Lecture 1: Introduction and Maths Review

COMP90049
Introduction to Machine Learning
Semester 1, 2025

Jean Honorio, CIS

Acknowledgement: Lea Frermann, Lida Rashidi

Roadmap

This lecture

- About COMP90049
- Machine Learning
- Expected Background
- Linear Algebra Review

About COMP90049

COMP90049 – Subject staff / Contact

Lecturers:	Jean Honorio [Week 1, Week 5 to Week 9 (1st lecture), and Week 12 (2nd lecture)] **Khris Ehinger** [Week 2 to 4] **Michael Kirley** [Week 9 (2nd lecture) to Week 12 (1st lecture)]
Tutors:	Hasti Samadi, Mojgan Kouhounestani (Head Tutors), Behzad Moradi, Jiayang Ao, Kazi Adnan, May Le, Rena Gao, Tom Lin, Viktoria Schram, Vincent Barboza *See Canvas for latest list and contact details.*
Ed discussion	*Already open*
Contacting staff	*Combined staff email* comp90049-2025s1-staff@lists.unimelb.edu.au
Readings	There is no particular textbook for this subject, but some lectures have recommended readings

On completion of this subject, students are expected to be able to:

- Apply elementary mathematical concepts used in machine learning
- Derive machine learning models from first principles
- Design, implement, and evaluate machine learning systems for real-world problems
- Identify the correct machine learning model for a given real-world problem

Programming assignment worth 20%, applying and evaluating specific machine learning method(s) on given dataset(s).

- From Week 4 to Week 7.
- *(Late submissions incur 10% penalty per day or fraction. Late submissions after 3 days not allowed.)*

A final, open-ended (in terms of methodology) research project worth 30%.

- From Week 8 to Week 12.
- *(Late submissions incur 10% penalty per day or fraction. Late submissions after 3 days not allowed.)*

2hr closed-book end-of-semester **on-campus exam** worth 50%.

Hurdle: to pass the subject hurdle, students must achieve 50% or higher of their final exam.

This subject is offered **on campus only**

- The lectures are **on campus**
- Tutorials are **on campus**
- All lectures will be recorded. All recordings and other materials will be made available online through Canvas
- Attending the lectures and tutorials is **highly recommended**.

Tutorials started on Week 1

COMP90049 – Subject Content

- **Topics**:
  - supervised learning (regression as well as classification such as k-nearest neighbors, decision trees, naive Bayes classifiers, logistic regression)
  - unsupervised learning (clustering)
  - semi-supervised learning (active learning)
  - evaluation (bias/variance tradeoff and generalization) * Important
  - ensemble learning
  - neural networks
  - generative models
  - anomaly detection
  - more...

- Refreshers on maths and programming basics
- Theory in the lectures
- Hands-on experience in tutorials and projects

Ed Discussion

- Actively engage by **asking and answering** questions. Peer teaching is the most effective way of learning!
- All questions will be answered usually within 2 weekdays
- Please do not post questions too close to an assessment deadline
- (Of course, no assignment solutions should be given away. Doing – or asking for – this is academic misconduct.)

Personal/private concerns: Email Jean, for instance

- With specific assignment questions
- With private or personal concerns
- Please include COMP90049 in the email subject

We need 2 or 3 student representatives

- Communication channel between class and teaching team
- Collect and pass on (anonymous) feedback or complaints
- Attend a student-staff meeting during the semester (TBD)
- Represent the **diversity** of the class

Interested? Send Jean an email.

Lecturer's Consultation

- **Tuesday and Thursday (the hour just after the lecture)**, go to ask the lecturer for questions just **after** the class, or wait for the lecturer in their office.

More information on Canvas

- Tutorials location and time
- Academic integrity
- University-wide support student services

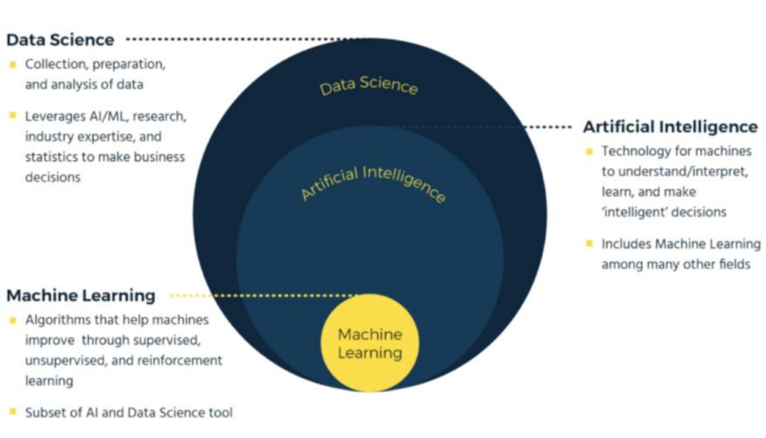
Machine Learning

What is learning and how do we learn?

Entity	Class
Baby	Cute
Sports car	Cool
Tiger	Cool
Hello Kitty	Cute
Koala	???
Water	???

Learning from examples
The goal is to accurately label new entities not seen before

What is Machine Learning?



- Machine learning is a method of teaching software to learn from data and make decisions on their own, without being explicitly programmed.

Definitions of Machine Learning

Learning what?

- **Concept**
  - Assign a label to inputs (for instance, finding letter A)
  - Group inputs into known classes (email → {spam, no-spam})
  - Understand patterns in the data

Learning from what?

- **Data**
- Where do the data come from? Is it reliable? Representative?

How do we learn?

- define a **model** that explains how to get from input to output → Big Part
- derive a **learning algorithm** to find the best model parameters

How do we know learning is happening?

- The algorithm improves at its task with exposure to more data
- We need to be able to **evaluate** performance objectively

THE UNIVERSITY OF MELBOURNE

Three ingredients for machine learning

… and related questions

1. Data

- Discrete vs continuous vs …
- Big data vs small data
- Labeled data vs unlabeled data
- Public vs sensitive data

Three ingredients for machine learning

... and related questions

2. Models

- function mapping from inputs to outputs
- probabilistic machine learning models
- geometric machine learning models
- parameters of the function are unknown ?

Three ingredients for machine learning

... and related questions

3. Learning

- Improving (on a task) after data is taken into account
- Finding the best model parameters (for a given task)
- Supervised vs. unsupervised learning

ML Example Problem

ML Example Problem

- Scenario 1

  You are an archaeologist in charge of classifying a mountain of fossilized bones, and want to quickly identify any “finds of the century” before sending the bones off to a museum

- Solution:

ML Example Problem

- Scenario 1

  You are an archaeologist in charge of classifying a mountain of fossilized bones, and want to quickly identify any “finds of the century” before sending the bones off to a museum

- Solution:

  Identify bones which are of different size/dimensions/characteristics to others in the sample and/or pre-identified bones

CLUSTERING/OUTLIER DETECTION

ML Example Problem

- Scenario 2:

  You are an archaeologist in charge of classifying a mountain of fossilized bones, and want to come up with a consistent way of determining the species and type of each bone which doesn't require specialist skills

- Solution:

ML Example Problem

- Scenario 2:

  You are an archaeologist in charge of classifying a mountain of fossilized bones, and want to come up with a consistent way of determining the species and type of each bone which doesn't require specialist skills

- Solution:

  Identify some easily measurable properties of bones (size, shape, number of "lumps", ...) and compare any new bones to a pre-classified database of bones

SUPERVISED CLASSIFICATION

ML Example Problem

- Scenario 3:

  You are in charge of developing the next “release” of Coca-Cola, and you want to be able to estimate how well-received a given recipe will be

- Solution:

- Scenario 3:

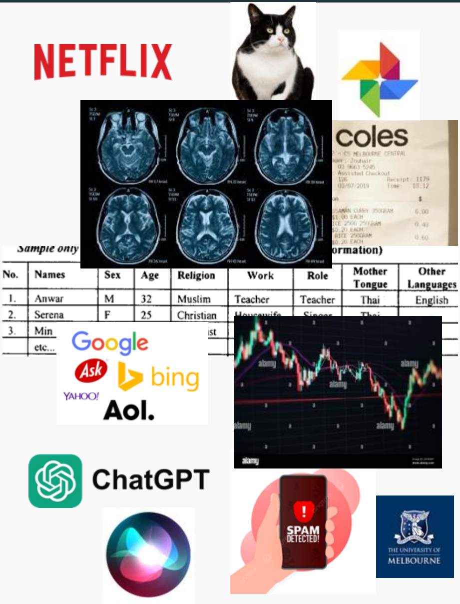
  You are in charge of developing the next “release” of Coca-Cola, and you want to be able to estimate how well-received a given recipe will be

- Solution:

  Carry out taste tests over various “recipes” with varying proportions of sugar, caramel, caffeine, phosphoric acid, coca leaf extract, ... (and any number of “secret” new ingredients), and estimate the function which predicts customer satisfaction from these numbers

  REGRESSION

- natural language processing
- image classification
- stock market prediction
- movie recommendation
- web search
- medical diagnoses
- spam/malware detection
- ...



Expected Background

Expected Background

Programming concepts

- We will be using **Python** and **Jupyter Notebooks**
- Basic familiarity with libraries (NumPy, Scikit-learn, SciPy, pandas)
- You need to be able to write code to process your data, apply different algorithms, and evaluate the output
- Tutorials use Jupyter Notebooks through Ed Lessons

Mathematical concepts

- Formal maths notation
- Basic probability, statistics, calculus, geometry, linear algebra

Some statistics about students in this class

62%	Master Information Technology
13%	Master Software Engineering
12%	Master Computer Science
10%	Master Mechatronics Engineering
3%	other Master programs

Some know a bit more programming than others
Some know a bit more math than others

What Level of Maths are we Talking?

$$\ln \frac{P(y = \mathrm{true}|x)}{1 - P(y = \mathrm{true}|x)} = w \cdot x$$

$$\frac{P(y = \mathrm{true}|x)}{1 - P(y = \mathrm{true}|x)} = e^{w \cdot x}$$

$$P(y = \mathrm{true}|x) = e^{w \cdot x} - e^{w \cdot x} P(y = \mathrm{true}|x)$$

$$P(y = \mathrm{true}|x) + e^{w \cdot x} P(y = \mathrm{true}|x) = e^{w \cdot x}$$

$$P(y = \mathrm{true}|x) = \frac{e^{w \cdot x}}{1 + e^{w \cdot x}} = \frac{1}{1 + e^{-w \cdot x}}$$

$$P(y = \mathrm{false}|x) = \frac{1}{1 + e^{w \cdot x}} = \frac{e^{-w \cdot x}}{1 + e^{-w \cdot x}}$$

Linear Algebra Review

Why?

Follow-up lectures that use linear algebra notation:

- linear regression
- logistic regression
- support vector machines
- perceptron, and more...

You only need to be familiar with the notation. This is not a linear algebra subject!

- vector
- norms/distances (you will see more in week 2)
- matrix
- inverse

Other math knowledge

- Probability theory and probabilistic modeling on week 2
- Python and Jupyter: during tutorials

Matrices

- A matrix is a 2-D array of numbers:

$$\begin{array}{cc} \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix}. & \begin{matrix} \leftarrow \text{Row} \\ \end{matrix} \\ \begin{matrix} \uparrow \\ \text{Column} \end{matrix} & \end{array}$$

A[row][column]

- Example notation for type and shape:

$$\boldsymbol{A} \in \mathbb{R}^{m \times n}$$

A and B have the same number of rows and columns

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 0 \\ 0 & 4 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 1 & 0 \\ 5 & 7 & 2 \\ -5 & 3 & 1 \end{bmatrix}$$

Add corresponding entries in A and B

$$(A+B)_{i,j} = A_{i,j} + B_{i,j}$$

$$A + B = \begin{bmatrix} 7 & 4 & 1 \\ 6 & 9 & 2 \\ -5 & 7 & 6 \end{bmatrix}$$

2+5

5+1

Subtract corresponding entries in A and B

$$(A-B)_{i,j} = A_{i,j} - B_{i,j}$$

$$A - B = \begin{bmatrix} -3 & 2 & 1 \\ -4 & -5 & -2 \\ 5 & 1 & 4 \end{bmatrix}$$

2-5

5-1

Number of columns of A = number of rows of B

$$(AB)_{i,j} = \sum_k A_{i,k} B_{k,j}$$

Example:

$$A = \begin{bmatrix} 3 & 1 & -2 & 4 \\ -2 & 4 & 2 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 2 & 1 \\ 4 & 5 & -3 \\ 2 & 3 & 2 \\ -1 & 2 & -4 \end{bmatrix}$$

$$AB = \begin{bmatrix} 5 & 13 & -20 \\ 14 & 22 & -10 \end{bmatrix}$$

$3 \times 2 + 1 \times 5 - 2 \times 3 + 4 \times 2 = 13$

BA ×

A scalar c is a real value

Multiply/divide all entries of matrix A by the scalar c

$$(cA)_{i,j} = cA_{i,j}$$

$$(A/c)_{i,j} = A_{i,j}/c$$

Example:

$$A = \begin{bmatrix} 4 & 5 \\ 0 & -2 \\ 3 & 6 \end{bmatrix}, \quad 3A = \begin{bmatrix} 12 & 15 \\ 0 & -6 \\ 9 & 18 \end{bmatrix}, \quad A/2 = \begin{bmatrix} 2 & 2.5 \\ 0 & -1 \\ 1.5 & 3 \end{bmatrix}$$

Rows become columns, columns become rows

$$(A^T)_{i,j} = A_{j,i}$$

Example:

$$A = \begin{bmatrix} 3 & 1 & -2 & 4 \\ -2 & 4 & 2 & 0 \end{bmatrix}, \quad A^T = \begin{bmatrix} 3 & -2 \\ 1 & 4 \\ -2 & 2 \\ 4 & 0 \end{bmatrix}$$

Multiplication property: $(AB)^T = B^T A^T$

If $A = A^T$ then A is called **symmetric** $A = \begin{bmatrix} 1 & 3 & 5 \\ 3 & -2 & 0 \\ 5 & 0 & 4 \end{bmatrix}$

A vector is a matrix with several rows and one column

$$a = \begin{bmatrix} 5 \\ 7 \\ 1 \\ 4 \end{bmatrix} = (5,7,1,4)^T$$

Notation: $a \in \mathbb{R}^m$

a is a vector with m Real numbers.

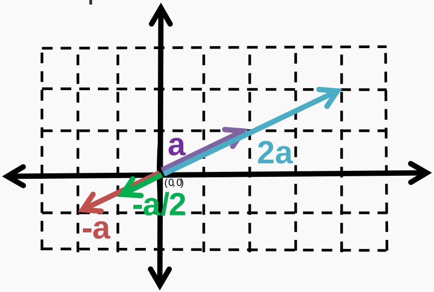
A scalar *c* is a real value

Multiply/divide all entries of vector *a* by the scalar *c*

$$(ca)_i = ca_i$$

$$(a/c)_i = a_i/c$$

Example:



$$a = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad 2a = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

$$-a = \begin{bmatrix} -2 \\ -1 \end{bmatrix}, \quad -a/2 = \begin{bmatrix} -1 \\ -0.5 \end{bmatrix}$$

a and *b* have the same number of rows

$$a = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 7 \\ 3 \end{bmatrix}$$

Add corresponding entries in *a* and *b*

$$(a+b)_i = a_i + b_i$$

$$a + b = \begin{bmatrix} 4 \\ 9 \\ 7 \end{bmatrix}$$

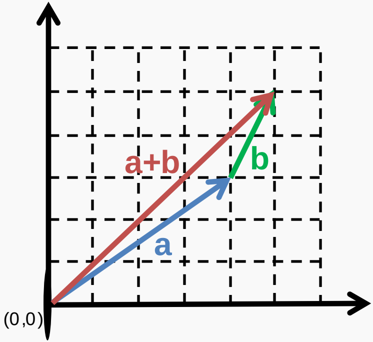
Subtract corresponding entries in *a* and *b*

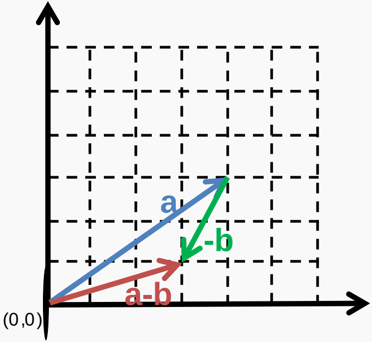
$$(a-b)_i = a_i - b_i$$

$$a - b = \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}$$

Geometrically…

$$a = \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad a + b = \begin{bmatrix} 5 \\ 5 \end{bmatrix}, \quad a - b = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$





Vector: inner product

(dot)

Using matrix multiplication notation:

dot product

$$a \cdot b = a^T b = \sum_{k=1}^{m} a_k b_k$$

a and *b* have the same number of rows: $a \in \mathbb{R}^m, b \in \mathbb{R}^m$

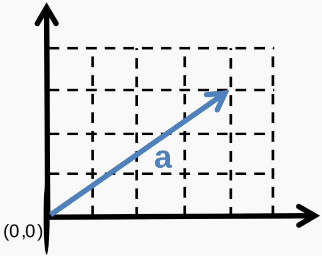
$$a = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ -7 \\ 3 \end{bmatrix}$$

$$a \cdot b = 3 \times 1 + 2 \times (-7) + 4 \times 3 = 1$$

Vector: Euclidean norm

The norm of $a \in \mathbb{R}^m$ is $\|a\| = \sqrt{a \cdot a} = \sqrt{a_1^2 + a_2^2 + \ldots + a_m^2}$

Example



$$a = \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \qquad \|a\| = \sqrt{3^2 + 4^2} = 5$$

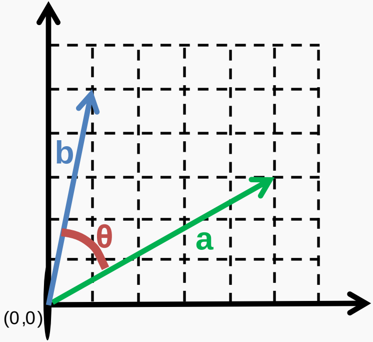
Distance between two vectors *a* and *b* is $\|a - b\|$

If $\|a\| = 1$ then a is called **unitary** $\quad a = \begin{bmatrix} 4/5 \\ 3/5 \end{bmatrix}$

If $a \neq (0,0,\ldots,0)$ then $a/\|a\|$ is a unitary vector

The cosine of the angle between two vectors can be found by using norms and the inner product

$$cos\theta = \frac{a \cdot b}{\|a\| \times \|b\|} = \left(\frac{a}{\|a\|}\right) \cdot \left(\frac{b}{\|b\|}\right)$$



$$a = \begin{bmatrix} 5 \\ 3 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

Identity matrix has 1s in the diagonals and 0s everywhere else

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

For any vector *x*, we have $Ix = x$ → I is akin to the number 1.

Matrix **inverse**: $A^{-1}A = I$

A matrix cannot be inverted if:

- More rows than columns
- More columns than rows
- Redundant rows/columns (linear independence)

Example

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 2 & 4 & 1 \\ -2 & 1 & 7 \end{bmatrix}, \qquad A^{-1} = \begin{bmatrix} -27 & 19 & 5 \\ 16 & -11 & -3 \\ -10 & 7 & 2 \end{bmatrix}$$

Several languages provide functions/methods for computing the inverse ***(We will not go into these details.)***

Summary

Today

- COMP90049 Overview
- What is machine learning?
- Why is it important? Some use cases.
- Linear algebra review

Next lecture: Concepts in machine learning