

Kaggle Project 계획안

Presenter :

기초학부 한현영

Index

- I. 주제
- II. 팀 소개
- III. 일정 정리
- IV. 데이터 설명
- V. 예측

주제

KBO 정규시즌 팀별 승률, 타율 및 방어율(평균자책점) 예측

- 스포츠투아이에서 제공하는 야구데이터(팀 통산 전적자료, 연도별 전적자료, 팀 및 선수 기록 등)를 활용
- 정규시즌 잔여 경기에 대한 각 팀별 승률, 타율 및 방어율(평균자책점) 예측

팀 소개

- Team name
 - 현풍데이타(Hyeonpung Data)
- Members
 - 고낙헌(19')
 - 김주형(19')
 - 정회성(19')
 - 최세영(19')
 - 한현영(19')
- Collaborate through Git

일정 정리

- Week1(8/3 ~ 8/9)
 - 주제 선정 및 탐색적 데이터 가공
- Week2(8/10 ~ 8/16)
 - 승률 예측 1(pythagorean method 이용)
- Week3(8/17 ~ 8/23)
 - 데이터 가공 및 승률 예측 2(K-mean algorithm, linear regression 이용)
- Week4(8/24 ~ 8/27)
 - 승률 예측 3(...)

데이터 설명

- 스포츠투아이에서 제공하는 데이터를 바탕으로 진행
 - 개인타자 (2016-2020)
 - 개인투수 (2016-2020)
 - 경기 (2016-2020)
 - 등록선수 (2016-2020)
 - 선수 (2016-2020)
 - 팀 (2016-2020)
 - 팀타자 (2016-2020)
 - 팀투수 (2016-2020)

예측

- 승률과 연관 있는 데이터를 바탕으로 데이터 전처리 진행
- 현재 다양한 ML algorithm을 사용해보며 높은 정확치를 도출하는 과정에 있음
- 예측 정확도가 높을 시에 서빙할 예정