Consulting System of Undergraduate Course Guidance

Masahiro KURODA*, Muneyoshi KIMURA*, Shigekazu NAKAGAWA**, Shinji KADOWAKI**, Hiroshi KIMURA** and Minoru ICHIMURA**

**Graduate School of Science,*
Okayama University of Science,
Okayama 700, Japan.
***Department of Applied Mathematics,*
Okayama University of Science,
Okayama 700, Japan.

(Received September 30, 1993)

Abstract

This paper mainly presents the reasoning method in our consulting system in order to predict how many credits the students can get. It is necessary to use a prior knowledge obtained by the grade data of the students. So we adopt the probabilistic logic model due to Niki[7] in this sense. Numerical results for actual grade data in our university are shown.

1. Introduction

It seems that there exist students who repeat a year by which choosing courses are not appropriate in many universities. If academic advisors give proper guidance to the students, then some of them can graduate smoothly for four years. To such students they want to predict the number of courses which are given credits and further advise to get the number of necessary credits or courses which each university stipulates.

We construct a consulting system that accomplishes their purpose, namely, it predicts the number of courses which are given credits on course cards that they turn in and advises them which courses they take.

The approaches for reasoning under uncertainty, Bayes rule[1], the certainty factor method[2] used in the system MYCIN, subjective Bayesian method[3] used in the system Prospector, Dempster-Shapfer theory[4], probabilistic logic[5], fuzzy theory[6] and so on, are well known. In particular we adopt probabilisitic logic from a viewpoint of making use of prior knowledge in our problem.

In this paper, we apply probabilistic logic due to Niki[7] based on the maximum entropy principle in the sense of Kullback-Leibler[8] to the prediction of the number of courses which are given credits. For 81 students who matriculate in 1987 in university, we predict the number of courses which are given credits and compare them with their actual grade data.

2. Model

A certain student turns in course cards at the beginning of the first semester or the second semester. It is our problem that we predict, on these cards, how many credits he can get. These cards also lead us to two kind of prior knowledge as follows:

1. prior knowledge about the trend that each teacher gives credits,
2. prior knowledge about his past grade.

The trend means whether the degree that each teacher gives credits is high or not.

Firstly we classify 54 teachers in our university into five groups from a viewpoint of the degrees that each of them gives credits. Cluster analysis based on further neighbor method with Mahalanobis distance is also applied to this classification and its results are shown in Table 1. If there are n teachers in group i and each of the degrees which these teachers in this group give credits is x_j $(j = 1,\cdots,n)$, then the degree giving credits is $\sum_{j=1}^{n} x_j/n$, and the ratios are each ratio of the number of teachers in group i to 54 teachers in our university.

Secondly his past grades given by the teachers in each group are known. Thus we look up the number of courses given by the teachers in each group and count up the number of courses which are given credits in every group.

Prior knowledge about his grade is updated with that he moves up to the next grade. Therefore we want to use this knowledge for prediction at the next grade, positively. From making use of the above prior knowledge we apply Niki's probabilistic logic model to our problem or to obtain the information how many credits a certain student can get.

3. Probabilistic logic

We define propositions X, Y and Z as follows:

$X \equiv$ "A certain student takes the course given by teachers in group $i(i = 1,\cdots,5)$",
$Y \equiv$ (A teacher in group i gives credits",
$Z \equiv$ "A certain student can get credits".

Under the circumstances C that a certain student turns in course cards, we estimate $u_2 = \Pr(Z|C)$ when $u_0 = \Pr(X|C)$ and $u_1 = \Pr(Y|C)$ are known. For example, if a certain student takes ten courses and the number of courses given by teachers in group 1 is two, then u_0 is $2/10 = 0.2$ and u_1 is 0.9127 from Table 1.

Table 1 The degree giving credits and the ratio of teachers in each group

	The degree giving credits	the ratio of teachers
The teachers in group 1	0.9127	0.13
The teachers in group 2	0.4406	0.11
The teachers in group 3	0.8685	0.34
The teachers in group 4	0.6903	0.14
The teachers in group 5	0.8758	0.28

(On 54 teachers in our university)

Table 2 Probabilistic logic model

	M_0	M_1	M_2	M_3	M_4	M_5	M_6	M_7	
X	0	1	0	1	0	1	0	1	u_0
Y	0	0	1	1	0	0	1	1	u_1
Z	0	0	0	0	1	1	1	1	u_2
Π	π_0	π_1	π_2	0	0	0	0	π_7	
P	p_0	p_1	p_3	0	0	0	0	p_7	

Table 2 represents all possible truth values (*true* is 1 and *false* is 0) for $^t(X\ Y\ Z)$. $\{M_0, \cdots, M_7\}$ are the sets of events corresponding to different vectors of values. A discrete distribution $P = \{p_0, \cdots, p_7\}$ is each component of which represents the probability that the actual event is a member of M_i and given by

$$p_i = \Pr\{M_i\ is\ true|C\} \qquad (i = 0, \cdots, 7) \tag{1}$$

and intervening knowledge is a discrete distribution Π defined on the set of M_i with the probabilities

$$\pi_i = \Pr\{M_i\ is\ true\} \qquad (i = 0, \cdots, 7). \tag{2}$$

This distribution Π approximates, as prior knowledge about P, to that by P. In our problem we may regard knowledge about his grade in the past as Π. In Table 2, the semanticklly consistent sets of truth values and corresponding sets of events are clearly $\{M_0\ M_1\ M_2\ M_7\}$. The set of event M_7, for example, can be interpreted into

$M_7 =$ "A certain student takes the course given by a teacher in group i and this teacher also gives credits of the course, so he can get this credits".

Then the distribution P is satisfied with

$$p_3 = p_4 = p_5 = p_6 = 0, \qquad p_0 + p_1 + p_2 + p_7 + 1. \tag{3}$$

In order to estimate u_2 represented by p_i, the maximum entropy estimate $\tilde{P}$ of P can be obtained by minimizing the Kullback-Leibler information

$$I(P;\Pi) = \sum_{i=0}^{7}(p_i \log p_i - p_i \log \pi_i) \tag{4}$$

for discrimination of P from Π, subject to the constraint (3), where we assume $0\log 0 = 0$.

4. Numerical results

This reasoning method is applied to a certain student who matriculate in 1987 and predicts the number of courses that he is given credits, on the course cards turned in by him at the first semester of third-grade for every group. In the case of this prediction, we make use of his grade from first-grade till second-grade. In Table 3 and Table 4, A means the number of courses which he takes, B means the number of

Table 3 Example 1

	A	B	C
Group 1	1	1	1.000
Group 2	2	2	0.461
Group 3	9	8	8.466
Group 4	1	1	0.790
Group 5	2	2	2.000
Total	15	14	12.717

Table 4 Example 1

	A	B	C
Group 1	0	0	0.000
Group 2	0	0	0.000
Group 3	5	4	4.796
Group 4	2	0	1.705
Group 5	4	2	3.744
Total	12	6	10.245

Table 5 The errors between grade data and the our predicting results

#	1	2	3	4	5	total
the number of students	59	12	7	2	1	81
ratio	0.72	0.15	0.09	0.03	0.01	1.00

(on 81 students who matriculate in 1987)

courses which he is given credits, and C means the product of u_2 and A or our predicting value. Table 4 represents that C does not fit B, when the ratio of B to A is low.

Next we apply to 81 students who maticulate in 1987 and predict the number of courses which each student is given credits at the first semester of third-grade. For each student, we use

$$\# = \text{int} \max_{1 \leq i \leq 5} |B_i - C_i|, \tag{5}$$

in order to look up the errors between B_i and C_i, where i is group i, B_i is the number of courses that he get and C_i is the product of u_2 and the number of courses which he takes, and its results are shown in Table 5.

If this error (= #) is one or two courses and it is valid as the predicting result to our problem, then 87 % students are included among this error.

5. Conclusion

We construct the model that predicts the number of courses given credits by applying probabilistic logic as the reasoning method on prior knowledge, that is, knowledge about the trend that each teacher gives credits and his past grade.

Our purpose is accomplished for 87% students, when we predict, for 81 students who matriculate in 1987 at the first semester of third-grade, the number of courses which each of them gets.

Furthermore we need to work the approaches to the case as Table 4 in order to get better predicting results. As one of the approaches to solve this problem we consider modifying prior knowledge with increasing grade data.

REFERENCES

1) Press, S. James (1989). *Bayesian statistics: principles, models, and applications,* John Wiley & Sons.
2) Buchman, B. G. and Shortliffe, E. H. (1984). *Rule-Based Expert Systems-The MYCIN Experiments of the Stanford Heuristic Programming Project,* Addison-Wesley.
3) Duda, R. O., Hart, P. and Nilsson, N. J (1976). *Subjective Bayesian methods for rule-based inference systems,* NCC.
4) Shafer, G. (1976). *A Mathematical Theory of Evidence,* Princeton Univ. Press
5) Nilson, N. J. (1986). Probabilistic logic, *Artificial Intelligence,* **28**, pp. 71-87.
6) Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Trans. Syst., Man & Cybern.,* **SMC-2**, *1,* pp. 28-44.
7) Nili, N. (1989) Maximum entropy principle in probabilistic logic, *Proc. 3rd Japan-China Symp. Statist.,* pp. 283-287.
8) Kullback, S. (1959). *Information Theory and Statistics,* John Wiley, London.