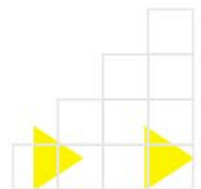


Curso 2 – CD, AM e DM

Profa. Roseli Ap. Francelin Romero

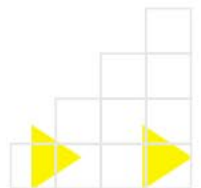
MBA em Inteligência Artificial e BigData

**Depto. de Ciências de Computação
ICMC - USP**



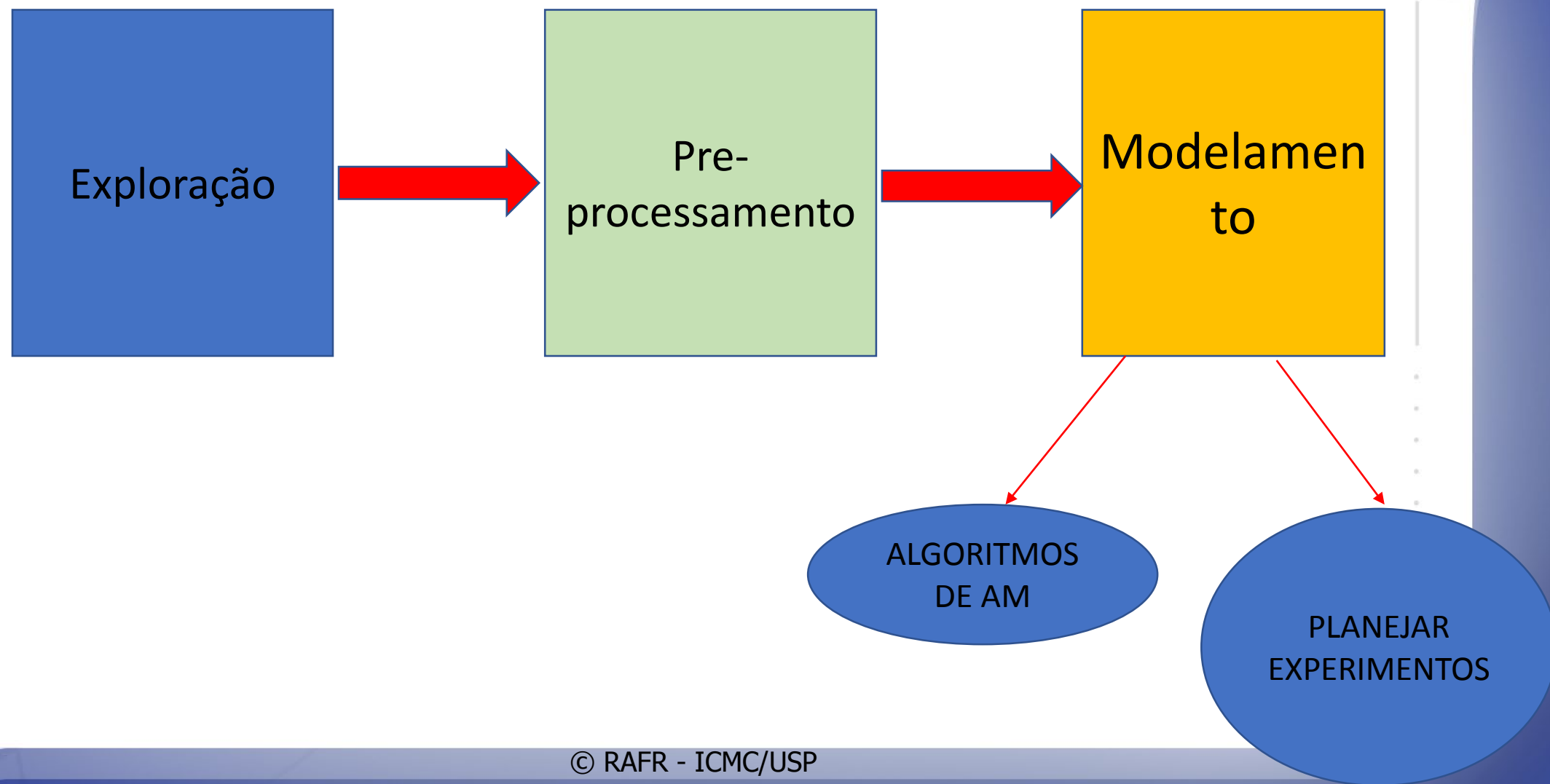
Semana 3

- Paradigmas de Aprendizado.
- Modelos preditivos.
- Partição dos dados.
- Reamostragem (Holdout, bootstrap, K-fold cross validation).
- Modelamento de Dados





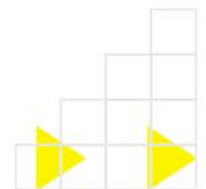
3º. ETAPA



APRENDIZADO DE MÁQUINA

- ÁRVORES DE DECISÃO
- RANDOM FOREST
- GRADIENTE BOOSTING MACHINES
- COMITÊS DE CLASSIFICADORES
- SVMs e KERNELS
- CLASSIFICADOR BAYES

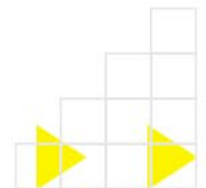
~~PERCEPTRON
MLP
DEEP LEARNING~~





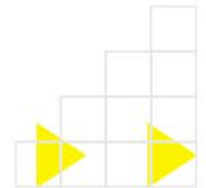
Sistemas de Aprendizado

- Os diversos sistemas de Aprendizado possuem características particulares e comuns que permitem sua classificação:
- Modo (on-line, off-line)
- Paradigma (supervisionado ou não-supervisionado)
- Forma de aprendizado utilizado (indutivo ou descritivo)



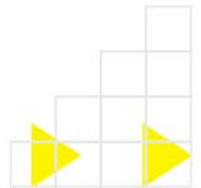
Sistemas de Aprendizizado

- AM é uma ferramenta poderosa para a aquisição automática de conhecimento,
- Porém, deve ser observado que **não existe** um único algoritmo que apresente o melhor desempenho para todos os problemas.



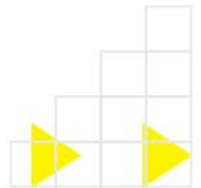
Sistemas de Aprendizado

- É importante compreender o poder e a limitação dos diversos algoritmos de AM
- Utilizar alguma metodologia que permita avaliar os conceitos induzidos por esses algoritmos em determinados problemas.



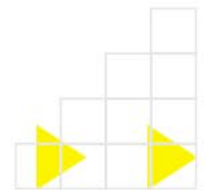
INDUÇÃO

- A indução é a forma de inferência lógica que permite obter conclusões genéricas sobre um conjunto particular de exemplos.
- Na indução, um conceito é aprendido efetuando-se inferência indutiva sobre os exemplos apresentados.
- Portanto, as hipóteses geradas podem ou não preservar a verdade.



INDUÇÃO

- Apesar disto, ela é um dos met. mais utilizados para derivar conhecimento novo e predizer eventos futuros.
- Exemplos:
- Arquimedes descobriu a 1a. Lei da hidrostática e o principio da alavanca
- Kepler descobriu as leis do mov. Planetário
- Darwin: as leis da seleção natural das espécies.



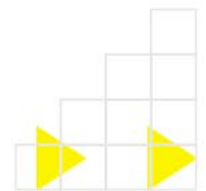
INDUÇÃO



- Cuidado:

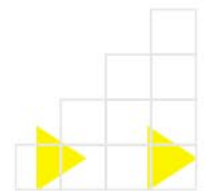
Apesar de ser o recurso mais utilizado nos algoritmos para obter conhecimento novo, temos que:

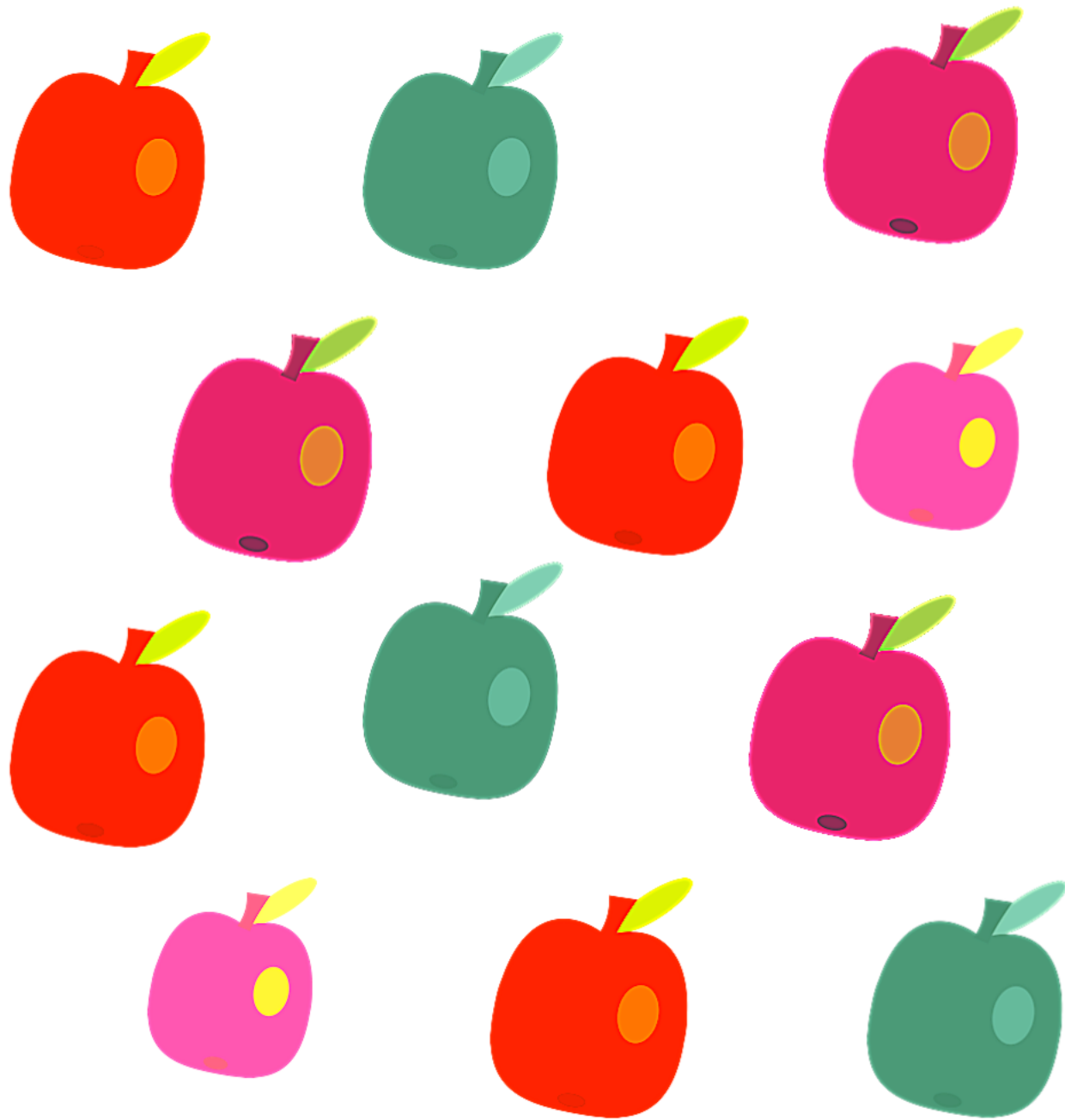
- se o no. de exemplos for insuficiente,
- se os exemplos não forem bem escolhidos,
- se os exemplos não refletirem as características do domínio ==> as hipóteses obtidas podem ser de muito pouco valor.



Tipos de Aprendizado Indutivo

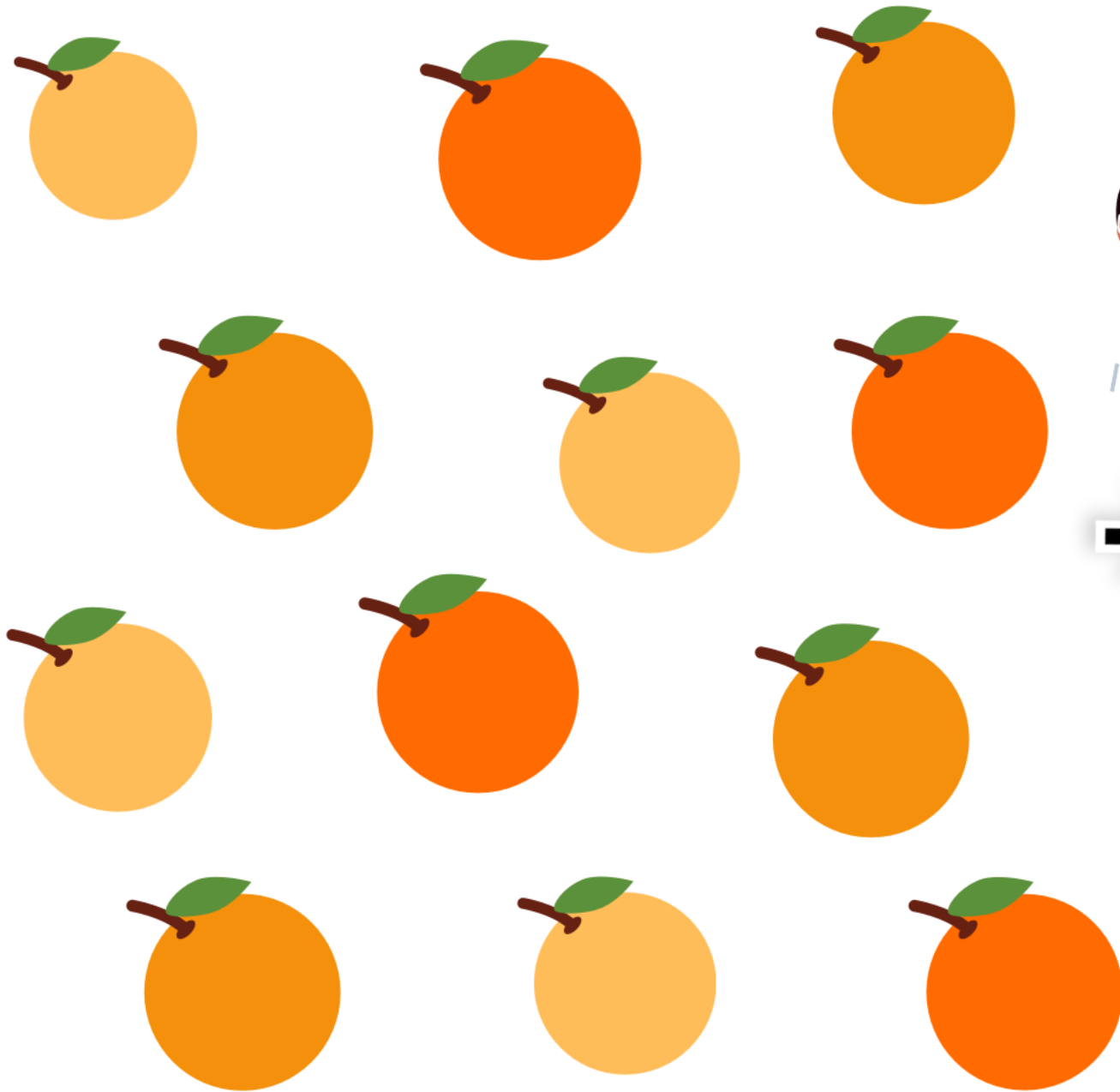
- O aprendizado indutivo pode ser:
Aprendizado Supervisionado
Aprendizado Não-supervisionado





Ensinado...

MAÇÃ



Ensinado...

LARANJA



Não ensinado...

×

?

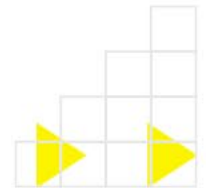
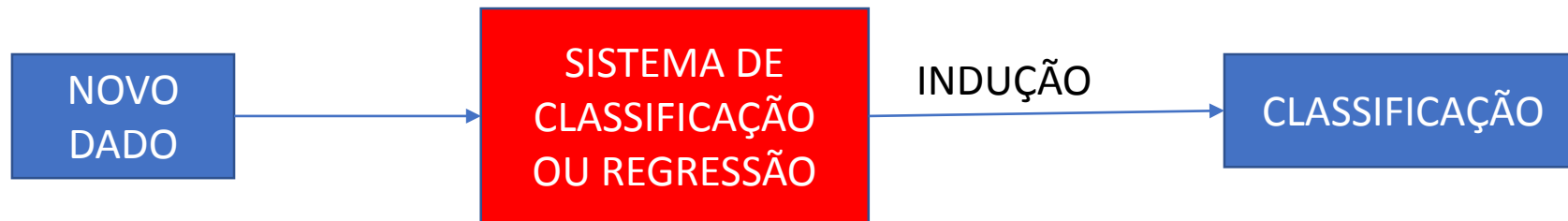


Aprendizado Indutivo

- Supervisionado: conjunto de exemplos de treinamento para os quais o rótulo da classe associada é conhecido.

Cada exemplo: vetor de características ou atributos e o rótulo da classe associada.

Objetivo: construir um classificador \Rightarrow classe a que pertence um novo exemplo ainda não rotulado.



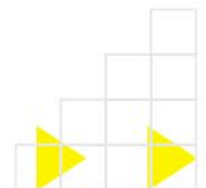
Aprendizado Indutivo

- **Aprendizado Supervisionado:**

Um conjunto de exemplos de treinamento para os quais o rótulo da classe associada é conhecido.

$\underline{\mathbf{x}}_1$	t_1	atributos+rótulo (pares de E/S)
$\underline{\mathbf{x}}_2$	t_2	
$\underline{\mathbf{x}}_3$	t_3	
$\underline{\mathbf{x}}_N$	t_N	

$\underline{\mathbf{x}}$ denota um vetor de p componentes



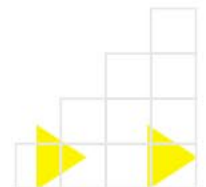
Aprendizado Indutivo

- **Aprendizado Não-supervisionado:**

- O indutor analisa os exemplos de treinamento e tenta determinar se alguns deles podem ser agrupados de algum modo formando **agrupamentos ou clusters**.
- Após este processo, uma análise dos resultados é necessária para determinar o que cada agrupamento significa.

\underline{x}_1	atributos (Entradas)
\underline{x}_2	
\underline{x}_3	
...	
\underline{x}_N	

\underline{x} denota um vetor de p componentes

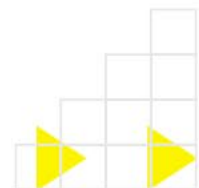


APRENDIZADO NÃO SUPERVISIONADO



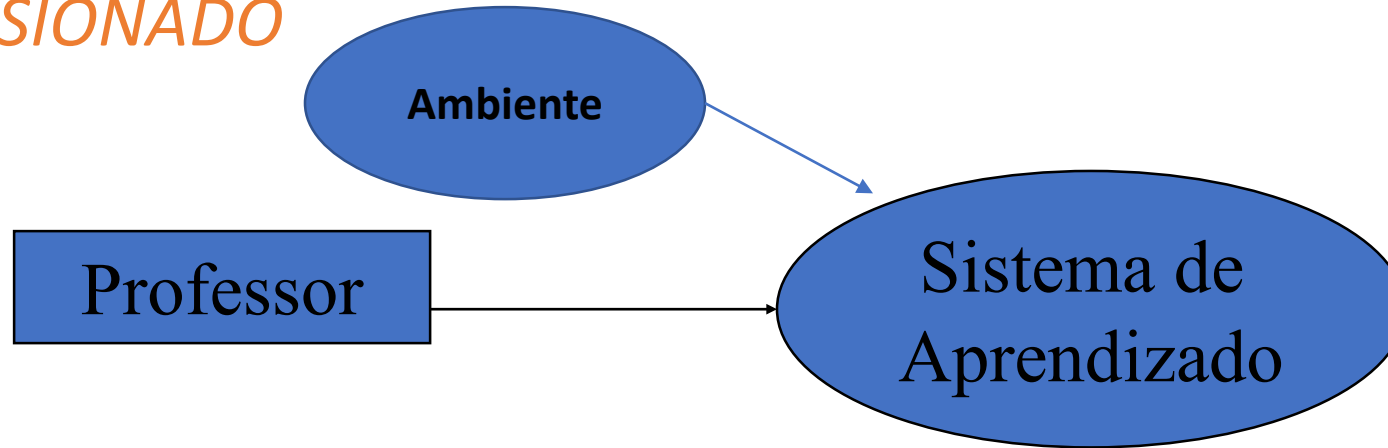
NÃO EXISTE O PROFESSOR

EXTRAIR AS
CARACTERÍSTICAS DOS
DADOS

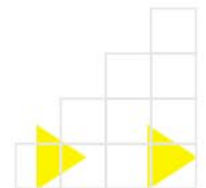
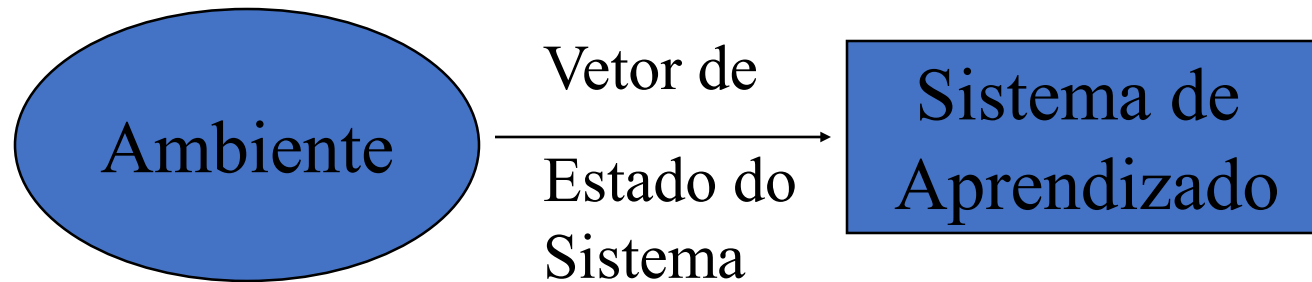


Paradigmas de Aprendizado

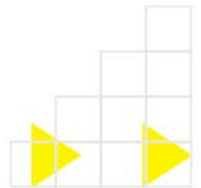
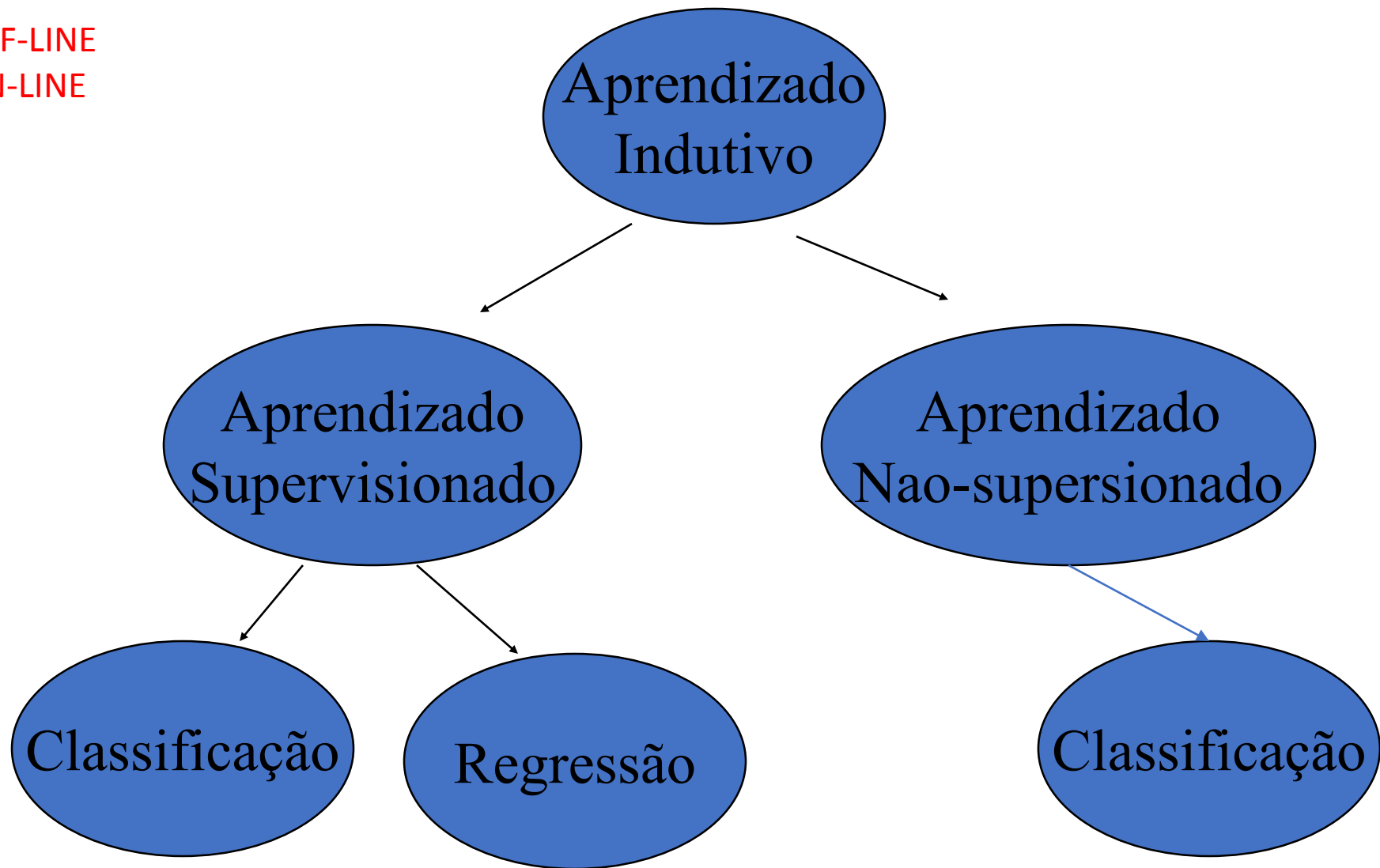
- *SUPERVISIONADO*



- *NAO-SUPERVISIONADO*



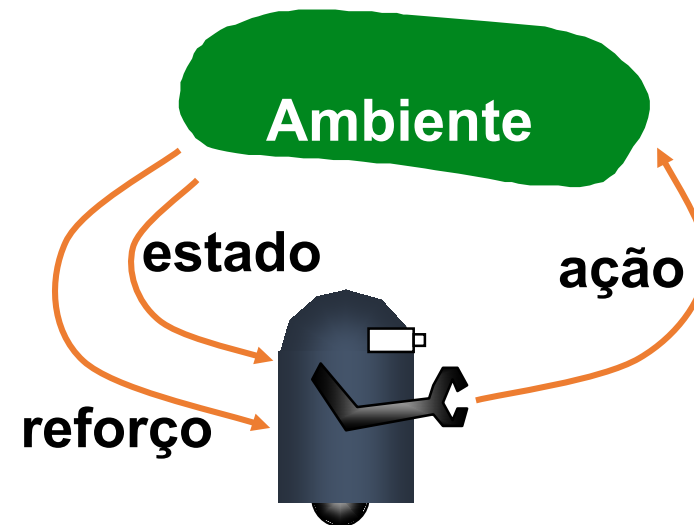
Modo: OFF-LINE
Modo: ON-LINE



Paradigmas de Aprendizado

- *Aprendizado por REFORÇO*

É um aprendizado on-line de um mapeamento de E/S através de processo de Tentativa e Erro.



TAREFAS PREDITIVAS E DESCRITIVAS

