

Curso de extensão em Inteligência Artificial



Agradecimentos

Pró-Reitoria de Pesquisa e Inovação da USP

Pró-Reitoria de Cultura e Extensão Universitária da USP

Projeto Synestech.AI: InovaUSP, Softex, MCTI e Motorola

Centro de Inteligência Artificial da USP (C4AI): FAPESP e IBM

Instituto de Ciências Matemáticas e de Computação
(ICMC-USP/São Carlos)

Quem somos

Professores

- Thiago Alexandre Salgueiro Pardo (ICMC-USP)
- Nina Sumiko Tomita Hirata (IME-USP)
- Roberto Hirata Junior (IME-USP)
- Roberto Marcondes Cesar Junior (IME-USP)



Quem somos

Monitores

- Alexandre Benatti
- Artur André Almeida de Macedo Oliveira
- Germano Antônio Zani Jorge
- Gustavo Sampaio Lima
- Pietro Guarinello Cariola
- Rafael Jeferson Pezzuto Damaceno

Nosso propósito com esse curso

Compartilhar com a sociedade um pouco do conhecimento acumulado na USP na área de IA

- Melhor entendimento do tópico
- Engajamento mais informado nas discussões atuais
- Início do caminho para cada um poder seguir seu trajeto na área
- Apoio para decisões de carreira (estudos, empregos, etc.)

Momento especial

90 anos da USP

- Grandes iniciativas e projetos de IA

IA no centro das atenções no mundo

1934



No dia 25 de janeiro, o interventor federal do Estado de São Paulo, Armando de Salles Oliveira, assina o decreto nº 6.283 que cria a **Universidade de São Paulo (USP)**

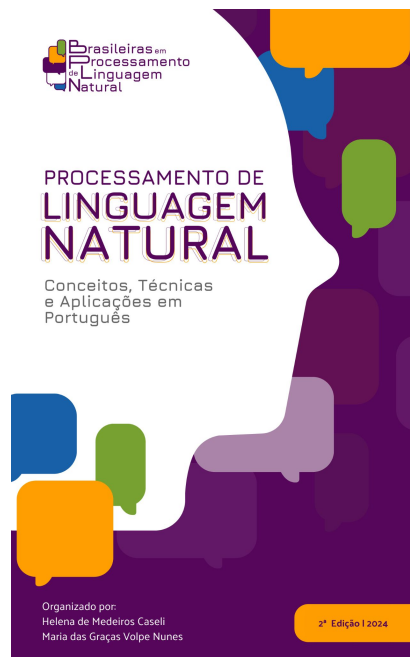
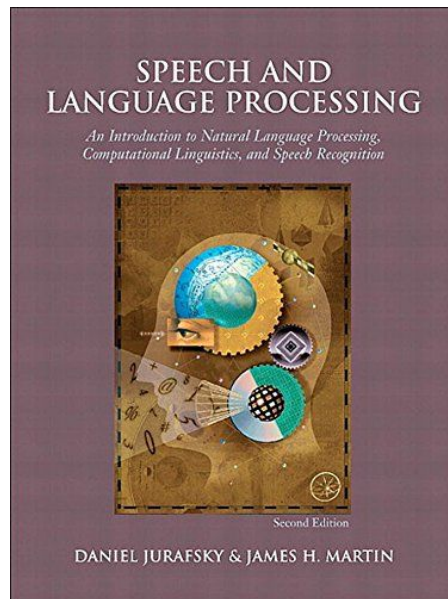
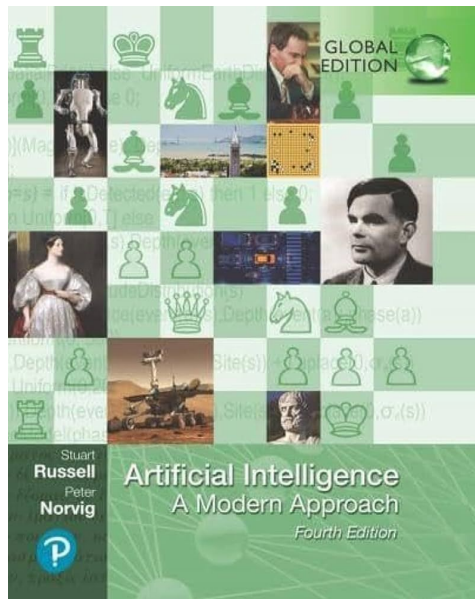
<https://jornal.usp.br/universidade/usp-90-anos-nossa-historia-em-uma-linha-do-tempo/>

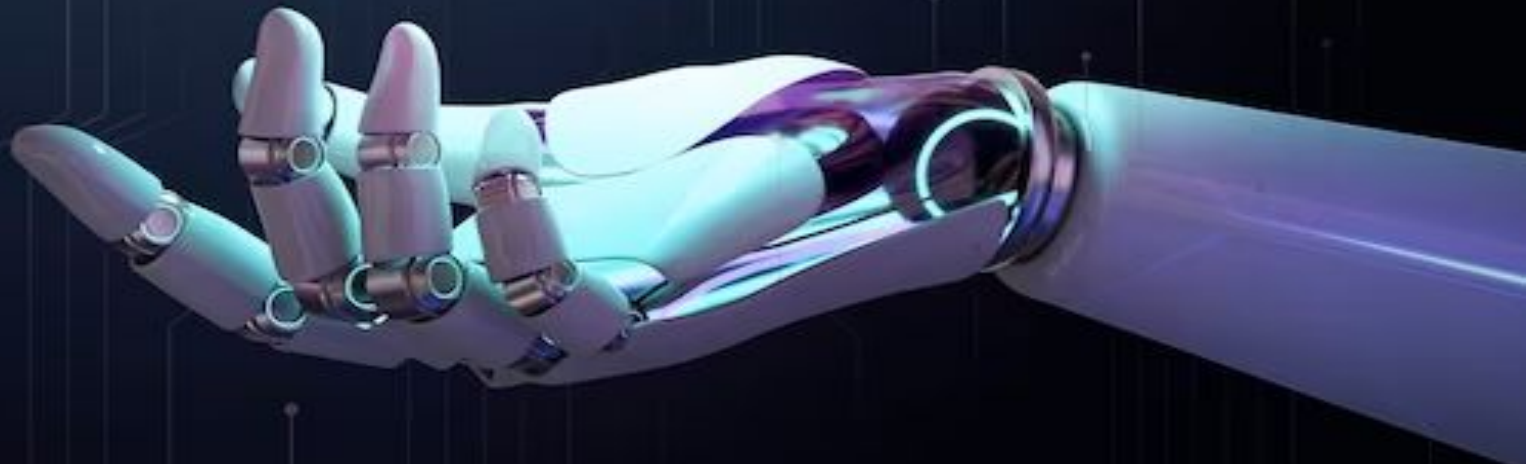
Visão geral do curso

1. Introdução à Inteligência Artificial
2. Aprendizado de Máquina
3. Processamento de Linguagem Natural
4. Visão Computacional
5. Reflexões e perspectivas da área

Oportunidades extras: semana de bate-papos de Inteligência Artificial com pesquisadores renomados da USP (8 a 12/abril)

As bases principais desse curso





Acesso a todo o material e links

Site do curso:

<https://cursosextensao.usp.br/course/view.php?id=3867>

- Slides das aulas
- Link para gravação
- Link para formulário de dúvidas (algumas serão respondidas ao vivo, outras pelos monitores e outras nos bate-papos)
- Link para sala virtual dos bate-papos da próxima semana
- Link para formulário de avaliação do aluno

Avaliação e certificado

Para alunos regularmente inscritos (via sistema Apolo da USP)

- Presença de manhã e de tarde (registrada via formulário online)
 - Avaliação com questões de múltipla escolha, com **nota mínima 7**
-
- Certificado da USP (enviado por e-mail em até 4 meses)

Sistema Apolo e comunicação

Toda a comunicação é feita por e-mail (o que foi informado na hora da inscrição)

Não é preciso criar usuário ou fazer login no sistema Apolo da USP

- Mas o sistema ainda deve enviar algumas mensagens a todos, pois os dados dos inscritos estão sendo processados em lotes

Formulário para postar dúvidas e comentários

<https://forms.gle/3CJhiHRyuQmEmqLs9>



Formulário para presença (manhã)

<https://forms.gle/44BKVrv1L8QwNHHV6>



Introdução à Inteligência Artificial

Thiago A. S. Pardo

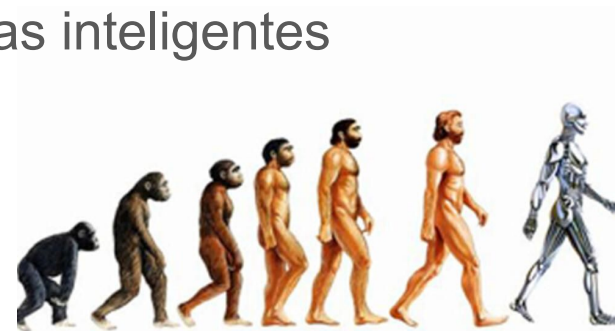
A importância da inteligência

Como *homo sapiens*, valorizamos a **inteligência**

- Esforço em desvendar como pensamos e agimos
- Relação entre inteligência e humanidade
- O poder de nosso cérebro em perceber, entender, prever e manipular um mundo complexo

Na **Inteligência Artificial (IA)**

- Além do entendimento, a construção de sistemas inteligentes
- Máquinas que podem computar como agir de maneira efetiva e segura nas mais diversas situações



O sistema do momento

OpenAI



ChatGPT

As percepções... um dos motivos para o curso :-)

A IA chegou!

Só mais um sistema. Logo todos esquecem.

É o apocalipse tecnológico!



As percepções... um dos motivos para o curso :-)



A IA chegou!

Só mais um sistema. Logo todos esquecem.

É o apocalipse tecnológico!



Apenas alguns exemplos

- Entretenimento: sistemas de recomendação (filmes, séries, músicas, conteúdos de redes sociais), jogos (xadrez, Go, RPGs)
- Texto: chatbots, sumarização de textos, tradução automática, revisão gramatical, análise de sentimentos
- Fala: reconhecimento e síntese de voz, interfaces com sistemas computacionais
- Matemática: prova de teoremas

Apenas alguns exemplos

- Trânsito: apoio ao motorista, direção autônoma, rotas
- Saúde: apoio ao diagnóstico, descoberta de novos medicamentos, rastreamento de epidemias
- Direito: busca por leis e jurisprudências, argumentação
- Clima: apoio a previsão do tempo, predição de catástrofes
- Etc.

Nossos primeiros passos

O que significa ser inteligente?

Como se faz um sistema inteligente?

Um pouco da história

Tendências

Desafios



Definição de IA

Campo tão relevante, mas **difícil de definir**



O que é inteligência?

- Fidelidade à forma que o humano faz? Ou basta ter ação de maneira racional?
 - Não quer dizer que humanos sejam irracionais, mas que as decisões humanas não são sempre “matematicamente perfeitas” ;-)

Definição de IA

Campo tão relevante, mas **difícil de definir**



O que é inteligência?

- Propriedade dos processos de pensamento e raciocínio? Ou basta comportamento inteligente (de caracterização externa)?
- É única ou é um aglomerado de “inteligências específicas”?

Definição de IA

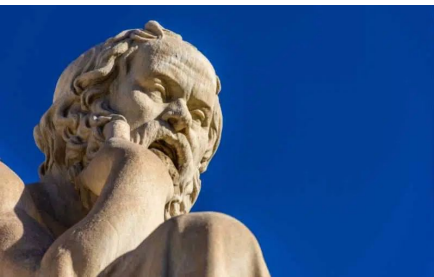
Campo tão relevante, mas **difícil de definir**



- Podem parecer tópicos muito abstratos, mas tiveram **grande impacto na história** da área
- Algumas **decisões** facilitam a sistematização dos conhecimentos envolvidos e o trabalho em IA

Algumas comparações interessantes

- O caso da Lógica: pensamento humano dedutivo
 - Sócrates é um homem e todos os homens são mortais ⇒ Sócrates é mortal
 - Difícil de generalizar para todos os aspectos da vida
 - Dificuldade em lidar com incertezas, subjetividade e vagueza



*Sócrates adora legumes e verduras,
mas odeia brócolis!*



Algumas comparações interessantes

- O caso do avião: apesar de não fazerem como na natureza, aviões voam



- Uma questão de engenharia

As perguntas difíceis da IA

(mas não só da IA)

Discussões históricas

- **Inteligência implica consciência e livre arbítrio?**



- Dualismo de René Descartes (1596-1650): a mente é mais do que matéria, sendo que há uma parte (o espírito?) que não segue as leis da Física
 - *If the mind is governed entirely by physical laws, then it has no more free will than a rock “deciding” to fall downward*

As perguntas difíceis da IA

(mas não só da IA)

Discussões históricas

- **Inteligência implica consciência e livre arbítrio?**
 - Materialismo: o funcionamento do cérebro segundo as leis da Física constitui a mente; o livre arbítrio relaciona-se a como a percepção das opções aparece para o indivíduo

As perguntas difíceis da IA (mas não só da IA)

Discussões históricas

- Queremos consciência e livre arbítrio para as máquinas?



- Discussão científica fascinante (que pode facilmente extrapolar a Ciência)

Os limites da IA

Até onde a IA deve ir para atingir seu objetivo? (ou deveriam ser “nossos” objetivos?)



Carros inteligentes

- Efetividade: chegar ao destino a qualquer custo, mesmo que machuque pessoas (outros motoristas ou o próprio passageiro) ou danifique o carro? Em caso de acidente, a quem responsabilizar?
- Segurança: se maximizada, talvez o carro não devesse sair da garagem

Os limites da IA

Até onde a IA deve ir para atingir seu objetivo? (ou deveriam ser “nossos” objetivos?)

- Mesmo em situações “mais controladas” como no xadrez
 - A máquina pode chantagear ou fazer barulho para atrapalhar o oponente?
 - Um dos primeiros livros de xadrez (Lopez, 1561): *“Always place the board so the sun is in your opponent’s eyes.”*



Os limites da IA

Até onde a IA deve ir para atingir seu objetivo? (ou deveriam ser “nossos” objetivos?)

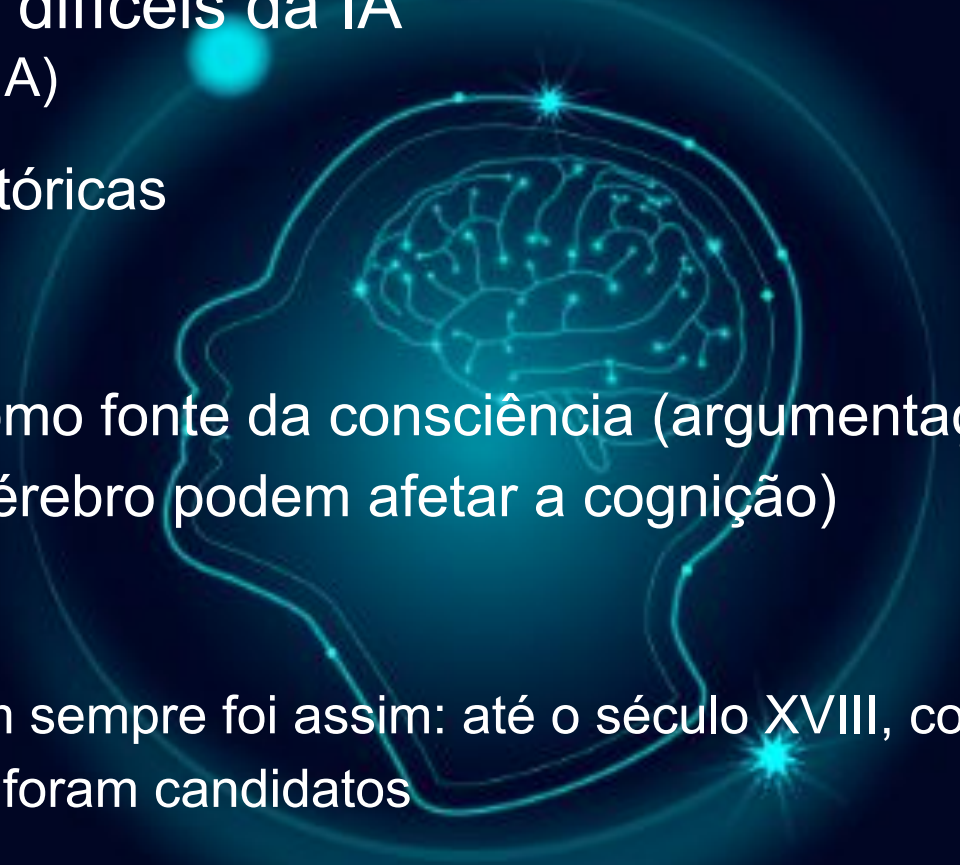
- ChatGPT, grandes modelos de língua e cia: apoio ou substituição?
quais os impactos para indivíduos e sociedade?
 - Ensino e aprendizagem
 - Medicina
 - Psicologia
 - Direito
 - Etc.



As perguntas difíceis da IA (mas não só da IA)

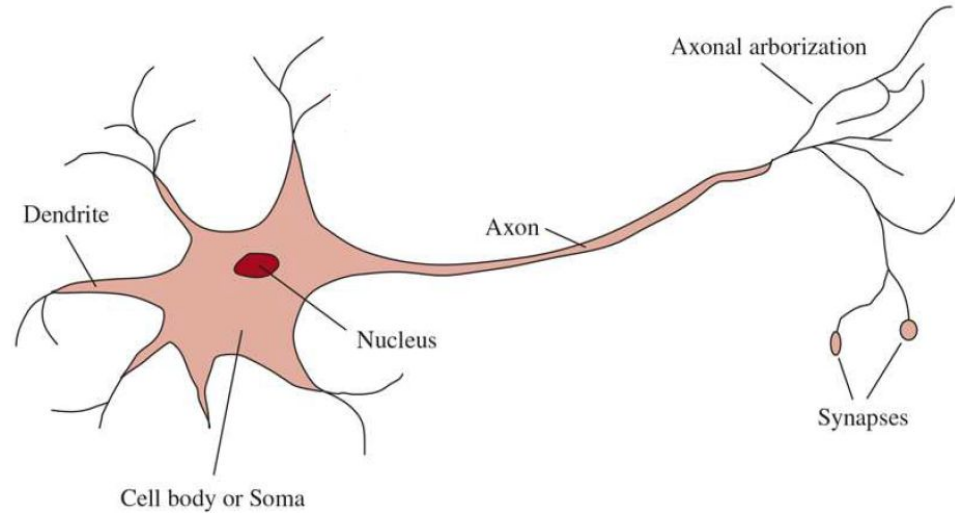
Discussões históricas

- Cérebro como fonte da consciência (argumentação óbvia: danos ao cérebro podem afetar a cognição)
 - Mas nem sempre foi assim: até o século XVIII, coração e baço também foram candidatos



As perguntas difíceis da IA (mas não só da IA)

Discussões históricas

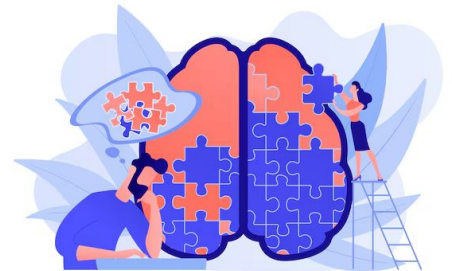


- Searle (1992): *brains cause minds*
 - Uma coleção de células “simples”, com reações eletroquímicas propagadas entre si, pode levar a pensamento, ação e consciência
 - Neurônios são a base da abordagem predominante na IA atualmente

As perguntas difíceis da IA (mas não só da IA)

Discussões históricas

- Armazenamento de memórias e funções cognitivas superiores: apesar de conhecermos regiões ativadas por determinados estímulos (graças à Neurociência), **não sabemos exatamente como esses processos acontecem**
 - Alguns desses processos são as bases de modelos de IA



Máquinas e inteligência

Luger (2013): O problema de definir o campo inteiro da inteligência artificial é semelhante ao de definir a própria inteligência: ela é uma única faculdade ou é apenas um nome para a coleção de capacidades distintas e não relacionadas? Até que ponto a inteligência é aprendida e não existe desde o nascimento? O que acontece exatamente quando ocorre o aprendizado? O que é criatividade? O que é intuição? A inteligência pode ser deduzida do comportamento observável ou ela requer evidências de um mecanismo interno em particular? Como o conhecimento é representado no tecido nervoso de um ser humano e que lições isso traz para o projeto de máquinas inteligentes? O que é autopercepção? Que papel ela desempenha na inteligência? Além disso, o conhecimento sobre a inteligência humana é necessário para construir um programa inteligente, ou uma técnica estritamente de “engenharia” é suficiente para tratar o problema? É possível conseguir inteligência em um computador, ou uma entidade inteligente requer a riqueza de sensações e experiências que só poderiam ser encontradas em uma existência biológica?

MIND

A QUARTERLY REVIEW

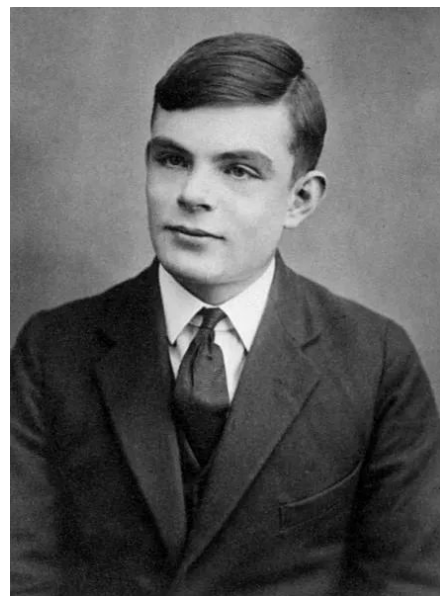
OF

PSYCHOLOGY AND PHILOSOPHY



I.—COMPUTING MACHINERY AND
INTELLIGENCE

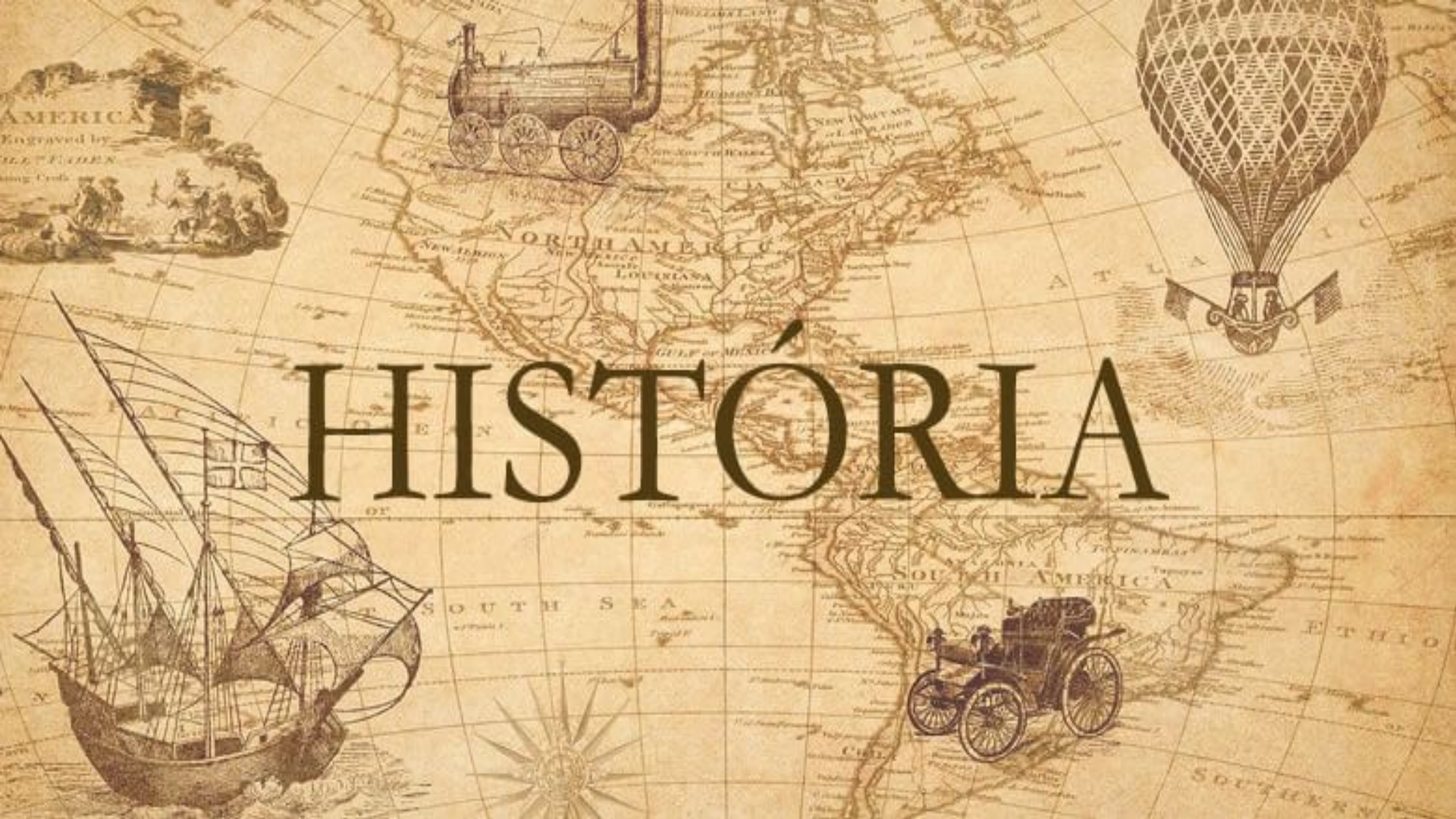
BY A. M. TURING



**A questão é
antiga!**

1. The Imitation Game

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

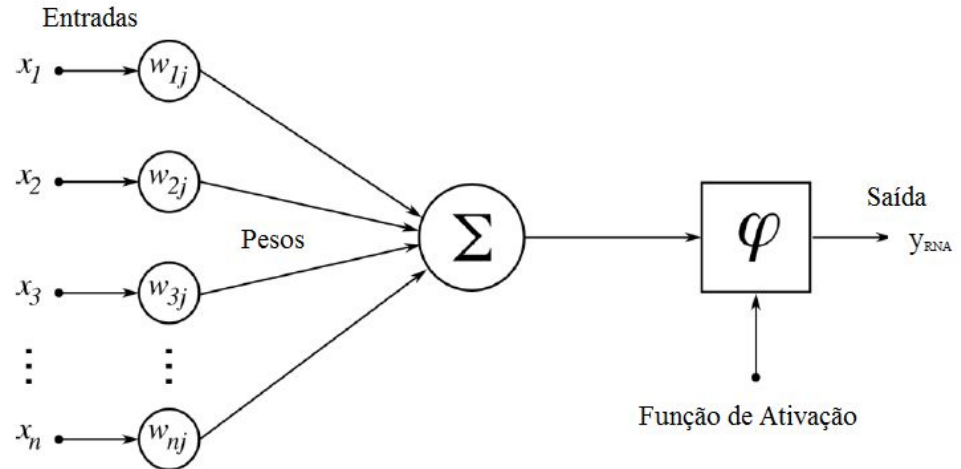
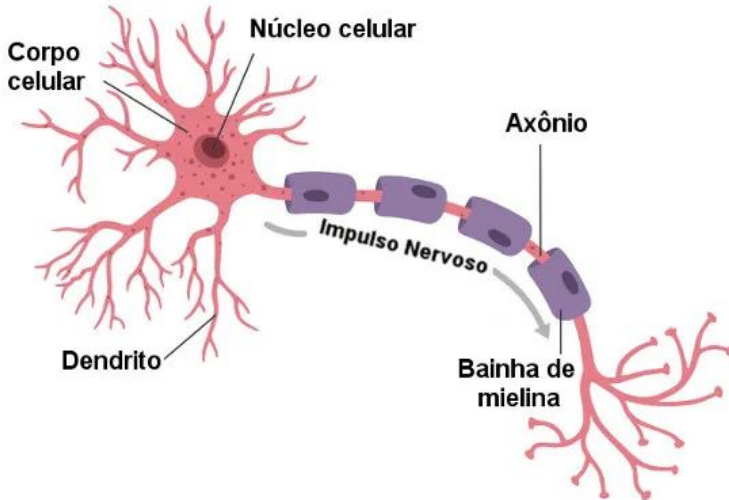


HISTÓRIA

Um pouco de história

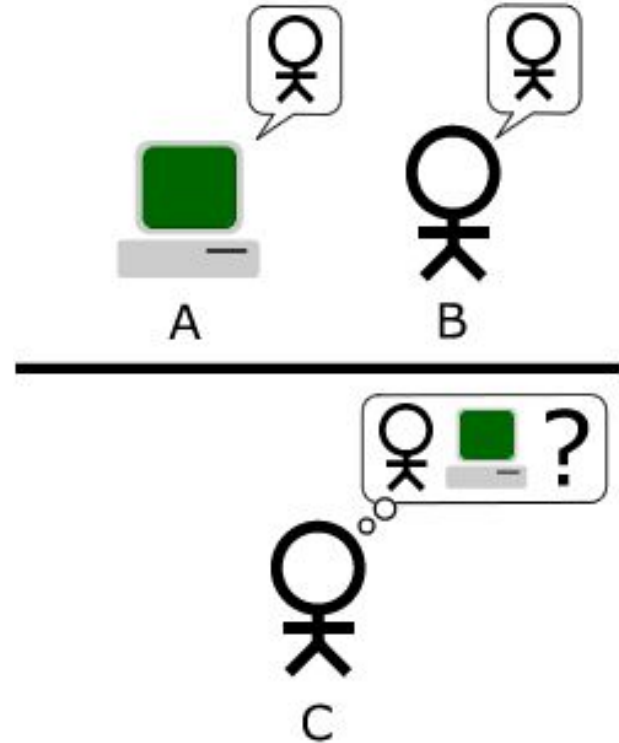
- **Gestação da IA (1943-1956)**

- Warren McCulloch e Walter Pitts (1943): a proposta do primeiro neurônio artificial
 - Considerado por muitos o primeiro trabalho de IA



Um pouco de história

- **Gestação da IA (1943-1956)**
 - Sistemas jogadores de damas
 - Alan Turing: teste de Turing e as ideias iniciais do aprendizado de máquina

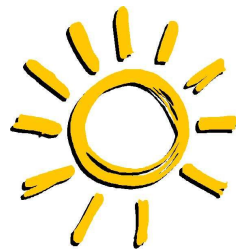


Um pouco de história

- **Gestação da IA (1943-1956)**

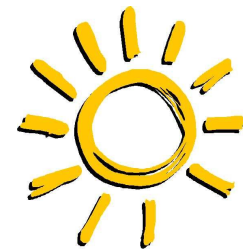
- Workshop de dois meses no Dartmouth College, nos EUA (1956), promovido por John McCarthy e colegas: o termo “Inteligência Artificial” é cunhado

- *We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.*



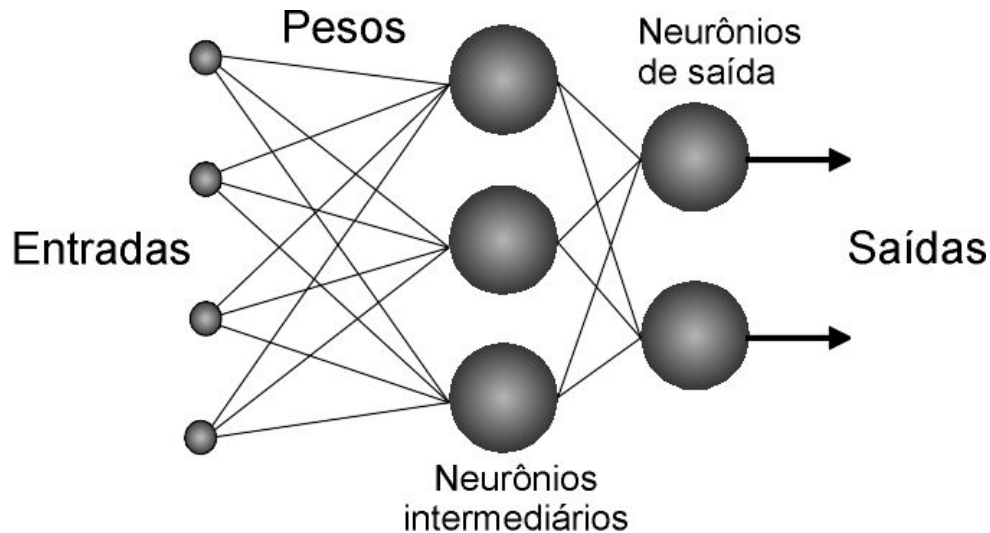
Um pouco de história

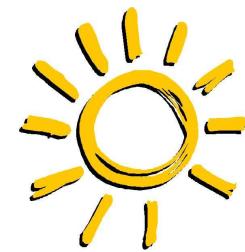
- **Entusiasmo e grandes expectativas (1952-1969)**
 - Resolução de jogos, quebra-cabeças e problemas matemáticos
 - Allen Newell e Herbert Simon propõem o *General Problem Solver* (GPS) para resolver quebra-cabeças da “mesma forma que humanos”
 - Arthur Samuel e seu sistema jogador de dama (amador) que aprendia
 - John McCarthy cria a linguagem de programação LISP: a representação de conhecimento e o raciocínio



Um pouco de história

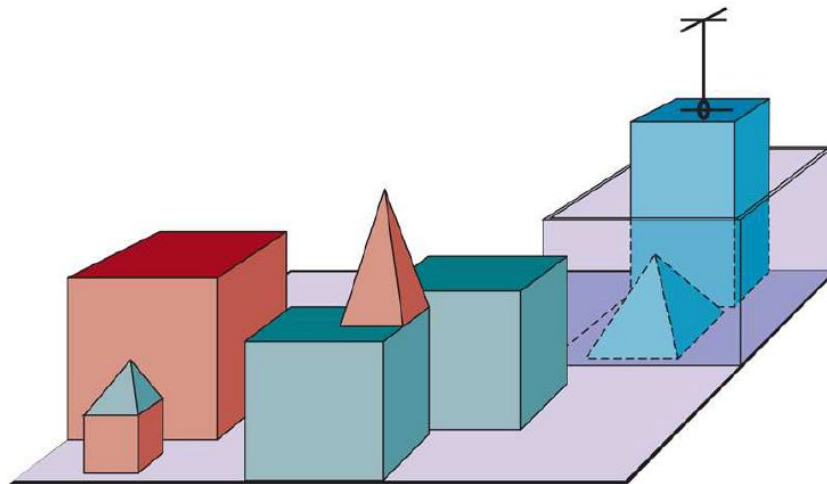
- **Entusiasmo e grandes expectativas (1952-1969)**
 - Os avanços nas construções de redes neurais artificiais





Um pouco de história

- **Entusiasmo e grandes expectativas (1952-1969)**
 - Marvin Minsky e a resolução de problemas limitados que requerem inteligência: o “mundo dos blocos” e a base de trabalhos em visão, planejamento, aprendizado e processamento de linguagem natural



*Pegue o bloco
mais alto e
coloque na caixa.*

Um pouco de história

- **Uma dose de realidade (1966-1973)**
 - A falha dos sistemas para problemas mais complexos
 - Dificuldades no entendimento e modelagem das tarefas
 - Intratabilidade de problemas e a dificuldade de escalar as soluções
 - A falha dos neurônios artificiais em lidar com um problema relativamente simples (o famoso ou-exclusivo)
 - Cortes de financiamento

Um pouco de história

- **Sistemas especialistas (1969-1986)**
 - **Conhecimento de domínio**: mais poder e raciocínio
 - O programa DENDRAL e a inferência da estrutura de moléculas
 - MYCIN e o diagnóstico de infecções sanguíneas: na época, tão bom quanto alguns especialistas e melhor do que médicos recém-formados
 - O **sucesso comercial** dos sistemas especialistas: equipes de especialistas de IA nas grandes empresas



Um pouco de história

- Sistemas especialistas (1969-1986)
 - Técnicas de representação de conhecimento: regras, lógica, frames e cia

Se febre alta, dor no corpo e dor nos olhos, então pode ser dengue
 - Tentativas de criação de computadores “inteligentes”:
linguagem Prolog, projeto de chips e design de interfaces



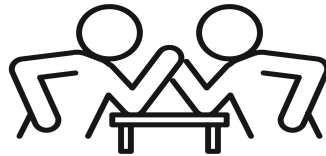
Um pouco de história

- **Sistemas especialistas (1969-1986)**
 - Novas frustrações e “inverno da IA”: promessas extravagantes não cumpridas, dificuldades em lidar com incerteza e a falta de aprendizado



Um pouco de história

- O retorno das redes neurais (1986 até os tempos atuais)
 - Novos esforços
 - Métodos para treinamento das redes
 - Os modelos conexionistas rivalizam com métodos simbólicos e lógicos
- A discussão de símbolos versus números



Um pouco de história

- Raciocínio probabilístico e aprendizado de máquina (1987 até os tempos atuais)
 - **Mudança de paradigma**
 - Probabilidade em vez de lógica
 - Aprendizado automático em vez de codificação manual do conhecimento
 - Resultados experimentais em vez de discussões filosóficas
 - Aplicações reais em vez de exemplos controlados (*toy*)



Um pouco de história

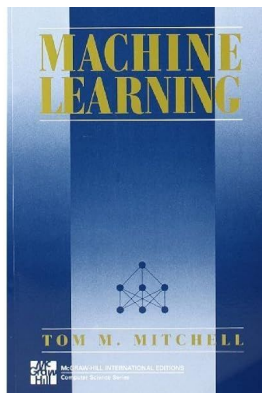
- Raciocínio probabilístico e aprendizado de máquina (1987 até os tempos atuais)
 - Desenvolvimento de conjuntos de dados e competições: a importância dos dados
 - A conexão com outras áreas: Estatística, Pesquisa Operacional, Controle e Automação, etc.
 - A reunificação da área: Processamento de Linguagem Natural, Visão Computacional, Robótica, Sistemas Multiagente



O livro clássico de Tom Mitchell (1997)

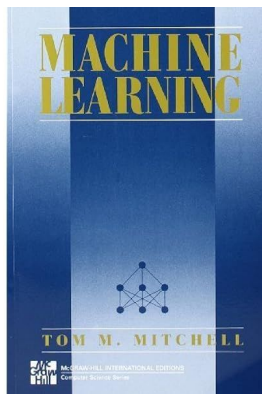
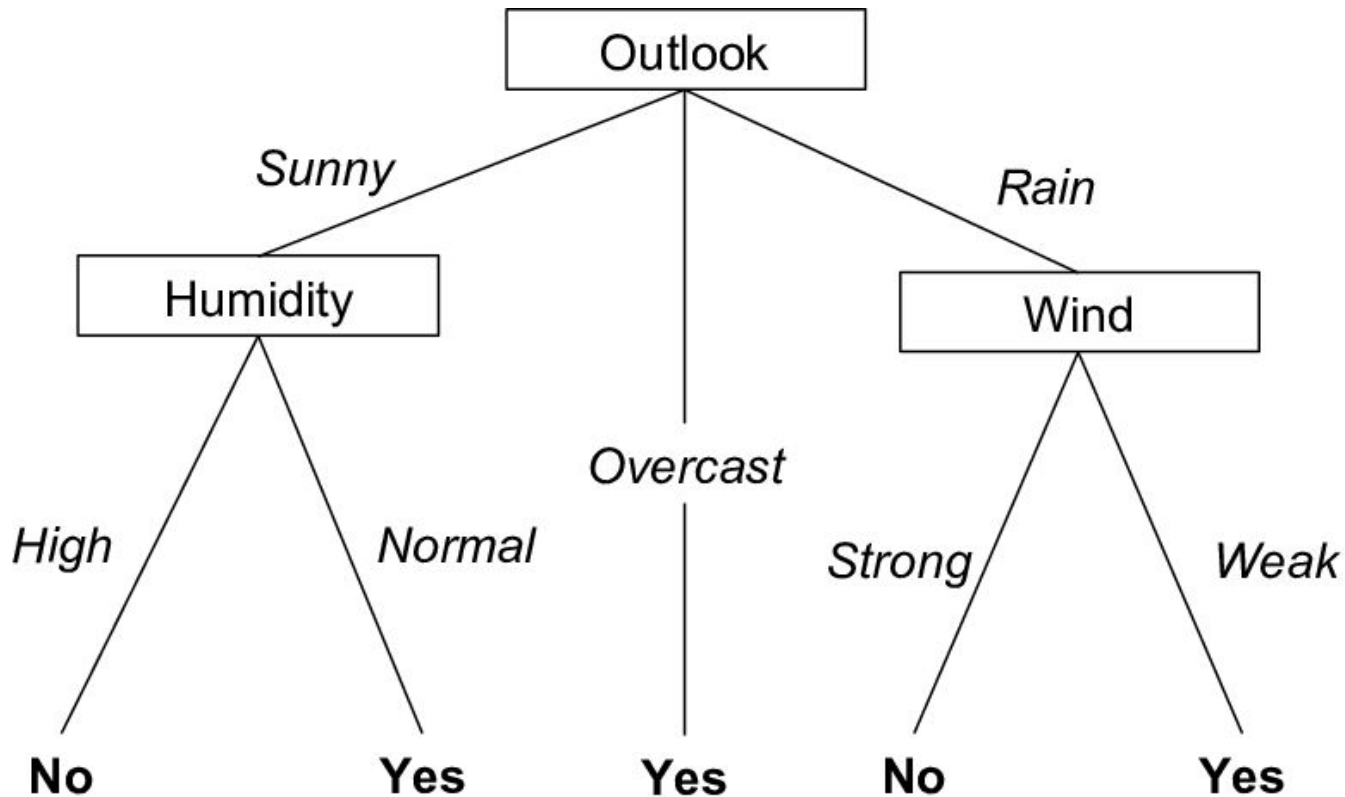
Aprendendo dos dados: quando jogar tênis?

Day	<i>Outlook</i>	<i>Temperature</i>	<i>Humidity</i>	<i>Wind</i>	<i>PlayTennis</i>
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No



O livro clássico de Tom Mitchell (1997)

Aprendendo dos dados: quando jogar tênis?



Um pouco de história

- *Big data (2001 até os tempos atuais)*
 - Web e muitos dados
 - Muitos tipos: textos, imagens, áudios, vídeos
 - Dados específicos de medicina, leis, trânsito, redes sociais, etc.
 - Aumento do poder computacional
 - Aprendizado de máquina (especialmente sem supervisão humana)
 - Impacto comercial

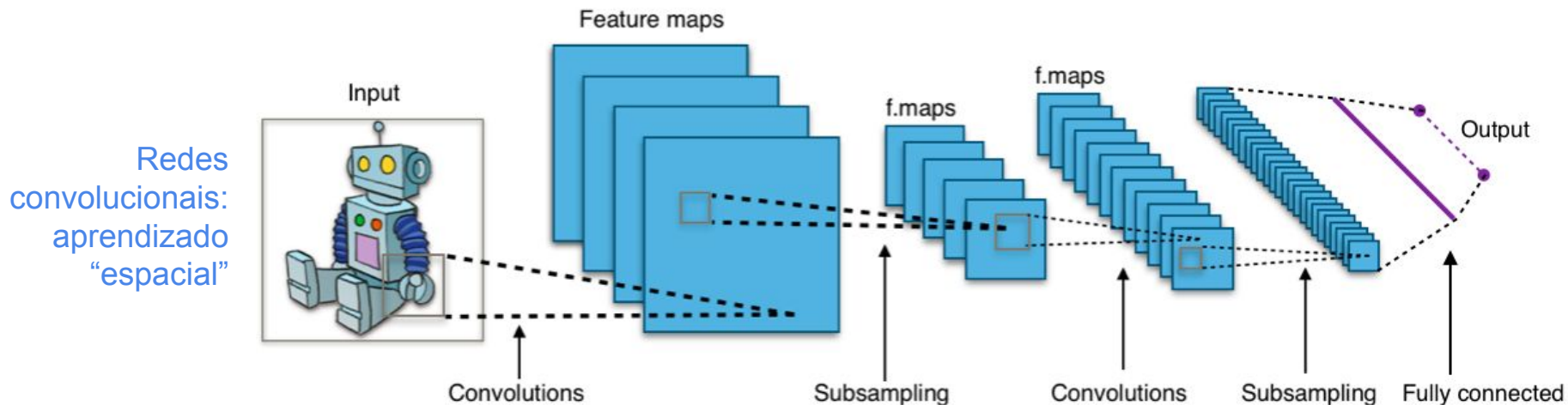
Um pouco de história

- *Big data* (2001 até os tempos atuais)
 - IBM Watson vence campeões humanos no Jeopardy! (2011)



Um pouco de história

- **Aprendizado profundo (2011 até os tempos atuais)**
 - Apesar de datarem dos anos 1970, grandes avanços e impactos na última década em muitas tarefas e frentes da IA

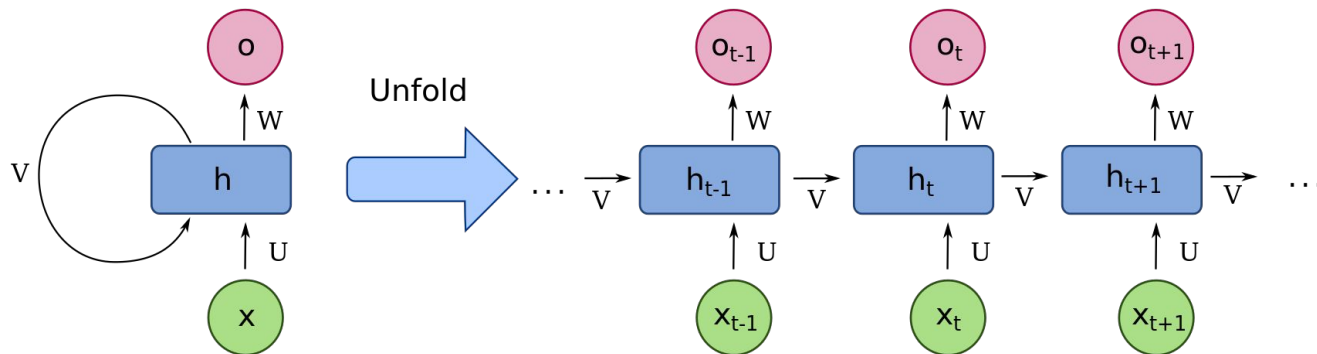


Um pouco de história

- **Aprendizado profundo (2011 até os tempos atuais)**

- Apesar de datarem dos anos 1970, grandes avanços e impactos na última década em muitas tarefas e frentes da IA

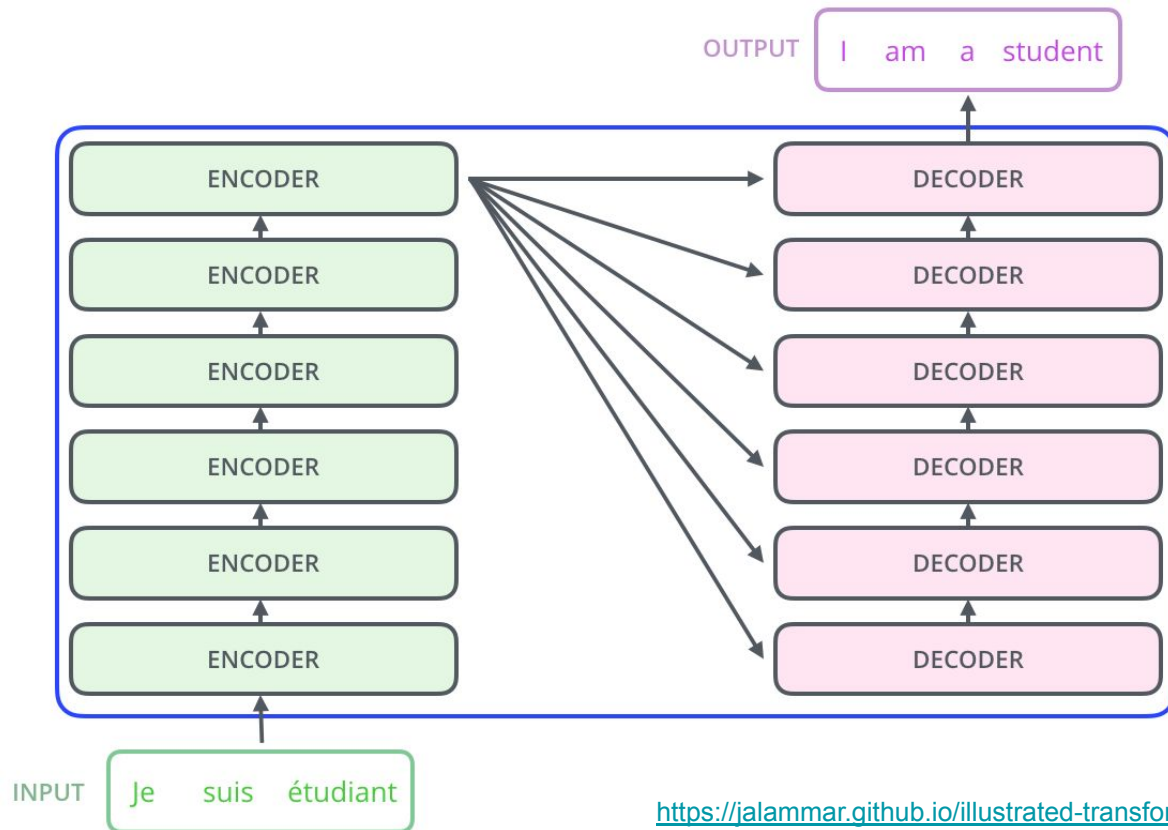
Redes
recorrentes:
processando
seqüências



Um pouco de história

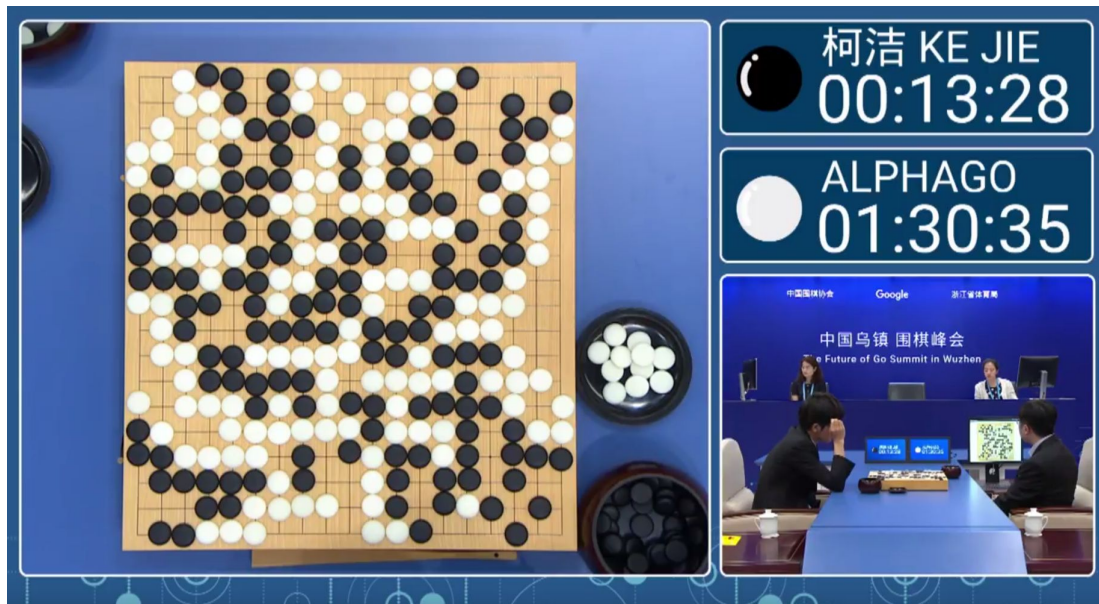
Transformers:
atenção & pré-treino e
ajuste-fino

(e o que está por trás da
maioria dos tradutores
automáticos recentes,
dos grandes modelos de
língua, etc.)



Um pouco de história

- **Aprendizado profundo (2011 até os tempos atuais)**
 - AlphaGo vence campeões humanos no jogo Go (2016)



Um pouco de história

- **Aprendizado profundo (2011 até os tempos atuais)**

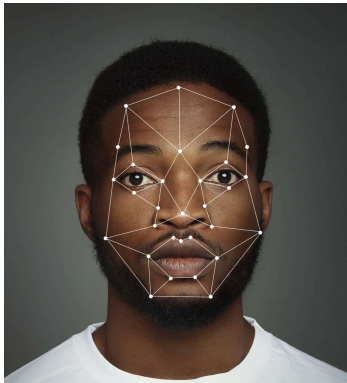
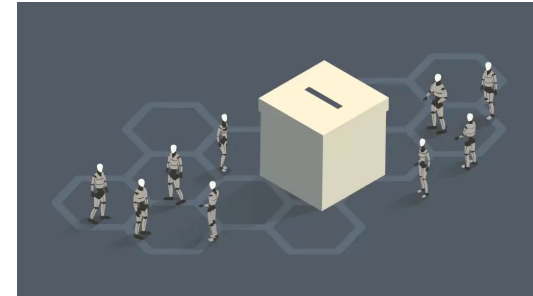
- Entrada de grandes empresas no cenário



- Aplicações surpreendentes (e o medo de um novo “**inverno**”)
- Dependência de dados e grande poder computacional

Um pouco de história

- **Aprendizado profundo (2011 até os tempos atuais)**
 - Público em geral com grande interesse na área
 - A **preocupação** de governos e sociedades
 - O uso inapropriado da tecnologia: relatos mais frequentes!



ONU adota primeira resolução global sobre uso da inteligência artificial

Texto, aprovado por unanimidade, enfatiza a necessidade de se promover sistemas de IA "seguros e confiáveis", diante de seu rápido desenvolvimento e uso na sociedade

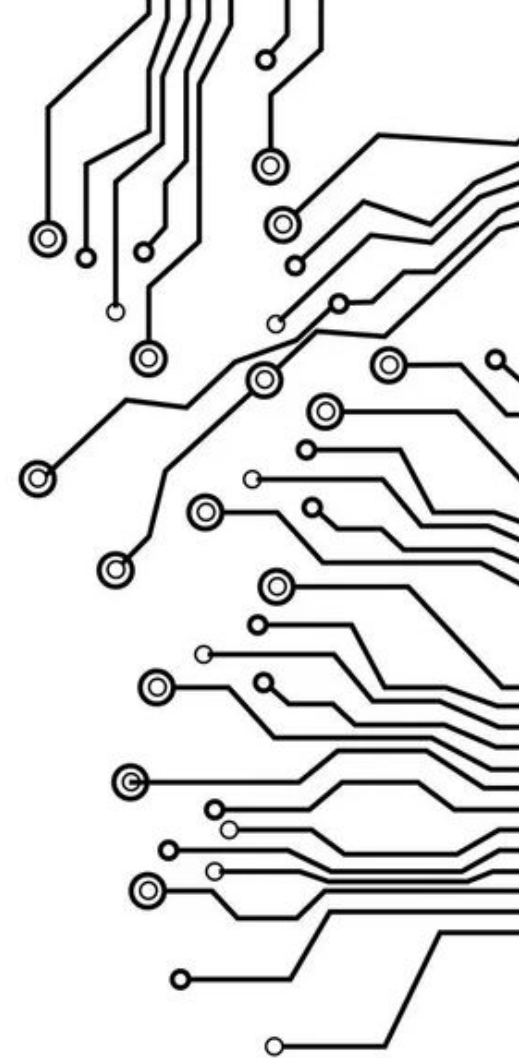


Março/2024

<https://oglobo.globo.com/economia/tecnologia/noticia/2024/03/22/onu-adota-primeira-resolucao-global-sobre-uso-da-inteligencia-artificial.ghtml>

Dados atuais

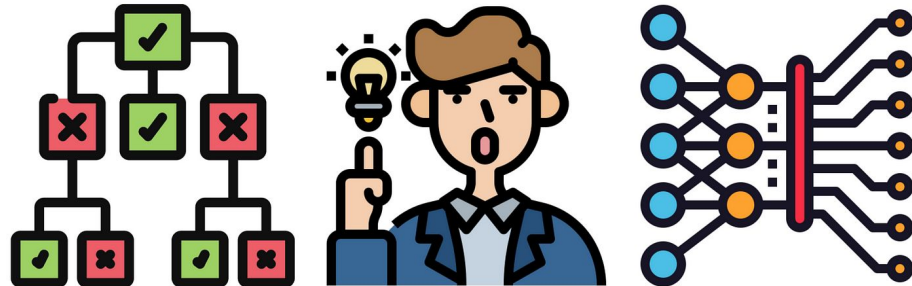
- [AI Index Annual Report](#)
 - **Balço de 2023**
 - O avanço das grandes empresas
 - Processo de saturoação dos dados utilizados para avaliação
 - Uso da IA para o progresso científico: novos medicamentos, processos matemáticos melhores, etc.



Dados atuais

- [AI Index Annual Report](#)
 - **Balanco de 2023**
 - Incidentes com o uso da IA
 - A importância ainda maior de questões éticas e de explicabilidade
 - Interesse público na regulação da IA

<https://towardsdatascience.com/interperable-vs-explainable-machine-learning-1fa525e12f48>



Dados atuais



- *AI Index Annual Report*
 - **Balanço de 2023**
 - Alto impacto ambiental da IA, especialmente no treinamento de modelos de língua

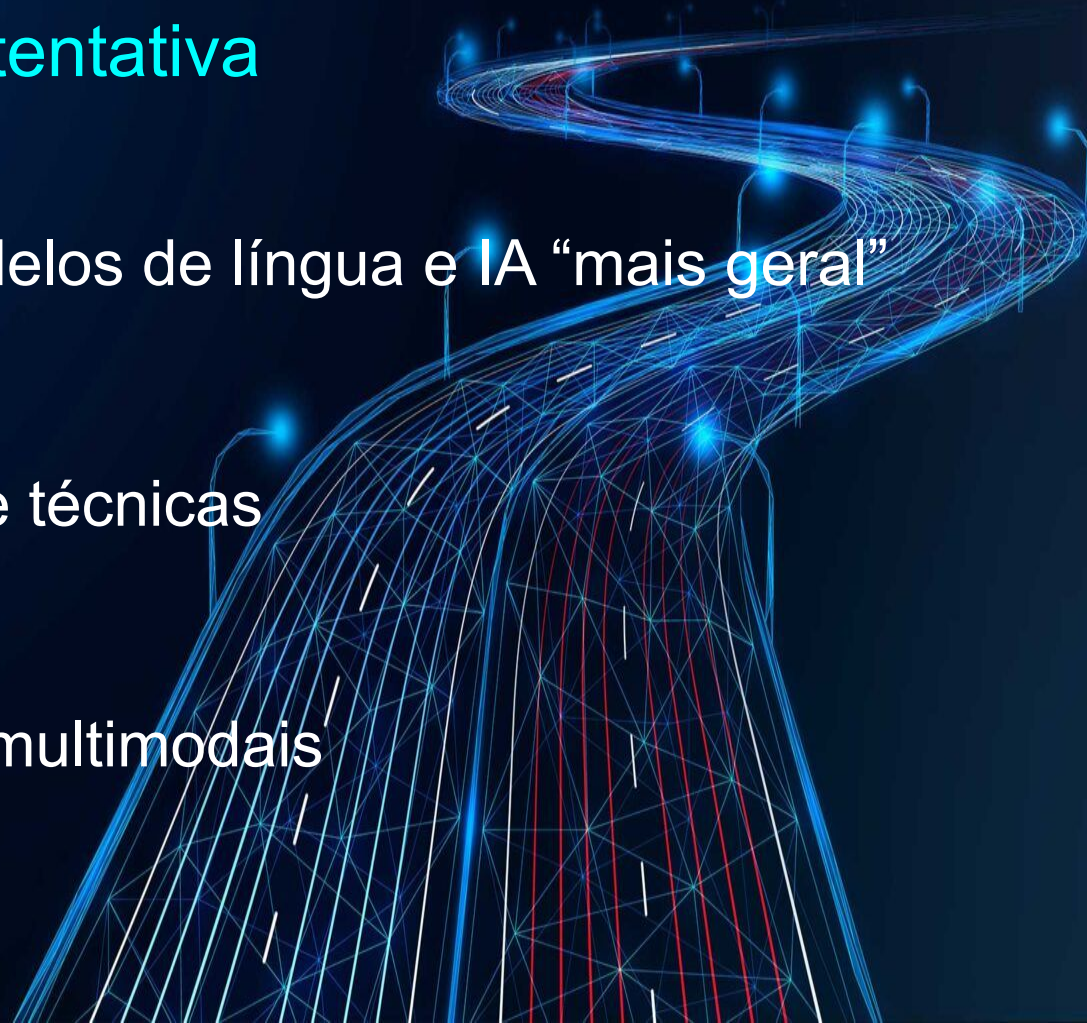
O futuro

- A história nos mostra que somos limitados em prever o futuro
 - *“Conhecimento especializado é o caminho”*
 - *“Probabilidades são o principal caminho”*
 - *“Aprendizado de máquina pode não ser teoricamente sólido”*
 - *“Modelos neurais não são explicáveis”*
 - *“Grandes modelos de língua são o caminho”*
- A característica da IA de usufruir do passado
 - Nenhuma técnica foi completamente abandonada
 - Há tarefas que demandam técnicas específicas



O futuro: uma tentativa

- Grandes modelos de língua e IA “mais geral”
- Hibridismo de técnicas
- Abordagens multimodais





Formulário para presença (manhã)

<https://forms.gle/44BKVrv1L8QwNHHV6>

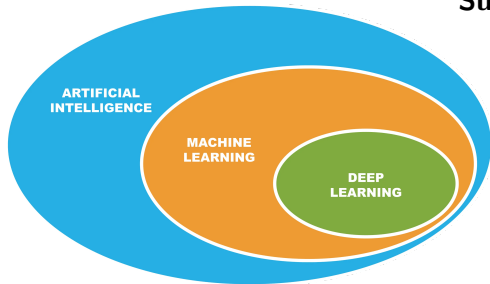




Aprendizado de Máquina

Nina S. T. Hirata

IA – Técnicas baseadas em dados



Subáreas em destaque

Machine learning

Deep learning

Fatores relevantes

Capacidade de transmissão,
armazenamento e
processamento de dados

Abundância de dados

Vamos começar falando sobre

Machine Learning (Aprendizado de Máquina)

Máquina: **Dispositivos COMPUTACIONAIS**

SÃO, ESSENCIALMENTE, MÁQUINAS QUE PROCESSAM DADOS

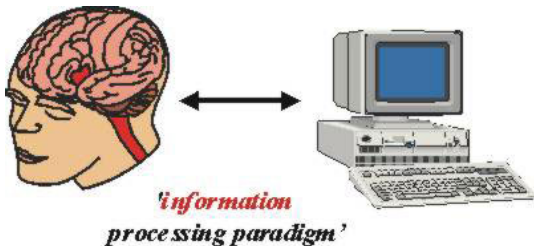


PERSPECTIVA COMPUTACIONAL de Machine Learning

Por trás da aparente autonomia ou inteligência,

há um **sistema computacional**

que processa dados



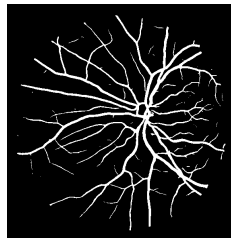
COMPUTADORES PROCESSAM DADOS: transformam dados ou geram alguma resposta

TAREFA: segmentação de vasos sanguíneos



Entrada

⇒ Processamento ⇒



Saída

COMPUTADORES PROCESSAM DADOS: transformam dados ou geram alguma resposta

TAREFA: adição de números

5, -8, 13, 2 \Rightarrow Processamento \Rightarrow 12

Entrada

Saída

COMPUTADORES PROCESSAM DADOS: transformam dados ou geram alguma resposta

TAREFA: tradução

Chinese (Simplified)



Portuguese

虽然我不懂歌词 但
是他唱得太好听了 ×

Suīrán wǒ bù dǒng gēcí dànshì tā
chàng dé tài hǎotīngle

Entrada

Embora eu não
entenda a letra, ele
cantou tão
lindamente

Saída

Computadores realizam processamento de dados

Processamentos podem ser caracterizado em termos de **entrada** e **saída**

Na prática, precisamos programá-los para que realizem o processamento correto/esperado

Um programa para calcular a soma

5, -8, 13, 2 \Rightarrow Processamento \Rightarrow 12

```
sum = 0
```

```
Repeat
```

```
    num = read next number
```

```
    sum = sum + num
```

```
Until there is no more numbers to be added
```

```
print(sum)
```

OUTRA TAREFA: reconhecimento de placa de veículos



Entrada

⇒ Processamento ⇒ **HQW - 5678**

Saída


ENTRADA é uma **foto** (imagem)

Precisamos reconhecer cada um dos caracteres na placa ...

Vamos simplificar e pensar sobre
como reconhecer imagens de **letra A**?

Podemos treinar uma “máquina” ?

Ideia central: “O QUE FAZER” ao invés de “COMO FAZER”
Por meio de exemplos de entrada-saída da tarefa-alvo!

 ⇒ 'a'

 ⇒ 'X'

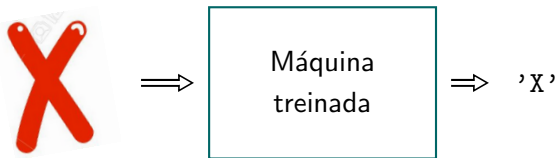
 ⇒ 'A'

 ⇒ 'd'

 ⇒ 'b'

 ⇒ 'M'

Uma vez bem treinada, espera-se que a máquina gere respostas corretas

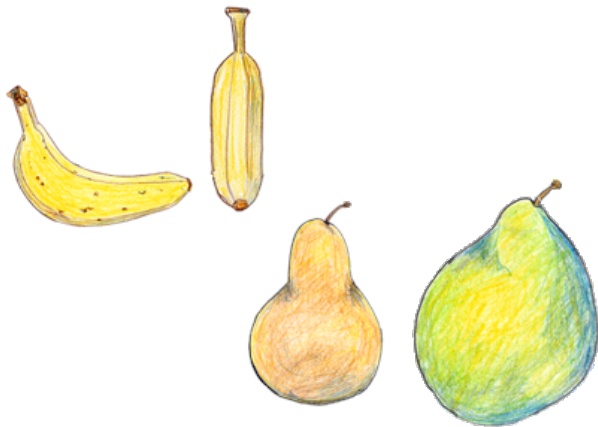


O que significa “treinar” uma máquina e como fazemos isso?

Ilustração considerando dois problemas de predição

- Classificação
- Regressão

Problema de classificação: separar bananas de peras



Simplificação: em vez da foto (imagem)
vamos considerar 2 atributos

- *yellowness*: cor (quão amarela ela é)
- *assymetry*: forma (quão simétrica ela é)

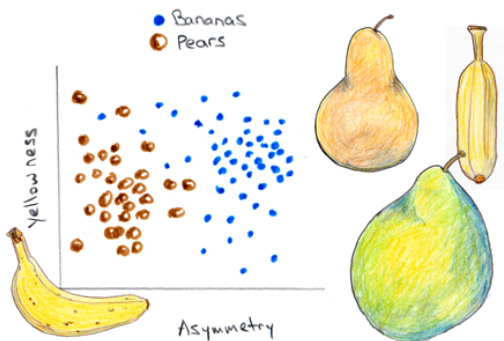


Simplificação: em vez da foto (imagem)
vamos considerar 2 atributos

- *yellowness*: cor (quão amarela ela é)
- *assymetry*: forma (quão simétrica ela é)



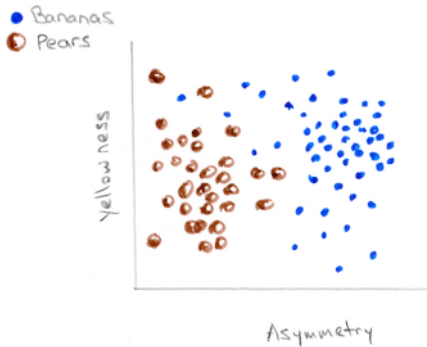
Considerando-se os dois atributos, podemos visualizar os frutos em um gráfico



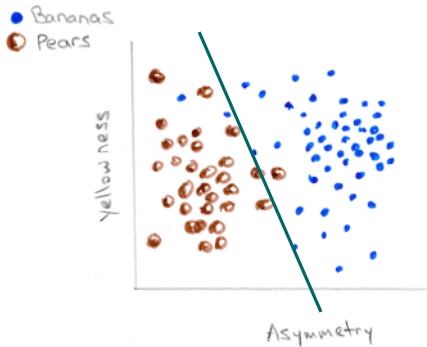
Peras: menos assimétricas, menos amarelas

Bananas: mais assimétricas, mais amarelas

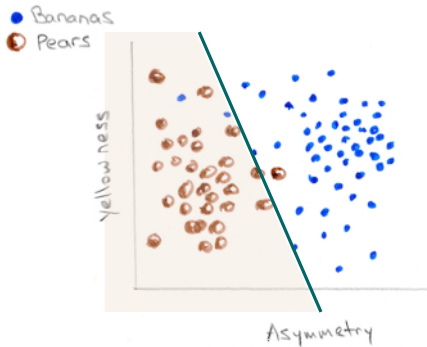
Uma estratégia para separar bananas e peras



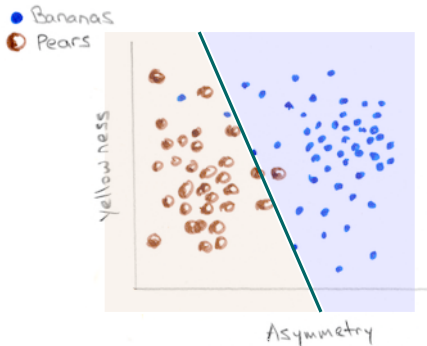
Uma estratégia para separar bananas e peras



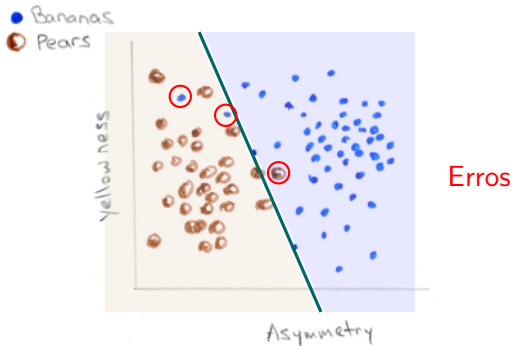
Uma estratégia para separar bananas e peras



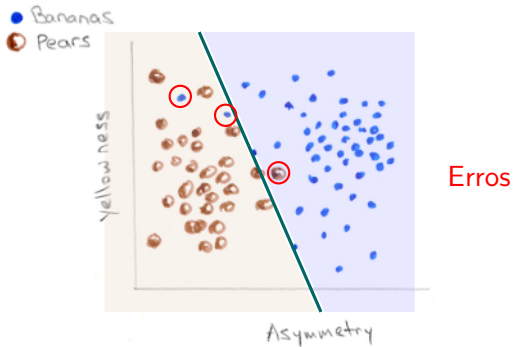
Uma estratégia para separar bananas e peras



Uma estratégia para separar bananas e peras

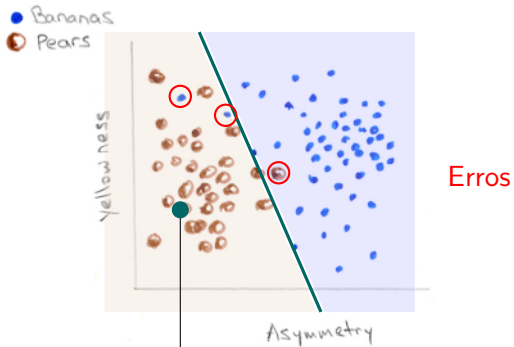


Uma estratégia para separar bananas e peras



Que fruta é essa?

Uma estratégia para separar bananas e peras

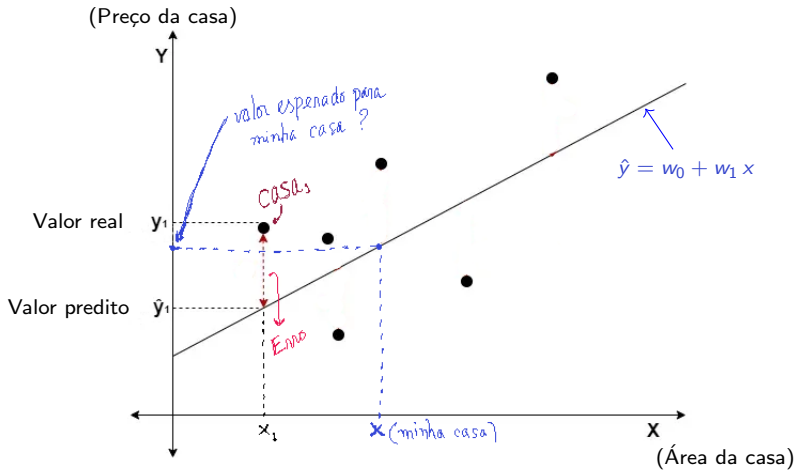


simétrica (baixa assimetria)
mais verde que amarela

Que fruta é essa?

Problema de regressão

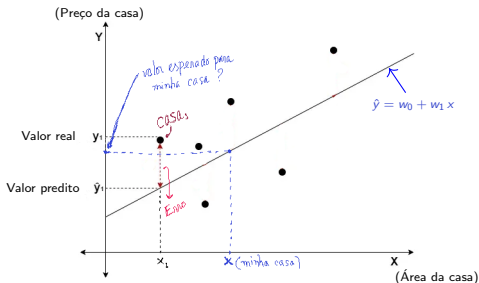
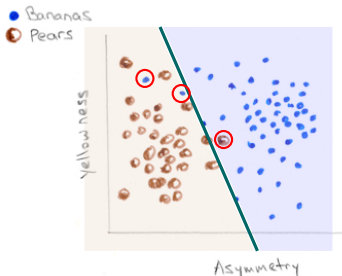
Quero vender uma casa e quero ter uma ideia de qual preço cobrar



<https://blog.gopenai.com/linear-regression-understanding-the-basics-1b1511499911>

Princípio do treinamento

Dentre todas as soluções possíveis, encontrar a que erra menos



$$\text{Erro: } Err(w_0, w_1) = \frac{1}{N} \sum_{i=1}^N \left(y_i - \underbrace{\hat{y}_i}_{w_0 + w_1 x_i} \right)^2$$

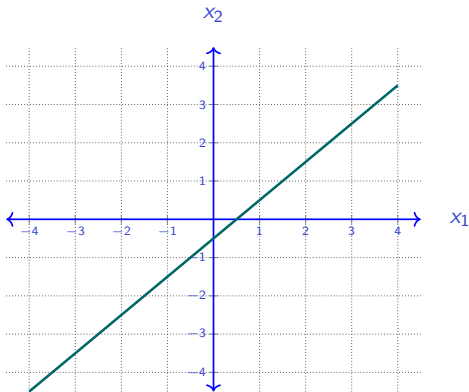
No caso de retas, ao alterar os coeficientes w_0 , w_1 e w_2 temos diferentes separadores

$$f(x_1, x_2) = w_0 + w_1 x_1 + w_2 x_2$$

No caso de retas, ao alterar os coeficientes w_0 , w_1 e w_2 temos diferentes separadores

$$f(x_1, x_2) = w_0 + w_1 x_1 + w_2 x_2$$

$$-0.5 + 1x_1 - 1x_2 = 0$$

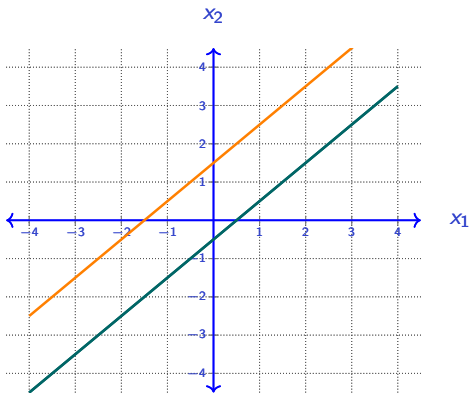


No caso de retas, ao alterar os coeficientes w_0 , w_1 e w_2 temos diferentes separadores

$$f(x_1, x_2) = w_0 + w_1 x_1 + w_2 x_2$$

$$-0.5 + 1x_1 - 1x_2 = 0$$

$$+1.5 + 1x_1 - 1x_2 = 0$$



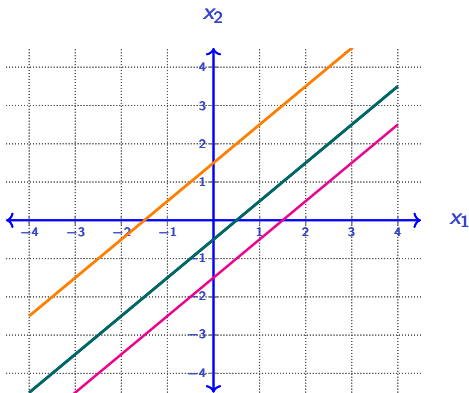
No caso de retas, ao alterar os coeficientes w_0 , w_1 e w_2 temos diferentes separadores

$$f(x_1, x_2) = w_0 + w_1 x_1 + w_2 x_2$$

$$-0.5 + 1x_1 - 1x_2 = 0$$

$$+1.5 + 1x_1 - 1x_2 = 0$$

$$-1.5 + 1x_1 - 1x_2 = 0$$



No caso de retas, ao alterar os coeficientes w_0 , w_1 e w_2 temos diferentes separadores

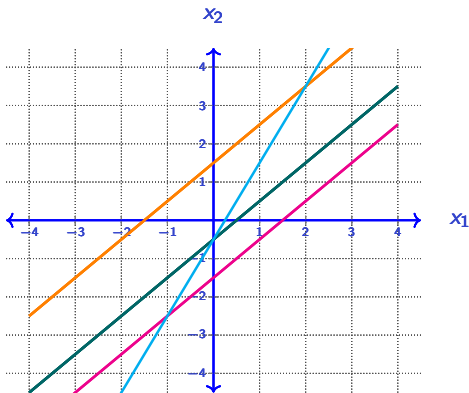
$$f(x_1, x_2) = w_0 + w_1 x_1 + w_2 x_2$$

$$-0.5 + 1x_1 - 1x_2 = 0$$

$$+1.5 + 1x_1 - 1x_2 = 0$$

$$-1.5 + 1x_1 - 1x_2 = 0$$

$$-0.5 + 2x_1 - 1x_2 = 0$$



No caso de retas, ao alterar os coeficientes w_0 , w_1 e w_2 temos diferentes separadores

$$f(x_1, x_2) = w_0 + w_1 x_1 + w_2 x_2$$

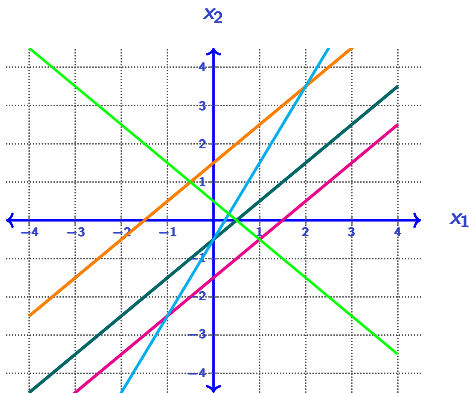
$$-0.5 + 1x_1 - 1x_2 = 0$$

$$+1.5 + 1x_1 - 1x_2 = 0$$

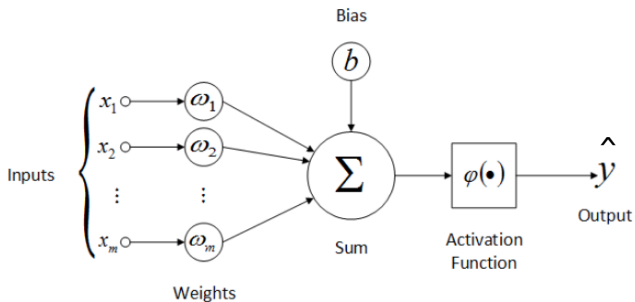
$$-1.5 + 1x_1 - 1x_2 = 0$$

$$-0.5 + 2x_1 - 1x_2 = 0$$

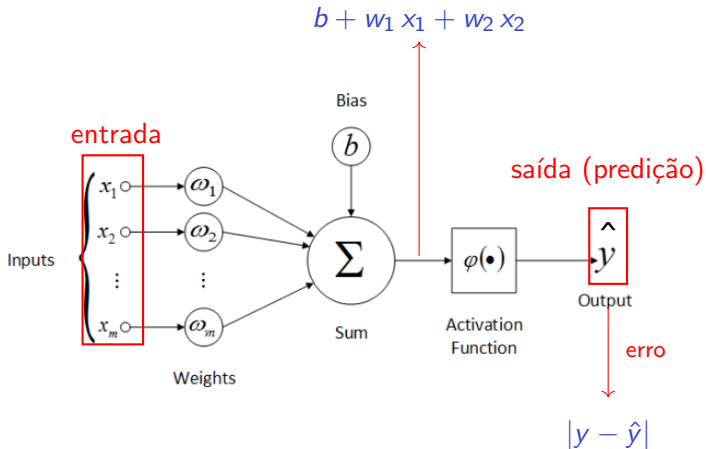
$$-0.5 + 1x_1 + 1x_2 = 0$$



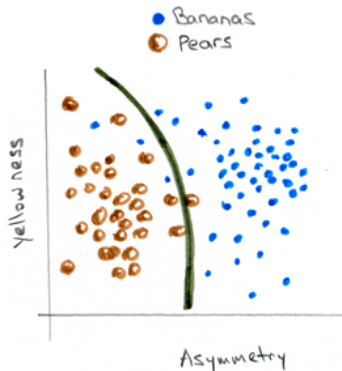
Neurônios correspondem a fronteiras lineares



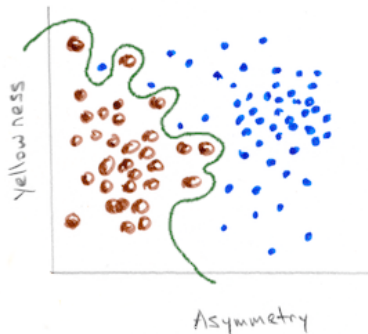
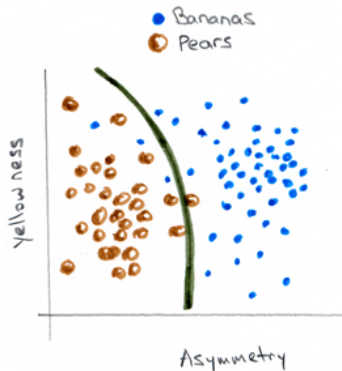
Neurônios correspondem a fronteiras lineares



Podemos pensar em curvas mais complexas, não retilíneas

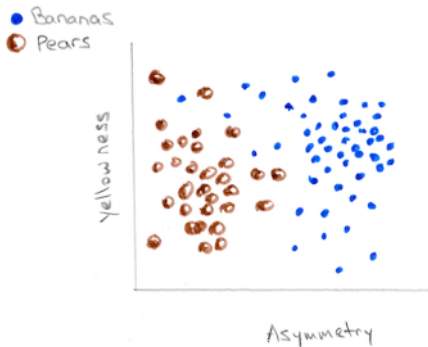


Podemos pensar em curvas mais complexas, não retilíneas

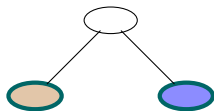
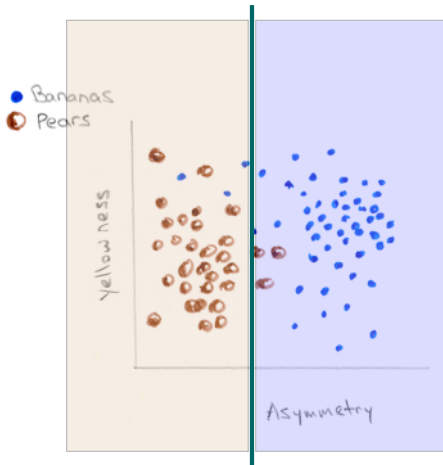


Sobreajuste (overfitting)!

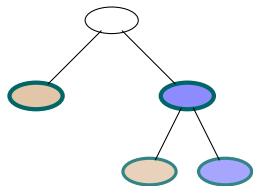
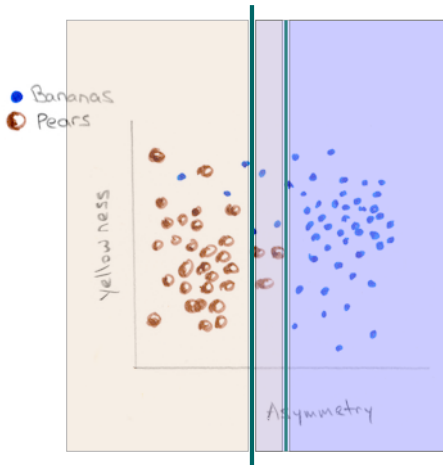
Ou fronteiras de separação mais complexas



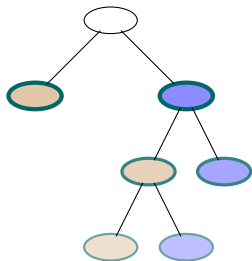
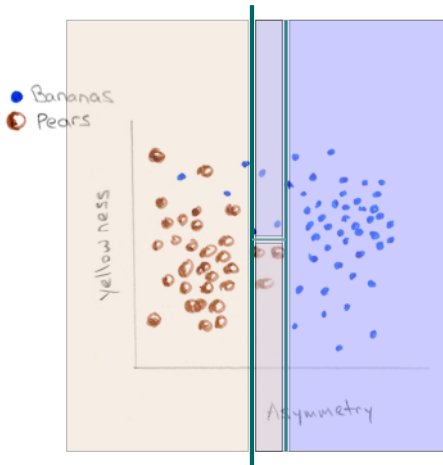
Ou fronteiras de separação mais complexas



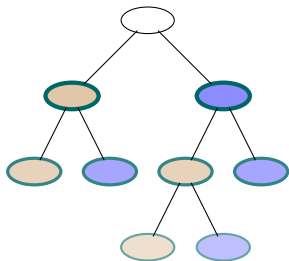
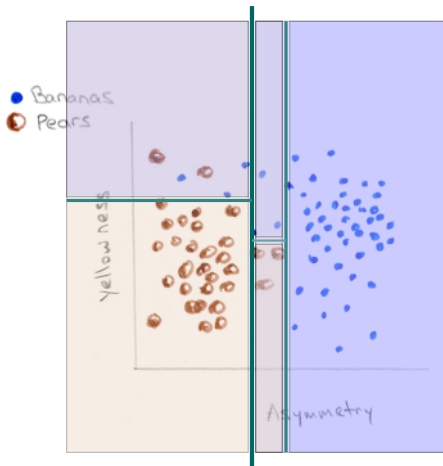
Ou fronteiras de separação mais complexas



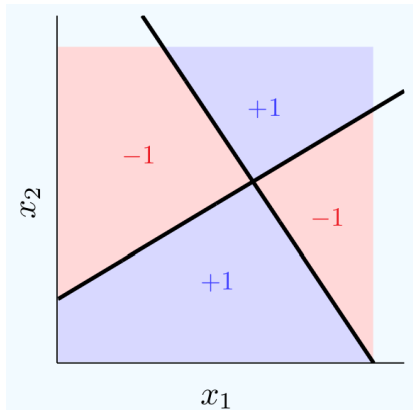
Ou fronteiras de separação mais complexas



Ou fronteiras de separação mais complexas

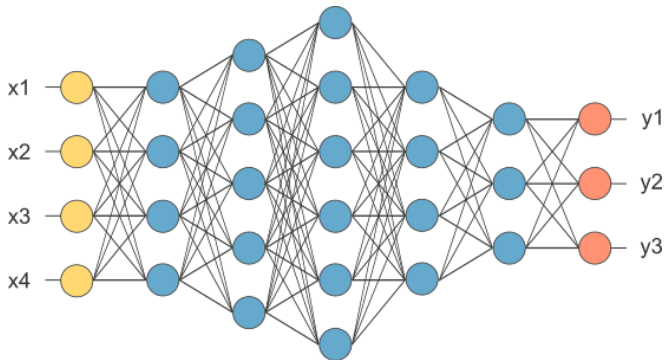


Fronteiras de separação elaboradas combinando múltiplos neurônios

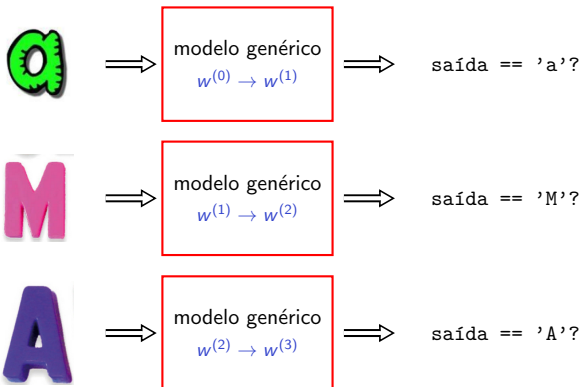


Redes neurais

Conseguem representar uma grande família de curvas, e podem ter bilhões de parâmetros. O treinamento delas é um processo iterativo.



Treinamento: **Ajuste iterativo dos parâmetros**



Visando sempre reduzir o erro!

Como aplicar ML

Definir uma tarefa

Preparar exemplos entrada-saída que ilustram a tarefa-alvo

Escolher um modelo genérico de processamento

Definir como um modelo genérico de processamento será ajustado para a tarefa-alvo

Machine learning

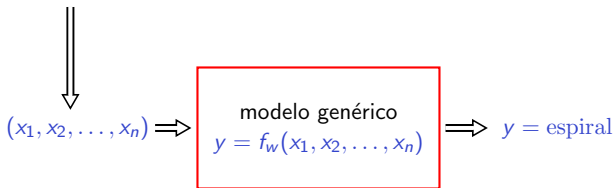
Deep learning

Machine Learning tradicional



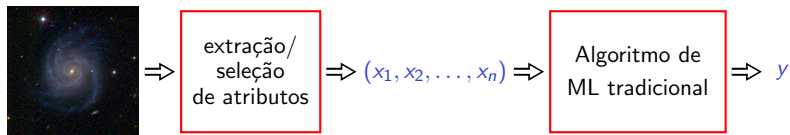
Atributos (características) do objeto

- concentração
- assimetria
- suavidade
- entropia
- espiralidade, etc

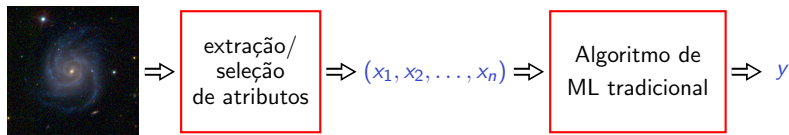


**Um grande desafio
para aplicação de machine learning
costumava ser a
extração / seleção de atributos**

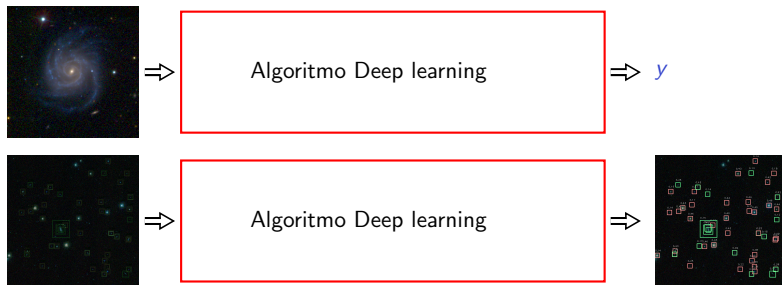
ML Tradicional



ML Tradicional



Deep learning: “aprende” representações!!



Redes neurais modernas

- incluem neurônios de diferentes tipos
- utilizam mecanismos de atenção
- possuem arquiteturas mais complexas
- possuem muito mais parâmetros
- processam diferentes tipos de dados, sem necessidade de extração de atributos
Imagem, vídeo, áudio, texto, combinação destes

Mais sobre isso na SESSÃO da TARDE !!!

Evolução de ML ao longo dos anos

quase nenhum dado
regras

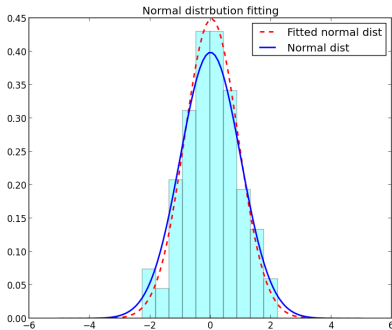
```
if age > 40:
    if is_home_owner:
        print("give a credit")
    else:
        if income > 5000:
            print("give a credit")
        else:
            print("to refuse")
else:
    if education == "university":
        print("...")
    else:
        print("...")
```

Evolução de ML ao longo dos anos

quase nenhum dado
regras

pouco dado

estimação paramétrica
modelos específicos



Evolução de ML ao longo dos anos

quase nenhum dado

regras

pouco dado

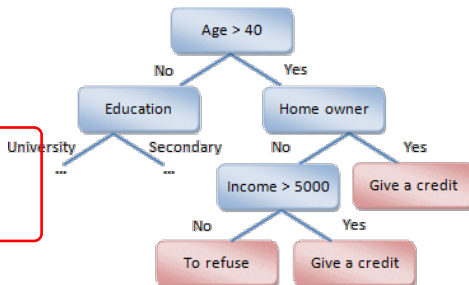
estimação paramétrica

modelos específicos

mais dados

indução de modelos (genéricos)

(algoritmos de ML)



Evolução de ML ao longo dos anos

quase nenhum dado

regras

pouco dado

estimação paramétrica

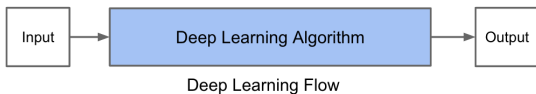
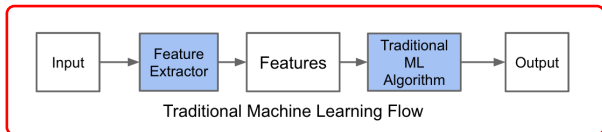
modelos específicos

mais dados

indução de modelos (genéricos)
(algoritmos de ML)

bastante dado

extração de features



Evolução de ML ao longo dos anos

quase nenhum dado

regras

pouco dado

estimação paramétrica

modelos específicos

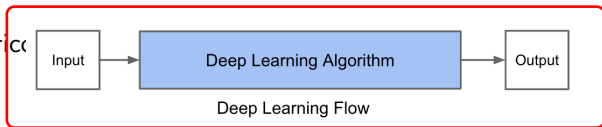


Traditional Machine Learning Flow

mais dados

indução de modelos (genéricos)

(algoritmos de ML)



Deep Learning Flow

bastante dado

extração de features

Big Data

aprendizado de representações

(deep learning, ponta-a-ponta)

Evolução de ML ao longo dos anos

quase nenhum dado

regras

pouco dado

estimação paramétrica

modelos específicos

mais dados

indução de modelos (genéricos)

(algoritmos de ML)

bastante dado

extração de features

Big Data

aprendizado de representações

(deep learning, ponta-a-ponta)

Baseado em dados

Resumindo ...

- A **tarefa computacional** precisa ser bem definida
- Precisamos de **exemplos de entrada-saída**
Não se deve subestimar o trabalho para a preparação adequada dos dados
- Faz sentido para problemas nos quais, dada a entrada, dizer qual é a saída é “simples”, enquanto dizer como mapear a entrada para a saída não é
- O **modelo genérico** deve ser **expressivo** suficiente (capaz de representar o mapeamento)
- **Validação do modelo**: funciona igualmente bem para instâncias não usadas no treinamento?

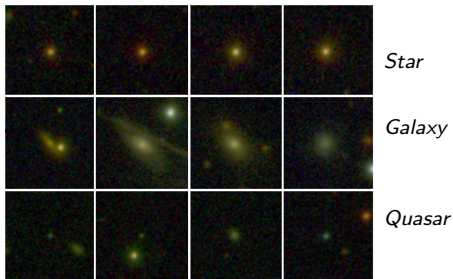
Materials para estudo

- Abu-Mostafa, Y. S., Lin, H.-T., and Magdon-Ismail, M. (2012). *Learning From Data*. AMLBook.
<https://work.caltech.edu/telecourse>
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press.
<https://www.deeplearningbook.org/>
- Michael A. Nielsen (2015). *Neural Networks and Deep Learning*. Determination Press.
<http://neuralnetworksanddeeplearning.com/>

Exemplos de processamento e análise de imagens

Reconhecimento de objetos astronômicos

Em colaboração
com IAG/USP



Spiral



Elliptic

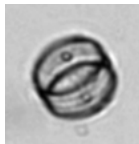
Exemplos de processamento e análise de imagens

Classificação de espécies

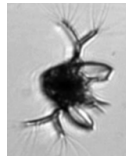
Plâncton – em colaboração com IO/USP



Cnidaria



Coscinodiscus



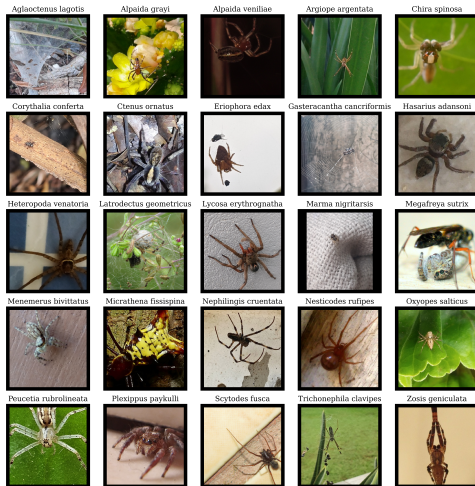
Penilia

Exemplos de processamento e análise de imagens

Classificação de espécies

Plâncton – em colaboração com IO/USP

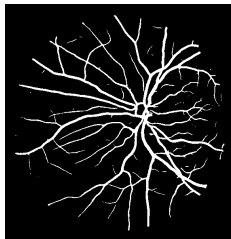
Aranhas



Exemplos de processamento e análise de imagens

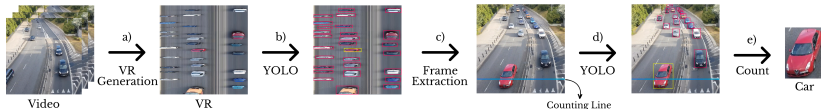
Análise de imagens de retina

Em colaboração
com FM/USP




Exemplos de processamento e análise de imagens

Detecção de objetos em imagens



Exemplos de processamento e análise de imagens

Reconhecimento de texto/manuscrito em imagens

Figure → 

Formula →
$$f(x) = \text{sgn}(g(x)), \quad (1)$$

where

$$g(x) = \left(\sum_{i=1}^n \alpha_i K(x_i, x) \right) + b, \quad (2)$$

Text →

components, with high interclass and low intraclass variation.

We use support vector machines (SVM) to classify the data vectors resulting from the Haar wavelet representation of the components. SVMs were proposed by Vapnik [25] and have yielded excellent results in various data classification tasks, including people detection [16], [14] and text classification [9]. Traditional training techniques for classifiers like multilayer perceptrons use empirical risk minimization and lack a solid mathematical justification. The SVM algorithm uses structural risk minimization to find the hyperplane that optimally separates two classes of objects. This is equivalent to minimizing a biased on generalization error. The optimal hyperplane is computed as a decision surface of the form

Fig. 3. It is very important to place geometric constraints on the location and scale of component detectors. Even though a detector may be triggered in a particular window examined, it might not be at the proper location. In this figure, the shadow of the person's head is detected with a higher score than the head itself. If we did not check for proper configuration and scale, component detectors like these would lead to false detections and false-alarmed detections.

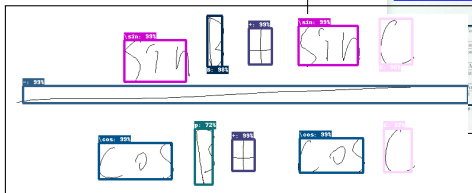
(1) K is one of many possible kernel functions. $y_i \in \{+1, -1\}$ is the class label of the data point x_i , and $\{x_i^+ \}_{i=1}^n$ is a subset of the training data set. The x_i^+ are called support vectors and are the points from the data set that define the separating hyperplane. Finally, the coefficients α_i and b are determined by solving a large-scale quadratic programming problem. One of the appealing characteristics of SVMs is that there are just two tunable parameters, C_{pos} and C_{neg} , which are penalty terms for positive and negative pattern misclassifications, respectively. The kernel function K that is used in the component classifiers is a quadratic polynomial and is $K(x, x') = \|x - x'\|^2$.

(2) In (2), $K(x, x') \in \{+1, -1\}$ is referred to as the binary class of the data point x which is being classified by the SVM. As (1) shows, the binary class of a data point is the sign of the raw output $g(x)$ of the SVM classifier. The raw output of an SVM classifier is the distance of a data point from the decision hyperplane. In general, the greater the magnitude of the raw output, the more likely a classified data point belongs to the binary class it is grouped into by the SVM classifier.

TABLE 1
Geometric Constraints Placed on Each Component

Component	Center	Size	Other Criteria	
Shoulders	Row: 3 Column: 32 ± 2	28 × 28	42 × 62	
Head	Row: 32 ± 2	42 × 28	48 × 48 Bottom Edge: Row: 120 ± 4	
Arm Extended	54 ± 5	46 ± 2	31 × 25	42 × 31
Arm Bent	46 ± 2	31 × 25	42 × 31	Top Edge: Row: 31 ± 3
Arm Extended	54 ± 5	17 ± 2	31 × 25	42 × 31
Arm Bent	46 ± 2	17 ± 2	31 × 25	42 × 31 Top Edge: Row: 31 ± 3

© 2004 IEEE. Published online in IEEE Xplore.





Formulário para postar dúvidas e comentários

<https://forms.gle/3CJhiHRyuQmEmqLs9>



Formulário para presença (tarde)

<https://forms.gle/RCHKwtBaQmiHgvMTA>



Processamento de Linguagem Natural

Thiago A. S. Pardo

Língua Natural

- Língua humana



- Em oposição às linguagens artificiais
 - Matemática, lógica, linguagens de programação



Processamento de Língua Natural (PLN)

- Habilitar a máquina a lidar com as línguas humanas
 - Ensinar a máquina a “ler e escrever”
- Diversas tarefas
 - Tradução automática
 - Sumarização de textos
 - Auxílio à escrita e à leitura
 - Análise de sentimentos
 - Reconhecimento e síntese de fala
 - Sistemas de perguntas e respostas e de diálogos (*chatbots*)
 - Etc.

Linguagem & língua: um erro histórico

- **Linguagem**: capacidade humana de comunicação e suas manifestações, de forma verbal ou não
 - Fala, gestos, música, dança, pintura, *um sorriso*
 - Envolve nosso aparato físico e mental/cognição
- **Língua**: código de comunicação utilizado por uma comunidade, com suas regras específicas
 - Português, Inglês, LIBRAS, etc.

PLN: um pouco de história

- Nascimento na 2ª guerra mundial
 - Tradução automática
- Possíveis nomes
 - *Computational Linguistics*
 - *Mechanolinguistics*
 - *Automatic Language Data Processing*
 - *Natural Language Processing*

PLN: um pouco de história

- Trajetória da Inteligência Artificial
 - Primeiros trabalhos □ grande expectativa □ resultados pobres □ desilusão e hibernação da área □ novos horizontes
 - Relatório da ALPAC (*Automatic Language Processing Advisory Committee*)

PLN: um pouco de história

- Globalização, internet, tecnologia da informação, Google e demais *big techs*
- Mais recentemente, **smartphones, redes sociais, modelos distribucionais, aprendizado profundo, big data** e ciência de dados, internet das coisas

PLN: um pouco de história



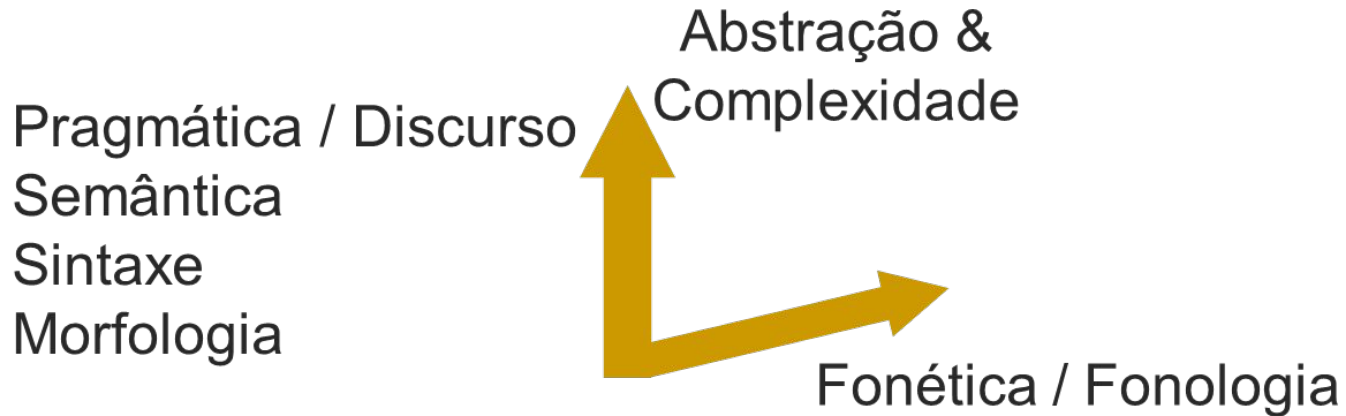
- Modelos distribucionais & aprendizado profundo: um casamento que abalou o mundo
 - A “numerificação” do texto: de símbolos para números (vetores)
 - Passamos a poder “fazer contas” com palavras:
rei-homem+mulher≈rainha
 - Redes neurais e o aprendizado de tarefas

PLN: um pouco de história

- Novos entendimentos sobre PLN
- Infraestrutura computacional em rápida evolução
- Novas demandas
- E os grandes desafios da língua

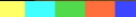
A estrutura da língua

- Níveis de representação e processamento linguístico





Um tour pelos produtos de PLN



Tradução simbólica

- [Apertium](#): a free and open-source platform for developing rule-based machine translation systems

The screenshot shows the Apertium web interface. At the top, the Apertium logo is displayed with the tagline "A free/open-source machine translation platform". A language dropdown menu is set to "English". Below the header, there are two text input areas. The left area has a language dropdown set to "Spanish" and contains the text "Finalmente caiu a ficha de que estava na hora de pendurar as chuteiras.". The right area has a language dropdown set to "Portuguese" and contains the translated text "Finalmente cayó la ficha de que estaba en la hora de colgar las chuteiras.". Below the input areas are two buttons: "Translate a document" and "Translate a webpage". On the right side, there are two checkboxes: "Mark unknown words" (unchecked) and "Instant translation" (checked). At the bottom of the page, there are links for "About", "Download", "Documentation", and "Contact". A footer message says "Spotted a mistake? Help us improve Apertium!".

Eliza (Jackson e Moulinier, 2002)

Eliza: como vai você hoje?

P: não muito bem

Eliza: por que você diz não muito bem?

P: eu estou com alergia

Eliza: interessante

P: eu fico com manchas vermelhas

Eliza: acho que não sei do que você está falando

P: manchas vermelhas no meu rosto

Eliza: seu rosto

P: isso mesmo

Eliza: é isso mesmo?

Eliza (Jackson e Moulinier, 2002)

P: você é um programa, não é?

Eliza: te agrada pensar que eu sou um programa?

P: com certeza

Eliza: eu percebo

P: ok

Eliza: um pouco antes você falou de seu rosto

P: e sobre o seu rosto?

Eliza: o que é que você realmente quer saber?

P: você é meu tipo de garota?

Eliza (Jackson e Moulinier, 2002)

Eliza: talvez em suas fantasias eu seja seu tipo de garota

P: em minhas fantasias você é loira

Eliza: o que te faz pensar que sou loira?

P: não sei

Eliza: você realmente não sabe?

P: não

Eliza: você está dizendo não apenas para ser negativo?

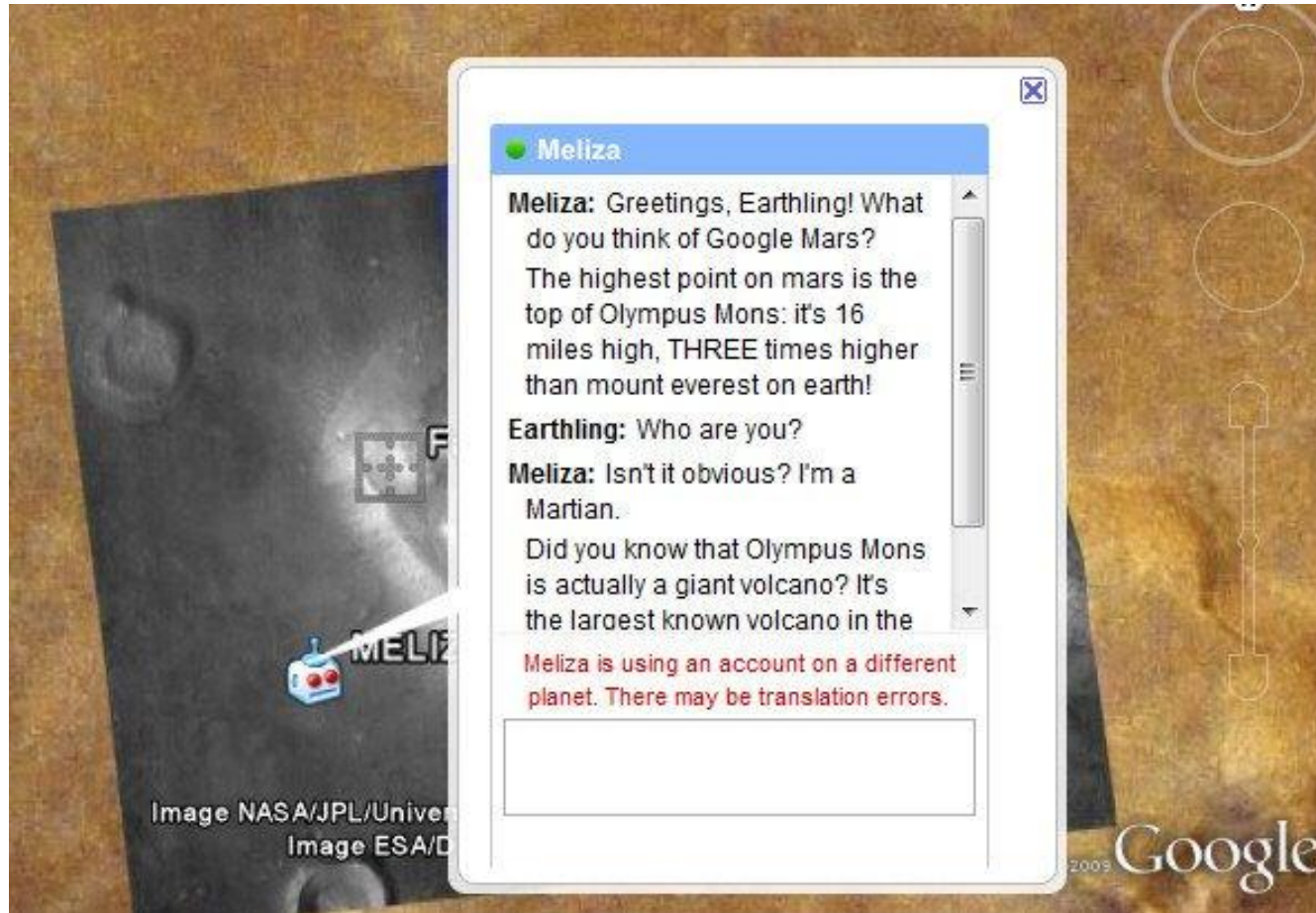
P: não

Eliza: você está sendo um pouco negativo

P: não, não estou

Eliza: por quanto tempo você não tem estado?

Eliza e Google Mars



START

START Natural Language Question Answering System

What South-American country has the largest population?



Ask Question >

==> What South-American country has the largest population?

Brazil has the highest population among countries in South America.

Brazil



Population:
204,259,812 (July 2015 est.)

Source: The World Factbook


Auxílio à escrita

The image shows the LanguageTool website and its browser extension interface. The website header includes the LanguageTool logo, the tagline "Seu corretor inteligente", and navigation links for "Aplicativos e extensões", "Premium -50%", "Para empresas", "API de revisão", "Blog", "Fazer login", and a "Registre-se" button. The main content area features the headline "Aprimore sua escrita instantaneamente." and a sub-headline "O corretor de gramática e estilo multilíngue do LanguageTool é usado por milhões de pessoas em todo o mundo". A green button labeled "Adicionar ao Chrome - É gratuito" is prominently displayed.

The browser extension interface is shown below, displaying the text "As menina comprou um carru." with a grammar error highlighted. The error message reads: "Grammar: Possível erro de concordância de número." and offers suggestions: "A menina", "As meninas", "As, menina", and "As de menina". A right-hand panel lists corrections: "As menina - Erro gramatical" and "carru - Erro ortográfico". The bottom status bar shows "Caracteres 27 Palavras 5 Reformulações 0/3" and a red notification badge with the number "2".

Auxílio à escrita

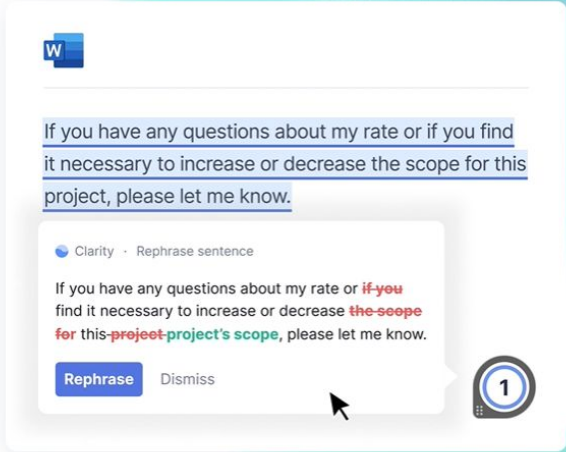
Go beyond words when generative AI comes to Grammarly

 **grammarly** Why Grammarly ▾ For Work ▾ For Education ▾ Compare Plans Tools & Guides ▾

Great Writing, Simplified

Compose bold, clear, mistake-free writing with Grammarly's new AI-powered desktop Windows app.

[Get Grammarly](#) It's free



W

If you have any questions about my rate or if you find it necessary to increase or decrease the scope for this project, please let me know.

Clarity · Rephrase sentence

If you have any questions about my rate or **if-you** find it necessary to increase or decrease **the-scope** for this-**project-project's** scope, please let me know.

Rephrase Dismiss

1

II

Sumarização automática

linguakit


Sobre Linguakit Legal PT


Ferramentas Linguísticas

 Análise completa


 Resumidor


Análise Linguístico

 Etiquetador morfossintático

 Analisador sintático

Análítica Textual

 Analisador de sentimento

 Extrator de palavras-chave

 Extrator multipalavra

 Extractor de tripletas

Resumidor

Todos los análisis lingüísticos en un solo lugar. El análisis completo informa del número de palabras y frases del texto, y de su tipología. Ofrece un resumen y su valoración. Además, te facilita las cinco palabras y multipalabras más relevantes del texto, las entidades más importantes que allí se mencionan, las palabras más frecuentes y el contexto en el que aparece la palabra clave escogida.

3946/5000

BELO HORIZONTE Trabalhadores entraram em greve e mais de 65 mil pessoas foram às ruas na Grécia na maior demonstração pública de indignação contra a tragédia ferroviária ocorrida no país na última semana.

Em 28 de fevereiro, um acidente de trens próximo à cidade de Larissa — a cerca de 350 km ao norte da capital, Atenas— matou 57 pessoas. Os trens percorreram o mesmo trilho por quilômetros, sem que ninguém percebesse, e colidiram frontalmente. Das 350 pessoas que estavam nos veículos, muitas eram jovens e estudantes, e 14 continuam hospitalizadas.

Nesta quarta (8), foram registrados confrontos com a polícia em Atenas, onde ocorreu o maior protesto, com cerca de 40 mil pessoas. Manifestantes lançaram coquetéis molotov em frente ao Parlamento e incendiaram uma van e caixotes de madeira.

"Nós sentimos raiva porque o governo não fez nada por todos esses jovens. O transporte público está uma bagunça", disse Níkomathi Vathi, 19, à agência de notícia Reuters.

Em Tessalônica, segunda maior cidade do país, a manifestação se concentrou do lado de fora da estação ferroviária. Lá, um grupo chegou a arremessar pedras contra um edifício público.

Servidores fizeram uma paralisação por 24 horas, assim como professores do ensino fundamental, médicos, motoristas de ônibus, maquinistas dos metrô e capitães de embarcações utilizadas no transporte entre as ilhas gregas.

"Estou aqui para expressar minha homenagem aos mortos e minha indignação e frustração", disse em Atenas Niki Siouta, engenheiro civil de 54 anos, à agência de notícias AFP. "Este governo tem que sair."

Trabalhadores da área ferroviária, que já estavam em situação de greve desde a última quinta-feira (2), afirmam que o acidente da última semana denuncia o estado em ruínas da malha do país. Eles

Porcentagem de texto a resumir

 Resumidor

Trabalhadores da área ferroviária, que já estavam em situação de greve desde a última quinta-feira (2), afirmam que o acidente da última semana denuncia o estado em ruínas da malha do país. Eles apontam negligência do governo, que teria falhado ao ignorar os avisos das graves falhas técnicas da linha feitos muito antes da tragédia. "Nós, motoristas, apresentamos queixas sobre essas questões, fizemos greve para isso, alertamos, protestamos", disse o representante Kostas Genidounias à Reuters. "Eles disseram que estávamos mentindo, que éramos caluniosos e tínhamos outros interesses. No final, [a tragédia] mostrou que os trabalhadores estavam certos". O subinvestimento nas linhas é também legado da longa crise de dívida do país europeu. Nas manifestações, muitos gregos reclamaram da decadência dos serviços públicos desde que planos de austeridade foram impostos pelos credores da Grécia. Horas após o acidente, o primeiro-ministro grego Kyriakos Mitsotakis, cujo mandato termina em julho deste ano, afirmou que o ocorrido foi um "trágico erro humano". A fala foi duramente criticada. O premiê pediu desculpas às famílias das vítimas pelo acidente no domingo (5), cinco dias após a tragédia. Também solicitou ajuda à União Europeia, que deve enviar especialistas de sua agência ferroviária ao país ainda nesta semana. O governo, de orientação conservadora, anunciou a suspensão do serviço ferroviário enquanto analisa a segurança das linhas. "Nenhum trem partirá novamente se não tivermos segurança no máximo nível possível", disse em entrevista coletiva o novo ministro dos Transportes, George Gerapetrite. Ele assumiu a chefia da pasta na semana passada após a renúncia de Kostas Karamanlis e afirmou que vai esclarecer as causas do acidente. O novo titular da pasta também se comprometeu a investir fundos na atualização da infraestrutura e na contratação de funcionários. Em 2017, durante a crise que assolou o país, a Grécia vendeu suas operações ferroviárias, hoje chamadas Hellenic Train, para a Ferrovie Dello Stato, da Itália. A empresa privada, também apontada como culpada pelos manifestantes, respondeu às acusações dizendo que a responsabilidade pela manutenção da rede seria da empresa pública grega OSE. O chefe da estação de Larissa, que reconheceu sua responsabilidade no acidente, está em prisão provisória e pode ser condenado a prisão perpétua.

ORIGINAL

BELO HORIZONTE Trabalhadores entraram em greve e mais de 65 mil pessoas foram às ruas na Grécia na maior demonstração pública de indignação contra a tragédia ferroviária ocorrida no país na última semana.

Em 28 de fevereiro, um acidente de trens próximo à cidade de Larissa —a cerca de 350 km ao norte da capital, Atenas— matou 57 pessoas. Os trens percorreram o mesmo trilho por quilômetros, sem que ninguém percebesse, e colidiram frontalmente. Das 350 pessoas que estavam nos veículos, muitas eram jovens e estudantes, e 14 continuam hospitalizadas.

Nesta quarta (8), foram registrados confrontos com a polícia em Atenas, onde ocorreu o maior protesto, com cerca de 40 mil pessoas. Manifestantes lançaram coquetéis molotov em frente ao Parlamento e incendiaram uma van e caixotes de madeira.

...

ORIGINAL

BELO HORIZONTE Trabalhadores entraram em greve e mais de 65 mil pessoas foram às ruas na Grécia na maior demonstração pública de indignação contra a tragédia ferroviária ocorrida no país.

Em 28 de fevereiro, um trem colidiu frontalmente com um trem de passageiros de 350 km ao norte da capital. O acidente ocorreu no mesmo trilho por quase duas horas. Das 350 pessoas que estavam no trem, 14 continuam em estado crítico. Nesta quarta (8), foram realizadas manifestações em várias cidades. Ocorreu o maior protesto em 10 anos, com coquetéis molotov em frente ao Parlamento de Atenas e queimada de pneus e madeira.

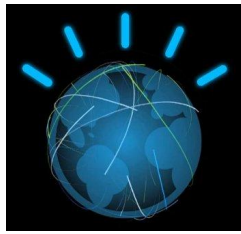
RESUMO

Trabalhadores da área ferroviária, que já estavam em situação de greve desde a última quinta-feira (2), afirmam que o acidente da última semana denuncia o estado em ruínas da malha do país. Eles apontam a negligência do governo, que teria falhado ao ignorar os avisos das graves falhas técnicas da linha feitos muito antes da tragédia.

...

...

Watson (IBM)



- Venceu os melhores participantes humanos no show de perguntas e respostas Jeopardy!
 - *“more than 100 different techniques are used to analyze natural language, identify sources, find and generate hypotheses, find and score evidence, and merge and rank hypotheses”*
 - *“sources of information include encyclopedias, dictionaries, thesauri, newswire articles, and literary works. Watson also used databases, taxonomies, and ontologies. Specifically, DBPedia, WordNet, and Yago were used”*

Assistentes com processamento de fala

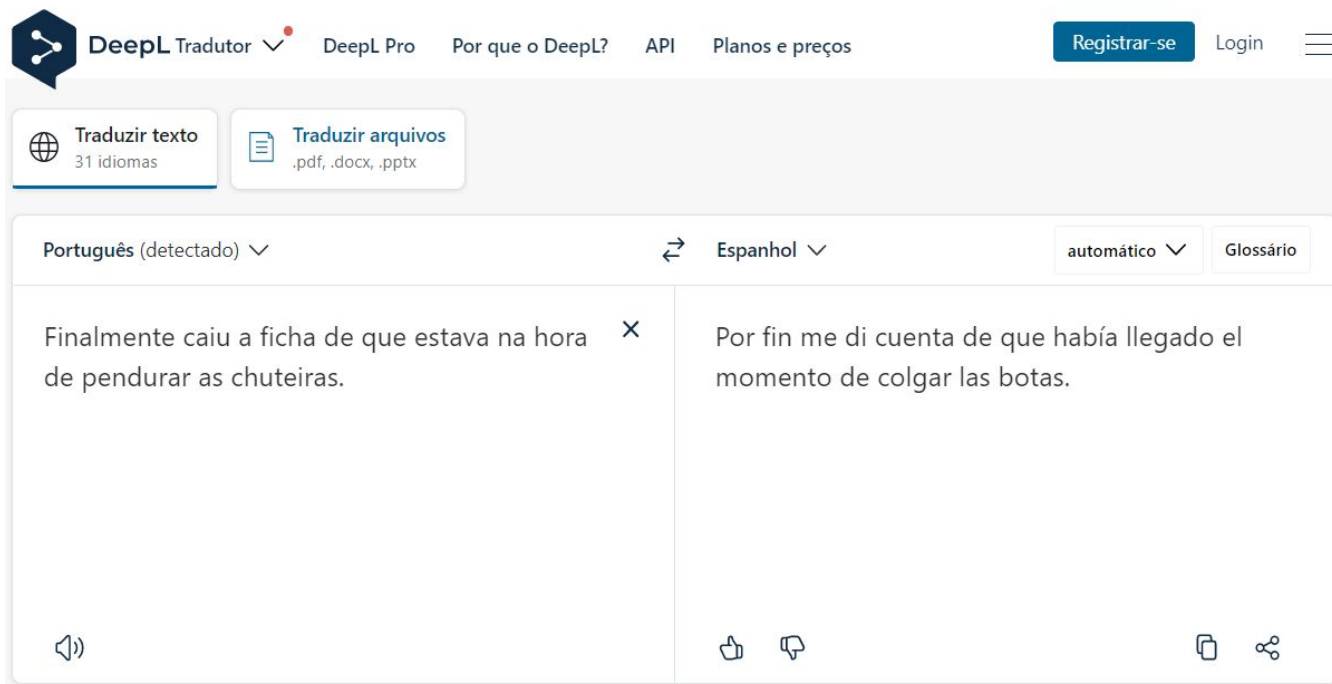


Hey Siri



Tradução “subsimbólica”

- [Google Tradutor](#), [DeepL](#), etc.



The screenshot displays the DeepL Translator web interface. At the top, the navigation bar includes the DeepL logo, the text "DeepL Tradutor" with a dropdown arrow, and links for "DeepL Pro", "Por que o DeepL?", "API", and "Planos e preços". On the right side of the navigation bar, there are buttons for "Registrar-se" (in a dark blue box) and "Login", along with a hamburger menu icon.

Below the navigation bar, there are two main options: "Traduzir texto" (with a globe icon and "31 idiomas") and "Traduzir arquivos" (with a document icon and ".pdf, .docx, .pptx").

The main translation area shows the source language as "Português (detectado)" and the target language as "Espanhol". There are dropdown menus for "automático" and "Glossário".

The text being translated is:
Portuguese: "Finalmente caiu a ficha de que estava na hora de pendurar as chuteiras."
Spanish: "Por fin me di cuenta de que había llegado el momento de colgar las botas."

At the bottom of the translation area, there are icons for audio playback (speaker icon), thumbs up/down for feedback, and copy/share icons.

Decodificação

- **How Revolutionary Tools Cracked a 1700s Code**
 - *Now a team of Swedish and American linguists has applied statistics-based translation techniques to crack one of the most stubborn of codes: the Copiale Cipher, a hand-lettered 105-page manuscript that appears to date from the late 18th century. They described their work at a meeting of the Association for Computational Linguistics in Portland, Ore.*
 - *Kevin Knight, a computer scientist at the Information Sciences Institute at the University of Southern California, collaborated with Beata Megyesi and Christiane Schaefer of Uppsala University in Sweden to decipher the first 16 pages. They turn out to be a detailed description of a ritual from a secret society that apparently had a fascination with eye surgery and ophthalmology.*



Tradução “portátil”

- *Novos Google Pixel Buds mostram por que a tecnologia é maravilhosa*



Análise de sentimentos

GOTIT AI

[🏠](#) [Soluções](#) [Preços](#) [Documentação](#) [Contato](#) [Crie sua conta grátis!](#) [Acessar](#)

Experimente agora mesmo

Esse aplicativo é muito bom! Só não gostei que ele é pago.

Linguagem Texto Original:

(Detectar Automaticamente) ▾

Analisar

Texto analisado JSON

[Pegue sua API Key Gratis!](#)

Linguagem Detectada:
PtBr (Português Brasileiro)

Sentimento	58,00% POSITIVE
------------	-----------------

Emoções	SADNESS: 42,00% JOY: 100,00% FEAR: 0,00% DISGUST: 0,00% ANGER: 0,00%
---------	--

Detecção de notícias falsas

FakeCheck

Detector de Fake News

Como funciona?

Copie o texto de uma notícia, cole na caixa abaixo e clique em "Enviar". O sistema irá processar o texto para identificar características de escrita, como palavras usadas ou classes gramaticais mais essas características em um modelo de aprendizado de máquina que classificará a notícia como verdadeira ou falsa. Para mais informações sobre como o sistema funciona e sua taxa de precisão, você também pode utilizar o nosso [bot do WhatsApp](#).

Resultado:

Essa notícia pode ser falsa. 😨 Busque fontes confiáveis.

Insira o texto completo da notícia! O texto deve ter pelo menos 100 palavras. O sistema não funciona corretamente com apenas partes de notícias.

Suco de limão e caldo de cana curam dengue

Esprema 200 ml do limão taiti. Dá mais ou menos um copo de limão. Beba todo o copo de uma vez só. Basta apenas uma dose. A carga viral vai acabar, bem como todos os

Modelagem de língua

[GPT-3: o mais poderoso sistema de inteligência artificial já criado](#)

- *O desenvolvimento das tecnologias de inteligência artificial se tornou cada vez mais sofisticado ao longo dos últimos anos. Computadores com maior poder de processamento vêm adquirindo repertórios abrangentes, dominando diversas áreas do conhecimento e capacidades diferentes. Um exemplo deste desenvolvimento é o GPT-3: sistema desenvolvido pela OpenAI baseado em machine learning (aprendizado de máquinas) que possui a capacidade de escrever diversos tipos de gêneros textuais com grande verossimilhança a qualquer trabalho executado por um humano... Da poesia a literatura, do jornalismo ao direito: as possibilidades de criação original por parte do GPT-3 demonstram que seu poder produção é vasto.*

Modelagem de língua “acessível”

- [ChatGPT](#): o “empoderamento” dos *chatbots*



E muito mais

- Muitos objetivos

- Estudo computacional da evolução da língua
- Investigação e modelagem computacional de mecanismos linguísticos cognitivos
- Criação e formalização de teorias e modelos linguístico-computacionais
- Aplicações variadas em educação, saúde, direito, etc.
- Etc.



E muito mais

- Também pode envolver PLN
 - Recuperação de informação
 - Sistemas de recomendação
 - Detecção de spam
 - Etc.
- Várias plataformas
 - Desktop, web, *mobile*, etc.





Paradigmas de trabalho em PLN

Classificação das abordagens

Cr�terios	Paradigmas
Uso de conhecimento lingu�stico	Superficial, profundo e h�brido
Representa�o do conhecimento	Simb�lico, sub/n�o-simb�lico e h�brido
Obten�o do conhecimento	Manual, autom�tica e h�brida

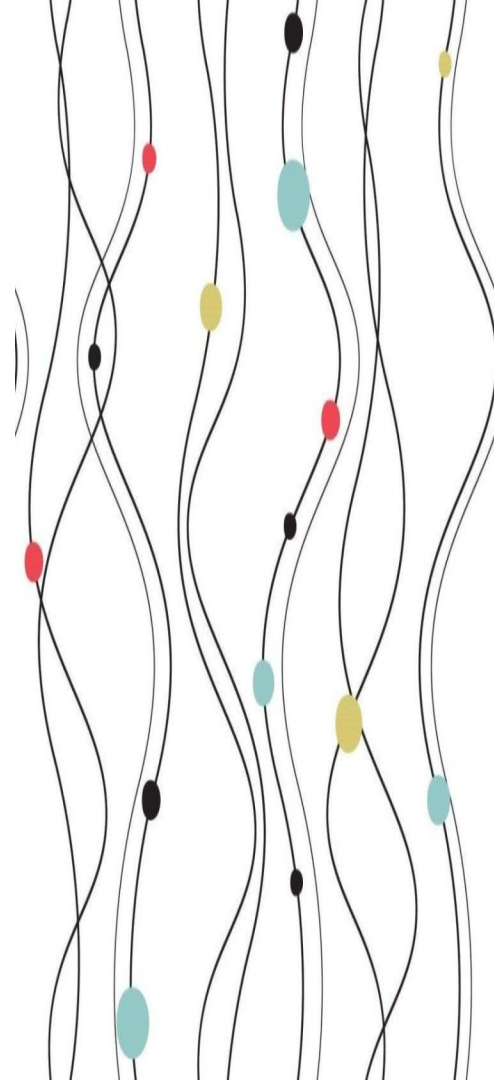
Os movimentos em PLN

Abordagens simbólicas

⇒ Abordagens estatísticas

⇒ Abordagens neurais

... e as estratégias se misturam!



Racionalismo

- 1960-1985
 - Crença de que parte significativa do conhecimento humano não vem dos sentidos, mas é herdada geneticamente
- Noam Chomsky e o *gerativismo*
 - Linguagem inata
- PLN: sistemas com muito conhecimento manualmente fornecido e com mecanismos de inferência



Empirismo

- 1920-1960
 - Mente não vem com princípios e procedimentos pré-determinados
 - Mas vem com operações gerais de associação, reconhecimento de padrões e generalizações
 - Importância do estímulo sensorial para o aprendizado da língua



- Retorno triunfal na atualidade
 - Aprendizado de máquina
 - Conjuntos de dados textuais: os córpus (ou *corpus* e *corpora*)

Exemplo: revisão gramatical

O carros quebrou

Exemplo: revisão gramatical

O carros quebrou

Léxico:

- o: artigo (masc, sing),
pronome, ...
- carros: substantivo (masc,
sing)
- ...

& regras de desambiguação: se
“o” antes de substantivo, então
“artigo”

O_art carros_subst quebrou_verbo

Exemplo: revisão gramatical

O carros quebrou

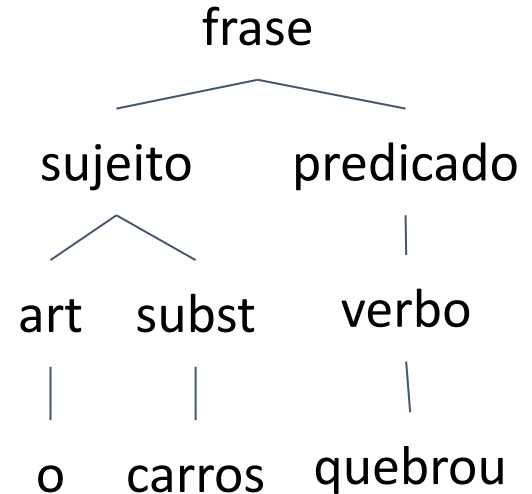
Léxico:

- o: artigo (masc, sing),
pronome, ...
- carros: substantivo (masc,
sing)
- ...

& regras de desambiguação: se
"o" antes de substantivo, então
"artigo"

O_art carros_subst quebrou_verbo

frase → sujeito, predicado
sujeito → artigo, substantivo
...



Exemplo: revisão gramatical

O carros quebrou

Léxico:

- o: artigo (masc, sing), pronome, ...
- carros: substantivo (masc, sing)
- ...

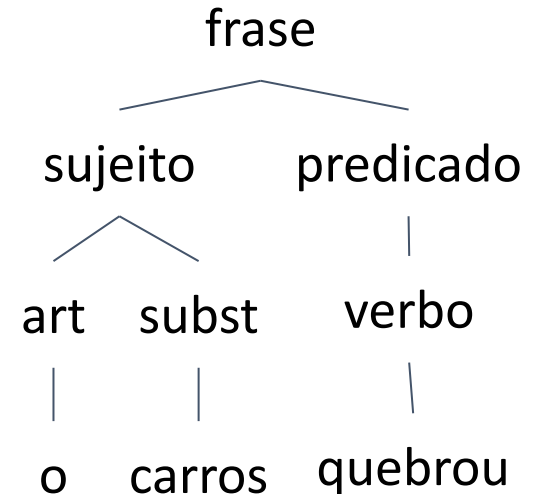
& regras de desambiguação: se "o" antes de substantivo, então "artigo"

O_art carros_subst quebrou_verbo

frase → sujeito, predicado
sujeito → artigo, substantivo
...

Sugerir alteração do artigo
ou do substantivo:
o carro ou os carros

Regra de correção: artigo e substantivo devem concordar dentro do sujeito



Exemplo: filtragem de spams

Golpe da herança nigeriana ganha fôlego em rede social

“Antes da morte dele, no dia 12 de dezembro de 2007, em um hospital de Abidjan, ele me chamou e pediu que eu guardasse US\$ 6,5 milhões em uma conta bancária também em meu nome. Por causa de uma disputa com seu sócio, eu devo encontrar um parceiro de negócios no país de minha escolha, para onde eu possa transferir o valor e usá-lo para investimento”

Exemplo: filtragem de spams

Golpe da herança nigeriana ganha fôlego em rede social

“Antes da morte dele, no dia 12 de dezembro de 2007, em um hospital de Abidjan, ele me chamou e pediu que eu guardasse US\$ 6,5 milhões em uma conta bancária também em meu nome. Por causa de uma disputa com seu sócio, eu devo encontrar um parceiro de negócios no país de minha escolha, para onde eu possa transferir o valor e usá-lo para investimento”

$P(\text{é spam} | \text{mensagem}) = ?$

$P(\text{não é spam} | \text{mensagem}) = ?$

O que é mais provável?

Exemplo: filtragem de spams

Golpe da herança nigeriana ganha fôlego em rede social

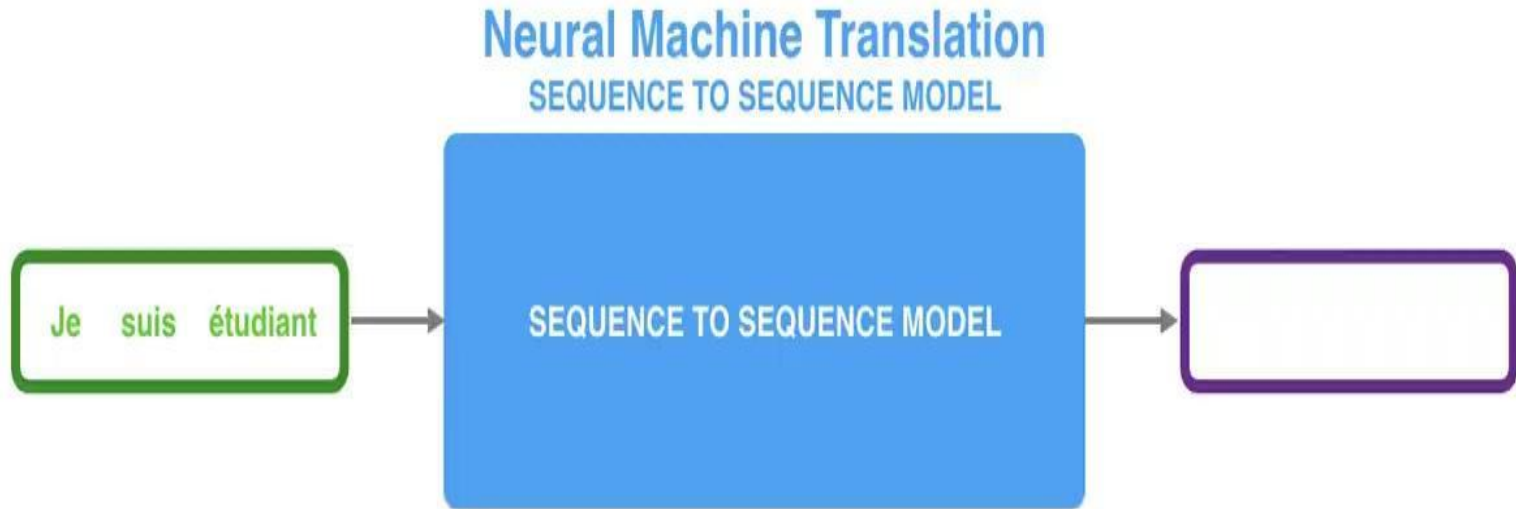
“Antes da morte dele, no dia 12 de dezembro de 2007, em um hospital de Abidjan, ele me chamou e pediu que eu guardasse US\$ 6,5 milhões em uma conta bancária também em meu nome. Por causa de uma disputa com seu sócio, eu devo encontrar um parceiro de negócios no país de minha escolha, para onde eu possa transferir o valor e usá-lo para investimento”

$P(\text{é spam} | \text{mensagem}) \Rightarrow P(\text{“conta”} | \text{é spam}) * P(\text{“país”} | \text{é spam}) * \dots$

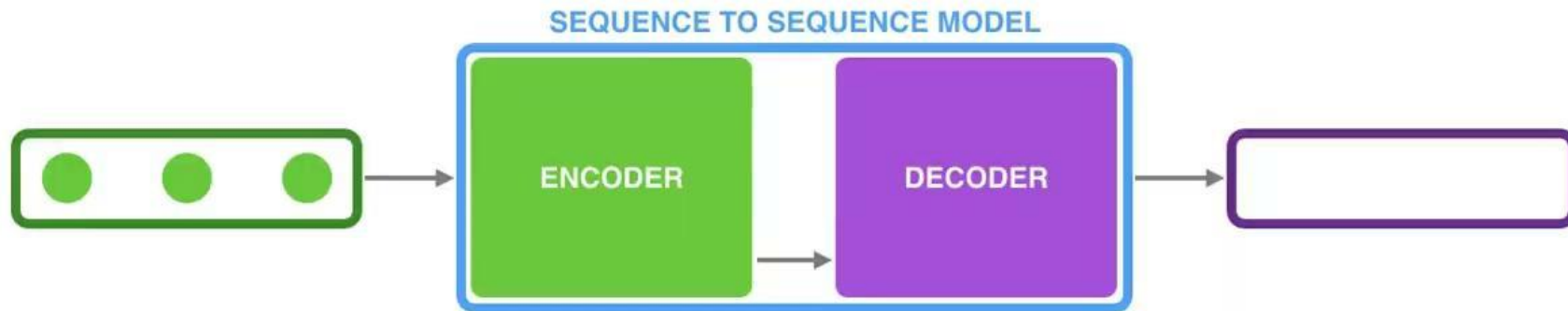
$P(\text{não é spam} | \text{mensagem}) \Rightarrow P(\text{“conta”} | \text{não é spam}) * P(\text{“país”} | \text{não é spam}) * \dots$

...

Exemplo: tradução automática



Exemplo: tradução automática





Tendências

Tópicos de pesquisa

- E-mails, mensagens, redes sociais e *User Generated Content* (UGC)
- Mineração de opiniões
- Assistentes/agentes inteligentes
- Abordagens multimodais

⇒ entrada da indústria no cenário

Ambições

- Aplicações *cross-language*
 - Apesar de possíveis limitações de PLN
- Robustez, escalabilidade e independência de língua
 - “Deve funcionar para qualquer coisa na web”

Ambições

- Atenção aos **minoritários**



ILLC-NLP 2024

First Workshop on NLP for Indigenous Languages of Lusophone Countries

March 12, 2024

Os desafios do português



BRASIL



GUINÉ-BISSAU



MOÇAMBIQUE



MACAU



ANGOLA



CABO VERDE



PORTUGAL



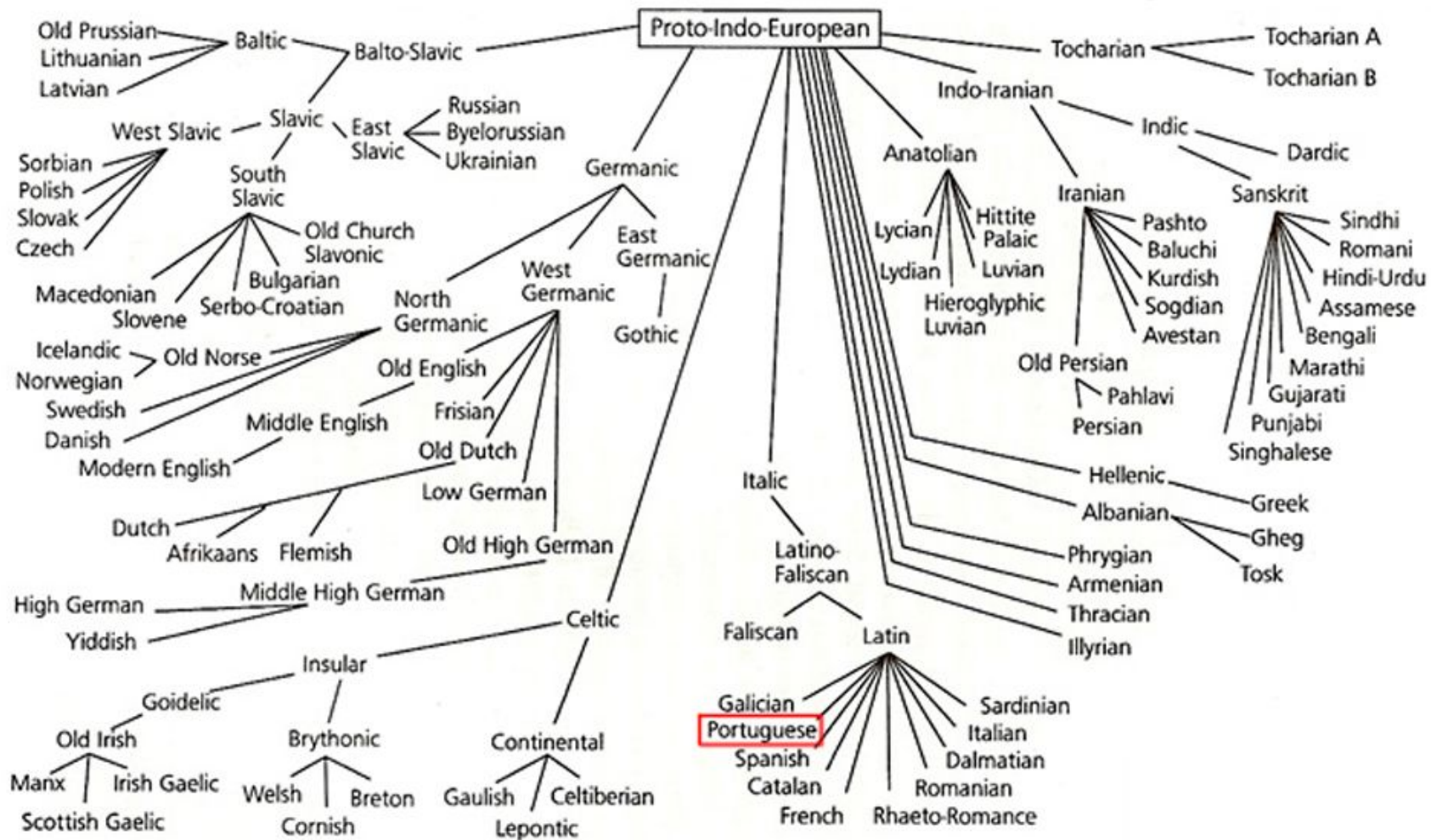
SÃO TOMÉ E PRÍNCIPE



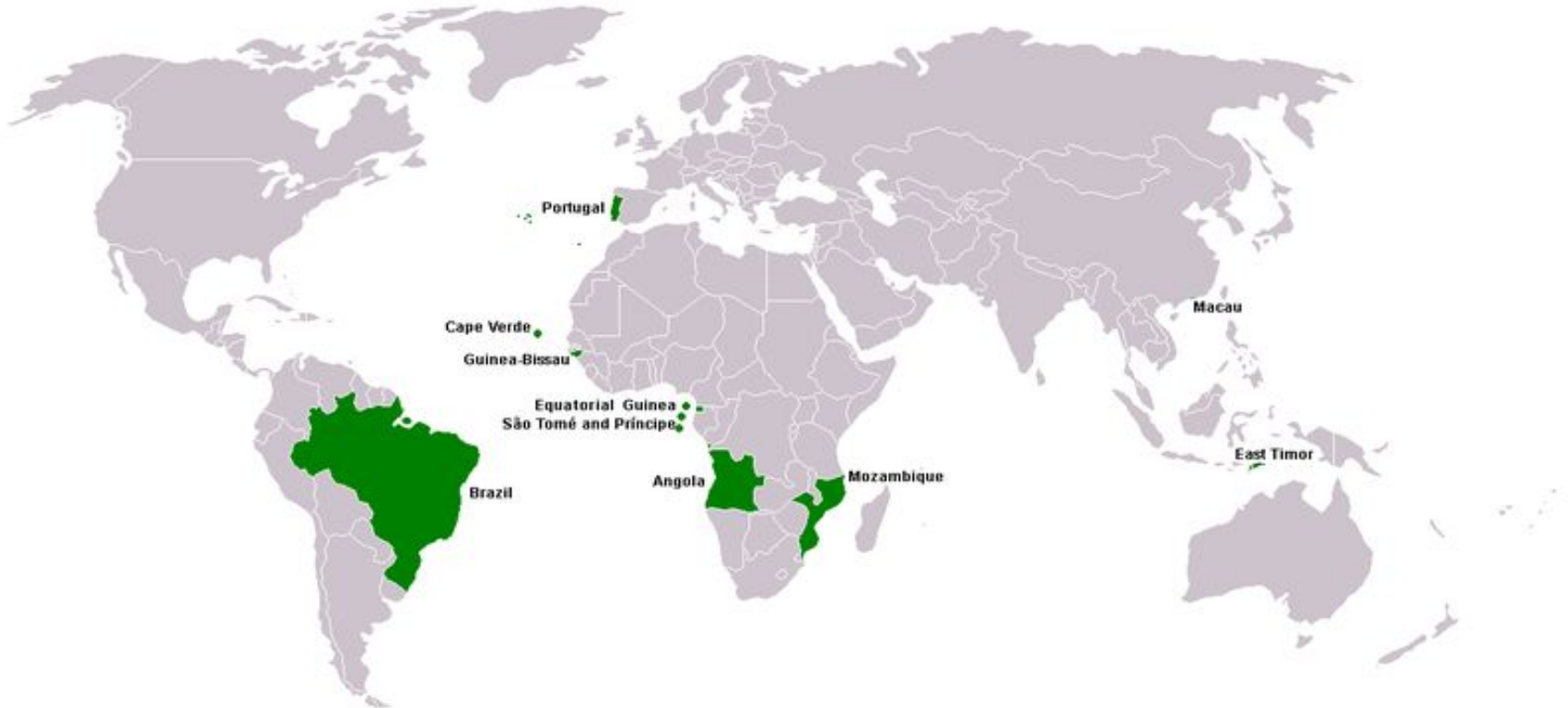
TIMOR LESTE



GUINÉ EQUATORIAL



Falado em 10 territórios



Dados sobre a língua

- 6ª língua mais falada no mundo
- [Instituto Internacional da Língua Portuguesa \(IILP\)](#)
 - Divulgação e promoção da língua portuguesa
- Grande número de vocábulos: segundo Academia Brasileira de Letras, mais de 356.000 palavras

Desafios para processamento

- **Variações** nos países em que é falado
 - Pronúncias variadas
 - Ortografias variadas, apesar do acordo ortográfico
 - Diferenças nas construções sintáticas mais usuais
 - Sentidos diferentes de palavras
 - “Perfis de usuários” diferentes: hábitos, expectativas e cultura variados

Livro Branco da Língua Portuguesa

([Branco et al., 2012](#))

Em Portugal, a divisão geográfica dos dialetos [13] distingue os dialetos do Centro-Sul, os dialetos do Norte e os dialetos das ilhas atlânticas. Os dialetos do Norte podem ser identificados pela ausência da distinção fonológica entre /b/ e /v/, com prevalência do /b/, pela preservação de antigos ditongos, e pela existência de fricativas ápicoalveolares. As diferenças entre estes dialetos encontram-se sobretudo ao nível da fonética e fonologia e ao nível lexical, sendo todos eles mutuamente compreensíveis de forma imediata (possivelmente com a exceção de alguns dialetos das ilhas).

A situação das variedades africanas do português é variada: enquanto em Angola e Moçambique o número de falantes de português tem vindo a aumentar desde a independência destes países, noutras casos, como São Tomé e Príncipe ou Cabo Verde, em muitas circunstâncias utiliza-se amplamente o crioulo e o português é adquirido como língua segunda.

Quanto ao Brasil, dada a dimensão geográfica deste país, não é viável apresentar aqui as suas variedades linguísticas. Por razões geográficas, políticas e sociais, não é possível falar de uma variedade padrão do português do Brasil. Os especialistas tendem a mencionar “normas urbanas cultas”.

Desafios para processamento

- [Zuchini \(2011\)](#)

- Elevado número de vocábulos existentes
- Elevado número de sinônimos entre vocábulos
- Elevado número de flexões verbais
- Diversas possibilidades de construção sintática
- Elevado número de flexões em gênero, número e grau
- Grande número de exceções para praticamente todas as regras

PLN para o português

Perspectiva histórica

- Fortalecimento a partir da década de 1990
- Alinhamento aos tópicos de trabalhos para outras línguas (inglês, principalmente), mesmo que com algum atraso
- Valorização do tratamento “individualizado” do português



DISPONIBILIZAR REUTILIZAR IMPULSIONAR

"O português está para ficar"



the latest	one way	flight	from	dallas	
NP/N	N/N	N	$N \setminus N / NP$	NP	$N \setminus N / NP$
$\lambda f. \arg \max(f, y. depart_time(y))$	$\lambda f. \lambda x. f(x)$ $\lambda one_way(x)$	$\lambda x. flight(x)$	$\lambda y. \lambda f. \lambda x. f(x)$ $\wedge from(x, y)$	dallas	$\lambda y. \lambda f. \lambda x. f(x)$
	N		$N \setminus N$		
	$\lambda x. one_way(x) \wedge flight(x)$		$\lambda f. \lambda x. f(x) \wedge from(x, dallas)$		$\lambda f. f(x)$
			$N \setminus N$		
			$\lambda f. \lambda x. f(x) \wedge from(x, dallas) \wedge to(x, v)$		
			N		
			$\lambda x. one_way(x) \wedge flight(x) \wedge from(x, dallas) \wedge to(x, washing)$		



Serviços e dados para investigadores, inovadores, estudantes e profissionais da **linguagem**.

<https://www.linguateca.pt/>

Apresentação

- [Historial](#)
- [Produção escrita](#)
- [Encontros da Linguateca](#)
- [Bastidores](#)
- [Estatísticas de acesso](#)
- [Como colaborar](#)

[Acesso a recursos](#)

[Avaliação conjunta](#)

[Catálogo de recursos](#)

[Catálogo de ferramentas](#)

[Catálogo de atores](#)

[Catálogo de publicações](#)

[Informação interessante](#)

Linguateca

O objectivo da Linguateca, um **centro de recursos -- distribuído -- para o processamento computacional da língua portuguesa**, é servir a comunidade que se dedica ao processamento da nossa língua. Em particular,

- facilitar o acesso aos recursos já existentes, através do desenvolvimento de serviços de acesso na rede, e mantendo um portal com informação útil,
- desenvolver, de forma harmoniosa, em colaboração com os interessados, os recursos considerados mais prementes,
- organizar avaliações conjuntas que envolvam a comunidade como um todo.



Formulário para presença (tarde)

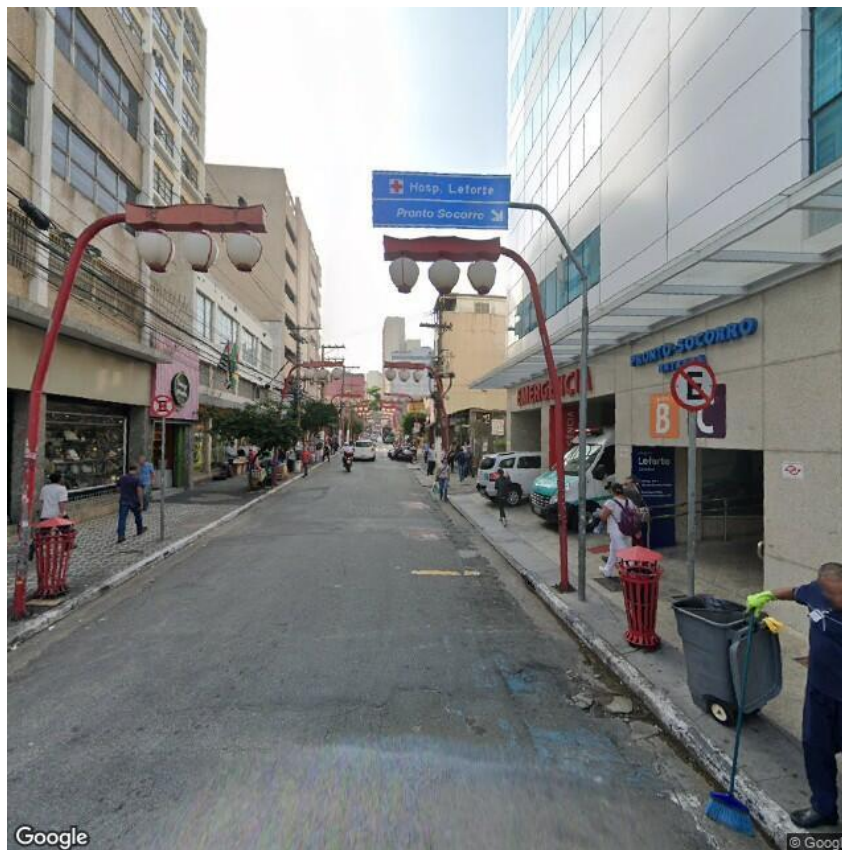
<https://forms.gle/RCHKwtBaQmiHgvMTA>



Visão Computacional

Roberto Hirata Jr.

Em 2003 um estudante me disse: “Eu queria...”



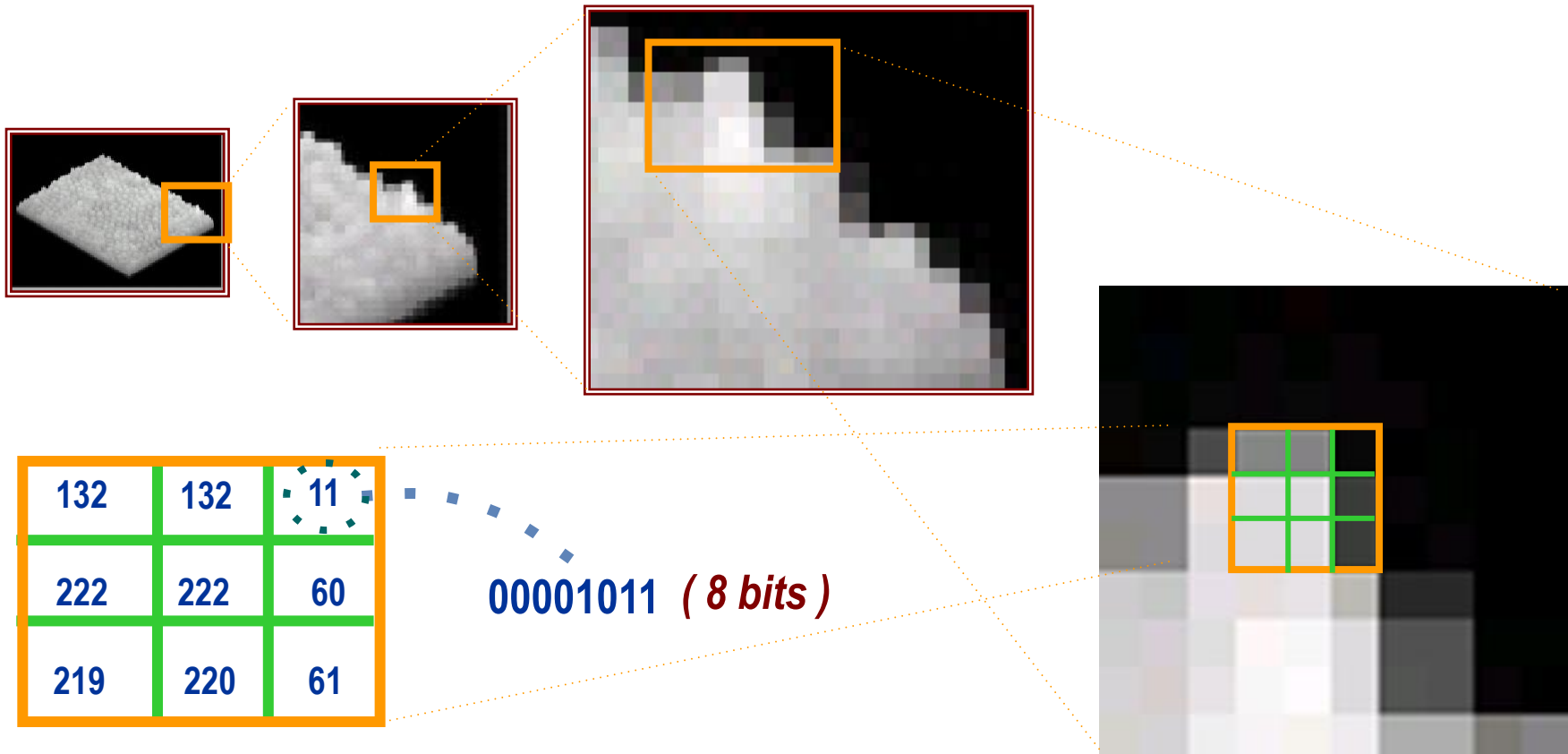
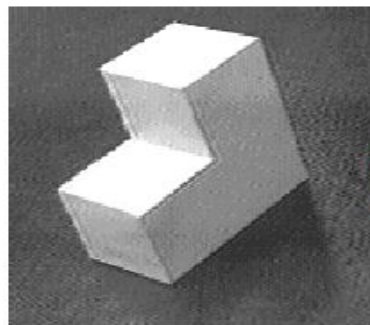


Imagem digital

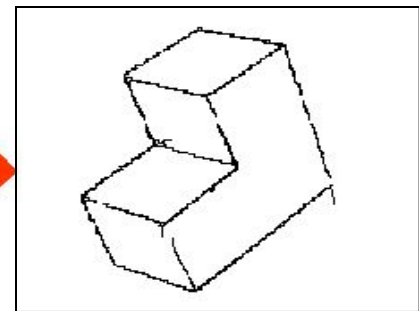
Pixel □ Picture Element

Processos

Processamento de Imagens



IMAGEM

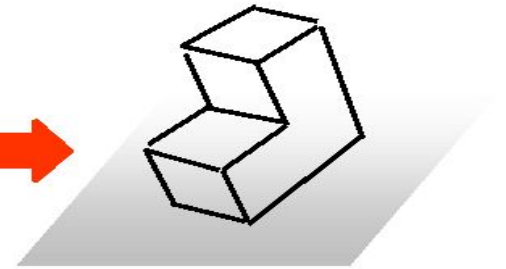


IMAGEM

Processos

- *Computação Gráfica*

bloco
poliédrico
8 faces
pos. observador
pos. iluminação
fundo
etc

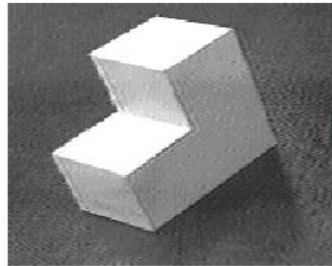


DESCRIÇÃO
FORMAL

IMAGEM
(sintética)

Processos

- *Visão Computacional*



IMAGEM

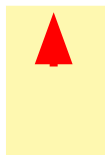


bloco
poliédrico
8 faces
pos. observador
pos. iluminação
fundo
etc

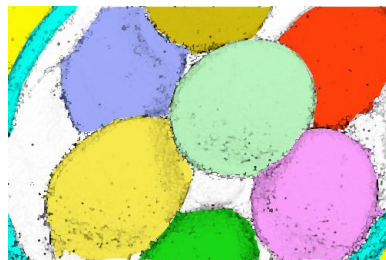
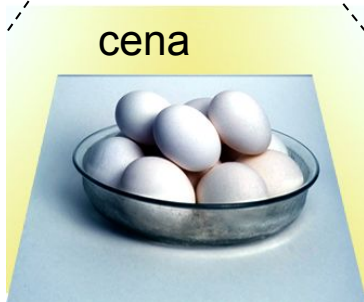
DESCRIÇÃO
FORMAL

Algumas tarefas em Visão Computacional

- Descrição de cenas
- Detecção de objetos (animais, aviões, bicicletas, carros, drones, faces, pessoas etc)
- Identificação/reconhecimento de caracteres impressos e manuscritos
- Identificação/reconhecimento de impressão digital
- Identificação/reconhecimento de objetos
- Rastreamento de objetos em movimento
- Vigilância
- etc



câmera



Visão Computacional

- Ovos
- Tigela
- Borda
- Fundo

Visão Computacional

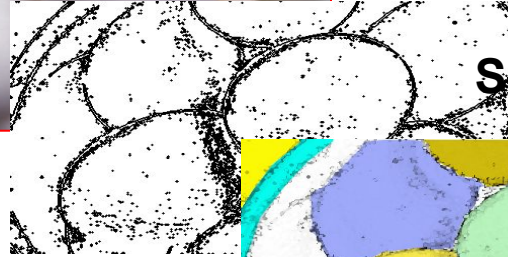
IMAGEAMENTO



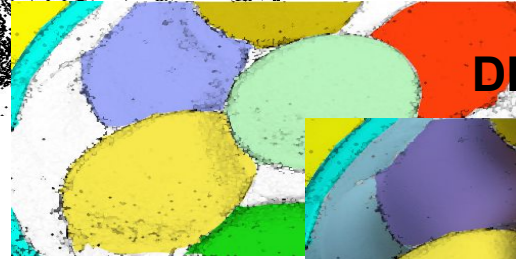
PRÉ-PROCESSAMENTO



SEGMENTAÇÃO



DESCRIÇÃO



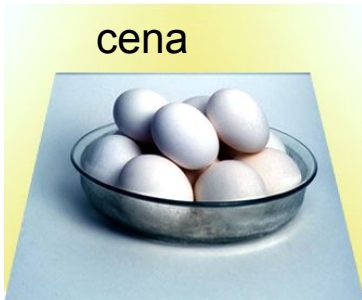
RECONHECIMENTO



câmera



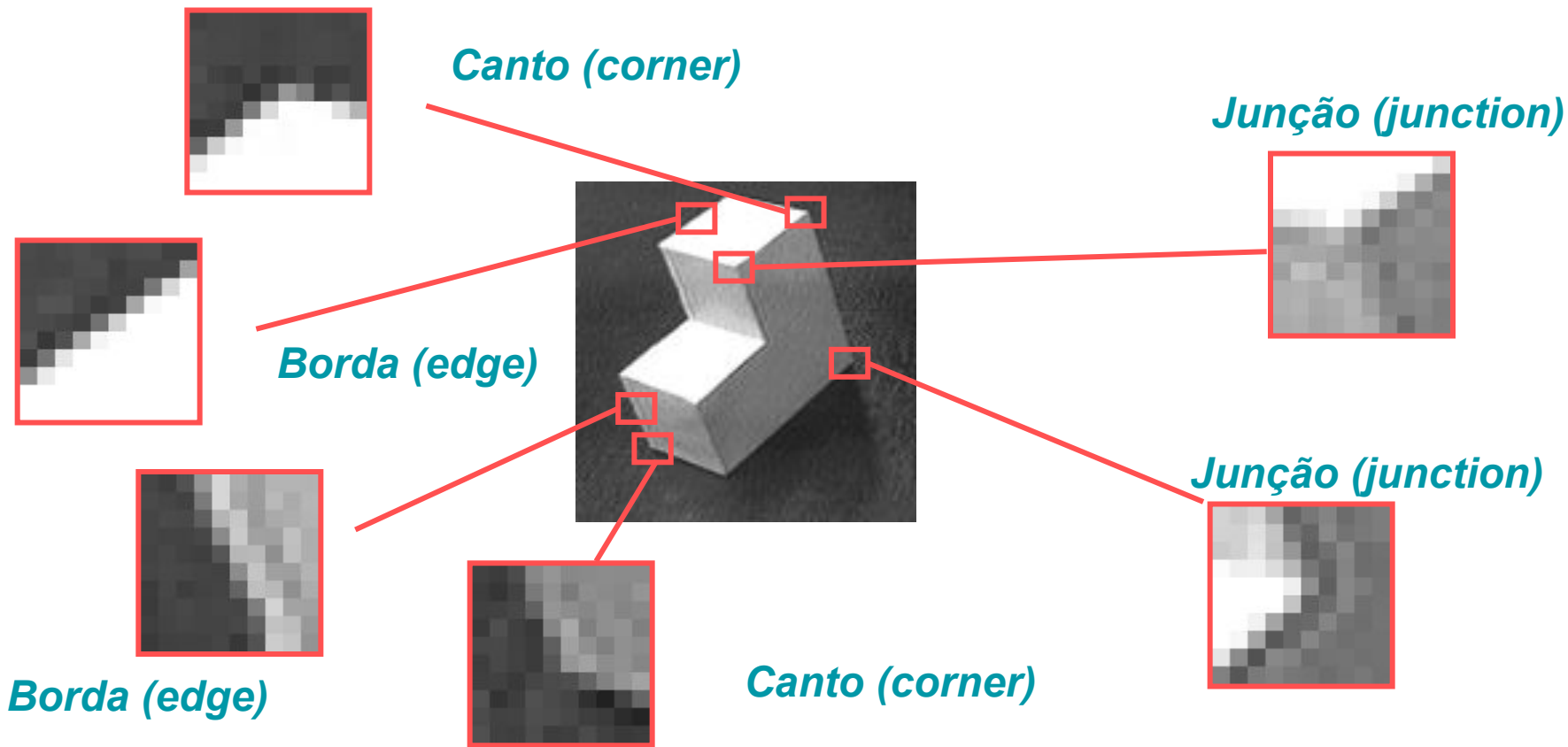
cena



- Ovos
- Tigela
- Borda
- Fundo

INTERPRETAÇÃO

Atributos locais (*Local features*)



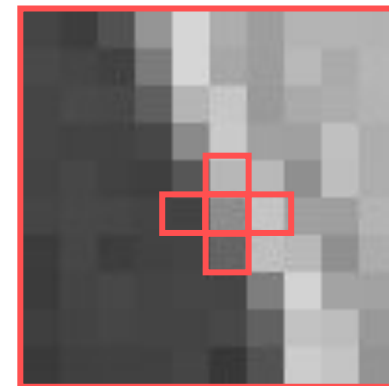
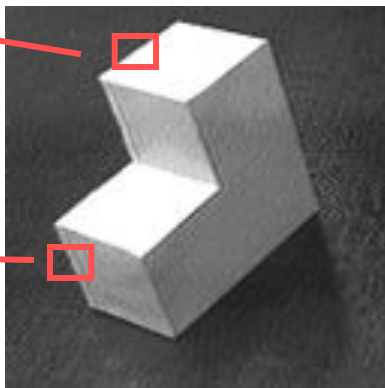
Atributos locais (*Local features*)



Borda (edge)



Borda (edge)



Bordas podem ser representadas com vizinhanças - 4 ou 8 de 1ª ordem

Filtragem convolucional

- É o nome generalizado para uma multiplicação de um vetor multidimensional por outro vetor multidimensional normalmente conhecido por máscara de tamanho menor, ou igual, ao primeiro vetor.
- Os pesos são os elementos que constituem a máscara de convolução.
- A filtragem convolucional é uma operação linear.
- A primeira palavra do nome “Convolutional Neural Network” deve-se ao fato que a máscara “transladada” sobre a imagem
- No contexto de processamento de imagens, a convolução é na verdade uma correlação, mas isso não é realmente relevante

Filtragem convolucional

6	16	15	12	11	11	11	10	10	12	10
8	10	9	10	12	13	13	11	10	9	9
8	8	9	12	10	9	8	10	11	8	9
61	16	12	10	6	8	9	10	10	6	8
175	129	78	52	22	15	13	7	9	12	8
149	172	151	154	128	66	20	10	8	12	9
40	65	114	139	127	61	14	12	14	11	10
6	5	18	27	31	19	12	13	10	9	11
8	9	7	7	7	6	7	8	9	12	12
11	11	8	10	9	8	10	11	10	6	12
13	13	9	9	9	12	10	7	9	10	10

Imagem original - f

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Máscara - g

0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Imagem convoluída - f * g

Filtragem convolucional

6	16	15	12	11	11	11	10	10	12	10
8	10	9	10	12	13	13	11	10	9	9
8	8	9	12	10	9	8	10	11	8	9
61	16	12	10	6	8	9	10	10	6	8
175	129	78	52	22	15	13	7	9	12	8
149	172	151	154	128	66	20	10	8	12	9
40	65	114	139	127	61	14	12	14	11	10
6	5	18	27	31	19	12	13	10	9	11
8	9	7	7	7	6	7	8	9	12	12
11	11	8	10	9	8	10	11	10	6	12
13	13	9	9	9	12	10	7	9	10	10

Imagem original - f

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Máscara - g

0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Imagem convoluída - $f * g$

Filtragem convolucional

6	16	15	12	11	11	11	10	10	12	10
8	10	9	10	12	13	13	11	10	9	9
8	8	9	12	10	9	8	10	11	8	9
61	16	12	10	6	8	9	10	10	6	8
175	129	78	52	22	15	13	7	9	12	8
149	172	151	154	128	66	20	10	8	12	9
40	65	114	139	127	61	14	12	14	11	10
6	5	18	27	31	19	12	13	10	9	11
8	9	7	7	7	6	7	8	9	12	12
11	11	8	10	9	8	10	11	10	6	12
13	13	9	9	9	12	10	7	9	10	10

Imagem original - f

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Máscara - g

0	0	0	0	0	0	0	0	0	0	0
0	9,89	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Imagem convoluída - $f * g$

Filtragem convolucional

6	16	15	12	11	11	11	10	10	12	10
8	10	9	10	12	13	13	11	10	9	9
8	8	9	12	10	9	8	10	11	8	9
61	16	12	10	6	8	9	10	10	6	8
175	129	78	52	22	15	13	7	9	12	8
149	172	151	154	128	66	20	10	8	12	9
40	65	114	139	127	61	14	12	14	11	10
6	5	18	27	31	19	12	13	10	9	11
8	9	7	7	7	6	7	8	9	12	12
11	11	8	10	9	8	10	11	10	6	12
13	13	9	9	9	12	10	7	9	10	10

Imagem original - f

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Máscara - g

0	0	0	0	0	0	0	0	0	0	0
0	9,59	11,22	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Imagem convoluída - $f * g$

Filtragem convolucional

6	16	15	12	11	11	11	10	10	12	10
8	10	9	10	12	13	13	11	10	9	9
8	8	9	12	10	9	8	10	11	8	9
61	16	12	10	6	8	9	10	10	6	8
175	129	78	52	22	15	13	7	9	12	8
149	172	151	154	128	66	20	10	8	12	9
40	65	114	139	127	61	14	12	14	11	10
6	5	18	27	31	19	12	13	10	9	11
8	9	7	7	7	6	7	8	9	12	12
11	11	8	10	9	8	10	11	10	6	12
13	13	9	9	9	12	10	7	9	10	10

Imagem original - f

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Máscara - g

0	0	0	0	0	0	0	0	0	0	0
0	9,89	11,22	11,11	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Imagem convoluída - $f * g$

Filtragem convolucional

6	16	15	12	11	11	11	10	10	12	10
8	10	9	10	12	13	13	11	10	9	9
8	8	9	12	10	9	8	10	11	8	9
61	16	12	10	6	8	9	10	10	6	8
175	129	78	52	22	15	13	7	9	12	8
149	172	151	154	128	66	20	10	8	12	9
40	65	114	139	127	61	14	12	14	11	10
6	5	18	27	31	19	12	13	10	9	11
8	9	7	7	7	6	7	8	9	12	12
11	11	8	10	9	8	10	11	10	6	12
13	13	9	9	9	12	10	7	9	10	10

Imagem original - f

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Máscara - g

0	0	0	0	0	0	0	0	0	0	0
0	9,89	11,22	11,11	11,11	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Imagem convoluída - $f * g$

Filtragem convolucional

6	16	15	12	11	11	11	10	10	12	10
8	10	9	10	12	13	13	11	10	9	9
8	8	9	12	10	9	8	10	11	8	9
61	16	12	10	6	8	9	10	10	6	8
175	129	78	52	22	15	13	7	9	12	8
149	172	151	154	128	66	20	10	8	12	9
40	65	114	139	127	61	14	12	14	11	10
6	5	18	27	31	19	12	13	10	9	11
8	9	7	7	7	6	7	8	9	12	12
11	11	8	10	9	8	10	11	10	6	12
13	13	9	9	9	12	10	7	9	10	10

Imagem original - f

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Máscara - g

0	0	0	0	0	0	0	0	0	0	0
0	9,89	11,22	11,11	11,11	10,89	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Imagem convoluída - $f * g$

Filtragem convolucional

6	16	15	12	11	11	11	10	10	12	10
8	10	9	10	12	13	13	11	10	9	9
8	8	9	12	10	9	8	10	11	8	9
61	16	12	10	6	8	9	10	10	6	8
175	129	78	52	22	15	13	7	9	12	8
149	172	151	154	128	66	20	10	8	12	9
40	65	114	139	127	61	14	12	14	11	10
6	5	18	27	31	19	12	13	10	9	11
8	9	7	7	7	6	7	8	9	12	12
11	11	8	10	9	8	10	11	10	6	12
13	13	9	9	9	12	10	7	9	10	10

Imagem original - f

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Máscara - g

0	0	0	0	0	0	0	0	0	0	0
0	9,89	11,22	11,11	11,11	10,89	10,67	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Imagem convoluída - $f * g$

Filtragem convolucional

6	16	15	12	11	11	11	10	10	12	10
8	10	9	10	12	13	13	11	10	9	9
8	8	9	12	10	9	8	10	11	8	9
61	16	12	10	6	8	9	10	10	6	8
175	129	78	52	22	15	13	7	9	12	8
149	172	151	154	128	66	20	10	8	12	9
40	65	114	139	127	61	14	12	14	11	10
6	5	18	27	31	19	12	13	10	9	11
8	9	7	7	7	6	7	8	9	12	12
11	11	8	10	9	8	10	11	10	6	12
13	13	9	9	9	12	10	7	9	10	10

Imagem original - f

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Máscara - g

0	0	0	0	0	0	0	0	0	0	0
0	9,89	11,22	11,11	11,11	10,89	10,67	10,44	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Imagem convoluída - $f * g$

Filtragem convolucional

6	16	15	12	11	11	11	10	10	12	10
8	10	9	10	12	13	13	11	10	9	9
8	8	9	12	10	9	8	10	11	8	9
61	16	12	10	6	8	9	10	10	6	8
175	129	78	52	22	15	13	7	9	12	8
149	172	151	154	128	66	20	10	8	12	9
40	65	114	139	127	61	14	12	14	11	10
6	5	18	27	31	19	12	13	10	9	11
8	9	7	7	7	6	7	8	9	12	12
11	11	8	10	9	8	10	11	10	6	12
13	13	9	9	9	12	10	7	9	10	10

Imagem original - f

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Máscara - g

0	0	0	0	0	0	0	0	0	0	0	
0	9,89	11,22	11,11	11,11	10,89	10,67	10,44	10,1	1	0	0
0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	

Imagem convoluída - $f * g$

Filtragem convolucional

6	16	15	12	11	11	11	10	10	12	10
8	10	9	10	12	13	13	11	10	9	9
8	8	9	12	10	9	8	10	11	8	9
61	16	12	10	6	8	9	10	10	6	8
175	129	78	52	22	15	13	7	9	12	8
149	172	151	154	128	66	20	10	8	12	9
40	65	114	139	127	61	14	12	14	11	10
6	5	18	27	31	19	12	13	10	9	11
8	9	7	7	7	6	7	8	9	12	12
11	11	8	10	9	8	10	11	10	6	12
13	13	9	9	9	12	10	7	9	10	10

Imagem original - f

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Máscara - g

0	0	0	0	0	0	0	0	0	0	0
0	9,89	11,22	11,11	11,11	10,89	10,67	10,44	10,1	9,78	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Imagem convoluída - $f * g$

Filtragem convolucional

6	16	15	12	11	11	11	10	10	12	10
8	10	9	10	12	13	13	11	10	9	9
8	8	9	12	10	9	8	10	11	8	9
61	16	12	10	6	8	9	10	10	6	8
175	129	78	52	22	15	13	7	9	12	8
149	172	151	154	128	66	20	10	8	12	9
40	65	114	139	127	61	14	12	14	11	10
6	5	18	27	31	19	12	13	10	9	11
8	9	7	7	7	6	7	8	9	12	12
11	11	8	10	9	8	10	11	10	6	12
13	13	9	9	9	12	10	7	9	10	10

Imagem original - f

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Máscara - g

0	0	0	0	0	0	0	0	0	0	0
0	9,89	11,22	11,11	11,11	10,89	10,67	10,44	10,1	9,78	0
0	15,67	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Imagem convoluída - $f * g$

Filtragem convolucional

6	16	15	12	11	11	11	10	10	12	10
8	10	9	10	12	13	13	11	10	9	9
8	8	9	12	10	9	8	10	11	8	9
61	16	12	10	6	8	9	10	10	6	8
175	129	78	52	22	15	13	7	9	12	8
149	172	151	154	128	66	20	10	8	12	9
40	65	114	139	127	61	14	12	14	11	10
6	5	18	27	31	19	12	13	10	9	11
8	9	7	7	7	6	7	8	9	12	12
11	11	8	10	9	8	10	11	10	6	12
13	13	9	9	9	12	10	7	9	10	10

Imagem original - f

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Máscara - g

0	0	0	0	0	0	0	0	0	0	0
0	9,89	11,22	11,11	11,11	10,89	10,67	10,44	10,1	9,78	0
0	15,67	10,67	10,00	10,00	9,78	10,11	10,22	9,44	8,89	0
0	55,11	36,22	23,44	16,00	11,11	9,89	9,67	9,22	9,00	0
0	104,78	86,00	68,11	51,22	31,89	17,56	10,67	9,33	9,11	0
0	119,22	117,11	107,22	84,89	51,78	24,22	11,89	10,56	10,33	0
0	80,00	93,89	98,78	83,56	53,11	25,22	12,56	11,00	10,44	0
0	30,22	43,44	53,00	47,11	31,56	16,89	11,00	10,89	10,89	0
0	9,22	11,33	13,78	13,78	12,11	10,44	10,00	9,78	10,11	0
0	9,89	9,22	8,33	8,56	8,67	8,78	9,00	9,11	0	0
0	0	0	0	0	0	0	0	0	0	0

Imagem convoluída - f * g

Filtragem convolucional

6	16	15	12	11	11	11	10	10	12	10
8	10	9	10	12	13	13	11	10	9	9
8	8	9	12	10	9	8	10	11	8	9
61	16	12	10	6	8	9	10	10	6	8
175	129	78	52	22	15	13	7	9	12	8
149	172	151	154	128	66	20	10	8	12	9
40	65	114	139	127	61	14	12	14	11	10
6	5	18	27	31	19	12	13	10	9	11
8	9	7	7	7	6	7	8	9	12	12
11	11	8	10	9	8	10	11	10	6	12
13	13	9	9	9	12	10	7	9	10	10

Imagem original - f

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Máscara - g

0	0	0	0	0	0	0	0	0	0	0
0	9,89	11,22	11,11	11,11	10,89	10,67	10,44	10,1	9,78	0
0	15,67	10,67	10,00	10,00	9,78	10,11	10,22	9,44	8,89	0
0	55,11	36,22	23,44	16,00	11,11	9,89	9,67	9,22	9,00	0
0	104,78	86,00	68,11	51,22	31,89	17,56	10,67	9,33	9,11	0
0	119,22	117,11	107,22	84,89	51,78	24,22	11,89	10,56	10,33	0
0	80,00	93,89	98,78	83,56	53,11	25,22	12,56	11,00	10,44	0
0	30,22	43,44	53,00	47,11	31,56	16,89	11,00	10,89	10,89	0
0	9,22	11,33	13,78	13,78	12,11	10,44	10,00	9,78	10,11	0
0	9,89	9,22	8,33	8,56	8,67	8,78	9,00	9,11	10	0
0	0	0	0	0	0	0	0	0	0	0

Imagem convoluída - f * g

Filtragem convolucional

6	16	15	12	11	11	11	10	10	12	10
8	10	9	10	12	13	13	11	10	9	9
8	8	9	12	10	9	8	10	11	8	9
61	16	12	10	6	8	9	10	10	6	8
175	129	78	52	22	15	13	7	9	12	8
149	172	151	154	128	66	20	10	8	12	9
40	65	114	139	127	61	14	12	14	11	10
6	5	18	27	31	19	12	13	10	9	11
8	9	7	7	7	6	7	8	9	12	12
11	11	8	10	9	8	10	11	10	6	12
13	13	9	9	9	12	10	7	9	10	10

Imagem original - f

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Máscara - g

0	0	0	0	0	0	0	0	0	0	0
0	10	11	11	11	11	11	10	10	10	0
0	16	11	10	10	10	10	10	9	9	0
0	55	36	23	16	11	10	10	9	9	0
0	105	86	68	51	32	18	11	9	9	0
0	119	117	107	85	52	24	12	11	10	0
0	80	94	99	84	53	25	13	11	10	0
0	30	43	53	47	32	17	11	11	11	0
0	9	11	14	14	12	10	10	10	10	0
0	10	9	8	9	9	9	9	9	10	0
0	0	0	0	0	0	0	0	0	0	0

Imagem convoluída - $f * g$


Filtragem convolucional

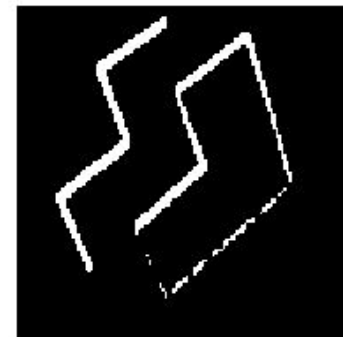
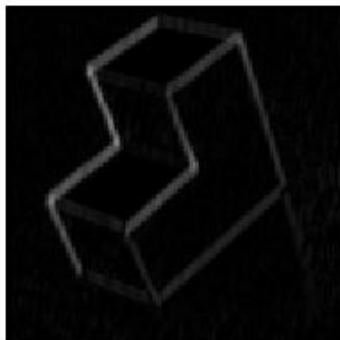
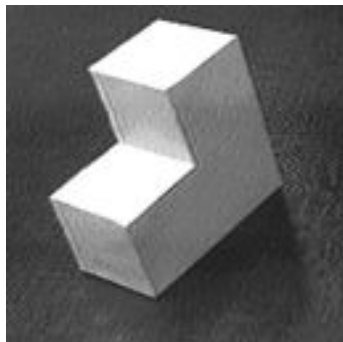
- De acordo com os valores dos pesos os filtros realizarão diferentes tipos de ações sobre a imagem
 - Suavização
 - Filtragem de ruído
 - Realce de contraste
 - Detecção de bordas
 - Detecção de atributos locais

- Exemplo: bloco
 - Aplicando-se o núcleo à imagem abaixo e
 - depois a limiarização

Filtros direcionais

$$K := \begin{pmatrix} -1 & -2 & 0 & 2 & 1 \\ -1 & -2 & 0 & 2 & 1 \\ -1 & -2 & 0 & 2 & 1 \\ -1 & -2 & 0 & 2 & 1 \\ -1 & -2 & 0 & 2 & 1 \end{pmatrix}$$

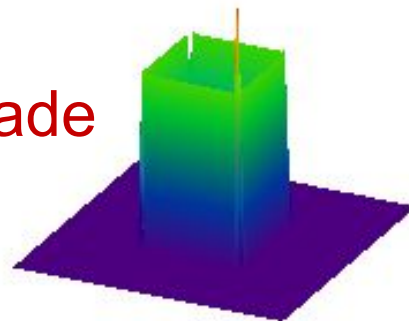

gradiente



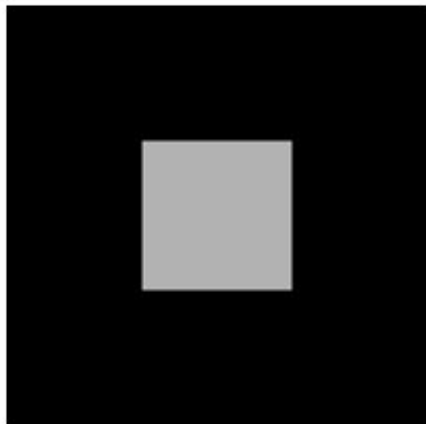
Máscara para detecção de cantos

$$\begin{pmatrix} -1 & -1 & -1 \\ -1 & 0 & 0 \\ -1 & 0 & 5 \end{pmatrix}$$

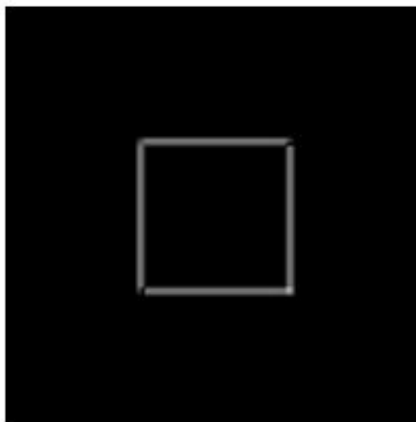
Intensidade
Filtrada



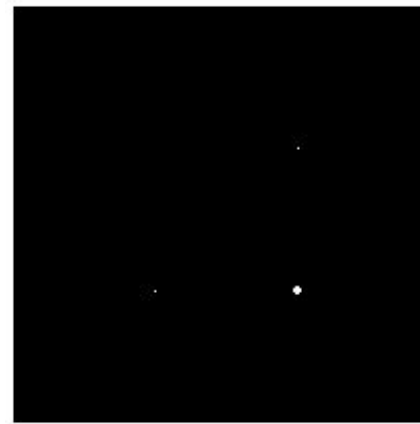
Original



Filtrada



Limiarizada

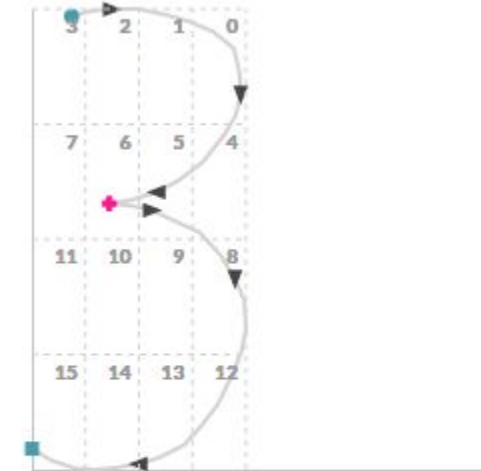
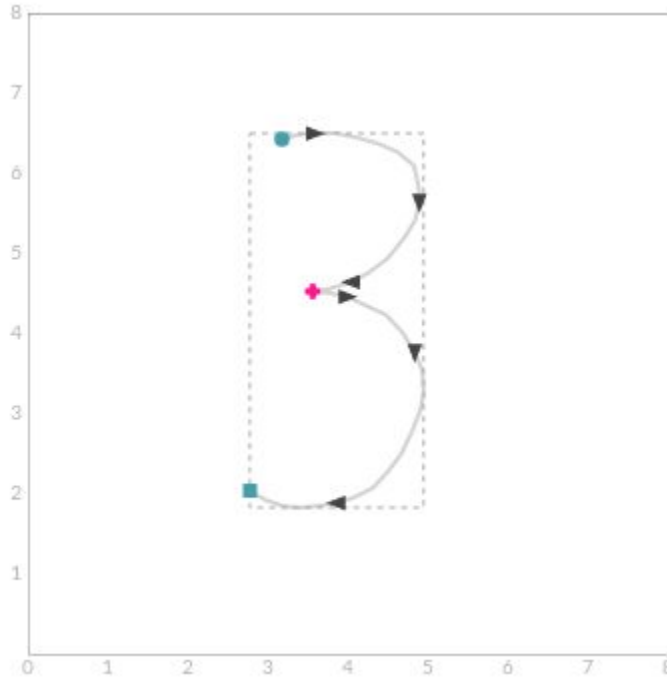


Exemplo: identificação via atributos locais

- A identificação de caracteres manuscritos é um dos problemas mais antigos e difíceis em Visão Computacional.
- Em 1966, um grupo de cientistas que trabalhavam para a RAND Corporation publicaram um artigo interno sobre um reconhecedor de caracteres para um dispositivo do tipo “tablet” baseado em atributos locais como **cantos, curvatura, posicionamento, razão proporcional, tamanho e terminações**.
- O classificador da RAND era baseado em regras que analisavam esses atributos locais e classificavam os caracteres.



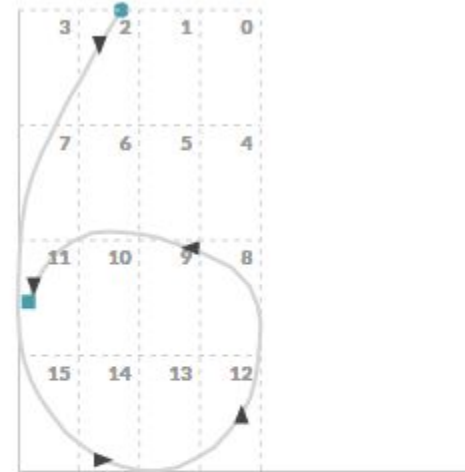
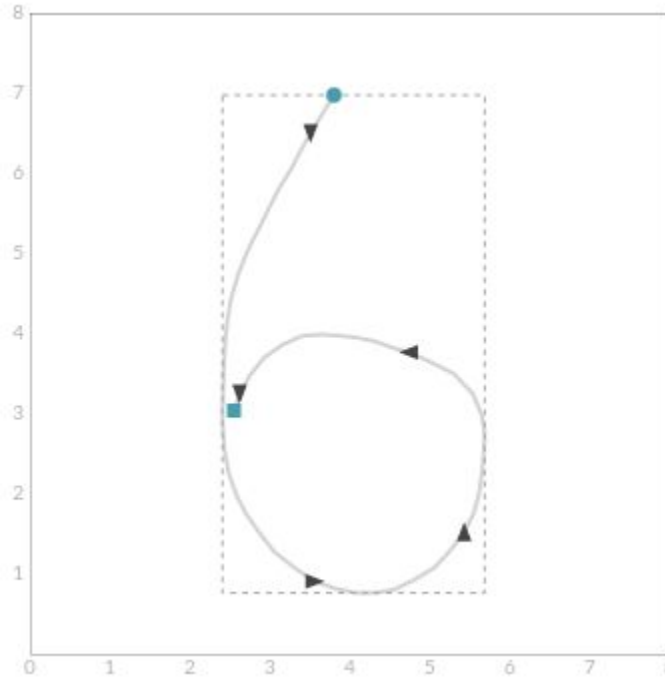
Identificação via atributos locais - exemplos



Center: (3.86, 4.16)
Stroke Start: 3
Stroke End: 15
Directions: ▶ ▼ ◀ ▶ ▼ ◀
Corners: 6

Width: 2.17
Height: 4.67
Aspect Ratio: 2.15

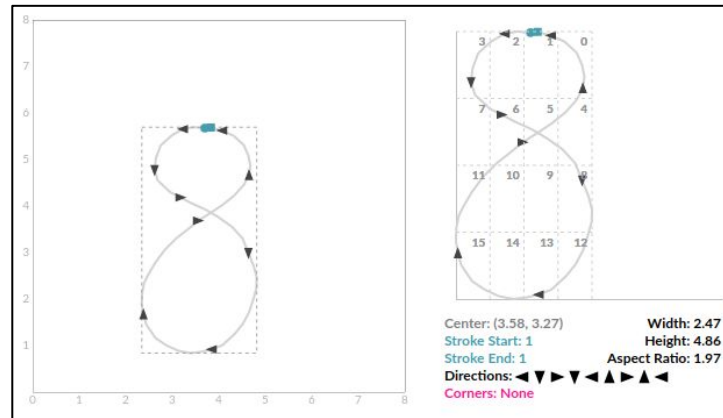
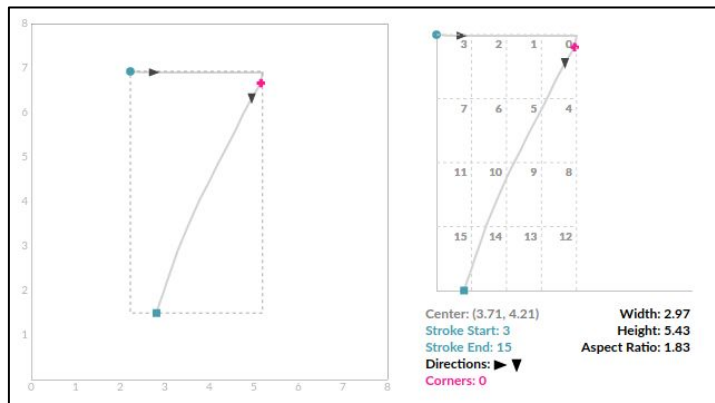
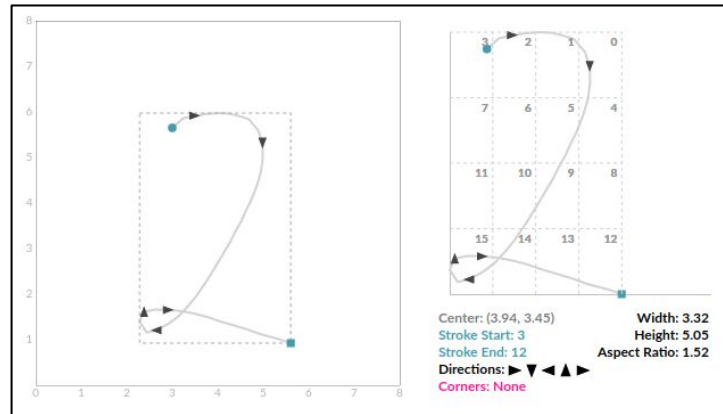
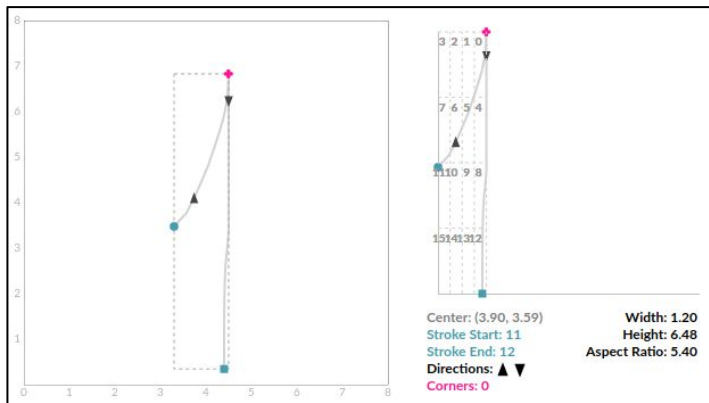
Identificação via atributos locais - exemplos



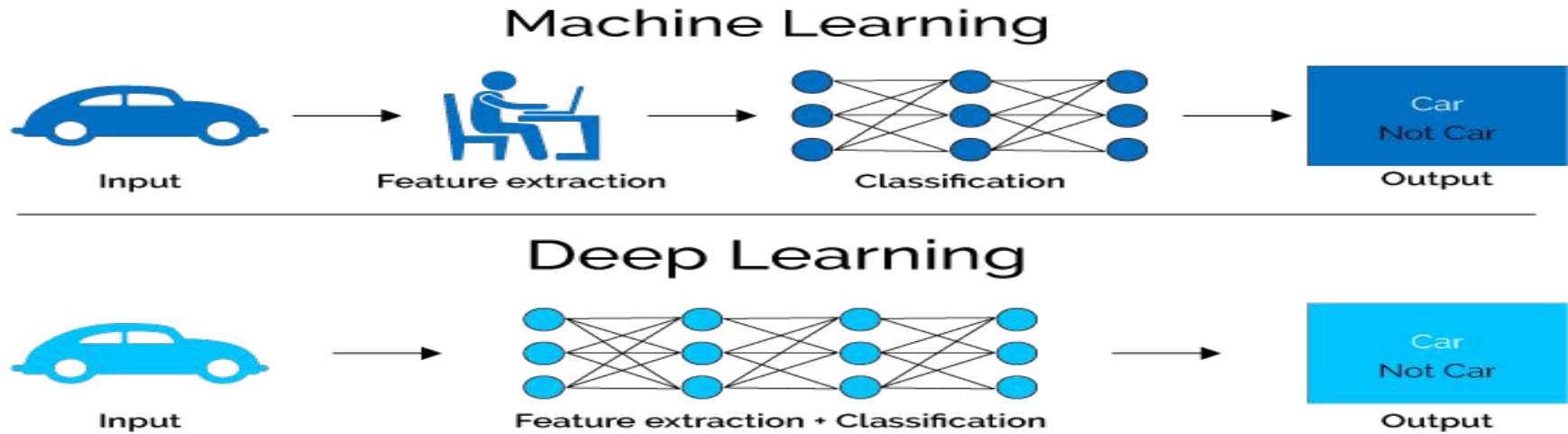
Center: (4.05, 3.87)
Stroke Start: 2
Stroke End: 11
Directions: ▼▶▲◀▼▼
Corners: None

Width: 3.28
Height: 6.21
Aspect Ratio: 1.90

Identificação via atributos locais - exemplos



ML convencional X CNN



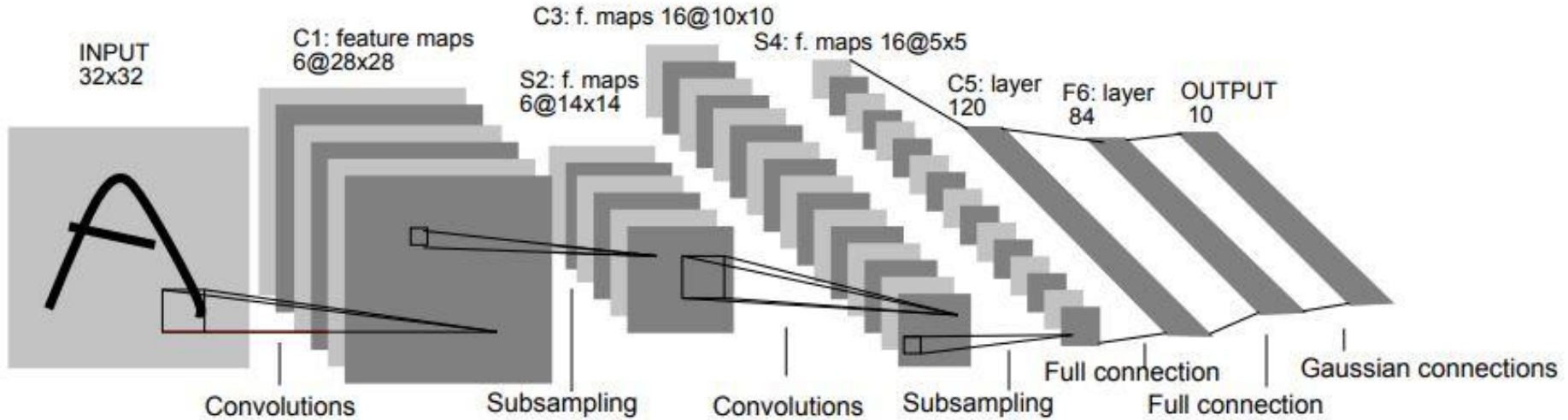
- E se os núcleos fossem criados no processo de aprendizado?

Deep learning

- Prêmio Turing da ACM - 2018



Rede neural convolucional - LeNet 5



Principais técnicas:

- “average pooling” - espécie de filtro da média, mas com mudança de resolução;
- função de ativação - sigmoid ou tangente hiperbólica (tanh);
- camadas completamente conectadas no final;
- treinado no conjunto MNIST, um conjunto de dígitos com 60K exemplos.

Propriedades físicas da imagem - resolução espacial



512x512



512x512

Propriedades físicas da imagem - resolução espacial



512x512



256x256

Propriedades físicas da imagem - resolução espacial



512x512



128x128

Propriedades físicas da imagem - resolução espacial



512x512



64x64

Propriedades físicas da imagem - resolução espacial



512x512



32x32

Propriedades físicas da imagem - resolução espacial



512x512



16x16

Propriedades físicas da imagem - resolução representação



8 bits



8 bits

Propriedades físicas da imagem - resolução representação



8 bits



7 bits

Propriedades físicas da imagem - resolução representação



8 bits



6 bits

Propriedades físicas da imagem - resolução representação



8 bits



5 bits

Propriedades físicas da imagem - resolução representação



8 bits



4 bits

Propriedades físicas da imagem - resolução representação



8 bits



3 bits

Propriedades físicas da imagem - resolução representação



8 bits

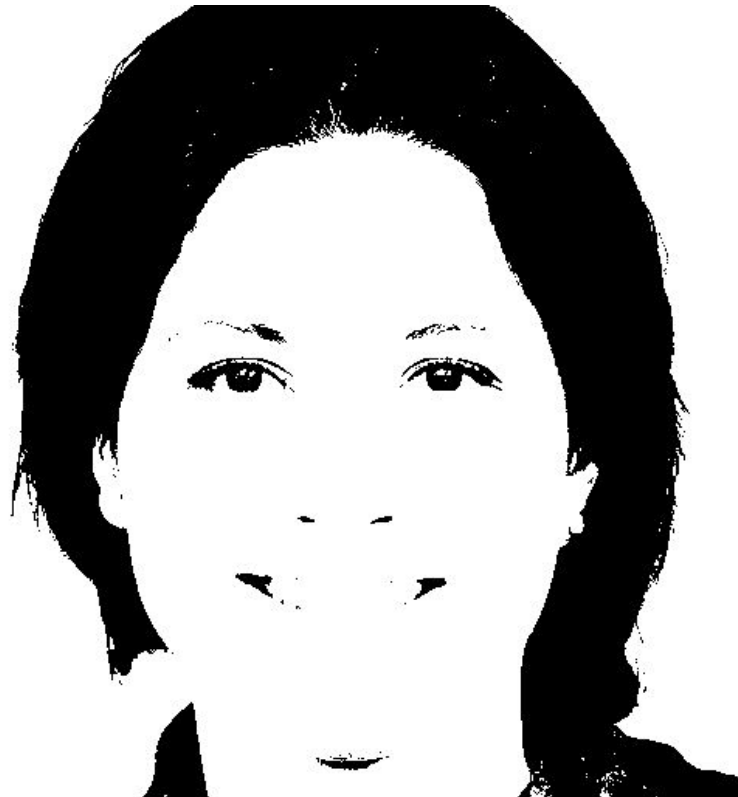


2 bits

Propriedades físicas da imagem - resolução representação

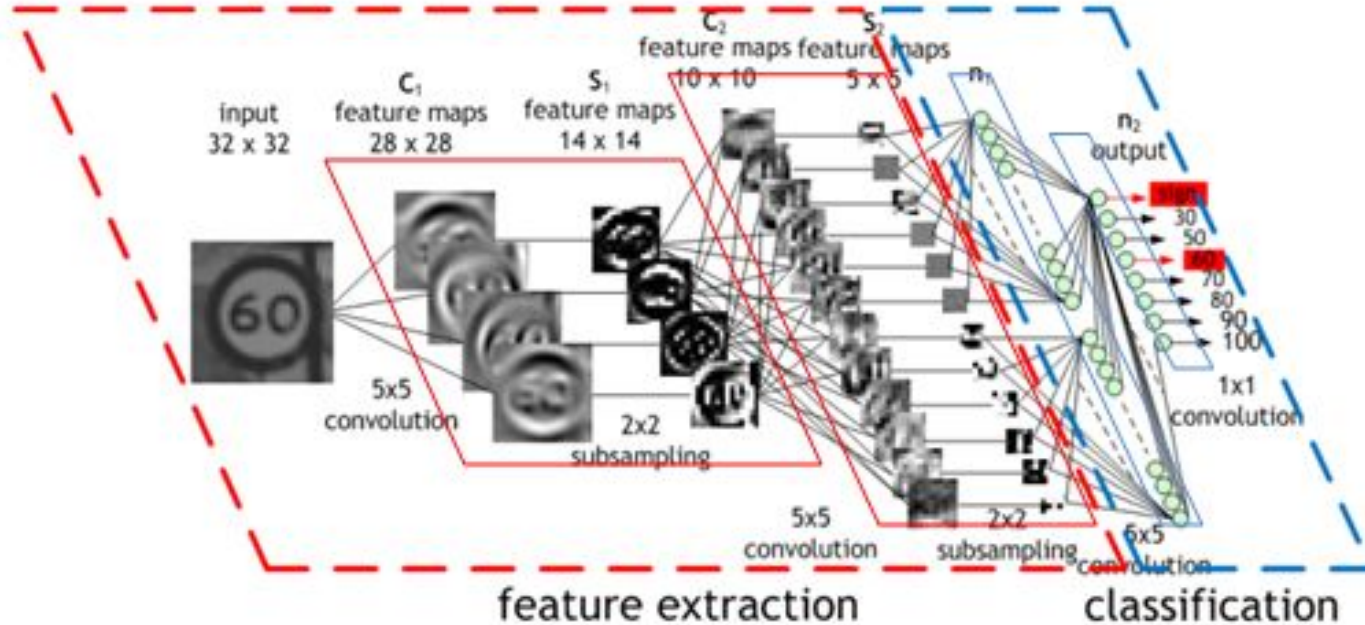


8 bits



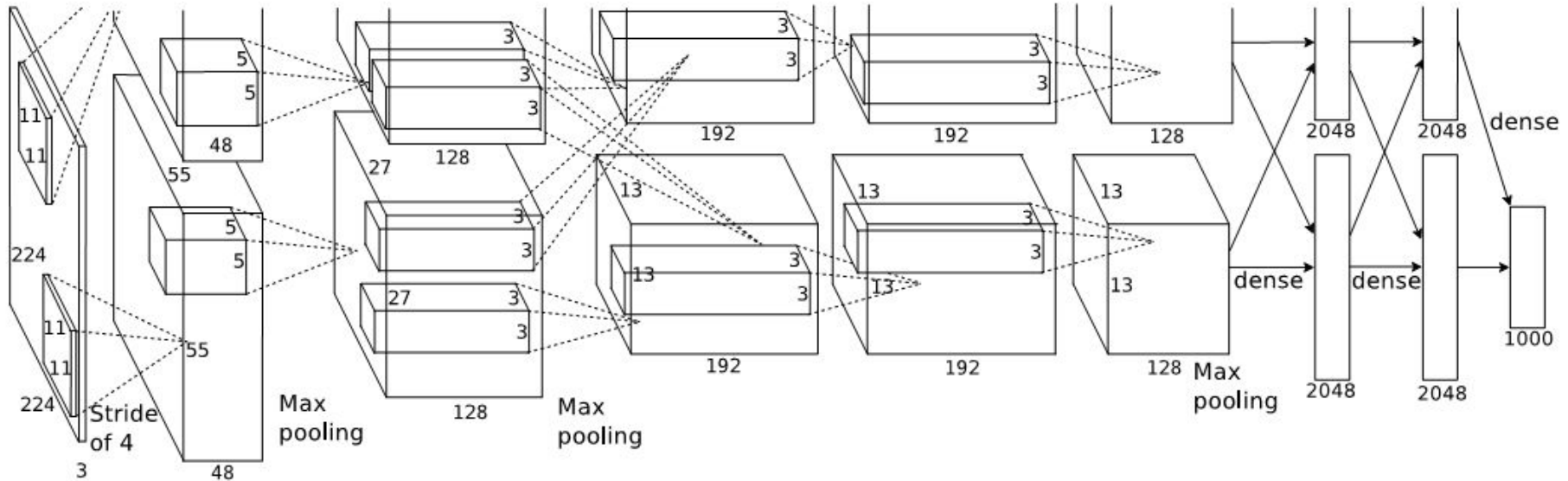
1 bits

Convolutional neural network (CNN)



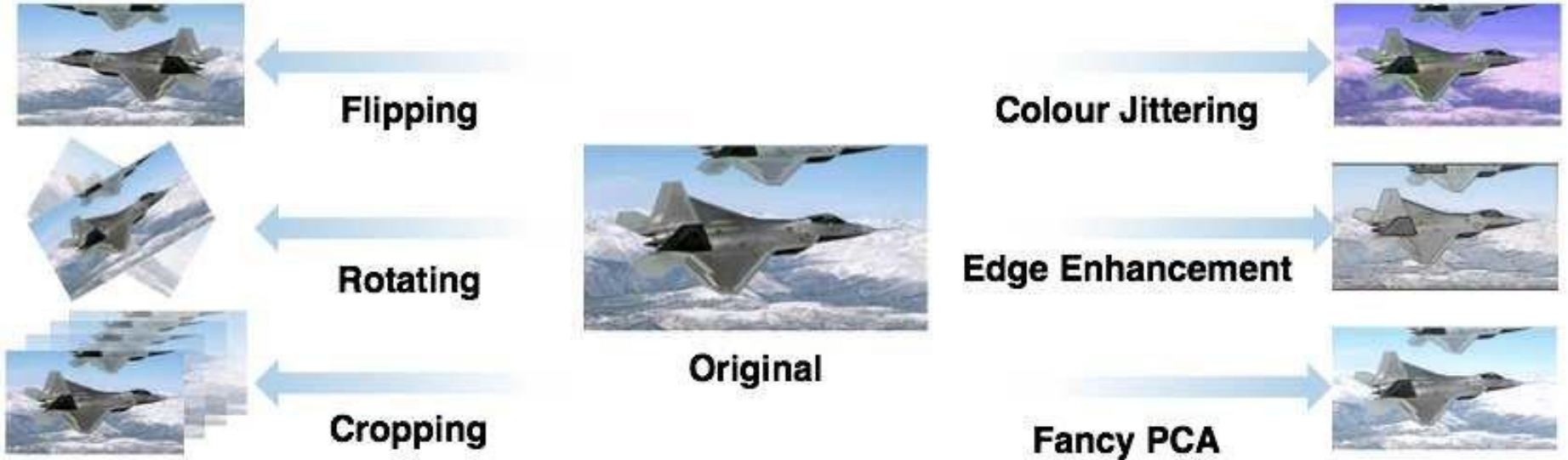
Uma CNN simples para classificar placas de limite de velocidade

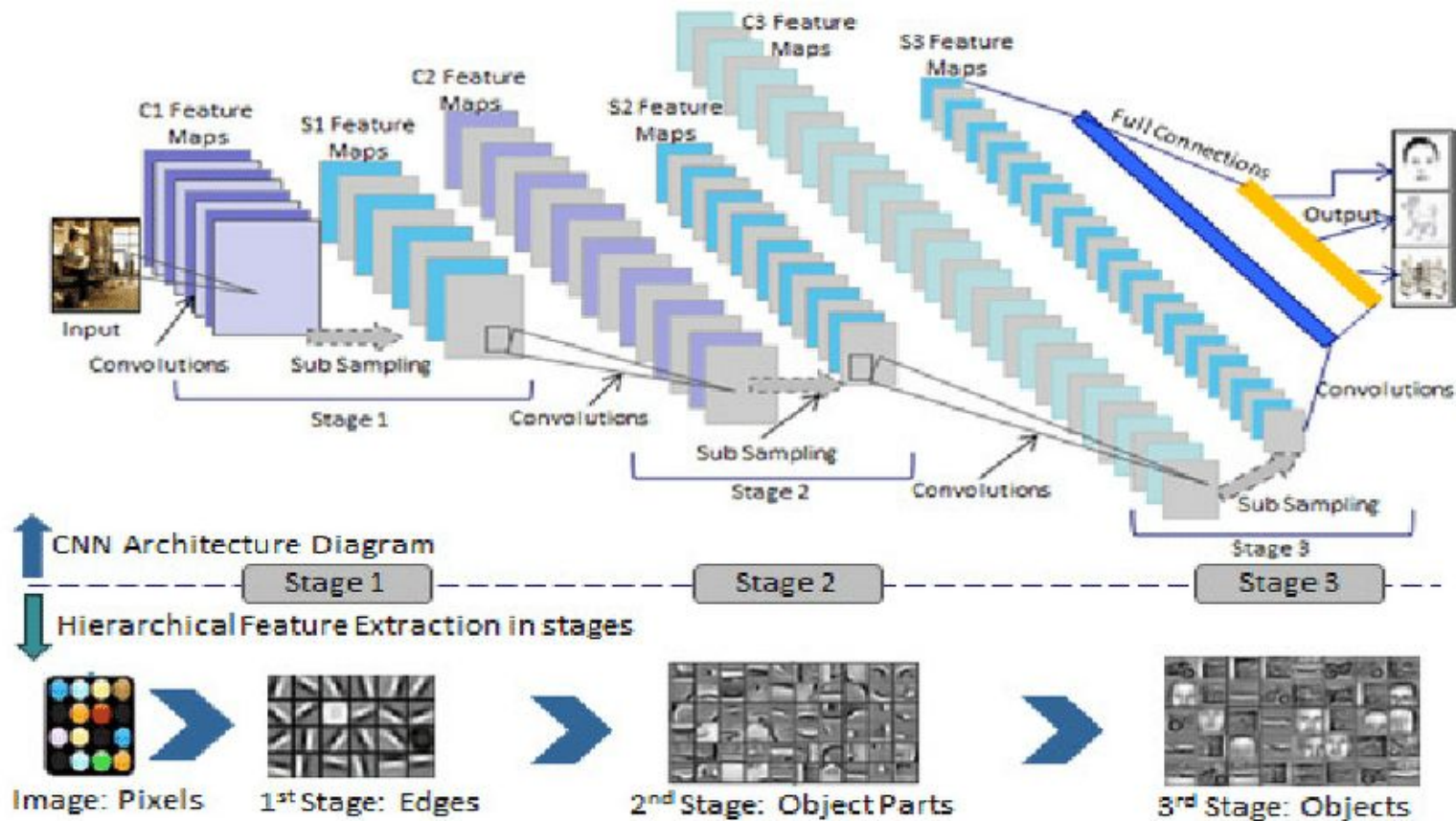
Alex Net - 2012



- Primeira rede CNN a ganhar a competição ImageNet (> 14 milhões de imagens, 20k categorias)
- Cinco camadas convolucionais + três camadas completamente conectadas + classificador softmax (generalização da regressão logística);
- Função de ativação ReLU, “dropout” e aumento de dados;
- 61 milhões de parâmetros, 650 mil neurônios

Aumentação de dados





VGG Net - 2014

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Aumenta o número de camadas de 8 para 11 até 19

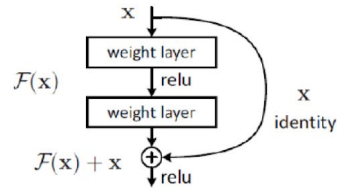
Diminui o tamanho das máscaras de convolução

As versões D e E são usadas para muitas tarefas e são conhecidas como VGG 16 e VGG 19

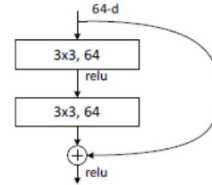
A VGG 16 tem aproximadamente 138 milhões de parâmetros e a VGG 19 quase 143 milhões

Fonte da tabela: Very Deep Convolutional Networks for Large-Scale Image Recognition - K. Simonyan e A. Zisserman

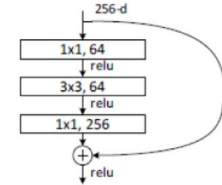
ResNet - 2015



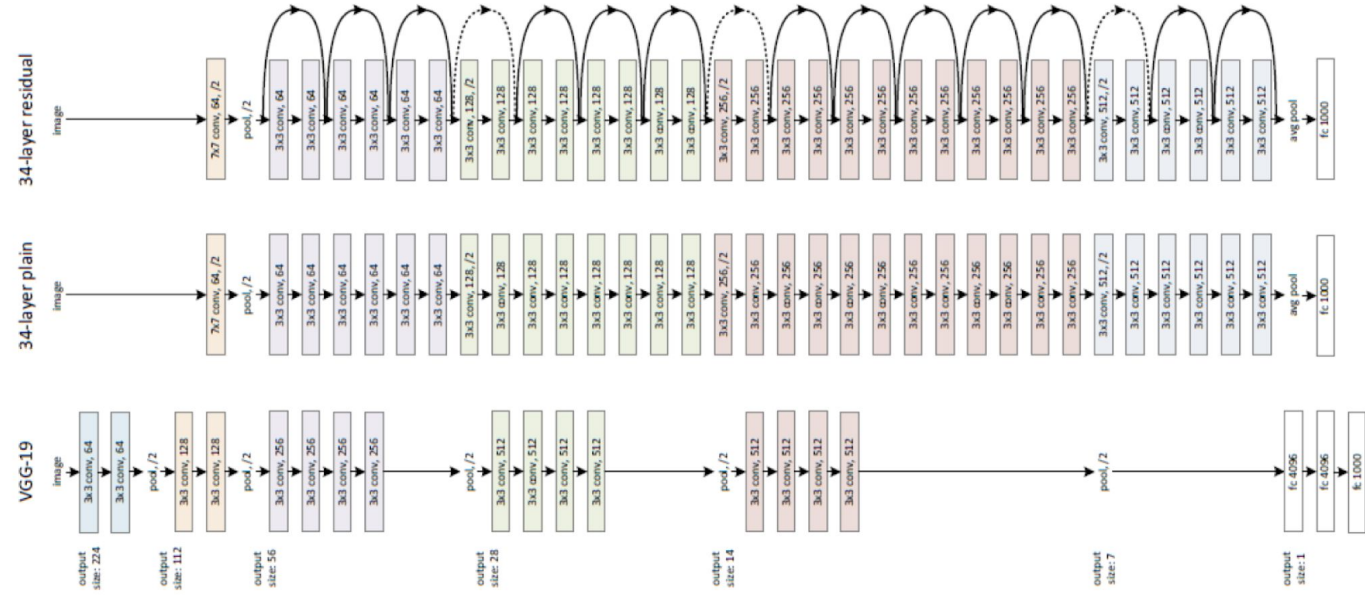
Residual learning: a building block



A deeper residual function F for ImageNet



Example network architectures for ImageNet



Diferenciais: Atinge o patamar humano para classificação no ImageNet; usa a ideia de conexões residuais; > de 1000 camadas.

Fonte: Atlas Wang

Uma miríade de arquiteturas em 2016!

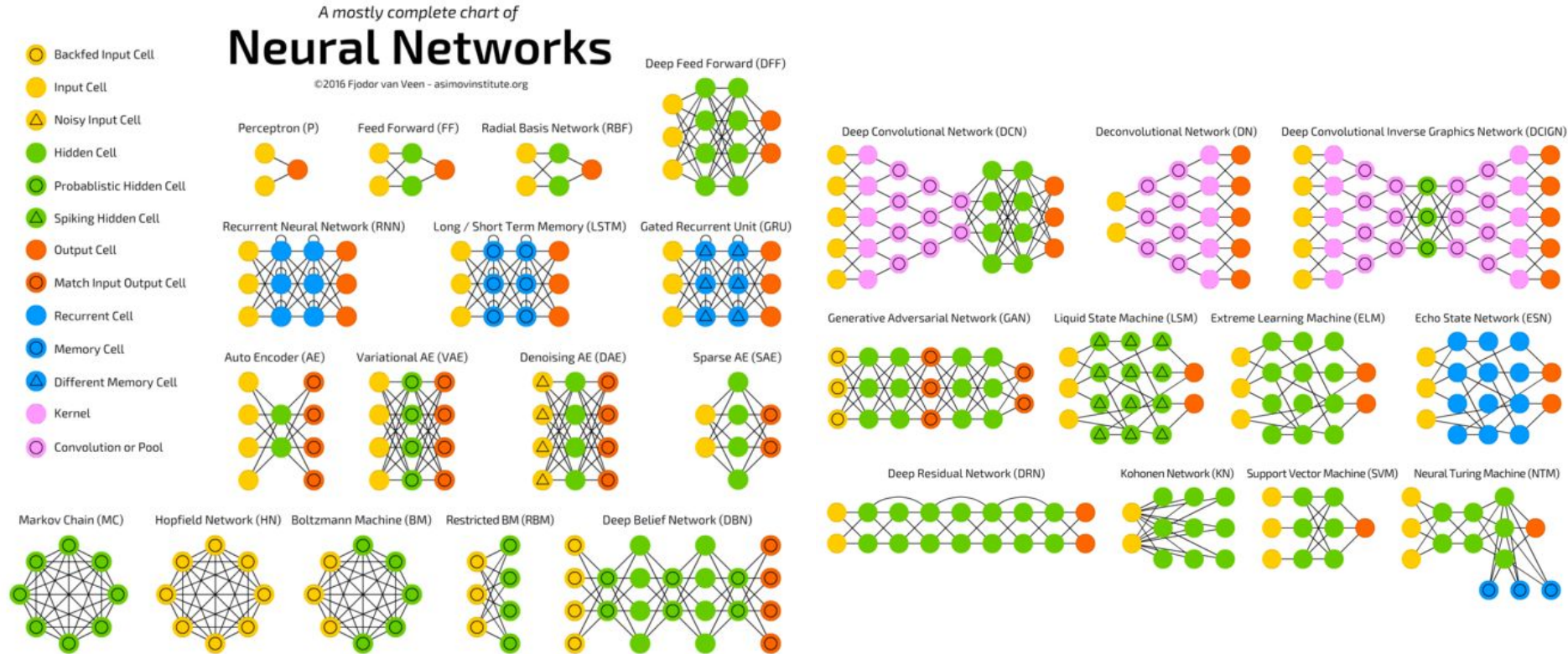
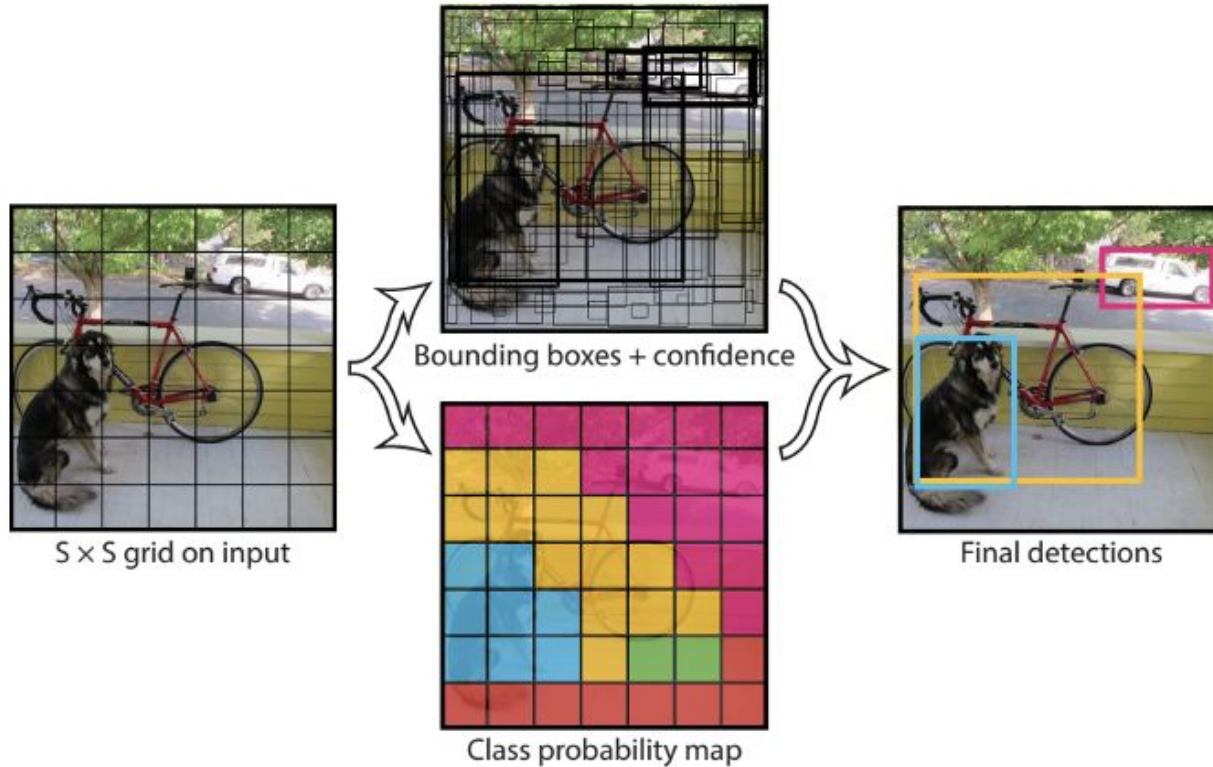
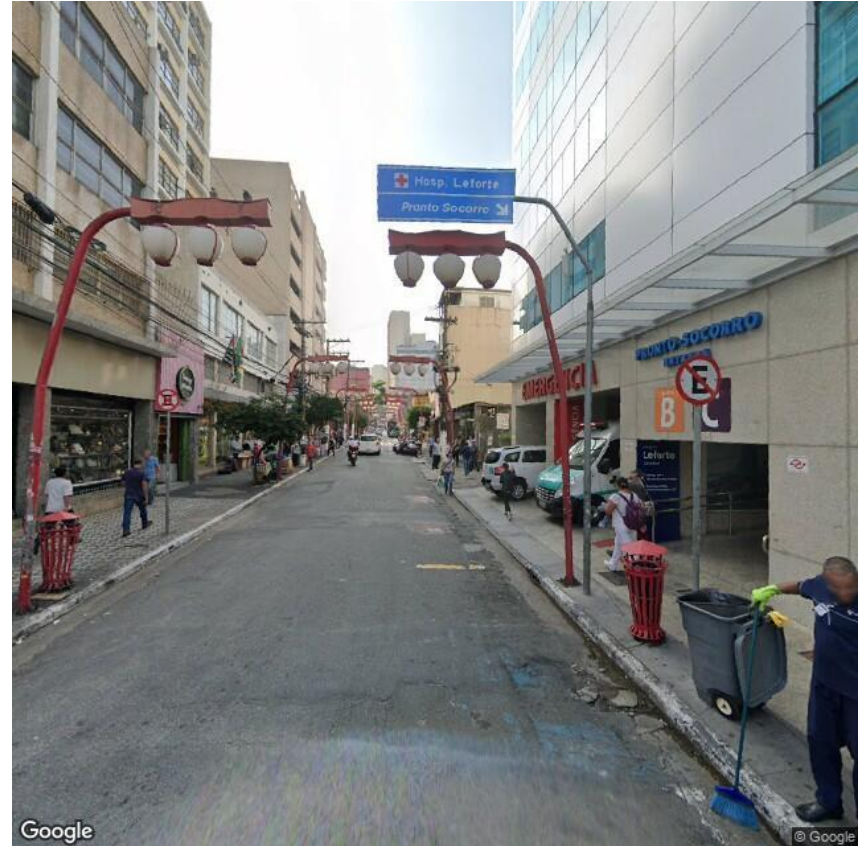
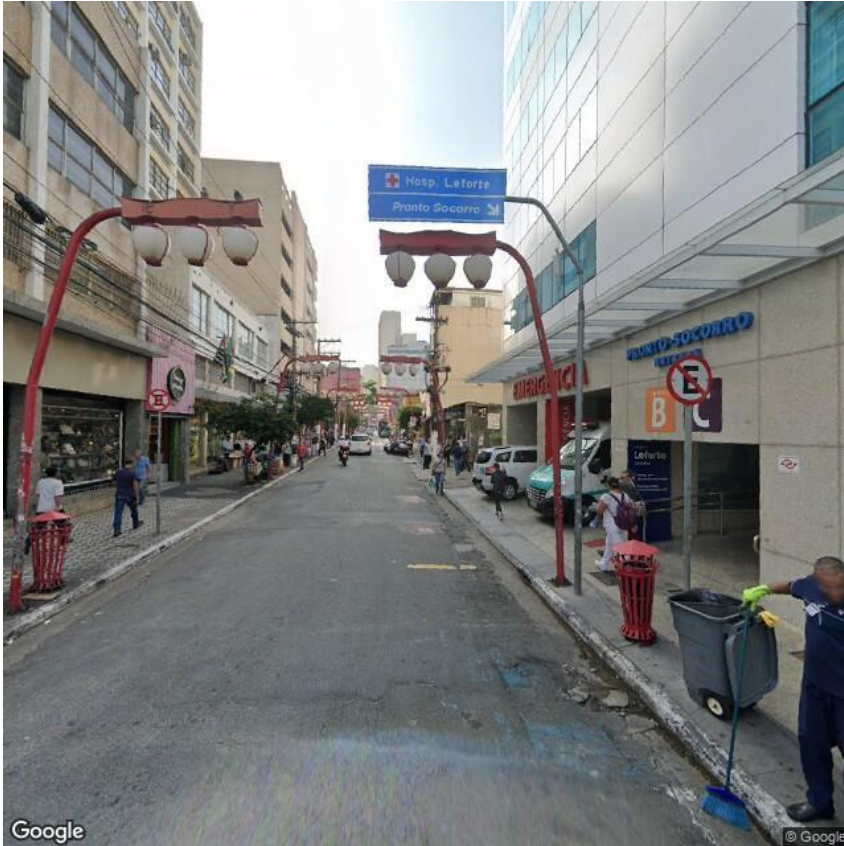


Diagrama feito por [Fjodor van Veen](https://www.asimovinstitute.org/)

You only look once - YOLO - 2016



YOLO v3

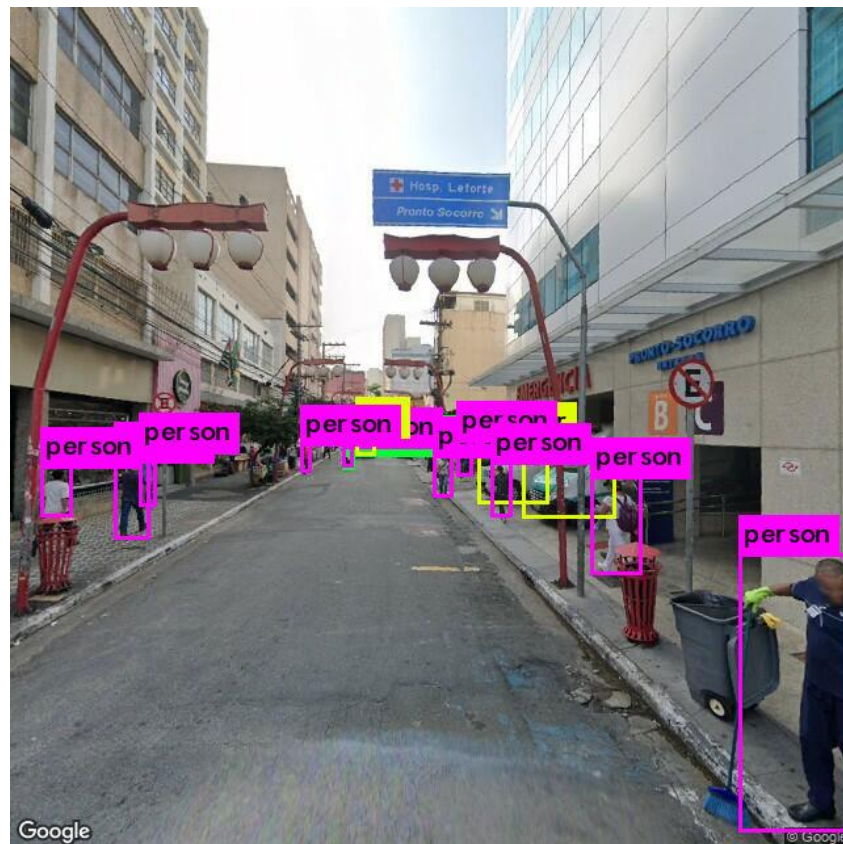


Software: <https://pjreddie.com/darknet/yolo/>

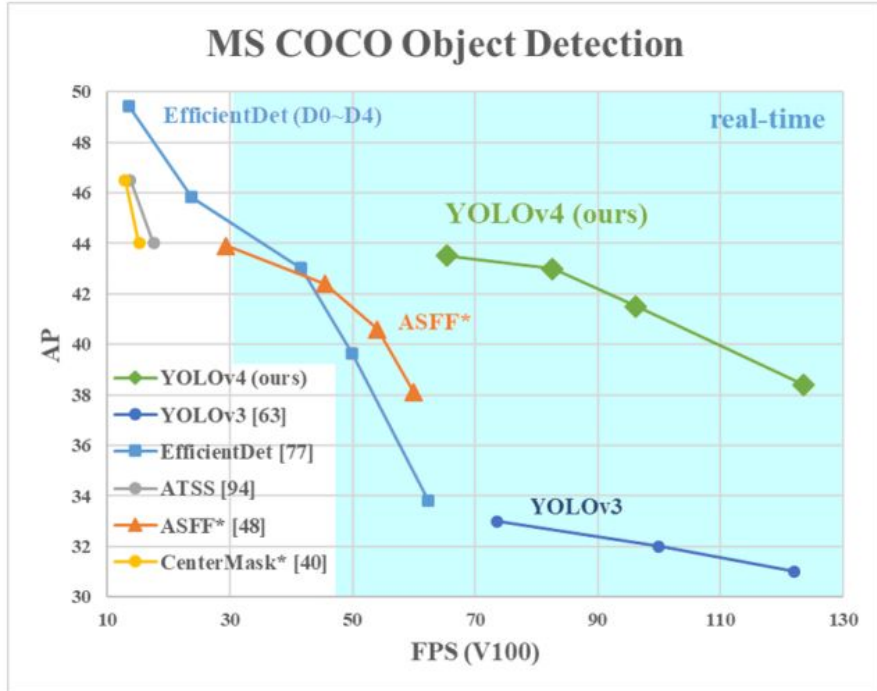
YOLO v3

GalvaoBueno.jpeg: Predicted in 27.216096 s

motorbike: 81%
person: 79%
car: 93%
car: 80%
car: 61%
person: 100%
person: 99%
person: 95%
person: 83%
person: 82%
person: 73%
person: 67%
person: 58%
person: 57%



Tarefa: localização e reconhecimento de objetos



YOLO v5 e variantes - max 140 FPS

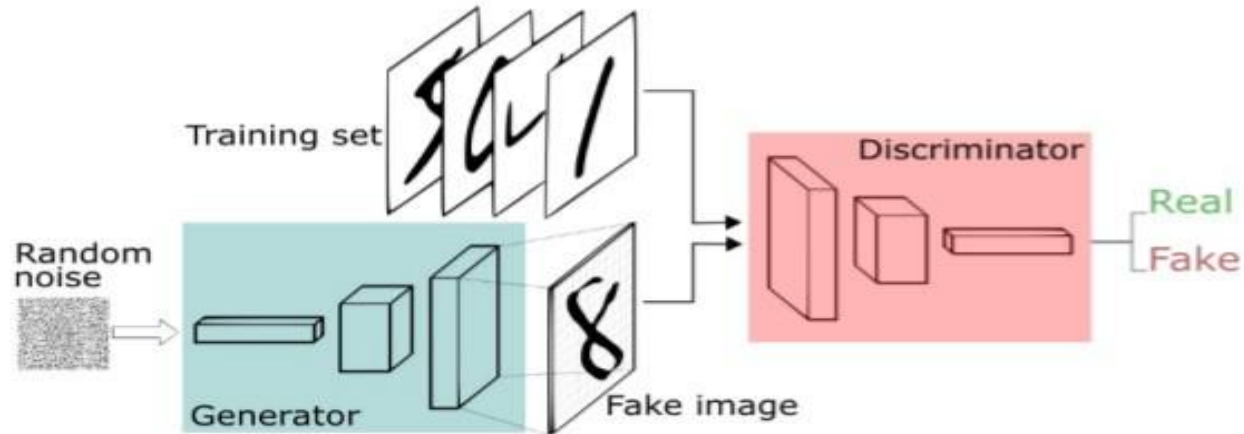
Fonte: <https://blog.roboflow.ai/yolov5-is-here/>

<https://blog.roboflow.ai/yolov5-improvements-and-evaluation/>

Generative Adversarial Networks

Goodfellow et. al. - NIPS - 2014

GAN Architecture

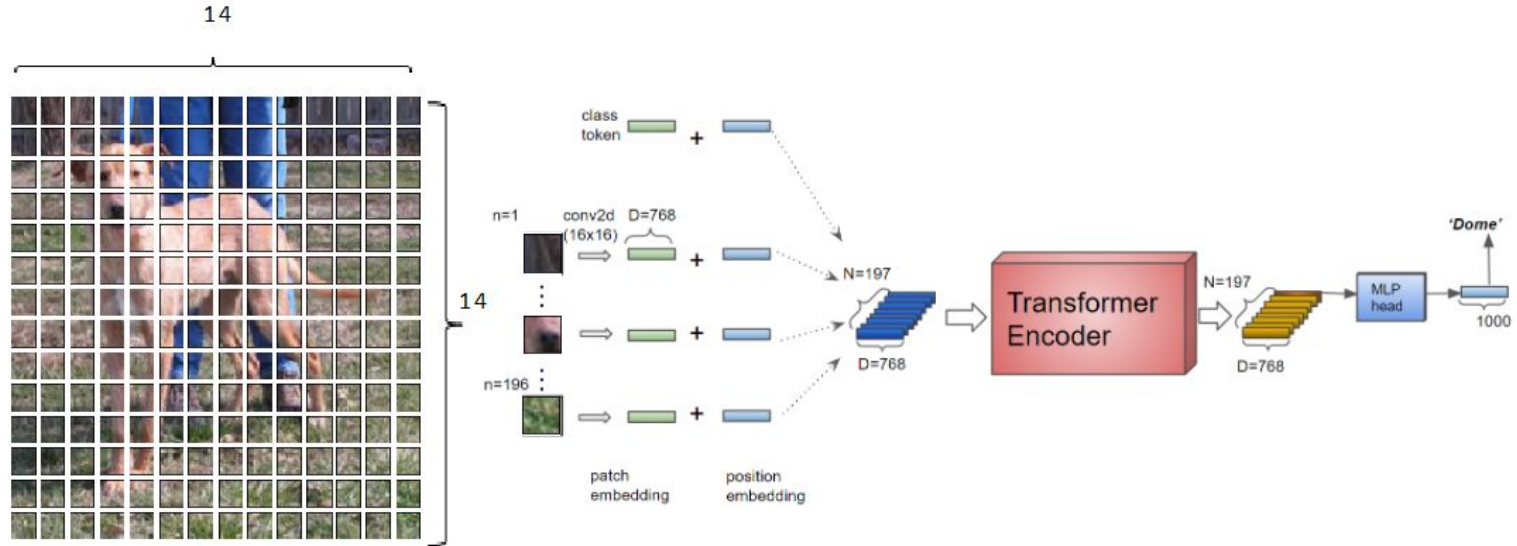


Style GAN

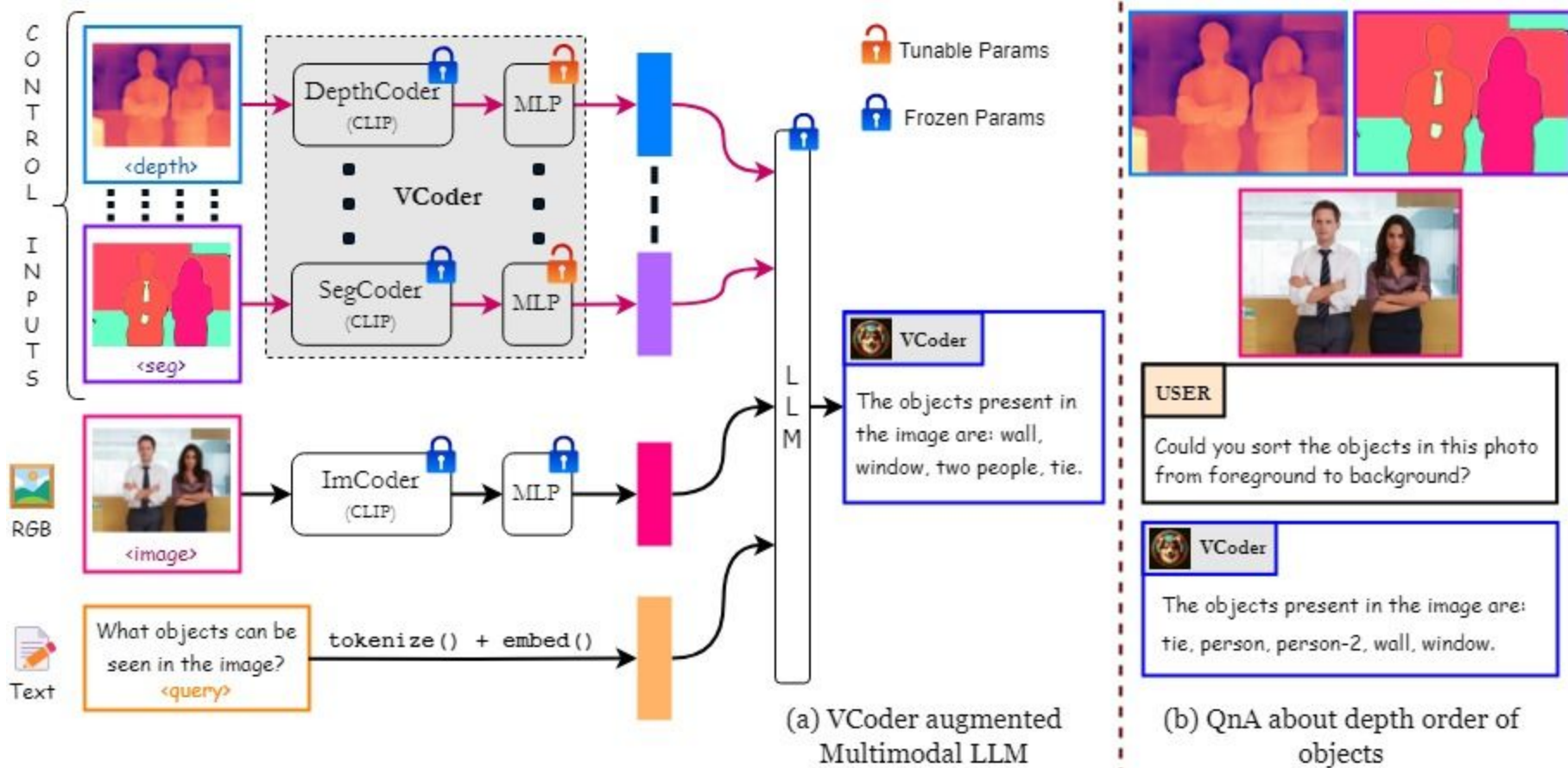


Nenhuma destas pessoas existe!!!

Vision Transformer (ViT) - 2020/2021



Modelos mistos - VCoder - Jain et. al. - 2023



O problema da cesta de ovos com chatGPT 4

Você pode descrever esta imagem?



chatGPT 4: A imagem parece um close-up de um grupo de ovos brancos. Os ovos parece agrupados e o foco está no ovo central. Os ovos ao redor estão desfocados, dando a impressão de profundidade. O fundo é uniforme, como se eles estivessem numa bandeja. A impressão geral é de simplicidade.

O problema da cesta de ovos com chatGPT 4

Quantos ovos há na imagem?



chatGPT 4: A imagem não é clara o suficiente para contar o número exato de ovos...

Quantos ovos inteiros há na imagem?

chatGPT 4: Na imagem, eu consigo distinguir no mínimo quatro ovos inteiros que se destacam do fundo. Pode haver mais, mas estão fora de foco, ou apenas parcialmente visíveis.

O problema da cesta de ovos com chatGPT 4

Esta é uma imagem melhor. Você pode descrever esta imagem?



chatGPT 4: A imagem melhor que mostra um certo número de ovos brancos. Eles estão agrupados e, deste ângulo, você pode ver mais ovos e seu arranjo, sugerindo que eles estão numa bandeja. Há sombras na bandeja, o que ajuda a distinguir os ovos individualmente.

O problema da cesta de ovos com chatGPT 4

Quantos ovos há na imagem?



chatGPT 4: Nesta imagem, eu posso ver seis ovos...

vcoder: Há seis ovos na imagem

vcoder: Há oito ovos na imagem



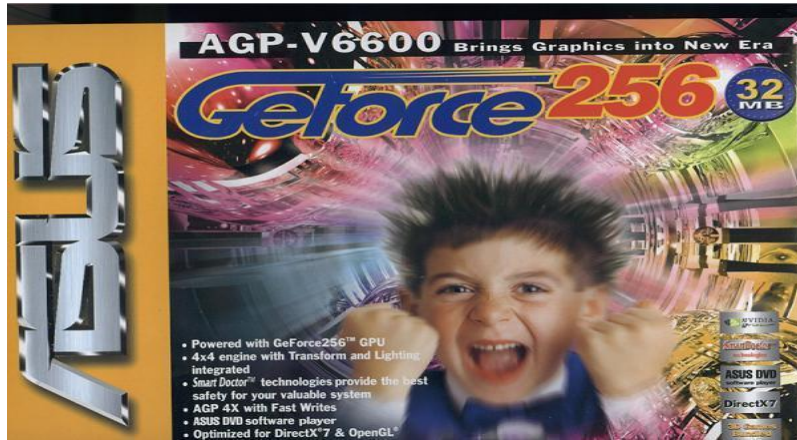
+



Razões do sucesso

- GPU
- Python
- Grande quantidade de dados rotulados
- Competições
- Software Livre

GPU – CG a serviço dos jogos



NVIDIA – GeForce 256 - 1999

Considerada a primeira GPU - 256MB!

10 milhões de polígonos por segundo!

GPGPU – a serviço dos jogos e ciência



NVIDIA – GeForce GT 8800 – 2007 – 256M/512M/1024M

4 anos de desenvolvimento, quase US\$500M, 800M transistores

Melhoramentos significativos na arquitetura – 128 processadores

GPU – a serviço do aprendizado

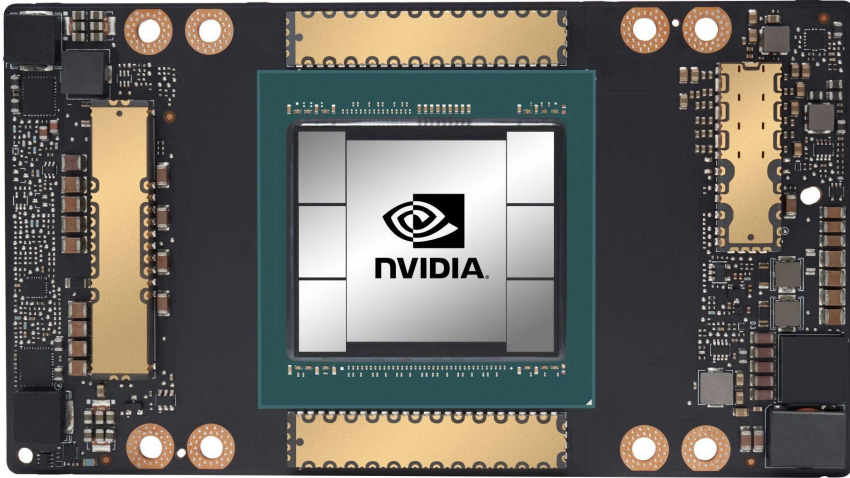


NVIDIA – GeForce Titan X – 2015 – 12GB

3072 processadores de Stream, 96 renderizadores

192 texturizadores, 24 tesseladores, 6 rasterizadores

NVIDIA A100 GPU

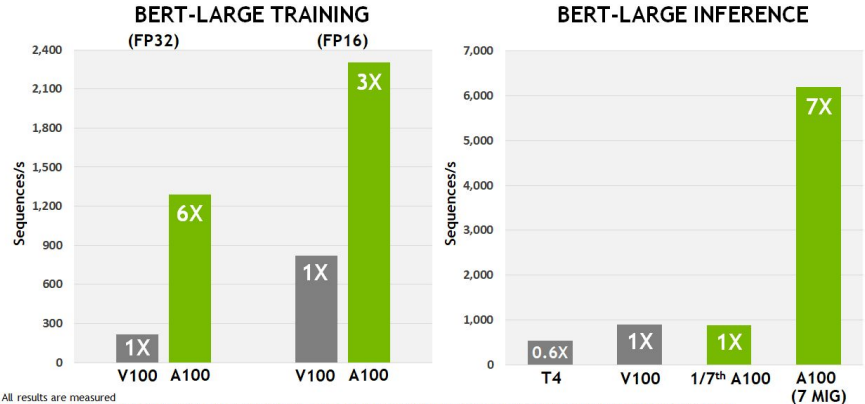


- **54 bilhões de transistores!!!**

- 8 GPCs, 8 TPCs/GPC, 2 SMs/TPC, 16 SMs/GPC, 128 SMs per full GPU
- 64 FP32 CUDA Cores/SM, 8192 FP32 CUDA Cores per full GPU
- 4 Third-generation Tensor Cores/SM, 512 Third-generation Tensor Cores per full GPU
- 6 HBM2 stacks, 12 512-bit Memory Controllers

Fonte: nvidia-ampere-architecture-whitepaper.pdf

UNIFIED AI ACCELERATION

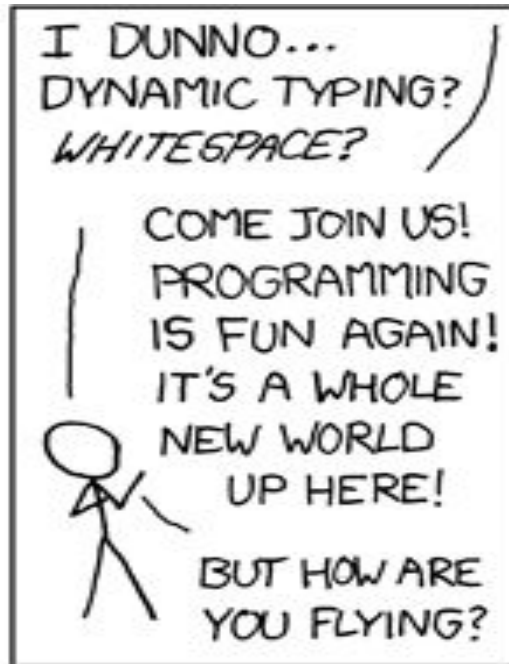


All results are measured
BERT Large Training (FP32 & FP16) measures Pre-Training phase, uses PyTorch including (2/3) Phase1 with Seq Len 128 and (1/3) Phase 2 with Seq Len 512, V100 is DGX1 Server with 8xV100, A100 is DGX A100 Server with 8xA100, A100 uses TF32 Tensor Core for FP32 training
BERT Large Inference uses TRT 7.1 for T4/V100, with INT8/FP16 at batch size 256. Pre-production TRT for A100, uses batch size 94 and INT8 with sparsity

Python - xkcd



Python - xkcd



Python

```
class SimpleCNN(nn.Module):
    def __init__(self, data_size, input_var, output_size):
        super(SimpleCNN, self).__init__()
        self.conv1 = nn.Conv2d(data_size[0], 6, kernel_size=5)
        self.pool = nn.MaxPool2d(2, 2)
        self.conv2 = nn.Conv2d(6, 10, kernel_size=5)
        self.fc1 = nn.Linear(10*((data_size[1]-4)//2-4)//2*((data_size[2]-4)//2-4)//2,100)
        self.drop1 = nn.Dropout(0.5)
        self.fc2 = nn.Linear(100, output_size)

    def forward(self, x):
        x = self.pool(F.relu(self.conv1(x)))
        x = self.pool(F.relu(self.conv2(x)))
        x = x.view(x.size(0), -1)
        x = F.relu(self.fc1(x))
        x = self.drop1(x)
        x = self.fc2(x)
        return F.log_softmax(x, dim=1)
```

Competições

- Várias conferências têm feito competições
 - NIPS
 - CVPR
 - ICDAR
 - ICIP, etc
- Hackatons de ML
- Kaggle

Kaggle

- Wikipedia's Participation Challenge – US\$ 10000
- Predicting a Biological Response – US\$ 20000
- Heritage Health Prize – US\$ 500000
- MasterCard - Data Cleansing Competition – US\$ 100000
- Galaxy Zoo - The Galaxy Challenge – US\$ 16000
- Diabetic Retinopathy Detection – US\$ 100000
- Planet: Understanding the Amazon from Space – US\$ 60000
- Passenger Screening Algorithm Challenge – US\$ 1500000
- Zillow Prize: Zillow's Home Value Prediction (Zestimate) - US\$ 1200000

Reflexões e Perspectivas

Roberto Marcondes C. Jr

É um longo caminho ao topo



AGC: It's a Long Way to the Top (BBC4 Mind Week, Full HD, 90mins, Repeat and Update)

Ada Lovelace (1840)

Had the expansion been $\Delta^n u, \Delta^{n+1} u, \Delta^{n+2} u, \dots$ instead of $\Delta^n u, \Delta^{n+1} u, \Delta^{n+2} u, \dots$, it would then be clear that if the denominator diminished without limit would approach to the denominator, the fraction itself would approach without limit to 0. But as it is, I see no reason assumed by a comparison with $(\Delta^n u)^2$.

But I do not perceive the fault itself either. When in the proof that $\Delta^n u, \Delta^{n+1} u, \Delta^{n+2} u, \dots$ diminish without limit would depend to $(\Delta^n u)^2$.

Page 115 (at the top): There is a slight misprint $C_0 R^k A^k$ instead of $C_0 R^{2k} A^k$.

Page 156: (line 9 from the top): no Circle Circle of kind (explain this step?)



(quite clear)
A. Lovelace

Gödel (1931)

Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I¹⁾.

Von Kurt Gödel in Wien.

I.

Die Entwicklung der Mathematik in der Richtung zu größerer Exaktheit hat bekanntlich dazu geführt, daß weite Gebiete von ihr formalisiert wurden, in der Art, daß das Beweisen nach einigen wenigen mechanischen Regeln erfolgen werden kann. Die umfassendsten derzeit aufgestellten formalen Systeme sind das System der Principia Mathematica (PM)²⁾ einerseits, das Zermelo-Fraenkel'sche (von J. v. Neumann weiter ausgebildete) Axiomensystem der Mengenlehre³⁾ andererseits. Diese beiden Systeme sind so weit, daß alle Sätze in der Mathematik anerkannter Beweismethoden in ihnen

Es gelten folgende Sätze:
I. Jede aus rekursiven Funktionen (Relationen) durch Einsetzung rekursiver Funktionen an Stelle der Variablen entstehende Funktion (Relation) ist rekursiv; ebenso jede Funktion, die aus rekursiven Funktionen durch rekursive Definition nach dem Schema (2) entsteht.

II. Wenn R und S rekursive Relationen sind, dann auch $R \vee S$ (daher auch $R \& S$).
III. Wenn die Funktionen $\varphi(x), \psi(y)$ rekursiv sind, dann auch die Relation: $\varphi(x) \leq \psi(y)$.

IV. Wenn die Funktion $\varphi(x)$ und die Relation $R(x, y)$ rekursiv sind, dann auch die Relationen S, T

$$S(x, y) \approx (Ez) [x \leq \varphi(z) \& R(x, y)]$$

$$T(x, y) \approx (Ex) [x \leq \varphi(x) \rightarrow R(x, y)]$$

sowie die Funktion ψ

$$\psi(x, y) = Ex [x \leq \varphi(y) \& R(x, y)],$$

wobei $z \leq F(x)$ bedeutet: Die kleinste Zahl z , für welche $F(x)$ gilt und 0, falls es keine solche Zahl gibt.

Satz I folgt unmittelbar aus der Definition von „rekursiv“. Satz II und III beruhen darauf, daß die den $\neg, \vee, =$ entsprechenden zahlentheoretischen Funktionen

$$\alpha(x), \beta(x, y), \gamma(x, y)$$

nämlich:

$$\alpha(0) = 1; \alpha(x) = 0 \text{ für } x \neq 0$$

$$\beta(0, x) = \beta(x, 0) = 0; \beta(x, y) = 1, \text{ wenn } x, y$$



Turing (1936)

ON COMPUTABLE NUMBERS, WITH AN APPLICATION TO THE ENTSCHIEDUNGSPROBLEM

By A. M. TURING.

[Received 28 May, 1936.—Read 12 November, 1936.]

The "computable" numbers may be described briefly as the real numbers whose expressions as a decimal are calculable by finite means. Although the subject of this paper is ostensibly the computable numbers, it is almost equally easy to define and investigate computable functions of an integral variable or a real or computable variable, computable predicates, and so forth. The fundamental problems involved are,

- $q_1(\xi)$
 - None \in
 - $q_1(\xi, a)$
 - $a \in$
 - not $a \in$
- $pr_1(\xi, \alpha, \beta)$
 - $pr(\xi, \beta, \alpha)$
- $ct_1(\xi, \alpha, \beta)$
 - $ct(\xi, \beta, \alpha)$
- $ct_2(\xi, \alpha, \beta, \gamma)$
 - $ct(\xi, \beta, \gamma, \alpha)$
- $e(\xi)$
 - o R
 - Not o L
 - Any R, E, L
 - None \in

$pr_1(\xi, \alpha, \beta)$. The machine prints $\alpha \beta$ at the end.

$ct_1(\xi, \alpha, \beta, \gamma)$. The machine copies down at the end first the symbols marked α , then those marked β , and finally those marked γ ; it erases the symbols α, β, γ .

From $e(\xi)$ the marks are erased from all marked symbols. $\rightarrow \in$.

A computable sequence which computes γ . Thus by the table on p. 234, and,



HBL 2015

REVIEW

Deep learning

Yann LeCun¹⁾, Yoshua Bengio²⁾ & Geoffrey Hinton³⁾

Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep learning discovers intricate structure in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer. Deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent nets have shown great power in processing sequential data such as text and speech.

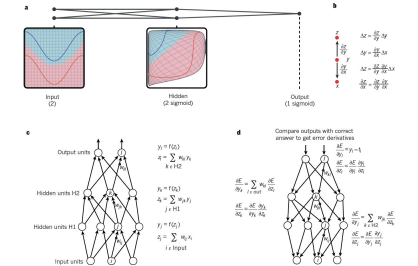


Figure 1: Multi-layer neural networks and backpropagation. A multi-layer neural network (shown by the connected dots) can distort the input space to make the classes of data (examples of which are on the red and blue lines) linearly separable. Note how a regular grid (shown on the left) is bent over to the transformed classes in the middle when the hidden units which are backpropagating gradients. At each layer, we first compute the total input z to each unit, which is a weighted sum of the outputs of the units in the layer below. Then a non-linear function f is applied to z to get the output of the unit. For simplicity, we have omitted bias terms. The non-linear function used is a rectified linear function.



Considerações finais

- Longo caminho de pesquisa fundamental e tecnológica
- Desenvolvimento e Futuro da IA
- Ética da IA
- Impacto da IA no mercado de trabalho
- Aprendendo sobre IA





Lembretes finais

Avaliação

Questões de múltipla escolha disponíveis a partir de segunda-feira: link para formulário na página do curso

<https://cursosextensao.usp.br/course/view.php?id=3867>

⇒ submissão até 12 de abril

⇒ nota mínima 7 para ser aprovado no curso e receber certificado

Bate-papos de IA

Atividade extra: bate-papos diários com pesquisadores renomados de IA da USP

⇒ 8 a 12 de abril, das 14 às 15h

⇒ Sala virtual do Google Meet:

<https://meet.google.com/mnw-qotv-vzo>

(usuários aceitos até o limite da sala)

Curso de extensão em Inteligência Artificial

