

연관 규칙을 이용한 방학 중의 생활 습관 분석¹⁾

김혜숙^o 문양세 노희영 김진호

강원대학교 컴퓨터학과

(hskim, ysmoon, rohhy, jhkim)@kangwon.ac.kr

Association Rule-based Analysis to Living Style during Vacation

Hea-Suk Kim^o, Yang-Sae Moon, Hi-Young Roh, and Jinho Kim

Dept. of Computer Science, Kangwon National University

요 약

본 논문에서는 학생들의 방학 중 생활 습관을 분석하기 위해 연관규칙을 사용한 접근법을 제안한다. 이를 위해, 우선 학생들이 할 수 있는 방학 중의 생활 습관에 대한 다양한 요소를 도출한다. 다음으로, 연관규칙을 적용하기 위한 데이터 변환 및 분석 방법을 제안한다. 마지막으로, 설문조사를 통해 수집한 현실의 구체적 데이터에서 연관규칙을 추출한다. 그 결과, 학원수강을 하는 학생들은 국어, 영어, 수학 세과목 모두를 수강한다는 일반적인 규칙을 얻을 수 있었다.

1. 서 론

최근 정보화 시대가 열리면서 정보와 지식이 기업 경영의 경쟁력을 결정하는 중요한 요소가 되었다. 또한 생산자 위주의 시장 환경에서 소비자의 주도의 시장 환경으로 변화하면서 기업이 생존하고 발전하기 위해서는 소비자의 동향이나 경쟁사의 경영 전략 등을 지속적이고 신속하게 분석 할 수 있는 능력이 필요하게 되었다. 이에 맞물려 정보 처리 속도와 정보 저장 기술이 빠른 속도로 발전되면서 데이터 분석 기법들 또한 발전하게 되었다. 결국 기업들은 컴퓨터를 사용하여 방대한 양의 데이터 속에서 쉽게 드러나지 않는 유용한 정보나 지식을 찾아내는 방법이 필요하게 되었다[1].

이러한 정보 및 지식 발견에 있어서 대표적인 데이터 분석 기법 중 하나가 데이터 마이닝이다. 데이터 마이닝은 정보기술의 발달과 비즈니스 요구에 의해 등장하게 되었다고 볼 수 있다. 1980년대 이후 급속한 성장을 이룬 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]. 그 결과, SAS Enterprise Miner[2], IBM Intelligent Miner[3], DBMiner[4] 등의 데이터 마이닝 제품이 상용화되었으며, 금융, 통신, 경영 등의 많은 분야에서 데이터 마이닝이 활용되고 있다.

본 연구는 데이터 마이닝 기법중 연관 규칙을 통한 학생들의 방학 중 생활 습관 분석에 초점을 둔다. 최근 까지 교육 분야에서 진행되어 온 학생생활 관련 연구들은 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 생활을 체계적으로 분석한다.

학생들의 방학 중 생활 습관을 분석하기 위하여 본 논문에서는 연관규칙을 적용한다. 이를 위해, 다음과 같은 세 단계로 연구 및 분석을 수행한다. 첫째, 학생들이 방학 중 할 수 있는 생활 습관들의 항목을 도출하고, 이를 항목을 기반으로 데이터 분석 방법을 구체화한다. 둘째, 방학 중 생활 습관을 분석하기 위해 1) 과거성적, 2) 방학 중 생활 습관, 3) 현재성적으로 구성되는 구체적 데이터를 설문조사를 통해 수집한다. 셋째, 수집한 데이터를 대상으로, 데이터 변환 및 정제를 거쳐, 상용 데이터 마이닝 도구인 SAS社의 Enterprise Miner(간략히 *E-Miner*라 한다)를 사용하여 의미있는 규칙들을 추출하고 해석한다.

본 논문의 구성은 다음과 같다. 제2장의 관련연구에서는 교육 분야의 기존 연구와 교육 분야의 데이터 마이닝 연구를 설명한다. 제3장에서는 방학 중 생활 습관 분석을 위해 연관 규칙을 제안한다. 제4장에서는 설문조사 결과를 기반으로 추출한 연관 규칙 결과를 제시하고 그 의미를 해석한다. 마지막으로, 제5장에서 결론을 맺는다.

2. 관련 연구

교육 분야에서는 주로 학생들의 학업성취도에 대한 여러 연구들이 진행되었다[5-8]. 참고문헌 [5]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 [6]에서는 MBTI(Myers-Briggs Type Indicator) 성격 유형론을 토대로 학생의 성격 유형에 따른 학업성취도를 분석하였다. 이 연구에서는 성격의 기본 유형과 학업성취도 사이에는 밀접한

¹⁾ 2005년도 강원대학교 학술연구조성비로 연구하였음.

관계가 있다는 결과를 도출하였다. 다음으로, 참고문헌 [7]에서는 교사의 효능감과 학생의 학업적 자기 효능감이 학업성취도에 미치는 영향을 분석하였다. 통계 분석 기법인 2수준 다중 통계 모형을 적용하여 교사의 효능감은 학생의 학업성취도에 큰 영향을 미치지 못하지만 학생의 학업적 자기 효능감은 학업성취도에 큰 영향을 미친다는 결과를 얻었다. 또한, 참고문헌 [8]에서는 식사 패턴이 학업성취도에 미치는 영향을 설문 데이터를 바탕으로 분석하여, 식생활과 학업성취도 사이에는 일정한 관계가 있다는 결과를 얻었다. 그러나, 지금까지의 학업성취도 관련 연구들은 주로 정규 학기의 학업성취도 분석에 연구의 초점이 맞추어져 있었다. 즉, 본 연구에서 다루는 방학 기간 동안의 학생 생활 습관 분석은 현재까지 연구된 바가 없다.

교육 분야에 데이터 마이닝을 적용하려는 연구 또한 활발히 진행되어 왔다[9-13]. 이 중 학업성취도와 관련하여 마이닝을 적용한 연구는 참고문헌 [9]에서 수행되었다. 참고문헌 [9]의 연구에서는 고등학생을 대상으로 1학기 성적과 학원수강 여부, 수업시간의 학습태도를 이용해 2학기 성적을 예측하는 의사결정트리를 생성하였다. 학업성취도 외에 데이터 마이닝을 교육 분야에 적용한 연구는 다음과 같다. 첫째, 참고문헌 [10]에서는 연관 규칙을 사용한 학생 정보 마이닝 시스템을 구현하였다. 둘째, 참고문헌 [11]에서는 실업계 고등학생의 진로 분석에 의사결정트리를 이용하였다. 셋째, 참고문헌 [12]에서는 의사결정트리, 신경망 등의 분석 기법을 이용해 고등학교 중도 탈락자 사례를 분석하였다. 넷째, 참고문헌 [13]에서는 의사결정트리를 이용하여 Vygotsky의 근접 발달 영역에 기반 한 동료(친구) 교수형 프로그램을 설계하였다. 그러나 이들 연구는 데이터 마이닝 기법을 사용했다는 특징이 있으나, 본 연구에서 초점을 맞추 학생의 생활 습관과는 특별한 관련이 없다.

3. 방학 중 생활 습관의 연관성 분석

3.1 마이닝 항목 도출

학생들의 방학 중 생활습관을 분석하기 위해 본 절에서는 학생들이 방학 중에 할 수 있는 활동들을 도출한다. 먼저, 방학 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 학습 요소들과 일기 작성, TV 시청 시간 등과 같은 생활 요소들을 선정한다. 마지막으로, 이와 같이 선정된 요소들을 대상으로 설문 데이터를 작성한다.

학습관련항목은 직관적으로 생각했을 때 학업성취도에 영향을 줄 가능성이 높은 요소들로 선정한다. 즉, 학생들이 학업성취도 상승을 위하여 의도적으로 실시하는 학습 방법을 학습관련항목이라 한다. 학습관련항목에는 학원수강, 집단 또는 개인 대상의 사실과외, 교육방송 시청, 학습지에 의한 학습, 인터넷 학습사이트 이용 등이 있다. 학습관련항목에 대해서는 방학 전후의 성적과 함께 국어, 영어, 수학 등 각 과목별로 데이터를 수집한다.

생활관련항목은 학생들이 방학 중 수행하는 일반적인 활동들 중에서 학업성취도와 간접적 관련이 있는 요소들을 사용한다. 즉, 학업성취도 상승이 직접적인 목적이 아닌 일상적인 생활습관이나 취미 등을 생활관련항목으로 사용한다. 이러한 생활관련항목으로는 우선 생활습관 측면에서 컴퓨터 사용 시간, TV 시청시간, 운동 시간, 봉사활동 여부, 이성 친구 유무, 아침식사 여부, 부모님과의 대화 여부 등을 사용한다. 또한, 특정 과목의 학업성취도와 간접적으로 연관이 있을 가능성이 있는 일기 작성 여부, 편지나 메일 작성 여부, 홈페이지(미니홈피나 블로그 포함) 운영 여부, 독서 여부, 어학연수 여부 등을 선정한다. 그리고 기타 항목으로서 목표 학교, 여행 여부, 다른 국가 거주 여부 등을 사용한다.

3.2 연관 규칙 분석

학생들의 방학 중 생활 습관 사이에 특이한 연관 관계가 있는지를 분석하기 위해서, 본 논문에서는 대표적 데이터 마이닝 기법인 연관 규칙[1,14,15]을 사용한다. 연관 규칙의 추출은 우선 빈발항목집합(frequent itemset)을 찾고, 이들 빈발항목집합에서 연관 규칙을 찾는 과정으로 이루어진다[14]. 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[15].

연관 규칙 추출에는 제3.1절에서 설명한 학습관련항목과 생활관련항목 모두를 사용한다. 연관 규칙을 추출하기 위해서는 설문 결과를 정제하는 과정이 필요하다. 그 이유는 연관 규칙에서는 특정 항목이 트랜잭션에 나타나지 않거나 나타나지 않는지만을 고려하는 반면에, 설문 항목에는 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 필요하지 않지만, "TV시청 시간"과 같이 답변이 "1시간/1시간~2시간/2시간~3시간"으로 주어지는 연속형 항목의 경우 불리언형으로의 변환이 필요하다.

표 1. 연관 규칙 추출을 위한 입력 트랜잭션 예제.

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

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

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

4. 실험 및 분석

4.1 데이터 수집

학생들의 방학 중 생활습관과 활동이 학업성취도에 미치는 영향을 분석하기 위해 설문조사를 통해 현실의 구체적인 데이터를 수집하였다. 설문조사서는 제3.1절에서 도출한 항목들을 기반으로 작성하였다. 설문조사는 그림 1과 같이 온라인 인터넷 사이트에서 진행하였다. 온라인 설문조사 전문업체인 리서치중앙[16]에 설문을 의뢰하였으며, 사이트는 여자중학교를 중심으로 홍보하였다. 데이터 수집을 위한 설문조사 기간은 2005년 5월 17일부터 2005년 5월 31일까지의 15일로 하였다.

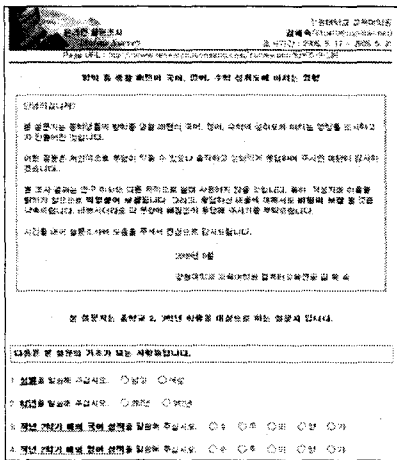


그림 1. 온라인 설문 조사.

설문에 응답한 학생은 총 536명으로, 이들의 성별 및 학년 분포는 표 3과 같다. 응답한 내용 중 질문의 일부 항목에만 대답한 불완전 응답 데이터를 제거하였고, 남학생과 여학생 비율이 크게 차이가 나서 남학생 응답 데이터 또한 데이터 구성에서 제외하였다. 설문조사 사이트가 여자중학교에 중점 소개된 관계로 여학생들의 설문 응답이 많은 반면 남학생들의 응답은 저조하였다.. 이에 따라, 실제 실험에서는 불완전 데이터와 남학생 데이터를 제외한 여학생 462명의 데이터를 사용하였다.

4.2 연관 규칙 추출 방법

본 연구는 데이터 분석 및 마이닝을 위하여 E-Miner를 사용한다. E-Miner는 분석흐름도(PFD: process flow diagram)와 노드(node)를 이용하여 데이터 마이닝의 전 과정을 GUI환경에서 쉽게 수행할 수 있고, 다양한 DBMS에 접근이 용이하다. 또한, 다양한 통계분석, 신경망, 의사결정트리 등의 여러 분석 기법을 제공하며, 데이터 추출에서 결과 반영까지의 전 과정이 프로그램 코드로 생성되므로 재실행이 용이하다.

E-Miner를 사용하여 데이터 마이닝을 수행하는 절차는 표본추출, 데이터 탐색, 데이터 변환, 모형화, 평가의 다섯 단계로 구성된다. 그림 2는 E-Miner의 초기화면을 나타낸다. E-Miner에서는 프로젝트 생성, 라이브러리 등록, 데이터 가져오기, 분석 흐름도 작성, 신뢰도와 지지도 설정, 연관규칙 추출의 여섯 단계를 거쳐 의사결정트리를 생성한다. 이들 각 단계에 대한 설명은 다음과 같다[3].

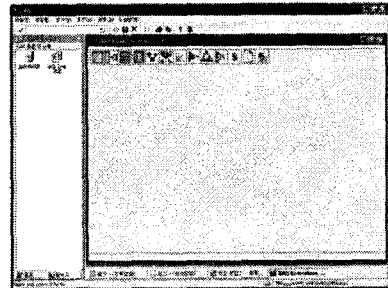


그림 2. E-Miner 초기화면.

단계 1) 프로젝트 생성: E-Miner에서는 데이터의 생성 및 추출, 탐색 및 변형, 모형화, 평가 등으로 이루어지는 작업을 프로젝트(project)라고 부른다. 따라서, E-Miner에서의 첫 번째 작업은 그림 3과 같이 프로젝트를 생성하는 것이다. 그림 3은 'DM프로젝트'라는 프로젝트를 등록하는 방법을 나타낸다.

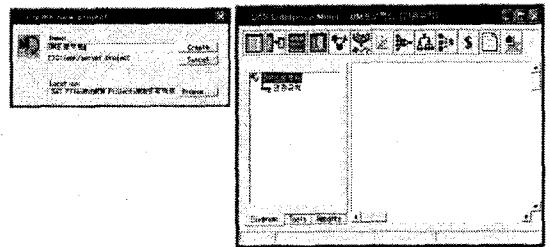


그림 3. E-Miner의 프로젝트 등록 예제.

단계 2) 라이브러리 등록: E-Miner에서는 특정 폴더에 대해서 그 이름을 직접 사용할 수가 없으며, '라이브러리(library)'라고 불리는 별칭을 사용해야 한다. 라이브러리를 생성한 후 분석할 데이터 세트(data set)를 그 안에 등록할 수 있다.

단계 3) 데이터 가져오기: 라이브러리 창에서 다양한 형식의 외부 파일들을 라이브러리에 등록할 수 있다. 그림 4는 Microsoft의 Excel 형식 파일을 가져오는 방법을

나타낸다.

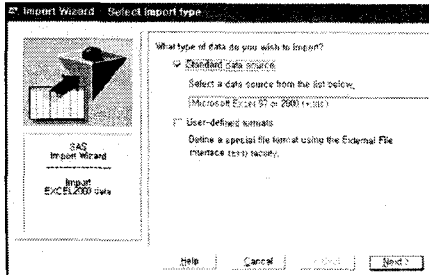


그림 4. E-Miner에서 Microsoft Excel 데이터 가져오기 예제.

단계 4) 분석흐름도 작성: 그림 5는 네 번째 단계로서 연관 규칙에 대한 분석 흐름도를 작성하는 단계의 예제 화면이다. 분석 흐름도 작업 영역(diagram workspace)은 프로젝트의 일부 작업을 그림 5와 같이 흐름도의 형태로 작성하는 작업공간이다.

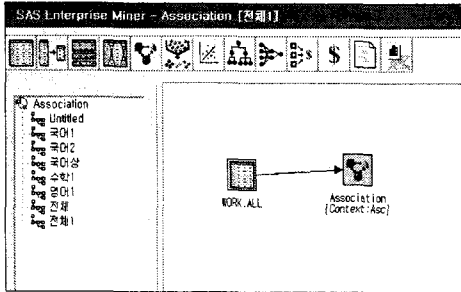


그림 5. E-Miner 연관 규칙 추출 분석 흐름도 예제.

단계 5) 신뢰도와 지지도 설정: 다섯 번째 단계는 연관 규칙 노드의 일반탭에서 신뢰도와 지지도, 하나의 규칙에 포함될 수 있는 항목의 개수를 설정하는 과정으로서 그림 6이 이 단계의 예제 화면이다.

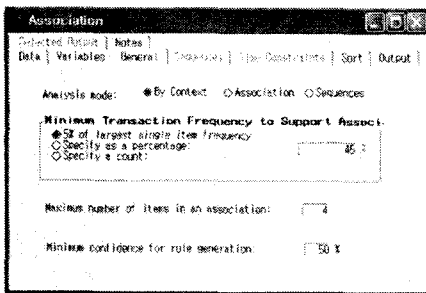


그림 6. E-Miner 연관 규칙 노드의 일반탭 화면.

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, 학습 및 생활관련항목을 모두 사용하였으며, 연속형 항목을 불리언형으로 변환하여 실험하였다.

그림 7은 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.

Relations	Lift	Support(%)	Confidence(%)	Transaction Count	Rule
1	1.06	52.16	33.71	246.00	영학=>국학
2	1.06	49.75	30.84	225.00	국학=>영학
3	1.07	46.10	30.71	212.00	국학=>수학
4	1.07	41.13	27.23	183.00	수학=>국학
5	1.08	33.53	23.88	154.00	국학=>수영
6	1.08	34.61	24.30	153.00	수영=>국학
7	1.06	29.19	22.02	151.00	국학=>영어
8	1.06	29.59	22.72	139.00	영어=>국학
9	1.06	30.43	24.00	136.00	국학=>수학
10	1.06	30.43	24.00	136.00	수학=>국학
11	1.06	30.43	24.00	136.00	국학=>수영
12	1.06	30.43	24.00	136.00	수영=>국학
13	1.06	30.43	24.00	136.00	국학=>수학
14	1.06	30.43	24.00	136.00	수학=>국학
15	1.06	30.43	24.00	136.00	국학=>수영
16	1.06	30.43	24.00	136.00	수영=>국학
17	1.06	30.43	24.00	136.00	국학=>수학
18	1.06	30.43	24.00	136.00	수학=>국학
19	1.06	30.43	24.00	136.00	국학=>수영
20	1.06	30.43	24.00	136.00	수영=>국학
21	1.06	30.43	24.00	136.00	국학=>수학
22	1.06	30.43	24.00	136.00	수학=>국학
23	1.06	30.43	24.00	136.00	국학=>수영
24	1.06	30.43	24.00	136.00	수영=>국학

그림 7. 신뢰도 90%, 지지도 30%인 경우의 연관 규칙 결과.

다음으로, 그림 8은 지지도를 30%에서 45%로 높은 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.

Relations	Lift	Support(%)	Confidence(%)	Transaction Count	Rule
1	2	1.05	53.05	91.00	295.00 (여행=>아침)

그림 8. 신뢰도 90%, 지지도 45%인 경우의 연관 규칙 결과.

실험 1)과 실험 2)의 결과를 종합해 보면, 방학 중 생활습관 사이에서는 특이한 규칙이 발견되지 않았다. 특히, 방학 중 생활습관이 학업성취도에는 별다른 영향을 미치지 않는 것으로 나타났다.

5. 결론

본 연구에서는 학생들의 방학 중 생활 습관을 데이터 마이닝 기법의 연관규칙을 사용하여 분석하였다. 현재까지 진행된 교육 분야의 관련 연구에서는 방학 중 생활 지도에 관한 구체적인 접근법이 없었다. 본 연구에서는 데이터 마이닝 기법의 연관규칙을 사용하여 다음과 같은 구체적 분석 방법을 제안하였다. 우선, 방학 중 학습방법과 생활습관에 대한 다양한 요소들을 도출하였다. 또한, 현실의 구체적 데이터 수집을 위해 중학교 학생들을 대상으로 방학 전후 성적과 방학 중 활동에 대한 설문조사를 수행하였다. 마지막으로, 방학 중 학습방법과 생활습관들을 대상으로 연관 규칙을 추출하고 그 의미를 분석하였다. 그 결과, 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 방학 중의 학습방법과 생활습관을 분석하였다. 이와 같은 연구 결과는 교사나 학부모가 학생의 생활 및 학습 지도를 하는데 유용하게 사용될 수 있다. 또한 학생들에게는 방학 중 계획을 세우는 자료로 활용될 수 있다.

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] <http://www.sas.com/korea>.
- [3] <http://www-306.ibm.com/software/data/iminer>.
- [4] <http://www.dbminer.com>.
- [5] 김동화, 조영재, "수학 교수, 학습 태도에 따른 학업 성취도의 수준에 관한 분석," *교육이론과 실천*, 제14권, 제1호, pp. 115-132, 2004.
- [6] 김병식, 김민정, "청소년의 성격유형이 학업성취도에 미치는 영향에 관한 연구," *한국인간관계학보*, 제7권, 제1호, pp. 1-20, 2002.
- [7] 김아영, 차정은, "교사효능감 및 학생의 학업적 효능감이 학업성취도에 미치는 영향에 따른 다층분석," *교육심리학회*, 제17권, 제2호, pp. 25-43, 2003.
- [8] 하정순, 이호지, "중학생의 식행동이 건강상태와 학업성취도에 미치는 영향," *대한가정학회지*, 제33권, 제3호, pp. 225-242, 1995.
- [9] 배재호, 데이터 마이닝을 이용한 학업성취도 분석, 경희대학교 교육대학원 석사학위논문, 2001.
- [10] 공현성, 김영, "학생정보마이닝 시스템의 설계 및 구현," *한국컴퓨터교육학회논문지*, 제6권, 제1호, p. 58-63, 2003.
- [11] 김재은, 데이터 마이닝을 이용한 실업계 고등학교 학생의 진로분석, 경기대학교 교육대학원 석사학위논문, 2002.
- [12] 송미현, 데이터 마이닝을 활용한 교육정보화 실천에 관한 연구-중도탈락 사례분석을 중심으로, 전주대학교 석사학위논문, 2001.
- [13] 성경숙, Data Mining을 이용한 잠재적 학습능력 예측 및 P2P기반 동료교수형 프로그램 설계, 연세대학교 교육대학원 석사학위논문, 2004.
- [14] 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, D C, pp. 207-216, May 1993.
- [15] 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.
- [16] <http://www.research.joongang.com>.