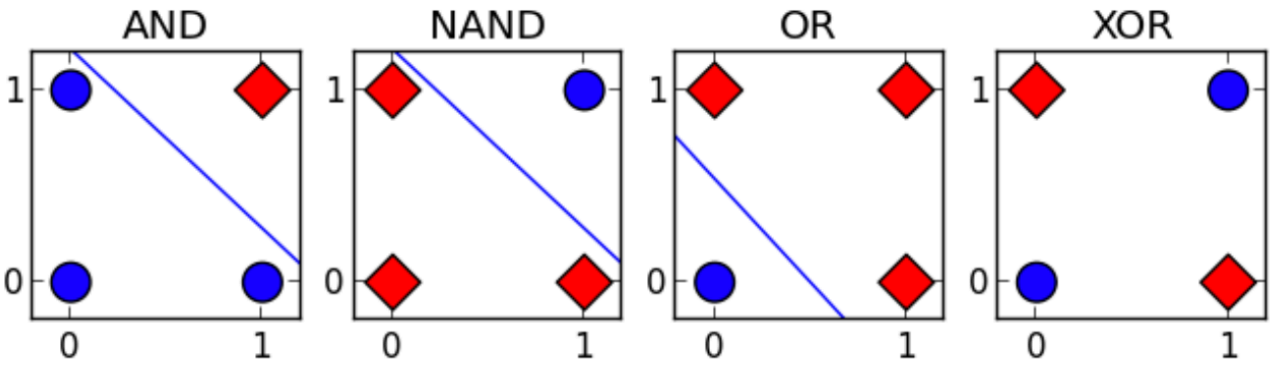
Multiclass Classifications

기존에는 개vs고양이 ,말vs사람 처럼 두 가지로 분류하는 과정을 하였습니다.

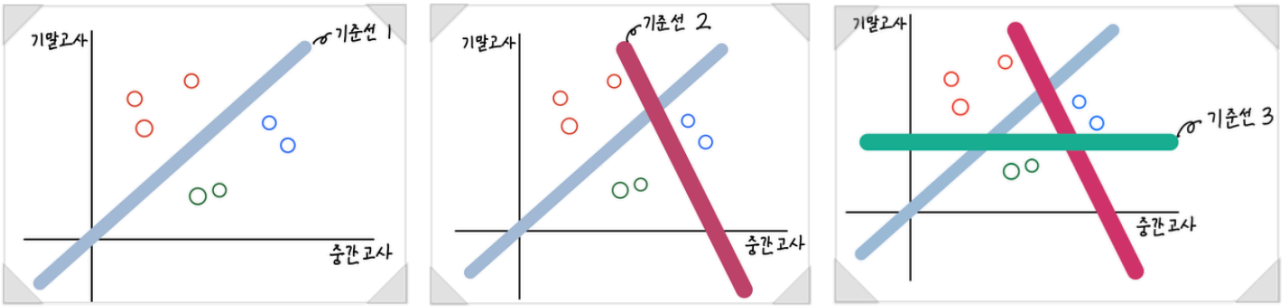
다중 분류는 말 그대로 다양하게 분류하는 것으로,

그 예시로는 성적을 A~F로 나누는 것이 있습니다.



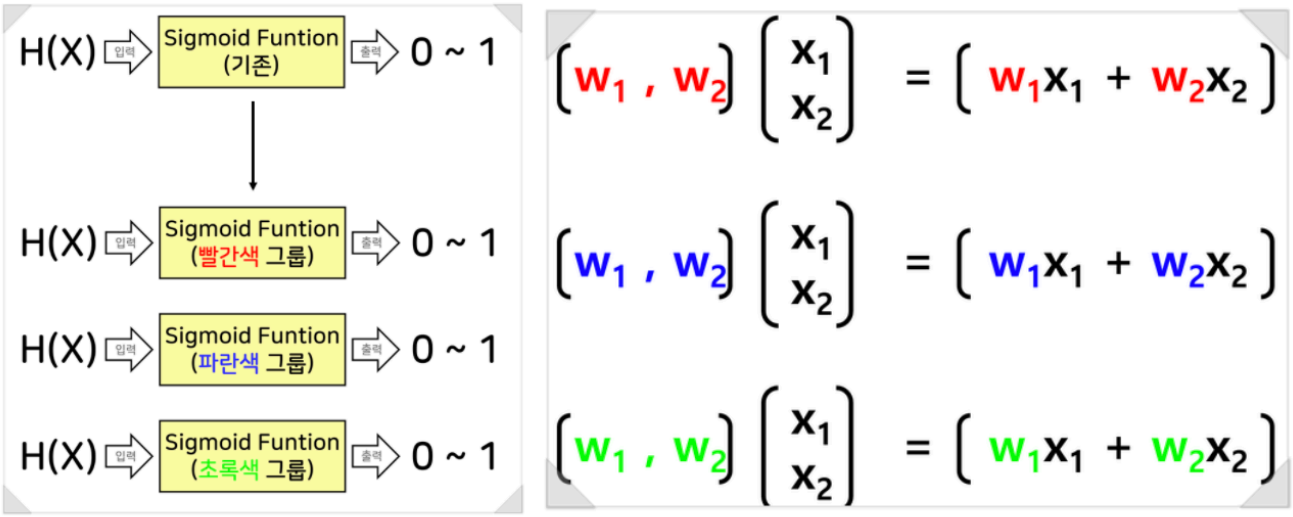
우리는 단층 퍼셉트론을 통해 AND, NAND, OR을 해결할 수 있습니다. 적절한 기준선 하나만 찾으면 되었기 때문입니다.

그러나 XOR을 푸는 것은 기준선 하나만으로는 해결할 수 없습니다. 여기서 다층 퍼셉트론을 적용할 수 있습니다. 퍼셉트론은 일종의 기준선을 의미합니다.



구분해야 할 카테고리(Category) 또는 클래스(Class)가 3개니까, 기준선도 3개를 그렸다

다중 분류 문제를 해결하기 위해 기준선을 세 개 그린 것 입니다. 각각의 기준선이 하는 역할은 명확합니다. 빨간색 그룹에 속하는 학생과 그 외 / 파란색 그룹에 속하는 학생과 그 외. 이렇게 구분하고 있습니다. 최적의 기준선을 찾을 수만 있다면, 어떤 입력이 들어오더라도 세 가지 카테고리로 분류할 수 있습니다.



서로 다른 임무를 가진 수식, 즉 분류기(Classifier)

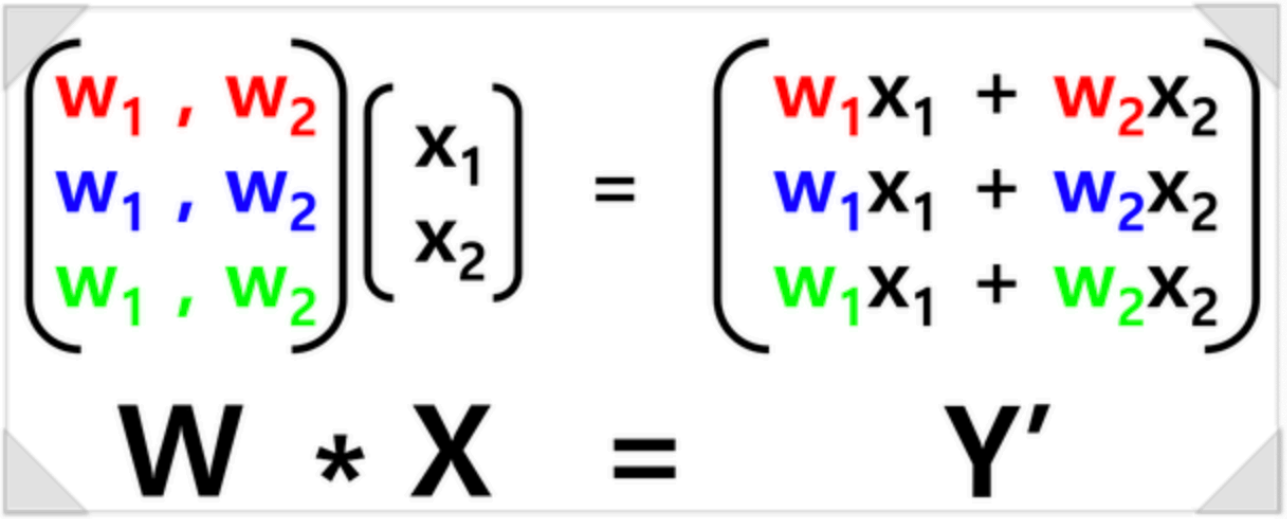
선을 여러 개 사용한다는 것은 같은 수식을 여러 번 사용한다는 뜻입니다. 물론 수식마다 맡은 임무는 조금씩 차이가 있지만 수식 자체에는 변화가 없으므로, 논리 회귀 문제를 해결하기 위해서 사용했던 수식을 그대로 사용하겠습니다.

먼저 가설 형태부터 살펴봅시다.

기본 가설 H(x)의 모양은 W*X + b입니다.

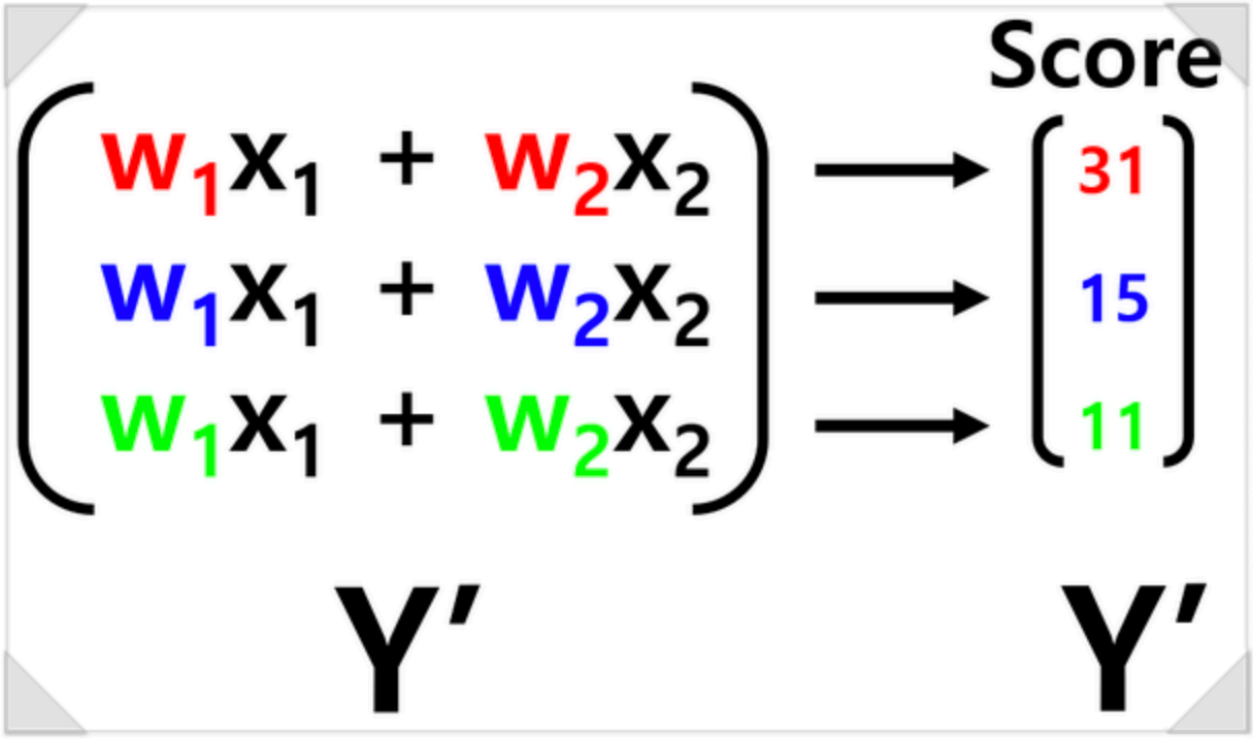
이를 간단하게 생각하기 위해 bias는 없는 것으로 가정했고, 그 결과 가설을 W*X 형태로 나타낼 수 있었습니다. 위에서 입력값을 중간고사 점수와 기말고사 점수, 이렇게. 두 가지로 취했기 때문에 가중치 W도 두 개를 취합니다.

위의 식을 정리하면 다음과 같습니다.



수식이 세 개였으므로 출력값도 세 개가 나올 것입니다.

즉, 위 그림에서는 (3,1)행렬로 결과가 나옵니다.

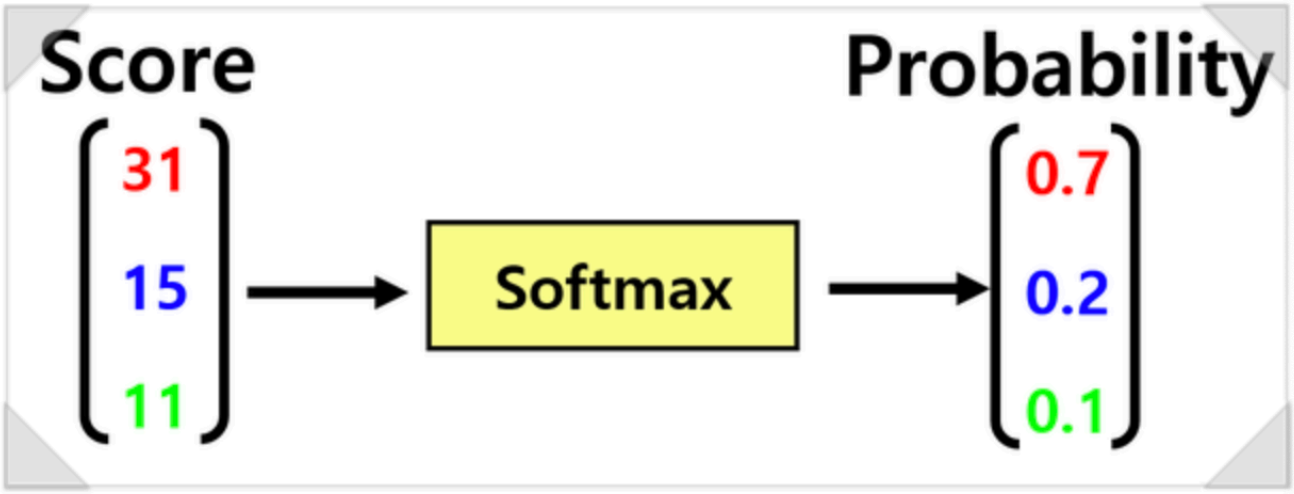


학생 성적을 입력하니까 값이 나왔다.

수식에 학생의 성적을 입력해보았습니다. 이 결과값이 의미하는 바가 무엇일까요?

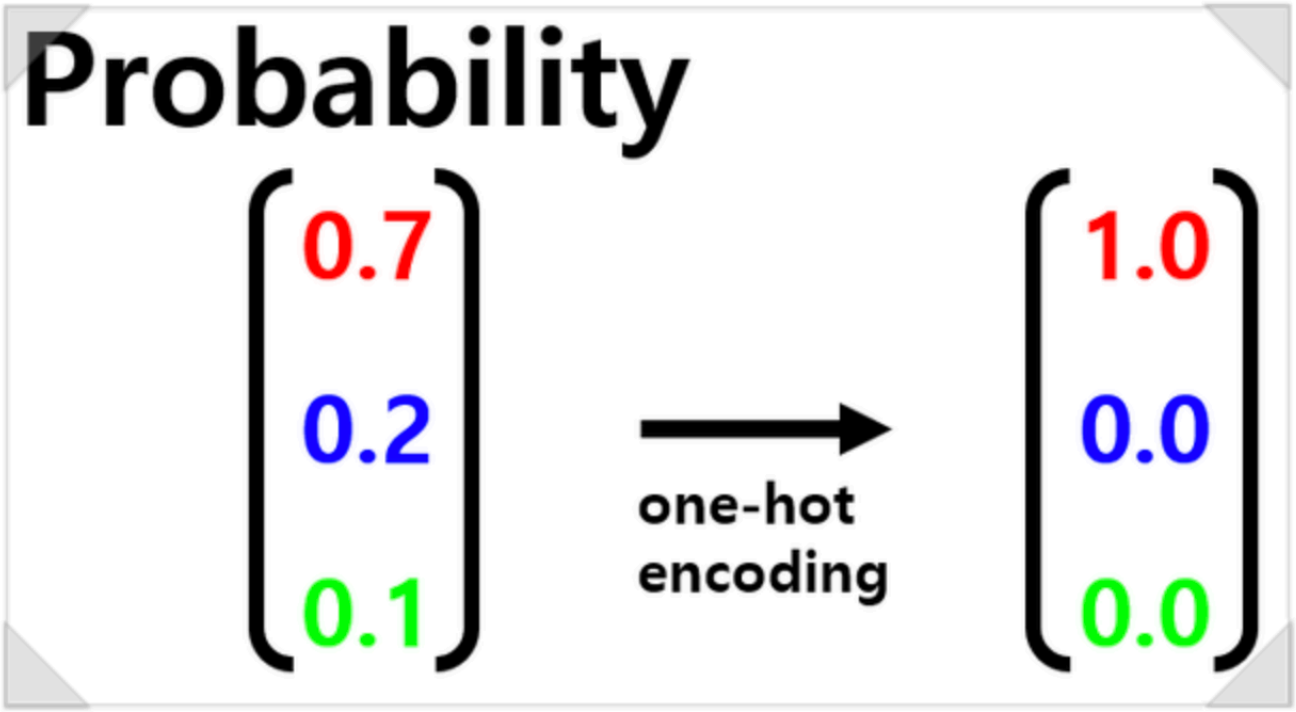
빨간색 수식에서 나온 값이 가장 크기 때문에 이 학생은 빨간색 그룹에 속할 확률이 높다는 것을 의미합니다. 이 값이 유용한 의미를 가질 수 있도록 후처리를 해봅시다.

- 활성화함수(Activation function) - Softmax



활성화 함수, 소프트맥스(Softmax)

단순히 0과 1로 분류할 때는 Sigmoid 함수를 사용해도 되지만 소프트맥스를 사용하는 것이 효과적입니다. 이 함수의 목적은 입력값을 모두 확률로 바꾸어주는 것입니다. 따라서 소프트맥스는 출력값을 모두 더했을 때, 총합이 1이 되도록 개개별 입력을 0~1 범위 내로 압축시킵니다.



다만 왼쪽 행렬은 확률만을 나타내고 있고, 모델은 아직 결과를 예측하지 않았습니다. 따라서 확률값을 한 번 더 가공해야합니다. 대표적인 방법에는 one-hot-encoding 기법이 있는데, 이것은 가장 확률이 높은 값을 1로, 나머지는 0으로 만드는 방법입니다. 텐서플로우에서 제공하는 argmax() 메서드를 사용하면, 가장 높은 확률 값의 인덱스를 반환합니다. 위의 예시에서는 0.7이 가장 높은 확률이므로 인덱스0번을 리턴받습니다.