University of Stavanger

Stavanger, January 2, 2020

Theoretical exercise 2

ELE520 Machine learning

A PDF version of the student solution of the exercise shall be submitted on CANVAS.

Problem 1

A company producing sea food plans to start using an automated classification system to detect possible toxic content in blue mussels. The classification will be based on chemical analyses of one mussel from each container in a batch of incoming containers of newly caught blue mussels, and the concentration x of a specific chemical substance in the mussels will be used as feature for classification. If this concentration is above a specific threshold value, the whole container will be considered poisonous and rejected. The company's gain for each container is NOK 250,-, but the company has also issued a guarantee to its customers that if they buy a container that shows itself to be poisonous the company will compensate costs up to NOK 100000,-.

a) Let ω_1='toxic mussels' og ω_2='non toxic mussels'. Formulate the loss functions $\lambda(\alpha_i|\omega_j), i = 1, 2, j = 1, 2$, where α_i corresponds to the decision ω_i (risk will be the same as cost in this setting) when we do not include the possibility of rejecting classifications.

   (You do not have to take into consideration that *correct classification* of non toxic mussels represents a gain which can be regarded as a "negative cost". Set the cost associated to this event to zero.!)

b) Dependent on the value of x we classify as ω_1 or ω_2. Express the conditional loss functions, $R(\alpha_i|x)), i = 1, 2$?

Trygve Eftestøl, Professor
Faculty of Science and Technology
Department of Electrical and Computer Engineering
Kjølv Egelands hus

University of Stavanger
N-4036 Stavanger.

Telephone: +47 51 83 10 00.
Telefax: +47 51 83 17 50.
E-mail: trygve.eftestol@uis.no
www.ux.his.no/~trygve-e

It is given that we know in advance that one of 25 containers contains toxic mussels, and that in toxic and non-toxic mussels the concentration x has a gaussian distribution $N(0.4, 0.0001)$ and $N(0.2, 0.0001)$ respectively.

c) Determine the decision boundary that minimises the overall loss (average cost) upon classification (R).

d) Determine the minimum average cost R upon classification in NOK.

Problem 2

We want to design a pattern recognition system that classifies 2-dimensional feature vectors $\{\boldsymbol{x}\}$ to class ω_1, or class ω_2.

It is known that the distribution between the two classes are 1/2 and 1/2 respectively for class ω_1 and class ω_2. Furthermore the feature vectors of the two classes are normally distributed around $\boldsymbol{\mu}_1 = (3\ 3)^t$ with covariance matrix

$$\boldsymbol{\Sigma}_1 = \begin{pmatrix} 1/2 & 0 \\ 0 & 2 \end{pmatrix} \tag{1}$$

for class ω_1, og around $\boldsymbol{\mu}_2 = (3\ \ -2)^t$ with covariance matrix

$$\boldsymbol{\Sigma}_2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \tag{2}$$

for class ω_2.

The eigenvalue and eigenvector matrices $\boldsymbol{\Lambda}_i$ and $\boldsymbol{\Phi}_i$, $i = 1, 2$ for the covariance matrices are

$$\boldsymbol{\Lambda}_1 = \begin{pmatrix} 1/2 & 0 \\ 0 & 2 \end{pmatrix} \tag{3}$$

and

$$\boldsymbol{\Phi}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{4}$$

for class ω_1, and

$$\boldsymbol{\Lambda}_2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \tag{5}$$

and

$$\boldsymbol{\Phi}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{6}$$

for class ω_2 respectively.

a) Sketch the contour lines for the class specific probability density functions for the two classes in the same diagram.

b) Formulate Bayes decision rule for this problem. Determine the decision boundary and decision regions in the same diagram as in a).

Problem 3

Show that with

$$g_i(\boldsymbol{x}) = -\frac{\|\boldsymbol{x} - \boldsymbol{\mu}_i\|^2}{2\sigma^2} + \ln P(\omega_i) \tag{1}$$

as starting point where $i = 1, 2$, we can reach an expression defining the decision border between the two classes:

$$\boldsymbol{\theta}^t(\boldsymbol{x} - \boldsymbol{x}_0) = 0 \tag{2}$$

where

$$\boldsymbol{\theta} = \boldsymbol{\mu}_i - \boldsymbol{\mu}_j \tag{3}$$

and

$$\boldsymbol{x}_0 = \frac{1}{2}(\boldsymbol{\mu}_i + \boldsymbol{\mu}_j) - \frac{\sigma^2}{\|\boldsymbol{\mu}_i - \boldsymbol{\mu}_j\|^2} \ln \frac{P(\omega_i)}{P(\omega_j)}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) \tag{4}$$