



Exercício 1

No exercício proposto, foram amostrados 400 dados ao todo (360 para treinamento e 40 para testes) em torno de dois pontos $(2, 2, \sigma)$ e $(4, 4, \sigma)$, considerados dados de classes balanceadas **vermelha**, obtido considerando o desvio padrão igual a $\sigma = 0.8$, e **azul**, obtido considerando o desvio padrão igual a $\sigma = 0.4$, como mostra a figura 1. Os dados em formato de cruz, representados pela cor **verde**, são os dados que serão utilizados para teste do algoritmo. Os dados de treinamento representam 90% dos dados totais, enquanto os dados de teste representam os 10% restantes.

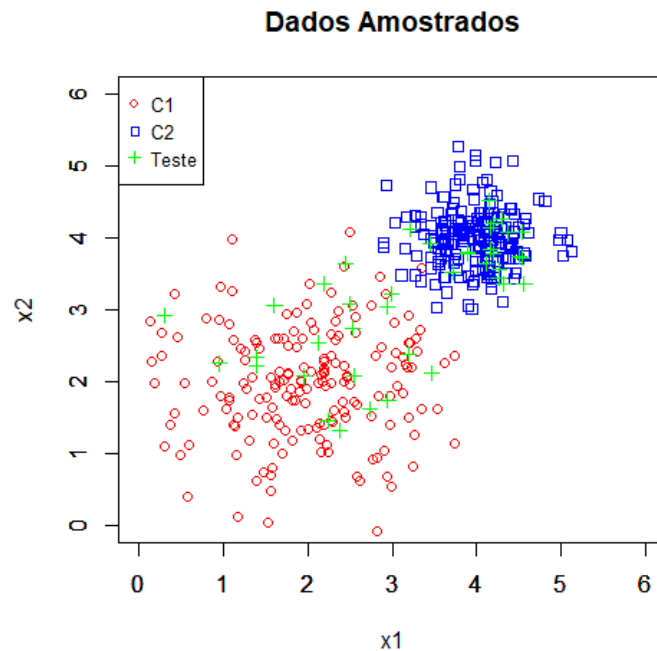


Figura 1: Dados amostrados para o exercício 1.

Utilizando um Classificador de Bayes, o algoritmo classificou corretamente 100% das amostras, obtendo um erro nulo. O resultado da superfície de contorno é mostrada na figura 2.

O resultado mostra claramente que, devido ao desvio padrão maior nos dados da classe C1, sua distribuição ocorreu de forma menos uniforme. Com isso, o algoritmo reservou uma área maior à essa classe. Assim, mesmo que a probabilidade de pertencer à qualquer uma das classes seja a mesma (50%, já que as classes são balanceadas), a influência do desvio padrão é perceptível no Classificador Bayesiano. A figura 3 mostra um exemplo de como seria a superfície de contorno caso o desvio padrão das classes fosse o mesmo ($\sigma = 0.4$), que se assemelha mais a uma reta.

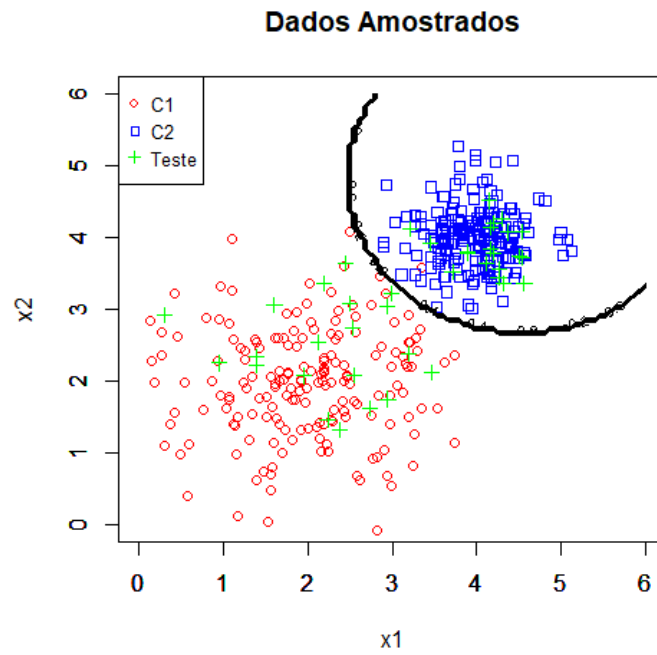


Figura 2: Resultado da curva de separação das classes.

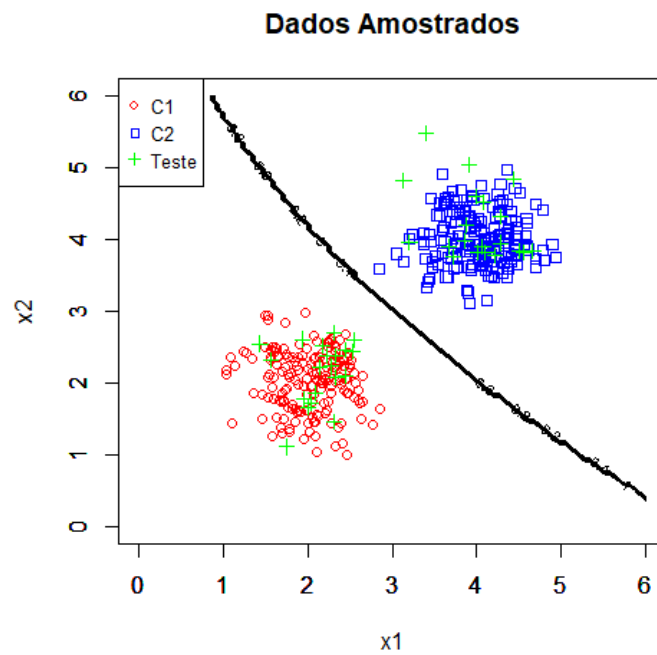


Figura 3: Resultado da curva de separação das classes com desvio padrão idêntico.

Exercício 2

No segundo exercício as amostras foram feitas em torno de uma gaussiana, como mostra a figura 4. Estas amostras foram feitas em torno dos pontos (2,2) e (4,4) para a classe **vermelha** e (2,4) e (4,2) para a classe **azul**, ambas com o desvio padrão $\sigma = 0.3$. Assim como no exercício 1, as amostras em **verde** são as amostras de teste, que correspondem a 10% do total de amostras, enquanto as demais são amostras

de treinamento. O total de amostras foi de 800, de forma balanceada entre as duas classes e as quatro gaussianas formadas.

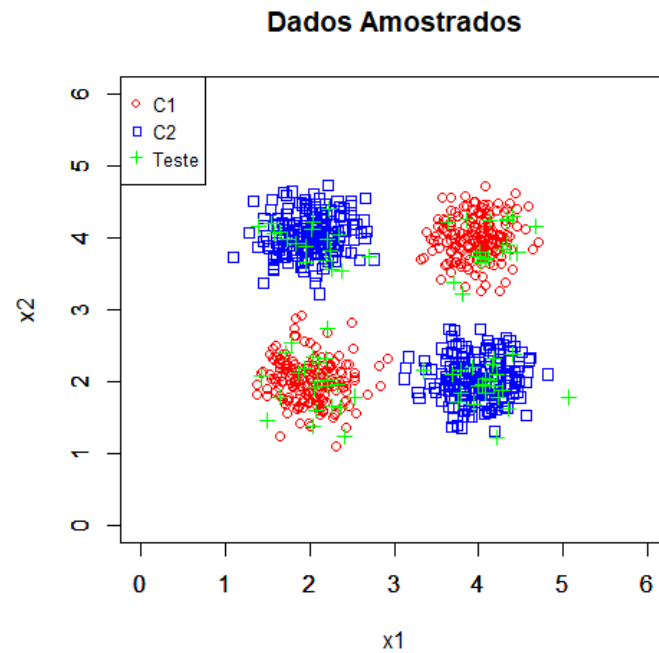


Figura 4: Dados amostrados do exercício 2.

Após a aplicação de um classificador Bayesiano, obteve-se novamente um resultado perfeito, em que 100% das amostras de teste retornaram resultados precisos, logo, obtendo-se erro nulo. A figura 5 mostra o resultado do exercício 2.

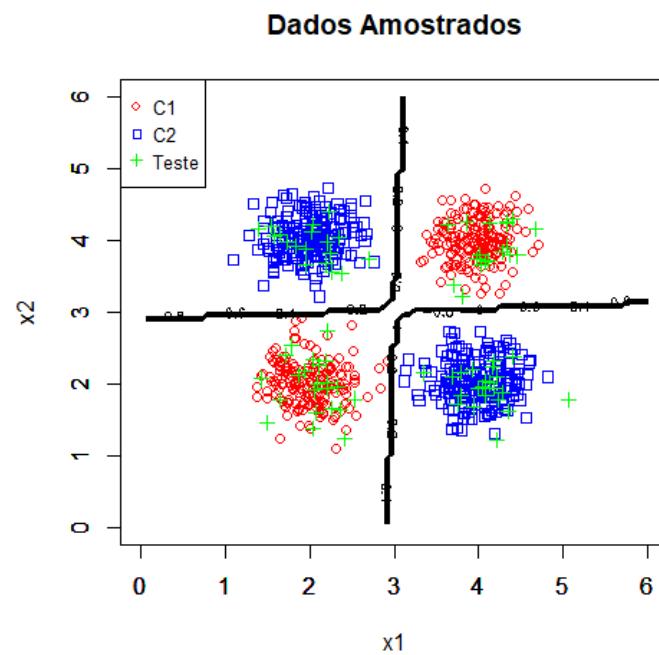


Figura 5: Resultado do exercício 2.

Ainda que o problema tenha obtido uma solução totalmente correta, pode-se perceber a dificuldade de

classificar um dado que encontra-se em torno do ponto $(3, 3)$, em que o classificador considerou pertencer apenas a região da classe C1. Ao repetir o experimento algumas vezes para distribuições aleatórias distintas, foi possível fazer um gráfico contrário ao da figura 5, como mostra a figura 6. Isto mostra a fragilidade do classificador Bayesiano, que não depende apenas da probabilidade à priori de cada classe, mas também da distribuição e do balanceamento dos dados utilizados.

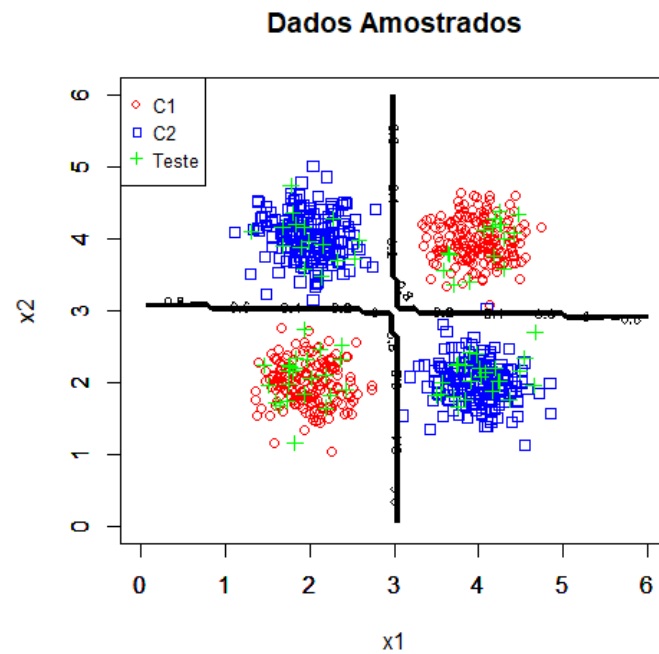


Figura 6: Resultado do exercício 2 em um segundo teste.