Chapter Six

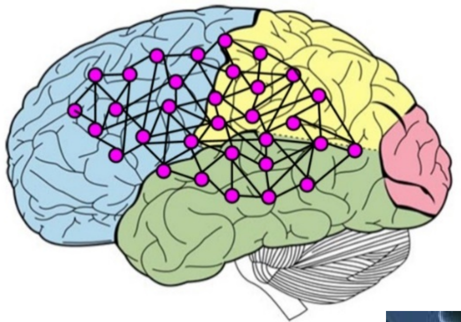
Unsupervised Learning

Zhang Yi, *IEEE Fellow*
Autumn 2019

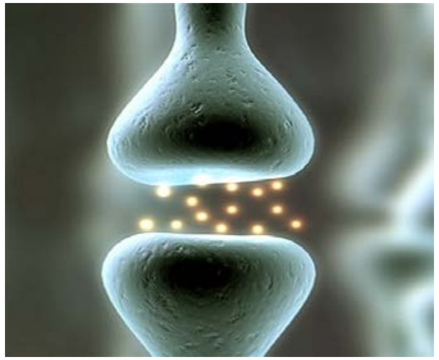
Outline

- Learning in Neural Networks
- Supervised Learning
- Unsupervised Learning
- Autoencoder Neural Networks
- Assignment

Learning in Neural Networks

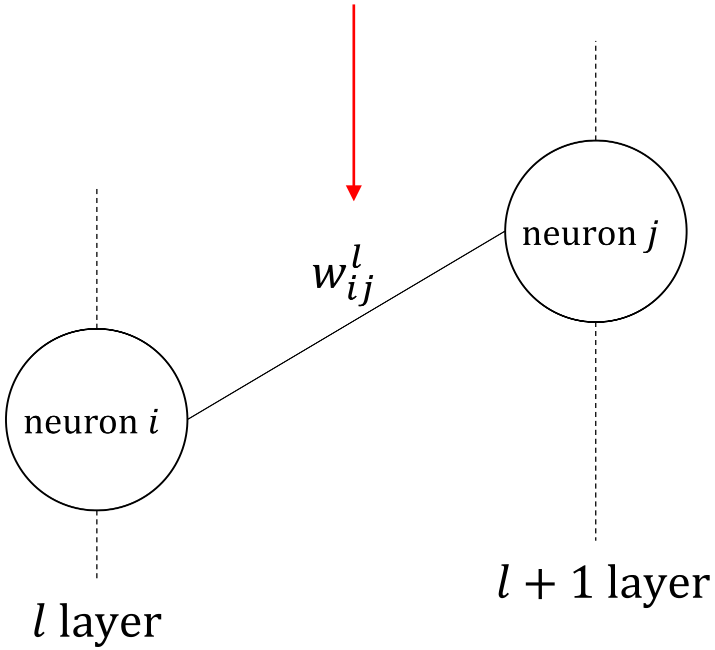


The brain is a learning system. The brain can learn by some supervisor or by itself. Thus, there are Supervised Learning and Unsupervised Learning.



Neural Network

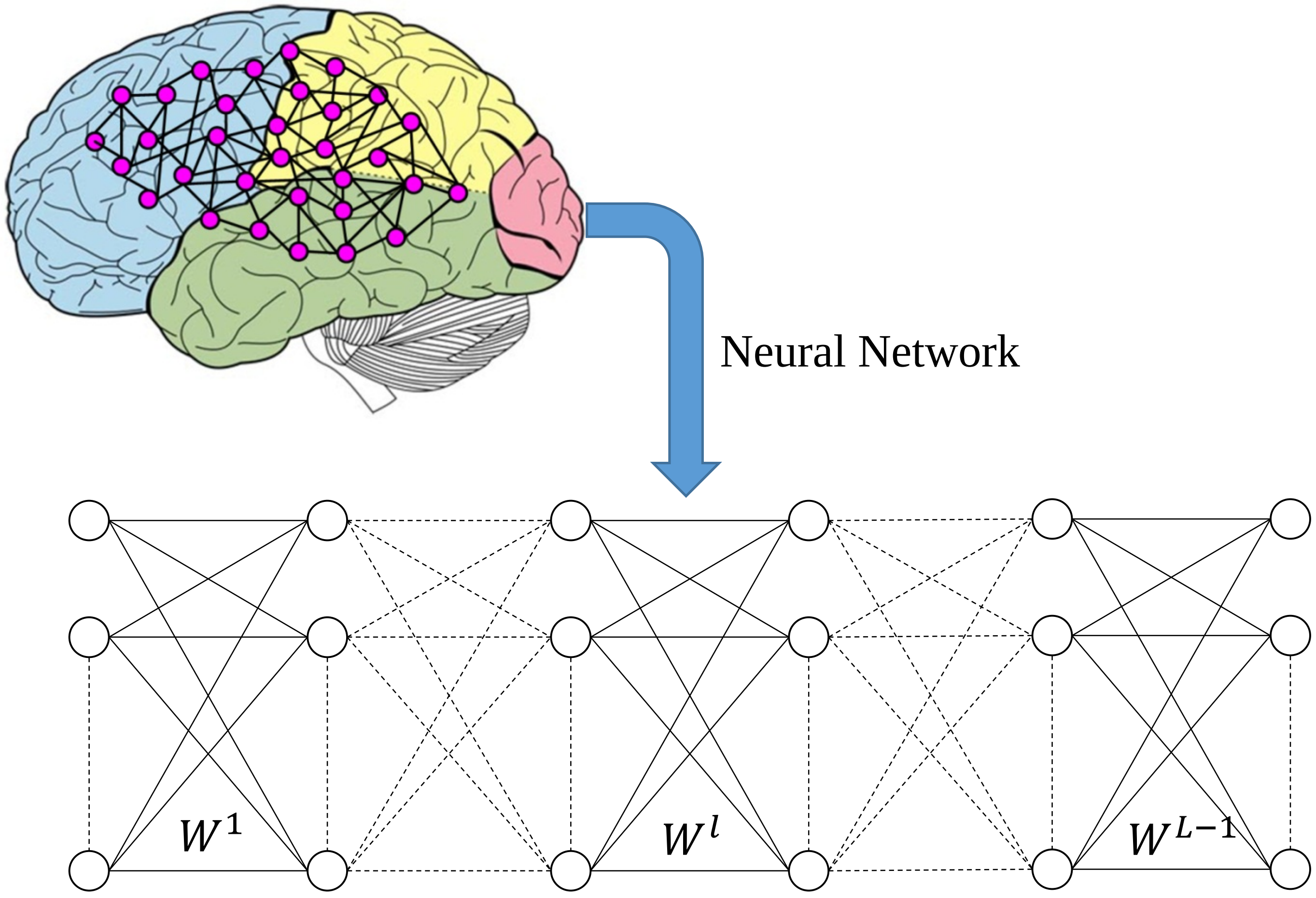
The knowledge is stored in connections weight.



Learning is the changing of connection strength between two connected neurons

$$w_{ij}^{l}(new) \leftarrow w_{ij}^{l}(old)$$

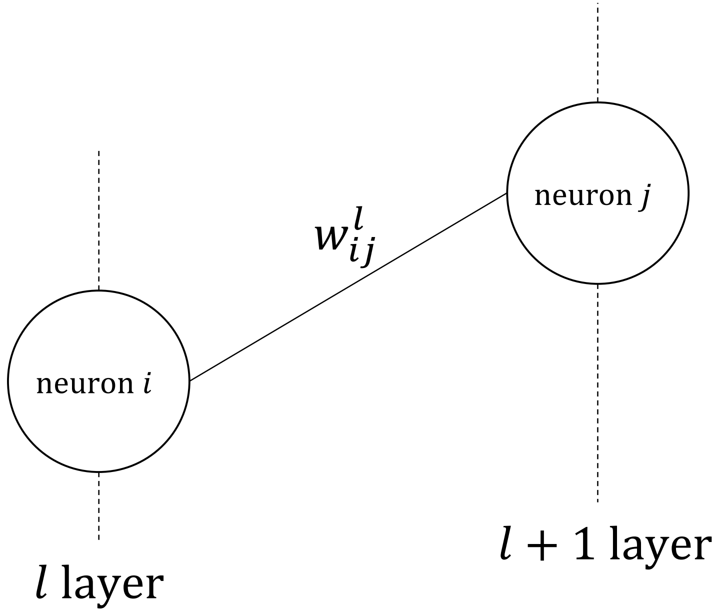
Learning in Neural Networks



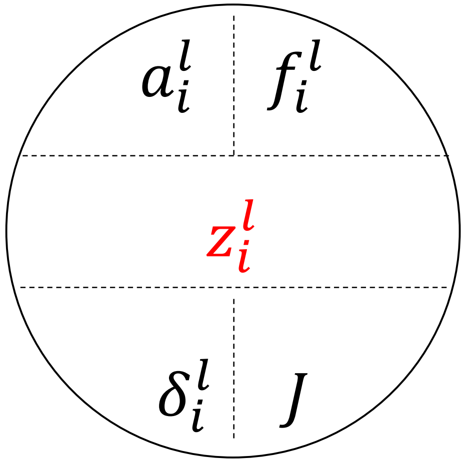
The network stores only knowledge, it does not store original data.

Learning is the updating of connection weight between two connected neurons

$$w_{ij}^{l}(new) \leftarrow w_{ij}^{l}(old)$$

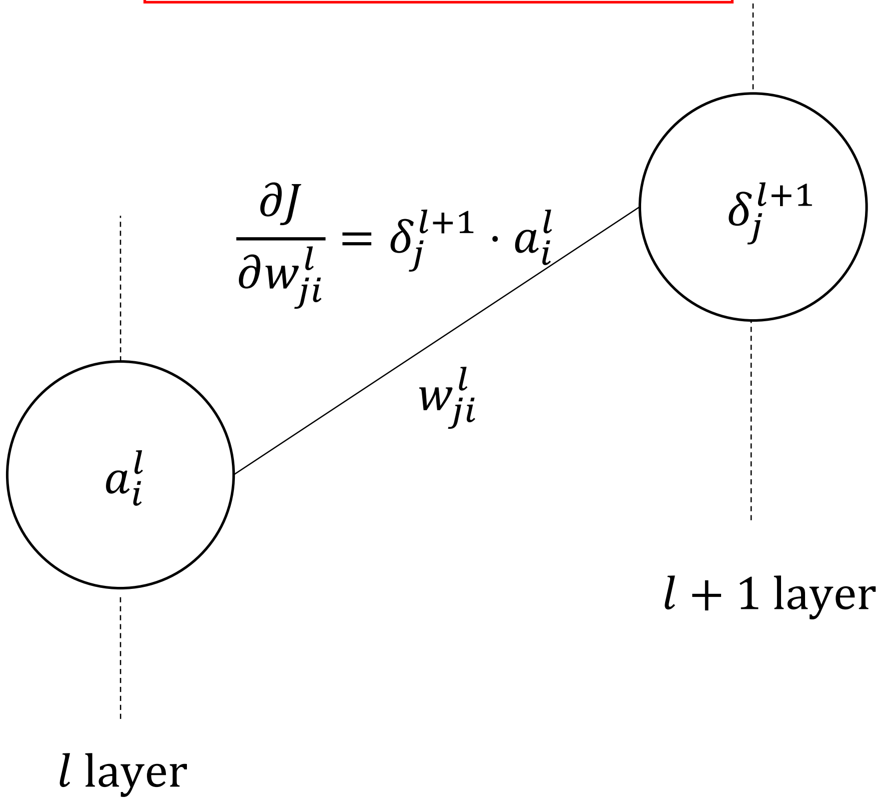


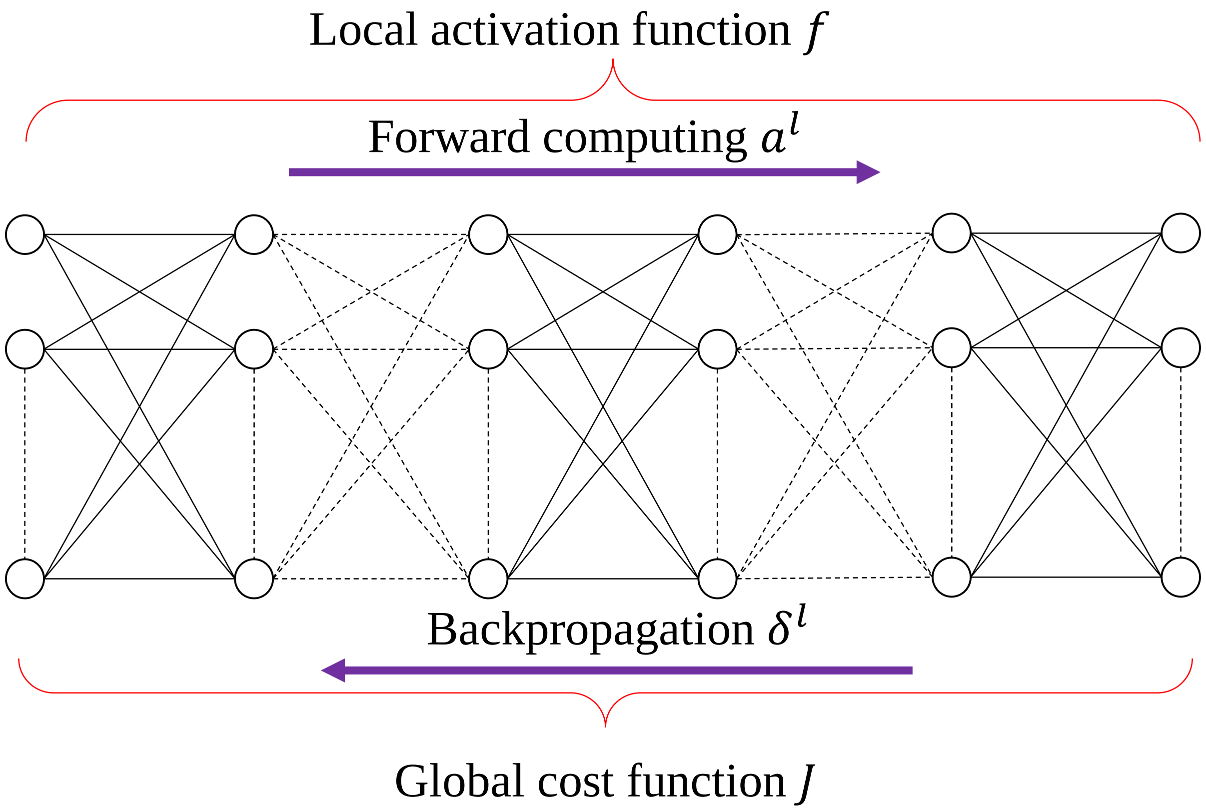
Learning in Neural Networks



BP Learning Algorithm

$$w_{ij}^l \leftarrow w_{ij}^l - \alpha \cdot (\delta_j^{l+1} \cdot a_i^l)$$





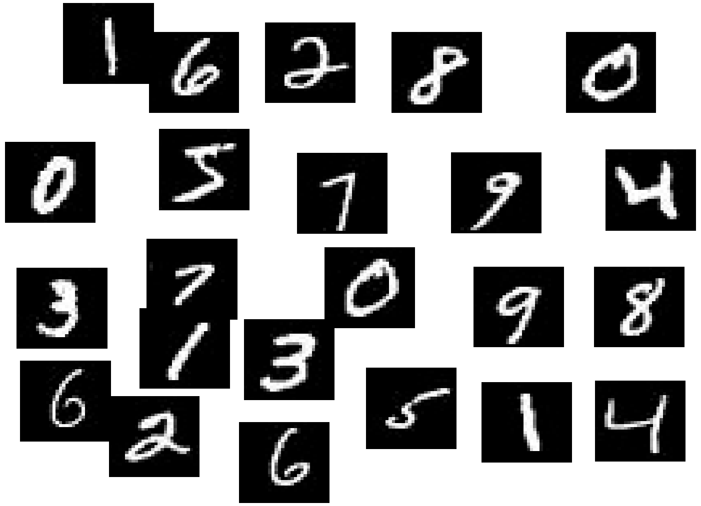
Outline

- Learning in Neural Networks
- Supervised Learning
- Unsupervised Learning
- Autoencoder Neural Networks
- Assignment

Supervised Learning

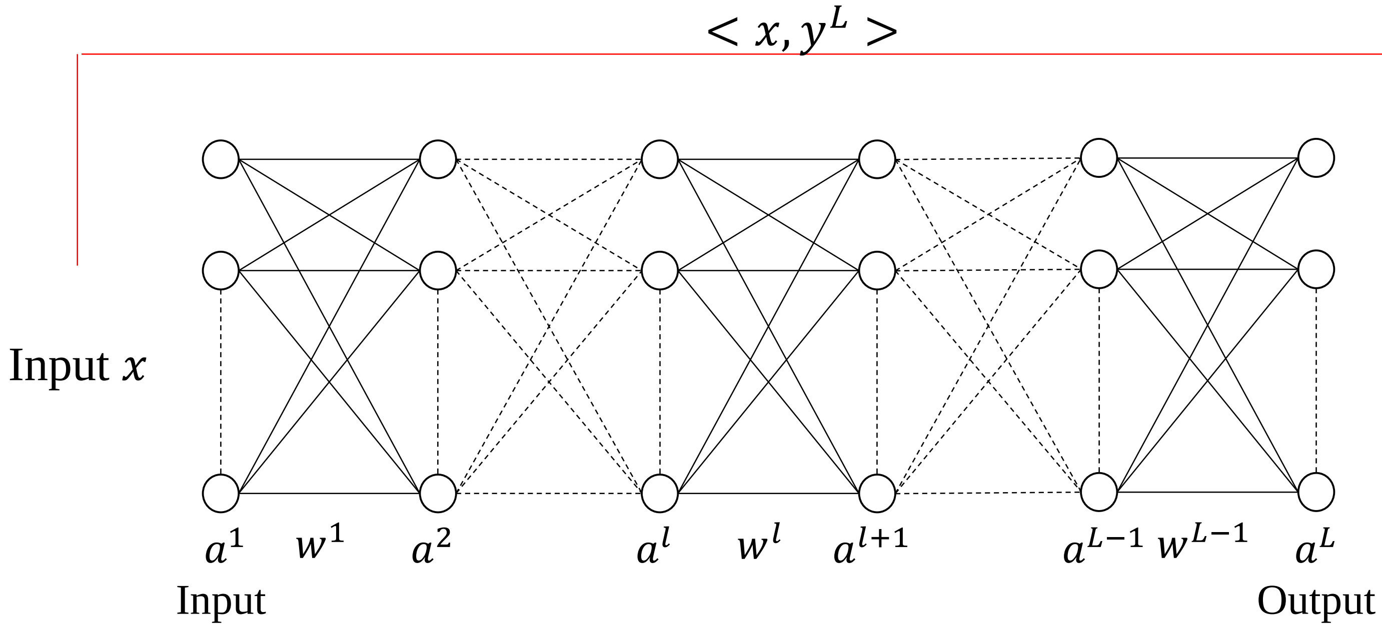
- Supervised Learning
  - Learning with a supervisor
  - The supervisor knows the correct answer
  - Each training sample must contain input and target

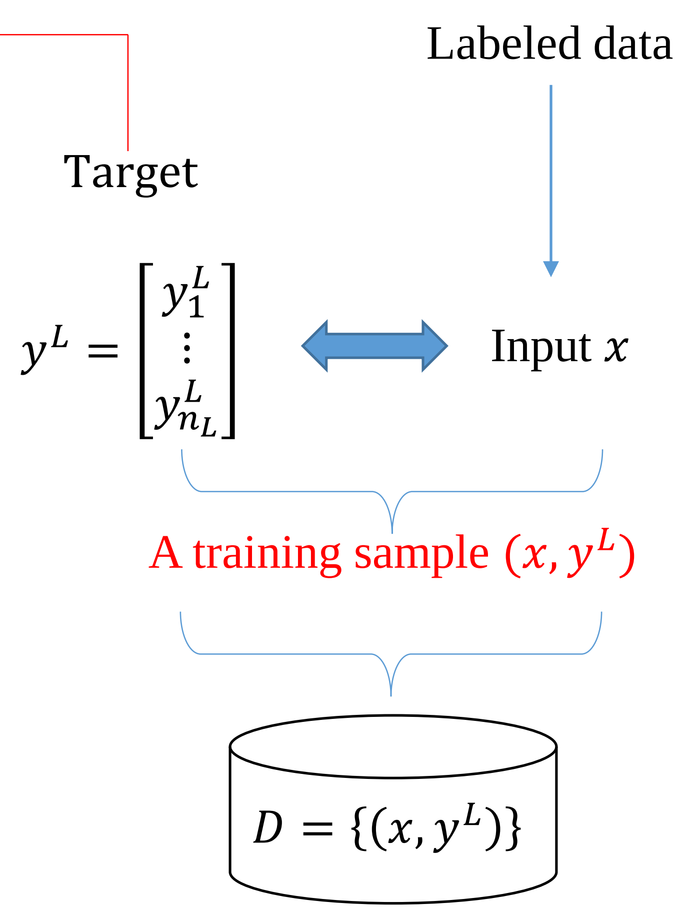
Learn from teacher



Learn from parent

Supervised Learning





In supervised learning, each training sample contains input and the associated target output.

With only training data is not sufficient, you need a learning algorithm to update the knowledge, i.e., updating the connection weights in the network.

A training data set is a set composed by training samples

Supervised Learning

The well known BP algorithm is a supervised learning algorithm.

$$D = \{(x, y^L)\}$$

Network prediction | Target

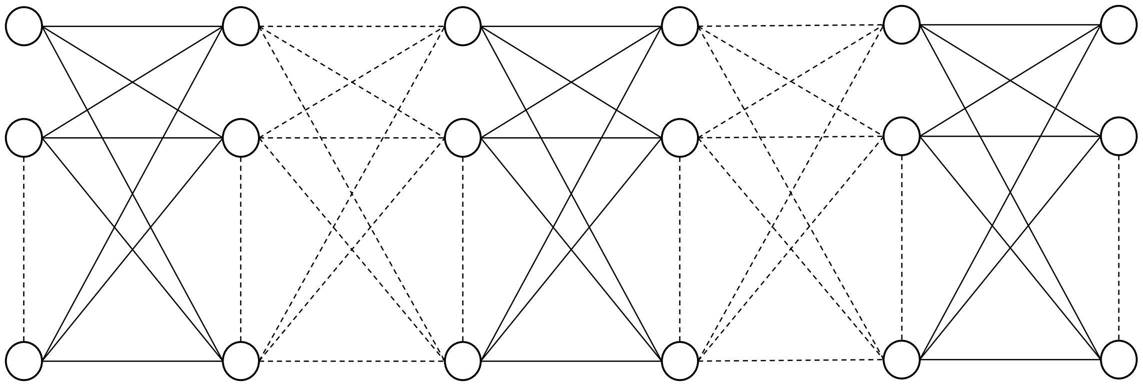
$$a^L = \begin{bmatrix} a_1^L \\ \vdots \\ a_{n_L}^L \end{bmatrix} \qquad y^L = \begin{bmatrix} y_1^L \\ \vdots \\ y_{n_L}^L \end{bmatrix}$$

Cost function (Energy Function)

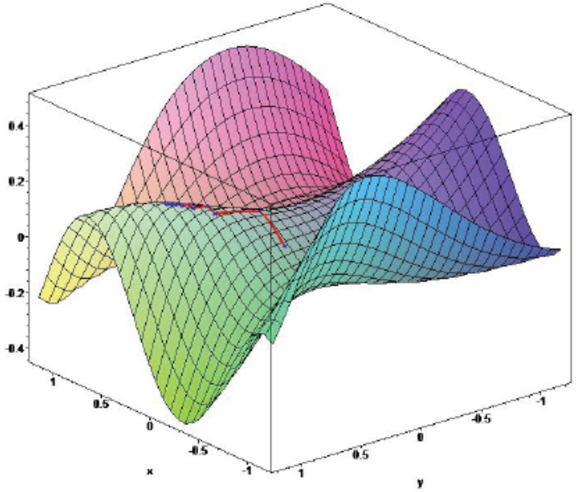
$$J = J(w^1, \cdots, w^{L-1})$$

$$\frac{\partial J}{\partial w_{ji}^l} = \delta_j^{l+1} \cdot a_i^l$$

Forward computing a^l



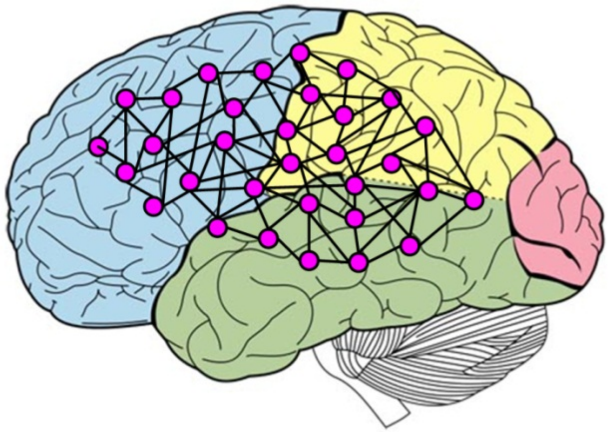
Backpropagation δ^l



Outline

- Learning in Neural Networks
- Supervised Learning
- Unsupervised Learning
- Autoencoder Neural Networks
- Assignment

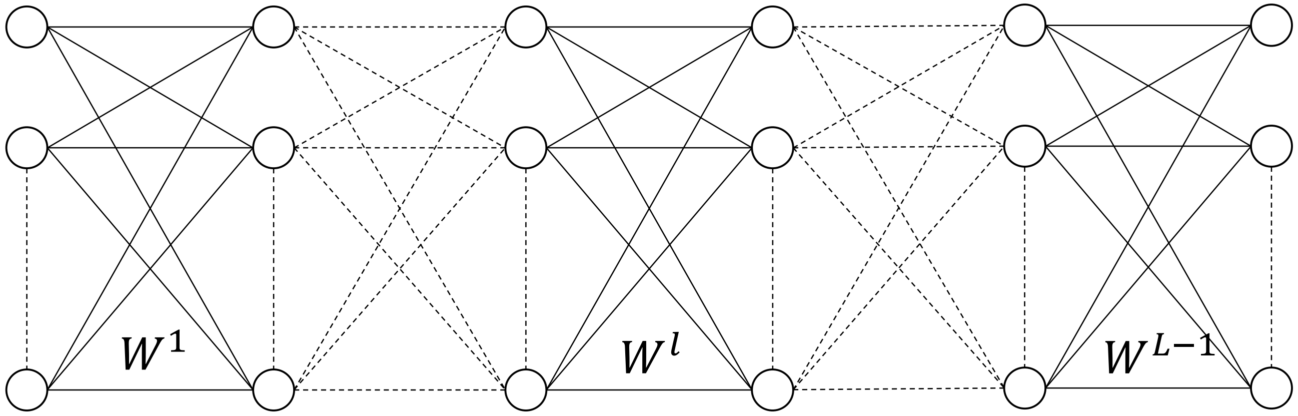
Unsupervised Learning



- Unsupervised Learning
  - Learning without supervisor
  - Each training sample do not have any explicit target

Neural Network

Input x



No target

Problem: How can I learn without a supervisor?

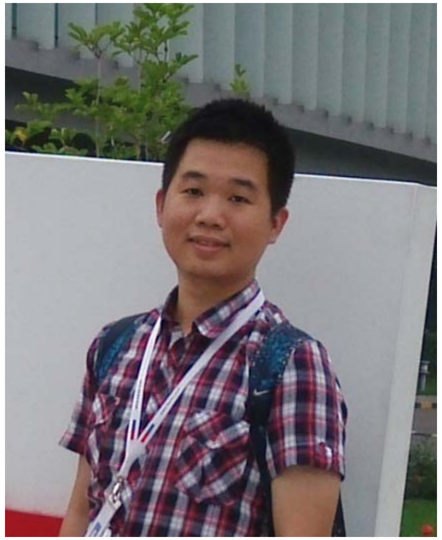
Unsupervised Learning: Feature Extraction

Can we learn something without a supervisor?

What's this?!

No body can tell the student about this one.

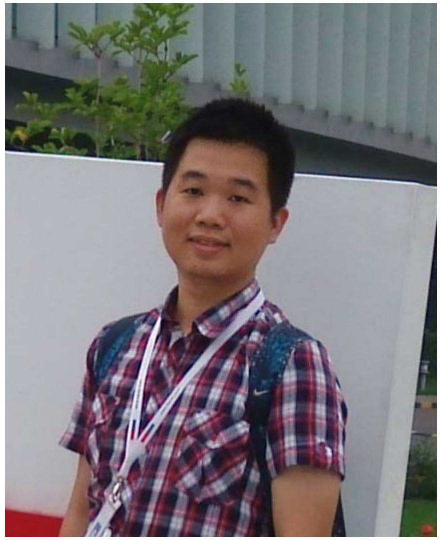
Unsupervised Learning: Feature Extraction

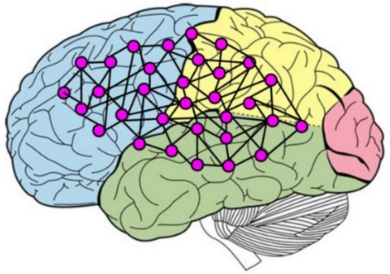


Haha, I saw it yesterday.

Unsupervised Learning: Feature Extraction

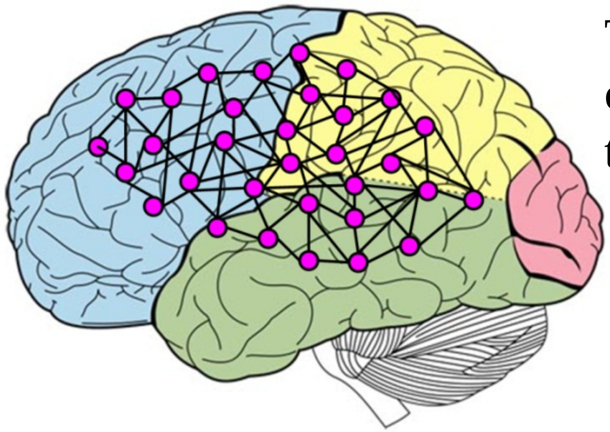
They are not the same one. The student successfully learnt the features of this unknown fruit.





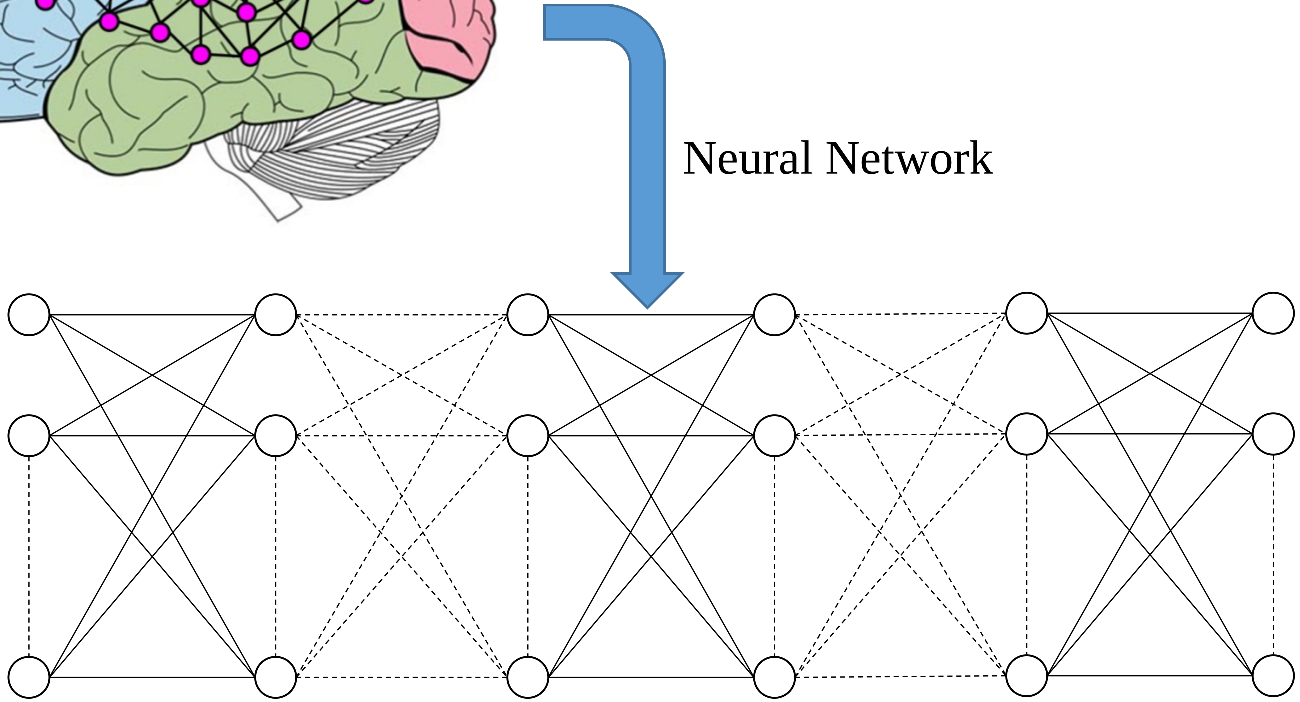
Without a teacher, how can the student learn that the fruits in these pictures are in fact the same? There must exist some learning mechanism in his brain neural network?

Unsupervised Learning: Feature Extraction



There must exist some learning mechanism can do feature extraction without supervisor in the brain neural network.

Input x



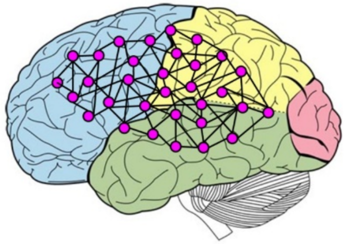
No target

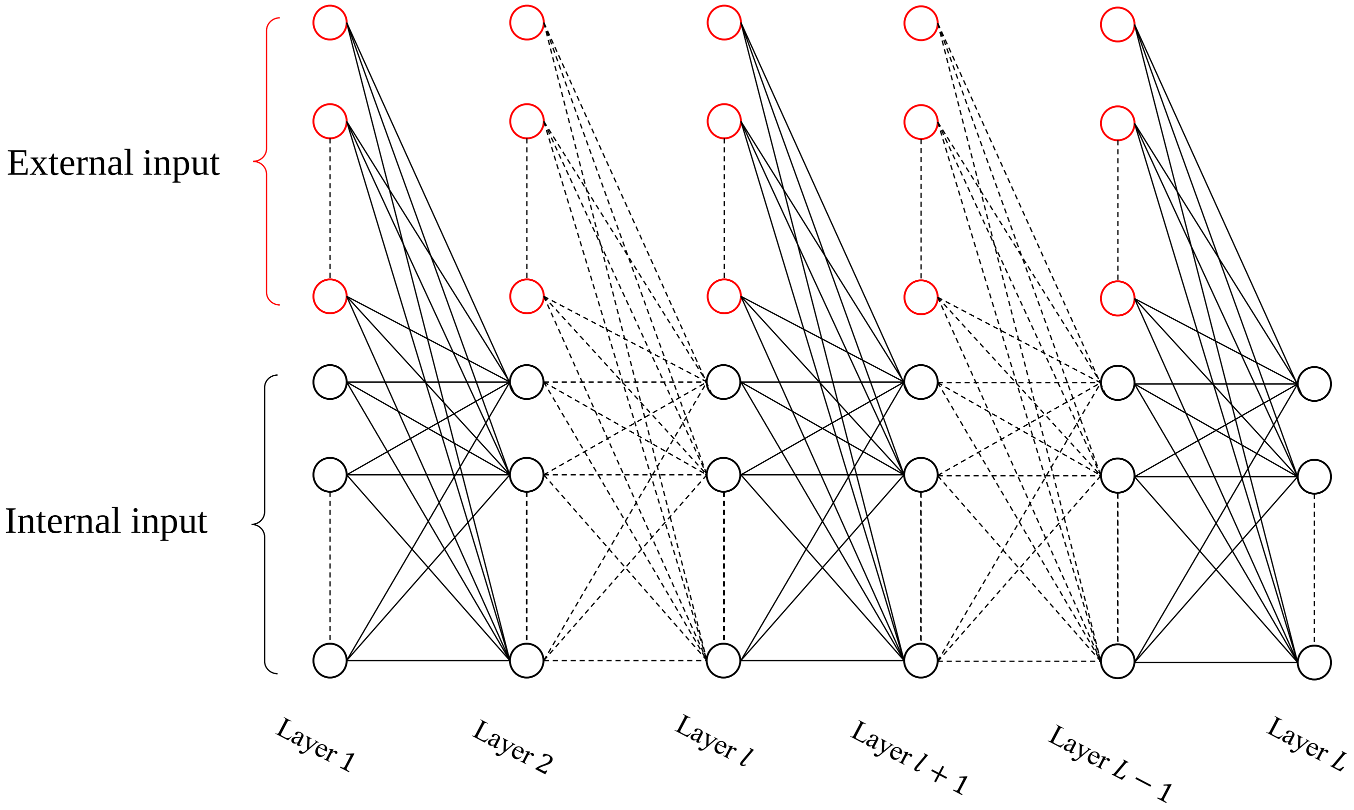
Problem: How to develop algorithms such that an artificial network can learn without any supervisor?

Outline

- Learning in Neural Networks
- Supervised Learning
- Unsupervised Learning
- Autoencoder Neural Networks
  - The Network Structure
  - Compressive Representation
  - Sparse Representation
  - Feature Learning
  - Application to Supervised Learning
- Assignment

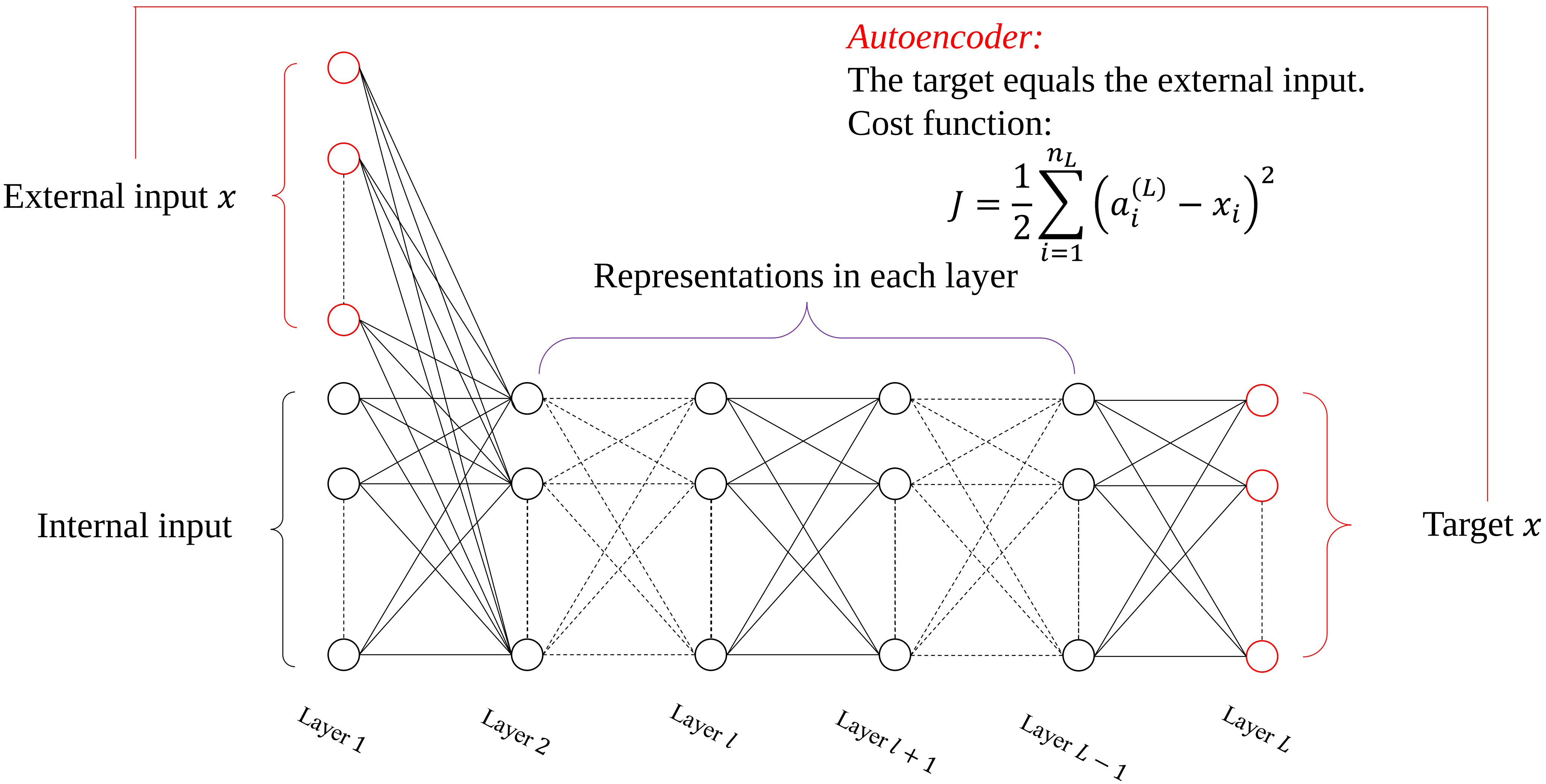
Autoencoder Neural Networks



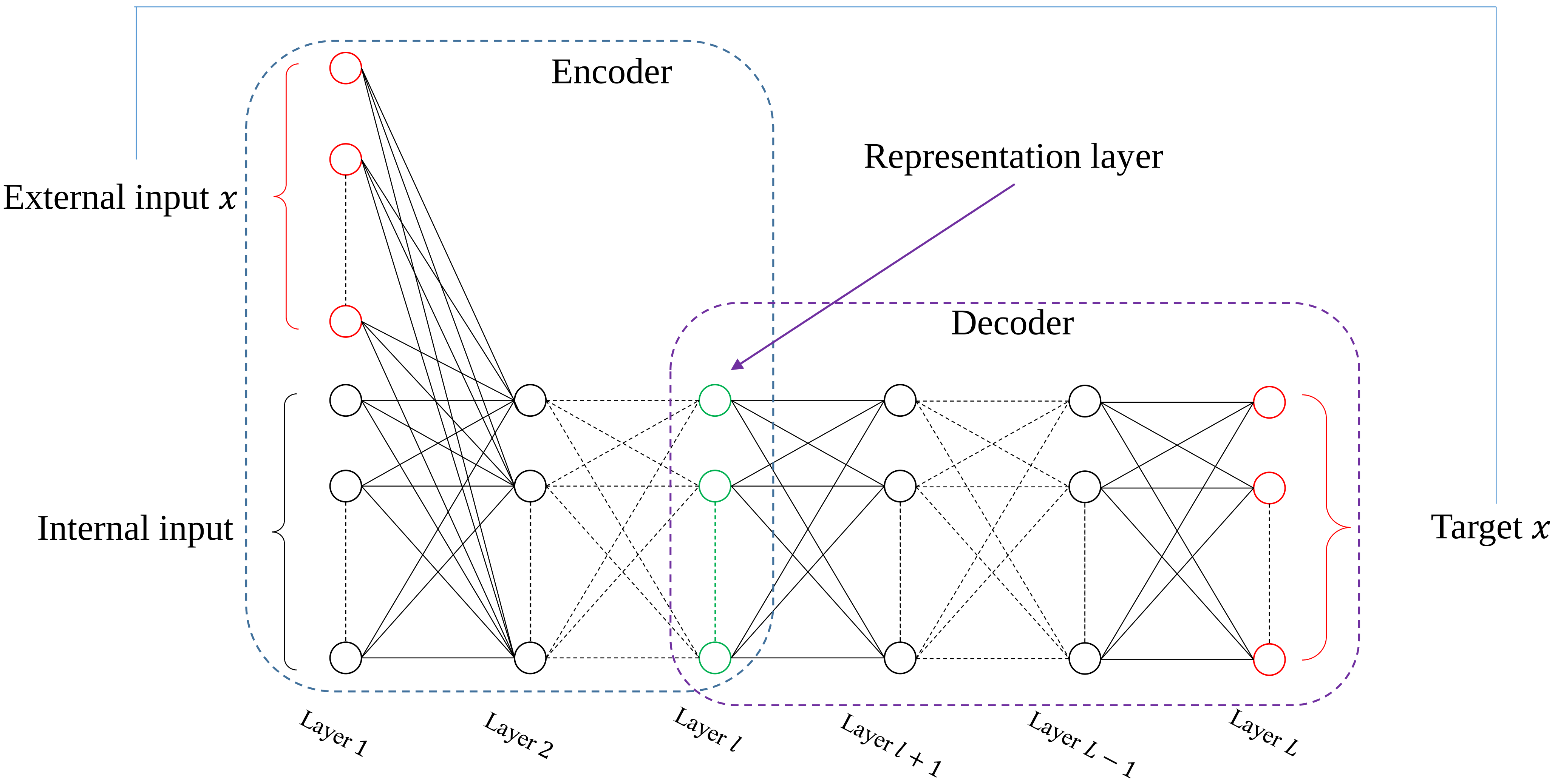


The well-known network structure

Autoencoder Neural Networks



Autoencoder Neural Networks



Autoencoder Neural Networks

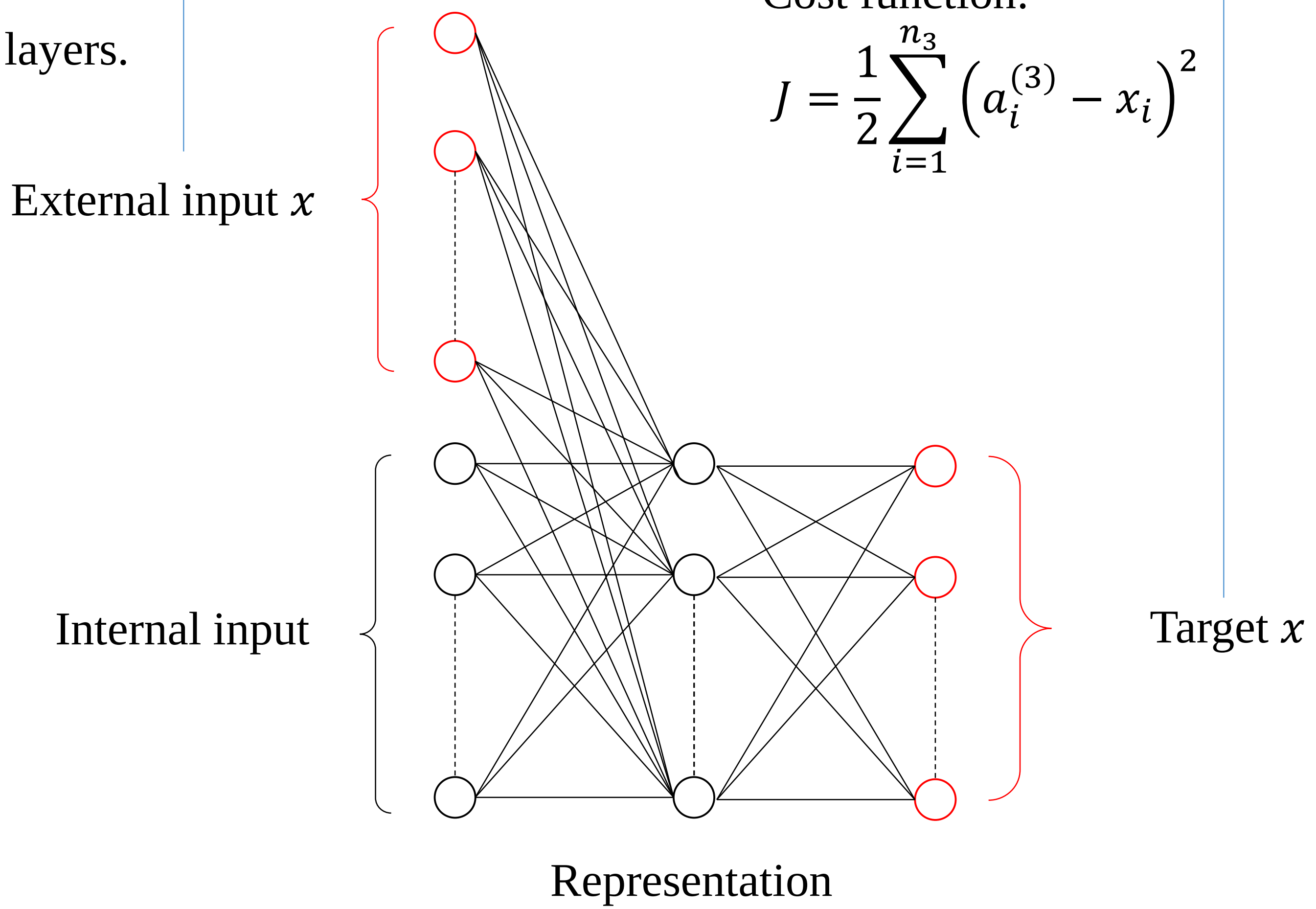
Simplest Autoencoder:
The simplest autoencoder contains three layers.

$$a_i^{(2)} = f\left(\sum_{j=1}^{n_1} w_{ij}^{(1)} a_j^{(1)}\right)$$

$$a_i^{(3)} = f\left(\sum_{j=1}^{n_2} w_{ij}^{(2)} a_j^{(2)}\right)$$

Cost function:

$$J = \frac{1}{2}\sum_{i=1}^{n_3}\left(a_i^{(3)} - x_i\right)^2$$

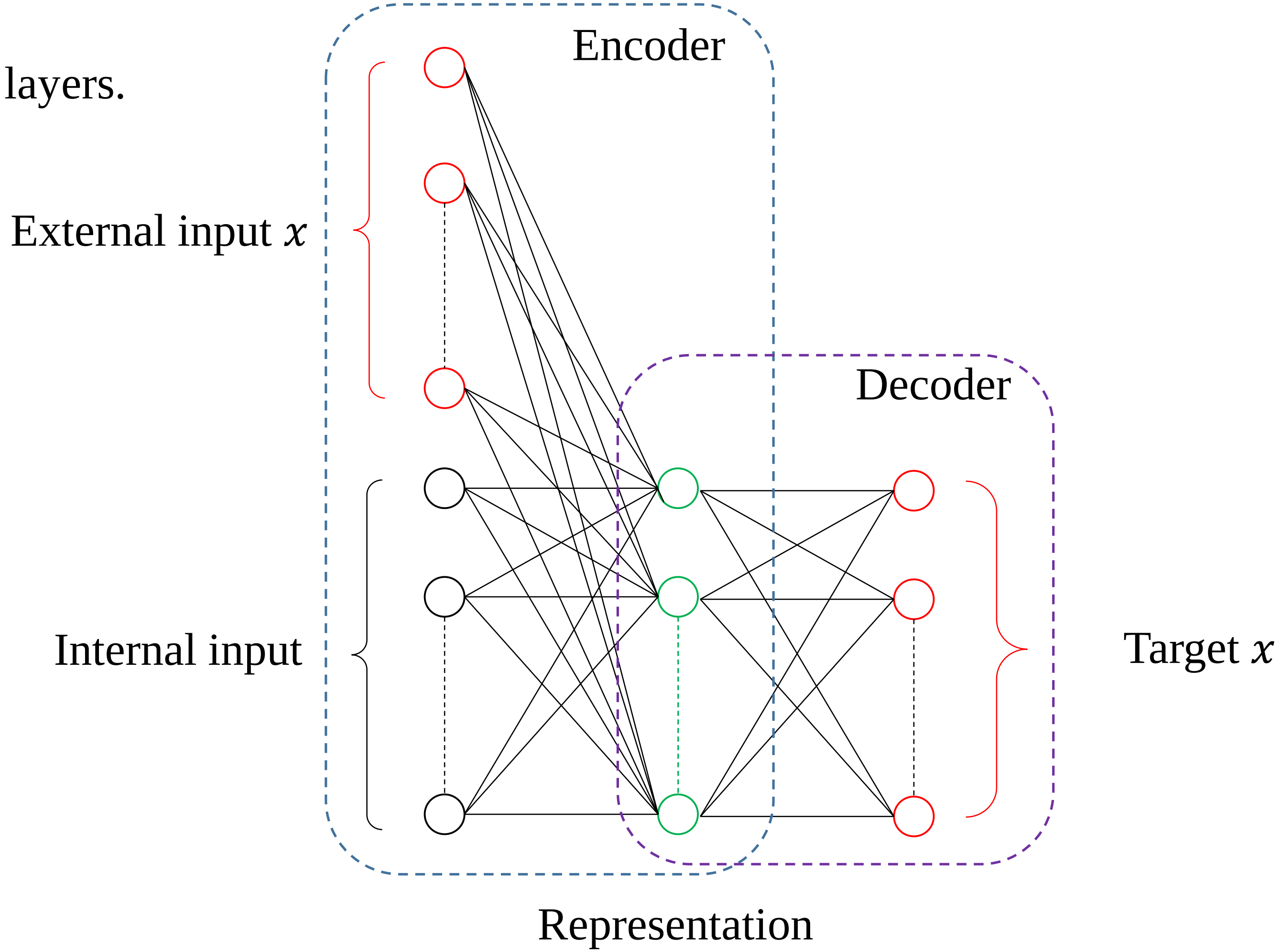


Autoencoder Neural Networks

Simplest Autoencoder:
The simplest autoencoder contains three layers.

$$a_i^{(2)} = f\left(\sum_{j=1}^{n_1} w_{ij}^{(1)} a_j^{(1)}\right)$$

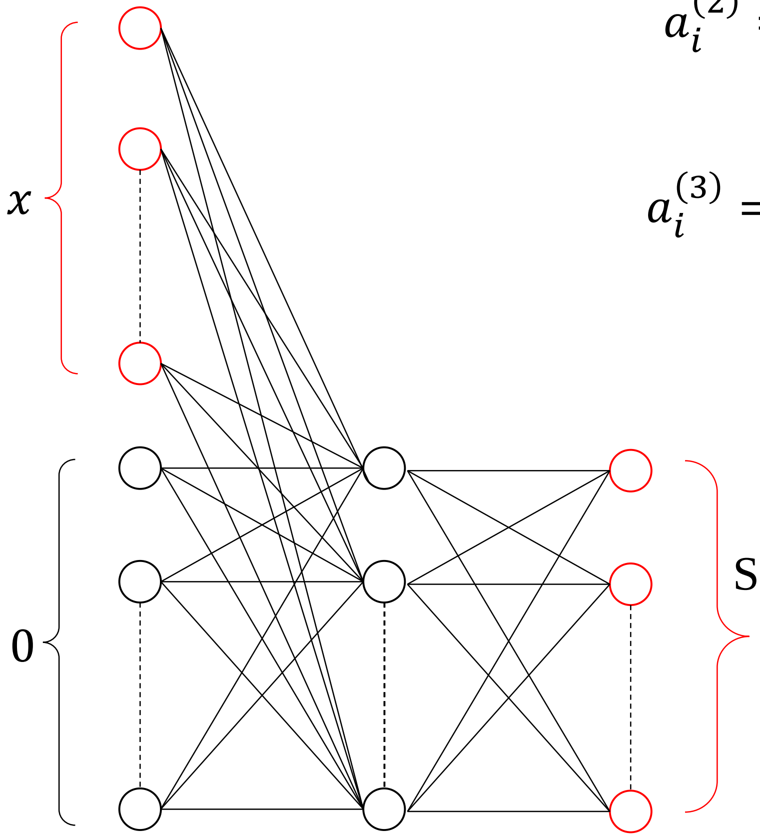
$$a_i^{(3)} = f\left(\sum_{j=1}^{n_2} w_{ij}^{(2)} a_j^{(2)}\right)$$



Autoencoder Neural Networks

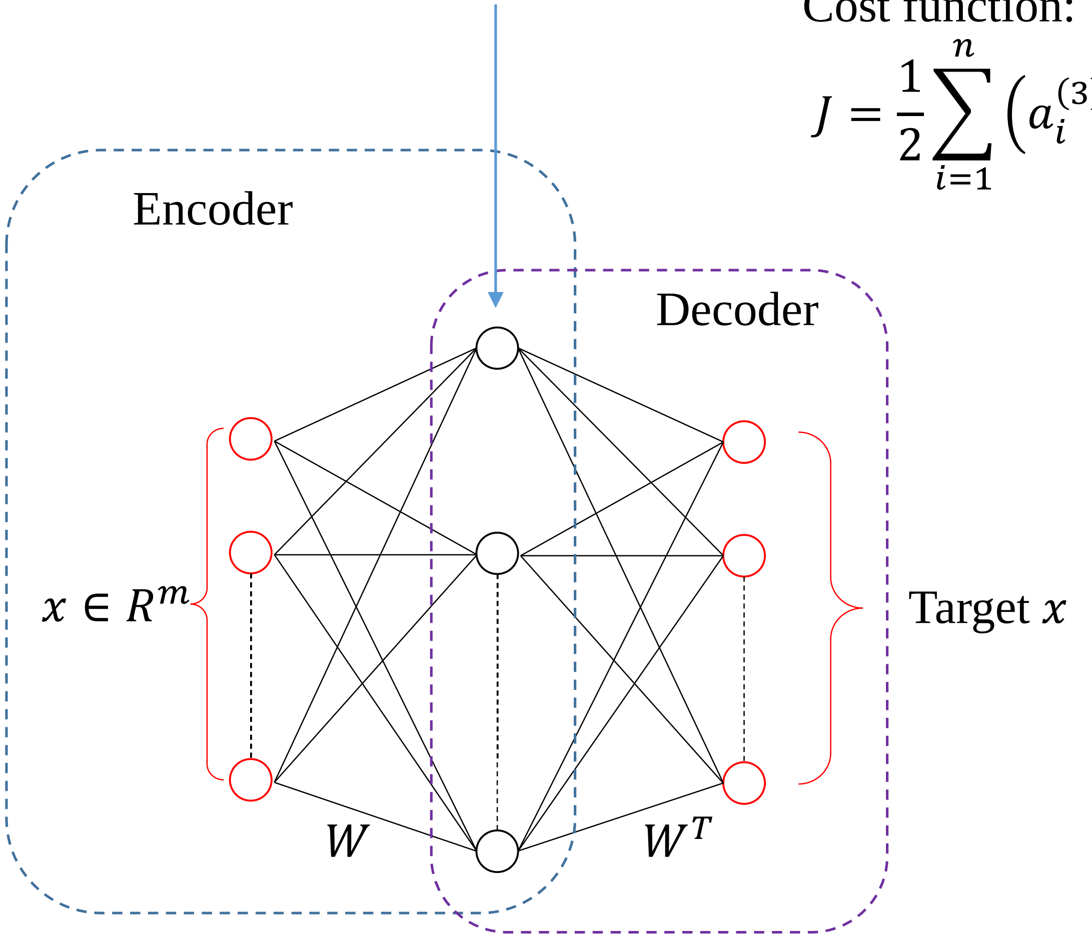
$$a_i^{(2)} = f\left(\sum_{j=1}^{m} w_{ij} x_j\right)$$

$$a_i^{(3)} = f\left(\sum_{j=1}^{n} w_{ji} a_j^{(2)}\right)$$



Simplest Autoencoder

Representation in R^n



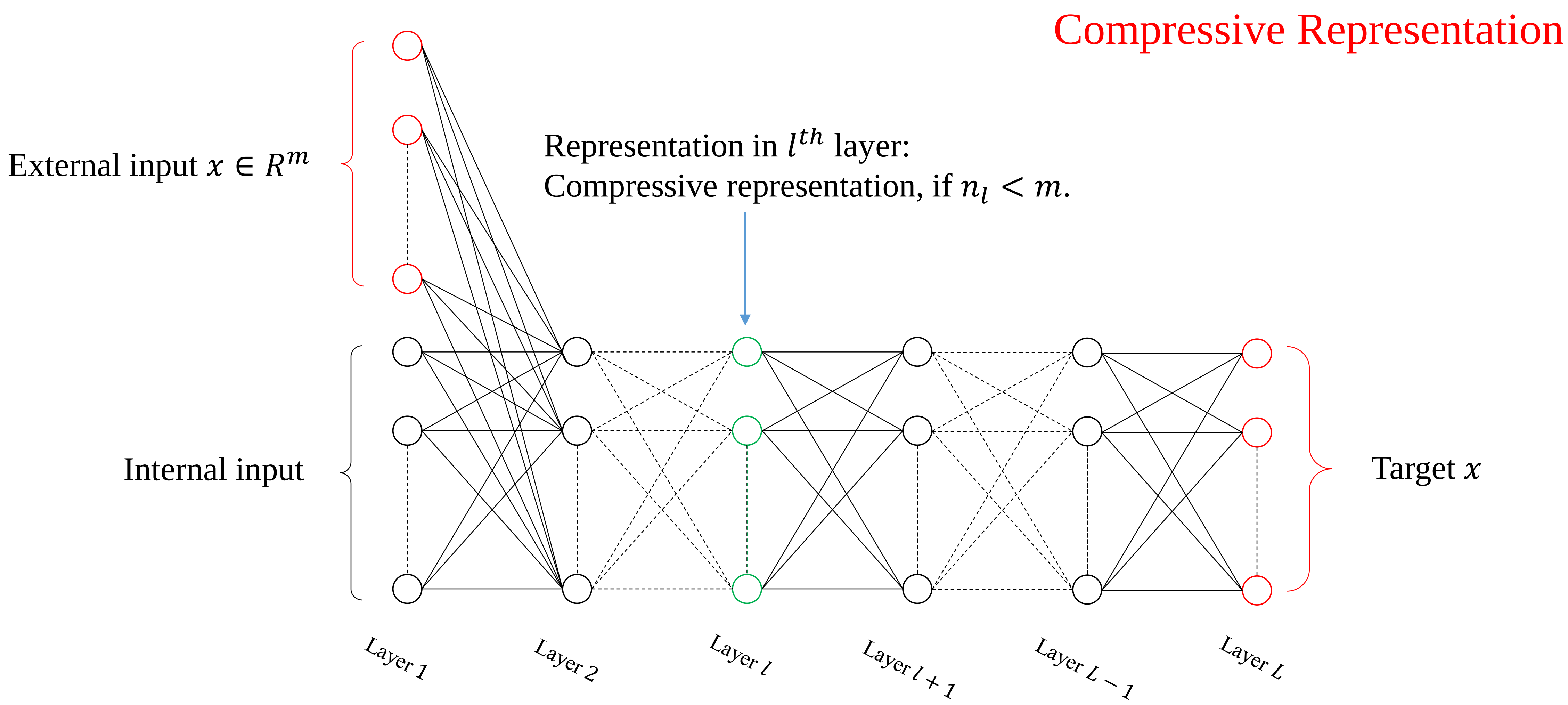
Cost function:

$$J = \frac{1}{2}\sum_{i=1}^{n}\left(a_i^{(3)} - x_i\right)^2$$

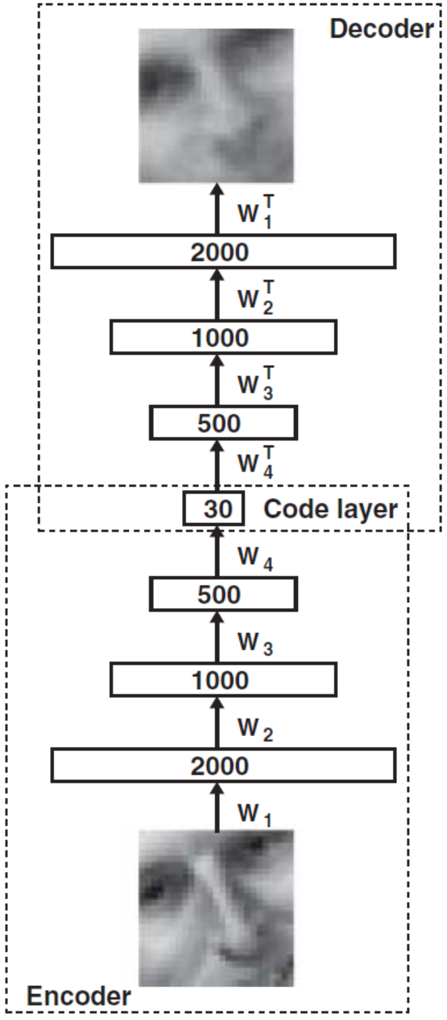
Outline

- Learning in Neural Networks
- Supervised Learning
- Unsupervised Learning
- Autoencoder Neural Networks
  - The Network Structure
  - Compressive Representation
  - Sparse Representation
  - Feature Learning
  - Application to Supervised Learning
- Assignment

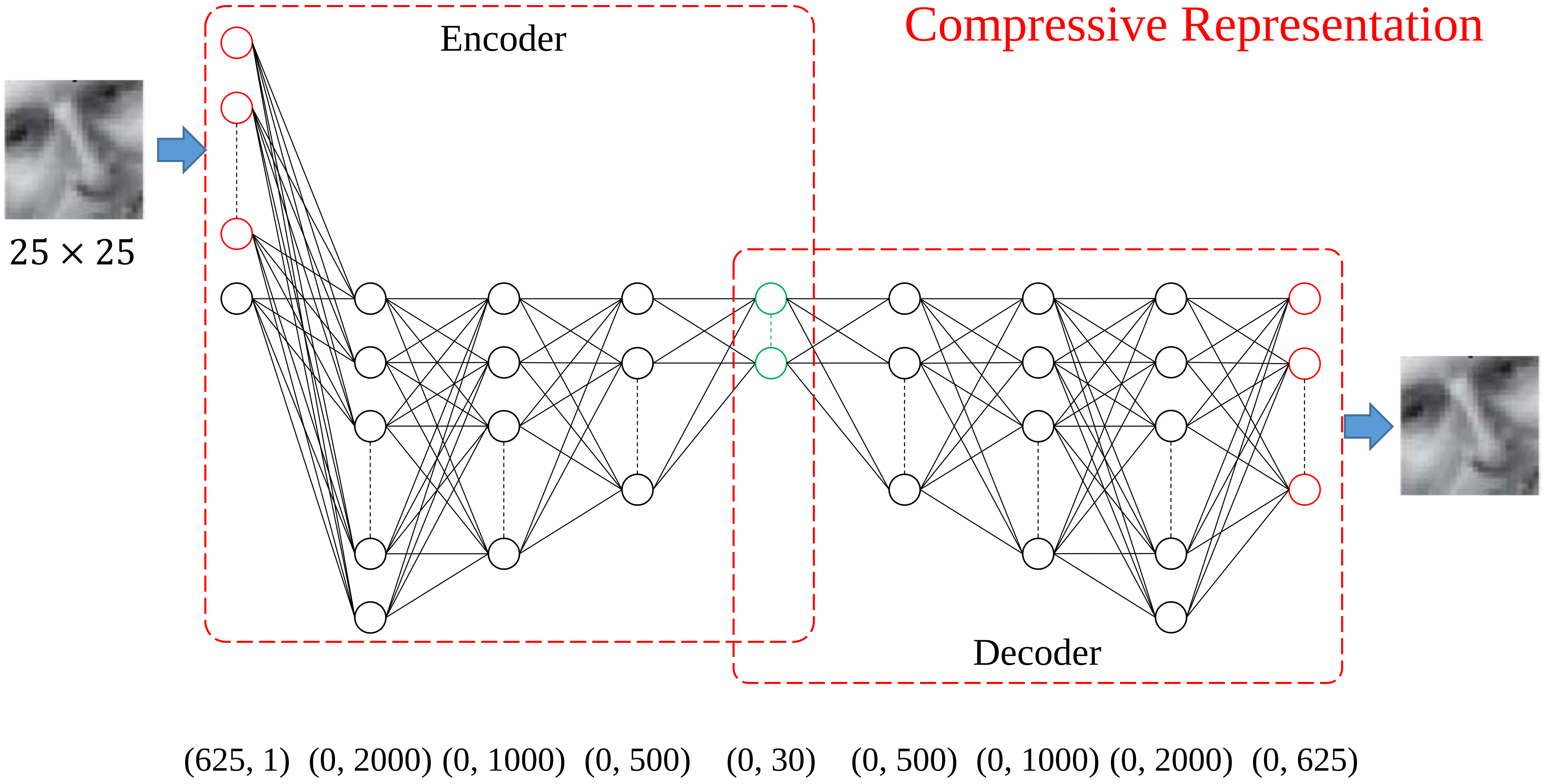
Autoencoder Neural Networks



Autoencoder Neural Networks



Unrolling

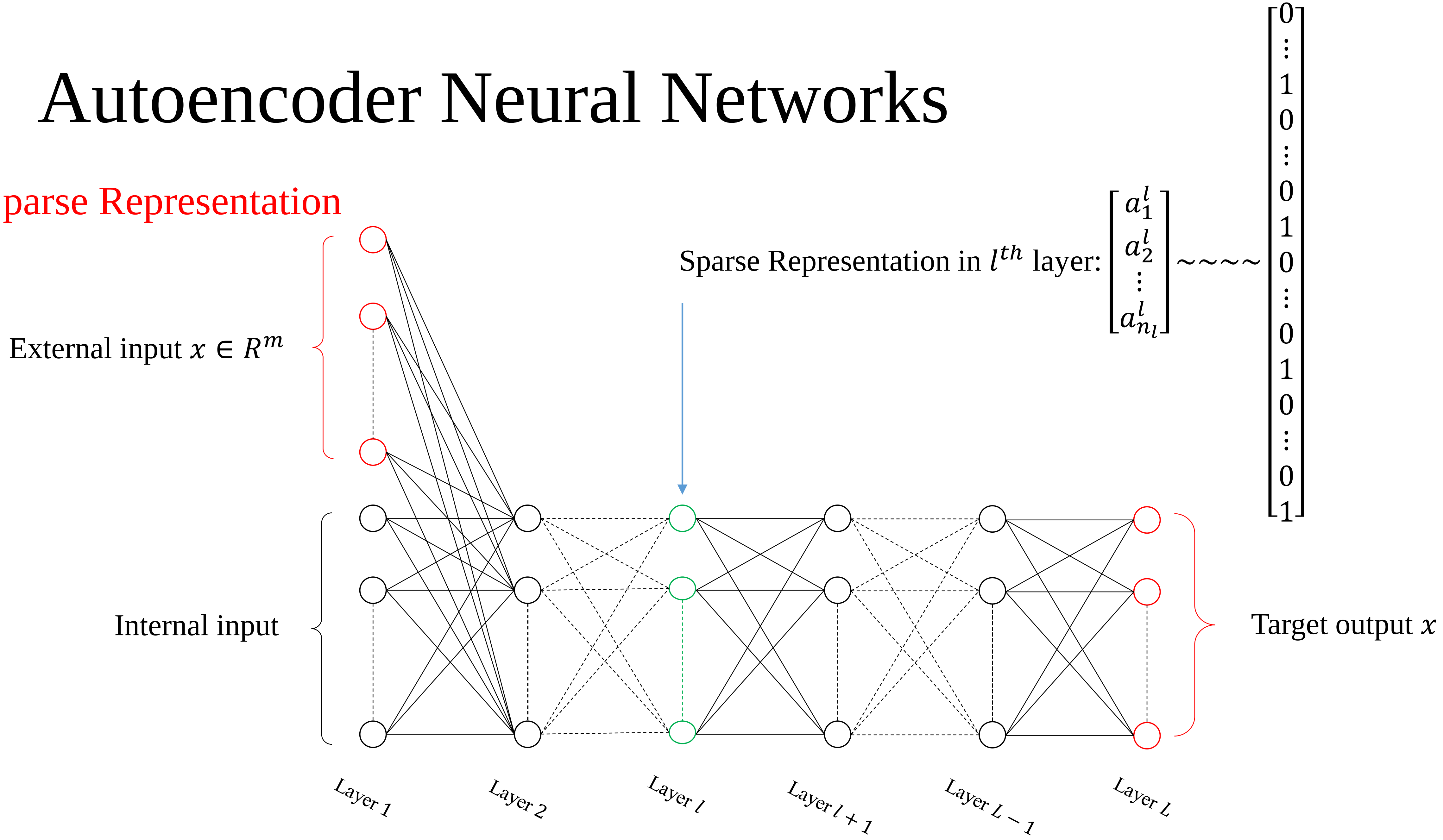


Outline

- Learning in Neural Networks
- Supervised Learning
- Unsupervised Learning
- Autoencoder Neural Networks
  - The Network Structure
  - Compressive Representation
  - Sparse Representation
  - Feature Learning
  - Application to Supervised Learning
- Assignment

Autoencoder Neural Networks

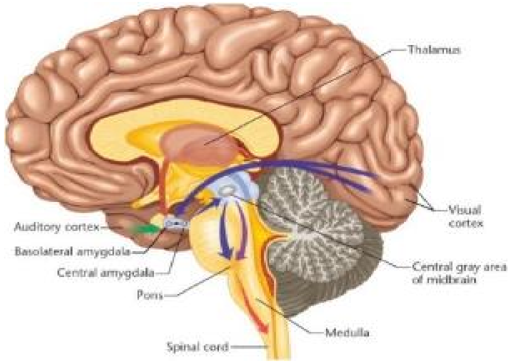
Sparse Representation

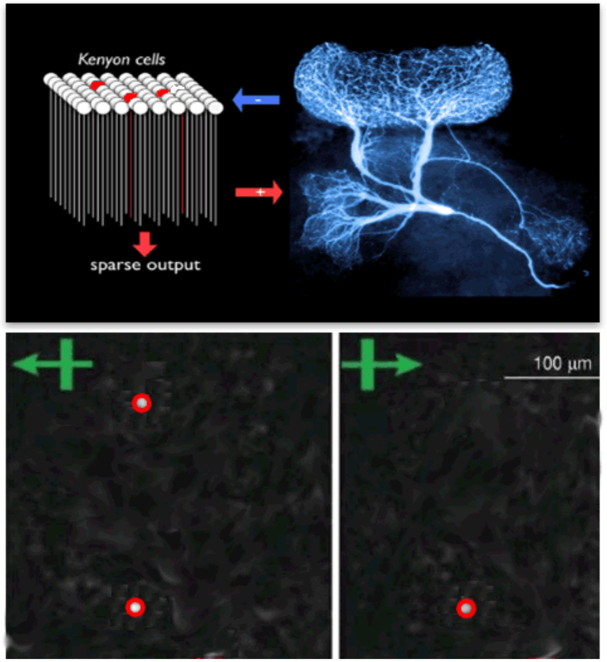


Autoencoder Neural Networks

Sparse Representation

The brain represents information in sparsity way.





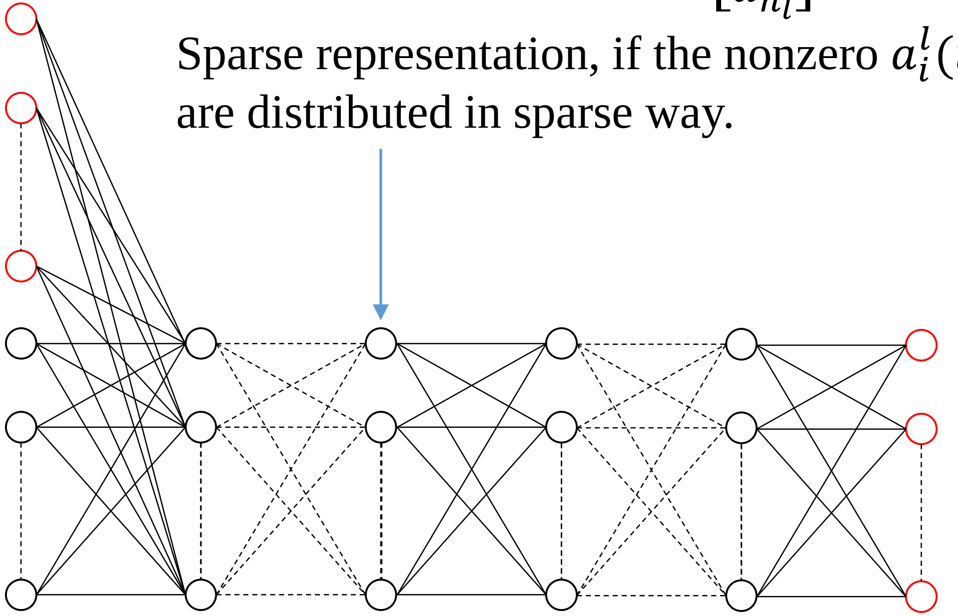
nature

Emergence of simple-cell receptive field properties by learning a sparse code for natural images
Bruno A. Olshausen and David J. Field
Nature **381**, 607 - 609 (13 June 1996);
doi:10.1038/381607a0

Problem:
How to get sparse representation?

Representation in l^{th} layer: $\begin{bmatrix} a_1^l \\ a_2^l \\ \vdots \\ a_{n_l}^l \end{bmatrix}$

Sparse representation, if the nonzero $a_i^l(i = 1, \cdots, n_l)$ are distributed in sparse way.



Autoencoder Neural Networks

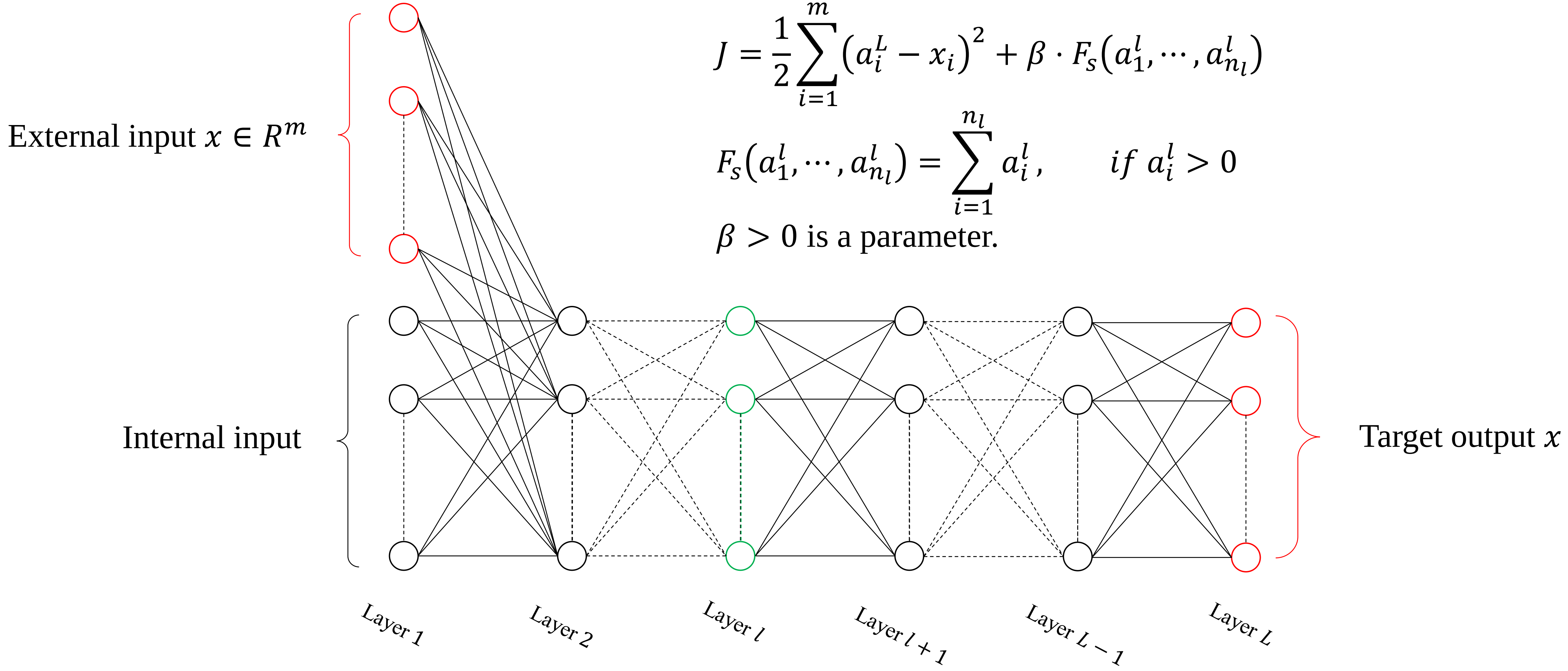
Sparse Representation

Cost function:

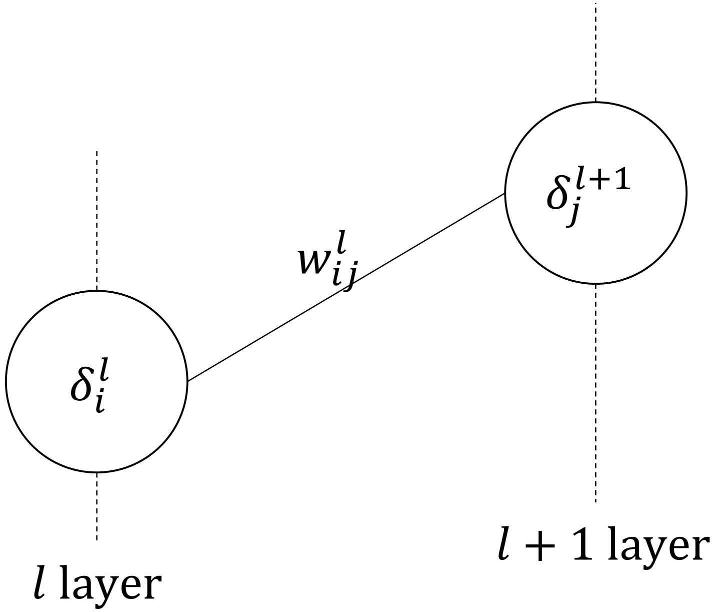
$$J = \frac{1}{2}\sum_{i=1}^{m}\left(a_i^L - x_i\right)^2 + \beta \cdot F_s\left(a_1^l, \cdots, a_{n_l}^l\right)$$

$$F_s\left(a_1^l, \cdots, a_{n_l}^l\right) = \sum_{i=1}^{n_l} a_i^l, \qquad if\ a_i^l > 0$$

$\beta > 0$ is a parameter.



Autoencoder Neural Networks



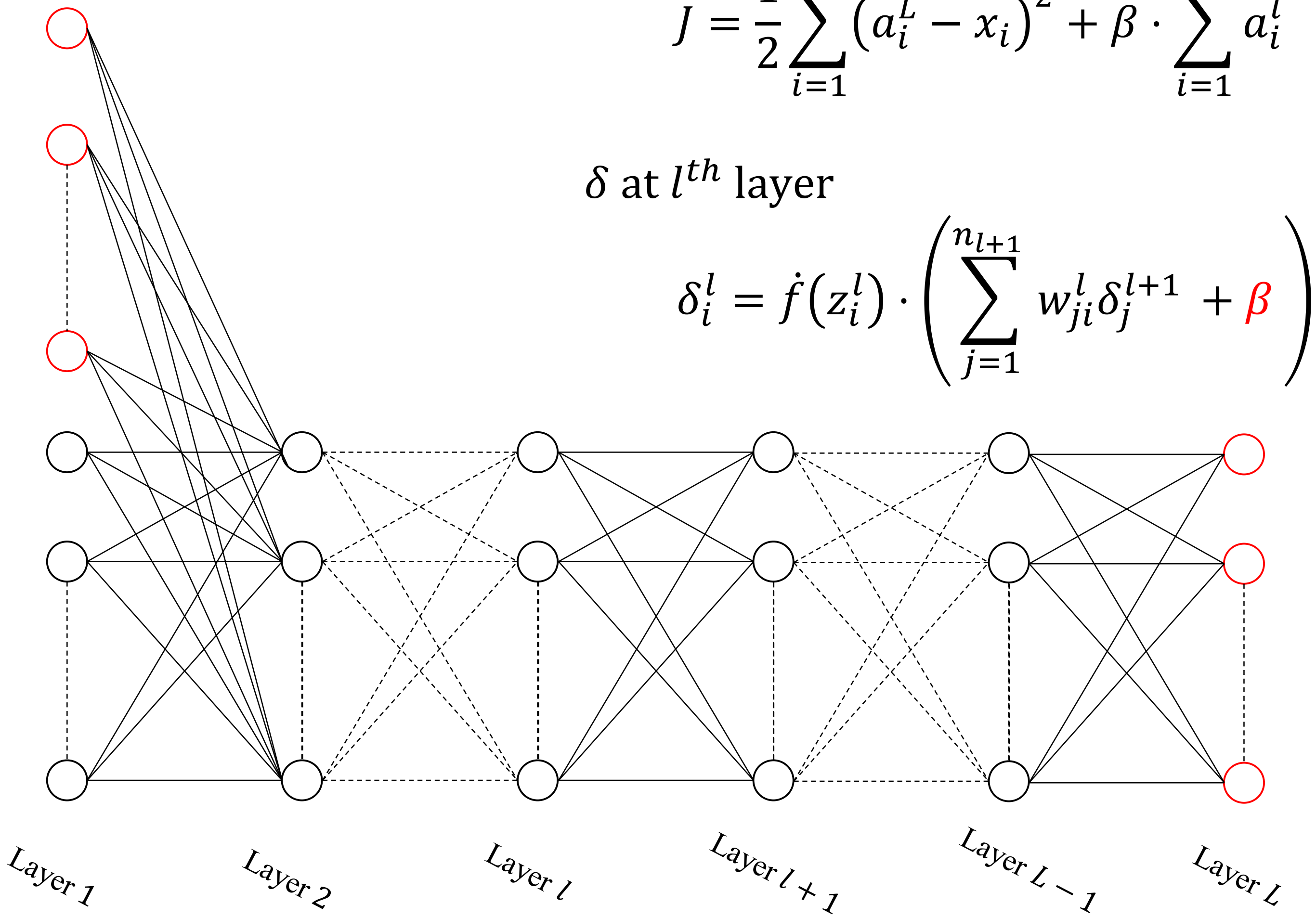
Sparse Representation

Cost function

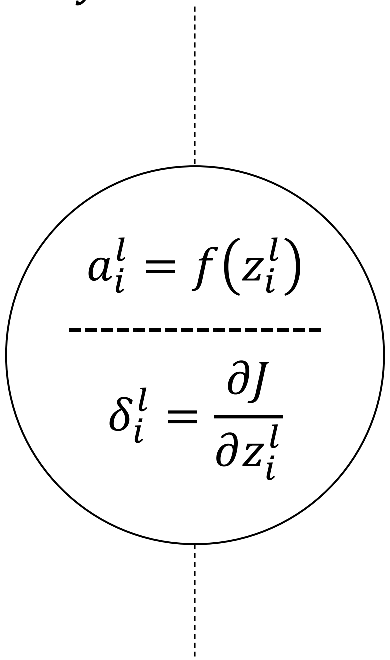
$$J = \frac{1}{2}\sum_{i=1}^{m}\left(a_i^L - x_i\right)^2 + \beta \cdot \sum_{i=1}^{n_l} a_i^l$$

δ at l^{th} layer

$$\delta_i^l = \dot{f}\left(z_i^l\right) \cdot \left(\sum_{j=1}^{n_{l+1}} w_{ji}^l \delta_j^{l+1} + \beta\right)$$



l layer ith *neuron*

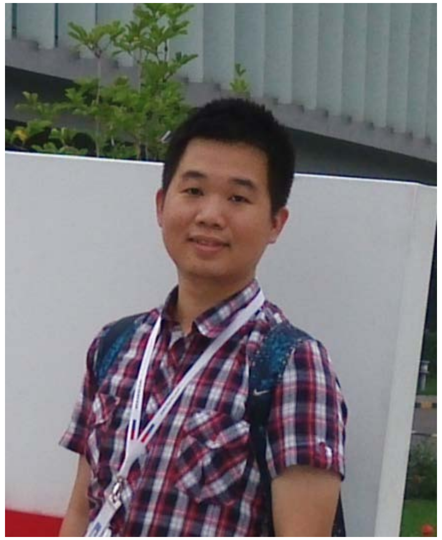


Outline

- Learning in Neural Networks
- Supervised Learning
- Unsupervised Learning
- Autoencoder Neural Networks
  - The Network Structure
  - Compressive Representation
  - Sparse Representation
  - Feature Learning
  - Application to Supervised Learning
- Assignment

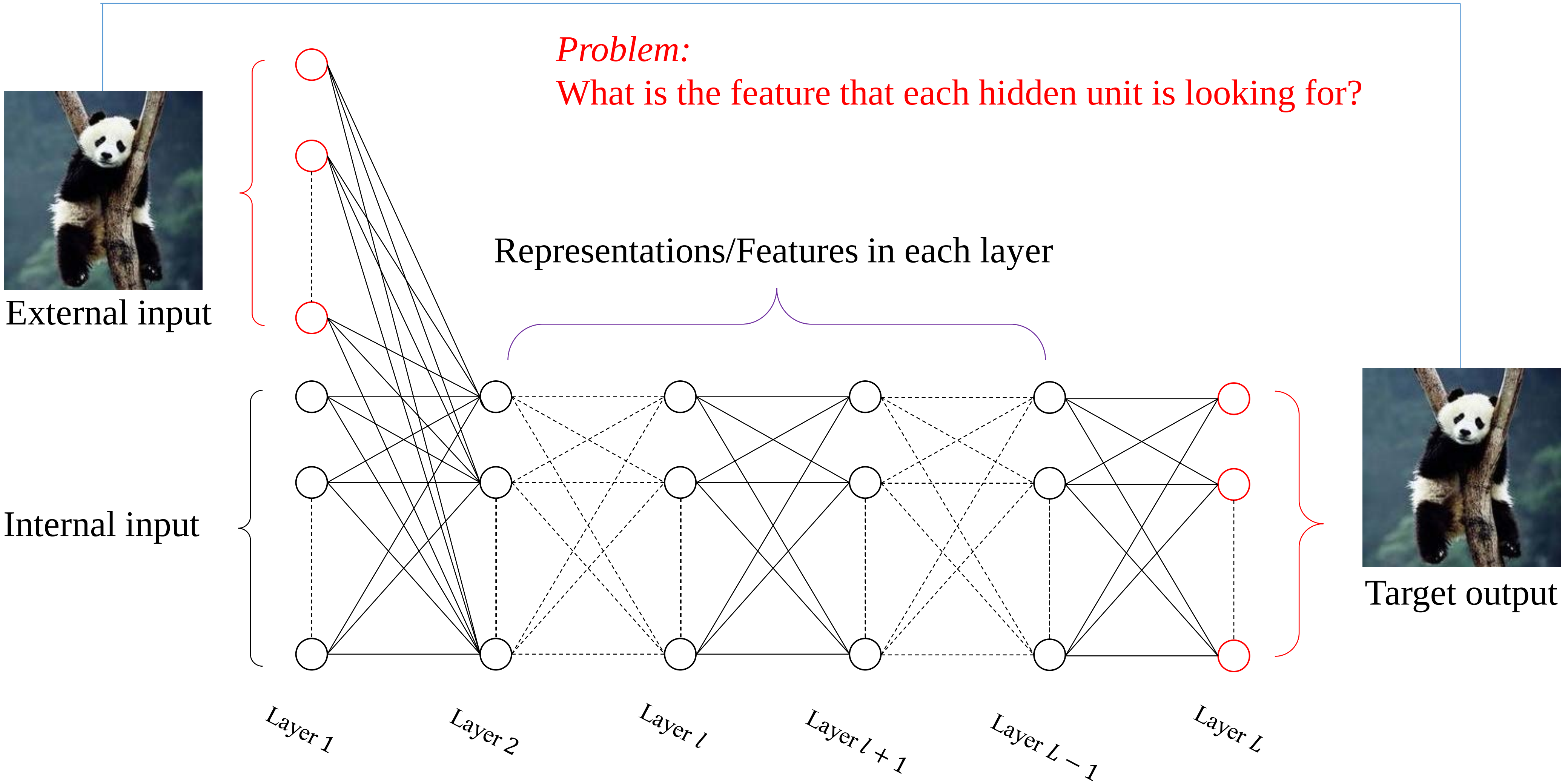
Autoencoder Neural Networks

They are not the same one. The student successfully learnt the features of this unknown fruit.



Without a teacher, how can the student learn that the fruits in these pictures are in fact the same? There must exist some learning mechanism in his brain neural network?

Autoencoder Neural Networks



Autoencoder Neural Networks

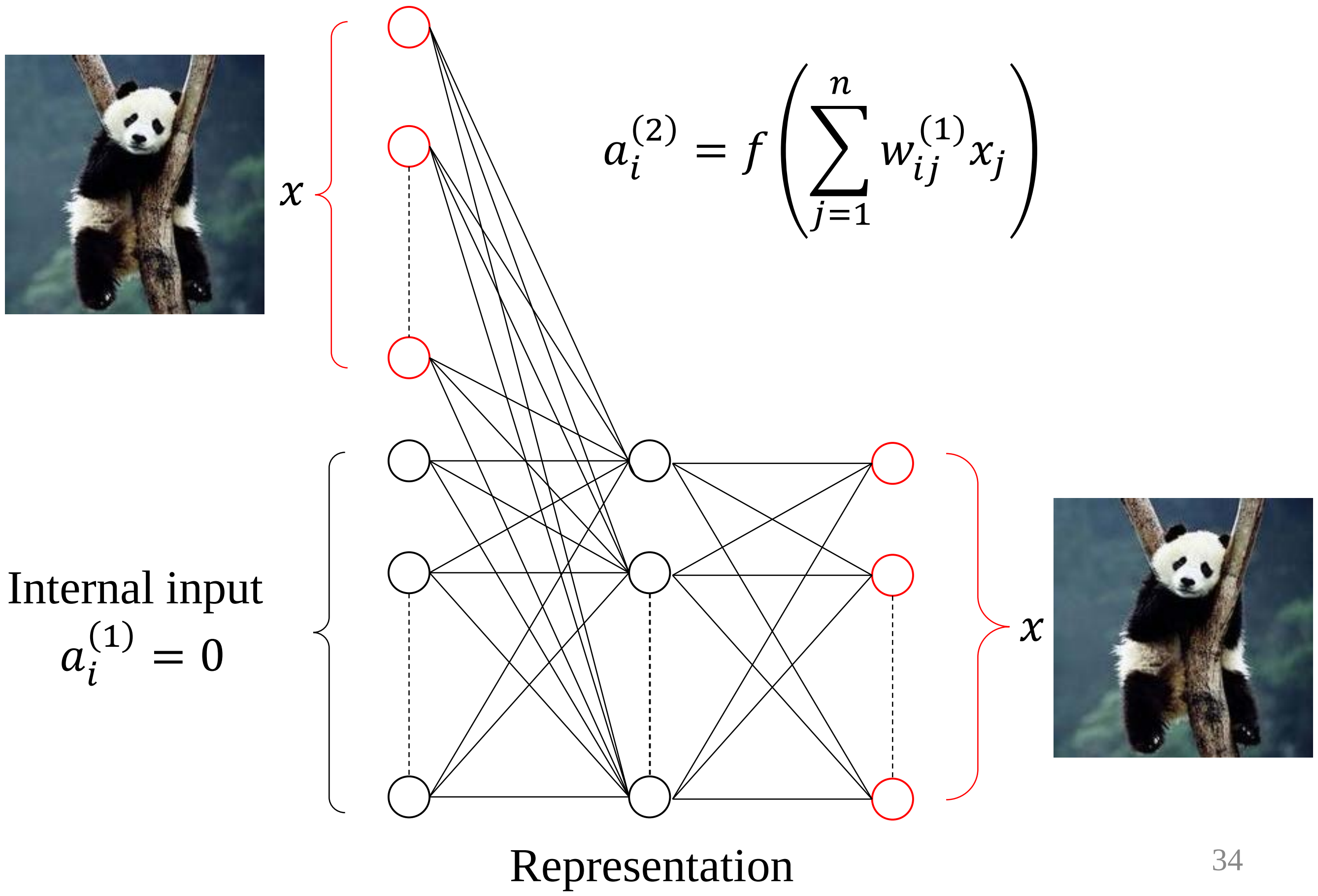
Problem:

What is the feature that hidden neuron i is looking for?

In other words, what input image x cause $a_i^{(2)}$ to be maximally activated?

$$\begin{cases} \max \sum_{i=1}^{n} w_{ij}^{(1)} x_j \\ s.t. \sum_{i=1}^{n} x_i^2 \leq 1 \end{cases}$$

$$x_j = \frac{w_{ij}^{(1)}}{\sqrt{\sum_{j=1}^{n} \left(w_{ij}^{(1)}\right)^2}}, (j = 1, \cdots, n)$$



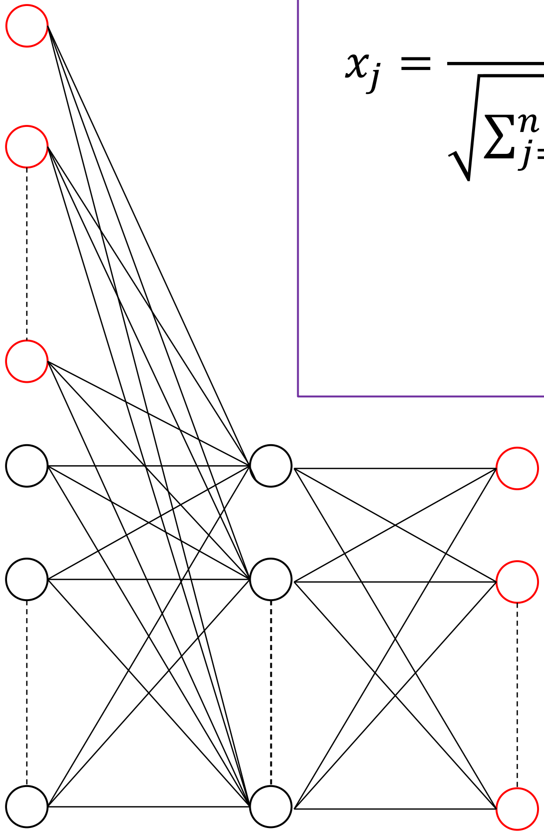
Exercise: How to solve this problem?

Autoencoder Neural Networks

Problem:

What is the feature that hidden neuron i is looking for?

Training

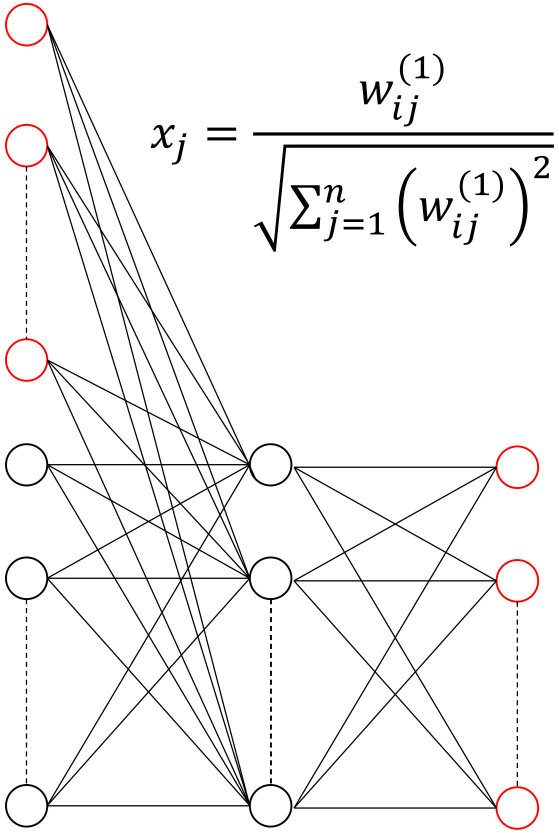


$$x_j = \frac{w_{ij}^{(1)}}{\sqrt{\sum_{j=1}^{n}\left(w_{ij}^{(1)}\right)^2}}, (j = 1, \cdots, n)$$

?

Autoencoder Neural Networks

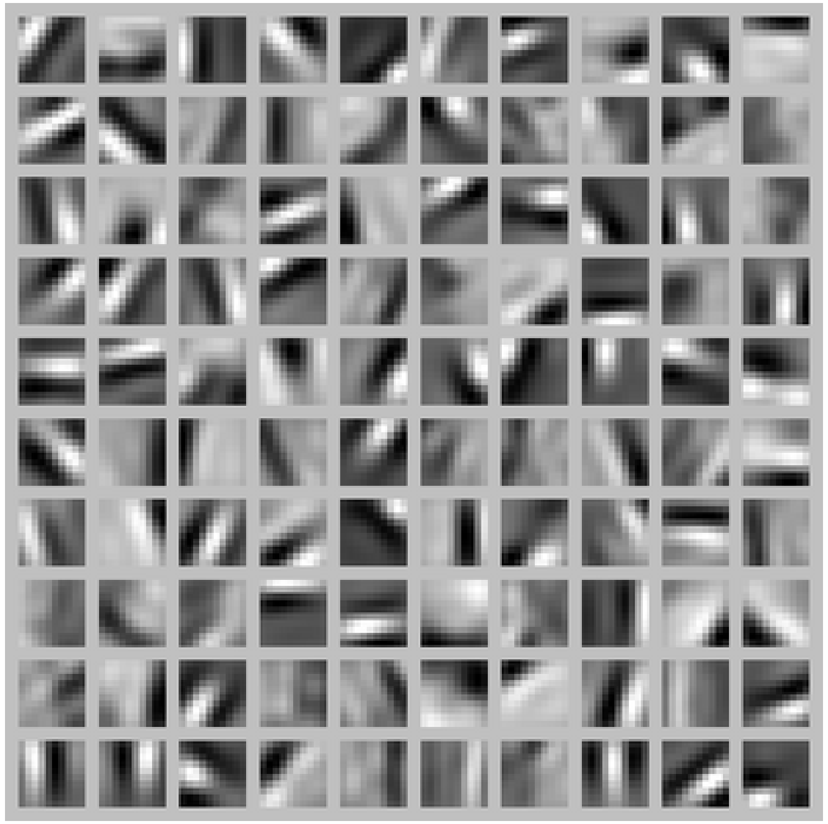
Problem:
What is the feature that hidden neuron i is looking for?



$$x_j = \frac{w_{ij}^{(1)}}{\sqrt{\sum_{j=1}^{n}\left(w_{ij}^{(1)}\right)^2}}$$

100

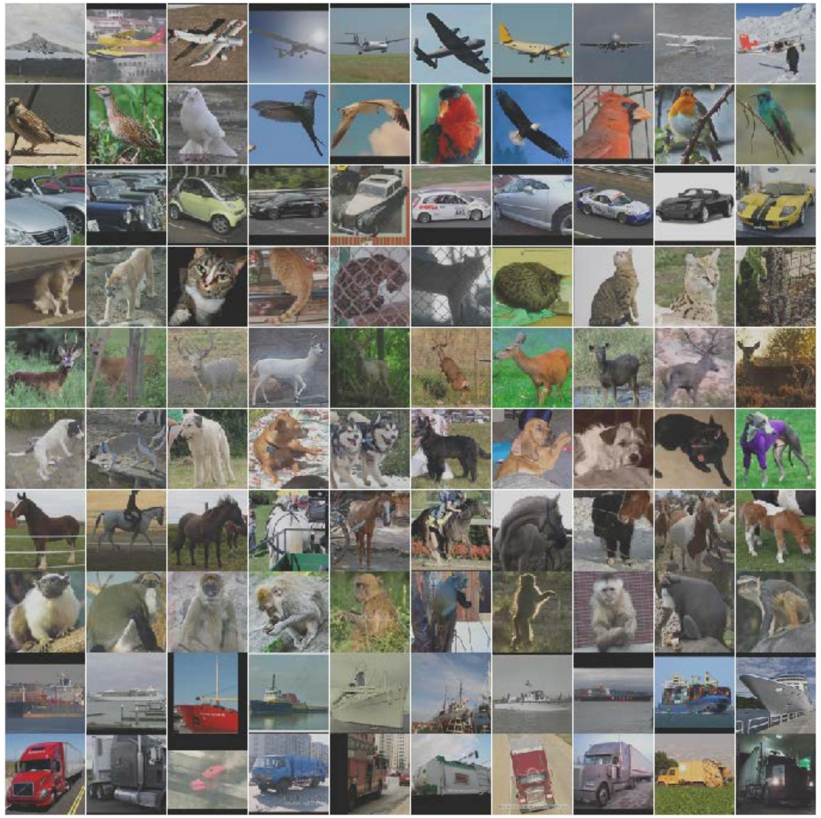
Edges at different positions and orientations



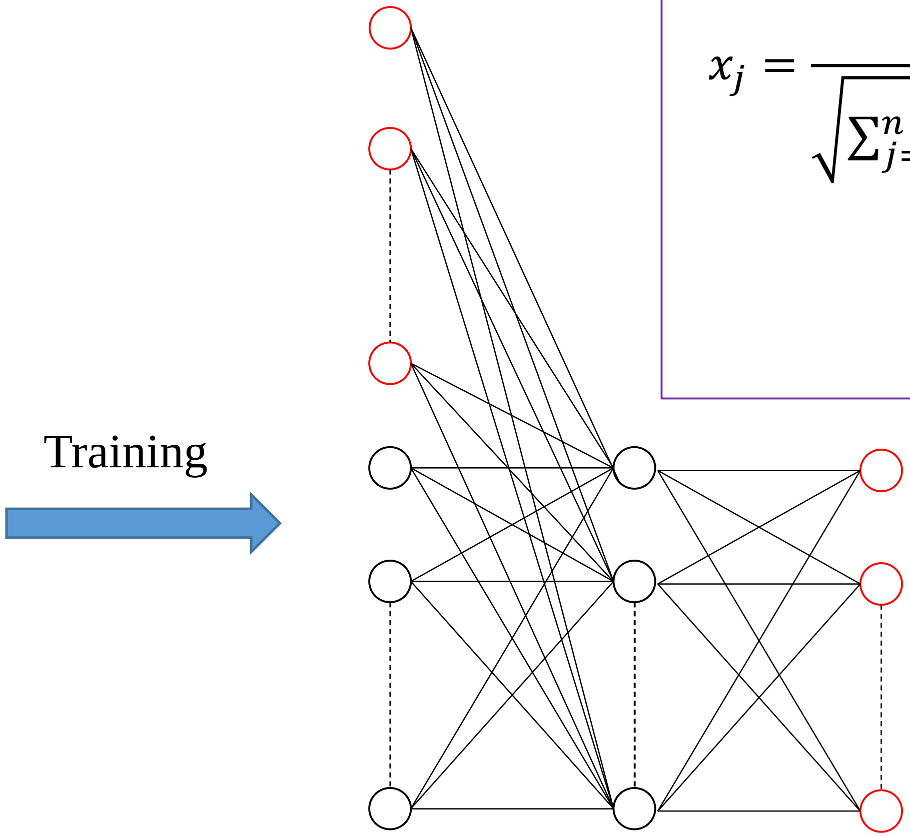
Autoencoder Neural Networks

Problem:

What is the feature that hidden neuron i is looking for?



Training



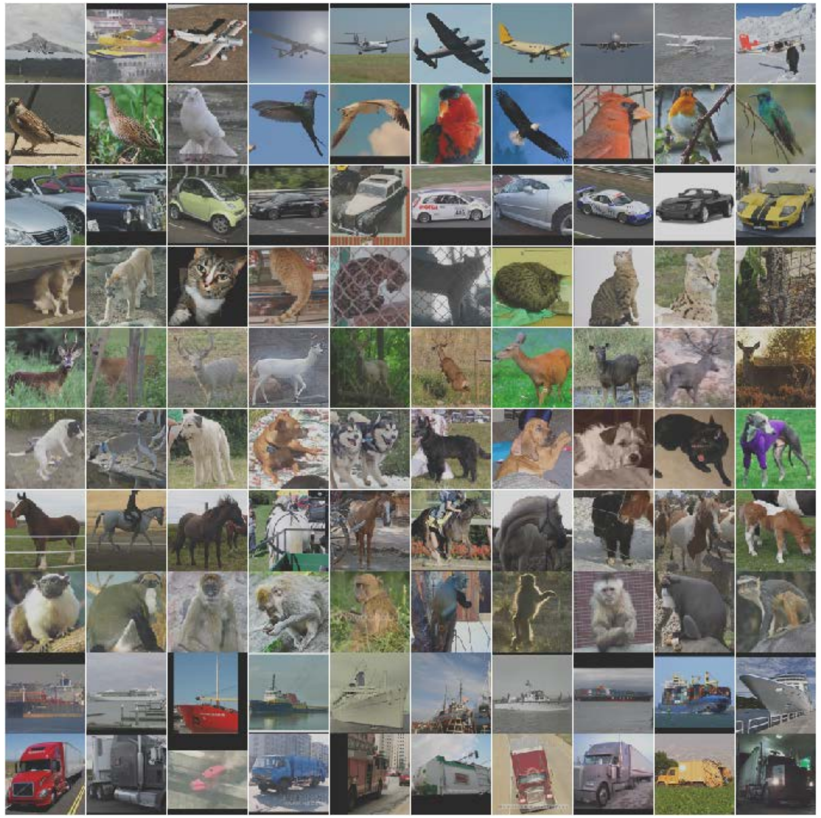
$$x_j = \frac{w_{ij}^{(1)}}{\sqrt{\sum_{j=1}^{n}\left(w_{ij}^{(1)}\right)^2}}, (j = 1, \cdots, n)$$

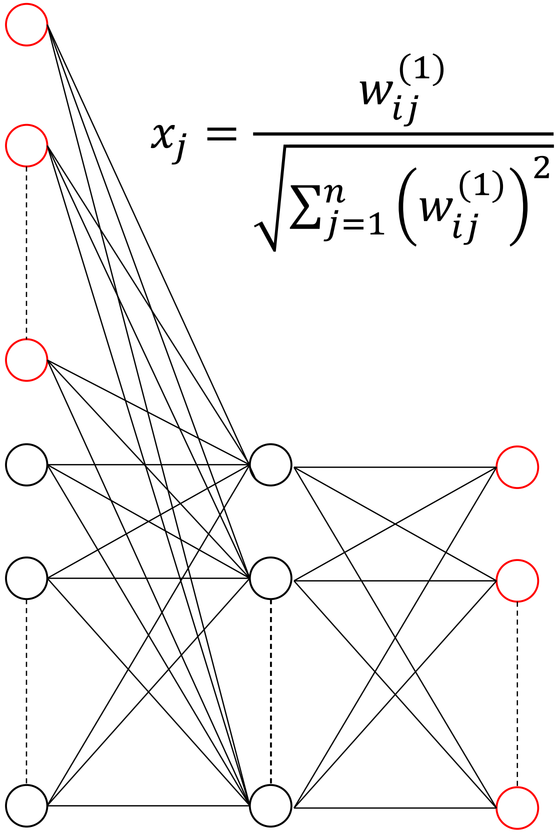
?

Autoencoder Neural Networks

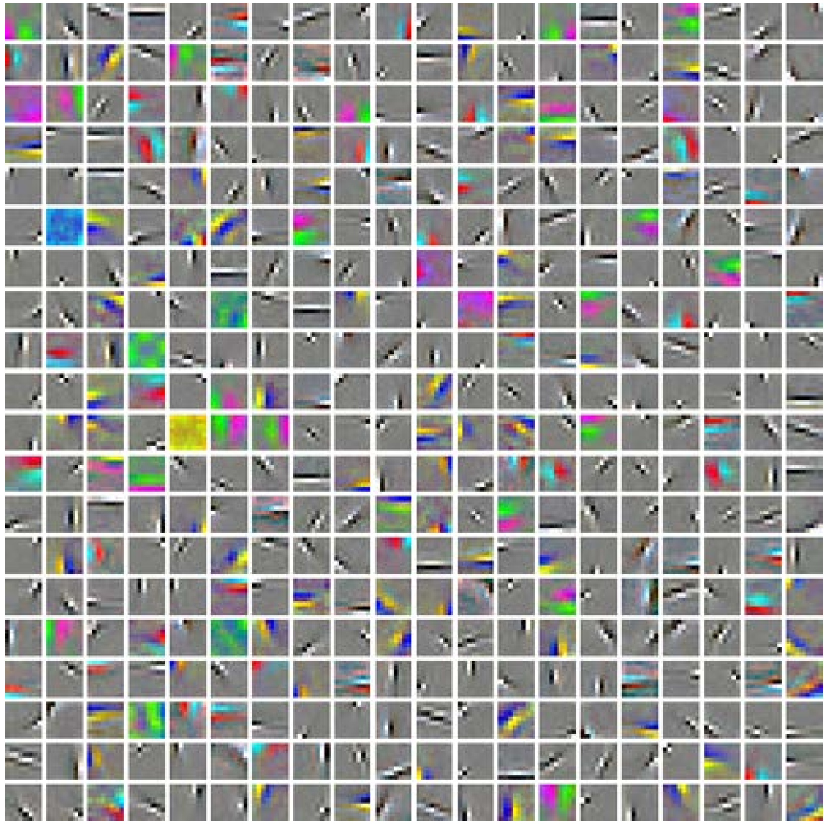
Problem:
What is the feature that hidden neuron i is looking for?

Edges at different positions and orientations





$$x_j = \frac{w_{ij}^{(1)}}{\sqrt{\sum_{j=1}^{n}\left(w_{ij}^{(1)}\right)^2}}$$

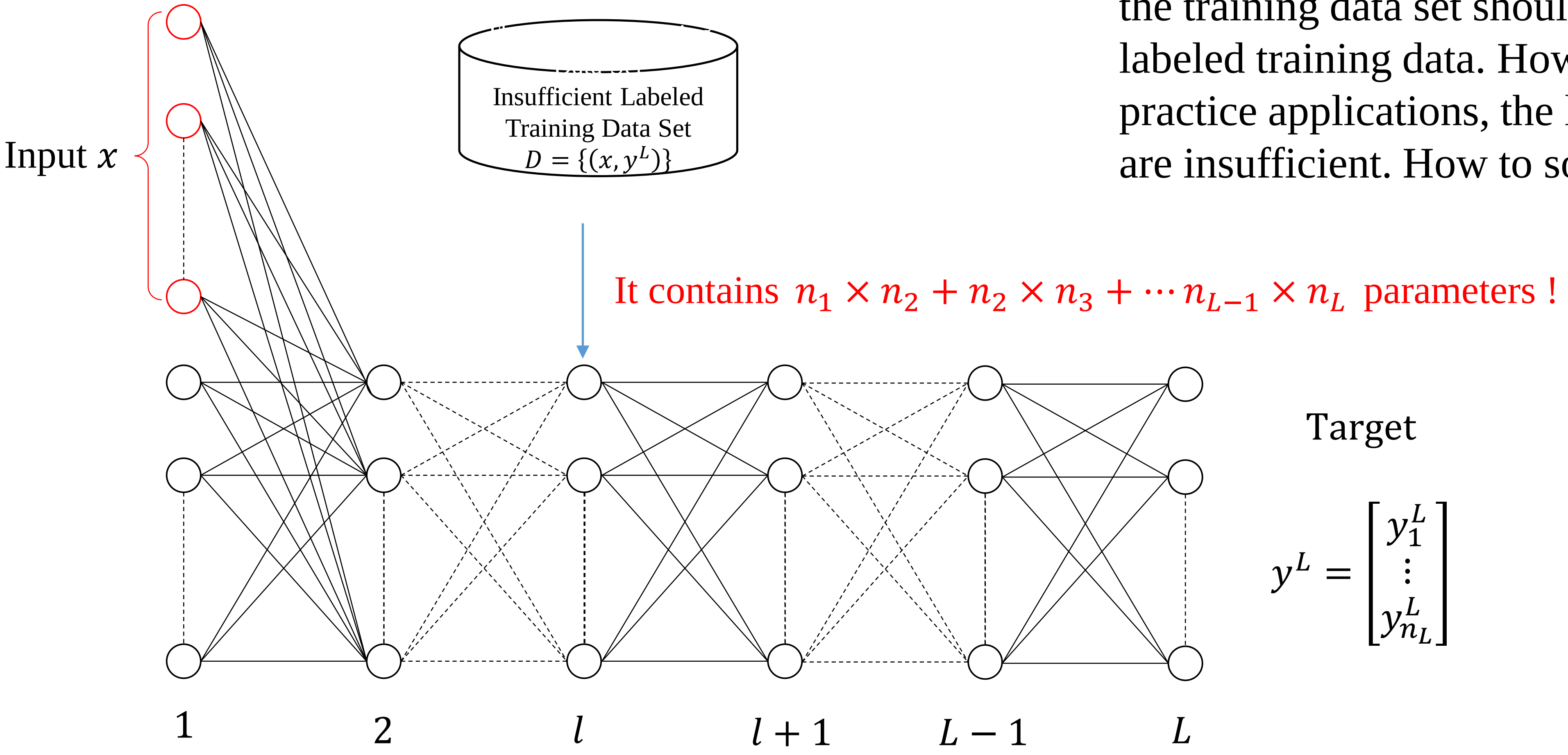


Outline

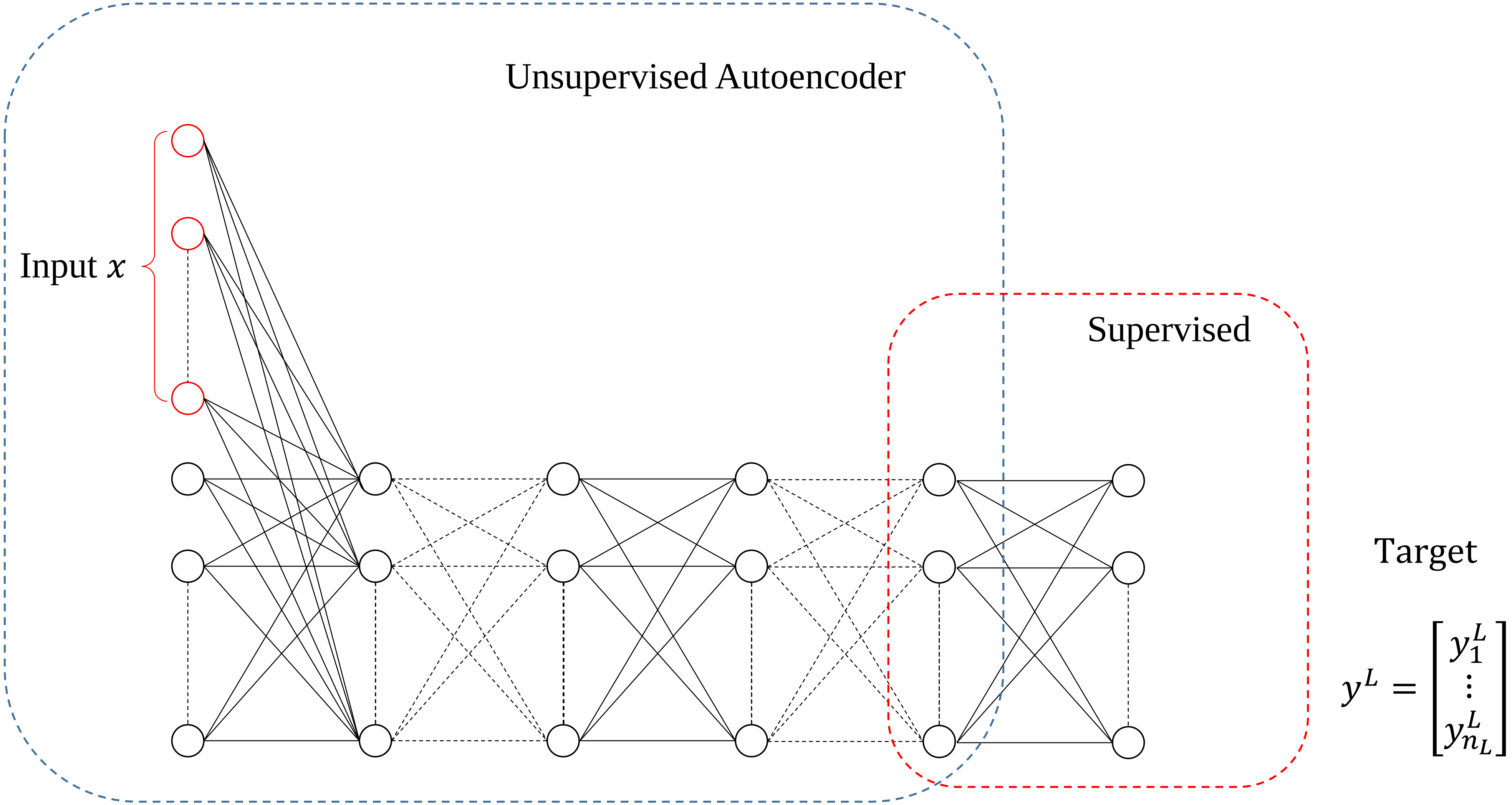
- Learning in Neural Networks
- Supervised Learning
- Unsupervised Learning
- Autoencoder Neural Networks
  - The Network Structure
  - Compressive Representation
  - Sparse Representation
  - Feature Learning
  - Application to Supervised Learning
- Assignment

Application to Supervised Learning

In supervised learning, to train the network, the training data set should contain sufficient labeled training data. However, in many practice applications, the labeled training data are insufficient. How to solve this problem?

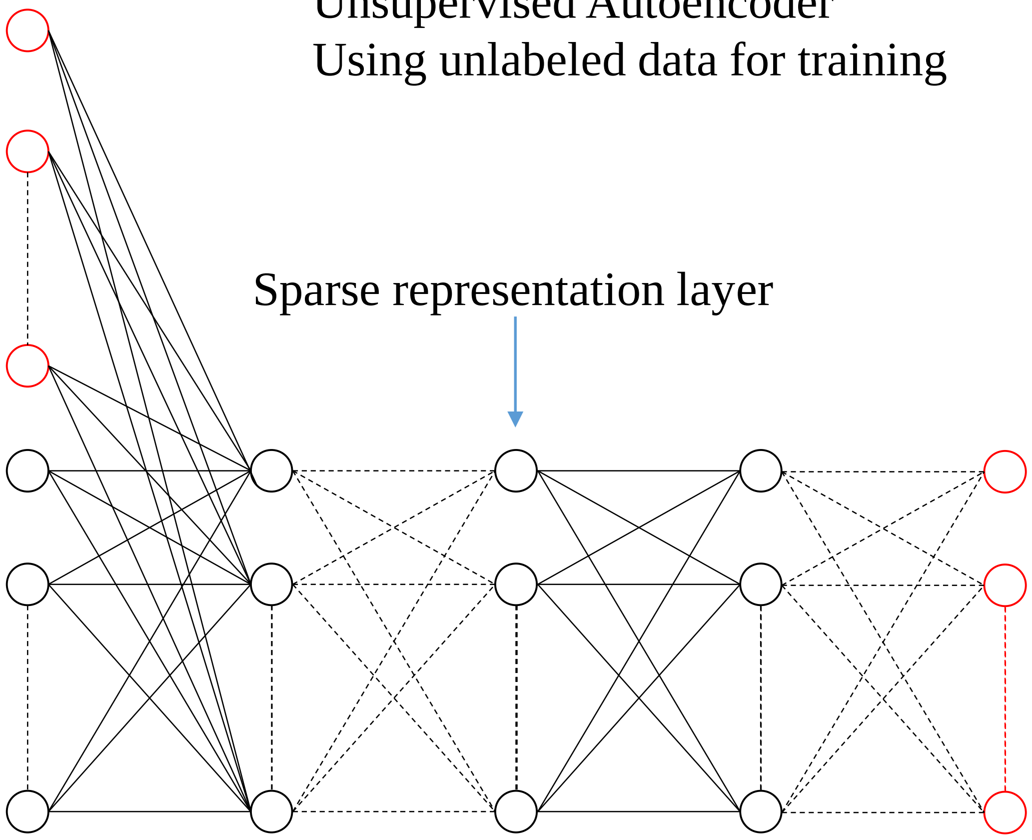


Application to Supervised Learning



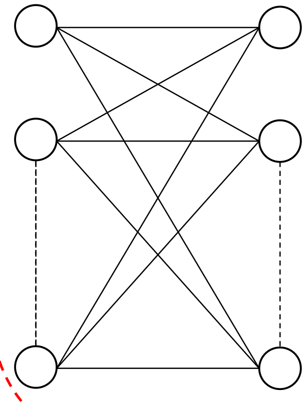
Application to Supervised Learning

Unsupervised Autoencoder
Using unlabeled data for training



l layer

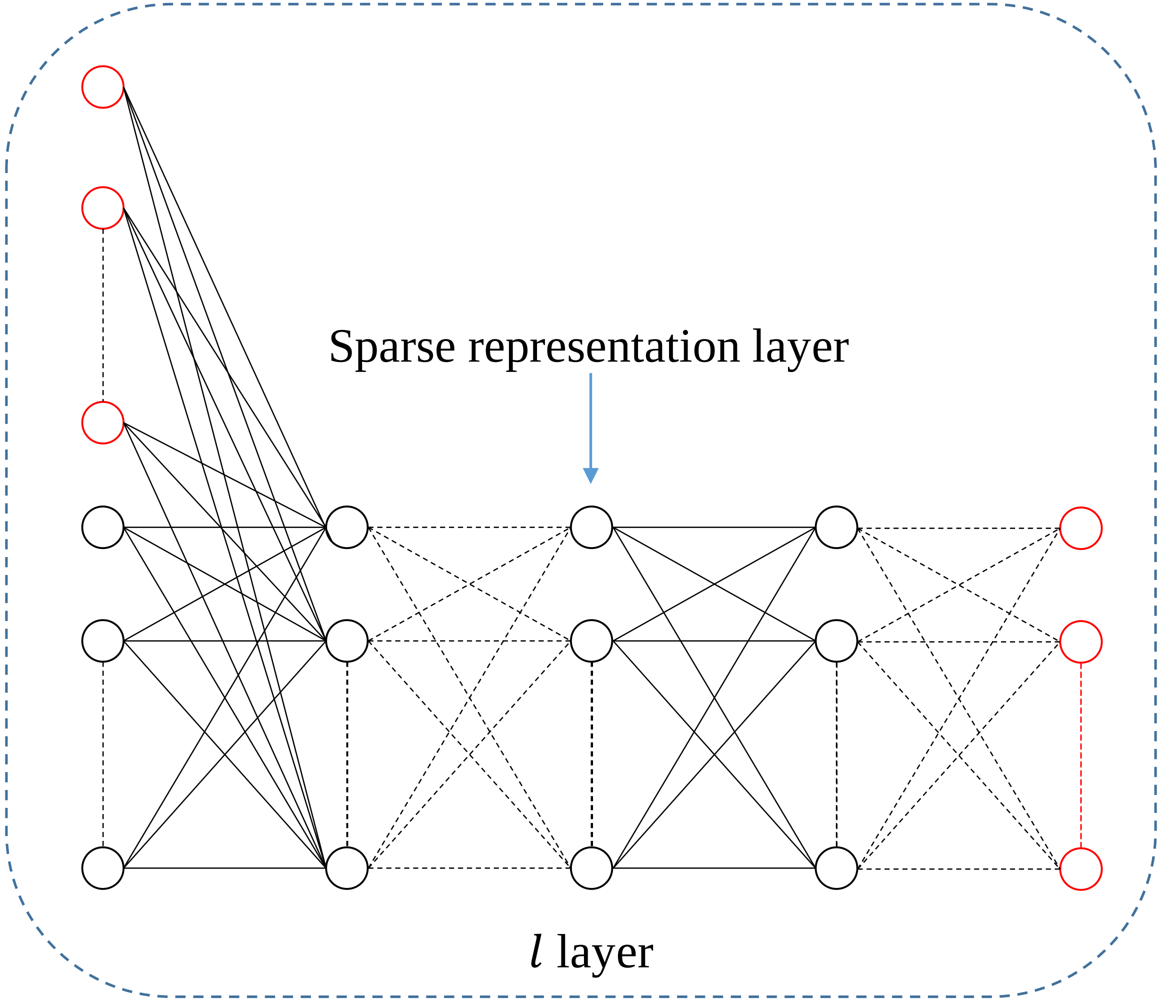
Supervised Two Layer Network
It contains $n^{L-1} \times n^{L}$parameters.
Using labeled data for training



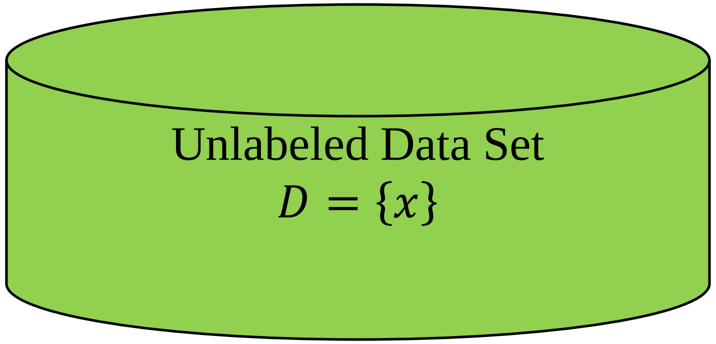
Target

$$y^L = \begin{bmatrix} y_1^L \\ \vdots \\ y_{n_L}^L \end{bmatrix}$$

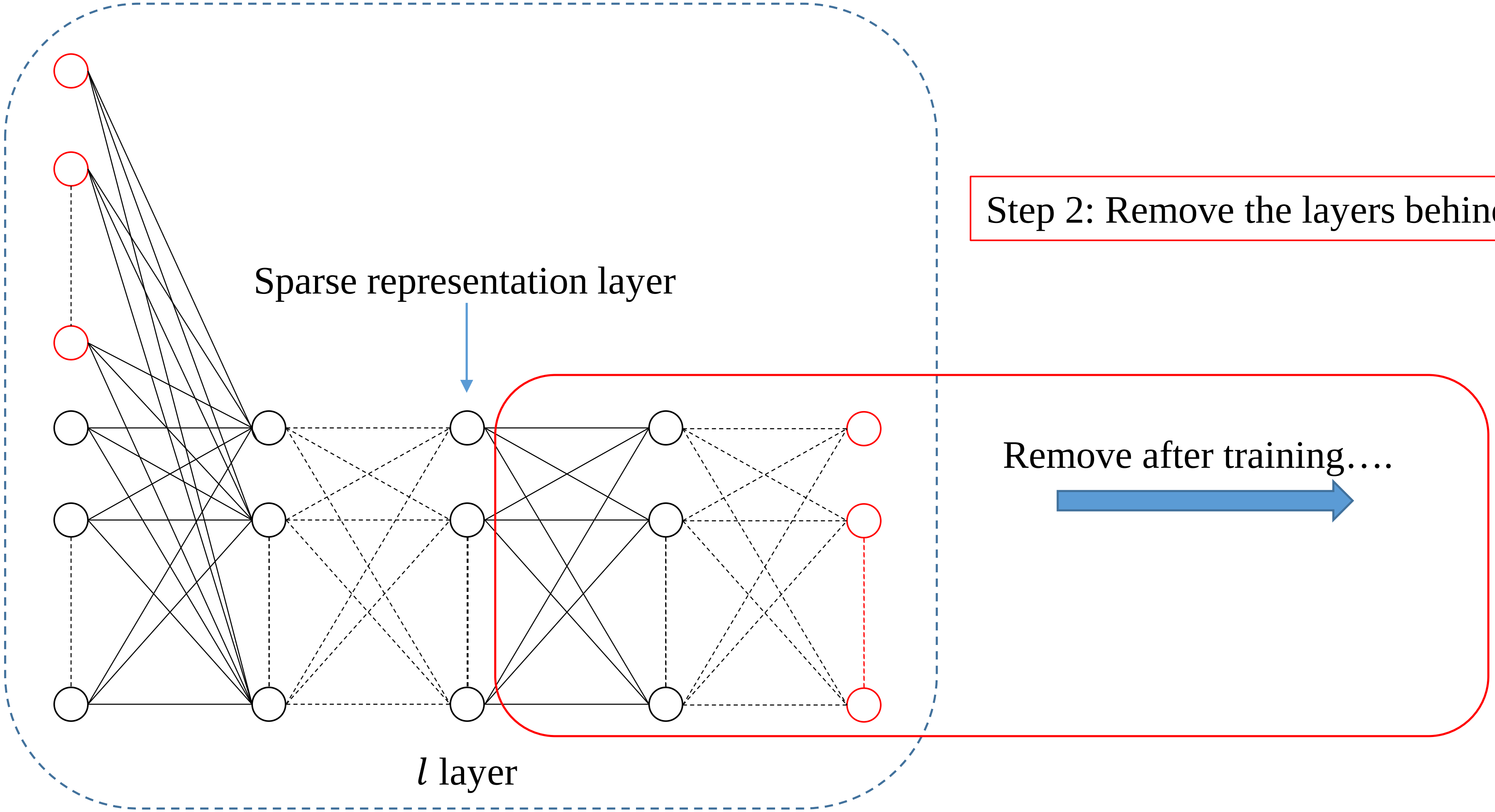
Application to Supervised Learning



Step 1: Train the autoencoder by using unlabeled data.



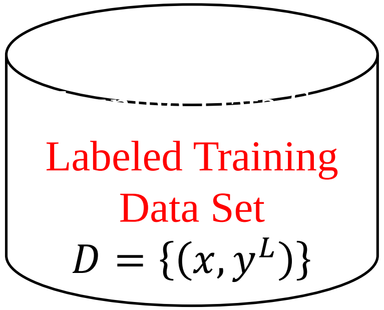
Application to Supervised Learning

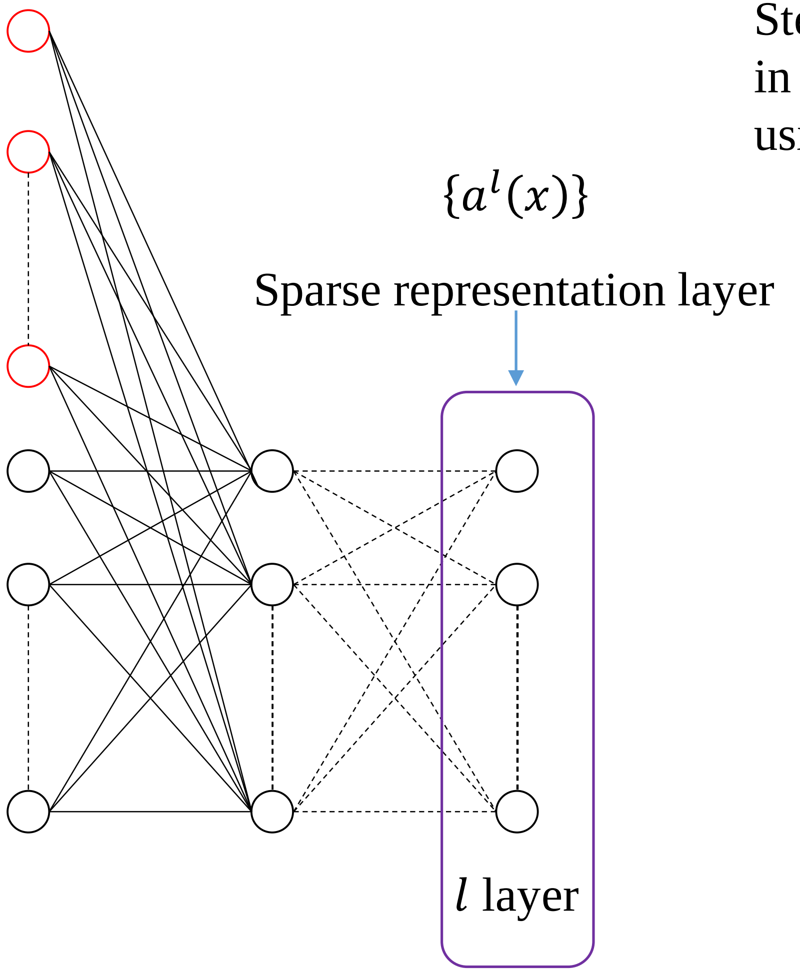


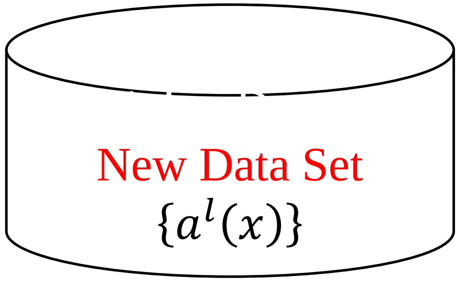
Step 2: Remove the layers behind l^{th} layer after training.

Application to Supervised Learning

Step 3: Form a new data set $\{a^l(x)\}$ in sparse representation layer by using the labeled data set.

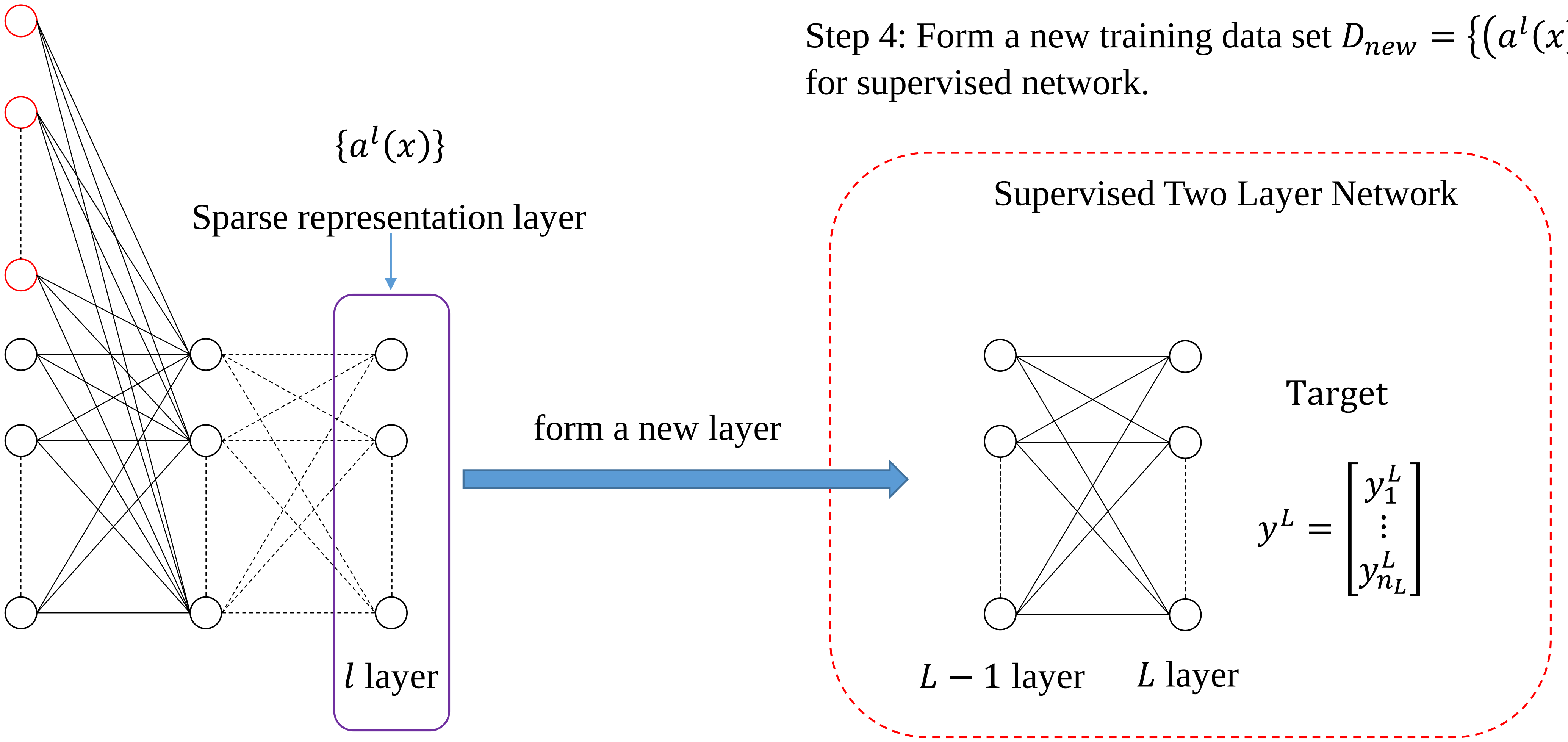




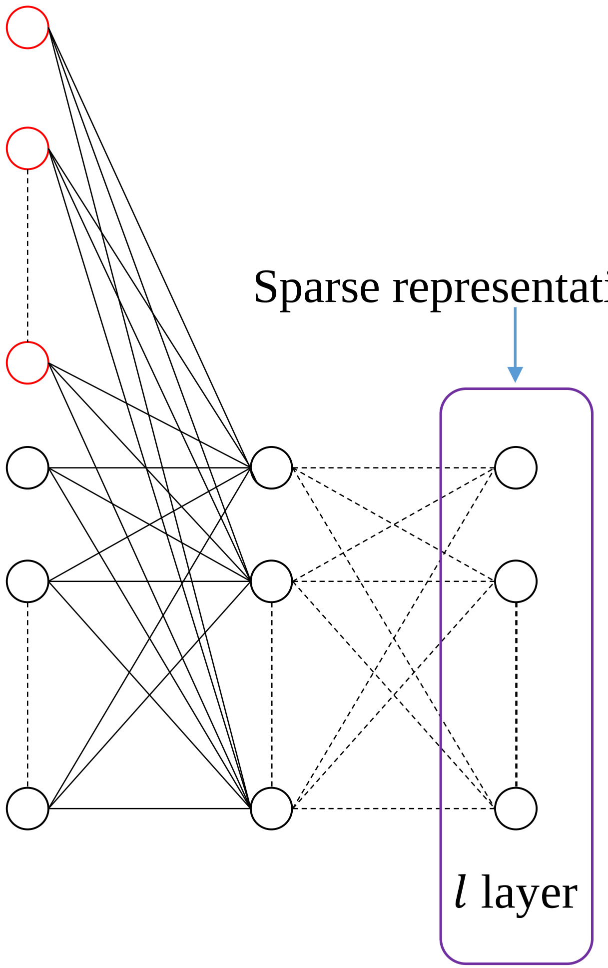


Application to Supervised Learning

Step 4: Form a new training data set $D_{new} = \{(a^l(x), y^L)\}$ for supervised network.



Application to Supervised Learning

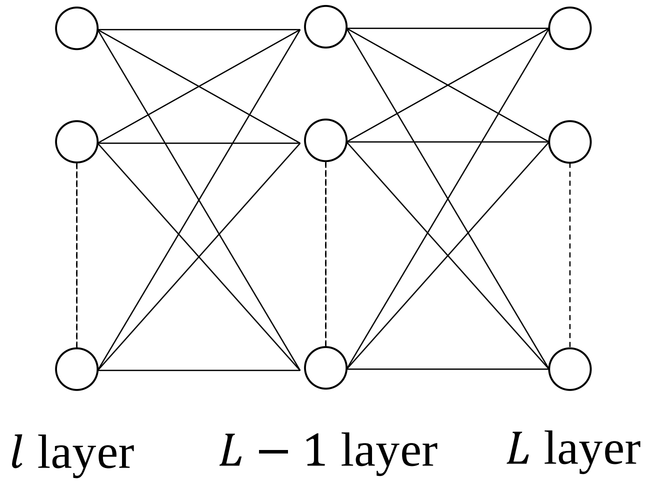


Step 5: Training the network by using the new data set

$$D_{new} = \{(a^l(x), y^L)\}$$

Supervised Three Layers Network

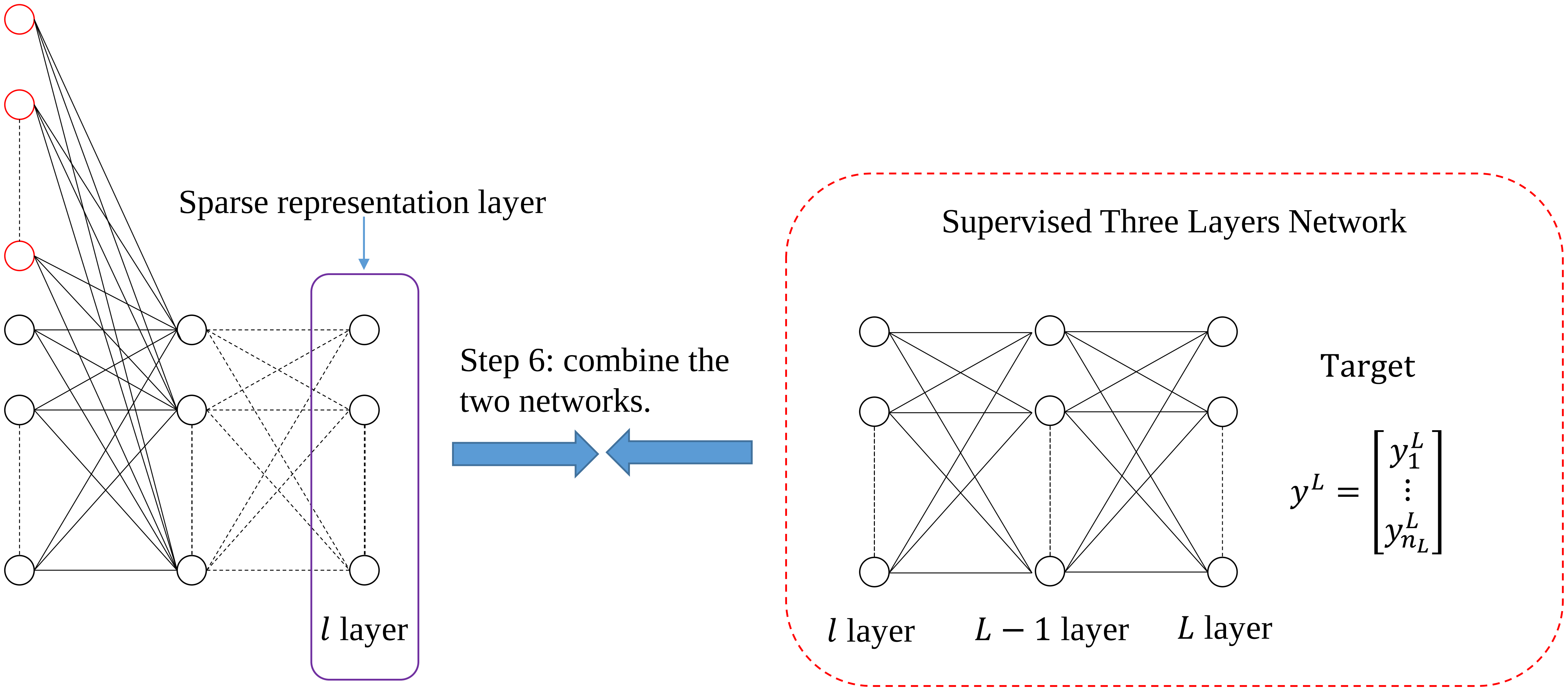
It contains $n_l \times n_{L-1} + n_{L-1} \times n_L$ parameters.



Target

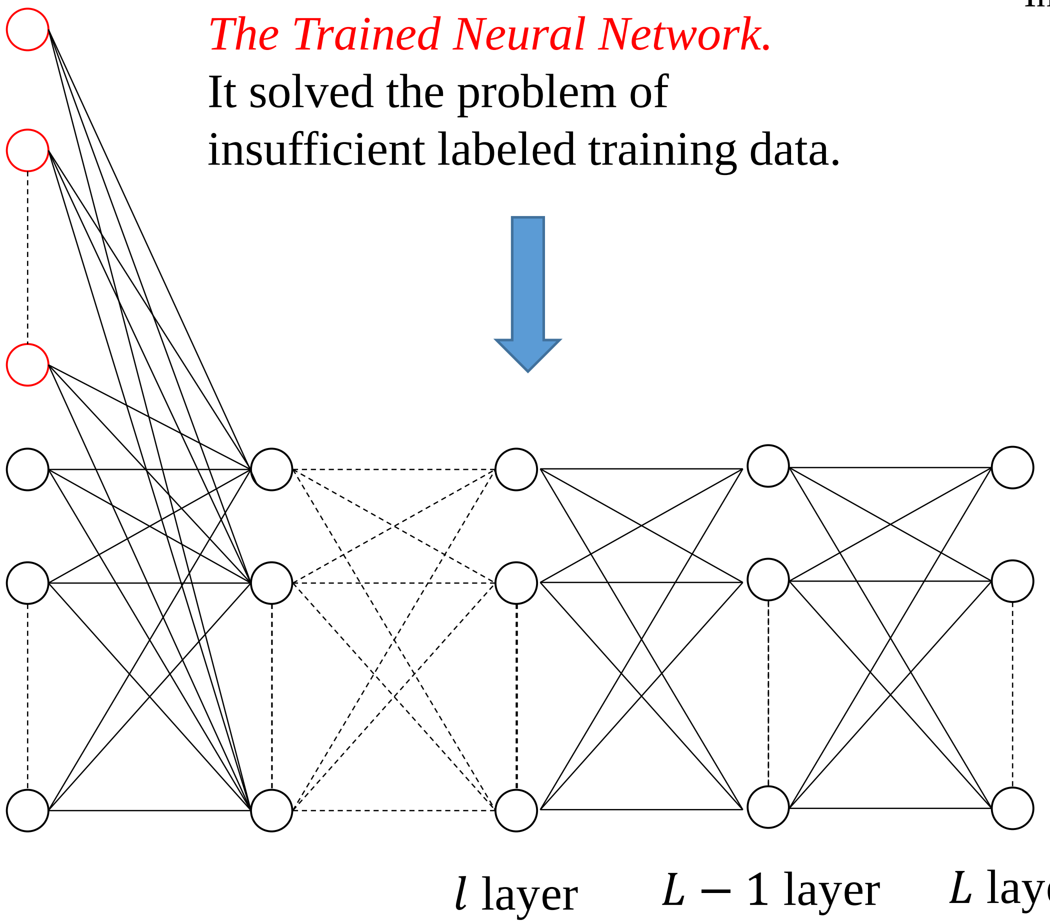
$$y^L = \begin{bmatrix} y_1^L \\ \vdots \\ y_{n_L}^L \end{bmatrix}$$

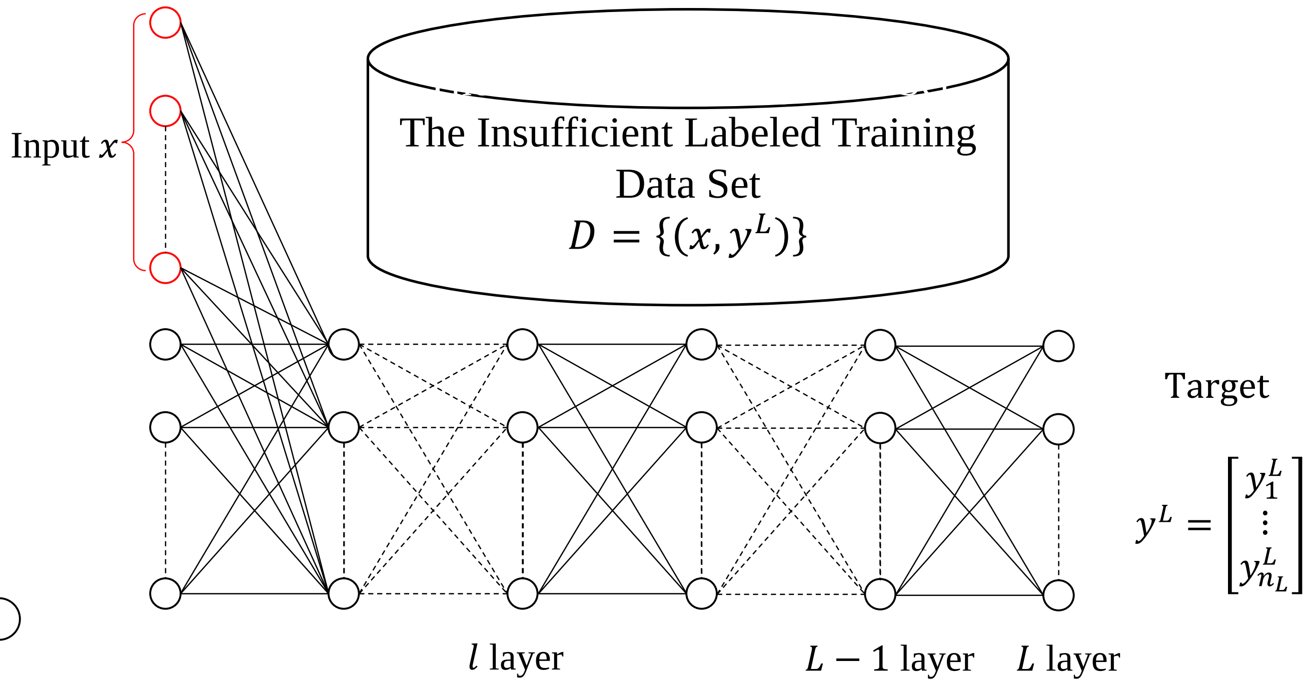
Application to Supervised Learning



Application to Supervised Learning

The Trained Neural Network.
It solved the problem of
insufficient labeled training data.





The Original Neural Network

Outline

- Learning in Neural Networks
- Supervised Learning
- Unsupervised Learning
- *Autoencoder Neural Networks*
  - The Network Structure
  - Compressive Representation
  - Sparse Representation
  - Feature Learning
  - Application to Supervised Learning
- Assignment

Assignment 1

1. Given the cost function

$$J = \frac{1}{2}\sum_{i=1}^{m}\left(a_i^L - x_i\right)^2 + \beta \cdot \sum_{i=1}^{n_l} a_i^l$$

Prove that

$$\delta_i^l = \dot{f}\left(z_i^l\right) \cdot \left(\sum_{j=1}^{n_{l+1}} w_{ji}^l \delta_j^{l+1} + \beta\right)$$

2. Given the optimization problem

$$\begin{cases} \max \sum_{i=1}^{n} w_{ij}^{(1)} x_j \\ s.t. \sum_{i=1}^{n} x_i^2 \leq 1 \end{cases}$$

Prove that

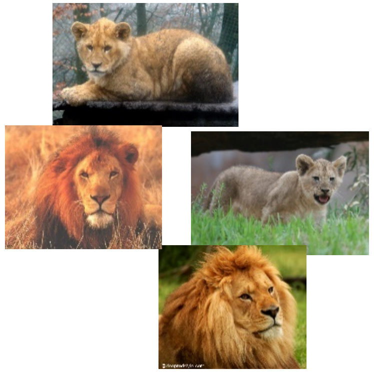
$$x_j = \frac{w_{ij}^{(1)}}{\sqrt{\sum_{j=1}^{n}\left(w_{ij}^{(1)}\right)^2}}, (j = 1, \cdots, n)$$

Assignment 2

Assignment:

In this example, the labeled training data are insufficient to train a classifier by using BP directly. However, a good classifier can be developed by using the autoencoder method. Please do it.

Unlabeled data



Label: Lion : 0

Label: Tiger : 1

Q & A

The End