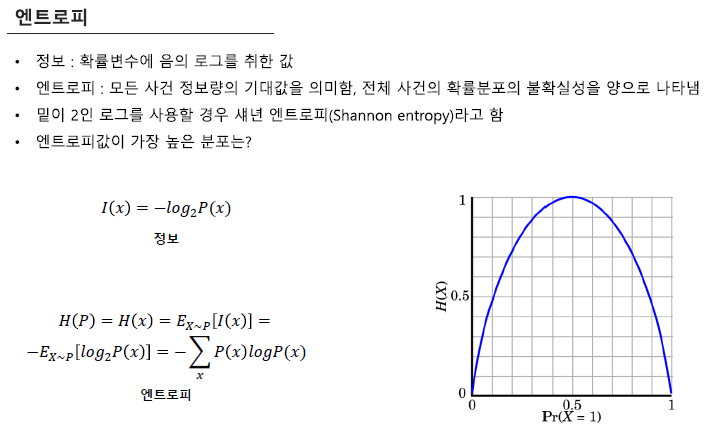
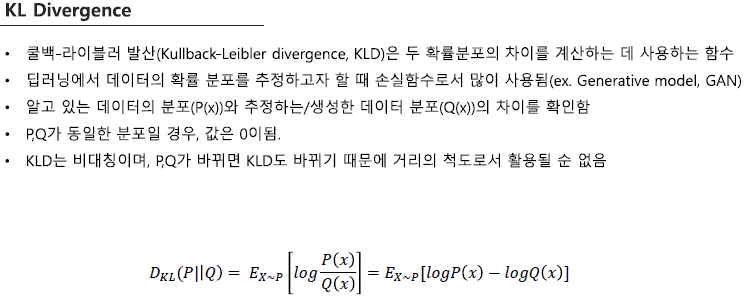
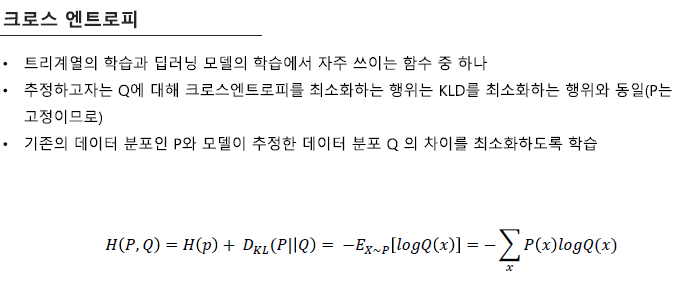


잘 일어나지 않는 사건>자주 발생하는 사건

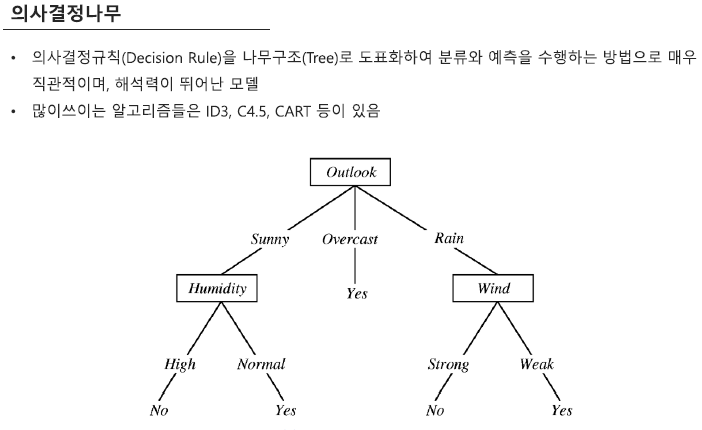






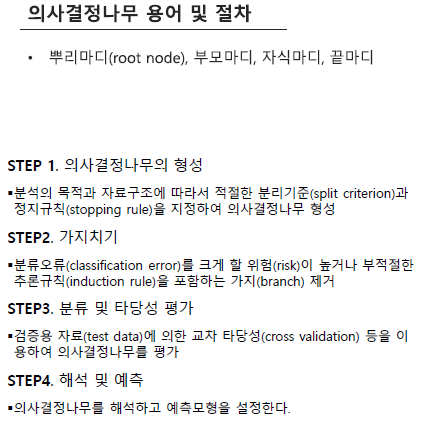
기존에 존재하는 P라는 분포를 Q가 P를 따라가도록 학습한다.

(이 공식이 중요 딥 러닝에서 자주사용\*\*\*\*\*\*)



해석하는 것이 직관적이기 때문에 Decision Tree는 많이 사용한다.(현업하는 분들을 설득하는데 굿)

퍼포먼스 측면에서는 Neural network를 사용 but, 블랙박스이기 때문에 해석할 수 없다.

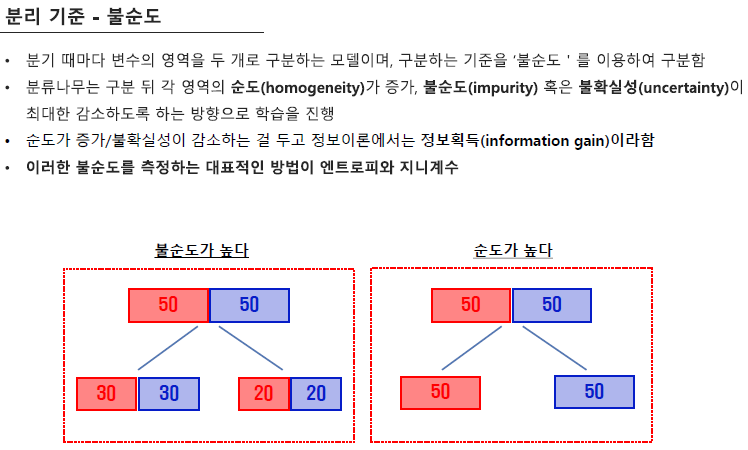


분리기준 – 나누는 기준 점

정지규칙 – 중간에 Cut하거나 마지막까지 내려가기까지 하는 규칙

(마지막까지 갈 경우 트리의 깊이가 너무 깊어짐 🡪 Overfitting 유발)

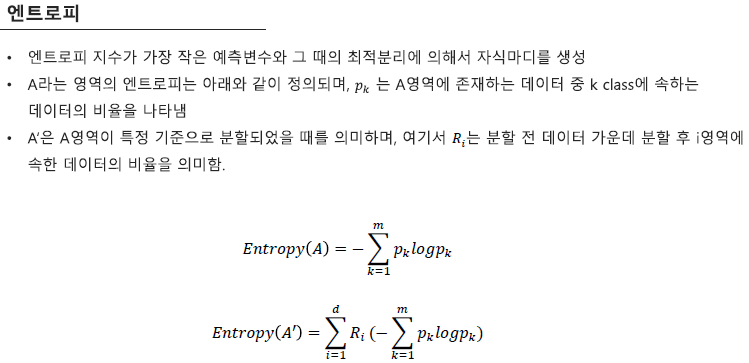
가지치기는 Overfitting을 줄임



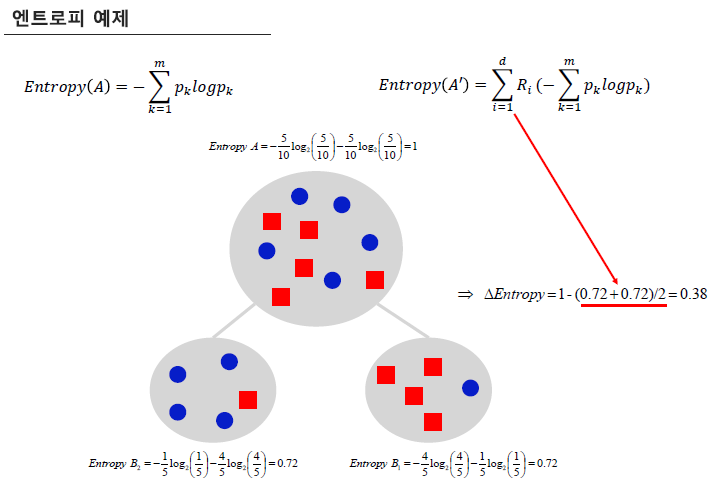
불순도 - 섞여 있는 정도

예를 들어 나무를 조건을 가지고 쪼갰을 때 분기 전보다 후에 불순도가 낮아져야 한다.

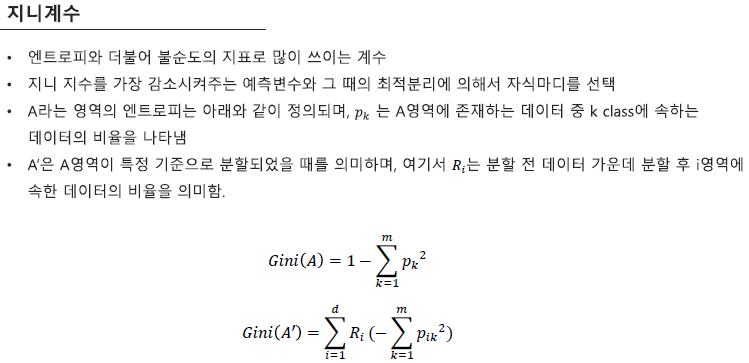
트리를 학습시킬 때, 불순도는 최대한 감소, 순도를 최대한 증가하는 방향으로 학습을 진행

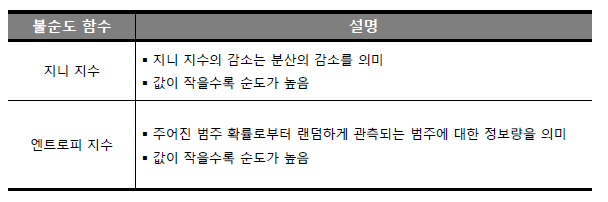


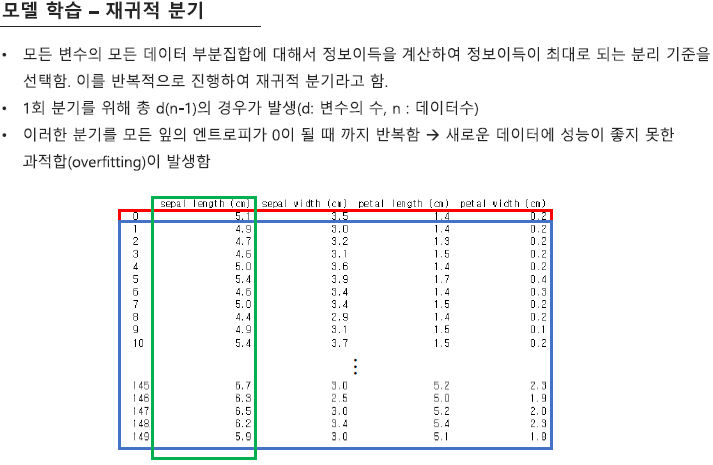
엔트로피는 모든 사건 정보량의 기대 값을 의미함, 전체 사건의 확률 분포의 불확실성을 양으로 나타냄

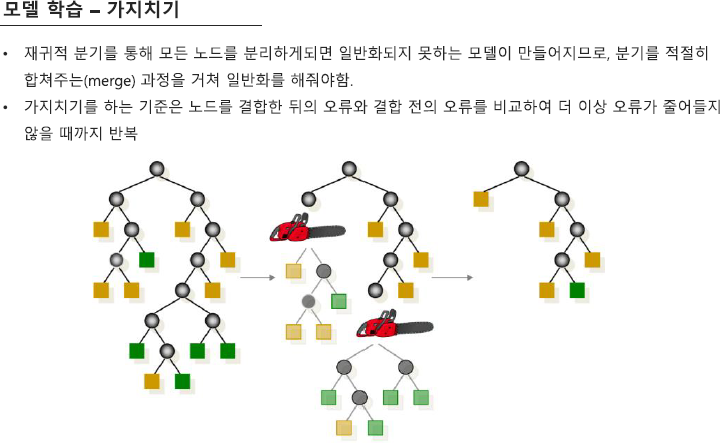


델타 엔트로피는 불순도의 감소량이 가장 큰 것을 찾는 과정에서 사용된다.





엔트로피가 0이라는 의미는 불순도가 0 🡪 단 하나의 클래스로 구성된 것 🡪 Overfitting이 발생 🡪 Overfitting을 방지하기 위해 가지치기를 행함



비용함수 – 아래의 비용함수를 최소로하는 분기를 찾도록 학습됨

