RBF Networks

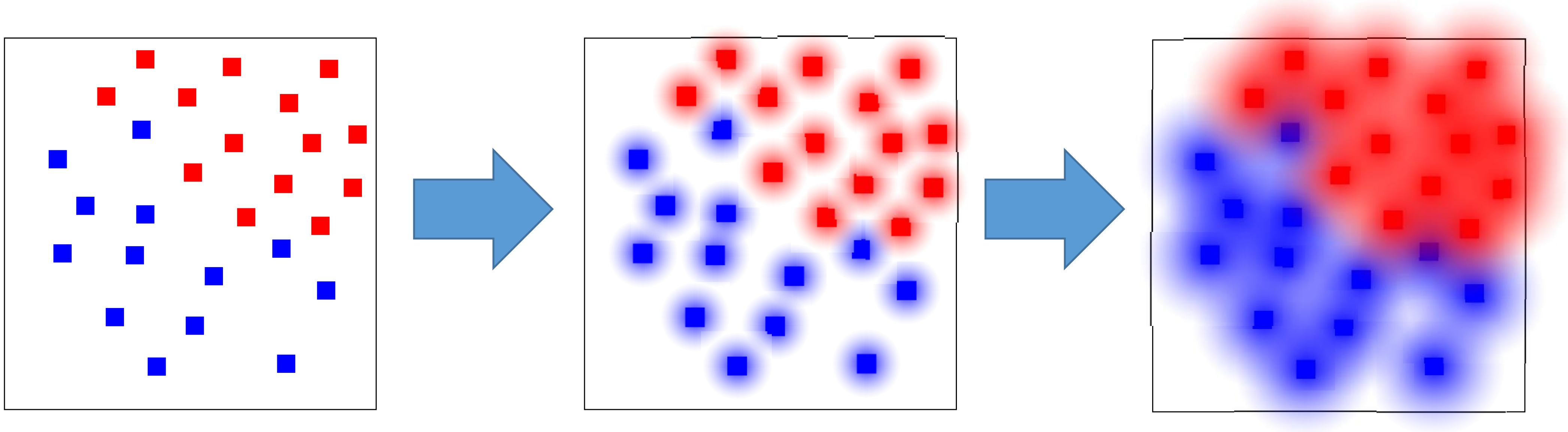
Intuition

Retours sur l'apprentissage « par cœur »

Intuition

Retours sur l'apprentissage « par cœur »

Conserver les exemples et attribuer une 'zone d'influence'



Principes

Régression RBF Naïf :

$$output(x) = \sum_{n=1}^{N} w_n e^{-\gamma||x-x_n||^2}$$

Classification RBF Naïf :

$$output(x) = sign(\sum_{n=1}^{N} w_n e^{-\gamma\left||x-x_n|\right|^2})$$

Principes

Régression RBF Naïf :

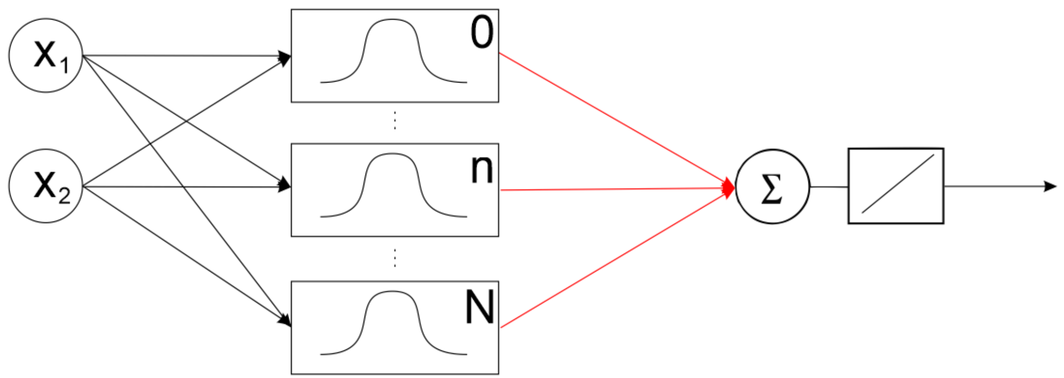
$$output(x) = \sum_{n=1}^{N} w_n e^{-\gamma ||x - x_n||^2}$$

Classification RBF Naïf :

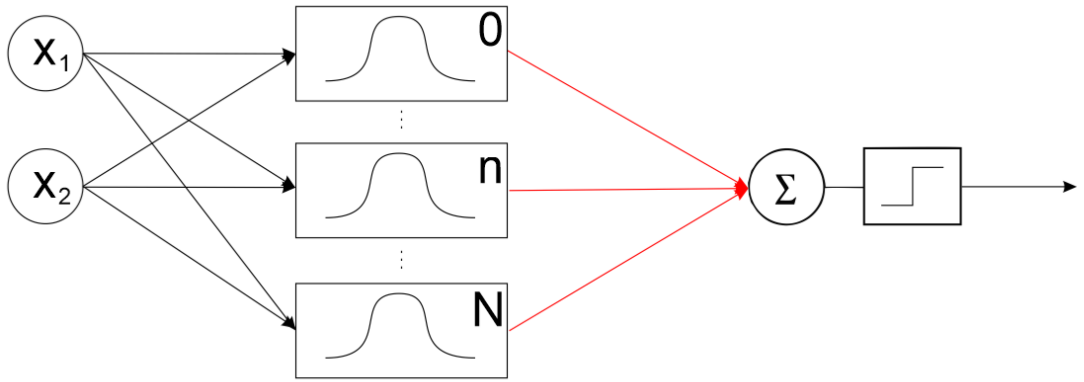
$$output(x) = sign(\sum_{n=1}^{N} w_n e^{-\gamma ||x - x_n||^2})$$

Principes

Régression RBF Naïf :



Classification RBF Naïf :



Principes :

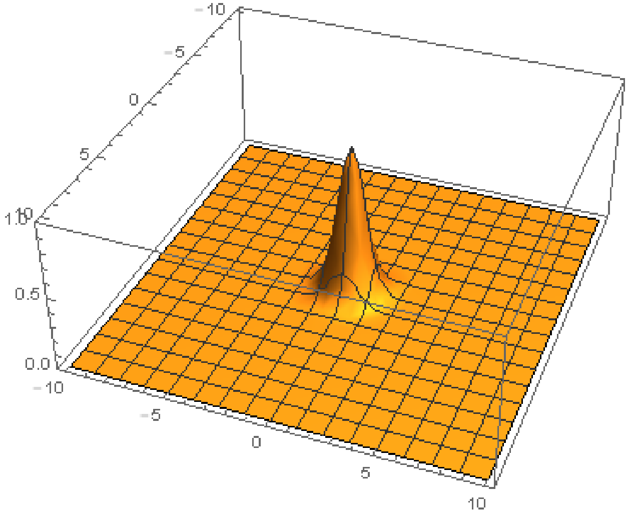
Trouver W pour un RBF naïf :

Soit $\phi = \begin{bmatrix} e^{-\gamma||x_1 - x_1||^2} & \cdots & e^{-\gamma||x_1 - x_N||^2} \\ \vdots & \ddots & \vdots \\ e^{-\gamma||x_N - x_1||^2} & \cdots & e^{-\gamma||x_N - x_N||^2} \end{bmatrix}$ et $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$

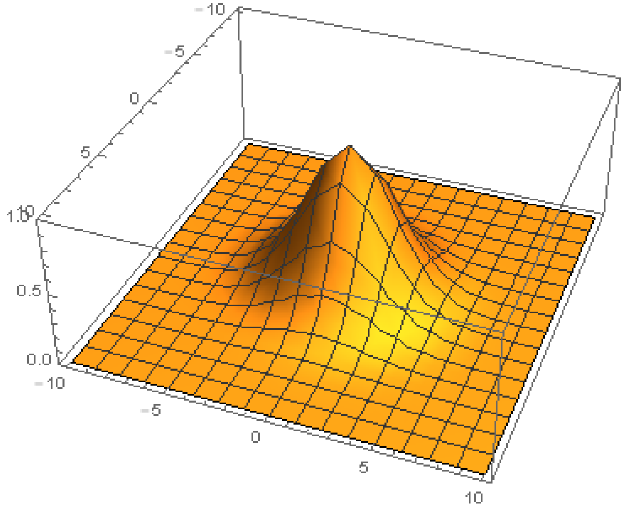
Alors $W = \phi^{-1} Y$

Principes :

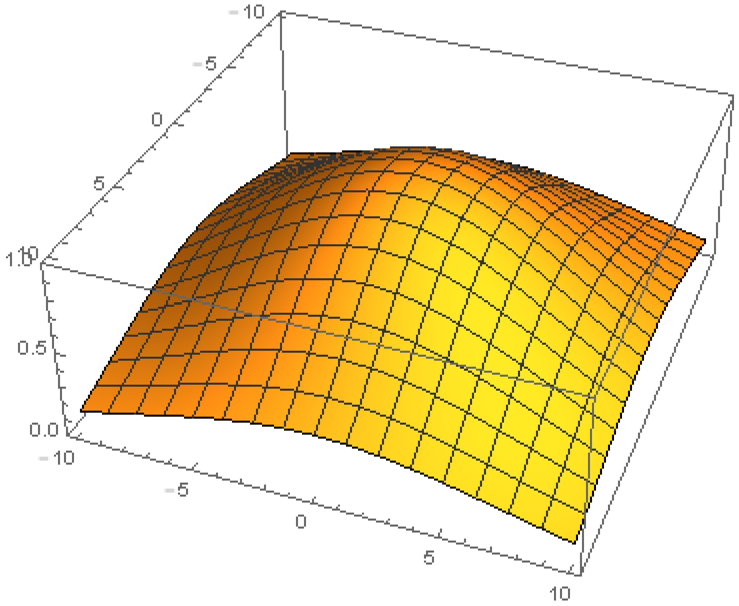
Impact du choix de Gamma :



$\gamma = 1$



$\gamma = 0{,}1$



$\gamma = 0{,}01$

Principes

n a d'exem… …isposition, mieux c'est ?

Principes

Plus on a d’exemples à disposition, mieux c’est ?

Nombre de w_i = nombre d’exemples !

Principes

Plus on a d’exemples à disposition, mieux c’est ?

Nombre de w_i = nombre d’exemples !

Mauvais signe pour la généralisation.

Intuition

Ne pas prendre tous les exemples !

Intuition

Ne pas prendre tous les exemples !

Elire des ‘représentants’

Principes

KNN (k-Nearest Neighbours)

Méthode exacte : NP-Difficile !

Algorithme de LLoyd

Répéter :

1 : $\mu_k = \frac{1}{|S_k|} \sum_{x_n \in S_k} x_n$

2 : $S_k = \{x_n tq\ \forall l, \|x_n - \mu_k\| \leq \|x_n - \mu_l\|\}$

RBF utilisant K centres

Trouver W pour un RBF utilisant K Centres :

Soit $\phi = \begin{bmatrix} e^{-\gamma||x_1-\mu_1||^2} & \cdots & e^{-\gamma||x_1-\mu_K||^2} \\ \vdots & \ddots & \vdots \\ e^{-\gamma||x_N-\mu_1||^2} & \cdots & e^{-\gamma||x_N-\mu_K||^2} \end{bmatrix}$ et $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$

Alors $W = (\phi^T\phi)^{-1}\phi^T Y$