

# Apprentissage par renforcement de stratégies d'apprentissage actif : une évaluation

Louis Desreumaux\*, Vincent Lemaire\*\*

\*Université de Technologie de Compiègne, \*\*Orange Labs

**Résumé.** L'apprentissage actif est le domaine qui regroupe l'ensemble des stratégies permettant de construire itérativement l'ensemble d'apprentissage d'un modèle en interaction avec un expert humain, l'objectif étant de sélectionner les exemples les plus informatifs pour minimiser le coût d'étiquetage. Bien que ce domaine soit assez ancien, la mise en œuvre de stratégies d'apprentissage actif en contexte industriel est encore difficile. En particulier, il n'existe pas de stratégie dans l'état de l'art qui soit meilleure que les autres dans tous les contextes. Dans cet article, nous évaluons une méthode représentative d'une orientation récente de la recherche en apprentissage actif vers des méthodes de méta-apprentissage permettant d'apprendre de nouvelles stratégies d'apprentissage actif (Konyushkova et al., 2019). Nous comparons notamment les performances d'une stratégie apprise *via* cette méthode avec celles d'une stratégie d'échantillonnage basé sur l'incertitude, lauréate de plusieurs benchmarks (Yang et Loog, 2018; Pereira-Santos et al., 2019).

## 1 Introduction

L'apprentissage actif regroupe l'ensemble des stratégies de sélection d'exemples permettant de construire itérativement l'ensemble d'apprentissage d'un modèle en interaction avec un expert humain, aussi appelé oracle. Il s'agit de sélectionner les exemples les plus informatifs pour minimiser le coût d'étiquetage.

Dans cet article, nous nous plaçons dans le cadre de l'échantillonnage sélectif, dans lequel les stratégies manipulent un ensemble d'exemples  $\mathcal{D} = \mathcal{L} \cup \mathcal{U}$  de taille constante, où  $\mathcal{L} = \{(\mathbf{x}_i, y_i)\}_{i=1}^l$  est l'ensemble des exemples étiquetés et  $\mathcal{U} = \{\mathbf{x}_i\}_{i=l+1}^n$  est l'ensemble des exemples non étiquetés. Dans ce cadre, l'apprentissage actif est un processus itératif qui s'exécute jusqu'à l'épuisement d'un budget d'étiquetage ou l'atteinte d'un seuil de performance prédéfini. Chaque itération débute par la sélection de l'exemple le plus informatif  $\mathbf{x}^* \in \mathcal{U}$ . Cette sélection s'effectue généralement à partir des informations collectées lors des itérations précédentes (prédictions d'un classifieur, mesures de densité, etc.). L'exemple  $\mathbf{x}^*$  est alors soumis à l'oracle qui retourne la classe  $y^*$  correspondante, et le couple  $(\mathbf{x}^*, y^*)$  est incorporé dans  $\mathcal{L}$ . Le nouvel ensemble d'apprentissage est alors utilisé pour améliorer le modèle et les nouvelles prédictions sont exploitées à l'itération suivante.

Les mesures d'utilité définies par les stratégies d'apprentissage actif de la littérature (Settles, 2012) diffèrent dans leur positionnement selon un dilemme entre l'exploitation du classifieur courant et l'exploration des données d'apprentissage. La sélection d'un exemple non