

# **Inteligência Artificial (grad) Sistemas Inteligentes (pós)**

---

**Prof. Flávio Varejão  
Informática  
Universidade Federal do Espírito Santo**

# Sumário

---

- Objetivos
  - Conteúdo
  - Bibliografia
  - Material Didático
  - Avaliação
  - Calendário
  - História
  - Inteligência Artificial, Filosofia e Ética
  - Problemas de Interesse
-

# Objetivos

---

- Proporcionar uma visão geral sobre Inteligência Artificial
- Apresentar várias técnicas utilizadas em Sistemas com Inteligência Artificial

# Conteúdo

---

- Introdução
- Aprendizado de Máquina
  - Classificação, Regressão, Agrupamento, Associação
- Busca
  - Tabu, Resfriamento Simulado, Grasp, Genético
- Representação de Conhecimento
  - Sistemas Baseados em Conhecimento
  - Sistemas de Produção

# Metodologia de Ensino

---

- Aulas Teóricas em Sala de Aula
- Aulas Práticas em Laboratório
  - Roteiros de Atividades Semanais
- Presença requerida
- Avaliação
  - Duas Provas
  - Dois Trabalhos

$$\text{Média Parcial} = 0.4 P1 + 0.4 P2 + 0.1 T1 + 0.1 T2$$

# Bibliografia

---

- Disponíveis em  
<https://bibliotecas-digitais.ufes.br/> Minha  
Biblioteca
  - FACELLI, K.; Lorena, A.C.; Gama, J.;  
Carvalho, A.C.P.L.F.; Inteligência Artificial  
– Uma Abordagem de Aprendizado de  
Máquina, 2011.
  - NORVIG, Peter; Russel, Stuart. Inteligência  
Artificial. 3a. ed. Rio de Janeiro: Campus,  
2013.

# Bibliografia

---

- Complementares
  - MULLER, Andreas C. and Guido, Sarah. Introduction to Machine Learning with Python: A Guide to Data Scientists, 2017
  - RASCHKA, Sebastian; Python Machine Learning, 2016
- Artigos Diversos

# Material Didático

---

- Slides, Gravações de Aulas, Video Aulas e Roteiros de aulas disponibilizados
- Sites Internet
  - <https://towardsdatascience.com/>
- Cursos online
  - Andrew Ng
  - <https://youtube.com/playlist?list=PLxfEOJXRm7eZKJyovNH-IE3ooXTsOCvfC>

# Material Didático

---

- Python
  - Pré-requisito
  - Jupyter IDE
  - Suas bibliotecas para ciência de dados e aprendizado de máquina
    - Matplotlib
    - Seaborn
    - Pandas
    - Numpy
    - Scipy
    - Scikit-learn

# Calendário

---

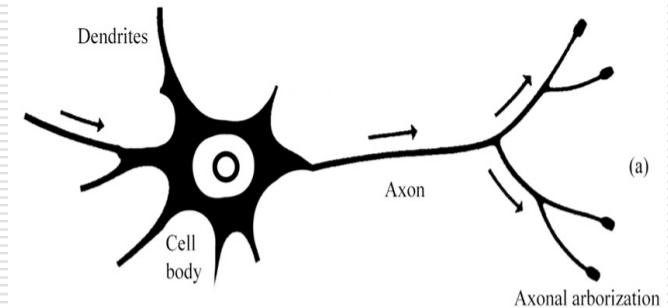
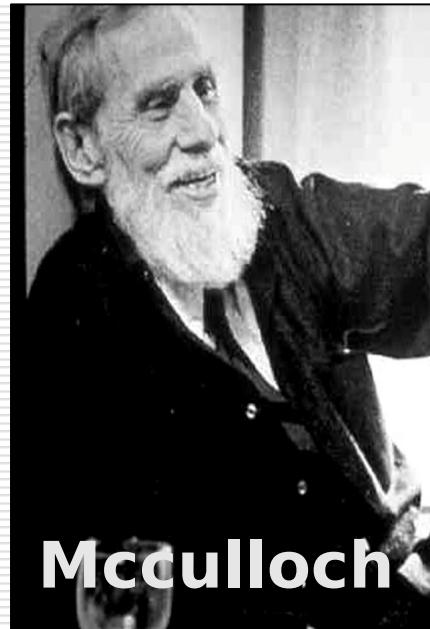
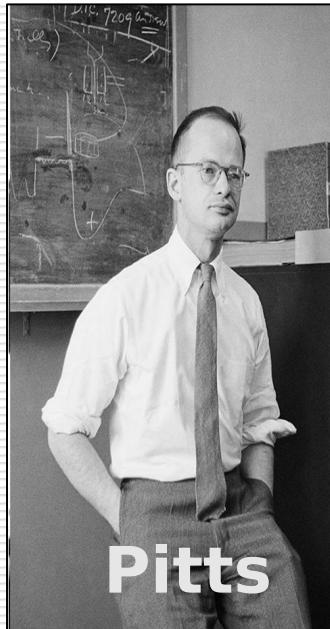
- T1: 29/04 → 03/06
- T2: 03/06 → 03/07
- P1: 29/05
- P2: 08/07
- PF: 15/07

# Breve Histórico da IA

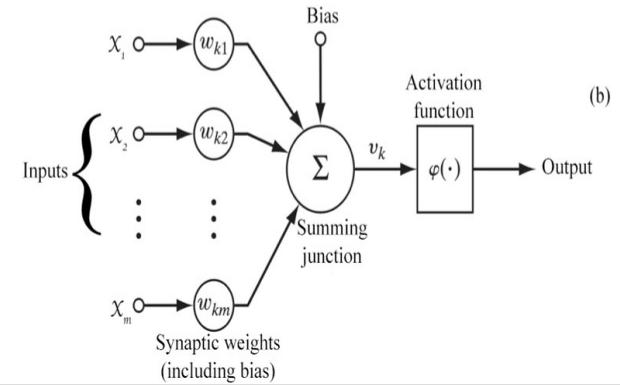
---

- 1936: tese de Church-Turing
  - Computadores digitais podem simular qualquer processo de raciocínio formal
- 1943: Neurônio artificial de McCulloch e Pitts
- 1950: Turing publica Computing Machinery and Intelligence
- 1956: John McCarthy cria o termo Inteligência Artificial
  - Primeiras aplicações (toy problems)
  - Foco em abordagem simbólica

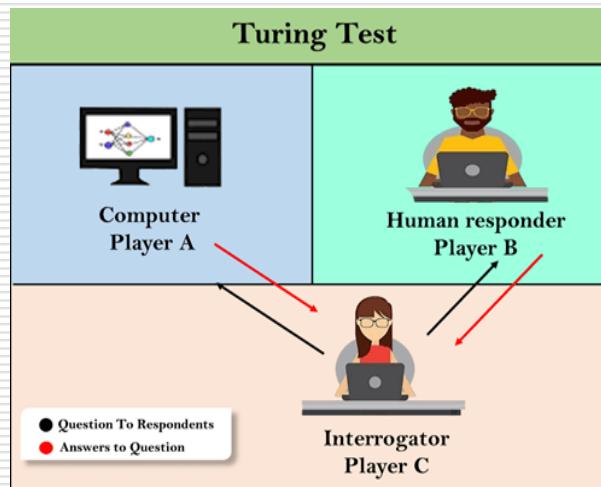
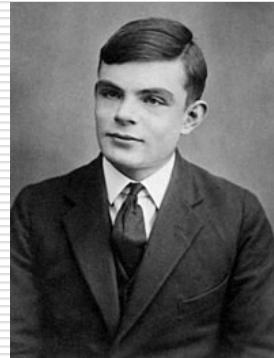
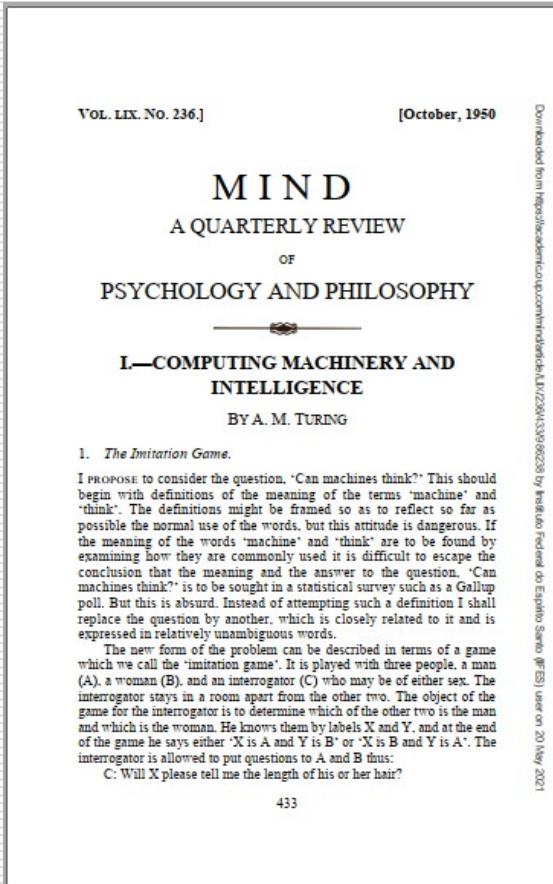
# Breve Histórico da IA (1943)



McCulloch-Pitts Neuron



# Breve Histórico da IA (1950)



# Breve Histórico da IA (1956)

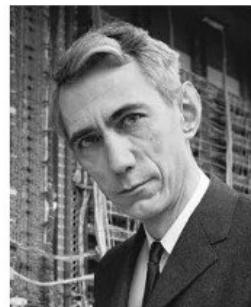
## 1956 Dartmouth Conference:



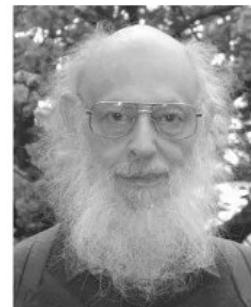
John MacCarthy



Marvin Minsky



Claude Shannon



Ray Solomonoff



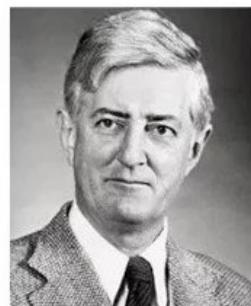
Alan Newell



Herbert Simon



Arthur Samuel



Oliver Selfridge



Nathaniel Rochester



Trenchard More

# Breve Histórico da IA (1962)

---



## Perceptrons are awesome

Redes neurais capazes de operar sobre dados perceptuais complexos (e.g., imagens), alcançaram resultados iniciais promissores e possuem garantias de convergência.

**Frank Rosenblatt**

---

# Breve Histórico da IA (1969)

---



**Perceptrons  
are not that  
awesome,  
man.**

O livro Perceptrons escrito por Minsky e Papert mostrou que perceptrons só eram capazes de resolver problemas linearmente separáveis.

# Breve Histórico da IA

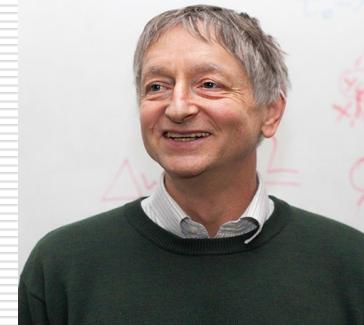
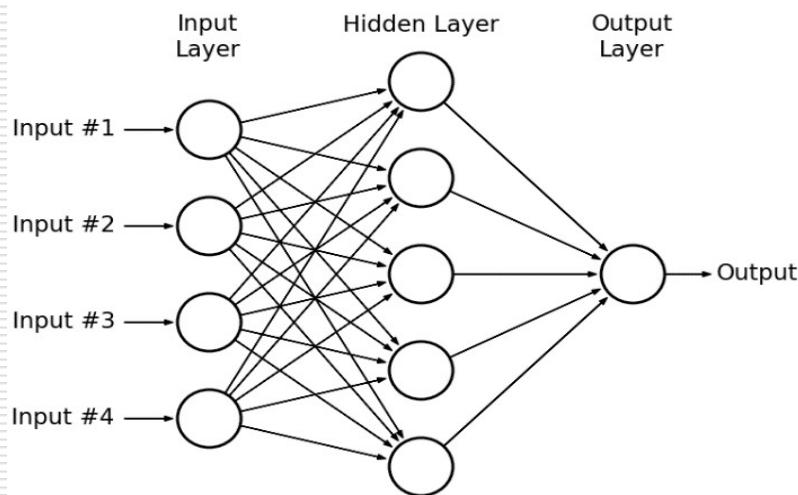
---

- 1971/72: Teoria da Complexidade (Cook e Karp)
  - 1975-2000: Inverno da IA
    - Sistemas Especialistas
  - 1995-2010: Machine Learning
    - Estatística
    - Backpropagation/Redes Neurais Artificiais
    - Inteligência Computacional...
  - 2010-: Deep Learning + GPUs
  - 2015-: Verão da IA
-

# Breve Histórico da IA (1985)

Perceptrons com múltiplas camadas podem ser treinados por propagação de erros para trás e eles são capazes de resolver problemas não linearmente separáveis

Rumelhart, Hinton & Williams



# Sistemas Lógicos e Sistemas Especialistas

## Base de Conhecimento:

1. consulta\_clinico => ped\_hemograma
2. ped\_hemograma and fez\_hemograma => tem\_hemograma
3. tem\_hemograma and consulta\_anestesista => tem\_laudo

Conhecimento and  
Observação liberada  
consulta\_clinico  
fez\_hemograma  
consulta\_anestesista

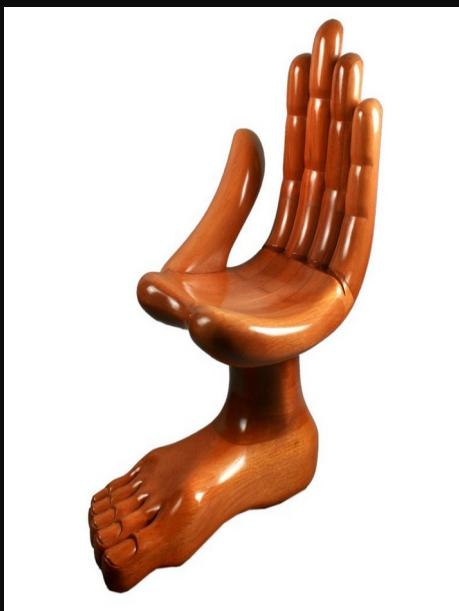
Pergunta (Query):>  
tem\_hemograma and  
cirurgia\_liberada?  
**Resposta:**  
True

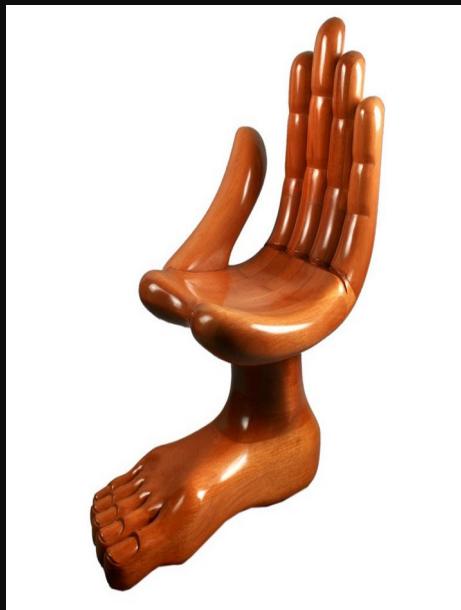
# Sistemas Lógicos e Sistemas Especialistas

- Requer conhecimento de especialistas
- Requer tempo e esforço de especialistas em modelagem
- Se torna muito complexo em alguns contextos como dados perceptuais complexos

O que é uma  
cadeira?

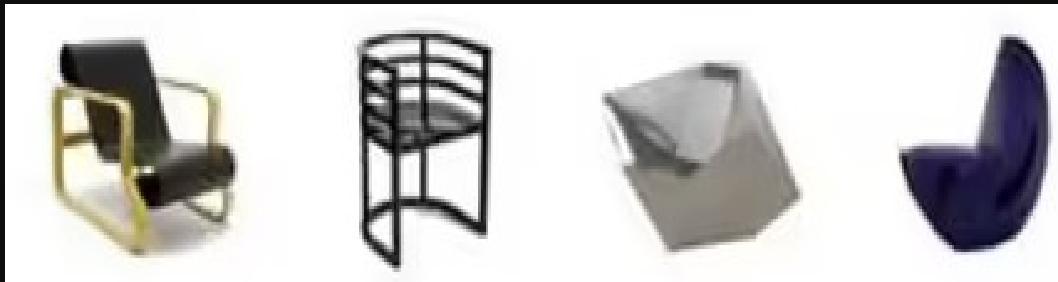
1) algo com 4  
pernas e uma  
superfície plana  
em cima



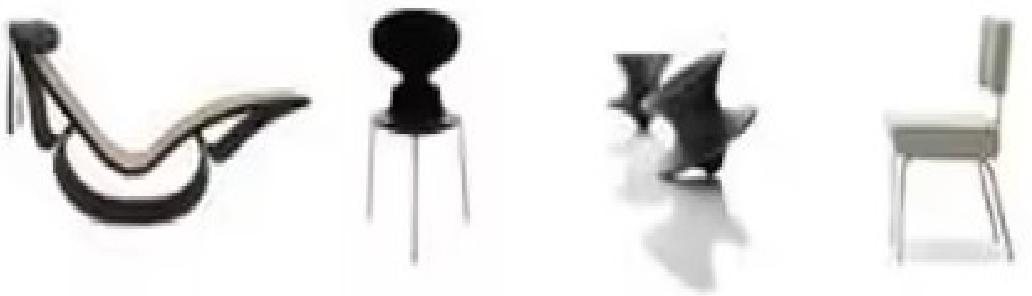


2) São coisas  
usadas para  
sentar em cima





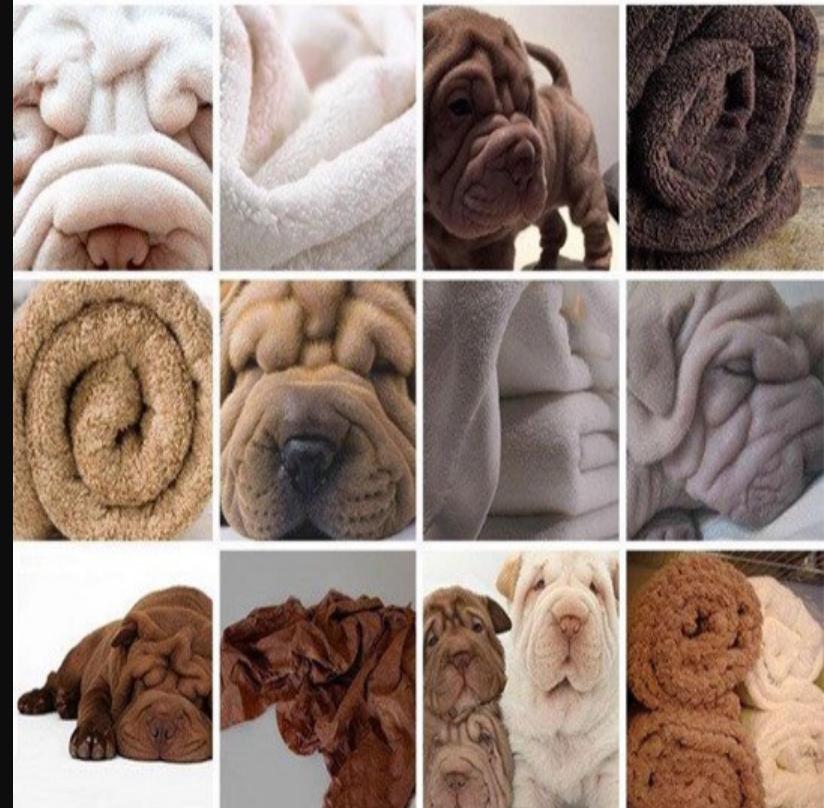
# CHAIRS



# CAT OR CROISSANT?



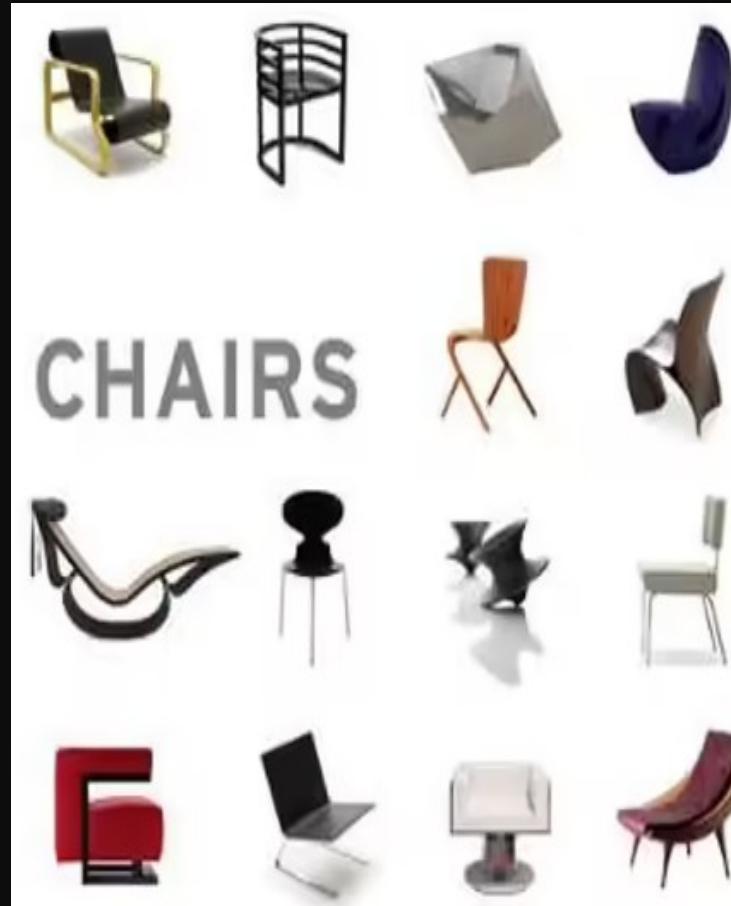
Shar Pei or Towel?



**É MUITO DIFÍCIL CRIAR REGRAS PARA LIDAR COM  
DADOS PERCEPTUAIS (E.G. VISÃO)**

# Como crianças aprendem?

cadeiras



# Como crianças aprendem?

cadeiras?

não!



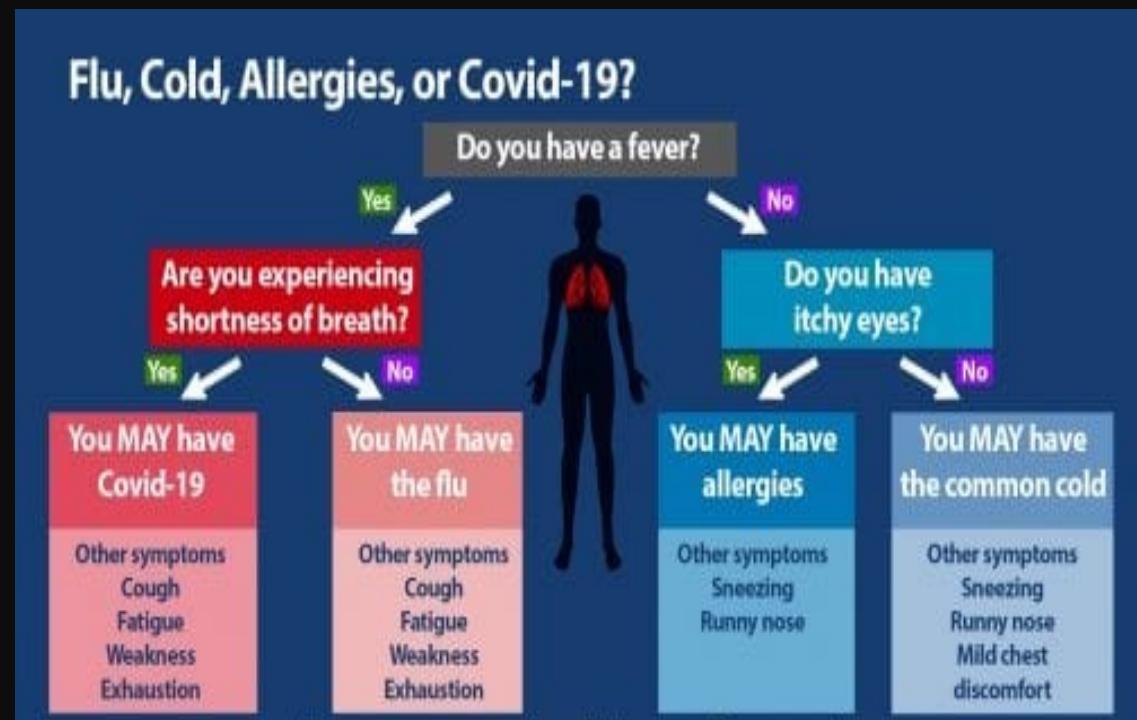
# Aprendizado de Máquina

- Conhecimento nos dados
- Identificação de regularidades / padrões
- Tarefas como problemas de otimização
- Fundamentos em Estatística
- Utiliza Álgebra Linear e Cálculo



# Exames e Prontuários Médicos Passados

# Algoritmo de ML



# A Revolução do Deep Learning

Redes neurais são utilizadas em diversas áreas, tanto na academia quanto em empresas.  
Investimentos são abundantes.

# 2016

AlphaGo wins  
4/5 games  
against Mr Lee  
Sedol, the  
winner of 18  
world titles of  
in the game of  
Go.

<https://deepmind.com/research/case-studies/alphago-the-story-so-far>



Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... & Dieleman, S. (2016). Mastering the game of Go with deep neural networks and tree search. *nature*, 529(7587), 484.



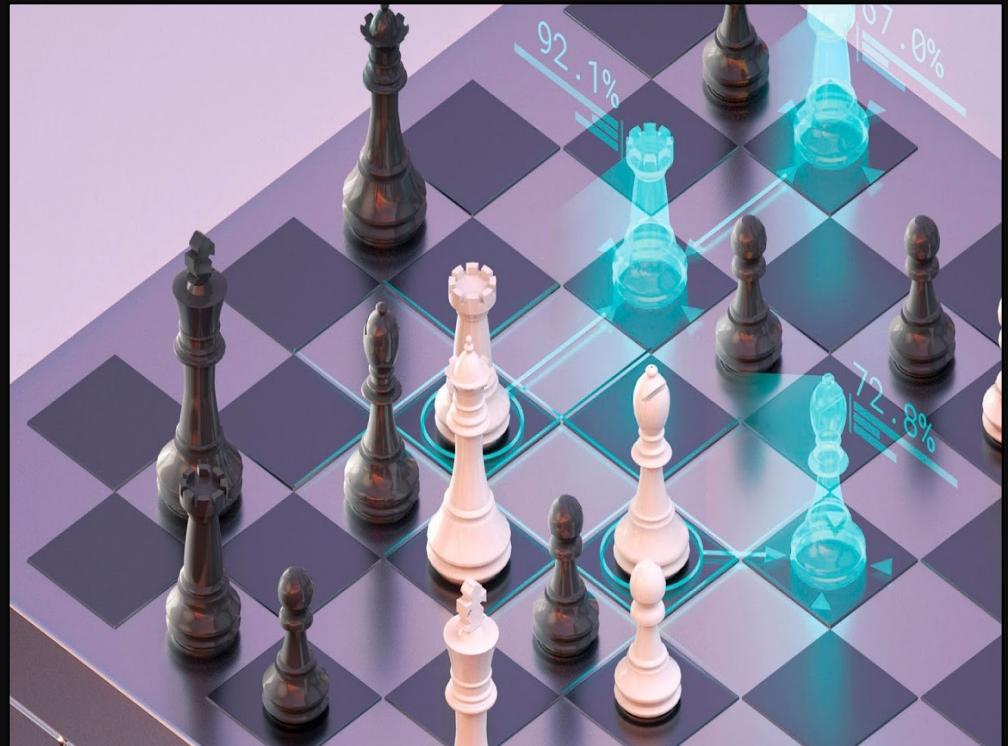
<https://www.youtube.com/watch?v=qhUvQiKec2U>

<https://www.nvidia.com/pt-br/self-driving-cars/drive-platform/>

201  
6  
NVIDIA  
presented  
a  
neural network

AlphaZero  
taught  
itself  
chess,  
shogi, and  
Go,  
beating a  
world-  
champion  
program in  
each case.

2017



[https://deepmind.com/research/case-studies/alpha  
go-the-story-so-far](https://deepmind.com/research/case-studies/alpha-go-the-story-so-far)

## Language Models are Few-Shot Learners

Tom B. Brown<sup>\*</sup> Benjamin Mann<sup>\*</sup> Nick Ryder<sup>\*</sup> Melanie Subbiah<sup>\*</sup>  
 Jared Kaplan<sup>†</sup> Prafulla Dhariwal Arvind Neelakantan Pranav Shyam Girish Sastry  
 Amanda Askell Sandhini Agarwal Ariel Herbert-Voss Gretchen Krueger Tom Henighan  
 Rewon Child Aditya Ramesh Daniel M. Ziegler Jeffrey Wu Clemens Winter  
 Christopher Hesse Mark Chen Eric Sigler Mateusz Litwin Scott Gray  
 Benjamin Chess Jack Clark Christopher Berner  
 Sam McCandlish Alec Radford Ilya Sutskever Dario Amodei  
 OpenAI

### Abstract

Recent work has demonstrated substantial gains on many NLP tasks and benchmarks by pre-training on a large corpus of text followed by fine-tuning on a specific task. While typically task-agnostic in architecture, this method still requires task-specific fine-tuning datasets of thousands or tens of thousands of examples. By contrast, humans can generally perform a new language task from only a few examples or from simple instructions – something which current NLP systems still largely struggle to do. Here we show that scaling up language models greatly improves task-agnostic, few-shot performance, sometimes even reaching competitiveness with prior state-of-the-art fine-tuning approaches. Specifically, we train GPT-3, an autoregressive language model with 175 billion parameters, 10x more than any previous non-sparse language model, and test its performance in the few-shot setting. For all tasks, GPT-3 is applied without any gradient updates or fine-tuning, with tasks and few-shot demonstrations specified purely via text interaction with the model. GPT-3 achieves strong performance on many NLP datasets, including translation, question-answering, and cloze tasks, as well as several tasks that require on-the-fly reasoning or domain adaptation, such as unscrambling words, using a novel word in a sentence, or performing 3-digit arithmetic. At the same time, we also identify some datasets where GPT-3’s few-shot learning still struggles, as well as some datasets where GPT-3 faces methodological issues related to training on large web corpora. Finally, we find that GPT-3 can generate samples of news articles which human evaluators have difficulty distinguishing from articles written by humans. We discuss broader societal impacts of this finding and of GPT-3 in general.

<sup>\*</sup>Equal contribution  
<sup>†</sup>Johns Hopkins University, OpenAI

Author contributions listed at end of paper.

# 2020 - GPT-3

Title: United Methodists Agree to Historic Split

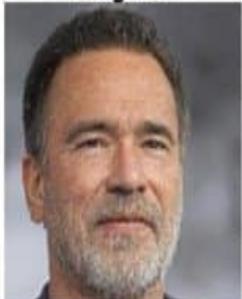
Subtitle: Those who oppose gay marriage will form their own denomination

Article: After two days of intense debate, the United Methodist Church has agreed to a historic split – one that is expected to end in the creation of a new denomination, one that will be “theologically and socially conservative,” according to The Washington Post. The majority of delegates attending the church’s annual General Conference in May voted to strengthen a ban on the ordination of LGBTQ clergy and to write new rules that will “discipline” clergy who officiate at same-sex weddings. But those who opposed these measures have a new plan: They say they will form a separate denomination by 2020, calling their church the Christian Methodist denomination.

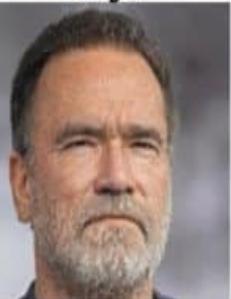
The Post notes that the denomination, which claims 12.5 million members, was in the early 20th century the “largest Protestant denomination in the U.S.,” but that it has been shrinking in recent decades. The new split will be the second in the church’s history. The first occurred in 1968, when roughly 10 percent of the denomination left to form the Evangelical United Brethren Church. The Post notes that the proposed split “comes at a critical time for the church, which has been losing members for years,” which has been “pushed toward the brink of a schism over the role of LGBTQ people in the church.” Gay marriage is not the only issue that has divided the church. In 2016, the denomination was split over ordination of transgender clergy, with the North Pacific regional conference voting to ban them from serving as clergy, and the South Pacific regional conference voting to allow them.

Figure 3.14: The GPT-3 generated news article that humans had the greatest difficulty distinguishing from a human written article (accuracy: 12%).

Original



Anger



Disgust



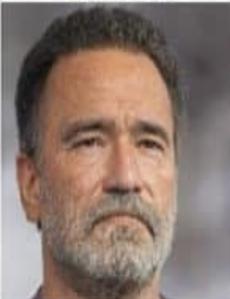
Fear



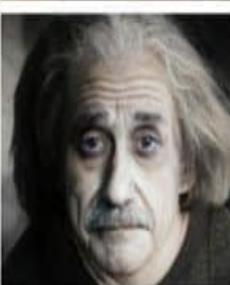
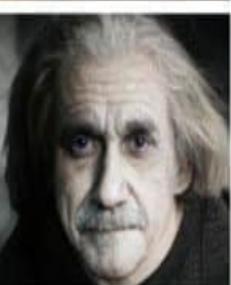
Happiness



Sadness



Surprise



TEXT PROMPT

an armchair in the shape of an avocado [...]

AI-GENERATED IMAGES



[Edit prompt or view more images ↓](#)

TEXT PROMPT

a store front that has the word 'openai' written on it [...]

AI-GENERATED IMAGES



# DALL·E: Creating Images from Text

We've trained a neural network called DALL-E that creates images from text captions for a wide range of concepts expressible in natural language.



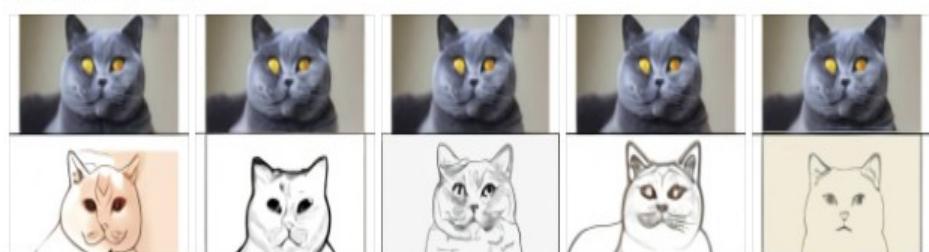
January 5, 2021

27 minute read

TEXT AND IMAGE PROMPT

the exact same cat on the top as a sketch on the bottom

AI-GENERATED IMAGES



[Edit prompt or view more images ↓](#)

# Questões Filosóficas

---

- O que é Inteligência?
- Inteligência é propriedade humana?
- Máquinas podem pensar?
- Máquinas podem ter consciência
- Máquinas podem aprender?

# O que é Inteligência?

---

- Definições dos Alunos
- Aurélio
  - Capacidade de aprender, raciocinar, pensar, etc.
- McCarthy
  - Parte computacional da habilidade de atingir objetivos no mundo
- Ricardo Machado
  - Capacidade que alguém possui de fazer algo que não sabemos como fazer

# Atividades Inteligentes

---

- Quais das seguintes atividades é inteligente?
  - Matemático deduzindo teoremas
  - Rato encontrando saída em labirinto
  - Ser humano explicando alguma ação
  - Máquina de calcular operando
  - Humano fazendo programas de computador
  - Vírus fugindo de anticorpos
  - Castor construindo represa

# Inteligência Humana

---

- Inteligência é exclusividade humana?
  - Constrarstar definições de inteligência com atividades inteligentes
  - Ações decorrentes de instinto, adestramento, inconsciente, intuição são inteligentes?
  - Inteligência é simbólica?
    - Visão Neural X Visão Newell
  - Inteligência é relacionada a linguagem?
  - Uso de linguagem produz conhecimento?
    - O que é conhecimento?

# Máquinas Inteligentes

---

- Máquinas podem pensar?
  - Computing Machinery and Intelligence
    - Alan Turing, 1950
  - Teste de Turing
    - Um investigador humano fazendo perguntas
    - Dois investigados desconhecidos respondendo sem serem visualizados
    - Se investigador não puder diferenciar, prova-se que máquinas podem pensar
  - Lista de Objeções
  - Máquinas que Aprendem

# Máquinas Inteligentes

---

- Lista de Objeções
  - Teológica
    - Pensar é função exclusiva da alma humana
    - Animais não possuem almas?
  - Temerosa
    - Seria muito temeroso se máquinas pudessesem pensar
    - Ser humano tem de ser necessariamente superior
      - Perderíamos humanidade
      - Narcisismo humano?
    - Choques no Ego: Copérnico, Darwin e Freud

# Máquinas Inteligentes

---

- Lista de Objeções
  - Matemática
    - Teorema da Incompletude de Godel
    - Sistemas formais sempre apresentam sentenças indecidíveis
    - Questões significativas para julgar inteligência?
    - Limitação não se aplica também ao intelecto humano?

# Máquinas Inteligentes

---

- Lista de Objeções
  - Ausência de Consciência
    - Máquina não se emocionaria ao escrever um soneto
    - Não podemos provar que outras pessoas são conscientes ou sentem algo. Porque exigir isso de uma máquina?
  - Teste da Câmara Chinesa de Searle
    - Máquina não “sabe” realmente o que escreve
    - Apenas faz o que a gente manda elas fazerem
  - Pessoas só fazem o que os seus neurônios dizem para fazer
  - ChatGPT Consciente

# Máquinas Inteligentes

---

- Lista de Objeções
  - Ausência de Habilidades
    - Máquinas nunca poderão fazer atividade tal!
      - Mas podem fazer atividades que não fazemos
      - Quais são as atividades significativas?
    - Argumento baseado em indução e pouco desenvolvimento das máquinas atuais?
  - Continuidade dos sistemas neurais
    - Máquinas de estados discretos e finita
    - Máquinas podem simular comportamento contínuo?

# Máquinas Inteligentes

---

- Lista de Objeções
  - Ausência de Originalidade
    - Máquinas nunca serão capazes de criar algo realmente novo!
    - Máquinas são incapazes de nos surpreender?
      - Jogo do Pato no ChatGpt
      - Pinturas criadas por Midjourney
        - » A painting mixing art nouveau and cubism of a bedroom
        - » A painting of a sunflower field in impressionist style
        - » A painting representing the pain and shame of brazilian carandiru episode in cubist style

# Máquinas Criativas e Originais?

---



# Máquinas Inteligentes

---

- Lista de Objeções
  - Informalidade do Comportamento
    - Não é possível produzir um conjunto de regras que descrevam antecipadamente o que o ser humano deve fazer em todas as possíveis circunstâncias
    - Em última instância, não somos regidos pelas leis da natureza, as quais determinam nosso comportamento?
  - Ausência de Percepção Extrasensorial
    - Máquinas nunca serão capazes de fazer telepatia, clarividência, etc!
    - Quem acredita nisso não deveria acreditar em tudo?

# Máquinas Inteligentes

---

- Lista de Objeções
  - Intratabilidade Computacional
    - Não consta na lista de objeções do artigo de Turing
    - Problemas intratáveis quando tempo de resolução cresce exponencialmente com tamanho da instância do problema
      - NP
    - Achamos a solução ótima desses problemas???
    - Exigência maior da máquina do que de nós mesmos???

# Previsão de Turing

---

- Em 50 anos máquinas passarão pelo teste!
  - Não se confirmou!!!
  - Pode vir a se confirmar?
  - Quão longe estamos?
    - ChatGPT (2022)

# Questões Éticas

---

- Desemprego
  - Computação reduz empregos pouco especializados
  - IA reduz/reduzirá empregos especializados
- Perda de privacidade
- Responsabilidade pela falha da máquina
- Máquinas tomarem o controle
- Perda de Humanidade
  - Somos diferentes?
  - Podemos desligá-las?

# O que Máquinas já fazem

---

- Jogar tênis de mesa
  - Dirigir de forma segura em uma estrada
  - Fazer compra semanal de supermercado
  - Jogar xadrez, bridge, etc.
  - Descobrir e provar um novo teorema
  - Escrever intencionalmente uma estória engraçada
  - Consultoria jurídica
  - Realizar uma cirurgia complexa
  - Tradução de linguagem natural
  - Conversar com uma pessoa durante uma hora
  - Fazer uma pesquisa científica
-

# Abordagens

---

- IA Fraca
  - Focada em uma tarefa específica
  - Muitos exemplos de sucesso
- IA Forte
  - Máquina capaz de aplicar inteligência a qualquer problema
  - Ainda em desenvolvimento?

# Problemas de Interesse de IA/SI

---

- Busca
- Classificação
- Regressão
- Agrupamento
- Associação

# Busca

---

- Satisfação de Restrições
  - $X = \{X_1, \dots, X_n\}$ : conjunto de variáveis
  - $D = \{D_1, \dots, D_n\}$ : domínio de valores
  - $C = \{C_1, \dots, C_m\}$ : conjunto de restrições
    - $C_i = (t_i, R_i)$ 
      - $t_i$  é subconjunto de  $X$
      - $R_i$  é uma relação envolvendo variáveis em  $t_i$
    - $C_i$  é satisfeita se valores atribuídos a  $t_i$  fazem relação  $R_i$  ser verdadeira

# Mochila 0/1

---

- Mochila de tamanho  $s$
- Conjunto de itens de tamanho  $t_i$ 
  - $t = \{t_1, \dots, t_n\}$
- $X = \{X_1, \dots, X_n\}$ : conjunto de variáveis binárias
- $D = \{D_1, \dots, D_n\}$ : domínio de valores
  - $D_i = \{0, 1\}$
- $C = \{C_1\}$ : uma única restrição
  - $C_1 = \sum (x_i \cdot t_i) \leq s$

# Mochila 0/1

---

## □ Exemplo

- $s = 8$
- $t = \{1, 7, 4\}$
- Soluções:
  - $\{0, 0, 0\} \Rightarrow t = \{\}, s = 0$
  - $\{1, 0, 0\} \Rightarrow t = \{1\}, s = 1$
  - $\{0, 1, 0\} \Rightarrow t = \{7\}, s = 7$
  - $\{0, 0, 1\} \Rightarrow t = \{4\}, s = 4$
  - $\{1, 1, 0\} \Rightarrow t = \{1, 7\}, s = 8$
  - $\{1, 0, 1\} \Rightarrow t = \{1, 4\}, s = 5$

# Busca

---

- Otimização
  - Especialização de Satisfação de Restrições
  - Adiciona funções objetivos  $F$ 
    - $F = \{f_1, \dots, f_k\}$ : conjunto de funções objetivo
  - Objetivo é minimização ou maximização de  $f_j$ 
    - $M = \{M_1, \dots, M_k\}$ : conjunto de operadores de minimização ou maximização
    - $M_j = \{\min, \max\}$
  - $k > 1 \Rightarrow$  Otimização Multiobjetivo

# Otimização

---

- Combinatória
  - $X_i$  são todas discretas
- Contínua
  - $X_i$  são todas contínuas
- Mista
  - $X_i$  pode ser discreta ou contínua

# Mochila 0/1

---

- Exemplo
  - $s = 8$
  - $T = \{1, 7, 4\}$
  - $C_1 = \sum (x_i \cdot t_i) \leq s$
  - $F = \{f(X) = \sum (x_i \cdot t_i)\}$
  - $M = \{\max\}$
  - Solução:
    - $\{1, 1, 0\} \Rightarrow t = \{1, 7\}, s = 8$

# Classificação

---

- Variáveis Descritoras de Instância
- Variáveis de Decisão Binárias
  - Rótulos Nominais designam Classes
- Exemplos
  - Diagnóstico Médico
  - Descritores: Sintomas
  - Decisão: Doença
  - Detecção de Fraude
  - Descritores: Dados da Operação
  - Decisão: Existência de Fraude

# Classificação

---

- $Y = h(X)$ 
  - $X = \{X_1, \dots, X_n\}$ : conjunto de variáveis
  - $D = \{D_1, \dots, D_n\}$ : domínio de valores
  - $L = \{L_1, \dots, L_m\}$ : conjunto de rótulos nominais
  - $Y = \{Y_1, \dots, Y_m\}$ : conjunto de variáveis binárias
  - $h$ : Classificador
    - Mapeia instância  $i$  do espaço  $n$  dimensional de  $X$  para espaço  $m$  dimensional de  $Y$
    - $\{y_{i1}, \dots, y_{im}\} = h(x_{i1}, \dots, x_{in})$ 
      - $x_{ij} \in D_j$

# Classificação

---

## □ Tipos

- Binária
  - $m = 1$
- Multiclasse
  - $m \geq 2$
  - $\sum y_{ij} = 1$
- Multirótulo
  - $m \geq 2$
  - $\sum y_{ij} \geq 0$

# Classificação

---

- Exemplo
  - Decisão sobre ir à praia
  - $X = \{\text{Aparência, Ventando, Umidade}\}$
  - $D_a = \{\text{Sol, Chuva, Nublado}\}$
  - $D_v = \{\text{Sim, Não}\}$
  - $D_u = \{0..100\}$
  - $L = \{\text{Ir}\}$
  - $Y = \{Y_1\}$
  - $Y_{i1} = h(x_{ia}, x_{iv}, x_{iu})$

# Ir à praia

---

## □ Exemplo de classificador

- $Y_{i1} = h_j(x_{ia}, x_{iv}, x_{iu})$
- $h_j$

Se  $x_{ia} = \text{Sol} \wedge x_{iv} = \text{Não}$  Então

$$Y_{i1} = 1$$

– Se  $x_{ia} = \text{Nublado} \wedge x_{iu} > 60$  Então

$$Y_{i1} = 1$$

Senão

$$Y_{i1} = 0$$

# Regressão

---

- Semelhante a Classificação
- Variáveis de Decisão Contínuas
- Exemplo
  - Série Temporal
    - Cotação Diária do Dólar
    - Relação de Variáveis em Fenômenos Físicos
    - Pressão e Temperatura

# Regressão

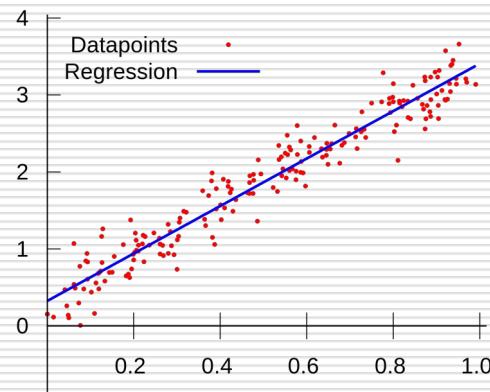
---

- $Y = r(X)$ 
  - $X = \{X_1, \dots, X_n\}$ : conjunto de variáveis
  - $D = \{D_1, \dots, D_n\}$ : domínio de valores
  - $Y = \{Y_1, \dots, Y_m\}$ : conjunto de variáveis contínuas
  - $r$ : Regressor
    - Mapeia instância  $i$  do espaço  $n$  dimensional de  $X$  para espaço  $m$  dimensional de  $Y$
    - $\{y_{i1}, \dots, y_{im}\} = r(x_{i1}, \dots, x_{in})$ 
      - $x_{ij} \in D_j$

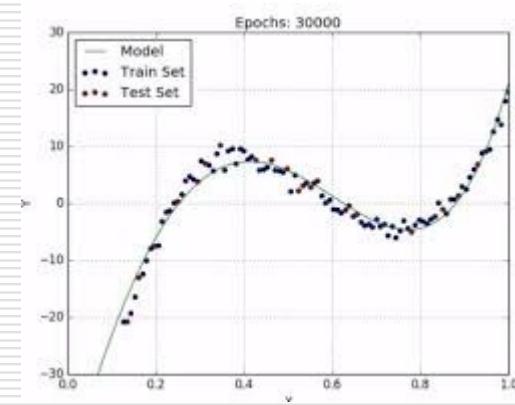
# Tipos de Regressão

---

- Univariada
  - $n = 1$
- Multivariada
  - $n > 1$
- Linear

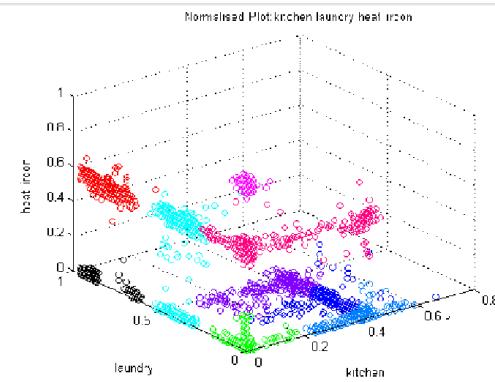
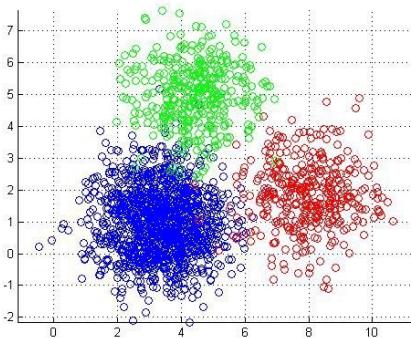


Não Linear



# Análise de Agrupamento de Dados

- Semelhança entre instâncias
- Exemplo
  - Perfil de Clientes
  - Sugestão de compra
  - Clientes de seu perfil se interessam por



# Análise de Agrupamento de Dados

---

## □ $Y = c(Z)$

- $X = \{X_1, \dots, X_n\}$ : conjunto de variáveis
- $D = \{D_1, \dots, D_n\}$ : domínio de valores
- $Z = \{z_i \mid z_i = (x_{i1}, \dots, x_{in}), x_{ij} \in D_j\}$ : conjunto de elementos em espaço n dimensional de X
- $Y = \{Y_1, \dots, Y_k \mid Y_j \subseteq Z\}$ : conjunto de subconjuntos de dados de Z
- $c$ : Agrupador
  - Mapeia elementos de Z (em espaço n dimensional de X) para espaço k dimensional de Y
  - $\{Y_1, \dots, Y_k\} = c(Z)$

# Análise de Agrupamento de Dados

---

## □ Tipos

- Número de Partições
  - Previamente conhecidas
    - $k$  é entrada do problema
  - Desconhecidas
    - $k$  é determinado pela solução
- Conjuntos da Solução
  - Disjuntos
    - $z_i$  pertence a um único  $Y_j$
  - Difusos
    - $z_i$  pode pertencer a mais de um  $Y_j$

# Análise de Agrupamento de Dados

---

- Tipos
  - Cobertura
    - Completa
      - $\bigcup Y_j = Z$
    - Incompleta
      - $\bigcup Y_j \subset Z$
- Mais comum
  - Partições Previamente Conhecidas
  - Soluções Disjuntas
  - Cobertura Completa

# Análise de Agrupamento de Dados

---

## □ Exemplo

- $k = 3$
- $Z = \{7.3, 3.8, 9.4, 8.7, 9.9, 8.8, 1.0, 2.1, 7.2\}$
- $Y = \{ \{1, 2.1\}, \{3.8\}, \{7.2, 7.3, 8.7, 8.8, 9.4, 9.9\} \}$

# Associação

---

- Relações entre variáveis
- Exemplo
  - Sistemas de Recomendação
  - Sugestão de compra
    - Quem compra tal item também compra este

Transações	Itens comprados
1	calça, camiseta, tênis
2	camiseta, tênis
3	bermuda, tênis
4	calça, sandália

Solução: Quem compra camiseta compra tênis

---