

# Inteligência Computacional

Projeto de redes Neurais

Erich Daniel Morais

# Dataset escolhido

- Wine Quality Data Set – Disponível no repositório UCI machine Learning
  - <https://archive.ics.uci.edu/ml/datasets/Wine+Quality>
- No repositório é possível encontrar dois datasets, um relacionado a vinhos brancos e outros a vinhos vermelhos, o escolhido foi o dataset [“winequality-red.csv”](#)

Número de Atributos	11 + Atributo Classificador
Atributo Classificador	score entre 0 e 10 (10 Classes)
Instâncias	1599



## 7. Attribute information:

For more information, read [Cortez et al., 2009].

Input variables (based on physicochemical tests):

- 1 - fixed acidity
- 2 - volatile acidity
- 3 - citric acid
- 4 - residual sugar
- 5 - chlorides
- 6 - free sulfur dioxide
- 7 - total sulfur dioxide
- 8 - density
- 9 - pH
- 10 - sulphates
- 11 - alcohol

Output variable (based on sensory data):

- 12 - quality (score between 0 and 10)

# Carga do dataset

- Foram utilizados todos os atributos disponíveis no dataset

	fixed acidity	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH	sulphates	alcohol	quality
0	7.4	0.70	0.00	1.9	0.076	11.0	34.0	0.9978	3.51	0.56	9.4	5
1	7.8	0.88	0.00	2.6	0.098	25.0	67.0	0.9968	3.20	0.68	9.8	5
2	7.8	0.76	0.04	2.3	0.092	15.0	54.0	0.9970	3.26	0.65	9.8	5
3	11.2	0.28	0.56	1.9	0.075	17.0	60.0	0.9980	3.16	0.58	9.8	6
4	7.4	0.70	0.00	1.9	0.076	11.0	34.0	0.9978	3.51	0.56	9.4	5
5	7.4	0.66	0.00	1.8	0.075	13.0	40.0	0.9978	3.51	0.56	9.4	5
6	7.9	0.60	0.06	1.6	0.069	15.0	59.0	0.9964	3.30	0.46	9.4	5
7	7.3	0.65	0.00	1.2	0.065	15.0	21.0	0.9946	3.39	0.47	10.0	7
8	7.8	0.58	0.02	2.0	0.073	9.0	18.0	0.9968	3.36	0.57	9.5	7
9	7.5	0.50	0.36	6.1	0.071	17.0	102.0	0.9978	3.35	0.80	10.5	5
10	6.7	0.58	0.08	1.8	0.097	15.0	65.0	0.9959	3.28	0.54	9.2	5
11	7.5	0.50	0.36	6.1	0.071	17.0	102.0	0.9978	3.35	0.80	10.5	5

# Testes visando a melhor configuração

- Todos os testes foram realizados pelo método Cross-Validation com 10 partições e seguindo o exemplo disponibilizado.
- **TESTE 1 – NÚMERO DE CAMADAS DA REDE**
  - Foi definido um número fixo de 300 neurônios por camada
  - 2000 iterações
  - Momentun default : 0.9
  - Demais características: Default

```
#Armazena as predições
from sklearn.metrics import accuracy_score
from sklearn.neural_network import MLPClassifier
from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report
from sklearn.model_selection import train_test_split
from imblearn.over_sampling import SMOTE

mlp = MLPClassifier(hidden_layer_sizes=(300, 300), max_iter=2000, momentum=0.9)

# podemos alterar o activation, layers, iterações e o momentum
mlp.fit(X_train, y_train.values.ravel())

predictions = cross_val_predict(mlp,X_resampled,y_resampled,cv=10)
#Calcula acurácia do treino
print("acurácia do treino")
print(accuracy_score(y_resampled,predictions)*100)

#Gera a matriz de confusão do treino
print("matriz de confusão")
print(confusion_matrix(y_resampled,predictions))
```

# Testes visando a melhor configuração

## 2 Camadas

```
acurácia do treino
75.23250122369065
matriz de confusão
[[675  6  0  0  0  0]
 [ 11 613 35 17  5  0]
 [ 16  94 341 173 43 14]
 [ 15  49 215 241 125 36]
 [  2  5  21  71 557 25]
 [  0  2  1  9 22 647]]
Report
```

	precision	recall	f1-score	support
3	0.94	0.99	0.96	681
4	0.80	0.90	0.85	681
5	0.56	0.50	0.53	681
6	0.47	0.35	0.40	681
7	0.74	0.82	0.78	681
8	0.90	0.95	0.92	681
accuracy			0.75	4086
macro avg	0.73	0.75	0.74	4086
weighted avg	0.73	0.75	0.74	4086

## 4 Camadas

```
acurácia do treino
75.72197748409202
matriz de confusão
[[670  9  0  1  1  0]
 [ 21 589 39 27  4  1]
 [ 21  59 339 202 43 17]
 [  8  31 223 277 113 29]
 [  1  2  24  52 575 27]
 [  0  0  6 12 19 644]]
Report
```

	precision	recall	f1-score	support
3	0.93	0.98	0.96	681
4	0.85	0.86	0.86	681
5	0.54	0.50	0.52	681
6	0.49	0.41	0.44	681
7	0.76	0.84	0.80	681
8	0.90	0.95	0.92	681
accuracy			0.76	4086
macro avg	0.74	0.76	0.75	4086
weighted avg	0.74	0.76	0.75	4086

## 3 Camadas

```
acurácia do treino
76.89672050905531
matriz de confusão
[[671  9  1  0  0  0]
 [ 11 616 33 16  3  2]
 [ 16  73 331 215 32 14]
 [ 11  38 186 287 127 32]
 [  1  7  15  42 597 19]
 [  0  0  6 13 22 640]]
Report
```

	precision	recall	f1-score	support
3	0.95	0.99	0.96	681
4	0.83	0.90	0.87	681
5	0.58	0.49	0.53	681
6	0.50	0.42	0.46	681
7	0.76	0.88	0.82	681
8	0.91	0.94	0.92	681
accuracy			0.77	4086
macro avg	0.75	0.77	0.76	4086
weighted avg	0.75	0.77	0.76	4086

## 5 Camadas

```
acurácia do treino
75.64855604503181
matriz de confusão
[[658 10  8  3  2  0]
 [ 10 613 34 19  4  1]
 [ 22  71 312 213 46 17]
 [  7  32 238 274  90 40]
 [  0  5  17  54 584 21]
 [  0  0  4 15 12 650]]
Report
```

	precision	recall	f1-score	support
3	0.94	0.97	0.96	681
4	0.84	0.90	0.87	681
5	0.51	0.46	0.48	681
6	0.47	0.40	0.44	681
7	0.79	0.86	0.82	681
8	0.89	0.95	0.92	681
accuracy			0.76	4086
macro avg	0.74	0.76	0.75	4086
weighted avg	0.74	0.76	0.75	4086

# Testes visando a melhor configuração

```
acurácia do treino
76.89672050905531
matriz de confusão
[[671  9  1  0  0  0]
 [ 11 616 33 16  3  2]
 [ 16 73 331 215 32 14]
 [ 11 38 186 287 127 32]
 [  1  7 15  42 597 19]
 [  0  0  6 13  22 640]]
Report
```

	precision	recall	f1-score	support
3	0.95	0.99	0.96	681
4	0.83	0.90	0.87	681
5	0.58	0.49	0.53	681
6	0.50	0.42	0.46	681
7	0.76	0.88	0.82	681
8	0.91	0.94	0.92	681
accuracy			0.77	4086
macro avg	0.75	0.77	0.76	4086
weighted avg	0.75	0.77	0.76	4086

Resultado: De acordo com as métricas o melhor número de camadas é 3

# Testes visando a melhor configuração

- Todos os testes foram realizados pelo método Cross-Validation com 10 partições e seguindo o exemplo disponibilizado.
- **TESTE 2 – NÚMERO DE NEURONIOS POR CAMADAS DA REDE**
  - Foi definido um número fixo de 3 camadas
  - 2000 iterações
  - Momentun Default : 0.9
  - Demais características: Default

# Testes visando a melhor configuração

250 Neurônios em  
todas as camadas

```
acurácia do treino
77.58198727361723
matriz de confusão
[[680  0  1  0  0  0]
 [ 10 616 34 18  3  0]
 [ 21  60 337 212 37 14]
 [ 16  29 207 300 94 35]
 [  1  3  16  59 582 20]
 [  0  0  4  11 11 655]]
Report
```

	precision	recall	f1-score	support
3	0.93	1.00	0.97	681
4	0.87	0.90	0.89	681
5	0.56	0.49	0.53	681
6	0.50	0.44	0.47	681
7	0.80	0.85	0.83	681
8	0.90	0.96	0.93	681
accuracy			0.78	4086
macro avg	0.76	0.78	0.77	4086
weighted avg	0.76	0.78	0.77	4086

300 Neurônios em  
todas as camadas

```
acurácia do treino
76.89672050905531
matriz de confusão
[[671  9  1  0  0  0]
 [ 11 616 33 16  3  2]
 [ 16  73 331 215 32 14]
 [ 11  38 186 287 127 32]
 [  1  7  15  42 597 19]
 [  0  0  6  13 22 640]]
Report
```

	precision	recall	f1-score	support
3	0.95	0.99	0.96	681
4	0.83	0.90	0.87	681
5	0.58	0.49	0.53	681
6	0.50	0.42	0.46	681
7	0.76	0.88	0.82	681
8	0.91	0.94	0.92	681
accuracy			0.77	4086
macro avg	0.75	0.77	0.76	4086
weighted avg	0.75	0.77	0.76	4086

400 Neurônios em  
todas as camadas

```
acurácia do treino
76.70093000489476
matriz de confusão
[[674  3  2  2  0  0]
 [ 10 618 36 13  2  2]
 [ 18  68 302 230 50 13]
 [ 15  29 193 292 105 47]
 [  1  7  12  48 596 17]
 [  0  0  5  7 17 652]]
Report
```

	precision	recall	f1-score	support
3	0.94	0.99	0.96	681
4	0.85	0.91	0.88	681
5	0.55	0.44	0.49	681
6	0.49	0.43	0.46	681
7	0.77	0.88	0.82	681
8	0.89	0.96	0.92	681
accuracy			0.77	4086
macro avg	0.75	0.77	0.76	4086
weighted avg	0.75	0.77	0.76	4086



# Testes visando a melhor configuração

300 nas camadas 1 e 3,  
350 na camada  
intermediaria

```
acurácia do treino
76.65198237885463
matriz de confusão
[[669 10 0 2 0 0]
 [ 13 626 22 18 2 0]
 [ 17 91 292 218 42 21]
 [ 12 30 190 320 97 32]
 [ 4 5 14 60 571 27]
 [ 0 0 1 7 19 654]]
Report
```

	precision	recall	f1-score	support
3	0.94	0.98	0.96	681
4	0.82	0.92	0.87	681
5	0.56	0.43	0.49	681
6	0.51	0.47	0.49	681
7	0.78	0.84	0.81	681
8	0.89	0.96	0.92	681
accuracy			0.77	4086
macro avg	0.75	0.77	0.76	4086
weighted avg	0.75	0.77	0.76	4086

250 nas camadas 1 e 3,  
350 na camada  
intermediaria

```
acurácia do treino
77.11698482623592
matriz de confusão
[[673 6 1 1 0 0]
 [ 6 640 19 14 1 1]
 [ 20 89 312 200 48 12]
 [ 8 40 189 290 117 37]
 [ 0 5 16 55 582 23]
 [ 0 0 4 9 14 654]]
Report
```

	precision	recall	f1-score	support
3	0.95	0.99	0.97	681
4	0.82	0.94	0.88	681
5	0.58	0.46	0.51	681
6	0.51	0.43	0.46	681
7	0.76	0.85	0.81	681
8	0.90	0.96	0.93	681
accuracy			0.77	4086
macro avg	0.75	0.77	0.76	4086
weighted avg	0.75	0.77	0.76	4086

250 nas camadas 1 e 3,  
400 na camada  
intermediaria

```
acurácia do treino
77.75330396475772
matriz de confusão
[[672 8 0 1 0 0]
 [ 6 637 24 13 1 0]
 [ 13 75 325 216 38 14]
 [ 14 33 191 304 112 27]
 [ 1 9 13 42 592 24]
 [ 0 0 2 15 17 647]]
Report
```

	precision	recall	f1-score	support
3	0.95	0.99	0.97	681
4	0.84	0.94	0.88	681
5	0.59	0.48	0.53	681
6	0.51	0.45	0.48	681
7	0.78	0.87	0.82	681
8	0.91	0.95	0.93	681
accuracy			0.78	4086
macro avg	0.76	0.78	0.77	4086
weighted avg	0.76	0.78	0.77	4086

Resultado: De acordo com as métricas o melhor número de neurônios é 250 nas camadas 1 e 3, 400 na camada intermediaria

# Testes visando a melhor configuração

TESTE 3 – Foi definido um número fixo 250 nas camadas 1 e 3, 400 na camada intermediária

- 2000 iterações
- Demais características: Default

## Ativação: Relu

```
acurácia do treino
77.75330396475772
matriz de confusão
[[672  8  0  1  0  0]
 [ 6 637 24 13  1  0]
 [13 75 325 216 38 14]
 [14 33 191 304 112 27]
 [ 1  9  13  42 592 24]
 [ 0  0  2  15 17 647]]
Report
```

	precision	recall	f1-score	support
3	0.95	0.99	0.97	681
4	0.84	0.94	0.88	681
5	0.59	0.48	0.53	681
6	0.51	0.45	0.48	681
7	0.78	0.87	0.82	681
8	0.91	0.95	0.93	681
accuracy			0.78	4086
macro avg	0.76	0.78	0.77	4086
weighted avg	0.76	0.78	0.77	4086

## Ativação: logistic

```
acurácia do treino
74.0332843857073
matriz de confusão
[[655 21  3  1  1  0]
 [21 594 32 21  8  5]
 [23 93 296 210 41 18]
 [10 40 184 282 109 56]
 [ 2  9  11  72 540 47]
 [ 0  0  2  4 17 658]]
Report
```

	precision	recall	f1-score	support
3	0.92	0.96	0.94	681
4	0.78	0.87	0.83	681
5	0.56	0.43	0.49	681
6	0.48	0.41	0.44	681
7	0.75	0.79	0.77	681
8	0.84	0.97	0.90	681
accuracy			0.74	4086
macro avg	0.72	0.74	0.73	4086
weighted avg	0.72	0.74	0.73	4086

## Ativação: Tanh

```
acurácia do treino
78.87909936368087
matriz de confusão
[[678  3  0  0  0  0]
 [ 6 640 19 11  3  2]
 [19 73 312 222 43 12]
 [16 25 205 303 93 39]
 [ 0  2  7  28 631 13]
 [ 0  0  6  7  9 659]]
Report
```

	precision	recall	f1-score	support
3	0.94	1.00	0.97	681
4	0.86	0.94	0.90	681
5	0.57	0.46	0.51	681
6	0.53	0.44	0.48	681
7	0.81	0.93	0.86	681
8	0.91	0.97	0.94	681
accuracy			0.79	4086
macro avg	0.77	0.79	0.78	4086
weighted avg	0.77	0.79	0.78	4086

Resultado: De acordo com as métricas melhor método é a tanh

# Testes visando a melhor configuração

- **TESTE 4** – Foi definido um número fixo de 10 neurônios por camada em 7 camadas
  - 2000 iterações
  - Demais características: Default

## Momentun 0.7

```
acurácia do treino
79.2462065589819
matriz de confusão
[[672  5  0  3  1  0]
 [ 2 647 15 16  1  0]
 [18 80 314 218 40 11]
 [15 36 193 318 94 25]
 [ 1  4  15  29 620 12]
 [ 0  0  4  2  8 667]]
```

Report		precision	recall	f1-score	support
	3	0.95	0.99	0.97	681
	4	0.84	0.95	0.89	681
	5	0.58	0.46	0.51	681
	6	0.54	0.47	0.50	681
	7	0.81	0.91	0.86	681
	8	0.93	0.98	0.96	681
	accuracy			0.79	4086
	macro avg	0.78	0.79	0.78	4086
	weighted avg	0.78	0.79	0.78	4086

## Momentun 0.8

```
acurácia do treino
79.00146842878121
matriz de confusão
[[678  1  2  0  0  0]
 [ 6 643 17 14  1  0]
 [17 63 321 226 43 11]
 [12 29 201 304 102 33]
 [ 0  4  11  26 623 17]
 [ 0  0  2  9 11 659]]
```

Report		precision	recall	f1-score	support
	3	0.95	1.00	0.97	681
	4	0.87	0.94	0.90	681
	5	0.58	0.47	0.52	681
	6	0.53	0.45	0.48	681
	7	0.80	0.91	0.85	681
	8	0.92	0.97	0.94	681
	accuracy			0.79	4086
	macro avg	0.77	0.79	0.78	4086
	weighted avg	0.77	0.79	0.78	4086

## Momentun 0.9

```
acurácia do treino
78.87909936368087
matriz de confusão
[[678  3  0  0  0  0]
 [ 6 640 19 11  3  2]
 [19 73 312 222 43 12]
 [16 25 205 303 93 39]
 [ 0  2  7  28 631 13]
 [ 0  0  6  7  9 659]]
```

Report		precision	recall	f1-score	support
	3	0.94	1.00	0.97	681
	4	0.86	0.94	0.90	681
	5	0.57	0.46	0.51	681
	6	0.53	0.44	0.48	681
	7	0.81	0.93	0.86	681
	8	0.91	0.97	0.94	681
	accuracy			0.79	4086
	macro avg	0.77	0.79	0.78	4086
	weighted avg	0.77	0.79	0.78	4086

Resultado: De acordo com as métricas melhor  
Momentun é 0.7

# Conclusão

- As melhores configurações de acordo com as métricas são:
  - 250 nas camadas 1 e 3, 400 na camada intermediária
  - **Método de ativação "tanh"**
  - Momentun ideal 0.7

- Com estas configurações é possível obter 79.24% de acurácia, como pode ser visto na matriz de confusão ao lado.

```
acurácia do treino
79.2462065589819
matriz de confusão
[[672  5  0  3  1  0]
 [ 2 647 15 16  1  0]
 [ 18 80 314 218 40 11]
 [ 15 36 193 318 94 25]
 [  1  4  15  29 620 12]
 [  0  0  4  2  8 667]]
Report
```

	precision	recall	f1-score	support
3	0.95	0.99	0.97	681
4	0.84	0.95	0.89	681
5	0.58	0.46	0.51	681
6	0.54	0.47	0.50	681
7	0.81	0.91	0.86	681
8	0.93	0.98	0.96	681
accuracy			0.79	4086
macro avg	0.78	0.79	0.78	4086
weighted avg	0.78	0.79	0.78	4086