

19-1 기말고사 데이터 사이언스 개론(003) 이현자 교수님

1. 보편적인 분류 기법에서 일어나는 2가지 프로세스를 서술하시오.
2. Parametric 함수를 일반화할 때 사용하는 함수는?
(답: logistic regression)
3. Support Vector에 대해 서술하시오.
4. (표를 제공) tree의 루트 노드로 들어갈 데이터의 속성을 구하시오.
이때, Information Gain을 구하는 식을 이용해 데이터 속성을 구하는 과정을 보이시오.
5. K-mean Algorithm에 대해 서술하시오.
6. (ROC 그래프를 제공)이 그래프에서 FP, TP, recall, accuracy, precision를 구하시오.
7. (이익곡선과 예산, 전체 인원, 홍보 비용을 제공) 이 중 분류기의 성능이 제일 좋은 것을 고른 후, 몇 명에게 홍보하는 것이 가장 효율이 좋은 지 구하시오.
8. holdup을 이용하는 이유
9. flexity가 높은 모델의 장단점을 작성하시오.