Plastic-Type Classification for Sorting System Based on Digital Image using Multinomial Logistic Regression with k-Fold Cross Validation

Irsyadi Yani, Ismail Thamrin, Dewi Puspitasari, Barlin, Yulia Resti, and Dendy Adanta, *Member, IAENG*

***Abstract*— Recycling plastics has become increasingly prevalent in recent years. Recycling reduces plastic waste through the reuse of materials as opposed to their disposal. Additionally, this method aids in the reduction of pollution caused by the emission of greenhouse gases during the production of new plastic from basic materials. The initial phase of the plastic waste recycling procedure involves sorting plastic to various types of material. Accurately identifying the type of plastic is exceptionally beneficial for developing sifting systems in the recycling industry. This study aimed to test how well multinomial logistic regression with k-fold cross-validation can determine the difference among varying types of plastic. This method is a frequently employed statistical learning technique due to its generally satisfactory performance compared with alternative methods. Results showed that multinomial logistic regression performed well in identifying the type of plastic in all performance metrics. The performance average measures of five folds were 86.08% accuracy, 79.11% recall-μ, 79.08% recall-M, 89.56% specificity-μ and 89.59% specificity-M.**

***Index Terms*— Plastic type, classification method, k-fold cross-validation, classification.**

I. Introduction

PLASTIC, an inorganic substance, is extensively utilised globally, particularly in nations undergoing substantial economic expansion [1]. Annual plastic production has increased by over two-thirds worldwide, from 234 million tons (Mt) in 2000 to 460 Mt in 2019. Since 2000, plastic waste has increased by more than twofold, from 156 Mt to 353 Mt in 2019 [2]. One of the most significant problems the world is currently experiencing is plastic waste, which has far-reaching effects. It can harm species, disrupt ecosystems and hurt human health [3].

When considering reducing and managing plastic waste, recycling has emerged as a viable alternative to landfills and incineration [4]. Material type separation is a critical initial procedure within the plastic recycling industry. Inadequate classification practices may lead to cross-contamination among different types of plastic, resulting in escalated operational costs for an industrial facility [5]. As mentioned previously, the procedure often faces challenges in differentiating among varying types of plastic [6]. The capacity to precisely predict the nature of plastic is exceedingly advantageous in the context of sorting system advancements within the recycling sector. Given that sorting plastic by hand is inefficient, automatic sorting has become a possible solution [7]. Applying machine learning techniques to a digital image as a dataset can create an economical automated plastic sorting system [8].

The use of digital images for object classification or identification has been prevalent in the last decade because of their low cost [9], especially those transformed into red, green and blue (RGB) colour space models [10]. Digital image-based object classification has demonstrated satisfactory performance in numerous applications, including weed damage [11], histopathology medical records [12], plant disease and pest [13], [14], [15], [16], [17], [18], land cover and land use [19], [20] and privacy-preserving images [21]. Researchers have applied several statistical machine learning methods, including multinomial naive Bayes [22], decision trees [23], Fisher discriminant analysis [8] and artificial neural network backpropagation [24], to classify plastic types based on digital images. We need to conduct further in-depth exploration to identify the optimal plastic-type classification method.

Logistic regression can be employed to organise data in machine learning. It works well in many situations, such as auto insurance portfolios [25], medical records [26], identifying patients with COVID-19 [27], classifying corn diseases and pests [28], sorting new-born weights [29], industrial tomography [30], determining if someone has diabetes mellitus [31] and sorting different kinds of trash

Manuscript received September 23, 2024; revised February 11, 2025.

This work was supported by Universitas Sriwijaya.

Irsyadi Yani is a lecturer in the Department of Mechanical Engineering, Faculty of Engineering, Universitas Sriwijaya, Indralaya 30662, South Sumatera, Indonesia (corresponding author to provide e-mail: irsyadiyani@ft.unsri.ac.id).

Ismail Thamrin is a lecturer in the Department of Mechanical Engineering, Faculty of Engineering, Universitas Sriwijaya, Indralaya 30662, South Sumatera, Indonesia (e-mail: ismailthamrin@ft.unsri.ac.id).

Dewi Puspitasari is a lecturer in the Department of Mechanical Engineering, Faculty of Engineering, Universitas Sriwijaya, Indralaya 30662, South Sumatera, Indonesia (e-mail: dewipuspitasari@unsri.ac.id).

Barlin is a lecturer in the Department of Mechanical Engineering, Faculty of Engineering, Universitas Sriwijaya, Indralaya 30662, South Sumatera, Indonesia (e-mail: barlin@ft.unsri.ac.id).

Yulia Resti is a lecturer in the Department of Mathematic, Faculty of Math and Natural Science, Universitas Sriwijaya, Indralaya 30662, South Sumatera, Indonesia (e-mail: yulia_resti@mipa.unsri.ac.id).

Dendy Adanta is a lecturer in the Department of Mechanical Engineering, Faculty of Engineering, Universitas Sriwijaya, Indralaya 30662, South Sumatera, Indonesia (e-mail: dendyadanta@ymail.com).

[32]. This method assumes that the objective variable has a Bernoulli distribution, and the predicted value is a probability. Consequently, the results are simple to interpret. This method allows us to determine the probability that an observation belongs to a particular category (class) of the target variable [25]. Multinomial logistic regression is unique in that it is applicable to more than two classes [33]. Logistic regression or multinomial logistic regression does not consistently perform well in various classification or identification tasks, but it often achieves accuracy beyond 90%. To evaluate the validity of model performance, one can employ the sideways method. A resampling method with minimal bias is k-fold cross-validation with k=5 [34]. This resampling method is recommended as a result of the enhanced accuracy of the derived classification performance, which is determined as the mean of the k datasets that are generated [35]. The goal of this study is to test how well k-fold cross-validation method with k=5 and multinomial logistic regression on digital images can classify the type of plastic.

II. Methodology

The following procedures were used to predict the type of plastic based on a digital image, using resampling k-fold cross-validation with k=5, in conjunction with multinomial logistic regression.

1) Pre-processing the digital image. The initial step in this phase comprised two operations: cropping the image and extracting the RGB colour characteristics, including entropy and variance.
2) We divided the data into training and test sets via k-fold cross-validation to resample them, as shown in Table 1. We partitioned the data into five folds of comparable size, designating four for training and one for testing, resulting in five distinct data compositions.
3) The initial (k-1) fold data were modelled by multinomial logistic regression, whilst the remaining data were tested employing a onefold approach. We sampled the specified fold five times and calculated the prediction performance as the average of the five calculated sizes [34], [35], [36].
4) The performance of plastic-type classification was assessed based on the confusion matrix [8], [37], [38]. Suppose TP_j is the true classification of the j-th plastic-type; TN_j is the true classification of not the j-th plastic-type; FP_j is the false classification of the j-tj plastic-type, where it is not the j-th plastic-type but is classified as the j-th plastic-type; and FN_j is the false classification of the j-th for j=1, as shown in Table 1. For the j-th plastic-type, j=1,2,3 for TP_j, FP_j, TB_j, and FN_j counts, respectively.

Performance measures accuracy, recall-micro (μ), recall-macro (M), specificity-micro (μ) and specificity-macro (M). The performance measurements specified in (1)–(5) pertain to the first plastic type; in the same manner as other types.

$$\text{accuracy}=\frac{\sum_{j=1}^{3}\frac{TP_j+TN_j}{TP_j+FP_j+FN_j+TN}}{3} \tag{1}$$

$$\text{recall}_\mu=\frac{\sum_{j=1}^{3}TP_j}{\sum_{j=1}^{3}(TP_j-FN_j)} \tag{2}$$

$$\text{recall}_M=\frac{\sum_{j=1}^{3}\frac{TP_j}{TP_j+FN_j}}{3} \tag{3}$$

$$\text{specificity}_\mu=\frac{\sum_{j=1}^{3}TN_j}{\sum_{j=1}^{3}(FP_j-TN_j)} \tag{4}$$

$$\text{specificity}_M=\frac{\sum_{j=1}^{3}\frac{TN_j}{FP_j+TN_j}}{3} \tag{5}$$

The greater the values of these metrics, the superior the classification performance of the classification method.

Let Y be a multinomial distributed random variable with parameter (j, π), where j is several trials, and π is the probability of each trial. Variable Y takes a particular value $y=(y_1,y_2,\ldots,y_j)$. The y has a probability density function:

$$f(y)=\frac{j}{y_1!\,y_2!\cdots y_j!}\pi_1^{y_1}\pi_2^{y_2}\cdots\pi_j^{y_j} \tag{6}$$

For modelling the classification of the plastic type via multinomial logistic regression, let the predictor variable be X with the number of p, and the target variable with a multinomial distribution with j category is as follows:

$$Y=\begin{cases}1 \text{ if plastic type is PET with probability } \pi_1(x)\\ 2 \text{ if plastic type is HDPE with probability } \pi_2(x)\\ 3 \text{ if plastic type is PP with probability } \pi_3(x)\end{cases}$$

The probability $\pi_j(x)=P(y=j|x)$, j=1,2,3 and $\sum_{j=1}^{3}\pi_j(x)=1$. Let the last category be the reference in the model, where probability $\pi_j(x)$ for each j with logit function $g_j(x)=\beta_{0j}+\beta_{1j}x_{1j}+\cdots+\beta_{pj}x_{kj}$ can be defined as

$$\pi_1(x)=\frac{e^{g_1(x)}}{1+e^{g_1(x)}+e^{g_2(x)}} \tag{7}$$

$$\pi_2(x)=\frac{e^{g_2(x)}}{1+e^{g_1(x)}+e^{g_2(x)}} \tag{8}$$

$$\pi_3(x)=\frac{1}{1+e^{g_1(x)}+e^{g_2(x)}} \tag{9}$$

The parameter β in the logit function can be estimated using maximum likelihood estimation and Newton–Raphson. For this research, the likelihood function derived from observations under the assumption of independence between each pair of observations can be written as follows:

$$l(\beta)=\prod_{i=n}^{n}f(y_i)=\prod_{i=n}^{n}\pi_1(x_i)^{y_{1i}}\pi_2(x_i)^{y_{2i}}\pi_3(x_i)^{y_{3i}} \tag{10}$$

TABLE I
K-FOLD CROSS-VALIDATION RESAMPLING TECHNIQUE FOR K=5

Composition	Fold				
	1	2	3	4	5
1					
2					
3					
4					
5					

modeling, testing

TABLE II
CONFUSION MATRIX FOR PLASTIC-TYPE, J = 1

	j	Actual		
		1	2	3
Classification	1	True-Positive (TP)	False-Negative (FN)	False-Negative (FN)
	2	False-Positive (FP)	True-Negative (TN)	True-Negative (TN)
	3	False-Positive (FP)	True-Negative (TN)	True-Negative (TN)

TABLE III
RESEARCH DATA

Type	Predictor Variable														
	Red			Green			Blue			Entropy			Variance		
	(X_1)			(X_2)			(X_3)			(X_4)			(X_5)		
	min	mean	max	min	mean	max	min	mean	max	min	mean	max	min	mean	max
PET	0	0.45	0.99	0.13	0.49	1	0.12	0.56	0.99	0.03	0.76	0.95	0	0.14	0.98
HDPE	0	0.90	1	0	0.93	1	0.03	0.91	1	0	0.24	0.99	0	0.85	1
PP	0.01	0.58	1	0.11	0.60	1	0	0.61	1	0.19	0.63	1	0	0.21	0.87

TABLE IV
UNITS FOR MAGNETIC PROPERTIES

Data	Resampling				
Testing	Fold				
	1	2	3	4	5
PET	34	27	26	36	27
HDPE	29	29	31	27	34
PP	27	34	33	27	29
Sum	90	90	90	90	90
Learning	Except Fold				
	1	2	3	4	5
PET	116	123	124	114	123
HDPE	121	121	119	123	116
PP	123	116	117	123	121
Sum	360	360	360	360	360

TABLE V
SIMULTANEOUS TEST

$-2ln(l_0)$	$-2ln(l_k)$	G	$X^2_{(0.05;10)}$	p-value
790.78	423.81	366.98	18.31	0.00

TABLE VI
GOODNESS OF FIT TEST

deviance	DF	p-value
423.81	694	1.00

TABLE VII
PARAMETER ESTIMATION

Plastic type		β	SE (β)	Wald	p-value	Exp (β)
PET	Intercept	2.93	1.23	5.64	0.02	
	X_1	-5.79	1.41	16.80	0.00	0.00
	X_2	-11.62	2.47	22.22	0.00	8.98×10^{-6}
	X_3	-1.08	1.34	0.65	0.42	0.34
	X_4	6.12	1.26	23.69	0.00	455.94
	X_5	14.77	2.51	34.66	0.00	2.61×10^{6}
HDPE	Intercept	2.83	1.61	3.07	0.08	
	X_1	-7.64	2.37	10.42	0.00	0.00
	X_2	-17.96	3.60	24.85	0.00	1.58×10^{-8}
	X_3	-1.60	1.77	0.82	0.37	0.20
	X_4	6.76	1.85	13.40	0.00	860.26
	X_5	26.98	3.40	63.13	0.00	5.21×10^{9}

TABLE VIII
THE BEST MODEL OF MULTINOMIAL LOGISTIC REGRESSION

Plastic type (j)	π_i (x)
PET	$\dfrac{\exp(2.93-5.79x_1-11.62x_2+6.12x_4+14.77x_5}{1+\exp(2.93-5.79x_1-11.62x_2+6.12x_4+14.77x_5}$
HDPE	$\dfrac{\exp(2.83-7.64x_1-17.96x_2+6.76x_4+26.98x_5}{1+\exp(2.83-7.64x_1-17.96x_2+6.76x_4+26.98x_5}$

TABLE IX
PERFORMANCE COMPARISON OF PLASTIC TYPE CLASSIFICATION

Research	Classification method	Measurement				
		Accuracy	Recall-μ	$Recall_M$	$Specificity_\mu$	$Specificity_M$
Yani et al., 2021 [8]	FDA	87.11	91.67	80.97	90.33	90.38
Yani & Resti, 2024a [22]	MNB	97.34	96.00	96.07	98.00	98.02
Yani & Resti, 2024b [23]	DTID3	82.74	74.11	74.68	87.06	87.22
Proposed method	MLR	86.08	79.11	79.08	89.56	89.59

(a) (b) (c)

Fig. 1. Image sample for each plastic type: (a) PET, (b) HDPE, and (c) PP

The parameter β in the logit function can be estimated using maximum likelihood estimation and Newton–Raphson. For this research, the likelihood function derived from observations under the assumption of independence between each pair of observations can be written as follows:

$$l(\beta) = \prod_{i=n}^{n} f(y_i)$$
$$= \prod_{i=n}^{n} \pi_1(x_i)^{y_{1i}} \pi_2(x_i)^{y_{2i}} \pi_3(x_i)^{y_{3i}} \quad (10)$$

The logarithm of Equation (10) is given in Equation (11):

$$L(\beta) = \ln[l(\beta)]$$
$$= \sum_{i=1}^{n} y_{1i} \ln[\pi_1(x_i)] + y_{2i} \ln[\pi_2(x_i)] + y_{3i} \ln[\pi_3(x_i)] \quad (11)$$

The Newton–Raphson method was employed to derive an explicit solution for the second differential L(β) to β:

$$\frac{\partial^2 L(\beta)}{\partial\beta_{3k}\partial\beta_{3k'}} = \sum_{i=1}^{n} x_{1i}x_{2i}x_{3i}\pi_1(x_i)\pi_2(x_i)\pi_3(x_i) \quad (12)$$

To simultaneously assess the significance of the entire impact of the independent variable on the dependent variable, we utilised the G test. Let l_0 be the likelihood without an independent variable, and l_k be the likelihood with independent variables with k=1,2,…,p. The G test is written as

$$G = -2 \ln\left[\frac{l_o}{l_k}\right] \quad (13)$$

The null hypothesis posits that there is no statistically significant independent variable that influences the dependent variable ($\beta_1 = \beta_2 =…= \beta_p=0$). By contrast, the alternative hypothesis asserts that at least one independent variable significantly impacts the dependent variable (at least one ($\beta_k \neq 0$). Reject criteria for the null hypothesis when $G>X^2_{a,df}$, where the degree of freedom (DF) is the number of independent variables.

$\hat{\beta}_k$ is the estimator of β_k, and SE ($\hat{\beta}_k$) is the standard error of $\hat{\beta}_k$ with $k = 1,2,\ldots p$. The null hypothesis posits that there is no partial influence of an independent variable on the dependent variable ($\beta_k = 0$), whereas the alternative hypothesis suggests that such a variable exerts a partial influence ($\beta_k \neq 0$). The null hypothesis is rejected when the p-value is less than α, where α represents the significance level. To analyse the relationship between a dependent variable and an independent variable among individuals, we calculated Wald statistics for partial testing, as follows:

$$W_k = \left(\frac{\hat{\beta}_k}{SE(\hat{\beta}_k)}\right)^2 \quad (14)$$

For the goodness-of-fit test using deviance, let $\hat{\mu}_{ij} = n_i\hat{\pi}_{ij}$, and j is the number of classes in target variable. The deviance is written as

$$deviance = -2\sum_{i=1}^{n}\sum_{j=1}^{J} y_{ij} \log \frac{y_{ij}}{\hat{\mu}_{ij}} \quad (15)$$

Utilising the null hypothesis as the model is appropriate. The criterion by which the null hypothesis is to be rejected is deviance>X^2(α;(n-p)(r-1)) or p-value<α.

III. Result and Discussion

The dataset of this research comprised 450 images, with 150 images allocated to each type of plastic: polyethylene terephthalate (PET/PETE), high-density polyethylene (HDPE) and polypropylene (PP). These plastic types are commonly utilised in society and possess the capacity to degrade into waste. Sample images for each plastic type for this study are presented in Figure 1.

The digital image of this dataset was obtained from the three different types of plastic and subsequently processed into the RGB colour format. Each colour component was represented in this format by eight bits, resulting in a scale of 28 = 256 or a pixel value range of 0–255. RGB colour components, entropy and variance were predictor variable values normalised to the interval (0.1). The target variable was the plastic type, which was denoted by Y. The predictor variables were represented by X_d, d=1,2,…,5. A summary of research data for this study is presented in Table 3.

HDPE had the highest mean across all variables, except for the variable where PET had the highest mean. By contrast, PET had the lowest mean for all variables, except for the variable where HDPE had the highest mean. The maximum values for HDPE and PP were greater than PET for all variables, except for one variable where PP had the lowest maximum value.

Table 4 details the composition of the training and test datasets. The data were divided into five folds of comparable size through a random process [34], [35], [36]. The test data comprised one fold of the computation for each fold, whilst the training data comprised the remaining four folds.

Tables 5–8 present the results of testing, estimation and the best model for the first resampling, where fold 1 was test data and the remaining folds provided the training data. For the other four, resampling was carried out in the same way. To determine whether the independent variables significantly influence the dependent (target) variable concurrently, we performed a simultaneous test, as presented in Table 4. If the G value was greater than $X^2_{(0.05:10)}$, then the null hypothesis could be rejected. Each image characteristic influenced the plastic type.

As shown in Table 6, the goodness-of-fit test by deviance indicated that the null hypothesis could not be rejected because the p-value exceeded the predetermined significance level. The obtained model could be utilised with a confidence level of 95%.

Table 7 presents the parameter estimation for the initial data composition, encompassing the partial test and odds ratio. In all datasets, we utilised the PP plastic-type class as a reference for modelling purposes. The table indicated that all of the image features within each plastic type class significantly affected the model, except for the pixel value of the blue variable (X_3). The relative tendency of classifying a digital image as a plastic type of PET decreased by 5.79 pixels when the pixel value of red (variable X_1) increased by one and the pixel values of all other variables remained constant. This relationship applied to digital images of plastic types of PET. In the given scenario where all other variables remained constant and the pixel value of the green (variable X_1) in the plastic type of PET increased by one, the relative risk or tendency of the digital image being classified as belonging to the PET class decreased by 11.62 pixels. The significance of the estimated parameters of other variables and parameters pertaining to other plastic types was determined.

The odds ratio of the green pixel value (variable X_2) at 8.98×10^{-6} in the plastic type of PET indicated that a digital image had a smaller tendency at 8.98×10^{-6} times to be classified into the PET type compared with the PP plastic type based on the variable. A digital image had a greater tendency at 455.94 times to be classified into the PET type as compared with the PP type based on entropy (variable X_4). Likewise, the interpretation of the odds ratio for other significant variables and the odds ratio. Table 8 presents the optimal model for the initial dataset composition of k-fold cross-validation.

The aggregate performance of the proposed multinomial logistic regression classification model using fivefold cross-validation was equal to the mean performance of the five sets of models. Table 9 shows the average performance of the proposed model, consisting of accuracy, recall-µ, recall-M, specificity-µ and specificity-M.

The proposed model showed that 86.08% of digital images of plastic waste were accurately classified across all categories of plastic. This model also succeeded in classifying the number of observations in a class, which effectively comprised the classified class. In particular, 79.11% pertained to the number of decisions made for each object, and 79.08% referred to the average decisions made per class. This model exhibited a degree of accuracy in classifying a digital image observation that did not correspond to the correct type of plastic compared with the complete dataset of plastic images that were also incorrect types. The number of decisions made for each object and the average of decisions made by each class were 89.56% and 89.59%, respectively.

Table 9 compares the performance of the model proposed in this paper with the performance of other models. All of these studies used the fivefold cross-validation technique as a resampling technique.

MNB had the highest performance measure compared with the others, whilst the lowest used the DTID3 method. However, the accuracy with DTID3 was unsatisfactory [39], [40]. The lowest accuracy obtained by the DTID3 method [23] was due to the inaccurate transformation of numerical variables into categorical variables. The transformation to categorical variables in the DT method aimed to improve the performance of the classification model. Inappropriate interval limits may result in low accuracy. The performance of the method proposed in this paper was not the best compared with the classification performance using statistical methods such as multinomial naïve Bayes [22] and Fisher discriminant analysis [8]. Nevertheless, the performance of the proposed method was better than that of the decision tree method with the ID3 algorithm [23]. Further research is needed to achieve enhanced plastic-type classification performance.

IV. Conclusion

The cost-effectiveness of digital images has made them increasingly popular over the past decade for object identification and classification, mainly when converted to RGB colour space models. In machine learning, multinomial logistic regression is a classification technique that is frequently effective in various contexts. Implementing this method using the fivefold cross-validation technique to classify three plastic types based on digital images transformed into RBG colour space models revealed satisfactory performance. The performance average measures of fivefold cross-validation were 86.08% accuracy, 79.11% recall-µ, 79.08% recall-M, 89.56% specificity-µ and 89.59% specificity-M. Compared with the classification performance achieved with statistical methods like multinomial naïve Bayes and Fisher discriminant analysis, the performance of the proposed method in this paper was not optimal. Nevertheless, the results of this study revealed an improvement over the performance achieved with the decision tree with the ID3 algorithm. Additional research is required to improve the efficacy of plastic-type classification.

Acknowledgment

Thanks to Dr. Dendy Adanta from the Department of Mechanical Engineering, Universitas Sriwijaya, Indonesia, for help in writing and reviewing this manuscript.

References

[1] Srigul, W., Inrawong, P., & Kupimai, M. (2016). Plastik classification base on the correlation of RGB color. 2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, ECTI-CON 2016, 1–5. https://doi.org/10.1109/ECTICon.2016.7561304

[2] OECD, 2022. Policy Highlights. Global Plastics Outlook Economic Drivers, Environmental Impacts and Policy Options.

[3] Abubakar. I. R., Maniruzzaman, K. M., Dano, Ul. L., AlShihri, F. S., AlShammari, M. S., Ahmed, S. M. S., Al-Gehlani, W. A. G., and Alrawaf, T. I. 2022. Environmental Sustainability Impacts of Solid Waste Management Practices in the Global South. Int. J. Environ. Res. Public Health 19, 12717. https://doi.org/10.3390/ijerph191912717

[4] Chow C, So WMW, and Cheung T-Y, 20.16. Research and development of a new waste collection bin to facilitate education in plastic recycling," Journal Applied Environmental Education & Communication Volume 15, 2016 - Issue 1

[5] Pivnenko K, Eriksen MK, Martín-Fernández JA, Eriksson E, & Astrup TF, 2016, Recycling of plastic waste: Presence of phthalates in plastics from households and industry. Waste Management, 54, 44–52. https://doi.org/10.1016/j.wasman.2016.05.014

[6] Ruj B, Pandey V, Jash P, and Srivastava VK. 2015, Sorting of plastic waste for effective recycling, International Journal of Applied Sciences and Engineering Research, Vol. 4, Issue 4.

[7] Yani, I., Firmansyah, B., Resti , Y., Arnas, Y., Kartika, R. B., Nasution, T. M. R, Hendro, W., and Endrawijaya, I. 2021. Automatic Identification of Plastic Waste by HSV Colour. Lecture Notes in Mechanical Engineering, Human-Centered Technology for a Better Tomorrow, Springer.

[8] Yani, I., Resti, Y., Burlian, F., and Ansyori, 2021, Prediction of Plastic-Type for Sorting System using Fisher Discriminant Analysis", Science and Technology Indonesia, 6(4), pp. 313–318.

[9] Ngugi, L.C.; Abelwahab, M.; Abo-Zahhad, M. Recent Advances in Image Processing Techniques for Automated Leaf Pest an Diseas Recognition—A Review. Inf. Process. Agric. 2021, 8, 27–51

[10] Yani I, Rosiliani D, Khonaáh B, Almahdini FA. 2020, Identification and plastic-type and classification of PET, HDPE, and PP using RGB

method, IOP Conference Series: Material Science and Engineering 857 012015
[11] Wu, Y., He, Y., Wang, Y. 2023. Multi-Class Weed Recognition Using Hybrid CNN-SVM Classifier. Sensors 2023, 23, 7153. https://doi.org/10.3390/s23167153
[12] Fan, J., Lee, J. H., Lee, Y. K. 2021. A Transfer Learning Architecture Based on a Support Vector Machine for Histopathology Image Classification. Appl. Sci. 2021, 11, 6380. https://doi.org/10.3390/app11146380
[13] Mishra, A., Chaurasia, P., Arya, V., & Jos´e Garc´ıa Pe˜nalvo, F. (2023). Plant Disease Detection using Image Processing. In International Conference on Cyber Security, Privacy and Networking (ICSPN 2022) (pp. 227-235). (Lecture Notes in Networks and Systems; Vol. 599 LNNS). Advance online publication.https://doi.org/10.1007/978-3-031-22018-0_21
[14] Resti, Y., Irsan, C., Neardiaty, A., Annabila, C., and Yani, I. Fuzzy Discretization on the Multinomial Naïve Bayes Method for Modeling Multiclass Classification of Corn Plant Diseases and Pests. Mathematics, vol. 11, no.1761, 2023. https://doi.org/10.3390/math11081761
[15] Domingues, T.; Brandão, T.; Ferreira, J.C. Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive Survey. Agriculture 2022, 12, 1350. https://doi.org/10.3390/agriculture12091350
[16] Resti, Y.; Irsan, C.; Amini, M.; Yani, I.; Passarella, R.; Zayanti, D.A. Performance Improvement of Decision Tree Model using Fuzzy Membership Function for Classification of Corn Plant Diseases and Pests. Sci. Technol. Indones. 2022, 7, 284–290. https://doi.org/10.26554/sti.2022.7.3.284-290
[17] Haque, M.A.; Marwaha, S.; Deb, C.K.; Nigam, S.; Arora, A.; Hooda, K.S.; Soujanya, P.L.; Aggarwal, S.K.; Lall, B.; Kumar, M. et al. Deep learning-based approach for identification of diseases of maize crop. Sci. Rep. 2022, 12, 1–14. https://doi.org/10.1038/s41598-022-10140-z
[18] Almadhor, A.; Rauf, H.T.; Lali, M.I.U.; Damaševičius, R.; Alouffi, B.; Alharbi, A. Ai-driven framework for recognition of guava plant diseases through machine learning from dslr camera sensor based high resolution imagery. Sensors 2021, 21, 3830. https://doi.org/10.3390/s21113830
[19] Yousefi, S., Mirzaee, S., Almohamad, H., Al Dughairi, A. A., Gomez, C., Siamian, N., Alrasheedi, M., and Abdo, H. G. 2022. Image Classification and Land Cover Mapping Using Sentinel-2 Imagery: Optimization of SVM Parameters. Land 2022, 11, 993. https://doi.org/10.3390/land11070993
[20] Razaque, A., Frej, M. B. H., Almi'ani, M., Alotaibi, M., and Alotaibi., B. 2021. Improved Support Vector Machine Enabled Radial Basis Function and Linear Variants for Remote Sensing Image Classification. Sensors 2021, 21, 4431. https://doi.org/10.3390/s21134431.
[21] Senekane, M. 2019 Differentially Private Image Classification Using Support Vector Machine and Differential Privacy. Mach. Learn. Knowl. Extr. 2019, 1, 483–491; doi:10.3390/make1010029
[22] Yani & Resti, 2024a. Plastic-Type Prediction Based on Digital Image using Multinomial Naïve Bayes Method, AIP Proceeding of Conference on Fundamental and Applied Science for Advanced Technology
[23] Yani & Resti, 2024b. Performance of Plastic-Type Prediction using Decision Tree Approaches. AIP Proceeding of Sriwijaya International Conference on Basic and Applied Sciences
[24] Khona'ah B, Rosiliani, Yani I. 2019, Identification and Classification of Plastic Color Images based on the RGB Method. Journal of Multidisciplinary Engineering Science and Technology, 6(6), 10170 – 10174
[25] Tzougas, G., Kutzkov, K. 2023. Enhancing Logistic Regression Using Neural Networks for Classification in Actuarial Learning. Algorithms 2023, 16, 99. https://doi.org/10.3390/a16020099
[26] Awad, F.H., Hamad, M. M., and Alzubaidi., L. 2023. Robust Classification and Detection of Big Medical Data Using Advanced Parallel K-Means Clustering, YOLOv4, and Logistic Regression. Life 2023, 13, 691. https://doi.org/10.3390/life13030691
[27] Nopour R, Shanbehzadeh, Kazemi-Arpanahi H (2022). Using logistic regression to sdevelo[a diagnostic model for covid-19: A single-center study. Journal of Education and Health Promotion. http://www.jehp.net on Tuesday, November 1, 2022, IP:30.72.197.125
[28] Resti, Y., Desi, H.S., Zayanti, D.A., Eliyati, N. (2022c). Classification of Diseases Aand Pests of Maize using Multinomial Logistic Regression Based on Resampling Technique of K-Fold Cross-Validation. Indones. J. Eng. Sci. 3, 69–76
[29] Sari PS, Noya van Delsen MS, Lesnussa YA, Binary Logistic Regression Model to Identify Factors Associated with Low Birth Weight, Case Study: Baby Data at Dr. M. Haulussy Hospital Ambon. Barekeng: Journal of Mathematics and Its Application, 16:3, pp. 985-994. http://doi.org/10.30598/barekengvol16iss3pp985-994
[30] Rymarczyk, T., Kozłowski, E., Kłosowski, G., and Niderla, K. 2019. Logistic Regression for Machine Learning in Process Tomography. Sensors 2019, 19, 3400; doi:10.3390/s19153400
[31] Resti, Y., Kresnawati, E., S., Dewi, N. R., Zayanti, D.A., Eliyati, N., 2021. Diagnosis of Diabetes Mellitus in Women of Reproductive Age using The Prediction Methods of Naïve Bayes, discriminant Analysis, and Logistic Regression. Science and Technology Indonesia, Vol.6, No.2, April, pp. 96-104
[32] Resti , Y., Mohruni, A. S., Burlian, F., Yani, I., and Amran., A. 2017. A probability approach in cans identification. MATEC Web of Conferences 101, 03012. DOI: 10.1051/matecconf/201710103012.
[33] James G, Witten D, Hastie T, Tibshirani R. 2013, An Introduction to Statistical Learning with Application in R, Springer
[34] Rodrıguez, J. D., A. Perez, and J. A. Lozano (2009). Sensitivity analysis of k-fold cross validation in prediction error estimation. IEEE transactions on pattern analysis and machine intelligence, 32(3); 569–575
[35] Bengio, Y. and Y. Grandvalet (2004). No unbiased estimator of the variance of k-fold cross-validation. Journal of machine learning research, 5; 1089–1105
[36] Kuhn, M., Johnson, K., 2013. Applied Predictive modelling, Springer
[37] Dinesh, S., Dash, T. 2016. Reliable Evaluation of Neural Network for Multiclass Classification of Real-world Data. Arxiv, 1612.00671v1, 30 Nov
[38] Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. Information Processing and Management, 45(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002
[39] Aronoff, S. 1982. Classification accuracy: a user approach. Photogramm. Eng. Remote Sensing, vol. 48, no. 8, pp. 1299–1307.
[40] Aronoff, S. 1985. The Minimum Accuracy Value as an Index of Classification Accuracy. Photogrammetric Engineering and Remote Sensing, Vol. 51, No. 1, January, Pp. 99-111.

Irsyadi Yani (Corresponding author) was born in South Sumatera, Indonesia on December 25th 1971. Degree: B.Eng. (1996) in Mechanical Engineering from Universitas Sriwijaya, South Sumatera - Indonesia. M.Eng. (2004) in Mechanical Engineering, Toyohashi University of Technology, Japan. Ph.D. (2015) in Faculty of Engineering, Universiti Kebangsaan Malaysia, Malaysia.
He has been working as a lecturer for the Department of Mechanical Engineering, Faculty of Engineering, Universitas Sriwijaya, South Sumatera - Indonesia. Irsyadi Yani, B.Eng., M.Eng., Ph.D. have publication had written in: Science and Technology Indonesia; Mathematics; Biomedicines; Engineering Letters; CFD Letters; Membranes; Lecture Notes in Mechanical Engineering; ARPN Journal of Engineering and Applied Sciences; and several seminar proceedings.
E-mail corresponding: irsyadiyani@ft.unsri.ac.id

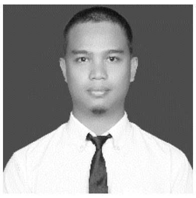
Ismail Thamrin was born in South Sumatera, Indonesia on September 2nd 1972. Degree: B.Eng. (1996) in Mechanical Engineering from Universitas Sriwijaya, South Sumatera - Indonesia. M.Eng. (2004) in Mechanical Engineering, from Universitas Sriwijaya, South Sumatera - Indonesia. Dr. (2020) in Engineering Sciences, Universitas Sriwijaya, South Sumatera - Indonesia.
He has been working as a lecturer for the Department of Mechanical Engineering, Faculty of Engineering, Universitas Sriwijaya, South Sumatera, Indonesia. Dr. Ismail Thamrin, B.Eng., M.Eng. have publication had written in: Journal of Advanced Research in Fluid Mechanics and Thermal Sciences; International Journal of Thermofluids; Civil Engineering Journal; CFD Letters; and several seminar proceedings.

Dewi Puspitasari was born in South Sumatera, Indonesia on January 15th 1970. Degree: B.Eng. (1993) in Mechanical Engineering from Universitas Sriwijaya, South Sumatera - Indonesia. M.Eng. (2000) in Mechanical Engineering from Universitas Gajah Mada, Yogyakarta - Indonesia. Dr. (2013) in Mechanical Engineering, Universitas Gajah Mada, Yogyakarta - Indonesia.

She has been working as as a lecturer for the Department of Mechanical Engineering, Faculty of Engineering, Universitas Sriwijaya, South Sumatera, Indonesia. Dr. Dewi Puspitasari, B.Eng., M.Eng. have publication had written in: Frontiers in Heat and Mass Transfer; CFD Letters; Journal of Mechanical Engineering and Sciences; International Journal of Rotating Machinery; and several seminar proceedings.

Barlin was born in South Sumatera, Indonesia on June 30th 1981. Degree: B.Eng. (2004) in Mechanical Engineering from Universitas Sriwijaya, South Sumatera - Indonesia. M.Eng. (2009) in Mechanical Engineering from Universitas Gajah Mada, Yogyakarta - Indonesia. Dr. (2020) in Mechatronic Science and Technology, Southern Taiwan University of Science and Technology - Taiwan.
He has been working as a lecturer for the Department of Mechanical Engineering, Faculty of Engineering, Universitas Sriwijaya, South Sumatera, Indonesia. Barlin, B.Eng., M.Eng., Ph.D. have publication had written in: Journal of Advanced Research in Micro and Nano Engineering; Journal of Advanced Research in Fluid Mechanics and Thermal Sciences; Science and Technology Indonesia; CFD Letters; International Journal of Advanced Manufacturing Technology; International Journal of Technology; and several seminar proceedings.

Yulia Resti was born in South Sumatera, Indonesia on July 19th 1973. Degree: B.Sc. (1996) in Department of Mathematic from Universitas Sriwjaya, South Sumatera-Indonesia. M.Si. (2003) in Actuary, Institut Teknologi Bandung, West Java - Indonesia. Dr. (2013) in Statistic, Universiti Kebangsaan Malaysia, Malaysia.
She has been working as a lecturer for the Department of Mathematic, Faculty of Mathematic and Natural Science, Universitas Sriwijaya, South Sumatera, Indonesia. Prof. Yulia Resti, B.Si., M.Si., Ph.D. have publication had written in: IAES International Journal of Artificial Intelligence; Mathematics Education Journal; Science and Technology Indonesia; Symmetry; Mathematics; International Journal of Applied Mathematics and Statistics; Journal of Mathematics and Statistics; and several seminar proceedings.



Dendy Adanta was born in South Sumatera, Indonesia, on June 5th, 1993. Degree: M.Eng. (2017) in Mechanical Engineering from Universitas Indonesia, Jakarta - Indonesia. Dr. (2020) in Mechanical Engineering, Universitas Indonesia, Jakarta - Indonesia.
He has been working as a lecturer for the Department of Mechanical Engineering, Faculty of Engineering, Universitas Sriwijaya, South Sumatera, Indonesia. Dr. Dendy Adanta, M.Eng. some publication had written in: International Journal of Technology; International Journal on Advanced Science, Engineering and Information Technology; Civil Engineering Journal; CFD Letters; Journal of Advanced Research in Fluid Mechanics and Thermal Science; Energy Reports; Journal of Mechanical Engineering and Sciences; Engineering Letter; International Energy Journal; International Review of Mechanical Engineering; International Journal on Energy Conversion; International Journal of Fluid Machinery and Systems; Journal of the Brazilian Society of Mechanical Sciences and Engineering; International Journal of Thermofluid; Renewable Energy Focus; Malaysian Journal of Medicine and Health Sciences; Journal of Advanced Research in Numerical Heat Transfer; and several seminar proceedings.