

Decision Tree

2022년 6월 3일 금요일 오후 9:32

1. 개념

- 기준을 만들어 샘플을 분류해서 집단의 성질로 추정
- 장점 : 해석력이 높음
- 단점 : 변동성이 높음(샘플의 조그만 변화에 기준이 잘 바뀜)

2. Entropy

- impurance
- entropy가 1이라는 의미는 데이터가 잘 섞여 있다는 의미

3. information Gain

- 모델 이전 엔트로피 - 모델 이후엔트로피
- 분류가 되면서 엔트로피가 낮아짐