

Arief_Hermawan_The Dimensional REduction to Improve the Speed adn accuracy of NN

by Arief Hermawan

Submission date: 04-Apr-2023 07:50PM (UTC-0400)

Submission ID: 2056076123

File name: 1_43909_-_IJAER_ok_4758-4762_new1.pdf (244.78K)

Word count: 2497

Character count: 14470

The Dimensional Reduction to Improve the Speed and Accuracy of Neural Network in Identifying the Senior High School Students' Major

Arief Hermawan

University Technology of Yogyakarta, Indonesia

E-mail: ariefjb@uty.ac.id

Abstract

The appropriateness of choosing the major for the Senior High School students is considered as the most important part of the academic process to be considered in order to determine the right major for their further study in the highest level. The presence of back propagation neural network with momentum and adaptive learning rate make it possible for the system itself to determine the appropriate major for the Senior High School students based on their academic potential test result with the level of accuracy about 97.5%. In order to be able to do so, the system iterates in the amount of 28,356. The objective of the present study is to perform the dimensional reduction to decrease the iteration times and to improve the accuracy of choosing the major. The researcher had tried out and examined the neural network by reducing the input variable. A series of simulation had been conducted by the researcher using Matlab software. Upon completion the simulation the researcher found out that neural network was convergent when general competence variable was reduced with the constant accuracy about 97.5%, while the iteration had decrease in the amount of 1043.

Keywords: Dimensional Reduction, Major, Neural Network, Senior High School.

INTRODUCTION

The government regulation of the republic of Indonesia no 17 year 2010 on education management and implementation [1] mentions that Senior High Schools must perform majoring. This majoring is meant to facilitate learning and competence required by learning participants to continue to pursue higher education.

Major selection accuracy for senior high school students is vital since it is the foundation to determine major's at their higher education. Many failures in students study continuity are caused by mistakes in choosing a major. In the 2013 curriculum which has been enforced since 2013, aptitude tests have taken a vital part in students major selection. [2] finds that back error propagation neural network can be employed to perform majoring for high school students by using aptitude test results. Backpropagation neural network processes aptitude test result, namely general competence, mechanical competence, numerical competence, verbal competence, and language competence to become majors for senior high school students. In his finding Hermawan [2] states that the various simulated neural network configurations used, have not been able to effectively identify majors for senior high school students. A simulated neural network could identify senior high school students majors after it has been iterated for 28,356 times with identification accuracy of 97,5%. This

finding result indicates that a simulated neural network requires iteration for thousands of times and its pattern identification accuracy has not been able to reach 100%. It is indicated that out of the 5 variables used in senior high school students majoring, which are general competence, mechanical competence, numerical competence, verbal competence, and language competence, there is a variable acting as the disturbing variable.

This research will conduct improvement on senior high school students majoring selection system. Improvement will be performed by reducing input dimensionality of a neural network so that variables used to determine senior high school majors are the variables which are truly influential toward major selection. It is expected that by employing the truly influential variables, senior high school major selection system using neural network could be more effective. The success indicator is that neural network could identify senior high school students' majors with fewer iteration and the major identification could be 100%. The success of the system will ensure that high school students could choose majors in accordance to their aptitude so that students' failure in learning due to mistakes in tracking selection could be prevented. In general, the success of the system implementation would reduce the number of dropouts caused by major selection mistakes.

The objective that this research wishes to achieve is to find the disturbing variable in senior high school students majoring system by using neural network. The discovery of the disturbing variable would be accomplished by training and testing a neural network by reducing input dimensionality of a simulated neural network.

RELATED LITERATURES

A neural network is a computational model inspired by human neural network system. A neural network solves problems by learning. A neural network learns to make all data entered to it could provide output as expected [3].

Numerous simulated neural network application researches at educational field show that a neural network could effectively solve problems related to education. Oladokun, Adebajo, & Charles-Owaba [4] state that neural network could predict students' academic competence well. In that research, the developed neural network could predict students' competence with the accuracy of more than 70%. In technical education, a neural network with back error propagation algorithm could predict education quality up to 90% accuracy [5]. Budhi, Liliana, & Haryanto [6] state that a neural network with Self Organizing Map Algorithm could predict the talents of basketball players. NBA player data is used for training while Indonesian player data is used for test. Based on the test, the

neural network could predict the players' talent as much as 95%.

One problem frequently encountered in classification using a neural network is the convergence duration of a neural network to identify patterns trained and the low accuracy of neural network in identifying the patterns tested. Muknahallipatna & Chowdhury [7] state that those problems could be solved by reducing input variable dimensionality of a neural network.

Numerous studies related to dimensionality reduction prove that increase in neural network is successfully achieved by reducing input variable dimensionality of a neural network. Blume et al. [8] finds that dimensionality reduction with neural network could lower pattern identification mistakes by 21%. Hinton & Salakhutdinov [9] concludes that neural network is very effective for dimension reduction. Compared to other methods such as Principal Component Analysis (PCA), dimensionality reduction with neural network could produce better classification. Perantonis & Virvilis [10] state that dimensionality reduction could increase neural network classification accuracy from 65% to 92,5%.

Research on dimensionality reduction is also performed by Hermawan [11]. In his research, a neural network is used to classify consumers characteristics in buying which are based on seven (7) variables, namely: delivery speed, price rate, price flexibility, company image, service, seller image, and product quality. In his conclusion, Hermawan [11] states that the seven variables used to classify consumer characteristics could be reduced to four (4) variables. The variables determining the classification of consumer characteristics in buying products are: delivery speed, price rate, price flexibility, and product quality.

RESEARCH METHOD

The steps below are taken to accomplish the research objectives.

Preparing training and testing data

1. Training and testing data are taken from the data of aptitude test result of a senior high school in Yogyakarta. The number of the whole data used in the research is 140 data. One hundred (100) data are used for training and the other 40 are used for test.

The 100 training data comprise of 50 students from science major and 50 students from social studies major. The 40 test data comprise of 20 students from science major and 20 students from social studies major. All of the available data are duplicated according to the provision as seen in table 1.

Table 1: Data Duplication Provision

Type of Duplication	Variables reduced/ not included
D1	General Competence (GC)
D2	Mechanical Competence (MC)
D3	Numerical Competence (NC)
D4	Verbal Competence (VC)
D5	Language Competence (LC)

2. Develop a software to simulate back propagation neural network, with adaptive learning rate and momentum as seen in figure 1. Simulation is performed by using Neural Network Toolbox Matlab 7.1 device. The number of hidden layers is 20, momentum value is 0.6, initial learning rate value is 0.5, and permitted sum square error is 0.1.
3. Train and test neural network by using type D1. Recording iteration number required to identify senior high school students' major and the pattern identification accuracy.
4. Train and test neural network by using type D2. Recording iteration number required to identify senior high school students' major and the pattern identification accuracy.
5. Train and test neural network by using type D3. Recording iteration number required to identify senior high school students' major and the pattern identification accuracy.
6. Train and test neural network by using type D4. Recording iteration number required to identify senior high school students' major and the pattern identification accuracy.
7. Train and test neural network by using type D5. Recording iteration number required to identify senior high school students' major and the pattern identification accuracy.

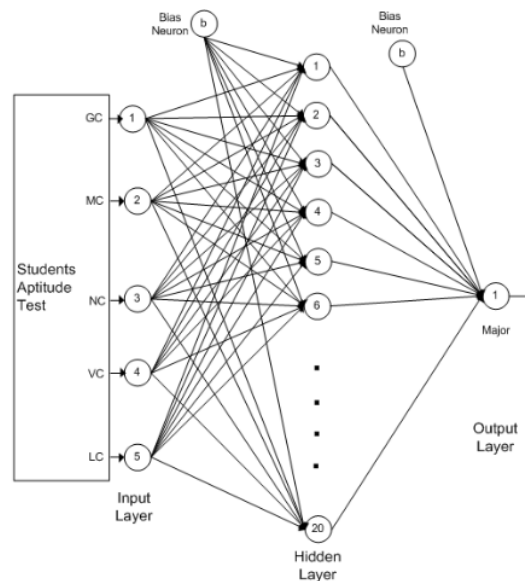


Figure 1: Neural Network Topology

RESEARCH FINDING AND DISCUSSION

Figure 2-6 show the process of neural network training by reducing General Competence variable, Mechanical Competence variable, Numerical Competence variable, Verbal Competence variable, and Language Competence

variable. After neural network training has been conducted in which various variables are reduced, convergence simulated neural network result is only obtained when General Competence variable is reduced. When Mechanical Competence, Numerical Competence, Verbal Competence and Language Competence are reduced, the simulated neural network could not achieve convergence until the 50.000th iteration.

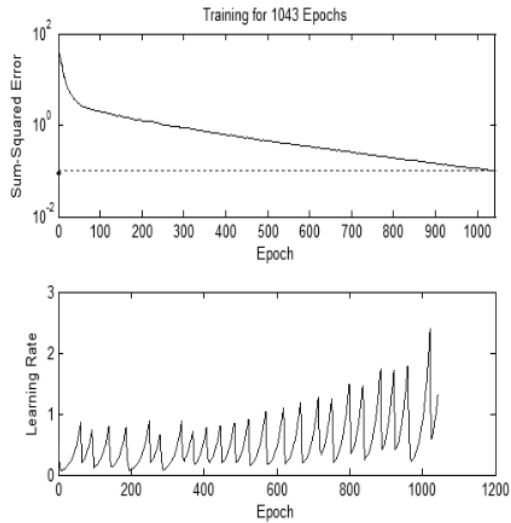


Figure 2: Neural Network Training Process by Reducing General Competence Variable.

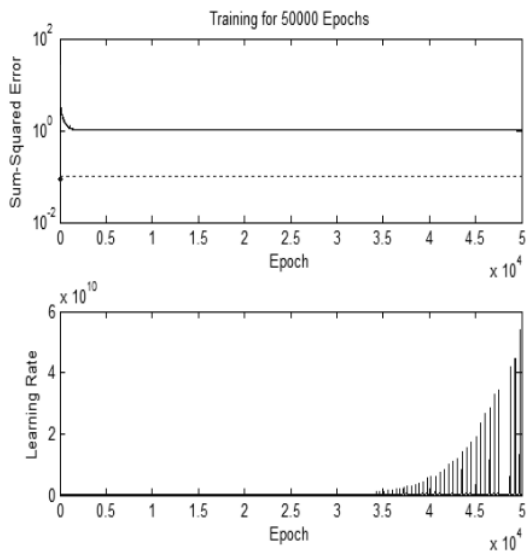


Figure 3: Neural Network Training Process by Reducing Mechanical Competence variable.

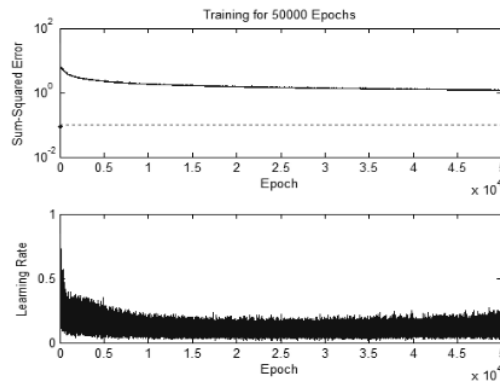


Figure 4: Neural Network Training Process by Reducing Numerical Competence variable.

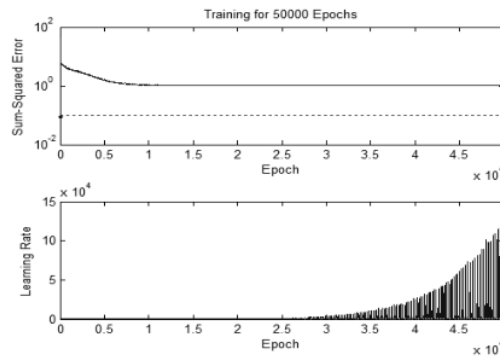


Figure 5: Neural Network Training Process by Reducing Verbal Competence variable.

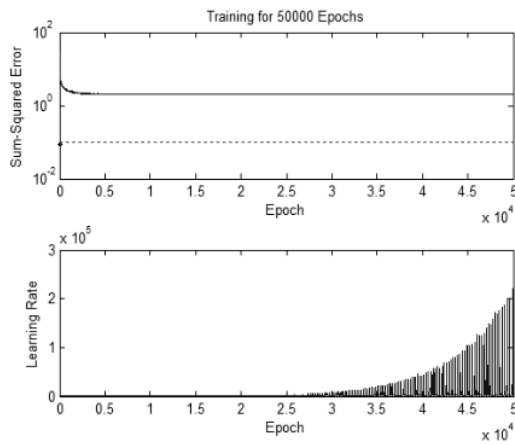


Figure 6: Neural Network Training Process by Reducing Language Competence variable.

Table 2 shows neural network training and testing result by using reduced variables. When General Competence variable is reduced, convergence network could complete its training at the 1043th iteration and could identify pattern tested with the accuracy of 97.5%. Based on the result, if the dimensionality reduction performed is compared to Hermawan's finding (2014), it can be seen that it is only capable of reducing the number of the iteration, while the identification accuracy of the pattern tested still remains the same which is 97.5%. This shows that General Competence is the disturbing variable, meaning that it is a variable included in the training but does not provide any contribution in major selection. However, by reducing the general competence variable, it is known that simulated neural network is not capable to enhance its capabilities to identify majors provided, which still remains at 97.5%. The 97.5% accuracy in identifying majors by using the four variables (General Competence is excluded) demonstrates that there is another variable which should be included in senior high school students' majoring to afford 100% accuracy.

Table 2: Training and Testing Result Which Variables are Reduced

	Reducing Variable				
	KU	KM	KN	KV	KB
Iteration	1043	*	*	*	*
Accuracy of Training	100%	NA	NA	NA	NA
Accuracy of Testing	97.5%	NA	NA	NA	NA

*not convergent up to the iteration of 50.000

The variable which should be included to reach 100 % major identification accuracy must contain additional information which could provide contribution in determining majoring for senior high school students. Although it has not been able to reach 100% accuracy, the variable employed by Nugroho & Retnawati [12] and Yuyun & Munawir [13] to identify the majoring for senior high school students is possible to be included. That variable is called subject grade.

CONCLUSION AND FUTURE WORKS

Backpropagation neural network with adaptive learning constant and momentum could identify students majors based on the General, Mechanical, Numerical, Verbal, and Language Competence variables. By performing dimensionality reduction process, it could be known that the General Competence variable is the disturbing variable. After the disturbing variable is removed the convergence neural network could complete faster iteration, which is at the 1043th iteration. The students major identification accuracy in which a variable has been reduced remains the same, which is 97.5%.

The future works of the research can be done by adding another variable to enhance major identification accuracy. The added variable must contain additional information which could increase major identification accuracy.

REFERENCES

- [1] Indonesia government regulation no 17/2010.
- [2] Hermawan, "High School Majors Selection Using Neural Network", Proceeding ICERI, UNY, 2014.
- [3] A.Hermawan, "Jaringan Saraf Tiruan", Penerbit Andi, 2006.
- [4] V. O. Oladokun, A. T. Adebanjo, and O. E. Charles-Owaba, "Predicting Students Academic Performance using Neural Network: a Case Study of an Engineering Course", Pasific Journal of Science and Technology. vol 9, no 1, pp. 72-79, 2008.
- [5] S. S. Mahapatra, and M.S. Khan, "A Neural Network Approach for Assesing Quality in Technical Education: an Empirical Study", International Journal of Productivity Quality Management, vol 2, no 3, pp. 287-306, 2007.
- [6] Budhi, Liliana, and S. Harryanto, "Cluster Analysis untuk Memprediksi Talenta Pemain Basket Menggunakan Jaringan Saraf Tiruan Self Organizing Maps (SOM)", Jurnal Informatika. vol 9, no 1 May, pp. 23-32, 2008.
- [7] S. Muknahallipatna, and B. Chowdhury, "Input Dimension Reduction in Neural Network Training – case Study in Transient Stability Assessment of Large System", Proceedings ISAP, pp 50-54, 1996.
- [8] M. G. B. Blum, M. A. Munes, and S. A. Sisson, "A Comparative Review of Dimension Reduction Methods in Approximate Bayesian Computation", Statistical Science, vol 28, no 2, pp 189-208, 2013.
- [9] G. E. Hinton, and Salakhutdinov, " Reducing the Dimensionality of Data With Neural Network", Science, vol 313, pp 504-507, 2006.
- [10] S. J. Perantonis, and V. Virvilis, "Dimensionality Reduction using a Novel Neural Network Based Feature Extraction Method", IJCNN, pp 1195-1198, 1999.
- [11] Hermawan, "Reduksi Dimensi dengan Menggunakan Jaringan Saraf Tiruan, Jurnal Sistem Cerdas", vol 1, no 1, 2002.
- [12] Nugroho and Retnawati, "Sistem Pendukung Keputusan Penjurusan di SMA Menggunakan Metode Neural Network Backpropagation (Studi Kasus SMA Islam Kepanjen Malang)", <http://ejournal.unikama.ac.id/index.php/JFTI/article/view/831/519>, 2015.
- [13] Yuyun and Munawir, , "Neural Network Algorithm for Determining a Majors for Senior High School", Prosiding Konferensi Nasional Komputer (KNN), ISBN 078-602-98563-0-9, STMIK Kharisma Makasar, 2011.

Authors' information



Dr. Arief Hermawan was born in Pekalongan (Indonesia) on 11th May 1970. He received his S1 and MS in Electrical Engineering Gadjah Mada University, Yogyakarta on 1994 and 2001. He received the Ph.D in Information Technology Education on 2013 from Yogyakarta State University. His Research focus on neural network, information system and information technology for education.

He is currently an Ass Professor at University Technology of Yogyakarta. He has published several research paper in conferences and journals. He also published book entitled neural network: theory and application.

Dr. Hermawan is currently IACSIT senior member and Indonesian Engineers Association member.

Arief_Hermawan_The Dimensional REducation to Improve the Speed adn accuracy of NN

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5
