

# Machine Learning

## Chapter 2 지도 학습(Supervised Learning)



START



**Smart Media**  
스마트미디어인재개발원

- 선형 분류모델을 이해하고 사용 할 수 있다.
- 예측의 불확실성 개념을 이해 할 수 있다.



# Linear Model

## (Classification)



**Smart Media**  
스마트미디어인재개발원

## 분류용 선형 모델

$$\hat{y} = w_1x_1 + w_2x_2 + w_3x_3 + \cdots + w_px_p + b > 0$$

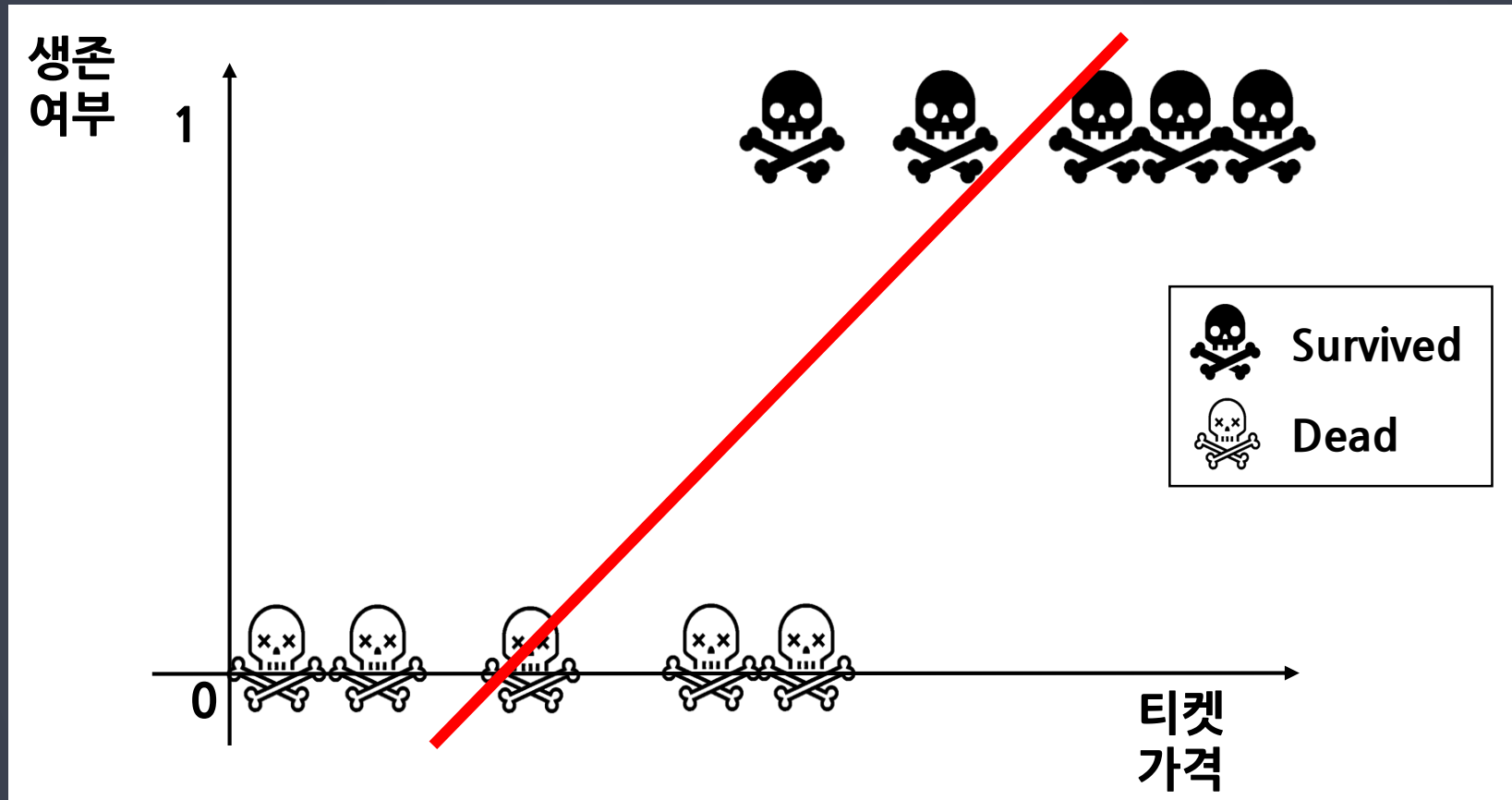
- 특성들의 가중치 합이 0보다 크면 class를 +1 (양성클래스)로 0보다 작다면 클래스를 -1 (음성클래스)로 분류한다.
- 분류용 선형모델은 결정 경계가 입력의 선형함수

## 분류용 선형 모델

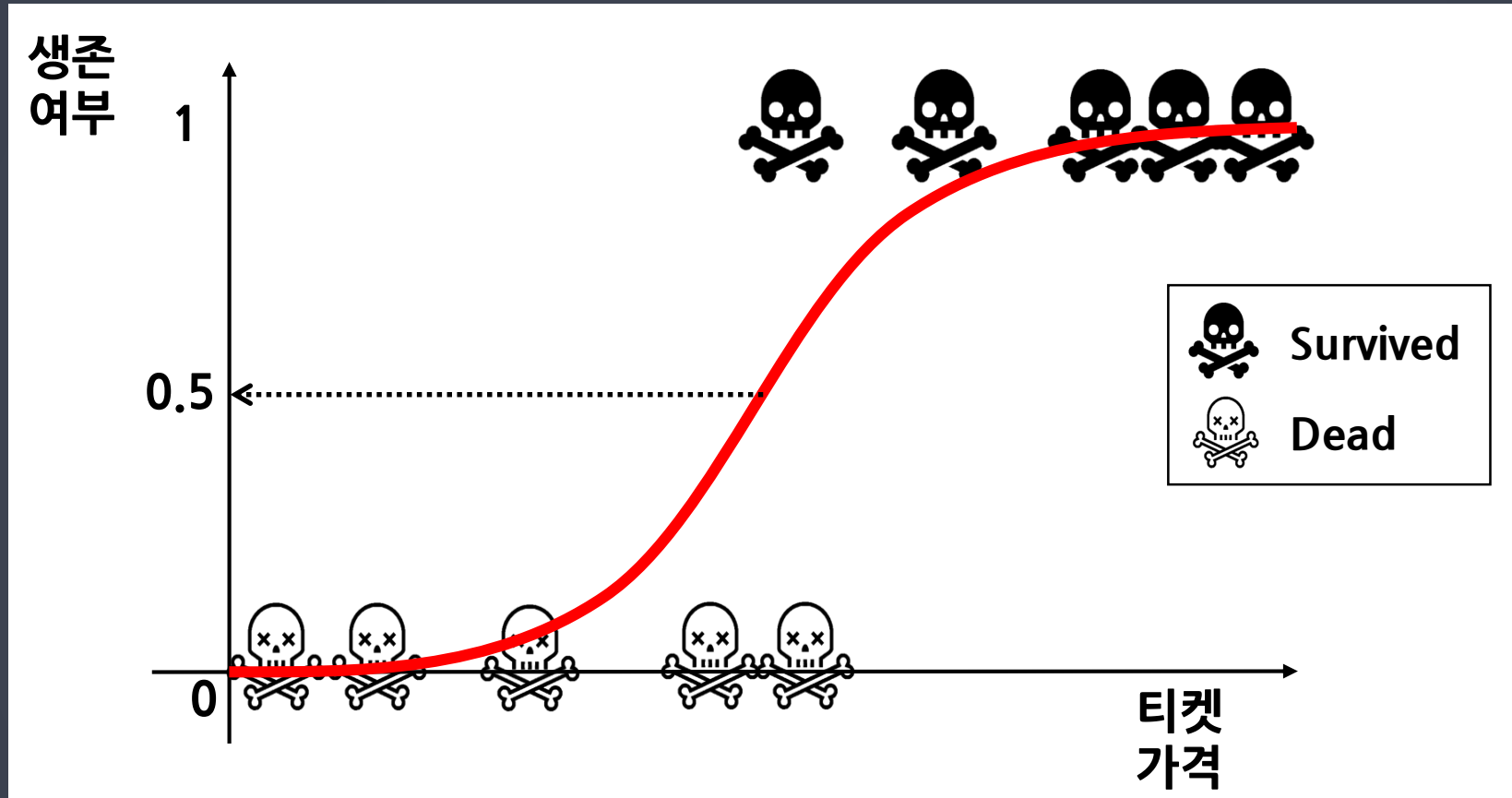
- Logistic Regression  
(Regression 단어가 붙지만 분류용 모델)
- Linear Support Vector Machine



# Linear Model – Classification

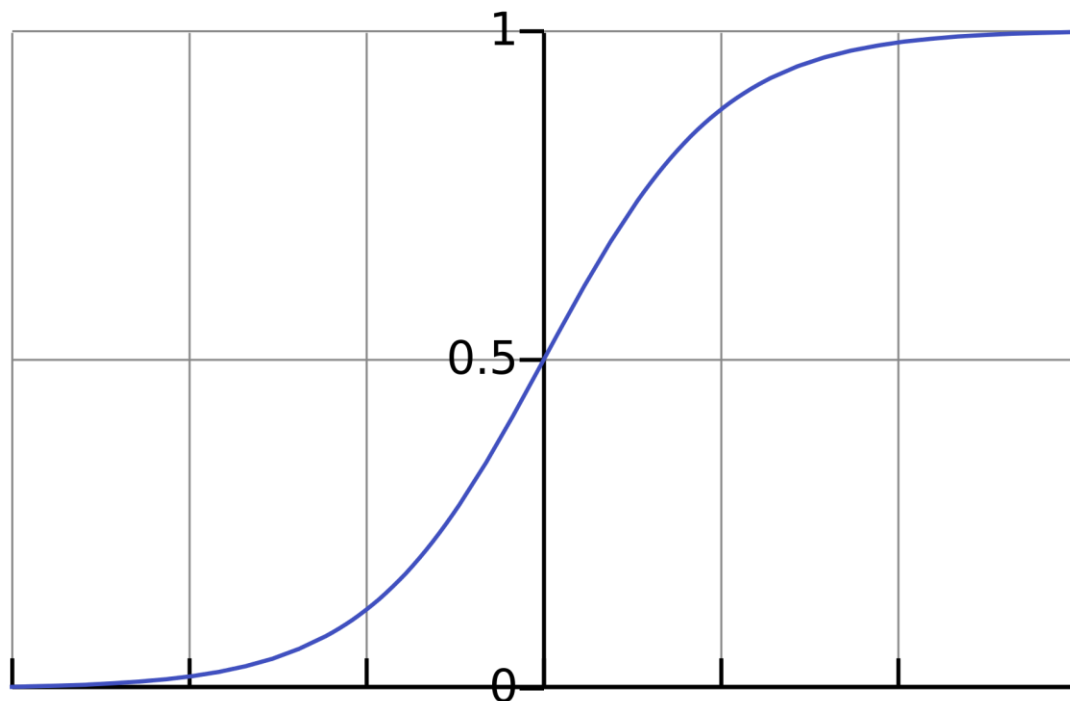


# Linear Model – Classification



## Logistic Regression

- 선형 함수의 결과 값을 Sigmoid Function(Logistic Function)을 이용해 0과 1로 변환한다.

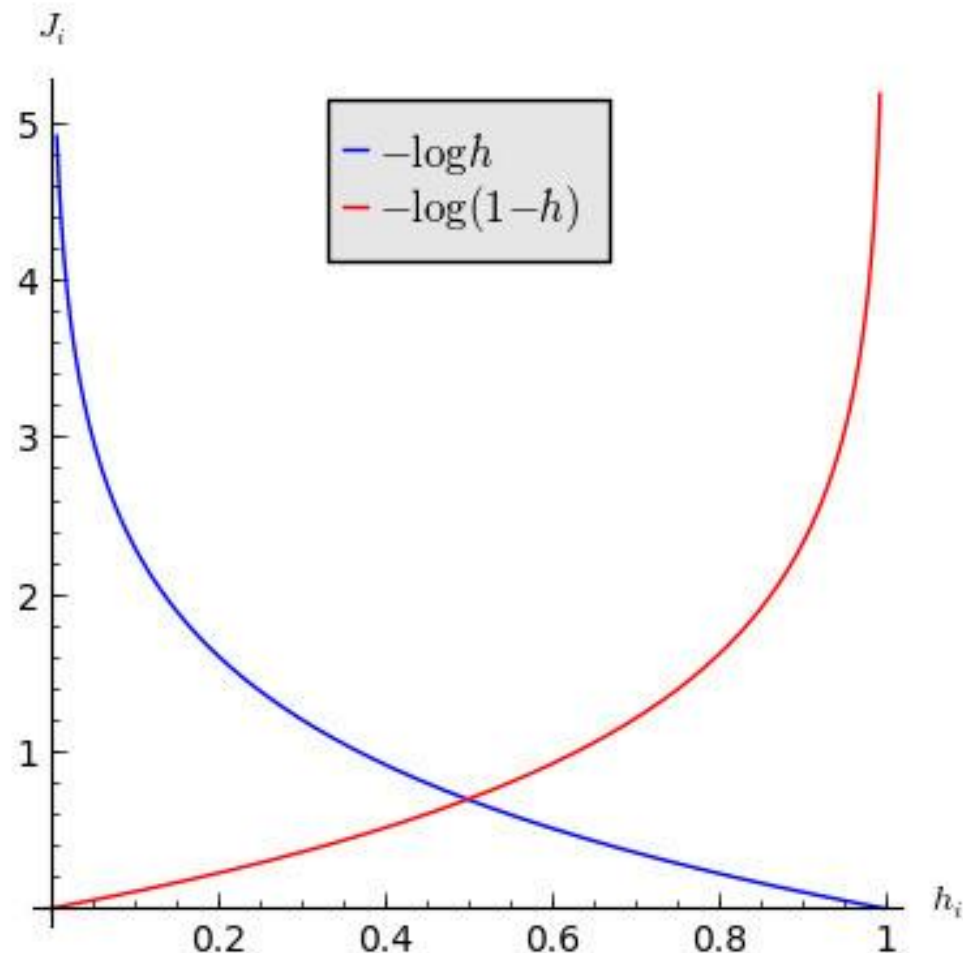


$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}}$$



# Linear Model – Classification

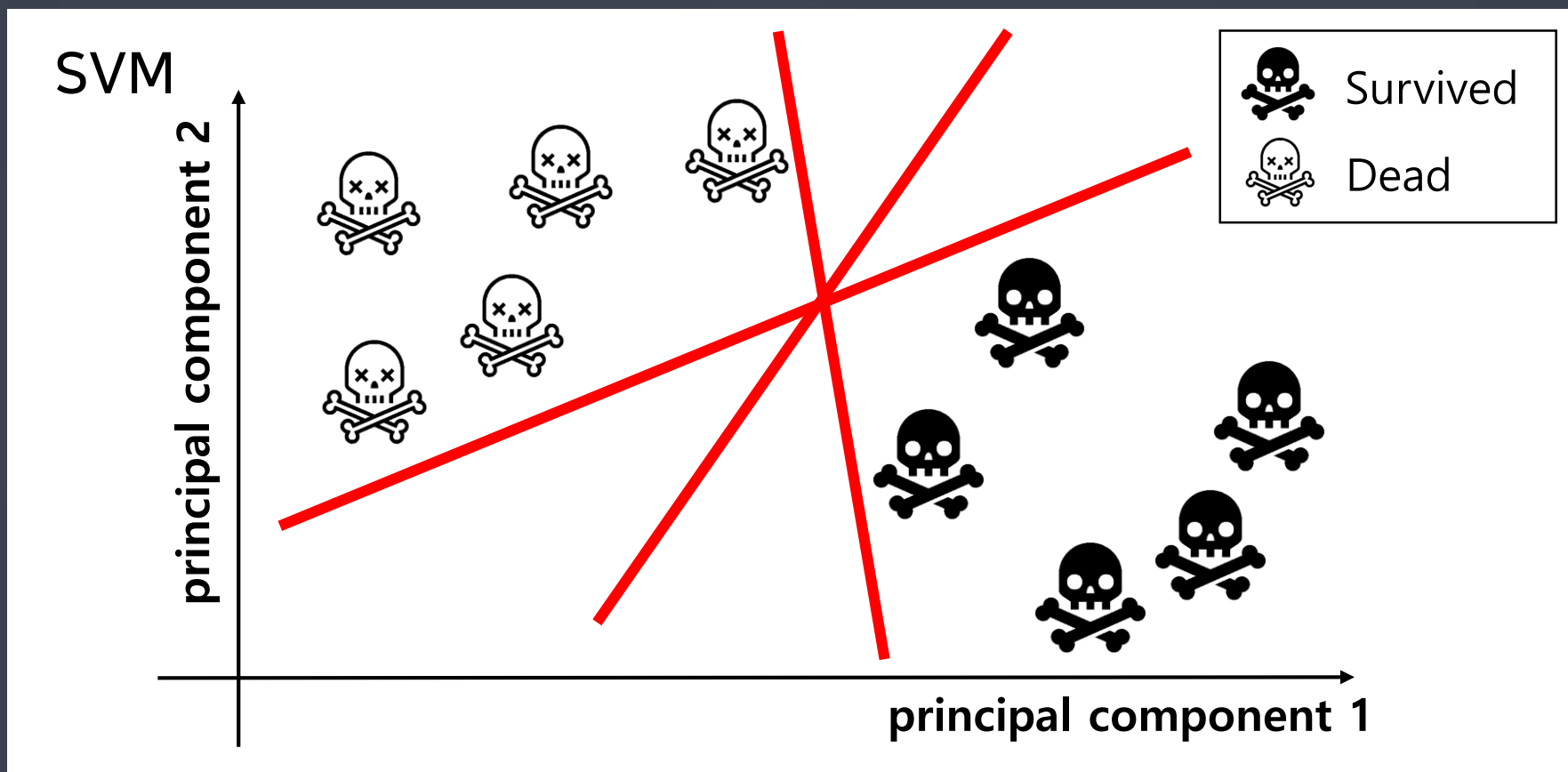
$$\begin{aligned}\text{Cost}(h_{\theta}(x), y) &= -\log(h_{\theta}(x)) && \text{if } y = 1 \\ \text{Cost}(h_{\theta}(x), y) &= -\log(1 - h_{\theta}(x)) && \text{if } y = 0\end{aligned}$$



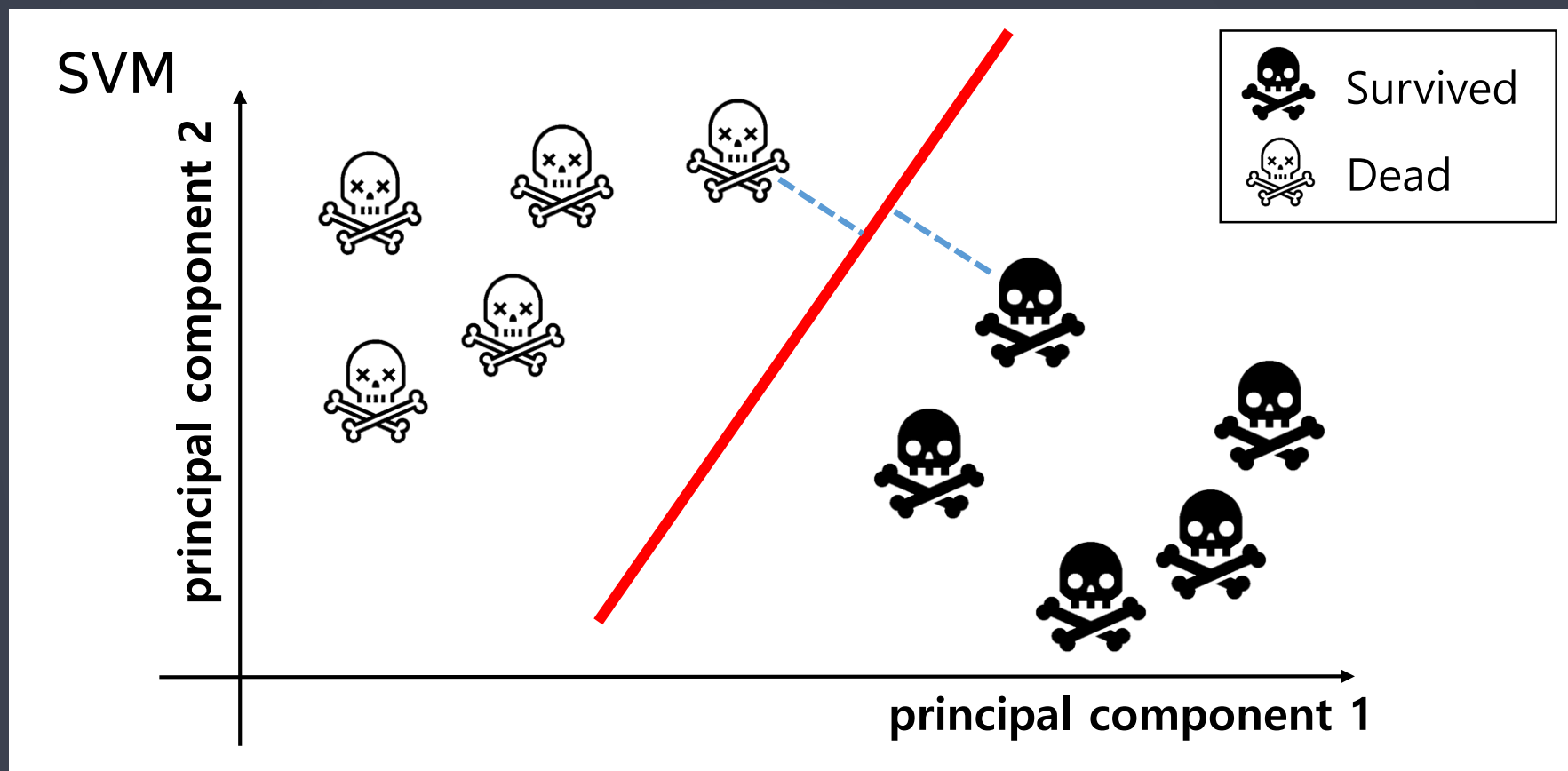
## Binary Cross-Entropy Loss

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m [y^{(i)} \log(h_{\theta}(x^{(i)})) + (1 - y^{(i)}) \log(1 - h_{\theta}(x^{(i)}))]$$

## Support Vector Machine (SVM)



## Support Vector Machine (SVM)



## 장단점 및 주요 매개변수(Hyperparameter)

- 회귀 선형 모델 :  $\alpha$   
(값이 클수록 규제가 강해진다.)
- 선형 분류 모델 :  $C$   
(값이 클수록 규제가 약해진다.)
- 기본적으로 L2규제를 사용, 하지만 중요한 특성이 몇 개 없다면 L1규제를 사용해도 무방  
(주요 특성을 알고 싶을 때 L1 규제를 사용하기도 한다.)

## 장단점 및 주요 매개변수(Hyperparameter)

- 선형 모델은 학습 속도가 빠르고 예측도 빠르다.
- 매우 큰 데이터 세트와 희소 (sparse)한 데이터 세트에서도 잘 동작한다.
- 특성이 많을 수록 더욱 잘 동작한다.
- 저차원(특성이 적은)데이터에서는 다른 모델이 더 좋은 경우가 많다.