

Kahoot 5

6 – Machine Learning is a new way of building computer programs because

- a) We use data and output to build the program
- b) We use data and a program to build the output
- c) We are able to learn the input data from the program without using output
- d) We solve new problems using software engineering

7 – O que são outliers?

- a) Dados extremos que diferem da distribuição habitual dos restantes dados
- b) Dados com ruído que nunca podem ser usados para ML
- c) Dados errados
- d) Dados com distribuição normal mas classificação errada

8- Eliminar dados, Estimar valores, Ignorar indivíduos/variáveis durante a análise

- a) São métodos para Lidar com Ruído
- b) São métodos para Explorar Dados
- c) São métodos para Lidar com Valores Omissos
- d) São métodos para analisar duplicados

9 – As seguintes são Metodologias para extração de conhecimento

- a) KDD, KAPS, CREEP
- b) KAPS, CREEP, CRISP-DM
- c) KDD, CRISP-DM, SEMMA
- d) SEMMA, Analysys+, DataMark

Kahoot 6

1 - Types of machine learning

- a) Supervised Labeled
- b) Reinforcement, Supervised, Unsupervised
- c) Supervised, Undupervised, Mixed, Global
- d) Reinforcement, Deep

2 – Types/applications of Supervised Learning

- a) Reinforcement and Unsupervised
- b) Classification and Labeling
- c) Regression and Classification
- d) Unsupervised, Labeling Regression

3 – In _____ we try to find the decision boundary which can divide the dataset into different groups.

- a) Classification
- b) Machine Learning
- c) Supervised Learning
- d) Regression

4 – _____ Learning can be used for those cases where we have labelled input data

- a) Supervised
- b) Unsupervised
- c) Machine
- d) Reinforcement

5 – O método k-Fold Cross Validation

- a) Em cada experiencia (run) usa: k-1 folds para treino e 1 fold para teste
- b) Em cada experiencia (run) usa: 2 folds para treino e 1 fold para teste
- c) Em cada experiencia (run) usa: 1 folds para treino e 1 fold para teste
- d) Em cada experiencia (run) usa: 1 folds para treino e k-1 fold para teste

6 – Which of the following is characteristic of decision trees?

- a) Decision trees do factor analysis
- b) Decision trees are very robust to outliers
- c) Decision trees are prone to be overfit
- d) Decision trees use synapses

9 – Which of the following is not associated with Neural Networks

- a) Training
- b) Weights
- c) DNA strands
- d) Input/ Output Layer

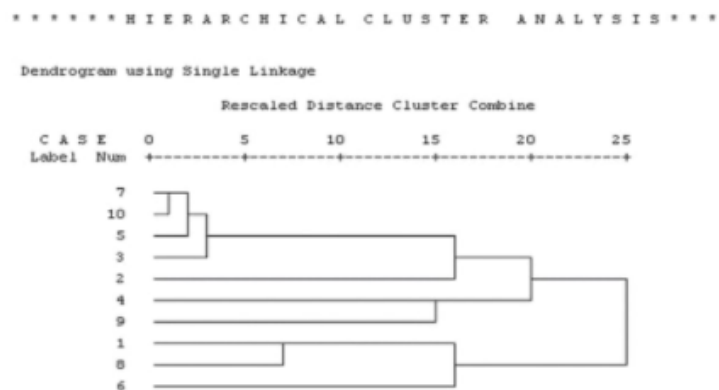
10 – Which Python Libraries are useful for Machine Learning

- a) Clion, PyTorch, PyGame
- b) Bear, NumPy, MathLib, MathPlot2
- c) Pandas, NumPy, Scikit Learn
- d) MathPlot, MathLib, LearnPy

Kahoot 7

9 – Supondo a geração final de três
quais os pontos que ficam em clusters

- a) 7 e 10
- b) 3 e 2
- c) 2 e 4
- d) 1 e 8



clusters
distintos

10 – Que Algoritmo de Clustering está representado na figura

- a) Algoritmo Hierarquico Aglomerativo
- b) Algoritmo K-Means
- c) Algoritmo Hierarquico Divisivo
- d) Algoritmo k-Nearest Neighbour

Kahoot 8

1 – Reinforcement Learning

- a) Is the most close to human learning of the machine learning main types
- b) Is a type of unsupervised learning
- c) Is a type of supervised learning
- d) Is a type of artificial networks learning

2 – Reinforcement learning is focused on

- a) Classification based learning with neural networks
- b) Goal-directed learning from interaction
- c) Value based learning from pre-classified examples
- d) Regression base learning from individuals

3 – In the exploration-exploitation trade-off in Reinforcement Learning, “Exploration” refers to:

- a) Agent do actions that it knows to be very effective
- b) Agent try new actions not completely analysed before
- c) Agent learns by using supervised data
- d) Agent do actions from its policy that are the optimal ones

4 – A Markov Decision Process includes:

- a) Actions, Policy, Algorithms and Type
- b) States, Actions, Reward Function and Transitions (Action's Effects)
- c) Actions base on : Q-Learning, SARSA, PPO
- d) Discount, Reward, Penalty, Update and Relation factors

5 – In RL, what is a Policy π ?

- a) A way of discussing with Supervised Algorithms
- b) The goal/objective function of the RL problem
- c) A mapping from perceived states to actions
- d) A policy specifies what is good in the long run and its evaluation function

6 – In the cumulative reward function what is represented by ' γ ' and multiplies by each term?

- a) The discounting factor
- b) The multiplicative factor
- c) The reward factor
- d) The penalty factor

7 – Which of the following are aimed at helping the researcher on implementing RL applications?

- a) RL Agents and SARSA tool
- b) Gym and RL-Tools2
- c) Q-Learning and AgentNet
- d) Open AI Gym and ML-Agents

8 – These are known Reinforcement learning algorithms

- a) SARSA, Neares Neighbour, StateAction
- b) Neural Networks, Q-Learning, SVMs
- c) SARSA, Q-Learning, PPO
- d) PPO, Neural Networks, KMeans

9 – In reinforcement learning

- a) The search is guided by a dataset of pre-classified/pre-evaluated actions
- b) The reward tells us the exact sequence of actions we should do to solve it
- c) The greedy action selection strategy uses both exploit and explore (50/50)
- d) The expected reward of a policy is the “value” of the policy

10 – Reinforcement Learning Achived Very Good Results on:

- a) Classifying Images for person identification
- b) Playing Games and performing Robotic Tasks

- c) Performing Robotic Tasks mainly Perception (vision and sound)
- d) Natural Language Processing

Kahoot 9

1 – A tarefa de tokenização consiste em:

- a) dividir um texto em frases
- b) dividir um texto em palavras ou partes de palavras
- c) dividir um texto em paragrafos
- d) determinar a categoria gramatical das palavras

2 – A tarefa de part-of-speech (POS) tagging consiste em :

- a) construir a árvore sintática da frase
- b) distinguir a categoria gramatical das palavras (nome verbo adjetivo) <-
- c) determinar correspondencia entre diferentes menções a um mesmo objeto
- d) desambiguar o significado das palavras

3 – A resolução de correferências consiste em :

- a) desambiguar o significado das palavras
- b) determinar a categoria gramatical das palavras
- c) determinar correspondencia entre diferentes menções a um mesmo objeto <-
- d) reconhecer entidades mencionadas no texto

5 – Dos seguintes classificadores, qual não é do tipo discriminativo

- a) SVM
- b) Decision tree
- c) Naïve bayes <-
- d) logistic regression

6 – A forma de representação conhecida por bag-of-words

- a) tem em conta a ordem das palavras no texto
- b) contabiliza o numero de ocorrências de cada palavra no texto <-
- c) permite capturer o contexto de cada palavra no text
- d) lida bem com fenómenos linguísticos com a negação

7 – A utilização de léxicos de sentimento:

- a) exige que todas as palavras do dicionário estejam etiquetadas
- b) limita-se à polaridade (positiva/negativa) das palavras
- c) pode permitir diferentes graus de sentimento nas palavras do léxico <-
- d) só faz sentido de o vocabulário dos conjuntos de treino e teste for igual

8 – Sobre word embeddings NÃO é verdade que:

- a) permitem representar palavras como vetores densos de números reais
- b) capturaram o significado das palavras com base no contexto onde ocorrem
- c) permitem explorar grandes quantidades de dados não anorados
- d) só podem ser utilizados com arquiteturas de deep learning <-

Kahoot 11 2021

1 – Strong AI refers to a kind of AI capable of solving

- a) simple tasks that demand for a strong power
- b) complex tasks like planning a trip
- c) a single narrow task
- d) any intellectual task that a human being can <-

2 – Used for: Understand functioning, Performance optimization, Testing and validation, Decision making, Training, etc

- a) Coordination
- b) Simulation <-
- c) Robotics
- d) Robotic Soccer

3 – What was the name of the agent that won at jeopardy against top human players

- a) bond
- b) holmes
- c) watson <-
- d) turing

4 – What was the algorithm used for this running skill?

- a) q-learning
- b) sarsa
- c) PPO <-
- d) Deep Q-learning

5 – These are Robotic soccer league

- a) Middle size, big size and small size
- b) middle size, simulation 3D and small size <-

6 – What is a country with an AI ministry in the government

- a) japan
- b) china
- c) united arab emirates <-
- d) brazil

Kahoot 10 - 2021

1 – Why is Natural language understanding difficult

- a) too much noise in speech-to-text
- b) no one really knows
- c) natural language is inherently ambiguous <-
- d) there are many languages

2 – A generative type of classifier

- a) gets the class that most likely produced a given observation <-
- b) classifies words based on neural networkd
- c) learns which features are most useful for discriminating between classes
- d) generates sequences of features to be analyzed

3 – Regarding word embeddings, it is NOT true that:

- a) they allow representing words as dense vectores of real numbers
- b) they capture the meaning of words based on the context on which they occur

- c) they can only be used with deep learning architectures <-
- d) they allow you to exploit large amounts of unlabeled data

Kahoot 9 - 2021

1 – these are interesting measures of node impurity that may be used in node splitting algorithms for decision trees

- a) Gini Index, C4.5, ID3
- b) Gini Index, Entropy, Classification Error <-
- c) Accuracy, precision, classification error
- d) accuracy precision, recall, F measure

2 – A typical value of k for the number of folds in k-fold cross validation is

- a) k=10 <-
- b) k=2
- c) k= number of training instances
- d) k=100

3 – When the model is too complex after some training the training error is small but the test error is large and increases

- a) Cross validation
- b) overfitting <-
- c) underfitting
- d) Error validation

4 – What is the accuracy for the example

- a) 50%
- b) 70% <-
- c) 10%
- d) 30%

	PREDICTED CLASS	
	Class=Yes	Class=No
	Class=Yes	Class=No
ACTUAL CLASS	30 (TP)	20 (FN)
	10 (FP)	40 (TN)

6 – using k-nn (with k=1) in which class is the example Par3=8 classified?

- a) Class B
- b) Class A <-
- c) It is not classified
- d) Class C

Treino	Par1	Par2	Par3	Classe
Ind1	3	8	9	A
Ind2	4	7	8	A
Ind3	2	2	5	B
Ind4	1	3	6	B
Ind5	8	6	5	C
Ind6	9	5	6	C
Ind7	2	1	6	B
Ind8	1	3	5	B
Ind9	4	9	8	A
Ind10	3	8	7	A
Ind11	9	4	6	C
Ind12	8	5	5	C
Ind13	1	1	1	C

Par1=3, Par2=8,

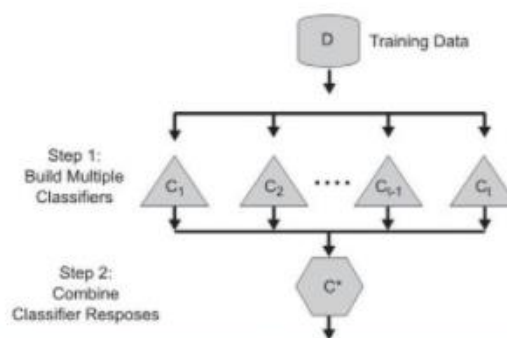
7 – Which of the following is not associated with

- a) Node split <-
- b) Layer
- c) Weights
- d) Backpropagation

neural networks?

9 – which learning method is depicted in the

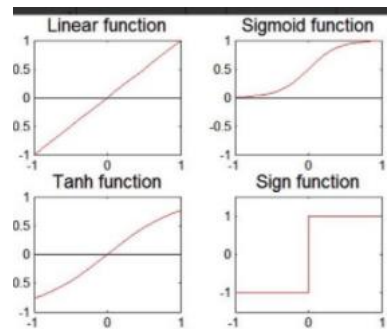
- a) Naïve Bayes
- b) Decision tree
- c) Neural network
- d) ensemble learning <-



figure?

10 – In which algorithm these “active functions” may be used

- a) Naïve Bayes
- b) Decision tree
- c) Neural network <-
- d) SVM



analysis and

Kahoot 8 - 2021

1 – Fast, powerful, flexible and easy to use open-source data manipulation tool

- a) Seaborn
- b) NumPy
- c) Keras
- d) Pandas <-

2 – Churn prediction for telephone customers was presented as ____ application

- a) Clustering
- b) Classification <-
- c) Regression
- d) Reinforcement learning

3 – Unify statistics, data analysis, ML and related methods to “understand/analyse actual phenomena” with data

- a) data science <-
- b) reinforcement learning
- c) supervised learning
- d) machine learning

4 – ____ works by interacting with the environment

- a) Data mining
- b) Reinforcement learning <-
- c) Machine learning
- d) Data science

5 – Knowledge extraction methodologies

- a) Kdd, kextract, orange, crisp-dm
- b) crisp-dm, semma, statisticaDM
- c) kdd, statisticaDM, Crisp-dm
- d) Semma, crisp-dm, kdd <-

6 – Examples of data quality

- a) ordinal, continuous, discrete
- b) Noise, outliers, missing values, duplicate data <-

Kahoot 7 - 2021

1 – Forward-chaining

- a) Is data-driven reasoning: starts from the known data <-
- b) Derives conclusions from incoming percepts, with a specific query in mind
- c) Is goal-directed reasoning: starts from a query
- d) Work backwards from a query

2 – First-order logic

- a) Uses facts, relations and times
- b) uses facts, objects and relations <-
- c) Is the same as propositional logic
- d) Uses facts with a degree of belief

3 – A knowledge-based system

- a) Uses machine learning and knowledge to show intelligent behaviour
- b) is particular type of expert system
- c) Is a type of Intelligent system <-
- d) Is a generalization of an intelligent system

4 – An expert system

- a) Does not include a knowledge base
- b) encodes expert knowledge in a knowledge base <-
- c) can only be used for diagnosis
- d) can only be built to reason through forward-chaining

2 – Why do we act under uncertainty

- a) Environments are static and deterministic
- b) environments are only partially observable and non-deterministic <-
- c) Environments are dynamic and continuous
- d) Environments are of different types several are dynamic and multi-agent

2 – Making rational decisions under uncertainty

- a) Is not possible
- b) requires one of the possible outcomes to be certain
- c) Requires considering preferences over uncertain outcomes <-
- d) Is only possible in deterministic environments

2 – Why is naive bayes naive?

- a) It doesn't work in practice
- b) Bayes was a naïve man
- c) It assumes variables are conditionally independent, even when they are not <-
- d) It assumes variables are conditionally independent, and they actually are

Exame 2020

13. O Mundo de Wumpus pode ser resolvido por um agente puramente reativo? Justifique

Não! O ambiente do mundo de Wumpus não é totalmente acessível e precisa de ser explorado. É preciso que o agente se lembre de estados anteriores, onde é que o "stench" e "breeze" estiveram, etc. Por isso, um agente reativo simples, embora, com sorte possa resolver algumas instâncias do problema, não pode com segurança resolver o mundo do wumpus.

14. Suponha os seguintes conjuntos de treino (indivíduos 1 a 13) e de teste (indivíduos 14, 15 e 16). Sem necessidade de fazer contas, supondo a distância Euclidiana, indique, justificando, quais as classes previstas para os indivíduos 14, 15 e 16 utilizando o algoritmo Nearest Neighbour, versões 1NN e 3NN.

1NN: 1 vizinho mais próximo

14: C ; 15: B ; 16: A

3NN: 3 vizinhos mais próximos

14: B ; 15: B ; 16: A

Treino	Par1	Par2	Par3	Classe
Ind1	3	8	9	A
Ind2	4	7	8	A
Ind3	2	2	5	B
Ind4	1	3	6	B
Ind5	8	6	5	C
Ind6	9	5	6	C
Ind7	2	1	6	B
Ind8	1	3	5	B
Ind9	4	9	8	A
Ind10	3	8	7	A
Ind11	9	4	6	C
Ind12	8	5	5	C
Ind13	1	1	1	C

Teste	Par1	Par2	Par3	Classe 1NN	Classe 3NN
Ind14	1	1	3	?	?
Ind15	2	4	6	?	?
Ind16	5	9	9	?	?

15. Suponha que dispõe de um ficheiro CSV com dados relativos à dimensão (em m2), localização (concelho), ano de construção (valor inteiro) e tipo de acabamentos (escala com 5 valores possíveis) de 20000 imóveis portugueses. Para cada imóvel está ainda disponível no CSV uma classificação do seu "tipo para venda" (simples, médio, luxo e super_luxo). Suponha que pretende usar aprendizagem supervisionada para treinar um classificador para, a partir da dimensão, localização, ano e tipo de acabamentos ser capaz de classificar o "tipo para venda" de novos imóveis. Apresente o código essencial (resumido) para, utilizando Python e as bibliotecas Pandas e SciKit Learn, ler o CSV, treinar um classificador com Nearest Neighbour ou SVMs utilizando 10 fold cross validation e apresentar os resultados de três das seguintes medidas: accuracy, precision, sensitivity e/ou f-measure. Nota: Não necessita de apresentar o código absolutamente correto bastando uma aproximação.

```
import pandas as pd
import sklearn as sk
data=pd.read_csv("ficheiro.csv")
treino=['direcao','localizacao','ano']
dados_treino=data[treino]
dados_teste=data['tipo']
x_treino,x_teste,y_treino,y_teste=sk.split(dados_treino,dados_teste,0,2)
modelo=sk.KNeighborsClassifier()
modelo.fit(x_treino,y_treino)
model.accuracy(x_teste,y_teste)
model.sensitivity(x_teste,y_teste)
model.precision(x_teste,y_teste)
```

16. Num problema de aprendizagem por reforço, ao chegar a um estado um agente tem ao seu dispor quatro ações, cujos valores Q são os seguintes: Q(A)=5, Q(B)=5, Q(C)=1, Q(D)=6. Seguindo uma política de seleção de ações do tipo soft-max, qual é a probabilidade, arredondada a 2 casas decimais, de escolha da ação B quando a temperatura é de 0.5?

$$P = (e^{5/0.5}) / (e^{5/0.5} + e^{5/0.5} + e^{1/0.5} + e^{6/0.5}) = 0.11$$

17. Considere o texto desta pergunta, que tem um vocabulário de tamanho 27 (número de sequências diferentes de caracteres alfanuméricos). Qual é a probabilidade de ocorrência da palavra "de", usando Laplace smoothing?

$$P = (\text{count}(\text{"de"},c)+1) / (\text{count}(w,c)+\text{voc_size}) = (5+1) / (27+31) = 0.10$$

Exame 2019

15 . Caracterize o conceito de overfitting e indique que cuidados devem ser tidos em conta na aplicação de redes neurais de modo a evitar este problema.

O conceito de overfitting surge quando o modelo que foi induzido se ajusta demasiado (sobreajuste) ao conjunto que foi utilizado para treino. No caso das redes neurais, para estas serem capazes de generalizar com base nos inputs, deve ser utilizado um número de exemplos de treino elevado, na ordem de 10 vezes o número de graus de liberdade da rede neuronal. Este elevado número de exemplos de treino vai permitir forte diversidade no conjunto de treino e, consequentemente, diminuir o grau de overfitting na rede neuronal.

16. Explique em que consistem e para que servem os métodos Holdout e Cross-Validation, distinguindo-os devidamente. Identifique um exemplo em que seja mais indicado utilizar Holdout e outro em que seja mais indicado utilizar Cross-Validation.

Os métodos Holdout e Cross-Validation servem para, no contexto de avaliação de modelos, efetuar a estimação da taxa de erro. O método Holdout, utilizado quando há muitos exemplos (> 1000 exemplos, ordem dos milhares), consiste em dividir 2/3 dos dados existentes para a fase de treino e os restantes 1/3 para a fase de teste. O método Cross-Validation, utilizado para conjuntos de tamanho intermédio (aproximadamente 1000 exemplos), consistem em dividir os dados em k partições disjuntas. Ao utilizar k-fold cross validation, o conjunto de dados é dividido em k partições (k-folds) e em cada experiência (run), usa k-1 folds como conjunto de treino e a restante (1 fold, atual) como conjunto de teste (iterativamente). Ao utilizar leave-one-out cross validation, um caso específico de k-fold em que k = N, usa, em cada iteração-experiência, N-1 exemplos para conjunto de treino e o exemplo restante para conjunto de teste. Por exemplo, num caso em que temos 15000 exemplos deveremos utilizar Holdout (devido ao elevado número de exemplos, cross-validation seria computacionalmente muito pesado) e num caso em que temos cerca de 800 exemplos deveremos utilizar cross-validation.

17. Considere a seguinte matriz de confusão que traduz os resultados de uma marca de teste de gravidez.

Identifique os valores dos Verdadeiros Positivos (VP); Falsos Negativos (FN); Falsos Positivos (FP) e Verdadeiros Negativos (VN). Calcule a taxa de acerto; a precisão; a sensibilidade e a medida F.

Resultado do Teste	Teste Positivo	Teste Negativo
Estado		
Grávida	45	5
Não grávida	20	30

Verdadeiro Positivo = Grávida e deu teste positivo; VP=45

Verdadeiro Negativo = Não-Grávida e deu teste negativo; VN=30

Falso Positivo = Não-Grávida e deu teste positivo; FP=20

Falso Negativo = Grávida e deu teste negativo; FN=5

Taxa de Acerto = $(VP+VN) / (VP+VN+FP+FN) = (45+30)/(45+30+20+5) = 0.75 = 75\%$

(Percentagem de Acerto do Teste de Gravidez – acertou em 75% dos casos)

Esta medida não funciona bem em conjuntos não balanceados)

Precisão = $VP / (VP+FP) = 45 / (45 + 20) = 0.692 = 69.2\%$

(Percentagem de Casos realmente Corretos de entre os 65 Testes Positivos)

Sensibilidade = $VP / (VP+FN) = 45 / (45 + 5) = 0.9 = 90\%$

(Percentagem de casos Corretos detetada pelo Teste de Gravidez de entre todas as Grávidas)

Medida F = $2TP / (2TP + FP + FN) = (2*45) / (2*45 + 5 + 20) = 0.783 = 78.3\%$

(Média harmónica da precisão e da sensibilidade)

Construa uma DCG que permita validar sintática e semanticamente frases do tipo [X, subiu/desceu, de, N1, valor(es) para, N2, valor(es)]. Por exemplo, a frase “Ana subiu de 14 valores para 12 valores” possui erro semântico, e a frase “Rui desceu de 5 valor para 1 valores” possui erro sintático.

Simple, but incomplete solution:

frase --> sintagma_nominal, sintagma_verbal.

sintagma_nominal --> nome.

nome --> ['Ana'].

nome --> ['Rui'].

sintagma_verbal --> verbo(N1, N2), proposição, numero(N1), valor(N1), proposição, numero(N2), valor(N2).

verbo(subir) --> ['subiu'], {N1 > N2 ; write('erro semantico')}.

verbo(descer) --> ['desceu'], {N1 < N2 ; write('erro semantico')}.

proposição --> ['de'].

proposição --> ['para'].

valor(N1) --> ['valor'], {N1 = 1 ; write('erro sintatico')}.

valor(N1) --> ['valores'], {N1 != 1 ; write('erro sintatico')}.

numero(12) --> ['12'].

numero(14) --> ['14'].

numero(5) --> ['5'].

numero(1) --> ['1'].

Frase --> SN, SV

SN --> Nome

Nome --> [Rui, Ana, Pedro, Maria, ...]

SV --> Verbo(n1, n2), De, Valor(n1), PalavraValor(n1), Para, Valor(n2), PalavraValor(n2)

Verbo: n1 > n2 --> subir

Verbo: n1 < n2 --> descer

Valor --> [0,1,2,3,...,18,19,20]

PalavraValor(x) --> x = 1 --> valor

PalavraValor(x) --> x != 1 --> valores

De --> [de]

Para --> [para]

Exame 2018

1. Seja S uma coleção de exemplos distribuídos por 3 classes. Há duas vezes mais exemplos da classe A do que da B, e o mesmo número de exemplos entre B e C. Calcule a informação média relativamente à classificação da coleção S.

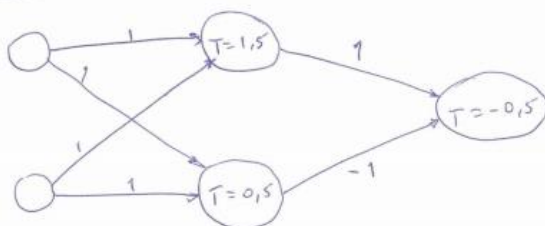
Class A: 0.5 / classB: 0.25 / classC: 0.25 / $\text{info}(S) = -0.5 \log_2(0.5) - 0.25 \log_2(0.25) - 0.25 \log_2(0.25) = 1.5$

2. Um agente movimentando-se num espaço percebe o ambiente através de um conjunto de sensores, a partir dos quais infere sobre a segurança da sua posição atual. Indique se este tipo de inferência se baseia em regras causais ou de diagnóstico. Justifique.

3. Regras de diagnóstico, uma vez que parte de efeitos observados em direção às causas desses efeitos.

Desenhe uma rede neuronal simples que consiga implementar a função NXOR (Não Ou Exclusivo).

NXOR



Name	Hair	Height	Weight	Lotion	Result
Sarah	blonde	average	light	no	sunburned
Dana	blonde	tall	average	yes	none
Alex	brown	short	average	yes	none
Annie	blonde	short	average	no	sunburned
Emily	red	average	heavy	no	sunburned
Pete	brown	tall	heavy	no	none
John	brown	average	heavy	no	none
Katie	blonde	short	light	yes	none

4. Chegado o Verão, há que ter cuidado com o sol. A tabela ao lado mostra o que aconteceu a cada uma de 8 pessoas, com base nos seus dados fisionómicos, e sabendo se usou ou não protetor solar.

a) Calcule a informação média relativamente à classificação desta coleção.

$E(S) = -3/8 \log_2(3/8) - 5/8 \log_2(5/8) = 0.954$

b) Usando o algoritmo ID3, determine que atributo deve ser colocado na raiz da árvore de decisão, justificando com cálculos.

$E(\text{res}/\text{hair}) = 4/8 [-2/4 \log_2(2/4) - 2/4 \log_2(2/4)] + 3/8 [-3/3 \log_2(3/3)] + 1/8 [-1/1 \log_2(1/1)] = 0.5$

$$E(\text{res/height}) = 3/8 [-2/3 \log_2(2/3) - 1/3 \log_2(1/3)] + 2/8 [-2/2 \log_2(2/2)] + 3/8 [-2/3 \log_2(2/3) - 1/3 \log_2(1/3)] = 0.689$$

$$E(\text{res/weight}) = 2/8 [-1/2 \log_2(1/2) - 1/2 \log_2(1/2)] + 3/8 [-2/3 \log_2(2/3) - 1/3 \log_2(1/3)] + 3/8 [-1/3 \log_2(1/3) - 2/3 \log_2(2/3)] = 0.939$$

$$E(\text{res/lotion}) = 5/8 [-3/5 \log_2(3/5) - 2/5 \log_2(2/5)] + 3/8 [-3/3 \log_2(3/3)] = 0.607$$

O atributo escolhido é "hair".

c) O algoritmo C4.5 usa o critério da razão do ganho para construir a árvore de decisão. Neste caso, determine qual o atributo que deve ser colocado na raiz da árvore de decisão, justificando com cálculos.

$$\text{infosep}(\text{hair}) = -4/8 \log_2(4/8) - 3/8 \log_2(3/8) - 1/8 \log_2(1/8) = 1.406$$

$$\text{infosep}(\text{height}) = -3/8 \log_2(3/8) - 2/8 \log_2(2/8) - 3/8 \log_2(3/8) = 1.561$$

$$\text{infosep}(\text{weight}) = -2/8 \log_2(2/8) - 3/8 \log_2(3/8) - 3/8 \log_2(3/8) = 1.561$$

$$\text{infosep}(\text{lotion}) = -5/8 \log_2(5/8) - 3/8 \log_2(3/8) = 0.954 \quad \text{GainRatio}(\text{hair}) = (0.954 - 0.5) / 1.406 = 0.321$$

$$\text{GainRatio}(\text{height}) = (0.954 - 0.689) / 1.561 = 0.1699 \quad \text{GainRatio}(\text{weight}) = (0.954 - 0.939) / 1.561 = 0.0098$$

$$\text{GainRatio}(\text{lotion}) = (0.954 - 0.607) / 0.954 = 0.364 \quad \text{O atributo escolhido é "lotion"}.$$

d) Calcule a razão de erro em cada uma das folhas da árvore que contém apenas o nó de decisão identificado na alínea anterior.

Aplicando o atributo "lotion":

- se lotion="no" então result="sunburned": 2 erros em 5 $\rightarrow (e+1)/(n+2) = 3/7 = 0.429$

- se lotion="yes" então result="none": 0 erros em 3 $\rightarrow (e+1)/(n+2) = 1/5 = 0.2$

5. Um técnico de informática decidiu recuperar computadores antigos para criar uma cloud. A sua confiança na solução construída pode ser representada pelas seguintes regras: SE sw recente ENTÃO funciona (FC=0.8); SE computadores em_bom_estado ENTÃO funciona (FC=0.7); SE checkup ok ENTÃO computadores em_bom_estado (FC=0.75). O software (sw) utilizado é recente, e o check up parece ter corrido bem (FC=0.9). O que concluir, e com que fator de certeza?

$$R3: \text{computadores em_bom_estado} = 0.9 * 0.75 = 0.675$$

$$R2: \text{funciona com FC} = 0.675 * 0.7 = 0.4725$$

$$R1+R2: \text{funciona com FC} = 0.8 + 0.4725 * (1-0.8) = 0.8945$$

6. Ao aplicar um algoritmo genético a um problema de maximização, obteve-se uma população com os cromossomas C1, C2, C3 e C4, com os valores 10, 15, 27 e 30, respetivamente. Determine a probabilidade de seleção de cada cromossoma.

$$10+15+27+30 = 82$$

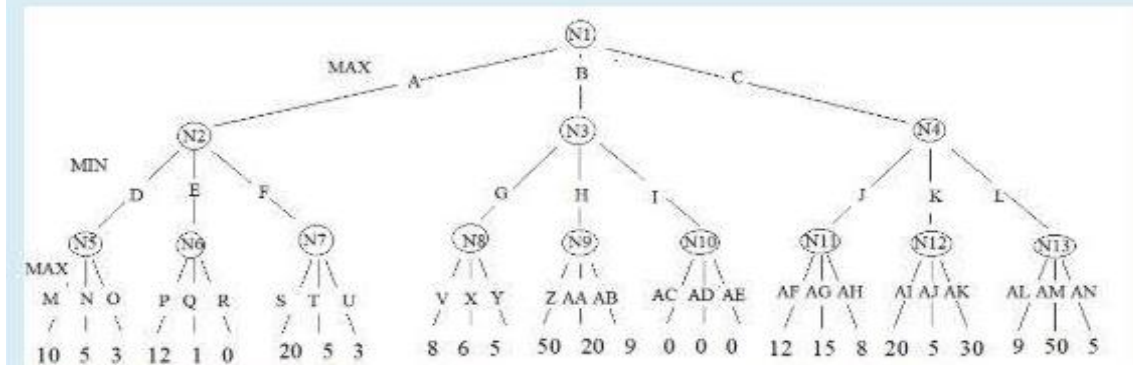
$$p(C1) = 10/82 = 0.122 \quad / \quad p(C2) = 15/82 = 0.183 \quad / \quad p(C3) = 27/82 = 0.329 \quad / \quad p(C4) = 30/82 = 0.366$$

7. Um agente tem conhecimento completo do seu ambiente, aplicando este conhecimento num sistema baseado em regras. Indique se tais regras são causais ou de diagnóstico. Justifique.

Estas regras são causais, uma vez que o agente conhece todo o ambiente, daí sabendo quais as causas de cada perceção que tem do mesmo.

Aplique o algoritmo minimax com cortes alfa-beta à seguinte árvore, indicando:

- Qual o valor final dos Nós N1, N2, N3, ... N13 (Nota: dado que são usados cortes alfa-beta o valor pode ser um número, ex: 8 ou uma condição ex: ≥ 20 ou ≤ 8).
- Quais os ramos cortados pelos cortes alfa-beta (de entre os ramos A, B, C, ... AN).



Suponha o seguinte dataset em que é analisado o facto de alguém emigrar para os EUA em função de três outras variáveis: a sua classe etária, nacionalidade e situação familiar.

Idade	Nacionalidade	Situação Familiar	Emigra EUA
jovem	alemã	solteiro	sim
media	francesa	solteiro	sim
media	alemã	solteiro	sim
jovem	italiana	solteiro	sim
media	alemã	casado	não
media	italiana	solteiro	sim
media	italiana	casado	não
jovem	alemã	casado	não

Sem necessidade de fazer contas indique qual o atributo de entropia mínima e quais as regras de decisão que resultam da aplicação do algoritmo ID3.

Selecione uma opção de resposta:

- ☐ a. Situação Familiar pois tem entropia baixa (entre 0.1 e 0.5)
- ☐ b. Nacionalidade pois tem a entropia mais baixa
- ☒ c. Situação Familiar pois tem entropia 0
- ☐ d. São todos iguais pelo que é indiferente
- ☐ e. Idade pois tem entropia 0

Figura P14 E2020 - Sem necessidade de fazer contas, supondo a distancia Euclidiana, indique quais as classes previstas para os indivíduos 14,15 e 16 utilizando o algoritmo Nearest Neighbour, versão 3NN

- Ind14 – B; Ind15 – B; Ind16 -A <- 3NN
- Ind14 – A; Ind15 – C; Ind16 – B
- Ind14 – B; Ind15 – A; Ind16 – A
- Ind14 – B; Ind15 – A; Ind16 – C
- Ind14 – C; Ind15 – B; Ind16 – A <- 1NN

Que individuo parece ser um outlier no exercício anterior?

- Individuo 13 pois tem valores dos parâmetros muito distantes dos restantes membros da classe C <-
- Individuo 6 pois tem todos os valores muito elevados

- c) Indivíduo 2 pois é o único com valor 7 no parâmetro 2
- d) Indivíduo 10 pois tem valores dos parâmetros muito distantes dos restantes membros da classe A

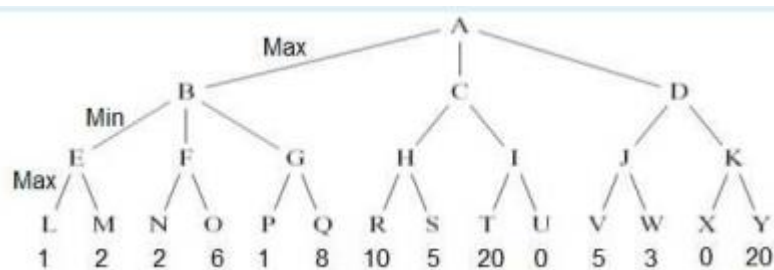
Compare o algoritmo KNN, a nível de tempo para realizar o treino, relativamente a outros algoritmos possíveis tais como as Redes Neurais e Support Vector Machines

- a) É mais leve computacionalmente, i.e mais rápido no treino, mas unicamente relativamente a redes neurais pois o treino de SVMs é sempre mais rápido
- b) É mais pesado computacionalmente pois o KNN tem de realizar N iterações
- c) É mais leve computacionalmente, i.e mais rápido no treino, pois nem precisa de treino <-
- d) É mais pesado computacionalmente pois o KNN tem de treinar todas as distancias a todos os indivíduos

10. Aplique o algoritmo minimax com cortes alfa-beta à seguinte árvore, supondo que joga primeiro MAX, depois MIN e novamente MAX, indicando:

a) Qual o valor final dos Nós A, B, C e D? (Nota: dado que são usados cortes alfa-beta o valor pode ser um número, ex: 8 ou uma condição ex: ≥ 20 ou ≤ 8)

b) Quais os nós folha (de entre L a Y) que não chegam a ser avaliados pelo algoritmo minimax com cortes alfa-beta?



a) A = 10, B = 2, C = 10, D ≤ 5 b) O, U, X, Y