

러프집합과 정보이론을 이용한 대학생역량강화 진단

박인규

중부대학교 컴퓨터학과

Diagnosis by Rough Set and Information Theory in Reinforcing the Competencies of the Collegiate

In-Kyoo Park

Dept. of Computer Science Joongbu University

요약 본 논문은 대학생을 대상으로 학생들의 학업 및 취업경쟁력 강화에 필요한 핵심 역량의 도출과 진단에 관한 연구이다. 이러한 데이터의 처리에는 매우 많은 변수로 인한 차원의 증가로 인하여 계산상의 어려움이 수반되어 지고 변수의 중복성과 중요도에 있어서 다양한 통계적 관계가 존재한다. 따라서 범주형 데이터의 분류에서 발생하는 애매함이나 불확실성을 처리하기 위하여 러프집합과 정보 엔트로피를 기반으로 불확실성의 척도를 정의하여 학생들의 유사행동을 분석하고, 기존의 통계적인 방법과의 비교우위를 위하여 속성간의 변별력을 비교하였다. 도출된 공통 핵심역량과 전공핵심역량을 이용하여 학생들이 가지고 있는 역량의 정성적인 보유수준과 부족한 역량을 파악할 수 있기 때문에, 대학생활지도와 취업진로지도의 보조자료로 활용이 가능할 뿐만 아니라 대학 적응을 높이고 취업 활성화에 부합될 수 있다고 사료된다.

주제어 : 역량강화, 정보이론, 불확실성, 엔트로피, 러프집합

Abstract This paper presents the core competencies diagnosis system which targeted our collegiate students in an attempt to induce the core competencies for reinforcing the learning and employment capabilities. Because these days data give rise to a high level of redundancy and dimensionality with time complexity, they are more likely to have spurious relationships, and even the weakest relationships will be highly significant by any statistical test. So as to address the measurement of uncertainties from the classification of categorical data and the implementation of its analytic system, an uncertainty measure of rough entropy and information entropy is defined so that similar behaviors analysis is carried out and the clustering ability is demonstrated in the comparison with the statistical approach. Because the acquired and necessary competencies of the collegiate is deduced by way of the results of the diagnosis, i.e. common core competencies and major core competencies, they facilitate not only the collegiate life and the employment capability reinforcement but also the revitalization of employment and the adjustment to college life.

Key Words : Competency Reinforcement, Information Theory, Uncertainty, Entropy, Rough Set

Received 16 May 2014, Revised 7 July 2014

Accepted 20 August 2014

Corresponding Author: In-Kyoo Park(Joongbu Univ.)

Email: fip2441g@gmail.com

ISSN: 1738-1916

© The Society of Digital Policy & Management. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서론

최근 사회는 지식과 정보가 끊임없이 변화하는 지식 기반 사회로 변모하고 있다. 이러한 사회변화에 따라 세계 각국의 교육현장에서는 다양한 맥락에서 새로운 지식을 창출할 수 있는 역량을 지닌 창의적 인재양성의 필요성이 대두되고 있다[1]. 대학은 사회의 변화에 대응하고 역량을 가진 인재를 양성하기 위해 다양한 방안을 마련하고 있다. 또한 대학교육성과의 객관적인 측정에 대한 관심이 커져 가고 있는 상황에서 이를 구체적으로 측정할 수 있는 지표개발에 대한 요구에 부응하여 학문분야에 따른 학생들의 기초능력 수준에 대한 평가 자료를 제공할 필요가 있다. 2006년 교육과학기술부와 한국직업능력개발원이 개발한 대학생 핵심역량 진단도구(K-CESA)가 현재 많은 대학에서 학생들의 핵심역량을 측정하는데 활용되고 있으나, 핵심역량의 개념구성, 진단 방법, 소요시간 및 비용, 원 데이터의 DB구축 및 활용 등에 있어서 여러 가지의 제약요소가 있어 보다 효율적인 지표개발과 활용에 대한 필요성이 제기되고 있는 것이 현실이다. 현재의 역량진단 도구들은 직종이나 직위에 상관없이 일반적인 직무를 성공적으로 수행하는데 공통적으로 요구되는 지식, 기술, 태도 등을 측정하기 때문에, 실질적으로 본교와 같이 재학생과 졸업 후 취업 현황의 특수성이 반영된 역량 측정과 학문 분야별 특성에 맞는 역량 파악에 대한 요구가 필요한 실정이다.

이러한 맥락에서 방대한 양의 데이터 속에서 쉽게 드러나지 않는 유용한 정보를 찾아내는 데이터 마이닝을 통하여 어떠한 목적의 의사결정을 하는데 유용한 정보로 활용하고자 한다. 이는 데이터에 내재하는 데이터간의 관계, 패턴, 규칙을 발견하는 데 그 목적이 있으며, 작업 유형에 따라서 사용되는 분석기법으로 통계적인 방법은 모든 데이터를 대상으로 평균이나 표준편차, 선형적인 정규분포를 이용하여 유도된 결론에는 어느 정도의 정보의 손실이 발생할 수 있다. 또한 범주형 데이터에서는 속성의 개수가 많고 그 값이 다양하기 때문에 하나의 객체가 여러 개의 집단으로 분류되는 불확실성으로 인하여 기존의 알고리즘에서는 애매함을 완전하게 처리하지 못하고 있는 실정이다. 이러한 일환으로 불확실하고 애매 모호한 성격을 가지는 데이터를 분석하는데 있어서 합리적이고 효과적인이고 새로운 연구역량을 강화하는 프로

그램을 개발하기에 러프 집합이론(rough set theory)이 매우 적합한 방법론이 될 수 있다[2,3].

본 연구에서는 연구역량 강화 진단의 본래의 목적을 달성하기 위하여 불확실성의 척도를 기반으로 더욱 체계적이고 효과적인 창의적 연구역량 진단시스템을 개발하고자 한다.

2. 러프집합 이론

부정확한 데이터로 부터의 데이터간의 관계를 기반으로 하고 있으며 지식베이스 시스템 $S=(U, A, V)$ 에서 객체들의 유한집합 $U=\{x_1, x_2, \dots, x_n\}$, $U \neq \emptyset$ 이며, A는 기본 속성들의 유한집합이다. A에 있는 속성들은 조건속성 C와 결정속성 D로 분류되며, $A=C \cup D$, $V=\cup_{a \in A} V_a$ 이다. 속성들의 모든 부분집합 $B(B \subseteq A)$ 와 임의의 원소 $x_i, x_j \in U$ 라면, 식별불가능 관계(indiscernibility relation)인 이진 관계 $IND(B)$ 를 정의하면 다음과 같다.

$$IND(B) = \{(x_i, x_j) \in U^2 : \forall a \in B, a(x_i) = a(x_j)\} \quad (1)$$

여기서 x_i, x_j 는 지식베이스 시스템 S에서 속성 B의 집합에 의하여 식별 불가능하다고 말한다. 그리고 $B(x)$ 는 객체 x에 할당된 속성 B의 값으로서, $IND(B)$ 는 모든 $B \subseteq A$ 에 대해 U에서 식별불가능 동치관계(equivalence relation)가 된다. 지식베이스 시스템 $S=(U, A, V)$ 이고, $R \subseteq A$ 가 동치관계라면 근사공간(approximate space)은 $AS=(U, R)$ 이다. 또한 U의 원소 x_i 에 대하여 $IND(B)$ 에서 x_i 의 동치 클래스는 식(2)와 같다.

$$[x_i]_{IND(B)} = \cap_{R \in B} [x_i]_R \quad (2)$$

$$\underline{R}(X) = \{x_i \in U | [x_i]_R \subseteq X\} \quad (3)$$

$$\overline{R}(X) = \{x_i \in U | [x_i]_R \cap X \neq \emptyset\} \quad (4)$$

$$\alpha_R(X) = \text{card } \underline{R} / \text{card } \overline{R}, X \neq \emptyset \quad (5)$$

$X \subseteq U$ 라 하면, AS에서 X의 하한근사와 상한근사는 각각 식(3)과 (4)와 같다. 상한과하한의 차이인 집합 $BND(X)$ 를 X의 R-경계(R-boundary)라고 한다. 객체들

의 집합 X에 대한 정확도는 식(5)와 같다.

여기서 $0 \leq \alpha_R(X) \leq 1$ 이고, 부정확성 척도 $\alpha_R(X) = 1 - \alpha_R(X)$ 는 X의 “R-대략적 정의”가 된다. 이러한 부정확성에 대한 수치적 표현은 근사개념에 근거하고 있으며, 지식베이스 시스템의 제한된 분류 증력을 명시적으로 나타낸다.

3. 역량 진단시스템

3.1 불확실성과 러프엔트로피

확률이론이 불확실성에 대한 일어날 가능성을 모델링하는 것이고, 이러한 불확실한 상황에서 추론에 근거해 결정을 내리는 것을 결정이론(decision theory)라 한다. 따라서 정보이론(information theory)은 불확실성을 평가하는 것으로 어떤 확률변수의 불확실성에 대한 측정이 엔트로피라고 할 수 있다[4]. 따라서 정보이론은 불확실성에 대한 평가라고 할 수 있고 정보이론의 기본적인 단위인 엔트로피(entropy)의 계산에도 확률변수가 사용된다. 엔트로피는 Claude Shannon에 의한 정보량으로써 확률변수 x의 불확실성을 나타내는 엔트로피는 각 확률 값에 로그를 취한 것의 기대값(즉, 확률 값에 대한 가중 평균)으로 식(6)과 같이 정의 할 수 있다[5].

$$H(x) = E_X[I(X)] \quad (6)$$

$$= - \sum_{x \in X} p(x) \log_2 p(x)$$

결국 엔트로피는 랜덤변수 x가 가질 수 있는 모든 값(사건)에 대해 정보량을 평균한 것이다[6-8]. 본 논문에서는 기존의 정보이론에서 사용되는 엔트로피에 대한 척도를 변형하여 식(7)과 같이 러프 엔트로피(rough entropy)를 제안하였다.

$$RE(X, Y) = -K \sum_{i=1}^m L \sum_{j=1}^n \log_2 \frac{|X_i \cap Y_j|}{|X_i|} \quad (7)$$

$$K = \text{card}(X_i) / \text{card}(U)$$

$$L = \text{card}(Y_j) / \text{card}(U)$$

결국 불확실성이 높아질수록 정보의 양은 더 많아지

고 엔트로피는 더 커진다. 또한 엔트로피가 커진다는 의미는 에너지 무질서도가 높아지고 에너지의 질이 떨어진다는 의미가 된다. 본 논문에서는 러프집합에서의 지식에 존재하는 속성들의 중복성에 관한 불확실성에 대한 문제를 설정하여 임의의 러프집합에서 러프 엔트로피에 의하여 동치류 X와 Y의 Y에 대한 중복성의 비율과 U에 대한 Y의 비율의 곱으로써 조건부 엔트로피는 항상 음이 아니며, Y값을 알고 있을 때 X값의 무작위한 정도의 양으로 간주하였다.

3.2 MRERS 클러스터링

대학생들의 핵심역량은 많은 요소로 구성되어 있다. 이러한 요소들은 핵심역량에 기여하는 정도에는 차이가 있다. 따라서 핵심역량과 무관하거나 커다란 의미가 없는 역량들은 핵심역량의 도출에 방해요소로 작용할 수 있다. 따라서 제안한 러프엔트로피를 이용하여 요소들 간의 정확한 관계를 측정하기 위하여 변형된 러프 엔트로피 기반 러프집합모형(modified rough-entropy based rough set : MRERS)을 제안한다. 이 방법은 속성에 대하여 러프 엔트로피를 적용하여 속성이 가지는 근사화의 정확도가 항상 0이 아니기 때문에 속성간의 변별력을 유지하는데 보다 안정적이다. 기존의 방법에서는 속성이 가지는 개념간의 중복성을 측정하는 방법으로 속성간의 오류율을 정의하는 방법으로써 이는 개념에 대한 중복성이 일부 동일하게 평가되어 속성간의 변별력이 떨어진다. 따라서 제안된 러프 엔트로피를 군집화 분석에 적용하기 위하여 <Table 1>의 예를 고려하자.

<Table 1> Information system

U / A	C1	C2	C3	C4	D
1	0	2	0	1	0
2	0	1	1	0	1
3	1	0	0	0	0
4	1	1	0	1	0
5	2	2	1	0	1
6	2	0	0	0	1
7	2	0	1	1	0
8	1	0	1	1	1

각각의 속성간의 의존성의 관계를 조사하여 가장 신뢰도가 높은 속성을 군집화를 수행하는 기준이 되는 속성으로 간주하게 된다. 예를 들어 ‘C1’ 속성에 대하여 ‘C3’ 속성의 러프 엔트로피는 다음과 같이 계산된다.

$$\begin{aligned}
 RE(\{1,2\}|\{2,5,7,8\}) &= -(2/8) * (4/8) \log_2(1/2) = 0.0866, \\
 RE(\{3,4,8\}|\{2,5,7,8\}) &= -(3/8) * (4/8) \log_2(1/3) = 0.2060, \\
 RE(\{5,6,7\}|\{2,5,7,8\}) &= -(3/8) * (4/8) \log_2(2/3) = 0.0760, \\
 RE(\{1,2\}|\{1,3,4,6\}) &= -(2/8) * (4/8) \log_2(1/2) = 0.0866, \\
 RE(\{3,4,8\}|\{1,3,4,6\}) &= -(3/8) * (4/8) \log_2(1/3) = 0.0766, \\
 RE(\{5,6,7\}|\{1,3,4,6\}) &= -(3/8) * (4/8) \log_2(1/3) = 0.0766
 \end{aligned}$$

따라서 엔트로피는 속성간의 불확실성을 나타내기 때문에 가장 작은 값을 취함으로써 가장 신뢰도가 높은 속성간의 러프정도(roughness)는 식(8)에 의하여 구할 수 있다.

$$MRE_{a_j}(a_i = \alpha) = \min(RE_{Y_{a_j}}(X|a_i = \alpha)) \quad (8)$$

MRE('C3'='1') = min(0.0866, 0.2060, 0.0760) = 0.076, MRE('C3'='0') = min(0.0866, 0.0766, 0.0766) = 0.0766이다. 다섯 가지의 속성에서 각 속성의 평균 불완전성의 평균을 식(9)에 의하여 구할 수 있다. <표 3>에서 가로의 'C3'과 세로의 'C1'의 중복성은 0.0766이 된다.

$$\begin{aligned}
 MR_{a_j}(a_i) &= K \frac{1}{|V(a_i)|} \quad (9) \\
 K &= MRE_{a_j}(a_i) + \dots + MRE_{a_j}(a_{i_{|V(a_i)|}})
 \end{aligned}$$

MR('C3') = ('C1'(X|'C3'='0') + 'C1'(X|'C3'='1')) / V'C1' = (0.0766+0.0766)/2 = 0.0766이다. 결국 'C1'에 대한 'C3'소성의 전체적인 불완전성은 0.152로 나타났다. 이러한 방법으로 모든 속성간의 관계가 <표 3>에 나타나 있다. 또한 식(10)을 이용하여 다섯 가지의 속성에서 발생하는 신뢰도의 평균적인 러프정도(mean roughness)을 구할 수 있다.

$$MMR(a_i) = \frac{MR(a_1) + \dots + MR(a_i)}{|V(a_i)|} \quad (10)$$

결국 주어진 다섯 개의 속성에서 군집화속성인 MMR은 식(11)로 정의되고 모든 속성의 최소값에 해당하는 속성이 결정된다.

$$MMMR(a_i) = \min(MMR_{a_1}(a_i), \dots, MMR_{a_j}(a_i)) \quad (11)$$

<Table 2> MMR accuracy of <Table 2>

attr.	rough entropy mean roughness					mean
	C1	C2	C3	C4	D	
C1	0	0.0577	0.1444	0.1444	0.1444	0.1227
C2	0.0542	0	0.1732	0.1732	0.1732	0.1435
C3	0.0760	0.0866	0	0.1732	0.0719	0.1020
C4	0.0760	0.0866	0.1732	0	0.0719	0.1020
D	0.0760	0.0866	0.0719	0.0719	0	0.0766

<Table 1>의 지식 시스템에 대하여 [Fig. 1]의 알고리즘을 이용하여 속성간의 최소평균 정확도를 계산한 결과가 <Table 2>에 나타나 있다. 결과에서 알 수 있듯이 속성에 대한 중요도가 가장 높은 속성은 'D'이다. 따라서 이 속성은 지식 테이터를 분류하기 위한 군집속성으로 간주하게 되고 U의 지식을 {1,3,4,7}와 {2,5,6,8}으로 분할할 수 있게 된다.

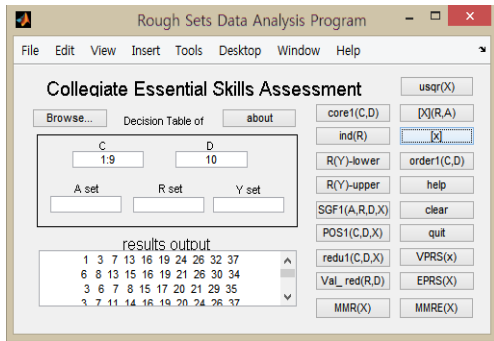
Input : A complete knowledge (U, A)
Output: A selected attribute subset

- (1) Algorithm MRERS(U,k)
- (2) Set current number of cluster(CNC)=1
- (3) Set Parent node PNode=U
- (3) Loop1:
- (4) If CNC<k and CNC≠1
- (5) Then PNode = Procedure PNode(C)
- (6) End
- (7) For all $a_i \in A(i=1, \dots, n)$
- (8) Determine $[X_i]_{IND(a_i)}$
- (9) For $\forall a_i \in A$
- (10) $RE(a_j, a_i) = - \frac{card(a_i)}{card(U)} \sum_{i=1}^m \frac{card(a_j)}{card(U)} \sum_{j=1}^n \log_2 \frac{|a_i \cap a_j|}{|a_i|}$
- (11) End
- (12) $MMR(a_i) = \text{mean}(RE(a_j, a_i))$
- (13) End
- (14) $MMMR = \min(MMR(a_i)) // i=1, \dots, n$
- (15) Determine clustering attribute a_i
- (16) Do binary partition by the clustering attribute a_i
- (17) CNC = number of leaf nodes
- (18) goto Loop 1
- (19) Procedure PNode(cnc)
- (20) $i=1$
- (21) Do Until $i < cnc$
- (22) $size(i) = \text{elements of cluster } i$
- (23) $i=i+1$
- (24) Loop
- (25) Determine $\max(size(i))$
- (26) Return (sets of elements in cluster i)

[Fig. 1] MRERS clustering algorithm

3.3 시스템 구현

제안된 기법을 대학생들의 역량강화 진단에 적용하기 위해 [Fig. 2]와 같이 Matlab(Version R2010A)을 이용하여 역량강화진단 의사결정 시스템(decision support system)을 구현하였다.



[Fig. 2] Diagnosis System of MRERS

4. 결과 고찰

4.1 불확실성과 러프엔트로피

설문조사는 본 대학의 학생 수준과 전공별 특성에 맞는 역량을 도출하기 위해 5개 단과대학, 57개 학과 약 200여명의 교수의 설문조사와 분석에 사용된 자료는 5개 단과대학, 53개 학과 152명이 응답하였다.

<Table 3> Competencies of Questionnaire

no.	competencies	no.	competencies
1	analytic thinking	20	self-management
2	conceptual thinking	21	planning and execution
3	creative thinking	22	information/sharing
4	strategic thinking	23	process management
5	systematic thinking	24	speciality
6	making decision	25	insight of job
7	self-development	26	global mind
8	self-confirmation	27	client-directed
9	flexibility/opening	28	emotional bondage
10	trigger role	29	conflict management
11	morality/work ethic	30	cooperation and control
12	achievement-directed	31	persuasion power
13	impromptu action	32	team work/cooperation
14	problem tackling	33	presentation skill
15	determination	34	oral communication
16	responsibility	35	document communication
17	self-control	36	open communication
18	action-directed	37	leadership
19	positive mind		

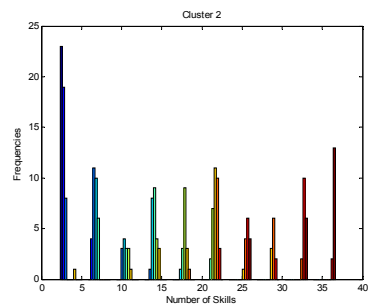
자료 분석은 역량에 대한 “중요도”, “보유도”, “필요도”, “역량별 우선순위”를 알아보기 위하여 기존 방법으로 빈도분석과 평균값을 이용하였으며, 역량별 유사성을 검증하기 위하여 요인분석을 실시하였다[9]. 따라서 기존의 방법에서는 단순히 빈도수를 고려하여 역량을 도출하지만 제안된 방법에서는 역량간의 상호연관 관계의 불확실성을 고려하여 역량을 평가하는데 중점을 두었다. 현재 대학들이 학생들의 역량평가에 가장 많이 사용하고 있는 ‘대학생 핵심 역량 진단도구(K-CESA)와 대학들의 역량측정 사례, 주요기업의 인재상을 기초로 선별된 37개 역량은 <Table 3>과 같다.

4.2 핵심역량의 진단

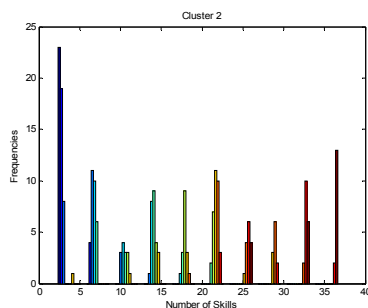
37개의 역량에 대하여 핵심적인 10개의 역량의 선택에 대한 분석을 위하여 제안된 알고리즘을 이용하여 <Table 4>와 같이 교수들의 선택의 유형을 8개의 유사성에 의해 분류할 수 있었고 각각의 유형에 대한 분포를 [Fig. 3]부터 [Fig. 10]까지 나타내었다.

<Table 4> Clustering of 10 competencies

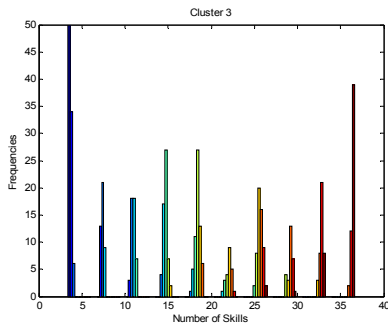
clusters	1	2	3	4	5	6	7	8
professors	46	23	50	5	7	10	9	2



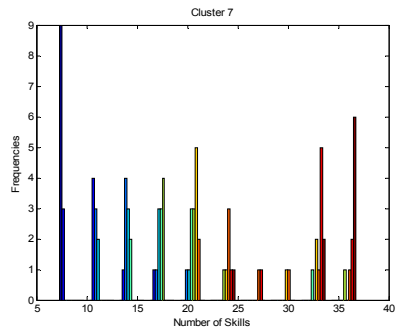
[Fig. 3] Competencies in Cluster 1



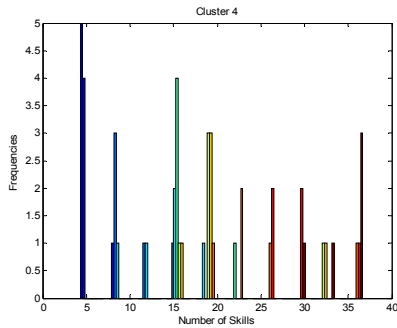
[Fig. 4] Competencies in Cluster 2



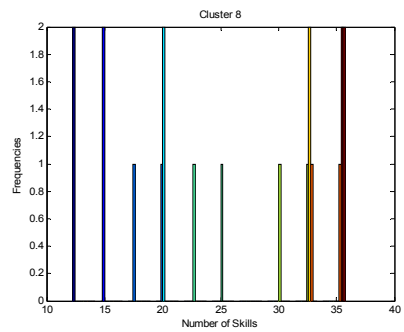
[Fig. 5] Competencies in Cluster 3



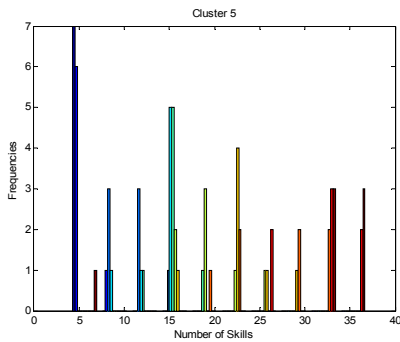
[Fig. 9] Competencies in Cluster 7



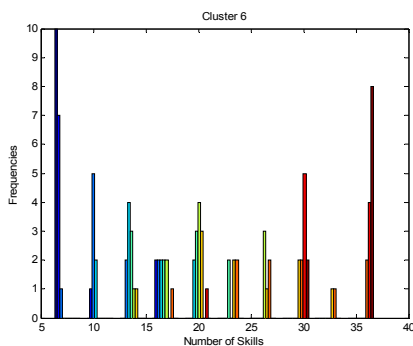
[Fig. 6] Competencies in Cluster 4



[Fig. 10] Competencies in Cluster 8



[Fig. 7] Competencies in Cluster 5



[Fig. 8] Competencies in Cluster 6

역량의 우선순위에 대한 전체적인 평가에서는 기존방법의 경우에 ‘책임감’이 가장 높게 나타났으며 ‘창의적 사고’, ‘자기개발’, ‘자기관리’, ‘전문성’순으로 나타났다. 반면에 제안된 방법에서는 ‘책임감’, ‘창의적 사고’, ‘자기개발’, ‘리더십’, ‘의사결정/판단력’순으로 나타났다. 결과적으로 기존의 방법에서는 동일한 평가치가 많았다. 또한 <Table 5>에서는 단과대학별 역량의 우선순위를 도출하였다. 사회과학대학의 경우 ‘프레젠테이션 스킬’과 ‘창의적 사고’가 동일한 평가치를 가지고 있고 경찰경호대학의 경우 ‘긍정적 마인드’, ‘문서의사소통’과 ‘리더십’, 관광보건대학의 경우 공과대학의 경우 ‘분석적 사고’, ‘자기관리’와 ‘문제해결’이 동일한 평가치를 가지고 있었다. 예술체육대학의 경우 ‘창의적 사고’와 ‘문서의사소통’이 동일한 평가치를 가지고 있다. 따라서 이러한 불확실성을 극복하기 위하여 러프 엔트로피를 적용한 방법의 경우에는 서로 동일한 평가치가 상대적으로 적게 나타났다. 따라서 기존방법과의 비교에서 동일한 값을 가지는 역량이 비교적 적게 추출되어 제안된 방법이 역량간의 변별력에서 상대적으로 양호하다고 할 수 있다.

<Table 5> Reference competencies

colleges	methods	no. of competencies				
		probability	16	33	3	32
social science	rough set	16	3	7	20	21
	probability	10	19	35	37	7
police&guard	rough set	13	30	5	6	11
	probability	24	7	16	20	35
tour&health	rough set	3	20	23	7	36
	probability	3	1	20	14	16
engineering	rough set	16	20	34	29	3
	probability	16	3	35	19	6
art&physica	rough set	3	17	7	31	19

전체 역량에서 학생들에게 요구되는 10개 역량의 선택순위를 <Table 6>과 같이 리프엔트로피에 의하여 산출하였다. 첫 번째로 선택한 역량이 역량간의 연관관계의 측면에서 우선순위가 가장 높았고, 다음으로 네 번째, 일곱 번째, 아홉 번째, 다섯 번째의 역량으로 나타났다.

<Table 6> Priority of Competencies

order	normalized mean values										accu.	rank
	1	0.35	0.29	0.15	0.09	0.08	0.09	0.13	0.29	0.15		
2	0.89	0.17	0.17	0.06	0.14	0.12	0.17	0.22	0.39	0.26	10	
3	0.93	0.45	0.24	0.24	0.15	0.11	0.08	0.22	0.31	0.30	6	
4	0.92	0.46	0.22	0.15	0.20	0.09	0.13	0.20	0.49	0.32	2	
5	0.70	0.34	0.18	0.17	0.23	0.13	0.12	0.18	0.36	0.27	5	
6	0.66	0.30	0.21	0.18	0.12	0.14	0.15	0.21	0.28	0.25	7	
7	0.67	0.32	0.18	0.18	0.11	0.24	0.14	0.17	0.43	0.27	3	
8	0.74	0.34	0.24	0.23	0.14	0.20	0.12	0.24	0.51	0.31	8	
9	0.67	0.39	0.25	0.23	0.15	0.21	0.19	0.18	0.67	0.33	4	
10	0.57	0.31	0.21	0.20	0.11	0.15	0.10	0.10	0.15	0.21	9	

현재 학생역량의 보유도에 대한 평가에서는 <Table 7>과 같이 기존방법에서는 ‘팀워크 및 협동’이 가장 높게 나타났으며 ‘긍정적 마인드’, ‘정서적 유대형성’, ‘윤리성 직업의식’, ‘유연성 개방성’순으로 나타났다. 제안된 방법에서는 ‘성취지향’, ‘문제해결’, ‘자기관리’, ‘전략적 사고’, ‘갈등관리’순으로 나타났다. 단과대학별로는 경찰경호 대학의 경우 ‘자기조절’, ‘책임감’, ‘계획수립’이 동일한 평균 값을 가지고 관광보건 대학의 경우 ‘윤리성 직업의식’, ‘구두의사소통’이 동일하였고 공과대학의 경우 ‘윤리성 직업의식’, ‘의사결정 판단력’이 동일하였다. 예술체육대학의 경우 ‘팀워크 및 협동’, ‘계획수립’이 동일하였다.

<Table 7> Acquired Competencies

colleges	methods	no. of competencies				
		probability	32	19	28	29
social science	rough set	32	24	7	19	2
	probability	10	17	16	21	22
police&guard	rough set	8	36	15	21	27
	probability	9	11	34	19	37
tour&health	rough set	24	7	16	3	37
	probability	28	22	11	6	32
engineering	rough set	7	21	33	2	19
	probability	20	32	21	36	22
art&physica	rough set	26	32	34	14	5

또한 필요한 역량의 전체적인 평가에서는 <Table 8>과 같이 기존방법에서는 ‘창의적 사고’가 가장 높게 나타났으며 ‘분석적 사고’, ‘정서적 유대형성’, ‘전문성’, ‘자기개발’과 ‘문제해결’순으로 나타났다. 반면에 제안된 방법에서는 ‘문제해결’, ‘글로벌 마인드’, ‘체계적 사고’, ‘자기개발’과 ‘윤리성 직업의식’순이고, 단과대학별로는 사회과학대학의 경우 ‘성취지향’과 ‘주도성’이 동일하였고 경찰경호 대학의 경우 네 개의 역량이 동일하였다. 공과대학의 경우 ‘자기관리’, ‘자기조절’이 동일하였다.

<Table 8> Necessary Competencies

colleges	methods	no. of competencies				
		probability	3	1	12	10
social science	rough set	33	21	7	19	35
	probability	24	8	7	19	15
police&guard	rough set	8	36	15	33	21
	probability	24	3	14	7	6
tour&health	rough set	14	34	5	33	12
	probability	3	20	17	14	4
engineering	rough set	14	32	5	12	31
	probability	16	6	1	4	7
art&physica	rough set	16	3	32	37	18

예술체육대학의 경우 ‘의사결정 판단력’, ‘분석적 사고’가 동일하였다. 따라서 학생들의 보유역량에서 기존 방법에 의해 도출된 역량은 ‘팀워크 및 협동’인 반면에 제안된 방법에서는 ‘설득력’로 나타났고 조사대상자들의 분류는 11명, 44명, 69명, 26명과 2명으로 각각 분류되었다.

또한 학생들에게 부족한 필요 역량에서 기존 방법에 의해 도출된 역량은 ‘창의적 사고’인 반면에 제안된 방법에서는 ‘자기개발’로 나타났고 조사대상자들의 분류는 3명, 3명, 30명, 48명과 68명으로 각각 분류되었다.

<Table 9> Diagnosis of Core Competencies

colleges	methods	core competencies		
		probability	32	1
social science	rough set	16	25	3
	probability	10	19	35
police&guard	rough set	13	10	7
	probability	35	34	21
tour&health	rough set	20	17	6
	probability	32	1	5
engineering	rough set	20	3	16
	probability	35	19	11
art&physica	rough set	3	36	2

결국 <Table 5>~<Table 8>을 통하여 ‘자기개발’, ‘창의적 사고’, ‘문제해결’, ‘책임감’과 ‘글로벌 마인드’가 적합한 전체적인 공통역량으로 도출되었고, <Table 9>와 같이 단과대학별 전공역량에서는 사회과학대학의 경

우 '책임감', '업에 대한 통찰'과 '창의적 사고'이고 경찰경호대학의 경우 '즉각적 대처', '주도성'과 '자기개발'로 도출되었다. 또한 관광보건대학의 경우 "자기관리", '자기조절'과 '의사결정/판단력'이고 공과대학의 경우 '자기관리', '창의적 사고'와 '책임감'으로 도출되었으며 예술체육대학의 경우 '창의적 사고', '열린의사 소통/중재'와 '개념적 사고'로 도출되었다. 따라서 역량의 부족한 역량을 우선 순위로 선정하여 핵심역량은 공통역량 5개와 단과대학별 전공역량 11개로 총 16개로 도출되었다. 결국 속성간의 애매함은 속성의 개수와 선형적인 관계를 가지고 있다는 것을 확인할 수 있었다. 따라서 학생들에게 기업체에 대한 취업지도를 수행하는 과정에서 보다 정확하고 정성적인 정보를 가지고 학생들을 지도할 수 있다.

4. 결론

본 논문에서는 정보이론과 러프집합을 결합하여 범주형 데이터를 유사성에 따라 분류하는 불확실성의 척도를 제안하고 이를 기반으로 하는 진단 시스템을 구현하였다. 제안된 척도는 속성간의 변별력을 향상시키기 위하여 속성간의 연관관계를 강화하였다. 기존방법과의 비교우위를 통하여 역량 값의 동일 여부에 대한 변별력에서 단과대학별로 20~30%의 향상을 보였다. 제안된 진단도구를 활용하여 학생들의 역량 보유수준과 상황을 파악할 수 있고 대학생활지도 뿐만 아니라 취업진로지도의 자료로 활용이 가능하다. 따라서 단과대학별로 특성을 고려하여 도출된 전공핵심역량은 향후 학생들의 진단결과를 토대로 그에 맞는 프로그램이나 교과목의 개설, 교육과정 개편에 이르기 까지 실질적인 교육서비스 향상에 기초자료로써 활용될 수 있을 것으로 기대된다. 또한, 현재 대학들이 교육의 효과나 학생수준에 대한 객관적 진단 및 평가에 일조할 것으로 사료된다.

REFERENCES

[1] M.H. Son, The meaning of competency on daily attribute of pragmatic knowledge: knowledge-based society vs society-based society, Institute of Korean

education process, Vol. 24, No. 4, pp. 1-25, 2006

[2] Pawlak, Z. Rough sets - Theoretical Aspects for Reasoning about Data, Klumer, 1991

[3] Pawlak, Z. Rough sets, International Horland of Information and Computer Sciences, Vol.11, No. 5, pp. 341-356, 1982

[4] Shannon, C., L., The mathematical theory of communication, Bell System Technical Journal, Vol. 27, 1948

[5] Beaubouef, T., Petry, F. E. and Arora, G., Information-theoretic measures of uncertainty for rough sets and rough relational databases, Information Science, Vol. 109, No. 1-4, pp. 185-195, 1998.

[6] R. Vashish, M.L. Garg, Rule Generation based on Reduct and Core: A Rough Set Approach, International Journal of Computer Applications, Vol. 29, No. 9, pp. 0975-8887, 2011

[7] J. Aheng, R. Yan, Attribute Reduction based on Cross Entropy in Rough Set Theory, Journal of Information and Computational Science, Vol. 9, No. 3, pp. 745-750, 2012

[8] L. Sun, J. Xu, Z. Xue, L. Zheng, Rough Entropy-based Feature Selection and Its Application, Journal of Information and Computational Science, Vol. 8, No. 9, pp. 1525-1532, 2011

[9] TaeJin Choi et. al., Diagnosis Tool for Joongbu Core Competency, Teaching and Study Support Center, 2014

박인규(Park, In Kyoo)



· 1985년 2월 : 연세대학교 공학석사
 · 1997년 2월 : 원광대학교 공학박사
 · 1997년 2월 ~ 현재 : 중부대학교 컴퓨터학과 교수
 · 관심분야 : 데이터마이닝, 러프집합
 · E-Mail : fip2441g@gmail.com