Prediction conforme

Hazar HAMOUDA - Mohamed MEGDICHE

2 mai 2025

1 Résultat de la Prédiction Conforme générale

Théorème : Soit pour $i = 1, ..., n$, $(X_i, Y_i) \sim P_{\mathcal{X}\mathcal{Y}}$ une suite de couples de variables aléatoires, indépendates et identiquement distributées. (X_i, Y_i) sont à valeurs dans X*Y. on definit egalement alpha, le niveau d'erreur pour notre prédiction On construit C chapeau tel que :

C chapeau : ksi -> ensemble de partie de Y

ayant pour propriété, pour une nouvelle paire de test (X$_test, Y_test) sim P_k si Y$

C(X$_test) = y appartenant \mathcal{Y}, s(X_test, y) < q$

on a donc probabilité(Y$_t est appartient C chapeau(X_test)) >= 1 - alpha$

On définit alors les scores de cal * q le (n+1)(1)quantile des scores de calibration s1 = s(X1, Y1), ..., sn = s(Xn, Yn)

2 Fonction score s

On definit tout d'abord les objets mathematiques qui nous seront utile dans l'ennonce du theoreme de la prediction conforme :

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

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

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.

3 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 $s_1 = s(X_1, Y_1), \ldots, s_n = s(X_n, Y_n)$.

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

avec :

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

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.

5 application du theoreme de la prédiction conforme

[Prédiction conforme pour la régression linéaire] Soit un ensemble de calibration $(X_1, Y_1), \ldots, (X_n, Y_n)$ constitué de paires i.i.d. selon une distribution quelconque sur $\mathcal{X} \times \mathbb{R}$, et soit $(X_{\text{test}}, Y_{\text{test}})$ un point test indépendant issu de la même distribution.

Soit $\hat{f} : \mathcal{X} \to \mathbb{R}$ un estimateur déterministe (par exemple une régression linéaire entraînée sur un autre jeu de données), et définissons :

— Les scores de non-conformité : $s_i = |Y_i - \hat{f}(X_i)|$ pour $i = 1, \ldots, n$;

— Le quantile empirique conforme :

$$\hat{q} = \text{score au rang } \lceil (n+1)(1-\alpha) \rceil \text{ parmi les } s_i;$$

— L'ensemble prédictif conforme :

$$\hat{C}_n(x) = \left[\hat{f}(x) - \hat{q},\ \hat{f}(x) + \hat{q}\right].$$

Alors, sous l'hypothèse d'échangeabilité des données de calibration et du point test, l'ensemble $\hat{C}_n(x)$ satisfait la propriété de validité marginale :

$$\mathbb{P}\left(Y_{\text{test}} \in \hat{C}_n(X_{\text{test}})\right) \geq 1 - \alpha$$

où la probabilité est prise conjointement sur les données de calibration et le point test.

[Prédiction conforme pour la classification] Soit un ensemble de calibration $\{(X_1, Y_1), \dots, (X_n, Y_n)\}$ composé de paires i.i.d. selon une distribution inconnue sur $\mathcal{X} \times \mathcal{Y}$, où $\mathcal{Y} = \{1, \dots, K\}$ est un ensemble fini de classes.
Soit $(X_{\text{test}}, Y_{\text{test}})$ une paire test indépendante, issue de la même distribution.

Supposons qu'un modèle probabiliste $\hat{f} : \mathcal{X} \to \Delta_K$ ait été entraîné, où Δ_K est le simplexe de dimension $K-1$ (i.e. l'ensemble des vecteurs de probabilités de classe).
On note $\hat{f}(x) = (p_1(x), \dots, p_K(x))$ les probabilités prédites pour chaque classe.

Définissons :

— Le score de non-conformité pour chaque point de calibration :

$$s_i = 1 - \hat{f}_{Y_i}(X_i), \quad \text{pour } i = 1, \dots, n,$$

où $\hat{f}_{Y_i}(X_i)$ est la probabilité attribuée par le modèle à la vraie classe Y_i.

— Le seuil conforme comme le quantile empirique :

$$\hat{q} = \text{score au rang } \lceil (n+1)(1-\alpha) \rceil \text{ parmi les } \{s_i\}_{i=1}^n.$$

— L'ensemble de prédiction conforme :

$$\hat{C}_n(x) = \left\{ y \in \mathcal{Y} : 1 - \hat{f}_y(x) \leq \hat{q} \right\}.$$

Alors, sous l'hypothèse d'échangeabilité des $n+1$ exemples, on a la garantie suivante :

$$\mathbb{P}\left(Y_{\text{test}} \in \hat{C}_n(X_{\text{test}})\right) \geq 1 - \alpha,$$

où la probabilité est prise sur les données de calibration et le point test.

6 Conclusion

deux fcts qui prennent en entrée les models et ensemble de test (1 ere calcule la taille de l'intervalle moyen(pour la regression et la cardinalite pour la classification)(Éeme clacule empiriquement que ds 1-alpha des cas on a bien ce quii est prédit est ds l'intervalle)) on utilise np.quantile (il faut utiliser la bonne method(regarder dcoumentation)) et il faut appliquer prediction conforme sur le model (o caclcule scores sur calibration) intervel x+$_quantileetl'autreonprendceuxduisoftmaxenaqequationaveclequantile$