Feature extraction and classification of learners using neural networks

Tomohiro Hayashida
Graduate School of Engineering
Hiroshima University
Higashihiroshima, JAPAN
hayashida@hiroshima-u.ac.jp

Toru Yamamoto
Graduate School of Engineering
Hiroshima University
Higashihiroshima, JAPAN
yama@hiroshima-u.ac.jp

Shin Wakitani
Graduate School of Engineering
Hiroshima University
Higashihiroshima, JAPAN
wakitani@hiroshima-u.ac.jp

Ichiro Nishizaki
Graduate School of Engineering
Hiroshima University
Higashihiroshima, JAPAN
nisizaki@hiroshima-u.ac.jp

Shinya Sekizaki
Graduate School of Engineering
Hiroshima University
Higashihiroshima, JAPAN
sekizaki@hiroshima-u.ac.jp

Yusuke Tanimoto
Graduate School of Engineering
Hiroshima University
Higashihiroshima, JAPAN
g180046@hiroshima-u.ac.jp

***Abstract*—The aim of this study is to predict the achievement degree of each student at the end of a lecture, based on a simple questionnaire result which regularly surveys degree of the subjective understanding conducted to students in a class. In this study, the feedforward neural networks (FNNs) and decision tree are used for the prediction and the classification. A FNN which is with a multiple input and multiple output structure is well known that it has high performance for multidimensional data prediction or classification. Therefore, FNNs are considered to be suitable for the problem dealt with in this study such that student classification based on multiple questionnaire results. Additionally, it is possible to analyze students' learning process in detail by using a decision tree that can obtain student classification rules in an explicit form. This study conducts an experiment using data of a classification six times questionnaire surveys and a final examination and constructed a system for student classification based on the the answers of three questionnaire. Experiment has succeeded in roughly classifying learners into three clusters based on achievement degree. It means that the proposed method is predict the potential comprehension degree of a student. Sequentially, by providing additional education for the students who are classified as a low degree, it is expected to be able to take countermeasures for not becoming "dropout" in early stage.**

***Index Terms*—Learning system, neural networks, feature extraction, classification of learners**

I. Introduction

Due to the rapid development of advanced information and communication technology, diversification of education, especially individual learning such as home learning and distance learning progresses. Such situation makes the role of the learning support system important. In a learning support system, it is necessary to appropriately describe the contents of education received by the learners and a learner model indicating the change in proficiency degree. At present, however, it can not be said that any appropriate learner model has been constructed. Therefore, it is difficult for the learners to improve motivation to learn or to attract sustained and aggressive learning behaviors.

In order to improve learners proficiency or motivation, it is desirable for the instructor to grasp proficiency degree of the learners in real time and provide individual learning support corresponding to each learner. The aim of this study is constructing a system to provide such learning support suitable for each learner, and proposes a method for real-time grasp of learners' proficiency level. The proposed method of this study predicts the results of the final examination conducted at the end of a lecture using the neural networks and a decision tree based on information observed from individual learners during the corresponding class. By this scheme, it is expected that it will be possible to prevent lowering of motivation for learning or proficiency in advance, by providing additional education to the learners who are predicted to have poor results before the examination. A FNNs have been applied to many multidimensional data classification or prediction [1], [4], for example, classification and feature extraction of nonlinear time-series data in economic field [2], [3]. This study conducts numerical experiments using a cluster of learners who have taken a same class and the experimental result indicates the usefulness of the proposed method.

II. Classification of the learners

In consideration of raising the mastery degree of all students who take a class to a certain degree or higher, it is generally necessary to support learning according to the learning degree of each learner. However, it is difficult to grasp the proficiency degree of all learners, for example, as a result of examinations or mini-exams, supplementary classes for students who are considered to have low proficiency level, some kinds of measures are often taken not to become so-called "fallout". Additionally, it becomes possible to provide learning support suitable for each learner, and it is expected to greatly improve learning efficiency by offering learning support not only based on proficiency of the learners at present but also on the degree of proficiency in the future. This study

978-1-5386-1174-6/18/$31.00 ©2018 IEEE

assumes classes to be handled multiple times by dealing with syllabus contents in elementary school, junior high school, high school, university or other educational institutions. In the target classes, it is assumed that some not only an final examination but also multiple reports, questionnaires, mini-exams in order to grasp the learners' proficiency degree are conducted. Here, a plurality of prior reports, questionnaires, or mini-exams are called *preliminary survey*, in this study.

Periodic participation of the class make learners' proficiency degree increase, however, the growing process may be different depending on learning characteristics of each learner. If the results of the final examination of each learner, that is, future proficiency levels based on learning characteristics can be roughly predicted, then learning support for the learners with low future proficiency degree becomes possible. Result of the preliminary survey can be observed as time-series data representing the learning characteristics of each learner, if preliminary surveys are performed periodically. An instructor can infer the learning characteristics of each learner if the result of the final examination can be estimated based on such data of the preliminary surveys. In other words, by letting the result of preliminary surveys be input and the result of final examination be output, the learning characteristics can be estimated by modeling the input-output relation between them.

Since such input-output relationships are modeled by non-linear systems of multiple inputs and multiple outputs, a neural network is appropriate for predicting the learners' future proficiency with high accuracy. This study constructs a learner's model using a 3-layer feedfoward neural network (FNN) as shown in Fig. 1.

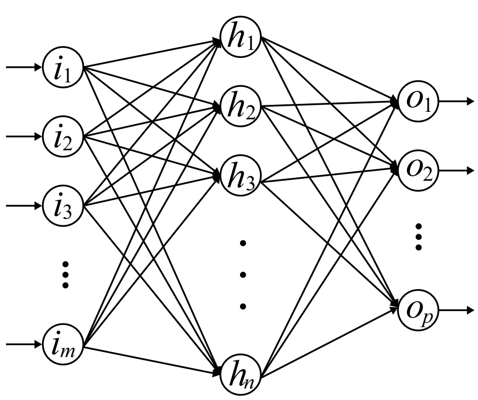

Fig. 1. A FNN for prediction of the learners

The number of units in the input layer m is same as number of dimensions of the system input and is the number of dimensions of the result of the preliminary survey. Let the number of units in the middle layer be n, the number of units in the output layer be p which corresponds to the result of the final examination. In case that the output value of the neural network is taken as the predicted value of the score of the final examination, the output value in the range of $[0, 1]$ can be interpreted as the score rate of the final examination with $p = 1$. In another case such that the learners are classified into 5 clusters of "very good", "good", "normal", "bad", and "very bad", p indicates number of clusters and each node of the output layer corresponds to each clusters. Let $o_l \in [0, 1], l = 1, 2, \ldots, p$ be the output values from the nodes of the output layer, where o_l can be interpreted as the priority of cluster l. The learners are clustered based on output values $o_l, l = 1, 2, \ldots, p$ for input information $(x_1, x_2, \ldots, x_m)$.

III. Numerical Experiments

This study conducts numerical experiments using the records of 205 learners of taking the same lecture which held for 5 days. In this lecture, as preliminary survey, questionnaire surveys on proficiency degree which is evaluated by each learner by herself/himself are conducted after each class, and a final examination is conducted after all classes. Questionnaire survey consists of 10 questions in the form of "how much can you explain about . . ." and self-evaluation are done in 4 stages. As the proficiency degree is higher, the learners are instructed to answer higher numerical values. The questionnaire survey is conducted 6 times in total, that is, before the 1st class starts and after the class of each day of 5 days finishes.

Fig. 2 shows a histogram of the overall score of the final examination.

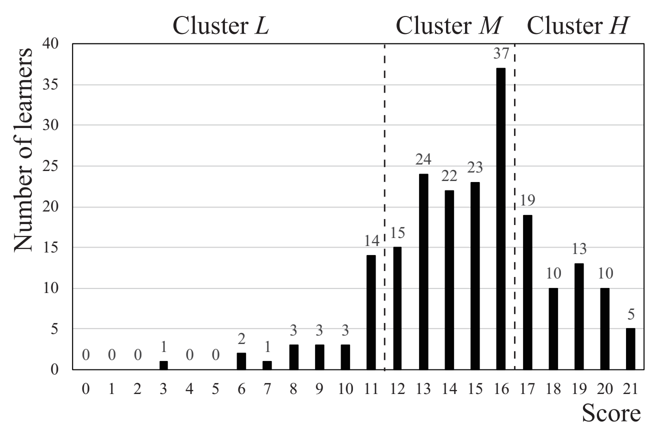

Fig. 2. A histogram of the final examination

Based on the results shown in Fig. 2, the learners are divided into 3 clusters: cluster H with high score, cluster M with medium score, and cluster L with low score. The experiment in this section aims to classify each learner into one of these 3 clusters based on the answers to the questionnaire of preliminary survey. From the viewpoint of learning support, clustering of the learners should be done as accurately as possible and early in time. Therefore, in this study, clustering of the learners is performed based on answers of as few questionnaires as possible.

A. Classification using the Neural Network

In the lecture targeted in this experiment, questionnaires consisting of 10 questions asking the subjective proficiency degree of the learners are conducted 6 times. The contents of the questionnaire of each time are same. By preliminary experiments, it is known that clustering accuracy of the learners gets higher if the system can observe not only the subjective proficiency degree but also the amount of change of them. Therefore, this study uses the following 6 types of data $S_0, S_1, \ldots, S_5$ as input which consist of not only the questionnaire responses but also the change of answers, and verify the classification accuracy for each input data.

$$S_0^i \equiv X_0^i,$$
$$S_t^i \equiv S_{t-1}^i \cup X_t^i \cup \Delta X_t^i \ (t = 1, 2, \ldots, 5),$$

where, $X_t^i \equiv \{x_{t,k}^i, k = 1, 2, \ldots 10\}, i = 1, 2, \ldots, 205, t = 1, 2, \ldots, 5$ indicates the response of the questionnaire on the t-th day of the learner i. Also, the answer of the questionnaire conducted before the class on the first day is represented as $x_{0,k}^i$, and $\Delta X_t^i \equiv X_t^i - X_{t-1}^i, t = 1, 2, \ldots, 5$ indicates the change amount of the answers.

Back propagation (BP) [5] is applied for updating the connecting weights and thresholds of an FNN. The teacher signal consists of the questionnaire answer $S_0^i, S_1^i, S_2^i, S_3^i, S_4^i$ and S_5^i, or apart of them as input value and the result of final examination as the output value. The learning rate and the number of iteration of BP are set as 0.025 and 1000, respectively. The number of nodes in the output layer is set as $p = 3$ and the output value can be interpreted as priority for clusters H, M and L. Here, let $(o_H, o_M, o_L) = (o_1, o_2, o_3)$. That is, if the questionnaire answer of a new learner is used as input to the FNN, the output value is interpreted as the priority of each cluster, and the learner is classified based on the output value. The output value of the teacher signal is $(1, 0, 0)^T$ for cluster H, $(0, 1, 0)^T$ for cluster M, and $(0, 0, 1)^T$ for cluster L. It can be said that the FNN, in which learning process using such teacher signals sufficiently performed, approximates the input-output relationship of learning characteristics of learners with high accuracy.

Fig. 3 shows a graph that the horizontal axis shows input information and the vertical axis shows average output error.

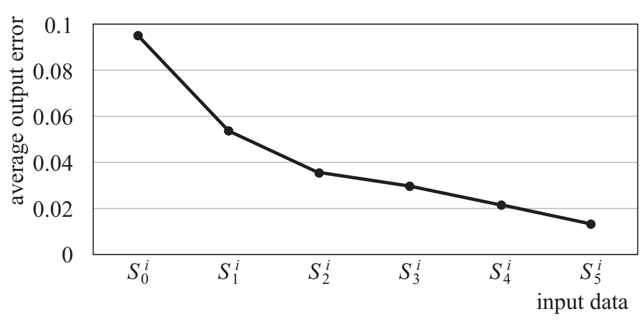

Fig. 3. Preliminary experiments for choosing input data

Either one of $S_0^i, S_1^i, \ldots, S_5^i$ is used as the input value of the teacher signals. As the information given increases, the error between the output value of the neural network and the output value of the teacher signal becomes small. From Fig. 3, if the input information is increased as S_0^i, S_1^i, S_2^i, the average output error decreases sharply. The average output error for the input data S_2^i is less than 5%, and even if the input information is increased more than S_2^i, the average output error does not decrease greatly. This means that the learners can be roughly classified by the result of the final examination based on the preliminary survey up to the 2nd day. Therefore, this study employs S_2^i as input information for executing clustering of the learners.

B. Selection of learner records

In order for the neural network to acquire learning characteristics of the learners of all clusters, it is necessary to select so that the number of learners belonging to each cluster is almost the same number. However, as shown in Fig. 2, cluster H consists of 57 learners, cluster M consists of 121 and cluster L consists of 27, and number of learners belonging to M is largest. Therefore, this study replicates each number so that the number of cluster H is doubled, L is quadrupled, as shown in Fig. 4.

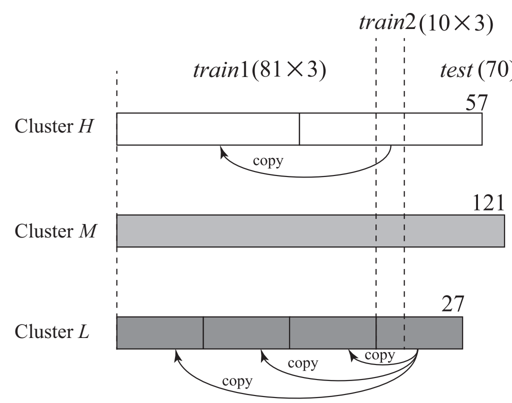

Fig. 4. Data generation for training of the neural network

Of the new generated 343 learner records, 81 learners in each cluster, total 243 learner records are selected as the teacher signal, and let this dataset be $train1$. Here, for clusters H and L, 138 (57+27×3) replicated learner records which are added by replication are preferentially selected as $train1$, and the remaining 105 learner records are randomly selected from the original data. The dataset $train1$ is used for learning of the neural network, which updates weights and thresholds by BP. Of 100 learner records which are not selected as $train1$, the 10 learner records are randomly selected from each cluster (total 30 learner records), and let this dataset be $train2$. Let the remaining 70 learner records which are not selected as either $train1$ or $train2$ be a dataset $test$.

C. Classifying rules using neural networks

As mentioned above, the output value (o_H, o_M, o_L) of the neural network is interpreted as the priority indicating which cluster the input data belongs to. Generally, assigning to the cluster having the highest priority among 3 clusters, allocating with a probability proportional to the magnitude of the output value, or a decision tree based on the priority, are used for clustering based on the neural networks. This study chooses appropriate method for clustering by using the dataset $train2$ which is not used for neural network learning. As a result the appropriate method is that uses the decision tree shown in Fig. 5.

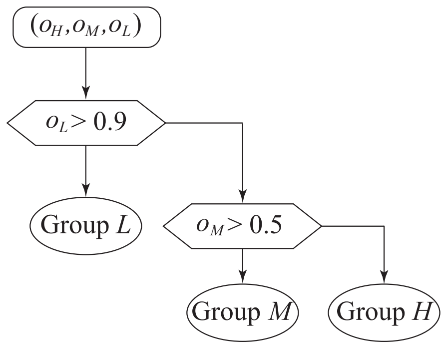

Fig. 5. Decision tree for classification

D. Experimental result

This study applies FNN shown in Fig. 1, and the dataset $S_2^i, i \in train1 \cup train2 \cup test$ of the preliminary survey are given as input information. Based on the output of FNN (o_H, o_M, o_L), the learners are classified into one of the clusters H, M or L using the decision tree shown in Fig. 5. Table I shows the result of the classification of 70 learners belonging to $test$.

TABLE I
CLASSIFICATION OF THE LEARNERS USING THE INPUT DATA $S_2^i \in test$

		Prediction H	Prediction M	Prediction L
True Cluster	H	**18**	5	0
	M	13	**14**	3
	L	0	0	**17**

From Table I, it is found that the result of the questionnaire until the 2nd day can predict the result of the final examination of the unknown learner with high accuracy. In particular, it succeeds in extracting all 17 learners belonging to cluster L. This means that the results of the final examination performed after the lecture are bad, that is, learners whose future proficiency is low are extracted. The experimental result indicates that the learners who are subject to additional learning support can be discovered at early stage by the proposed method.

IV. CONCLUSION

This study proposes a method to classify the learners participating in a lecture according to future proficiency degree based on the preliminary survey other than the final examination such as quizzes, reporting tasks, questionnaire, and so forth. A FNN is employed for the classification of the learners, and the result of preliminary survey is used for input information of FNN and classification of future proficiency degree is used for output. A numerical experiments using data collected from learners in a single kind of lecture is conducted. As the result, it succeeded in extracting learners with low proficiency in the future.

As one of the future works, Experiments using the preliminary surveys other than questionnaires are necessary to confirm the usefulness of the proposed method.

REFERENCES

[1] R. Asadi and S.A. Kareem, "Review of feed forward neural network classification preprocessing techniques," Proceedings of the 3rd International Conference on Mathematical Sciences, vol. 1602, pp. 567–573, 2014.

[2] A. Bahrammirzaee, "A comparative survey of artificial intelligence applications in finance: Artificial neural networks, expert system and hybrid intelligent systems," Neural Computing and Applications, vol. 19, pp. 1165–1195, 2010.

[3] P.G. Michaelides, E.G. Tsionas, A.T. Vouldis, K.N. Konstantakis, and P. Patrinos, "A semi-parametric non-linear neural network filter: Theory and empirical evidence," Computational Economics, vol. 51, pp. 637–675, 2018.

[4] K.J. Niky, O.K. Samrat, and C.V. Maheshwari, "Feed forward neural network classification for INDIAN Krishna Kamod rice," International Journal of Computer Applications (0975-8887), vol. 134, pp. 38–42, 2016.

[5] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, "Learning representations by back-propagating errors," Nature, vol. 323, pp. 533–536, 1986.