

프로세스 마이닝을 활용한 온라인 교육 오픈 플랫폼 내 학습 패턴 분석 방법 개발*

김태영
광운대학교 정보융합학부
(kty4119@naver.com)

김효민
광운대학교 정보융합학부
(khmiee@naver.com)

조민수
광운대학교 정보융합학부
(mcho@kw.ac.kr)

비대면 교육의 중요성 및 필요에 따른 수요가 증가함에 따라 국내의 온라인 교육 오픈 플랫폼이 활성화되고 있다. 본 플랫폼은 대학 등 교육 전문기관과 달리 학습자의 자율성이 높은 특징을 가지며 이에 따라 개인화된 학습 도구를 지원하기 위한 학습 행동 데이터의 분석 연구가 중요시 되고 있다. 실제적인 학습 행동을 이해하고 패턴을 도출하기 위하여 프로세스 마이닝이 다수 활용되었지만 온라인 교육 플랫폼과 같이 자기 관리형(Self-regulated) 환경에서의 학습 로그를 기반한 사례는 부족하다. 또한, 대부분 프로세스 모델 도출 등의 모델 관점에서의 접근이며 분석 결과의 실제적인 적용을 위한 개별 패턴 및 인스턴스 관점에서의 방법 제시는 미흡하다. 본 연구에서는 온라인 교육 오픈 플랫폼 내 학습 패턴을 파악하기 위하여 프로세스 마이닝을 활용한 분석 방법을 제시한다. 학습 패턴을 다각도로 분석하기 위하여 모델, 패턴, 인스턴스 관점에서의 분석 방법을 제시하며, 프로세스 모델 발견, 적합도 검사, 군집화 기법, 예측 알고리즘 등 다양한 기법을 활용한다. 본 방법은 국내 오픈 교육 플랫폼 내 기계학습 관련 강좌의 학습 로그를 추출하여 분석하였다. 분석 결과 온라인 강의의 특성에 맞게 비구조화된 프로세스 모델을 도출할 수 있었으며 구체적으로 한 개의 표준 학습 패턴과 세 개의 이상 학습 패턴으로 세분화할 수 있었다. 또한, 인스턴스별 패턴 분류 예측 모델을 도출한 결과 전체 흐름 중 초기 30%의 흐름을 바탕으로 예측하였을 때 0.86의 분류 정확도를 보였다. 본 연구는 프로세스 마이닝을 활용하여 학습자의 패턴을 체계적으로 분석한다는 점에서 기여점을 가진다.

주제어 : 학습 패턴 분석, 프로세스 마이닝, 온라인 교육, 프로세스 모델 발견, 오토인코더

논문접수일 : 2023년 6월 12일 논문수정일 : 2023년 6월 22일 게재확정일 : 2023년 6월 22일
원고유형 : Regular Track 교신저자 : 조민수

1. 개요

팬데믹으로 인한 사회 및 환경의 변화뿐만 아니라 디지털 전환, 메타버스, 인공지능 등의 ICT 기술은 온라인 교육의 발전 및 실제적 활용의 증대를 가능하게 하였다(손세창, 윤희영, 2022). 온라인 교육은 고등학교, 대학교 등 전문 교육기관 외 다양한 공공, 민간 기관에서 교육 콘텐츠를

개발 및 보급하여 지속적인 발전을 이룩하고 있다. 특히, 전 국민을 대상으로 무료 공개강좌를 제공하는 온라인 교육 오픈 플랫폼으로 인하여 이를 바탕으로 지식과 기술을 습득하려는 수요가 증가하고 있다(허묘연, 2020).

온라인 교육은 비대면적 특성으로 인하여 학습 환경 및 과정에 대한 통제가 가능한 집체 교육과 달리 학습자의 자율성이 크게 높아진다(허묘연,

* 본 연구는 네이버 커넥트재단 지원 및 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2021R1G1A1094019).

2020). 온라인 교육 오픈 플랫폼 내에서의 온라인 교육은 구체적인 학습 성취가 불분명하기 때문에 학습자의 자기 주도(Self-regulated) 학습이 더욱 중요시된다. 이를 위해, 학습자의 학습 과정을 관리하고 학습을 돕기 위한 보조 도구의 개발이 수행되고 있으며, 추가적으로 학습 패턴을 측정하고 이를 분석하여 학습자의 학습 성과를 개선하기 위한 연구 개발이 병행되고 있다.

교육공학 분야에서 학습자의 학습 개선을 주목적으로 한 다양한 연구 개발이 선행되었다. 대표적인 학습 양식 측정 도구인 Felder-Silverman 모델은 학습자가 정보를 수용하고 처리하는 방식과 관련하여 처리, 지각, 입력, 이해 네 개의 차원에 대한 세분화된 학습양식 정의 방법을 제시하였다(Graf et al., 2007). 이는 정규화된 질의응답을 통해 학습자의 학습양식을 측정하는 정성적 분석 방법의 대표적인 도구이며, 이를 바탕으로 학습 패턴 유형과 학업성취도, 수업 만족도 등 다양한 연구가 수행되었다(김혜경, 배성아, 2019).

학습 패턴에 대한 정성적 분석과 더불어, 학습 관리 시스템(LMS, Learning Management Systems)을 활용함에 따른 누적 학습 행동 데이터를 기반한 정량적 학습 분석 관련 연구도 다수 수행되었다(Siemens, 2013). 학습 행동 데이터란 학습자가 LMS를 통해 참여한 영상 시청, 과제 수행, 교수자와의 질의응답 등의 학습과 관련된 누적 데이터를 의미한다. 이를 바탕으로 누적 접속 수, 평균 학습시간, 토의 참여 횟수 등 주요 변수를 설계하여 학습 패턴 분류 및 성과 예측을 위한 기계학습 기반 모델링 방안이 다수 연구되었다(Black et al., 2008).

이와 더불어, 학습자의 행동 로그를 그대로 활용하는 데 효과적인 프로세스 마이닝 기술을 활용한 연구들도 수행되었다(Mukala et al., 2015,

Hachicha et al., 2021). 프로세스 마이닝은 정보 시스템에 누적된 로그 데이터를 기반으로 프로세스를 모델링, 분석, 개선하는 것을 목적으로 한다(van der Aalst, 2016). 프로세스 모델 발견, 적합도 검사, 객체 중심 프로세스 마이닝, 성과 분석 및 예측 등 프로세스 관점에서의 다양한 기법들이 포함되며, 이를 기반으로 온라인 교육에서의 학습자 행동을 이해하기 위한 노력이 수행되었다.

온라인 교육에서의 학습자의 행동을 이해하기 위한 다양한 시도가 있었지만 다음과 같은 한계 및 미비점이 있다. 먼저, 많은 연구에서 제시된 행동 양식은 전문 교육기관의 교과과정에 해당하는 강좌이며 온라인 교육 오픈 플랫폼에서의 오픈 강좌를 대상으로 하지 않았다. 즉, 오픈 강좌에서 나타나는 학습자 행동 양식을 분석하여 정리한 연구 결과가 부재한 상황이다. 또한, 대부분 모델 관점에서의 프로세스 전반을 파악하기 위한 목적으로 학습 행동 관련 프로세스 모델을 주로 도출하였으며, 프로세스 패턴, 학습자 관점에서의 학습 패턴을 파악하는 분석은 미흡하다.

이를 극복하기 위하여, 본 연구에서는 온라인 교육 오픈 플랫폼 내 학습 패턴을 파악하기 위한 분석 방법을 제시한다. 특히, 학습 패턴을 다각도로 분석하기 위하여 프로세스 모델, 패턴, 인스턴스 관점에서의 분석 방법을 제시한다. 모델 관점에서는 학습자의 전체적인 학습 행동 파악을 목적으로 하며 프로세스 모델 발견 알고리즘, 모델 평가 기법을 적용한다. 패턴 레벨에서는 학습 패턴 도출 및 표준화된 패턴과 특이 패턴을 정의하고자 하며, 이를 위해 오토인코더 기반 자취 군집화를 활용한다. 마지막으로, 인스턴스 관점에서는 개별 학습자별 학습 패턴 및 성과를 예측한다. 본 연구 방법의 검증은 위하여, 국내 한 오픈 교육 플랫폼 내 기계학습 관련 강좌의 학습

로그를 추출하여 분석하였다.

본 논문의 나머지 부분은 다음과 같이 구성된다. 2장은 기존 연구의 동향을 파악하고, 3장은 본 연구에서 제시하는 연구 방법을 소개한다. 4장에서는 사례 연구를 통해 실제 데이터 분석 결과를 제시하며, 5장에서는 본 논문의 내용을 요약하고 향후 연구에 대해 소개한다.

2. 관련 연구

2.1. 데이터 기반 학습 분석

학습 분석은 학습자의 학습에 대한 이해와 최적화를 목적으로 학습 행동의 수집, 분석 등을 목적으로 한다(권영욱, 2013, 정준희 2012). 데이터마이닝, 기계학습, 인공지능 등의 기술 발전에 따라 학습 분석 분야는 데이터를 기반한 증거기반 연구가 주를 이루고 있다. 데이터 기반 학습 분석은 분석 대상, 분석 목적 및 방법으로 구분될 수 있다.

데이터 기반 학습 분석의 대상은 강의의 종류 즉, 오프라인, 온라인 여부에 따라 구분될 수 있다. 오프라인 강의의 경우 학습자의 설문조사 결과를 분석 데이터로 하여 기초 통계량 분석 및 통계 검정 분석을 활용한다. 대표적으로 설문조사를 통해 튜터와 튜티의 학업성취도와 만족도 간 관계 연구(김혜경과 배성아, 2019), 학습 방식을 향상시키기 위한 교육 데이터 사이언스 프로그램 내 교수자의 경험 파악 연구(McCoy and Shih, 2016) 등이 있다. 온라인 강의의 경우 Moodle과 같은 학습 관리 시스템을 통해 학습자의 학습 관련 모든 행동을 데이터 로그로 활용하여 학습 패턴을 분석한다 (Black et al., 2008). 로그를 기반

으로 학습자의 학습 행동을 표현하는 파생 변수를 정의하여 이를 통해 예측, 분류, 군집화 등을 수행하는 기계학습, 딥러닝 기반 연구(Lockyer et al., 2013, Gašević et al., 2016)와 로그를 그대로 활용하여 프로세스 모델, 마르코프 모델 등의 네트워크를 도출하는 연구로 구분할 수 있다(Mukala et al., 2015, Hachicha et al., 2021). 이 밖에, 학습 분석 대상은 강의 제공 주체가 고등 교육 기관 혹은 오픈 플랫폼 인지에 따라 구분될 수 있다. 대학 등 고등 교육 기관과 달리 오픈 플랫폼은 Personal Learning Environments, Massive Open Online Courses 등이 포함되며 강의의 대상이 되는 학습자가 다양하고 이에 따라 다양한 학습 패턴이 도출된다는 특징을 갖는다(Muslim et al., 2020).

분석 목적의 관점에서는 묘사적 분석을 통한 모니터링, 예측적 분석을 통한 성과 예측, 처방적 분석을 통한 개인화로 구분할 수 있다. 학습자의 학습 패턴 및 성과를 시각화하기 위한 분석 및 시스템이 개발되었으며, 평가 점수, 참여 수준 등의 주요 지표들이 활용되었다(Chen et al., 2019). 예측 관련 연구는 학습자의 최종 성과를 예측하는 것이 주를 이루었다. 학습자간 비교 분석을 통해 각 학습자의 성과 예측, 강의 수료에 위험이 있는 학생 파악을 위한 Graduate At-Risk 모델링의 연구가 수행되었다(Fleur et al., 2020, Baneres et al., 2019). 또한, Gašević et al. (2016)은 학습 패턴을 분석하여 학습 성과를 예측하고 이를 해석하는 방법을 제시하였으며, 교육 조건 및 학습 환경이 성과에 미치는 영향을 파악하여 9개 교과목별 예측 모델과 전체 성과 예측 모델을 개발하였다. 처방적 분석 관련 연구에서는 강의에 대한 개별 학습자별 콘텐츠 및 활동 추천, 메시지 및 face-to-face 대시보드 개발 등이 연구되었다(Han et al., 2021, Kokoc and Altun, 2021).

분석 방법 관점에서는 대부분 회귀 및 분류를 포함한 예측 방법이 대다수를 이루었다. 이는 앞서 제시한 예측적 분석 목적에 대응되는 부분이다. 이 밖에, 학습 패턴을 기반으로 학습자의 군집화를 수행하는 군집화 방법도 다수 발견되었다. Khalil and Ebner (2017)은 군집화를 적용하여 학습을 위한 외부, 내부요인의 중요도를 파악한다. 더불어, 학습 패턴을 파악하기 위한 연관 규칙 학습법 등을 포함한 관계 마이닝 기법도 활용되었다 (Viberg et al., 2018).

2.2. 프로세스 마이닝 기반 학습 분석

프로세스 마이닝 기술은 이벤트 로그로부터 프로세스 관련 지식을 추출하는 것을 목적으로 한다(강영식 등, 2023). 프로세스 마이닝은 프로세스 발견, 적합도 검사, 성과 분석, 비교 분석, 예측 분석 등의 세부 기술로 구성된다. 실제적 프로세스 도출의 장점을 바탕으로 프로세스 마이닝은 의료, 제조, 서비스 등 다양한 분야에 활용되었으며 교육 분야에서도 연구 성과가 존재한다(Bogarín et al, 2017).

관련 연구 중 대다수는 이벤트 로그를 바탕으로 학습 관련 프로세스 모델의 도출을 목적으로 하였다. Bogarín et al. (2017)은 Moodle 로그로부터 학습자의 학습 행동을 모델링하였으며, 이를 세분화하기 위하여 클러스터링을 같이 활용하였다. 또한, Mukala et al. (2015)은 로그 데이터를 활용한 프로세스 마이닝 분석 방법을 제시하였다. Dotted Chart, 프로세스 발견을 통해 학생 그룹들 간의 행동 차이를 찾고 분류하여 학생들의 실제 학습 행동을 시각화하였다. Hachicha et al. (2021)은 Moodle 학습 관리 시스템에서 추출한 데이터를 기반으로 이탈할 가능성이 있거나 초기 단계

에서 실패할 가능성이 있는 학습자를 식별하고 학습자의 학업 성과를 향상시킬 수 있는 방법을 제시하였다. 또한, Schoor and Bannert (2012)는 CSCL(Computer-Supported Collaborative Learning)에서 수집된 로그에 Fuzzy Miner를 사용하여 채팅 로그에서 발생한 이벤트 시퀀스를 분석하였다.

또한, 프로세스 모델을 평가하기 위하여 적합도 검사 및 선형 시제 논리를 활용하는 관련 연구도 수행되었다. Trcka and Pechenizkiy (2009)는 학습자의 수강, 시험, 시작, 종료 등 다양한 학습 패턴을 모델 형태로 도출하고 이를 적합도 검사를 통해 평가하는 방법론을 개발하였다. Bannert et al. (2014)은 자기 규제 학습관련 이벤트를 도출하여 학습 성과가 좋은 그룹과 그렇지 않은 그룹간 적합도 비교 분석을 수행하였다. 그 결과 모니터링 후 평가, 검색 후 읽기 및 평가 패턴에서 두 그룹 간 차이가 명확한 것을 파악할 수 있었다.

이 밖에, Cairns et al. (2014)는 프로세스 관점이 아닌 작업자 관점으로 사회 연결망 분석을 통해 교수자와 학습자 간 상호작용을 파악하였다. 또한, Juhaňák et al. (2019)는 다양한 유형의 온라인 퀴즈 활동 관련 행동 및 상호작용 패턴 시간적 측면인 성과 관점에서 분석하여 새로운 분류의 행동 패턴 파악 및 비표준적 퀴즈 응시 행동 감지 방법을 제시하였다.

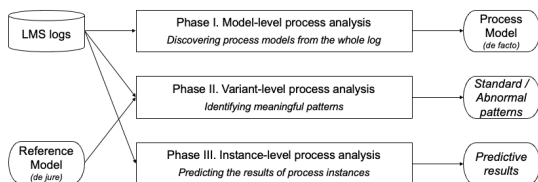
프로세스 마이닝 기술은 이와 같이 다양한 학습 행동 로그에 적용되어 학습 패턴 파악을 주목적으로 한 다양한 문제를 해결하였다. 그러나, 대부분 대학 등 관리된 환경에서의 학습 로그를 기반하며, 온라인 교육 플랫폼과 같이 자기 관리형(Self-regulated) 환경에서의 학습 로그를 기반한 분석은 부족하다. 이 밖에, 대부분 프로세스 모델 관점에서의 문제 해결을 중점으로 다루어 전반적인 흐름을 파악하고 있다. 그러나, 분석 결

과의 실제적인 적용을 위해서는 세밀한 형태 즉, 개별 패턴 혹은 인스턴스 관점에서의 분석 및 예측 등의 학습 기반 방법 제시가 필요하다. 이를 위해, 본 연구에서는 아래와 같이 온라인 교육 플랫폼 내 프로세스 분석 방법을 제시한다.

3. 온라인 교육 플랫폼 내 학습 프로세스 분석 방법

3.1. 연구 프레임워크

본 연구에서는 프로세스 마이닝 기법을 적용하여 온라인 교육 플랫폼 내 학습 프로세스를 분석하는 방법을 제시한다. 연구 프레임워크는 <그림 1>과 같다. 본 연구에서는 LMS로부터 누적된 이벤트 로그를 주요 데이터로 활용하며, 강의자가 정의한 레퍼런스 모델(*de jure model*)을 보조 데이터로 활용한다. 프로세스 분석은 세 단계로 구성되어 있으며 모델, 패턴, 프로세스 인스턴스의 서로 다른 관점을 중심으로 분석한다. 모델 레벨은 LMS 로그를 바탕으로 전체 학습 행동을 파악할 수 있는 프로세스 모델 도출을 목적으로 하며, 패턴 레벨은 세부적으로 표준 및 이상 패턴을 도출하는 것을 목적으로 한다. 추가적으로, 인스턴스 레벨에서는 개별 학습자별로 학습에 대한 성과를 예측하는 것을 목적으로 한다.



<그림 1> LMS 로그 기반 프로세스 분석 절차

3.2. LMS 로그

LMS 로그는 프로세스 마이닝에서 활용하는 이벤트 로그의 형식과 동일하다. 이벤트 로그는 프로세스 마이닝을 위한 기초 자료이며 이벤트의 집합으로 정의된다. 각 이벤트는 케이스(Case), 작업(Activity), 시간기록(Timestamp)의 필수 요소를 포함하며 작업자(Resource), 비용(Cost) 등의 추가적인 정보를 보유할 수 있다(Cho et al., 2017). LMS 로그는 학습자가 온라인 수업을 수강함에 있어 발생하는 학습 기록과 관련된 이벤트를 수집한 정보이다. 즉, 각 학습자별 영상 시청기록, 퀴즈 참여 기록, 포럼 등 상호작용 활동 기록을 포함한 LMS 내 모든 이벤트가 포함된다. <표 1>은 LMS 로그의 간단한 예시를 나타낸다.

표에 제시된 것과 같이, LMS 로그 내 작업에는 Video(영상 시청), Quiz(퀴즈 참여), Project(프로젝트 참여), Forum(상호작용) 등의 다양한 LMS 내 활동이 포함될 수 있다. 이때, Video의 경우 Course(강좌), Chapter(장), Lecture(강의)의 계층을 갖고 있는 교육 도메인의 특성을 활용하여 작업을 세분화할 수 있다. 예를 들어, Event 1의 Activity1(상위 수준)은 Video로 기록되지만 Chapter와 Lecture를 활용하여 Video_C1(Activity2), Video_C1_L1(Activity3)으로 세분화할 수 있다. 이와 더불어, 강의 시청 지속(continued, c), 완료(finished, f) 여부를 바탕으로 동일한 강의를 구분할 수 있으며, 예를 들어, Activity4에서 Event 1은 video_C1_L1_c, Event 3은 video_C1_L1_f로 기록된다.

3.3. Model-level 프로세스 분석

모델 레벨 프로세스 분석은 먼저 프로세스 모델을 도출하기 위하여 프로세스 발견 알고리즘을 적용한다. 프로세스 발견(Automated discovery)은

〈표 1〉 LMS 이벤트 로그 데이터 예시

Event	User	Timestamp	Progress	Activity1	Activity2	Activity3	Activity4
1	U1	2023-04-05 15:18:47	30	video	video_C1	video_C1_L1	video_C1_L1_c
2	U1	2023-04-05 15:25:07		project			
3	U1	2023-04-05 15:30:23	100	video	video_C1	video_C1_L1	video_C1_L1_f
4	U1	2023-04-05 15:33:32	45	video	video_C2	video_C2_L1	video_C2_L1_c
5	U2	2023-04-08 11:13:01	100	video	video_C2	video_C2_L1	video_C2_L1_f
6	U2	2023-04-08 11:17:50		forum			
7	U2	2023-04-08 11:39:25	100	video	video_C2	video_C2_L1	video_C2_L1_f
8	U2	2023-04-19 17:02:29		quiz			
...

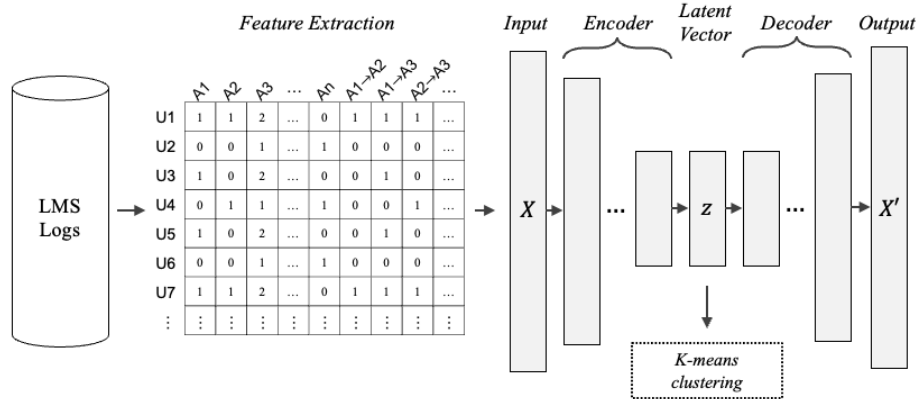
이벤트 로그로부터 순차(Sequence), 병렬 분기(Parallel), 배타 분기(Exclusive choice), 순환(Loop) 등 제어 흐름(Control-flow)을 도출하여 실제적인 프로세스 모델 분석을 목적으로 한다(Augusto et al., 2018). 이를 위해, Alpha mining, Fuzzy Mining, Heuristic Mining, Inductive Mining, Evolutionary Tree Mining, Split Mining, ILP Mining 등 다양한 알고리즘을 적용할 수 있다. 본 연구에서는 Fuzzy Mining을 활용한다.

프로세스 모델 도출 후에는 도출된 프로세스 모델을 평가하는 과정이 필요하다. 본 연구에서는 적합도 검사(Conformance Checking)를 기반한 Fitness 측정, 복잡성 판단을 위한 Simplicity 측정으로 모델을 평가한다(van der Aalst, 2016). 적합도 검사는 프로세스 모델 도출에 활용된 LMS 로그와 프로세스 모델을 비교하여 정량적 결과를

도출하는 것을 목적으로 한다. 일반적으로, 프로세스 모델이 이벤트 로그를 얼마나 설명할 수 있는지를 평가하는 Fitness가 주로 활용된다. 이와 더불어, 프로세스 모델은 사용자의 이해를 돕기 위해 가능한 한 간단하게 모델링 되는 것이 중요하며, 이를 측정하기 위해 Cardoso Metric과 Cyclomatic Metric을 활용한다(Cardoso et al., 2006).

3.4. Variant-level 프로세스 분석

패턴(Variant) 관점에서의 프로세스 분석은 모델에 포함된 여러 패턴 중 표준 패턴과 이상 패턴을 감지하는 것을 목적으로 한다. 이를 위해, 먼저 패턴 도출을 위해 <그림 2>와 같이 Autoencoder-based Trace clustering을 적용한다. 먼저, LMS 로그로부터 각 학습자별 Trace를 표현하는 Feature extraction



〈그림 2〉 Autoencoder-based Trace clustering

을 적용한다. 이때, 작업의 빈도, 작업 선후행 관계 빈도 등의 Feature가 적용된다. 도출된 인풋을 Autoencoder 아키텍처를 활용하여 잠재 벡터를 도출 후 이를 활용하여 군집화를 수행한다. 본 연구에서는 K-means clustering을 활용한다.

군집화 후, 각 군집에 해당하는 세부 패턴 로그별 프로세스 모델을 도출한다. 도출된 여러 패턴에 대해 교수가 정의한 프로세스 모델과 비교 및 유사도를 계산하여 표준 패턴과 이상 패턴을 정의하여 이를 분석한다.

3.5. Instance-level 프로세스 분석

프로세스 인스턴스 관점에서의 분석은 각 학습자별 학습 흐름이 추후 표준 혹은 이상 패턴에 속하는지 예측하는 것을 목적으로 한다. 이를 위해, 각 케이스별 전체 기록을 대상으로 하는 것이 아닌 일정 부분을 마스킹하여 활용한다. 전체 기록을 활용할 경우 특정한 학습자가 표준 혹은 이상 패턴에 속하는 지를 쉽게 판단할 수 있다. 그러나, 이는 예측의 의미를 갖지 않는다. 따라서, 해당 학습자가 현재 수강 중이라는 가정 하에 특정한

시점까지의 학습 흐름만을 활용하고 이후 부분은 마스킹한 상태로 표준 혹은 이상 패턴 여부를 예측한다.

표준 혹은 이상의 이진 분류 모델을 구축하기 위해, 학습 및 테스트 데이터를 8:2로 분리한다. 또한, 나이브 베이즈 분류기, 서포트 벡터 머신(SVM), 로지스틱 회귀 등 다양한 기계학습 모델을 활용한다. 모델의 성능을 평가하기 위하여, 정확도(Accuracy), F1-Score, AUC-ROC를 활용한다.

4. 연구 결과

4.1. 연구 대상

본 연구는 국내 전 국민 대상 온라인 교육 플랫폼에서 수집된 학습자 로그를 활용하였다. 해당 플랫폼은 소프트웨어개발, 인공지능, 데이터 사이언스, 언어, 수학 등 다양한 분야의 강좌를 제공한다. 이 중 본 연구에서는 기계학습 관련 강좌 중 학습자가 많은 한 강좌를 선택하여 연구 대상으로 설정하였으며, 해당 강좌는 총 12개의

세부 챗터로 구성되며, 각 챗터는 최소 한 개부터 최대 17개의 강의를 포함하여 총 49개의 강의로 구성된다.

본 강좌의 수강 로그에 관한 요약은 <표 2>에 나타나 있다. 본 강좌가 개설된 2018년 3월부터 2022년 5월까지의 약 4년간의 수강 기록을 추출하였으며, 총 9,448명의 학습자가 본 강좌를 수강하였다. 해당 학습자는 약 13만 개의 학습 관련 이벤트를 수행하였으며 총 49개의 작업으로 구성되어 있다. 본 연구에서는 다양한 학습 작업 중 강의 시청과 관련된 작업만 추출하였고 <표 1> 중 Activity3인 챗터, 강의를 포함한 형태를 활용하였다. 이를 기준으로, 학습자별 평균적으로 14개의 강의를 시청하였고, 최소 1개, 최대 약 1200개의 강의를 시청한 경우도 포함되었다. 이와 더불어, 본 연구에서는 프로세스 마이닝 기법을 적용하기 위하여 Fluxicon사에서 만든 Disco와 오픈소스 소프트웨어인 ProM Framework를 활용하였다.

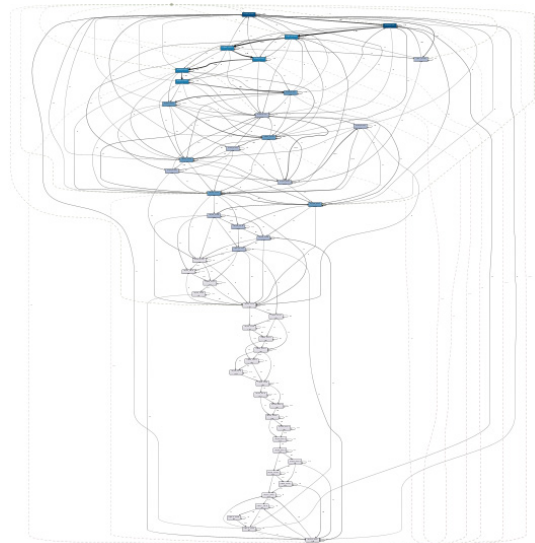
<표 2> 연구 대상 로그 요약

전체 학습자 수 (명)	9,448					
이벤트 수(개)	131,795					
학습자별 평균 이벤트 수(개)	평균	최소	최대	Q1	Q2	Q3
	14	1	1,206	2	5	16
작업 수(개)	49개					
데이터 수집기간	2018년 3월 ~ 2022년 5월					

4.2. Model-level 프로세스 분석

프로세스 모델 도출을 목적으로 추출된 학습 로그에 Fuzzy Mining을 적용하였으며, 도출된 프로세스 모델은 <그림 3>과 같다. 모델의 노드는 학습 작업을 나타내며 간선은 작업 간 선후행 관계

를 의미한다. 모델의 가독성을 위하여 노드는 전체 (100%), 간선은 20%만 포함하였다. 필터링에도 불구하고, 프로세스 모델 도출 결과 복잡한 구조를 가진 전형적인 스파게티 형태를 갖는 것이 확인되었다. 또한, 모델 내 빈도는 노드와 간선 색의 짙음으로 구분되며, 짙을수록 빈도가 높다는 것을 의미한다. 그 결과, 프로세스 시작 후 초반에는 많은 학습자들이 유사한 학습 작업들을 수행하는 반면, 강의 종료에 해당하는 학습 작업들은 일부 학습자만 참여하는 것을 확인할 수 있다.



<그림 3> Fuzzy Miner 기반 프로세스 모델

구체적으로, <표 3>은 각 강의 별 시작, 종료 빈도의 비율을 나타낸다. 대부분의 학습자가 본 강좌의 소개 및 개요에 해당하는 C1, C2 내 강의로 시작하였다. 이와 반대로, 종료의 경우 C1, C2, C3, C6, C12 등 학습자별로 서로 다른 단계에서 본 강좌를 종료하였다. 특히, C1, C2 내 종료 빈도가 35%를 웃도는 것을 보아 많은 학생이

시작 단계에서 본 강좌를 포기한 것을 알 수 있고, C12에서 강의를 종료한 학습자는 4% 미만으로 나타났다.

〈표 3〉 강의 별 시작, 종료 빈도 비율 (3% 이상)

Lecture	First in case(%)	Last in case(%)	Lecture	First in case(%)	Last in case(%)
C1_L1	63.1	17.5	C2_L6	0.5	3.2
C2_L1	16.7	11.4	C2_L7	0.2	4.4
C2_L2	3.1	6.1	C3_L3	0.3	3.8
C2_L3	1.0	5.0	C6_L1	1.4	4.2
C2_L4	0.6	3.5	C12_L1	0.6	3.8
C2_L5	0.8	3.5	Others	11.7	33.6

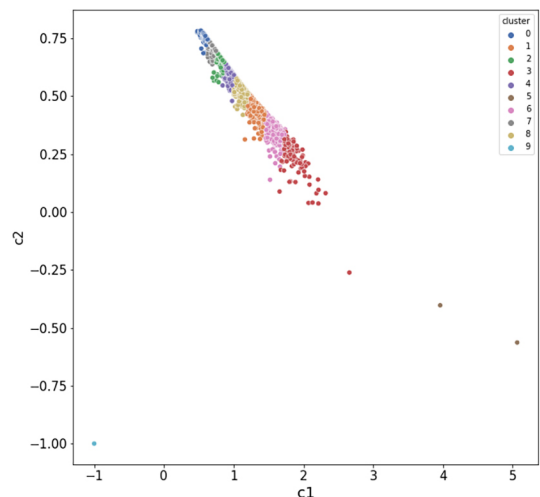
도출된 프로세스 모델의 품질 검증을 위하여 Conformance Checking을 통한 Fitness와 Cardoso Metric 및 Cyclomatic Metric을 기반한 Simplicity를 평가하였다. 분석 결과, Fitness의 수치는 0.785, Cardoso Metric과 Cyclomatic Metric은 각각 65, 458로 계산되었다. Fitness의 경우 [0, 1]의 값을 갖는 수치로 상대적으로 높아 보이지만, 실제로 많은 학생이 5개 미만의 학습활동만 수행하고 본 강좌를 포기한 것을 고려할 때 값이 과대평가된 경향이 있다. 두 개의 Simplicity 측정 결과값의 경우 직접적인 비교 및 평가가 어렵다. 다만, Cardoso Metric의 수치가 평균적으로 특정한 작업 후에 접근 가능한 작업의 수를 의미하고 이를 전체 작업 수와 비교하였을 때, 거의 유사하다는 것은 모델이 상당히 복잡하다는 것을 나타낸다.

4.3. Variant-level 프로세스 분석

패턴 관점의 프로세스 분석을 위해 먼저 Autoencoder-based trace clustering을 적용하였다.

각 학습자별 Activity profile을 만드는 Feature extraction 단계에서는 전체 작업에 대한 빈도수를 바탕으로 수행하였다. 인코더 단계에서는 98차원 (작업의 진행, 종료 고려)의 Activity profile을 축소시키기 위하여 각 layer의 차원을 각각 64, 16, 4, 2로 설정하였다. 반대로, 디코더 단계에서는 각 layer의 차원을 각각 4, 16, 64, 98로 설정하여 다시 기존 벡터의 차원으로 복원한다. 본 구조를 통해 학습한 Autoencoder의 loss 값은 0.16이며, 학습 과정에서 고차원 데이터를 가장 잘 표현하는 latent vector를 추출하여 이를 클러스터링에 활용하였다.

클러스터링 단계에서는 K-means clustering을 적용하였으며 전체 데이터의 실루엣 계수 및 군집에 대한 정성적 평가 결과를 고려하여 군집의 개수를 10으로 선정하였다. <그림 4>는 군집 결과를 반영하여 차원 축소한 데이터를 2차원에 시각화한 그림을 나타낸다. 총 8개의 군집은 여러 학생들이 고루 분포하였고 두 개의 군집은 특이한 패턴을 보이는 소규모의 학생들이 분포하였다.



〈그림 4〉 군집화 분석에 대한 시각화 결과

각 군집별 특징 파악을 위하여, 전체 강의를 대상으로 학습자별 강의 수강 패턴 분석을 수행하였다. <그림 5>는 전체 강의에 대한 각 군집별 수강 패턴 결과를 나타낸다. 본 그림의 색은 각 군집별 전체 학습자 수 대비 해당 강의의 수강 완료 학습자 수를 의미한다. 즉, 특정한 강의에 대한 수치가 1일 경우 군집 내 모든 학습자가 해당 강의를 수강하였음을 의미하며, 반대로 0일 경우 모든 학습자가 해당 강의를 수강하지 않았음을 의미한다. 본 그림 내에서 값이 1에 가까울수록 색이 짙고, 0에 가까울수록 색이 옅게 표현되었다. 분석 결과, 군집 1, 3, 5, 6은 Chapter 1부터 12에 포함된 강의를 대부분 수강한 군집으로 파악되며, 이외 군집은 이 중 일부만을 수강한 것으로 분석된다. 특히, 군집 0, 7의 경우 Chapter 1, 2의 극 초반 단계의 강의만 수강한 것을 알 수 있다.

군집화 결과에 따른 총 10개에 군집에 대하여 기초 통계 분석 결과 및 강의 완료 비율 분석 결과를 바탕으로 각 군집의 특징을 파악하였으며 이에 따른 결과는 다음과 같다.

- 군집 3, 6 (570명, 6.0%) : 본 강좌 내 모든 강의를 성실히 수강한 학습자 그룹
- 군집 1, 8 (1,058명, 11.2%) : 본 강좌 내 초반부 강의의 수강정도 대비 후반부에는 다소 떨어지는 학습자 그룹

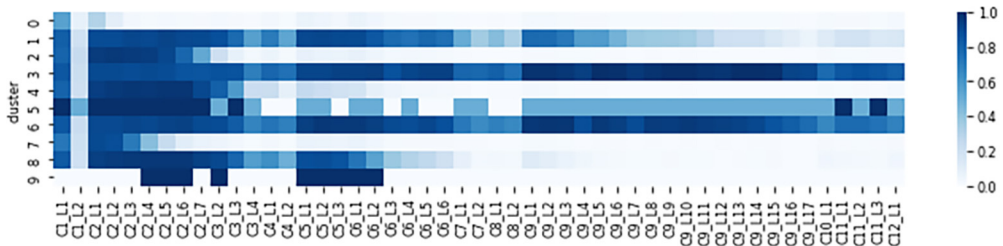
- 군집 0, 2, 4, 7 (7,817명, 82.7%) : 대다수의 학습자들이 포함되는 그룹이며 Chapter 1, 2의 극초반 강의만 수강한 후 본 강좌를 포기하는 그룹
- 군집 5, 9 (3명, 0.0%) : 특정 강의에 대한 학습 횟수가 굉장히 높은 학습자 그룹

이후, 표준 및 이상 패턴을 정의하기 위해 각 군집별로 교수자가 정의한 레퍼런스 모델과의 유사도 분석을 수행하였다. <표 4>는 이에 대한 상세 유사도 분석 결과를 나타낸다. 이를 통해, 군집 6, 3 내 학습자의 학습 흐름이 교수자가 유도한 흐름과 가장 유사하며, 이에 반해 군집 0, 5, 7 내 학습자는 거의 따르지 않음을 알 수 있다.

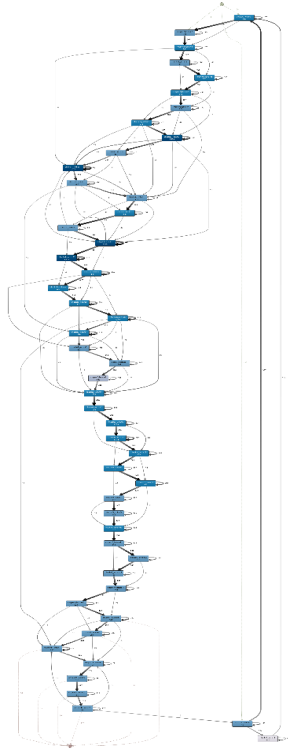
<표 4> 각 군집별 교수자 정의 모델과의 유사도 분석 결과

군집	0	1	2	3	4	5	6	7	8	9
유사도	0.03	0.43	0.14	0.54	0.18	0.09	0.67	0.09	0.27	0.18

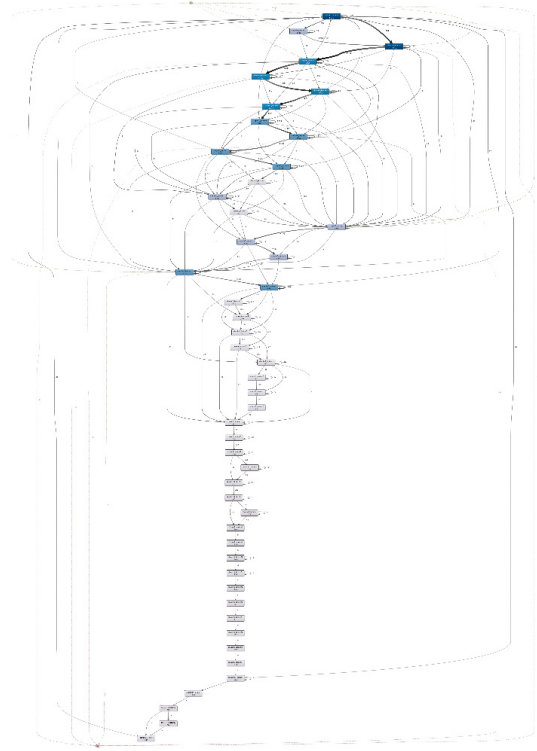
표준 패턴 분류를 위하여 한계값을 0.5로 설정하였으며, 이에 따라 군집 3, 6을 표준 패턴으로 나머지 8개의 군집을 이상 패턴으로 정의하였다. 표준 패턴 및 이상 패턴에 따른 프로세스 모델 도출 결과(노드 100%, 간선 13% 한계값 적용)는 각각 <그림 6>, <그림 7>과 같다. 두 프로세스



<그림 5> 군집 별 강의 완료 비율



〈그림 6〉 표준 패턴 기반 모델



〈그림 7〉 이상 패턴 기반 모델

모델을 비교하면 <그림 6>이 강의를 더 순서대로 수강한 경향이 있고, <그림 7>의 경우 모델이 더 복잡하고 유연한 경향이 있는 것을 알 수 있다.

4.4. Instance-level 프로세스 분석

각 학습자별 표준 혹은 이상 패턴 예측 분류 모델 구축을 위해, 표준 및 이상 패턴에 해당하는 데이터를 구축하였다. 앞서 분석한 결과를 활용하여 군집 3, 6에 해당하는 570명을 표준 패턴으로 설정하였고, 비율을 동일하게 수행하기 위해 나머지 학습자 중 570명을 랜덤 추출하여 이상 패턴으로 데이터로 정의하였다. 또한, 앞서 제시한 것처럼 마스킹을 진행하였으며 각각

10%, 20%, 30%, 40%, 50%, 100%의 흐름만 추출하였다. 이는 각 학습자별 전체 흐름 중 시작 작업부터 해당 비율만큼의 일부 흐름만을 바탕으로 분류 모델을 정의하는 것을 의미한다. 본 연구에서는 타 기법 대비 높은 모델 정확도를 보인 로지스틱 회귀 모델을 활용하였으며, <표 5>는 이에 대한 모델 구축 결과를 나타낸다.

〈표 5〉 각 학습자별 표준/이상 패턴 분류 모델 구축 결과

	10%	20%	30%	40%	50%	100%
Accuracy	0.71	0.83	0.86	0.93	0.93	1.00
F1 score	0.73	0.84	0.86	0.93	0.93	1.00
AUC-ROC	0.72	0.83	0.92	0.96	0.96	1.00

기본적으로 학습자별 이벤트 기록을 더 많이 활용할수록 더 높은 정확도를 갖는 것을 알 수 있으며, 모든 흐름을 활용할 경우 모델의 정확도 1.0 임을 알 수 있다. 전체 흐름 중 시작 기준 10%의 적은 양의 데이터만 활용할 경우 약 0.71 정도의 정확도를 갖는다. 이는 각 학습자별로 전체 학습과정 대비 10%의 학습이 진행되었을 때, 해당 학습자가 표준 패턴 혹은 이상 패턴을 따르는 지를 약 0.71의 정확도로 분류할 수 있음을 나타낸다. 이러한 정보를 활용하여 학습자의 학습에 대한 개입 시점을 파악할 수 있으며, 예를 들어 본 강좌의 경우, 0.85의 분류 정확도가 의미 있는 수치로 가정할 때 학습자별로 약 30%의 시점에서 이상 패턴으로 분류되는 학습자에게 개입하여 원활한 학습의 가이드라인을 제공할 수 있게 된다.

5. 결론 및 시사점

본 연구에서는 온라인 교육 오픈 플랫폼 내 학습 패턴을 파악하기 위하여 프로세스 마이닝을 활용하였으며, 모델, 패턴, 인스턴스 관점으로 이를 분석하였다. 모델 관점에서는 학습 패턴에 대한 전반적인 이해 및 전체 흐름을 표현하는 프로세스 맵 도출, 패턴 관점에서는 군집화를 통한 세부 패턴 도출 및 표준, 이상 패턴 정의 및 특징 파악, 인스턴스 관점에서는 개별 학습자별 표준, 이상 패턴 분류 모델 개발을 각각 목적으로 하였다. 또한, 실제 약 1만 명의 학습 로그를 바탕으로 본 연구에서 제시하는 방법에 적용하여 각 관점 별 유의미한 분석 결과를 도출하였으며, 이에 대한 시사점은 다음과 같다.

모델 관점 분석에서는 온라인 교육 오픈 플랫폼의 주요 특징인 자기 규제형 학습(Self-regulated learning)을 확인할 수 있었다. 특히, 많은 학습자

가 소개 강의부터 시작하는 것에 반해 마지막 강의는 서로 다른 형태를 보였으며, 이는 강의를 처음부터 끝까지 모두 듣는 것이 아닌 본인이 필요로 하는 일부만 수강하는 것을 의미한다. 이와 더불어, 도출된 프로세스 모델이 상당히 복잡한 것을 보면 상대적으로 자율성이 떨어지는 대학 등의 비대면 강의와 매우 다른 특징을 갖는 것을 알 수 있다. 패턴 관점 분석에서는 군집화를 통해 학습자들의 학습 패턴을 구체화할 수 있었으며, 대다수의 학습자가 교수가 설정한 흐름을 따르지 않는 것을 확인할 수 있었다. 특히, 약 83%의 학습자가 극초반 강의만 수강한 후 본 강좌를 포기하는 것을 고려하였을 때, 학습자의 학습 성취를 올릴 수 있는 다른 방법이 고려될 필요가 있다. 인스턴스 관점 분석에서는 서로 다른 마스킹에 따른 예측 모델을 도출할 수 있었으며, 많은 학습자가 극초반만 강의를 듣는 이상 패턴에 해당하여 상대적으로 높은 정확도를 보였다. 그럼에도 불구하고, 본 연구 결과는 학습자의 학습에 대한 개입 시점을 파악한다는 점에서 의의를 갖는다.

본 연구는 아래와 같은 한계점을 갖는다. 먼저, 학습 패턴에 대한 다각도 분석을 주목적으로 하여 분석 결과 및 문제점에 대한 개선점 도출 방법을 제시하지는 않는다. 즉, 시뮬레이션 모델 구축 및 분석을 통한 시나리오 분석과 Action-oriented process mining을 통한 문제 해결 등 방법 제시가 필요하다. 또한, 본 연구는 기계학습 관련 한 강의에만 적용하여 다양한 사례 연구 결과를 도출하지 못하였다. 현재 확보된 데이터로 컴퓨터 공학, 인문학 등 다양한 데이터를 보유하고 있어 이를 적용하여 추가적인 사례 연구를 도출할 필요가 있다. 이와 더불어, 각 분야별 학습 패턴 도출을 통해 하나의 강의를 아닌 크로스 도메인 관점에서의 접근이 필요시 된다.

참고문헌(References)

[국내 문헌]

- 권영욱. (2013). 빅데이터를 활용한 맞춤형 교육 서비스 활성화 방안연구. *지능정보연구*, 19(2), 87-100.
- 정준희. (2012). 통계적 접근법을 기초로 하는 지능형 교육 지원 시스템. *지능정보연구*, 18(1), 109-123.
- 강영식, 정진우, 심선영. (2023). RPA 로그 마이닝 기반 프로세스 자동화 현황 분석 - 중소기업 대상 실증 연구. *지능정보연구*, 29(1), 265-288.
- 김혜경, 배성아. (2019). Felder-Silverman 학습양식에 따른 영어 튜터링 프로그램의 튜터와 튜티의 학습성취도와 만족도 간의 관계 연구. *교양교육연구*, 3(2), 115-140.
- 손세창, 윤한영. (2022). 포스트 코로나19 시대 공항의 디지털 트랜스포메이션을 통한 공항 운영에서 혁신적 비즈니스 모델 적용 연구. *한국산학기술학회논문지*, 23(4), 483-492.
- 허묘연. (2020). 언택트 교육의 시대: 온라인 교육을 위한 준비. *차세대컨버전스정보서비스기술 논문지*, 9(3), 315-325.

[국외 문헌]

- Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist*, 57(10), 1439-1459.
- Graf, S., Viola, S.R. & Leo, T. (2007). In-Depth Analysis of the Felder-Silverman Learning Style Dimensions *Journal of Research on Technology in Education*, 40(1), 79 - 93. <https://doi.org/10.1080/15391523.2007.10782498>
- Bogarín, A., Cerezo, R. & Romero, C. (2017) A survey on educational process mining *WIREs*

Data Mining and Knowledge Discovery, 8(1). <https://doi.org/10.1002/widm.1230>

- Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline *American Behavioral Scientist*, 57(10), 1380-1400
- Gašević, D., Dawson, S., Rogers, T. & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success *Internet and Higher Education*, 68-84. <https://doi.org/10.1016/j.iheduc.2015.10.002>
- Mukala, P., Buijs, J., Leemans, M. & Van Der Aalst, W. (2015). Exploring Students' Learning Behaviour in MOOCs Using Process Mining. *BPM Reports*, 1510, Eindhoven University of Technology, Eindhoven, The Netherlands. <https://research.tue.nl/nl/publications/exploring-students-learning-behaviour-in-moocs-using-process-mini>
- Hachicha, W., Ghorbel, L., Champagnat, R., Zayani, C.A. & Amous, I. (2021). Using Process Mining for Learning Resource Recommendation: A Moodle Case Study. *Procedia Computer Science*, 853-862. <https://doi.org/10.1016/j.procs.2021.08.088>
- Schoor, C. & Bannert, M. (2012). Exploring regulatory processes during a computer-supported collaborative learning task using process mining *Computers in Human Behavior*, 1321-1331. <https://doi.org/10.1016/j.chb.2012.02.016>
- Juhaňák, L., Zounek, J. & Rohlíková, L. (2019). Using process mining to analyze students' quiz-taking behavior patterns in a learning management system *Computers in Human Behavior*, 496-506. <https://doi.org/10.1016/j.chb.2017.12.015>
- Van Der Aalst, W. (2016). *Process mining: data science in action* (Vol. 2). Heidelberg: Springer.
- Augusto, A., Conforti, R., Dumas, M., La Rosa,

- M., Maggi, F. M., Marrella, A., ... & Soo, A. (2018). Automated discovery of process models from event logs: review and benchmark. *IEEE transactions on knowledge and data engineering*, 31(4), 686-705.
- Cho, M., Song, M., Comuzzi, M., & Yoo, S. (2017). Evaluating the effect of best practices for business process redesign: An evidence-based approach based on process mining techniques. *Decision Support Systems*, 104, 92-103.
- Cardoso, J., Mendling, J., Neumann, G., & Reijers, H. A. (2006, September). A discourse on complexity of process models. In *Business process management workshops* (Vol. 4103, pp. 117-128).
- McCoy, C. & Shih P. C. (2016). Teachers as Producers of Data Analytics: A Case Study of a Teacher Focused Educational Data Science Program, *Journal of Learning Analytics*, 3(3), 193-214.
- 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.
- Muslim, A., Chatti, M. A., & Guesmi, M. (2020). Open learning analytics: a systematic literature review and future perspectives. *Artificial Intelligence Supported Educational Technologies*, 3-29.
- Wang, Y., Mu, C., Li, X., & Yang, Y. (2019, August). An Applied Research on Big Data Analysis and Mining Technology in Education. In *2019 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC)* (pp. 116-120). IEEE.
- Chen, L., Lu, M., Goda, Y., & Yamada, M. (2019). Design of Learning Analytics Dashboard Supporting Metacognition. *International Association for Development of the Information Society*.
- Fleur, D. S., van den Bos, W., & Bredeweg, B. (2020). Learning analytics dashboard for motivation and performance. In *Intelligent Tutoring Systems: 16th International Conference, ITS 2020, Athens, Greece, June 8 - 12, 2020, Proceedings 16* (pp. 411-419). Springer International Publishing.
- Baneres, D., Rodríguez-Gonzalez, M. E., & Serra, M. (2019). An early feedback prediction system for learners at-risk within a first-year higher education course. *IEEE Transactions on Learning Technologies*, 12(2), 249-263.
- Han, J., Kim, K. H., Rhee, W., & Cho, Y. H. (2021). Learning analytics dashboards for adaptive support in face-to-face collaborative argumentation. *Computers & Education*, 163, 104041.
- Kokoç, M., & Altun, A. (2021). Effects of learner interaction with learning dashboards on academic performance in an e-learning environment. *Behaviour & Information Technology*, 40(2), 161-175.
- Khalil, M., & Ebner, M. (2017). Clustering patterns of engagement in Massive Open Online Courses (MOOCs): the use of learning analytics to reveal student categories. *Journal of computing in higher education*, 29, 114-132.
- Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in human behavior*, 89, 98-110.
- Trcka, N., & Pechenizkiy, M. (2009, November). From local patterns to global models: Towards domain driven educational process mining. In *2009 Ninth international conference on intelligent systems design and applications* (pp. 1114-1119). IEEE.

Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and learning*, 9, 161-185.

Cairns, A. H., Gueni, B., Fhima, M., Cairns, A., David, S., & Khelifa, N. (2014, July). Towards

custom-designed professional training contents and curriculums through educational process mining. In *The fourth international conference on advances in information mining and management* (pp. 53-58).

Abstract

Toward understanding learning patterns in an open online learning platform using process mining

Taeyoung Kim* · Hyomin Kim* · Minsu Cho**

Due to the increasing demand and importance of non-face-to-face education, open online learning platforms are getting interests both domestically and internationally. These platforms exhibit different characteristics from online courses by universities and other educational institutions. In particular, students engaged in these platforms can receive more learner autonomy, and the development of tools to assist learning is required. From the past, researchers have attempted to utilize process mining to understand realistic study behaviors and derive learning patterns. However, it has a deficiency to employ it to the open online learning platforms. Moreover, existing research has primarily focused on the process model perspective, including process model discovery, but lacks a method for the process pattern and instance perspectives. In this study, we propose a method to identify learning patterns within an open online learning platform using process mining techniques. To achieve this, we suggest three different viewpoints, e.g., model-level, variant-level, and instance-level, to comprehend the learning patterns, and various techniques are employed, such as process discovery, conformance checking, autoencoder-based clustering, and predictive approaches. To validate this method, we collected a learning log of machine learning-related courses on a domestic open education platform. The results unveiled a spaghetti-like process model that can be differentiated into a standard learning pattern and three abnormal patterns. Furthermore, as a result of deriving a pattern classification model, our model achieved a high accuracy of 0.86 when predicting the pattern of instances based on the initial 30% of the entire flow. This study contributes to systematically analyze learners' patterns using process mining.

Key Words : Learning pattern analysis, process mining, online education, process model discovery, autoencoder

Received : June 12, 2023 Revised : June 22, 2023 Accepted : June 22, 2023

Corresponding Author : Minsu Cho

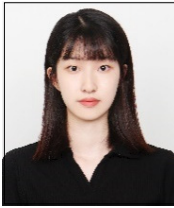
* School of Information Convergence, Kwangwoon University
** Corresponding author: Minsu Cho
School of Information Convergence, Kwangwoon University
20 Kwangwoon-ro, Nowon-gu, Seoul 01897, Korea
Tel: +82-2-940-8454, Fax: +82-2-940-5651, E-mail: mcho@kw.ac.kr

저 자 소개



김태영

광운대학교 정보융합학부 학사과정에 재학중이며, 주요 관심분야는 프로세스 마이닝, 텍스트 마이닝, 딥러닝 등이다.



김효민

광운대학교 정보융합학부 학사과정에 재학중이며, 주요 관심분야는 프로세스 마이닝, 텍스트 마이닝, GNN 등이다.



조민수

울산과학기술원에서 박사학위를 취득하였으며, 현재 광운대학교 정보융합학부 조교수로 재직중이다. Decision Support Systems, International Journal of Medical Informatics, IEEE Transactions on Semiconductor Manufacturing 등 우수 저널에 논문을 게재하였으며, 주요 관심분야는 프로세스 마이닝, 딥러닝, 프로세스 재설계, 교육 데이터 분석, RPA 등이다.