|  |  |
| --- | --- |
| **| 문제의도** | * 성능 평가 지표를 사용하여 모델의 성능을 다각도로 평가하는 방법을 익히는 것이 목적입니다. 성능 평가 지표로 \*\*혼동 행렬(Confusion Matrix), 정밀도(Precision), 재현율(Recall), F1 점수(F1 Score), 정확도(Accuracy)\*\*를 사용합니다. |
| **| 사전지식** | * K-최근접 이웃(KNN) 알고리즘의 개념 * pandas, numpy, scikit-learn 사용법 * confusion\_matrix, classification\_report, accuracy\_score 함수의 활용 방법 |
| **| 문제핵심** | * KNN 모델 학습: KNeighborsClassifier를 사용해 모델을 학습하고, 테스트 데이터로 예측을 수행하는 과정 * 모델 평가: confusion\_matrix, classification\_report, accuracy\_score를 사용하여 모델 성능을 평가하고, 혼잡 여부 예측의 성능을 확인하는 과정 |

|  |  |
| --- | --- |
| **| 문제해설** | |
| **| 해설** | 1. 모델 성능 평가  - 혼동 행렬(Confusion Matrix)를 사용해 모델이 예측한 값과 실제 값을 비교하여 True Positive, True Negative, False Positive, False Negative 값들을 확인합니다.  - 분류 성능 지표(Classification Report)에서는 정밀도(Precision), 재현율(Recall), F1 점수(F1 Score)가 출력되며, 이를 통해 모델의 성능을 더 깊이 평가할 수 있습니다.  - 정확도(Accuracy)는 전체 데이터에서 모델이 얼마나 정확하게 예측했는지를 나타내는 지표로, accuracy\_score()를 사용하여 계산합니다. |