

# 의사결정트리 기반의 분석을 통한 청소년의 컴퓨터 사용 유형별 관련 변수 추출

이혜주<sup>†</sup> · 정의현<sup>††</sup>

## 요 약

본 연구에서는 청소년의 컴퓨터 사용 유형별 관련변수의 추출 알고리즘 과정을 조사하였다. 이를 위해 한국청소년 패널 조사(KYPS)의 중 2학년의 1차년도 자료(총 3409명, 남: 1704명, 여: 1705명)를 활용하여 데이터마이닝의 의사결정트리분석을 하였다. 그 결과, 오락추구는 성별, 컴퓨터 사용시간, 비행친구 수, 부모감독, 주변문제아인식, 부모공부기대, 자아통제력, 교사애착, 형제자매사이 등, 관계추구는 성별, 사이버동아리 참여, 컴퓨터 사용시간, 자기신뢰, 온라인 비행 등, 정보추구는 학업에 대한 열중, 개인공부 시간, 낙관적 성향, 학기 중 공부시간과 여가시간, 본인희망교육, 사이버동아리 참여, 자기신뢰, 주변사람 들비난 등의 변수들이 조합되는 다양한 조건에 따라 증가하는 것으로 나타났다. 본 연구의 결과로 청소년의 컴퓨터 사용과 관련된 다양한 조건들을 심층적으로 이해하고 컴퓨터를 좀 더 효율적으로 활용할 수 있음을 제언하였다.

**주제어** : 의사결정트리, 컴퓨터 사용 유형, 데이터마이닝

## A Study of Extraction of Variables Affecting the Adolescents' Computer Use Type with Decision Tree

Hyejoo Lee<sup>†</sup> · Euihyun Jung<sup>††</sup>

### ABSTRACT

This study investigated the extraction algorithm fitting for variables of adolescents' computer use type with the sample from KYPS data (3409 in the second grade of the junior high school; 1704 boys and 1705 girls). The results of the decision tree model revealed that : (1) Gender, computer use time, misdeed friends, parent supervision, other agreement of misdeed, parent study expectation, self-control, teacher attachment, and sibling relation were significant for entertainment type. (2) Gender, cyberclub, computer use time, self-belief, online misdeed were significant for relation type. (3) Study enthusiasm, personal study time, optimistic disposition, study and spare time, cyberclub, self-belief, and other people criticism were significant for information type. These results suggest that adolescents' diverse conditions should be considered for using computer more efficiently.

**Keywords** : Decision Tree, Computer Use Type, Data Mining

---

<sup>†</sup> 정 회 원: 중앙대학교 교육학과 강사  
<sup>††</sup> 정 회 원: 안양대학교 컴퓨터학과 조교수(교신지자)  
논문접수: 2011년 12월 11일, 심사완료: 2012년 02월 24일, 게재확정: 2011년 03월 15일

## 1. 서론

현대사회에서 컴퓨터는 일상생활에서 필수적이며, 사회진반에 큰 영향력을 미치는데, 이를 어떻게 사용하는가에 따라 유익하기도 하고, 부작용을 유발하기도 한다. 컴퓨터를 통한 정보검색이나 학습활동은 능동적인 정보처리와 컴퓨터 활용능력을 향상시키고, 컴퓨터 게임은 통쾌감, 흥분, 스트레스 해소 및 자신감을 증진시킨다. 또한 카페, 이메일, 동호회, 블로그 등을 통해 친교 및 의사소통 능력이 상승된다. 그러나 컴퓨터의 과다사용, 인터넷 중독, 사이버비행 등 부정적인 행동의 가능성도 있다[1][2]. 청소년은 다른 연령층보다 컴퓨터에 많이 노출되어 있으며, 발달단계의 특성상 스스로 환경을 통제·조절하기보다는 가족, 교사, 친구, 주변 환경과의 상호작용에 의해 영향을 받는다. 따라서 컴퓨터의 영향과 관련하여 관심을 가져야 할 대상이며, 컴퓨터 사용에 미치는 여러 변수들에 대한 폭넓은 이해가 필요하다[3].

청소년의 컴퓨터 사용유형에 대해 연구마다 차이가 있지만, 대체적으로 컴퓨터 게임 등의 오락추구, 카페, 이메일, 동호회, 블로그 등의 관계추구, 공부 및 학습관련 정보검색 및 열람 등의 정보추구로 구분되며, 특징이나 관련 변수 등에서 차이가 밝혀지고 있다[4][5][6]. 그러나 지금까지의 청소년의 컴퓨터 사용에 대한 연구들은 실태 파악에 그치거나, 컴퓨터 과다사용, 인터넷 중독 등의 부작용에 집중하여, 일반적인 컴퓨터 사용유형과 관련 변수들을 분석한 연구는 드물다. 관련 변수들을 분석했다라도 몇몇 변수들의 연관성을 연구하거나 연구초점에 따라 선택적으로 변수들을 투입하여 각기 다른 결과를 제시하였다. 따라서 다양한 변수간의 가능한 상호작용을 모두 탐색할 수 없었고, 따라서 심층적인 이해에 한계가 있었다[2][4].

데이터마이닝(Data Mining)은 대용량의 데이터에 내재되어 있는 유의미한 정보를 찾아내기 위해 자동적이거나 반자동적으로 자료를 탐색, 분석하는 방법이다. 데이터마이닝 기법 중 의사결정트리(Decision Tree)는 변수 간 상호작용에서 조합 가능한 알고리즘을 모두 추출한다. 가설에 근거하여 이론적 모형을 세우고, 검증하여 일반화하

는 기존 방법과는 달리, 의사결정트리는 개별적으로 의미 있는 조건과 특성을 파악하여, 보다 구체적인 설명이 가능하다는 점에서 효율적이다[8][9].

이에 본 연구에서는 의사결정트리 기반의 분석을 활용하여 청소년의 컴퓨터 사용 유형별 관련 변수들의 상호작용과 조합관계를 밝히고자 한다. 이를 통해 청소년의 컴퓨터 사용에 대한 심층적인 이해와 올바른 컴퓨터 사용에 유용한 기초자료를 제시할 수 있을 것으로 기대한다.

## 2. 이론적 배경

### 2.1 컴퓨터 사용 유형 및 관련 변수들

컴퓨터를 선택하고, 사용하는 목적은 매우 광범위하고 다양한데, 대체적으로 컴퓨터 오락추구 유형, 다른 사람들과의 상호작용을 통한 관계추구 유형, 그리고 실용적인 정보추구 유형으로 구분된다. 이에 관련된 변수로 개인, 부모, 환경 등 여러 측면이 고려되었으며, 사용유형에 따라 다소 차이를 나타냈다[4][5]. 전반적으로 남학생은 오락추구를, 여학생은 관계추구나 정보추구를 보였고[4][10], 자기효능감과 자기 통제력이 낮은 경우에는 오락추구를 나타냈다[7][11][12]. 또한 대체적으로 부모와의 관계가 좋지 않으면 오락추구와 관계추구가, 부모와의 관계가 좋으면 정보추구가 증가하는 것으로 보고되었다[4][7][10]. 뿐만 아니라 대인관계에 대한 요구나 현실세계의 좌절된 욕구를 대리만족하기 위해서 오락추구가 증가하여 또래에착이 향상되기도 하였고, 의사소통이나 사회적 관계에 대한 요구로 인해 관계추구가 증가하기도 하였다[13]. 청소년이 현실 도피나 재미를 위해 오락추구 사용을 하거나, 현실에서 부족한 사회자본을 충족하기 위해 관계추구 사용을 하는 경우에는 컴퓨터 사용시간이 증가하여 과도한 수준에 이를 수 있지만, 정보추구 사용은 오락추구나 관계추구와 대조적인 양상을 가지며, 과도한 컴퓨터 사용으로 이어지지 않는다고 보고되었다[4]. 이와 같이, 컴퓨터 사용유형에 대해 다양한 변수들이 관련되어 있음이 제시되었으나 대부분의 연구들이 회귀분석이나 구조방정식 등 이론적 근거에 따라 몇몇 변수에 초점을 맞춰 가설이나

모형을 설정하고, 이를 검증하여 일반화하려는 기존의 통계적 기법을 사용하고 있어서 종속변수에 영향을 미치는 다양한 변수들간의 숨어있는 상호작용이나 조합관계를 모두 탐색하는데 한계가 있다[4][7][11]. 뿐만 아니라 이러한 한계점을 극복하기 위해 데이터마이닝의 의사결정트리 기반의 기법을 사용한 연구도 있었지만, 주로 인터넷 게임 중독에만 초점을 맞추고 있어서 오락추구, 관계추구, 정보추구의 사용 유형별로 영향을 미치는 변수들을 충분히 고려한 조사가 부족하다[10][12].

## 2.2 데이터마이닝 기반의 분석을 통한 관련 변수 추출 알고리즘

데이터마이닝은 방대한 양의 자료를 탐색하여, 의미있는 패턴과 규칙을 찾아내고 모형화하는 방법이다[8]. 데이터마이닝의 여러 방법 중에서 본 연구에서는 분류 모형인 의사결정트리를 활용하였다. 이 방법은 의사결정규칙에 따라 변수의 조건들이 구체적으로 달라질 때 종속변수의 발생 가능성을 실시간의 확률(%)로 계산하고 규칙의 형태로 자료를 나눌 수 있는 노드(node)를 구성한다. 따라서 각 조건이나 변수가 추가될 때마다 모형들을 연속적으로 비교가능하다[14]. 본 연구에서는 의사결정트리 중에서 CHAID(Chi-Squared Automatic Interaction Detection) 알고리즘을 채택하였는데, 이는 적합성 검정에 근거하여 마디내의 동질성을 다지 분리(multiway-split)를 하며, 가능한 모든 상호작용을 자동적으로 탐색해내는 데 효율적이다[14].

## 3. 연구 방법

### 3.1 분석자료 및 분석대상

본 연구에서는 한국청소년개발원의 한국청소년 패널자료(Korea Youth Panel Survey: KYPS)를 분석하였다. KYPS는 2003년부터 2008년까지 매년 청소년과 학부모(초 4패널과 중 2패널)를 대상으로 자기보고식으로 조사한 자료이다. 본 연구에서는 중 2패널의 1차년도(2003년) 자료를 사용하였으며, 불성실하거나 누락 문항이 많은 응답자료

를 제외한 후, 총 3409명의 자료를 분석하였다. 본 연구대상의 인구사회학적 특성을 살펴보면, 남학생이 50%(1704명), 여학생이 50%(1705명)이었고, 부의 학력은 고졸 42.9%(1464명), 대졸(29.7%, 1013명), 중졸(8.0%, 272명), 대학원석사(5.5%, 189명) 순으로, 모의 학력은 고졸 56.3%(1920명), 대졸(18.0%, 612명), 중졸(12.4%, 422명), 초졸(3.7%, 125명) 순이었다. 서울지역의 거주자는 17%(580명), 서울 제외 도시지역 거주자는 74.2%(2529명), 농촌지역 거주자는 7.0%(239명)이었다. 친부모와 함께 사는 경우는 93%(3171명)이었고, 자기집 68.3%(2330명), 전세 21.4%(729명), 월세 5.6%(192명)로 나타났다.

### 3.2 조사도구

데이터마이닝은 대용량의 자료 중에서 종속변수를 설명하는 변수들의 상호작용과 조합관계를 최대한 많이 추출하고자 하므로 예측변수의 가능성이 있는 변수들을 가급적 많이 포함시켜야 한다. 그러나 변수들의 수가 많을 때 후보 변수들간 정보가 상당히 중첩될 가능성이 있으므로 중복성이 크게 나타나는 변수들을 찾아내 제거해야 한다[14]. 본 연구에서는 패널자료에서 중복성이 큰 변수들을 제거한 후, 개인, 가정, 부모, 또래, 학교, 학업, 컴퓨터 등과 관련된 45개의 변수들을 독립변수로 설정하였다. 변수 중 5점 리커트 척도로 응답된 경우는 연구결과의 간명성을 위해 ‘낮음’, ‘보통’, ‘높음’ 또는 ‘그렇지 않다’, ‘보통’, ‘그렇다’의 세 가지 등급 분류로 사전그룹화 하여 사용하였다[9]. 각 변수에 대한 구체적인 설명은 <표 1>과 같다.

본 연구에서의 종속변수는 컴퓨터 사용유형으로, 선행연구에 따라 오락추구, 관계추구, 정보추구로 구분하였다[4][5][6]. 오락추구는 컴퓨터 게임 이용정도를 묻는 1문항을 사용하였고, 관계추구는 컴퓨터를 채팅/메신저/전자우편, 동호회/카페/커뮤니티 활동, 게시판 이용정도를 묻는 4문항으로 측정하였다(Cronbach's  $\alpha$ =.784), 정보추구는 공부 및 학습관련 또는 기타 정보검색 및 열람 이용정도를 묻는 2문항으로 측정하였다(Cronbach's  $\alpha$ =.734). 각 문항은 5점 리커트 척도로 측정되었으

<표 1> 컴퓨터 사용 유형에 대한 독립변수들

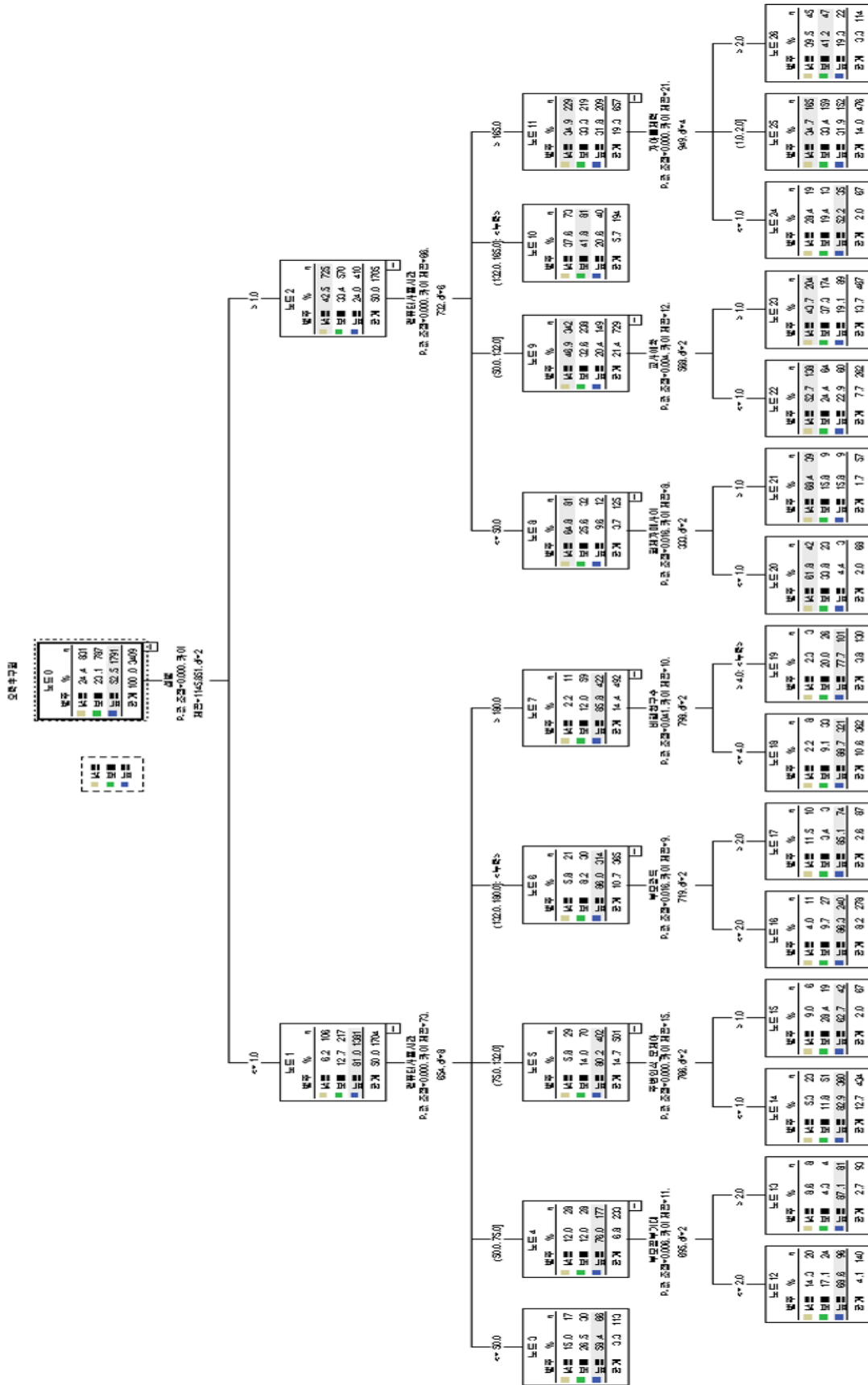
변수명	변수 설명
성별	남, 여
자아존중감	자신이 생각하는 자신의 능력, 가치(6문항, Chronbach's $\alpha=.765$ ); 낮음, 보통, 높음
자기신뢰	결정, 문제해결, 삶의 주관에 대한 신뢰(3문항, Chronbach's $\alpha=.825$ ); 낮음, 보통, 높음
자아통제력	본인 자신의 통제능력(6문항, Chronbach's $\alpha=.751$ ); 낮음, 보통, 높음
낙관적 성향	일이 잘 될 것이라는 낙관적 성향(3문항, Chronbach's $\alpha=.747$ ); 낮음, 보통, 높음
생활만족도	생활에 대한 만족도 수준(1문항); 낮음, 보통, 높음
공격성	다른 개인에 대한 공격적 성향(6문항, Chronbach's $\alpha=.764$ ); 낮음, 보통, 높음
오프라인비행	흡연, 음주, 무단결석, 가출, 성관계, 폭행, 패싸움, 돈/물건 훔치기, 도둑질, 협박, 왕따 시키기, 성폭행/성희롱 등의 경험의 횟수
스트레스수준	성적, 부모, 친구, 외모, 물질 관련 스트레스(16문항; Cronbach's $\alpha=.881$ ); 낮음, 보통, 높음
자기인식 문제아	본인을 문제아, 비행청소년으로 인식(2문항; Cronbach's $\alpha=.779$ ); 낮음, 보통, 높음
주변인식 문제아	주위사람들이 본인을 문제아, 비행청소년으로 인식(2문항; Cronbach's $\alpha=.884$ ); 낮음, 보통, 높음
주변비난걱정	나쁜 일을 했을 경우, 주위사람들의 비난걱정(2문항; Cronbach's $\alpha=.838$ ); 낮음, 보통, 높음
한달 평균 용돈	한달 평균 용돈( $M: 30567.92, SD: 24529.11$ )
가족구성형태	친부모와 있을 경우, 그 외
가구소득	월평균 가구소득( $M: 300.32, SD: 217.48$ )
거주지소유	자기집 소유, 그 외(전세, 월세, 무상, 사택, 임대 등)
거주 지역	서울지역, 서울 외 도시지역, 농촌지역
집안경제곤란	현재 집안 경제 곤란 정도(1문항); 그렇지 않다, 보통, 그렇다
가정불만정도	현재 부모의 가정불화, 이혼 등으로 느끼는 가정에 대한 불만 정도(1문항); 낮음, 보통, 높음
형제자매와의 관계	형제자매와의 관계 좋음(1문항); 해당사항 없음, 그렇지 않다, 보통, 그렇다
부모애착	부모와의 애착관계(6문항, Cronbach's $\alpha=.858$ )
부학력	무학, 초졸, 중졸, 고졸, 전문대졸, 대졸, 대학원석사, 대학원박사
모학력	무학, 초졸, 중졸, 고졸, 전문대졸, 대졸, 대학원석사, 대학원박사
부모감독	부모의 자녀에 대한 통제정도(4문항, Cronbach's $\alpha=.852$ ); 낮음, 보통, 높음
부부간폭력	부부간의 폭행 및 욕설 정도(2문항; Cronbach's $\alpha=.747$ ); 낮음, 보통, 높음
부모폭력	자녀에 대한 부모의 폭행 및 욕설 정도(2문항; Cronbach's $\alpha=.756$ ); 낮음, 보통, 높음
부모공부기대	자녀의 공부에 대한 부모 기대 감당여부(1문항); 낮음, 보통, 높음
부모희망교육	자녀에 대한 부모의 희망교육; 중졸, 고졸, 초대졸(2년제), 대졸(4년제), 대학원졸(석, 박사)
본인희망교육	본인의 희망교육; 중졸, 고졸, 초대졸(2년제), 대졸(4년제), 대학원졸(석사, 박사)
학교공부흥미	현재 학교 공부에 대한 흥미(1문항); 낮음, 보통, 높음
학업에 대한 열중	국어/영어/수학/사회/과학에서의 학업에 대한 열중(5문항, Cronbach's $\alpha=.780$ ); 낮음, 보통, 높음
학업성적수준	국어/영어/수학/사회/과학의 학업성적수준(5문항, Cronbach's $\alpha=.790$ ); 낮음, 보통, 높음
공부고민	공부 때문에 많은 고민하고 있는 정도(1문항); 낮음, 보통, 높음
개인공부시간	일주일 평균 국어/영어/수학/사회/과학 개인공부시간(시간)
학기중공부시간	학기 중 평일과 주말 동안의 공부시간(학교수업 및 개인학습)(시간)
학기중여가시간	학기 중 평일과 주말 동안의 여가시간(시간)
교사애착	교사와의 애착관계(3문항, Cronbach's $\alpha=.703$ ); 낮음, 보통, 높음
또래애착	또래와의 애착관계(4문항, Cronbach's $\alpha=.761$ ); 낮음, 보통, 높음
매우친한친구수	매우 친한 친구수 합산
비행친구수	친한 친구 중 처벌(근신/정학/퇴학), 술, 담배, 무단결석, 폭행, 도둑질한 친구의 수 합산
컴퓨터사용시간	하루 평균 컴퓨터 사용시간(분)
온라인일탈	지난 1년 동안 인터넷상 비윤리적 행위(허위정보유포, 불법 소프트웨어 복제/사용, 타인 ID/주민번호 도용, 채팅시 성/나이 속이기, 해킹, 온라인상 욕설/폭언 등)의 횟수 합산
오프라인 동아리참여	교내공식동아리 참여여부; 그렇지 않다, 그렇다
사이버동아리 참여	사이버 동아리 참여여부; 그렇지 않다, 그렇다
한달사교육비	평균 한 달동안 지불하는 사교육비

나, 연구결과의 간명성을 위해 ‘낮음’, ‘보통’, ‘높음’의 세 등급 분류로 사전그룹화하였다.

### 3.3 분석방법

본 연구에서는 분석을 위해 데이터마이닝의 의사결정트리의 CHAID 알고리즘을 채택하였으며, SPSS 14.0 Classification Tree를 활용하였다. 이 방법은 구조화를 통해 분리 규칙을 추론하며, 변수적용 시 각 변수의 마디에서 실제빈도와 기대

빈도 간에 차이가 없다는 영가설을  $\chi^2$  검증한다. 이 값이 기각되면( $p<.05$ ) 각 마디의 구간 특성에 따라 분리되고, 이 과정을 반복하면서 분류트리를 형성한다. CHAID 알고리즘을 위한 독립변수들의 분리와 병합의 기준은 .05 수준이었으며, 자동적으로 집단 내의 구간 변경을 감안해 값에 대한 Bonferroni 조정을 거쳤다[15]. 결측치는 특정마디에서 예측이 일치하는 순서에 따라 대체규칙(surrogate rule)을 설정하였고, 정지규칙으로 최대한의 트리깊이(maximum tree depth)는 5수준



<그림 1> 청소년의 오락추구형 의사결정트리모양

이었으며, 부모마디(parent node)와 자식마디(child node)의 사례수는 각각 100과 50으로 지정하였다 [14].

#### 4. 분석 결과

<그림 1>은 오락추구형의 CHAID 알고리즘을 나타내는 다중트리구조(multi-tree structure)이다. 가장 중요한 변수는 성별이며, 총 9개의 독립변수가 오락추구형의 의사결정모형을 구성하였다. 9개의 변수는 모두 17개의 가지의 알고리즘으로 형성되었고 각 조건은 최대 5개의 변수를 가지고 있었다.

오락추구는 전체의 52.5%(3409명)이었고, 변수들의 조합에 따라 기대빈도가 달라졌다. 남학생이고 컴퓨터 사용시간이 하루 평균 180분 초과며 비행친구수가 4명이하면 88.7%의 가장 높은 기대빈도를 보였다(노드 18). 남학생이고 컴퓨터 사용시간이 132분초과~180분이하, 부모감독이 보통이하면 86.3%, 남학생이고 컴퓨터 사용시간이 75분초과~132분이하, 주변이 본인을 문제아로 낮게 인식하면 기대빈도가 82.9%로 나타났다. 또한 남학생이고 컴퓨터 사용시간이 50분초과~75분이하, 부모공부기대가 높으면 기대빈도가 87.1%로 나타났다. 여학생의 경우, 컴퓨터 사용시간이 165분초과, 자아통제력이 낮으면 기대빈도가 52.2%이었지만, 여학생의 컴퓨터 사용시간이 50분초과~132분이하, 교사애착이 낮으면 기대빈도가 22.9%이었다. 여학생의 컴퓨터 사용시간이 50분이하, 형제자매사이가 보통이상이면 15.8%이었다. 산출된 의사결정모형의 적합성 여부를 판별하기 위하여 단순임의추출법에 의하여 훈련집단과 타당화집단을 70:30으로 구분한 후, 교차타당성 평가를 실행하였다. 그 결과, 훈련집단(위험추정치=0.373,  $SE=.010$ )과 타당화집단(위험추정치=0.360,  $SE=.015$ )이 거의 차이가 없는 것으로 나타나 모형의 일반화에 무리가 없는 것으로 판단되었다. 산출된 모형의 효율성을 판단하고 알고리즘의 타당성 검토를 위해 <표 2>의 이익도표를 작성하였다. <표 2>의 가장 상단에 있는 18번은 가장 높은 기대빈도를 나타낸다. 즉, 아무런 자료 없이 오락추구형을 판단하는 노드 0번의 52.5%보다

168.8% 향상된 의사결정을 의미한다. 이런 판단에 성별, 컴퓨터 사용시간, 비행친구수 등의 변수들이 활용되었다. 반면, 노드 20은 문항 내 전체 응답자 68명 중 4.4%로 8.4%의 낮은 이득률을 보였으며, 성별, 컴퓨터 사용시간, 형제자매사이 등의 변수들이 활용되었다. 이익도표에서 모형의 효율성을 높이기 위해서는 개별이익지수가 100% 이하가 되는 24, 25, 22, 10, 26, 23, 21, 20번의 마디를 가지치기할 수 있다.

위와 같은 방식으로 관계추구형과 정보추구형의 의사결정 알고리즘을 추출하고, 이익도표를 작성하였으며, 이를 요약하여 제시하면 <표 3>과 같다. 관계추구형은 전체 중 28.1%(958명)이며 가장 중요한 변수는 성별로 나타났지만, 오락추구형과 달리, 여학생이 남학생보다 더 많은 기대빈도를 나타냈다. 여학생이고(39.5%), 사이버동아리에 참여하며(62.0%), 컴퓨터 사용시간이 235분초과면 79.2%로, 그리고 235분이하이면 57%의 의사결정을 하였다. 사이버동아리에 참여하지 않아도(32.9%), 컴퓨터 사용시간이 235분초과면 47.7%, 120분초과~235분미만이면 36.4%로 나타났다.

또한 남학생이고(16.8%), 사이버동아리에 참여하며(36.8%), 자기신뢰가 높으면 기대빈도가 53.6%로 나타났다. 사이버동아리에 참여하지 않아도(11.5%), 온라인 비행이 7회초과~8회이하면 12.5%, 8회초과면 17.3%로 나타났다. 단순임의추출법에 의한 훈련집단과 타당화집단을 대상으로 교차타당성 평가한 결과, 훈련집단(위험추정치=0.481,  $SE=.010$ )과 타당화집단(위험추정치=0.480,  $SE=.016$ )의 차이가 거의 없는 것으로 나타나 모형의 적합성을 나타내었다.

정보추구형은 전체 중 13.0%(441명)로, 학업에 대한 열증이 가장 중요한 변수로 추출되었다. 학업에 대한 열증이 높고(20.1%), 일주일평균 개인 공부시간이 10시간초과면 기대빈도가 26.4%였지만, 4시간초과~10시간, 4시간이하이면 19.0%, 8.9%로 나타났다. 개인공부시간이 10시간초과일 때, 낙관적 성향이 높으면 기대빈도가 36.1%였지만, 보통이하면 20.4%로 나타났다. 개인공부시간이 4시간초과~10시간이어도 주말을 포함한 학기 중 공부시간이 15시간이상이면 기대빈도가 26.0%로 나타났다. 학업에 대한 열증이 보통이고

<표 2> 오락추구형을 판별해주는 각 마디의 이익도표

Node	Node: n(%)	Resp: n(%)	Gain(%)	Index(%)	Cum Index(%)
18	362(10.6)	321(17.9)	88.7	168.8	168.8
13	93(2.7)	81(4.5)	87.1	165.8	168.2
16	278(8.2)	240(13.4)	86.3	164.3	166.7
17	87(2.6)	74(4.1)	85.1	161.9	166.2
14	434(12.7)	360(20.1)	82.9	157.9	163.3
19	130(3.8)	101(5.6)	77.7	147.9	161.9
12	140(4.1)	96(5.4)	68.6	130.5	159.0
15	67(2.0)	42(2.3)	62.7	119.3	157.3
3	113(3.3)	66(3.7)	58.4	111.2	154.3
24	67(2.0)	35(2.0)	52.2	99.4	152.2
25	476(14.0)	152(8.5)	31.9	60.8	132.8
22	262(7.7)	60(3.4)	22.9	43.6	123.5
10	194(5.7)	40(2.2)	20.6	39.2	117.5
26	114(3.3)	22(1.2)	19.3	36.7	114.2
23	467(13.7)	89(5.0)	19.1	36.3	103.1
21	57(1.7)	9(0.5)	15.8	30.1	101.9
20	68(2.0)	3(0.2)	4.4	8.4	100.0

Node: 마디번호  
 Node: n(%): 해당마디번호에서의 사례수와 전체 사례수에 대한 백분율  
 Resp: n(%): 목표범주의 사례수와 전체에서 목표범주의 사례수에 대한 백분율  
 Gain(%): 목표범주의 사례수/사례수  
 Index(%): 목표범주의 비율/전체 목표범주의 비율  
 Cum Index(%): Cumulative Index(%)로 노드 및 사례수의 누계에 따른 Index(%)

<표 3> 관계추구형과 정보추구형의 의사결정모형 요약표

컴퓨터 사용	Depth1				Depth2				Depth3			
	변수	Rs	Gain (%)	Index (%)	변수	Rs	Gain (%)	Index (%)	변수	Rs	Gain (%)	Index (%)
관계추구형 (958명, 28.1%)	성별	M	16.8	59.8	사이버 동아리 참여	A	36.8	131.0	자기신뢰	A	53.6	190.4
						BC	24.0	85.3	BC	24.0	85.3	
						8VD	17.3	61.6	8VD	17.3	61.6	
						7VD~8V	12.5	44.3	7VD~8V	12.5	44.3	
						7VH	9.1	32.5	7VH	9.1	32.5	
						235KD	79.2	281.3	컴퓨터 사용시간	235KD	79.2	281.3
	F	39.5	140.6	사이버 동아리 참여	C	11.5	40.9	온라인 비행	235KH	57.0	202.4	
					235KD	47.7	169.4	235KD	47.7	169.4		
					120KD~ 235KE	36.4	129.2	컴퓨터 사용시간	120KD~ 235KE	36.4	129.2	
					75KD~ 120KE	28.4	101.0	75KD~ 120KE	28.4	101.0		
					50KD~ 75KE	21.5	76.3	50KD~ 75KE	21.5	76.3		
					50KH	11.2	39.8	50KH	11.2	39.8		
정보추구형 (441명, 13.0%)	학업에 대한 열중	A	20.1	154.6	개인 공부 시간	10JD	26.4	203.1	낙관적 성향	A	36.1	277.8
						4JD~ 10J	19.0	146.2	학기 중 공부시간	BC	20.4	157.0
						4JH	8.9	68.2	4JH	8.9	68.2	
						S	19.2	147.7	사이버동아리 참여	A	23.3	179.7
						T	9.4	72.3	사이버동아리 참여	C	17.6	135.6
						UH	7.9	60.8	학기 중 여가시간	A	12.9	99.6
	B	10.3	79.2	본인 희망 교육	C	8.4	64.3	C	8.4	64.3		
					29JD	1.8	14.0	29JD	1.8	14.0		
					29JH	11.5	88.2	29JH	11.5	88.2		
					A	13.2	101.3	A	13.2	101.3		
					A	6.9	53.1	자기 신뢰	A	6.9	53.1	
					BC	5.1	39.2	주변비난걱정	B	4.8	37.3	
C	4.1	31.2	C	4.1	31.2							

Rs: 문항에 대한 응답의 종류로서 'M'은 '남학생', 'F'는 '여학생', 'A'는 '그렇다' 또는 '높음', 'B'는 '보통', 'C'는 '그렇지 않다' 또는 '낮음', 'D'는 '초과', 'E'는 '미만', 'G'는 '이상', 'H'는 '이하', 'J'는 '시간', 'K'는 '분', 'S'는 '대학원졸(석사, 박사)', 'T'은 '대학졸(2년제, 4년제)', 'U'는 '고졸', 'V'는 '횡수'

(10.3%), 본인희망교육이 대학원졸(석·박사)이면 19.2%이었다. 여기에 사이버동아리에 참여하면 기대빈도가 23.3%로 나타났다. 본인희망교육이 대학졸이고(9.4%), 사이버동아리에 참여하면 12.9%이었다. 본인희망교육이 고졸이하이고(7.9%), 학기 중 여가 시간이 29시간 이하면 기대빈도가 11.5%로 나타났다. 또한 학업에 대한 열중이 낮아도(6.9%) 자기신뢰가 높으면 13.2%로 나타났다. 자기신뢰가 보통이하여도(5.1%), 주변비난걱정이 높으면 기대빈도가 6.9%이었다. 모형에 대한 훈련집단과 타당화집단의 교차타당성 평가 결과, 훈련집단과 타당화집단의 위험추정치가 각각 0.464( $SE=0.010$ ), 0.461( $SE=0.015$ )로써 두 집단 차이가 거의 없는 것으로 나타나 일반화에 무리가 없는 것으로 판단되었다.

## 5. 결 론

본 연구에서는 데이터마이닝의 의사결정트리분석을 활용하여 청소년의 컴퓨터 사용 유형별 관련 변수의 추출을 위한 알고리즘 과정을 산출하고, 관련 변수들의 상호작용과 조합관계를 조사하였다. 그 결과 컴퓨터 사용 유형별로 관련 변수들의 다양한 조합들을 발견하였다.

첫째, 청소년의 오락추구와 관련된 변수는 성별, 컴퓨터 사용시간, 비행친구수, 부모감독, 주변문제아인식, 부모공부기대, 자아통제력, 교사애착, 형제자매사이로 나타났다. 이러한 변수들의 조합에 따라 오락추구의 기대빈도가 높아졌다. 즉, 남학생-컴퓨터 사용시간(180분초과)-비행친구수(4명 이하), 남학생-컴퓨터 사용시간(132분초과~180분 이하)-낮은 부모감독, 남학생-컴퓨터 사용시간(75분초과~132분 이하)-보통이하의 주변문제아인식, 남학생-컴퓨터 사용시간(50분초과~75분 이하)-높은 부모공부기대, 여학생-컴퓨터 사용시간(165분초과)-낮은 자아통제력, 여학생-컴퓨터 사용시간(50분초과~132분 이하)-낮은 교사애착, 여학생-컴퓨터 사용시간(50분 이하)-보통이상의 형제자매사이의 경우 오락추구의 기대빈도가 증가하였다. 이러한 결과는 남학생의 오락추구 사용률이 높고, 자아통제력과 부모감독이나 교사애착이 낮을 때 오락추구가 증가한다는 선행연구들의 결과와 일치

한다[4][5][7][10][11][12]. 그러나 많은 컴퓨터 사용시간은 오락추구의 기대빈도가 커진다는 본 연구의 결과는 청소년이 현실로부터의 도피와 재미로 컴퓨터를 사용하면 컴퓨터 사용시간이 과도하게 된다는 연구와[4] 역방향의 관련성을 제시한다. 또한, 부모공부기대는 부모가 자녀에 대해 가지는 포부수준으로 자녀가 해주기 바라는 공부나 교육적 수준을 말한다. 본 연구에서는 부모공부기대가 높으면 오락추구의 기대빈도가 높게 나타났는데, 이는 부모공부기대가 높으면 성취압력으로 작용하여, 스트레스가 가중되고[16], 이를 해소하기 위해 오락추구가 증가하는 역효과를 가져올 수 있음을 보여준다. 또한 한편, 본 연구에서는 비행친구가 4명 이하일 때, 오락추구의 기대빈도가 높게 나타났는데, 이는 남학생의 경우 컴퓨터 게임을 통해 또래애착이 증가했다는 연구결과와 다소 상반된다[13]. 또한 비행친구수와 오락추구간의 관계는 두 변수 중 원인적 요소와 결과적 요소를 무엇으로 하느냐에 따라 해석이 달라질 수 있다. 본 연구에서는 오락추구를 결과적 요소로 설정하여 분석하였으므로, 본 연구의 결과에 대해 비행친구가 많으면 컴퓨터 사용보다는 실제로 비행을 저지를 확률이 높아지므로 비행친구수가 오히려 적을 때, 오락추구가 증가하였다고 해석가능하다. 이와 같은 이유로 주변문제아 인식이 높을 때 오락추구의 기대빈도가 낮게 나타난 결과도 해석할 수 있다. 그러나 비행친구수나 주변문제아인식과 같은 변수와 오락추구형과의 관계의 원인적·결과적 요소와 관련된 다양한 해석이 가능한데 비해 이에 대한 구체적이고 실질적인 연구가 부족하므로 본 연구결과에 대한 후속연구가 이루어져야 할 것이다.

둘째, 관계추구형은 성별, 사이버동아리 참여, 컴퓨터 사용시간, 자기신뢰, 온라인 비행 등의 변수들의 상호작용과 조합으로 기대빈도가 높게 나타났다. 즉, 여학생-사이버동아리 참여-컴퓨터 사용시간(235분 이상), 여학생-사이버동아리 비참여-컴퓨터 사용시간(120분초과), 남학생-사이버동아리 참여-높은 자기신뢰, 남학생-사이버동아리 비참여-온라인비행 횟수(7회초과)일 때, 관계추구의 기대빈도가 높게 나타났다. 이러한 결과는 사이버동아리 참여가 관계추구를 더 증가시켰다는 선행



연구와 일치한다[4][10]. 특히, 사이버동아리 참여는 청소년의 의사소통이나 사회적 관계에 대한 욕구의 표현으로 볼 수 있다. 컴퓨터 사용시간이 증가할수록 관계추구형이 증가한다는 본 연구의 결과는 청소년이 현실에서 부족한 사회자본을 충족하기 위해 컴퓨터를 사용하며, 그 결과로 컴퓨터 사용시간이 증가하였다는 선행연구[4]와 역방향의 관련성을 제시한다. 오락추구형에서의 마찬가지로 추후연구를 통해 두 변수를 충분히 탐색할 필요가 있다.

셋째, 정보추구형은 학업에 대한 열중, 개인공부시간, 낙관적 성향, 학기 중 공부시간과 여가시간, 본인희망교육, 사이버동아리 참여, 자기신뢰, 주변비난걱정 등의 변수들의 조합적 관계로 기대빈도가 높게 나타났다. 높은 학업에 대한 열중-개인공부시간(10시간이상)-높은 낙관적 성향, 개인공부시간(4시간초과~10시간)-학기공부시간(15시간이상), 보통의 학업에 대한 열중-본인희망교육(대학졸), 대학원졸(석·박사)-사이버동아리 참여, 본인희망교육(고졸이하)-학기여가시간(29시간이하), 낮은 학업에 대한 열중-높은 자기신뢰, 보통이하의 자기신뢰-높은 주변비난걱정의 경우 정보추구의 기대빈도가 높게 나타났다. 오락추구나 관계추구와 달리 정보추구를 설명하는 변수들에서 대조적인 양상을 보였는데, 이는 정보추구를 학습 목적으로 보는 선행연구와도 일치한다[4].

이처럼 본 연구에서는 표집의 특성과 척도의 제약이라는 방법론에 구애받지 않는 데이터마ining의 의사결정트리 기법을 활용해 오락추구, 관계추구, 정보추구를 증가시키는 변수 추출을 위한 알고리즘을 제시하여 지금까지 컴퓨터 사용유형과 관련된 단일변수들의 확인에 치중했던 기존 연구와는 달리, 컴퓨터 사용유형을 증가시키는 조합들을 추출할 수 있었다. 이를 통해 청소년의 컴퓨터 사용 형태를 보다 정확하게 이해하고 판단하며, 아울러 상황에 따라 요구되는 컴퓨터 사용유형을 활성화시키는 등 효율적인 컴퓨터 사용을 위한 유용한 기초자료로 활용할 수 있을 것이다.

## 참 고 문 헌

- [1] 김순규 (2009). 인터넷 사용이 청소년의 학교 생활 부적응에 미치는 영향. **청소년학연구**, 16(1), 1-22.
- [2] 손진희·손은령 (2007). 가정환경, 컴퓨터 과다사용 및 인터넷 관련 비행과의 관계. **한국청소년연구**, 18(1), 135-162.
- [3] Gross, E. F. (2004). Adolescent Internet use: What we expect, what teens report. *Applied Developmental psychology*, 25, 633-649.
- [4] 이영글·김아래미·김주일 (2011). 청소년의 사회자본이 인터넷 사용에 미치는 영향. **한국청소년연구**, 22(3), 231-259.
- [5] 장현미·김반야 (2008). 인터넷 이용 유형이 초등학생의 인터넷 일탈에 미치는 영향: 잠재성장모형을 적용하여. **한국청소년연구**, 19(4), 199-225.
- [6] Weiser, E. (2001). The functions of Internet use and their psychological consequences. *CyberPsychology and Behavior*, 4, 723-744.
- [7] 박선희, 전경자(2007). 중학생들의 인터넷 사용수준 증가에 영향을 미치는 요인들. **청소년학연구**, 14(2), 71-93.
- [8] 김영숙 (2009). 아동사례관리의 욕구사정 정확도 향상을 위한 사정도구 개발과 욕구추출 알고리즘 과정 연구: 데이터 마이닝 분석기법을 활용하여. **사회복지정책**, 36(2), 167-189.
- [9] 이주리 (2009). Data Mining을 이용한 초등학생의 삶의 만족도에 대한 보호요인 및 위험요인 탐색. **아동학회지**, 30(1), 11-25.
- [10] 김기숙·김경희 (2010). 의사결정나무 분석기법을 이용한 청소년의 인터넷 게임중독 영향요인 예측모형 구축. **대한간호학회지**, 40(3), 378-388.
- [11] 이경남 (2003). 아동이 지각한 어머니와의 의사소통과 자기통제가 게임중독에 미치는 영향. **대한가정학회지**, 4(1), 77-91.
- [12] 주석진, 좌동훈(2011). 청소년들의 인터넷 게임중독에 미치는 예측모형: 인구사회학적 특성과 가족관련 특성을 중심으로. **청소년학연구**, 18(5), 165-190.

- [13] Suler, J. (2004). Computer and cyberspace addition. *International Journal of Applied Psychoanalytic Studies*, 1(4), 359-362.
- [14] Witten, I. H., & Frank, E. (2005). *Data mining: Practical machine learning tools and techniques*. San Francisco, MA: Morgan Kaufmann.
- [15] Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 29(2), 119-127.
- [16] Campbell, J. R. (1994). Developing cross-cultural/cross-national instrument: Using cross-national methods and procedures. *International Journal of Educational Research*, 21, 675-684.



## 이혜주

1992 동덕여자대학교 아동학과(이학사)  
 1995 동덕여자대학교 아동학과(이학석사)  
 2005 이화여자대학교 초등교육과(문학박사)  
 2005 한국교육개발원 부연구위원대우  
 2004~현재 중앙대학교 교육학과 외래강사  
 관심분야: 창의성교육, u-러닝, 교수방법  
 E-Mail: ladyzen@naver.com



## 정의현

1992 한양대학교 전자공학과(공학사)  
 1994 한양대학교 전자공학과(공학석사)  
 1999 한양대학교 전자공학과(공학박사)  
 1999~2002 대우통신 선임연구원  
 2002~2003 SCT 연구소장  
 2004~현재 안양대학교 컴퓨터학과 조교수  
 관심분야: 데이터마이닝, 시맨틱웹, 식별체계 등  
 E-Mail: jung@anyang.ac.kr