

## Sélection des caractéristiques dans des espaces à grandes dimensions et applications

Djemel ZIOU

Le développement des technologies d'information et de communication a conduit à une forte production de données de tout genre. L'exploitation de ces données apporte une valeur ajoutée aux différents secteurs d'activités tels que la recherche, la santé, la culture, l'éducation et la sécurité. Ainsi, nous assistons ces dernières années à un développement formidable de méthodologies, de modèles, d'algorithmes et de systèmes pour l'analyse de données. Cependant, la dimension de l'espace de données et la sélection de caractéristiques sont toujours des obstacles. L'objectif de la conférence est d'introduire des méthodologies statistiques et des algorithmes pour analyser les données dans des espaces à grandes dimensions en choisissant les caractéristiques pertinentes. Après un tour d'horizon des critères de sélection existants, nous développerons une méthodologie d'apprentissage non-supervisée de mixture de densités de Dirichlet généralisées en identifiant les variables indépendantes. L'apprentissage non supervisé se base sur le paradigme du plus court message (MML). Grâce à la flexibilité de la densité de Dirichlet, la méthodologie proposée s'applique à la fois aux cas des données gaussiennes et non-gaussiennes. Cette méthodologie et les algorithmes résultants seront illustrés dans le contexte d'une nouvelle problématique à savoir la suggestion par le contenu de documents visuels.

La conférence s'appuiera sur nos travaux récents. Voici quelques exemples.

R. Ksantini, D. Ziou, B. Colin, and F. Dubeau. Weighted Pseudometric Discriminatory Power Improvement Using a Bayesian Logistic Regression Model Based on a Variational Method. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30(2), pp. 253-266, 2008.

S. Boutemedjet, and D. Ziou. A Graphical Model for Context-Aware Visual Content Recommendation. *IEEE Transactions on Multimedia* 10(1), pp. 52-62, 2008.

R. Ksantini, D. Ziou, B. Colin, and F. Dubeau. A Bayesian Kernel Logistic Discriminant Model: An Improvement to the Kernel Fisher's Discriminant. *AAAI*, 2008.

S. Boutemedjet, D. Ziou, and N. Bouguila (2007). Unsupervised Feature Selection for Accurate Recommendation of High-Dimensional Image Data. *NIPS*, 2007.

R. Ksantini, D. Ziou, B. Colin, F. Dubeau. Logistic Regression Models for a Fast CBIR Method Based on Feature Selection. *IJCAI*, pp. 2790-2795, 2007.

S. Boutemedjet, D. Ziou, and N. Bouguila. A Graphical Model for Content Based Image Suggestion and Feature Selection. ECML/PKDD, pp. 30-41, 2007.

S. Boutemedjet and D. Ziou (2007). Content-based Collaborative Filtering Model for Scalable Visual Document Recommendation, IJCAI, pp. 11-18, 2007.

N. Bouguila and D. Ziou. High-Dimensional Unsupervised Selection and Estimation of a Finite Generalized Dirichlet Mixture Model Based on Minimum Message Length, IEEE Transactions on Pattern Analysis and Machine Intelligence 19(10), pp. 1716-1731, 2007.

A. El Zaart and D. Ziou. Statistical Modeling of Heterogeneous MultiModel Image Histogram Using Parametric Distribution. Int. Journal of Remote Sensing, 28(10), pp. 2277-2294, 2007.

ML Kherfi and D. Ziou. Image Collection Modeling and Application to Indexing, Browsing and Semantic Retrieval. IEEE Transactions on Multimedia, 9(4), pp. 893-900, 2007.

N. Bouguila and D. Ziou. A Hybrid SEM Algorithm for High-Dimensional Unsupervised Learning Using a Finite Generalized Dirichlet Mixture. IEEE Transactions on Image Processing, vol. 15, pp. 2657- 2668, 2006.

N. Bouguila and D. Ziou Unsupervised Selection of a Finite Dirichlet Mixture Model: An MML-Based Approach. IEEE Transactions on Knowledge and Data Engineering , vol. 18, pp. 993-1009, 2006.

M.L. Kherfi and D. Ziou. Relevance Feedback for CBIR: A New Approach Based on Probabilistic Feature Weighting with Positive and Negative Examples. IEEE Transactions on Image Processing, vol. 15, pp. 1017- 1030, 2006.

N. Bouguila, D. Ziou, and J. Vaillancourt. Maximum Likelihood Estimation of the Generalized Dirichlet Mixture. IEEE Transactions on Image Processing (13), 1533- 1543, 2004.

M.L. Kherfi, D. Ziou, and A. Bernardi. Image Retrieval from the World Wide Web : Issues, Strategies, and Systems. ACM Computing Surveys (36), 35-67, 2004.