Artificial Neural Network For Estimation Nutrient Utilization Based on Chemical Composition on Ruminant Animal Feed

Toto Haryanto[1,2*]
[1]*Department of Computer Science IPB University*
Bogor, 16680 Indonesia
[2]*Faculty of Computer Science Universitas Indonesia*
Depok ,16424 Indonesia
totoharyanto@apps.ipb.ac.id

Dias Febrisahrozi
Department of Computer Science IPB University
Bogor, 16680 Indonesia
dias.febrisahrozi@gmail.com

Anuraga Jayanegara
Department of Nutrition and Feed Technology IPB University
Bogor, 16680 Indonesia
anuragaja@apps.ipb.ac.id

Aziz Kustiyo
Department of Computer Science IPB University
Bogor, 16680 Indonesia
azizku@apps.ipb.ac.id

***Abstract*— The production of ruminant livestock is used to meet the needs of human food. The productivity of ruminants is determined by the feed given. Therefore, a proper feeding strategy is needed. This research aims to estimate the nutrient utilization of ruminant forage plants based on chemical content using Artificial Neural Network (ANN). Chemical content used as the estimator is the main analysis category; meanwhile, ruminant nutritive values are predicted from our model. A total of 136 species of forage plants were obtained from *http://www.feedipedia.org*. Data were divided into training and testing data using K-fold cross-validation with k=3. The process of estimation was performed using the ANN structure with several variation numbers of neurons in the hidden layer from 4, 6, and 8. As a result, the ANN structure built could estimate the nutrient properly, indicated by small values of RMSE (3.33) revealed from ANN structure with four neurons. The determination coefficient (R^2) from this ANN model more than 0.50 for all output, except in the production of Metabolizable energy, which was 0.49.**

Keywords— Artificial Neural Network, forage, nutrient, ruminant

I. Introduction

Humans need the production of ruminant livestock, especially to meet food needs, for example, meat and milk. The protein contained in meat is essential because it contains all essential amino acids, including those containing S minerals that are not possessed by vegetable protein and are needed for growth. In addition, beef is also a significant source of minerals Ca, P, Zinc, Fe, and vitamins B2, B6, and B12 [1]. Because of the importance of livestock products for humans, the action needs to be taken to maintain the quality and quantity of ruminant livestock production, namely by providing suitable feed and containing balanced nutrition.

The nutritional quality of animal feed is a very important factor in efforts to improve the productivity and performance of ruminants. Feed consumption by livestock will be used to fulfill basic living needs, and if the energy produced from feed exceeds the basic living needs, the animal will use the excess for growth and production [2]. So, although it comes from superior seeds and has superior genetic traits, it is not balanced with proper quality feeding, the various advantages will not provide significant added value.

The chemical content of feed plants as livestock consists of four categories, namely main analysis, minerals, amino acids, and secondary metabolites. Among the components influence the accumulation of nutrients contained in these plants [3].

In the field of animal husbandry, general feed given to livestock is sourced from plants. Ruminant animal feed plant Species, along with variations in chemical content, are numerous. This makes it very difficult for farmers to choose the right type of feed. During this time, farmers provide feeding based on experience and habits, so feeding in this way certainly cannot guarantee the optimal value of nutrient intake for livestock.

Currently, the process of making animal feed formulas is still input based. Livestock is fed with a number of nutritional parameters. From these nutrients, it is hoped that livestock will be able to get a nutritional value comparable to that contained in the feed. However, the assumption is not always true. Not all of the chemical contents in the feed are absorbed by ruminants. Efficient livestock productivity can be achieved by providing food efficiently, in accordance with the purpose of its maintenance [4].

Some modern breeders have adopted feeding methods by focusing on the nutritional value actually received by livestock, rather than focusing on the nutritional content of the feed. However, farmers have difficulty finding the right feed formulation to obtain optimal nutrient values for their animals. So far, the steps taken have been to try various formulations of feed to livestock directly, then analyze the nutrient values in animals in the laboratory. In an effort to help farmers determine a suitable feed, it is necessary to estimate nutrient values based on the chemical components of the feed.

In animal husbandry, the artificial neural network (ANN) has a potential role for animal science applications as mention in [5]. ANN has success in predicting optimal livestock feed blend for animals in terms of price, nutrient, and water [6]. Another research related to ANN in the field of animal production is for predicting the breeding value of Kermani sheep [7]. A similar study was carried out by Fadare and Babayemi [8] to estimate livestock gas production due to certain forage feed using an Artificial Neural Network method with an input layer in the form of the main analysis value contained in plants.

In this study, the method used is an Artificial Neural Network (ANN). The choice of method is because the relationship between the chemical composition and nutrient value is not a linear problem. In addition, the contribution of the ANN architecture model is multi-output ANN. So all nutrient components will be predicted in one ANN structure. ANN is designed so that it gets results quickly by using a set of weights for all nutrient elements. The estimation results are used as a guide for preparing feed according to the nutritional needs of ruminants.

The benefit of this research is to be able to assist livestock feed makers in preparing feed as well as for extension workers in considering farmers related to ruminant animal feeding. By implementing selection based on estimated nutrient utilization, ruminant livestock productivity can be optimized.

II. METHODOLOGY

Our research was carried out through several stages, starting from data selection, pre-processing, training to test as presented in Fig. 1

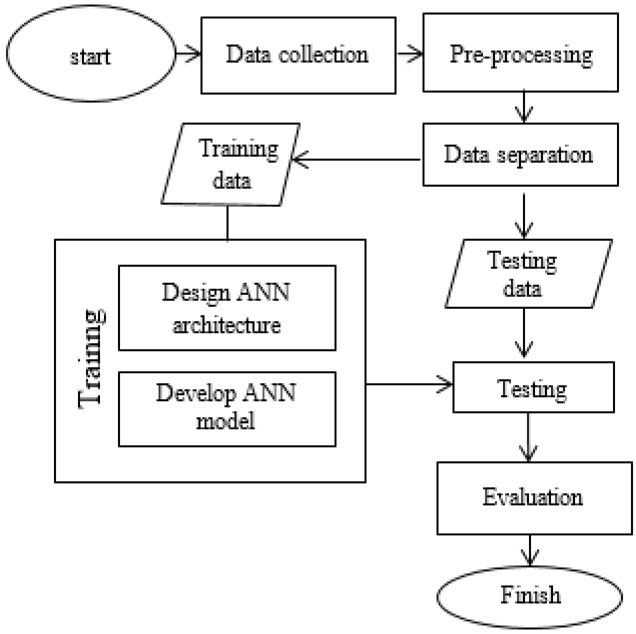


Fig 1. Research flowchart

A. Data Collection

This research utilizes secondary data compiled from animal feed data on the website http://www.feedipedia.org, which is the world animal feed database [9]. Data on the chemical content and nutrient value analyzed were taken from all categories of ruminant animal feed plants. Only plant species included in the database have been investigated for their chemical substance in the laboratory, marked by the "finished" attribute in the "Completion status" filter. A total of 193 datasets are collected.

B. Pre-processing

Pre-processing of data is done by removing components that do not significantly affect the accumulation of nutrients for livestock or only for certain species. Besides, removal is also carried out for species that contain an outlier value. Data that has been cleared of noise is then normalized using the minimum-maximum method with the formula (1) refers to [10].

$$z = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

After pre-processing, we obtain 136 data records for the training process. List of the datasets can be requested on github: https://github.com/totoharyantoui/annanimalnutrien.

C. Data Separation

The formed animal feed nutrition database is divided into training data and test data using the K-fold cross-validation method [11]. Training data is data that will be used as input for the ANN model, while test data is used to verify the estimation results. The scheme of K-Fold cross-validation shown in Table I.

TABLE I. K-FOLD SCHEME FOR DATASET

Fold th	Data th		
	1 - 45	46 - 90	91 - 136
1	training data	training data	**testing data**
2	training data	**testing data**	training data
3	**testing data**	training data	training data

D. ANN Training

The designed ANN architecture consists of an input layer, one hidden layer, and one output layer with four neurons. Visually the architecture is presented in Fig 2.

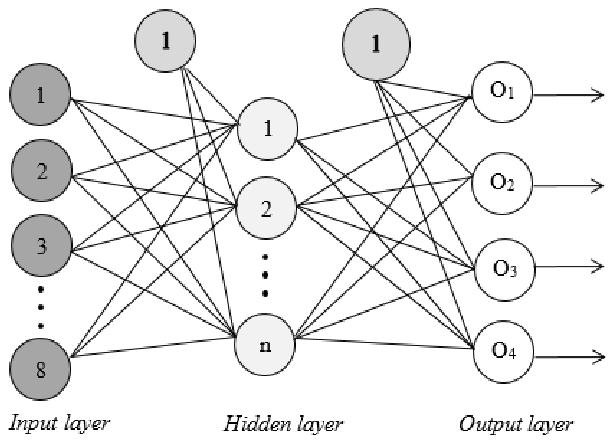


Fig 2. Design of ANN architecture

The training process is conducted by defining the parameters of an ANN structure first. The input layer consists of eight neurons representing chemical components in plants, including crude protein, crude fiber, neutral-detergent fiber (NDF), acid detergent fiber (ADF), lignin, crude fat (ether extract), ash (ash), and gross energy (gross energy). The output layer consists of four neurons that represent the four estimated components of animal nutrition for animal feed, namely organic matter digestibility, energy digestibility, digestible energy, and metabolizable energy. The input of ANN is the main analysis that can be found on plants.

We use two training functions on the ANN structure, namely Levenberg-Marquardt (LM) and Scaled Conjugate Gradient (SCG). The transfer function used in the hidden layer is the sigmoid-binary function, while the output layer is a linear function. The training was conducted for several variations in the number of hidden layers, from 4, 6, and 8 neurons. This scenario uses to find out the ANN structure that produces the best estimation of the tested data. The detailed scenario of the ANN structure in this study is shown in Table II.

TABLE II. DETAIL OF ANN STRUCTURE

ANN structure	Information
Input layer	8 neuron
Hidden layer	1 layer with 4, 6 , and 8 neuron
Output layer	4 neuron
training function	*Levenberg-Marquardt* (LM) *Scaled Conjugate Gradient* (SCG)
Transfer function	*Hidden layer* : sigmoid-biner *Output layer* : linear
Bias	1 bias per *layer*
Epoch	1000 (maximum)

The input dataset is main analysis of plant for livestock. Main analysis is the main compound in plants in term of animal feed. The chemical composition of real data before cleaning and pre-processing presented in Table III.

TABLE III. DETAIL OF INPUT DATASET BEFORE CLEANING

Input	Unit	Min	Max
Dry matter	% as fed	0	95
Crude protein	% DM	0	32
Crude fiber	% DM	0	72
NDF	% DM	0	89
ADF	% DM	0	74
Lignin	% DM	0	22
Ether extract	% DM	0	15
Ash	% DM	3	25
Gross energy	MJ/kg DM	14	22
Water-soluble carbohydrates	% DM	1	14
Starch (enzymatic)	% DM	1	8
Starch (polarimetry)	% DM	2	50
Total sugars	% DM	2	29

After the pre-processing step, cleaning and normalization are applied to the data, we only use eight as input of the ANN architecture. The new dataset after cleaning and pre-processing is described in Table IV. The explanation about pre-processing was in discussion section.

TABLE IV. FINAL INPUT DATASET

Input	Unit	Min	Max
Crude protein	% DM	0	1
Crude fiber	% DM	0	1
NDF	% DM	0	1
ADF	% DM	0	1
Lignin	% DM	0	1
Ether extract	% DM	0	1
Ash	% DM	0	1
Gross energy	MJ/kg DM	0	1

Ruminants nutritions value will be determined by the main analysis consumption of the animal. Proper feed will result in good nutritional intake for livestock. There are four parameter that completely absorbed by ruminants based on main analysis input. Table V illustrates the output of ANN architecture.

TABLE V. RUMINANT NUTRIENT AS THE OUTPUT OF ANN

Output	Unit
Organic matter digestibility	(%)
Energy digestibility	(%)
Digestible Energy	(MJ/kg)
Metabolizable Energy	(MJ/kg)

E. Testing and Evaluation

The testing is applied to chemical content data, which has been divided into training data and test data using K-fold cross-validation. ANN testing is based on the K scheme shown in Table I. Testing will be conducted on the number of neurons from 4, 6, and 8; both are using the Levenberg-Marquardt (LM) and Scaled Conjugate Gradient (SCG) training functions. The results of each test are recorded and compared with the results of laboratory observations for evaluation.

ANN performance in the estimation process is measured based on the level of error and accuracy. Testing errors are measured using Root Mean Square Errors (RMSE). RMSE is the root of the average square of the difference between the actual value and the estimated value. RMSE value is said to be good if the value is close to 0, which means there are only a few estimation errors. The RMSE calculation [12] is formulated in (2)

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{n}(X_t - F_t)^2}{n}} \quad (2)$$

Where X is the actual value for data, and F is the prediction value of the model for t=1 to n number of the data. The results of experiments with the smallest RMSE values are then calculated on the coefficient of determination. This is to measure the strength of the relationship with the value of laboratory observations. The coefficient of determination is said to be good if the value is close to 1 [13]. The coefficient of determination is formulated in (3)

$$R^2 = \frac{\left[\sum_{i=1}^{n}(\hat{y}_1 - \bar{\hat{y}})(y_i - y)\right]^2}{\sum_{i=1}^{n}(y_i - \bar{\hat{y}})^2 \sum_{i=1}^{n}(y_1 - \bar{y})^2} \quad (3)$$

Where R is a coefficient of determination, y is actual values, and ŷ is prediction values.

III. RESULT AND DISCUSSION

A. Data Collection

Database of chemical contents of animal feed and nutrient values compiled from the website http://www.feedipedia.org. The chemical content of fodder plants is taken from the main analysis table, while the nutrient values received by ruminants are taken from the ruminant nutritive values table.

B. Pre-processing

At this stage, the elimination of species whose chemical component values and/or nutrient content is incomplete is eliminated. The elimination process is carried out on species that have a chemical component of outliers. As a part of the pre-processing is the removal of some chemical components and nutrient content of fodder plants. From the main analysis table, five columns of chemical components must be removed. The dry matter column is removed because the columns of other chemical components represented the composition. Elimination is applied to chemical elements that are only owned by a few species of fodder plants, namely water-soluble carbohydrates, starch (enzymatic), starch (polarimetry), and sugar.

In the ruminant nutritive values table, there are six columns of nutrient content removed because a few species of fodder plants only own them. These components are Nitrogen digestibility, a(N), b(N), c(N), Nitrogen degradability (effective, k = 4%), and Nitrogen degradability (effective, k = 6%). From the elimination process that has been done, found 136 species that are ready to be used as training data and test data.

The chemical components contained in the main analysis table have different units and ranges of values. The unit of value for crude protein (crude protein), crude fiber (crude fiber), neutral-detergent fiber (NDF), acid detergent fiber (ADF), lignin, crude fat (ether extract), and ash (ash) is the percentage in dry matter (% DM). The gross energy (gross energy) has a unit of MJ / kgDM. Also, the size of the concentration of the main analysis component does not necessarily represent its contribution in influencing the output. For example, crude fat, which is a relatively small value, but has a significant role in shaping the value of ruminant nutritive values.

C. ANN Training

In our research, the hardware specifications used for this study are the Intel Core i5 M450 2.4 GHz, 4GB memory, 500 GB hard drive. We use MS Excel 2010 for cleaning and pre-processing data. Meanwhile, nnet library from Matlab 2008b is used to design, train, and evaluate neural network architecture [14].

D. Evaluation by RMSE

The nutrient estimation results are calculated the difference with the results of laboratory tests. This difference value is squared to get the MSE value. From the MSE that has been obtained, the roots are drawn to obtain RMSE. The average RMSE of each fold dataset is presented in Table VI.

TABLE VI. RMSE AVERAGE OF THE 3-FOLD DATASETS

# neuron *hidden layer*	***Training function***	
	***Levenberg-Marquardt* (LM)**	***Scaled Conjugate Gradient* (SCG)**
4	4.23	3.78
6	4.77	4.29
8	4.67	4.49

ANN performance is relatively stable for all experimental scenarios. This is indicated by the small difference in the value of RMSE from each experiment. From all the experiments that have been revealed, the smallest average RMSE of three fold is 3.78 obtained from ANN using SCG training function with four neurons in the hidden layer. The detailed RMSE test results using the LM and SCG training functions from each fold dataset are shown in Table VII.

TABLE VII. DETAIL RMSE AVERAGE OF THE THREE FOLD DATASETS

***Fold*th**	**# Neuron *hidden layer***	***Training Function***	
		***Levenberg-Marquardt* (LM)**	***Scaled Conjugate Gradient* (SCG)**
1	4	4.78	4.28
	6	4.91	4.74
	8	5.08	5.20
2	4	3.97	3.73
	6	5.25	4.52
	8	4.48	4.36
3	4	3.93	3.33
	6	4.15	3.62
	8	4.44	3.90

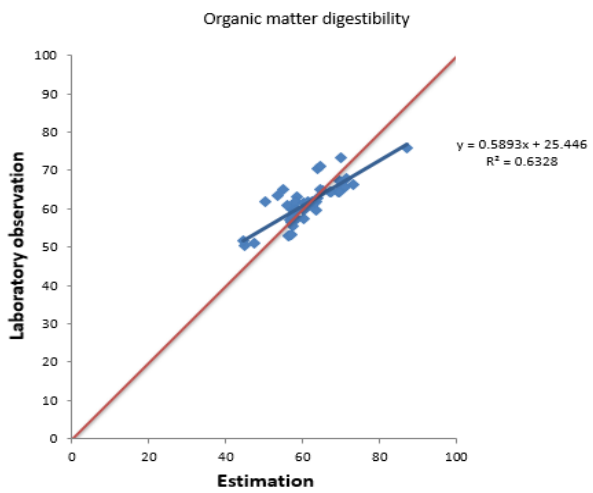
The smallest RMSE value was obtained from the ANN model with four hidden layer neurons that implemented the SCG training function using the 3rd fold dataset with RMSE value of 3.33. The ANN model with the smallest RMSE is then measured its coefficient of determination (R^2). This is to see the strength of the relationship between the estimated nutritional value of feed with laboratory observations. The coefficient of determination (R^2) is calculated for each estimated nutritional value. The results of the calculation of the coefficient of determination (R^2) are presented in Table VIII.

TABLE VIII. THE COEFFICIENT OF DETERMINATION FOR THE ANN MODEL WITH THE SMALLEST RMSE

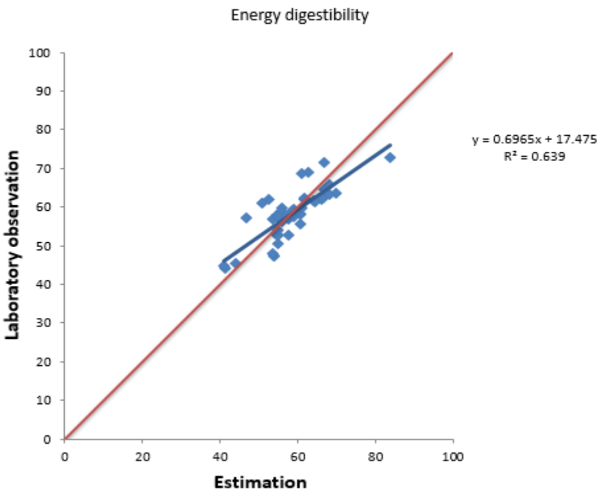
The nutritional value of feed	**coefficient of determination (R^2)**
Organic matter digestibility	0.63
Energy digestibility	0.64
Digestible energy	0.63
Metabolizable energy	0.49

The calculation result of the coefficient of determination (R^2) shows a relatively constant value for the four estimated outputs. The coefficient of determination (R^2) is around 0.63, except Metabolizable energy (ME) ruminants, which is 0.49. These results are consistent with the approach of animal science in determining the value of ME ruminants that require further experiments in the laboratory so that the results are accurate. Comparison graphs between estimated results using ANN and laboratory observations are visualized in Fig 3

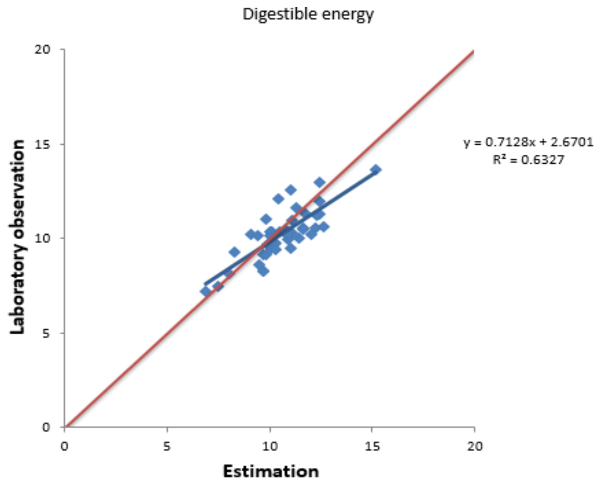
.



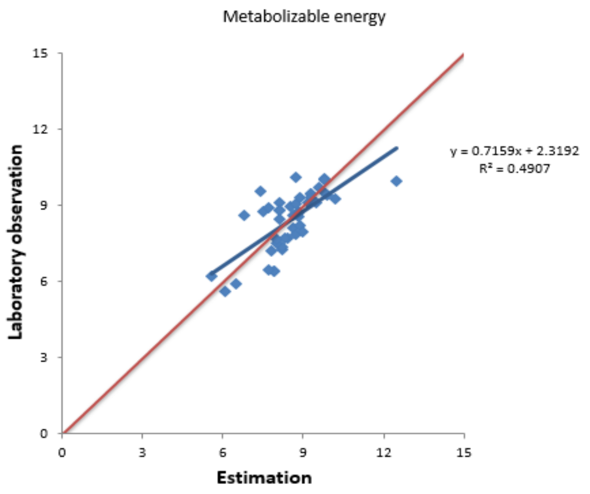
(a)



(b)



(c)



(d)

Fig 3. Graph of Comparison R^2 between estimated data using ANN and laboratory observations: (a) Organic matter digestibility, (b) Energy digestibility, (c) Digestible energy, and (d) Metabolizable energy

Fig. 3 shows the deviation from the estimated value of the actual value in the ANN model with the smallest RMSE. The point above the identity line indicates that the estimated nutrient value is lower than the nutrient value from laboratory observations. On the other hand, the score below the identity line indicates that the estimated nutrient value is higher than the nutrient value from laboratory observations.

Based on the graphs in Fig 3, we obtain the plot equation results for each estimated nutrient value. The further the deviation from the estimation line, it means that the estimation model is not successful. Conversely, if the estimation line coincides with the identity line, we can say that the model is successful to estimate. From the four graphs in Fig. 3, the line deviation looks relatively small from the identity line. This shows that the model of ANN good enough to estimate. To increase the value of R^2, some additional scenarios such as exclude some variables [15] can be able as the future direction of this research.

IV. Conclusion

This research has succeeded in designing an ANN model that can estimate nutrient values well, indicated by the relatively small RMSE value. The smallest RMSE was obtained from the ANN model with the SCG training function with four hidden layer neurons, with an RMSE value of 3.33. Our ANN model shows good accuracy, shown by the coefficient of determination (R^2) for the estimated nutrient value of more than 0.50; except for the fourth output (Metabolizable energy) which is worth 0.49. In terms of application, this ANN model can be used by feed mills to formulate feed suitable for livestock. In addition, it is also possible to predict the nutrients that livestock will get if they are fed with certain feed ingredients (plants). Additional datasets and deep neural network (DNN) architecture also become potential research in the future.

References

[1] C. T. D. Y. Guntara, “Beef Cattle Production in Supporting Indonesian Food Safety and Security,” in Seminar Nasional Teknologi Peternakan Dan Veteriner, 2008, no. 17, pp. 44–51.

[2] A. Setyawan, K. Setyaningsih, G. Mahesti, E. Rianto, Sunarso, and A. Purnomohadi, “Different Proportion of Muscle, Fat and Bone in Thin Tailed Sheep Fed at Maintenance and Production Level),” in Seminar Nasional Teknologi Peternakan dan Veteriner 2008, 2008, pp. 395–400.

[3] D. Givens, E. Owen, R. F. Axford, and H. Omed, "Forage Evaluaion in Ruminat Nurtition", New York: CABI Publishing, 2000.

[4] H. Hartadi, S. Reksohadiputro, S. Lebdosukojo, and A. D. Tillman, Table of Feed Composition for Indonesia, no. November 1980. Utah: International Feedstuffs Institute Utah Agricultural Experiment Station, Utah State University, 1980.

[5] C. Fernández, E. Soria, J. D. Martín, and A. J. Serrano, “Neural networks for animal science applications: Two case studies,” Expert Syst. Appl., vol. 31, no. 2, pp. 444–450, 2006, doi: 10.1016/j.eswa.2005.09.086.

[6] P. Saxena and Y. Parasher, “Application of artificial neural network (ANN) for animal diet formulation modeling,” Procedia Comput. Sci., vol. 152, pp. 261–266, 2019, doi: 10.1016/j.procs.2019.05.018.

[7] H. Ghotbaldini, M. Mohammadabadi, H. Nezamabadi-Pour, O. I. Babenko, M. V. Bushtruk, and S. V. Tkachenko, “Predicting breeding value of body weight at 6-month age using artificial neural networks in Kermani sheep breed,” Acta Sci. - Anim. Sci., vol. 41, no. 1, pp. 1–9, 2019, doi: 10.4025/actascianimsci.v41i1.45282.

[8] D. A. Fadare and O. J. Babayemi, “Modelling the association between in vitro gas production and chemical composition of some lesser known tropical browse forages using artificial neural network,” African J. Biotechnol., vol. 6, no. 18, pp. 2184–2192, 2007.

[9] G. Tran, V. Heuzé, D. Bastianelli, H. Archimède, and D. Sauvant, “Tables of nutritive value for farm animals in tropical and Mediterranean regions: an important asset for improving the use of

local feed resources," Adv. Anim. Biosci., vol. 1, no. 2, 2011, doi: 10.1017/s2040470011002615.

[10] C. Saranya and G. Manikandan, "A study on normalization techniques for privacy preserving data mining," Int. J. Eng. Technol., vol. 5, no. 3, pp. 2701–2704, 2013, doi: https://pdfs.semanticscholar.org/35a8/7b51f7441a87adee91e12eb4d22cd2565556.pdf.

[11] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," Int. Jt. Conf. Artif. Intell., no. 0, pp. 0–6, 1995.

[12] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature," Geosci. Model Dev., vol. 7, no. 3, pp. 1247–1250, 2014, doi: 10.5194/gmd-7-1247-2014.

[13] D. Zhang, "A Coefficient of Determination for Generalized Linear Models," Am. Stat., vol. 71, no. 4, pp. 1–20, 2017, doi: https://doi.org/10.1080/00031305.2016.1256839.

[14] R. Walpole and R. Myers, "Probailidad y Estadistica para Ingenieros," Prentice Hall, pp. 1–752, 1999, doi: 10.1017/CBO9781107415324.004.

[15] Supriyanto et al., "Artificial neural networks model for estimating growth of polyculture microalgae in an open raceway pond," Biosyst. Eng., vol. 177, pp. 122–129, 2019, doi: 10.1016/j.biosystemseng.2018.10.002.