



GIRLS SUPPORT GIRLS BH

INTRO MACHINE LEARNING

Introdução

Entendendo o Aprendizado de Máquina

Projeto

Resolvendo um problema na prática!

Estão preparadas?

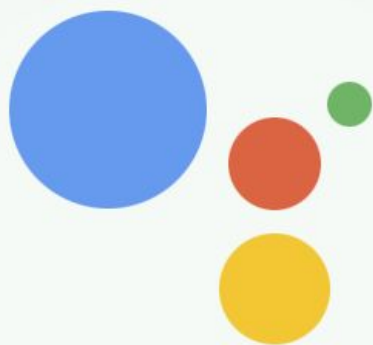
Computadores estão com Jupyter Notebook instalado?

Todo mundo com o arquivo enviado por email?

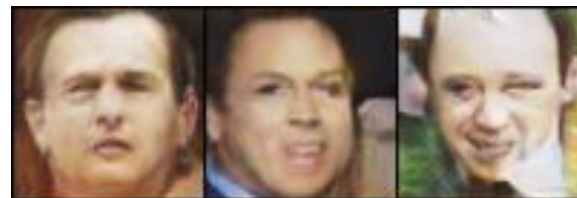
Requerimentos

- Linux ou bash para windows
- Python3
- pip
- jupyter notebook
- scikit-learn
- numpy
- pandas

O quê e pra quê



*“Hi, I'm calling to book a
women's haircut for a client.”*



O que é Aprendizado de Máquina?

É a ciência de fazer computadores tomarem
decisões sem terem sido explicitamente
programados para isso

exemplo

Queremos descobrir se amanhã chove, sabendo a temperatura e a umidade

Explicitamente programado

Tarefa: Decidir se amanhã vai chover

Input: temperatura e umidade de hoje

Processo: Após longo estudo científico dos meteorologistas sobre a relação entre a umidade, a temperatura e a probabilidade de chover, você coloca isso na forma de código

```
func probabilidade_de_chover ( temperatura, umidade )  
    if.....  
        [faça isso]  
    else  
        [faça isso]  
  
    retorna se chove ou não
```

Aprendizado de Máquina

Programa: Decidir se amanhã vai chover

Input: Dataset com **várias** temperaturas e umidades, bem como a **resposta** de se choveu ou não no dia seguinte

Processo: O algoritmo vai modelar uma função que relaciona cada dupla (umidade, temperatura) à resposta certa

$$f(u, t) = \{chove, \text{ não chove} \}$$

Explicitamente programado

Precisamos de: a regra que mapeia umidade e temperatura à resposta certa

O programa entrega: a resposta certa

Aprendizado de máquina

Precisamos de: a resposta certa

O programa entrega: a regra que mapeia umidade e temperatura à resposta certa

Explicitamente programado

Precisamos de: a regra que mapeia umidade e temperatura à resposta certa

O programa entrega: a resposta certa

DETERMINÍSTICO

Aprendizado de máquina

Precisamos de: a resposta certa

O programa entrega: a regra que mapeia umidade e temperatura à resposta certa

PROBABILÍSTICO

WHO NEEDS COPERNICUS IF YOU HAVE MACHINE LEARNING?

It took humanity centuries to decide that Earth orbits the sun. Now a neural network has come to the same conclusion, using the same data, in just a few hours.

Aprendizado de máquina:

- Automatiza a descoberta de padrões matemáticos
- Reduz a dependência do conhecimento de domínio

Como?

$$f(u, t) = \{\text{chove, não chove}\}$$

- 1 **Cria Hipótese:** Atribui pesos para as variáveis umidade e temperatura
- 2 *hipótese = P1, P2, como em P1*umidade + P2*temperatura*
- 3
- 4 Para cada linha do dataset:
- 5 Calcula a saída da função hipótese
- 6 **Função de ativação:** Transforma a saída da hipótese no palpite de chover ou não
- 7 Verifica a resposta
- 8 **Função erro:** Avalia o quanto ele errou
- 9 **Função de otimização:** Altera os pesos da hipótese

$$f(u, t) = \{\text{chove}, \text{não chove}\}$$

- 1 **Cria Hipótese:** Atribui pesos para as variáveis umidade e temperatura
- 2 *hipótese = P1, P2, como em P1*umidade + P2*temperatura*
- 3
- 4 Para cada linha do dataset:
- 5 Calcula a saída da função hipótese
- 6 **Função de ativação:** Transforma a saída da hipótese no palpite de chover ou não
- 7 Verifica a resposta
- 8 **Função erro:** Avalia o quanto ele errou
- 9 **Função de otimização:** Altera os pesos da hipótese

Ao final do treino, o algoritmo terá tido 'x' oportunidades de melhorar os pesos

Por último, devemos avaliar o desempenho do modelo por meio de uma **Métrica de avaliação**

Features

Cada tipo de variável que vai ser analisada

f1	f2	f3	f4
Outlook	Temperature	Humidity	Windy
Rainy	Hot	High	False
Rainy	Hot	High	True
Overcast	Hot	High	False

Parâmetros

Pesos atribuídos a cada feature

P1 * f1;

P2 * f2;

P3 * f3

Hiperparâmetro

Decisões que a programadora toma

Quanto tempo o algoritmo deve rodar

Qual função erro utilizar

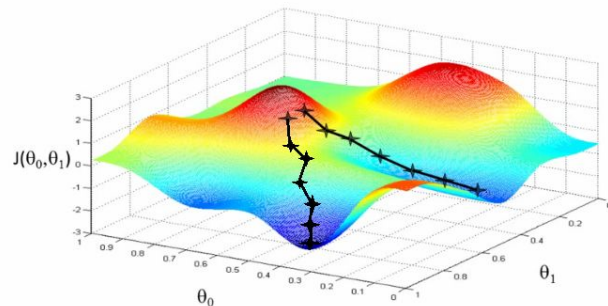
...

Função Hipótese

Função que o modelo dá como saída. A hipótese muda cada vez que a função de otimização é acionada

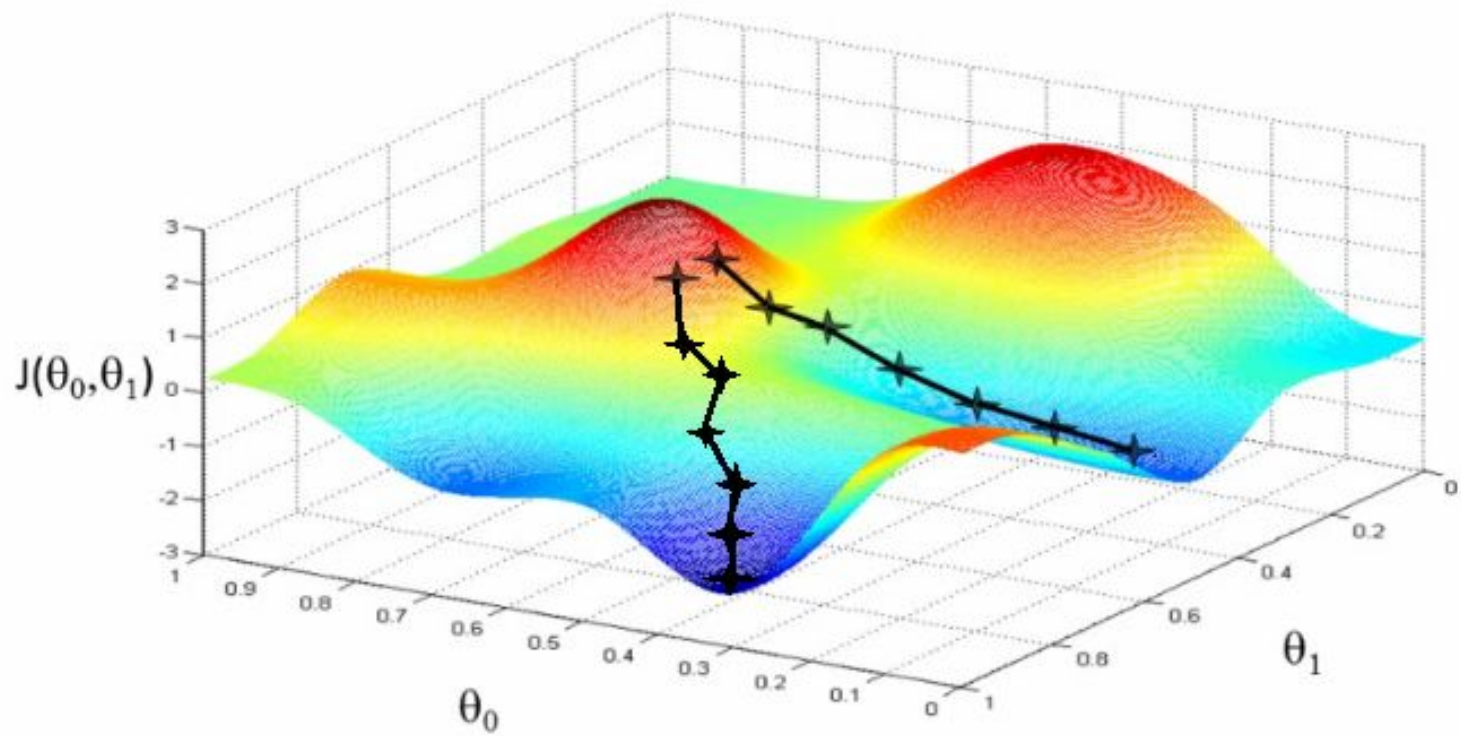
Função erro ou função custo

Métrica que avalia o quanto a função hipótese está longe do correto. É essa função que queremos matematicamente minimizar.



Função de Otimização

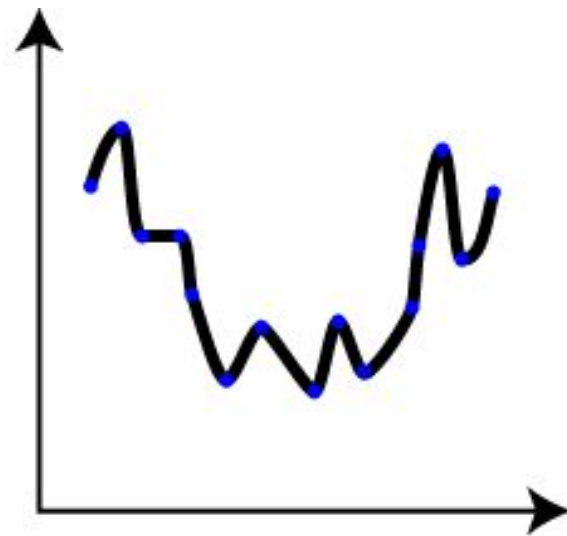
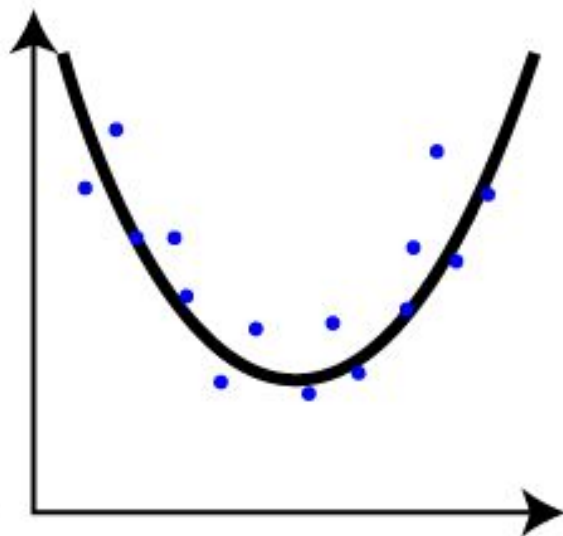
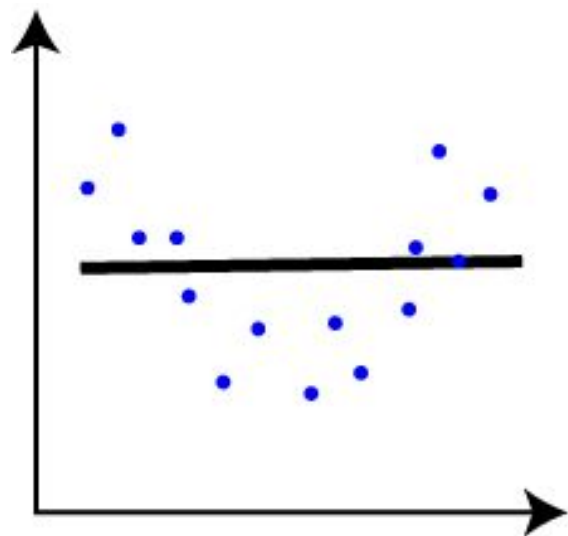
Estratégia utilizada para corrigir os pesos. O objetivo dessa função é achar um mínimo da função erro.

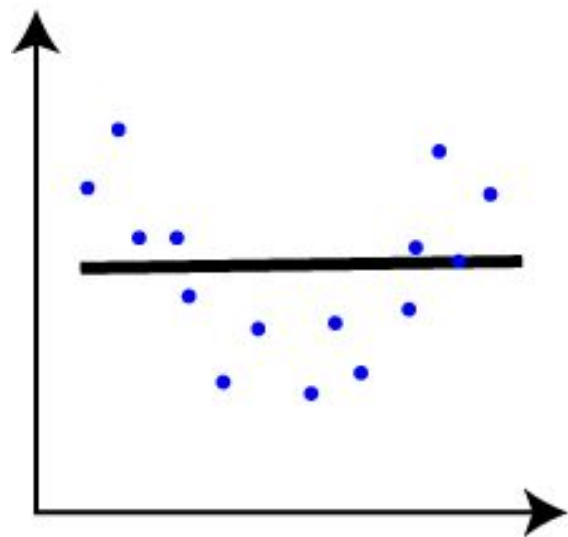


Atenção

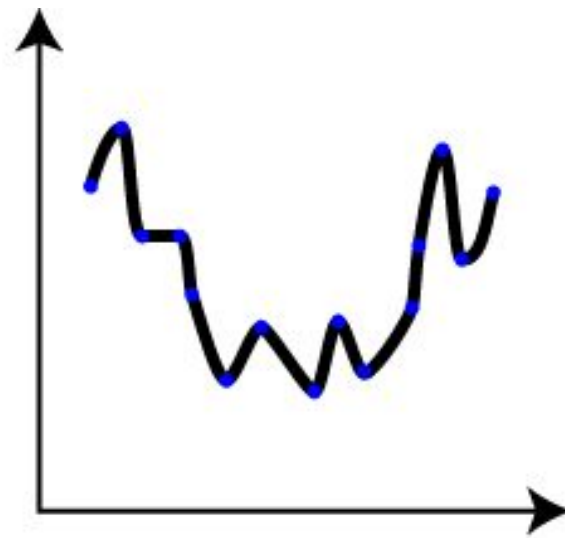
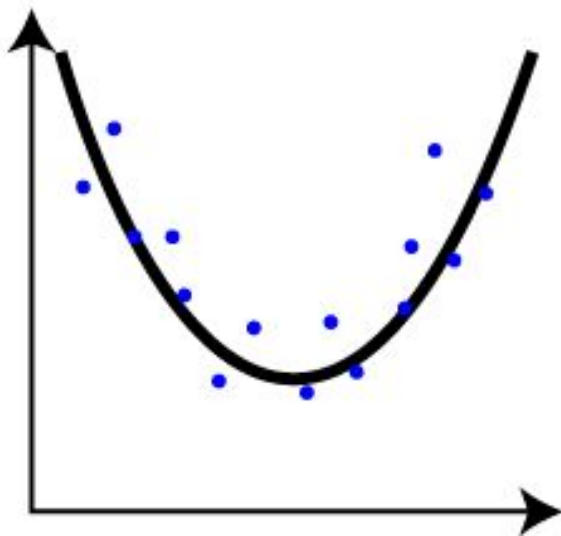
O processo de ajustes sucessivos dos pesos (parâmetros) para achar uma função que modela os dados é chamado de **“fitting”**

Do inglês: ajustar, modelar

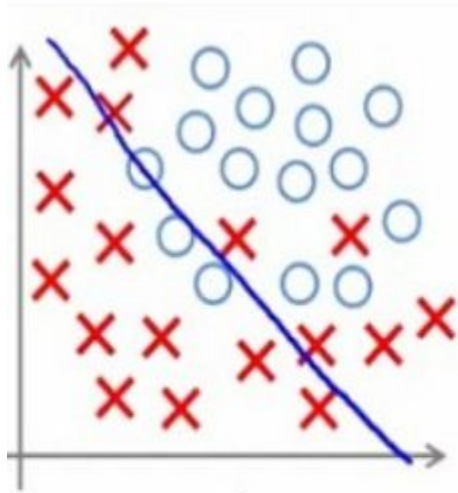




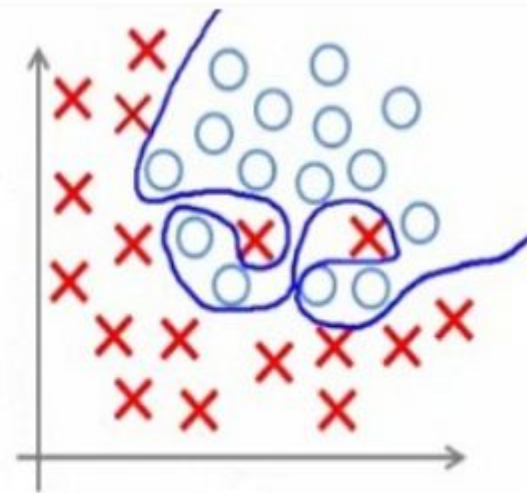
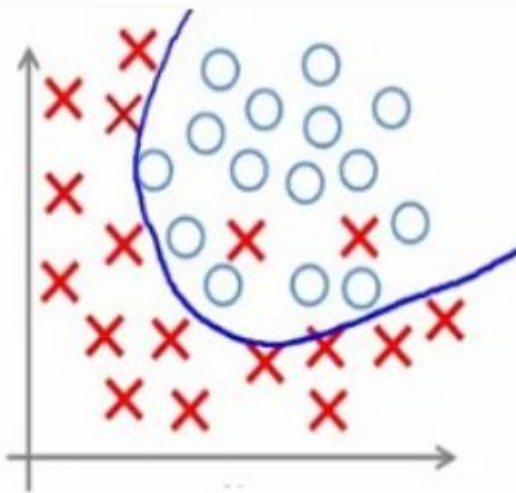
underfitting



overfitting

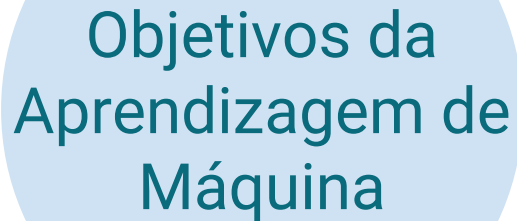


underfitting



overfitting

Tipos de Machine Learning



Objetivos da Aprendizagem de Máquina

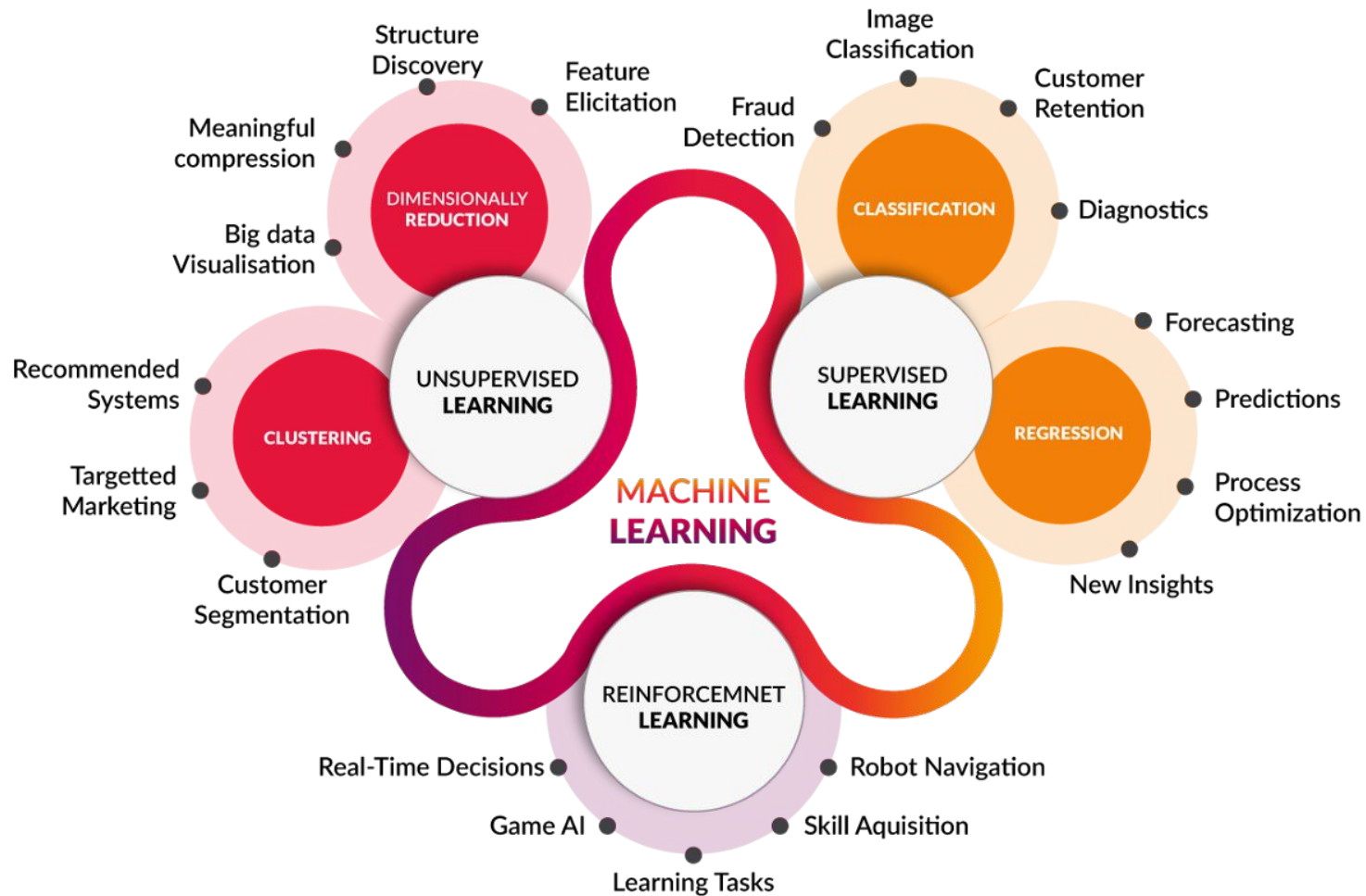
Identificar a
importância de cada
feature e simplificar
dados

Classificar imagem,
texto

Fazer previsões
de valores futuros

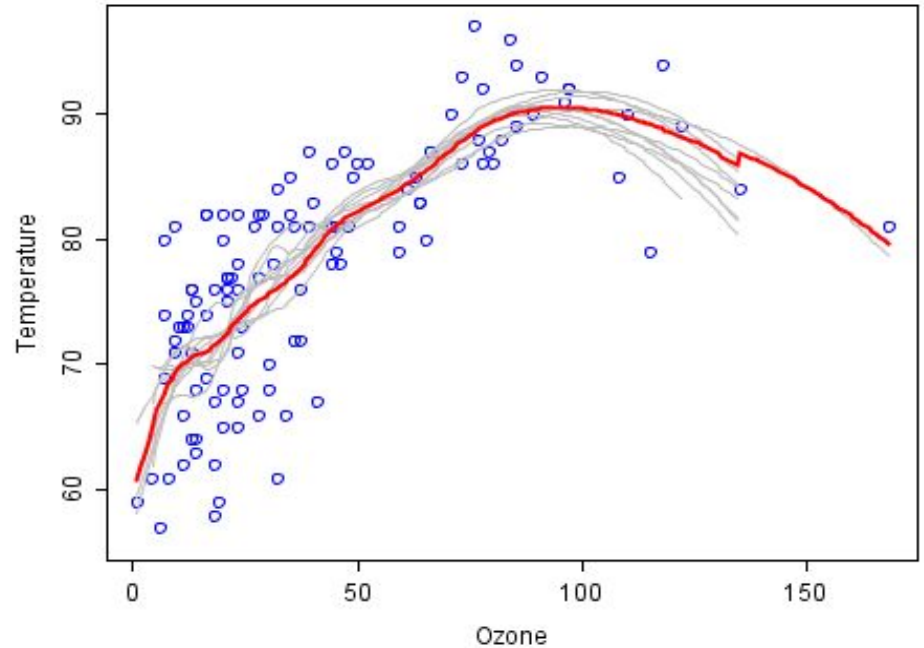
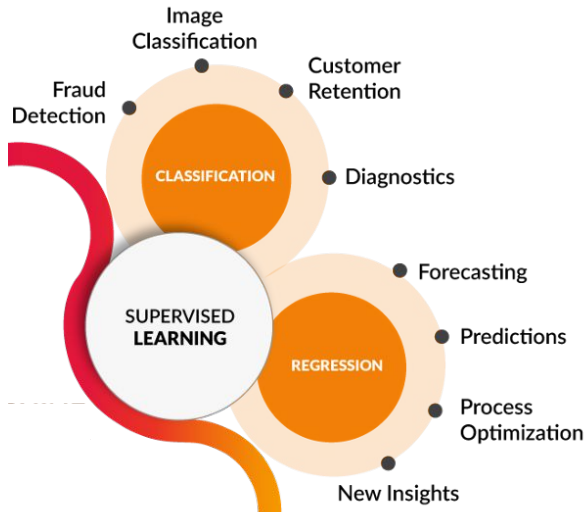
Gerar novos objetos
e comportamentos
virtuais

Identificar padrões nos
dados



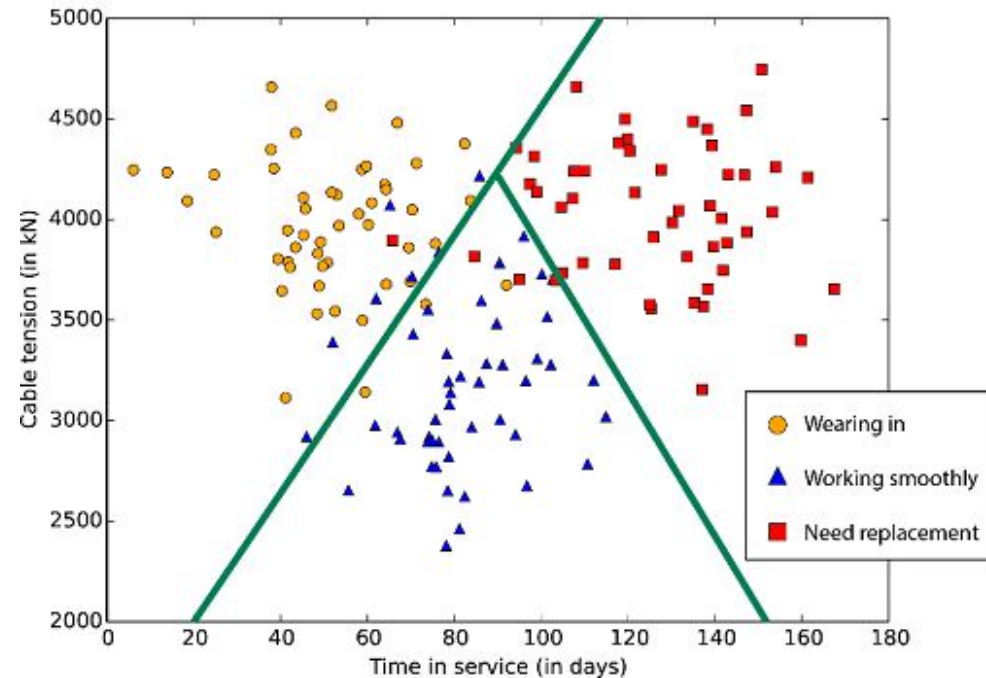
Regressão

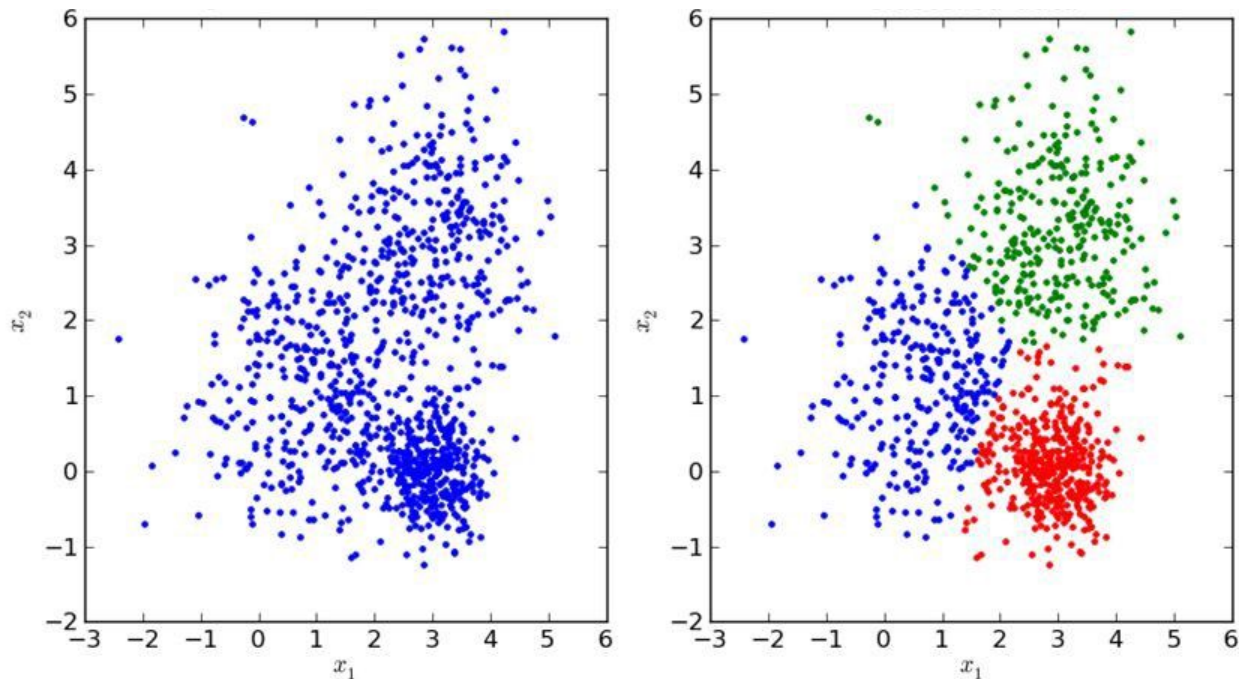
A partir das respostas fornecidas, modela uma função que aproxima como as features se relacionam com a variável dependente. Objetivo: Fazer predição de valores futuros.



Classificação

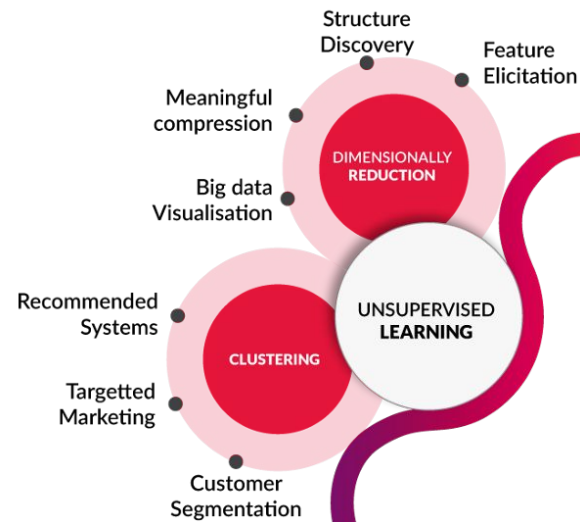
A partir das respostas fornecidas, modela uma função que aproxima como as features se relacionam com cada categoria possível.
Objetivo: Classificar itens

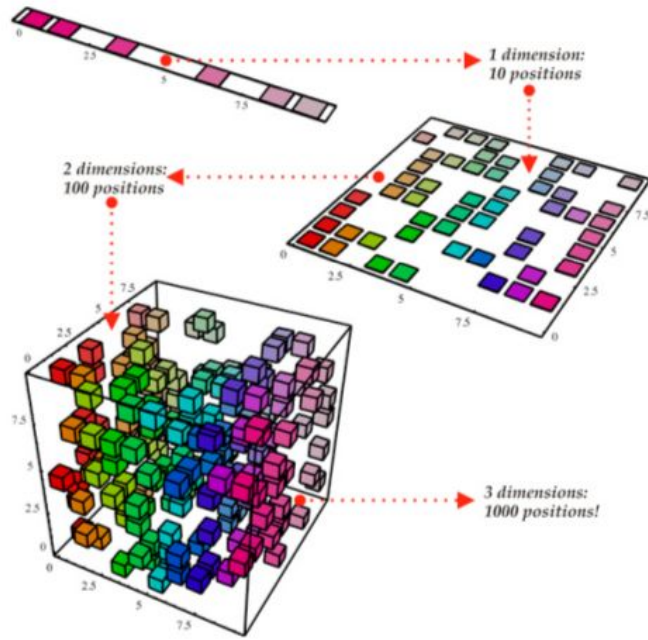




Clustering

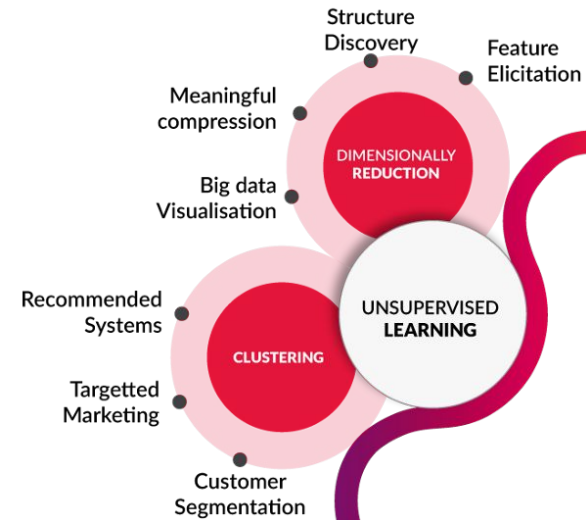
Identifica sozinho padrões nos dados gerando suas próprias categorias. Reage à presença das características em novos dados. Objetivo: descobrir padrões naturais ou identificar anomalias nos dados



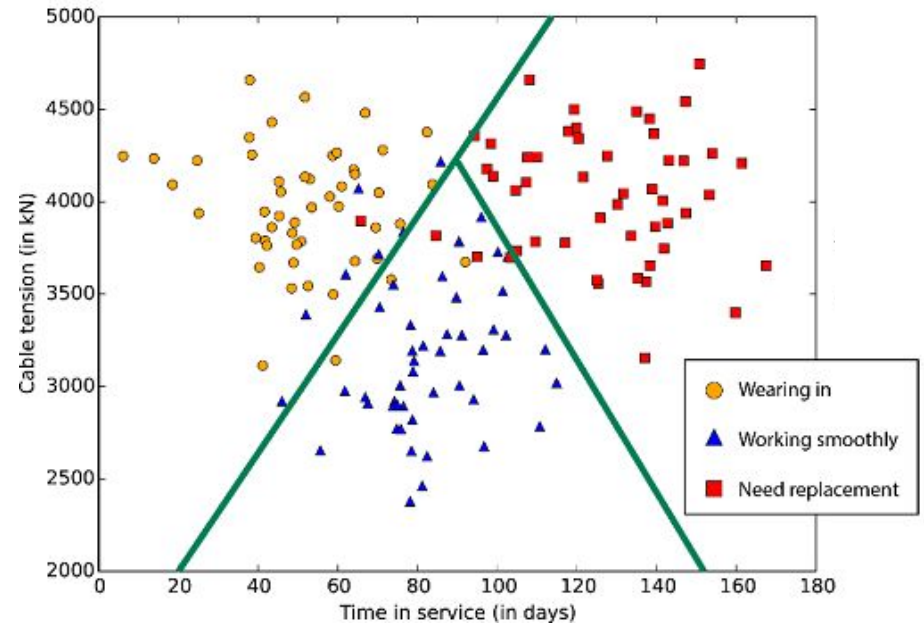
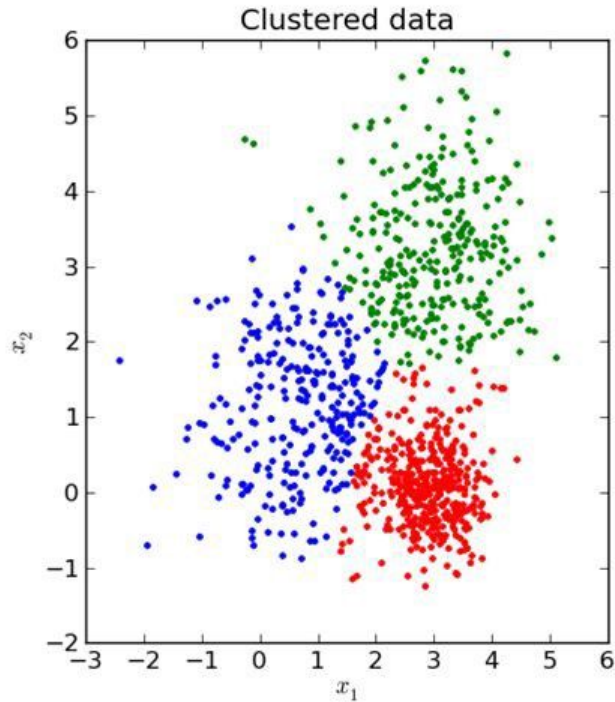


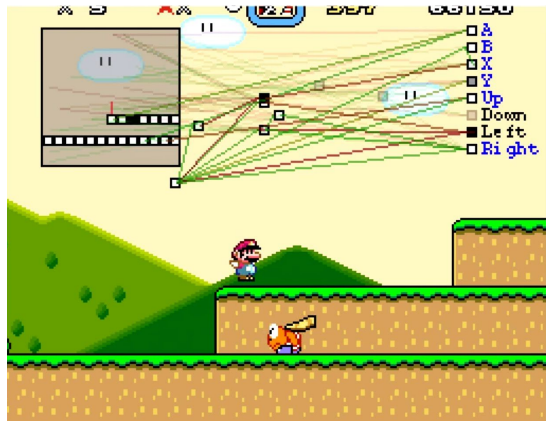
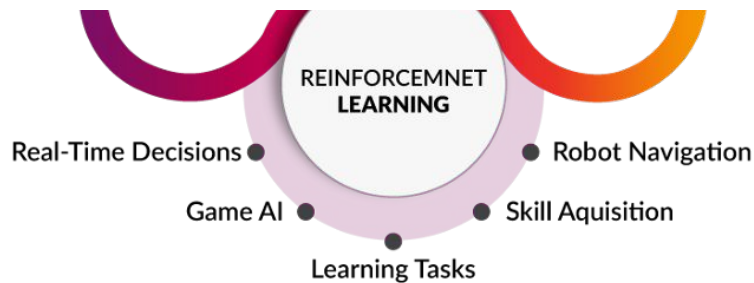
Dimensionality Reduction

Identifica sozinho quais features melhor representam os dados, reduzindo a complexidade dos dados, mas mantendo representatividade. Objetivo: processamento mais rápido sem perda de qualidade; visualização de dados acima de 3D



Qual é a diferença?





Aprendizagem por Reforço

A partir de um **Objetivo** dado ao programa, bem como **regras de punição e recompensa**, o agente decide sozinho as **Ações** a serem executadas.

Jogar videogames

Investir em ações

Carros autônomos

[Todo tipo de ação ao vivo]

Pipeline



1

**Entender o
problema**

2

**Obter e Tratar
dados**

3

**Decidir modelo de
ML a ser usado**

4

**Treinar, testar e
avaliar**

1

Entender o problema

2

Obter e Tratar dados

3

Decidir modelo de ML a ser usado

4

Treinar, testar e avaliar

Obter e Tratar Dados

- Valores nulos, incoerentes
- Transformar dados categóricos
- Seleção e extração de features
- Redução de dimensionalidade

1

Entender o
problema

2

Obter e Tratar
dados

3

**Decidir modelo de
ML**

4

Treinar, testar e
avaliar

Decidir modelo de ML

- Qual tipo de Machine Learning?
- Quais técnicas testar?
- Decidir funções de **ativação**, função **erro** e função de **otimização**
- Decidir hiperparâmetros específicos do seu modelo

1

Entender o
problema

2

Obter e Tratar
dados

3

Decidir modelo de
ML a ser usado

4

**Treinar, testar e
avaliar**

Treinar, testar e avaliar

- Técnicas de medir e reduzir overfitting
- Decidir métricas de **avaliação**
- Rodar o modelo

Vamos ver na prática?