

Sujet de Stage Recherche

Master 2 Génie Informatique

Codes Correcteurs d'Erreurs et Combinaison de Classifieurs Binaires

Laurent HEUTTE
PSI – FRE CNRS 2645

Contexte

Les codes correcteurs d'erreurs (en anglais ECOC pour Error-Correcting Output Codes) sont utilisés depuis quelques années pour concevoir des ensembles de classifieurs divers. L'idée de base derrière l'utilisation de ces méthodes est d'éviter de concevoir directement un classifieur multi-classes en décomposant le problème de classification en un ensemble de classifieurs binaires (fonctions discriminantes). Chaque classifieur dans l'ensemble est utilisé pour discriminer deux classes, ces classes pouvant être des regroupements des classes d'origine. Les performances de ce type de méthodes dépendent essentiellement de la façon dont on construit le code correcteur. Un certain nombre de méthodes existent pour construire un bon ECOC (codes exhaustifs, sélection, optimisation, BCH). Toutefois, la conception d'un code optimal pour un problème de classification donné reste un question ouverte.

Etude proposée

Le sujet proposé s'inscrit dans ce cadre de la combinaison de classifieurs binaires par ECOC. Le travail du stagiaire consistera en l'étude, la mise en œuvre et l'évaluation de quelques procédures de construction d'ECOC. Le travail débutera par une analyse bibliographique des approches récentes proposées dans ce domaine. De cette analyse, le stagiaire devra dégager les spécificités des différentes méthodes de construction pour ne retenir et ne développer que les plus robustes selon des critères qu'il reste à déterminer. Plus particulièrement, l'étude portera sur le rôle que joue la diversité des sorties d'ECOC sur les performances des méthodes de combinaison. Pour évaluer la robustesse des différentes solutions retenues, le stagiaire disposera de différentes bases de données (chiffres manuscrits, caractères manuscrits, classes morphologiques, ...) sur lesquelles seront testées les classifieurs binaires (SVM). Les solutions retenues devront être développées en C/C++ ou sous MATLAB.

Lieu

Le travail se déroulera au sein du laboratoire PSI – FRE CNRS 2645 à l'UFR des Sciences, Département de Physique, 1^{er} étage.

Contact

Laurent HEUTTE; Bureau EP104, Bâtiment de Physique, UFR Sciences

tel : 02.35.14.68.77 – email: Laurent.Heutte@univ-rouen.fr - url: <http://www.univ-rouen.fr/psi/heutte>