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

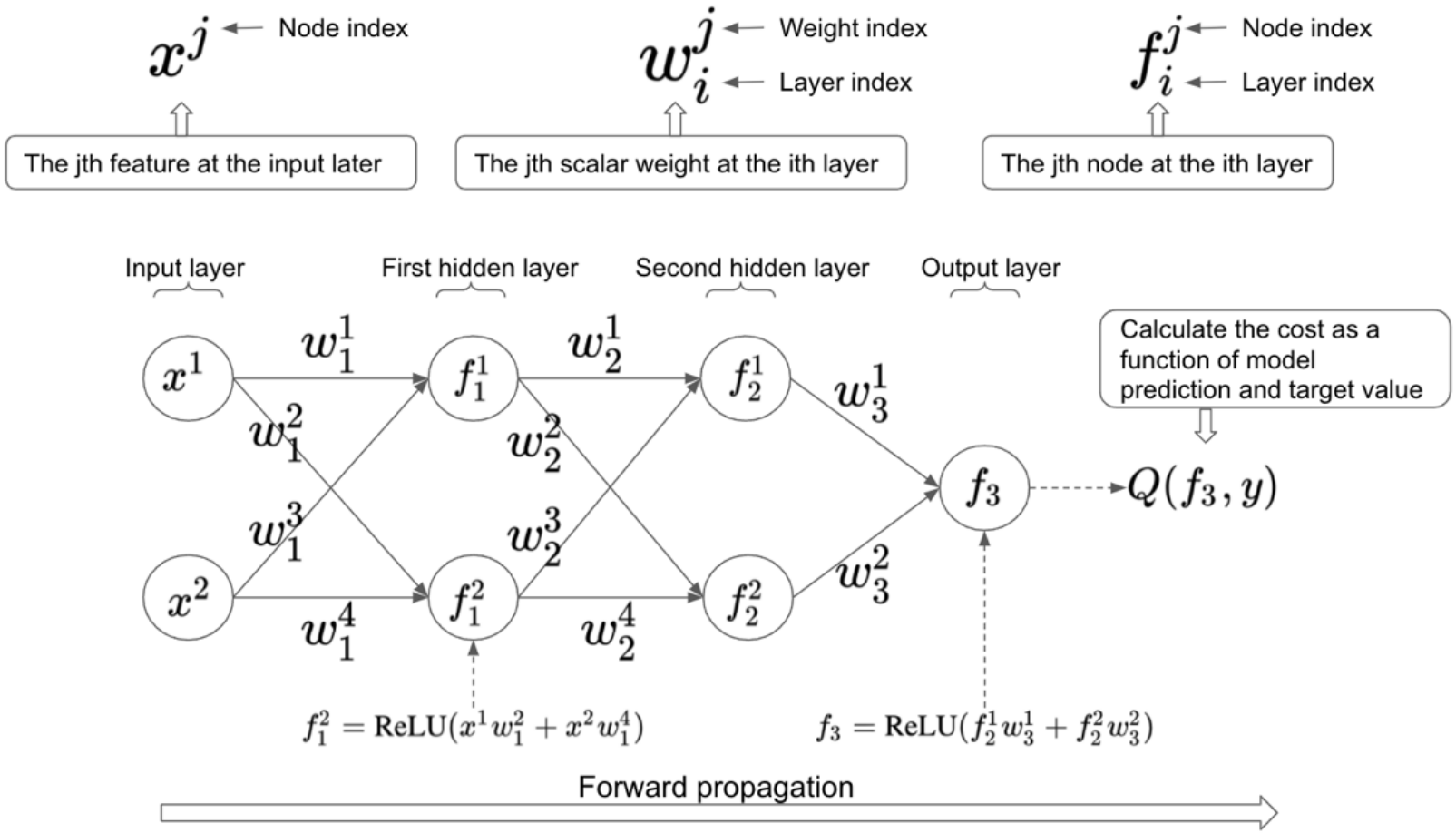
Activation function

Layer

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

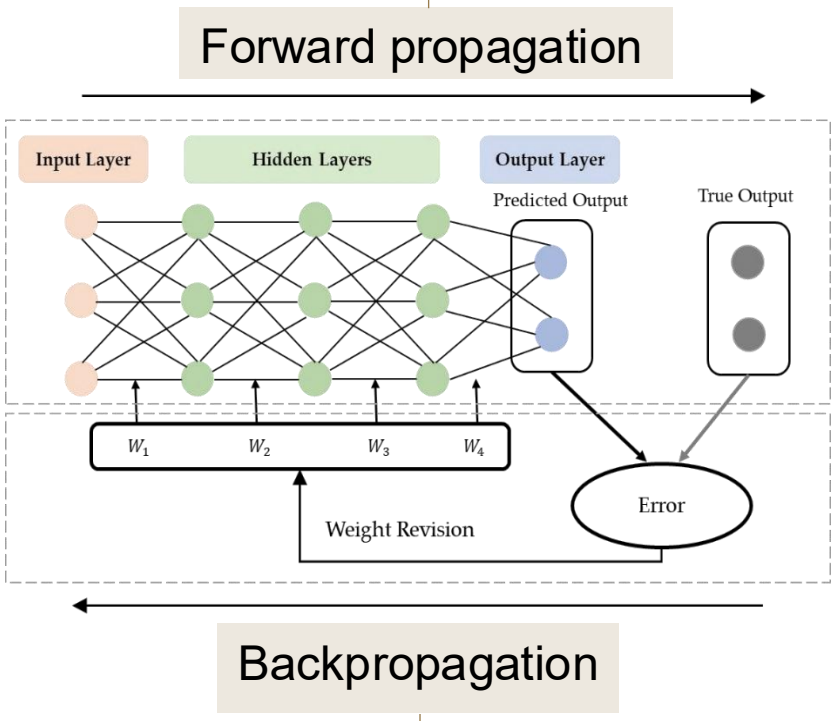
- Batch size
- Regularization
- Dropout
- Learning rate
- CNN/RNN

Forward Propagation



Workflow of ANN Algorithm

- Each input node receives the input variables' values from the data set and sends weighted signals to the hidden nodes
- Each hidden node and output node combines the received signals and pass it through an activation function to generate the output
- The output value of the output nodes are the predicted values



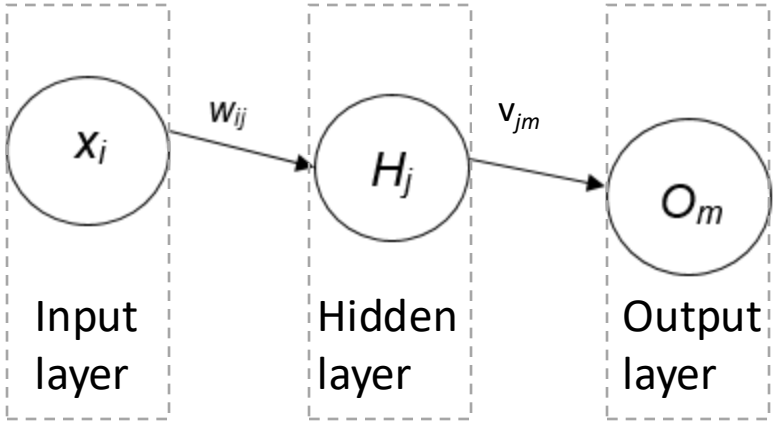
- Compute the prediction error at the output nodes
- Update the weights between the last hidden layer and output layer
- Compute the signal error at the hidden layer
- Update the weights between the last hidden layer and the second last hidden layer
- Similarly repeat for all the hidden layers until we update weights from input layer to first hidden layer

More on Data Set for the ANN Model

How is the training data set fed into the process?

- Randomly initialize all weights
- Select an observation
- Forward propagation to calculate the output signals
- Backpropagation to revise all weights
- Repeat until all observations are fed into the ANN

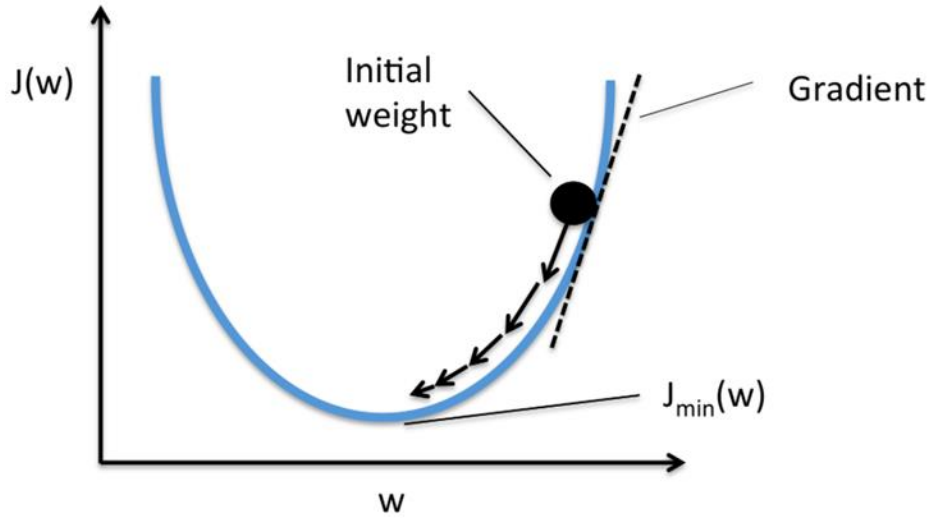
Notations for a Two-layer ANN



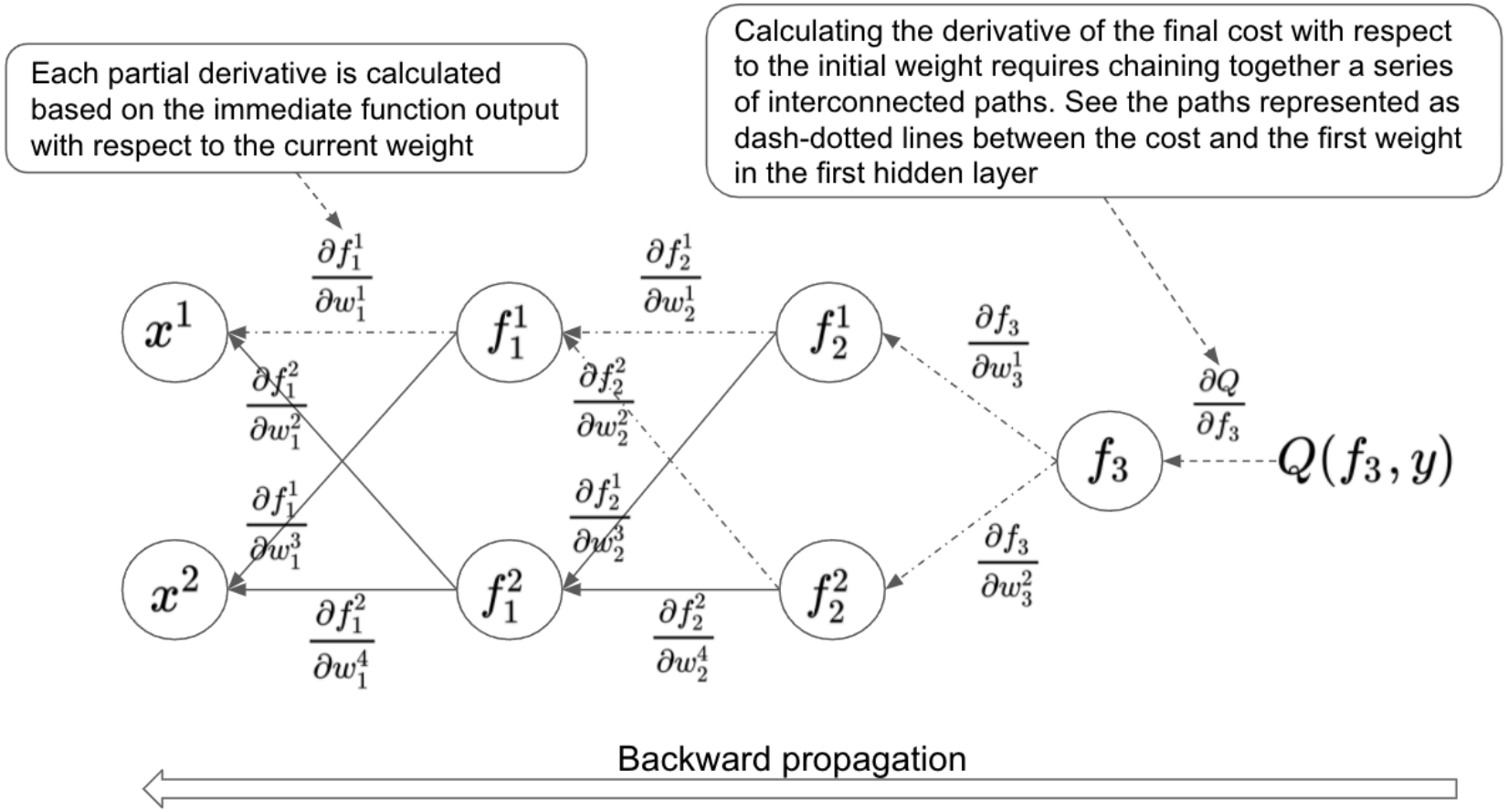
- x_i : *i-th* input variable node, $i = 1, 2, \ldots, d$
- H_j : *j-th* hidden node, $j = 1, 2, \ldots, p$
- O_m : *m-th* output variable node, $m = 1, 2, \ldots, c$
- w_{ij} : weight between x_i and H_j
- v_{jm} : weight between H_j and O_m

- Each node is an autonomous computational unit
- Each layer has a unique **activation function**
- Each connection has a **weight** indicating the **importance** of each input signal

Gradient Descent

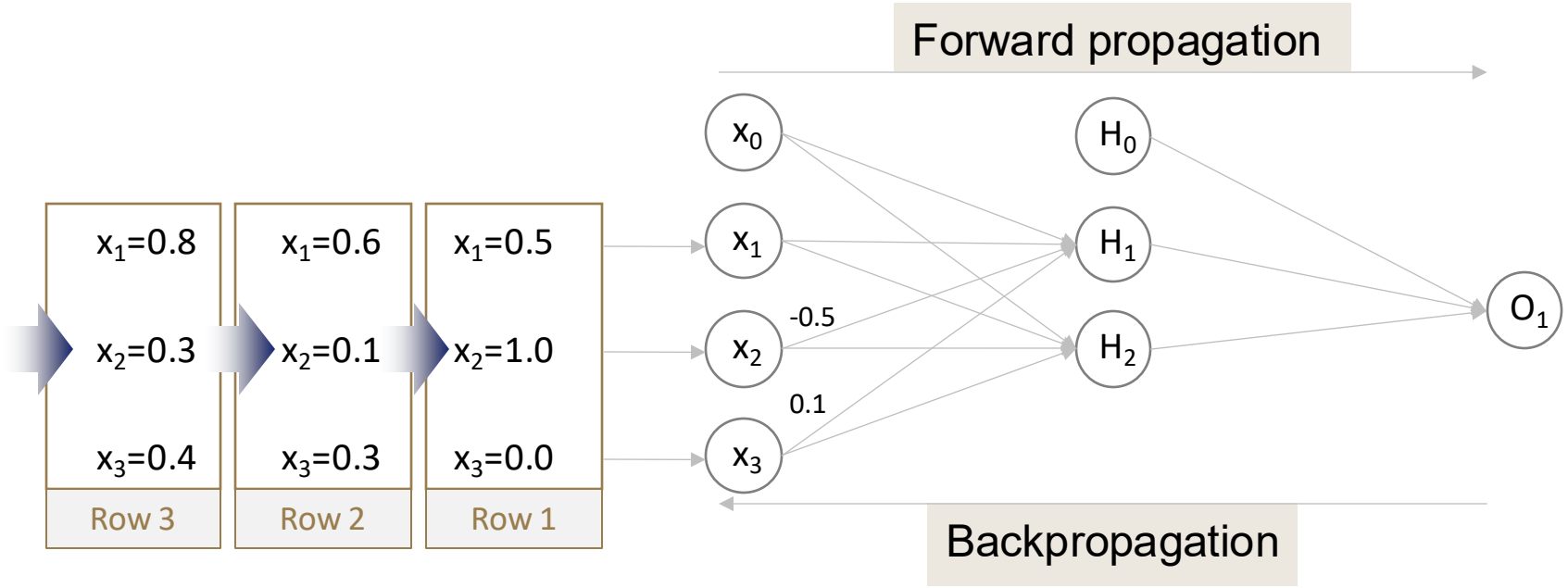


Backward Propagation



What is next after one iteration?

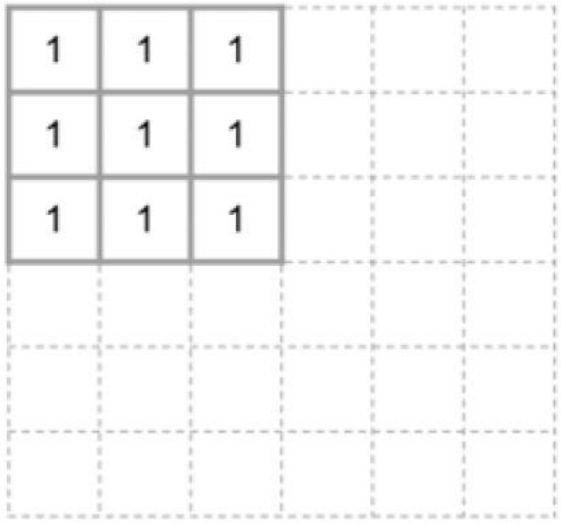
- One forward propagation and one backpropagation conducted with one observation
- New weights are not too different from original weights
  - predictive error E_1 is not large
  - learning rate η is intentionally set as 0.5 and not large
- Next observation can be fed into the ANN and update all weights again slowly



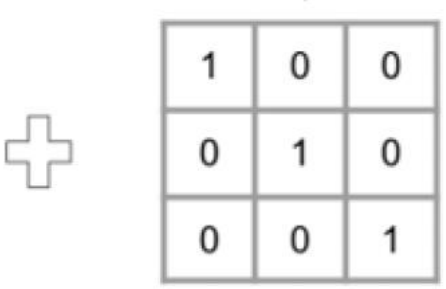
Convolutional Neural Network (CNN)

Convolution operation

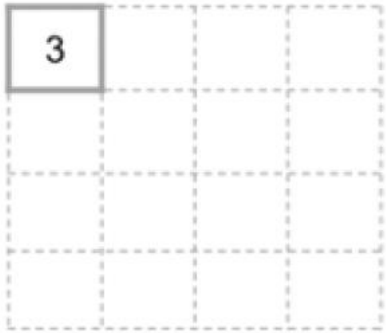
A specific local region (patch) of the grayscale input image data shown in solid line, with the rest of the image shown in dashed line. Each cell holds a pixel value between 0 and 255.



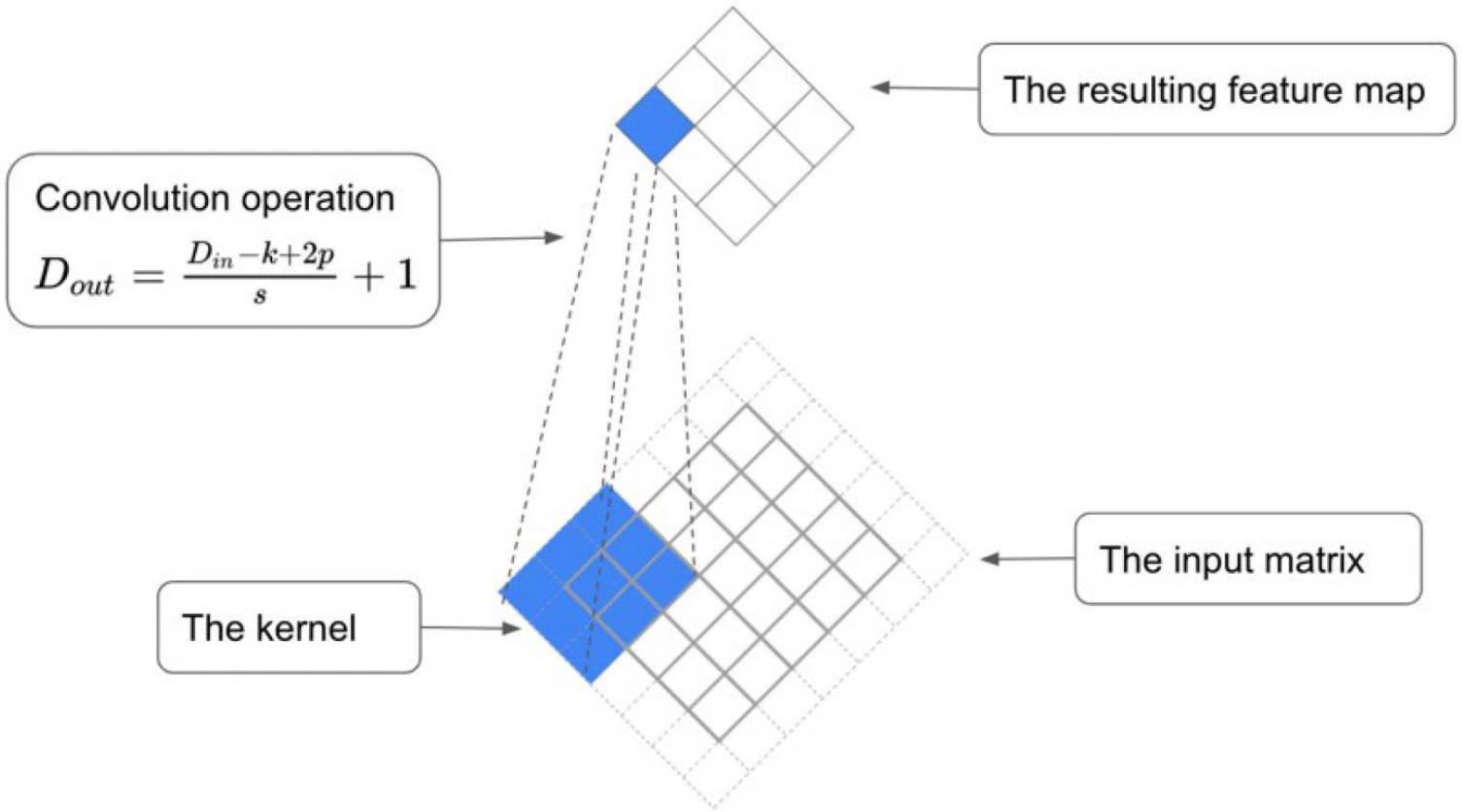
A 3x3 kernel defined by a convolution layer. Each cell in the kernel holds a weight parameter. The kernel will convolve with the input patch via a weighted sum operation, i.e., element-wise multiplication and summation.



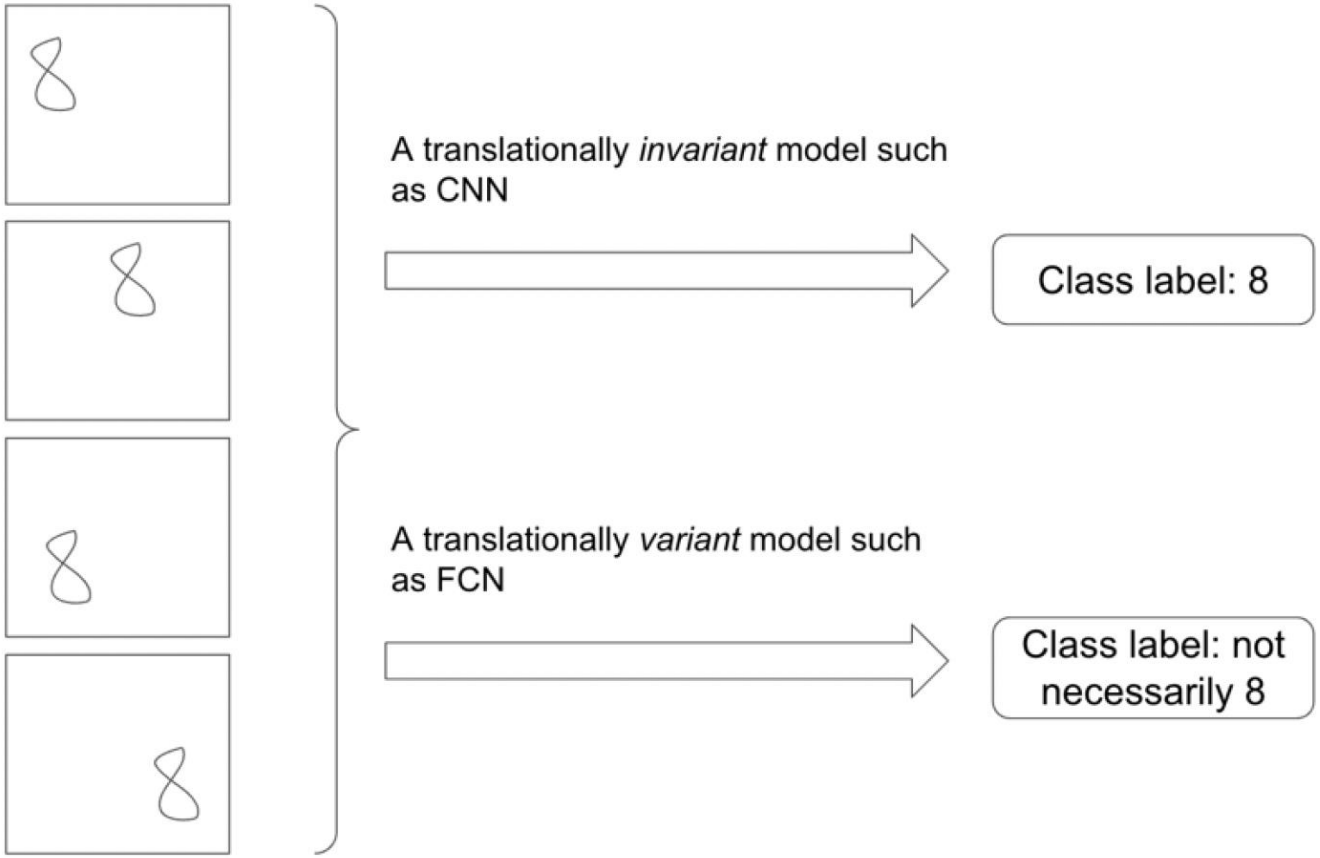
The resulting feature map, with the example convoluted feature in solid bold and others in dashed line.



Sliding the kernel to produce feature maps



Translational invariance




Coding session

In-class quiz

Homework

Post learning reflections and questions if any

Complete group assignment and submit as a group before week 7 lecture

Continue to work on final group project