

## Data Mining et Statistique Décisionnelle L'intelligence des données

S. Tufféry

1 Vol., 554 pages, Technip  
2007, ISBN 978-2-7108-0888-6

Cette nouvelle édition revue et augmentée de plus de 150 pages garde une structure identique à la précédente avec les mêmes chapitres. Parmi les nombreux compléments répartis tout au long de l'ouvrage, citons en quelques-uns : vue d'ensemble des techniques de classification, exemple de régression linéaire montrant l'intérêt de l'étude des résidus et des visualisations graphiques, exemple détaillé de modélisation par régression logistique, régression de Poisson, développements plus importants sur les SVM, amélioration des performances d'un modèle prédictif, etc. Rappelons les différents chapitres de cet ouvrage dont la première édition a été analysée dans le numéro 1 du volume 54 de la Revue de Statistique Appliquée :

1. Panorama du Data Mining ; 2. Le déroulement d'une étude de Data Mining ; 3. L'exploration et la préparation des données ; 4. L'utilisation des données commerciales ; 5. Aperçu sur les techniques de Data Mining ; 6. L'analyse factorielle ; 7. Les réseaux de neurones ; 8. Les techniques de classification automatique ; 9. La recherche d'associations ; 10. Les techniques de classement et prédiction ; 11. Une application du Data Mining : le scoring ; 12. Les facteurs de succès d'un projet de Data Mining ; 13. Les logiciels de statistique et Data Mining ; 14. Le Text Mining ; 15. Le Web Mining ; Annexe A : Rappels de statistique ; Annexe B : Data Mining informatique et liberté ; Bibliographie ; Index.

Dans ce livre, les aspects méthodologiques vont de la conduite des projets jusqu'aux facteurs de réussite et aux pièges à éviter, en passant par l'évaluation et la comparaison de modèles, l'intégration dans le marketing de bases de données, le calcul du retour sur investissements, les interfaces informatiques, sans oublier les contraintes juridiques dès que l'on traite des données à caractère personnel.

Cet ouvrage de référence, agréable à lire malgré sa technicité, et qui comporte beaucoup d'exemples traités avec les logiciels du marché, s'adresse aux statisticiens, aux utilisateurs et gestionnaires de bases de données dans tous les domaines d'activité, ainsi qu'aux étudiants de masters et à leurs enseignants.

*Journal de la Société Française de Statistique*

*Revue de Statistique Appliquée*

105

*Vol 148, N°3, 3<sup>e</sup> trimestre 2007.*