Exercício Reconhecimento de Padrões KDE

Lucas Ribeiro da Silva - 2022055564

Universidade Federal de Minas Gerais Belo Horizonte - Minas Gerais - Brasil

lucasrsilvak@ufmg.br

1 Introdução

Nesse relatório consta a realização e implementação da estimação de densidades a partir do método KDE, a aplicação do classificador de Bayes na resolução de um problema de classificação univariado.

2 Criação de Conjuntos

Foram criados dois conjuntos, de centros em (2,0) e (4,0) respectivamente e com desvio padrão de 0.75. As classes apresentam superposição e cada uma possui 1000 elementos.

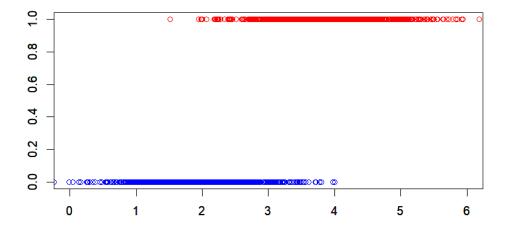


Figura 1: Amostras

Depois, separou-se o modelo em treinamentos e testes, na proporção 10:1.

3 KDE para diferentes h

 ${\rm Em}$ seguida, observou-se a variação do parâmetro h
 para observar o h ${\rm com}$ a melhor acurácia.

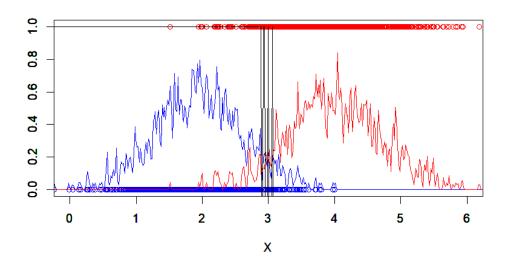


Figura 2: h = 0.01 | Acurácia: 92,5%

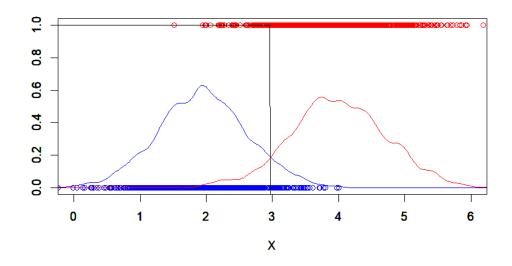


Figura 3: h = 0.1 | Acurácia: 92,5%

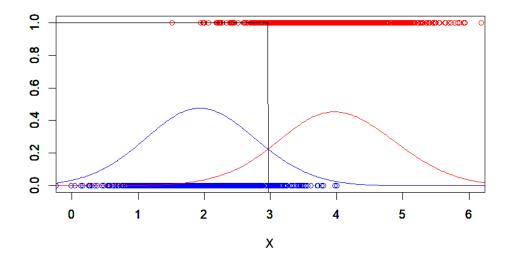


Figura 4: h = 0.5 | Acurácia: 92,5%

Percebeu-se que alterações em h para valores maiores que 0.01 não produziram alterações no resultado. A função de densidade torna-se progressivamente mais suave conforme o parâmetro h aumenta.

4 Variando Conjuntos de Treinamento

Alterando a proporção dos conjuntos de treinamento de 10:1 para 5:1, 10:3 e 2:1 e reduzindo-se para 100 amostras obteve-se as seguintes acurácias para os valores de h verificados anteriormente.

h	10%	20%	30%	50%
0.01	95.00	92.50	86.67	92.00
0.1	100.00	92.50	88.33	94.00
0.5	100.00	97.50	93.33	91.00

Tabela 1: Acurácia (%) em função da largura de banda h e da proporção de dados de teste

5 Conclusões

O h ótimo tende a suavizar a função de densidade de probabilidade, e no caso do problema observado, o melhor h foi aquele que suavizou melhor a curva de probabilidade. Entretanto, é sempre necessário prestar atenção nas configurações do problema e na relação teste/treinamento/aplicação para definir a melhor utilização do hiperparâmetro h em relação a sua aplicação.