

## Machine Learning 중간고사 Take-Home Exam

### < 개요 >

중간고사 제출 시간: 5월 3일 수업시간 시작 시간. 출석체크시

제출 방법: 인쇄본 제출 1부

코드 제출 이메일 ([hoyoung23@uos.ac.kr](mailto:hoyoung23@uos.ac.kr))

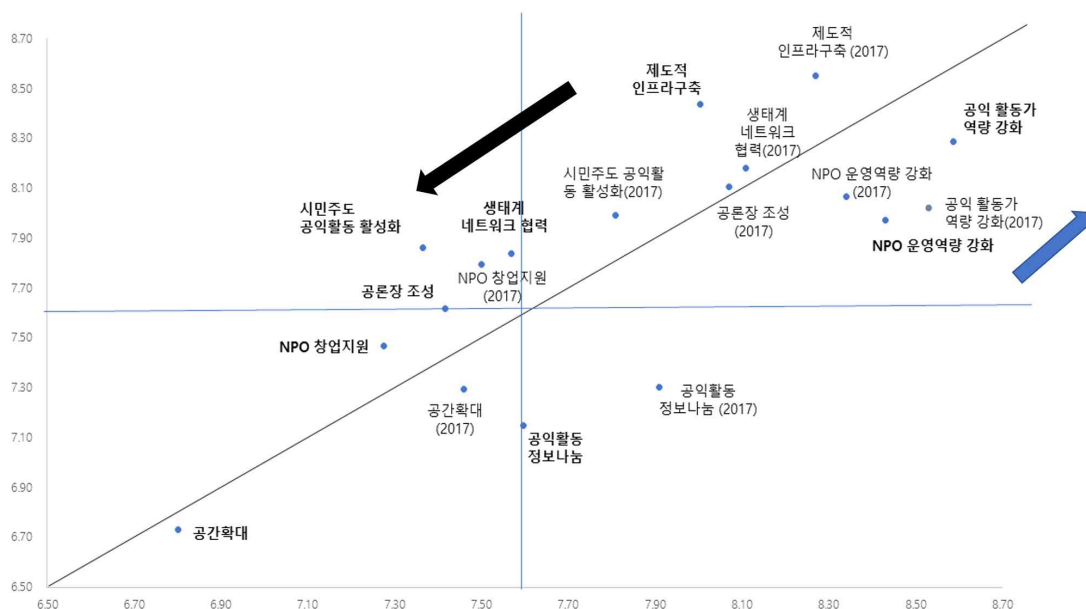
작성 분량: 10쪽을 넘지 않을 것 (표, 그래프, 서술 모두 포함)

중간고사 배점: 전체 성적의 30%

본 중간고사 Take-Home Exam은 시험이므로 해당 마감시간 이후에는 제출을 받지 않습니다. 마감시간 이후 제출할 경우, 시험을 보지 않은 것으로 간주합니다. 본 Take-Home Exam은 엄연한 시험이므로, 시험마감일과 시간에 관해 엄격히 적용합니다. 더불어서, 충분한 시간이 주어지는 만큼, 개인 사정이 있을 경우에 대비 미리미리 완료하시어 백업본을 클라우드 서비스에 보관하시길 바랍니다. 코드는 해당 사항을 실행하였을 시 실제 인쇄본 내용이 그대로 재현되어야 인정합니다. 수업에 관한 시험이므로 R코드만 인정합니다. 본 시험은 Take-Home Exam은 시험이므로 선제출에 대한 가산점은 없습니다.

### 비영리기관 지원센터 사업 프로젝트에 관한 설문 과제

다음은 비영리기관(NPO) 지원센터가 실시한 2017년과 2018년의 설문 조사의 결과로, 단기적으로 센터가 중요하게 추진해야 할 사업과 중장기적으로 센터가 추진해야 할 사업의 중요도를 표시한 그래프이다. 2018년은 굵은 글씨체로 2017년은 2017년 항목으로 표기되어 있다.



해당 그래프는 다음의 내용을 담고 있다.

- 대각선에 가까울수록 단기적 중요도와 중장기적 중요도가 일치하게 된다.
- 대각선 아래에 위치하면 단기적 중요도가 장기적 중요도 보다 더 높게 평가된다
- 대각선 위쪽으로 위치하면 중장기적 중요도가 단기적 중요도보다 더 높게 평가된다.
- X축의 값 오른쪽으로 움직여 갈수록 단기적 중요도가 높아진다.
- Y축의 값 위쪽으로 올라갈수록 중장기적 중요도가 높아진다.

해당 그래프를 보면, 2017년과 비교하여 보았을 때, 2018년의 변화가 감지되고 있다. 가장 큰 특징은 첫째, 역량 강화와 관련된 항목들의 변동 둘째, 전반적인 중장기적인 중요도가 단기적인 중요도보다 높은 사업들의 중요도 하락으로 요약된다. 주어진 데이터는 2018년도에 실시한 설문조사 데이터이다. 이러한 배경을 가지고 다음을 분석하여 보시오.

1. 위의 그래프를 그리는데 사용된 단기적 사업 중요도 문항과 중장기적 사업 중요도 문항 18개 문항에 대한 응답을 사용하여 응답자를 분류하여 보시오. (10점)
2. 1)에서 응답자들을 응답 패턴에 따라 응답자들을 분류한 프로파일링을 사업 중요도 문항이 아닌 다른 설문 문항에 대한 응답에 적용하여, 어떤 특징을 가진 응답자들이 해당 분류에 속하는지 응답자들 분류에 따른 응답자들의 성격을 특정하여 보시오. (10점)
3. 해당 결과를 가지고 인사이트를 도출하시오 (10점)

즉, 이 과제는 사업 중요도에 대한 평가를 통해 응답자들의 응답 패턴을 발견하여 응답자들을 구분하고, 구분된 응답자들이 어떠한 특징을 가지는 지 다른 문항으로 예측, 추론하는 2가지 내용으로 구성되어 있음.

본 과제 작성시 다음 사항에 주의해서 작성해야 하는 제약사항이 두 가지 있음

- 첫째. 수업시간에 배운 기계학습 방법론을 적어도 2가지는 활용해야 함
- 둘째. 분석 결과 제시와 더불어, 해당 분석 결과에 대한 명확한 설명이 가능해야 함.
- 셋째. 중간고사 기간전까지 다른 분석방법으로 한정하여 모든 사람에 대한 동일한 기준에 기반한 평가가 가능하도록 함.

본 과제 작성시 수업시간에 배우지 않은 “통계적 분석 방법”의 활용하는데에는 제약이 없음. 다만, 해당 방법론은 기계학습과 유기적인 연결관계를 가지고 있어야 하며, 통계적 분석 방법으로만 분석이 채워질 경우, 성적을 인정하지 않음.

## 변수 설명

ID: 설문 응답자 ID

outsider : 설문 응답자의 속성 - 1: 외부 이해관계자 0: 내부이해관계자

사업 중요도 10점만점 문항(단기적 중요도는 문항 앞에 a, 중장기적 중요도는 b가 붙음. 예를 들어, NPO 운영 역량 강화에 대한 단기적 중요도 평가 문항은 a1, 중장기적 평가 문항은 b1)

a1, b1: NPO 운영역량 강화

a2, b2: 공익활동가의 역량 강화

a3, b3: NPO 창업지원

a4, b4: 시민 주도공익활동 활성화

a5, b5: 공익활동 공론장 조성

a6, b6: 공익활동 상태계 네트워크 협력

a7, b7: 공익활동 제도적 인프라 구축

a8, b8: 정보 나눔

a9, b9: 공간 확대

recommend: 지원센터를 다른 이에게 추천했다 (Yes,No)

future\_recommend: 향후에 추천할 가능성 (10점 만점)

lookfor: 정보 우선성.

NPO관련 정보가 필요할 때 센터 정보를 우선적으로 찾아보는 정도(10점만점)

future\_lookfor: 향후 정보 우선성

NPO관련 정보가 필요할 때 센터 정보를 향후 우선적으로 찾아볼 가능성(10점만점)

trust: 지원센터에 대한 신뢰도 (10점만점)

firstknow: 지원센터를 알게 된 시점 (년도)

knowyear: firstknow를 해수로 계산한 변수

experience: NPO 영역 경력(년차)

times: 설문 응답한 횟수( 1은 이번 년도 최초 응답, 3은 3번째로 응답중).