

데이터마이닝 기법을 활용한 대졸자 취업요인 분석

이재 화 (계명대 학교 경영대학 경영정보학과 석사과정, 주저자 gampo@nate.com)

이충 권 (계명대 학교 경영대학 경영정보학과 조교수, 교신저자 ckleee@kmu.ac.kr)

An analysis of college graduates' employment factors by using data mining

*Jae-Hwa Lee (the master's course, Dept. of Management Information
Systems, College of Business, Keimyung University)*

*Choong-Kwon Lee (Assistant Professor, Dept. of Management Information Systems,
College of Business, Keimyung University)*

-원고매수: 13 페이지

[교신저자 연락처]

© 이충권

- ◆ 주소 : 대구광역시 달서구 달구벌대로 1095 계명대학교 경영대학 경영정보학과
- ◆ 전화번호: 053-580-6416
- ◆ 휴대폰: 010-5580-3713
- ◆ E-mail주소: gampo@nate.com

데이터마이닝 기법을 활용한 대졸자 취업요인 분석

An analysis of college graduates' employment factors
by using data mining

• 목차 •

- | | |
|------------|-----------|
| I. 서론 | IV. 연구 결과 |
| II. 관련연구 | V. 결론 |
| III. 연구 방법 | 참고문헌 |
-

… Abstract …

The unemployment crisis of college graduates is being exacerbated and becoming a social issue, even though they can get support by various government policies on employment. College students spend a lot of time and expense to build up qualifications for job acquisition. Colleges offer supporting policies and diverse programs for helping them reinforce employment competency. However, there are not enough empirical studies to demonstrate that employment supporting programs have a positive effect on college students' job acquisition apparently. This study inquires the effectiveness of measurement factors evaluating students' competency from a program implemented in K university in Daegu. We used decision tree model and logistic regression analysis among data mining models to analyze the influence of measurement factors on job acquisition.

Key Words : Data Mining, Employment factor, Employment rate, University graduates

I. 서론

대학교 졸업자들의 취업난은 시간이 흐를수록 심화되어 가고 있고 사회적으로도 많은 관심을 받는 문제가 되었다. 통계청에 따르면, 2009년에 68.2%였던 4년제 대학교 졸업자의 취업률은 2014년에 54.8%까지 하락했다. 이는 불과 5년 만에 약 14%가 하락한 수치이고 대졸자 미취업의 심각성은 매년 심화되고 있다. 대졸자 뿐만 아니라 청년층의 실업률은 2014년 10%를 돌파하였고, 2016년에는 12.5%로 역대 최고치를 경신하였다. 이러한 상황 속에서 대학생들은 높아진 취업의 문턱을 넘기 위하여 전공이나 직무에 대한 역량 강화보다는 학점, 어학, 자격증 등 자신들의 ‘스펙 관리’에 대부분의 시간을 쏟고 있으며, 각종 구직활동에 다양한 노력을 기울이고 있다(이만기, 김효원, 2015).

정부나 대학 차원에서도 모두 대졸자 취업률에 촉각을 곤두세우고 있으며, 다양한 정책 및 프로그램 등을 통하여 취업률 제고에 힘쓰고 있는 상황이다. 정부는 대학의 운영에 있어서 투명한 정보공시를 통해 교육의 질을 높이고 정확한 정보를 전달하기 위하여 2007년부터 대학정보공시제를 도입했다. 교육부에서는 학력인구 급감에 대비하고, 대학 교육의 질을 제고하기 위하여 ‘대학 구조개혁 평가’를 발표하였다(최영욱, 2015). 같은 이유로 정부뿐만 아니라 전국의 대학교들은 취업역량강화 프로그램과 취업지원 활동 등 취업률 향상에 노력하고 있다. 하지만 이러한 프로그램들의 실효성과 취업 역량을 측정하는 항목들에 대한 검증은 이루어지지 않고 있다. 대학생들은 정부나 학교 측의 지원프로그램에 치중하여 구체적이고 객관적인 자료나 근거 없이 무분별한 스펙 쌓기에 시간과 노력을 낭비하고 있는 실정이다.

따라서 본 연구에서는 대구에 위치한 K대학교 IT관련 학과의 2010과 2015년 사이에 졸업한 학생들을 대상으로 학점과 재학 중에 누적한 취업역량강화 프로그램 참여를 분석하였다. 이를 통해 대졸자의 취업역량 측정 항목들이 취업에 실질적으로 어떠한 영향을 미치는지 알아보고, 취업에 가장 많은 영향을 미치는 핵심 요인들은 어떠한 것들인지 파악하고자 한다.

II. 관련 연구

대학생들의 취업에 영향을 미치는 요인에 관한 다양한 연구가 있었다. 박순영(2015)은 소위 얘기하는 ‘스펙’이라는 것을 쌓기 위하여 투입되는 비용과 얻어지는 보상이

무엇인지, 대학생들이 주로 어떠한 스펙들을 쌓고 있으며 이를 통해 무엇을 기대하는지 연구하였다. 조성인과 김은애(2014)의 연구에서는 구직요건(SPEC)에 대해 기업의 인사담당자가 갖는 인식과 취업준비생이 갖는 인식의 차이를 검증하고자 하였다. 이 결과를 통해서 흔히 얘기하는 8대 스펙(학력, 학점, 토익, 자격증, 어학연수, 인턴 경험, 자원봉사, 공모전 입상 경험) 중 7개의 구직 요건에서 통계적인 인식의 차이가 있음을 밝혀내었다.

데이터마이닝 기법을 활용하여 취업 관련 요인을 찾고자 하는 연구도 있었다. 최영욱(2015)은 대학에서 시행하고 있는 경력개발관리시스템의 참여자와 비참여자의 취업률을 비교했을 뿐만 아니라, 취업에 영향을 미치는 요인들을 통계적 방법과 랜덤 포레스트(Random Forest) 기법으로 검증함으로써 프로그램에 적극적으로 참여하는 것이 취업에 긍정적인 영향을 미친다고 하였다. 또한 조장식(2011)은 입학 정보, 학사 정보, 개인 특성 등의 변수들을 이용하여 데이터마이닝 기법을 적용하여 취업에 미치는 영향력을 분석하였다. 또한 장애인 취업과 관련된 요인을 찾아내는데 있어서도 데이터마이닝 기법이 활용되고 있다. 양수정(2014)은 데이터마이닝 기법을 이용하여 장애인의 취업 모형에 대한 연구를 수행하였고, 임은정 등(2013)과 신현욱(2011)은 대상을 한정지어 각각 시각장애인과 지체장애인의 취업결정요인을 분석하기도 하였다.

III. 연구 방법

1. 데이터 처리

시행되고 있는 대학교 취업역량강화 프로그램 항목의 실효성과 취업에 영향을 미치는 핵심 요인을 분석하기 위해 K대학교 IT관련 학과의 졸업생들을 대상으로 하였다. 표 1에서 보듯이 해당 학과에서 2010년과 2015년 사이에 졸업한 490명중 취업역량강화 프로그램 점수를 보유하지 않은 112명을 제외하고 378명의 샘플을 추출하였다. 이 과정에서 ‘래피드마이너’에서 제공하는 데이터 필터링과 분석을 이용하였다. ‘Read Excel’을 통하여 490명의 졸업생들의 정보가 저장된 엑셀 파일을 불러왔는데 이 중 취업역량평가 점수를 보유하고 있지 않은 사례들(결측치)을 제거하기 위하여, ‘Filter Example’을 통해 데이터를 필터링하는 작업을 수행하였다. 이 필터링 작업을 통해서 점수가 없는 112명의 데이터가 제거 되었고, 나머지 378명의 사례가 모델을 생성하기 위한 학습 데이터로 활용되었다.

취업역량점수를 바탕으로 모수적인 로지스틱 회귀분석과 비모수적인 기법인 의사결

정나무를 적용하였다. 회귀 분석은 독립변수와 종속 변수간의 관계를 모형으로 설정하여 두 변수 간의 상관관계를 예측, 분류하는 방법이다. 연속형 변수의 예측은 선형 회귀분석 기법을 이용할 수 있는데 이러한 선형 회귀 분석은 취업 혹은 미취업 두 개의 값만 갖는 범주형 종속변수를 설명하지 못한다. 따라서 목표변수가 이분형(Binary)인 경우에는 로지스틱 회귀분석 기법을 이용하여 분류할 수 있다(이재현, 2007). 의사결정나무는 데이터마이닝에서 분류(Classification)의 한 종류로서 과정이 나무 형태로 표현되기 때문에 직관적인 이해가 쉽고 설명력이 우수하여 탐색적인 분석 작업에 용이한 기법이다(이상호, 2007).

본 연구는 취업에 영향을 미치는 요인을 규명하기 위한 목적을 가지고 있으므로 이해가 쉽고 높은 설명력을 필요로 한다. 따라서, 학습 데이터의 종속 변수가 ‘취업’ 혹은 ‘미취업’인 이분형 변수이기 때문에 의사결정나무 분석 기법과 로지스틱 회귀분석 기법을 활용하였다. 취업요인 분석에 데이터마이닝 기법 적용을 위해 오픈소스(Open-Source) 데이터마이닝 도구인 ‘Rapidminer Studio(이하 래피드마이너)’를 활용하였다. 래피드마이너는 다양한 지도학습기법(Supervised Learning)과 비지도학습기법(Unsupervised Learning), 텍스트 분석 등을 포함한 다양한 데이터 분석 기법을 제공하고 있다.

<표 1> 분석 대상

		총 합계	주간			야간		
			남	여	합계	남	여	합계
2010	취업	34	16	4	20	12	2	14
	미취업	27	14	8	22	3	2	5
2011	취업	34	15	7	22	11	1	12
	미취업	42	21	10	31	8	3	11
2012	취업	67	29	13	42	19	6	25
	미취업	32	22	9	31	0	1	1
2013	취업	45	20	11	31	10	4	14
	미취업	26	14	5	19	4	3	7
2014	취업	60	23	17	40	14	6	20
	미취업	35	14	11	25	8	2	10
2015	취업	35	16	13	29	6	0	6
	미취업	53	23	10	33	15	5	20
합계	취업	275	119	65	184	72	19	91
	미취업	214	108	53	161	38	15	53
	합계	490	227	118	345	110	35	145

K대학교의 취업역량강화 프로그램은 목표관리역량, 전공 역량, 국제화 역량, 봉사 역량, 창의 역량, IT 역량, 진로/취업 역량 등 6가지의 카테고리와 모든 항목의 환산 점수를 더한 합계점수인 한도적용총점으로 구성되어 있다. 목표관리 역량에는 자기소개서 항목과 교수 상담 항목이 포함되고, 이 두 항목의 환산점수를 더한 점수가 목표 관리 역량 점수가 된다. 이처럼 전공 역량에도 2가지 항목, 국제화 역량은 5가지 측정 항목으로 평가하고 있다. IT 역량은 IT 자격증 취득 시 점수가 주어지고, 봉사 역량과 창의 역량은 각각 3가지 측정항목을 기준으로 점수를 부여하고 있다. 마지막으로 진로/취업 역량은 교내외 취업관련 프로그램이나 활동, 과목 수강 시에 점수가 부여되는데 총 4가지 항목으로 이를 평가한다. 각 항목별로 얻을 수 있는 점수한도가 적용된 후에 얻어진 합계가 개인의 한도적용총점이다. 각 항목들은 <표 2>에 정리되어 있다.

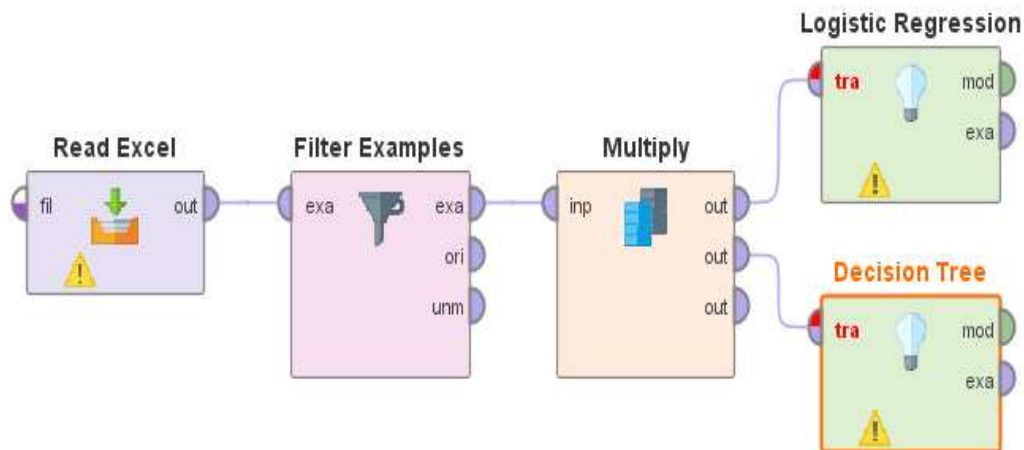
<표 2> 취업역량 측정항목

카테고리	측정 항목
목표관리역량	자기소개서 교수상담
전공역량	성적 국내외인턴쉽
국제화역량	어학자격증 교환학생 국외어학연수 전공영어교과목 다문화교양교과목
봉사역량	국내외봉사 교내간부활동 헌혈
IT역량	IT자격증
창의역량	공모전대회 발명특허 창업교양교과목 수강
진로/취업 역량	전공취업관련자격증 취득 전공취업관련교육프로그램 이수 전공취업관련대외공동활동 취업교양교과목이수
합계	한도적용총점

2. 데이터 분석

본 연구에서는 실제 취업에 도움이 되는 요인들이 각 역량 별로 잘 반영이 되어 있는지 검증해 보기 위하여 각 카테고리에 속해 있는 측정항목들의 총점을 합산하여 해당 역량의 총점을 입력 데이터로 로지스틱 회귀분석과 의사결정 나무 알고리즘을 적

용하여 카테고리별 분석을 진행하였다. 그리고 21개의 모든 측정 항목들을 독립적으로 전체 학습데이터로 활용하였다. 전체 21개 변수 중에 어떠한 변수들이 취업에 가장 긍정 또는 부정적인 영향을 미치는지 알아보고자 한다. 그림 1은 래피드마이너에서의 분석과정을 요약한 것이다. 학습 데이터를 통해서 로지스틱 회귀분석 모델과 의사결정나무 모델을 생성하였다. 분석의 순서는 변수의 범위가 큰 것부터 수행한 뒤에 보다 세밀한 분석을 진행하는 하향식으로 진행하였다. 먼저, 각 역량별 점수를 합산한 6가지 카테고리 점수에 한도적용총점을 포함하여 7개의 변수로 모델을 생성하는 역량별 분석을 앞서 수행하였다. 이후 21가지의 모든 측정 항목들을 각각의 독립적인 변수로 처리하여 분석을 수행함으로써 취업에 영향을 미치는 요인들을 파악하고자 하였다.



<그림 1> 래피드마이너를 활용한 분석 프로세스

IV. 연구 결과

1. 카테고리변수 분석결과

1) 로지스틱 회귀 분석

점수를 합산한 역량 카테고리별 분석결과는 표 3과 같다. 즉 취업역량 프로그램을 측정하는 7가지 역량점수를 학습 데이터로 로지스틱 회귀분석을 수행한 결과인 것이다. 진로/취업 역량 점수가 목표변수인 취업여부에 미치는 영향력이 0.329로 가장 높은 것으로 나타났다. 이는 진로/취업 역량 점수가 높을수록 취업에 성공할 확률이 다른 역량들의 점수가 높을 때 취업에 성공할 확률보다 높다는 것을 의미한다. 그 뒤를 이어서 전공역량과 봉사역량 순으로 높은 것으로 나타났다. 또한, 개인의 취업역량강화 프로그램의 총 합산 점수인 한도적용총점은 취업과는 가장 상관관계가 없는 것으로 나타났으며 오히려 음(-)의 상관관계가 나타났다. 이외에도 목표관리역량과 IT역량은 음의 상관관계를 나타내고 있는 것을 알 수 있다. 그러나, 결과값이 $-0.3 \sim 0.3$ 수준이므로 매우 신뢰성이 높다고 판단하기는 어렵다.

<표 3> 카테고리별 로지스틱 회귀분석 결과

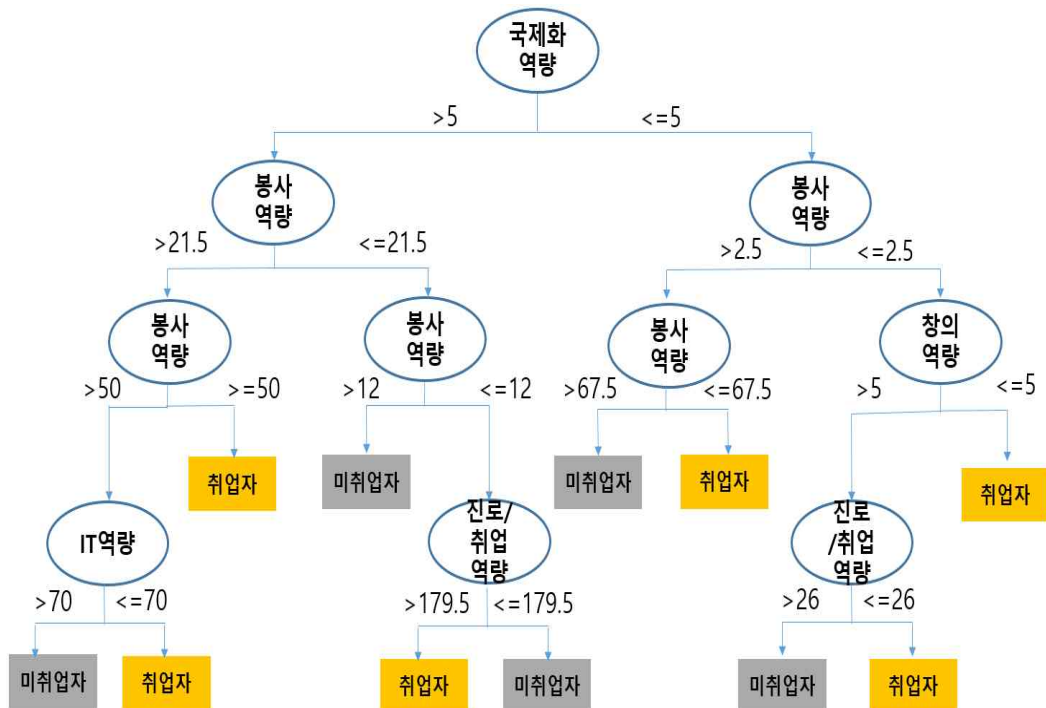
역량	가중치
진로/취업 역량	0.329
전공 역량	0.263
봉사 역량	0.238
창의 역량	0.120
국제화 역량	-0.094
목표관리 역량	-0.122
IT 역량	-0.188
한도적용총점	-0.357

2) 의사결정나무 분석

취업 여부에 가장 큰 영향요인을 찾는 로지스틱 회귀분석 뿐만 아니라 어떠한 특성을 갖는 사람들이 취업을 하였는지 알아보기 위해 의사결정나무 분석을 수행하였다. 의사결정나무 분석은 특정 기준을 바탕으로 취업자와 미취업자를 구분해주기 때문에 취업에 성공한 사람들이 어떠한 공통적인 특성을 가지고 있는지 알 수 있다. 그림 2는 취업역량강화 프로그램의 7가지 카테고리를 기준으로 의사결정나무 분석 기법을 통해 생성된 모델이다. 7가지의 역량 중 취업자와 미취업자를 구분하는 가장 큰 변수는 국제화 역량이었다. 좌측의 나뭇가지를 보면, 국제화 역량이 5보다 높은 사람들 중, 봉사 역량 점수가 21.5점보다는 높고 50점보다는 같거나 낮은 졸업생들은 모두 취업에 성공하였음을 보여준다.

우측 가지를 해석해본다면, 국제화 역량이 5점보다 같거나 낮고, 봉사역량이 2.5점

보다 높으면서 창의역량이 5점 이하인 그룹은 취업에 성공한 사람들의 사례가 많은 그룹이라고 해석할 수 있다. 하지만 이 의사결정나무 분석 결과도 취업에 성공한 그룹의 사례 수가 5명 이하인 경우가 많아서 일반화에는 어려움이 있고 1차 분석은 다소 포괄적인 형태로 진행되었으므로 추가적인 분석이 필요하다고 판단하였다. 따라서 본 연구에서는 보다 구체적인 분석을 위해서 2차 분석을 수행하였다.



<그림 2> 카테고리별 의사결정나무 결과

2. 측정 항목변수 분석결과

1) 로지스틱 회귀분석

표 4는 측정항목별 로지스틱 회귀분석의 2차 결과를 보여준다. 21개의 모든 변수가 상관 계수와 함께 나열되고 있는데, 취업 여부와 가장 큰 상관 관계를 가지고 있는 변수는 교내간부활동 변수인 것으로 나타났다. 교내간부활동은 봉사 역량에 포함되는 항목으로, 직책에 따라 다르게 점수가 부여되는데 교내간부활동 점수가 높은 사람일

수록 취업에 성공할 확률이 높은 것으로 해석할 수 있다. 상관 계수는 0.840으로 압도적인 수치를 나타내고 있는데, 이는 매우 높은 신뢰도를 나타낸다고 할 수 있다. 실제 사례 수를 통해 검증해본 결과 교내간부활동 점수를 보유하고 있는 21명 중 18명의 졸업생이 취업을 하여 약 86%가 취업에 성공하였던 것을 확인할 수 있었다. 이 결과에서 확인할 수 있듯이 378명의 분석대상 졸업생 중에서는 교내간부활동 여부가 취업에 매우 큰 영향을 미치는 것을 알 수 있다.

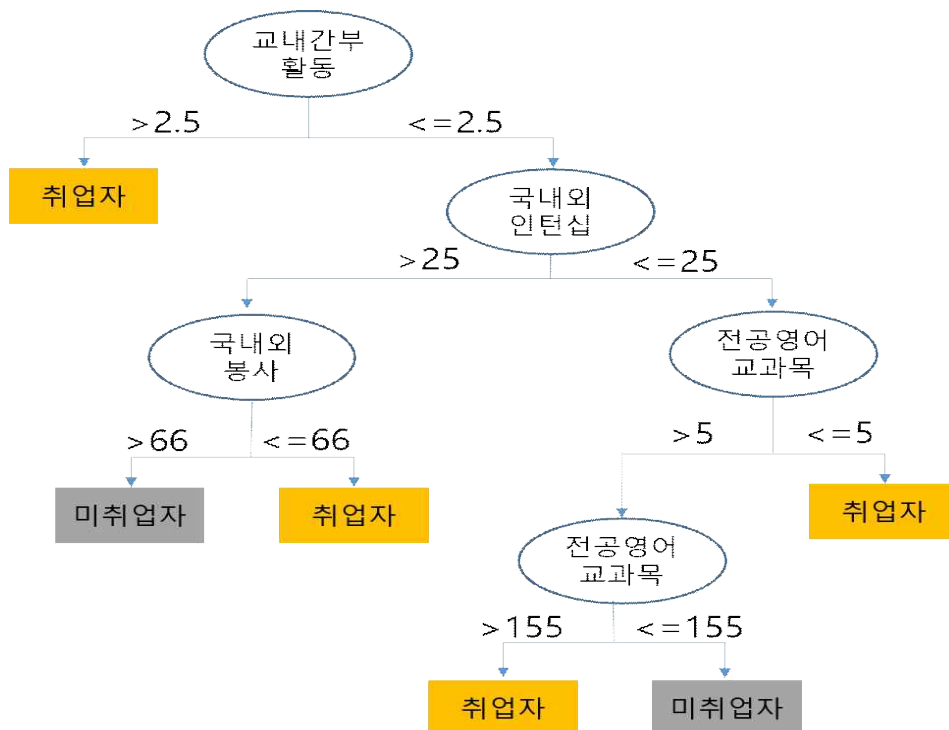
<표 4> 측정항목 로지스틱 회귀분석 결과

프로그램	가중치
교내간부활동	0.840
국내외인턴십	0.564
국내외봉사	0.252
진공취업관련교육프로그램	0.239
발명특허	0.216
창업교양교과목	0.123
공모전대회	0.119
진공취업관련자격증	0.112
성적	0.106
다문화교양교과목	0.058
국외어학연수	0.046
교수상담	0.021
한도적용총점	0.014
현열	-0.056
교환학생	-0.133
IT자격증	-0.156
취업교양교과목이수	-0.162
어학자격증	-0.191
진공취업관련 대외공동활동	-0.398
진공영어교과목	-0.480
자기소개서	-0.506

다음으로는 국내외인턴십 변수가 0.564로 2번째로 높은 상관 계수를 보였는데, 매우 유의한 결과라고 판단할 수 있다. 이 또한 실제 사례 수를 통해 검증해본 결과 한 학기 동안 인턴십을 수행하고 15학점을 인정받는 경영정보학기인턴십 경험이 있는 졸업생들 30명 중 26명이 취업을 하여 87%의 취업률을 보였고, 3학점을 인정받는 단기인턴십을 수행한 졸업생 36명 중 23명이 취업에 성공함으로써 64%의 취업률을 나타냄으로써, 국내외인턴십 여부가 취업 여부에 매우 유의한 영향을 미치고 있음을 알 수 있다. 반대로, 진공취업관련 대외공동활동과 진공영어교과목 수강, 자기소개서 등 등록 등의 항목들은 높은 음(-)의 상관관계를 나타내고 있으므로, 이 항목들이 취업역량 강화 프로그램의 평가 항목으로 적절한지 실효성을 검증해볼 필요가 있을 것으로 판단된다.

2) 의사결정나무 분석

측정항목들에 대한 의사결정나무 분석은 로지스틱 회귀분석과 동일하게 21개의 측정항목들을 대상으로 수행하였다. 변수가 많은 만큼 공통적인 특성을 갖는 졸업생들의 사례 수가 많을 것으로 예측하고 분석을 진행하였다. 그림 3은 의사결정나무 분석 결과를 보여주는데, 로지스틱 회귀분석 결과와 동일하게 취업자와 미취업자의 분류 기준 중 가장 중심이 되는 항목은 교내간부활동 변수였다. 교내간부활동 점수가 2.5점보다 큰 9명의 졸업생들은 전원 취업에 성공함으로써 100%의 취업률을 나타냈다. 교내간부활동 점수를 보유하고 있지만 2.5점보다 낮은 그룹의 경우, 국내외인턴십 점수가 25점보다 높고 국내외봉사 점수가 66점보다 낮거나 같은 졸업생은 전원 취업에 성공한 반면에, 66점보다 높은 사람들은 미취업자 그룹으로 분류되었다. 또한 국내외인턴십 점수가 25점보다 낮거나 같으며, 전공영어교과목 점수가 5점보다 높고 전공취업관련자격증 점수가 155점보다 높은 졸업생들은 취업에 성공한 그룹인 것을 알 수 있다.



<그림 3> 측정항목별 의사결정나무 분석 결과

V. 결론

본 연구는 K대학교의 IT관련 학과의 졸업생들을 대상으로 취업에 영향을 미치는 요인을 파악하고자 하였다. 대학에서 취업역량 강화를 위하여 시행되고 있는 다양한 프로그램들을 대상으로 실제로 취업에 도움이 되는 요인을 찾기 위하여 데이터마이닝 기법을 적용하였다. 먼저 21개의 측정 항목들을 7개의 카테고리로 분류하여 카테고리별 점수 합계를 학습 데이터로 하여 로지스틱 회귀분석 모델과 의사결정나무 모델을 생성하였다. 진로/취업 역량(0.329)이 취업 여부와 가장 높은 긍정적인 영향력을 나타내고 있었고, 전공역량(0.263)과 봉사역량(0.230)의 순서로 나타났다. 이는 진로/취업과 관련된 활동을 많이 수행한 대학생들이 실제로도 취업 가능성이 높아진다고 볼 수 있다. 또한 전공관련 점수와 봉사역량 점수도 다른 변수들에 비해 상대적으로 취업에 영향을 주는 것으로 나타났다.

카테고리별 분석의 의사결정나무는 국제화점수가 5보다 높은 사람들 중에서 봉사점수가 21.5점보다 높고 50점보다는 낮거나 같은 사람들은 모두 취업에 성공하였다. 우측 가지를 해석해 본다면 국제화점수가 5점보다 낮거나 같고, 봉사역량이 2.5점보다 높으면서 창의역량이 5점보다 낮거나 같은 사람들은 취업에 성공한 사람들의 사례가 많은 그룹이라는 결과가 나타났다. 그러나 각 그룹의 사례들이 5명 이하인 경우가 많아서 일반화에 어려움이 있을 것으로 판단되어 카테고리 내의 21개 측정항목들을 대상으로 2차 분석을 수행하였다. 측정항목 21개를 각각 하나의 변수로 가정하여 학습 데이터로 활용하였다. 로지스틱 회귀분석 결과는 교내간부활동 항목이 0.840으로 압도적인 영향을 미치고 있는 것으로 나타났다. 실제 사례 수에서도 교내간부활동 점수를 보유하고 있는 21명 중 18명(87%)이 취업에 성공한 것을 확인할 수 있었다. 다음으로 국내외인턴십 여부도 0.564로 매우 높은 상관 계수를 보였고, 단기 및 장기 인턴십 경험이 있는 66명 중 49명이 취업에 성공하였다.

의사결정나무 분석을 통해 생성된 모델을 살펴보면, 교내간부활동 점수가 2.5점보다 높은 9명의 졸업생들은 전원 취업에 성공함으로써 100% 취업에 성공하였다. 교내간부활동 점수를 보유하고 있지만 2.5점보다 낮은 그룹의 경우, 국내외인턴십 점수가 25점보다 높고 국내외봉사 점수가 66점보다 낮거나 같은 사람은 전원 취업하였지만 66점보다 높은 사람들은 미취업자 그룹으로 분류되었다. 이 결과에서도 로지스틱 회귀분석과 마찬가지로 교내간부활동과 국내외인턴십 그리고 국내외봉사 점수가 가장 중요한 취업에 있어서 중요한 요인이었음을 알 수 있다.

본 연구의 결과는 이론과 실무적으로 시사점을 가지고 있다. 이론적으로는 연구자

가 임의로 취업역량 측정 항목을 선정한 것이 아니라 실제 대학교의 취업역량 평가항목과 점수를 사용하여 실효성을 검증해보았다는 것이다. 실무적으로는 실제 졸업자들의 데이터를 활용함으로써 현재 대학생들의 취업에 도움이 되는 요인들을 탐색하여 제시하였다는 것이다. 그러나 결과가 비교적 명확하게 나왔음에도 본 연구는 몇 가지 한계점을 가지고 있다. 첫째, 학습 데이터로 활용되었던 데이터세트의 사례가 K대학에 속한 것이어서 연구의 결과를 다른 대학에 일반화하기에는 무리가 있다. 향후 연구에서는 타 대학으로 부터 보다 많은 데이터를 확보하여 일반화할 수 있는 모델을 생성할 필요가 있다. 둘째, 각 전공마다 취업에 영향을 미치는 핵심 요인들이 크게 차이가 있을 것임에도 불구하고 본 연구는 단일 학과 졸업생들만을 대상으로 분석을 수행하였으므로, 분석 결과를 다양한 전공에서 참고하기는 어려움이 있을 수 있다. 마지막으로, 매우 다양한 데이터마이닝 기법들이 존재하는데 그중 로지스틱 회귀분석과 의사결정나무 두 가지만을 활용하였다. 따라서 다양한 모델을 생성하여 그 성과를 측정·비교하는 부분도 포함되지 않았으므로 향후 연구에서는 다양한 기법들을 적용해보고 설명력이 높은 모델을 통해서 취업 여부를 예측할 수 있는 모델을 생성하려는 시도가 필요하다.

참고문헌

- 박순영, “대학생들의 취업을 위한 스펙관리에 관한 연구”, 전남대학교 석사학위 논문, 2007.
- 이만기, 김효원, “대졸자의 괜찮은 일자리에 영향을 미치는 취업준비활동의 영향분석”, 취업진로연구, 제 5권, 제 4호(통권 12호), 한국취업진로학회, 2015, pp.115~138.
- 이상호, “C4.5 알고리즘을 이용한 알고리즘을 이용한 의사결정나무 구성 및 탐색 키폴론트 개발”, 울산대학교 석사학위 논문, 2007.
- 이재현, “로지스틱 회귀모형에서의 로버스트 주성분 추정법에 관한 연구“, 숙명여자대학교 석사학위 논문, 2007.
- 조성인, 김은애, “구직요건(SPEC)에 대한 기업 인사담당자와 취업준비생의 인식 차이”, 취업진로연구, 제4권 1호(통권 7호), 2014, pp.41~51
- 최영옥, “대학의 경력개발관리시스템 참여자와 비참여자의 취업률 비교 분석 : Y대학 사례”, 연세대학교 석사학위 논문, 2015..