

Introdução ao Aprendizado de Máquina



Gabriel P. Oliveira

Inteligência Artificial
2026/1

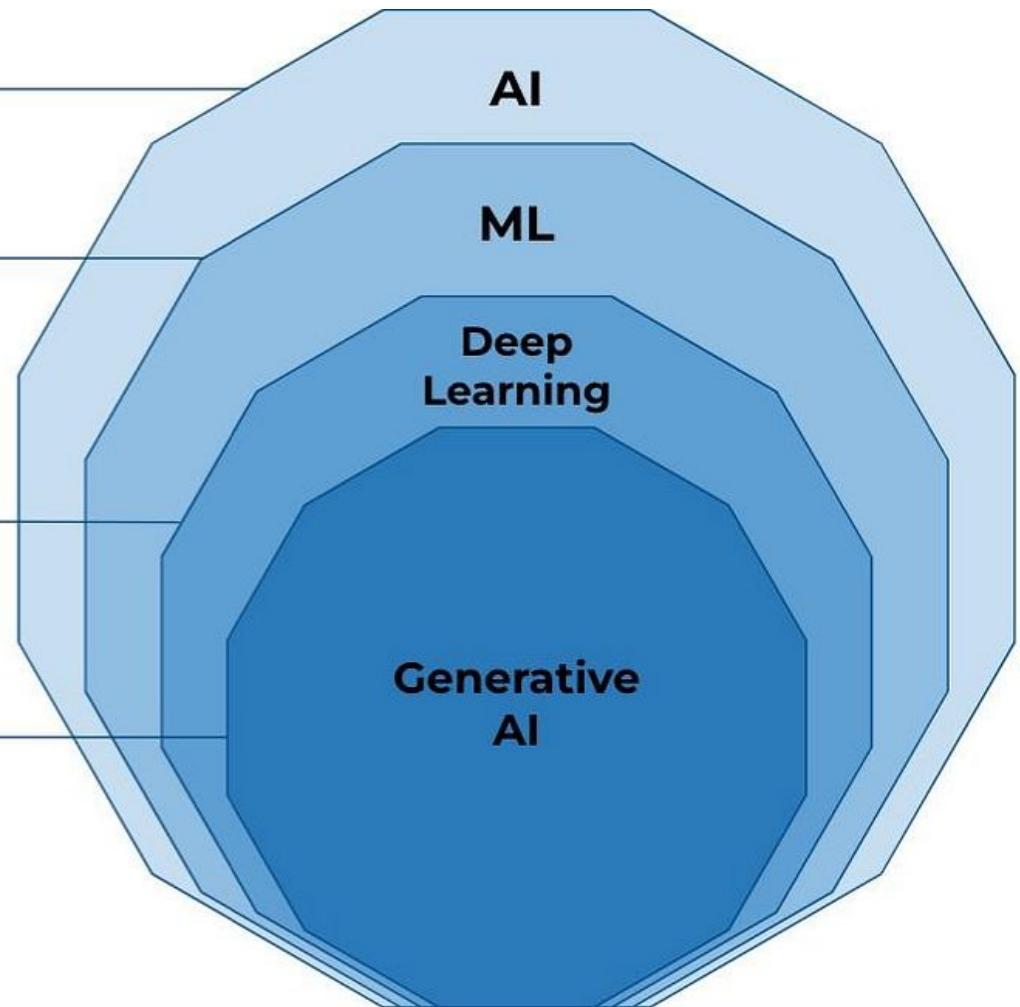
AI? ML? DL? Generative AI?

Artificial Intelligence

Machine Learning

Deep Learning

Generative AI



AI? ML? DL? Generative AI?

Artificial Intelligence

Is the field of study

Machine Learning

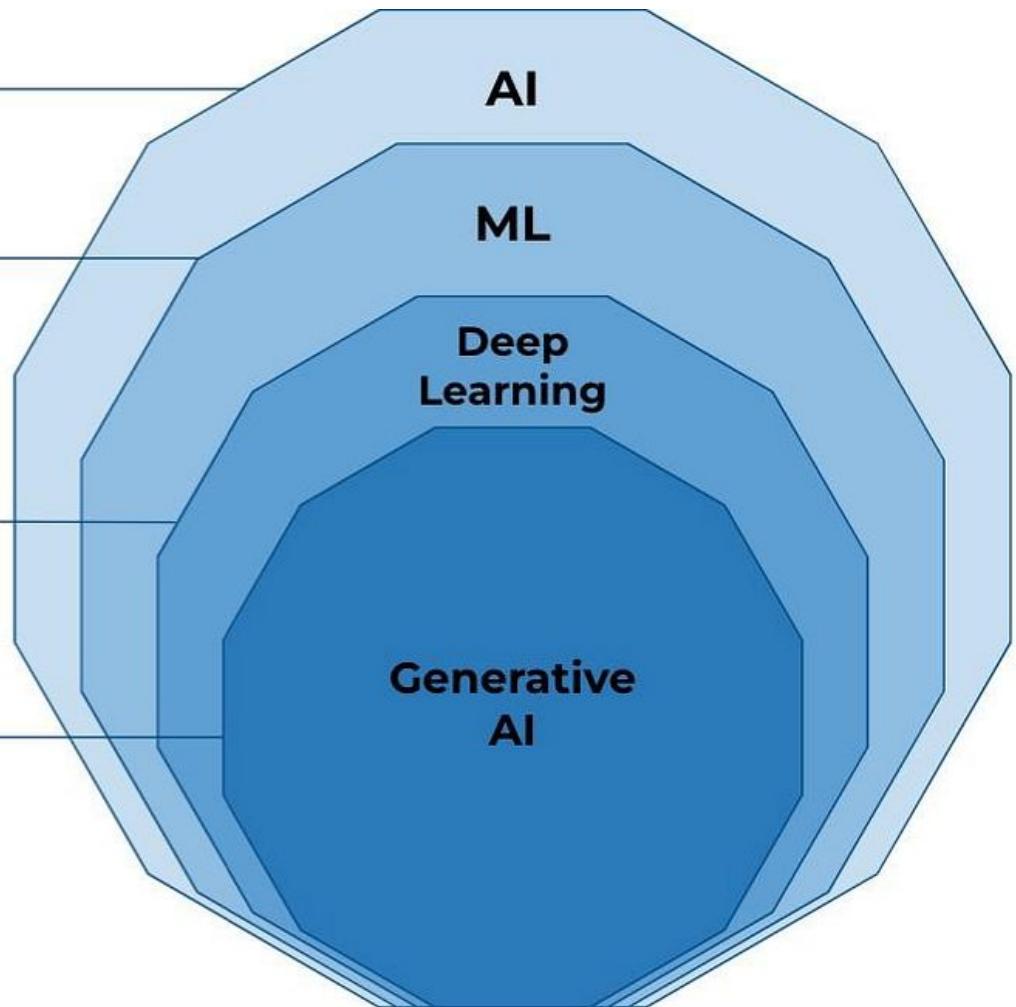
Is a branch of AI that focus on the creation of intelligent machines that learn from data. Another very well known branch inside AI is **Optimization**.

Deep Learning

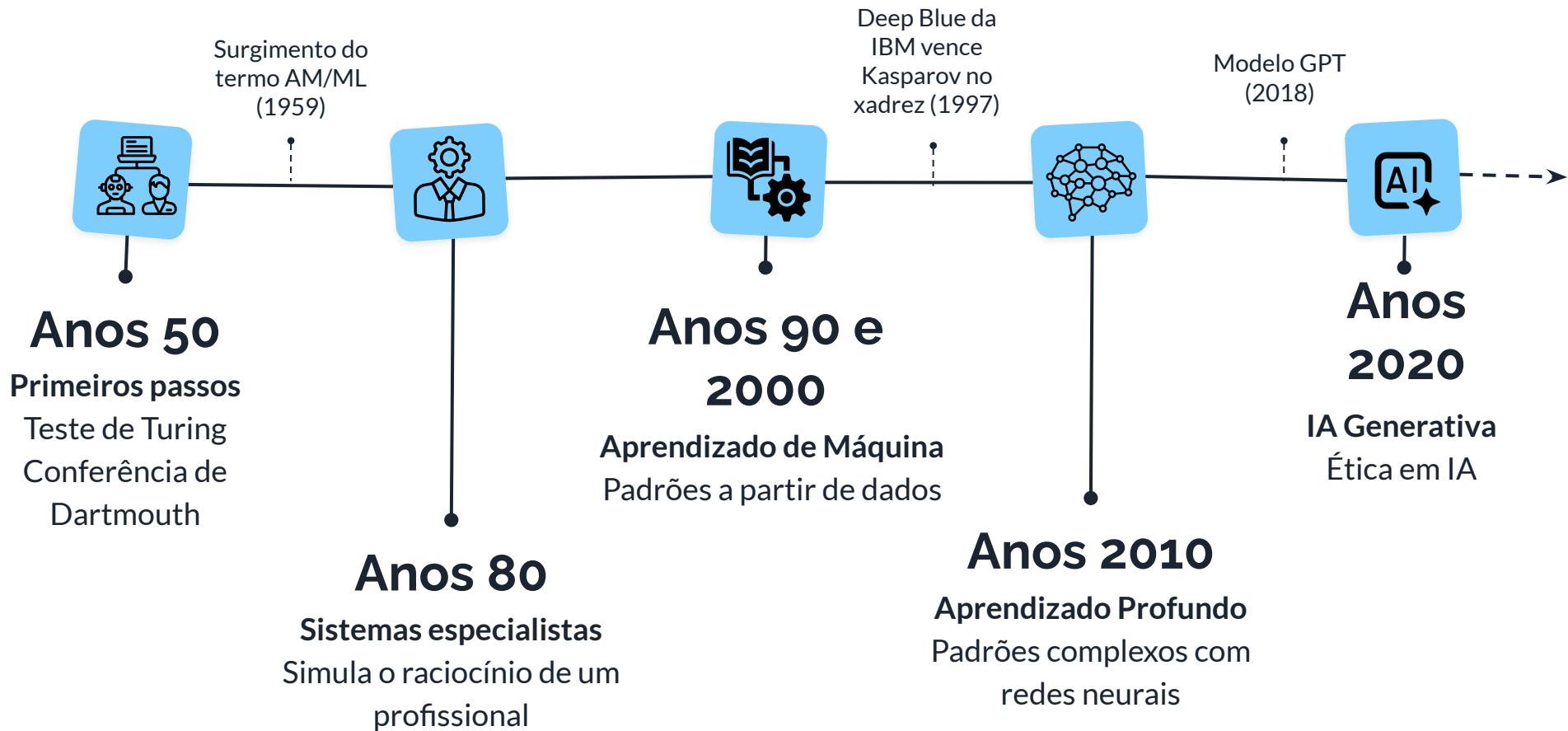
Is a subset of Machine Learning methods, based on **Artificial Neural Networks**. Examples: CNNs, RNNs

Generative AI

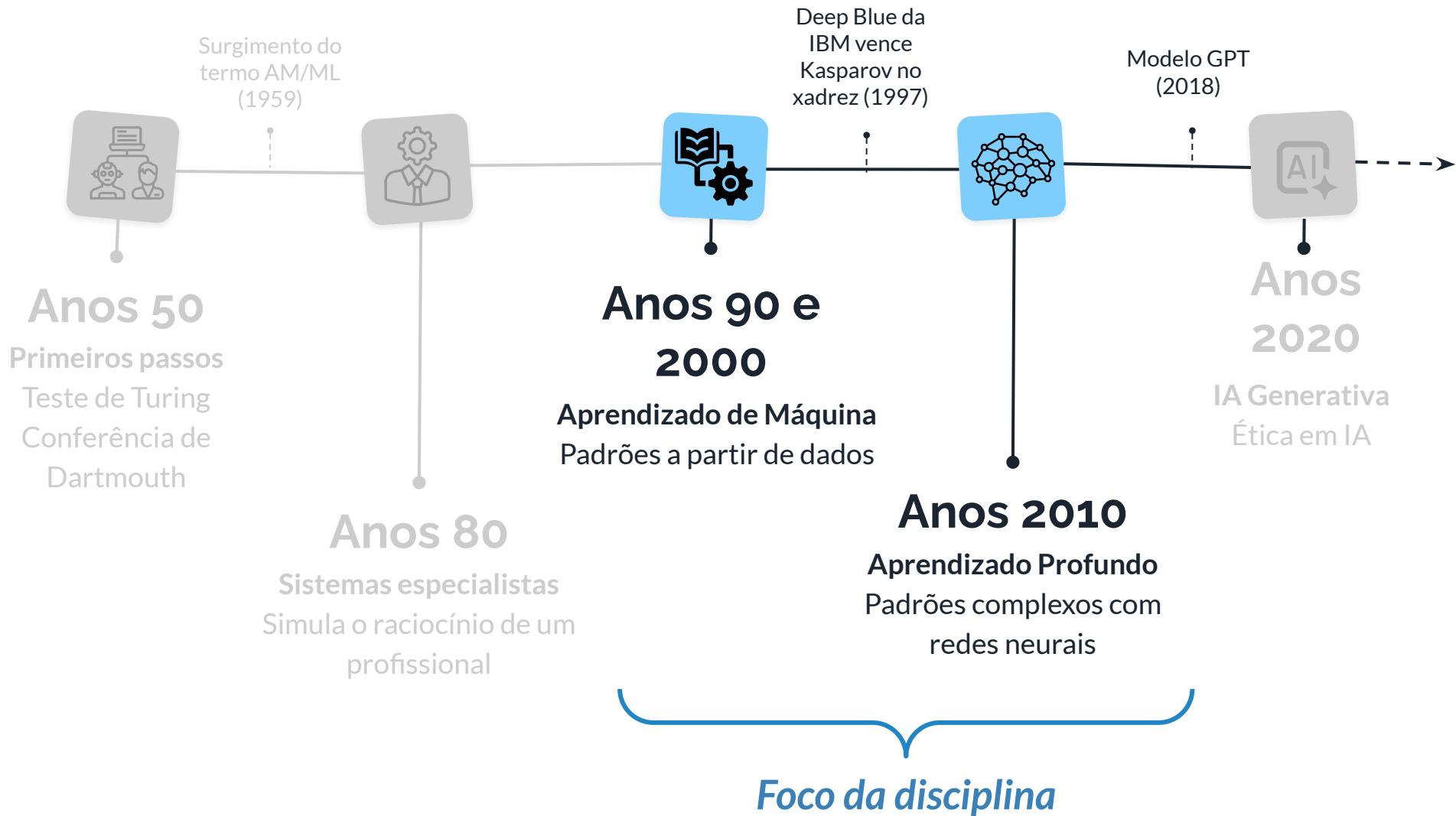
A type of ANNs that generate data that is similar to the data it was trained on. Examples: GANs, LLMs



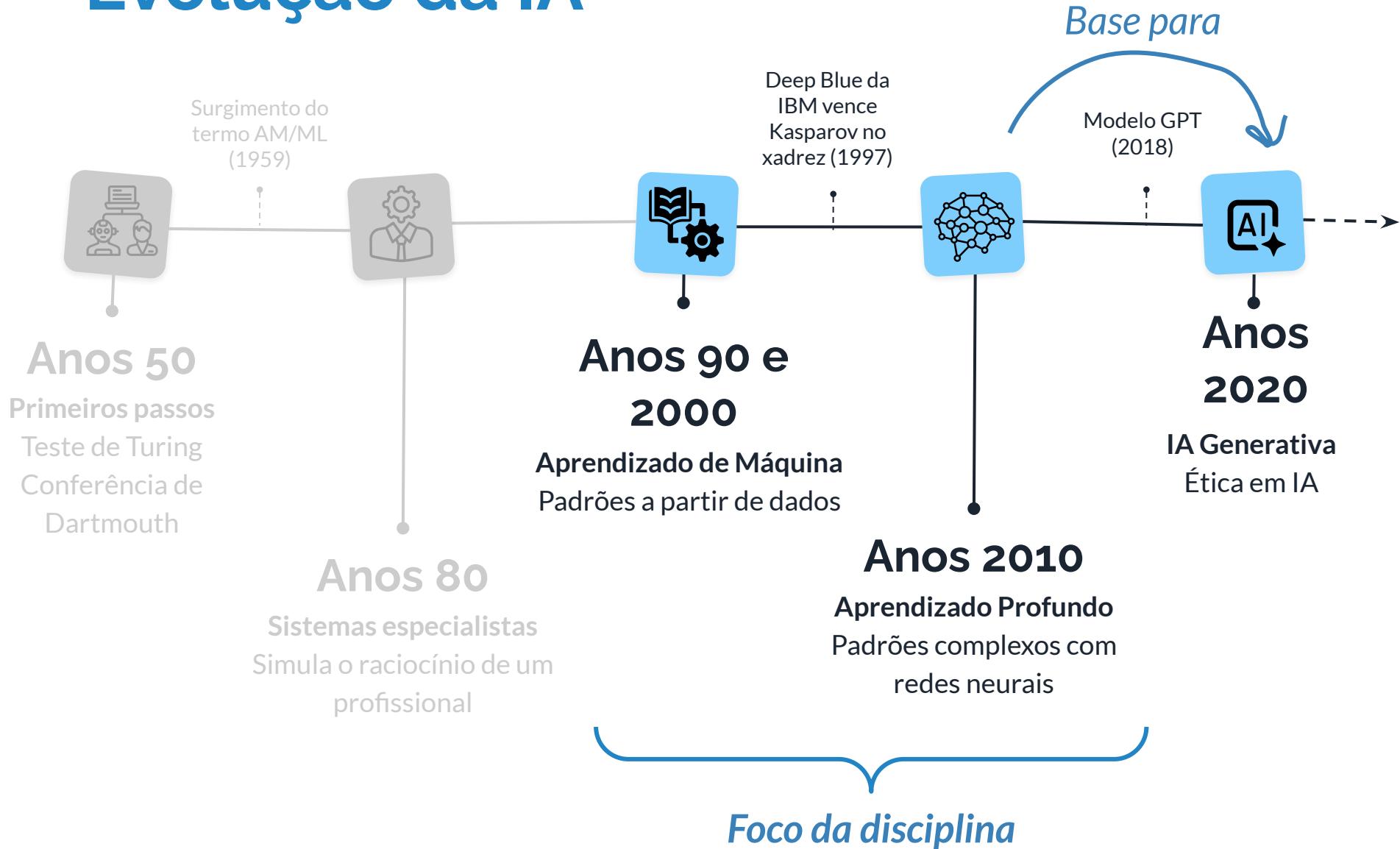
Evolução da IA



Evolução da IA



Evolução da IA



1.

O que é Aprendizado de Máquina (AM)?

Definições e tarefas

Algumas definições de AM

Termo introduzido por Arthur Samuel (1959)

“A capacidade de melhorar o desempenho na realização de alguma **tarefa** por meio da **experiência**.” Mitchell (1997)

“Processo que permite aos computadores aprenderem **autonomamente**, identificando **padrões** e tomando decisões **baseadas em dados**.” Encyclopédia Britannica (2026)

Algumas definições de AM

Termo introduzido por Arthur Samuel (1959)

“A capacidade de melhorar o desempenho na realização de alguma **tarefa** por meio da **experiência**.” Mitchell (1997)

“Processo que permite aos computadores aprenderem **autonomamente**, identificando **padrões** e tomando decisões **baseadas em dados**.” Encyclopédia Britannica (2026)

Indução de Hipóteses

P: Como aprender autonomamente?

Indução de Hipóteses

P: Como aprender autonomamente?

Indução de hipótese (regra ou função) a partir de dados

*Treinamento
Exemplos*

Indução de Hipóteses

P: Como aprender autonomamente?

Indução de hipótese (regra ou função) a partir de dados

*Treinamento
Exemplos*

É desejável que a hipótese seja válida para outras instâncias do problema que não foram vistas ainda

*Teste
Novidade*

Na prática

Aprovação de empréstimo

Imagine que um banco quer prever se deve **aprovar ou rejeitar** um pedido de empréstimo com base em alguns atributos do cliente.

Na prática

Aprovação de empréstimo

Imagine que um banco quer prever se deve **aprovar ou rejeitar** um pedido de empréstimo com base em alguns atributos do cliente.

Quais atributos seriam importantes de serem avaliados?

Na prática

Aprovação de empréstimo

Imagine que um banco quer prever se deve **aprovar ou rejeitar** um pedido de empréstimo com base em alguns atributos do cliente.

Quais atributos seriam importantes de serem avaliados?

Valor da dívida, renda, idade, histórico, etc.

Na prática

Aprovação de empréstimo – Treinamento

Cliente	Idade	Renda	Histórico de Crédito	Valor da Dívida	Aprovado?
c1	25	3000	Bom	500	Sim
c2	45	8000	Bom	1000	Sim
c3	23	2000	Ruim	2500	Não
c4	35	5000	Bom	700	Sim
c5	52	9000	Ruim	4000	Não
c6	28	3500	Ruim	1800	Não
c7	40	6500	Bom	900	Sim

Na prática

Aprovação de empréstimo – Treinamento

Instâncias

Cliente	Idade	Renda	Histórico de Crédito	Valor da Dívida	Aprovado?
c1	25	3000	Bom	500	Sim
c2	45	8000	Bom	1000	Sim
c3	23	2000	Ruim	2500	Não
c4	35	5000	Bom	700	Sim
c5	52	9000	Ruim	4000	Não
c6	28	3500	Ruim	1800	Não
c7	40	6500	Bom	900	Sim

Na prática

Aprovação de empréstimo – Treinamento

Instâncias

Cliente	Idade	Renda	Histórico de Crédito	Valor da Dívida	Aprovado?
c1	25	3000	Bom	500	Sim
c2	45	8000	Bom	1000	Sim
c3	23	2000	Ruim	2500	Não
c4	35	5000	Bom	700	Sim
c5	52	9000	Ruim	4000	Não
c6	28	3500	Ruim	1800	Não
c7	40	6500	Bom	900	Sim

Atributos de entrada
Features

Na prática

Aprovação de empréstimo – Treinamento

Instâncias

Cliente	Idade	Renda	Histórico de Crédito	Valor da Dívida	Aprovado?
c1	25	3000	Bom	500	Sim
c2	45	8000	Bom	1000	Sim
c3	23	2000	Ruim	2500	Não
c4	35	5000	Bom	700	Sim
c5	52	9000	Ruim	4000	Não
c6	28	3500	Ruim	1800	Não
c7	40	6500	Bom	900	Sim

Atributo de saída
Target (alvo)

Na prática

Aprovação de empréstimo – Teste

Cliente	Idade	Renda	Histórico de Crédito	Valor da Dívida	Aprovado?
c1	25	3000	Bom	500	Sim
c2	45	8000	Bom	1000	Sim
c3	23	2000	Ruim	2500	Não
c4	35	5000	Bom	700	Sim
c5	52	9000	Ruim	4000	Não
c6	28	3500	Ruim	1800	Não
c7	40	6500	Bom	900	Sim

Novo cliente!

c8	30	4000	Bom	800	???
----	----	------	-----	-----	-----

Na prática

Aprovação de empréstimo – Teste

Cliente	Idade	Renda	Histórico de Crédito	Valor da Dívida	Aprovado?
c1	25	3000	Bom	500	Sim
c2	45	8000	Bom	1000	Sim
c3	23	2000	Ruim	2500	Não
c4	35	5000	Bom	700	Sim
c5	52	9000	Ruim	4000	Não
c6	28	3500	Ruim	1800	Não
c7	40	6500	Bom	900	Sim

Novo cliente!

c8	30	4000	Bom	800	Sim
----	----	------	-----	-----	------------

Com base na hipótese aprendida

Indução de Hipóteses

Generalização → A hipótese deve ser válida para outras instâncias do problema que não foram vistas ainda

Indução de Hipóteses

Generalização → A hipótese deve ser válida para outras instâncias do problema que não foram vistas ainda

- ***Overfitting*** → hipótese memorizou/especializou nos dados de treinamento
- ***Underfitting*** → hipótese apresenta baixa taxa de acerto até mesmo no treinamento

Tarefas de Aprendizado

Preditivas → Encontrar uma função (hipótese) a partir dos dados de treinamento para prever um rótulo ou valor com base nos atributos de entrada

Algoritmos seguem o paradigma do **aprendizado supervisionado**

- *Existe um “supervisor externo” que conhece a saída desejada para cada exemplo – o atributo de saída*

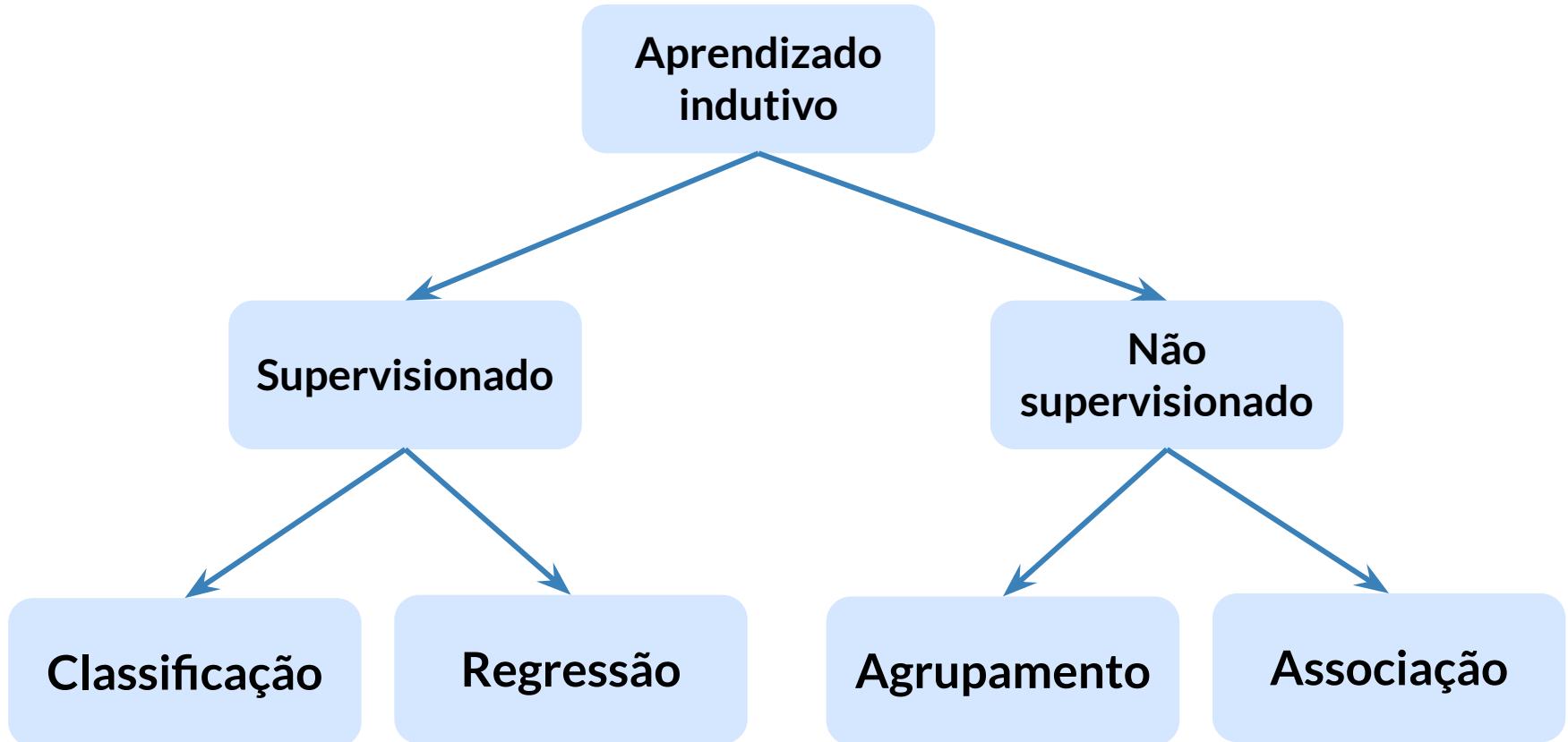
Tarefas de Aprendizado

Descritivas → Explorar ou descrever um conjunto de dados

Algoritmos seguem o paradigma do **aprendizado não supervisionado**

- *Não fazem uso o atributo de saída*

Tarefas de Aprendizado



Existem outras tarefas?

Aprendizado semi-supervisionado → Envolve conjuntos de dados de treinamento rotulados e não-rotulados

Ideal quando o projeto possui muitos dados de treinamento, mas a maioria não estão rotulados

Existem outras tarefas?

Aprendizado por reforço → Reforçar ou recompensar uma ação considerada positiva e punir uma ação considerada negativa

Exemplo: ensinar um robô a encontrar a melhor trajetória entre dois pontos

2. Aprendizado Supervisionado

Classificação e regressão

Classificação

Objetivo → dado um exemplo não rotulado, atribui uma das possíveis **classes** (rótulos)

Classes são sempre atributos **categóricos** (nominais)

É uma das tarefas mais comuns em AM

Classificação

O cliente vai ter valor aprovado?

Cliente	Idade	Renda	Histórico de Crédito	Valor da Dívida	Aprovado?
c1	25	3000	Bom	500	Sim
c2	45	8000	Bom	1000	Sim
c3	23	2000	Ruim	2500	Não
c4	35	5000	Bom	700	Sim
c5	52	9000	Ruim	4000	Não
c6	28	3500	Ruim	1800	Não
c7	40	6500	Bom	900	Sim

Novo cliente!

c8	30	4000	Bom	800	???
----	----	------	-----	-----	-----

Qual será a classe?

Regressão

Semelhante à classificação

Objetivo → dado um exemplo não rotulado, atribui uma dos possíveis **valores** (rótulos)

Classes são sempre atributos **numéricos**

Regressão

Qual o valor máximo que o cliente pode receber?

Cliente	Idade	Renda	Histórico de Crédito	Valor da Dívida	Valor Concedido
c1	25	3000	Bom	500	8000
c2	45	8000	Bom	1000	25000
c3	23	2000	Ruim	2500	2000
c4	35	5000	Bom	700	15000
c5	52	9000	Ruim	4000	6000
c6	28	3500	Ruim	1800	4000
c7	40	6500	Bom	900	20000

Novo cliente!

c8	30	4000	Bom	800	???
----	----	------	-----	-----	-----

Qual será o valor liberado para o novo cliente?

2.

Aprendizado Não Supervisionado

Agrupamento e associação

Agrupamento (*clustering*)

Objetivo → agrupar as instâncias de acordo com a similaridade entre as instâncias (a partir das *features*)

Não existe rótulo/atributo de saída

É muito interessante para descobrir perfis dentro do conjunto de dados

Agrupamento (*clustering*)

Que tipos de clientes existem na base?

Cliente	Idade	Renda	Histórico de Crédito	Valor da Dívida
c1	25	3000	Bom	500
c2	45	8000	Bom	1000
c3	23	2000	Ruim	2500
c4	35	5000	Bom	700
c5	52	9000	Ruim	4000
c6	28	3500	Ruim	1800
c7	40	6500	Bom	900

Agrupamento (*clustering*)

Que tipos de clientes existem na base?

Cliente	Idade	Renda	Histórico de Crédito	Valor da Dívida
c1	25	3000	Bom	500
c2	45	8000	Bom	1000
c3	23	2000	Ruim	2500
c4	35	5000	Bom	700
c5	52	9000	Ruim	4000
c6	28	3500	Ruim	1800
c7	40	6500	Bom	900

Grupo A: baixa renda, maior risco

Agrupamento (*clustering*)

Que tipos de clientes existem na base?

Cliente	Idade	Renda	Histórico de Crédito	Valor da Dívida
c1	25	3000	Bom	500
c2	45	8000	Bom	1000
c3	23	2000	Ruim	2500
c4	35	5000	Bom	700
c5	52	9000	Ruim	4000
c6	28	3500	Ruim	1800
c7	40	6500	Bom	900

Grupo B:
média/alta renda,
bom histórico

Agrupamento (*clustering*)

Que tipos de clientes existem na base?

Cliente	Idade	Renda	Histórico de Crédito	Valor da Dívida
c1	25	3000	Bom	500
c2	45	8000	Bom	1000
c3	23	2000	Ruim	2500
c4	35	5000	Bom	700
c5	52	9000	Ruim	4000
c6	28	3500	Ruim	1800
c7	40	6500	Bom	900

Grupo C: alta renda, mas alto endividamento

Associação

Objetivo → busca de **padrões frequentes** de associações entre atributos de um conjunto de dados

Features são sempre **categóricas**

Exemplo: em sistemas de e-commerce, descobrir regras do tipo “quem compra produto A, também compra produto B”

Associação

Quais características costumam aparecer juntas?

Cliente	Idade	Renda	Histórico de Crédito	Valor da Dívida
c1	25	3000	Bom	500
c2	45	8000	Bom	1000
c3	23	2000	Ruim	2500
c4	35	5000	Bom	700
c5	52	9000	Ruim	4000
c6	28	3500	Ruim	1800
c7	40	6500	Bom	900

Associação

Quais características costumam aparecer juntas?

Cliente	Idade	Renda	Histórico de Crédito	Valor da Dívida
c1	Jovem	Baixa	Bom	Baixa
c2	Adulto	Alta	Bom	Baixa
c3	Jovem	Baixa	Ruim	Média
c4	Adulto	Média	Bom	Baixa
c5	Adulto	Alta	Ruim	Alta
c6	Jovem	Baixa	Ruim	Média
c7	Adulto	Alta	Bom	Baixa

Associação

Quais características costumam aparecer juntas?

Cliente	Idade	Renda	Histórico de Crédito	Valor da Dívida
c1	Jovem	Baixa	Bom	Baixa
c2	Adulto	Alta	Bom	Baixa
c3	Jovem	Baixa	Ruim	Média
c4	Adulto	Média	Bom	Baixa
c5	Adulto	Alta	Ruim	Alta
c6	Jovem	Baixa	Ruim	Média
c7	Adulto	Alta	Bom	Baixa

(Alta renda, bom histórico) → Baixa dívida

Algoritmo descobre
padrões e regras

3. Técnicas e Algoritmos

Para cada tarefa

Exemplos de algoritmos clássicos

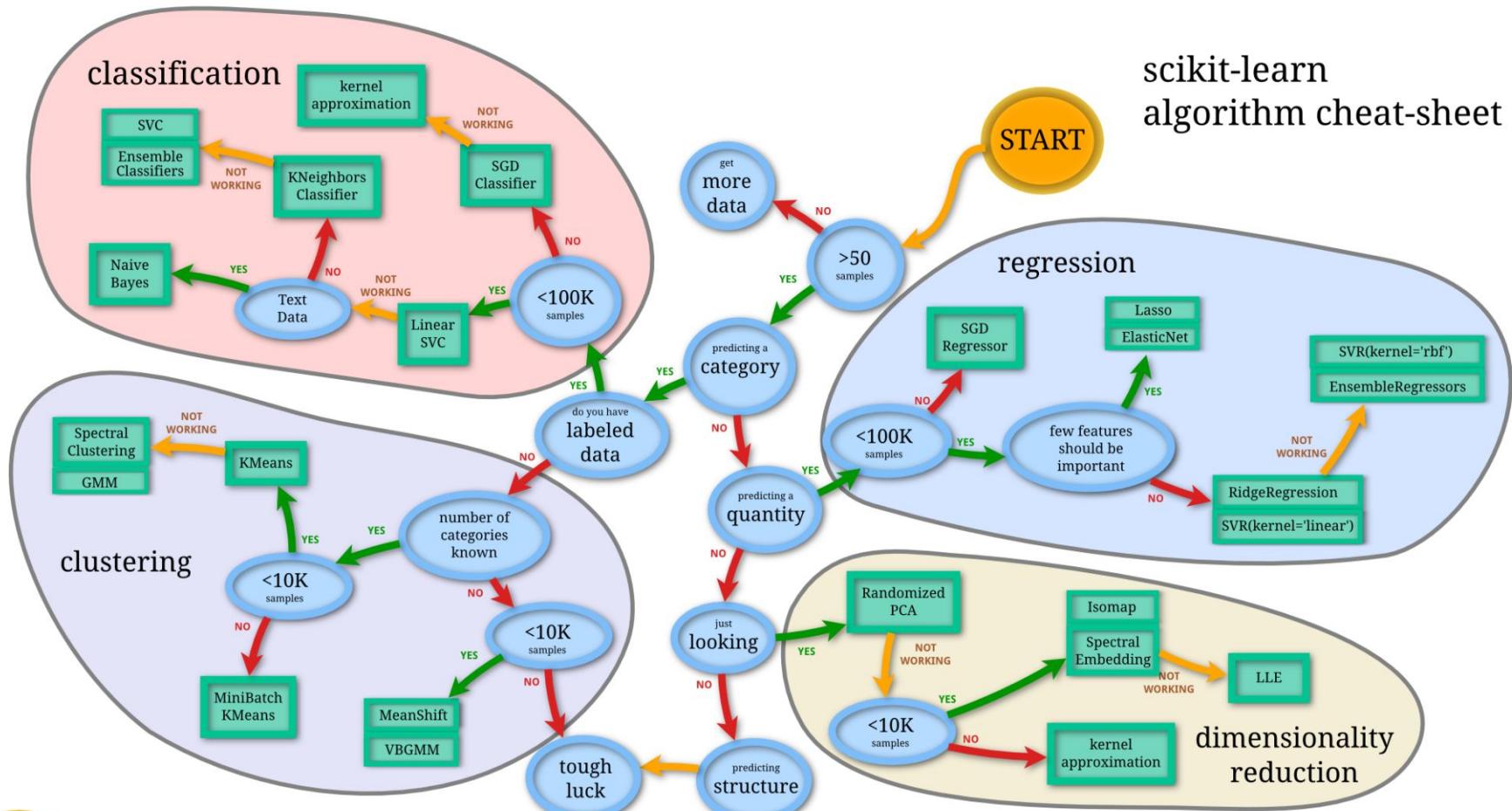
Supervisionado		Não supervisionado	
Classificação	Regressão	Agrupamento	Associação
Regressão Logística	Regressão Linear	K-Means	Apriori
Árvore de Decisão	Árvore de Regressão	DBSCAN	FP-Growth
Random Forest	Random Forest	Hierárquico	Eclat
k-NN	k-NN		
SVM	SVR		
Naive Bayes			

Exemplos de algoritmos clássicos

Supervisionado		Não supervisionado	
Classificação	Regressão	Agrupamento	Associação
Regressão Logística	Regressão Linear	K-Means	Apriori
Árvore de Decisão	Árvore de Regressão	DBSCAN	FP-Growth
Random Forest	Random Forest	Hierárquico	Eclat
k-NN	k-NN		
SVM	SVR		
Naive Bayes			

Mesmas técnicas com
pequenas adaptações de
acordo com a tarefa

scikit-learn algorithm cheat-sheet



Back

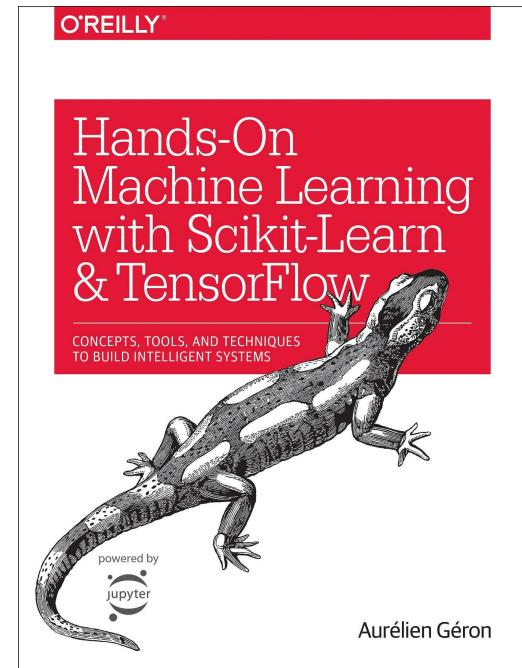


4. Ferramentas

Para a parte prática

Principais ferramentas

Linguagens de programação:



Bibliotecas/frameworks (Python):

Manipulação de dados



Algoritmos



Redes Neurais



TensorFlow

5. Exercícios

Vamos testar os conhecimentos?

Exercício #1

Uma universidade quer prever se um aluno tem alta chance de abandonar o curso no próximo semestre para que possa intervir antecipadamente.

- Qual a tarefa?
- Possíveis atributos de entrada (*features*)?
- Existe atributo de saída (alvo/target)? Qual?

Exercício #1

Uma universidade quer prever se um aluno tem alta chance de abandonar o curso no próximo semestre para que possa intervir antecipadamente.

- Qual a tarefa? *Classificação*
- Possíveis atributos de entrada (*features*)?

Curso, frequência, nota, tempo de deslocamento, etc.

- Existe atributo de saída (*alvo/target*)? Qual?

Sim. Evasão (sim/não)

Exercício #2

Uma empresa de bicicletas compartilhadas quer prever quantas bicicletas serão alugadas na Lagoa da Pampulha ao longo do dia.

- Qual a tarefa?
- Possíveis atributos de entrada (*features*)?
- Existe atributo de saída (alvo/target)? Qual?

Exercício #2

Uma empresa de bicicletas compartilhadas quer prever quantas bicicletas serão alugadas na Lagoa da Pampulha ao longo do dia.

- Qual a tarefa? Regressão
- Possíveis atributos de entrada (*features*)?

Dia da semana, temperatura, chuva, feriado, evento, etc.

- Existe atributo de saída (*alvo/target*)? Qual?

Sim. Número de bicicletas alugadas.

Exercício #2

Uma empresa de bicicletas compartilhadas quer prever quantas bicicletas serão alugadas na Lagoa da Pampulha ao longo do dia.

Podemos transformar em um problema de classificação?

- Qual a tarefa? Regressão
- Possíveis atributos de entrada (*features*)?

Dia da semana, temperatura, chuva, feriado, evento, etc.

- Existe atributo de saída (*alvo/target*)? Qual?

Sim. Número de bicicletas alugadas.

Exercício #2

Uma empresa de bicicletas compartilhadas quer prever quantas bicicletas serão alugadas na Lagoa da Pampulha ao longo do dia.

- Qual a tarefa? **Classificação**
- Possíveis atributos de entrada (**features**)?

Podemos transformar em um problema de classificação?

Dia da semana, temperatura, chuva, feriado, evento, etc.

- Existe atributo de saída (**alvo/target**)? Qual?

Sim. Quantidade de bicicletas alugadas (alta/baixa).

Exercício #2

Uma empresa de bicicletas compartilhadas quer prever quantas bicicletas serão alugadas na Lagoa da Pampulha ao longo do dia.

- Qual a tarefa?
- Possíveis atributos de entrada (*features*)?
- Existe atributo de saída (alvo/target)? Qual?

Exercício #3

Um hospital quer descobrir quais medicamentos costumam ser prescritos juntos para determinados perfis de pacientes.

- Qual a tarefa?
- Possíveis atributos de entrada (*features*)?
- Existe atributo de saída (alvo/target)? Qual?

Exercício #3

Um hospital quer descobrir quais medicamentos costumam ser prescritos juntos para determinados perfis de pacientes.

- Qual a tarefa? Associação
- Possíveis atributos de entrada (*features*)?

Medicamento A, medicamento B, Idade > 60, diabetes, etc.

- Existe atributo de saída (alvo/target)? Qual?

Não (descoberta de regras).

Referências

FACELI, Katti et al. **Inteligência artificial: uma abordagem de aprendizado de máquina.** Rio de Janeiro, RJ: LTC, 2011.
(Capítulo 1 e Introdução das Partes II e III)