



G. Dreyfus · J.-M. Martinez · M. Samuelides
M. B. Gordon · F. Badran · S. Thiria · L. Hérault

Sous la direction de Gérard Dreyfus

Réseaux de neurones

Méthodologie et applications

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

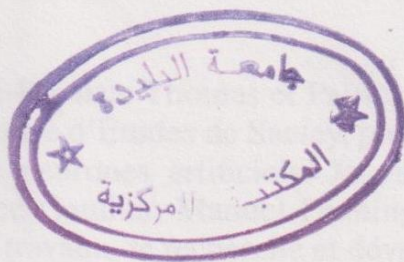
2-005-275-1

Réseaux de neurones

Méthodologie et applications

G. Dreyfus • J.-M. Martinez • M. Samuelides
M. B. Gordon • F. Badran • S. Thiria • L. Hérault

Sous la direction de Gérard Dreyfus



Deuxième tirage 2002

EYROLLES

Avant-propos	XV
Guide de lecture	XVI
Détail des contributions	XVII
1 Les réseaux de neurones : pourquoi et pour quoi faire ?	1
Les réseaux de neurones : définitions et propriétés	2
Les réseaux de neurones	3
L'apprentissage des réseaux de neurones	9
Propriété fondamentale des réseaux de neurones non bouclés à apprentissage supervisé : l'approximation parcimonieuse	10
À quoi servent les réseaux de neurones non bouclés à apprentissage supervisé ?	
Modélisation statique et discrimination (classification)	12
À quoi servent les réseaux de neurones à apprentissage non supervisé ?	
Analyse et visualisation de données	17
À quoi servent les réseaux de neurones bouclés à apprentissage supervisé ?	
Modélisation dynamique « boîte noire » et « semi-physique » ; commande de processus	17
À quoi servent les réseaux de neurones bouclés sans apprentissage ?	
L'optimisation combinatoire	19
Quand et comment mettre en œuvre des réseaux de neurones à apprentissage supervisé ?	19
Quand utiliser les réseaux de neurones ?	19
Comment mettre en œuvre ces réseaux de neurones ?	20
Conclusion	26
Réseaux de neurones à apprentissage supervisé et discrimination (classification)	26
Qu'est-ce qu'un problème de classification ?	26
Quand est-il opportun d'utiliser un classifieur statistique tel qu'un réseau de neurones ?	27
Classification probabiliste et formule de Bayes	29
Règle de décision de Bayes	32
Classification et régression	34
Quelques exemples d'applications des réseaux de neurones à divers domaines des sciences de l'ingénieur	39
Introduction	39
Une application en reconnaissance de formes : la lecture automatique de codes postaux	40
Une application en contrôle non destructif : la détection de défauts dans des rails par courants de Foucault	44

Une application en prévision : l'estimation de la probabilité de succès aux élections législatives	44
Une application en fouille de données : le filtrage de documents	45
Une application en bio-ingénierie : relations structure-activité (QSAR) pour la prédiction de propriétés chimiques de molécules	49
Une application en formulation : la prédiction de la température de liquidus de verres	50
Une application en modélisation de procédé industriel : la modélisation du soudage par points	51
Une application en robotique : la modélisation de l'actionneur hydraulique d'un bras de robot	53
Une application de la modélisation semi-physique à un procédé manufacturier	54
Deux applications en contrôle de l'environnement : pollution par l'ozone et hydrologie urbaine	55
Une application en robotique mobile : le pilotage automatique d'un véhicule autonome	58
Conclusion	59
Compléments théoriques et algorithmiques	59
Quelques types de neurones usuels	59
L'algorithme de Ho et Kashyap	61
Bibliographie	62

2 Modélisation à l'aide de réseaux de neurones : principes et méthodologie de conception de modèles 65

Qu'est-ce qu'un modèle ?	65
De la boîte noire au modèle de connaissance	65
Modèles statiques et modèles dynamiques	66
Comment traiter des variables incertaines ?	
Le contexte statistique de la modélisation	66
Notions élémentaires de statistiques	67
Qu'est-ce qu'une variable aléatoire ?	67
Espérance mathématique d'une variable aléatoire	68
Estimateur non biaisé d'une variable aléatoire	69
Variance d'une variable aléatoire	69
Intervalle de confiance	70
Tests d'hypothèse	71
Modélisation statique « boîte noire » : généralités	71
Formalisation du problème : la régression	71
Introduction à la méthodologie de conception	72
Sélection des entrées d'un modèle statique boîte noire	73
Réduction de l'espace de représentation des variables du modèle	73
Élimination des variables non pertinentes pour la grandeur à modéliser	74
Conclusion sur la sélection des entrées	79

Estimation des paramètres (apprentissage) d'un modèle statique	79
Apprentissage des modèles linéaires par rapport à leurs paramètres	81
Apprentissage non adaptatif de modèles statiques non linéaires par rapport à leurs paramètres (réseaux de neurones non bouclés)	84
Apprentissage adaptatif de modèles non linéaires par rapport à leurs paramètres	93
Apprentissage avec régularisation	94
Conclusion sur l'apprentissage de modèles statiques	100
Sélection de modèle	100
Préliminaire : élimination de modèles surajustés par calcul du rang de la matrice jacobienne	102
Approche globale de la sélection de modèles : validation croisée et « leave-one-out »	103
Moindres carrés locaux : effet du retrait d'un exemple sur le modèle et « leave-one-out virtuel »	105
Méthodologie de sélection de modèle par combinaison de l'approche globale et de l'approche locale	109
Modélisation dynamique « boîte noire »	115
Représentations d'état et représentations entrée-sortie	115
Les hypothèses concernant le bruit et leurs conséquences sur la structure, l'apprentissage et l'utilisation du modèle	117
Apprentissage non adaptatif des modèles dynamiques sous forme canonique	124
Que faire en pratique ? Un exemple réel de modélisation « boîte noire »	127
Mise sous forme canonique des modèles dynamiques	130
Modélisation dynamique « boîte grise »	133
Principe de la modélisation semi-physique	133
Conclusion : quels outils ?	141
Compléments théoriques et algorithmiques	142
Complément sur les intervalles de confiance : démarche et exemple	142
Exemple de test d'hypothèse	143
Complément : distributions de Pearson, de Student et de Fisher	143
Compléments sur la sélection des entrées : test de Fisher ; calcul de la fonction de distribution du rang du descripteur sonde	144
Complément algorithmique : méthodes d'optimisation de Levenberg-Marquardt et de BFGS	146
Complément algorithmique : méthodes de recherche unidimensionnelle pour le paramètre d'apprentissage	148
Complément théorique : distance de Kullback-Leibler entre deux distributions gaussiennes	149
Complément algorithmique : calcul des leviers	150
Bibliographie	151

3 Compléments de méthodologie pour la modélisation : réduction de dimension et ré-échantillonnage	155
Pré-traitements	156
Pré-traitements des entrées	156
Pré-traitement des sorties pour la classification supervisée	156
Pré-traitement des sorties pour la régression	157
Réduction du nombre de composantes	158
Analyse en composantes principales	158
Principe de l'ACP	158
Analyse en composantes curvilignes	162
Formalisation de l'analyse en composantes curvilignes	163
Algorithme d'analyse en composantes curvilignes	164
Mise en œuvre de l'analyse en composantes curvilignes	165
Qualité de la projection	166
Difficultés présentées par l'analyse en composantes curvilignes	166
Application en spectrométrie	167
Le bootstrap et les réseaux de neurones	168
Principe du bootstrap	169
Algorithme du bootstrap pour calculer un écart-type	169
L'erreur de généralisation estimée par bootstrap	170
La méthode NeMo	171
Test de la méthode NeMo	173
Conclusions	175
Bibliographie	175
4 Identification « neuronale » de systèmes dynamiques commandés et réseaux bouclés (récurrents)	177
Formalisation et exemples de systèmes dynamiques commandés à temps discret	178
Formalisation d'un système dynamique commandé par l'équation d'état	178
Exemple d'un système dynamique à espace d'état discret	179
Exemple d'un oscillateur linéaire	179
Exemple du pendule inversé	180
Exemple d'un oscillateur non linéaire : l'oscillateur de Van der Pol	181
Introduction d'un bruit d'état dans un système dynamique à espace d'état discret : notion de chaîne de Markov	181
Introduction d'un bruit d'état dans un système dynamique à états continus : modèle linéaire gaussien	183

Modèles auto-régressifs	183
Limites des modélisations des incertitudes sur le modèle par un bruit d'état	185
Identification de systèmes dynamiques commandés par régression	185
Identification d'un système dynamique commandé par régression linéaire	185
Identification d'un système dynamique non linéaire par réseaux de neurones non bouclés	189
Identification adaptative (en ligne) et méthode de l'erreur de prédiction récursive	191
Estimateur récursif de la moyenne empirique	191
Estimateur récursif de la régression linéaire	193
Identification récursive d'un modèle AR	194
Méthode générale de l'erreur de prédiction récursive	195
Application à l'identification neuronale d'un système dynamique commandé	196
Filtrage par innovation dans un modèle d'état	197
Introduction d'une équation de mesure et problème du filtrage	197
Filtrage de Kalman	199
Extension du filtre de Kalman	203
Apprentissage adaptatif d'un réseau de neurones par la méthode du filtrage de Kalman	204
Réseaux neuronaux récurrents ou bouclés	206
Simulateur neuronal d'un système dynamique commandé en boucle ouverte	206
Simulateur neuronal d'un système dynamique commandé en boucle fermée	207
Quelques réseaux bouclés particuliers	207
Mise sous forme canonique des réseaux bouclés	210
Apprentissage des réseaux de neurones récurrents ou bouclés	210
Apprentissage dirigé (teacher forcing)	211
Dépliage de la forme canonique et rétropropagation à travers le temps	212
Apprentissage en temps réel des réseaux bouclés	214
Application des réseaux neuronaux bouclés à l'identification de systèmes dynamiques commandés mesurés	215
Compléments algorithmiques et théoriques	216
Calcul du gain de Kalman et propagation de la covariance	216
Importance de la distribution des retards dans un réseau récurrent	218
Bibliographie	219
5 Apprentissage d'une commande en boucle fermée	221
Généralités sur la commande en boucle fermée des systèmes non linéaires	221
Principe de la commande en boucle fermée	221
Commandabilité	222
Stabilité des systèmes dynamiques commandés	223

Synthèse d'une commande « neuronale » par inversion du modèle du processus	225
Inversion directe	225
Utilisation d'un modèle de référence	228
Commande avec modèle interne	229
Utilisation des réseaux récurrents	230
Programmation dynamique et commande optimale	231
Exemple de problème déterministe à espace d'états discret	231
Exemple de problème de décision markovienne	233
Définition d'un problème de décision markovienne	234
Programmation dynamique à horizon fini	237
Programmation dynamique à horizon infini et à coût actualisé	238
Problèmes de décision markovienne partiellement observés	240
Apprentissage par renforcement et programmation neuro-dynamique	240
Évaluation d'une politique par la méthode de Monte-Carlo et apprentissage par renforcement	240
Présentation de l'algorithme TD d'évaluation d'une politique	241
Apprentissage par renforcement : méthode du Q-learning	244
Apprentissage par renforcement et approximation neuronale	246
Bibliographie	249
 6 La discrimination	 251
Apprentissage de la discrimination	252
Erreurs d'apprentissage et de généralisation	253
Surfaces discriminantes	254
Séparation linéaire : le perceptron	254
Géométrie de la classification	256
Algorithmes d'apprentissage pour le perceptron	258
Exemple d'application : la classification de signaux de sonar	266
Algorithmes d'apprentissage adaptatifs (« en ligne »)	267
Interprétation de l'apprentissage en termes de forces	268
Au-delà de la séparation linéaire	269
Perceptron sphérique	269
Heuristiques constructives	270
Algorithme constructif NetLS	271
Machines à vecteurs supports (Support Vector Machines)	272
Problèmes à plusieurs classes	274
Questions théoriques	275
Formulation probabiliste de l'apprentissage	276

Théorie statistique de Vapnik	279
Prédiction du comportement typique	282
Compléments théoriques	283
Bornes au nombre d'itérations de l'algorithme du perceptron	283
Nombre de dichotomies linéairement séparables	284
Bibliographie	285
7 Cartes auto-organisatrices et classification automatique	287
Notations et définitions	289
Méthode des k-moyennes	290
Présentation de l'algorithme	290
Version stochastique des k-moyennes	292
Interprétation probabiliste des k-moyennes	295
Carte topologique auto-organisatrice	298
Les cartes auto-organisatrices	298
L'algorithme d'optimisation non adaptative des cartes topologiques	301
L'algorithme de Kohonen	307
Discussion	308
Architecture neuronale et carte topologique	309
Architecture et carte topologique évolutive	310
Interprétation de l'ordre topologique	311
Carte topologique probabiliste	313
Classification et carte topologique	316
Étiquetage de la carte par données expertisées	316
Recherche d'une partition adaptée aux classes recherchées	317
Étiquetage et classification	319
Applications	320
Une application en télédétection satellitaire	321
Classification et PRSOM	329
Carte topologique et recherche documentaire	332
Bibliographie	335
8 Réseaux de neurones sans apprentissage pour l'optimisation	337
Modélisation d'un problème d'optimisation	337
Exemples	338
Problème du voyageur de commerce	339

Complexité d'un problème d'optimisation	339
Exemple	340
Approches classiques des problèmes combinatoires	340
Introduction aux métaheuristiques	341
Les techniques dérivées de la physique statistique	342
Analyse canonique	342
Analyse microcanonique	346
Exemple : problème du voyageur de commerce	348
Les approches neuronales	352
Les neurones formels utilisés pour l'optimisation	352
Architectures de réseaux de neurones pour l'optimisation	353
Fonctions d'énergie pour l'optimisation combinatoire	353
Les réseaux de neurones récurrents de Hopfield	354
Améliorations des réseaux de Hopfield	360
La recherche tabou	368
Les algorithmes génétiques	369
Vers des approches hybrides	369
Conclusion	370
Le choix d'une technique	370
Bibliographie	370
 Bibliographie commentée	 375
 Index	 379

Réseaux de neurones

Gérard Dreyfus dirige le laboratoire d'électronique de l'École supérieure de physique et de chimie industrielles de la ville de Paris (ESPCI) où il enseigne notamment les méthodes de modélisation par apprentissage. Il dispense des formations continues à l'usage des ingénieurs dans ce domaine.

Manuel Samuelides dirige le département de mathématiques appliquées de l'ENSAE (Supaéro); il y enseigne les probabilités, l'optimisation et les techniques probabilistes de l'apprentissage. Il effectue des recherches au département de traitement de l'information et modélisation de l'ONERA.

Jean-Marc Martinez enseigne les méthodes d'apprentissage statistique à l'INSTN de Saclay et à Evry en collaboration avec le LSC, unité mixte CEA - Université. Il effectue des recherches dans le domaine des méthodes adaptées à la supervision de la simulation.

Mirta B. Gordon, physicienne et directeur de recherches au CNRS, effectue des recherches sur les algorithmes d'apprentissage. Elle travaille au laboratoire Leibniz de l'Institut de mathématiques appliquées de Grenoble (IMAG) ainsi qu'au groupe Théorie du département de recherche fondamentale du CEA Grenoble.

Fouad Badran, maître de conférences au CNAM, y enseigne les réseaux de neurones.

Sylvie Thiria, professeur à l'université de Versailles Saint-Quentin-en-Yvelines, effectue des recherches sur la modélisation neuronale et sur ses applications, notamment à la géophysique, au Laboratoire d'océanographie dynamique et de climatologie (LODYC).

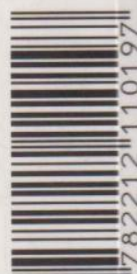
Laurent Hérault est chef de projet au CEA - LETI où il coordonne des recherches sur les réseaux de neurones et l'optimisation combinatoire appliqués à des problèmes industriels, notamment le traitement du signal pour les télécommunications sans fil.

- ▶ Les réseaux de neurones constituent une technique de traitement de données bien comprise et maîtrisée, désormais indispensable dans la boîte à outils de tout ingénieur soucieux d'extraire le maximum d'informations pertinentes des données dont il dispose : effectuer des prévisions, de la fouille de données, élaborer des modèles, reconnaître des formes ou des signaux, etc.
- ▶ Joignant fondements théoriques et applications pratiques dans un langage accessible, cet ouvrage a pour but de permettre aux décideurs, aux ingénieurs et aux chercheurs de bénéficier de méthodologies claires pour mettre en œuvre avec profit les réseaux de neurones dans des applications industrielles, financières ou bancaires, dont de nombreux exemples sont présentés.
- ▶ Le lecteur désireux d'acquérir des bases solides trouvera, pour chaque chapitre, des compléments théoriques qui lui permettront d'approfondir les concepts, et des compléments algorithmiques destinés à en faciliter l'implantation informatique.
- ▶ À qui s'adresse ce livre ?
 - Aux ingénieurs, informaticiens, industriels et décideurs ayant à résoudre des problèmes de modélisation, de reconnaissance, de prévision, de commande, etc.
 - Aux étudiants et élèves-ingénieurs des disciplines scientifiques et économiques.
 - Aux enseignants de ces disciplines à la recherche d'un ouvrage complet et d'études de cas sur les réseaux de neurones.

▶ Sommaire

Panorama des réseaux de neurones et de leurs applications · Exemples d'applications : reconnaissance de formes, fouille de données, prévision, prédiction de propriétés chimiques et physiques, modélisation et commande de procédés industriels, robotique · **Méthodologie de conception de modèles :** Notions de statistiques - Modélisation statique et dynamique - Modélisation semi-physique · **Bootstrap et rééchantillonnage · Simulation et commande de processus :** Réseaux récurrents · **Discrimination :** Réseaux classifieurs - Machines à vecteurs supports · **Analyse et visualisation de données :** Cartes de Kohonen · **Optimisation combinatoire et recherche opérationnelle :** Compléments théoriques et algorithmiques · Bibliographie commentée.

Code éditeur : G11019 • ISBN : 2-212-11019-7



9 782212 110197

50 €



Eyrolles