

청소년 인터넷 중독 검사 방법의 통계적 고찰 및 개선 - Young의 척도를 중심으로

박찬정[†] · 김형철^{††} · 고영민^{††} · 현정석^{†††} · 김철민^{††††}

요 약

국내외에서 이루어진 기존의 인터넷 중독 검사에 대한 연구는 다양하며 방대하지만, 대부분의 연구가 중독도 검사를 위한 항목추출이거나 항목들을 이용한 중독도 분석이었지 항목간의 분석을 통한 차별화는 이루어지지 않았다. 본 논문에서는 리즈렐과 데이터마이닝 기법을 이용하여 Young의 인터넷 중독 검사방법을 고찰하고 문제점을 지적한 후, 검사 방법에 대한 대안을 제안한다. 이를 위해 Young의 척도를 이용하여 청소년 440명을 대상으로 설문을 실시하고, 요인분석을 통해 Young 척도의 문제점을 분석한다. 또한, 데이터 마이닝 알고리즘인 J48 및 PART를 이용하여 설문문항들 중에서 인터넷 중독도를 구분 지을 수 있는 항목들을 선별하고 이를 기반으로 인터넷 중독 학생들을 진단하고 처방하는데 활용하는 방법을 제안한다.

주제어 : 인터넷 중독, 데이터마이닝, 리즈렐, 분류분석, 사이버 윤리

Statistical Examination and Improvement of the Internet Addiction Testing Method for Adolescent - Focused on Young's Measurement

Chan-Jung Park[†] · Hyung-Chul Kim^{††} · Young-Min Ko^{††} ·
Jung-Suk Hyun^{†††} · Cheol-Min Kim^{††††}

ABSTRACT

So far, the researches about the Internet addiction testing are diverse and enormous. However, since the testing factors have the same importance, there is no discrimination among the factors. In our paper, we analyze the problems of the existing researches by using Lisrel and data mining algorithms. Next, we propose an alternative for the Internet addiction testing. In order to accomplish these, we make an questionnaire consisted of Young's test and then, we asked to 440 primary and secondary school students. In addition, we analyzed the problems of Young's test with factor analysis and then, we distinguish more important items from others with J48 and PART. Finally, based on this analysis, we propose a method to diagnose addicted students and prescribe them.

Keywords : Internet Addiction, Data Mining, Lisrel, Classification Analysis, Cyber Ethics

† 종신회원: 제주대학교 컴퓨터교육과 부교수(교신저자)

†† 준회원: 제주대학교 교육대학원 컴퓨터교육전공 석사과정

††† 정회원: 제주대학교 경영정보학과 부교수

†††† 정회원: 제주대학교 컴퓨터교육과 부교수

논문접수: 2010년 6월 11일, 심사완료: 2010년 7월 5일

1. 서 론

인터넷이 보편화되면서 청소년의 인터넷 중독에 대한 문제 심각성이 증가하고 있으며 청소년 인터넷 중독과 관련된 국내 논문만도 400편에 이른다[1]. 그 중, 인터넷 중독 장애(Internet Addiction Disorder)를 처음 제안했던 골드버그(Goldberg)나 영(Young)은 인터넷 중독에 대해 내성이나 금단 증상에 기초하여 척도를 제안하였다[2]. 대부분의 국내연구는 Young의 척도를 사용해 왔다. Young의 척도를 다각적으로 연구한 연구들은 신뢰도를 나타내는 크론바흐 계수가 높아 신뢰할 수 있다고 간주해 왔다. 한편, 국내에서도 국내 상황에 맞게 K-척도[3]와 같은 인터넷 중독도를 검사하는 검사문항이 개발되었다.

인터넷 중독의 요인에 대한 연구도 많다. 손경애[4] 등의 연구에 의하면 사회인구학적 특성과는 상관없이 인터넷 중독에 영향을 미치는 요인으로 인터넷 활용정도(인터넷 활용기술, 충동성, 익명성), 인터넷 태도, 인터넷 효능감이 있다. 즉, 인터넷 활용기술이 탁월하고 오랜 시간동안 빈번하게 인터넷을 활용하는 사람일수록 인터넷에 중독될 경향이 높은 것으로 나타났다. 하지만 이들 요인 간에 어떤 요인이 상대적으로 더 큰 영향을 미치는지 알 수 없었다.

강만철의 연구[5]에서도 인터넷 중독을 측정하기 위한 문항으로 집착, 강박적 사용, 재발, 내성 및 의존성을 제시하였지만, 이 연구에서도 역시 요소간 우월성은 알 수 없다.

박경호의 연구[6]에서도 Young의 문항을 포함하여 중독척도, 신체 증상, 준거, 인적 배경 문항 등을 추가하였다. 이 연구는 Young의 문항을 확장하여 척도로 제안하였으며, 분석한 결과 Young과 비슷한 신뢰도를 갖는 것으로 나타났다.

이와 같이 인터넷 중독 측정을 위한 척도 개발에 대한 연구가 많이 진행되었고, 비교 대상으로는 항상 Young의 연구가 있어왔다. 하지만, 최근 Young의 척도는 많은 결함이 있음이 지적되고 있다. 모든 척도들은 문항들을 몇 개의 요인 그룹으로 구분하여 다양한 통계적 분석을 통해 의미를 찾고 있을 뿐, 각 요인이 중독도에 얼마나 영향을 미치는지 상대적 중요도에 대한 분석은 이

루어지지 않았다.

보다 정확한 인터넷 중독을 진단하기 위해 최근 데이터 마이닝 기법을 도입한 연구[7]가 시도되었다. 김은주의 연구[7]에서는 인터넷 중독도를 검사하기 위해 설문조사 방식이 아닌 사용자의 인터넷사용 데이터를 분석하는 방식을 사용함으로써 보다 객관적인 자료를 추출하려고 힘썼다. 이를 위해 데이터 마이닝 기법을 사용하였으며 주중사용시간, 선호게임장르, 주 사용목적 등을 주요한 요인으로 제시하였다. 또한, 사용자 특징을 파악하기 위해 K-척도[8]를 중심으로 간단한 설문을 실시하였다[7]. 하지만, 이 연구에서도 척도의 독립성 가정에 대한 통계적 검증은 이루어지지 않았다.

본 논문에서는 우선 인터넷 중독을 검사하는 Young의 척도에 대한 요인분석의 의미를 기존연구[5]를 바탕으로 분석하여 Young 척도의 문제점을 지적한다. 또한, Young의 척도를 여러 통계적 기법을 사용하여 통계적으로 검증 및 분석한다. 본 논문은 데이터 마이닝 기법 중 분류분석을 중심으로 인터넷 중독을 검사하는 척도들 중에서 청소년들에게 보다 큰 영향을 미치는 요인을 찾아 우선순위를 부여함으로써 인터넷 중독의 예방과 치료에 도움을 줄 수 있다. 본 논문에서는 Young의 척도를 중심으로 그 문제점을 기술하지만, 실제로 Young 척도와 유사하게 리커트 척도 방식으로 설문을 실시한 후, 응답자의 답을 모두 합산하는 방식에 대한 문제점을 지적하는 연구이다.

본 논문의 구성은 다음과 같다. 2장에서는 인터넷 중독 검사지와 데이터 마이닝에 대한 이론적 배경을 기술하고, 3장에서는 선정된 연구대상과 적용된 리즈렐 및 데이터마이닝과 같은 통계적 분석 기법을 개괄한다. 그리고 데이터 분석 결과를 4장에서 논한 후, 5장에서 결론을 맺는다.

2. Young의 인터넷 중독 검사지 문제점

2.1 Young의 인터넷 중독도 검사지

Young의 검사지는 연구 [9]에 기술되어 있다. Young의 검사지는 지나친 몰두, 의존성, 조절능력 상실 등을 표현하며 20개의 문항으로 이루어

진다. 각 문항은 동등하게 1~5점까지 값을 가지며, 최대 점수는 100점이 된다. 점수 합계가 20~49점은 평균적인 온라인 이용자로, 웹을 오래 사용하지만 자신의 통제 하에 인터넷을 사용할 수 있는 상태라고 판정했다. 한 단계 더 높은 점수인 50~79점은 인터넷 때문에 문제가 있었을 것으로 간주하고, 자신의 인터넷 중독 가능성에 대해 준비할 필요가 있다고 판정했다. 가장 높은 점수인 80~100점은 인터넷 때문에 생활에 중대한 문제가 발생하고 있다고 판정, 당장의 조치로서 전문가 상담이나 예방, 혹은 치료 등을 반드시 받도록 권고하였다[10][11]. 하지만, Young의 척도는 통계적으로 문제를 지니고 있다. 이는 강만철의 연구[5]에서 일부 지적이 되었다. Young의 문제점은 다음 절에서 다룬다.

2.2 요인분석을 통한 Young 척도의 문제점

인터넷 중독 척도 개발을 위한 기초 연구[5]에서 Young 척도의 문제점으로 우리 상황에 맞지 않는다는 점과 통계적 분석절차가 수행되지 못했다는 점을 지적하고 있다.

Young의 척도는 요인분석을 통해 문제의 심각성을 알 수 있다. [5]의 연구에서 여러 연구들이 Young 척도에 대한 요인분석을 실시하여 하위 구조를 밝히는 내용을 담고 있다(<표 1> 참조)

<표 1> Young 척도에 대한 요인분석 결과[5]

연구자	조사대상	추출요인
강지선(1999)	10~40대 277명	<ul style="list-style-type: none"> ◦ 생활상 부적응 (10문항) ◦ 조절능력 상실(7문항) ◦ 강박적 생각이나 기대감(3문항)
김교현(2000)	중학생 385명	<ul style="list-style-type: none"> ◦ 통제력 부족(7문항) ◦ 심리적 의존(5문항) ◦ 사회적 고립(2문항) ◦ 강박적 집착(3문항) ◦ 사용 은폐(3문항)
안정임과 김동규(2000)	중등학생 696명	<ul style="list-style-type: none"> ◦ 가상 현실 지향 (10문항) ◦ 절제 실패(6문항) ◦ 피해 경험(4문항)

[5]의 연구는 <표 1>의 연구자마다 요인분석 결과가 달라지는 문제점을 지적하고 있다. 이 사실을 통해 Young의 척도가 하나의 단일 척도가 아니며, 하위척도 간에 상관관계가 있을 수 있는 가

능성을 짐작할 수 있다. Young의 척도가 갖는 문제점을 나열하면 다음과 같다.

첫째, Young의 척도는 20문항 각각에 대해 1~5점 사이로 부여하여 합산하는 형태를 띠고 있다. 단순합산 방식은 모든 문항이 통계적으로 서로 독립적임을 내포한다. 하지만, <표 1>에 나타난 바와 같이 문항들은 몇 개의 요인으로 그룹된다. 이런 결과는 문항 간 상관관계가 있기 때문에 나타났다고 볼 수 있다.

둘째, <표 1>과 같이 요인분석 후, 추출된 요인의 수와 각 요인에 적재된 측정항목의 수가 일정하지 않은 문제점이 있다. 각 요인에 적재된 측정항목의 수가 다르다는 것은 Young의 척도가 연구자마다 상이한 가중치를 주었다는 것을 알 수 있다. 예를 들어, 강지선의 연구에서 생활상 부적응의 요인에는 10개의 측정문항이 적재되었고 강박적 생각이나 기대감이라는 요인에는 3개의 측정문항이 적재된 것으로 보고하였다. 이런 접근법은 중독도를 계산할 때 생활상 부적응의 요인을 강박적 생각이나 기대감의 요인보다 훨씬 큰 요인으로 간주하게 되는 결과를 초래하게 된다. 즉, 측정항목의 수를 중복하여 계산하는 이중 계산편기(double counting bias)가 개입될 가능성이 생긴다. 또한 연구자마다 요인분석의 결과가 다르다는 것은 연구자마다 상이한 가중치를 부여하는 것을 의미하므로 연구의 일반화를 저해하는 문제가 있다.

셋째, <표 1>에 의하면, 요인분석 결과가 표본에 따라 모두 다르며, 중요하게 생각하는 요인도 다르다. 즉, 특정 항목이 인터넷 중독도에 높은 판별력을 가질 수 있음에도 불구하고, 요인분석 결과로부터 Young이 문항 가중치에 대한 어떠한 가정을 두지 않았음을 알 수 있다.

상기와 같은 논의에도 불구하고 Young의 척도에 대해 요인분석으로 안정적으로 통계적 검증을 하기 어렵다. 왜냐하면, 요인분석은 군집분석과 유사하게 분석하는 방식을 직교 또는 사각으로 선택하느냐에 따라 분석 결과가 달라지기 때문이다. 또한 요인추출모형을 선정하는 방법에 따라서도 결과가 달라지며, 요인추출을 아이겐 값 기준과 설명된 분산기준에 의해서도 결과가 달라진다.

결국 20개의 각 문항들에 대해서 구조방정식 모

형인 리즈렐을¹⁾ 통한 확인적 요인 분석을 통해 통계적 검증을 거치고 분류분석기법을 이용하여 20개의 각 문항중에서 판별력이 높은 변수를 추출하여 인터넷 중독도를 계산하는 방법이 필요하다.

3. 연구대상 및 분석기법

3.1 연구대상 및 도구

본 논문에서의 조사대상자 특성은 다음 <표 2>와 같다. 초등 5, 6학년과 중·고등으로 구분하여 모두 440명에 대해서 설문조사를 실시하였다.

<표 2> 조사대상자 분포

학년 \ 성별	남성	여성	전체
초등학생	51명	55명	106명
중학생	59명	67명	126명
고등학생	90명	118명	208명
전체	200명	240명	440명

설문 문항은 인터넷 중독도 조사를 위한 20개의 문항으로 구성된 Young의 척도가 주어졌으며, 선호하는 교과목, 하루 인터넷 사용시간, 하루 게임시간, 경험한 게임의 개수, 게임을 시작한 시기에 대한 문항을 추가하였고 기본 정보로는 성별, 학년이 주어졌다.

3.2 분석기법

본 논문에서는 리즈렐을 이용하여 Young 척도의 문제점을 분석한다. 우선, Young 척도는 20 문항에 대한 검사치를 모두 더하여 총합을 기준으로 중독도를 결정하는 방법을 취하고 있기 때문에 20개의 검사 항목들은 통계적으로 모두 독립적이라는 가정을 가지고 있다. 이와 같은 가정하에 확인적 요인분석 결과는 자유도 = 190, 카이제곱 = 3427.6, GFI(제안 모델의 적합도 설명력) = 0.313 이다. 즉, 기존 Young의 방식대로 20개 항목의 점수를 단순히 합산하여 중독도를 계산한다면, 중독도에 대한 설명력이 31.3%정도밖에는 미치지 못함을 의미한다.

반면, 확인적 요인분석에서 상관관계가 매우 높은 몇 개의 문항을 (예를 들면, c1과 c2, c1과 c16 등) 파악할 수 있었다. 이들 측정 항목 간에 상관관계를 허용하여 실시한 확인적 요인분석의 결과, 자유도는 172로 줄어든 반면에 카이제곱 값 = 1935.7, GFI = 0.534로 향상되는 것으로 나타났다. 측정항목 간 독립성을 가정한 모형과 상관관계를 반영한 모형 간 카이제곱 차이검증결과, 카이제곱 = 1497.9 ($p < .01$)로 유의한 차이를 보였다. 앤더슨과 거빙의 연구[13]에 따라서 상관관계를 반영한 모형이 독립성을 가정한 모형에 비해 데이터에 보다 부합되는 것으로 나타났다.

이러한 분석결과를 통해 Young의 20개 항목에 대해 독립성 가정에 기초하여 동등하게 간주하고 각 항목에 대한 대답을 단순히 합산하여 중독도를 처리하는 방법에 문제가 있음을 알 수 있다.

실용적인 측면에서 Young의 측정항목 중에서 어떤 변수가 인터넷 중독도에 대한 판별력이 높은지 조사할 필요가 있다. 이를 위해 본 논문은 Weka 3.7.0[14]을 이용해 데이터 마이닝 기법인 분류분석을 실시하였다. 의사결정 트리는 Weka의 J48 알고리즘을 이용하여 인터넷 중독도에 대한 검사항목을 분석하는 데 적용하였다[14]. J48은 C4.5 알고리즘을 개선한 알고리즘으로 수치형(numeric) 데이터 타입뿐만 아니라 명목형(nominal) 데이터 타입도 처리하며 의사결정 트리로 결과 모델이 생성되는 알고리즘이다[14].

규칙기반 분류분석은 Weka의 PART 알고리즘을 이용하여 인터넷 중독도와 관련된 규칙 리스트를 생성하였다. 분류분석은 다수의 속성을 갖는 객체를 사전에 정해진 그룹 또는 범주중의 하나로 분류하는 것으로 J48와 같은 의사결정 트리와는 다르게 특정 범주를 결정할 수 있는 주요 규칙을 찾아낸다[14].

위의 데이터 마이닝 분석 모두 신뢰성 있는 결과를 얻기 위하여 k-fold cross-validation을 적용하여 10회(k=10) 실험한 결과의 평균값을 측정했으며, 실제 시스템이 전체 데이터 중 올바르게 예측한 비율과 예측한 값들 중 정확하게 예측한 비율인 Recall(재현율)과 Precision(정확도)을 측정하였다. 마지막으로 인터넷 중독 레벨별로 Young의 20개 문항의 평균을 구하고, 이를 통해 3개의

1) 구조방정식모델인 리즈렐은 사전적인 성격을 가지며 연구자가 이론을 토대로 사전에 설정한 모델이 자료에 의해 지지되는가를 검증한다. 따라서, 탐색적 접근법이라기 보다는 확인적 접근법이라 할 수 있다. 리즈렐은 모델에 포함된 관측변수의 측정오차를 고려하여 보다 정확한 결론에 도달하게 돕는다[12].

그룹(일반 사용자군, 잠재적 위험군, 고위험군) 간의 특성을 살펴본다.

4. 연구결과

4.1 인터넷 중독도

총 440명에 대한 인터넷 중독 레벨은 <표 3>과 같다. 잠재적 위험군이 53.4%로 설문 응답자 중 가장 많았다.

<표 3> 인터넷 중독 레벨

중독도	빈도(명)	비율(%)
일반사용자군	182	41.4
잠재적 위험군	235	53.4
고위험군	23	5.2
전체	440	100

성별 인터넷 중독 레벨은 <표 4>와 같다. 인터넷 중독도가 높은 잠재적 위험군과 고위험군을 더한 수치는 남학생이 전체의 29.8%로 여학생 비율인 28.9% 보다 높게 측정 되었다. 설문자 중 여학생의 수가 남학생보다 많은 것을 감안하면 남학생의 인터넷 중독도는 여학생보다 높았다.

<표 4> 성별 인터넷 중독 레벨

		중독도			전체
		일반	잠재적 위험	고위험	
성 별	남자	69	122	9	200
	여자	113	113	14	240
전체		182	235	23	440

학년별 인터넷 중독 레벨은 <표 5>와 같다. 학년이 높을수록 인터넷 레벨이 높아지는 것을 알 수 있다.

<표 5> 학년별 인터넷 중독 레벨

		중독도			전체
		일반	잠재적 위험	고위험	
학 년	초	78	26	2	106
	중	51	71	4	126
	고	53	138	17	208
전체		182	235	23	440

4.2 요인분석

본 논문에서는 다음(c1~c20)과 같이 Young의 인터넷 중독도 검사 문항[10] 20개를 가지고 분석을 실시하였다. 또한, 인터넷 중독도에 관한 요인 분석을 실시하였으며 결과는 다음 <표 6>과 같다.

c1	계획보다 더 오랫동안 인터넷에 접속한 경험
c2	인터넷 때문에 숙제를 소홀히 한 경험
c3	가족보다 인터넷에 더 흥미를 느낀 경험
c4	온라인상의 친구를 만들어 본 경험
c5	온라인 접속으로 다른 사람이 불평한 경험
c6	온라인 접속시간으로 학교생활에 문제 있는 경험
c7	다른 일 보다 먼저 전자우편을 점검한 경험
c8	인터넷 때문에 학업에 문제가 있었던 경험
c9	인터넷에서 무엇을 했느냐고 물었을 때 숨기거나 변명을 하며 얼버무린 경험
c10	인터넷 때문에 생활상의 어려운 문제를 생각 하지 못한 경험
c11	인터넷 사용 후 다시 온라인에 접속할 때까지의 기간을 기다린 경험
c12	인터넷이 없는 생활은 따분하고 공허하며 재미없을 것이라고 두려워한 경험
c13	온라인에 접속했을 때 누군가가 방해할 한다면 소리를 지르거나 화를 내거나 귀찮아 한 경험
c14	밤늦게까지 접속해 있느라 잠을 못 잔 경험
c15	오프라인 상태일 때 인터넷에 정신이 팔려 있거나 다시 온라인에 접속해 있는 듯한 환상을 느낀 경험
c16	온라인에 접속해 있을 때 "몇 분만 더" 라고 하며 시간을 허비한 경험
c17	온라인 접속 시간을 줄이는데 실패한 경험
c18	온라인 접속시간을 숨긴 경험
c19	친구나 가족과 밖으로 외출하려고 하기보다 온라인 상태에 더 머무르기 위해 접속하려고 한 경험
c20	오프라인 상태일 때에는 우울하고 침울하며 신경질이 났다가 다시 온라인 상태로 오면 이런 감정들이 모두 사라진 경험

다른 연구에서와 유사하게 4가지 요인으로 그루핑할 수 있었으며, 각 요인간에는 4~6개 사이의 하위 요인들이 속함을 알 수 있었다. 추출된 요인으로는 통제력 부족, 집착, 실제생활 부적응, 가상현실 지향으로 요약되었다. 하지만, 다른 연구와 마찬가지로 본 논문에서 추출된 4 가지 요인에 대한 하위 요인들이 동등하게 그루핑이 되지 않아 요인간의 중요도간에 차이가 있음을 알 수 있다.

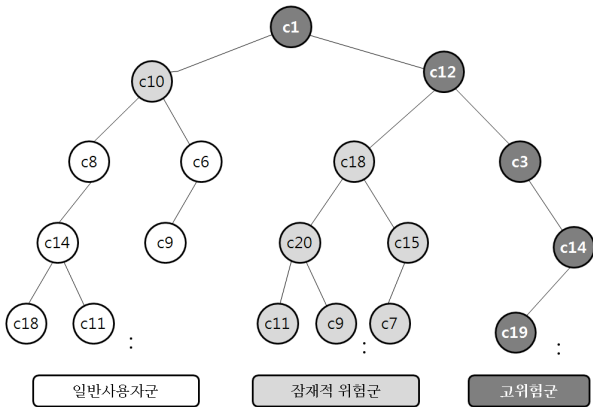
<표 6> 본 논문에서 진행한 요인분석 결과

요인	하위요인
통제력 부족	4개(c1, c2, c16, c17)
집착	5개(c5, c7, c9, c10, c18)
실제생활 부적응	6개(c11, c12, c13, c15, c19, c20)
가상현실 지향	5개(c3, c4, c6, c8, c14)

4.3 의사결정 트리를 이용한 분류분석

Weka[14]를 이용하여 J48알고리즘 기반 의사결정 트리를 구축한 결과는 다음과 같다. 구축된 트리의 크기는 63이며, 단말의 수는 총 32개이다. 교차 검증(cross-validation) 후 총 440개의 인스턴스 중 정확한 인스턴스는 364개(82.7%)로 높게 측정 되었다. 전체 데이터 중 제대로 예측한 비율

인 재현율(Recall)은 평균 84.1%, 예측한 값들 중 정확하게 예측한 비율인 정확도(Precision)는 83.8%로 둘 다 신뢰도가 높게 측정 되었다. 유도된 의사결정트리를 <그림 1>과 같이 일반사용자군, 잠재적 위험군, 고위험군으로 요약하였다.



<그림 1> 사용자군별 분류한 의사결정트리

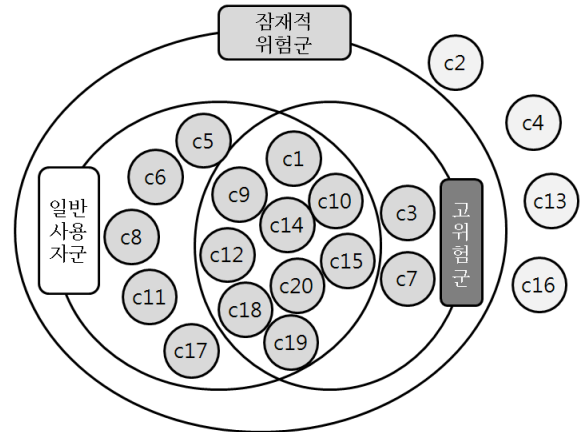
<그림 1>과 같이 사용자군별 분류한 의사결정트리를 바탕으로 세부적인 요인 및 경로를 각 사용자군별 경로의 총 개수(PathNumber), 최단경로의 길이(MinDepth), 최장경로의 길이(MaxDepth), 경로의 평균길이(AvgDepth)를 분석한 결과는 <표 7>과 같다.

<표 7> 사용자군별 분석 결과

	일반 사용자군	잠재적 위험군	고위험군
PathNumber	8	19	5
MinDepth	4	3	4
MaxDepth	10	10	6
AvgDepth	7.1	5.9	5.0

우선, PathNumber를 보면 잠재적 위험군이 가장 많았고, 다음으로는 일반사용자군, 고위험군 순으로 나타났다. MinDepth는 잠재적 위험군이 가장 낮았고, 다음으로는 일반사용자군과 고위험군이 동일하게 나타났다. MaxDepth는 고위험군이 가장 낮았고, 다음으로는 일반사용자군과, 잠재적 위험군이 동일하게 나타났다. 그리고 종합적으로 AvgDepth는 고위험군이 가장 낮았으며, 다음으로는 잠재적 위험군, 일반사용자군 순으로 나타났다. 특히, 고위험군은 PathNumber가 가장 작고, Depth 또한 낮기 때문에 다른 사용자군에 비해 보다 빠르게 인터넷 중독을 진단할 수 있음을 알 수 있다.

또한, 각 사용자군별 공통적으로 포함된 문항과 어느 군에도 포함되지 않은 문항을 추출한 결과 <그림 2>와 같이 c1, c9, c10, c14, c15, c18, c19, c20 총 9개의 문항들이 공통적으로 들어있었고, c2, c4, c13, c16 총 4개의 문항들이 어느 군에도 포함되지 않았다.



<그림 2> 사용자군별 문항 포함관계

이를 분석하면, 일반사용자군과 고위험군은 잠재적 위험군의 부분집합임을 알 수 있다. 결국 인터넷 중독도를 분류할 때 일반사용자군과 고위험군을 한번에 나누지 않는다는 것을 알 수 있다. 또한, 일반사용자군을 나누는 문항은 c5, c6, c8, c11, c17 항목들이, 고위험군은 c3, c7 항목이 관여한다는 것을 알 수 있다.

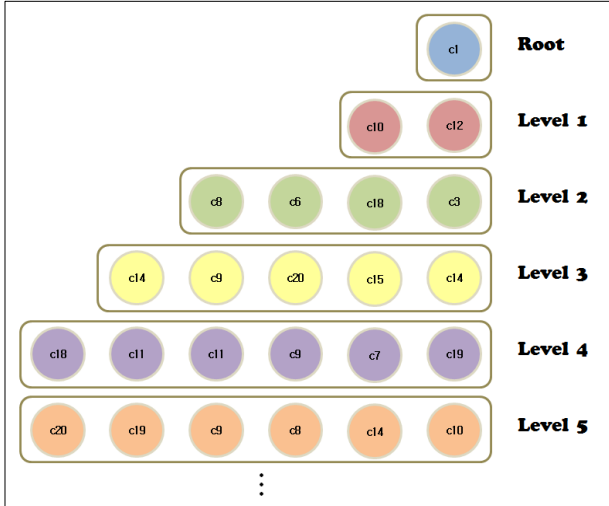
포함되지 않은 c2, c4, c13, c16 문항에 대해 응답별 평균 비율을 분석한 결과를 <표 8>에 분석하였다. 각 수치들이 한쪽에 치우치지 않고 대부분 균등하게 분포되어 있는 것으로 나타나 있다. 이는 응답 총 4개의 문항들은 인터넷 중독을 진단하는데 있어 필요하지 않음을 알 수 있다.

<표 8> 문항별 응답 비율 (단위 : %)

문항	c2	c4	c13	c16
1	22.0	23.0	31.6	14.3
2	19.1	18.0	29.5	23.2
3	31.6	28.6	21.1	28.0
4	17.3	17.3	11.6	18.4
5	10.0	13.2	6.1	16.1

<그림 1>을 바탕으로 설문 문항 중 사용자군을 결정짓는 문항의 중요도를 분석한 결과 <그림 3>과 같다. Root에 문항 c1이 위치하고, 다음으로는

Level 1에 문항 c10, c12, Level 2에 문항 c8, c6, c18, c3, Level 3에 문항 c14(2), c9, c20, c15, Level 4에 문항 c18, c11(2), c9, c7, c19, Level 5에 문항 c20, c19, c9, c8, c14, c10이 위치하고 있다.



<그림 3> 계층별 중요 설문 문항

즉, 20개의 인터넷 중독도 검사 문항에서 자기 조절을 하지 못해 인터넷을 생각했던 것보다 더 오랫동안 사용하는 ‘계획적이지 못한 온라인 접속 시간’과 관련한 문항 c1이 가장 중요한 요인이라는 것을 알 수 있다. 그리고 일상생활의 어려운 점을 인터넷을 통해 해결하려고 하고, 인터넷이 없는 생활은 생각할 수도 없는 ‘과도한 인터넷의 의존성’과 관련한 문항 c10과 문항 c12가 다음으로 중요한 요인이라는 것을 알 수 있다. 특히 문항 c10과 문항 c12는 각각 일반사용자군과 잠재적 위험군 또는 고위험군, 잠재적 위험군과 고위험군을 결정짓는 요인이기도 하다.

4.4 규칙기반 분류분석

Weka를 이용하여 규칙기반 분류분석을 실시한 결과는 다음과 같다. 추출된 규칙 중 중요도를 고려한 결과는 <표 9>와 같다. 또한 <표 10>은 규칙에서 사용된 문항의 빈도수를 나타낸 것이다. 교차 검증 후, 총 440개의 인스턴스 중 정확한 인스턴스는 370개(85.1%)로 높게 측정 되었다. 전체 데이터 중 제대로 예측한 비율인 재현율(Recall)은 평균 82.7%, 예측한 값들 중 정확하게 예측한 비율인 정확도(Precision)는 82.5%로 둘 다 신뢰

도가 높게 측정 되었다.

<표 9> 규칙 리스트

R1: $c1 \leq 3$ & $c10 > 2$ & $c6 > 1$ → 잠재적 위험군
R2: $c1 \leq 3$ & $c14 > 2$ & $c11 > 1$ → 잠재적 위험군
R3: $c1 \leq 3$ & $c8 \leq 3$ & $c18 \leq 2$ & $c20 \leq 2$ & $c9 \leq 1$ & $c15 \leq 1$ → 일반사용자군
R4: $c1 > 2$ & $c14 \leq 3$ & $c11 > 2$ & $c5 \leq 4$ & $c7 \leq 4$ → 잠재적 위험군
R5: $c11 > 3$ & $c8 > 2$ & $c3 < 3$ & $c17 > 2$ & $c19 > 2$ → 고위험군
R6: $c8 > 2$ & $c20 \leq 3$ & $c9 < 3$ & $c14 < 4$ & $c3 > 1$ → 잠재적 위험군
R7: $c8 > 3$ & $c5 > 2$ → 고위험군
R8: $c3 > 4$ & $c18 > 3$ → 고위험군
R9: $c4 > 4$ → 잠재적 위험군
R10: $c7 < 4$ & $c17 < 3$ & $c8 < 1$ & $c10 < 1$ → 일반사용자군

<표 10> 규칙 항목 사용빈도

문항	빈도수	문항	빈도수
c1	4	c11	4
c2	1	c12	1
c3	2	c14	4
c5	4	c15	2
c6	2	c17	2
c7	4	c18	3
c8	3	c19	3
c9	2	c20	1
c10	2		

규칙의 주요 분석결과를 해석하면 다음과 같다. 검사지 항목 중에서 중독단계를 결정하는 가장 상위의 규칙에서 사용한 문항은 c1임을 알 수 있다. 이는 의사결정 트리의 루트 항목인 c1과 일치한다. 또한 규칙에서 사용된 문항의 빈도수 역시 c1이 가장 많이 사용되었음을 알 수 있다. 결국 문항 c1의 내용인 자기조정체계 및 계획(Albert Bandura의 사회학습이론)이 인터넷 중독 단계를 구분하는 가장 중요한 요소임을 알 수 있다.

규칙기반 분류분석은 의사결정 트리와는 달리 인터넷과 게임 시간과 관련된 문항이 빈도수가 높은 것으로 나타났다. 특히 c15를 분석하면 인터넷을 하루 평균 2시간 이상 사용한 학생은 잠재적 위험군일 가능성이 높다.

전체 문항 중 c4, c13, c16은 한 번도 사용되지 않았으며, 문항 c2는 우선순위가 낮은 규칙에서 단 1회 사용되었다. 이는 의사결정 트리의 분석결과와 유사하며 이 문항들은 삭제할 필요가 있으며, 인터넷과 미디어의 발전에 대응하여 문항 내용에 대한 수정이 검토되어야 한다.

4.5 클래스별 문항 평균치를 통한 분석

이 절에서는 각 인터넷 중독 레벨별, 중독도를 판별하는 문항에 대한 평균치를 통해 분석한다. 결과는 다음 <표 11>과 같다. 표 안의 각 셀 안에는 문항에 대한 평균값이 저장되었다.

<표 11> 인터넷 중독 레벨별 문항들 평균

문항	레벨	일반사용자군	잠재적 위험군	고위험군
c1		2.52	3.81	4.78
c2		1.93	3.22	4.13
c3		1.43	2.40	3.60
c4		2.20	3.12	4.08
c5		1.53	2.51	3.69
c6		1.46	2.70	3.91
c7		1.76	2.20	3.39
c8		1.36	2.49	4.13
c9		1.24	2.04	3.26
c10		1.17	1.88	3.34
c11		1.51	2.62	4.17
c12		1.49	2.38	4.30
c13		1.58	2.71	3.95
c14		1.26	2.16	4.08
c15		1.07	1.59	2.82
c16		2.10	3.49	4.91
c17		1.57	2.88	3.95
c18		1.20	2.06	3.39
c19		1.25	2.33	3.65
c20		1.18	1.92	3.47

고위험군은 문항 중에서 “계획보다 더 오랫동안 인터넷에 접속한 경험”인 c1 문항과 “온라인에 접속해 있을 때 “몇 분만 더” 라고 하며 시간을 허비한 경험”인 c16문항이 특히 높았다. 두 문항 모두 시간에 대한 조절에서 어려움을 겪은 것으로 드러났다. 하지만, “오프라인 상태일 때 인터넷에 정신이 팔려 있거나 다시 온라인에 접속해 있는 듯한 환상을 느낀 경험”인 c15번 문항은 모든 레벨에서 낮은 점수를 가져 청소년들에게는 그 다지 영향을 미치지 않음을 알 수 있었다. 이 결과는 군집분석을 실시한 [15]의 연구와 일관성이 있음을 알 수 있다.

반면, 이 결과는 이전 절의 결과와 일관성이 없는 부분도 있다. 차이를 분석하면, 인터넷 중독도 분석 시, 응답값이 높은 것만이 의미를 가지는 것이 아니라 중독 레벨간 값의 차이도 의미를 가짐을 알 수 있다.

4.6 Young의 실험 중 일부문항을 이용한 인터넷 중독 검사 방법

이 절에서는 Young의 문항 중에서 분류 분석을 통해 나온 결과와 연관분석을 통해 나온 결과를 바탕으로 새롭게 인터넷 중독도를 실험하는 알고리즘을 제시한 후 어느 정도 Young의 인터넷 중독도 실험에 접근하는지를 살펴본다.

Alg_c는 분류분석에 근거를 둔 인터넷 중독도 검사 알고리즘이고 Alg_a는 규칙기반 분석에 근거를 둔 인터넷 중독도 검사 알고리즘이다.

Alg_c:

```

if (c1 > 3) {
    if (c12 >4 && c3 > 2)
        adlevel[1][j]=3;
    else adlevel[1][j]=2;
}
else {
    if (c1 >= 3 && c10 >2 && c6 >1)
        adlevel[1][j]=2;
    else if (c1 <= 3 && c14 > 2 && c11 >1)
        adlevel[1][j]=2;
    else if (c1 >= 3 && c8 > 2 && c11 > 3)
        adlevel[1][j]=3;
    else adlevel[1][j]=1;
}
    
```

Alg_a:

```

if (c1<=3 && c10 > 2 && c6 > 1)
    adlevel[1][j]=2;
if (c1<=3 && c14 > 2 && c11 > 1)
    adlevel[1][j]=2;
if (c14<=3 && c11>2 && c1>2 && c5<=4 && c7<=4)
    adlevel[1][j]=2;
if (c11 > 3 && c8 > 2 && c3 <=3)
    adlevel[1][j]=3;
if (c18>3 && c12 > 3)
    adlevel[1][j]=3;
if (c5>3 && c10 <= 2) adlevel[1][j]=2;
if (c5>3 && c7 > 2) adlevel[1][j]=3;
if (c8 > 2 && c3>1)adlevel[1][j]=2;
    
```

분류기준들 중에서 6개의 조건으로 이루어진 알고리즘 Alg_c는 Young의 결과와 86%의 일치도를 나타내었으며, 8개의 조건으로 이루어진 Alg_a는 Young의 결과와 81%의 일치도를 나타내었다. 조건의 수가 적음에도 불구하고 분류분석이 일치도 면에서는 다소 높음을 알 수 있었는데, 이는 자기

통제에 속하는 c1 항목을 우선적으로 기준을 삼은 Alg_c의 특징 때문인 것으로 드러났다.

최근 김광수의 연구[16]에서도 인터넷 중독에 가장 영향을 미치는 변인으로 인터넷 사용시간을 지적하고 있으며 그 중 자기 통제에 대한 교육을 실시해야 한다고 주장하고 있는데, 본 연구의 결과도 이와 일맥상통한다. 또한, 좀 더 중독을 확장하여 휴대폰 중독에 관한 연구[17]에서 개인요인 중에서 성별을 제외하고 자기통제력이 가장 영향을 미치는 것으로 드러났는데 이러한 연구 결과와도 관련이 된다.

5. 결 론

본 논문에서는 인터넷 중독도를 판별하는 20개의 항목에 대해서 동일한 가중치를 부여하는 Young의 척도의 문제점을 지적하고, 척도간 구조와 척도를 구성하는 항목 중에서 보다 중독도를 판별하는데 중요한 항목을 찾아내었다.

인터넷 중독예방 상담센터에서 개발한 K-척도 역시 Young과 마찬가지로 문항별로 4점에서 1점까지 점수를 주고 이들의 총점을 이용하여 고위험군, 잠재적 위험군, 일반 사용자군으로 나누고 있어 Young이 지닌 문제점을 그대로 갖고 있다.

이를 위해서 청소년 440명을 대상으로 설문을 실시하고 요인분석을 통해 Young 척도의 문제점을 기술하였다. 그 다음은 데이터 마이닝 기술을 이용하여 인터넷 중독도를 판별하는 항목들 중에서 보다 의미를 갖는 항목을 찾아내었다. 이를 통해, 인터넷 중독도 측정 시, 기존의 방법이 아닌 항목간 차별화 전략이 필요함을 알게 되었다.

향후에는 인터넷 중독도를 계산하기 위한 체계적인 통계방법을 제시하여 연구의 일반화를 높이고 보다 정확하면서도 예방까지도 고려할 수 있는 척도를 개발하고자 한다. 또한, K-척도에 대한 비교연구도 함께 병행하여 인터넷 중독 검사 방법의 완성도를 높이고자 한다. 현재 가중치에 대한 연구를 진행 중이며 추후 연구결과를 발표할 것이다.

참 고 문 헌

- [1] 박승민, 김창대, 천명재(2005). 청소년 인터넷 중독에 관한 국내 연구의 동향과 과제. **청소년상담연구**. 13(2). 3-14.
- [2] Young, K. S.(1996). Internet Addiction: emergence of a new clinical disorder. *Cyber Psycho & Behavior*. 1(3). 237-244.
- [3] <http://www.nia.or.kr>
- [4] 손경애, 고종욱, 이덕로(2003). 인터넷 중독의 결정변인들에 대한 탐색적 연구. **한국사회학**. 37(5). 177-207.
- [5] 강만철, 오익수(2001). 인터넷 중독 척도 개발을 위한 기초 연구. **청소년상담연구**. 9. 114-135.
- [6] 박경호, 강만철, 오익수, 김건웅(2001). KIAS:한국형 인터넷 중독 척도와 이를 이용한 국내 실태 조사. **한국통신학회 논문지**. 26(12). 292-304.
- [7] 김은주, 송원문, 김명원(2008). 데이터 마이닝을 이용한 인터넷 중독 진단. **한국컴퓨터종합학술대회 논문집**. 35(1). 359-363.
- [8] 한국정보문화진흥원(2006). **인터넷 이용에 대한 자가 진단**.
- [9] 박찬정(2007). 컨조인트 분석을 이용한 초·중등학교 정보통신윤리교육 수업 설계 요소 분석. **컴퓨터교육학회논문지**. 10(1). 9-20.
- [10] http://www.netaddiction.com/resources/internet_addiction_test.htm
- [11] 이종원(2004). 청소년 인터넷중독의 실태와 문제. **한국사회학회 사회학대회 논문집**. 559-565.
- [12] 배병렬(2002). **구조방정식모델 이해와 활용**. 도서출판 대경.
- [13] Anderson, James C. and Gerbing, W. (1988). "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach," *Psychological Bulletin*. 103(3). 411-423.

[14] Witten, I. and Frank, E. (2005). *Data Mining : Practical Machine Learning Tools and Techniques*. Elsevier.

[15] Park, C. and Hyun, J.(2010). "A Study on Young's Internet Overindulgence Testing Method with Lisrel and Clustering Algorithms," *Proceedings of EU-Korea Conference on Science and Technology*.

[16] 김광수(2009). 농어촌지역 청소년의 인터넷 중독에 영향을 미치는 생태체계 변인 연구. *사회연구*. 18(2). 143-174.

[17] 이연미, 이선정, 신효식(2009). 청소년의 휴대폰 중독성에 영향을 미치는 개인, 가족, 학교환경 변인. *한국가정교육학회지*. 21(3). 29-43.



박 찬 정

1988 서강대학교
전자계산학과(공학사)
1990 한국과학기술원
전산학과(공학석사)

1998 서강대학교 대학원 전자계산학과(공학박사)
1999 ~ 현재 제주대학교 컴퓨터교육과 부교수
관심분야: e-러닝, 협동학습, IT융합기술, e-윤리
E-Mail: cjpark@jejunu.ac.kr



김 형 철

2009 제주대학교
컴퓨터교육과(이학사)
2009 ~ 현재 제주대학교
교육대학원 컴퓨터교육전공
(석사과정)

관심분야: 컴퓨터교육, 수업설계, EPL,
Computational Thinking
E-Mail: k2youngc@naver.com



고 영 민

2007 제주대학교
컴퓨터교육과(이학사)
2009 ~ 현재 제주대학교
교육대학원 컴퓨터교육전공
(석사과정)

관심분야: 컴퓨터교육, EPL, Web2.0
E-Mail: libra109@hanmail.net



현 정 석

1991. 8. 서강대학교
경영학과(경영학사)
1993. 8. 서강대학교 대학원
경영학과(경영학석사)

1998. 2. 서강대학교 대학원 경영학과(경영학박사)
2002 ~ 현재 제주대학교 부교수
2007 제주대학교 연구업적 우수교수상 수상
2008 제주대학교 대학을 빛낸 교수상 수상
관심분야: 마케팅, 행동의사결정론, 트리즈
E-Mail: jshyun@jejunu.ac.kr



김 철 민

1988 서울대학교
전자계산기공학과(공학사)
1990 서울대학교 대학원
컴퓨터공학과(공학석사)

1996 서울대학교 대학원 컴퓨터공학과(공학박사)
1996 ~ 1997 서울대학교 컴퓨터신기술공동연구소
특별연구원
1997 한국학술진흥재단 부설 첨단학술정보센터
선임연구원
1997 ~ 현재 제주대학교 컴퓨터교육과 부교수
관심분야: 운영체제, EPL, Computational Thinking
E-Mail: cmkim@jejunu.ac.kr