

Nome: _____ Matrícula: _____
Data: __/__/__

EXERCÍCIO PRÁTICO REDES NEURAIS SUPERVISIONADAS - CLASSIFICAÇÃO

Considere a seguinte situação:

- A fraude de vinhos é uma situação muito real. Pretende-se com um algoritmo de *Machine Learning* de Redes Neurais, classificar automaticamente o **produtor** de um vinho com base em dados de 13 características químicas de cada vinho.
- Temos uma base de dados de 178 vinhos diferentes divididas em 3 **produtores** diferentes. Esses dados são os resultados de uma análise química de vinhos cultivados na mesma região da Itália, mas derivados de três **produtores** diferentes.

produtor	álcool	ácido málico	...
1	14.23	1.71	...

Colunas do *dataset*:

- Produtor
- Álcool
- ácido málico
- cinza
- Alcalinidade das cinzas
- Magnésio
- Fenóis totais
- Flavonóides
- Fenóis não flavonóides
- Proantocianinas
- Intensidade da cor
- Matiz
- OD315 de vinhos diluídos
- Prolina

Os dados para treino e testes do modelo encontram-se no arquivo CSV: **wine.csv**

EXERCÍCIO:

1) Crie um programa em Python que leia os dados do arquivo csv e utilize um algoritmo de Rede Neural Perceptron de Múltiplas Camadas para efetuar futuras previsões de possível produtor do vinho com base nas 13 características químicas. Apresente um menu com 3 opções:

```
####REDE NEURAL PARA PREVISÃO DE PRODUTORES DE VINHO####  
1 - Exibir dados originais e estatísticas  
2 - Imprimir gráfico de Teor Alcoólico / Magnésio / Fenóis por Produtor  
3 - Exibir previsão da Rede Neural
```

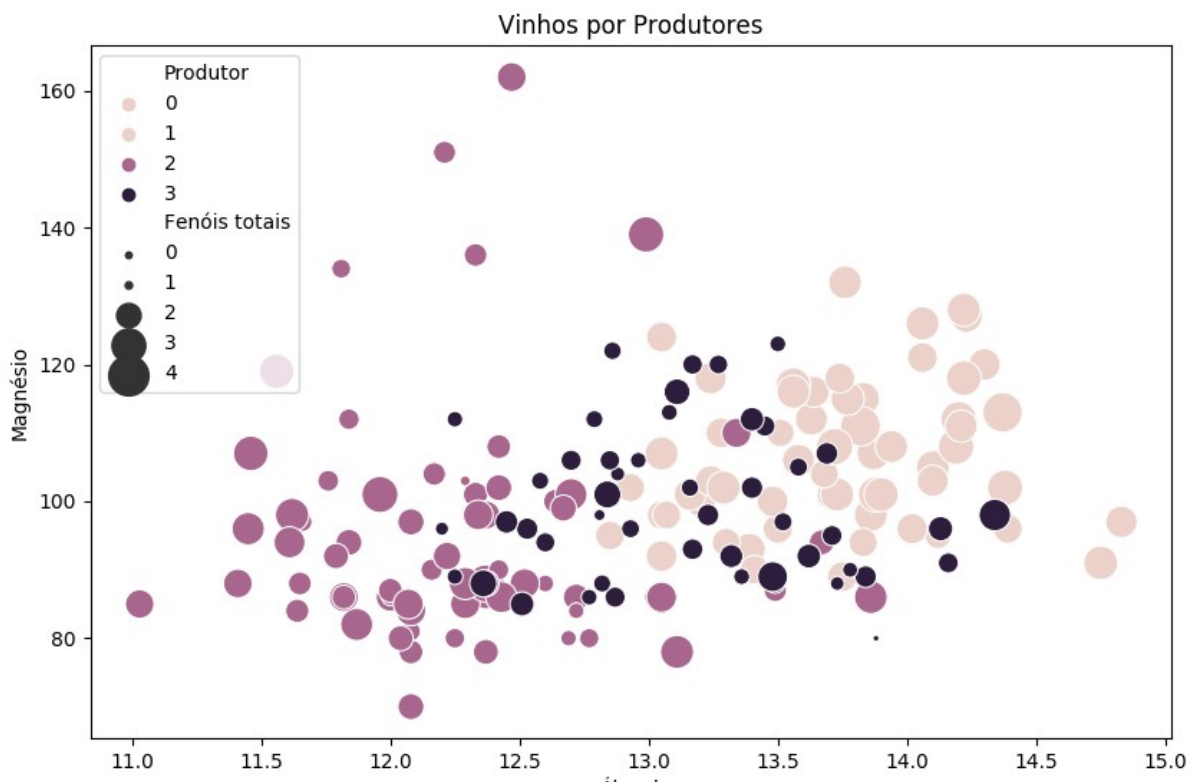
a) 1 – Exibir dados originais e estatísticos

- Esta opção deve exibir o formato do arquivo de dados. 5 primeiras linhas e colunas.
- Deve ser exibida quantidade de linhas e colunas
- Para cada coluna, exibir dados estáticos: valor médio, mediana, mínimo, máximo, desvio padrão, etc.

b) 2 – Imprimir gráfico de teor alcoólico / magnésio / fenóis - por produtor

- **Aqui temos um gráfico para representar 4 dimensões de dados.**
- Esta opção deve imprimir um gráfico de dispersão para os dados acima para uma boa visualização dos dados.
- Eixo X: Teor Alcoólico, Eixo Y: Magnésio.
- Cada produtor deverá ser identificado por uma cor diferente no gráfico.
- A quantidade de fenóis deverá ser identificada pelo tamanho do ponto no gráfico (quanto maior o ponto, maior a quantidade do ponto e vice-versa).
- Quais conclusões se chega com a representação visual?

Gráfico de Saída:



c) 3 – Exibir previsão da Rede Neural

- Efetue treino da rede com 75% dos dados
- Teste com 25% dos dados.
- Deixe os dados de treino e testes balanceados para não viciar o aprendizado. Fixe um *random_state* para início do *split*.
- Efetue vários aprendizados da rede e verifique qual a melhor opção alterando as possibilidades:
 - Número de camadas ocultas e neurônios por camada
 - Função de Ativação
 - Tipo de otimização dos pesos

Exiba:

- Dados reais de teste: Quantidade de dados para cada produtor.
- Resultado da previsão da rede: Previsão da rede para cada produtor.
- Taxa de Acerto da rede.
- Matriz de confusão.