|  |  |
| --- | --- |
| **| 문제의도** | * Python에서 **선형 회귀**와 **다항 회귀** 모델을 사용하여 교통량 데이터를 예측하고, 두 모델의 성능을 비교하는 과정을 학습하는 것이 목표입니다. * 이를 통해 선형과 비선형 데이터 관계를 처리하는 방법을 익히고, 두 회귀 모델의 예측 성능을 정량적으로 평가하는 능력을 키우는 것이 목적입니다. |
| **| 사전지식** | * pandas와 numpy를 사용하여 데이터를 처리하는 방법 * 선형 회귀(Linear Regression)와 다항 회귀(Polynomial Regression)에 대한 기본 개념 * train\_test\_split을 사용한 데이터 분리 및 모델 학습 * 평균 제곱 오차(MSE)와 루트 평균 제곱 오차(RMSE)를 통한 모델 성능 평가 |
| **| 문제핵심** | * 선형 회귀와 다항 회귀의 차이를 이해하고, 각 모델을 학습시키는 방법 * 두 모델의 예측 성능을 MSE 및 RMSE로 평가하고, 성능을 비교하는 과정 * PolynomialFeatures를 사용하여 데이터를 다항식으로 변환하는 방법 |

|  |  |
| --- | --- |
| **| 문제해설** | |
| **| 해설** | 1. 데이터 준비 및 전처리   - pandas를 사용하여 데이터를 불러오고, 날짜 인덱스(독립 변수)와 8시 교통량(종속 변수) 데이터를 설정합니다.  - numpy의 arange()로 날짜 인덱스를 생성하고, train\_test\_split을 사용해 훈련 데이터와 테스트 데이터를 나눕니다.   1. 선형 회귀 모델 학습   - LinearRegression 객체를 생성하고, 훈련 데이터를 학습시켜 선형 회귀 모델을 완성합니다.  - 테스트 데이터를 사용해 예측을 수행하고, MSE와 RMSE로 성능을 평가합니다.   1. 다항 회귀 모델 학습   - PolynomialFeatures로 2차 다항식을 생성하고, 다항 회귀 모델을 학습합니다.  - 다항 회귀 모델도 동일하게 성능을 평가합니다.   1. 성능 비교 및 평가   - 두 모델의 MSE와 RMSE를 계산한 뒤, 두 값의 크기를 비교하여 어느 모델이 더 나은 예측 성능을 보였는지 확인합니다.  - 선형 회귀와 다항 회귀의 차이점을 분석하고, 데이터에 대한 더 적합한 모델을 도출합니다. |