

동영상 기반 학습 환경에서 머신러닝을 활용한 행동로그의 학업성취 예측 모형 탐색

Exploration of Predictive Model for Learning Achievement of Behavior Log Using Machine Learning in Video-based Learning Environment

이정은[†] · 김다솜^{††} · 조일현^{†††}
Jungeun Lee[†] · Dasom Kim^{††} · Il-Hyun Jo^{†††}

요 약

동영상 강의 중심의 온라인 학습 형태가 보편화 되고 지속적으로 증가됨에 따라 다양한 교육방법을 적용한 동영상 기반 학습 환경도 학습 효과성을 높이기 위해 변화, 발전하고 있다. 온라인 학습 환경에서의 교육 효과성 측정을 위해 학습자 로그 데이터가 대두되었으며, 학습자 맞춤형 학습 처방을 위해 로그 데이터의 다양한 분석 방법이 중요하다. 이를 위해 본 연구에서는 동영상 기반 학습 환경에서의 학습자 행동 데이터 분석, 머신러닝 기법에 따른 학업성취 예측을 실험을 통해 분석하였다. 분석 결과 각 모델에서 공통적으로 동영상 탐색과 코멘트 작성과 같은 상호작용 행동, 학습자 주도적 학습 행동이 학업성취를 예측하였다. 연구 결과를 토대로 동영상 학습 환경 설계에 있어 시사점을 제공하였다.

주제어: 동영상 학습, 행동로그, 머신러닝, 예측모형

ABSTRACT

As online learning forms centered on video lectures become more common and constantly increasing, the video-based learning environment applying various educational methods is also changing and developing to enhance learning effectiveness. Learner's log data has emerged for measuring the effectiveness of education in the online learning environment, and various analysis methods of log data are important for learner's customized learning prescriptions. To this end, the study analyzed learner behavior data and predictions of achievement by machine learning in video-based learning environments. As a result, interactive behaviors such as video navigation and comment writing, and learner-led learning behaviors predicted achievement in common in each model. Based on the results, the study provided implications for the design of the video learning environment.

Keywords: Video-based Learning, Behavior Logs, Machine Learning, Predictive Model

1. 서론

ICT(Information and Communication Technology) 기반 인프라 발달로 웹 기반의 온라인 학습, 이러닝(e-learning)이라고 불리는 학습 형태가 지속적으로 증가하고 있다 [1]. 동영상 강의 중심의 온라인 학습은 MOOC, Blended Learning, Flipped Learning 등과 같이 다양한 형태로 학습 효과성을 높이기 위해 발전하고 있다. 하지만 온라인

환경에서 의미 있는 학습자 정보를 추출하고 학습 효과성을 평가하는 것에는 어려움이 존재한다[2][3]. 이에 온라인 기반 학습을 통해 학습자들로부터 생성되는 디지털 데이터가 주목되었고, 이를 활용한 학습 문제 해결과 평가를 시행하는 방법 등이 도입되고 있다[2][4].

학습분석학 맥락에서 온라인 학습 환경에서는 학습자들의 학습 활동이 실시간으로 기록되어 분석 가능한 형태의 로그 데이터로 저장되며, 학습 시점에 저장되는 데이

[†]정 회 원: 이화여자대학교 교육대학원 교육공학·HRD전공 석사과정

^{††}정 회 원: 이화여자대학교 일반대학원 교육공학과 석사

^{†††}정 회 원: 이화여자대학교 교육공학과 교수(교신저자)

터는 해당 시점의 학습자 심리와 행동 반응 등을 포함하기 때문에 로그 데이터 분석을 통해 개별 학습자들에게 맞춤형 학습 처방을 제공할 수 있다[5]. 더불어, 학습자의 동기 부여를 위한 개입 및 의사결정과 학습 촉진을 위해서도 로그 데이터를 통한 다양한 분석 방법이 중요하다.

동영상 학습 환경의 유용성과 학습자와 매체 간 상호작용은 학습에 영향을 미치기에[6][7][8][9], 학습자의 상호작용과 행위를 유발하는 어포던스의 속성이 중요하다[10]. 행위유발성을 의미하는 어포던스는 주체의 행동에 영향을 주는 객체의 속성이라고 할 수 있다[11]. Norman은 물리적 행동인 실제 어포던스와, 인지적 행동의 지각된 어포던스를 구분하였다[12]. 지각된 어포던스는 물리적 환경이나 매체가 인지적 자극을 통해 행동을 유도하는 것으로 학습자 태도 형성에 영향을 줄 수 있는 요소이다[10][12]. 따라서 동영상 학습 환경 설계에 있어 지각된 어포던스를 높일 수 있는 학습자 중심의 분석이 필요하다[10][13]. 동영상 학습 상황에서 학습자의 동영상 탐색 행동이 학업성취에 영향을 미친다는 선행연구와 함께 학업성취를 예측하는 다양한 학습자 행동 요소를 도출할 필요성이 제기되었다[14]. 따라서 동영상 학습 환경에서 어포던스에 기반한 학습자의 세부 행동 요소가 학업성취에 어떤 관계성을 갖는지 연구할 필요가 있다.

학습자 행동 데이터는 그 정보가 방대하고 다양한 학습자들의 행동에 대한 규칙성을 찾기가 어려워 특정 행동과 학업성취에 대한 관계를 찾기가 어렵지만, 머신러닝 기법을 활용하면 행동 로그에 따른 학업성취를 예측할 수 있다. 머신러닝은 특정한 의도 없이 수집된 다양한 형태의 자료 분석이 가능하고[15], 학습 알고리즘을 만들어 패턴을 파악하고 예측하는 것에 초점을 맞추고 있어[16] 학습자 행동에 따른 학업성취를 분석하기 적절하다. 예를 들어, 랜덤 포레스트 모델은 비모수적 방법으로 설명 변수가 많고 변수 간 상호작용이 복잡한 자료에서의 예측률이 높아[17][18][19], 학습자 행동 데이터를 분석하기에도 적절하다. 직관적으로 해석하기 어렵다는 단점이 있지만[20], 랜덤화로 예측의 일반화 성능이 강하고 각 변수 중요도에 순위를 매겨 설명력을 더한다는 장점이 있다[17]. 랜덤 포레스트뿐만 아니라 각 머신러닝 기법은 모델에 따라 장단점이 있다. 예측률과 주요 변수를 예측하는 부분에서도 조금씩 차이가 있어 단점 보완과 해석을 위해 비교 분석이 필요하다. 각 모델에서 도출하는 공통된 주요 변수는 추후 교수설계에서 중요한 요소로 고려할 수 있다.

본 연구에서는 동영상 기반 학습 환경에서 학습자들의 인지적, 행동적 과정에 따라 발생하는 행동 로그 데이터

를 분석하고, 머신러닝 기법에 따라 학습자 행동 로그가 학업성취를 어떻게 예측하는지 분석하고자 한다. 분석 결과를 통해서 학습에 영향을 미치는 학습자의 공통적 행동 요인을 도출하고 실제 동영상 학습 환경 설계에서의 적용점을 모색하고자 한다.

2. 이론적 배경

2.1 동영상 기반 학습

동영상 기반 학습은 동영상을 활용한 학습으로, 교육 맥락에서의 동영상은 문자와 그림, 소리 등의 시청각 자료를 통합하여, 단일 문자나 정적인 그림만 제공함으로써 인해 내용 표현에 어려움이 있는 부분까지 전달해 주는 매체를 의미한다[21]. 이는 학습자의 집중과 몰입을 유발하고 인지적 실재감과 만족도를 향상시키는 효과적인 교수 학습 매체라고 할 수 있다[22][21].

MOOCs, OCW(Open Course Ware), YouTube 등 개방형 학습 모델의 확산으로 동영상 기반 학습의 양적·질적 향상과 더불어 온라인 플랫폼을 통한 학습자들의 다양한 로그 정보 수집도 가능하게 되었다. 동영상 기반 교육 콘텐츠를 제공하는 학습 환경은 디지털 테크놀로지를 기반으로 기술적 상호작용이 수반된다. 동영상과 학습자 간 양방향적 상호작용에 대한 연구들은 크게 도구적 행동과 심리적 행동 연구로 나뉜다[23]. 도구적 행동은 키보드 또는 마우스와 같은 인공물을 매개로 객체를 변형 또는 조작하는 표면적 행동을 의미하고, 심리적 행동은 시각적 주의분산을 예로 들 수 있다.

동영상 기반 학습은 주로 교수자와 분리된 환경에서 진행되기에 학습자의 동기와 의지에 영향을 많이 받는다[24]. 또한 동영상 기반 학습 상황에서는 학습자와 학습 콘텐츠와 플랫폼과 같은 매체와의 상호작용에 의해 학습 활동이 유발되고, 학습 목표를 달성하게 된다. 이를 위해서는 학습자 행동을 유발할 수 있는 환경 설계가 필요하다[25][26].

2.2 지각된 어포던스

Gibson에 의해 제안된 개념인 어포던스는 객체와 주체 간 상호작용에서 발생하는 행위유발성을 의미하는 용어로 객체가 가진 환경적 속성으로 인해 주체의 행동을 유발하는 속성을 의미한다[11]. 이후 Norman이 논한 지각된 어포던스의 개념은 인간과 컴퓨터와의 상호작용에 관한 것으로 사물의 지각된 특성 또는 환경이 어떻게 사용될 것인지에 대한 실제적 특성이다[12]. 지각된 어포던스는 도구나 환경의 물리적 형태가 가진 내재적 특성이다.

인간의 시각적 인지를 통해 도구를 활용하거나 해당 정보를 암시하거나 이해할 수 있게 하기에 동영상 기반 학습 환경과 학습자 간 상호작용의 행동 유발 측면에서 중요하게 다루어지고 있다[25][26][12]. 학습자는 지각된 어포던스를 통해 도구나 환경을 어떻게 다루면 될 것인지에 대해 생각할 수 있고, 그 심리에 의해 어떤 행동을 결정할 수 있다[25][27][12].

지각된 어포던스는 학습자 중심의 학습 상황과 다양한 멀티미디어 환경에서 인간과 환경의 상호작용 맥락에서 중요한 요소로 여겨진다[28]. 동영상 학습 환경에서 학습자와 학습 플랫폼 및 콘텐츠와의 상호작용으로 학습이 이루어지기에 지각된 어포던스는 중요하다[26]. 특히 동영상 학습 환경에서 사용자 인터페이스가 학업성취에 주요한 영향을 미치기에[29], 동영상 환경의 지각된 어포던스를 높이는 것은 학습 효과성을 높이며 학업성취에 기여할 수 있다[26].

2.3 행동로그

학습자 디지털 데이터는 주로 LMS와 VLE와 같은 학습 환경에서 수집할 수 있다. 교수설계자는 디지털 데이터 분석을 통해 학습자 특성과 학습 패턴 등을 도출하고, 이를 학습 촉진을 위한 적절한 개입 또는 의사결정을 하는데 활용할 수 있다.

온라인 학습 환경에서 수집할 수 있는 학습자 데이터는 학습자 행동로그, 맥락, 제스처, 표정, 생리심리반응과 같은 생체 데이터 등 매우 다양하다[30]. 행동로그 데이터는 학습자의 클릭, 재생, 글쓰기와 같은 행동을 그대로 나타내는 데이터로, 교수자와 학습자가 분리되어 직접 관찰하기 힘든 학습자의 온라인 학습 행동 양상을 간접적으로 확인할 수 있다는 점에서 가장 유망한 자원으로 여겨진다[31]. 학습자와 학습 환경과의 상호작용을 통해 발생하는 행동은 인지적 정보처리 과정과 의도성을 내포하며, 특정 행동에 따른 로그 데이터는 환경에 대한 행위 유발성을 표상한다. 학습자 행동 로그 데이터를 수집하고 분석하는 것은 특정 학습 환경에 따른 학습자 행동을 규명할 수 있게 한다[32]. 동영상 기반 학습 상황에서 행동로그는 보다 직접적으로 학습자 행동을 보여주는 지표로 볼 수 있으며, 행동 유형과 수준을 포함하는 행동로그 데이터를 통해 지각된 어포던스를 구체화할 수 있다.

2.4 머신러닝

동영상 기반 학습과 같은 디지털 학습 환경에서 생성되는 학습자 데이터는 정형 또는 비정형의 다양한 형태

로 수집된다. 교육 분야에서의 머신러닝은 대학을 중심으로 로그 데이터와 같은 비정형 데이터를 분석하여 학습 처방을 제공하거나 교수설계에 따른 학업성취 예측을 위해 활용되고 있다.

머신러닝은 변수 간 주요 관계 파악, 데이터 학습을 통한 패턴 평가 및 예측에 초점을 맞추기에[33] 다양한 문제 상황에 유연하게 적용할 수 있다. 사례 수 보다 변수 수가 더 많은 wide data는 전통적인 통계학에서는 분석하기 힘들다는 한계가 있지만, 머신러닝에서는 wide data 분석과 더불어 특정한 의도 없이 수집된 다양한 형태의 자료도 분석 가능하다[15]. 머신러닝은 비선형(nonlinear) 모형과 변수들 사이의 복잡한 상호작용 또한 분석할 수 있다는 장점이 있다[34].

온라인 학습과 웹 기반 학습자 관리 시스템이 보편화된 현대 교육환경에서 머신러닝 기법은 교육 연구에 더욱 유용하게 활용할 수 있다. 학습자의 디지털 로그 파일이 빅데이터로 수집되기에 전통적인 통계기법으로는 분석이 어렵기 때문이다. 소수의 변수만을 이용한 모형 설정뿐만 아니라, 새로운 변수를 탐색할 수 있다는 점에서 머신러닝 기법은 학습분석학 연구에 있어서도 유용하게 활용할 수 있다[35]. 교육 연구에서 머신러닝은 지도학습과 비지도학습으로 나누어 활용되고 있다. 지도학습 기법에서는 별점회귀모형과 의사결정나무, 랜덤 포레스트 등이 많이 활용된다. 비지도학습 기법은 K-평균 군집이나 위계적 군집 분석, 텍스트 마이닝 위주로 활용된다[35].

머신러닝 중 분류 문제를 다루는 주요 모델 별 장단점은 <표 1>과 같다.

3. 연구방법

3.1 연구대상

본 연구는 수도권 소재 대학교 학부생 총 78명을 대상으로, 남자 33명(42.31%), 여자 45명(57.69%)으로 진행되었다. 실험에 참여한 연구대상자들은 인문계열 36명(46.15%), 자연계열 42명(53.85%)으로 구성되었다.

3.2 연구도구

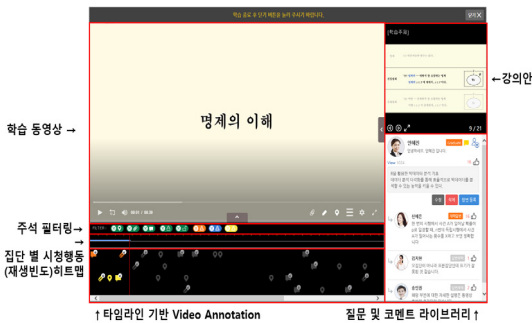
3.2.1 동영상 학습 플레이어

본 연구에서 사용된 동영상 학습 플레이어는 소셜 클라우딩을 통한 동영상 기반 학습을 중심으로 양방향적 상호작용이 가능한 동영상 기반 학습[36][37]을 기술적으로 지원하기 위해 설계 및 개발되었다[38]. 학습자는 본

표 1. 머신러닝 주요 분류 모델 장단점

모델	장점	단점
K-근접이웃	이해가 쉽고 단순 데이터 가정 필요 없음	비교적 느림
의사결정 나무	간단하고 해석 가능 변수 중요도 파악 가능 전처리 과정 적음 이상치 영향력 적음	과적합 가능성 데이터에 따른 성능 차 다소 낮은 모델 예측력
인공신경망	이론적 성능 높음 모든 변수 처리 가능 변수 중요도 파악 가능	학습 과정이 김 결과 해석이 어려움 최적 모델 결정 어려움
서포트 벡터 머신	높은 예측력 최적값의 이론적 보장 이진분류에 강함 계산량 다소 적음	해석 어려움 연속형 데이터만 가능 표준화 필요 사용자 파라미터 민감 커널 선택 민감
랜덤 포레스트	높은 성능 변수 중요도 파악 가능 표준화 과정 불필요	설명력 약함 과적합 가능성

동영상 학습 플레이어를 통해 기본 영상 제어와 코멘트, 북마크, URL 링크 첨부 기능 등을 활용하여 학습 활동을 기록할 수 있고, 목록화된 형태로 열람할 수 있다. 모든 연구대상자들은 독립된 실험환경에서 해당 동영상 학습 플레이어를 이용하여 자기주도적 학습을 진행하였다. 동영상 학습 플레이어 인터페이스 화면은 [그림 1]과 같다 [38].



[그림 1] 동영상 학습 플레이어 인터페이스[38]

3.2.2 학습자 행동로그 데이터

학습자 행동로그 데이터는 지각된 어포던스의 직접적인 지표 변인으로서 앞서 언급한 동영상 학습 플레이어를 통해 수집되었다. 동영상 학습 플레이어는 재생, 동영상 탐색, 주석 생성, 주석 탐색, 외부환경 조정 활동의 5가지 기능을 지원한다. 플레이어 내 구동버튼은 총 21개로 학습활동에 따른 버튼 별 기능은 <표 2>와 같다.

표 2. 동영상 학습 플레이어의 구동 버튼

학습활동	종류	구동버튼 기능
재생	2	재생/정지, 현재 시청 위치
동영상 탐색	3	재생표시줄, 10초 뒤로, 강의안 넘기기
주석 생성	6	링크 달기, 코멘트 작성/등록, 북마크 등록, 북마크 목록 (동료 학습자 학습 기록 내)코멘트 작성, 코멘트 등록
주석 탐색	3	주석 필터, 등록된 주석, 좋아요
외부환경 조정	7	(학습동영상 및 강의안) 화면 키우기/줄이기, 전체화면 음량, 설정, 플레이어 종료

본 실험에서는 외부 웹페이지 방문을 제한하였기 때문에 링크 달기 기능은 제외하고 동영상 학습 플레이어를 활용하였다.

3.2.3 학업성취

본 연구의 목적은 동영상 학습 상황에서의 학습자들의 지각과 인지적 과정에 따라 발생하는 학습 행동에 대한 결과가 학업에 미치는 영향을 분석하는 데 있다. 동영상 기반 학습 상황에서 학습자들의 지각된 어포던스가 학업 성취에 미치는 영향을 규명하고자 하며, 시험 결과를 동영상 학습의 결과로서 학업성취를 측정하는 지표로 활용하였다.

시험은 동영상 학습 전과 후를 기준으로 사전시험과 사후시험으로 나누었으며, 동형검사로 진행하였다. 시험

내용은 ‘수리논리학의 기초 증명’에 대한 것으로 시험 문항은 수학교육과 교수와 교육공학과 전문가가 협력하여 개발하였다. 본 시험은 총 44문항, 44점 만점으로 구성하였으며, 시험 문제 풀이를 위해 총 15분의 제한시간이 주어졌다. 모든 연구대상자는 제한시간 내에 시험을 응시하고, 모르는 문항은 공란으로 남겨두도록 안내받았다.

3.3 연구절차

본 연구는 크게 준비와 실험으로 구분되며, 준비 단계에서 생명윤리위원회의 심의와 결과에 따라 진행하였다 [38]. 연구대상자는 2주간의 온라인 공고를 통해 수도권 소재 대학에 재학 중인 학부생을 모집하였다. 학교 홈페이지 게시판, 학교 커뮤니티 및 학생회 메신저를 통해 공고물을 배포하고, 연구 희망 참가자들의 일정을 조율하여 회신 후 최종적으로 실험 일정을 배정하였다. 실험 배정이 완료된 참가자에게는 주의사항, 실험 장소 안내, 참여 확인 내용을 제시하고 관련 공지 사항을 설명하였다. 실험에 앞서 실험과 동일한 연구 환경에서 모의실험을 실행하고, 진행과 관련된 세부사항을 확인하고 개선하였다.

본 실험은 19일간 개별적으로 진행되었으며, 각 실험 참가자들에게는 실험 환경 준비와 개별 안내를 포함하여, 사전 및 사후 시험, 학습, 휴식시간까지 약 90분을 할당하며, 하루 최대 6명의 실험을 진행하였다. 참가자의 실험 준비가 완료되면 설계된 실험 절차에 따라 사전시험, 동영상 기반 학습, 사후시험의 순서로 진행하였다.

3.4 분석방법

분석 자료는 행동로그 데이터와 학업성취 데이터로 나누어 수집하였다. 행동로그는 동영상 학습 플레이어를 개발한 개발사의 데이터베이스에 저장된 행동로그 데이터를 추출하였으며, 학업성취 예측모형을 개발한 선행연구[38]를 기준으로 변수를 동일하게 8가지로 재범주화하고, 최종적으로 산출된 16개의 행동로그 변수에 따라 자료를 분석하였다. 학업성취는 사전시험과 사후시험의 점수 결과를 이용하였다.

수집한 데이터는 RStudio가 제공하는 패키지와 프로그램을 활용하여 분석과 시각화를 하였다. 전처리 과정에서 연구대상자별로 취합한 행동로그 데이터 중 사전 및 사후시험 결과값이 결측치였던 2명의 데이터를 이상치로 보고 제외하여 최종 76명의 연구대상자의 실험 데이터를 활용하였다. 전처리 과정 후, 행동로그와 학업성취 데이터에 대한 기술통계 분석을 진행하고, 머신러닝

기법을 통한 예측 모형을 도출하고 모형별 변수 중요도를 비교 분석하였다. 본 연구에서는 동일한 행동 로그 데이터를 토대로 머신러닝의 주요 기법으로 제시되는 각 알고리즘인, K-근접 이웃(K-Nearest Neighbors; KNN), 인공신경망(Artificial Neural Network), 서포트 벡터 머신(Support Vector Machine; SVM), 랜덤 포레스트(Random Forest) 기법을 적용하여 모형을 도출하고, 분류 분석 결과를 위주로 주요 변수별 중요도를 확인하였다.

4. 연구결과

4.1 기술통계

총 76명의 연구대상자들의 전반적인 행동로그 및 학업성취 데이터의 일반적 경향을 파악하기 위해 기술통계 분석을 실시하였다. 설명변수로 활용된 행동로그 지표는 동영상 학습 플레이어 구동에 따른 행동으로 분류되었으며, 총 21가지로 구성되었다. 행동로그 지표는 <표 3>과 같이 행동 빈도와 및 행동 시간으로 나뉘었다. 행동 빈도 지표는 12가지로 분류할 수 있었으며, 행동 시간 지표는 9가지로 분류할 수 있었다.

표 3. 행동로그 지표 분류

분류	변수명	지표 의미
행동 빈도	Annotation.Freq	북마크 및 코멘트 열람 빈도
	Bookmark.Freq	북마크 설정 및 확인 빈도
	Comment.Freq	코멘트 작성 및 확인 빈도
	Condition.Freq	외부환경 조정 빈도
	Filtering.Freq	주석 필터링 빈도
	Pause.Freq	일시정지 빈도
	Play.Freq	재생 빈도
	Seek.Freq	동영상 탐색 빈도
	Slide.Freq	슬라이드 조정 빈도
	ReplyIn.Freq	동료 학습자 코멘트 답글 작성 빈도
	Likeln.Freq	좋아요 누른 빈도
	ViewIn.Freq	동료 학습자 코멘트 확인
행동 시간	Paused.TTime	총 일시정지 시간
	Play.TTime	총 재생 시간
	Annotation.Time	북마크 및 코멘트 열람 시간
	Comment.Time	코멘트 작성 시간
	Condition.Time	외부환경 조정 시간
	Filtering.Time	주석 필터링 시간
	Seek.Time	동영상 탐색 시간
	Slide.Time	슬라이드 조정 시간
	Bookmark.MTime	북마크 설정 시간

행동로그에 따른 반응변수인 학업성취 지표로는 사전 및 사후시험 점수가 사용되었다. 행동로그와 학업성취 기술통계 결과는 <표 4>와 같다.

표 4. 행동로그 기술통계 결과

변인	단위	Mean	SD	Min	Max	
Annotation.Freq	회	8.09	14.03	0	79	
Bookmark.Freq		2.47	2.42	0	16	
Comment.Freq		5.97	4.92	0	21	
Condition.Freq		1.38	2.62	0	15	
Filtering.Freq		4.00	6.61	0	35	
Pause.Freq		14.26	13.12	0	77	
Play.Freq		31.32	23.04	2	145	
Seek.Freq		30.91	37.94	0	220	
Slide.Freq		2.30	6.19	0	37	
ReplyIn.Freq		0.11	0.39	0	2	
LikeIn.Freq		0.05	0.28	0	2	
ViewIn.Freq		2.63	4.35	0	27	
Paused.TTime		초	293.12	258.08	0	1311.55
Play.TTime	1237.97		203.16	703	1902	
Annotation.Time	53.28		99.32	0	483.67	
Comment.Time	150.04		135.25	0	553.92	
Condition.Time	13.81		52.37	0	434	
Filtering.Time	20.29		33.00	0	158.15	
Seek.Time	25.59		26.95	0	102	
Slide.Time	0.43		1.24	0	6	
Bookmark.MTime	157.03		162.54	0	722.52	
Pre	점		24.54	7.07	6	36
Post			30.07	6.17	6	41
Improved			5.53	4.56	-7	19

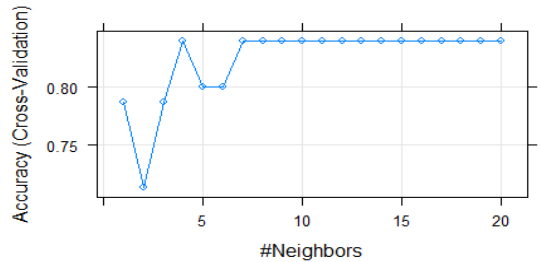
행동로그에서 빈도와 관련된 지표에서는 재생 빈도와 탐색 빈도가 가장 높은 평균 빈도를 보였다. 반대로 다른 학습자와의 상호작용을 나타내는 동료학습자 코멘트 확인, 답 코멘트 작성 행동과 좋아요 클릭 행동 빈도가 낮았으며, 개인학습에 있어서는 재생 환경 조정 빈도가 가장 낮았다. 행동로그의 지속시간과 관련해서는 총 재생 시간이 가장 높은 평균 지속시간을 나타냈으며, 슬라이드 조정 지표에 대해서는 가장 낮은 지속 시간이 나타났다.

학업성취 지표로서 44점이 만점인 사전시험과 사후시험 결과는 Pre와 Post에서 확인할 수 있는데, 각각 평균 24.54(SD=7.07), 30.07(SD=6.17)로 나타났다. 'Improved' 변수는 사전시험 대비 사후시험에서 성적 향상 정도를 나타내는 것으로 실험에서 대부분의 연구대상자들은 동

영상 학습을 진행한 후 평균 5.53(SD=4.56)점 향상된 결과를 성취하였다. 성취결과 최대 19점이 향상되기도 하였으나, 사전 성취 결과보다 7점 하락하기도 하였다.

4.2 K-근접 이웃

사전시험 대비 사후시험 점수가 향상했는지에 따라 학업성취 유무 집단으로 구분하고 행동로그가 학업성취를 예측하는지 K-근접 이웃(K-Nearest Neighbor, KNN) 모형을 통해 예측력을 확인하였다. 모형을 도출하기 위해 k값을 1부터 10까지 설정하고 분석한 결과, 최적의 결과는 k=20이었지만(Accuracy=0.840000, $\kappa=0.00000000$), Kappa 계수가 0으로 신뢰할 수 없었다. Kappa계수가 0.2 이상이면 동의할 수 있음을 고려하여[39], k=4 (Accuracy=0.8400000, $\kappa=0.20454545$)일 때, 약간 동의할 수 있었다. k=4일 때와, k=20일 때의 정확도는 [그림 2]에서와 같이 동일하게 나타났다. 하지만 보다 정합성 있는 분석 결과를 도출하기 위해 k=4일 때 모델의 성능을 평가하였다.

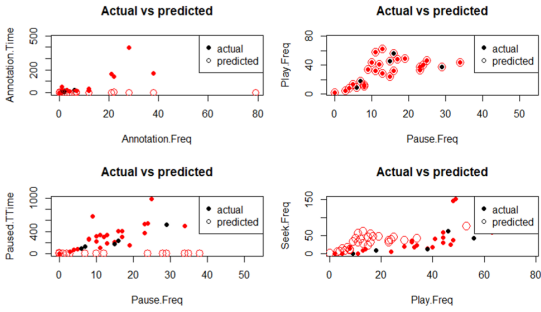


[그림 2] K-근접 이웃 모형에서의 최적의 k

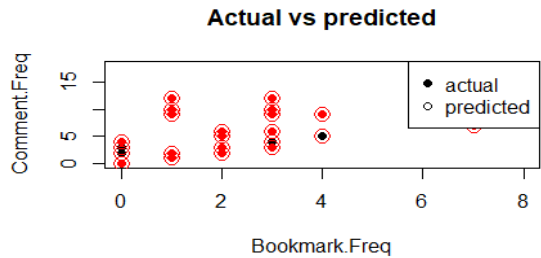
실제 값과 예측한 값에 대한 분류 모델 평가를 위해 교차분석(CrossTable)과 혼동행렬(Confusion Matrix)을 사용하였다. 교차분석에서는 총 관측치는 30개로 정밀도는 82.1%, 재현율은 92.0% 정확도는 93.3%의 결과로 모델 성능이 높은 것으로 확인되었다. 혼동행렬 결과에서도 모델에 대한 정확도가 83.3%의 높은 결과로 학업성취를 예측했다.

모델 성능 평가를 위해 21가지 행동로그 지표 간 상관을 분석하고, 지표 간 상관이 높았던 예측 결과를 시각화 하면 [그림 3]과 같다.

북마크 및 코멘트 열람 빈도와 시간 92.4%, 일시정지 빈도와 총 시간 86.6%, 일시정지 빈도와 재생 빈도는 86.3%, 재생 빈도와 동영상 탐색 빈도가 82.6%로 높은 상관을 나타냈다. 대부분의 지표가 동일한 행동에 대한 빈도와 시간의 상관이 높게 평가되었으나 북마크와 코멘트 열람, 일시정지와 같은 행동을 한 실제 학습자의 수가



[그림 3] 행동로그 지표 별 예측

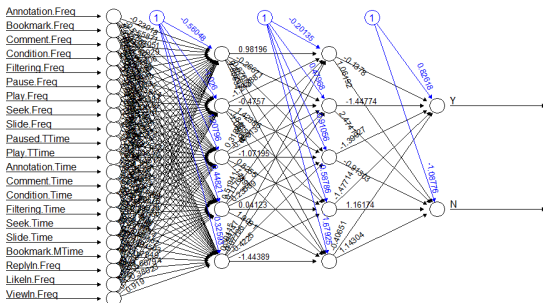


[그림 4] 북마크 빈도와 코멘트 빈도 간 행동 예측

한정적이었기에 실제 데이터와 상이한 예측 결과를 도출하였다. 일시정지 빈도와 재생 빈도는 모든 학습자 행동에 따른 높은 상관을 나타내는 지표이기에 동일한 예측 결과를 나타냈다. 북마크 빈도와 코멘트 빈도는 지표 간 상관성이 55.9%로 높다고 할 수 없지만 [그림 4]에서와 같이 실제 데이터와 유사한 예측 결과가 나타났다.

4.3 인공신경망

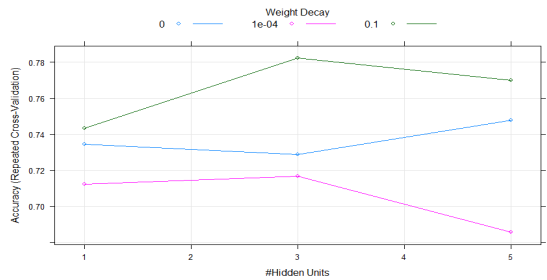
인공신경망(Artificial Neural Network; ANN) 분석 기법에 따라 21가지 행동로그 지표에 따른 학업성취 예측 결과는 [그림 5]와 같이 나타났다. 그림에서 검정 실선과 파란 실선은 각각 가중치와 편차(bias)를 의미한다.



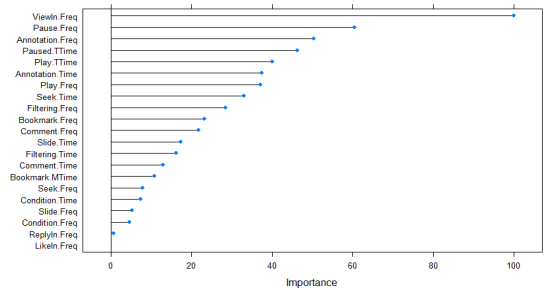
[그림 5] 인공신경망 분석 결과

분석결과 오차(Error)는 52.15544이었으며, 모형 도출을 위해 15번의 단계(Steps)를 거쳤다.

실제 데이터에 대해 학업성취를 예측하는지를 인공신경망 기법을 통해 분석한 결과 학업성취가 거의 없을 것으로 예측하였다. 하지만 95% 신뢰구간(0.0305, 0.3634)에서 정확도가 0.1429로 분류 추정 값이 불확실했다. 관찰된 정확도가 단순히 예측한 NIR(No Information Rate)와 유의미한 차이를 보이는지 나타내는 p-value 검정에 있어서도 정확도는 14.29%인데 NIR은 85.71로 차이는 있으나, Kappa계수가 0으로 신뢰하기 어려웠다. 학습률(Learning rate)을 0.5, 5개의 은닉 뉴런을 가지도록 설정한 훈련 데이터에 대한 인공신경망 분석 모델은 정확도 기준의 분류 문제로 nnet 방법을 사용하였다.



[그림 6] ANN 모델에서의 최적 모델 값 시각화



[그림 7] ANN 학업성취 예측 변수 중요도

최적 모델을 도출한 결과 [그림 6]과 같이 모델 크기가 3이며 가중치는 70, 학습률에 대한 감쇠율(decay)을 0.1로 조정할 때, 정확도가 0.7822222($\kappa=0.0983622$), 정확도에 대한 표준편차 0.1997604($\kappaSD=0.3639474$)로 최적 조정값이라고 제시되었다.

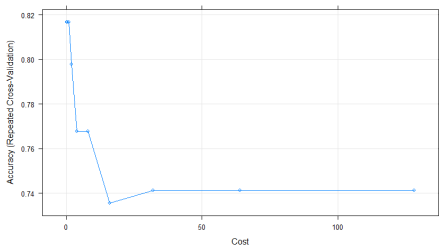
최적 모델에 대한 예측 결과, 95% 신뢰구간(0.3402, 0.7818)에서 정확도는 0.5714로 실제 데이터와 달리 상당 부분 상승했으며, 학업성취를 예측하는 것으로 나타났다. 인공신경망 기법에서 학업 성취를 예측하는 중요 행동로그 지표는 동료 학습자 코멘트 확인 빈도와 일시정지 빈

도, 북마크 및 코멘트 열람 빈도, 총 일시정지 시간, 총 재생시간이 학업성취에 가장 영향을 많이 미치는 것으로 나타났다[그림 7].

4.4 서포트 벡터 머신

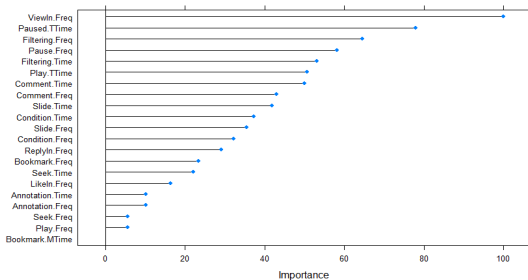
서포트 벡터 머신(Support Vector Machine; SVM) 기법으로 학습자 행동로그에 따른 성취 유무를 예측하기 위해 RStudio에서 제공하는 e1071 패키지를 사용하여 분석을 진행하였다. 분류 유형으로 41개의 서포트 벡터 수를 가지고 결과를 예측한 결과, 95% 신뢰구간(0.6366, 0.9695)에서 정확도가 0.8571 에서 학업성취가 있을 것으로 분류하였다. tune.svm으로 예측률이 가장 좋은 파라미터를 확인한 결과, gamma=0.125, cost=0.125였다. 최적의 값을 넣었을 때의 서포트 벡터 수는 45이었다. 최종 SVM 모델에서도 85.71%의 정확도로 학업성취가 있을 것으로 분류하였다.

kernellab 패키지를 사용하여 SVM 분류를 한 결과 최적 모델을 위한 파라미터는 sigma= 0.103328, cost=0.25였다 (Accuracy=0.8166667, κ =0.0000000)[그림 8].



[그림 8] SVM 모델에서의 최적 모델 값

SVM 분류 모델에서 학업성취를 예측하는 중요 행동 로그 지표는 [그림 9]와 같이 동료 학습자 코멘트 확인 빈도, 총 일시정지 시간, 주석 필터링 시간, 일시정지 빈도, 주석 필터링 시간 순으로 나타났다.

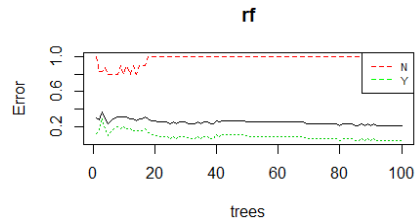


[그림 9] SVM 학업성취 예측 변수 중요도

SVM 최종 모델에 적용된 서포트 벡터 수는 43이었고, 85.71%의 동일한 정확도로 학업성취가 있을 것으로 예측하였다.

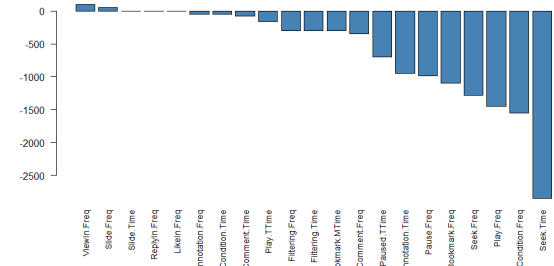
4.5 랜덤 포레스트

의사결정 나무의 과적합과 불안정성 문제를 양상블 기법으로 개선한 랜덤 포레스트는 설명 변수가 반응 변수를 예측하는 상대적인 중요도를 도출하여 실무적인 유용성이 크다고 할 수 있다. 최대 500개의 트리 분류 결과를 적용하여 랜덤 포레스트 분석 결과 최적의 mtry는 값은 1일 때 OOB error 18.18%, 2일 때 OOB error 20%이었다. 랜덤 포레스트 트리 수에 따른 오차율은 [그림 10]과 같이 나타난다.



[그림 10] 랜덤 포레스트 트리 수에 따른 오차

랜덤 포레스트에서 설명변수의 중요도를 분석하였고, 변수 간 상관에 따라 시각화하여 도출한 결과 동료 학습자 코멘트 확인 빈도와 슬라이드 조정 빈도가 중요한 것으로 나타났다[그림 11].



[그림 11] 관련성 기반 변수 중요도 시각화

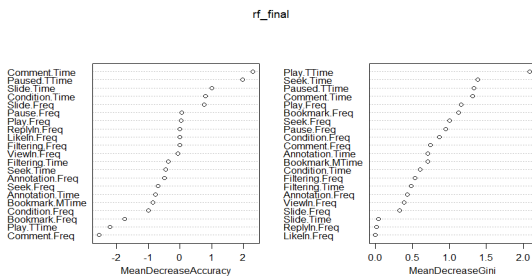
랜덤 포레스트에서는 설명변수의 중요도를 측정하기 위해 MDA(Mean Decrease Accuracy)와 MDG(Mean Decrease Gini)를 사용한다[40]. 정확도로 변수의 중요성을 판단하는 MDA는 분류 정확도를 높이는데 영향을 많이 준 변수일수록 특정 변수를 제거 후 재구축 했을 때 정확도 감소

량이 커진다. MDG는 값이 높은 변수일수록 해당 변수를 가지고 개체를 분류하면 불순도를 감소시킬 수 있다. 두 가지 지표는 모두 값이 커질수록 변수 중요도가 높으며, 랜덤 포레스트 분류 결과 변수 중요도는 <표 5>와 같이 나타났다.

표 5. 랜덤 포레스트 변수 중요도

행동변수	분류		변수 중요도 지표	
	N	Y	MDA	MDG
Annotation.Freq	-0.1280	-0.2769	-0.4947	0.4287
Bookmark.Freq	-2.2553	-1.1513	-1.7349	1.1250
Comment.Freq	-1.4139	-2.2713	-2.5461	0.7433
Condition. Freq	-1.9248	0.2682	-0.9890	0.8652
Filtering. Freq	-0.8192	0.2339	-0.0098	0.5375
Pause.Freq	-1.4986	0.8847	0.0578	0.9536
Play.Freq	-2.4634	1.2784	0.0314	1.1605
Seek.Freq	-1.4610	0.515	-0.6971	1.0016
Slide.Freq	0.2000	0.9184	0.7674	0.3255
Puased.TTime	-1.4139	2.7118	1.9796	1.3292
Play.TTime	-0.3150	-2.3362	-2.2033	2.0838
Annotation.Time	-2.1882	-0.1238	-0.7626	0.7078
Comment.Time	-0.0982	3.8738	2.2988	1.3119
Condition.Time	-0.0576	0.8633	0.7959	0.6048
Filtering.Time	-1.0050	0.2307	-0.3579	0.4778
Seek.Time	-3.2907	0.7166	-0.4451	1.3851
Slide.Time	0.0000	1.0050	1.0050	0.0385
Bookmark.MTime	-1.4286	-0.5467	-0.8410	0.7072
ReplyIn.Freq	0.0000	0.0000	0.0000	0.0160
LikeIn.Freq	0.0000	0.0000	0.0000	0.0000
ViewIn.Freq	0.2209	-0.0846	-0.0601	0.3867

도출된 변수 중요도 결과를 시각화 하면 [그림 12]와 같이 나타난다. 두 지표 모두 그래프 상단에 위치한 변수일수록 중요도가 높다고 해석할 수 있으며, 하위에 위치한 변수들은 학업 성취를 분류함에 있어 영향력이 약하다고 할 수 있다.



[그림 12] 랜덤 포레스트 변수 중요도 시각화

MDA 그래프 상 중요도가 높다고 분류된 행동변수는 코멘트 작성 시간, 총 일시정지 시간, 슬라이드 조정 시간, 외부환경 조정시간, 슬라이드 조정 빈도 순이었다.

MDG 그래프에서는 총 재생시간, 동영상 탐색 시간, 총 일시정지 시간, 코멘트 작성 시간, 재생 빈도 순으로 중요하다고 나타났다. 두 지표에서 동일하게 중요도가 높다고 분류된 변수는 코멘트 작성 시간, 총 일시정지 시간이었다. 두 변수가 학업성취를 분류하는데 중요하게 작용한 결과에 따르면, 코멘트 작성과 동영상을 일시정지 하는 행동이 학업성취 분류를 위한 주요 속성으로 해석된다. 혼동행렬을 통해 본 모델의 정확도를 분석한 결과 95% 신뢰구간(0.5809, 0.9455)에서 정확도가 0.8095에서 학업성취가 있을 것으로 분류하였다.

4. 결론 및 제언

본 연구에서는 상호작용이 가능한 동영상 기반 학습 환경에서 학습자들의 인지적, 행동적 과정을 묘사하는 시각된 어포던스에 대한 지표로 행동로그 데이터를 활용하였다. 행동로그가 학업성취를 예측하는지, 학업성취에 영향을 미치는 공통 행동이 무엇인지 탐색함으로써 동영상 학습 환경 설계 및 개발에 시사점을 제시하고자 하였다. 이를 위해 행동로그를 산출 및 수집하고, 기술통계와 머신러닝 기법으로 분석하였다. 각각의 머신러닝 모델이 가지는 단점을 보완하여 해석하기 위해 주요 분류 모델인 K-근접 이웃, 인공신경망, 서포트 벡터 머신, 랜덤 포레스트를 활용하였다. 행동로그에 대한 학업성취를 분류하고 주요 행동들을 분석함으로써 각 모델에서 공통적으로 학업성취에 영향을 미치는 주요 행동을 도출하였다.

분석결과 모든 모델에서 상호작용이 가능한 동영상 학습 환경에서의 학업성취는 있을 것으로 예측하였다. 혼동행렬 분석 결과를 기준으로 예측력이 가장 높게 나온 알고리즘은 서포트 벡터 머신 모형이었으며, 인공신경망 모형의 분석력이 가장 낮았다. 동영상 학습 환경에서 학습자들이 가장 많이 하는 행동은 재생 활동과 같은 기본 행위를 포함하여 동영상 탐색과 일시정지, 북마크 설정 및 코멘트 작성에 시간을 많이 할애하였다.

변수 중요도를 통해 도출한 학업성취에 영향을 가장 많이 미치는 학습자 행동은 서포트 벡터 머신, 인공신경망, 랜덤 포레스트 모형 모두 학습 중 동료 학습자가 작성한 코멘트를 확인하는 것으로 나타났다. 다음으로 학업성취를 높게 예측하는 행동은 총 일시정지 시간, 일시정지 빈도, 슬라이드 조정 빈도로 나타났다. 이는 학업성취가 있었던 대부분의 학생들이 성취가 없었던 학생들에 비해 코멘트 및 북마크 설정과 같이 학습에 있어 적극적인 행동을 보였던 점을 참고하면 동영상 학습에 있어 자신의 생각을 정리하는 일련의 행동을 위해 동영상을 일

시적으로 정지시키거나, 학습자 스스로에게 적절한 학습 환경을 조성하기 위해 슬라이드를 조작하는 행동을 했음을 추측할 수 있다.

학습자 행동지표 간 상관을 통해 학업성취를 예측한 결과, 지표 간 상관이 높더라도 실제 데이터에 비해 예측 결과는 다르게 도출되거나, 특정 행동지표에 있어서는 상관성이 비교적 낮더라도 실제 데이터와 예측 결과가 유사하게 나타나기도 하였다. 이는 머신러닝 분석 활용에 있어 각 모델마다 가장 적합한 계수 적용과 같은 데이터 처리 조건과 과적합 과정에서 분석 수행 결과나 예측에 상이한 차이가 나타난 것으로 해석되었다. 또한 데이터 측면에서 개별 학습자들마다 상이한 학습 행동 패턴을 가지고 있어서 실제 데이터를 통해 학습자 행동을 일괄적으로 예측하기가 어렵다고 추측할 수 있다. 하지만 학습자들의 적극적 태도를 반영하는 코멘트 작성 또는 북마크 설정과 같은 학습 행동은 일반적으로 강의를 듣기만 하는 학습자 보다 자기주도적 학습이 익숙한 학습자들의 특성으로 고려되며, 이와 같은 학습자들의 행동 양식은 실제 데이터와 예측 데이터가 거의 동일하게 도출된다고 볼 수 있다. 따라서 학습자 행동에 따라 학업성취를 보다 세밀하게 예측하기 위해서는 전체 학습자 대상 분석보다 학습자 특성을 고려한 군집을 배분하여 분석할 필요가 있다. 학습자들의 성적 및 사전지식 수준 또는 학습 태도와 학습인지전략 수준 별로 군집을 나누고 군집별 특징과 주요 행동에 따른 학업성취를 예측함으로써 모형에 대한 예측력 또한 높일 수 있을 것이다.

본 연구는 학습자 배경과 특성의 다양성을 충분히 고려하지 못한 한계를 가진다. 하지만 현대 교육에서 동영상 학습이 보편화 되었으나 일반적인 동영상 학습이 학업성취에 항상 기여하지 못한다는 한계에 따라 보다 효과적인 방법으로 양방향적 동영상 학습이 제안되었다 [37]. 본 연구에서도 양방향적 동영상 학습이 가능한 학습 환경에서 학습자 상호작용 행동이 학업성취에 영향을 미치는 결과를 도출하였다. 학습자 간 상호작용이라고 할 수 있는 동료 학습자가 작성한 코멘트를 확인하는 행동이 학업성취에 유의미하다는 결과를 도출했다. 더불어 코멘트 작성과 북마크 설정 등 학습 중 내용 정리 및 중요도 파악과 같은 사고 활동을 많이 한 학습자의 학업성취가 높았다. 따라서 학습자 간 상호작용이 가능한 동영상 학습 환경 설계가 필요하며, 학습자들이 학습 활동에 적극적으로 참여할 수 있도록 학습 환경에서의 지각된 어포던스를 부각시킬 필요가 있을 것이라는 시사점을 도출하였다.

기존의 머신러닝을 활용한 교육연구는 대부분 LMS

상의 로그파일을 기반으로 대용량 패널 데이터를 중심으로 학습자 요구분석 또는 학업성취를 예측하는 것이다 [35]. 다양한 머신러닝 모델을 활용한 교육 연구로는 이러닝 강의에 참여한 학습자의 출결, 과제 제출, 시험 성적 기록 등을 토대로 학습자 성취 수준을 예측한 것이 있다[41]. 본 연구는 기존의 빅 데이터를 중심으로 머신러닝을 활용한 연구와 달리 통제된 환경에서 일회적으로 진행된 실험 상황에서 적은 수의 학습자 데이터를 수집하였다. 하지만 피험자 내 설계(within-subject design)로 1인 당 150건 이상의 개별 학습자 데이터로 머신러닝 연구가 가능하도록 보완하였다. 본 연구를 통해 출결 및 시험 성적 데이터 중심의 연구에서 학습 과정 중 데이터를 중심으로 학업성취에 주요한 영향을 주는 요인을 규명하고자 하였다. 본 연구에서 활용한 학습자 행동 데이터는 동영상 학습 상황에서의 학습자 행동 및 매체 사용에 대한 태도를 나타내기 때문에 플립러닝, 블렌디드 러닝과 같은 동영상 학습을 활용하는 학습 환경에 대한 예측에도 공통적으로 적용할 수 있는 틀을 제시한다. 특히 머신러닝 기법을 활용하여 다양한 특성을 가진 학습자 행동을 분석할 수 있었다. 보다 우수한 예측률과 학습자 행동의 일반화된 경향성을 도출하기 위해서는 중장기적 관찰을 통한 데이터 수집이 필요하다. 그리고 학습자 성향과 수준을 고려하여 세부 집단별 데이터를 추가 확보한 후, 머신러닝 기법을 활용하면 학습자 특성에 따른 행동을 분류하고 학업성취 예측과 더불어 학습자의 학습 효과성을 증대시킬 수 있는 맞춤형 교수설계를 제공할 수 있을 것이다.

참고문헌

- [1] 산업통상자원부·정보통신산업진흥원 (2019). **2018년 이러닝 산업 실태조사**. 세종:산업통상자원부.
- [2] 조일현·김정현 (2013). 학습분석학을 활용한 e-러닝 학업성과 추정 모형의 통계적 유의성 확보 시점 규명. **교육공학연구**, 29(2), 285-306.
- [3] Zorrilla, M. E., Menasalvas, E., Marin, D., Mora, E., & Segovia, J. (2005). *Web Usage Mining Project for Improving Web-Based Learning Sites*. In *Web Mining Workshop*. Cataluna (pp.1-22).
- [4] Romero, C., & Ventura, S. (2010). Educational data mining: A Review of The State-of-the-art. *IEEE Transactions on Systems, Man and Cybernetics*, 40(6), 601-618.
- [5] 성한울·조일현 (2018). 온라인 학습 상황에서 행동 로그, 생리심리반응 및 시험불안을 통한 멀티모달 (Multimodal) 학업성취 예측모형 개발. **교육공학연구**, 34(2), 287-308.

- [6] 강혜영·이장익·김성완 (2012). 스마트교육 수용의도에 영향을 미치는 요인 간의 관계 분석. **한국컴퓨터정보학회 하계학술대회논문집**, 20(2), 141-144.
- [7] 김성현 (2009). **이러닝 학습자의 지각된 유용성과 용이성이 학습 몰입과 학습 만족도에 미치는 영향**. 석사학위논문, 숙명여자대학교.
- [8] 박성익·김연경 (2006). 온라인 학습에서 학습몰입요인, 몰입수준, 학습성취간의 관련성 탐구. **열린교육연구**, 14(1), 93-115.
- [9] 이은환·이종연 (2013). 스마트교육 환경에서 학습자 특성, 상호작용, 몰입, 지각된 유용성 및 학습만족도의 구조적 관계 분석-초등학교를 중심으로-. **교육정보미디어연구**, 19(3), 573-603.
- [10] 김선영 (2016). 융합(STEAM)형 디자인 교육을 위한 행동유도성 학습환경 고찰-노먼의 인지적 어포던스 디자인 이론을 중심으로-. **디자인융복합연구**, 15(5), 105-118.
- [11] Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin, New York.
- [12] Norman, D. A. (1988). *The Psychology of Everyday Things*. New York: Basic Books.
- [13] Mishra P. & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record*, 108(6), 1017e1054.
- [14] Brinton, C. G., & Chiang, M. (2015, April). MOOC performance prediction via clickstream data and social learning networks. In *2015 IEEE conference on computer communications (INFOCOM)* (pp. 2299-2307). IEEE.
- [15] Yoo, J. E. (2018). TIMSS 2011 Student and Teacher Predictors for Mathematics Achievement Explored and Identified via Elastic Net. *Frontiers in Psychology*, 9, 317. Retrieved from <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.00317/full>
- [16] Bzdok, D., Altman, N., & Krzywinski, M. (2018). Statistics Versus Machine Learning. *Nature Methods*, 15(4), 233-234.
- [17] 유진은 (2015). 랜덤 포레스트. **교육평가연구**, 28(2), 427-448.
- [18] Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, 88, 2783-2792.
- [19] Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, 14, 323-348.
- [20] Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction (2nd ed.)*. New York: Springer.
- [21] Mayer, R. E. (2009). *Multimedia Learning* (Vol. 2nd ed). Cambridge: Cambridge University Press. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=e000xww&AN=400712&lang=ko&site=ehost-live>
- [22] Maniar, N., Bennett, E., Hand, S., & Allan, G. (2008). The Effect of Mobile Phone Screen Size on Video Based Learning. *Journal of software*, 3(4), 51-61.
- [23] 김지현 (2019). **동영상 기반 학습 환경에서 학습자의 시각 행동에 의한 입체적 학습과정 분석 및 학습성과 예측모형 개발**. 석사학위 논문. 이화여자대학교.
- [24] 임철일 (2001). 웹 기반 자기조절학습환경을 위한 설계 전략의 특성과 효과. **교육공학연구**, 17(3), 53- 83.
- [25] 고호경 (2016). 스마트 기기 활용 시니어 수학 앱 개발을 위한 어포던스 설계 전략. **E-수학교육 논문집**, 30(1), 85-99.
- [26] 박형주 (2012). **이러닝 콘텐츠 사용자 인터페이스의 교육적 사용성 증진을 위한 어포던스 설계 전략 개발**. 박사학위논문. 중앙대학교.
- [27] 민병택 (2007). **스마트 공간 환경에 있어 어포던스 디자인의 요인과 구조**. 석사학위논문. 홍익대학교.
- [28] 김성훈·김인수 (2013). 인지적 어포던스 기반의 인터랙티브 e-퍼블리싱 콘텐츠 디자인에 관한 연구. **디자인 지식저널**, 25, 355-366.
- [29] 박성익 (2010). 공공기관 교육용 콘텐츠의 학습효과 증진을 위한 디자인 전략. **한국콘텐츠학회논문지**, 10(3), 444-453.
- [30] Bower, M., & Hedberg, J. G. (2010). A Quantitative Multimodal Discourse Analysis of Teaching and Learning in a Web-conferencing Environment—the Efficacy of Student-centred Learning Designs. *Computers & education*, 54(2), 462-478.
- [31] Black, E. W., Dawson, K., & Priem, J. (2008). Data for Free: Using LMS Activity Logs to Measure Community in Online Courses. *The Internet and Higher Education*, 11(2), 65-70.
- [32] Mesgari, M., & Okoli, C. (2015). Ecological Approach to User Sensemaking of Technology. In *Proceeding of International Conference on Information Systems*, 1-12.
- [33] Bzdok, D., Altman, N., & Krzywinski, M. (2018). Statistics Versus Machine Learning. *Nature Methods*, 15(4), 233-234.
- [34] Yoo, J. E., & Rho, M. (2018). *Exploration of Predictors to Teacher Job Satisfaction Using TALIS 2013 Korea Data*. Presented at the 19th ICER(International Conference on Educational Research). Seoul, South Korea.
- [35] 유진은 (2019). 기계학습: 대용량/패널자료와 학습분석학 자료 분석으로의 활용. **교육공학연구**, 35(2), 313-338.
- [36] Klefodimos A., & Evangelidis G. (2016) An Interactive Video-Based Learning Environment Supporting Learning Analytics: Insights Obtained from Analyzing Learner

Activity Data. In *State-of-the-Art and Future Directions of Smart Learning*, 471-481.

- [37] Zhang, D., Zhou, L., Briggs, R. O., & Nunamaker Jr, J. F. (2006). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & management*, 43(1), 15-27.
- [38] 김다솜 (2019). **동영상 기반 학습 환경에서 지각된 어포던스를 통한 학습성취 예측모형 개발: 행동 로그와 안구 운동을 중심으로**. 석사학위 논문. 이화여자대학교.
- [39] Altman, D. G. (1991). *Practical Statistics for Medical Research*. London: Chapman and Hall.
- [40] Archer, K. J., & Kimes, R. V. (2008). Empirical characterization of random forest variable importance measures. *Computational Statistics & Data Analysis*, 52(4), 2249-2260.
- [41] 조한국 (2018). 머신 러닝을 활용한 이러닝 학습 환경에서의 학습자 성취 예측 모형 탐색. **학습자중심교과교육연구**, 18(21), 553-572.



김 다 솜

2016년 이화여자대학교
교육공학과(교육학석사)
2019년 이화여자대학교 대학원
교육공학과(교육학석사)
관심분야: 에듀테크, 학습분석학,
인적자원개발
E-Mail: pumiliosomnium@gmail.com



조 일 현

1987년 서울대학교
농경제학과(경제학사)
1994년 연세대학교 교육학과
(교육학석사)
2001년 Florida State University
교육공학과(교육학박사)

2004년~2008년 춘천교육대학교 사범대학 교육학과 교수
2008년~현재 이화여자대학교 사범대학 교육공학과 교수
2018년~현재 이화여자대학교 에듀테크융합연구소장
2018년~현재 국가공무원인재개발원 겸임교수
관심분야: 학습분석학, 인적자원개발, 빅데이터분석
E-Mail: ijo@cwaha.ac.kr



이 정 은

2013년 영남대학교 국제통상학과
(상학사)
2018년~현재 이화여자대학교
교육대학원 교육공학·HRD전공
석사과정
관심분야: 에듀테크, 학습분석학,
인적자원개발
E-Mail: jungeunjoan.lee@gmail.com