



1

Logistic Regression

Julien ABOUTARD

Marwin LAUNAY

DevIA#2 – 24/10/2022

Présentation de la Logistic Regression

- Avant d'être un Algorithme c'est un Modèle Mathématique.
- Elle permet de calculer la probabilité d'un certain résultat en fonction des relations entre les divers caractéristiques.
- Deux grand type de Logistic Regression : Binaire et Multinomiale puis La "Ordinale" qui est un sous type de la Multinomiale

Secteur d'application fréquent



Médecine



Politique



Commercial



Finance



Test de produit



E-commerce

Les Variables de Logistic Regression

un modèle de régression dont la variable dépendante est dichotomique/binaire.

variables indépendantes peuvent être quantitatives (continues ou discrète) ou qualitative (catégorielle)

Leur lien est appelé "relation"

5

exemple



Les Hyperparamètres en Machine Learning

C : force de régularisation doit être un flottant positif (compris généralement entre 0.001 et 1)

Solver : possède plusieurs paramètres en son sein

-lbfgs

-sag

-saga

-newton-cg

-liblinear

Penalty : "L1", "L2" ou les deux (elasticnet) on peut aussi ne pas lui en assigner avec "None"

L1 ou L2 ?

7

- Pénaliser un algorithme

(Attention ! certaine Penalty ne fonctionne pas avec certains solvers, documentation recommandée !!)

-Système de "weight"

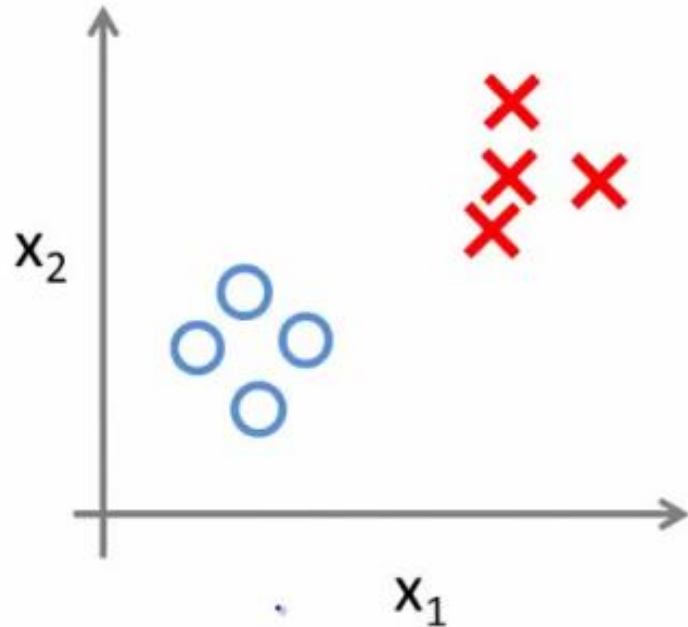
L1(ou Lasso) : combat l'overfitting en rendant les weight à 0

L2(ou Ridge): combat l'overfitting en forçant les weights à être minime, mais sans les rendre exactement 0.



Un peu de math

Binary classification:



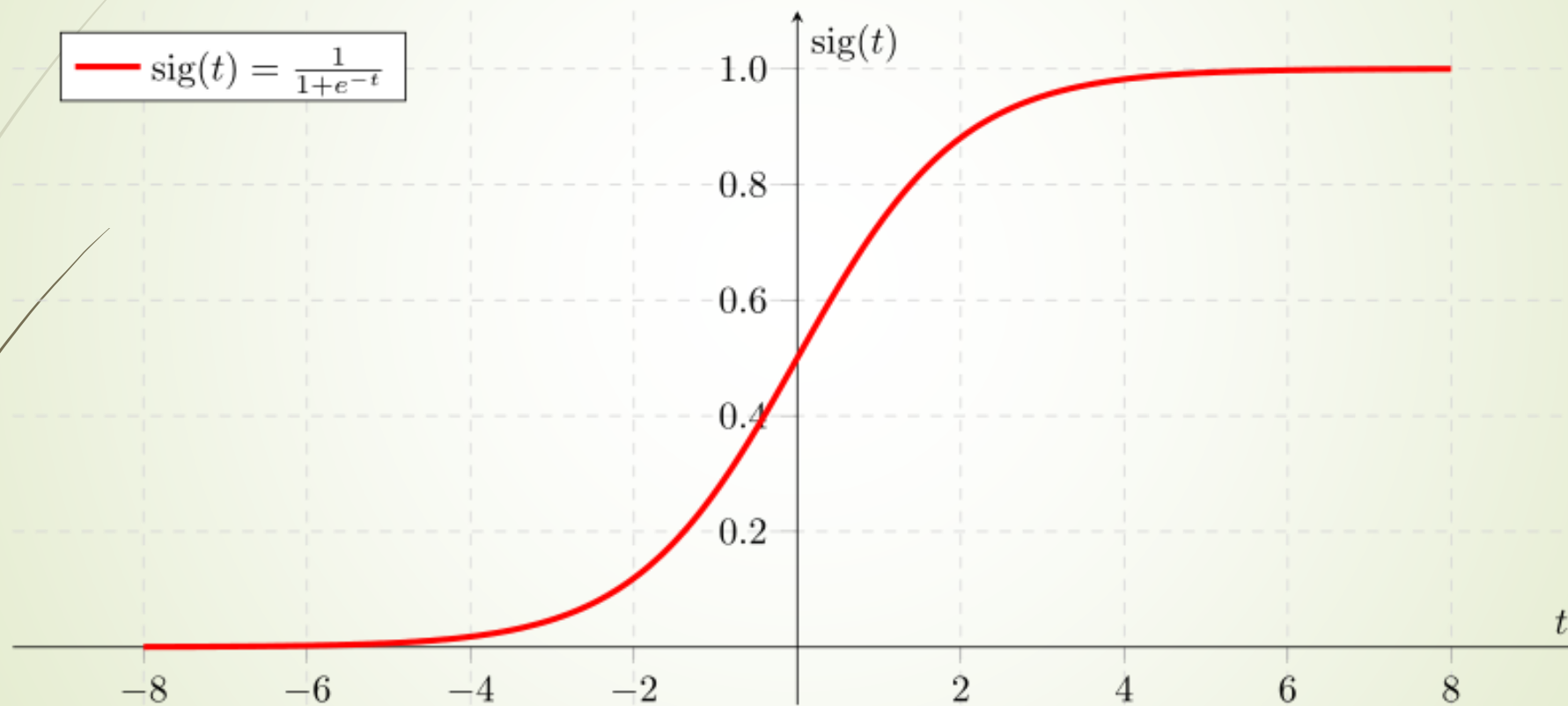
$$S(X) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \dots + \theta_n x_n$$

$$S(X) = \theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \dots + \theta_n x_n$$

$$S(X) = \sum_{i=0}^n \theta_i x_i$$

$$S(X) = \Theta X$$

Un peu de math



Un peu de math

$$\text{Sigmoid}(t) = \frac{1}{1 + e^{-t}}$$

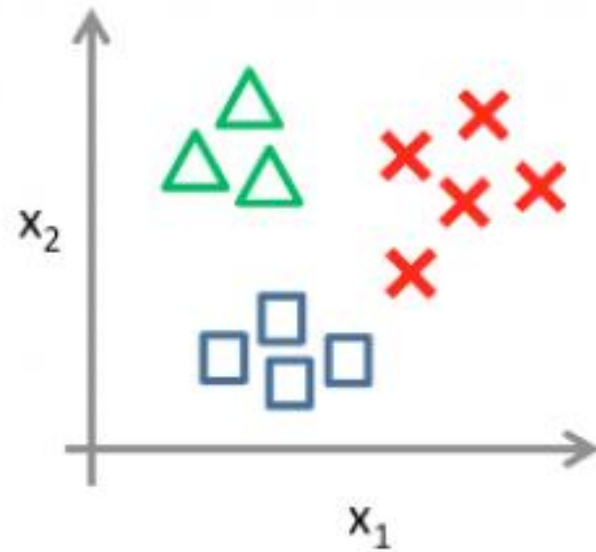
$$H(X) = \text{Sigmoid}(S(X)) = \frac{1}{1 + e^{-\Theta X}}$$




$$H(X) = P(y = 1 \parallel X; \Theta)$$

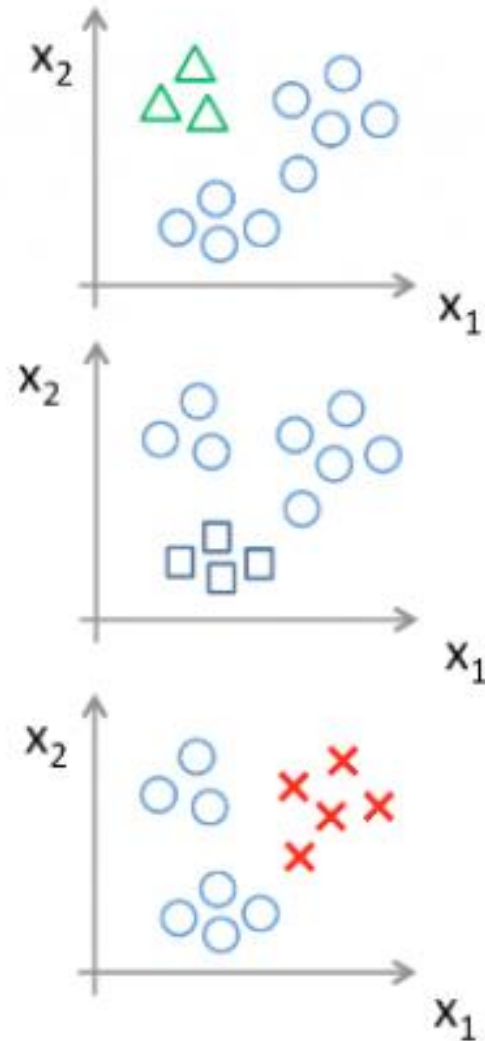
$$P(y = 0 \parallel X; \Theta) = 1 - P(y = 1 \parallel X; \Theta)$$

Un peu de math

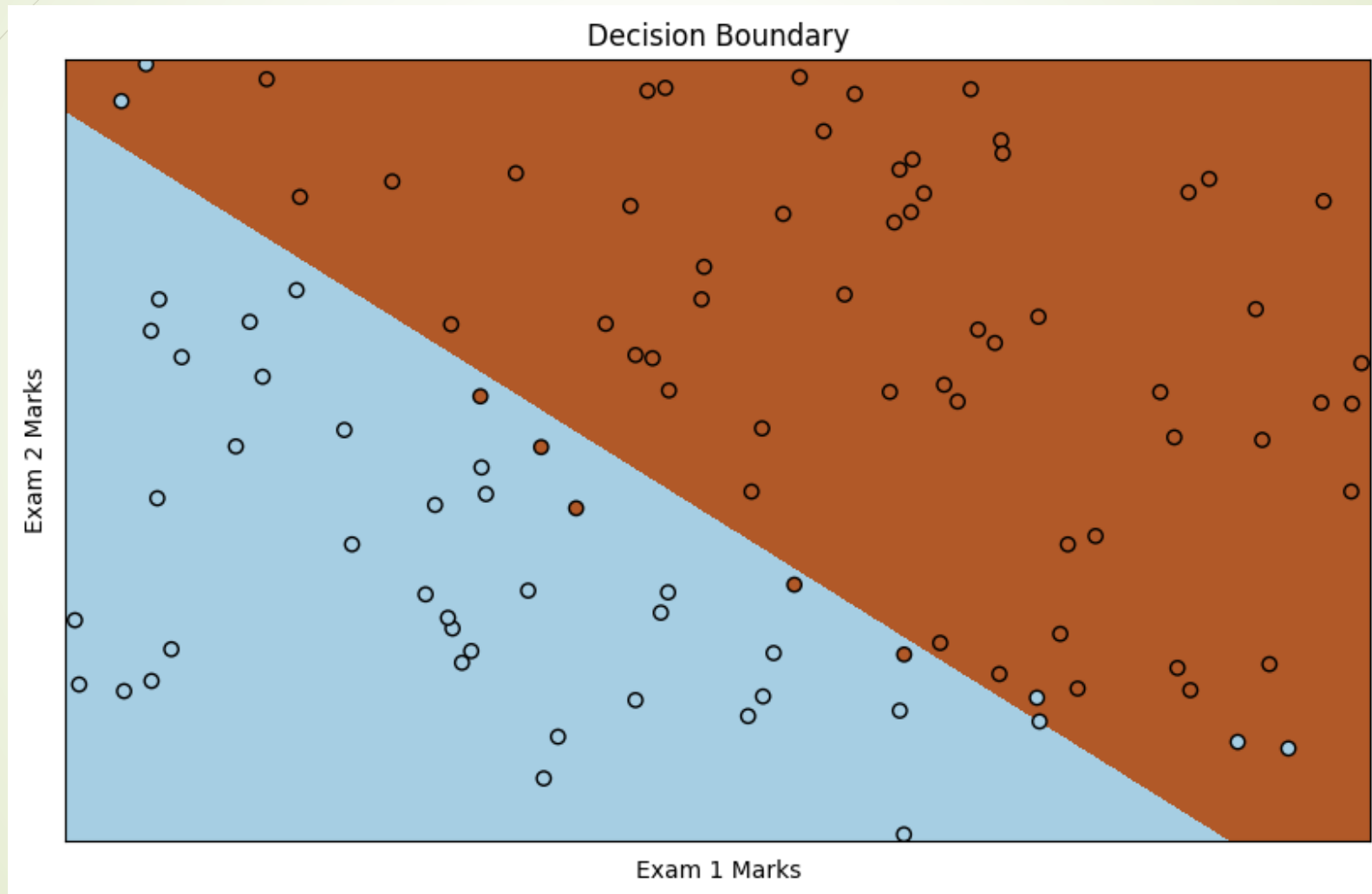
One-vs-all (one-vs-rest):



Class 1: 
Class 2: 
Class 3: 



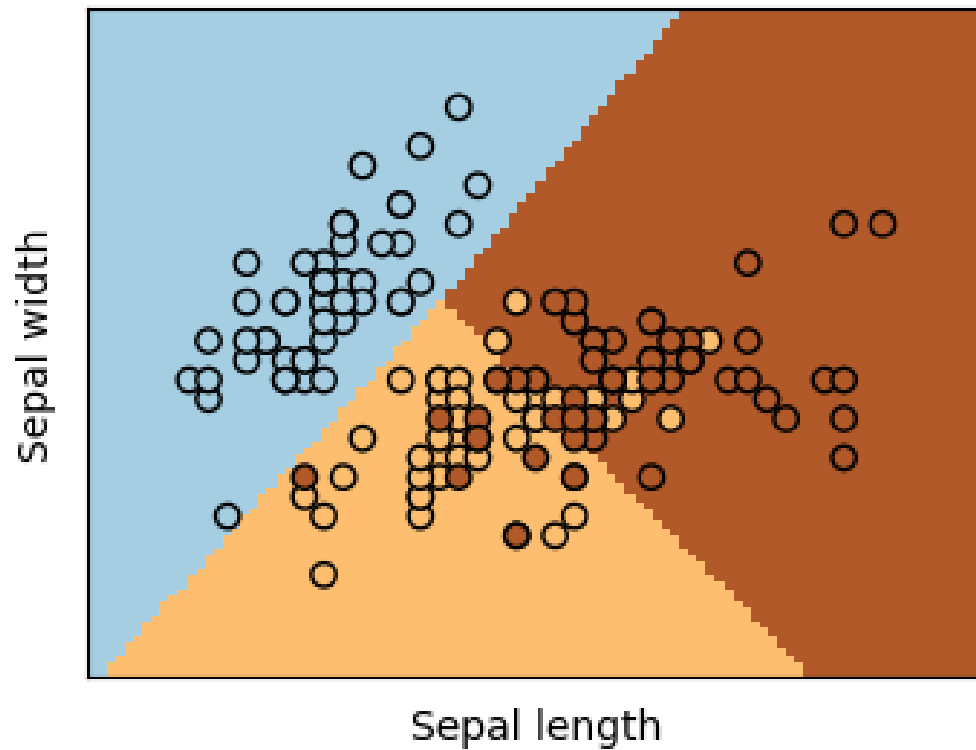
Exemples de Logistic Regression



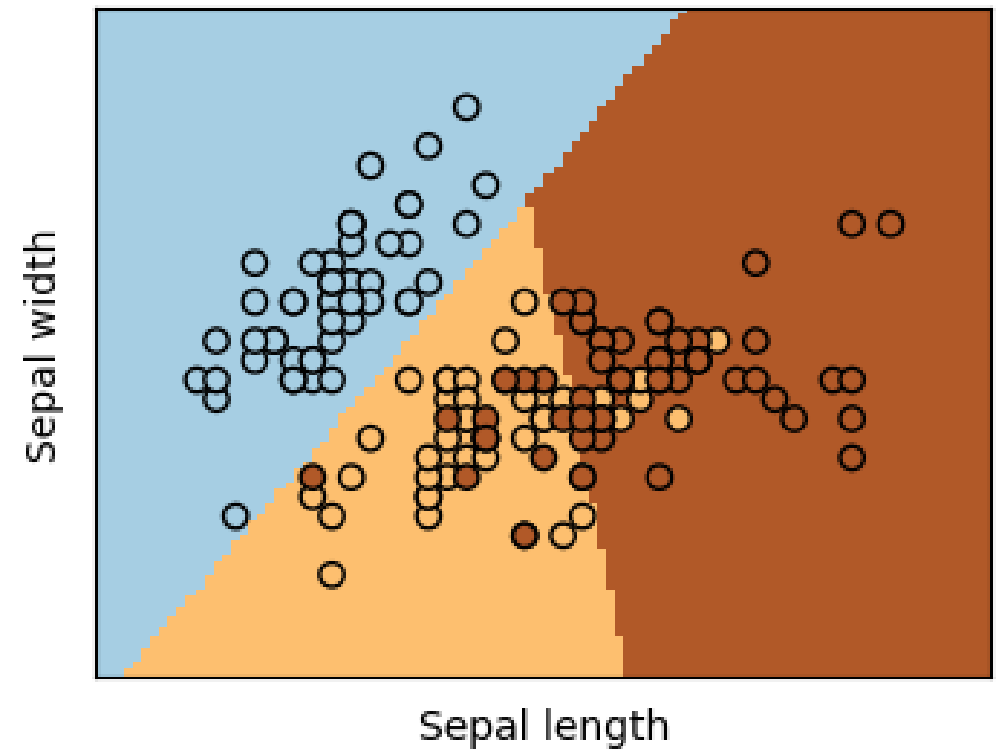
Crédit pour le code : **Satish Gunjal**

Exemples de Logistic Regression

Logistic Regression OVR



Logistic Regression multinomial



Avantages

- Extrêmement efficace
- Faible besoin puissance de calcul
- Simple à régulariser
- Pas de mise à l'échelle nécessaire

Inconvénients

- Utile que pour les problèmes linéaires
- Algorithmes alternatives sont plus puissant
- Vulnérable à l'overfitting si trop peu d'observation