

Calculo de Entropia:

2,32 / 2,45

Exemplo	Alternativo	Bar	Sex/Sab	fome	Cliente	Preço	Chuva	Res	Tipo	Tempo	Vai esperar?
X1	Sim	Não	Não	Sim	Alguns	RRR	Não	Sim	Francês	0-10	Sim
x2	Sim	Não	Não	Sim	Cheio	R	Não	Não	Tailandês	30-60	Não
x3	Não	Sim	Não	Não	Alguns	R	Não	Não	Hambúrguer	0-10	Sim
x4	Sim	Não	Sim	Sim	Cheio	R	Sim	Não	Tailandês	10-30	Sim
X5	Sim	Não	Sim	Não	Cheio	RRR	Não	Sim	Francês	>60	Não
X6	Não	Sim	Não	Sim	Alguns	RR	Sim	Sim	Italiano	0-10	Sim
X7	Não	Sim	Não	Não	Nenhum	R	Sim	Não	Hambúrguer	0-10	Não
X8	Não	Não	Não	Sim	Alguns	RR	Sim	Sim	Tailandês	0-10	Sim
X9	Não	Sim	Sim	Não	Cheio	R	Sim	Não	Hambúrguer	>60	Não
X10	Sim	Sim	Sim	Sim	Cheio	RRR	Não	Sim	Italiano	10-30	Não
X11	Não	Não	Não	Não	Nenhum	R	Não	Não	Tailandês	0-10	Não
X12	Sim	Sim	Sim	Sim	Cheio	R	Não	Não	Hambúrguer	30-60	Sim

classe

Entropia classe: $E(7/12, 2/12, 3/12)$

$$= -\frac{7}{12} \times \log_2(7/12, 3) - \frac{2}{12} \times \log_2(2/12, 3) - \frac{3}{12} \times \log_2(3/12, 3) = 0,87$$

2-R

BASE 3, POIS TEM 3 ATRIBUTOS

3-RRR

atributo

Entropia tipo: $E(2/12, 4/12, 4/12, 2/12)$

2-Franc

$$\rightarrow 2/12 \times E(0/2, 0/2, 2/2) + 4/12 \times E(3/4, 1/4, 0/4) + 4/12 \times E(4/4, 0/4, 0/4) + 2/12 \times E(0/2, 1/2, 1/2)$$

4-Tai.

4-Hamb.

2-Itali

Francés = 0

Hambúr. = 0

quantidade de tailandeses.

$$\text{Tailand.} = -\frac{3}{4} \times \log_2(3/4, 3) - \frac{1}{4} \times \log_2(1/4, 3) \rightarrow x^{4/12} = 0,170$$

$$\text{Itali.} = -0 \times \log_2(0, 3) - \frac{1}{2} \times \log_2(1/2, 3) - \frac{1}{2} \times \log_2(1/2, 3) \rightarrow x^{2/12} = 0,105$$

quantidade de Italianos.

ganhos(atributo) = Entropia(classe) - Entropia(atributo)

* Note formar a entropia da 0.

$$\text{Entropia tipo: } 0,17 + 0,105 = 0,275$$

$$\text{Ganhos tipo: } 0,87 - 0,275 = 0,595$$

Ex: $\log_2(5/11) =$

$$\log 5 / \log 2 = x$$

$$\log 11 / \log 2 = y$$

$y - x = \text{resposta}$

	osa	11	0	0
olor		0	13	0
nica		0	1	5

Iris-setosa Iris-versicolor Iris-virginica

	Recall	Precisão	F1	TVP	TFN	TFP	TVN / TYN
Setosa	11/11 = 1	11/11 = 1	1	1	0	0/(11+5+1)	1-0=1
Versicolor	13/13 = 1	13/14 = 0,92	0,96	1	0	1/(11+5+1)	1-0,05=0,95
Virginica	5/6 = 0,8	5/5 = 1	0,90	0,8	0,2	0/(11+13)	1-0=1

Recall: Analisa a linha.

Precisão: Analisa a coluna.

F1: Fórmula: $2 \times \text{recall} \times \text{precisão} / (\text{recall} + \text{precisão})$

TVP: Mesma regra do recall.

TFN: $1 - \text{TVP}$ \rightarrow analisa pela coluna.

TFP: A / B . A = quantos foram classificados de forma indevida. B = o resto.

TVN: $1 - \text{TFP}$

\rightarrow Da 1º/ base direta, subtraindo da soma das linhas de resto, o nº dos que foram classificados de forma indevida.

	Recall	Prec	F1	TVP	TFN	TFP	TVN
A	10/11	10/11	2x0,91	10/17	7/11	7/105	98/105
B	=	=	=	=	=	=	=

Recall = $\frac{VP}{VP+FN}$ = $\frac{10}{10+7} = \frac{10}{17}$

Foi classificado como

	A	B	C	D	
A	10	4	2	1	18
B	1	15	2	0	30
C	2	3	20	5	
D	4	1	2	50	57

3)

	Precisão	Recall	F1Score	TVP	TFN	TFP	TVN
A	10/17 ^{0.58}	10/17 ^{0.58}	0,58	10/17	7/17	7/105	98/105
B	15/23 ^{0.65}	15/18 ^{0.83}	0,72	15/18	3/18	8/104	96/104
C	20/26 ^{0.76}	20/30 ^{0.66}	0,71	20/30	10/30	6/92	86/92
D	50/56 ^{0.89}	50/57 ^{0.87}	0,88	50/57	7/57	6/65	59/65

4) O algoritmo CART utiliza essa métrica (Gini) para decidir como dividir os dados em subgrupos em cada nó da árvore. A ideia é escolher a divisão que minimize a impureza nos subgrupos resultantes.

A fórmula é:

$$\text{Gini}(D) = 1 - \sum_{i=1}^c p_i^2$$

Em que:

D é o conjunto de dados em um nó da árvore.

c é o número de classes no conjunto de dados.

p_i é a proporção de registros da classe i no conjunto de dados D.

Um valor de Gini próximo de 0 indica uma pureza alta, enquanto um valor de Gini próximo de 1 indica uma impureza alta.

1)

Setosa, varicolor, virginica
39 37 44

Esg = V
Dir = F

(a) A ou C.

C) Versicolor, Setosa, Versicolor, Virgínica

→ Alternativa correta.

2)

Alternativa 1 está correta.

Alternativa 2 está correta

Alternativa 3 está errada pois a menor cobertura de classe correspondente é de 2%.

C) 1 e 2 apenas.

NAVIE BAYES

Introdução

História do crédito	Dívida	Garantias	Renda Anual	Risco
Ruim	Alta	Nenhuma	<15000	Alto
Desconhecida	Alta	Nenhuma	>=15000 a <=35000	Alto
Desconhecida	Baixa	Nenhuma	>=15000 a <=35000	Moderado
Desconhecida	Baixa	Nenhuma	>35000	Alto
Desconhecida	Baixa	Nenhuma	>35000	Baixo
Desconhecida	Baixa	Adequada	>35000	Baixo
Ruim	Baixa	Nenhuma	<15000	Alto
Ruim	Baixa	Adequada	>35000	Moderado
Boa	Baixa	Nenhuma	>35000	Baixo
Boa	Alta	Adequada	>35000	Baixo
Boa	Alta	Nenhuma	<15000	Alto
Boa	Alta	Nenhuma	>=15000 a <=35000	Moderado
Boa	Alta	Nenhuma	>35000	Baixo
Ruim	Alta	Nenhuma	>=15000 a <=35000	Alto

Questão 07 (8,5 pontos) Considere a seguinte base de dados, em que se deseja avaliar se o cogumelo é comestível ou não.

RUIM			
C - D	B - W - G - U	BOM	
C = 5 D = 6	B = 5 W = 4 G = 1 U = 1	UIM = 3 OH = 5 TIMO = 3	
5/11	6/11	3/11	
= 6/11			
Forma	Cor	Odo	Comestível
C	(B)	Ruim	Yes
D	(B)	Ruim	Yes
B	W	Ruim	Yes
G	W	(Bom)	Yes
U	W	(Bom)	Yes
OH	G	(Bom)	No
TIMO	G	(Bom)	No
	C	U	(Bom)
	C	U	No
	C	W	Ótimo
	C	W	No
	D	W	Ótimo
	D	W	No

Usando Naive Bayes, qual a probabilidade de o cogumelo ser comestível (YES) ou não comestível (NO), respectivamente, para o seguinte registro:



Como avaliar um novo registro, dada esta tabela?

Registro de teste:

História: Boa
Dívida: Alta
Garantias = Nenhuma
Renda = >35000

$$P(\text{alto}) = 6/14 * 1/6 * 4/6 * 6/6 * 1/6$$

$$P(\text{alto}) = 0,0079$$

$$P(\text{alto}) = 0,0079 / 0,0645 * 100 = 12,24\%$$

$$\begin{aligned} \text{Soma:} \\ 0,0079 + 0,0052 + 0,0514 = \\ 0,0645 \end{aligned}$$

FAZER TESTE

$$P(\text{baixo}) = 5/14 * 3/5 * 2/5 * 3/5 * 5/5$$

$$P(\text{baixo}) = 0,0514$$

$$P(\text{baixo}) = 0,0514 / 0,0645 * 100 = 79,68\%$$

Risco de crédito	História do crédito			Dívida	Garantia	Renda anual
	Desconhe cida	Ruim	Baixa			
Alto 6/14	5	4	1/6	2/6	6/6	0/6
Moderado 3/14	1/3	1/3	1/3	2/3	2/3	1/3
Baixo 5/14	3/5	2/5	0/5	2/5	3/5	2/5

Tabela de produtividade

$$\rightarrow \text{FORMA=D} \mid \text{COR=B} \mid \text{ODOR=BOM}$$

FORMA	COR	ODOR
5/11 Sim 2/5	3/5 3/5	2/5 0/5 3/5
6/11 Não 3/6	3/6 2/6	2/6 1/6 1/6 0/6 3/6

$$\text{Reporto: YES: } 5/11 \times 3/5 \times 3/5 \times 2/5 = 0,06545$$

$$\text{NO: } 6/11 \times 3/6 \times 2/6 \times 3/6 = 0,04545$$

$$\begin{aligned} \text{total: } 0,06545 + 0,04545 \\ = 0,1109 \end{aligned}$$

$$\text{Probabilidade: YES: } 0,06545 / 0,1109 \times 100 = 59,01\%$$

$$\text{NO: } 0,04545 / 0,1109 \times 100 = 40,98\%$$

$$\text{C) } 59,01\% \text{ e } 40,98\%$$

ALGORITMO ID3

- Só resolve problemas de questão Nominal.
- Utiliza O GANHO de informação pra avaliar a importância do atributo

ALGORITMO C4.5

- trabalha com base de dados com dados ausentes
- Ele avalia a RAZÃO DE GANHO de informação para avaliar a importância do atributo.

GRIDSEARCH e RANDOMSEARCH

- Encontram os melhores hiperparâmetros de algoritmos de aprendizado, desde que o usuário forneça a faixa de variação a ser investigada.
- O Random, não testa todas as combinações possíveis de hiperparametros, explorando e selecionando de forma aleatória.
- O Grid, busca um conjunto pré definidos de valores para cada hiperparametro, explorando todas as combinações possíveis, não selecionando aleatoriamente, como o Random.

RANDOM FOREST

- Utiliza múltiplas árvores de decisão e faz as decisões, combinando as previsões de cada árvore, individualmente -> método “essemble learning”.

CART

- É simples e fácil de lidar com conjuntos de dados grandes e complexos.
- Começa com todo o conjunto de dados e divide eles em subconjuntos homogêneos.
- Utiliza o índice de Gini para saber quais atributos e valor de corte deve usar. utiliza esse método para saber como dividir os dados em subgrupos.

PROVA ANTIGA:

① TFP da classe C. $\frac{16}{197} = 0,083$

② 1 - correto

2 - correto

3 - Errado

→ A B C D E

$5/25$ $8/38$ $33/23$ $30/50$ $11/78$

* As classes não aquelas que possuem filtro.

* A leitura por classe, englobam apenas os filtro.

LISTA 4

1) Probabilidad de fumar ou não || aparência = chave, temp = fria, humidade = normal, ventando = sim

aparência	temp	humidade	ventando						
S 9/14 não 5/14	N 3/9 2/5	C 3/9 2/5	Q 2/9 2/5	A 4/9 1/5	F 3/9 1/5	A 3/9 1/5	N 6/9 3/5	S 3/9 2/5	N 6/9 2/5

$$\text{Probabilidad de fumar } \frac{9}{14} \times \frac{3}{9} \times \frac{3}{9} \times \frac{6}{9} \times \frac{3}{9} = 0,0158$$

$$\text{Roma} = 0,0158$$

$$\text{Probabilidad de não fumar } \frac{5}{14} \times \frac{2}{5} \times \frac{1}{5} \times \frac{1}{5} \times \frac{3}{5} = 0,00342$$

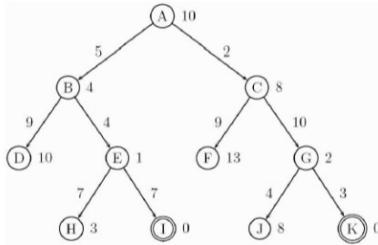
$$\text{PROBABILIDADES: SIM } 0,0158 / 0,0158 \times 100 = 82,29\%$$

$$\text{NÃO } 0,00342 / 0,0158 \times 100 = 17,81\%$$

PROVA 2

1-

No caso de escolhas equivalentes entre diferentes nodos, prefira o nodo mais próximo da raiz, seguido pelo nodo mais à esquerda na árvore. O algoritmo pára a busca quando encontra o I ou o K. Ou seja, não é necessário encontrar os dois objetivos.



- 1) Algoritmo de Busca em Largura
- 2) Algoritmo de Busca em Profundidade
- 3) Custo Uniforme
- 4) Algoritmo de Busca Gulosa
- 5) Algoritmo A*

1- **Largura** A - B - C - D - E - F - G - H - I

2- **Profundidade** A - B - D - E - H - I

3- **Custo uniforme** A - C - B - E - F - G - D - K

↳ avalia o custo da aresta (Varia dependendo do valor das arestas) *seleção sempre o menor valor.*

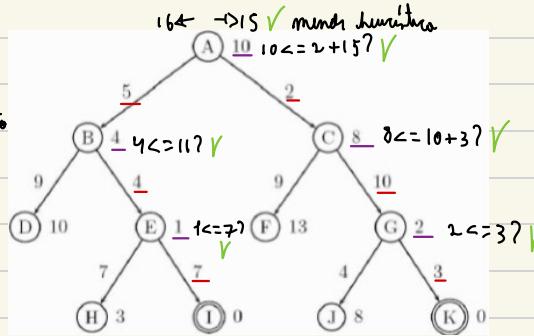
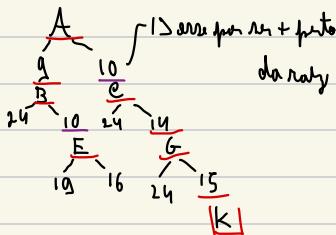
4- **Gulosa** A - B - E - I

↳ avalia o custo da célula (heurística) sempre var f/p/ o lado da menor heurística.

5- **A*** A, B, C, E, G, K

↳ se a heurística for admissível, tem收敛性 (convergência), caso contrário, não terá

↳ Como calcular uma heurística admissível.



É admissível

Questão 03

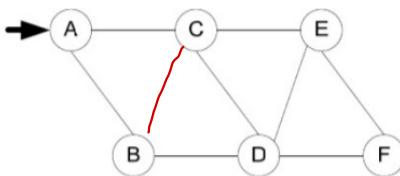
Julgue os itens a seguir, relativos a métodos de busca com informação (busca heurística) e sem informação (busca cega), aplicados a problemas em que todas as ações têm o mesmo custo, o grafo de busca tem fator de ramificação finito e as ações não retornam a estados já visitados.

- ✓ I. A primeira solução encontrada pela estratégia de busca em largura é uma solução ótima.
- ✗ II. A primeira solução encontrada pela estratégia de busca em profundidade é uma solução ótima.
- ✓ III. As estratégias de busca com informação usam funções heurísticas que, quando bem definidas, permitem melhorar a eficiência da busca.
- ✗ IV. A estratégia de busca gulosa é eficiente porque expande apenas os nós que estão no caminho da solução.
 - Isso realmente acontece, mas não necessariamente expande APENAS os nós do caminho de solução.

Estão certos apenas os itens

- a) I e II.
- ✗ b) I e III.
- c) I e IV.
- d) II e IV.
- e) III e IV.

* Se todos as ações tem o mesmo custo, a largura encontra a solução ótima



- ✗ A B C D E F
- B) A B D C E F
- C) A C D B F E
- D) A B C E D F
- E) A B D F E C

A - B - C - D - E - F

-> Como não passa parâmetro nenhum, seguimos por ordem alfabética, por isso a escolha de B antes de C. Apesar disso, os filhos do B, que é a letra D. Depois disso, o filho do C, que é E, depois isso volta para o filho do B, que é F.

- I. A estratégia de busca em largura encontra a solução ótima quando todos os operadores de mudança de estado têm o mesmo custo.
- II. A estratégia de busca em profundidade sempre expande um menor número de nós que a estratégia de busca em largura, quando aplicadas ao mesmo problema.
- III. A estratégia de busca heurística encontra sempre a solução de menor custo.
- IV. A estratégia de busca heurística expande um número de nós em geral menor que o algoritmo de busca em largura, mas não garante encontrar a solução ótima.
- V. O algoritmo de busca heurística que utiliza uma função heurística admissível encontra a solução ótima.

A esse respeito, pode-se concluir que

- (a) apenas a afirmativa V é correta.
- (b) todas as afirmativas são corretas.
- (c) todas as afirmativas são falsas.
- (d) apenas as afirmativas II e V são corretas.
- ✗ (e) apenas as afirmativas I, IV e V são corretas.

Questão 06 - POSCOMP 2007

[TE] Considerando que $h(n)$ é o custo estimado do nó n até o objetivo, em relação à busca informada, pode-se afirmar que

- (x) a busca *gulosa* minimiza $h(n)$. \rightarrow A busca gulosa sempre entra no menor valor de H . \rightarrow *percebe sempre pelos menores custos.*
- (b) a busca A^* minimiza $h(n)$. $\rightarrow F(n) = G(n) + H(n)$
- (c) a busca de custo uniforme minimiza $h(n)$. \rightarrow é igualdade, é i) Heurística, é ii) Cega, minimizando a $G(n)$. $F(n) = G(n)$
- (d) a busca *gulosa* minimiza $h(n)$ somente se a heurística for admissível.
- (e) a busca A^* minimiza $h(n)$ somente se a heurística for admissível.

F

F

Questão 07 - POSCOMP 2005

Considere $h(x)$ como uma função heurística que define a distância de x até a meta; considere ainda $h^r(x)$ como a distância real de x até a meta. $h(x)$ é dita admissível se e somente se:

- (a) $\exists n \ h(n) \leq h^r(n)$.
- (x) $\forall n \ h(n) \leq h^r(n)$.
- (c) $\forall n \ h(n) > h^r(n)$.
- (d) $\exists n \ h(n) > h^r(n)$.
- (e) $\exists n \ h(n) < h^r(n)$.

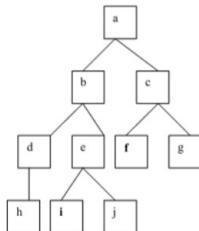
* Heurística é sempre \leq

$\exists n = \text{admissível}$

$\forall n = \text{inadmissível}$

Questão 8

59. Seja a árvore binária abaixo a representação de um espaço de estados para um problema p, em que o estado inicial é a, e i e f são estados finais.



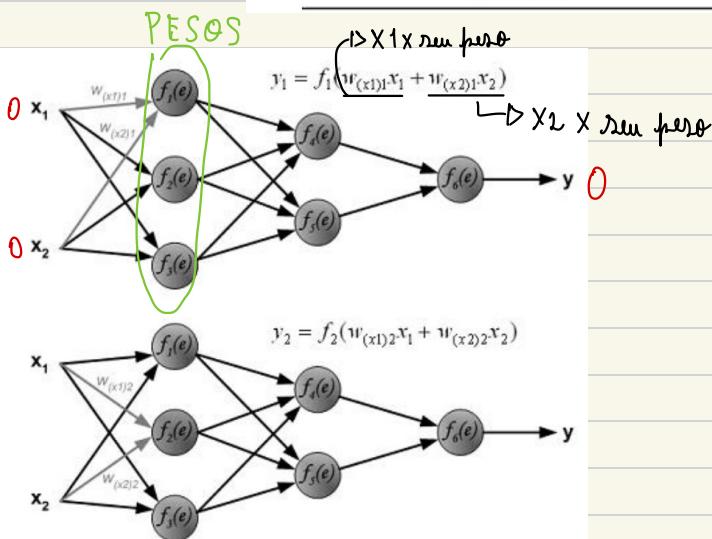
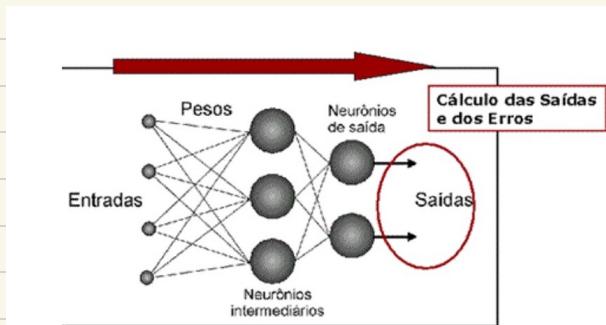
Um algoritmo de busca em largura-primeiro forneceria a seguinte seqüência de estados como primeira alternativa a um caminho-solução para o problema p:

- a) a b d h e i
x) a b c d e f
c) a b e i
d) a c f
e) a b d e f

RNA's

- Conjunto de neurônios que processam as informações
- Não dão explicações sobre as decisões tomadas
 - ↳ Solução p/ isso não é alguma metade de extração de conhecimento dos algoritmos
- São necessárias de muitas simulações p/ encontrar parâmetros adequados

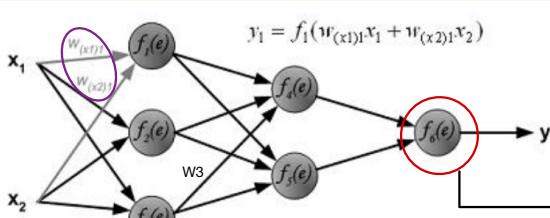
BACKPROPAGATION



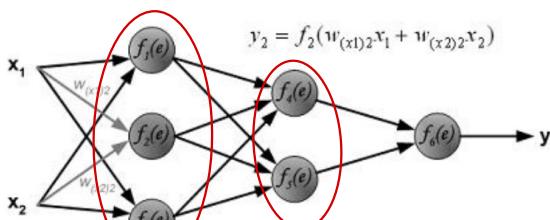
Calculo de erros

\triangleright de raiz da

- De o erro estiver no último neurônio, (desejado - raizda) \times derivada da função
 - \hookrightarrow se tende à 0, nenhuma mudança
- De o erro estiver na camada intermediária, (erro do neurônio anterior \times o seu atual peso)
 - \hookrightarrow neurônios intermediários



(desejado - raizda) \times derivada da função



(erro do neurônio anterior \times o seu atual peso)

Repetindo isso, p/ cada um dos 3 neurônios $\left((\text{erro } F_4 \times \text{peso } F_1) + (\text{erro } F_5 \times \text{peso } F_1) \right)$

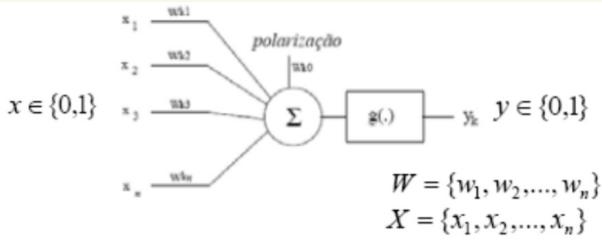
$\hookrightarrow F_1, F_2, F_3$

Ajuste das pesos:

Ex: w_2 peso $w_2 +$ taxa aprendizagem \times erro $F_1 \times x_2$

w_3 peso anterior + taxa aprendizagem \times erro $F_4 \times$ (a. entrada de F_4)

\hookrightarrow o que F_1 fez em p/ frente



x_i é a excitação de entrada na sinapse i ;
 y_k é a resposta (ou saída) do neurônio k ;
 w_{ki} é o peso sináptico da entrada i do neurônio k ;
 $g(\cdot)$ é a função de ativação do neurônio.

* Rede neural não trabalha c/ atributos nominais →

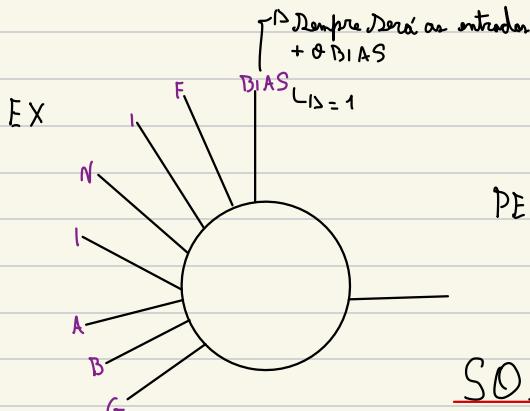
	Gênero	Raça	Idade	Faixa salarial	Saída
1	F	BRANCA	45	1000-3000	1
2	M	AMARELO	100	500-999	0
3	F	INDÍGENA	78	5001-7000	1
4	M	Negra	56	500-999	1
5	F	AMARELA	34	4000-5000	0

genero	branca	amarela	indigena	negra	idade	faixa salarial	Saída
1	1	1	0	0	45	1000-3000	1
2	0	0	1	0	100	500-999	0
3	1	0	0	1	78	5001-7000	1
4	0	0	0	0	56	500-999	1
5	1	0	1	0	34	4000-5000	0

Transformar a sequência de atributo nominal ordinal em sequência numérica.

1	2	3	4	5	6	7	Saída
genero	branca	amarela	indigena	negra	idade	faixa salarial	
1	1	1	0	0	45	2	1
2	0	0	1	0	100	1	0
3	1	0	0	1	78	4	1
4	0	0	0	0	56	1	1
5	1	0	1	0	34	3	0

Quando se tem valores muito diferentes, fazemos a normalização:
Fórmula: (valor atual) - (menor valor coluna) / (maior valor coluna) - (menor valor coluna)



PESOS. 0,2 0,1 0,001 0,02 -0,03 0,4 0,7 -0,1
BIAS

SOMATÓRIO

(Pegar o produto de cada entrada x x seu peso) + a mesma coisa do restante. Por linha.

EX:

$$1 - 1 \times 0,2 + 1 \times 0,1 + 0 \times 0,001 + 0 \times 0,02 + 0 \times -0,03 + 1,7 \times 0,4 + 0,3 \times 0,7 + 0,1 \times -0,1 = 0,5$$

erro será fute e/ todas as linhas

- Para os resultados dos somatórios, > 0 , saída = 1
 ≤ 0 , saída = 0

	genero	branca	amarela	indigena	negra	idade	faixa salarial	Saída	→ desejada	Saída Somatório	Saída final
1	1	1	0	0	0	0,17	0,33333333	1	0,5	1	
2	0	0	1	0	0	1	0	0	0,301	1	
3	1	0	0	1	0	0,67	1	1			
4	0	0	0	0	1	0,33	0	1			
5	1	0	1	0	0	0	0,66666667	0			

- Se a saída final for diferente da desejada, ajustar os pesos e gerar um novo tabela com os pesos, se quiser alterar.

Como alterar esses pesos? peso anterior + TA x produto da entrada x erro

→ deseja - saída

* Após ajustar todos os pesos, continuar c/ os cálculos de saída e/ os novos pesos

→ deseja

	genero	branca	amarela	indigena	negra	idade	faixa salarial	Saída	→ deseja	Saída Somatório	Saída final
1	1	1	0	0	0	0,17	0,33333333	1	0,5	1	
2	0	0	1	0	0	1	0	0	0,301	1	
3	1	0	0	1	0	0,67	1	1	0,59	1	
4	0	0	0	0	1	0,33	0	1	-0,2	0	
5	1	0	1	0	0	0	0,66666667	0			

- Como deu um novo erro, temos alterar novamente os pesos

→ deseja

	genero	branca	amarela	indigena	negra	idade	faixa salarial	Saída	→ deseja	Saída Somatório	Saída final
1	1	1	0	0	0	0,17	0,33333333	1	0,5	1	
2	0	0	1	0	0	1	0	0	0,301	1	
3	1	0	0	1	0	0,67	1	1	0,59	1	
4	0	0	0	0	1	0,33	0	1	-0,2	0	
5	1	0	1	0	0	0	0,66666667	0	0,32	1	

- Como deu um novo erro, temos alterar os pesos novamente

pesos originais	0,2	0,1	0,001	0,02	-0,03	0,4	0,7	-0,1
pesos alterados 1	0	-0,1	0,001	0,02	-0,03	0,366666667	0,633333333	-0,3
pesos alterados 2	0	-0,1	0,001	0,02	0,17	0,433333333	0,633333333	-0,1
pesos alterados 3	-0,2	-0,1	-0,199	0,02	0,17	0,433333333	0,5	-0,3

→

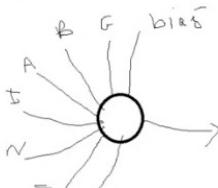
	genero	branca	amarela	indigena	negra	idade	faixa salarial	Bias	Saida desejada	Saidas somatória	Saída Final
1	1	1	0	0	0	0,1666666667	0,3333333333	1	1	0,5	1
2	0	0	1	0	0	1	0	1	0	0,301	1
3	1	0	0	1	0	0,6666666667	1	1	1	0,59	1
4	0	0	0	0	1	0,3333333333	0	1	1	-0,21	0
5	1	0	1	0	0	0	0,6666666667	1	0	0,32	1

s	0,2	0,1	0,001	0,02	-0,03	0,4	0,7	-0,1			
os 1	0	-0,1	0,001	0,02	-0,03	0,3666666667	0,6333333333	-0,3			
os 2	0	-0,1	0,001	0,02	0,17	0,4333333333	0,6333333333	-0,1			
os 3	-0,2	-0,1	-0,199	0,02	0,17	0,4333333333	0,5	-0,3			

função de ativação de limiar

>0 então saída 1

<0 então saída 0



Somatório do produto das entradas pelo peso

- 1 0,5
- 2 0,301
- 3 0,597777778
- 4 -0,207777778
- 5 0,323222222

$$w_{jl}(t+1) = w_{jl}(t) + \eta x^j \delta_l$$

CNN

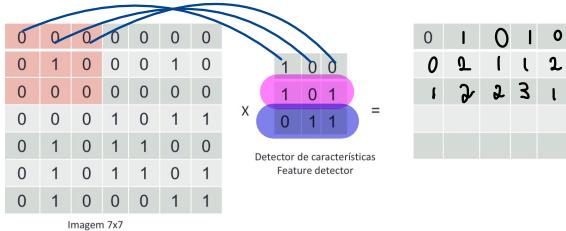
Rede neural convolucional

- Extrai características na etapa inicial Ex foto de pássaro

Funciona em 4 etapas.

- Operador de convolução.
- Pooling
- Flattening
- Rede neural densa.

Etapa 1 - Operador de convolução



→ depois disso, arredondamos uma linha.



$$0*1 + 0*0 + 0*0 + 0*1 + 1*0 + 0*1 + 0*0 + 0*1 + 0*1 = 0$$

Etapa 1 - Operador de convolução

0	0	0	0	0	0	0	0
0	1	0	0	0	1	0	
0	0	0	0	0	0	0	
0	0	0	1	0	1	1	
0	1	0	1	1	0	0	
0	1	0	1	1	0	1	
0	1	0	0	0	1	1	

X 1 0 0
 1 0 1
 0 1 1

Detector de características
Feature detector

=

0	1					
0	1					
0	1					
0	1					
0	1					
0	1					
0	1					

Imagen 7x7

$$0*1 + 0*0 + 0*0 + 1*1 + 0*0 + 0*1 + 0*0 + 0*1 + 0*1 = 1$$

→ Vamos arredondando de linha em linha até não dar mais

Etapa 1 - Operador de convolução

0	0	0	0	0	0	0	0
0	1	0	0	0	1	0	
0	0	0	0	0	0	0	
0	0	0	1	0	1	1	
0	1	0	1	1	0	0	
0	1	0	1	1	0	1	
0	1	0	0	0	1	1	

X 1 0 0
 1 0 1
 0 1 1

Detector de características
Feature detector

=

0	1	0	1	0		
0	2	1	1	2		
1	2	2	3	1		
1	3	3	3	2		
1	3	1	3	5		

Imagen 7x7

$$1*1 + 0*0 + 0*0 + 1*1 + 0*0 + 1*1 + 0*0 + 1*1 + 1*1 = 5$$

Mapa de características
Feature map

→ P/ levar só a mesma cara de 1 em 1 linha.

Transformar a tabela da esquerda na da direita.

Etapa 1 - Operador de convolução (RELU)

0	0	0	0	0	0	0	0
0	1	0	0	0	1	0	
0	0	0	0	0	0	0	
0	0	0	1	0	1	1	
0	1	0	1	1	0	0	
0	1	0	1	1	0	1	
0	1	0	0	0	1	1	

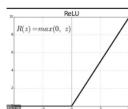
X 1 0 0
 1 0 1
 0 1 1

Detector de características
Feature detector

=

0	1	0	1	0		
0	2	1	1	2		
1	2	2	3	1		
1	3	3	3	2		
1	3	1	3	5		

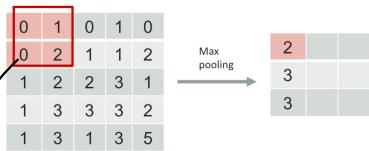
Imagen 7x7



→ Se algum valor na tabela for negativo, ele gera esse valor.

Etapa 2 - Pooling

- Seleciona as características mais relevantes da imagem (reduz overfitting e ruídos)
- Pode-se usar o máximo, o mínimo a média dos valores, dentre outros

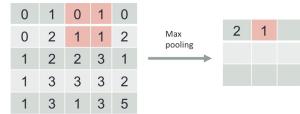


Pegamos o maior valor e colocamos na tabela.

- Andaremos p/ o lado, o tanto da marcação (neste caso, de 2 em 2)

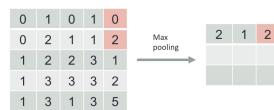
Etapa 2 - Pooling

- Seleciona as características mais relevantes da imagem (reduz overfitting e ruídos)
- Pode-se usar o máximo, o mínimo a média dos valores



Etapa 2 - Pooling

- Seleciona as características mais relevantes da imagem (reduz overfitting e ruídos)
- Pode-se usar o máximo, o mínimo a média dos valores

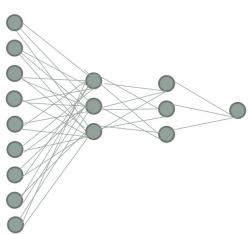
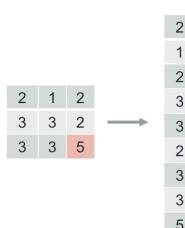


Etapa 2 - Pooling

- Seleciona as características mais relevantes da imagem (reduz overfitting e ruídos)
- Pode-se usar o máximo, o mínimo a média dos valores



Etapa 3 - Flattening



Rede neural densa

Dataser

A 1000

B 1000

Train test split

Treino	Teste
800A	200A
800B	200B

chamar a cross-validation e $k=5$

dividir em 5 partes, dividindo os 800 para 5 e completando os últimos

160

1	160 A e 160B				
2	160 A e 160B				
3	160 A e 160B				
4	160 A e 160B				
5	160 A e 160B				

	800 A 800 B				Treino	rodar a árvore
	vamos chamar a cross-validation com K=5				validação	
1	160 A e 160B	160 A e 160B	160 A e 160B	160 A e 160B	160 A e 160B	com quantas instâncias eu vou usar criar a árvore
2	160 A e 160B	160 A e 160B	160 A e 160B	160 A e 160B	160 A e 160B	1280 treino — Azul linha
3	160 A e 160B	160 A e 160B	160 A e 160B	160 A e 160B	160 A e 160B	320 validação — Verde linha
4	160 A e 160B	160 A e 160B	160 A e 160B	160 A e 160B	160 A e 160B	1600 — Soma da linha. 800 + 800
5	160 A e 160B	160 A e 160B	160 A e 160B	160 A e 160B	160 A e 160B	

- A parte verde nos dá uma acurácia, precisão e F1 score

* Agrupamento → quanto menor próximo de 1 a silhueta, melhor o agrupamento

1- Carregar a base de dados

← Avaliar a qualidade do agrupamento

2- Codificar os atributos

3- Eliminar redundâncias

4- Remoção outliers

5- Imputação de dados ausentes

6- Normalização

7- Seleção de atributos

8- Agrupar os dados

9- Avaliar a qualidade do agrupamento

10- Caracterizar os grupos

3 questões de Léxico

Mineração de Texto

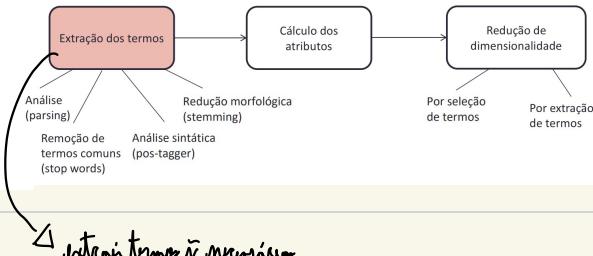
Mineração de textos

O objetivo na fase inicial do projeto é “transformar textos em números (índices significativos)”, que podem então ser incorporados em outras análises tais como problemas **supervisionados** ou **não supervisionados**.



Mineração de textos

O que deverá ser feito na etapa de pré-processamento do texto?



Exemplo

FREQUÊNCIA

- TFxIDF de alguns termos
 - Um termo comum: “Este”

$$\begin{aligned} \text{TF}(\text{"Este"}, D1) &= 1/5 = 0,2 \\ \text{TF}(\text{"Este"}, D2) &= 1/4 = 0,25 \\ \text{TF}(\text{"Este"}, D3) &= 1/6 = 0,17 \end{aligned}$$

$$\begin{aligned} (\text{TF}(\text{"Este"}, D1) \times \text{IDF}(\text{"Este"})) &= 0,2 \times 0 = 0 \\ (\text{TF}(\text{"Este"}, D2) \times \text{IDF}(\text{"Este"})) &= 0,25 \times 0 = 0 \\ (\text{TF}(\text{"Este"}, D3) \times \text{IDF}(\text{"Este"})) &= 0,17 \times 0 = 0 \end{aligned}$$

Colocar esses valores na tabela de palavras

uma palavra em todos os textos, sendo irrelevantes

• TF(i, D): número de vezes que o termo i aparece em relação ao total de termos do texto D.

• IDF(i): log do número de textos no corpus dividido pelo número de textos que o termo i aparece.

D1:	Este é um exemplo A.
D2:	Este é um mostruário.
D3:	Este é outro A, exemplo A.

Exemplo

TFxIDF de alguns termos

- Um termo de frequência mais variada: “A”

$$\begin{aligned} \text{TF}(\text{"A"}, D1) &= 1/5 = 0,2 \\ \text{TF}(\text{"A"}, D2) &= 0/4 = 0 \\ \text{TF}(\text{"A"}, D3) &= 2/6 = 0,33 \end{aligned}$$

$$\begin{aligned} (\text{TF}(\text{"A"}, D1) \times \text{IDF}(\text{"A"})) &= 0,2 \times 0,18 = 0,036 \\ (\text{TF}(\text{"A"}, D2) \times \text{IDF}(\text{"A"})) &= 0 \times 0,18 = 0 \\ (\text{TF}(\text{"A"}, D3) \times \text{IDF}(\text{"A"})) &= 0,33 \times 0,18 = 0,06 \end{aligned}$$

• TF(i, D): número de vezes que o termo i aparece em relação ao total de termos do texto D.

• IDF(i): log do número de textos no corpus dividido pelo número de textos que o termo i aparece.

D1:	Este é um exemplo A.
D2:	Este é um mostruário.
D3:	Este é outro A, exemplo A.

- TFxIDF igual a zero indica termo não relevante.
- TFxIDF maior indica termo mais relevante.

Questão 08 (4 pontos)

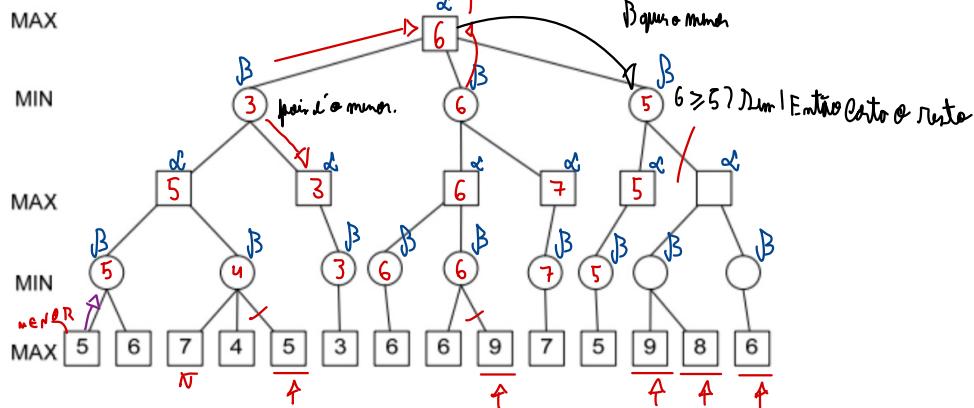
Considere a árvore minimax abaixo, representando um jogo onde queremos maximizar o valor da função de avaliação estática:

4/8

A pergunta deve ser se é maior ou igual a B. De fato, o valor de B está certo e pode cortar travesseiros para o que é maior.

$\text{MIN} = B$

$\text{MAX} = \infty$



Assinale a alternativa que apresenta a quantidade de folhas que não deverão ser visitados em uma busca da melhor jogada se a estratégia de poda alfa-beta for utilizada.

Importante: Resposta sem marcação na árvore não será considerada. Assim, favor anotar os valores alfa, beta e os cortes realizados na árvore acima.

- 5
- b) 8
- c) 9
- d) 10
- e) 11

Questão 06 (2 pontos)

Considerando o seguinte **Corpus**:

T1: Estamos chegando ao final do semestre. Que Alegria!

T2: Devemos estar felizes! Afinal de contas, foi um semestre pesado! Mas, calma! As férias estão chegando! Que Alegria! Que Alegria!

T3: Em agosto a gente vai ser ver novamente! Beleza?

T4: E que Alegria! Vocês acabam de finalizar metade do curso de Computação. Parabéns!

A frequência **TF x IDF** da palavra **Alegria**, dos quatro textos do Corpus são respectivamente:

Importante:

1. É necessário apresentar os cálculos. Questões sem os cálculos apresentados, serão desconsideradas.
2. Considere 3 casas decimais para os cálculos

$$TF = 1/8 = 0,125$$

$$TF = 2/10 \approx 0,1$$

$$TF = 0/9 = 0$$

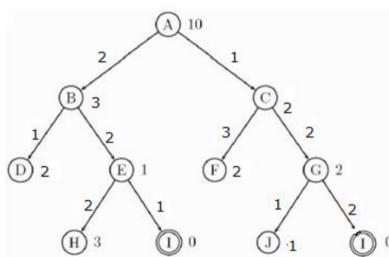
$$TF = 1/13 = 0,07$$

$$IDF = 4/3 = 1,33$$

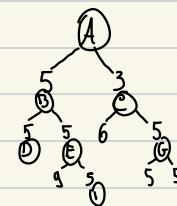
- a) 0,016 – 0,012 – 0,00 – 0,010
 b) 0,010 – 0,012 – 0,00 – 0,016
 c) 0,010 – 0,012 – 0,016 – 0,00
 d) 0,010 – 0,012 – 0,00 – 0,001
 e) 0,018 – 0,015 – 0,00 – 0,001

Questão 07 (5 pontos)

Considere o espaço de busca a seguir. Cada nó é rotulado por uma letra. O nó objetivo é representado por um círculo duplo. Existe uma heurística estimada para cada dado nó (indicada por um valor ao lado do nó). Arcos representam os operadores e seus custos associados.



A-C-B-D-E-G-I



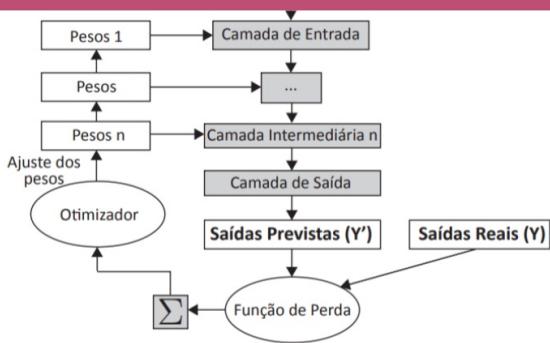
Para o algoritmo **A***, pede-se os nós visitados na ordem em que eles são examinados, começando pelo nó A.

No caso de escolhas equivalentes entre diferentes nodos, prefira o nodo mais próximo da raiz, seguido pelo nodo mais à esquerda na árvore. O algoritmo pára a busca quando encontra o nó I.

Importante:

- 1) Você precisa mostrar as análises feitas para responder à questão. Ou seja, resposta sem justificativa não será considerada.

- a) A, C, G, I
 b) A, C, B, D, E, G, I
 c) A, C, B, E, I
 d) A, B, D, E, I
 e) A, C, B, D, E, I



CHOLLET, F. Deep Learning with Python. New York: Manning Publications, 2017 (adaptado).

Após uma tentativa, notaram-se duas dificuldades: 1) o tempo de treinamento da rede estava muito longo e 2) a acurácia da rede treinada não estava no patamar aceito pela empresa.

Diante deste contexto, avalie as afirmações a seguir:

I. Aumentar o número de camadas é uma alternativa que pode levar a uma melhora na acurácia, além de diminuir o tempo de treinamento.

I aumenta o tempo de treinamento

2/8

II. Fazer uso de redes convolucionais é uma alternativa que pode levar a uma melhora na acurácia, no entanto, pode exigir uso de máquinas com maior poder de processamento.

III. Aumentar o número de unidades de processamento (neurônios) nas camadas pode levar a uma piora na acurácia, além de diminuir o tempo de treinamento da rede.

II aumento

IV. Aumentar o número de amostras de treinamento é uma alternativa que pode levar a uma melhora na acurácia, apesar de aumentar o tempo de treinamento da rede.

IV aumentando as/ as

V. Fazer uso de redes recorrentes é uma alternativa que pode levar a uma melhora na acurácia, no entanto, pode exigir uso de máquinas com maior poder de processamento.

É correto apenas o que se afirma em:

- a) I e IV
- b) I e V.
- c) II e III.
- d) II e IV.
- e) III e V

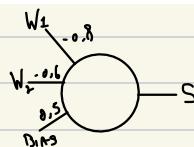
Questão 09 (3 pontos)

Dado um perceptron simples de duas entradas e um bias, cujos pesos são $w_1 = -0,8$, $w_2 = -0,6$ e $w_{bias} = 0,5$, respectivamente, pede-se:

Qual a função booleana que este perceptron realiza?

- a) AND
- b) OR
- c) NAND
- d) NOR
- e) Nenhuma das alternativas

Importante: Você deve apresentar a solução da questão na prova. Ou seja, resposta sem justificativa não será considerada. **Coloquei aqui a sua solução.**



w_1	w_2	s_p	s_f	OR	NOR
0	0	0,5	1	0	1
0	1	-0,5	0	0	0
1	0	-0,3	0	0	0
1	1	-0,9	0	1	0

$\triangleright 0 \times -0,8 + 0 \times -0,6 + 1 \times 0,5 = 0,5$

* Qual a negação da frase?

A) a) Existem alunos de gostam de IA
→ todos os alunos n gostam de IA

B) b) todo aluno gosta de IA
→ Existe aluno que n gosta de IA

C) c) nenhum aluno gosta de IA
→ algum aluno gosta de IA

$$\neg(A \wedge B)$$

$$\neg A \vee \neg B$$

A

Questão 03

QUESTÃO 12 – Considere a seguinte afirmação: "Há uma sorveteria onde todos os sorvetes são doces, mas não contém adoçantes." B

A negação da afirmação acima é logicamente equivalente à afirmação:

- A) Não há sorveteria que faz sorvetes doces e com adoçantes.
- B) Há uma sorveteria em que sorvete algum é doce ou contém adoçante.
- C) Em toda sorveteria, há um sorvete que não é doce, mas contém adoçante.
- D) Em toda sorveteria, há sempre algum sorvete que não é doce ou que contém adoçante.
- E) Há uma sorveteria em que há algum sorvete que não é doce ou que contém adoçante.

Questão 01 (1 ponto)

O processamento de linguagem natural pode ser entendido como uma subárea da inteligência artificial que estuda a compreensão automática de línguas naturais.

Nesse contexto, pode-se afirmar que as seguintes tarefas são utilizadas em problemas de processamento de linguagem natural:

- a) Stopwords e segmentação semântica
- b) Quantização e Part-of-Speech
- c) Extração de entidades e tokenização
- d) Bigramas e memórias recorrentes
- e) Espectrogramas e realce

Questão 02 (1 ponto)

Sabendo que a principal tarefa de um sistema será de classificação em **domínios complexos**, um gerente de projetos precisa decidir como vai incorporar essa capacidade em um sistema computacional a fim de torná-lo inteligente.

Existem diversas técnicas de inteligência computacional/artificial que possibilitam isso.

Nesse contexto, a técnica de **inteligência artificial** mais indicada para o gerente é:

- a) Lógica nebulosa.
- b) Árvores de decisão.
- Redes neurais artificiais. → p/ resolver problemas complexos,
- d) Regras de associação.
- e) Algoritmos genéticos

Questão 03 (1 ponto)

Dentre os algoritmos para mineração de dados, um exemplo de algoritmo para o particionamento de dados pelo aprendizado não supervisionado, que não usa uma amostra de treinamento pré-classificada, é denominado algoritmo de:

↳ Agrupamento

- a) Crescimento padrão frequente
- Agrupamento k-means
- c) Regra de Associação

1/8

- d) Redes Neurais
- Rede neural convolucional

Questão 04 (1 ponto)

Avalie a seguinte afirmação com Verdadeiro ou Falso no contexto das **Redes Neurais Convolucionais (CNN)**.

Afirmiação: A função de ativação **RELU**, muito utilizada em Redes Neurais Convolucionais, é aplicada a cada pixel da imagem e substitui todos os valores de pixels positivos por zero.

- a) Verdadeiro
- Falso

Caso a afirmativa seja falsa, justifique aqui: *Substitui os negativos por 0 e NÂO é aplicado a cada pixel da imagem*

Questão 11 (1 ponto)

Considere a seguinte proposição:

Em todos os cursos de Computação, existe, pelo menos, uma disciplina de Lógica.

A **negação** da proposição acima é logicamente equivalente à proposição:

- a) Em nenhum curso de Computação, há alguma disciplina de Lógica.
- ✗ b) Há, pelo menos, um curso de Computação no qual não há disciplina de Lógica.
- c) Em cada um dos cursos de Computação, não há disciplina de Lógica.
- d) Não há curso de Computação no qual tenha disciplina de Lógica.
- e) Há um curso de Computação no qual há, no máximo, uma disciplina de Lógica.

NEGACÃO

$$\neg(P \wedge Q) = \neg P \vee \neg Q$$

Ex

Bruno é professor e Bruno sabe matemática = Bruno não é professor e Bruno não sabe matemática.

→ Se/então

$$\neg(P \rightarrow Q) = P \wedge \neg Q$$

Se Bruno vai à praia então Bruno se molha = Bruno vai à praia e Bruno não se molha.

$$P \rightarrow Q = \neg P \vee Q$$

Se 2 é par então 2+1 é ímpar = Se 2+1 é par então 2 é ímpar

$$\neg(P \leftarrow Q) = P \vee \neg Q$$

$$\neg(P \leftarrow Q) = (P \wedge \neg Q) \vee (\neg P \wedge Q)$$

Bruno é carioca se e somente se 1 não é primo

(Bruno é carioca ou 1 não é primo)

Bruno é carioca e 1 é primo ou 1 não é primo e Bruno não é carioca

Questão 02

QUESTÃO 11 – Sejam m , n , p , q e r proposições lógicas tais que p é falsa e a proposição composta $((m \rightarrow n) \wedge (n \rightarrow p) \wedge (p \rightarrow q) \wedge (q \rightarrow r))$ é verdadeira, qual preposição abaixo é necessariamente verdadeira?

- A) $n \rightarrow r$
- B) $m \wedge r$
- C) $q \rightarrow n$
- D) $m \vee r$
- E) $r \rightarrow q$

QUESTÃO 15 – Considere as seguintes premissas sobre os alunos de uma universidade:

- I. Algum aluno que é estagiário não recebe bolsa.
- II. Todos aqueles alunos que estão no último período recebem bolsa.

Portanto,

- A) algum aluno do último período é estagiário.
- B) todos os alunos do último período não são estagiários.
- C) algum aluno que é estagiário não está no último período.
- D) algum aluno do último período não é estagiário.
- E) todos os alunos que são estagiários não estão no último período.

QUESTÃO 21

Na lógica proposicional, definem-se regras para determinar o valor-verdade (verdadeiro ou falso) de sentenças em relação a um modelo particular. Essas regras permitem representar raciocínios lógicos comuns das linguagens naturais.

Nesse contexto, considere a sentença e as proposições lógicas a seguir.

“Um veículo que é elétrico (E) pode ser um robô (R) se for autônomo (A), caso contrário não é um robô (R)”.

$$P1 = (E \wedge R) \leftrightarrow A;$$

$$P2 = E \rightarrow (R \leftrightarrow A);$$

$$P3 = E \rightarrow ((A \rightarrow R) \vee \neg R).$$

A sentença pode ser representada pela(s) expressão(ões) lógica(s)

- A) P2, apenas.
- B) P3, apenas.
- C) P1 e P2, apenas.
- D) P1 e P3, apenas.
- E) P1, P2 e P3.