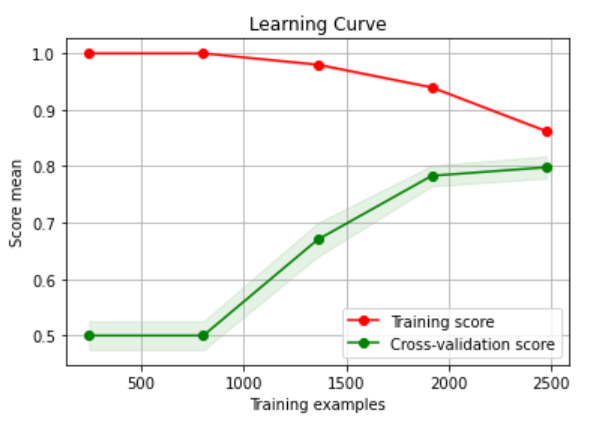
신어감지모델

Data의 Imbalanced Problem을 해결하기 위해서 UnderSampling을 진행하고 모델 학습 및 평가를 진행하였다.

KNN, Logistic Regression, Decision Tree, Random Forest, XGBClassifier간의 GridSerchCV 결과XGBClassifier의 성능이 가장 좋게 나와 최종 Model은 XGBClassifier 사용하였다. 또한, 각각에 대해 RFE를 적용해본 결과 더 좋은 성능을 보이지 못해서 RFE는 사용하지 않았다.

K-Fold Cross Validation 결과 Accuracy는 약 80%, F1-Score도 역시 약 80% 정도로 측정되었다.



Learning Curve를 그려본 결과, Overfitting은 일어나지 않은 것으로 보인다.