

❖ 분류 문제의 사용

- 분류 문제 풀기 위해 Tree 모양의 Map을 만드는 것이 효과적임
- 나무 모양을 만들어내는 알고리즘을 의사결정트리라고 함

❖ 엔트로피의 이해

- 의사결정트리를 만드는 과정
 - 1) 엔트로피가 가장 많이 줄어드는 방법으로 데이터를 나눔
 - 2) 다른 속성 기준으로 다시 데이터를 나눔
 - 3) 엔트로피가 0이 될 때까지 과정을 반복함
- 정보엔트로피를 구하는 공식

$$-\sum_{i=1}^n p(x_i) \log_2 p(x_i)$$

❖ 의사결정트리

- 의사결정트리 알고리즘의 특징
 - 엔트로피 이용
 - 지니계수 이용
 - 데이터를 나누는 방법에 대한 고민
 - 가장 영향력이 큰 속성을 찾는 방법
- 카이제곱스퀘어 검정의 용도
 - 2개 범주의 데이터를 나눌 때 상관관계가 0에 가까운지 검정하는 '독립성 검정'의 용도로 쓰임