

# Chapter 4

Simon Ward-Jones  
simonwardjones16@gmail.com

September 6, 2020

[Link to chapter](#)

## 1 Multilayer Perceptrons

- Incorporating Hidden Layers to network architecture
- Adding non linearity with activation functions

## 2 Implementation of Multilayer Perceptrons from Scratch

Similar to logistic/linear regression implementation involving

- Define the model (net) and initialise params
- Define the loss
- Define the optimizer (stepping in the direction of neg gradient)
- Implement training loop (using efficient mini batch data loaders)

## 3 Model Selection, Underfitting, and Overfitting

- Training Error and Generalization Error
- Model Selection and validation error
- K-Fold Cross-Validation

## 4 Weight decay

- Weight decay reduces weights to reduce complexity
- Commonly  $L_2$  decay is used

## 5 Dropout

### References

Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2020). *Dive into deep learning*.  
(<https://d2l.ai>)