

# 데이터 마이닝을 사용한 방학 중 학습방법과 학업성취도의 관계 분석

(Effective Studying Methods during a School Vacation: A  
Data Mining Approach)

김혜숙<sup>†</sup>    문양세<sup>\*\*</sup>    김진호<sup>\*\*\*</sup>    노웅기<sup>\*\*\*\*</sup>  
(Hea-Suk Kim)    (Yang-Sae Moon)    (Jinho Kim)    (Woong-Kee Loh)

**요약** 학업성취도 향상을 위해 정규 수업 이외에도 과외, 학원수강, 교육방송 시청 등 다양한 교육이 이루어지고 있다. 본 논문에서는 방학 중 학습방법과 생활습관이 학업성취도 변화에 미치는 영향을 분석하기 위한 데이터 마이닝 접근법을 제안한다. 우선, 학업성취도에 영향을 미치는 방학 중의 학습방법과 생활습관에 대한 다양한 요소를 도출한다. 다음으로, 마이닝 기법 중 의사결정트리와 연관 규칙을 사용하기 위한 데이터 변환 및 분석 방법을 제안한다. 마지막으로, 설문조사를 통해 수집한 현실의 구체적 데이터에서 의사결정트리를 생성하고 연관 규칙을 추출하는 방법을 제안한다. 중학생들에 대한 설문조사를 분석한 결과, 의사결정트리의 경우 네 가지 의미있는 결과를 도출하였다. 첫째, 상위권 학생들의 경우 학원수강이 성적을 향상시키는 것으로 나타났다. 둘째, 대부분 학생들의 경우 인터넷 학습사이트 이용은 성적을 하락시키는 것으로 나타났다. 셋째, 성적 변화에 큰 영향을 미칠 것으로 예상했던 과외는 실제로 큰 영향을 미치지 못하는 것으로 나타났다. 넷째, 다양한 학습방법의 병행은 오히려 성적 하락의 요인이 되는 것으로 파악되었다. 다음으로, 연관 규칙 추출 결과, 방학 중 활동 사이에는 특이한 규칙이 없는 것으로 나타났다. 본 논문에서 제시한 데이터 마이닝 접근법 및 결과는 학생들의 방학 중 생활 지도나 학습 계획 수립에 많은 도움이 될 수 있다고 사료된다.

**키워드** : 학업성취도 분석, 효율적 학습방법, 데이터 마이닝, 의사결정트리, 연관 규칙

**Abstract** To improve academic achievement, the most students not only participate in regular classes but also take various extra programs such as private lessons, private institutes, and educational TV programs. In this paper, we propose a data mining approach to identify which studying methods or usual life patterns during a school vacation affect changes in the academic achievement. First, we derive various studying methods and life patterns that are thought to be affecting changes in the academic achievement during a school vacation. Second, we propose the method of transforming and analyzing data to apply them to decision trees and association rules, which are representative data mining techniques. Third, we construct decision trees and find association rules from the real survey data of middle school students. We have discovered four representative results from the decision trees. First, for students in the higher rank, there is a tendency that private institutes give a positive effect on the academic achievement. Second, for the most students, the Internet learning sites may give a negative effect on the achievement. Third, private lessons that have thought to be making a large impact to the achievement, however, do not make a positive effect on the achievement. Fourth, taking several studying methods in parallel may give a negative effect on the achievement. In association rules, however, we cannot find any meaningful relationships between academic achievement and usual life patterns during a school vacation. We believe that our approach will be very helpful for teachers

· 본 연구는 강원대학교 학술연구조성비와 첨단정보기술연구센터(AITrc)를 통하여 한국과학재단(KOSEF)의 지원을 받았다

† 학생회원 : 강원대학교 컴퓨터학과  
hskim@kangwon.ac.kr

\*\* 정회원 : 강원대학교 컴퓨터학과 교수  
ysmoon@kangwon.ac.kr

\*\*\* 종신회원 : 강원대학교 컴퓨터학과 교수  
jhkim@kangwon.ac.kr

\*\*\*\* 정회원 : Univ. of Minnesota, Dept. of CSE 연구원  
loh@cs.umn.edu

논문접수 : 2006년 1월 31일  
심사완료 : 2006년 11월 20일

and parents to give a good direction both in preparing a studying plan and in selecting studying methods during a school vacation.

**Key words** : Analysis of academic achievement, Efficient studying method, Data Mining, Decision Tree, Association Rules

## 1. 서론

데이터 마이닝은 대용량 데이터에서 정보 및 지식을 발견하는 대표적 분석 기법 중 하나이다[1]. 정보 기술의 발전에 따라 기업들은 수십 GB 이상의 방대한 데이터를 관리하기에 이르렀고, 기업의 관심은 데이터를 관리하는 개념을 넘어 데이터에서 보다 가치있는 정보나 지식을 찾고자 하는 부분으로 집중되고 있다. 이러한 요구에 따라 1990년대 초반부터 데이터 마이닝 연구가 활발하게 진행되었다[1,2]. 그 결과, SAS Enterprise Miner[3], IBM Intelligent Miner[4], DBMiner[5] 등의 데이터 마이닝 제품이 상용화되었다. 또한, 금융, 통신, 경영, 마케팅 등의 많은 분야에서 데이터 마이닝이 활용되고 있다[2,6].

본 논문에서는 교육 분야의 학업성취도 분석에 데이터 마이닝 기법을 활용한다. 교육에서의 학업성취도(academic achievement)란 학습의 결과로 지식과 기능을 습득하는 과정이나 결과를 의미한다[7]. 현재 우리나라에서는 학업성취도 상승을 위해 학교 수업 이외에 과외, 학원수강, 교육방송 등의 다양한 교육 방법들이 동원되고 있다. 이러한 현실에서, 교육평가의 기능이 제대로 발휘되기 위해서는 다양한 교육 방법만큼 다양한 관점에서의 학업성취도 분석이 이루어져야 한다. 그런데, 최근까지 진행되어 온 학업성취도 연구는 주로 정규학기 중 학습방법이나 활동에 집중되어 왔다. 특히, 1년 중 4개월에 해당하는 방학과 학업성취도의 관계에 대한 연구는 미비한 형편이다. 더구나, 최근 방학은 학업의 연장선이라는 관점에서 교사들이나 학부모들의 높은 관심을 받고 있다. 이에 따라 본 연구에서는 방학 중의 학습방법이나 생활이 학업 성취도에 미치는 영향을 체계적으로 분석한다.

학생들의 방학 중 학습 및 활동에 따른 학업성취도 변화를 분석하기 위하여 본 논문에서는 데이터 마이닝 기법을 적용한다. 이를 위해, 다음과 같은 네 단계로 연구 및 분석을 수행한다. 첫째, 여러 데이터 마이닝 기법들 중에서 방학 중 효과적인 학습방법 및 연관성 있는 생활습관을 찾아내기 위한 마이닝 기법을 선택한다. 둘째, 선택한 데이터 마이닝 기법을 적용하기 위해 학업성취도에 변화를 줄 수 있는 항목을 도출하고, 이들 항목을 기반으로 데이터 분석 방법을 구체화한다. 셋째, 방학 중 학습방법이 학업성취도에 영향을 미치는 요소를

도출하기 위해 1) 과거성적, 2) 방학 중 학습방법 및 생활습관, 3) 현재성적으로 구성되는 구체적 데이터를 설문조사를 통해 수집한다. 넷째, 수집한 데이터를 대상으로, 데이터 변환 및 정제를 거쳐, 상용 데이터 마이닝 도구인 SAS社의 Enterprise Miner(간략히 E-Miner라 한다)를 사용하여 의미있는 규칙과 패턴을 추출하고 해석한다.

본 논문에서 사용한 데이터 마이닝 기법은 의사결정트리(decision tree)[8-10]와 연관 규칙(association rule)[11,12]이다. 먼저, 의사결정트리는 방학 중 학습방법에 따른 학업성취도 변화를 분석하기 위해 사용한다. 다음으로, 연관 규칙은 방학 중 학습 방법이나 생활습관 사이에 특이한 규칙이 있는지를 파악하기 위해 사용한다. 방학 중 학습방법 및 생활습관이 학업성취도에 미치는 영향을 분석한 본 연구 결과는 교사나 학부모가 학생의 생활 및 학습 지도를 하는데 유용하게 사용될 수 있다. 또한 학생들에게는 방학 중 학습 계획을 세우는 자료로 활용될 수 있다.

본 논문의 구성은 다음과 같다. 제2장의 관련연구에서는 학업성취도 관련 기존 연구와 교육 분야의 데이터 마이닝 연구를 설명한다. 제3장에서는 방학 중 학업성취도 분석을 위한 데이터 마이닝 접근법을 의사결정트리, 연관 규칙의 순으로 제안한다. 제4장에서는 설문조사 결과를 기반으로 생성 및 추출한 의사결정트리와 연관 규칙 결과를 제시하고 그 의미를 해석한다. 마지막으로, 제5장에서 결론을 맺는다.

## 2. 관련 연구

학생들의 학업성취도에 관한 여러 연구가 교육 분야에서 진행되어 왔다[13-15]. 참고문헌 [13]에서는 수학 과목에서 교수 학습 태도에 따른 학업성취도 변화를 분석하였다. 이 연구에서는 설문조사 분석을 통해, 수학 과목의 성취도가 높은 학생은 수학에 대한 학습 습관과 학습 동기가 뚜렷하고 끈기를 가지고 열심히 학습한다는 결과를 도출하였다. 그리고 참고문헌 [14]에서는 MBTI(Myers-Briggs Type Indicator) 성격 유형론을 토대로 학생의 성격 유형에 따른 학업성취도를 분석하였다. 이 연구에서는 성격의 기본 유형과 학업성취도 사이에는 밀접한 관계가 있다는 결과를 도출하였다. 다음으로, 참고문헌 [15]에서는 ICT(Information Commu-

nication Technology)를 활용한 학습에서의 학습자 특성에 따른 학업성취도를 분석하였다. 이 연구에서는 ICT 교수-학습 환경에서 정서적 요인, 사회적 요인 등이 학업성취도에 미치는 영향을 분석하였다.

그러나, 지금까지의 학업성취도 관련 연구들은 주로 정규 학기의 학업성취도 분석에 연구의 초점이 맞추어져 있었다. 즉, 본 연구에서 다루는 방학 기간 중의 학업성취도 변화 분석은 현재까지 연구된 바가 없다. 경험적으로 볼 때, 방학 중 학습방법이나 생활습관은 학업성취도 향상에 큰 영향을 미친다. 따라서, 본 논문은 방학 전후의 학업성취도 변화 요인을 체계적으로 분석한다는 점에서 기존 연구와 차별성을 가진다.

교육 분야에 데이터 마이닝을 적용하려는 연구 또한 활발히 진행되어 왔다[8,9,16]. 이 중 학업성취도와 관련하여 마이닝을 적용한 연구는 참고문헌 [8]에서 수행되었다. 참고문헌 [8]의 연구에서는 고등학생을 대상으로 1학기 성적과 학원수강 여부, 수업시간의 학습태도를 이용해 2학기 성적을 예측하는 의사결정트리를 생성하였다. 그 결과, “상위권 학생들의 경우, 다음 학기 성적에 수업태도나 학원수강이 영향을 미치지 않지만, 중하위권 학생들의 경우는 성적에 영향을 미친다.”는 지식을 찾아내기도 하였다. 그러나, 이 연구에서는 학원수강 여부와 수업시간의 학습태도라는 두 가지 단편적 정보만 사용했기 때문에, 학업성취도에 영향을 줄 수 있는 과외나 인터넷 학습 등의 여러 요소들이 고려되지 않은 문제점이 있다. 또한, 학기와 학기 사이의 방학 중 학습방법이나 생활습관이 전혀 고려되지 않은 문제점이 있다. 반면에, 본 논문에서는 학원수강 이외에 과외, 교육 방송 시청 등 여러 요소들을 고려하고, 정규 학기가 아닌 방학 기간을 학업성취도가 변화하는 주요 기간으로 삼는다. 또한, 데이터 마이닝 방법도 의사결정트리 이외에 연관 규칙을 함께 사용한다.

학업성취도 외에 데이터 마이닝을 교육 분야에 적용한 연구는 다음과 같다. 첫째, 참고문헌 [16]에서는 연관 규칙을 사용한 학생 정보 마이닝 시스템을 구현하였다. 다음으로, 참고문헌 [9]에서는 실업계 고등학생의 진로 분석에 의사결정트리를 이용하였다. 그러나 이들 연구는 데이터 마이닝 기법을 사용했다는 공통적인 특징이 있으나, 본 연구에서 초점을 맞춘 학업성취도와는 특별한 관련이 없다.

### 3. 방학 중 학습방법과 학업성취도의 관계 분석

#### 3.1 마이닝 항목 도출

방학 중 학습방법과 생활습관이 학업성취도 변화에 미치는 영향을 분석하기 위하여 본 절에서는 학업성취도에 영향을 미칠 수 있는 요소들을 도출한다. 먼저, 방

학 전후의 학업성취도 변화를 판단하기 위한 요소로 학생 개개인의 방학 전 성적과 개학 후 성적을 사용한다. 그리고 과외, 학원수강 등과 같이 학업성취도에 영향을 줄 가능성이 높은 요소들을 직접관련항목으로 선정한다. 또한, 일기 작성, TV 시청 시간 등과 같이 학업성취도에 간접적인 영향을 줄 수 있는 요소들을 간접관련항목으로 선정한다. 마지막으로, 이와 같이 선정한 직접 및 간접관련항목들을 대상으로 설문 데이터를 작성한다. 직접관련항목은 현재까지 시도된 바가 없어 교육관련 교수 및 현직 교사들의 비공식적 설문 조사를 바탕으로 하였다.

학업성취도 변화를 측정하는 요소로는 **상승, 유지, 하락**을 사용한다. 즉, 방학 전 성적과 개학 후 성적을 비교하여 성적이 높아진 학생을 상승으로 분류한다. 예를 들어, 방학 전 성적이 ‘미’였던 학생이 개학 후 ‘우’로 높아졌으면, 그 학생은 방학을 전후로 학업성취도가 상승한 것으로 본다. 같은 방법으로, 방학 전 성적과 개학 후 성적을 비교하여, 성적이 동일한 학생을 유지로, 성적이 낮아진 학생을 하락으로 설정한다.

직접관련항목은 직관적으로 생각했을 때 학업성취도에 영향을 줄 가능성이 높은 요소들로 선정한다. 즉, 학생들이 학업성취도 상승을 위하여 의도적으로 실시하는 학습방법을 직접관련항목이라 하며, 이들 직접관련항목은 다음과 같다.

- 학원수강
- 집단 또는 개인 대상의 사설과의
- 교육방송 시청
- 학습지에 의한 학습
- 인터넷 학습사이트 이용

상기 언급된 직접관련항목에 대해서는 방학 전후의 성적과 함께 국어, 영어, 수학 등 각 과목별로 데이터를 수집한다.

간접관련항목은 학생들이 방학 중 수행하는 일반적인 활동들 중에서 학업성취도와 간접적 관련이 있는 요소들을 사용한다. 즉, 학업성취도 상승이 직접적인 목적이 아닌 일상적인 생활습관이나 취미 등을 간접관련항목으로 사용하며, 이러한 간접관련항목은 다음과 같다.

- 생활습관 관련항목
  - 컴퓨터 사용시간, TV 시청 시간, 운동 시간
  - 봉사활동 여부, 이성 친구 유무, 부모님과의 대화 여부
  - 아침식사 여부
- 특정 과목의 학업성취 관련항목
  - 일기 작성 여부, 편지나 메일 작성 여부
  - 홈페이지(미니홈피나 블로그 포함) 운영 여부, 독서 여부

- 어학연구 여부 등을 선정
- 기타 항목
  - 목표 학교
  - 여행 여부
  - 다른 국가 거주 여부

**3.2 의사결정트리 분석**

방학 중 학습방법이 학업성취도 변화에 미치는 영향을 분석하기 위해서 분류(classification)의 대표적 기법인 의사결정트리(decision tree)를 사용한다. 의사결정트리는 수집된 과거의 데이터 레코드들을 분석하여 이들 사이에 존재하는 패턴이나 분류 모델을 트리 구조로 대표화한 후, 이를 사용하여 분류와 예측을 수행하는 방법이다[1,8-10].

의사결정트리는 방학 중 어떠한 학습방법들이 학업성취도 변화에 영향을 미치는지 파악하기 위해 사용한다. 즉, “방학 중 어떤 학습방법을 사용하면 성적이 향상되는가?”와 같은 형태의 분석을 위해 의사결정트리를 사용한다. 본 연구에서는 제3.1절에서 도출된 요소들 중 직접관련항목을 사용해 의사결정트리를 생성한다. 직접관련항목만을 사용하는 이유는 간접관련항목까지 모두 고려하면 트리가 너무 복잡해져 분석이 오히려 어렵기 때문이다. 의사결정트리의 목표변수는 방학 전후의 성적 변화, 즉, 상승, 유지, 하락으로 두고 입력변수를 변화시키면서 트리를 생성한다. 그리고 각 과목별로 두 개의 의사결정트리를 생성한다. 첫 번째 트리는 학원수강, 과외, 교육방송 시청, 학습지 학습, 인터넷 학습사이트 이용을 입력변수로 설정하여 방학 중 학습방법들이 학업성취도에 미치는 영향을 분석한다. 이와 같이 방학 전 성적에 관계없이 모든 학생을 대상으로 생성한 트리를 **전체 트리**라 한다. 두 번째 트리는 첫 번째 트리의 입력변수에 지난 학기 성적을 추가하여 성적 별로 학생들을 구분한 후, 방학 전 성적에 따른 효과적인 학습방법을 분석한다. 이와 같이 학생들을 우선 성적 별로 구분한 후 생성한 의사결정트리를 **세분 트리**라 한다. 즉, 전체 트리는 방학 전 성적에 관계없이 모든 학생들을 대상으로 생성한 트리이고, 세분 트리는 방학 전 성적으로 학생들을 그룹(수, 우, 미, 양, 가)으로 구분한 후 구분된 그룹에 대해 생성한 트리이다.

본 연구에서는 의미있는 의사결정트리를 생성하기 위하여 E-Miner가 생성한 트리에 두 가지 전지 규칙(pruning rule)을 적용한다. 다음 두 가지의 전지 규칙 사용 목적은 E-Miner에 의해 생성된 트리에서 의미없는 노드들을 제거하기 위해서이다.

**전지 규칙 1:** 단말 노드에 속하는 개체 수는 전체 입력 데이터 수의 5%이상이어야 한다. □

전지 규칙 1을 적용하는 이유는 단말 노드의 개체 수

가 5% 미만으로 너무 작은 경우, 전체 학생에서 해당 학생들이 차지하는 비율이 너무 작아 규칙으로서 의미가 없기 때문이다. 즉, 5% 미만에 해당하는 일부 학생에게만 적용되는 규칙은 일반적인 규칙이라 보지 않는다.

**전지 규칙 2:** 루트 노드와 단말 노드의 성적 변화율 차이는 5%이상이어야 한다. □

전지 규칙 2를 적용하는 이유는 루트에서 단말까지의 성적 변화율 차이가 5% 미만으로 너무 작은 경우, 성적의 변화가 확실하다고 판단하기 어렵기 때문이다. 예를 들어, 성적이 ‘수’였던 학생의 5% 미만이 개학 후 성적이 ‘우’로 바뀐 것을 의미있는 규칙으로 보지 않는다.

**3.3 연관 규칙 분석**

학생들의 방학 중 생활습관 사이에 특이한 연관 관계가 있는지를 분석하기 위해서, 본 논문에서는 대표적 데이터 마이닝 기법인 연관 규칙[1,11,12]을 사용한다. 연관 규칙의 추출은 우선 빈발항목집합(frequent itemset)을 찾고, 이들 빈발항목집합에서 연관 규칙을 찾는 과정으로 이루어진다[17]. 발견된 연관 규칙을 평가하는 척도로는 지지도(support degree)와 신뢰도(confidence degree)가 있다. 지지도란 해당 규칙이 얼마나 자주 나타나는가를 의미하며, 신뢰도는 해당 규칙이 어느 정도의 의미를 가지느냐를 나타내는 척도로 사용된다[18].

연관 규칙 추출에는 제3.1절에서 설명한 직접관련항목과 간접관련항목 모두를 사용한다. 의사결정트리와 달리, 직접관련항목과 간접관련항목 모두를 사용하는 이유는 연관 규칙의 경우 항목 수가 많더라도 분석 및 해석이 복잡하지 않기 때문이다. 다시 말해서 연관 규칙 자체가 수천에서 수만 개의 항목들 중에서 궁극적으로는  $A \Rightarrow B$  형태의 간단한 규칙을 추출하므로, 본 연구에서 다루는 수십 개 항목을 분석하는 데 아무런 문제가 없기 때문이다.

하지만 연관 규칙을 추출하기 위해서는 설문 결과를 정제하는 과정이 필요하다. 그 이유는 연관 규칙에서는 특정 항목이 트랜잭션에 나타나지 않거나 나타나지 않는지만을 고려하는 반면에, 설문항목에는 어떤 작업이나 활동을 몇 시간 했는지 등의 수량 정보를 포함하기 때문이다. 즉, 연관 규칙의 추출을 위해서는 표 1의 트랜잭션 예제와 같이 어떤 활동을 “몇 시간을 했는지”가 아닌 그 활동을 “했는지, 하지 않았는지”의 불리언 정보만이 사용된다. 표 1의 학생 3을 보면, “TV시청을 했다”는 정보는 사용하나 “몇 시간 동안 시청했는지”는 사용하지 않는다. 결국, 분석 대상 항목을 이러한 불리언형으로 나타내기 위해서는 일부 항목에 대한 변환 작업이 필요하다. 즉, 간접관련항목 중 “과외 여부”와 같이 답변이 ‘예/아니오’인 경우 이미 불리언형이므로, 별도의 변환이

표 1 연관 규칙 추출을 위한 입력 트랜잭션 예제

학생	방향 중 활동
1	국어 과외, 수학 학원수강, TV시청, 여행
2	여행, 영어 교육방송 시청, 운동
3	수학 과외, TV시청, 운동, 여행
4	국어 학원수강, 영어 과외, 운동, 여행
5	국어 과외, 영어 인터넷 학습, 운동, 여행
⋮	⋮

표 2 연속형 항목의 불리언형 변환 표기법

Symbols	Definitions
$Q_i$	$i$ 번째 설문항목 ( $1 \leq i \leq m$ , $m$ 은 설문항목 개수)
$R_j$	$j$ 번째 응답자 ( $1 \leq j \leq n$ , $n$ 은 응답자 수)
$ans(Q_i, R_j)$	$i$ 번째 설문항목 $Q_i$ 에 대한 $j$ 번째 응답자 $R_j$ 의 응답
$avg(Q_i)$	연속형 설문항목 $Q_i$ 에 대한 응답 데이터의 평균

필요하지 않지만, “TV시청 시간”과 같이 답변이 “1시간/1시간~2시간/2시간~3시간”으로 주어지는 연속형 항목의 경우 불리언형으로의 변환이 필요하다.

연속형 항목을 불리언형으로 변환하기 위해서는 평균을 기준으로 구분하는 방법을 사용한다. 즉, 연속형인 경우 해당 항목 값의 평균을 구한 후, 주어진 값이 평균 값보다 작으면 False(해당 항목이 트랜잭션에 나타나지 않음)로, 크면 True(해당 항목이 트랜잭션에 나타남)로 변환하는 방법을 사용한다. 이를 좀 더 정확히 표현하기 위해, 표 2의 표기법과 식 (1)을 사용한다. 식 (1)에서  $Trans(Q_i, R_j)$  는  $i$  번째 설문항목  $Q_i$  의  $j$  번째 응답자  $R_j$  가 답한 응답 데이터를 불리언형으로 변경하는 함수이

다. 식 (1)을 보면 응답 값이 평균보다 낮으면 False로, 크면 True로 변환함을 알 수 있다.

$$Trans(Q_i, R_j) = \begin{cases} true, & \text{if } ans(Q_i, R_j) \geq avg(Q_i) \\ false, & \text{if } ans(Q_i, R_j) < avg(Q_i) \end{cases} \quad (1)$$

### 4. 실험 및 분석

#### 4.1 데이터 수집

학생들의 방향 중 생활습관과 활동이 학업성취도에 미치는 영향을 분석하기 위해 설문조사를 통해 현실의 구체적 데이터를 수집하였다. 설문조사서는 제3.1절에서 도출한 항목들을 기반으로 작성하였다.

설문조사는 강원도 춘천지역 중등학교 2학년 및 3학년 학생을 대상으로 하였다.<sup>1)</sup> 모집단의 학생 수는 10,790명이었으며 신뢰도 90%, 표본오차범위 5%인 경우의 표본 집단수는 371명이었다. 본 논문의 설문조사에는 500여명 이상의 학생들이 참여하였으므로, 신뢰도 90% 및 표본오차범위 5% 수준을 만족한다.

설문조사는 그림 1과 같이 온라인 인터넷 사이트에서 진행하였다. 온라인 설문조사 전문업체인 리서치중앙 [19]에 설문을 의뢰하였으며, 사이트는 여자중학교를 중심으로 홍보하였다. 데이터 수집을 위한 설문조사 기간은 2005년 5월 17일부터 2005년 5월 31일까지의 15일로 하였다.

설문에 응답한 학생은 총 536명으로, 이들의 성별 및 학년 분포는 표 3과 같다. 응답한 내용 중 질문의 일부

1) 강원도 춘천지역은 2005년 현재 고교 비평준화 지역으로 평준화와 관련이 적은 중학교 학생을 설문 대상으로 하였다.

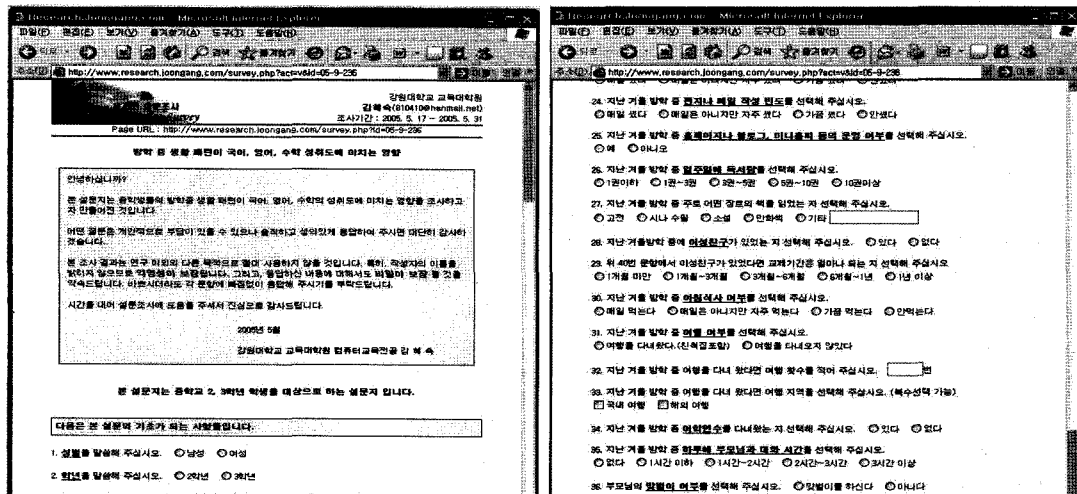


그림 1 온라인 설문조사 화면 예제

표 3 응답 데이터의 전체 통계 요약

	2학년		3학년		총계	
	응답자수(명)	백분율(%)	응답자수(명)	백분율(%)	응답자수(명)	백분율(%)
남학생	66	22	8	8	74	14
여학생	232	78	230	92	462	86
총 계	298	100	238	100	536	100

항목에만 대답한 불완전 응답 데이터를 제거하였고, 남학생과 여학생 비율이 크게 차이가 나서 남학생 응답 데이터 또한 데이터 구성에서 제외하였다.<sup>2)</sup> 이에 따라, 실제 실험에서는 불완전 데이터와 남학생 데이터를 제외한 여학생 462명의 데이터를 사용하였다.

그림 2는 설문 응답 데이터의 과목별 성적 변화율을 그래프로 나타낸 것이다. 그림을 보면, 국어, 영어, 수학적 과목 모두 약 50%의 학생은 방학 전후의 성적이 유지되었고, 나머지 50%의 학생들은 상승이나 하락 등 성적의 변화가 있었다. 다음으로, 그림 3은 학원수강, 과외, 교육방송 시청, 학습지 학습, 인터넷 학습 사이트 등의 직접관련항목에 대한 응답 비율 통계를 그래프로 표현한 것이다. 그림을 보면, 국어, 영어, 수학 모든 과목에 있어서 타 학습방법에 비해 학원수강을 가장 많이 하고 있음을 알 수 있다. 또한, 수학은 다른 과목에 비해 과외와 교육방송 시청의 비율도 높게 나타났음을 알 수 있다.

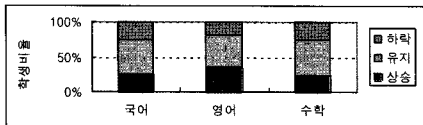


그림 2 방학 전후의 성적 변화 비율

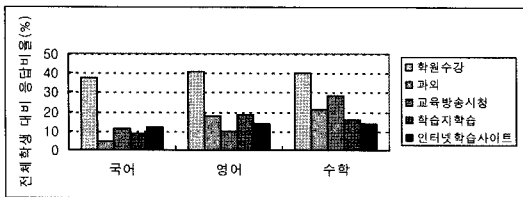


그림 3 직접관련항목 설문 데이터의 응답 비율

간접관련항목에 대한 설문조사 주요 결과를 요약하면 다음과 같다. 먼저, 대부분의 학생들이 인문계 학교를 목표로 하였다. 그리고 50% 이상의 학생이 1시간에서 3시간 정도 컴퓨터를 사용하며, TV를 3시간 이상 시청하는 학생도 50% 이상으로 나타났다. 반면에, 운동 시

간은 1시간 미만인 학생이 67%로, 많은 학생들이 운동을 하지 않는 것으로 나타났다. 또한, 많은 학생들이 봉사활동, 독서, 여행을 하고 일기, 편지, 메일을 작성하는 것으로 나타났다. 그리고 홈페이지 등을 운영하는 학생도 50% 정도가 되는 것으로 파악되었다.

4.2 의사결정트리 생성 결과

본 절에서는 전체 트리와 세분 트리 생성 결과를 제시하고, 그 의미를 해석한다. 제3.2절에서 설명한 바와 같이, 전체 트리는 학원수강, 과외, 교육방송 시청, 학습지 학습, 인터넷 학습 사이트 이용 등의 직접관련항목을 입력 변수로 생성한 의사결정트리이다. 그리고 전체 트리의 입력 변수에 지난 학기 성적을 추가하여 생성한 트리가 세분 트리이다. 실험에서는 각 과목에 따른 학습성취도를 분석하기 위해, 각 과목별로 전체 트리와 세분 트리의 두 가지 의사결정트리를 생성하였다.

국어 과목에 대한 전체 트리 생성 결과는 그림 4와 같다. 그림을 보면, E-Miner에 의해 생성된 트리에서 네 개의 단말 노드가 전지 규칙 1과 전지 규칙 2에 의해 삭제되었음을 알 수 있다. 결국, 그림 5에서는 두 개의 의미있는 단말 노드가 생성되고, 이들 단말 노드를 논리곱(logical AND) 형태의 지식으로 나타내면 표 4와 같다. 다음으로, 그림 5는 학생들을 방학 전 성적<sup>3)</sup>으로

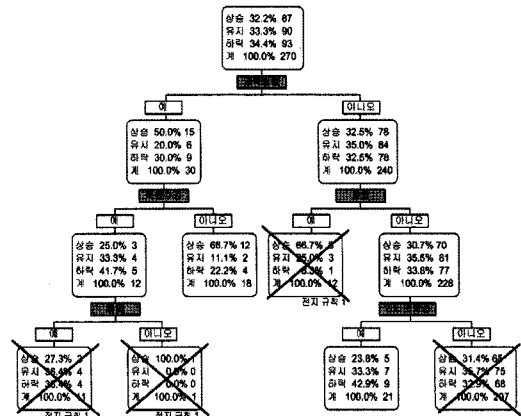


그림 4 국어 과목의 전체 트리

2) 설문조사 사이트가 여자중학교에 중점 소개된 관계로 여학생들의 설문 응답이 많은 반면 남학생들의 응답은 저조하였다.

3) 국어의 경우, '양'과 '가'의 학생 수가 적어 이를 합해 하나의 그룹으로 취급하였다. 반면에 영어와 수학은 학생분포가 고른 관계로 '양'과 '가'를 분리하여 트리를 생성하였다.

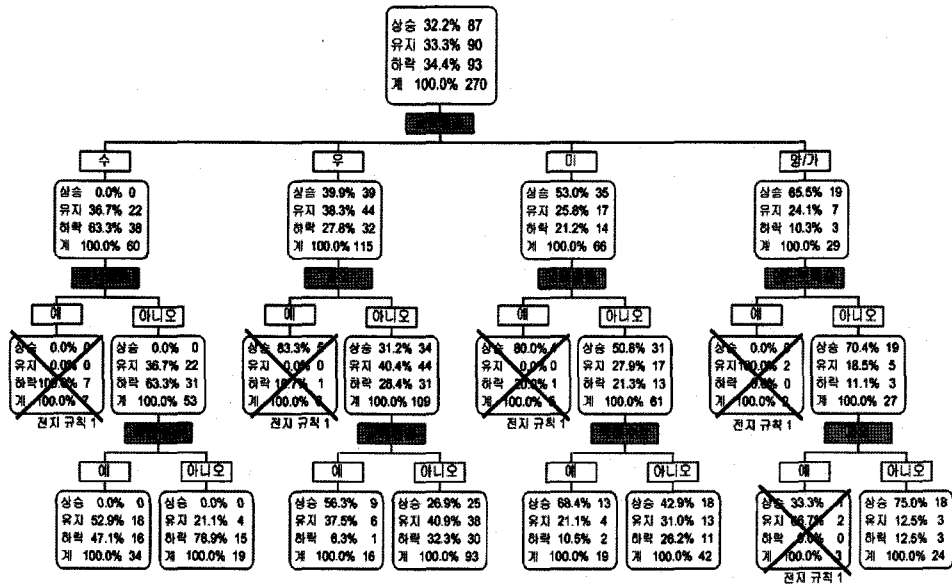


그림 5 국어 과목의 세분 트리

표 4 국어 전체 트리에 대한 논리곱 형태의 지식

분류	조건
상승	(교육방송=예) ∧ (학원수강=아니오)
하락	(교육방송=아니오) ∧ (과외=아니오) ∧ (인터넷학습=예)

표 5 국어 세분 트리에 대한 논리곱 형태의 지식

작년 성적	분류	조건
수	유지	(교육방송=아니오) ∧ (학원수강=예)
	하락	(교육방송=아니오) ∧ (학원수강=아니오)
우	상승	(과외=아니오) ∧ (교육방송=예)
	유지	(과외=아니오) ∧ (교육방송=아니오)
미	상승	(교육방송=아니오) ∧ (학원수강=예)
	유지	(교육방송=아니오) ∧ (학원수강=아니오)
양/가	상승	(인터넷학습=아니오) ∧ (과외=아니오)

먼저 구분한 후 생성한 세분 트리이다. 전체 트리와 마찬가지로 전지 규칙 1과 2를 사용하여 의미 없는 단말 노드를 삭제한 후, 이들 단말 노드를 논리곱 형태의 지식으로 나타내면 표 5와 같다.

국어 과목의 의사결정트리에서 나온 표 4와 표 5의 주요 지식을 해석하면 다음과 같다.

- 전체적으로 보았을 때, 교육방송은 국어 성적의 상승을, 인터넷 학습 사이트 이용은 국어 성적의 하락을 가져오는 것으로 나타났다.
- 방학 전 국어 성적이 '수'인 학생이 학원수강을 하지 않았을 경우 성적은 하락하는 것으로 나타났다.
- 방학 전 국어 성적이 '우'인 학생이 교육방송을 시청할 경우 성적은 상승한 것으로 나타났다.

• 방학 전 국어 성적이 '미'인 학생은 학원수강이 성적 상승에 큰 영향을 주지 않는 것으로 나타났다.

다음으로, 그림 6과 7은 영어 과목에 대해 생성한 전체 트리와 세분 트리를 각각 나타낸다. 국어 과목과 마찬가지로 트리를 생성한 후 두 가지 전지 규칙을 사용하여 의미 없는 단말 노드를 제거하였다. 그리고 표 6과 7은 이들 전체 트리와 세분 트리의 단말 노드들을 대상으로 논리곱 형태의 지식을 표현한 것이다.

표 6 영어 전체 트리에 대한 논리곱 형태의 지식

분류	조건
상승	(과외=아니오) ∧ (학원수강=아니오) ∧ (인터넷학습=아니오)
유지	(과외=예) ∧ (학습지=아니오) ∧ (인터넷학습=아니오)
하락	(과외=예) ∧ (학습지=아니오) ∧ (인터넷학습=예)
	(과외=아니오) ∧ (학원수강=예) ∧ (학습지=예)

표 7 영어 세분 트리에 대한 논리곱 형태의 지식

작년 성적	분류	조건
수	유지	(인터넷학습=아니오) ∧ (과외=아니오)
	상승	(과외=아니오) ∧ (학원수강=예)
우	유지	(과외=예) ∧ (학습지=아니오)
	하락	(교육방송=아니오) ∧ (과외=아니오)
미	상승	(인터넷학습=아니오)
가	상승	(과외=아니오)

영어 과목의 의사결정트리에서 나온 주요 지식을 해석하면 다음과 같다.

- 전체적으로 보았을 때, 과외를 하면 영어 성적이 유지

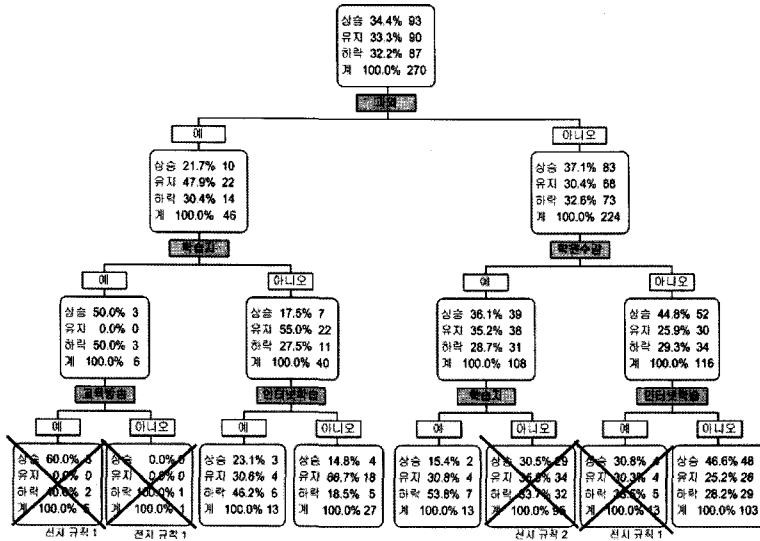


그림 6 영어 과목의 전체 트리

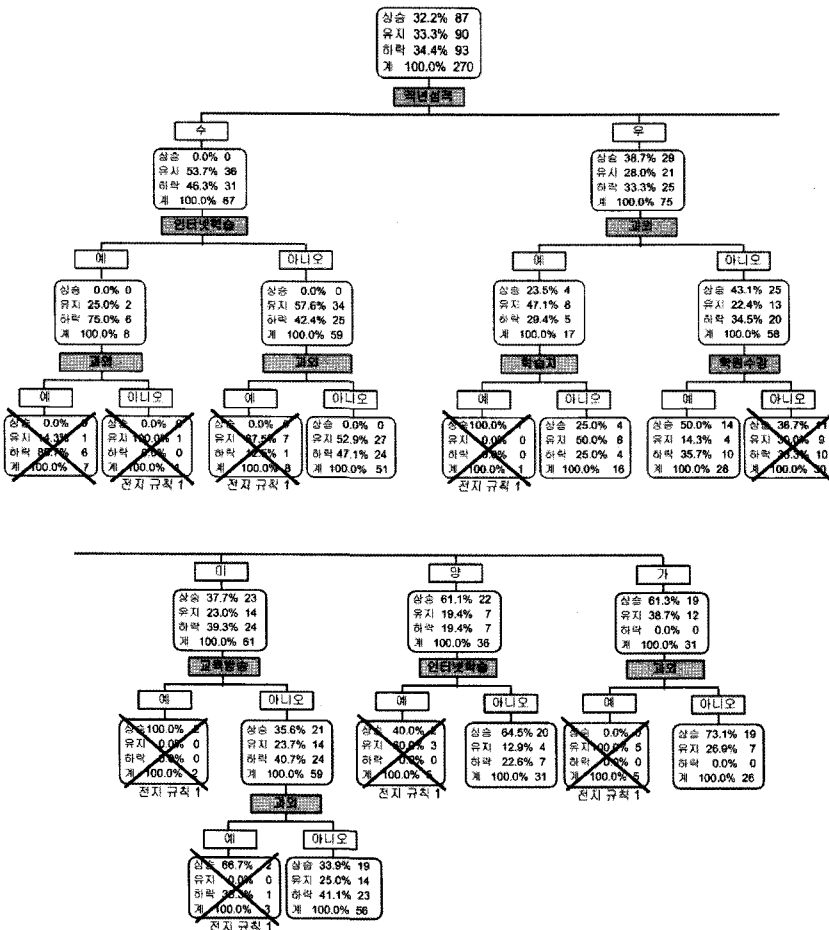


그림 7 영어 과목의 세분 트리



되며, 과외, 학원수강, 인터넷 학습 사이트 이용의 세 가지를 모두 하지 않으면 성적은 오히려 상승하는 것으로 나타났다.

- 방학 전 영어 성적이 '수'인 학생은 인터넷 학습 사이트 이용이나 과외를 하지 않아도 성적이 유지되는 것으로 나타났다.
- 방학 전 영어 성적이 '우'인 학생의 성적 상승에는 학원수강이, 성적 유지에는 과외가 영향을 미치는 것으로 나타났다.
- 방학 전 영어 성적이 '미'인 학생이 교육방송과 과외들 모두를 하지 않을 경우 성적이 하락하는 것으로 나타났다.
- 방학 전 영어 성적이 '양'인 학생이 인터넷 학습 사이트 이용을 하지 않으면 성적은 상승하는 것으로 나타났다.
- 방학 전 영어 성적이 '가'인 학생이 과외를 하면 성적이 상승하는 것으로 나타났다.

마지막으로, 그림 8과 9는 수학 과목에 대해 생성한 전체 트리과 세분 트리를 각각 나타낸다. 마찬가지로, 전지 규칙 1 및 2를 사용하여 의미 없는 단말 노드를 제거하였다. 표 8과 9는 이들 트리를 대상으로 논리곱 형태의 지식을 추출한 것이다.

수학 과목의 의사결정트리에서 나온 주요 지식을 해

석하면 다음과 같다.

- 전체적으로 보았을 때, 학원수강을 하면 수학 성적이 유지되는 것으로 나타났다.
- 방학 전 수학 성적이 '수/우/미'인 학생들의 경우, 학원수강이 성적 유지에 큰 영향을 미치는 것으로 나타났다.
- 방학 전 수학 성적이 '미'인 학생이 과외를 하면 성적이 오히려 하락하는 것으로 나타났다.
- 방학 전 수학 성적이 '양'인 학생은 인터넷 학습 사이트 이용과 과외를 하지 않으면 성적이 상승하는 것으로 나타났다.
- 방학 전 수학 성적이 '가'인 학생은 학습지 학습과 학원수강을 하지 않으면 성적이 상승하는 것으로 나타났다.

지금까지 살펴 본 국어, 영어, 수학 세 과목의 전체 트리에서 추출된 지식을 종합하면 다음과 같은 결과를 얻을 수 있다.

• 학원수강과 인터넷 학습 사이트 이용은 성적 상승에 부정적인 영향을 미치는 것으로 나타났다(중상위권은 예외로서 다음의 세분 트리 결과를 참조한다).

• 교육방송 시청, 학원수강, 과외를 모두 할 경우, 오히려 성적이 하락하는 결과를 보이고 있다.

다음으로, 세 과목의 세분 트리에서 추출된 공통된 내

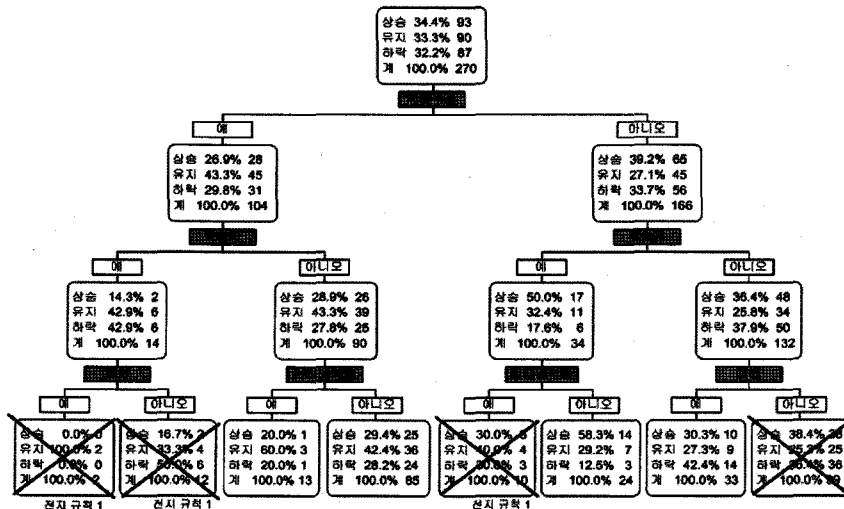


그림 8 수학 과목의 전체 트리

표 8 수학 전체 트리에 대한 논리곱 형태의 지식

분류	조건
상승	(학원수강=아니오) ∧ (학습지=예) ∧ (인터넷학습=아니오)
유지	(학원수강=예) ∧ (학습지=아니오) ∧ (인터넷학습=예)
	(학원수강=예) ∧ (학습지=아니오) ∧ (인터넷학습=아니오)
하락	(학원수강=아니오) ∧ (학습지=아니오) ∧ (과외=예)

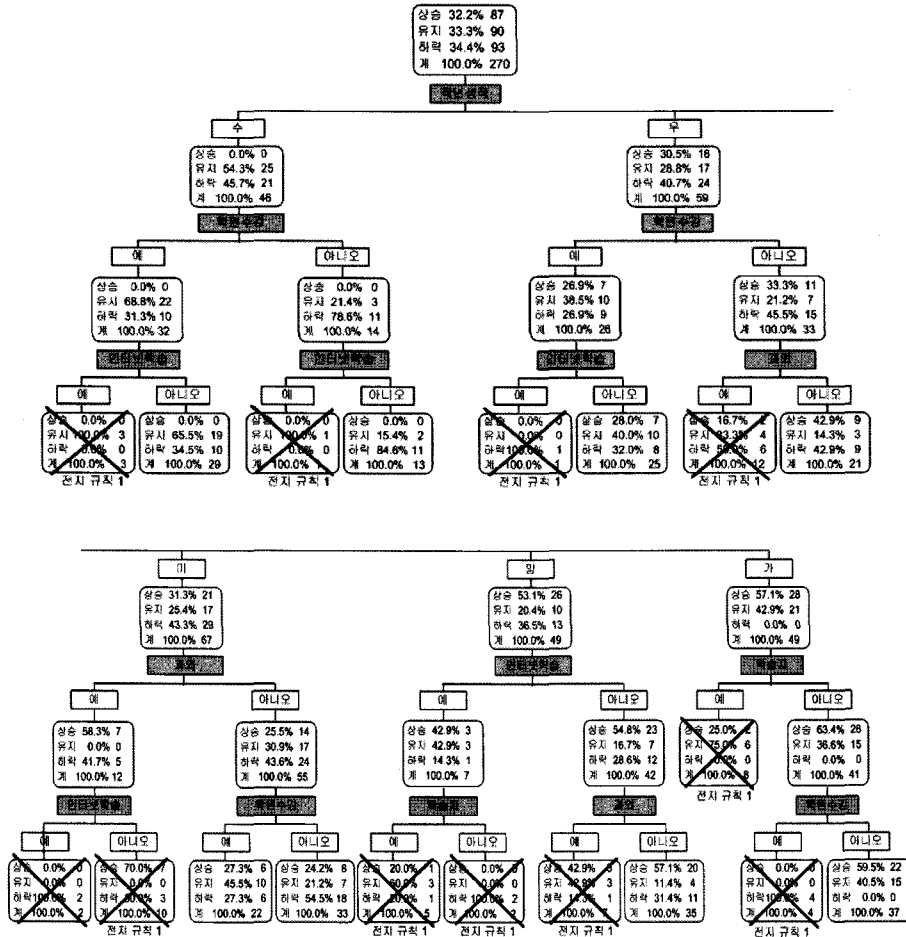


그림 9 수학 과목의 세분 트리

표 9 수학 세분 트리에 대한 논리곱 형태의 지식

작년 성적	분류	조건
수	유지	(학원수강=예) ∧ (인터넷학습=아니오)
	하락	(학원수강=아니오) ∧ (인터넷학습=아니오)
우	상승	(학원수강=아니오) ∧ (과외=아니오)
	유지	(학원수강=예) ∧ (인터넷학습=아니오)
미	유지	(과외=아니오) ∧ (학원수강=예)
	하락	(과외=아니오) ∧ (학원수강=아니오)
양	상승	(인터넷학습=아니오) ∧ (과외=아니오)
	가	상승

용을 요약하면 다음과 같다.

- 방학 전 성적이 '수'일 경우, 학원수강은 성적 유지에 긍정적인 영향을 미치는 반면 인터넷 학습 사이트 이용은 부정적인 영향을 미치는 것으로 나타났다.
- 방학 전 성적이 '우'일 경우, 과외보다는 교육방송이나

학원수강이 성적 향상에 도움이 되는 것으로 나타났다.

- 방학 전 성적이 '미'인 경우, 학원수강은 성적 상승 및 유지에 긍정적인 영향을 미치는 것으로 나타났다.
- 방학 전 성적이 '양'이나 '가'인 경우, 인터넷 학습 사이트 이용이 성적 상승에 부정적인 영향을 미치는 것으로 나타났다.

지금까지의 결과를 간략히 요약하면 다음과 같다. 첫째, 상위권 학생의 경우, 학원수강이 성적 향상에 큰 도움을 준다. 둘째, 인터넷 학습 사이트 이용은 전체 학생에 있어 성적 하락의 원인으로 작용하고 있다. 셋째, 성적 향상에 큰 영향을 줄 것으로 예상한 과외가 실제로는 큰 영향을 미치지 못하며, 다양한 학습방법(교육 방송 시청, 학원수강, 과외 등)을 병행할 경우 오히려 성적이 하락한다. 이와 같은 분석 결과는 학생들의 방학 중 학습 지도를 위한 유용한 자료로 활용할 수 있다고 사료된다.

4.3 연관 규칙 추출 결과

연관 규칙 분석은 신뢰도와 지지도를 달리하면서 실험하였다. 실험 1)에서는 신뢰도를 90%, 지지도를 30%로 설정하고 실험하고, 실험 2)에서는 신뢰도는 90%로 동일하게 하되, 지지도를 45%로 높여 실험하였다. 제3.3절에서 설명한 바와 같이, 직접 및 간접관련항목을 모두 사용하였으며, 연속형 항목을 불리언형으로 변환하여 실험하였다.

그림 10은 실험 1)의 연관 규칙 생성 결과를 나타낸다. 그림을 보면, 추출된 연관 규칙은 총 24개로 나타났다. 추출된 규칙 중에서 의미를 가진다고 생각되는 규칙은 “수학과 영어 학원수강을 하는 학생은 국어도 학원수강을 한다”이다. 그런데, 이러한 연관 규칙은 학원수강을 하는 학생들은 국어, 영어, 수학 세 과목 모두를 수강한다는 지극히 상식적 수준의 지식이라 할 수 있다.

LHS	RHS	Support	Confidence
1.수학	2.영어	0.10	0.33
2.영어	3.국어	0.10	0.33
1.수학 & 2.영어	3.국어	0.10	0.33
1.수학	3.국어	0.10	0.33
2.영어	3.국어	0.10	0.33
1.수학 & 3.국어	2.영어	0.10	0.33
2.영어 & 3.국어	1.수학	0.10	0.33
1.수학 & 2.영어 & 3.국어		0.10	0.33
1.수학	2.영어 & 3.국어	0.10	0.33
2.영어	1.수학 & 3.국어	0.10	0.33
3.국어	1.수학 & 2.영어	0.10	0.33
1.수학 & 2.영어	3.국어	0.10	0.33
2.영어 & 3.국어	1.수학	0.10	0.33
1.수학 & 3.국어	2.영어	0.10	0.33
2.영어 & 1.수학	3.국어	0.10	0.33
3.국어 & 1.수학	2.영어	0.10	0.33
3.국어 & 2.영어	1.수학	0.10	0.33
1.수학 & 2.영어 & 3.국어		0.10	0.33

그림 10 신뢰도 90%, 지지도 30%인 경우의 연관 규칙 결과

다음으로, 그림 11은 지지도를 30%에서 45%로 높인 실험 2)의 결과를 나타낸다. 그림에서와 같이, 실험 2)에서는 총 네 개의 연관 규칙이 추출되었다. 그러나, 이들 규칙 모두는 “여행을 하는 학생은 아침 식사를 한다”와 같이 결론부가 “아침 식사를 한다”로 나타난다. 이러한 결과가 나온 이유는 “아침 식사를 한다”고 응답한 학생의 비율이 전체의 86%로 매우 높기 때문이다. 즉, 실험 2)에서 추출된 규칙은 아침 식사 항목이 빈발하게 나타났기 때문이며 특별한 의미를 부여하기는 어렵다.

실험 1)과 실험 2)의 결과를 종합해 보면, 방학 중 생활습관 사이에서는 특이한 연관 규칙이 발견되지 않았다. 결국 연관 규칙의 결과를 종합해 볼 때, 방학 중 생

LHS	RHS	Support	Confidence
1.수학	2.영어	0.10	0.33
2.영어	3.국어	0.10	0.33
1.수학 & 2.영어	3.국어	0.10	0.33
1.수학	3.국어	0.10	0.33
2.영어	3.국어	0.10	0.33
1.수학 & 3.국어	2.영어	0.10	0.33
2.영어 & 3.국어	1.수학	0.10	0.33
1.수학 & 2.영어 & 3.국어		0.10	0.33
1.수학	2.영어 & 3.국어	0.10	0.33
2.영어	1.수학 & 3.국어	0.10	0.33
3.국어	1.수학 & 2.영어	0.10	0.33
1.수학 & 2.영어	3.국어	0.10	0.33
2.영어 & 3.국어	1.수학	0.10	0.33
1.수학 & 3.국어	2.영어	0.10	0.33
2.영어 & 1.수학	3.국어	0.10	0.33
3.국어 & 1.수학	2.영어	0.10	0.33
3.국어 & 2.영어	1.수학	0.10	0.33
1.수학 & 2.영어 & 3.국어		0.10	0.33

그림 11 신뢰도 90%, 지지도 45%인 경우의 연관 규칙 결과

활습관이 학업성취도에는 별다른 영향을 미치지 않는 것으로 해석할 수 있다.

5. 결론

본 연구에서는 방학 중 학습방법 및 생활습관과 학업성취도의 관계를 데이터 마이닝 기법을 사용하여 분석하였다. 현재까지 진행된 학업성취도 관련 연구에서는 학생의 학업성취도 향상을 위한 방학 중 생활 지도에 관한 구체적인 접근법이 없었다. 본 연구에서는 데이터 마이닝 기법을 사용하여 다음과 같은 구체적 분석 방법을 제안하였다. 우선, 학업성취도에 영향을 미치는 방학 중 학습방법과 생활습관에 대한 다양한 요소들을 도출하였다. 그리고 이들 요소와 학업성취도 변화의 관계 분석을 위해서 데이터 마이닝 기법들 중 의사결정트리와 연관 규칙을 사용하였다. 또한, 현실의 구체적 데이터 수집을 위해 중학교 학생들을 대상으로 방학 전후 성적과 방학 중 활동에 대한 설문조사를 수행하였다. 다음으로, 의사결정트리를 생성하여 방학 중 효과적인 학습방법을 분석하였다. 마지막으로, 방학 중 학습방법과 생활습관들을 대상으로 연관 규칙을 추출하고 그 의미를 분석하였다.

방학 중 효과적인 학습방법 분석을 위해 생성한 의사결정트리에서 얻은 지식을 요약하면 다음과 같다. 첫째, 학원수강은 상위권 학생들의 성적 향상이 도움을 주는 것으로 나타났다. 둘째, 인터넷 학습사이트 이용은 일반적으로 성적을 하락시키는 것으로 나타났다. 셋째, 성적 변화에 큰 영향을 미칠 것으로 예상했던 과외는 실제로 큰 영향을 미치지 못하는 것으로 나타났다. 넷째, 다양한 학습방법(과외, 학원수강, 교육 방송 시청)을 병행할 경우 오히려 성적이 하락하는 것으로 나타났다. 다음으로, 방학 중 학습방법 및 생활습관 사이의 연관 규칙을 추출하였는데, 규칙의 해석 결과 특별한 의미를 부여하기는 어려운 것으로 파악되었다.

본 논문에서는 설문조사와 데이터 마이닝 기법을 사용하여 학업성취도에 영향을 미치는 방학 중의 학습방법과 생활습관을 분석하였다. 이와 같은 연구 결과는 추후 학생들의 방학 중 학습방법 및 활동 분석에 매우 유용한 모형을 제시한다고 사료된다.

참고 문헌

[1] Chen, M.-S., Han, J., and Yu, P. S., "Data Mining: An Overview from a Database Perspective," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 8, No. 6, pp. 866-883, Dec. 1996.

[2] Moon, Y.-S., Whang, K.-Y., and Han, W.-S.,

"General Match: A Subsequence Matching Method in Time-Series Databases Based on Generalized Windows," In *Proc. Int'l Conf. on Management of Data*, ACM SIGMOD, Madison, Wisconsin, pp. 382-393, June 2002.

[3] <http://www.sas.com/korea>.  
 [4] <http://www-306.ibm.com/software/data/iminer>.  
 [5] <http://www.dbminer.com>.  
 [6] Hwang, J.-J., Whang, K.-Y., Moon, Y.-S., and Lee, B.-S., "A Top-down Approach for Density-Based Clustering Using Multidimensional Indexes," *Journal of Systems and Software*, Vol. 73, Issue 1, pp. 196-180, Sept. 2004.  
 [7] 김신영, 백중근, 채선희, 국가교육과정에 근거한 평가 기준 및 도구 개발, 한국교육과정평가원, 1998.  
 [8] 배재호, 데이터 마이닝을 이용한 학업성취도 분석, 경희대학교 교육대학원 석사학위논문, 2001.  
 [9] 김재은, 데이터 마이닝을 이용한 실업계 고등학교 학생의 진로분석, 경기대학교 교육대학원 석사학위 논문, 2002.  
 [10] Agrawal, R., et al., "An Interval Classifier for Database Mining Applications," In *Proc. the 18th Int'l Conf. on Very Large Data Bases*, Vancouver, Canada, pp. 560-573, Aug. 1992.  
 [11] Agrawal, R., Imielinski, T., and Swami, A. N., "Mining Association Rules between Sets of Items in Large Databases," In *Proc. Int'l Conf. on Management of Data*, ACM SIGMOD, Washington, DC, pp. 207-216, May 1993.  
 [12] Agrawal, R. and Srikant, R., "Fast Algorithms for Mining Association Rules in Large Databases," In *Proc. the 20th Int'l Conf. on Very Large Data Bases*, Santiago, Chile, pp. 487-499, Sept. 1994.  
 [13] 김동화, 조영재, "수학 교수, 학습 태도에 따른 학업성취도의 수준에 관한 분석", *교육이론과 실천*, 제14권, 제1호, pp. 115-132, 2004.  
 [14] 김병식, 김민정, "청소년의 성격유형이 학업성취도에 미치는 영향에 관한 연구", *한국인간관계학보*, 제7권, 제1호, pp. 1-20, 2002.  
 [15] 김정경, "정보통신기술(ICT)을 활용한 교수-학습에서 학습자 특성이 학업성취도에 미치는 영향", *한국컴퓨터교육학회 논문지*, 제7권, 제2호, pp. 47-56, 2004.  
 [16] 공현성, 김병, "학생정보마이닝 시스템의 설계 및 구현", *한국컴퓨터교육학회 논문지*, 제6권, 제1호, pp. 58-63, 2003.  
 [17] 장남식, 홍성환, 장재호, 데이터 마이닝, 대청, 1999.  
 [18] 박종수, 유원경, 홍기형, "연관 규칙 탐사와 그 응용", *한국정보과학회논문지*, 제16권, 제9호, pp. 37-38, 1998.  
 [19] <http://www.research.joongang.com>.



김혜숙

2003년 2월 강원대학교 컴퓨터정보통신공학부 학사. 2006년 2월 강원대학교 교육대학원 컴퓨터교육전공 석사. 2006년 3월~현재 강원대학교 컴퓨터과학전공 박사과정. 2005년 3월~현재 한국과학기술원 첨단정보기술연구소 연구원. 관심분야는 컴퓨터 교육, 데이터 마이닝, 데이터 웨어하우스



문양세

1991년 2월 한국과학기술원 과학기술대학 전산학과 학사. 1993년 2월 한국과학기술원 전산학과 석사. 2001년 8월 한국과학기술원 전자전산학과 전산학전공 박사. 1993년 2월~1997년 2월 현대전자산업(주) 통신사업본부 주임연구원. 2001년 9월~2002년 2월 ㈜현대시스콤 호처리개발실 선임연구원. 2002년 2월~2005년 2월 ㈜인프라벨리 기술연구소 기술위원(이사). 2005년 3월~현재 한국과학기술원 첨단정보기술연구소 연구원. 2005년 3월~현재 강원대학교 컴퓨터과학과 조교수. 관심분야는 Data Mining, Knowledge Discovery, Stream Data, Storage System, Database Applications, Mobile/ Wireless Communication Services & Systems



김진호

1982년 2월 경북대학교 전자공학과 학사. 1985년 2월 한국과학기술원 전산학과 석사. 1990년 2월 한국과학기술원 전산학과 박사. 1995년 8월~1996년 7월 미국 미시간 대학교 객원 교수. 2003년 2월~2004년 2월 미국 Drexel University 객원 교수. 1999년 3월~현재 한국과학기술원 첨단정보기술연구소 연구원. 1990년 8월~현재 강원대학교 컴퓨터과학과 교수. 관심분야는 Data warehouse, OLAP, Data Mining, Real-time/Embedded Database, Main-memory database, Data Modeling, Web Database Technology



노웅기

1991년 2월 한국과학기술원 전산학과 학사. 1993년 2월 한국과학기술원 전산학과 석사(멀티미디어 전공). 2001년 2월 한국과학기술원 전산학과 박사(데이터 마이닝 전공). 2001년 2월~2003년 9월 ㈜티맥스 소프트웨어 연구원(미들웨어 개발). 2003년 10월~2005년 3월 ㈜티맥스데이터 수석연구원(DBMS 개발). 2005년 4월~2006년 5월 한국과학기술원 전산학과 초빙교수. 2006년 6월~현재 미국 University of Minnesota 방문연구원. 관심분야는 데이터 마이닝/데이터 웨어하우스, 정보 검색, 멀티미디어 데이터베이스, 멀티미디어 내용기반 검색, 데이터베이스 시스템, 임베디드 데이터베이스