Exercícios de Reconhecimento de Padrões

- **Objetivos**: Representação do KNN com ponderação de distâncias e entendimento desta abordagem como uma combinação de funções radiais.

- **Tarefa**: estender a implementação anterior incluindo a ponderação de distâncias para a tomada de decisão da classificação. De acordo com esta abordagem a resposta do KNN é representada na forma

$$\hat{y} = sinal(\sum_{i=1}^{N} \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i))$$

em que $\alpha_i = 1 \ \forall \mathbf{x}_i \in V_k$, $\alpha_i = 0 \ \forall \mathbf{x}_i \notin V_k$, V_k é o conjunto dos vizinhos mais próximos e $K(\mathbf{x}, \mathbf{x}_i)$ é um função de kernel radial na forma de uma função Gaussiana.

Avaliar os efeitos dos parâmetros do modelo e da função no desempenho do modelo, considerando a resposta observada da superfície de separação. Para o exemplo de dados sintéticos implementado, a função Gaussiana (Normal) é de duas variáveis, no entanto, sugere-se que ela seja implementada em sua forma geral:

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \mathbf{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})\right) \tag{1}$$

em que n é a dimensão de $\mathbf{x}$, $\mathbf{\Sigma}$ é a matriz de covariâncias, $|\mathbf{\Sigma}|$ o seu determinante e $\boldsymbol{\mu}$ é o vetor de médias das distribuições marginais, conforme apresentado a seguir nas Equações 2 e 3.

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \rho_{21}\sigma_2\sigma_1 & \sigma_2^2 & \cdots & \rho_{2n}\sigma_2\sigma_n \\ \vdots & \vdots & \cdots & \vdots \\ \rho_{n1}\sigma_n\sigma_1 & \rho_{n2}\sigma_n\sigma_2 & \cdots & \sigma_n^2 \end{bmatrix} \tag{2}$$

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} \tag{3}$$

Para facilitar a implementação, considere a seguinte implementação em R da função de densidade Normal multivariada:

```
> pdfnvar<-function(x,m,K,n) ((1/(sqrt((2*pi)^n*(det(K)))))*
+                              exp(-0.5*(t(x-m) %*%
+                              (solve(K)) %*% (x-m))))
```

em que $\mathbf{K} = h\mathbf{I}$, sendo assim h o raio da função Gaussiana.

Nesta etapa o desempenho do classificador será avaliado de acordo com os parâmetros k e h.

- **Apresentar**: Gráficos diversos com diferentes respostas do modelo para valores diferentes de k e h. Apresentar em um parágrafo de no máximo 10 linhas uma discussão dos resultados.