Cognitive Classification Based on Revised Bloom's Taxonomy Using Learning Vector Quantization

1rd Eko Subiyantoro
Departement of Technoy Information
BBPPMPV-BOE Malang
Malang, Indonesia
ekovedc@ gmail.com

2st Ahmad Ashari
Departement of Computer Sciences and Electronics
Universitas Gadjah Mada
Yogyakarta, Indonesia
ashari@ugm.ac.id

3nd Suprapto
Departement of Computer Sciences and Electronics
Universitas Gadjah Mada
Yogyakarta, Indonesia
sprapto@ugm.ac.id

***Abstract*— The cognitive dimensions of the new bloom taxonomy consist of six categories, namely C1, C2, C3, C4, C5, and C6. The difference is using verbs at each cognitive level. Based on the cognitive level, it appears that the C1 level is the lowest thinking level while C6 is the highest thinking level. However, cognitive classification to develop students' knowledge towards high-level cognitive skills has not been applied to managing students in learning. The focus of this study is to determine the cognitive classification structure using the bloom taxonomy. Learning Vector Quantization (LVQ) is used to classify cognitive levels into three classes, namely Low Cognitive (CL), Medium Cognitive (CM), and High Cognitive (CH). The results showed that the cognitive classification of LVQ succeeded in classifying the cognitive domains into three cognitive classes, namely CL, CM, and CH with an accuracy of 97% through a learning rate of 0.3.**

***Keywords*— Bloom taxonomy, LVQ, cognitive low, cognitive medium, cognitive high.**

I. Introduction

The cognitive domain published by Bloom in 1956 consists of six levels, namely knowledge, understanding, application, analysis, synthesis, and evaluation [1]. In its development, the cognitive taxonomy is known as Bloom's taxonomy, which is the classification of various goals and skills set by educators for students. Bloom's Taxonomy aims to make it easier for teachers to classify what their students should learn at a certain time. David R. Krathwohl and Anderson in 2001, proposed the Revised Taxonomy fortyfive years later, known as the Revised Bloom's Taxonomy (RBT) [2].

Furthermore, Anderson and Krathwoh explain that students are said to understand if they can construct the meaning of learning messages, whether oral, written, or graphic, conveyed through teaching, books, or computer screens. Cognitive processes in the understanding category include: 1) cognitive processes of interpreting; 2) exemplifying cognitive processes; 3) cognitive processes of classifying; 4) summarizes cognitive processes; 5) concluding cognitive processes; 6) comparing cognitive processes, and; 7) cognitive processes explained.

The order of the levels of human thinking corresponds to the six cognitive levels in the taxonomy bloom. Given (C1) is the lowest level of thinking, while creating (C6) is the highest level of thinking. One of the main focuses of 21st century thinking skills in achieving learning goals is Higher Order Thinking Skills (HOTS) [3] [4]. Thinking skills are divided into two, namely LOTs and HOTS[5]. Thinking skills are important skills to be achieved during the learning process, apart from other skills such as communication skills, social skills, and socialization skills in a global society. HOTs can occur when someone rearranges information to expand knowledge in order to achieve goals to be achieved. Thus HOTs is a thinking process of students at a higher cognitive level that is developed from various concepts and learning taxonomies [6].

The current learning paradigm had changed into a new one. The teacher-centered transfer of knowledge had altered into the construction of knowledge. It puts students at the center of learning activities [7]. The new paradigm in learning focuses on the view that students are human beings who have the potential to learn and develop. In fact, learning in a class is classically to reach all students, teachers often design the same learning and classroom activities. This situation can be remedied by providing personal guidance, so that students' needs and difficulties can be overcome. However, given the number of students a class has, it is not easy for a teacher to deal with the varying levels and needs of students all the time [8].

Based on the background that has been presented, the focus of the problem is that RBT produces a taxonomic table which is an interrelation between cognitive processes and knowledge that can measure the depth and breadth of the learning objectives to be achieved. But the diversity of characteristics of students' cognitive abilities in a class is always a problem that is often faced by a teacher.

This research uses the Learning Vector Quantization (LVQ) method to classify cognitive into three cognitive classes (Cognitive Low (CL), Cognitive Medium (CM), and Cognitive High (CH)). LVQ input vector data is a pretest that is designed using RBT approach which consists of twenty questions. The target class in the LVQ network is defined as three classes, namely CL, CM, and CH so that three competitive layers are needed.

II. Material

A. Revised Bloom' Taxonomy (RBT)

RBT is prepared by taking into account contemporary developments in education-related fields. The fields of intersection include psycho-education, neurosciences, education, and socio-culture. According to Conklin [9], the application of meta-analysis in RBT is a contemporary idea, because the word has been used very often in the last decade. Research in the field of neuroscience shows the existence of a corporation between the activities of certain neurons in thought processes. The more mature corporations point toward metacognition. RBT contains the word 'metacognitive' as part of the knowledge dimension as shown in Fig. 1.

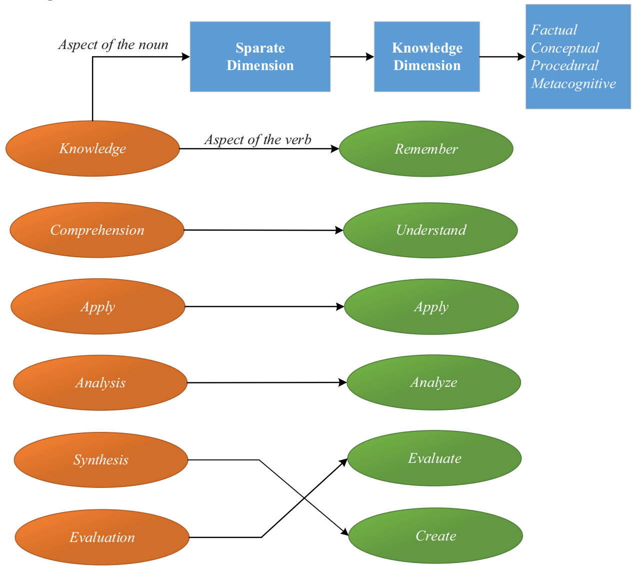

Fig. 1. Aplication Revised Bloom's Taxonomy

In the old Bloom's Taxonomy, the category 'knowledge' became the main category of the first level. RBT "take" this category of "knowledge" from the Taxonomy and turn it into a measure to be achieved. That is, "knowledge" is the attainment of cognition itself. The term 'knowledge' is further divided into sub-categories which correspond to developments in the field of neuroscience and research in the field of psychoeducation as follows knowledge: factual, conceptual, procedural, and metacognitive.

B. Learning Vector Quantization (LVQ)

Learning Vector Quantization was first developed by Kohonen in 1996 as a supervised Artificial Neural Network (ANN) [10]. Learning Vector Quantization is a learning method in supervised competitive layers that will automatically learn to classify input vectors [11]. The distance between the input vectors will determine the class or category. If the input vector distances have very close values, the input vectors are grouped in the same class.

The LVQ network architecture can be grouped into three parts, namely: the input layer, the competitive layer, and the output layer. The input layer contains input vectors that describe the features of the pattern to be recognized. In the Competitive layer, each unit layer will group the input vectors. The clustering results are obtained from the calculation of the distance between the input vector and the unit layer. The LVQ method can be used for a grouping process where the number of groups has been determined according to the target architecture design or predefined classes.

The LVQ algorithm is used to find output units that have a matching pattern with the input vector. The winning unit index is compared to the target. The results of the comparison are used to improve the weights as shown by the algorithm in Fig. 2 [13]. The notation used in the LVQ algorithm is

X is a training vector $(X_1, \ldots, X_i, \ldots, X_n)$.
$Target$ is the correct category.
w_j is the weight vector
Cog_j is the class represented.
α is the learning rate.
$\| X - w_j \|$ is the distance between the cognitive level and weight vector.

Algortihm 1 Algorithm for Learning Vector Quantization
Initialisation:
vectors x_i , w_i
learning rate α
LOOP
While (true) , **do**
  for i = x to maxIter, **do**
    $Cog_{(j)} = \sum (w_{ij} - x_i)^2$
    $j = \min (Cog_{(j)})$
      if (Target $= c_j$)
        $w_{j(nw)} = w_{j(old)} + \alpha[x - w_{j(old)}]$
      if (Tartget $\neq c_j$)
        $w_{j(nw)} = w_{j(old)} - \alpha[x - w_{j(old)}]$
      end if
  end for
$\alpha = \alpha$ - dec α
end While condition

Fig. 2. LVQ Algorithm

Learning rate (α) serves to converge the value of weight change. The value of the learning rate coefficient ranges from <α <1.The value of α deduction is β with (1) [14].

$$\beta = 0.7 \times (p)^{1/n} \quad (1)$$

where p = number of hidden units and n = number of input units. Learning rate upated (α) by (2)

$$\alpha_i = \alpha_i \times e^{(-\beta \times i)} \quad (2)$$

III. METHODS

The selection of the LVQ method in the cognitive classification process based on RBT with several considerations, among others, can summarize the pretest data set into an input vector with a small code for classification, the error value is smaller than artificial neural networks such as backpropagation, and the resulting model can be updated gradually.

A. LVQ Architecture on Cognitive Classification RBT

The LVQ method introduces the input data in the form of answers to pretest questions which are a combination of six cognitive levels, namely (C1, C2, C3, C4, C5, C6) which are presented in vector form. . The processing of each neuron is to find the distance between an input vector to the corresponding weights W_{iL}, W_{iM} and W_{iH}. W_{iL} is the weight vector for each neuron in the input layer to the first neuron, W_{iM} is the weight vector in the input layer to the second neuron, and W_{iH} is the weight vector at the input layer to the third neuron in the output layer, which is connected to the activation function N_{CL}, N_{CM}, and N_{CH}. The determination of the cognitive class (CC) is based on the optimum value of the L function, which is the minimum value of the activation function. The classification structure and function model using the Learning Vector Quantization in Fig. 3.

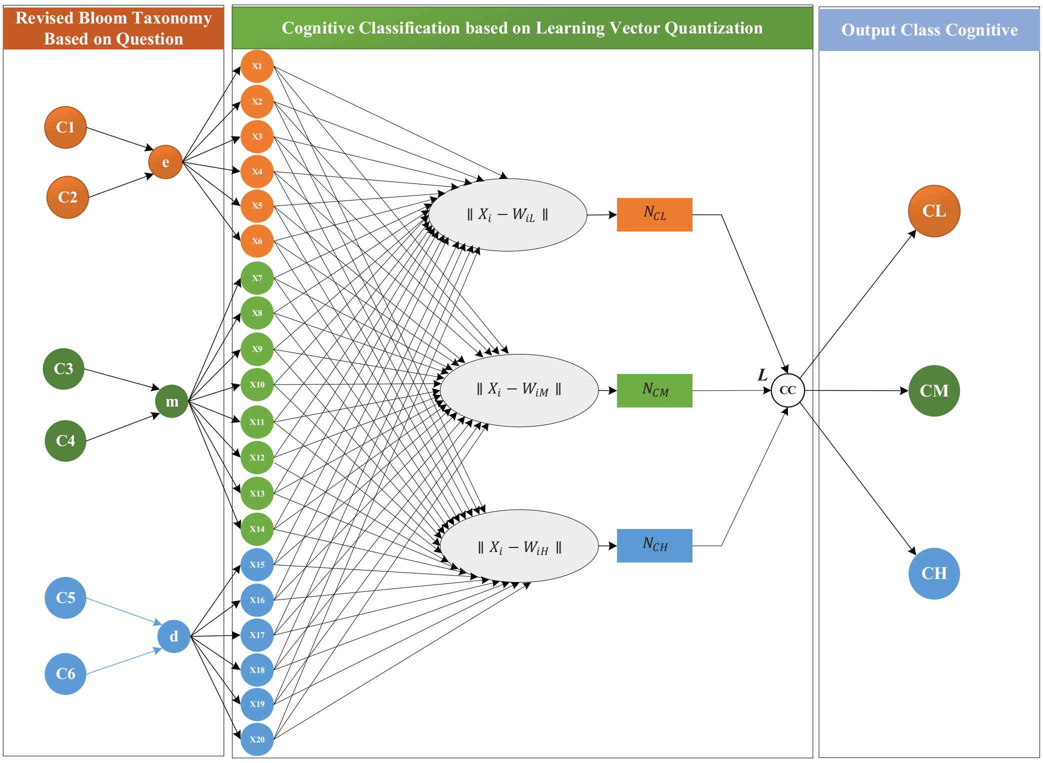

Fig. 3. LVQ Architecture on Cognitive Classification RBT

The criteria for the number of comparisons of questions based on the criteria for easy, medium, and difficult questions are 3: 4: 3. [15]. The number of cognitive sets in the form of questions is N = {e, m, d}. Where N is twenty questions, e is the number of easy category questions of thirty percent as a cognitive representation of C1 and C2, m is the number of medium category questions forty percent as a cognitive representation of C3 and C4, and d is the number of difficult category questions of thirty percent as a cognitive representation of C5 and C6. Table I shows the percentage of questions for each RBT level cognitive.

TABLE I. QUESTION MAKING CRITERIA BASED ON RBT

No	Cognitive Level						Number of Question	(%)	Category
	C1	C2	C3	C4	C5	C6			
1.	√						3	15	Easy
2.		√					3	15	Easy
3.			√				4	20	Medium
4.				√			4	20	Medium
5.					√		3	15	Defficult
6.						√	3	15	Defficult

B. Cognitive Classification Model

The process of training and learning the LVQ method uses vector operations. The patterns will be presented in vector form, where the processing that occurs for each neuron is to find the distance between an input vector to the corresponding weights W_1, W_2 and W_3. W_1 is a weight vector that connects each neuron in the input layer to the first neuron in the output layer, the activation function (F)used in the LVQ network is linear. The goal is to obtain the same output as the input, according to the linear function formula, namely $y = x$. The function F_1will map $y_{\_in1}$ to $Y_1 = 1$ if $| X - W_1 |$ smaller than $| X - W_2 |$, $| X - W_1 |$ smaller than $| X - W_3 |$, and $Y_1 = 0$ otherwise. Likewise, the activation functions F_2 dan F_3 will map with the same conditions. Values of N_{CL}, N_{CM}, and N_{CH} are by calculating the distance between the input vector with the weight vector shown in (3) (4) and (5).

$$e \rightarrow X_1 \dots X_6, m \rightarrow X_7 \dots X_{14}, d \rightarrow X_{15} \dots X_{20}$$
$$W_L \rightarrow N_{CL}, W_M \rightarrow N_{CM}, W_H \rightarrow N_{CH}$$

$$N_{CL} = \sqrt{\sum_i (X_i - W_{iL})^2} \tag{3}$$

$$N_{CM} = \sqrt{\sum_i (X_i - W_{iM})^2} \tag{4}$$

$$N_{CH} = \sqrt{\sum_i (X_i - W_{iH})^2} \tag{5}$$

If i is 20 then the input vector can be written as $X_i = (X_1, X_2, X_3, \dots X_{20})$ and the weight vectors W_L, W_L, W_L are weight vectors that connect each neuron in the input layer to the first neuron in the output layer $W_{iL} = (W_{11}, W_{21}, W_{31}, \dots, W_{20(1)})$, $W_{iM}(W_{11}, W_{21}, W_{31}, \dots, W_{20(1)})$ dan $W_{iH} = (W_{11}, W_{21}, W_{31}, \dots, W_{20(1)})$. Then the N_{CL} value on cognitive low, N_{CM} value on cognitive medium, and N_{CH} value on cognitive high can indicate differences in cognitive level for each class.

The optimum method based on LVQ is used by some researchers [16] [17] to state that L is a classification of the optimum condition for CS. L is defined in three probabilities of the optimum conditions, namely; 1) CL, 2) CM, and 3) CH are shown in shown in (6) (7)

$$L = arg\ min \parallel \{N_{CL}, N_{CM}, N_{CH}\} \parallel, \tag{6}$$

$$CC = \begin{cases} CL, if \quad L = N_{CL} \\ CM, if \ L = N_{CM} \\ CH, if \ L = N_{CH}. \end{cases} \tag{7}$$

IV. Result

A. Determination of training data with RBT and cognitive targets

Teachers were selected as cognitive references to obtain ideal cognitive characteristics because teachers are the best evaluators of cognitive skills and teachers have qualifications as pedagogical assessors shown by their diploma, certificate, and teaching experience [18]. The reference weight given by the teacher can affect cognitive grade scores (CC). Table II is a cognitive reference from the teacher as training data in the LVQ network.LVQ Network Training on Cognitive Classification

TABLE II. Cognitive reference as training data on the LVQ algorithm

No	Vector Input																				Target
	C1			C2			C3				C4				C5			C6			$Cl, Cm,$
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}	Ch
1.	1	1	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	Cl
2.	1	1	1	1	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	1	Cl
3.	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	0	0	0	1	0	Cl
4.	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	Cl
5.	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	1	0	0	0	Cl
..																					Cl
16.	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	Cm
17.	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	1	Cm
18.	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	1	0	Cm
19.	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0	0	1	0	0	Cm
20.	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	0	0	0	Cm
...																					Cm
36.	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	Ch
37.	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	Ch
38.	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	Ch
39.	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	Ch
40.	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	Ch
..																					Ch
50.	0	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	Ch

The initial learning rate is determined to be 0.3 based on equation (2), in the first iteration the α value becomes 0.182, and the α value will continue to decrease so that at the maximum iteration (100) the α value becomes 0. Figure 4 shows the learning rate test with a value ranging from 0.01 to 0.07.

The learning rate test shows that the accuracy drops to 70% starting at a value of 0.065 to 0.1, while the highest accuracy is 100% at a learning rate of 0.025 to 0.04. When the learning rate is below 0.02, the accuracy has decreased by 86%. Thus a learning rate that is too small results in faster convergence, while a learning rate that is too large results in diverging which affects accuracy.

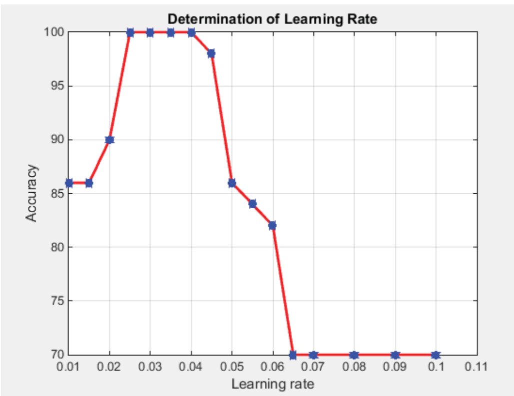

Fig. 6. Graph of learning rate testing on cognitive classification

Based on the learning rate test used in this study is 0.03 to produce a fixed weight value of W_1, W_2, W_3, the complete comparison of the initial and final weight values is presented in Fig. 7. These weight values are embedded in the LVQ algorithm to produce three cognitive classes of students, namely cognitive low, cognitive medium, and cognitive high.

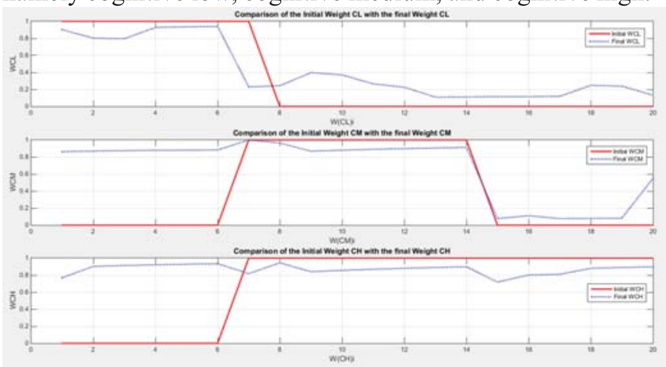

Fig. 7. Comparison graph of initial weight and final weight

B. Cognitive Classification Results

The results of the cognitive classification of 32 students on the LVQ algorithm show that the results of the experiment are in accordance with the target class that was designed are shown in Fig. 8. This research has succeeded in developing a cognitive classification model based on LVQ-based revised bloom taxonomy called RBT Cognitive Classification Structure. This modeling algorithm has embedded in the system test application. Through the cognitive LVQ network, students had classified into three cognitive domains, namely; Cognitive Low (CL), Cognitive Medium (CM), and Cognitive High(CH). The teacher's cognitive ability through RBT is the training data in the LVQ network

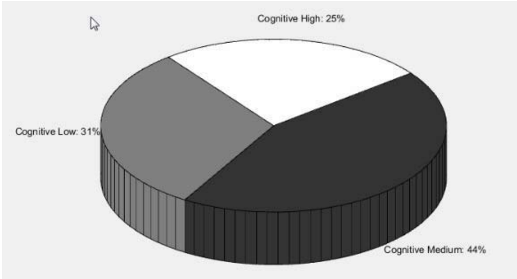

Fig. 8. Results of cognitive classification with LVQ networks

V. Conclusion and Future Work

This RBT cognitive classification research results that this developed model can identify cognitive students with high accuracy can reach 97% through the determination of the right learning rate on LVQ networks. 32 students had classified into three cognitive groups namely; 1) Cognitive low, 2) Cognitive medium, and 3) Cognitive high, with 31% results are low cognitive, 44% medium cognitive, and 25% high cognitive.

As the future work of this work, making learning path models based on cognitive classification and assessment models that have been developed. Thus each student knows the profile and sequence of learning objects that must be learned.

Acknowledgment

Director General of Higher Education of the Ministry of Education and Culture Indonesia, National Competitive Research Program for Doctoral Dissertation Research (PDD) in 2020.

References

[1] Bloom, B. S. (1956). Taxonomy of education objectives. 207.

[2] Krathwohl, *et al.*, "A Taxonomy For Learning, Teaching, And Assessing: A Revision Of Bloom's Taxonomy Of Educational Objectives," *New York Longman*, 41(4), 302, 2011.

[3] Saido, *et al.*, "Teaching strategies scale for promoting higher order thinking skills among students in science," *Proceedings of ISER 5th , International Conference,* 91–94, 2015.

[4] D. Sukla and A.P. Dungsungneon, "Students Perceived Level and Teachers Teaching Strategies of Higher Order Thinking Skills A Study on Higher Educational Institutions in Thailand," *Journal of Education and Practkice*, 7(12), 211–219, 2016.

[5] Coffman, D. M. Thinking about thinking: An explanatory of preservice teachers' views about higher order thinking skills. Dissertation,The University of Kansas, Kansas. Retrieved from https://kuscholarworks.ku.edu/handle/1808/15086?show=full, 2013.

[6] Lewis, A., & Smith, D. Defining higer order thinking: Theory into practice. Journal of Teaching for Higher Order Thinking, 1993.

[7] Jonassen, D., "Designing Constructivist Learning Environment". In C.M.Reigeluth (Ed.), Instructional-Design Theories and Models, Volume II: A New Paradigm of Instructional Theory. Mahwah, NJ, Lawrence Erlbaum, 1999.

[8] Santos, Á., Gomes, A., & Mendes, A. (2013). A taxonomy of exercises to support individual learning paths in initial programming learning. Proceedings - Frontiers in Education Conference, FIE, 87–93.

[9] Conklin, J. (2005, Spring). Book Reviews : A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of L. Educational Objectives. Educational Horizons, 83, 154-159.

[10] E. M. Soleiman and A. Fetanat, "Using Learning Vector Quantization (LVQ) in Intrusion Detection Systems," *International Journal of Innovative Research in Advanced Engineering (IJIRAE)*, 1(10), 15–19, 2014.

[11] M. F. Umer and M. S. H. Khiyal, "Classification of Textual Documents Using Learning Vector Quantization," *Information Technology Journal*, 2007.

[12] C.Chen, *et al.*, "Supervised Learning Vector Quantization for Projecting Missing Weights of Hierarchical Neural Networks, "*WSEAS Transactions on Information Science and Applications* 7(6):799-808, 2010.

[13] Fausett, "Fundamentals of Neural Networks Architectures, Algorithms, and Applications," *New Jersey: Prentice-Hall, Inc.*, 1994

[14] Marcin Blachnik and Włodzisław Duch, "LVQ algorithm with instance weighting for generation of prototype-based rules," Elsevier, 2011.

[15] Z.Giani, "Analisis Tingkat Kognitif Soal-Soal Buku Teks Matematika Kelas VII Berdasarkan Taksonomi Bloom," *Jurnal Pendidikan Matematika, Department of Master Program on Mathematics Education, Sriwijaya University,* 2015.

[16] H-H. Song and S-W. Lee, "LVQ Combined with Simulated Annealing for Optimal Design of Large-set Reference Models," *Neural Networks Elsevier Science Ltd*, vol. 9, no. 2, pp. 329-336, 1996.

[17] D. K Kim, *et al.*, "Generalized LVQ for Optimal Reference Vectors using a Differentiable MIN Module," *Proceedings of International Conference on Neural Information Processing*, pp. 1937-1942, 1994.

[18] M.A. Syufagi, *et al.*, "Petri Net Model for Serious Games based on Motivation Behavior Classification," *International Journal of Computer Games Technology (IJCGT)*, 2013.