

# Machine Learning and Financial Applications

## Lecture 8 Deep Neural Networks

---

Video tutorial:

<https://youtu.be/zKN9HOnAByQ>

Liu Peng

liupeng@smu.edu.sg

# A Review of Terminology

Neuron

Weight

Bias

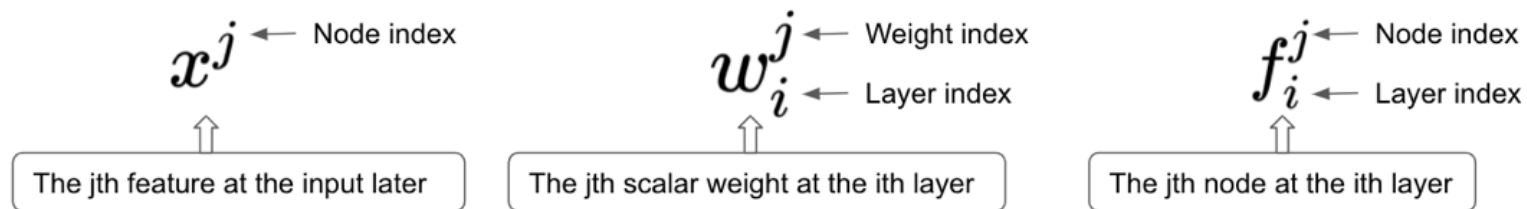
Activation function

Layer

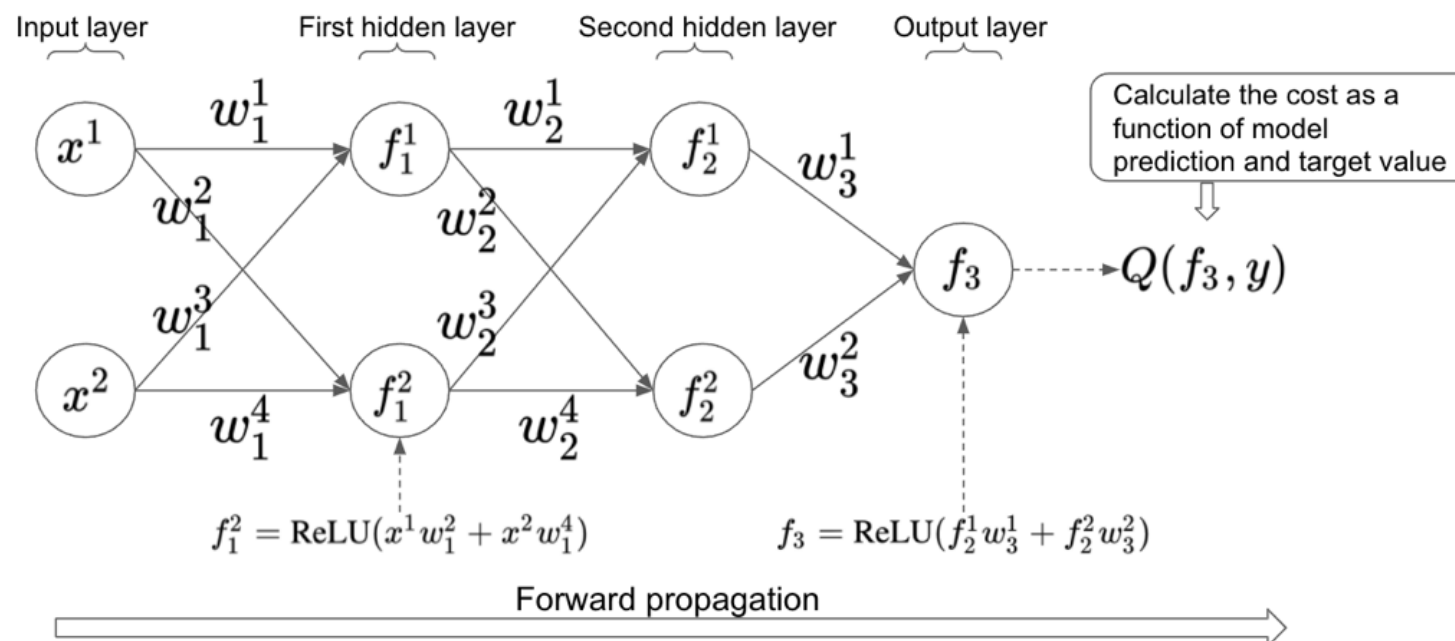
- Feedforward neural network
- Backpropagation
- Loss function
- Optimizer
- Epoch

- Batch size
- Regularization
- Dropout
- Learning rate
- CNN/RNN

# Forward Propagation

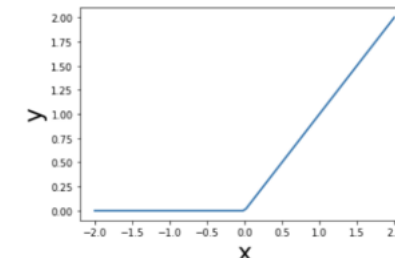


fully connected



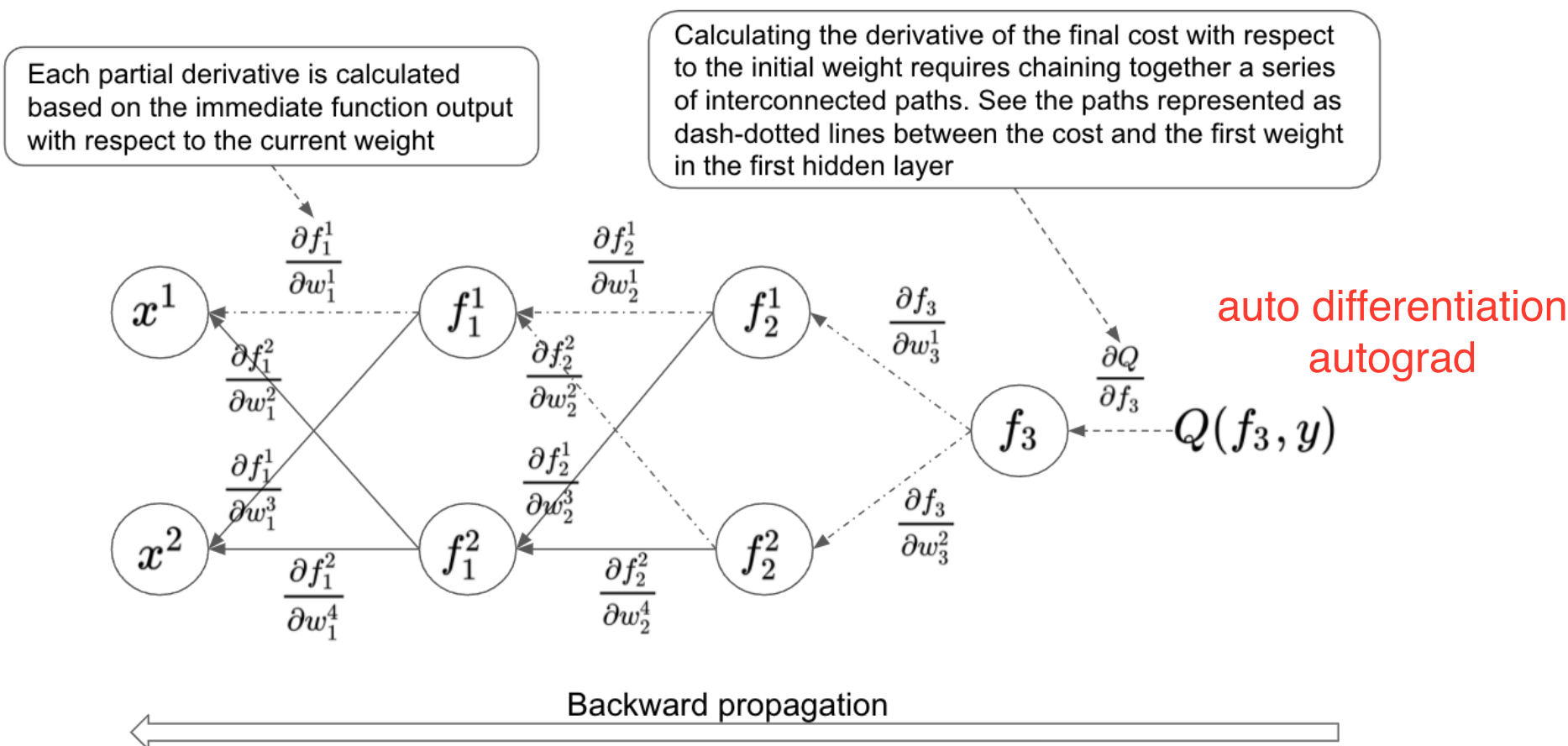
ReLU

$$\phi(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases}$$



# Backward Propagation

he initialization  
xavier initialization



# The Overall Training Process

- Randomly initialize all weights
- Select observation(s)
- Forward propagation to calculate the output signals and compute loss
- Backpropagation to revise all weights
- Repeat until all observations are fed into the ANN, completing an epoch

# In-class quiz

Q1-3

# Gradient Descent

• Forward propagation to calculate the output signals and compute loss  
 • Backpropagation to revise all weights.  
 • Repeat until all observations are fed into the ANN, completing an epoch

In-class quiz

Q1-3

learning rate/step size  
 $\downarrow$   
 $\nabla$   
 expression:  $w = w - \eta \times \text{gradient}$

Gradient Descent

minimization

if just 1  $w$ , shape of graph is

$\|y - f(w)\|^2$

random initial weight.

gradient line, which is partial derivative

if  $\frac{\partial Q}{\partial w} > 0$

subtract from this partial derivative

So  $w$  becomes smaller, and point moves left on the curve.

if  $\frac{\partial Q}{\partial w} < 0$  on the left, the gradient will be negative. So subtract a negative will increase  $w$  and move right.

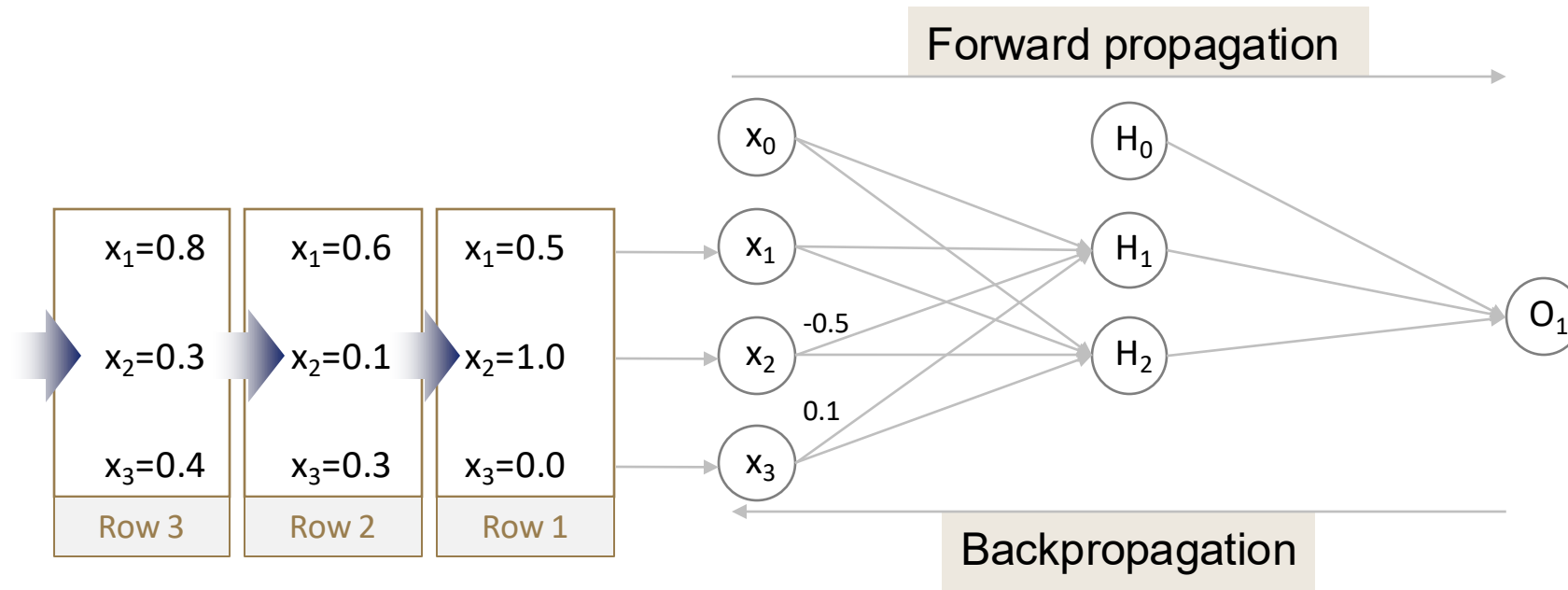
What is next after one i

- One forward propagation and one backpropagation
- New weights are typically not too different from old weights
- Next observation can be fed into the ANN and

$x_1=0.8$	$x_1=0.6$	$x_1=0.5$
$x_2=0.3$	$x_2=0.1$	$x_2=1.0$
$x_3=0.4$	$x_3=0.3$	$x_3=0.0$
Row 3	Row 2	Row 1

# What is next after one iteration?

- One forward propagation and one backpropagation conducted with one observation
- New weights are typically not too different from original weights
- Next observation can be fed into the ANN and update all weights again slowly





# Convolutional Neural Network (CNN)

## Convolution operation

A specific local region (patch) of the grayscale input image data shown in solid line, with the rest of the image shown in dashed line. Each cell holds a pixel value between 0 and 255.

1	1	1			
1	1	1			
1	1	1			

A 3x3 kernel defined by a convolution layer. Each cell in the kernel holds a weight parameter. The kernel will convolve with the input patch via a weighted sum operation, i.e., element-wise multiplication and summation.

+

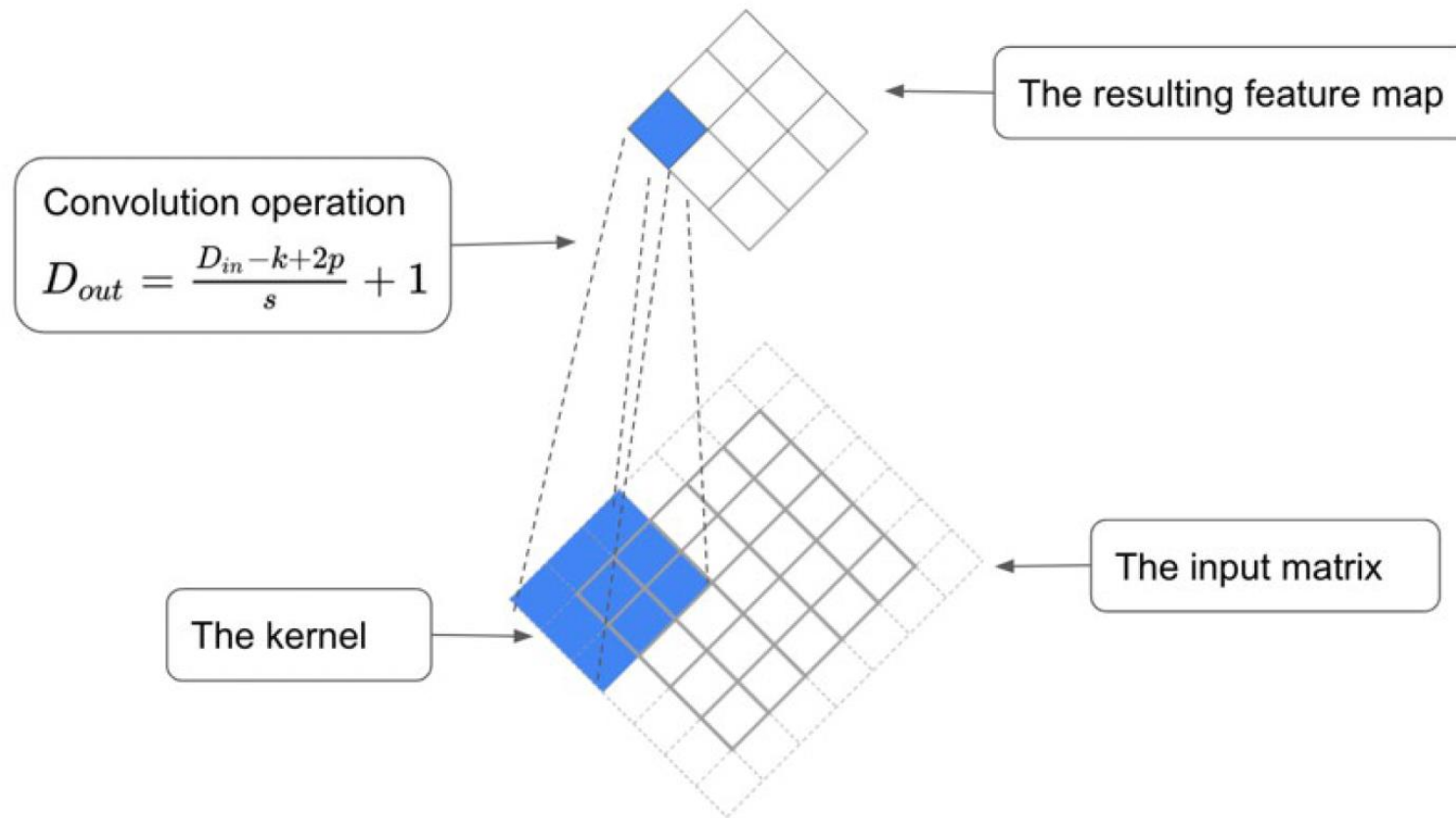
1	0	0
0	1	0
0	0	1



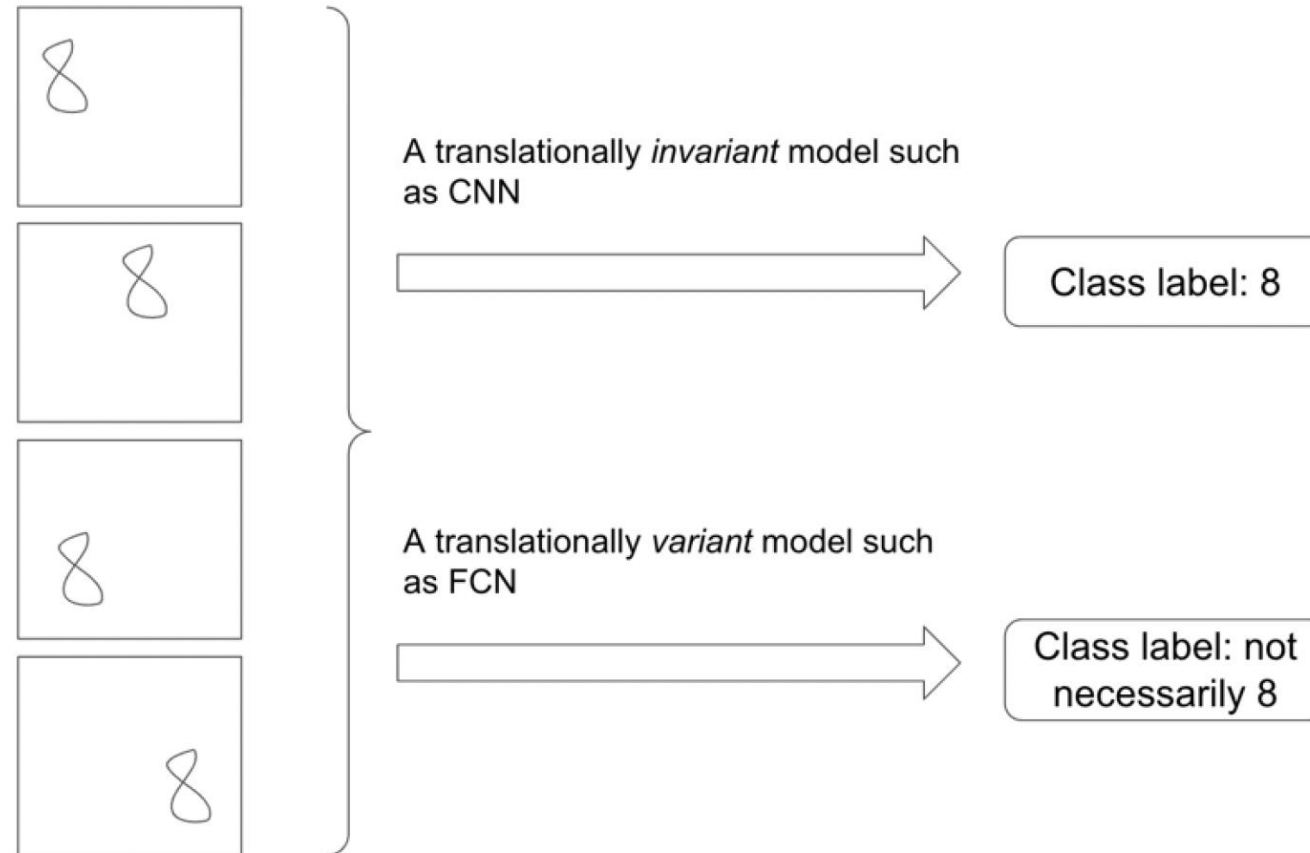
The resulting feature map, with the example convolved feature in solid bold and others in dashed line.

<b>3</b>			

# Sliding the kernel to produce feature maps



# Translational invariance



# In-class quiz

Q4-7



Coding session