



# Machine Learning

## La Régression logistique

Présenté par:

AYMANE MAGHOUTI (Étudiant en Ingénierie des Données (ID2))

# Outline:

- 1- Introduction à la classification
- 2- Regression vs Classification
- 3- Les types de classification
- 4- L'algorithme de régression logistique
- 5- Evaluation d'un modèle de classification:
  - Matrice de confusion
  - Exactitude (Accuracy)

## 6- Mini - Project:

- Préparation des données
- Entraînement du modèle (Régression logistique)
- Evaluation de modèle
- Déploiement du modèle
- Faire des test

# Introduction à la classification

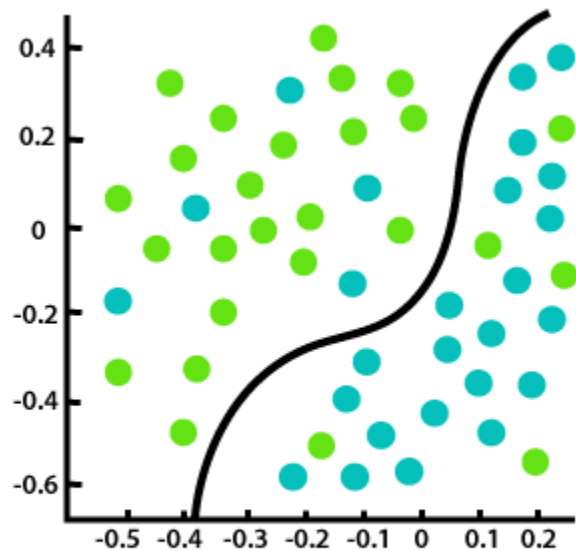
➤ **Variable discrète** : La variable à prédire peut prendre une valeur d'un ensemble fini et non ordonné  $L$  (qu'on appelle des classes ou catégories), tel que :

couleur des yeux  $\in \{\text{marron, bleu, vert}\}$

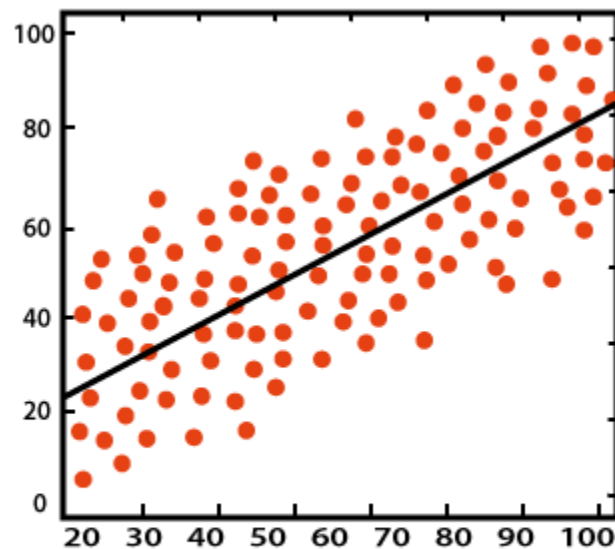
email  $\in \{\text{spam, ham}\}$

➤ Dans l'apprentissage supervisé, quand la variable à prédire prend une valeur discrète, on parle d'un problème de classification.

# Regression vs Classification

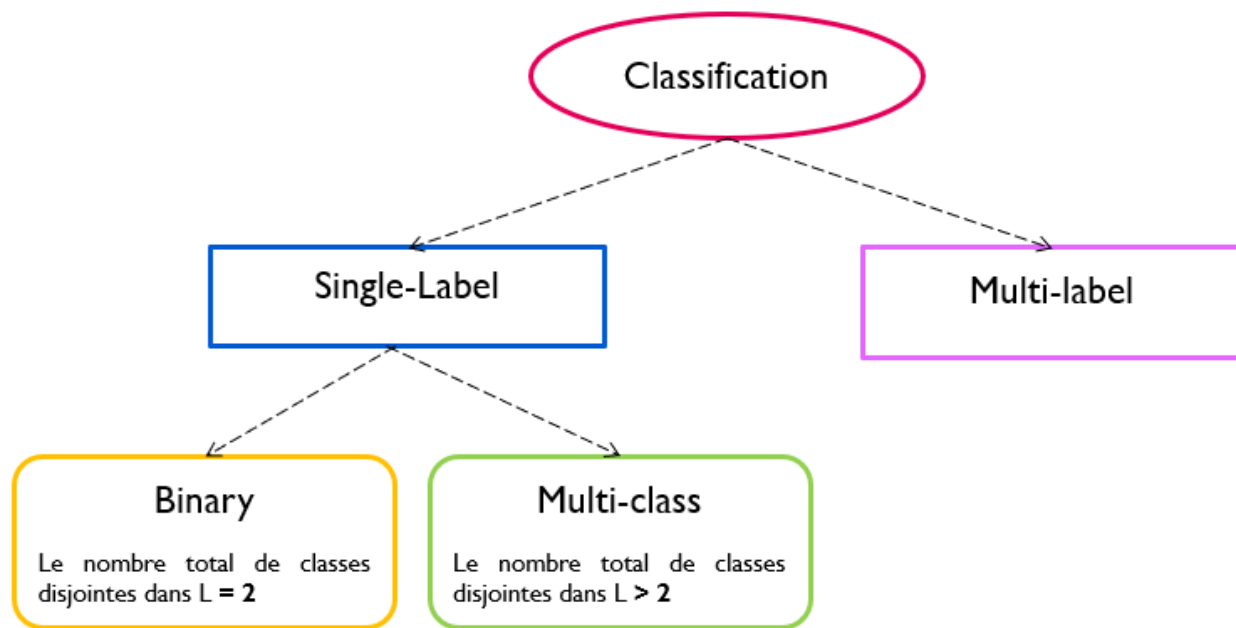


Classification



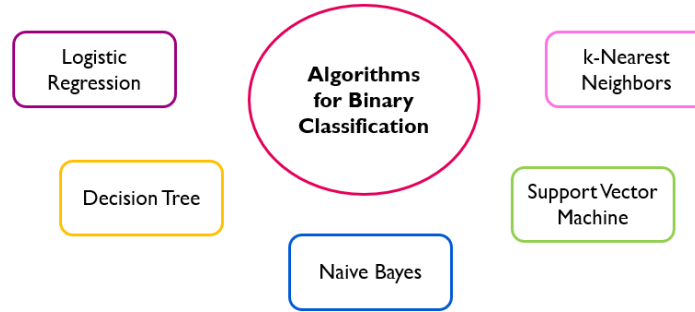
Regression

# Les types de classification

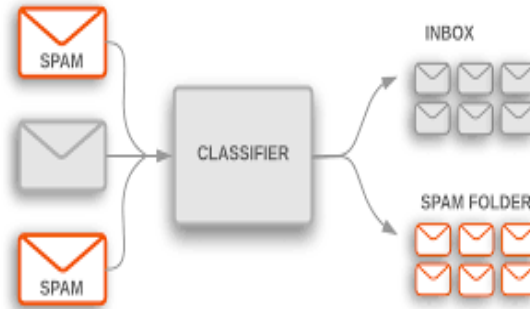


# Single-Label : Binary Classification

➤ Classification binaire fait référence aux tâches de classification qui ont deux classes.



➤ Exemple:



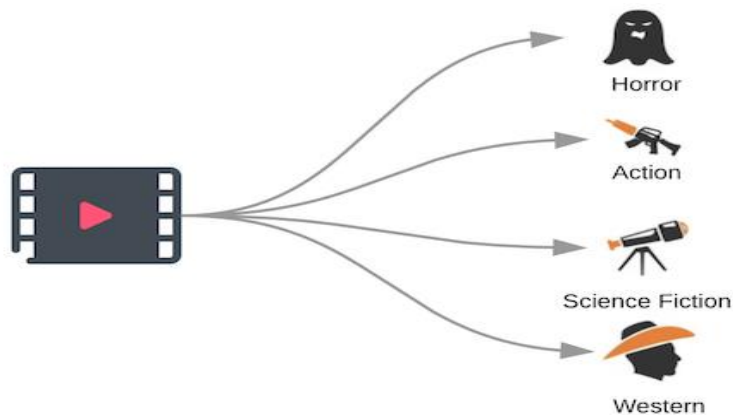
# Single-Label : Multi-Class Classification

- La classification multi-classes est un problème de classification des instances dans l'une des trois classes ou plus.
- Elle suppose que le choix est entre plus de deux classes, mais l'instance d'entrée doit être affectée à une et une seule étiquette de classe.
- Exemple:



# Multi-Label Classification

- La classification multi-étiquettes consiste à prédire zéro ou plusieurs étiquettes de classes.
- Le nombre d'étiquettes par instance n'est pas fixe.
- Exemple:





# La Régression logistique

- Le Modèle de régression linéaire

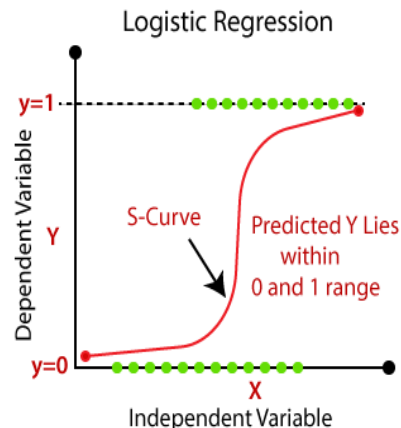
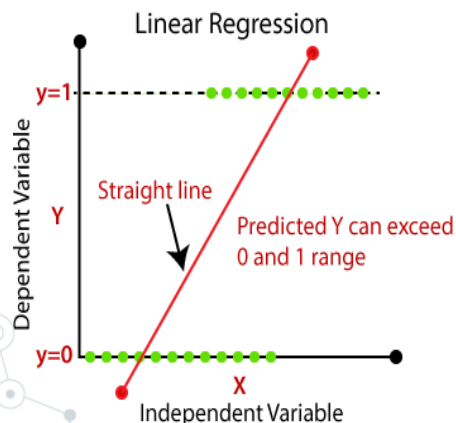
$$h_{\theta}(x) = \theta^T x.$$

- Le modèle de régression logistique :

$$0 \leq h_{\theta}(x) \leq 1, \text{ où } h_{\theta}(x) = g(\theta^T x).$$

- $g(.)$  est la fonction logistique, (fonction sigmoïde)

$$z = \theta^T x ; \quad g(z) = \frac{1}{1 + e^{-z}}$$



➤ Pour prédire une valeur discrète 0 ou 1, la sortie de la fonction d'hypothèse est traduite comme suit :

$$h_{\theta}(x) \geq 0.5 \rightarrow y = 1$$

$$h_{\theta}(x) < 0.5 \rightarrow y = 0$$

➤ Si notre entrée à  $g$  est  $\theta^T x$ , alors cela signifie :

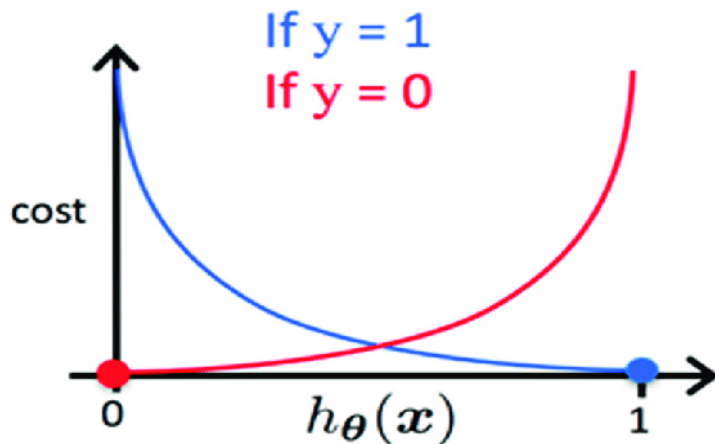
$$\begin{array}{l} h_{\theta}(x) = g(\theta^T x) \geq 0.5 \\ \text{when } \theta^T x \geq 0 \end{array} \quad \longrightarrow \quad \begin{array}{l} \theta^T x \geq 0 \Rightarrow y = 1 \\ \theta^T x < 0 \Rightarrow y = 0 \end{array}$$

## La fonction de coût pour la régression logistique:

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

Avec:

$$\text{Cost}(h_{\theta}(x), y) = -y \log(h_{\theta}(x)) - (1 - y) \log(1 - h_{\theta}(x))$$



## Minimisation du fonction de coût :

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

Descente du gradient

Répéter {

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta) \text{ (mettre à jour les } \theta_j \text{ simultanément)}$$

}

Descente du gradient

(mettre à jour les  $\theta_j$  simultanément)

Répéter {

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

}

# Evaluation d'un modèle de classification:

## Confusion Matrix

Confusion matrix		Reality	
		Negative : 0	Positive : 1
Prediction	Negative : 0	True Negative : TN	False Negative : FN
	Positive : 1	False Positive : FP	True Positive : TP

## Accuracy (exactitude)

$$\text{Accuracy} = \frac{TP + TN}{\text{total}}$$

## Exemple:

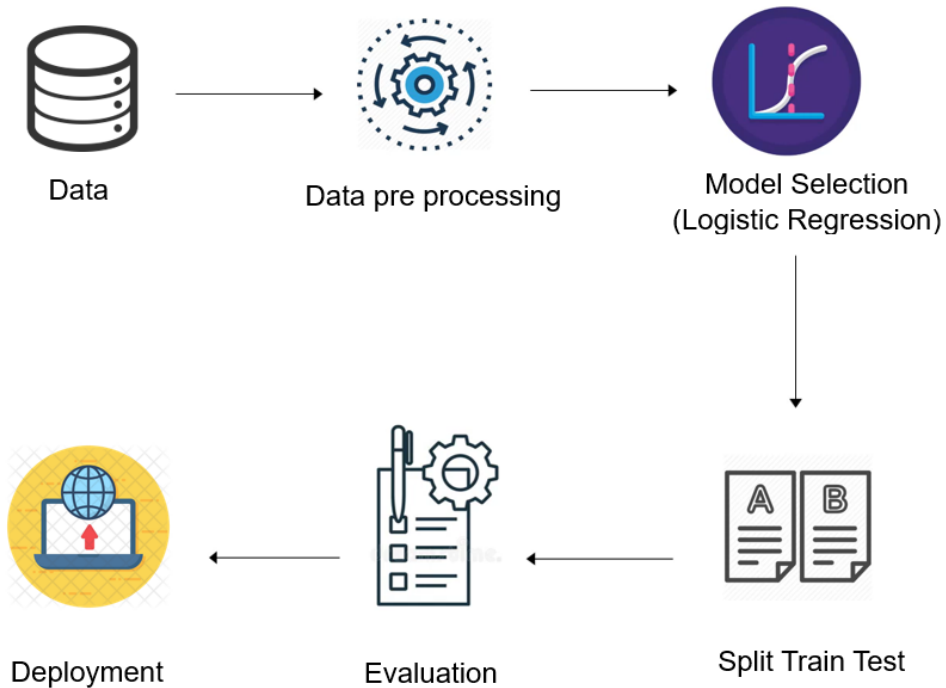
		Predicted Class	
		Spam	Non-Spam
Actual Class	Spam	TP=45	FN=20
	Non-Spam	FP=5	TN=30

Total = 100

$$\text{Accuracy} = (45+30) / 100 = 0.75$$

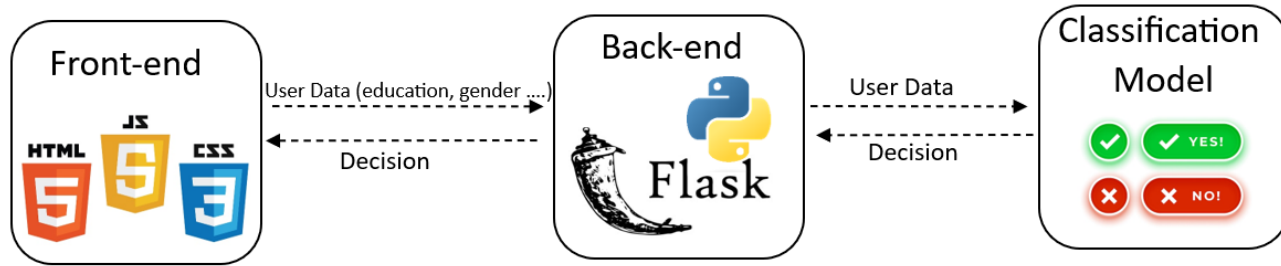
# Project : Loan Credit Classification

## Work Flow



# Architecture

- Web app pour l'estimation du crédit



By Aymane-MG





**Merci pour votre attention!**

