

기계학습 기말고사 정리 2 (20241519 조예성)

질문에 따라 데이터를 구분하는 알고리즘:

- 노드 구분 3가지: , ,
- 최대 깊이 제한:
- 각 분할에 사용할 속성 개수:
- 분할 위한 최소 샘플 수 조건:
- 리프 노드가 가져야할 최소 샘플 수:
- 가중치가 부여된 전체 샘플에서의 비율:
- 리프 노드의 최대 수:
- 분할로 얻어질 최소한의 불순도 감소량:
- 객체 명:

트리 분할의 기준:

- 두 종류: ,

간접 투표 위하여 변경: `voting_clf.` = " "

훈련세트 샘플링 방식 2가지: ,

랜덤포레스트에서 각 특성이 얼마나 중요한지 측정한 점수:

여러 모델 사용 학습:

- 합친 결과에 한 번 더 학습:

■ ,

주성분 분석 활용하여 차원을 2차원으로 줄이는 코드: (=2)

PCA 2 종류: ,

샘플 간 거리 보존하며 차원축소: ;

각 샘플을 가장 가까운 이웃과 연결:

비슷한 샘플은 가까이, 비슷하지 않은 샘플은 멀리:

훈련 과정에서 클래스 사이를 가장 잘 구분하는 축 학습:

k-평균에서 중앙 기준:

최적 클러스터 개수 찾는 그래프 2종류: ,

- 엘보:
- 높을수록 좋은거:

능동학습 서술: