

희소 은닉 상태 Deep Recurrent Neural Network을 이용한 웨어러블 라이프로그 학습

곽동현¹ 이충연² 곽하늬² 장병탁^{1,2}
 뇌과학 협동과정¹, 컴퓨터공학부², 서울대학교
 {dhkwak, cvlee, hnkwak, btzhang}@bi.snu.ac.kr

Wearable Lifelog Learning Using Sparse Hidden State Deep Recurrent Neural Network

Dong Hyun Kwak^{1*} Chung-Yeon Lee^{2**} Hanock Kwak^{2**} Byoung-Tak Zhang^{1,2**}

Brain Science Program^{*}, School of Computer Science & Engineering^{**}, Seoul National University

요 약

웨어러블 라이프로그는 사람의 일상을 기록한 시계열 데이터로 이를 학습하는 것은 실생활에 밀접한 지능형 서비스를 구현하기 위해서 필수적으로 풀어야 하는 문제이다. 기존에 시계열 데이터를 다루기 위한 인공신경망 모델은 RNN 계열의 알고리즘이 대표적이다. 그러나 RNN은 병렬화가 힘들고, 긴 시간 의존성을 학습하기 힘든 문제가 있다. 이러한 문제를 해결하기 위해 모든 instance가 은닉 상태를 갖지 않고, N개의 instance 당 1개의 은닉 상태를 갖는 희소 은닉 상태를 설정한 DRNN을 제안한다. 이를 토이 문제 및 웨어러블 라이프로그 데이터에서 기존의 DRNN과의 비교 실험 수행하여 희소 은닉 상태 가정의 유효함을 보이고, 시간을 고려하지 않는 모델인 DNN과 비교하여 시계열 데이터에서 시간 의존성 학습의 중요성을 보인다.

1. 서 론

웨어러블 라이프로그는 피험자의 시선 녹화 및 주변 환경 소리 녹음, GPS 정보, 문자 및 이메일 정보, 가속도 센서 등 개인의 일상생활에서 얻을 수 있는 다양한 센서 정보를 수집한 데이터이다[1]. 이를 학습할 경우 다음 이동경로와 행동, 감정 상태, 스케줄 등 생활에 밀접한 예측이 가능해 개인 맞춤형 비서 연구에 시작이 될 수 있다. 본 논문에서는 직접 수집한 라이프로그 데이터와 기존의 DRNN을 개선한 희소 은닉 상태 DRNN 모델을 이용하여 개인의 일상 생활을 학습하고 예측하는 연구를 수행하였다.

RNN은 DNN 모델에서 은닉 노드에 자기 자신을 가리키는 weight가 추가된 알고리즘이다. 또한 은닉 노드가 갖는 활성 값을 은닉 상태로 사용하며, 직전의 은닉 상태로부터 현재의 은닉 상태가 영향 받는 마코프 속성을 담고 있는 모델이다. 이러한 RNN 계열의 모델은 시간 정보가 담긴 시계열 데이터를 학습하는데 적합하여 음성인식, 언어모델, 신호처리, 자연어처리 등의 다양한 분야에서 연구되고 있다[2].

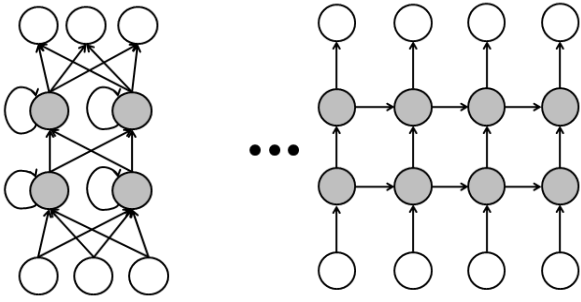
DRNN은 RNN의 은닉 층이 2개 이상인 깊은 구조를 가진 모델로 시간에 대한 더욱 고수준의 표현력을 학습할 수 있다[3]. DRNN을 이용해 라이프로그를 학습할 경우 사람의 행동 상태를 포함한 복잡한 잠재 요소와 표현을 학습할 수 있다. 그러나 기존의 DRNN은 직전 은닉 상태로부터만 영향을 받고, 보통

k번 까지만 back-propagation을 수행하여 긴 시간 의존성을 학습이 잘 안되었다. 병렬화 또한 어려워 GPU 컴퓨팅을 통한 빠른 연산도 힘든 문제가 있었다. 본 논문에서는 이를 해결하기 위해 은닉 상태가 매 instance 아니라 일정한 n개의 단위로 존재하는 희소 은닉 상태 DRNN을 실험하였다. 또한 이 모델을 기존에 널리 알려진 DNN 및 DRNN과 비교하여 수행 속도뿐만 아니라 성능에서도 우수함을 보이는 실험을 수행하였다.

2. 희소 은닉 상태 Deep Recurrent Neural Network

2.1 Deep Recurrent Neural Network

DRNN은 [그림 1]과 같이 모델의 복잡도가 높아 더욱 다양하고 깊은 표현력이 학습 가능한 모델이다. 그러나 이 DRNN을 적절히 학습시키기 위해서는 최적화된 하이퍼 파라미터 세팅이 필요해 학습이 어렵고 학습시간 또한 오래 걸리는 것으로 알려져 있다. 게다가 학습을 하더라도 단기간의 시간 의존성을 주로 학습하기 때문에 긴 은닉 상태들 간의 관계는 잘 학습하지 못하였다. 이는 신경망 알고리즘의 특성상 가장 최근에 학습한 요소가 이전에 학습한 요소들을 덮어버리는 현상에 의해 발생한다.



[DRNN의 공간적 구조] [DRNN의 시간적 구조]

그림 1. 두 개의 은닉 층으로 구성된 DRNN은 공간적으로는 왼쪽 그림처럼 일반 DNN에서 은닉 층에서 은닉 층을 향하는 완전 연결 가중치(fully connected weight)가 추가된 구조이다. 시간적으로는 오른쪽 그림과 같이 은닉 층이 가상의 과거 시간에 대해 펼쳐지는 구조로 시간에 대해 깊은 은닉 층을 갖는 구조이다.

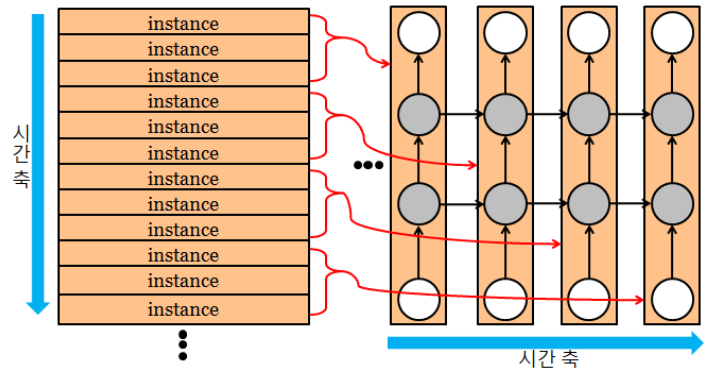
또한 컴퓨팅 효율에 있어서도 RNN 계열의 모델은 병렬화가 힘든 시간 의존성이 존재하여 GPU 컴퓨팅 등을 통한 빠른 연산도 어려웠다. 병렬화는 서로간의 의존성이 약한 부분에 대해서 여러 개의 계산 유닛을 사용해 개별 알고리즘을 수행하고, 이를 다시 합치는 방식으로 이루어 지는데 기존의 RNN 계열의 알고리즘은 반드시 직전 instance로부터 다음 instance가 영향을 받는 시간 의존성이 있기 때문에 병렬화가 이루어지기 힘든 구조였다.

이러한 긴 시간 의존성을 학습하지 못하는 문제를 해결하기 위해 Long-Short Term Memory와 같이 gating을 이용해 데이터에 따른 선택적인 학습과 추론을 통한 알고리즘이 현재 시계열 모델에서 가장 널리 쓰이는 방식이다[4].

2.2 희소 은닉 상태 Deep Recurrent Neural Network

희소 은닉 상태 DRNN은 [그림 2]와 같이 은닉 상태가 매 instance마다 존재하지 않고, 시간 축으로 펼쳐진 은닉 상태들이 여러 개씩 묶인 instance의 은닉 상태를 나타내도록 설계한 모델이다. 이러한 방식은 위에서 제기된 기존 DRNN의 문제인 긴 시간 의존성의 학습이 가능하고, 또한 동일한 은닉 상태를 갖는 수에 비례하여 병렬화된 처리가 가능하게 해준다. 동일한 은닉 상태를 가정한 크기가 n 이고, k 번의 시간 축에 대한 학습을 하는 k -truncated back-propagation through time을 적용할 경우 기존 DRNN은 오직 k 개의 시간을 학습시키는 반면, 희소 은닉 상태 DRNN은 $k*n$ 개의 시간을 학습시킬 수 있다.

또한 이 n 의 크기가 확률적 경사 하강(stochastic



[1/3의 희소 은닉 상태를 가정한 DRNN의 시간적 구조]

그림 2. 희소 은닉 상태 DRNN의 경우 공간적인 구조는 일반 DRNN과 동일하다. 시간적인 구조는 과거 시간에 대해 펼쳐진 층들에 대해 희소한 은닉 상태를 가정하여, 직전 상태로부터 영향을 받는 마코프 속성이 희소 은닉 상태들 사이에 존재한다.

gradient descent)에서의 미니배치 크기와 동일하게 처리할 경우 DNN과 동등한 병렬도로 매우 빠른 연산이 가능한 이점이 있다.

단, 동일한 은닉 상태의 크기가 커질수록 단기적인 시간 의존성의 학습이 잘 이루어지지 않기 때문에 time-window가 적용된 MFCC feature와 같은 데이터를 사용하여 단기적인 시간 요소가 추출된 상태로 학습을 시켜야 한다. 또는 RNN의 기계 번역 모델을 이용해 단기적인 시간 의존성이 반영된 요소를 먼저 추출한 뒤 이를 DRNN으로 학습하는 방법도 고려해 볼 수 있다.

3. 실험

3.1 토이 문제 실험

희소 은닉 상태 DRNN의 기본적인 성능을 비교하기 위해 [표 1]과 같은 실험을 수행하였다. 모델은 모두

표 1. 두 종류의 토이 문제를 학습한 후 테스트 정확도를 측정된 결과(동일한 크기의 은닉 층, 은닉 노드, 학습률 그리고 반복횟수를 갖도록 설정함). 1번 토이 문제의 경우 시간정보가 고려되지 않은 데이터로 [1,1000]사이의 자연수에는 클래스 1을, [1001,2000]사이의 자연수에는 클래스 2를 지정한 문제이다. 2번 토이 문제의 경우 [1:10]으로 1씩 증가하는 벡터에는 클래스 1을, [10:1]로 1씩 감소하는 벡터에는 클래스 2를 지정한 벡터들로 구성되어 있다.

데이터	DNN	DRNN	희소 은닉 상태 DRNN
1번 토이 문제	0.874	0.823	0.926
2번 토이 문제	0.5	0.976	0.985

[50, 25]개의 은닉 노드를 갖는 2개의 은닉 층을 갖도록 구성하였으며, 두 DRNN 모델은 1-truncated bptt를 사용하였다. 희소 은닉 상태는 2개의 instance가 같은 은닉 상태를 갖는다고 가정하였다. 실험 결과 시간이 고려되지 않은 1번 문제는 물론, 바로 직전 instance의 영향을 받는 2번 문제에서도 모두 희소 은닉 상태 DRNN의 성능이 더 높게 측정 되었다.

1번 문제의 경우 은닉 상태를 나타내는 가중치가 은닉 상태 간의 연관성을 학습하기 보다는 단순히 선형 분류 문제를 풀기에 더 좋은 스케일을 갖도록 해주는 역할을 수행하여 DNN보다 DRNN 및 희소 은닉 상태 DRNN이 더 많은 가중치를 사용해서 성능이 높게 나왔다고 분석된다.

2번 토이 문제의 경우 시간 요소를 고려하지 않은 DNN에서는 예상대로 전혀 학습이 이루어지지 않았다. 그러나 1씩 숫자가 증가하는 문제임에도 희소 은닉 상태 미세하게 DRNN이 더 좋은 성능을 보였다. 이는 우선 1-truncated bptt를 수행하였기 때문에 DRNN이 10개의 instance단위로 구성된 은닉 상태를 잘 학습하지 못한 반면 희소 은닉 상태 DRNN은 2배 더 많은 은닉 상태를 학습하였기 때문으로 해석된다

3.2 웨어러블 라이프로그 데이터 실험

웨어러블 센서로 수집한 실생활 데이터인 웨어러블 라이프로그(wearable lifelog)는 인공지능 비서의 구현을 위해 필수적인 연구이다. 본 실험에서는 구글 클래스를 활용하여 피험자가 보는 시선 동영상과 주변 환경의 소리를 매일 8시간 씩 녹화하는 방식으로 14일간 2명의 피험자를 통해 수집되었다. 직접 수집한 라이프로그

표 2. 웨어러블 라이프로그를 세가지 모델로 학습하여 크로스 엔트로피와 테스트 정확도를 측정한 결과(동일한 크기의 은닉 층, 은닉 노드, 학습률 그리고 반복횟수를 갖도록 설정함). 웨어러블 라이프로그 데이터는 데이터 특성상 같은 행동이 연속적으로 긴 시간 측정되므로 같은 종류의 클래스를 계속해서 학습하게 되어 overfitting이 쉽게 일어난다. 이를 방지하기 위해 데이터를 50개 단위로 뒤섞고, 앞부분의 데이터 75%로 학습을, 뒷 부분의 25%로 테스트를 수행하였다.

성능 평가	크로스 엔트로피	테스트 정확도
DNN	-10.62	0.3220
DRNN	-4.49	0.3421
희소 은닉 상태 DRNN	-8.19	0.3501

동영상에서 소리정보를 추출하고 이를 MFCC 13 bands 로 전처리 한 데이터를 사용해 실험하였다.

실험 결과 [표 2]와 같이 희소 은닉 상태 DRNN이 테스트 정확도에 있어서 약간 더 좋은 성능을 보였다. 이는 MFCC feature를 통해 단기적인 시간을 학습하고, 시간 속성은 희소 은닉 상태를 통해 보다 긴 의존성까지 학습한 것으로 해석된다.

4. 논의 및 결론

본 논문에서는 시간 요소를 학습할 수 있는 인공신경망 모델인 DRNN을 개선한 희소 은닉 상태 DRNN을 여러 실험을 통해 비교 분석하였다. 또한 제안하는 모델이 토이 문제가 아닌 실제 문제를 해결하는 데에도 적용 가능성을 보이고, DNN처럼 시간을 고려하지 않은 모델은 학습이 불가능한 문제들을 다루어 시간을 고려한 모델의 중요성을 살펴보았다.

희소 은닉 상태 DRNN은 기존 DRNN에 비해 학습이 빠르고 긴 시간 의존성을 학습할 수 있는 장점을 가진 모델이다. 단, 본 실험은 supervised sequence labeling문제에 해당함에도 불구하고 connectionist temporal classification과 같은 목표함수를 설정하지 않아 잘못된 방향으로 학습이 일어나는 경우가 있었다. 후속 연구를 통해 이를 개선하여, 웨어러블 라이프로그 데이터를 더욱 잘 학습하고 처리할 수 있는 알고리즘을 연구하여 이를 활용한 인공지능 비서 서비스의 가능성을 확인하고자 한다.

감사의 글

이 논문은 2015년도 정부(미래창조과학부)의 재원으로 정보통신기술진흥센터의 지원(R0126-15-1072-SW스타랩, 10035348-mLife, 10044009-HRI.MESSI)과 한국연구재단의 지원(NRF-2010-0017734-Videome)을 받아 수행된 연구임.

참고문헌

- [1] 이충연, 이상우, 광동현, 장병탁, “웨어러블 라이프로그 기반 상황 인식과 응용 시나리오”, 한국정보과학회 동계학술발표회 논문집, pp. 337-339, 2014
- [2] Sutskever, Ilya. “Training recurrent neural networks.”, Diss. University of Toronto, 2013.
- [3] PASCANU, Razvan, et al, “How to construct deep recurrent neural networks.” arXiv preprint arXiv:1312.6026, 2013.
- [4] HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. “Long short-term memory”. Neural computation, vol. 9, no.8: pp 1735-1780, 1997.