

Comparison Of Optimization Of Algorithm Particle Swarm Optimization And Genetic Algorithm With Neural Network Algorithm For Legislative Election Result

Mohammad Badrul

Dept. Information System
STMIK Nusa Mandiri Jakarta
Jakarta, Indonesia
mohammad.mbl@nusamandiri.ac.id

Frieyadie

Dept. Information System
STMIK Nusa Mandiri Jakarta
Jakarta, Indonesia
frieyadie @nusamandiri.ac.id

Akmaludin

Dept. Information System
STMIK Nusa Mandiri Jakarta
Jakarta Indonesia
akmaludin.akm@nusamandiri.ac.id

Dwi Arum Ningtyas

Dept. Information System
STMIK Nusa Mandiri Jakarta
Jakarta, Indonesia
dwi.dnt@nusamandiri.ac.id

Daning Nur Sulistyowati

Dept. Information System
STMIK Nusa Mandiri Jakarta
Jakarta, Indonesia
dns9321@gmail.com

Nurajijah

Dept. Information System
STMIK Nusa Mandiri Jakarta
Jakarta, Indonesia
nurazizah6878@gmail.com

Abstract: General Election is one of the characteristics of a democratic country. One of the countries that embrace the democratic system is the state of Indonesia. Elections are a party of democracy in Indonesia to elect representatives of the people who will sit in parliament and provide great opportunities for the people of Indonesia to compete to appoint themselves to become members of the legislature. Research related to the election has been done by researchers is by using decision tree method or by using neural network. each method has its own weaknesses and advantages, but neural network methods can cover the weaknesses of the decision tree. The result of research using neural network method in predicting election result has accurate result value is still less accurate. In this research, we create neural network algorithm model and optimization with particle swarm optimization algorithm to increase attribute weight to all attributes or variables used, select attributes, and feature selection. whereas the Genetic Algorithm for predicting the performance of generalizations based on static properties of networks such as activation function and hidden neurons will be strong enough to find solutions. After testing with neural network algorithm to produce accurate value of 98.50% and AUC value of 0.982, further optimization done with particle swarm optimization obtained an accuracy of 98.85% and AUC value of 0.996. and then done the optimization testing with genetic algorithm obtained an accuracy value of 96.56% and AUC value of 0.925 So that both methods have a difference of accuracy that is equal to 0,35% and difference of AUC value equal to 0,14.

Keywords: *General Election, Neural Network Algorithm, Particle Swarm Optimization, Genetic Algorithm.*

I. INTRODUCTION

General Election is one of the characteristics of democracy in the world. Indonesia is one of the world's most democratic systems[25]. It is seen at the democracy party every five years that serves to elect the future leaders of the nation who will represent the people sitting in parliament through the general election. General Election is a means of implementing the sovereignty of the people within the unitary state of the Republic of Indonesia based on Pancasila and the 1945 Constitution[22].

Elections are a very important tool for the implementation of a democratic political system. Therefore, it is not surprising that many countries wishing to be called democratic states use elections as a mechanism to build legitimacy. The general election aims to elect members in parliament which are conducted under an open proportional system[22]. With a direct electoral system and a large number of parties then legislative elections provide a great opportunity for the people of Indonesia to compete to become members of the legislature. A proportional scenario is a type of parliamentary election that will represent the people in parliament. The problem with proportional selection is the difficulty of evaluating the exact number of seats (vacancies) won by each candidate. Because there is no guarantee that the ratio between the number of votes and the number of seats is an integer number[5].

Predicted results of the general election need to be predicted. Such predictions or forecasts stimulate and require a theoretical framework to explain the regularities found in the data[21]. This encourages non-political scientists to make predictions or predictions about election outcomes in the future. For economic actors, political events such as elections

can not be underestimated, as they can lead to both positive and negative risks to business continuity.

The predicted method of election results has been conducted by researchers, [5] predicted elections using the Bayesian inference method, [21] predicted using the Partido Revolucionario Institucional method, [3] research by predicting the variables that determine the election of the presidential candidate [20] predicted the election results using the Bayesian Estimator method. [17] conducted a study to predict the US presidential election using a decision tree. [18] predicted the election results with the model classification tree and neural network. [2] who predicted the election results using the neural network method.

Decision tree has advantages that have advantages in prediction because algorithm structure is easy to understand and the error rate is quite small while the weakness of decision tree algorithm is lower branch reliability becomes worse than the branch above it, decision tree produced is not optimal and can not use the sample larger [23], because it's not easy to understand the big decision tree and the problem of overfitting data can happen with the limited targeted data set.

Neural networks can solve decision tree problems because they have a non-linear prediction, have excellent performance in parallel processing and the ability to tolerate errors [27]. This is very appropriate for the characteristics of predictive data on election results in this study. Neural network is a method often used to predict the results of legislative elections because the data presented for this method must be large and non linear [8]. The most popular technique in neural network methods is the backpropagation algorithm that is widely used to solve many real-world problems by building well-trained models that show good performance in some non-linear problems [19].

The most popular neural network algorithm is the backpropagation algorithm. backpropagation algorithms have too slow convergent speeds that ultimately backpropagation algorithms are highly dependent on initial parameters such as number of inputs, hidden nodes, outputs, learning rates and network connection weights [19]. Another common problem is that there is a weakness in the need for large training data and less efficient optimizations [27]. This can be solved because the amount of training data in this research is 2268 records.

Particle swarm optimization (PSO) is an effective optimization algorithm that resolves the problems of a neural network algorithm that generally uses backpropagation algorithms [19]. Particle swarm optimization has more comparisons for feature selection and has superior performance for many optimization problems with faster and more stable convergence rates [13]. Characteristics of particle swarm optimization are social interactions that promote the sharing of information between particles that will assist in finding the optimal solution [19]. Particle swarm optimization has several parameters such as position, speed, maximum speed, acceleration of the constant and the weight of inertia. In Particle swarm optimization techniques there are several ways to optimize: increase attribute weight of all attributes or variables used, select attributes, and feature selection. While

Genetic Algorithm is a technique to predict the performance of generalizations based on static properties of networks such as activation function and hidden neurons will be strong enough to find a solution [11]. This can solve the existing problems in the neural network method of optimization resulting less than optimal.

In this research, particle swarm optimization will be applied to solve problems that occur in neural network by selecting feature on attribute weight to maximize performance of the generated model, while genetic algorithm will be applied for parameter selection in neural network that is neuron size, hidden layer and activation function which is appropriate and optimal so that the results of the legislative election predictions of DKI Jakarta are more accurate.

II. THEORETICAL REVIEW

Elections are one of the main pillars of a democratic country. The position of elections in differentiating a democratic political system or not is evident from some of the definitions put forward by some researchers. In a democratic country, elections are one of the main pillars to elect leaders who will later represent the people to sit in government.

Academic studies on democracy recognize two major categories of meaning, namely the concept of minimalism and maximalism. Minimalist or procedural democracy is imposed on political systems that carry out regular leadership changes through a free, open and independent electoral mechanism involving universal voting masses. While the concept of maximization is the implementation of elections is not enough for a political system to get the title of democracy because this concept requires respect for wider civil rights. General elections in Indonesia are divided from two parts, namely [25]:

1. Elections of new era

The election system is proportionally impure, which means the number of seat determinations is not determined by the number of people but also based on the administrative area. The election of the new order began in 1955 as the first election held in the country of Indonesia.

2. Elections of reform era

It is said to be a reformation election due to the election process in 1999 before the end of elections in 1997. The elections of the reform era began because the electoral products in 1997 were considered by governments and other institutions to be unreliable.

The electoral system of members of the People's Legislative Assembly based on the provisions of Law No. 7 of 2017 article 5 paragraph 1 of the system used in legislative elections is a proportional system with open lists, the electoral system of the Regional Representative Council is carried out by the district system representing many of the Law number 7 years 2017 article 5, paragraph 2. According to Law no. 7 year 2017, Participant election of member of representative of the people is political party participating in General Election, while election member of Regional Representative Council is individual. Electoral political parties may nominate as many

as 120 per cent of the seats contested in any democratic and open election area and may nominate candidates with regard to women's representation of at least 30%. Political Parties Participants in the General Election shall be required by law to submit a list of candidates by sequence number (to obtain a Chair). Therefore, in terms of nomination of Law No. 7 of 2017 adopted a closed candidate list system.

2.1. Neural Network Algorithm

Neural network is an information processing system that has characteristics resembling a biological neural network in humans[6]. Neural networks are defined as computational systems in which architecture and operations are inspired by the knowledge of biological neurons in the brain, which is one of the artificial representations of the human brain that always tries to stimulate the learning process of the human brain[9]. The smallest component of a neural network is a unit commonly called a neuron that will transform the information received into another neuron[6].

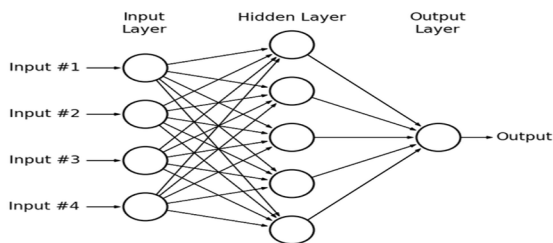


Fig. 1. Neural Network Model [6]

Neural networks consist of two or more layers, although most networks consist of three layers: input layer, hidden layer, and output layer[7]. The neural network approach is motivated by biological neural networks. Roughly speaking, a neural network is a set of input / output units, where each connection has a weight associated with it. Neural networks have several properties that make them popular for clustering[15]. First, the neural network is a parallel and distributed inherent processing architecture. Second, the neural network learns by adjusting the weights of interconnection with the data. This allows the neural network to "normalize" the pattern and act as a feature (attribute) extractors for different groups. Third, the neural network processes numerical vectors and requires an object pattern to be represented by a quantitative feature only. Neural network consists of a collection of nodes (neurons) and relationships. There are three types of nodes (neurons) namely, input, hidden and output. Each relation connects two nodes with a certain weight and there is also a direction that shows the flow of data in the process[15].

The node input is at the first layer in the neural network. In general each input node represents an input parameter such as age, gender, or income[16]. Hidden node is a node in the middle. This hidden node receives input from the input node on the first layer or from the hidden node of the previous layers. Hidden nodes combine all inputs based on the weights

of connected relations, calculate, and provide output for the next layer. The node output represents predicted attributes[24].

One advantage of using a neural network is that the neural network is strong enough with respect to the data. Because the neural network contains many nodes (artificial neurons) with weights assigned to each connection.

Neural network algorithms have other characteristics such as[9]:

1. Inputs can be either discrete or real values that have many dimensions.
2. The output is a vector consisting of several discrete or real values.
3. Can know the problem in black box, with only know the value of input and output only.
4. Able to handle learning to data that has noise (noise).
5. The shape of the target instructional function is unknown because it is simply the weight of the input value of each neuron.
6. Because it has to change a lot of weight value in the learning process, then the learning time becomes longer, so it is not suitable for problems that require fast time in learning.
7. Neural network of artificial learning results can be executed appropriately.

The most popular algorithm for neural network algorithms is the backpropagation algorithm[6]. The backpropagation method was originally designed for neural network feedforward, but in its development, this method was adapted for learning in other neural network models[9]. Backpropagation algorithm method is widely applied widely. backpropagation has been successfully applied in various fields, including financial fields, handwriting recognition, voice recognition, control systems, image processing. Backpropagation has become one of the most powerful computing methods[7]. The backpropagation algorithm has a very simple relation that is: if the output gives wrong result, then the weight is corrected so that its error can be minimized and the next network response is expected to be close to the correct value. The algorithm is also capable of fixing the weigher (hidden layer).

Langkah pembelajaran dalam algoritma *backpropagation* adalah sebagai berikut[9]:

1. Initialize network weights at random (usually between -0.1 to 1.0).
2. For each data in the training data, calculate the input for the node based on the input value and the current network weight, using the formula:

$$\text{Input } j = \sum_{i=1}^n O_i W_{ij} + \theta_j \quad (1)$$

3. Based on the input of step two, then generate output. to node using sigmoid activation function:

$$\text{Output} = \frac{1}{1 + e^{-\text{input}}} \quad (2)$$

4. Calculate Error value between the predicted value and the real value using the formula:

$$\text{Error}_j = \text{output}_j \cdot (1 - \text{output}_j) \cdot (\text{Target}_j - \text{Output}_j) \quad (3)$$

5. After Error value is calculated, then back to the previous screen (backpropagation). To calculate the Error value on the hidden layer, use the formula:

$$\text{Error}_j = \text{Output}_j(1 - \text{Output}_j) \sum_{k=1}^n \text{Error}_k W_{jk} \quad (4)$$

6. Generated Error value from the previous step is used to update the relation weights using the formula:

$$W_{ij} = W_{ij} + l \cdot \text{Error}_j \cdot \text{Output}_i \quad (5)$$

3.2 Genetic Algorithm

The genetic algorithm is a heuristic search algorithm based on the mechanism of biological evolution[14]. The diversity in biological evolution is a variation of the chromosomes between individual organisms. This chromosome variation will affect the rate of reproduction and the level of ability of organisms to survive[10].

Basically there are 4 conditions that greatly affect the evaluation process, namely as follows:

- a. ability of an organism to reproduce
- b. Presence of a population of reproductive organisms
- c. Existence of organisms in a population
- d. Differences ability to survive.

A stronger individual (fit) will have a higher survival rate and reproductive rate when compared with less fit individuals. Over a period of time (often known as generation), the population as a whole will contain more life-sustaining organisms[14].

In the genetic algorithm, the search technique is carried out simultaneously on a number of possible solutions known as the population. Individuals in one population are called chromosomes. This chromosome is a solution that is still a symbol. The initial population is randomly constructed, while the next population is the result of the evolution of chromosomes through an iteration called by the generation[10]. In each generation, the chromosomes will go through an evaluation process using a measuring device called a fitness function. The fitness value of a chromosome will show the quality of the chromosome in that population. The next generation known as the term child (off spring) is formed from a combination of 2 current generation chromosomes that act as parent by using crossover operator. In addition to the cross carrier, a chromosome may also be modified using a mutation operator. The population of the new generation is formed by selecting the fitness value of the child's chromosome (off spring), and rejecting the other chromosomes so that the population size (number of chromosomes in a population) is constant. After doing various generations, then this algorithm will converge to the best chromosome of life[10].

Let P (generation) be the population of a generation, then a simple genetic algorithm consists of the following steps[10]:

- Step 0: initialization Assume that the data is encoded in a string of bits (1 and 0). Determine the probability of crossover or pc Crossover rate and the probability of a mutation or muta mutation rate. Typically, pc is selected to be quite high (eg, 0.7), and pm is chosen very low (eg, 0.001)

- Step 1: The selected population, consisting of a set of n chromosomes each length i. Step 2: match the f (x) calculated for each chromosome in the population.
- Step 3: repeat through the following steps until the offspring have been generated
- Step 4: New chromosome population replaces the current population
- Step 5: Check if the termination criterion has been met. For example, is the average fitness change from generation to generation smaller? If convergence is reached, stop and report the results, otherwise go to step 2.

3.2 Particle Swarm Optimization Algorithm

Particle Swarm Optimization (PSO) is a population-based optimization technique developed by Eberhart and Kennedy in 1995, inspired by the social behavior of birds or fish[11]. Particle swarm optimization can be assumed as a group of birds in search of food in a region. The bird does not know where the food is, but they know how far the food is, so the best strategy to find the food is to follow the nearest bird from the food[26]. Particle swarm optimization is used to solve optimization problems.

Similar to the genetic algorithm (GA), the Particle swarm optimization performs a search using the population (swarm) of the individual (particles) to be updated from the iteration. Particle swarm optimization has several parameters such as position, speed, maximum speed, acceleration constant, and inertia weight[11]. Particle swarm optimization has more or better performance-search comparison for many optimization problems with faster and higher convergence rates more stable.

To find the optimal solution, each particle moves towards its previous best position and the best position globally. For example, the *i*th particle is expressed as: $x_i = (x_{i1}, x_{i2}, \dots, x_{id})$ in the *d*-dimensional space. The previous best position of the *i*th particles is stored and expressed as $pbest_i = (pbest_{i1}, pbest_{i2}, \dots, pbest_{id})$. The best particle index among all the particles in the group herd is expressed as $gbest_d$. Particle velocity is expressed as: $v_i = (v_{i1}, v_{i2}, \dots, v_{id})$. Modification of particle velocity and position can be calculated using the current velocity and $pbest_i$ distance, $gbest_d$ as shown in the following equation[26]:

$$v_{i,d} = w * v_{i,d} + c1 * R * (pbest_{i,d} - x_{i,d}) + c2 * R * (gbest_d - x_{i,d}) \quad (6)$$

$$x_{i,d} = x_{i,d} + v_{i,d} \quad (7)$$

Where:

V_i, d = speed of the *i*-th particles in the *i*-iteration

w = Inertial weight factor

$c1, c2$ = acceleration constants (learning rate)

R = Random number (0-1)

X_i, d = current position of the *i*th particle in the *i*-iteration

$pbest_i$ = Best previous position of the *i*-particles

$gbest_i$ = best particle among all the particles in one group or population

n = Number of particles in the group

d = Dimension

Equation (6) calculates the new velocity for each particle (potential solution) based on the previous velocity (V_i , m), the location of the particle where the best fitness value has been achieved (pbest), and the global population location (gbest for global version, lbest for local version) or local neighborhood on localized algorithm where the best fitness value has been reached.

Equation (7) renews the position of each particle in the solution space. Two random numbers c_1 and c_2 are self-generated. The heavy use of w inertia has provided increased performance in a number of applications. The result of particle counting is particle velocity between intervals $[0,1]$ [26].

3.3 Confusion Matrix

Confusion matrix provides the decisions obtained in training and testing, confusion matrix provides an assessment of classification performance by object correctly or wrongly[12]. Confusion matrix contains actual information (actual) and predicted (predicted) on the classification system. The following table explanation about confusion matrix.

TABLE I
Confusion Matrix [12]

Classification	Predicted Class		
	Class = Yes	Class = No	
Observed Class	Class = Yes	A (True Positif- tp)	B (False negatif- fn)
	Class = No	C (False positif- fp)	D (true negative-tn)

Information:

True Positive (tp) = positive proportion in the data set that is classified positively

True Negative (tn) = negative proportion in the data set that is classified negative

False Positive (fp) = negative proportion in potentially classified data sets

FalseNegative (fn) = negative proportions in data sets that are classified negatively

Here is the equation of confusion matrix model:

a. accuracy value (acc) is the proportion of the correct number of predictions. Can be calculated using the equation:

$$acc = \frac{tp+tn}{tp+tn+fp+fn}$$

b. Sensitivity is used to compare the proportion of tp to positive tuples, which is calculated using the equation:

$$Sensitivity = \frac{tp}{tp + fn}$$

c. Specificity is used to compare the proportion of tn to negative tuples, which is calculated using the equation:

$$Specificity = \frac{tn}{tn + fp}$$

d. PPV (positive predictive value) is the proportion of cases with a positive diagnosis, calculated using the equation:

$$PPV = \frac{tp}{tp + fp}$$

e. NPV (negative predictive value) is the proportion of cases with a negative diagnosis, calculated using the equation:

$$NPV = \frac{tn}{tn + fn}$$

3.4 ROC Curve

ROC Curve (Receiver Operating Characteristic) is another way to evaluate the accuracy of the classification visually[1]. An ROC graph is a two-dimensional plot with a false positive proportion (fp) on the X axis and a true positive proportion (tp) on the Y axis. Point (0,1) is a perfect classification of all positive and negative cases [Bruce, 2016] . The false positive value is none (fp = 0) and the true positive value is high (tp = 1). Point (0,0) is a classification that predicts each case to be negative {-1}, and point (1,1) is a classification that predicts each case to be positive {1}. The ROC graph illustrates the trade-off between benefits ('true positives') and cost ('false positives'). Here are two types of curve ROC (discrete and continuous).

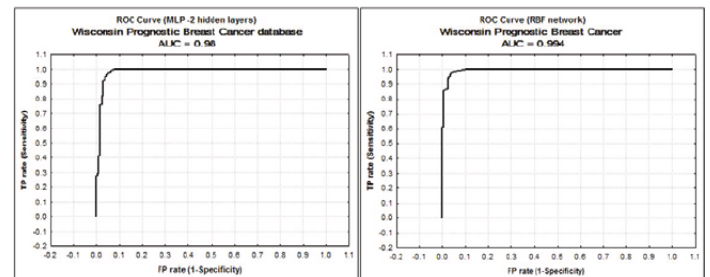


Fig. 2. ROC Grafix (discrete dan continuous) [4]

In Figure 2, the diagonal line divides the ROC space, ie:

1. (a) points above the diagonal line are the result of good classification.
2. (b) the point below the diagonal line is a poor classification result.

It can be concluded that, one point on the ROC curve is better than the other if the direction of the transverse line from the lower left to the top right in the graph. Accuracy rates can be diagnosed as follows [4]:

Accuracy 0.90 - 1.00 = Excellent classification

Accuracy 0.80 - 0.90 = Good classification

Accuracy 0.70 - 0.80 = Fair classification

Accuracy 0.60 - 0.70 = Poor classification

Accuracy 0.50 - 0.60 = Failure

III. IMPLEMENTATION AND RESULTS

3.1 Neural Network Method

Neural network algorithm is an algorithm for supervised training and is designed for operations on multilapis feed forward. The neural network algorithm can be described as follows: when the network is given an input pattern as a training pattern then the pattern goes to the units in the hidden layer to be forwarded to the outermost layer units.

The best results in the experiment is with accuracy produced by 98.50 and AUCnya 0.982. From the best experiments above we get the neural network architecture by generating six hidden layers with seven input layer attributes

and two output layers. Picture of neural network architecture shown in Figure 3 below:

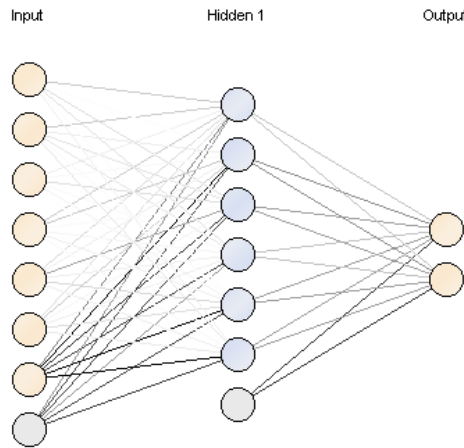


Fig. 3. Neural Network Architecture

3.2 PSO-based Neural Network Method

Particle Swarm Optimization has more or more superior performance search for more optimization problems with faster and more stable convergence rates. To find the optimal solution, each particle moves towards its previous best position and the best position globally. The best results in the experiment above is with accuracy produced by 98.85 and AUCnya 0.996.

Next step is to select the attributes used are gender, no. sequence of political parties, legitimate vote of the party, number of seats, election area, no. sequential caleg, legitimate voice caleg and 1 attribute as label that is result of election. From the experimental results by using the neural network algorithm based on particle swarm optimization obtained some attributes attributes that have an effect on the attribute weights: Juml. Acquired seats with weight of 0.143, no. sequential caleg with weight 0.344 and legitimate voice caleg with weight 1. Whereas other attributes such as: gender, party serial number, legitimate vote of party, election area and legitimate vote caleg no effect to attribute weight.

From the best experiments above we get the neural network architecture by producing fifteen hidden layers with two input layer attributes and two output layers. Picture of neural network architecture shown in Figure 4 below

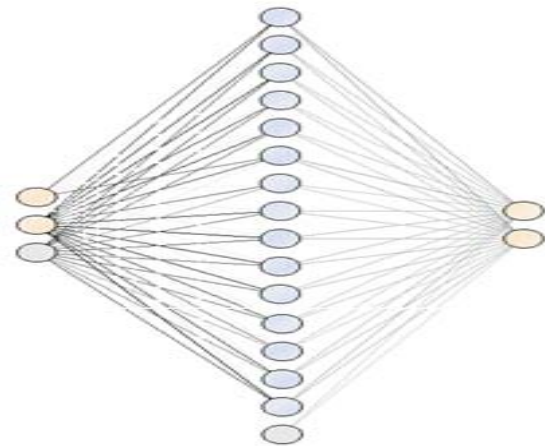


Fig. 4. PSO based neural network architecture

3.3 Neural Network based Genetic Algorithm

Neural network algorithm can be described as follows: when the network is given an input pattern as a training pattern then the pattern goes to the units in the hidden layer to be forwarded to the outermost layer units. While the genetic algorithm is a heuristic search algorithm based on the mechanism of biological evolution. The diversity in biological evolution is a variation of the chromosomes between individual organisms. This chromosome variation will affect the rate of reproduction and the level of the organism's ability to survive. Testing using genetic algorithm based neural network obtained accuracy value 96.56% with value and value of AUC is 0.925.

3.4 Evaluation and Validation Analysis Model

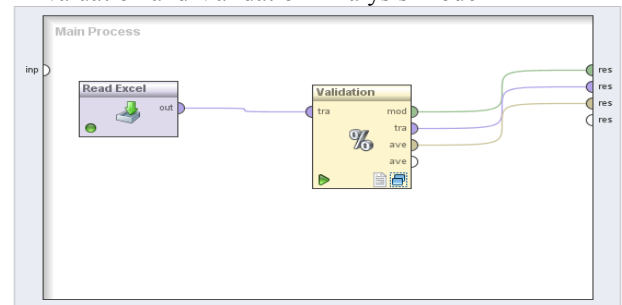


Fig. 5. Cross validation Test

From the above test results, both evaluations using counfusion matrix and ROC curve proved that the results of algorithm-based neural network algorithm particle swarm optimization algorithm has a higher accuracy value than neural network algorithm and genetic algorithm. The accuracy value for the neural network algorithm model is 98.50%.

The test using the particle swarm optimization neural network algorithm obtained accuracy value of 98.85 and the accuracy value for Genetic Algorithm model based on 96.56%. While the evaluation using ROC curve so as to produce the value of AUC (Area Under Curve) for neural network algorithm model to produce the value of 0.982 with Excellent Classification diagnostic value, and the AUC curve

for optimization testing with particle swarm optimization of 0.996 with the value of Excellent Classification diagnosis, while the optimization algorithm genetics yields a value of 0.925 with an Excellent Classification diagnostic value.

IV. CONCLUSION

Based on experimental results conducted from the optimization analysis of neural network algorithm model and optimization with particle swarm optimization algorithm and genetic algorithm. The resulting model is tested to get accuracy, precision, recall and AUC value of each algorithm so that the test obtained by using neural network obtained accuracy value is 98.50% with precision value 91.20% and the value of AUC is 0.982. then testing by using the neural network algorithm based on particle swarm optimization to get accuracy value generated equal to 98.85 and AUC curve equal to 0.996. while testing by using genetic algorithm based neural network obtained accuracy value 96.56% with 91.28% precision value and the value of AUC is 0.925. it can be concluded that the testing of Jakarta legislative election model using neural network based on particle swarm optimization has accurate value and AUC curve better than the test using neural network algorithm based on genetic algorithm or just neural network itself.

REFERENCES

- [1]. Aggarwal, C. C. (2015). *Data Mining The Textbook*. Springer.
- [2]. Borisyuk, R., Borisyuk, G., Rallings, C., & Thrasher, M. (2013). Forecasting the 2005 General Election: A Neural Network Approach. *The British Journal of Politics & International Relations*, 145-299.
- [3]. Bradberry, L. A. (2016). The Effect of Religion on Candidate Preference in the 2008 and 2012 Republican Presidential Primaries. *Plos One*, 1-2.
- [4]. Bruce, P. C., Patel, N. R., & Shmueli, G. (2016). *Data Mining for Business Analytics : Concepts, Techniques, and Applications with XLMiner 3rd Edition*. New York, United States: John Wiley & Sons Inc.
- [5]. Brunello, G. H., & Nakano, E. Y. (2015). Bayesian Inference on Proportional Elections. *PLOS ONE*, 1-2.
- [6]. Casey, A. (2016). *Soft Computing : Developments, Methods & Applications*. New York, United States: Nova Science.
- [7]. Demuth, H. B., Hagan, M. T., & Beale, M. H. (2014). *Neural Network Design (2nd Edition)*. Martin Hagan.
- [8]. Gill, G. S. (2005). Election Result Forecasting Using two layer Perceptron Network. *Journal of Theoretical and Applied Information Technology Volume.4 No.11*, 144-146.
- [9]. Goldberg, Y. (2017). *Neural Network Methods for Natural Language Processing*. San Rafael, CA, United States: Morgan & Claypool Publishers.
- [10]. Jones, S. (2015). *Genetic Algorithms : Practical Applications*. Clanrye International.
- [11]. kiranyaz, S., Ince, T., & Gabbouj, M. (2013). *Multidimensional Particle Swarm Optimization for Machine Learning and Pattern Recognition*. Springer.
- [12]. Lichtendahl, K. C., Bruce, P. C., Patel, N. R., Shmueli, G., Torgo, L., & Yahav, I. (2017). *Data Mining for Business Analytics: Concepts, Techniques, and Applications in R*. John Wiley & Sons Inc.
- [13]. Ling, S. H., Nguyen, H. T., & Chan, K. Y. (2009). A New Particle Swarm Optimization Algorithm for Neural Network Optimization. *Network and System Security, third International Conference*, 516-521.
- [14]. Mahmud, F., & Zuhori, S. T. (2012). *Genetic Algorithm*. Germany: LAP Lambert Academic Publishing.
- [15]. Mark, E. F., Witten, I., Pal, H. C., & Hall, M. (2017). *Data Mining 4th Edition Practical Machine Learning Tools and Techniques*. Elshiever Science & Technology.
- [16]. Meira, W., & Zaki, M. J. (2015). *Data Mining and Analysis: Fundamental Concepts and Algorithms*. Cambridge, United Kingdom: Cambridge University Press.
- [17]. Moscato, P., Mathieson, L., Mendes, A., & Berreta, R. (2015). The Electronic Primaries: Prediction The U.S. Presidential Using Feature Selection with safe data. *ACSC '05 Proceeding of the twenty-eighth Australian conference on Computer Science Volume 38*, (pp. 371-379).
- [18]. Nagadevara, & Vishnuprasad. (2013). Building Predictive models for election result in india an application of classification trees and neural network. *Journal of Academy of Business and Economics*.
- [19]. Park, T. S., Lee, J. H., & Choi, B. (2009). Optimization for Artificial Neural Network with Adaptive inertial weight of particle swarm optimization. *Cognitive Informatics, IEEE International Conference*, 481-485.
- [20]. Rigdon, S. E., Jacobson, S. H., Sewell, E. C., & Rigdon, C. J. (2013). A Bayesian Prediction Model For the United State Presidential Election. *American Politics Research*, 700-724.
- [21]. Saldana, & Hernandez, H. (2013). Result on Three Predictions for July 2012 Federal Elections in Mexico Based on Past Regularities. *Plos One*, 1-2.
- [22]. Sekretariat Negara, R. (2017). *Pemilihan Umum Anggota Dewan Perwakilan Rakyat, Dewan Perwakilan Daerah, Dan Dewan Perwakilan Rakyat Daerah*. Jakarta: Setneg.
- [23]. Sug, H. (2009). An Empirical Determination of Samples for Decision Trees. AIKED'09 Proceeding of the 8th WSEAS international conference on Artificial intelligence. *Knowledge engineering and data bases*, 413-416.
- [24]. Tan, P. N., Steinbach, M., Karpatne, A., & Kumar, V. (2018). *Introduction to Data Mining 2nd Edition*. United States: Pearson Education.
- [25]. Thoha, M. (2014). *Birokrasi Politik Pemilihan Umum di Indonesia*. Jakarta: Kencana.
- [26]. Walker, B. (2017). *Particle Swarm Optimization (PSO) : Advances in Research & Applications*. New York, United States: Nova Science Publishers Inc.
- [27]. Xiao, & Shao, Q. (2011). Based on two Swarm Optimized algorithm of neural network to prediction the switch's traffic of coal. *ISCCS '11 Proceeding of the 2011 International Symposium on Computer Science and Society*, 299-302.