

Lista 4

Classificadores Estatísticos

Pedro Henrique G. Carvalho (362982)

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 ex4data1.data.

O arquivo contém uma matriz de dados. Esta matriz é composta de 150 linhas e 5 colunas. As 4 primeiras colunas representam 4 atributos e a coluna 5 representa a classe a qual pertence o exemplo. Nestes dados, existem 3 classes, sendo 50 exemplos de cada classe.

Os dados pertencem a um problema de reconhecimento flores (íris dataset). Os 4 atributos são tamanho e espessura da sépala e da pétala de cada flor. As três classes referem-se as flores 1-setosa, 2-versicolor e 3-virginica.

- Divida o conjunto de dados entre treino e teste. Para este problema, utilize 90 (30 de cada classe) dados de cada classe para treino, 60 (20 de cada classe) para teste.

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

Apresentar: As matrizes de confusão para os dois classificadores.

- 1) Para o Naive Bayes, obtive:

20	0	0
0	20	0
0	1	19

2) Para o Discriminante Quadrático Gaussiano, obtive:

20	0	0
0	19	1
0	1	19

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

=> Note que, para o Naive Bayes, a taxa de erro foi de 1,67% e, para o Discriminante quadrático gaussiano, 3,33%. Se os conjuntos não fossem linearmente separáveis, o algoritmo iria errar bem mais por não se adequar suficiente às propriedades de cada classe. Porém, como a classificação obteve altas taxas de acerto, podemos crer que as classes são linearmente separáveis.