Session 1: Recognition and Detection

Application of Structural Case-Based Reasoning to Activity Recognition in Smart Home Environments .................................................................1
    Steven Satterfield, Thomas Reichherzer, John Coffey, and Eman El-Sheikh

Detecting Web Robots Using Resource Request Patterns ........................................7
    Derek Doran and Swapna S. Gokhale

Automatically Detecting Avalanche Events in Passive Seismic Data ..........................13
    Marc J. Rubin, Tracy Camp, Alec Van Herwijnen, and Jürg Schweizer

Sparse Dictionary Reconstruction for Textile Defect Detection ...................................21
    Jian Zhou, Dimitri Semenovich, Arcot Sowmya, and Jun Wang

Convolutional Neural Support Vector Machines: Hybrid Visual Pattern Classifiers for Multi-robot Systems ...............................................................27
    Jawad Nagi, Gianni A. Di Caro, Alessandro Giusti, Farrukh Nagi,
    and Luca M. Gambardella

Introducing Subspace Grids to Recognise Patterns in Multidimensional Data .............33
    M. Arif Wani
Session 2: Face Recognition

Face Recognition in the Virtual World: Recognizing Avatar Faces .................................................... 40
Roman V. Yampolskiy, Brendan Klare, and Anil K. Jain

Subject-Independent Facial Expression Recognition with Biologically Inspired Features .............................................................................................................. 46
Weifeng Liu, Caifeng Song, and Yanjiang Wang

Sparse Representation Based Discriminative Canonical Correlation Analysis for Face Recognition ........................................................................................................... 51
Naiyang Guan, Xiang Zhang, Zhigang Luo, and Long Lan

Adaptive Extended Local Ternary Pattern (AELTP) for Recognizing Avatar Faces ................................................................. 57
Abdallah A. Mohamed and Roman V. Yampolskiy

Session 3: Machine Learning in Bio-informatics and Computational Biology I

Mean Aggregation versus Robust Rank Aggregation for Ensemble Gene Selection ...................................................................................................................................................... 63
Randall Wald, Taghi M. Khoshgoftaar, and David Dittman

SVM-based Framework for the Robust Extraction of Objects from Histopathological Images Using Color, Texture, Scale and Geometry ........................................................................................................................................... 70
Antoine Veillard, Stéphane Bressan, and Daniel Racoceanu

Effective Enrichment of Gene Expression Data Sets ......................................................................................................................... 76
Utku Sirin, Utku Erdogdu, Mehmet Tan, Faruk Polat, and Reda Alhajj

Determining the Number of Iterations Appropriate for Ensemble Gene Selection on Microarray Data .................................................................................................................................................. 82
David J. Dittman, Taghi M. Khoshgoftaar, Randall Wald, and Amri Napolitano

A Novel Hierarchical Level Set with AR-boost for White Matter Lesion Segmentation in Diabetes .................................................................................................................................................. 90
Baidya N. Saha, Sriraam Natarajan, Gopi Kota, Christopher T. Whitlow, Donald Bowden, Jasmin Divers, Barry I. Freedman, and Joseph A. Maldjian
Session 4: Machine Learning in Bio-informatics and Computational Biology II

Prediction of Protein Essentiality by the Support Vector Machine with Statistical Tests .........................................................................................................................96
Chiou-Yi Hor, Chang-Biau Yang, Zih-Jie Yang, and Chiou-Ting Tseng

Incremental Mitosis: Discovering Clusters of Arbitrary Shapes and Densities in Dynamic Data ...................................................................................................................102
Rania Ibrahim, Naglaa Ahmed, Noha A. Yousri, and Mohamed A. Ismail

Network Transformation of Gene Expression for Feature Extraction ..............................................................................................................................108
Rémynicolle, Mohamed Elati, et François Radvanyi

Combining Gene Expression Profiles and Protein-Protein Interactions for Identifying Functional Modules ........................................................................................................114
Dingding Wang, Mitsunori Ogihara, Erling Zeng, and Tao Li

Building Synthetic Networks of the Budding Yeast Cell-Cycle Using Swarm Intelligence ..................................................................................................................120
Gonzalo A. Ruz, Tania Timmermann, and Eric Goles

Incorporating Gene Significance in the Impact Analysis of Signaling Pathways .........................................................................................................................126
Calin Voichita, Michele Donato, and Sorin Draghici

Session 5: Kernel and Re-inforcement Learning

Differentiable Kernels in Generalized Matrix Learning Vector Quantization .........................................................................................................................132
Marika Kästner, David Nebel, Martin Riedel, Michael Biehl, and Thomas Villmann

Monte Carlo Tree Search for Bayesian Reinforcement Learning .........................................................................................................................138
Ngo Anh Vien and Wolfgang Ertel

An Inverse Reinforcement Learning Algorithm for Partially Observable Domains with Application on Healthcare Dialogue Management ...............................................................................144
Hamid R. Chinaei and Brahim Chaib-Draa

Increasing Efficiency of Evolutionary Algorithms by Choosing between Auxiliary Fitness Functions with Reinforcement Learning .........................................................................................150
Arina Buzdalova and Maxim Buzdalov

pRBF Kernels: A Framework for the Incorporation of Task-Specific Properties into Support Vector Methods ........................................................................................................156
Antoine Veillard, Daniel Racoceanu, and Stéphane Bressan
**Session 6: Structure, Manifold and Bayesian Learning**

Deep Structure Learning: Beyond Connectionist Approaches .................................................................162  
*Ben Mitchell and John Sheppard*

Balancing Public Cycle Sharing Schemes Using Independent Learners ......................................................168  
*Jeremiah Smith, Luke Dickens, and Krysia Broda*

Semi-supervised Non-negative Patch Alignment Framework .................................................................174  
*Long Lan, Xuhui Huang, Naiyang Guan, Zhigang Luo, and Xiang Zhang*

**Session 7: Machine Learning Methods for Bio-medical Application**

Estimating Hospital Admissions with a Randomized Regression Approach ..............................................179  
*Kleber A. Garcia and Philip K. Chan*

Coupled Directional Level Set for MR Image Segmentation ........................................................................185  
*Xianjing Qin, Yang Liu, Hongbing Lu, Xuelong Li, and Pingkun Yan*

Stereo Matching Based on Random Speckle Projection for Dynamic 3D Sensing ........................................191  
*Jun Jiang, Jun Cheng, and Haifeng Zhao*

Semantic Data Types in Machine Learning from Healthcare Data ..........................................................197  
*Janusz Wojtusiak*

A Machine Learning Pipeline for Three-Way Classification of Alzheimer Patients from Structural Magnetic Resonance Images of the Brain .................................................................203  
*Sriram Natarajan, Saket Joshi, Baidya N. Saha, Adam Edwards, Tushar Khot, Elizabeth Moody, Kristian Kersting, Christopher T. Whitlows, and Joseph A. Maldjian*

**Session 8: Classification Methods**

Minimal Norm Support Vector Machines for Large Classification Tasks ..................................................209  
*Robert Strack and Vojislav Kecman*

Fast Time Series Classification Based on Infrequent Shapelets ..................................................................215  
*Qing He, Zhidong, Fuzhen Zhuang, Tianfeng Shang, and Zhongzhi Shi*

Cost-Sensitive Universum-SVM ..................................................................................................................220  
*Sauptik Dhar and Vladimir Cherkassky*

A Hybrid Transfer Learning Mechanism for Object Classification across View ........................................226  
*Yi Mo, Zhaoxiang Zhang, and Yunhong Wang*

Location Type Classification Using Tweet Content .......................................................................................232  
*Haibin Liu, Bo Luo, and Dongwon Lee*
Session 9: Machine Learning Application

Topological Indoor Localization & Navigation for Autonomous Industrial Mobile Manipulator ..................................................................................................................238
  Hongtai Cheng, Heping Chen, and Yong Liu

Estimating the Convection Heat Transfer Coefficient of a Run-Out Cooling Table in a Steel-Making Process by Neural Networks .................................................................244
  Gláucio Barros Barcelos, Douglas Alexandre Gomes Vieira,
  Rodney Rezende Saldanha, and Luciano Lellis Miranda

Writing with Style: Venue Classification ......................................................................250
  Zaihan Yang and Brian D. Davison

Incorporating Incremental and Active Learning for Scene Classification ..................256
  Xianglin Li, Runqiu Guo, and Jun Cheng

Augmented Coupled Dictionary Learning for Image Super-Resolution ....................262
  Muhammad Rushdi and Jeffrey Ho

Studying the Effect of Class Imbalance in Ocean Turbine Fault Data on Reliable State Detection ..................................................................................................................268
  Janell Duhaney, Taghi M. Khoshgoftaar, and Amri Napolitano

Session 10: Data Categorization, Data Visualization and Numerical Optimization

Cellular Differentiation Algorithm for High Dimensional Numerical Function Optimization ...............................................................................................................................276
  Yanjiang Wang, Chengna Yuan, and Weifeng Liu

Online Recovery of Missing Values in Vital Signs Data Streams Using
Low-Rank Matrix Completion ......................................................................................281
  Shiming Yang, Konstantinos Kalpakis, Colin F. Mackenzie, Lynn G. Stansbury,
  Deborah M. Stein, Thomas M. Scalea, and Peter F. Hu

Analysis of Causality in Stock Market Data .................................................................288
  Chathra Hendahewa and Vladimir Pavlovic

An Efficient Algorithm for top-k Queries on Uncertain Data Streams .......................294
  Caiyan Dai, Ling Chen, Yixin Chen, and Keming Tang

Perceptual User Study for Combined Treemap ............................................................300
  Jie Liang, Mao Lin Huang, and Quang Vinh Nguyen
Session 11: Feature Selection and Fusion

Feature Mapping and Fusion for Music Genre Classification .................................................................306
  Haythem Balti and Hichem Frigui

Multi-atlas Based Image Selection with Label Image Constraint ............................................................311
  Yihui Cao, Xuelong Li, and Pingkun Yan

An Empirical Study on the Stability of Feature Selection for Imbalanced Software Engineering Data ........................................317
  Huanjing Wang, Taghi M. Khoshgoftaar, and Amri Napolitano

Supervised Low Rank Matrix Approximation for Stable Feature Selection ........................................324
  Salem Alelyani and Huan Liu

Applying Feature Selection to Short Time Wavelet Transformed Vibration Data for Reliability Analysis of an Ocean Turbine ................................................................................330
  Janell Duhaney, Taghi M. Khoshgoftaar, and Randall Wald

Session 12: Machine Learning Methods for Data Analysis

Backtracking for More Efficient Large Scale Dynamic Programming .........................................................338
  Charles Tripp and Ross Shachter

Gradient-Based Algorithms for the Automatic Construction of Fuzzy Cognitive Maps ......................................344
  Salomao Sampaio Madeiro and Fernando Jose Von Zuben

Randomized Sampling for Large Data Applications of SVM ........................................................................350
  Erik M. Ferragut and Jason Laska

Obtaining Full Regularization Paths for Robust Sparse Coding with Applications to Face Recognition ........356
  Jan Chorowski and Jacek Zurada

Session 13: Probabilistic and Model Based Learning

Online Variational Learning for a Dirichlet Process Mixture of Dirichlet Distributions and its Application ........362
  Wentao Fan and Nizar Bouguila

Personalized Recommendation in Folksonomies Using a Joint Probabilistic Model of Users, Resources and Tags 368
  Muzaffer Ege Alper and Şule Gündüz Öğüdüğü

Pattern Identification Using Reconstructed Phase Space and Hidden Markov Model ........................................374
  Wenjing Zhang and Xin Feng
Session 14: Clustering Methods

Compressive Clustering of High-Dimensional Data ................................................................. 380
   Andrzej Ruta and Fatih Porikli

Binary Function Clustering Using Semantic Hashes ................................................................. 386
   Wesley Jin, Sagar Chaki, Cory Cohen, Arie Gurfinkel, Jeffrey Havrilla,
   Charles Hines, and Priya Narasimhan

Automated Storage Tiering Using Markov Chain Correlation Based Clustering .................. 392
   Malak Alshawabkeh, Alma Riska, Adnan Sahin, and Motasem Awwad

Session 15: Graph-Based Learning

A Comparison of Graph Embedding Methods for Vertex Nomination ................................. 398
   Ming Sun, Minh Tang, and Carey E. Priebe

Graph Based Semi-supervised Non-negative Matrix Factorization for Document Clustering ................................................................. 404
   Naiyang Guan, Xuhui Huang, Long Lan, Zhigang Luo, and Xiang Zhang

Can Frequent Itemset Mining Be Efficiently and Effectively Used for Learning from Graph Data? ............................................................................................................................. 409
   Thashmee Karunaratne and Henrik Boström

Bias Selection Using Task-Targeted Random Subspaces for Robust Application of Graph-Based Semi-supervised Learning .................................................................................................................. 415
   Christopher T. Symons, Ranga Raju Vatsavai, Goo Jun, and Itamar Arel

Session 16: Multi-strategy, Multi-label, Multi-objective Learning

Cost-Sensitive Multi-strategy Active Annotation for Text Classification ............................... 421
   Shilpa Arora, Pinar Donmez, and Eric Nyberg

Multi-label Collective Classification Using Adaptive Neighborhoods ............................... 427
   Tanwistha Saha, Huzefa Rangwala, and Carlotta Domeniconi

Meta-learning and Model Selection in Multi-objective Evolutionary Algorithms .............. 433
   Martin Pilát and Roman Neruda

Poster Presentation Session

Supervised Dictionary Learning via Non-negative Matrix Factorization for Classification ................................................................. 439
   Yifeng Li and Alioune Ngom

A Series Inspired CPG Model for Robot Walking Control .................................................. 444
   Jiaqi Zhang, Xianchao Zhao, and Chenkun Qi
CPG and Reflexes Combined Adaptive Walking Control for AIBO ..................................................448
   Xianchao Zhao, Jiaqi Zhang, and Chenkun Qi

Integrating Machine Learning Into a Medical Decision Support System
to Address the Problem of Missing Patient Data .................................................................454
   Atif Khan, John A. Doucette, Robin Cohen, and Daniel J. Lizotte

Semi-Supervised Learning on Single-View Datasets by Integration of Multiple
Co-trained Classifiers ..................................................................................................................458
   Jelena Silvka, Ping Zhang, Aleksandar Kovačević, Zora Konjović, and Zoran Obradović

Scalable Overlapping Co-clustering of Word-Document Data ..................................................464
   Fabrício Olivetti De França

Real-Time Statistical Background Learning for Foreground Detection
under Unstable Illuminations ....................................................................................................468
   Dawei Li, Lihong Xu, and Erik Goodman

Ordinal Data Classification Using Kernel Discriminant Analysis: A Comparison
of Three Approaches ................................................................................................................473
   Jaime S. Cardoso, Ricardo Sousa, and Inês Domingues

Interval-Valued Centroids in K-Means Algorithms ....................................................................478
   Benjamine Nordin, Chenyi Hu, Bernard Chen, and Victor S. Sheng

A Novel Approach to Detect Smile Expression ........................................................................482
   Yuanzhi Zhang, Li Zhou, and Tao Sun

Structure Learning for Bayesian Networks Using the Physarum Solver ....................................488
   Torsten Schön, Martin Stetter, and Elmar W. Lang

An Integrated Machine Learning and Case-Based Reasoning Approach
to Answer Validation ..................................................................................................................494
   Ingo Glöckner and Karl-Heinz Weis

Using Local Regression in Monte Carlo Tree Search ...............................................................500
   Arisoa S. Randrianasolo and Larry D. Pyeatt

An Artificial Immune System Based on Holland's Classifier as Network
Intrusion Detection .....................................................................................................................504
   Arisoa S. Randrianasolo and Larry D. Pyeatt

Taxonomic Dimensionality Reduction in Bayesian Text Classification ....................................508
   Richard McAllister and John Sheppard

Artificial Neural Network for Prediction of Ethnicity Based on Iris Texture ................................514
   Anahita Zarei and Duxing Mou

Semantic Indexing of Video Simulations for Enhancing Medical Care During
Crises ........................................................................................................................................520
   Shuangshuang Jiang, Hichem Frigui, and Aaron W. Calhoun
Unsupervised Anomaly Detection in Transactional Data ..........................................................526
  Mohamed Bouguessa

Riemannian Shape Analysis Based on Meridian Curves ........................................................532
  Shuisheng Xie, Jundong Liu, and Charles D. Smith

The Assessment of the Quality of Sugar Using Electronic Tongue and Machine Learning Algorithms ...........................................................................................................538
  Tiemi C. Sakata, Katti Faceli, Tiago A. Almeida, Antonio Riul Júnior, and Wanessa M. D. M. F. Steluti

A Normalized Criterion of Spatial Clustering in Model-Based Framework ................................542
  X.Z. Wang, E. Grall-Maës, and P. Beuseroy

Muscle Categorization Using Quantitative Needle Electromyography: A 2-Stage Gaussian Mixture Model Based Approach ...............................................................548
  Meena Abdelmaseeh, Pascal Poupart, Benn Smith, and Daniel Stashuk

Physics-Based Constraints in the Forward Modeling Analysis of Time-Correlated Image Data .........................................................................................................................554
  James L. Carroll and Christopher D. Tomkins

Regularized Probabilistic Latent Semantic Analysis with Continuous Observations ........560
  Hao Zhang, Richard Edwards, and Lynne Parker

A Statistical Associative Classifier with Automatic Estimation of Parameters on Computer Aided Diagnosis ..........................................................................................................564
  Carolina Y. V. Watanabe, Marcela X. Ribeiro, Agma J. M. Traina, and Caetano Traina Jr

Mining Web to Detect Phishing URLs ..................................................................................568
  Ram B. Basnet and Andrew H. Sung

Macro-action Discovery Based on Change Point Detection and Boosting ..........................574
  Leonidas Lefakis and François Fleuret

A Multi-label and Adaptive Genre Classification of Web Pages ............................................578
  Chaker Jebari and M. Arif Wani

Identifying Accuracy of Social Tags by Using Clustering Representations of Song Lyrics .................................................................................................................................582
  Yajie Hu and Mitsunori Ogihara

Enhanced Multiagent Multi-objective Reinforcement Learning for Urban Traffic Light Control ...................................................................................................................586
  Mohamed Abdelaziz Khamis and Walid El-Sayed Ali Gomaa

On the Use of SVMs to Detect Anomalies in a Stream of SIP Messages ...............................592
  Raihana Ferdous, Renato Lo Cigno, and Alessandro Zorat
Decremental Learning of Evolving Fuzzy Inference Systems Using a Sliding Window ................................................................. 598

Manuel Bouillon, Eric Anquetil, and Abdullah Almaksour

Online Time Series Segmentation Using Temporal Mixture Models and Bayesian Model Selection ................................................. 602

Allou Samé and Gérard Govaert

Occluded Face Recognition Using Correntropy-Based Nonnegative Matrix Factorization ............................................................... 606

Tolga Ensari, Jan Chorowski, and Jacek M. Zurada

Reinforcement Learning for Production Ramp-Up: A Q-Batch Learning Approach ................................................................. 610

Stefanos Doltsinis, Pedro Ferreira, and Niels Lohse

Interaction Trees: Optimizing Ensembles of Decision Trees for Gene-Gene Interaction Detections ............................................................... 616

Amin Assareh, L. Gwenn Volkert, and Jing Li

Scalable Ensemble Learning by Adaptive Sampling .......................................................................................................................... 622

Jianhua Chen

Tensor-Based Temporal Behavior Analysis in Pain Medicine ........................................................................................................... 626

Andy Hall, Guangzhi Qu, Ishwar K. Sethi, and Craig Hartrick

Prediction of Protein Structures Using GPU Based Simulated Annealing ....................................................................................... 630

Hui Li and Chunmei Liu

Assessing Encoding Techniques through Correlation-Based Metrics ......................................................................................... 634

Prof. Giuliano Armano and Dr. Emanuele Tamponi

Predicting Congressional Votes Based on Campaign Finance Data ................................................................. 640

Samuel Smith, Jae Yeon (Claire) Baek, Zhaoyi Kang, Dawn Song, Laurent El Ghaoui, and Mario Frank

Polynomial Correlation Filters for Human Face Recognition ........................................................................................................... 646

Mohamed Alkanhal and Ghulam Muhammad

Block Level Video Steganalysis Scheme ................................................................................................................................. 651

Kesav Kancherla and Srinivas Mukkamala

Estimation of Susceptibility to Landslides Using Neural Networks Based on the FALCON-ART Model ............................................ 655

A. Viloria, C. Chang, M.C. Pineda, and J. Viloria

Prediction of the Calcium and Magnesium Content in Soils through a Generalized Regression Neural Networks and Genetic Algorithms ........................................................................................................... 661

Y. Labrador, C. Chang, and J. Viloria

A New Continuum Regression Method for Quantitative Analysis of Raman Spectrum ................................................................. 667

Shuo Li, Jean Gao, James O. Nyagilo, and Digant P. Dave
A Sequential Multi-task Learning Neural Network with Metric-Based Knowledge Transfer ................................................................. 671
   Simeng Yue and Seiichi Ozawa

Numerical Solution of Dirichlet Boundary Value Problems for Partial Differential Equations Using Quantum-Behaved Particle Swarm Optimization with Random Gaussian Function .................................................. 675
   Youngmin Ha

Venue Recommendation: Submitting Your Paper with Style ........................................................................................................... 681
   Zaihan Yang and Brian D. Davison

A Sampling-Based Approach to Reducing the Complexity of Continuous State Space POMDPs by Decomposition Into Coupled Perceptual and Decision Processes .................................................................................. 687
   Rassool Fakoor and Manfred Huber

Age-Group Classification of Facial Images ........................................................................................................................................ 693
   Li Liu, Jianming Liu, and Jun Cheng

A Game Theoretic Framework for Communication in Fully Observable Multiagent Systems ........................................................................................................... 697
   Tummalapalli Sudhamsh Reddy, Gergely Zaruba, and Manfred Huber

Author Index - Volume 1 ........................................................................................................................................................................ 703
2012 11th International Conference on Machine Learning and Applications

(ICMLA 2012)

Boca Raton, Florida, USA
12 – 15 December 2012

Volume 2
Pages 1-598
2012 11th International Conference on Machine Learning and Applications

ICMLA 2012

Table of Contents

Volume - 2

Preface - Volume 2.....................................................................................................................................xv
Conference Organization - Volume 2.................................................................xvii
Program Committee - Volume 2........................................................................xix
Workshops and Special Sessions - Volume 2....................................................xxiii
Keynotes - Volume 2..............................................................................................................................xxviii

Workshop A: Machine Learning Algorithms, Systems and Application

Workshop A: Regular Papers

Prognosis Based on Handling Drifts in Dynamical Environments: Application to a Wind Turbine Benchmark .................................................................1
  Antoine Chammas, Eric Duviella, and Stéphane Leceouche

A Machine Learning Based Topic Exploration and Categorization on Surveys .................................................7
  Clint P. George, Daisy Zhe Wang, Joseph N. Wilson, Liana M. Epstein,
  Philip Garland, and Annabell Suh

Error-Driven Adaptive, Virtual Machine Model-Based Control with High Availability Platform ..........................................................13
  Aman H. Bura, Bo Chen, and Li Yu

Improving Image Segmentation Using Genetic Algorithm .........................................................18
  Huynh Thi Thanh Binh, Mai Dinh Loi, and Nguyen Thi Thuy

The Importance of Outlier Relationships in Mobile Call Graphs .................................................24
  Derek Doran, Veena Mendiratta, Chitra Phadke, and Huseyn Uzunalioglu
Toward a Sequential Approach to Pipelined Image Recognition ...............................................................30
Derek Rose and Itamar Arel
Combining Parameter Space Search and Meta-learning for Data-Dependent Computational Agent Recommendation ...........................................................................................................36
Ondřej Kazík, Klára Pešková, Martin Plát, and Roman Neruda
Estimating Software Effort Using an ANN Model Based on Use Case Points ...............................................42
Ali Bou Nassif, Luiz Fernando Capretz, and Danny Ho

Workshop A: Short Papers
Feature Extraction and Selection in Ground Penetrating Radar with Experimental Data Set of Inclusions in Concrete Blocks ........................................................................................................48
F.A.A. Queiroz, D.A.G. Vieira, X.L. Travassos, and M.F. Pantoja
Recurrent Clustering for Unsupervised Feature Extraction with Application to Sequence Detection ...............................................................................................................................54
Steven Robert Young and Itamar Arel
Extraction of Strong Associations in Classes of Similarities ...........................................................................56
Ismaïl Biskri, Louis Rompré, Steve Descoteaux, Abdelghani Achouri, and Boucif Amar Benasaber
Generation of Tests for Programming Challenge Tasks on Graph Theory Using Evolution Strategy ............................................................................................................................................62
Maxim Buzdalov
Adaptive Selection of Helper-Objectives with Reinforcement Learning ............................................................66
Arina Buzdalova and Maxim Buzdalov
Online Ensemble Learning Approach for Server Workload Prediction in Large Datacenters ..................................................................................................................................................68
Nidhi Singh and Shrisha Rao

Workshop B: Class Imbalances on Past, Present and Future (CIPPF 2012)

Workshop B: Regular Papers
A Sequential Ensemble Classification (SEC) System for Tackling the Problem of Unbalance Learning: A Case Study ............................................................................................................................................72
Samaneh Sheikh-Nia, Gary Grewal, and Shawki Areibi
Using SVM with Adaptively Asymmetric MisClassification Costs for Mine-Like Objects Detection .................................................................................................................................78
Xiaoguang Wang, Hang Shao, Nathalie Japkowicz, Stan Matwin, Xuan Liu, Alex Bourque, and Bao Nguyen
Disturbing Neighbors Ensembles of Trees for Imbalanced Data ..........................83
Juan J. Rodríguez, José F. Diez-Pastor, Jesús Maudes, and César García-Osorio

Evaluation of SMOTE for High-Dimensional Class-Imbalanced Microarray Data .................................................................89
Rok Blagus and Lara Lusa

An Experimental Design to Evaluate Class Imbalance Treatment Methods .................................................................95
Gustavo Enrique De Almeida Prado Alves Batista, Diego Furtado Silva, and Ronaldo Cristiano Prati

One-Class versus Binary Classification: Which and When? ................................................102
Colin Bellinger, Shiven Sharma, and Nathalie Japkowicz

Multilevel Regression Models for Learning in the Presence of Rare Data ..........................................................107
Srinath Ravindran and Dennis Bahler

A Dynamic Sampling Framework for Multi-class Imbalanced Data ........................................113
B. Debowski, S. Areibi, G. Grewal, and J. Tempelman

Workshop B: Short Papers

A Boosting Method for Learning from Uneven Data for Improved Face Recognition ..........................................................119
Xiaohui Yuan and Mohamed Abouelenien

The Class-Imbalance Problem for High-Dimensional Class Prediction ................................................123
Lara Lusa and Rok Blagus

Workshop C: Machine Learning and Applications in Health Informatics

Workshop C: Regular Papers

PQL: Protein Query Language .................................................................127
Sherif Elfayoumy and Paul Bathen

Risk Estimation in Spatial Disease Clusters: An RBF Network Approach .........................................................133
Fernanda C. Takahashi and Ricardo H. C. Takahashi

Novel Margin Features for Mammographic Mass Classification ........................................................139
A. Bagheri-Khaligh, A. Zarghami, and M. T. Manzuri-Shalmani

Analyzing the Performance of Hierarchical Binary Classifiers for Multi-class Classification Problem Using Biological Data .........................................................145
Salma Begum and Ramazan S. Aygun

First Order Statistics Based Feature Selection: A Diverse and Powerful Family of Feature Selection Techniques ..............................................................151
Taghi Khoshgoftaar, David Dittman, Randall Wald, and Alireza Fazelpour
Sampling-Based Subnetwork Identification from Microarray Data and Protein-Protein Interaction Network .................................................................158  
  *Xiao Wang, Jinghua Gu, Jianhua Xuan, Ayesha N. Shajahan, Robert Clarke, and Li Chen*

Establishment of a Diagnostic Decision Support System in Genetic Dysmorphology ...............................................................................................................................164  
  *Kaya Kuru, Mahesan Niranjan, and Yusuf Tunca*

A New Fixed-Overlap Partitioning Algorithm for Determining Stability of Bioinformatics Gene Rankers ..........................................................................................170  
  *Randall Wald, Taghi Khoshgoftaar, and David Dittman*

Distributed Privacy Preserving Decision Support System for Predicting Hospitalization Risk in Hospitals with Insufficient Data ..........................................................................................................................178  
  *George Mathew and Zoran Obradovic*

Comparing Two New Gene Selection Ensemble Approaches with the Commonly-Used Approach ..................................................................................................................184  
  *David J. Dittman, Taghi M. Khoshgoftaar, Randall Wald, and Amri Napolitano*

Predicting Patient Outcomes from a Few Hours of High Resolution Vital Signs Data ................................................................................................................................192  
  *Tim Oates, Colin F. Mackenzie, Lynn G. Stansbury, Bizhan Aarabi, Deborah M. Stein, and Peter F. Hu*

The Effect of Number of Iterations on Ensemble Gene Selection .................................................................................................................................198  
  *Wael Awada, Taghi M. Khoshgoftaar, David Dittman, and Randall Wald*

**Workshop C: Short Papers**

Clinical Report Classification Using Natural Language Processing and Topic Modeling ................................................................................................................204  
  *Efsun Sarioglu, Hyeong-Ah Choi, and Kabir Yadav*

**Special Session I: Machine Learning in Information and System Security Issues**

**Special Session I: Regular Papers**

A New, Principled Approach to Anomaly Detection .................................................................................................................................210  
  *Erik M. Ferragut, Jason Laska, and Robert A. Bridges*

Web Spam: A Study of the Page Language Effect on the Spam Detection Features ................................................................................................................216  
  *Abdulrahman Alariti and Mansour Alsaleh*

A Web Based Face Counter Prediction System from Only Fingerprints ......................................................................................................................222  
  *Seref Sagiroglu and Uraz Yavanoglu*
An Analysis of Machine Learning Methods for Spam Host Detection ..................................................227  
Renato M. Silva, Akebo Yamakami, and Tiago A. Almeida

A New Intelligent Steganalysis Method for Waveform Audio Files ......................................................233  
Uraz Yavanoglu, Burak Ozcakmak, and Ozlem Milletsever

On the Validity of a New SMS Spam Collection ..........................................................................................240  
José María Gómez Hidalgo, Tiago A. Almeida, and Akebo Yamakami

Special Session II: Information Visualization in Machine Learning and Applications

Special Session II: Regular Papers
Interactive Visual Classification of Multivariate Data ................................................................................246  
Ke-Bing Zhang, Mehmet A. Orgun, Rajan Shankaran, and Du Zhang

A New Axes Re-ordering Method in Parallel Coordinates Visualization ..................................................252  
Liang Fu Lu, Mao Lin Huang, and Tze-Haw Huang

An Interactive Scatter Plot Metrics Visualization for Decision Trend Analysis .........................................258  
Tze-Haw Huang, Mao Lin Huang, and Kang Zhang

Fast Insight into High-Dimensional Parametrized Simulation Data ..........................................................265  
Daniel Butnaru, Benjamin Peherstorfer, Hans-Joachim Bungartz, and Dirk Pflüger

Special Session II: Short Papers
Cluster Analysis Using Improved Fuzzy c-Means Clustering for Object Localization: Vision Based III Cattle Localization System ..................................................................................................................271  
Jieun Kim and Woo Young Jung

Special Session III: Machine Learning Applications in Software Engineering

Special Session III: Regular Papers
Optimal Tile Size Selection Problem Using Machine Learning ..................................................................275  
Abid M. Malik

A Hybrid Approach to Coping with High Dimensionality and Class Imbalance for Software Defect Prediction ..............................................................................................................................................281  
Kehan Gao, Taghi M. Khoshgoftaar, and Amri Napolitano

Active Learning of Markov Decision Processes for System Verification ..................................................289  
Yingke Chen and Thomas Dyhre Nielsen

Machine Learning Module to Improve Communication between Agents in Multi-agent System .................295  
Shimaa M. El-Sherif, Behrouz Far, and Armin Eberlein
### Special Session IV: Machine Learning with Multimedia Data

#### Special Session IV: Regular Papers

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross-Domain Facial Expression Recognition Using Supervised Kernel Mean</td>
<td>Yun-Qian Miao, Rodrigo Araujo, and Mohamed S. Kamel</td>
</tr>
<tr>
<td>Automatic Identification of Crisis-Related Sub-events Using Clustering</td>
<td>Daniela Pohl, Abdelhamid Bouchchia, and Hermann Hellwagner</td>
</tr>
<tr>
<td>Classification of Urban Scenes from Geo-referenced Images in Urban Street-View Context</td>
<td>Corina Iovan, David Picard, Nicolas Thome, and Matthieu Cord</td>
</tr>
<tr>
<td>Composer Classification in Symbolic Data Using PPM</td>
<td>Antonio Deusany De Carvalho Junior and Leonardo Vidal Batista</td>
</tr>
<tr>
<td>Melodic Segmentation Using the Jensen-Shannon Divergence</td>
<td>Marcelo, Enrique, Rodríguez, López, and Anja Volk</td>
</tr>
<tr>
<td>Rethinking Automatic Chord Recognition with Convolutional Neural Networks</td>
<td>Eric J. Humphrey and Juan P. Bello</td>
</tr>
<tr>
<td>A Minimum Frame Error Criterion for Hidden Markov Model Training</td>
<td>Taemin Cho, Kibeom Kim, and Juan P. Bello</td>
</tr>
<tr>
<td>Classification, Segmentation and Chronological Prediction of Cinematic Sound</td>
<td>Pedro Silva</td>
</tr>
</tbody>
</table>

### Special Session V: Learning on the Web (LW)

#### Special Session V: Regular Papers

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animal Cognition, Epistemic Fluency, Social Networks and the Scientific Habit of Mind</td>
<td>Donald M. Morrison and Hu Xiangen</td>
</tr>
</tbody>
</table>
Hierarchical Classification Approach to Emotion Recognition in Twitter .............................................................381

Ahmed A.A. Esmin, Roberto L. De Oliveira Jr., and Stan Matwin

Predicting Dark Triad Personality Traits from Twitter Usage and a Linguistic Analysis of Tweets ...........................................386

Chris Sumner, Alison Byers, Rachel Boochever, and Gregory J. Park

Using Twitter Content to Predict Psychopathy ..............................................................................................................394

Randall Wald, Taghi M. Khoshgoftaar, Amri Napolitano, and Chris Sumner

Large Scale URL-based Classification Using Online Incremental Learning .................................................................402

Nidhi Singh, Harsimrat Sandhawalia, Nicolas Monet, Herve Poirier, and Jean-Marc Coursimault

Machine Learning and Text Mining of Trophic Links .................................................................................................410

Ghazal Afroozi Milani, David Bohan, Stuart Dunbar, Stephen Muggleton, Alan Raybould, and Alireza Tamaddoni-Nezhad

Special Session VI: Adaptive and Dynamic Modeling in Non-stationary Environments

Special Session VI: Regular Papers

Spatial Feature Extraction for Classification of Nonstationary Myoelectric Signals ..........................................................416

David Hofmann and J. Michael Herrmann

Maintaining Prior Distributions across Evolving Eigenspaces: An Application to Portfolio Construction ........................................422

Kevin R. Keane and Jason J. Corso

Adaptive Soft Sensor for Online Prediction Based on Moving Window Gaussian Process Regression ................................................428

Ratko Grbić, Dražen Slišković, and Petr Kadlec

Abrupt and Drift-Like Fault Diagnosis of Concurrent Discrete Event Systems ............................................................434

Moamar Sayed-Mouchaweh and Patrice Billaudel

Evolving Neural Fuzzy Network with Adaptive Feature Selection ................................................................................440

Alisson Marques Silva, Walmir Matos Caminhas, André Paim Lemos, and Fernando Gomide

Moving Object Detection in Aerial Video .......................................................................................................................446

Yunfei Wang, Zhaoxiang Zhang, and Yunhong Wang
Special Session VI: Short Papers
An Extension of the Consensus-Based Bundle Algorithm for Multi-agent Tasks with Task Based Requirements ................................................................................................................451
S. Hunt, Q. Meng, and C.J. Hinde

Special Session VII: Machine Learning in Energy Applications

Special Session VII: Regular Papers

Artificial Intelligence Based Hybrid Structures for Short-Term Load Forecasting Without Temperature Data ..................................................................................................................457
İdil Işikli Esener, Tolga Yüksel, and Mehmet Kurban

A Cooperative Learning Scheme for Energy Efficient Routing in Wireless Sensor Networks .................................................................................................................................463
Sami S. Alwakeel and Najla A. Al-Nabhan

The Prediction of Monthly Average Solar Radiation with TDNN and ARIMA ..........................................................................................................................469
Ji Wu and C.K. Chan

Measuring the Spatial Error in Load Forecasting for Electrical Distribution Planning as a Problem of Transporting the Surplus to the In-Deficit Locations .............................................................................................................475

Finding Motifs in Wind Generation Time Series Data ..........................................................................................................................481
Chandrika Kamath and Ya Ju Fan

Position Control of a DC Motor Used in Solar Panels with Artificial Neural Network .........................................................................................................................487
Murat Sahin, H. Ibrahim Bulbul, and Ilhami Colak

Approach to Control of the Output Voltage in Renewable Energy Sources on the Basis of AE-method Using Genetic Algorithm ..........................................................................................................................493
V. Ten, N. Isembergenov, Y. Akhmetbekov, D. Sarbassov, A. Iglikov, and B. Matkarimov

Application of Adaptive Artificial Neural Network Method to Model the Excitation Currents of Synchronous Motors ..........................................................................................................................498
Ramazan Bayindir, Ilhami Colak, Seref Sagiroglu, and Hamdi Tolga Kahraman

A Novel Neural Network Based Control Method with Adaptive On-Line Training for DC-DC Converters ..........................................................................................................................503
Hidenori Maruta, Masashi Motomura, and Fujio Kurokawa

Multivariate Assessment of a Repair Program for a New York City Electrical Grid ..........................................................................................................................509
Rebecca J. Passonneau, Ashish Tomar, Somnath Sarkar, Haimonti Dutta, and Axinia Radeva
Unsupervised Disaggregation for Non-intrusive Load Monitoring ..........................................................515
   Sundeep Pattem

Excitation Current Forecasting for Reactive Power Compensation in Synchronous Motors: A Data Mining Approach ..........................................................521
   Ramazan Bayindir, Mehmet Yesilbudak, Ilhami Colak, and Seref Sagiroglu

Monitoring and Determination of Wind Energy Potential by Web Based Wireless Network ..........................................................526
   Onur Keskin, İSmet Ateş, Ziya Haktan Karadeniz, Alpaslan Turgut, and Zeki KiRal

**Special Session VIII: Machine Learning Ensemble Methods and Applications**

**Special Session VIII: Regular Papers**

Decision Level Fusion of Wavelet Features for Ocean Turbine State Detection ..................................................531
   Janell Duhaney and Taghi M. Khoshgoftaar

Exploiting Representational Diversity for Time Series Classification ..........................................................538
   Tim Oates, Colin F. Mackenzie, Deborah M. Stein, Lynn G. Stansbury,
   Joseph Dubose, Bizhan Aarabi, and Peter F. Hu

Finding Survival Groups in SEER Lung Cancer Data ..................................................................................545
   Iryna Skrypnyk

A Novel Noise-Resistant Boosting Algorithm for Class-Skewed Data ..........................................................551
   Jason Van Hulse, Taghi M. Khoshgoftaar, and Amri Napolitano

Jensen-Shannon Divergence in Ensembles of Concurrently-Trained Neural Networks ..........................................................558
   Aaron Mishtal and Itamar Arel

O-linked Glycosylation Site Prediction Using Ensemble of Graphical Models ..........................................................563
   Aditya Sriram and Feng Luo

A Hybrid Method for Estimating the Predominant Number of Clusters in a Data Set ..........................................................569
   Jamil Alshaqsi and Wenjia Wang

**Special Session IX: ICMLA 2012 Avatar Face Recognition Challenge**

Face Recognition Challenge: Object Recognition Approaches for Human/Avatar Classification ..........................................................574
   Toshihiko Yamasaki and Tsuhan Chen

Convolutional Neural Networks Applied to Human Face Classification ..........................................................580
   Brian Cheung
Learning Visual Features for the Avatar Captcha Recognition Challenge ...........................................584
  Mohammed Korayem, Abdallah A. Mohamed, David Crandall,
  and Roman V. Yampolskiy

Ensemble Feature Selection in Face Recognition: ICMLA 2012 Challenge ...........................................588
  Salem Alelyani and Huan Liu

ICMLA Face Recognition Challenge – Results of the Team Computational Intelligence Mittweida .................................................................592
  Thomas Villmann, Marika Kästner, David Nebel, and Martin Riedel

Author Index - Volume 2 ..................................................................................................................................596