

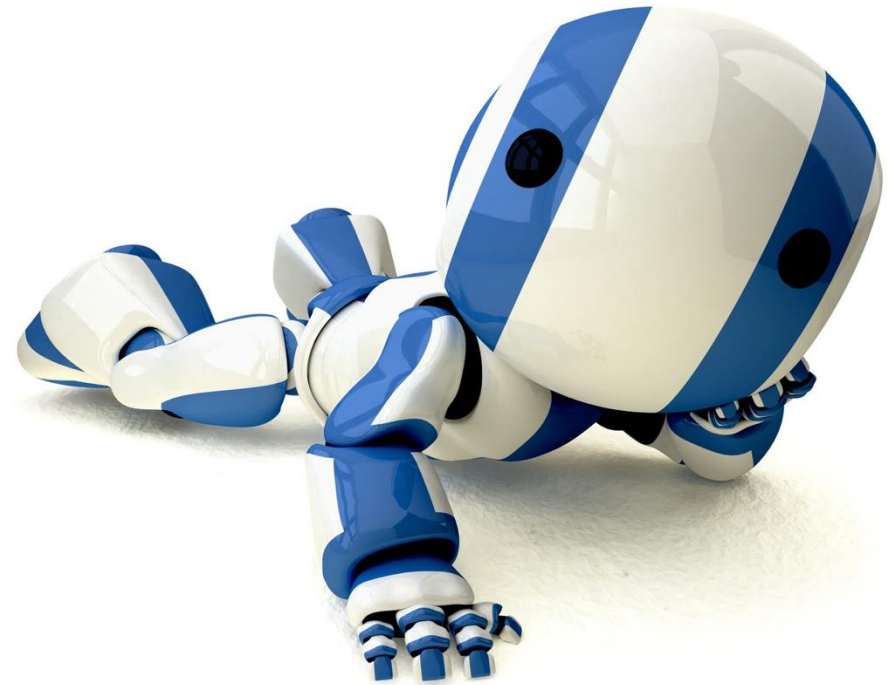


PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
ESCOLA POLITÉCNICA

Aprendizado de Máquina

Conceitos Básicos

Prof. Me. Otávio Parraga



MALTA

Machine Learning Theory
and Applications Lab

Aula de Hoje

- Motivação
- Definições de AM
- Paradigmas de AM
- Projetando um Sistema de Aprendizado
- Informações Úteis
 - Repositórios de Dados
 - Toolboxes
 - Periódicos
 - Conferências

Motivação

- Desde que os computadores foram inventados, é perguntado se eles podem **aprender**

A. M. Turing (1950) Computing Machinery and Intelligence. *Mind* 49: 433-460.

COMPUTING MACHINERY AND INTELLIGENCE

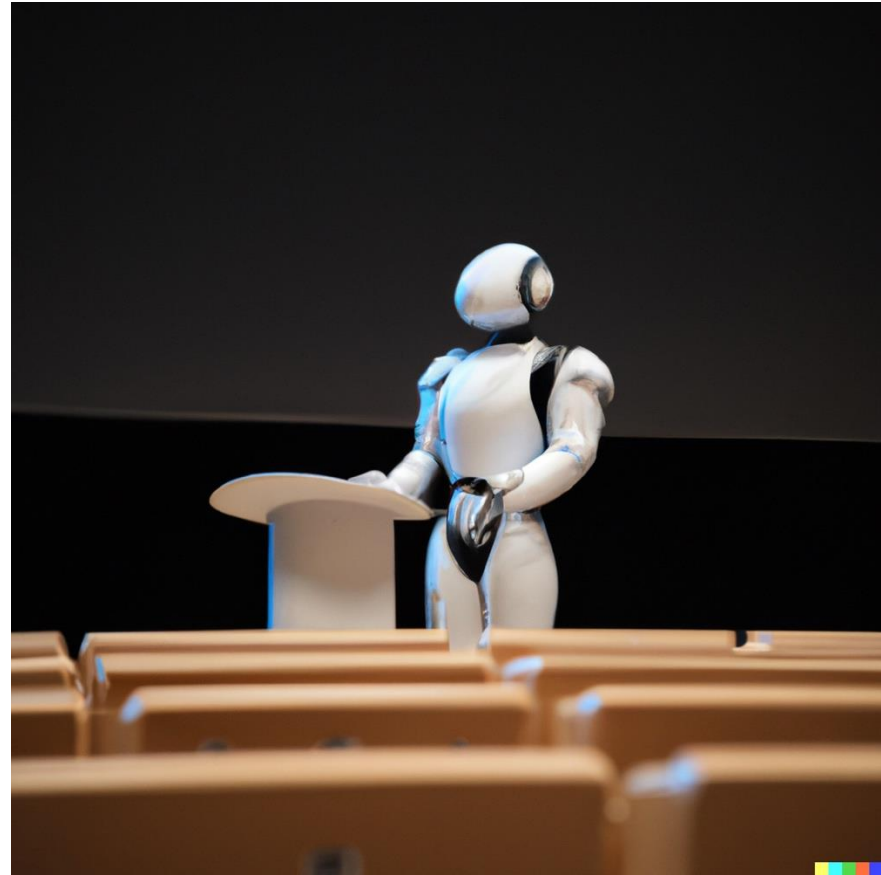
By A. M. Turing

- Se fosse possível entender **como** **programá-los para aprender**, o impacto seria enorme



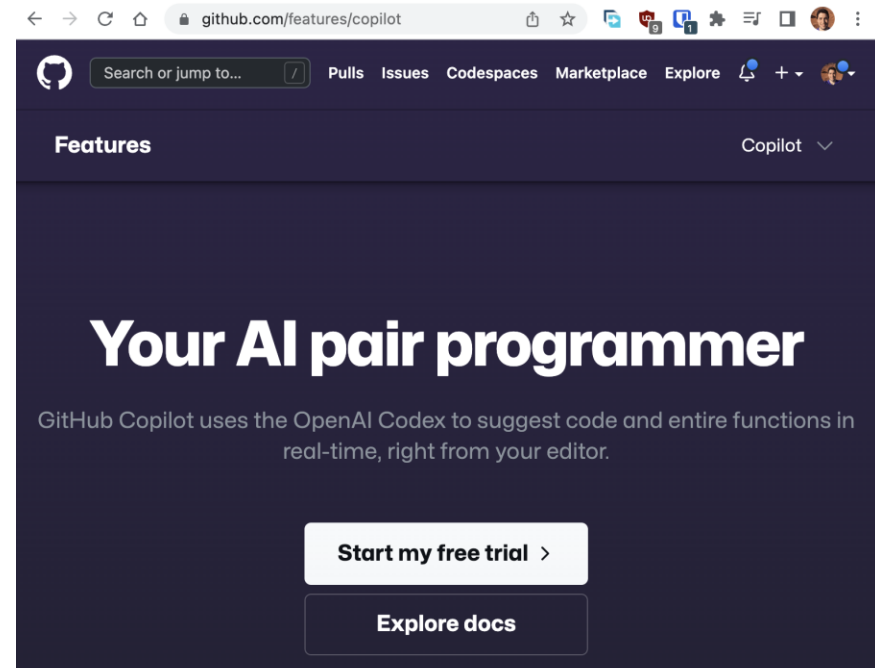
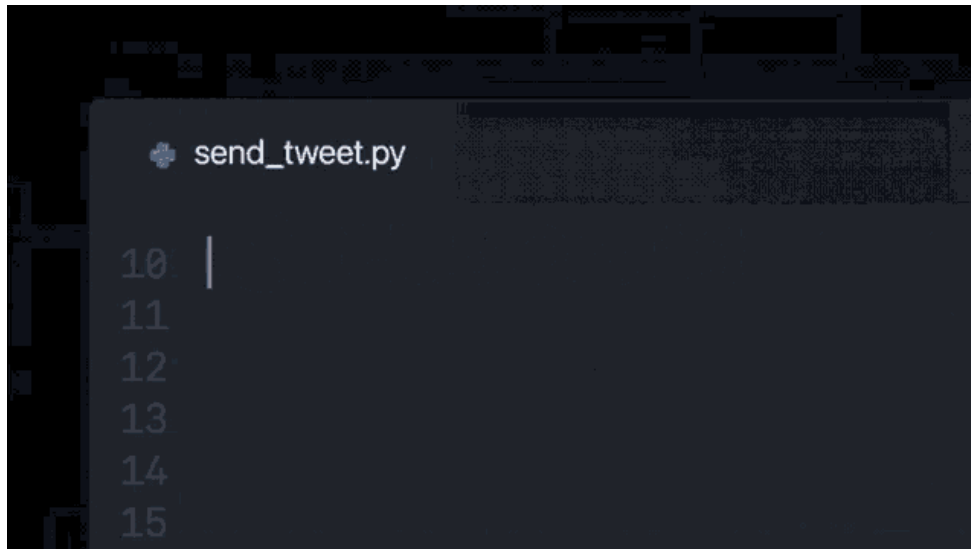
Motivação

- Computadores capazes de aprender poderiam:
 - Ajudar Engenheiros de Software a programar
 - Recomendar bons cursos para aprendermos AM
 - Gerar imagens para melhorar nossos slides de aula
 - Dirigir veículos sem qualquer necessidade de supervisão humana



Motivação

- Ajudar Engenheiros de Software a programar



Motivação

— Recomendar bons cursos para aprendermos AM

ChatGPT



Examples

"Explain quantum computing in simple terms" →

"Got any creative ideas for a 10 year old's birthday?" →

"How do I make an HTTP request in Javascript?" →



Capabilities

Remembers what user said earlier in the conversation

Allows user to provide follow-up corrections

Trained to decline inappropriate requests



Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

RO

Tell me about the best Machine Learning courses available out there



There are many great machine learning courses available online, both paid and free, that cater to individuals with different levels of experience and learning styles. Here are some of the most popular ones:



1. Coursera's Machine Learning by Andrew Ng: This is a highly regarded course that covers the basics of machine learning, including supervised and unsupervised learning, and has been taken by millions of people worldwide.
2. Fast.ai: Fast.ai offers a free, cutting-edge course on deep learning that covers topics such as computer vision and natural language processing. It's designed for individuals who already have a basic understanding of programming and machine learning.
3. Udacity's Intro to Machine Learning: This course is designed for beginners with no prior experience in machine learning, and covers the fundamental concepts and algorithms of the field.
4. MIT's Introduction to Deep Learning: This course provides a comprehensive overview of deep learning, including its history, applications, and technical details.
5. Kaggle Learn: Kaggle is a well-known platform for data science and machine learning competitions. Its Kaggle Learn section offers courses and tutorials on a variety of machine learning topics, including deep learning, computer vision, and time series analysis.

These are just a few of the many great machine learning courses available online. I would recommend you to explore these courses and others to find the one that best fits your



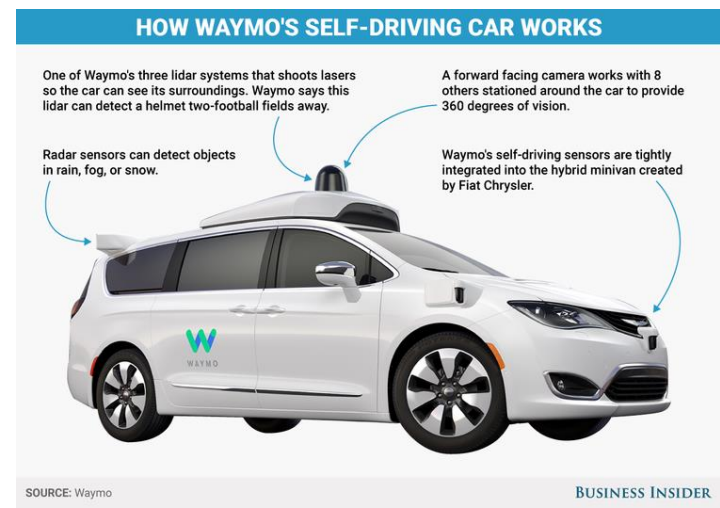
Motivação

- Gerar imagens para melhorar nossos slides de aula



Motivação

- Dirigir veículos sem qualquer necessidade de supervisão humana



Motivação

“Na medida que os computadores se tornam mais sofisticados, parece inevitável que Aprendizado de Máquina exerça um papel central em Ciência da Computação e tecnologia de computadores.”

Tom Mitchell

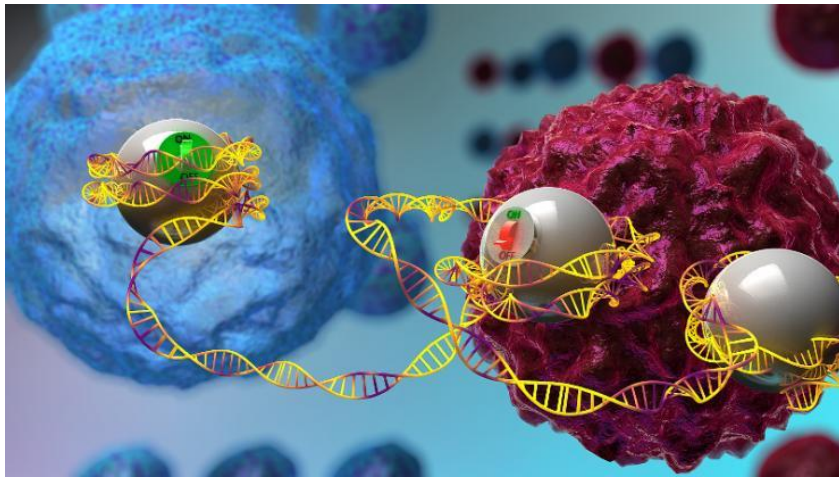
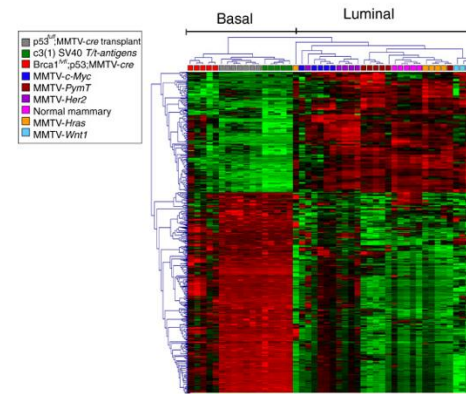
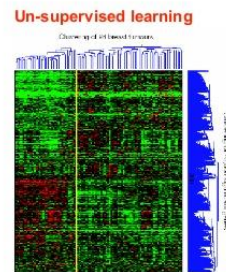
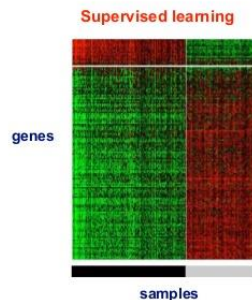
Aplicações de AM

- Algoritmos de AM têm sido bem sucedidos para:

Aplicações de AM

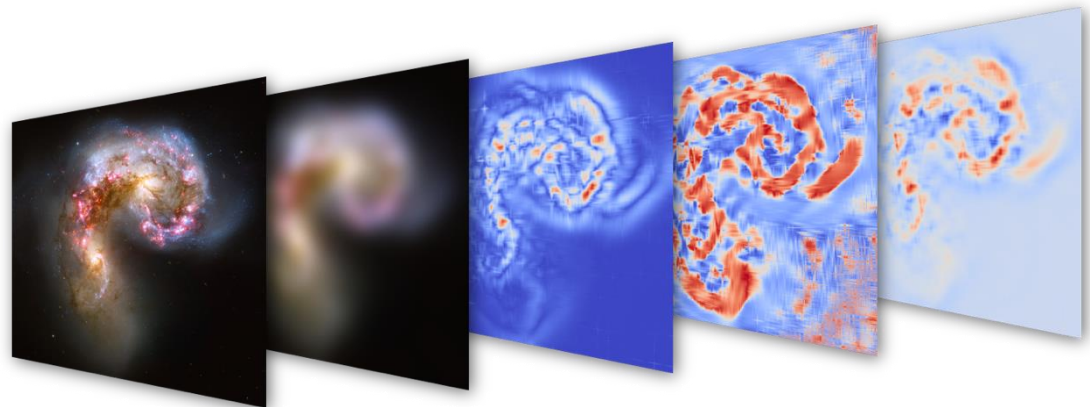
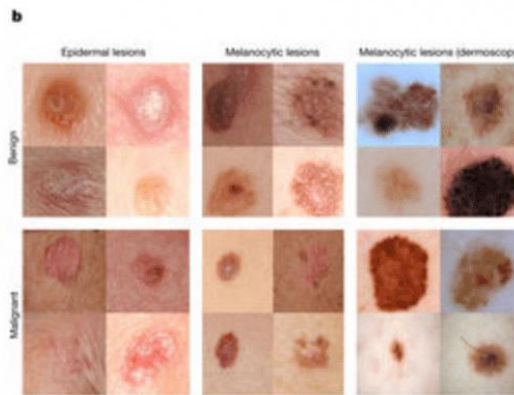
- Algoritmos de AM têm sido bem sucedidos para:
 - Identificar genes associados a determinadas doenças

Analyzing gene expression data



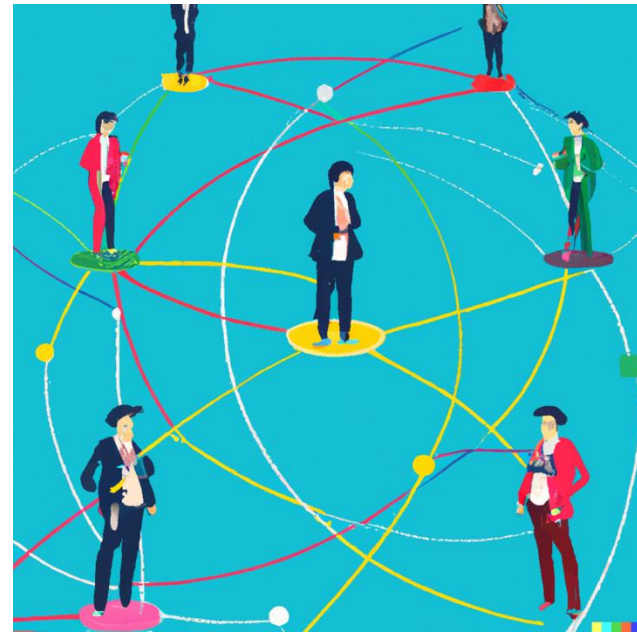
Aplicações de AM

- Algoritmos de AM têm sido bem sucedidos para:
 - Discriminar tecidos (saudáveis/doentes), objetos celestiais, ...



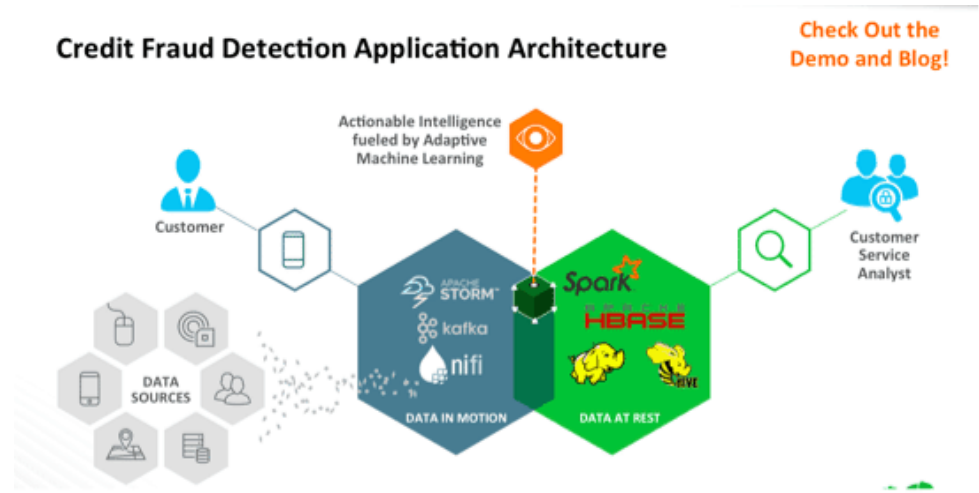
Aplicações de AM

- Algoritmos de AM têm sido bem sucedidos para:
 - Identificar nichos de mercado



Aplicações de AM

- Algoritmos de AM têm sido bem sucedidos para:
 - Detectar uso fraudulento de cartões de crédito



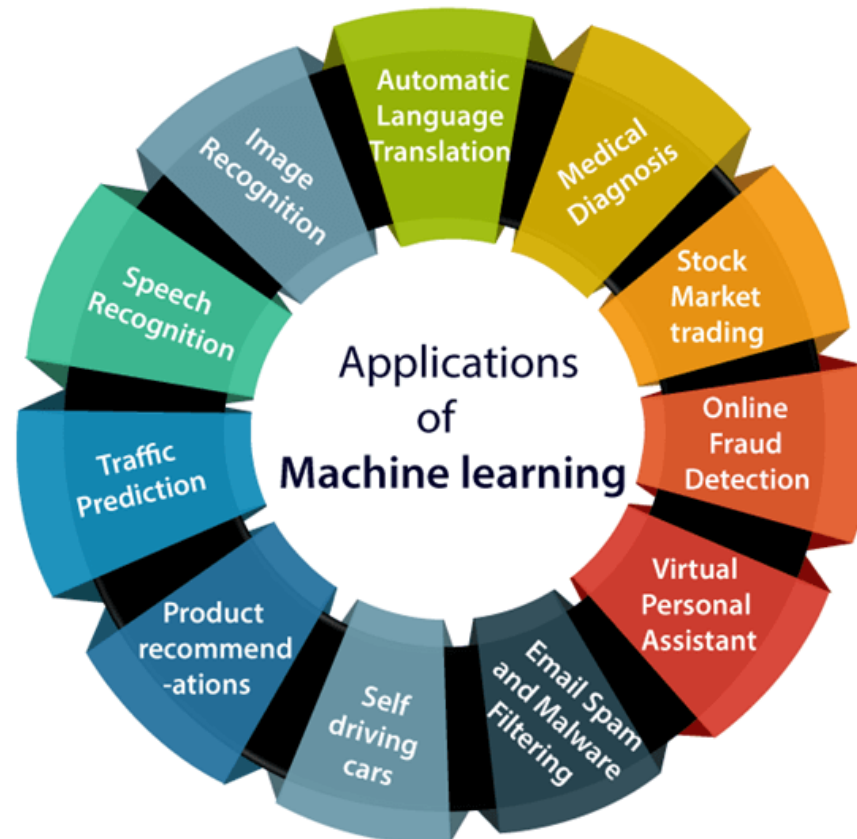
Aplicações de AM

- Algoritmos de AM têm sido bem sucedidos para:
 - Reconhecimento facial, de voz, de assinaturas



Aplicações de AM

- Algoritmos de AM têm sido bem sucedidos para:
 - e tantas (!!!) outras aplicações



Exemplo: ALVINN



Dean Pomerleau (1989)
CMU

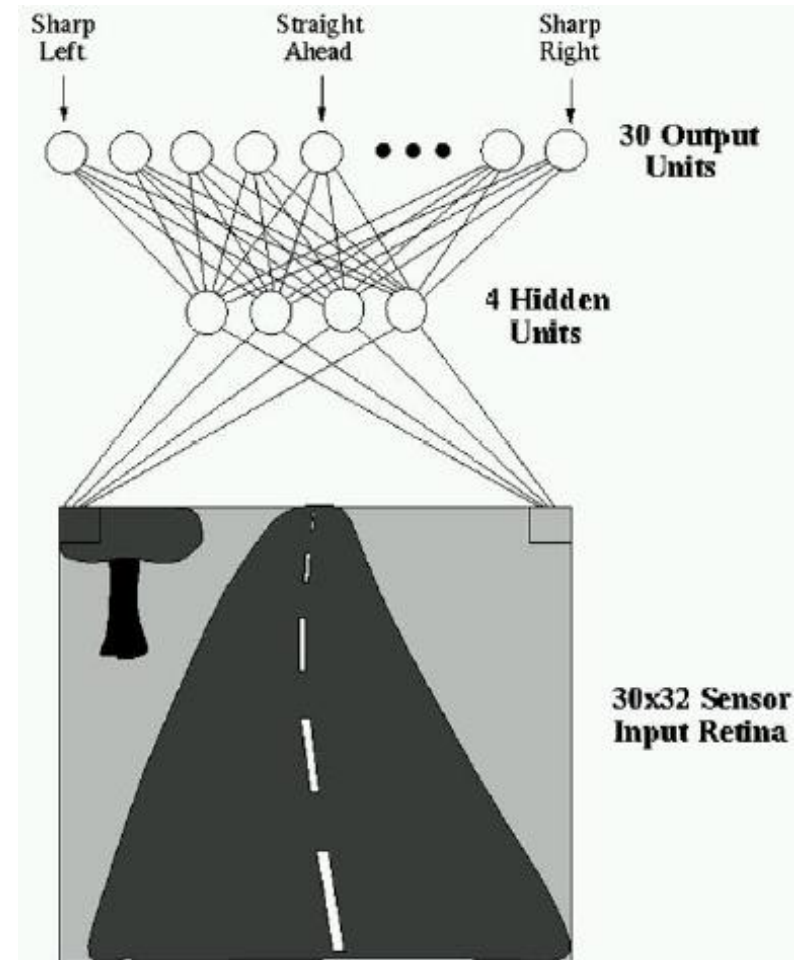
ALVINN

- Sistema automático de navegação para automóveis
 - Baseado em imagens de uma câmara *onboard*
 - Dirigiu a 110km/h em rodovia pública americana
 - De costa a costa em 1989 por 2850 milhas



ALVINN

- Utiliza uma Rede Neural
 - 960 entradas
 - Matrix 30x32 derivada dos pixels da imagem
 - 4 neurônios intermediários
 - 30 neurônios de saída
 - Cada um representando um comando de direção



Tesla *Self-Driving* Car

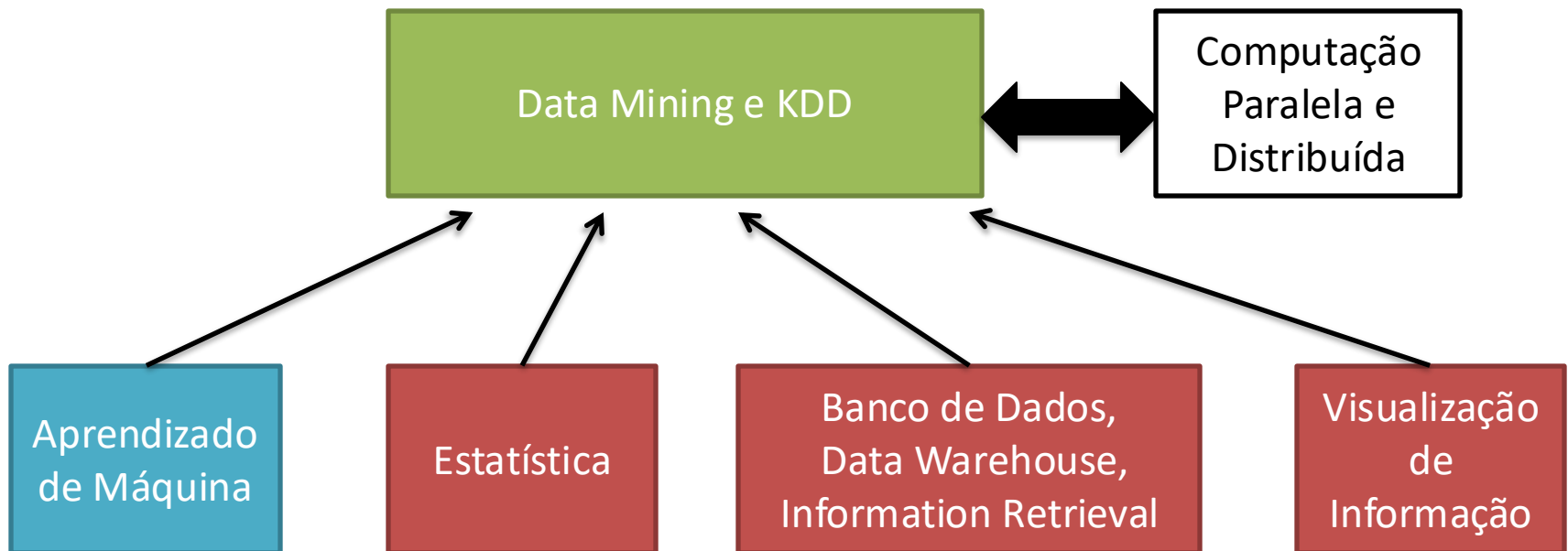


Aplicações de AM

- Uma das áreas de aplicação mais tradicionais de aplicação de AM é descoberta de conhecimento em bases de dados (KDD)
 - Mineração de Dados (*Data Mining*, DM)
 - Exemplos:
 - Registros de compras em grandes supermercados
 - Registros de empréstimos financeiros
 - Registros de transações financeiras
 - Registros médicos
 - Projeto genoma
 - ...

AM, DM e KDD

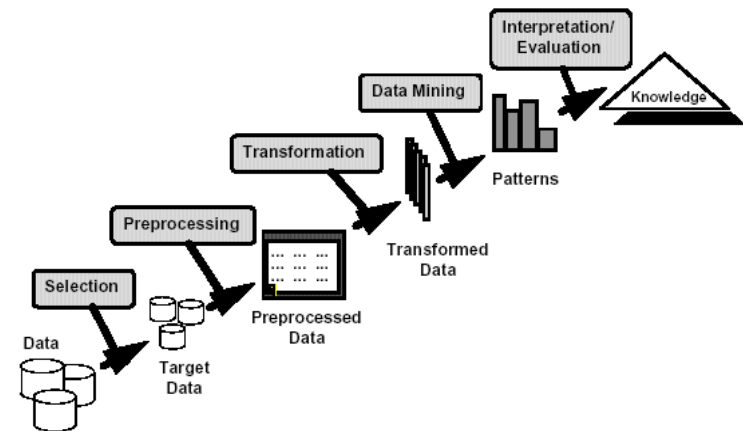
- Os termos *Aprendizado de Máquina* e *Data Mining* são muitas vezes utilizados de maneira indiscriminada, porém se referem a conceitos diferentes



- DM é geralmente feita utilizando AM (mas nem sempre!)

AM, DM e KDD

- Já KDD é usualmente definido como um ciclo que envolve DM
 - Formalmente, KDD é o processo não-trivial de identificar padrões válidos, novos, potencialmente úteis e compreensíveis em dados
 - Advances in Knowledge Discovery and Data Mining, Fayyad, Piatetsky-Shapiro, Smyth, and Uthurusamy, (Chapter 1), MIT Press, 1996
- Mas AM não é um sub-conjunto de DM ou KDD!
 - Reconhecimento de fala, voz, assinaturas, controle, robótica, visão computacional,...



(Tan et al., 2006)

Por que estudar AM?

- ***Esse é o momento!***
 - Vários algoritmos [efetivos](#) e [eficientes](#) estão disponíveis
 - Grande quantidade de **dados disponíveis online**
 - Elevada capacidade dos [recursos computacionais](#) disponíveis

“80% of executives believe that AI (machine learning) will revolutionize the way companies will proceed to acquire information from their customers”

http://www.insuranceup.it/en/scenarios/artificial-intelligence-and-machine-learning-40-impact-on-insurance-companies_1854.htm

“AI (Machine Learning) Will Add \$15.7 Trillion to the Global Economy”

<https://www.bloomberg.com/news/articles/2017-06-28/ai-seen-adding-15-7-trillion-as-game-changer-for-global-economy>

Por que estudar AM?

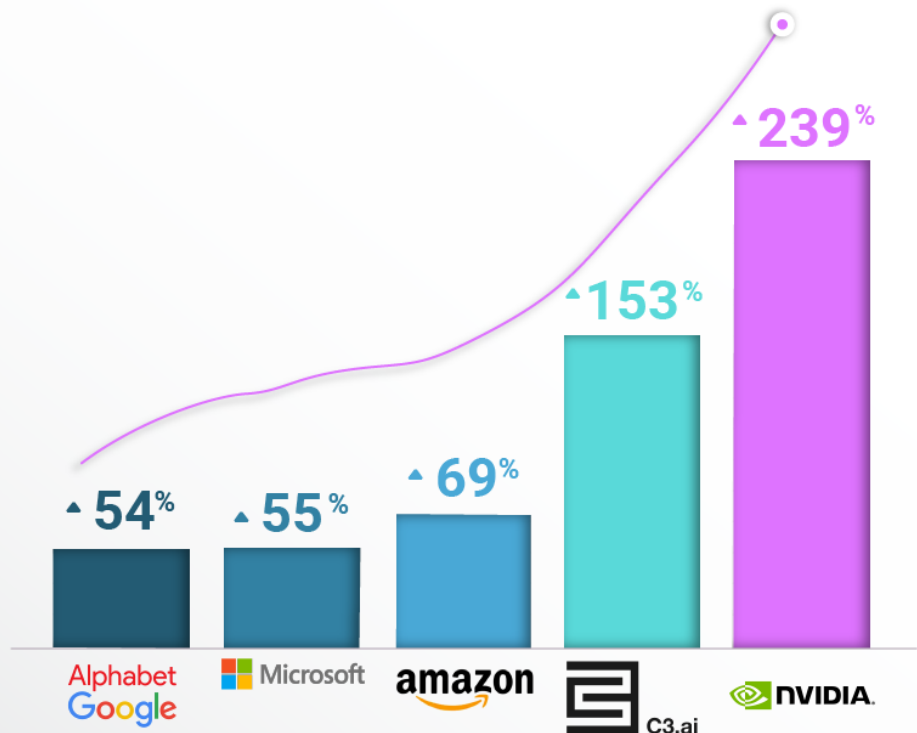
A year after ChatGPT's release, the AI revolution is just beginning

By Catherine Thorbecke, CNN

🕒 8 minute read · Updated 10:32 AM EST, Thu November 30, 2023

AI's 2023 Stock Price Impact

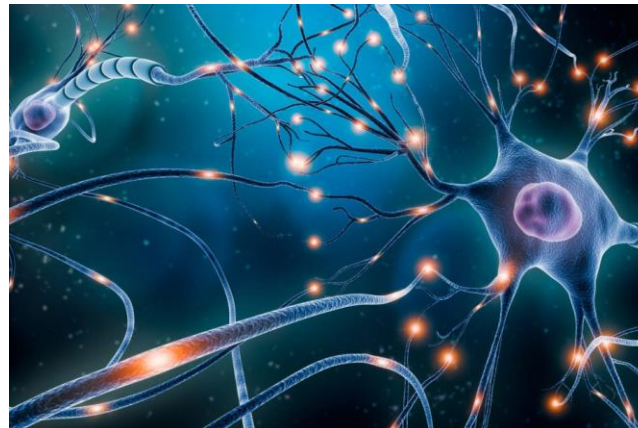
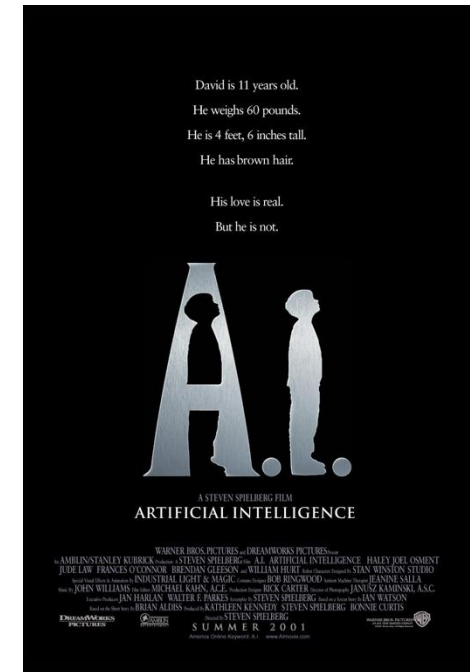
One year return; price change through January, 2024



<https://www.linkedin.com/pulse/emerging-ai-january-february-2024-roundup-sabic/>

Disciplinas Relacionadas

- Inteligência Artificial
- Mineração de Dados
- Teoria da Informação
- Estatística Bayesiana e Frequentista
- Otimização
- Estimação, Filtragem e Controle
- Teoria da Complexidade Computacional
- Neurociências
- ...



Definições

- Definições relacionadas a aprendizado e AM
 - Uma máquina de aprendizado, definida de maneira ampla, é qualquer dispositivo cujas ações são influenciadas por **experiências anteriores** (Nilsson 1965)
 - Qualquer mudança em um sistema que o permite ter um melhor desempenho na segunda vez em que ele **repete uma mesma tarefa** (Simon 1983)
 - Modificação de uma tendência comportamental **por meio de experiência** (Webster 1984)
 - Uma melhoria na capacidade de processar informação a partir da atividade de processar informação (Tanimoto 1990)

Aprendizado de Máquina

- Definição clássica (Mitchell, 1997)

“Um programa de computador é dito **aprender** a partir de uma experiência **E** com respeito a alguma classe de tarefas **T** e medida de desempenho **P**, se seu desempenho em tarefas de **T**, medido por **P**, melhora com a experiência **E**”



Exemplo 1

- **Problema de Aprender a Jogar Xadrez**
 - **Tarefa T:** jogar xadrez
 - **Medida de desempenho P:** percentagem de jogos vencidos contra adversários
 - **Experiência de treinamento E:** praticar jogando contra si próprio ou contra adversários humanos (p.ex., pela internet)

Exemplo 2

- **Problema de Filtrar SPAMs**
 - **Tarefa T:** categorizar mensagens de e-mail como spam ou legítima
 - **Medida de desempenho P:** percentagem de mensagens corretamente classificadas
 - **Experiência de treinamento E:** conjunto de e-mails manualmente rotulados por seres humanos



Exercício 1

- Problema de reconhecimento de escrita manual
 - Tarefa T:
 - Medida de desempenho P:
 - Experiência de treinamento E:

Exercício 2

- Problema do Veículo Autônomo
 - Tarefa T:
 - Medida de desempenho P:
 - Experiência de treinamento E:

Exercício 3

- Problema de Diagnóstico Médico
 - Tarefa T:
 - Medida de desempenho P:
 - Experiência de treinamento E:

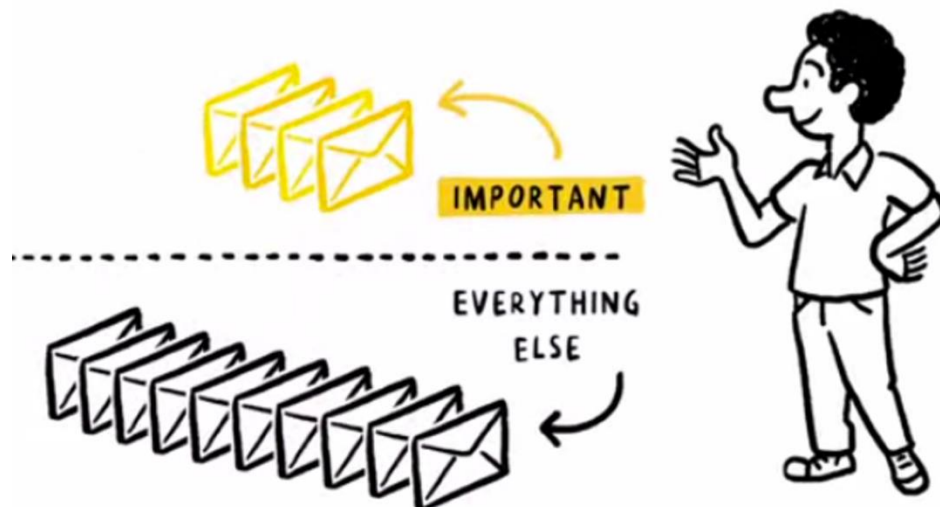
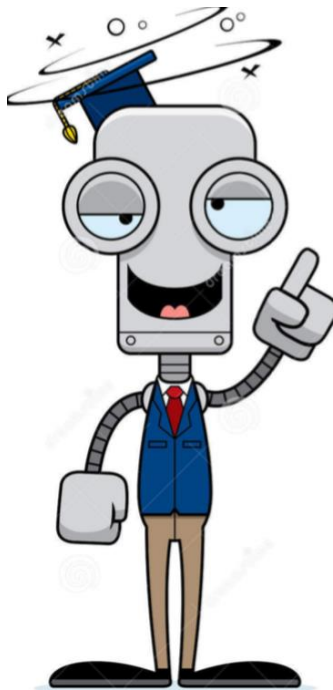
Paradigmas de AM

- O treinamento de um sistema de aprendizado pode ser:
 - Supervisionado
 - Semi-supervisionado
 - Auto-supervisionado
 - Não supervisionado
 - Reforço

Aprendizado Supervisionado

Guiado por professor externo

- Professor possui conhecimento sobre a tarefa
- Representado por conjuntos de pares (\mathbf{x}, y)
- Algoritmo de AM gera modelo que busca **reproduzir comportamento do professor**
- Parâmetros do modelo são ajustados por apresentações sucessivas dos pares (\mathbf{x}, y) : **fase de treinamento**
 - Após o treinamento, o desempenho do sistema deve ser testado com dados não-vistos: **fase de teste**



Aprendizado Supervisionado

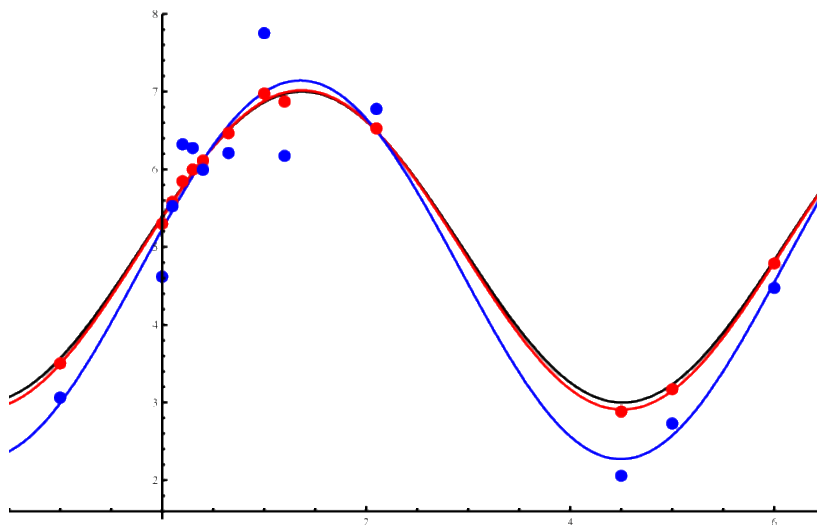
- Exemplos de tarefas supervisionadas

- Classificação de padrões

- Categorizar objetos

- Regressão

- Previsão de valores contínuos



Aprendizado por Reforço

- Guiado por um crítico externo

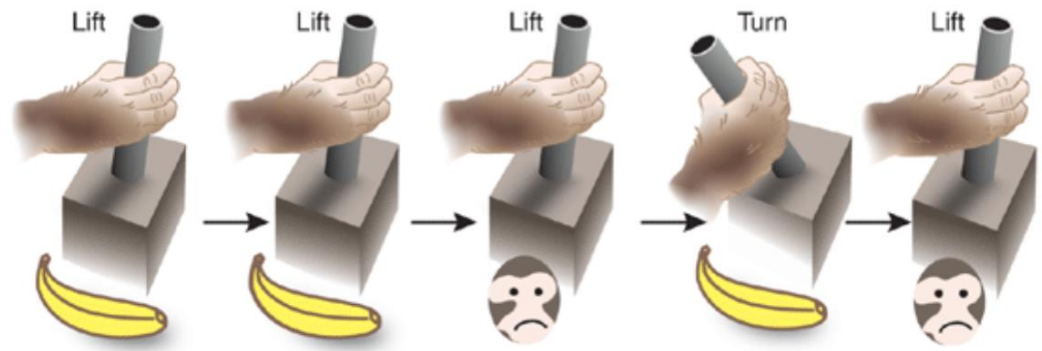
- Processo de tentativa e erro
- Procura maximizar sinal de reforço

- Se ação tomada por sistema é seguida por estado satisfatório, o sistema é fortalecido. Caso contrário, o sistema é enfraquecido (Lei de Thorndike)



- Tipos de reforço

- Positivo = recompensa
- Negativo = punição
- Nulo

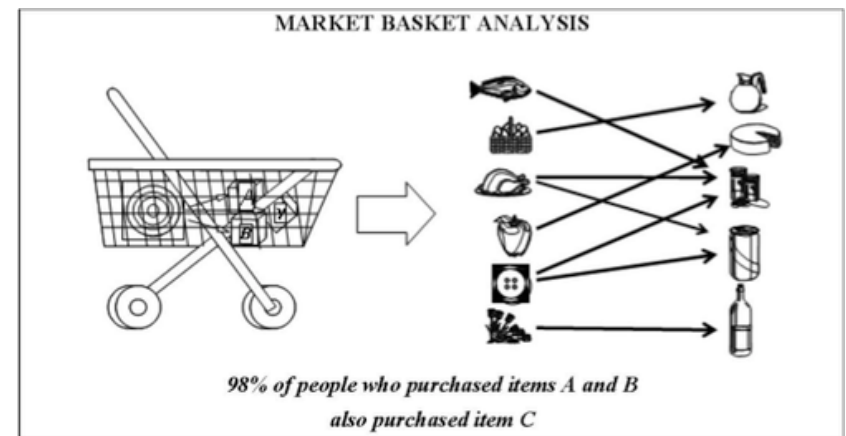
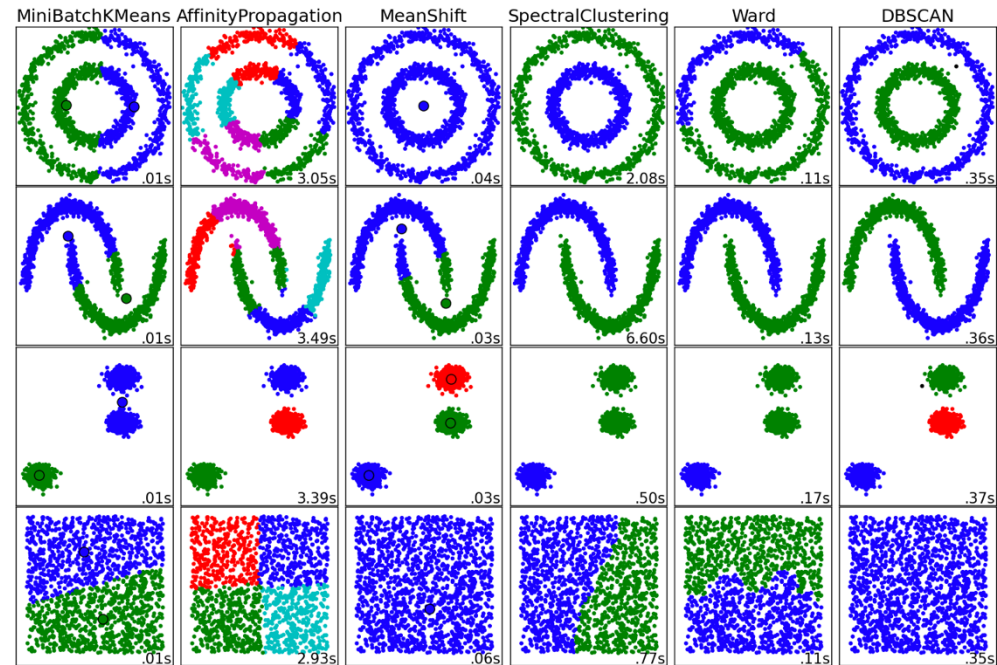
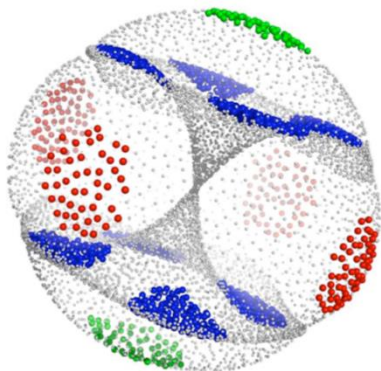


Supervisionado vs. Reforço

Treinamento Supervisionado	Treinamento por Reforço
Professor	Crítico
Sistema de Feedback	Sistema de Feedback
É dito o que fazer	Faz e vê o que acontece
Mais rápido	Mais lento
Supervisão é cara	Reforço é barato

Aprendizado Não Supervisionado

- Não há crítico ou professor externo
- Exemplos:
 - Clustering: descobre categorias automaticamente
 - Associação: descobre relacionamentos entre variáveis
 - Redução de dimensionalidade

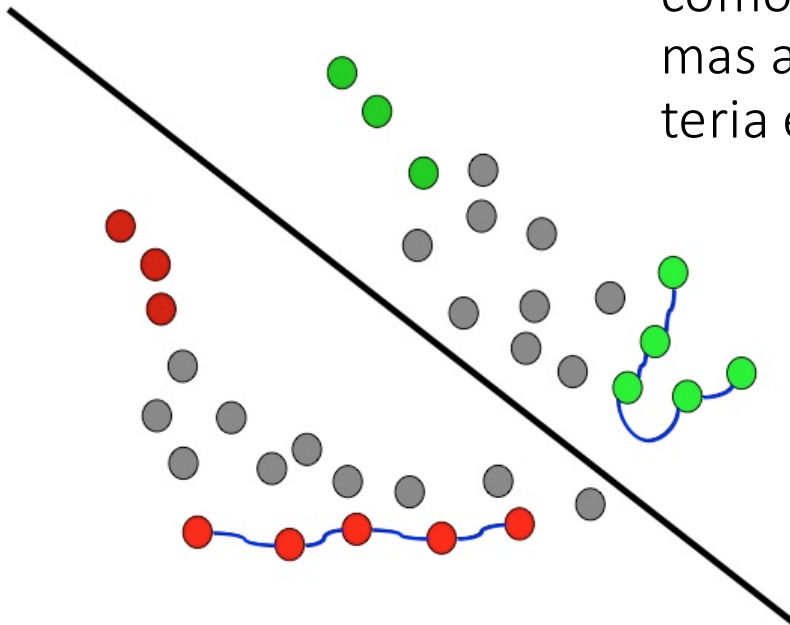


Aprendizado Semi-Supervisionado

- Professor externo apenas para parte dos exemplos de treinamento

– Exemplo:

- *Web mining*: usuários podem rotular páginas como pertencentes a determinadas categorias, mas apenas uma parcela ínfima de webpages teria essa informação associada



Sup. vs. Não Sup.

Treinamento Supervisionado	Treinamento Não-Supervisionado
Previsão	Descoberta
Dados e rótulos	Dados
É dito o que fazer	Faz e vê o que acontece
Fácil avaliação	Difícil avaliação

Aprendizado Auto-Supervisionado

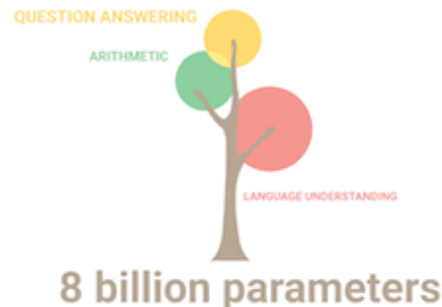
- Aprendizado supervisionado com rótulos “gratuitos”
 - Exemplos:
 - Language modeling: geração de textos onde o rótulo (supervisão) é parte do texto já existente
 - Colorir imagens: o rótulo é a cor original da imagem
 - Reconstruir imagens: o rótulo é a parte da imagem que foi escondida/danificada durante o treinamento

The cat

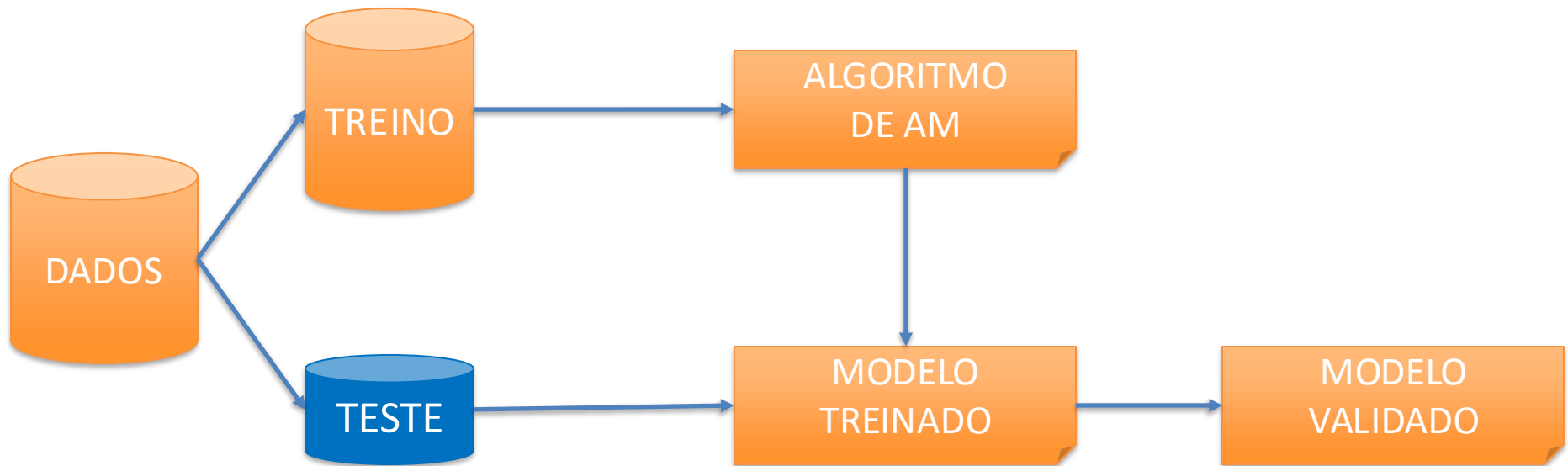


Aprendizado Auto-Supervisionado

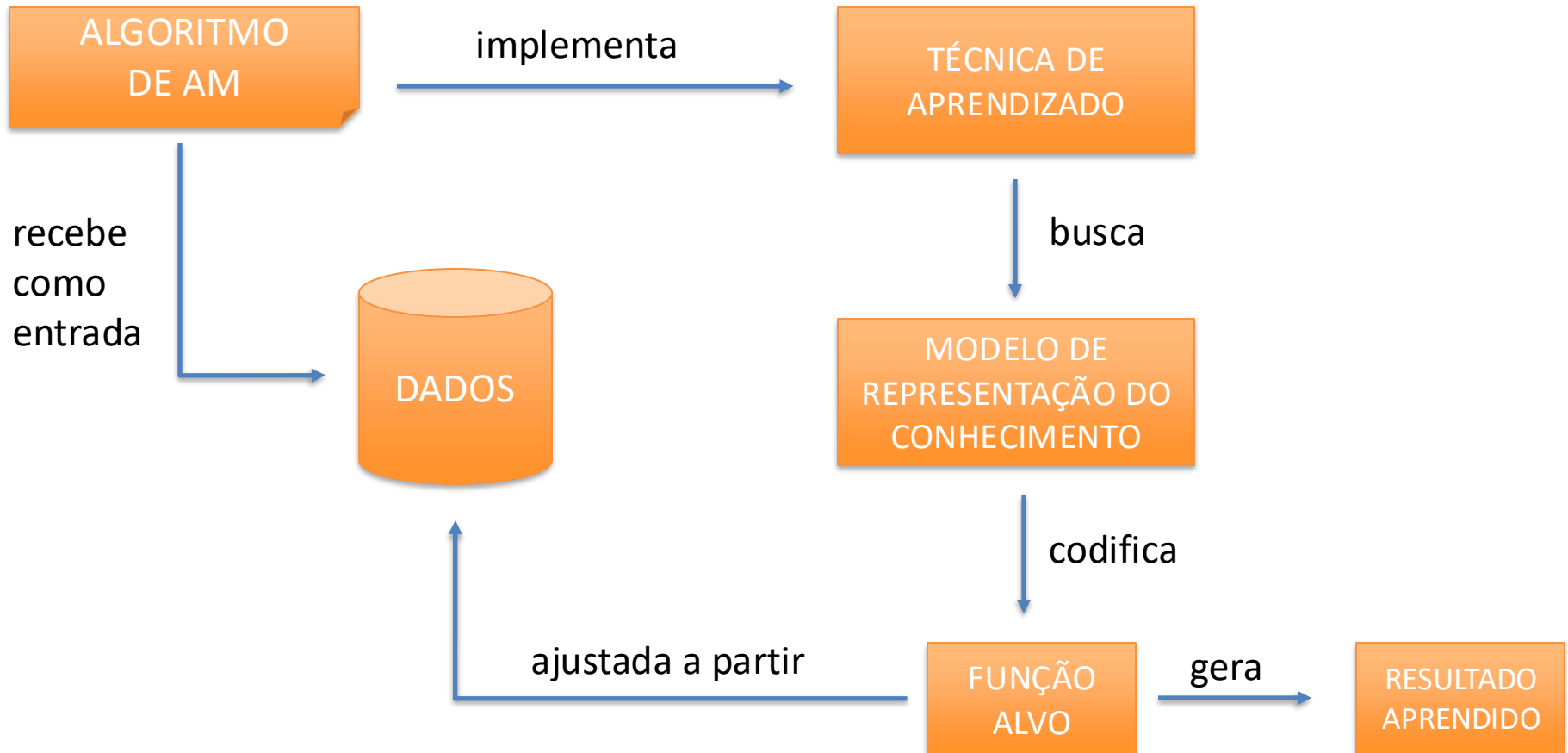
- A maioria dos modelos geradores são treinados com Auto-Supervisão
 - GPT-Like
 - Stable Diffusion e Midjourney
- Fácil obter mais dados



Fluxo de um Sistema de Aprendizado



Projetando um Sistema de Aprendizado



Projetando um Sistema de Aprendizado

TÉCNICA DE APRENDIZADO

Como procurar pela solução?

busca

MODELO DE REPRESENTAÇÃO DO CONHECIMENTO

Como representar/gravar a solução encontrada?

codifica

FUNÇÃO ALVO

Representação matemática do foco que temos!

Modelo de Representação do Conhecimento

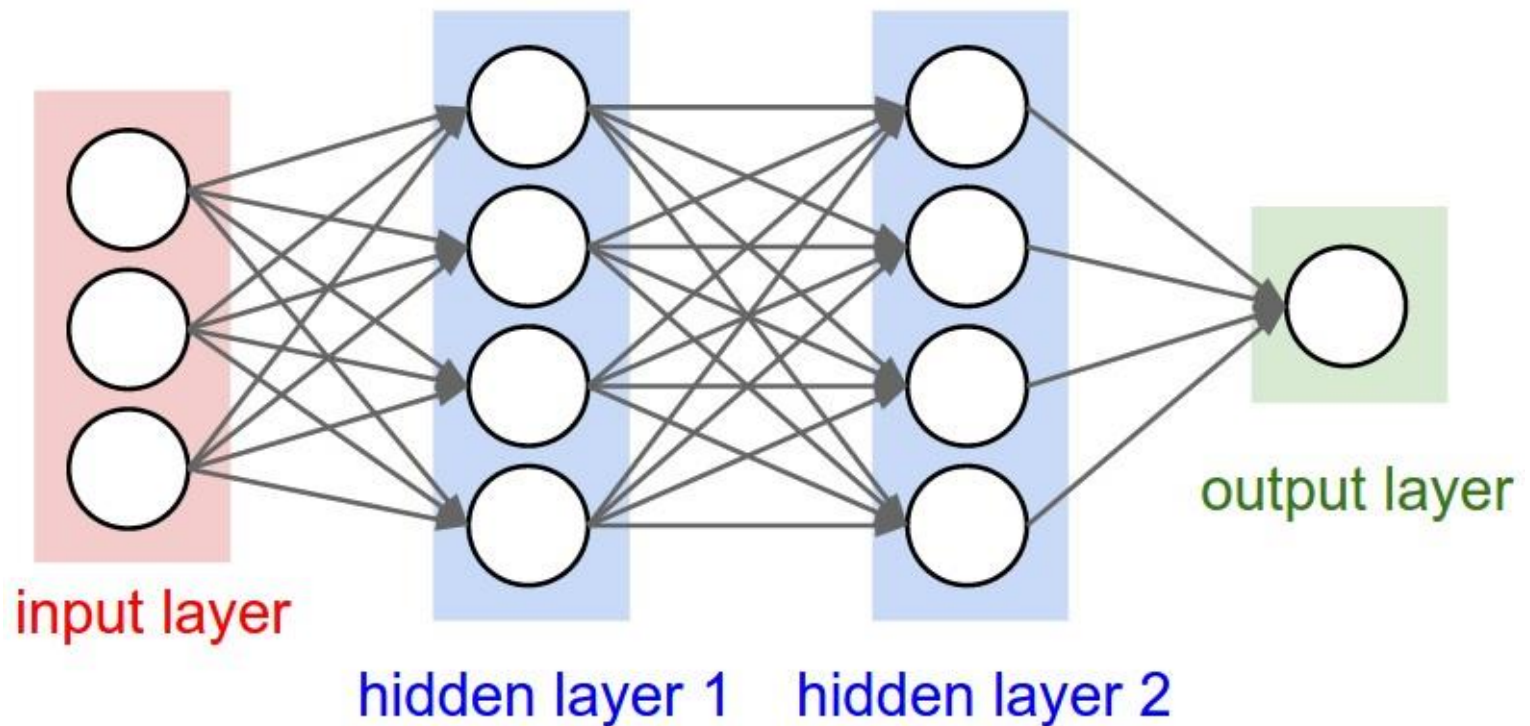
- Modelos Matemáticos
 - Regressão linear / logística
 - Redes Neurais
 - Máquinas de Vetores de Suporte, ...
- Modelos Simbólicos
 - Árvores de Decisão
 - Regras em lógica proposicional ou de 1ª ordem
 - Redes Semânticas, ...

Modelo de Representação do Conhecimento

- Modelos “*Lazy*”
 - k -NN
 - Raciocínio Baseado em Casos (CBR),
 - ...
- Modelos Probabilísticos
 - Naïve Bayes
 - Redes Bayesianas
 - Misturas de Gaussianas
 - Modelos Ocultos de Markov (HMMs),
 - ...

Exemplos de Modelos

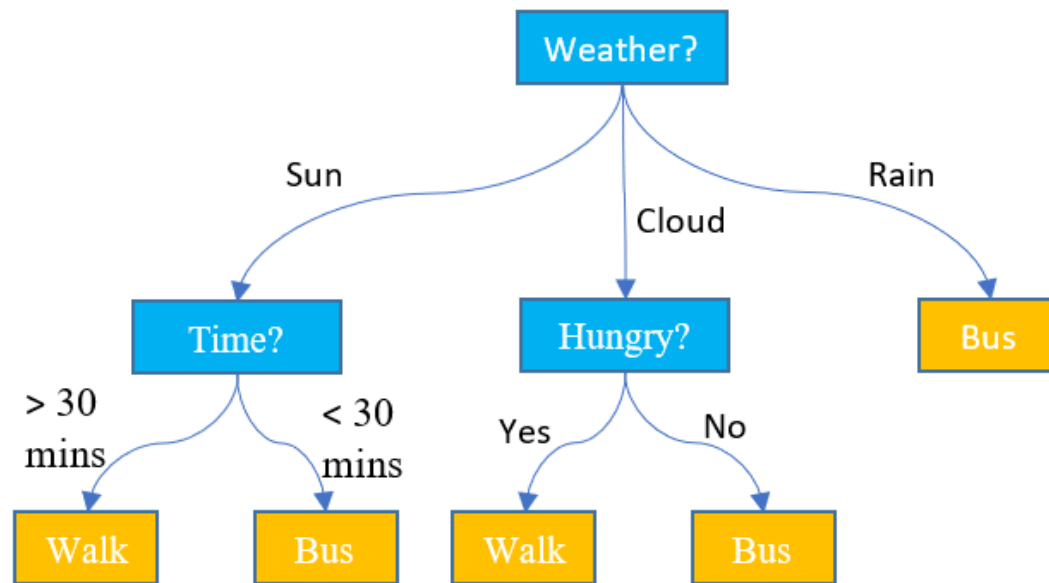
Rede Neural



$$h = o(((xw^1 + b^1)w^2 + b^2)w^3 + b^3)$$

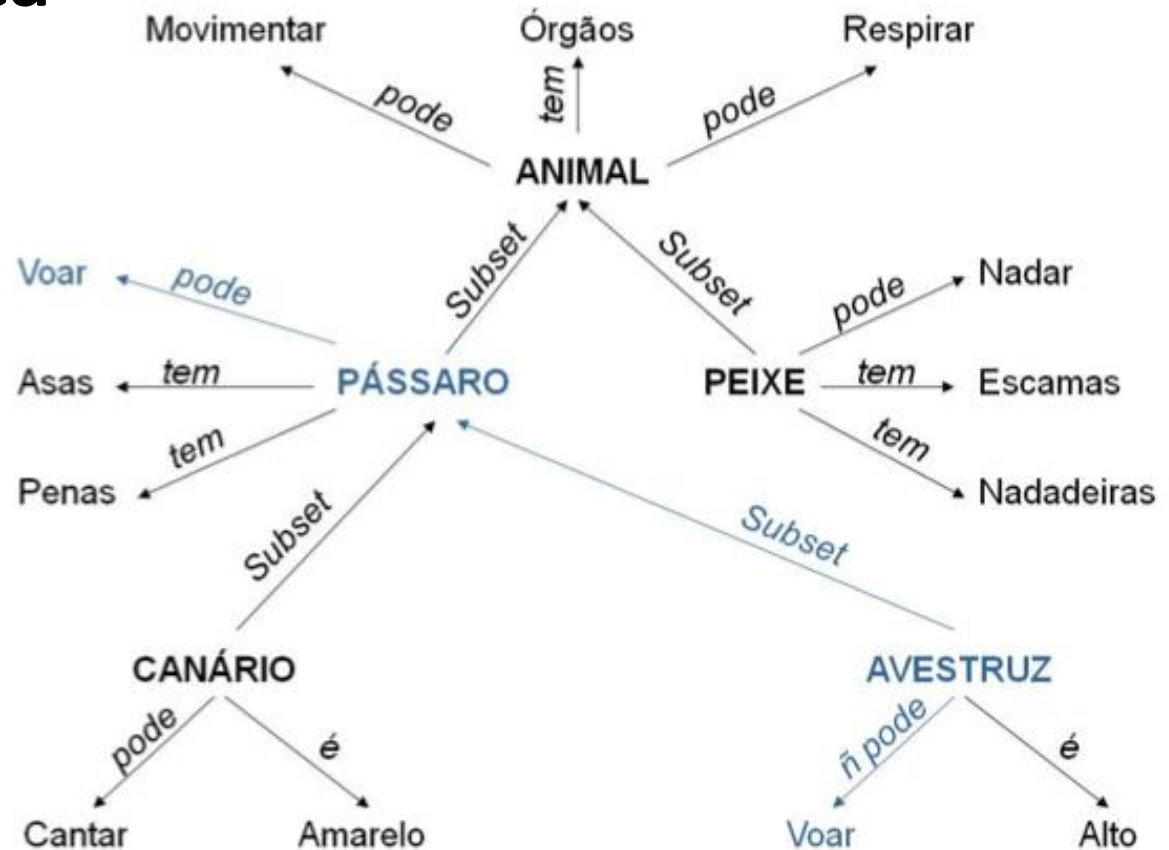
Exemplos de Modelos

Árvore de Decisão



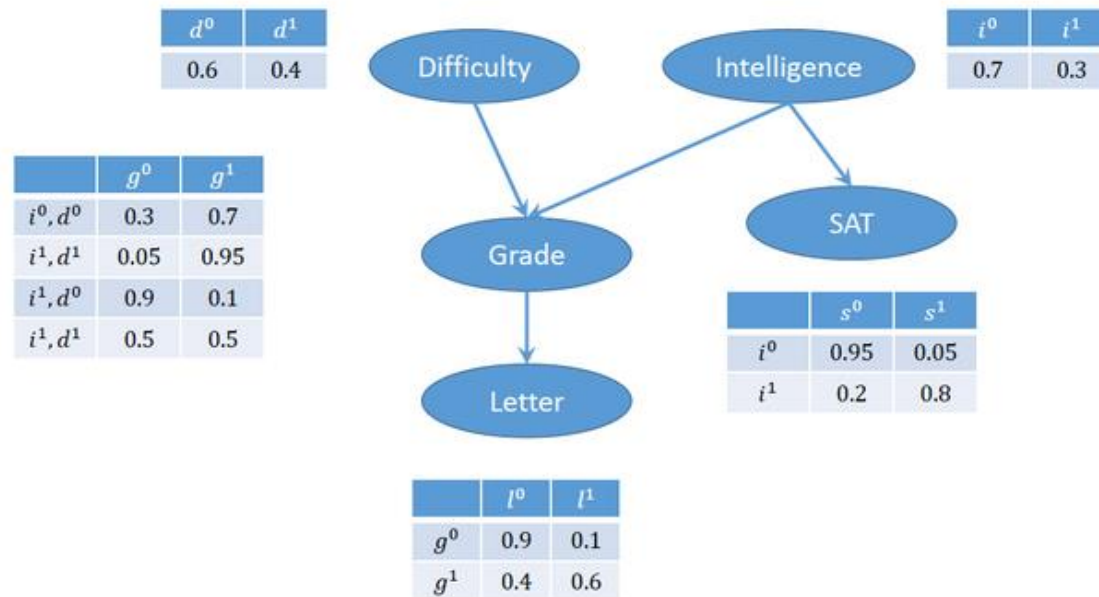
Exemplos de Modelos

Rede Semântica



Exemplos de Modelos

Rede Bayesiana



Técnicas de Aprendizado

- Dado um tipo de modelo, uma função alvo e um conjunto de objetos de treinamento, é preciso algum mecanismo para obter um modelo específico que represente bem a função alvo
 - Esse mecanismo consiste basicamente de uma técnica de busca
 - Busca-se, no espaço de modelos plausíveis de um determinado tipo, aquele que melhor represente a função alvo



Técnicas de Aprendizado

- Algoritmos Baseados em Otimização via Gradiente Descendente
 - Regressão linear/logística, redes neurais, ...
- Algoritmos baseados em Programação Dinâmica
 - HMMs, ...
- Algoritmos baseados em Divisão e Conquista
 - Indução de árvores e regras de decisão
- Algoritmos baseados em Probabilidades
 - Naïve Bayes, Redes Bayesianas, ...
- Algoritmos baseados em Computação Evolutiva
 - Aplicáveis a vários modelos
- ...

Modelos, Técnicas e *Bias Indutivo*

- Cada tipo de modelo é mais apropriado para determinada classe de problemas (*no free-lunch*)
 - É parte importante do estudo de AM aprender a identificar os cenários mais apropriados para cada modelo e técnica de aprendizado
- O modelo e a técnica estabelecem algo **fundamental** em Aprendizado de Máquina
 - *Bias Indutivo*

Bias Indutivo

- Informalmente, o *bias indutivo* de um sistema de AM é uma tendência a **privilegiar** um dado conjunto de **hipóteses** em detrimento a outras
 - Assuma que “hipótese” neste caso se refere a uma realização (ou instanciiação) particular de um modelo

Informações Úteis

- Repositório de Dados na Web
 - UCI data repository
 - <http://archive.ics.uci.edu/ml/>
 - Kaggle
 - <https://www.kaggle.com/>
- Cursos
 - Machine Learning (Coursera – Andrew Ng)
 - <https://www.coursera.org/learn/machine-learning>
 - Introduction to Machine Learning (Udacity)
 - <https://br.udacity.com/course/intro-to-machine-learning--ud120/>

Informações Úteis

- Weka
 - <http://www.cs.waikato.ac.nz/ml/weka/>
- Scikit Learn (Python)
 - <http://scikit-learn.org/>
- Tensorflow
 - <https://www.tensorflow.org/>
- Pytorch
 - <https://pytorch.org/>
- Matlab Toolbox for Pattern Recognition
 - <http://www.prtools.org>
- R
 - <http://cran.r-project.org/>

Informações Úteis

- Tom Mitchell's webpage
 - <http://www.cs.cmu.edu/~tom/>
- Kdnuggets
 - <http://www.kdnuggets.com/>
- Perfis a seguir no Twitter
 - Gregory Piatetsky
 - @kdnuggets
 - Nando de Freitas
 - @NandoDF
 - Yann LeCun
 - @ylecun
 - François Chollet
 - @fchollet

Informações Úteis

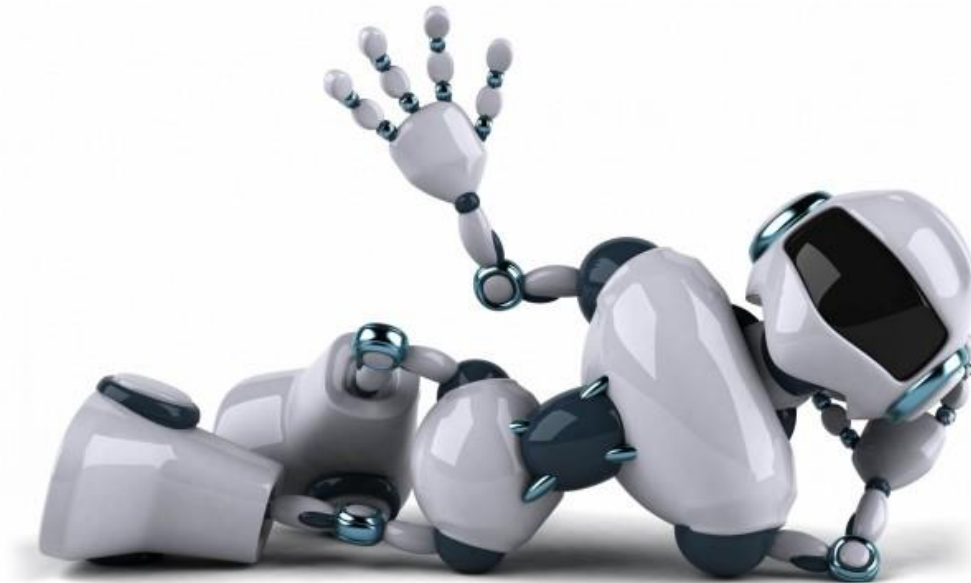
- Journals
 - Journal of Machine Learning Research
 - Transactions on Machine Learning Research
 - Machine Learning (Springer)
 - IEEE Transactions on Pattern Analysis and Machine Intelligence
 - Information Sciences (Elsevier)
 - IEEE Transactions on Neural Networks and Learning Systems
 - Data Mining and Knowledge Discovery (Springer)
 - Pattern Recognition (Elsevier)
 - Neurocomputing (Elsevier)

Informações Úteis

- Conferences
 - ICML (International Conference on Machine Learning)
 - NeurIPS (Neural Information Processing Systems)
 - ICLR (International Conference on Learning Representations)
 - ECML / PKDD (European Conference on Machine Learning)
 - CVPR (Computer Vision and Pattern Recognition)
 - ICDM (International Conference on Data Mining)
 - IJCNN (International Joint Conference on Neural Networks)
 - ICPR (International Conference on Pattern Recognition)
 - ICANN (International Conference on Artificial Neural Networks)
 - BRACIS (Brazilian Conference on Intelligent Systems)

Leituras Recomendadas

- Capítulo 1 (Mitchell 1997)
- Capítulo 1 (Tan et al. 2006)



Créditos e Referências

- Slides adaptados dos originais gentilmente cedidos pelos Profs. André Carvalho e Ricardo Campello, do ICMC-USP, e Rodrigo Coelho Barros (PUCRS)
- TAN, P. N. STEINBACH, M. KUMAR, V. **Introduction to Data Mining**. Addison-Wesley, 2005. 769 p.
- Mitchell, T. **Machine Learning**. McGraw-Hill, 1997. 432 p.