
Manipulation des modèles

1. Télécharger les deux fichiers "crx.names" et "crx.data" présents sur la page "<http://archive.ics.uci.edu/ml/datasets/Credit+Approval>"
2. Subdiviser le fichier "crx.data" en trois fichiers :
 - (a) Un fichier "train" contenant les 500 premiers exemples,
 - (b) Un fichier "test" contenant les 150 exemples suivants,
 - (c) Un fichier "use" contenant les 40 derniers exemples avec l'attribut classe à "?"
3. Construire un modèle Bayésien naïf en l'entraînant sur le fichier "train" et en le testant sur le même fichier. Donner le taux de reconnaissance et la matrice de confusion.
4. Construire un modèle Bayésien naïf en l'entraînant sur le fichier "train" et en le testant par une validation croisée de 5 parties. Donner le taux de reconnaissance et la matrice de confusion.
5. Construire un modèle Bayésien naïf en l'entraînant sur le fichier "train" et en le testant sur le fichier "test". Donner le taux de reconnaissance et la matrice de confusion.
6. Comparer les résultats des trois modèles.
7. Construire un modèle par arbre de décision en utilisant la méthode J48 en l'entraînant sur le fichier "train" et en le testant sur le fichier "test"
8. Donner l'arbre construit et son taux de reconnaissance.
9. Enregistrer le modèle construit.
10. Utiliser ce modèle pour retrouver les classes des exemples du fichier "use".
11. Construire un modèle par vote majoritaire combinant les classifieurs bayésien naïf, J48 et ZeroR. Le modèle est testé par une validation croisée de 8 parties.
12. Donner son taux de reconnaissance.

Responsable de la matière :
Dr A.Djeffal
