Table des figures

1 Processus d'Extraction des Connaissances à partir des Données. xii

1.1 Illustration d'une ACP sur des données en dimension 2 8
1.2 Exemple de graphe des voisins relatifs dans $\mathbb{R}^2$. 11
1.3 Exemple de graphe de Gabriel dans $\mathbb{R}^2$. 12
1.4 Schéma d'un X-arbre libre valué 18
1.5 Analyse arborée selon le vocabulaire du corpus cohabitation Chirac-Jospin (1997-2002) 19
1.6 Illustration des étapes du *neighbor joining* 20
1.7 Volume d'une hypersphère de rayon 1 en fonction du nombre de dimensions 25
1.8 Rapport entre les volumes d'une hypersphère de rayon $0,9$ et d'une hypersphère de rayon 1 en fonction du nombre de dimensions 26
1.9 Densité de probabilité d'une gaussienne 29
1.10 Probabilités relatives des distances inter-points 30

2.1 Illustration de la rétinotopie 38
2.2 (a) Hypercolonne du cortex visuel et (b) colonne d'orientation 39
2.3 Fonction de voisinage gaussienne des cartes auto-organisatrices 41
2.4 Pavage de Voronoï d'une carte en dimension 2 42
2.5 Région d'influence d'un modèle dans la version par lot des cartes auto-organisatrices 46
2.6 Illustration des phases d'apprentissage et de prédiction du modèle LASSO . 54
2.7 Arêtes coupées sur une carte étiquetée avec deux classes. 64
2.8 Projection par la méthode de Sammon d'une carte apprise sur les données *optical digit* 66

2.9 Visualisation du nombre d'exemples représentés par chaque modèle d'une carte apprise sur les données *optical digit* 68

2.10 Taux d'erreur en généralisation en fonction de $2J''/S_0$ et de $2J'''/S_0$ 74

3.1 Régions d'erreur associées aux approximations des probabilités *a posteriori*. 87

3.2 Typologie des méthodes ensemblistes. 94

3.3 Les différentes étapes de l'optimisation d'un comité par un AG. 99

3.4 Taux d'erreur des comités de cartes supervisés et non-supervisés comparés à ceux d'une carte seule apprise sur la totalité des attributs, du boosting et des forêts aléatoires. 105

3.5 Taux d'erreur d'une carte synthétique comparé à celui d'une carte seule apprise sur la totalité des attributs, d'un comité de cartes et à la moyenne des taux d'erreur des cartes du comité. 105

3.6 Taux d'erreur en fonction du niveau de bruit sur 4 jeux de l'UCI. 109