

1. Sendo  $x$ , uma lista de inteiros, que tipo de objeto  $y$  é gerado pelo código Python  $y=(i/2 \text{ for } i \text{ in } x \text{ if } i\%2=1)$ ?
  - a. objeto gerador
  - b. função Lambda
  - c. lista
  - d. tuplo
2. Modelos de regressão logística podem ser usados quer em classificação quer em regressão
  - a. apenas em problemas de classificação
  - b. apenas em problemas de regressão
  - c. apenas em problemas de classificação binária
  - d. quer em classificação quer em regressão
3. Qual das seguintes qualidades NÃO é correto associar à linguagem Python?
  - a. baixo nível
  - b. interpretada
  - c. orientada a objetos
  - d. case-sensitive (sensível a maiúsculas e minúsculas)
4. Alguns dos algoritmos de ML não conseguem lidar com variáveis preditivas do tipo categórico
  - a. Esse não é o caso das SVM (máquinas de vetores de suporte)
  - b. nenhuma está correta
  - c. É precisamente o que acontece com as SVM
  - d. Ainda que as SVM não tenham, pela sua natureza, essa limitação, ela está presente nas implementações do tipo de algoritmo disponibilizadas pelo Scikit-learn
5. Em Python, um set é um conjunto
  - a. não ordenado de objetos que poderão estar repetidos
  - b. não ordenado de objetos não repetidos
  - c. ordenado de objetos que poderão estar repetidos
  - d. ordenado de objetos não repetidos
6. Dois dos parâmetros de configuração das SVM (máquinas de vetores de suporte) são os seguintes:
  - a. Termo de regularização. Função de Kernel
  - b. Número de árvores. Função de ativação
  - c. Termo de regularização. Função de ativação
  - d. Taxa de aprendizagem. Função de ativação
7. Em Python, as diferentes formas de inicialização das instâncias de uma classe poderão ser asseguradas usando
  - a. vários métodos construtores (ou inicializadores) com parâmetros opcionais
  - b. o construtor (ou inicializador) por defeito
  - c. parâmetros opcionais no método construtor (ou inicializador)
  - d. vários métodos construtores (ou inicializadores)
8. Em que tipo de variáveis categóricas existe uma relação de ordem entre os seus valores?
  - a. apenas nas nominais
  - b. em todas as variáveis categóricas
  - c. apenas nos ordinais
  - d. em nenhuma das variáveis categóricas

9. Qual dos seguintes packages do Python se revela de maior utilidade na visualização gráfica dos dados?
- Numpy
  - Pandas
  - Seaborn
  - Scikit-learn
10. O valor Micro AUC usa-se na avaliação de modelos
- de classificação binária
  - de classificação binária e multi-classe
  - de classificação multi-classe
  - de regressão
11. Na avaliação dum modelo classificador, a Especialidade representa
- proporção dos exemplos positivos que foram corretamente classificados pelo modelo
  - a taxa de acerto desse classificador entre os exemplos que classificou como positivos
  - a proporção dos exemplos negativos que foram corretamente classificados pelo modelo
  - a proporção de exemplos que foram corretamente classificados pelo modelo
12. O algoritmo K-Means tem por base uma técnica de aprendizagem
- por reforço
  - supervisionada
  - nenhuma das opções está correta
  - não supervisionada
13. Na classificação multi-label, cada exemplo do conjunto de dados pode assumir
- um valor contínuo
  - apenas uma entre várias categorias
  - diferentes categorias em simultâneo
  - nenhuma das opções está correta
14. Após a atribuição `x = ['vermelho', 'verde', 'azul', 'branco', 'preto']` a operação de 'slicing' `x[-3:3]` produz o seguinte resultado:
- `[]`
  - `['vermelho', 'verde', 'azul']`
  - `['azul']`
  - `['azul', 'branco']`
15. O algoritmo KNN (vizinhos mais próximos) pode ser usado
- apenas em problemas de regressão
  - quer em classificação quer em regressão
  - apenas em problemas de classificação
  - apenas em problemas de classificação binária
16. Com a instrução Python `x = [10, 20, 30]`, a variável `x` passa a referenciar uma estrutura de dados do tipo
- set (conjunto)
  - lista
  - tuplo
  - dicionário

17.	Classe	+	19	1
	Verdadeira	-	80	900

- a. boa sensibilidade, mas fraca precisão
  - b. boa precisão, mas fraca sensibilidade
  - c. fraca precisão, e fraca sensibilidade
  - d. boa precisão e boa sensibilidade
18. Alguns dos parâmetros de configuração das redes neuronais artificiais são os seguintes:
- a. nº de camadas escondidas, termo de regularização, função de kernel
  - b. taxa de aprendizagem, nº camadas escondidas, função de ativação
  - c. nº camadas escondidas, função de kernel, função de ativação
  - d. nº árvores, taxa de aprendizagem, nº camadas escondidas
19. Num modelo de classificação binária, designam-se por falsos positivos
- a. os exemplos da classe negativa, classificados pelo modelo como pertencentes à classe positiva
  - b. os exemplos da classe positiva, classificados pelo modelo como pertencentes à classe positiva
  - c. os exemplos da classe negativa, classificados pelo modelo como pertencentes à classe negativa
  - d. os exemplos da classe positiva, classificados pelo modelo como pertencentes à classe negativa
20. Uma variável que assuma valores do conjunto {'vermelho', 'verde', 'azul'}, por quantas variáveis dum ()?
- a. 1
  - b. 3
  - c. 2
  - d. 4
21. Em qual das seguintes fases da KDD se aplicam aos dados os ... algoritmos de ML?
- a. Seleção de dados
  - b. Transformação de dados
  - c. Prospeção de dados (Data Mining)
  - d. Pré-Processamento
22. Qual dos seguintes do Python disponibiliza a estrutura Data Frame?
- a. Pandas
  - b. Numpy
  - c. Seaborn
  - d. Scikit-learn
23. Após a atribuição `x = numpy.array([[10, 20, 30], [-30, -20, -10]])`, qual o valor devolvido por `x.argmin()`?
- a. 10
  - b. 3
  - c. -30
  - d. 4
24. Após a atribuição `x = numpy.array([[10, 20, 30], [-40, -50, -60]])`, qual o valor devolvido por `x.argmax()`?
- a. 30
  - b. -60
  - c. 3
  - d. 2

25. Na aprendizagem não supervisionada, o método de clustering (ou de agrupamento) tenta, garante a:
- heterogeneidade dentro de cada cluster e homogeneidade entre diferentes clusters
  - similaridade entre elementos de diferentes clusters e diferenciação entre elementos de um mesmo cluster
  - heterogeneidade entre todos os elementos
  - homogeneidade dentro de cada cluster e heterogeneidade entre diferentes clusters
26. Para que serve o método `fit()` do modelo gerado por um dos algoritmos de ML do Scikit-learn?
- para criar o modelo
  - para treinar o modelo
  - para afinar o modelo
  - para testar o modelo
27. Qual das seguintes métricas é usada em problemas de regressão?
- F-medida
  - Sensibilidade (*recall*)
  - Valor AUC da curva ROC
  - Coeficiente de determinação ( $R^2$ )
28. Qual das seguintes métricas é utilizada nos problemas de regressão?
- Sensibilidade
  - F-medida
  - AUC (Área sob a Curva ROC)
  - Erro médio ao quadrado (MSE)
29. Na construção de métricas multiclasse, que tipos de média são utilizados?
- Nenhuma das opções estão corretas
  - macro, micro ou *highted*
  - macro ou micro
  - macro, micro ou *weighted*
30. Qual do seguinte hiper parâmetro não faz parte da rede neuronal
- taxa de aprendizagem
  - números de neurónios em cada uma das camadas da rede
  - número de camadas ocultas
  - número de árvores
31. Qual das seguintes estruturas de dados NÃO é apresentada na linguagem Python?
- dicionário
  - tuplo
  - lista
  - stack
32. Qual é o método que nos permite converter uma *String* em minúsculas em Python?
- `lower()`
  - `upper()`
  - `ToLower()`
  - `low()`

33. Na avaliação de um classificador binário, a Taxa de Falsos Negativos representa a ...
- taxa de erro na classe negativa. É uma medida da proporção dos exemplos de classe negativos classificados incorretamente pelo preditor
  - taxa de erro na classe positiva. É uma medida da proporção dos exemplos de classe positivos classificados incorretamente pelo preditor
  - proporção de exemplos negativos que foram corretamente classificados pelo modelo
  - proporção de exemplos que foram corretamente classificados pelo modelo
34. Alguns dos algoritmos ML não conseguem lidar com variáveis preditivas de tipo categórico
- Esse não é o caso das Random Forests
  - é precisamente o que acontece com as Random Forests
  - nenhuma das opções está correta
  - ainda que as Random Forests não tenham, pela sua natureza, essa limitação, ela está presente nas implementações desse tipo de algoritmo disponibilizadas pelo Scikit-learn
35. Na avaliação de um classificador binário, Verdadeiros Positivos representa o número
- de previsões negativas que estejam corretas
  - de previsões negativas que são incorretas
  - de previsões positivas que sejam corretas
  - de previsões positivas que são incorretas
36. Na ML, quando os algoritmos aprendem a partir do feedback que recebem sobre a qualidade das soluções que vão produzindo, estamos perante a aprendizagem
- nenhuma está correta
  - não supervisionada
  - supervisionada
  - por reforço
37. Em Python, os Métodos são, por defeito
- virtual
  - abstrato
  - não polimórfico
  - estático
38. Que funções são utilizadas para inicializar um objeto?
- função definida pelo utilizador
  - construtor
  - função de geração
  - função sem parâmetros
39. Na classificação multiclasse, cada exemplo do conjunto de dados pode assumir
- uma categoria entre duas categorias possíveis <-
  - categorias diferentes ao mesmo tempo
  - um valor contínuo
  - nenhuma das opções está correta

40. Em Python o processo de *pickling* inclui:
- a. a conversão de uma tabela de objetos numa lista de dados
  - b. conversão de uma lista de objetos numa tabela de dados
  - c. conversão de uma tabela de dados numa lista de objetos
  - d. conversão de uma hierarquia de objetos python em um fluxo de bytes
41. Ao avaliar um classificador binário, a métrica F1:
- a. é a média linear da precisão e sensibilidade
  - b. é a média harmónica ponderada de precisão e sensibilidade
  - c. é a média harmónica de precisão e sensibilidade
  - d. é a média ponderada de precisão e sensibilidade
42. Que palavra-chave é usada para definir uma função na Língua Python?
- a. def
  - b. fun
  - c. define
  - d. function
43. Qual dos seguintes índices nos permite obter a matriz `[[30, 32, 34], [50, 52, 54]]`:
- a. `a[2::2,::2]`
  - b. `a[3::5]`
  - c. `a[3::2, ::2]`
  - d. `a[:2]`
44. Qual dos seguintes pontos não faz parte do *GridSearchCV*?
- a. cv
  - b. param\_grid
  - c. estimator
  - d. score
45. Qual dos seguintes aspetos NÃO representa uma vantagem da *Random Forest*?
- a. ser resistente a *overfitting*
  - b. lidar bem com atributos categóricos com um elevado número de classes distintas
  - c. gerar estimativas de enviesamento baixo
  - d. a capacidade de lidar com dados categóricos e contínuos
46. O algoritmo PCA (Principal Component Analysis) faz uso das funções do núcleo?
- a. sim, mas só usa sigmoidal ou polinomial
  - b. não
  - c. sim
  - d. nenhuma das opções está correta
47. Suponha que A é a subclasse de B. Qual das seguintes afirmações é a correta forma de invocar o construtor de B de um objeto de classe A?
- a. `super().init(self)`
  - b. `B.init([args])`
  - c. `super().__init__(self)`
  - d. `B.__init__(self,[args])`

48. Qual dos seguintes conjuntos de valores não deve ser associado a Variáveis Categóricas Nominais?
- nenhuma das opções
  - 'mau', 'suficiente', 'muito bom', 'insuficiente', 'bom'
  - 'vermelho', 'verde', 'azul', 'branco', 'preto'.
  - 'laranja', 'maçã', 'morango', 'kiwi', 'manga'
49. Qual dos seguintes aspetos não pretende reduzir a dimensionalidade?
- decomposição
  - redução por transformação de tipo
  - projeção/extração de características
  - selecção de características
50. Que operador aritmético não pode ser usado com operandos de Strings em Python?
- +
  - 
  - \*
  - todos mencionados
51. Quais das seguintes configurações de arquitetura de Redes Neurais Artificiais NÃO existem?
- recorrente
  - multi hierárquica
  - multicamadas
  - uma camada oculta
52. Qual dos seguintes aspetos não representa uma vantagem das Máquinas Vetoriais de Apoio (SVM)?
- produzir regras de classificação fáceis de interpretar
  - a precisão não depende do tamanho e da dimensionalidade dos dados
  - ter uma boa tolerância ao ruído
  - boa capacidade de generalização
53. Qual das seguintes não faz parte do ciclo de vida da Descoberta do Conhecimento da Base de Dados (KDD)?
- exploração de dados
  - selecção de dados
  - pré-processamento
  - todas as tarefas mencionadas fazem parte do ciclo de vida do kdd
54. Qual das seguintes métricas é utilizada nos problemas de classificação binária?
- erro quadrado médio (MSE)
  - matriz de confusão
  - coeficiente de determinação ( $R^2$ )
  - erro quadrado médio de raiz (RMSE)
55. Podem ser utilizados *K-Nearest Neighbors*
- Apenas em problemas de classificação binária
  - Em classificação ou regressão
  - apenas em problemas de classificação
  - apenas em problemas de regressão

56. Relativamente ao Valor AUC, qual das seguintes afirmações é verdadeira?
- a. esta é a área abaixo da curva roc <-
  - b. produz valores entre -1 e 1
  - c. produz valores entre -100 e 100
  - d. [nenhuma das opções está correcta]
57. O Python é *case-sensitive* quando se trata de identificadores?
- a. depende da máquina
  - b. sim
  - c. não
  - d. nenhum dos mencionados
58. Qual das seguintes técnicas não se aplica à aprendizagem supervisionada?
- a. K-Nearest Neighbors
  - b. Redes Neurais
  - c. K-Means
  - d. Regressão Linear
59. O ciclo de vida da KDD (descoberta de conhecimento em base de dados) compreende uma sequência de 5 fases, realizadas, tendencialmente, pela seguinte ordem:
- a. seleção de dados, prospeção de dados, pré-processamento, transformação dos dados e interpretação dos resultados
  - b. pré-processamento, seleção dos dados, prospeção de dados, transformação de dados e interpretação dos resultados
  - c. seleção dos dados, pré-processamento, transformação dos dados, prospeção dos dados e interpretação dos resultados
  - d. pré-processamento, prospeção dos dados, seleção dos dados, transformação dos dados e interpretação dos resultados



60. Considere a seguinte declaração Python, qual é o resultado da impressão?

```
1 x = {'idade': 21, 'ano': 2022, 'IADone': Verdade}
2 print("{}".format(type(x)))
```

- a. <classe 'dict'>
- b. <classe 'list'>
- c. <classe 'set'>
- d. <classe 'tuple'>

61. Considere a seguinte declaração Python:

```
1 len(['ola'], 2, 4, 6)
```

- a. 3
- b. 4
- c. 6
- d. Error

62. Considere os seguintes problemas de classificação binária do trecho de código, o que será o *stdout* do print(x)?

```
1 x = ['Red', 'Green', 'Blue', 'White', 'Black']
2 x = x[-3:-1]
3 print(x)
```

- a. ['vermelho', 'verde', 'azul']
- b. []
- c. ["azul"].
- d. ["azul", "branco"]

63. Considere a seguinte expressão Python, qual é o valor de x?

```
1 x = 3 + 6 % 7
```

- a. 1
- b. 2
- c. 7
- d. 9

64. Dado o seguinte código, que método da classe *foo* é invocado na última declaração?

```
1 class foo:
2     def __init__(self, weight):
3         self.weight = weight
4 a = foo(50)
5 b = foo(60)
6 print(a < b)
```

- a. `__lt__`
- b. `__eq__`
- c. `__lt__`
- d. `__less__`

65. Qual é a saída do seguinte trecho de código Python?

```
1 my_string = "hello world"
2 k = [(i.upper()), len(i)] for i in my_string]
3 print(k)
```

- a. ['olá mundo'], 11
- b. [('h', 1), ('e', 1), ('l', 1), ('l', 1), ('o', 1), (' ', 1), ('w', 1), ('o', 1), ('r', 1), ('l', 1), ('l', 1), ('d', 1)]
- c. nenhum das opções
- d. [('olá', 5), ('mundo', 5)]

66. Ao avaliar um classificador binário, a métrica F1

```
1 x = 'abcd'
2 for i in range(len(x)):
3 print(i)
```

- a. 1 2 3 4
- b. 0 1 2 3
- c. a b c d
- d. erro

67. Considere o seguinte trecho de código, este gera 15 valores aleatórios entre:

```
1 import numpy as np
2 val = np.random.random((3,5))*10+10
3 Print("{}".format(val))
```

- a. 40 e 60
- b. 30 e 50
- c. 3 e 5
- d. 10 e 20