

‘Discussion sur les prédicteurs conformes proposés par Alex Gammerman et Vladimir Vovk’

Alexey Chervonenkis

Research Institute of Control Problems, Russian Academy of Sciences
Computer Learning Research Centre,
Royal Holloway, University of London, UK
chervnks@ipu.rssi.ru

Mots clés : apprentissage automatique, prédicteurs conformes, complexité de Kolmogorov, approche bayésienne, étrangeté d’une prédiction

Abstract: Conformer predictors approach seems to be new and powerful. Its main advantage is that it is non-parametric and based only on the i.i.d. assumption. In comparison to the Bayesian approach, no prior distribution is used. The main theoretical result is the proof of validity of proposed conformal predictors. The second result is that asymptotically the relative number of cases when the real output value is within confidence interval converges to the average value of conformal predictors. The proposed technique is now applied to a large variety of practical problems. Two drawbacks of the approach are still mentioned in this discussion

Keywords: Machine Learning, conformer predictor, Kolmogorov complexity, Bayesian approach, prediction strangeness

Il y a aujourd’hui de nombreux algorithmes d’Apprentissage Automatique très diversifiés qui sont développés et appliqués dans les différents domaines scientifiques et industriels. Mais cette nouvelle approche comportait jusqu’à présent un certain inconvénient : on ne peut calculer un degré de confiance à accorder à la prédiction d’une valeur pour les nouveaux objets.

L’idée principale de l’article des professeurs Alex Gammerman et Vladimir Vovk est de considérer toutes les étiquettes possibles pour un nouvel élément et d’évaluer l’étrangeté de chaque appellation en comparaison de celles des objets de l’ensemble d’apprentissage. Tout le problème est de trouver une mesure intéressante pour cette ‘étrangeté’. Dans un premier temps les auteurs tentent d’appliquer les idées de la complexité selon Kolmogorov pour estimer l’étrangeté des étiquettes. Mais la complexité n’est d’une part pas calculable, de plus elle n’est définie qu’à une constante près, et enfin elle n’a de sens que pour l’entière séquence des objets, et non pas pour l’un d’entre eux en particulier.

Les auteurs proposent alors une autre idée (toujours en liaison avec la complexité chez Kolmogorov) pour estimer cette étrangeté des étiquettes. C’est pour chaque algorithme particulier d’apprentissage, qu’on envisage de trouver une mesure raisonnable de l’étrangeté de l’étiquette attribuée à un objet. Pour la régression (ou la régression Ridge) cette mesure peut être choisie comme la différence absolue entre le résultat de la régression et le résultat réel : plus grande est la différence, plus ‘étranger’ est l’objet. Dans une approche SVM de reconnaissance de formes, cela