Using Machine Learning Tools PG

Week 1 – Introduction to Machine Learning

COMP SCI 7317

Trimester 2, 2024

THE UNIVERSITY of ADELAIDE

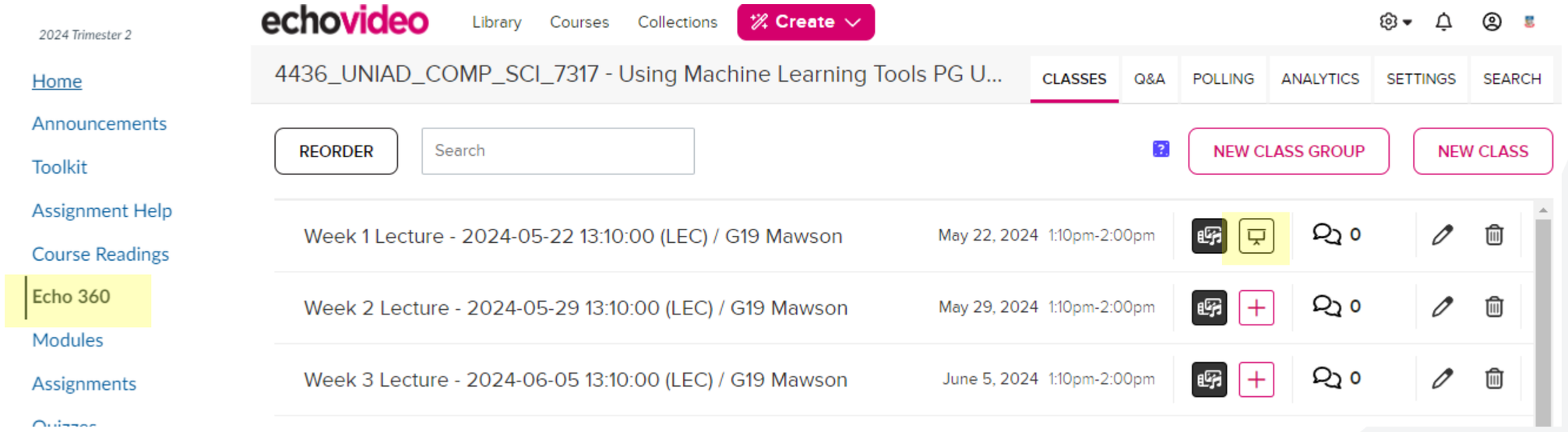
We acknowledge and pay our respects to the Kaurna people,
the traditional custodians whose ancestral lands we gather on.

We acknowledge the deep feelings of attachment and relationship of the Kaurna people to country and we respect and value their past, present and ongoing connection to the land and cultural beliefs.

Before we start…

There will be some **interactive elements** throughout this lecture.

Please participate by heading to **MyUni** > **4436_COMP_SCI_7317** > **Echo360** > **Lecture Week 1** > 



THE UNIVERSITY of ADELAIDE
150 YEARS

Let’s get to know each other

Which degree are you currently completing?

A) Master of Data Science

B) Master of AI & Machine Learning

C) Master of Computer Science

D) Master of Cyber Security

E) Other

THE UNIVERSITY of ADELAIDE

150 YEARS

Teaching team

Course co-ordinators

Prof. Mark Jenkinson

Email: mark.jenkinson@adelaide.edu.au

Dr. Dhani Dharmaprani

Email: dhani.Dharmaprani@adelaide.edu.au

Tutors

Shubham Bhatia

Email: a1844883@adelaide.edu.au

Saurabh Prashar

Email: a1844883@adelaide.edu.au

Luke Whitbread

Email: luke.whitbread@adelaide.edu.au

THE UNIVERSITY of ADELAIDE

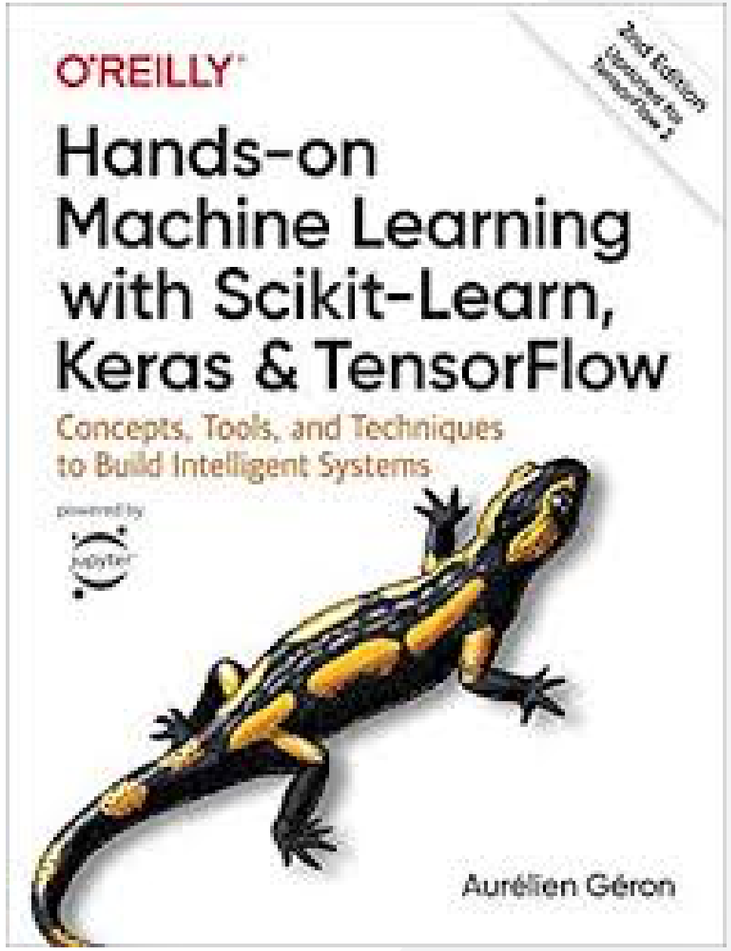
150 YEARS

Textbook

Course Textbook:

'*Hands-on Machine Learning with Scikit-Learn, Keras & Tensor Flow*', Aurélïen Géron

You can find online copies in course readings on MyUni

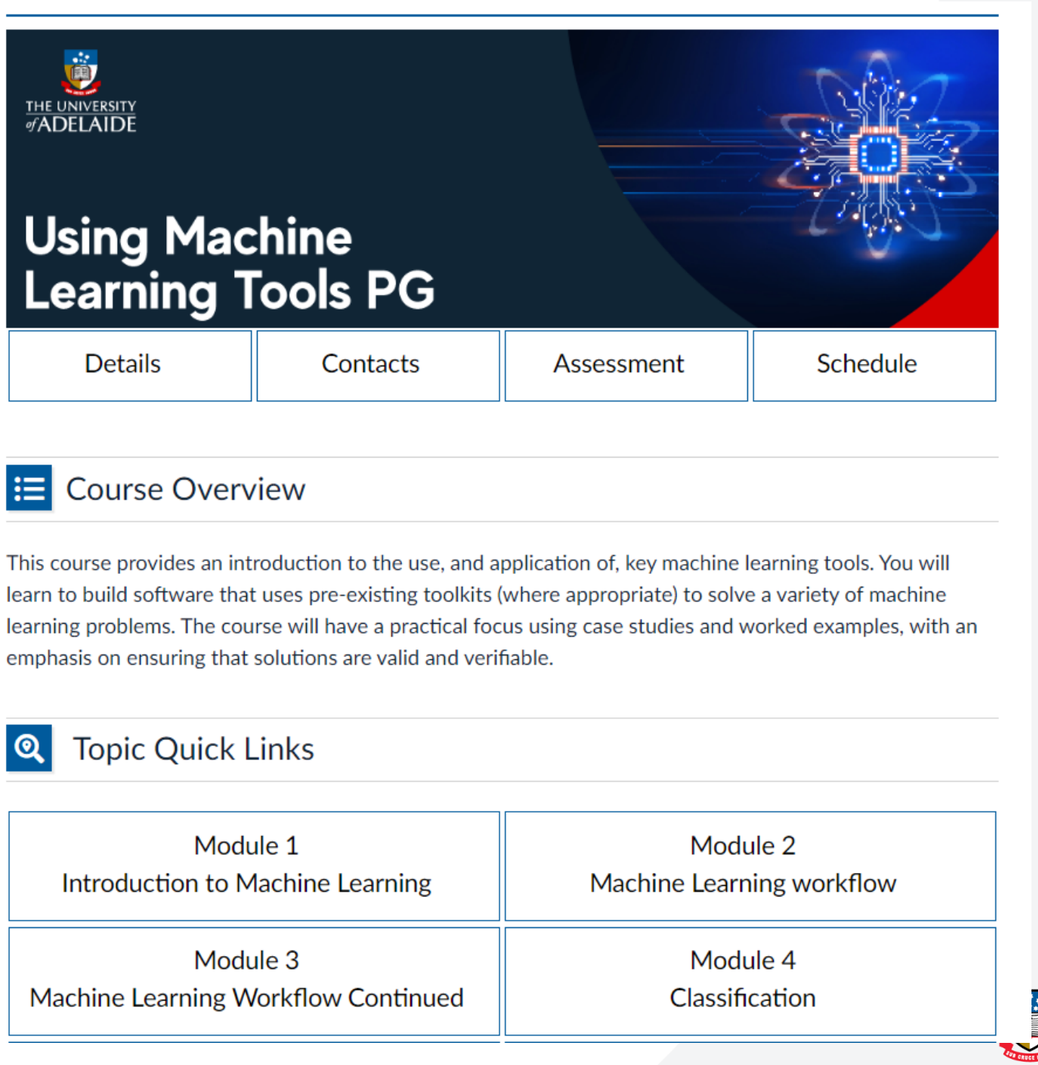


THE UNIVERSITY of ADELAIDE
150 YEARS

MyUni Page

Outlines the course timetable, lectures, workshops, assignments, announcements, discussions etc.

Each week has a dedicated module which covers one topic.



THE UNIVERSITY of ADELAIDE
150 YEARS

Assessments, procedures & policies

Contact hours and weekly course outline

You will have:

- 1 lecture (1hr) each week
- 1 workshop (2hr) in weeks 1 - 3; 5 – 12

Week 1 workshop optional (help with setup)
No workshops in week 4

Workshops are **very helpful** for your assignments (worth 80% of your topic grade) and in-class graded quizzes (20% of topic grade).

- Opportunity to go through guided examples/code
- Discussions about conceptual ideas, critical thinking
- ‘Answers’ to workshops are not provided outside of the class, so you are encouraged to attend

THE UNIVERSITY of ADELAIDE
150 YEARS

Assessments

- This course has 2 assessed components:
  - 3 practical assignments:
    - Assignment 1, worth 20%, due in week 5
    - Assignment 2, worth 30%, due in week 8
    - Assignment 3, worth 30%, due in week 12/13
  - 2 in-class quizzes, worth 10% each (20% in total)
    - These will be held at the start of your workshop in weeks 3 and 10

- There are also weekly multiple-choice quizzes that do not count towards your final grade
- You are expected to participate in all activities, **view all** lectures, do the quizzes and submit your assignments on time.

THE UNIVERSITY of ADELAIDE
150 YEARS

Grading

Grade	Mark	Description
FNS		Fail No Submission
F	1-49	Fail
P	50-64	Pass
C	65-74	Credit
D	75-84	Distinction
HD	85-100	High Distinction
CN		Continuing
NFE		No Formal Examination
RP		Result Pending

- There are no minimum performance hurdles in this course
- To pass the course, you must obtain a passing mark overall, based on the total from the three assignments and the two assessed multiple-choice quizzes.

THE UNIVERSITY of ADELAIDE
150 YEARS

Late submissions

- If you hand in your work late, your mark will be capped, based on how many days late it is.
  - up to 1 day late — **mark reduced to 75%, marks below 75% not affected**.
  - up to 2 days late — **mark reduced to 50%, marks below 50% not affected.**
  - up to 3 days late — **mark reduced to 25%, marks below 25% not affected.**
  - More than 3 days late — **mark is reduced to 0.**

- If you handed in something on time, and it is worth more than something that you handed in late, you will get the higher mark.
- Hand in early!

THE UNIVERSITY of ADELAIDE
150 YEARS

Assignment extensions

- A student may be eligible for assignment extensions based on **medical, compassionate, extenuating** circumstances

- A student will be ineligible if their circumstances:
  - **Were avoidable** and the student had opportunity to make alternative arrangements;
  - **Relate to balancing workloads** from other units of study, disciplines or faculties;
  - **Were personal commitments** or events such as international travel, holidays or weddings;
  - **Relate to temporary minor ailments** (colds, minor respiratory infections, headaches)
  - **Relate to stress or anxiety normally associated** with study
  - As a result of **misreading or misunderstanding** the examination timetable.

THE UNIVERSITY of ADELAIDE
150 YEARS

Assignment extensions

- A

- A

For extensions, please contact course coordinator

dhani.dharmaprani@adelaide.edu.au

Students who deliberately submit false or fraudulent documentation may be referred to the Student Misconduct Tribunal.

- As a result of **misreading or misunderstanding** the examination timetable.

THE UNIVERSITY of ADELAIDE

150 YEARS

Academic Integrity

Academic Integrity

The University takes academic integrity **very seriously**. For the most serious types of misconduct students can be **suspended** or completely **excluded** from the University.

Please refer to the University's academic integrity policy:
https://www.adelaide.edu.au/student/success/academic-integrity-for-students

THE UNIVERSITY of ADELAIDE
150 YEARS

Academic Integrity

- **Plagiarism**: students present work for assessment or publication that is not their own, without attribution or reference to the original source.
- **Collusion**: students present work as independent when it has been produced in whole or in part with others, unless prior permission for joint or collaborative given.
- **Contract Cheating**: student submits completed or partially completed work that a third party has completed for them.
- **Misrepresentation**: student presents untrue information with the intention of deceiving or misleading the assessor.
- **Solicitation**: student offers or gives money or any item or service to a University staff member or any other person to gain academic advantage for the student or another person.

THE UNIVERSITY of ADELAIDE

150 YEARS

What about the use of AI tools?

- **AI tools** like ChatGPT are now available.
- We recognise the value of using AI tools, so will aim to teach you how to evaluate the outputs and think critically about its use in your learning.
- The University has put together a reference document for the use of AI tools, including how to cite its use, found here: https://libguides.adelaide.edu.au/c.php?g=959585&p=6965636

THE UNIVERSITY of ADELAIDE
150 YEARS

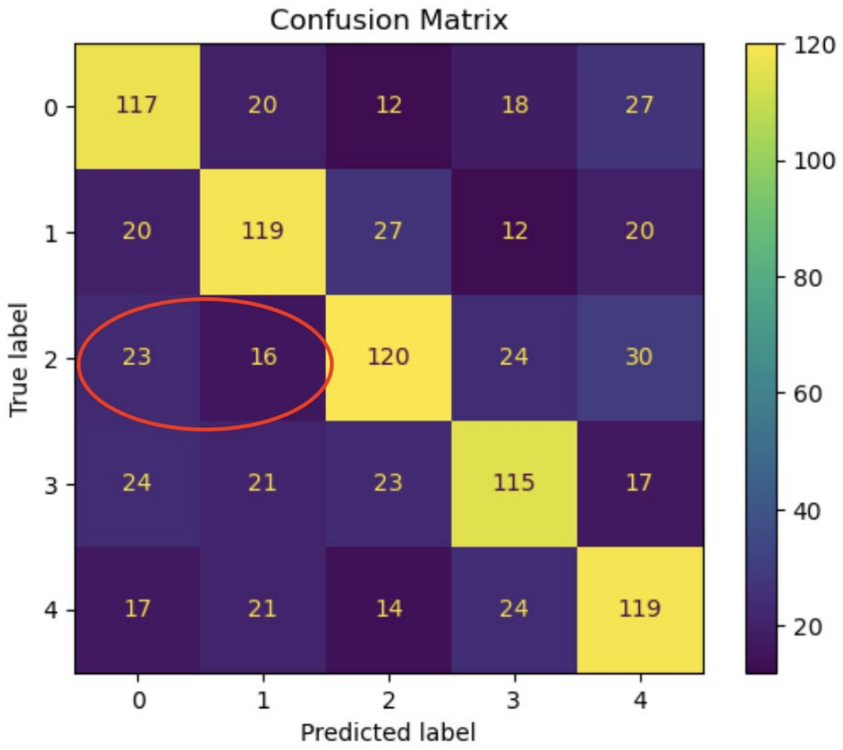
Guidelines for using AI tools

- **Supplement not replace:** These tools should only assist your learning and should not be used to substitute it without your own thoughts and analysis.
- **Cite appropriately:** Always give credit to sources, including AI tools. There are examples on how to do so in the University's AI tools reference document.
- **Check reliability:** Verify the information provided by the AI tools before using it in your work!

THE UNIVERSITY of ADELAIDE
150 YEARS

Here’s an example…

Input into ChatGPT



Give me a text version of the numbers in this matrix.

Output from ChatGPT

ChatGPT

Here's the text version of the numbers in the confusion matrix provided:

Copy code

```
117, 20, 12, 18, 27
20, 119, 27, 12, 20
16, 21, 120, 24, 30
24, 21, 23, 115, 17
17, 21, 14, 24, 119
```

THE UNIVERSITY of ADELAIDE

150 YEARS

What is the most common type of error ChatGPT makes?

A) Factual inaccuracies: Incorrect or outdated information

B) Misunderstanding the query: Incorrect understanding of user's intended question

C) Overgeneralisation: Providing responses that are overly broad

D) Bias: Biases that avoid neutrality of accuracy of responses

E) Failure to maintain context: Inability to recall previous detail

It is important to check for information that is

Incorrect (though it will appear confident)

Out of date

Biased or offensive

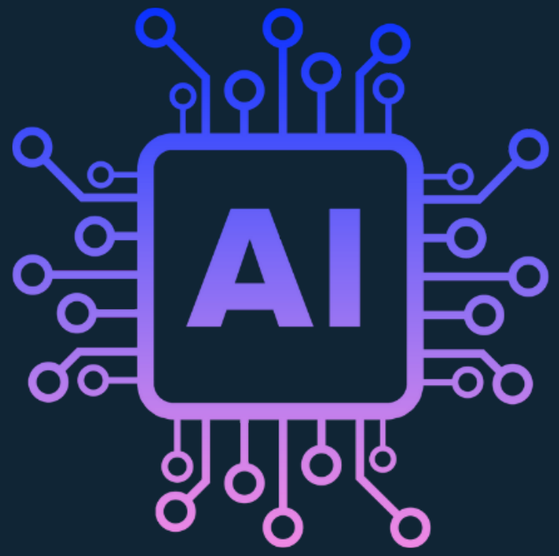
Showing poor judgement

Lacking originality

When is it appropriate to use LLMs?

- Not all sectors/workplaces embrace the use of AI tools like Large Language Models (LLMs).
- The use of LLMs may raise substantial concerns in the following contexts:
  - **Privacy and Confidentiality:** When there is potential risk of leaking sensitive information.
  - **Bias and Fairness:** When there is risk of perpetuating biases.
  - **Copyright Infringement:** When using copyrighted material without permission.
  - **Academic integrity:** When used to make intellectual contributions without proper attribution.

THE UNIVERSITY of ADELAIDE
150 YEARS



An Introducion to Machine Learning

Aims of the course

To teach you **how to use** machine learning tools, rather than to develop them.

- Data pre-processing and manipulation
- Dealing with missing/corrupt data
- How to select the right tool
- How to tune existing algorithms for best performance
- How to evaluate a machine learning tool
- Understand the strengths and weaknesses of each model
- Using statistical testing to make improvements/differences

Course topics

- **Scope of machine learning**
  - Typical workflow, practical issues
  - Use in real applications
- **Classification and regression problems (traditional ML)**
  - Training and validating models
  - Analyzing results
- **Unsupervised learning**
- **Deep Learning (deep neural networks)**
  - Fully connected and Convolutional Neural Nets
  - Effective training and testing

What is machine learning?

How do you differentiate a cat from a dog?

Have someone describe the differences to you in significant detail, including all features that relate to cats and those that relate to dogs.

Since I've seen hundreds of cats and dogs over time, I intrinsically know which is which.

What is machine learning?

Explicit or imperative programming is the first example:

Have someone describe the differences to you in significant detail, including all features that relate to cats and those that relate to dogs

- You define all conditions and steps necessary to complete a task, leaving little room for the program to operate outside the explicitly defined instructions.

From Géron, Hands-On Machine Learning

What is machine learning?

Machine learning!

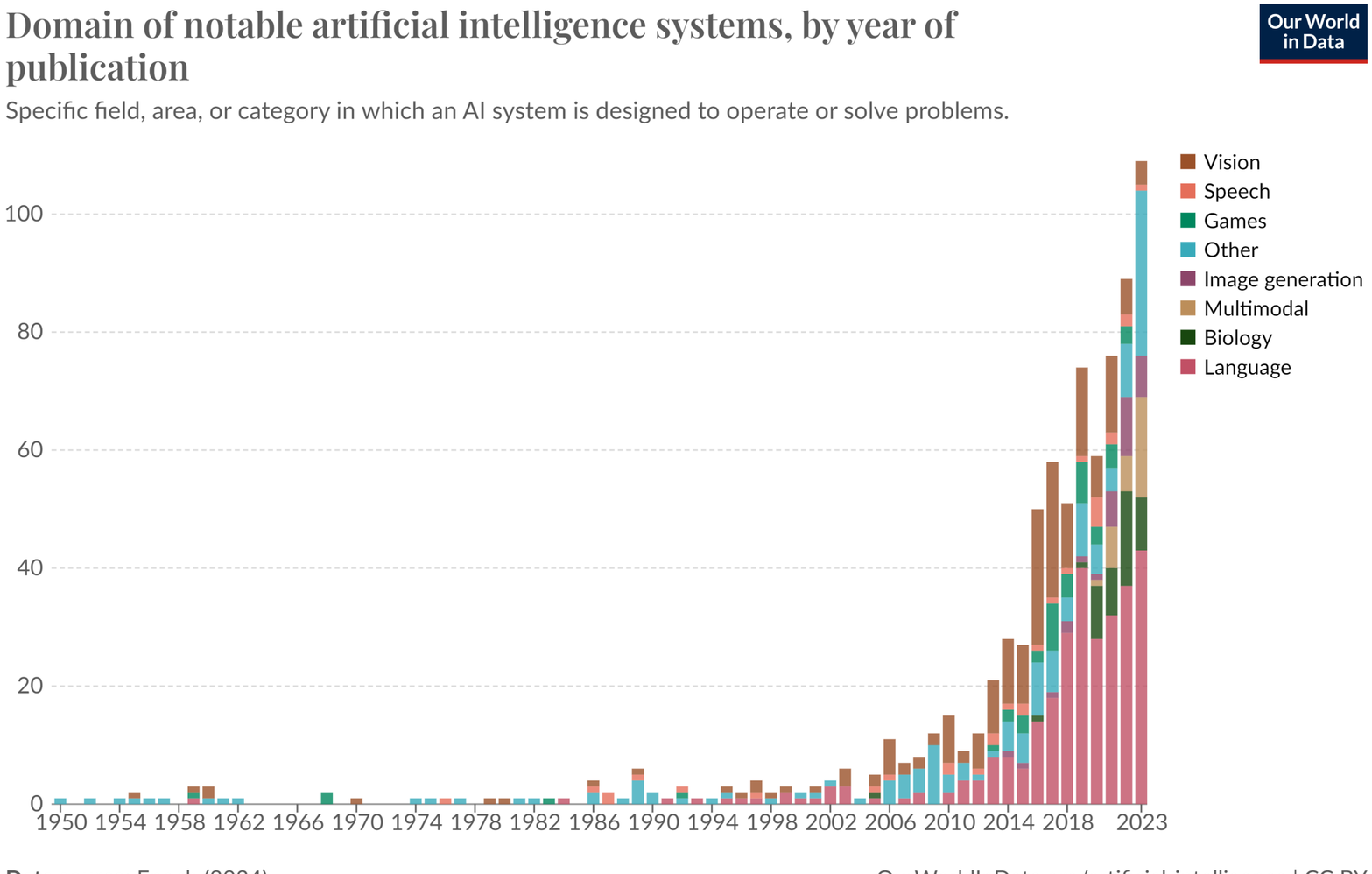
Pattern-based learning is the second example:

Since I've seen hundreds of cats and dogs over time, I intrinsically know which is which.

- An algorithm identifies patterns found in the data and makes decisions based on those patterns.

From Géron, Hands-On Machine Learning

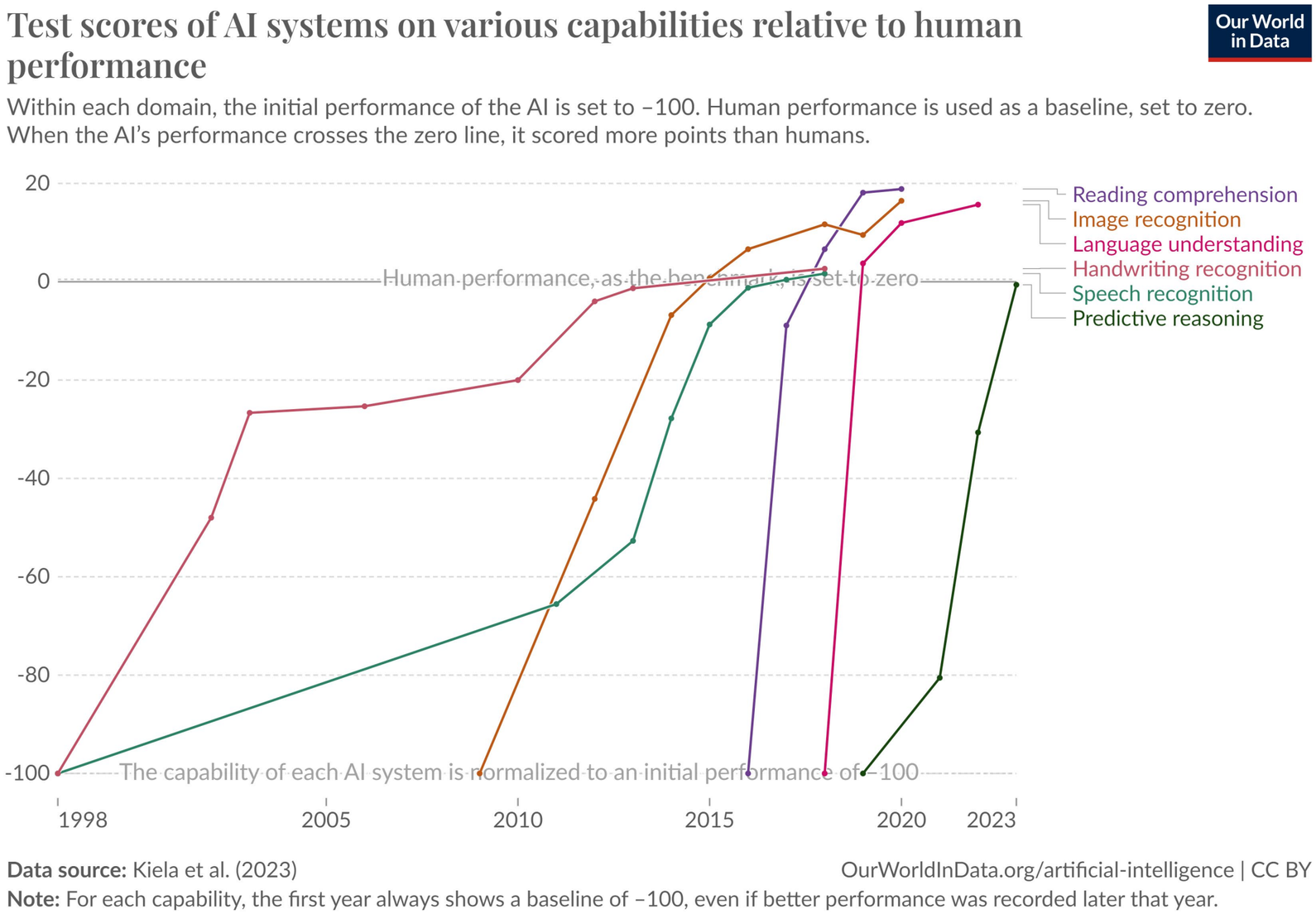
Machine learning in today's world



THE UNIVERSITY of ADELAIDE
150 YEARS

Charlie Giattino, Edouard Mathieu, Veronika Samborska and Max Roser (2023) - "Artificial Intelligence" Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/artificial-intelligence' [Online Resource]

Machine learning in today's world



THE UNIVERSITY of ADELAIDE

150 YEARS

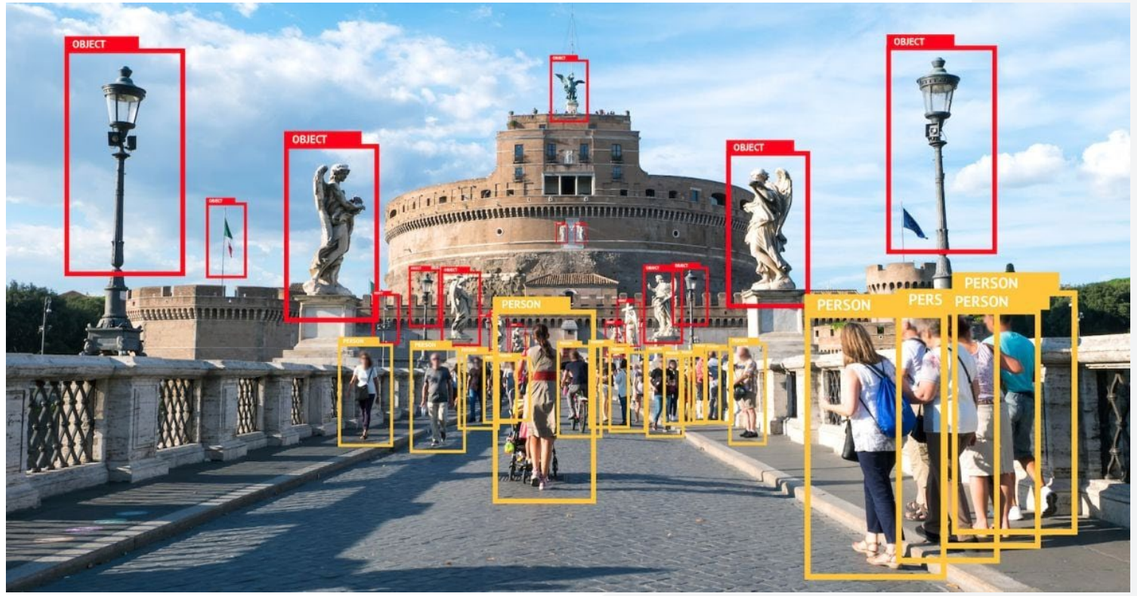
Charlie Giattino, Edouard Mathieu, Veronika Samborska and Max Roser (2023) - "Artificial Intelligence" Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/artificial-intelligence' [Online Resource]

Machine learning examples

Image recognition

Application: Used in various technologies from facial recognition systems in smartphones to automated image tagging in social media platforms.

How it works: Machine learning models (like Convolutional Neural Networks, CNNs) are trained on vast datasets of images that are labelled with information about their content. The model learns to recognise patterns and features in the images and uses this to recognise new images.



Machine learning examples

Recommendation systems

Application: Used by streaming services like Netflix or shopping platforms like Amazon to suggest media and products.

How it works: These systems use algorithms to predict user preferences based on past behaviour and similar user profiles.

Types of Machine Learning: Supervised

In the context of machine learning, **supervised** learning means:

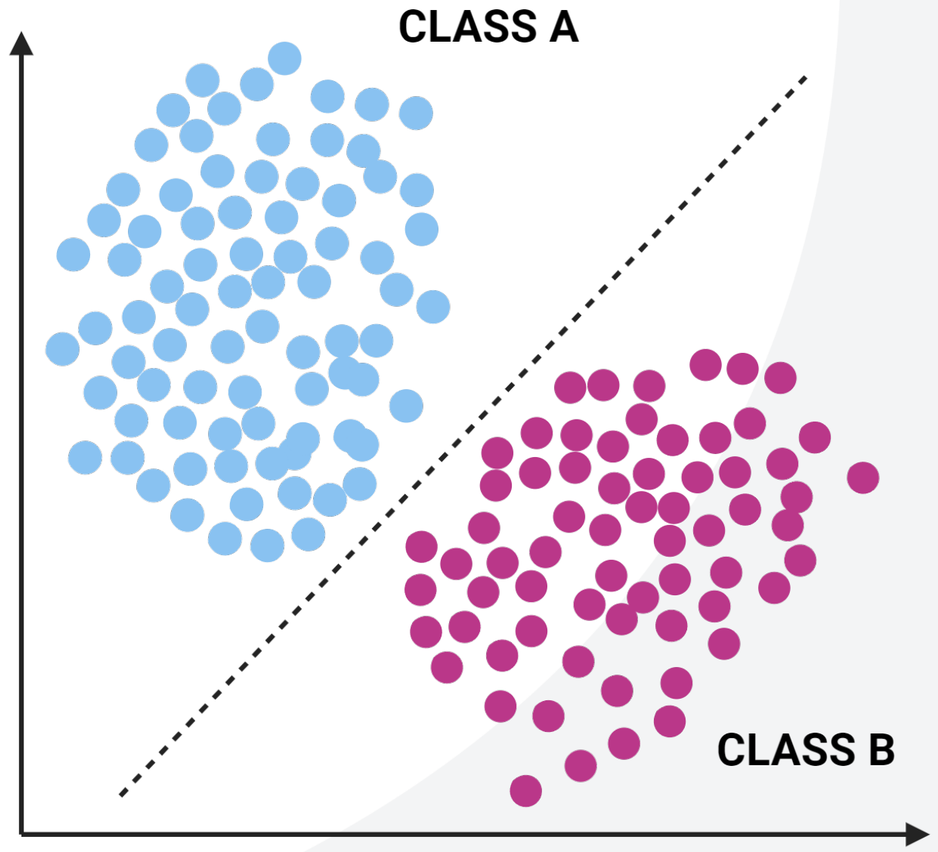
- Training data includes labels or values where we know the ground truth of the thing we are trying to predict (i.e. the desired result)

Some common types of supervised machine learning include **classification** and **regression** models

Classification

Classification algorithms use discrete labels and try to assign a ‘type’ to something

- Predict if new mail is junk
- Predict the type of animal in an image
- Predict whether a person has a disease or not (i.e. through medical images)



Regression

Regression models use real values to assign a continuous value to something, based on training data that contains values and possibly other features

- Predict life expectancy, based on historical/medical records
- Predict next month's sales
- Predict tomorrow's temperature

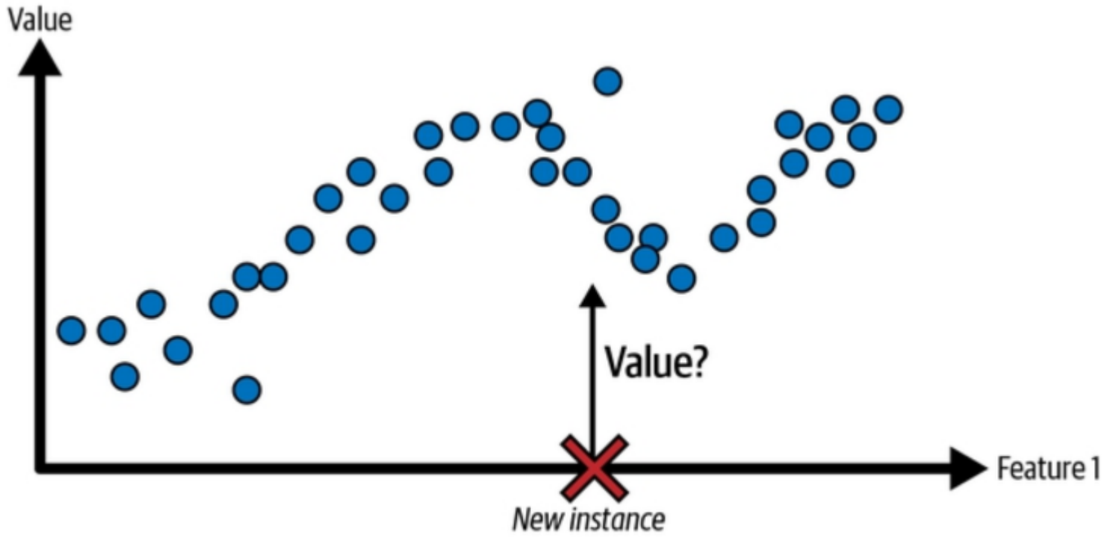


Figure 1-6. A regression problem: predict a value, given an input feature (there are usually multiple input features, and sometimes multiple output values)

From Géron, Hands-On Machine Learning

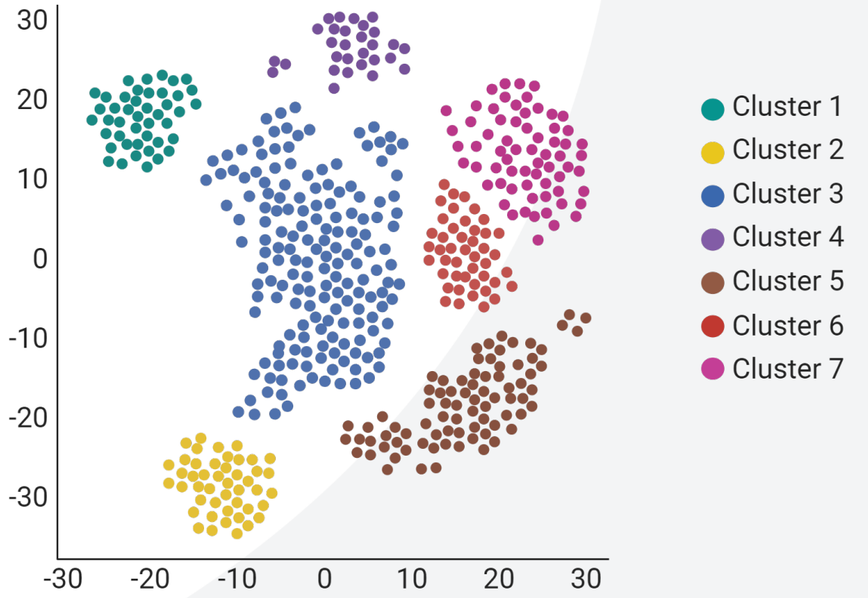
Types of Machine Learning: Unsupervised

In the context of machine learning, **unsupervised** learning means:

- Training data includes ***does not include*** labels or values where we know the ground truth of the thing we are trying to predict

Common applications of unsupervised learning include:

- **Clustering**
  - Marketing, sales etc.
- **Dimension Reduction**
  - Visualisation
  - Pre-processing
- **Anomaly detection**
  - Fraud detection
  - Security
  - Pathologies in medical data



Types of ML: Reinforcement learning

In the context of machine learning, **reinforcement** learning means:

- Decisions are learnt by performing actions in an environment and receiving feedback in the form of rewards/penalties. This helps to decide actions that lead to better outcomes over time.

This differs from supervised and unsupervised machine learning, which try to tries to learn relationships, patterns or structures from data.

Applications:

- Robotics, gaming, autonomous vehicles, finance, healthcare

Types of ML: Reinforcement learning

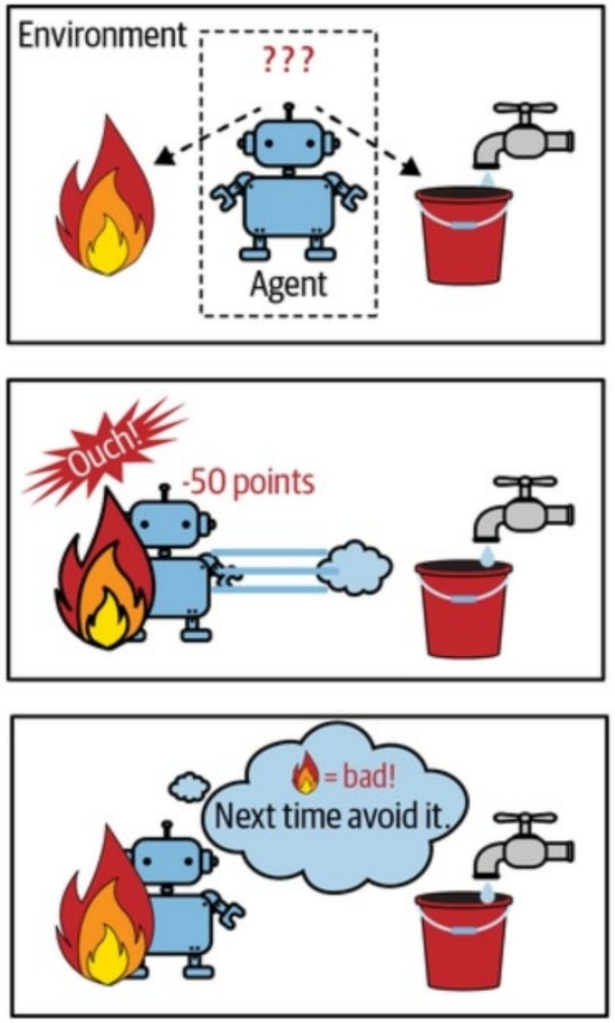


Figure 1-13. Reinforcement learning

From Géron, Hands-On Machine Learning

THE UNIVERSITY of ADELAIDE

150 YEARS

Which statement is correct?

A) Supervised ML – uses labels; unsupervised ML – learns from actions

B) Supervised ML – does not use labels; unsupervised ML – receives rewards/penalties

C) Unsupervised ML – does not use labels; reinforcement ML – receives rewards/penalties

D) Unsupervised ML – uses labels; reinforcement ML – uses ground truth

Considerations when using machine learning

Data quality and availability

- Machine learning models are only as good as the data they're trained on. Data quality and the amount of data available can significantly influence the accuracy and reliability.

- **Implication:** Poor data can lead to biased or inaccurate models that perform well in training but poorly in real-world applications.

Considerations when using machine learning

Ethics and Bias

- AI systems can inadvertently learn and perpetuate biases present in their training data. This is a significant ethical concern, particularly in applications like hiring, law enforcement, and lending.

- **Implication:** Developers must actively seek to identify and mitigate biases in datasets and model predictions.

Considerations when using machine learning

Model Complexity

- There's often a trade-off between model accuracy and interpretability. More complex models can achieve high accuracy but can also become 'black boxes’.

- **Implication:** Understanding why a model makes a certain decision is crucial for trust and accountability, especially in critical applications.

Considerations when using machine learning

Computational Resources

- Training sophisticated machine learning models requires significant computational power and energy, which can be costly and environmentally impactful.

- **Implication:** Optimising model architecture and training processes for efficiency can reduce costs and environmental footprint.

Summary

- Course administration, assessments, procedures & policies
- Academic integrity & use of AI tools
- Introduction to Machine Learning
  - There are different types and models, suited to different problems
  - There are various aspects to consider when using ML

Next week: Machine Learning Workflows

Further reading: Textbook chapter 1

Questions?

dhani.dharmaprani@adelaide.edu.au