

CONE : Un algorithme d'optimisation de la F-Mesure par pondération des erreurs de classification

Kévin Bascol^{1,2}, Rémi Emonet², Elisa Fromont³, Amaury Habrard², Guillaume Metzler^{2,4},
and Marc Sebban²

¹Bluecime inc., France

²Univ Lyon, UJM-Saint-Etienne, CNRS, Institut d'Optique Graduate School, Laboratoire Hubert Curien UMR 5516, F-42023, Saint-Etienne, France

³Univ. Rennes 1, IRISA/Inria, 35042 Rennes cedex, France

⁴Blitz inc., France

9 novembre 2018

Résumé

Nous proposons un algorithme d'optimisation de la F-Mesure avec des garanties théoriques utilisable avec toute méthode d'apprentissage par pondération des erreurs. **CONE**, notre algorithme, génère itérativement un ensemble de coûts à partir de l'ensemble d'entraînement de telle sorte que le classifieur final ait une F-Mesure proche de l'optimale. L'optimalité de la F-Mesure obtenue est exprimée à l'aide d'une borne supérieure plus fine que celle présentée dans [Parambath et al. 2014] De plus, nous montrons que les coûts obtenus à chaque itération de **CONE** permettent de réduire drastiquement l'espace de recherche et ainsi de converger rapidement vers les paramètres optimaux. L'efficacité de notre méthode est montrée à la fois en terme de F-Mesure mais aussi de vitesse de convergence sur plusieurs jeux de données déséquilibrés.

Mots-clef : Algorithm, Supervised Learning, Theory.