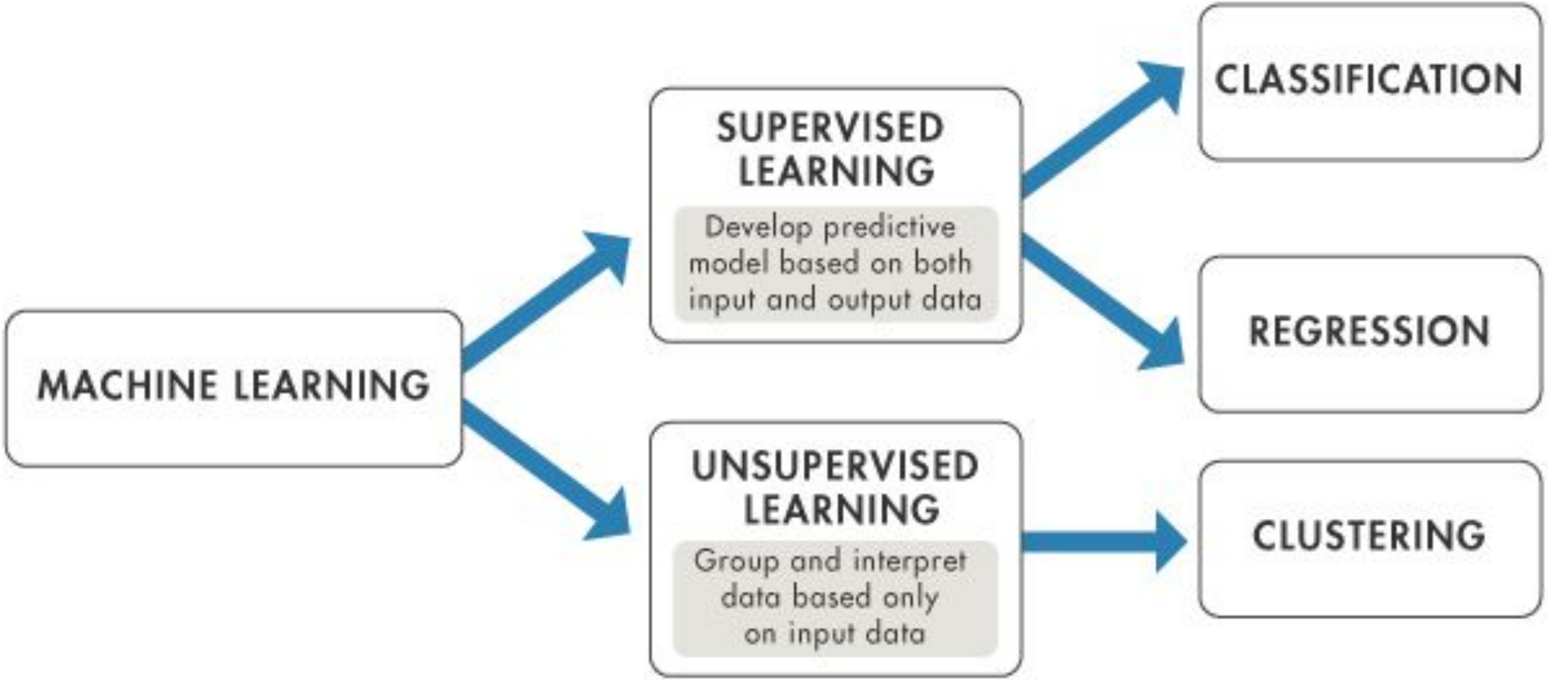
MIDS W207
Applied Machine Learning

Summer 2022

Week 4


MACHINE LEARNING
SUPERVISED LEARNING
Develop predictive model based on both input and output data
UNSUPERVISED LEARNING
Group and interpret data based only on input data
CLASSIFICATION
REGRESSION
CLUSTERING

enjoy algorithms

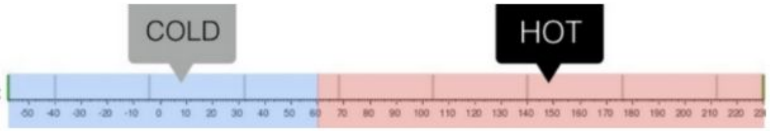
Regression

What will be the temperature tomorrow?

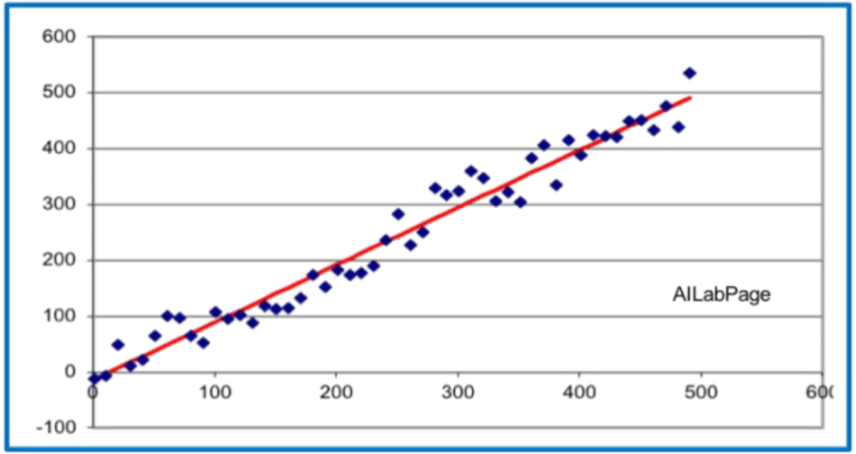
Fahrenheit

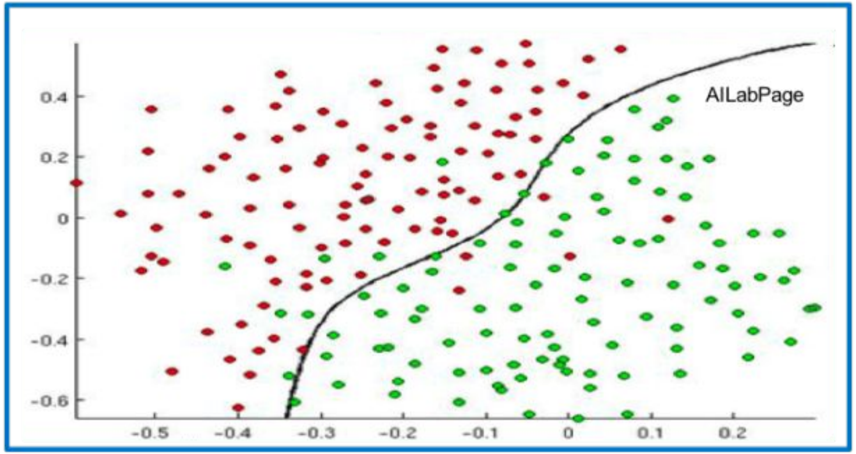
Classification

Will it be hot or cold tomorrow?



Fahrenheit



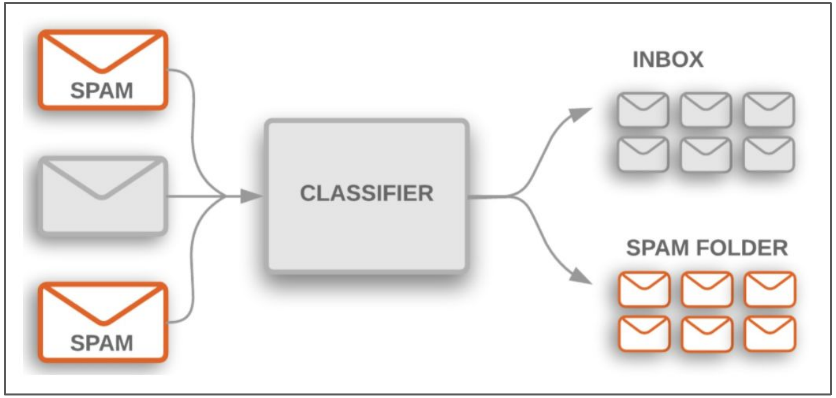


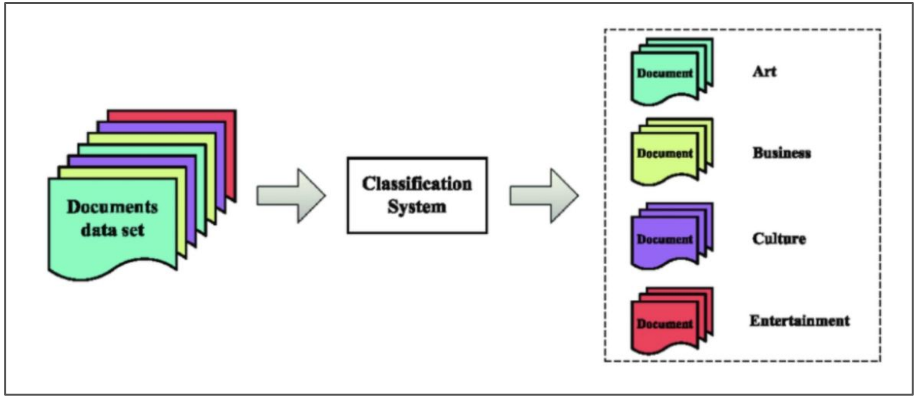
Regression

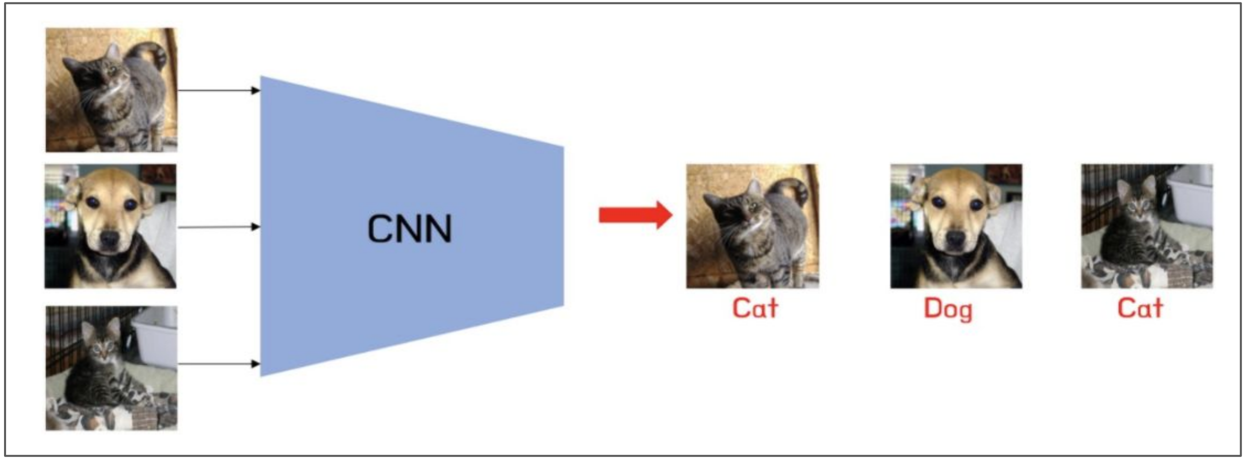
1. The system attempts to predict a value for an input based on past data.
2. Real number / Continuous numbers – Regression problem
3. Example – 1. Temperature for tomorrow

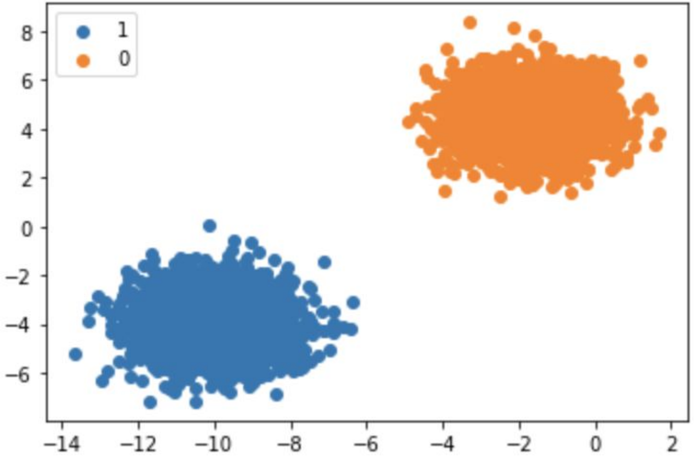
Classification

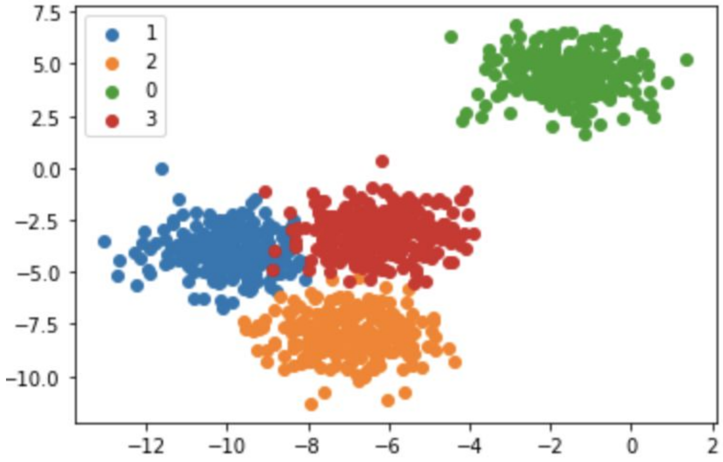
1. In classification, predictions are made by classifying them into different categories.
2. Discreate / categorical variable – Classification problem
3. Example – 1. Type of cancer 2. Cancer Y/N


SPAM
SPAM
CLASSIFIER
INBOX
SPAM FOLDER


Documents data set
Classification System
Document
Art
Document
Business
Document
Culture
Document
Entertainment


CNN
Cat
Dog
Cat





In [1]:
```
import pandas as pd
```

In [3]:
```
# Read a comma-separated values (csv) file into DataFrame.
# filepath_or_bufferstr, path object or file-like object
df = pd.read_csv("diabetes.csv")
```

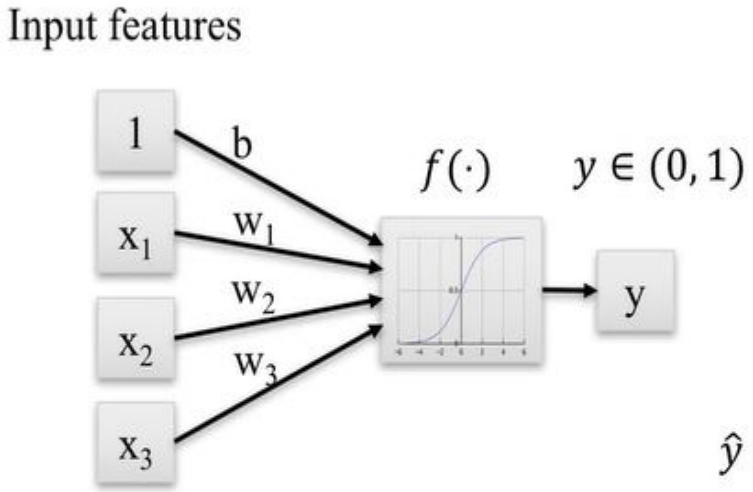
In [4]:
```
df.head(3)
```

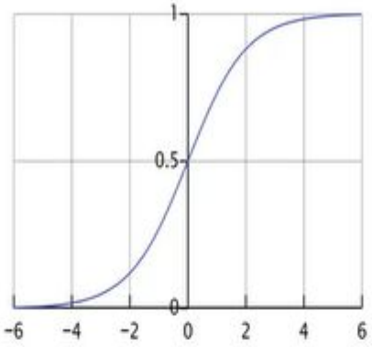
Out[4]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1

SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
6.8	3.2	5.9	2.3	Iris-virginica
6.9	3.1	5.1	2.3	Iris-virginica
4.9	3.0	1.4	0.2	Iris-setosa
5.6	3.0	4.5	1.5	Iris-versicolor
4.8	3.1	1.6	0.2	Iris-setosa
5.8	2.8	5.1	2.4	Iris-virginica
7.2	3.6	6.1	2.5	Iris-virginica
5.1	3.5	1.4	0.3	Iris-setosa
4.7	3.2	1.6	0.2	Iris-setosa
6.6	3.0	4.4	1.4	Iris-versicolor

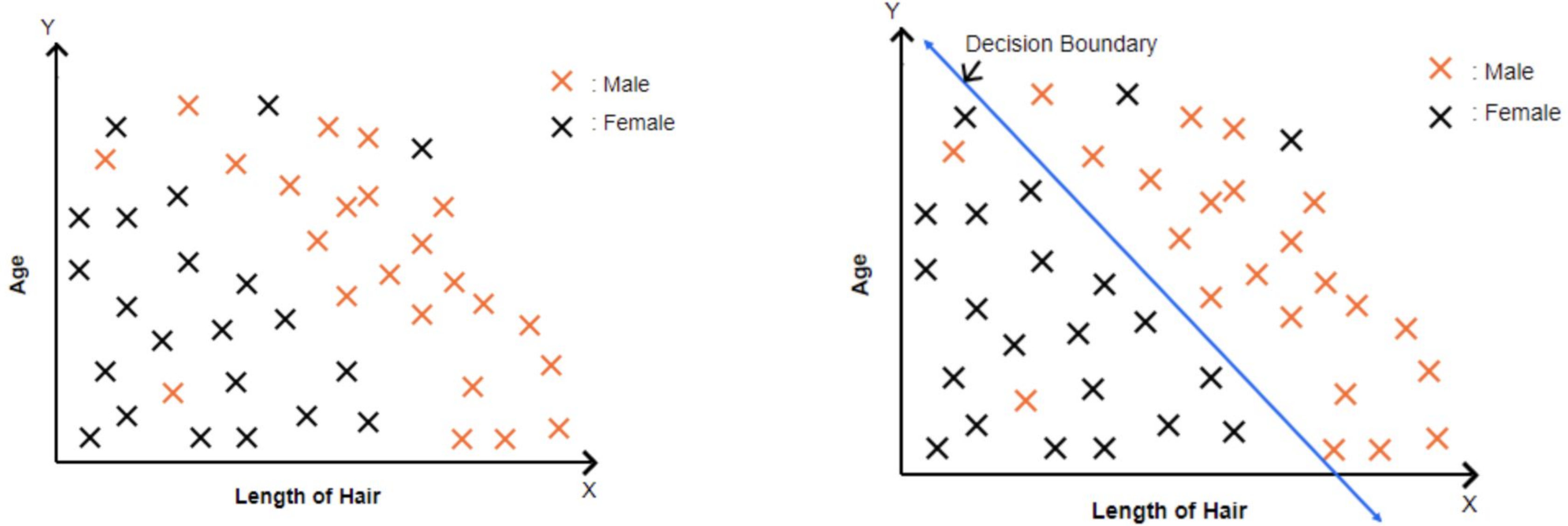
Fig.1: Iris dataset having three categories





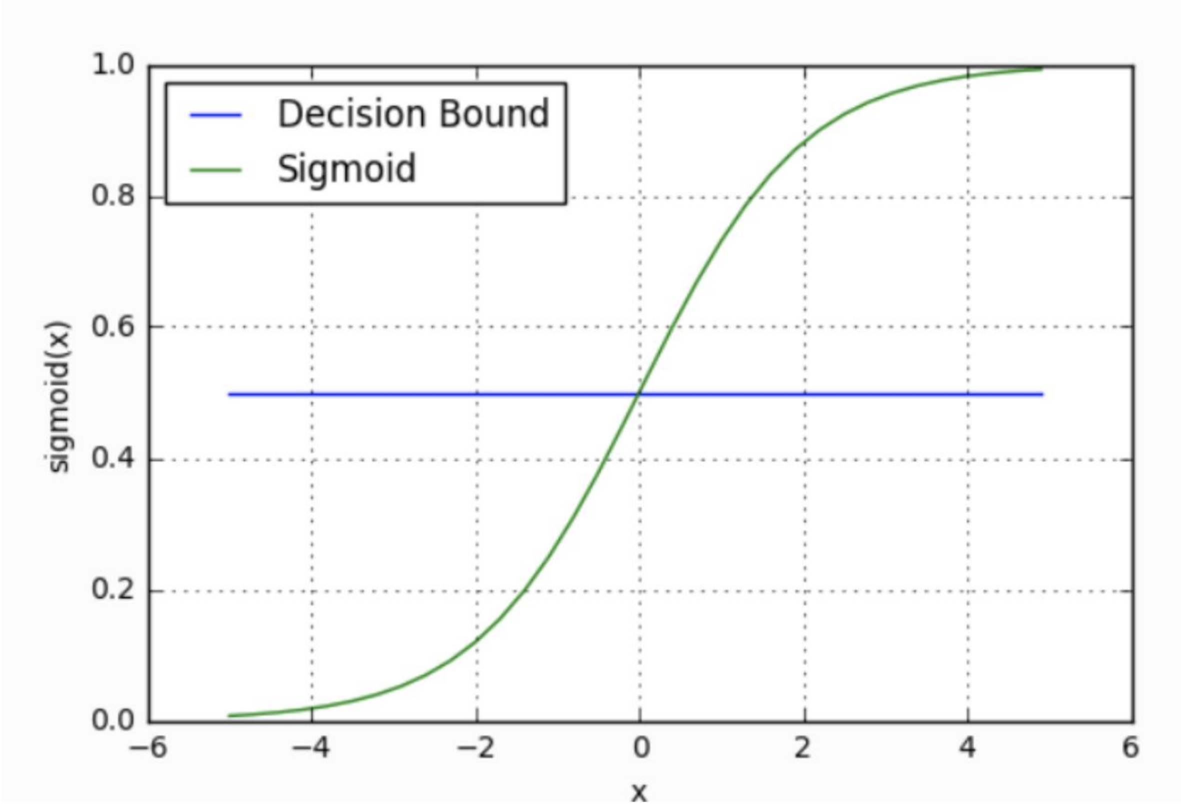
$$\hat{y} = \text{logistic}(\hat{b} + \hat{w}_1 \cdot x_1 + \cdots \hat{w}_n \cdot x_n)$$

$$= \frac{1}{1 + \exp\left[-(\hat{b} + \hat{w}_1 \cdot x_1 + \cdots \hat{w}_n \cdot x_n)\right]}$$



$$w_0 + w_1 x_1 + w_2 x_2 = 0$$

Decision Boundary



Log Loss with Gradient Descent

$$y = g(z) = \frac{1}{1+e^{-z}} = \frac{e^z}{1+e^z}$$

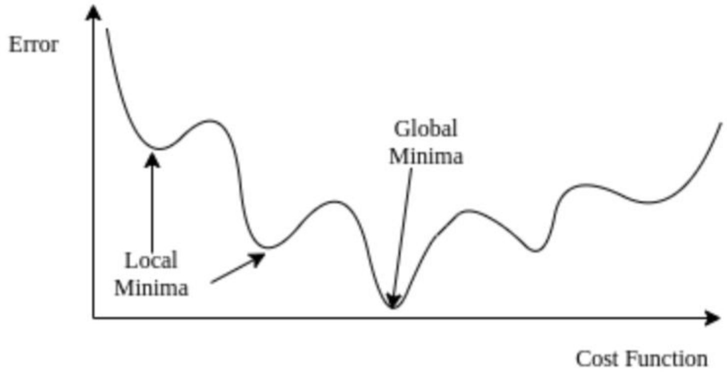
$$h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$$

$$P(y=1|\theta,x) = g(z) = \frac{1}{1+e^{-\theta^T x}}$$

$$P(y=0|\theta,x) = 1 - g(z) = 1 - \frac{1}{1+e^{-\theta^T x}} = \frac{1}{1+e^{\theta^T x}}$$

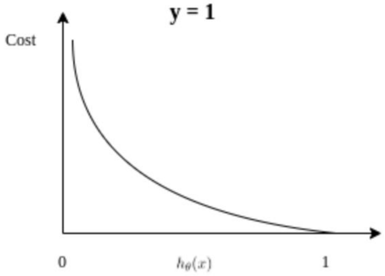
$$P(y|\theta,x) = \left(\frac{1}{1+e^{-\theta^T x}}\right)^y \times \left(1 - \left(\frac{1}{1+e^{\theta^T x}}\right)\right)^{1-y}$$

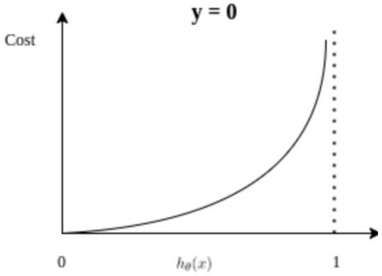
$$cost(h_\theta(x), y) = \begin{cases} -log(h_\theta(x)) & \text{, if } y = 1 \\ -log(1 - h_\theta(x)) & \text{, if } y = 0 \end{cases}$$



$$cost(h_\theta(x), y) = -y^{(i)} \times log(h_\theta(x^{(i)})) - (1 - y^{(i)}) \times log(h_\theta(x^{(i)}))$$

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





$$\theta_j \leftarrow \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

$$\frac{\partial}{\partial \theta_j} J(\theta) = \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right) x_j^{(i)}$$

$$\theta_j \leftarrow \theta_j - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right) x_j^{(i)}$$

Code Review