

## Table of Contents – Part II

### Clustering

Semi-supervised Clustering Ensemble Evolved by Genetic Algorithm for Web Video Categorization .....	1
<i>Amjad Mahmood, Tianrui Li, Yan Yang, and Hongjun Wang</i>	
A Scalable Approach for General Correlation Clustering .....	13
<i>Yubo Wang, Linli Xu, Yucheng Chen, and Hao Wang</i>	
A Fast Spectral Clustering Method Based on Growing Vector Quantization for Large Data Sets .....	25
<i>Xiuju Wang, Xiao Zheng, Feng Qin, and Baohua Zhao</i>	
A Novel Deterministic Sampling Technique to Speedup Clustering Algorithms .....	34
<i>Sanguthevar Rajasekaran and Subrata Saha</i>	
Software Clustering Using Automated Feature Subset Selection .....	47
<i>Zubair Shah, Rashid Naseem, Mehmet A. Orgun, Abdun Mahmood, and Sara Shahzad</i>	
The Use of Transfer Algorithm for Clustering Categorical Data .....	59
<i>Zhengrong Xiang and Lichuan Ji</i>	
eDARA: Ensembles DARA .....	71
<i>Chung Seng Kheau, Rayner Alfred, and HuiKeng Lau</i>	

Efficient Mining Maximal Variant and Low Usage Rate Biclusters without Candidate Maintenance in Real Function-Resource Matrix: The DeCluster Algorithm .....	83
<i>Lihua Zhang, Miao Wang, Zhengjun Zhai, and Guoqing Wang</i>	

### Association Rule Mining

MEIT: Memory Efficient Itemset Tree for Targeted Association Rule Mining .....	95
<i>Philippe Fournier-Viger, Espérance Mwamikazi, Ted Gueniche, and Usef Faghhihi</i>	

### Pattern Mining

Mining Frequent Patterns in Print Logs with Semantically Alternative Labels .....	107
<i>Xin Li, Lei Zhang, Enhong Chen, Yu Zong, and Guandong Xu</i>	

Minimising $K$ -Dominating Set in Arbitrary Network Graphs .....	120
<i>Guangyuan Wang, Hua Wang, Xiaohui Tao, Ji Zhang, and Jinhua Zhang</i>	

## Regression

Logistic Regression Bias Correction for Large Scale Data with Rare Events .....	133
<i>Zhen Qiu, Hongyan Li, Hanchen Su, Gaoyan Ou, and Tengjiao Wang</i>	
An Automatical Moderating System for FML Using Hashing Regression .....	145
<i>Peichao Zhang and Minyi Guo</i>	
Batch-to-Batch Iterative Learning Control Based on Kernel Independent Component Regression Model .....	157
<i>Ganping Li, Jun Zhao, Fuyang Zhang, and Zhizhen Ni</i>	

## Prediction

Deep Architecture for Traffic Flow Prediction .....	165
<i>Wenhai Huang, Haikun Hong, Man Li, Weisong Hu, Guojie Song, and Kunqing Xie</i>	
Compact Prediction Tree: A Lossless Model for Accurate Sequence Prediction .....	177
<i>Ted Gueniche, Philippe Fournier-Viger, and Vincent S. Tseng</i>	
Generalization of Malaria Incidence Prediction Models by Correcting Sample Selection Bias .....	189
<i>Orlando P. Zacarias and Henrik Boström</i>	
Protein Interaction Hot Spots Prediction Using LS-SVM within the Bayesian Interpretation .....	201
<i>Juhong Qi, Xiaolong Zhang, and Bo Li</i>	

Predicting the Survival Status of Cancer Patients with Traditional Chinese Medicine Symptom Variation Using Logistic Regression Model .....	211
<i>Min Wan, Liying Fang, Mingwei Yu, Wenshuai Cheng, and Pu Wang</i>	

## Feature Extraction

Exploiting Multiple Features for Learning to Rank in Expert Finding .....	219
<i>Hai-Tao Zheng, Qi Li, Yong Jiang, Shu-Tao Xia, and Lanshan Zhang</i>	
Convolution Neural Network for Relation Extraction .....	231
<i>ChunYang Liu, WenBo Sun, WenHan Chao, and WanXiang Che</i>	

Extracting Fuzzy Rules from Hierarchical Heterogeneous Neural Networks for Cardiovascular Diseases Diagnosis . . . . .	243
--	-----

*YuanLian Cui and MingChui Dong*

<i>kDMI: A Novel Method for Missing Values Imputation Using Two Levels of Horizontal Partitioning in a Data set . . . . .</i>	250
---	-----

*Md. Geaur Rahman and Md Zahidul Islam*

## Identification

Traffic Session Identification Based on Statistical Language Model . . . . .	264
--	-----

*Xinyan Lou, Yang Liu, and Xiaohui Yu*

Role Identification Based on the Information Dependency Complexity . . . . .	276
--	-----

*Weidong Zhao, Haitao Liu, and Xi Liu*

Detecting Professional Spam Reviewers . . . . .	288
---	-----

*Junlong Huang, Tieyun Qian, Guoliang He, Ming Zhong, and Qingxi Peng*

Chinese Comparative Sentence Identification Based on the Combination of Rules and Statistics . . . . .	300
--	-----

*Quanchao Liu, Heyan Huang, Chen Zhang, Zhenzhao Chen, and Jiajun Chen*

## Privacy Preservation

Utility Enhancement for Privacy Preserving Health Data Publishing . . . . .	311
---	-----

*Lengdong Wu, Hua He, and Osmar R. Zaïane*

Optimizing Placement of Mix Zones to Preserve Users' Privacy for Continuous Query Services in Road Networks . . . . .	323
---	-----

*Kamenyi Domenic M., Yong Wang, Fengli Zhang, Yankson Gustav, Daniel Adu-Gyamfi, and Nkatha Dorothy*

## Applications

Comparison of Cutoff Strategies for Geometrical Features in Machine Learning-Based Scoring Functions . . . . .	336
--	-----

*Shirley W.I. Siu, Thomas K.F. Wong, and Simon Fong*

Bichromatic Reverse Ranking Query in Two Dimensions . . . . .	348
---	-----

*Zhao Zhang, Qiangqiang Kang, Cheqing Jin, and Aoying Zhou*

Passive Aggressive Algorithm for Online Portfolio Selection with Piecewise Loss Function . . . . .	360
--	-----

*Li Gao, Weiguo Zhang, and Qiang Tang*

Mining Item Popularity for Recommender Systems . . . . .	372
<i>Jilian Zhang, Xiaofeng Zhu, Xianxian Li, and Shichao Zhang</i>	
Exploring an Ichthyoplankton Database from a Freshwater Reservoir in Legal Amazon . . . . .	384
<i>Michel de A. Silva, Daniela Queiroz Trevisan, David N. Prata, Elineide E. Marques, Marcelo Lisboa, and Monica Prata</i>	
A Pre-initialization Stage of Population-Based Bio-inspired Metaheuristics for Handling Expensive Optimization Problems . . . . .	396
<i>Muhammad Marwan Muhammad Fuad</i>	
A Hybrid-Sorting Semantic Matching Method . . . . .	404
<i>Kan Li, Wensi Mu, Yong Luan, and Shaohua An</i>	
Improving Few Occurrence Feature Performance in Distant Supervision for Relation Extraction . . . . .	414
<i>Hui Zhang and Yuanhao Zhao</i>	
Cluster Labeling Extraction and Ranking Feature Selection for High Quality XML Pseudo Relevance Feedback Fragments Set . . . . .	423
<i>Minjuan Zhong, Changxuan Wan, Dexi Liu, Shumei Liao, and Siwen Luo</i>	
Informed Weighted Random Projection for Dimension Reduction . . . . .	433
<i>Jaydeep Sen and Harish Karnick</i>	
Protocol Specification Inference Based on Keywords Identification . . . . .	443
<i>Yong Wang, Nan Zhang, Yan-mei Wu, and Bin-bin Su</i>	
An Adaptive Collaborative Filtering Algorithm Based on Multiple Features . . . . .	455
<i>Yan-Qiu Zhang, Hai-Tao Zheng, and Lan-Shan Zhang</i>	
<b>Machine Learning</b>	
Ensemble of Unsupervised and Supervised Models with Different Label Spaces . . . . .	466
<i>Yueyun Jin, Weilin Zeng, Hankz Hankui Zhuo, and Lei Li</i>	
Cost-Sensitive Extreme Learning Machine . . . . .	478
<i>Enhui Zheng, Cong Zhang, Xueyi Liu, Huijuan Lu, and Jian Sun</i>	
Multi-Objective Optimization for Overlapping Community Detection . . . . .	489
<i>Jingfei Du, Jianyang Lai, and Chuan Shi</i>	
Endmember Extraction by Exemplar Finder . . . . .	501
<i>Yi Guo, Junbin Gao, and Yanfeng Sun</i>	

EEG-Based User Authentication in Multilevel Security Systems . . . . .	513
<i>Tien Pham, Wanli Ma, Dat Tran, Phuoc Nguyen, and Dinh Phung</i>	
A New Fuzzy Extreme Learning Machine for Regression Problems with Outliers or Noises . . . . .	524
<i>Enhui Zheng, Jinyong Liu, Huijuan Lu, Ling Wang, and Le Chen</i>	
<b>Author Index . . . . .</b>	<b>535</b>