TP RF 2: LibSVM

Partie 1 : Commandes en ligne

- Vous allez récupérer l'archive TP2 : LibSVM qui est sur Célène.
- Dans un shell windows, vous allez dans le répertoire .\libsvm-3.22\windows
- Essayez svm-toy.exe en modifiant les paramètres c et g pour visualiser leurs effets. Si vous n'êtes pas sous windows, svm-toy existe sous forme d'applet sur le site : https://www.csie.ntu.edu.tw/~cjlin/libsvm/.
- Essayez les exécutables sym-train et sym-predict sur les données iris (déjà reformatées dans iris.app pour l'apprentissage et iris.test pour les tests).
- Comprendre les différents paramètres de sym-train, pour cela vous allez créer un fichier batch (sym.bat par exemple) qui va contenir les 2 lignes de commandes (symtrain et sym-predict) avec des paramètres fixes (noms des fichiers) et des paramètres variables qui seront sur la ligne de commande du batch. Le premier paramètre s'appelle %1, le deuxième %2, etc.

Partie 2: Programmation

- Vous allez utiliser les librairies de LibSVM dans le langage de votre choix (C++ ou Java ou python). Pour C++, il faut inclure les fichiers svm.cpp et svm.h. Pour Java, il faut inclure libsvm.jar.
- L'objectif initial va être de refaire l'équivalent du batch précédent, en utilisant les fonctions sym-train et sym-predict.
- Ensuite on va chercher à trouver automatiquement les 2 principaux paramètres par grid-search (voir guide.pdf).
- Enfin, on va appliquer la procédure de cross-validation pour découper la base d'apprentissage afin de déterminer les paramètres en généralisation.