

Aprendizado de Máquina

Aula 2: Aprendizado

André C. P. L. F de Carvalho
ICMC/USP

andre@icmc.usp.br



Tópicos

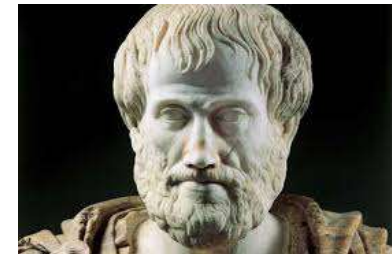
- Aprendizado de Máquina
- O que faz?
- Aplicações
- Tarefas
- Abordagens

Aprendizado de Máquina (AM)

- Revolução industrial automatizou trabalho manual
- Revolução da informação automatizou trabalho mental
- Revolução de aprendizado de máquina automatiza a própria automação

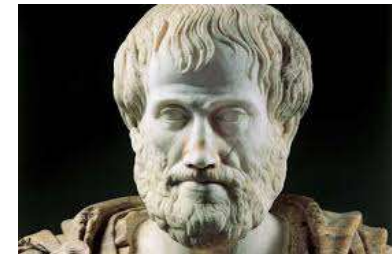
Quando começou?

- Filósofo grego Aristóteles criou a lógica moderna (384-322 AC)
 - Compôs Órganon (instrumento): coleção de seis textos em que reconhecia duas formas de argumentação
 - Dedução (raciocínio dedutivo)
 - Começa com uma teoria geral e usa observações para torná-la mais específica
 - *Top-down* (geral para específico)
 - Indução (raciocínio indutivo)
 - Começa com observações e busca uma teoria para explicá-las
 - Bottom-up (específico para geral)
 - **Aprendizado**



Quando começou?

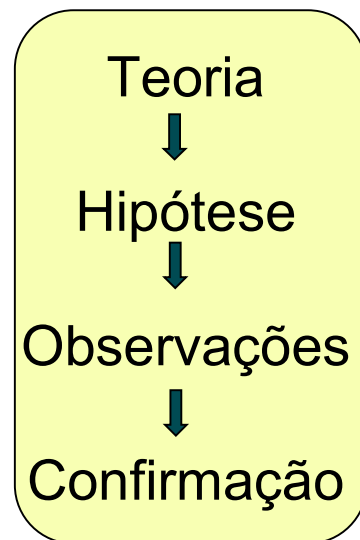
- Filósofo grego Aristóteles criou a lógica moderna (384-322 AC)
 - Compôs Órganon (instrumento): coleção de seis textos Reconhecia duas formas de argumentação
 - Dedução (raciocínio dedutivo)
 - Começa com uma teoria geral e usa observações para torná-la mais específica
 - *Top-down* (geral para específico)
 - Indução (raciocínio indutivo)
 - Começa com observações e busca uma teoria para explicá-las
 - Bottom-up (específico para geral)
 - **Aprendizado**



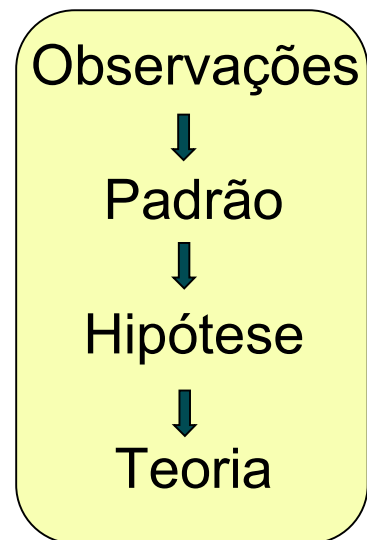
Sr. X foi assassinado. Polícia suspeita de Sr. Y. Polícia busca evidências para deduzir que foi realmente Sr. Y.

Sr. X foi assassinado. Polícia olha trajetória da bala, cinzas de cigarro, fios curtos de cabelo, distância entre pegadas. Das observações, a polícia prediz que o assassino é Sr. Y.

Dedução x Indução



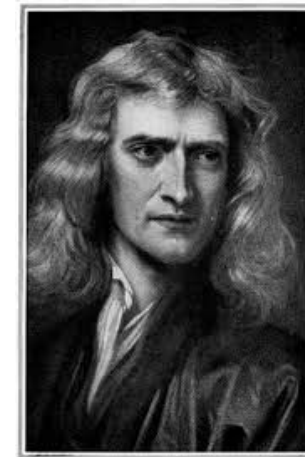
**Raciocínio
Dedutivo**



**Raciocínio
Indutivo**

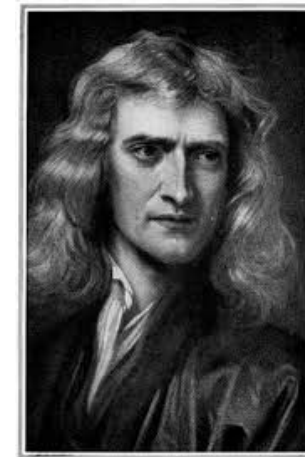
Um dos marcos de AM

- Isaac Newton publica *Principia* em 1687
 - Princípios Matemáticos da Filosofia Natural
 - Conjunto de 3 livros
 - Livro 1: 3 leis de movimento
 - Lei da inércia
 - Princípio fundamental da dinâmica
 - Lei da ação e reação



Um dos marcos de AM

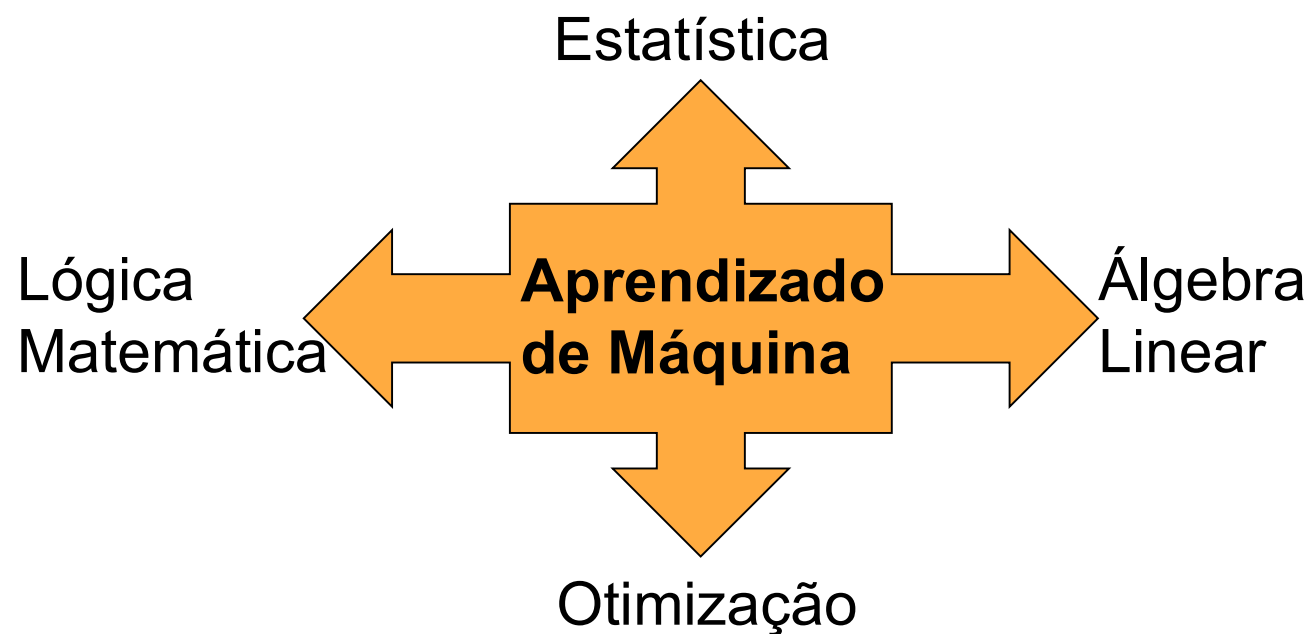
- Isaac Newton publica *Principia* em 1687
 - Princípios Matemáticos da Filosofia Natural
 - Conjunto de 3 livros
 - Livros 1 : propõe 3 leis de movimento
 - Lei da inércia
 - Princípio fundamental da dinâmica
 - Lei da ação e reação
 - Livro 3: propõe 4 regras de indução
 - 3ª regra: O que é verdade para tudo que nós vimos é verdade para tudo no Universo



Quando o termo foi criado?

- Termo Machine Learning (Aprendizado de Máquina) foi usado pela primeira vez em 1952, por Artur Samuel, programador da IBM
- Desenvolveu um programa para jogar Xadrez
 - Para lidar com a pouca memória, escreveu o programa de poda alpha-beta
 - Deu origem ao algoritmo MinMax
 - Construiu vários mecanismos para o programa lembrar das posições já percorridas, que era combinado com valores de uma função de recompensa
 - Quanto mais o programa jogava, mais melhorava
 - Aprendizado de Máquina

É só computação?



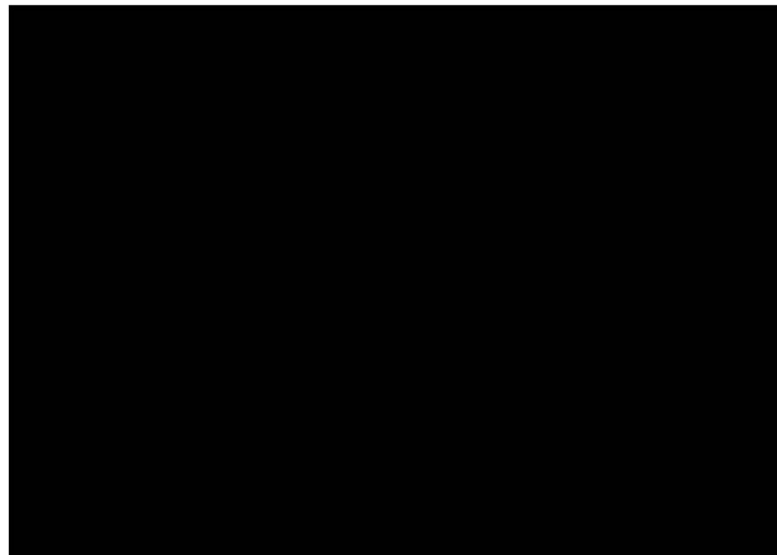
Aplicações de AM

- AM esta presente em várias atividades do nosso dia-a-dia, sendo utilizado para:
 - Recomendar que mensagens mostrar em aplicativos de redes sociais
 - Filtrar spams de seus *emails*
 - Decidir que resultados (e anúncios), e em que ordem, mostrar após uma busca na internet
 - Sugerir filmes ou livros que alguém vai gostar de assistir
 - Diagnosticar se você pode ter uma determinada doença
 - Ex.: Covid 19

Aplicações clássicas de AM

- Aprender a ler em voz alta
 - NETtalk (Terrence Sejnowski e Rosenberg, 1986)
- Aprender a reconhecer palavras faladas
 - SPHINX (Lee, 1989)
- Aprender a conduzir um automóvel
 - ALVINN (Pomerleau, 1989)
- Aprender a jogar gamão
 - TD-GAMMON (Tesauro 1992)

NETtalk



<https://www.youtube.com/watch?v=gakJlr3GecE>

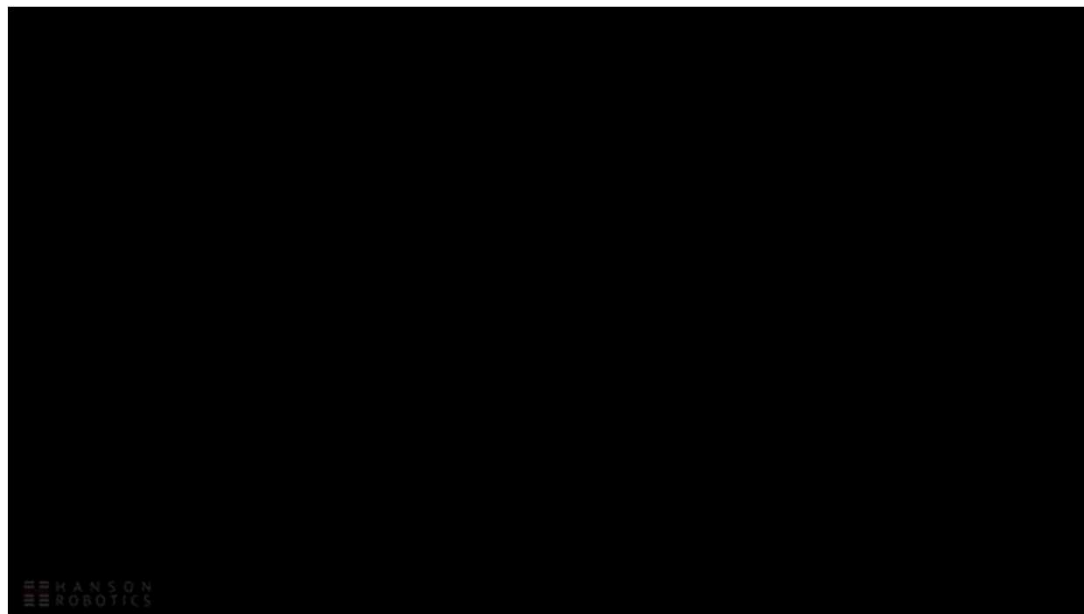
Copyright © 2020. Todos os direitos reservados ao CeMEAI-USP. Proibida a cópia e reprodução sem autorização



NETtalk

- Desempenho
 - 95% para os dados de treinamento e 78% para os dados de teste
 - Comparado com Dectalk
 - Sistema especialista desenvolvido por linguistas e baseado em regras
 - Dectalk apresentou um desempenho melhor
 - Foi desenvolvido em cerca de dez anos utilizando análises feitas por linguistas

Sophia



<https://www.youtube.com/watch?v=XrSAQoetF0A>

Copyright © 2020. Todos os direitos reservados ao CeMEAI-USP. Proibida a cópia e reprodução sem autorização



ALVINN



Dean Pomerleau
CMU



Copyright © 2020. Todos os direitos reservados ao CeMEAI-USP. Proibida a cópia e reprodução sem autorização



ALVINN

- *Autonomous Land Vehicle In a Neural Network*
 - Sistema automático de navegação para automóveis baseado em Redes Neurais
 - Tese de doutorado da CMU
 - Coletava imagens por meio de uma câmera montada no topo do veículo
 - Dirigiui em 1989 a 110 Km/h em uma rodovia pública americana
 - De costa a costa por 4500 Km (com exceção de 80 Km)

ALVINN



<https://www.youtube.com/watch?v=WPexu1mUH5s>

Copyright © 2020. Todos os direitos reservados ao CeMEAI-USP. Proibida a cópia e reprodução sem autorização



Carros Tesla



<https://www.youtube.com/watch?v=tIThdr3O5Qo>

Copyright © 2020. Todos os direitos reservados ao CeMEAI-USP. Proibida a cópia e reprodução sem autorização



Principais preocupações



Cyber-Attacks

One cyber-attack affected

143 million

Americans in 2017. Will driverless cars be secure?



Untested Technology

Autonomous driving involves too many variables that have yet to be fully tested.



Computer Error

Software malfunctions in a driverless car endangers lives.



Liability Issues

Who is responsible when a **driverless car fails?**

Conjuntos de dados

- Rotulados
 - Cada objeto recebe um rótulo
 - Classe
 - Valor real
- Não rotulados
 - Objetos não recebem rótulos
- Parcialmente rotulados
 - Alguns objetos recebem rótulos

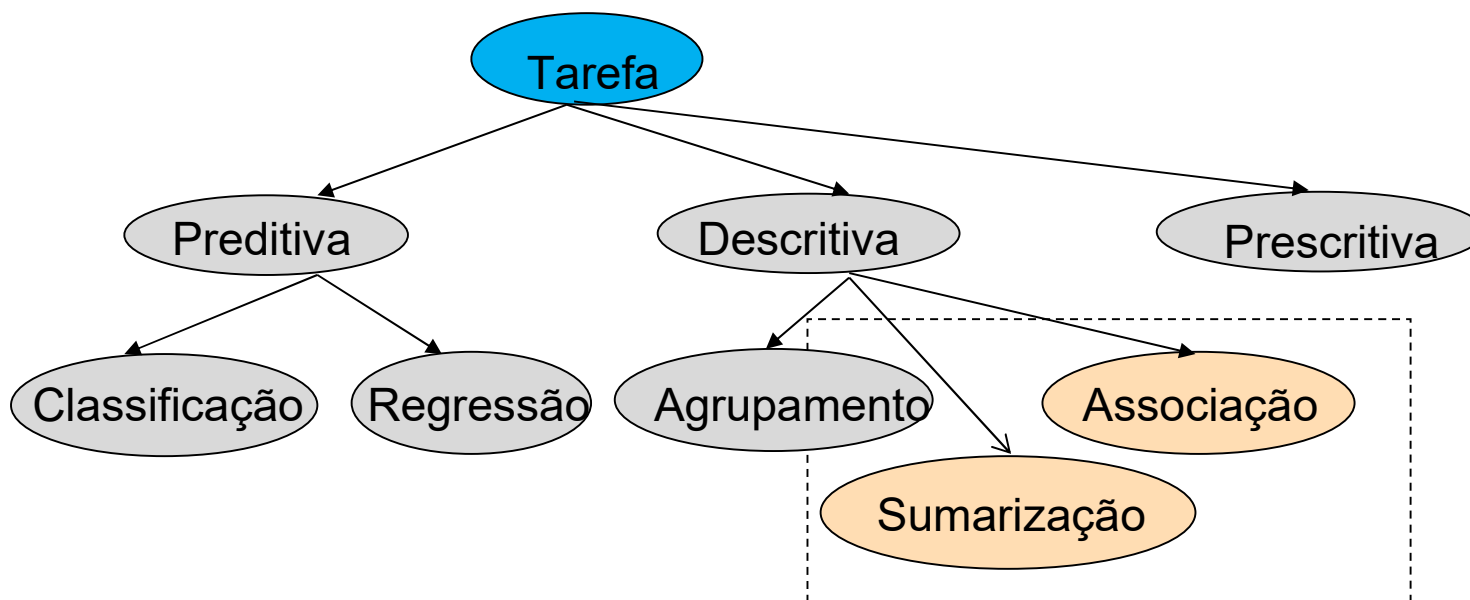
Dados estruturados rotulados

Atributos de entrada (preditivos)				
Altura	TamanhoRabo	Textura	Classe	
500	110	Manchas	Girafa	Exemplos (objetos, instâncias)
440	90	Manchas	Girafa	
240	45	Listas	Zebra	
520	115	Manchas	Girafa	
260	50	Listas	Zebra	
230	50	Listas	Zebra	
				Atributo alvo

Dados estruturados não rotulados

Atributos de entrada (preditivos)			
	Altura	TamanhoRabo	Textura
Exemplos (objetos, instâncias)	500	110	Manchas
	440	90	Manchas
	240	45	Listas
	520	115	Manchas
	260	50	Listas
	230	50	Listas

Tarefas de aprendizado



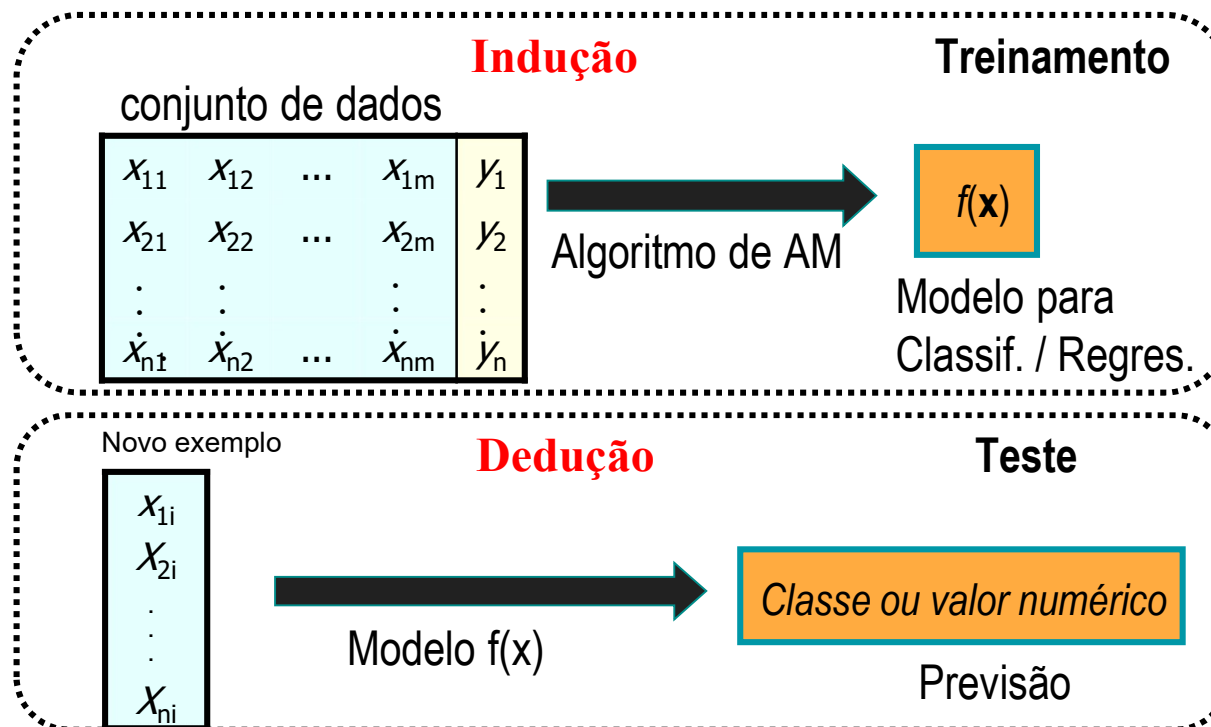
Algoritmo de aprendizado

- Supervisionado
 - Tarefa preditiva (mais comum) ou descritiva
 - Aprendizado usa atributos preditivos e alvo
- Não supervisionado
 - Tarefa descritiva (mais comum) ou preditiva
 - Aprendizado usa apenas atributos preditivos
- Semi-supervisionado
- Aprendizado ativo
- Por reforço

Algoritmos para tarefas preditivas

- Induzem modelos (funções) preditivas
 - Dados de treinamento rotulados
- Modelo pode ser aplicado a novos dados (predição)
 - Rotula dados de teste
- Principais tarefas:
 - Classificação
 - Regressão

Algoritmos para tarefas preditivas



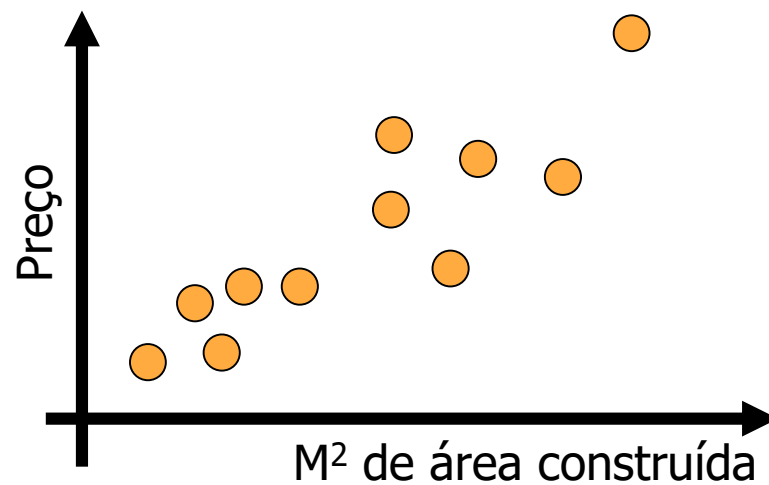
Regressão

- Objetivo: aprender uma função que mapeia descrição de um exemplo em um valor real
- Exemplos:
 - Prever valor de mercado de um imóvel
 - Prever o lucro de um empréstimo bancário
 - Prever tempo de internação de paciente

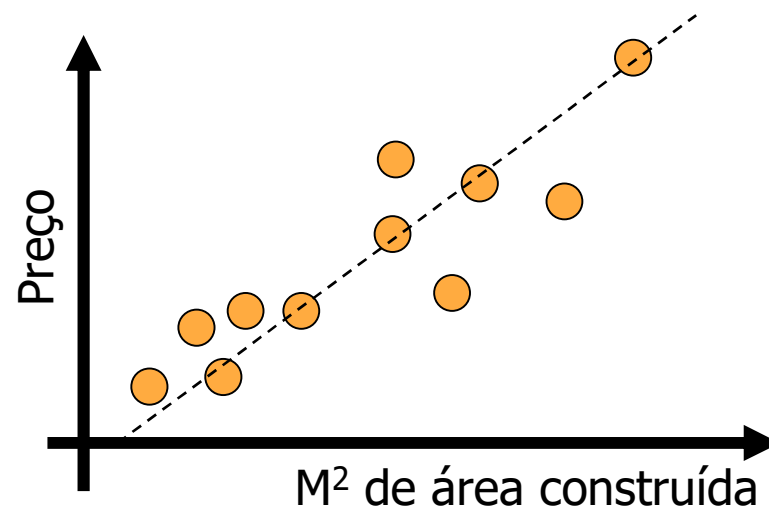
Regressão

- Imobiliária vendenada
 - Um corretor é o especialista em dar preços
 - Já vendeu várias casas
 - Usa uma idéia simples: preço é igual a 10.000 vezes o número de minutos que demora para ver a casa
 - Filho do dono acha que pode achar valores parecidos usando área construída

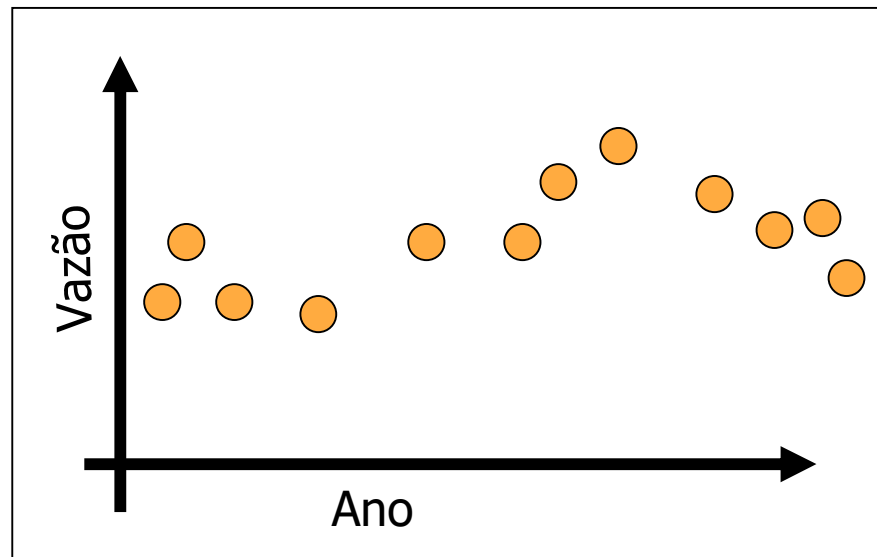
Regressão



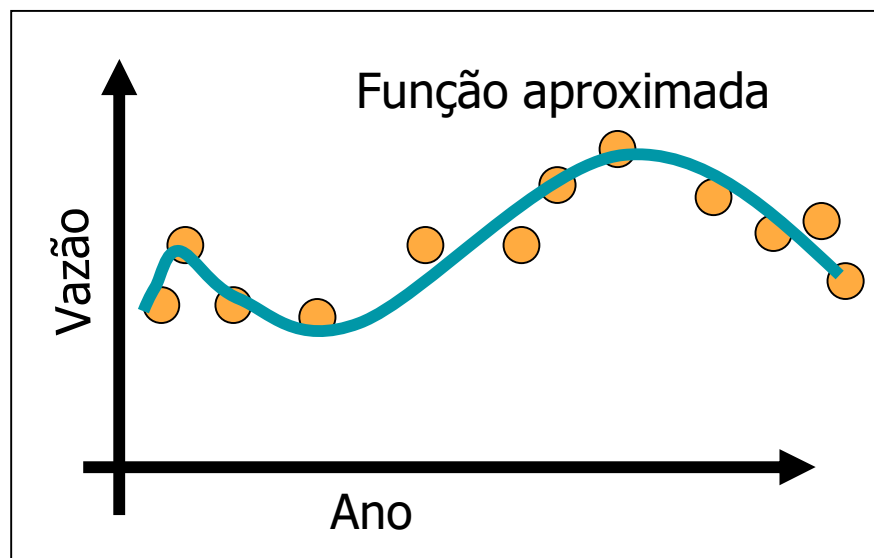
Regressão



Nem tudo é tão simples

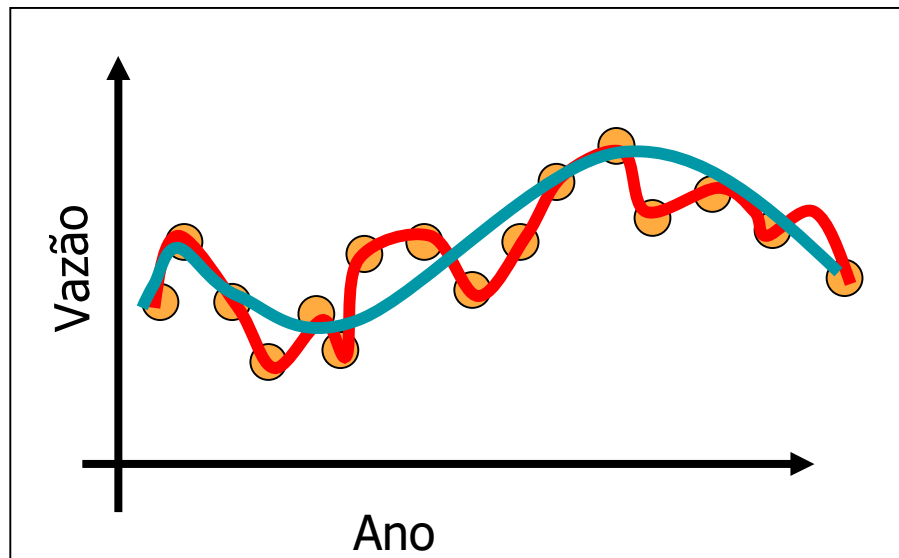


Nem tudo é tão simples



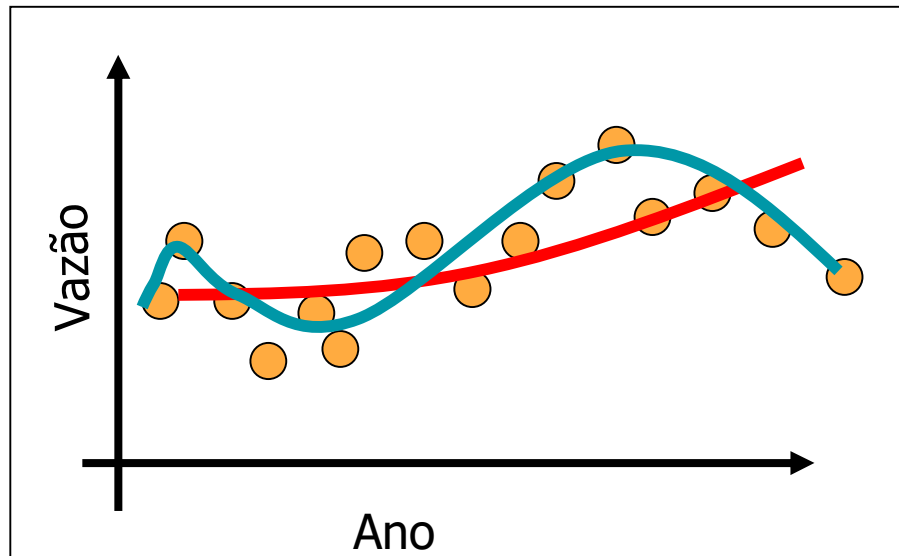
Nem tudo é tão simples

- Overfitting



Nem tudo é tão simples

- Underfitting



Algoritmos de regressão

- Árvores de Regressão
- *Backpropagation* (redes neurais)
 - *Deep learning*
- Máquinas de Vetores de Suporte
- Regressão Linear
- Lasso
- ...

Classificação

- Objetivo: aprender função que associa descrição de um objeto a sua classe
- Exemplos:
 - Definir a função de uma proteína
 - Distinguir *emails* entre *spam* e *ham*
 - Definir se um paciente tem ou não uma determinada doença

Classificação

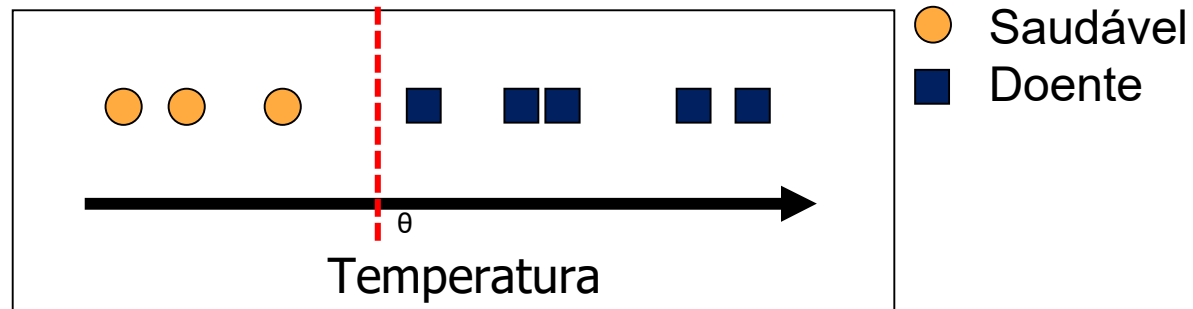
- Posto médico Daquiproceu
 - Tem um histórico de vários atendimentos e diagnósticos
 - João, ao sentir alguns sintomas, vai ao posto para uma consulta médica
 - O único médico, faltou
 - Mas o estagiário pode anotar os sintomas
 - É possível fazer um pré-diagnóstico?

Classificação

- Sintomas coletados pelo estagiário:
 - Temperatura

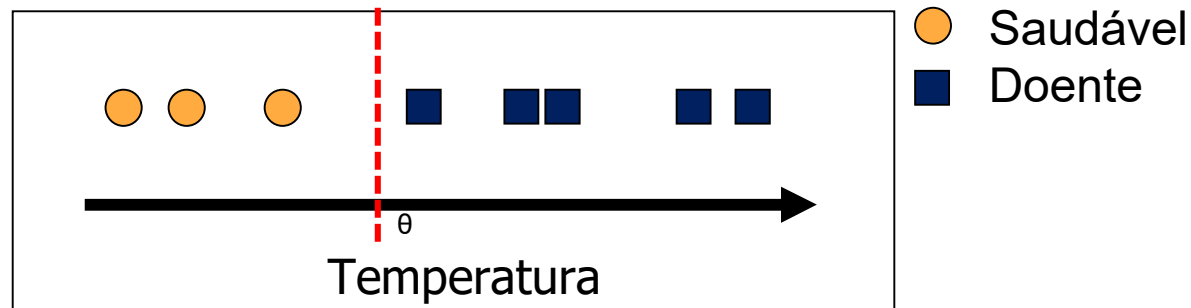
Classificação

- Forma mais simples



Classificação

- Forma mais simples



Função estimada: diagnóstico = $f(\text{temperatura})$

Se temperatura $> \theta$

Então doente

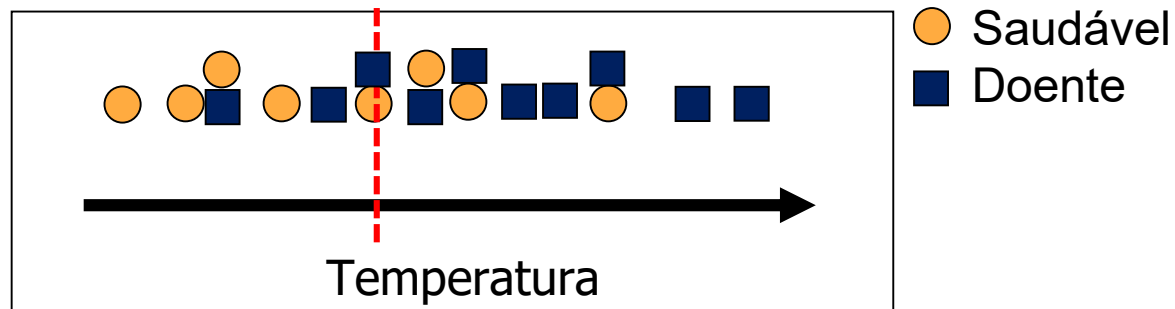
Senão saudável

Classificação

- Basta encontrar um valor (limiar) de temperatura que separa
 - Doentes
 - Saudáveis
- Mas todo problema de classificação é simples assim?
 - Uso apenas da temperatura gera um bom modelo preditivo?

Classificação

- Problema pode não ser tão simples



- Alternativa: considerar outros sintomas para o diagnóstico

Classificação

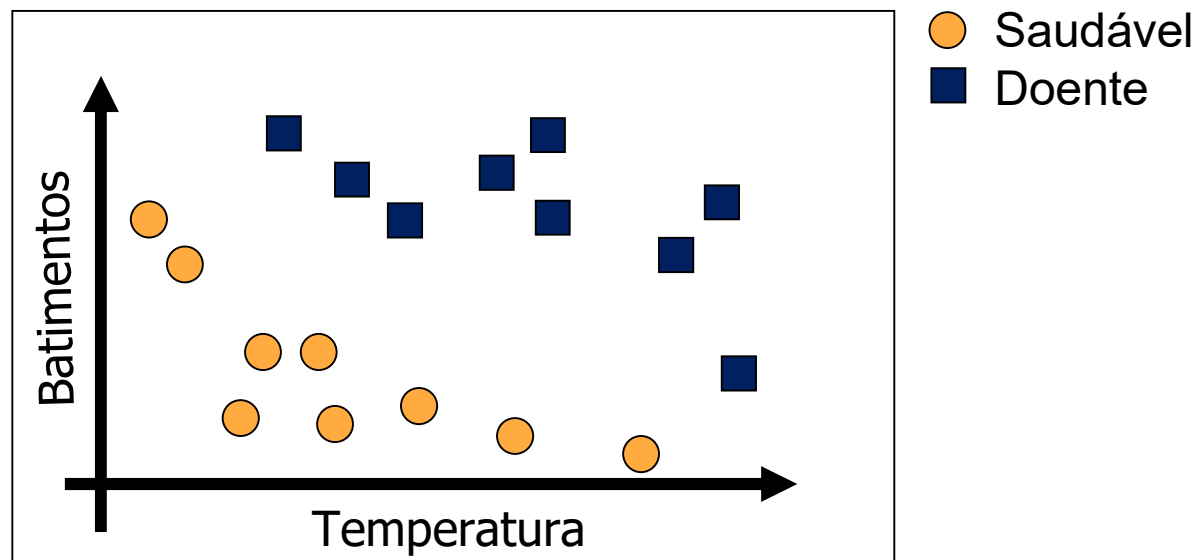
- Sintomas coletados pela enfermeira:
 - Batimentos cardíacos
 - Temperatura

Classificação

- Incluir número de batimentos

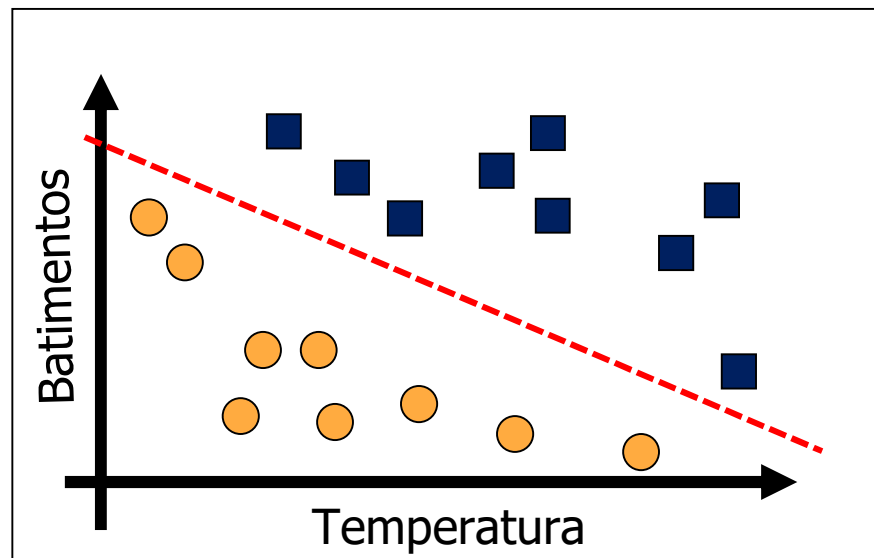
Classificação

- Incluir número de batimentos



Classificação

- Função linear permite diagnóstico



● Saudável
■ Doente

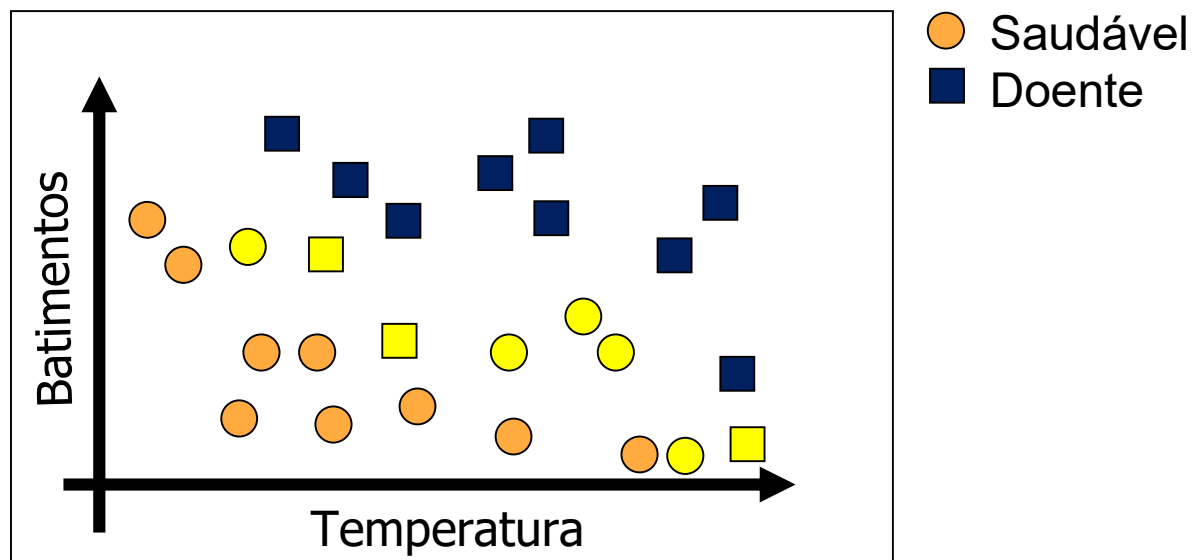
Nova função:
Se $a \cdot t + b > 0$
Então doente
Senão saudável

Classificação

- Basta encontrar uma função linear que separa pacientes doentes de saudáveis
 - Inclinação da reta e ponto onde cruza o eixo da ordenada
- Espaço de pacientes
 - Ordenada: número de batimentos cardíacos
 - Abscissa: temperatura
- Mas toda tarefa de classificação é simples assim?

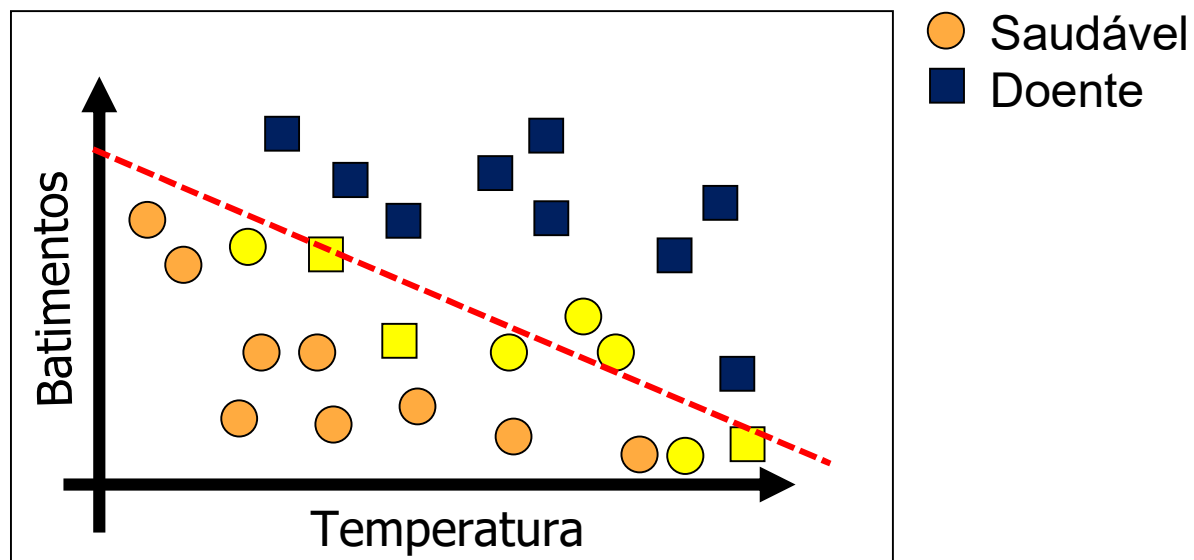
Classificação

- Supor inclusão de outros pacientes



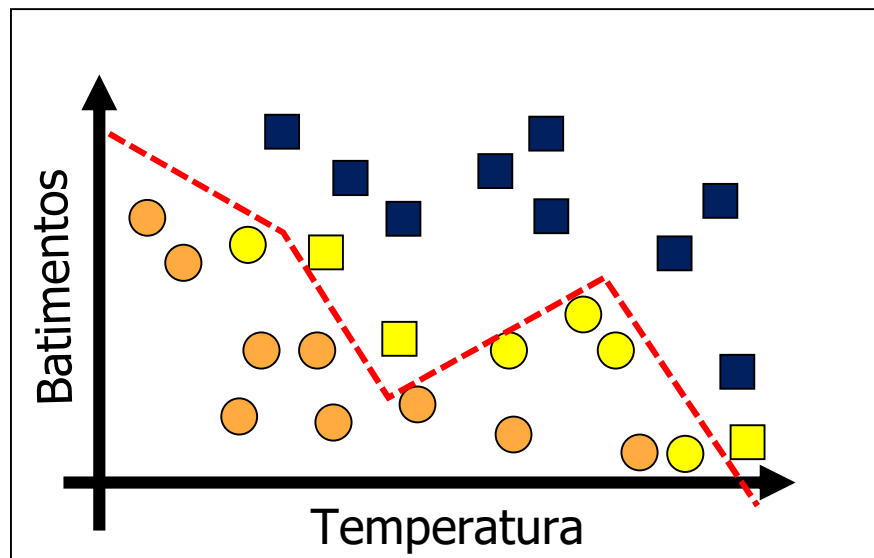
Classificação

- Função linear agora não permite diagnóstico



Classificação

- Função não linear permite diagnóstico

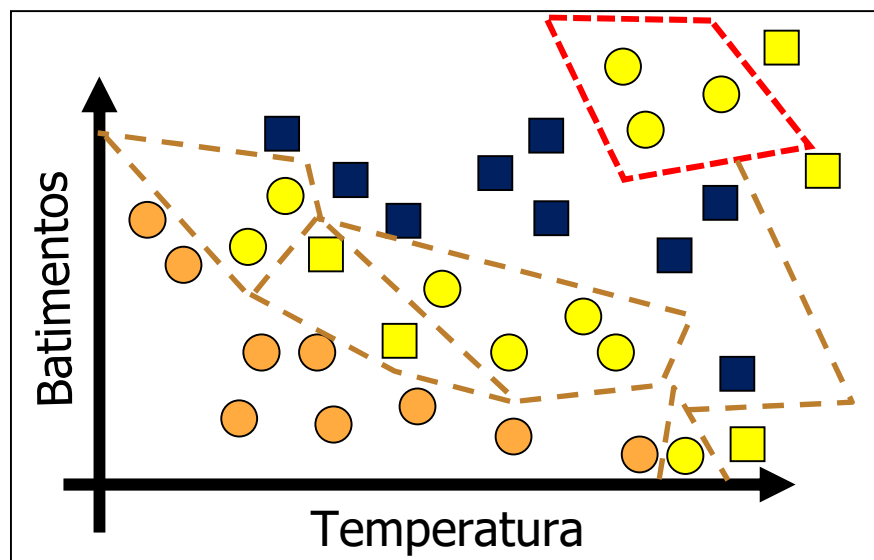


● Saudável
■ Doente

Nova função:
Mais complexa

Classificação

- Função não linear permite diagnóstico



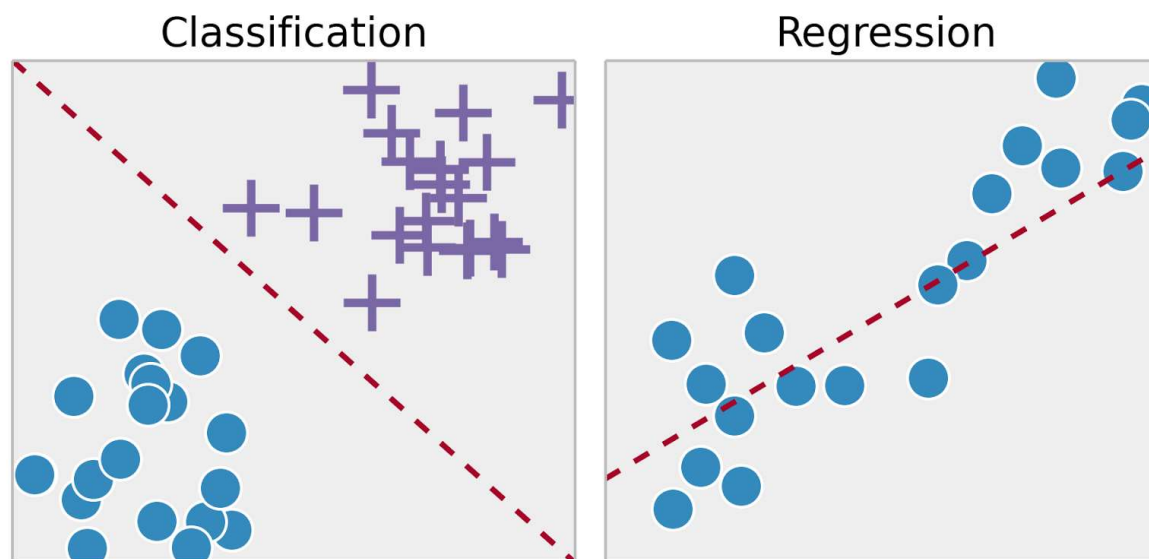
● Saudável
■ Doente

Nova função:
Muito mais complexa

Algoritmos de classificação

- Árvores de Classificação
- *Backpropagation* (redes neurais)
- Máquinas de Vetores de Suporte
- Regressão Logística
- K-NN
- Redes Bayesianas

Classificação vs regressão



Algoritmos para tarefas descritivas

- Também geram modelos em um processo de treinamento
 - Descrevem ou sumarizam dados
 - Treinamento utiliza todos o conjunto de dados
 - Ex.: Agrupamento de dados
- Alguns algoritmos não utilizam treinamento
 - Ex.: Regras de associação e sumarização

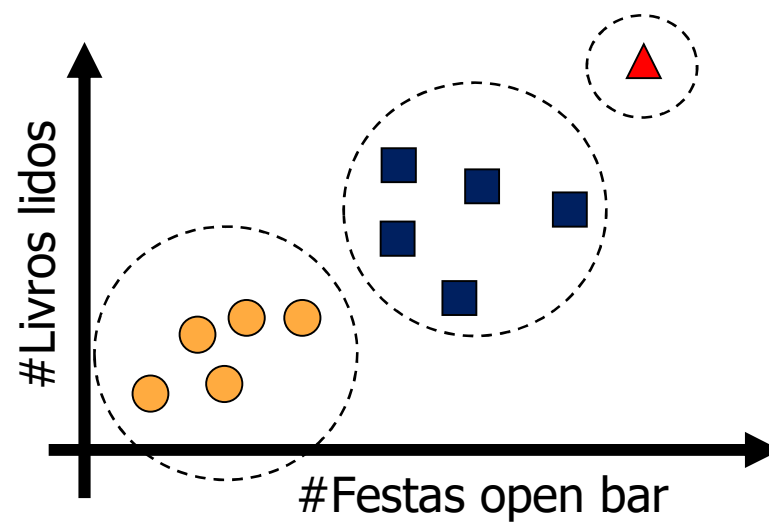
Agrupamento (Clustering)

- Objetivo: organizar exemplos não rotulados em grupos (clusters)
 - De acordo com uma medida de similaridade ou correlação entre eles
 - Aprendizado não supervisionado
- Não existe conhecimento anterior sobre:
 - Número de grupos (várias vezes)
 - Significado dos grupos

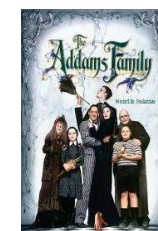
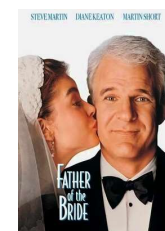
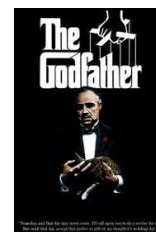
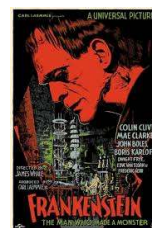
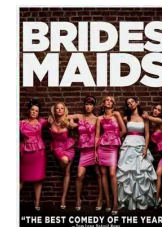
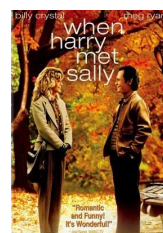
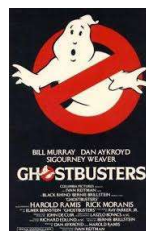
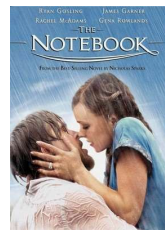
Agrupamento

- Colégio Nãovaidar
 - Tem um grande número de alunos
 - Gostariam que eles formassem grupos que compartilhassem interesses
 - Só sabe duas coisas de cada aluno
 - Quantos livros leu no ano passado
 - Quantas vezes foi para uma festa open bar no ano passado
 - Dá para sugerir grupos?

Agrupamento



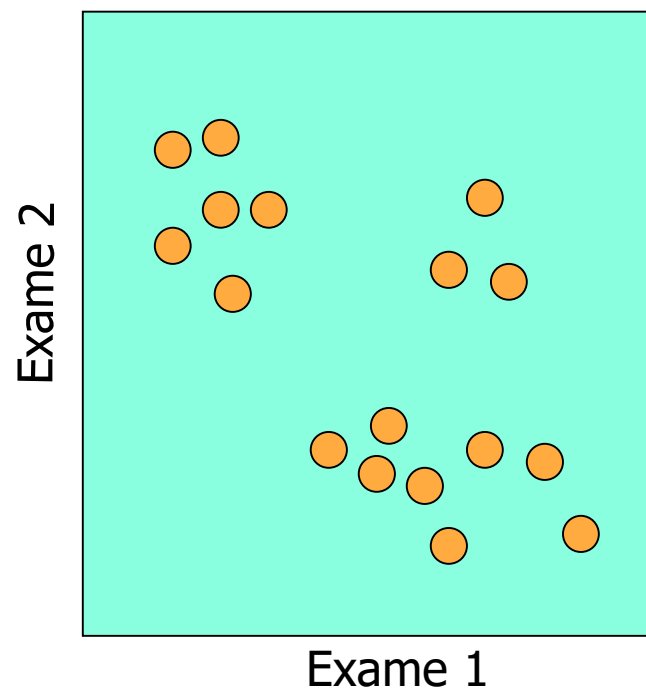
Agrupamento



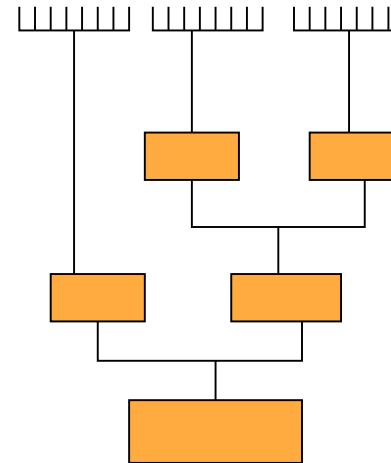
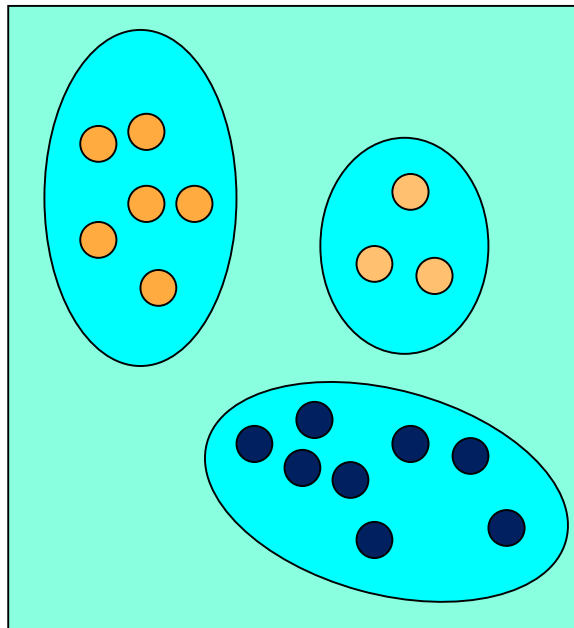
Copyright © 2020. Todos os direitos reservados ao CeMEAI-USP. Proibida a cópia e reprodução sem autorização



Agrupamento



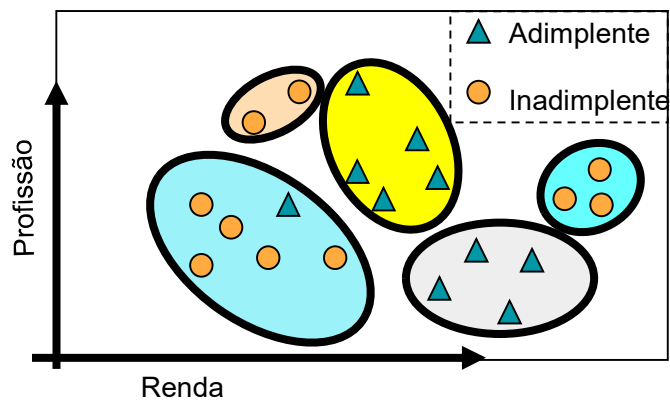
Particional vs hierárquico



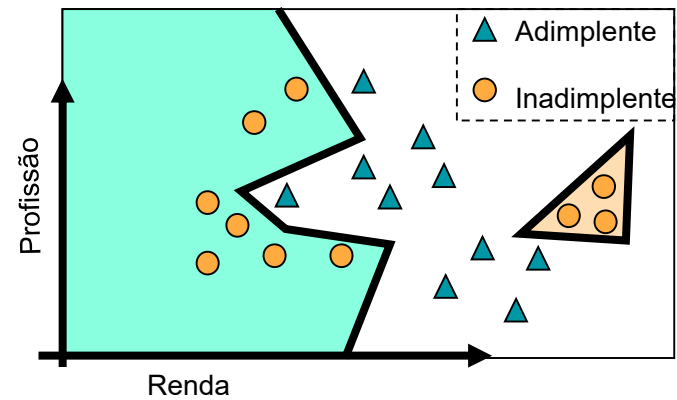
Algoritmos de agrupamento

- Redes Neurais SOM
- K-médias
- K-medianas
- FCM
- DBSCAN
- Single-Link

Modelagem por AM



Descritivo
Agrupamento



Preditivo
Classificação

Sumarização

- Objetivo: encontrar descrição simples e compacta para um conjunto de dados
- Frequentemente utilizada para:
 - Exploração interativa de dados
 - Geração automática de relatórios
 - Exemplo:
 - Definir perfis de compras feitas nos finais de semana em um supermercado

Sumarização

Nome	Idade	Sexo	Altura	Tem filhos
João	32	M	180	S
Maria	30	F	-----	N
Pedro	23	M	160	S
José	45	M	170	S
Sueli	18	F	175	N

Sumarização

Nome	Idade	Sexo	Altura	Tem filhos
João	32	M	180	S
Maria	30	F	-----	N
Pedro	23	M	160	S
José	45	M	170	S
Sueli	18	F	175	N

Maior altura: 180
Sexo mais frequente: M
Mediana da idade: 30
Idade média: 29.6

Regras de Associação

- Objetivo: dado um conjunto de itens e uma base de dados de transações
 - Encontrar um conjunto de regras de associação entre os itens
- Exemplo:
 - Procurar por itens que são frequentemente comprados juntos
 - Itens frequentes

Regras de Associação

Transação	Itens comprados
1	pão, queijo, manteiga, massa
2	pão, geléia, suco
3	queijo, arroz, massa
4	queijo, vinho
5	massa, queijo, pão

Regras de Associação

Transação	Itens comprados
1	pão, queijo, manteiga, massa
2	pão, geléia, suco
3	queijo, arroz, massa
4	queijo, vinho
5	massa, queijo, pão

40% dos clientes compram pão e queijo
75% dos clientes que compram queijo
também compram massa

Memória associativa

- Algoritmo associa um dado padrão de entrada a um padrão de saída
 - Hetero-associativa
 - Auto-associativa
 - Associa um vetor (padrão) a ele mesmo
- Exemplos
 - Associar voz à imagem, recuperar itens de um BD utilizando eles mesmos como endereços
 - Codificar um vetor (autoencoder)

Aprendizado prescritivo

- Prescreve que entrada é necessária para gerar uma dada saída
 - Contrário do aprendizado preditivo
 - Ao invés de prever o que vai acontecer, sugerir o que fazer para que algo aconteça
- Exemplo
 - Controle de processos químicos
 - Controle de robôs
 - Gera entrada de controle para que um sistema siga uma trajetória especificada por um modelo de referência

Tópicos

- Aprendizado de Máquina
- O que faz?
- AM x computação convencional
- Aplicações
- Tarefas
- Abordagens

Final da Apresentação

Copyright © 2020. Todos os direitos reservados ao CeMEAI-USP. Proibida a cópia e reprodução sem autorização

