 <small>ESCOLA SUPERIOR DE TECNOLOGIA E GESTÃO</small>	Tipo de Prova Exame Teórico – Época Normal	Ano letivo 2021/2022	Data 23-06-2022
	Curso Licenciatura em Engenharia Informática	Hora 10:00	
	Unidade Curricular Inteligência Artificial	Duração 2:30 horas	

Observações:

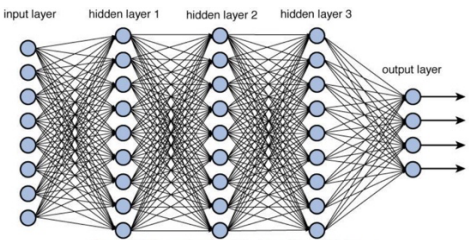
- Pode trocar a ordem das questões, desde que as identifique convenientemente.
- Qualquer tentativa de fraude implica a anulação do exame.
- A Parte 1 deste exame é constituída por questões de escolha múltipla. As mesmas devem ser respondidas na folha de resposta. Cada resposta errada desconta 0.25 valores da Parte 1.
- O enunciado deve ser entregue juntamente com a folha de resposta.


Número: \_\_\_\_\_ Nome: \_\_\_\_\_

PARTE I – Escolha Múltipla (10V)

1. (1V) Numa árvore de decisão:
  - A. Uma folha pode ser constituída por instâncias com diferentes valores na variável dependente
  - B. A profundidade da árvore não tem qualquer relação com a sua tendência para *overfitting*
  - C. Não é possível ter folhas com apenas uma instância
  - D. Os nós representam decisões tomadas com base numa instância dos dados de treino
2. (1V) Assinale o algoritmo que, quando aplicado com a mesma configuração aos mesmos dados, resulta sempre no mesmo modelo:
  - A. Rede Neuronal
  - B. Random Forest
  - C. Naïve Bayes
  - D. Nenhum dos restantes
3. (1V) Numa Random Forest de regressão:
  - A. A previsão é calculada através do valor mais frequente nas previsões de cada árvore
  - B. A previsão é calculada através do valor médio das previsões de cada árvore
  - C. A previsão é calculada somando os valores das funções de ativação
  - D. A previsão é calculada utilizando o teorema de Bayes
4. (1V) Assinale o algoritmo em que a complexidade de fazer previsões cresce com o nº de instâncias do dataset de treino:
  - A. Árvores de Decisão
  - B. Random Forest
  - C. Nenhuma das restantes, a complexidade de fazer previsões depende sempre e apenas do tamanho do dataset para o qual se está a prever, e não do dataset de treino
  - D. K-Nearest Neighbours
5. (1V) Considere a rede neuronal que se apresenta à direita:
 

- A. O dataset utilizado no seu treino tem, no máximo, 12 variáveis
  - B. O dataset utilizado no seu treino tem, no máximo, 9 variáveis
  - C. O dataset utilizado no seu treino tem, no máximo, 8 variáveis
  - D. O dataset utilizado no seu treino tem, no



 <small>ESCOLA SUPERIOR DE TECNOLOGIA E GESTÃO</small>	Tipo de Prova Exame Teórico – Época Normal	Ano letivo 2021/2022	Data 23-06-2022
	Curso Licenciatura em Engenharia Informática	Hora 10:00	
	Unidade Curricular Inteligência Artificial	Duração 2:30 horas	


- máximo, 4 variáveis
6. (1V) Considere que, para um determinado problema de Machine Learning, tem uma quantidade de dados considerada mais que suficiente, e pouco tempo disponível para treinar o modelo. Das seguintes opções, assinale a metodologia de treino e avaliação de modelos mais adequada para este cenário:
- A. N-fold cross validation
  - B. Train-test split ou hold-out method**
  - C. Backpropagation
  - D. Gradient Descent
7. (1V) Assinale a métrica de performance mais adequada para avaliar a qualidade de um modelo de regressão:
- A. Nenhuma das restantes**
  - B. AUC
  - C. Precision
  - D. Recall
8. (1V) Considere a base de conhecimento Prolog apresentada à direita. A questão `passa(joao, Z)`:
- A. Falha, pelo princípio do mundo fechado**
  - B. Falha, porque a questão está mal formulada
  - C. Tem sucesso, com  $Z = 12$
  - D. Tem sucesso, com  $Y = 12$ .
- ```

passa(X, Y):- aluno(X),
                nota(X,Y),
                Y > 9.5.

aluno(baião).
aluno(maria).
aluno(carlos).

nota(joao, 12).
nota(maria, 8).
nota(carlos, 15).

```
9. (1V) Uma das vantagens das redes neurais profundas (Deep Learning), quando comparadas com outros modelos mais simples é que:
- A. Permite automatizar o processo de treino de um modelo
  - B. Permite automatizar o processo de *feature extraction***
  - C. Permite treinar modelos de forma mais rápida
  - D. Permite treinar ensembles constituídos por conjuntos de redes neurais
10. (1V) Indique qual das seguintes é uma característica dos Algoritmos Genéticos:
- A. Precisam de um grande conjunto de dados de treino
  - B. O tempo de treino do modelo é elevado
  - C. Não garantem uma solução ótima**
  - D. Assumem que as variáveis são independentes

|                                                                                                                                                        |                                                 |                         |                    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------|-------------------------|--------------------|
| <br><small>ESCOLA<br/>SUPERIOR<br/>DE TECNOLOGIA<br/>E GESTÃO</small> | Tipo de Prova<br>Exame Teórico – Época Normal   | Ano letivo<br>2021/2022 | Data<br>23-06-2022 |
|                                                                                                                                                        | Curso<br>Licenciatura em Engenharia Informática | Hora<br>10:00           |                    |
|                                                                                                                                                        | Unidade Curricular<br>Inteligência Artificial   | Duração<br>2:30 horas   |                    |

## PARTE II – Prolog (5V)


11. Considere o famoso jogo Tic-Tac-Toe, conhecido em Portugal como o jogo do galo, que se joga num tabuleiro de 3x3. Os jogadores vão colocando as suas peças ou símbolos sucessivamente, com o objetivo de conseguir preencher uma linha, coluna ou diagonal com os seus símbolos. O jogo termina quando isso acontecer, com a vitória desse jogador, ou quando todas as posições estiverem preenchidas, terminando assim num empate.
- Pretende-se modelar o funcionamento do jogo do galo em Prolog, utilizando os seguintes factos:
- `simbolo(n, s)` – representa que o jogador `n` está a utilizar o símbolo `s` para jogar
  - `jogada(n, l, c)` – representa que o jogador `n` jogou o seu símbolo na posição (`l` -> linha, `c` -> coluna)
  - `proximo(n)` – representa que o próximo jogador é o `n`
- Apresenta-se de seguida o estado de um jogo, a sua representação em Prolog, e a sua descrição. Assuma que as coordenadas começam em (1,1) no canto superior esquerdo.
- |   |   |  |
|---|---|--|
| X | O |  |
| O | X |  |
|   |   |  |
- ```

simbolo(maria, 'O').
simbolo(joao, 'X').

jogada(maria, 1, 2).
jogada(joao, 1, 1).
jogada(maria, 2, 1).
jogada(joao, 2, 2).

proximo(maria).

```
- O jogo é entre a Maria e o João.
  - A Maria usa o símbolo 'O'.
  - O João usa o símbolo 'X'.
  - Já foram feitas 4 jogadas, sendo que foi a Maria que começou a jogar.
  - O próximo jogador a jogar é a Maria.
- 11.1 (1.5V) Implemente em Prolog a regra `valida/4` que, dado o nome de um jogador, um símbolo, e a posição (linha e coluna) em que pretende jogar, determina se essa jogada é válida. Considere as seguintes regras:
- Um jogador apenas pode jogar se for a sua vez
  - Um jogador apenas pode jogar numa posição livre
  - Um jogador apenas pode jogar utilizando o seu símbolo
- 11.2 (1.5V) Implemente em Prolog a regra `terminou/0` que determina se o jogo já terminou
- 11.3 (2V) Implemente em Prolog a regra `vencedor/1` que determina o nome do vencedor do jogo. Caso o jogo ainda não tenha terminado ou tenha terminado em empate, a regra deve falhar.

 <small>ESCOLA SUPERIOR DE TECNOLOGIA E GESTÃO</small>	Tipo de Prova Exame Teórico – Época Normal	Ano letivo 2021/2022	Data 23-06-2022
	Curso Licenciatura em Engenharia Informática	Hora 10:00	
	Unidade Curricular Inteligência Artificial	Duração 2:30 horas	

### PARTE III – Desenvolvimento (5V)

12.  
(2.5V)

Considere o algoritmo Random Forest. Descreva brevemente o seu funcionamento, indicando nomeadamente as diferenças no seu funcionamento entre problemas de regressão e de classificação. Indique ainda se este algoritmo é adequado para evitar overfitting, e que configurações podem ser utilizadas para que tal aconteça.

13.  
(2.5V)

Considere o dataset cars\_origin, utilizado nas aulas práticas de Inteligência Artificial, e cujo excerto se apresenta abaixo. A tarefa de Machine Learning associada consiste em tentar adivinhar a origem de um carro (US, Europe ou Japan) dadas algumas das suas características, tais como o número de cilindros, nº de cavalos ou ano de fabrico.

mpg	cylinders	cubicinches	hp	weightlbs	time-to-60	year	origin
14	8	350	165	4209	12	1972	US.
31.9	4	89	71	1925	14	1980	Europe.
17	8	302	140	3449	11	1971	US.
15	8	400	150	3761	10	1971	US.
30.5	4	98	63	2051	17	1978	US.
23	8	350	125	3900	17	1980	US.
13	8	351	158	4363	13	1974	US.
14	8	440	215	4312	9	1971	US.
25.4	5	183	77	3530	20	1980	Europe.
37.7	4	89	62	2050	17	1982	Japan.
34	4	108	70	2245	17	1983	Japan.

Desenhe, justificando, uma possível arquitetura de uma Rede Neuronal para o problema proposto. Note que é suficiente desenhar cada uma das camadas que considerar necessárias e os seus neurónios, não sendo necessário desenhar as ligações entre os neurónios.

Indique ainda que métricas e/ou outros elementos poderiam ser utilizados para avaliar a qualidade do modelo resultante.