# EXAMPLES SHEET

David Wong

October 2020

These are some exercises designed to test your understanding of k-nearest neighbours and SVMs.

These exercises are of a similar difficult to what you will encounter in the final assessment. We have chosen to make these exercises *long-form* (rather than multiple choice), to encourage you to work through calculations carefully.

*answers are provided in a separate sheet, but try to attempt the questions first before looking at the solutions*

## Question 1

| Age ($x_0$) | Income ($x_1$) | Cats ($x_2$) | Dogs ($x_3$) | Married? ($t$) |
|---|---|---|---|---|
| 15 | 1000 | 0 | 0 | No |
| 26 | 20000 | 1 | 0 | No |
| 38 | 50000 | 0 | 1 | Yes |
| 41 | 48000 | 0 | 1 | Yes |
| 55 | 60000 | 0 | 1 | Yes |
| 60 | 30000 | 2 | 0 | No |
| 72 | 10000 | 4 | 0 | No |

This question uses the data in the table, above

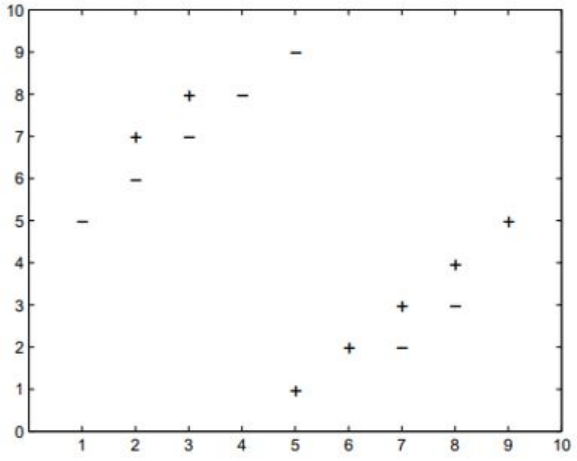
1. Briefly describe the K-nearest Neighbour learning method
2. What *training* error would a nearest neighbour classifier with $K = 1$ have for this data and why?
3. calculate the training error when $K = 3$
4. for the example: $[x_0, x_1, x_2, x_3] = [38, 38000, 0, 1]$ what is the predicted output, $t$, when $K = 1$? (assuming that we use a euclidean distance metric)
5. for the same example, what is the predicted output when $K = 3$?

1. (1) start with a data set with known categories
(2) add a new cell with unknown category
(3) classify the new cell by looking at the nearest annotated cells.

2. 0. because it is overfitted and it will choose itself.
Actually when k = 1, it produce decision boundaries and only this point in the boundary

3.0 (Income)
4. No.
5. Yes.

## Question 2




In the following questions, you will consider a K-nearest neighbour that uses Euclidean distance to classify the data as '+' or '-'.

1. What value of $K$ minimises the training set error for this dataset?
2. What is the resulting training error?
3. What value(s) of K minimise the leave-one-out cross-validation error for this data set?
4. Sketch the decision boundary for K-nearest neighbours in the case of $K = 1$

1. K = 1 can make any error = 0
2. 0
3. K=5,7. You should be able to understand that this will lead to an error of 4/14

## Question 3

1. What is the main difference between the decision boundary induced by a perceptron and the decision boundary induced by a *linear* SVM with hard margins?

2. Consider a set of training data, $\mathbf{x}$ with two features $x_1$ and $x_2$, with corresponding labels, $\mathbf{y}$. If the training data are linearly separable, show that the margin is $2/|\mathbf{w}|$

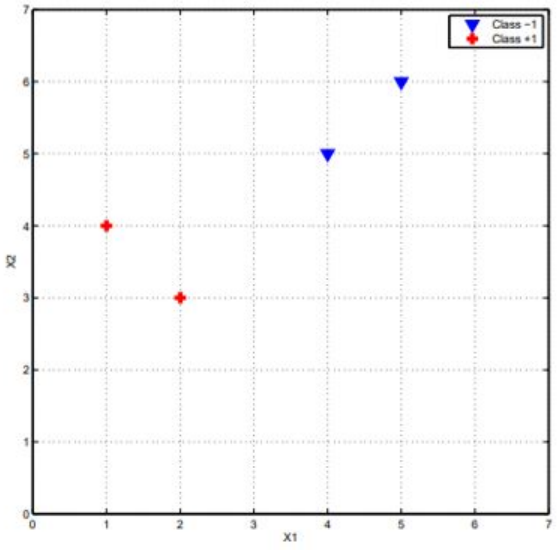
3. The SVM error function with soft margins can be written as:

$$E = \sum_{i=1}^{N} max\{0, 1 - y_i f(\mathbf{x}_i - \xi_i)\} + \frac{1}{2}\sum_{j=1}^{d} w_j^2 + C\sum_{i=1}^{N} \xi_i$$

   Define the terms of this expression, i.e. state what is meant by $w_j$, $f(\mathbf{X}_i)$,$C$,$\xi_i$. State also what is achieved by minimising the first part, versus the second part, versus the third part of the expression (i.e. before and after the '+' symbols)

1. with hard margins it will seperate the data accurately,
 by a perceptron it allows some wrongs.
2.

## Question 4




You are given the training data shown in the plot. The dataset consists of two examples with class label -1 and two with class label +1

1. Find the weight vector $\mathbf{w}$ and bias $b$.
2. circle the support vectors and draw the decision boundary