

논문 2019-56-11-6

설명 가능한 인공지능 기반의 프로세스 마이닝 분석 자동화 연구

(Explainable Artificial Intelligence based
Process Mining Analysis Automation)

정 찬 일*, 이 후 진**

(Chanyil Jung and Hoojin Lee[©])

요 약

업무 혁신을 위한 경영기법 중 하나인 프로세스 마이닝은 정보 시스템 로그에서 실행 프로세스 모델을 작성하는 것을 자동화하였으나, 프로세스 내용을 분석하여 문제에 대한 원인을 분석하는 영역에서는 여전히 분석가의 관련 경험과 업무 도메인에 대한 지식에 많이 의존하고 있다. 특히 대부분의 도출된 프로세스 모델이 매우 복잡하기 때문에 사람의 인식 능력의 한계로 인해 정확한 분석이 불가능하여 프로세스 마이닝 알고리즘에서는 추상화 과정을 거치는데, 이 과정에서 왜곡이 발생하거나 정작 중요한 이슈를 놓치게 될 가능성이 있다. 본 연구에서는 먼저 머신러닝 모델 중 LSTM 알고리즘을 이용한 프로세스 예측 모델을 개발하여 프로세스의 수행결과를 예측하였고 다시 이 예측모델에 설명 가능한 인공지능 기법을 적용하여 분석함으로써 최종적으로 프로세스 수행결과에 가장 큰 영향을 미치는 원인을 탐색하는 작업을 자동화하였다. 이를 통해 프로세스 마이닝에 대한 원인 분석 과정에서 인간의 직관과 업무지식에 대한 의존을 획기적으로 낮추었으며 정량적이고 신속한 프로세스 분석이 가능하게 되었다.

Abstract

Process mining is one of the business management techniques for business innovation. While it automates the creation of execution process models in the information system log, it still relies on the analyst's relevant experience and knowledge of the business domain in analyzing the process content and the root cause of the problem. Since most of the derived process models are very complicated, an accurate analysis is impossible due to the limitations of human cognition ability, and most of the process mining algorithms go through an abstraction process. There is a possibility that distortions may occur in this process, or important issues may be missed. Hence, in this study, we attempted to predict the process results by using the process prediction model incorporating a LSTM algorithm of machine learning, and then to apply an explainable artificial intelligence(XAI) technique. Thus, this study has drastically reduced the reliance on the human intuition and knowledge in the root cause analysis of process mining, enabling the quantitative and rapid analysis.

Keywords : Process mining, Process analysis automation, Machine learning, eXplainable AI

I. 서 론

업무에서 정보시스템의 활용이 높아지고 분석 관련

기술이 발전하면서 업무혁신을 위한 경영기법으로써 프로세스 마이닝의 가치가 증가하고 있다^[1~2]. 프로세스 마이닝은 시스템의 로그를 분석하여 실행 프로세스를 도식화하는 프로세스 발견(Discovery)과 이 프로세스를 분석하는 절차로 진행된다^[2]. 그중 프로세스 발견은 다양한 알고리즘과 도구가 개발되어 상당부분 자동화되어 있으나, 발견된 프로세스를 분석하여 이슈가 발생하는 요소를 찾아내어 개선방안을 도출할 때는 분석가의 판

* 학생회원, ** 평생회원, 한성대학교 스마트융합컨설팅학과
(Department of Smart Convergence Consulting,
Hansung University)

© Corresponding Author(E-mail : hjlee@hansung.ac.kr)

Received ; August 4, 2019

Revised ; September 2, 2019

Accepted ; October 20, 2019

단을 지원하기 위한 기능 위주로 구성되어 있기 때문에 프로세스 마이닝에 대한 경험과 해당 업무 도메인에 대한 지식이 필수적이다^[3]. 시스템 로그를 사용해서 도출해낸 프로세스 모델은 기본적으로 매우 복잡^[3~4]하므로 프로세스 마이닝 도구들은 사람이 인식 가능한 형태로 추상화할 수 있는 다양한 필터와 통계 기능을 제공하고 있다. 그러나 이런 추상화 과정에서 왜곡이 발생하거나 정작 중요한 예외 프로세스가 제외되기도 하고 프로세스 분석은 사람의 직관에 의지할 수밖에 없기 때문에 분석가의 역량에 따라 다른 결과가 도출되는 한계가 있었다.

본 연구에서는 머신러닝 중 LSTM(Long Short Term Memory) 알고리즘^[5] 기반으로 프로세스 결과를 예측하고, 설명 가능한 인공지능(eXplainable Artificial Intelligence)^[6] 알고리즘을 활용하여 전체 프로세스를 대상으로 수행결과에 영향을 주는 주요 업무 단계(Activity)와 원인이 되는 속성(Attribute)을 자동으로도출함으로써 보다 빠르고 정확하게 프로세스 이슈에 대한 원인을 파악하였다.

본 연구에서 제안한 인공지능 기술에 기반을 둔 이슈 근본 원인 파악 자동화는 프로세스 분석 시 개개인의 역량에 따라 다른 결과가 나오는 것을 사전에 배제할 수 있으며, 사람의 인식 능력 한계 때문에 주요 프로세스만을 대상으로 추상화^[4]하여 원인을 분석할 수밖에 없는 문제를 해결하고 프로세스 분석에 걸리는 시간과 노력을 대폭 감소시킬 것으로 기대된다.

II. 선행 연구 및 한계

1. 프로세스 마이닝

프로세스 마이닝은 표 1과 같이 개별 정보시스템에서 기록된 로그 데이터를 수집하여 프로세스 모델을 자동으로 생성(Process Discovery)하고 실제 프로세스가 모델을 기반으로 수행되는지를 평가(Conformance

표 1. 프로세스 마이닝 구성 요소
Table 1. Components of Process Mining.

Component	Explanation	Automation rate
Discovery	Generate a process model from event log data	High
Conformance	Compare with process model and execution process log	Medium
Enhancement	Improve the performance of the existing model	Low

Checking)하여 프로세스의 품질을 개선(Enhancement)하는 기법^[1~2]으로, 프로세스 발견에 비해 적합성 평가 및 확장을 위한 분석 기능은 비교적 자동화 비율이 낮다.

특히 프로세스 발견을 통해 도식화된 프로세스 모델은 그림 1처럼 스파게티 모델이 되어 사람이 확인할 수 없으므로 프로세스 마이닝 도구인 ProM, DISCO, PPM^[1] 등에서 제공하는 다양한 필터 기능을 이용하여 프로세스를 추상화하는 작업을 수행하고 있다. 그림 1은 BPI Challenge 2019^[7]의 데이터 전체를 모델화한 것이고 그림 2는 그중 가장 자주 출현하는 5%의 Activity만으로 추상화한 모델이다.

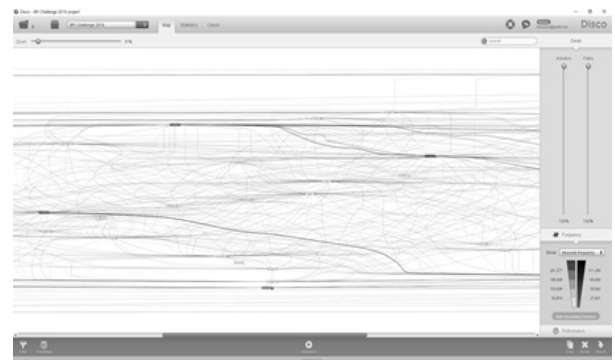


그림 1. 프로세스 마이닝 결과 모델의 복잡성
Fig. 1. Complex of Process Mining Model.

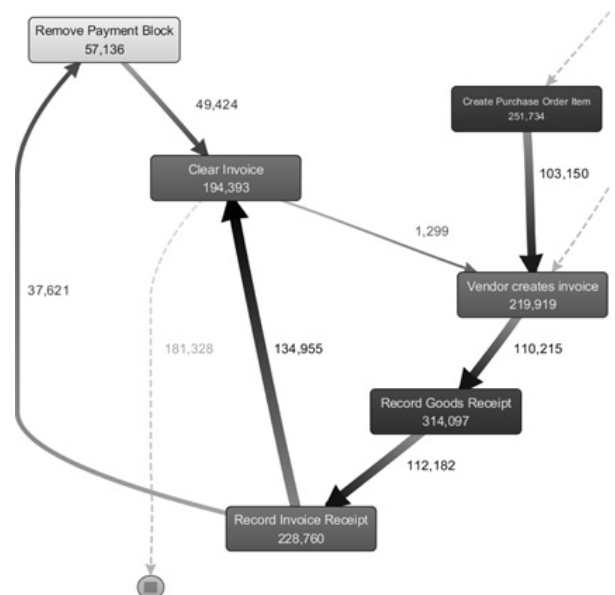


그림 2. 프로세스 모델 추상화 (Activity 필터 5% 적용)
Fig. 2. Process Model Abstraction (5% Activity Filter).

이렇게 추상화된 모델의 경우 사람의 직관을 통한 분석에는 유리하나 주로 문제가 되는 예외 프로세스를 놓치게 되고 데이터가 왜곡되어 정확한 분석에 한계가 발

생할 수 있다는 문제가 있다. 이렇게 기존의 프로세스 마이닝 분석 방식은 업무에 대한 지식과 프로세스 마이닝 경험이 있는 전문가가 꼭 참여하여야 하며^[3~4] 분석가의 개인적인 역량에 따라 다른 결과가 나오기도 하며 분석 기간도 오래 걸린다는 문제점이 있다.

2. 프로세스 예측

프로세스 마이닝을 통해 도출된 실행 프로세스를 기반으로 새로운 프로세스가 발생될 때 해당 프로세스의 진행 상황을 모니터링하여 향후 프로세스의 진행 결과를 사전에 예측하려는 연구는 초기에 통계적인 기법을 통해 남은 프로세스 수행 시간을 예측^[8]하는데서 Naïve Bayes, SVM 등의 다양한 알고리즘을 기반으로 발전^[9~10]하였으며, 최근에는 사용량 예측이나, 데이터 프리패치 등 다양한 분야의 예측문제에서 정확도가 높다고 알려진 LSTM 알고리즘^[11~12] 기반의 프로세스 예측^[13~16]을 통해 정확도가 높아졌다.

많은 프로세스 예측 알고리즘 관련 논문에서 표준적으로 사용된 BPI Challenge 2012 데이터 셋을 기반으로 알고리즘별 정확도를 계산했을 때 표 2와 같이 현재 89.5% 까지 정확도가 향상^[15]되었으며, 데이터 셋은 다르지만 국내 모 화장품 기업의 프로세스를 기준으로 정확도가 90% 이상^[14]으로 발표되기도 하였다.

표 2. 알고리즘별 프로세스 예측 정확도^[13, 15]
Table 2. Process prediction accuracy by algorithm.

Algorithm		Accuracy	Precision	Recall	F1-Score
SVM		0.817	0.856	0.822	0.817
Naïve Bayes		0.617	0.633	0.612	0.555
LSTM	Evermann et al. ^[13]	0.824	0.852	0.824	0.817
	J. Theis et al. ^[15]	0.895	0.895	0.895	0.887

※ BPI Challenge 2012 데이터 기준

한편 프로세스의 다음 단계가 아니라 Case 별로 KPI를 예측하는 연구^[16]가 진행되기도 하였다.

3. 설명 가능한 인공지능(XAI)

설명 가능한 인공지능 기술은 “인공지능의 개별 의사결정에 대한 설명을 제공하고, 사용자가 인공지능 시스템의 전반적인 강점 및 약점을 이해하도록 도와주기 위해 고안”^[6, 17, 18]되었으며, 인공지능 기반의 예측에 대해 주요 영향을 미치는 요소를 찾아내어 예측결과에 대한

설명을 해주는 용도^[19]로도 유용하게 사용되고 있다.

XAI를 위한 알고리즘 및 도구들은 현재 DARPA에서 진행 중인 Explainable Learners, Psychology of Explanation 프로젝트^[6] 외에도 IBM의 AI Fairness 360^[20], LIME(Local Interpretable Model-agnostic Explanations)^[17] 등이 있으며, 본 연구에서 사용한 LIME 알고리즘은 예측 모델 자체를 해석하는 것이 아니라 표본 입력값을 교란해서 예측이 어떻게 바뀌는지 확인하는 방식이다^[17]. 다시 말하면 입력값을 조금씩 바꿨을 때 모델의 예측결과가 크게 바뀌면, 해당 변수는 중요한 변수라고 인식할 수 있다고 가정한다.

III. 연구 방법 및 결과

본 연구는 그림 3과 같이 ① 프로세스 마이닝에서 사용되는 XES 표준 포맷 데이터를 기반으로 전처리를 수행하고, ② 머신러닝 중 LSTM 알고리즘을 이용하여 예측모델을 개발하여 프로세스 수행 결과를 예측하고, ③ 이 예측 모델을 XAI 모델 중 LIME 알고리즘으로 분석하여 최종 프로세스 결과에 가장 큰 영향을 주는 요소를 파악하는 절차를 통해 프로세스 이슈 원인 분석을 자동화하였다.

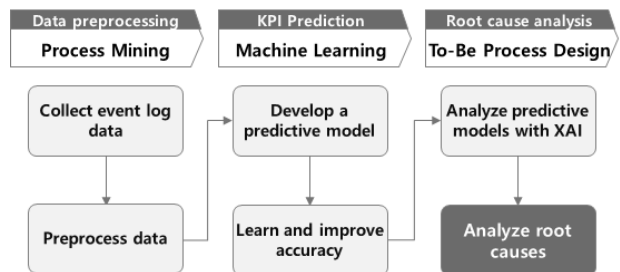


그림 3. 프로세스 이슈 원인 분석 자동화 절차
Fig. 3. Process Issue Cause Analysis Automation.

그림 4와 같이 A, B, C 라는 프로세스가 수행되었을 때 LSTM 알고리즘을 통해 최종 프로세스 수행결과가 D라는 것을 높은 정확도로 예측할 수 있다면, 그림 5와 같이 XAI 알고리즘을 통해 그 예측에 가장 큰 영향을 준 핵심 원인이 B라는 것을 알 수 있다.

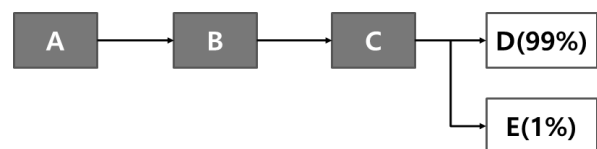


그림 4. 프로세스 예측
Fig. 4. Process Predict.

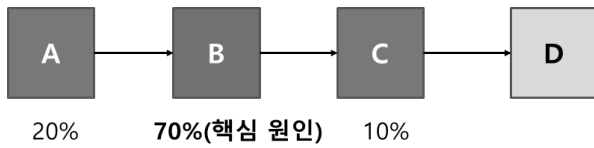


그림 5. 프로세스 결과에 대한 원인 파악

Fig. 5. Process Root Cause Analysis.

이때 기본적인 전제는 프로세스 예측 알고리즘의 정확도가 충분히 신뢰 가능한 수준이어야 한다는 것이다.

1. 사용 데이터 전처리

선행 연구^[13, 15]에서는 공통적으로 주요 데이터로 사용하고 있는 BPI Challenge 2012 데이터 셋을 사용하여 예측 정확도에 대한 검증을 수행하고 프로세스 결과에 대한 원인을 분석하였다.

BPI 2012 데이터는 네덜란드의 한 지자체에서 발생한 업무 프로세스로 개인정보는 비식별화 처리 되어 있다^[21]. XES 포맷^[22]으로 구성된 데이터를 변환하여 Case 별 Activity 명을 순서대로 정렬한 후 LSTM 모델에 적용하기 위하여 자연어 처리와 같이 워드 임베딩을 진행하였다. 이때 각 Activity를 1개 워드로 맵핑하였으며 Embedding Dimension은 2개만 지정하였다. 또한, 추가적인 프로세스 분석을 위해 프로세스 마이닝 솔루션인 DISCO에서 제공하는 제조업체의 Sales 프로세스 예제를 이용하였다.

2. 프로세스 마지막 Activity 예측 및 원인 분석

BPI 2012 데이터를 Python 기반의 Tensorflow와 Keras를 활용하여 LSTM 모델을 개발한 후 학습을 진행하였다. Evermann의 선행연구^[13]는 각 단계에서 다음 Activity를 연속하여 예측하도록 모델을 설계하였으나 본 연구의 목적은 각 Case 별로 최종 프로세스 결과 1개만을 예측하는 것이므로 Input 데이터는 마지막 프로세스 결과 이전까지의 전체 Activity이며, Validation 데이터는 Case 별 최종 Activity 결과로 설계하였다.

표 3과 같이 실험에 사용된 LSTM 모델은 Memory Cell 개수를 가장 긴 Case 숫자에 맞춰 174개로 지정하였고, 활성화 함수는 Softmax를 손실함수는 binary crossentropy를 사용하였으며 학습한 결과 그림 6 및 그림 7과 같이 Accuracy가 0.994로 매우 양호하게 나타났다.

이 예측 모델에 설명 가능한 인공지능 기법 중 LIME 알고리즘을 적용하여 임의의 프로세스 Case를

표 3. 예측 모델 요약

Table 3. Prediction Model Summary.

Layer	Output Shape	Number of Parameter
Embedding	(None, 172, 1)	25
LSTM	(None, 174)	122,496
Dense	(None, 25)	4,375
Total parameters : 126,896		
Trainable parameters : 126,896		

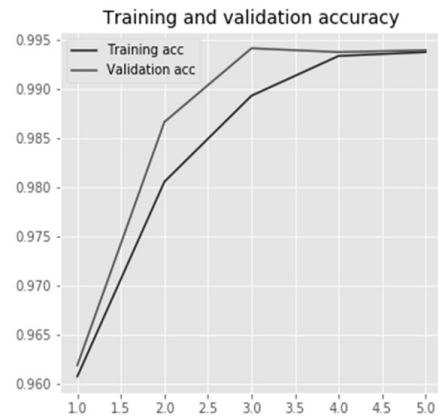


그림 6. 예측 모델 정확도 학습 곡선

Fig. 6. Predict Model's Accuracy Learning Curve.

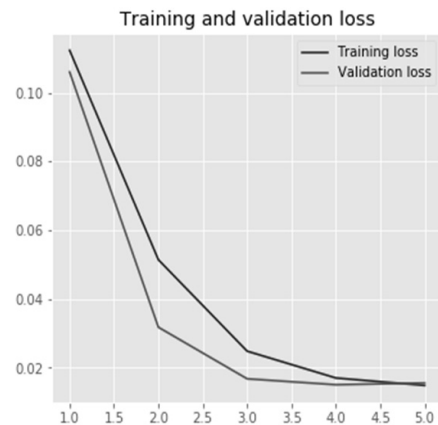


그림 7. 예측 모델 손실값 학습 곡선

Fig. 7. Predict Model's Loss Learning Curve.

분석한 결과 그림 8과 같이 마지막 Activity가 ‘신청서 처리’(네덜란드어, afhandelen)일 확률이 0.99로 거의 확정적이었으며(실제 마지막 Activity와 일치함) 가장 높은 영향을 미친 Activity는 바로 전 단계인 ‘신청서 확인’(네덜란드어, valideren aanvraag)으로 0.09의 영향 정도를 가지며 이전 Activity로 내려갈수록 그 영향을 미치는 정도가 낮아짐을 보인다.

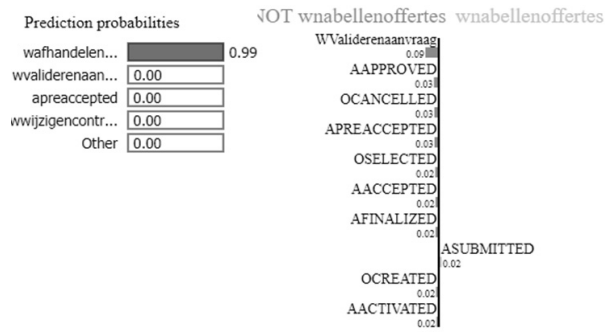


그림 8. 프로세스 결과 원인
Fig. 8. Process Result Cause.

3. 성과지표 예측 및 원인 파악

프로세스의 마지막 Activity를 예측하는데 사용되는 주요 원인이 그 직전 Activity에 가장 큰 영향을 받는다는 점에서 프로세스의 마지막 Activity 예측은 해당 프로세스를 분석하는데 별다른 도움이 되지 못하였다. 이에 프로세스 마이닝 전문 기업인 PMIG에서 제공한 제조업체의 Sales 프로세스를 이용하여 마지막 Activity가 아닌 성과지표를 예측하고 그 원인을 파악하는 실험을 진행하였다.

프로세스 마이닝을 위한 기본적인 데이터 구조인 XES 표준에서는 그림 9와 같이 Activity 명칭과 더불어 Case ID, Time stamp, Resource(수행자)가 필수 항목이며 데이터 특성에 따라 제품코드, 고객번호 등 부가적인 데이터를 갖는다^[1, 2, 22].

Case ID	Activity	End Time	Resource	제품코드(CLTMC)	고객번호
S2016053S0050-6	등록중	2016/05/20 17:42:40.000	2010007	109001052	0
S2016053S0050-6	등록중	2016/05/20 17:42:55.000	2010007	109001052	0
S2016053S0050-5	등록중	2016/05/20 17:42:40.000	2010007	109001051	0
S2016053S0050-5	등록중	2016/05/20 17:42:55.000	2010007	109001051	0
S2016053S0050-4	등록중	2016/05/20 17:42:40.000	2010007	109001050	0
S2016053S0050-4	등록중	2016/05/20 17:42:55.000	2010007	109001050	0
S2016051S0015-2	등록중	2016/05/21 03:05:40.000	2011003	104401009	10156
S2016051S0015-2	등록완료	2016/05/21 03:05:47.000	2011003	104401009	10156
S2016051S0015-2	주문확정	2016/05/21 03:05:49.000	2011003	104401009	10156
S2016051S0015-2	등록중	2016/05/21 03:05:59.000	2011003	104401009	10156
S2016051S0015-2	등록완료	2016/05/21 03:07:03.000	2011003	104401009	10156

그림 9. 프로세스 마이닝 데이터
Fig. 9. Process Mining Data.

프로세스 마이닝 도구인 DISCO를 사용해 프로세스를 확인하면 그림 10과 같이 1,348개의 Case와 9,036개의 Event를 가지며 평균 수행일은 16일임을 알 수 있다. 여기에서는 Sales 프로세스의 성과지표를 적기처리율로 선정하고 각 Case 중 출하 완료까지 30일 이내에 완료된 건은 합격(Good), 30일을 초과할 경우 불합격(Bad)으로 판정하여 지도학습을 위한 Validation 데이터로 사용하였으며, 데이터의 속성 중 Activity, Resource, Client, Product Model 등을 One Hot Encoding을 통해 전처리하여 전체 830개의 Feature를 구성하였다.

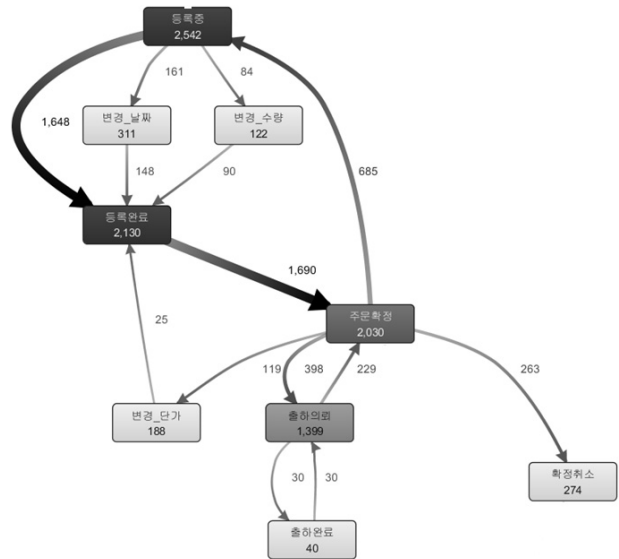


그림 10. 판매 프로세스
Fig. 10. Sale Process.

표 4. 판매 프로세스에 영향을 주는 항목 분석 결과
Table 4. Result of Analysis of the Sales Process.

Rank	No.	Type	Value	Impact
1	t-91	Employee	2016039	0.0010012
2	t-39	Activity	Change date	-0.0009252
3	t-44	Product	100001149	-0.0008557
4	t-47	Customer	10166	-0.0008454
5	t-59	Employee	2015041	0.0008348
6	t-1	Customer	10250	0.0008340
...	...		...	
338	t-18	Product	100001175	0.00000082
339	t-25	Customer	10026	-0.00000069

이 데이터를 이전 실험에서 진행한 LSTM 기반 예측 모델을 통해 학습한 결과 정확도가 0.947로 합리적으로 나타났다. 이를 다시 LIME 알고리즘을 통해 분석한 결과 표 4와 같이 339개의 결과에 영향을 주는 특징을 도출하였으며 91번째 Activity를 수행한 작업자 사번이 '2016039'가 아닐 경우 결과가 Good(기한내 수행)이 되는데 0.00100123의 영향을 주고 있으며, 39번째 Activity가 '날짜 변경'인 것이 Bad(기한 경과)로 판단하는데 0.0009252 만큼의 영향을 끼치고 있는 것을 알 수 있다.

해당 프로세스는 총 94개의 단계로 수행되는데 각 단계별 Activity, 고객번호, 작업자, 제품 코드의 값이 예측결과에 어떤 영향을 미치고 있는지 확인해보면 표 5와 같다. 각 단계별 양의 값은 결과가 Good일 확률을

표 5. 실행 단계별 영향도 분석

Table 5. Analysis impact by Process Steps.

No.	Activity	Customer	Employee	Product
0	0.0001316	0.0004146	0.0004371	-0.0013154
1	0.0002056	0.0008775	-0.0002436	0.0002996
2			-0.0006521	0.0000884
3	0.0002146	-0.0007512	-0.0011972	0.0007904
4	-0.0005618	-0.0001290	-0.0001325	0.0012798
5	0.0005183	0.0006102		-0.0007852
6		-0.0009863		0.0009447
7	-0.0000578		0.0001634	-0.0001420
...				
93		-0.0003568		
94			0.0003262	
Total	-0.0046516	0.0043283	0.0009988	-0.0003569

높이고 음의 값은 결과가 Bad일 확률을 높이는데 예측 결과에 대해 전체적으로 가장 많은 영향을 미치는 것은 절댓값이 0.0046516으로 가장 높은 “Activity명”과 0.0043283인 “고객번호” 순인 것으로 확인되었다.

IV. 연구 결과 및 의미

첫 번째 실험을 통해 프로세스의 마지막 Activity를 예측하는 알고리즘에 영향을 주는 요소를 파악하였으나 예외 없이 마지막 직전 Activity의 상태에 매우 크게 영향을 받기 때문에 구체적인 업무 프로세스의 원인 분석에는 활용하기 어려웠다.

그러나 판매 프로세스에 대한 두 번째 실험을 통해서 구체적으로 프로세스의 성과지표를 선정하고 모든 속성을 대상으로 이 성과지표에 미치는 영향을 평가함으로써 구체적인 원인을 빠르게 파악할 수 있었다.

분석 결과에서 91번째 Activity를 수행한 작업자 사번과 39번째 수행된 Activity명 등이 결과에 가장 큰 영향을 미친다는 것과 전체적으로 Activity와 고객번호가 영향을 준다는 것을 확인하였다. 이 결과 자체만으로는 업무와의 구체적인 인과관계까지 파악하긴 힘들지만, 사번이 ‘2016039’인 작업자의 업무 패턴을 다른 작업자와 비교해보거나 업무 담당자와 협의하여 ‘날짜 변경’ Activity를 수행하는 화면에 문제가 있는지를 확인하는 식으로 활용할 수 있다. 경험이 있는 전문가라도 프로세스 이슈의 원인이 되는 Activity를 찾기까지 많은 시간

이 걸린다는 것을 고려한다면 이 기법은 프로세스 분석에 대한 시간 및 품질을 대폭 향상한 것으로 판단된다.

V. 결 론

본 연구에서는 LSTM 알고리즘 통한 프로세스 예측 기법과 XAI 알고리즘의 결합으로부터 프로세스 결과에 가장 많은 영향을 주는 실제 요인을 효과적으로 파악할 수 있다는 것을 확인하였다. 이를 통해 기존 프로세스 마이닝 방식으로는 해결하기 어려웠던 업무지식이 없거나 프로세스 분석 경험이 적은 경우에도 알고리즘 적용으로 쉽고 빠르게 프로세스 이슈에 대한 원인 분석이 가능하게 되었다.

향후에는 실제 프로세스 마이닝 기반의 이슈 원인 분석 및 업무개선 사례 데이터를 이용하여 ‘설명 가능한 인공지능 기반의 프로세스 이슈 원인 분석 자동화’ 기법이 업무지식과 경험을 갖춘 사람이 분석한 것과 의미하게 일치하는 결과를 도출하는지 추가적인 실험을 진행하여 그 실효성을 보다 명확하게 입증하고자 한다.

REFERENCES

- [1] W. M. P. van der Aalst, A. Adriansyah and M. Song et al., “Process mining manifesto”, Lecture Notes in Business Information Processing, Springer, vol. 99, pp. 169-194, Aug. 2011.
- [2] W. M. P. van der Aalst, “Process mining: discovery, conformance and enhancement of business processes”, Springer-Verlag, Apr. 2011.
- [3] Y. Kang and B. Kim, “Process mining for business”, Hannarae, 2016.
- [4] D. Jeon and H. Bae, “Detection of API(anomaly process instance) based on distance for process mining”, Journal of the Korean Institute of Industrial Engineers, vol. 41, no. 6, pp. 540-550, Nov. 2015.
- [5] F. A. Gers, J. Schmidhuber, and F. Cummins, “Learning to forget: Continual prediction with LSTM”, Neural Computation, vol. 12, no. 10, pp. 2451-2471, Oct. 2000.
- [6] D. Gunning, “Explainable artificial intelligence (XAI)”, Defense Advanced Research Projects Agency (DARPA), n.d. Web, <https://www.darpa.mil/program/explainable-artificial-intelligence>, [Accessed: Jul. 15, 2019]
- [7] B. F. van Dongen, “Dataset BPI challenge 2019. 4TU.Centre for research data”, <https://doi.org/>

- 10.4121/uuid:d06aff4b-79f0-45e6-8ec8-e19730c248f1, [Accessed: Jul. 25, 2019]
- [8] W. M. P. Van Der Aalst, M. H. Schonenberg and M. Song, "Time prediction based on process mining", *Information Systems*, vol. 36, no. 2, pp. 450-475, Jun. 2011.
- [9] G. T. Lakshmanan, D. Shamsi and Y. N. Doganata et al., "A markov prediction model for data-driven semi-structured business processes", *Knowledge and Information Systems*, vol. 42, no. 1, pp. 97-126, Jan. 2015.
- [10] M. Ceci, P. F. Lanotte and F. Fumarola et al. "Completion time and next activity prediction of processes using sequential pattern mining", In *Proceedings of the 17th International Conference on Discovery Science, Lecture Notes in Computer Science*, Springer, vol. 8777, pp. 49-61, Oct. 2014.
- [11] J. Choi and J. Kim, "Analysis of water consumption data from smart water meter using machine learning and deep learning algorithms", *Journal of The Institute of Electronics and Information Engineers*, vol. 55, no. 7, pp. 31-39, Jul. 2018.
- [12] K. Song, K. Kim, and S. Choi, "Implementation of hardware data prefetching technique using machine learning technique", *Journal of The Institute of Electronics and Information Engineers*, vol. 56, no. 3, pp. 11-23, Mar. 2019.
- [13] J. Evermann, J. R. Rehse and P. Fettke, "A deep learning approach for predicting process behavior at runtime," *PRAISE Workshop at the 14th International Conference on BPM*, Rio de Janeiro, Brazil, Sep. 2016.
- [14] J. Kim, S. Yoon and B. Lee, "Exploring process prediction based on deep learning: Focusing on dynamic recurrent neural networks" *The Journal of Information Systems*, vol. 27, no. 4, pp. 115-128, Dec. 2018.
- [15] J. Theis and H. Darabi, "DREAM-NAP: Decay replay mining to predict next process activities", *arXiv preprint arXiv:1903.05084*, Mar. 2019.
- [16] Y. Kang, H. Lee and B. Kim, "Exploring the prediction of timely stocking in purchasing process using process mining and deep learning", *Information Systems Review*, vol. 20, no. 4, pp. 25-41, Dec. 2018.
- [17] M. T. Ribeiro, S. Singh and C. Guestrin, "Why should I trust you?: Explaining the predictions of any classifier", In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1135-1144, San Francisco, California, USA, Aug. 2016.
- [18] W. Samek, T. Wiegand and K. R. Müller, "Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models", *arXiv preprint arXiv:1708.08296*, Aug. 2017.
- [19] I. Giurciu and A. Schumann, "Explainable failure predictions with RNN classifiers based on time series data", *arXiv preprint arXiv:1901.08554*, Jan. 2019.
- [20] R. K. E. Bellamy, K. Dey and M. Hind et al., "AI Fairness 360: An extensible toolkit for detecting, understanding, and mitigating unwanted algorithmic bias", *arXiv preprint arXiv:1810.01943*, Oct. 2018.
- [21] B. F. van Dongen, "Dataset BPI challenge 2012. 4TU.Centre for research data", <https://doi.org/10.4121/uuid:3926db30-f712-4394-aebc-75976070e91f>, [Accessed: Jul. 25, 2019]
- [22] XES Standard, IEEE 1849-2016, 2016.

저 자 소 개



정 찬 일(학생회원)

1997년 동국대학교 윤리학과
학사 졸업

2009년 고려대학교 디지털정보
공학과 석사 졸업

2017년~현재 한성대학교 스마트
융합컨설팅학과 박사 과정

<주관심분야: 프로세스 마이닝, 인공지능, XAI, RPA>



이 후 진(평생회원)

1997년 서울대학교 전기공학부
학사 졸업

2002년 The University of Texas
at Austin ECE 석사 졸업

2007년 The University of Texas
at Austin ECE 박사 졸업

2009년~현재 한성대학교 IT융합공학부 및 스마
트융합컨설팅학과 교수

<주관심분야: 통신 및 네트워크, 멀티미디어 신호
처리, 사이버보안>