#### 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

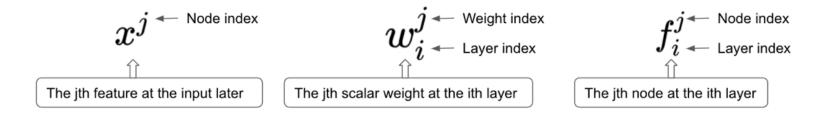
**Activation function** 

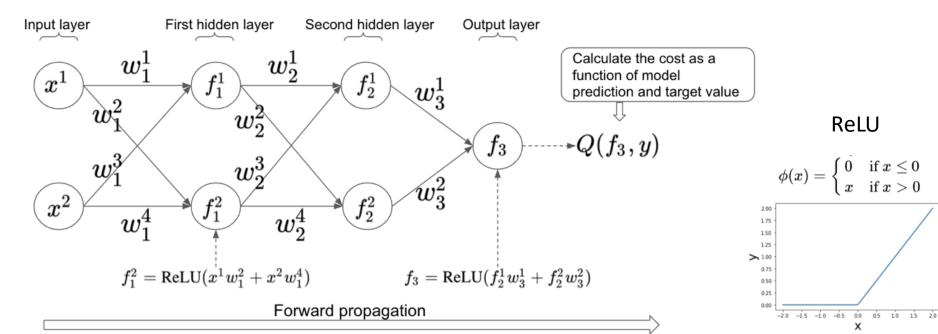
Layer

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

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

## Forward Propagation





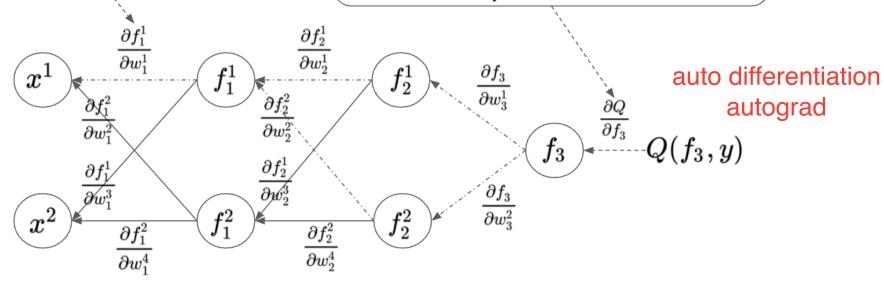
fully connected

### **Backward Propagation**

### he initialization xavier initialization

Each partial derivative is calculated based on the immediate function output with respect to the current weight

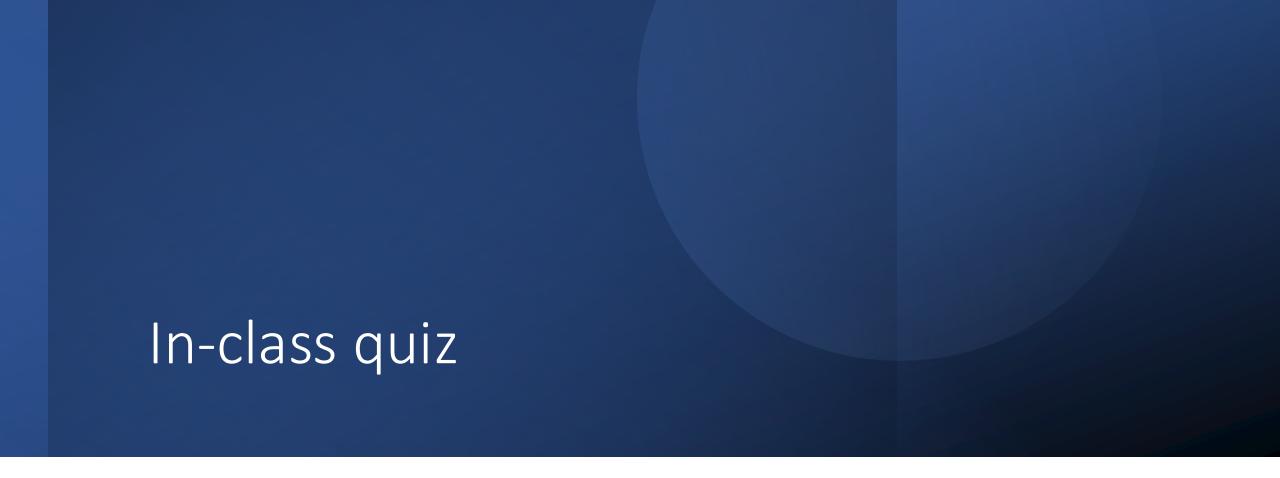
Calculating the derivative of the final cost with respect to the initial weight requires chaining together a series of interconnected paths. See the paths represented as dash-dotted lines between the cost and the first weight in the first hidden layer



Backward propagation

### The Overall Training Process

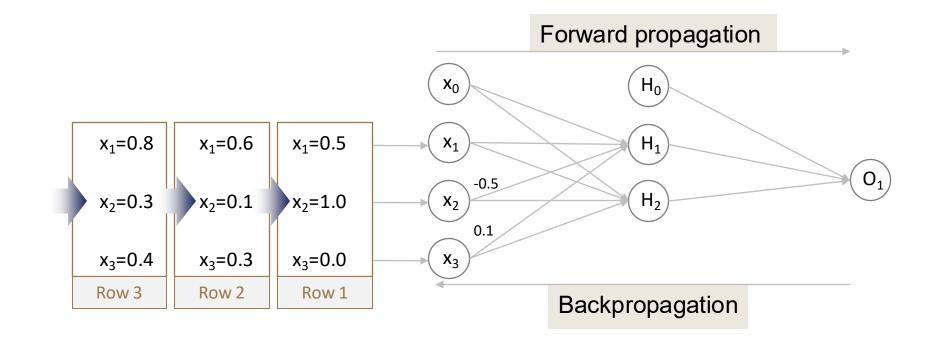
- Randomly initialize all weights
- Select observation(s)
- Forward propagation to calculate the output signals and compute loss
- Backpropagation to revise all weights
- Repeat until all observations are fed into the ANN, completing an epoch



#### Gradient Descent

#### What is next after one iteration?

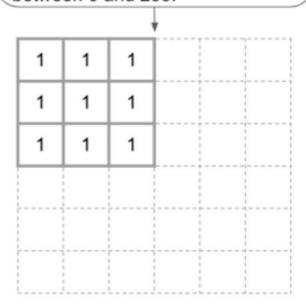
- One forward propagation and one backpropagation conducted with one observation
- New weights are typically not too different from original weights
- Next observation can be fed into the ANN and update all weights again slowly



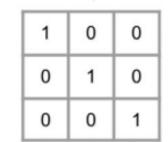
#### Convolutional Neural Network (CNN)

#### Convolution operation

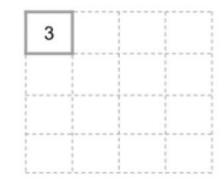
A specific local region (patch) of the grayscale input image data shown in solid line, with the rest of the image shown in dashed line. Each cell holds a pixel value between 0 and 255.



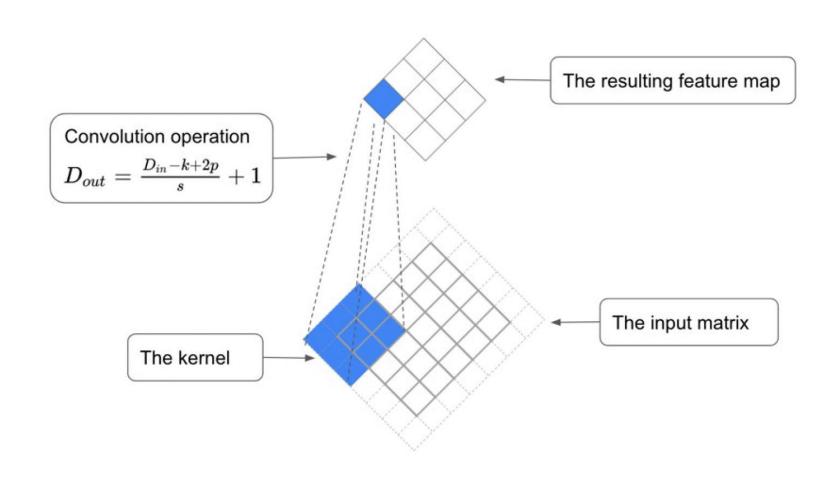
A 3x3 kernel defined by a convolution layer. Each cell in the kernel holds a weight parameter. The kernel will convolve with the input patch via a weighted sum operation, i.e., element-wise multiplication and summation.



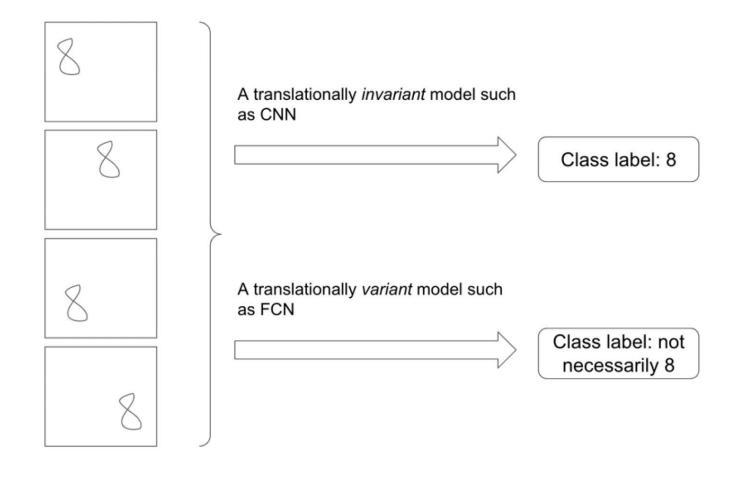
The resulting feature map, with the example convoluted feature in solid bold and others in dashed line.

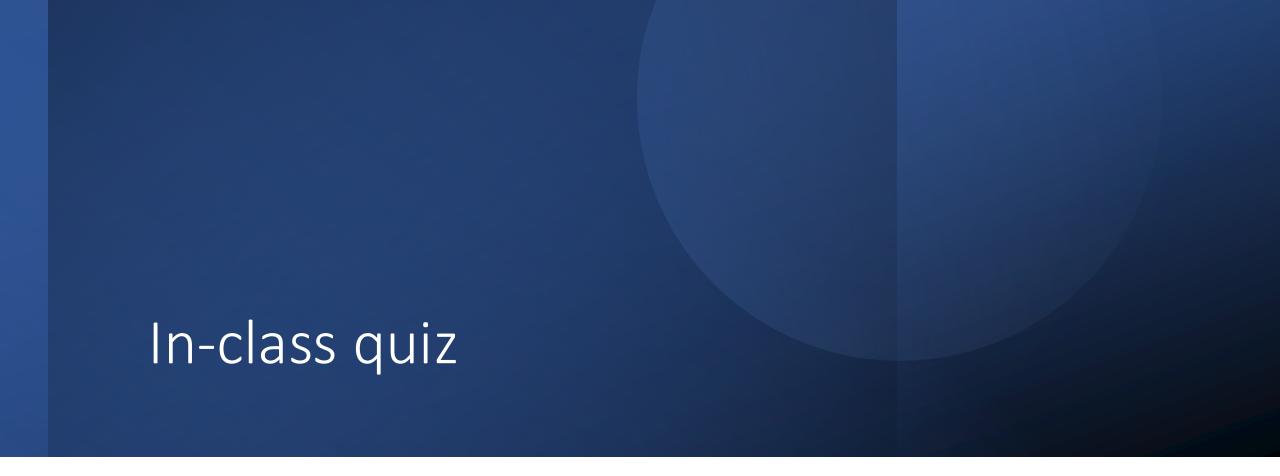


#### Sliding the kernel to produce feature maps



#### Translational invariance





# Coding session