

# FIT5201 - Data Analysis Algorithms

## Assignment 1

Name: Nisheeth Shah

Id: 29599644

Email: nsha0020@student.monash.edu

### Section C. Probabilistic Machine Learning

In this section, you show your knowledge about the foundation of the probabilistic machine learning (i.e. probabilistic inference and modeling) by solving two simple but basic statistical inference problems. Solve the following problems based on the probability concepts you have learned in Module 1 with the same math conventions.

#### Question 4 [Bayes Rule, 20 Marks]

Recall the simple example from Appendix A of Module 1. Suppose we have one red and one blue box. In the red box we have 2 apples and 6 oranges, whilst in the blue box we have 3 apples and 1 orange. Now suppose we randomly selected one of the boxes and picked a fruit. If the picked fruit is an apple, what is the probability that it was picked from the blue box?

Note that the chance of picking the red box is 40% and the selection chance for any of the pieces from a box is equal for all the pieces in that box. Please show your work in your PDF report.

Answer:

$$P(\text{Red}) = 4/10$$

$$P(\text{Blue}) = 6/10$$

$$P(\text{Apples} \mid \text{Red}) = 1/4$$

$$P(\text{Oranges} \mid \text{Red}) = 3/4$$

$$P(\text{Apples} \mid \text{Blue}) = 3/4$$

$$P(\text{Oranges} \mid \text{Blue}) = 1/4$$

$$\begin{aligned} P(\text{Blue} \mid \text{Apples}) &= \frac{P(\text{Apples} \mid \text{Blue}) * P(\text{Blue})}{P(\text{Apples} \mid \text{Blue}) * P(\text{Blue}) + P(\text{Apples} \mid \text{Red}) * P(\text{Red})} \\ &= \frac{(3/4) * (6/10)}{(3/4) * (6/10) + (1/4) * (4/10)} \end{aligned}$$

$$= \frac{9}{11} = 0.818$$

The probability that fruit picked from blue box given that fruit is apple is 82%.

## Section D. Ridge Regression

In this section, you develop Ridge Regression by adding the L2 norm regularization to the linear regression (covered in Activity 1 of Module 2). This section assesses your mathematical skills (derivation) and programming skills.

### Question 6 [Ridge Regression, 25 Marks]

- I. Given the gradient descent algorithms for linear regression (discussed in Chapter 2 of Module 2), derive weight update steps of stochastic gradient descent (SGD) as well as batch gradient descent (BGD) for linear regression with L2 regularisation norm. Show your work with enough explanation in your PDF report; you should provide the steps of SGD and BGD, separately.

Answer: