EasyChair Preprint

№ 2013

Optimization of Artificial Neural Networks to Improve the Accuracy in Predicting the Selection of Competency Competencies of Vocational Students Using Nguyen-Widrow

I. Gusti Putu Christyaditama, I Made Candiasa and
I Gede Aris Gunadi

November 21, 2019

OPTIMIZATION OF ARTIFICIAL NEURAL NETWORKS TO IMPROVE THE ACCURACY IN PREDICTING THE SELECTION OF COMPETENCY COMPETENCIES OF VOCATIONAL STUDENTS USING NGUYEN-WIDROW.

I. G. P. Christyaditama, S.Pd.Kom[1)], Prof. Dr. I Made Candiasa, M.I.Kom[2)],
Dr. I Gede Aris Gunadi, S.Si, M.Kom[3)]

[1)2)3)] Study Program of Computer Science, Postgraduate Program
Ganesha University of Education
Singaraja, Indonesia

e-mail: Christyaditama@gmail.com[1)], made.candiasa@pasca.undiksha.ac.id[2)],
igedearisgunadi@undiksha.ac.id[3)]

Abstrak

The use Neural Network Backpropagation combined with Nguyen-Widrow to optimize the disadvantages of ANN, which is the difficulty in initializing initial weights. The test was conducted on a dataset of values in semesters 1 and 2. The test results show that the best performance in training model of artificial neural networks with Nguyen-widrow is the smallest average MSE error of 0.002 and the highest average accuracy of 96.38%. Training on artificial neural network model training data with Nguyen-widrow has the smallest MSE error, 0.000996 and the highest accuracy is 97.49% on architecture ANN 9-9.1 with training function parameters: traingdx, epoch: 1000, learning rate: 0.1, and error: 0.001 . The best performance was also seen in testing the testing of artificial neural network models with Nguyen-widrow with the smallest average error-MSE of 0.026 and the highest average accuracy of 87.85%. Training data testing on artificial neural network models with Nguyen-widrow has the smallest error-MSE which is 0.004436 and the highest accuracy is 94.50% on architecture ANN 9-9.1 with training function parameters: traingdx, epoch: 1000, learning rate: 0.1, and error: 0.001 . The artificial neural network model with Nguyen-widrow has an accuracy difference of 8.53% smaller than the artificial neural network model with an accuracy difference of 9.28%. It can be concluded that the Artificial Neural Network with Nguyen-Widrow can overcome the ANN problem in determining initial weights so that it gives an increase in the accuracy of the prediction of students' competency selection better than the Artificial Neural Network without Nguyen-Widrow.

Keywords: *Selection of specialization of vocational students' expertise, artificial neural networks, nguyen-widrow*

1. INTRODUCTION

The process of determining the competency of expertise that has been used at SMK Wira Harapan is only based on student specialization and interview results. According to students' specialization guidelines based on Government Regulation Number 17 of 2010 concerning Implementation and Management of Education for the selection of specialization in the field of expertise and expertise programs conducted in semester 3, based on report card grades and / or recommendations in SMK / MAK and / or placement test results (placement test) by a psychologist. In addition to non-test techniques, it can also use test techniques, such as psychological tests carried out by testers or specialization tests that can be carried out by BK / Counselor teachers. In the specialization guidelines for students based on the 2013 curriculum convey data obtained from test and non-test techniques (documentation, questionnaires, interviews, observations, etc.) complement each other. Can be interpreted that the more data collected and can be analyzed correctly, the accuracy of determining the interest of students will be higher.

Based on the background above, the research will propose using the backparagation neural network method to assist in supporting students' competency selection decisions. The method is used after looking at the accuracy comparison of the 4 other methods used in predicting the selection of Competency Competencies. The first method is K-nearest Neighbor, with an accuracy level of 79.68% tested on 160 data which is divided into 64 test data and classified with 96 training data [1]. The disadvantages of K-nearest Neighbor are the value of k bias, complex computation, memory limitations and are easily fooled by irrelevant links [2]. The second method is Naive Bayes is better than the other methods with a value of 77.51 [3]. The weakness of the Naïve Bayes method is the probability of not being able to measure the magnitude of the accuracy of a prediction [4]. The second method is Decisioan Tree C4.5 for determining the suitability of a Student Competency to get an accuracy of 93.31% [5]. The disadvantage of C4.5 decision tree is overlap, especially when the classes and criteria used are very large and the quality of the decisions obtained depends very much on how the tree is designed [6]. Seeing the results of the accuracy of the 3 methods, the accuracy is still smaller compared to journals to predict student achievement with artificial neural networks having 95.6% training error accuracy and 100% testing error [7]. The use of artificial neural network methods using backpropagation algorithms in data processing is still experiencing weakness. The weakness of the artificial neural network method using the backpropagation algorithm is influenced by the initial weight chosen randomly [8]. As an optimization to overcome the minimum local weaknesses that occur in the initial weighting of artificial neural networks using the Nguyen-widrow method can provide a more convergent value and is used as an initial weighting in the training process in the backpropogation method.

The application of this method is done by looking at semester 1 and 2 data on group A and B subjects including; Religion, PPKN, Indonesian Language, Mathematics, Indonesian History, English, Penjaskes (Physical Education and Health Sciences) and Kewirausahaan (Entrepreneurship) to predict competence in appropriate expertise. In this study do not consider other factors as a supporter of decision support in choosing competency expertise, for example the factors of skill, talent, interest and other factors. Based on the description above, the authors raise the title of artificial neural network optimization research to improve the accuracy of the prediction of the competency selection of vocational students' expertise using Nguyen-Widrow

2. SUBJECT OF THE STUDY

This subject was based on the dapodik data in SMK Wira with the total amount of data is 611 in 2017. The dataset that are the grade of students in all competent skills, based on a 2017 grade historical.

3. RESEARCH METHOD

The research method used was experimental research, with the research stage testing the accuracy of artificial neural network algorithms with a backpropagation model that uses parameters in the training and testing process to influence the values, namely; training functions: traingdx, epoch: 1000, learning rate: 0.1, and error: 0.001. The dataset used is the values of Wira Harapan Vocational School students of all expertise competencies, based on the 2017 value history. Model of Artificial Neural Networks and Artificial Neural Networks with Nguyen-widrow is tested to compare the results of Error-MSE and Smallest Accuracy seen from hidden layer: 9, 6 and 3 on artificial neural network architecture. So the architecture in the model of artificial neural networks and artificial neural networks with Nguyen-widrow are: 9-9-1, 9-6-1, and 9-3-1. The determination of the model will be seen from the highest prediction accuracy with the rules of the number of neurons, hidden layers and the training function used. The model with the highest accuracy will be used to predict the competency of students' expertise so that they will get a model with more accurate prediction results. Prediction results obtained will be validated and measured by comparing the error value of the predicted algorithm.

ANN design to predict expertise competency for vocational students is done by determining the architecture of the artificial neural network model and artificial neural network with Nguyen-widrow which will be used to get the model with the highest prediction accuracy, with the following process:

1. Data Pra-process

Data pra-processing is carried out in this study to prepare student grade data so that it can be processed or applied to the ANN model to be developed, there are several stages that are carried out in data processing preprocess:

a) Data Correction
Select semester 1 and 2 data on group A and B subjects including; Religion, PPKN, Indonesian, Mathematics, Indonesian History, English, Penjaskes and Kewirausahaan. Furthermore, by giving an empty value with an average value.

b) Data Normalization
The data used will use the sigmoid activation function, so the data will be transformed to a smaller interval that is a value that lies between 0 and 1 with the formula. This transformation process is called data normalization which aims to simplify the calculation process. Normalization is done by using the transformation formula

$$X = \frac{0.8(X - Xmin)}{Xmax - Xmin} + 0.1$$

2. Nguyen-widrow

The Nguyen-widrow algorithm is used to make weighting initialization as a method to increase accuracy in artificial neural network methods. The Nguyen-widrow algorithm can be demonstrated using the following formula:

a. Set:
N = number of input units
P = number of hidden units
β = scale factor = $0.7(p)^{1/n} = 0.7\sqrt[n]{p}$
b. For each hidden unit $(j = 1, .., p)$, do steps $(c) - (f)$
c. For $i = 1, .., n$ (all input units), $v_{ij}(old)$ = random numbers between -0.5 and 0.5
d. Calculate values with formulas

$$\|v_j(old)\| = \sqrt{v_j^2 j + v_1^2 ja + \cdots + v_n^2 j}$$

e. Calculate new weights by formula

$$v_{ij} = \frac{\beta v_{ij(old)}}{\|v_j\|}$$

f. Bias used as initialization:

v_{oj} = random number between $-\beta \; and \; \beta$

3. Training and Test Data

The average value in semester 1 will be used as output for later compared with the output of the model of artificial neural network and artificial neural network with Nguyen-widrow produced, so that from the comparison of the reality value and the value of the model results will be obtained its accuracy value. The average value of Semester 2 will be the target for later compared to the outputs from the model of artificial neural networks and artificial neural networks with Nguyen-widrow produced, so that from the comparison of target values and the results of the model results will be obtained the highest accuracy value approaching 0.

4. Model Design with Artificial Neural Networks (ANN)

The design of artificial neural network models and artificial neural networks with Nguyen-widrow uses parameters in the training and testing process to influence the values namely; training functions: traingdx, epoch: 1000, learning rate: 0.1, and error: 0.001. Architecture in the model of artificial neural networks and artificial neural networks with Nguyen-widrow are: 9-9-1, 9-6-1, and 9-3-1.

5. Model Testing

Tests on the ANN and ANN models with Nguyen-Widrow are performed to determine the level of accuracy of the predictions of the ANN and ANN algorithms with Nguyen-Widrow against the actual data.

6. Forecast Model

Artificial neural network models and Nguyen-widrow artificial neural networks were tested with hidden layers: 9, 6 and 3 with predetermined parameters. So that it can compare models from the results of the smallest Error-MSE average results and the average percentage of the highest accuracy to predict the competency skills of SMK Wira Harapan students.

4. RESULT AND DISCUSSION

In this test explains that there is an increase in the accuracy of training data training using artificial neural network models and artificial neural networks with Nguyen-widrow

Table 1. Error-MSE and accuracy in training data training

MAJOR	HIDDEN LAYER	WITHOUT *NGUYEN-WIDROW*		*NGUYEN-WIDROW*	
		Error-MSE	Average Accuracy	*Error-MSE*	Average Accuracy
TKJ	9	0.001	97.15%	0.001	97.49%
	6	0.001	97.44%	0.002	96.77%
	3	0.002	96.55%	0.003	96.58%
RPL	9	0.001	97.03%	0.002	96.75%
	6	0.003	95.93%	0.002	96.42%
	3	0.009	91.79%	0.002	96.11%
MM	9	0.002	96.66%	0.002	96.57%
	6	0.001	97.38%	0.002	96.49%

	3	0.005	93.97%	0.002	96.73%
UPW	9	0.002	95.95%	0.001	96.97%
	6	0.004	95.51%	0.003	95.33%
	3	0.010	92.00%	0.007	93.18%
AP	9	0.001	97.51%	0.003	95.91%
	6	0.002	96.63%	0.001	97.46%
	3	0.003	95.94%	0.004	95.32%
JB	9	0.001	97.57%	0.002	96.57%
	6	0.002	96.95%	0.001	97.11%
	3	0.002	96.18%	0.001	97.12%
TOTAL AVERAGE		0.003	96.01%	0.002	96.38%

The artificial neural network model with Nguyen-widrow has an average Error-MSE of 0.002 smaller than the artificial neural network model of 0.003. In addition to the smaller error value, the average accuracy of the neural network model with Nguyen-widrow is 96.38% higher than the artificial neural network model with an average accuracy of 96.01%. In the test results get the same comparison as shown in table 2.

Table 2. Error-MSE and accuracy in testing data testing

MAJOR	HIDDEN LAYER	WITHOUT *NGUYEN-WIDROW*		*NGUYEN-WIDROW*	
		Error-MSE	Average Accuracy	*Error-MSE*	Average Accuracy
TKJ	9	0.009	92.63%	0.073	74.42%
	6	0.298	85.30%	0.088	74.83%
	3	0.043	81.25%	0.012	90.49%
RPL	9	0.067	77.77%	0.097	69.89%
	6	0.086	73.21%	0.013	90.73%
	3	0.240	60.96%	0.027	85.81%
MM	9	0.004	94.65%	0.006	94.30%
	6	0.010	91.83%	0.014	89.56%
	3	0.037	82.00%	0.015	89.20%
UPW	9	0.033	84.01%	0.029	85.41%
	6	0.015	90.05%	0.015	90.01%
	3	0.047	81.93%	0.041	83.02%
AP	9	0.006	93.96%	0.004	94.50%
	6	0.005	94.41%	0.005	94.40%
	3	0.002	96.04%	0.006	93.61%
JB	9	0.003	95.65%	0.005	94.12%
	6	0.013	90.26%	0.006	93.52%
	3	0.006	95.14%	0.006	93.45%
TOTAL AVERAGE		0.051	86.73%	0.026	87.85%

The artificial neural network model with Nguyen-widrow has an average Error-MSE of 0.026 smaller than the artificial neural network model of 0.051. In addition to the smaller error value, the average

accuracy of the neural network model with Nguyen-widrow is 87.85% higher than the artificial neural network model with an average accuracy of 86.73%

In the training of artificial neural network model training data with Nguyen-widrow there is the smallest MSE error, 0.000996 and the highest accuracy is 97.49% on the architecture of ANN 9-9.1 with training function parameters: traingdx, epoch: 1000, learning rate: 0.1, and error: 0.001. Comparison of the average total Error-MSE and accuracy in training data training can be represented as in Figure 1

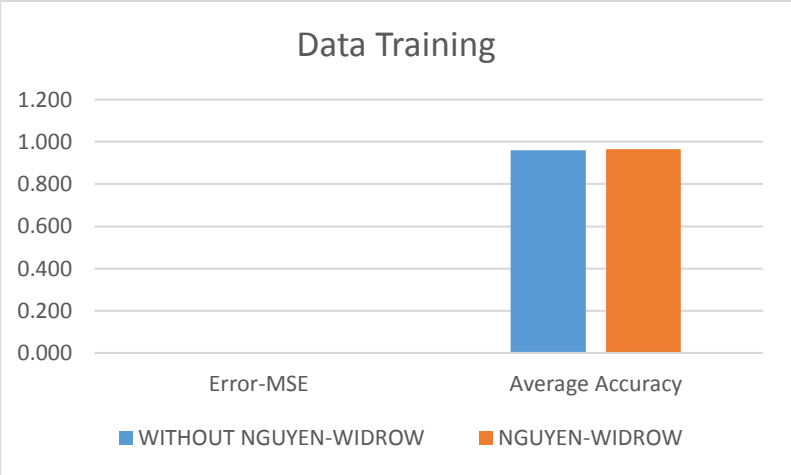

Figure 1. Comparison of average error-MSE and accuracy in training data training

The best performance was also seen in testing the testing model of artificial neural networks with Nguyen-widrow with the smallest average error-MSE of 0.026 and the highest average accuracy of 87.85% shown in table 3

Table 3. Average Total Error-MSE and accuracy in data testing

DATA TESTING	Error-MSE	Average Accuracy
WITHOUT NGUYEN-WIDROW	0.051	86.73%
NGUYEN-WIDROW	0.026	87.85%

Training of artificial neural network model training data with Nguyen-widrow has the smallest MSE error, the value of 0.004436 and the highest accuracy of 94.50% on the architecture of ANN 9-9.1 with training function parameters: traingdx, epoch: 1000, learning rate: 0.1, and error: 0.001 . Comparison of the average total Error-MSE and accuracy in training data testing can be represented as in Figure 2.

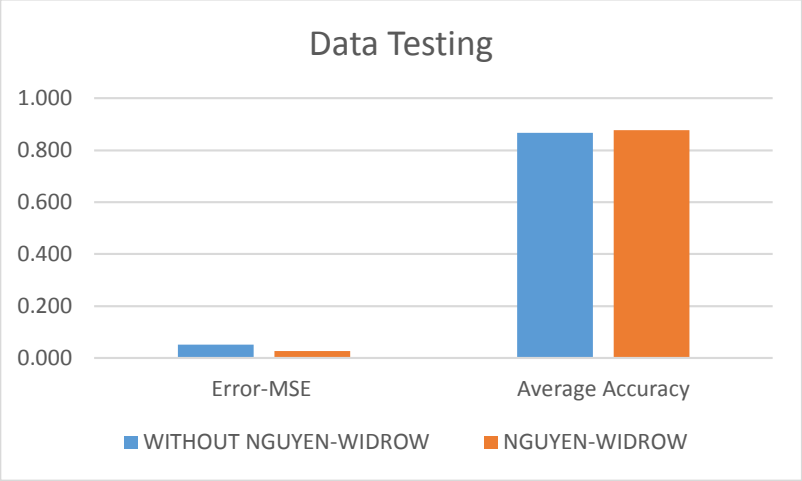

Figure 2. Comparison of average Error-MSE and accuracy in data testing

The difference in accuracy can also be seen from the difference between training data and testing of artificial neural networks and with artificial neural networks with Nguyen-widrow. The artificial neural network model with Nguyen-widrow has an accuracy difference of 8.53% smaller than the artificial neural network model with an accuracy difference of 9.28%.

5. CLOSING

Based on the test results and analysis of the results of the Matlab, we get several conclusions, namely:

1. The results of experiments on training data with artificial neural network models obtained different accuracy values in each of the expertise competencies of SMK Wira Harapan. Competence of TKJ expertise with the highest accuracy of 97.44% and error-MSE as low as 0.001 on architecture ANN 9-6-1. RPL expertise competency with the highest accuracy is 97.03% and error-MSE is as low as 0.001 on the architecture of ANN 9-9-1. MM competence with the highest accuracy is 97.38% and error-MSE is as low as 0.001 on architecture ANN 9-6-1. UPW expertise competence with the highest accuracy of 95.95% and error-MSE 0.002 on ANN 9-9-1 architecture. AP expertise competence with the highest accuracy is 97.51% and error-MSE is as low as 0.001 on the architecture of ANN 9-9-1. JB's competence with the highest accuracy is 97.57% and error-MSE is as low as 0.001 on the architecture of ANN 9-9-1. The results of this accuracy and error-MSE use parameters; training function: traingdx, epoch: 1000, learning rate: 0.1, and error: 0.001 in the process of training data training to influence grades.
2. The experimental results on the training data model of artificial neural networks with Nguyen-widrow obtained different values of accuracy in each of the expertise competencies of SMK Wira Harapan. Competence of TKJ expertise with the highest accuracy of 97.49% and error-MSE as low as 0.001 on the architecture of ANN 9-9-1. RPL expertise competency with the highest accuracy is 96.75% and error-MSE is as low as 0.002 on the architecture of ANN 9-9-1. MM expertise competence with the highest accuracy is 96.73% and error-MSE is as low as 0.002 on architecture ANN 9-3-1. UPW expertise competency with the highest accuracy is 96.97% and error-MSE is as low as 0.001 on the architecture of ANN 9-9-1. AP expertise competence with the highest accuracy is 97.46% and error-MSE is as low as 0.001 on architecture ANN 9-6-1. JB's competence with the highest accuracy is 97.12% and error-MSE is as low as 0.001 on architecture ANN 9-3-1.
3. The results of experiments on testing data with neural network models obtained different accuracy values in each of the expertise skills of SMK Wira Harapan. Competence of TKJ expertise with the highest accuracy of 92.63% and error-MSE as low as 0.009 on the architecture of ANN 9-9-1.

RPL expertise with the highest accuracy is 77.77% and error-MSE is as low as 0.067 on the architecture of ANN 9-9-1. MM competence with the highest accuracy is 94.65% and error-MSE is as low as 0.004 on the architecture of ANN 9-9-1. UPW expertise competencies with the highest accuracy is 90.05% and error-MSE is as low as 0.015 on architecture ANN 9-6-1. AP expertise competency with the highest accuracy is 96.04% and error-MSE is as low as 0.002 on architecture ANN 9-3-1. JB's competence with the highest accuracy is 95.65% and error-MSE is as low as 0.003 on the architecture of ANN 9-9-1.

4. Experimental results on testing the neural network model data with Nguyen-widrow obtained different values of accuracy in each of the expertise competencies of SMK Wira Harapan. TKJ competence with the highest accuracy is 90.49% and error-MSE is as low as 0.012 on architecture ANN 9-3-1. RPL expertise with the highest accuracy is 90.73% and error-MSE is as low as 0.013 on architecture ANN 9-6-1. MM competence with the highest accuracy is 94.30% and error-MSE is as low as 0.006 on ANN 9-9-1 architecture. UPW expertise competencies with the highest accuracy is 90.01% and error-MSE is as low as 0.015 on architecture ANN 9-6-1. AP expertise competencies with the highest accuracy is 94.50% and error-MSE is as low as 0.004 on the 9-9-1 ANN architecture. JB's competence with the highest accuracy is 94.12% and error-MSE is as low as 0.005 on ANN 9-9-1 architecture.
5. Experimental results on training and testing data with artificial neural network models and artificial neural networks with Nguyen-widrow obtained 9-9-1 architecture which has the best accuracy performance. Can be seen from the 24 highest accuracy and lowest error-MSE, architecture 9-9-1 appears 14 times, 9-6-1 appears = 6 times, and 9-3-1 appears 4 times
6. Based on the comparison of Error-MSE values and the average accuracy between artificial neural network models and artificial neural networks with Nguyen-widrow, then training and network testing using artificial neural network models with Nguyen-widrow can be an option to improve the accuracy of prediction of competency selection student expertise.

The results of this study need to get more developments to get better and perfect results. Development that can be given, among others, by; 1). Implement new parameters to be able to improve the accuracy of prediction of the competency selection of Wira Harapan Vocational School students' competency. So that it can contribute to the world of education and science until it finally benefits society; 2) Implementing architectures that vary from the application of hidden layers so as to provide the smallest accuracy and error-MSE increase for research using artificial neural networks. 3) Adding methods or changing methods so that performance is seen from the accuracy and error-MSE will be better for speeding up the weighting process or determining the bias value.

REFERENCES

[1] A. Sulistiyo, "Penentuan Jurusan Sekolah Menengah Atas Menggunakan Metode K-Nearest Neighbor Classifier Pada SMAN 16 Semarang," pp. 1–5, 2015.

[2] M. Mutrofin, Siti., Izzah, Abidatul., Kurniawardhani, Arrie., Masrur, "Optimasi teknik klasifikasi modified k nearest neighbor menggunakan algoritma genetika," *J. Gamma*, no. September, pp. 130–134, 2014.

[3] S. N. Nugroho, Yusuf S., Haryati, "Klasifikasi dan Klastering Penjurusan Siswa SMA Negeri 3 Boyolali," 2015.

[4] H. Muhamad, C. A. Prasojo, N. A. Sugianto, L. Surtiningsih, and I. Cholissodin, "Optimasi Naïve Bayes Classifier Dengan Menggunakan Particle," vol. 4, no. 3, pp. 180–184, 2017.

[5] L. Swastina, "Penerapan Algoritma C4 . 5 Untuk Penentuan Jurusan Mahasiswa," vol. 2, no. 1, 2013.

[6] Azmi, Zulfian. and M. Dahria, "Decision Tree Berbasis Algoritma Untuk Pengambilan Keputusan," *Saintikom*, vol. 12, pp. 157–164, 2013.

[7] F. Zola, G. W. Nurcahyo, and T. K. Jaringan, “Jaringan syaraf tiruan menggunakan algoritma backpropagation untuk memprediksi prestasi siswa,” vol. 1, no. 1, pp. 58–72, 2018.

[8] N. M. Nawi, A. Khan, and M. Z. Rehman, “A New Back-Propagation Neural Network Optimized,” *A New Back-Propagation Neural Netw. Optim. New Back-Propagation Neural Netw. Optim.*, pp. 413–426, 2013.