

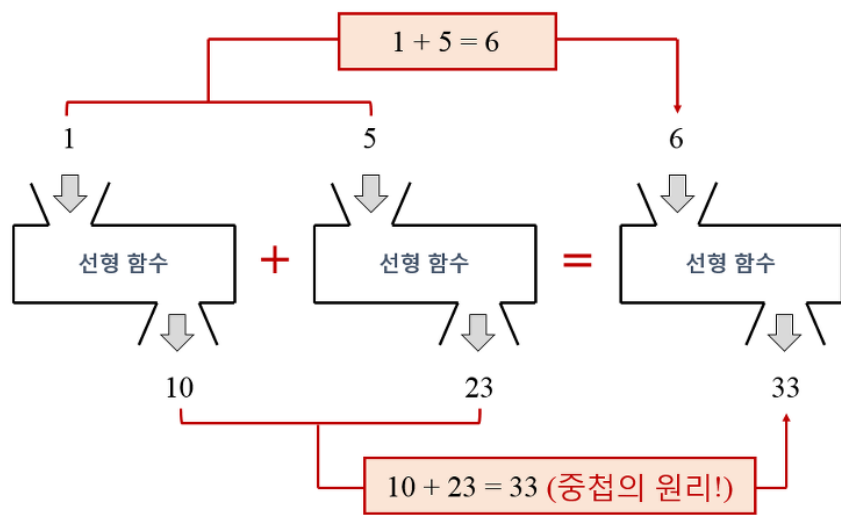
선형 회귀(다중선형회귀) 로지스틱 곡선 의사결정나무

목차

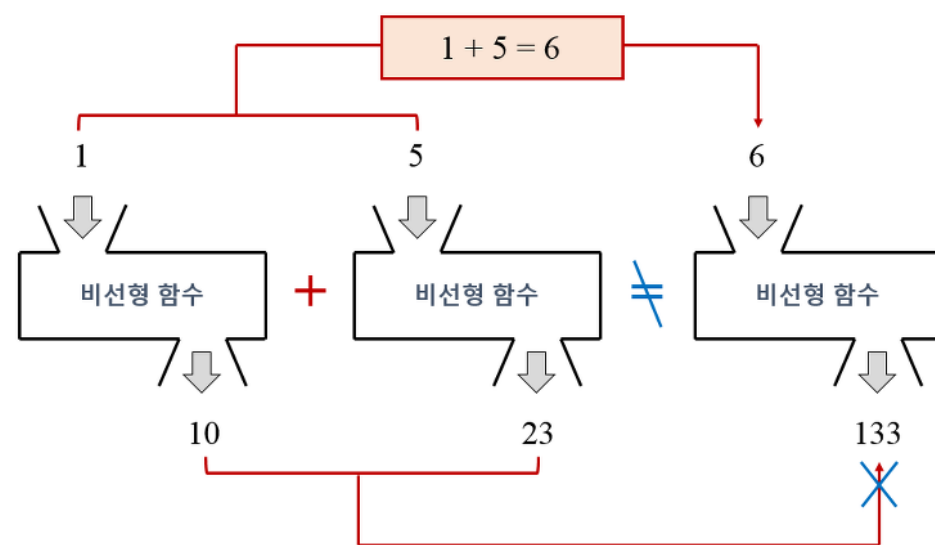
- 선행회귀(다중선행회귀)
- 로지스틱곡선
- 의사결정나무

선형회귀(다중선형회귀)

선형 vs 비선형



선형함수



비선형함수

선형회귀(다중선형회귀)

- 선형회귀?

종속 변수 하나와 한 개 이상의 독립 변수 사이의 선형 관계(직선의 형태를 띄는가)를 모델링하는 방법

여기서 가장 중요한 식

$$y = mx + b$$

x = 독립변수

m = 기울기

b = 절편

선형회귀(다중선형회귀)

- 다중선형회귀?

독립 변수가 여러 개로 이뤄진 선형회귀!

여기서 나오는 식

$$y = m_1x_1 + m_2x_2 + m_3x_3 + m_4x_4 + \cdots + b$$

로지스틱 곡선

• 로지스틱곡선

벨기에의 수학자 베르홀스트가 제안한 곡선으로 처음 고안되었을 때는 생물체의 개체군 성장 모델로써 사용되었다. 식으로는 $1/(1+e^{-x})$ 로 나타낼 수 있고, 그래프의 점근선은 $y = 0$, $y = 1$ 두개로 그래프 내 값은 0 초과 1 미만으로 나타나게 된다.

로지스틱 곡선은 두 개의 값만을 가지는 종속변수와 독립변수의 인과관계를 알아보기 위해 로지스틱 곡선을 통해 추정하게 된다. 로지스틱 곡선의 사용 예시로는 금융권의 고객 신용도 평가, 질병 예측 등, 그리고 활성화함수 중 sigmoid 곡선이 로지스틱 곡선의 형태를 띠고 있다.

의사결정나무

데이터마이닝의 예측 기법중 하나

데이터를 분석하여 이들 사이에 존재하는 패턴을 예측 가능한 규칙들의 조합으로 나타내는 기법으로, 만든 모양이 마치 나무와 같아 의사결정나무라는 이름으로 지어졌다. 맨 처음 해결할 문제에서 체크리스트들이 하나하나 뻗어가고 또 그 리스트에서 다시 체크리스트로 뻗어나가는 과정을 반복해나가는 것을 의미한다. 마치 스무고개와 비슷한 구조를 가진다.