

Avant-propos

L'information n'est pas la connaissance. À mesure que se développent les technologies permettant de stocker, d'échanger de l'information et d'y accéder, la question de l'analyse et de la synthèse de ces informations devient essentielle. Le développement de technologies facilitant le passage de l'information à la connaissance est déterminant pour que la société annoncée de l'information ne soit pas en réalité la société du bruit.

Deux types d'approches connaissent donc tout naturellement un intérêt croissant. Les méthodes statistiques tout d'abord, parce qu'elles sont précisément conçues pour permettre le passage de l'observation à la loi, fût-elle loi de probabilité. Les technologies de l'intelligence artificielle ensuite, parce que leur vocation est de permettre aux ordinateurs de traiter de la connaissance plutôt que de l'information.

Les réseaux bayésiens sont le résultat d'une convergence entre ces deux disciplines et constituent aujourd'hui l'un des formalismes les plus complets et les plus cohérents pour l'acquisition, la représentation et l'utilisation de connaissances par des ordinateurs. Encore du domaine de la recherche au début des années 1990, cette technologie connaît de plus en plus d'applications, depuis le contrôle de véhicules autonomes à la modélisation des risques opérationnels, en passant par le *data mining* ou la localisation des gènes.

Les réseaux bayésiens, qui doivent leur nom aux travaux de Thomas Bayes au XVIII^e siècle sur la théorie des probabilités, sont le résultat de recherches effectuées dans les années 1980, dues à J. Pearl à UCLA et à une équipe de recherche danoise à l'université de Aalborg.

L'objectif initial de ces travaux était d'intégrer la notion d'incertitude dans les systèmes experts. Les chercheurs se sont rapidement aperçus que la construction d'un système expert nécessitait presque toujours la prise en compte de l'incertitude dans le raisonnement.

En effet, dans la plupart des domaines complexes, un expert humain est capable de porter un jugement sur une situation, même en l'absence de

toutes les données nécessaires. En médecine, par exemple, une même combinaison de symptômes peut être observée dans différentes pathologies.

Il n'y a donc pas de règle stricte qui permette de passer systématiquement d'un ensemble d'observations à un diagnostic. De plus, les informations pertinentes ne sont pas toujours observables. Pour que des systèmes experts puissent être utilisés dans de tels domaines, il faut donc qu'ils soient capables de raisonner sur des faits et des règles incertains. Dans le cadre des systèmes experts, les réseaux bayésiens constituent une approche possible pour intégrer l'incertitude dans le raisonnement. D'autres méthodes existent, mais les réseaux bayésiens présentent l'avantage d'être une approche quantitative.

D'un autre côté, imaginons à présent un statisticien, qui s'efforce d'analyser un tableau de mesures de plusieurs variables sur une population donnée. Il va pour cela essayer de démêler les relations pertinentes entre les variables, les dépendances ou indépendances entre plusieurs groupes de variables. L'utilisation de réseaux bayésiens va lui permettre d'extraire de ce tableau une représentation compacte, sans perte d'information, à partir de laquelle il va être beaucoup plus facile de raisonner.

Le lien entre ces deux problématiques est clairement celui de la connaissance. D'un côté, un expert dispose d'une connaissance présentant certaines incertitudes. Pour la formaliser, il va utiliser des descriptions causales : A a une influence sur B ; en général, si B est observé, il y a de fortes chances que C se produise, etc. Pour rendre cette connaissance opérationnelle, il lui faut quantifier ses incertitudes, c'est-à-dire les convictions plus ou moins précises que l'expert a des liens entre les faits.

D'un autre côté, un ensemble de données contient lui aussi de la connaissance, mais qui n'est pas directement accessible à un analyste, car elle est noyée dans les chiffres. Pour rendre cette connaissance interprétable, il faut la transformer en *modèle de causalité*, mettant en évidence les liens entre les variables observées.

C'est grâce à la notion mathématique de probabilité que les réseaux bayésiens vont permettre de résoudre ces deux problèmes duaux : transformer en chiffres une connaissance subjective, et transformer en modèle interprétable une connaissance contenue dans des chiffres.

L'expert formalise sa connaissance sous forme de modèle de causalité, indiquant les liens entre les variables. Cette description graphique est transformée en une loi de probabilité équivalente. Cette loi de probabilité permet de faire des calculs, et donc en particulier des raisonnements prenant en compte des aspects incertains.

Réciproquement, à partir des données, on va mettre en évidence des

propriétés (indépendances, causalités) de la relation entre les différentes variables observées.

Cette relation est transformée en graphe de causalités, qui peut alors être lu et interprété par un analyste, beaucoup plus facilement que les données initiales. Ces deux opérations ne sont possibles que grâce aux trois propriétés suivantes :

- Les probabilités subjectives (celles que l'expert utilise pour décrire les liens entre les variables) sont assimilables à des probabilités mathématiques (H1).
- Les fréquences observées (tableau de mesures) sont assimilables à des probabilités mathématiques (H2).
- Le graphe de causalités est une représentation fidèle d'une loi de probabilité sous-jacente : il est alors possible de raisonner sur le graphe sans revenir aux chiffres.

Les deux premières propriétés sont des hypothèses de travail, et leur discussion peut être considérée comme relevant de la philosophie. La dernière, en revanche, est un résultat très important, qui garantit que tout ce qui peut être déduit du graphe est également vrai dans la distribution de probabilité sous-jacente. Ce résultat sera étudié en détail et démontré dans la suite du livre.

Ce livre est organisé de la façon suivante.

La première partie, *Introduction aux réseaux bayésiens*, est une présentation intuitive de la construction des réseaux bayésiens à partir de quelques exemples simples. Dans cette partie nous abordons également l'étude des algorithmes, mais là encore de façon relativement intuitive. Cette partie se conclut par des exercices simples, qui permettent de manipuler les concepts introduits, ou encore de prendre en main un outil informatique de réseaux bayésiens.

La deuxième partie, *Cadre théorique et algorithmes*, présente une formalisation complète des réseaux bayésiens, ainsi que l'étude détaillée des algorithmes les plus importants, aussi bien pour l'utilisation de ces modèles (inférence) que pour leur construction à partir de données (apprentissage). Cette partie est très technique, car nous avons choisi de démontrer certains des résultats annoncés. Le lecteur rebuté par les longs développements techniques pourra survoler cette partie.

Dans la troisième partie, *Méthodologie de mise en œuvre et études de cas*, nous abordons l'aspect pratique de cette technologie. Le premier chapitre de cette partie, le chapitre 7 page 187, est consacré aux aspects méthodologiques en tentant de répondre aux trois questions suivantes : pourquoi, où (dans quelles applications), et comment utiliser des réseaux bayésiens ? Nous présentons ensuite plusieurs exemples d'application ayant fait l'objet

de publications, suivis de six études de cas réelles, auxquelles nous avons directement participé.

Un ensemble d'annexes (*Théorie des graphes, Probabilités, et Outils*) ainsi qu'une bibliographie et un index complètent le livre.

Écrit par une équipe combinant les points de vue de l'enseignant, du chercheur, de l'ingénieur, et de l'utilisateur final, ce livre s'adresse à un large public.

Il s'adresse aux ingénieurs et décideurs dans l'un des nombreux domaines d'application des réseaux bayésiens : santé, industrie, banque, marketing, informatique, défense, etc. Pour ce profil de lecteur, nous recommandons surtout une lecture de la première partie, éventuellement en omettant le chapitre 3 (*Exercices*), et de la troisième partie. Cette première lecture leur permettra de se faire rapidement une idée sur les possibilités d'utilisation de cette technique dans leur domaine, et leur offrira des points de comparaison avec d'autres techniques.

L'ouvrage s'adresse également aux étudiants et chercheurs du niveau deuxième ou troisième cycle dans plusieurs disciplines : statistiques, mathématiques de la décision, analyse de risque, intelligence artificielle, ainsi qu'à tous les élèves ingénieurs. Ils y trouveront une présentation intuitive des réseaux bayésiens, un développement théorique complet sur les algorithmes les plus récents, ainsi qu'une base pour des investigations complémentaires. Les exercices présentés dans la première partie leur permettront d'évaluer leur compréhension des concepts et algorithmes. Pour ce profil de lecteur, nous recommandons une lecture progressive en fonction du niveau d'approfondissement requis. Pour une prise de contact et une compréhension des concepts de base, la première partie, en incluant les exercices, pourra être suffisante. Pour une étude plus poussée des algorithmes, la deuxième partie sera un compagnon utile des notes de cours, ou un bon point de départ pour des recherches personnelles. Enfin, pour développer un projet applicatif basé sur cette technique, les exemples et études de cas présentés dans la troisième partie seront une bonne source d'inspiration.