

Industrialiser le data mining : enjeux et perspectives

Françoise Fogelman-Soulié¹

KXEN, 25 quai Galliéni, 92 158 SURESNES Cedex, France
francoise@kxen.com

Résumé. L'informatique décisionnelle est un secteur en forte croissance dans toutes les entreprises. Les techniques classiques (reporting simple & Olap), qui s'intéressent essentiellement à présenter les données, sont aujourd'hui très largement déployées. Le data mining commence à se répandre, apportant des capacités de prévision à forte valeur ajoutée pour les entreprises les plus compétitives. Ce développement est rendu possible par la disponibilité croissante de masses de données importantes et la puissance de calcul dorénavant disponible. Cependant, la mise en IJuvre industrielle des projets de data mining pose des contraintes tant théoriques (quels algorithmes utiliser pour produire des modèles d'analyses exploitant des milliers de variables pour des millions d'exemples) qu'opérationnelles (comment mettre en production et contrôler le bon fonctionnement de centaines de modèles). Je présenterai ces contraintes issues des besoins des entreprises ; je montrerai comment exploiter des résultats théoriques (provenant des travaux de Vladimir Vapnik) pour produire des modèles robustes ; je donnerai des exemples d'applications réelles en gestion de la relation client et en analyse de qualité. Je conclurai en présentant quelques perspectives (utilisation du texte et des réseaux sociaux).

Mots clés : Data mining ; robustesse ; Structural Risk Minimization ; scalabilité.

¹Françoise Fogelman Soulié est Vice President Strategic Business Development chez KXEN, après avoir été Directeur Associé & Directeur de l'Agence CRM de Business & Decision ; Directeur de l'équipe "data warehouse, data mining" de Atos Ingénierie ; Directeur Scientifique & Directeur Général de Mimetics. Précédemment, elle était Professeur au Laboratoire d'Informatique de Paris 11 Orsay. Ancienne élève de l'Ecole Normale Supérieure, elle est agrégée de mathématiques et Docteur d'Etat en Informatique (réseaux neuronaux). Elle a exercé de nombreuses fonctions d'expertise scientifique : membre du Bureau Exécutif du RNTL & évaluateur pour le programme Technologies logicielles de l'ANR ; membre du Conseil Scientifique du Groupe France Telecom ; expert auprès de la Commission Européenne. Elle a été Président de la European Neural Network Society et membre du Board of Governors de l'International Neural Network Society. Elle est auteur ou éditeur de 12 ouvrages et plus de 100 articles scientifiques. Elle a encadré une vingtaine de thèses (réseaux neuronaux). Elle est Chevalier des Palmes Académiques.