취업은 따논 당상!!!!

본격 취업생 취직도와주기 PROJECT



소프트웨어학과

201220982 김동규

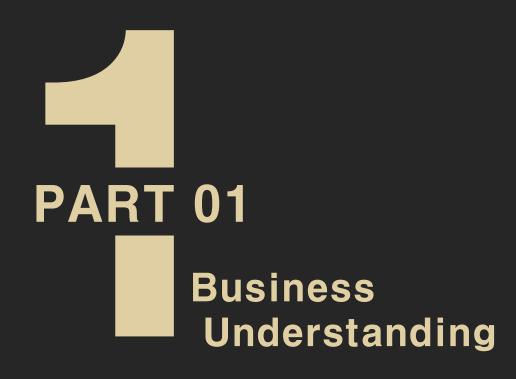
201221015 전준형

201220890 오준배

201220952 이덕명



- 01. Business Understanding
- 02. Data understanding
- 03. Data Preparation
- 04. Modeling
- 05. Evaluation



Business Understanding

Background

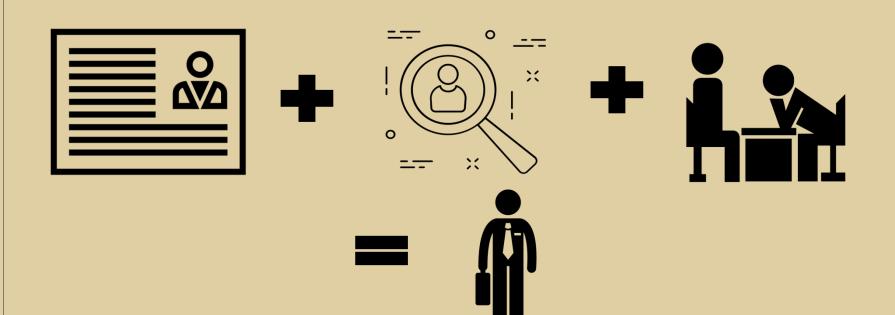
실제의 스펙 + 우리의 프로젝트

→ 극대화된 효율

Assumption

시뮬레이션으로 나타나는 합격 여부 정확도

→ 실제와 거의 비슷하다고 가정





Constraints Risks requirement

Constraints

기업의 모든 data가 아님

→ 실제와는 다른 기준의 합격점수

Risk

트레이닝 데이터를 통해 산출한 결과

물합격 판단의 데이터가 합격의 결과



Requirement

취업 정보 공유 사이트에서 얻어 온 자료

Business Understanding

Goals and Success criteria

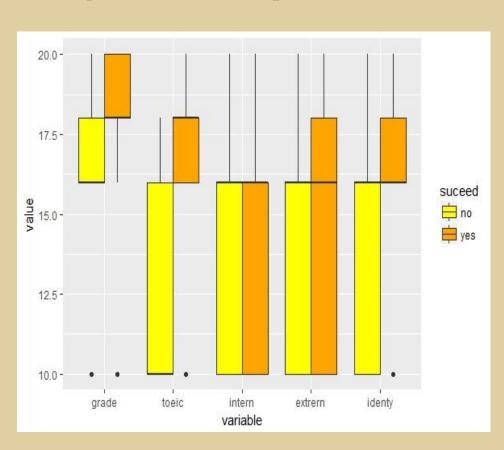
기업에서 판단하는 합격 여부의 결과 값 = > 오차를 줄이기





Data understanding

Exploratory Data Analysis

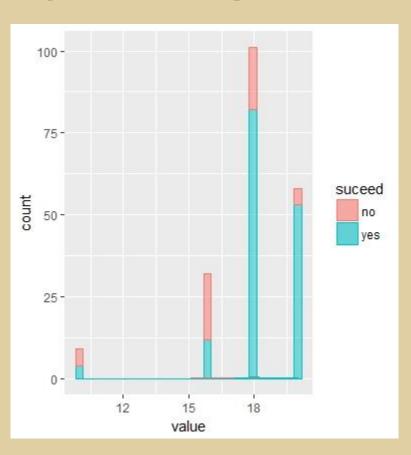


인턴을 제외하고는

합격자가 높은 범위의 평균점수의 합일 때 합격



Exploratory Data Analysis



총 6개의 attribute

합격자와 불합격자의평균점수



Data quality report

실제 합격 데이터 150개 산출

실제 불합격 데이터 50개 산출







Data cleaning / pre-processing

	Α	В	С	D	E	F	
1	grade	toeic	intern	extrern	identy	suceed	
2	3.4	805	2	0	4	yes	
3	3.7	930	1	1	4	yes	
4	3.7	О	1	1	1	yes	
5	3.3	O	1	1	3	yes	
6	3.9	O	O	1	O	yes	
7	3.8	O	O	1	O	yes	
8	4	935	1	1	1	yes	
9	4.3	O	O	O	O	yes	
10	4.2	960	1	1	O	yes	
11	4	850	O	O	O	yes	
12	3.8	O	1	O	O	yes	
13	3.9	950	O	O	5	yes	
14	3.9	950	O	O	2	yes	
15	4	920	O	1	O	yes	
16	4	875	O	1	2	yes	
17	3.8	900	O	O	1	yes	
18	4	О	1	0	1	yes	
19	3.7	955	1	O	1	yes	
20	4.1	875	О	О	1	yes	
21	3.1	700	0	1	1	no	
22	3.2	600	1	1	2	no	
23	3.8	685	1	1	1	no	
24	4.4	550	0	0		no	

각 사이트 수집한 DATA

=> 엑셀파일로 정리



Construct the new data

Welcome to EarthPro ■ X							
User							
Password							
login	register						

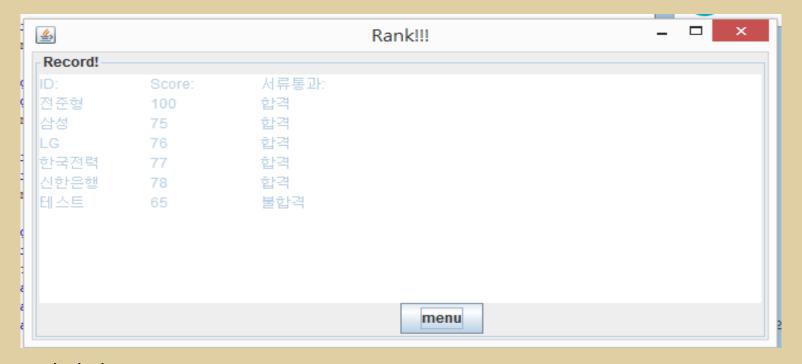
입력				
?	학점을 입력하세요.(4.5만점 기준입L 확인 취소	ICH)		

새로운 데이터 → 회원가입을 통해 받아옴

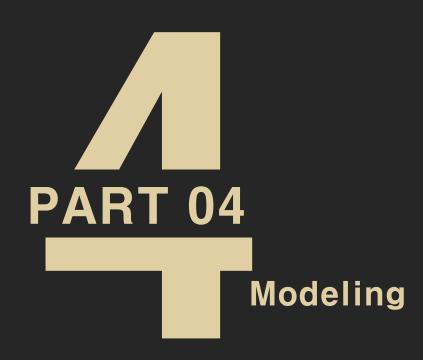
입력 된 데이터는 기존 데이터를 포함하고 있는 join.csv 파일에 새롭게 저장



Construct the new data



합격여부는 학점, 토익점수, 대외활동, 인턴경험, 자격증 입력 시 위와 같이 한번에 확인 가능



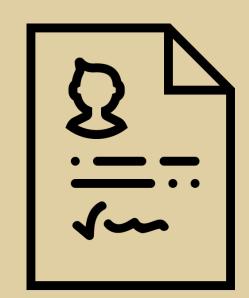


Model / Algorithm description

파라미터 설정은 서류통과 기준에 가장 많이 쓰이는 항목

각 각의 항목은 기준에 따라 10~20점으로 환산 되며, 5개의 총합의 최대가 100점

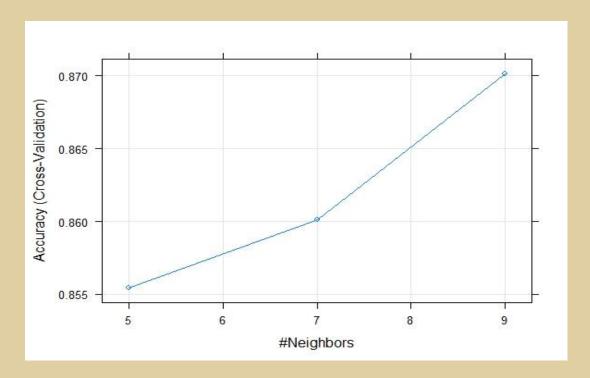
평균 점수는 계속해서 지원자에 의해 변화







Evaluation



합격인데 불합격으로 분석되거나 불합격인데 합격으로 분석되지 않게 하기 위한 e valuation