

Ciência de Dados: Aprendizagem de Máquina usando Python

Joyce Quintino

Mestranda em Ciência da Computação - MDCC/UFC

Analista de Sistemas - GREat

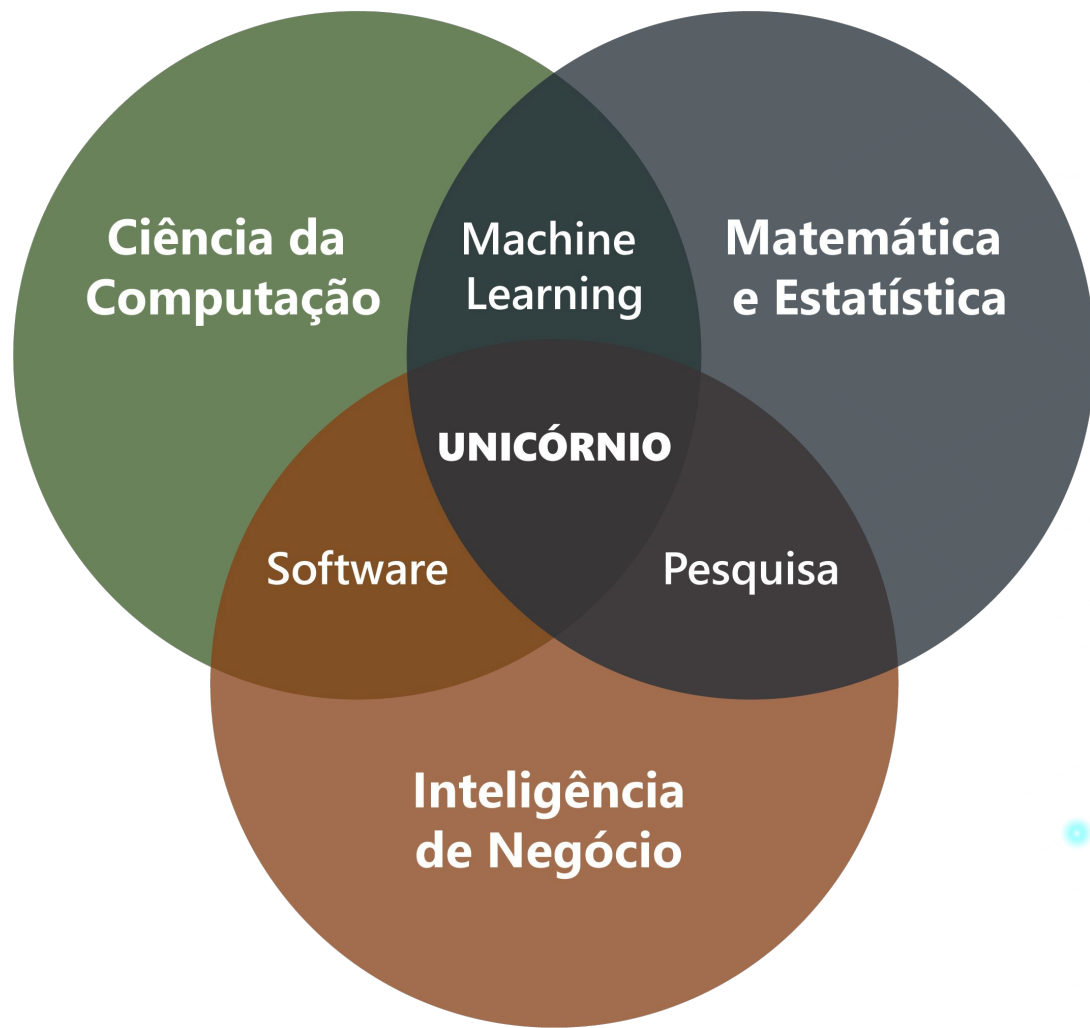


Roteiro

- ⬡ O que é Ciência de Dados?
- ⬡ Principais Etapas da Ciência de Dados
- ⬡ Aplicações
- ⬡ Cientista de Dados
- ⬡ Linguagens de programação e ferramentas
- ⬡ Inteligência Artificial
- ⬡ Aprendizagem de Máquina
- ⬡ Algoritmos de Aprendizagem de Máquina
- ⬡ Tipos de Variáveis
- ⬡ Vamos ao código?!

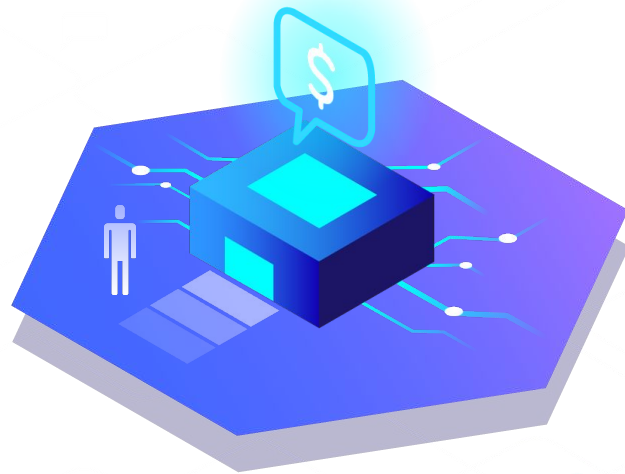
1. O que é Ciência de Dados?

- ⬡ Engloba extração, limpeza, preparação e análise de dados
- ⬡ Identificar os possíveis *insights* nos dados
- ⬡ Obter informações relevantes para empresas definirem melhor suas estratégias de negócio



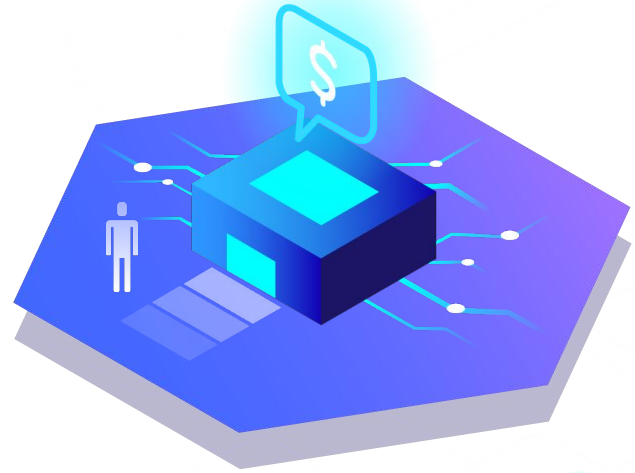
2. Principais etapas da Ciência de Dados

- ⬡ Importação
- ⬡ Organização
- ⬡ Transformação
- ⬡ Visualização
- ⬡ Modelagem
- ⬡ Comunicação



3. Aplicações

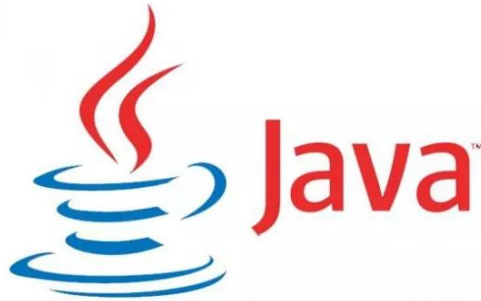
- ⬡ Sistemas de recomendação
- ⬡ Comparadores de preços
- ⬡ Logística de entrega
- ⬡ Fraude e risco
- ⬡ Reconhecimento de imagem e fala



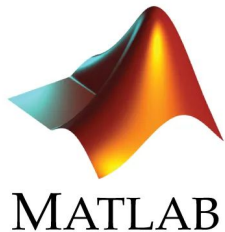
4. Cientista de Dados

- ⬡ Faça uma auto avaliação
- ⬡ Estatística e Matemática
- ⬡ Big Data
- ⬡ Linguagem de Programação e Aprendizagem de Máquina
- ⬡ Engenharia de Dados
- ⬡ Conhecimento de Negócios
- ⬡ Técnicas de Apresentação e Visualização de Dados
- ⬡ Pratique!!

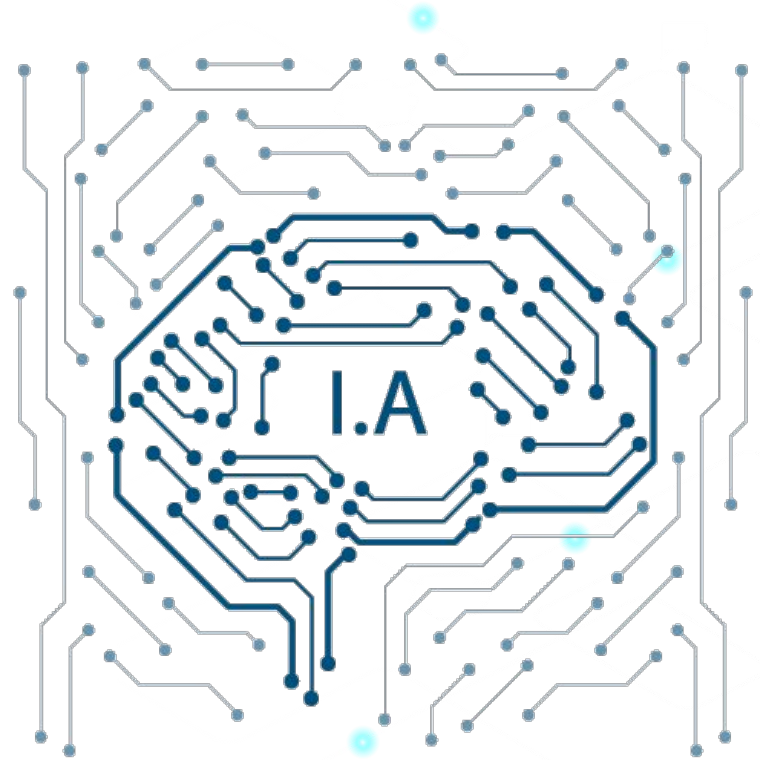
5. Linguagens de programação



5. Ferramentas



Inteligência Artificial

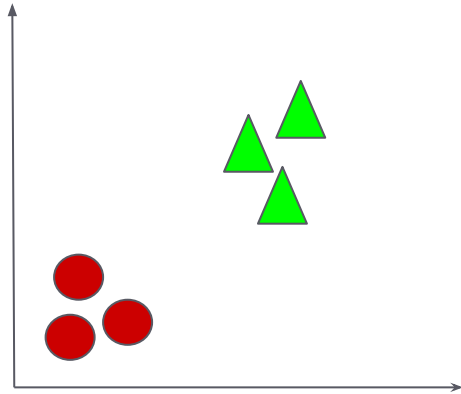


7. Aprendizagem de Máquina

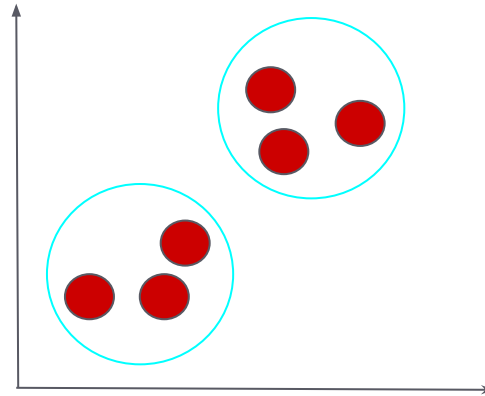
- ⬡ Capacidade de uma máquina aprender
- ⬡ A máquina aprende interagindo com o ambiente
- ⬡ O ambiente é representado por dados
- ⬡ A máquina aprende com os dados e cria um modelo que representa sua inteligência



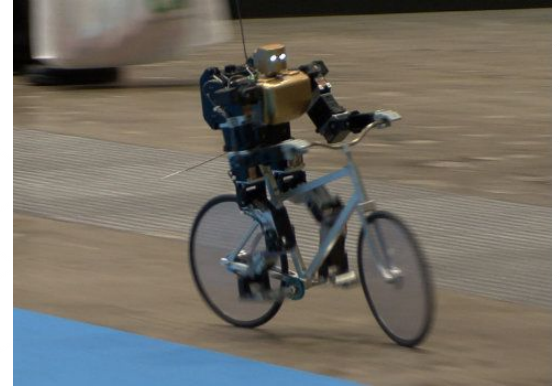
Tipos de Aprendizagem de Máquina



Aprendizado Supervisionado

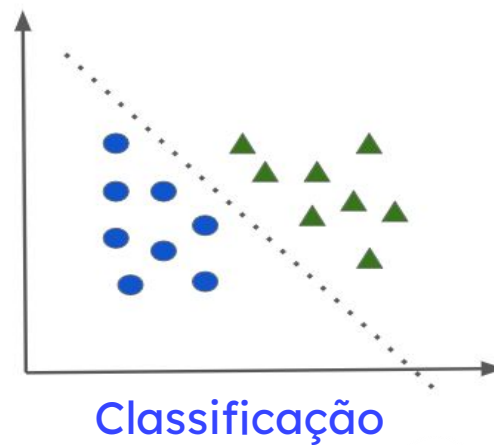
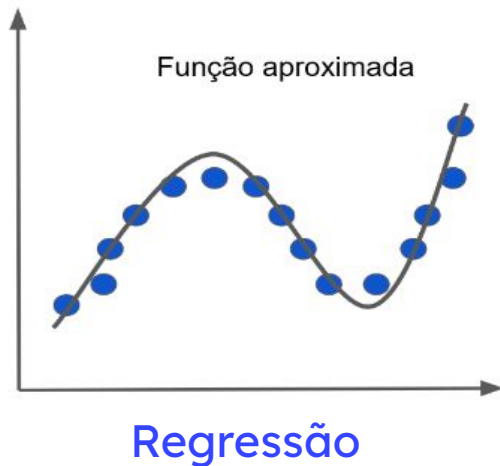


Aprendizado Não Supervisionado

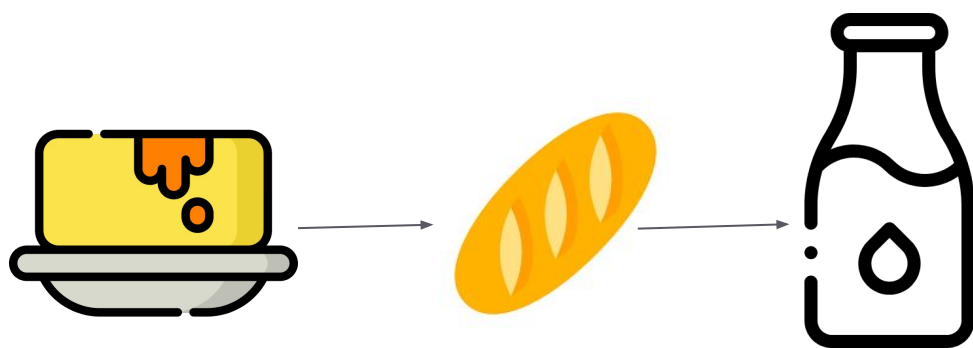


Aprendizado Por Reforço

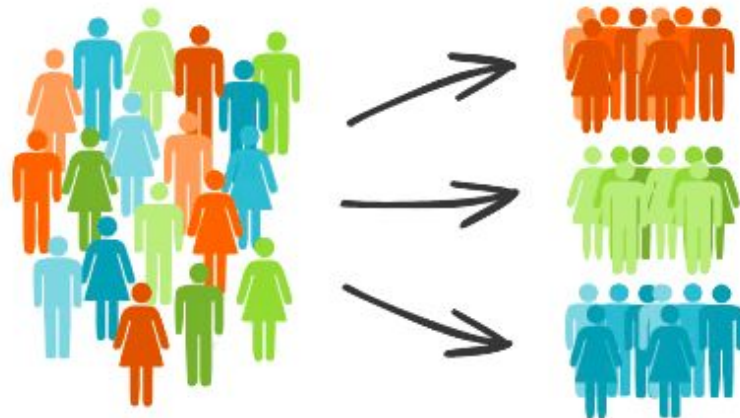
Aprendizado Supervisionado



Aprendizado Não Supervisionado

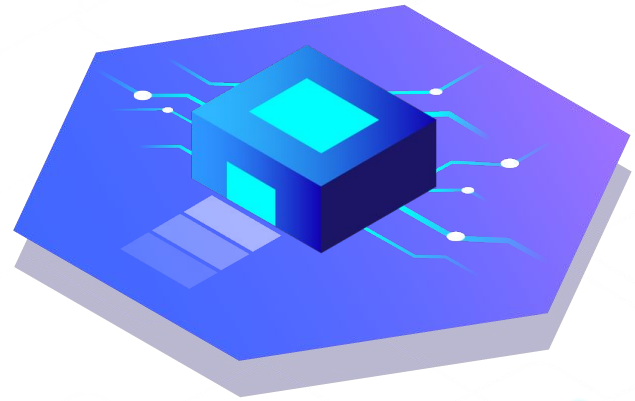


Regras de Associação

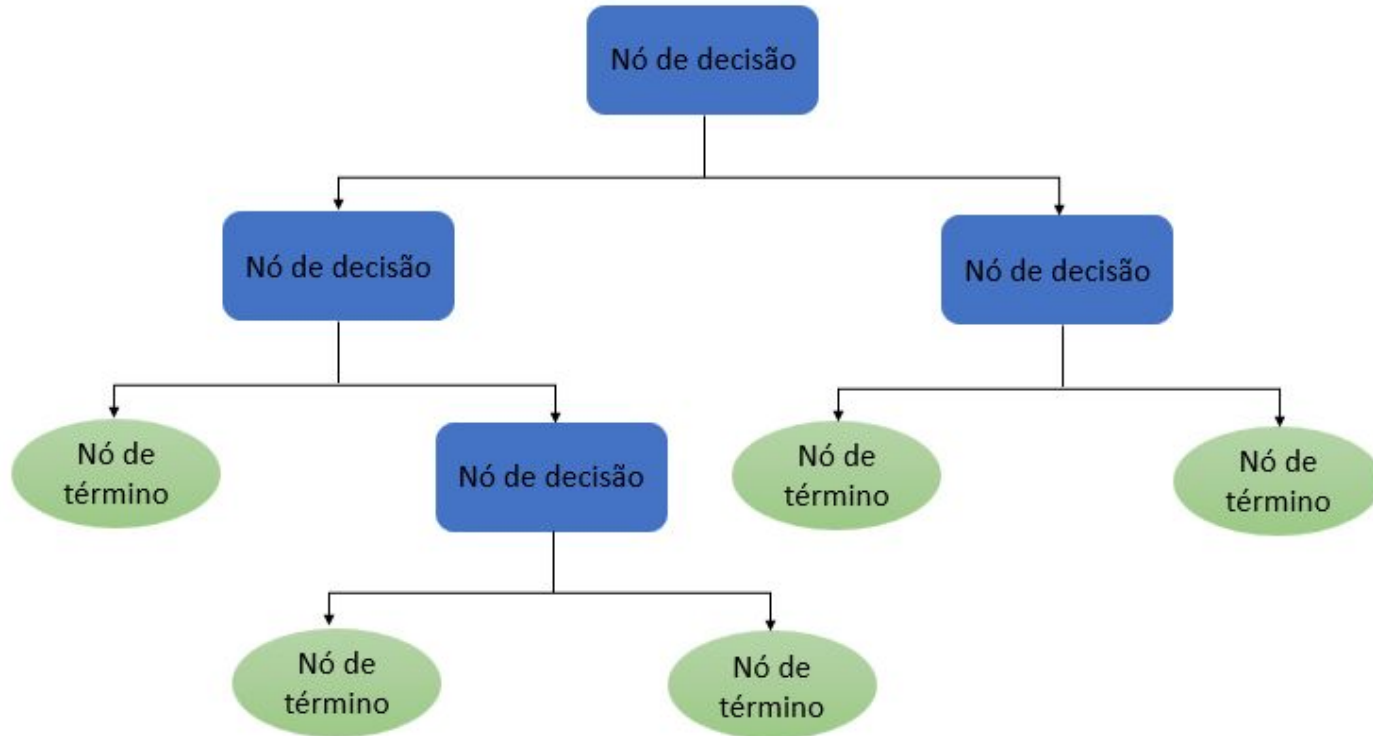


Agrupamento

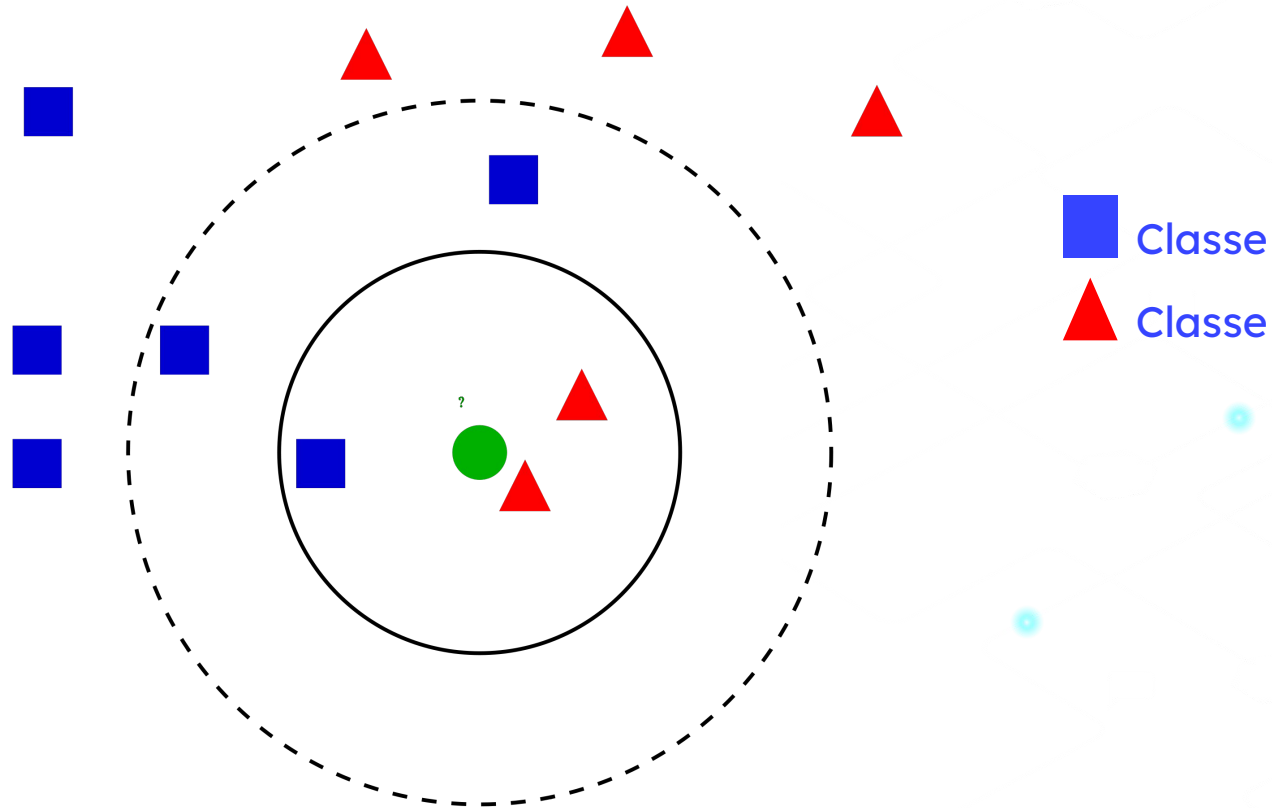
8. Algoritmos do Aprendizado Supervisionado



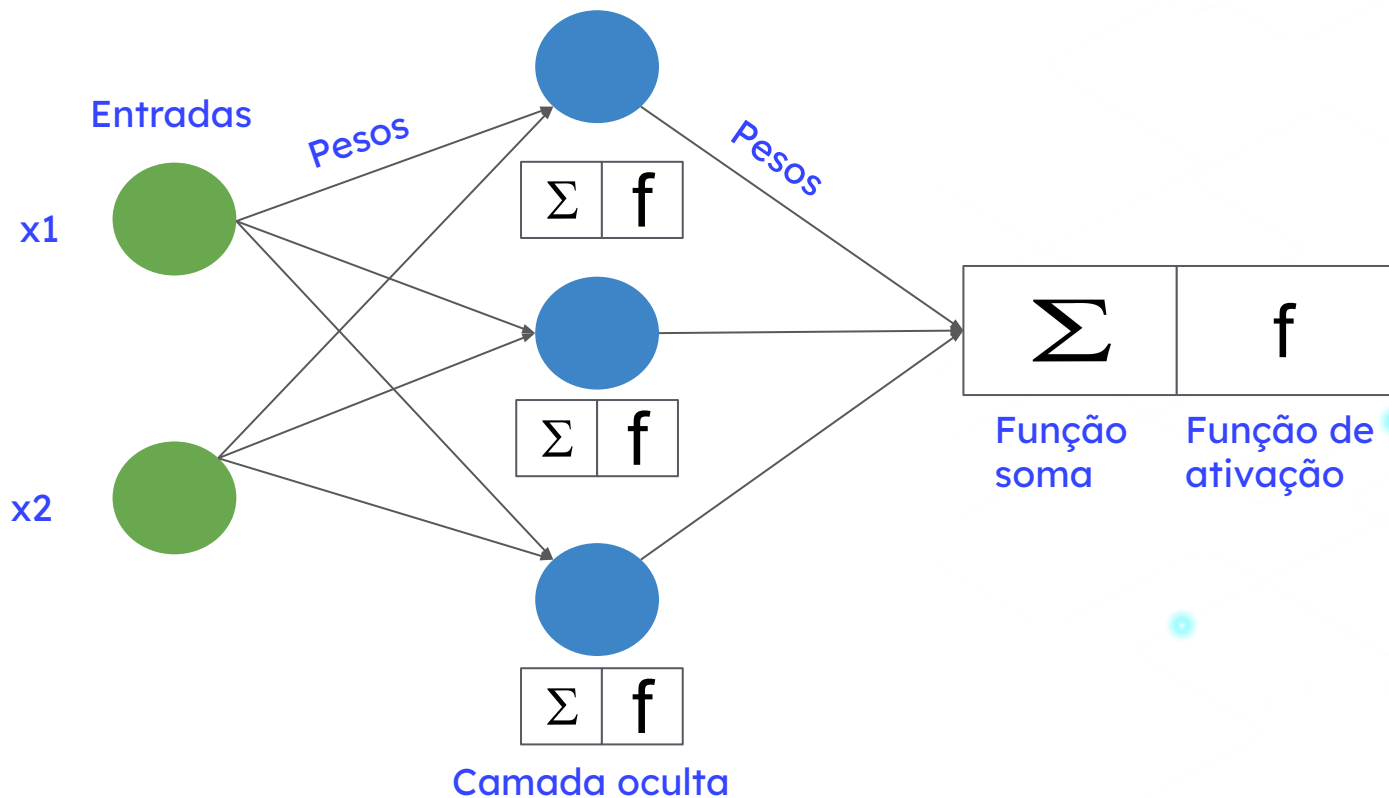
Árvore de Decisão



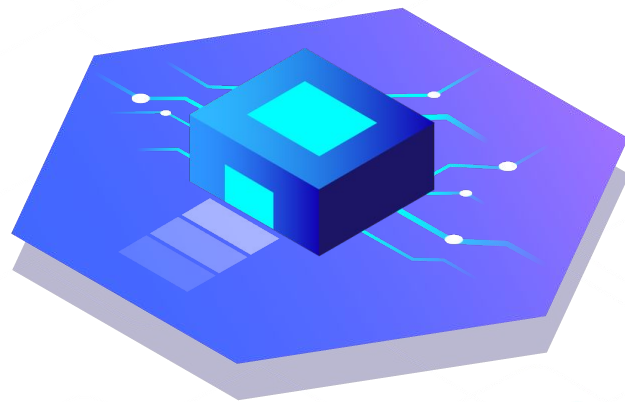
K-Vizinhos Mais Próximos (KNN)

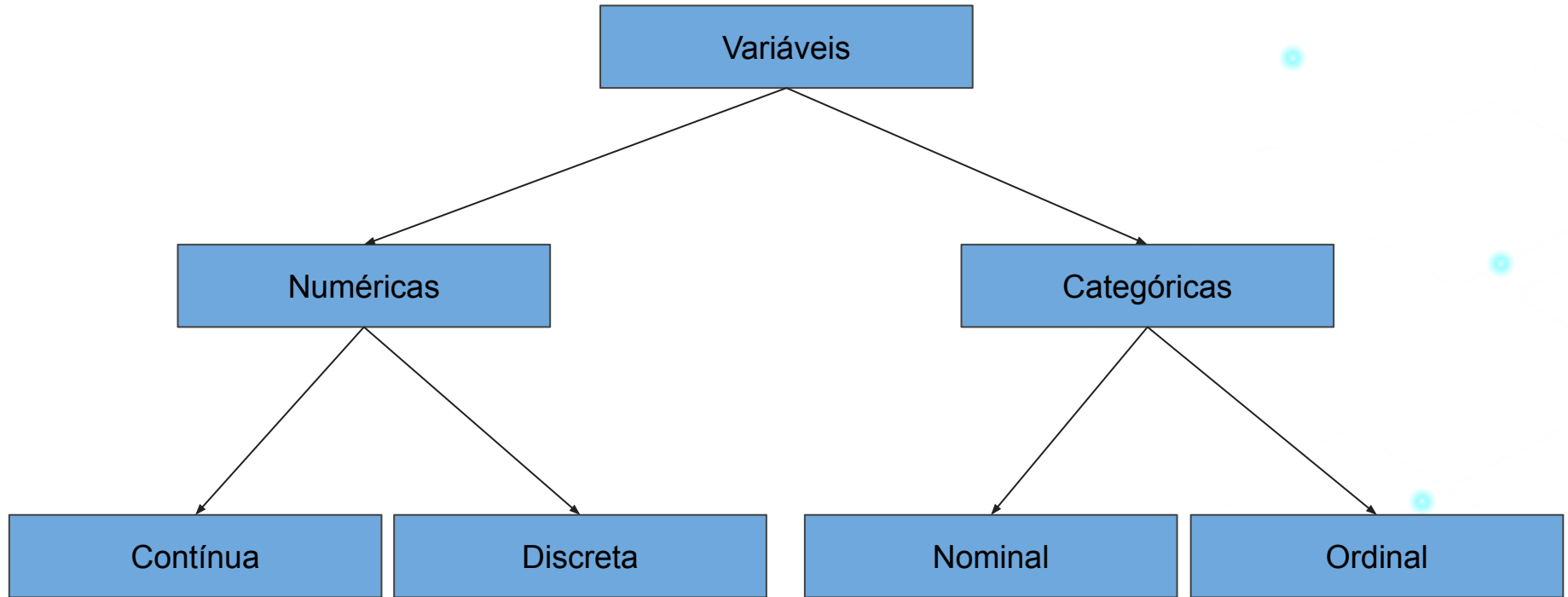


Rede Neural Perceptron



9. Tipos de Variáveis





Números reais:
Temperatura, altura, peso,
salário

Conjunto de valores finito
(inteiro):
Contagem de alguma coisa,
1,2,3

Dados não mensuráveis
sem ordenação: cor dos
olhos, gênero

Categorizado sob uma
ordenação:
Tamanho P, M e G

Vamos ao código?!

