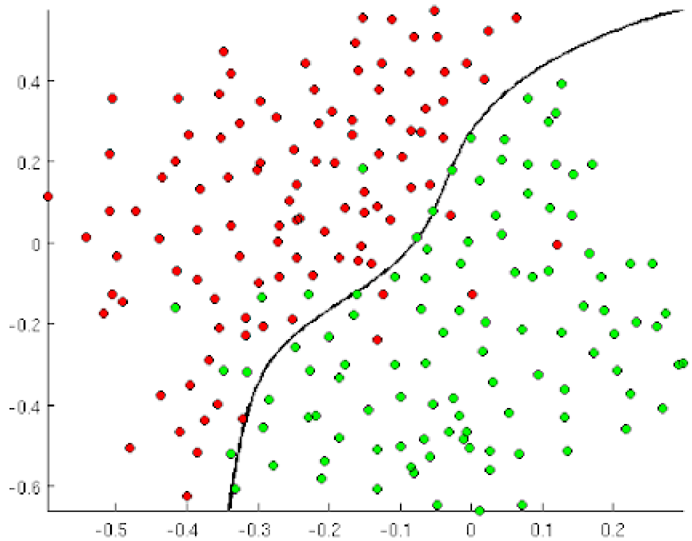
MSE491: Application of Machine Learning in Mechatronic Systems

Classification

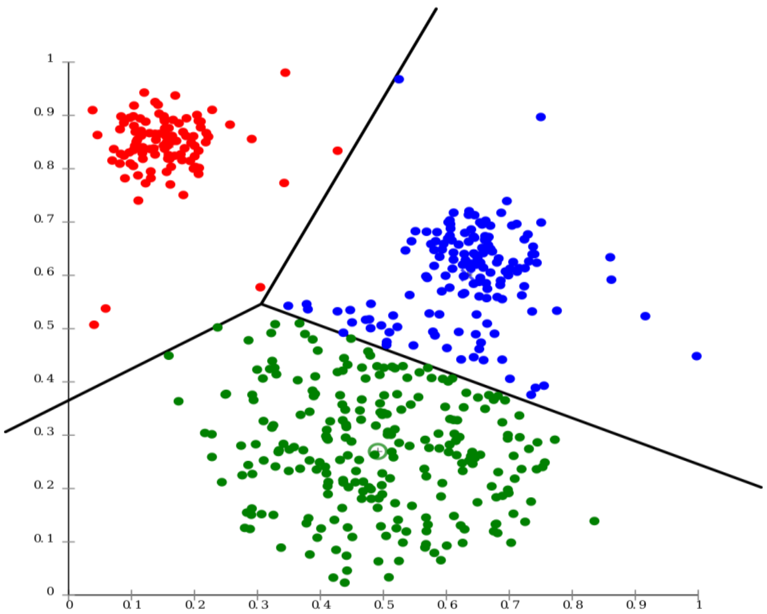
Mohammad Narimani, *Ph.D., P.Eng.*
Lecturer
School of Mechatronic Systems Engineering
Simon Fraser University

Classification

- **What is classification:** Classification is a is a supervised machine learning approach for determining which class the dependent belongs to based on one or more independent variables.



Binary classification



Multi-class classification

Classification

- **Binary Classification:** a training example is assigned to one of two classes.

- **Multi-class classification:** a multi-class classification problem is divided into multiple binary classification datasets.
  - **One-Vs-One (OVO):**

    - it requires one model to be created for each class. The dataset is divided into one dataset for each class versus all other classes e.g. color selection in {red, blue, green} is problem of red vs [blue, green] or blue vs [red, green].

    - this could be an issue for large datasets

  - **One-Vs-Rest (OVR):**

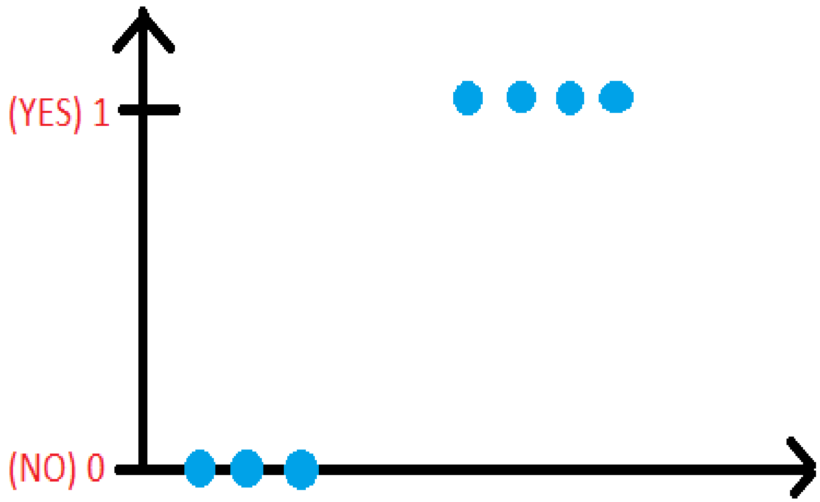
    - the dataset is divided into one dataset for each class versus every other class, e.g. red vs blue, red vs green, and blue vs green

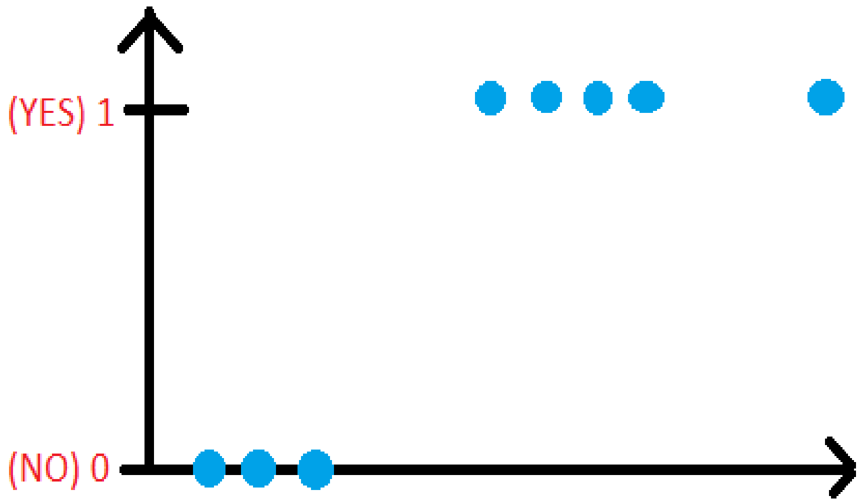
Classification

- **Mechatronic Examples:**
  - Mechanical diagnosis
  - Image recognition
  - Tumour classification (benign vs malignant tumours)
  - Speech recognition

Classification using Linear Regression

- Consider binary classification
  - $y \in \{0,1\}$ 0: "No"
    1: "Yes"
- Logistic Regression is considered as a probability problem

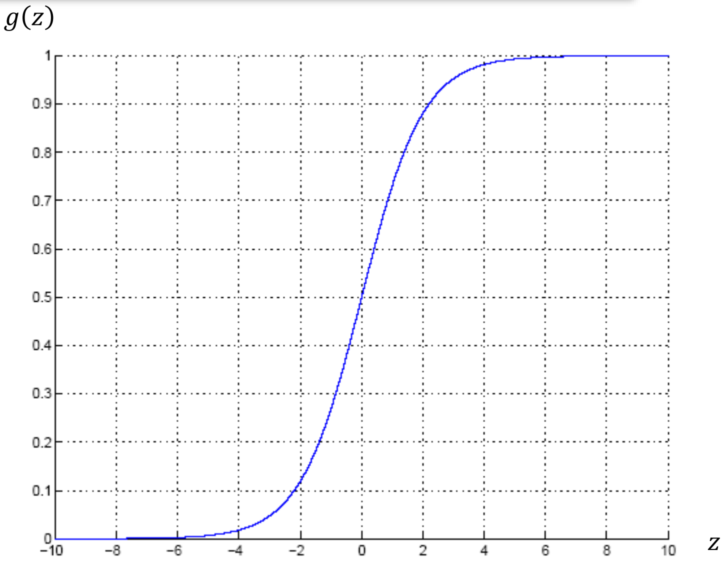




Logistic Regression

- **Logistic Regression Model:**

We want to have $0 \leq h_\theta\ (x) \leq 1$

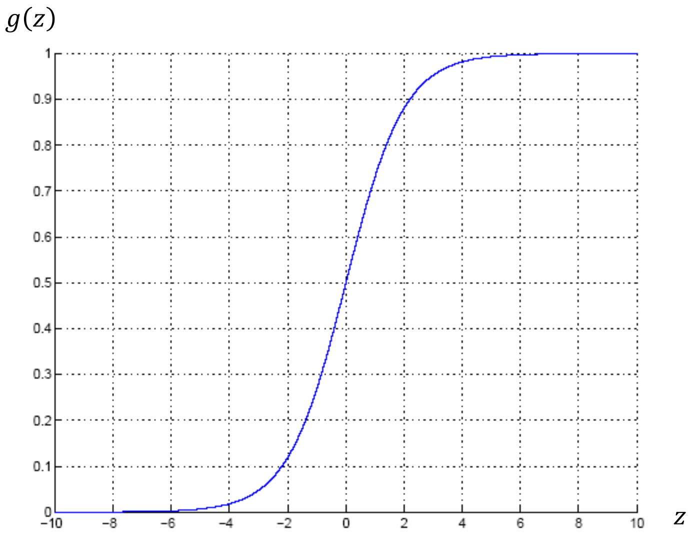


sigmoid (logistic) function:
$$g(z) = \frac{1}{1+e^{-z}}$$

- **Logistic Regression is considered as a probability problem**

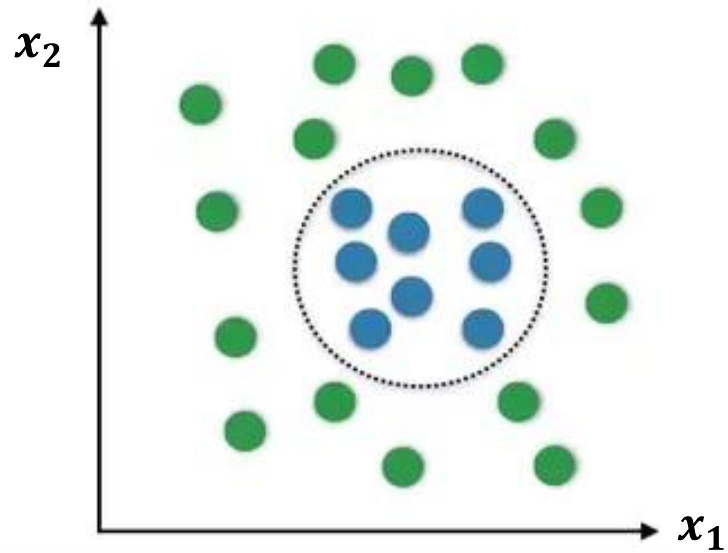
Logistic Regression

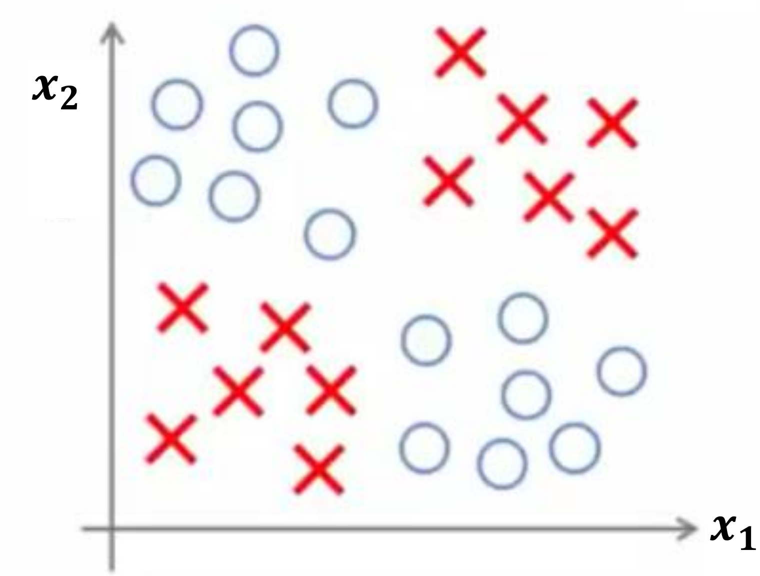
- **Decision Boundary:**



Logistic Regression

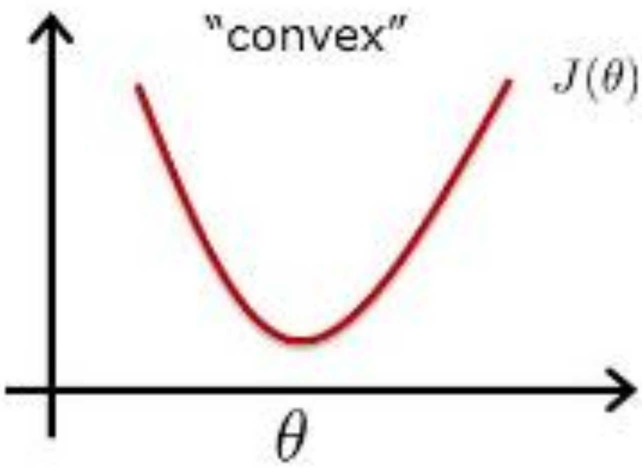
- **Nonlinear Decision Boundary:**

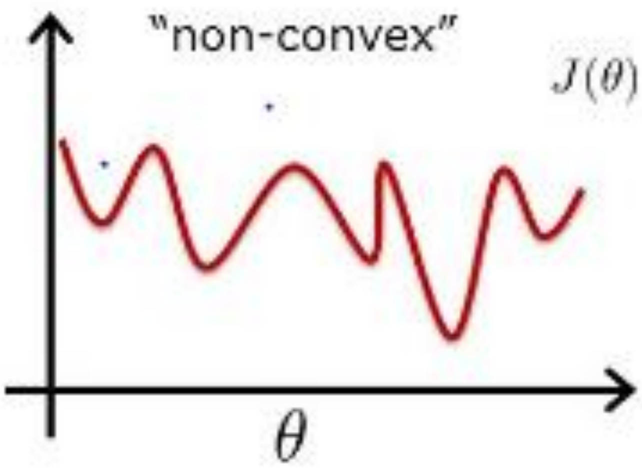




Logistic Regression

- **Cost Function: Preliminary**





Logistic Regression

- **Cost Function: Preliminary**
  - $-\log(w)$

  - $-\log(1-w)$

Logistic Regression

- **Cost Function: Review**

Training set: $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)})\}$

Output for Binary classification: $y \in \{0,1\}$

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

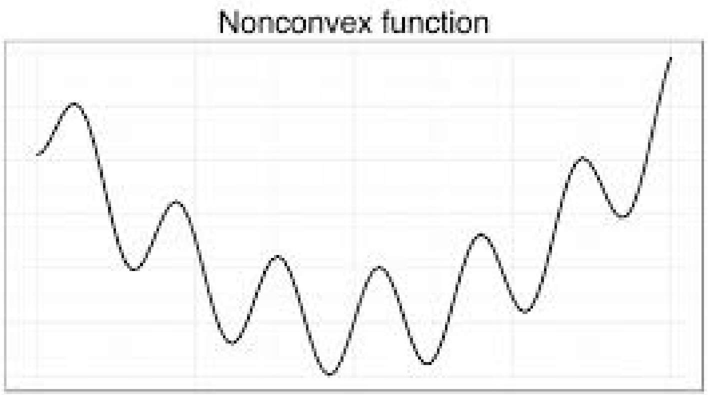
Therefore, θ should be selected in a way to reduce the cost function as the difference between $h_\theta\left(x^{(i)}\right)$ and y.

Logistic Regression

- **Cost Function: cost function selection**

- If Mean Squared Error is a cost functi

$$J(\theta) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)^2$$



- We are looking for a cost function $J(\theta)$ that:

Logistic Regression

- **Cost Function: cost function selection**

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

Logistic Regression

- **Cost Function: Gradient Decent**

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

The same update as linear regression!!

- Matrix format:

Logistic Regression

- **Cost Function: Multu-class Classification**

(Next lecture)