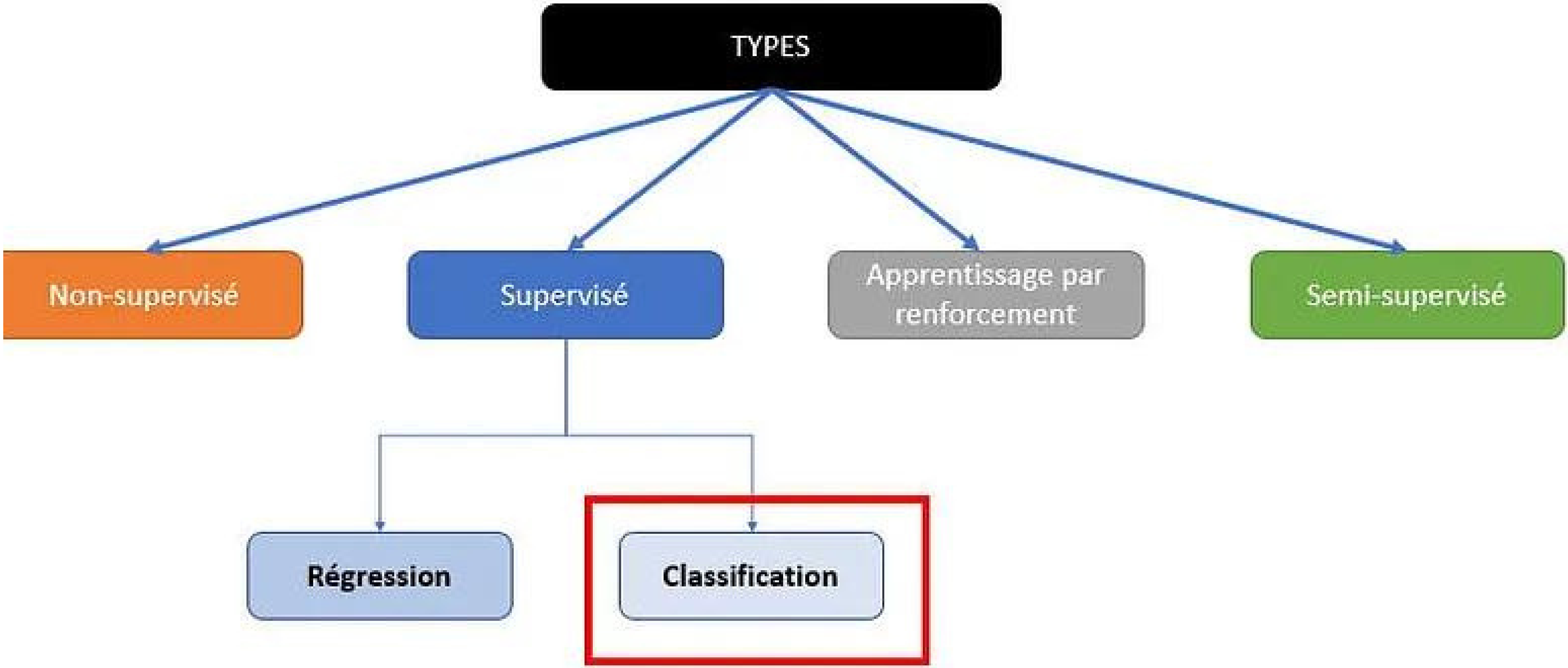
RECAP REGRESSION

La régression est une technique d'apprentissage supervisé utilisée pour prédire des valeurs continues à partir de variables indépendantes. Par exemple, elle peut être utilisée pour prédire le prix d'une maison en fonction de ses caractéristiques telles que la superficie, le nombre de chambres, etc

- La **problématique** se pose lorsque nous devons attribuer des données à des catégories discrètes plutôt que de prédire des valeurs continues
- Prenons l’exemple :

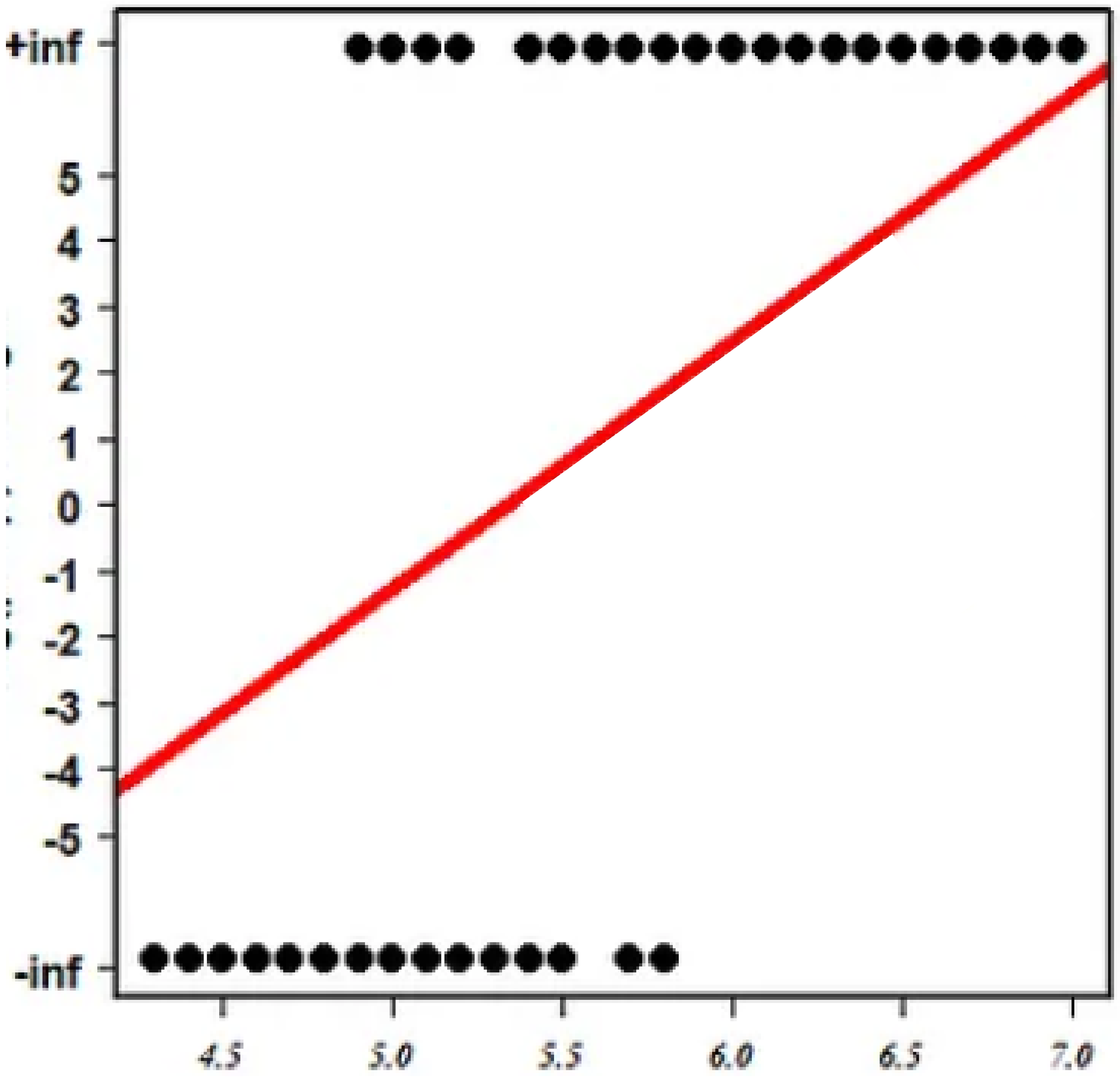
décider si un email est du spam ou non est une tâche de classification, car les résultats sont binaires (spam ou non-spam), pas continus. La régression, qui prédit des valeurs continues, n'est pas adaptée à cette problématique, d'où le recours à la classification pour traiter de telles situations.

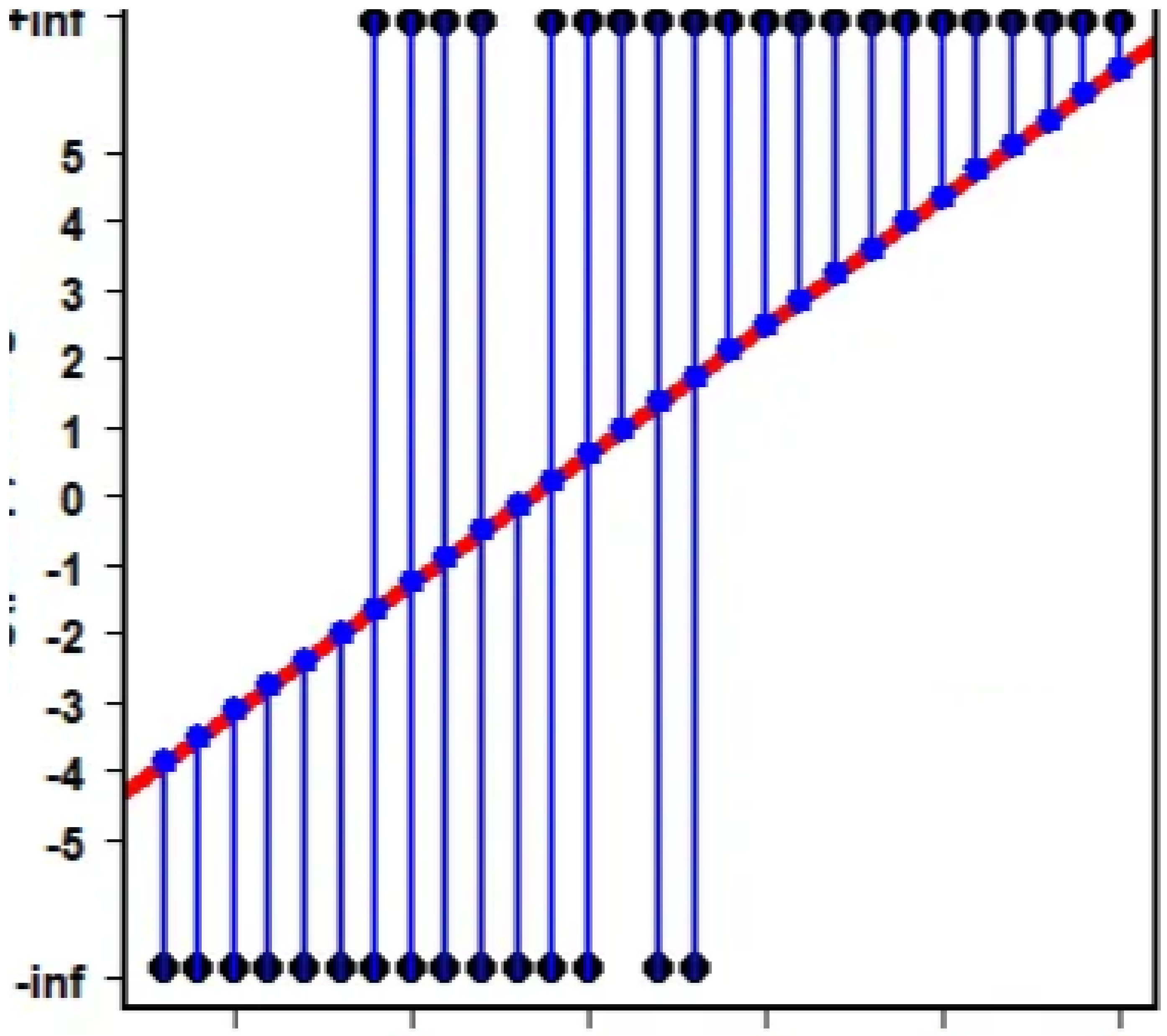

TYPES
Non-supervisé
Supervisé
Apprentissage par renforcement
Semi-supervisé
Régression
Classification

La classification consiste a assigner des étiquettes ou des catégories à des données en fonction de certaines caractéristiques spécifiques. L'objectif de la classification est de développer un modèle qui peut prédire la classe d'un nouvel exemple en se basant sur des exemples d'entraînement préalablement étiquetés.

On ne peut pas trouver une fonction qui prédit de manière certaine un résultat de 0 ou 1 à 100%, mais plutôt on opte pour une approche probabiliste où les prédictions sont exprimées sous forme de probabilités, indiquant la confiance du modèle dans chaque classe.

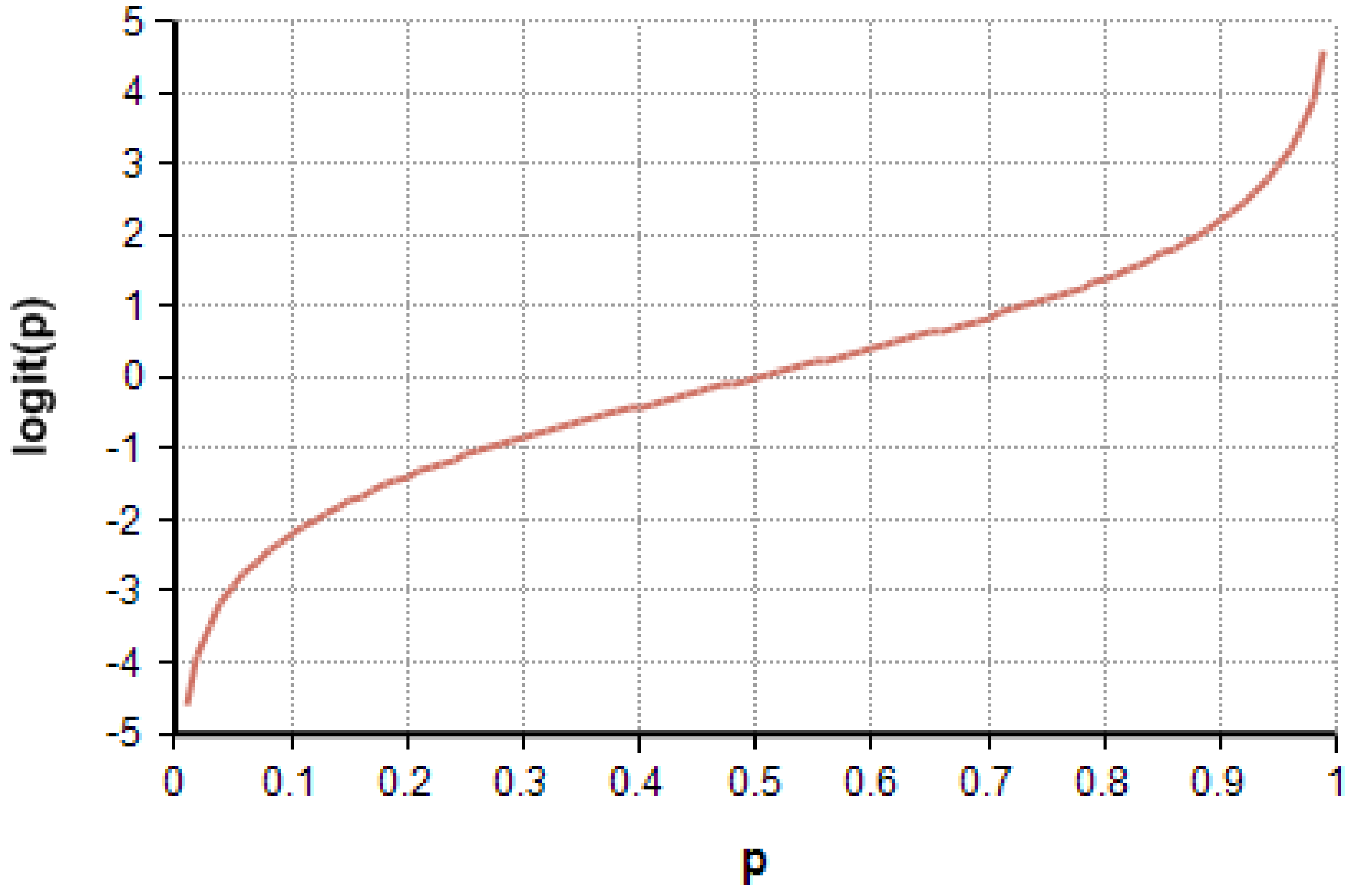
La régression logistique est appelée ainsi en raison de son origine historique et de sa similarité conceptuelle avec la régression linéaire. Bien que la régression logistique soit en réalité utilisée pour la classification, son nom découle du fait qu'elle est basée sur une fonction de régression qui modélise la relation entre les variables d'entrée et la probabilité d'appartenance à une classe particulière.

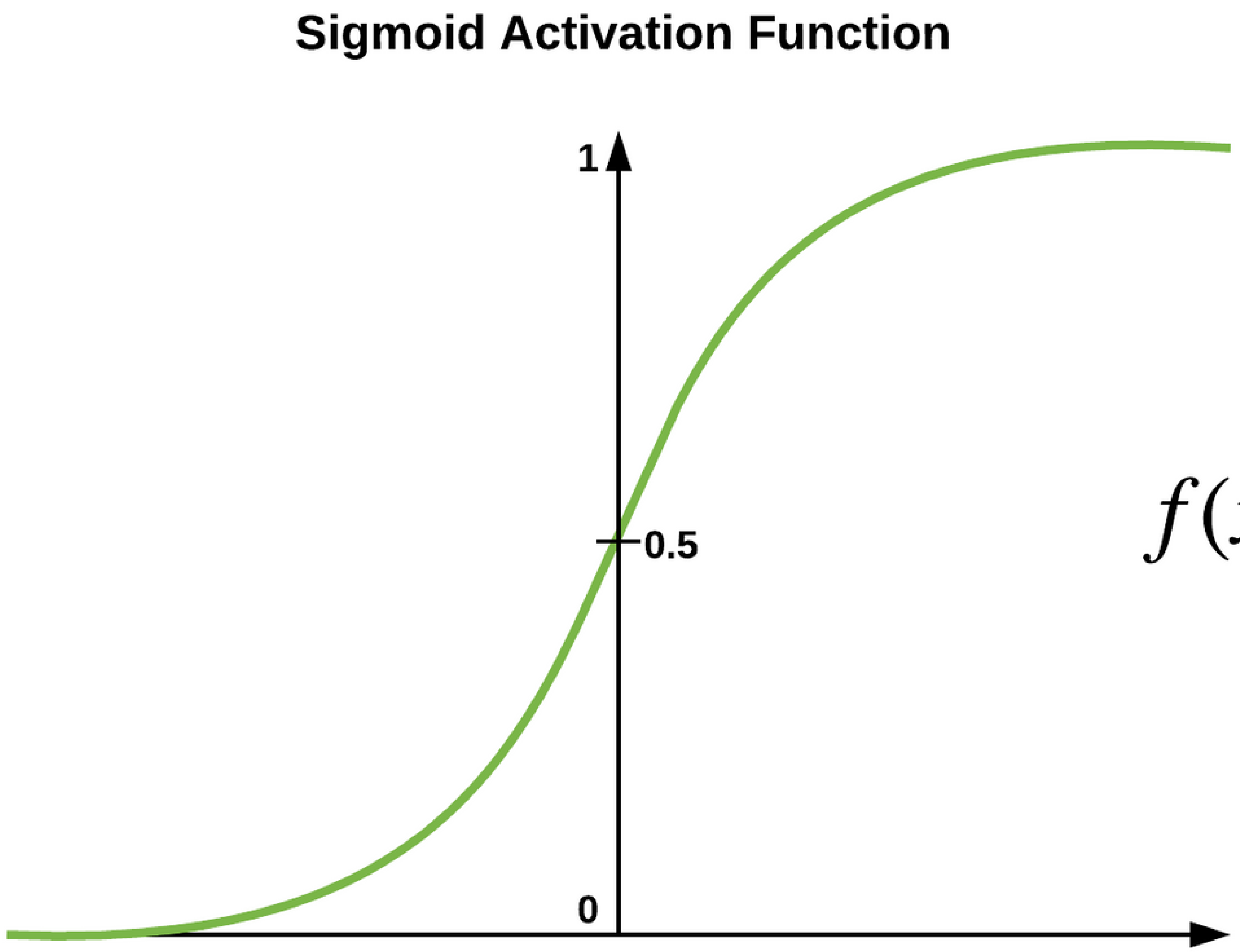

+inf
5
4
3
2
1
0
-1
-2
-3
-4
-5
-inf
4.5
5.0
5.5
6.0
6.5
7.0


+inf
5
4
3
2
1
0
-1
-2
-3
-4
-5
-inf

$$\mathrm{logit}(p) = \ln\left(\frac{p}{1-p}\right)$$

[0 ,1] => [-inf ,+inf]





$$f(x) = \frac{1}{1+e^{-\beta x}}$$

[-inf ,+inf] => [0 ,1]

Bernoulli Distribution

$$X \sim \text{Bernoulli}(p) \qquad P_X(k) = \begin{cases} p & k = 1 \\ 1-p & k = 0 \end{cases}$$

$$\mathbb{E}(X) = p \qquad Var(X) = p(1-p)$$

$$P(Y = y|X = x) = \sigma(\beta^T x)^y \cdot [1 - \sigma(\beta^T x)]^{(1-y)}$$

Vraisemblance:

$$L(\beta) = {}^n\prod_{i=1} P(Y = y^{(i)} \mid X = x^{(i)}) \ldots\ldots\ldots\ldots\ldots\ldots$$

$$L(\beta) = {}^n\prod_{i=1} \sigma(\beta^T x^{(i)})^{y(i)} \cdot [1 - \sigma(\beta^T x^{(i)})]^{(1-y(i))} \ldots.$$

log Vraisemblance:

$$LL(\beta) = {}^n\sum_{i=1} y^{(i)} \log \sigma(\beta^T x^{(i)}) + (1 - y^{(i)}) \log[1 - \sigma(\beta^T x^{(i)})].$$

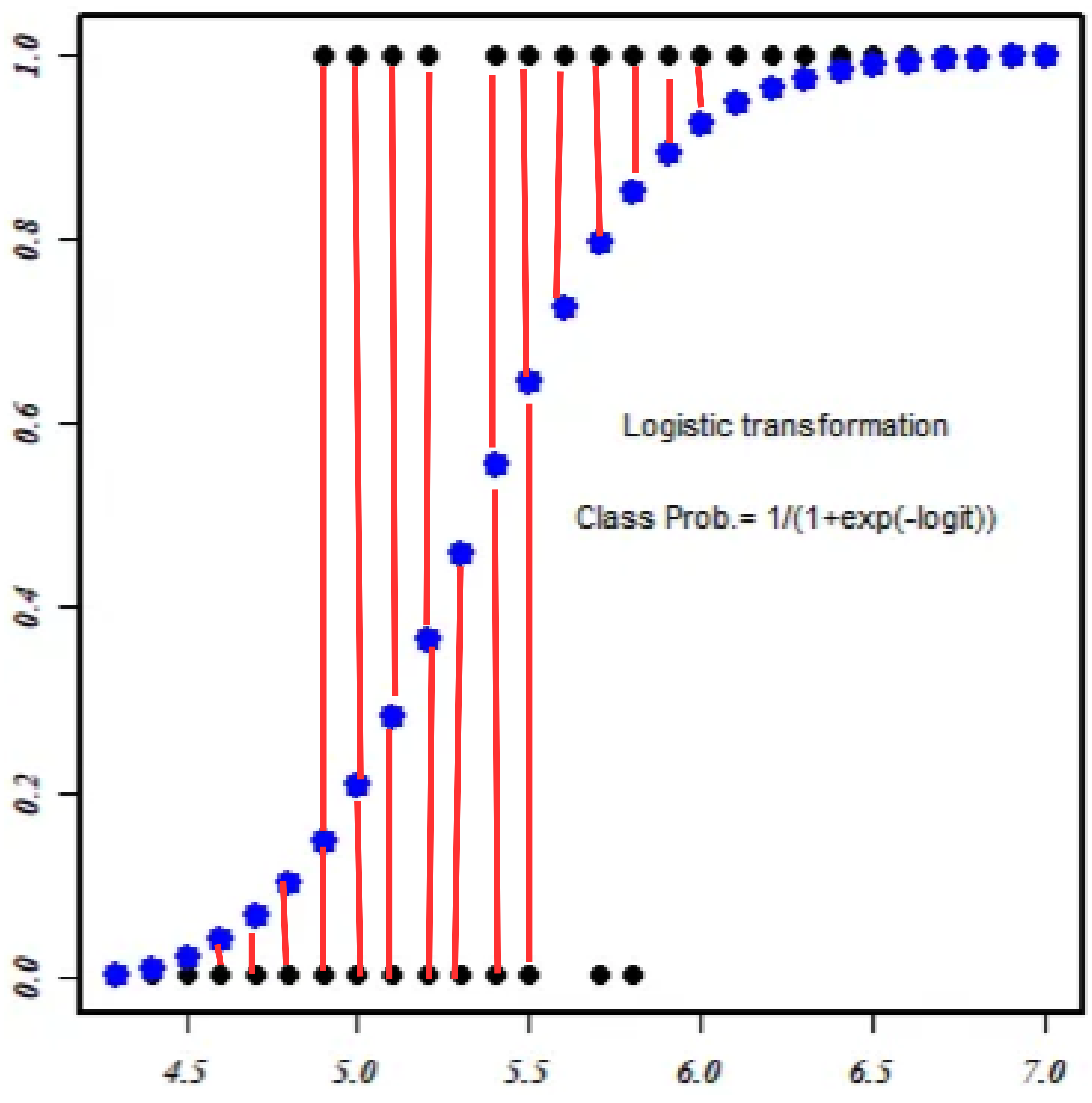
Gradient Ascent

$$
\begin{aligned}
\frac{\partial LL(\theta)}{\partial \theta_j} &= \frac{\partial}{\partial \theta_j} y \log \sigma(\theta^T \mathbf{x}) + \frac{\partial}{\partial \theta_j} (1-y) \log[1-\sigma(\theta^T \mathbf{x}] \\
&= \left[\frac{y}{\sigma(\theta^T \mathbf{x})} - \frac{1-y}{1-\sigma(\theta^T \mathbf{x})}\right] \frac{\partial}{\partial \theta_j} \sigma(\theta^T \mathbf{x}) \\
&= \left[\frac{y}{\sigma(\theta^T \mathbf{x})} - \frac{1-y}{1-\sigma(\theta^T \mathbf{x})}\right] \sigma(\theta^T \mathbf{x})[1-\sigma(\theta^T \mathbf{x})] x_j \\
&= \left[\frac{y-\sigma(\theta^T \mathbf{x})}{\sigma(\theta^T \mathbf{x})[1-\sigma(\theta^T \mathbf{x})]}\right] \sigma(\theta^T \mathbf{x})[1-\sigma(\theta^T \mathbf{x})] x_j \\
&= \left[y-\sigma(\theta^T \mathbf{x})\right] x_j
\end{aligned}
$$

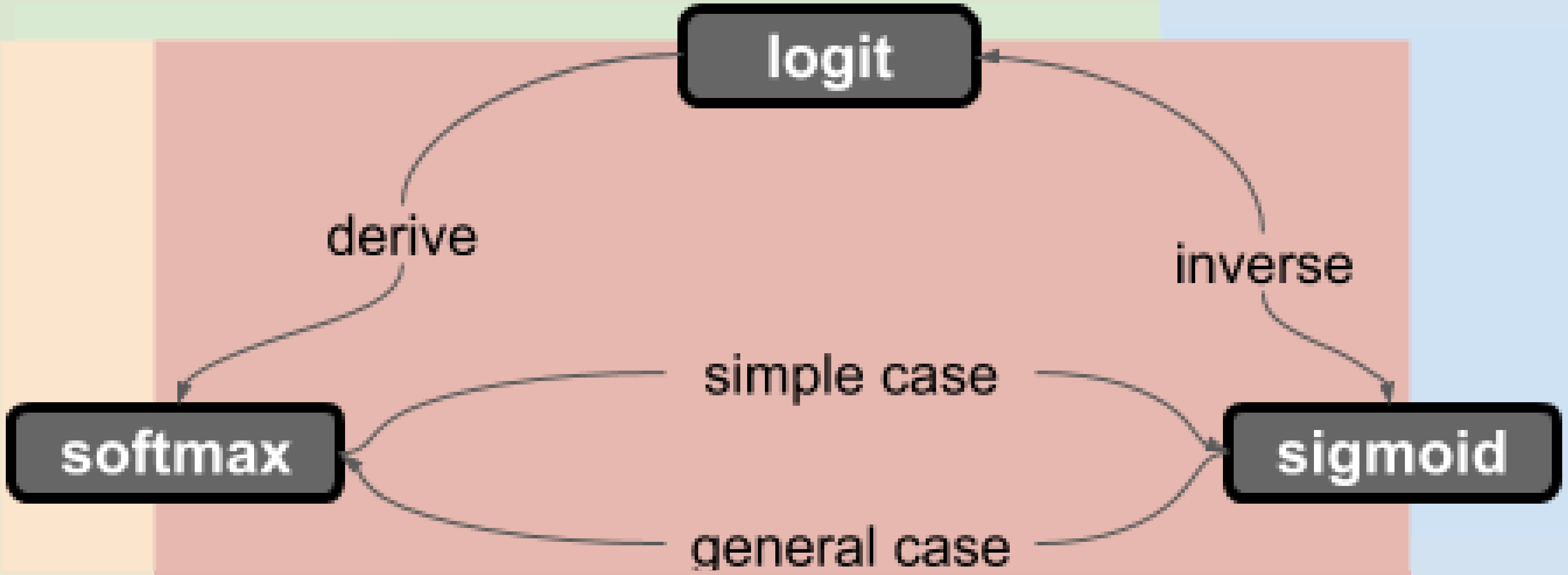
$$
\frac{\partial LL(\theta)}{\partial \theta_j} = \sum_{i=1}^{n} \left[y^{(i)} - \sigma(\theta^T \mathbf{x}^{(i)})\right] x_j^{(i)}
$$

$$
\begin{aligned}
\theta_j^{\text{new}} &= \theta_j^{\text{old}} + \eta \cdot \frac{\partial LL(\theta^{\text{old}})}{\partial \theta_j^{\text{old}}} \\
&= \theta_j^{\text{old}} + \eta \cdot \sum_{i=1}^{n} \left[y^{(i)} - \sigma(\theta^T \mathbf{x}^{(i)})\right] x_j^{(i)}
\end{aligned}
$$

Comparaison:



Géneralisation pour les problémes de Multi-classifcation



Levez la main si vous avez des questions !