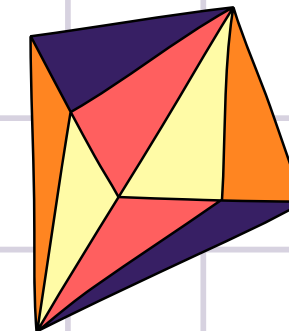
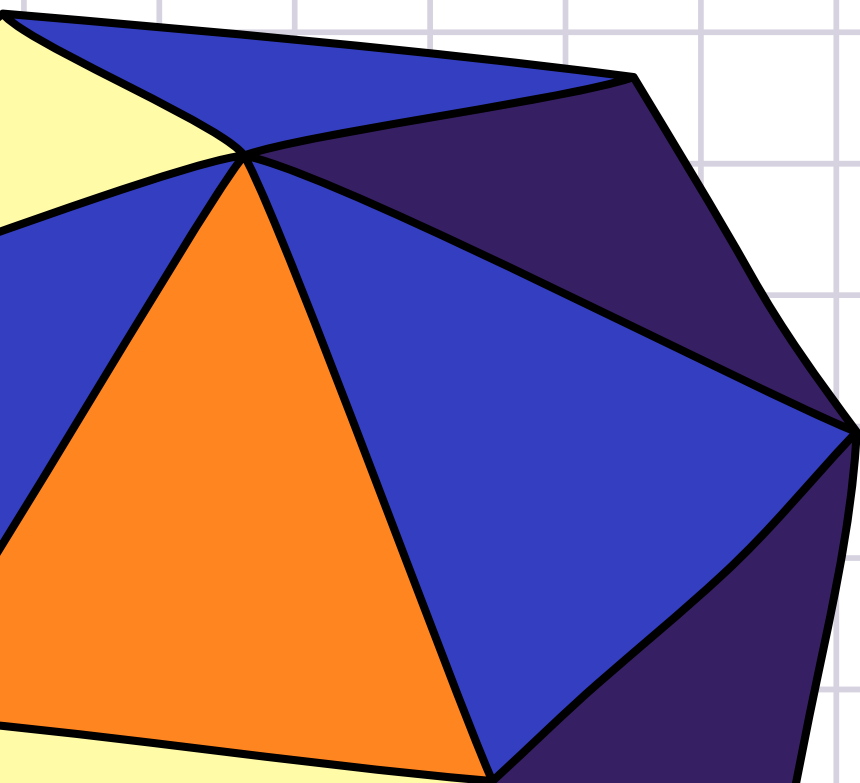
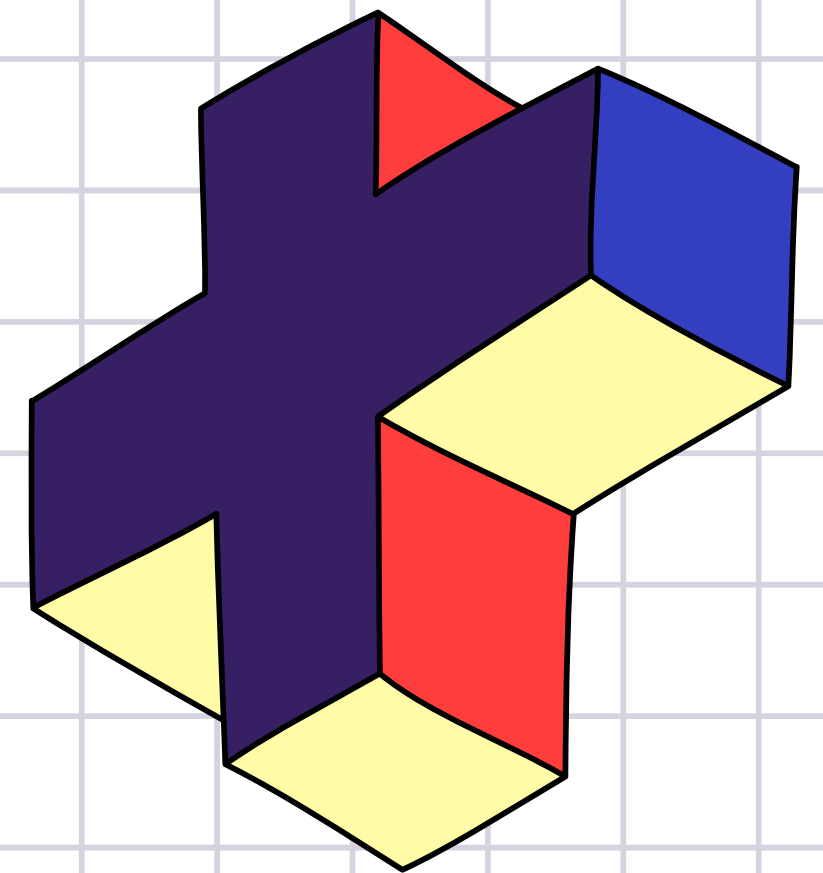


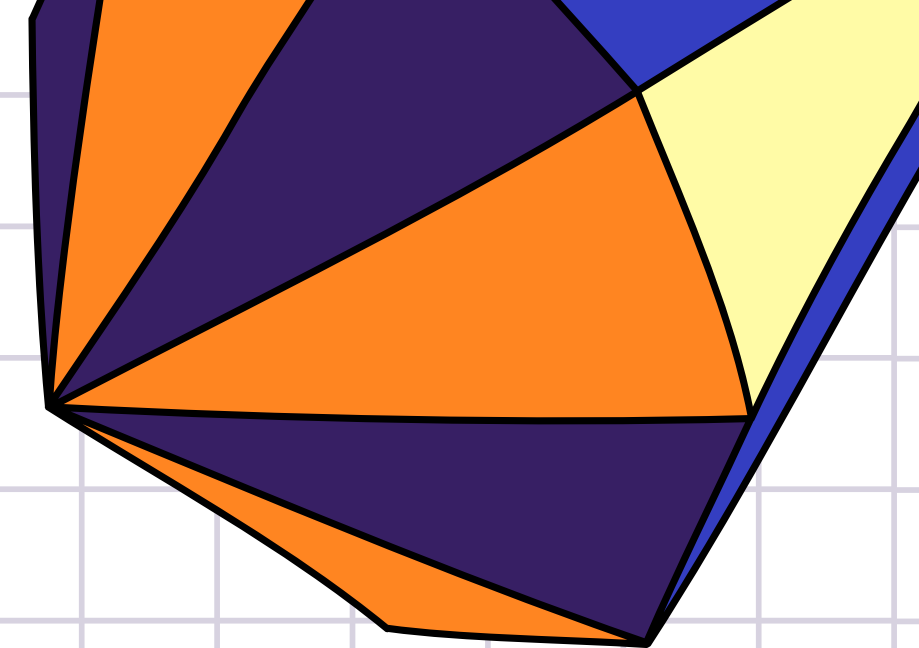
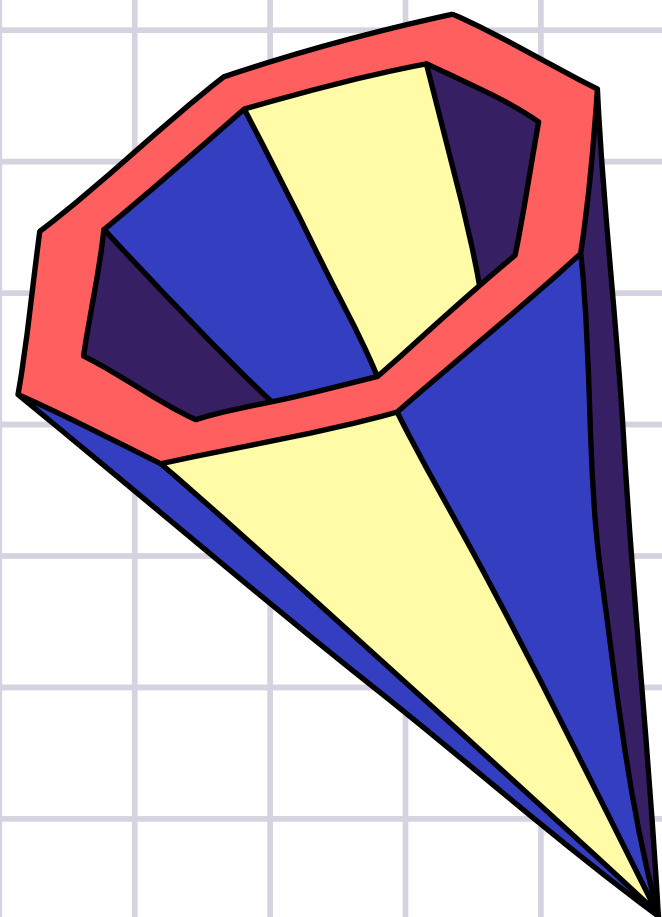
INTRODUÇÃO À CIÊNCIA DE DADOS



03

INTRODUÇÃO AO MACHINE LEARNING E AO DEEP LEARNING





MACHINE LEARNING

É UM SUBCAMPO DA IA

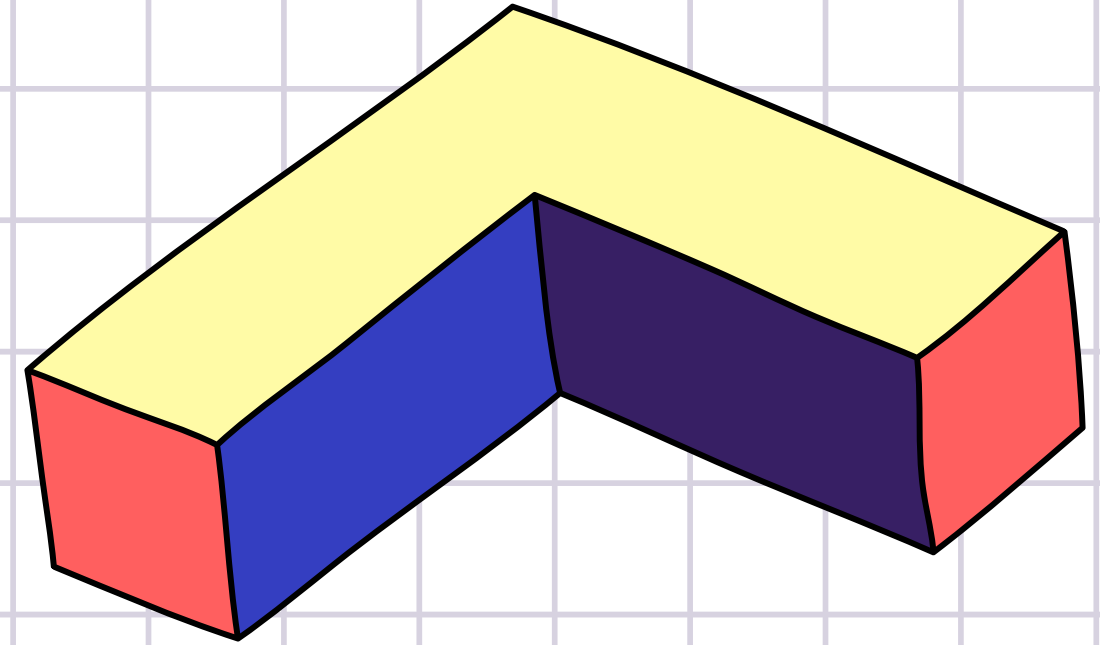
que permite que computadores, a partir de algoritmos e modelos, aprendam em cima de dados, sem terem instruções explícitas.

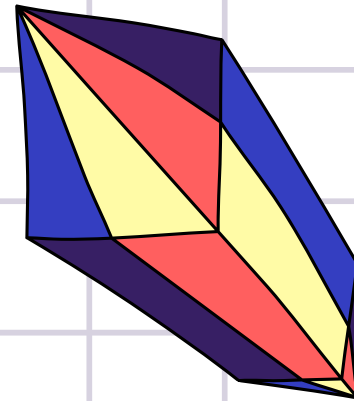
RECONHECEM PADRÕES

nesses dados e tomam decisões de ajustes, sem a nossa intervenção a cada passo.

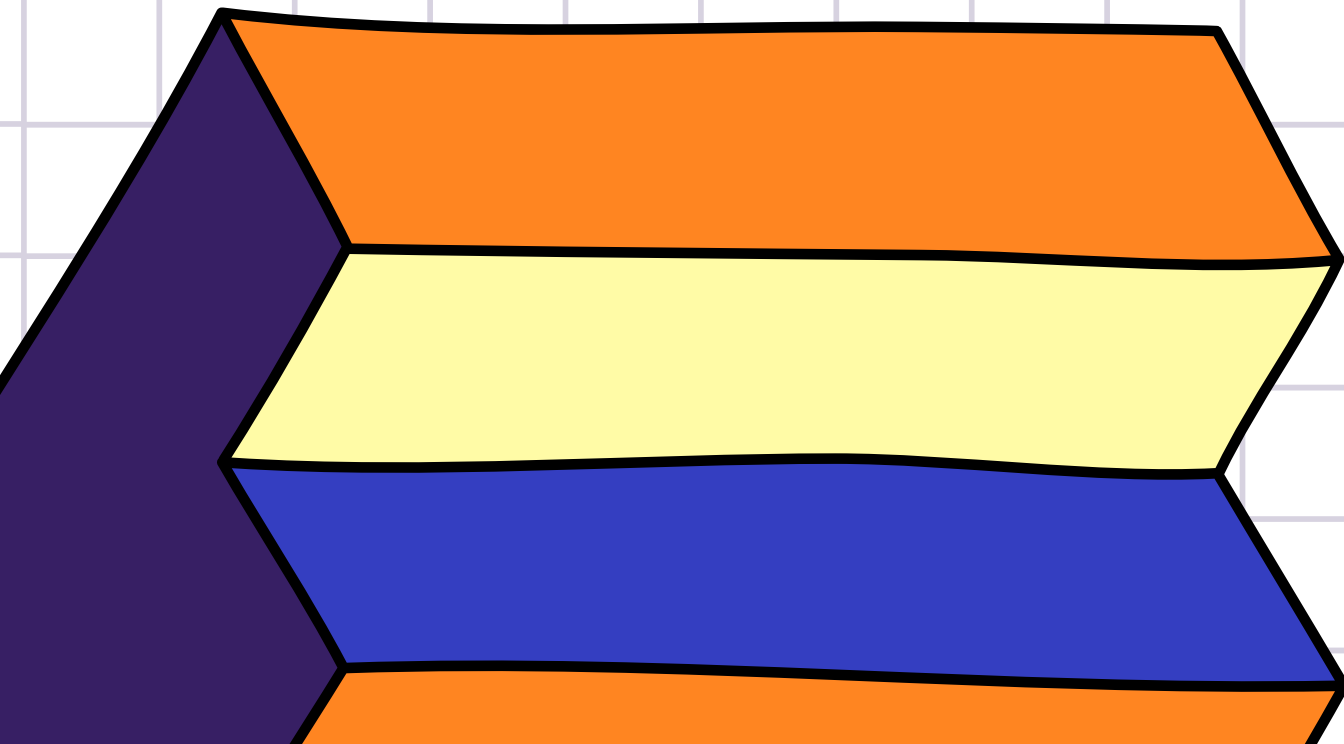
AO INVÉS DE PROGRAMAR

uma função específica, damos muitos dados ao algoritmo e ele aprende a desempenhar essa função a partir dos seus padrões.





TIPOS DE APRENDIZADO



SUPERVISIONADO

- Aprende em cima de dados pré-classificados ou já preditos, comparando o dado com o seu rótulo/tipo/classe.

NÃO-SUPERVISIONADO

- Busca similaridades entre os dados para, em geral, identificar classes ou segmentos, sem um rótulo pré-existente pra isso.

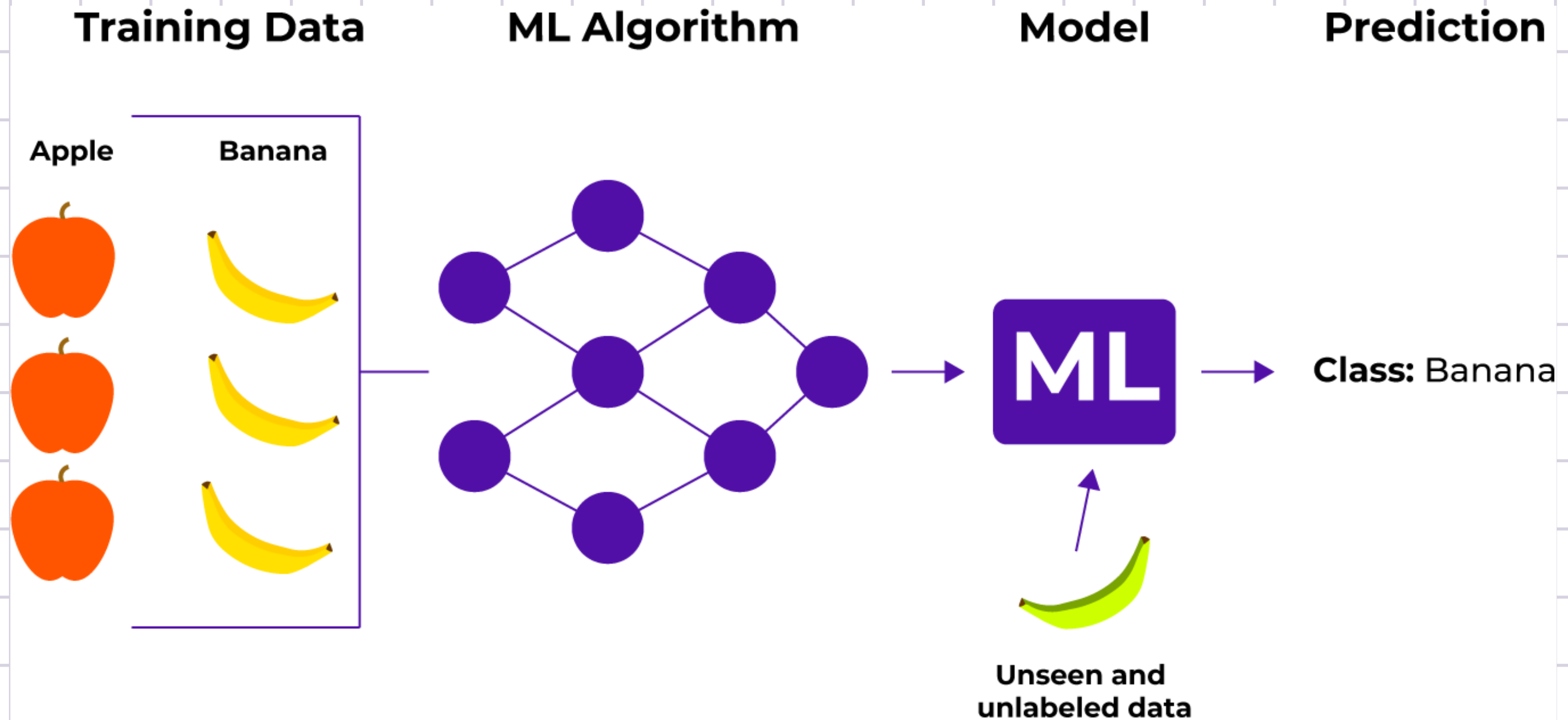
SEMI-SUPERVISIONADO

- O meio termo: alguns dados são rotulados, e outros não. Assim, o modelo busca rotular os faltantes.

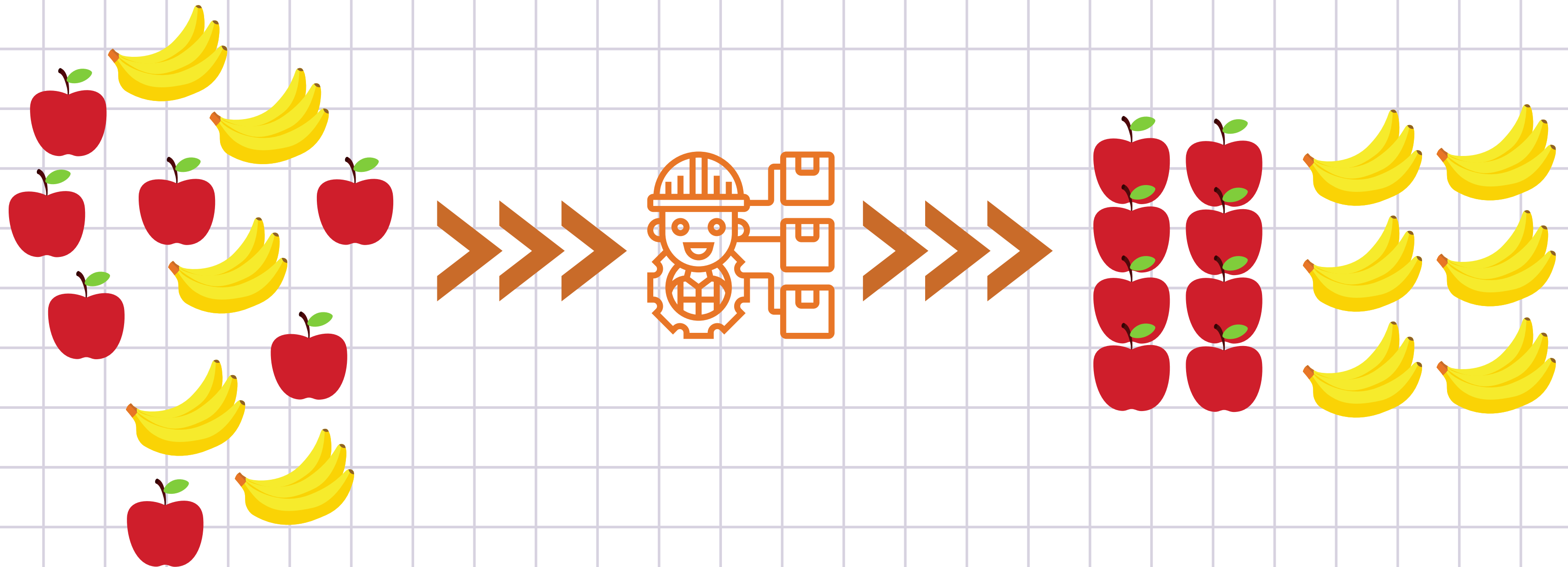
POR REFORÇO

- Nesse aprendizado, existe um agente que executa ações e, por elas, recebe recompensas ou punições, ajustando seu "comportamento".

APRENDIZADO SUPERVISIONADO



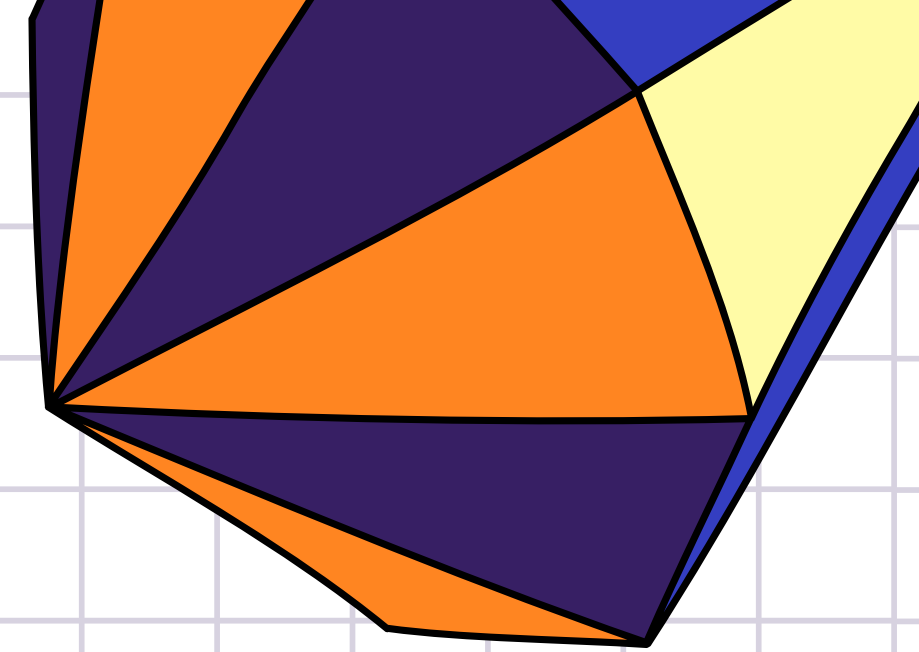
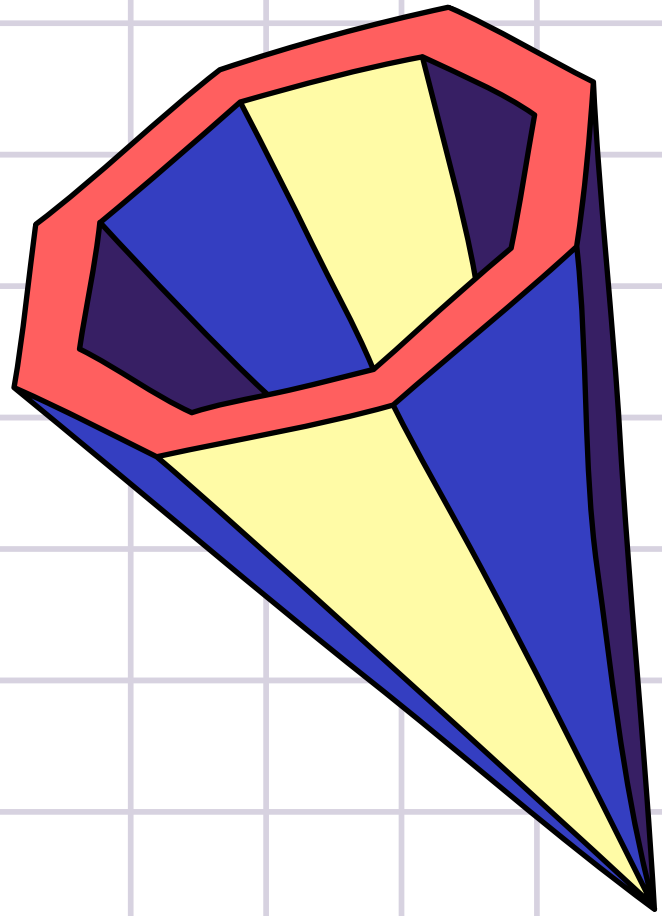
APRENDIZADO NÃO-SUPERVISIONADO



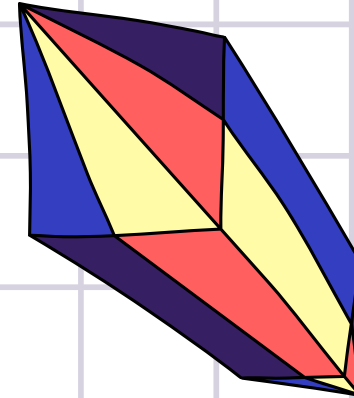
APRENDIZADO POR REFORÇO



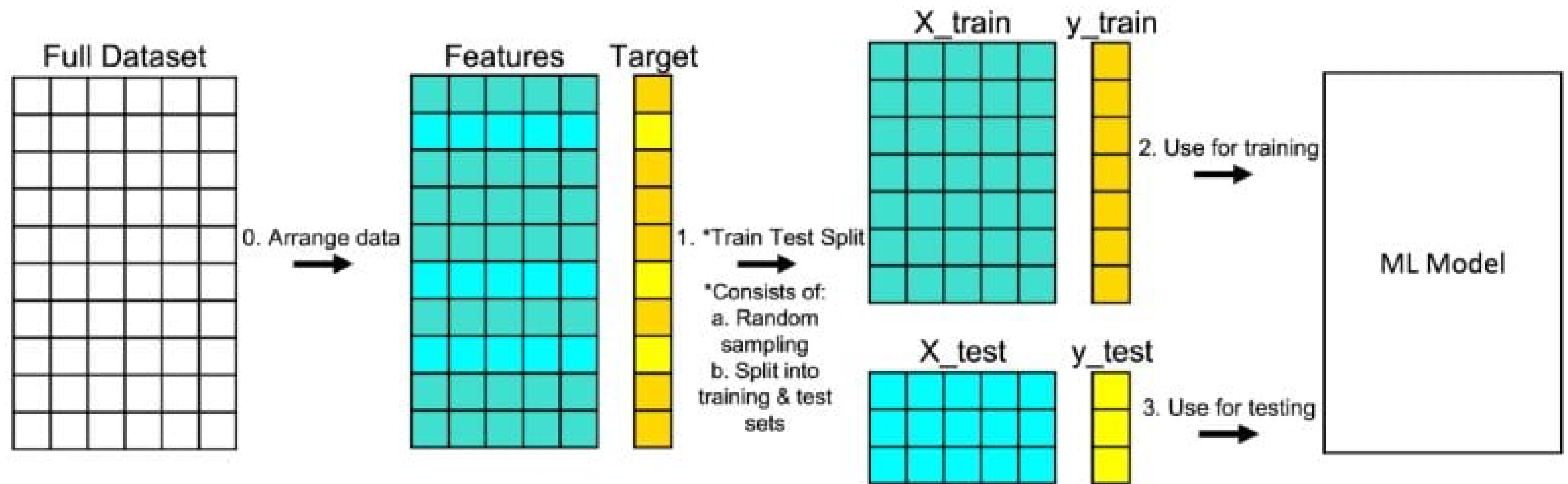
APRENDIZADO SEMI-SUPERVISIONADO



COMO AS MÁQUINAS APRENDEM?



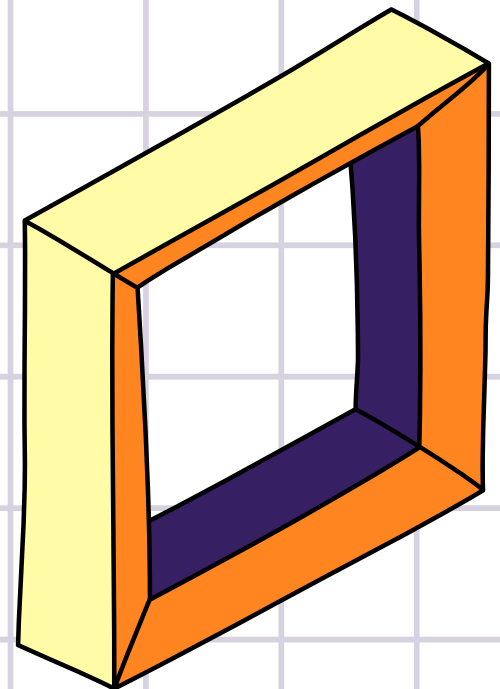
O PROCESSO DE APRENDIZAGEM



TREINO E TESTE

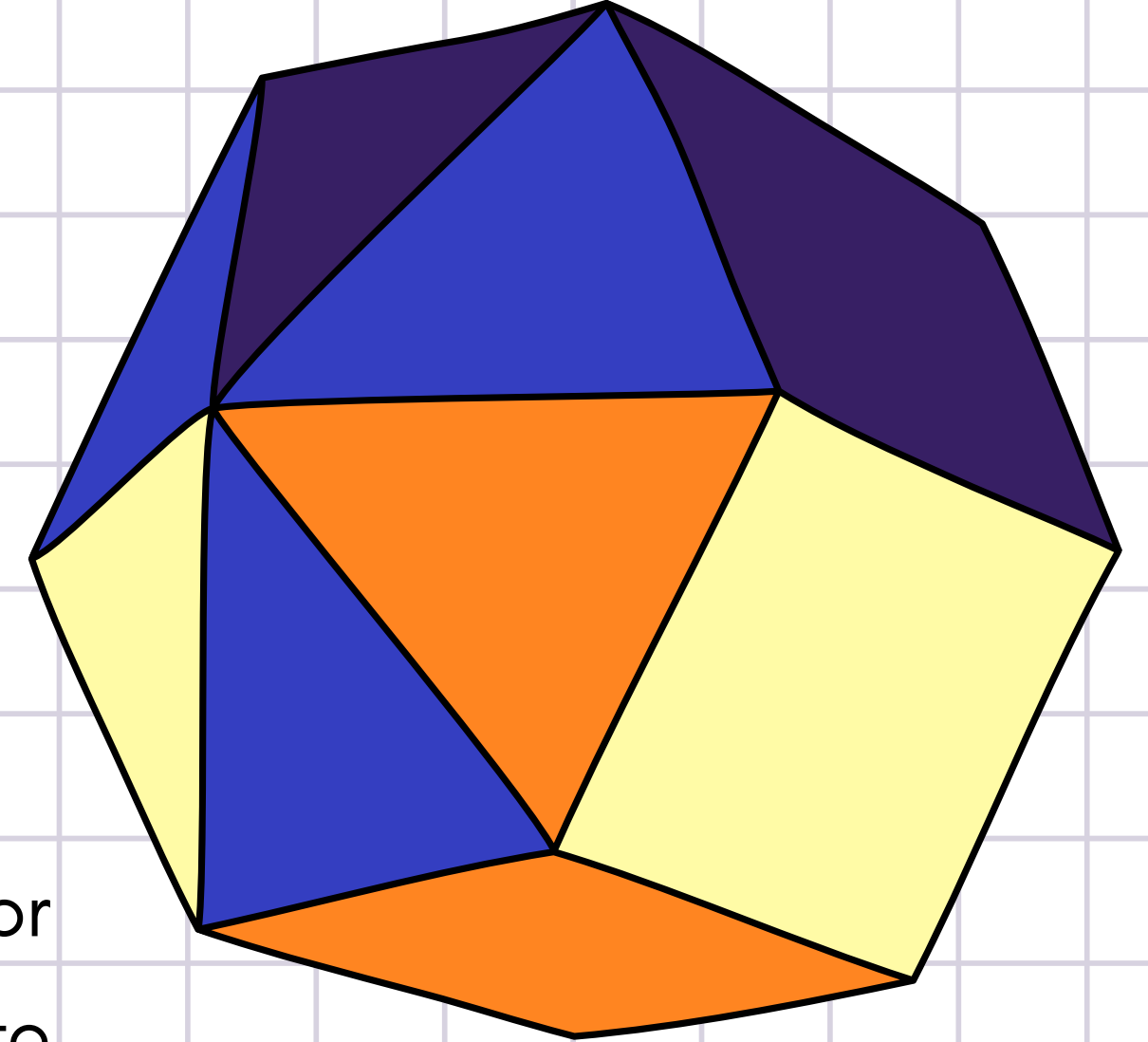
PROCESSO DE APRENDIZAGEM

- Dividimos um dataset em duas partes: treino e teste
- Geralmente o teste corresponde a 20%-30% do total
- Geramos o modelo em cima do treino, e aplicamos o modelo em cima do teste para ver como ele performa com dados novos
- Esse método é essencial para garantir um modelo que não sofra de *overfitting*
- E se ele não performar bem? Mudamos nossas variáveis, aumentamos os dados, mudamos o algoritmo...

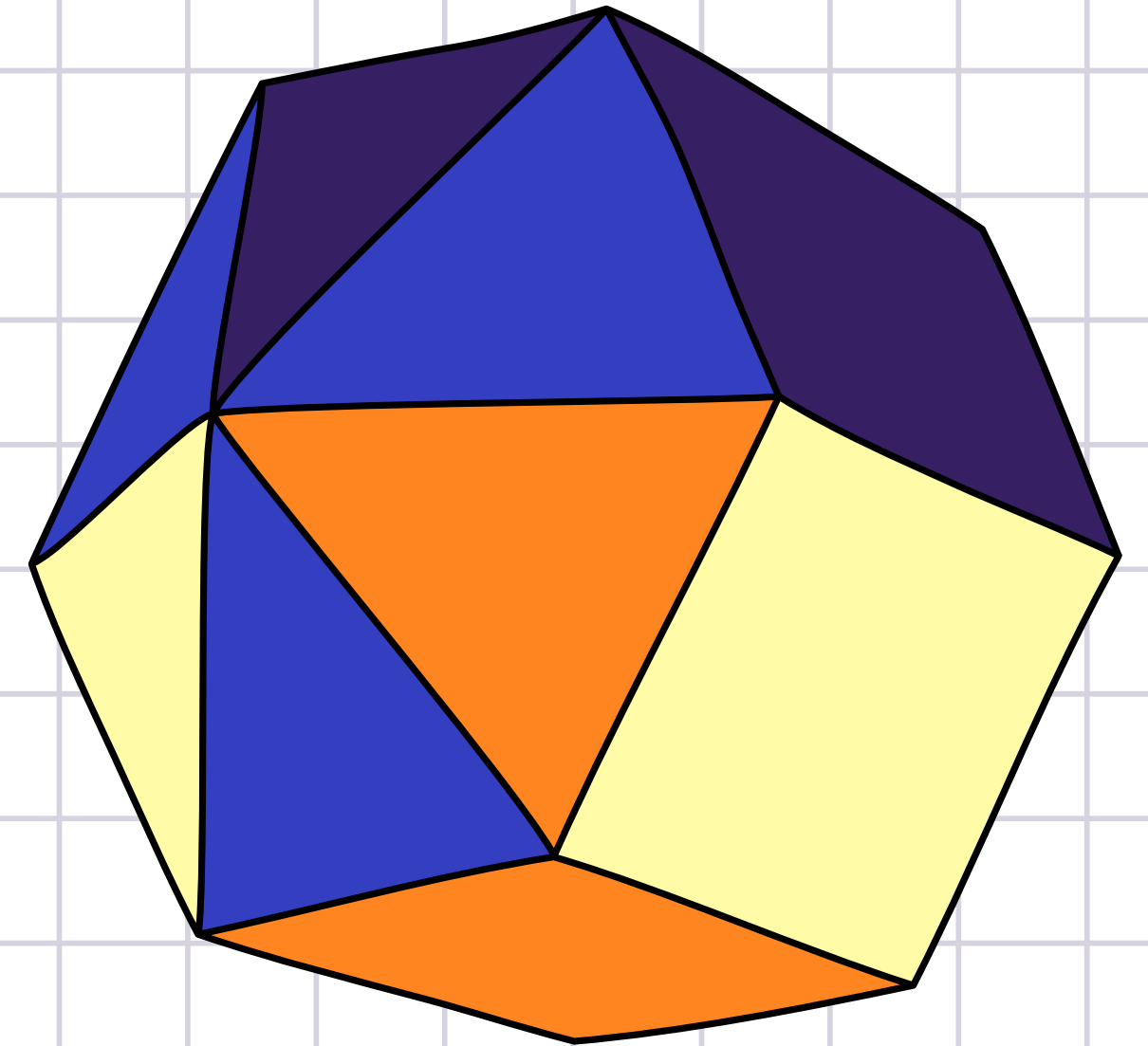


REPETIÇÕES/ ITERAÇÕES

Muitos algoritmos de aprendizado de máquina aprendem por meio de um processo iterativo. Começamos com um conjunto aleatório de parâmetros do modelo e, em seguida, iteramos repetidamente, ajustando os parâmetros para minimizar a diferença (o "erro") entre as previsões do modelo e os verdadeiros valores nos dados de treinamento.



REPETIÇÕES/ITERAÇÕES

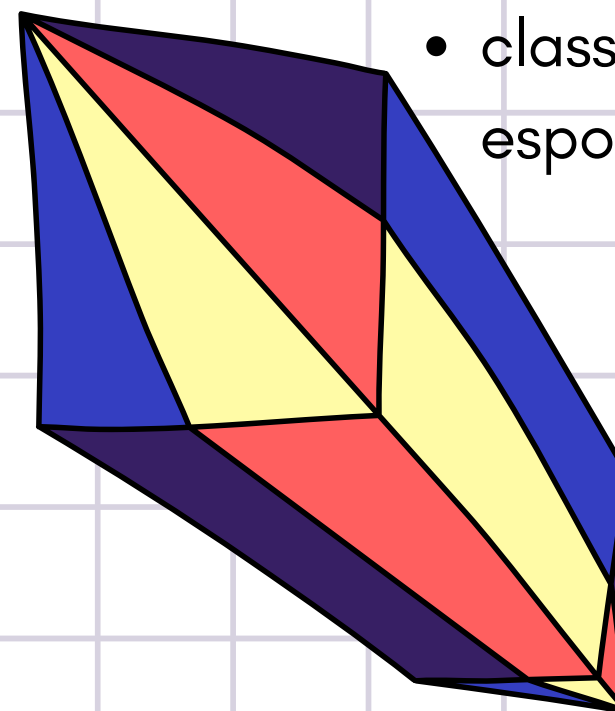


TIPOS DE PROBLEMA

REGRESSÃO

O objetivo é prever um valor contínuo:

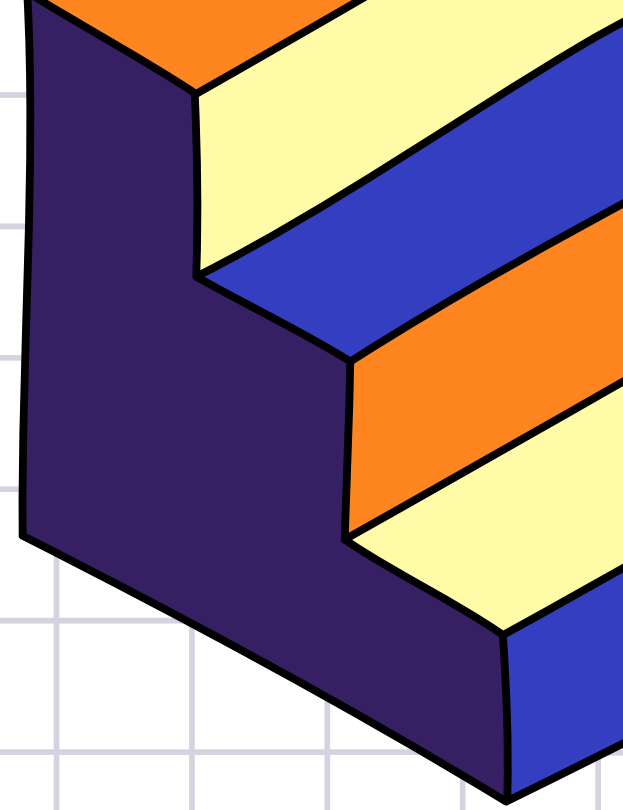
- preço de casas pelas suas características
- prever o valor de uma ação no futuro
- estimar expectativa de vida
- prever as vendas de um produto no próximo trimestre
- prever o tempo que demorará para um carro chegar ao destino, baseado nas condições do trânsito, na velocidade e na distância
- estimar a idade de uma pessoa pela foto do rosto



CLASSIFICAÇÃO

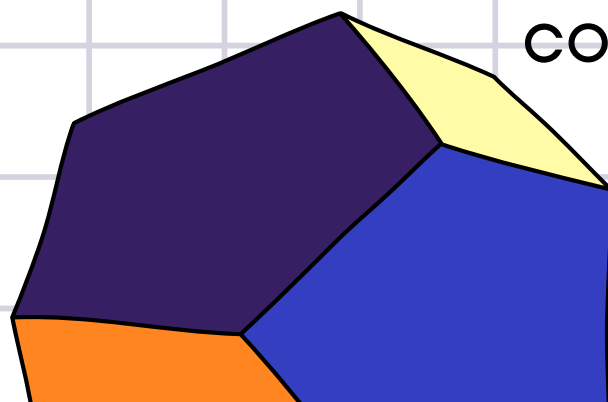
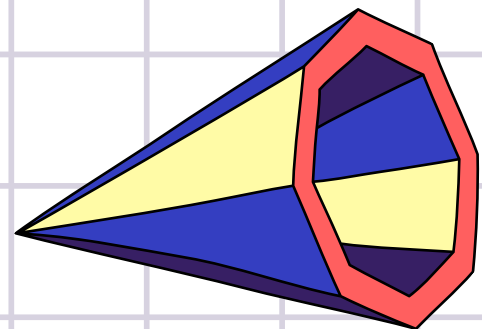
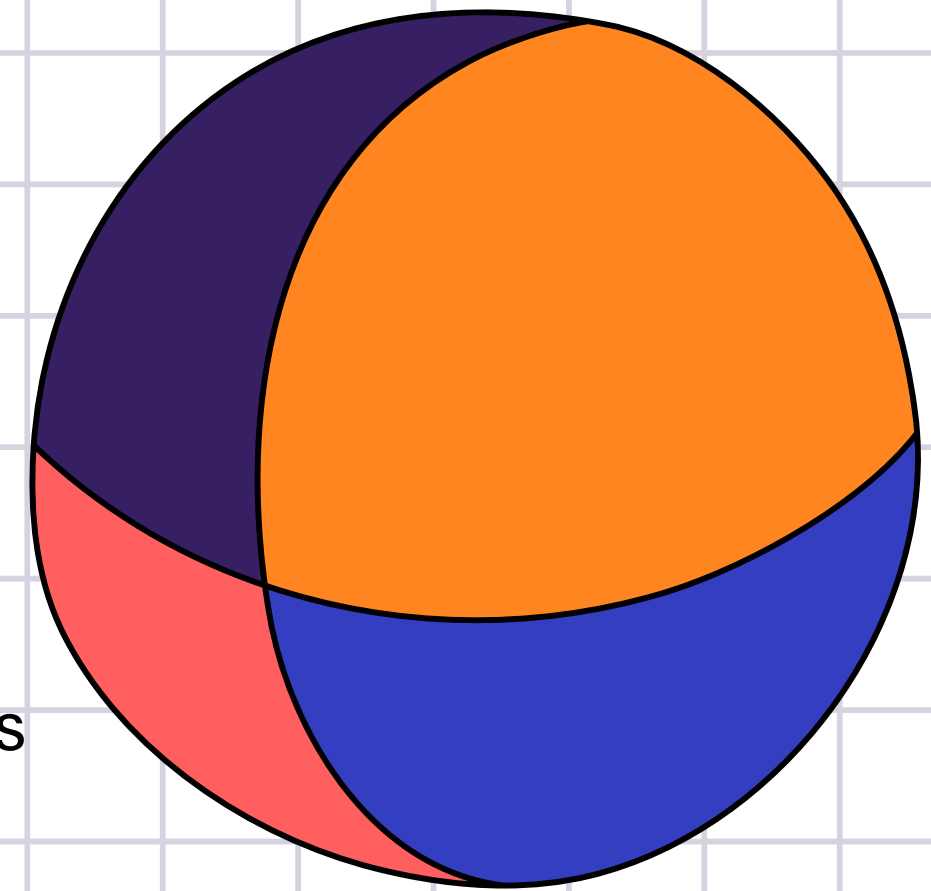
O objetivo é prever uma classe ou categoria:

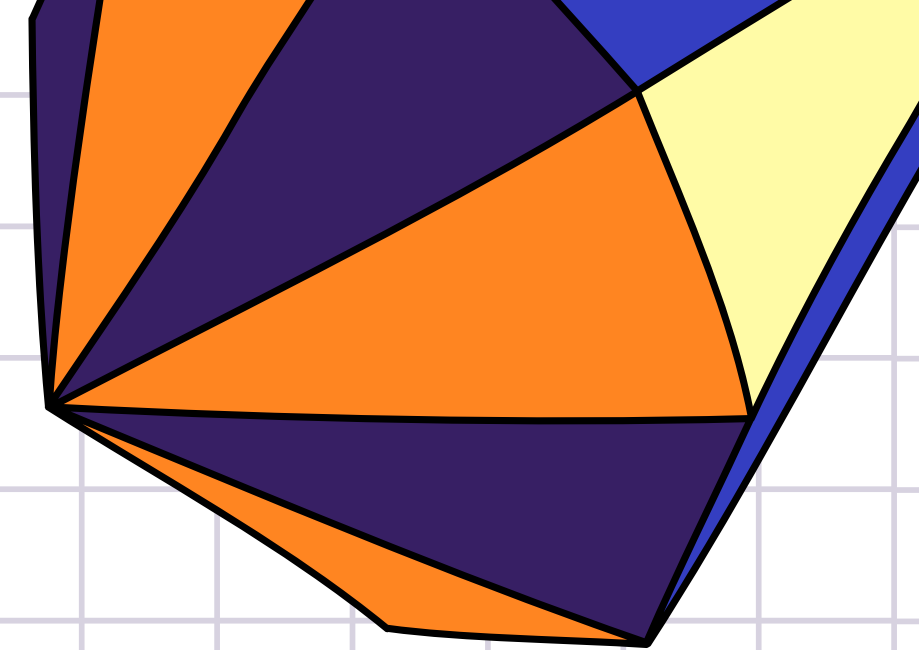
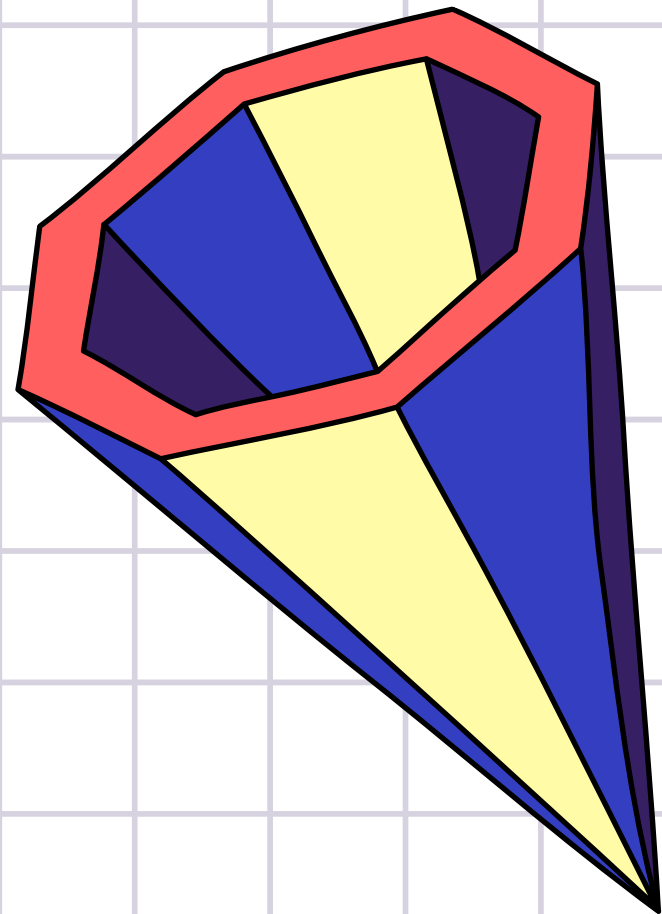
- dizer se um email é Spam ou não
- prever a raça de um cão pela foto
- se uma mensagem é positiva, negativa ou neutra
- classificar um tweet para identificar um discurso de ódio
- classificar uma notícia entre política, esporte, entretenimento



SCIKIT-LEARN

- É a biblioteca que utilizamos no Python para criar modelos de aprendizado de máquina
- Possui diversos algoritmos: Regressão Linear, Logística, Máquinas de Vetor Suporte, Naïve-Bayes, RandomForest, K-Nearest Neighbors, etc.
- Também possui ferramentas para fazer a separação de treino e teste, permite trabalhar com textos, escolher os melhores parâmetros de modelos, dentre outros.
- A documentação é ampla, a interface é simples e consistente.





DEEP LEARNING

DEEP LEARNING

O CÉREBRO HUMANO

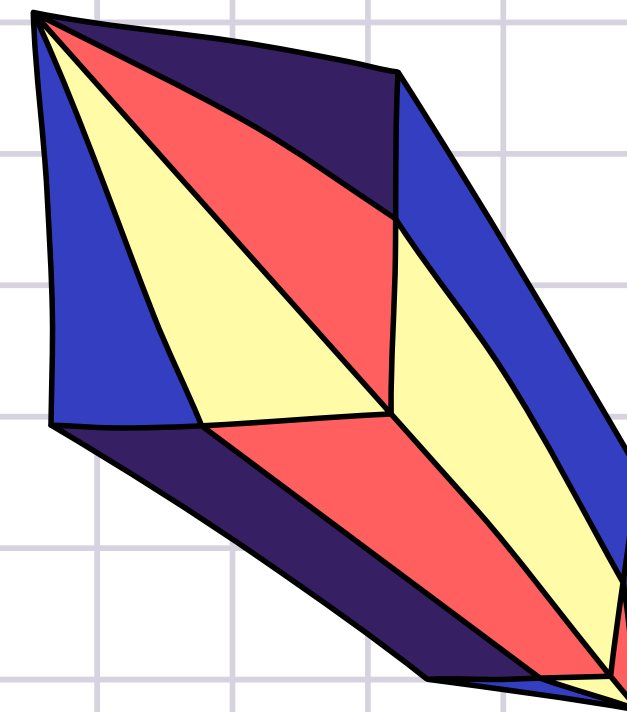
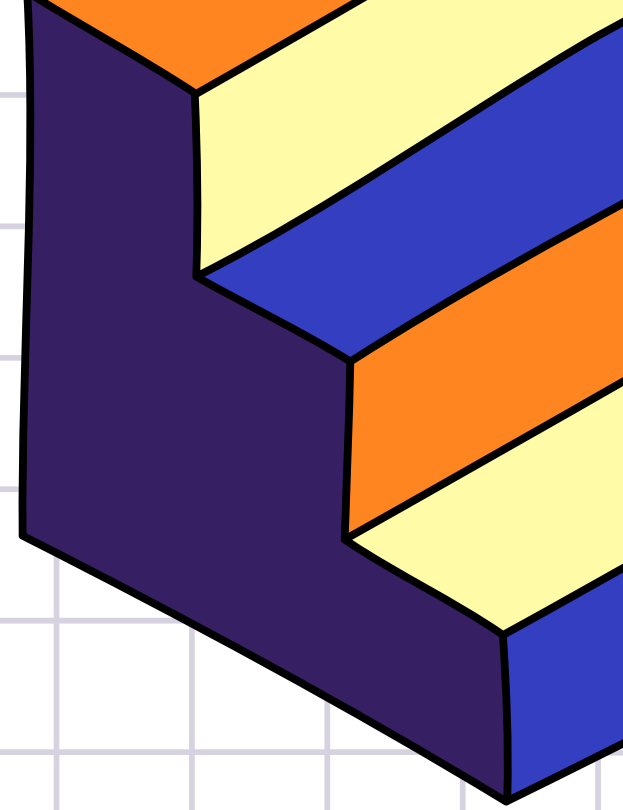
é a inspiração para esse subcampo da IA: as redes neurais artificiais.

É PROFUNDO

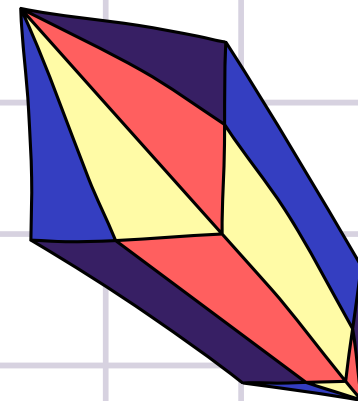
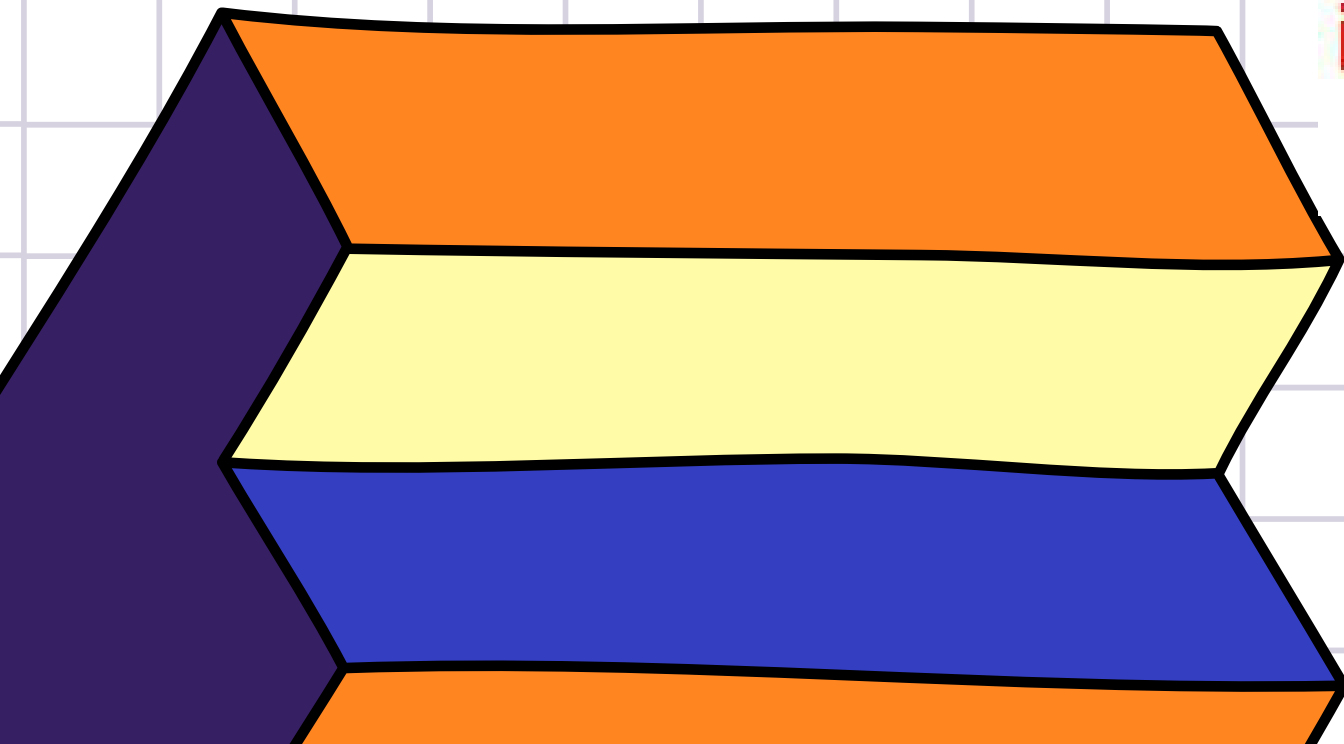
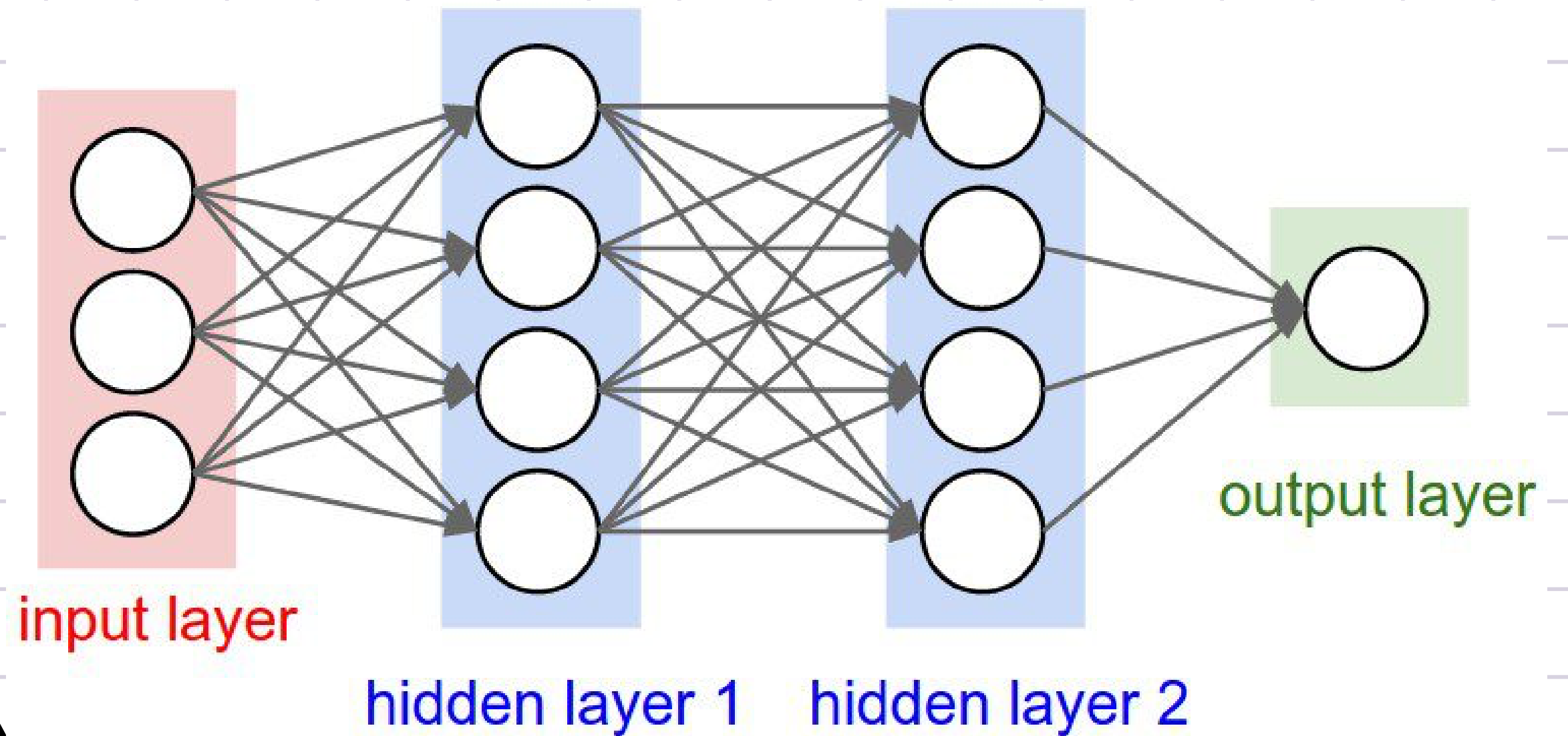
pois a rede neural possui diversas camadas, que aprofundam a análise.

ATUALMENTE É UTILIZADO

em diversos sistemas: reconhecimento de imagem, processamento de linguagem natural, reconhecimento e geração de voz, geração de imagens e carros autônomos.



REDES NEURAIS



REDES NEURAIS

