

THE UNIVERSITY OF
SYDNEY

Advanced Machine Learning

(COMP 5328)

School of Computer Science
Introduction to Machine Learning Problems

Tongliang Liu



THE UNIVERSITY OF
SYDNEY

Acknowledgement of Country

We recognize and pay respect to the Elders and communities-past, present, and emerging - of the lands that the University of Sydney's campuses stand on. For thousands of years, they have shared and exchanged knowledges across innumerable generations for the benefit of all.



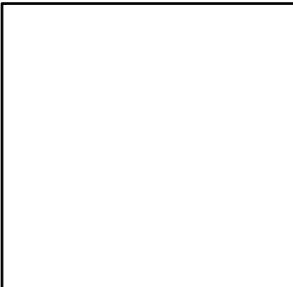
THE UNIVERSITY OF
SYDNEY

WHS INDUCTION

School of Computer Science



THE UNIVERSITY OF
SYDNEY



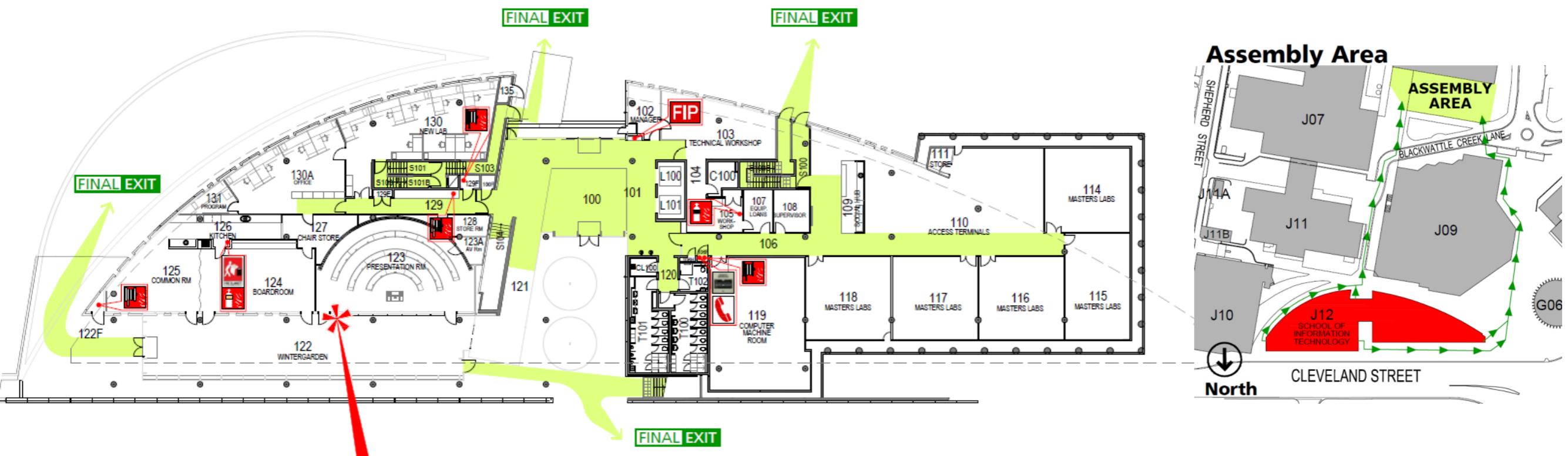
General Housekeeping – Use of Labs

- › Keep work area clean and orderly
- › Remove trip hazards around desk area
- › No food and drink near machines
- › No smoking permitted within University buildings
- › Do not unplug or move equipment without permission



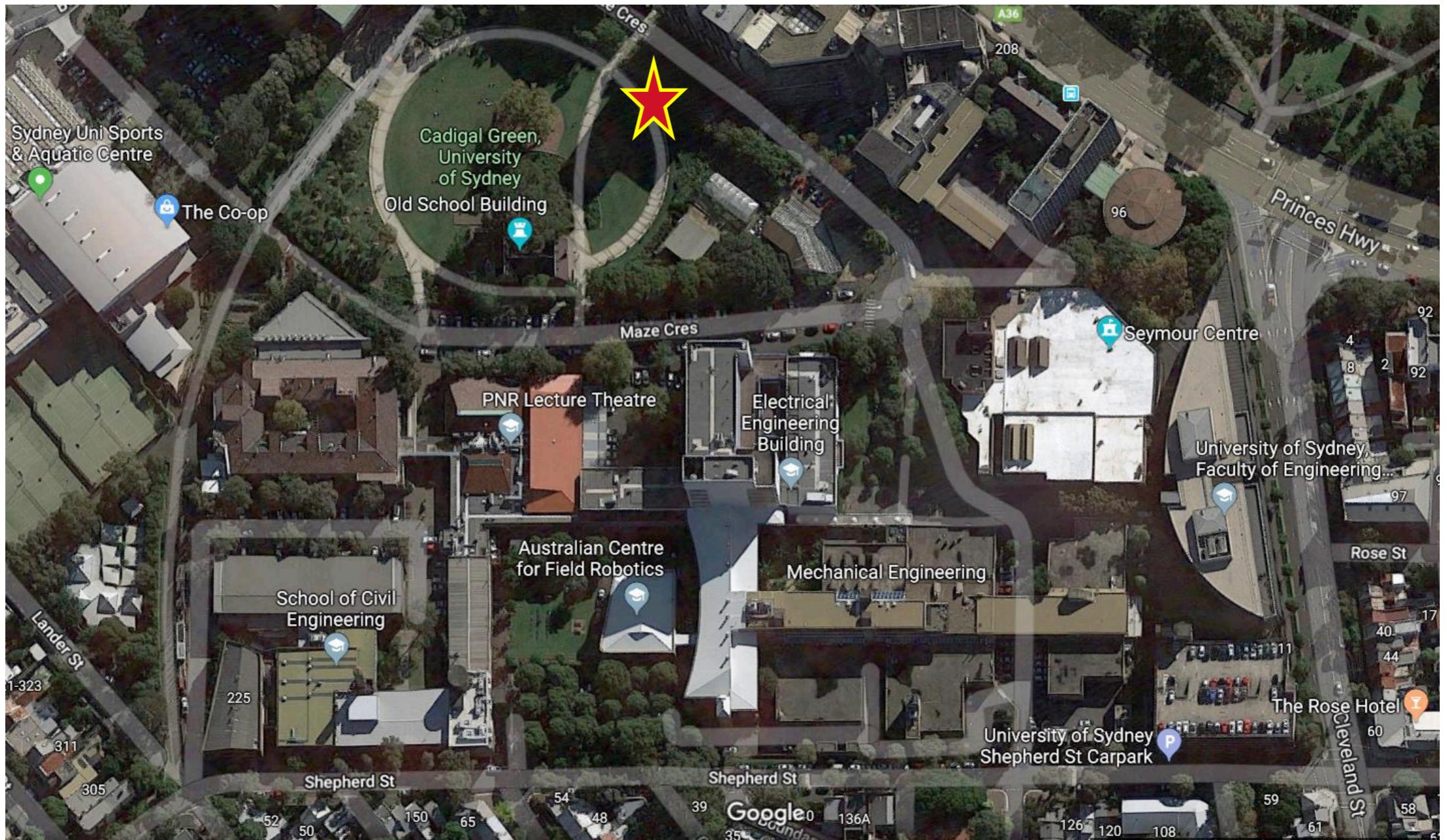


WHERE IS YOUR CLOSEST SAFE EXIT ?





Follow directions of wardens



› If a person is seriously ill/injured:

1. **call an ambulance 000**
2. **notify the closest Nominated First Aid Officer**

If unconscious— send for Automated External Defibrillator (AED)
AED [locations](#).

NEAREST to SIT Building (J12)

- Electrical Engineering Building, L2 (ground) near lifts
- Seymour Centre, left of box office
- Carried by all Security Patrol vehicles



3. **call Security - 9351-3333**
4. **Facilitate the arrival of Ambulance Staff (via Security)**



Nearest Medical Facility

University Health Service in Level 3, Wentworth Building

First Aid kit – SCS Building (J12)

kitchen area adjacent to Lab 110

CHIEF WARDEN

Name: Greg Ryan
Mobile: : +61 411 406 322

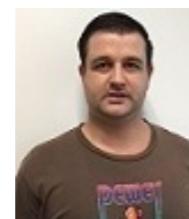


FIRST AID OFFICERS

Name: Julia Ashworth
Location: 2E 237
Phone: 8627 9058



Name: Will Calleja
Location: 1 West 103
Phone: 9036 9706
0422 001 964



Name: Muhammad
Sajjad Akbar
Location: 2E-237
Phone: 0470 257 750

**Orally REPORT all
INCIDENTS
& HAZARDS
to your SUPERVISOR**

OR the admin team

Undergraduates: cs.ug@sydney.edu.au

Postgraduate HDR: cs.research@sydney.edu.au

Postgraduate Coursework:
cs.pg@sydney.edu.au

School Manager: Priyanka Magotra
8627 4295



Places

- Lecture:
 - Every Friday, 5 to 7pm, Wallace Lecture Theatre 200
- Tutorials:
 - Check your timetable
 - Onsite
 - Go to the room you are scheduled for
 - If for some reason you miss it, you can attend a later lab session *if there is space and the tutor agrees*, but ask the tutor before taking a seat
- Do not miss classes, except for illness, emergencies, etc
- Get help from staff if you feel you are falling behind



Team

- Lecturer:
 - Dr Tongliang Liu (coordinator)
- Teaching Assistant:
 - Ziming Hong
- Tutors:
 - Li He
 - Ziming Hong
 - Anjin Liu
 - Vincent Qu
 - Jun Wang
 - Yuhao Wu
 - Haimei Zhao



Resources

- Canvas
 - Login using Unikey and password
 - Link to the unit website: <https://canvas.sydney.edu.au/courses/66429>
 - list of learning outcomes, etc
 - Official schedule
 - Copies of slides
 - Tutorial instructions
 - Assignment instructions
 - Lecture videos
 - We intend to record the lectures, but the technology is not reliable
 - ***Submit official assignments in Canvas;***
 - See your grades; etc
 - Ed discussion board

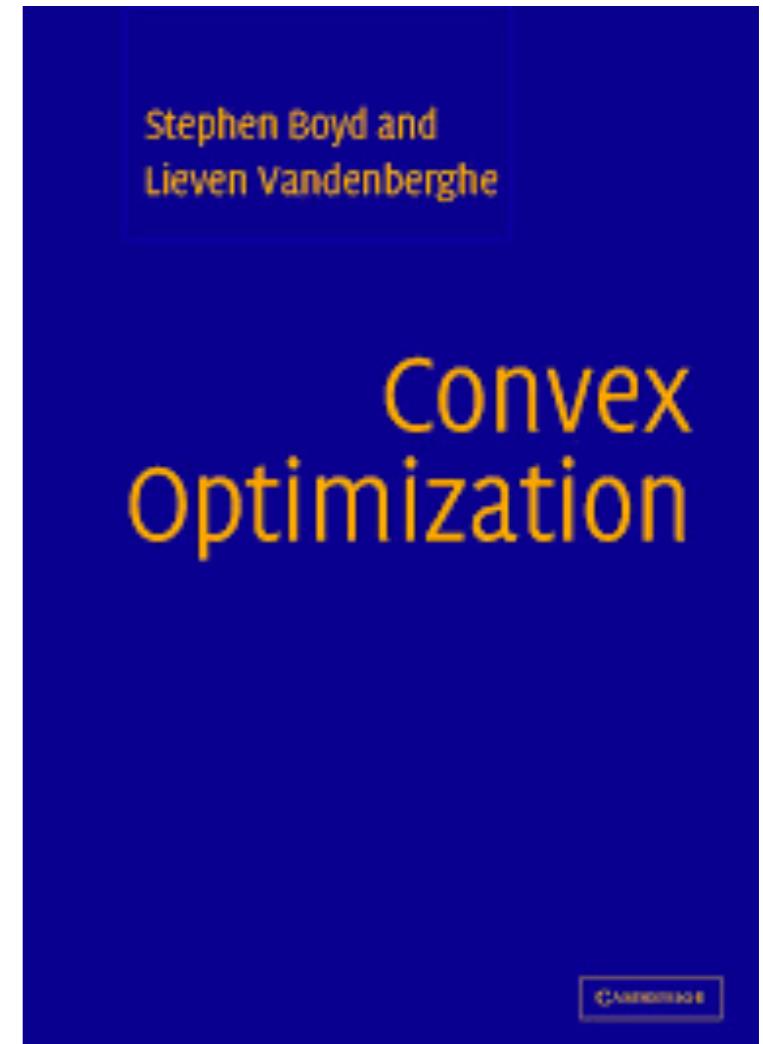
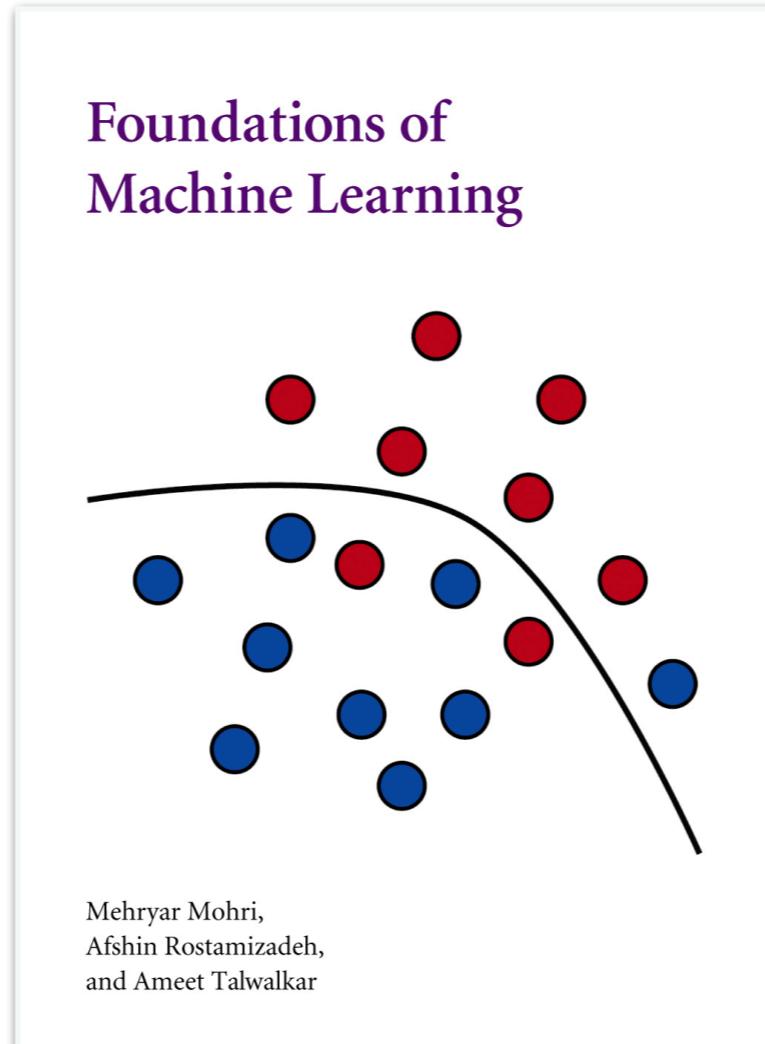
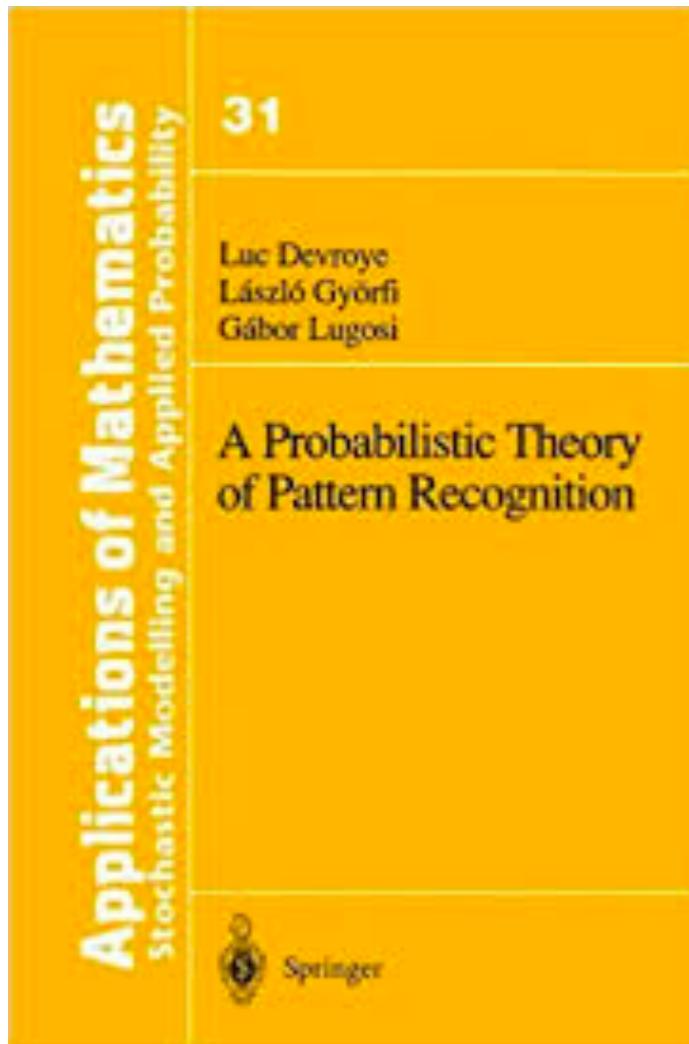


Resources

- Ed discussion board:
- **General** (student helpline 1800 793 864 or contact the admin team) and **Technical**
- Mail technical questions to tutors
- Consultation: 2-3pm Friday (office 315 at J12; appointment needed at least one day in advance)



Best books available





Assessment overview

- Quiz: 0%
 - Week 4
 - Individual
 - Contents in the first three weeks
 - Lower than 60%
 - The census date is on 1 September 2025
- Assignment 1: 25%
 - Due: Week 9 (9/10), 11:59pm
 - Groups of 3 or 4 students
 - Method comparison and analysis for feature noise
- Assignment 2: 25%
 - Due: Week 13 (6/11), 11:59pm
 - Groups of 3 or 4 students
 - Classification with noisy labels



Late submissions

- Consistent penalty of 5% per day late
 - e.g.,
 - a) A “good” assignment that would normally get 9/10, and is 2 days late, loses 10% of the full 10 marks, ie, new mark = 8/10.
 - b) An average assignment, that would normally get 5/10, that is 5 days late, loses 25% of the full 10 marks, ie new mark = 2.5/10
 - Assignments more than 5 days late get 0.



Assessment overview

In order to pass the course, the School requires at least 40% in the written final exam, at least 40% in the other assessment components together and an overall final mark of 50 or more. This means that students who score less than 40% in the final exam will fail the course regardless of their marks during the semester.



Expectations

- Students attend scheduled classes, and devote an extra 6-9 hrs per week
 - doing assessments
 - preparing and reviewing for classes
 - revising and integrating the ideas
 - practice and self-assess
- Students are responsible learners
 - Participate in classes, constructively
 - Respect for one another (criticise ideas, not people)
 - Humility: none of us knows it all; each of us knows valuable things
 - Check Canvas site at least once a week!
 - Notify academics whenever there are difficulties
 - Notify group partners honestly and promptly about difficulties
 - Communicate respectfully and efficiently with your group partners.



Topics

Week	Lecture	Tutorial
1	Introduction to ML Problems	No tutorial
2	Loss Functions and Convex Optimisation	Tutorial 1 (take home)
3	Hypothesis Complexity and Generalisation	Tutorial 2
4	Dictionary Learning and NMF	Quiz
5	Sparse Coding and Regularisation	Tutorial 3
6	Learning with Noisy Data	Tutorial 4
7	Domain Adaptation and Transfer Learning	Tutorial 5
8	Learning with Noisy Data II: Label Noise	Tutorial 6
9	Reinforcement Learning	Tutorial 7
10	Causal Inference	Tutorial 8
11	Multi-task Learning	Tutorial 9
12	Guest Lecturer (Google)	Tutorial 10
13	Review	Tutorial 11



THE UNIVERSITY OF
SYDNEY

Assumed knowledge

- Linear algebra, calculus
- Basics of probability and statistics
- Programming skills



Labs: Python

- Python is a high-level programming language designed to enforce good coding practices.
- Interactive and very natural to use.
- Extremely versatile and excellent for prototyping.
- Great libraries for machine learning eg. scikit-learn, TensorFlow, Keras, Edward

www.python.org



Special Consideration (University Policy)

- If your performance on assessments is affected by illness or misadventure
- Follow proper bureaucratic procedures
 - Have professional practitioner sign special USYD form
 - Submit application for special consideration online, upload scans
 - Note you have only a quite short deadline for applying
 - <https://www.sydney.edu.au/students/special-consideration.html>
- Also, notify the coordinator and TA by email as soon as *anything begins to go wrong*
- There is a similar process if you need special arrangements eg for religious observance, military service, representative sports



Special Consideration (University Policy)

- <https://www.sydney.edu.au/students/special-consideration.html>

Reason	Deadline
Special consideration Illness, injury or misadventure	<p>No longer than three working days after:</p> <ul style="list-style-type: none">• the original assessment due date• the sitting date of the exam• the date of the missed class or missed placement <p>For example, if your assessment was due on Monday, you have until Thursday, 11.59pm (Sydney local time) to apply.</p>
Special arrangement For an assessment or exam during semester	If you are aware of your essential commitment at the start of the semester, you need to apply at the start of semester. Otherwise, you need to apply as soon as you become aware.
Special arrangement For an assessment or exam in the formal University exam periods	<p>No longer than 14 days after we publish the exam timetable.</p> <p>If you're not aware of the commitment at this time, you'll need to apply within three working days after you know the special commitment date.</p> <p>For example, if on Monday you receive a summons to appear at court, you have until Thursday, 11.59pm (Sydney local time) to apply.</p>



Academic Integrity (University Policy)

- Please read the University policy on Academic Honesty carefully:
- http://sydney.edu.au/elearning/student/EI/academic_honesty.shtml
- All cases of academic dishonesty and plagiarism will be investigated
- There is a new process and a centralised University system and database
- Three types of offences:
 - **Plagiarism** – when you copy from another student, website or other source. This includes copying the whole assignment or only a part of it.
 - **Academic dishonesty** – when you make your work available to another student to copy (the whole assignment or a part of it). There are other examples of academic dishonesty.
 - **Misconduct** - when you engage another person to complete your assignment (or a part of it), for payment or not. This is a **very serious** matter and the Policy requires that your case is forwarded to the University Registrar for investigation.



Academic Integrity (University Policy)

- The penalties are **severe** and include:
 - 1) a permanent record of academic dishonesty, plagiarism and misconduct in the University database and on your student file
 - 2) mark deduction, ranging from 0 for the assignment to Fail for the course
 - 3) expulsion from the University and cancelling of your student visa
- When there is copying between students, note that **both students are penalised** – the student who copies and the student who makes his/her work available for copying
- Advertising or providing contract cheating services is against the law. Penalties can include fines of up to \$100,000 or imprisonment.
- The number of students reported for contract cheating in 2023 rose 111% compared to 2022.



Academic Integrity (University Policy)

Signs you might be dealing with a *contract cheating service*

- You're approached or solicited to use the service in person, or online
- You're asked to pay for the service
- You're asked to share your Unikey or other personal details
- Advertisements for the service mention a specific USYD unit of study
- The service provides a 'guarantee' that you won't get caught
- The service is not acting in an open, transparent way
- You're added to a private group chat by the service
- You're given step-by-step instructions for how to complete an assessment
- The service is aggressive or threatening
- You feel that by using the service you are getting an unfair advantage



Academic Integrity (University Policy)

- We will use the similarity detection software TurnItIn and MOSS to compare your assignments with these of other students (current and previous) and the Internet
 - Turnitin is for text documents: http://www.turnitin.com/en_us/higher-education
 - MOSS is for programming code: <https://theory.stanford.edu/~aiken/moss/>
- These tools are **extremely good!**
 - e.g. MOSS cannot be fooled by changing the names of the variables or changing the order of the conditions in if-else statements
- Examples of plagiarism in programming code:
 - http://www.upenn.edu/academicintegrity/ai_computercode.html



Academic Integrity (University Policy)

- All these are cases of **plagiarism** and **academic dishonesty** we have seen in our school and the student excuses are not acceptable:
 - *Plagiarism means presenting another person's work as one's own work by presenting, copying or reproducing it without appropriate acknowledgement of the source.*
 - *I sat the test and then posted the questions and solutions to my friends whose test was later in the week. I only wanted to help them understand the concepts that are examinable.*
 - *I posted parts of my code on my web page (group discussion forum) because my solution was cool (or I wanted to help them). I didn't expect them to copy it.*
 - *I tried to do the assignment on my own but I had problems with the extension part that I couldn't fix, so I submitted my core part and his extension part. I didn't cheat.*
 - *I finished my assignment but my friend had family problems. I felt sorry for her, so I gave her my assignment as an example. She said she only wanted to have a look and promised not to copy it.*
 - *The test has finished but the tutor hasn't collected the papers yet. I showed my answer to my friend. I didn't expect him to copy it.*
 - *He is my best friend. I had no choice but to let him copy my assignment.*



Academic Integrity (University Policy)

- Plagiarism and any form of academic dishonesty will be dealt with, and the penalties are severe
- We use plagiarism detection systems such as MOSS and TurnItIn that are extremely good. If you cheat, the chances you will be caught are very high.
- If someone asks you to see or copy your assignment, or to complete the assignment instead of them, just say: *I can't do this. This is against the University policy. I will not risk my future by doing this.*

Be smart and don't risk your future by engaging in plagiarism and academic dishonesty, misconduct!



Academic Integrity Practice

Source of help

- Encouraged
- Attribution required
- Not acceptable
- Ask Lecturer/Coordinator

Lecturer	Teaching Assistants / Tutors	Classmates	Online forums/ Online tutors	Students outside course/UoS	Hired coders Tutorial Company outside University	Relatives	Other

- A student needs to gain an understanding of high level knowledge/skills
- A student needs to gain skills to find, evaluate and apply existing knowledge/solutions



Academic Integrity Practice

Types of help

- Encouraged
- Attribution required
- Not acceptable
- Ask Lecturer/Coordinator

Understanding General Concepts	Explained using similar material (not assignment)	Sharing approach/concept to derive assignment solution	Designing code/solution	Implementing code/solution
--------------------------------	---	--	-------------------------	----------------------------

- A student needs to gain an understanding of fundamental knowledge/skills
- It is important to master the knowledge/skills themselves
- Students are encourage to obtain help through relevant teaching material and practices



Self-test

- When is the first assessment due?
- How much work will you be devoting to this unit, each week?
- Do you need to buy a book?
- What do you do if you get sick during semester?
- What help can you use when answering assessments?
- How do you find out about assignment instructions?
- How do you submit your work?
- What is Turnitin?
- What language will you be coding in?



Disability Services

Do you have a disability?

- You may not think of yourself as having a ‘disability’ but the definition under the **Disability Discrimination Act** is broad and includes temporary or chronic medical conditions, physical or sensory disabilities, psychological conditions and learning disabilities.
- The types of disabilities we see include:
 - anxiety, arthritis, asthma, asperger's disorder, ADHD, bipolar disorder, broken bones, cancer, cerebral palsy, chronic fatigue syndrome, crohn's disease, cystic fibrosis, depression, diabetes, dyslexia, epilepsy, hearing impairment, learning disability, mobility impairment, multiple sclerosis, post traumatic stress, schizophrenia , vision impairment, and much more.
- Students needing assistance must register with Disability Services –
 - it is advisable to do this as early as possible.
 - <https://www.sydney.edu.au/students/health-wellbeing/inclusion-and-disability.html>



Other support

- Learning support
 - <http://sydney.edu.au/study/academic-support/learning-support.html>
- International students
 - <http://sydney.edu.au/study/academic-support/support-for-international-students.html>
- Aboriginal and Torres Strait Islanders
 - <http://sydney.edu.au/study/academic-support/aboriginal-and-torres-strait-islander-support.html>
- Student organisation (can represent you in academic appeals etc)
 - <http://srcusyd.net.au/> or <http://www.supra.net.au/>

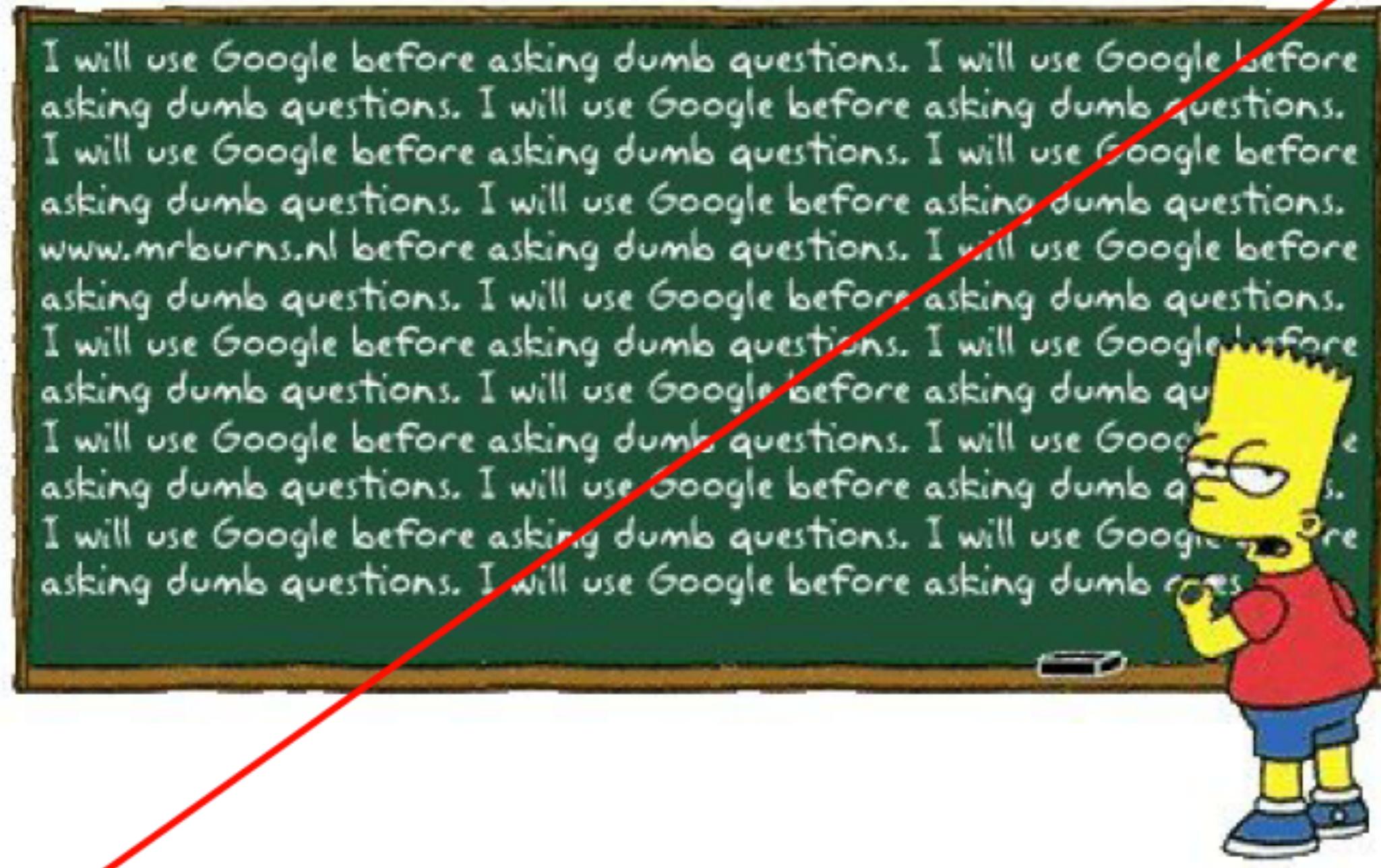


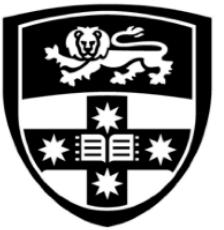
Advice

- Metacognition
 - Pay attention to the learning outcomes
 - Self-check that you are achieving each one
 - Think how each assessment task relates to these
- Time management
 - Watch the due dates
 - Start work early, submit early
- Networking and community-formation
 - Make friends and discuss ideas with them
 - Know your tutor, lecturer, coordinator
 - Keep them informed, especially if you fall behind
 - Don't wait to get help
- Enjoy the learning!



Ask questions sooner!





THE UNIVERSITY OF
SYDNEY

**Let's talk about
some machine learning
problems**



Midjourney



Unitree
Robotics



Google™

Bai du 百度

amazon.com®



SONY

SIEMENS

腾讯
Tencent

IBM

商汤
sense time

Microsoft®

Research

Alibaba Group

Research



HRI Europe
Honda Research Institute

facebook.

YAHOO!

Commonwealth Bank
COMP5328 - Tongliang Liu



39

Rio Tinto



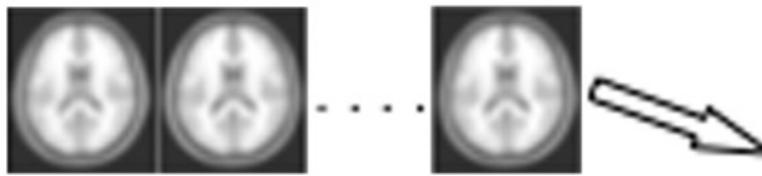
No Free Lunch Theorem

- The “No Free Lunch” theorem states that there is no one model that works best for every problem.
- There are numerous applications. We cannot learn their solutions independently.
- We should know how machine learning algorithms work. How can we improve them?

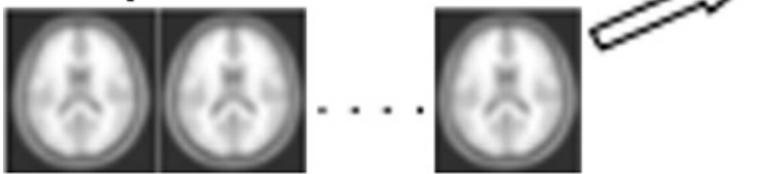
Elements of Machine Learning Algorithms

Input training data

Group 1

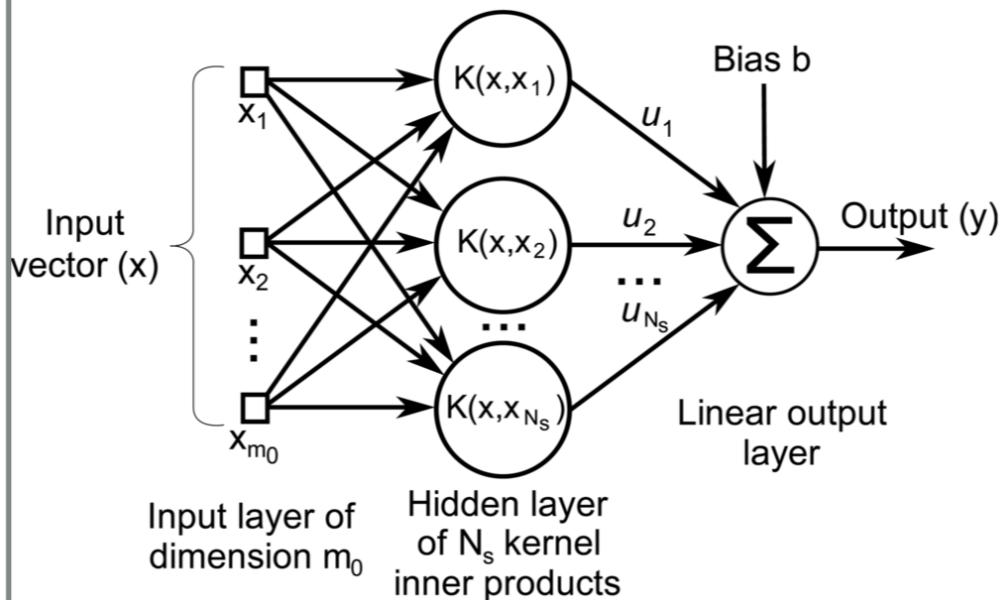


Group 2

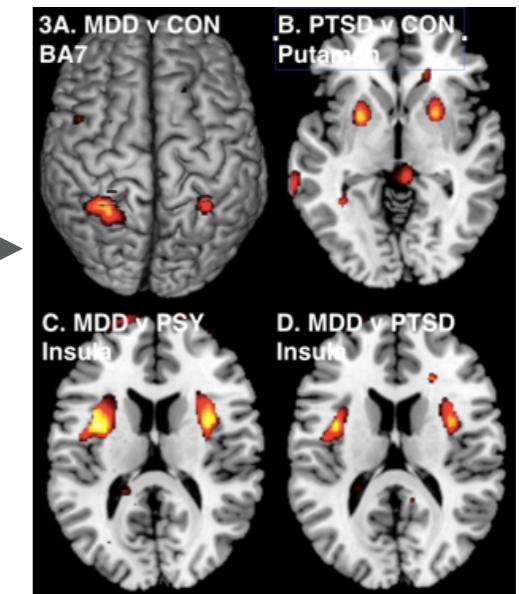


Data

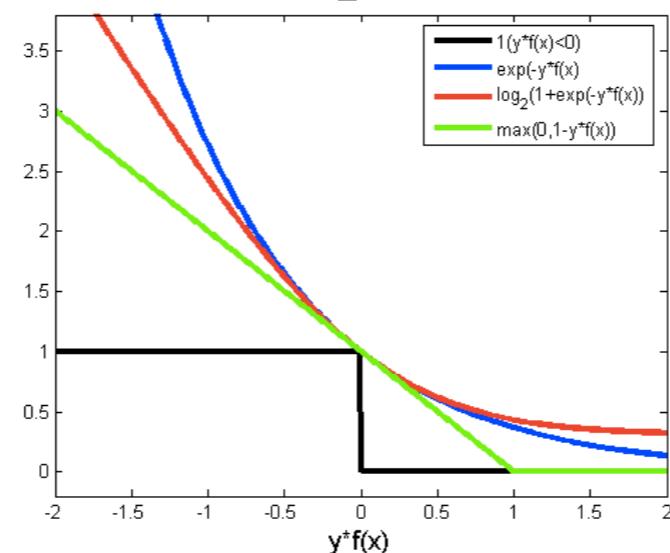
Mathematical Model



Output hypothesis
(Predictions)



Input predefined
hypothesis/function class



Objective function

Optimisation
method



Elements of Machine Learning Algorithms

- I. Input training data
- II. Predefined hypothesis class
- III. Objective function
- IV. Optimisation method
- V. Output hypothesis



THE UNIVERSITY OF
SYDNEY

What is Machine Learning? (COMP5318)

Informally: Making predictions from data

Formally: The construction of a statistical model that is an underlying distribution from which the data is drawn from.



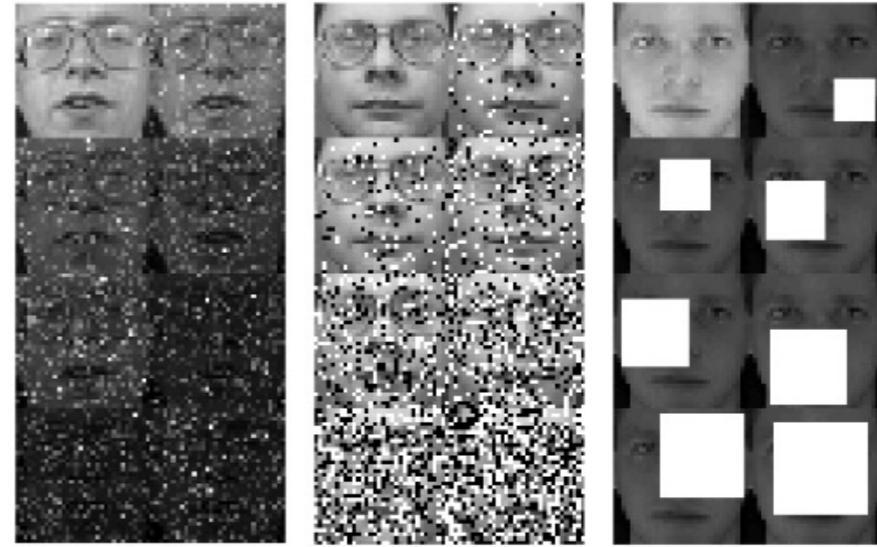
What is Machine Learning? (COMP5328)

- Input training data: $S = \{(X_1, Y_1), \dots, (X_n, Y_n)\}$
- Input predefined hypothesis class: $H = \{h_1, h_2, \dots\}$
- The objective function and optimisation method together make up a mapping: $\mathcal{A} : (\mathcal{X} \times \mathcal{Y})^n \rightarrow H$
- Output hypothesis: h_S
- The overall learning algorithm is a mapping:

$$\mathcal{A} : S \in (\mathcal{X} \times \mathcal{Y})^n \mapsto h_S \in H$$

I. Input data: feature noise

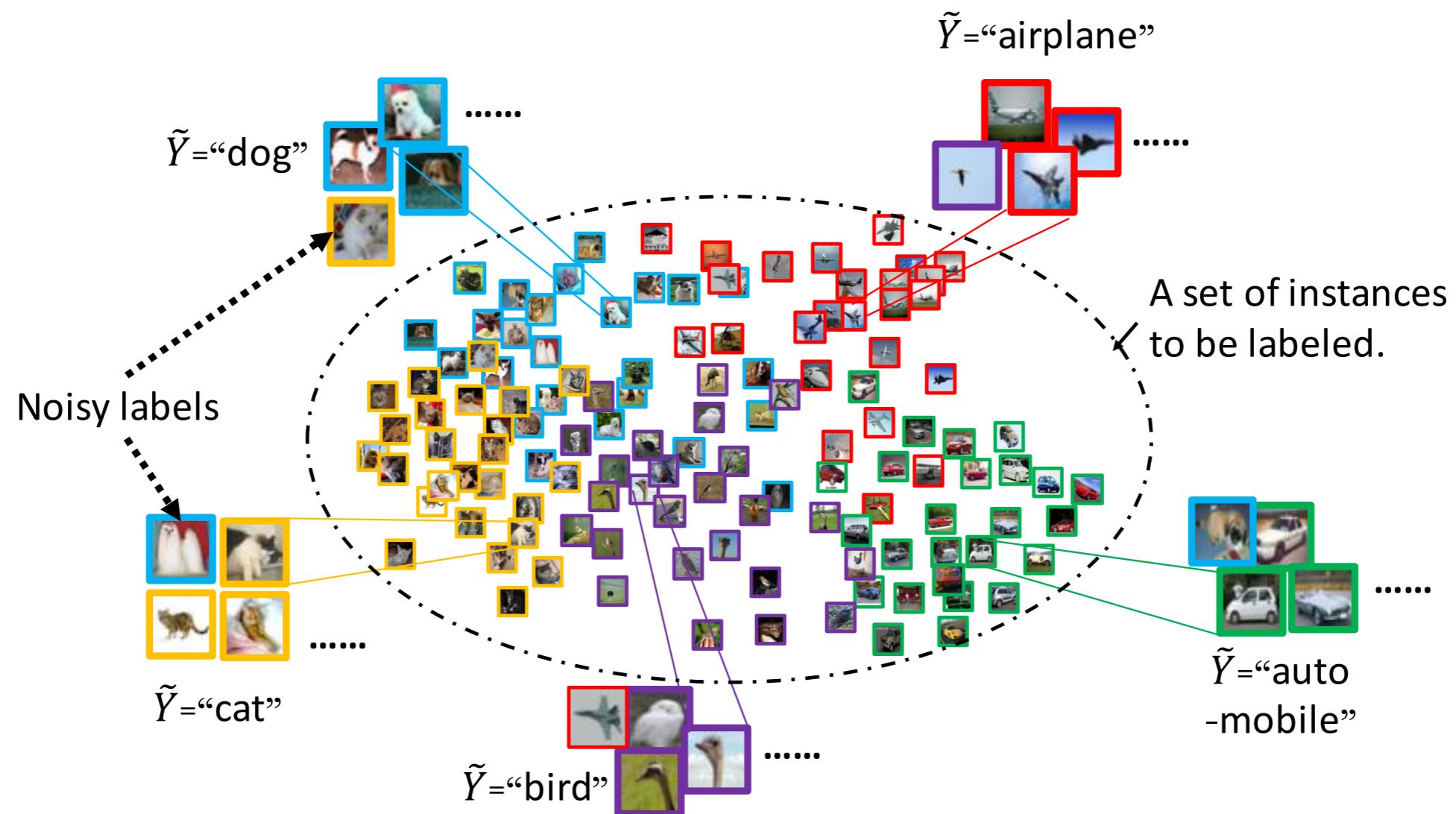
The input data are contaminated by noise. How can we deal?



Guan, Naiyang, et al. "Truncated Cauchy Non-negative Matrix Factorization for Robust Subspace Learning." IEEE Transactions on Pattern Analysis and Machine Intelligence (2017).

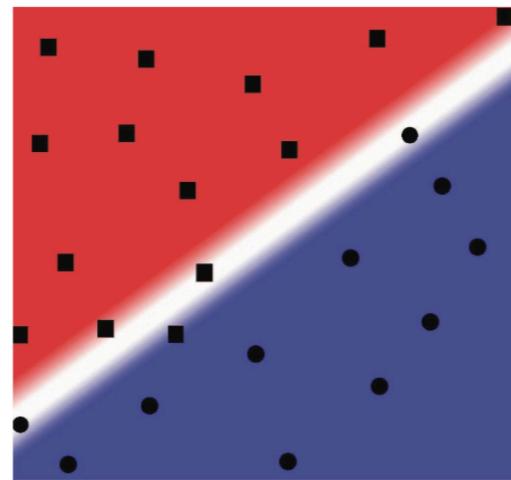
I. Input data: label noise

The labels of the input training data are noisy or contaminated.
How can we deal?

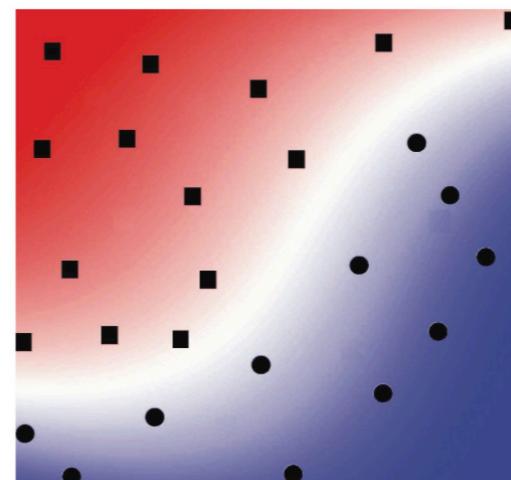


Yu, Xiyu, et al. "An Efficient and Provable Approach for Mixture Proportion Estimation Using Linear Independence Assumption."
Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.

II. Input Predefined Hypothesis class



Linear SVM



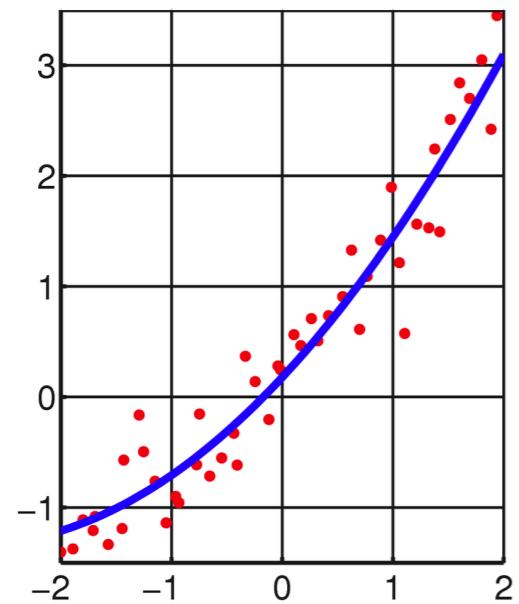
Kernel SVM

How to find a proper predefined hypothesis class?

III. Objective function

An objective is a numerical value to be minimised or maximised.

The objective function of Least squares:



$$\arg \min_h \frac{1}{n} \sum_{i=1}^n (y_i - h(x_i))^2$$

III. Objective function

$$\arg \min_h \frac{1}{n} \sum_{i=1}^n (y - h(x))^2$$

- Why the objective function of Least squares is defined in this way?
- What is the best regressor?
- Can we obtain the best regressor?

III. Objective function

- Given a classification task, we should firstly defined which hypothesis or classifier is the best.
- One intuitive way to defined the best classifier: the classifier that has the minimum classification error on the all possible data generated from the task.



Best classifier

- For a given data point (X, Y) , the classification error for a hypothesis h is measured by the 0-1 loss function:

$$1_{\{Y \neq \text{sign}(h(X))\}} = \begin{cases} 0 & Y = \text{sign}(h(X)) \\ 1 & Y \neq \text{sign}(h(X)) \end{cases}$$

- The best classifier can be mathematically defined as:

$$\arg \min_h \frac{1}{|D|} \sum_{i \in D} 1_{\{Y_i \neq \text{sign}(h(X_i))\}}$$

where D is the set of indices of **all possible data** points of the task, and $|D|$ denotes the size of the set D .



The law of large numbers

LLN describes the result of performing the same experiment a large number of times.

The average of the results obtained from a large number of independent trials should converge to the expected value.

$$\frac{1}{|D|} \sum_{i \in D} 1_{\{Y_i \neq \text{sign}(h(X_i))\}} \xrightarrow{|D| \rightarrow \infty} \mathbb{E}[1_{\{Y \neq \text{sign}(h(X))\}}]$$



Best classifier

- The best classifier can be mathematically defined as:

$$\arg \min_h \mathbb{E}[1_{\{Y \neq \text{sign}(h(X))\}}]$$

- Some problems: 1, the distribution of data is unknown. We cannot calculate the expectation. 2, the objective function is not convex or smooth, hard to optimise. 3, what kind of hypothesis h should we employ to fit the data?

III. Objective function

- Given a classification task, we want to find a classifier such that the following is minimised:

$$\mathbb{E}[1_{\{Y \neq \text{sign}(h(X))\}}]$$

- We don't have the distribution of data. Fortunately, we have some examples (or a training sample) draw from the distribution:

$$S = \{(X_1, Y_1), \dots, (X_n, Y_n)\}$$

- Because of the law of large numbers, we can use

$$\frac{1}{n} \sum_{i=1}^n 1_{\{Y \neq \text{sign}(h(X))\}}$$

(unbiased estimator)

to estimate $\mathbb{E}[1_{\{Y \neq \text{sign}(h(X))\}}]$

III. Objective function

- The estimator is unbiased because

$$\frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{Y_i \neq \text{sign}(h(X_i))\}} \xrightarrow{n \rightarrow \infty} \mathbb{E}[\mathbb{1}_{\{Y \neq \text{sign}(h(X))\}}]$$

- This also explains why big data is very helpful.

IV. Optimisation method

- How to obtain the hypothesis that minimises the objective function, i.e.,

$$\arg \min_h \frac{1}{n} \sum_{i=1}^n 1_{\{Y_i \neq \text{sign}(h(X_i))\}}$$

- Pick one from the predefined hypothesis class H to minimise the objective, i.e.,

$$\arg \min_{h \in H} \frac{1}{n} \sum_{i=1}^n 1_{\{Y_i \neq \text{sign}(h(X_i))\}}$$

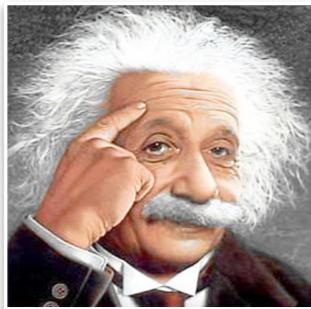
IV. Convex optimisation

- Pick one from the predefined hypothesis class H to minimise the objective, i.e.,

$$\arg \min_{h \in H} \frac{1}{n} \sum_{i=1}^n \ell(X_i, Y_i, h)$$

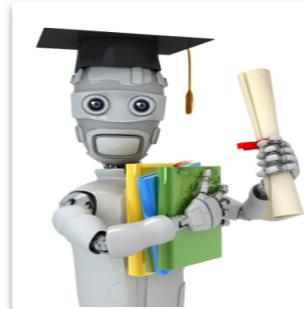
where the loss function ℓ is a convex surrogate for the 0-1 loss function.

V. Output hypothesis h_S



Human

- Experience
- Choose a rule
- Do classification



Machine

- A sample
- Choose a hypothesis
- Do classification



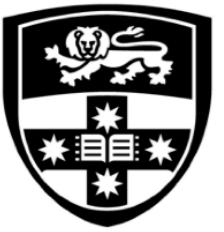
Human vs Machine learning

- Human can be trained to be qualified for a job.
How about machine?
- As human grows older, more reliable decision
will be made. How about machine?



Generalisation error of h_S

- Informally: generalisation error is defined by the difference between training and test error.
- Formally: will be introduced in Week 3.
- Generally: the more the training data, the smaller the generalisation error.



THE UNIVERSITY OF
SYDNEY

Thank you!