

Régression Logistique

1. Introduction

La régression logistique est un algorithme de classification supervisée utilisé pour prédire la probabilité qu'une observation appartienne à une certaine classe. Contrairement à la régression linéaire qui prédit une valeur continue, la régression logistique prédit une probabilité (entre 0 et 1).

2. Fonction d'activation : Sigmoid

La sortie du modèle est une probabilité estimée par la fonction logistique (sigmoid) :

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

3. Le Modèle

On modélise la probabilité par :

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

Où :

- x est le vecteur d'entrée (features)
- θ est le vecteur des paramètres
- $\sigma(z)$ est la fonction sigmoïde

4. Fonction Coût

La fonction de coût pour la régression logistique est définie comme :

$$J(\theta) = -\frac{1}{m} \times \sum [y_i \log(\theta(x_i)) + (1 - y_i) \log(1 - \theta(x_i))]$$

Où :

- m est le nombre d'exemples d'entraînement
- $y(i)$ est la vraie étiquette pour l'exemple i

5. Descente de Gradient

Pour minimiser $J(\theta)$, on utilise la descente de gradient. La mise à jour des paramètres est :

$$\theta := \theta - \alpha * \nabla J(\theta)$$

Avec :

$$\nabla J(\theta) = \left(\frac{1}{m}\right) \times X^T(\theta(x) - y)$$

Et α est le taux d'apprentissage (learning rate).

6. Algorithme d'entraînement

1. Initialiser θ aléatoirement
2. Répéter jusqu'à la convergence :
 - a. Calculer $\theta(x)$
 - b. Calculer le gradient $\nabla J(\theta)$
 - c. Mettre à jour θ
3. Retourner θ final

7. Applications et Avantages

La régression logistique est particulièrement efficace pour :

- La classification binaire (spam/non-spam, malade/sain, etc.)
- Elle est simple à implémenter, rapide à entraîner et interprétable.

8. Conclusion

La régression logistique est une méthode simple mais puissante pour les problèmes de classification binaire. Sa compréhension mathématique repose sur la fonction sigmoïde, la log-loss et l'optimisation par descente de gradient.