

LISTA I-APRENDIZADO DE MÁQUINA

Nome: Guatano Costa

01- Leia o artigo 'A Survey on Machine Learning Approaches and Its Techniques' que está no CANVAS e responda as seguintes questões:

a) Explique o que é aprendizado supervisionado, não supervisionado, semi-supervisionado e reforço

I- Aprendizado supervisionado é quando no conjunto de dados cada instância já possui uma correspondência no conjunto de variáveis dependente classificadas (y), em outras palavras, cada instância possui um rótulo. Portanto, algoritmos supervisionados são aplicados em tarefas de classificação e regressão.

II- Aprendizado não supervisionado em contrapartida ao supervisionado, não possui rótulos para as instâncias, tomando o algoritmo responsável por entender os dados e separá-los em conjuntos diferentes baseados nas features de cada instância.

III- Quando falamos de dados, em certas situações, nem sempre é viável fornecer ao algoritmo dados rotulados, seja pelo volume massivo de instâncias ou porque não é bem custoso. Para essa caso, o aprendizado semi-supervisionado é bem útil, primeiro ele usará o aprendizado não supervisionado para separar as instâncias em diferentes grupos e posteriormente aplicar técnicas supervisionadas para tarefas de classificação, predição.

IV- Aprendizado por Reforço é quando o algoritmo que será nomeado para 'Agente' é inserido em um ambiente com diversas formas de tomar

Jandaia

decisão até atingir um objetivo final. Para cada estratégia tomada, que é chamada também de 'política', o agente recebe recompensas positivas ou negativas a cada decisão tomada. No final, o algoritmo deve ter encontrado a política que maximiza as recompensas positivas.

b) dê exemplos de algoritmos de aprendizado pertencentes a cada um destas tipos de aprendizado

I- Aprendizado supervisionado: Regressão Linear, Regressão Logística

II- Aprendizado não supervisionado: K-means, PCA, DBSCAN

III- Aprendizado Semi supervisionado: Label Propagation, Self-Training

IV- Aprendizado por Reforço: Deep Q-Network, Q-Learning

02- Explique o significado das seguintes tipos de problemas de aprendizado de máquina e forneça pelo menos 1 exemplo para cada tipo de problema

a) Classificação

Problemas de Classificação consistem em conseguir classificar coisas em uma determinada classe, por exemplo, classificar mensagens de e-mail recebidas como um spam ou como não spam. Outro exemplo seria classificar imagem de animais como sendo cachorros ou gatos.

b) Regressão

Problemas de Regressão envolvem a previsão de valores numéricos contínuos, i.e. tentar prever um valor que uma variável assumirá. Por exemplo, qual é o preço de um imóvel de 150 m², 3 quartos e 3 banheiros no Jandaia? Ou ainda, qual é o valor esperado para o salário de um contador de dados que trabalha em uma empresa de porte médio?

c) Agrupamento

Problemas de agrupamento é quando temos como objetivo agrupar dados semelhantes entre si com o objetivo de aprender mais padrões, similaridades, etc. Um caso comum é realizar uma segmentação de clientes para fazer marketing mais efetivo para cada tipo de problema, outro caso comum é o uso de processamento de linguagem natural para agrupar documentos de mesmo tema.

d) Associação

Problemas de associação geralmente envolvem inferir novas informações baseados em outros conhecimentos prévios. Por exemplo, dado que uma pessoa compra feijão e talco, qual a probabilidade dela comprar leite junto?

03 - Considere a seguinte base de dados, em que se deseja analisar se o consumo é comestível ou não

TABELA

Rede - 288

a) Decida qual regra o algoritmo PRISM gerará a partir dessa base de dados

I - Fixei a classe COMESTIVEL = 1ea

Forma	Abstração	Ode
'C'	2	Ruim 3
'D'	3	Bom 2
		Ótimo 0

Cor	
'B'	3
'W'	2
'G'	0
'U'	0

Se Forma = 'D' Então é Comestível

D	S	T	Q	Q	S	S
D	L	M	M	J	V	S

→ Se forma = D então é comestível

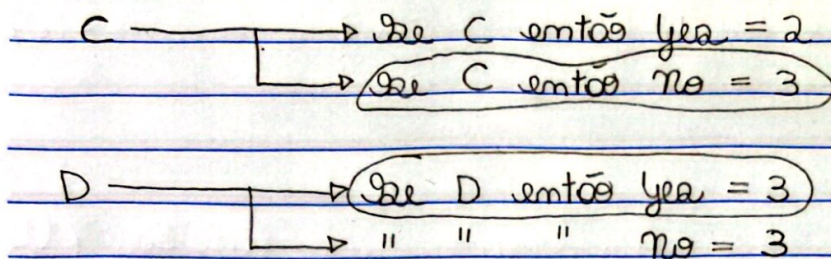
Por		Oder	
B	1	Rum	2
W	2	Bom	1
G	0	Time	0
U	0		

Se forma = 'D' e Por = 'W' então é comestível

Oder		
Rum	1	Se forma = 'D' e Por = 'W' e Oder = 'Bom' então
Bom	1	é comestível //
Time	0	

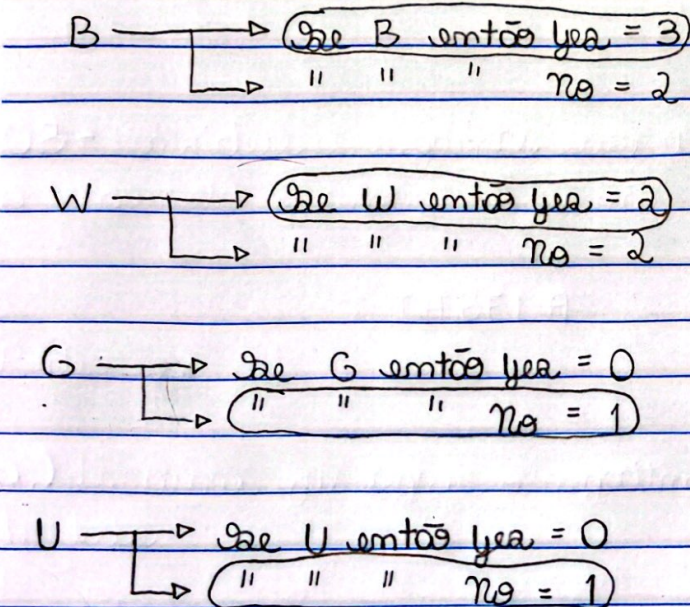
b) descreva que regras que o algoritmo ONE R gerará a partir dessa base de dados

I-Forma



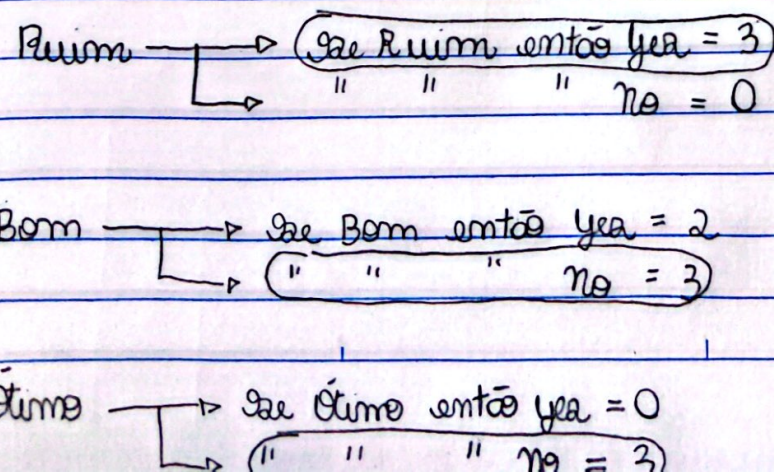
Forma	Euro	Euro Total
C	$\frac{2}{5}$	
D	$\frac{3}{6}$	$\frac{5}{11}$

II- Cor



Cor	Euro	Euro Total
B	$\frac{2}{5}$	
W	$\frac{2}{4}$	4
G	0	11
U	0	

III- Odor



Odor	Euro	Euro Total
Ruim	0	
Bom	$\frac{2}{5}$	$\frac{2}{11}$
Ótimo	0	

Regra One R
memoriza