

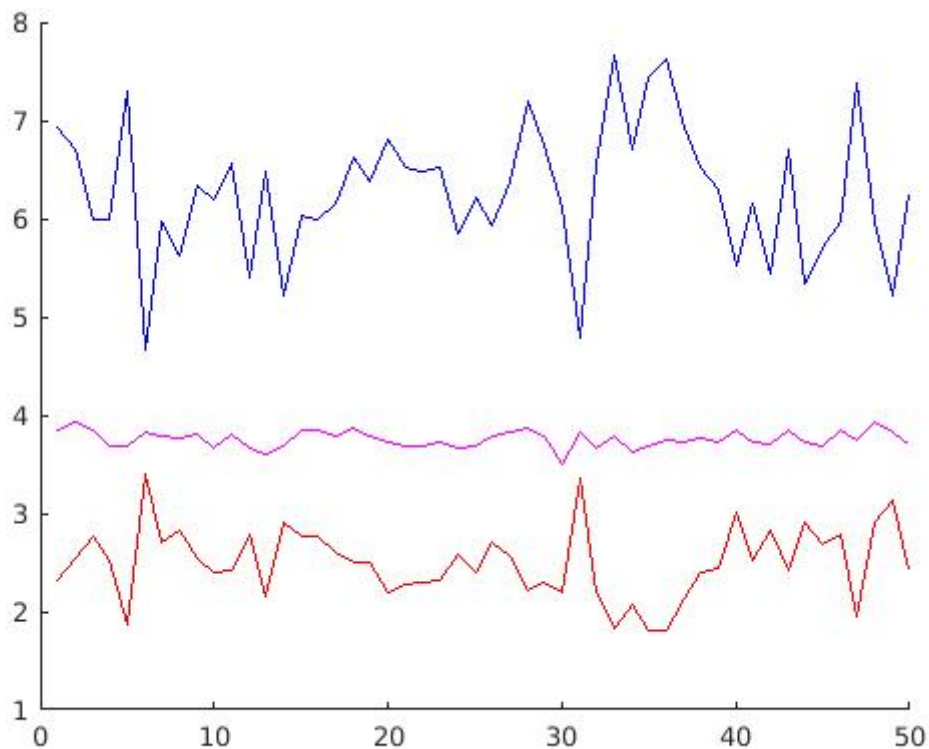
Analyse, classification et indexation des données

CC2 - Rapport

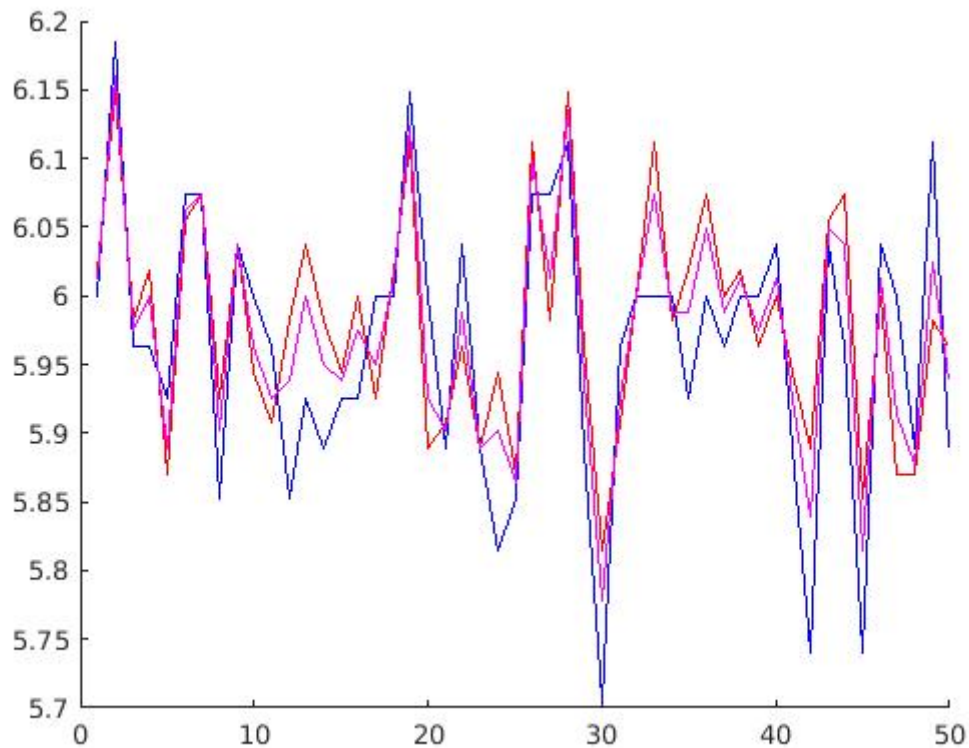
Les résultats suivants sont obtenus au bout de 50 itérations sur des échantillons aléatoires générés à partir des classes données et sont affichés en pourcentage d'erreurs.

Exercice 1 :

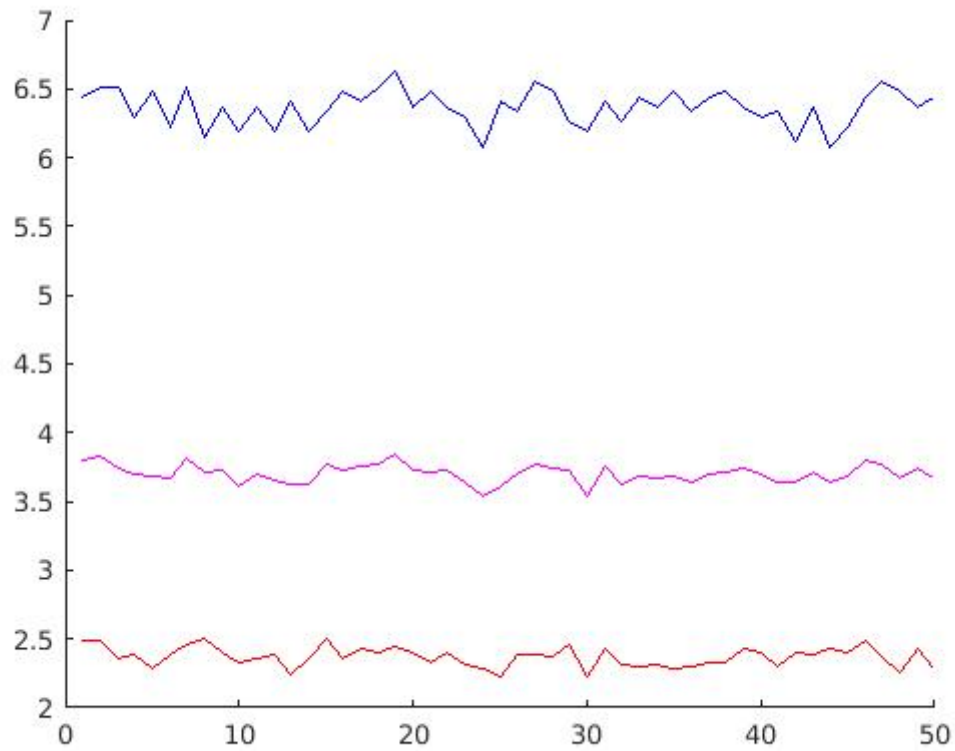
1) On effectue une classification par maximum a posteriori sur C1 et C2. On observe un taux d'erreur de **2.50 % sur C1**, de **6.27 % sur C2** et de **3.76 % au total**. Ce taux d'erreur est plutôt faible mais reste non négligable dû à la proximité des deux ensembles. On remarque que C2 ayant plus de points isolés dans la masse de C1 que C1 n'en a dans C2, il a donc plus d'erreurs.



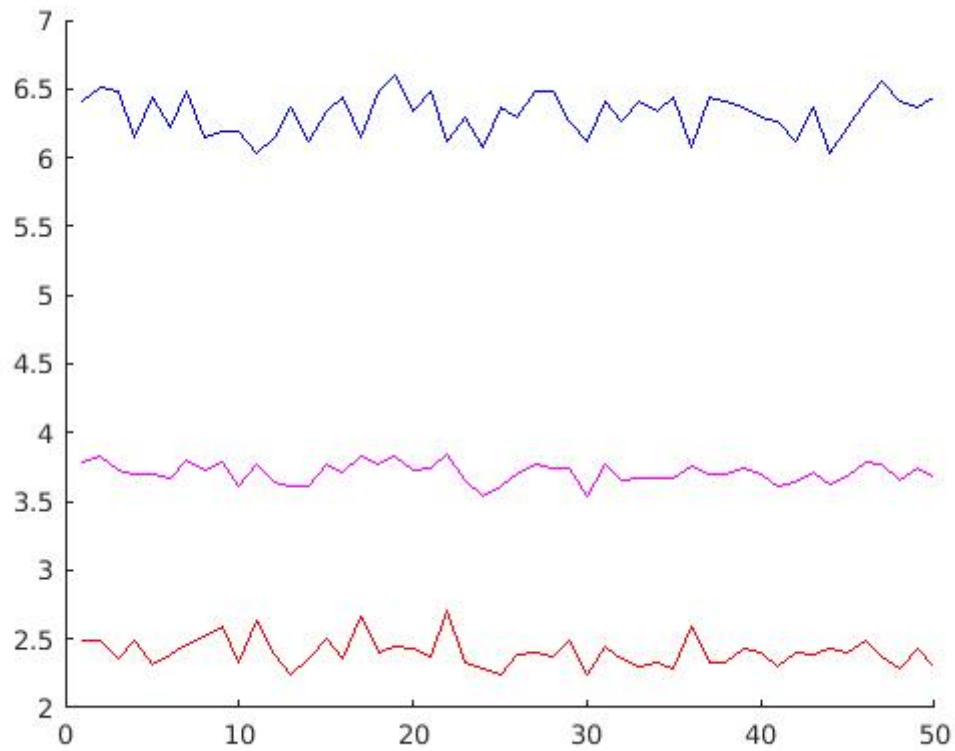
2) On effectue une classification par perceptron sur C1 et C2. On observe un taux d'erreur de **5.98 % sur C1**, de **5.96 % sur C2** et de **5.97 % au total**. On remarque que les taux d'erreurs des deux classes sont très proches. Cela est dû à la nature du perceptron de chercher un plan linéaire qui les sépare de manière équitable. Ces résultats sont cependant plus élevés que ceux obtenus par MAP car le plan de séparation utilisé ne suffit pas à séparer les ensembles de données de manière suffisamment distincte.



3) On effectue une classification basée sur les moindres carrés sur C1 et C2. On observe un taux d'erreur de **2.37 % sur C1**, de **6.37 % sur C2** et de **3.70 % au total**. Ces résultats sont très proches de ceux précédemment obtenus par MAP. C'est la classification qui offre le moins d'erreurs parmi celles testées dans cet exercice.



4) On effectue une classification basée sur les moindres carrés avec optimisation sur C1 et C2. On observe un taux d'erreur de **2.40 % sur C1**, de **6.32 % sur C2** et de **3.71 % au total**. L'optimisation implémentée semble inefficace car le taux d'erreur total a très légèrement augmenté par rapport à la classification par moindres carrés originale.



Exercice 2 :

- 1) On effectue une classification par maximum a posteriori sur C3 et C4. On observe un taux d'erreur de **14.43 % sur C3**, de **3.95 % sur C4** et de **8.14 % au total**.
- 2) On effectue une classification par perceptron sur C3 et C4. On observe un taux d'erreur de **14.35 % sur C3**, de **13.48 % sur C4** et de **13.83 % au total**.
- 3) On effectue une classification basée sur les moindres carrés sur C3 et C4. On observe un taux d'erreur de **23.70 % sur C3**, de **3.87 % sur C4** et de **11.80 % au total**.
- 4) On effectue une classification basée sur les moindres carrés avec optimisation sur C3 et C4. On observe un taux d'erreur de **12.65 % sur C3**, de **16.69 % sur C4** et de **15.07 % au total**.

Tous les résultats obtenus sont supérieurs à ceux sur C1 et C2 car les données des classes sont beaucoup plus mélangées.

