



I N S E A





2

5

Regularization

Machine learning classic: zero-to-hero

Tout modèle de machine learning (supervisé) cherche une fonction de prédiction f .

Supposons qu'elle est paramétrée par $\theta \in \mathbb{R}^p$.

Tout modèle de machine learning cherchera à apprendre:

$$\min_{\theta \in \mathbb{R}^p} \frac{1}{n} \sum_{i=1}^n \text{loss}(f_{\theta}(\mathbf{x}_i), y_i)$$

Minimiser l'erreur de prédiction
sur les données “train”

*C*ontrôle la complexité du modèle

$$+ \frac{1}{C} \text{pénalité}(\theta)$$

des paramètres “simples” pour
généraliser à des données
nouvelles test (éviter l’overfitting)

La fonction ‘pénalité’ est aussi appelée ‘régularisation’: elle vient simplifier (régulariser) la fonction de prédiction

Ce type de régularisation (+ pénalité) est dit: régularisation de Tikhonov

Comment choisir la pénalité ?

Comment choisir C ?

Tout modèle de machine learning (supervisé) cherche une fonction de prédiction f .

Supposons qu'elle est paramétrée par $\theta \in \mathbb{R}^p$.

Tout modèle de machine learning cherche un compromis entre:

$$\min_{\theta \in \mathbb{R}^p} \frac{1}{n} \sum_{i=1}^n \text{loss}(f_{\theta}(\mathbf{x}_i), y_i) \quad + \quad \frac{1}{C} \text{pénalité}(\theta)$$

Minimiser l'erreur de prédiction
sur les données "train"

des paramètres "simples" pour
généraliser à des données
nouvelles test (éviter l'overfitting)

C contrôle la complexité du modèle

La fonction "pénalité" est aussi appelée "régularisation": elle vient simplifier (régulariser) la fonction de prédiction

Ce type de régularisation (+ pénalité) est dit: régularisation de Tikhonov

Comment choisir la pénalité ?

Comment choisir C ?



Comment choisir la pénalité ?

