**Overview**

This assessment aims to assess and address the following learning outcome(s):

• understanding the roles of linear algebra, probability theory and optimisation in the realm of machines learning; understanding algorithms underpinning various optimisation methods;

• applying and implementing concepts in linear algebra, probability theory and optimisation using R.

**Submission**

You will need to submit the following:

• A PDF file clearly shows the assignment question, the associated answers, any relevant R outputs, analyses and discussions.

• **R or Rmarkdown** script file to reproduce your work.

• Format of PDF file: single line space, 12 pt Calibri (body) style.

**A word on plagiarism:**

Plagiarism is the act of using another’s words, works or ideas from any source as one’s own, this includes the use of large language model, such as ChatGPT. Plagiarism has no place in a business submission. Business work containing plagiarised material will be subject to formal penalty.

**1 Probabilities**

**Question 1**

A survey of holiday accommodation in New Zealand shows that in Auckland, excluding caravan parks and camping grounds, 25% were full (i.e. had no vacancies). A sample of 90 holiday accommodation facilities (of the same capacity) in Auckland is selected at random. If the survey result is reliable,

(a) What is the probability that no more 60 of these are full?

(b) What is the probability that exactly 35 of these are full?

(c) How many of the 90 holiday accommodations are expected to be full? (i.e. mean) Provide a detailed description to demonstrate your understanding

**Question 2**

In this question, we consider the marketing dataset from datarium package in R.







1

The data contains 200 observations and 4 variables. The response variable is sales, denoted as *Y*. The explanatory variables—measured in thousands of dollars—are advertising budget spent on youtube, newspapers and facebook, respectively, which are denoted as *X*1*, X*2 and *X*3, respectively.

To model the impact of the three media on sales, a researcher uses the following multiple linear regression:

*Y* = *β*0 + *β*1*X*1 + *β*2*X*2 + *β*3*X*3 + *ϵ,* (1)

= **X***β* + *ϵ* (2)

where *Y, X*1*, X*2 and *X*3 is a vector of *n* × 1 (*n* is the number of observations in the dataset), **X** = (**1***n,X*1*,X*2*,X*3) and *β* = [*β*0*,β*1*,β*2*,β*3]′.

The parameter *β* in Equation (2) and its standard error (*s.e*()) can be estimated by using the function lm() in R. Alternatively, *β* can be estimated by minimising the following loss function - mean squared errors:

(3)

It is well-known that the optimal solution of *β* in Equation (3), denoted as *β*ˆ, has the following form:

(4)

and its standard error is

where (5)

Your tasks are to:

(a) use equations (4) and (5) to estimate *β* and its standard error in R.

(b) compare the results obtained in Question 3(a) with those obtained from the function lm() in R.

**2 Optimisation**

**Question 3**

Another approach to estimate *β* in Equation (3) is to use Classical Gradient Descent.

2

(a) Write down a step-by-step procedure of Classical Gradient Descent to estimate *β* in Equation (3)

(b) Write an R code to implement the Classical Gradient Descent procedure provided in Question 2(a).

(c) Discuss the results obtained from Question 3(b) and compare it with that obtained from Question 2(a).

3