

# **2013 12th International Conference on Machine Learning and Applications**

**(ICMLA 2013)**

**Miami, Florida, USA  
4-7 December 2013**

**Volume 1  
Pages 1-449**



**IEEE Catalog Number: CFP13592-POD  
ISBN: 978-1-4799-4155-1**

# 2013 12th International Conference on Machine Learning and Applications

## ICMLA 2013

### Table of Contents

### Volume - 1

Preface - Volume 1 .....	xii
Organizing Committee - Volume 1 .....	xiii
Program Committee - Volume 1 .....	xiv
Keynotes - Volume 1 .....	xvii

---

#### **Ensemble-Based Methods I**

Ecoensembles: A Rapidly Deployable Image Classification System Using Feature-Views .....	1
<i>Adrian Rosebrock, Tim Oates, and Jesus Caban</i>	
An Empirical Comparison of Spectral Learning Methods for Classification .....	9
<i>Adam Drake and Dan Ventura</i>	
Informative Projection Recovery for Classification, Clustering and Regression .....	15
<i>Madalina Fiterau and Artur Dubrawski</i>	
Locally Linear Minimum Spanning Trees for Manifold Learning .....	21
<i>Carlos A. Quintero and Fernando Lozano</i>	

#### **Ensemble-Based Methods II**

Improving Software Quality Estimation by Combining Boosting and Feature Selection .....	27
<i>Kehan Gao, Taghi Khoshgoftaar, and Amri Napolitano</i>	
SetNet: Ensemble Method Techniques for Learning Regulatory Networks .....	34
<i>I. Chhbil, M. Elati, C. Rouveinol, and G. Santini</i>	
An Ensemble Learning Framework for Online Web Spam Detection .....	40
<i>Cailing Dong and Bin Zhou</i>	

Rot-SiLA: A Novel Ensemble Classification Approach Based on Rotation Forest and Similarity Learning Using Nearest Neighbor Algorithm .....	46
<i>Muhammad Shaheryar, Mehrosh Khalid, and Ali Mustafa Qamar</i>	

## **Kernel-Based Methods**

Distributed Kernel Matrix Approximation and Implementation Using Message Passing Interface .....	52
<i>Taher A. Dameh, Wael Abd-Almageed, and Mohamed Hefeeda</i>	
Approximate I-Fold Cross-Validation with Least Squares SVM and Kernel Ridge Regression .....	58
<i>Richard E. Edwards, Hao Zhang, Lynne E. Parker, and Joshua R. New</i>	
Multi-modal Tree-Based SVM Classification .....	65
<i>Cecille Freeman, Dana Kulic, and Otman Basir</i>	
Kernel SODA: A Feature Reduction Technique Using Kernel Based Analysis .....	72
<i>Yinan Yu, Tomas Mckelvey, and S.Y. Kung</i>	

## **Supervised Learning Methods I**

INDiC: Improved Non-intrusive Load Monitoring Using Load Division and Calibration .....	79
<i>Nipun Batra, Haimonti Dutta, and Amarjeet Singh</i>	
Automatic Grading of Computer Programs: A Machine Learning Approach .....	85
<i>Shashank Srikant and Varun Aggarwal</i>	
Empirical Co-occurrence Rate Networks for Sequence Labeling .....	93
<i>Zhemin Zhu, Djoerd Hiemstra, Peter Apers, and Andreas Wombacher</i>	
Applying Machine Learning and Audio Analysis Techniques to Insect Recognition in Intelligent Traps .....	99
<i>Diego F. Silva, Vinícius M.A. De Souza, Gustavo E.A.P.A. Batista, Eamonn Keogh, and Daniel P.W. Ellis</i>	
How Learning Enables Intelligence Analysts to Rapidly Develop Practical Cognitive Assistants .....	105
<i>Gheorghe Tecuci, Mihai Boicu, Dorin Marcu, and David Schum</i>	

## **Supervised Learning Methods II**

Learning Decision Trees from Uncertain Data with an Evidential EM Approach .....	111
<i>Nicolas Sutton-Charani, Sébastien Destercke, and Thierry Denoeux</i>	
A Sequential Testing Procedure for Multiple Change-Point Detection in a Stream of Pneumatic Door Signatures .....	117
<i>Nicolas Cheifetz, Allou Samé, Patrice Aknin, Emmanuel De Verdalle, and Damien Chenu</i>	

Linear Online Learning over Structured Data with Distributed Tree Kernels .....	123
<i>Simone Filice, Danilo Croce, Roberto Basili, and Fabio Massimo Zanzotto</i>	

## **Machine Learning in Bioinformatics and Computational Biology I**

eQTL Mapping Study via Regularized Sparse Canonical Correlation Analysis .....	129
<i>Mingon Kang, Shuo Li, Dongchul Kim, Chunyu Liu, Baoju Zhang, Xiaoyong Wu, and Jean Gao</i>	
Blood Glucose Level Prediction Using Physiological Models and Support Vector Regression .....	135
<i>Razvan Bunescu, Nigel Struble, Cindy Marling, Jay Shubrook, and Frank Schwartz</i>	
Mining Biomedical Ontologies and Data Using RDF Hypergraphs .....	141
<i>Haishan Liu, Dejing Dou, Ruoming Jin, Paea Lependu, and Nigam Shah</i>	
A Fast Multi-component Latent Variable Regression Framework for Quantitative Analysis of Surface-Enhanced Raman Spectra .....	147
<i>Shuo Li, James O. Nyagilo, Digant P. Dave, Baoju Zhang, Xiaoyong Wu, and Jean Gao</i>	

## **Machine Learning in Bioinformatics and Computational Biology II**

A Comparative Study of Food Intake Detection Using Artificial Neural Network and Support Vector Machine .....	153
<i>Muhammad Farooq, Juan M. Fontana, Akua F. Boateng, Megan A. Mccrory, and Edward Sazonov</i>	
Random Forest with 200 Selected Features: An Optimal Model for Bioinformatics Research .....	154
<i>Randall Wald, Taghi Khoshgoftaar, David J. Dittman, and Amri Napolitano</i>	
Zebrafish Larva Locomotor Activity Analysis Using Machine Learning Techniques .....	161
<i>Hao Zhang, Scott C. Lenaghan, Michelle H. Connolly, and Lynne E. Parker</i>	
Machine Learning Techniques for LV Wall Motion Classification Based on Spatio-temporal Profiles from Cardiac Cine MRI .....	167
<i>Juan Mantilla, Mireille Garreau, Jean-Jacques Bellanger, and José Luis Paredes</i>	

## **Clustering**

Variable-Length Protein Sequence Motif Extraction Using Hierarchically-Clustered Hidden Markov Models .....	173
<i>Cody Hudson and Bernard Chen</i>	

A Latent Semantic Approach to XML Clustering by Content and Structure Based on Non-negative Matrix Factorization .....	179
<i>Gianni Costa and Riccardo Ortale</i>	
Terrain Classification for a Quadruped Robot .....	185
<i>Jonas Degraeve, Robin Van Cauwenbergh, Francis Wyffels, Tim Waegeman, and Benjamin Schrauwen</i>	

## **Machine Learning and Applications**

Community Event Prediction in Dynamic Social Networks .....	191
<i>Nagehan Ilhan and İlule Gündüz Öğüdücü</i>	
Learning to Track Multi-target Online by Boosting and Scene Layout .....	197
<i>Guang Chen, Feihu Zhang, Daniel Clarke, and Alois Knoll</i>	
A First Step towards the Runtime Analysis of Evolutionary Algorithm Adjusted with Reinforcement Learning .....	203
<i>Maxim Buzdalov, Arina Buzdalova, and Anatoly Shalyto</i>	
Partially-Sparse Restricted Boltzmann Machine for Background Modeling and Subtraction .....	209
<i>Rui Guo and Hairong Qi</i>	

## **Posters - Main Conference**

The SLS-Generated Soft Robotic Hand - An Integrated Approach Using Additive Manufacturing and Reinforcement Learning .....	215
<i>Arne Rost and Stephan Schädle</i>	
WMCD: A Situation Aware Multicast Congestion Detection Scheme Using Support Vector Machines in MANETs .....	221
<i>Xiaoming Liu, Henry Nyongesa, and James Connan</i>	
Coordinated Reinforcement Learning Agents in a Multi-agent Virtual Environment .....	227
<i>William Sause</i>	
Natural Image Segmentation Based on Precise Edge Detection .....	231
<i>Wenya Feng, Yilin Guo, Xiaoyu Shi, and Yonggan Hou</i>	
Applications of Class-Conditional Conformal Predictor in Multi-class Classification .....	235
<i>Fan Shi, Cheng Soon Ong, and Christopher Leckie</i>	
Affect Detection and Classification from the Non-stationary Physiological Data .....	240
<i>Omar Alzoubi, Davide Fossati, Sidney D'mello, and Rafael A. Calvo</i>	
Application of Neural Networks Based SANARX Model for Identification and Control Liquid Level Tank System .....	246
<i>Juri Belikov, Sven Nömm, Eduard Petlenkov, and Kristina Vassiljeva</i>	

Arm Space Decomposition as a Strategy for Tackling Large Scale Multi-armed Bandit Problems .....	252
<i>Neha Gupta, Ole-Christoffer Granmo, and Ashok Agrawala</i>	
On the Behaviour of Scalarization Methods for the Engagement of a Wet Clutch .....	258
<i>Tim Brys, Kristof Van Moffaert, Kevin Van Vaerenbergh, and Ann Nowé</i>	
Random Forest Classification for Hippocampal Segmentation in 3D MR Images .....	264
<i>Rosalia Maglietta, Nicola Amoroso, Stefania Bruno, Andrea Chincarini, Giovanni Frisoni, Paolo Inglese, Sabina Tangaro, Andrea Tateo, and Roberto Bellotti</i>	
Ordered Segment for Classification of Big Data .....	268
<i>A. Fatholahzadeh</i>	
Indicative Support Vector Clustering with Its Application on Anomaly Detection .....	273
<i>Huang Xiao and Claudia Eckert</i>	
Computed Data-Geometry Based Supervised and Semi-supervised Learning in High Dimensional Data .....	277
<i>Elizabeth P. Chou, Fushing Hsieh, and John Capitanio</i>	
Collective Classification Using Semantic Based Regularization .....	283
<i>Claudio Saccá, Michelangelo Diligenti, and Marco Gori</i>	
Graph and Manifold Co-regularization .....	287
<i>Claudio Saccá, Michelangelo Diligenti, and Marco Gori</i>	
Segmental Analysis and Evaluation of User Focused Search Process .....	291
<i>Chathra Hendahewa and Chirag Shah</i>	
Dynamic Distributed Genetic Algorithm Using Hierarchical Clustering for Flight Trajectory Optimization of Winged Rocket .....	295
<i>S. Miyamoto, T. Matsumoto, and K. Yonemoto</i>	
Adversarial Spam Detection Using the Randomized Hough Transform-Support Vector Machine .....	299
<i>Dave Debarr, Hao Sun, and Harry Wechsler</i>	
Machine Learning Techniques Applied to Sensor Data Correction in Building Technologies .....	305
<i>Matt K. Smith, Charles C. Castello, and Joshua R. New</i>	
Eigenfaces for Face Detection: A Novel Study .....	309
<i>Salaheddin Alakkari and John James Collins</i>	
Discriminative Apprenticeship Learning with Both Preference and Non-preference Behavior .....	315
<i>Dingsheng Luo, Yi Wang, and Xihong Wu</i>	

BEMI Biclust Ensemble Using Mutual Information .....	321
<i>Geeta Aggarwal and Neelima Gupta</i>	
Using Support Vector Machines to Classify Student Attentiveness for the Development of Personalized Learning Systems .....	325
<i>Manus Ross, Corey A. Graves, John W. Campbell, and Jung H. Kim</i>	
A Temporal Difference GNG-Based Algorithm That Can Learn to Control in Reinforcement Learning Environments .....	329
<i>Davi C.L. Vieira, Paulo J.L. Adeodato, and Paulo M. Gonçalves Júnior</i>	
Cost Sensitive Credit Card Fraud Detection Using Bayes Minimum Risk .....	333
<i>Alejandro Correa Bahnsen, Aleksandar Stojanovic, Djamila Aouada, and Björn Ottersten</i>	
Event Causality Identification Using Conditional Random Field in Geriatric Care Domain .....	339
<i>Saeed Mehrabi, Anand Krishnan, Eric Tinsley, Jon Sligh, Natalie Crohn, Heather Bush, Jason Depasquale, Jean Bandos, and Mathew Palakal</i>	
Movie Recommendation Using Unrated Data .....	344
<i>Dong Nie, Lingzi Hong, and Tingshao Zhu</i>	
Targeted Action Rule Discovery .....	348
<i>Tom Johnsten, Samy Alihamad, Ashwin Kannalath, and Ryan G. Benton</i>	
Parallel Coordinate Descent for the Adaboost Problem .....	354
<i>Olivier Fercoq</i>	
Hybrid Ontology-Based Information Extraction for Automated Text Grading .....	359
<i>Fernando Gutierrez, Dejing Dou, Adam Martini, Stephen Fickas, and Hui Zong</i>	
Empirical Evaluation of Profile Characteristics for Gender Classification on Twitter .....	365
<i>Jalal S. Alowibdi, Ugo A. Buy, and Philip Yu</i>	
Determining Potential Yeast Longevity Genes via PPI Networks and Microarray Data Clustering Analysis .....	370
<i>Bernard Chen, Roshan Doolabh, and Fusheng Tang</i>	
Higher-Order Regularized Kernel CCA .....	374
<i>Md. Ashad Alam and Kenji Fukumizu</i>	
Protein Local Tertiary Structure Prediction Using the Adaptively-Branching FGK-DF Model .....	378
<i>Bernard Chen, Cody Hudson, Aaron Crawford, and Minwoo Kim</i>	
Data Analysis, Discharge Classifications, and Predictions of Hydrological Parameters for the Management of Rawal Dam in Pakistan .....	382
<i>Maqbool Ali, Ali Mustafa Qamar, and Bilal Ali</i>	
Evolutionary Content Pre-fetching in Mobile Networks .....	386
<i>Omar K. Shoukry and Magda B. Fayek</i>	

Learning the Dynamic Process of Inhibition and Task Switching in Robotics Cognitive Control .....	392
<i>Matteo Menna, Mario Gianni, and Fiara Pirri</i>	
Scalable Expert Selection When Learning from Noisy Labelers .....	398
<i>Chirine Wolley and Mohamed Quafafou</i>	
Incorporating Categorical Information for Enhanced Probabilistic Trajectory Prediction .....	402
<i>Juergen Wiest, Felix Kunz, Ulrich Kreßel, and Klaus Dietmayer</i>	
Online Processing of Social Media Data for Emergency Management .....	408
<i>Daniela Pohl, Abdelhamid Bouchachia, and Hermann Hellwagner</i>	
Deep Multiple Kernel Learning .....	414
<i>Eric V. Strobl and Shyam Visweswaran</i>	
Comparative Analysis on the Stability of Feature Selection Techniques Using Three Frameworks on Biological Datasets .....	418
<i>Randall Wald, Taghi Khoshgoftaar, Ahmad Abu Shanab, and Amri Napolitano</i>	
Learning-Based Incremental Creation of Web Image Databases .....	424
<i>Marian George, Nagia Ghanem, and M.A. Ismail</i>	
An Incremental Parallel Particle Swarm Approach for Classification Rule Discovery from Dynamic Data .....	430
<i>Kaveh Hassani and Won-Sook Lee</i>	
A Classifier Ensembling Approach for Imbalanced Social Link Prediction .....	436
<i>Jose Hurtado, Napati Taweewitchakreeya, Xue Kong, and Xingquan Zhu</i>	
Phrase Based Topic Modeling for Semantic Information Processing in Biomedicine .....	440
<i>Zhiguo Yu, Todd R. Johnson, and Ramakanth Kavuluru</i>	
<b>Author Index - Volume 1</b> .....	<b>446</b>



# **2013 12th International Conference on Machine Learning and Applications**

**(ICMLA 2013)**

**Miami, Florida, USA  
4-7 December 2013**

**Volume 2  
Pages 1-581**



**IEEE Catalog Number: CFP13592-POD  
ISBN: 978-1-4799-4155-1**

# 2013 12th International Conference on Machine Learning and Applications

## ICMLA 2013

### Table of Contents Volume - 2

#### Special Sessions and Workshops

Organizations and Committees - Volume 2.....	xiv
--	-----

---

#### **01. Special Session: Adaptive and Dynamic Modelling in Non-stationary Environments**

Generalized Flexible Fuzzy Inference Systems .....	1
<i>Edwin Lughofer, Carlos Cernuda, and Mahardhika Pratama</i>	
Design of Chaos-Based Communication System with Use of the Derivative-Free Nonlinear Kalman Filter .....	8
<i>Gerasimos G. Rigatos</i>	
Evolving Dynamic Forecasting Model for Foreign Currency Exchange Rates Using Plastic Neural Networks .....	15
<i>Gul Muhammad Khan, Durre Nayab, S. Ali Mahmud, and Haseeb Zafar</i>	
Stable On-Line Learning with Optimized Local Learning, But Minimal Change of the Global Output .....	21
<i>Andreas Buschermöhle and Werner Brockmann</i>	
On-Line Incremental Learning for Unknown Conditions during Assembly Operations with Robots .....	28
<i>Navarro Gonzalez Jose Luis, Lopez Juarez Ismael, and Ordaz Hernandez Keny</i>	

## 02. Special Session: Machine Learning Challenges in Cyber Security Applications

Active Learning for Alert Triage .....	34
<i>Justin E. Doak, Joe Ingram, Jeffery Shelburg, Joshua Johnson, and Brandon R. Rohrer</i>	
Aiding Intrusion Analysis Using Machine Learning .....	40
<i>Loai Zomlot, Sathya Chandran, Doina Caragea, and Xinming Ou</i>	
Streaming Malware Classification in the Presence of Concept Drift and Class Imbalance .....	48
<i>W. Philip Kegelmeyer, Ken Chiang, and Joe Ingram</i>	
An Evaluation of Machine Learning Methods to Detect Malicious SCADA Communications .....	54
<i>Justin M. Beaver, Raymond C. Borges-Hink, and Mark A. Buckner</i>	
PACE: Pattern Accurate Computationally Efficient Bootstrapping for Timely Discovery of Cyber-security Concepts .....	60
<i>Nikki McNeil, Robert A. Bridges, Michael D. Iannacone, Bogdan Czejdo, Nicolas Perez, and John R. Goodall</i>	

## 03. Special Session: Machine Learning Applications in Software Engineering

DECOBA: Utilizing Developers Communities in Bug Assignment .....	66
<i>Shadi Banitaan and Mamdouh Alenezi</i>	
Predicting Parallelization of Sequential Programs Using Supervised Learning .....	72
<i>Daniel Fried, Zhen Li, Ali Jannesari, and Felix Wolf</i>	
Identifying Effective Test Cases through K-Means Clustering for Enhancing Regression Testing .....	78
<i>Yulei Pang, Xiaozhen Xue, and Akbar Siami Namin</i>	
An Empirical Study on Wrapper-Based Feature Selection for Software Engineering Data .....	84
<i>Huanjing Wang, Taghi M. Khoshgoftaar, and Amri Napolitano</i>	
Learning Finite-State Machines: Conserving Fitness Function Evaluations by Marking Used Transitions .....	90
<i>Daniil Chivilikhin and Vladimir Ulyantsev</i>	
Class Diagram Retrieval Using Genetic Algorithm .....	96
<i>Hamza Onoruoiza Salami and Moataz Ahmed</i>	
Objective Re-weighting to Guide an Interactive Search Based Software Testing System .....	102
<i>Bogdan Marculescu, Robert Feldt, and Richard Torkar</i>	

Worst-Case Execution Time Test Generation for Augmenting Path Maximum	
Flow Algorithms Using Genetic Algorithms .....	108
<i>Viktor Arkhipov, Maxim Buzdalov, and Anatoly Shalyto</i>	

## **04. Special Session: Machine Learning for Predictive Models**

Bug Reports Prioritization: Which Features and Classifier to Use? .....	112
<i>Mamdouh Alenezi and Shadi Banitaan</i>	
Fuzzy Model Tree for Early Effort Estimation .....	117
<i>Mohammad Azzeh and Ali Bou Nassif</i>	
Ensemble Feature Selection Methods for a Better Regularization of the Lasso	
Estimate in $P \gg N$ Gene Expression Datasets .....	122
<i>Adel Aloraini</i>	
Personalized Spam Filtering with Natural Language Attributes .....	127
<i>Rushdi Shams and Robert E. Mercer</i>	
First-Order Probabilistic Model for Hybrid Recommendations .....	133
<i>Julia Hoxha and Achim Rettinger</i>	
Learning from Multiple Graphs Using a Sigmoid Kernel .....	140
<i>Thomas Ricatte, Gemma Garriga, Rémi Gilleron, and Marc Tommasi</i>	
Methods and Applications for Distance Based ANN Training .....	146
<i>Christoph Lassner and Rainer Lienhart</i>	
Very Short Term Load Forecasting Using Cartesian Genetic Programming	
Evolved Recurrent Neural Networks (CGPRNN) .....	152
<i>Gul Muhammad Khan, Faheem Zafari, and S. Ali Mahmud</i>	
Use of a Feedforward Neural Network for Predicting the Development	
Duration of Software Projects .....	156
<i>Cuahtémoc López-Martín, Arturo Chavoya, and María Elena Meda-Campaña</i>	
A Genetic Algorithm to Optimize Lazy Learning Parameters for the Prediction	
of Customer Demands .....	160
<i>Mirko Kück and Bernd Scholz-Reiter</i>	
DroidMLN: A Markov Logic Network Approach to Detect Android Malware .....	166
<i>Mahmuda Rahman</i>	
Preprocessing in Fuzzy Time Series to Improve the Forecasting Accuracy .....	170
<i>Fábio José Justo Dos Santos and Heloisa De Arruda Camargo</i>	
Virtual Metrology in Semiconductor Manufacturing by Means of Predictive	
Machine Learning Models .....	174
<i>Benjamin Lenz, Bernd Barak, Julia Mührwald, Carolin Leicht, Benjamin Lenz, and Julia Mührwald</i>	

Selective Sampling Designs to Improve the Performance of Classification Methods .....	178
<i>Soroosh Ghorbani and Michel C. Desmarais</i>	
Pairwise Clustering by Minimizing the Error of Unsupervised Nearest Neighbor Classification .....	182
<i>Yingzhen Yang, Xinqi Chu, and Thomas S. Huang</i>	

## **05. Special Session: Machine Learning in Energy Applications**

Permeability Parametrization Using Higher Order Singular Value Decomposition (HOSVD) .....	188
<i>Sardar Afra and Eduardo Gildin</i>	
Quality Classification of Green Pellet Nuclear Fuels Using Radial Basis Function Neural Networks .....	194
<i>Benyamin Kusumoputro, Akhmad Faqih, Dede Sutarya, and Lina</i>	
Approach to Control of Hybrid Renewable Power System on the Basis of AE-Method Using Genetic Algorithm .....	199
<i>V. Ten, B. Matkarimov, and N. Isembergenov</i>	
Price Forecasting in the Spanish Day-Ahead Electricity Market Using Preconditioned Wind Power Information .....	203
<i>C. Geidel and H. Zareipour</i>	
A Hybrid Feature Selection and Generation Algorithm for Electricity Load Prediction Using Grammatical Evolution .....	211
<i>Anthony Mhirana De Silva, Farzad Noorian, Richard I.A. Davis, and Philip H.W. Leong</i>	
An Intelligent Tool for ANN Based Power Factor Correction .....	218
<i>Omer Sesveren and Ramazan Bayindir</i>	
Intelligent Decision Support System for Energy Investments .....	224
<i>Uraz Yavanoglu, Orhan Kaplan, Gizem Tanis, Hacer Atli, Ozlem Milletsever, and Ugur Inal</i>	
Operation Planning of Hydroelectric Systems: Application of Genetic Algorithms and Differential Evolution .....	232
<i>Priscila C. Berbert Rampazzo, Akebo Yamakami, and Fabrício O. De França</i>	
Artificial Neural Networks Controller Algorithm Developed for a Brushless DC Motor .....	238
<i>Ilhami Colak, Murat Sahin, and Zafer Esen</i>	
A Neural Network Approach to Multi-step-ahead, Short-Term Wind Speed Forecasting .....	243
<i>Julian L. Cardenas-Barrera, Julian Meng, Eduardo Castillo-Guerra, and Liuchen Chang</i>	

An Evaluation Study on Circuit Parameter Conditions of Neural Network Controlled DC-DC Converter .....	249
<i>Hidenori Maruta, Masashi Motomura, and Fujio Kurokawa</i>	

## **06. Special Session: Machine Learning in Information and System Security Issues**

A Comparative Analysis of Cryptographic Algorithms and Transformation Functions for Biometric Data .....	255
<i>Isaac De L. Oliveira Filho, Otaciana G.R. Santiago, Anne M.P. Canuto, and Benjamín R.C. Bedregal</i>	
Learning to Block Undesired Comments in the Blogosphere .....	261
<i>Tiago A. Almeida and Túlio C. Alberto</i>	
Predicting Next Location of Twitter Users for Surveillance .....	267
<i>Sedef Gunduz, Uraz Yavanoglu, and Seref Sagiroglu</i>	
Google Penguin: Evasion in Non-English Languages and a New Classifier .....	274
<i>Abdulrahman Alarifi, Mansour Alsaleh, Abdulmalik Al-Salman, Abdulmajeed Alswayed, and Ahmad Alkhaledi</i>	
Intelligent Approach for Identifying Political Views over Social Networks .....	281
<i>Uraz Yavanoglu, Medine Colak, Busra Caglar, Semra Cakir, Ozlem Milletsever, and Seref Sagiroglu</i>	
A New Approach to Detecting Content Anomalies in Wikipedia .....	288
<i>Duygu Sinanc and Uraz Yavanoglu</i>	

## **07. Special Session: Machine Learning with Multimedia Data**

A Riemannian Stopping Criterion for Unsupervised Phonetic Segmentation .....	294
<i>Ciro Gracia Pons, Xavier Anguera, and Xavier Binefa</i>	
Two-Level Clustering towards Unsupervised Discovery of Acoustic Classes .....	299
<i>Ciro Gracia Pons, Xavier Anguera, and Xavier Binefa</i>	
Unsupervised Video Summarization via Dynamic Modeling-Based Hierarchical Clustering .....	303
<i>Karim M. Mahmoud, Nagia M. Ghanem, and Mohamed A. Ismail</i>	
Can We Minimize the Influence Due to Gender and Race in Age Estimation? .....	309
<i>Xiaolong Wang, Vincent Ly, Guoyu Lu, and Chandra Kambhamettu</i>	
EPG Content Recommendation in Large Scale: A Case Study on Interactive TV Platform .....	315
<i>Dávid Zibriczky, Zoltán Petres, Márton Waszlavik, and Domonkos Tikk</i>	
A Multimodal Approach to Song-Level Style Identification in Pop/Rock Using Similarity Metrics .....	321
<i>Ching-Hua Chuan</i>	

## **08. Special Session: Learning on the Web (LW)**

Learning Relevance of Web Resources across Domains to Make Recommendations .....	325
<i>Julia Hoxha, Peter Mika, and Roi Blanco</i>	
Future Clients' Requests Estimation for Dynamic Resource Allocation in Cloud Data Center Using CGPANN .....	331
<i>Jawad Ali, Faheem Zafari, Gul Muhammad Khan, and S. Ali Mahmud</i>	
Incorporating Game-Theoretic Rough Sets in Web-Based Medical Decision Support Systems .....	335
<i>Jingtao Yao and Nouman Azam</i>	

## **09. Workshop: Machine Learning Algorithms, Systems, and Applications**

Gradient Hyper-parameter Optimization for Manifold Regularization .....	339
<i>Cassiano O. Becker and Paulo A.V. Ferreira</i>	
Fault Diagnosis in Railway Track Circuits Using Support Vector Machines .....	345
<i>Shangpeng Sun and Huibing Zhao</i>	
A Framework towards the Unification of Ensemble Classification Methods .....	351
<i>Mohammad Ali Bagheri, Qigang Gao, and Sergio Escalera</i>	
Clustering Based Classification in Data Mining Method Recommendation .....	356
<i>Ondřej Kazík, Klára Pešková, Jakub Šmíd, and Roman Neruda</i>	
Sampling Adaptively Using the Massart Inequality for Scalable Learning .....	362
<i>Jianhua Chen and Jian Xu</i>	
Advanced Pattern Recognition Approach for Fault Diagnosis of Wind Turbines .....	368
<i>Houari Toubakh, Moamar Sayed-Mouchaweh, and Eric Duviella</i>	
Improved Helper-Objective Optimization Strategy for Job-Shop Scheduling Problem .....	374
<i>Irina Petrova, Arina Buzdalova, and Maxim Buzdalov</i>	
Evolving Hybrid Neural Fuzzy Network for System Modeling and Time Series Forecasting .....	378
<i>Raul Rosa, Fernando Gomide, and Rosangela Ballini</i>	

## 10. Workshop: Machine Learning and Applications in Health Informatics

A Multiagent Approach to Ambulance Allocation Based on Social Welfare and Local Search .....	384
<i>Dean Shaft and Robin Cohen</i>	
A Comprehension Approach for Formalizing Privacy Rules of HIPAA for Decision Support .....	390
<i>Imran Khan, Moheeb Alwarsh, and Javed I. Khan</i>	
Simplifying the Utilization of Machine Learning Techniques for Bioinformatics .....	396
<i>David J. Dittman, Taghi M. Khoshgoftaar, Randall Wald, and Amri Napolitano</i>	
Contrasting Undersampled Boosting with Internal and External Feature Selection for Patient Response Datasets .....	404
<i>Taghi M. Khoshgoftaar, David J. Dittman, Randall Wald, and Amri Napolitano</i>	
Gene Expression Profile Classification Using Random Projection and Sparse Representation .....	411
<i>X. Hang</i>	
Random Forests on Ubiquitous Data for Heart Failure 30-Day Readmissions Prediction .....	415
<i>Michael A. Vedomské, Donald E. Brown, and James H. Harrison</i>	
Scalable and Locally Applicable Measures of Treatment Variation That Use Hospital Billing Data .....	422
<i>Michael A. Vedomské, Matthew S. Gerber, Donald E. Brown, and James H. Harrison</i>	
Automatic Behavior Learning for Personalized Assisted Living Systems .....	428
<i>E. Kařitoch and P. Augustyniak</i>	
Technical Verification of Integrating Wearable Sensors into BSN-Based Telemedical Monitoring System .....	432
<i>Eliasz Kařitoch</i>	
Learning to Rank Drug Combinations .....	436
<i>Fernando Lozano</i>	
Automated Pap Smear Cell Analysis: Optimizing the Cervix Cytological Examination .....	441
<i>Daniela M. Ushizima, Alessandra H. Gomes, Claudia M. Carneiro, and Andrea G.C. Bianchi</i>	



## 11. Workshop: Big Data and Data Analytics Applications

Recommending Messages to Users in Social Networks: A Cross-Site Study .....	445
<i>R. Cohen, N. Sardana, K. Rahim, D.Y. Lam, M. Li, O. Maccarthy, E. Woo, J. Zhang, and G. Guo</i>	
Massive GIS Database System with Autonomic Resource Management .....	451
<i>Yun Lu, Ming Zhao, Guangqiang Zhao, Lixi Wang, and Naphtali Rishe</i>	
Comparison of Stability for Different Families of Filter-Based and Wrapper-Based Feature Selection .....	457
<i>Randall Wald, Taghi Khoshgoftaar, and Amri Napolitano</i>	
Survey of Clinical Data Mining Applications on Big Data in Health Informatics .....	465
<i>Matthew Herland, Taghi M. Khoshgoftaar, and Randall Wald</i>	
Constrained Motion Particle Swarm Optimization and Support Vector Regression for Non-linear Time Series Regression and Prediction Applications .....	473
<i>Nicholas I. Sapankevych and Ravi Sankar</i>	
Survey of Data Cleansing and Monitoring for Large-Scale Battery Backup Installations .....	478
<i>Liz Aranguren Pachano, Taghi M. Khoshgoftaar, and Randall Wald</i>	
Epidemiological Data Analysis in TerraFly Geo-spatial Cloud .....	485
<i>Huibo Wang, Yun Lu, Yudong Guang, Erik Edrosa, Mingjin Zhang, Raul Camarca, Yelena Yesha, Tajana Lucic, and Naphtali Rishe</i>	
Hybrid Method for Fast SVM Training in Applications Involving Large Volumes of Data .....	491
<i>M. Arif Wani</i>	

## 12. Special Session: Machine Learning in Visual Information Processing

SksOpen: Efficient Indexing, Querying, and Visualization of Geo-spatial Big Data .....	495
<i>Yun Lu, Mingjin Zhang, Shonda Witherspoon, Yelena Yesha, Yaacov Yesha, and Naphtali Rishe</i>	
Dual Tree Complex Wavelet Transform Based Multiclass Object Classification .....	501
<i>Ashish Khare, Manish Khare, and Rajneesh Kumar Srivastava</i>	
A Robust Unsupervised Feature Learning Framework Using Spatial Boosting Networks .....	507
<i>Nam Do-Hoang Le and Minh-Triet Tran</i>	
LIPID: Local Image Permutation Interval Descriptor .....	513
<i>Tian Tian, Ishwar Sethi, Delie Ming, Yun Zhang, and Jiayi Ma</i>	

Weak Segmentations and Ensemble Learning to Predict Semantic Ratings of Lung Nodules .....	519
<i>Ethan Smith, Patrick Stein, Jacob Furst, and Daniela Stan Raicu</i>	
Visual Speech Detection Using an Unsupervised Learning Framework .....	525
<i>Rameez Ahmad, Syed Paymaan Raza, and Hafiz Malik</i>	
Automatic Binding Point and Surface Helix Angle Measurement in Historic Weft-Faced Compound Weave Figured Silks .....	529
<i>Nilesh Patel, Michael Flynn, Andrey Semenov, and Julia Galliker</i>	

### **13. Special Session: Machine Learning Applications in Education**

The Estimation of Students' Academic Success by Data Mining Methods .....	535
<i>Hanife Göker, Halil Ibrahim Bülbül, and Erdal Irmak</i>	
The Determination of Socio-economic Factors Affecting Student Success by Data Mining Methods .....	540
<i>Veli Atalay, Süleyman Üstün, and Selin Bülbül</i>	
Approach to Cold-Start Problem in Recommender Systems in the Context of Web-Based Education .....	543
<i>Reginaldo Aparecido Gotardo, Estevam Rafael, Hruschka Junior, Sergio Donizetti Zorzo, and</i>	
An Artificial Neural Networks Based Software System for Improved Learning Experience .....	549
<i>Utku Kose</i>	
Hindi Language Graphical User Interface to Database Management System .....	555
<i>Mohit Dua, Sandeep Kumar, and Zorawar Singh Virk</i>	
Improving the Transcription of Academic Lectures for Information Retrieval .....	560
<i>Audrey Mbogho and Stephen Marquard</i>	
Enhancing Online Music Lessons with Applications in Automating Self-Learning Tutorials and Performance Assessment .....	568
<i>Nashlie H. Sephus, Temiloluwa O. Olubanjo, and David V. Anderson</i>	
Scenario Based Functional Regression Testing Using Petri Net Models .....	572
<i>Farooq Ahmad and Zahid Hussain Qaisar</i>	
<b>Author Index - Volume 2</b> .....	<b>578</b>