

Week01 발표

RNN
이경선

순환 신경망 개념

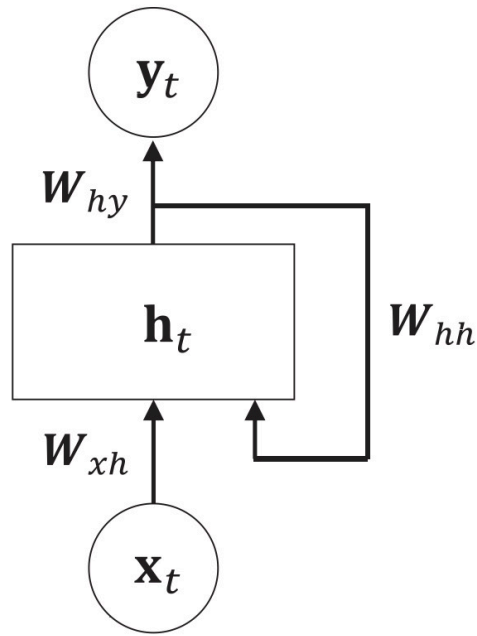


그림 12-71 순환 신경망 구조

출력 결과를 다음 시점까지 기억했다가 사용하는 방법

순환 신경망 개념

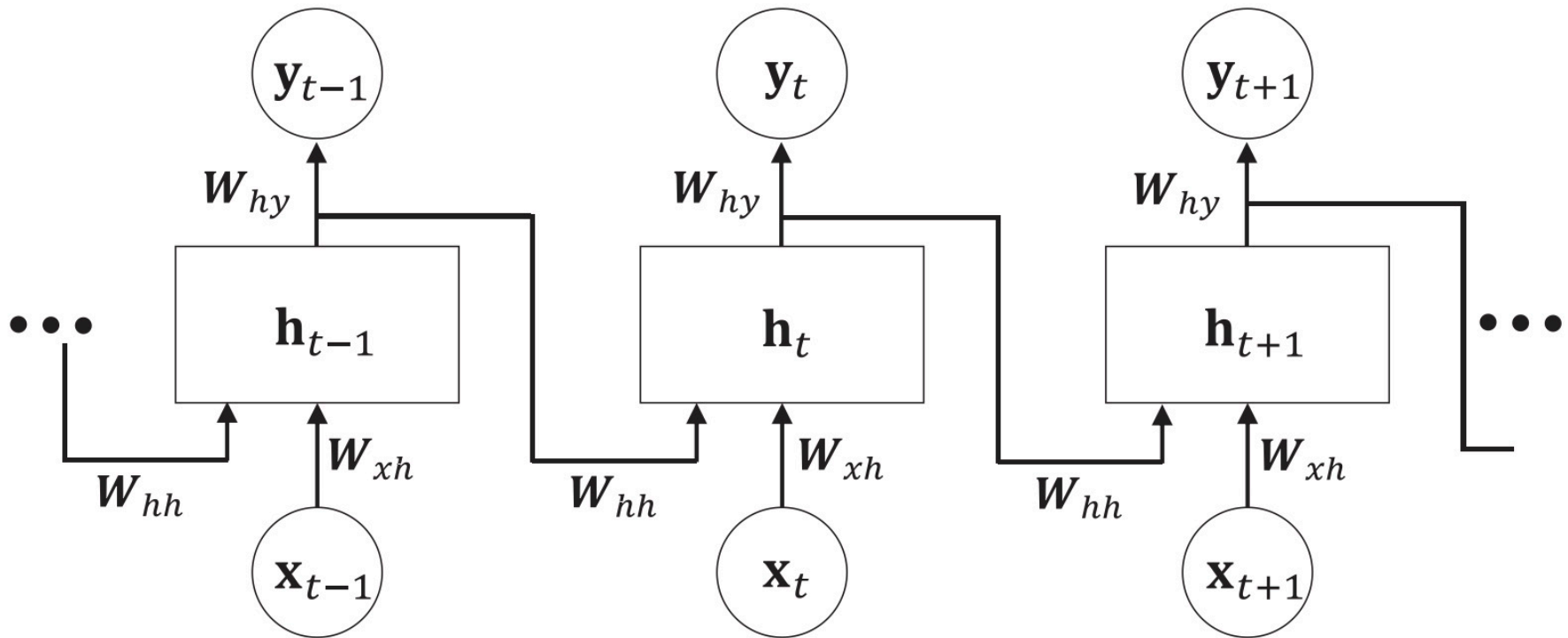


그림 12-72 순환 신경망 구조(펼침)

시간 순서로 순환 신경망을 표현한 그림.

순환 신경망 개념

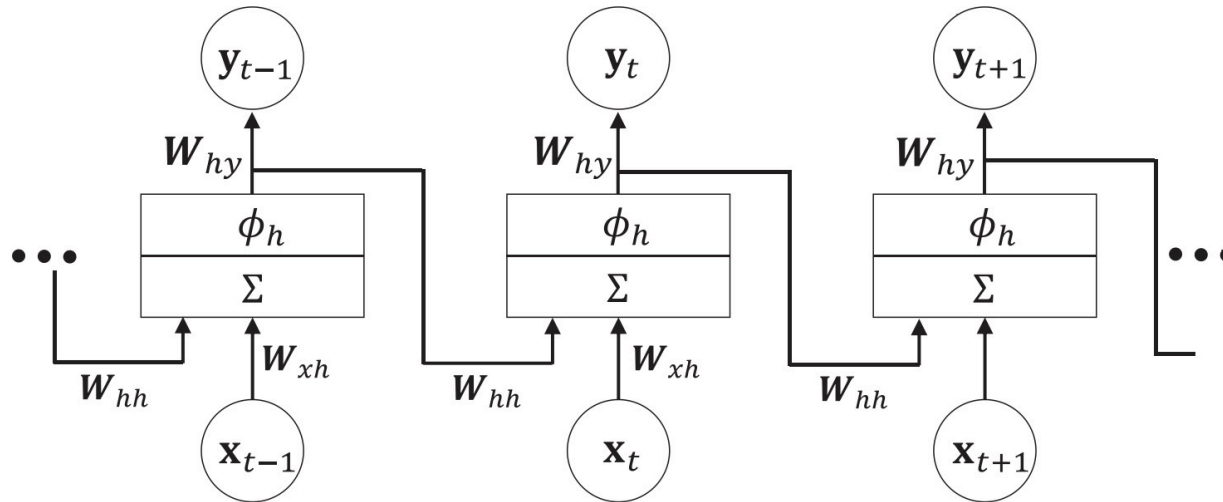


그림 12-73 순환 신경망 내부

$$\mathbf{h}_t = \phi_h(W_{xh}\mathbf{x}_t + W_{hh}\mathbf{h}_{t-1} + \mathbf{b}_h)$$

$$\mathbf{h}_t = \phi_h \left(\begin{matrix} W_{xh} & \mathbf{x}_t \\ W_{hh} & \mathbf{h}_{t-1} \\ \mathbf{b}_h \end{matrix} \right)$$

The diagram shows the calculation of the hidden state \mathbf{h}_t at time step t . It is represented as a vertical vector of four cells. The calculation is performed by the activation function ϕ_h on the sum of three components: the product of the input weight matrix W_{xh} (a 4x3 grid) and the input vector \mathbf{x}_t (a 3x1 vector), the product of the hidden state weight matrix W_{hh} (a 4x4 grid) and the previous hidden state vector \mathbf{h}_{t-1} (a 4x1 vector), and the bias vector \mathbf{b}_h (a 4x1 vector).

그림 12-74 은닉층 계산 (1)

순환 신경망 개념

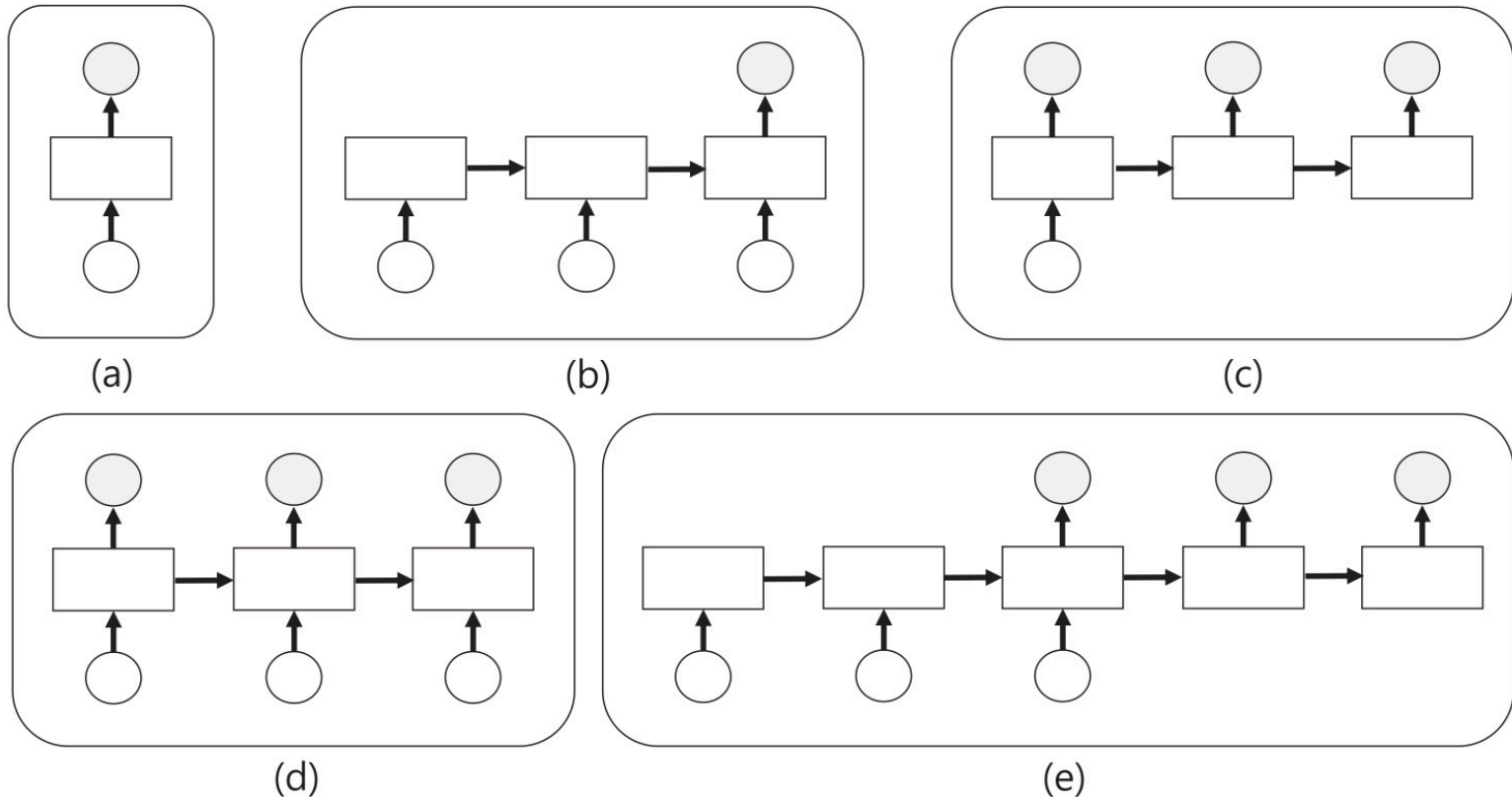
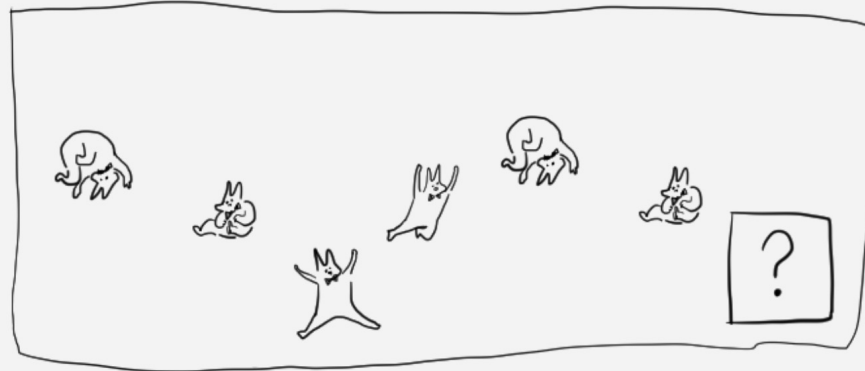


그림 12-76 순환 신경망 종류

RNN의 용도

연속적인 이벤트 :

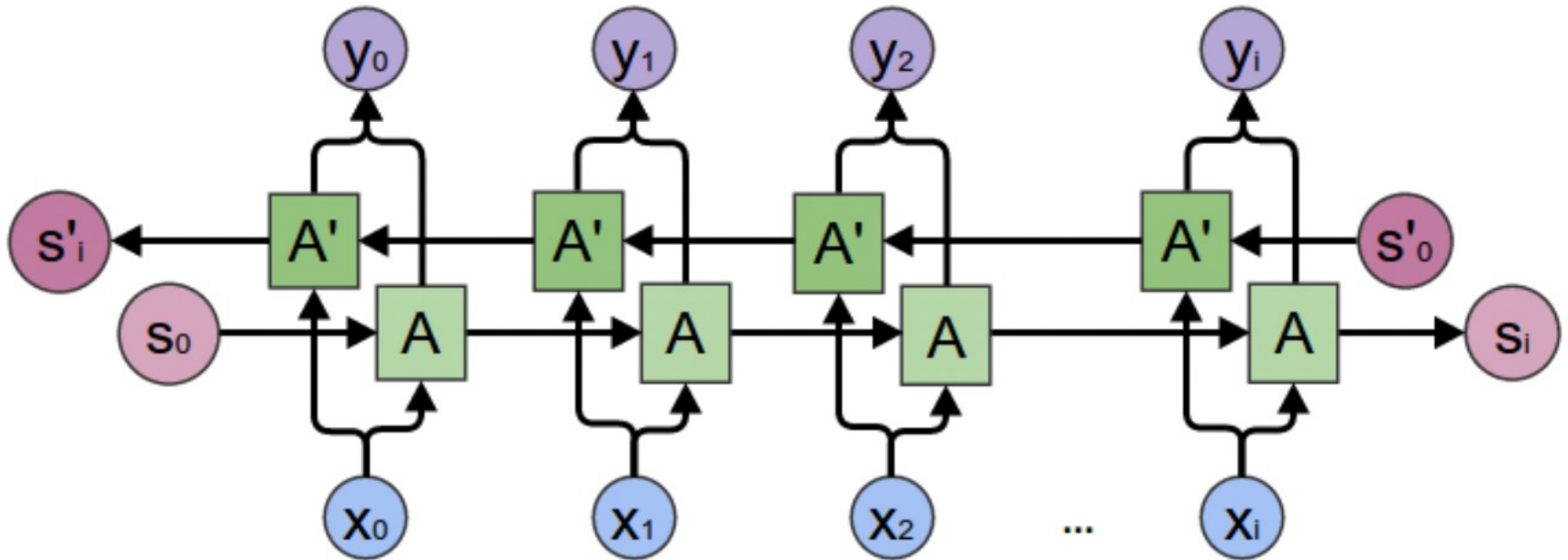


자연어 처리 :



이전 단계의 결과를 다음 단계의 입력으로 사용하므로,
과거의 일을 토대로 미래를 예측하는 데에 이용할 수 있다.

Bidirectional RNN

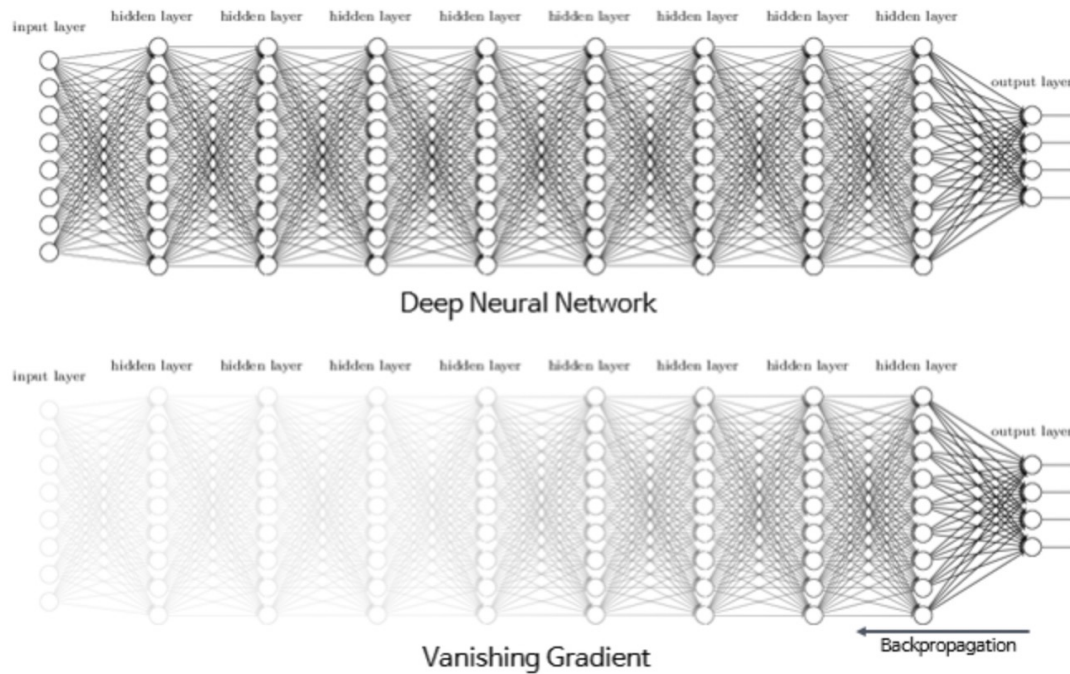


"오늘 엄청 맛있는 고기를 먹었는데 기분이 -----, 아니 고기가 상했지 뭐야!"
빈 칸 이전의 문장만 보고서는 긍정적인 답을 생각하게 된다.

RNN의 문제

Vanishing gradient problem

Exploding gradient problem



역전파 알고리즘이 입력층으로 전달됨에 따라 그래디언트가 점점 작아져 결국 가중치 매개변수가 업데이트 되지 않는 경우가 발생하게 된다.