The 4th Asian Conference on Machine Learning

November 4–6, 2012

Singapore Management University

Singapore

ACML2012

http://acml12.comp.nus.edu.sg/

ACML 2012 PROGRAM AT A GALANCE

		November 4, 2012 (Sunday)
08:30am - 18:30pm	SMU School of Accountancy	Registration
08:30am - 10:10am	Auditorium	**Tutorial 1:** Bandit Games *Sebastien Bubeck*
	Seminar Room 2-2	**Workshop 1:** LAWS'12
10:10am - 10:30am	Outside Auditorium	Coffee Break
10:30am - 11:20am	Auditorium	**Tutorial 1:** Bandit Games (cont')
11:20am - 12:30pm	Auditorium	**Tutorial 2:** Domain Adaptation in Real World Applications *Fei Sha, Ivor W. Tsang, Sinno Pan*
10:30am - 12:30pm	Seminar Room 2-2	**Workshop 1:** LAWS'12 (cont')
12:30pm - 13:30pm	Outside Auditorium	Lunch
13:30pm - 15:10pm	Auditorium	**Tutorial 2**: Domain Adaptation in Real World Applications (cont')
	Seminar Room 2-2	**Workshop 2:** FDMA'12
15:10pm - 15:30pm	Outside Auditorium	Coffee Break
15:30pm - 17:00pm	Auditorium	**Tutorial 3:** Probabilistic Modeling of Ranking *Jose A. Lozano, Ekhine Irurozki*
	Seminar Room 2-2	**Workshop 2:** FDMA'12 (cont')
17:00pm - 17:20pm	Outside Auditorium	Coffee Break
17:20pm - 18:30pm	Auditorium	**Tutorial 3:** Probabilistic Modeling of Ranking (cont')
18:30pm - 21:00pm	Reception Venue	**Reception**

ACML 2012 PROGRAM AT A GALANCE

		November 5, 2012 (Monday)
08:30am - 18:30pm	SMU School of Accountancy	**Registration**
08:30am - 09:00am	Auditorium	**Opening Session**
09:00am - 10:10am	Auditorium	**Keynote Talk 1:** Behavior Imaging and the Study of Autism *James Rehg (Georgia Tech)*
10:10am - 10:30am	Outside Auditorium	Coffee Break
10:30am - 12:00pm	Auditorium	**Oral Session 1:** Unsupervised Learning and Deep Learning
12:00pm - 14:00pm	SMU Library	**Poster Session 1** + Lunch
14:00pm - 15:10pm	Auditorium	**Keynote Talk 2:** Convex Methods for Representation Learning *Dale Schuurmans (University of Alberta)*
15:10pm - 15:30pm	Outside Auditorium	Coffee Break
15:30pm - 17:00pm	Auditorium	**Oral Session 2:** Feature Selection and Dimension Reduction
17:00pm - 17:20pm	Outside Auditorium	Coffee Break
17:20pm - 18:30pm	Auditorium	**Oral Session 3:** Learning in Graphs, Networks and Structures
18:30pm - 21:00pm	Banquet Venue	**Banquet**

ACML 2012 PROGRAM AT A GALANCE

		November 6, 2012 (Tuesday)
08:30am - 18:30pm	SMU School of Accountancy	**Registration**
08:30am - 09:00am	Auditorium	**Best Paper Session:** Variational Bayesian Matching *Arto Klami*
09:00am - 10:10am	Auditorium	**Keynote Talk 3:** Multiclass Losses and Multidistribution divergences *Bob Williamson (ANU and NICTA)*
10:10am - 10:30am	Outside Auditorium	Coffee Break
10:30am - 12:00pm	Auditorium	**Oral Session 4:** Classification and Ranking
12:00pm - 14:00pm	SMU Library	**Poster Session 2** + Lunch
14:00pm - 15:10pm	Auditorium	**Oral Session 5:** Supervised and Semi-supervised Learning
15:10pm - 15:30pm	Outside Auditorium	Coffee Break
15:30pm - 17:00pm	Auditorium	**Oral Session 6:** Learning Theory, Reinforcement and Online Learning
17:00pm - 17:20pm	Auditorium	**Closing Session**
17:20pm - 21:00pm		**Social Events**

WELCOME MESSAGE

The Asian Conference on Machine Learning aims to provide a leading venue for publication and dissemination of new ideas and achievements in machine learning. The fourth in the series, the conference this year is held in the city of Singapore, following the first three conferences in Nanjing in China, Tokyo in Japan, and Taoyuan in Taiwan.

The conference has grown from strength to strength. While located in Asia, the conference has always drawn wide international participation. This year, we received submissions from 30 countries from Asia, Australasia, Europe, North America and South America. There were 138 submissions, of which 36 were accepted into the proceedings, published by JMLR Workshop and Conference Proceedings series, giving an acceptance rate of 26%. We experimented with a two-cycle submission process this year, where each cycle had its own submission deadline. Papers that could not be accepted in the first cycle, but had the potential to be acceptable after careful improvements, were encouraged for resubmission in the second cycle, allowing the referees' concerns to be addressed. Fresh submissions that were too late to make the first deadline were also welcomed for the second cycle. We think that the new process is successful but would welcome hearing your feedback on it.

We are keeping the ACML tradition of being a single track conference, with some papers presented as short presentations. This year we invited all accepted papers to give poster presentations, in addition to oral presentations. We would like to thank the program chairs, Wray Buntine and Steven Hoi, together with the program committee, for their outstanding effort in putting together the program. We believe that the papers are of high quality and hope that you enjoy the conference.

We have an exciting program for this conference with invited talks from Jim Rehg, Dale Schuurmans and Bob Williamson, three tutorials on bandit problems, domain adaptation and probabilistic ranking, and two workshops on weak supervision as well as fraud detection. We would like to thank the invited speakers, the tutorial presenters and the workshop organizers for their time and effort, for helping the conference and the machine learning community in general. We are also grateful to our sponsors, the Air Force Office of Scientific Research and Asian Office of Aerospace Research and Development, Huawei, and Lee Foundation for their financial support. We would also like to thank the Machine Learning Journal for sponsoring the best student paper award. We are also grateful to the steering committee for invaluable guidance. Last but not least, we would like to thank the organizing committee for all the hard work that they have put into making the conference a success.

Welcome to Singapore!

ACML 2012 General Co-Chairs
Wee Sun Lee (National University of Singapore) and Zhi-Hua Zhou (Nanjing University)

TABLE OF CONTENTS

NOVEMBER 4 (SUNDAY)

08:30AM – 11:20AM (Auditorium, School of Accountancy, SMU)

TUTORIAL 1

Title: Bandit Games

Speaker: Sebastien Bubeck (Princeton University)

In the recent years the multi-armed bandit problem has attracted a lot of attention in the theoretical learning community. This growing interest is a consequence of the large number of problems that can be modeled as a multi-armed bandit: ad placement, website optimization, packet routing, ect. Furthermore the bandit methodology is also used as a building block for more complicated scenarios such as reinforcement learning, model selection in statistics, or computer game-playing. While the basic stochastic multi-armed bandit can be traced back to Thompson (1933) and Robbins (1952), it is only very recently that we obtained an (almost) complete understanding of this simple model. Moreover many extensions of the original problem have been proposed in the past fifteen years, such as bandits without a stochastic assumption (the so-called adversarial model), or bandits with a very large (but structured) set of arms.

The tutorial will be divided into three parts:

- In the first part we discuss the state-of-the-art results on the basic multi-armed bandit problem (both stochastic and adversarial).
- In the second part the focus will be on continuously-armed stochastic bandits, with a Lipschitz assumption on the mean-payoff.
- Finally in the third part we consider the case of adversarial bandits, with a linear loss and a very large set of arms with some combinatorial structure.

11:20AM – 15:10PM (Auditorium, School of Accountancy, SMU)

TUTORIAL 2

Title: Domain Adaptation in Real World Applications

Speakers: Fei Sha (University of Southern California), **Ivor W. Tsang** (Nanyang Technological University), **Sinno Pan** (Institute for Infocomm Research)

Domain adaptation (a.k.a., cross-domain learning or transfer learning) has become an increasingly important research direction in machine learning and many other application areas. It is well-known that features sampled from different domains may differ tremendously in their distributions, such as mean, intra-class/inter-class variance. Domain adaptation methods have been developed to cope with such mismatch among domains. Recently they have also been successfully used in a broad range of real-world applications, in which the (target) domain of interest contains very few labeled samples, while an existing (auxiliary) domain is often available with a large number of labeled examples. We propose this tutorial with three specific objectives: 1) provide an overview of this fast-growing research area; ii) describe representative methods and their real-world applications to natural language processing, speech recognition and computer vision; iii) raise awareness of research challenges and engage discussions on future directions in this area.

NOVEMBER 4 (SUNDAY)

15:30PM – 18:30PM (Auditorium, School of Accountancy, SMU)
TUTORIAL 3
Title: Probabilistic Modeling of Ranking
Speakers: Jose A. Lozano (University of the Basque Country), **Ekhine Irurozki** (University of the Basque Country)
Rankings and permutations have become, nowadays, ubiquitous. They appear in numerous areas of computer systems: information retrieval, recommender systems, identity tracking or chemical compound classification, etc. Dealing with rankings, and particularly with rankings of many objects is a complex computational task as the number of permutations of n objects scales factorially in n. Recently a number of approaches have come to the machine learning arena to address this kind of data. Most of these approaches are based on the building of a probability distribution over the space of rankings. However, given the complexity of storing, learning and making inference on this kind of models, different simplifying assumptions have been considered: the use of parametric models, models based on low-order statistics, models based on kernels and the definition and use of notions of independence and conditional independence in the space of permutations. In this tutorial we will review the literature on the topic, explaining the different approaches in detail that have emerged in the literature, putting them in relation with other non-probabilistic ranking models and giving a collection of open problems in the area. In addition we will present the most relevant applications in the field as well as the most common benchmark datasets and software.

18:30PM – 21:00PM (Reception Venue, SMU)
RECEPTION
Workshop Award Presentation
Social Events
Other Events

NOVEMBER 4 (SUNDAY)

08:30AM – 12:30PM (Seminar Room 2-2, School of Accountancy, SMU)

WORKSHOP 1

The First International Workshop on Learning with Weak Supervision (LAWS'12)

Chairs: Min-Ling Zhang (Southeast University), **Dell Zhang** (University of London), **Jing Jiang** (SMU)

The First International Workshop on Learning with Weak Supervision (LAWS'12) will be held in conjunction with ACML12 on 4th November 2012. Topics of interest at the workshop include semi-supervised learning, PU learning, multi-instance learning, constrained clustering, multi-label learning, partial label learning (learning from candidate labeling sets), multi-instance multi-label learning (MIML), domain adaptation, transfer learning and multi-task learning.

For more information on the workshop including submission deadlines, please see

http://cse.seu.edu.cn/conf/LAWS12/

13:30PM – 18:30PM (Seminar Room 2-2, School of Accountancy, SMU)

WORKSHOP 2

International Workshop on Fraud Detection in Mobile Advertising (FDMA'12)

Chairs: Richard J. Oentaryo (SMU), **Ee-Peng Lim** (SMU), **Feida Zhu** (SMU), **David Lo** (SMU), **Kok Fung Lai** (BuzzCity Pte Ltd)

Online advertising flourishes as the ideal choice for both small and large businesses to target their marketing campaigns to the appropriate customers on the fly. An advertiser provides an advertising commissioner with its advertisements, plans a budget, and sets a commission for each customer action. The content publishers, on the other hand, make a contract with the commissioner to display advertisements on their websites. However, since publishers earn revenue based on impressions and clicks they drive to advertisers, there is an incentive for dishonest publishers to inflate the number of impressions/clicks their sites generate-a phenomenon known as click fraud. Click fraud hinders the reliability of online advertising system, and the market for online advertising will eventually contract in a long-term. Moreover, it may lead to expensive litigations from unsatisfied advertisers and bad reputation for the commissioner. It is important for the commissioner to proactively prevent click fraud so as to convince their advertisers the fairness of their accounting practices. Accordingly, a reliable click fraud detection system is needed to help identify dishonest publishers and maintain the commissioner's credibility.

The goal of this competition is to build effective data-driven models for proactive detection of fraudulent publishers. A workshop for the competition results will be held at the Asian Conference on Machine Learning (ACML) 2012. For more information on the competition and workshop including submission deadlines, please see

http://palanteer.sis.smu.edu.sg/fdma2012/

NOVEMBER 5 (MONDAY), 08:30AM – 12:00PM

08:30AM – 09:00AM (Auditorium, School of Accountancy, SMU)
OPENING SESSION
Welcome from the Conference Chairs
Session Chair: Wee Sun Lee and Zhi-Hua Zhou

09:00AM – 10:10AM (Auditorium, School of Accountancy, SMU)
KEYNOTE TALK 1
Title: Behavior Imaging and the Study of Autism
Speaker: James Rehg (Georgia Tech)
Session Chair: Steven C.H. Hoi
In this talk I will describe current research efforts in Behavior Imaging, a new research field which encompasses the measurement, modeling, analysis, and visualization of social and communicative behaviors from multi-modal sensor data. Beginning in infancy, individuals acquire the social and communicative skills which are vital for a healthy and productive life, through face-to-face interactions with caregivers and peers. However, children with developmental delays face great challenges in acquiring these skills, resulting in substantial lifetime risks. Autism, for example, affects 1 in 88 children in the U.S. and can lead to substantial impairments, resulting in a lifetime cost of care of $3.2M per person. The goal of our research in Behavior Imaging is to develop computational methods that can support the fine-grained and large-scale measurement and analysis of social behaviors, with the potential to positively impact diagnosis and treatment. I will present an overview of our research efforts in Behavior Imaging, with a particular emphasis on the use of machine learning methods to extract behavior measurements from weakly-annotated video data. Specifically, I will describe a new approach to video analysis based on the concept of temporal causality, which leverages a novel representation of video events as multiple point processes. Our method provides a new bottom-up approach to video segmentation based on the temporal structure of video events. I will present results for retrieving and categorizing social interactions in collections of real-world video footage. I will also highlight our recent efforts in recognizing activities in video which is acquired by a wearable camera (also known as egocentric vision). This is joint work with Alireza Fathi, Yin Li, Karthir Prabhakar, and Sangmin Oh.

NOVEMBER 5 (MONDAY), 08:30AM – 12:00PM

10:30AM – 12:00PM (Auditorium, School of Accountancy, SMU)
ORAL PRESENTATION SESSION 1
Topic: Unsupervised Learning and Deep Learning
Session Chair: Wray Buntine
Recovering Networks from Distance Data (25 min) *Best Student Paper Award*
Sandhya Prabhakaran (University of Basel), Karin Metzner (University Hospital Zurich), Alexander Böhm (LOEWE-Zentrum für Synthetische Mikrobiologie), Volker Roth (University of Basel)
Statistical Models for Exploring Individual Email Communication Behavior (20 min)
Nicholas Navaroli (University of California, Irvine), Christopher DuBois (University of California, Irvine), Padhraic Smyth (University of California, Irvine)
Multiresolution Mixture Modelling using Merging of Mixture Components (10 min)
Prem Adhikari (Aalto University), Jaakko Hollmén (Aalto Univeristy)
Cumulative Restricted Boltzmann Machines for Ordinal Matrix Data Analysis (10 min)
Truyen Tran (Curtin University), Dinh Phung (Deakin University), Svetha Venkatesh (Deakin University)
Learning Latent Variable Models by Pairwise Cluster Comparison (10 min)
Nuaman Asbeh (Ben-Gurion University of the Negev), Boaz Lerner (Ben-Gurion University of the Negev)
A Note on Metric Properties for Some Divergence Measures: The Gaussian Case (10 min)
Karim Abou-Moustafa (McGill University), Frank Ferrie (McGill University)

NOVEMBER 5 (MONDAY), 12:00PM – 14:00PM

12:00PM – 14:00PM (Library, SMU)	
POSTER PRESENTATION SESSION 1	
Session Chair: Hai Leong Chieu	
Poster ID	**Paper Titles and Authors**
P1	**Recovering Networks from Distance Data** *(Best Student Paper Award)*
	Sandhya Prabhakaran (University of Basel), Karin Metzner (University Hospital Zurich), Alexander Böhm (LOEWE-Zentrum für Synthetische Mikrobiologie), Volker Roth (University of Basel)
P2	**Statistical Models for Exploring Individual Email Communication Behavior**
	Nicholas Navaroli (University of California, Irvine), Christopher DuBois (University of California, Irvine), Padhraic Smyth (University of California, Irvine)
P3	**Multiresolution Mixture Modelling using Merging of Mixture Components**
	Prem Adhikari (Aalto University), Jaakko Hollmén (Aalto Univeristy)
P4	**Cumulative Restricted Boltzmann Machines for Ordinal Matrix Data Analysis**
	Truyen Tran (Curtin University), Dinh Phung (Deakin University), Svetha Venkatesh (Deakin University)
P5	**Learning Latent Variable Models by Pairwise Cluster Comparison**
	Nuaman Asbeh (Ben-Gurion University of the Negev), Boaz Lerner (Ben-Gurion University of the Negev)
P6	**A Note on Metric Properties for Some Divergence Measures: The Gaussian Case**
	Karim Abou-Moustafa (McGill University), Frank Ferrie (McGill University)
P7	**Topographic Analysis of Correlated Components**
	Hiroaki Sasaki (University of Electro-Communication), Michael Gutmann (University of Helsinki), Hayaru Shouno (University of Electro-Communications), Aapo Hyvarinen (Helsinki Institute for Information Technology)
P8	**Sparse Additive Matrix Factorization for Robust PCA and Its Generalization**
	Shinichi Nakajima (Nikon Corporation), Masashi Sugiyama (Tokyo Institute of Technology), S. Derin Babacan (UIUC)
P9	**Spatial Locality-Aware Sparse Coding and Dictionary Learning**
	Jiang Wang (Northwestern University), Junsong Yuan (Nanyang Technological University), Zhuoyuan Chen (Northwestern University), Ying Wu (Northwestern University)
P10	**Local Kernel Density Ratio-Based Feature Selection for Outlier Detection**
	Fatemeh Azmandian (Northeastern University), Jennifer Dy (Northeastern University), Javed Aslam (Northeastern University), David Kaeli (Northeastern University)
P11	**Supervised Dimension Reduction with Topic Models**
	Khoat Than (JAIST), Tu Bao Ho (JAIST), Duy Khuong Nguyen (JAIST), Ngoc Khanh Pham (JAIST)
P12	**Key Instance Detection in Multi-Instance Learning**
	Guoqing Liu (Nanyang Technological Univeristy), Jianxin Wu (Nanyang Technological Univeristy), Zhi-Hua Zhou (Nanjing University)

NOVEMBER 5 (MONDAY), 12:00PM – 14:00PM

12:00PM – 14:00PM (Library, SMU)	
POSTER PRESENTATION SESSION 1 (cont')	
Session Chair: Hai Leong Chieu	
Poster ID	**Paper Titles and Authors**
P13	**A Convex-Concave Relaxation Procedure Based Subgraph Matching Algorithm**
	Zhiyong Liu (Chinese Academy of Sciences), Hong Qiao (Chinese Academy of Sciences)
P14	**Learning and Model-Checking Networks of I/O Automata**
	Hua Mao (Aalborg University), Manfred Jaeger (Aalborg University)
P15	**Learning Temporal Association Rules on Symbolic Time Sequences**
	Mathieu Guillame-Bert (INRIA), James Crowley (INRIA)
P16	**Improved Sequence Classification Using Adaptive Segmental Sequence Alignment**
	Shahriar Shariat (Rutgers), Vladimir Pavlovic (Rutgers)
P17	**A Coupled Indian Buffet Process Model for Collaborative Filtering**
	Sotirios Chatzis (Cyprus University Technology)
P18	**Two-way Parallel Class Expression Learning**
	An C. Tran (Massey University), Jens Dietrich (Massey University), Hans W. Guesgen (Massey University), Stephen Marsland (Massey University)
P19	**Data Thresholding for Large-scale Sparse Linear Classification**
	Gia Vinh Anh Pham, Laurent El Ghaoui (UC Berkeley)
P20	**Exploiting Rank-Learning Models to Predict the Diffusion of Preferences on Social Networks**
	Chin-Hua Tsai, Hung-Yi Lo, Shou-De Lin (National Taiwan University)
P21	**Efficient AUC Maximization by Approximate Reduction of Ranking SVMs**
	Daiki Suehiro, Daiki Suehiro, Daiki Suehiro (Kyushu University)

NOVEMBER 5 (MONDAY), 14:00PM – 17:00PM

14:00PM – 15:10PM (Auditorium, School of Accountancy, SMU)
KEYNOTE TALK 2
Title: Convex Methods for Representation Learning
Speaker: Dale Schuurmans (University of Alberta)
Session Chair: Zhi-Hua Zhou
Automated feature discovery is a fundamental problem in data analysis. Although classical feature learning methods fail to guarantee optimal solutions in general, convex reformulations have been developed for a number of such problems. Most of these reformulations are based on one of two key strategies: approximating pairwise representations or exploiting induced matrix norms. Despite their use of relaxation, convex reformulations can demonstrate significant improvements in solution quality by eliminating local minima. I will discuss several convex reformulations for representation learning problems, including clustering, subspace learning, multi-view learning, and hidden-layer network training---demonstrating how feature discovery can co-occur with parameter optimization while admitting globally optimal solutions.

15:30PM – 17:00PM (Auditorium, School of Accountancy, SMU)
ORAL PRESENTATION SESSION 2
Topic: Feature Selection and Dimension Reduction
Session Chair: Dell Zhang
Topographic Analysis of Correlated Components (20 min)
Hiroaki Sasaki (University of Electro-Communication), Michael Gutmann (University of Helsinki), Hayaru Shouno (University of Electro-Communications), Aapo Hyvarinen (Helsinki Institute for Information Technology)
Sparse Additive Matrix Factorization for Robust PCA and Its Generalization (20 min)
Shinichi Nakajima (Nikon Corporation), Masashi Sugiyama (Tokyo Institute of Technology), S. Derin Babacan (UIUC)
Spatial Locality-Aware Sparse Coding and Dictionary Learning (10 min)
Jiang Wang (Northwestern University), Junsong Yuan (Nanyang Technological University), Zhuoyuan Chen (Northwestern University), Ying Wu (Northwestern University)
Local Kernel Density Ratio-Based Feature Selection for Outlier Detection (10 min)
Fatemeh Azmandian (Northeastern University), Jennifer Dy (Northeastern University), Javed Aslam (Northeastern University), David Kaeli (Northeastern University)
Supervised Dimension Reduction with Topic Models (10 min)
Khoat Than (JAIST), Tu Bao Ho (JAIST), Duy Khuong Nguyen (JAIST), Ngoc Khanh Pham (JAIST)
Key Instance Detection in Multi-Instance Learning (10 min)
Guoqing Liu (Nanyang Technological Univeristy), Jianxin Wu (Nanyang Technological Univeristy), Zhi-Hua Zhou (Nanjing University)

NOVEMBER 5 (MONDAY), 17:20PM – 21:00PM

17:20PM – 18:30PM (Auditorium, School of Accountancy, SMU)
ORAL PRESENTATION SESSION 3
Topic: Learning in Graphs, Networks, and Structures
Session Chair: Bernhard Pfahringer
A Convex-Concave Relaxation Procedure Based Subgraph Matching Algorithm (20 min)
Zhiyong Liu (Chinese Academy of Sciences), Hong Qiao (Chinese Academy of Sciences)
Learning and Model-Checking Networks of I/O Automata (10 min)
Hua Mao (Aalborg University), Manfred Jaeger (Aalborg University)
Learning Temporal Association Rules on Symbolic Time Sequences (10 min)
Mathieu Guillame-Bert (INRIA), James Crowley (INRIA)
Improved Sequence Classification Using Adaptive Segmental Sequence Alignment (10 min)
Shahriar Shariat (Rutgers), Vladimir Pavlovic (Rutgers)
A Coupled Indian Buffet Process Model for Collaborative Filtering (10 min)
Sotirios Chatzis (Cyprus University Technology)
Two-way Parallel Class Expression Learning (10 min)
An C. Tran (Massey University), Jens Dietrich (Massey University), Hans W. Guesgen (Massey University), Stephen Marsland (Massey University)

18:30PM – 21:00 PM (Banquet Venue)
BANQUET
ACML 2012 Best Paper Awards
ACML 2013 Presentation
Other Social Events

NOVEMBER 6 (TUESDAY), 08:30AM – 12:00PM

08:30AM – 09:00AM (Auditorium, School of Accountancy, SMU)
BEST PAPER SESSION
Title: Variational Bayesian Matching
Author: Arto Klami (Aalto University)
Session Chair: Wray Buntine and Steven C.H. Hoi

09:00AM – 10:10AM (Auditorium, School of Accountancy, SMU)
KEYNOTE TALK 3
Title: Multiclass Losses and Multidistribution divergences
Speaker: Bob Williamson (Australian National University and NICTA)
Session Chair: Wee Sun Lee
Binary prediction problems (and their associated loss functions) are perhaps the simplest machine learning problems and have been extensively studied. Similarly, divergence measures between two probability distributions are well understood, for example the classical Csiszar f-divergences. There is a natural bridge between binary proper losses and f-divergences via the Bayes risk of a binary learning problem induced by the loss. Multiclass prediction problems and multiclass loss functions are less well understood. It is not even clear (at first) what the "divergence" between k distributions even means when k>2 In this talk I will show how the binary "bridge" extends to the multiclass case and allows simple proofs of the properties of multidistribution f-divergences which are analogous to those satisfied by the classical f-divergences. I will also outline the theory of composite multiclass losses, which are the composition of a proper loss with a link function, including a characterisation of when they are convex. (Joint work with Dario Garcia-Garcia, Mark Reid, and Elodie Vernet)

NOVEMBER 6 (TUESDAY), 08:30AM – 12:00PM

10:30AM – 12:00PM (Auditorium, School of Accountancy, SMU)
ORAL PRESENTATION SESSION 4
Topic: Classification and Ranking
Session Chair: Tu Bao Ho
Multi-Stage Classifier Design (20 min)
Kirill Trapeznikov (Boston University), Venkatesh Saligrama (Boston University), David Castanon (Boston University)
QBoost: Large Scale Classifier Training with Adiabatic Quantum Optimization (20 min)
Hartmut Neven (Google), Vasil Denchev (Purdue University), Geordie Rose (D-Wave Systems, Inc.), William Macready (D-Wave Systems, Inc.)
Practical Large Scale Classification with Additive Kernels (10 min)
Hao Yang (Nanyang Technological Univeristy), Jianxin Wu (Nanyang Technological Univeristy)
Max Margin Ratio Machine (10 min)
Suicheng Gu (Temple University), Yuhong Guo (Temple University)
A Ranking-based KNN Approach for Multi-Label Classification (10 min)
Tsung-Hsien Chiang (National Taiwan University), Hung-Yi Lo (Academia Sinica), Shou-de Lin (National Taiwan University)
Learning From Ordered Sets and Applications in Collaborative Ranking (10 min)
Truyen Tran (Curtin University), Dinh Phung (Deakin University), Svetha Venkatesh (Deakin University)

NOVEMBER 6 (TUESDAY), 12:00PM – 14:00PM

12:00PM – 14:00PM (Library, SMU)	
POSTER PRESENTATION SESSION 2	
Session Chair: Hai Leong Chieu	
Poster ID	**Paper Titles and Authors**
P22	**Variational Bayesian Matching** *(Best Paper Award)*
	Arto Klami (Aalto University)
P23	**Multi-Stage Classifier Design**
	Kirill Trapeznikov (Boston University), Venkatesh Saligrama (Boston University), David Castanon (Boston University)
P24	**QBoost: Large Scale Classifier Training with Adiabatic Quantum Optimization**
	Hartmut Neven (Google), Vasil Denchev (Purdue University), Geordie Rose (D-Wave Systems, Inc.), William Macready (D-Wave Systems, Inc.)
P25	**Practical Large Scale Classification with Additive Kernels**
	Hao Yang (Nanyang Technological Univeristy), Jianxin Wu (Nanyang Technological Univeristy)
P26	**Max Margin Ratio Machine**
	Suicheng Gu (Temple University), Yuhong Guo (Temple University)
P27	**A Ranking-based KNN Approach for Multi-Label Classification**
	Tsung-Hsien Chiang (National Taiwan University), Hung-Yi Lo (Academia Sinica), Shou-de Lin (National Taiwan University)
P28	**Learning From Ordered Sets and Applications in Collaborative Ranking**
	Truyen Tran (Curtin University), Dinh Phung (Deakin University), Svetha Venkatesh (Deakin University)
P29	**On Using Nearly-Independent Feature Families for High Precision and Confidence**
	Omid Madani (Google), Manfred Georg (Google), David Ross (Google)
P30	**More Is Better: Large Scale Partially-supervised Sentiment Classification**
	Yoav Haimovitch (Technion), Koby Crammer (Technion), Shie Mannor (Technion)
P31	**Active Learning with Hinted Support Vector Machine**
	Chun-Liang Li (National Taiwan University), Chun-Sung Ferng (National Taiwan University), Hsuan-Tien Lin (National Taiwan University)
P32	**Multi-view Positive and Unlabeled Learning**
	Joey Tianyi Zhou (Nanyang Technological University), Sinno Jialin Pan (A*STAR), Qi Mao (Nanyang Technological University), Ivor Wai-Hung TSANG (Nanyang Technological University)
P33	**Frustratingly Simplified Deployment in WLAN Localization by Learning from Route Annotation**
	Ryoma Kawajiri (University of Tokyo), Masamichi Shimosaka (University of Tokyo), Rui Fukui (University of Tokyo), Tomomasa Sato (University of Tokyo)

NOVEMBER 6 (TUESDAY), 12:00PM – 14:00PM

12:00PM – 14:00PM (Library, SMU)	
POSTER PRESENTATION SESSION 2 (cont')	
Session Chair: Hai Leong Chieu	
Poster ID	**Paper Titles and Authors**
P34	**Conditional validity of inductive conformal predictors**
	Vladimir Vovk (Royal Holloway)
P35	**Multi-objective Monte-Carlo Tree Search**
	Weijia Wang (LRI), Michele Sebag (CNRS)
P36	**A Stochastic Bandit Algorithm for Scratch Games**
	Raphaël Féraud (Orange Labs), Tanguy Urvoy (Orange Labs)
P37	**AIC and BIC based approaches for SVM parameter value estimation with RBF kernels**
	Sergey Demyanov (University of Melbourne), James Bailey (University of Melbourne), Ramamohanarao Kotagiri (University of Melbourne), Christopher Leckie (University of Melbourne)
P38	**Online Rank Aggregation**
	Shota Yasutake (Kyushu University), Kohei Hatano (Kyusyu University), Eiji Takimoto (Kyushu University), Masayuki Takeda (Kyushu University)
P39	**Online Learning of a Dirichlet Process Mixture of Generalized Dirichlet Distributions for Simultaneous Clustering and Localized Feature Selection**
	Wentao Fan (Concordia University), Nizar Bouguila (Concordia University)
P40	**Distribution Changes Based Financial Time Series Forecasting using Data Mining and Machine Learning Methods**
	Goce Ristanoski, James Bailey (University of Melbourne)
P41	**Learning Interpretable Models from Distributed Data**
	Artur Andrzejak, Felix Langner, Silvestre Zabala, See-Kiong Ng (Ruprecht-Karls-University of Heidelberg)
P42	**A generalized dependent normalized measure framework for Bayesian nonparametric modelling**
	Changyou Chen (Australian National University)

NOVEMBER 6 (TUESDAY), 14:00PM – 17:00PM

14:00PM – 15:10PM (Auditorium, School of Accountancy, SMU)
ORAL PRESENTATION SESSION 5
Topic: Supervised and Semi-supervised Learning
Session Chair: Xiaoli Li
On Using Nearly-Independent Feature Families for High Precision and Confidence (20 min)
Omid Madani (Google), Manfred Georg (Google), David Ross (Google)
More Is Better: Large Scale Partially-supervised Sentiment Classification (10 min)
Yoav Haimovitch (Technion), Koby Crammer (Technion), Shie Mannor (Technion)
Active Learning with Hinted Support Vector Machine (10 min)
Chun-Liang Li (National Taiwan University), Chun-Sung Ferng (National Taiwan University), Hsuan-Tien Lin (National Taiwan University)
Multi-view Positive and Unlabeled Learning (10 min)
Joey Tianyi Zhou (Nanyang Technological University), Sinno Jialin Pan (A*STAR), Qi Mao (Nanyang Technological University), Ivor Wai-Hung TSANG (Nanyang Technological University)
Frustratingly Simplified Deployment in WLAN Localization by Learning from Route Annotation (10 min)
Ryoma Kawajiri (University of Tokyo), Masamichi Shimosaka (University of Tokyo), Rui Fukui (University of Tokyo), Tomomasa Sato (University of Tokyo)

15:30PM – 17:00PM (Auditorium, School of Accountancy, SMU)
ORAL PRESENTATION SESSION 6
Topic: Learning Theory, Reinforcement and Online Learning
Session Chair: Bob Williamson
Conditional validity of inductive conformal predictors (20 min)
Vladimir Vovk (Royal Holloway)
Multi-objective Monte-Carlo Tree Search (20 min)
Weijia Wang (LRI), Michele Sebag (CNRS)
A Stochastic Bandit Algorithm for Scratch Games (20 min)
Raphaël Féraud (Orange Labs), Tanguy Urvoy (Orange Labs)
AIC and BIC based approaches for SVM parameter value estimation with RBF kernels (10min)
Sergey Demyanov (University of Melbourne), James Bailey (University of Melbourne), Ramamohanarao Kotagiri (University of Melbourne), Christopher Leckie (University of Melbourne)
Online Rank Aggregation (10 min)
Shota Yasutake (Kyushu University), Kohei Hatano (Kyusyu University), Eiji Takimoto (Kyushu University), Masayuki Takeda (Kyushu University)
Online Learning of a Dirichlet Process Mixture of Generalized Dirichlet Distributions for Simultaneous Clustering and Localized Feature Selection (10 min)
Wentao Fan (Concordia University), Nizar Bouguila (Concordia University)

ABSTRACTS OF ACCEPTED PAPERS (NOVEMBER 5, MONDAY)

Recovering Networks from Distance Data *(Best Student Paper Award)*

Sandhya Prabhakaran, Karin Metzner, Alexander Böhm, Volker Roth

A fully probabilistic approach to reconstructing Gaussian graphical models from distance data is presented. The main idea is to extend the usual central Wishart model in traditional methods to using a likelihood depending only on pairwise distances, thus being independent of geometric assumptions about the underlying Euclidean space. This extension has two advantages: the model becomes invariant against potential bias terms in the measurements, and can be used in situations which on input use a kernel- or distance matrix, without requiring direct access to the underlying vectors. The latter aspect opens up a huge new application field for Gaussian graphical models, as network reconstruction is now possible from any Mercer kernel, be it on graphs, strings, probabilities or more complex objects. We combine this likelihood with a suitable prior to enable Bayesian network inference. We present an efficient MCMC sampler for this model and discuss the estimation of module networks. Experiments depict the high quality and usefulness of the inferred networks.

Statistical Models for Exploring Individual Email Communication Behavior

Nicholas Navaroli, Christopher DuBois, Padhraic Smyth

As digital communication devices play an increasingly prominent role in our daily lives, the ability to analyze and understand our communication patterns becomes more important. In this paper, we investigate a latent variable modeling approach for extracting information from individual email histories, focusing in particular on understanding how an individual communicates over time with recipients in their social network. The proposed model consists of latent groups of recipients, each of which is associated with a piecewise-constant Poisson rate over time. Inference of group memberships, temporal changepoints, and rate parameters is carried out via Markov Chain Monte Carlo (MCMC) methods. We illustrate the utility of the model by applying it to both simulated and real-world email data sets.

Multiresolution Mixture Modelling using Merging of Mixture Components

Prem Adhikari, Jaakko Hollmén

Observing natural phenomena at several levels of detail results in multiresolution data. Extending models and algorithms to cope with multiresolution data is a prerequisite for wide-spread exploitation of the data represented in the multiple resolutions. Mixture models are widely used probabilistic models, however, the mixture models in their standard form can be used to analyze the data represented in a single resolution. In this paper, we propose a multiresolution mixture model based on merging of the mixture components across models represented in different resolutions. Result of such an analysis scenario is to have multiple mixture models, one mixture model for each resolution of data. Our proposed solution is based on the idea on the interaction between mixture models. More specifically, we repeatedly merge component distributions of mixture models across different resolutions. We experiment our proposed algorithm on the two real-world chromosomal aberration datasets represented in two different resolutions. Results show an improvement on the compared multiresolution settings.

ABSTRACTS OF ACCEPTED PAPERS (NOVEMBER 5, MONDAY)

Cumulative Restricted Boltzmann Machines for Ordinal Matrix Data Analysis

Truyen Tran, Dinh Phung, Svetha Venkatesh

Ordinal data is omnipresent in almost all multiuser-generated feedback - questionnaires, preferences etc. This paper investigates modelling of ordinal data with Gaussian restricted Boltzmann machines (RBMs). In particular, we present the model architecture, learning and inference procedures for both vector-variate and matrix-variate ordinal data. We show that our model is able to capture latent opinion profile of citizens around the world, and is competitive against state-of-art collaborative filtering techniques on large-scale public datasets. The model thus has the potential to extend application of RBMs to diverse domains such as recommendation systems, product reviews and expert assessments.

Learning Latent Variable Models by Pairwise Cluster Comparison

Nuaman Asbeh, Boaz Lerner

Identification of latent variables that govern a problem and the relationships among them given measurements in the observed world are important for causal discovery. This identification can be made by analyzing constraints imposed by the latents in the measurements. We introduce the concept of pairwise cluster comparison PCC to identify causal relationships from clusters and a two-stage algorithm, called LPCC, that learns a latent variable model (LVM) using PCC. First, LPCC learns the exogenous and the collider latents, as well as their observed descendants, by utilizing pairwise comparisons between clusters in the measurement space that may explain latent causes. Second, LPCC learns the non-collider endogenous latents and their children by splitting these latents from their previously learned latent ancestors. LPCC is not limited to linear or latent-tree models and does not make assumptions about the distribution. Using simulated and real-world datasets, we show that LPCC improves accuracy with the sample size, can learn large LVMs, and is accurate in learning compared to state-of-the-art algorithms.

A Note on Metric Properties for Some Divergence Measures: The Gaussian Case

Karim Abou-Moustafa, Frank Ferrie

Multivariate Gaussian densities are pervasive in pattern recognition and machine learning. A central operation that appears in most of these areas is to measure the difference between two multivariate Gaussians. Unfortunately, traditional measures based on the Kullback-Leibler (KL) divergence and the Bhattacharyya distance do not satisfy all metric axioms necessary for many algorithms. In this paper we propose a modification for the KL divergence and the Bhattacharyya distance, for multivariate Gaussian densities, that transforms the two measures into distance metrics. Next, we show how these metric axioms impact the unfolding process of manifold learning algorithms. Finally, we illustrate the efficacy of the proposed metrics on two different manifold learning algorithms when used for motion clustering in video data. Our results show that, in this particular application, the new proposed metrics lead to boosts in performance (at least 7%) when compared to other divergence measures.

ABSTRACTS OF ACCEPTED PAPERS (NOVEMBER 5, MONDAY)

Topographic Analysis of Correlated Components

Hiroaki Sasaki, Michael Gutmann, Hayaru Shouno, Aapo Hyvärinen

Independent component analysis (ICA) is a method to estimate components which are as statistically independent as possible. However, in many practical applications, the estimated components are not independent. Recent variants of ICA have made use of such residual dependencies to estimate an ordering (topography) of the components. Like in ICA, the components in those variants are assumed to be uncorrelated, which might be a rather strict condition. In this paper, we address this shortcoming. We propose a generative model for the source where the components can have linear and higher order correlations, which generalizes models in use so far. Based on the model, we derive a method to estimate topographic representations. In numerical experiments on artificial data, the new method is shown to be more widely applicable than previously proposed extensions of ICA. We learn topographic representations for two kinds of real data sets: for outputs of simulated complex cells in the primary visual cortex and for text data.

Sparse Additive Matrix Factorization for Robust PCA and Its Generalization

Shinichi Nakajima, Masashi Sugiyama, S. Derin Babacan

Principal component analysis (PCA) can be regarded as approximating a data matrix with a low-rank one by imposing sparsity on its singular values, and its robust variant further captures sparse noise. In this paper, we extend such sparse matrix learning methods, and propose a novel unified framework called sparse additive matrix factorization (SAMF). SAMF systematically induces various types of sparsity by the so-called model-induced regularization in the Bayesian framework. We propose an iterative algorithm called the mean update (MU) for the variational Bayesian approximation to SAMF, which gives the global optimal solution for a large subset of parameters in each step. We demonstrate the usefulness of our method on artificial data and the foreground/background video separation.

Spatial Locality-Aware Sparse Coding and Dictionary Learning

Jiang Wang, Junsong Yuan, Zhuoyuan Chen, Ying Wu

Nonlinear encoding of SIFT features has recently shown good promise in image classification. This scheme is able to reduce the training complexity of the traditional bag-of-feature approaches while achieving better performance. As a result, it is suitable for large-scale image classification applications. However, existing nonlinear encoding methods do not explicitly consider the spatial relationship when encoding the local features, but merely leaving the spatial information used at a later stage, e.g. through the spatial pyramid matching, is largely inadequate. In this paper, we propose a joint sparse coding and dictionary learning scheme that take the spatial information into consideration in encoding. Our experiments on synthetic data and benchmark data demonstrate that the proposed scheme can learn a better dictionary and achieve higher classification accuracy.

ABSTRACTS OF ACCEPTED PAPERS (NOVEMBER 5, MONDAY)

Local Kernel Density Ratio-Based Feature Selection for Outlier Detection

Fatemeh Azmandian, Jennifer Dy, Javed Aslam, David Kaeli

Selecting features is an important step of any machine learning task, though most of the focus has been to choose features relevant for classification and regression. In this work, we present a novel non-parametric evaluation criterion for filter-based feature selection which enhances outlier detection. Our proposed method seeks the subset of features that represents the inherent characteristics of the normal dataset while forcing outliers to stand out, making them more easily distinguished by outlier detection algorithms. Experimental results on real datasets show the advantage of this feature selection algorithm compared to popular and state-of-the-art methods. We also show that the proposed algorithm is able to overcome the small sample space problem and perform well on highly imbalanced datasets.

Supervised Dimension Reduction with Topic Models

Khoat Than, Tu Bao Ho, Duy Khuong Nguyen, Ngoc Khanh Pham

We consider supervised dimension reduction (SDR) for problems with discrete variables. Existing methods are computationally expensive, and often do not take the local structure of data into consideration when searching for a low-dimensional space. In this paper, we propose a novel framework for SDR which is (1) general and flexible so that it can be easily adapted to various unsupervised topic models, (2) able to inherit scalability of unsupervised topic models, and (3) can exploit well label information and local structure of data when searching for a new space. Extensive experiments with adaptations to three models demonstrate that our framework can yield scalable and qualitative methods for SDR. One of those adaptations can perform better than the state-of-the-art method for SDR while enjoying significantly faster speed.

Key Instance Detection in Multi-Instance Learning

Guoqing Liu, Jianxin Wu, Zhi-Hua Zhou

The goal of traditional multi-instance learning (MIL) is to predict the labels of the bags, whereas in many real applications, it is desirable to get the instance labels, especially the labels of key instances that trigger the bag labels, in addition to getting bag labels. Such a problem has been largely unexplored before. In this paper, we formulate the Key Instance Detection (KID) problem, and propose a voting framework (VF) solution to KID. The key of VF is to exploit the relationship among instances, represented by a citer kNN graph. This graph is different from commonly used nearest neighbor graphs, but is suitable for KID. Experiments validate the effectiveness of VF for KID. Additionally, VF also outperforms state-of-the-art MIL approaches on the performance of bag label prediction.

ABSTRACTS OF ACCEPTED PAPERS (NOVEMBER 5, MONDAY)

A Convex-Concave Relaxation Procedure Based Subgraph Matching Algorithm

Zhiyong Liu, Hong Qiao

Based on the convex-concave relaxation procedure (CCRP), the (extended) path following algorithms were recently proposed to approximately solve the equal-sized graph matching problem, and exhibited a state-of-the-art performance (Zaslavskiy et al., 2009; Liu et al., 2012). However, they cannot be used for subgraph matching since either their convex or concave relaxation becomes no longer applicable. In this paper we extend the CCRP to tackle subgraph matching, by proposing a convex as well as a concave relaxation of the problem. Since in the context of CCRP, the convex relaxation can be viewed as an initialization of a concave programming, we introduce two other initializations for comparison. Meanwhile, the graduated assignment algorithm is also introduced in the experimental comparisons, which witness the validity of the proposed algorithm.

Learning and Model-Checking Networks of I/O Automata

Hua Mao, Manfred Jaeger

We introduce a new statistical relational learning (SRL) approach in which models for structured data, especially network data, are constructed as networks of communicating finite probabilistic automata. Leveraging existing automata learning methods from the area of grammatical inference, we can learn generic models for network entities in the form of automata templates. As is characteristic for SRL techniques, the abstraction level afforded by learning generic templates enables one to apply the learned model to new domains. A main benefit of learning models based on finite automata lies in the fact that one can analyse the resulting models using formal model-checking techniques, which adds a dimension of model analysis not usually available for traditional SRL modeling frameworks.

Learning Temporal Association Rules on Symbolic Time Sequences

Mathieu Guillame-Bert, James Crowley

We introduce a temporal pattern model called Temporal Interval Tree Association Rules (Tita rules or Titar). This pattern model can express both uncertainty and temporal inaccuracy of temporal events. Among other things, Tita rules can express the usual time point operators, synchronicity, order, and chaining, as well as temporal negation and disjunctive temporal constraints. Using this representation, we present the Titar learner algorithm that can be used to extract Tita rules from large datasets expressed as Symbolic Time Sequences. The selection of temporal constraints (or time-frames) is at the core of the temporal learning. Our learning algorithm is based on two novel approaches for this problem. This first one is designed to select temporal constraints for the head of temporal association rules. The second selects temporal constraints for the body of such rules. We discuss the evaluation of probabilistic temporal association rules, evaluate our technique with two experiments, introduce a metric to evaluate sets of temporal rules, compare the results with two other approaches and discuss the results.

ABSTRACTS OF ACCEPTED PAPERS (NOVEMBER 5, MONDAY)

Improved Sequence Classification Using Adaptive Segmental Sequence Alignment

Shahriar Shariat, Vladimir Pavlovic

Traditional pairwise sequence alignment is based on matching individual samples from two sequences, under time monotonicity constraints. However, in some instances matching two segments of points may be preferred and can result in increased noise robustness. This paper presents an approach to segmental sequence alignment based on adaptive pairwise segmentation. We introduce a distance metric between segments based on average pairwise distances, which addresses deficiencies of prior approaches. We then present a modified pair-HMM that incorporates the proposed distance metric and use it to devise an efficient algorithm to jointly segment and align the two sequences. Our results demonstrate that this new measure of sequence similarity can lead to improved classification performance, while being resilient to noise, on a variety of problems, from EEG to motion sequence classification.

A Coupled Indian Buffet Process Model for Collaborative Filtering

Sotirios Chatzis

The dramatic rates new digital content becomes available has brought collaborative filtering systems in the epicenter of computer science research in the last decade. In this paper, we propose a novel methodology for rating prediction utilizing concepts from the field of Bayesian nonparametrics. The basic concept that underlies our approach is that each user rates a presented item based on the latent genres of the item and the latent interests of the user. Each item may belong to more than one genre, and each user may belong to more than one latent interest class. The number of existing latent genres and interests are not known beforehand, but should be inferred in a data-driven fashion. We devise a novel hierarchical factor analysis model to formulate our approach under these assumptions. We impose suitable priors over the allocation of items into genres, and users into interests; specifically, we utilize a novel scheme which comprises two coupled Indian buffet process priors that allow the number of latent classes (genres/interests) to be automatically inferred. We experiment on a large set of real ratings data, and show that our approach outperforms four common baselines, including two very competitive state-of-the-art approaches.

Two-way Parallel Class Expression Learning

An C. Tran, Jens Dietrich, Hans W. Guesgen, Stephen Marsland

In machine learning, we often encounter datasets that can be described using simple rules and regular exception patterns describing situations where those rules do not apply. In this paper, we propose a two-way parallel class expression learning algorithm that is suitable for this kind of problem. This is a top-down refinement-based class expression learning algorithm for Description Logic (DL). It is distinguished from similar DL learning algorithms in the way it uses the concepts generated by the refinement operator. In our approach, we unify the computation of concepts describing positive and negative examples, but we maintain them separately, and combine them at the end. By doing so, we can avoid the use of negation in the refinement without any loss of generality. Evaluation shows that our approach can reduce the search space significantly, and therefore the learning time is reduced. Our implementation is based on the DL-Learner framework and we inherit the Parallel Class Expression Learning (ParCEL) algorithm design for parallelisation.

ABSTRACTS OF ACCEPTED PAPERS (NOVEMBER 6, TUESDAY)

Variational Bayesian Matching *(Best Paper Award)*

Arto Klami

Matching of samples refers to the problem of inferring unknown co-occurrence or alignment between observations in two data sets. Given two sets of equally many samples, the task is to find for each sample a representative sample in the other set, without prior knowledge on a distance measure between the sets. Recently a few alternative solutions have been suggested, based on maximization of joint likelihood or various measures of between-data statistical dependency. In this work we present an variational Bayesian solution for the problem, learning a Bayesian canonical correlation analysis model with a permutation parameter for re-ordering the samples in one of the sets. We approximate the posterior over the permutations, and demonstrate that the resulting matching algorithm clearly outperforms all of the earlier solutions.

Multi-Stage Classifier Design

Kirill Trapeznikov, Venkatesh Saligrama, David Castanon

In many classification systems, sensing modalities have different acquisition costs. It is often unnecessary to use every modality to classify a majority of examples. We study a multi-stage system in a prediction time cost reduction setting, where the full data is available for training, but for a test example, measurements in a new modality can be acquired at each stage for an additional cost. We seek decision rules to reduce the average measurement acquisition cost. We formulate an empirical risk minimization problem (ERM) for a multi-stage reject classifier, wherein the stage kclassifier either classifies a sample using only the measurements acquired so far or rejects it to the next stage where more attributes can be acquired for a cost. To solve the ERM problem, we factorize the cost function into classification and rejection decisions. We then transform reject decisions into a binary classification problem. We construct stage-by-stage global surrogate risk, develop an iterative algorithm in the boosting framework and present convergence results. We test our work on synthetic, medical and explosives detection datasets. Our results demonstrate that substantial cost reduction without a significant sacrifice in accuracy is achievable.

QBoost: Large Scale Classifier Training with Adiabatic Quantum Optimization

Hartmut Neven, Vasil Denchev, Geordie Rose, William Macready

We introduce a novel discrete optimization method for training in the context of a boosting framework for large scale binary classifiers. The motivation is to cast the training problem into the format required by existing adiabatic quantum hardware. First we provide theoretical arguments concerning the transformation of an originally continuous optimization problem into one with discrete variables of low bit depth. Next we propose QBoost as an iterative training algorithm in which a subset of weak classifiers is selected by solving a hard optimization problem in each iteration. A strong classifier is incrementally constructed by concatenating the subsets of weak classifiers. We supplement the findings with experiments on one synthetic and two natural data sets and compare against the performance of existing boosting algorithms. Finally, by conducting a quantum Monte Carlo simulation we gather evidence that adiabatic quantum optimization is able to handle the discrete optimization problems generated by QBoost.

ABSTRACTS OF ACCEPTED PAPERS (NOVEMBER 6, TUESDAY)

Practical Large Scale Classification with Additive Kernels

Hao Yang, Jianxin Wu

For classification problems with millions of training examples or dimensions, accuracy, training and testing speed and memory usage are the main concerns. Recent advances have allowed linear SVM to tackle problems with moderate time and space cost, but for many tasks in computer vision, additive kernels would have higher accuracies. In this paper, we propose the PmSVM-LUT algorithm that employs Look-Up Tables to boost the training and testing speed and save memory usage of additive kernel SVM classification, in order to meet the needs of large scale problems. The PmSVM-LUT algorithm is based on PmSVM (Wu, 2012), which employed polynomial approximation for the gradient function to speedup the dual coordinate descent method. We also analyze the polynomial approximation numerically to demonstrate its validity. Empirically, our algorithm is faster than PmSVM and feature mapping in many datasets with higher classification accuracies and can save up to 60% memory usage as well.

Max Margin Ratio Machine

Suicheng Gu, Yuhong Guo

In this paper, we investigate the problem of exploiting global information to improve the performance of SVMs on large scale classification problems. We first present a unified general framework for the existing min-max machine methods in terms of within-class dispersions and between-class dispersions. By defining a new within-class dispersion measure, we then propose a novel max-margin ratio machine (MMRM) method that can be formulated as a linear programming problem with scalability for large data sets. Kernels can be easily incorporated into our method to address non-linear classification problems. Our empirical results show that the proposed MMRM approach achieves promising results on large data sets.

A Ranking-based KNN Approach for Multi-Label Classification

Tsung-Hsien Chiang, Hung-Yi Lo, Shou-de Lin

Multi-label classification has attracted a great deal of attention in recent years. This paper presents an interesting finding, namely, being able to identify neighbors with trustable labels can significantly improve the classification accuracy. Based on this finding, we propose a k-nearest-neighbor-based ranking approach to solve the multi-label classification problem. The approach exploits a ranking model to learn which neighbor's labels are more trustable candidates for a weighted KNN-based strategy, and then assigns higher weights to those candidates when making weighted-voting decisions. The weights can then be determined by using a generalized pattern search technique. We collect several real-word data sets from various domains for the experiment. Our experiment results demonstrate that the proposed method outperforms state-of-the-art instance-based learning approaches. We believe that appropriately exploiting k-nearest neighbors is useful to solve the multi-label problem.

ABSTRACTS OF ACCEPTED PAPERS (NOVEMBER 6, TUESDAY)

Learning From Ordered Sets and Applications in Collaborative Ranking

Truyen Tran, Dinh Phung, Svetha Venkatesh

Ranking over sets arise when users choose between groups of items. For example, a group may be of those movies deemed 5 stars to them, or a customized tour package. It turns out, to model this data type properly, we need to investigate the general combinatorics problem of partitioning a set and ordering the subsets. Here we construct a probabilistic log-linear model over a set of ordered subsets. Inference in this combinatorial space is highly challenging: The space size approaches (N!=2)6:93145N+1 as N approaches infinity. We propose a split-and-merge Metropolis-Hastings procedure that can explore the state-space efficiently. For discovering hidden aspects in the data, we enrich the model with latent binary variables so that the posteriors can be efficiently evaluated. Finally, we evaluate the proposed model on large-scale collaborative filtering tasks and demonstrate that it is competitive against state-of-the-art methods.

On Using Nearly-Independent Feature Families for High Precision and Confidence

Omid Madani, Manfred Georg, David Ross

Often we require classification at a very high precision level, such as 99%. We report that when very different sources of evidence such as text, audio, and video features are available, combining the outputs of base classifiers trained on each feature type separately, aka late fusion, can substantially increase the recall of the combination at high precisions, compared to the performance of a single classifier trained on all the feature types i.e., early fusion, or compared to the individual base classifiers. We show how the probability of a joint false-positive mistake can be upper bounded by the product of individual probabilities of conditional false-positive mistakes, by identifying a simple key criterion that needs to hold. This provides an explanation for the high precision phenomenon, and motivates referring to such feature families as (nearly) independent. We assess the relevant factors for achieving high precision empirically, and explore combination techniques informed by the analysis. We compare a number of early and late fusion methods, and observe that classifier combination via late fusion can more than double the recall at high precision.

More Is Better: Large Scale Partially-supervised Sentiment Classification

Yoav Haimovitch, Koby Crammer, Shie Mannor

We describe a bootstrapping algorithm to learn from partially labeled data, and the results of an empirical study for using it to improve performance of sentiment classification using up to 15 million unlabeled Amazon product reviews. Our experiments cover semi-supervised learning, domain adaptation and weakly supervised learning. In some cases our methods were able to reduce test error by more than half using such large amount of data.

ABSTRACTS OF ACCEPTED PAPERS (NOVEMBER 6, TUESDAY)

Active Learning with Hinted Support Vector Machine

Chun-Liang Li, Chun-Sung Ferng, Hsuan-Tien Lin

The abundance of real-world data and limited labeling budget calls for active learning, which is an important learning paradigm for reducing human labeling efforts. Many recently developed active learning algorithms consider both uncertainty and representativeness when making querying decisions. However, exploiting representativeness with uncertainty concurrently usually requires tackling sophisticated and challenging learning tasks, such as clustering. In this paper, we propose a new active learning framework, called hinted sampling, which takes both uncertainty and representativeness into account in a simpler way. We design a novel active learning algorithm within the hinted sampling framework with an extended support vector machine. Experimental results validate that the novel active learning algorithm can result in a better and more stable performance than that achieved by state-of-the-art algorithms.

Multi-view Positive and Unlabeled Learning

Joey Tianyi Zhou, Sinno Jialin Pan, Qi Mao, Ivor Wai-Hung Tsang

Learning with Positive and Unlabeled instances (PU learning) arises widely in information retrieval applications. To address the unavailability issue of negative instances, most existing PU learning approaches require to either identify a reliable set of negative instances from the unlabeled data or estimate probability densities as an intermediate step. However, inaccurate negative-instance identification or poor density estimation may severely degrade overall performance of the final predictive model. To this end, we propose a novel PU learning method based on density ratio estimation without constructing any sets of negative instances or estimating any intermediate densities. To further boost PU learning performance, we extend our proposed learning method in a multi-view manner by utilizing multiple heterogeneous sources. Extensive experimental studies demonstrate the effectiveness of our proposed methods, especially when positive labeled data are limited.

ABSTRACTS OF ACCEPTED PAPERS (NOVEMBER 6, TUESDAY)

Frustratingly Simplified Deployment in WLAN Localization by Learning from Route Annotation

Ryoma Kawajiri, Masamichi Shimosaka, Rui Fukui, Tomomasa Sato

Recently wireless LAN (WLAN) localization systems are gaining popularity in pervasive computing, machine learning and sensor networks communities, especially indoor scenarios where GPS coverage is limited. To accurately predict location, a large amount of fingerprints composed of received signal strength values is necessary. Moreover, standard supervised or semi-supervised approaches also require location information to each fingerprint, where annotation work is rather tedious and time consuming. To reduce the efforts and time required to build calibration data, we present a novel calibration methodology \route-annotation" and a self-training algorithm for learning from route information effectively. On the proposed calibration methodology, an annotator walks around while measuring fingerprints, then occasionally stops to annotate fingerprints with route from previous location to current location. This calibration reduces work time even compared to partially annotation, while routes have richer information for learning. The proposed learning algorithm comprises following two iterative steps: 1) inferring locations of each fingerprint under route constraints and 2) updating parameters. Experimental results on real-world datasets demonstrate learning from route-annotated data is comparable to state-of-the-art supervised and semi-supervised approaches trained with large amount of calibration data.

Conditional validity of inductive conformal predictors

Vladimir Vovk

Conformal predictors are set predictors that are automatically valid in the sense of having coverage probability equal to or exceeding a given confidence level. Inductive conformal predictors are a computationally efficient version of conformal predictors satisfying the same property of validity. However, inductive conformal predictors have been only known to control unconditional coverage probability. This paper explores various versions of conditional validity and various ways to achieve them using inductive conformal predictors and their modifications.

Multi-objective Monte-Carlo Tree Search

Weijia Wang, Michele Sebag

Concerned with multi-objective reinforcement learning (MORL), this paper presents MO-MCTS, an extension of Monte-Carlo Tree Search to multi-objective sequential decision making. The known multi-objective indicator referred to as hyper-volume indicator is used to define an action selection criterion, replacing the UCB criterion in order to deal with multi-dimensional rewards. MO-MCTS is firstly compared with an existing MORL algorithm on the artificial Deep Sea Treasure problem. Then a scalability study of MO-MCTS is made on the NP-hard problem of grid scheduling, showing that the performance of MO-MCTS matches the non RL-based state of the art albeit with a higher computational cost.

ABSTRACTS OF ACCEPTED PAPERS (NOVEMBER 6, TUESDAY)

A Stochastic Bandit Algorithm for Scratch Games

Raphaël Féraud, Tanguy Urvoy

Stochastic multi-armed bandit algorithms are used to solve the exploration and exploitation dilemma in sequential optimization problems. The algorithms based on upper confidence bounds offer strong theoretical guarantees, they are easy to implement and efficient in practice. We considers a new bandit setting, called "scratch-games", where arm budgets are limited and reward are drawn without replacement. Using Serfling inequality, we propose an upper confidence bound algorithm adapted to this setting. We show that the bound of expectation to play a suboptimal arm is lower than the one of UCB1 policy. We illustrate this result on both synthetic problems and realistic problems (ad-serving and emailing campaigns optimization).

AIC and BIC based approaches for SVM parameter value estimation with RBF kernels

Sergey Demyanov, James Bailey, Ramamohanarao Kotagiri, Christopher Leckie

We study the problem of selecting the best parameter values to use for a support vector machine (SVM) with RBF kernel. Our methods extend the well-known formulas for AIC and BIC, and we present two alternative approaches for calculating the necessary likelihood functions for these formulas. Our first approach is based on using the distances of support vectors from the separating hyperplane. Our second approach estimates the probability that the SVM hyperplane coincides with the Bayes classifier, by analysing the disposition of points in the kernel feature space. We experimentally compare our two approaches with several existing methods and show they are able to achieve good accuracy, whilst also having low running time.

Online Rank Aggregation

Shota Yasutake, Kohei Hatano, Eiji Takimoto, Masayuki Takeda

We consider an online learning framework where the task is to predict a permutation which represents a ranking of n fixed objects. At each trial, the learner incurs a loss defined as Kendall tau distance between the predicted permutation and the true permutation given by the adversary. This setting is quite natural in many situations such as information retrieval and recommendation tasks. We prove a lower bound of the cumulative loss and hardness results. Then, we propose an algorithm for this problem and prove its relative loss bound which shows our algorithm is close to optimal.

ABSTRACTS OF ACCEPTED PAPERS (NOVEMBER 6, TUESDAY)

Online Learning of a Dirichlet Process Mixture of Generalized Dirichlet Distributions for Simultaneous Clustering and Localized Feature Selection

Wentao Fan, Nizar Bouguila

Online algorithms allow data instances to be processed in a sequential way, which is important for large-scale and real-time applications. In this paper, we propose a novel online clustering approach based on a Dirichlet process mixture of generalized Dirichlet (GD) distributions, which can be considered as an extension of the finite GD mixture model to the infinite case. Our approach is built on nonparametric Bayesian analysis where the determination of the number of clusters is sidestepped by assuming an infinite number of mixture components. Moreover, an unsupervised localized feature selection scheme is integrated with the proposed nonparametric framework to improve the clustering performance. By learning the proposed model in an online manner using a variational approach, all the involved parameters and features saliencies are estimated simultaneously and effectively in closed forms. The proposed online infinite mixture model is validated through both synthetic data sets and two challenging real-world applications namely text document clustering and online human face detection.

WORKSHOP PROGRAM

LAWS'12	
08:30am - 08:35am	**Workshop Opening**
08:35am - 09:30am	**Keynote Talk** Learning with Weak Supervision: Charting the Territory *Bernhard Pfahringer*
09:30am - 10:10am	**Invited Talk** Positive Unlabelled Learning and Its Applications *Xiaoli Li*
10:10am - 10:30am	Coffee Break
10:30am - 10:55am	A Functional-Space Approach to Budgeted Knowledge Transfer Among RL Agents *Farbod Farshidian, Zeinab Talebpour, Majid Nili Ahmadabadi*
10:55am - 11:20am	A Novel Classification Method with Cross-View Constraints *Hui Xue, Songcan Chen, Jie Liu, Jijian Huang*
11:20am - 11:45am	Constrained Modularity for Semi-Supervised Community Detection *Steve Poulson, Dell Zhang, Mark Levene*
11:45am - 12:10pm	Cost-Sensitive Semi-Supervised Software Defect Prediction *Ming Li, Wei-Wei Tu, Chen-You Fan, Zhi-Hua Zhou*
12:10pm - 12:35pm	Active Semi-Supervised Concept Hierarchy Refinement *Michał Dereziński*

WORKSHOP PROGRAM

FDMA'12	
13:30pm - 14:00pm	**Introduction & Prize Giving to Competition Winners**
14:00pm	**Winners Presentation Session** Presentation by Winner 1 Feature Engineering for Click Fraud Detection Speaker: Dr. Clifton PHUA Institute for Infocomm Research, Singapore
15:00pm	Presentation by Winner 2 A Novel Approach Based on Ensemble Learning for Fraud Detection in Mobile Advertising Speaker: Kasun S. PERERA and Bijay NEUPANE Masdar Institute of Science and Technology, Abu Dhabi, UAE
15:10pm - 15:30pm	Coffee Break
15:30pm	**Winners Presentation Session (cont')** Presentation by Winner 3 Hybrid Models for Click Fraud Detection in Mobile Advertising Speaker: Wei CHEN National University of Singapore, Singapore
16:00pm	**Runner-ups Presentation Session** Hierarchical Committee Machines for Fraud Detection in Mobile Advertising Speaker: Manoj Prasanna KUMAR Ericsson India Global Services Pvt. Ltd., Tamil Nadu, India
16:30pm - 17:30pm	**Poster Presentation Session** by Top 10 Teams

CONFERENCE ORGANIZATION

General Co-Chairs

- Wee Sun Lee (National University of Singapore)
- Zhi-Hua Zhou (Nanjing University)

Program Co-Chairs

- Steven C.H. Hoi (Nanyang Technological University)
- Wray Buntine (NICTA)

Local Arrangement Committee

- Jing Jiang (Co-chair, Singapore Management University)
- Ivor Tsang (Co-chair, Nanyang Technological University)
- Yi Wang (National University of Singapore)
- Sintiani Dewi Teddy (Institute for Infocomm Research)

Sponsorship Chair

- David Lo (Singapore Management University)

Finance Chair

- Jianxin Wu (Nanyang Technological University)

Tutorial Co-Chairs

- Hai Leong Chieu (DSO National Laboratories)
- Stanley Kok (Singapore University of Technology and Design)

Workshop Co-Chairs

- David Hardoon (SAS)
- Huan Xu (National University of Singapore)

Publication Chair

- Sinno Jialin Pan (Institute for Infocomm Research)

Local Advisory Commitee

- Ee Peng Lim (Singapore Management University)
- Ah Hwee Tan (Nanyang Technological University)
- Yew Soon Ong (Nanyang Technological University)

CONFERENCE ORGANIZATION

ACML Steering Committee

- Tom Dietterich (Oregon State University, USA)
- Tu Bao Ho (JAIST, Japan)
- Hiroshi Motoda, Chair (Osaka University, Japan)
- Bernhard Pfahringer (Waikato University, New Zealand)
- Masashi Sugiyama (Tokyo Institute of Technology, Japan)
- Takashi Washio (Osaka University, Japan)
- Geoff Webb (Monash University, Australia)
- Qiang Yang (Hong Kong University of Science and Technology, Hong Kong)
- Zhi-Hua Zhou, Co-Chair (Nanjing University, China)

Senior Program Committee

- Edward Chang (Google Research Asia)
- Wei Fan (IBM TJ Watson Lab)
- Xiaofei He (Zhejiang University)
- Tu Bao Ho (Japan)
- Chun-Nan Hsu (University of Southern California)
- Aapo Hyvarinen (Helsinki Institute for Information Technology)
- Rong Jin (Michigan State University)
- Kristian Kersting (University of Bonn)
- Kee-Eung Kim (KAIST)
- Irwin King (Chinese University of Hong Kong)
- James Kwok (Hong Kong University of Science and Technology)
- Pavel Laskov (University of Tubingen)
- Yuh-Jye Lee (NTUST)
- YoonKyung Lee (Ohio State University)
- Ping Li (Cornell University)
- Hang Li (Microsoft Research Asia)
- Phil Long (Google)
- Klaus-Robert Muller (Technical University Berlin)
- Bernhard Pfahringer (University of Waikato)
- Scott Sanner (Australian National University)
- Shirish Shevade (Indian Institute of Science)
- Masashi Sugiyama (Tokyo Institute of Technology)
- Jieping Ye (Arizona State University)
- Dell Zhang (Birbeck College University of London)
- Min-Ling Zhang (Southeast University)
- Xiaojin Jerry Zhu (University of Wisconsin-Madison)

CONFERENCE ORGANIZATION

Program Committee

- Albert Bifet (Yahoo! Research)
- Alain Rakotomamonjy (University of Rouen)
- Yuan Qi (Purdue University)
- Bin Gao (Microsoft Research Asia)
- Ulf Brefeld (University of Bonn)
- Wei Bian (University of Technology Sydney)
- Li Cheng (Bioinformatics Institute)
- Chong Wang (Princeton University)
- Alfredo Cuzzocrea (Italian National Research Council)
- Jun Zhu (Tsinghua University)
- Decheng Dai (Google Research)
- Edwin Bonilla (NICTA)
- Shuicheng Yan (NUS)
- Francesco Dinuzzo (Max Planck Institute for Intelligent Systems)
- Gang Niu (Tokyo Institute of Technology)
- Tsuyoshi Ide (IBM)
- Grigorios Tsoumakas (Aristotle University of Thessaloniki)
- Haixuan Yang (Royal Holloway University of London)
- Kohei Hatano (Kyusyu University)
- Hirayama Junichiro (ATR and University of Kyoto)
- Tuyen N. Huynh (SRI International)
- Haiqin Yang (Chinese University of Hong Kong)
- Hsuan-Tien Lin (National Taiwan University)
- Yang Hui (CMU)
- Hui Xue (Southeast University)
- Issei Sato (University of Tokyo)
- Jianhui Chen (GE Global Research)
- Jianke Zhu (Zhejiang University)
- Johan Suykens (KU Leuven)
- Jun Sakuma (Tsukuba University)
- Jinzhu Jia (Peking University)
- Masayuki Karasuyama (Kyoto University)
- Kaizhu Huang (Chinese Academy of Sciences)
- Hisashi Kashima (University of Tokyo)
- Motoaki Kawanabe (ATR)
- Kevin Duh (NAIST)

- Koji Tsuda (National Institute of Advanced Industrial Science and Technology)
- Stefan Kramer (University of Mainz)
- Kai Zhang (Siemens Corporate Research, Princeton)
- Lei Wang (The Australian National University)
- Lei Yuan (Arizona State University)
- Ming Liu (Nanjing University)
- Xu-Ying Liu (Southeast University, China)
- Liang Sun (Opera Solutions)
- Luis Torgo (University of Porto)
- Marc Deisenroth (University of Washington)
- Marc Deisenroth (TU Darmstadt)
- Mark Reid (Australian National University)
- Yiming Ying (Univerity of Exeter)
- Yuji Matsumoto (Nara Institute of Science and Technology)
- Ding-Xuan Zhou (City University of Hong Kong)
- Gutmann Michael (University of Helsinki)
- Mingmin Chi (Fudan University)
- Shinichi Nakajima (Nikon)
- Sriraam Natarajan (University of Wisconsin)
- Jinfeng Zhuang (Microsoft Inc.)
- Peter Andreae (Victoria University)
- Peter Sunehag (ANU)
- Pingkun Yan (Chinese Academy of Sciences)
- Quanquan Gu (UIUC)
- Achim Rettinger (Karlsruher Institute of Technology)
- Rudy Setiono (National University of Singapore)
- Sanjay Jain (National University of Singapore)
- Shou-de Lin (National Taiwan University)
- Shalini Ghosh (SRI International)
- Shengbo Guo (Xerox Research Centre Europe)
- See-Kiong Ng (institute for infocomm research)
- Shohei Shimizu (Osaka University)
- Taiji Suzuki (University of Tokyo)
- Ichiro Takeuchi (Nagoya Institute of Technology)
- Tao Qin (Microsoft Research Asia)
- Tetsuro Morimura (IBM)
- Thomas J. Walsh (University of Kansas)

- Tohgoroh Matsui (College of Life and Health Sciences, Chubu University)
- Ryota Tomioka (University of Tokyo)
- Wei Tong (CMU)
- Naonori Ueda (NTT)
- Vikas Sindhwani (IBM T.J. Watson Research Center)
- Volker Roth (University of Basel)
- Jun Wang (IBM)
- Liwei Wang (Peking University)
- Wai Lam (Chinese University of Hong Kong)
- Wei Liu (Columbia University)
- Junjie Wu (Beihang University)
- Lei Wu (GE Global Research)
- Wei Wu (Microsoft Research Asia)
- Xingquan Zhu (University of Technology Sydney)
- Xinhua Zhang (University of Alberta)
- Xiaoli Li (Institute for Infocomm Research)
- Xiaotong Yuan (Rutgers University)
- Zenglin Xu (Purdue University)
- Xuelong Li (Chinese Academy of Sciences)
- Xu Sun (Cornell University)
- Makoto Yamada (Tokyo Institute of Technology)
- Tianbao Yang (Michigan State University)
- Yanjun Qi (NEC)
- Yisong Yue (Carnegie Mellon University)
- Yun Sing Koh (University of Auckland)
- Yang Yu (Nanjing University)
- Changshui Zhang (Tsinghua University)
- De-Chuan Zhan (Nanjing University)
- Zhao Xu (Fraunhofer IAIS)
- Zhao Zheng (SAS Institute)
- Zhengdong Lv (Microsoft Research Asia)
- Zhiyong Liu (Chinese Academy of Sciences)
- Zhihua Zhang (Zhejiang University)
- Ziyu Guan (UCSB)

CONFERENCE ORGANIZATION

Best Paper Award Committee

- Wray Buntine, Chair (NICTA)
- Steven C.H. Hoi (Nanyang Technological University)
- Chun-Nan Hsu (University of Southern California)
- Hiroshi Motoda (Osaka University)
- Bob Williamson (Australian National University and NICTA)

SPONSORS

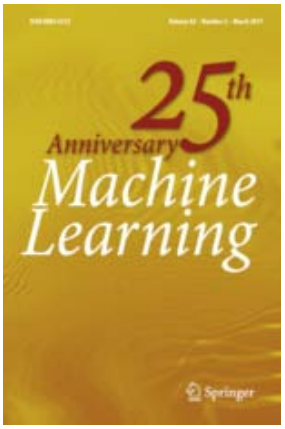
Acknowledgements

We would like to thank the Air Force Office of Scientific Research, Asian Office of Aerospace Research and Development, Huawei, Lee Foundation, and the Machine Learning journal for their generous support.

Gold Sponsors

Silver Sponsors

Best Student Paper Award Sponsor



SUPPORTING ORGANIZATIONS

VENUE INFORMATION

Main Conference

Ngee Ann Kongsi Auditorium, Level 2, School of Accountancy,
Singapore Management University, 60 Stamford Road, Singapore 178900

Workshop Venue

Seminar Room 2-2, Level 2, School of Accountancy,
Singapore Management University, 60 Stamford Road, Singapore 178900

Poster Session

Reading Room, Level 5, Li Ka Shing Library,
Singapore Management University, 70 Stamford Road, Singapore 178901

Reception Venue

U-Lounge, Level 6, Administration Building,
Singapore Management University, 81 Victoria Street, Singapore 188065

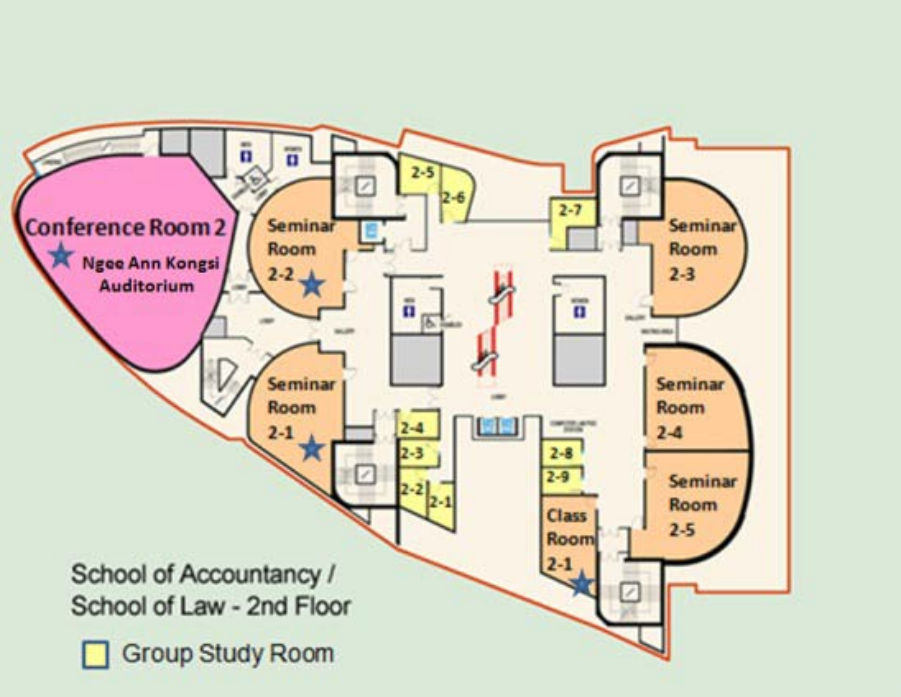
Banquet Venue (to be confirmed during the conference)

Carlton Hotel Singapore
76 Bras Basah Road, Singapore 189558

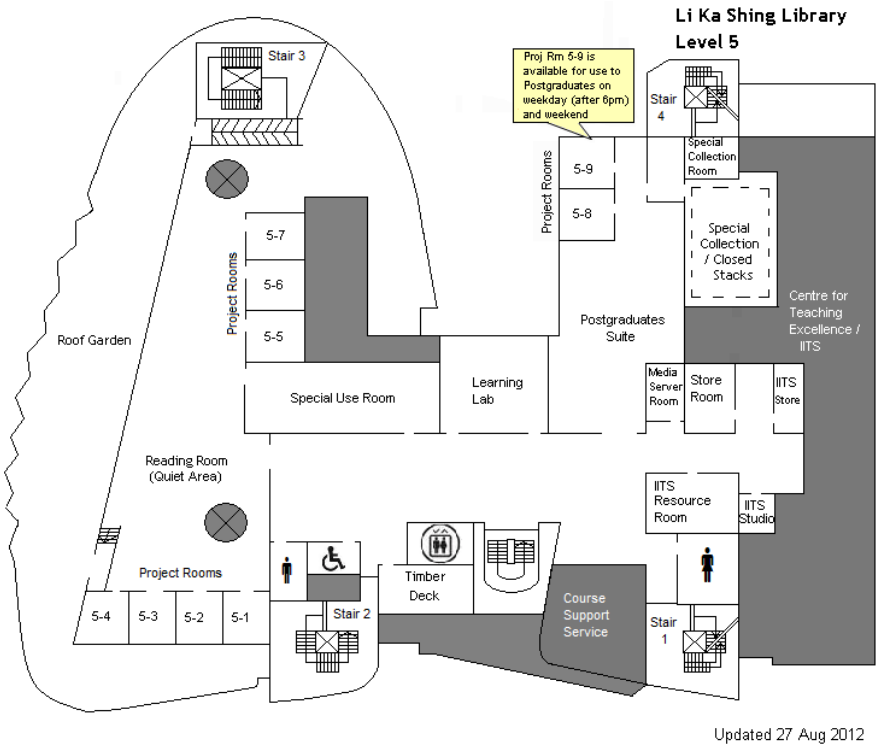
Please see floor plans and maps in the next pages.

VENUE INFORMATION

Floor Plan of Level 2, School of Accountancy, SMU



Floor Plan of Level 5, Li Ka Shing Library, SMU



VENUE INFORMATION

Map of SMU

Map of Carlton Hotel

Link to Google maps

TRANSPORTATION

MRT (Train) Stations Near SMU

1. Bras Basah Station, Circle Line

   Exit the station at the concourse (basement 1) and follow the overhead signage to reach the School of Accountancy (approximately 5 minutes).

2. Dhoby Ghaut Station, North-South Line

   Leave the station at exit A and walk past School of Economics/Social Science, School of Information System and Li Ka Shing Library to reach the School of Accountancy (approximately 10 minutes).

3. City Hall Station, East-West Line

   Leave the station at exit A and walk towards the School of Business to reach the School of Accountancy (approximately 10 minutes).

Bus Services Near SMU

1. Victoria Street

   Bus stop number: 04151
   Bus services: 130, 133, 145, 197, 851, 960

2. SMU, Stamford Road

   Bus stop number: 04121
   Bus services: 7, 14, 16, 36, 77, 106, 111, 124, 128, 131, 147, 14E, 162, 166, 167, 171, 174, 175, 190, 502, 700, 857, NR7, 162M, 174E, 502A, 502B, 700A

3. NTUC Income Centre, Bras Basah Road

   Bus stop number: 04179
   Bus services: 7, 14, 16, 36, 77, 106, 111, 128, 131, 14E, 162, 167, 171, 175, 502, 518, 587, 590, 598, 700, 857, NR7, 162M, 502A, 502B, 518A, 518B, 700A

RESTAURANTS AND CAFÉS

1. Straits Café at Rendezvous

   Offers international buffet breakfast, lunch & dinner (7am to 11pm daily)
   Buffet from $32++ per person

2. Kopitiam Food Court

   Indulge in a variety of local delights (opens 24 hours daily)
   Chicken rice from $4.8++, Noodles from $3.5++, Coffee & Toast from $2++

3. Wendy's Fast Food

   Grab your burgers and fries (7 am to 11pm daily)
   Combo meal from $7++

4. The Coffee Bean & Tea Leaf

   Chill with coffee & cakes (7.30am to 11pm Mon - Sat, 9am to 11pm Sun & PH)
   Coffee from $4++, Cake from $6++

5. Dome Cafe

   Fresh coffee, cakes & full meals (8.30am to 10.30pm daily)
   Salad from $7.5++, Pasta from $12.8++, Pizza from $12.50++

6. Wahlok Restaurant in Carlton Hotel

   Offers fine Cantonese cuisine (11.30am to 10.30pm daily)
   8 course set menu start from $88 per person

7. CHIJMES

   A range of Westerns/Chinese/Japanese cuisine are available
   Japanese BBQ about $30++ per person, Grill & ribs about $20++ per person

8. Koufu Food Court

   Indulge in a variety of local delights (7.30am to 8pm Mon - Fri)
   Western delight from $4.5++, Fish soup from $4++, Economic rice from $3++

9. Metro-Y Restaurant

   The best of local porridge buffet (7:15am to 10pm Mon - Fri)
   Buffet from $13.80++ per person

10. National Museum of Singapore

    Offers a selection of fine dining restaurant and cozy cafes
    Chef Chan set menu from $38++ per person, Novus Bar & Café set lunch from $32++

RESTAURANTS AND CAFÉS