

Lista 4

Classificadores Estatísticos

Instruções

Deverá ser enviado ao professor, um arquivo texto contendo os gráficos, resultados e comentários requeridos em cada item.

1. Classificação utilizando Naive Bayes e DQG

- Carregue os dados contidos no arquivo DadosLista4.mat.

O arquivo contém uma matriz de dados e o rótulo associado a cada dado. A matriz de dados (Dados) é composta de 1500 linhas e 2 colunas. A matriz dos rótulos (y) apresenta os rótulos das classes. Nestes dados, existem 3 classes, sendo 500 exemplares de cada classe.

- Divida aleatoriamente o conjunto de dados entre treino e teste. Para este problema, utilize 600 (200 de cada classe) dados para treino, 900 (300 de cada classe) dados para teste.

- Implemente os seguintes classificadores: Naive Bayes e Discriminante Quadrático Gaussiano

Apresentar: As matrizes de confusão para os dois classificadores.

Naive Bayes - Matriz de Confusão

| classes | 1 | 2 | 3 |
|---------|-----|-----|-----|
| 1 | 177 | 23 | 0 |
| 2 | 3 | 197 | 0 |
| 3 | 0 | 0 | 200 |

Discriminante Quadrático Gaussiano (DQG) - Matriz de Confusão

| classes | 1 | 2 | 3 |
|---------|-----|-----|-----|
| 1 | 170 | 30 | 0 |
| 2 | 4 | 196 | 0 |
| 3 | 0 | 0 | 200 |

Comentários: Com base nos resultados, comente sobre a hipótese de cada uma das classes ser linearmente separável.

Devido a não-linearidade na separação dos exemplos da classe 1 e classe 2 (Figura 1), os dois classificadores classificaram alguns falsos positivos entres essas duas classes. Como o DQG é um classificador mais eficiente para dados linearmente separáveis, ele teve um resultado um pouco abaixo do Naive Bayes para a classe 1 e classe 2. Já o Naive Bayes, por conseguir classificar superfícies mais complexas, classificou melhor os exemplos das mesmas classes. A classe 3, como já está mais linearmente separada da classe 1 e classe 2 (Figura 1), foi bem classificada tanto no Naive Bayes como no DQG.

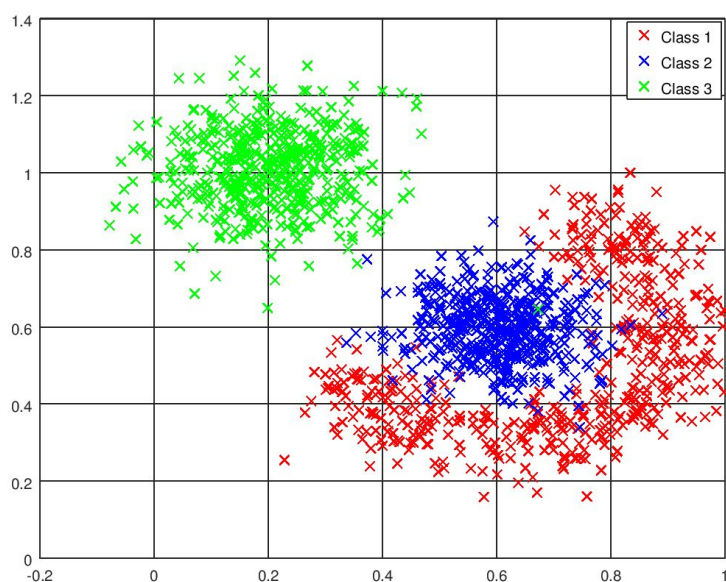


Figura 1 - Dados de exemplos e suas classes