

클러스터링에 기반한 학업성적의 등급화 방법

A Grading Method for Student's Achievements Based on the Clustering Technique

박은진 · 정 흥 · 장덕성

Eun-Jin Park, Hong Chung and Duk-Sung Jang

계명대학교 공학부

요 약

학업성적 평가방법에는 일반적으로 절대평가와 상대평가가 있다. 이들 방법은 각기 장점들을 가지고 있으나 평가기준이 너무 일률적이거나 아니면 학습자간에 지나친 경쟁을 유발하는 등 일부 한계점들을 내포하고 있다. 본 논문에서는 일률적인 기준이 아닌 학습자들의 점수분포와 발생빈도를 감안한 등급화를 통해 학업성적을 평가하는 방법을 제안한다. 즉, 전체 학습자의 학업성적을 적합도(goodness)를 감안하여 몇 개의 클래스로 분할하고 이를 등급으로 하는 것이다. 이를 위해 본 논문은 RE(Relaxation Error)를 적용하여 적합도를 구하고 이를 바탕으로 성적을 등급화 한다. 제안된 방법은 학업성적의 분포와 발생빈도를 고려한 적합도 추정에 의해 등급기준을 정하고 이에 따라 학업성적을 등급화하므로 근소한 차로 등급이 구분되는 사례가 거의 없게 되며, 학교간 학생들의 학업 성취도를 차별화 할 수 있고, 실제적인 우수자나 부진자 선발 및 수준별 이동식 수업을 위한 평가에 유용하므로 기존의 두 방법을 보완하는 특성을 갖고 있다.

Abstract

There are two methods in evaluating student's achievement. The two evaluation methods are absolute evaluation and relative evaluation. They have much advantages respectively, but also have some limitations such as being too stereotyped or causing overcompetition among learners.

This paper suggests a new evaluation method which evaluates student's achievements by considering the score distribution and the frequency. The proposed method classifies the scores into several clusters considering the goodness. This approach calculates the goodness by applying the RE(relaxation error), and grades the achievement scores based on the goodness.

The suggested method can avoid the problem of grading caused by the narrow gap of scores because it sets a standard for grading by the calculated goodness considering the score distribution and frequency of occurrence. The method can differentiate achievements of a school from those of others, and that it is useful for selecting advanced students and dull ones, and for evaluation of classes based on student's achievement.

Key words : Evaluation, Clustering, Grading, Relaxation Error

1. 서 론

오늘날 학교에서 교과목 성적을 평가하는 유형에는 두 가지가 있다. 준거지향평가(準據指向評價)라고도 하는 절대평가와 기준지향평가(基準指向評價)라고도 하는 상대평가가 그것이다[1, 2, 3].

절대평가는 개인이 어떤 준거에 비추어 어느 위치에 있는지를 밝히는 평가로 설정된 목표에 어느 정도 도달하였는지를 평가하므로 학습자 전원이 설정된 목표에 도달하였으면 전원에게 만점을 주는 것이다. 이것은 교육의 수월성을 만족시키고 학습자 개인의 목표 도달정도를 알 수 있어 개별학습이 가능하지만 성취도를 가지고 사 람을 선발하고자 할 경우 제도적 형평성을 만족시키지

못한다는 비판을 받고 있다. 예를 들어 8명의 성적 {91, 90, 89, 83, 80, 78, 72, 70}이 있다고 할 때, 이를 학업 성취도에 따라 A, B, C로 3등분하려고 한다. 평가기준이 90이상 A, 80이상 B, 70이상 C라고 할 때, {A=(91, 90), B=(89, 83, 80), C=(78, 72, 70)}으로 3등분된다. 여기서 89가 B, 78이 C로 분류되는 것은 형평성에 있어서 합리적 이지 못하다.

반대로 상대평가는 어떤 기준에 비추어 본 개인의 상대적 위치를 평가하여 석차를 매기는 방식이다. 현재 고등학교 3학년까지의 평가방법으로 대학입학전형 자료로 제출하는 내신성적이 이에 해당된다. 이것은 학교별 수준차를 인정한 평가 방법으로 제도적 형평성은 만족시키 지만, 지나친 경쟁을 유발하여 인성교육의 저해를 가져 오고 학습자의 학습목표 달성정도를 알 수 없다는 비판 을 받고 있다. 위의 예에서 평가기준이 A 25%, B 50%, C 25%라고 할 때, {A=(91, 90), B=(89, 83, 80, 78), C=(72, 70)}으로 3등분된다. 여기서도 89가 B로 분류되는

접수일자 : 2001년 8월 29일
완료일자 : 2002년 2월 22일

것은 적합하다고 볼 수 없다. 이를 성취도에 따라 경험적으로 분류한다면, 성적 분포를 감안하여 {A=(91, 90, 89), B=(83, 80, 79), C=(72, 70)}으로 3등분하는 것이 가장 바람직하다.

본 논문에서는 이와 같은 일률적인 기준이 아닌 점수의 분포와 발생빈도를 고려한 등급화를 통해 학업의 성취도를 평가하고자 한다. 즉, 전체 학습자의 학업성적에 대해 적합도(goodness)를 감안하여 원하는 몇 개의 클래스로 분할하고 이를 등급화 하는 것이다. 이를 위해 본 논문은 클러스터링의 한 방법인 RE(relaxation error) 기법[5, 6, 7]을 적용하여 적합도를 구하고 이를 바탕으로 등급화 하고자 한다.

2. 등급화를 위한 클러스터링 방법

자료를 클러스터링 하는 방법에는 통계적 클러스터링과 개념적 클러스터링이 있다. 전자는 자료간 유사도(similarity)를 사용하여 자료를 클러스터링 하는 방법이고, 후자는 자료가 속한 분할의 적합도 측정(goodness-measure)을 통하여 클러스터링 하게 되는데, 후자가 좀더 적합한 자료의 클러스터링을 유도할 수 있다.

본 논문에서는 개념적 클러스터링의 적합도 측정방법인 RE에 대해서 기술한다[5, 6, 7]. 그리고, 클러스터링 과정의 등급 분할(N-ary partition)에 있어서 시간 복잡도(computation complexity)를 줄이기 위해 경험적인(heuristic) 방법인 이진분할(binary cut) 알고리즘[5, 7]을 사용한다.

클래스 C를 N개의 배타적인 클래스 C_1, \dots, C_N 으로 분할한 경우, 일반적으로 개념적 클러스터링에는 CU(category utility)를 사용하나 이는 비수치 자료인 범주(categorical) 자료만을 다룬다. 그러므로, 자료값에는 관심이 없고 단지 자료의 그룹화에만 집중하게 된다. 즉 클러스터링시 자료의 발생 빈도만을 고려하고 자료의 분포는 고려하지 않게 된다. 이런 경우 예를 들어 두 클래스 $C_1=(0, 1)$ 과 $C_2=(0, 100)$ 을 동일하게 취급하게 되는데 이는 클러스터링 측면에서 볼 때 적절하다고 볼 수 없다.

클러스터링에서 수치자료를 다루기 위해 본 논문에서는 RE를 사용한다. RE는 클래스 C의 자료간의 절대적 차이 즉, 자료의 분포와 자료의 발생빈도를 동시에 고려한 것으로 식(1)과 같이 정의된다¹⁾.

$$RE(C) = \sum_{i=1}^N \sum_{j=1}^N P(x_i)P(x_j)|x_i - x_j| \quad (1)$$

x_i, x_j : 속성값

$P(x_i), P(x_j)$: 전체(C)에서 x_i, x_j 의 발생 확률

RE(C)는 클래스 C에서 클러스터링 시 자료 값간의 차이에 의해 반영하는 에러(error)로서 이 값이 크면 클러스터링의 적합도는 떨어진다.

C의 분할이 $P=\{C_1, C_2, \dots, C_N\}$ 일 때, 클래스 C의 감축 에러인 RE(C)보다 각 서브클러스터(sub-cluster)의 감축 에러의 합인 RE(P)가 값이 작아지므로 이를 이용하면

최적 분할을 판단할 수 있다. C의 분할이 P일 때, RE(P)는 식(2)와 같다.

$$RE(P) = \sum_{k=1}^N P(C_k)RE(C_k) \quad (2)$$

n개의 자료 값을 가진 클러스터 C가 있을 때, 분할의 수는 n에 지수적으로 증가하므로 그중 가장 좋은 분할을 발견하는데 걸리는 시간도 지수적이다. 따라서 시간 복잡도를 줄이기 위해 이진분할을 사용한다. 이진분할 알고리즘은 다음과 같다.

Binary_Partition(C)

$C_{copy} = C;$ /* $C = \{x_1, \dots, x_n\}$ */

for i = 1 to Grade - 1 /* Grade: 등급 수 */

/* 분할지점 선택 */

for h = first to n - 1

/* $C_{copy} = \{x_{first}, \dots, x_n\}$ 일 때 각 분할지점 계산 :

C_{copy} 를 C_1 와 C_2 로 나누기 위한 best cut을 찾기 위해 각 요소에 대한 RE(P) 계산*/

Let P is partition with clusters $C_1 = \{x_1, \dots, x_h\}$ and

$C_2 = \{x_{h+1}, \dots, x_n\};$

compute RE(P);

/* RE(P)가 가장 작은 partition지점을 best cut으로 선택 */

if RE(P) < MinRE then

MinRE = RE(P);

cut = h; /* the best cut */

endif;

endfor;

let C_1 and C_2 be two sub-cluster of C_{copy} ;

/* C_i 선택 : C의 sub-cluster중 RE(C_i)가 가장 큰 C_i 를 선택*/

let C_i be the sub-cluster of C with maximum RE(C_i);

$C_{copy} = C_i;$

endfor;

즉, 클러스터 $C = \{x_1, x_2, \dots, x_n\}$ 를 서브클러스터 $C_1 = \{x_1, \dots, x_h\}$ 와 $C_2 = \{x_{h+1}, \dots, x_n\}$ 로 이진분할 하여 서브클러스터 C_i 중 RE(C)가 가장 큰 C_i 를 선택하고, 선택된 서브클러스터에 대해 다시 이진분할을 적용하는 것을 반복하여 사용자가 분할하고자 하는 수만큼 클러스터 C를 분할하게 된다. 결과적으로 RE(C)의 계산에 $O(n)$ 이 걸리고 이러한 계산이 (분할수 - 1)회 반복되므로 이진분할의 시간 복잡도는 $O(n^2)$ 이 된다.

3. 학업성적의 클러스터링

학업성적 클러스터 C의 등급화를 위한 분할지점 선택을 위해 C의 자료 각각을 분할지점으로 가정하여 각 분할에 대한 RE(P)를 계산한다. 이 중에서 RE(P)값이 가장 작은 경우의 자료를 분할지점으로 선택한 후, 이

1) 한 항목의 분류에 사용할 때는 RE_i로 표시해야 하나 편의를 위해 RE로 표현한다.

분할지점을 기준으로 분할된 서브클러스터 중에서 RE(C)가 가장 큰 서브클러스터를 재분할할 클러스터로 정하는 과정을 반복한다.

예를 들어 학생 10명의 성적이 {60, 46, 88, 72, 100, 92, 88, 32, 60, 40}이고 상·중·하 세 등급으로 구분하고자 할 때, 첫 번째 분할지점을 구하는 과정은 다음과 같다.

1) 성적 정렬

C(성적, 발생빈도) = {(100, 1), (92, 1), (88, 2), (72, 1), (60, 2), (46, 1), (40, 1), (32, 1)}

2) 첫 번째 분할지점 계산

정렬된 성적에 대한 이진분할은 7가지의 경우가 있는데, 각각 경우에 대한 RE(P)는 다음과 같이 계산된다.

$P_1 = \{C_1, C_2\} = \{\{100\}, \{92, 88, 72, 60, 46, 40, 32\}\}$ 인 경우, C_1 은 하나의 값만 있으므로 $RE(C_1) = 0$ 이다. C_2 는 7개의 값을 가지고 있으므로 식(1)에 의해 다음과 같이 RE를 계산한다.

$$RE(92) = (1/9) \times (|92-92|) + (2/9) \times (|92-88|) + (1/9) \times (|92-72|) + (2/9) \times (|92-60|) + (1/9) \times (|92-46|) + (1/9) \times (|92-40|) + (1/9) \times (|92-32|) = 27.28$$

$$RE(88) = (1/9) \times (|88-92|) + (2/9) \times (|88-88|) + (1/9) \times (|88-72|) + (2/9) \times (|88-60|) + (1/9) \times (|88-46|) + (1/9) \times (|88-40|) + (1/9) \times (|88-32|) = 24.67$$

⋮

$$RE(32) = (1/9) \times (|32-92|) + (2/9) \times (|32-88|) + (1/9) \times (|32-72|) + (2/9) \times (|32-60|) + (1/9) \times (|32-46|) + (1/9) \times (|32-40|) + (1/9) \times (|32-32|) = 32.33$$

따라서,

$$RE(C_2) = (1/9) \times 27.28 + (2/9) \times 24.67 + (1/9) \times 19.33 + (2/9) \times 18 + (2/9) \times 22.67 + (1/9) \times 26.44 + (1/9) \times 32.33 = 23.71$$

그리고, 식(2)에 의해,

$$RE(P_1) = 1/10 \times RE(C_1) + 9/10 \times RE(C_2) = 0 + 9/10 \times 23.71 = 21.34$$

같은 방법으로 나머지 분할에 대하여 계산하면 다음과 같다.

$$P_2 = \{C_1, C_2\} = \{\{100, 92\}, \{88, 72, 60, 46, 40, 32\}\}$$
인 경우: $RE(P_2) = 20.97$

$$P_3 = \{C_1, C_2\} = \{\{100, 92, 88\}, \{72, 60, 46, 40, 32\}\}$$
인 경우: $RE(P_3) = 17.58$

$$P_4 = \{C_1, C_2\} = \{\{100, 92, 88, 72\}, \{60, 46, 40, 32\}\}$$
인 경우: $RE(P_4) = 11.33$

$$P_5 = \{C_1, C_2\} = \{\{100, 92, 88, 72, 60\}, \{46, 40, 32\}\}$$
인 경우: $RE(P_5) = 10.62$

$$P_6 = \{C_1, C_2\} = \{\{100, 92, 88, 72, 60, 46\}, \{40, 32\}\}$$
인 경우: $RE(P_6) = 16.8$

$$P_7 = \{C_1, C_2\} = \{\{100, 92, 88, 72, 60, 46, 49\}, \{32\}\}$$
인 경우: $RE(P_7) = 20.78$

P_1 에서 P_7 중 RE(P)가 가장 작은 지점이 P_5 이므로 분할 지점은 60이 되고, 클러스터 C는 서브클러스터 $C_1 = \{100, 92, 88, 72, 60\}$ 과 $C_2 = \{46, 40, 32\}$ 로 분할된다.

3) 두 번째 분할지점 계산

서브클러스터 C_1 과 C_2 중 재분할할 클러스터를 선택하는 과정은 다음과 같다.

$RE(C_1)$

$$= [1/7 \times \{|100-100| + 1/7 \times |92-100| + 2/7 \times |88-100| + 1/7 \times |72-100| + 2/7 \times |60-100|\} + 1/7 \times \{|100-92| + 1/7 \times |92-92| + 2/7 \times |88-92| + 1/7 \times |72-92| + 2/7 \times |60-92|\} + 2/7 \times \{|100-88| + 1/7 \times |92-88| + 2/7 \times |88-88| + 1/7 \times |72-88| + 2/7 \times |60-88|\} + 1/7 \times \{|100-72| + 1/7 \times |92-72| + 2/7 \times |88-72| + 1/7 \times |72-72| + 2/7 \times |60-72|\} + 2/7 \times \{|100-60| + 1/7 \times |92-60| + 2/7 \times |88-60| + 1/7 \times |72-60| + 2/7 \times |60-60|\}] = 16$$

$RE(C_2)$

$$= [1/3 \times \{|46-46| + 1/3 \times |40-46| + 1/3 \times |32-46|\} + 1/3 \times \{|46-40| + 1/3 \times |40-40| + 1/3 \times |32-40|\} + 1/3 \times \{|46-32| + 1/3 \times |40-32| + 1/3 \times |32-32|\}] = 6.22$$

서브클러스터 C_1 과 C_2 중 RE가 큰 서브클러스터가 C_1 이므로 재분할할 클러스터는 C_1 이 된다.

두 번째 분할지점 선택을 위해 클러스터 $C_1 = \{100, 92, 88, 72, 60\}$ 에 대한 이진분할은 P_{11} 에서 P_{14} 까지 4가지 경우가 있으며, 각 분할에 대한 RE(P)를 계산하면 다음과 같다.

$$P_{11} = \{\{100\}, \{92, 88, 72, 60\}\}$$
인 경우: $RE(P_{11}) = 12.95$

$$P_{12} = \{\{100, 92\}, \{88, 72, 60\}\}$$
인 경우: $RE(P_{12}) = 10.94$

$$P_{13} = \{\{100, 92, 88\}, \{72, 60\}\}$$
인 경우: $RE(P_{13}) = 5.33$

$$P_{14} = \{\{100, 92, 88, 72\}, \{60\}\}$$
인 경우: $RE(P_{14}) = 6.63$

P_{11} 에서 P_{14} 중 RE가 가장 작은 지점이 P_{13} 이므로 분할지점은 88이 된다. 따라서 클러스터 C_1 은 서브클러스터 $C_{11} = \{100, 92, 88\}$, $C_{12} = \{72, 60\}$ 로 분할된다.

위와 같은 과정을 거쳐 성적 = {60, 46, 88, 72, 100, 92, 88, 32, 60, 40}은 성적의 분포와 빈도를 고려한 상 = {100, 92, 88}, 중 = {72, 60}, 하 = {46, 40, 32}라는 세 등급으로 구분된다. 만약 4등급으로 나누려고 한다면 상, 중, 하중 가장 큰 RE를 가지는 클러스터를 택하여 상기와 같은 방법으로 분할한다.

4. 실험 및 평가

4.1 실험

클러스터링을 위한 이진분할 알고리즘을 구현하여 여러 가지 성적 데이터로 실험한다. 이 때 클러스터링을

통해 생성된 등급 결과 뿐 아니라 동일한 자료에 대해 상대평가와 절대평가를 적용했을 경우 생성될 등급 결과를 함께 제시한다. 그림 1은 40명의 성적이 {90, 90, 90, 90, 90, 85, 85, 85, 85, 80, 80, 80, 80, 70, 70, 70, 70, 70, 65, 65, 65, 65, 60, 60, 60, 60, 60, 60, 55, 55, 55, 50, 50, 50, 35, 35, 32, 20}인 경우 5등급으로 분류한 결과 출력 화면이다. 괄호 안의 숫자는 해당 점수의 발생빈도를 의미한다.

상대평가는 학생들이 받은 성적의 평균 의거하여 미리 정해 놓은 비율에 따라 분할되어 있고, 절대평가는 점수대에 따라 일률적으로 등급이 분할되어 있음에 비해, 클러스터링 방법은 유사한 점수대에 따라 분할되어 있음을 볼 수 있다.

1	90(5)	85(4)	80(4)						
2	70(6)								
3	65(4)	60(6)	55(3)	50(4)					
4	35(2)	32(1)							
5	20(1)								

그림 1. 성적의 등급화
Fig. 1. Grading of score

그림 2는 50명의 성적이 {100, 97, 95, 95, 93, 93, 91, 89, 87, 87, 85, 85, 83, 83, 81, 81, 79, 79, 77, 77, 77, 75, 75, 75, 73, 73, 73, 71, 71, 69, 69, 67, 67, 67, 65, 65, 65, 63, 63, 63, 61, 61, 59, 59, 57, 57, 57, 55, 55, 55}인 비교적 고르게 분포되어 있는 경우 등급화 한 결과 출력이다.

1	100(1)	97(1)	95(2)	93(2)	91(1)	89(1)	87(2)		
2	85(2)	83(2)	81(2)						
3	79(2)	77(3)	75(3)	73(3)	71(2)				
4	69(2)	67(3)	65(3)						
5	63(3)	61(2)	59(2)	57(3)	55(3)				

그림 2. 균일하게 분포된 성적의 등급화
Fig. 2. 균일하게 분포된 성적의 등급화

성적이 고르게 분포된 경우 상대평가와 절대평가 그리고 본 논문의 등급화 방법에 의한 등급 결과가 유사함을 볼 수 있다.

그림 3은 70명의 성적이 {100, 100, 100, 99, 99, 99, 98, 98, 98, 97, 97, 97, 96, 96, 96, 95, 95, 95, 94, 94, 94, 93, 93, 93, 92, 92, 91, 91, 91, 91, 91, 90, 90, 90, 89, 89, 89, 88, 88, 88, 87, 87, 87, 86, 86, 86, 85, 85, 85, 84, 84, 84, 83, 83, 83, 82, 82, 82, 81, 81, 81, 81, 80, 80, 80, 79, 79, 79, 78, 78, 78}인 높은 점수대에 치중하여 분포된 경우 등급화 한 결과 출력 화면이다. 클러스터링 방법에서는 성적 분포와 빈도에 따라 성취도에 따른 최적의 5등급 분류가 되나 상대평가와 절대평가는 일률적 기준에 의하여 4, 5 등급은 발생하지 않는다.

1	100(3)	99(3)	98(3)	97(3)	96(3)	95(3)	94(3)	93(3)	92(2)	91(4)	90(3)
2	89(3)	88(3)	87(3)	86(3)	85(3)	84(3)	83(3)	82(3)	81(4)	80(3)	
3	79(3)	78(3)									
4											
5											

그림 3. 높은 성적으로 분포된 성적의 등급화
Fig. 3. 높은 성적으로 분포된 성적의 등급화

위 3가지 경우 모두 상대평가와 절대평가가 같은 결과가 나온 이유는 본 실험에서 사용한 평가 기준이 현재 고등학교 3학년에 적용되고 있는 상대평가 기준인 교과별 성취도 평정환산표[4]와 절대평가 기준인 목표지형 평가기준[4]을 적용했기 때문이다.

4.2 평가

학습자 중심의 교육, 학습자 개인의 학업성취도를 고려한 개별학습의 중요성이 대두되면서 수준별 이동학습과 부진아의 방과후 교육 등이 시행되고 있다. 이러한 프로그램의 올바른 시행을 위해서 학생들의 학업 성취도를 고려하여 적절하게 등급화 할 평가방법이 요구되고 있다. 현행 학업성적의 평가에 사용되고 있는 유형으로는 절대평가와 상대평가 있다. 절대평가는 준거 즉 설정된 목표에 도달정도를 평가하고, 상대평가는 어떤 기준에 비추어 본 개인의 상대적 위치를 평가하는 것으로 서로 평가 목적이 다르다. 그러나, 점수의 분포와 빈도와는 상관없이 일률적인 기준을 정해두고 학습자의 성취도를 평가하고 있다는 점에서는 동일하다. 다시 말하면 절대평가는 100%-90%는 '수', 89%-80%는 '우'라는 식의 절대적이면서 일률적인 기준을 적용하고, 상대평가는 중등교육의 경우 교과소점(평균점)을 기준으로 상대적이면서도 일률적인 평가기준을 적용하고 있다. 이것은 학습자의 절대적 학업성취정도를 알아보기나 대학입학전형 자료로 사용하기 위한 평가방법으로는 적절하나 수준별 이

동수업을 하기 위해 그룹을 나누고자 할 경우의 평가방법으로 적절하다고 볼 수 없다. 그러므로 점수들의 분포와 빈도를 고려한 클러스터링에 의한 등급화를 통해 학업성취도를 평가함으로써 기존의 평가방법의 단점을 보완할 수 있다.

그리고 영재아 또는 부진아를 선발하여 개별학습을 실시하고자 할 때, 기존에는 '성적 우수자 10명' 또는 '학습 부진아 10명' 등 인위적으로 선발 인원을 선정하는 방법을 사용하고 있는데, 클러스터링을 이용한 등급화를 사용하면 학업성취도가 유사한 집단으로 분할해주므로 실제적인 우수자나 부진아를 정확하게 선발할 수 있다는 장점이 있다.

또한 절대평가의 경우 90점은 '수'가 되지만, 1점 차이인 89점은 '우'가 된다. 예를 들어 성적클래스가 {90, 90, 89, 88, 76, 60, 66, 66, 64, 51, 50, 50}일 때 90점은 '수'가 되고 89점과 88점은 '우'가 되므로 적절한 등급화라고 보기 어렵다. 클러스터링을 이용한 등급화 방법은 등급간에는 점수 차가 많이 나고 같은 등급 내에서는 비슷한 성적들이 분포하게 되어 근소한 차이로 등급이 분할되지 않는다. 그러므로 본 등급화 방법을 성적 클래스에 적용하면 90점, 89점, 88점이 동일 등급으로 분할되므로 이러한 문제점이 해결된다.

5. 결 론

기존의 학업성적 등급화 방법으로 상대평가와 절대평가가 있다. 이 두 방법은 학업성적 평가의 목적이 다르므로 실제적용에 있어 평가 과정과 장·단점 또한 다르다. 그러나 절대평가는 점수대에 따른 절대적 기준을, 상대평가는 점수와는 상관없이 평균이나 백분율을 기준으로 한 일률적인 기준을 적용하고 있다는 면에서는 유사하다. 본 연구에서는 일률적인 기준이 아닌 점수의 분포와 발생빈도를 감안한 평가기준 선정에 의한 등급화를 통해 학업의 성취도를 평가하고자 했다. 즉, 전체 학습자의 학업성적에 대해 적합도를 감안하여 필요한 몇 개의 클래스로 분할하고 이를 등급으로 정했다.

본 논문에서 연구한 등급화 방법은 다음과 같은 점에 의의를 부여할 수 있다.

첫째, 기존의 일률적 평가기준이 아닌 학습자의 성적 분포를 기준으로 하여 학업 성취도를 평가하므로 전체 학생들의 성취도가 높으면 상위 등급이 많을 수도 있고, 반대로 성취도가 낮으면 하위 등급이 많을 수도 있다. 이것은 수준별 이동식 수업이나 영재아나 부진아를 선발하여 개별학습을 하고자 할 경우 유용한 평가방법이다.

둘째, 적합도 측정에 의한 등급을 판단하므로 근소한 차이의 성적으로 등급이 구분되는 사례가 거의 없게 되며, 따라서 등급간에는 점수차가 많이 나게 되고 같은 등급 내에서는 비슷한 성적들이 분포하게 된다.

셋째, 학교간 학생들의 성취도를 차별화 할 수 있다. 즉, 시험문제의 난이도가 공정하게 적용되거나 혹은 학교간 공동 시험문제로 시험을 칠 경우 우수 학생이 많은 학교와 적은 학교에 있어서 등급 분포의 차별화가 이루어 질 수 있다. 따라서 현재 특수 목적고등학교와 같은 문제점 해결에 도움을 줄 수 있을 것이다. 이를 고등학교에서 현재 사용하고 있는 평정환산표를 적용한 상대평가 혹은 평가기준을 적용한 절대평가에 의하여 학업성적을

등급화 하면 어느 고등학교나 똑 같은 정규분포의 등급화가 이루어지게 되어 학업 성취도에 대한 학교간 차별화를 알 수가 없다.

앞으로 보완되어야 내용으로는 학업성적이 필기, 실기 등 둘 이상으로 구성될 경우 자료의 조직화 과정에서 단순히 합계만을 사용할 것이 아니라 각 항목의 분포까지 고려하는 방법이 더 연구되어야 한다.

참 고 문 헌

- [1] 김기완, "퍼스날 컴퓨터를 利用한 相對評價 處理業務 設計에 관한 연구", 단국대학교 경영대학원 경영학과 석사학위논문, pp. 4-10, 1986.
- [2] 김재복, "絕對評價를 위한 教授目標設定과 尺度에 관한 研究", 계명대학교, pp. 5-10, 1977.
- [3] 정종진, 「학교학습의 극대화를 위한 교육평가의 이해」, 양서원, pp. 53-59, 1999.
- [4] 울산광역시교육청, 「초·중고등학교 학교생활기록부 전산 처리 지침」, pp. 6-7, 1998.
- [5] W. Chu, and K. Chiang, "Abstraction of High Level Concepts from Numerical Values in Databases", *Proceedings of the AAAI Workshop on Knowledge Discovery in Databases*, p. 12, July, 1994. (<http://cobase-www.cs.ucla.edu/refs.html>)
- [6] Wesley W. Chu, Kuorong Chiang, and Chih-Cheng Hsu, "An Error-based Conceptual Method of Providing Approximate Query Answers", *Communications of ACM*, p. 18, 1996. (<http://cobase-www.cs.ucla.edu/refs.html>)
- [7] Wesley W. Chu, et al. "CoBase: A Scalable and Extensible Cooperative Information System", *Journal of Intelligence Information Systems*, vol 6, no. 2/3, pp. 223-259, 1996.

저 자 소 개



정 흥(Hong Chung)

1972년 : 한양대학교 원자력공학과 (공학사)
 1976년 : 고려대학교 경영대학원 생산관리 (경영학석사)
 1996년 : 대구카톨릭대학교 전산통계학과 (이학석사)
 1999년 : 대구카톨릭대학교 전산통계학과 (이학박사)

1972년~1981년 : 한국과학기술연구원 선임연구원
 2000년~2001년 : 미국 Washington State University 연구 교수
 1981년~현재 : 계명대학교 컴퓨터공학과 부교수

관심분야 : 지능정보시스템, 소프트웨어공학
 E-mail : jhong@kmu.ac.kr



박은진 (Eun-Jin Park)

1996년 : 계명대학교 컴퓨터공학과 (공학사)
2000년 : 계명대학교 교육대학원 전산교육
(교육학석사)
1999년 : 울산 서여자상업고등학교 교사
2000년 ~ 현재 : 대구 제일여자정보고등학교
교사

관심분야 : 교육공학



장덕성 (Duk-Sung Jang)

1979년 : 경북대학교 컴퓨터공학과 (공학사)
1981년 : 서울대학교 전산학과 (이학석사)
1988년 : 서울대학교 컴퓨터공학과 (공학박
사)
1982년 ~ 1985 : 동아대학교 전산공학과 조
교수

1992년 ~ 1993년 : 미국 University of Colorado 연구교수
1985 ~ 현재 : 계명대학교 컴퓨터공학과 교수

관심분야 : 컴파일러, 시각프로그래밍, 자연어처리, 정보검색,
에이전트

E-mail : dsjang@kmu.ac.kr