

Avant-propos et guide de lecture

En une vingtaine d'années, l'apprentissage artificiel est devenu une branche majeure des mathématiques appliquées, à l'intersection des statistiques et de l'intelligence artificielle. Son objectif est de réaliser des modèles qui apprennent « par l'exemple » : il s'appuie sur des données numériques (résultats de mesures ou de simulations), contrairement aux modèles « de connaissances » qui s'appuient sur des équations issues des premiers principes de la physique, de la chimie, de la biologie, de l'économie, etc. L'apprentissage statistique est d'une grande utilité lorsque l'on cherche à modéliser des processus complexes, souvent non linéaires, pour lesquels les connaissances théoriques sont trop imprécises pour permettre des prédictions précises. Ses domaines d'applications sont multiples : fouille de données, bio-informatique, génie des procédés, aide au diagnostic médical, télécommunications, interface cerveau-machines, et bien d'autres.

Cet ouvrage reflète en partie l'évolution de cette discipline, depuis ses balbutiements au début des années 1980, jusqu'à sa situation actuelle ; il n'a pas du tout la prétention de faire un point, même partiel, sur l'ensemble des développements passés et actuels, mais plutôt d'insister sur les principes et sur les méthodes éprouvés, dont les bases scientifiques sont sûres. Dans un domaine sans cesse parcouru de modes multiples et éphémères, il est utile, pour qui cherche à acquérir les connaissances et principes de base, d'insister sur les aspects pérennes du domaine.

Cet ouvrage fait suite à *Réseaux de neurones, méthodologies et applications*, des mêmes auteurs, paru en 2000, réédité en 2004, chez le même éditeur, puis publié en traduction anglaise chez Springer. Consacré essentiellement aux réseaux de neurones et aux cartes auto-adaptatives, il a largement contribué à populariser ces techniques et à convaincre leurs utilisateurs qu'il est possible d'obtenir des résultats remarquables, à condition de mettre en œuvre une méthodologie de conception rigoureuse, scientifiquement fondée, dans un domaine où l'empirisme a longtemps tenu lieu de méthode.

Tout en restant fidèle à l'esprit de cet ouvrage, combinant fondements mathématiques et méthodologie de mise en œuvre, les auteurs ont élargi le champ de la présentation, afin de permettre au lecteur d'aborder d'autres méthodes d'apprentissage statistique que celles qui sont directement décrites dans cet ouvrage. En effet, les succès de l'apprentissage dans un grand nombre de domaines ont poussé au développement de très nombreuses variantes, souvent destinées à répondre efficacement aux exigences de telle ou telle classe d'applications. Toutes ces variantes ont néanmoins des bases théoriques et des aspects méthodologiques communs, qu'il est important d'avoir présents à l'esprit.

Le terme d'apprentissage, comme celui de réseau de neurones, évoque évidemment le fonctionnement du cerveau. Il ne faut pourtant pas s'attendre à trouver ici d'explications sur les mécanismes de traitement des informations dans les systèmes nerveux ; ces derniers sont d'une grande complexité, résultant de processus électriques et chimiques subtils, encore mal compris en dépit de la grande quantité de données expérimentales disponibles. Si les méthodes d'apprentissage statistique peuvent être d'une grande utilité pour créer des modèles empiriques de telle ou telle fonction réalisée par le système nerveux, celles qui sont décrites dans cet ouvrage n'ont aucunement la prétention d'imiter, même vaguement, le fonctionnement du cerveau. L'apprentissage artificiel, notamment statistique, permettra-t-il un jour de donner aux ordinateurs des capacités analogues à celles des êtres humains ? Se rapprochera-t-on de cet objectif en perfectionnant les techniques actuelles d'apprentissage, ou bien des approches radicalement nouvelles sont-elles indispensables ? Faut-il s'inspirer de ce que l'on sait, ou croit savoir, sur le fonctionnement du cerveau ? Ces questions font l'objet de débats passionnés, et passionnantes, au sein de la communauté scientifique : on n'en trouvera pas les réponses ici.

Les objectifs de ce livre sont, plus modestement :

- de convaincre les ingénieurs, chercheurs, et décideurs, de l'intérêt et de la grande efficacité de l'apprentissage statistique ;
- de leur permettre de le mettre en œuvre de manière simple et raisonnée dans des applications.

Guide de lecture

La variété des motivations qui peuvent amener le lecteur à aborder cet ouvrage justifie sans doute un guide de lecture. En effet, les applications de l'apprentissage statistique ne nécessitent pas toutes la mise en œuvre des mêmes méthodes.

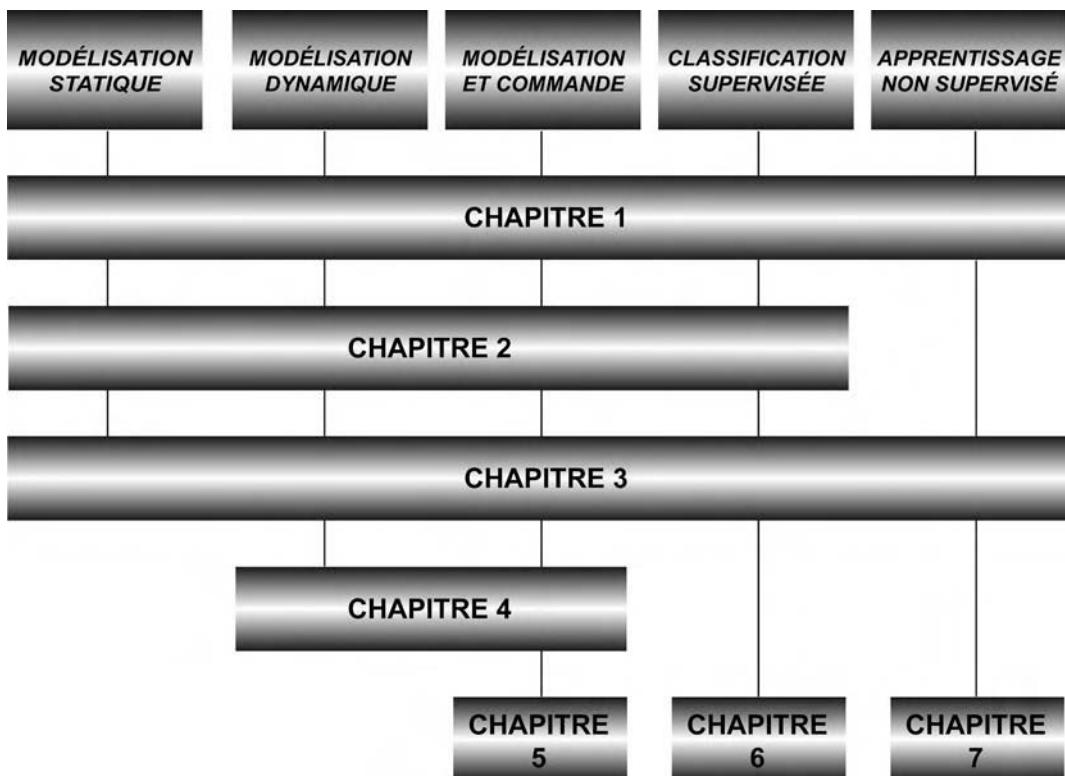
Le premier chapitre (« L'apprentissage statistique : pourquoi, comment ? ») constitue une présentation générale des principes de l'apprentissage statistique et des problèmes fondamentaux à résoudre. À partir d'exemples académiques très simples, le lecteur est amené à découvrir les problèmes que pose la conception de modèles par apprentissage. Ces problèmes sont ensuite formalisés par la présentation de quelques éléments de la théorie de l'apprentissage. La conception des modèles les plus simples – les modèles linéaires en leurs paramètres – est décrite. Enfin, les différentes étapes de la conception d'un modèle par apprentissage statistique sont détaillées : sélection de variables, apprentissage, sélection de modèle, test du modèle sélectionné.

Le chapitre 2 est entièrement consacré aux réseaux de neurones, qui constituent une des familles de modèles les plus utilisés. Les lecteurs qui s'intéressent à un problème de *modélisation statique* liront ce chapitre jusqu'à la section « Techniques et méthodologie de conception de modèles statiques (réseaux non bouclés) » incluse. Ils tireront également profit de la lecture du chapitre 3 (« Compléments de méthodologie pour la modélisation : réduction de dimension et validation de modèle par ré-échantillonnage »).

Les lecteurs qui se posent un problème de *modélisation dynamique* liront le chapitre 2 en entier, le chapitre 3 et le chapitre 4 (« Identification “neuronale” de systèmes dynamiques commandés et réseaux bouclés (récurrents) »). S'ils veulent utiliser ce modèle au sein d'un dispositif de commande de processus, ils liront ensuite le chapitre 5 (« Apprentissage d'une commande en boucle fermée »).

Les lecteurs qui s'intéressent à un problème de *classification supervisée* (ou discrimination) liront le chapitre 1, la section « Réseaux de neurones à apprentissage supervisé et discrimination » du chapitre 2, puis le chapitre 3 (« Compléments de méthodologie pour la modélisation : réduction de dimension et validation de modèle par ré-échantillonnage ») et surtout le chapitre 6 (« Discrimination »), qui introduit, de manière originale, les machines à vecteurs supports.

Enfin, les lecteurs qui cherchent à résoudre un problème qui relève de l'*apprentissage non supervisé* passeront du chapitre 1 au chapitre 3, puis au chapitre 7 (« Cartes auto-organisatrices et classification automatique »).



Détail des contributions

Chapitres 1 et 2	<p>Gérard Dreyfus est professeur à l'École Supérieure de Physique et de Chimie Industrielles (ESPCI-ParisTech), et directeur du Laboratoire d'Électronique de cet établissement. Il enseigne l'apprentissage statistique à l'ESPCI, ainsi que dans plusieurs masters et mastères. Depuis 1988, il organise chaque année deux sessions de formation continue pour ingénieurs, consacrées à l'apprentissage statistique et à ses applications industrielles et financières. Depuis 1982, les recherches de son laboratoire sont entièrement consacrées à la modélisation et à l'apprentissage, pour l'ingénierie et la neurobiologie.</p> <p>ESPCI, Laboratoire d'Électronique, 10 rue Vauquelin, F – 75005 Paris – France</p>
Chapitre 3	<p>Jean-Marc Martinez, ingénieur au Centre d'Études de Saclay, effectue des recherches dans le domaine des méthodes adaptées à la supervision de la simulation. Il enseigne les méthodes d'apprentissage statistique à l'INSTN de Saclay et à Évry en collaboration avec le LSC, unité mixte CEA – Université.</p> <p>DM2S/SFME Centre d'Études de Saclay, 91191 Gif sur Yvette – France</p>

Chapitres 4 et 5	<p>Manuel Samuelides, professeur à l'École Nationale Supérieure de l'Aéronautique et de l'Espace (Supaéro), et chef du département de Mathématiques Appliquées de cette école, enseigne les probabilités, l'optimisation et les techniques probabilistes de l'apprentissage et de la reconnaissance des formes. Il effectue des recherches sur les applications des réseaux de neurones au Département de Traitement de l'Information et Modélisation de l'ONERA.</p> <p>École Nationale Supérieure de l'Aéronautique et de l'Espace, département Mathématiques Appliquées, 10 avenue Édouard Belin, BP 4032, 31055 Toulouse Cedex – France</p>
Chapitre 6	<p>Mirta B. Gordon, physicienne et directrice de recherches au CNRS, est responsable de l'équipe « Apprentissage: Modèles et Algorithmes » (AMA) au sein du laboratoire TIMC-IMAG (Grenoble). Elle effectue des recherches sur la modélisation des systèmes complexes adaptatifs, et sur la théorie et les algorithmes d'apprentissage. Elle enseigne ces sujets dans différentes écoles doctorales.</p> <p>Laboratoire TIMC – IMAG, Domaine de la Merci – Bât. Jean Roget, 38706 La Tranche – France</p>
Chapitre 7	<p>Fouad Badran, professeur au CNAM (CEDRIC), y enseigne les réseaux de neurones.</p> <p>Mustapha Lebbah est maître de conférences à l'université de Paris 13.</p> <p>Laboratoire d'Informatique Médicale et Bio-Informatique (LIMBIO), 74, rue Marcel Cachin 93017 Bobigny Cedex – France</p> <p>Sylvie Thiria est professeur à l'université de Versailles Saint-Quentin-en-Yvelines, chercheur au LODYC (Laboratoire d'Océanographie DYnamique et de Climatologie). Elle effectue des recherches sur la modélisation neuronale et sur son application à des domaines comme la géophysique.</p> <p>Laboratoire d'Océanographie Dynamique et de Climatologie (LODYC), case 100, Université Paris 6, 4 place Jussieu 75252 Paris cedex 05 – France</p>