

📌 모델 베이스라인 요약 (2025-07-07 기준)

✔ 데이터셋

- 총 샘플 수: 약 20만 건
- 사용 데이터: 전처리된 df\_processed (dummy encoding, robust scaling 등 적용)

◆ 분류 모델 (XGBoostClassifier)

Target	Accuracy	Macro F1	비고
post_charge_departure_range	0.45	0.14	클래스 불균형 존재
usage_departure_range	0.47	0.13	대부분 클래스 1로 예측됨 일부 분류 개선됨

◆ 회귀 모델 (XGBoostRegressor)

Target	MSE	Mean (정답)	Std (정답)	설명
kwh_per_usage_time	62.65	27.36	28.08	평균 예측 대비 MSE 90%↓ 성능

⚙ 기타 정보

- 피처 중요도 기준으로 station\_location, evse\_name 등 제거 후 성능 소폭 향상
- 군집 기반 피처(cluster), 시계열 파생 피처 활용
- 모델: XGBoost (sklearn pipeline + ColumnTransformer로 구성)