

## 딥러닝 기반의 프로세스 예측에 관한 연구: 동적 순환신경망을 중심으로\*

김정연\*\* · 윤석준\*\*\* · 이보경\*\*\*\*

### <목 차>

I. 서론	IV. 실험방법과 결과
II. 관련 연구 고찰	4.1 실험방법
2.1 이벤트 로그의 구조	4.2 실험결과
2.2 딥러닝과 순환신경망	V. 결 론
2.3 딥러닝을 활용한 프로세스 예측 연구	참고문헌
III. 접근법	<Abstract>
3.1 동적 순환신경망을 활용한 구조	
3.2 접근법 구현	

### I. 서론

예측 분석을 자사의 비즈니스 프로세스에 활용할 수 있는 기업은 차별화된 가치를 창출할 수 있다(LaValle et al., 2011). 이에 따라, 프로세스의 미래 행동을 예측할 수 있는 능력은 기업의 핵심역량이 되고 있다(Breuker et al., 2016). 예를 들어, 구매한 자재의 지연 또는 조기 입고를 예측할 수 있다면 기업은 사전에 필요한 조치를 취함으로써 지연 입고에 따른 생

산 결품과 조기 입고에 따른 장기 재고의 발생을 막을 수 있다.

최근에 정적 순환신경망(RNN: Recurrent Neural Network) 형태의 딥러닝을 이용한 프로세스 예측 연구가 큰 관심을 받고 있고, 우수한 결과를 내고 있다(예, Evermann et al., 2016; 2017a; 2017b; Tax et al., 2016). 관련 연구는 문장에서 다음 단어를 예측하기 위해 인공신경망을 자연어 처리에 적용한 방식(예, Sutskever et al., 2011)을 채택하고 있다. 구체적으로 말하면 관련 연구는 ‘이벤트 로그’를 ‘문서’, ‘케이

\* 본 연구는 서울시 산학연 협력사업(SC170043) 지원에 의해 수행되었음.

\*\* (주)피엠아이, jykim@pmig.co.kr(주저자)

\*\*\* (주)피엠아이, roy941124@pmig.co.kr

\*\*\*\* (주)피엠아이, bklee@pmig.co.kr(교신저자)

스(프로세스 인스턴스)’를 ‘문장’, ‘이벤트’를 ‘단어’에 대응시켜 정적 순환신경망을 이용한 프로세스 예측 접근법을 개발하고 검증했다.

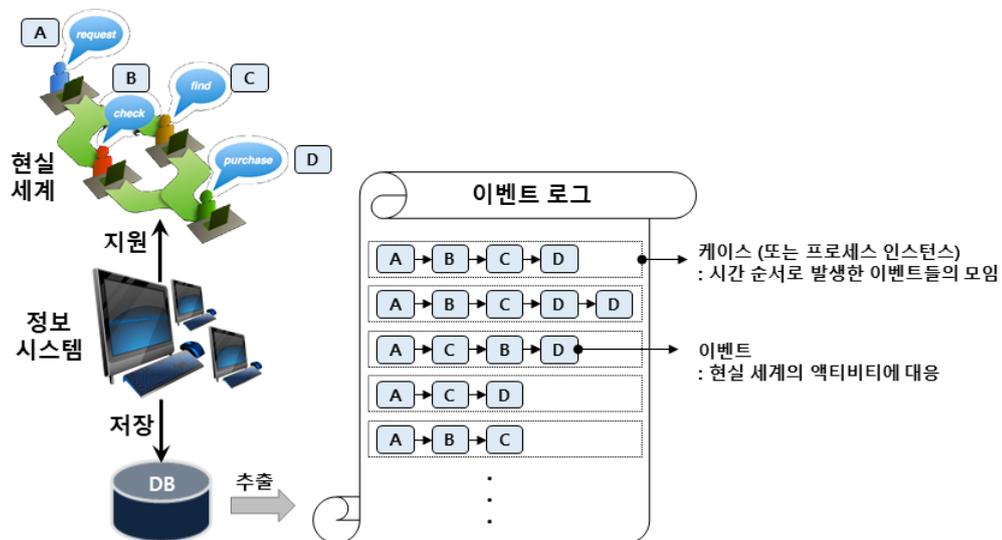
그러나 정적 순환신경망을 활용한 프로세스 예측 접근법은 프로세스의 실제 행동을 제대로 반영하지 못하고, 특정 프로세스에 적용할 때마다 많은 노력과 시간을 요구한다는 한계를 지닌다. 이러한 한계를 극복하기 위해 본 연구는 동적 순환신경망 형태의 딥러닝을 활용하는 차별화된 접근법을 제안할 것이다. 제안된 접근법은 국내 중소기업의 판매 프로세스를 기록한 실제 데이터를 활용하여 검증되었다. 특정 프로세스 인스턴스의 마지막 이벤트(즉, 액티비티)를 예측하는 작업에서 제안된 접근법을 활용한 예측 정확도(96.29%)는 기존 접근법을 활용한 예측 정확도보다 테스트 데이터 기준으로 1.62% 더 높았다. 기존 접근법을 활용한 예측 정확도가 이미 94.67%를 달성한 가운데 본 연구의 접근법이 이를 1 ~ 2% 정도 더 개선했다는 것은 매우 우수한 결과라고 판단된다.

본 논문의 구성은 다음과 같다. 2장에서 순환 신경망의 입력 데이터로 활용될 이벤트 로그의 구조를 설명하고, 딥러닝과 순환신경망의 개념 및 딥러닝(특히, 정적 순환신경망)을 활용한 프로세스 예측 연구를 고찰할 것이다. 3장에서는 본 연구가 제안하는 접근법을 설명하고 이를 구현한 방안을 제시할 것이다. 4장에서는 실험 방법과 결과를 설명할 것이다. 본 연구는 딥러닝의 문제점으로 언급되는 과적합(overfitting)을 막고, 실험 결과의 신뢰성을 확보하기 위해 10겹 교차검증(10-fold cross validation) 기법을 적용했다. 마지막으로, 본 연구의 시사점과 한계점, 향후 연구방향을 5장에서 토론할 것이다.

## II. 관련 연구 고찰

### 2.1 이벤트 로그의 구조

ERP(Enterprise Resource Planning), WMS



<그림 1> 이벤트 로그의 구조

(Workflow Management System), ITSM(IT Service Management) 등의 정보시스템은 현실 세계에서 업무를 수행할 때마다 발생하는 이벤트 데이터를 기록하고 있다(한진영 등, 2016). 이벤트는 수행된 활동이나 의사결정, 또는 관심 사안의 발생 등을 나타낸다(Breuker et al., 2016). 각 이벤트는 발생시점을 나타내는 타임스탬프와 특정 프로세스 인스턴스(또는 케이스)에 속함을 표시하는 케이스 아이디 속성을 가진다(van der Aalst, 2016). 또한 각 이벤트는 현실 세계에서 수행된 특정 액티비티에 대응한다. <그림 1>에 나타난 것처럼, 이벤트 로그는 시간 순서로 발생한 이벤트들의 모임인 케이스들의 집합으로 구성된다(van der Aalst, 2016).

흥미로운 점은 프로세스 마이닝은 이러한 이벤트 로그를 활용해서 실제 발생했던 업무흐름을 반영하는 프로세스 모델을 발견하거나 프로세스가 의도한 대로 수행되었는가를 검증하는 회고 분석(retrospective analysis)에 초점을 두고 있다(강영식, 이보경, 2016; van der Aalst, 2016). 이에 반해, 프로세스 예측의 핵심 기술인 순환신경망은 다음 단계의 이벤트를 예측하기 위해 동일한 이벤트 로그를 학습에 활용한다(Evermann et al., 2016).

## 2.2 딥러닝과 순환신경망

### 2.2.1 딥러닝

인공신경망(artificial neural networks)은 인간 뇌의 정보처리 과정을 수학적 모델링을 통해 모사한 모형이다(Schmidhuber, 2015). 인공신경망은 입력층(input layer), 은닉층(hidden layer), 출력층(output layer)으로 구성된다. 각

층에 포함된 뉴런(또는 노드)들은 가중치(weight, 연결강도를 의미함)를 통해 연결되어 있다(안성만, 2016; 트란 광 카이, 송사광, 2017; 송현정, 이석준, 2018; 김대희 등, 2018). 은닉층과 출력층의 각 노드는 가중치와 이전 단계 노드 값의 선형결합에 적용한 비선형 활성화 함수(activation function)를 통해 출력 값을 결정한다. 인공신경망 학습의 목적은 출력층에서 계산한 출력 값과 실제 출력 값 간의 차이를 최소화하는 가중치를 찾는 데 있다(LeCun et al., 2015).

기본적으로 딥러닝은 기존 인공신경망에 은닉층의 수를 증가시킨 심층신경망(deep neural networks) 혹은 심층망(deep networks) 구조를 지닌다. 이러한 심층망의 장점은 기존의 신경망에 비해 더 많은 은닉층을 사용함으로써 데이터에 대한 표현 능력을 크게 증가시킨다는 것이다(최희열, 민윤홍, 2015). 사실, 심층망에 관한 아이디어는 이전부터 존재해 왔다. 그러나 심층망을 위한 효과적인 학습방법의 부재로 신경망은 큰 주목을 받지 못하고 있었다. 2006년 사이언스지에 발표된 Hinton과 Salakhutdinov (2006)의 논문은 사전학습(pre-training)이라는 개념을 제안함으로써 심층망의 학습 가능성을 보여주었다. 그 이후 다양한 학습 알고리즘들이 제안되고, 학습에 활용될 수 있는 데이터의 크기가 커지고, 클라우드나 GPU(Graphic Processing Unit) 기반의 컴퓨팅 역량이 폭발적으로 증가했다. 이에 따라, 최근에 학계와 산업계는 딥러닝에 큰 관심을 기울이고 있다.

심층망에서 은닉층을 여러 개 가져가는 것은 기존에 활용된 신경망의 단순 확장으로 보일 수도 있다. 그러나 이러한 확장은 이미지나 문

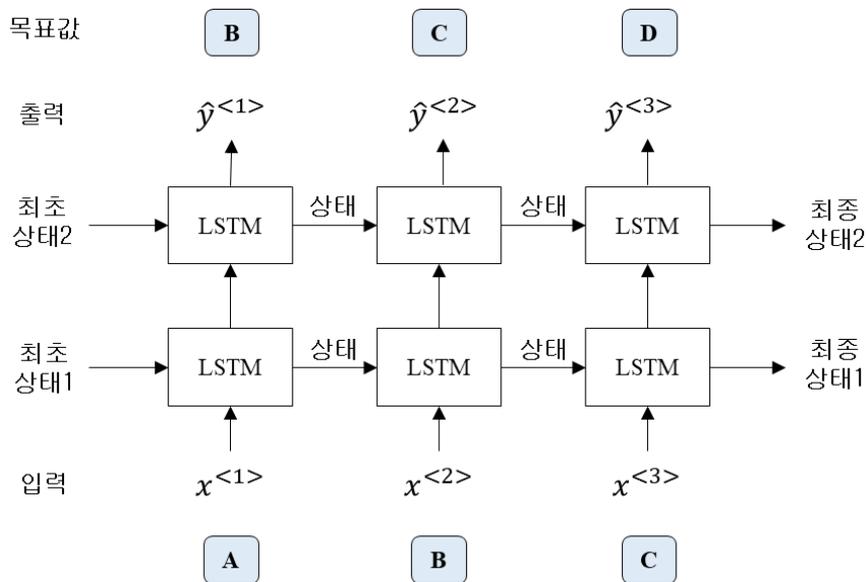
자 인식 및 자연어 처리 분야 등에서 패러다임의 변화를 가져올 만큼 크나큰 혁신으로 간주되고 있다(김진백, 김유일, 2000; 최희열, 민운홍, 2015). 딥러닝의 혁신은 크게 두 가지로 요약될 수 있다. 첫째, 해당 분야에 속한 전문가의 지식 없이도 딥러닝은 데이터로부터 자동적으로 필요한 특징들(features)을 추출할 수 있다. 둘째, 딥러닝은 특징 추출과 분류를 하나의 모델에 통합됨으로써 예측 성능을 극대화할 수 있다(Kang and Park, 2018).

### 2.2.2 순환신경망

딥러닝 기술인 순환신경망(RNN: Recurrent Neural Network)은 음성, 영상, 문서와 같은 순서 데이터(sequential data) 처리에 적합한 구조를 가지고 있다(Graves, 2012). 그러나 자연어 처리 등에서 상호 중요한 관계를 갖는 정보가 시간적으로 멀리 떨어지는 경우가 자주 발생하

면 순환신경망은 이러한 상관관계를 효과적으로 학습하지 못한다는 한계를 지닌다(Bengio et al., 1994). RNN의 이러한 한계를 극복하기 위해 새로운 순환신경망인 장단기 메모리(LSTM: Long Short-term Memory)(Hochreiter and Schmidhuber, 1997)와 GRU(Gated Recurrent Unit)(Chung et al., 2014)가 개발되었다. LSTM과 GRU는 최근 음성인식, 필기인식, 기계번역을 비롯한 자연어 처리에서 탁월한 성능을 보이며 딥러닝의 핵심 방법론으로 자리 잡고 있다(Evermann et al., 2017). 이런 이유로 본 연구도 RNN 대신에 LSTM과 GRU를 채택했다.

순환신경망 구조의 예가 <그림 2>에 제시되어 있다. 이 구조는 2층으로 구성되어 있고, 3단계로 펼쳐진 LSTM 셀을 가지고 있다(즉, 시퀀스 길이가 3임). LSTM 셀 대신에 GRU나 RNN 셀을 활용할 수도 있다. 사실, 층의 수나 시퀀스 길이(sequence length), 셀에 적용한 순환신경



<그림 2> 순환신경망 구조의 예 (2층의 3단계로 펼쳐진 LSTM 셀을 가진 다대다의 RNN 구조)

망의 종류(예, LSTM이나 GRU, RNN) 등과 같은 하이퍼파라미터(hyperparameter)의 값을 최적으로 결정하는 것은 쉬운 일이 아니다. 최적의 하이퍼파라미터 값을 결정하기 위해서는 수많은 반복 실험 과정을 거쳐야 한다(LeCun et al., 2015).

순환신경망은 일대다, 다대일, 다대다 등의 유형을 가질 수 있다. 예를 들어, 모차르트가 작곡한 악보를 학습한 순환신경망은 음악을 만들 수 있다. 이 경우에는 일대다 유형의 순환신경망을 적용한다. 문장에 포함된 단어가 사람의 이름인가를 찾는 순환신경망은 <그림 2>와 같은 다대다 유형을 가진다. 많은 단어로 구성된 사용자 후기로부터 평점을 예측하는 순환신경망은 다대일 유형을 가진다.

한편, <그림 2>의 순환신경망 구조는 펼쳐진 단계(즉, 시퀀스 길이)가 고정된 정적 순환신경망이다. 최근에 입력 데이터의 각 사례(example)마다 시퀀스 길이를 변화시킬 수 있는 동적 순환신경망이 제공되어 이를 활용한 신경망 설계와 구현이 가능하게 되었다<sup>1)</sup>. 동적 순환신경망을 활용한 국내외 연구가 부재한 상황이나, 이 신경망은 음성, 영상, 문서와 같은 순서 데이터 처리에서 우수한 결과를 낼 것으로 기대된다.

### 2.3 딥러닝을 활용한 프로세스 예측 연구

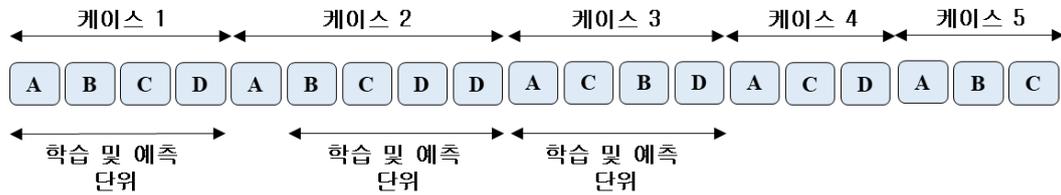
대부분의 프로세스 예측 연구는 다양한 기법을 활용하여 프로세스 결과(예, 완료까지 남은

시간)를 예측하는 것에 초점을 두고 있었다(예, van der Aalst et al., 2011). 그런데 최근에 딥러닝을 활용하지 않고 프로세스의 다음 단계를 예측하고자 한 일부 연구(예, Breuker et al., 2016; Lakshmanan et al., 2015; Unuvar et al., 2016)가 수행되었다. 이러한 연구는 명확한 모델 표현에 기반을 두고 있고, 회귀나무(regression trees)와 같은 선형기법을 활용하고 있다. 이에 반해, 정적 순환신경망 형태의 딥러닝을 활용한 프로세스 예측 연구(예, Evermann et al., 2016; 2017a; 2017b; Tax et al., 2016)는 순환신경망 내에 암묵적으로 학습된 프로세스를 활용하고, 예측변수와 목표변수 간 비선형 관계를 가정하고 있다. 그러므로 딥러닝을 활용한 프로세스 예측 연구는 (프로세스) 모델 표현을 위한 제약을 줄일 수 있고, 예측 정확도를 개선할 수 있다. 이런 이유로 딥러닝을 활용한 프로세스 예측 방법이 새롭게 혁신적인 접근법으로 간주되고 있다(Evermann et al., 2017a).

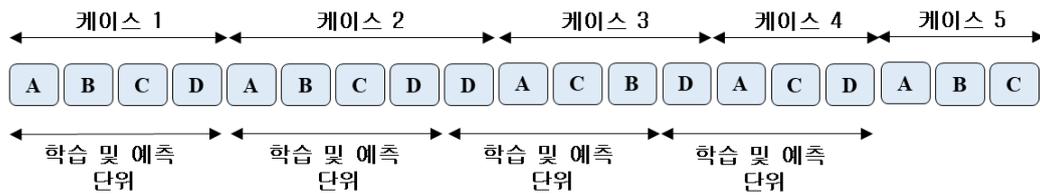
본 연구의 목표인 각 케이스의 마지막 이벤트(즉, 액티비티)를 예측한다고 가정해 보자. 정적 순환신경망을 활용한 프로세스 예측 방법은 학습과 예측을 수행하기 위해 <그림 1>에 표시된 5개의 케이스 수행과 관련된 이벤트 기록을 <그림 3>과 같이 변환<sup>2)</sup>한다. 만약 <그림 2>의 순환신경망 구조를 활용한다면 시퀀스 길이가 3이므로 예측될 마지막 이벤트를 포함하여 마지막 이벤트보다 먼저 수행된 3개의 이벤트가 학습과 예측에 활용된다. 그러므로 케이스 1과 3의 이벤트 기록은 학습과 예측에 온전히 활용

1) [https://www.tensorflow.org/api\\_docs/python/tf/nn/dynamic\\_rnn](https://www.tensorflow.org/api_docs/python/tf/nn/dynamic_rnn)

2) 정적 순환신경망을 이용한 학습과 예측을 위한 데이터 변환 시에 각 케이스의 끝을 나타내는 <EoC>(즉, End of Case)를 별도의 이벤트로 추가할 수 있음. 그러나 본 연구는 각 케이스의 마지막 이벤트를 예측하는 것을 목표로 하므로 <EoC>를 추가하지 않음.



<그림 3> 정적 순환신경망을 이용한 학습과 예측을 위한 첫 번째 데이터 변환 방법



<그림 4> 정적 순환신경망을 이용한 학습과 예측을 위한 두 번째 데이터 변환 방법

될 수 있다. 그러나 케이스 2의 경우에는 첫 번째 이벤트인 A가 빠져야만 한다. 왜냐하면 마지막 이벤트인 D를 예측하기 위해 시퀀스 길이('3')에 맞는 이벤트 B, C, D만 활용될 수 있기 때문이다. 케이스 4와 5의 경우에는 예측될 이벤트 D나 C를 제외하면 시퀀스 길이와 맞지 않기 때문에 케이스 4와 5의 전체 이벤트 기록이 학습이나 예측에서 빠져야만 한다.

사실 기존 연구(예, Evermann et al., 2016; 2017a; 2017b)는 <그림 1>에 표시된 5개의 케이스 수행과 관련된 이벤트 기록을 <그림 4>와 같이 변환한 후에 이를 이용하여 학습하고 예측한다. 이 경우에도 케이스 5의 전체 이벤트 기록이 학습이나 예측에서 빠져야만 한다. 또한 학습 및 예측의 단위가 프로세스 수행 단위인 케이스가 아니라, 임의로 설정한 시퀀스 길이이므로 마지막 이벤트 예측과 같은 예측 작업에서 이러한 데이터 변환 방법은 예측 성능을 저하시킬 수 있다. 본 연구의 실험 결과, <그림 4>의 데이터 변환 방법을 이용하면 마지막 이벤트 예측의 정

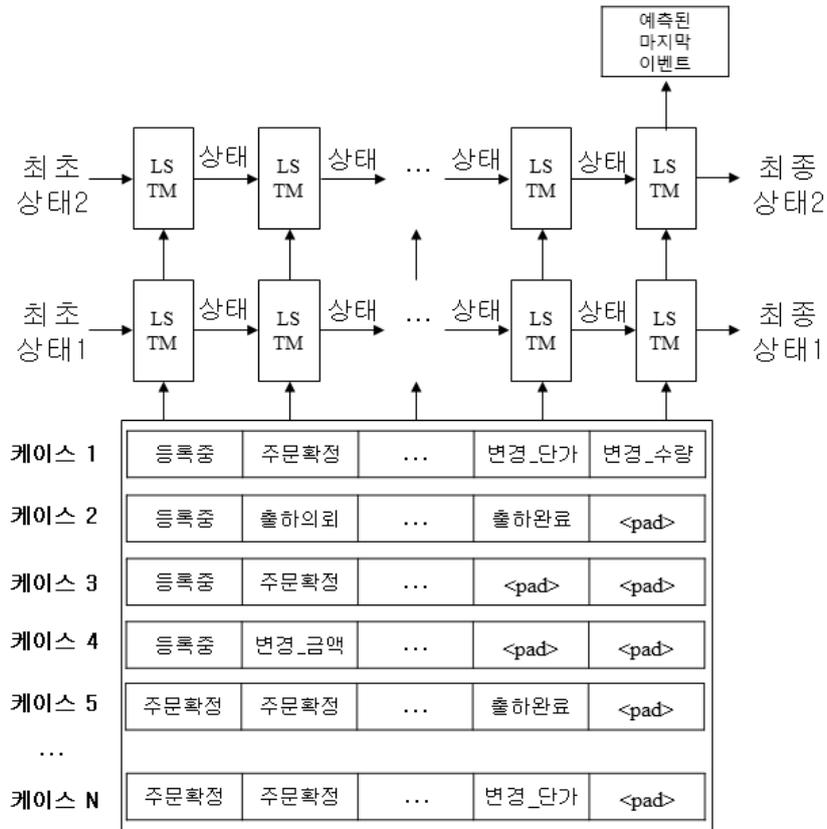
확도가 크게 떨어지는 것으로 확인되었다.

예측 정확도를 높일 수 있는 첫 번째 데이터 변환 방법(<그림 3> 참조)을 따르는 경우에도 정적 순환신경망 모형에 대한 학습을 진행할 때 케이스 4와 5의 수행을 기록한 전체 이벤트 기록과 케이스 2의 수행을 기록한 일부 이벤트 기록을 학습 데이터에 포함할 수 없었다. 그러므로 이러한 데이터를 이용하여 학습한 모형은 현실 세계의 실제 프로세스 행동을 정확히 반영하지 못한다는 한계를 지닌다. 정적 순환신경망에 기반을 둔 프로세스 예측 모형의 이러한 한계는 결국 예측 정확도에도 영향을 줄 것이다.

### Ⅲ. 접근법

#### 3.1 동적 순환신경망을 활용한 구조

앞에서 언급한 것처럼, 정적 순환신경망 형태의 딥러닝을 활용한 프로세스 예측 방법이



<그림 5> 동적 순환신경망을 이용한 구조

새롭고 혁신적인 접근법으로 간주되고 있다. 그러나 이러한 접근법은 실제 프로세스의 행동을 정확히 반영하지 못한다는 한계를 지닌다. 이러한 한계를 극복하기 위해 본 연구는 동적 순환신경망을 활용하는 구조를 제안하고 구현했다.

본 연구에서 제안한 동적 순환신경망을 활용하는 구조는 <그림 5>와 같다. 정적 순환신경망과 달리, 동적 순환신경망은 시퀀스 길이에 상관없이 모든 케이스를 활용할 수 있다. 다만 전체 케이스 중에서 이벤트 수가 가장 많은 케이스에 맞추어 나머지 케이스들에 대해서는 <그림 5>와 같은 패딩(padding) 작업을 해 주어야

한다. 본 연구는 각 케이스의 마지막 이벤트를 예측하는 것이 목표이기 때문에 다대일 유형의 순환신경망을 채택했다.

예측된 마지막 이벤트는 대상 프로세스의 전체 액티비티 종류의 수만큼 가지는 차원으로 선형 투영시켜 결과물을 출력하고, 교차 엔트로피(cross-entropy)를 통해 손실 값(loss)을 계산한다. 손실 값을 계산한 이후에는 역전파 알고리즘(back propagation algorithm)을 통해 오류를 수정하면서 학습을 진행한다.

<그림 5>의 구조를 활용하는 접근법이 기존 접근법에 비해 가지는 장점을 정리하면 다음과

같다. 첫째, 모든 케이스의 업무 수행과 관련된 이벤트 데이터가 학습과 예측에 활용될 수 있다. 둘째, 임의의 시퀀스 길이 설정에 따라 각 케이스의 특정 부분을 학습하는 것이 아니라, 실제 프로세스의 행동을 반영한 케이스 단위로 학습할 수 있다. 마지막으로, 최적의 예측 정확도를 달성할 수 있는 임의의 시퀀스 길이를 찾기 위한 노력과 시간을 줄일 수 있다. 특히, 마지막 장점은 제안한 접근법을 다양한 실세계 비즈니스 프로세스의 행동 예측에 적용할 때 큰 도움을 줄 수 있다.

### 3.2 접근법 구현

우리는 오픈소스 딥러닝 프레임워크인 텐서플로우<sup>3)</sup>에 기반을 두고 파이썬 프로그래밍 언어를 이용하여 <그림 5>에 제시한 구조를 구현했다. 텐서플로우는 손실함수 측면에서 그래프를 구성하는 모든 텐서들의 경사도(gradient)를 계산하고 다양한 최적화 함수를 제공한다(Evermann et al., 2017b). 학습의 목적은 손실함수를 최소화 하는 것이다. 텐서플로우를 활용한 응용프로그램은 텐서 연산을 사용하여 최적화 함수를 통해 손실함수를 줄인다. 손실함수는 최소화될 그래프의 노드이다. 예를 들어, 범주형 변수의 교차 엔트로피와 같은 계산된 출력을 대상 목표변수의 실제 값과 비교하여 손실(오차)를 계산하고, 이를 줄여나가는 연산을 수행한다.

개발된 응용프로그램은 크게 CSV 형식의 데이터를 파싱하는 모듈과 하이퍼파라미터 설정(configuration) 모듈, 동적순환신경망 학습 모

듈, 예측 모듈로 구성된다. 파싱 모듈은 학습에 사용할 데이터와 제외할 데이터를 구분한다. 예를 들어, 각 케이스는 최소 하나의 입력 이벤트와 예측에 활용될 목표 이벤트를 가져야 하므로 프로세스 길이가 최소 2가 되어야 한다. 따라서 프로세스 길이가 1인 케이스는 학습 및 예측에 사용될 수 없다. 이와 함께, 최대의 프로세스 길이보다 작은 길이를 가지는 케이스에 대해서는 패딩 처리를 해야 한다(<그림 5> 참조).

하이퍼파라미터 설정 모듈은 배치 크기, 에포크 수, 활성화 함수와 최적화 함수 등의 설정에 활용된다. 하이퍼파라미터의 설정에 따라 결과가 크게 달라질 수 있다. 그러므로 가장 효과적인 하이퍼파라미터 설정을 찾기 위해서는 반복 과정을 거쳐야 한다. 학습 모듈은 패딩 처리된 입력데이터를 활용하여 학습을 진행한다. 10겹 교차검증 기법을 사용하기 위해 입력데이터와 목표데이터를 인덱싱하는 과정이 함께 진행된다. 예측된 마지막 이벤트(즉, 액티비티)는 대상 프로세스의 전체 액티비티 종류의 수만큼 가지는 차원으로 선형 투영시켜 결과물을 출력하고, 교차 엔트로피(cross-entropy)를 통해 손실 값(loss)을 계산한다. 손실 값을 계산한 이후에는 역전파 알고리즘(back propagation algorithm)을 통해 오류를 수정하면서 학습을 진행한다. 마지막으로, 예측 모듈은 학습된 모형을 테스트하기 위해 학습에 사용하지 않은 30%의 데이터를 이용하여 학습된 모형의 정확도를 계산하여 제공한다.

본 연구는 예측 정확도를 개선하기 위해 최신의 기법들을 적용했다. 예를 들어, 입력 데이터를 점검하고, 예측 모형의 하이퍼파라미터

3) <https://www.tensorflow.org/>

(예, 배치크기, 레이어 수, 학습률, 숨겨진 차원 등)를 조정했다. 또한 새로운 초기화 함수 (Xavier와 He 초기화 함수)를 적용했다(Glorot, Bengio, 2010). 이와 함께, 다양한 최적화 함수 (Adam, Momentum, RMSProp 등)를 테스트했다.

케이스를 무작위로 나누어서 학습 데이터(70%, 7,933개의 케이스)와 테스트 데이터(30%, 3,400개의 케이스)를 구성했다.

## IV. 실험 방법과 결과

### 4.1 실험 방법

#### 4.1.1 활용 데이터

구현한 접근법을 검증하기 위해 본 연구는 국내 화장품 제조업체인 I사의 판매 프로세스 데이터를 활용했다. 이 데이터는 70,240개의 이벤트와 11,632개의 케이스를 포함하고 있다. 계절적 요인 등을 제거하기 위해 1년 이상의 판매 프로세스 행동을 기록한 데이터가 본 연구에 활용되었다. 단일 이벤트만을 포함하는 299개의 케이스가 제외되어 실제 사용된 데이터는 11,333개의 케이스를 포함하고 있다. 이 케이스들은 최소 2개부터 최대 51개의 이벤트를 포함하고 있다(<표 1> 참조). 파이썬 프로그래밍 언어에 내장된 기계학습 라이브러리인 사이킷런(Scikit-learn)의 함수를 이용하여 11,333개의

#### 4.1.2 학습 방법

기존 연구에서 적용한 방식처럼(예, Evermann et al., 2017a), 우리는 케이스를 구성하는 액티비티 이름을 순환신경망의 각 셀에 입력하기 위해 워드 임베딩(word embedding) 기법을 사용했다. 액티비티 이름도 문자열이므로 단어를 벡터로 변환하는 워드 임베딩 기법을 적용한 것은 적절하다고 판단된다.

다음으로 우리는 배치(batch) 크기를 20으로 설정하고, 7,933개의 케이스를 포함한 학습 데이터가 100번 반복하여 학습되도록 에포크(epoch)의 수를 100으로 설정했다. 배치 크기에 따른 정확도의 차이는 크지 않지만 배치 크기가 20일 때 가장 우수한 예측 정확도를 얻을 수 있었다(<표 2> 참조). 마지막으로 과적합(overfitting)을 방지하기 위해 드롭아웃과 10겹 교차검증 기법을 적용했다. 특히, 교차검증 기법을 적용할 때 첫 번째 겹(fold)에서는 7,933개의 학습 데이터를 10개의 부분집합(Sub1 ~ Sub10)으로 임의로 나눈 후에 처음 9개의 부분집합(Sub1 ~ Sub9)을 학습에 사용하고, 마지막 1개의 부분집합(Sub10)을 학습된 모형의 검

<표 1> 활용 데이터 요약

구분	값	구분	값
전체 이벤트 수	70,240개	전체 케이스 수	11,632개
액티비티 종류의 수	9개	제외 케이스 수	299개
케이스의 최대 이벤트 수	51개	케이스의 최소 이벤트 수	2개
시작일시	2016/05/19 14:31:29	완료일시	2017/10/30 11:21:03

증에 사용했다. 두 번째 겹에서는 검증에 사용된 부분집합(Sub10)과 Sub1 ~ Sub8까지의 부분집합이 다시 학습에 사용하고, Sub9 부분집합은 학습된 모형의 검증에 사용했다. 이러한 과정을 반복하면 예측모형에 대한 학습과 검증이 총 10번 수행되고, 각 겹마다 학습과 검증 정확도를 도출할 수 있었다.

<표 2> 배치 크기에 따른 예측 정확도

배치 크기	예측 정확도(%)
10	96.411
20	96.441
40	96.290
60	96.209
80	95.000
100	95.700
<b>평균</b>	<b>96.002</b>

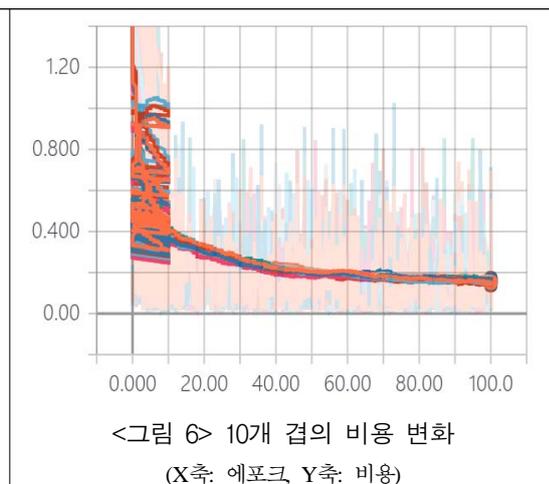
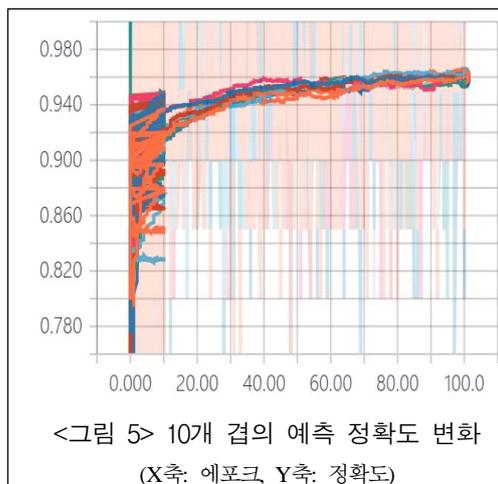
## 4.2 실험 결과

<표 3>과 <그림 5>에 제시된 것처럼, 학습 데이터를 활용한 예측 정확도의 평균은 95.79% 이었고, 테스트 데이터를 활용한 예측 정확도는

96.29%이었다. 이와 함께, <그림 6>을 보면 각 겹(fold)의 비용(cost)이 0.2에 근접하게 줄어들고 있음을 확인할 수 있다. 이러한 사실을 통해 우리는 학습이 제대로 수행되었다고 판단할 수 있었다. 한편, 기존 연구에서 제안된 정적 순환 신경망 형태의 딥러닝을 활용한 접근법에서는 학습 데이터를 활용한 예측 정확도의 평균은 94.49%이었고, 테스트 데이터를 활용한 예측 정확도의 평균은 94.67%이었다.

<표 3> 10겹 교차검증 적용에 따른 예측 정확도

겹(fold)	예측 정확도(%) (학습 데이터 기준)
1	97.5
2	99.5
3	100.0
4	95.499
5	90.999
6	92.5
7	96.499
8	93.0
9	93.5
10	98.999
<b>평균</b>	<b>95.79</b>



## V. 결론

본 연구는 전 세계에서 처음으로 동적 순환 신경망을 활용한 프로세스 예측 접근법을 성공적으로 구현하고 검증했다. 동적 순환신경망 형태의 딥러닝을 활용한 접근법의 예측 정확도는 정적 순환신경망 형태의 딥러닝을 활용하는 기존 접근법보다 테스트 데이터 기준으로 1.62% 더 높았다. 정적 순환신경망 형태의 딥러닝을 활용한 예측 정확도가 이미 94.67%를 달성한 가운데 이를 1.62% 더 개선했다는 것은 매우 우수한 결과라고 판단된다. 이와 함께, 본 연구가 제안한 접근법은 프로세스 수행을 기록한 모든 이벤트 데이터를 학습과 예측에 활용할 수 있는 방안을 제공하고 있다. 이러한 방안을 통해 본 연구가 제안한 접근법은 현실 세계의 실제 프로세스 행동을 정확히 학습할 수 있게 되었다.

한편, 정적 순환신경망의 시퀀스 길이를 잘못 설정하면 예측 정확도가 현저하게 떨어졌다. 예를 들어, 시퀀스 길이를 5로 설정하면 테스트 데이터를 활용한 예측 정확도가 84.89%로 떨어졌다. 정적 순환신경망 형태의 딥러닝을 활용하는 접근법은 최적의 시퀀스 길이를 찾는 데 많은 노력과 시간을 요구했다. 그러나 동적 순환신경망 형태의 딥러닝을 활용하는 접근법을 채택하면 최적의 시퀀스 길이를 찾기 위한 별도의 노력을 들일 필요가 없었다. 그러므로 본 연구에서 제안한 접근법은 실무에서 다양한 프로세스의 행동 예측에 매우 효율적으로 적용될 수 있을 것이다.

향후 연구는 본 연구에서 제안한 접근법을 다양한 프로세스의 행동과 성과지표 예측에 적

용할 필요가 있다. 또한 딥러닝 분야에서 제안되고 있는 새로운 알고리즘을 반영하여 예측 정확도를 더욱 개선할 필요가 있다. 이와 함께, 향후 연구는 회고 분석에 초점을 맞춘 프로세스 마이닝에서 발견한 통찰력을 프로세스 예측에 연결할 수 있는 방안을 연구할 필요가 있다.

## 참고문헌

- 강영식, 이보경, (경영자와 실무전문가를 위한) 프로세스 마이닝, 한나래 출판사, 2016.
- 김대회, 최승완, 곽수영, “딥러닝 기반의 가짜 얼굴 검출,” 한국산업정보학회지, 제 23 권, 제5호, 2018, pp. 9-17.
- 김진백, 김유일, “인공 신경망의 학습에 있어 가중치 변화방법과 은닉층의 노드수가 예측정확성에 미치는 영향,” 정보시스템 연구, 제9권, 제1호, 2000, pp. 27-44.
- 안성만, “딥러닝의 모형과 응용사례,” 지능정보 연구, 제22권, 제2호, 2016, pp. 127-142.
- 송현정, 이석준, “딥러닝을 활용한 실시간 주식 거래에서의 매매 빈도 패턴과 예측 시점에 관한 연구: KOSDAQ 시장을 중심으로,” 정보시스템연구, 제27권, 제3호, 2018, pp. 123-140.
- 최희열, 민윤홍, “딥러닝 소개 및 주요 이슈,” 정보처리학회지, 제22권, 제1호, 2015, pp. 7-21.
- 트란 광 카이, 송사광, “딥러닝 기반 침수 수위 예측: 미국 텍사스 트리니티강 사례연구,” 정보과학회논문지, 제44권, 제6호,

- 2017, pp. 607-612.
- 한진영, 조철현, 손인수, “기업의 빅데이터 활용에 관한 실증연구 : A 쇼핑사의 빅데이터 기반 통합로그 시스템 사례,” 인터넷전자상거래연구, 제15권 제6호, 2016, pp. 1-19.
- Bengio, Y., Simard, P., and Frasconi, P., “Learning long-term dependencies with gradient descent is difficult,” *Journal of IEEE Transactions on Neural Networks*, Vol. 5, No. 2, 1994, pp. 157-166.
- Breuker, D., Matzner, M., Delfmann, P., and Becker, J., “Comprehensible predictive models for business process,” *MIS Quarterly*, Vol. 40, No. 4, 2016, pp. 1009-1034.
- Chung, J., Gulcehre, C., Cho, K., and Bengio, Y., Empirical evaluation of gated recurrent neural networks on sequence modeling, *CoRR*, abs/1412.3555, 2014.
- Evermann, J., Rehse, J.-R., and Fettke, P., “A deep learning approach for predicting process behavior at runtime,” *PRAISE Workshop at the 14th International Conference on BPM*, 2016.
- Evermann, J., Rehse, J.-R., and Fettke, P., “Predicting process behaviour using deep learning,” *Decision Support Systems*, Vol. 100, 2017a, pp. 129-140.
- Evermann, J., Rehse, J.-R., and Fettke, P., “XES TensorFlow - Process prediction using the TensorFlow deep-learning framework,” *Forum of the Conference on Advanced Information Systems Engineering(CAiSE)*, Essen, Germany, 2017b.
- Glorot, X. and Bengio, Y., “Understanding the difficulty of training deep feedforward neural networks,” *International Conference on Artificial Intelligence and Statistics*, 2010.
- Graves, A., “Supervised Sequence Labelling with Recurrent Neural Networks,” Berlin Heidelberg, *Springer*, 2012.
- Hinton, G. and Salakhutdinov, R., “Reducing the dimensionality of data with neural networks,” *Science*, Vol. 313, No. 5786, pp. 504-507, 2006.
- Hochreiter, S. and Schmidhuber, J., “Long short-term memory,” *Neural Computation*, Vol. 9, No. 8, 1997, pp. 1735-1780.
- Jung, H. I., Park, I. S., and Ahn, H., “Identifying the key success factors of massively multiplayer online role playing game design using artificial neural networks,” *The Journal of Society for e-Business Studies*, Vol. 17, No. 1, 2012, pp. 23-38.
- Kang, Y. S. and Park, S., “Predicting the relationship between corporate financial information and credit rating using deep learning,” *Korean Journal of Business Administration*, Vol. 31, No. 7, 2018, pp. 1253-1275.

Lakshmanan, G., Shamsi, D., Doganata, Y. N., Unuvar, M., and Khalaf, R., "A markov prediction model for data-driven semi-structured business processes," *Knowledge and Information Systems*, Vol. 42, No. 1, 2015, pp. 97-126.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., and Kruschwitz, N., "Big data, analytics and the path from insights to value," *MIT Sloan Management Review*, Vol. 52, 2011, pp. 21-32.

LeCun, Y., Bengio, Y., and Hinton, G., "Deep learning," *Nature*, Vol. 521, pp. 436-444. 2015.

Schmidhuber, J., "Deep learning in neural networks: An overview," *Neural Networks*, Vol. 61, 2015, pp. 85-117.

Sutskever, I., Martens, J., and Hinton, G. E., "Generating text with recurrent neural networks," *ICML*, Omnipress, 2011, pp. 1017-1024.

Tax, N., Verenich, I., Rosa, M. L., and Dumas, M., "Predictive business process monitoring with LSTM neural networks," *CoRR*, abs/1612.02130, 2016.

Unuvar, M., Lakshmanan, G. T., and Doganata, Y. N., "Leveraging path information to generate predictions for parallel business processes," *Knowledge and Information Systems*, Vol. 47, No. 2, 2016, pp. 433-461.

van der Aalst, W. M. P., "Process Mining: Data

Science in Action," *Springer*, 2016.

van der Aalst, W. M. P., Schonenberg, M. H., and Song, M., "Time prediction based on process mining," *Information Systems*, Vol. 36, No. 2, 2011, pp. 450-475.

#### 김 정 연 (Kim, Jung-Yeon)



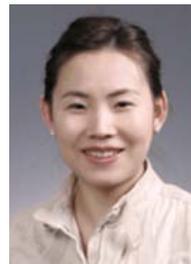
명지대학교 경영정보학과를 졸업하고, 현재 프로세스 마이닝 전문기업인 (주)피엠아이지에 재직 중이다. 주요 관심분야는 프로세스 마이닝과 딥러닝을 활용한 프로세스 예측 기술 개발이다.

#### 윤 석 준 (Yoon, Seok-Joon)



현재 프로세스 마이닝 전문기업인 (주)피엠아이지에 재직 중이다. 주요 관심분야는 딥러닝을 활용한 프로세스 예측 기술 개발과 프로세스 마이닝 등이다.

#### 이 보 경 (Lee, Bo-Kyoung)



서울여자대학교에서 컴퓨터학 학사학위 및 석사학위를 취득하고, 동대학원에서 경영정보학 박사학위를 취득하였다. 현재 (주)피엠아이지에서 수석컨설턴트로 재직 중이다. 주요 관심분야는 프로세스 마이닝을 활용한 데이터 분석, RPA 등이다.

<Abstract>

## **Exploring process prediction based on deep learning: Focusing on dynamic recurrent neural networks**

Kim, Jung-Yeon · Yoon, Seok-Joon · Lee, Bo-Kyoung

### **Purpose**

The purpose of this study is to predict future behaviors of business process. Specifically, this study tried to predict the last activities of process instances. It contributes to overcoming the limitations of existing approaches that they do not accurately reflect the actual behavior of business process and it requires a lot of effort and time every time they are applied to specific processes.

### **Design/methodology/approach**

This study proposed a novel approach based using deep learning in the form of dynamic recurrent neural networks. To improve the accuracy of our prediction model based on the approach, we tried to adopt the latest techniques including new initialization functions(Xavier and He initializations). The proposed approach has been verified using real-life data of a domestic small and medium-sized business.

### **Findings**

According to the experiment result, our approach achieves better prediction accuracy than the latest approach based on the static recurrent neural networks. It is also proved that much less effort and time are required to predict the behavior of business processes.

**Keyword:** Deep Learning, Recurrent Neural Networks, Process Prediction, Process Intelligence, Event Log

\* 이 논문은 2018년 11월 16일 접수, 2018년 12월 3일 1차 심사, 2018년 12월 15일 게재 확정되었습니다.