

## 15장 RNN과 CNN 활용: 순차 데이터 처리

## 감사의 글

자료를 공개한 저자 오렐리앙 제롱과 강의자료를 지원한 한빛아카데미에게 진심어린 감사를 전합니다.

# 소개

- 시계열 데이터 분석 활용
  - 주식가격 예측
  - 차의 이동경로 예측
- 임의의 길이를 갖는 순차데이터(시퀀스) 처리 가능
  - 문장, 문서, 오디오 샘플 등
- 자동번역, 받아쓰기 등 자연어처리(NLP) 분야에 매우 유용함.

## 주요 내용

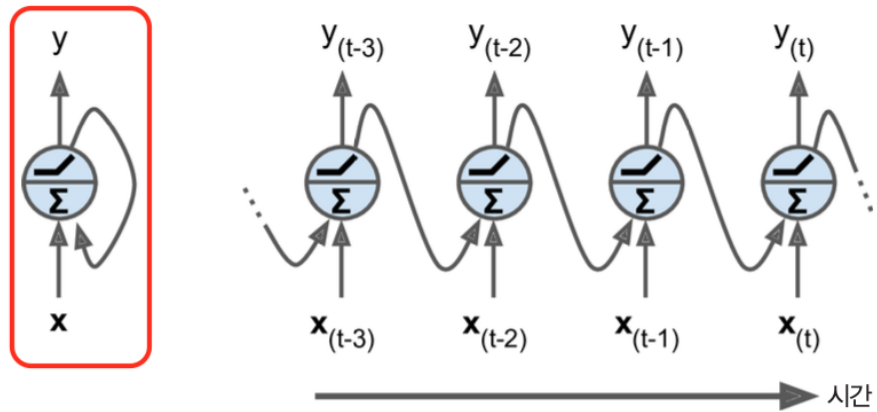
- 순환신경망(RNN) 기본개념 및 역전파 네트워크 훈련 방법 소개
- 시계열 예측 수행
- RNN의 주요 난제 설명
  - 불안정한 그레이디언트 다루기: 순환 드롭아웃 및 순환층 정규화
  - 매우 제한적인 단기기억 문제 다루기: LSTM과 GRU 셀 활용

**순환뉴런과 순환층**

- 지금까지 살펴본 신경망 모델은 아래층에서 위층으로 학습결과를 전달하는 **전방향전달** (feedforward) 방식 사용
- 순환신경망(recurrent neural network, RNN)도 비슷하게 작동
- 차이점: 후방향으로 전달하여 학습결과를 순환시키는 기능 활용

**순환뉴런과 타이스텝**

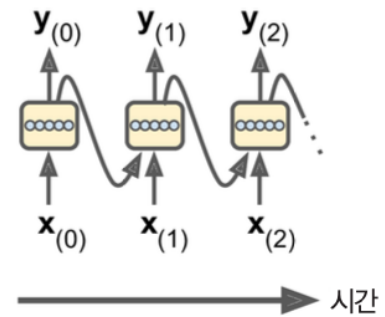
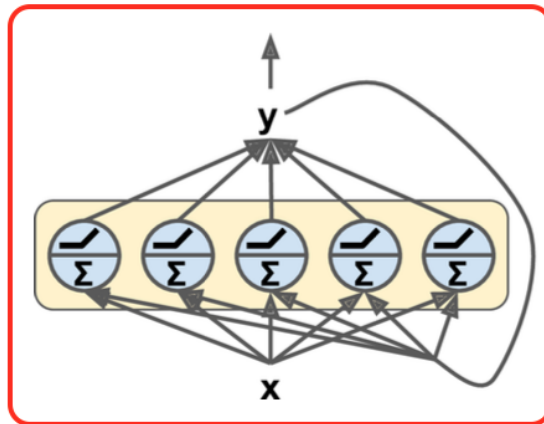
- 순환뉴런: 입력을 받아 출력을 만든 후 자신에게도 출력결과를 전달





순환층

- 순환뉴런으로 구성된 층



## 하나의 샘플에 대한 순환층의 출력

- 가중치 행렬
  - $\mathbf{W}_x$ : 현재 타임스텝의 입력에 대한 연결가중치 행렬
  - $\mathbf{W}_y$ : 이전 타임스텝의 출력에 대한 연결가중치 행렬
  - 타임스텝에 의존하지 않음.

$$\mathbf{y}_{(t)} = \phi(\mathbf{W}_x^T \mathbf{x}_{(t)} + \mathbf{W}_y^T \mathbf{y}_{(t-1)} + \mathbf{b})$$

## 배치에 대한 순환층의 출력

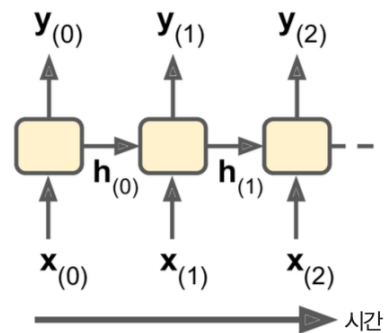
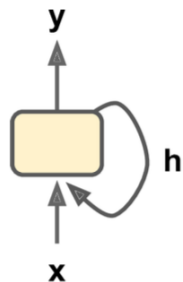
$$\begin{aligned} \mathbf{Y}_{(t)} &= \phi(\mathbf{X}_{(t)} \mathbf{W}_x + \mathbf{Y}_{(t-1)} \mathbf{W}_y + \mathbf{b}) \\ &= \phi([\mathbf{X}_{(t)} \ \mathbf{Y}_{(t-1)}] \mathbf{W} + \mathbf{b}) \quad \text{여기에서} \quad \mathbf{W} = \begin{bmatrix} \mathbf{W}_x \\ \mathbf{W}_y \end{bmatrix} \end{aligned}$$

**(메모리) 셀**

- 메모리 셀: 타임스텝에 걸쳐서 특정 상태를 보존하는 신경망의 구성요소
- 예제: 하나의 순환뉴런 또는 순환뉴런층
  - 10 스텝 정도의 짧은 패턴을 학습하는 기본적인 셀

- 상태  $\mathbf{h}_{(t)}$ : 타임스텝  $t$  에서의 셀의 상태

$$\mathbf{h}_{(t)} = f(\mathbf{h}_{(t-1)}, \mathbf{x}_{(t)})$$



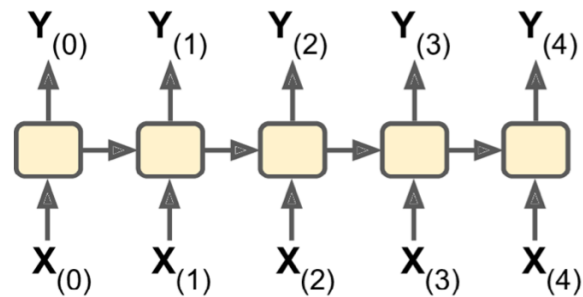
**입력/출력 순차데이터**



- 목적에 따라 입력/출력 데이터의 형식을 벡터 또는 순차데이터(시퀀스)로 지정할 수 있음.

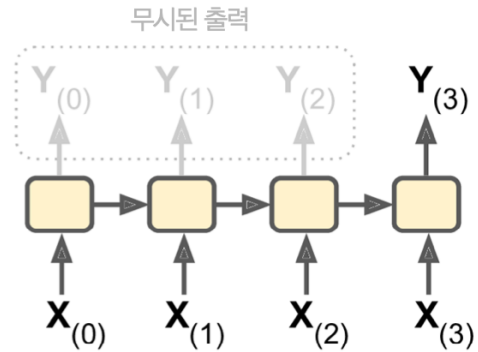
## seq-to-seq

- 입력: 순차데이터
- 출력: 순차데이터
- 예제: 주식가격 시계열 데이터 활용에 유용
  - 최근 N일 동안의 주식가격을 입력하면 각 입력값보다 하루 이후의 가격을 출력



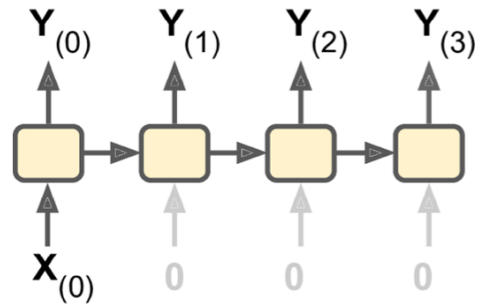
## seq-to-vector

- 입력: 순차데이터
- 출력: 벡터
- 예제: 영화리뷰
  - 영화리뷰에 있는 연속된 단어를 주입하면 감성점수 출력(-1에서 1사이)



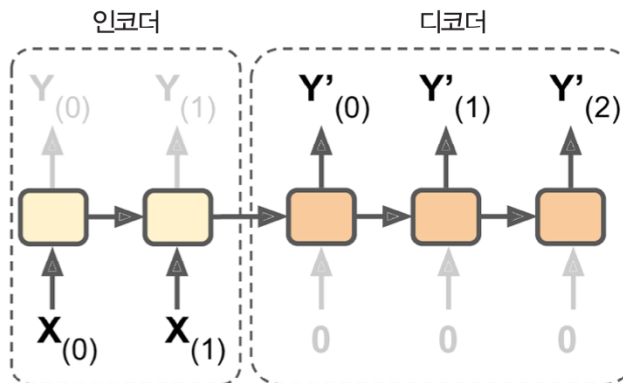
## vector-to-seq

- 입력: 벡터
- 출력: 순차데이터
- 예제: 이미지를 입력하면 이미지 캡션 출력



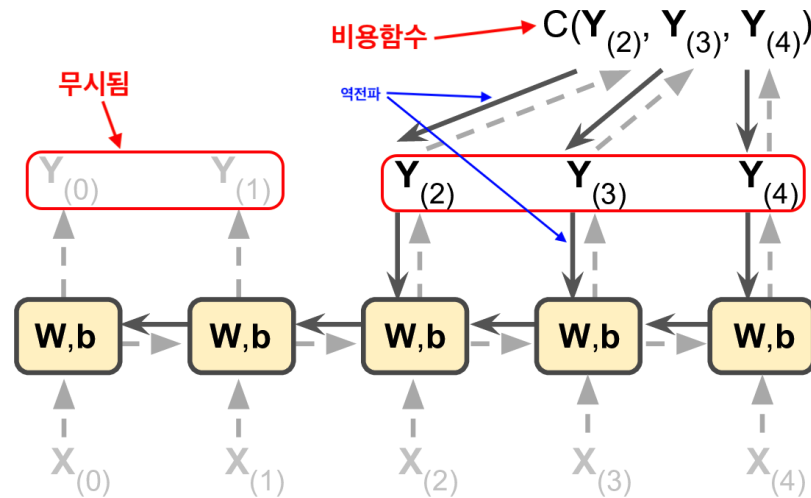
## encoder-decoder

- 인코더와 디코더의 조합을 사용하는 이중단계모델
  - 인코더: seq-to-vector 신경망
  - 디코더: vector-to-seq 신경망
- 예제: 언어 번역
  - seq-to-seq 모델보다 성능 좋음.
  - 이유: "한국말은 끝까지 들어야 한다" 와 동일한 이유임.
  - 16장 참조



# RNN 훈련

- BPTT(backpropagation through time) 전략: 타임스텝을 통과하는 역전파 전략



**시계열 예측**

## 단변량 시계열

- 타임스텝마다 하나의 값을 가지는 순차데이터
- 예제
  - 웹사이트의 시간당 접속자 수
  - 도시의 날짜별 온도



## 다변량 시계열

- 타임스텝마다 여러 개의 값을 가지는 순차데이터
- 예제
  - 회사의 수입, 부채 등 여러 지표를 사용한 기업의 분기별 재정안정성

## 활용법

- 예측: 미래의 값 추정하기
- 결측치 대체: 과거 데이터에서 누락된 값 추정하기

## 예제

- 아래 그림: 3개의 단변량 시계열
  - 타임스텝: 50개
  - 목표: x 로 표시된 값 예측하기

