 2022. Terdapat 43 baris data pada dataset tersebut. Data yang digunakan mempunyai beberapa variabel sebagaimana tercantum pada Tabel 1.

Pada artikel ini, model yang dikembangkan memiliki variabel dependen yaitu suhu dan variabel independent yaitu variabel lainnya selain variabel suhu.

TABEL 1. KAMUS DATA PENELITIAN

Kolom	Keterangan
nibs_capacity	kapasitas nibs (kg)
solution_load	jumlah air yang dibutuhkan (L)
beans_source	wilayah kebun dari biji tersebut
is_alkalized	alkilasi atau tidak
product_type	tipe produk
durasi_roasting	durasi <i>roasting</i>
suhu	suhu pengaturan
pH_0	pH awal biji kakao
pH_N	pH akhir biji kakao
moist_0	kadar air (%) awal biji kakao
moist_N	kadar air (%) akhir biji kakao

B. Alur Metodologis

Terdapat beberapa tahapan yang ditempatkan pada penelitian ini, yaitu:

- a. Mengoleksi Data
- b. Praproses Data
- c. Seleksi Fitur
- d. Membagi Data Latih dan Uji
- e. Melatih Model menggunakan Data Latih
- f. Menguji Model dengan Data Uji
- g. Membandingkan Model Data Mining
- h. Mendapatkan Model yang Terbaik

C. Praproses Data

Praproses data dilakukan untuk membersihkan data, serta untuk memastikan data siap atau layak di-inputkan ke dalam model. Beberapa tahap yang dilakukan pada pra proses data pada penelitian ini adalah sebagai berikut:

1) *One-Hot Encoding*

One-Hot Encoding digunakan untuk kolom kategorik, yaitu `beans_source` dan `product_type`. *One-Hot Encoding* adalah skema pengkodean yang paling banyak digunakan. *One-Hot Encoding* membandingkan setiap tingkat variabel kategori dengan tingkat referensi tetap. Satu hot encoding mengubah satu variabel dengan n observasi dan d nilai berbeda, menjadi d variabel biner dengan masing-masing n observasi. Setiap pengamatan menunjukkan ada (1) atau tidak adanya (0) dari variabel biner dikotomis [6].

2) *MinMax Normalization*

MinMax Normalization digunakan untuk melakukan normalisasi pada data sehingga memiliki skala yang sama (0 sampai 1), hal ini diharapkan mengurangi bias pada pemodelan. *MinMax Normalization* adalah metode normalisasi menggunakan transformasi linear pada data asli sehingga menghasilkan data baru yang variasinya seimbang antara satu kolom dengan kolom lainnya [7]. *MinMax*

Normalization dapat ditulis dengan perumusan sebagaimana berikut :

$$X_{new} = \frac{X_{old} - \text{Min}(X_{old})}{\text{Max}(X_{old}) - \text{Min}(X_{old})} \quad (1)$$

D. Seleksi Fitur

Seleksi fitur adalah salah satu tahapan dalam proses data mining untuk memilih variabel yang relevan dengan variabel target dan membuang variabel yang tidak diperlukan atau redundan. Tujuan dari seleksi fitur adalah untuk meningkatkan performa, mempercepat waktu, dan mengurangi beban pada proses prediksi. Terdapat 3 metode populer yang digunakan dalam seleksi fitur yaitu *embedded*, *filter*, dan *wrapper* [8]. Seleksi fitur yang digunakan pada penelitian ini adalah seleksi fitur berbasis *filter*. Teknik seleksi fitur berbasis *filter* menggunakan metode statistik seperti kemiripan, ketergantungan, informasi, jarak untuk menunjukkan ketergantungan atau korelasi penting antara variabel input dan target [9]. Apabila variabel target numerik dan variabel input numerik, korelasi *pearson* dapat digunakan. Jika variabel target numerik dan variabel input kategorik maka dapat digunakan ANOVA F-test [10].

E. Support Vector Regression (SVR)

Model pertama yang digunakan pada penelitian ini adalah SVR. SVR adalah penerapan algoritma Support Vector Machine untuk kasus regresi. Dalam kasus regresi, output adalah bilangan real atau kontinu. SVR adalah metode yang dapat mengatasi *overfitting*. Sehingga akan menghasilkan kinerja yang baik. Misalnya N adalah data latih (X, y) dengan menggunakan SVR, pengguna dapat menentukan fungsi $f(X)$. Fungsi tersebut memiliki deviasi terbesar E dari target sebenarnya untuk semua data pelatihan. Jika nilai E sama dengan 0 maka regresi dianggap sempurna. SVR disini bertujuan untuk menemukan fungsi regresi $f(X)$ yang dapat mendekati output ke target aktual, dengan toleransi kesalahan E dan kompleksitas minimal. Fungsi regresi $f(X)$ dapat dituliskan sebagai berikut.

$$f(X) = w^T \varphi(X) + b \quad (2)$$

$\varphi(X)$ menunjukkan suatu titik dalam ruang fitur berdimensi lebih tinggi dan hasil pemetaan input vektor X dalam ruang fitur berdimensi lebih rendah. Koefisien w dan b akan diestimasi menggunakan persamaan [11] :

$$\min \frac{1}{2} \|w\|^2 + C \frac{1}{N} \sum_{i=1}^N L_E(y_i, f(X_i)) \quad (3)$$

$$y_i - w\varphi(X_i) - b \leq E \quad (4)$$

$$w\varphi(X_i) - y_i + b \leq E, i = 1, 2, 3, \dots, N$$

dengan,

$$L_E(y_i, f(X_i)) = |y_i - f(X_i)| - E|y_i - f(X_i)| \quad (5)$$

Pada umumnya, terdapat 3 kernel yang digunakan pada SVR atau SVM. Berikut adalah 3 kernel yang biasa digunakan pada SVR [11] :

- Linear Kernel

$$k(x, y) = x^T y + C \quad (6)$$

- Polynomial Kernel

$$k(x, y) = (ax^T y + C)^d \quad (7)$$

- Radial Basis Function (RBF) Kernel

$$k(x, y) = \exp\left(-\gamma \|x - y\|^2\right)^d \quad (8)$$

Langkah pertama yang dilakukan pada pemodelan menggunakan model SVR adalah melakukan *hyperparameter tuning* menggunakan metode *grid search* untuk mencari parameter SVM yang optimal. *Grid search* menguji semua kombinasi dari hyperparameter yang diberikan pada konfigurasi model machine learning [12]. *Grid search* membagi rentang parameter yang akan dioptimalkan ke dalam grid dan melintasi semua titik untuk mendapatkan parameter yang optimal. *Grid search* mengoptimalkan parameter SVM (dalam kasus ini SVR) menggunakan teknik validasi silang sebagai metrik kinerja [13]. *Grid search* dapat diimplementasikan menggunakan library *sklearn* dengan menggunakan perintah *GridSearchCV*. Beberapa keuntungan menggunakan *grid search* sebagai metode optimasi adalah penerapannya mudah, dapat menemukan λ yang jauh lebih baik daripada pengoptimalan sekuensial manual, dan keandalan dan dimensinya yang rendah [13].

3. Hasil dan Pembahasan

A. Seleksi Fitur

Untuk variabel numerik, seleksi fitur dilakukan menggunakan metode statistik pearson sementara untuk variabel kategorik, seleksi fitur dilakukan menggunakan metode statistika ANOVA F-score, pada kedua metode tersebut dilakukan uji hipotesis dengan melihat p-value. Penelitian ini menggunakan tingkat kepercayaan 0.01 sehingga apabila p-value < 0.01 maka hipotesis 0 ditolak dan hipotesis 1 diterima. Berikut adalah pernyataan dari hipotesis 0 dan hipotesis 1 :

- H0 : variabel bebas x tidak memiliki hubungan dengan variabel target y (suhu)
- H1 : variabel bebas x memiliki hubungan dengan variabel target y (suhu)

TABEL 2. HASIL UJI STATISTIK VARIABEL NUMERIK

Variabel	r_suhu	pvalue_suhu
nibs_capacity	-0.5089	0.0009
solution_load	-0.2962	0.0671
pH_0	0.0392	0.8125
pH_N	0.1251	0.4480
delta_pH	-0.1771	0.2807
moist_0	-0.0248	0.8809
moist_N	-0.5860	0.0001
delta_moist	0.2155	0.1876

Berdasarkan uji statistik variabel numerik pada Tabel 2 untuk variabel dependen y yaitu suhu, terdapat dua variabel independen yang secara signifikan berhubungan dengan variabel dependen suhu yaitu **nibs_capacity** dan **moist_N**. Dengan demikian, variabel numerik yang akan dipilih pada pemodelan menggunakan algoritma data mining adalah **nibs_capacity** dan **moist_N** serta **durasi_roasting**, karena penelitian ini juga digunakan untuk mengoptimalisasi suhu dan **durasi_roasting**. untuk mengoptimalisasi

durasi roasting dan suhu roasting, maka durasi dan suhu juga dihubungkan pada model.

Tabel 3. HASIL UJI STATISTIK VARIABEL KATEGORIK

Variabel	r_suhu	pvalue_suhu
beans_source	29.522289	0.000004
product_type	0.512261	0.478655
is_alkalized	6.490727	0.015131

Berdasarkan uji statistik pada variabel kategorik pada Tabel 3, didapati bahwa terdapat satu variabel independen yang secara signifikan mempunyai hubungan dengan variabel dependen suhu yaitu *beans_source*. Dengan demikian, **variabel kategorik yang akan digunakan** pada tahap pemodelan adalah variabel *beans_source* saja. Dari sekian banyak variabel yang ada pada data yang tersedia, variabel independen yang mempunyai tingkat hubungan yang tinggi dengan variabel suhu adalah sebagai berikut : **nibs_capacity**, **moist_N**, **durasi_roasting** (untuk mengoptimalisasi durasi *roasting* dan suhu *roasting*, maka durasi dan suhu dihubungkan pada model), **beans_source**. Variabel tersebut akan dijadikan variabel independen pada model dan yang akan bertindak sebagai variabel dependen atau variable target adalah suhu.

B. Pemodelan menggunakan SVR

Tahap pertama yang dilakukan adalah mencari *hyperparameter* optimal menggunakan *grid search*. *Hyperparameter* yang dicari adalah C, gamma, epsilon, dan kernel. Berikut adalah *range* atau kumpulan nilai *hyperparameter* yang akan dicari nilai optimalnya :

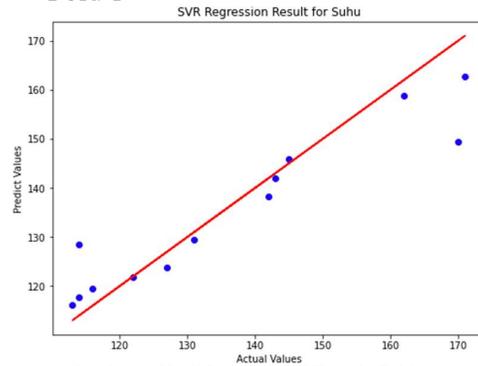
parameter = {'C': [0.1, 0.001, 1, 10, 12, 14, 16, 18, 20, 22],
 'gamma': [0.001, 0.01, 0.1, 1, 2, 5],
 'epsilon': [0.001, 0.01, 0.1, 1, 2, 4],
 'kernel': ("rbf", "poly", "linear")}

Setelah dilakukan pencarian *hyperparameter* yang optimal menggunakan *grid search* dengan 3-fold validasi silang, didapat *hyperparameter* yang optimal sebagai berikut:

{'C': 22, 'epsilon': 1, 'gamma': 1, 'kernel': 'rbf'}

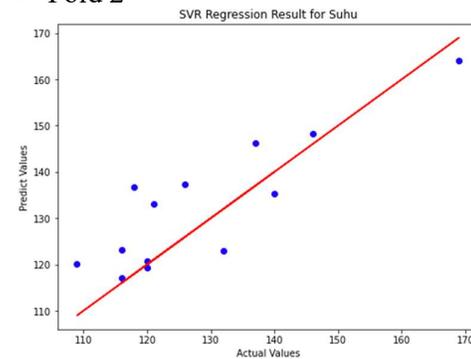
Skor rata-rata RMSE terbaik yang didapat dengan parameter tersebut sebesar 9.17 dan untuk skor rata-rata MAPE sebesar 4.76%. Berikut adalah perbandingan nilai prediksi dan aktual dari model tersebut untuk setiap *fold*-nya :

• Fold 1



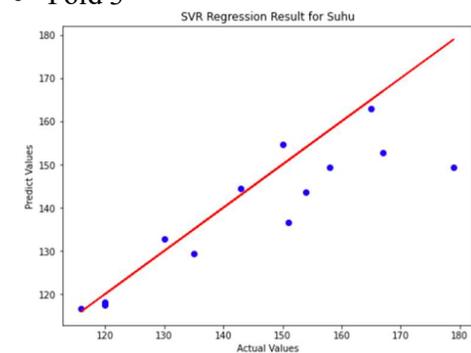
Gambar 1. Hasil Pengujian SVR pada Fold 1

• Fold 2



Gambar 2. Hasil Pengujian SVR pada Fold 2

• Fold 3



Gambar 3. Hasil Pengujian SVR pada Fold 3

Gambar 4, 5, dan 6 menunjukkan perbandingan nilai prediksi dan aktual pada data tes untuk setiap fold pada pemodelan suhu menggunakan SVR. Untuk fold pertama, didapat nilai MAPE sebesar 3.71% dan nilai RMSE sebesar 7.72, pada fold kedua didapat

nilai MAPE sebesar 5.73% dan nilai RMSE sebesar 8.85, dan pada fold ketiga didapat nilai MAPE sebesar 4.84% dan nilai RMSE sebesar 10.94. Apabila dirata-ratakan, pada pemodelan menggunakan SVR didapatkan rata-rata dan standar deviasi MAPE sebesar 4.76 (1.01) dan RMSE sebesar 9.17 (1.63). Berikut adalah tabel lengkap dari hasil pengujian model tersebut:

TABEL 4. HASIL PENGUJIAN MODEL SVR

SVR		
Ukuran	MAPE	RMSE
Fold 1	3.71	7.72
Fold 2	5.73	8.85
Fold 3	4.84	10.94
Rata-rata	4.76	9.17
Std Dev	1.01	1.63

4. Kesimpulan

Suhu optimal dan durasi roasting nibs biji kakao dipengaruhi oleh variabel numerik, yang didasarkan pada data kuantitatif, dan variabel kategorik yang berasal dari data kualitatif. Berdasarkan analisis data penyangraian nibs kakao di CTLI, terdapat dua variabel numerik independen yang secara signifikan berhubungan dengan suhu yaitu **kapasitas nibs** dan **kadar air nibs**. Adapun berdasarkan uji statistik pada variabel kategorik, didapati bahwa terdapat satu variabel kategorik independen yang secara signifikan mempunyai hubungan dengan variabel dependen suhu adalah **tempat asal biji**. Algoritma data *mining* SVR dapat digunakan untuk merekomendasikan suhu dan durasi *roasting* pada proses *roasting* biji kakao dengan rata-rata dan standar deviasi MAPE sebesar 4.76 (1.01) dan RMSE sebesar 9.17 (1.63).

Untuk mendapatkan hasil yang lebih komprehensif, disarankan untuk menguji algoritma data mining lainnya, seperti Random Forest Regression, Neural Networks, atau Gradient Boosting, untuk membandingkan performa prediksi suhu dan

durasi roasting. Selain itu, penelitian juga dapat diperluas dengan mempertimbangkan lebih banyak variabel, seperti komposisi kimia biji kakao, kondisi penyimpanan, dan metode pasca-panen. Penelitian masa depan juga dapat mengintegrasikan metode optimasi multi-objektif untuk mengevaluasi hubungan antara suhu roasting, durasi, dan kualitas produk akhir berdasarkan parameter organoleptik seperti rasa dan tekstur.

5. Daftar Pustaka

- [1] Aziziah, Sabila Aulia, and Nyoman Djinar Setiawina. "Analisis Pengaruh Produksi, Harga Dan Nilai Tukar Terhadap Ekspor Biji Kakao Indonesia Ke Belanda." *Cerdika: Jurnal Ilmiah Indonesia* 1.4 (2021): 448-455.
- [2] Izzah, Nidaul, and Dennysa Damayanti. "Pengaruh Jumlah Produksi dan Harga terhadap Nilai Ekspor Kakao Indonesia Tahun 2017-2020." *Transparansi: Jurnal Ilmiah Ilmu Administrasi* 6.1 (2023): 78-85.
- [3] Mulyo, Panca Rahadi, and Yuli Hariyati. "Dinamika perkembangan perkebunan kakao rakyat di Indonesia." *Agriekonomika* 9.1 (2020): 48-60.
- [4] Nabila, Salsa Olivia, and Nur Wijayanti. "Study of Cocoa Powder Production Process and Packaging Technology in UGM Cocoa Teaching and Learning Industry." *Indonesian Journal of Food Technology* 3.1 (2024): 87-100.
- [5] S. Wijanarti, A. M. Rahmatika, and R. Hardiyanti, "Pengaruh lama penyangraian Manual Terhadap Karakteristik Kakao Bubuk," *Jurnal Nasional Teknologi Terapan (JNTT)*, vol. 2, no. 2, p. 212, 2019

- [6] K. Potdar, T. S., and C. D., "A comparative study of categorical variable encoding techniques for neural network classifiers," *International Journal of Computer Applications*, vol. 175, no. 4, pp. 7–9, 2017
- [7] O. A. Akanbi, I. S. Amiri, and E. Fazeldehkordi, *A machine learning approach to phishing detection and Defense*. Amsterdam: Elsevier, 2015
- [8] F. F. Firdaus, H. A. Nugroho, and I. Soesanti, "A review of feature selection and classification approaches for heart disease prediction," *IJITEE (International Journal of Information Technology and Electrical Engineering)*, vol. 4, no. 3, p. 75, 2021
- [9] N. Sánchez-Marroño, A. Alonso-Betanzos, and M. Tombilla-Sanromán, "Filter methods for feature selection – A comparative study," *Intelligent Data Engineering and Automated Learning - IDEAL 2007*, pp. 178–187, 2007
- [10] Brownlee, Jason. "How to choose a feature selection method for machine learning." *Machine Learning Mastery* 10 (2019): 1-7
- [11] A. Shirzad, M. Tabesh, and R. Farmani, "Performance Comparison between Support Vector Regression and Artificial Neural Network for Prediction of Oil Palm Production," *Jurnal Ilmu Komputer dan Informasi (Journal of Computer Science and Information)*, vol. 9, no. 1, pp. 1-8, 2016.
- [12] D. M. Belete and M. D. Huchaiah, "Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results," *International Journal of Computers and Applications*, vol. 44, no. 9, pp. 875–886, 2021.
- [13] Sulistiana and M. A. Muslim, "Support Vector Machine (SVM) optimization using grid search and UNIGRAM to improve e-commerce review accuracy," *Journal of Soft Computing Exploration*, vol. 1, no. 1, 2020.