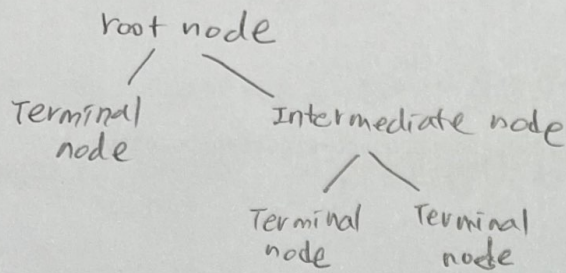


- 남수연 -

< Decision Tree >

Def 의사 결정 규칙을 나무구조로 나타냄. (분류, 예측)



prop 1) 한쪽에 몰려 있을수록 좋은 decision.

2) 불순도를 기준 잡는다!

⊗ 노드 기준 : ID3 & CART

↳ entropy & Gini index.

① entropy.

entropy 감소 (= 불순도 감소 = 순도 증가)가 목표.

↓ 이용한 알고리즘

ID3 : Information Gain이 크게 나쁜 변수 A를 기준으로 선택.
상위노드 entropy - 하위노드 entropy

② Gini index.

불순도 정도를 정량화 해서 표현한 값.

Gini index 감소 = 불순도 감소 = 순도 증가 → 우리의 목표.

↓ 이용한 알고리즘

CART : 데이터를 split 했을 때의 불순한 정도. (Binary split을 전제로 분석)

3) 연속형 변수 의사결정 나무.

* Label의 class가 바뀌는 수를 기준으로 한다.

↓ 문제점 해결 방안.

앙상블