

한국심리학회지: 건강
The Korean Journal of Health Psychology
2003, Vol. 8, No. 1, 133-146

중퇴에 관한 위험 및 보호요인의 신경망 모형

구본용 유제민[†]

강남대학교 교육대학원 교양학부

학업 중단 문제는 발달 과업의 성취에 부정적인 계기로 작용할 가능성이 크다. 본 연구에서는 중퇴를 야기시키는 개인, 가정, 사회, 학교 등의 전 영역을 포괄적으로 고찰하고자 그 관련 변인들을 위험 요인과 보호요인으로 구별한 후 신경망 모형을 구성하였다. 개발된 모형에 의하면 위험 요인이 증가할수록 중퇴 반응도 대체로 증가하고 있었지만 보호요인이 증가할수록 중퇴 반응이 감소하는 것은 일관된 경향이 아니었다. 모형은 비 선형적이었으며, 중학생보다는 고등학생 집단에 그리고 여학생보다는 남학생집단에 더 설명력이 있는 것으로 나타났다. 신경망 모형은 위험 요인과 보호요인의 시기적 변환을 파악할 수 있게 해주고 예방에 관한 맞춤형 접근을 개발하는데 유용하게 활용될 수 있다. 추후 연구에서는 중퇴의 유형과 지속 기간 등을 고려하고 장기적인 후유증을 추적하는 연구가 필요할 것으로 제안되었다.

주요어: 학교중도탈락, 위험요인과 보호요인, 신경망 모형

학생시절에 학업을 포기하고 학교를 떠나는 중도탈락은 학생들의 교육기회를 박탈하고, 진로 개발과 선택 및 사회성 기술의 발달 등 청소년 시기에 달성해야 할 발달과업을 성취할 수 있는 기회를 잃게 하며, 장기적으로는 안정적인 성인생활을 유지하는데 어려움을 주는 중요한 선행 요인이 되기도 한다(Natriello, Pallas, & McDill,

1986). 더욱이 가족해체 현상의 증가, 청소년 유해 환경의 확대, 입시 위주의 교육 및 극기·절제력의 부족 등으로 인한 부적응 학생의 지속적 증가 등으로 잠재적 학업중단 학생의 비율이 높아지고 있음에 따라 학업중단 문제는 학생 개인 혹은 가정, 학교 차원의 문제가 아니라 국가·사회적 차원에서 다루어져야 할 문제가 되었다.

[†] 교신저자(corresponding author): 유제민 (449-702) 경기도 용인시 기흥읍 구갈리 산 6-2강남대학교 교양학부, E-mail : pavlov2@kangnam.ac.kr

우리나라 중·고등학교에서는 지난 1980년도부터 최저 1.3%(1980년)에서 최고 2.1%(1988년)의 학생들이 학업을 중단하고 학교를 떠나, 매년 1,000명 가운데 13명 내지 21명이 중도탈락하였다(이혜영, 2001). 2001년에는 가정사정, 품행문제, 학교생활부적응 등으로 인한 초·중·고 학업 중단자가 55,528명에 달하여 전체 재학생 7,821,754명 가운데 0.71%를 차지하였다. 학업중단 중학생의 수는 좀처럼 감소하지 않았으나 실업계 고교생은 크게 감소하고 있고 유학·이민에 의한 학생은 크게 증가한다는 점이 최근의 추세이다(한국청소년상담원, 2002). 우리나라의 경우 OECD 국가 중에서는 학업중단 비율이 낮은 편이지만 학업중단 청소년을 위한 사회적 안전망이 미흡한 상황임을 고려하면 결코 낮은 비율이라고 볼 수 없다.

미국에서 구축된 전국적 규모의 교육관련 데이터베이스(National Education Longitudinal Study, Common Core of Data, High School & Beyond, Current Population Survey, National Center for Educational Statistics, 1994)에서는 중도탈락과 관련된 변인들을 행동적 범주(결석, 징계, 학교활동에의 저조한 참여 등), 심리적 범주(낮은 포부수준, 낮은 자아 존중감, 정서적 문제, 소외감 등), 학업적 범주(학습부진이나 학업수행상의 어려움), 사회경제적 범주(가정의 경제적 어려움, 부모의 낮은 교육수준, 편부모 가정 등), 환경적 범주(임신, 아르바이트 등 학교공부와 병행하기 어려운 기타 상황 변인들)의 다섯 가지로 구분하고 있다. 현재 학교중도탈락은 이러한 요인들이 복합적으로 작용하는 사회일탈 현상으로 간주되고 있다(교육인적자원부, 1996; Levine, 1986). 또한 어느 한 시점에서 발생하는 하나의 사건

(event)이라기보다는 학교 안팎에서 학생들이 일련의 긍정적, 부정적 경험을 겪으면서 일어나는 역동적이며 진행적인 과정(process)이라고 정의되기도 한다(Finn, 1989, Institute on Community Integration[ICI], 1994).

그럼에도 불구하고 지금까지의 선행연구들은 주로 중도 탈락생의 개별적인 특성이나 가정, 학교 및 지역사회 환경특성을 구체화하는데 초점을 두어왔다. 이러한 연구들을 통해 다수의 변인들이 각각 학교중도탈락에 어떠한 영향을 미치는지 밝혀졌지만 중도탈락과 관련된 많은 변인들 가운데 어떤 변인들이 어떻게 상호 작용하여 어떤 시기에 학교를 그만두게 되는지는 확인하기 어려웠다. 무엇보다도 중도탈락의 원인을 밝히기 위해서는 학생들의 중도탈락에 직·간접으로 영향을 미치는 다양한 변인들과 이들 변인들간의 조합적인 관계를 밝히는 것이 중요하다. 최근에 구본용, 신현숙 및 유제민(2002)은 학교중도탈락에 관한 이론을 토대로 중퇴를 결정하는 변인들을 개인, 가족, 학교, 또래, 지역사회 관련 영역으로 구분한 후 데이터 마이닝을 통해 성별(남, 여)과 시기별(중, 고)로 변화하는 상호작용 모형을 제안한 바 있다.

이들의 연구(구본용 등, 2002)에서 밝혀진 중학생의 학교중도탈락 모형에서 중학생의 중퇴유무를 가장 잘 변별해주는 변인은 자녀의 최종학력에 대한 부모의 교육적 기대수준이었으며, 어울리는 또래들에 대한 교사의 평가, 결석이나 지각으로 인한 수업결손의 변인 등 학업 수행에서 파생되는 일관된 변인들이 크게 작용하는 것으로 나타났다. 이에 비해 고등학생의 학교중도탈락에는 현재 재학중인 학교의 계열(인문계, 실업계),

흡연, 무기력 및 우울, 미래 진로에 대한 기대, 가정의 경제적 수준과 같은 심리·사회적 변인들이 다양하게 영향을 주고있었다. 그리고 남학생의 경우에는 학력에 대한 본인의 희망사항과 현실적인 실제상황과의 괴리가 학교중도탈락 유무를 변별해주는 강력한 변인으로 작용하고 있었다. 가족 갈등이나 친구 갈등은 이러한 불일치에서 파생되는 부차적인 현상이었다. 이 같은 결과는 남학생들이 학력을 중시하는 사회적 기대와 현실에서 크게 갈등하고 있으며, 학력에 대한 기대마저도 완전히 포기할 경우에는 학교중도탈락과 직결될 수 있음을 시사한다. 여학생의 경우에는 부모-자녀관계와 학년 변화와 같은 과도기의 부적응이 중퇴에 크게 작용하는 변인들이었다.

현재 우리나라에서는 중퇴에 관한 체제적 접근은 어느 정도 이루어지고 있지만 임상적 접근은 대단히 부진한 상태이다(박창남 2001). 본 연구에서는 구분용 등의 연구결과(2002)를 중심으로 학교 중도 탈락을 일종의 심리, 사회적인 문제 행동(problem behavior)으로 규정하였고 상기한 다섯 가지 범주의 변인들을 위험 요인(risk factors)과 보호 요인(protective factors)으로 양분하였다. 위험요인은 문제 행동이 발생할 가능성을 높여주는 모든 조건이나 상황을 말한다(Sroufe, Cooper, & De Hart, 1992). 단일한 위험요인으로 추후의 문제 행동을 예언하기란 쉽지 않지만 여러 개의 위험요인들이 중복되면(cumulative) 부 적응적 증상으로의 발달을 더 신뢰 톱게 예언할 수 있다. 이에 비해 증상으로의 발달을 억제하고 건강한 발달을 유지, 촉진시키는 조건이나 상황을 보호요인이라고 지칭한다(Masten, Best, & Garnezy, 1990). 특정한 외상적 사건이 모든 개인에게 동일

한 부 적응적 행동을 유발하는 것이 아니다(Pynoos, Steinberg, & Wraith, 1995). 특정 시기에서 각 개인에게 작용하는 위험 요인과 보호요인의 불균형적 배합은 적응 혹은 부적응의 이질적인 결과로 나타날 수 있다(Cicchetti & Rogosch, 1996). 따라서 중퇴에 관한 연구는 잘 알려진 위험요인들이나 몇몇 보호요인의 독립적인 작용만으로는 그 현상을 설명하기가 어렵다(Battin-Pearson et al., 2000). 그러나 중퇴 관련요인들을 위험 및 보호 요인으로 구분하여 특정 시기나 구체적 조건에서의 중퇴와 연결시키려는 시도는 이론적으로는 시사하는 바가 크지만 방법론적으로는 대단히 복잡한 문제일 수 있다(Rumberger, 1995).

본 연구는 선행한 구분용 등(2002) 연구를 근거로 하고있다. 데이터 마이닝 연구는 일반적으로 두 단계로 진행된다. 의사 결정 나무 분석(decision tree analysis)에서는 변인의 축소와 요약 그리고 그 변인들의 상호 작용을 분석하게 되고, 인공 신경망(artificial neural network) 분석에서는 그 축소된 변인들을 중심으로 자가 수정된 최적의 예측 모형을 구성하기 위해 수행된다. 본 연구자들은 중퇴 관련 변인 수가 158개나 되기 때문에 이 절차가 필수적이라고 판단하였다. 인공 신경망은 의사 결정 나무분석같이 그 진행 단계를 잘 파악하기 어렵다는 문제가 있기는 하지만 연구자의 입장에서 가용한 소수의 한정된 변인들로 결과의 효율성을 극단적으로 최대화시키는데 큰 장점이 있다. 의사 결정 나무분석에서는 특정한 변인의 특정한 조건 구간에서의 중퇴 가능성이 순서도 방식으로 제시되지만 신경망 분석에서는 특정한 변인의 바로 그 조건에서의 중

되 확률이 실시간으로 비교되는 구체성이 있다.

이 기법은 CRM을 근간으로 광범위하게 적용되고 있고 특히, 건강정보학(Health Informatics) 분야에서 대단히 유망하게 활용되고 있다. 후자의 경우 잘 알려진 건강 관련 위험 요인과 보호 요인들이 보편적으로 적용되지 않을 수 있다는 점을 전제해 이 기법을 이용하여 개별화된 맞춤형 치료를 시도하고 있다. 최근에 당뇨병의 예측(Park & Edington, 2001), 외상 후 진행과정에서 나타나는 병발성의 확인(Marble & Healy, 1999), 임상적 경과의 시계열 분석(Lin, Chou, Pan, & Chen, 2001)등에서 이러한 시도의 유용성이 확인되었다.

본 연구에서는 선행 연구에서 요약된 변인들을 중퇴의 위험 요인과 보호요인들이 지속적으로 상호작용 한다는 선행 연구 결과를 토대로 이를 X축과 Y축의 양 지표 상에서 각각 추정하게 한 후 그 교차지점에서 중퇴 예측율(z)을 모형화하고자 하였다. 이러한 방식의 분석 절차를 통해 위험 및 보호요인의 상호 작용과 상대적 영향력 그리고 변화 양상과 누적된 효과를 수량화시키는 것이 가능하다. 이 모형은 중퇴 위험 군의 변별 및 맞춤형 예방 전략을 구체화하기 위한 추후 목적을 염두에 두고 구성한 것이다.

방법 및 절차

연구대상

연구대상은 경기도 소재 중·고등학교에 재학하는 도중에 학교를 중도 탈락한 청소년 291명과 경기도에 위치한 중·고등학교에 재학하고 있는 학생 374명을 포함하여 총 665명이었다. 연구 대

상의 구성은 아래와 같다.

	성별구분		Total
	남자	여자	
현재 또는 중학교	157	62	219
중단시 다 인문계 고등학교	202	117	319
나던 학교 실업계 고등학교	85	42	127
Total	444	221	665

자료의 수집

경기도 교육청으로부터 통보 받은 중도탈락 학생의 주소록에 기초하여, 2001년 8월 20일부터 8월 31일 사이에 이들의 거주지 주소로 1,010부의 설문지를 우편 발송하였다. 설문지 회수율이 저조하여(65부 우편 회수), 2001년 9월 3일부터 10월 30일 사이에 중도 탈락생을 대상으로 2차 조사를 실시하였다. 2차 조사에서는 본 연구자들로부터 면접절차에 관한 사전훈련을 받은 경기도 내 청소년상담 관련자들과 청소년 관련학과 대학생들이 중도탈락 청소년들을 대상으로 전화면접과 대면면접을 실시하여, 226부에 해당하는 자료를 추가로 수집하였다.

재학생 대상의 조사는 경기도 내 중학교와 고등학교에 재학중인 400명을 대상으로 2001년 11월에 학급별로 담임교사나 교과담당교사에 의한 집단설문조사의 형태로 이루어졌다. 학생들은 자기 보고형 질문지를 자신의 속도대로 읽고 답하도록 하였다. 응답에 시간제한을 두지 않았으며, 전체 질문지를 완성하는데 소요된 시간은 약 30분 정도였다. 재학생 대상으로 수집된 자료 가운데 응답에 누락이 있거나 불성실한 응답이 있는 설문지 26부의 자료는 분석에서 제외되었다.

변인의 선정

본 연구의 독립변인들은 인구 통계학적인 상태들과 심리·사회적인 제반 변인들이며 종속변인은 중퇴유무(재학생/중퇴자 구분)였다. 본 연구에서는 명명 척도와 서열 척도 등이 혼합된 158개의 독립변인들을 데이터 마이닝한 구분용 등(2002)의 연구 결과를 토대로 중/고/남/여학생의 의사결정 나무모형에서 나타난 변인들을 일차적으로 선정하였다. 여기에 구분용 등(2002)의 연구 결과에서는 나타나지 않았으나 전통적으로 확인되어온 중퇴 관련 위험요인과 보호 요인들(McWhirter et al., 1998; Rumberger et al., 1995; Krohn et al., 1995; Janosz et al., 1997)을 실증적 연구 결과들을 토대로 추가하였다. 이 과정을 통해 최종적으로 선정된 변인들은 개인적 범주(교육적 기대 수준, 우울 정도, 자기 존중감, 자기 유능감, 10년 후 희망 직업, 흡연정도, 음란물 시청정도, 임신여부), 가족 범주(동거 가족, 부모의 자녀 교육 기대 수준, 가정의 경제적 수준, 부부 관계, 부모-자녀 관계), 또래 범주(교우와의 관계, 친구에 대한 교사의 평가), 학교 범주(학업성적, 무단결석, 지각 빈도, 재학 시 다니던 학교, 전학 경험)로 구분할 수 있었다.

본 연구에서는 이 20개의 변인들을 다시 위험요인과 보호요인으로 재분류하였다. 정의에 따라 의사 결정 나무 모형의 각 수준에서 중퇴의 가능성을 높여주는 변인을 위험요인으로 그리고 이 변인이 있는 조건에서 특정한 변인이 조합되었을 때 중퇴의 가능성이 감소하는 경우 보호요인으로 규정할 수 있다(구분용 등 2002). 이 기준에 따라 20개의 변인 중 보호요인 군에는 높은 교육적 기

대수준, 자기 존중감, 자기 유능감, 10년 후 희망 직업의 구체성, 갈등 없는 부모의 관계, 원만한 부모-자녀관계, 학업 성적의 7개 변인이 포함되었고, 나머지 13개 변인은 모두 위험 요인 군으로 분류되었다.

자료의 처리

본 연구에서는 위험요인과 보호요인을 동시에 투입하여 중퇴유무를 가장 효과적으로 변별하는 비 선형 방정식을 도출하고자 하였다. 투입된 변인들은 연속형과 명목형이 혼재되어있다. 또한 재학생들과 중퇴생들의 반응특성이 다소 이질적이었고 결측치들도 많았기 때문에 비모수 통계를 기초로 하는 데이터 마이닝을 실시하였다. 적용한 기법은 신경망 분석(neural networks)이며, 사용 프로그램은 Neural Connection 3.0 v (2002)이었다.

신경망 모형은 요인간 또는 그 하위 측정 변인들간에 발생할 수 있는 모든 상관유형을 함수에 포함시키는 병렬형 방식이며, 인간두뇌의 학습 과정을 시냅스 작용을 통한 연산과정으로 재현한 데이터 마이닝 기법의 일종이다. 이 방법은 실제 출력 값과 기대 출력 값을 비교하고 그 차이를 시냅스 역할을 하는 노드(node)에 가중치를 주어 조정된 후 그 과정을 신경망 구조가 안정화될 때까지 반복함으로써 예측 및 분류 모형을 자가 창출한다. 스스로 피드백을 주어가면서 수정하는 역전파망 방식(Back-Propagation Network)으로 경험을 통한 학습 과정을 기계적으로 단축시킨 분석 기법이다.

본 연구에서는 전체 665명의 자료중 모형 개

발(development)에 400명, 모형 예측(prediction)에 265명을 각각 무선 할당하였다. 모형 개발에 사용된 400명은 신경망 분석 절차에 따라 훈련(train) 표집에 70%, 검증(test) 표집에 20%, 타당화(validation) 표집에 10%씩 각각 무선 할당되었다. 모형 예측에 사용된 265명은 중학생과 고등학생, 남학생과 여학생의 집단으로 구분한 후 개발된 모형을 적용하는 절차를 거쳤다. 신경망 분석에서 자료를 분석하는 방법은 예측변인의 특성에 따라 일곱 가지가 있는데, 그 중 단층인식모형(Radial Basis Function; RBF)과 다층인식모형(Multiple Layer Perception; MLP)이 가장 많이 사용된다. RBF는 비 선형적 고차 함수를 산출하며, MLP는 신경망 구조에서 자료를 훈련시키는 과정에서 여러 개의 은닉 층을 거치면서 모형을 최적화시킬 때 사용하는 방법이다. 본 연구에서는 위험 요인과 보호 요인이 반복적으로 상호 작용하면서 비 선형적으로 중퇴에 영향을 미치고 있다고 가정하였기 때문에 RBF와 MLP를 통합하여 적용하였다.

결 과

중퇴와 관련되는 20개의 변인들은 13개의 위험 요인과 7개의 보호요인으로 분류된다. 인공 신경망 분석은 가설을 검증하는 것이 아니라 스스로 자료(데이터)를 적합시켜 나갈 수 있도록 조건을 설정해주는 사전 절차를 거쳐야 한다. 데이터는 이 조건(흐름도)을 따라 이동하면서 최종적인 결과를 최대한 유사하게 추정할 수 있는 가중치를 스스로 찾아내 이를 근거로 방정식을 도출하고 예측모형을 창출한다. 위험 및 보호 요인에 포함된 중퇴 관련 변인들은 오차거리를 Euclidean 분포로 변환(filter)하여 산출한 RBF에 따라 각각 중퇴의 유무를 예측하게 된다. 다음 단계에서는 결합도구(combiner)를 사용하여 위험 요인과 보호 요인을 통합한 후, 다 축 변환(simulation)을 거쳐 다층인식모형(MLP)을 구성하였다. 최종적인 결과값은 네 가지 형태로 출력하였다. 다 축 변환은 위험 요인과 보호요인을 각각 X, Y축으로 하고 결과변인인 중퇴 유무를 등고선형 그래프의 Z축으로 설정하였으며, 예측변인의 변화단위는 RBF

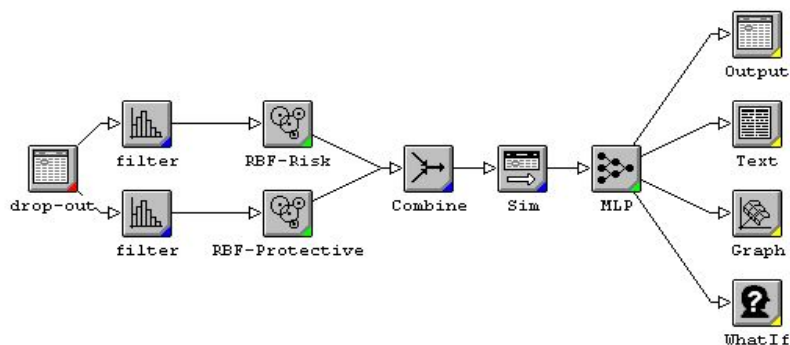


그림 1. 중퇴 예측의 신경망 위상

로 처리된 최하 값에서 최상 값을 스물 한 개의 구간으로 할당하였다. 다층인식 모형에서는 가중치 부여 방식으로 Gaussian분포에 맞추어 예측모형을 최적화하는 방식을 선택하였다. 전 과정에서 결측 값에 대한 대체규칙(surrogate rule)이 설정되었고 단순한 이진(binary) 분류에 따른 오류를 감소시키고자 중퇴 유무를 포함한 문자(symbol)형 변인들을 실수(float)형 변인으로 변경시켰다. 아래의 <그림 1>은 중퇴를 예측하는 신경망 분석의 흐름도이다.

<그림 1>의 신경망 흐름도에 따라 위험 요인과 보호 요인으로 결과변인인 중퇴를 예측하는 RBF를 각각 산출하였으며, 그 결과는 다음의 <표 1>과 같다.

표 1. 위험 및 보호요인의 다층인식모형

		중퇴여부
위험 요인	Best Fit(%)	70.10
	RMS Error	4.91
	Mean Absolute	4.08
	Mean Absolute(%)	17.10
보호 요인	Best Fit(%)	58.34
	RMS Error	5.46
	Mean Absolute	4.94
	Mean Absolute(%)	20.88

<표 1>은 모형 개발에 사용된 사례의 70%로 구성된 훈련모형의 결과로서 Best Fit는 각 요인이 종속변인을 설명하는 예측 값이고, Mean Absolute는 예측 값과 실제 값의 차이를 절대값으로 표현한 것이며, Mean Absolute(%)는 그것의 백분율이다. RMS Error는 이 두 값을 토대로 얻어진 모형의 정확도를 나타내는데 오차 값이 0이라는 것은 모형이 100%의 예측 정확도를 갖는 것

을 의미한다. 결과에서 보여지듯이 보호 요인(58.34%)보다 위험 요인(70.10%)이 설명력이 훨씬 높았다. 20%의 사례로 구성된 검증모형(20%)의 결과, 위험 요인과 보호 요인의 Best Fit도 각각 67.13%와 56.71%로 나타나 훈련모형이 과잉적합(over fit)하게 구성되었을 가능성은 그리 크지 않았다.

다음 단계로 RBF로 예측된 위험 요인과 보호 요인의 함수를 결합하여 위험 요인과 보호요인의 영향을 각각 X축과 Y축으로 하고 중퇴 유무를 Z축으로 하는 다 축 변환을 실시하였다. 이 과정을 통해 구성된 다층인식모형(MLP)의 결과는 아래의 <표 2>에 나타나 있다.

표 2. 위험 및 보호 요인을 결합한 다층인식 모형

		중퇴여부
위험 요인	Best Fit(%)	74.18
	RMS Error	1.88
	Mean Absolute	1.45
	Mean Absolute(%)	7.40

다층인식모형은 RMS 오차가 0.001미만이거나 예측정확도가 95%이상이면 자동으로 정지된다. <표 2>에서 알 수 있듯이, 위험 요인과 보호요인을 결합하여 다 축 변환시킨 중퇴의 다층인식모형은 74.18%의 높은 예측력을 보여주고 있다. 검증자료와 타당화 자료에서는 각각 75.33%와 72.31%로 나타났다. RBF에서 산출한 값에 비해 다층인식모형에서의 예측력이 더 높았으며 과잉적합의 위험성도 보다 줄어든 것으로 나타났다.

위험 요인과 보호 요인을 다 축 변환(simulation)시킨 3D 등고선형 그래프는 <그림 2>

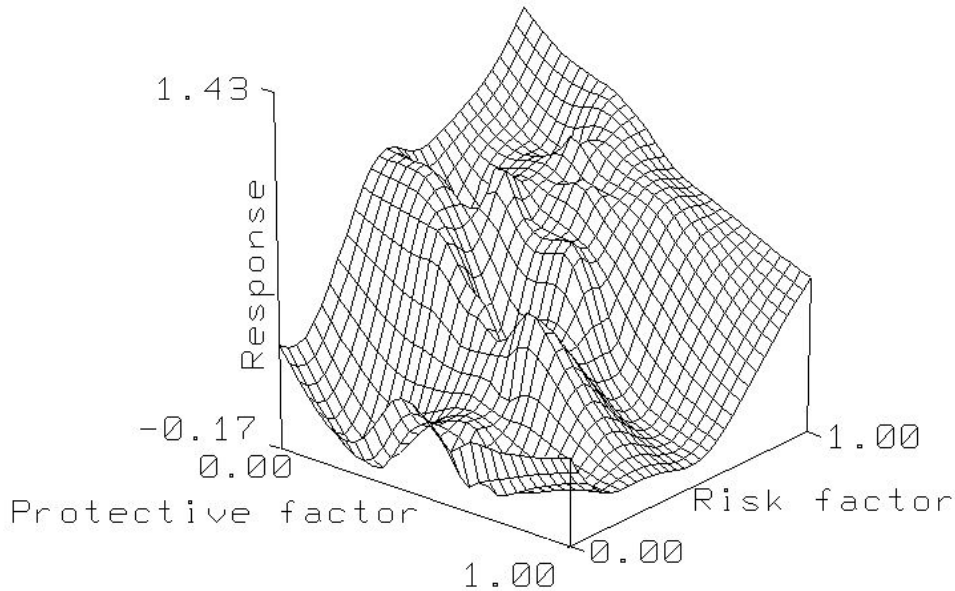


그림 2. 위험 및 보호 요인을 다 축 변환(simulation)시킨 3D 등고선형 그래프

와 같다.

<그림 2>에서 보여지는 바와 같이 위험 요인이 증가할수록(0에서 1로) 중퇴 반응(response)도 대체로 증가하고 있었으나 보호요인이 증가할수록 중퇴 반응이 감소하는 것은 일관된 경향이 아니었다. 이것은 <표 1>에 나타난 결과와 일치하는 것이다. 그러나 3D그래프는 X축과 Y축의 접점을 기준으로 Z 축의 반응을 확인해야 한다. 이 그래프 상에서 보면 위험요인이 가장 높고(1), 보호요인이 가장 낮을 때(0) 예상했던 대로 중퇴반응이 가장 높게 나타났다. 전체적으로 볼 때 중퇴 비율을 떨어뜨리는 것은 보호요인의 증가보다는 위험 요인의 감소에 의해 주도되는 양상이었다.

신경망의 whatIF 출력에서는 3차원 공간의 그래프를 2차원의 형태로 투영시키고 3차원 등고선의 굴곡을 색의 진함인 밀도로 표현하거나 등고선으로 표현된 두 축의 평균 좌표로 비 선형성

의 정도를 확인할 수 있다. 모형은 연구자가 원하는 값(X, Y축에)을 대입하면 3D 그래프에 근거하여 변화되고 그 값에 따라 중퇴의 확률도 추적하여 계산되기 때문에 다분히 동태적이다. 그러나 <그림 3>에서 보여지는 바와 같이 밀도나 평균 좌표로 볼 때 위험 요인과 보호요인으로 구성된 본 중퇴 모형은 비 선형성의 경향이 비교적 뚜렷하였다.

마지막으로 이 신경망 모형을 모형 예측에 사용된 265명에 적용시켜 보았다. 표집을 중학생과 고등학생, 남학생과 여학생의 집단으로 구분한 후 개발된 모형을 적합화 시키는 절차를 거쳤다. <표 3>에는 개발된 신경망 모형이 집단에 따라 얼마나 예측력이 있는지 제시되어있다.

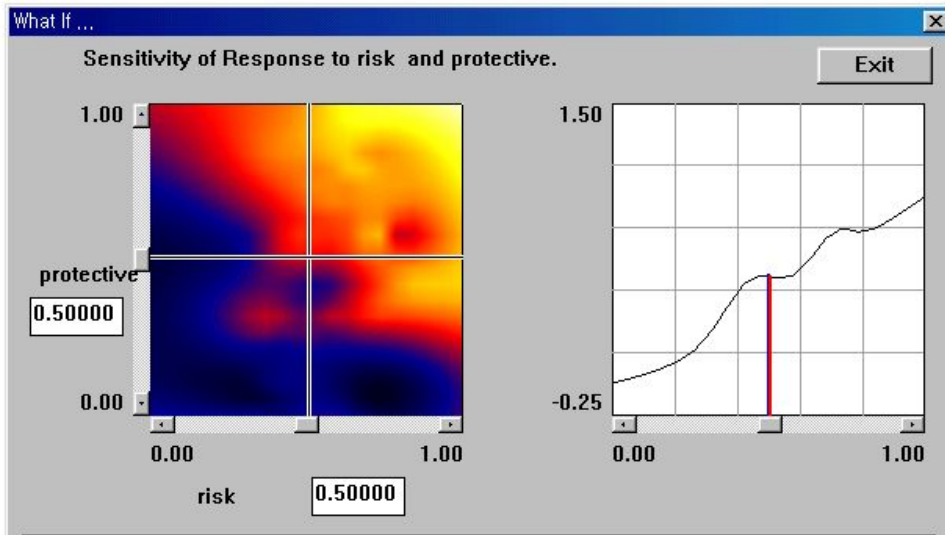


그림 3. 민감도 분석과 평균 좌표축으로 조망한 위험 및 보호 요인의 중퇴 예측 모형

<표 3>에서 보듯이 개발된 신경망 모형은 중학생(67.13%)보다는 고등학생(77.17%) 집단에 그리고 여학생(69.36%)보다는 남학생(74.32%) 집단에 더 설명력이 있는 것으로 나타났다. 전체적으로 볼 때 위험 요인 군과 보호 요인 군을 병합시켜 중퇴를 예측하는 신경망 모형은 평균 70% 이상의 설명력을 가지는 우수한 모형이었다.

논 의

본 연구는 구분용 등(2002)에 의한 중퇴 관련 변인의 데이터 마이닝을 기초로 하였다. 본고에서는 기존의 이론적 배경과 의사결정 나무분석에

의해서 요약, 축소된 중퇴 관련 변인을 위험 요인과 보호 요인으로 양분한 후 인공 신경망 분석을 통해 중퇴를 가장 효율적으로 설명할 수 있는 모형을 구성하고자 하였다.

<표 1>의 단층 인식 모형(RBF)의 결과 보호 요인(58.34%)보다 위험 요인(70.10%)이 설명력이 훨씬 높았으며, 위험 요인과 보호 요인을 다 축 변환(simulation)시킨 3D 등고선형 그래프<그림 2>에서도 보호요인보다는 위험 요인에 의한 중퇴여부가 보다 일관되게 예측되는 것으로 나타났다. 이러한 결과는 중퇴를 예방하고자 할 때 보호요인을 개발시키는 것보다는 위험 요인을 감소시키는 것이 보다 효과적인 전략임을 시사하는 것이다.

표 3. 집단에 따른 신경망 모형의 예측도

	중학생	고등학생	남학생	여학생
Best Fit(%)	67.13	77.17	74.32	69.36
RMS Error	3.19	2.64	2.88	3.07
Mean Absolute	4.17	2.91	3.02	3.38
Mean Absolute(%)	18.26	13.96	15.25	17.44

RBF 모형과 MLP 모형을 비교한 결과, 개별 요인의 상대적 효과만을 산출한 RBF 모형보다 두 요인간의 상호작용을 함수에 포함시킨 MLP 모형의 예측력이 더 높았다. 위협 요인과 보호요인으로 구성된 다층 인식 모형(MLP)은 중퇴 여부에 대해 74% 이상의 예측력을 보여주었다. 일반적으로 '의미 있는'(significant) 정도라고 인정되는 중다 회귀분석의 설명력(R^2)을 감안한다면 신경망 모형의 설명력은 극적이고 인상적인 수준이라고 말할 수 있다.

발달정신 병리학에서는 위협 요인과 보호요인의 관계에 대해 크게 세 가지의 입장이 제안되어 있다. 위협 요인과 보호 요인을 독립적인 차원으로 다루거나, 양극단의 일차원적인 속성으로 간주해 위협 요인이 없는 것 자체가 보호요인으로 보든지 그 반대의 경우를 가정한다. 마지막으로 위협 요인이 상존해 있더라도 이를 막는 심리적, 환경적 특성을 소유했을 경우에만 보호요인의 존재를 가정한다(Cicchetti & Rogosch, 1996). 세 번째 견해에 의하면 위협 요인이 높은 수준인 상태에서 추가된 특정한 변인이 보호요인이기 위해서는 문제 행동이나 증상의 비율이 감소되어야 한다. 본 신경망 모형에서는 위협 요인의 감소와 보호요인의 증가가 이루어지는 동안 일관되게 중퇴율을 감소시키지 않음이 확인되었다. <그림 2>에서 나타나듯이 위협요인이 중간 수준일 때보다 매우 낮은 수준일 때 오히려 중퇴율이 더 높았으며, 보호요인의 경우는 아주 낮은 수준일 때뿐만 아니라 높은 수준일 때도 중퇴율이 다소 증가하는 것으로 나타났다. 이러한 결과는 중퇴에 관한 위협 요인과 보호요인이 양극단의 상반된 개념이 아님을 시사하는 것이다. 또한 위협 요인이

높은 수준에서 보호요인이 높은 경우 오히려 중퇴율이 약간이나마 증가하는 것으로 나타나 세 번째 견해보다는 중퇴에 관한 한 위협 요인과 보호 요인이 독립된 차원이라는 Battin-Pearson 등의 주장(2000)이 보다 설득력이 있는 것으로 확인되었다.

특히, 본 연구의 중퇴 모형에서는 위협 요인과 보호요인이 중간 수준일 때 오히려 위협 요인은 중퇴율을 낮추고 보호 요인은 중퇴율을 증가시키는 것으로 나타나 위협 요인과 보호요인의 상호작용이 특정한 수준에서 중퇴에 다른 영향을 끼칠 가능성이 시사되었다. 구본용 등의 선행 연구(2002)에서 가난과 같은 위협요인이 고등학생 시기의 중퇴에는 오히려 보호 요인으로 작용한다는 사실이 확인된 바 있다. 그러나 본 연구에서는 중간 수준의 보호요인이 중퇴율을 증가시키는 경향도 있었으므로 신경망 모형의 이러한 결과가 선행 연구에서와 같이 역경(adversity)에 대한 반작용이라고 결론짓기는 어렵다.

<표 3>에서 보듯이 본 연구에서 구성한 신경망 모형은 중학생보다는 고등학생에 그리고 여학생보다는 남학생에 상대적으로 더 적합하였다. 구본용 등(2002)의 연구에서도 중학생의 경우 다양한 심리, 사회적 변인이 작용하기보다는 학교 및 학업 생활에 관련된 단순한 변인들로 중퇴가 결정되었고, 여학생의 중퇴도 부모-자녀관계와 전환기의 심리적 부적응에 한정되는 경향이 있었기 때문에 본 연구에서와 같이 위협 요인과 보호요인을 망라한 포괄적인 신경망 모형에는 적합성이 다소 줄어든 것이라고 해석할 수 있다. 그러나 이것은 중학생이나 여학생 중퇴의 예방적 개입 전략은 오히려 특정 영역(학교관련 영역과 가정 관

런 영역)에 집중하는 일관된 방식이 더 효과적일 수 있음을 시사하는 결과이기도 하다.

발달 정신 병리학에서는 특정한 연령 대에서의 위험 요인이나 보호요인이 대상자의 발달 시점에 따라 그 역할을 달리하는 소위 '시기적 변환(timinal turn)' 현상을 중요한 연구 과제로 삼고 있다(Loeber, 1990). 예를 들어, 10대의 결혼은 여성의 우울증에 위험 요인이지만 20대 후반 여성의 결혼은 우울증의 보호 요인이며, 직업이 있을 경우 30대 초반까지의 미혼은 우울증의 보호 요인이지만 직업 여부와 관계없이 30대 후반의 미혼은 우울증에 위험요인이다(Capaldi, 1992). 위험 요인과 보호요인으로 문제 행동이나 장애를 설명하는 신경망 모형은 이러한 변화과정에 대한 중요한 시사점을 제공할 수 있을 것으로 보인다.

즉, 본 연구에서의 신경망 모형은 다음과 같은 활용이 가능하다. 특정한 대상자가 중퇴를 할 것인지 예측하기 위해 그를 대상으로 중퇴에 관련된 정보들(본 모형에서는 20 개변인)을 수집한 후 이미 개발된 모형에 대입해 중퇴가능성을 산출할 수 있다. 또한 관련 변인들을 중심으로 그에게 적용 가능한 영역에 개입하여 그 수준을 어떻게 변화시켰을 때 중퇴율이 얼마나 변하는지 실시간으로 확인할 수 있기 때문에(신경망 위상에서의 What If는 특정변인의 수준을 다 축 상에서 인위적으로 대입하여 종속변인의 설명량을 시시각각으로 산출하는 도구이다) 특정 변인의 변화 수준으로 개입의 효과를 판정하고 그 우선 순위를 정할 수 있다. 특정 대상자가 20개의 변인 중 몇 가지 정보밖에 없는 경우라도 가용한 정보를 피드백하여 스스로 적합시켜 나가는 소위 '기계 학습 과정'을 통해 20개 변인을 모두 사용한 것 못지

않게 예측력을 유지할 수도 있다. 이 신경망 모형을 통해 중퇴 가능성이 있는 위험 군을 대상으로 모의 실험(simulation)을 행할 수 있는데 가상적인 위험군을 관리하는 이러한 방식은 현재 CRM분야(Ball, 2000)나 건강정보학 분야(Ammenwerth, 2001)에서 탁월한 효과가 증명되고 있다. 이러한 방식을 임상이나 건강 관리 분야에 적용할 경우 각 개인적 특성에 따른 맞춤형 접근(tailored approach)도 충분히 가능할 것으로 보인다.

중퇴는 그 유형(push-outs, drop-ins, stay-outs 등)에 따라 동기나 결과가 상이하게 나타날 수 있다(Rumberger & Thomas, 2000). 또한 학업 중단은 그 발생이나 가능성을 예측하는 변인과 지속기간을 예측하는 변인이 다를 가능성이 있다(Trusty & Dooley-Dickey 1993). 위험 요인과 보호 요인 중에서 안정적인 특질인 취약성(vulnerability)과 탄력성(resilience)요인을 변별하는 것도 의미가 있을 것이다. 추후 연구에서는 이와 더불어 중퇴의 장기적 후유증을 추적하는 연구가 필요할 것으로 보인다.

참고문헌

- 구본용, 신현숙, 유제민(2002). 데이터마이닝을 이용한 중퇴 모형에 관한 연구. 청소년 상담 연구, 10권(2), 35-57.
- 교육인적자원부 (1996). 교육복지종합대책(I). 서울: 교육부.
- 박창남(2001). 중도탈락 청소년 예방 및 사회적응을 위한 종합대책 연구, 교육인적 자원부 & 한국 청소년 개발원.
- 이혜영 (2001). 학업중도탈락의 실태와 학교에서의

- 대처방안. 한국청소년복지학회 추계학술 발표대회: 학업중도탈락과 청소년복지(pp. 7-23). 서울: 한국청소년복지학회.
- 한국청소년상담원 (2002). 학업중단 청소년사회적응을 위한 지역단위 관련기관간 연계협력체제구축방안. 서울: 한국청소년상담원.
- Ammenwerth, E., Knaup, P., & Meier(2001). Digital libraries and recent medical informatics. *Informatics*, 40, 163-168.
- Ball, M.J.(2000). *International Efforts in Informatics: Creating a Global Villiage for Healthcare*. M.D. Computing May/June 2000, 50-53.
- Battin-Pearson, S., Newcomb, M. D., Abbott, R. D., Hill, K. G., Catalano, R. F., & Hawkins, J. D. (2000). Predictors of early high school dropout: A test of five theories. *Journal of Educational Psychology*, 92(3), 568-582.
- Capaldi, D. M.(1992). Co-occurrence of conduct problems and depressive symptoms in early adolescent boys. PartII: A 2-year follow-up at Grade 8. *Development and Psychopathology*, 4, 125-144.
- Cicchetti, D., & Rogosch, F. A.(1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology*, 8, 597-600.
- Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research*, 59(2), 117-142.
- Institute on Community Integration[ICI] (1994). *Policy research brief: Are we pushing students in special education to drop out of school?* Minneapolis, MN: University of Minnesota.
- Janosz, M., Le Blanc, M., Boulerice, B., & Tremblay, R. E. (1997). Disentangling the weight of school dropout predictors: A test on two longitudinal samples. *Journal of Youth and Adolescence*, 26, 733-762.
- Krohn, M. D., Thornberry, T. P., Collins-Hall, L., & Lizotte, A. J. (1995). School dropout, delinquent behavior, and drug use: An examination of the causes and consequences of dropping out of school. In H. B. Kaplan (Ed.), *Drugs, crime, and other deviant adaptations: Longitudinal studies* (pp. 163-183). New York: Plenum Press.
- Levine, H. M. (1986). *Educational reform for disadvantaged students: An emerging crisis*. West Heaven, CT: National Educational Association.
- Lin, F., Chou, S., Pan, S., & Chen, Y.(2001).Mining time dependency patterns in clinical pathways. *International Journal of Medical Informatics* 62, 11-25.
- Loeber, R.(1990). Antisocial behavior: More enduring than changeable? *Journal of the American Academy of Child and*

- Adolescent Psychiatry*, 30, 393-397.
- Marble, R.P. & Healy, J.C.(1999). A neural network approach to the diagnosis of morbidity outcomes in trauma care. *Artificial Intelligence in Medicine*, 15, 299-307.
- Masten, A.S., Best, K.M., & Garmezy, N.(1990). Resilience and development: Contributions from the study of children who overcome adversity. *Development and Psychopathology*, 2, 425-424.
- McWhirter, J. J., McWhirter, B. T., McWhirter, A. M., & McWhirter, E. H. (1998). At-risk youth: A comprehensive response. Pacific Grove, CA: Brooks/Cole.
- National Center for Education Statistics[NCES] (1994). *Dropout rates in the United States: 1993*. Washington, DC: U. S. Government Printing Office.
- Natriello, G., Pallas, A. M., & McDill, E. L. (1986). Taking stock: Renewing our research agenda on the causes and consequences of dropping out. *Teachers College Record*, 87(3), 430-440.
- Park, J & Edington, D.E.(2001). A sequential neural network model for diabetes prediction. *Artificial Intelligence in Medicine* 23, 277-293.
- Pynoos, R.S., Steinberg, A.M., & Wraith, R.(1995). A developmental model of childhood traumatic stress. In D. Cicchetti & D.J. Cohen(Eds.), *Developmental Psychopathology*(vol. 2, pp. 72-96). New York: John Wiley & Sons, Inc.
- Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal*, 32(2), 583-625.
- Rumberger, R. W., & Thomas, S. L. (2000). The distribution of dropout and turnover rates among urban and suburban high schools. *Sociology of Education*, 73, 39-67.
- Sroufe, L.A., Cooper, R.G., & DeHart, G.B.(1992). *Child Development: Its nature and course*(2nd ed.). New York: McGraw-Hill.
- Trusty, J., & Dooley-Dickey, K. (1993). Alienation from school: An exploratory analysis of elementary and middle school students' perceptions. *Journal of Research and Development in Education*, 26(4), 232-242.

원고 접수: 2003년 2월 26일

수정원고 접수: 2003년 4월 2일

게재 결정: 2003년 4월 7일

The neural network model of risk and protective factors on school drop out

Bon-Yong Khu Jae-Min Yoo

Kangnam University

The present study was conducted to develop neural network model on school drop out, with two primary steps in mind. They were : First step, to differentiate between risk and protective factors by applying the decision tree analysis on the data about individual, family, peer, school and community-related variable. Second step, to integrate risk and protective factors that would simulated on the multi-axis. According to the neural network model, the more risk factors is cumulated, the more rate of school drop out is increased. But in the case of protective factors, its tendency is not clear. The neural network model showed that risk and protective factor were simulated on the multi-axis would predict school drop out mainly by non-linearity function. Because the various sub-dimensions was involved, the neural network model was fit male and high school student samples very well. Tailored approach of school drop out is available by applying neural network model. The future study on risk and protective factors model of school drop out is desirable to simultaneous consider patterns and duration of school drop out.

Keywords: school dropout, risk factors & protective factors, artificial neural networks