

사이버교육활성화를 위한 CRM방법의 적용에 관한 연구

- A Study on the CRM Application for Activation of
Cyber Education -

김 한 신*

Jin Han Xin

이 공 섭*

Lee Gong Seob

이 창 호*

Lee Chang Ho

Abstract

Nowdays cyber education based on the internet is actively developed. But the management of the customers in the cyber education field is not enough. Then, in this paper, we provide the learner with the proposals of lectures to be extremely matched by analyzing the learning capacity and the greatest concern of him(her) using the methods of data mining, such as RFM, prediction, stickness, association rule, classification, and so on.

1. 서론

1.1 사이버교육

1.1.1 사이버교육의 정의

사이버교육은 인터넷을 기반으로 다양한 멀티미디어 정보기술을 활용하여 사용자가 원하는 시간과 장소인 사이버 공간에서 이루어지는 교육 방식으로 특히 강사와 학습자간의 시간적, 공간적 제약 없이 이루어지는 학습활동을 통하여 수요자 중심의 교육서비스가 제공되는 새로운 개념의 교육체계를 칭한다.

* 인하대학교 산업공학과

** 유한대학 산업시스템경영과

1.1.2 사이버교육의 장점

- 1) 유연성(flexibility).
- 2) 교육비용이 저렴하며 비용대비 효율성이 높다.
- 3) 교육 기회와 대상을 최대화 할 수 있다.
- 4) 심리적 부담감이 해소된다.

1.1.3 사이버교육의 단점

- 1) 상호작용 (interactivity)의 제한이다.
- 2) 첨단 미디어를 활용할 경우 초기 개발 비용이 높다는 점이다.
- 3) 실험, 실습과 같은 동작 훈련 등의 교육에는 적합하지 않다는 점이다. [2]

1.2 연구 내용 및 방법

최근 교육업체들은 고객 개개인의 학습능력 분석을 통하여 학습자의 눈높이에 맞춘 서비스를 제공하기 위해 CRM에 대한 관심이 높아지고 있다. 하지만 지금까지 대부분 사이버강의프로그램은 단순한 강의 듣기수준에 그치고 있다. 본 연구에서는 사이버강의분야에서 CRM을 추진하기 위해 데이터 마이닝기법을 활용한 두개 전략을 제안한다. 하나는 기존학습자에 대한 추천전략이고, 하나는 신규학습자에 대한 추천전략이다.

본 연구는 인터넷을 기반으로 음성, 화상, Tablet등 다양한 멀티미디어 정보기술을 활용하여 분산 환경 구축을 위한 훌륭한 솔루션으로서 마이다스를 활용하여 서버부분, 강사프로그램부분과 학습자프로그램부분으로 구성 되어 있고 개발Tool로는 Borland사의 delphi를 사용하였고 데이터베이스는 Microsoft사의 SQL 2000서버를 사용하였다.

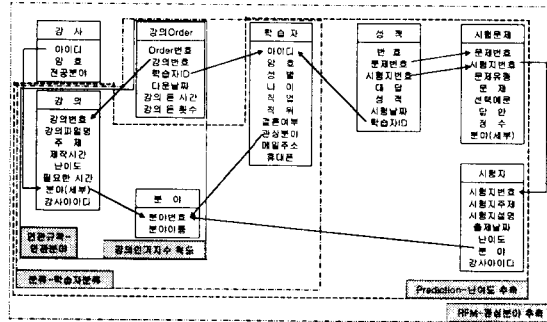
기존의 강사프로그램에서 강의 만들기 등 단순한 기능만 갖고 있지만, 본 연구에서 구축된 강사프로그램은 학생의 학습상황을 정확하게 파악하기 위한 출제기능을 제공하였다. 또한 기존의 학생프로그램은 단순히 강의를 듣는 기능만 있지만, 본 연구에서 구축된 학생프로그램은 CRM을 추진하기 위하여 서버프로그램부분에서 데이터 마이닝 기법을 활용하여 학습자가 로그인 할 때마다 학습자자료를 분석하고 학습자의 수준에 맞는 강의내용을 추천한다. 이에 따라 학습자의 충성도를 강화할 수 있다.

2. 사이버교육 프로그램 구축

2.1 사이버교육프로그램

2.1.1 데이터베이스 구조

본 연구에서 구축된 데이터베이스의 구조는 [그림 1]과 같다.



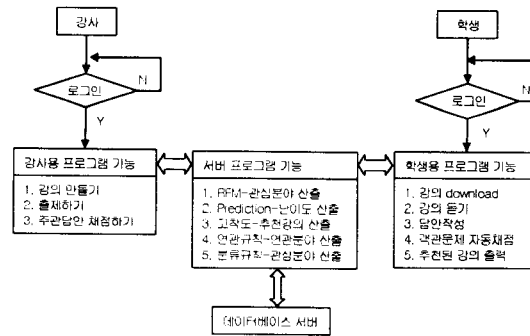
[그림 1] 데이터베이스 구조

그리고 본 연구에서는 가상 데이터를 만들었다. 각 테이블에 대해 만드는 코더개수는 아래와 같다.

학습자레코더: 300개; 강의레코더: 50개; 강의 Order레코더: 3000개; 시험지레코더: 50개; 시험문제레코더: 500개; 성적레코더: 30600개.

2.1.2 사이버교육 프로그램 흐름도

본 연구에서 구축된 강의프로그램은 서버프로그램 부분, 강사용 프로그램부분과 학습자용 프로그램부분으로 구성 되어 있으며 전체적인 프로세스는 [그림 2]와 같다. 가장 핵심적인 부분은 CRM을 추진하고 있는 서버프로그램 부분이다.



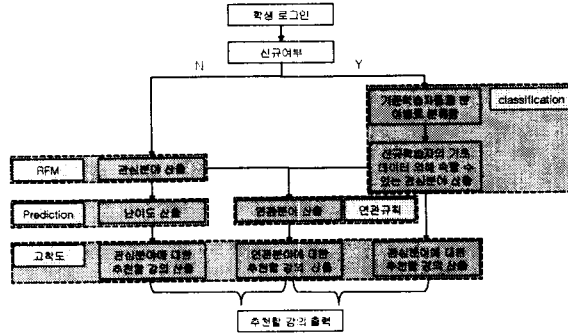
[그림 2] 사이버교육프로그램 흐름도

2.2 서버프로그램

2.2.1 강의 추천 흐름도

서버프로그램은 client프로그램과 데이터베이스사이의 중계자로서 크게 로그인과 CRM 부분으로 나뉘어 있다. CRM부분은 신규학습자와 기존학습자를 구별하고 데이터 마이닝기법을 활용하여 학습자의 학습능력을 분석하며 학습자의 수준에 맞는 학습내

용을 학습자에게 추천한다. 전체 강의 추천 흐름도는 [그림 3]과 같다.



[그림 3] 강의 추천 흐름도

2.2.2 기존 학습자에 대한 강의 추천 전략

기존 학습자에 대해 데이터 마이닝 기법을 활용하여 학습자가 듣는 강의, 시험결과 등에 의하여 강의 추천 전략은 아래와 같은 5 단계를 포함하고 있다.

1. RFM 기법을 활용하여 관심 분야를 산출

RFM을 활용하여 학습자에 대해 분야별로 point를 계산하고 Point의 최고 값을 갖고 있는 분야를 학습자의 관심 분야로 결정한다. point를 계산하는 방법은 아래와 같다.

1) 분야별로 point 계산함 : $Point = R + F + M$

①R: 얼마나 최근에 우리 강의를 download했는가?

②F: 얼마나 빈번하게 우리 강의를 download했나?

③M: 취약분야와 관심분야의 총합

2) 관심 분야를 결정 : Point 최고 값을 갖고 있는 분야를 학습자의 관심 분야로 결정

2. Prediction을 활용하여 추천할 강의의 난이도를 결정

Predication을 활용하여 차후의 시험결과를 예측하고 예측결과와 일정규칙에 의해 추천할 강의의 난이도를 결정한다. 차후의 시험결과를 예측하는 방법은 아래와 같다.

1) 관심 분야에 대해 난이도별로 차후의 시험결과를 예측한다.

$$\text{시험성적} = \alpha + \beta * \text{시험횟수}$$

$$\alpha = \text{평균점수} - \beta * \text{평균시험횟수}$$

$$\beta = \frac{\sum_{i=1}^n (i - \text{평균시험횟수}) * (i\text{-번째 시험 성적} - \text{평균성적})}{\sum_{i=1}^n (i - \text{평균시험횟수})^2}$$

2) 일정 규칙에 따라 추천할 강의의 난이도 결정한다.

3. 고착도 기법을 활용하여 추천할 강의를 산출

상위 1, 2방법으로 결정된 분야 및 난이도에 속하는 모든 강의들에 대해 고착도 기법을 활용하여 그들의 인기지수를 계산하여 고착도가 높은 순서로 학습자에게 강의들을 추천한다. 강의의 인기지수는 아래와 같은 식들에 의해 계산할 수 있다.

강의의 인기지수 = 강의를 Download하는 빈도 * 본 시간 * 총 download 횟수

$$\text{강의를 download 하는 빈도} = \frac{\text{기간 } T \text{ 동안의 강의 order 수}}{\text{기간 } T \text{ 동안 방문한 유일 이용자의 총수}}$$

$$\text{강의를 본 시간} = \frac{\text{모든 강의의 총 강의시간}}{\text{기간 } T \text{ 동안 방문한 유일 이용자의 총수}}$$

$$\text{강의를 download 하는 총 횟수} = \frac{\text{기간 } T \text{ 동안 방문한 유일 이용자의 총수}}{\text{총 유일 이용자의 수}}$$

4. 연관규칙을 활용하여 관심분야의 연관분야를 산출

강의들을 교차 추천하여 학습자당 수익을 극대화시키기 위해 연관규칙을 활용해서 관심분야에 대해 다른 모든 분야와의 지지도, 신뢰도, 향상도를 구하여 연관분야를 산출한다. 다른 분야와의 지지도, 신뢰도, 향상도는 아래의 식들에 의해서 계산된다.

$$\text{지지도 (관심분야} \Rightarrow \text{분야 } i) = \frac{\text{관심분야의 강의와 분야 } i \text{의 강의를 같이 order 횟수}}{\text{총 Order 수}}$$

$$\text{신뢰도 (관심분야} \Rightarrow \text{분야 } i) = \frac{\text{관심분야의 강의와 분야 } i \text{의 강의를 같이 order 횟수}}{\text{관심분야의 강의를 order 총 횟수}}$$

$$\text{향상도 (관심분야} \Rightarrow \text{분야 } i) = \frac{\text{confidence (관심분야} \Rightarrow \text{분야 } i)}{\text{분야 } i \text{의 강의 order 횟수}} \\ \text{총 Order 수}$$

그리고 지지도, 신뢰도, 향상도들은 일정한 최소값 보다 커야 연관관계가 의미있다.

5. 고착도기법을 활용하여 연관분야에 대해 추천할 강의를 산출

관심분야와 연관분야에 대해 추천할 강의들을 산출하고 학습자에게 전송한다.

앞에서 언급한 데이터 마이닝기법들과 가상 데이터에 의해 컴퓨터에서 산출한 결과는 [그림 4]와 같다.

A. 가상 기초 데이터.

B. RFM을 활용하여 학습자에 대해 분야별로 point를 계산한 결과물이고 국어는 Point의 최고 값을 갖고 있기 때문에 학습자의 관심분야로 결정한다.

C. Predication를 활용하여 예측한 차후의 시험결과이다.

D. 상위 B, C방법으로 결정된 국어 및 난이도5의 강의들에 대해 강의인기지수를 구하고 인기지수가 제일 많은 강의인 "200201010101hanxu"를 학습자에게 추천한다.

E. 연관규칙을 활용해서 관심분야 국어에 대해 다른 모든 분야와의 지지도, 신뢰도, 향상도를 구하였고 영어는 관심분야의 연관분야로 결정한다.

F. 연관분야의 강의들에 대해 인기지수를 구하는 결과물이다.

마지막은 프로그램에서 추천한 강의리스트이다.

나머지 3개 단계는 기존학습자의 3)부터 5)까지와 같다.

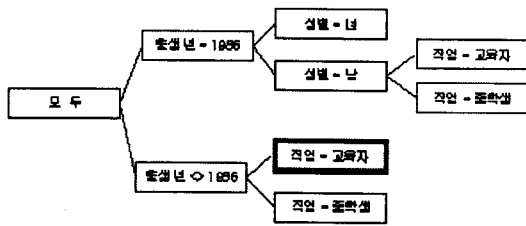
앞에서 언급한 데이터 마이닝기법들과 가상 데이터에 의해 컴퓨터에서 산출한 결과는 [그림 6]과 같다.

- A. 가상 기초 데이터.
 - B. SQL 2000서버를 활용하여 추측된 의사결정나무이다.
 - C. 학습자의 기초데이터와 의사결정나무에 의해 관심분야는 수학으로 결정한다.
 - D. 고착도 기법을 활용하여 강의인기지수를 구하는 결과물이다.
 - E. 관심분야 수학에 대해 다른 모든 분야와의 지지도, 신뢰도, 향상도를 구하였고 미리 정의된 최소 값을 보다 작기 때문에 수학은 연관분야가 없다.
- 마지막은 프로그램에서 추천한 강의리스트이다.

A 기본 데이터

이름	이희희	관심분야	성별	출생년	직업	결혼여부	학력
교용유	good	수학	남	1975	교육자	미혼	대학卒

B 의사결정나무 구축 - SQL 2000의 기능을 활용해서 구축된 의사결정나무



노드	노드명	노드경로(D)
001	모두	출생년 < 1986 그리고 직업 = 교육자
002	출생년 < 1986	
003	성별 = 남	
004	성별 = 여	
005	출생년 > 1986	
006	직업 = 교육자	
007	직업 = 중학생	

C 관심분야 산출: 001002002003 - 수학

D 강의인기지수 (관심분야 = 수학, 유일이용자수 302, 총 학습자수 303)

강의번호	시간(초)	강의 down 수	인기지수
20020101010126hannu	500	47	0.6002
20020101010126hannu	500	44	0.2885

E 연관분야 (수학(A)를 포함하는 order의 건수120, 총 order건수 902)

분야(B)	A, B 같이	B포함한 order수	지지도	신뢰도	향상도
국어	29	76	0.096	0.245	0.975
영어	58	147	0.165	0.467	0.959
사회	69	225	0.228	0.772	0.772
과학	85	120	0.116	0.282	0.745

연관분야 없음

프로그램에서 추천한 강의는:

추천한 강의:

[그림 6] 컴퓨터에서 산출한 결과

2.3 강사프로그램

강사용 프로그램은 강의제작, 출제하기, 채점등 여러 가지 기능을 갖고 있다.

강의 작성에는 기존 문서파일, 그림파일, 스크린 캡처 화면을 불러와서 이미지화 하고 교안을 만든다. 강의를 녹화할 때 여러 가지 기본 기능들을 갖고 있다. 음성 녹화, 화상 녹화, 간단한 도형 그리기, 문자 쓰기, 마우스포인터 추적하기, 지우기, 전체 지우기, 또한 화상녹화여부는 선택할 수 있고 Tablet을 활용하여 풍부한 강의내용을 입력할 수 있다.

출제는 시험분야, 난이도를 선택하고 시험문제 작성한다. 시험문제는 주관식과 객관식문제를 포함하며 객관식에는 단항선택과 다중선택이 있고, 또한 본 시험문제 세부분야를 설정할 수 있다.

채점은 자동과 수동으로 이루어지는데 객관식문제는 학습자가 시험지를 제출하는 시점에서 자동으로 채점되어 주관식 문제가 없으면 시험총점으로 DB에 적장된다. 주관식문제가 있을 때에는 강사가 주관식문제를 채점하여 올리면 객관식 문제의 점수와 결합하여 총점으로 DB에 저장된다.

2.4 학생프로그램

학생용 프로그램은 주로 강의 다운로드하기, 강의 듣기, 시험보기 등 여러 가지 기본 기능을 갖고 있고, 또한 서버에서 제공한 여러 가지 데이터마이닝기법들을 활용하여 학습자의 수준에 맞는 강의를 추천, 출력할 수 있다.

3. 결론 및 추후연구과제

본 연구는 먼저 훌륭한 솔루션으로서 마이다스를 활용하여 사이버강의프로그램을 구축하며 CRM을 추진하기 위해 이를 기반으로 고객편에 서서 데이터 마이닝기법을 활용하여 맞춤형 사이버교육시스템을 구축하였다.

첫째, 기존학습자에 대해 데이터 마이닝기법인 RFM, Prediction, 고착도등을 활용하여 학습자가 들은 강의, 시험결과 등을 분석하며 학습자의 수준에 맞는 강의를 추천하며 학습자의 학습능률을 높이고 충성도를 강화할 수 있다. 둘째, 신규학습자에 대하여 분류규칙을 활용하여 기초데이터를 분석하며 학습자의 심리를 파악하고 학습자가 흥미를 갖고 있는 맞춤 강의를 제공함으로써 대량 학습자를 유지할 수 있다. 셋째, 고착도기법을 활용하여 항상 수준이 높은 강의들을 학습자에게 추천함으로써 서비스 만족도를 높이고 학습자의 충성도를 강화할 수 있다. 이런 기법들을 활용하여 교육업체의 사이버교육을 활성화시키고 수익성 증대를 기대한다. 또한 연관규칙을 활용하여 관심분야의 연관분야를 산출하고 강의의 교차추천을 진행함으로써 학습자당 수익을 극대화시킬 수 있다.

본 연구에서도 여러 가지 한계점이 있다. 강사와 학생의 상호작용의 제한점, 관리자가 강사와의 상호작용의 제한점, 강의평가의 부족함과 강사와 학부모와의 상호작용의 제한점 등이 있다.

이 같은 사항들을 보완하기 위해서 추후 연구과제로는 학부모들이 학생의 수강정황, 능력정황 등 학생관련 정보들을 잘 파악하고 학습자를 잘 관리하기 위해 학부모모드를 구축하고, 사이버교육의 단점인 상호작용의 제한을 해결하여 기존 고객의 유지 강화 실현을 위한 화상면담 시스템도 구축할 것이다. 또한 학습자들을 지속적으로 유지하기 위해 다양하고 질이 높은 강의가 필요하기 때문에 강사 평가시스템도 구축할 것이다.

4. 참고문헌

- [1] 김봉관 저, 남두도서, 데이터베이스 마케팅을 위한 고객관리, 2000
- [2] 리넷 R. 포터 지음 / 이성은 옮김, 가상교실 만들기 -- 인터넷을 이용한 사이버학습, 한울, 2000
- [3] 백운기, 한상홍, 박준후 저, delphi 5 contact, www.daerimbook.co.kr, 2000
- [4] 한국통신 하이텔 전성훈 / 한국IBM 최현희 지음, E-CRM 실무지침, 삼각형프레스, 2001
- [5] Claude Seidman 저/ 권오주. 이정무 역, Microsoft SQL Server 2000 데이터 마이닝 Technical reference, 정보문화사, 2001
- [6] Jiawei Han, Micheline Kamber, Data Mining Concepts and Techniques, An imprint of Academic press, a Harcourt Science and Technology Company, 2000

참고 웹 사이트

- [7] www.fooledison.com, 교육용 CRM 솔루션 제공업체
- [8] www.liveshare.co.kr, 사이버교육

저 자 소 개

김한신: 인하대학교 산업공학과 대학원 석사. 관심분야는 OR, 전자상거래 등.

이공섭: 인하대학교 섬유공학과 졸업. 동대학원 산업공학과 석사 및 박사학위를 취득. 현 유한대학 산업시스템경영과 교수로 재직중. 관심분야는 의사결정론, 경제성공학, 응용통계학 및 전산활용 등이다.

이창호: 인하대학교 산업공학과 졸업(공학사)하고, 한국과학기술원 산업공학과 공학석사, 박사학위를 취득하였으며, 현재 인하대학교 산업공학과 교수로 재직중이다. 주요관심분야는 인천항만과 관련된 물류최적화(항만 운영, 하역 효율화), 인천항만과 관련된 Simulation/Simulator 개발, 물류네트워크 구축 물류정보화 구축(DataBase/Data Warehouse/Data Mining 기술 활용), 중소기업 정보화 구축(Microsoft Tool의 활용), 화상회의시스템 구축과 활용 등이다.