

G. Dreyfus, J.-M. Martinez, M. Samuelides  
M. B. Gordon, F. Badran, S. Thiria  
Sous la direction de Gérard Dreyfus

# A Apprentissage statistique

Réseaux de neurones • Cartes topologiques  
Machines à vecteurs supports



- Cinq exemples de modèles, avec données et code source
    - Neuro One 6.10.7\*, outil de création de modèles neuronaux
    - Compilateur C pour Windows.
    - Bibliothèque non linéaire MonaEx70.dll, niveau 0.
- \* Version d'évaluation de 6 semaines pour MS-Windows NT, 2000, 2003, XP

- Prévision
- Data mining
- Bio-ingénierie
- Reconnaissance de formes
- Robotique et commande de processus

EYROLLES

---

<b>Avant-propos et guide de lecture</b>	XI
Guide de lecture	XII
Détail des contributions	XIII
<b>1 L'apprentissage statistique : pourquoi, comment ?</b>	1
Introduction	1
<b>Premier exemple : un problème élémentaire d'apprentissage statistique</b>	2
Point de vue algorithmique	3
Point de vue statistique	4
<b>Quelques définitions concernant les modèles</b>	5
Modèles statiques	5
Modèles dynamiques	6
<b>Deux exemples académiques d'apprentissage supervisé</b>	7
Un exemple de modélisation pour la prédiction	7
Un exemple de classification	11
Conclusion	16
<b>Éléments de théorie de l'apprentissage</b>	16
Fonction de perte, erreur de prédiction théorique	17
Dilemme biais-variance	22
<b>De la théorie à la pratique</b>	25
Remplacer des intégrales par des sommes	26
Bornes sur l'erreur de généralisation	27
Minimisation du risque structurel	30
<b>Conception de modèles en pratique</b>	30
<b>Collecte et prétraitement des données</b>	30
Les données sont préexistantes	30
Les données peuvent être spécifiées par le concepteur	30
Prétraitement des données	31
Sélection des variables	31
Apprentissage des modèles	32
Sélection de modèles	32
<b>Sélection de modèles</b>	32
Validation simple (hold-out)	32

Validation croisée (« cross-validation »)	33
Leave-one-out	34
<b>Sélection de variables</b>	35
Cadre théorique	36
Méthode de la variable sonde	37
<b>Résumé : stratégies de conception</b>	47
<b>Conception de modèles linéaires par rapport à leurs paramètres (régression linéaire)</b>	48
Sélection de variables pour les modèles linéaires en leurs paramètres	48
Apprentissage de modèles linéaires en leurs paramètres : la méthode des moindres carrés	49
Propriétés de la solution des moindres carrés	51
Estimation de la qualité de l'apprentissage	52
Interprétation géométrique	53
Dilemme biais-variance pour les modèles linéaires	54
Sélection de modèles linéaires	56
Moindres carrés par orthogonalisation de Gram-Schmidt	59
<b>Éléments de statistiques</b>	60
Qu'est-ce qu'une variable aléatoire ?	60
Espérance mathématique d'une variable aléatoire	62
Estimateur non biaisé	63
Variance d'une variable aléatoire	64
Autres distributions utiles	65
Intervalles de confiance	66
Tests d'hypothèse	68
<b>Conclusion</b>	70
<b>Bibliographie</b>	70

## 2 Les réseaux de neurones 73

---

<b>Introduction</b>	73
<b>Réseaux de neurones : définitions et propriétés</b>	73
Les neurones	74
Les réseaux de neurones	75
Propriété fondamentale des réseaux de neurones statiques (non bouclés) : l'approximation parcimonieuse	82
À quoi servent les réseaux de neurones non bouclés à apprentissage supervisé ?	
Modélisation statique et discrimination (classification)	84
À quoi servent les réseaux de neurones à apprentissage non supervisé ? Analyse et visualisation de données	87

À quoi servent les réseaux de neurones bouclés à apprentissage supervisé ? Modélisation dynamique « boîte noire » et « semi-physique » ; commande de processus	87
<b>Quand et comment mettre en œuvre des réseaux de neurones à apprentissage supervisé ?</b>	88
Quand utiliser les réseaux de neurones ?	88
Comment mettre en œuvre les réseaux de neurones ?	89
Conclusion	93
<b>Réseaux de neurones à apprentissage supervisé et discrimination (classification)</b>	93
Quand est-il opportun d'utiliser un classifieur statistique ?	93
Classification statistique et formule de Bayes	95
Classification et régression	96
<b>Modélisation et classification de données structurées : les « graph machines »</b>	103
Définitions	104
Apprentissage	105
Deux exemples académiques	106
<b>Exemples d'applications</b>	107
Introduction	107
Reconnaissance de formes : la lecture automatique de codes postaux	107
Une application en contrôle non destructif : la détection de défauts dans des rails par courants de Foucault	111
Fouille de données : le filtrage de documents	112
Aide à la découverte de médicaments : prédiction de propriétés chimiques et d'activités thérapeutiques de molécules	116
Une application en formulation : la prédiction de la température de liquidus de verres	118
Modélisation d'un procédé de fabrication : le soudage par points	118
Application en robotique : modélisation de l'actionneur hydraulique d'un bras de robot	121
Modélisation semi-physique d'un procédé manufacturier	122
Contrôle de l'environnement : hydrologie urbaine	123
Une application en robotique mobile : le pilotage automatique d'un véhicule autonome	124
<b>Techniques et méthodologie de conception de modèles statiques (réseaux non bouclés)</b>	125
Sélection des variables	126
Estimation des paramètres (apprentissage) d'un réseau de neurones non bouclé	126

Sélection de modèles	143
<b>Techniques et méthodologie de conception de modèles dynamiques (réseaux bouclés ou récurrents)</b>	156
Représentations d'état et représentations entrée-sortie	157
Les hypothèses concernant le bruit et leurs conséquences sur la structure, l'apprentissage et l'utilisation du modèle	158
Apprentissage non adaptatif des modèles dynamiques sous forme canonique	166
Que faire en pratique ? Un exemple réel de modélisation « boîte noire »	172
Mise sous forme canonique des modèles dynamiques	175
Modélisation dynamique « boîte grise »	179
Principe de la modélisation semi-physique	179
<b>Conclusion : quels outils ?</b>	188
<b>Compléments théoriques et algorithmiques</b>	189
Quelques types de neurones usuels	189
Algorithme de Ho et Kashyap	191
Complément algorithmique : méthodes d'optimisation de Levenberg-Marquardt et de BFGS	191
Complément algorithmique : méthodes de recherche unidimensionnelle pour le paramètre d'apprentissage	193
Complément théorique : distance de Kullback-Leibler entre deux distributions gaussiennes	194
Complément algorithmique : calcul des leviers	196
<b>Bibliographie</b>	197

### **3 Compléments de méthodologie pour la modélisation : réduction de dimension et ré-échantillonnage** 203

<b>Pré-traitements</b>	204
Pré-traitements des entrées	204
Pré-traitement des sorties pour la classification supervisée	204
Pré-traitement des sorties pour la régression	205
<b>Réduction du nombre de composantes</b>	206
<b>Analyse en composantes principales</b>	206
Principe de l'ACP	206
<b>Analyse en composantes curvilignes</b>	210
Formalisation de l'analyse en composantes curvilignes	211
Algorithme d'analyse en composantes curvilignes	212
Mise en œuvre de l'analyse en composantes curvilignes	213
Qualité de la projection	214
Difficultés présentées par l'analyse en composantes curvilignes	214

Application en spectrométrie	215
<b>Le bootstrap et les réseaux de neurones</b>	216
Principe du bootstrap	217
Algorithme du bootstrap pour calculer un écart-type	218
L'erreur de généralisation estimée par bootstrap	218
La méthode NeMo	219
Test de la méthode NeMo	221
Conclusions	223
<b>Bibliographie</b>	224
<b>4 Identification « neuronale » de systèmes dynamiques commandés et réseaux bouclés (récurrents)</b>	225
<hr/>	
<b>Formalisation et exemples de systèmes dynamiques commandés à temps discret</b>	226
Formalisation d'un système dynamique commandé par l'équation d'état	226
Exemple d'un système dynamique à espace d'état discret	227
Exemple d'un oscillateur linéaire	227
Exemple du pendule inversé	228
Exemple d'un oscillateur non linéaire : l'oscillateur de Van der Pol	229
Introduction d'un bruit d'état dans un système dynamique à espace d'état discret : notion de chaîne de Markov	229
Introduction d'un bruit d'état dans un système dynamique à états continus : modèle linéaire gaussien	231
Modèles auto-régressifs	231
Limites des modélisations des incertitudes sur le modèle par un bruit d'état	233
<b>Identification de systèmes dynamiques commandés par régression</b>	233
Identification d'un système dynamique commandé par régression linéaire	233
Identification d'un système dynamique non linéaire par réseaux de neurones non bouclés	237
<b>Identification adaptative (en ligne) et méthode de l'erreur de prédiction récursive</b>	239
Estimateur récursif de la moyenne empirique	239
Estimateur récursif de la régression linéaire	241
Identification récursive d'un modèle AR	242
Méthode générale de l'erreur de prédiction récursive	243
Application à l'identification neuronale d'un système dynamique commandé	244
<b>Filtrage par innovation dans un modèle d'état</b>	245
Introduction d'une équation de mesure et problème du filtrage	245
Filtrage de Kalman	247
Extension du filtre de Kalman	251