School of Computing and Information Systems
The University of Melbourne
COMP30027 MACHINE LEARNING (Semester 1, 2019)

Tutorial exercises: Week 7

1. What is the **Gradient Descent** method, and why is it important?
2. What is **Regression**? How is it similar to **Classification**, and how is it different?
   (a) What is **Linear Regression**? In what circumstances is it desirable, and it what circumstances is it undesirable?
   (b) How do we build a (linear) regression model? What is **RSS** and what advantages does it have over (some) alternatives?
3. Recall that the update rule for Gradient Descent with respect to RSS is as follows:

$$\beta_k^{i+1} \quad := \quad \beta_k^i + 2\alpha \sum_{j=1}^{N} x_{jk}(y_j - \hat{y_j^i})$$

   Build a Linear Regression model, using the following instances:

x	y
1	1
2	2
2	3

4. What is **Logistic Regression**?
   (a) How is Logistic Regression similar to **Naive Bayes** and how is it different? In what circumstances would the former be preferable, and in what circumstances would the latter?
   (b) What is "logistic"? What are we "regressing"?
   (c) How do we train a Logistic Regression model? In particular, what is the significance of the following:

$$\text{argmax}_\beta \sum_{i=1}^{n} y_i \log h_\beta(x_i) + (1 - y_i) \log(1 - h_\beta(x_i))$$