

Programme de cotutelles U. Libanaise - UT INSA

Description du sujet (merci de vous conformer aux recommandations indiquées sur le site web)

Nom : Prénom :

Fonction (prof., MdC) :

Laboratoire : Adresse web :

Etabliss^t : Adresse web :

Compétence scientifique:

Analyse de données, Apprentissage, Classification multi-label.

2 publications importantes en relation avec le sujet proposé :

- Omrani H., Abdallah F., Charif O., Longford N. Multi-label class assignment in land-use modelling. International Journal of Geographical Information Science, Vol. 29, 6, 2015. A23

-Sawsan Kanj, Fahed Abdallah, Thierry Denoeux, Kifah Tout, Editing training data for multi-label classification with the k-nearest neighbor rule, In Pattern Analysis and Applications, Springer Verlag, 2016, 19 (1), pp.145-161

Adresse web de votre page personnelle :

Adresse mail :

Description du sujet de thèse proposé n° du thème :

Titre :

Sujet :

La classification multi-label est un concept en apprentissage automatique qui diffère de celui de la classification habituelle par le fait qu'une observation peut appartenir simultanément à plusieurs classes. Le problème de classification multi-label devient de plus en plus répandu grâce à de nombreuses applications modernes telles que la classification génomique fonctionnelle, le diagnostic médical, l'indexation de documents et la catégorisation de texte, la classification d'images, ou de voix et musiques. Le but du travail de thèse demandé est d'analyser et développer des méthodes d'apprentissage multi-label en prenant en compte la géométrie des observations en fonction de leurs étiquettes multi-label. Il s'agit d'utiliser des concepts développés dans le champ des règles d'association pour trouver une meilleure association entre les différentes classes multi-label. Les algorithmes développés doivent faire face au problème des classes déséquilibrées lorsque les classes ont des proportions significativement différentes dans la base d'apprentissage. La suite des travaux consiste à analyser et à implémenter les méthodes d'apprentissage multi-label dans le cas d'une absence partielle ou d'une incertitude d'étiquetage multi-label.

mots clés :

Sélection de caractéristiques, Règles d'association, Apprentissage, Classification.

Collaborations attendues :

Encadrement Libanais : Fahed Abdallah et Rim Kaddah(UL) + Collaborations avec des collègues de l'UTT

Compétences nécessaires du candidat :

Informatique appliquée, Mathématique appliquée. analyse de données, apprentissage automatique.

Existence d'un fichier pdf détaillant le sujet (oui-non) :

(respecter les indications données sur le site web)

