

정보공시 자료를 이용한 교육/연구성과 영향요인 추출 및 대학의 군집 분석

신선미[†] · 김현철^{††}

요 약

본 연구의 목적은 통계 분석과 데이터마이닝 기법을 이용하여 대학정보공시 자료 속의 새로운 패턴이나 의미 있는 결과를 도출함으로써 대학의 교육역량 및 경쟁력을 나타내는 지표 개선에 유용한 지식을 제공하는 것이다. 이를 위해 교육/연구성과를 나타내는 취업률, 기술이전건수, 전임교원 1인당 논문수 지표의 영향요인 탐색 및 의사결정나무 모형 도출과 대학평가 관련지표를 이용한 대학의 군집분석을 실시하였다. 연구 결과, 대학 교육/연구성과 지표의 공통 영향요인은 신입생충원율, 재학생충원율, 전임교원 1인당 학생수로 나타났다. 군집분석에서는 전체 대학, 대학규모별, 소재지별로 각각 실시하였을 때 유명대학, 예체능·비이공계·종교지도자양성 대학, 그 외 대학으로 군집이 형성되는 양상을 보였고, 그 주요 영향요인으로는 대학 교육/연구성과 지표인 취업률, 기술이전건수 등으로 나타났다.

주제어 : 대학정보공시, 데이터마이닝, 교육/연구성과

Data Mining Analysis of Educational and Research Achievements of Korean Universities Using Public Open Data Services

Sun Mi Shin[†] · Hyeon Cheol Kim^{††}

ABSTRACT

The purpose of this study is to provide useful knowledge for improving indicators that represent competitiveness and educational competency of the university by deriving a new pattern or the meaningful results from the data of information disclosure of universities using statistical analysis and data mining techniques. To achieve this, a model of decision tree was made and various factors that affect education/research performance such as employment rate, the number of technology transfer and papers per full-time faculty were explored. In addition to this, the cluster analysis of universities was conducted using attributes related to evaluation of university. According to the analysis, common factors affecting higher education/research performance are following indicators ; incoming student recruitment rate, enrollment rate, and the number of students per full-time faculty. In the cluster analysis, when performed by the entire university, the size, location of the university respectively, clusters are mainly formed by well-known universities, art · physical · non-science and engineering · religious leaders training universities, and others. The main influencing factors of this cluster are higher education/research performance indicators such as employment rate and the number of technology transfer.

Keywords : Information Disclosure of Universities, Data Mining, Higher Education/Research Performance

† 정 회 원: 고려대학교 교육대학원 석사과정
†† 종신회원: 고려대학교 컴퓨터교육과 교수 (교신저자)
논문접수: 2014년 1월 10일, 심사완료: 2014년 1월 21일, 게재확정: 2014년 1월 28일
* 본 논문은 2012년 고려대학교 사범대학 특별연구비 지원으로 수행되었음

1. 서론

최근 교육부 자문기구인 대학 구조개혁위원회는 학령인구 감소에 따라 2023년까지 입학정원을 16만 명 감축하고 부실대학을 퇴출하는 대규모 대학 구조개혁안을 발표하였다[1]. 대학은 취업률 등 교육역량을 나타내는 지표가 곧 대학의 존폐와 직결되는 상황에 처해 있으며, 이로 인해 지표개선을 위한 자구 노력을 기울이고 있다[3][4]. 그 예로 교육부의 13학년도 정부재정지원제한대학으로 지정되었던 대학 43개교 중 26개교(약 60%)가 전임교원확보율과 장학금 지급률을 높이는 지표개선 노력 등을 통해 정부재정지원제한대학에서 벗어났음을 알 수 있다[2]. 대학은 각종 대학 평가에 대비하여 대학알리미 사이트를 통해 공개되는 지표들의 향상과 대학 교육역량을 강화하기 위한 노력이 불가피하다. 심인선, 정성희(2007)는 대학의 지표 향상 노력이 교육과 연구에 도움이 되므로 필요하다고 하였으며[3], 김순우 등(2010)은 대학정보공시를 통해 공개되는 지표가 대학의 수준을 평가하므로 교육역량강화에 힘써야 한다고 하였다[4].

한편 빅데이터 시대를 맞이하면서 데이터의 활용, 분석 등이 각광을 받고 있다[5][6]. 데이터의 분석이나 데이터를 통한 미래상황 예측을 위해 데이터마이닝은 오랜 기간 동안 과학적이고 객관적인 데이터 분석방법으로 활용되어 왔다[7]. 현재까지 교육부와 대학정보공시 총괄관리기관에서 실시한 공시자료 분석의 결과물은 보도자료이다. 보도자료는 공시자료의 정확한 통계치를 대국민에게 보도하는 데에 목적이 있다. 공시되는 각 항목의 합계, 평균값 등 기초 통계를 산출하고, 전년 대비 증감률, 항목별 상위대학 리스트 등이 주요 내용으로 구성된다[8]. 구 대학정보공시 총괄관리기관이었던 한국교육개발원에서는 보도자료 이외에 분석자료집을 통해 공시자료의 분석결과를 제공하였다. 분석자료집의 내용은 주요 지표의 설립유형별, 계열별, 소재지별 합계, 평균 등의 통계 결과이다[9].

이러한 통계 결과는 여러 항목과의 관계나 데이터 속의 숨어 있는 의미를 찾는 데에 한계점이 있다. 공시자료의 여러 항목 간에는 어떤 연관성이 있는지, 대학 간에는 어떤 데이터의 유사성이

있는지 데이터 속에 숨은 의미를 찾고, 예측 가능한 모형을 만들어본다면 대학의 지표 개선에 도움이 되는 지식을 마련할 수 있을 것이다.

기존의 연구에서는 지표에 영향을 미치는 요인을 찾는 것이 주를 이루었으나, 이 연구에서는 다양한 방법과 분석도구를 이용하여 영향요인 탐색과 함께 의미 있는 규칙이나 패턴을 찾고 지표개선 전략을 도출한다는 점에서 차별성이 있다.

따라서 본 연구에서는 통계 분석과 데이터마이닝의 여러 기법을 활용하여 대학알리미 공시 지표들 간, 그리고 대학들 간의 새로운 패턴이나 의미 있는 지식을 도출하고자 한다. 이를 통해 대학의 지표 개선을 위한 전략 수립 시 유용한 지식을 제공할 것이다.

2. 이론적 배경

2.1 데이터마이닝

데이터마이닝은 대량의 데이터 속의 숨은 지식이나 패턴을 찾아내는 과정이다[7]. 데이터마이닝의 기법에는 회귀 분석, 분류 분석, 연관규칙 분석, 특징선택(Feature selection), 군집 분석 등으로 다양하다[10].

회귀 분석은 하나 이상의 독립변수가 하나의 목표변수에 미치는 영향을 알아보고자 할 때 사용된다. 이 때 독립변수가 하나이면 단순회귀분석, 두 개 이상이면 다중회귀분석이라 일컫는다[11]. 다중회귀분석은 회귀식을 통해 독립변수의 값의 변화에 따른 목표변수의 값을 예측해 볼 수 있으며, 표준화된 회귀계수를 비교해봄으로써 독립변수들의 영향력을 파악할 수 있다[11].

분류 분석은 데이터마이닝에서 가장 많이 사용되는 분석방법으로, 데이터로부터 규칙을 분석하여 모델을 생성하고, 새로운 데이터에 적용하여 분류, 예측하는 방법이다. 분류분석 방법 중 의사결정나무(Decision Tree)가 많이 사용되며, 이것은 변수 영역의 반복적 분할을 통해 규칙을 생성하는 기법이다[12]. 의사결정나무 기법은 영향요인의 탐색은 물론, 계층적으로 관계 구조를 파악할 수 있으며 결과의 적용이 용이하다는 장점이 있다. 또한 선형성, 정규성, 등분산성의 세 가지 가정이 불

필요한 비모수적인 방법이다[12]. 그러나 다른 데이터마이닝 기법에 비해 예측력이 대체로 떨어지며, 데이터의 약간의 변화에도 상이한 결과를 나타낼 수도 있어 불안정한 단점이 있다[10].

특징 선택(Feature Selection)은 원본 데이터 집합에서 목표변수에 대해 연관성이 큰 독립변수를 찾는 방법으로 데이터의 차원(Dimension)을 감소시켜 빠른 연산을 가능하게 한다. 또한 관계없는 데이터를 제거함으로써 분류기(classifier)의 성능을 향상시킬 수 있으며, 복잡도를 낮추어 해석을 쉽게 하는 장점이 있다[13][14].

군집 분석이란 다양한 특성을 지닌 대상으로부터 유사한 성격의 대상들을 그룹화하여 의미를 탐색하는 분석 방법이다[15]. 군집분석은 명확한 분류기준이 없거나 사전정보가 없어도 사용될 수 있으며 이상치나 특이값을 발견하고 결측치를 보정하는 등에 활용될 수 있는 장점이 있다. 하지만 사전정보가 없는 자료에 대해 탐색적으로 시행착오를 반복하면서 연구자의 주관이 개입될 수 있고, 결과나 해석이 연구자에 따라 달라질 수 있다는 단점이 있다[10].

데이터마이닝 기법마다 장·단점이 존재하며, 서로 다른 형태의 결과물을 보여준다. 따라서 다양한 기법을 적용해보고 이를 종합하여 연구문제를 해결하고자 한다.

2.2 대학정보공시

대학정보공시제도는 2008년 「교육관련기관 정보공개에 관한 특례법」에 근거하여 대학 운영, 학생, 교원, 교육여건, 재정 등에 관한 대학의 정보를 대국민에게 공개하는 것이다[16]. 대학 정보의 공개 방법은 '대학알리미'라 일컫는 대국민포털사이트를 통해 이루어지고 있다[17].

2013년 기준 대학정보공시의 대상학교는 고등교육법 제2조와 그 밖의 다른 법률에 의해 설립된 총 431개교이며[16], 공시항목은 14개 분야, 67개 항목, 103개의 세부 항목(2013년 기준)으로 이루어져 있으며 매년 정기공시는 4월, 6월, 8월, 10월에 진행되고 있다[8].

대학정보공시제도가 도입된 이후 정부, 대학평가기관에서는 대학평가의 공정성 확보를 위해 공

시자료를 활용하고 있다. 대학역량강화사업, 학부교육 선도대학지원사업, 정부재정지원제한대학평가, 외국인 유학생 유치·관리 역량 인증제, 산학협력선도대학육성사업, 국가장학금사업, 기관인증평가 등에 공시자료를 토대로 지표값이 재산출되어 활용되고 있다[18][19][20][21][22].

2.3 관련 연구

대학정보공시 자료 중에서 교육/연구성과 지표로 정의되어 있는 취업률, 기술이전실적, 논문실적의 영향요인에 관한 선행연구를 살펴보았다.

취업률의 영향요인과 관련한 선행연구에서는 학점이 취업률에 양의 상관관계가 있다는 다수의 연구 결과가 있었다[23][24][25][26][27][28]. 수도권 대학의 경우 재학생충원율이 취업률의 영향요인으로 나타났고[29], 전임교원 강의담당비율이 취업률과 상관관계가 있는 것으로 나타났다[30]. 김경일(2013)은 산학협력의 활성화 수준이 교육성과에 영향을 미치는 것으로 분석하였다[31].

기술이전성과 영향요인과 관련한 연구에서 대규모 대학은 전임교원수, 사립대학은 특허관련 정부지원금이 영향요인으로 나타났다[32]. 김경진(2010)은 과학기술분야의 수주연구비, 특허 출원건수, 소병우(2009)는 연구비 규모, 김철희, 이상돈(2007)은 SCI급 논문수, 국내특허 등록수, 국제특허 등록수를 기술이전성과의 영향요인으로 분석하였다[33][34][35]. 또한 대학 연구활동비와 산학협력 성과 간 관계분석 연구에서 대학 연구비가 늘어날수록 기술이전 성과가 늘어날 가능성이 높다고 하였다[36].

교원연구성과 영향요인에 관한 선행연구에서는 연구를 선호하는 교수의 개인적 특성과 대학의 연구 환경이 영향요인으로 나타났다[37]. 김훈호, 박환보(2011)는 전임교원 1인당 연구비와 강의담당비율[37], 김화영(2013)의 연구는 전임교원 1인당 연구비가 전임교원 1인당 논문수에 양의 상관관계가 있는 것으로 나타났다[30]. 이정미, 김민희(2010)의 연구에서는 대학의 연구성과에 교수1인당 학생수, 정부지원금이 영향이 큰 것으로 나타났다[38].

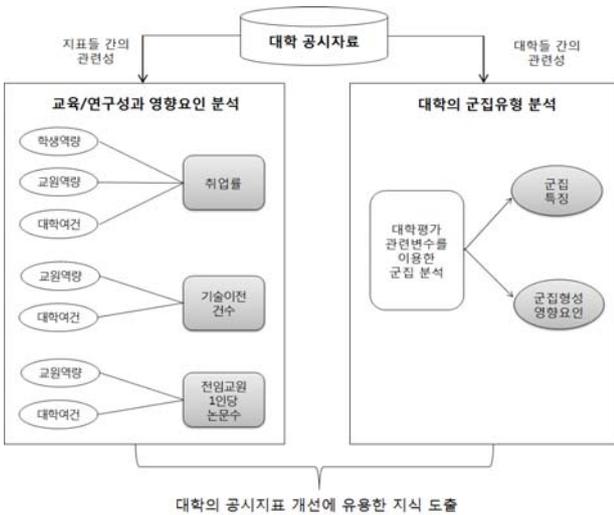
다음으로 대학의 군집분석과 관련한 선행연구

를 살펴보았다. 박혜정, 김종태(2013)는 대구, 경북지역의 24개 대학정보공시 자료를 이용하여 재학생증원율과 취업률에 대해 계층적 군집분석과 서포트벡터 군집분석의 방법으로 대학을 분류하였다. 그 결과 1그룹은 4년제 소규모대학, 2그룹은 4년제 대학 중 안정적인 대학, 3그룹은 신성대학, 4그룹은 특성화대학이라는 특징을 찾아내었다[39]. 서민원, 배성근(2012)은 국립대학 4그룹, 사립대학 3그룹으로 분류하였다[15].

3. 연구 방법

3.1 연구 모형

본 연구는 <그림 1>과 같이 대학알리미 사이트에서 추출한 대학정보공시 자료로 크게 두 가지의 분석을 실시하였다.



<그림 1> 연구 모형

첫째로 교육/연구성과를 나타내는 취업률, 기술이전건수, 전임교원 1인당 논문수 지표에 대해 영향요인을 분석하였다. 취업률은 학생역량, 교원역량, 대학여건 관련 지표를 변수로 선택하였으며, 기술이전건수와 전임교원 1인당 논문수는 교원역량, 대학여건 관련 지표를 변수로 선택하여 영향요인을 알아보았다. 둘째로 대학정보공시 자료로 대학을 군집화하면 어떠한 유사성을 가진 그룹이 형성되는지 알아보고, 그 영향요인에 대해 분석하였다. 대학 평가에 쓰이는 지표를 선택하여 1차로

전체 연구대상 대학을 대상으로, 2차로 대학규모별, 소재지별로 군집분석을 실시하였다. 이 두 가지 분석을 종합하여 대학의 경쟁력을 나타내는 지표 개선에 유용한 지식을 도출하였다.

3.2 연구 대상

본 연구의 대상은 2012년 대학정보공시 대상학교 중 고등교육법 제2조에 근거한 4년제 일반대학(194개교)의 공시자료이다. 4년제 일반대학 중 산업대학에서 일반대학으로 전환되거나 전문대학에서 일반대학으로 전환되는 등 학제가 변경된 대학 및 통폐합 대학(13개교)을 제외하고 최종적으로 181개교의 공시자료를 연구의 대상으로 선정하였다. 기술이전건수의 영향요인 분석 시에는 예술, 체육, 종교지도자양성 대학 등 이공계열이 없는 학교(31개교)를 제외한 150개교를 대상으로 분석을 추가적으로 진행하였다.

3.3 자료의 수집 및 전처리

대학알리미 사이트로부터 통합비교검색과 주요 지표검색을 통해 PDF공시 항목을 제외한 공시년도 2012년의 전체 항목을 대상으로 자료를 추출하였다. 공시자료의 추출시점은 2013년 9월 20일부터 23일까지이며, 추출 당시 2013년의 공시자료의 경우 미공시된 항목(10월 공시)이 존재하여 2012년 공시자료를 대상으로 하였다.

회계항목의 경우 설립구분(국공립, 사립)별 회계방식이 상이하고, 분교(건국대, 고려대, 단국대, 동국대, 상명대, 연세대, 한국외대, 한양대, 홍익대)의 경우 회계자료가 본교에 통합 공시되어 분교의 값이 존재하지 않으므로 제외하였다. 또한 수익용기본재산확보현황의 경우 사립대학만 해당되는 항목이므로 제외하였고, 산학협력단, 창업현황의 경우 자료가 미존재하는 대학이 많아 분석에서 제외하였다.

대학알리미에서 추출된 자료를 바탕으로 인스턴스는 대학, 변수는 공시 지표로 하여 데이터 집합(data set)을 생성하였다. 수집한 데이터 집합으로 전처리를 진행하였다. 기술통계분석으로 최댓값, 최솟값을 구하여 극단치 여부를 확인하였고, 극단치

가 존재하는 지표의 경우 제외하였다. 결측치의 경우 실적이 없어 null값이면 0으로 입력하였다. 공시 지표 중 비율로 산출이 가능한 경우에는 비율데이터로 변환하였으며, 비슷한 정보를 가진 지표의 경우 데이터를 축소하기 위해 하나의 지표로 병합하였다. 또한 공시자료가 본교, 제2캠퍼스, 제3캠퍼스로 분리된 대학(17개교)의 경우 본교에 통합하였다.

3.4 변수의 선정

3.4.1 대학 교육/연구성과 영향요인 분석을 위한 변수 선정

교육/연구성과 영향요인 분석을 위해 선행연구, 논리적 타당성, 독립변수들 간의 상관성을 토대로 변수를 선정하였다. 교육/연구성과 영향요인 분석에 사용되는 독립변수와 종속변수를 정리하면 <표 1>과 같다.

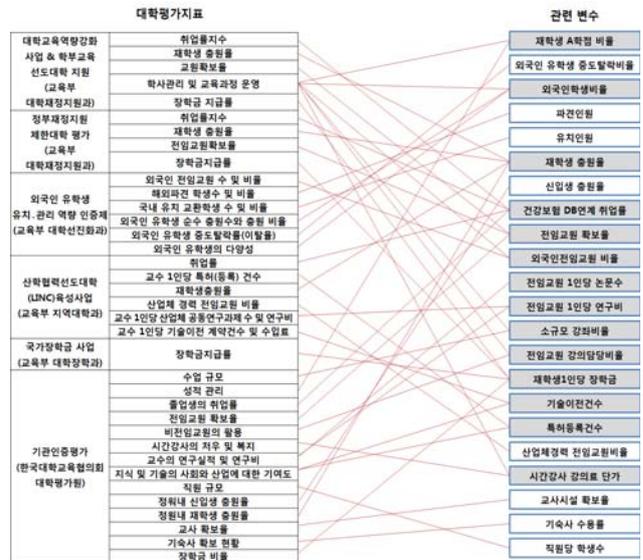
<표 1> 독립변수와 종속변수

독립변수		종속변수
학생	졸업생평점포준환산점수, 현장학습 이수학생 비율, 캡스톤디자인 이수학생 비율	건강보험 DB연계 취업률
교원	전임교원 1인당 논문수, 전임교원 1인당 연구비, 기술이전건수, 특허등록건수	
대학여건	신입생충원율, 재학생충원율, 전체교원대비 외국인전임교원비율, 산업체경력 전임교원비율, 전임교원 1인당 학생수(편제정원기준), 전임교원강의담당비율, 재정지원사업금액	기술이전건수
교원	전임교원 1인당 연구비, 전임교원 1인당 논문수, 특허등록건수	
대학여건	신입생충원율, 재학생충원율, 전임교원 1인당 학생수(편제정원기준), 전임교원확보율(편제정원기준), 산업체경력 전임교원비율, 전임교원 강의담당비율, 교사시설확보율, 재정지원사업금액	전임교원 1인당 논문수
교원	전임교원 1인당 연구비, 기술이전건수, 특허등록건수	

3.4.2 대학평가지표를 이용한 군집 분석을 위한 변수 선정

군집분석에 쓰일 변수의 선정은 <그림 2>와 같이 각종 대학평가에 사용되는 공시지표 중 주요 지표인 재학생 A학점 비율, 외국인학생비율, 재학생충원율, 취업률, 전임교원확보율, 외국인전

임교원비율, 전임교원 1인당 논문수, 전임교원 1인당 연구비, 소규모강좌비율, 전임교원 강의담당비율, 재학생 1인당 장학금, 기술이전건수, 특허등록건수, 시간강사강의료단가를 최종적으로 선택(14개 변수)하였다. 평가지표 중 공시자료를 활용하여 산출하지 않는 것은 제외하였으며, 본 연구의 변수는 대학정보공시 기준이므로 각 평가에서 산출하는 식과는 상이한 점이 있다.



<그림 2> 군집분석을 위한 대학평가 관련 변수 선정

3.5 자료의 분석방법

대학 교육/연구성과(취업률, 기술이전건수, 전임교원 1인당 논문수)의 영향요인을 탐색하기 위해 상관관계 분석을 실시하였고, 데이터마이닝 기법으로 회귀분석, 특징 선택(Feature selection), 분류분석(의사결정나무)을 실시하였다. 상관관계 분석은 SPSS statistics 19로 Pearson 상관계수를 이용하여 각 변수들 간의 상관관계를 살펴보았다.

회귀분석은 SPSS statistics 19, WEKA 3.6, Rattle 2.6.26(R Data Miner)을 이용하여 다중회귀분석을 실시하였다. SPSS에서는 단계선택방법으로 진행하였고, WEKA에서는 Classifier - functions - LinearRegression을 이용하였으며, Rattle에서는 Model- Linear로 진행하였다. 특징선택은 WEKA로 필터기법을 사용하였고, CFS알고리즘을 선택

하였다. AttributeSeletion - CfsSubsetEval 에서 검색방법으로는 BestFirst, GeneticSearch, GreedyStepwise, LinearForwardSelection, ScatterSearchV1로 각각 수행하고 결과를 비교해보았다. 분류분석은 의사결정나무방법으로 Classifier - trees - J48(C4.5) 알고리즘을 사용하였다. 설명력이 낮은 변수들을 제거하고, minNumObj값을 조정하면서 반복 수행하였고, 모형의 타당성 검사를 위해 test options 에서 Use training set, 10 fold cross validation, Percentage split 90%의 3가지 방법으로 평가하였다.

대학평가에 쓰이는 주요 지표를 이용하여 대학 간에는 어떤 유사성이 있는지 파악하기 위하여 SPSS로 k-means 군집분석을 실시하였다. 이 때 0과 1 사이의 값으로 표준화하여 변환된 값을 분석에 사용하였다. 그리고 군집형성에 영향을 미치는 요인과 군집의 특징을 알아보기 위하여 군집 번호를 목표변수로 하여 WEKA에서 분류분석(의사결정나무)을 실시하였다.

자료의 분석 방법, 도구, 옵션을 정리하면 <표 2>와 같다.

<표 2> 자료의 분석 방법

목적	분석 방법	분석 도구 및 옵션
대학 교육/연구성과 영향요인 추출	상관 분석	· SPSS : Pearson상관계수
	회귀 분석	· SPSS : 선형회귀분석 - 단계선택방법 · WEKA : Classifier - functions - LinearRegression · Rattle : Model - Linear
	특징 선택	· WEKA : AttributeSeletion - CfsSubsetEval - 검색방법 : BestFirst, GeneticSearch, GreedyStepwise, LinearForwardSelection, ScatterSearchV1
	분류 분석	· WEKA : Classifier - trees - J48 (C4.5) 알고리즘
대학평가지표를 이용한 대학의 군집유형, 특징, 군집형성 영향요인 분석	군집 분석	· SPSS : K-means
	분류 분석	· WEKA : Classifier - trees - J48 (C4.5) 알고리즘

4. 연구 결과

4.1 대학 교육/연구성과 영향요인 분석

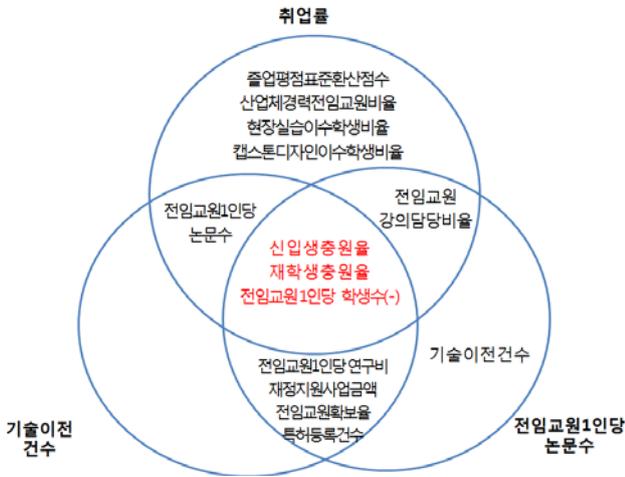
취업률에 영향을 미치는 요인은 분석방법별로 약간 상이하나, 정원내 신입생충원율, 전임교원 강의담당비율, 산업체경력 전임교원비율, 현장실습 이수학생비율, 재학생충원율, 전임교원 1인당 논문수, 전임교원 1인당 학생수(-), 캡스톤디자인 이수학생비율, 졸업생평점 표준환산점수로 나타났다. 즉 학생역량, 교수역량, 대학여건 지표가 취업률에 영향을 미치는 것을 알 수 있다.

기술이전건수에 영향을 미치는 요인은 분석방법별로 약간의 차이가 존재하나 전임교원 1인당 논문수, 재정지원사업금액, 특허등록건수, 전임교원 1인당 연구비, 정원내 신입생충원율, 재학생충원율, 전임교원 1인당 학생수(-), 전임교원확보율로 나타났다. 즉 교수역량과 대학여건 지표가 기술이전건수에 영향을 미치는 것을 알 수 있다.

전임교원 1인당 논문수의 영향요인은 분석방법별로 차이를 보이나, 전임교원 1인당 연구비, 기술이전건수, 재학생충원율, 특허등록건수, 정원내 신입생충원율, 전임교원확보율, 전임교원 강의담당비율, 재정지원사업금액, 전임교원 1인당 학생수(-)로 나타났다. 즉 교수역량과 대학여건 지표가 전임교원 1인당 논문수에 영향을 미치는 것으로 나타났다.

<표 3> 분석방법별 대학 교육/연구성과의 영향요인

분석 방법	취업률	기술이전건수	전임교원 1인당 논문수
상관 분석	정원내 신입생충원율 재학생충원율 전임교원 1인당 논문수 전임교원 강의담당비율 산업체경력전임교원비율 현장실습이수학생비율 캡스톤디자인이수학생비율	전임교원 1인당 논문수 전임교원 1인당 학생수(-) 전임교원 1인당 연구비 특허등록건수 정원내 신입생충원율 재학생충원율 전임교원확보율 재정지원사업금액	전임교원 1인당 학생수(-) 전임교원 1인당 연구비 기술이전건수 특허등록건수 정원내 신입생충원율 재학생충원율 재정지원사업금액
회귀 분석	정원내 신입생충원율 전임교원강의담당비율 재학생충원율	재정지원사업금액 전임교원 1인당 논문수 정원내 신입생충원율	재학생충원율 기술이전건수 전임교원 1인당 연구비 전임교원 강의담당비율
특징 선택	정원내 신입생충원율 현장실습 이수학생비율 산업체경력 전임교원비율 전임교원 1인당 논문수	재학생충원율 특허등록건수 전임교원 1인당 논문수	재학생충원율 기술이전건수
분류 분석 (의사 결정 나무)	정원내 신입생충원율 전임교원 강의담당비율 전임교원 1인당 학생수(-)	특허등록건수 재정지원사업금액 전임교원 1인당 연구비 전임교원 1인당 논문수	전임교원 1인당 연구비 정원내 신입생충원율 특허등록건수 전임교원확보율



<그림 3> 대학 교육/연구성과 영향요인 종합

종합적으로 대학의 교육/연구성과에 모두 영향을 미치는 요인은 <그림 3>과 같이 정원내 신입생충원율, 재학생충원율, 전임교원 1인당 학생수(-)임을 확인하였다.

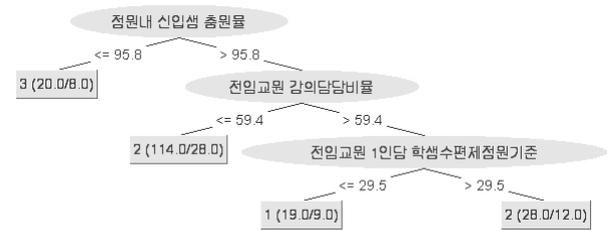
정원내 신입생충원율, 재학생충원율과 같은 지표가 교육/연구성과에 중요한 영향요인으로 나타난 것은 학생 충원율이 좋은 학교, 안정된 학교의 조건임에 기인한다.

4.2 의사결정나무 분석으로부터 추출된 규칙

4.2.1 취업률에 대한 의사결정나무 분석

연속형 변수인 취업률 지표에 대해 상위 20% 대학을 1, 하위 20% 대학을 3, 그 외 대학을 2로 변환하여 분석하였다.

목표변수를 취업률로 하고, 10개의 독립변수(졸업생 평점표준환산점수, 정원내 신입생충원율, 재학생충원율, 전임교원확보율, 전임교원 1인당 연구비, 전임교원 강의담당비율, 재정지원사업금액, 기술이전건수, 전임교원 1인당 논문수, 전임교원 1인당 학생수)로 분류기를 실행하였을 때(binarySplits = True, confidenceFactor = 0.25, minNumObj = 15) <그림 4>와 같은 나무의 크기는 7, 잎의 수는 4인 의사결정나무가 형성되었다.



<그림 4> 취업률에 대한 의사결정나무

※ 1 : 취업률 상위 20% 대학, 2 : 취업률 상, 하위 20% 외 대학, 3 : 취업률 하위 20% 대학

취업률에 영향을 미치는 요인은 정원내 신입생충원율, 전임교원 강의담당비율, 전임교원 1인당 학생수 순으로 나타났다. 이 의사결정나무 모형의 결과를 IF-THEN 분류규칙으로 해석해보면 정원내 신입생충원율이 95.8%보다 크고, 전임교원강의담당비율이 59.4%보다 크고, 전임교원 1인당 학생수가 29.5명 이하이면 취업률 상위 20% 대학이 될 수 있지만, 전임교원 1인당 학생수가 29.5명보다 크면 취업률 상위 20% 대학에서 벗어남을 알 수 있었다. 정원내 신입생충원율이 95.8%이하일 경우 취업률 하위 20% 대학으로 분류되었다.

의사결정나무의 타당성은 3가지 옵션(Use training set, 10 fold cross validation, Percentage split 90%)에서 평가했을 때 55.6%-68.5%의 정확성을 보였다.

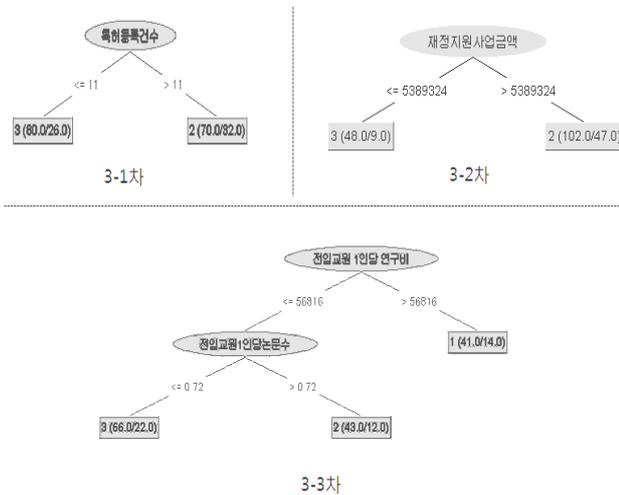
4.2.2 기술이전건수에 대한 의사결정나무 분석

연속형 변수인 기술이전건수 지표에 대해 상위 20% 대학을 1, 기술이전 건수가 0인 대학을 3, 그 외 대학을 2로 변환하였다.

1차 실험에서 목표변수를 기술이전건수로 하고, 11개의 독립변수(정원내 신입생충원율, 재학생충원율, 전임교원 1인당 학생수, 전임교원확보율, 전임교원 1인당 연구비, 전임교원 1인당 논문수, 전임교원 강의담당비율, 교사시설확보율, 재정지원사업금액, 산업체경력 전임교원비율, 특허등록건수)로 분류기를 실행하였을 때(binarySplits = True, confidenceFactor = 0.25, minNumObj = 50) 기술이전건수에 특허등록건수가 큰 영향을 미치는 것으로 나타났으며 이 때 의사결정나무의 타당성은 66.3-66.7%로 평가되었다. 특허등록건수

이외의 영향요인을 살펴보기 위하여 2차 실험에서 목표변수를 기술이전건수로 하고, 특허건수를 제외하여 분류기를 실행하였을 때(binarySplits = True, confidenceFactor = 0.5, minNumObj = 35) 재정지원사업금액, 전임교원 1인당 연구비가 기술이전건수의 영향요인으로 나타났다. 이 의사결정나무의 타당성은 70.2%-79.6%로 평가되었다.

3차 실험에서는 181개의 연구대상 중 이공계열 학과가 없는 예술, 체육, 종교지도자양성 대학(총 31개교)을 제외한 150개의 대학에 대해 동일한 방법으로 분석하였다.



<그림 5> 3차 실험 기술이전건수에 대한 의사결정나무 모형 (이공계열이 없는 대학 제외)

※ 1 : 기술이전건수 상위 20%대학, 2 : 기술이전건수는 존재하나 상위 20%에는 들지 못하는 대학, 3: 기술이전건수가 0인 대학

그 결과 <그림 5>를 보면 1차 실험에서와 같이 특허등록건수가 큰 영향요인으로 나타났고(3-1차), 특허건수를 제외하고 재실험하였을 때 재정지원사업금액이 큰 영향요인으로 나타났다(3-2차). 재정지원사업금액을 제외하고 다시 실험하였을 때 전임교원 1인당 연구비, 전임교원 1인당 논문수가 기술이전건수에 영향을 미치는 요인으로 나타났다(3-3차). 이공계열이 없는 예술, 체육, 종교지도자양성 대학을 포함하였을 때(1, 2차)와 포함하지 않았을 때(3차) 영향요인에 큰 차이가 없는 것을 알 수 있었다.

3-3차 의사결정나무 모형의 결과를 IF-THEN 분류규칙으로 해석해보면 전임교원 1인당 연구비가 56,816천원보다 크면 기술이전건수 상위 20% 대학으로 분류되었다. 전임교원 1인당 연구비가

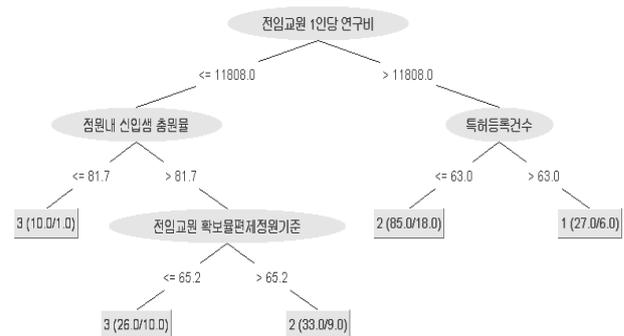
56,816천원 이하이고, 전임교원 1인당 논문수가 0.72건 이하이면 기술이전건수가 0인 대학으로 분류되었으나, 전임교원 1인당 논문수가 0.72건보다 크면 기술이전건수 상위 20% 대학에는 들지 못하더라도 기술이전건수가 존재하는 대학으로 분류되었다.

3차 실험 의사결정나무의 타당성은 3가지 옵션(Use training set, 10 fold cross validation, Percentage split 90%)에서 평가했을 때 53.3%-68%의 정확성을 보였다.

4.23 전임교원 1인당 논문수에 대한 의사결정나무 분석

연속형 변수인 전임교원 1인당 논문수 지표에 대해 전임교원 1인당 논문수 상위 20% 대학을 1, 하위 20% 대학을 3, 그 외 대학을 2로 변환하였다.

목표변수를 전임교원 1인당 논문수로 하고, 10개의 독립변수(정원내 신입생총원율, 재학생총원율, 전임교원확보율, 전체교원대비 외국인 전임교원비율, 전임교원 1인당 연구비, 전임교원 강의당비율, 특허등록건수, 교사시설확보율, 재정지원사업금액, 기술이전건수)로 분류기를 실행하였을 때(binarySplits = True, confidenceFactor = 0.25, minNumObj = 10) <그림 6>과 같은 나무의 크기는 9, 잎의 수는 5인 의사결정나무가 형성되었다. 이 의사결정나무 모형의 타당성을 평가하기 위해 3가지 옵션에서 평가했을 때 64.1%-75.7%의 정확성을 보였다.



<그림 6> 전임교원 1인당 논문수에 대한 의사결정나무 모형

※ 1 : 전임교원 1인당 논문수 상위 20% 대학, 2 : 전임교원 1인당 논문수 상, 하위 외 대학, 3 : 전임교원 1인당 논문수 하위 20% 대학

전임교원 1인당 논문수에 가장 영향을 미치는 요인은 전임교원 1인당 연구비로 나타났고, 정원

내 신입생충원율, 특허등록건수, 전임교원확보율이 전임교원 1인당 논문수의 영향요인으로 나타났다.

이 의사결정나무 모형의 결과를 IF-THEN 분류규칙으로 해석해보면 전임교원 1인당 연구비가 11,808천원보다 크고, 특허건수가 63건보다 크면 전임교원 1인당 논문수 상위 20% 대학으로 분류되었다. 전임교원 1인당 연구비가 11,808천원 이하이고, 정원내 신입생충원율이 81.7%이하이면 전임교원 1인당 논문수 하위 20%로 분류되었고, 정원내 신입생충원율이 81.7%보다 큰 경우에는 전임교원확보율이 65.2%보다 크면 하위 20% 대학에서 벗어날 수 있지만, 65.2%이하이면 하위 20% 대학으로 분류되는 것으로 나타났다.

4.3 대학평가지표를 이용한 대학의 군집유형 분석

4.3.1 전체 대학에 대한 군집 분석

전체 대학을 대상으로 하였을 때 <표 4>와 같이 4개의 군집이 형성되었다. 1군집은 중·소규모의 지방 사립대학, 2군집은 유명대학(대규모 수도권대학과 대규모 지방 국공립대학)으로, <표 5>를 보면 타군집에 비해 14개의 지표값 중 8개 지표 평균이 높았다. 3군집은 주로 소규모 사립의 종교지도자양성대학으로 14개의 지표 중 10개 지표 평균이 낮았으며, 4군집은 그 외 대학으로 나타났다.

<표 4> 대학평가 관련 변수를 이용한 대학의 군집유형과 그 특징

군집	대학 수	대학 특징	지표의 특징	의사결정나무로부터 추출된 특징
1	70	중·대규모 대학	타군집에 비해 재학생 1인당 장학금 평균이 낮음	기술이전건수가 27건 이하이고, 취업률이 38.9%보다 크며, 전임교원 1인당 논문수가 0.64건 이상인 대학일 확률 93.9%
2	21	유명대학 (대규모 수도권대학과 대규모 지방 국공립대학)	타군집에 비해 재학생충원율, 외국인학생비율, 전임교원확보율, 전임교원 1인당 논문수, 전임교원 1인당 연구비, 기술이전건수, 특허등록건수, 시간당 시간강사강의료단가 평균이 높고, 소규모장좌비율 평균이 낮음	기술이전건수가 27건보다 큰 대학일 확률 87.5%
3	69	중·소규모 대학	타군집에 비해 취업률, 외국인전임교원비율, 전임교원강의담당비율 평균이 높고, 재학생 A학점비율, 전임교원확보율 평균이 낮음	기술이전건수가 27건 이하이고, 취업률이 38.9%보다 크며, 전임교원 1인당 논문수가 0.64건 이하인 대학일 확률 87.8%
4	21	소규모 사립대학	타군집에 비해 재학생 A학점 비율, 소규모장좌비율, 재	기술이전건수가 27건 이하이고, 취업률

군집	대학 수	대학 특징	지표의 특징	의사결정나무로부터 추출된 특징
		(종교지도자 양성대학)	학생 1인당 장학금 평균이 높고, 재학생충원율, 외국인 학생비율, 취업률, 외국인전임교원비율, 전임교원 1인당 논문수, 전임교원 1인당 연구비, 전임교원강의담당비율, 기술이전건수, 특허등록건수, 시간당 시간강사강의료단가 평균이 낮음	이 38.9%이하인 대학일 확률 100%

<표 5> 군집유형별 변수의 평균값

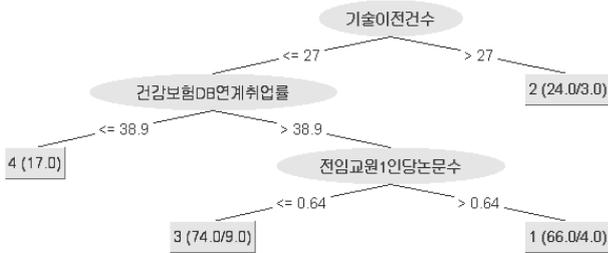
변수	전체대학 평균	군집유형별 평균			
		1군집	2군집	3군집	4군집
재학생 A학점 비율	37.3	36.3	39.5	36.0	42.4
재학생충원율	105.0	113.0	126.0	98.7	78.1
외국인학생비율	2.3	2.3	5.1	1.8	1.3
취업률	53.1	54.8	57.9	60.8	17.7
전임교원확보율	77.1	76.5	88.1	74.6	76.7
외국인전임교원비율	3.2	2.5	2.3	4.6	2.2
전임교원 1인당 논문수	0.7	0.9	1.1	0.5	0.3
전임교원 1인당 연구비(천원)	40,943	46,041	143,329	13,924	10,338
소규모장좌비율	40.2	36.1	33.9	40.4	59.5
전임교원강의담당비율	55.9	4.4	52.7	59.9	51.4
재학생 1인당 장학금(천원)	1,599	1,361	1,842	1,525	2,398
기술이전건수	9	8.5	44.8	1.4	0.1
특허등록건수	38.6	27.5	225.3	4.8	0.6
시간당시간강사강의료 단가(천원)	43.2	47.6	54.9	37.8	34.8

※ 푸른 음영 : 평균이 높은 군집, 붉은 음영 : 평균이 낮은 군집

4개의 군집을 이루는데 영향을 미치는 요인이 무엇인지 알아보기 위하여 군집유형을 목표변수로 하고, 대학평가 관련변수를 독립변수로 하여 분류분석(의사결정나무)을 실시하였다. 그 결과 <그림 7>과 같이 기술이전건수가 군집형성에 가장 큰 영향요인으로 나타났고, 건강보험 DB연계 취업률, 전임교원 1인당 논문수가 주요 영향요인으로 나타났다. 즉 대학 교육/연구성과 지표가 대학의 군집형성에 주요 요인으로 나타남을 확인하였다. 이 의사결정나무 모형은 10 fold cross validation 테스트에서 85.2%의 타당성을 보였다.

의사결정나무 모형으로부터 군집의 특징을 추출해보면 <표 4>와 같다. 1군집(중·대규모 대학)은 주로 기술이전건수가 27건 이하이고, 취업률이 38.9%보다 크며, 전임교원 1인당 논문수가 0.64건 이상인 대학이었다. 2군집(유명대학)은 대부분 기술이전건수가 27건보다 큰 대학이었고, 3군집(중·소규모 대학)은 기술이전건수가 27건 이

하이고, 취업률이 38.9%보다 크며, 전임교원 1인당 논문수가 0.64건 이하인 대학이었다. 4군집(종교지도자양성대학)은 주로 기술이전건수가 27건 이하이고, 취업률이 38.9%이하인 대학이었다.



<그림 7> 군집형성의 영향요인에 대한 의사결정나무

전체 대학에 대한 군집분석의 타당성을 검증하기 위해 판별분석을 실시한 결과 91.7%가 올바르게 분류되었음을 확인하였다.

4.3.2 대학 규모에 따른 군집 분석

대학 규모별 군집 분석에서는 <표 6>과 같이 대체로 유명대학, 종교지도자양성대학, 그 외 대학으로 3군집이 형성되는 양상을 보였다. 대규모 1군집은 주로 지방 사립대학으로 <표 7>을 보면 14개 지표 중 11개의 지표 평균이 타군집에 비해 낮게 나타났다. 대규모 2군집은 주요 국공립, 사립대학이었다. 대규모 3군집은 수도권 유명대학으로 14개 지표 중 12개의 지표 평균이 타군집에 비해 높게 나타났다. 중규모 1군집은 주로 지방의 사립대학으로 14개의 지표 평균 중 9개의 지표가 낮게 나타났다. 중규모 2군집은 수도권, 지방의 사립대학이었다. 중규모 3군집은 수도권 사립대학과 지방 국공립대학으로 14개 지표 중 8개의 지표가 타군집에 비해 높게 나타났다. 소규모 1군집은 수도권, 지방의 사립대학이었다. 소규모 2군집은 종교지도자양성대학으로 14개 지표평균 중 9개 지표가 타군집에 비해 낮게 나타났다. 소규모 3군집은 유명대학으로 14개의 지표 중 10개의 지표가 타군집에 비해 높게 나타났다.

<표 6> 대학평가 관련 변수를 이용한 대학규모별 군집 유형과 그 특징

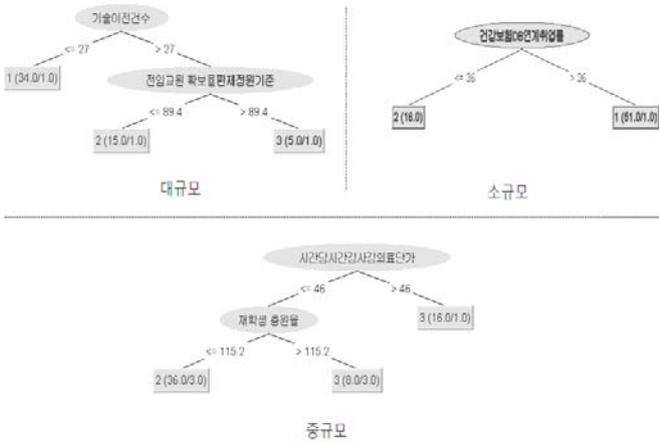
규모	군집	대학 수	대학 특징	지표의 특징	의사결정나무로부터 추출된 특징
----	----	------	-------	--------	------------------

규모	군집	대학 수	대학 특징	지표의 특징	의사결정나무로부터 추출된 특징
대규모	1	34	지방 사립대학	대규모 타군집에 비해 외국인전임교원비율, 전임교원강의담당비율 평균이 높고, 재학생 A학점 비율, 재학생충원율, 외국인학생비율, 취업률, 전임교원확보율, 전임교원 1인당 논문수, 전임교원 1인당 연구비, 재학생 1인당 장학금, 기술이전건수, 특허등록건수, 시간당 시간강사강의료단가 평균이 낮음	기술이전건수가 27건 이하인 대학일 확률 97%
	2	16	주요 대학	대규모 타군집에 비해 외국인전임교원비율, 소규모강좌비율, 전임교원강의담당비율 평균이 낮음	기술이전건수가 27건보다 크고, 전임교원확보율이 89.4%이하인 대학일 확률 93.3%
	3	4	수도권 유명대학	대규모 타군집에 비해 재학생 A학점 비율, 재학생충원율, 외국인학생비율, 취업률, 전임교원확보율, 전임교원 1인당 논문수, 전임교원 1인당 연구비, 소규모강좌비율, 재학생 1인당 장학금, 기술이전건수, 특허등록건수, 시간당 시간강사강의료단가 평균이 높음	기술이전건수가 27건보다 크고, 전임교원 확보율이 89.4%보다 큰 대학일 확률 80%
중규모	1	4	지방 사립대학	중규모 타군집에 비해 외국인학생비율, 취업률, 외국인전임교원비율, 소규모강좌비율 평균이 높고, 재학생 A학점 비율, 전임교원확보율, 전임교원 1인당 논문수, 전임교원 1인당 연구비, 전임교원강의담당비율, 재학생 1인당 장학금, 기술이전건수, 특허등록건수, 시간당 시간강사강의료단가 평균이 낮음	-
	2	36	수도권, 지방 사립대학	중규모 타군집에 비해 전임교원강의담당비율, 재학생 1인당 장학금 평균이 높고, 재학생충원율, 외국인학생비율 평균이 낮음	시간당 시간강사강의료 단가가 46천원 이하이고, 재학생충원율이 115.2%이하인 대학일 확률 91.7%
	3	20	수도권 사립, 지방 국공립대학	중규모 타군집에 비해 재학생 A학점 비율, 재학생충원율, 전임교원확보율, 전임교원 1인당 논문수, 전임교원 1인당 연구비, 기술이전건수, 특허등록건수, 시간당 시간강사강의료단가 평균이 높고, 취업률, 외국인전임교원비율, 소규모강좌비율 평균이 낮음	시간당 시간강사강의료 단가가 46천원 이하이고, 재학생충원율이 115.2%보다 큰 대학일 확률 62.5% 시간당 시간강사강의료단가가 46천원보다 큰 대학일 확률 93.8%
소규모	1	50	수도권, 지방 사립대학	소규모 타군집에 비해 외국인학생비율, 취업률, 전임교원강의담당비율 평균이 높고, 재학생 A학점 비율, 전임교원확보율, 소규모강좌비율, 재학생 1인당 장학금 평균이 낮음	취업률이 36%보다 큰 대학일 확률 98%
	2	16	종교 지도자양성대학	소규모 타군집에 비해 소규모강좌비율 평균이 높고, 재학생충원율, 취업률, 외국인전임교원비율, 전임교원 1인당 논문수, 전임교원 1인당 연구비, 전임교원강의담당비율, 기술이전건수, 특허등록건수, 시간당 시간강사강의료 단가 평균이 낮음	취업률이 36%이하인 대학일 확률 100%
	3	1	유명대학	소규모 타군집에 비해 재학생 A학점 비율, 재학생충원율, 전임교원확보율, 외국인전임교원비율, 전임교원 1인당 논문수, 전임교원 1인당 연구비, 재학생 1인당 장학금, 기술이전건수, 특허등록건수, 시간당 시간강사강의료 단가 평균이 높음	-

<표 7> 대학 규모에 따른 군집유형별 변수의 평균값

대학평가 관련 변수	중규모								
	대규모			중규모			소규모		
	1군집	2군집	3군집	1군집	2군집	3군집	1군집	2군집	3군집
재학생A학점비율	34.4	38.5	41.7	31.2	36.6	37.1	37.9	41.8	50.9
재학생총원율	113.1	123.9	134.4	104.9	104.6	116.4	96.2	75.5	115.9
외국인학생비율	2.0	4.1	8.6	3.6	1.9	2.7	1.8	1.1	1.1
취업률	53.8	55.5	65.4	64.4	57.8	56.4	60.3	6.7	50.5
전임교원확보율	76.4	80.2	109.0	64.8	73.4	83.7	72.8	81.2	134.5
외국인전임교원비율	3.3	2.1	2.4	12.2	4.3	1.6	2.9	2.7	4.1
전임교원 1인당 논문수	0.8	1.1	1.2	0.5	0.6	0.9	0.5	0.3	1.4
전임교원 1인당 연구비(천원)	43046	88918	171855	10445	19753	68530	14110	13167	79660
소규모강좌비율	35.5	32.2	38.4	43.8	36.2	34.8	45.6	56.0	47.2
전임교원강의 담당비율	55.7	51.5	55.6	48.5	58.4	53.3	58.4	53.0	57.3
재학생 1인당 장학금(천원)	13792	14388	23829	12045	14010	13820	16698	23362	53893
기술이전건수	11.0	38.4	66.0	0.8	2.8	10.7	0.5	0.2	40.0
특허등록건수	35.0	163.5	463.0	4.0	7.2	32.3	2.0	0.8	302.0
시간당시간강사 강의료단가(천원)	44.2	54.6	56.8	34.8	40.4	54.2	38.2	36.5	52.0

* 푸른 음영 : 평균이 높은 군집, 붉은 음영 : 평균이 낮은 군집



<그림 8> 규모별 군집형성 영향요인 의사결정나무

<그림 8>을 보면 대규모대학 군집형성의 영향요인은 기술이전건수, 전임교원확보율로 나타났고, 중규모대학 군집형성의 영향요인은 시간당 시간강사강의료단가, 재학생총원율로 나타났으며, 소규모 대학 군집형성의 영향요인은 취업률로 나타났다.

4.3.3 소재지에 따른 군집분석

소재지별 군집 분석에서는 <표 8>과 같이 주로 유명대학, 종교지도자양성 대학, 그 외 대학으로 3 군집이 형성되었다. 수도권 1군집은 주로 소규모

사립의 종교지도자 양성대학으로, <표 9>를 보면 타군집에 비해 평균값이 낮은 지표가 14개 중 10개로 나타났다. 수도권 2군집은 대, 중, 소 사립대학이 혼재되어 소속되었다. 수도권 3군집은 대규모 유명 사립대학으로 타군집에 비해 14개 지표 중 11개 지표가 높게 나타났다.

또한 지방 1군집은 대, 중, 소 사립대학이 혼재되어 소속되었고, 지방 2군집은 외국어, 예술, 신학대학 등 소규모 사립대학이었다. 지방 3군집은 주로 중, 대규모 국공립대학으로 타군집에 비해 14개의 지표 중 7개 지표 평균이 높게 나타났다.

<표 8> 대학평가 관련 변수를 이용한 소재지별 군집 유형과 그 특징

규모	군집	대학 수	대학 특징	지표의 특징	의사결정나무로부터 추출된 특징
수도권	1	13	소규모 사립의 종교 지도자 양성대학	수도권 타군집에 비해 재학생 A학점 비율, 소규모강좌비율 평균이 높고, 재학생총원율, 외국인학생비율, 취업률, 외국인전임교원비율, 전임교원 1인당 논문수, 전임교원 1인당 연구비, 전임교원강의담당비율, 기술이전건수, 특허등록건수, 시간당 시간강사강의료단가 평균이 낮음	기술이전건수가 22건 이하이고, 전임교원 1인당 논문수가 0.33건 이하인 대학일 확률 100%
	2	43	대, 중, 소 사립대학	수도권 타군집에 비해 외국인전임교원비율 평균이 높고, 재학생A학점비율, 전임교원확보율, 재학생 1인당 장학금 평균이 낮음	기술이전건수가 22건 이하이고, 전임교원 1인당 논문수가 0.33건보다 큰 대학일 확률 95.6%
	3	13	유명 사립대학	수도권 타군집에 비해 재학생총원율, 외국인학생비율, 취업률, 전임교원확보율, 전임교원 1인당 논문수, 전임교원 1인당 연구비, 전임교원강의담당비율, 재학생 1인당 장학금, 기술이전건수, 특허등록건수, 시간당 시간강사강의료단가 평균이 높음	기술이전건수가 22건보다 큰 대학일 확률 100%
지방	1	85	대, 중, 소 사립대학	지방 타군집에 비해 취업률, 외국인전임교원비율, 전임교원강의담당비율 평균이 높고, 재학생A학점비율, 외국인학생비율, 전임교원확보율, 재학생1인당장학금 평균이 낮음	특허등록건수가 73건 이하이고, 소규모강좌비율이 61.6% 이하이고, 기술이전건수가 22건 이하이고, 취업률이 43.1%보다 큰 대학일 확률 100%
	2	14	소규모 사립대학 (외국어, 예술, 신학)	지방 타군집에 비해 재학생A학점비율, 전임교원확보율, 소규모강좌비율, 재학생1인당장학금 평균이 높고, 재학생총원율, 취업률, 전임교원1인당논문수, 전임교원1인당연구비, 기술이전건수, 특허등록건수, 시간당시간강사강의료단가 평균이 낮음	특허등록건수가 73건 이하이고, 소규모강좌비율이 61.6%보다 큰 대학일 확률 100%

규모	군집	대학 수	대학 특징	지표의 특징	의사결정나무로부터 추출된 특징
	3	13	중,대규모 국공립 대학	지방 타군집에 비해 재학생총원율, 외국인학생비율, 전임교원1인당논문수, 전임교원1인당연구비, 기술이전건수, 특허등록건수, 시간당시간강사강의료단가 평균이 높고, 외국인전임교원비율, 소규모강좌비율, 전임교원강의담당비율 평균이 낮음	특허등록건수가 73건보다 큰 대학일 확률 100% 특허등록건수가 73건 이하이고, 소규모 강좌 비율이 61.6%이하이고, 기술이전건수가 22건보다 큰 대학일 확률 60%

<그림 9>를 보면 수도권대학 군집형성의 영향 요인은 기술이전건수와 전임교원 1인당 논문수, 지방대학 군집형성의 영향요인은 특허건수, 소규모강좌비율, 기술이전건수, 취업률 순으로 나타났다.

5. 결론 및 제언

대학정보공시 자료의 데이터마이닝 결과를 통해 도출할 수 있는 결론 및 시사점은 다음과 같다.

첫째, 대학에서는 교육/연구성과 지표의 개선 시 학생총원율을 높이고, 전임교원 1인당 학생수를 낮추는 방안에 대해 고려해볼 필요가 있다. 구체적으로 취업률 지표에 신입생총원율이 여러 가지 분석 방법에서 공통적으로 나타난 요인이었으며, 이는 좋은 학교, 안정된 학교가 되기 위한 기본 조건이기 때문이다. 그리고 취업률의 영향요인으로 나타난 현장실습 이수학생비율, 산업체경력 전임교원비율, 캡스톤디자인 이수학생비율 지표의 공통점은 산학협력과 관련된다. 대학이 산학협력을 강화할 필요성이 있다. 연구성과를 높이기 위해서는 연구비, 재정지원사업금액 등의 재정적인 확보가 중요하며, 대학정보공시 자료에서 연구비, 재정지원사업금액, 특허건수, 기술이전건수, 논문실적은 특히 연관성이 높은 것으로 나타나 이 지표들의 상관성을 고려하여 관리한다면 연구성과 지표 향상에 도움이 될 것이다.

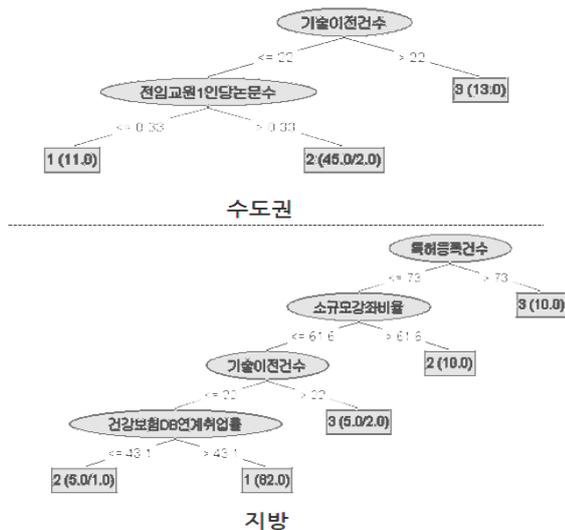
또한 의사결정나무 모형으로 지표별 상위그룹에 포함되기 위한 분류기준치를 확인할 수 있을 것이며, 군집분석에 나타난 결과를 통해 유사한 그룹 내의 평균과 다른 그룹 간의 평균값 비교가 가능하다. 또한 낮은 지표값, 경쟁력을 갖고 있는 지표값이 무엇인지 발견할 수 있고, 이에 따라 지표 개선 전략을 수립할 수 있을 것이다.

둘째, 총괄관리기관에서는 공시자료의 신뢰성을 높이기 위한 방안으로 데이터마이닝의 기법을 활용하여 자료의 극단치, 결측치 등을 확인할 수 있을 것이다. 그리고 빅데이터 시대의 흐름에 맞추어 데이터를 보유, 관리, 서비스하는 것에서 더욱 발전하여 누적되는 데이터의 분석을 통한 미래 예측이나 의미 있는 지식을 도출하여 대학에 제공하는 역할을 한다면 고등교육 발전에 기여할 수 있을 것이다.

<표 9> 소재지에 따른 군집유형별 변수의 평균값

대학평가 관련 변수	수도권			지방		
	1군집	2군집	3군집	1군집	2군집	3군집
재학생 A학점 비율	45.9	37.1	38.9	35.1	40.8	38.4
재학생총원율	91.8	115.2	131.7	101.9	70.8	115.4
외국인학생비율	0.3	2.5	6.9	1.9	2.0	2.2
취업률	33.9	54.3	61.2	59.5	22.5	51.4
전임교원확보율	79.3	72.2	89.8	74.4	91.0	81.9
외국인전임교원비율	2.1	3.4	3.2	3.7	2.6	0.9
전임교원 1인당 논문수	0.2	0.8	1.1	0.6	0.3	1.0
전임교원 1인당 연구비 (천원)	5,669	36,707	124,503	24,706	17,059	138,547
소규모강좌비율	41.7	40.1	33.1	37.4	66.9	35.3
전임교원 강의담당비율	44.3	51.8	52.7	60.6	56.2	54.2
재학생 1인당 장학금 (천원)	1,691.0	1,353.6	1,889.5	1,432.7	3,042.3	1,568.3
기술이전건수	0.0	4.0	47.0	5.0	0.0	36.0
특허등록건수	1.0	18.0	261.0	13.0	1.0	133.0
시간당시간강사강의료단가 (천원)	40.0	41.0	53.0	42.0	34.0	59.0

※ 푸른 음영 : 평균이 높은 군집, 붉은 음영 : 평균이 낮은 군집



<그림 9> 소재지별 군집형성 영향요인 의사결정나무

셋째, 정부 및 각 대학평가기관에서는 평가 시 국공립/사립, 교육/산업/일반대학, 학교규모라는 구분 이외에 비이공계·예체능·종교지도자양성 대학의 군집에 대한 별도의 재정지원사업이나 평가 정책 수립에 대해 고려해 볼 필요성이 있다.

앞으로 대학은 학령인구 감소에 따른 구조조정이 불가피한 시대에 경쟁력을 갖추기 위해서 통합적이고 체계적인 지표 관리가 필수적이다. 각종 평가에 대비한 지표 개선을 위한 의사 결정 및 전략 수립 시 본 연구의 결과를 참고할 수 있을 것이다.

이 연구의 결과가 단순히 지표값을 높이는 데에 활용되는 것이 아니라 실질적인 교육성과와 연구성과를 창출하고, 그 질을 제고하기 위한 지표 개선에 도움이 되는 지식으로 활용되기를 기대해본다.

참 고 문 헌

- [1] 배상훈·고석규·최석식 (2013) **대학구조개혁 토론회 자료**. 한국대학교육협의회.
- [2] 교육부 (2013). 14학년도 정부재정지원제한대학 평가결과.
- [3] 심인선·정성희 (2007). 대학정보공시제와 대학의 발전과제. 경남정책 Brief 제9호, 1-8, 경남발전연구원.
- [4] 김순우·조병량·한상필 (2010). 대학평가가 대학의 이미지 변화에 미치는 영향에 관한 연구. **언론과학연구**, 10(2), 139-178.
- [5] 백인수·박지혜 (2013). **데이터 시대 : 데이터 분석의 중요성**. IT&Future Strategy 제9호. 한국정보화진흥원.
- [6] 한국정보화진흥원 (2013). **새로운 미래를 여는 빅데이터 시대**. 한국정보화진흥원.
- [7] Witten, Ian H., Frank, Eibe, Holmes, Geoffrey(2013). **데이터마이닝 Data Mining 데이터 속 숨은 의미를 찾는 기계 학습의 이론과 응용**(이승현 역). 경기: 에이콘출판주식회사. (원서출판 2011).
- [8] 교육부, 한국대학교육협의회 (2013), 10월공시 보도자료.
- [9] 유현숙·임후남 외 (2010). **2009 대학 공시 정보 분석자료집**. 통계자료 SM 2010-02, 한국교육개발원.
- [10] 박창이·김용대·김진석·송종우·최호식 (2011). **R을 이용한 데이터마이닝**. 서울: 교우사.
- [11] 이학식·임지훈 (2013). **SPSS 20.0 매뉴얼**. 서울: 집현재.
- [12] 김완섭·이수원 (2006). 데이터마이닝을 이용한 설문조사의 심층 분석. **한국공학교육학회**, 9(4), 71-82.
- [13] 김정호·인주호·채수환 (2012). 의미 기반 유전 알고리즘을 사용한 특징 선택. **한국인터넷정보학회**, 13(4), 1-10.
- [14] Liu, H. and Yu, L. (2005). Toward Integrating Feature Selection Algorithms for Classification and Clustering. **IEEE Transactions on Knowledge and Data Engineering**, 17(4), 491-502.
- [15] 서민원·배성근 (2012). 대학교육역량 평가지표의 요인구조와 대학의 군집유형 분석. **교육평가연구**, 25(1), 117-144.
- [16] 교육부, 한국대학교육협의회 (2013), **대학정보공시 지침서**.
- [17] 임후남 (2009). **교육통계 및 공시정보의 관리·활용 방안 연구 : 고등교육을 중심으로**. 연구보고 RR 2009-29, 한국교육개발원.
- [18] 교육과학기술부 (2012). 2012년도 외국인유학생 인증제 공고문.
- [19] 교육과학기술부 (2012). 산학협력 선도대학(LINC) 육성사업 기본계획.
- [20] 교육부 (2013). 2013년 국가장학금사업 기본계획.
- [21] 교육부(2013). 대학역량강화사업 기본계획.
- [22] 한국대학교육협의회 (2013), **2013 대학기관 평가인증 편람**. RM 2012-33-532.
- [23] 조장식 (2011). 학생정보를 이용한 대졸 취업에 미치는 영향력 분석. **한국데이터정보과학회지**, 22(5), 849-856.
- [24] 채구묵 (2007). 신규대졸자의 취업 및 임금수준 결정요인 분석. **한국사회복지학**, 59, 35-61.
- [25] 정태영·이기엽 (2005). 대졸자 취업결정요인에 관한 연구 - 학점이 취업확률에 미치는 영향을 중심으로. **한국경영학회**, 8(2), 159-184.
- [26] 황여정·백병부(2008). 대졸 청년층의 노동시장 성과 결정요인. **직업능력개발연구**. 11(2), 1-23.

- [27] 채창균·김태기(2009). 대졸 청년층의 취업 성과 결정 요인 분석. **직업교육연구**, 28(2), 89-107.
- [28] 김안국·유한구·한상근·이영대·장수명 (2010). **대졸자의 노동시장 이행 연구**. 한국직업능력개발원.
- [29] 이병식·최정운 (2009). 대학졸업자의 취업성과에 대한 개인 및 대학특성 영향요인. **한국교육**, 36(1), 191-215.
- [30] 김화영 (2013). **DEA와 MPI를 이용한 대학의 효율성 분석 : 2008-2012년 대학정보공시자료를 중심으로**. 박사학위논문, 이화여자대학교 대학원.
- [31] 김정일 (2013). **대학교육역량의 성과 영향요인에 관한 연구**. 박사학위논문, 중앙대학교 대학원.
- [32] 이창학·이철규·이동명 (2011). 대학의 기술이전성과 영향요인 분석. **Journal of the Korea Academia-Industrial Cooperation Society**, 12(9), 3936-3951.
- [33] 김정진 (2010). **대학기술이전 성과에 미치는 영향요인에 관한 실증연구**. 박사학위논문, 부산대학교 대학원.
- [34] 소병우(2009). **대학의 지식재산경영이 기술이전 성과에 미치는 영향에 관한 실증연구**. 박사학위논문, 호서대학교 대학원.
- [35] 김철희·이상돈(2007). 산학협력성과와 대학의 역량요인의 관계에 관한 연구. **기술혁신학회지**, 10(4), 629-653.
- [36] 교육부·한국대학교육협의회 (2013), **2013년도 상반기 대학정보공시 Issue Papers**.
- [37] 김훈호·박환보 (2011). 교수의 개인 특성과 대학의 연구 환경이 연구 성과에 미치는 영향. **한국교육행정학회**, 29(2), 185-210.
- [38] 이정미·김민희(2010). 대학재정의 효율성 분석. **교육행정학연구**, 28(4), 163-186.
- [39] 박해정·김종태 (2013). 서포트벡터 군집분석을 이용한 대구·경북지역 대학의 분류. **한국데이터정보과학회지**, 24(4), 783-791.



신 선 미

2010 인천대학교 컴퓨터공학과 (공학사)
 2011~현재 고려대학교 교육대학원
 교육정보전공 석사과정
 2010 한국교육개발원 대학정보공시센터 팀원
 2011~2012 한국대학교육협의회
 대학정보공시센터 전문원
 관심분야: 교육정보화, 고등교육, 데이터마이닝
 E-Mail: sunmi1338@korea.ac.kr



김 현 철

1988 고려대학교 전산학과 (이학사)
 1990 Univ. of Missouri-Rolla, (전산학 석사)
 1998 Univ. of Florida (전산정보학 박사)
 1999~현재 고려대학교 사범대 컴퓨터교육과 교수
 2005~2006 Univ. of Florida 대우부교수
 2010 홋카이도대학 특임교수
 2012 아리조나주립대학 학습과학연구소 방문교수
 관심분야: 컴퓨터교육, 스마트러닝, 기계학습
 E-Mail: harrykim@korea.ac.kr