

Definitions

Hazar HAMOUDA - Mohamed MEGDICHE

30 avril 2025

1 Fonction score s

On définit tout d'abord les objets mathématiques qui nous seront utiles dans l'énoncé du théorème de la prédiction conforme :

Une *fonction de score* (ou fonction de non-conformité) est une application

$$s : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$$

qui mesure à quel point une paire (x, y) est atypique ou en désaccord avec les données d'entraînement ou le modèle prédictif.

Un score élevé $s(x, y)$ indique que la valeur y est peu plausible compte tenu de x , selon le modèle ou une heuristique choisie. Cette fonction est utilisée pour comparer de nouveaux exemples avec les exemples de calibration, indépendamment de la distribution sous-jacente.

2 Le quantile \hat{q}

avec \hat{q} comme le quantile d'ordre $\left\lceil \frac{(n+1)(1-\alpha)}{n} \right\rceil$ des scores de calibration $s_1 = s(X_1, Y_1), \dots, s_n = s(X_n, Y_n)$.

3 Ensemble de prédiction conforme $\hat{C}_n(x)$

avec :

$$\hat{C}_n : \mathcal{X} \rightarrow \{\text{sous-ensembles de } \mathcal{Y}\}$$

qui associe à chaque entrée $x \in \mathcal{X}$ un ensemble $\hat{C}_n(x) \subseteq \mathcal{Y}$ représentant les valeurs plausibles de sortie y .

Cet ensemble est défini par :

$$\hat{C}_n(x) = \{y \in \mathcal{Y} : s(x, y) \leq \hat{q}\}$$

Autrement dit, pour une nouvelle observation x , on considère toutes les sorties y dont le score est inférieur ou égal au seuil \hat{q} , déterminé à partir des données de calibration.