

# 과제 #1 완료 보고서

## 모델 선택 및 성능 비교

### 1. SVR 모델

- 하이퍼파라미터:  $C=30000$ ,  $\text{kernel}=\text{"linear"}$
- RMSE: 70158.80248353248
- 교차 검증 RMSE 점수: [68813.3993301, 83178.17296742, 70915.38578773]
- 평균 RMSE: 74302.31936175143
- 표준 편차: 6334.570213729821

### 2. 다른 알고리즘 - 그레디언트 부스팅 회귀 (Gradient Boosting Regressor)

- RMSE: 50651.5088944892
- 교차 검증 RMSE 점수: [51840.72867638, 53489.21682285, 54089.99141937]
- 평균 RMSE: 53139.97897286685
- 표준 편차: 950.8841600156443

## 기타 모델 성능 비교

- 결정 트리 회귀 (Decision Tree Regressor)
  - 평균 RMSE: 71629.89009727491
  - 표준 편차: 2914.035468468928
- 선형 회귀 (Linear Regression)
  - 평균 RMSE: 69104.07998247063
  - 표준 편차: 2880.3282098180666
- 랜덤 포레스트 회귀 (Random Forest Regressor)
  - 평균 RMSE: 50435.58092066179
  - 표준 편차: 2203.3381412764606

## 새로운 모델이 SVR 모델보다 우수한 이유

### 1. 더 낮은 RMSE:

그레디언트 부스팅 회귀 모델의 평균 RMSE가 SVR 모델보다 21162.340 낮아졌습니다. 이는 그레디언트 부스팅이 전반적인 예측 정확도에서 SVR 모델보다 우수하다는 것을 나타냅니다.

### 2. 더 작은 표준 편차:

그레디언트 부스팅 회귀 모델의 표준 편차는 950.8841600156443이며, SVR 모델의 표준 편차는 6334.570213729821입니다. 표준 편차가 작을수록 모델의 예측 결과가 안정적입니다. 그레디언트 부스팅 모델의 결과는 더 일관적이고 신뢰할 수 있습니다.

### 3. 앙상블 학습의 장점:

그레디언트 부스팅 회귀 모델은 앙상블 학습의 장점을 활용하여 여러 번의 반복을 통해 모델을 점진적으로 개선하여 편향과 분산을 크게 줄입니다. 이 방법은 복잡한 비선형 관계를 처리할 때 특히 뛰어납니다.

### 4. 강건성:

그레디언트 부스팅 회귀 모델은 이상값과 잡음 데이터에 대해 높은 강건성을 가지고 있습니다. 이는 각 단계에서 가중치를 통해 오류의 영향을 줄이려고 시도하기 때문입니다. 이로 인해 모델이 실제 응용에서 더 실용적이고 신뢰할 수 있습니다.

## 제출 내용

- **코드:** SVR 모델과 그레디언트 부스팅 회귀 모델의 구현 코드 포함
- **보고서:** 두 모델의 성능을 비교하고 자세한 RMSE 점수 및 통계 데이터 제공
- **실행 환경:**
  - sklearn 버전: 1.5.0
  - python 버전: 3.12.4
  - 실행 환경: 로컬 환경 (Windows 64-bit)