

## 2022-1학기 정보통신대학원 머신러닝1 실시간 비대면 중간고사 문제

시험시간: 2022.04.26.(화) 18:35~19:35 (총1시간)

### 유의사항

- 실시간 비대면 중간고사 안내파일을 다시 한번 더 읽어보고, 2022년 04월 26일(화) 오후 7시35분(시간엄수)까지 학번\_이름을 기입한 답안지 파일을 사이버캠퍼스 과제란을 통해 제출해 주길 바랍니다.
- 답안지 파일 형식은 pdf, ppt, 한글(.hwp), MSword(.doc 또는 .docx) 파일, 또는 수기로 작성하셔서 스캔이나 사진 찍어서 파일로 보내셔도 됩니다. 패드로 작성하셔서 pdf로 변환하여 보내셔도 됩니다. 어떤 형식이든 채점을 할 수 있도록 해 주시길 바랍니다.
- 문제는 큰 문제 기준으로 총 2문제입니다.

1. 행렬  $A = \begin{bmatrix} \frac{3}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{3}{5} \end{bmatrix}$  에 대하여 물음에 답하라.

- (1) 행렬  $A$ 의 역행렬  $A^{-1}$ 을 구하시오.(10점)
- (2) 행렬  $A^{-1}$ 의 고유값과 각 고유값에 대응하는 고유벡터를 구하시오.(15점)
- (3) 행렬  $A^{-1}$ 을 스펙트럼 분해하여 행렬의 합으로 표현하시오.(15점)

2. 다음의 설명들에 대해서 참(True) 또는 거짓(False)으로 답하시오. 거짓(False)일 경우에는 그 이유를 설명하시오. (여기에서의 참과 거짓의 판단은 본 강의의 강의자료 또는 강의동영상의 내용을 기준으로 판단하는 것을 말합니다). (문제당 각 10점)

- (1) 데이터 유형에서 성별, 혈액형은 명목형(nominal) 자료이며 연속형(continuous) 자료에 속한다.
- (2) 데이터마이닝 기법은 목표변수(Y값, 반응변수, 종속변수)의 존재여부에 따라 regression (목표변수 존재함)과 classification(목표변수 존재하지 않음)으로 나눈다.
- (3) 일반화 능력이란 훈련집합(training set)에 좋은 성능을 보이는 것을 말한다.
- (4) 바이어스와 분산의 관계에서 기계학습의 목표는 바이어스는 낮추고 분산을 크게 하는 것이다.

- (5) 어떤 모델이 좋은 모델인지 모를 경우 모델 집합의 여러 모델을 독립적으로 학습시킨 후 가장 좋은 모델을 선택하는데, 이 때 필요한 데이터를 검증집합(validation set)이라고 한다.
- (6) 모델의 일반화 능력을 높이기 위해 사용하는 가중치 감쇠 (weight decay) 기법에서 가중치가 작을수록 더 매끄러운(smooth) 곡선을 가진다.

(문제 끝.)