

Probabilistic neural network (PNN)

Consider the problem of multi-class classification. We are given a set of data points from each class. The objective is to classify any new data sample into one of the classes.

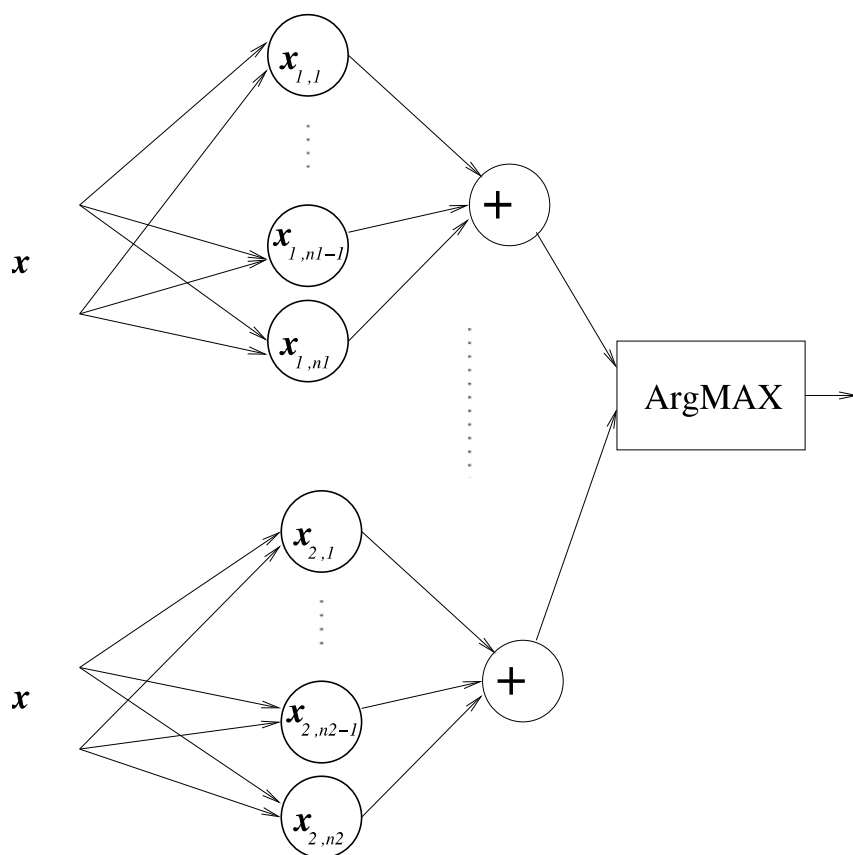


Figure above: A schematic illustration of a PNN.

Lecture 3

Probabilistic neural network (PNN) is closely related to Parzen window pdf estimator. A PNN consists of several sub-networks, each of which is a Parzen window pdf estimator for each of the classes.

The input nodes are the set of measurements.

The second layer consists of the Gaussian functions formed using the given set of data points as centers.

The third layer performs an average operation of the outputs from the second layer for each class.

The fourth layer performs a vote, selecting the largest value. The associated class label is then determined.

Lecture 3

Suppose that for class 1, there are five data points $x_{1,1} = 2$, $x_{1,2} = 2.5$, $x_{1,3} = 3$, $x_{1,4} = 1$ and $x_{1,5} = 6$. For class 2, there are 3 data points $x_{2,1} = 6$, $x_{2,2} = 6.5$, $x_{2,3} = 7$. Using the Gaussian window function with $\sigma = 1$, the Parzen pdf for class 1 and class 2 at x are

$$y_1(x) = \frac{1}{5} \sum_{i=1}^5 \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_{1,i} - x)^2}{2}\right)$$

and

$$y_2(x) = \frac{1}{3} \sum_{i=1}^3 \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_{2,i} - x)^2}{2}\right)$$

respectively. The PNN classifies a new x by comparing the values of $y_1(x)$ and $y_2(x)$. If $y_1(x) > y_2(x)$, then x is assigned to class 1; Otherwise class 2.

For this example $y_1(3) = 0.2103$ (see Lecture 2).

Lecture 3

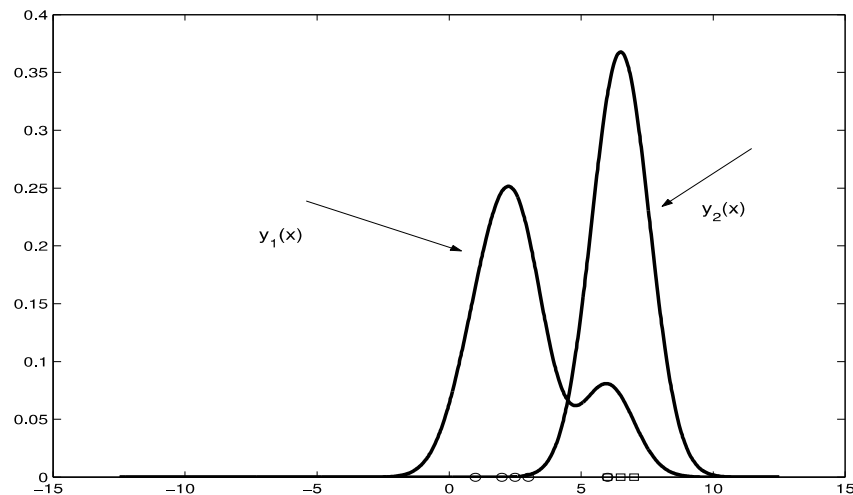


Figure above, Parzen window pdf for two classes.

$$\begin{aligned} y_2(3) &= \frac{1}{3\sqrt{2\pi}} \left\{ \exp\left(-\frac{(6-3)^2}{2}\right) \right. \\ &\quad + \exp\left(-\frac{(6.5-3)^2}{2}\right) \\ &\quad \left. + \exp\left(-\frac{(7-3)^2}{2}\right) \right\} \\ &= 0.0011 < 0.2103 = y_1(x) \end{aligned}$$

so the sample $x = 3$ will be classified as class 1 using PNN.

Lecture 3

The decision boundary of the PNN is given by

$$y_1(x) = y_2(x).$$

So

$$\begin{aligned} & \frac{1}{3\sqrt{2\pi}} \left\{ \exp\left(-\frac{(6-x)^2}{2}\right) + \exp\left(-\frac{(6.5-x)^2}{2}\right) \right. \\ & \quad \left. + \exp\left(-\frac{(7-x)^2}{2}\right) \right\} \\ &= \frac{1}{5\sqrt{2\pi}} \left\{ \exp\left(-\frac{(2-x)^2}{2}\right) + \exp\left(-\frac{(2.5-x)^2}{2}\right) \right. \\ & \quad \left. + \exp\left(-\frac{(3-x)^2}{2}\right) + \exp\left(-\frac{(1-x)^2}{2}\right) \right. \\ & \quad \left. + \exp\left(-\frac{(6-x)^2}{2}\right) \right\} \end{aligned}$$

The solution of x can be found numerically. e.g. a grid search. This is an optimal solution, minimizing misclassification rate.

Lecture 3

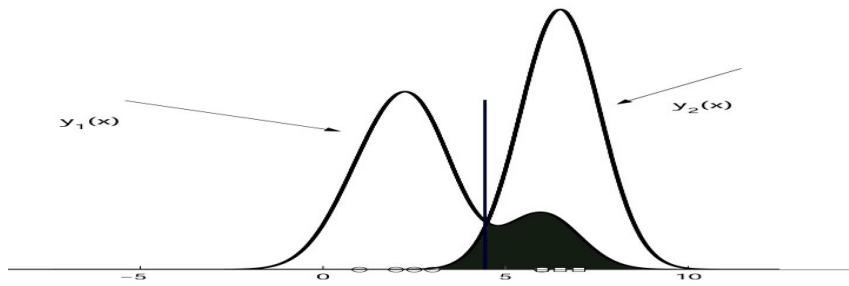
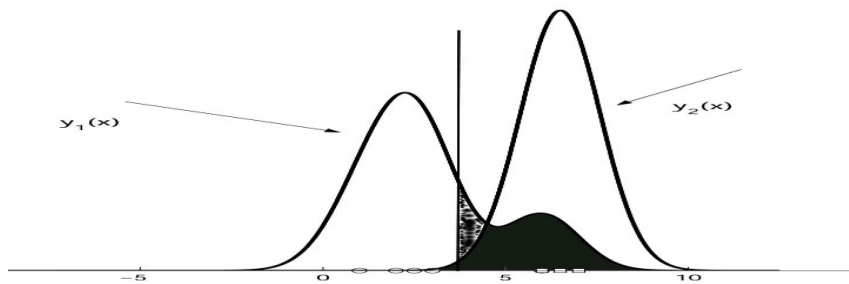
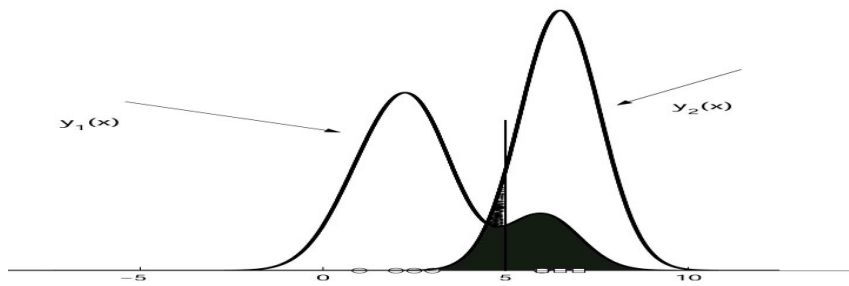


Figure above: Decision boundary and the error probability of PNN (the shaded area).

Moving the decision boundary to either side would increase the misclassification probability.



Lecture 3

Since the term $\frac{1}{\sqrt{2\pi}}$ is a common factor in both $y_1(x)$ and $y_2(x)$, it can be dropped out without changing the classification result. We can use

$$y_1(x) = \frac{1}{5} \sum_{i=1}^5 \exp \left(-\frac{(x_{1,i} - x)^2}{2} \right)$$

and

$$y_2(x) = \frac{1}{3} \sum_{i=1}^3 \exp \left(-\frac{(x_{2,i} - x)^2}{2} \right)$$

In general, a PNN for M classes is defined as

$$y_j(\mathbf{x}) = \frac{1}{n_j} \sum_{i=1}^{n_j} \exp \left(-\frac{(\|\mathbf{x}_{j,i} - \mathbf{x}\|)^2}{2\sigma^2} \right)$$

$$j = 1, \dots, M$$

where n_j denotes the number of data points in class j . The PNN assigns \mathbf{x} into class k if $y_k(\mathbf{x}) > y_j(\mathbf{x})$, $j \in [1, \dots, M]$. $\|\mathbf{x}_{j,i} - \mathbf{x}\|^2$ is calculated as the sum of squares. e.g. if $\mathbf{x}_{j,i} = [2, 4]^T$, $\mathbf{x} = [3, 1]^T$, then

$$\|\mathbf{x}_{j,i} - \mathbf{x}\|^2 = (2 - 3)^2 + (4 - 1)^2 = 10$$

Lecture 3

Example 1: Determine the class label for the data sample $\mathbf{x} = [0.5, 0.5]^T$ by using a two class PNN classier, with $\sigma = 1$, based on the two class data sets given in the following Table.

$\mathbf{x}_{1,1}$	1	0
$\mathbf{x}_{1,2}$	0	1
$\mathbf{x}_{1,3}$	1	1

class 1

$\mathbf{x}_{2,1}$	-1	0
$\mathbf{x}_{2,2}$	0	-1

class 2

Solution:

$$\begin{aligned} y_1(\mathbf{x}) &= \frac{1}{3} \left\{ \exp \left(-\frac{(1 - 0.5)^2 + (0 - 0.5)^2}{2} \right) \right. \\ &\quad + \exp \left(-\frac{(0 - 0.5)^2 + (1 - 0.5)^2}{2} \right) \\ &\quad \left. + \exp \left(-\frac{(1 - 0.5)^2 + (1 - 0.5)^2}{2} \right) \right\} \\ &= 0.7788 \end{aligned}$$

Lecture 3

$$\begin{aligned} y_2(\mathbf{x}) &= \frac{1}{2} \left\{ \exp \left(-\frac{(-1 - 0.5)^2 + (0 - 0.5)^2}{2} \right) \right. \\ &\quad \left. + \exp \left(-\frac{(0 - 0.5)^2 + (-1 - 0.5)^2}{2} \right) \right\} \\ &= 0.4724 \end{aligned}$$

Because $y_2(\mathbf{x}) < y_1(\mathbf{x})$, so $\mathbf{x} = [0.5, 0.5]^T$ is classified as Class 1.

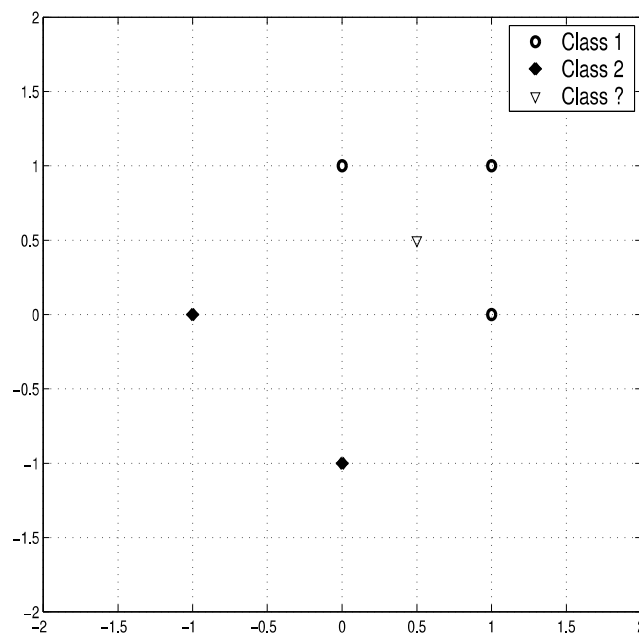


Figure above, Data points in Example 1.