Machine Learning – July 21, 2020

Time limit: **2 hours 15 minutes**.

EXERCISE 1

The data collected over several years led real estate agents to conclude that:

- The probability that a house has a high price, given that it is located downtown, is

$$P(HIGH|DOWN) = .9;$$

- The probability that a house both has a high price and is located outside downtown is

$$P(HIGH, OUT) = .1;$$

- the probability that a house is located outside downtown is

$$P(OUT) = .8.$$

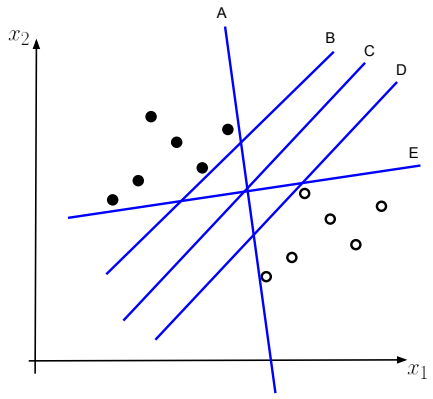
$HIGH$ indicates that the price of a house is high, $DOWN$ that the house is located downtown, and OUT that the house is not located downtown. Based on this:

1. What is the probability that a house has a low price (LOW, i.e., not $HIGH$)?
2. What is the probability that a house both is located downtown and has a low price?
3. If you know that a house has a low price, what is the probability that it is located downtown?

EXERCISE 2

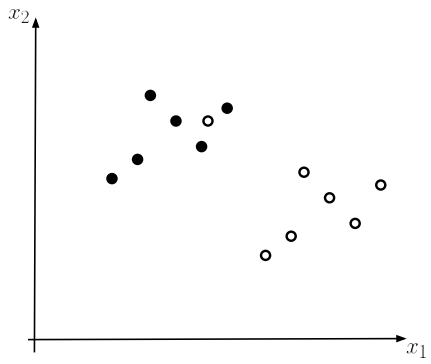
Consider Binary Classification.

1. Provide the definition of *linear discriminant function* and explain when its use is suitable for binary classification.
2. The following figure shows a dataset for binary classification with two classes (black and white).



The blue lines represent possible separation boundaries between the two classes. Which one among A, B, C, D, and E would you expect to be obtained through SVM? Explain why (indicating only the letter will not be accepted as an answer).

3. Consider the following dataset (notice, with respect to the previous dataset, the additional white point in the black cloud).



Would SVM, possibly adapted, be a suitable training method in this case? If so, define the corresponding mathematical model showing, in particular, the obtained error function and discussing the constraints over its variables.

EXERCISE 3

Consider a classification problem $f : \Re^2 \rightarrow \{0, 1\}$ with a dataset $D = \{(x_n, t_n)_{n=1}^N\}$

1. Describe the difference between parametric and non-parametric models.
2. Provide an example of a parametric model for the classification problem above.
3. Provide an example of a non-parametric model for the classification problem above.

EXERCISE 4

Consider a regression problem $f : X \rightarrow \Re$ with a dataset $D = \{(x_n, t_n)_{n=1}^N\}$, where f is known to be non-linear in x.

1. Describe a linear model for this problem and determine the trainable parameters.
2. Describe a solution of the problem in terms of least square error minimization.

EXERCISE 5

Consider the problem of estimating the age of a person given an image of his/her face. Design a Convolutional Neural Network (CNN) for addressing the problem by replying to the following questions:

1. What type of data the network takes as input and what is its output? (Provide data type, e.g. integer, float etc. and indicative dimensions)
2. Provide the list of layers you would use and for each layer describe its type (convolutional, pooling, fully-connected, etc.), the values of the layer's parameters and the corresponding activation function.

3. Explain what output activation function you would use.

EXERCISE 6

1. Explain the goal of dimensionality reduction. Give an example of dimensionality reduction using a problem of your choice.
2. Is using an Autoencoder equivalent to applying Principal Component Analysis for achieving dimensionality reduction? Explain why.
3. Which deep learning method can be used for learning the structure of a dataset in an unsupervised way?