

Trabalho 3

**Data de Entrega: 24/out/2018**

1) Dado o dataset artificial em base\_trabalho3.txt, mostrado na figura abaixo, escolha e compare três métodos de classificação usando o weka.

a) Insira o cabeçalho no arquivo base\_trabalho3.txt convertendo o arquivo em .arff - Attribute-Relation File Format (ARFF).

```
@RELATION teste2
```

```
@ATTRIBUTE value1 NUMERIC
@ATTRIBUTE value2 NUMERIC
@ATTRIBUTE value3 NUMERIC
@ATTRIBUTE class {1,2,3}
```

```
@data
```

b) Mostre a matriz de confusão de cada método.

lbk

```
=== Confusion Matrix ===
```

```
  a   b   c  <-- classified as
1987   8   5 |   a = 1
   8 1992   0 |   b = 2
   5   0 1995 |   c = 3
```

MultilayerPerceptron

```
=== Confusion Matrix ===
```

```
  a   b   c  <-- classified as
1990   6   4 |   a = 1
   6 1994   0 |   b = 2
   6   0 1994 |   c = 3
```

Logistic

```
=== Confusion Matrix ===
```

```
  a   b   c  <-- classified as
1991   4   5 |   a = 1
   5 1995   0 |   b = 2
   5   0 1995 |   c = 3
```

## Simple Logistic

=== Confusion Matrix ===

```
a  b  c  <-- classified as
1993  3  4 |  a = 1
 6 1994  0 |  b = 2
 5   0 1995 |  c = 3
```

c) Faça uma tabela comparativa informando o método e a acurácia obtida.

Método	Parâmetros	Acurácia
IBK	KNN-3	99.5667 %
MultilayerPerceptron	5-node sigmoid	99.6333 %
Logistic		99.6833 %
Simple Logistic		99.7 %

d) Faça uma apresentação sobre cada método (sugestão, usar o help do weka, e outras fontes da internet).

**IBK:** KNN é um método de aprendizagem na qual usa todo o banco para classificar, usando um método de votação usando os N mais pertos para classificar o que você quer....

**MultilayerPerceptron:** Usa uma rede neural com vários neurônios ligados entre si, com camada de entrada, saída e intermediárias....

Cada neurônio tem entradas que são mescladas com os pesos + bias e gerando uma saída e ele passando em uma função de ativação como as sigmoid

Fórmula do neurônio:

$\text{output} = \text{sig}(w_1 * X_1 + w_2 * X_2 + \dots + B)$

**Logística:** Usa gradiente descendente aplicado a uma fórmula, para resolver problemas de classificação, sendo o mais clássico:

$$\hat{y} = \frac{1}{1 + e^{-\theta X}}$$

e) Monte a mesma tabela do experimento anterior, desta vez removendo a coluna X2 na aba de pré-processamento do weka.

Método	Parâmetros	Acurácia
IBK	KNN-3	88.5333 %
MultilayerPerceptron	5-node sigmoid	90.3333 %
Logistic		90.7833 %
Simple Logistic		90.7833 %

