[개강했조] 멘토멘티 7주차 보고서

멘토멘티 정보			
팀 명	개강했조	활동 차시	7주차
이번 차시 활동 요약 (학습목표, 달성도 등)	캐글 모의대회 참여 및 데이터 분석 능력 증진 빅데이터 분석기사 실기 작업형 2 기반 베이스라인 작성	성 연습	
모임 일시 및	2024.11.08 금요일 11:00 - 13:00		
진행장소	컴퓨터공학관 2층 창의마루 회의실 A		

구성원 명단 및 참석여부				
연번	구분	이름	참석 여부(○/×)	불참 시 작성 (예비군,병결 등 증빙자료첨부)
1	멘토	신병근	О	불참사유:
2	멘티	김진우	О	불참사유:
3	멘티	김대욱	О	불참사유:
4	멘티	송민정	О	불참사유:
5	멘티	이서진	О	불참사유:
6	멘티	정유진	0	불참사유:

학습 주제 및 내용		
학습 목표	베이스라인 작성 능력 기르기 및 분석 능력 다각화	
학습 내용	- [House Prices - Advanced Regression Techniques Data], - 2024-3 Big Data Analytics Certification KR Competition ● 집값 예측 모델 만들기 대회 ● 베이스라인 제작을 위해 문자형 컬럼은 건드리지 않고 수치형만 랜덤포레스트로 학습시킴. ● target값의 분포를 확인하고 왜도>0이고 우측꼬리가 긴 histogram을 보아 log1p 변환을 수행하여 정규분포로 변형 ● Imputer 활용한 결측치 채우기 ● RandomForestRegressor을 이용하여 집 가격 같은 회귀형 변수 예측 Scatter Plot of Predicted vs. Actual Sale Prices With Regression Line Pegression Line Actual Sale Prices Fregression Line 100000 Actual Sale Prices Fregression Line 100000 Pata Prices Catter Plot of Predicted vs. Actual Sale Prices With Regression Line 100000 Catter Plot of Predicted vs. Actual Sale Prices Negression Line 100000 Actual Sale Prices With Regression Line 100000 Actual Sale Prices Actual Sale Prices Actual Sale Prices Actual Sale Prices 100000 Actual Sale Prices Actual Sale Prices 100000 Actual Sale P	
	● 이상치의 케이스를 분리하여 모델에 편입하는 것을 이용한 (리 샘플링) 아웃라이어 처리, MinMaxScaler를 이용한 정규화	

• RandomForestClassifier 분류기 학습

멘토 자체 평가

이번 스터디를 통해 데이터 분석에서의 베이스라인 작성과 모델링의 중요성을 체감했음. House Prices 프로젝트에서는 target값의 왜도를 조 정해 모델 성능을 높였고, Home Credit Default Risk에서는 불균형 데이 터를 다루며 리샘플링과 PCA의 필요성을 경험함.

학습 참고자료 출처

https://www.kaggle.com/competitions/big-data-analytics-certification-kr-2024-

https://www.kaggle.com/c/house-prices-advanced-regression-techniques

https://www.kaggle.com/competitions/home-credit-default-risk/ https://www.kaggle.com/competitions/big-data-analytics-certification-kr-

2024-2/code

멘티	성찰	내용
----	----	----

(1차활동-3주차, 6주차 2차 활동-9주차,12주차에만 작성)		
연번	이 름	성찰 내용
1		
2		
3		
4		

활동사진(첨부)

활동을나누어 진행했을 경우, 추가적으로 작성

시작 시간 11:00

종료 시간 13:00

[개강했조 주차보고서],		