Image Classification of Seasoning Package Completeness in Noodle Products Using WEKA Analysis

Rendi Priyatna
Fakultas Teknologi Industri
Universitas Ahmad Dahlan
Yogyakarta,Indonesia
2307057017@webmail.uad.ac.id

Sunardi
Fakultas Teknologi Industri
Universitas Ahmad Dahlan
Yogyakarta,Indonesia
sunardi@ee.uad.ac.id

Anton Yudhana
Fakultas Teknologi Industri
Universitas Ahmad Dahlan
Yogyakarta,Indonesia
eyudhana@ee.uad.ac.id

Abstract—This research develops an intelligent system that utilizes vision camera technology to detect completeness of noodle packages consisting of noodle blocks, oil, and seasoning. Multi-Layer Perceptron (MLP) and Naïve Bayes are used to classify images in recognizing the shape and color of the seasoning that should be present in noodle packages using Weka. The system's input is the captured data of noodle package completeness taken in real-time with randomly positioned oil and seasoning. A total of 486 random data points were used, with 70% for training and 30% for testing. The testing results show that MLP outperforms Naïve Bayes in almost all evaluation metrics, with an accuracy of 98.48% for MLP, compared to 74.32% for Naïve Bayes. In terms of construction time, Naïve Bayes is superior with a construction time of 0.01 seconds.

Keywords—*Seasoning*, Naïve bayes, MLP, & Weka.

I. INTRODUCTION

Food industry, particularly noodle production, is experiencing increased consumer demand. The industry is not only faced with the need to increase production capacity but also to ensure consistent product quality. One crucial quality aspect is the completeness of the seasoning in noodles, which plays a significant role in providing flavor and added value to the product. However, despite the importance of seasoning completeness, the quality inspection process is still largely conducted manually, prone to human error, and unable to efficiently handle high production volume. Therefore, innovation is needed in the form of an intelligent system that can automatically detect and analyze the completeness of the seasoning in noodles.

In the initial stage, the classification of the completeness of seasoning in noodle packages is carried out. Classification is the object recognition stage [1]. Humans can easily and quickly recognize new objects they see, but this is not the case with machines [2].

Classifying shapes and colors in images, there are many methods available for processing, one of which is Neural Network (NN). CNN architecture has been used by many researchers for effective object representation [3], [4]. CNN is one of the algorithms of deep learning, which is a part of machine learning [5], which studies image representation and segmentation to achieve accurate results in predicting or detecting an object using a camera, consisting of several processing layers [6].

Several previous studies have yielded significant results by applying NN to various objects and cases [7], such as classifying the ripeness level of sweet oranges based on texture and color brightness using Deep Learning Convolutional Neural Network with Multilayer Perceptron algorithm. From 100 data sheet images, a testing accuracy graph of 92% was obtained, able to classify oranges into good quality (ripe) and poor quality (rotten) [8].

Research on image identification of textile fabrics using digital microscopy for image data acquisition with fabric algorithm showed an accuracy rate of 93% for original image size 600x800, using feature extraction method histogram and classification technique Squared Chi Squarede metric distance [2].

Another study classified tilapia and tilapia fish using backpropagation. From 254 data points, although the types of tilapia and tilapia fish appear similar physically, the results achieved a relatively good accuracy rate with a success rate of 74.04% [9].

Another study classified diseases in the healthcare field to help early detection of hypertension in patients using Naïve Bayes method and data discretization CART, achieving an accuracy rate of 84.28% [10].

The research applied to classify and identify types of vehicles such as cars and motorcycles, specifically four-wheeled vehicles [11], [12], [13]. Achieved a high accuracy rate averaging above 95%.

Based on the description above, this research aims to develop a system to classify the completeness of seasoning packages in noodle products using WEKA software. The objective of this study is to find the best method for classifying image characteristics of seasoning completeness in noodle packages using Multi-Layer Perceptron (MLP) and Naïve Bayes algorithms. The selection of these two algorithms, MLP representing complex data classification, adopts the neural functioning of living organisms [3], [14]. Meanwhile, Bayes represents a simple classification method that applies Bayes' theorem, assuming all features are independent of each other [15], [16]. This algorithm uses probability theory in its classification process [17], [18]. By collecting 486 randomly sampled data in the field, with 70% used for training data and 30% for testing data.

II. METODOLOGI

Research process flow can be seen in Figure 2.1. The study is conducted in several phases, starting from data collection, which will later become training and testing data, data classification, and finally, analysis stage.

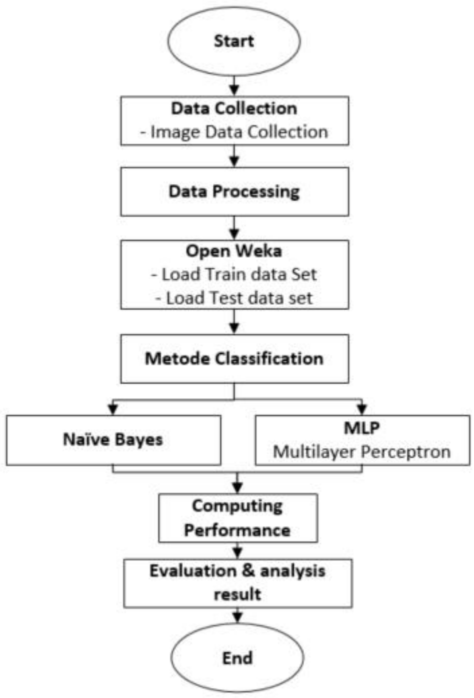


Figure 1. Diagram Alir Penelitian

A. Data Collection

Image data collection is directly taken from the noodle manufacturing process with the assistance of a series of electronic tools, as shown in Figure 2.2. The camera captures the object when the trigger sensor detects the noodle block. The image resolution is 1088 x 672 using a webcam camera. Picture properties can be seen in Figure 2.3.

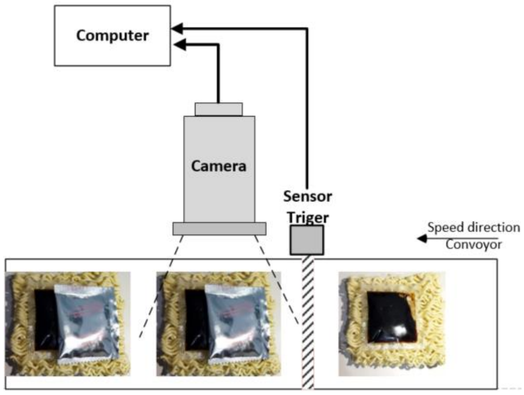
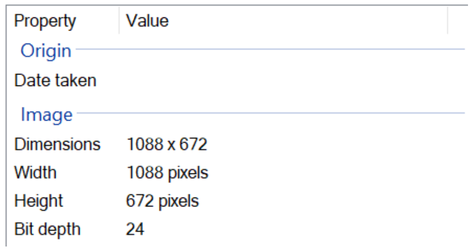


Figure 2. Image Capture Methode



Property	Value
Origin	
Date taken	
Image	
Dimensions	1088 x 672
Width	1088 pixels
Height	672 pixels
Bit depth	24

Figure 3. Properties Picture

B. Data Processing

At this stage, the collected data will be processed to determine the average RGB values in the captured images. This processing is assisted using the Python program with the help of the Python Imaging Library (PIL) and pandas for exporting results to a folder population. Figure 2.4, shows a snippet of the Python program using the PIL and Pandas libraries.

```
image_files = [f for f in os.listdir(folder_path) if f.endswith('.jpg') or f.endswith('.png')]
all_rgb_data = []
for image_file in image_files:
    image_path = os.path.join(folder_path, image_file)
    rgb_data = convert_to_rgb(image_path)
    all_rgb_data.extend(rgb_data)
save_to_csv(all_rgb_data, csv_file)
```

Figure 4. Snippet of Pytjon code

The output from the Python program provides the average RGB values of the images in a folder population and saves them in an Excel file, as shown in Figure 2.5.

Image	R_avg	G_avg	B_avg
scene00016.png	133.232	133.3823	132.5797
scene00031.png	117.7343	115.1691	114.3882
scene00046.png	123.1728	125.7065	125.2489
scene00061.png	124.417	122.0264	118.2709
scene00076.png	129.2516	130.725	128.6836
scene00091.png	125.5694	127.6158	122.2069
scene00106.png	124.8499	122.2304	119.2686
scene00121.png	134.4353	131.9959	129.7993
e00136.png	115.573	113.8244	112.6206
e00151.png	119.1868	116.595	120.213
scene	95.15982	95.65146	89.68632
scene	147.6139	149.8635	152.6105
scene00196.png	108.9645	106.0622	104.628
scene00211.png	101.0059	102.0616	100.1422
scene00226.png	117.1985	115.2324	110.9888
scene00241.png	128.4827	132.5244	124.3896
scene00256.png	123.7312	121.1965	120.4839
scene00271.png	121.2487	118.8289	114.4783
scene00286.png	127.1253	125.7223	122.9284

Figure 5. Results of converting images to RGB values

C. Tools dan data set

Waikato Environment for Knowledge Analysis (WEKA, version 3.8.6) is software that provides a data mining interface with numerous machine learning algorithms, commonly used to compare various data mining techniques to determine the best one [17], [19]. WEKA operates based on predictive data described by specific attribute types [20], using datasets sourced from a food factory that processes noodles. The data includes two categories: NG (Not Good) and OK, ss shown in Figure 2.6.

(a) Complete Noodle Package "OK"

(b) Incomplete Noodle Package "NG"

Figure 6. Noodle package completeness

D. Image Classification Method

2 image classification methods will be used: MLP and Bayes.

1. MLP (Multilayer Perceptron)
   MLP algorithm is used for regression and classification. MLP trains artificial neural networks by passing inputs through weighted layers of neurons and adjusting these weights to replicate patterns in the output. Each neuron calculates its output as the weighted sum of inputs plus bias, which is then processed by an activation function to introduce non-linearity. This function allows MLP to capture complex relationships in data, making it effective for various applications [6].

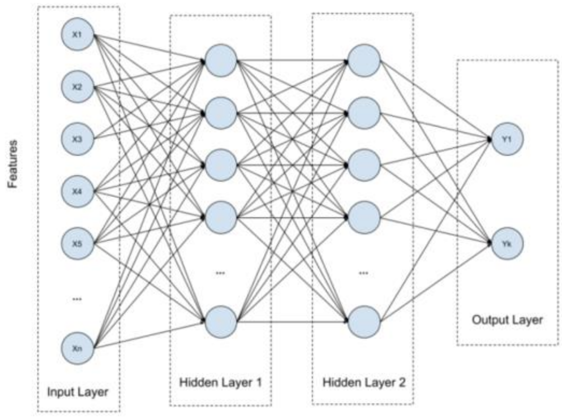


Figure 7. MLP Neural Network Topology

Figure 2.7, shows the neural network topology consisting of hidden layers. The hidden layers are activated using activation functions.

$$y_{Out} = f\left(b + \sum_{i=1}^{n} w_i \, . \, x_i\right) \tag{1}$$

Where :
y_{Out} : output neuron
f : activation function (e.g., Relu, Sigmoid, tanh, Softmax)
b : bias neuron
w : weight connecting neuron I to j
x : input data x =(x1,x2,x3,,,,,,,xn)

2. Naïve Bayes
   Bayesian networks are a combination of probability and graphical models. Probabilistic classification algorithms use the assumption of independent features [15], as shown in Figure 2.8.

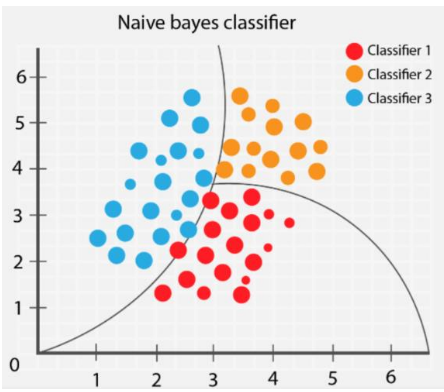


Figure 8. Bayes Algorithm

Naive Bayes algorithm is a statistical method in machine learning for classification that calculates prior probabilities, likelihood probabilities, and posterior probabilities using rules [15], [16], [17], [21]. Bayes' theorem states the relationship between conditional probability and marginal probability as follows:

$$\mathrm{P(C|X)} = \frac{P(X|C.P(C)}{P(X)} \tag{2}$$

Where :

P(C|X) : posterior probability of class C given feature X
P (X|C) : proabilitas likelihood dari fitur X diberikan kelas C
P(C) : probabilitas prior dari kelas C
P(X) : probabilitas ecivence dari fitur X

III. RESULT & DISCUSSION

A. Preprocessed Data

This stage processes captured data into RGB values using Python V.3. There are four categories of datasets as shown in Table 3.1.

Tabel 3.1.
Preprocessing data object

Noodle Block	*Seasoning* powder	*Seasoning* Oil	Category
-	-	-	NG
✓	-	-	NG
✓	✓	-	NG
✓	-	✓	NG
✓	✓	✓	**OK**

B. Computational Results and Performance Comparison

Image classification results in Figure 3.1, and Table 3.2, show that the accuracy rate of MLP is significantly higher than Naïve Bayes, with lower classification errors in MLP. This indicates that MLP is more effective in correctly classifying the data.

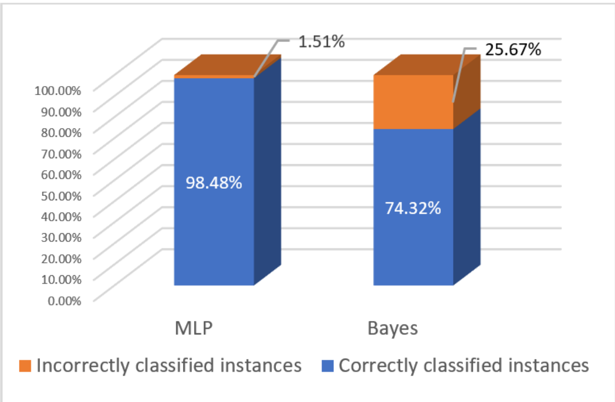


Figure 9. Comparison of Classification Accuracy between MLP and Bayes Methods

The higher Kappa statistic value in MLP (0.96) compared to Naïve Bayes (0.47) indicates that MLP has better agreement between predicted results and actual classification. The Mean Absolute Error (MAE) in MLP (0.02) is lower than in Naïve Bayes (0.28), suggesting that MLP predictions are closer to the actual values. Additionally, the Root Mean Squared Error (RMSE) in MLP (0.12) is also lower than Naïve Bayes (0.44), meaning that the average squared prediction error in MLP is smaller, as shown in Table 3.2.

Tabel 3.2.
Comparison of Computational Results of Image Data

Parameter	**Algorithm**	
	MLP (Multilayer Perceptron)	**Naïve Bayes**
Kappa statistic	0.96	0.47
Mean absolute error	0.02	0.28
Root mean squared error	0.12	0.44
Relative absolute error	4.59%	57.79%
Root relative squared error	24.21%	89.79%
Construction Time	8.9 Sec	0.01 Sec

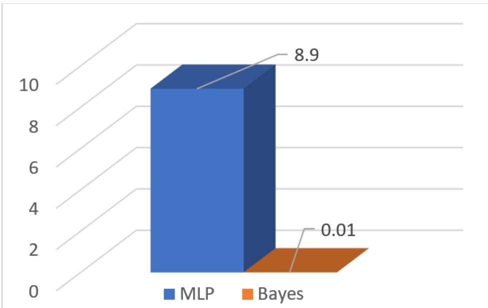


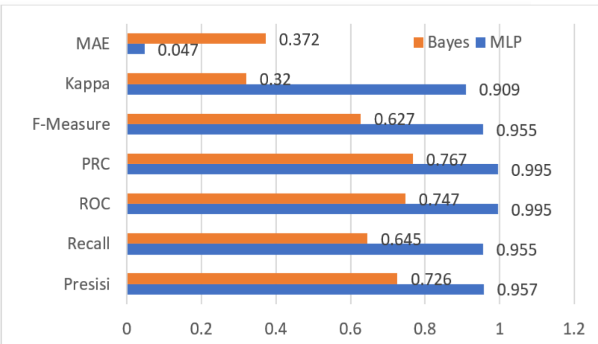
Figure 10.Time Comparason

However, in terms of learning time, Naïve Bayes has the advantage with a construction time of only 0.01 seconds, much faster than MLP, which requires 8.9 seconds. The faster construction speed is suitable for applications that require quick processing with simpler data.

Despite Naïve Bayes being more time-efficient, MLP's superior predictive performance compared to Naïve Bayes makes it a better choice for applications that require high accuracy.

C. Testing

In the testing phase, predictions were made for 155 image testing data that were not used in training. The test results can be shown in Figure 3.3, and Figure 3.4.



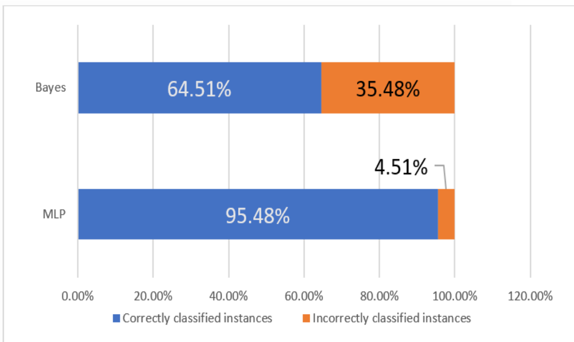


Figure 11.
Comparison of Classification Accuracy Test Results between MLP and Bayes Methods

Comparison results based on Mean Absolute Error (MAE) evaluation parameter show that MLP is significantly lower (0.047) compared to Naïve Bayes (0.372). Looking at the Kappa statistic, which measures the agreement between predicted results and actual outcomes, MLP demonstrates a significant advantage with a value of 0.909, while Naïve Bayes only reaches 0.32. Similarly, other parameters such as F Measure, Precision-Recall Curve (PRC), and Receiver Operating Characteristic (ROC) consistently show higher values for MLP compared to Naïve Bayes. With MLP achieving an accuracy rate of 95.48% and Naïve Bayes 64.51%, the decrease in performance from training classification may be due to variations in captured product outcomes in the field, which could be improved by increasing the amount of training data.

IV. CONCLUSION

Overall, MLP shows superior performance compared to Naïve Bayes across nearly all evaluation metrics, except for construction time. Therefore, for ensuring the quality of noodle products with accurate seasoning completeness, using the MLP algorithm is more effective than Naïve Bayes, especially when dealing with complex data and requiring higher accuracy, lower errors, and better consistency in predictions, except for data processing issues.

Further research will be conducted to explore the correlation between processing time and accuracy in seasoning completeness image data analysis. This study will include actual speed tests in the field to gain a more comprehensive understanding of algorithm performance under real operational conditions.

REFERENCES

[1] M. Aldossary, H. A. Alharbi, and C. Anwar Ul Hassan, “Internet of Things (IoT)-Enabled Machine Learning Models for Efficient Monitoring of Smart Agriculture,” *IEEE Access*, vol. 12, pp. 75718–75734, 2024, doi: 10.1109/ACCESS.2024.3404651.

[2] A. F. Program, S. Teknik, E. Universitas, A. Dahlan, J. Soepomo, and J. Yogyakarta, “SISTEM PENGENALAN CITRA JENIS-JENIS TEKSTIL.”

[3] E. R. Albasiouny, A. F. A. Heliel, H. E. Abdelmunim, and H. M. Abbas, “Multilayer Perceptron Generative Model via Adversarial Learning for Robust Visual Tracking,” *IEEE Access*, vol. 10, pp. 121230–121248, 2022, doi: 10.1109/ACCESS.2022.3222867.

[4] K. Endo, M. Tanaka, and M. Okutomi, “CNN-Based Classification of Degraded Images Without Sacrificing Clean Images,” *IEEE Access*, vol. 9, pp. 116094–116104, 2021, doi: 10.1109/ACCESS.2021.3105957.

[5] N. Zuzzaifa Rianto Rianto Fakultas Sains dan Teknologi, “Convolutional Neural Network Untuk Perbandingan Optimizer Pada Citra Batang Pohon,” 2023.

[6] O. Oktay *et al.*, “Anatomically Constrained Neural Networks (ACNNs): Application to Cardiac Image Enhancement and Segmentation,” *IEEE Trans Med Imaging*, vol. 37, no. 2, pp. 384–395, Feb. 2018, doi: 10.1109/TMI.2017.2743464.

[7] A. Muis, A. Yudhana, and A. Dahlan, “COMPARISON ANALYSIS OF BRAIN IMAGE CLASSIFICATION BASED ON THRESHOLDING SEGMENTATION WITH CONVOLUTIONAL NEURAL NETWORK.”

[8] B. Yanto *et al.*, “Klasifikasi Tekstur Kematangan Buah Jeruk Manis Berdasarkan Tingkat Kecerahan Warna dengan Metode Deep Learning Convolutional Neural Network,” vol. 6, no. 2, p. 2021.

[9] J. Homepage, S. Windharta Oei, L. Yahya, I. Wiranto, and P. Korespondensi, “Research in the Mathematical and Natural Sciences Deteksi Citra Ikan Nila dan

Mujair Menggunakan Metode Jaringan Syaraf Tiruan Propagasi Balik," *Res. Math. Nat. Sci*, vol. 2, no. 1, pp. 32–36, 2023, doi: 10.55657/rmns.v2i1.85.

[10] Y. Munali, "Classification of Hypertension Using Naïve Beyes Method with Risk Factor Data Discretization Approach," 2024.

[11] A. Peryanto, A. Yudhana, and D. R. Umar, "Rancang Bangun Klasifikasi Citra Dengan Teknologi Deep Learning Berbasis Metode Convolutional Neural Network," 2019. [Online]. Available: https://www.mathworks.com/discovery/convolutional-neural-network.html

[12] F. Ramadhani, A. Satria, and S. Dewi, "Identifikasi Kendaraan Bermotor pada Dashcam Mobil Menggunakan Algoritma YOLO," *Hello World Jurnal Ilmu Komputer*, vol. 2, no. 4, pp. 199–206, Mar. 2024, doi: 10.56211/helloworld.v2i4.466.

[13] M. Irfan, B. Alldino, A. Sumbodo, and I. Candradewi, "Sistem Klasifikasi Kendaraan Berbasis Pengolahan Citra Digital dengan Metode Multilayer Perceptron," *IJEIS*, vol. 7, no. 2, pp. 139–148, 2017.

[14] B. Heuju and D. Sagar Baral, "Anomaly Detection in Computer Networks using Multilayer Perceptron," no. 10, pp. 2350–8906.

[15] Dewi Widyawati and Amaliah Faradibah, "Comparison Analysis of Classification Model Performance in Lung Cancer Prediction Using Decision Tree, Naive Bayes, and Support Vector Machine," *Indonesian Journal of Data and Science*, vol. 4, no. 2, pp. 80–89, Jul. 2023, doi: 10.56705/ijodas.v4i2.76.

[16] T. Kim and J. S. Lee, "Exponential Loss Minimization for Learning Weighted Naive Bayes Classifiers," *IEEE Access*, vol. 10, pp. 22724–22736, 2022, doi: 10.1109/ACCESS.2022.3155231.

[17] B. Imran *et al.*, "DATA MINING USING RANDOM FOREST, NAÏVE BAYES, AND ADABOOST MODELS FOR PREDICTION AND CLASSIFICATION OF BENIGN AND MALIGNANT BREAST CANCER." doi: 10.33480/pilar.v18i1.2912.

[18] K. Eka Setyaputri and A. Fadlil, "Comparative Analysis of Certainty Factor Method and Bayes Probability Method on ENT Disease Expert System," *Scientific Journal of Informatics*, vol. 5, no. 2, pp. 2407–7658, 2018, [Online]. Available: http://journal.unnes.ac.id/nju/index.php/sji

[19] R. Yatabe *et al.*, "Odor Sensor System Using Chemosensitive Resistor Array and Machine Learning," *IEEE Sens J*, vol. 21, no. 2, pp. 2077–2083, Jan. 2021, doi: 10.1109/JSEN.2020.3016678.

[20] F. Aslay and N. S. Ting, "Machine Learning-Based Estimation of Output Current Ripple in PFC-IBC Used in Battery Charger of Electrical Vehicles: A Comparison of LR, RF and ANN Techniques," *IEEE Access*, vol. 10, pp. 50078–50086, 2022, doi: 10.1109/ACCESS.2022.3174100.

[21] N. Deepa, J. Sathya Priya, and T. Devi, "Towards applying internet of things and machine learning for the risk prediction of COVID-19 in pandemic situation using Naive Bayes classifier for improving accuracy," *Mater Today Proc*, vol. 62, pp. 4795–4799, Jan. 2022, doi: 10.1016/j.matpr.2022.03.345.