



Société Francophone de Classification

Bulletin numéro 26

Octobre 2014

Président : M. Nadif

Vice-président : P. Kuntz-Cosperec & M. Rémon

Secrétaire : M. Lebbah

Trésorier : N. Rogovschi

Éditeur : G. Cleuziou

Société Francophone de Classification, 45 rue des Saints Pères, 75270 Paris Cedex 06, France <http://www.sfc-classification.net/>

Correspondance : M. Nadif, LIPADE, Université Paris Descartes, 45 rue des Saints Pères, 75270 Paris Cedex 06, France

Ce bulletin est le vôtre, faites-le vivre en envoyant toutes les informations que vous jugerez utiles à la communauté de la SFC, à l'éditeur en utilisant l'adresse suivante : guillaume.cleuziou@univ-orleans.fr

Sommaire

Le mot du président	1
Compte-rendu de l'AG 2014	2
Thèses et habilitations	2
Rencontres SFC 2015	3
Conférences à venir	3
Appel Prix S. Régnier 2015	5
Parutions	5
Listes de diffusion SFC	7

communautés, il en va de l'intérêt de notre thématique qui a une large place dans les challenges actuels.

J'entame mon deuxième mandat et je souhaiterais que chaque année soit l'occasion de nous rencontrer au moins deux fois. Je suis convaincu que nous ne devrions pas attendre les prochaines journées de classification qui auront lieu à Nantes au mois de septembre 2015. L'an dernier, nous avons organisé une journée le 02 avril en hommage à Yves leChevallier sur les "stratégies de classification non-supervisés", ce fut une réussite. J'encourage donc vivement toutes ces actions et toute autre suggestion pour promouvoir la vie de notre société sera la bienvenue. En janvier 2015, une demi-journée sera consacrée aux systèmes de recommandation, soyez nombreux !

Pour conclure, je vous souhaite une très bonne lecture de ce bulletin et vous donne rendez-vous au plus tard à Nantes lors des XXIIèmes Rencontres de la Société Francophone de Classification 2015.

Mohamed Nadif
Président de la SFC

Le mot du président

Chers collègues, Chers amis,

Je tiens tout d'abord à adresser mes plus vifs remerciements à toutes celles et ceux qui participent directement en tant que membres du conseil d'administration ou indirectement en tant qu'adhérents à la vie de notre société. Et je vous remercie pour la confiance que vous accordez aux membres du bureau.

Comme vous le savez, les XXI journées ont eu lieu à Rabat au Maroc et ce fut un succès. Au cours de cette manifestation, nous avons eu une pensée pour Marc Csernel qui nous a quitté l'été dernier. L'organisation a été confiée à la nouvelle société marocaine de classification (SMC) qui vient d'être créée et qui fait désormais partie de l'IFCS. Nous avons beaucoup travaillé ensemble sur ce projet qui s'est concrétisé et nous espérons qu'un élargissement au Maghreb se fasse dès que possible. Nous devons donc œuvrer ensemble pour tisser des liens étroits avec la SMC voire d'autres sociétés de classification et d'autres

Compte-rendu de l'Assemblée Générale de la SFC

Rabat, Maroc - 10 septembre 2014

L'Assemblée Générale est ouverte sous la présidence de Mohamed Nadif. Un point sur la conférence SFC 2014 est présenté par les organisateurs locaux. Au niveau des statistiques il y a eu plus de soumission de papiers que pour l'édition 2013. Les assemblées générales de la SFC et de la SMC se tenaient au même moment (dans des salles différentes). Le secrétaire de la SFC présente les différents points de l'ordre du jour.

1. Rapport moral (M. Lebbah)

(a) Membres de la SFC

- En 2014, la Société Marocaine de Classification (SMC) a organisé sa première assemblée à l'occasion des rencontres de la SFC. Un protocole de collaboration scientifique sera signé entre la SFC et la SMC.
- La SFC est membre fondateur de l'IFCS. Des informations sur la SFC sont reprises dans chaque bulletin de l'IFCS. Des membres de la SFC participent au Conseil de l'IFCS.

(b) À l'occasion du départ en retraite de Yves Lechevallier, une journée thématique intitulée "Stratégies de Classification Non-Supervisées", a été organisée à l'IHP Paris le 2 avril 2014. <https://sites.google.com/site/journeesfc2014>.

(c) Site Web et liste de diffusion

- Le site internet de la SFC est pris en charge par Guillaume Cleuziou et Lydia Boudjeloud (<http://www.sfc-classification.net>). Un nouveau site est en cours de finalisation.
- Une seule liste de diffusion large (ouverte à toutes les personnes intéressées par la classification) : liste-sfc@sfc-classification.net. Cette liste est modérée par Guillaume Cleuziou.

La liste de diffusion a déjà été largement utilisée en 2014. Si vous n'êtes pas encore abonné, n'hésitez pas à contacter son responsable (Guillaume Cleuziou) afin de vous y inscrire. Vous pouvez également lui communiquer toutes les informations utiles que vous souhaitez diffuser via ces listes.

(d) Divers

Élections au Conseil d'Administration de la SFC :

- Président : Mohamed Nadif (Pr. Univ. Paris Descartes)
- Postes de vice-président : Pascale Kuntz (Pr. Univ. Nantes) et Marcel Remon (Pr. Univ. Namur)
- Trésorière : Nicoleta Rogovschi (MCF. Univ. Paris Descartes)

Les 4 membres du bureau ont été élus à l'unanimité.

Les mandats d'éditeur du bulletin de la SFC et de secrétaire de la SFC arrivent à échéance en 2015. Habituellement, l'Assemblée Générale de la SFC se tient lors des rencontres de la SFC.

2. Cotisation

Il est décidé que le montant de la cotisation SFC (et la cotisation conjointe SFC/SFdS) reste inchangée pour 2014 (16-€).

3. Rapport financier pour l'année 2014 (N. Rogovschi)

Recettes	
Cotisations perçues durant l'année 2013	453,00
TOTAL	453,00
Dépenses	
Frais bancaires (tenue de compte)	6,50
Hébergement de compte et obtention d'une adresse internet	47,68
Frais divers	100,00
TOTAL	154,18
Avoir de la SFC au 1er janvier 2013	14 485,67
Avoir de la SFC au 31 décembre 2013	14 784,49

4. Une journée thématique est prévue en janvier 2015 sur les "systèmes de recommandation". Elle sera organisée par Pascale Kuntz et Mohamed Nadif.

5. Prix Simon Régnier

La SFC lance l'appel à candidature pour le prix Simon Régnier qui sera décerné à l'occasion des rencontres de la SFC en 2015.

6. SFC 2015

Les 22èmes Rencontres de la Société Francophone de Classification (SFC) seront organisées par le Laboratoire d'Informatique de Nantes Atlantique (LINA) de l'université de Nantes au tout début du mois de septembre 2015. La présidente du comité de programme est Pascale Kuntz (LINA - Université de Nantes). Une ouverture interdisciplinaire en Sciences Humaines et Sociales est envisagée.

Mustapha LEBBAH
Secrétaire de la SFC

Thèses et Habilitations

Voici une liste non exhaustive de thèses et habilitations à diriger les recherches soutenues ces derniers mois dans des domaines liés à la classification.

- HDR de **Fatma Bouali** : "*Contribution à la fouille visuelle de données : de OLAP aux OPEN DATA*", soutenue en octobre 2014 (Université de Tours)
- Doctorat de **Dialla Azzedine** : "*A Contribution to Topological Learning and its Application in Social Networks*", soutenu en octobre 2014 (Dir. Djamel A. Zighed - Université de Lyon 2)
- Doctorat de **Vincent Brault** : "*Estimation et sélection de modèle pour le modèle des blocs latents*", soutenu en septembre 2014 (Dir. Gilles Celeux et Christine Keribin - Université Paris Sud)

- Doctorat de **Bruno Belatre** : "*Extraction, analyse et utilisation de relations spatiales entre objets d'intérêt pour une analyse d'images de télédétection guidée par des connaissances du domaine*", soutenu en septembre 2014 (Dir. Cédric Wemmert - Université de Strasbourg)
- Doctorat de **Ramazan gokberk Cinbis** : "*Classification d'images et localisation d'objets par des méthodes de type noyau de Fisher*", soutenu en juillet 2014 (Dir. Cordelia Schmid et Jakob Verbeek - Université de Grenoble)
- Doctorat de **Alina Dana Miron** : "*Multi-modal, Multi-Domain Pedestrian Detection and Classification : Proposals and Explorations in Visible over Stereo Vision, FIR and SWIR*", soutenu en juillet 2014 (Dir. Abdelaziz Bensrhair - Co-Tutelle INSA Rouen / Université Babes-Bolyai en Roumanie)
- Doctorat de **Minh Thuy Ta** : "*Techniques d'optimisation non convexe basée sur la programmation DC et DCA et méthodes évolutives pour la classification non supervisée*", soutenu en juillet 2014 (Dir. Hoai An Le Thi - Université de Lorraine)
- HDR de **Khalid Benabdeslem** : "*Contributions en apprentissage semi-supervisé : modélisation, classification et sélection*", soutenue en juin 2014 (INSA - Université de Lyon 1)
- Doctorat de **Chiheb-Eddine Ben N'Cir** : "*Efficient k-means based methods for overlapping clustering*", soutenu en juin 2014 (Dir. Nadia Essoussi - ISG - Université de Tunis)
- Doctorat de **Mory Ouattara** : "*Développement et mise en place d'une méthode de classification multi-blocs : application aux données de l'OQAI*", soutenu en mars 2014 (Dir. Fouad Badran, Ndèye Niang-Kéïta et Corinne Mandin - CNAM Paris)
- Doctorat de **Roberto D'Ambrosio** : "*Classification de bases de données déséquilibrées par des règles de décomposition*", soutenu en mars 2014 (Dir. Michel Barlaud et Giulio Iannello - Co-Tutelle Université de Nice / Université de Rome)
- Doctorat de **Florian Maire** : "*Détection et classification de cibles multispectrales dans l'infrarouge*", soutenu en février 2014 (Dir. Randal Douc - Telecom SudParis)
- Doctorat de **Zeynep Akata** : "*Contributions à l'apprentissage grande échelle pour la classification d'images*", soutenu en janvier 2014 (Dir. Cordelia Schmid - Université de Grenoble)

Rencontres de la SFC en 2015

La SFC organise, chaque année, les Rencontres de la Société Francophone de Classification, qui ont pour objectifs de présenter des résultats récents, des applications originales en classification ou dans des domaines connexes,

ainsi que de favoriser les échanges scientifiques à l'intérieur de la société et de faire connaître à divers partenaires extérieurs les travaux de ses membres. Durant ces rencontres qui rassemblent régulièrement une centaine de participants, est attribué le prix Simon Régnier, consacrant une contribution originale d'un jeune docteur à la classification.

En 2015, le Laboratoire d'Informatique de Nantes Atlantique (LINA) de l'université de Nantes se mobilise pour organiser les Rencontres de la Société Francophone de Classification qui se dérouleront du 9 au 11 septembre 2015.

- Présidente du Comité de Programme : Pascale Kuntz (LINA - Polytech' Nantes)

Conférences à venir

ICPRAM 2015

4th International Conference on Pattern Recognition Applications and Methods
January 10-12, 2015 in Lisbon, Portugal

The International Conference on Pattern Recognition Applications and Methods would like to become a major point of contact between researchers, engineers and practitioners on the areas of Pattern Recognition, both from theoretical and application perspectives.

Contributions describing applications of Pattern Recognition techniques to real-world problems, interdisciplinary research, experimental and/or theoretical studies yielding new insights that advance Pattern Recognition methods are especially encouraged.

ICPRAM 2015 will be held in conjunction with ICORES 2015 and ICAART 2015.

Important dates :

- Deadline for position papers : October 9, 2014
- Notification of acceptance : November 6, 2014
- Camera Ready and Registration : November 17, 2014

www.icpram.org/

EGC 2015

Conférence francophone sur l'Extraction et la Gestion des Connaissances
27 - 30 janvier 2015 au Luxembourg

La 15ème conférence "Extraction et Gestion des Connaissances" (EGC 2015) se déroulera au Luxembourg du 27 au 30 janvier 2015.

La conférence annuelle de l'association Extraction et Gestion des Connaissances (EGC) réunit chercheurs, industriels et étudiants dans un ensemble de domaines connexes regroupant notamment Fouille de Données, Gestion des Connaissances, Visualisation d'Informations, Sciences et Applications Cognitives, Apprentissage Automatique, Bases de Données, Statistiques.

La multiplication des producteurs d'information (humains, capteurs, ordinateurs ?) constitue en effet une des tendances de fond du monde actuel qui se connecte et se numérise à un rythme exponentiel. Cette évolution rend indispensable l'utilisation d'outils informatiques pour interagir avec ces masses de données complexes, dynamiques, hétérogènes et de qualité variable.

La conférence EGC est le lieu privilégié d'échange entre chercheurs et industriels en vue de discuter des innovations et des résultats scientifiques les plus récents, de préciser les besoins, de faire émerger les enjeux du domaine et de contribuer de manière plus générale à la formation d'une communauté d'excellence, en particulier dans le monde francophone, autour de la problématique de l'extraction et de la gestion des connaissances.

Les articles peuvent être soumis en français ou en anglais (pour les non-francophones). Une session spéciale sur les applications dans le domaine big data analytics sera cette année organisée totalement en anglais.

- Président d'Honneur : David Taniar, Université de Monash, Australie/Monash University, Australia
- Président du Comité de Programme : Jérôme Darmont, ERIC, Université de Lyon, France
- Présidents du Comité d'Organisation : Benoît Otjacques et Thomas Tamisier, CRP Gabriel Lippmann, Luxembourg

Dates importantes :

- Soumission des résumés : 10 octobre 2014
- Soumission des articles : 17 octobre 2014
- Notification aux auteurs : 22 novembre 2014
- Version finale des articles : 10 décembre 2014

<http://egc2015.lippmann.lu>

IFCS 2015

Conference of the International Federation of Classification Societies

July 6-8, 2015 in Bologna, Italy

The 2015 conference of the International Federation of Classification Societies (IFCS) will be organized jointly by the Department of Statistics Paolo Fortunati and the SIS-ClaDAG group at the University of Bologna, Italy from July 6 to July 8, 2015.

www2.stat.unibo.it/ifcs2015

ICML 2015

32nd International Conference on Machine Learning
July 6-11, 2015 in Lille, France

ICML is the leading international machine learning conference and is supported by the International Machine Learning Society (IMLS).

The 32nd International Conference on Machine Learning (ICML 2015) will be held in Lille, France, from July 6 to 11, 2015. The conference will consist of one day of

tutorials, followed by three days of main conference sessions, followed by two days of workshops. We invite submissions of papers on all topics related to machine learning for the conference proceedings, and proposals for tutorials and workshops.

This year, ICML will adopt a single reviewing cycle, with a single paper deadline on February 6th, 2015.

<http://icml.cc/2015/>

SDA 2015

5th Workshop in Symbolic Data Analysis
November 16th-19th, 2015 in Orléans, France

Symbolic Data Analysis is a relatively recent field (less than two decades old) that focuses on the analysis of datasets where individuals are described by variables that can represent internal variation and/or structure. Symbolic data values can be intervals, histograms, distributions, lists of values, taxonomies, etc.

Some datasets naturally consist of symbolic data (e.g. the dataset of the minimum and maximum temperatures can be naturally represented by intervals), while many other symbolic datasets result from the aggregation of large classical datasets (the dataset of the distribution incomes of the population in each country). It quickly becomes apparent that variation and structure must be taken into account when analysing the dataset.

The catalogue of symbolic data analysis methods is increasing with time covering new areas and offering more rigorous approaches. An extensive coverage of earlier symbolic data analysis methods can be found in Bock and Diday (2000), Billard and Diday (2006) and Diday and Noirhomme-Fraiture (2008). Many other methods can be found disseminated in journal articles and conference presentations. Noirhomme-Fraiture and Brito (2011) offer an up-to-date review of the field. Symbolic methods have been successfully applied in different fields such as Official Statistics, marketing, industry and finance to name a few. The present dynamic development of the field is mainly due to the growth of the symbolic community. Since its beginnings, many researchers with different backgrounds (mathematicians, statisticians, computer scientists, economists, among others) have contributed to the field, which has benefited from that heterogeneity and is now well-established. Many work has been done, but many challenges lie ahead.

The workshop presents itself as an excellent opportunity for crossover of ideas between the community members and to open the way to future cooperation between them.

Prix Simon Régnier 2015

Le Prix Simon Régnier 2015 sera décerné lors des 22èmes Rencontres de la SFC (Nantes) à un jeune chercheur de moins de 35 ans au 1er janvier 2015 et ayant soutenu une thèse dans le domaine de la classification, lequel devra exposer ses travaux durant les rencontres.

Son montant est de 450 euros. Le lauréat est défrayé de droits d'inscription au congrès SFC 2015 et de cotisation à la SFC pour 2016.

Pour candidater, il faut fournir un dossier composé de :

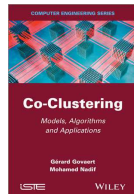
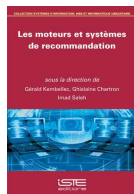
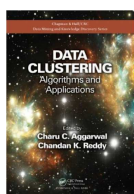
- un CV précisant la liste des membres du jury de la thèse et incluant la liste des titres et travaux,
- un article représentatif de la recherche effectuée,
- la communication présentée aux Rencontres de la SFC 2015.

Enfin, le lauréat devra être membre de la SFC en 2015 : inscription sur place possible durant le congrès.

Ce dossier doit être envoyé par courrier électronique à guillaume.cleuziou@univ-orleans.fr. Le sujet du courrier doit être "Prix Simon Régnier".

Nous remercions l'ensemble des collègues de la communauté scientifique d'encourager les jeunes docteurs méritants à candidater pour ce prix.

Parutions



Les moteurs et systèmes de recommandation

Ouvrage sous la direction de Gérard Kembellec, Ghislaine Chartron et Imad Saleh chez ISTE Éditions dans la collection "Systèmes d'information, web et informatique ubiquitaire".
2014, 228 p.
ISBN : 978-1-78405-041-2

Plébiscités par les différentes plateformes de contenus (livres, musiques, films) et par les sites de ventes en ligne, les systèmes de recommandation sont des éléments-clés des stratégies numériques. Si leur développement visait à l'origine la performance des systèmes d'information, les enjeux se sont aujourd'hui massivement déplacés sur des logiques d'optimisation de la relation client, avec comme objectif principal de maximiser les ventes potentielles.

D'approche transdisciplinaire, Les moteurs et systèmes de recommandation réunit des contributions liant sciences de l'information et de la communication, marketing, sociologie, mathématiques et informatique. Il traite de la compréhension des modèles sous-jacents aux systèmes de recommandation et détaille leur mise en perspective historique. Il analyse également leur développement dans les offres de contenus et évalue leur impact sur le comportement des usagers.

Mots-clés : Autorité, intelligence économique, légitimité, Systèmes de recommandation, web social.

Fouille de Données et Humanités Numériques

MASHS 2014. Revue des Nouvelles Technologies de l'Information vol.RNTI-SHS-2

Rédacteur invité : Nicolas Turenne

Éditions Hermann

2014, 232 p.

ISBN : 978-2-7056-8953-7

Le but de ce numéro spécial est de regrouper des contributions autour de la thématique de l'extraction et gestion de connaissances en interface avec les humanités numériques. Nous sommes spectateurs et acteurs d'un tournant sociétal qui est admis aujourd'hui, c'est le "tournant du numérique" transformant l'environnement politique, économique et culturel contemporain. L'implication de cette contribution du numérique à nos sociétés tend à fabriquer des espaces de données de plus en plus importants nécessitant à cet endroit des traitements spécifiques. Les disciplines des sciences de l'homme et de la société (SHS), qui placent l'homme au coeur des questionnements, ont jusque là peu profité des méthodes traitant des ensembles de données numériques et du potentiel de transformation de ce tournant. Quelles nouvelles questions peuvent se poser sur de nouvelles données, selon quels modes d'archivage, avec quels algorithmes de recherche d'information utile "dans une botte de documents" ? Pour ne citer que deux exemples, les humanités numériques s'invitent déjà en sociologie, qui aborde la démocratie d'opinion dans les réseaux sociaux, et en histoire où la lecture numérique change l'accès à la matérialité du passé. La sphère "humanités numériques" peut être vue comme un nouvel espace épistémique de réflexion qui invite à de nouvelles questions mais en incluent d'anciennes. Ainsi la question du traitement des données linguistiques s'est constituée en communauté de recherche depuis au moins les années cinquante sous l'auspice de la cybernétique naissante. La sociologie qualitative a intégré aussi, via les réseaux sociaux, des éléments d'analyse pour qualifier l'espace des controverses scientifiques ou bien procéder aux traitements d'enquêtes avec des catégories spécifiées. L'ampleur des media sociaux sur internet vient consolider cette sociologie digitale. Le trait d'union entre les différents univers abordés par les humanités numériques concerne la production de données, comprenant totalement ou partiellement du matériau textuel et issues du monde des SHS. Leur nature est culturelle, littéraire, philosophique, économique, psychologique, sociologique, d'opinion, ethnologique, démographique, géographique, historique, musicologique, archéologique, architecturale, juridique, théologique ou pédagogique. Des algorithmes doivent être mobilisés pour traiter cette masse d'information qui elle même doit être stockée de façon à faciliter ces traitements.

Mots-clés : Sociologie, Linguistique, Communication, Histoire, Archivage, Géographie.

Data Clustering : Algorithms and Applications

Ouvrage sous la direction de Charu C. Aggarwal et Chandan K. Reddy
Chapman and Hall/CRC Edts.
2013, 652 p.
ISBN : 978-1466558212

Research on the problem of clustering tends to be fragmented across the pattern recognition, database, data mining, and machine learning communities. Addressing this problem in a unified way, *Data Clustering : Algorithms and Applications* provides complete coverage of the entire area of clustering, from basic methods to more refined and complex data clustering approaches. It pays special attention to recent issues in graphs, social networks, and other domains. The book focuses on three primary aspects of data clustering : *Methods* , describing key techniques commonly used for clustering, such as feature selection, agglomerative clustering, partitional clustering, density-based clustering, probabilistic clustering, grid-based clustering, spectral clustering, and nonnegative matrix factorization *Domains* , covering methods used for different domains of data, such as categorical data, text data, multimedia data, graph data, biological data, stream data, uncertain data, time series clustering, high-dimensional clustering, and big data *Variations and Insights* , discussing important variations of the clustering process, such as semisupervised clustering, interactive clustering, multiview clustering, cluster ensembles, and cluster validation In this book, top researchers from around the world explore the characteristics of clustering problems in a variety of application areas. They also explain how to glean detailed insight from the clustering process including how to verify the quality of the underlying clusters through supervision, human intervention, or the automated generation of alternative clusters.

Probabilités pour l'ingénieur

des fondements aux calculs
Dominique Pastor - Christophe Sintès
Editions Lavoisier
Hermès
2014, 374 p.
ISBN : 978-2-7462-3820-6

Les probabilités et les statistiques occupent une place toujours plus importante dans tous les domaines scientifiques et technologiques. Cependant, l'enseignement des probabilités se heurte à deux difficultés principales : la familiarisation avec les phénomènes aléatoires et la mise en place de l'outillage mathématique spécifique au calcul des probabilités.

Cet ouvrage a pour objectif de fournir un ensemble d'outils mathématiques permettant d'écrire et de calculer les probabilités. Basé sur l'expérience pratique des auteurs, *Probabilités pour l'ingénieur* propose de nombreux exemples et discussions afin d'amener le lecteur à une connaissance des probabilités qui repose sur une version

très opératoire de la théorie de la mesure et de l'intégration de Lebesgue. En effet, l'intégrale de Lebesgue reste l'outil théorique le plus adapté à la manipulation des probabilités. Il permet de traiter efficacement des problèmes d'ingénierie et de recherche.

Mots-clés : Probabilités, processus stochastiques, statistiques.

Théories de l'incertain et fusion de données multicapteurs

Alain Appriou
Directeurs de Collection : SEE , DUBUISSON Bernard
Editions Lavoisier
Hermès
2014, 248 p.
ISBN : 9782746231986

En matière d'observation, l'exploitation coordonnée de capteurs multiples permet de répondre aux exigences croissantes des systèmes complexes, que ce soit en termes d'environnement ou de besoins opérationnels. Tirer profit de la diversité des informations, des capacités complémentaires de leurs différentes sources requiert néanmoins le développement de fonctionnalités spécifiques et l'intégration de données imparfaites. Dans ce contexte, cet ouvrage considère l'ensemble des problèmes soulevés par l'élaboration de traitements multicapteurs. Il propose une suite de modules génériques dont l'association permet de fournir une chaîne complète de traitement adaptable à chaque situation.

L'efficacité des méthodes dégagées repose sur l'aptitude des différentes théories de l'incertain à n'extraire et à n'utiliser que la partie pertinente des données disponibles, grâce au potentiel propre de chacune de ces théories, mais surtout grâce aux synergies qu'il est possible d'établir entre elles.

Co-Clustering : Models, Algorithms and Applications

Gérard Govaert, Mohamed Nadif
Computer Engineering Series
Wiley
2013, 256 p.
ISBN : 978-1-84821-473-6

Cluster or co-cluster analyses are important tools in a variety of scientific areas. The introduction of this book presents a state of the art of already well-established, as well as more recent methods of co-clustering. The authors mainly deal with the two-mode partitioning under different approaches, but pay particular attention to a probabilistic approach. Chapter 1 concerns clustering in general and the model-based clustering in particular. The authors briefly review the classical clustering methods and focus on the mixture model. They present and discuss the use of different mixtures adapted to different types of data. The algorithms used are described and related works with

different classical methods are presented and commented upon. This chapter is useful in tackling the problem of co-clustering under the mixture approach. Chapter 2 is devoted to the latent block model proposed in the mixture approach context. The authors discuss this model in detail and present its interest regarding co-clustering. Various algorithms are presented in a general context. Chapter 3 focuses on binary and categorical data. It presents, in detail, the appropriated latent block mixture models. Variants of these models and algorithms are presented and illustrated using examples. Chapter 4 focuses on contingency data. Mutual information, phi-squared and model-based co-clustering are studied. Models, algorithms and connections among different approaches are described and illustrated. Chapter 5 presents the case of continuous data. In the same way, the different approaches used in the previous chapters are extended to this situation.

Keywords : Cluster Analysis, Model-Based Co-Clustering.

Liste de diffusion de la SFC

La liste-sfc@sfc-classification.net compte actuellement plus de 120 membres, destinataires d'une dizaine de messages par mois. Nous nous efforçons de modérer au mieux cette liste en ne sélectionnant que les informations ayant trait au domaine de la classification, de manière à ne pas faire doublon avec d'autres listes de diffusion plus génériques.

Abonnement :

Vous pouvez vous abonner à la liste-sfc en envoyant un email (sans contenu ni sujet) à liste-sfc-subscribe@sfc-classification.net. Une demande de confirmation vous sera automatiquement adressée (il suffit simplement de faire un "reply" sur ce dernier email).

Diffusions :

Vous pouvez envoyer des informations (appels à communication, congrès...) sur la liste en écrivant un message à : liste-sfc@sfc-classification.net.

Retrouvez toutes ces informations dans la rubrique "Listes de diffusion" sur le site web de la SFC

<http://www.sfc-classification.net>