**2조 최종보고서**

(1) 2조 예비보고서

**1. 프로젝트 제목**

랜덤포레스트를 이용한 공공 자전거 수요 예측

**2. 프로젝트 시작 계기**

현재 전세계에는 500개가 넘는  공공 자전거 프로그램이 있다.  이러한 공공 자전거 프로그램은, 어플리케이션, 혹은 키오스크를 통해서, 회원가입, 대여, 반납 등을 진행할 수 있다. 이런 일련의 시스템을 초기에 구축을 한다고 해서, 이 프로그램이 지속적으로 원활히 돌아가는것은 아니다.  이 프로그램을 지속적으로 원활히 진행되기 위해서는 정확한  ‘자전거 수요량 예측’이 필요하다. 공공 자전거는 흔적을남긴다. 즉, 앞서 언급한 시스템에 의해서 다양한 데이터가 기록된다는 것이다. 예를 들면, 출발지, 도착 위치 및 경과 시간 등의 데이터가명시적으로 기록된다. 이렇게 기록된 과거의 데이터의 패턴과, 날씨 데이터를 결합하여, 자전거 대여 수요를 예측해보기로 했다.

**3. 프로젝트 개요**

* 이 프로젝트는 <https://www.kaggle.com/c/bike-sharing-demand/data> 에서 제공하는 데이터셋을 바탕으로 진행될 것이다.
* 이 프로젝트는 [https://github.com/corazzon/KaggleStruggle/blob/master/bike-sharing-demand/bike-sharing-demand-EDA.ipynb,](https://github.com/corazzon/KaggleStruggle/blob/master/bike-sharing-demand/bike-sharing-demand-EDA.ipynb) <https://github.com/corazzon/KaggleStruggle/blob/master/bike-sharing-demand/bike-sharing-demand-EDA.ipynb>을 바탕으로 진행될 것이다.
* 랜덤 포레스트 회귀를 이용하여, 자전거 수요 예측을 진행될 것이다.
* RMSLE function을 이용하여 과대평가 된 항목보다 과소평가 된 항목에   패널티를 주는 방식을 사용할 것이다.
* Cross Validation의 과정을 KFold 방식으로 진행할 것이다.

**4.기대효과**

* 학습자가 랜덤포레스트라는 기계 학습 기법을 직접 적용하여 학습할 수 있다.
* 정확한 수요 예측을 통해서, 수요에 맞는 적절한 자전거 공급을 할 수 있다.
* 이후에, 공공 자전거 시스템을 구축하고 도입하려는 공공 자전거 프로그램 기획자에게, 하나의 머신러닝 실습사례로서 인사이트를 줄 수 있다.

(2) 프로젝트 작업파일

* 작업파일 : [bike-sharing-demand-EDA\_.ipynb](/redirect/s3?bucket=uploads&prefix=paste%2Fklxd5q24dv14ta%2Fbc71a218e3e63321257163ecf599a30a42cdc552b6f62bc6b288c0279bfbf354%2Fbike-sharing-demand-EDA_.ipynb) / [bike-sharing-demand-random\_forestipynb.ipynb](/redirect/s3?bucket=uploads&prefix=paste%2Fklxd5q24dv14ta%2Fde61c2af2e604c781d4c4792da1195e86fe6fab93c171a4b8971435bcb50adf1%2Fbike-sharing-demand-random_forestipynb.ipynb)
* PPT :[\_\_\_\_2\_\_\_\_\_\_\_\_\_\_\_\_.pptx](/redirect/s3?bucket=uploads&prefix=paste%2Fksxrz24c4dz7lx%2F1ed6dbd56dd5443534d97138d8f6e5409a755db50a605dee7bb38dc028d9ff81%2F____2____________.pptx)
* Source code : <https://github.com/corazzon/KaggleStruggle/blob/master/bike-sharing-demand/bike-sharing-demand-EDA.ipynb>, <https://github.com/corazzon/KaggleStruggle/blob/master/bike-sharing-demand/bike-sharing-demand-rf.ipynb>
* Data: <https://www.kaggle.com/c/bike-sharing-demand/data>

(3) 동영상 발표

* <https://youtu.be/T0tbCZzZwxA>