

---

# DEEP LEARNING 2 – READING REVIEW – FEED FORWARD NEURAL NETWORKS AND REGULARIZATION.

## INTRODUCTION

This assignment is non-coding. Please read chapters 6 and 7 from the deep learning book: <https://www.deeplearningbook.org/>

## QUESTIONS

Deep feedforward networks:

1. Depth versus width of a neural network
  - a. Define the depth and width of a neural network in lay terms.
  - b. How might you choose depth/width of a network?
2. Forward propagation
  - a. What is the final output from forward propagation?
  - b. What is the input?
  - c. Define forward propagation in lay terms.
3. Back propagation
  - a. What is the goal of backpropagation? How does it differ from forward propagation?
  - b. Define backpropagation in lay terms.
4. Activation functions
  - a. Describe 3 different activation functions potentially used by a hidden layer. How do they differ?
5. Activation functions for output units versus hidden layers
  - a. Why is the choice of activation function for the output unit (output layer) more constrained compared to the activation functions for hidden layers? What considerations are there when choosing this activation function?
6. Universal approximation function theorem
  - a. Does the universal approximation function theorem mean we should prefer shallow, wide networks? Why or why not?

## Regularization:

7. When implementing parameter norm penalties, we could fit a different penalty strength for every neuron or layer in the network. Why might we decide against that and fit the same penalty across all layers?
8. When implementing a parameter norm penalty (like L1 or L2), do we penalize the bias terms for each layer, or just the weights? Why?
9. What are the different effects of L1 versus L2 regularization in deep learning?
10. Explain why early stopping can be considered a form of regularization. How do we know when to stop?
11. Pick one regularization technique covered in the chapter (other than parameter norm penalties like L1/L2 or early stopping). Briefly explain what the technique is and why it works to regularize the model.