

청년취업 결정요인에서 본 청년 취업대책 개선책 연구*

한성민**·문상호***·이숙종****(성균관대학교)

clon96@skku.edu ; smoon@skku.edu ; skjlee@skku.edu

최고치를 경신한 청년실업률로 인해 청년실업문제는 경제문제를 벗어나 사회문제로 까지 확산되고 있다. 국가경제에 부정적인 영향을 미치고 있음은 물론 많은 청년들이 미래설계를 포기하고 희망을 잃어가고 있다. 정부는 이러한 문제들을 해결하기 위해 공공부문 일자리 및 다양한 직장체험기회 제공, 직업훈련과 취업알선 등 각종 대책을 수립하고 시행해 왔지만 청년실업률은 좀처럼 낮아지지 않고 있다. 이에 따라 본 연구는 큰 가시적인 효과를 보여주지 못하고 있는 정부의 청년취업대책의 문제점을 살펴보고 기존 거시적 연구들이 결여하고 있는 청년층의 취업에 영향을 주는 요인들에 대해 탐색적 분석을 시도하였다. 의사결정나무 분석과 신경망분석을 혼합한 데이터마이닝 기법을 활용한 결과 개인 특성에서는 성별과 학력, 가족환경에서는 가족수입만족도, 취업준비에서는 직업훈련이 청년취업 결정요인으로 작용하는 것을 발견했다. 청년취업 대책의 효과성을 높이기 위해서는 위와 같은 요인들이 충분히 고려될 필요가 있다.

■ 주제어: 청년실업, 데이터마이닝, 취업결정요인, 취업대책

I. 서론

대한민국의 청년실업과 고용절벽은 갈수록 심각해지고 있다. 2016년 2월 청년실업률은 조사 이래 최고치를 기록했으며 이로 인해 청년들은 장년기와 노년기가 되어서 까지 빈곤에 영향을 받게 되었다. 청년실업문제는 국가경제에 부정적인 영향을 미치고 있음은 물론¹⁾ 출신배경에 따른 기회 불균등을 바탕으로 수저계급론²⁾과 ‘헬조선’³⁾ 같은 신조어를 탄생시

* 이 논문은 2013년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2013S1A3A2055042)

** 제1저자

*** 제2저자

**** 교신저자

1) 서울신문, “청년실업 경제성장률 까지 낮춘다”, 2015.10.22.

키며 사회문제로까지 확산되고 있다. 정부는 이를 해결하기 위해 공공부문 일자리 및 다양한 직장체험 기회 제공, 직업훈련과 취업알선 등 각종 대책을 수립하고 시행해 왔으나 이러한 노력에 비해 청년실업률은 쉽게 낮아지지 않고 있다. 또한 청년일자리 창출이 단기 일자리 중심으로 집중되어 왔으며 구직자에 대한 장려 보다 기업에 대한 지원금 형식으로 이루어진 것은 정책의 효과성에 대한 의구심을 가지게 한다. 청년일자리에 관한 예산 규모와 집행 및 배분에 대한 문제도 지적되고 있으며 향후 경제전망이 밝지 않을 것으로 예상됨에 따라 청년실업문제의 해결은 앞으로도 쉽지 않아 보인다. 본 연구는 이와 같이 심각한 경제 및 사회문제로 자리 잡은 청년실업문제의 해결을 위해 청년층의 취업에 영향을 주는 요인들에 대한 탐색적 분석을 시도하였다. 본고에서는 청년실업의 현황 및 원인과 함께 역대 정부의 청년실업에 대한 대책들을 살펴볼 것이다. 이후 의사결정나무분석과 신경망분석을 혼합한 데이터마이닝 기법을 활용하여 청년층의 취업에 영향을 미치는 요인들을 찾아내 청년 취업 성공률을 높일 수 있도록 정책적인 함의를 찾고자 한다. 데이터마이닝 기법이란 대용량의 자료로부터 이전에 알려지지 않은 유용한 정보들을 뽑아내 의사결정에 도달하는 경로를 설명하고 연결시키는 주요 정보기술이다⁴⁾. 데이터마이닝 모델인 의사결정나무분석, 신경망분석 두 모델을 비교하는 것은 물론 이를 통해 청년취업 영향요인 분석에 가장 적합한 예측모델은 무엇인지에 대해서도 함께 설명할 것이다.

II. 연구배경

1. 청년실업의 현황

청년층의 경제활동현황은 우리나라 전체와 비교했을 때 매우 심각한 수준이다. 2016년 2월 기준 우리나라의 만 15세부터 29세까지의 청년인구는 만 15세 이상 인구의 약 20%를 차지하고 있으나 경제활동인구는 전체 경제활동인구의 6.1% 밖에 되지 않는다.

2) 중앙일보, “농담인데 불편하네, 수저 계급론”, 2015.10.28.

3) 경향일보사, <주간경향> 1157호, 2015.12.

4) Brethenoux 외, 1996; 신현욱, 2011; 임은정 외, 2013.

〈표1〉 청년경제활동 현황 (단위: 천명, %)

구분	인구 (15세 이상)	경제활동인구			비경제 활동인구	경제활동 참가율	실업률	고용률
		계	취업자	실업자				
전체	43,274	26,734	25,418	1,317	16,540	61.8%	4.9%	58.7%
청년	9,451 (~29세)	4,477	3,916	560	4,974	47.4%	12.5%	41.4%

자료: 통계청, 2016년 2월 고용동향.

특히 〈표2〉에 제시한 바와 같이 청년실업률은 전체실업률과 비교하여 2.6배나 더 높다. 12.5%의 실업률은 역대 최고치를 갱신한 것이다. 우리나라 경우 2014년 OECD 평균 15~24세 청년실업률인 7.4% 보다는 조금 더 높은 수준으로 파악되나⁵⁾ 세대 간 위험배분 측면에서 청년층에게 부담이 치우쳐 있다는 특수성이 있다⁶⁾.

〈표2〉 전체실업률 대비 청년실업률 (단위: %)

	2012년	2013년	2014년	2015년	2016.2
전체실업률	3.2%	3.1%	3.1%	3.6%	4.9%
청년실업률	7.5%	8.0%	9.0%	9.2%	12.5%
전체실업률 대비 청년실업률	2.3배 높음	2.6배 높음	2.9배 높음	2.6배 높음	2.6배 높음

자료: 통계청, 2016년 2월 고용동향.

청년실업문제는 노동문제를 벗어나 이미 사회문제로 자리 잡은 것으로 보고는 파악한다. “사회전체의 문제라는 보편성”을 이미 지니고 있다는 의미이다⁷⁾. 청년들은 실업 및 고용절벽으로 인해 연애, 결혼, 출산은커녕 꿈과 희망까지 포기하고 있는데 이것은 생애주기의 문제를 넘어 어쩔 수 없는 포기이자 계산적인 선택이다⁸⁾. 몇 개만 특정 지을 수 없을 정도로 거의 모든 것을 포기한 세대라는 뜻의 “N포세대”라는 신조어가 회자되고 있는 현실은 이러한 시대의 트렌드를 반영하고 있는 것으로 볼 수 있다⁹⁾.

5) OECD, “OECD Employment Outlook 2015”, p.273.

6) Chmid, Gunther. 2013. “Youth Unemployment in Korea: From a German and Transitional Labour Market Point of View”, IZA policy paper no. 63.

7) 서울시 청년허브, (2015), “서울이 만드는 다음 세대의 여지 청년이 만드는 다음 사회의 공공”, 〈청년정책의 재구성 기획 연구〉: 76.

8) 한금윤, (2015), “대학생의 연애, 결혼에 대한 의식과 문화 연구”, 〈인간연구〉28: 24.

9) 김난도, (2015), 〈트렌드코리아 2016〉.

2. 정부 청년실업 대책

청년실업에 대한 대책은 사실상 2010년에 들어서야 본격적으로 시작되었다고 볼 수 있다. 그 이전에는 IMF이후 구조적인 문제에 초점을 두고 진행이 되어 왔다. 2010년 이후 정부의 대책들은 청년일자리 창출, 청년 내일 만들기, 청년 일자리 미스매치 완화, 청년 창업 진흥, 청년 해외취업 지원 등이다. 최근에는 특히 현장중심의 직업교육과 스펙월 채용시스템을 확산하고자 많은 노력을 기울이고 있다¹⁰⁾. 현 박근혜정부에 들어서는 공공기관 및 지방공기업에 정원 3%이상 미취업 청년을 의무고용하게 하고 교원들의 명예퇴직과 포괄간호서비스의 확대, 시간선택제 등이 새롭게 사업내용에 추가되었다. 이로 인해 공공부문 4만 명 이상 민간부문에서 상시고용과 인턴제도의 확대나 직업훈련 등을 통해 16만 명을 포함한 총 20만 명을 목표로 일자리를 창출하겠다는 것이다. 청년실업대책은 발표를 거듭할수록 지원대상과 예산이 증가하고 있다. 그러나 청년실업률에 대한 노동규제 완화를 제외하면 이들 대책이 청년실업문제의 해소에 어떠한 핵심적 역할을 하고 있는지 명확하지 않다. 게다가 이러한 대책들은 일시적 처방으로 이뤄져 있는 경우가 많은데 그것만으로는 실현가능성이 없는 정책에 그칠 가능성이 높다¹¹⁾. 2015년 감사원은 청년실업 대책에 대해 직업훈련과 고용서비스에 비중을 두라는 OECD의 권고와는 달리 예산의 상당부분이 직접적인 일자리로 투입되고 있음을 지적한 바 있다.¹²⁾ 이러한 직접적 일자리 지원의 경우 중장기적 효과는 오히려 부정적인 것으로 연구 된다¹³⁾. 또한 이렇게 투입되는 정부의 예산이 사업자위주의 고용보조금 형태로 지원되고 있기 때문에 정책체감 및 만족도 향상에 기여하지 못하고 있음은 물론 오히려 기업의 비용절감 수단으로 활용될 가능성이 있다¹⁴⁾. 최근 기획재정부와 고용노동부는 기업에 지원되는 고용보조금을 낮추고 개인에게 직접 주는 보조금을 높이는 방안을 검토 중이다¹⁵⁾. 기업을 통한 보조금이든 개인에게 주는 직접적인 보조금이든 이러한 처방이 청년실업을 줄이는 데 기여하려면 청년실업의 원인을 보다 복합적으로 이해한 다음에 설계되어야 정책 효과성을 높일 수 있을 것이다.

10) 이규용, (2015), “청년층 일자리정책의 방향모색”, <노동리뷰> 2015 7월호: 16.

11) 최용환, (2015), “청년 실업률의 영향요인과 정책방향 탐색”, <한국청소년연구>: 88-105.

12) 감사원, (2015), “재정지원 일자리 사업 추진실태”.

13) 김용성, (2013), “고용률 제고를 위한 재정지원 일자리사업 방향의 모색”, <KDI 정책연구> 2013-19: 26.

14) 국회예산정책처, (2016), <경제동향 & 이슈> 42호: 24.

15) 동아일보, “사업주에 주던 고용보조금, 청년 취업자에 직접 주기로”, 2016.2.29.

3. 청년실업의 원인

기존 연구들을 종합한 청년실업의 원인은 크게 <표3>과 같이 세 가지로 분류할 수 있다.

첫째는 고용흡수력의 감소와 기업규모의 불균형에 따른 ‘노동시장의 수요측면’이다. 고용흡수력은 실질국내총생산이 1% 상승할 때 고용이 얼마나 증감하는지를 보여주는데 경제성장의 둔화로 인해 고용흡수력이 저하되었기 때문에 청년일자리에 대한 수요도 감소하고 있다¹⁶⁾. 기업들은 채용규모를 줄이고 있다. 2016년 대기업 신입사원 공채의 경우 2015년에 비해 채용계획이 4.8% 줄었다¹⁷⁾. 공채를 하더라도 작년 수준의 규모를 유지하는 곳은 그나마 상황이 나은 편이다. 공채를 진행하는 곳보다 채용계획이 없는 기업이 더 많다. 중소기업의 경우 상황은 더욱 심각해서 중소기업 10곳 중 6곳이 2016년 인턴조차 채용하지 않을 계획이다¹⁸⁾. 중소기업이 대기업의 줄어든 일자리에 대한 대안이 되지 못하는 실정이다.

둘째는 청년층 학력과잉, 전공과 취업의 불일치로 인한 ‘노동시장의 공급측면’이다. 청년층의 학력과잉은 청년실업의 주요 원인으로 손꼽힌다. 한국의 높은 교육수준은 세계적으로 유명하다. 2015 OECD 교육지표에 따르면 우리나라의 대학진학률은 68%로 OECD 국가 중 1위를 기록했다. 그러나 기업의 대졸자 수요는 이에 미치지 못하고 있고 취업자의 전공과 일자리 내용이 불일치하여 학교교육과 산업수요 사이의 괴리가 발생하고 있다¹⁹⁾. 또한 고학력화에 따른 하향취업현상도 벌어지고 있다. 고등학교를 졸업한 인력은 실제 취업 인력이 필요인력보다 더 적으며 전문대학교를 졸업한 인력과 석사와 박사 인력은 실제 취업 인력이 필요인력보다 더 많다는 연구결과가 있다.²⁰⁾

셋째는 임금경직성과 고용경직성, 고용지원서비스 및 노동시장정보의 취약 등 ‘노동시장 기능측면’이다²¹⁾. 대부분의 기업이 갖고 있는 연공급 임금체계는 일정 근무기간이 초과할 경우 임금이 생산성을 초과하게 되어 기업의 임금부담을 증가시켜 채용기회를 감소시킨다. 또한 경직적인 고용보호법제는 정규직 근로자의 생산성에 비해 높은 임금비용으로 저임금의 비정규직을 회피하게 만들고 이는 청년실업의 가중을 초래한다. 마지막으로 노동시장에 대한 부족한 정보와 고용지원서비스의 미흡은 노동시장에 새로 참가할 청년들의 올바른

16) 이병희, (2004), “청년층 노동시장 구조변화”, <동아시아연구소 주최 ‘고용없는 성장기의 청년실업 해법 토론회’ 발표논문>.

17) 헤럴드경제, “올 상반기 대기업 신입공채 전년 대비 4.8% 감소”, 2016.2.27.

18) 헤럴드경제, “중소기업 10곳 중 6곳 ‘올해 인턴 채용도 힘들어서 안 한다’”, 2016.2.24.

19) 금재호, (2007), “청년실업의 현황과 원인 및 대책”, <숭실대 사회과학논총> 9집: 45-48.

20) 김안국·유한국, (2014), <대학 및 전문대학 졸업자의 직종별 수요추정>, 한국직업능력개발원.

21) 금재호, (2007), 앞의 글.

르고 신속한 판단을 저해함으로써 청년실업에 부정적 효과를 가져 온다.

〈표3〉 청년실업의 원인

노동시장의 수요측면	노동시장의 공급측면	노동시장의 기능측면
·고용흡수력 감소 ·기업규모 불균형	·청년층 학력과잉 ·전공과 취업의 불일치	·임금경직성 ·고용경직성 ·고용지원서비스 및 노동시장정보 취약

이상의 청년실업 원인에 대한 거시적 분석들은 청년층의 취업 동기나 취업 결정과정과 같은 미시적 분석들을 결여하고 있다. 청년실업을 동태적 시각으로 살펴보았을 때 청년실업의 원인이 직장을 못 구해서가 아니고 일자리에 있어서의 이탈이 심하게 일어나기 때문이라는 지적도 있는 만큼 취업 희망자인 청년 개인들의 취업결정 요인을 이해해야 보다 정교한 청년층 취업정책이 만들어 질 수 있을 것이다.

4. 청년층 취업결정 요인

취업결정요인은 사회적 배경, 취업준비행동 같이 취업에 영향을 미치는 요인들을 말한다. 사회적 배경은 부모의 학력이나 소득수준 등을 말하며 취업준비행동은 〈표4〉와 같이 진로준비행동, 구직활동 등 직업에 대한 정보를 수집하고 기초역량을 마련하는 행동들을 말한다(이수영·이영민, 2009).

〈표4〉 취업준비행동 유형

구분		주요변수 내용
취업 준비 행동	(요인1) 취업사전관리	학점관리, 아르바이트 경험, 인턴십 경험, 영어공부 경험, 제2외국어 경험, 어학연수 경험, 영어점수, 직업훈련 참여, 자격증 취득
	(요인2) 취업정보탐색	미래직업 결정(직업목표 결정), 구체적인 미래직업 계획, 미래직업 업무내용 인지(관심), 미래직업 전망 및 보수 인지도부, 미래 직업과 현재 전공과의 관련성 탐색, 직업에의 관심
	(요인3) 직접구직활동	구직활동경험 여부, 취업정보 검색여부, 취업부탁 경험여부, 학교지원서비스 이용 경험, 방문부탁 경험여부, 면접훈련 경험여부, 취업박람회 경험, 구직정보 인터넷 업로드, 인터넷검색경험, 구직상담 경험

출처: 이영민 외(2014), p.342.

청년층 취업결정과 관련해서는 인적자본론, 직업탐색이론, 통계적 차별이론, 노동시장분절론, 선별이론 등의 이론들이 있다. 어수봉 외(2010)는 위의 이론들을 토대로 청년취업에 영향을 주는 요인들을 세 가지 이론적 측면에서 분류하였다. 첫째는 노동공급의 측면을 중요시하는 이론이다. 개인의 기술, 능력, 일하는 태도 등 개별적 요인들을 청년취업 영향요인으로 본다. 따라서 청년취업을 위해 직업훈련과 학교교육을 통해 노동의 질을 개발하는 것을 강조한다. 둘째는 노동시장 분절론에 입각한 노동수요측면 이론이다. 고용이 안정적이고 고임금의 1차 시장, 고용이 불안정하고 저임금인 2차 시장 중 어떠한 시장에 진입하느냐에 따라 청년취업 영향요인이 달라진다는 입장이다. 많은 청년들이 1차 시장 진입을 위해 노력하고 있는 현실을 잘 설명할 수 있다. 셋째는 노동과 수요를 연계시키는 과정을 중요시 하는 입장의 이론이다. 위의 두 이론과 관련된 요인들 외 이 두 요인들이 서로 만나는 계기를 만들어 내는 사회적 관계망과 구직방법 같은 구직과정과 관련한 요인들이 중요한 역할을 한다고 본다.

청년층 취업결정요인 대한 연구는 IMF이후 청년실업이 사회적 이슈로 대두되자 본격적으로 이루어 졌는데 위의 이론을 바탕으로 선행연구들을 살펴보면 성별, 연령, 학력, 대학, 전공, 자격증, 어학능력, 사회적 배경을 요인들로 함을 알 수 있다.

김성훈(2015)은 대학생의 취업 준비 활동에 영향을 미치는 요인에 관한 기존 연구결과를 <표5>와 같이 정리한바 있다. 이를 통해 뚜렷한 유형으로 구분되는 것은 아니지만 일반적인 인식에서 취업에 유리한 영향과 불리한 영향을 주는 요인들을 찾아낼 수 있다. 남성, 4년제 대학, 수도권 대학, 부모의 높은 사회 경제적 지위 등을 유리한 영향을 주는 속성으로 간주하고 있는데 가구소득, 부모계급, 부모학력 등 가족배경이 어학연수, 외국어 사교육 등과 재학 중 노동에 영향을 미치는 중요한 요인임을 확인할 수 있다.

한편 공통적으로 노동공급 측면의 이론에서 ‘직업훈련’ 요인이 그 중요성에도 불구하고 영향요인에 포함되지 않았음을 알 수 있다. 뿐만 아니라 성별, 학력, 가족수입, 부모학력, 직업훈련, 자격증 등 속해있는 요인들 간의 상호작용에 대해서도 궁금증이 생긴다. 따라서 본 연구에서는 청년취업에 ‘직업훈련’ 요인을 포함시키고 각 요인들 간의 상호작용에 대해서도 살피고자 한다.

〈표5〉 대학생의 취업 준비 활동에 영향을 미치는 요인에 관한 기존 연구 결과

연구	취업준비 활동	영향요인
김정숙(2009)	어학연수, 영어사교육, 자격증 취득	성, 아버지학력, 가구소득, 전공계열, 대학소재지
신광영·문수영(2014)	어학연수	부모계급, 부모학력(+), 대학유형(소재지,전문대,일반대)
	외국어 사교육	부모학력(+), 남성(-), 대학유형(소재지,전문대,일반대)
	자격증 획득	부모계급, 부모학력
이상록(2004)	아르바이트	전공 만족도
	직장 체험	성, 학년, 전공계열
조혜영·김지경·전상진(2012)	재학 중 노동	인문사회계열(+), 어머니 임금노동(+, 기준:비경제활동), 국공립대(+), 4년제 대학(-), 비수도권 대학(-), 빛(+), 기숙사 거주(-), 장학금(+), 부모가 자녀에게 기대하는 경제적 자립수준(+), 자녀가 부모에게 학비와 용돈을 기대(-)
	재학 중 노동시간	가구소득(+), 4년제 대학(-), 장학금(-), 부모가 자녀에게 기대하는 경제적 지원수준(-), 교육관련 노동(-)
	학업 외 취업준비 비용	재학기간(+), 어머니 임금노동(+, 기준:비경제활동), 국공립대(+), 4년제 대학(-), 재학 중 노동(-)
채창균(2003)	취업준비	취업희망, 재학기간, 병역, 일반대(기준:전문대), 학점, 아르바이트 경험, 가구소득
채창균 외(2005)	취업준비, 정보획득	직장선정(+), 특정시험 준비(+), 영어점수(-), 아르바이트경험(+), 수도권대학(-), 4학년(+), 전공계열, 가구소득

출처: 김성훈(2015), “대학생의 학업 외 취업 준비의 영향 요인과 취업 성과”, p.117.

Ⅲ. 연구 설계 및 방법

1. 연구대상

본 연구는 한국 노동연구원에서 제공하는 ‘2012년 노동패널’ 자료를 활용하여 청년들이 취업하는데 영향을 미치는 요인을 탐색적으로 분석하였다. 한국 노동연구원은 비농촌지역에 거주하는 한국의 가구와 가구원을 대표하는 패널표본구성원(5,000 가구에 거주하는 가구원)을 대상으로 연 1회 경제활동 및 노동시장 이동, 소득활동 및 소비, 교육 및 직업훈련, 사회생활 등에 대한 조사를 실시해오고 있는데 2012년 조사의 경우 2012년 6월 실시되었으며 약 1만4000명을 설문대상자로 한다.²²⁾ 이중 청년고용촉진특별법에 따라 정부에서 실시하는 청년취업 인턴제, 청년취업 내일공제, 청년취업 성공패키지 등의 대상에 해당하는

만 15세 이상 만 34세 이하인 응답자 3801명을 연구대상으로 하였으며 대상자의 일반적인 특성은 <표6>과 같다.

<표6> 연구대상자 특성

특성	구분	빈도(%)
성별	남성	1181(49.5%)
	여성	1920(50.5%)
학력	미취학	-
	무학	1(0%)
	초등학교	3(0.1%)
	중학교	208(5.5%)
	고등학교	1186(31.2%)
	2년제 대학, 전문대학	931(24.5%)
	4년제 대학	1364(35.9%)
	대학원 석사	92(2.4%)
	대학원 박사	16(0.4%)
가족수입 만족도	매우 만족	20(0.5%)
	만족	878(23.1%)
	보통	2240(58.9%)
	불만족	582(15.3%)
	매우 불만족	49(1.3%)
직업훈련 경험여부	받아본 경험이 있다	156(4.1%)
	현재 받고 있다	19(0.5%)
	받아본 적이 없다	3621(95.3%)
자격증 보유 여부	있다	68(1.8%)
	없다	3729(98.1%)

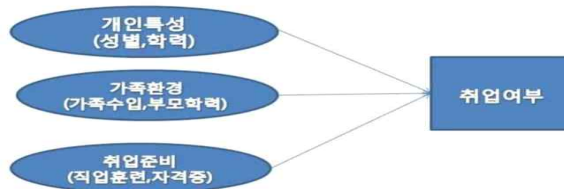
먼저 성비는 49.5%와 50.5%로 고르게 나타났다. 학력의 경우 4년제 대학 졸업자가 35.9%로 제일 높았고 다음으로 고등학교 졸업자가 31.2%, 24.5%인 2년제 및 전문대학 졸업자 순이었다. 대학졸업 여부로 구분하였을 때 2년제 대학 졸업 이상이 63.2%로 고졸이하 학력자보다 많았다. 가족수입에 대한 만족도는 58.9%가 보통으로 답하였으며 만족하는 대상자는 23.6%로 불만족하는 대상자 16.6%보다 많았다. 직업훈련의 경우 95.3%가 받은 경험이 없고 자격증은 1.8%만이 갖고 있는 매우 저조하게 나타났다.

2. 연구모형

선행연구를 바탕으로 본 연구에서는 2012년 한국 노동패널을 기준으로 청년층의 취업에 영향을 줄 수 있는 변인들을 선별하여 구성 하였다. 이영민 외(2014)는 대학생 선호직장의 취업 결정요인을 분석하면서 독립변수를 개인변수, 학교변수, 취업준비행동 으로 구분하였다. 김형래·전도홍(2012)은 대졸자 고용에 미치는 영향요인을 분석하면서 학교생활 요인, 개인환경 요인, 직장생활 요인으로 구분하였다.

이를 참고하여 포괄적이지는 않지만 독립변수를 개인특성 변수, 가족환경 변수, 취업준비 변수로 나누었으며 개인특성 변수로는 성별, 학력, 가족환경 변수로는 부모학력과 가구 소득에 대한 만족도, 취업준비 변수로는 직업훈련과 자격증 취득여부를 활용하였다. 종속 변수는 취업형태의 구분 없이 조사당시의 취업여부(임금근로자)로 설정하였다.

〈그림1〉 연구모형



3. 데이터마이닝 기법

본 연구는 청년층의 취업결정 요인을 분석하기 위해 데이터마이닝에서 주로 사용되는 기법인 의사결정나무 분석과 신경망 분석을 혼합하여 사용하였다. 데이터마이닝이란 패턴인식, 통계 및 수학적 기법, 인공지능 기법 등을 통해 대량의 데이터로부터 잘 알려지지 않고 숨겨져 있던 유용한 데이터들을 찾아 상호관련성 및 패턴, 경향 등 의미 깊은 정보를 추출하는 방법이다(이혜주 외, 2013; Vandamme, Meskens, & Superby, 2007). 즉 “데이터에 함축되어 있으나 이전에는 발견되지 않은 잠재적으로 유용한 정보를 새롭게 추출”하여 “과거의 행위를 분석한 것을 기초로 미래의 행위를 예측하는 의사결정지원을 위한 모형을 만들어 내는 것”이라 할 수 있다(강현철 외, 2015; Berson et al., 1999). 지금까지의 통계분석 방법이 이론적 근거에 따라 연구모형이 설정되고 이후 가설검정을 위한 실증분석이 이루어 졌다

면 데이터마이닝 분석방법은 자료들 속에서 나타나는 일정패턴을 바탕으로 탐색적으로 설명모형을 만들어 내는데 주안점을 둔다(김영숙, 2009; Vandamme et al., 2007). 기존 회귀 분석이나 로짓분석 같은 모수적 방법은 교호효과를 밝혀내는 것이 어려우나 이 방법은 선형성, 정규성, 등분산성 등의 가정이 필요 없는 비모수적 방법이다(진현정, 2010). 따라서 두 개 이상의 변수들의 결합으로 인해 발생하는 종속변수에 대한 영향을 쉽게 알 수 있다. 데이터마이닝의 주요 기법인 의사결정나무 분석은 변수들 간의 교호작용의 파악이 용이하여 분석결과와 분류와 예측이 효과적인 반면 독립변수가 종속변수에 미치는 개별효과를 파악할 수 없다는 단점이 있다. 반면 신경망 분석은 여러 독립 변수들의 상호결합이 종속변수에 미치는 효과를 매우 높은 예측력과 함께 파악할 수 있다는 장점이 있지만 결과에 대한 해석이 용이하지 않은 단점이 있다(강현철 외, 2015). 본 연구에서 사용한 의사결정나무 분석과 신경망 분석을 혼합한 방법은 <그림2>과 같이 의사결정나무 분석을 통해 생성된 변수들을 신경망 분석의 입력변수로 사용하는 방법으로 이 방법을 활용하면 의사결정나무 분석과 신경망 분석의 장단점을 잘 살려 결과에 대한 설명력을 강화시키는 한편 예측의 정확도를 향상시킬 수 있다²³⁾. 기존 연구들이 취업결정에 관한 다양한 요인들의 단편적인 관련성이나 역할을 검증하는데 그치고 있다면 본고에서는 위와 같은 장점들을 살려 청년취업에 보다 결정적인 영향을 미치는 요인들 간의 교호관계의 다각적인 탐색을 시도하였다.

〈그림2〉 의사결정나무와 신경망 혼합



의사결정나무 분석	신경망 분석
- 변수 간 교호작용 파악이 용이 - 분석결과 분류와 예측이 효과적 - 변수가 미치는 개별적 효과 파악 불가	- 변수 간 상호결합 효과 파악 가능 - 높은 결과 예측력 - 결과해석이 힘들
의사결정나무와 신경망 혼합	
- 의사결정나무 분석으로 생성된 변수를 신경망 분석에 사용 - 결과에 대한 설명력 강화 및 예측의 정확도 향상	

출처: 강진웅 외(2011)를 바탕으로 재구성²⁴⁾.

23) 실제 분석에서 의사결정나무 분석은 변수들 간의 관계를 나무도식으로 이해하기 쉽게 보여주었지만 변수가 개별적으로 미치는 영향력은 알 수 없었다. 반면 신경망 분석은 복잡한 parameter들로 인해 예측식을 직접작성하고 해석해야 하는 어려움은 있었지만 변수들 간의 영향력을 높은 설명력과 정확도로 파악할 수 있었다.

1) 의사결정나무 분석

데이터마이닝 방법 중 하나인 의사결정나무 분석은 “의사결정규칙을 나무구조로 도표화하여 분류와 예측을 수행하는 방법”으로 목표변수와 입력변수들 사이의 관련 정도에 따라 영향을 주는 변수를 선별하고, 의사결정나무를 생성하는 기법이다(강현철 외, 2015). 독립변수들이 청년취업에 미치는 개별적인 효과를 파악할 수 없는 대신 요인들 간 상호작용 효과를 해석할 수 있는 분석방법이기 때문에 측정 자료의 세분화, 결과 변인의 분류, 영향력 높은 변인을 선별하기 위한 차원축소 및 변수선택의 목적으로 사용하기에 적합하다(송태민·송주영, 2015; 임희진·유재민, 2007). 데이터마이닝에서는 데이터로의 특성을 파악하여 모형을 먼저 생성한 뒤 이를 토대로 새로운 레코드값의 예측을 통해 변수를 추출한다(이해주 외, 2013).

의사결정나무 분석을 위해서는 CHAID, CART와 같은 다양한 알고리즘이 제안 되는데 본 연구에서 사용된 SAS사의 E-Miner 14.1은 이러한 알고리즘을 결합하고 확장하여 의사결정나무분석을 다양하게 수행할 수 있는 장점이 있다²⁵⁾.

분리에 앞서 분리 규칙은 최대 가지²⁶⁾를 ‘4’로, 최대 깊이²⁷⁾는 ‘10’으로 설정하여 진행하였으며, 나머지 값들은 기준값을 사용하였다. 독립변수로는 본인 학력, 성별, 직업훈련 경험 여부, 자격증 보유 여부, 가족의 수입에 대한 만족도, 부모학력이 사용되었다. 사용된 독립변수에 대한 자세한 내용은 <표 7>과 같다. 학력의 경우 전체 미취학부터 대학원 박사이상 까지 9개의 변수값으로 설정하였으며, 직업훈련 경험 여부는 과거에 받은 경험이 있으면 1, 현재 받고 있으면 2, 받아본 적이 없는 경우 3으로 설정하였다. 자격증의 경우 보유 여부에 따라 1과 2로 나누었으며, 가족의 수입에 대한 만족에 대한 수준은 1부터 5까지 매우 만족, 만족, 보통, 불만족, 매우 불만족으로 설정하여 분석을 실시했다.

24) 앞의 글, p.169.

25) 강현철 외(2015), p.82.

26) 하나의 부모마디(상위 노드)에서 최대 몇 개까지 자식마디(하위 노드)가 형성될 수 있는지를 말한다.

27) 뿌리마디부터 끝마디까지 중간마디의 수를 의미한다.

〈표7〉 의사결정나무 분석에 사용된 독립변수와 변수값

변수명	응답 항목	변수값
성별	남성	1
	여성	2
본인 학력 및 부모학력	미취학	1
	무학	2
	초등학교	3
	중학교	4
	고등학교	5
	2년제 대학, 전문대학	6
	4년제 대학	7
	대학원 석사	8
	대학원 박사	9
직업훈련 경험 여부	받아본 경험이 있다	1
	현재 받고 있다	2
	받아본 적이 없다	3
자격증 보유 여부	있다	1
	없다	2
가족의 수입(생활 만족 수준)	매우 만족	1
	만족	2
	보통	3
	불만족	4
	매우 불만족	5

자료: 한국노동연구원, 15차년도 개인용 코드북 참조.

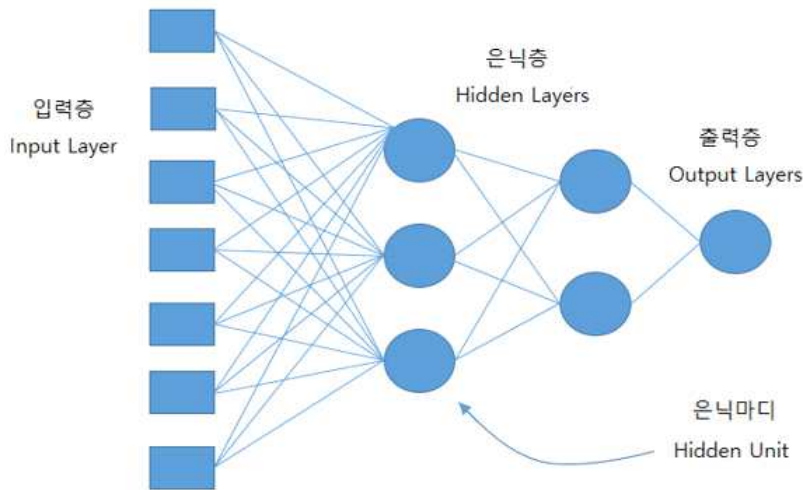
2) 신경망 분석

본 연구에서는 의사결정나무 분석에 이어 인공 신경망에 기반한 신경망 분석을 이용한다. 신경망 분석은 뇌 신경생리학에서 영감을 받은 것으로 1943년 맥컬러와 피츠가 제안한 신경모델에서 출발하였다. 입력정보가 일부 불완전하거나 정확하지 않더라도 올바른 결과를 얻을 수 있는 장점이 있다(김선주·이상엽, 2008). 그러나 해석상의 어려움이라는 단점도 함께 가지고 있다.

신경망 분석방법은 은닉마디를 통해 입력변수와 출력변수와의 관계를 설명한다. 은닉마디는 인간의 신경세포를 모형화한 것이다. 은닉마디들은 입력변수들의 결합을 수신하여 목표변수에 보내는 역할을 한다. 결합에 사용되는 계수들을 연결강도라고 하며 활성화수는

입력값의 변환을 통해 이를 입력으로 사용하는 다른 마디로 출력시킨다. 신경망에는 다양한 모형이 있는데 본 연구에서는 그중 가장 많이 사용되는 MLP(다층인식자)신경망을 사용한다. “MLP는 입력층, 은닉마디로 구성된 은닉층, 그리고 출력층으로 구성된 전방향(feed-forward) 신경망”으로 구조는 <그림3>과 같다²⁸⁾.

<그림3> MLP의구조



출처 : 강현철 외(2015), p.171.

예측에 사용될 신경망의 모형의 설정에는 다섯 단계가 필요하다. 첫째, 신경망의 입력변수와 출력변수를 수집한다. 둘째, 수집된 변수들은 표준화되고 조정되어야 하며, 셋째, 입력변수와 출력변수의 관계를 파악할 수 있게 신경망 모형이 설정되어야 한다. 넷째, 선택 가능한 추정모형들과 검정과 테스트용 자료가 조사되어야 하며, 마지막으로 최적의 모형을 선택해 예측에 사용하여야 한다.²⁹⁾

신경망 노드에서는 초기 설정으로 한 개의 은닉층을 가지는 다층 인식자를 통해 신경망 분석을 수행하게 되며, 은닉층과 은닉마디가 많아질수록 신경망이 더욱 복잡해지며, 최적화가 더 어려워지게 된다. 신경망 분석은 매우 유연한 분석 형태 이지만 앞서 지적과 같이 의사결정나무 분석이나 회귀분석에 비해 결과에 대한 해석에 어려움이 있는 것이 한계점이

28) 강현철 외, 2015.

29) 통계청, (1997), <통계분석연구>.

라 할 수 있다. 은닉마디 수는 1과 64 상의 정수값이 허용되며 기본 값은 3이다. 본 연구에서는 사용된 은닉마디 수를 ‘4’로 설정하였다.

IV. 분석 결과

1. 의사결정나무 분석을 활용한 청년층의 취업여부 결정요인 분석

분석결과 <표8>과 같이 청년층의 취업 여부에 가장 중요도가 큰 변수는 학력이며, 그 다음은 직업 훈련 여부, 가족수입 만족도, 성별 순으로 나타났다. 부모학력과 자격증 취득유무 변수는 영향력을 보이지 않았다.

<표8> 변수 중요도

순위	변수명	분석 중요도	평가 중요도	비율
1	학력	1.0000	1.0000	1.0000
2	직업훈련	0.2964	0.4421	1.4917
3	가족수입 만족도	0.2609	0.1905	0.7302
4	성별	0.2530	0.0000	0.0000

<그림5>는 청년층의 취업요인을 의사결정나무 분석으로 수행한 결과를 보여준다. 부모마디의 평가용 결과를 확인해보면, 취업자수(=1)는 44.61%, 미취업자수(=2)는 55.39%로 취업자보다 미취업자가 더 많은 것으로 나타났다. 먼저 전체 취업률 44.61% 보다 높은 취업률의 집단을 보면 2년제 또는 전문대학을 졸업한 집단(=6)과 대학원 석사를 졸업한 집단(=8), 4년제 대학을 졸업한 집단(=7)으로 나타났다. 2년제 또는 전문대학을 졸업한 집단(=6)과 대학원 석사를 졸업한 집단(=8)의 취업률은 61.98%로 가장 높았고 대학을 졸업한 집단(=7)의 취업률도 48.14%를 보였으나 전체 취업률 44.61%보다 조금 높은 수준이었다. 나머지 고등학교를 졸업한 집단(=5)과 대학원 박사를 졸업한 집단(=9)은 30.64%의 취업률을 보였으며 고등학교 이상을 나오지 않은 집단(=4 또는 결측)의 취업률은 2.45%로 극히 낮았다.

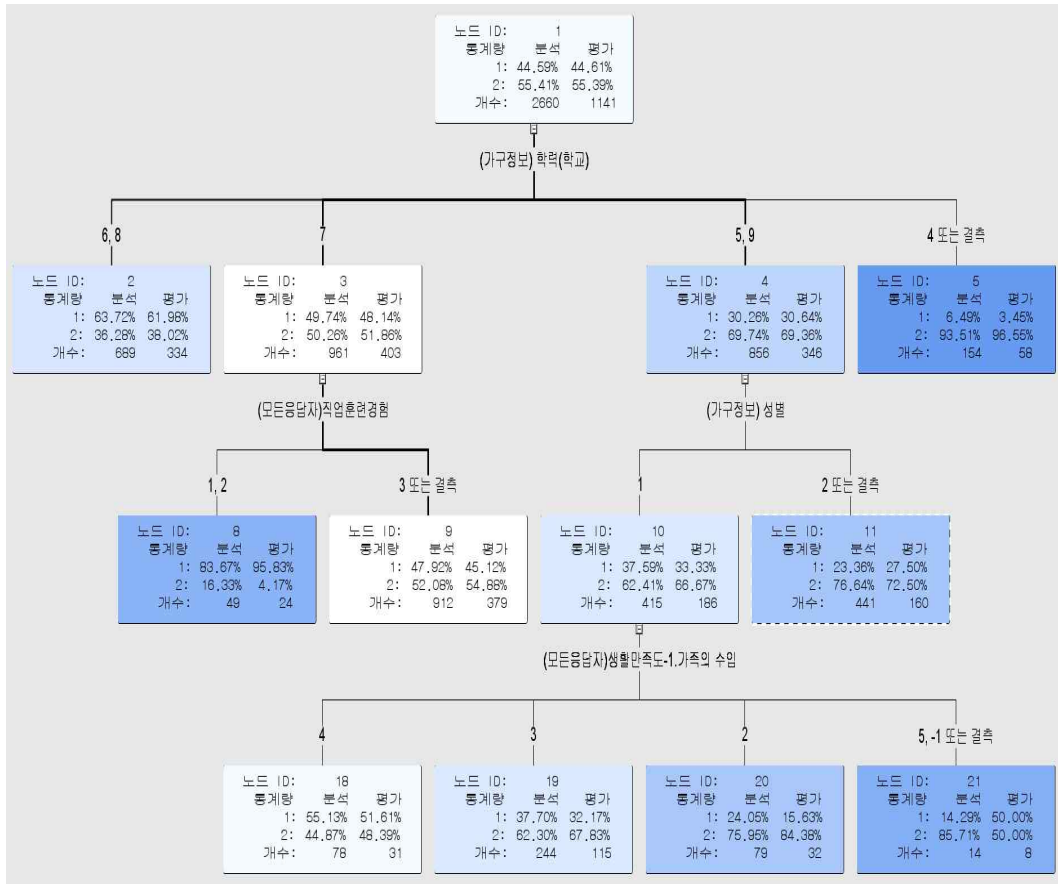
2년제 또는 전문대학을 졸업한 집단의 취업률이 높은 것은 이들의 직업탐색활동이 진로 결정수준을 높여 취업에 영향을 주었기 때문으로 볼 수 있다(황정희·임석빈 2015). 석사를

졸업한 집단의 취업률이 높은 것은 최근 청년취업난 속에서 대학원 석사과정에 진학하는 비율이 늘어난 것과, 대학원 진학을 취업의 발판으로 삼으려는 경향, 그리고 석사를 졸업한 후 대부분이 정규직이 아닌 상용직에 종사하거나 상위 과정에 진학하였기 때문으로 볼 수 있다. 이영민·임정연(2014)은 석사 취득자의 87.5%가 취업경로에 있으며, 그 이유가 대부분 상용직으로 근무 중이기 때문이라고 설명한다.

4년제 대학을 졸업한 집단을 보면, 직업훈련 경험을 보유하거나 훈련중인 경우 95.83%의 취업확률을 보였으며 훈련을 받지 않은 경우에는 취업률이 45.12%로 2배 이상의 확연한 차이를 보여 4년제 대학을 졸업한 집단의 경우 직업 훈련 경험 보유가 취업여부에 큰 영향을 미친다는 것이 뚜렷하게 나타난다. 그러나 앞서 연구대상자의 특성파악에서 직업훈련 참여율은 매우 저조한 것으로 나타났는데 취업률 향상을 위해서는 직업훈련 참여율을 재고할 필요가 있어 보인다.

고등학교를 졸업하거나 대학원 박사과정을 졸업한 집단의 경우 남성의 취업률이 33.33%로 27.60%를 보인 여성보다 높은 것으로 나타났다. 김태기 외(2005)에 따르면 이 같은 차이는 큰 차이는 아니지만 여성의 경우 취업에 있어 남성에 비해 정보 획득 능력이 부족하고 인맥관리와 같은 네트워크 구축에 어려움이 있기 때문으로 볼 수 있다. 또한 고등학교를 졸업하거나 대학원 박사과정을 졸업한 남성은 대체적으로 가족수입에 대한 만족도가 낮을수록 높은 취업률을 보이고 있다. 가족수입에 대한 만족도가 낮을수록 취업에 대한 절박함이 높기 때문임을 유추해 볼 수 있다. 가족수입에 대한 만족도가 보통인 경우의 취업률은 32.17%로 만족인 경우인 15.63%에 비해 2배 이상 취업률이 높았다. 만족도가 불만족인 경우의 취업률은 51.61%로 만족인 경우보다 약 3.3배 높은 것으로 나타났다. 가족수입에 대한 만족도가 매우 불만족인 경우의 취업률은 50.0%로 불만족인 경우보다 낮지만 경미한 수준이다.

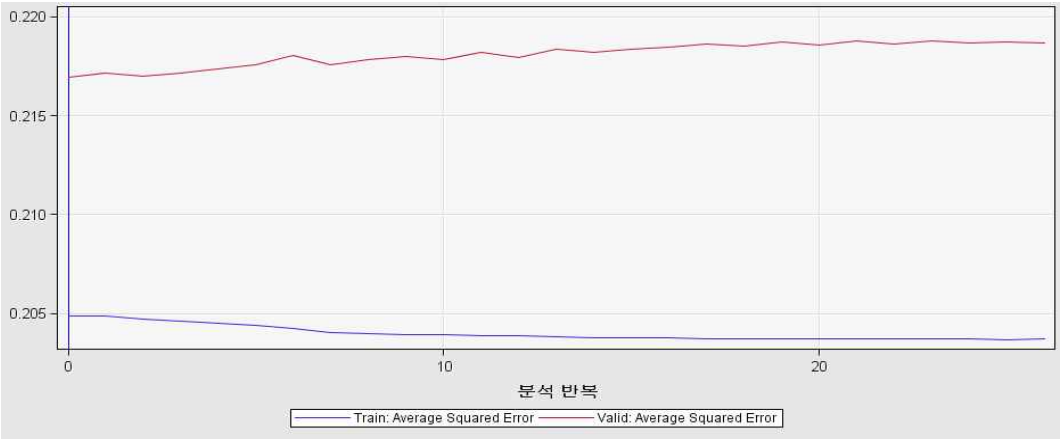
〈그림5〉 의사결정나무 분석에서 취업영향 요인



2. 신경망 분석을 활용한 청년층의 취업여부 결정요인 분석

〈그림6〉은 신경망 분석의 평균 제곱 오차(Average Squared Error)의 변화를 보여준다. 그림을 통해 26회의 반복적 최적화가 진행되었음을 알 수 있다. 자세히 살펴보면 반복 회수에 따라서 분석용(Training)과 평가용(Validation) 데이터에 대한 오차함수의 값을 살펴볼 수 있다. 분석용 데이터의 오차함수 값은 반복 횟수가 증가함에 따라 감소하고 있으나 평가용 데이터의 경우 시작과 함께 점차 증가한 것을 볼 수 있다. 신경망 노드에서는 평가용 데이터가 최소의 오차함수 값을 갖는 반복지점의 추정치를 선택하게 되는데 분석 결과 반복 회수 0에서 오차함수 값이 최소가 된 것으로 파악된다.

〈그림6〉 신경망 분석의 평균 제곱 오차 결과



신경망 분석의 결과 최종적인 모수들의 추정치들을 파악할 수 있었는데, 설정한대로 한 개의 은닉층에 4개의 은닉마디를 가진 다층 인식자 모형이 수행되었음을 〈표9〉를 통해 확인할 수 있다.

〈표9〉 신경망 분석의 출력 결과

N	Parameter	Estimate	Function
1	p1545011_H11 (직업훈련 경험이 있는 첫 은닉층 첫 번째 변수)	-2.480758	-0.000009945
2	p1545012_H11 (직업훈련 경험이 없는 첫 은닉층 첫 번째 변수)	1.283990	-0.000066635
3	p1545011_H12 (직업훈련 경험이 있는 첫 은닉층 두 번째 변수)	-0.774162	-0.000011820
4	p1545012_H12 (직업훈련 경험이 없는 첫 은닉층 두 번째 변수)	1.114523	-0.000030114
5	p1545011_H13 (직업훈련 경험이 있는 첫 은닉층 세 번째 변수)	-2.064440	0.000126
6	p1545012_H13 (직업훈련 경험이 없는 첫 은닉층 세 번째 변수)	-0.340644	0.000123
7	p1545011_H14 (직업훈련 경험이 있는 첫 은닉층 네 번째 변수)	0.881100	0.000015333
8	p1545012_H14 (직업훈련 경험이 없는 첫 은닉층 네 번째 변수)	-0.085463	0.000017920
9	BIAS_H11 (첫 은닉층 첫 번째 편향)	1.318463	0.000032418
10	BIAS_H12 (첫 은닉층 두 번째 편향)	3.049366	-0.000000851

본 결과는 81개의 모수가 존재하여 이를 반영한 실제 예측식을 작성하기에는 너무 복잡하므로 몇 개의 변수에 대해 간단히 요약한 결과를 보여주었으며 p_1545011와 p_1545012는 직업훈련 경험 보유 여부를 나타낸다. 예를 들어 p1545011_H11는 직업훈련 경험이 있는 경우 직업훈련 변수가 은닉층 첫 번째 첫 변수에 미치는 영향을 의미한다. 따라서 이 결과를 이용하여 직업훈련 경험 변수의 반응확률에 대한 예측식을 나타내면 다음과 같다(f는 쌍곡탄젠트 함수이다).

$$H_{11} = f(1.318463 - 2.480758 * \text{직업훈련 참여} + 1.083990 * \text{직업훈련 미참여})$$

$$H_{12} = f(3.049366 - 0.774162 * \text{직업훈련 참여} + 1.114523 * \text{직업훈련 미참여})$$

$$H_{13} = f(-1.752029 - 2.064440 * \text{직업훈련 참여} - 0.340644 * \text{직업훈련 미참여})$$

$$H_{14} = f(0.767388 + 0.881100 * \text{직업훈련 참여} - 0.085463 * \text{직업훈련 미참여})$$

이 예측식에 의하면 은닉층 첫 번째 마디에서 직업훈련 참여는 2.480758가 감소하는 만큼 청년층의 취업에 영향을 주며 직업훈련 미참여는 1.083990가 증가하는 만큼 영향을 준다고 할 수 있다. 은닉층 두 번째 마디에서 직업훈련 참여는 0.774162가 감소하는 만큼, 직업훈련 미참여는 1.114523가 증가하는 만큼 영향을 준다. 은닉층 세 번째 마디에서의 직업훈련 참여는 2.064440가 감소하는 만큼 직업훈련 미참여는 0.340644가 감소하는 만큼 영향을 준다. 은닉층 네 번째 마디에서 직업훈련 참여는 0.881100가 증가하는 만큼 직업훈련 미참여는 0.085463만큼 감소하는 만큼 영향을 주는 것으로 나타났다.

3. 모델 비교 분석 결과

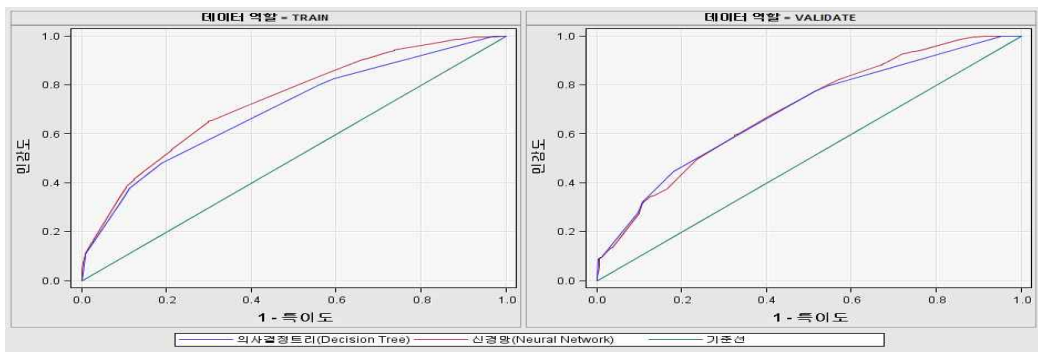
본 장에서는 위에 언급한 의사결정나무 분석, 신경망 분석에 대한 모델비교를 통해 취업 여부 결정요인 분석에 대한 가장 적합한 예측 모델을 설명하고자 한다. 본 연구에서 분석용은 70%, 평가용은 30%로 설정하였다.

〈그림7〉는 ROC(Receiver Operation Characteristic) 곡선으로서 판별분석에서 예측모델의 성능을 평가하기 위해 사용되는 것이다. 그래프에 나타나는 곡선이 볼록렌즈 모양처럼 두껍게 나타날수록 더 좋은 결과임을 나타낸다. 그래프에서 파란색 곡선이 의사결정나무 분석에 해당하며 빨간색 곡선이 신경망 분석 그리고 초록색 곡선이 기준선에 해당한다. 분석용 데이터와 평가용 데이터 경우 모두 모델 간에 큰 차이는 없으나 신경망 분석이 의사

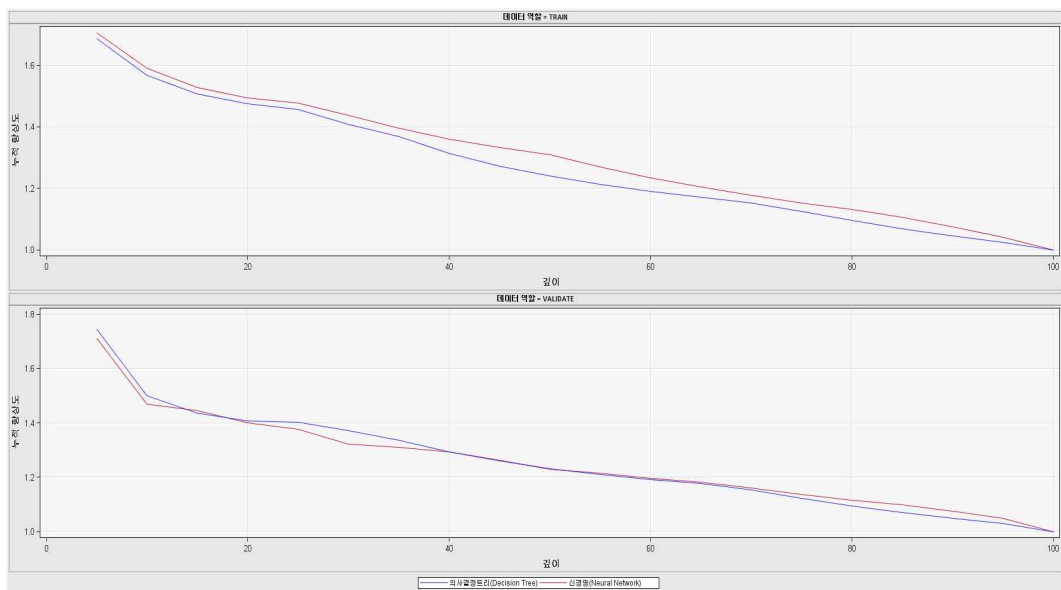
결정나무 분석에 비해 약간 좋은 편이라는 것을 볼 수 있다.

다음으로 모델 비교를 위해 누적향상도, 누적 반응률, 누적 반응 검출률을 살펴 볼 수 있는데 <그림8>은 평가용과 검증용의 누적향상도 그래프를 나타낸다. 그래프를 보면 상위등급에서는 신경망 노드의 향상도가 의사결정나무 노드보다 더 높고, 하위등급에서는 의사결정나무 노드가 다소 높게 위치 한 것으로 파악된다. 따라서 신경망 노드 예측모형이 더 좋은 성능을 갖고 있다고 할 수 있다.

<그림7> 모델비교 결과: ROC곡선



<그림8> 모델비교 결과: 스코어 순위 중첩그래프



〈표10〉은 세 모델의 적합 통계량 결과를 보여준다. 평가용 데이터의 오분류율 (misclassification rate), ROC 인덱스, Kolmogorov-Smirnov 통계량의 결과를 비교하여, 어떤 모델이 좋은 모델인지 파악할 수 있다.

〈표10〉 모델 적합 통계량

	의사결정나무 분석	신경망 분석
오분류율	0.36	0.37
ROC 인덱스	0.69	0.70
Kolmogorov-Smirnov	0.26	0.27

오분류율의 경우 낮을수록 좋은 모델이므로 0.37의 값을 갖는 신경망 분석은 0.36의 값을 갖는 의사결정나무 분석에 비해 약간 좋은 편이라고 할 수 있다. ROC 인덱스는 그 값이 높을수록 좋은 모델이라 할 수 있는데 마찬가지로 신경망 분석의 값이 0.70으로 0.69를 갖는 의사결정나무 분석에 비해 약간 좋은 모델이라고 평가할 수 있다. 마지막으로 Kolmogorov-Smirnov의 값도 높을수록 더 좋은 평가를 받게 되며 신경망 분석의 값이 0.27, 의사결정나무 분석의 값이 0.26로 앞의 두 통계량과 마찬가지로 신경망 분석이 더 좋은 편이라 할 수 있다.

〈표11〉은 모형평가 결과를 요약하고 있다. 분석용과 평가용으로 나누어 두 모델의 결과를 보여준다. Tree는 의사결정나무 분석, Neural은 신경망 분석을 나타낸다. 통계 결과를 확인해 보면, 오분류율은 분석용의 경우 신경망 분석이 평가용의 경우 의사결정나무가 더 좋게 나타났다. ROC 인덱스, Kolmogorov-Smirnov의 값은 분석용과 평가용 모두 신경망 분석이 더 좋은 모델인 것으로 분석되었다. 누적 반응률과 누적 반응 검출률은 분석용은 신경망 분석이 평가용은 의사결정나무 분석이 더 높은 것으로 나타났다.

〈표11〉 모형평가 결과의 요약

	역할	Tree	Neural
오분류율	분석용	0.36	0.33
	평가용	0.36	0.37
ROC 인덱스	분석용	0.70	0.73
	평가용	0.69	0.70
Kolmogorov-Smirnov	분석용	0.29	0.35
	평가용	0.26	0.27
상위 10% 누적 반응률	분석용	86.88%	88.10%
	평가용	83.07%	81.33%
상위 10% 누적 반응 검출률	분석용	15.68%	15.90%
	평가용	15.11%	14.80%

IV. 결 론

본 연구는 경제문제를 벗어나 심각한 사회문제로 자리 잡은 청년층 취업문제와 관련하여 청년들의 개인특성, 가족환경, 취업준비 중 어떠한 요인이 청년취업에 영향을 주고 있는지 데이터마이닝 기법을 활용하여 실증 분석한 것이다. 데이터마이닝 기법이란 방대한 자료들을 정제하여 의사결정에 까지 이르게 하는 기술이다. 데이터마이닝을 위해 한국 노동연구원에서 제공하는 제15차년도 노동패널(2012년)에서 만 15세 이상 만 34세 이하의 청년층 데이터를 추려내었으며 분석에는 SAS사의 E-MINER 14.1 프로그램이 사용되었다.

분석결과 개인특성에서는 성별과 학력, 가족환경에서는 가족수입만족도, 취업준비에서는 직업훈련이 청년취업 결정요인으로 작용하는 것으로 나타났다. 먼저 의사결정나무 분석결과 청년층의 취업 여부에 학력 영향력이 가장 크고 그 다음으로 직업훈련 여부가 큰 것으로 나타났다. 주목해야 할 것은 2년제 또는 전문대학을 졸업한 집단과 대학원 석사를 졸업한 집단의 취업률이 4년제 대학을 졸업한 집단보다 높다는 것이다. 반면 고등학교 졸업 이하의 학력집단은 매우 낮은 취업성공률을 보였다. 이를 통해 취업에 있어서는 전문대학에 진학하는 것과 대학원 석사과정을 졸업하는 것이 유리함을 알 수 있다. 흥미로운 사실은 4년제 대학을 졸업한 경우 직업훈련 경험을 갖고 있으면 95.83%의 취업성공률을 보인다는 것이다. 역대 정부의 청년실업대책을 살펴보면 청년인턴제도 등을 통해 단기일자리 마련을 최우선으로 두고 그 다음으로 직업훈련에 대한 내용이 포함되어 있는데 본 연구결과로 보

자면 4년제 대학 재학생이나 졸업자를 대상으로 직업훈련을 확대하고 적극적인 참여를 유도할 필요성이 있어 보인다. 한편 고등학교를 졸업하거나 대학원 박사과정을 졸업한 남성 은 가족수입에 대한 만족도가 낮을수록 높은 취업률을 보이고 있다. 가족수입에 대한 불만족이 취업의지를 높게 하였기 때문으로 예상해 볼 수 있다. 또한 이 집단에서 남성이 여성보다 더 높은 취업성공률을 보이고 있는데 이는 선행연구에서 여학생들이 네트워크 등의 부족으로 취업에 어려움이 있다는 지적과 일치한다.

신경망 분석결과도 의사결정나무 분석과 마찬가지로 입력변수들의 결합을 수신하여 목표변수에 보내는 역할을 하는 은닉층 첫째, 둘째, 셋째, 넷째 마디에서 직업훈련이 청년취업에 영향을 주는 것으로 나타났다. 의사결정나무 분석방법은 개별적으로 독립변수가 종속변수에 미치는 효과를 파악할 수는 없지만 신경망 분석방법은 은닉마디를 통해 입력변수와 출력변수와의 관계를 설명할 수 있다. 그러나 다양한 모형들과 검정과 테스트용 자료가 조사되어야 하며 마지막으로 최적의 모형이 선택되어야 한다. 또한 설명력은 높으나 해석이 용이하지 않다는 단점을 갖고 있다.

이제까지 살펴본 결과 우리는 학력과 직업훈련 등이 청년취업에 중요하며 자격증 보유 여부, 부모의 학력은 취업에 영향을 주지 않음을 알 수 있었다. 이는 선행연구와 일치한 결과이다. 정부의 청년실업 정책을 살펴보면 현장연수, 직업훈련 등이 빠짐없이 포함되어 집행되었음을 확인할 수 있었는데 이것들이 얼마나 효과적으로 행해졌는지에 대한 철저한 분석이 필요하다. 또한 최근 청년들 사이에서 취업에 있어 부모의 학력과 재력 등의 영향으로 기회의 균등이 보장되지 않는다는 인식이 퍼지고 있는 것과 관련해서 부모소득에 대한 만족도가 취업에 있어 일부 영향을 끼치고 있음을 확인하였다. 따라서 청년 취업대책 마련 시 가구소득이 낮은 계층의 청년을 대상으로 차별화된 정보제공과 직업훈련 프로그램이 집중될 필요가 있음을 시사한다. 마찬가지로 성별도 청년취업에 영향을 일부 미치고 있음을 확인하였다. 특히 여성 청년층이 취약한 연결망을 보완해 주어 보다 용이하게 여성 청년들이 취업관련 정보와 훈련이 가능하도록 성주류화적인 취업 프로그램이 절실하다.

향후 발표되는 청년취업 대책들에는 위와 같은 부분들이 충분한 논의를 거쳐 반영되어야 할 것이다. 그러나 무엇보다 중요한 것은 청년실업으로 인해 고통 받고 있는 청년들을 위한 따뜻한 사회적 분위기를 조성하는 것이다. 청년들의 비관과 상처를 감싸 안아 주고 그들에게 희망을 불어 넣어 주도록 낮은 임대료의 주택 서비스, 모금활동과 재능기부를 통한 자문활동 등 사회 각층의 참여가 필요하다.

【참고문헌】

- 강진웅 외. 2011. “의사결정나무와 신경망 모형 결합에 의한 운전자 우회결정요인 분석”, 〈한국도로학회논문집〉: 169-171.
- 강현철 외. 2015. 〈데이터마이닝 방법론〉, 파주:자유아카데미.
- 국회예산정책처. 2016. 〈경제동향 & 이슈〉 42호: 24.
- 금재호. 2007. “청년실업의 현황과 원인 및 대책”, 〈숭실대사회과학논총〉 9집: 45-48.
- 김난도. 2015. 〈트렌드코리아 2016〉, 서울:미래의 창.
- 김민정·류지영. 2011. “대학졸업자의취업에 영향을 미치는 요인 분석”, 〈제6회 한국교육고용패널 학술대회 논문집〉.
- 김선주·이상엽. 2008. “오피스 임대료 결정모형에 관한 연구”, 〈지역연구〉 제24권 2호: 6.
- 김성훈. 2015. “대학생의 학업 외 취업 준비의 영향 요인과 취업 성과”, 〈산업노동연구〉 제21권 3호: 117-118.
- 김성희. 2011. “역대 정부 청년고용대책 평가”, 〈민주당정책위원회 토론회〉.
- _____. 2015. “박근혜 정부 노동개혁과 청년실업”, 〈노동연구〉 제31권: 9.
- 김안국·유한구. 2014. 〈대학 및 전문대학 졸업자의 직종별 수요추정〉, 한국직업능력개발원.
- 김영숙. 2009. “아동사례관리의 욕구사정 정확도 향상을 위한 사정도구 개발과 욕구추출 알고리즘 과정 연구, 데이터마이닝 분석기법을 활용하여”, 〈사회복지정책〉 26(2): 167-189.
- 김용성. 2012. “고학력 청년층의 미취업 원인과 정책적 대응방안”, 〈한국개발연구〉 34: 70-72.
- _____. 2013. “고용률 제고를 위한 재정지원 일자리사업 방향의 모색”, 〈KDI 정책연구〉 2013-19: 26.
- 김형래·전도홍. 2012. “데이터마이닝 기법을 활용한 대졸자 고용에 미치는 영향요인 분석”, 〈한국컴퓨터정보학논문지〉 17(7): 170.
- 공수연. 2012. “청년구직자의 취업준비 과정에 대한 탐색적 분석”, 〈한국가정관리학회지〉: 48.
- 서울노동권익센터. 2015. “통계로 본 서울의 노동”, 〈동향과 이슈〉 2015-1: 12.
- 서울노동권익센터. 2015. “청년고용의 대안과 서울시 청년수당의 의미”, 〈동향과 이슈〉 2015-3: 4.
- 서울시 청년허브. 2015. “서울이 만드는 다음 세대의 여지 청년이 만드는 다음 사회의 공공”, 〈청년정책의 재구성 기획연구〉: 76.
- 성운현. 2000. “대학생의 경력준비행동”, 〈경영경제연구〉 23: 113-134.
- 송태민·송주영. 2015. 〈빅데이터연구 한권으로 끝내기〉, 서울:한나레아카데미.
- 신현욱. 2011. “데이터마이닝 기법을 이용한 지체장애인 취업결정요인 분석 연구”, 〈장애와 고용〉 21(3): 119-152.

- 어수봉 외. 2010. <대졸과 취업결정 요인분석>, 한국기술교육대학교: 117.
- 이수영·이영민. 2009. “이공계 대학생의 진로결정수준 및 진로준비행동 분석”, <한국공학교육학회지> 2: 73-79.
- 이병희. 2004. “청년층 노동시장 구조변화”, <동아시아연구소 주최 ‘고용없는 성장기의 청년실업 해법 토론회’ 발표논문>.
- 이영민·임정연. 2014. “일반대학원 석사졸업생의 취업현황 및 임금결정 요인분석”, <2014 고용패널 학술대회 자료집>.
- 이영민 외. 2014. “대학생 선호직장의 취업 결정요인 분석”, <사회과학연구> 53: 340-342.
- 이혜주·정의현. 2013. “데이터마이닝을 이용한 학업성취 결정요인 탐색”, <아동교육> 22(2): 6-8.
- 임은정 외. 2013. “데이터마이닝 기법을 이용한 시각장애인의 취업결정요인 분석연구”, <장애와 고용> 23(1): 274-275.
- 장외경. 1997. “다차원시계열 자료분석을 위한 신경망 모형에서 통계적 접근에 의한 적절한 구조 선택”, <통계분석연구> 제2권 제2호: 69-93.
- 진현정. 2010. “의사결정나무모형을 이용한 소비자 그룹 분석”, <한국농업경제연구지> 51(1): 19-40.
- 채창균. 2003. “대학생의 재학 중 일자리 체험과 취업준비”, <직업교육연구> 22: 169-184.
- 한국보건사회연구원. 2015. “한국사회의 사회 심리적 불안의 원인분석과 대응방안”, <연구보고서> 2015-22: 186.
- 한금윤. 2015. “대학생의 연애, 결혼에 대한 의식과 문화 연구”, <인간연구> 28: 24.
- 감사원. 2015. <재정지원 일자리 사업 추진실태>.
- 경향신문. “한국 2030세대 부모보다 가난한 첫 세대 됐다”, 2016.3.8.
- 경향일보사. <주간경향> 1157호, 2015.12.
- 고용노동부. 2015. <한권으로 통하는 2015 청년고용정책>.
- 국가법령정보센터. <청년고용촉진 특별법>.
- 동아일보. “사업주에 주던 고용보조금, 청년 취업자에 직접 주기로”, 2016.2.29.
- 서울신문. “청년실업 경제성장률 까지 낮춘다”, 2015.10.22.
- 중앙일보. “농담인데 불편하네 ‘수저계급론’”, 2015.10.28..
- 통계청. 1997. <통계분석연구>.
- 통계청. 2016. “2016년 1월 고용동향”.
- 통계청. 2016. “2015년 4분기 및 연간 가계동향조사”.
- 한국노동연구원. 2012. “15차년도 노동패널 조사”.
- 한국노동연구원. 2012. “15차년도 노동패널 조사 개인용 코드북”.

- Brethenoux, E., H. Drenser, K. Strange, J. Block. 1996. Strategic data management: Data warehouse, Data mining and Business intelligence: The hype stops here. Gartner Group Report, SDM: R-300-105.
- Chmid, Gunther. 2013. "Youth Unemployment in Korea: From a German and Transitional Labour Market Point of View". IZA policy paper no. 63.
- Vandamme, J. P., Meskens, N., & Superby, J. F. 2007. Predicting academic performance by data mining methods. *Education Economics*, 15(40), 405-419.
- OECD Health Data 2015.
- OECD Employment outlook 2015.
- OECD Education at a Glance 2015.

한 성 민: 성균관대학교 국정전문대학원 석박통합과정에 재학 중이다. (사)해위윤보선대통령 기념사업회에 재직하고 있으며 (사)한국청년유권자연맹 운영위원으로 활동하고 있다. 관심분야는 청년 및 사회정책, 사회자본, NGO, 민주시민교육, 빅데이터 분석 등이다. 주요논문으로는 "민주화세대 투표행태의 지속과 변화"(2013)가 있다(clon96@skku.edu).

문 상 호: 미국 University of Wisconsin-Madison에서 경제학 박사학위를 받은 후 테네시 주립대학교(Tennessee State University) 정책학 조교수로 근무하였다. 현재 성균관대학교 행정학과/국정전문대학원 교수로 재직 중이며 주요 관심분야는 재정학, 의료와 복지, 정책계량분석 등이다. 최근 논문으로 "노인 1인가구의 취업결정요인 분석: 탐색적 데이터마이닝의 활용과 적용(한국정책학회보, 2016)", "데이터마이닝 분석방법을 활용한 고령자의 자살위험 예측요인 분석(한국정책학회보, 2016)", "한국 장애인고용정책의 효과성 평가(한국정책학회보, 2016)", "근로장려세제(EITC)의 정책효과에 관한 연구(한국정책학회보, 2015)", "여성노인의 노동시장 참여에 대한 국민연금제도 유인효과와 계급효과(사회과학연구, 2015) 등이 있다(smoon@skku.edu).

이 숙 중: 미국 Harvard University에서 사회학 박사학위를 취득하고 현재 성균관대학교 행정학과/국정전문대학원 교수로 재직 중이다. 주요 관심분야는 민주주의, 동아시아 거버넌스, 정치경제, 여론 연구 등이다. 최근 저서로는 "세계화와 제 2막"(2010, 동아시아연구원), "Public Diplomacy and Soft Power in East Asia"(2011, Palgrave Macmillan), "글로벌 개발협력 거버넌스와 한국"(2012, 동아시아연구원)이 있으며 논문으로는 "정부에 대한 공정성 인식이 정부신뢰에 미치는 영향: 정부성과 만족도의 매개효과에 대한 다중집단 분석을 중심으로"(2016), "한국사회 갈등의 원인 및 관리에 대한 연구: 갈등유형별 특성을 중심으로"(2016), "한국의 종교적 신앙 기반 NGO(FBO)에 관한 탐색적 연구: 종교 유형별 현황분석과 조직 유형화 논의를 중심으로"(2015) 등이 있다(skjlee@skku.edu).

논문투고일: 2016.5.14 / 수정일: 2016.6.17 / 게재확정일: 2016.6.27