Prediction conforme

Hazar HAMOUDA - Mohamed MEGDICHE

30 avril 2025

1 Un objectif ambitieux ?

L'objectif fondamental de la prédiction conforme est le suivant. Soit $(X_i, Y_i) \sim P$, pour $i = 1, \dots, n$, une suite de paires i.i.d. (caractéristiques et réponses), issues d'une distribution P sur $\mathcal{X} \times \mathcal{Y}$. Par souci de clarté, on peut supposer que l'espace des caractéristiques est $\mathcal{X} = \mathbb{R}^d$, et que l'espace des réponses est $\mathcal{Y} = \mathbb{R}$, bien que cela ne soit pas nécessaire en général. Soit $\alpha \in (0, 1)$ un niveau d'erreur nominal. L'objectif est de construire une *bande de prédiction* :

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

telle que, pour une nouvelle paire $(X_{n+1}, Y_{n+1}) \sim P$, on ait :

$$\mathbb{P}(Y_{n+1} \in \hat{C}_n(X_{n+1})) \geq 1 - \alpha, \tag{1}$$

où la probabilité est prise sur l'ensemble des données (X_i, Y_i) pour $i = 1, \dots, n+1$.

D'une part, sans faire d'hypothèse sur P ni recourir à des approximations asymptotiques, cela peut sembler être un objectif très difficile à atteindre. D'autre part, on peut facilement construire une procédure triviale qui satisfait cette condition ; par exemple :

$$\hat{C}_n(X_{n+1}) = \begin{cases} \mathcal{Y} & \text{avec une probabilité } 1 - \alpha \\ \emptyset & \text{avec une probabilité } \alpha \end{cases}$$

aurait toujours une couverture exacte de $1 - \alpha$, c'est-à-dire que l'équation (1) serait satisfaite avec égalité.

La véritable question est donc la suivante (même si elle reste encore quelque peu vague) : peut-on satisfaire (1) avec un échantillon fini, sans faire d'hypothèse sur P, et de manière *non triviale* ? En particulier, on souhaiterait que notre stratégie s'adapte à la difficulté du problème : plus il est facile de prédire Y_{n+1} à partir de X_{n+1}, plus l'ensemble $\hat{C}_n(X_{n+1})$ devrait être petit.

2 Prédiction Conforme

Théorème : Soit pour $i = 1, ..., n$, $(X_i, Y_i) \sim P_{\mathcal{X}\mathcal{Y}}$ identiquement distribuées à valeurs dans ksi*Y1. De plus $(X_i, Y_i) Ondfinitgalementalphavaleursdans[0; 1], leniveaud'$ $qle(n+1)(1)quantiledesscoresdecalibrations1 = s(X1, Y1), ..., sn = s(Xn, Yn)$

3 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+$_{q}uantileetl'autreonprendceuxduisoftmaxenaqequationaveclequantile$