

Multi-output Hybrid GA-NN with Adaptive Mechanism

Faridah Sh Ismail

Malaysian Institute of Information Technology
Universiti Kuala Lumpur
Kuala Lumpur, MALAYSIA
faridah@miit.unikl.edu.my

Nordin Abu Bakar

Faculty of Computer & Mathematical Sciences
Universiti Teknologi MARA
Shah Alam, MALAYSIA
nordin@tmsk.uitm.edu.my

Abstract— This research presents a hybrid Genetic Algorithm Neural Network (GA-NN) model to simulate the physical tests procedures of Medium Density Fiberboard (MDF). Data included in the model are related to MDF properties and its fiber characteristics. Multilayer Perceptron NN is a reliable supervised machine learning model. The model learns from seven inputs fed to the network to produce prediction of three targets. In order to avoid result from local optimum scenario, GA optimizes synaptic weights of the network towards reducing prediction error. The research used a fixed probability rates for crossover and mutation for hybrid GA-NN model. GA-NN model is further improved using adaptive mechanism to help identify the best probability rates for crossover and mutation. Fitness value refers to Sum of Squared Error, which is the accumulation of network error in the Output Layer. The population fitness distribution will guide best rates for each epoch. Performance comparisons are among three models; namely NN with Back Propagation (BP), hybrid GA-NN and hybrid GA-NN with adaptive mechanism. Results show the hybrid GA-NN model perform much better than NN model with back propagation optimizer. Adaptive mechanism in GA helps increase capability to converge at zero sooner than the ordinary GA.

Keywords— *neural network; genetic algorithm; adaptive; prediction; hybrid model; MDF*

I. INTRODUCTION

Machine learning is a learning paradigm in computer science. The first kind of machine learning is the empirical learning, which accomplishes general rules or procedures by reasoning supplied examples. The two types of empirical learning are supervised learning and unsupervised learning. Supervised learning learns from previous trend to produce classification or prediction, while unsupervised learning creates new knowledge through concept clustering. One of the approaches in supervised learning is Artificial Intelligence (AI). It uses the nature of human intelligence and adapting it in computer science applications to solve various areas of automation problems. Among others, Neural Networks (NN)

and Genetic Algorithm (GA) are AI techniques well accepted in supervised learning environments.

A. Neural Network

NN is a classical tool for many predictive applications. It is a better pattern-learning instrument than decision trees and Naïve Bayes [1]. This method is a popular choice in making predictions in several areas, including in agriculture-based areas. To name a few, NN has been used for weather predictions [2], stock market forecasting [3] and medical diagnosis [4]. In the area of agricultural modelling, NN is used to maximize yield based on best combination of oil palm fertilizer by [5] and for lemon grass oil by [6]. NN approach has also implemented in engineered wood industry such as MDF predictions. Areas of research include moisture resistance by [7] and IB strength by [8]. Destruction time in laboratory can be reduced through IB prediction based on process parameters [9] and [10]. Physical properties too were discovered to be able to predict IB [11].

Multilayer perceptron (MLP) NN has at least three layers; Input Layer, Hidden Layer and Output Layer; connected by vectors (Fig. 1). The number of neurons in input layer simply refers to the number of inputs and as such, the number of neurons in output layer should refer to targets needed from the model. Different consideration needed for hidden layer where it can have more than one layers. However, one layer is more than sufficient [12] and proven able to produce excellent results [13]. Being a supervised learning machine, the network learns from previous knowledge; which is the input feed through the Input Layer. Next, Hidden Layer will process the input to match the targets in the Output Layer.

The training data used for learning will provide the validation task whereas data for testing is for verification. Training and testing patterns selected at random usually has 70:30 proportions. The network vectors contain weights, while each node in Hidden Layer and Output Layer contain bias values. Values for weights and biases are initially set with random. An optimizer later adjusts these values to suit the target values. The popular weight optimizer in NN is Back Propagation (BP) algorithm.

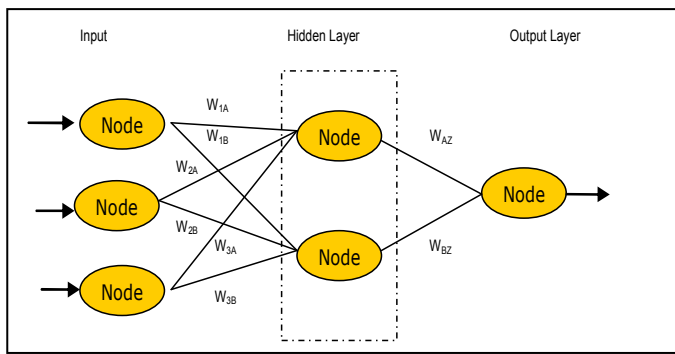


Fig. 1. Basic Multilayer Perceptron Neural Network Architecture

The BP algorithm is suitable and popular for supervised learning. The name back propagation comes from the nature of the training session on how data are input several times back to the hidden layer in search of best set of weights that produce closest prediction. This repeats until exceed certain number of epochs or until no more reduction in the SSE is reported.

There are several issues involved in designing an MLP network. Capability of NN to learn is very much depends on the topology, weight values, and transfer function including parameter values during training [14]. Learning rate and momentum parameter are applied to increase the sensitivity of BP to speed and robustness [15]. The learning rate value should not be constant; instead, it needs to evolve according to the number of iteration. The higher the iteration, the smaller it becomes. While momentum parameter should be within 0 and 1. Having too large a value will affect training capability.

Apart from that, BP algorithm has the possibility of being stuck in local minima. To overcome that limitation, NN replace BP algorithm with GA optimizer. The choice of GA is due to its searching ability towards global optimum.

B. Hybrid Genetic Algorithm Neural Network

Adapting from human genetic, GA consists of set of chromosomes in a population. Each chromosome contains random values representing genes. GA optimizes its chromosomes similar to how human reproduction works; through its reproduction operators, crossover and mutation. The purpose of reproduction is to produce new chromosomes; searching for better population. This is possible for offspring inherits good genes from the fittest parents chosen from the population. The chances of reproduction depend on the operators' probability rate, whereby, a higher rate has higher chances. However having too high a rate may cause searching to miss the optimum point. Therefore setting suitable rates is important.

Crossover and mutation operators contribute towards the quantity as well as the quality of offspring. The operators' probability rates controls the chances of reproduction, and the techniques used influence better offspring. For a non-fixed

rate, the changes are either with a constant rate or with a variable rate. A variable rate adapts with the current population environment and therefore increases the searching ability.

C. Hybrid GA-NN with Adaptive Mechanism

Most researchers [16][17][18] have agreed that the performance of GA mainly depend on the operators, the probability of the operators used, and the diversification of the population. The common practice is that the parameters are set at the beginning and the process will run according to these parameters until a good solution is found. Since GA is modelled after the evolution process, the longer the runs take place, the results will get better. However, in most cases GAs fail to produce good results in an acceptable period. Therefore, a new variant of GA should be available to give good results in a shorter time. By adapting one or several components of GAs, the quality of the results will be improved because the GA will quickly adapt its parameters to work with the problem at hand. The adaptation will provide a good environment to nurture good individuals and at the end of the process will give better results. Finding a good combination is a hard task [17][19][20]. By adjusting the values during the run and taking current performance as the index, the values for the parameters will be more reliable. Once the GA is running with optimum parameter values, the time taken to reach global optimum will be reduced considerably.

Previous researchers have suggested to focus on the values for GAs parameters (such population size, crossover rate and mutation rate) that worked well in their research [16][17][19]. The main drawback is that the problem solved might be different and the solutions found may be well below the optimum. Adaptation can also be found extensively in other works [21] [22].

D. Medium Density Fiberboard

Medium Density Fiberboard (MDF) is an engineered wood used in furniture industry as an alternative to solid wood. Prior to manufactured, it is compulsory to run for mechanical and physical tests procedures to ensure its quality. The purpose is to ensure that its formation is reliable; mechanically and physically according to the standard requirements. Mechanical tests are to determine the board strength capability. The properties are modulus of rupture (MOR), modulus of elasticity (MOE) and internal bonding (IB). The physical tests, on the other hand are for board resistance against water; whereby the properties are thickness swelling (TS), water absorption (WA) and moisture content (MC). These mechanical and physical tests consumed as short as 2 hours and as long as 48 hours per sample, conforming to the British standard (BS EN). In total, the procedures need several days to complete.

The main concern in these test-procedures is the time factor [23][24]. For that, this research aim to replace the lengthy tests with a simulation by a prediction model. All

other properties from testing results will be the input. Adding two more factors on fiber characteristics are significant and an advantage towards the prediction [25].

This paper reports a study on producing an intelligent prediction model, using hybrid GA-NN. Section II draws the methodology, begins with data analysis, techniques used by NN as the main architecture of the model and techniques used by the hybrid GA-NN. Adaptive mechanism enhances the GA search ability by guiding operators to suit the current population. Section III presents the results of all models and discusses the achievement of each. Finally, the paper ends with a conclusion and recommendation on future work.

II. METHODOLOGY

The main architecture is multilayer perceptron NN. The NN performance depends very much on the weights and therefore GA with an adaptive mechanism will help increase the model reliability. The multi-output prediction model contains the lengthy test properties, identified as targets; based on other tests, identified as the predictors.

A. Data

The data analysis was on the experimental tests data obtained from Malaysian Palm Oil Board (MPOB) Research Station. The data were taken from each of the tests done on mechanical and physical properties. The board density and the mixture percentage of oil palm biomass-rubber wood are two additional variable added to the data compilation. Ten variables with 192 observations involved in the compilation and analysis. Each observation is the average value of three test results done on the similar sample.

Data is stored in a Comma-separated Values (CSV) format. Firstly, missing values and extreme outliers are cleaned by replacing them with the average values. Secondly, rescaling of data is using normalized rescaling technique (Eq. 1) to reduce data large gap to a range between 0.0 and 1.0.

$$normalized = (x - min)/(max - min) \quad (1)$$

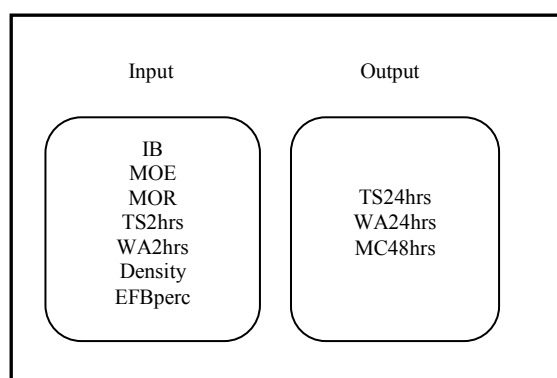


Fig. 2. Input and Output Variables

Eq. 1 is normalization scaling equation where x is the actual value, while min and max are the minimum and maximum values among all observations of that particular variable. Figure 2 shows variables used for input and output. The first five input variables are properties of non-lengthy procedures and the last two are fiber characteristics.

B. Neural Network Model

The first model formulates NN prediction using Multilayer Perceptron, which also known as multilayer feed forward. Figure 3 depicts the neural network architecture used by this research containing 7-node input layer, 4-node hidden layer and 3-node output layer. The adjoining lines between layers, called vectors, contain values, called weights. These weights determine the importance of each input towards the prediction of targets. The input layer corresponds to the attributes obtained for each training sample and fed to the hidden layer with the weighted values and consequently forwards results to output layer. The termination condition is set to 1000 epochs to allow improvements on training session. Training session by Back Propagation (BP) improves weight values in backwards manner several times until meet termination condition.

Network use Eq. 2 (for *OutputError*) as the output error calculation for each target in each pattern. This error considers the prediction gap as well as the proportion of prediction. Overall evaluation handles negative error by calculating Sum of Squared Error (SSE) for all patterns (Eq. 3). Finally, produces Root Mean Squared Error (RMSE) for each epoch by dividing with number of patterns and number of targets in output layer (Eq. 4).

$$OutputError = pred(1 - pred)(target - pred) \quad (2)$$

$$SSE = \sum (OutputError)^2 \quad (3)$$

$$RMSE = \sqrt{SSE / (N * 3)} \quad (4)$$

Parameters set for NN model with BP algorithm is as in Table 2. Among important parameter in BP are learning rate and momentum coefficient. Research is using an inverse value of epoch for learning rate; the higher the iteration, the smaller the learning rate is. Momentum coefficient has constant 0.5.

TABLE 1. SPECIFICATION FOR NN MODEL

Parameter	Value
Activation Function	Sigmoid
Hidden Nodes	4
Learning Rate	1/epoch
Momentum Rate	1.0
Training:Testing	80:20

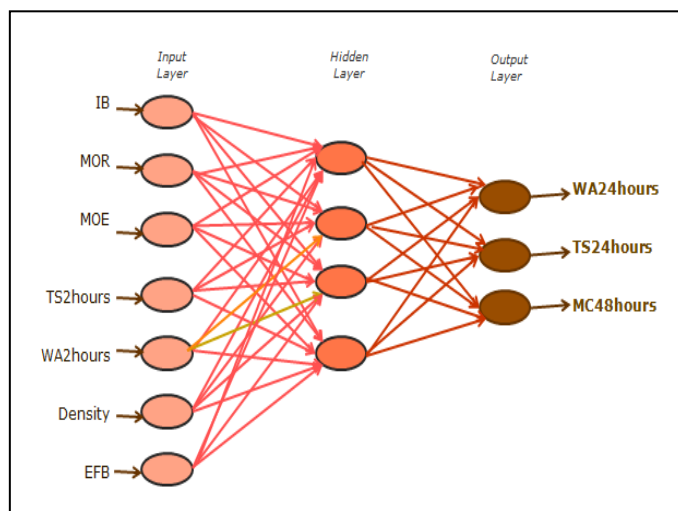


Fig. 3. A 7-4-3 NN model

Sigmoid activation function is suitable for data with positive values. Referring to [25], four hidden nodes is suggested for this network topology.

C. Hybrid Genetic Algorithm Neural Network Model

Classical NN uses BP as the weight optimizer. Embedding GA in NN will replace BP and optimizes network weights better.

Having a 7-4-3 network topology, meaning there are 40 synaptic weights and 7 biases created. Each weight and bias represents an allele in a chromosome. The chromosome contains real valued data in the range of -1.0 to 1.0. During the first generation, allele has random values obtained from initial population. Fitness value refers to Eq. 3. After fitness evaluation is done, population is sorted in ascending order, whereby the lowest error is the fittest and placed at the top of array of population. With that, parents are easily determined through rank selection for reproduction task. The probability rates for crossover (Pc) and mutation (Pm) is fixed at 0.3 and 0.01 respectively. Each training sample is iterated for 50 generations in search of best set of weights from 100 populations. Summary of GA-NN model is in Table 2. All the above steps repeat for 1000 epochs.

TABLE 2. SPECIFICATION FOR GANN MODEL

Parameter	Value
Pc	0.3
Pm	0.01
Alpha value for crossover	0.5
Generation:population	50: 100
Chromosome size	47
Fitness evaluation	SSE

TABLE 3. SPECIFICATION FOR GANN WITH ADAPTIVE MECHANISM

Parameter	Value
Adaptive operator	Pc, Pm
Pc	0.3
Pm	0.0145
Adaptive measure	Mean, Median

D. Hybrid Genetic Algorithm Neural Network Model with Adaptive Mechanism

The hybrid model is improved with an adaptive influence on the fitness performance of previous population. The adaptive mechanism is to ensure operators hold the best probability rates for reproduction based on the changes in population. Adaptive measure looks into the mean and median of fitness in current population and suggests on operators' updates significantly. When mean fitness lies smaller than median, increase Pc (Eq. 5) and decrease Pm (Eq. 6). The opposite takes place when the mean value is higher than median. The final Pm is 0.0145 while Pc remains unchanged at 0.3.

$$\text{Increase } Pc = 0.1 * (1/\text{epoch}) \tag{5}$$

$$\text{Decrease } Pm = 0.01 * (1/\text{epoch}) \tag{6}$$

III. RESULTS

The prediction models tested on the data are the NN model using BP algorithm and hybrid GA-NN models.

A. Results of Neural Network

Figure 4 plots the RMSE of 1000 epoch. The convergence was at 0.04 and not going any lower. This is a good prediction model, even better result as compared to result by [25]. The model has early convergence as early as epoch 50, however it seemed to stuck there until end of cycle. This local minima situation is an expected problem and usual scenario when using BP.

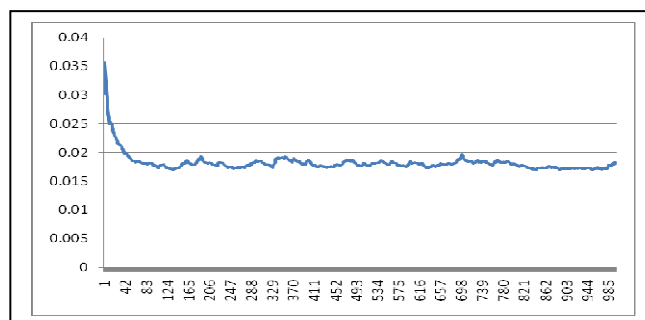


Fig. 4. RMSE for NN model

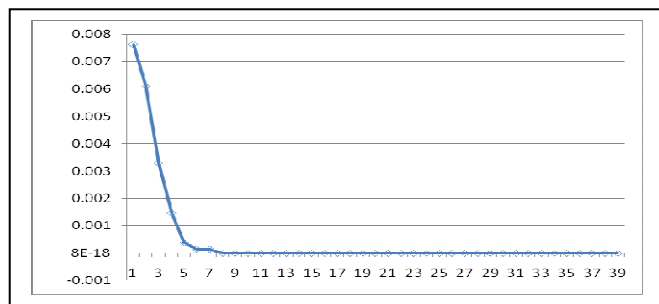


Fig. 5. RMSE for GANN model

B. Results of Hybrid Genetic Algorithm Neural Network

Figure 5 shows that RMSE remain between 0.001 and 0.0 for 50 epochs; then only it moves further down and converged. Even so, this is apparently better than NN model as it eventually resides at zero.

C. Results of Hybrid GANN with adaptive mechanism

Figure 5 shows that the research approach on adaptive mechanism works well with this data. RMSE slopes down and converged at epoch 10. This model has outperformed the ordinary GA by allowing a sooner convergence.

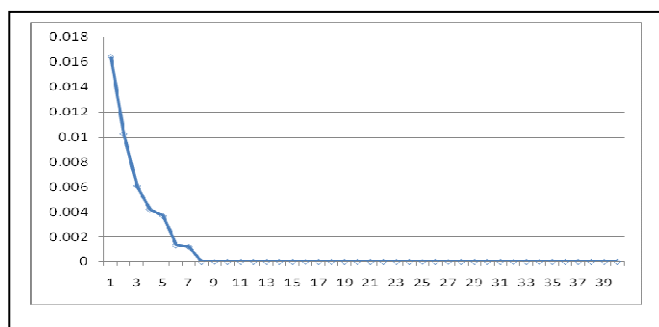


Fig. 6. RMSE for Adaptive GANN model

Table 4 summarizes the performance of all models. Both training and testing error reflects that none of the models experience overfitting; results of training explains testing data very well. Therefore, all models are reliable.

TABLE 4. ERROR COMPARISON AT EPOCH 50

Model	Training RMSE	Testing RMSE
NN with BP	0.017	0.018
Hybrid GA-NN	2.21E-97	0.0
Hybrid GA-NN with adaptive mechanism	0.0	0.0

IV. CONCLUSION

Three prediction models, BP-NN, hybrid GA-NN and adaptive GA-NN were presented and discussed. The targets are MDF testing variables; namely, TS24hrs, WA24hrs and MC48hrs. All models use NN as the main architecture. The results of prototype implementation were drawn. Parameter adjustments and considerations have contributed in producing multi output from each model. In order to make a reasonable comparison among performance of the models, the same network topology is applied to all models. Obvious improvement was seen when GA replaces BP in the NN model. Different methodologies applied to individual models have shown excellent results.

MDF plants require perfection in quality testing. With the results drawn, manufacturers can reduce testing time by reducing lengthy procedures while maintaining the quality standard set by BS EN. Furthermore, a multi-output model contributes to reduce simulation time.

For future research, this problem domain will be tested using other GA adaptive measures.

ACKNOWLEDGEMENTS

The authors would like to thank Ropandi Mamat and Suchirah Ratmin of Malaysian Palm Oil Board (MPOB) for supplying the data used in this research.

REFERENCES

- [1] A. A. Radwan, H. M. El-Bakry, H. M. El-Hadad, "A New Expert System for Repository Diseases by Using Neural Networks", *Recent Researches in Applied Informatics WSEAS 2011*, pp.296-306, 2011.
- [2] A. A. Munot and K. K. Kumar, "Long range prediction of Indian summer monsoon rainfall", *Journal Earth System Science*, 2007, 116(1) 73–79.
- [3] R. Lawrence, "Using Neural Networks to Forecast Stock Market Prices", University of Manitoba, 1997.
- [4] H. A. Abbass, "An evolutionary artificial neural networks approach for breast cancer diagnosis" *Artificial Intelligence in Medicine*, 2002, 25(3), 265-281
- [5] A. Khamis, Z. Ismail, K. Haron, A. T. Mohamed, "Neural Network Model for Oil Palm Yield Modeling", *Journal of Applied Sciences*, 2006, 6(2): 391-399.
- [6] K. K. Mishra, B. K. Singh, A. Punhani and U. Singh, "Combining NN and GA to Predict and To Maximize Lemon Grass Oil Production", *IEEE International Joint Conference on Computational Sciences and Optimization*, 2009.
- [7] L. G. Esteban, F. G. Fernandez, P. Palacios and B. G. Rodrigo, "Use of ANN as a predictive method to determine moisture resistance of particle and fiber boards under cyclic testing conditions", (UNE-EN 321) *Wood and Fiber Science*, 2010, 42(3), 1-11.
- [8] D. F. Cook and C. C. Chiu, "Predicting Internal Bond Strength of Particleboard utilizing a radial basis function Neural Network", *Engineering Application of Artificial Intelligence*, 1997, 10(2), 171-177.
- [9] T.M. Young, L.B. Shaffer, F.M. Guess, H. Bensmail, R.V. León, "A comparison of multiple linear regression and quantile regression for modeling the internal bond of medium density fiberboard", *Forest Products Journal* 58(4) (2008).
- [10] G. Painter, H. Budman and M. Pritzker, "Prediction of oriented strand board properties from mat formation and compression operating conditions. Part 2: MOE prediction and process optimization", *Wood Science Technology*, 2006, 40: 291–307 Springer

- [11] F. G. Fernandez, L. G. Esteban, P. D. Palacios, N. Navarro and M. Conde, "Prediction of standard Particleboard mechanical properties utilizing an artificial Neural Network and subsequent comparison with a multivariate regression model", *Investigacion Agraria: Sistemas y Recursos Forestales*, 2008, 17(2), 178-187 ISSN: 1131-7965.
- [12] J. Heaton, "Introduction to Neural Networks for Java", Second Edition, 2008
- [13] M. Sedighi and D. Afshari, "Creep Feed grinding optimization by an integrated GA-NN system", *Springer Journal Intell Manuf*, 2009, doi 10.1007/s10845-009-0243-4
- [14] D. Sarkar, "Methods to speed up error back-propagation learning algorithm", *ACM Computing Surveys*, 1995, 27(4), 519.
- [15] N. Abdul Hamid, N. Mohd Nawi, R. Ghazali, M.N. Mohd Salleh, "Improvements of Back Propagation Algorithm Performance Adaptively Changing Gain, Momentum and Learning Rate", *International Journal on New Computer Architectures and Their Applications (IJNCAA)*, 1(4):866-878, 2011.
- [16] Goldberg, David E. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, Massachusetts: Addison-Wesley, (1989).
- [17] De Jong, K. A., *Analysis of the behavior of a class of genetic adaptive systems*. Ph.D. Dissertation, Department of Computer and Communications Sciences, University of Michigan, Ann Arbor, MI, (1975).
- [18] Back, Thomas. *Self Adaption in genetic algorithms*. In *Towards a Practice of Autonomous Systems*, MIT Press, (1992), pp 263-271.
- [19] J.J. Grefenstette. *Optimization of Control Parameters for Genetic Algorithms Systems, Man and Cybernetics*, IEEE Transactions on Volume: 16, Issue: 1 Digital Object identifier: 10.1109/TSMC.1986.289288, (1986), pp 122 – 128.
- [20] L. Davis. *Adapting operator probabilities in genetic algorithms*. In J. David Schaffer, (ed.), *Proceedings of the Third International Conference on Genetic Algorithms*, San Mateo, CA: Morgan Kaufman Publishers, Inc., (1989), pp 61-69.
- [21] T. Back. *Evolutionary Algorithms in Theory and Practice*. Oxford University Press, New York (1996).
- [22] Z. Michalewicz. *Evolutionary computation: practical issues* *Evolutionary Computation*, 1996., *Proceedings of IEEE International Conference on* (1996), pp 30 – 39.
- [23] N. Andre, H.W. Choo, S.Y. Baek, M.K. Jeong, T.M. Young, "Prediction of internal bond strength in a medium density fibreboard process using multivariate statistical methods and variable selection", *Wood Science Technology*, Springer 42: (2008), pp 521-534.
- [24] T. M. Young, "Predictive Of The Physical Properties Of Wood Composites Using Genetic Algorithms", *Wood Utilization Research Project, University of Tennessee Forest Product Center*, 2004.
- [25] F. Sh Ismail, N. Abu Bakar, N. E. Abdul Khalid and R. Mamat, "Optimizing Oil Palm Fiberboard Properties Using Neural Network", *IEEE Conference on Data Mining and Optimization*, (2011), pp 271-275.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0

https://creativecommons.org/licenses/by/4.0/deed.en_US