

Analyse, classification et indexation des données: contrôle continu

Classification MAP vs classification linéaire

Archive fournie : `CC.zip` contenant les fichiers `donnees.mat` et `CC.m`.

Les réponses aux questions seront rendues sous la forme d'un court rapport au format de votre choix. Vous rendrez également le code matlab qui aura produit vos résultats. L'ensemble est à envoyer par mail à fabien.baldacci@u-bordeaux.fr à la fin du TP.

Dans ce sujet, on utilise d'une part un classifieur bayésien basé sur le maximum a posteriori et d'autre part les différentes versions de classifieurs linéaires vues en TD (Perceptron, Moindres carrés sans optimisation de \mathbf{b}). L'objectif est de mesurer et d'expliquer les différences de résultats entre les classifications sur un jeu de données. Vous utiliserez 10% des données pour construire vos modèles et le reste pour les tests. Les erreurs calculées seront données en pourcentages.

Le script `CC.m` remplit deux tableaux avec les données de deux classes `C1` et `C2` (contenues dans le fichier `donnees.mat`). Il permet également de visualiser ces données.

Exercice 1.

1. Calculez les résultats de classification MAP sur les classes `C1` et `C2` : erreur moyenne sur la classe `C1`, erreur moyenne sur la classe `C2`, erreur moyenne totale. Fournissez une représentation graphique des résultats obtenus et expliquez les.
2. Calculez les résultats de classification en utilisant au choix (ce choix devra être justifié) le perceptron ou le classifieur linéaire basé sur les moindres carrés (tous les coefficients de \mathbf{b} fixés à 1) sur les classes `C1` et `C2` : erreur moyenne sur la classe `C1`, erreur moyenne sur la classe `C2`, erreur moyenne totale. Fournissez une représentation graphique des résultats obtenus et comparez les avec les résultats précédents.
3. Refaire le travail en réduisant les données en 2D en ne conservant que les 2 descripteurs les plus discriminants (et en justifiant le choix)
4. Refaire le travail en réduisant les données en 2D en utilisant une des méthodes de réduction en dimension vue (et en justifiant le choix)

Exercice 2.

Refaire les mêmes questions que pour l'exercice 1 mais en n'utilisant que 10% des données.