|  |  |  |
| --- | --- | --- |
| **| 학습주제** | | * 선형 회귀 모델을 활용한 교통량 예측 |
| **| 학습목표** | | * 선형 회귀 모델을 사용하여 교통량 데이터를 학습하고 예측하는 방법을 익힌다. * 학습 데이터와 테스트 데이터를 나누는 방법을 이해한다. * R² 점수를 사용하여 모델의 예측 성능을 평가하는 방법을 배운다. |
| **| 학습 개념** | 선형 회귀 모델은 주어진 데이터를 기반으로 독립 변수(X)와 종속 변수(y) 간의 선형 관계를 학습하여 예측을 수행합니다. 이 모델은 새로운 입력값을 주었을 때, 기존 학습 데이터를 바탕으로 가장 적합한 선을 찾아 예측 결과를 반환합니다. scikit-learn의 LinearRegression 클래스를 사용하면 손쉽게 모델을 정의하고 학습할 수 있으며, 학습 데이터를 분할하는 train\_test\_split 함수와 함께 사용할 수 있습니다. 또한, R² 점수는 모델이 데이터를 얼마나 잘 설명하는지를 나타내는 지표로, 성능 평가에 사용됩니다. | |
| **| 학습 방향** | 서울시 교통량 데이터를 이용하여 요일을 독립 변수로, 0시 교통량을 종속 변수로 설정하여 선형 회귀 모델을 학습하고, 그 성능을 평가하세요. 이 과정에서 데이터를 필터링하고 학습용과 테스트용으로 분할한 후, R² 점수를 통해 모델의 예측 성능을 확인하세요. | |
| **| 요구 사항** | 1. 데이터 불러오기: pandas의 read\_excel 함수를 사용하여 '2023년 01월 서울시 교통량.xlsx' 파일을 불러옵니다.  2. 데이터 필터링: '성산로(금화터널)' 지점의 '유입' 방향 데이터만 사용하도록 필터링하세요. Boolean indexing을 사용하세요. 피처와 타깃 값 설정: '요일'을 독립 변수로, '0시' 시간대의 교통량을 종속 변수로 설정하세요.  3. 학습 데이터와 테스트 데이터 분할: train\_test\_split 함수를 사용하여 데이터를 8:2 비율로 나누세요. 8이 학습 데이터이고 2가 테스트 데이터입니다.  4. 선형 회귀 모델 정의 및 학습: LinearRegression 클래스를 사용하여 모델을 정의하고 학습 데이터를 사용해 모델을 학습시키세요. 모델 평가: R² 점수를 계산하여 모델의 성능을 평가하세요.  5. 선형 회귀 모델 정의 및 학습: LinearRegression 클래스를 사용하여 모델을 정의하고, 학습 데이터를 사용해 모델을 학습시키세요.  6. 모델 평가: 모델을 평가하기 위해 테스트 데이터를 사용하여 R² 점수를 출력하세요. | |