

Parrot 9기 2차 프로젝트: ML

1. Dataset

- ✓ Id : people's id
- ✓ year : year of study
- ✓ wave : from wave 1st in 2005 to wave 14th in 2018
- ✓ region : 1) Seoul 2) Kyeong-gi 3) Kyoung-nam 4) Kyoung-buk 5) Chung-nam 6) Gang-won & Chung-buk 7) Jeolla & Jeju
- ✓ income : yearly income in M KRW(Million Korean Won. 1100 KRW = 1 USD)
- ✓ family_member : number of family members
- ✓ gender : 1) male 2) female
- ✓ year_born
- ✓ education_level : 1) no education(under 7 yrs-old) 2) no education(7 & over 7 yrs-old) 3) elementary 4) middle school 5) high school 6) college 7) university degree 8) MA 9) doctoral degree
- ✓ marriage : marital status. 1) not applicable (under 18) 2) married 3) separated by death 4) separated 5) not married yet 6) others
- ✓ religion : 1) have religion 2) do not have
- ✓ occupation : this will be provided in separated code book
- ✓ company_size
- ✓ reason_none_worker : 1) no capable 2) in military service 3) studying in school 4) prepare for school 5) prepare to apply job 6) house worker 7) caring kids at home 8) nursing 9) giving-up economic activities 10) no intention to work 11) others

2. 특이사항

- ✓ 이 데이터셋은 동일한 사람을 일정 주기로 임금을 분석한 데이터입니다.
- ✓ 설명에 나와 있지 않은 이상한 변수들이 포함되어 있습니다. 해당 변수들을 어떻게 처리할지 조원들과 상의하여 전처리해주세요.
- ✓ 모르는 부분은 편하게 멘토에게 물어봐 주시기 바랍니다!

3. 발표 방법 및 날짜

- ✓ 발표에 정해진 양식은 없습니다. 원하는 형태로 자유롭게 발표해 주세요.

(ppt를 만들어서 발표하셔도 되고, .ipynb로 발표하셔도 됩니다.)

- ✓ **중간 발표** 때는 전처리를 어떻게 했는지 위주로 설명해 주세요! 추가적으로, 어떤 머신러닝 모델을 적용할 계획인지도 간략하게 말씀해 주세요.

(세션 때 배우지 않은 내용을 시도해도 괜찮습니다.)

- ✓ **최종 발표** 때는 머신러닝 모델을 전처리한 데이터에 적용하여 어떤 결과가 나왔는지를 설명해 주세요.

- ✓ 중간 발표: 11/1 (수)

- ✓ 최종 발표: 11/8 (수)