

# **2014 13th International Conference on Machine Learning and Applications**

**(ICMLA 2014)**

**Detroit, Michigan, USA  
3-6 December 2014**



**IEEE Catalog Number: CFP14592-POD  
ISBN: 978-1-4799-7416-0**

# 2014 13th International Conference on Machine Learning and Applications

## ICMLA 2014

### Table of Contents

Preface.....	xiv
Conference Organization.....	xv
Program Committees.....	xvii

---

#### Main Conference

Arctic Sea Ice Extent Forecasting Using Support Vector Regression .....	1
<i>Tyler G.R. Reid and Paul M. Tarantino</i>	
Budgeted Learning for Developing Personalized Treatment .....	7
<i>Kun Deng, Russ Greiner, and Susan Murphy</i>	
Modelling Mutual Information between Voiceprint and Optimal Number of Mel-Frequency Cepstral Coefficients in Voice Discrimination .....	15
<i>Kin Wah Edward Lin, Tian Feng, Natalie Agus, Clifford So, and Simon Lui</i>	
High Precision Screening for Android Malware with Dimensionality Reduction .....	21
<i>Britton Wolfe, Karim Elish, and Danfeng Yao</i>	
A Hybrid Genetic-Programming Swarm-Optimisation Approach for Examining the Nature and Stability of High Frequency Trading Strategies .....	29
<i>Andreea-Ingrid Funie, Mark Salmon, and Wayne Luk</i>	
Protein Conformation Motion Modeling Using Sep-CMA-ES .....	35
<i>Maxim Buzdalov, Sergey Knyazev, and Yury Porozov</i>	
LSH vs Randomized Partition Trees: Which One to Use for Nearest Neighbor Search? .....	41
<i>Kaushik Sinha</i>	
Bayesian Nonparametric Inverse Reinforcement Learning for Switched Markov Decision Processes .....	47
<i>Amit Surana and Kunal Srivastava</i>	

Varying Coefficient Models for Analyzing the Effects of Risk Factors on Pregnant Women's Blood Pressure .....	55
<i>Wenshuai Cheng, Liying Fang, Lin Yang, Han Zhao, Pu Wang, and Jianzhuo Yan</i>	
Ensemble Statistical and Heuristic Models for Unsupervised Word Alignment .....	61
<i>Mahsa Mohaghegh, Hossein Sarrafzadeh, and Mehdi Mohammadi</i>	
Reducing the Cost of Breaking Audio CAPTCHAs by Active and Semi-supervised Learning .....	67
<i>Malte Darnstädt, Hendrik Meutzner, and Dorothea Kolossa</i>	
LaCova: A Tree-Based Multi-label Classifier Using Label Covariance as Splitting Criterion .....	74
<i>Reem Al-Otaibi, Meelis Kull, and Peter Flach</i>	
A Cyclic Contrastive Divergence Learning Algorithm for High-Order RBMs .....	80
<i>Dingsheng Luo, Yi Wang, Xiaoqiang Han, and Xihong Wu</i>	
Learning to Rank with Only Positive Examples .....	87
<i>Mingzhu Zhu, Wei Xiong, and Yi-Fang Brook Wu</i>	
Learning Score Systems for Patient Mortality Prediction in Intensive Care Units via Orthogonal Matching Pursuit .....	93
<i>Aadirupa Saha, Chandrahas Dewangan, Harikrishna Narasimhan, Sriram Sampath, and Shivani Agarwal</i>	
Improving Robustness of Gaussian Process-Based Inferential Control System Using Kernel Principle Component Analysis .....	99
<i>Ali Abusnina, Daniel Kudenko, and Rolf Roth</i>	
State Abstraction in Reinforcement Learning by Eliminating Useless Dimensions .....	105
<i>Zhao Cheng and Laura E. Ray</i>	
A Knowledge Growth and Consolidation Framework for Lifelong Machine Learning Systems .....	111
<i>Fernando Martínez Plumed, Cèsar Ferri, José Hernández Orallo, and María José Ramírez Quintana</i>	
Improving Named Entity Recognition for Morphologically Rich Languages Using Word Embeddings .....	117
<i>Hakan Demir and Arzucan Özgür</i>	
Semi-supervised Kernel-Based Temporal Clustering .....	123
<i>Rodrigo Araujo and Mohamed S. Kamel</i>	
Adding Diversity to Rank Examples in Anytime Nearest Neighbor Classification .....	129
<i>Cristiano Inácio Lemes, Diego Furtado Silva, and Gustavo E.A.P.A. Batista</i>	
An Accurate, Fast Embedded Feature Selection for SVMs .....	135
<i>Tarfa Hamed, Rozita Dara, and Stefan C. Kremer</i>	

A Switch-and-Restart Algorithm with Exponential Restart Strategy for Objective Selection and its Runtime Analysis .....	141
<i>Maxim Buzdalov</i>	
Improving Spectral Learning by Using Multiple Representations .....	147
<i>Adam Drake and Dan Ventura</i>	
Using Spectral Features to Improve Sentiment Analysis .....	153
<i>Adam Drake and Dan Ventura</i>	
Human Action Recognition Based on MOCAP Information Using Convolution Neural Networks .....	159
<i>Earnest Paul Ijjina and C. Krishna Mohan</i>	
Activity Recognition Using Graphical Features .....	165
<i>Syeda Selina Akter and Lawrence B. Holder</i>	
Sequential Logistic Principal Component Analysis (SLPCA): Dimensional Reduction in Streaming Multivariate Binary-State System .....	171
<i>Zhaoyi Kang and Costas J. Spanos</i>	
Human Action Recognition Based on Recognition of Linear Patterns in Action Bank Features Using Convolutional Neural Networks .....	178
<i>Earnest Paul Ijjina and C. Krishna Mohan</i>	

## Short Papers

Reducing the Effects of Detrimental Instances .....	183
<i>Michael R. Smith and Tony Martinez</i>	
Adaptive Restructuring of Radial Basis Functions Using Integrate-and-Fire Neurons .....	189
<i>Jeremy A. Marvel</i>	
On-Line Signature Verification Using Symbolic Aggregate Approximation (SAX) and Sequential Mining Optimization (SMO) .....	195
<i>Rakesh Deivachilai and Tim Oates</i>	
Improved kNN Rule for Small Training Sets .....	201
<i>Sunsern Cheamanunkul and Yoav Freund</i>	
Extraction of Unexpected Rules from Twitter Hashtags and its Application to Sport Events .....	207
<i>Mariam Adedoyin-Olowe, Mohamed Medhat Gaber, Carlos Martin Dancausa, and Frederic Stahl</i>	
A Machine Learning Approach to Combining Individual Strength and Team Features for Team Recommendation .....	213
<i>Haibin Liu, Mu Qiao, Daniel Greenia, Rama Akkiraju, Stephen Dill, Taiga Nakamura, Yang Song, and Hamid Motahari Nezhad</i>	

Topic Detection in Instant Messages .....	219
<i>Han Zhang, Chang-Dong Wang, and Jian-Huang Lai</i>	
WiFi Localization for Mobile Robots Based on Random Forests and GPLVM .....	225
<i>Reda Elbasiony and Walid Gomaa</i>	
Geometric PDEs on Weighted Graphs for Semi-supervised Classification .....	231
<i>Mathieu Toutain, Abderrahim Elmoataz, and Olivier Lézoray</i>	
Visualising Singing Style under Common Musical Events Using Pitch-Dynamics Trajectories and Modified TRACLUS Clustering .....	237
<i>Kin Wah Edward Lin, Hans Anderson, Natalie Agus, Clifford So, and Simon Lui</i>	
Multi-variable Neural Network Forecasting Using Two Stage Feature Selection .....	243
<i>Rohit Rawat, Kunal Vora, Michael Manry, and Gautam Eapi</i>	
Genetically Supervised Self-Organizing Map for the Classification of Glass Samples .....	251
<i>Richard de Groof and Iren Valova</i>	
Q-Learning: From Computer Network Security to Software Security .....	257
<i>Arisoa S. Randrianasolo and Larry D. Pyeatt</i>	
Example-Dependent Cost-Sensitive Logistic Regression for Credit Scoring .....	263
<i>Alejandro Correa Bahnsen, Djamia Aouada, and Björn Ottersten</i>	
Time Warping Symbolic Aggregation Approximation with Bag-of-Patterns Representation for Time Series Classification .....	270
<i>Zhiguang Wang and Tim Oates</i>	
Variational Inference on Infinite Mixtures of Inverse Gaussian, Multinomial Probit and Exponential Regression .....	276
<i>S.K. Minhazul Islam and Arunava Banerjee</i>	
Recommendation Systems for Markets with Two Sided Preferences .....	282
<i>Anjan Goswami, Fares Hedayati, and Prasant Mohapatra</i>	
An Analysis of Instance Selection for Neural Networks to Improve Training Speed .....	288
<i>Xunhu Sun and Philip K. Chan</i>	
Concept Drift Awareness in Twitter Streams .....	294
<i>Joana Costa, Catarina Silva, Mário Antunes, and Bernardete Ribeiro</i>	
Automated Scoring of the Level of Integrative Complexity from Text Using Machine Learning .....	300
<i>Aardra Kannan Ambili and Khaled M. Rasheed</i>	
Feature Selections for Effectively Localizing Faulty Events in GUI Applications .....	306
<i>Xiaozhen Xue, Yulei Pang, and Akbar Siami Namin</i>	

Uncertainty Quantified Matrix Completion Using Bayesian Hierarchical Matrix Factorization .....	312
<i>Farideh Fazayeli, Arindam Banerjee, Jens Kattge, Franziska Schrod, and Peter B. Reich</i>	
Post-Processing Association Rules Using Networks and Transductive Learning .....	318
<i>Renan de Padua, Solange Oliveira Rezende, and Veronica Oliveira de Carvalho</i>	
Detection of Abnormal Human Behavior Using a Matrix Approximation-Based Approach .....	324
<i>Lijun Wang and Ming Dong</i>	
Dimensionality Reduction in Statistical Learning .....	330
<i>Alexander Bernstein and Alexander Kuleshov</i>	
Supervised Music Chord Recognition .....	336
<i>Imad Rida, Romain Hault, and Gilles Gasso</i>	
Computation of a Rejection Threshold Used for the Bayes Classifier .....	342
<i>Matthias Blankenburg, Christian Bloch, and Jörg Krüger</i>	
Combining Exact and Metaheuristic Techniques for Learning Extended Finite-State Machines from Test Scenarios and Temporal Properties .....	350
<i>Daniil Chivilikhin, Vladimir Ulyantsev, and Anatoly Shalyto</i>	
Applying Swarm Ensemble Clustering Technique for Fault Prediction Using Software Metrics .....	356
<i>Rodrigo A. Coelho, Fabrício dos R.N. Guimarães, and Ahmed A.A. Esmin</i>	
Expert Bayes: Automatically Refining Manually Built Bayesian Networks .....	362
<i>Ezilda Almeida, Pedro Ferreira, Tiago T.V. Vinhoza, Inês Dutra, Paulo Borges, Yirong Wu, and Elizabeth Burnside</i>	
Improving Performance on Problems with Few Labelled Data by Reusing Stacked Auto-Encoders .....	367
<i>Telmo Amaral, Chetak Kandaswamy, Luís M. Silva, Luís A. Alexandre, Joaquim Marques de Sá, and Jorge M. Santos</i>	
Learning Good Features to Track .....	373
<i>Raed Almomani, Ming Dong, and Zhou Liu</i>	
Implementation of Machine Learning for Classifying Hemiplegic Gait Disparity through Use of a Force Plate .....	379
<i>Robert LeMoyne, Wesley Kerr, Timothy Mastroianni, and Anthony Hessel</i>	
Using k-Nearest Neighbor and Speaker Ranking for Phoneme Prediction .....	383
<i>Muhammad Rizwan and David V. Anderson</i>	
One-Shot Periodic Activity Recognition Using Convolutional Neural Networks .....	388
<i>Earnest Paul Ijjina and C. Krishna Mohan</i>	

Facial Expression Recognition Using Kinect Depth Sensor and Convolutional Neural Networks .....	392
<i>Earnest Paul Ijjina and C. Krishna Mohan</i>	

## **Workshop: Machine Learning for Predictive Models**

Using Balanced Random Forests on Load Spectrum Data for Classifying Component Failures of a Hybrid Electric Vehicle Fleet .....	397
<i>Philipp Bergmeir, Christof Nitsche, Jürgen Nonnast, Michael Bargende, Peter Antony, and Uwe Keller</i>	
Multimodal Sparsity-Eager Support Vector Machines for Music Classification .....	405
<i>Kamelia Aryafar and Ali Shokoufandeh</i>	
A Better Case Adaptation Method for Case-Based Effort Estimation Using Multi-objective Optimization .....	409
<i>Ali Bou Nassif and Shadi Banitaan</i>	
Employing Markov Networks on Curriculum Graphs to Predict Student Performance .....	415
<i>Ahmad Slim, Gregory L. Heileman, Jarred Kozlick, and Chaouki T. Abdallah</i>	
Leveraging Machine Learning Algorithms to Perform Online and Offline Highway Traffic Flow Predictions .....	419
<i>Arezou Moussavi-Khalkhali, Mo Jamshidi, and Lutcher Brown Endowed Chair</i>	
OUPS: A Combined Approach Using SMOTE and Propensity Score Matching .....	424
<i>William A. Rivera, Amit Goel, and J. Peter Kincaid</i>	
A Comparison of Supervised Machine Learning Techniques for Predicting Short-Term In-Hospital Length of Stay among Diabetic Patients .....	428
<i>April Morton, Eman Marzban, Georgios Giannoulis, Ayush Patel, Rajender Aparasu, and Ioannis A. Kakadiaris</i>	

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

Diagnosis Using Incomplete Model in Fuzzy Discrete Event System: Application to Crisis Management .....	432
<i>Moussa Traore, Eric Chatelet, Eddie Soulier, and Hossam Gaber</i>	
Incremental SVD for Insight into Wind Generation .....	441
<i>Chandrika Kamath and Ya Ju Fan</i>	
Transient Characteristics of DC-DC Converter with PID Parameters Selection and Neural Network Control .....	447
<i>Hidenori Maruta, Daiki Mitsutake, and Fujio Kurokawa</i>	
Intelligent Crude Oil Price Forecaster .....	453
<i>Ardalan Tebyanian and Fares Hedayati</i>	

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

Next Generation Application-Layer DDoS Defences: Applying the Concepts of Outlier Detection in Data Streams with Concept Drift .....	456
<i>Dusan Stevanovic and Natalija Vlajic</i>	
TSD: Detecting Sybil Accounts in Twitter .....	463
<i>Mansour Alsaleh, Abdulrahman Alarifi, Abdul Malik Al-Salman, Mohammed Alfayez, and Abdulmajeed Almuhsin</i>	
An Intelligent Technique for Detecting Malicious Users on Mobile Stores .....	470
<i>Ramazan Terzi, Uraz Yavanoglu, Duygu Sinanc, Dogac Oguz, and Semra Cakir</i>	
Age Estimation from Fingerprints: Examination of the Population in Turkey .....	478
<i>Eyüp Burak Ceyhan, Şeref Sağroğlu, Ramazan Cesur, and Kadriye Öner</i>	

## **Special Session: Learning on the Web**

Comparative Study of Different Classification Techniques: Heart Disease Use Case .....	482
<i>Hanen Bouali and Jalel Akaichi</i>	

## **Special Session: Adaptive Data-Driven Modeling in Dynamic Environments and On-Line Processes**

Dynamic Inclusion of New Event Types in Visual Inspection Using Evolving Classifiers .....	487
<i>Edwin Lughofer, Eva Weigl, Wolfgang Heidl, and Thomas Radauer</i>	
Speeding Learning of Personalized Audio Equalization .....	495
<i>Bongjun Kim and Bryan Pardo</i>	
A Semi-supervised Clustering Approach for Semantic Slot Labelling .....	500
<i>Heriberto Cuayáhuitl, Nina Dethlefs, and Helen Hastie</i>	

## **Special Session: Machine Learning in Visual Information Processing**

Multimodal Music and Lyrics Fusion Classifier for Artist Identification .....	506
<i>Kamelia Aryafar and Ali Shokoufandeh</i>	
Machine-Sourced Segmentations vs. Expert-Sourced Segmentations for the Classification of Lung Nodules with Outlier Removal .....	510
<i>Mayra Alejandra Molina Puentes, Jacob Furst, and Daniela Stan Raicu</i>	
Tool Machines with Brains - Touchless Wheel Alignment with Neural Networks .....	516
<i>Karl-Heinz Weis</i>	



Performance Comparison of Major Classical Face Recognition Techniques .....521  
*Farooq Ahmad Bhat and M. Arif Wani*

Assessment of Different Image Clutter Metrics Using Multivariate Analyses  
and Neurofuzzy System .....529  
*Deok Hee Nam and Harpreet Singh*

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

A Genetic Algorithm Approach to Partitioning Clustering: A Case Study  
on M.Sc. Applicants .....535  
*Kittichai Lavangnananda and Ratipong Poolphol*

American Sign Language Recognition Using Leap Motion Sensor .....541  
*Ching-Hua Chuan, Eric Regina, and Caroline Guardino*

Learner Engagement Measurement and Classification in 1:1 Learning .....545  
*Sinem Aslan, Zehra Cataltepe, Itai Diner, Onur Dundar, Asli A. Esme,  
Ron Ferens, Gila Kamhi, Ece Oktay, Canan Soysal, and Murat Yener*

An Intelligent Tutoring System for Argument-Making in Higher Education: A  
Pilot Study .....553  
*Ching-Hua Chuan, Daniel Dinsmore, Joseph Schmuller, and Tyler Morris*

Automatically Filtering Irrelevant Words for Applications in Language  
Acquisition .....557  
*Gihad Sohsah, Emrah Akkurt, Ilkin Safarli, Muhammed Unal, and Onur Guzey*

A Clustering-Based Grouping Model for Enhancing Collaborative Learning .....562  
*Yulei Pang, Feiya Xiao, Huaying Wang, and Xiaozhen Xue*

Improving an Early Warning System to Prediction of Student Examination  
Achievement .....568  
*Hanife Goker and Halil Ibrahim Bulbul*

Investigating Sentimental Relation between Social Media Presence  
and Academic Success of Turkish Universities .....574  
*Sedef Gunduz, Fatih Demirhan, and Seref Sagiroglu*

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

Improved Selection of Auxiliary Objectives Using Reinforcement Learning  
in Non-stationary Environment .....580  
*Irina Petrova, Arina Buzdalova, and Maxim Buzdalov*

A New Algorithm for Adaptive Online Selection of Auxiliary Objectives .....584  
*Arina Buzdalova and Maxim Buzdalov*

Iterative Hard Thresholding for Keyword Extraction from Large Text Corpora .....	588
<i>Steve Yadlowsky, Preetum Nakkarin, Jingyan Wang, Rishi Sharma, and Laurent El Ghaoui</i>	
A Novel Bayesian Network Based Scheme for Finding the Optimal Solution to Stochastic Online Equi-partitioning Problems .....	594
<i>Sondre Glimsdal and Ole-Christoffer Granmo</i>	
Automatic Gender Classification System from Finger 2D: 4D Ratio and Comparison of Successes with Using Different Algorithm .....	600
<i>Eyüp Burak Ceyhan, Seref Sagiroglu, Ramazan Cesur, and Kadriye Öner</i>	
 <b>Workshop: Machine Learning of Graphical Models in Static and Dynamic Complex Environments</b>	
Causal Discovery from Spatio-Temporal Data with Applications to Climate Science .....	606
<i>Imme Ebert-Uphoff and Yi Deng</i>	
Graphical Model Based Approach for Fault Diagnosis of Wind Turbines .....	614
<i>Adel Aloraini and Moamar Sayed-Mouchaweh</i>	
A Directed Acyclic Graphical Approach and Ensemble Feature Selection for a Better Drug Development Strategy Using Partial Knowledge from KEGG Signalling Pathways .....	620
<i>Adel Aloraini</i>	
 <b>Author Index</b> .....	 625