

Recherche d'information

2^e édition

Massih-Reza Amini, professeur d'informatique à l'université Grenoble Alpes, est titulaire d'une thèse en Informatique de l'université Paris 6. Ses recherches portent sur l'apprentissage automatique appliqué aux problèmes d'accès à l'information à large échelle. Il est co-auteur de nombreux articles scientifiques notamment de l'ouvrage *Apprentissage machine paru aux éditions Eyrolles*. Il dirige actuellement l'équipe AMA dont les recherches se situent en analyse de données, modélisation et apprentissage automatique.

Éric Gaussier, professeur d'informatique à l'université Grenoble Alpes, est titulaire d'un diplôme en Mathématiques Appliquées de l'école Centrale Paris et d'une thèse en Informatique de l'université Paris 7. Ses travaux de recherche s'inscrivent dans la Science des données, au carrefour de l'apprentissage statistique, de la recherche d'information et du traitement automatique des langues. Il est co-auteur de nombreux articles et brevets dans ces domaines et dirige actuellement le laboratoire d'Informatique de Grenoble.

Le premier ouvrage francophone sur les algorithmes qui sous-tendent les technologies de big data et les moteurs de recherche !

Depuis quelques années, de nouveaux modèles et algorithmes sont mis au point pour traiter des données de plus en plus volumineuses et diverses. Cet ouvrage présente les fondements scientifiques des tâches les plus répandues en recherche d'information (RI), tâches également liées au data mining, au décisionnel et plus généralement à l'exploitation du big data.

La deuxième édition de cet ouvrage propose un exposé détaillé et cohérent des algorithmes classiques développés dans ce domaine, abordable par des lecteurs qui cherchent à connaître le mécanisme des outils quotidiens d'Internet. De plus, le lecteur approfondira les concepts d'indexation, de compression, de recherche sur le Web, de classification et de catégorisation, et pourra prolonger cette étude avec les exercices corrigés proposés en fin de chapitre.

Ce livre s'adresse tant aux chercheurs et ingénieurs qui travaillent dans le domaine de l'accès à l'information et employés de PME qui utilisent en profondeur les outils du webmarketing, qu'aux étudiants de Licence, Master, écoles d'ingénieurs ou doctorants qui souhaitent un ouvrage de référence sur la recherche d'information.

Sommaire

Représentation et indexation. Prétraitements linguistiques. Les deux lois de base en recherche d'information. Représentation documentaire. Index inversé. **Recherche d'information.** Modèles de recherche. Expansion de requêtes. Mesures d'évaluation. **Recherche sur le Web.** Architecture de la Toile. Trois inventions à la base du Web. Collecte et indexation des pages sur la Toile. Nouvelles stratégies de recherche. Calcul approché de similarités en grande dimension. **Catégorisation de documents.** Formalisme. Sélection de variables. Modèles génératifs. Modèles discriminants. Mesures d'évaluation. **Partitionnement de documents.** Les étapes du partitionnement. Principaux algorithmes de partitionnement. Évaluation. Applications à l'accès à l'information. **Réseaux de neurones profonds.** Neurone formel. Quelques réseaux. Applications en RI. **Recherche de thèmes latents.** Analyse sémantique latente. Analyse sémantique latente probabiliste. Le modèle LDA. **Considérations pratiques.** Logiciels libres pour la recherche d'information. Logiciels libres pour la catégorisation et le partitionnement.

Code éditeur : 667376
ISBN : 978-2-212-67376-0



39 €

www.editions-eyrolles.com
Groupe Eyrolles | Diffusion Geodif

Recherche d'information

M.-R. Amini
É. Gaussier



EYROLLES

Algorithmes

2^e édition

R

Recherche d'information

Applications, modèles et algorithmes

Massih-Reza Amini - Éric Gaussier

Préface de Stephen Robertson

2^e édition

Data mining, décisionnel et big data

