**Linear Regression Model과 Linear Classification Model의**

**차이점와 예시**

1. **Linear Regression Model**

선형 회귀는 한 개 이상의 독립 변수 x와 y의 선형 관계를 나타낸다. 독립 변수 x가 1개라면 이는 단순 선형 회귀라고 한다. 단순 선형 회귀는 다음과 같은 수식으로 표현할 수 있다.

위 식에서 독립 변수 x와 곱해지는 값 w를 머신러닝에서는 가중치(weight)라고 하며, 별도로 더해지는 값인 b에 대해서는 편향(bias)이라고 한다.

Y는 한 개이지만 x는 여러 개인 선형 회귀를 다중 선형 회귀라고 한다. 이에 대한 수식은 다음과 같다.

선형 회귀는 주어진 데이터로부터 y와 x의 관계를 가장 잘 나타내는 직선을 그리는 일이다. 어떤 직선이 그려질 지를 결정하는 값은 w와 b이므로 결국 선형 회귀에서 할 일은 적절한 w와 b를 찾아내는 것이다.

1. **Linear Classification Model**

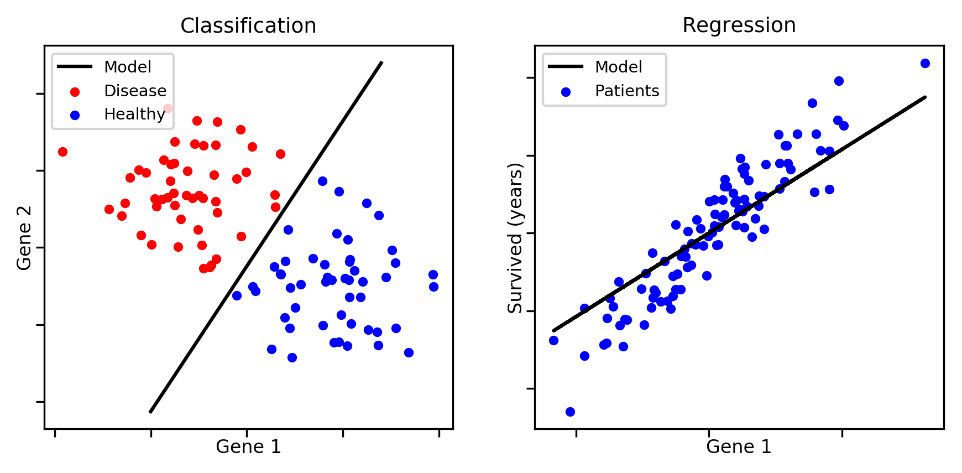
분류는 이산적인 값을 예측하는 프로세싱이다. 예를 들어 조류의 사진을 보고 해당 조류가 참새인지 까치인지 구분하는 문제가 있다. 이 예에서 참새나 까치 같은 것을 클래스라고 부른다.

분류 중에서도 선형 분류는 전체적인 데이터를 분류하는 경계선이 직선의 형태임을 가정하는 것이다. 쉽게 말해 선을 이용하여 집단을 분류하는 모델이 선형 분류라고 할 수 있다. 선형 분류에서는 x라는 하나의 데이터는 어떤 하나의 클래스에만 속하게 되며, 하나의 클래스는 다른 클래스를 부분적으로 포함하지 않는다.

1. **Linear Regression Model과 Linear Classification Model의 차이점**

선형 회귀와 선형 분류는 모두 예측을 위한 알고리즘이다. 그러나 선형 회귀는 연속적인 값을 예측하는 데에 사용되는 반면 선형 분류는 이산적인 값을 예측하는 데에 사용된다.

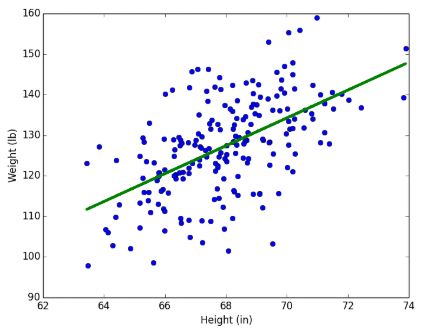
위 식은 선형 회귀와 선형 분류 모두의 식이라고 할 수 있다. 이처럼 둘의 식 자체는 동일하다. 그러나 예측하는 값이 다르게 때문에 선형 회귀와 선형 분류는 다른 목적을 지닌다. 선형 회귀는 데이터의 전체적인 추세를 아우를 수 있는 선을 찾는 것이 목적이다. 반면 선형 분류는 데이터를 구분할 수 있는 선을 찾는 것을 목적으로 한다. 이처럼 선형 회귀와 선형 분류는 차이를 보이는 데, 이는 <그림 1>를 통해 알 수 있다.



<그림 1> 선형 분류와 선형 회귀 그래프

1. **Linear Regression Model에 적용할 수 있는 예시**
   1. **키와 몸무게**

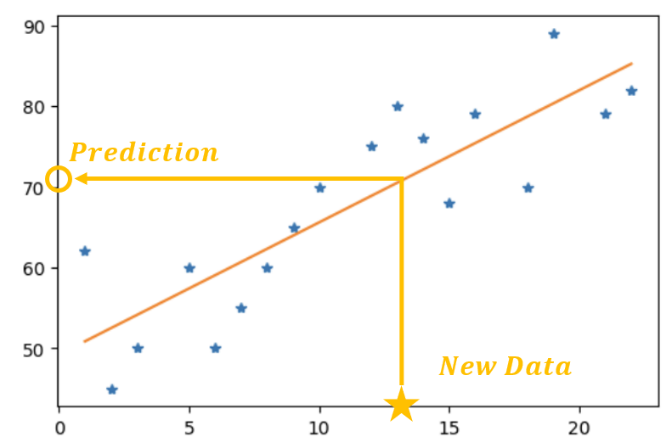
선형 회귀 모델을 통해 키 x와 몸무게 y의 상관관계를 알아볼 수 있다. 다음 그래프를 보면 키 x를 통해 몸무게 y를 예측하기 위한 선형 그래프를 확인할 수 있다.



<그림 2> 키와 몸무게의 선형 회귀 모델

* 1. **광고비와 판매 수익**

<그림 3>는 광고비 x와 판매 수익 y에 대한 데이터로 선형 회귀 모델을 만든 것이다. 이를 통해 추후 광고비 x를 투자하였을 떼 판매 수익 y를 예측할 수 있다.

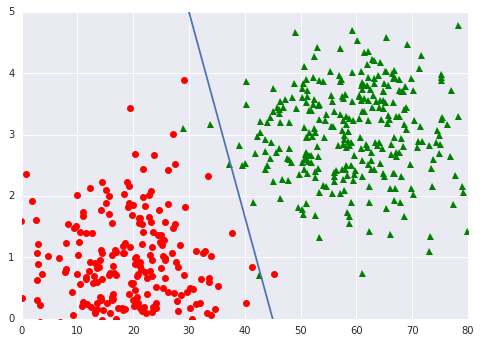


<그림 3> 광고비와 판매 수익의 선형 회귀 모델

* 1. **자녀의 나이와 자녀 조부모의 나이**

자녀의 나이 x와 자녀 조부모의 나이 y를 활용하여 선형 회귀 모델을 만들 수 있다.

1. **Linear Classification Model에 적용할 수 있는 예시**



<그림 4> 선형 분류 모델

* 1. **악성 종양 분류**

선형 분류에서는 <그림 4>와 같이 두 종류의 데이터를 나누는 직선을 그리게 된다. 악성 종양인지 아닌지를 분류할 경우 이를 사용할 수 있다.

* 1. **온라인 사기 거래 분류**

온라인에서의 거래가 사기인지 아닌지 분류하는 경우 선형 분류의 활용이 가능하다.

* 1. **스팸 분류**

이메일이 스팸인지 아닌지 분류하는 경우 선형 분류를 적용할 수 있다.