

CS2109S Tutorial 10

Unsupervised Learning with Neural Network + Review (Last Tutorial!)

(AY 25/26 Semester 2)

April 16, 2026

(Prepared by Benson)

Contents

Unsupervised Learning with Neural Networks

- Q1. Image Pre-training Strategies
- Q2. Next-Token Prediction

Review Questions

- Q3. Neural Networks
- Q4. Convolutional Neural Networks

Course Review

Q1. Image Pre-training Strategies

We have collected a massive dataset of **1.3 million high-resolution images from the ImageNet collection**. This dataset contains a wide variety of everyday objects, natural scenes, and vehicles, but it does not have any class labels. Our downstream task is to build a classifier for the **CIFAR-10** dataset. CIFAR-10 is a standard benchmark containing 60,000 small, low-resolution images sorted into 10 common categories, such as airplanes, cars, and ships. However, we only have a very small labeled data-set for these categories.

- (a) Is rotation prediction useful here as a pre-training task? How does this benefit the downstream task?
- ▶ Yes. The model must learn the natural orientation of the world (e.g. wheels belong on the ground, trees grow upward, etc.) to correctly predict whether an ImageNet photo has been rotated.

Q1. Image Pre-training Strategies

- (b) Instead of rotation prediction, we decide to use contrastive learning. How do we construct the training pairs? How does this benefit the downstream task?
- ▶ **Positive pair:** Two different augmented views of the same ImageNet photo.
 - ▶ **Negative pair:** Pair original image with an image of a completely different object from the dataset.
 - ▶ The model is trained to pull the embeddings of positive pairs together in the latent space while pushing negative pairs apart. This forces the model to learn consistency.

Q1. Image Pre-training Strategies

- (c) Now, we decide to use image inpainting. How is the training data for this task created, and how does it help the model learn about objects?
- ▶ "Mask" or cut out a random patch
 - ▶ Model predicts the pixel content of the missing area based on the surrounding context.
 - ▶ Model learns natural arrangement of object features like matching sides, how parts connect and continuing patterns

Q2. Next-Token Prediction

A tech company trains a massive transformer model entirely on the unsupervised task of next-token prediction using a massive dataset of unlabelled internet text (articles, forums, code, books).

- (a) The model is trained solely for next-token prediction. Explain in detail how it can be used to complete a full sentence.

Q2. Next-Token Prediction

- (b) An engineer provides the following input to the Base Model:

Input: “What is the capital of France”

Instead of answering the question with “Paris”, the model outputs:

Output: “and what is the currency of France?”

Explain why this behavior is possible.

- ▶ It’s only trained for text completion. It extends the question based on common textual patterns seen during pre-training.

- (c) To increase the probability that the model will output the correct answer without any fine-tuning, a friend suggests using the following as the input:

“Question: What is the capital of Japan? Answer: Tokyo.

Question: What is the capital of Australia? Answer: Canberra.

Question: What is the capital of France? Answer:”

Discuss why this input format may help the model generate the correct answer.

- ▶ There is a repeating structural pattern, encouraging the model to continue this structure. This technique is known as **in-context learning**.

Q3. Neural Networks

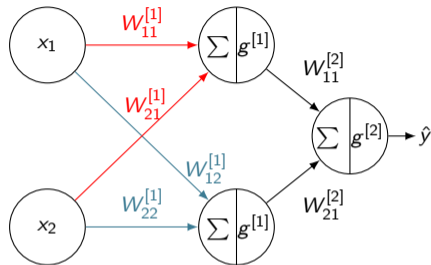
(a) If $x_1 = 1$, $x_2 = 1$, find \hat{y} .

$$\mathbf{f}^{[1]} = \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$\hat{y} = \text{ReLU} \left(\begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right) = \text{ReLU}(8) = 8$$

(b) Find the derivative of the loss w.r.t. $W_{11}^{[1]}$.

$$\begin{aligned} \frac{\partial J(W)}{\partial W_{11}^{[1]}} &= \frac{\partial J(W)}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial f^{[2]}} \cdot \frac{\partial f^{[2]}}{\partial f_1^{[1]}} \cdot \frac{\partial f_1^{[1]}}{\partial W_{11}^{[1]}} \\ &= \frac{1}{n} (\hat{y} - y) \cdot \begin{cases} 1 & \text{if } f^{[2]} > 0 \\ 0 & \text{otherwise} \end{cases} \cdot W_{11}^{[2]} \cdot x_1 \end{aligned}$$



$$\mathbf{W}^{[1]} = \begin{bmatrix} 2 & 3 \\ 1 & -1 \end{bmatrix}, \mathbf{W}^{[2]} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$g^{[1]}$: Identity, $g^{[2]}$: ReLU

x_1	x_2	y	\hat{y}
1	0	3	7
0	-1	2	0

Q4. Convolutional Neural Networks

Given a 3-channel image of height 3 and width 3.

- (a) Using all three channels as input for convolution, apply the 2D convolution operation as introduced in lecture to produce a 5-channel output using kernels of height and width 2. What is the total number of weights/parameters across all kernels in this convolution layer?

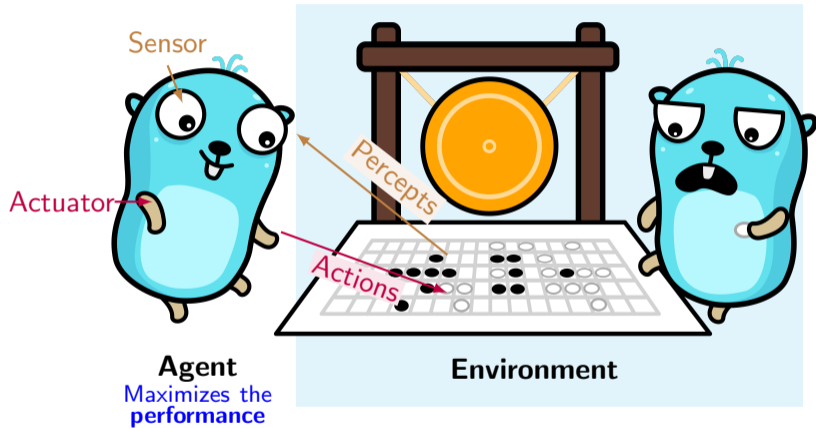
▶ $2 \times 2 \times 3 \times 5 = 60$.

- (b) Describe how to generate an output of shape 3×3 , where each element in the output is the average of the corresponding elements across the three input channels.

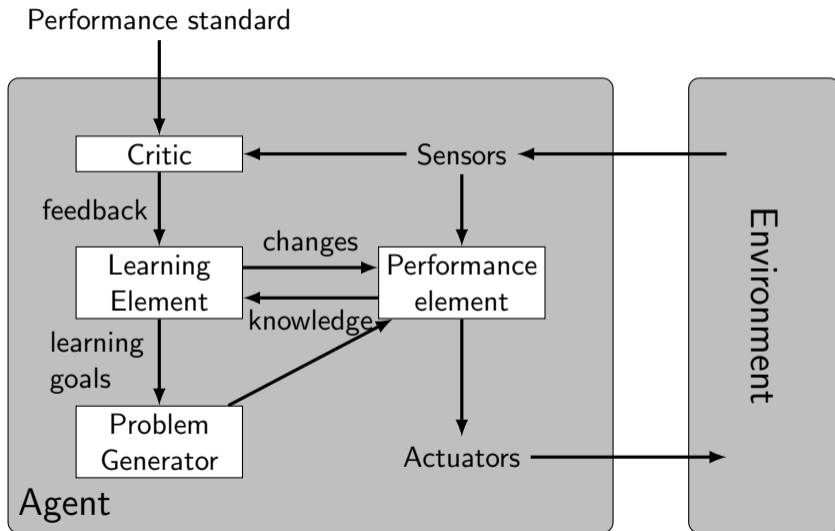
▶ $1 \times 1 \times 3$ kernel with all 3 weights being $\frac{1}{3}$.

▶ Stride: 1. No padding.

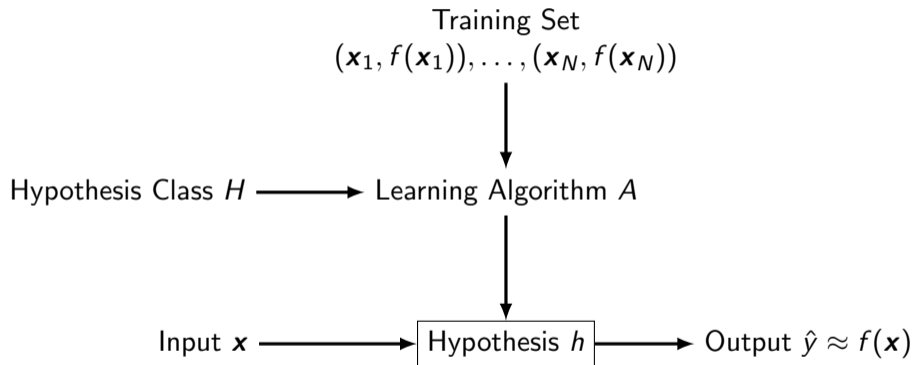
Course Review



Course Review



Course Review



Beyond CS2109S

Introductory

CS2109S Intro to AI and ML

Cool 5k mods

CS5339 Theory and Algo for ML

Algos, e.g. Perceptron, SVM, Kernels
ML Theory, e.g. Conc. Measures, VC-dim

Prerequisite: CS3264

CS5275 Algo Designer's Toolkit

Math Tools for Algo/ML: Randomized,
Optimization, Info Theory, and more

Prerequisite: CS3230

CS4262 Machine Learning Systems

Theoretical Foundations

CS3263 Foundations of AI

↪ **CS4246** AI Decision Making

↪ **CS5340** Uncertainty Modelling

CS3264 Foundations of ML

Sem 1 (Harold): Math-intensive version of CS2109S (more linear algebra?)

Sem 2 (Bryan): Focuses on theoreticals e.g. Proving convergence, Bayesian inference, Reinforcement learning

Applications

CS4243 Computer Vision

CS4248 Natural Language P.

Student Feedback Exercise:
Your Voice Matters!

Provide your feedback now >>

<https://blue.nus.edu.sg/blue/>

