
L'APPRENTISSAGE ARTIFICIEL

CONCEPTS ET ALGORITHMES.

(2ÈME ÉDITION)

Antoine CORNUÉJOLS
et
Laurent MICLET



Editions Eyrolles, 2009

Table des matières

Table des matières	iii
Avant-propos	ix
Quelques applications de l'apprentissage artificiel	x
Quelques définitions de base	x
Deux champs industriels de l'apprentissage artificiels : la reconnaissance des formes et la fouille de données	xi
Les caractéristiques de l'apprentissage artificiel	xiii
Trois exemples d'apprentissage	xiv
Organisation et plan de l'ouvrage	xvii
Guide de lecture	xix
Notations	xxv
I Les Fondements de l'Apprentissage	1
1 De l'apprentissage naturel à l'apprentissage artificiel	3
1 L'apprentissage artificiel	4
2 Deux exemples : apprendre à jouer, apprendre à lire	6
3 Deux approches : la cybernétique et les sciences cognitives	10
4 Les concepts de base de l'apprentissage	14
5 L'induction comme un jeu entre espaces	22
6 Retour sur l'organisation de l'ouvrage	29
2 Première approche théorique de l'induction	37
1 Poser un problème d'apprentissage	40
2 Approches bayésiennes et approche directe pour décider	46
3 Le critère inductif	49
4 Analyse du principe de Minimisation du Risque Empirique	60
5 Le lien entre le passé et le futur et le <i>no-free-lunch theorem</i>	73
6 Notes historiques et bibliographiques	78

3	L'environnement méthodologique de l'apprentissage	81
1	L'espace des données d'apprentissage	85
2	L'espace des hypothèses d'apprentissage	101
3	Les protocoles d'apprentissage	112
4	L'évaluation de l'apprentissage	113
5	La comparaison des méthodes d'apprentissage	124
6	Autres problèmes pratiques	129
II	Apprentissage par Exploration	135
4	Induction et relation d'ordre : l'espace des versions	137
1	Les concepts de base	140
2	La structuration de l'espace des hypothèses	144
3	La construction de l'espace des versions	152
4	La représentation des connaissances par un treillis de Galois	156
5	La programmation logique inductive	163
1	La programmation logique inductive : le cadre général	167
2	La logique des prédicats et les programmes logiques : terminologie	173
3	La structuration de l'espace des hypothèses en logique des prédicats	177
4	L'exploration de l'espace des hypothèses	184
5	Deux exemples de systèmes de PLI	188
6	La probabilisation de la PLI	192
7	Les domaines d'application de la PLI	193
8	Les chantiers de la PLI	195
6	Reformulation, transfert de connaissances et apprentissage par analogie	201
1	L'apprentissage en présence de théorie	203
2	L'apprentissage par examen de preuve (<i>EBL</i>)	204
3	Abstraction et reformulation des connaissances	211
4	Changement de repère, raisonnement par analogie et RaPC	213
5	L'apprentissage par proportion analogique	216
6	Bilan	220
7	L'inférence grammaticale	223
1	Définitions et notations	229
2	Les protocoles de l'inférence : quelques résultats théoriques	237
3	L'espace de recherche de l'inférence régulière	243
4	L'inférence régulière sans échantillon négatif	245
5	L'inférence régulière sous contrôle d'un échantillon négatif	250
6	L'inférence de grammaires algébriques.	254
7	L'inférence d'automates probabilistes	260
8	Quelques approches complémentaires	263

8	Apprentissage par évolution simulée	267
1	Trois espaces au lieu de deux	269
2	Un modèle formel simplifié de l'évolution	272
3	Les algorithmes génétiques	273
4	Les stratégies d'évolution	285
5	La programmation génétique	285
6	La coévolution	295
III Apprentissage par Optimisation		301
9	L'apprentissage de modèles linéaires	303
1	Introduction	304
2	Régression linéaire	305
3	Séparatrices linéaires	310
4	Modèles linéaires par morceaux et combinaisons de modèles locaux	321
5	La recherche des facteurs pertinents	324
10	L'apprentissage de réseaux connexionnistes	329
1	Les différents éléments d'un réseau connexionniste	332
2	L'architecture multicouche	334
3	L'algorithme d'apprentissage	339
4	Quelques résultats théoriques sur les réseaux connexionnistes	348
5	Comment choisir l'architecture d'un réseau ?	349
6	Les réseaux à architecture profonde	349
7	Réseaux et régime dynamique : le « <i>Reservoir Computing</i> »	351
11	L'apprentissage de réseaux bayésiens	357
1	Les modèles graphiques	359
2	Les réseaux d'inférence bayésiens	361
3	Les inférences dans les réseaux bayésiens	369
4	L'apprentissage des réseaux bayésiens	373
5	L'inférence de relations causales	382
6	Applications	383
7	Quelques logiciels	384
12	L'apprentissage de modèles de Markov cachés	387
1	Les modèles de Markov observables	390
2	Les modèles de Markov cachés (HMM)	391
3	Les HMM comme règles de classification de séquences	396
4	L'évaluation de la probabilité d'observation	397
5	Le calcul du chemin optimal : l'algorithme de Viterbi	399
6	L'apprentissage	402
7	Approfondissements	407
8	Applications	408
13	Apprentissage par inférence d'arbres	411
1	Les arbres de décision	413
2	Les arbres de régression	427

IV	Apprentissage par approximation et interpolation	431
14	Méthodes à noyaux	433
1	Trois voies vers les méthodes à noyau	435
2	Philosophie des méthodes à noyaux	445
3	Les Séparatrices à Vaste Marge (SVM)	447
4	Autres types d'induction avec fonctions noyau	457
5	Ingénierie des fonctions noyau	460
6	Les méthodes à noyaux en pratique	479
7	Bilan et perspectives	485
15	L'apprentissage bayésien et son approximation	489
1	L'apprentissage bayésien	492
2	Les méthodes paramétriques	501
3	Les méthodes non paramétriques	510
4	Les méthodes semi paramétriques	525
16	L'apprentissage de réflexes par renforcement	531
1	Description du problème	534
2	Si tout est connu : l'utilité de la fonction d'utilité	541
3	L'apprentissage des fonctions d'utilité quand l'environnement est connu	543
4	Sans modèle du monde : la méthode de Monte-Carlo	547
5	La méthode des différences temporelles	548
6	La généralisation dans l'apprentissage par renforcement	552
7	Contrôle optimal par recherche arborescente et algorithme UCT	558
8	Le cas des environnements partiellement observables	561
9	Exemples d'application	562
10	Bilan et perspectives	565
V	Au delà de l'apprentissage supervisé	569
17	Apprentissage de combinaisons d'experts	571
1	Principes des méthodes par combinaison	572
2	Le vote de plusieurs classificateurs	576
3	Les codes correcteurs de classes	577
4	Le <i>boosting</i> d'un algorithme d'apprentissage	582
5	Le <i>bagging</i>	590
6	L'apprentissage en cascade (<i>cascading</i>)	591
18	La classification non supervisée et la fouille de données	593
1	Introduction	596
2	Les méthodes de classification fondées sur les distances	597
3	Les méthodes de classification par des modèles probabilistes	608
4	Méthodes spectrales de catégorisation	610
5	La classification de données symboliques	611
6	La fouille de données	614
7	Les analyses en composantes	624

19	L'apprentissage semi-supervisé	633
1	Présentation	634
2	Les modèles génératifs : apprentissage dans l'espace joint $\mathcal{X} \times \mathcal{Y}$	638
3	L'auto-apprentissage	642
4	Le co-apprentissage	642
5	L'utilisation d'hypothèses fondamentales sur les données	644
6	Quelques directions pour l'apprentissage semi-supervisé	654
7	Conclusion	655
20	Vers de nouvelles tâches et de nouvelles questions	657
1	Apprentissage actif	659
2	Apprentissages en-ligne, incrémental et par transferts	667
3	Apprentissage à partir de flux de données	675
4	Apprentissage de sorties structurées	680
5	Apprentissage pour le filtrage collaboratif	684
21	Analyse de l'induction : approfondissements et ouvertures	693
1	Généralisation de l'analyse du principe MRE	694
2	Principes inductifs contrôlant l'espace des hypothèses	700
3	Prise en compte de l'algorithme d'apprentissage dans la théorie	713
4	Autres types d'analyses	718
VI	Annexes techniques	725
22	Annexes techniques	727
1	Exemples de fonctions de perte en induction	727
2	Le calcul de l'intervalle de confiance de l'estimation de la probabilité d'une règle de classification	730
3	Estimation d'une densité de probabilité en un point.	731
4	L'estimation des paramètres d'une distribution gaussienne.	732
5	Pourquoi et comment la règle du PPV converge-t-elle?	733
6	Pourquoi la règle de décision bayésienne est-elle optimale?	734
7	Apprentissage par Estimation-Maximisation.	734
8	Optimisation par descente de gradient	738
9	La rétropropagation du gradient de l'erreur	742
10	L'analyse de l'induction de Vapnik	745
11	L'induction par compression d'information	756
	Bibliographie	763
	Index	795