



CH7

7.1 시계열 문제

- 시계열 분석 : 시간에 따라 변하는 데이터 사용하여 추이를 분석하여 추세를 파악하거나 전망을 예측하는 테스트를 진행 (주가 변동, 기온 변화 등) =>

1. 불규칙 변동 : 규칙성이 없어 예측 불가, 우연적으로 발생하는 변동

예) 전쟁, 자연재해, 파업

2. 추세 변동 : 장기적인 변화 추세 → 단기에는 찾기 어려운 단점

예) GDP, 인구증가율

3. 순환 변동 : 2-3년 정도 일정한 기간을 주기로 나타나는 변동, 추세 변동에 따라 변동함

예) 경기 변동

4. 계절 변동 : 계절적 영향과 사회적 관습에 따라 1년 주기로 발생

⇒ 시계열 데이터 = 규칙적 시계열 + 불규칙적 시계열 (트렌드 혹은 분산의 변화 유무)

⇒ 시계열 분석 → 특정 기법이나 모델을 통해 불규칙한 데이터에서 규칙을 찾아내는 것

7.2 AR, MA, ARMA, ARIMA

시계열 분석에서는 시간이 독립 변수

AR(AutoRegressive)

이전 관측 값이 이후 관측 값에 영향을 준다는 아이디어 이용하여 자기 회귀 모델이라고도 부름

→

$$\underbrace{Z_t}_{\text{①}} = \underbrace{\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p}}_{\text{②}} + \underbrace{a_t}_{\text{③}}$$

- 1) 데이터 현재 시점

- 2) 과거가 현재에 미치는 영향을 나타낸 모수 Φ 에 과거 시점을 곱한 것

3) 오차항(백색잡음)

⇒ p시점을 기준으로 이전 데이터 **상태**에 의해 현 시점 데이터가 영향을 받는 모형

MA(Moving Average)

시계열을 따라 윈도우 크기만큼 슬라이딩되기에 이동 평균 모델이라고 부름

트렌드가 변화하는 상황에 적합

시간이 지날수록 어떠한 Variable의 평균 값이 지속적으로 감소하거나 증가하는 경향이 생길 수 있음

$$\underbrace{Z_t}_{\textcircled{1}} = \underbrace{\theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_p a_{t-p}}_{\textcircled{2}} + \underbrace{a_t}_{\textcircled{3}}$$

1) 데이터 현재 시점

2) 매개변수 θ 에 과거 시점의 오차를 곱한 것

→

3) 오차항 (정규 분포에서 도출되는 임의의 값)

⇒ 이전 데이터의 **오차**에서 현 시점 데이터의 상태를 추론

ARMA(AutoRegressive Moving Average)

AR과 MA을 섞은 모델 → 상태와 오차 두 관점에서 과거의 데이터를 사용

$$Z_t = a + \Phi_1 Z_{t-1} + \dots + \Phi_p Z_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} + a_t$$

ARIMA(AutoRegressive Integrated Moving Average)

자기 회귀와 이동 평균을 둘 다 고려하며 ARMA와 달리 과거 데이터의 선형 관계 + 추세가 고려한 모델

추세는 자기 자신(정상 데이터)의 추세만 반응하며 white_noise는 고려하지 않음

추세 관계는 공적분을 고려한 개념으로 X-Y간 cointegration >0 이면 X값이 이전 값보다 증가하면 Y값도 이전 값보다 증가함

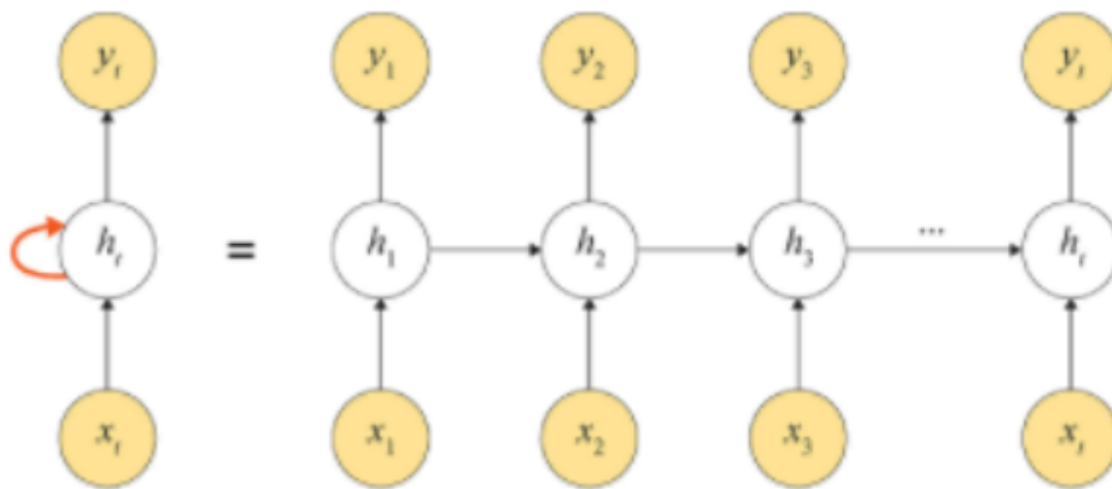
⇒

통계 분석 패키지인 statsmodels 라이브러리를 통해 사용 가능

ex) ARIMA(5,1,0) = 자기회귀 차수 5, 차분 차수 1, 이동평균 차수 0

7.3 순환 신경망(RNN)

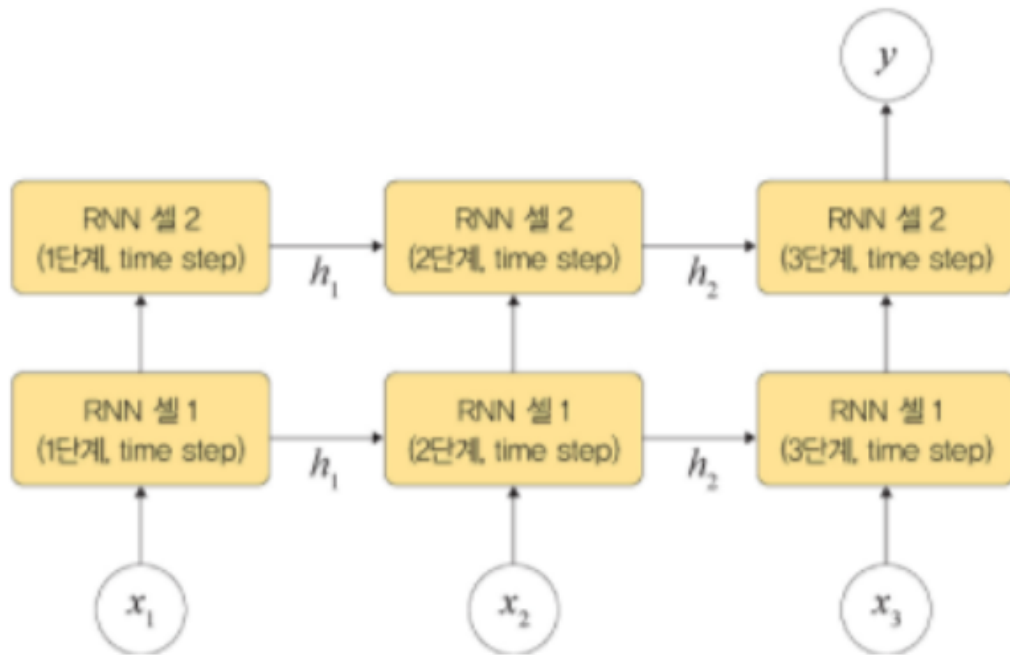
- RNN은 시간적으로 연속성이 있는 데이터를 처리하기 위해 고안된 인공 신경망
- 이전 은닉층이 현재 은닉층의 입력이 되는 “순환” 구조
- 현재까지 입력 데이터를 요약한 정보인 “기억”을 가진다는 것이 큰 특징 → 새로운 입력이 들어올 때마다 기억이 수정되며 최종 기억은 모든 입력을 요약한 정보가 됨



- 외부 입력과 자신의 이전 상태를 입력받아 현재 상태를 갱신

입출력에 따른 여러 유형

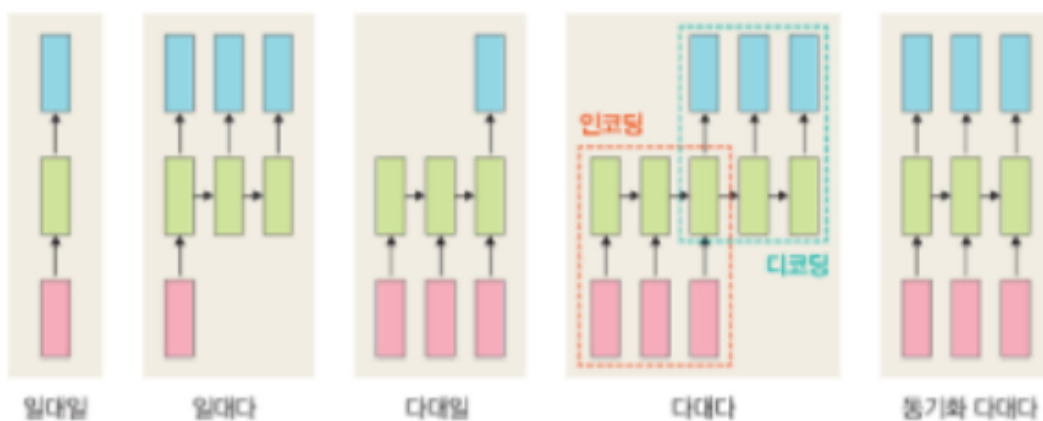
- 일대일 : 순환이 없으며 RNN 아님.
 - 예) 순방향 네트워크
- 일대다 : 입력 하나, 출력 다수.
 - 예) 이미지 캡션 : 이미지 입력을 통해 설명을 문장으로 출력
- 다대일 : 입력 다수, 출력 하나.
 - 예) 감성 분석기 : 문장 입력해서 긍부정 출력



▲ 그림 7-6 적층된 다대일 모델

⇒ 다대일 구조에 층을 쌓아 올라 적층 구조를 만들 수 있음

- 다대다 : 입력 다수, 출력 다수.
 - 예) 자동 번역기
- 동기화 다대다 : 입출력 다수.
 - 예) 문장 다음에 나올 단어 예측하는 언어 모델, 프레임 수준 비디오 분류

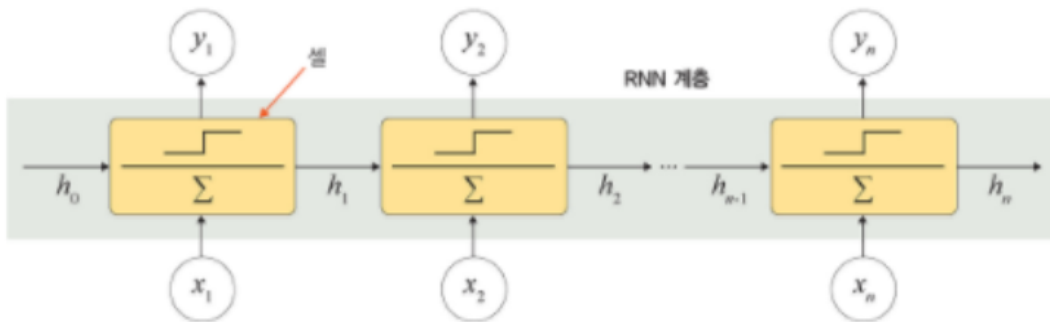


▲ 그림 7-8 RNN 모델 유형

RNN layer

RNN cell

- 하나의 단계만 처리 → RNN 계층의 for loop 구문을 갖는 구조



▲ 그림 7-9 RNN 계층과 RNN 셀

- RNN 계층은 셀을 래핑하여 같은 셀을 여러 단계에 적용
- 셀은 단일 입력과 과거 상태를 가져와 출력과 새로운 상태를 생성함
 - nn.RNNCell: SimpleRNN 계층에 대응되는 RNN 셀
 - nn.GRUCell: GRU 계층에 대응되는 GRU 셀
 - nn.LSTMCell: LSTM 계층에 대응되는 LSTM 셀
- 파이토치에서는 셀과 계층을 분리하여 구현가능

RNN 활용 분야

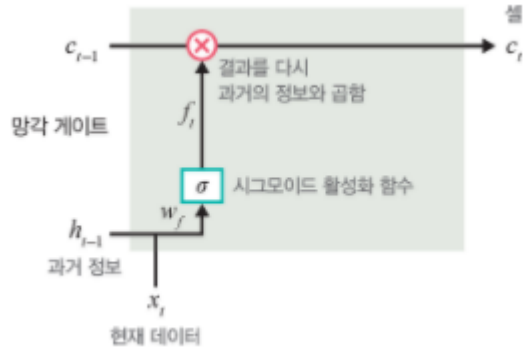
- 자연어 처리 → 음성 인식, 단어 의미 판단 및 처리
- 시계열 데이터 처리 → 센서 데이터

7.4 RNN 구조

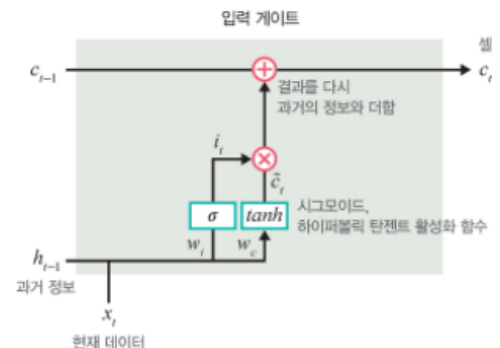
7.5 LSTM

- RNN의 기울기 소멸 문제 해결
1. 순전파
 - 망각게이트, 입력게이트, 출력게이트 은닉층 각 뉴런에 추가
 - 망각게이트 : 과거 정보를 어느 정도 기억할지 결정. 과거 정보 & 현재 데이터
 - 입력게이트 : 현재 정보 기억. 과거 정보 & 현재 데이터 입력받아 시그모이드

입력받아 시그모이드 취한 값을 과거 정보에 곱함 → 1이면 취하고 0이면 버림

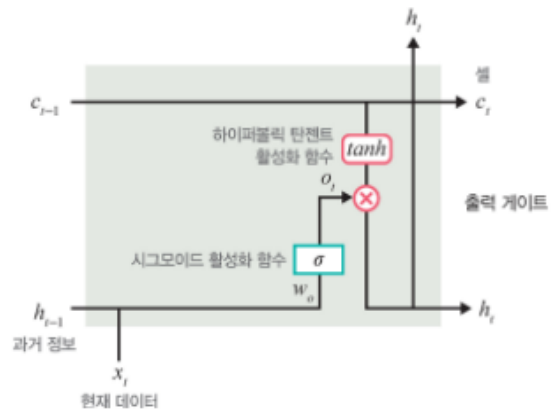
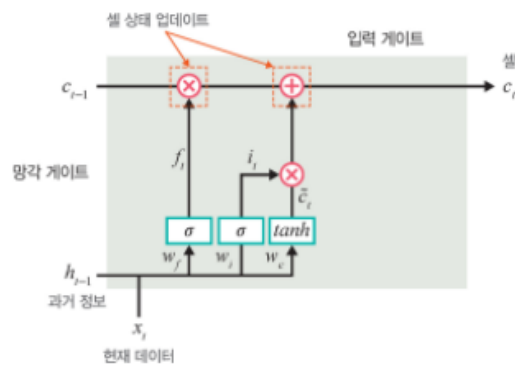


와 하이퍼볼릭 탄젠트 함수 기반으로 현 정보 보존량 결정. 1이면 새로운 정보 반영 허용

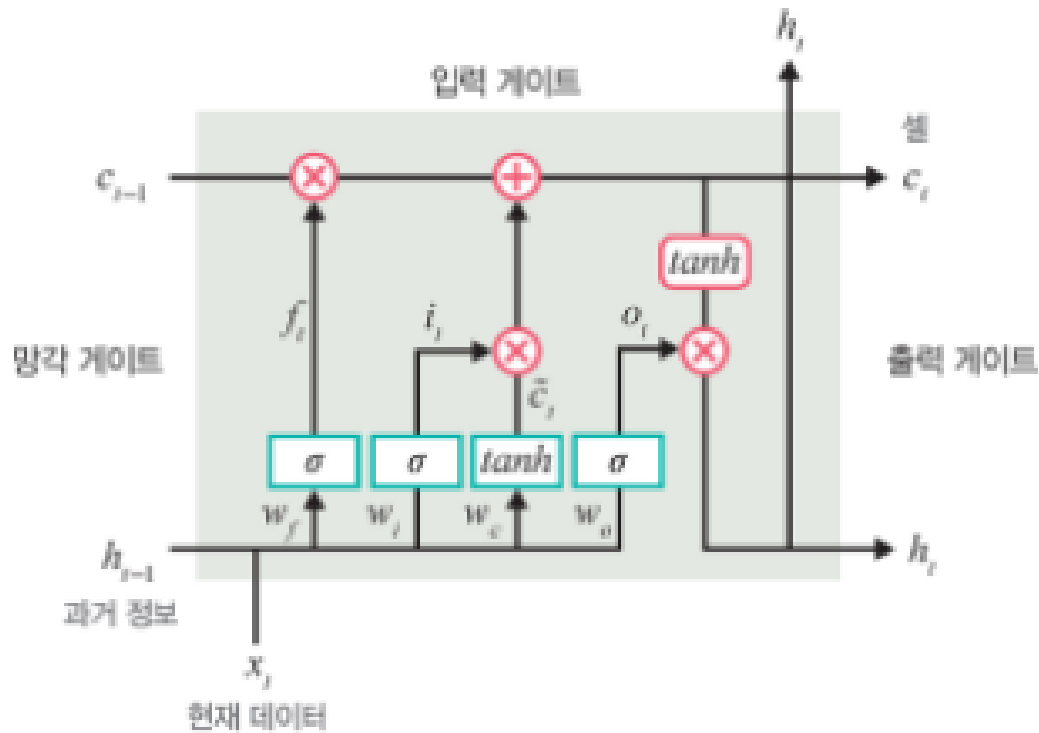


- 메모리 셀 : 은닉 노드. 총합을 통해 소멸 문제 해결
- ⇒ 망각 게이트와 입력 게이트의 전 단계 셀 정보 계산하여 업데이트

- 출력 게이트 : 과거 정보 & 현재 데이터로 뉴런 출력 결정
- 이전 은닉 상태와 현 시점 입력을 통해 다음 은닉 상태 계산. 1이면 의미있으므로 출력

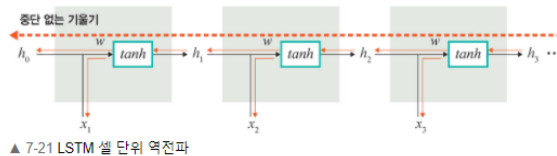


- LSTM 전체 게이트



2. 역전파

- 셀을 통해 수행. = 중단없는 기울기
- 최종 오차 모든 노드에 전파. 셀 내부적으로는 오차가 입력 방향으로도 전파됨



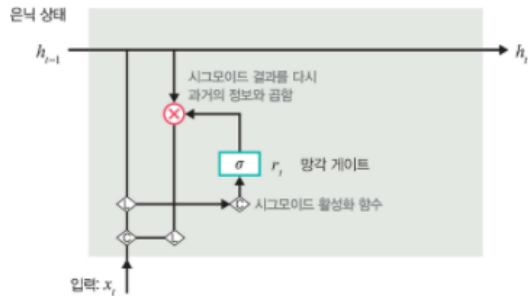
▲ 7-21 LSTM 셀 단위 역전파

$$\begin{aligned}
 t_t &= \tanh(w_{hh}h_{t-1} + w_{sh}x_t) \\
 &= \tanh((w_{hh} \quad w_{sh}) \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix}) \\
 &= \tanh(w \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix})
 \end{aligned}$$

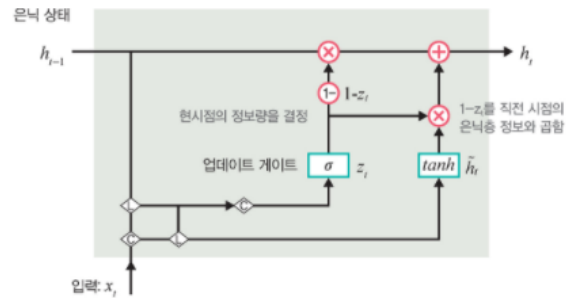
7.6 GRU_Gated Recurrent Unit

- LSTM의 망각&입력 게이트를 하나로 합치고 업데이트 게이트가 존재
- 게이트 컨트롤러가 망각&입력 게이트 제어. 1 출력시 망각 열고 입력 닫힘
- 이전 기억 저장될 때마다 단계별 입력 삭제됨
- 출력 게이트 없고, 매 단계마다 전체 상태 벡터 출력

- 망각 게이트 : 과거 정보 적당히 초기화. 시그모이드 출력값을 이전 은닉층에 곱함



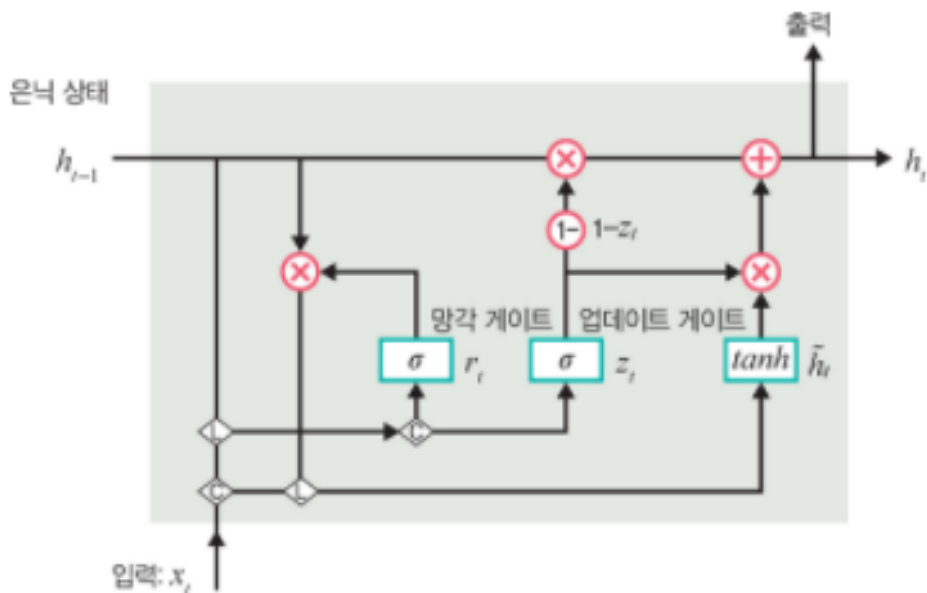
- 업데이트 게이트 : 과거 현재 정보의 최신화 비율 결정. 시그모이드 출력 결과는 현 정보량 결정하고, 1에서 뺀 값을 직전 은닉층 정보와 곱함



- 후보군 : 현 시점 정보에 대한 후보군 계산. 망각 게이트 결과 이용

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t])$$

- GRU 전체 내부 구조



7.7 양방향 RNN

- 이전 시점 & 이후 시점 데이터를 활용해 출력값 예측

- 두 메모리 셀 사용 : 이전 은닉 상태로 현 은닉 상태 계산 + 다음 시점 은닉 상태로 현 은닉상태 계산