

# Princípios de aprendizado de máquina

## Inteligência Artificial

Prof. Rodrigo Pedrosa

30 de Setembro de 2022

### 1 Questões teóricas

1. O que é aprendizado de máquina?
2. Descreva os seguintes tipos de aprendizado:
  - (a) Aprendizado supervisionado.
  - (b) Aprendizado não supervisionado.
  - (c) Aprendizado por reforço.
  - (d) Aprendizado semi-supervisionado.
  - (e) Aprendizado online.
  - (f) Aprendizado offline.
3. Para cada tipo de aprendizado listado abaixo, apresente DUAS aplicações.
  - (a) Aprendizado supervisionado.
  - (b) Aprendizado não supervisionado.
  - (c) Aprendizado por reforço
  - (d) Aprendizado semi-supervisionado.
  - (e) Aprendizado online.
  - (f) Aprendizado offline.
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.
  - (a) Erro 0/1.
  - (b) Erro absoluto.
  - (c) Erro quadrado médio.
  - (d) Erro de pior-caso.
  - (e) Verossimilhança (Likelihood).
  - (f) Log-verossimilhança (log-likelihood).
5. Qual a diferença entre um problema de regressão e um problema de classificação?
6. Explique o que significa dizer que "duas classes são linearmente separáveis".
7. Por quê dividimos erros em falso positivos e falso negativos? Esta divisão faz sentido para problemas de regressão? Explique.

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

- 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.
  - Derive o gradiente da função objetivo definida em (a) em relação aos pesos ( $w_{i_s}$ ) do modelo.
  - Apresente a formulação matemática para o problema de minimização do erro quadrado de um modelo de regressão logística (modelo que utiliza a função sigmoide).

## 2 Questões práticas

1. Tutorial<sup>1</sup> - Regressão Linear (<https://moodlepresencial.ufop.br/mod/url/view.php?id=704482>)

DATASET disponível em: ([https://raw.githubusercontent.com/rcpsilva/BCC325\\_ArtificialIntelligence/main/2022-1/Problem%20sets%20\(in%20Portuguese\)/datasets/candidatos.csv](https://raw.githubusercontent.com/rcpsilva/BCC325_ArtificialIntelligence/main/2022-1/Problem%20sets%20(in%20Portuguese)/datasets/candidatos.csv))

2. Tutorial<sup>1</sup> - Árvores de decisão ([https://drive.google.com/file/d/1DI2iVcy6jwT8Y\\_VZmF9xDqjl\\_C5GrqmQ/view?usp=sharing](https://drive.google.com/file/d/1DI2iVcy6jwT8Y_VZmF9xDqjl_C5GrqmQ/view?usp=sharing))

DATASET disponível em: ([https://raw.githubusercontent.com/rcpsilva/BCC325\\_ArtificialIntelligence/main/2022-1/%20sets%20\(in%20Portuguese\)/datasets/decision\\_tree.csv](https://raw.githubusercontent.com/rcpsilva/BCC325_ArtificialIntelligence/main/2022-1/%20sets%20(in%20Portuguese)/datasets/decision_tree.csv))

3. Tutorial<sup>1</sup> - Florestas Aleatórias ([https://drive.google.com/file/d/1SglH1MqJzaPHpjaQXo3ikJy6gC771H\\_f/view?usp=sharing](https://drive.google.com/file/d/1SglH1MqJzaPHpjaQXo3ikJy6gC771H_f/view?usp=sharing))

---

<sup>1</sup>Os tutorias só são acessíveis com o e-mail institucional