Inform : Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi
Vol.9 No.2 July 2024, P-ISSN : 2502-3470, E-ISSN : 2581-0367

Classification of Crude Palm Oil Quality Using Artificial Neural Networks Based on Chemical Components

Dimas Kurniawan[1], Armansyah[2]

[1,2]*Computer Science Department, Universitas Islam Negeri Sumatera Utara, Indonesia*
[1]dimaskurniawan40@uinsu.ac.id (*)
[2]armansyah@uinsu.ac.id

Received: 2024-06-13; Accepted: 2024-07-27; Published: 2024-07-31

Abstract— Palm oil, known as Crude Palm Oil (CPO), is a flagship product in the palm oil industry, playing a crucial role in the economies of various tropical countries such as Indonesia. This study aims to classify the quality of CPO based on chemical components using the Artificial Neural Networks (ANN) method with a backpropagation algorithm. The research data was obtained from PT. Perkebunan Lembah Bhakti (PLB) 2 from January to October 2023, consisting of 225 entries with five main chemical variables: impurity level, moisture content, free fatty acid (FFA) level, Deterioration of Bleachability Index (DOBI), and carotenoids. The data preprocessing stage involved transforming and normalizing the data using the Z-score method. The ANN model used has a 5-5-2 architecture with ReLU activation functions for the input and hidden layers and a Softmax function for the output layer. Model evaluation was conducted using accuracy metrics, which showed that the ANN model could classify CPO quality with an accuracy of 97.78% on the test data. The research results show that ANN can classify CPO quality with a high level of accuracy. This indicates that this method has great potential for use in the industry to improve CPO quality assessment. The improvement in accuracy and validity of the ANN classification results has significant implications for the industry. With high accuracy, ANN can reduce human errors in quality assessment, speed up the process, and increase the consistency of assessment results. This is crucial because consistent quality assessment can enhance operational efficiency and reduce production costs.

Keywords— Crude Palm Oil; Artificial Neural Network; Classification; Chemical Components; Backpropagation; Quality.

I. INTRODUCTION

Palm oil, often referred to as Crude Palm Oil (CPO), stands as a flagship product within the palm oil industry, playing a vital role in the economies of various tropical nations, such as Indonesia. According to the Ministry of Economic Affairs of the Republic of Indonesia in press release No. HM.4.6/82/SET.M.EKON.3/04/2021, the palm oil industry is recognized as one of the commodities contributing significantly to Indonesia's economic development. In 2018, palm oil production amounted to 40.57 million tons of Crude Palm Oil (CPO), sourced from Smallholder Plantations accounting for 16.8 million tons (35%), State-Owned Plantations contributing 2.49 million tons (5%), and Private Large-Scale Plantations yielding 29.39 million tons (60%). Within the palm oil industry, quality serves as a paramount factor in meeting market demand and sustaining product competitiveness in the global market.

The quality of CPO can be assessed based on several chemical and physical parameters [1], such as Free Fatty Acid (FFA), Moisture and Impurities (M&I), Iodine Value (IV), Peroxide Value (PV), Deterioration of Bleachability Index (DOBI), and colour. Each of these parameters is subject to specific standard limits that must be met for CPO to be considered of high quality. According to the Indonesian National Standard (SNI) 01-2901-2006, the maximum quality requirements for CPO are specified as follows: impurity level, free fatty acid content, and moisture content should not exceed 0.5% [2].

PT. PLB 2 produces CPO with a production capacity of 100 tons per day, diligently upholding the quality of its output. Presently, the determination of CPO quality is conducted through traditional means by factory operators, inspecting each quality parameter of CPO individually. This conventional quality assessment method is prone to errors due to the multitude of samples that need testing, as human capacity is limited in handling extensive data. Therefore, the implementation of machine learning technology utilizing artificial neural networks with backpropagation in classifying CPO quality holds the potential to address the limitations of human capability in processing vast data sets [3].

Artificial neural networks have been frequently employed for classification, as evidenced by previous studies utilizing neural network classification models. The research titled "Classification of Rice Quality based on Gas MQ Sensor Array Data using Artificial Neural Network Method based on Arduino" conducted by [4] achieved a classification accuracy of over 80% in classifying rice quality. Subsequent research titled "Classification of Banana Ripeness Levels in RGB Color Space Using Artificial Neural Networks (ANNs)" conducted by [5] demonstrated that artificial neural networks can be utilized as a method for classifying banana ripeness levels with an accuracy of 98.3%. The study developed by [6] under the title "Implementation of Backpropagation Artificial Neural Network in Black Tea Grade Classification" constructed an artificial neural network model for classifying black tea grades using a 4-5-3 architecture. This research yielded promising results with an error rate of 0.0096 after 1000 iterations.

Additionally, during the testing phase, an accuracy of 97.95% was achieved. A study by Erlangga et al. [7] successfully utilized ANNs in conjunction with electronic nose data to

DOI : https://doi.org/10.25139/inform.v9i2.8433

classify rice quality, demonstrating the applicability of this approach to different agricultural products. The classification results using ANN have an accuracy score of 99.84%. The ability of ANNs to learn from data and adapt to changing conditions makes them an attractive tool for improving the efficiency and accuracy of CPO quality assessment.

In this study, researchers will apply artificial neural network methods to classify the quality of CPO using chemical components. The outcome of this research is anticipated to provide benefits to the factory by enabling rapid and accurate determination of CPO quality.

II. Research Methodology

In this study, researchers will undertake a series of steps to design a classification model for CPO quality using artificial neural networks. This is conducted to ensure the smooth progress of the research and to achieve the desired outcomes. Refer to Figure 1 for further details.

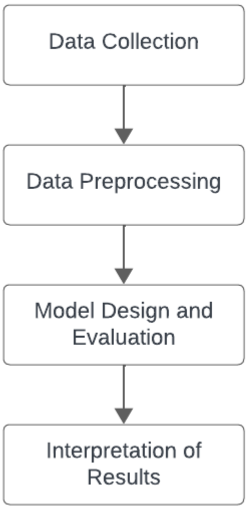

Figure 1. Research phases

A. *Data Collection*

Data obtained from PT. PLB 2 during the period of January to October 2023 consists of 225 entries. This data includes five chemical components related to CPO quality, which are used as variables. These five variables in Table I are impurity level, moisture content, fatty acid content, Deterioration of Bleachability Index (DOBI), and carotene.

TABLE I
CPO QUALITY DATA

No	Dirt	Moist	FFA	DOBI	Carotene	Quality
1	0.018	0.18	4.99	3.87	291	Regular CPO
2	0.015	0.16	2.9	3.3	322	Super CPO
3	0.015	0.17	2.74	3.02	322	Super CPO
…	…	…	…	…	…	…
224	0.015	0.21	4.21	2.61	366	Regular CPO
225	0.016	0.18	3.83	2.29	223	Regular CPO

B. *Data Preprocessing*

The obtained data will undergo the data preprocessing stage. Preprocessing is a step used to clean data that is not yet adequate to influence the classification process [8]. This step aims to clean, organize, and prepare the data before the classification process [9]. The data will be transformed and normalized to be utilized in artificial neural network learning. The data transformation process in this study involves converting character-form data into binary form (1 and 0) for the quality features. There are two types of quality features:" Regular CPO" (0) and" Super CPO" (1). Table I the collected data will be normalized using the Z-score method with the formula (1). Refer to Table II for details. Where the x', x, μ, σ variable is a normalized data, a value to be normalized, feature mean value, and a standard deviation value, respectively

$$x_i' = \frac{x_i - \mu}{\sigma} \quad (1)$$

TABLE II
TRANSFORMED AND NORMALIZED DATA

No	Dirt	Moist	FFA	DOBI	Carotene	Kualitas
1	1.6307	0.8420	3.4016	3.1749	-0.9121	0
2	-1.4184	-0.9885	-0.342	1.6533	-0.443	1
3	-1.4184	-0.0732	-0.6285	0.9059	-0.443	1
…	…	…	…	…	…	
224	-1.4184	3.5880	2.0044	-0.1885	0.2229	0
225	-0.4020	0.8420	1.32382	-1.0427	-1.9413	0

C. *Model Design and Evaluation*

The CPO quality classification model utilizes the Artificial Neural Network (ANN) method. Artificial Neural Networks (ANNs) are mathematical models inspired by the biological neurons of humans, mimicking how biological neurons signal and interact with each other [10]. The architecture of the artificial neural network used consists of 5 neurons in the input layer, 5 neurons in the hidden layer, and 2 neurons in the output layer. The structure of the neural network can be visualized in Figure 2.

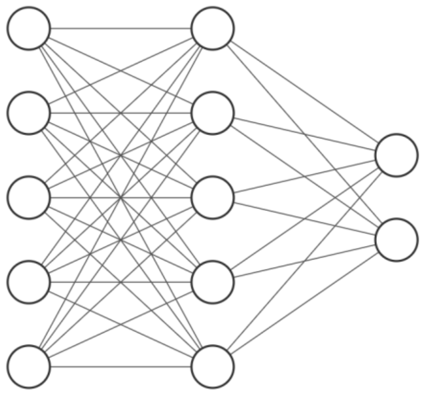

Figure 2. Architecture of the Artificial Neural Network

This artificial neural network model utilizes the backpropagation learning algorithm. Backpropagation is a supervised learning algorithm for neural networks that can minimize errors in the output by calculating the gradient of the error function with respect to the weights in the network and then iteratively adjusting these weights [11]. In this neural network model, the input layer and hidden layer use the ReLU activation function, while the output layer uses the Softmax activation function.

DOI : https://doi.org/10.25139/inform.v9i2.8433

Rectified Linear Unit (ReLU) is an activation function with a simple calculation process. When passing through ReLU, both in forward and backward propagation, it only involves an "if" condition. If the element's value is negative, it will be set to 0 without needing exponential operations, multiplication, or division [12]. The ReLU function is defined as equation 2.

$$f(x) = max(0, x) \tag{2}$$

Softmax is similar to logistic regression but is used in the context of classification of more than two classes. Below is the softmax activation equation [13]:

$$f(x) = \frac{e^{x_i}}{\sum_{j=1}^{K} e^{x_i}} \tag{3}$$

The classification model's performance is evaluated using the accuracy value obtained from calculating the True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN) values from the Confusion Matrix table [14]. Equations 4 to 5 are utilized to compute each value [15]:

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FN} \tag{4}$$

In this study, the learning rate is set to 0.01. This value is a common choice in ANN training and has been demonstrated to be effective in various applications [16]. Preliminary experiments conducted by the researchers also confirmed that a learning rate of 0.01 provides stable training and good convergence. Additionally, this learning rate is expected to help prevent overfitting by allowing the model to learn gradually and avoid overly adapting to the training data. It also contributes to a more stable training process and better convergence, as the model makes smaller weight updates at each iteration, reducing the risk of oscillations or divergence. Ultimately, the selection of 0.01 as the learning rate aims to strike a balance between model performance, training stability, and generalization capability.

The selection of 5 neurons in the hidden layer is a deliberate decision based on experiments with varying numbers of hidden layer neurons (e.g., 3, 5, 7, and 10), indicating that 5 neurons offered a good balance of performance without overfitting and strikes a balance between model performance and computational efficiency, as increasing the number of neurons would increase model complexity and training time.

III. Result and Discussion

The previously designed model is then implemented using the Python programming language. The first step involves using 180 training data samples to carry out the training process, which consists of 5 neurons in the input layer, 5 neurons in the hidden layer, and 2 neurons in the output layer. Figure 3 shows the reduction in loss of the backpropagation artificial neural network algorithm after 1000 epochs of training. From the first epoch to the 200th epoch, there is a significant decrease in loss value. However, from the 200th to the 1000th epoch, the reduction in error value becomes more gradual.

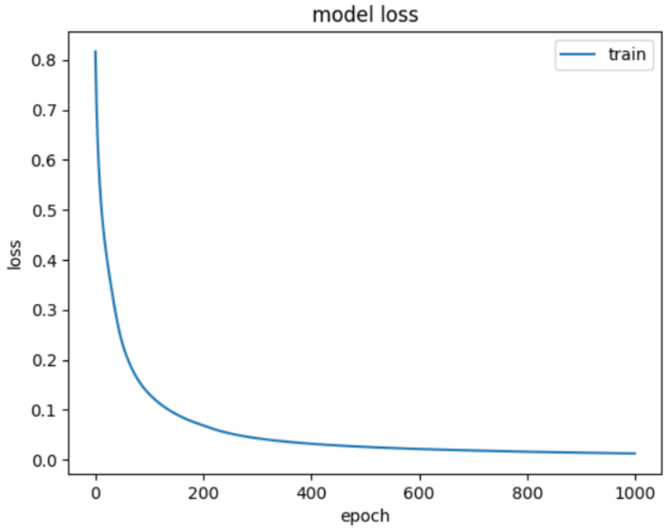

Figure 3. Graph of the Neural Network Model Training Results

After applying the backpropagation model, the loss value obtained for CPO quality is 0.0152 after 1000 epochs. Below is Table III, the result of a number of epochs against the loss value during the iteration process:

TABLE III
TRAINING DATA RESULTS

Epoch	Loss
1	0.8164
100	0.1319
200	0.0690
300	0.0429
400	0.0319
500	0.0256
600	0.0214
700	0.0183
800	0.0159
900	0.0140
1000	0.0124

After completing the training phase, the next step is data testing. Table IV below displays the prediction results and actual values for each iteration conducted.

TABLE IV
PREDICTION RESULTS USING TESTING DATA

No	Aktual	Prediksi
181	1	1
182	1	1
183	0	0
	…	…
216	0	1
217	0	0
218	1	1
219	1	1
220	1	1
221	0	0
222	1	1
223	0	0
224	1	1
225	1	1

DOI : https://doi.org/10.25139/inform.v9i2.8433

TABLE V
TESTING DATA RESULTS IN THE FORM OF A CONFUSION MATRIX

	Prediksi: Regular CPO	**Prediksi: Super CPO**
Aktual: Regular CPO	17	1
Aktual: Super CPO	0	27

From the Table V testing data results, it can be observed that all rows for Regular CPO have correct predictions, while for "Super CPO", there is one incorrect prediction, which is in row 216.

The accuracy can be calculated by summing the correct predictions (true positives) for both classes and dividing it by the total number of samples [17]: Total correct predictions = True Positives for Regular CPO + True Positives for Super CPO = 17 + 27 = 44 Total samples = Sum of all entries in the confusion matrix = 17 + 1 + 0 + 27 = 45. Hence, the accuracy of the model, based on the provided confusion matrix, is approximately 97.78%.

The research results show that ANN can classify CPO quality with a high level of accuracy. This indicates that this method has great potential for use in the industry to improve CPO quality assessment. The improvement in accuracy and validity of the ANN classification results has significant implications for the industry. With high accuracy, ANN can reduce human errors in quality assessment, speed up the process, and increase the consistency of assessment results. This is crucial because consistent quality assessment can enhance operational efficiency and reduce production costs.

The implementation of ANN can also reduce dependence on laboratory tests that are time-consuming and costly. With an integrated ANN system in the production process, quality assessment results can be obtained in real time, allowing companies to take corrective actions immediately if there are indications of quality decline. Using ANN not only improves efficiency but also reduces overall production costs.

Additionally, ANN allows companies to set more precise quality standards based on historical data analysis and chemical component patterns. Thus, companies can continuously improve stricter and more accurate quality standards, which in turn increases product reputation and customer satisfaction.

Having a tool that can efficiently classify CPO quality allows companies to focus more on product development and innovation. For instance, companies can explore new formulations or processing methods that can further enhance CPO quality. ANN classification results can be used for deeper research and development, potentially increasing production process efficiency and the quality of the final product.

Overall, the results of CPO quality classification using ANN based on chemical components not only provide insights into the capabilities of this technology in quality assessment but also demonstrate various practical implications that the industry can adopt to improve operational efficiency, reduce costs, and enhance overall product quality. Therefore, implementing ANN in the CPO production process is not only a technological innovation but also a strategic step to maintain competitiveness in the global market.

IV. Conclusion

This study utilized a dataset of CPO quality from PT. Perkebunan Lembah Bhakti (PLB) 2 during the period from January to October 2023, comprising a total of 225 records and 5 features/variables. The backpropagation artificial neural network model for the training process was designed with ReLU and softmax activation functions, 1000 iterations (epochs), a learning rate of 0.01, and 5 neurons in the hidden layer. The research results show that ANN can classify CPO quality with a high level of accuracy of 97.78% on 45 testing data. Proper parameter usage, along with appropriate data representation, will yield good accuracy values in model design. This indicates that this method has great potential for use in the industry to improve CPO quality assessment.

References

[1] L. Murjana, “Analisa Pengendalian Kualitas Crude Palm Oil (CPO) dengan menggunakan Metode Statistical Quality Control (SQC) pada PT. Sapta Karya Damai Kalimantan Tengah,” UPN Veteran Jawa Timur, 2022.
[2] D. Nurhasanah, D. A. Lestari, and S. Simatupang, “Pemilihan Kualitas Produk Kelapa Sawit Menggunakan Metode Naive Bayes di Labuhanbatu Selatan,” *J. Tek.*, vol. 3, no. 1, pp. 24–31, 2023.
[3] H. Aini, H. Haviluddin, E. Budiman, M. Wati, and N. Puspitasari, “Prediksi Produksi Minyak Kelapa Sawit Menggunakan Metode Backpropagation Neural Network,” *Sains, Apl. Komputasi dan Teknol. Inf.*, vol. 1, no. 1, pp. 24–33, 2019.
[4] D. A. Raihan, D. Syauqy, and B. H. Prasetio, “Klasifikasi Kualitas Beras berdasarkan Nilai Data Larik Sensor Gas MQ menggunakan Metode Jaringan Syaraf Tiruan berbasis Arduino,” *... Inf. dan Ilmu Komput. e-ISSN*, vol. 6, no. 6, pp. 2974–2981, 2022.
[5] Jusrawati, A. Futri, and A. B. Kaswar, “Klasifikasi Tingkat Kematangan Buah Pisang Dalam Ruang Warna RGB Menggunakan Jaringan Syaraf Tiruan (JST),” vol. 02, no. May, pp. 49–54, 2021.
[6] M. Ikhsan, A. Armansyah, and A. A. Tamba, “Implementasi Jaringan Syaraf Tiruan Backpropagation Pada Klasifikasi Grade Teh Hitam,” *J. Sist. Komput. dan Inform.*, vol. 4, no. 2, pp. 387–395, 2022.
[7] F. Erlangga, D. R. Wijaya, and W. Wikusna, "Electronic Nose Dataset for Classifying Rice Quality using Neural Network," in *2021 9th International Conference on Information and Communication Technology (ICoICT)*, 2021, pp. 462–466. doi: 10.1109/ICoICT52021.2021.9527423.
[8] M. Z. Sarwani and D. A. Sani, "Social Media Analysis Using Probabilistic Neural Network Algorithm to Know Personality Traits," *Editor. BOARD*, vol. 6, no. 1, 2020.
[9] S. Jesika, S. Ramadhani, and Y. P. Putri, “Implementasi Model Machine Learning dalam Mengklasifikasi Kualitas Air,” *J. Ilm. Dan Karya Mhs.*, vol. 1, no. 6, pp. 382–396, 2023.
[10] D. Monika, A. Ahmad, S. Wardani, and Solikhun, “Model Jaringan Syaraf Tiruan Dalam Memprediksi Ketersediaan Cabai Berdasarkan Provinsi,” *Teknika*, vol. 8, no. 1 SE-Articles, Jun. 2019, doi: 10.34148/teknika.v8i1.140.
[11] R. Rahmiyanti, S. Defit, and Y. Yunus, “Prediksi dan Klasifikasi Buku Menggunakan Metode Backpropagation,” *J. Inf. dan Teknol.*, pp. 109–114, 2021.
[12] I. Firmansyah and B. H. Hayadi, “Komparasi Fungsi Aktivasi Relu Dan Tanh Pada Multilayer Perceptron,” *JIKO (Jurnal Inform. dan Komputer)*, vol. 6, no. 2, pp. 200–206, 2022.
[13] R. Radikto, D. I. Mulyana, M. A. Rofik, and M. O. Z. Zakaria, “Klasifikasi Kendaraan pada Jalan Raya menggunakan Algoritma Convolutional Neural Network (CNN),” *J. Pendidik. Tambusai*, vol. 6, no. 1, pp. 1668–1679, 2022.
[14] D. Pardede and B. H. Hayadi, "MotoGP Mandalika 2022 Sentiment Classification Using Machine Learning," *J. Transform.*, vol. 20, no. 2, pp. 42–50, 2023.
[15] J. T. Samudra, B. H. Hayadi, and P. S. Ramadhan, “Komparasi 3 Metode Algoritma Klasifikasi Data Mining Pada Prediksi Kenaikan Jabatan,” *J. Teknol. Sist. Inf. dan Sist. Komput. TGD*, vol. 5, no. 2, pp. 127–133, 2022.

DOI : https://doi.org/10.25139/inform.v9i2.8433

[16] C. D. Suhendra and A. C. Saputra, "Penentuan parameter learning rate selama pembelajaran jaringan syaraf tiruan backpropagation menggunakan algoritma genetika," *J. Teknol. Inf. J. Keilmuan Dan Apl. Bid. Tek. Inform.*, vol. 14, no. 2, 2020.

[17] M. A. Ridho and M. Arman, "Analisis Serangan DDoS Menggunakan Metode Jaringan Saraf Tiruan," *J. Sisfokom (Sistem Inf. Dan Komputer)*, vol. 9, no. 3, pp. 373–379, 2020.

DOI : https://doi.org/10.25139/inform.v9i2.8433