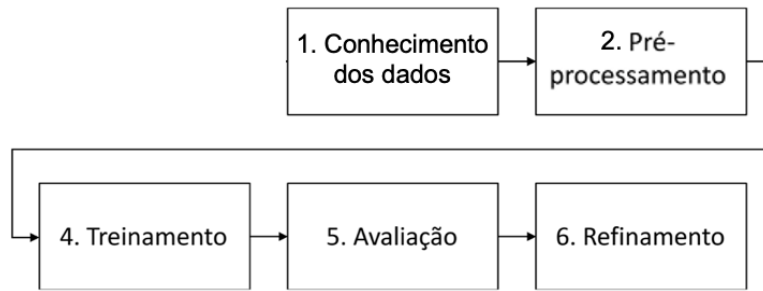


Lista de Exercícios

Introdução ao aprendizado de máquina

- 1) Defina as palavras “aprendizado”, “máquina” e “aprendizado de máquina” de forma breve, porém concisa.
- 2) O que nos motiva a querer um agente para aprender?
- 3) Segundo Tom Mitchell, diz-se que um computador **aprende** com uma experiência E ao realizar uma classe de tarefas T cujo desempenho é medido por P, **se** seu desempenho P nas tarefas T aumenta com a experiência E. Você acha essa definição suficiente para aprendizado de máquina? Justifique.
- 4) Defina aprendizado supervisionado, não supervisionado e por reforço. Qual o tipo de problema que cada um pode resolver? Que tipo de informação é necessária à priori?
- 5) Escreva sobre um problema real, que esteja presente no seu dia a dia, ao qual você possa modelar através de aprendizagem de máquina.
- 6) Classifique as tarefas abaixo em aprendizado supervisionado, aprendizado não supervisionado e aprendizado por reforço. Justifique seu raciocínio em cada item.
 - a) Prever o diagnóstico de diabetes em um paciente com base em testes clínicos.
 - b) Separar os usuários do Netflix em grupos similares para definir recomendações mais direcionadas.
 - c) Prever o consumo energético de um circuito integrado.
 - d) Aprender as ações mais adequadas para um veículo autônomo trabalhando para a Uber.
 - e) Agrupar clientes de uma empresa para definir estratégias de marketing específicas.
 - f) Classificar imagens de acordo com os objetos presentes nelas.
 - g) Detectar resultados de exames que estão muito além do esperado.
 - h) Prever a probabilidade de chuva de acordo com as medições meteorológicas.
- 7) Para cada tarefa do exercício 6, descrever possíveis entradas X e a saída Y (quando houver) que você acha adequado.
- 8) Diferencie os conceitos de *overfitting* e *underfitting*. De acordo com a literatura, qual deles acontece com mais frequência? Justifique sua resposta.
- 9) Como detectar ou ao menos considerar que houve *overfitting*? Quais os tipos de técnicas que podem ser aplicadas no treinamento para reduzir este fenômeno?
- 10) A figura abaixo ilustra um possível fluxo de implementação para sistemas baseados em aprendizado de máquina:



- Descreva cada componente explicando sua entrada, sua funcionalidade, um exemplo de técnica e a saída de cada um.
- Os dados aplicados para treinamento e teste nos passos 4 e 5 são os mesmos? Justifique a sua resposta.
- Digamos que estejamos trabalhando com um problema de classificação para prever se imagens contém ou não um gato. Quais métricas você usaria para treinamento e avaliação?
- Digamos agora que a tarefa consiste em detectar casos de câncer de pele com base em imagens. Quais métricas você usaria para treinamento e avaliação?

Aprendizado Supervisionado

- Quais são as principais diferenças entre problemas de classificação e problemas de regressão? Cite um exemplo de cada um deles.
- Suponha que você esteja trabalhando em uma empresa de investimentos na previsão do mercado de ações e gostaria de prever o preço de uma determinada ação amanhã. Você deseja usar um algoritmo de aprendizado para isso. Você trataria essa tarefa como uma tarefa de classificação ou regressão? Justifique sua resposta.
- O problema de detecção de *spams* tem sido tratado de diversas formas nos últimos anos. Discuta um classificador para detecção de *spam* considerando dois cenários distintos:
 - Muitos dados de ambas as classes (*spam* e não *spam*).
 - Poucos dados.

Árvores de decisão

- Qual o princípio básico de funcionamento das árvores de decisão? O que nos leva a classificar um atributo como forte dentre os vários disponíveis em uma árvore de decisão? Dê um exemplo.
- Construa o conjunto completo de árvores de decisão a partir dos seguintes dados:

A1	A2	Y
0	1	0
1	0	1
0	0	0
1	1	1

- 16) Suponha que queremos explorar o algoritmo de árvores de decisão com base no conjunto de dados apresentados na seguinte tabela:

A1	A2	A3	Y
0	1	1	1
1	0	1	1
0	0	0	0
1	1	1	0

- Calcule o ganho de informação de cada um dos atributos.
 - Qual atributo será escolhido para dividir os dados? Justifique.
 - Construa a árvore de decisão.
- 17) Os avaliadores de novos alunos de mestrado se baseiam em quatro critérios para aceitar (A) ou rejeitar (R) o ingresso dos alunos:

Nota atingida: {3.5, 3.7, 4.0}

Ranking da universidade de graduação: {top-10, top-20, top-30}

Houve publicação de artigos: {sim, não}

Qualidade da carta de recomendação: {boa, normal}

Apresente a árvore de decisão para prever se um aluno será aceito ou não no curso de mestrado.

Nota	Ranking	Publicou	Recomendação	Classe
4.0	top-10	sim	boa	A
4.0	top-10	não	boa	A
4.0	top-20	não	normal	A
3.7	top-10	sim	boa	A
3.7	top-20	não	boa	R
3.7	top-30	sim	boa	A
3.7	top-30	não	boa	R
3.7	top-10	não	boa	R
3.5	top-20	sim	normal	R
3.5	top-10	não	normal	R
3.5	top-30	sim	normal	R
3.5	top-30	não	boa	R

- 18) Uma pessoa, na sua apreciação de refeições em restaurantes, considera os seguintes atributos e respectivos valores possíveis:

Restaurante: {Copélia, Palma, Primavera}

Qualidade: {boa, má}

Preço: um valor inteiro

Refeição: {almoço, jantar, pequeno almoço}

Essa pessoa utiliza afirmações em português para exprimir o seu processo de classificação das refeições como satisfatórias ou não satisfatórias, ao invés de usar árvores de decisão. Indique árvores de decisão correspondentes a cada uma das seguintes afirmações:

- "Fico satisfeita com qualquer refeição de R\$ 50 ou menos, mas não existe nenhuma refeição do restaurante Palma por R\$ 30 ou menos que me agrade."
- "Fico satisfeita com um pequeno almoço de R\$ 30 ou menos ou com um jantar de R\$45 ou menos. Nunca fico satisfeito com uma refeição de má qualidade."

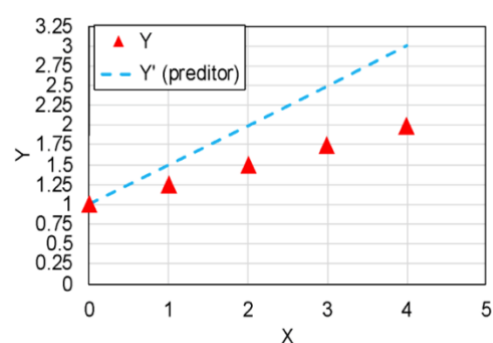
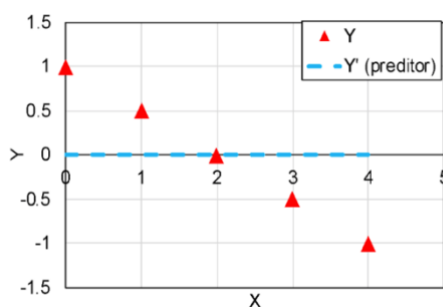
Regressão Linear

19) Sobre o algoritmo de regressão linear, responda:

- Qual a função do gradiente descendente? Ele é mesmo necessário?
- Qual a noção intuitiva por trás do funcionamento do algoritmo?
- O que acontece quando o gradiente é positivo, negativo e igual a zero?
- Qual o efeito da taxa de aprendizado (α ou β)?

20) As retas abaixo implementam modelos de regressão linear para os pontos de treino representados. Para cada reta apresente:

- O coeficiente linear e angular.
- O erro medido em Erro Quadrático Médio (MSE).



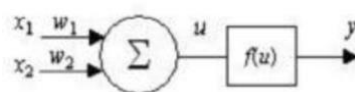
Redes Neurais

21) Em que consiste uma rede neural artificial? Qual sua composição básica? Quais classes de problemas da inteligência artificial podem ser utilizadas redes neurais?

22) Como funciona a aprendizagem do *perceptron*? É mais aconselhado utilizar uma taxa de aprendizagem mais alta ou mais baixa? Justifique sua resposta.

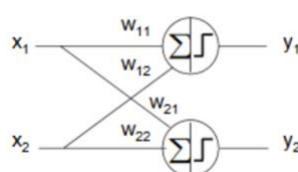
23) Disserte sobre a principal limitação do *perceptron* para a aprendizagem em redes neurais.

24) Dado o neurônio abaixo, determine valores para seus pesos, w_1 e w_2 , e para o limiar ϕ , tal que ele reproduza a porta lógica OR. Escolha um limiar entre 0.1 e 2.



x_1	x_2	$x_1 \text{ OR } x_2$
0	0	0
0	1	1
1	0	1
1	1	1

25) Sejam os neurônios abaixo com entradas x_1 e x_2 e saídas y_1 e y_2 . Determine as saídas para os valores indicados, supondo $w_{11}=1,66$; $w_{12}=0,83$; $w_{21}=-7$; $w_{22}=10$. Considere que ambos os neurônios têm limiar $\phi=1$.



x_1	x_2	y_1	y_2
0	0,5		
0,3	0,1		
0,7	0,3		
0,6	0,9		

Aprendizado não supervisionado e por reforço

- 26) Quais são as seis principais etapas do aprendizado não supervisionado? Explique brevemente cada uma delas. Por que é importante selecionarmos adequadamente os atributos?
- 27) Quais são as condições que precisam ser satisfeitas para formar um agrupamento no aprendizado não supervisionado? O que é esperado em relação ao conteúdo de cada um dos grupos formados?
- 28) Como que os algoritmos de agrupamento no aprendizado não supervisionado decidem se um dado elemento deve fazer parte de um determinado grupo ou de outro?
- 29) Qual o principal diferencial em relação ao treinamento do aprendizado por reforço? Qual a importância dos estados e das ações neste tipo de aprendizado? Como um agente determina se um estado é bom ou ruim?
- 30) Defina o termo “política de ações” no aprendizado por reforço. Dê um exemplo de um mapeamento de estado em ações.