

# When PAC-Bayesian Majority Vote meets Domain Adaptation

Emilie Morvant

April 29, 2019

## Résumé

De nos jours, de nombreuses données sont disponibles et de nombreuses applications requièrent l'utilisation de méthodes d'apprentissage automatique supervisées, capables de tirer avantage de différentes sources d'information. Une solution naturelle consiste à "combiner" ces sources. Nous nous concentrons ici sur une combinaison particulière : le vote à la majorité pondérée "PAC-Bayésien". C'est une méthode ensembliste, fondée théoriquement, où plusieurs modèles se voient attribuer un poids spécifique. Cette approche est motivée par l'idée qu'une combinaison peut potentiellement compenser les erreurs des modèles individuels et ainsi obtenir une meilleure robustesse et performance sur de nouvelles données tirées selon une certaine distribution de probabilité. En apprentissage statistique, la capacité d'un modèle à généraliser sur une telle distribution de données est mesurée par des bornes de généralisation. Dans cet exposé, après avoir rappelé le cadre théorique classique des bornes en généralisation PAC-Bayésiennes, nous l'étendons à la tâche d'adaptation de domaine où l'objectif est d'adapter un modèle à partir d'une distribution de données source à une distribution cible différente, mais liée.