

## TP8 – SVM

Yann.Esposito@lif.univ-mrs.fr  
2 décembre 2005

☞ DANS CE TP NOUS NOUS INTÉRESSERONS AUX SVM. NOUS ESSAYERONS DE COMPRENDRE L'INFLUENCE DES PARAMÈTRE DE CE CLASSIFIEUR.

## 1 Prise en main

**1.1.** Téléchargez les fichiers `train-lunettes.csv` et `tst-lunettes.arff`. Ces deux échantillons correspondent à un `crop` des images de visages sur mon site, correspondant à la partie autour de l'œil. Lancez le classifieur `NaiveBayes` en utilisant la méthode d'évaluation `Percentage split` à 66%.

(a) Quel est le taux de réussite? Recommencez en utilisant un `Percentage split` à 90%.

(b) Recommencez en utilisant le classifieur `SMO (SVM)` avec un `Percentage split` à 66% puis à 90%. Le taux de réussite est-il à la hauteur de vos espérances?

(c) Recommencez toutes ces expériences en appliquant le filtre `Randomize` aux instances. Les modifications sont-elles visible?

(d) Recommencez toutes les expériences en utilisant l'échantillon de test.

(e) Que pouvez-vous déduire de ces expériences sur le classifieur `SMO`?

## 2 Étude de l'influence des paramètres

**2.1.** Lancez l'`Experimenter` de Weka. Utilisez-le pour étudier l'influence de chaque paramètre (dans les cas où `SMO` est efficace).

(a) Remplissez le tableau suivant

paramètres	influence (+/0/-)	utilité	valeur optimale
buildLogisticModels			
c			
cacheSize			
debug			
epsilon			
exponent			
featureSpaceNormalization			
filterType			
gamma			
lowerOrderTerms			
gamma			
lowerOrderTerms			
numFolds			
randomSeed			
toleranceParameter			
useRBF			

**2.2.** Quels paramètres sont essentiels pour le bon fonctionnement des SVM ?

**2.3.** Quels paramètres sont peu utiles ?

**2.4.** Les tests que vous avez fait sont ils suffisant pour répondre précisément aux questions précédentes ?  
Faites les tests supplémentaires en conséquence.

**2.5.** Les réponses sont-elles identiques ?

