

Analyse, classification et indexation des données: contrôle continu 2

Classification MAP vs classification linéaire

Archive fournie : `CC2.zip` contenant les fichiers `donnees1.mat`, `donnees2.mat` et `CC2.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` avant le 6/12.

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 avec et sans optimisation de \mathbf{b}). L'objectif est de mesurer et d'expliquer les différences de résultats entre les classifications sur deux jeux de données. Vous utiliserez 10% des données pour construire vos modèles et le reste pour les tests. Les erreurs calculées seront données en pourcentages.

Comme dans la première évaluation le script `CC2.m` remplit deux tableaux avec les données de deux classes `C1` et `C2` (contenues dans le fichier `donnees1.mat`). Il permet également de visualiser ces données. Le fichier `donnees2.mat` contient deux autres classes `C3` et `C4`.

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. Oui, c'est le même travail que pour le premier contrôle continu !
2. Calculez les résultats de classification avec le perceptron 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. Calculez les résultats de classification avec 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.
4. Calculez les résultats de classification avec le classifieur linéaire basé sur les moindres carrés (avec optimisation de \mathbf{b}) 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.

Exercice 2.

Mêmes questions en utilisant cette fois les classes `C3` et `C4` contenues dans le fichier `donnees2.mat`. Comparez les résultats obtenus sur ce second jeu de données avec ceux obtenus sur le premier jeu de données.