

# ITS를 위한 개인화 학습코스 추천 모델 개발

한지원<sup>1</sup>, 조재춘<sup>2</sup>, 임희석<sup>3\*</sup>

<sup>1</sup>고려대학교 컴퓨터학과 박사수료, <sup>2</sup>고려대학교 컴퓨터학과 박사, <sup>3</sup>고려대학교 컴퓨터학과 교수

## Development of Personalized Learning Course Recommendation Model for ITS

Ji-Won Han<sup>1</sup>, Jae-Choon Jo<sup>2</sup>, Heui-Seok Lim<sup>3\*</sup>

<sup>1</sup>Ph. D. Candidate, Dept. of Computer Science and Engineering, Korea University

<sup>2</sup>Ph. D., Dept. of Computer Science and Engineering, Korea University

<sup>3</sup>Professor, Dept. of Computer Science and Engineering, Korea University

요 약 학습코스 선정에 많은 어려움과 시행착오를 겪고 있는 사용자들에게 수준별 학습코스를 제공하기 위해, ITS (Intelligence Tutoring System)를 위한 동적인 학습자 맞춤형 학습코스 추천 모델을 개발하였다. 이를 위해, 개인화 학습코스 추천모델에서는 먼저 학습자 프로파일을 분석하고, 단어별 가중치를 계산하여 핵심 키워드를 추출한다. 추출된 단어는 Cosine Similarity 기법을 통해 유사도를 측정하고, 최종적으로 유사도가 높은 상위 3개 과정이 학습자에게 추천된다. 추천 모델의 효과를 분석하기 위해, 경기도 소재 교육기관에 추천모델을 적용하였고, 만족도 조사를 통하여 설문 항목별 평균, 표준편차, 왜도, 첨도 값을 계산하였다. 실험결과, 정확성, 새로움, 자기참조, 유용성에서 높은 만족도를 보였으며, 추천모델의 실효성을 검증했다. 본 연구는 그동안 국내·외에서 충분히 다루지 않았던 기계학습 중심의 맞춤형 학습코스를 추천했다는 점에서 의미가 있다.

주제어 : TF-IDF, 유사성, 지능형교수시스템, 개인화, 추천시스템

**Abstract** To help users who are experiencing difficulties finding the right learning course corresponding to their level of proficiency, we developed a recommendation model for personalized learning course for Intelligence Tutoring System(ITS). The Personalized Learning Course Recommendation model for ITS analyzes the learner profile and extracts the keyword by calculating the weight of each word. The similarity of vector between extracted words is measured through the cosine similarity method. Finally, the three courses of top similarity are recommended for learners. To analyze the effects of the recommendation model, we applied the recommendation model to the Women's ability development center. And mean, standard deviation, skewness, and kurtosis values of question items were calculated through the satisfaction survey. The results of the experiment showed high satisfaction levels in accuracy, novelty, self-reference and usefulness, which proved the effectiveness of the recommendation model. This study is meaningful in the sense that it suggested a learner-centered recommendation system based on machine learning, which has not been researched enough both in domestic, foreign domains.

**Key Words** : TF-IDF, Similarity, Intelligence Tutoring System, Individualization, Recommendation System

### 1. 서론

추천 시스템은 사용자의 과거 행동 데이터나 로그 데

이터를 바탕으로 사용자에게 필요한 정보나 제품을 사용자의 성향에 맞춰 추천함으로써, 자신이 원하는 품목을 쉽게 획득할 수 있도록 돕고 있다. 이러한 추천 시스템은

\*Corresponding Author : Heui-Seok Lim (limhseok@korea.ac.kr)

Received August 7, 2018

Accepted October 20, 2018

Revised October 1, 2018

Published October 28, 2018

개인의 성향에 따라 효과가 다르게 나타나는 교육 환경에서도 필요한 부분이며 다양한 방법으로 교육 분야에 응용되고 있다[1]. 특히, 선택에 대한 어려움과 시행착오를 줄여 학습효과를 증진시킬 수 있도록 학습자 수준에 적합한 학습코스를 제공해 줄 수 있다는 점에서 추천시스템은 비형식 학습 환경에서 매우 유용하다[2]. 이는 학습자 개개인의 학습수준이나 학습내용 등에 따라 적합한 학습주제를 선정하여 제공해 줌으로써, 학습자 스스로가 자신의 학습과정을 계획하고 구성하는 것이 가능해지기 때문이다. 이를 위해, 기존의 CAI(Computer Assisted Instruction)에 인공지능(Artificial Intelligent)의 원리와 기법을 적용해서 학습자에게 더욱 지능적이고 적응력 있는 교육 시스템을 제공하기 위해 개발된 지능형 교육 시스템(ITS, Intelligent Tutoring System)[3]을 기반으로 추천시스템을 제안한다. ITS는 교수, 학습자, 시스템이 유기적으로 상호작용하고 여기에 효과적인 인터페이스가 가미되어 학습 효과를 높이고자 구성된 시스템이다. 즉, 인공지능의 개념을 교육 분야에서 적용하여 컴퓨터의 사고하는 기능과 자연스러운 대화능력, 인간교사와 유사한 방식으로 학습자의 문제를 진단하고 그에 적합한 개인별 특성 및 요구에 적용된 최적의 적응적 교수를 지원한다. ITS는 크게 전문가 모듈, 학습자 모듈, 교수 모듈, 인터페이스 모듈의 4가지 기본적인 모듈로 구성되어 있으며[4], 이를 기반으로 본 연구는, 사용자 수준에 적합

한 학습코스를 추천하여 학습효과를 향상시킬 수 있는 추천시스템을 제안하고자 한다. 또한, 학습자 개개인의 특성에 맞추어 교육을 제공하는 맞춤형 교육에 대한 수학, 과학 도메인의 기존 연구[5,6]들은 많지만, 현재 급부상하고 있는 SW교육에 대한 맞춤형 교육 추천 연구는 부족한 것이 현실이다. SW교육의 경우, 정해진 규칙을 통해 학습자가 제출한 답안의 정답만을 채점하고 결과를 제공하는 일률적인 교육만으로는 학습자에게 맞춤형 교육을 제공하는 것이 어렵다. 규칙기반이 아닌, 기존의 온라인 교육을 통해 누적된 데이터와 다양한 인공지능 기법 등을 함께 이용하여 학습자의 학습 상태나 평가결과를 분석하고 추천하는 모델을 제안하고자 한다. 추천시스템의 효과를 분석하고 검증하기 위해 만족도 조사를 실시하여 효과성을 검증했다.

## 2. 개인화 학습코스 추천모델(PLCR)

본 연구는 ITS 기반의 개인화 학습코스 추천(PLCR, Personalized Learning Course Recommendation) 모델을 제안한다. PLCR 모델은 학습자의 특성 및 학습 수준 등을 고려한 체계적인 학습코스 추천이 가능하도록, 학습자 모듈, 교수 모듈, 전문가 모듈, 인터페이스 모듈로 구성된다. 학습자의 특성 및 학습 수준을 파악하기 위한 학

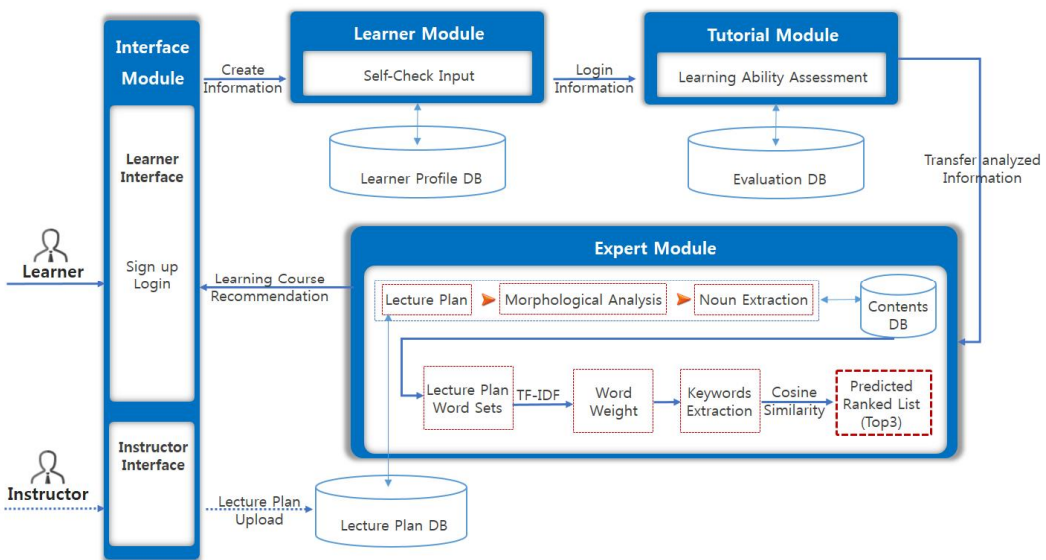


Fig. 1. A Structure of Personalized Learning Course Recommendation Model

습자 모듈과 기존의 학습자 수준을 알 수 있는 교수모듈을 통해 분석된 정보를 기반으로 개인화 학습코스를 추천한다. 이를 위해 전문가 모듈에서는 교수자가 업로드한 강의계획안을 기반으로 TF-IDF(Term Frequency - Inverse Document Frequency)[7]를 이용해 단어별 가중치를 계산하고, 핵심 키워드를 추출한다. TF-IDF를 통해 추출된 키워드를 기반으로 Cosine Similarity 기법을 이용해 유사도가 높은 상위 3개의 학습코스를 선정하여 학습자에게 추천한다.(2.2 참고) 제안하는 모델의 전체 구조도는 Fig. 1과 같다.

## 2.1 학습자 및 교수 모듈

학습자 모듈(LM, Learner Module)은 학습자들의 특성 및 학습 성향을 파악하고 학습목표에 따른 맞춤형 학습코스 추천을 위한 정보들을 학습자가 직접 입력하고 데이터베이스에 저장하여, 학습자 정보를 추출할 수 있도록 지원한다. 교수모듈(TM, Tutorial Module)은 학습자의 학습 상태를 평가하여 정량화 시키는 단계이다. 학습자가 자신의 개념 습득 정도나 학습 방향을 확인하여 학습자 스스로에게 동기가 부여되는 등의 ITS에서 중요한 역할을 한다. TM에서는 크게 두 유형으로 나누어 설계한다. 첫째, 초기 학습자의 학습자 모델을 위한 초기화 단계이다. 초기 학습자는 풀이한 문항 데이터가 존재하지 않기 때문에 학습자 모듈의 모든 요소를 0으로 초기화한다. 둘째, 학습자가 풀이한 문항의 정 오답 및 평가결과를 업데이트 하는 단계이다. 학습자에게 제공되는 각 문항의 정 오답 평가결과에 따라 학습자의 평가결과가 업

데이트 된다. 학습자 수준을 고려한 의미있는 평가를 위해서 문항반응이론(IRT, Item Response Theory)[8]을 적용하였다.

## 2.2 전문가 모듈

학습자의 특성을 파악하기 위한 학습자 모듈(LM)과 학습능력 평가로 학습자 수준을 알 수 있었던 교수모듈(TM)을 통해 분석된 정보를 기반으로 전문가 모듈(EM, Expert Module)에서 학습코스를 추천한다. 본 논문에서 제안하는 추천은 하이브리드 기반으로 ㉠과정에서는 협력기반의 추천과 ㉡과정에서는 콘텐츠 기반의 추천, ㉢과정에서는 사용자 기반의 추천기법을 사용하였다. ㉠, ㉡, ㉢의 결과를 종합하여 개인화 학습코스 추천 결과를 생성하게 된다. 추천과정에서 ㉡, ㉢는 해당 정보가 있는 경우에만 실행되며, LM의 프로파일 DB는 있으나, TM의 학습에 대한 평가를 거치지 않은 경우에만 진행된다.

또한, 추천과정에서 ㉠, ㉡, ㉢ 순으로 추천에 유용하다고 판단하여, 각각 다른 하이퍼파라미터(HP, HyperParameter)를 통해 가중치를 부여하였다. HP1은 0.5, HP2는 0.3, HP3은 0.2를 부여하고 HP가중치의 총합은 1이 되게 하였다. 즉 ㉠, ㉡, ㉢ 추천과정의 결과에 각각 0.5, 0.3, 0.2의 가중치를 부여하여 모든 결과를 합산한 후, 최종적으로 상위 3개의 학습코스를 선택하여 추천하게 된다. 만약, 추가 정보가 없어 ㉡, ㉢과정을 거치지 않았을 경우에는 ㉠과정의 결과가 가중치 1을 부여받고, 최종 추천결과가 된다. ㉠과정은, LM과 TM을 통해 Analyzed Information

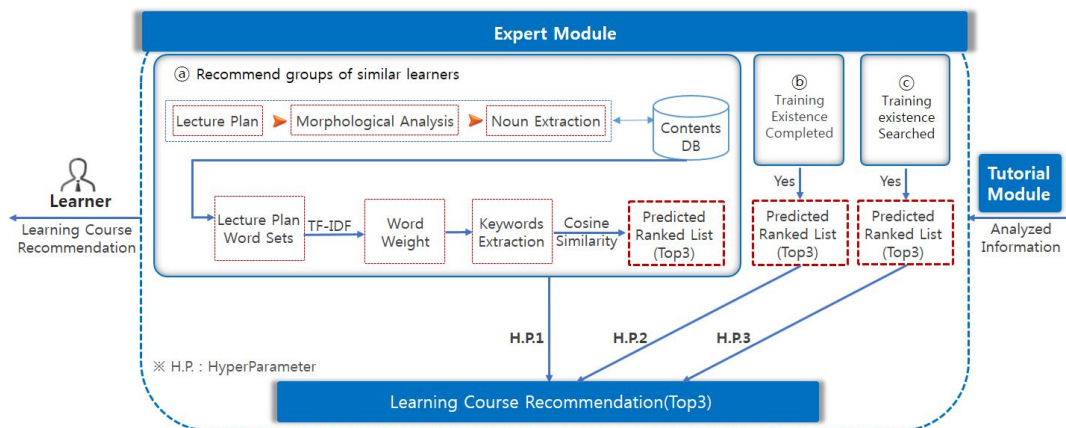


Fig. 2. Structure of the Expert Module

입력받아 Cosine Similarity 기법(2.2.2 참고)을 활용하여 추천한다. 추천 대상 학습자와 학습수준이 유사한 그룹을 찾고, 유사도가 높은 상위 3개의 학습코스를 선정하여 가중치 0.5를 부여한다. 만약, ⑤, ⑥, ⑦ 과정을 거치지 않았을 경우에는 가중치 1을 부여받고, 최종 추천결과가 된다. ⑤, ⑥, ⑦ 과정은 LM의 프로파일 DB는 있으나, TM의 학습에 대한 평가를 거치지 않은 경우에만 진행된다. 즉, 학습평가는 없고, LM의 Profile DB에 사용자가 수강했던 교육이 있거나, 검색했던 학습코스 정보가 있는 경우에만 진행된다. 각 과정에서는 최대 3개의 학습코스를 선정하고 가중치를 부여한다. Fig. 2는 전문가 모듈(EM)의 학습코스 추천 과정을 보여준다.

### 2.2.1 키워드 추출

교수자가 업로드한 강의계획안(LP, Lecture Plan)을 기반으로 강의안의 내용 중에서 스크립트를 추출하여 형태소 분석기[9]를 통해 각 단어 별로 형태소를 태그한다. 형태소가 태그된 단어들 중에서 명사 단어만 추출하여 Contents DB에 저장한다. 시스템은 DB로부터 단어를 호출하여 Word sets을 구성한다. 즉, 문서에서 많이 나타나는 명사가 그 문서를 대표하는 중요한 명사이므로 한 문서에 출현한 명사의 빈도수를 문서의 전체 길이로 정규화한 TF를 이용한다. 또한, 명사의 문서 간 가중치는 전체 문서집합 가운데 해당 명사가 나타난 문서가 어느 정도인지를 수치화한 IDF를 이용한다. 이는 특정 문서에만 나타나는 명사는 가중치를 높이고, 대부분의 문서에 나타나는 범용 명사는 가중치를 낮추는 역할을 한다. 위의 2가지 가중치를 함께 고려하기 위해 각 문서에 속한 명사의 최종 가중치는 TF-IDF로 결정되며, 이 값이 높은 명사들을 대표 키워드로 선택하게 된다. 자세한 계산식은 [Equation 1]과 같다.

$$Weight_{ij} = tf_{ij} \times \left\{ \log\left(\frac{N}{n}\right) + 1 \right\} \quad (1)$$

$Weight_{ij}$  : j번째 LP에서 i번째 단어의 가중치

$tf_{ij}$  : j번째 LP에서 i번째 단어의 빈도

$N$  : 전체 LP 수

$n$  :  $N$ 에서  $W_{ij}$  단어를 포함하고 있는 LP 수

즉, TF-IDF는 여러 강의계획안(LP<sub>1</sub>, LP<sub>2</sub>, ..., LP<sub>N</sub>)으로 이루어진 강의계획안 집합이 존재할 때, 각각의 LP에 포함

된 단어( $W_1, W_2, \dots, W_n$ )의 중요도를 산출하는 통계적 수치로써, LP 내 단어의 출현양을 나타내는 TF와 총 LP에서의 단어 출현비율의 역수를 취한 IDF를 활용한다. 따라서 단어의 중요도를 나타내는 TF-IDF 값은 특정 단어의 출현양이 많을수록 증가하지만, 모든 문서에서 빈번하게 사용되는 단어는 그 정도에 따라 값이 낮게 계산되는 것이다.

### 2.2.2 유사도 측정

TF-IDF가 높은 단어들은 문서의 Identity를 더 높게 반영한다고 할 수 있고, 이 값들이 유사하다면 유사한 문서라고 생각할 수 있다. 벡터간의 유사도를 측정하는 여러 방법이 있지만, 그 중 Cosine Similarity 기법을 이용했다. 코사인 유사도는 두 벡터간의 각도만을 이용하고 크기는 무시할 수 있기 때문에 선호도 정보가 없고, 컨텐츠 이용횟수 같은 암시적 정보로 벡터가 있는 경우에도 사용할 수 있는 장점이 있다[10]. 코사인 유사도의 계산식은 [Equation 2]와 같다.  $S$ 는 문서벡터와 관련된 값이고,  $h$ 는 검색어 벡터와 관련 값들이다.

$$\cos(\theta) = \frac{S \cdot h}{\|S\| \|h\|} = \frac{\sum_{i=1}^n S_i \times h_i}{\sqrt{\sum_{i=1}^n (S_i)^2} \times \sqrt{\sum_{i=1}^n (h_i)^2}} \quad (2)$$

즉, 추출된 키워드( $W_1, W_2, \dots, W_n$ )인 문서벡터와 LP<sub>N</sub> 검색어 벡터간의 유사도를 측정하기 위해 수식(2)를 적용한다. 유사도를 계산한 결과는 Table 3과 같다.

## 3. 실험 및 검증

### 3.1 실험설계 및 학습자 프로파일

학습자 특성과 학습자 수준을 고려한 PLCR 모델의 효과 분석을 위해, 실험집단(G1)과 비교집단(G2)을 선정하여 실험하였다. 두 집단 모두 경기도 소재 교육기관의 동일한 학습코스를 수료한 교육생으로 구성하였다. 실험에 참가한 집단의 Learning Proficiency(Mean values)을 비교해보면, 70점미만인 학습자가 G1 32%, G2 37%, 71점~89점 학습자 G1 51%, G2 47%, 90점 이상인 학습자 G1 16%, G2 15% 나타났다. 두 집단의 실험참가 날짜 및 인원은 다르지만, 학습수준 차이가 크지 않고, Gender(Female 100%), Age(대부분 30~39) 및 ICT

majors등이 유사한 것으로 보아 동일한 집단으로 가정할 수 있다.

동일한 두 집단을 대상으로, 먼저 학습자 능력수준 평가를 통한 Pre-Test를 수행한다. 다음으로, G2를 대상으로, 전문가 중심의 선행연구 추천[11]을 한다. G1을 대상으로, 본 연구에서 제안하는 기계학습 중심의 PLCR 추천한다. 마지막으로 PLCR모델의 효과분석을 위해 만족도 조사를 통한 Post Test를 수행하고, 결과를 비교분석한다.(3.3 참고) 이를 정리하면 Table 1과 같다. 교수자는 교실 당 1명씩 20명의 학습자를 통제한다. 학습자는 교육 센터의 특성상 미취업여성으로 평균 연령 40.5세이며, IT를 전공한 교육생이 61.3%, 대학이상의 교육을 받은 학습자가 92.6%, 추천시스템을 사용 해본 경험이 있는 학습자는 47.5%로 나타났다. G1 교실은 2017년 3월부터 7월까지 16주간(총 400시간) 매주 5회씩 1일 5시간, 80일 수업이 진행되었고, G2 교실은 2016년 3월부터 7월까지 16주간(총 400시간) 매주 5회씩 1일 5시간, 80일 수업이 진행되었다. 실험 기간 동안 매 수업 마다 운영현황을 파악하여 매주 보고서를 작성하여 통합 월간 운영 보고서를 작성하였다. 보고서에는 작성자, 일시, 과정명, 교사명, 운영 내용, 요구사항 등이 포함되어 있으며, 학습자 프로파일은 데이터베이스로 저장된다. 학습수준은 정량평가 70점, 정성평가 30점으로 총합 100점을 기준으로 하여 취득한 점수의 평균에 따라 프로파일을 생성했다[12]. 정량평가는 출석률 30점, 과제물 제출 20점, Final TEST 20점, 정성평가는 각 과정별 담당강사가 학습자의 교육 참여도 30점을 합산하여 평가점수를 도출한다. 학습기간 동안 교육 참여율 및 과제 미제출 등으로 교육을 수료하지 못한 교육생은 0점으로 처리했다. 또한, Delphi 기법을 활용하여, 실험대상에 적합한 IT전문교육 16개의 학습코스를 도출 하였다[11].

Table 1. Experimental Design

Group	Pre-test	Experiment	Post-test
G1	T1	X1	T2
G2		X2	

G1 : Experiment Group(in 2017, 4month(400h))  
 G2 : Comparison Group(in 2016, 4month(400h))  
 T1 : Learning Proficiency Evaluation(100points)  
 T2 : Effectiveness Evaluation(Survey)  
 X1 : Proposed method(PLCR model)  
 X2 : Previous method[11]

### 3.2 추천모델 비교 분석

선행연구[11]의 추천모델을 적용한 결과는 Table 2와 같다. 즉, 비교집단(G2) 180명의 능력수준 평가를 통해 학습수준 프로파일을 생성한 후, Pearson Correlation Coefficient을 이용해, 학습자 유사그룹을 형성한다. 그리고, 학습코스 16개를 학습 스킬간의 연관성을 고려하여, 전문가 그룹이 부여한 연관성 정도의 평균값과 학습자들간의 유사도값을 고려해 학습코스를 예측한다. 예측한 학습코스 중에서 상위 3개의 추천 목록을 학습자에게 제공한다.

Table 2. Learning courses by using the Human Recommendation(Previous method)[11]

Division	Recommendation Course			similarity	mean	
	14	9	10			
Similar Learner Group	51	8.74	8.74	8.74	0.87	60.00
	52	10.74	1.53	1.53	0.92	68.34
	74	1.51	10.57	1.51	0.90	78.34
	86	10.27	1.46	1.46	0.88	68.34
	95	8.68	8.68	0.00	0.86	90.00
	96	0.00	9.34	9.34	0.93	80.00
	107	7.78	-1.55	7.78	0.93	81.67
	121	9.20	0.00	0.00	0.92	80.00
	127	9.03	9.03	0.00	0.90	90.00
	165	9.10	9.10	1.42	0.91	90.00
Predictive value	78.35	76.82	73.53			
Ranking	1	2	3			

본 연구에서 제안하는 ITS 기반의 개인화 학습코스 추천(PLCR)모델을 통해서 예측된 상위 3개의 추천목록 결과는 Table 3와 같다. 본 연구의 실험집단(G1) 80명 모두 교수모듈(TM)을 통해 Learning Proficiency 분석하고, 전문가 모듈(EM) ①,②,③ 과정[그림 2] 중 ① 과정을 통해 추천목록을 생성했다. 즉, 강의계획안(LP)의 핵심 키워드 추출한 DB(2.2.1 참고)를 입력받아 Cosine Similarity기법(2.2.2 참고)을 활용하여 유사도가 높은 상위 3개의 학습코스를 추천한다.

결론적으로, 전문가 중심의 기존 추천모델과 본 연구에서 제안하는 PLCR 추천모델을 비교해본 결과, 예측된 추천 학습목록 중 상위 3개의 추천코스가 서로 다른 것으로 나타났다. 추천방식에 따라 서로 다른 예측 결과가 나온 실험결과를 바탕으로, PLCR 모델의 효과성을 분석한다.

Table 3. Learning courses by using the PLCR(Proposed method)

Division	Recommendation Course			similarity	mean	
	9	2	11			
Similar Learner Group	15	9.74	8.40	8.12	0.91	62.00
	19	1.30	1.50	1.38	0.95	70.44
	25	8.50	9.07	1.41	0.94	79.00
	34	1.42	1.36	1.47	0.89	71.34
	51	7.68	8.64	6.04	0.88	90.50
	58	9.31	7.34	8.31	0.94	84.00
	62	-1.00	-1.05	1.77	0.91	87.04
	73	0.02	0.10	0.11	0.97	83.14
	77	9.12	9.01	9.49	0.81	91.50
80	9.14	8.91	5.45	0.90	91.80	
Predictive value	78.82	77.85	75.41			
Ranking	1	2	3			

3.3 PLCR 추천모델 효과성 분석

선행연구[11]에서 개발된 설문지를 바탕으로 80명의 학습자를 대상으로 자기 기입식 설문조사 방법을 실시하여 PLCR 모델의 효과를 분석하고 검증했다. 만족도 조사를 실시한 결과, 80명의 학습자 중 72명이 응답하여 90%의 응답률을 보였다. 이중 의미 없는 답변을 한 2명을 제외한 70명의 만족도 조사결과를 바탕으로 분석하였다. 교육 만족도 분석을 위해 설문항목별 평균, 표준편차, 왜도, 첨도값을 계산하여 정량적 분석을 실시하였다[13]. 평균(Mean)은 각각의 설문항목에 대해 매우만족, 만족, 보통, 불만족, 매우불만족의 리커트 5점 척도를 활용하여 측정하였으며 3점 이상이면 설문항목에 대해 만족하는 것으로 본다[14]. 표준편차(Standard deviation)는 통계집단의 계량적 특성값에 관한 산포도를 나타내는 도수 특성값을 말한다. 표준편차가 0일 때는 관측값의 모두가 동일한 크기이고, 표준편차가 클수록 평균에서 떨어진 값이 많이 존재하는 것으로 보는데, 표준편차 값이 3이상이면 유의미한 결과가 될 수 없다. 왜도(skewness)는 분포의 비대칭의 정도, 즉 분포가 기울어진 방향과 그 기울어진 정도를 나타내는 척도이다. 왜도 계산값이 0이면 좌우 대칭분포를 가지고, 0보다 작으면 음(negative)의 왜도를 가지고, 0보다 크면 양(positive)의 왜도를 가진다. 왜도를 계산한 절대값이 클수록 분포의 비대칭 정도가 커진다는 것을 알 수 있는데 절대값이 3보다 크면 유의미한 결과가 될 수 없다[15]. 첨도(kurtosis)는 데이터의 분포

가 중심 경향값을 중심으로 집중적으로 분포되어 있는 정도 혹은 분포의 뾰족한 정도를 나타내는 척도이다. 정규분포의 첨도는 0이고, 첨도값이 0보다 크면 분포의 모양이 정규분포 보다 더 뾰족한 급첨(leptokurtic), 0보다 작으면 정규분포보다 더 평평한 평첨(platykurtic)이 된다. 첨도를 계산한 절대값이 8이상이면 유의미한 결과가 될 수 없다[16]. 기존의 전문가 그룹을 활용한 사람중심의 추천모델[11]이 가지고 있던 비효율적 한계점을 극복한 PLCR 모델의 만족도 설문 조사 결과, Mean 항목은 3.92, Standard deviation 항목은 0.805, Skewness 항목은 -0.335, Kurtosis 항목은 -0.313을 나타냈으며, 설문 응답 결과 모두 정규분포를 가지는 유의미한 결과를 나타냈다. 이것은 PLCR 모델이 학습자에게 교육만족도 및 학습효과를 향상 시킬 수 있는 요인을 제공한다는 것을 알 수 있으며, 설문 항목별 조사결과는 Table 4과 같다.

Table 4. Survey Result for each measured factor of the PLCR(Proposed method)

	Item	Sample No	Mean	Standard deviation	Skewness	Kurtosis
Accuracy	AC1	70	3.871	0.659	-0.510	-0.299
	AC2	70	3.799	0.671	-0.497	-0.301
	AC3	70	4.113	1.007	-0.401	-0.407
Novelty	NT1	70	3.852	1.010	-0.607	-0.403
	NT2	70	4.102	0.807	-0.402	-0.462
	NT3	70	3.881	0.641	-0.454	-0.573
Satisfaction	SF1	70	4.102	1.016	-0.203	-0.120
	SF2	70	4.192	1.071	-0.120	-0.176
	SF3	70	4.094	0.980	-0.200	-0.300
Self-Perfient	SR1	70	4.002	0.841	-0.204	-0.378
	SR2	70	3.891	0.762	-0.397	-0.307
	SR3	70	3.978	0.704	-0.217	-0.210
Usefulness	UF1	70	4.109	0.710	-0.308	-0.240
	UF2	70	3.942	0.607	-0.307	-0.207
	UF3	70	3.891	0.601	-0.201	-0.314
Total	15	70	3.987	0.805	-0.335	-0.313

또한, 선행연구[11]와 본 연구의 제안 추천모델 설문 결과를 비교해보니, PLCR 모델이 Mean 항목은 0.145 높게 나타나 리커트 5점 척도 중 만족항목에 좀 더 가까웠

다. Standard deviation 항목은 0.105 낮게 나타나 평균에서 떨어진 값이 선행연구 보다 적은 것으로 나타났다. Skewness 항목도 0.051 낮게 나타나 만족도 조사 결과 값에 대한 비대칭의 정도가 적었고, Kurtosis 항목도 0.059 낮게 나타나 데이터의 분포가 평균값에 집중적으로 분포되어 있다는 것을 알 수 있었다. 제안하는 PLCR 모델이 학습자에게 좀 더 높은 만족결과를 나타냈으며, 향후 학습자에게 학습코스를 추천할 때 보다 만족도 높은 서비스를 제공할 수 있다는 것을 알 수 있다.

Fig 3는 학습코스 추천모델의 만족도 평가결과를 선행연구와 비교하여 그래프로 표현한 것이다. 세부 항목별로 비교해보면, Novelty 항목의 세부문항들에 대해 만족도가 가장 높게 나타났다. 이는 그동안 국내·외에서 충분히 다루어지지 않았던 기계학습 중심의 맞춤형 학습코스를 추천해주는 PLCR모델이 학습자들에게 새롭고 흥미로움을 제공하는 결과임을 알 수 있다. 세부 문항별로 비교해보면, NT1(Satisfy my curiosity of the new learning courses), SR2(Provide right learning courses that considered to my interest), AC2(Helpful to make decisions for selecting learning courses)문항의 만족도 차이가 큰 것으로 나타났다. 이는 기계학습 중심으로 학습코스를 추천하는 PLCR 모델에 대해 학습자들이 Novelty, Self-Referent, Accuracy에 대한 만족도가 특히 더 높다는 것을 알 수 있는 유의미한 결과를 제공한다.

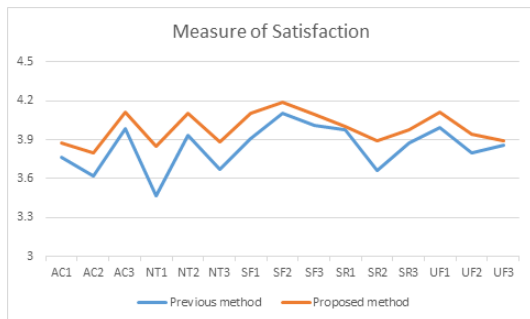


Fig. 3. Compare Proposed method(PLCR) and Previous method

#### 4. 결론

학습스킬 간의 연관성을 고려한 학습코스를 추천하기 위해 선행연구[11]에서는 전문가 집단을 활용하다 보니,

많은 시간과 비용이 소모되는 자원의 비효율적 한계점이 있었다. 이를 극복하기 위해, ITS 기반의 개인화 학습코스 추천 모델을 제안하고, 교육효과를 분석하고자 했다. 추천하는 학습코스는 학습자 개개인의 선행 학습 정도와 수준을 고려한 맞춤형 추천을 위해, 학습자 프로파일을 분석하고, TF-IDF를 이용해 단어별 가중치를 계산하였다. 고빈도 단어를 추출하여 Cosine Similarity 기법을 이용해 벡터간의 유사도를 측정하였으며, 제안하는 추천모델의 효과성을 검증하기 위한 실험을 설계하였다. 기 개발된 문항지를 바탕으로 만족도 조사를 실시하였으며, 항목별 평균, 표준편차, 왜도 첨도값을 계산하였다. 또한 정성 및 정량적 분석을 실시하여 제안 모델의 효과성을 검증하고자 했다. 그 결과, 제안하는 추천모델의 정확성, 새로움, 자기참조, 유용성에서 높은 만족도를 보였다. 특히, 학습코스를 추천하는 모델에 대한 호기심과 학습코스를 결정할 때 학습자에게 큰 도움이 된다는 결과가 나타났다.

본 연구는 학습코스 선정에 많은 어려움과 시행착오를 겪는 사용자들에게 수준별 학습코스를 제공할 수 있다는 장점이 있으며, 그동안 국내 및 국외에서 충분히 다루어지지 않았던 기계학습 중심의 맞춤형 학습코스 추천 모델을 개발했다는 점에서 의미가 있다. 하지만, 실험 데이터에 성별 및 지역적 한계는 존재한다. 성별 및 학습수준, 지역적인 특성에 따라 평가내용이 달라 질 수 있으나, 연구에 사용된 데이터의 경우, 경기도에 거주하는 여성만을 대상으로 실험하다보니, 연구결과에 한계점이 있다. 향후 연구에서는 한계점을 보완하고, 성능을 개선하여 사용자에게 좀 더 유용한 사용자 맞춤형 추천시스템이 될 수 있도록 추가적인 연구가 필요할 것으로 생각된다.

#### REFERENCES

- [1] Naser, S. A., Ahmed, A., Al-Masri, N. & Abu Sultan, Y.(2011). Human Computer Interaction Design of the LP-ITS: Linear Programming Intellignet Tutoring Systems, *International Journal of Artificial Intelligence & Applications (IJALA)*, 2(3), 60-70. DOI : 10.5121/ijaia.2011.2306
- [2] Wang, Z., Yu. X. & Feng. N.(2014). An improved collaborative movie recommendation system using computational intelligence, *Journal of Visual Languages and Computing*, 25(6), 667-675.



- DOI : 10.1016/j.jvlc.2014.09.011
- [3] S. S. A. Naser.(2012). A Qualitative Study of LP-ITS: Linear Programming Intelligent Tutoring Systems, *International Journal of Computer Science & Information Technology*, 4(1), 209-220.  
DOI : 10.5121/ijcsit.2012.4116
- [4] S. Y. Pi.(2015). Educational Utilization of Smart Devices in the Convergence Education Era, *Journal of digital Convergence*, 13(6), 29-37.  
DOI : 10.14400/JDC.2015.13.6.29
- [5] E. A. Rashid., S. B. Patnaik., & V. C. Bhattacharjee.(2014). Machine Learning and Software Quality Prediction: As an Expert System, *Journal of Information Engineering and Electronic Business*, 6(2), 9-27.  
DOI : 10.5815/ijeeb.2014.02.02
- [6] C. R. Beal., I. Arroyo., P. R. Cohen., B. P. Woolf., & C. R. Beal.(2010). Evaluation of AnimalWatch: An intelligent tutoring system for arithmetic and fractions. *J. Interact. Online Learn.*, 1(9), 64-77.
- [7] Baker, R. S.(2016). Stupid tutoring system, intelligent humans. *International Journal of Artificial Intelligence in Education*, 26(2), 600-614.  
DOI : 10.1007/s40593-016-0105-0
- [8] C. M. Chen. & C. J. Chung.(2007). Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle. *Computers & Education*. 51(2), 624-645.  
DOI : 10.1016/j.compedu.2007.06.011
- [9] K. S. Shim.(2014). Syllable-based Probabilistic Models for Korean Morphological Analysis, *Journal of KIISE*, 41(9), 642-651.  
DOI : 10.5626/JOK.2014.41.9.642
- [10] Linden, G., Smith, B., & York, J.(2003). Amazon.com recommendations: item-to-item collaborative filtering, *IEEE Internet Computing*, 7(1), 76-80.  
DOI : 10.1109/MIC.2003.1167344
- [11] W. Han, J. C. Jo, H. S. Ji, & H. S. Lim (2016). A collaborative recommender system for learning courses considering the relevance of a learner's learning skills. *Cluster Computing*, 19(4), 2273-2284.  
DOI : 10.1007/s10586-016-0670-x
- [12] K. B. Kim. & H. J. Cho.(2015). A Study on Smart Teaching Plan Production System Combined Education Profiling, *Journal of digital Convergence*, 13(3), 185-191.  
DOI : 10.14400/JDC.2015.13.3.185
- [13] Golding, J. F.(2006) Predicting individual differences in motion sickness susceptibility by questionnaire, *Personality and Individual Differences*, 41(2), 237-248.  
DOI : 10.1016/j.paid.2006.01.012
- [14] Dong Y., Sun Z. & Jia H.(2006). A Cosine Similarity-based Negative Selection Algorithm for Time Series Novelty Detection, *Mechanical Systems and Signal Processing*, 20(6), 1461-1472.  
DOI : 10.1016/j.ymsp.2004.12.006
- [15] Abuhay, T. M., et al.(2018). Analysis of publication activity of computational science society in 2001-2017 using topic modelling and graph theory, *Journal of Computational Science*, 26, 193-204.  
DOI : 10.1016/j.jocs.2018.04.004
- [16] Randall, W. Engle.(2016). Working Memory Capacity as Executive Attention. *Current Directions in Psychological Science*, 11(1), 19-23.  
DOI : 10.1111/1467-8721.00160

한 지 원(Han, Ji Won)

[정회원]



- 2011년 2월 ~ 현재 : 고려대학교 컴퓨터학과 박사수료
- 2016년 11월 ~ 현재 : 경기도일자리재단 청년사업팀장
- 관심분야 : 컴퓨터교육, Cognitive Engineering, Convergence

· E-Mail : ssukiya77@korea.ac.kr

조 재 춘(Jo, Jae Choon)

[정회원]



- 2018년 2월 : 고려대학교 컴퓨터학과 (이학박사)
- 2018년 2월 ~ 현재 : 고려대학교 정보창의교육연구소 연구교수
- 관심분야 : 컴퓨터교육, EDM, AI in Education

· E-Mail : jaechoon@korea.ac.kr

임 희 석(Lim, Heui Seok)

[정회원]



- 1997년 2월 : 고려대학교 컴퓨터학과 (이학박사)
- 2008년 3월 ~ 현재 : 고려대학교 컴퓨터학과 교수
- 관심분야 : 자연어처리(NLP), 뇌 신경 언어 정보처리

· E-Mail : limhseok@korea.ac.kr