PRACTICE 7

EXERCISE 1: 2D DENSITY ESTIMATION

Using the IRIS database, and use only second and third feature:

a) Estimate the density function of each class as a 2D Gaussian, and analyze the obtained boundary. Compute the errors obtained.

b) Estimate the density function using Parzen windows and analyze 2D boundaries obtained. Compute the errors obtained

c) Which method works best in this case? Why?

EXERCISE 2: PREDICTION RISK

Knowing that the Gaussian density function is defined as:

$$P(X \mid w_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}}$$

We can calculate the decision boundary matching the a posteriori probabilities of both classes: `P(X|w1)*P(w1) = P(X|w2)*P(w2)`, or applying logarithm (e base): `log(P(X|w1)) + log(P(w1)) = log(P(X|w2)) + log(P(w2))`

a) Modify the follow code to find the decisión boundary taking into account the prediction risk:

```
    A=s1*s1-s2*s2;
    B=2*(m1*s2*s2-m2*s1*s1);
    C=2*s1*s1*s2*s2*(log(Pw1)-log(Pw2)-
log(s1)+log(s2))+s1*s1*m2*m2-s2*s2*m1*m1;
    x1=(-B+sqrt(B*B-4*A*C))/2/A
    x2=(-B-sqrt(B*B-4*A*C))/2/A
```

To solve this problem you must analyze where it comes the formula above, and properly make the necessary terms of cost. Remember that:

$$r_j(x) = \sum_{i=1}^{M} L_{ij} \cdot p(x \mid w_i) \cdot p(w_i)$$

maria.guerrero@uca.es
Departamento Ingeniería Informática

b) If we consider the risk to choose w1 being really w2 like 0.8, and choose w2 being really w1 equal to 2, determine the decision boundary.