Atividade 3

Classificadores

Universidade Federal do Ceará

Aprendizado de Máquina

Prof. Victor Farias

V1.1

Entrega: 26/03/2020

Entrega Código + Relatório via Sigaa

**Peso 2**

**Conjunto de dados**

1. Usaremos o conjunto de dados **Iris** (arquivo em anexo).
2. Esse conjunto de dados contém 5 colunas. A 4 primeiras colunas são atributos das instâncias e a última coluna é a variável alvo discreta
3. Embaralhe o dado
4. Separe em 60% para treino e 40% para teste

**Implementação**

1. Implemente o Naive Bayes Gaussiano e o Discriminante Quadrático Gaussiano
2. Para a regressão logística, podemos usar o scikit-learn (<https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html>
3. Implemente uma função que receba um classificador e o conjunto de testes e exiba uma matriz de confusão:  
   assinatura: plot\_confusion\_matrix(X, y, clf)  
   Código exemplo: <https://scikit-learn.org/stable/auto_examples/model_selection/plot_confusion_matrix.html> (sem normalização)
4. Implemente uma função que receba o classificado e o conjunto de teste e exiba o conjunto de testes em um gráfico de dispersão juntamente com as fronteiras de separação.  
   assinatura: plot\_boundaries(X, y, clf)  
   Código exemplo: <https://scikit-learn.org/stable/auto_examples/linear_model/plot_iris_logistic.html>

**Relatório**

1. Para cada um das técnicas: Regressão Linear, Naive Nayes Gaussiano e o Discriminante Quadrático Gaussiano apresente:
   1. A porcentagem de predições corretas para o conjunto de treino
   2. A matriz de confusão
   3. O dado em um gráfico de dispersão com as fronteiras de separação. Nesse item, usar somente os dois primeiros atributos para treinar e gerar esse gráfico.
2. Quais dos métodos produziu fronteiras lineares?
3. O dado parece se linearmente separável ou não?