apts.ac.uk

Academy for PhD Training in Statistics

WEEK 4: UNIVERSITY OF OXFORD

9-13 SEPTEMBER 2024

Welcome to Oxford!

Workshop registration: Registration for the APTS week will take place between 11:30–13:00 on Monday 9 September 2024 in the Ground floor social area, Department of Statistics (see Figure 1). A map showing key locations for the week can be found on page 6 of this booklet.

Figure 1: Department of Statistic (left) and Balliol College (right).

Please note: The front door to the department will usually be on automatic open but, due to the St Giles' Fair taking place on Monday 9 and Tuesday 10 September, the front door will be kept locked as an extra safety measure. **Please can we ask you to use the back door to enter and leave the department on these days**. The following arrangements have been put in place for entering the building on Monday and Tuesday:

Monday 9 September

11.30 am – 1pm	Registration – Please come to the back door accessed via Blackhall Road. There will be clear signage outside the building and a member of staff will be waiting to let you in.
1.45 pm – 2.00 pm	Entrance via the back door – a member of staff will be on hand to let you in.

Tuesday 10 September

8.45 am – 9.00 am	Entrance via the back door – a member of staff will be on hand to let you in.
1.45 pm – 2.00 pm	Entrance via the back door – a member of staff will be on hand to let you in.

From Wednesday onwards, the front door will be open as usual.

You will receive your name badge from the registration desk. Please wear your badge at all times. This will help with security and also help you identify fellow participants.

Check-in / Luggage: Check-in for residential delegates is from 2.00 pm at the Lodge, Balliol College (see Figure 1) on Monday 9 September. Check-out is by 10 am on Friday 13 September. A room will be available at the Department of Statistics to store your luggage on Monday afternoon and until lunchtime on Friday.

IT: You are encouraged, if possible, to bring a laptop with R installed for taking part in the Statistical Modelling lab sessions. The Statistical Machine Learning module recommends the use of Github Codespaces to run R for the labs. Please follow the guide at **https://www.louisaslett.com/StatML/setup/** before the first lab to ensure you are setup with the necessary account. See notes on pages 13 for further details.

There will also be a computer lab with 50 desktop computers available for the practical sessions.

Wi-fi: We strongly advise that you set up Eduroam beforehand using **https://cat.eduroam.org/.** Alternatively, OWL accounts (a central wireless service for both University Members and Visitors) will be available at registration but please note you will need more than one OWL account if you plan to use multiple devices simultaneously. There will also be wifi provision for residential students staying at Balliol College. Wifi is available in the bedrooms via the OWL network and you will be given your code when you check-in.

Your room: Residential participants have rooms booked at Balliol College, Broad Street, Oxford OX1 3BJ. Unfortunately, there is no parking available at the College (if you have special requirements then please contact the local organizer) however the Pear Tree and Water Eaton Park and rides are very convenient as the buses stop near the College: **Park and Ride sites – Oxford City Council.** The College is a 15-minute walk from the railway station (taxis are available outside Oxford train station) and a 10-minute walk from the bus and coach station at Gloucester Green.

All residential delegates have been booked single room accommodation at Balliol College for 4 nights from Monday 9 September until Friday 13 September. Check-in is from 2.00 pm on Monday 9 September at the Lodge. Check-out is by 10:00 am on Friday 13 September. The Porter will give you a bedroom key and a code number to open the door to your staircase. All rooms are either ensuite or have washbasins. Linen and bedding is provided and tea and coffee making facilities are in all the rooms. Breakfast will be served in the Grand Hall, Broad Street, between 8:00 am and 9:00 am. You will have the option of a full English breakfast.

Balliol operates a non-smoking policy in all buildings and enclosed areas including outside entrances to buildings. Smoking is permitted in the Quads please use the outside receptacles provided.
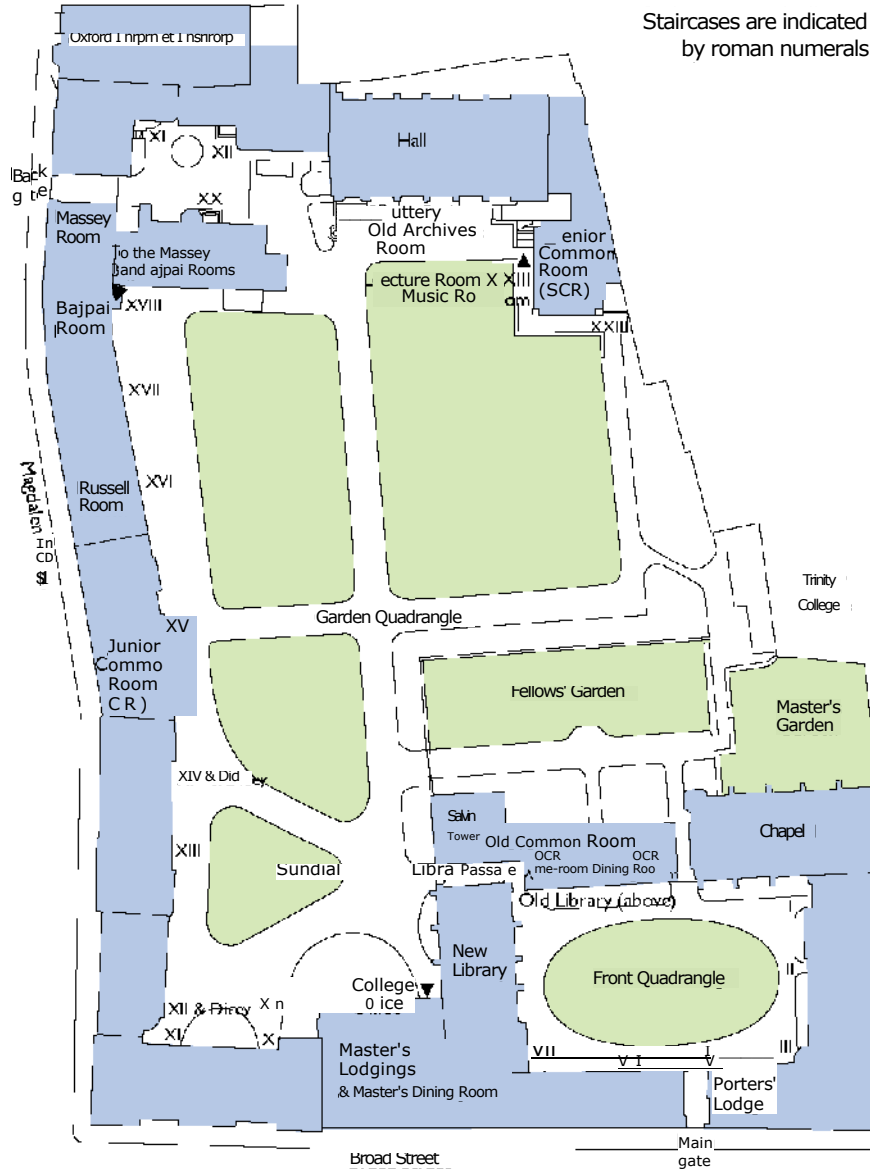
Meals: All meals, with the exception of the Academy Dinner and the Social Event will take place in the main Dining Hall of Balliol College. Breakfast will be served from 08:00–09:00, lunch from 13:00–14:00 and dinner from 18:00–19:30 on Monday and Wednesday. There will be a Social Event at the Cherwell Boathouse on Tuesday from 6.00 pm. The Academy Dinner will be held at Linacre College at 19:00 on Thursday.

Have a great week!

BALLIOL COLLEGE
Oxford OX1 3BJ

Staircases are indicated by roman numerals

Oxford Thrprn et Thshrorp

IX XI
XII
Hall

Back g te

XX

Massey Room

uttery
Old Archives Room

enior Common Room (SCR)

To the Massey band ajpai Rooms

ecture Room X X
Music Ro
om

XXII

Bajpai Room

XVIII

XXII

XVII

Magdalen

XVI

Russell Room

In CD

Trinity College

XV

Garden Quadrangle

Junior Commo Room (R)

Fellows' Garden

Master's Garden

XIV & Did ey

XIII

Salvin Tower

Old Common Room

Chapel

OCR
me-room

OCR
Dining Roo

Sundial

Libra Passa e

Old Library (above)

New Library

Front Quadrangle

College 0 ice

XII & Parcy

X n

Master's Lodgings
& Master's Dining Room

VII

V

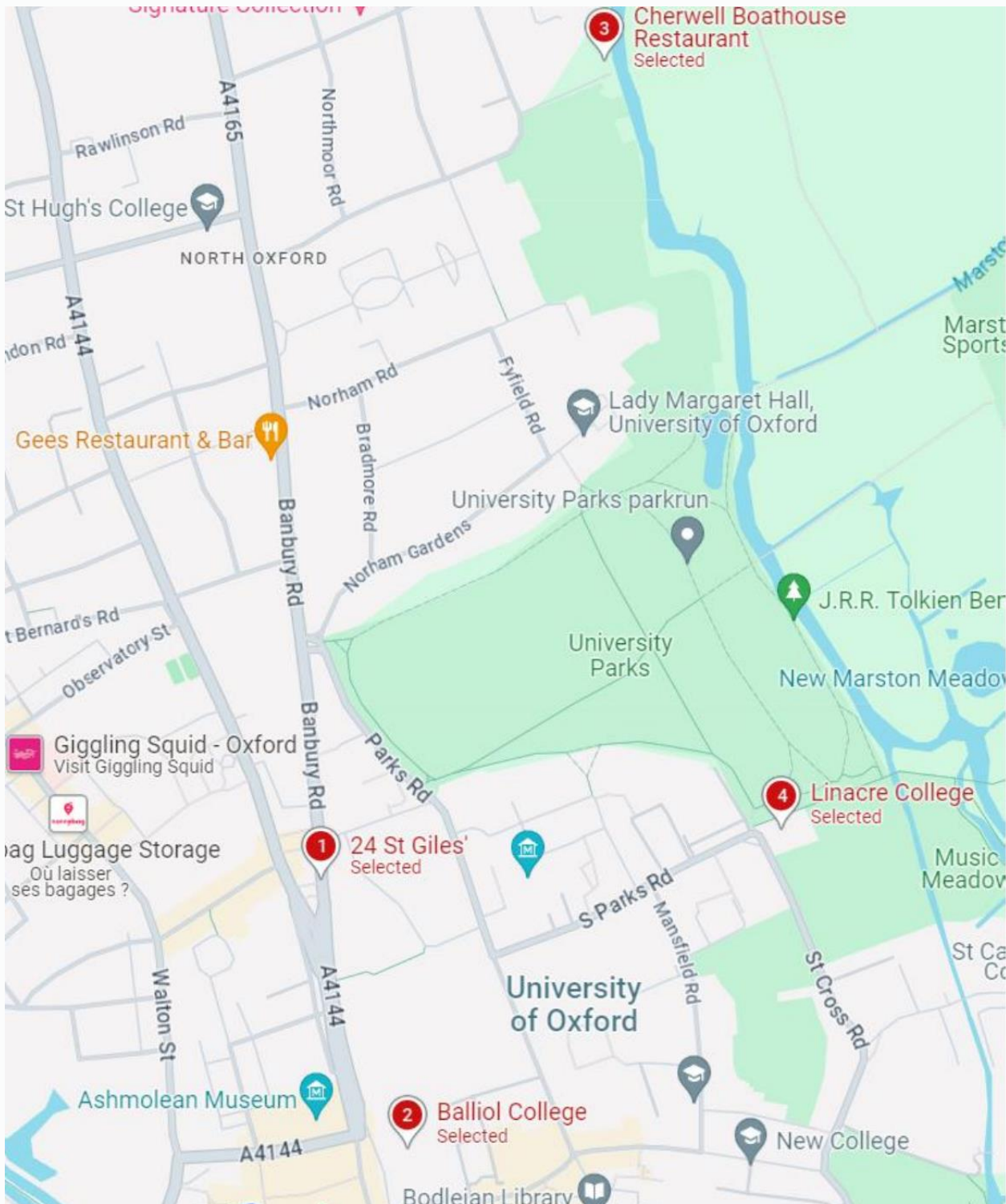III

XI
X

Porters' Lodge

Broad Street

Main gate

For disabled access, please see www.balliol.ox.ac.ulddisability

For a map of Oxford, other colleges and the University buildings, please see www.ox.ac.uldvisitors/map

N

APTS WEEK OXFORD 2024

1   Department of Statistics, 24-29 St Giles, Oxford OX1 3LB

2   Balliol College Porter's Lodge, Broad Street, Oxford OX1 3BJ

3   The Cherwell Boathouse, Bardwell Road, Oxford OX2 6ST

4   Linacre College, St Cross Road, Oxford OX1 3JA

APTS Timetable

	Monday 9th Sept	Tuesday 10th Sept	Wednesday 11th Sept	Thursday 12th Sept	Friday 13th Sept
08:00 – 09:00		Breakfast			
09:30 – 11:00		Lecture CI1	CI (Lab 1)	Lecture SML5	CI (Lab 2)
11:00 – 11:30		Tea & Coffee Break			
11:30 – 13:00	Registration	Lecture CI2	Lecture SML4	SML (Lab 2)	Lecture CI5
13:00 – 14:00	Lunch				
14:15 – 15:45	Lecture SML1	Lecture SML3	Free afternoon	Lecture CI3	
15:45 – 16:15	Tea & Coffee Break			Tea & Coffee	
16:15 – 17:45	Lecture SML2	SML (Lab 1)		Lecture CI4	
17:45 – 18:30	Free Time	Social Event: Punting followed by BBQ at The Cherwell Boathouse **18.15 to 22:00**		Free Time	
18:00 – 19:30	Dinner		Dinner		
Evening	Free Evening		Free Evening	Academy Dinner at Linacre College **19.00 (please arrive by 18:45)**	

Local information

Location of lectures: All APTS lectures will take place in the Large Lecture Theatre (LG.01) at the Department of Statistics, 24–29 St Giles', Oxford OX1 3LB. The practical sessions will take place in the Large Lecture Theatre (LG.01), IT Teaching Suite (LG.02) and Small Lecture room (LG.03) at the Department of Statistics.

Tea and coffee breaks: Refreshments will be served in the ground floor social area at the Department of Statistics.

Figure 2: Cherwell Boathouse (left) and Linacre College (right).

Evening events: Tuesday evening will start with punting at The Cherwell Boathouse, Bardwell Road, Oxford OX2 6ST, followed by a BBQ in their marquee. There will be a walking bus to the Cherwell Boathouse leaving the Department of Statistics at 6.00 pm. The Academy Dinner will take place at Linacre College at 7.00 pm on Thursday 12 September (see Figure 2). **Dress code**: Smart casual. Linacre College, St Cross Road, Oxford OX1 3JA is approx. 10 minutes' walk from the Department.

Sports facilities: Oxford University Sport **https://www.sport.ox.ac.uk/** offer a day pass structure for non-members for the Pulse Gym (£9 for one session) and Swim (£10 per session). They also have a number of fitness classes (£10 per class) available for non-members. Please note, prebooking is required for swimming, gym and fitness classes. You will need to register first, using the link below:
https://oxforduniversity.leisurecloud.net/JoinAtHome/landing.aspx?AppId=JHCL.
The Sports centre is located on Iffley Road, OX4 1EQ approx. 25 minutes' walk from Balliol College.

Things to do in Oxford: Oxford is a beautiful and historic city with many local attractions. Some suggestions on activities of interest are given below:

University of Oxford Colleges: Oxford University is the oldest University in the English-speaking world. Countless famous figures and great minds have studied here and you may wish to explore some of the colleges that they were a part of. There are 38 independent colleges, many of whom open their doors to visitors at least a few hours every day. Many are free to visit, but some charge a small fee. Further details can be found at the link below: **https://www.ox.ac.uk/visitors/ visiting-oxford/visiting-the-colleges.**

Museums, Libraries and places of interest: Oxford has a wealth of museums such as the Ashmolean Museum and the Oxford University Museum of Natural History, places of interest such as the Sheldonian Theatre as well as many tranquil areas to get away from the hustle and bustle, such as Christ Church Meadow and University Parks. For further details please see the link below: **http://www.ox.ac.uk/visitors/visiting-oxford/visiting-museums-libraries-places.**
TV & Film Locations Oxford is a great favourite of the film industry and has been featured as the backdrop for many TV series and films such as Inspector Morse and Harry Potter. Further details can be found at: **https://www.experienceoxfordshire.org/things-to-do/films-in-oxford-oxfordshire/**

Other places of interest There are many neighbouring tourist attractions that are easily accessible from Oxford such as Blenheim Palace and Bicester Village. For further details please see: **https://www.experienceoxfordshire.org/oxfordshire-daycation/**

Emergency details
In Office Hours

Medical Assistance: Please speak to Reception at the Department of Statistics if you need any medical assistance.

Messages: The telephone number for colleagues or family to leave an urgent message for you during office hours is +44 (0)1865 272860 or +44 (0)1865 281536.

Fire Procedures: The fire alarm is tested weekly, usually on Thursday mornings around 9:00 am. There is no need to evacuate the building then. If the fire alarm is sounded at any other time, you must evacuate the building. If you discover a fire, set the fire alarm off using the nearest red fire call point. In the event of the fire alarm sounding, evacuate the building safely and quickly. Do not use the lift. The assembly point is outside the Department of Physics, just across Keble Road. Full details of safety procedures will be provided at registration.

Out of Office Hours

Medical Assistance: Contact the Duty Porter of Balliol College at the Lodge or telephone +44 (0)1865 277777. If no reply is received, the porter may be on a security patrol. Ring for an ambulance on 999 or 112 asking them to come to the main college entrance on Broad Street. Having made the call, go to the main entrance yourself reporting the emergency to the Porter on duty and wait for the ambulance to arrive.

Messages: The Lodge, Broad Street, is open 24 hours and the Duty Porter will be glad to assist with all your queries. The Lodge can be contacted on +44 (0)1865 277777 or porter@balliol.ox.ac.uk

Fire Procedures: Upon check-in at Balliol College, you will be given an information sheet giving details of emergency procedures.

Causal Inference

Module leader: V Didelez and R Evans

Aims: Causal inference deals with investigating the effect(s) of (typically hypothetical) interventions from (typically 'imperfect') data; often the data is observational, i.e. not obtained from a 'perfect' randomised experiment. The aim of the course is to introduce the fundamental principles, concepts and basic methods of causal inference, including an outlook on current challenges and recent developments.

Learning outcomes: The participants will be able to distinguish research questions that are causal, as opposed to descriptive or predictive, in nature. They will be able to formalise research questions as causal estimands, state and evaluate the required assumptions in a variety of common settings. They will further be able to use causal diagrams to identify potential sources of bias and how to mitigate these. Finally, they will be able to carry out a basic causal analysis with standard software.

Prerequisites: Familiarity with regression models, multivariate distributions and their properties. It will be helpful to revisit, prior to the course, the general notions of conditional distributions and conditional (in)dependence. Some knowledge of methods of estimation beyond maximum likelihood, such as estimating equations will be useful.

Preliminary reading:

- [Didelez](). Causal Reasoning and Inference in Epidemiology. In: Handbook of Epidemiology, 3rd edition, eds. Wolfgang Ahrens & Iris Pigeot, Springer, 2024.

- Richardson and Robins. Single World Intervention Graphs: a primer. In Second UAI workshop on causal structure learning, Bellevue, Washington. 2013.

- Daniel, Cousens, De Stavola, Kenward and Sterne. Methods for dealing with time-dependent confounding. Statistics in Medicine, 32(9), pp.1584-1618, 2013.

Further reading:

- Didelez. Causal concepts and graphical models. In: Handbook of Graphical Models (eds. Maathuis, Drton, Lauritzen, Wainwright), Chapman Hall/CRC, 2018.

- Hernán and Robins. Causal Inference: What If, Chapman Hall/CRC, 2020 - online.

- Lauritzen. Causal inference from graphical models. In: Complex Stochastic Systems (eds. Barndorff-Nielsen, Cox, Klüppelberg), Chapman Hall/CRC, 2001.

- Peters, Janzing and Schölkopf. Elements of Causal Inference, Cambridge, 2017.

- Pearl. Causality: Models, Reasoning, and Inference, (3rd ed.) Cambridge, 2013.

Topics:

1. Basic concepts: causation vs association, randomisation, do-interventions, potential outcomes, causal estimands, identifiability (key assumptions), confounding and selection bias, target trial emulation.

2. Causal diagrams: conditional independence, directed acyclic graphs (DAGs), d-separation, causal Markov condition, backdoor criterion, confounding and selection bias revisited, SWIGs.

3. Estimating the causal effect of a point treatment: g-methods (propensity scores, inverse-probability weighting, g-formula, g-estimation), optimal adjustment, checking assumptions (positivity, negative controls), double-robustness and double-machine learning, survival outcomes.

4. Multiple or sequential treatments: causal interpretation (and misconceptions) of multiple regression, time-dependent confounding, marginal structural models.

5. Outlook on further topics: causal discovery. If there is time, we will discuss mediation, direct and indirect effects, and instrumental variables.

Assessment: A worksheet with questions based on the core topics above.

Statistical Machine Learning

Module leader: L Aslett

Please see the full Module Specifications *for background information relating to all of the APTS modules, including how to interpret the information below.*

Aims: This module introduces students to modern supervised machine learning methodology and practice, with an emphasis on statistical and probabilistic approaches in the field. The course seeks to balance theory, methods and application, providing an introduction with firm foundations that is accessible to those working on applications and seeking to employ best practice. There will be exploration of some key software tools which have facilitated the growth in the use of these methods across a broad spectrum of applications and an emphasis on how to carefully assess machine learning models.

Learning Outcomes: Students following this module will gain a broad view of the supervised statistical machine learning landscape, including some of the main theoretical and methodological foundations. They will be able to appraise different machine learning methods and reason about their use. In particular, students completing the course will gain an understanding of how to practically apply these techniques, with an ability to critically evaluate the performance of their models. Students will also have an insight into the extensive software libraries available today and their use to construct a full machine learning pipeline.

Prerequisites

To undertake this module, students should have:

- at least one undergraduate level course in probability and in statistics;

- standard undergraduate level knowledge of linear algebra and calculus;

- solid grasp of statistical computing in R;

- knowledge of statistical modelling, including regression modelling (eg. APTS Statistical Modelling course);

- some basic understanding of optimisation methods beneficial, but not essential.

As preparatory reading, the enthusiastic student may choose to browse *An Introduction to Statistical Learning* (James et al., 2013) (freely and legally available online), which covers some of the topics of the course at a more elementary and descriptive level.

Textbooks at roughly the level of the course include:

- The Elements of Statistical Learning (Friedman, Tibshirani, and Hastie) • Pattern Recognition and Machine Learning (Bishop)

- Machine Learning: A Probabilistic Perspective (Murphy)

- Deep Learning (Goodfellow, Bengio and Courville)

Topics

- Formulation of supervised learning for regression and classification (scoring/probabilistic, decision boundaries, generative/discriminative), loss functions and basic decision theory;

- Theory of model capacity, complexity and bias-variance decomposition;

- Curse of dimensionality;

- Overview of some key modelling methodologies (eg logistic regression, local methods, kernel smoothing, trees, boosting, bagging, forests);

- Model selection, ensembles, tuning and super-learning;

- Evaluation of model performance, validation and calibration and their reporting in applications;

- Reproducibility;

- Coverage of some key software frameworks for applying machine learning in the real world.

Assessment: An exercise set by the module leader involving practical use of some of the machine learning methods covered and critical evaluation of their performance.

Notes