

Knowledge-Based Method for Revising Misclassification on Learners' Comprehension

Hiroshi Murai

Abstract—This study is about Learning Support using classification of learners according to their type of learners' comprehension on IP education. The proposed system has an automatic recommendation function of learning contents. In order to support automatically, we use unsupervised classification method such as SOM. But classification procedure usually leaves us with a small number of poorly classified learners that are often located at the boundaries between two clearly assigned types. In such a case, homogenizing the classification by reassigning the learners to one or the other type is desirable. In order to support recommendation of suitable learning contents, I examine knowledge-based method for revising misclassification on learners' comprehension.

Keywords—Pattern Classification, Learning Support, e-Learning.

I. INTRODUCTION

Since an oral examination in law is proceeded on legal qualifying examination to know whether the learner understands IP law or not, we need individual guidance to a learner. But the individual guidance puts a large burden on a teacher. Therefore, we have attempted to support intellectual-property (IP) education automatically in the past researches [1]-[5]. The proposed system that has recommended learning contents suitable for a learner. The recommendation has been decided by the estimation of type of the learners' comprehension.

In the reference [5], we examined the similarity of the type of learners' comprehension classified by SOM classifier. The results show that some classes are similar to each other because of the characteristics of the learners' comprehension. Therefore, we need to divide the different types of learners classify in the same unit when we recommend the learning contents suitable to the learner. But, increasing of the variation of the type of learners' comprehension causes a large burden on the teacher to prepare for the individual guidance. Therefore, it is important that we should not increase the class by making rules of revising misclassification based on the teaching knowledge.

On the other hand, we proposed a pattern classification method which integrates the advantages of both the neural network and knowledge-based system on the remote sensing image analysis [7]. We found that the misclassification can be revised more easily because of introducing the geographical

knowledge into the system.

In this paper, we apply the knowledge-based method for revising misclassification on learners' comprehension and illustrate that we can revise easily because of introducing the teaching knowledge.

II. METHODOLOGY

A. Classification of Type of Learners' Comprehension

The proposed system consists of three functions:

- 1) execution of online test using categorized questions
- 2) measurement of four learner's academic skills
- 3) classification of learner's comprehension type to recommend suitable learning contents for the learner.

We defined six types of learners' comprehension based on their academic skills, namely, Master, Adept, Careless, Neophyte and Others illustrated in Table I. We can estimate the type of learners' comprehension based on their academic skills calculated by the measurement function. In addition, academic skills are measured for labeling the training data. After classification by unsupervised method (SOM), we examine similarity of competitive units.

TABLE I.
CATEGORIZATION OF LEARNER'S COMPREHENSION AND EXPLANATION OF CATEGORY

No	Category	Estimation Rules / Explanation of Category
1	Master	All skills were 0.8 and over, or average of correct answer ratio of all questions is 0.8 and over. Learner who achieved excellent results.
2	Adept	Lowest skill was Thinking among three skills, such as Memory, Judgement, and Thinking. Learner who acquired basic knowledge but was not good at the thinking.
3	Careless	Lowest skill was Judgement among three skills, such as Memory, Judgement, and Thinking. Learner who acquired basic knowledge but was not good at the judgement.
4	Apprentice	All skills were under 0.6. Learner who memorized the meaning of the legal terms.

5	Neophyte	Correct answer ratio of questions of legal terms was under 0.6. Learner who started learning.
6	Others	

A	Question Fill in the following blanks. A compilation that, by reason of the [selection or arrangement] of its [contents], constitutes an intellectual creation, is protected as a work.
	Answers a) selection / data b) arrangement / contents c) selection or arrangement / contents d) arrangement / data
B	Question Compare the following descriptions that relates to the case where you received a warning of infringing someone's patent right, then choose the most appropriate one.
	Answers a) Where the patentee doesn't show a patent registration that an examiner of the Patent Office prepared, the patentee cannot exercise the patent rights. b) Where a patented invention has not been used for three consecutive years or longer by the patentee, any person may file a request for a trial for rescission of such patent. c) Where a request for a trial for patent invalidation on the ground of the joint application violation may be filed by only the person who has the right to obtain a patent.

Fig. 1. Sample of Two Typical Learners' Mistakes

TABLE II.
SUMMARY OF ONLINE COURSE

Explanation	
Course Name	Local contents and Intellectual Property Management
Learning Style	on-demand e-Learning
No. of Learners	199
No. of Questions	30
	Terms: 3, Purposes: 6, Requirements: 4, Applications: 8, Proceedings: 7, Descriptions: 2

A. Type of Learners' Comprehension

Fig. 1 illustrates two typical learners' mistakes, Sample A is a mistake when they memorize the important words. The question is formed in Fill-in-Blanks. The average of correct answer ratio of this question is about 90.7%. The average for

each choice is follows. Most of the learners answered correctly. The choice of b) shows that about 6% of the learner remembered the keyword "contents", but didn't remember "selection".

Sample B shows the mistakes when the learners apply basic knowledge to solve a case problem. The average of correct answer ratio of the case problems is lower than that one of definition of legal terms, because they are more difficult. In this sample, both a) and b) are confusing Questions to confirm whether the learners understood the legal knowledge precisely or not. The average for each choice is distributed equally. Maybe, it was a difficult question for the learners.

The proposed system was implemented in an introductory course of on-demand e-Learning at five universities in Japan. The summary of the online course is illustrated in Table II. The number of learners is 199 and that of questions is 30 [3], [7]. Questions on "Legal Terms" are useful to memorize definition of legal terms at the first learning step. Questions on "Purposes" and "Requirements" ask knowledge whether the learner can use the important key phrases and enumerate the important requirements exactly. We made similar choices in the phase of making questions in order to ask whether the learner would be able to memorize or judge correct answers.

TABLE III.
SIMILARITY s_{ij} OF EACH TYPES OF LEARNERS' COMPREHENSION

Classification Results of Type j	Training Data of Type i					
	1	2	3	4	5	6
1	0.96	0.03	0.07	0.00	0.00	0.00
2	0.02	0.91	0.27	0.09	0.00	0.00
3	0.01	0.05	0.53	0.09	0.00	0.00
4	0.00	0.01	0.13	0.64	0.22	0.00
5	0.00	0.00	0.00	0.18	0.78	0.00
6	0.00	0.00	0.00	0.00	0.00	1.00

B. Classification of Learners by SOM

Table III shows the calculated similarity of each types. We count the number of learners in each units of competitive layer of the SOM. The similarity s_{ij} that illustrated in Table III is calculated by equation (1). n_{ij} means a training data of type i is mapped on a unit that a training data of type j is mapped, too. S_i that is illustrated in Table III means the total number of the training data categorized in type i .

$$s_{ij} = \frac{n_{ij}}{S_i} \tag{1}$$

For example, in the case of a training data of type 1 are mapped on a unit that a training data of type 2 is mapped, too, the similarity s_{12} is calculated by equation (2).

$$s_{12} = \frac{n_{12}}{S_1} = \frac{1}{81} = 0.02 \tag{2}$$

In addition, the Neophyte that couldn't get a good score on the online test, is classified by the score of the questions of legal

terms without considering their academic skills. Therefore, the Neophyte distributed in pieces. Taken together, we consider that the classification results well represent the relationship of each type.

C. Re-Classification Rules

In the recent study, we found that unsupervised classification method is useful to rough classification, but knowledge-based

TABLE IV.
RE-CLASSIFICATION RULES

IF (Mixed Patterns) ----->	THEN (Revised Category)
Master and Others	Master
Adept and Others (Exclude Master)	Adept
Careless and Apprentice	Apprentice
Apprentice and Neophyte	Apprentice

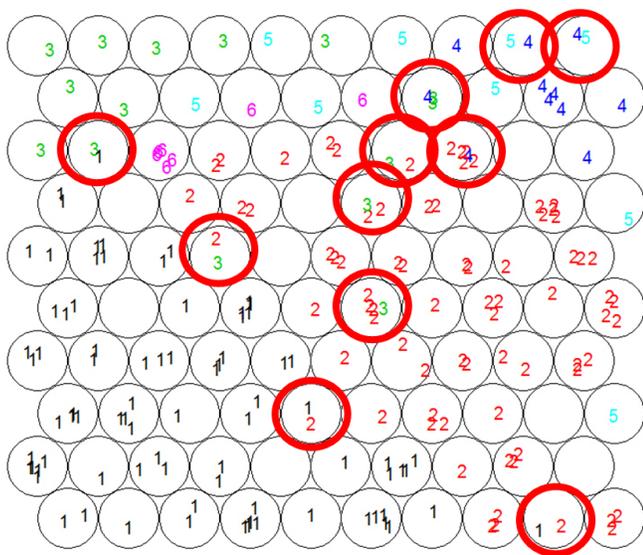


Fig. 2. Distribution of Learners' Comprehension Type Classified by SOM

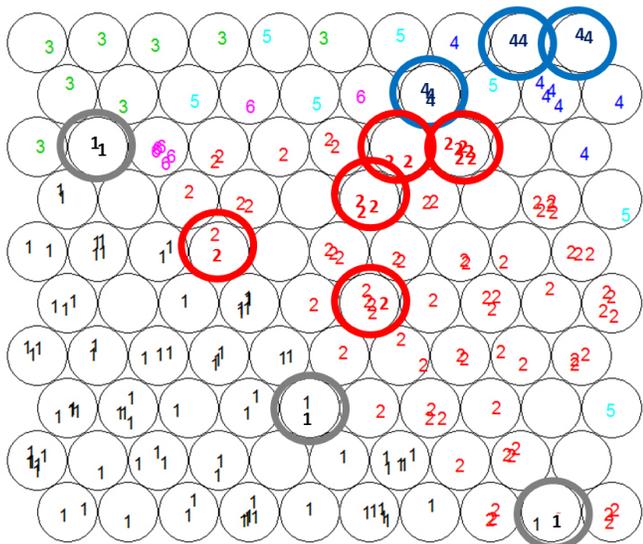


Fig. 3. Distribution of Learners' Comprehension Type Revised by Knowledge-Based Processing

processing easily revises misclassification [7]. Therefore, we introduce the teaching knowledge to revise the misclassification. The implementation of the rules is described below:

We describe the re-classification knowledge in IF THEN rules illustrated in Table IV. At first, learners who took high scores on average for all skills would be classified as a category Master. Since misclassification occurred in the category Adept or Careless as illustrated in Table III, we revise their category to Master. Next, the other learners whose lowest skill was Thinking nearly equals to Judgement would be classified as a category Adept. Therefore, we revise them to Adept. Finally, we revise the other misclassification with Apprentice to Apprentice to provide individual guidance conveniently.

III. RESULTS AND DISCUSSIONS

The experiment was proceeded to the learners in an introductory course of on-demand e-Learning at five universities in Japan. Fig. 2 shows the classification results of learners by SOM classifier. The training data of learners is distributed on the competitive map, named as numeric codes illustrated in Table I.

We employed one hundred units on the competitive layer. We used a "kohonen" package for R language. R is a free software environment for statistical computing and graphics. The units that represent the type of Neophyte are classified in an opposite corner to that of Master, and the type of Apprentice and Careless, too. The Adept is classified nearby the Master. We can read the relationship between the types of learners' comprehension from the map.

We can utilize the classification results to support the learning. For example, the Neophyte that is plotted in the cluster of the Adept, has a characteristic of not good at Thinking. We think it is suitable to recommend the learner learning contents that enhance the Thinking skill.

Fig. 3 and Table V show the revised distribution of learners' comprehension. The all misclassification illustrated in Table IV has been revised, so that we can provide the individual guidance relative to the type of learners' comprehension conveniently.

TABLE V.
REVISED DATA

Classification Results of Type j	Training Data of Type i					
	1	2	3	4	5	6
1	1.00	0.00	0.00	0.00	0.00	0.00
2	0.00	1.00	0.00	0.00	0.00	0.00
3	0.00	0.00	1.00	0.00	0.00	0.00
4	0.00	0.00	0.00	1.00	0.00	0.00
5	0.00	0.00	0.00	0.00	1.00	0.00
6	0.00	0.00	0.00	0.00	0.00	1.00

IV. CONCLUSION

In this paper, we apply the knowledge-based method for revising misclassification on learners' comprehension and

illustrate that we can revise easily because of introducing the teaching knowledge. Since the proposed method produces compact distribution of learners' comprehension, the teacher would make the individual guidance easily. In the future work, we will introduce the individual guidance to the proposed system.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 15K00483.

REFERENCES

- [1] H. Murai and H. Okumura, "Assistance of Understanding by e-Learning," *Journal of the educational application of information*, vol. 7, no. 1, pp. 31-35, 2004. [published in Japanese]
- [2] H. Murai, "Intellectual Property Learning System Supported by SOM," *IEICE Technical Report of NLP*, vol. 105, no. 547, pp. 113-116, 2006. [published in Japanese]
- [3] H. Murai, T. Hayashi, R. Yaegashi, K. Fujimoto and N. Gotoda, "A Study of Classification of Quizzes on Intellectual-Property Education", *Int'l Journal of Computing, Communications & Instrumentation Eng. (IJCCIE)*, vol. 3, no. 1, pp. 123-124, 2016. <http://dx.doi.org/10.15242/IJCCIE.AE01160012>
- [4] H. Murai, K. Fujimoto, T. Hayashi, N. Gotoda, R. Yaegashi, "Intellectual-Property Learning System Supported by SOM Classifier", *Proc. Int. Conf. on Artificial Intelligence Science and Technology*, pp.113-118, 2016.
- [5] H. Murai, "Similarity of Learners' Understanding Type Classified by SOM on Intellectual-Property Education", *Proc. 2016 Int. Conf. on Recent Development in Engineering and Technology (RDET-16)*, pp.61-62, 2016.
- [6] H. Murai and S. Omatu, "Remote Sensing Image Analysis using a Neural Network and Knowledge-Based Processing", *International Journal of Remote Sensing*, No. 18, pp.811-828, 1997.
- [7] H. Murai, T. Hayashi, R. Yaegashi, A. Iwaki, and K. Ura, "Development of Gaming App to Support Memorizing Keywords in Law Learning," *Technical Report of Japanese Society for Information and Systems in Education*, vol. 29, no. 3, pp. 3-6, 2014. [published in Japanese]