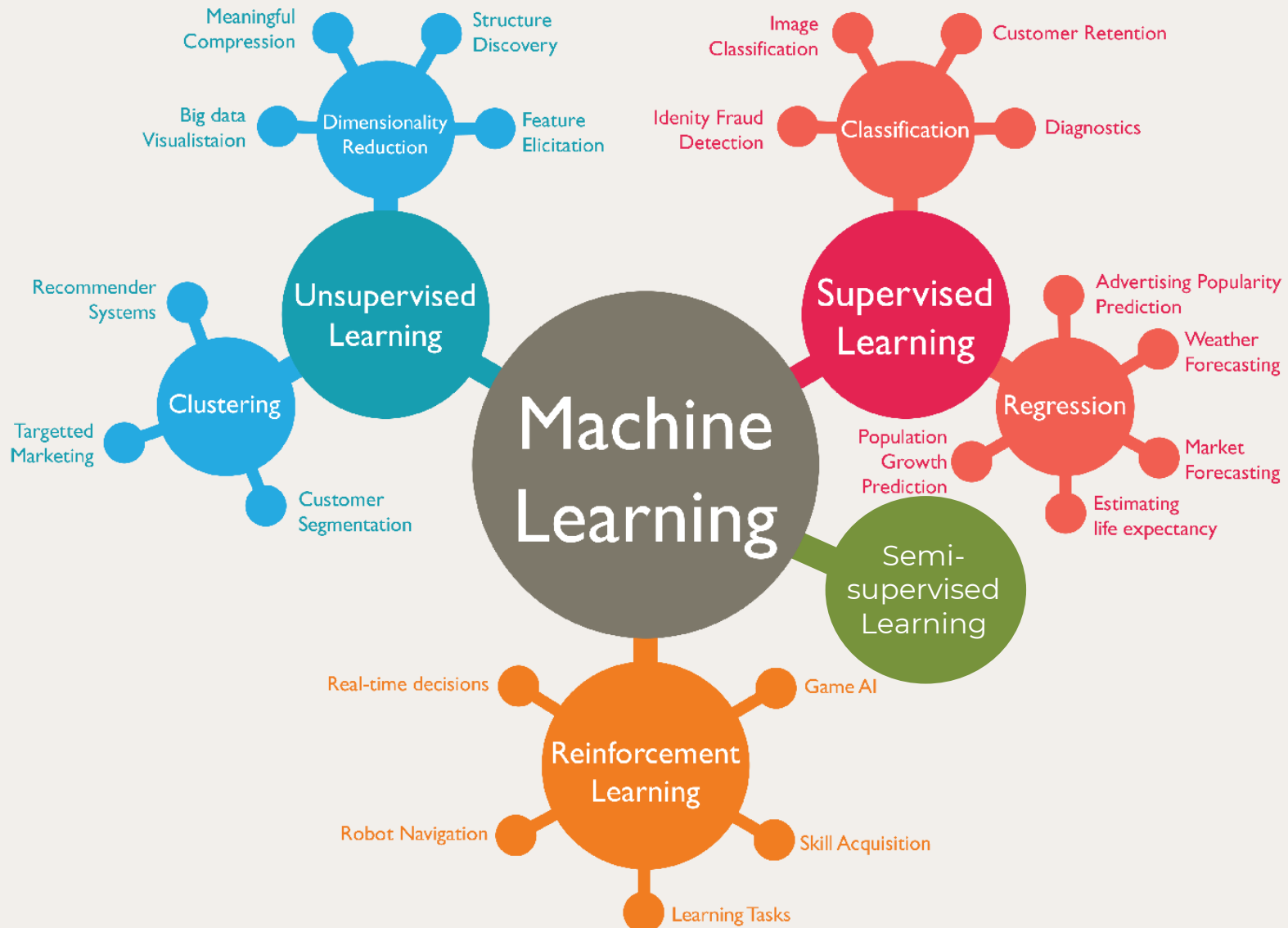

APRENDIZADO DE MÁQUINA

PROF. LETÍCIA RAPOSO
profleticiaraposo@gmail.com

DEFINIÇÃO



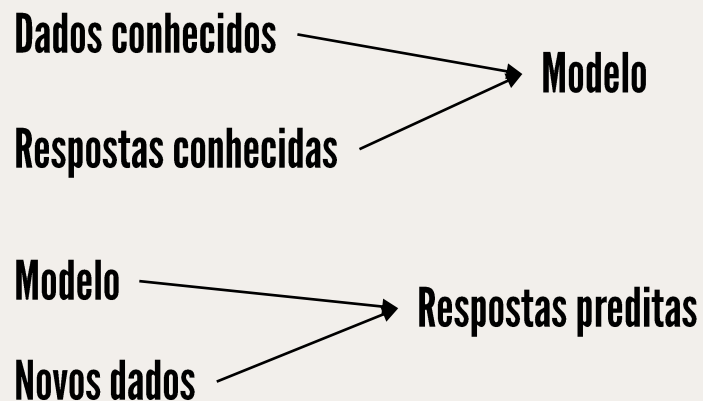
CONJUNTO DE REGRAS
E PROCEDIMENTOS
QUE PERMITE AOS
COMPUTADORES AGIR
E TOMAR DECISÕES
BASEADOS EM DADOS.



<https://becominghuman.ai/an-introduction-to-machine-learning-7db04da817c4>

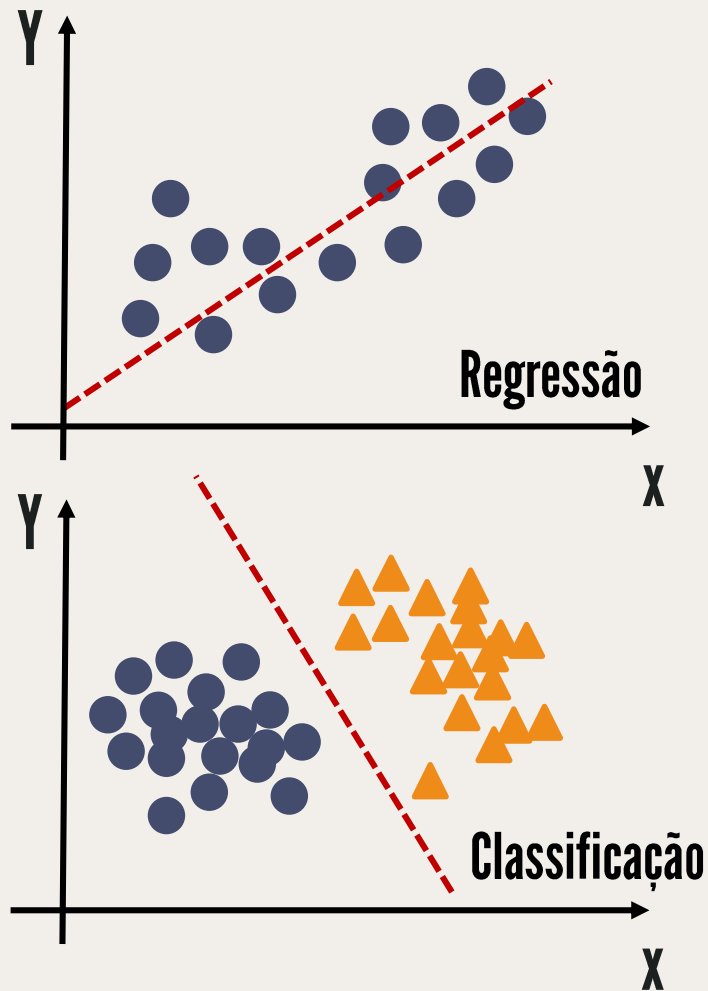
SUPERVISIONADO

O TERMO SUPERVISIONADO VEM DA SIMULAÇÃO DA PRESENÇA DE UM “SUPERVISOR EXTERNO”, QUE CONHECE A SAÍDA (RÓTULO) DESEJADA PARA CADA EXEMPLO.



- TAREFA DE ENCONTRAR UMA FUNÇÃO A PARTIR DE DADOS DE TREINAMENTO ROTULADOS.
- O OBJETIVO É ENCONTRAR OS PARÂMETROS ÓTIMOS QUE AJUSTEM UM MODELO QUE POSSA PREVER RÓTULOS DESCONHECIDOS EM OUTROS OBJETOS (O CONJUNTO DE TESTE).

SUPERVISIONADO

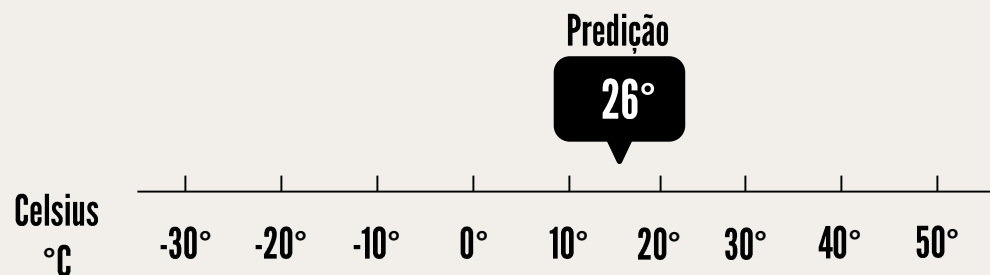


- REGRESSÃO: O OBJETIVO É PREVER UMA MEDIDA CONTÍNUA PARA UMA OBSERVAÇÃO. OU SEJA, AS VARIÁVEIS RESPOSTAS SÃO NÚMEROS REAIS.
- CLASSIFICAÇÃO: O OBJETIVO É ATRIBUIR UMA CLASSE (OU RÓTULO) DE UM CONJUNTO FINITO DE CLASSES A UMA OBSERVAÇÃO. OU SEJA, AS RESPOSTAS SÃO VARIÁVEIS CATEGÓRICAS.



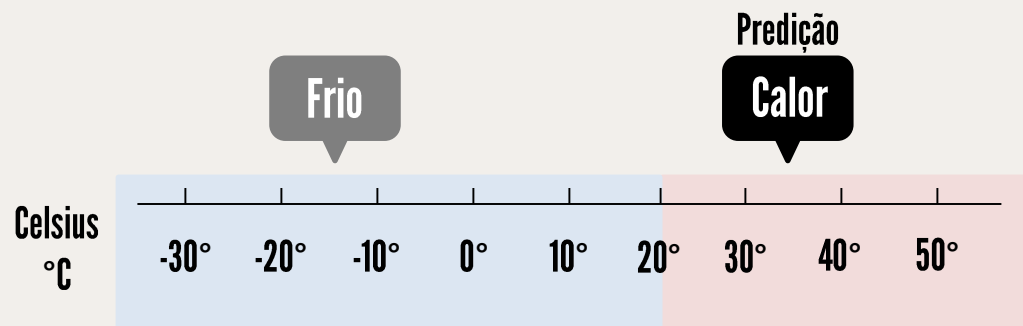
REGRESSÃO

Qual será a temperatura amanhã?

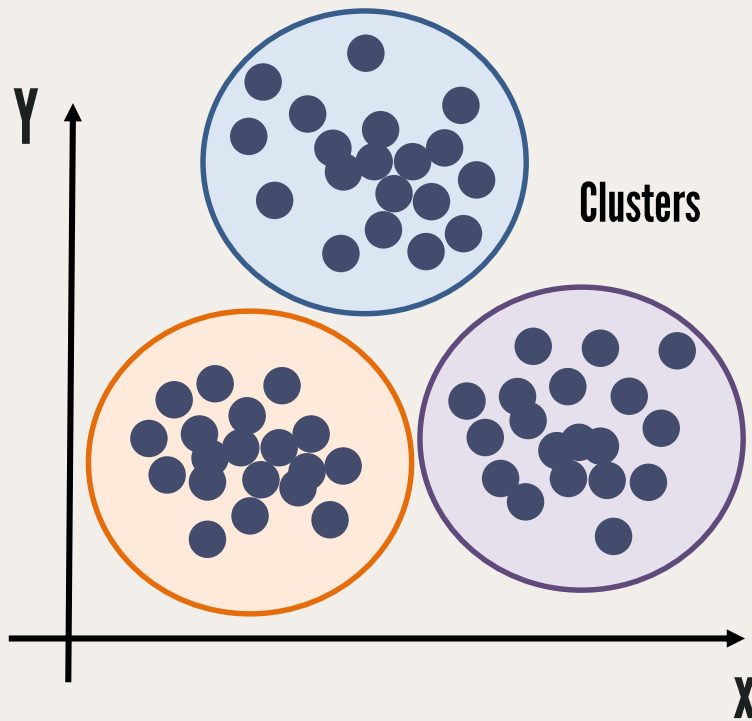


CLASSIFICAÇÃO

Vai fazer frio ou calor amanhã?

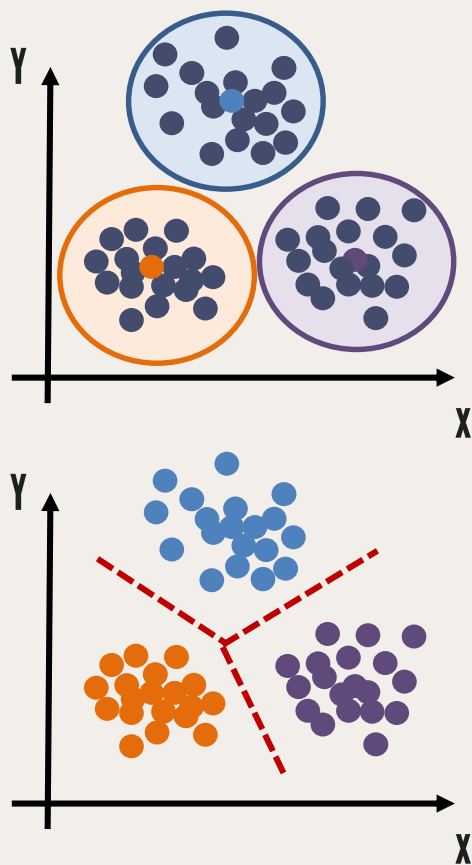


NÃO SUPERVISIONADO



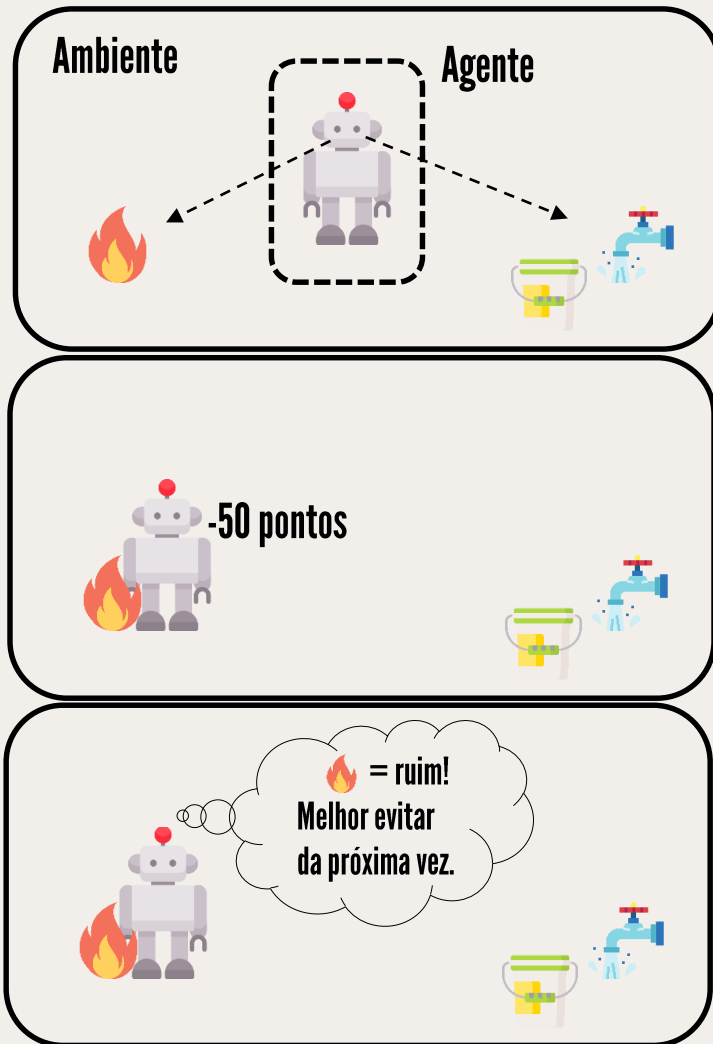
- O CONJUNTO DE TREINAMENTO NÃO É ROTULADO.
- NOSSO OBJETIVO É OBSERVAR ALGUMAS SIMILARIDADES ENTRE OS OBJETOS E INCLUI-LOS EM GRUPOS APROPRIADOS.
- TAMBÉM PODE SER USADA PARA REDUZIR O NÚMERO DE DIMENSÕES EM UM CONJUNTO DE DADOS PARA CONCENTRAR SOMENTE NOS ATRIBUTOS MAIS ÚTEIS.

SEMI-SUPERVISIONADO



- INCLUI AMBOS OS PROBLEMAS DISCUTIDOS ANTIORMENTE: ELA USA DADOS ROTULADOS E NÃO-ROTULADOS.
- A IDEIA É ROTULAR DADOS NÃO ROTULADOS POR UM APRENDIZADO NÃO SUPERVISIONADO E DEPOIS CONSTRUIR UM MODELO SUPERVISIONADO.

POR REFORÇO,



- UM AMBIENTE FORNECERÁ DADOS AOS QUAIS UM “AGENTE” IRÁ TOMAR DECISÕES DE ACORDO COM ELES.
- ESSE AMBIENTE IRÁ RESPONDER COM UMA RECOMPENSA/PUNIÇÃO QUE DETERMINARÁ A ALTERAÇÃO DE REGRAS DE DECISÃO FEITA PELO “AGENTE” PARA ATUAR MELHOR NO FUTURO.



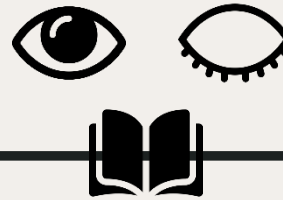
APRENDIZADOS



SUPERVISIONADO



NÃO
SUPERVISIONADO



SEMI-
SUPERVISIONADO



POR REFORÇO

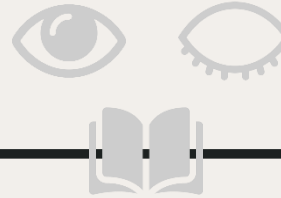
APRENDIZADOS



SUPERVISIONADO



NÃO
SUPERVISIONADO



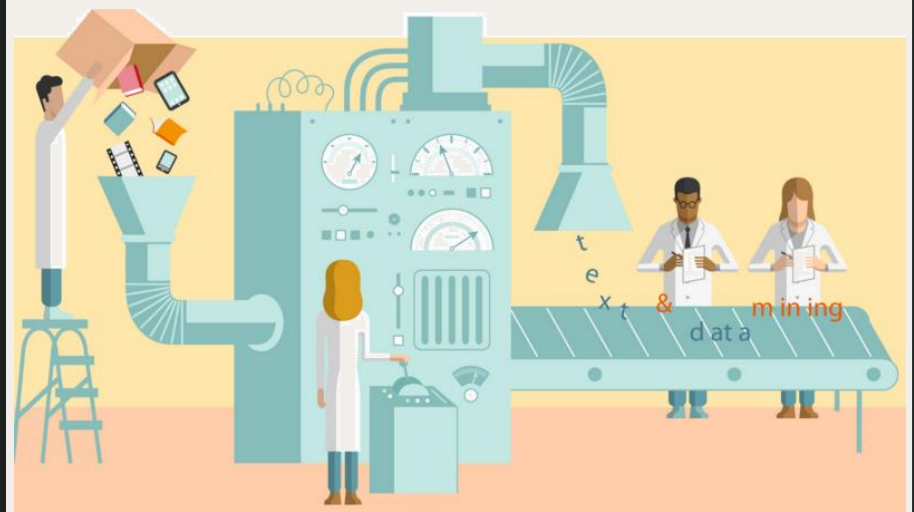
SEMI-
SUPERVISIONADO



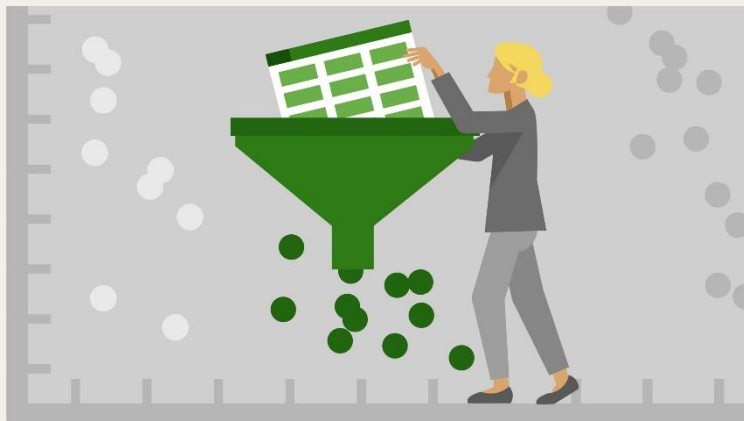
POR REFORÇO

- PREPARAÇÃO DOS DADOS
- ESCOLHA DE UM ALGORITMO
- TREINAMENTO
- AVALIAÇÃO
- AJUSTE DOS PARÂMETROS
- PREDIÇÃO

ETAPAS DO APRENDIZADO SUPERVISIONADO



PREPARAÇÃO DOS DADOS



- ELIMINAÇÃO MANUAL DE ATRIBUTOS
- INTEGRAÇÃO DE DADOS
- AMOSTRAGEM DE DADOS
- DADOS DESBALANCEADOS
- LIMPEZA DE DADOS
- TRANSFORMAÇÃO DE DADOS
- REDUÇÃO DE DIMENSIONALIDADE

ESCOLHA DE UM ALGORITMO



EXISTEM *TRADE-OFFS* ENTRE VÁRIAS CARACTERÍSTICAS DE ALGORITMOS, COMO:

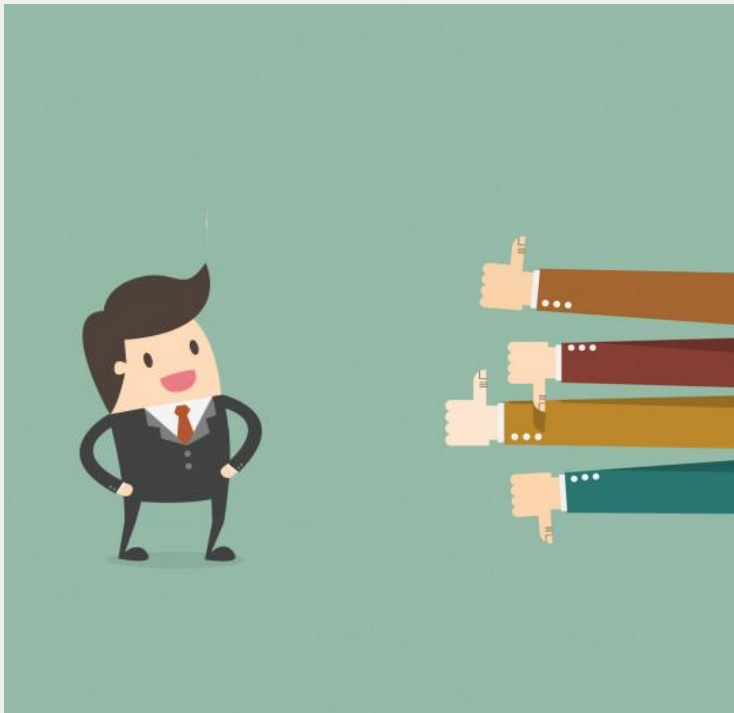
- VELOCIDADE DE TREINAMENTO
- USO DE MEMÓRIA
- ACURÁCIA PREDITIVA EM NOVOS DADOS
- TRANSPARÊNCIA OU INTERPRETABILIDADE (FACILIDADE PARA ENTENDER AS RAZÕES PELAS QUAIS UM ALGORITMO FAZ SUAS PREVISÕES)

TREINAMENTO



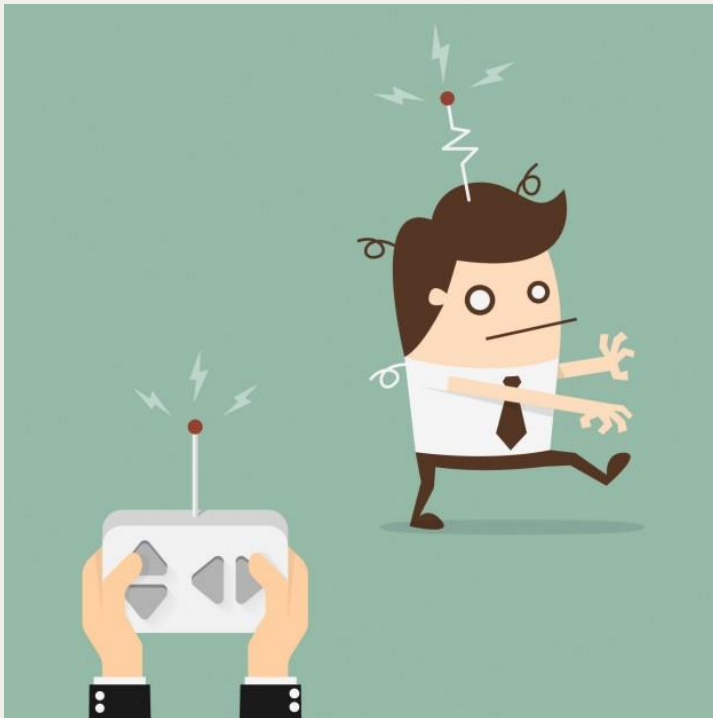
- REGRESSÃO LINEAR
- REGRESSÃO LOGÍSTICA
- K-NN
- NAÏVE BAYES
- ÁRVORES DE DECISÃO
- FLORESTA ALEATÓRIA
- REDES NEURAIIS
- MÁQUINAS DE VETORES DE SUPORTE
- ...

AVALIAÇÃO



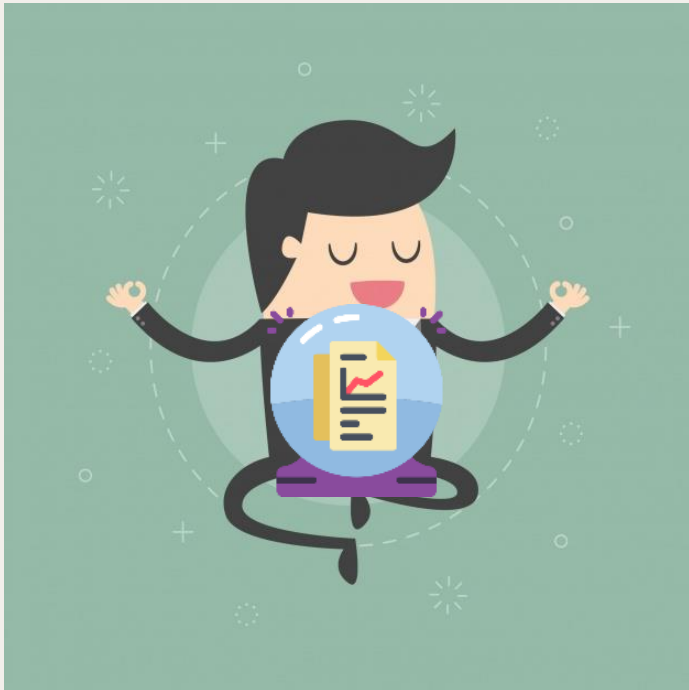
- QUANDO O TREINAMENTO ESTIVER CONCLUÍDO, É HORA DE VER SE O MODELO ESTÁ BOM.
- A AVALIAÇÃO NOS PERMITE TESTAR NOSSO MODELO EM RELAÇÃO A DADOS QUE NUNCA FORAM USADOS PARA TREINAMENTO.
- ISSO DEVE SER REPRESENTATIVO DE COMO O MODELO PODE FUNCIONAR NO MUNDO REAL.

AJUSTE DOS PARÂMETROS



- TENTAR MELHORAR AINDA MAIS O MODELO.
- É IMPORTANTE DEFINIR O QUE TORNA UM MODELO “BOM O SUFICIENTE”, CASO CONTRÁRIO, VOCÊ PODE SE ENCONTRAR AJUSTANDO PARÂMETROS POR UM TEMPO MUITO LONGO.
- ESSES PARÂMETROS SÃO GERALMENTE CHAMADOS DE “HIPERPARÂMETROS”.

PREDIÇÃO



- ETAPA EM QUE O MODELO IRÁ PREDIZER A RESPOSTA A NOVOS CASOS.
- TODO O TRABALHO REALIZADO TEM COMO OBJETIVO CHEGAR NESTA ETAPA.

REFERÊNCIAS



- Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina, por Katti Faceli, Ana Carolina Lorena, João Gama, André C. P. L. F. de Carvalho;
- Notas de aula do curso Mineração de Dados em Biologia Molecular, ministrado por André C. P. L. F. de Carvalho;
- FRIEDMAN, Jerome; HASTIE, Trevor; TIBSHIRANI, Robert. The elements of statistical learning. New York: Springer series in statistics, 2001.