INTRODUCTION TO DEEP LEARNING

Lecture 1

MAL2, Spring 2025

COURSE PLAN


INTRO
(1 lesson)


DETAILS
(1 lesson)


IMAGES
(3 lessons)

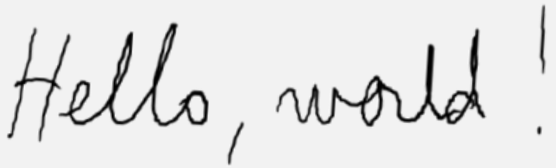

TIME
(3 lessons)


ACTION
(1 lesson)


ETHICS
(1 lesson)

PORTFOLIO ASSIGNMENTS

1. Recognizing speech



2. Generating handwriting

3. Composing music

4. Playing blackjack

FINAL PROJECT

A

machine learning system

involving at least one

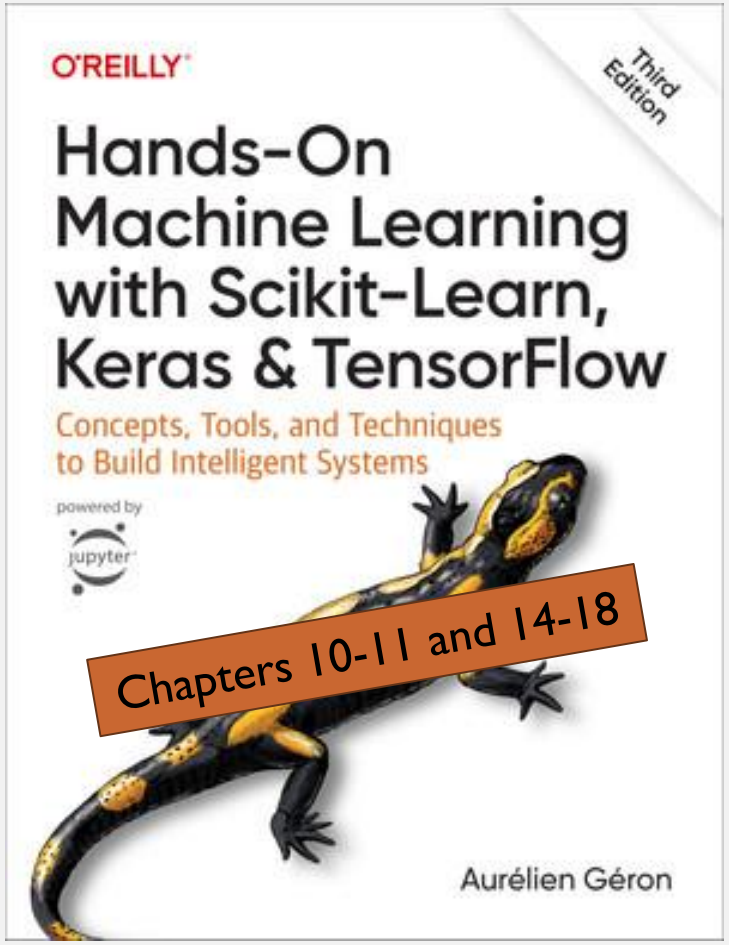
neural network

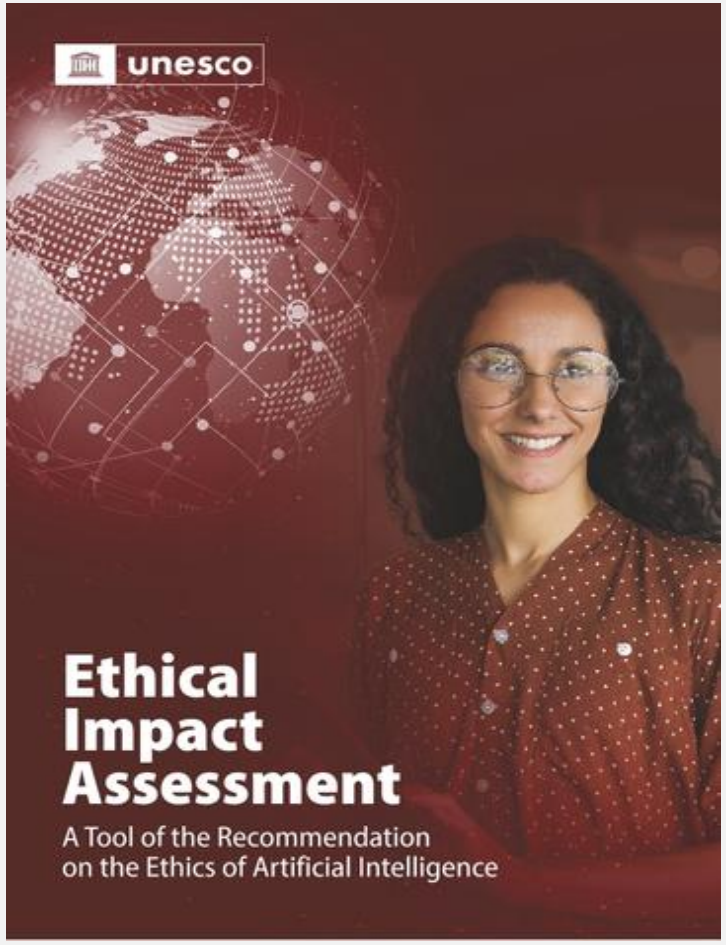
as well as an

ethical impact assessment

of that machine learning system.

LITERATURE





- Slides
- Notebooks
- Supplementary notes
- Portfolio assignments
- Your final project

COMPUTATIONAL RESOURCES

Login e-mail: mal2groupXX@via.dk

XX=01, 02, 03 …

Everything you run in Colab belongs to Google

"We need your permission if your intellectual property rights restrict our use of your content. You provide Google with that permission through this license."

The 3 rules for using Colab

1. Keep login details within the group.
2. Only use the license for MAL2 (and not BPR2).
3. Do not use with sensitive or confidential information (such as data covered by an NDA).

EXAM

Exam prerequisites:
None.

Exam type:
The course is evaluated based on an oral examination, which will take 20 minutes including everything.
At the exam, the student will randomly draw one of the portfolio assignments.
The exam will then take place as a discussion of this assignment, the students' group project and the curriculum in general.
Internal assessment.

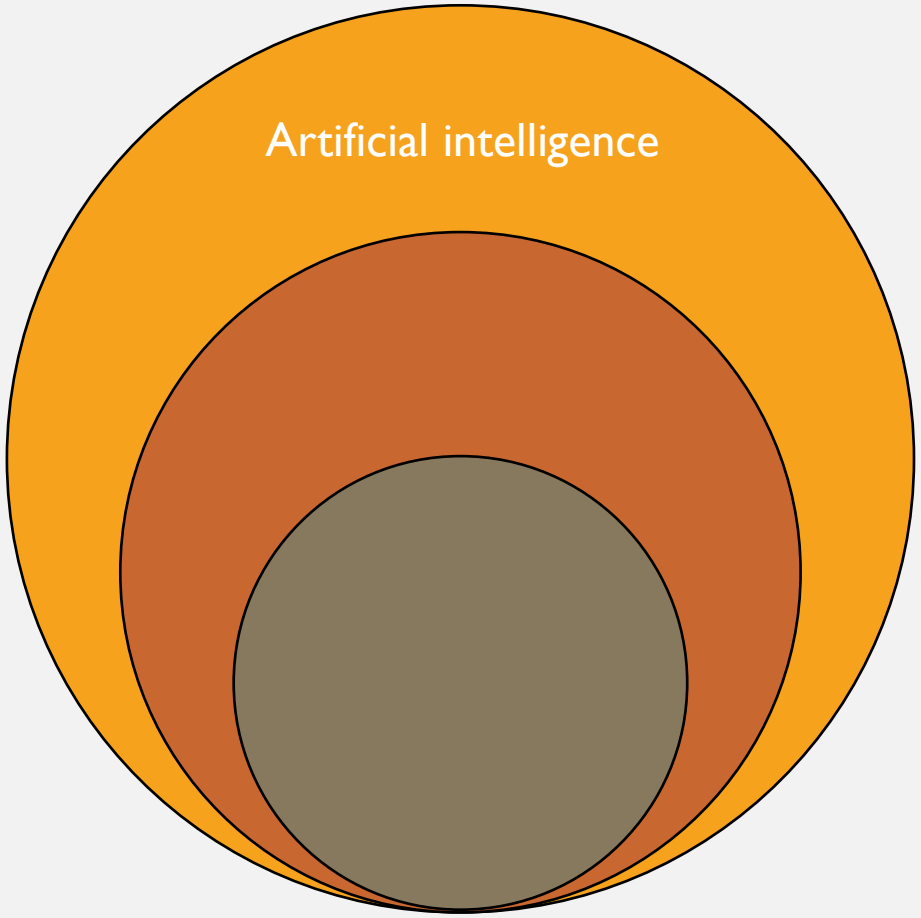
Tools allowed:
The student is expected to bring their portfolio assignments and their final project to the oral exam, such that they are able to display and run their code.

Re-exam:
The re-exam is the same as the ordinary exam.

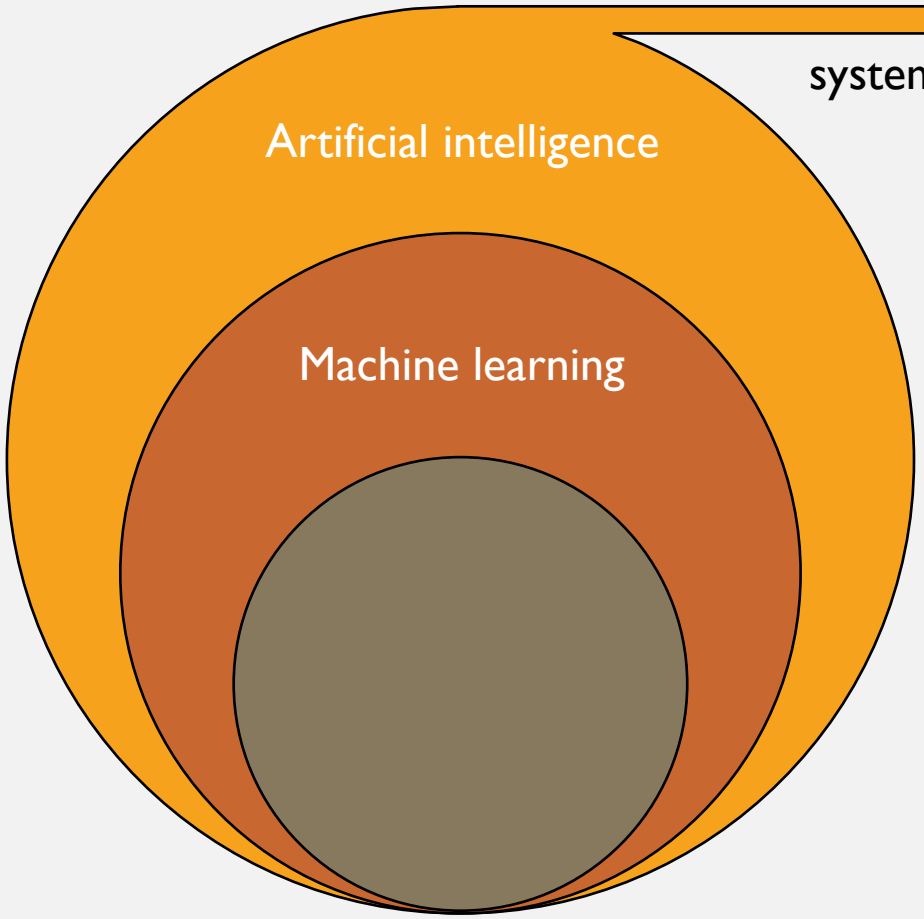
INTRODUCTION TO DEEP LEARNING

- What is deep learning?
- MALI recap
- Coding experiments
- Wrapping up & looking forward

WHAT IS DEEP LEARNING?

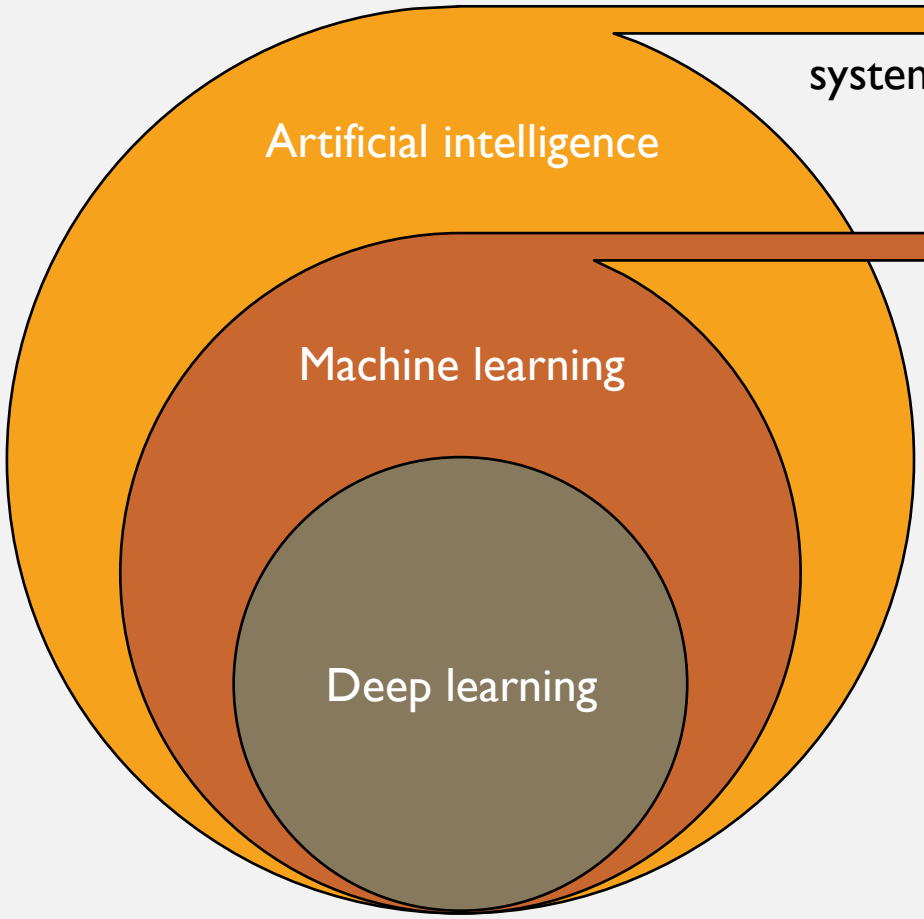


WHAT IS DEEP LEARNING?



systems that can perform tasks we would normally associate with human intelligence

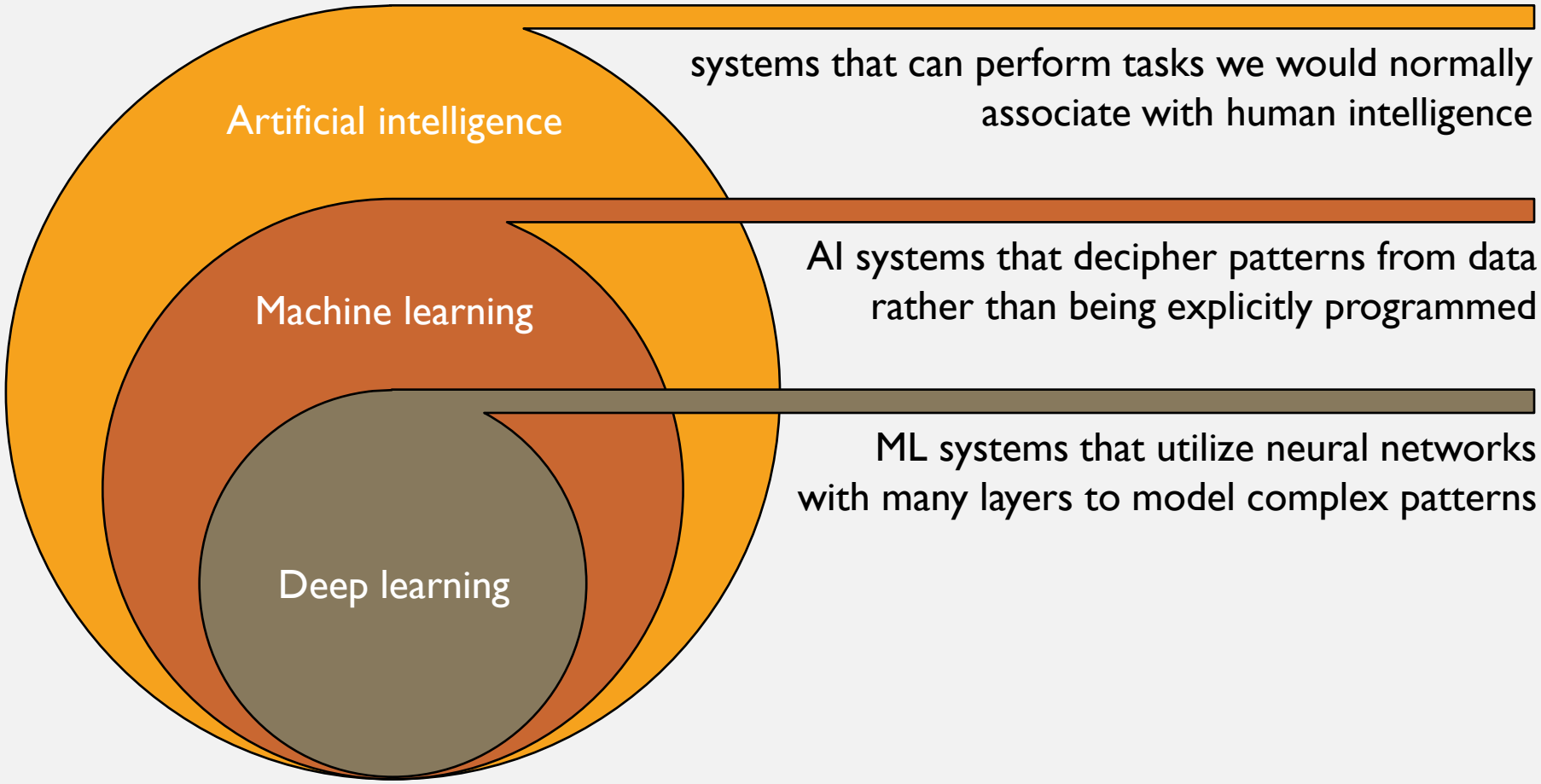
WHAT IS DEEP LEARNING?



systems that can perform tasks we would normally associate with human intelligence

AI systems that decipher patterns from data rather than being explicitly programmed

WHAT IS DEEP LEARNING?



INTRODUCTION TO DEEP LEARNING

- What is deep learning?
- MAL1 recap
- Coding experiments
- Wrapping up & looking forward

MALI RECAP

In groups of 2-3, **prepare a two-minute presentation including one illustration/diagram** on one of the following topics:

Neurons, weights and biases	**The structure of a neural network**	**Activation functions**	**Gradient descent**	**Loss functions**
How does a neuron work? What do weights and biases do?	*What types of layers do we have? How do layers connect?*	*What do they do? Why are they needed? What are some common ones?*	*How does it work? What is a learning rate?*	*What do they measure? How are they chosen?*

At most 2 groups can work on the same topic.

You have 10 minutes.

THE STRUCTURE OF A NEURAL NETWORK

during training, the weights and biases are updated using

GRADIENT DESCENT

find the direction in which the descent is steepest by calculating the gradient of the loss function:

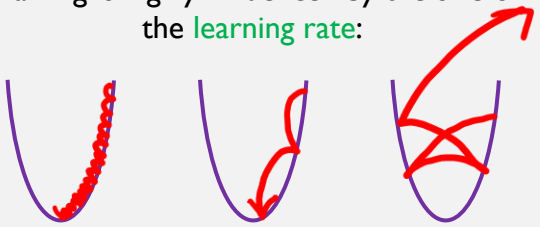
$$\nabla L(\theta) = \left(\frac{\partial L}{\partial \theta_1}, \frac{\partial L}{\partial \theta_2}, \frac{\partial L}{\partial \theta_3}, \ldots\right)$$

then update weights and biases by taking a step in that direction:

$$\theta \leftarrow \theta - \eta \nabla L(\theta)$$

and repeat until the gradient is 0.

Training is highly influenced by the size of the learning rate:

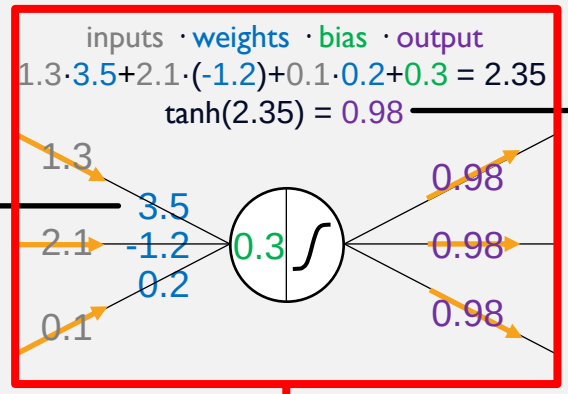


η too small
slow training

η just right
fast training

η too large
diverges

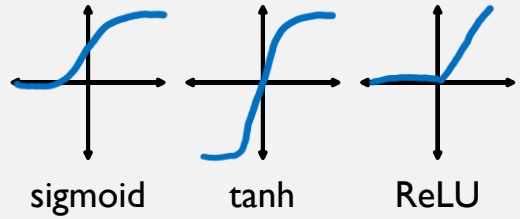
NEURONS, WEIGHTS AND BIASES



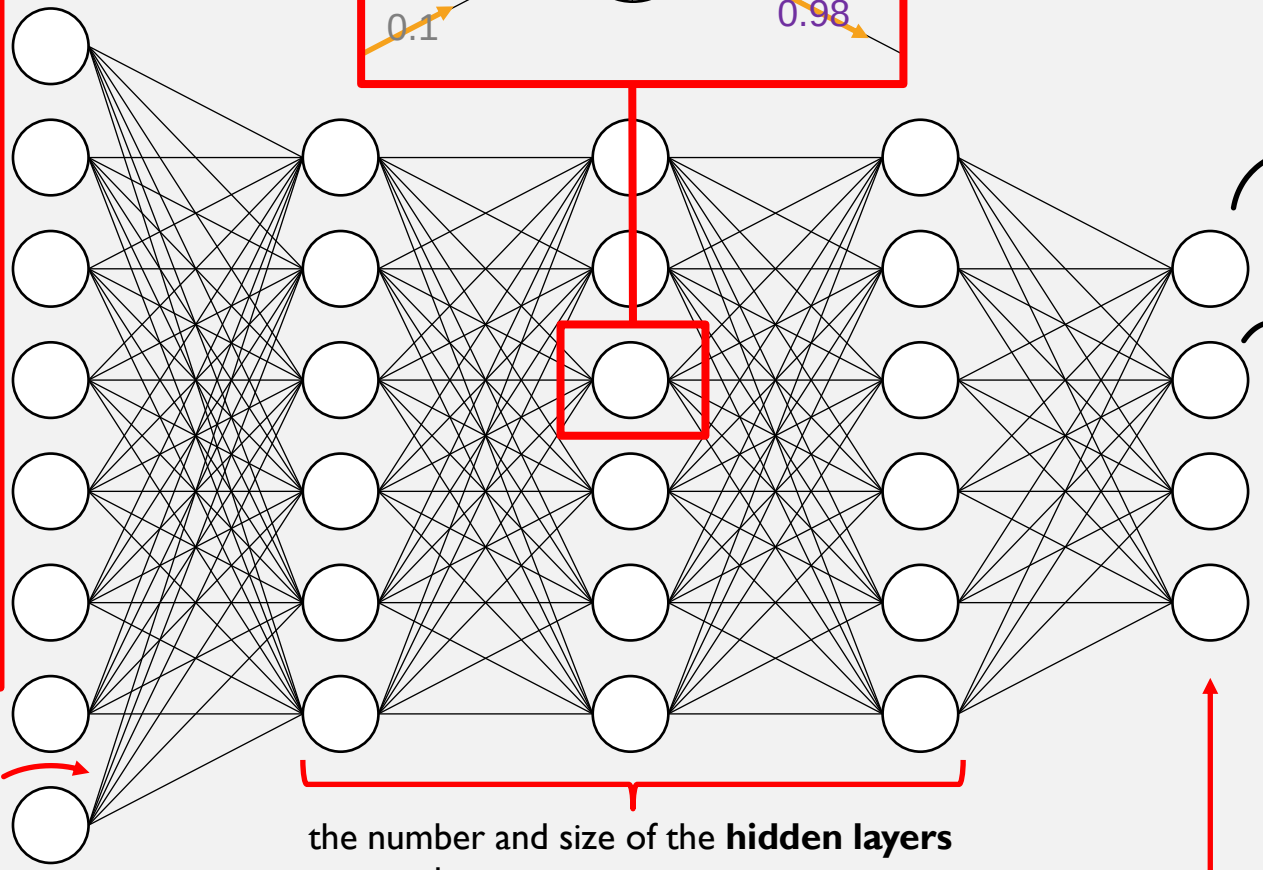
tanh is one of many possible

ACTIVATION FUNCTIONS

that introduce **non-linearity** in the neural network



in the output neurons, we typically use the **softmax** activation function to turn numbers into probabilities



in a **fully connected** neural network, every neuron in one layer is connected to every neuron in the next layer

the number and size of the **hidden layers** are hyperparameters we can tune

the difference between the predictions and the ground truth is measured by a

LOSS FUNCTION

such as mean squared error (MSE) for regression

$$\text{MSE} = \frac{1}{n}\sum (y_i - \hat{y}_i)^2$$

or cross-entropy for classification

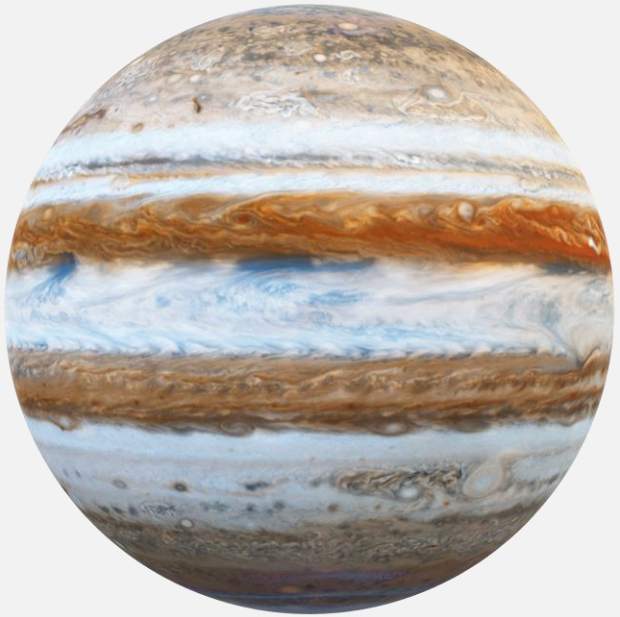
the size of the **input layer** is the number of features

the size of the **output layer** is the number of classes
(1 in a regression problem)

INTRODUCTION TO DEEP LEARNING

- What is deep learning?
- MALI recap
- Coding experiments
- Wrapping up & looking forward

CODING EXPERIMENTS



CODING EXPERIMENTS

In groups, play around with the neural network we just made.

- Modify the network architechture
- Experiment with different activation functions
- Adjust the learning rate
- Find out what a callback is and try the EarlyStopping callback
- …

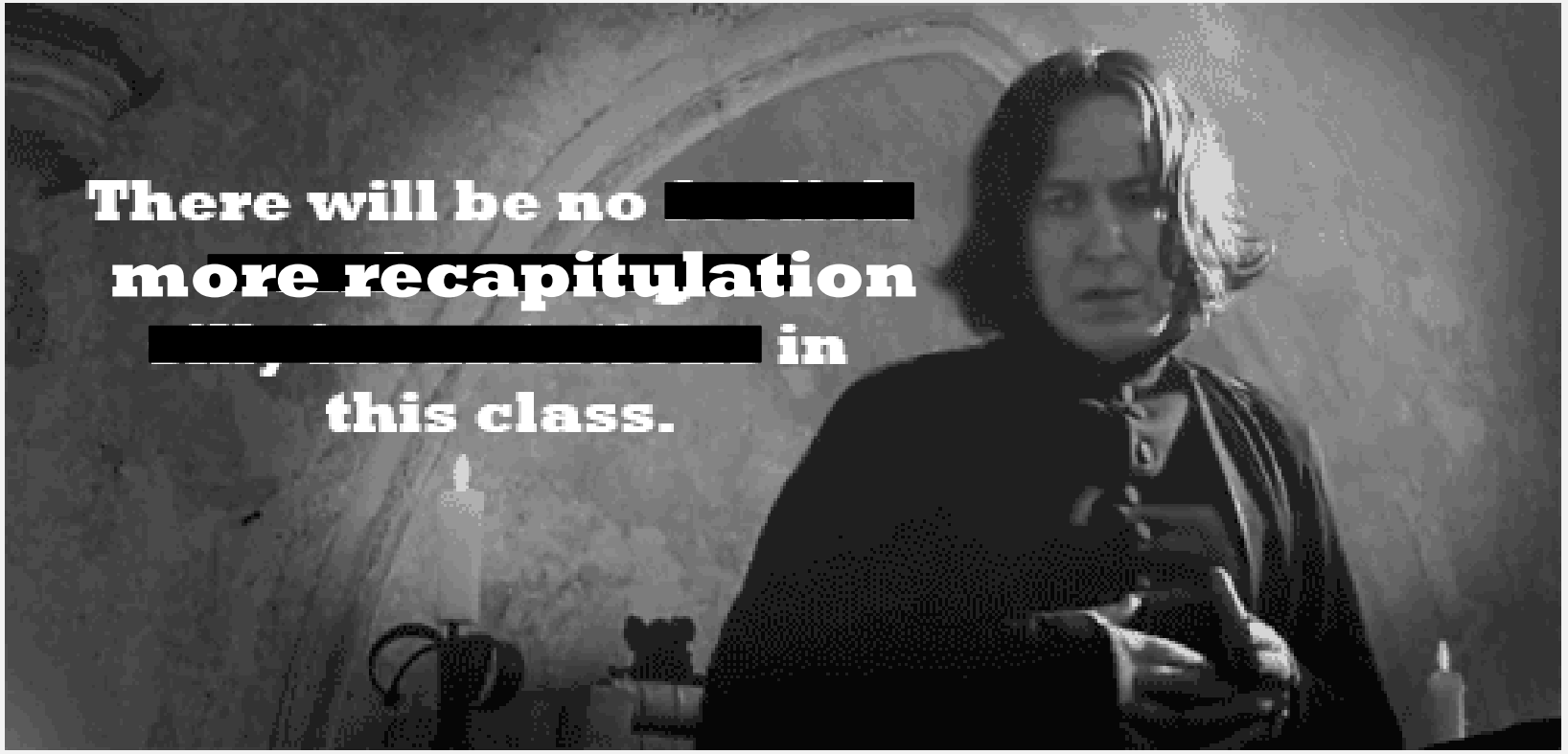
Try at least three different configurations and report what you see in terms of training time, overfitting, accuracy, etc.

You have 20 minutes.

INTRODUCTION TO DEEP LEARNING

- What is deep learning?
- MAL1 recap
- Coding experiments
- Wrapping up & looking forward

WRAPPING UP & LOOKING FORWARD




"That's all Folks!"
Isberg©