

Machine Learning Summer School

3-14 March 2008

Kioloa Coastal Campus, Australia

Avrim Blum, Carnegie Mellon, Learning in Games

Wray Buntine, NICTA, Models for Documents

Tiberio Caetano, NICTA, Introduction to Graphical Models

Nando de Freitas, U British Columbia, Sampling Methods

Marcus Hutter, ANU, Introduction to Machine Learning

Rao Kotagiri, Melbourne University, Data Mining

Simon Lucey, Carnegie Mellon, Learning in Computer Vision

Alexander Smola, NICTA, Introduction to Kernel Methods

Csaba Szepesvari, Alberta, Reinforcement Learning

Vishy Vishwanathan, NICTA, Machine Learning Laboratory

The Summer School is organized by the ANU and NICTA jointly with support from the PASCAL Network and the MPI Tübingen. For more details please contact:

Marcus.Hutter@nicta.com.au, Li.Cheng@nicta.com.au or Alex.Smola@nicta.com.au

Website: kioloa08.mlss.cc

Table of Contents

Machine Learning Summer School

- Preface
- Sponsors and Acknowledgements
- Schedule
- Speaker abstracts
- Social program
- Food Menu

- Marcus Hutter**
- Introduction to Statistical Machine Learning
 - Foundations of Machine Learning

- Alex Smola**
- Kernel methods and Support Vector Machines

- Nando De Freitas**
- Monte Carlo Methods in Learning

- Wray Buntine**
- Models for Document Analysis

- Csaba Szepesvari**
- Reinforcement Learning

- Tiberio Caetano**
- The Very Basics of Graphical Models

- Rao Kotagiri**
- Contrast Data Mining: Methods and Applications

- Simon Lucey**
- Object Registration and Tracking from a Learning Perspective

- Avrim Blum**
- Online Learning, Regret Minimization, and Game Theory
 - On Learning with Similarity Functions

- SVN Vishwanathan**
- Machine Learning Laboratory on Elephant
 - Machine Learning Laboratory on BNRM

- Kioloa**
- Getting to and from Kioloa
 - Maps of the Campus, Kioloa, New South Wales, Australia
 - Some Photos of the Campus
 - Checklist for Campus Usage
 - Safety Guidelines
 - Contact (registration, organizers, accommodation)
 - Conditions of Campus Use

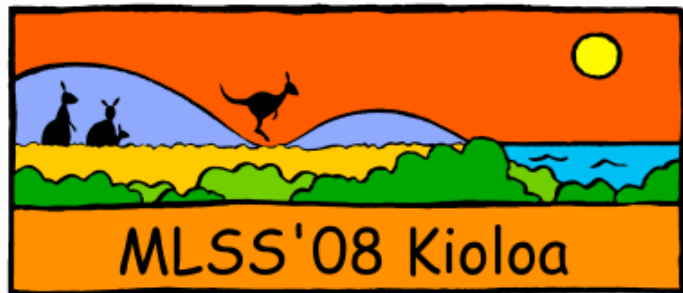
- Australia**
- Etymology, history, politics, states&territories, foreign relations, geography, economy, demography, culture.
 - The Australian National University (profile, scholarships, SML)
 - National ICT of Australia (SML, about NICTA)

- Local Sightseeing**
- Excursion Schedule
 - Local National Parks
 - River Cruises
 - Local History
 - Old Mogo Town

- Personal Notes**
- Empty pages

Machine Learning Summer School 2008

**Dates: Mon 3 Mar to Fri 14
Mar, 2008**



Venue:

[Kioloa Coastal Campus, Australian National University](#)

Overview

Machine Learning is a foundational discipline of the Information Sciences. It combines deep theory from areas as diverse as Statistics, Mathematics, Engineering, and Information Technology with many practical and relevant real life applications. The aim of the summer school is to cover the entire spectrum from theory to practice. It is mainly targeted at research students, IT professionals, and academics from all over the world.

This school is suitable for all levels, both for people without previous knowledge in Machine Learning, and those wishing to broaden their expertise in this area. It will allow the participants to get in touch with international experts in this field. Exchange of students, joint publications and joint projects will result because of this collaboration.

For research students, the summer school provides a unique, high-quality, and intensive period of study. It is ideally suited for students currently pursuing, or intending to pursue, research in Machine Learning or related fields. Limited scholarships are available for students to cover accommodation and registration costs. If funds are available partial travel support might also be provided.

IT professionals who use Machine Learning will find that the summer school provides relevant knowledge and exposure to contemporary techniques. In addition, they will benefit by direct interaction with top-notch researchers and knowledge workers. Previous experience indicates that personnel from both the industry as well as national laboratories like CSIRO, DSTO benefit immensely from the school.

For academics, the summer school is an excellent opportunity to help getting started in research on novel topics in Machine Learning. It provides an ideal forum for networking and discussions. Academics will also benefit from interaction with IT professionals which will lead to a deeper understanding of real life problems.

Venue

The summer school will be held on Kioloa Coastal Campus, 200km south of Sydney and 200km east of Canberra on the NSW south coast. The unique blend of living history and modern technology underlies the atmosphere of the Campus. Kioloa Coastal Campus has become the premier research field station on the South Coast. With a mild climate and unspoiled beach frontage, the Campus provides an ideal break from normal campus and work environments.

Organizers

This Summer School is organized by the Computer Sciences Laboratory of the Australian National University (CSL@ANU) and the Statistical Machine Learning program of the National ICT Australia (SML@NICTA), jointly with support from the Max-Planck-Institute for Biological Cybernetics in Tübingen and the Pascal Network. Please visit www.mlss.cc for more information about the [previous summer schools](#). [Local organizers](#) are Li Cheng, Marcus Hutter, and Alex Smola.

Sponsors

We would like to thank all sponsors below for their generous support.



[ANU](#)
[CSL@RSISE](#)
[Machine Learning Group](#)



[SML](#)
[NICTA](#)



[MPI Tuebingen](#)



[PASCAL](#)

The organizers would also like to thank all people who helped in organizing the Summer School, in particular the SML students and Michelle Moravec.

Schedule

Welcome Reception: Sunday, 19:00-22:00, 2nd of March, 2008.

Lectures: [Kiolea coastal campus](#).

Week 1: 3th of Mar 2008 - 7th of Mar 2008

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
07:30 – 08:30		Breakfast					
08:30 – 09:30		Marcus Hutter Introduction	Nando de Freitas Agents & Sampling Methods		Csaba Szepesvari Reinforcement Learning		
09:30 – 09:45		Coffee Break					
09:45 – 10:45		Marcus Hutter Introduction	Nando de Freitas Agents & Sampling Methods		Csaba Szepesvari Reinforcement Learning		
10:45 – 11:00		Coffee Break					
11:00 – 12:00		Marcus Hutter Introduction	Nando de Freitas Agents & Sampling Methods		Csaba Szepesvari Reinforcement Learning		
12:00 – 13:00		Lunch Break					
13:00 – 15:30		<i>Free Time</i>					Excursion
15:30 – 16:30		Alex Smola Kernels & SVMs		Wray Buntine Document Models		Vishwanathan Laboratory	
16:30 – 16:45		Coffee Break					
16:45 – 17:45	<i>Arrival</i>	Alex Smola Kernels & SVMs		Wray Buntine Document Models		Vishwanathan Laboratory	
17:45 – 18:00		Coffee Break					
18:00 – 19:00		Alex Smola Kernels & SVMs		Wray Buntine Document Models		Vishwanathan Laboratory	
19:00 – 21:00	Reception	Supper Break				Barbecue	
21:00 – 22:00		<i>Free Time</i>					

Schedule

Week 2: 10th of Mar 2008 - 14th of Mar 2008

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
07:30 – 08:30	<i>Free</i>	Breakfast					<i>Departure</i>
08:30 – 09:30		Marcus Hutter Foundations	Rao Kotagiri Data Mining	Avrim Blum Game Theory			
09:30 – 09:45		Coffee Break					
09:45 – 10:45		Marcus Hutter Foundations	Rao Kotagiri Data Mining	Avrim Blum Game Theory			
10:45 – 11:00		Coffee Break					
11:00 – 12:00		Marcus Hutter Foundations	Rao Kotagiri Data Mining	Avrim Blum Game Theory			
12:00 – 13:00		Lunch Break					
13:00 – 15:30		<i>Free Time</i>					
15:30 – 16:30		Tiberio Caetano Graphical Models	Simon Lucey Computer Vision	Vishwanathan Laboratory			
16:30 – 16:45		Coffee Break					
16:45 – 17:45		Tiberio Caetano Graphical Models	Simon Lucey Computer Vision	Vishwanathan Laboratory			
17:45 – 18:00		Coffee Break					
18:00 – 19:00		Tiberio Caetano Graphical Models	Simon Lucey Computer Vision	Vishwanathan Laboratory			
19:00 – 21:00		Supper Break					
21:00 – 22:00	<i>Free Time</i>						

Invited Speakers

[Avrim Blum](#)



Online Learning, Regret Minimization, and Game Theory [\[slides\]](#)

Avrim Blum is Professor of Computer Science in the Department of Computer Science at Carnegie Mellon University.

The first part of his tutorial will discuss adaptive algorithms for making decisions in uncertain environments (e.g., what route should I take to work if I have to decide before I know what traffic will like today?) and connections to central concepts in game theory (e.g., what can we say about how traffic will behave overall if everyone is adapting their behavior in such a way?). He will discuss the notions of external and internal regret, algorithms for "combining expert advice" and "sleeping experts" problems, algorithms for implicitly specified problems, and connections to game-theoretic notions of Nash and correlated equilibria.

The second part of his tutorial will be about some recent work on learning with similarity functions that are not necessarily legal kernels. The high-level question here is: if you have a measure of similarity between data points, how closely related does it have to be to your classification problem in order to be useful for learning?

[Wray Buntine](#)



Latent Variable Models for Document Analysis [\[slides\]](#)

Wray Buntine is a researcher in the Statistical Machine Learning group at NICTA's Canberra Laboratory.

He will consider various problems in document analysis (named entity recognition, natural language parsing, information retrieval), and look at various probabilistic graphical models and algorithms for addressing the problem. This will not be an extensive coverage of information extraction or natural language processing, but rather a look at some of the theory, methods and practice of particular cases, including the use of software environments.

[Tiberio Caetano](#)

Inference in Graphical Models [\[slides\]](#)

Tiberio Caetano is senior researcher at NICTA's Canberra Laboratory, where he is a member of the Statistical Machine Learning research program.

His short course will cover the basics of inference in graphical models. It will start by explaining the theory of probabilistic graphical models, including concepts of



Nando de Freitas



conditional independence and factorisation and how they arise in both Markov random fields and Bayesian Networks. He will then present the fundamental methods for performing exact probabilistic inference in such models, which include algorithms like variable elimination, belief propagation and Junction Trees. He will also briefly discuss some of the current methods for performing approximate inference when exact inference is not feasible. Finally, he will illustrate a range of real problems whose solutions can be formulated as inference in graphical models.

Monte Carlo Simulation for Statistical Inference, Model Selection and Decision Making [\[slides\]](#)

Nando de Freitas is Associate professor in the Department of Computer Science at the University of British Columbia .

The first part of his course will consist of two presentations. In the first presentation, he will introduce fundamentals of Monte Carlo simulation for statistical inference, with emphasis on algorithms such as importance sampling, particle filtering and smoothing for dynamic models, Markov chain Monte Carlo, Gibbs and Metropolis-Hastings, blocking and mixtures of MCMC kernels, Monte Carlo EM, sequential Monte Carlo for static models, auxiliary variable methods (Swendsen-Wang, hybrid Monte Carlo and slice sampling), and adaptive MCMC. The algorithms will be illustrated with several examples: image tracking, robotics, image annotation, probabilistic graphical models, and music analysis.

The second presentation will target model selection and decision making problems. He will describe the reversible-jump MCMC algorithm and illustrate it with application to simple mixture models and nonlinear regression with an unknown number of basis functions. He will show how to apply this algorithm to general Markov decision processes (MDPs). The course will also cover other Monte Carlo simulation methods for partially observed Markov decision processes (POMDPs) using policy gradients, common random number generation, and active exploration with Gaussian processes. An outline to some applications of these methods to robotics and the design of computer game architectures will be given. The presentation will end with the problem of Monte Carlo simulation for Bayesian nonlinear experimental design, with application to financial modeling, robot exploration, drug treatments, dynamic sensor networks, optimal measurement and active vision.

Marcus Hutter



Introduction and Foundations of Machine Learning [\[slides1\]](#) [\[slides2\]](#)

Marcus Hutter is Associate Professor in the RSISE at the Australian National University in Canberra and NICTA adjunct.

The first part of his tutorial provides a brief overview of the fundamental methods and applications of statistical machine learning. The other speakers will detail or build upon this introduction. Statistical machine learning is concerned with the development of algorithms and techniques that learn from observed data by constructing stochastic models that can be used for making predictions and decisions. Topics covered include Bayesian inference and maximum likelihood modeling; regression, classification, density estimation, clustering, principal component analysis; parametric, semi-parametric, and non-parametric models; basis functions, neural networks, kernel methods, and graphical models; deterministic and stochastic optimization; overfitting, regularization, and validation.

Machine learning is usually taught as a bunch of methods that can solve a bunch of problems (see above). The second part of the tutorial takes a step back and asks about the foundations of machine learning, in particular the (philosophical) problem of inductive inference, (Bayesian) statistics, and artificial intelligence. It concentrates on principled, unified, and exact methods.

Rao Kotagiri



Contrast Data Mining: Methods and Applications [\[slides\]](#)

Ramamohanarao (Rao) Kotagiri is Professor in the Department of Computer Science and Software Engineering at the University of Melbourne.

The ability to distinguish, differentiate and contrast between different datasets is a key objective in data mining. Such an ability can assist domain experts to understand their data, and can help in building classification models. His presentation will introduce the principal techniques for contrasting different types of data, covering the main dataset varieties such as relational, sequence, and graph forms of data, clusters, as well as data cubes. It will also focus on some important real world application areas that illustrate how mining contrasts is advantageous.

Simon Lucey

Learning in Computer Vision [\[slides\]](#)

Dr. Simon Lucey is Systems Scientist in the Robotics Institute at Carnegie Mellon University.



Alex Smola

In his tutorial he will cover some of the core fundamentals in vision and demonstrate how they can be interpreted in terms of machine learning fundamentals. Unbeknownst to most researchers in the field of machine learning, the fundamentals of object registration and tracking such as optical flow, interest descriptors (e.g., SIFT), segmentation and correlation filters are inherently related to the learning topics of regression, regularization, graphical models, generative models and discriminative models. As a result many aspects of vision can be interpreted as applied forms of learning. From this discussion on fundamentals we shall also explore advanced topics in object registration and tracking such as non-rigid object alignment/ tracking and non-rigid structure from motion and how the application of machine learning is continuing to improve these technologies.

Kernel methods and Support Vector Machines
[\[slides\]](#)



Alex Smola is leader of NICTA's Statistical Machine Learning Program in Canberra and Professor in the Computer Sciences Laboratory at the Australian National University.

His tutorial will introduce the main ideas of statistical learning theory, support vector machines, and kernel feature spaces. This includes a derivation of the support vector optimization problem for classification and regression, the v-trick, various kernels and an overview over applications of kernel methods.

Csaba Szepesvari



Introduction to Reinforcement Learning
[\[slides\]](#)

Csaba Szepesvari is Associate Professor in the Department of Computing Science at University of Alberta.

His tutorial will introduce Reinforcement Learning, that is, learning what actions to take, and when to take them, so as to optimize long-term performance. This may involve sacrificing immediate reward to obtain greater reward in the long-term or just to obtain more information about the environment. The first part of the tutorial will cover the basics, such as Markov decision processes, dynamic programming, temporal-difference learning, Monte Carlo methods, eligibility traces, the role of function approximation. In the second part we cover some recent developments, namely policy gradient and second order methods, such as LSPI and the modified

Bellman residual minimization algorithm.

[Vishy Vishwanathan](#)



Machine Learning Laboratory [\[slides\]](#)

Dr. Vishwanathan is senior researcher at NICTA's Canberra Laboratory, where he is a member of the Statistical Machine Learning research program.

The first laboratory on March 7 will feature some hands on experiments with Elefant

(<http://elefant.developer.nicta.com.au>) mainly concentrating on installing, using, and developing machine learning algorithms within the Elefant framework. We will walk through examples of implementing a simple stochastic gradient descent algorithm as a part of this tutorial.

The second session on March 14 will feature hands on experiments with BMRM (Bundle Methods for Regularized Risk Minimization) (<http://users.rsise.anu.edu.au/~chteo/BMRM.html>). The emphasis here will be on developing various loss function modules which can then be plugged into the BMRM solver.

CD's and USB sticks containing installation instructions for Elefant and BMRM will be handed out during the session. Preferably bring your own laptops, although some spare PC's might be made available. Students will also have a chance to interact with the leading developers of Elefant and BMRM.

Social Program

Reception

The welcome reception is on Sunday, 2 March, from 19:00 to 22:00.

Excursion

We have chartered a bus for a day's local excursion along the coastline on Saturday 8 March. It includes a boat cruise along the scenic Clyde River at [Batemans Bay](#), exploration of Mogo and Old Mogo Town, and a visit to [Murramarang National Park](#). The detailed plan described below is tentative and subject to change.

The excursion starts the morning of Saturday 8 March. Departing after breakfast, a bus will take us to Mogo where we will have a guided tour through Old Mogo Town, faithful re-creation of Mogo as it was believed to be in the Gold Rush days of the 1850s. After a bit of exploration and morning tea, we move on to the 3-hour boat trip (11:30-14:30) on the Clyde River. We will have lunch on the boat.

After the relaxation of the river cruise, we proceed to Murramarang National Park. After a short driving tour through the park, everyone will be dropped at [Pebbly Beach](#) area for the afternoon. People may walk along the coastal tracks in the forests, watch the wildlife (including kangaroos), or swim. At approximately 18:00, we meet together at the pick-up place and proceed to Batemans Bay for dinner.

07:30 to 08:30	Breakfast
08:30 to 11:15	Mogo and Old Mogo Town
11:30 to 14:30	Clyde River Luncheon Cruise
14:30 to 15:00	Drive through Murramarang National Park
15:00 to 18:00	Pebbly Beach
18:00 to 18:30	Travel to Batemans Bay
18:30 to 21:00	Dinner in Restaurant
21:00 to 21:30	Return to Kioloa Campus

Free Time

There will be plenty of free time. The long **lunch breaks** from 12:00-15:30 are ideal for recreation at the beach (5-minute walking distance). Don't forget sun protection and a hat.

Sunday 9 March is mainly free. You can explore the local area flora and fauna (best in the morning or late afternoon) or do your own sightseeing or excursion, or just relax. We are exploring the option of providing bushtucker tours of the local area.

WEEK ONE (Tentative Menu)

	Sunday 2nd	Monday 3rd	Tuesday 4th	Wednesday 5th	Thursday 6th	Friday 7th	Saturday 8th
	Cereals, Muesli, Toasts, Jams, Croissants, Fruit, Yoghurt, Juices Tea, Coffee	Cereals, Muesli, Toasts, Jams, Croissants, Fruit, Yoghurt, Juices Tea, Coffee	Cereals, Muesli, Toasts, Jams, Croissants, Fruit, Yoghurt, Juices Tea, Coffee	Cereals, Muesli, Toasts, Jams, Croissants, Fruit, Yoghurt, Juices Tea, Coffee	Cereals, Muesli, Toasts, Jams, Croissants, Fruit, Yoghurt, Juices Tea, Coffee	Cereals, Muesli, Toasts, Jams, Croissants, Fruit, Yoghurt, Juices Tea, Coffee	Cereals, Muesli, Toasts, Jams, Croissants, Fruit, Yoghurt, Juices Tea, Coffee
	Biscuits Coffee's, Tea's & Juice's	Biscuits Coffee's, Tea's & Juice's	Biscuits Coffee's, Tea's & Juice's	Biscuits Coffee's, Tea's & Juice's	Biscuits Coffee's, Tea's & Juice's	Biscuits Coffee's, Tea's & Juice's	
	Homemade Pies, Sausage Rolls, Mixed Quiches Refreshments	Mixed Wraps Fruit Platters Refreshments	Vegi Slices Light Salad Refreshments	Bread Rolls & a Variety Of Meats, Spreads & Salads Fruit Platters Refreshments	Mixed Wraps Fruit Platters Refreshments	Some Lunch Eat In Some T/Away Vegi	
ARRIVAL	Biscuits Refreshments	Biscuits Refreshments	Biscuits Refreshments	Biscuits Refreshments	Biscuits Refreshments	Biscuits Refreshments	
BBQ Kababs Kangaroo Lamb Seafood Vegetable With Light Salad Cake Selection Tea & Coffee	Thai Vegetable & Meat Curries With Steamed Jasmine Rice Thai Papaya Salad Lemon & Lime Tart & Cream	Pasta Night With Vegi Tortellini, Spaghetti Bolognaise & Garlic Bread With A Light Salad Apple Crumble & Ice Cream	Roast Chicken & Baked Potato's With Seasonal Vegetables Bread & Butter Pudding	Indian Dahl, Butter Chicken, Madras Beef, Mix Veg & Chic Pea Basmati Rice Cheese Cake Cream	BBQ Steaks, Sausage, Chicken Wings, Sweet Corn, Potato's In Foil With Sour Cream & Salads Chocolate Cake Fresh Cream	Chinese Spring Rolls, S/S Pork, Beef Veg Oyster Sauce & Rice Pavlova & Fruit Salad	

Note: Meals usually served buffet style. After completion of meal, plates returned to service area by guests. Vegetarian and other dietary needs catered for.

WEEK TWO (Tentative Menu)

Sunday 9th Monday 10th Tuesday 11th Wednesday 12th Thursday 13th Friday 14 Saturday 15th

Full Cooked Breakfast Bacon, Eggs, Tomato, Fruit, Yoghurt, Juices Tea, Coffee	Cereals, Muesli, Toasts, Jams, Croissants, Fruit, Yoghurt, Juices Tea, Coffee	Cereals, Muesli, Toasts, Jams, Croissants, Fruit, Yoghurt, Juices Tea, Coffee	Cereals, Muesli, Toasts, Jams, Croissants, Fruit, Yoghurt, Juices Tea, Coffee	Cereals, Muesli, Toasts, Jams, Croissants, Fruit, Yoghurt, Juices Tea, Coffee	Cereals, Muesli, Toasts, Jams, Croissants, Fruit, Yoghurt, Juices Tea, Coffee	Cereals, Muesli, Toasts, Jams, Croissants, Fruit, Yoghurt, Juices Tea, Coffee / Breakfast Bags
Biscuits Coffee's, Tea's & Juice's	Biscuits Coffee's, Tea's & Juice's	Biscuits Coffee's, Tea's & Juice's	Biscuits Coffee's, Tea's & Juice's	Biscuits Coffee's, Tea's & Juice's	Biscuits Coffee Tea's & Juice's	Biscuits Coffee Tea's & Juice's
Bread Rolls & a Variety Of Meats, Spreads & Salads Fruit Platters Refreshments	Mixed Wraps Fruit Platters Refreshments	Hot Dogs with Colslaw Salad Cheese & Mustard Fruit Platters Refreshments	Bread Rolls & a Variety Of Meats, Spreads & Salads Fruit Platters Refreshments	Slices & Quiche Light Salad Fruit Platter Refreshments	Bread Rolls & a Variety Of Meats, Spreads & Salads Fruit Platters Refreshments	Bread Rolls & a Variety Of Meats, Spreads & Salads Fruit Platters Refreshments
Biscuits Refreshments	Biscuits Refreshments	Biscuits Refreshments	Biscuits Refreshments	Biscuits Refreshments	Biscuits Refreshments	Biscuits Refreshments
Home Made Meat & Vegetarian Lasagne With Light salad	Pad Thai, Sate's, BBQ Chicken Thai Salad Baked Peach With Crumble To & Ice-Cream	Braised Beef & Red Wine Casserole With Potato & Vegetables Pecan Tart	Pasta Night With Vegi Tortellini, Spaghetti Bolognaise & Garlic Bread With A Light Salad Apple Crumble & Ice Cream	Roast Pork & Baked Potato with Seasonal Vegetables Pavlova & Fruit Salad	BBQ Steaks, Sausage, Chicken Wings, Sweet Corn, Potato's In Foil With Sour Cream & Salads Chocolate Broulee	Lunch bags

Note: Meals usually served buffet style. After completion of meal, plates returned to service area by guests. Vegetarian and other dietary needs catered for.