

7. 시계열

7.1 시계열 문제

- 정의:
 - 시간에 따라 관측된 데이터의 변화를 분석하거나 예측하는 문제. 데이터의 시간 순서가 중요하며, 순서를 고려한 모델링이 필요.
- 특징:
 - **자기상관성**: 과거 데이터와 현재 데이터 사이의 관계 존재.
 - **정상성(Stationarity)**: 평균, 분산, 자기공분산이 시간에 따라 일정한 데이터의 특성.
 - **비정상성(Non-Stationarity)**: 데이터의 평균 또는 분산이 시간에 따라 변하는 특성.
- 분해 기법:
 - 시계열 데이터는 추세(Trend), 계절성(Seasonality), 잔차(Residual)로 분해 가능.
- 활용 사례:
 - 금융 데이터: 주가 및 환율 예측.
 - 날씨 데이터: 온도, 강수량, 바람 예측.
 - 경제 데이터: 소비자 행동 예측, 수요 계획.

7.2 AR, MA, ARMA, ARIMA

- AR (Auto-Regressive):

- 현재 값이 과거 데이터의 선형 결합으로 표현.
- p : 고려할 과거 시점의 수 (차수).
- **적합 데이터**: 정상성을 가지는 데이터.
- **MA (Moving Average):**
 - 현재 값이 과거 오차의 선형 결합으로 표현.
 - q : 고려할 과거 오차의 수 (차수).
 - **적합 데이터**: 정상성을 가지는 데이터.
- **ARMA (Auto-Regressive Moving Average):**
 - AR과 MA를 결합한 모델.
 - 파라미터 (p, q) : AR의 차수와 MA의 차수.
 - **제한**: 데이터가 정상적이어야 함.
- **ARIMA (Auto-Regressive Integrated Moving Average):**
 - 비정상성을 가진 데이터에서 차분(differencing)을 통해 정상성을 확보한 후 ARMA 모델 적용.
 - 파라미터 (p, d, q) :
 - p : AR 차수.
 - d : 차분 횟수.
 - q : MA 차수.
 - **적용 과정**:
 1. 데이터가 비정상적이면 차분 수행.
 2. 정상성 확보 후 ARMA 모델링.

7.3 순환 신경망 (RNN)

- **정의**:
 - 시계열 데이터, 자연어, 음성과 같은 순차적 데이터를 학습하는 신경망 구조.
 - 이전 은닉 상태를 현재 입력과 결합하여 다음 상태를 계산.
- **특징**:

- 시점 간 관계 학습.
- 데이터의 시간 순서를 보존.
- 작동 원리:

은닉 상태 h_t : 현재 입력 x_t 와 이전 상태 h_{t-1} 의 함수.

$$h_t = f(W_h h_{t-1} + W_x x_t + b)$$

출력 y_t : 은닉 상태를 기반으로 계산.

$$y_t = g(W_y h_t + c)$$

- 문제점:
 - 기울기 소실/폭발 문제: 긴 시퀀스를 학습할 때 역전파 과정에서 기울기가 소실되거나 폭발.

7.4 RNN 구조

- 구성 요소:
 - 입력: 각 시간 단계에서의 입력 x_t .
 - 은닉 상태: 시간 단계별로 계산된 내부 상태 h_t .
 - 출력: 각 시간 단계에서 출력 y_t .
- 데이터 흐름:
 - 입력 데이터는 시간 축을 따라 순차적으로 처리.
 - 은닉 상태는 각 시간 단계에서 갱신되어 다음 단계로 전달.
- 한계:
 - 긴 시퀀스 데이터를 학습하기 어렵고, 장기 의존성을 잘 학습하지 못함.

7.5 LSTM (Long Short-Term Memory)

- 정의:

- RNN의 기울기 소실 문제를 해결하기 위해 설계된 신경망 구조.
- 장기 기억과 단기 기억을 관리하는 **게이트** 구조 포함.
- 구조:
 - **셀 상태**: 정보를 장기적으로 유지하는 경로.
 - **게이트**: 입력, 망각, 출력 게이트로 구성.
- 게이트 동작:

망각 게이트 (f_t):

- 셀 상태에서 제거할 정보 결정.

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

입력 게이트 (i_t):

- 셀 상태에 추가할 정보 결정.

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

셀 상태 갱신:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tanh(W_c[h_{t-1}, x_t] + b_c)$$

출력 게이트 (o_t):

- 다음 은닉 상태로 전달할 정보 결정.

$$h_t = o_t \odot \tanh(C_t)$$

7.6 게이트 순환 신경망 (GRU)

- 정의:
 - LSTM보다 간단한 구조를 가진 RNN으로, 계산 효율이 높음.
 - 은닉 상태와 셀 상태를 하나로 통합.
- 구조:

- 두 개의 게이트:
 - 업데이트 게이트 (z_t): 이전 상태를 유지할 비율 결정.

$$z_t = \sigma(W_z[h_{t-1}, x_t] + b_z)$$

- 리셋 게이트 (r_t): 과거 정보를 얼마나 잊을지 결정.

$$r_t = \sigma(W_r[h_{t-1}, x_t] + b_r)$$

- 은닉 상태 갱신:

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

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \tilde{h}_t$$

7.7 양방향 RNN (Bidirectional RNN)

- 정의:
 - 데이터를 순방향과 역방향으로 동시에 학습하는 RNN 구조.
 - 과거와 미래 정보를 모두 활용.
- 구조:
 - 순방향 RNN:

$$h_t^f = f(W_f h_{t-1}^f + W_x^f x_t + b_f)$$
 - 역방향 RNN:

$$h_t^b = f(W_b h_{t+1}^b + W_x^b x_t + b_b)$$
 - 출력:

$$y_t = g(W_y[h_t^f, h_t^b] + c)$$
- 장점:
 - 입력 데이터의 과거와 미래 정보를 모두 반영하여 정확한 결과 도출.
 - 자연어 처리, 음성 인식에서 성능 향상.
- 응용 사례:
 - 기계 번역, 감정 분석, 음성-텍스트 변환.