

1 Questões Teóricas

1. O que é aprendizado de máquina?

Aprendizado em máquina é um campo de estudo da inteligência artificial que estuda principalmente a criação de modelos computacionais que recebe um conjunto de dados como entrada e com base neles busca inferir uma certa informação.

A criação desse tipo de modelo é feita por meio de um algoritmo de aprendizado de máquina, que são algoritmos que recebem um conjunto de dados de treinamento como entrada e retorna um modelo que estabelece suas inferências com base nos dados de treinamento.

2. Descreva os seguintes tipos de aprendizado

- **Aprendizado supervisionado.**

O aprendizado supervisionado se caracteriza pelo fato de que os dados de treinamento incluem também o valor da variável a ser inferida, dessa forma o modelo iterativamente, durante seu treinamento, considerando os dados de entrada infere valores para a variável de saída, compara com seu valor real e realiza os ajustes necessários na tentativa de melhorar o modelo.

O aprendizado supervisionado é tipo de aprendizado em que as informações necessárias para o treinamento do modelo são conhecidas durante o treinamento do mesmo.

- **Aprendizado não supervisionado**

Nesse tipo de aprendizado o modelo não possui os valores referentes a variável de seu modelo e nenhum tipo de feedback que parte dessa variável, sendo assim o modelo busca identificar informações no conjunto de dados por conta própria.

- **Aprendizado por reforço (semi-supervisionado).** Funciona de forma similar ao aprendizado por reforço, porém não possui os valores exatos para os seus valores de saída. Dessa forma o agente aprende a partir de um feedback do ambiente, que não indica exatamente a ação que o agente deveria tomar.

- **Aprendizado online.**

Nesse tipo de aprendizado o agente é capaz de aprender mesmo depois da sua etapa de treinamento, podendo receber feedbacks durante seu funcionamento e ir ajustando seu modelo.

- **Aprendizado offline.**

O agente possui uma etapa de treinamento bem definida, ou seja, uma vez que o modelo é treinando ele somente recebe entradas e retorna saídas sem alterar o seu modelo.

3. Para cada tipo de aprendizado listado abaixo, apresente DUAS aplicações.

- **Aprendizado supervisionado.**

Classificação de Objetos e criar uma função que busca estimar uma função real.

- **Aprendizado não supervisionado**

Em técnicas de agrupamento de dados e redução de dimensões.

- **Aprendizado por reforço (semi-supervisionado).**

Em IA em jogos e Robótica.

- **Aprendizado online.**

Chatbots e Automação Residencial

- **Aprendizado offline.**

Reconhecimento de Objetos e Estimar uma função.

4. Para cada uma das métricas de avaliação de modelos abaixo apresente, (i) a definição matemática, (ii) casos de uso (em que situações devem ser utilizadas), (iii) vantagens e desvantagens.

- **Erro 0/1**

É uma métrica binária que indica acerto(1) ou erro(0).

Vantagem: Simples e eficiente em alguns casos.

Desvantagem: Não expressa o quão distante as observações estão dos valores desejados.

- **Erro absoluto.**

Indica as distâncias entre os valores desejados e os valores obtidos pelo agente.

Vantagem: Capaz de quantificar o quão distante uma predição está do valor real.

Desvantagem: Por esse erro ser modelado a partir de um função descontínua, não é possível utilizar o método do gradiente para minimizar-lo.

- **Erro quadrado médio.**

Indica o quadrados das distâncias entre os valores desejados e os valores obtidos pelo agente.

Vantagem: É possível utilizar o gradiente para minimizar e criar modelos que minimizam essa métrica.

Desvantagem: Pode ter ainda mais erros grandes e menos erros pequenos.

- **Erro de pior-caso.**

Indica qual o maior erro registrado.

Vantagem: Muito útil para casos em que é desejável manter todos os erros contidos em intervalos conhecidos.

Desvantagem: Expressa de forma limitada a eficiência do modelo.

- **Verossimilhança (Likelihood).**

Define a distância entre o valor de saída do agente e o valor real considerando a probabilidade.

Vantagem: Pode ser aplicado em problemas de classificação.

Desvantagem: Por sua equação de minimização ser definido como um produto de números decimais é difícil de ser computado.

- **Log-verossimilhança (log-likelihood)**

É definida pela aplicação da função de log à verossimilhança e que preserva as relações de ordem entre as funções a partir das quais são definidas.

Vantagem: Sua função de minimização pode ser representada como um somatório.

Desvantagem: Aplicável a modelos que lidam com probabilidade.

5. **Qual a diferença entre um problema de regressão e um problema de classificação ?**

O problema de regressão busca inferir valores contínuos enquanto o de classificação valores discretos para os dados.

6. **Explique o que significa dizer que "duas classes são linearmente separáveis".**

São classes que podem ser separadas por um hiperplano definido pelas variáveis de entrada do agente. Pode ser adotada como um critério por alguns métodos.

7. **Por que dividimos erros em falso positivos e falso negativos? Esta divisão faz sentido para problemas de regressão? Explique.**

Pois em alguns casos erros por falso positivos podem ser mais prejudiciais e vice-versa.

8. Considere a seguinte base de dados:

<i>Example</i>	<i>Author</i>	<i>Thread</i>	<i>Length</i>	<i>Where_read</i>	<i>User_action</i>
e_1	known	new	long	home	skips
e_2	unknown	new	short	work	reads
e_3	unknown	followup	long	work	skips
e_4	known	followup	long	home	skips
e_5	known	new	short	home	reads
e_6	known	followup	long	work	skips
e_7	unknown	followup	short	work	skips
e_8	unknown	new	short	work	reads
e_9	known	followup	long	home	skips
e_{10}	known	new	long	work	skips
e_{11}	unknown	followup	short	home	skips
e_{12}	known	new	long	work	skips
e_{13}	known	followup	short	home	reads
e_{14}	known	new	short	work	reads
e_{15}	known	new	short	home	reads
e_{16}	known	followup	short	work	reads
e_{17}	known	new	short	home	reads
e_{18}	unknown	new	short	work	reads
e_{19}	unknown	new	long	work	?
e_{20}	unknown	followup	short	home	?

- Apresente uma árvore de decisão para a classificação das User-actions.
 - De acordo com a árvore apresentada, qual a classificação dos exemplos e_{19} e e_{20} ?
 - Mostre como a árvore de decisão proposta pode ser representada como um conjunto de cláusulas definidas.
9. . Uma abordagem comum para treinamento de modelos em aprendizado supervisionado é formular o problema de aprendizado como um problema de otimização. Estes problemas, por sua vez, podem ser resolvidos com algoritmos de descida de gradiente. Neste contexto:
- Apresente a formulação matemática para o problema de minimização do erro quadrado de um modelo de regressão

linear.

$$f = \sum_{i=1}^m (\hat{y}_i - (w_0 + \sum_{j=1}^n x_{ij} * w_j))^2$$

m = numero de amostras.

n = numero de variáveis de entrada

- Derive o gradiente da função objetivo definida em (a) em relação aos pesos (w_{i_s}) do modelo.

$$\frac{df}{dw_k} = \sum_{i=1}^m 2 * (\hat{y}_i - (w_0 + \sum_{j=1}^n x_{ij} * w_j)) * (x_{ik})$$

- Apresente a formulação matemática para o problema de minimização do erro quadrado de um modelo de regressão logística (modelo que utiliza a função sigmoide).

$$g = \sum_{i=1}^m \left(\hat{y}_i - \frac{1}{1 + e^{w_0 + \sum_{j=1}^n x_{ij} * w_j}} \right)^2$$

- Derive o gradiente da função objetivo definida em (c) em relação aos pesos (w_{i_s}) do modelo.

$$\frac{dg}{dw_k} = \sum_{i=1}^m 2 \left(\hat{y}_i - \frac{1}{1 + e^{w_0 + \sum_{j=1}^n x_{ij} * w_j}} \right) * \frac{e^{w_0 + \sum_{j=1}^n x_{ij} * w_j}}{\left(1 + e^{w_0 + \sum_{j=1}^n x_{ij} * w_j} \right)^2} * (x_{ik})$$