



Marcelo Filipak
C.E.O. da FiliArt, Diretor de Inovação da AFRAC

I.A. NA PRÁTICA

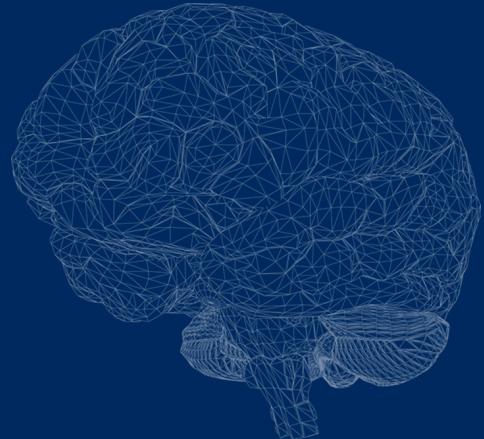
cases em automação comercial



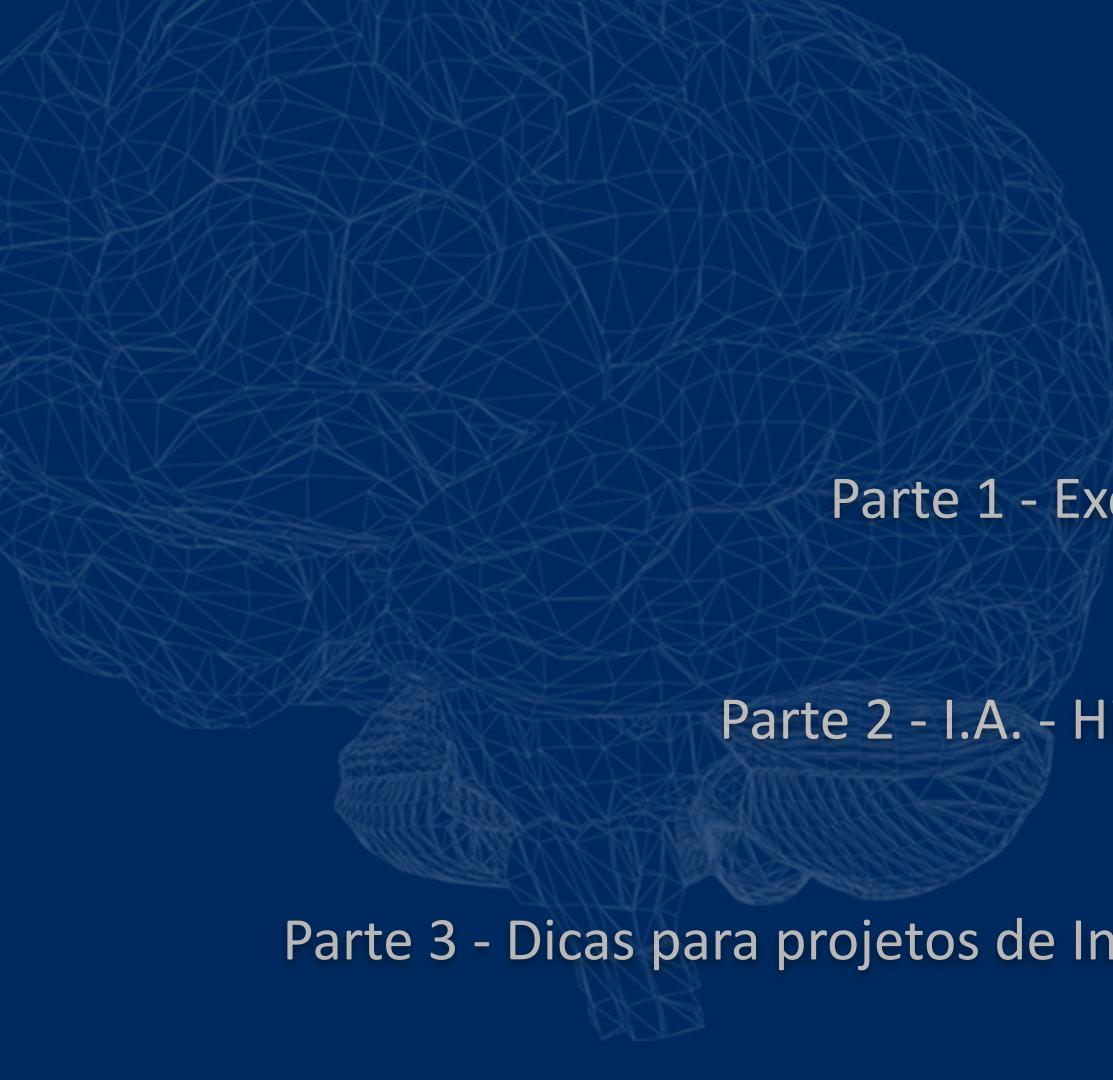
I.A. NA PRÁTICA

cases em automação comercial

Marcelo Filipak, Engenheiro e Cientista de dados.



- . Treino redes neurais há 7 anos.
- . Me especializei em criar modelos.
- . Desenvolvo projetos de modelos de deep learning com aplicações na área de saúde, segurança pública, controle de tráfego, logística, inteligência fiscal, controle aduaneiro, controle de qualidade ...
- . Trabalhei com hardware e software para automação comercial de 1990 a 2018 na BEMATECH e na TOTVS.
- . CEO na FiliArt - Projetos e Advisory em I.A

A large, semi-transparent graphic of a human brain is positioned on the left side of the slide. The brain is rendered with a complex network of white lines forming a mesh against a dark blue background, giving it a digital or scientific appearance.

Parte 1 - Exemplos, projetos e aplicações

Parte 2 - I.A. - História, fundamentos e futuro

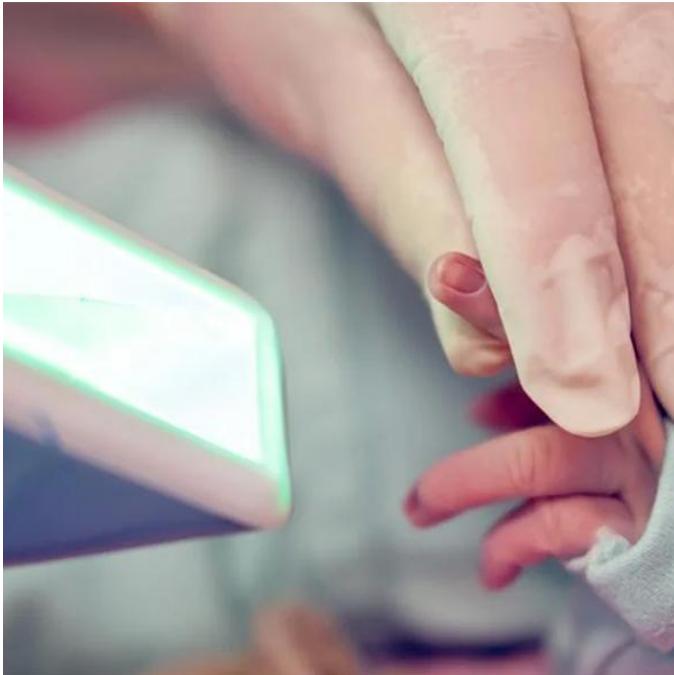
Parte 3 - Dicas para projetos de Inteligência Artificial e o futuro

—

Biometric protection
for children



Projeto - INFANT.ID



INFANT.ID

Construindo um mundo mais seguro para as nossas crianças

[CONHEÇA A PLATAFORMA INFANT.ID](#)



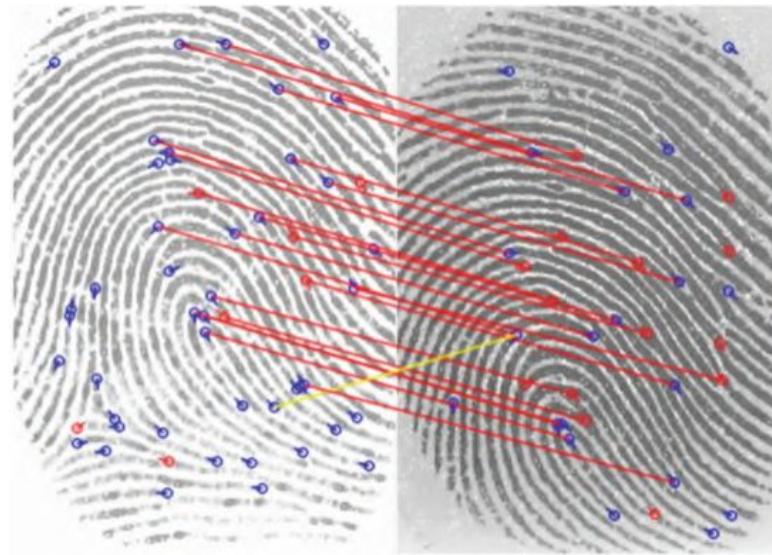
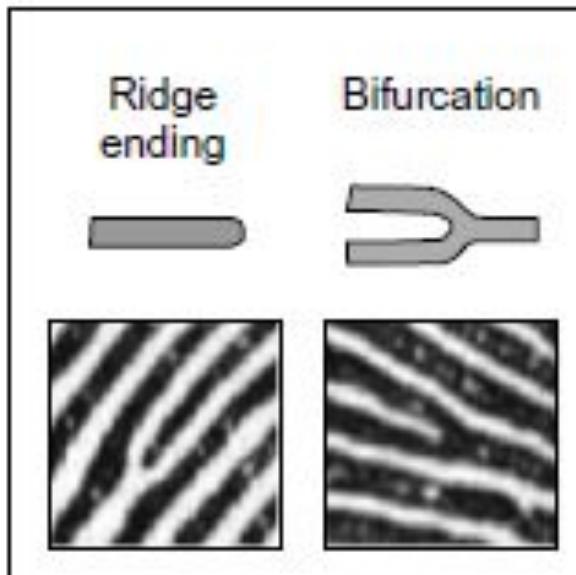
Projeto - INFANT.ID



- . Vincula a mãe com a criança.
- . Valida a digital da criança na saída da maternidade.
- . Autentica até mesmo o indivíduo adulto com a imagem de recém nascido.
- . Uso intenso de modelos de visão computacional.
- . Mais de 100.000 crianças registradas em diversas cidades.



Projeto - INFANT.ID



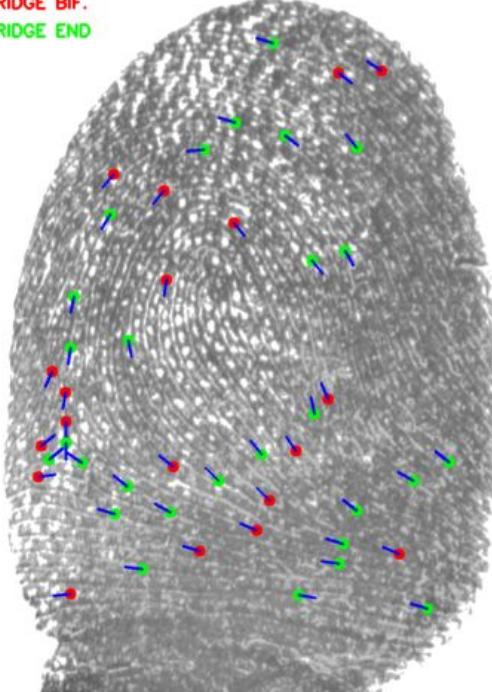
Para adultos é uma tecnologia madura e padronizada.



Projeto - INFANT.ID

MINUTIAE LABEL

- RIDGE BIF.
- RIDGE END



MINUTIAE PREDICTION

- RIDGE BIF.
- RIDGE END

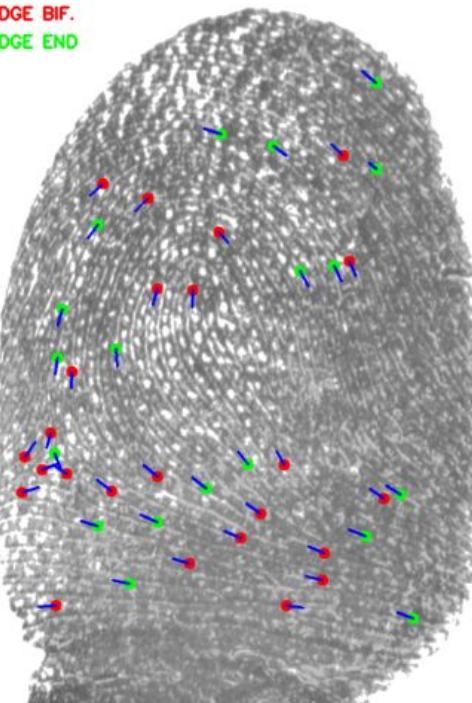
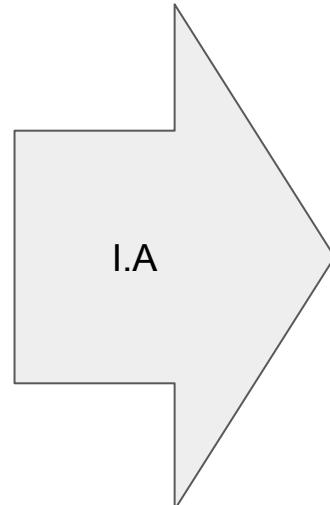
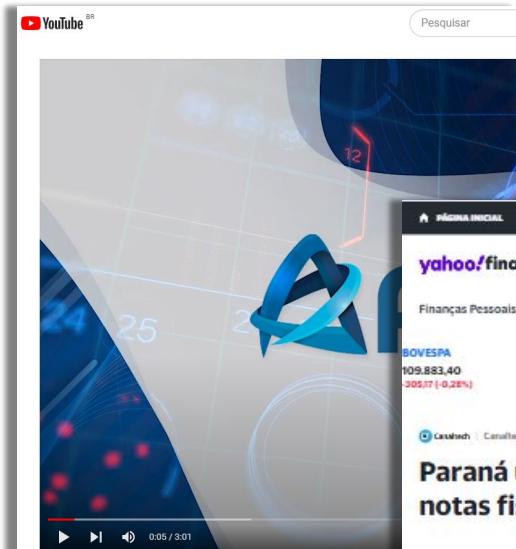


Imagen “boa” de
recém nascido.



Projeto - AFRAC Nota Certa



YouTube Pesquisar

AFRAC E SEFA-PR firmam um acordo de cooperação técnica!

0:05 / 3:01

AFRAC E SEFA-PR firmam um acordo de cooperação técnica!

Paraná usará inteligência artificial para preencher e corrigir notas fiscais

Kaique Lima | 25 de agosto de 2022 às 15h20

Finanças Pessoais | Investimentos | Empreendedorismo | Trabalho | Líderes | Agro | ...

BOVESPA: 109.883,40 (-305,77 [-0,28%])

Merval: 38.390,84 (+233,89 [+0,61%])

MXX: 51.020,65 (+221,19 [+0,44%])

PETROLEO CRU: 74,30 (+0,05 [+0,07%])

Yahoo! Finanças

Buscar por notícias, códigos de negociação ou empresas

Paraná usará inteligência artificial para preencher e corrigir notas fiscais

Por Kaique Lima | Editado por Cláudio Yuge | 25 de Agosto de 2022 às 15h20

compartilhar



Elements/twenty20photos

A Secretaria da Fazenda do Paraná (SEFA/PR) firmou uma parceria com a Associação Brasileira de Tecnologia para o Comércio e Serviços (AFRAC) para o desenvolvimento de um projeto-piloto que envolve o uso da inteligência artificial para preencher e corrigir notas fiscais. O objetivo do acordo é reduzir a complexidade fiscal para os varejistas e para o fisco.

Comer carne torrada a ouro faz mal

1. Bear market do Bitcoin está nos "estágios finais", diz analista da Bloomberg

2. Petrobras reduz diesel e gasolina nas refinarias após meses sem reajustes



Projeto - AFRAC Nota Certa

AFRAC NOTA CERTA

Descrição
Feijão Preto

Clear Submit

NCMs

NCM	Porcentagem
07133319	33%
07133399	14%
07133329	5%
20055100	5%
21069090	4%

Descrição do NCM

Produtos hortícolas, plantas, raízes e tubérculos, comestíveis. Legumes de vagem, secos, em grão, mesmo pelados ou partidos. Feijões (*Vigna* spp., *Phaseolus* spp.) -- Feijão comum (*Phaseolus vulgaris*) -- Preto Outros

https://huggingface.co/spaces/AFRAC/NCM_DEMO

https://github.com/mfilipak/AFRAC_IA



Projeto - AFRAC Nota Certa



```
import requests

url = "https://afrac-ncm-demo.hf.space/api/predict"

payload = {
    "data": ["Coca-Cola PET 21"]
}

headers = {
    "Content-Type": "application/json"
}

response = requests.post(url, json=payload, headers=headers)
result = response.json()

print(result)
```



Projeto - AFRAC Nota Certa



```
import requests

url = "https://afrac-ncm-demo.hf.space/api/predict"

payload = {
    "data": ["Coca-Cola PET 21"]
}

headers = {
    "Content-Type": "application/json"
}

response = requests.post(url, json=payload,
headers=headers)
result = response.json()

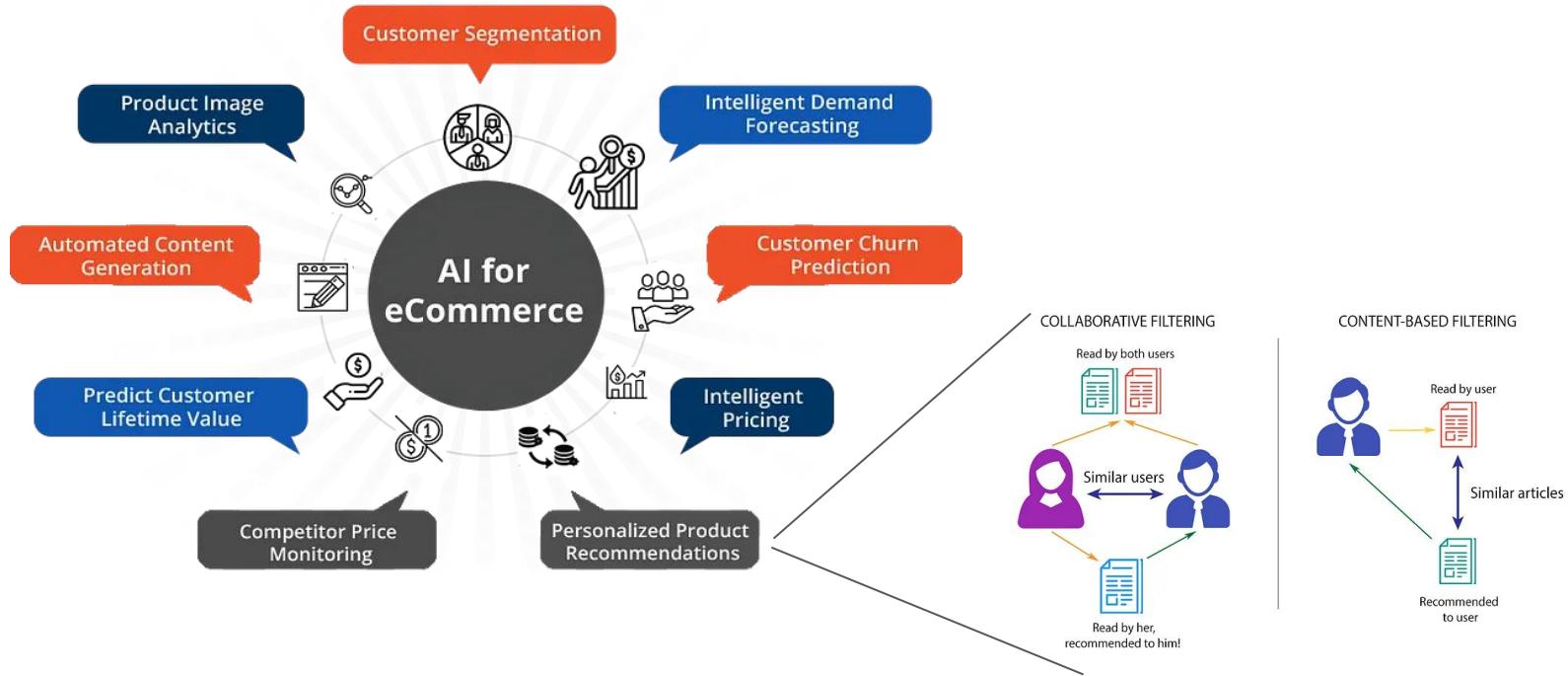
print(result)
```

The terminal window shows the command `result` followed by the API response. The response is a JSON object containing an array of data objects, each with a label and confidence score, and some descriptive text at the bottom.

```
{"data": [{"label': '22021000',  
'confidences': [{"label': '22021000', 'confidence': 0.5184096693992615},  
{'label': '22011000', 'confidence': 0.03791089728474617},  
{'label': '22030000', 'confidence': 0.03218604251742363},  
{'label': '21069000', 'confidence': 0.02866430953145027},  
{'label': '22029900', 'confidence': 0.02776830829679966}]],  
'Bebidas, líquidos alcoólicos e vinagres.Águas, incluindo as águas minerais e as águas (sumos) de fruta ou de produtos hortícolas da posição 20.09.- Águas, incluindo a  
'is_generating': False,  
'duration': 0.09473180770874023,  
'average_duration': 0.07237190791693515}
```



I.A. no e-commerce



<https://medium.com/@mahimaphalkey/ml-use-cases-in-e-commerce-ec507b5bd7fe>

CHAT-BOTS

Coletando dados
pela interação.

Extender o
suporte para 24x7

Recomendação de
produtos,
“upselling”

Suporte aos
vendedores.

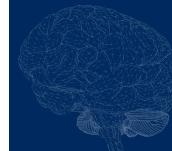
Facilitar tracking dos
pedidos.

Organizar FAQ e
perguntas
relacionadas.
perplexity.ai

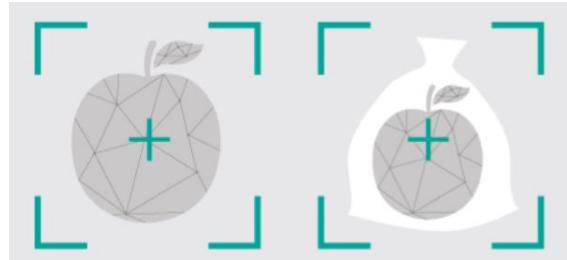


Encontrar uma loja
física.

Coletar feedbacks.



Balanças com I.A.



Contagem de produtos e lacunas

How to Use the Empty spaces in a supermarket hanger Detection API

Use this pre-trained Empty spaces in a supermarket hanger computer vision model to retrieve predictions with our hosted API or deploy to the edge. [Learn More About Roboflow Inference](#)

Switch Model:

v1.6 empty-spaces-in-a-supermarket-hanger-1upsp/16

Trained On: empty-spaces-in-a-supermarket-hanger-1upsp 1997 Images [View Version](#)

Model Type: Roboflow 2.0 Object Detection (Fast)

Checkpoint: COCOv6n

mAP 54.8% | Precision 75.4% | Recall 46.7% | [View Model Graphs](#)

Samples from Test Set

Upload Image or a Video File

Drop files here or [Select File](#)

Paste YouTube or Image URL

[Paste a link...](#)

Try With Webcam

</> Try On My Machine

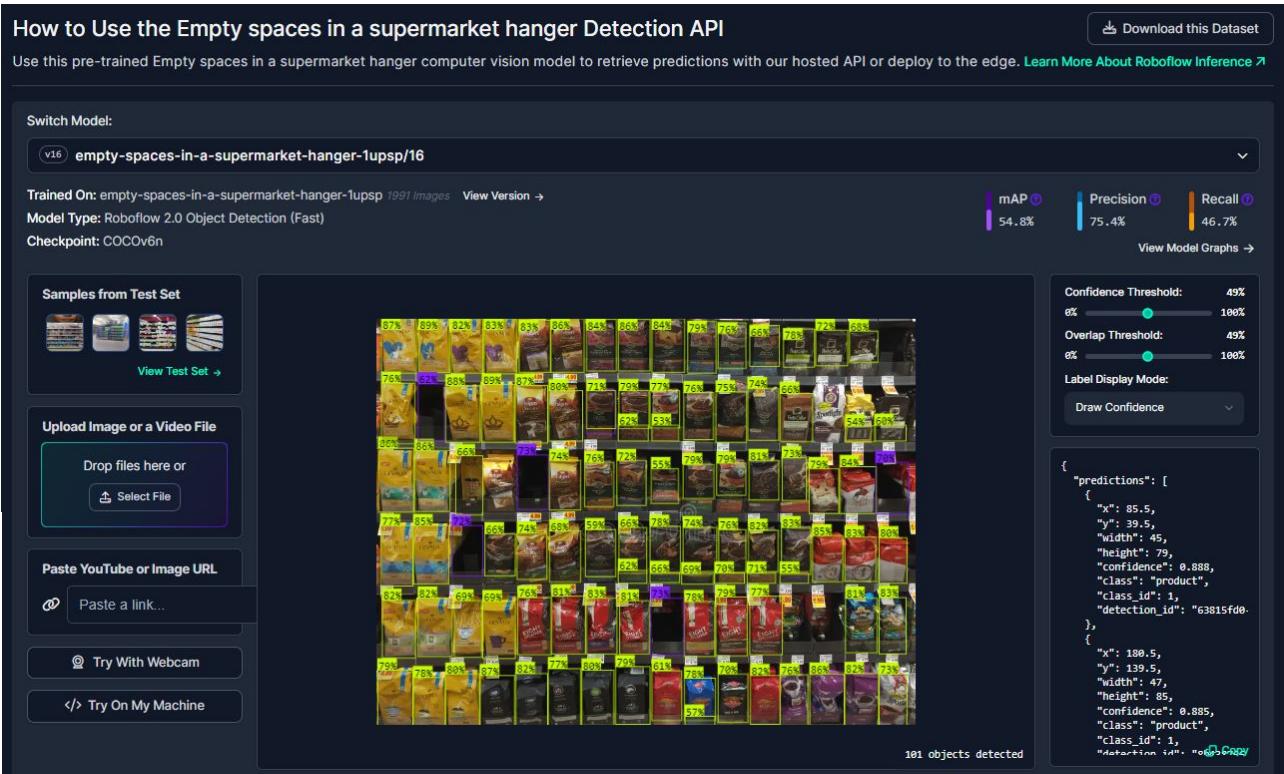
Confidence Threshold: 49% | 0% 100%

Overlap Threshold: 49% | 0% 100%

Label Display Mode: Draw Confidence

{ "predictions": [{ "x": 85.5, "y": 39.5, "width": 45, "height": 79, "confidence": 0.888, "class": "product", "class_id": 1, "detection_id": "63815fd0"}, { "x": 180.5, "y": 139.5, "width": 47, "height": 85, "confidence": 0.885, "class": "product", "class_id": 1, "detection_id": "63815fd1"}] }

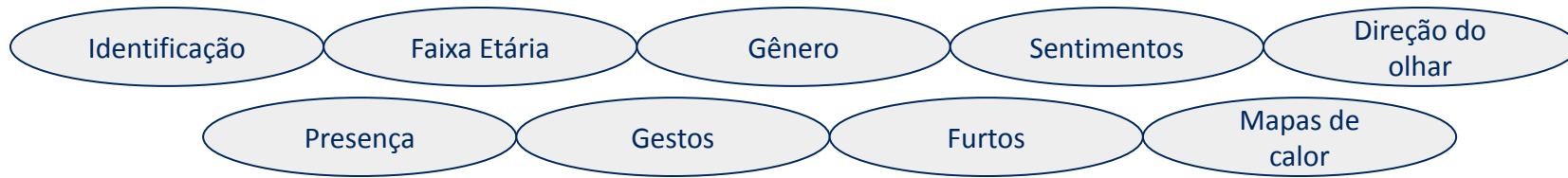
101 objects detected



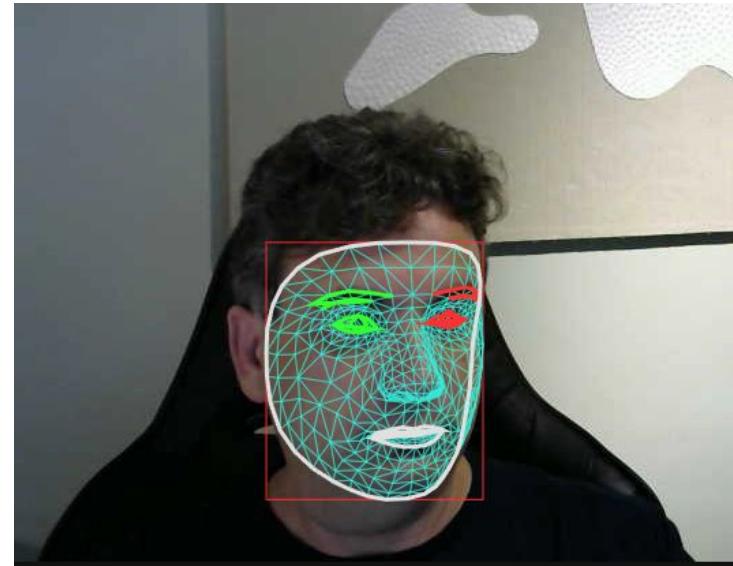
bit.ly/4c1NLv0



Modelos faciais e corporais



<https://bit.ly/3Y3kvjX>



Provador virtual



A screenshot of the DressX AI virtual try-on tool. At the top, a user's photo shows them in a grey t-shirt and orange pants, standing in front of the same "Avengers: Endgame" premiere backdrop. Below the photo are three buttons: "SAVE", "CHANGE PHOTO", and "SIZE GUIDE". To the right of the photo are several outfit suggestions in boxes: a red and blue patterned dress, a white t-shirt, black pants, and orange pants. At the bottom is a large "DRESSX ME" button and the text "BETWEEN 1MB AND 10MB".



Provador virtual

A screenshot of the DressX.ai website. At the top, the URL https://dressx.ai is visible in the browser bar. Below the photo, there are several options for different garments:

- A tan-colored, draped jumpsuit.
- A vibrant, multi-colored, ruffled dress.
- A black t-shirt.
- A pair of white pants with a circular icon containing an 'X' over them, indicating they are not selected.

At the bottom of the interface, there are buttons for "SAVE", "CHANGE PHOTO", "? SIZE GUIDE", and a large black button labeled "DRESSX ME".

Provador virtual



https://dressx.ai

↓ SAVE ↲ CHANGE PHOTO ? SIZE GUIDE

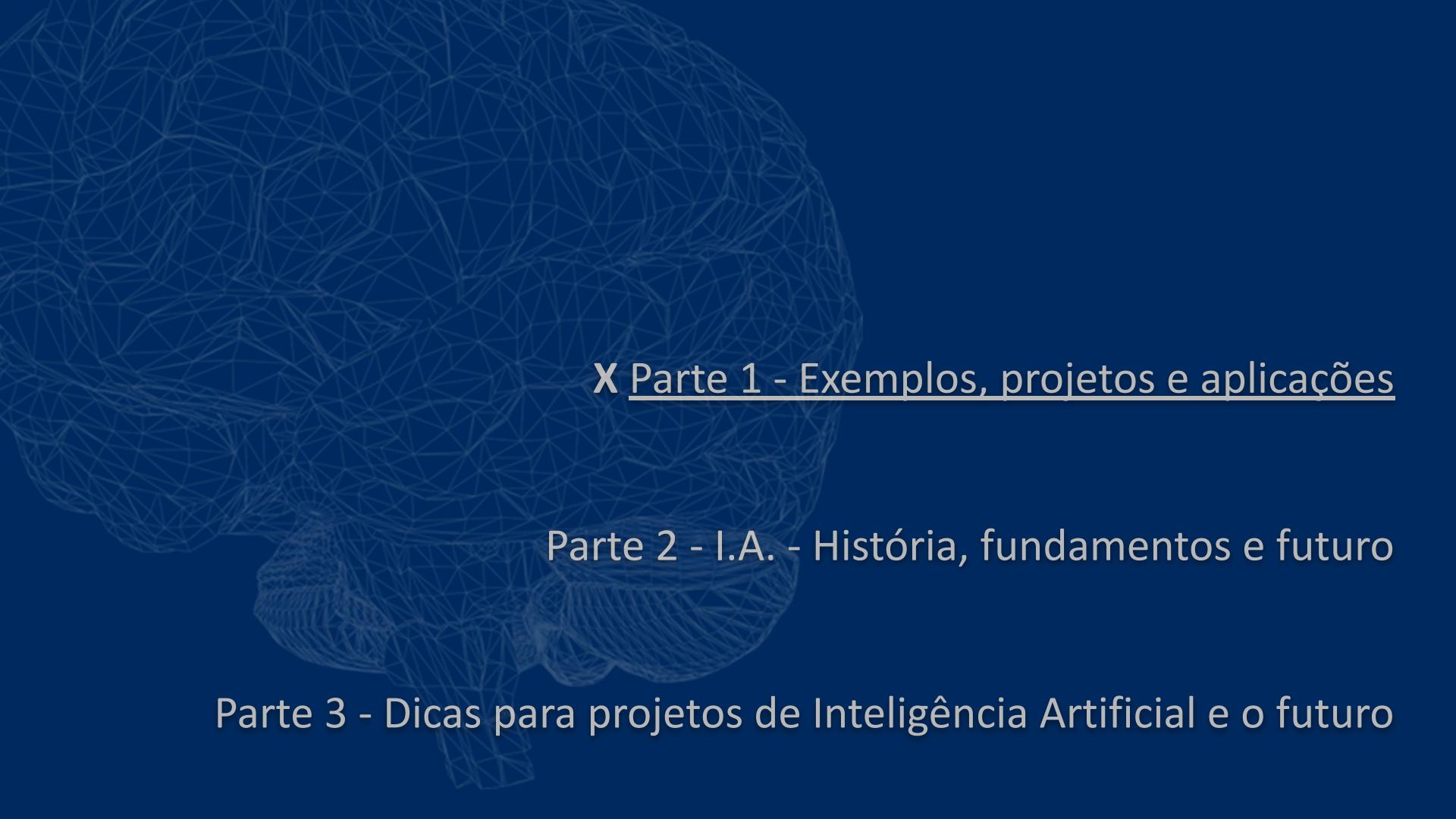
DRESSX ME

BETWEEN 1MB AND 10MB

Metaverso



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



X Parte 1 - Exemplos, projetos e aplicações

Parte 2 - I.A. - História, fundamentos e futuro

Parte 3 - Dicas para projetos de Inteligência Artificial e o futuro

<http://introtodeeplearning.com/>



Schedule

New 2024 lectures, slides, and labs!
Taught in-person at MIT — open-sourced to the world.

 Intro to Deep Learning Lecture 1 Apr. 29, 2024 [Slides] [Video]	 Deep Sequence Modeling Lecture 2 May 6, 2024 [Slides] [Video]	 Intro to TensorFlow; Music Generation Software Lab 1 [Code]
 Deep Computer Vision Lecture 3 May 13, 2024 [Slides] [Video]	 Deep Generative Modeling Lecture 4 May 20, 2024 [Slides] [Video]	 Facial Detection Systems Software Lab 2 [Paper] [Code]
 Deep Reinforcement Learning Lecture 5 May 27, 2024 [Slides] [Video]	 New Frontiers Lecture 6 June 3, 2024 [Slides] [Video]	 Large Language Models Software Lab 3 [Code] coming soon!
 Generative AI for Media Lecture 7 June 10, 2024 [Slides] [Video]	 Stories from Models in the Wild Lecture 8 June 17, 2024 [Slides] [Video]	 Final Project Work on final projects
 Final Project Work on final projects June 24, 2024	 Project Presentations Pitch your ideas! June 24, 2024	 Awards Ceremony Final awards and celebration!

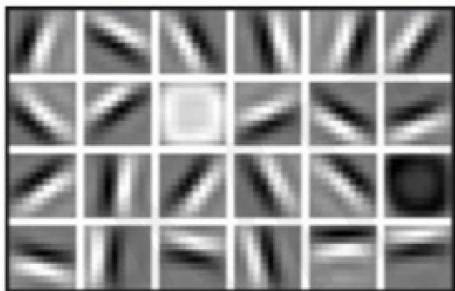


Why Deep Learning?

Hand engineered features are time consuming, brittle, and not scalable in practice

Can we learn the **underlying features** directly from data?

Low Level Features



Lines & Edges

Mid Level Features



Eyes & Nose & Ears

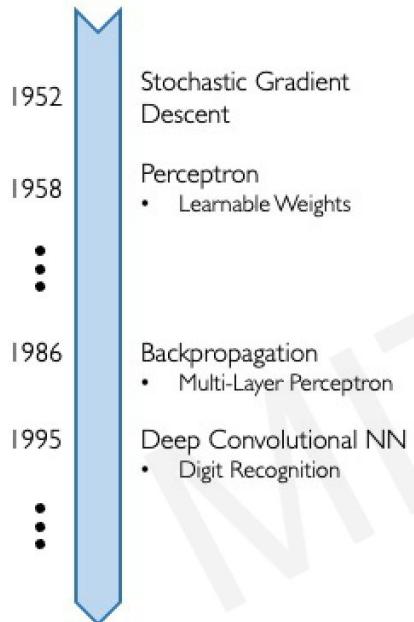
High Level Features



Facial Structure



Why Now?

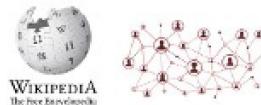


Neural Networks date back decades, so why the dominance?

1. Big Data

- Larger Datasets
- Easier Collection & Storage

IM²GENET



2. Hardware

- Graphics Processing Units (GPUs)
- Massively Parallelizable

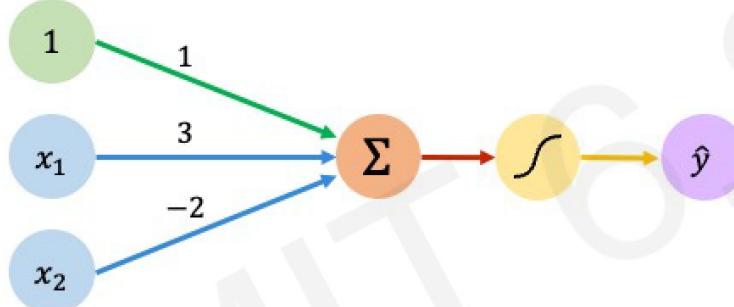


3. Software

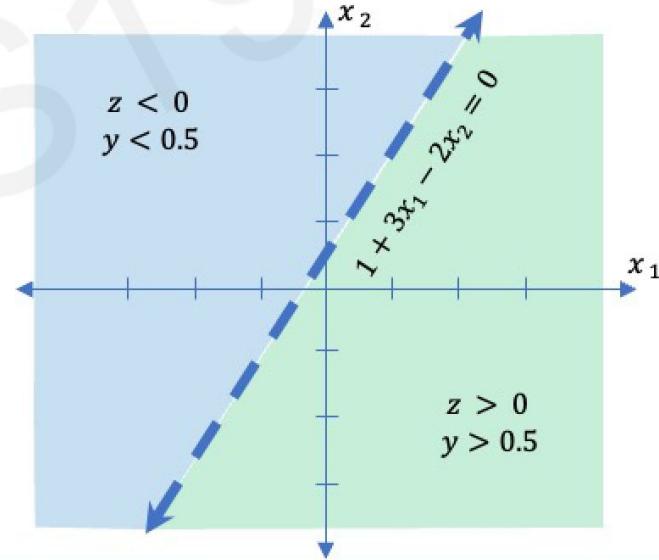
- Improved Techniques
- New Models
- Toolboxes



The Perceptron: Example



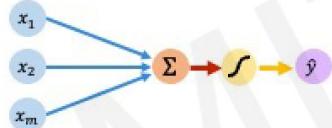
$$\hat{y} = g(1 + 3x_1 - 2x_2)$$



Core Foundation Review

The Perceptron

- Structural building blocks
- Nonlinear activation functions



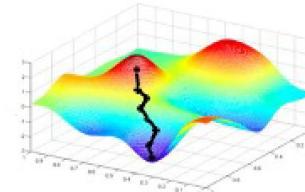
Neural Networks

- Stacking Perceptrons to form neural networks
- Optimization through backpropagation

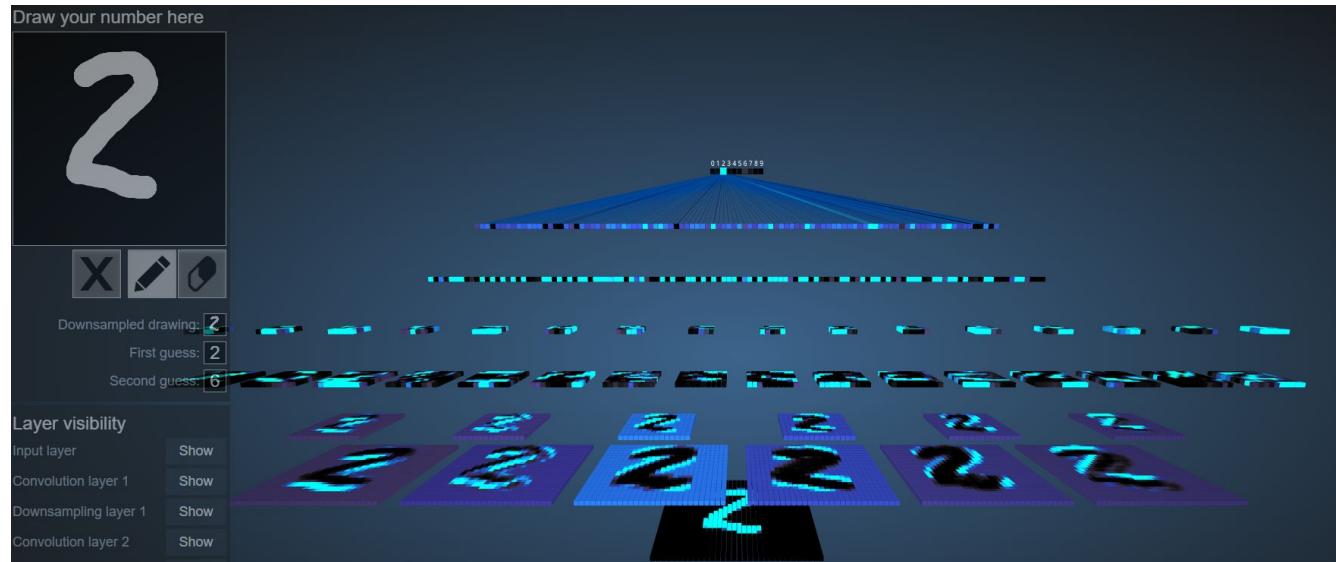


Training in Practice

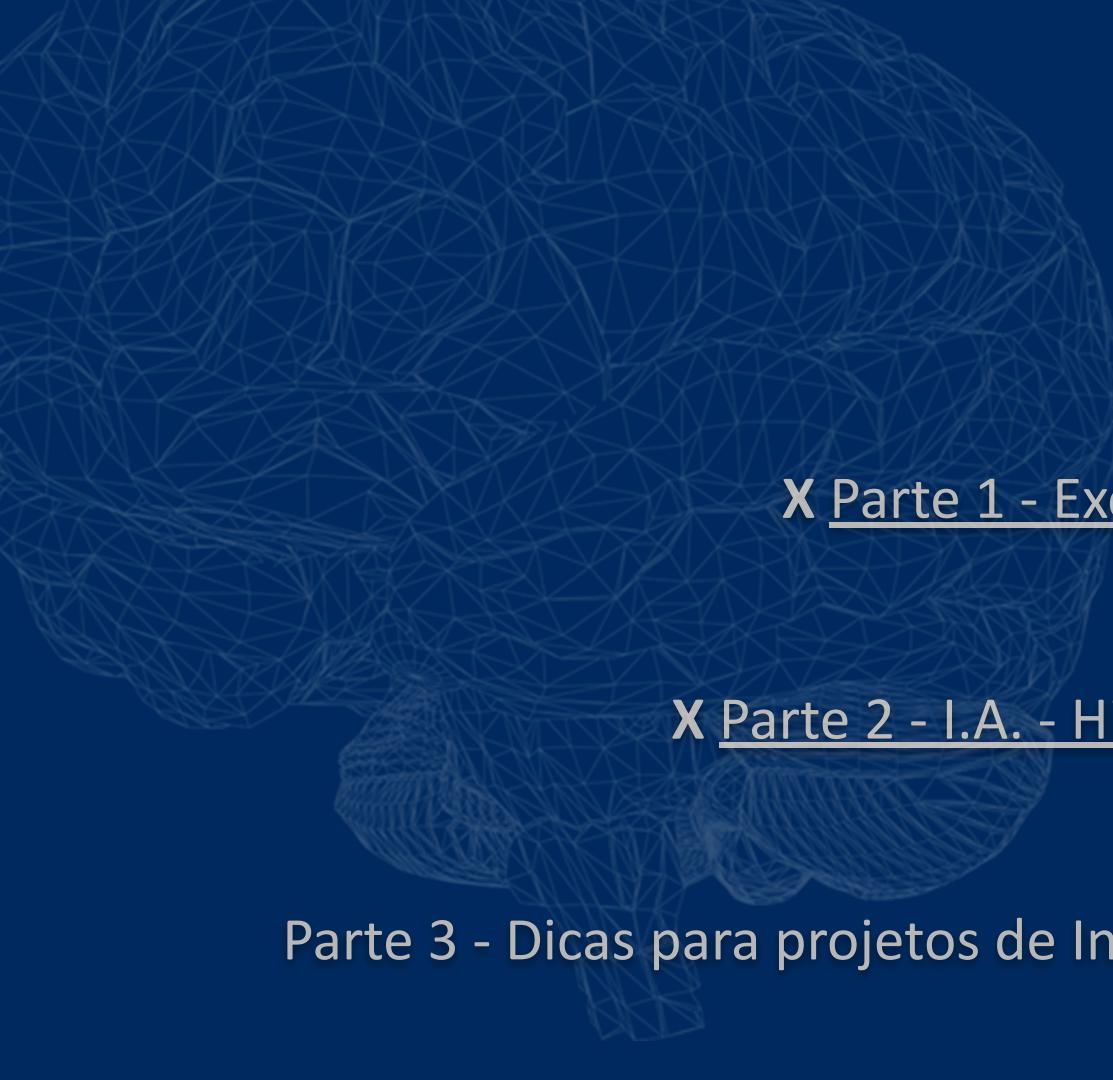
- Adaptive learning
- Batching
- Regularization



Visualização de uma rede neural



https://adamharley.com/nn_vis/cnn/3d.html

A large, semi-transparent graphic of a human brain is visible in the background, composed of a complex network of white lines forming triangles and polygons against a dark blue background.

X Parte 1 - Exemplos, projetos e aplicações

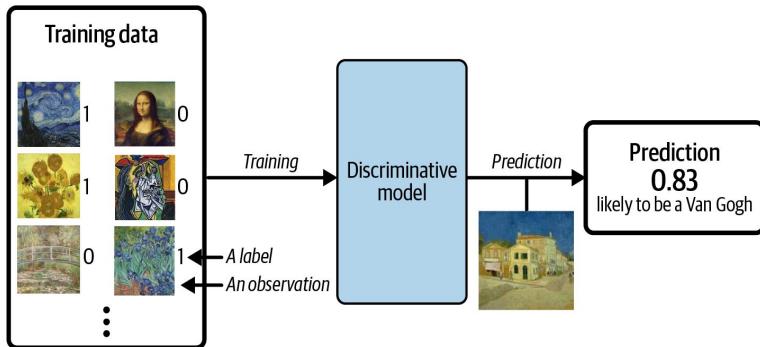
X Parte 2 - I.A. - História, fundamentos e futuro

Parte 3 - Dicas para projetos de Inteligência Artificial e o futuro



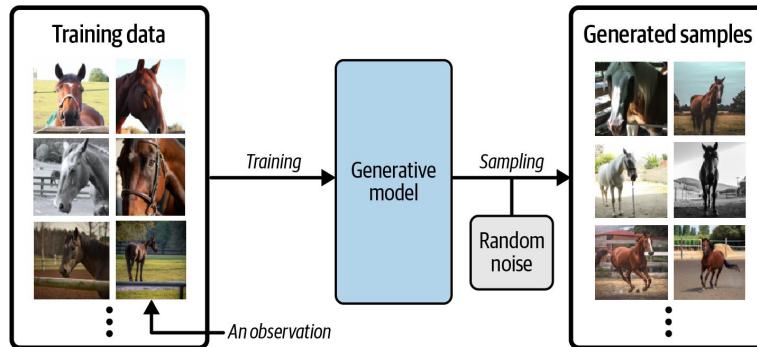
1 - Que tipo de I.A. eu preciso?

I.A. Discriminativa



Classificação, Regressão
Reconhecimento de imagens
Emoções em texto
Reconhecimento Facial
Contagem de objetos
...

I.A. Generativa



Tradução de texto
Geração de imagens
Chatbots (chat-gpt)
Geração de código
...

São especializadas em problemas diferentes, embora sua arquitetura interna tenha mais semelhanças que diferenças.



Que tipo de modelo eu preciso?

Shallow Learning (Ex: Random Forest)

. Classificação e Regressão

Simples: Previsão de vendas, diagnóstico de doenças.

. Análise de Crédito:

Avaliação de risco de crédito.

. Detecção de Fraudes:

Identificação de transações fraudulentas.

. Marketing Direto:

Segmentação de clientes.

Deep Learning

. Reconhecimento de Imagens:

Detecção de objetos, reconhecimento facial.

. Processamento de Linguagem Natural (NLP):

Tradução automática, análise de sentimento.

. Reconhecimento de Voz:

Conversão de fala para texto, assistentes virtuais.

Generative AI

. Criação de Conteúdo:

Geração de texto, imagens, música.

. Design e Arte:

Criação de designs gráficos, arte digital.

. Modelagem e Simulação:

Simulação de cenários, modelagem de personagens em jogos.

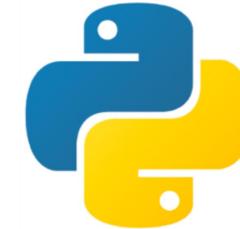
. Assistência Criativa:

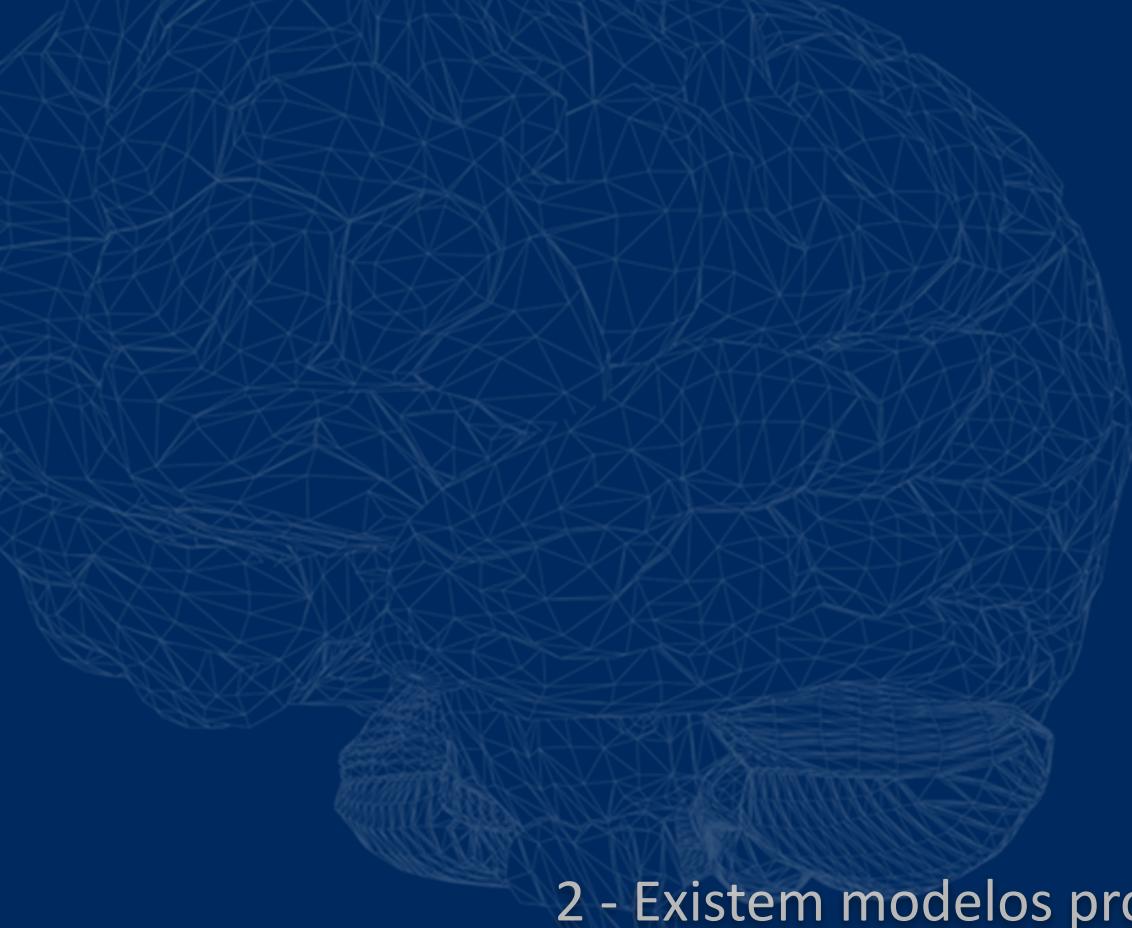
Ferramentas de apoio à escrita, composição musical.



Que tipo de ambiente eu preciso?

Google
colab





2 - Existem modelos prontos para o meu problema?
Tem muito modelo pré-treinado mas é provável que você precise melhorá-lo.

Modelos prontos



HUGGING FACE



kaggle

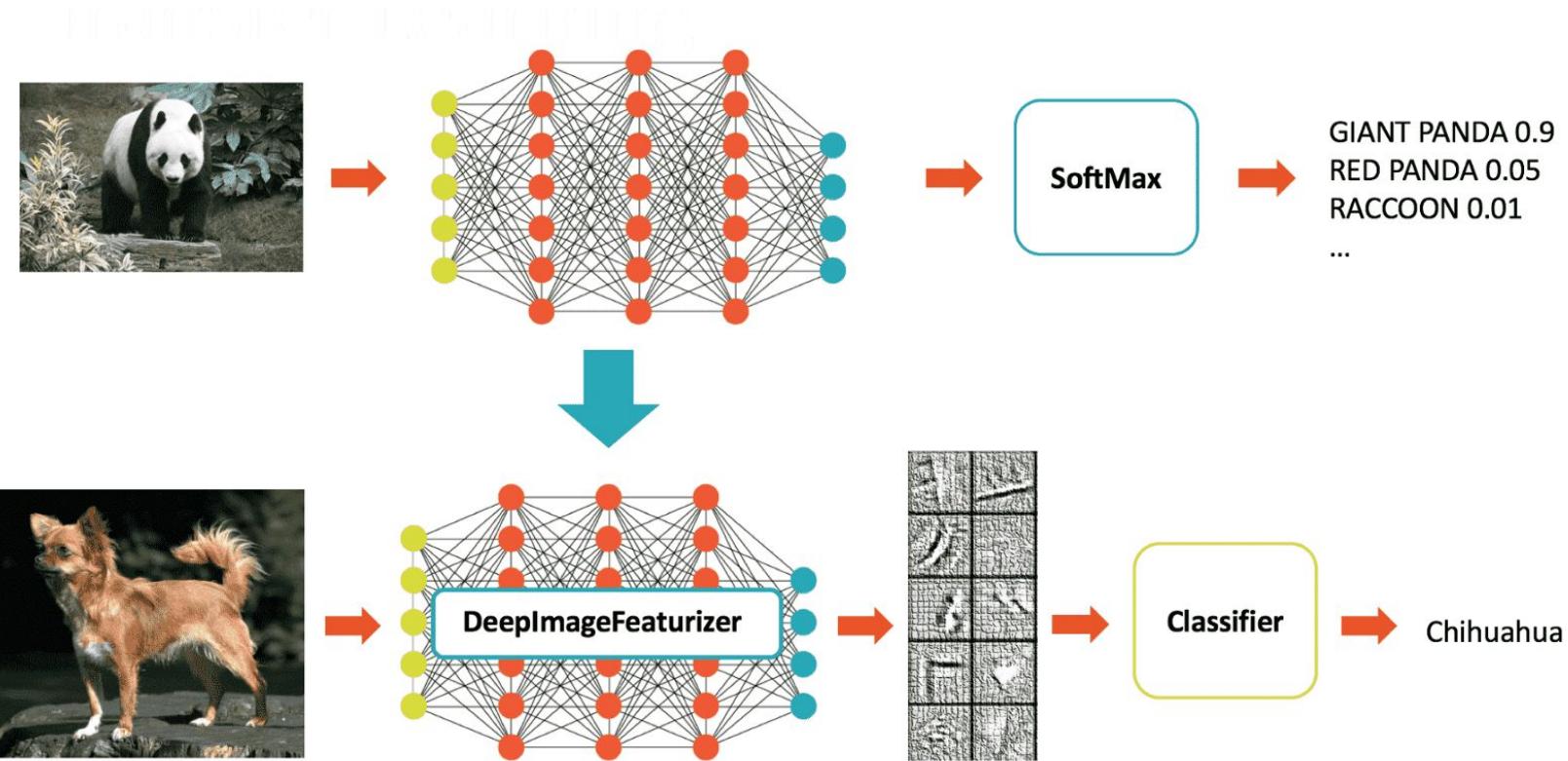
1 - Procure um modelo pronto para o seu problema (alguns como detecção de face, biometria facial, landscapes faciais, detecção de objetos, ...) são treinados em big data e seguramente vão ser melhores que treinar do zero.

2 - Se ele estiver no caminho, é possível fazer um “fine tuning” ou “transfer learning”. Aqui você pode congelar parte do modelo e treinar somente uma parte pequena e com relativamente poucos dados.

3 - Decida se está bom o suficiente ou se você vai treinar um modelo próprio. Nesse cenário você vai precisar obter, tratar ou construir dados e prepará-los para treinar modelos.



“Transfer learning”





3 - Vou precisar de mais dados?
COLETA, ORGANIZAÇÃO E LIMPEZA DOS DADOS.

Curando o dataset

- 1** - Treine ou obtenha um modelo que performe razoavelmente nos seus dados.
- 2** - Use esse modelo para ajudá-lo na anotação/organização dos dados. Por exemplo, separando casos fáceis de difíceis. Revise os dados difíceis e retreine.
- 3** - Use esse modelo para coletar dados onde o modelo tem baixa confiança.





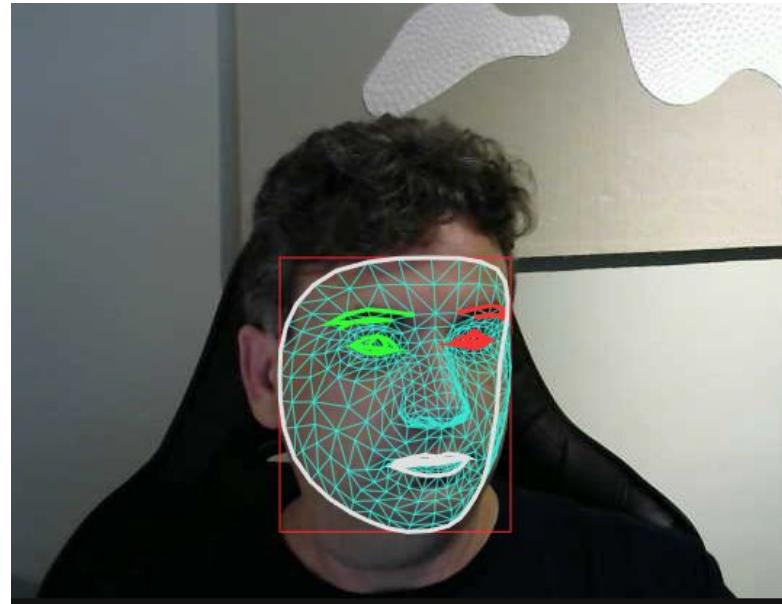
4 - Preciso rodar os modelos na nuvem ou num servidor com GPU?

Treinar uma rede é pesado, usá-la em geral é leve.

Considerações práticas

- . O mais comum é rodar em web services (facilidade de prototipação)
- . Muitos modelos complexos rodam até num smartphone. (Exceção: LLMs)
 - . Mesmo LLMs podem rodar num desktop.

<https://bit.ly/3Y3kvjX>

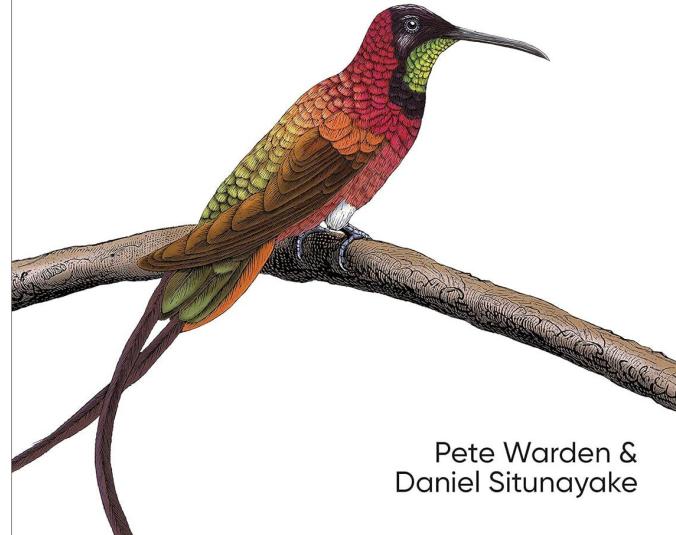




O'REILLY®

TinyML

Machine Learning with TensorFlow Lite on
Arduino and Ultra-Low Power Microcontrollers



Pete Warden &
Daniel Situnayake



E O FUTURO?

How Long Until Computers Have the Same Power As the Human Brain?

Lake Michigan's volume (in fluid ounces) is about the same as our brain's capacity (in calculations per second). Computing power doubles every 18 months. At that rate, you see very little progress for a long time—and suddenly you're finished.

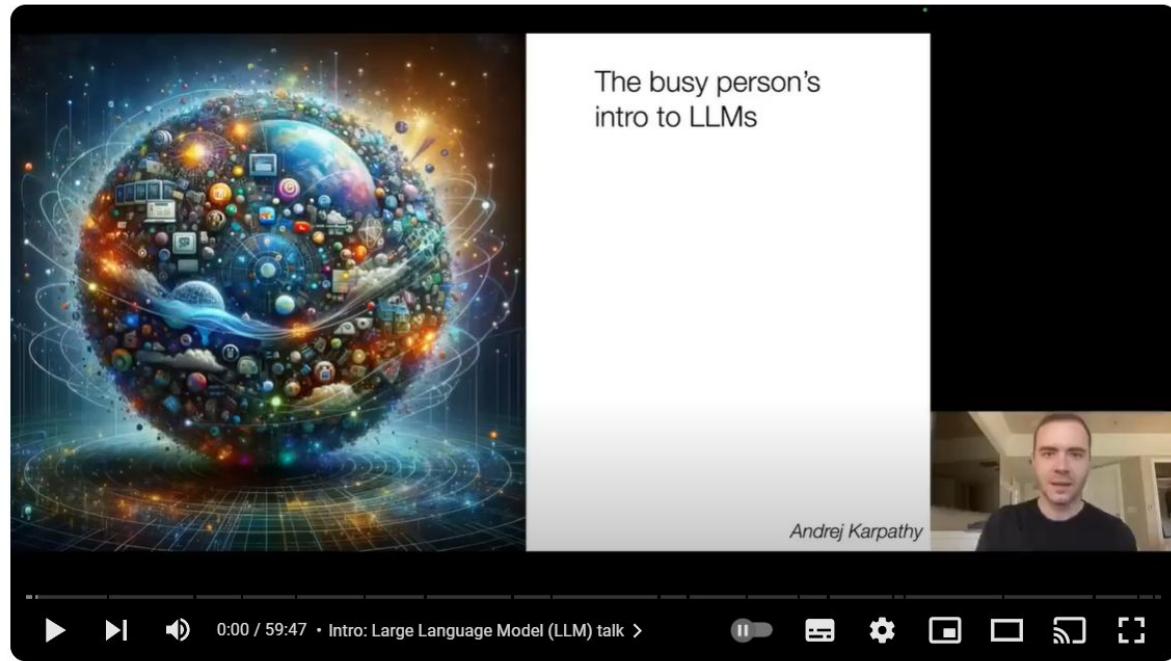


<https://www.motherjones.com/media/2013/05/robots-artificial-intelligence-jobs-automation/>



LLMs

[https://www.youtube.com/
watch?v=zjkBMFhNj_g&t](https://www.youtube.com/watch?v=zjkBMFhNj_g&t)



The busy person's intro to LLMs

Andrej Karpathy

0:00 / 59:47 • Intro: Large Language Model (LLM) talk >

[1hr Talk] Intro to Large Language Models

Andrej Karpathy
427 mil inscritos

Inscrito

59 mil

Compartilhar

...





LMSYS Chatbot Arena (Multimodal)

Chat now!

Expand to see the descriptions of 73 models

Model A

Pra quem o cachorro está olhando?

O cachorro está olhando para a pessoa sentada no sofá à direita.

Model B

Pra quem o cachorro está olhando?

Na imagem, o cachorro está olhando para a mulher que está sentada no sofá.

A is better

B is better

Tie

Both are bad





LMSYS Chatbot Arena (Multimodal)

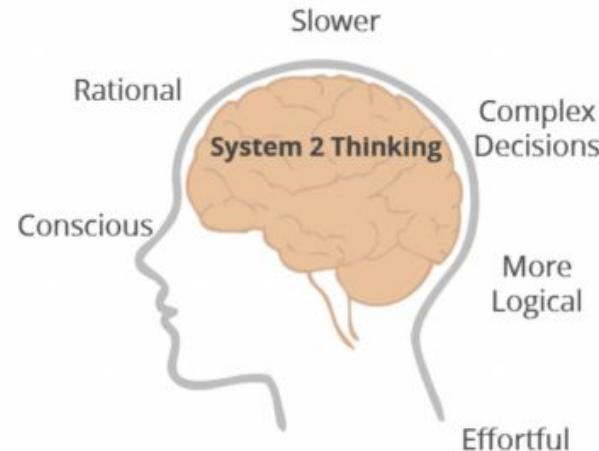
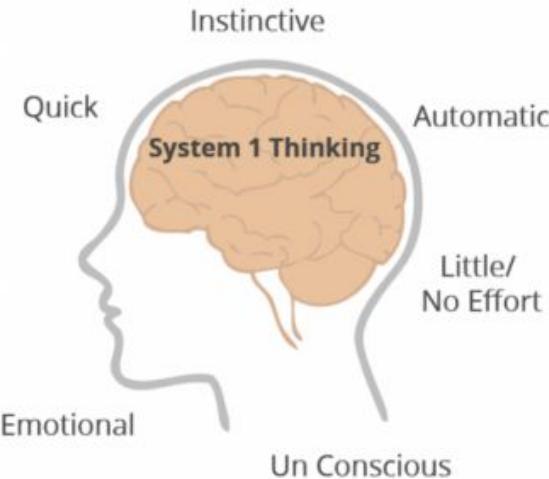
<https://chat.lmsys.org/>



Rank* (UB)	Model	Arena Score	95% CI	Votes	Organization	License	Knowledge Cutoff
1	GPT-4o-2024-05-13	1287	+3/-3	56905	OpenAI	Proprietary	2023/10
2	Claude 3.5 Sonnet	1272	+4/-4	24913	Anthropic	Proprietary	2024/4
2	Gemini-Advanced-0514	1267	+3/-3	42981	Google	Proprietary	Online
3	Gemini-1.5-Pro-API-0514	1262	+3/-3	49828	Google	Proprietary	2023/11
4	Gemini-1.5-Pro-API-0409-Preview	1258	+3/-3	55567	Google	Proprietary	2023/11

12	Nemotron-4-340B-Instruct	1208	+5/-5	19143	Nvidia	NVIDIA Open Model	2023/6
12	GLM-4-0520	1207	+5/-5	10248	Zhipu AI	Proprietary	Unknown
14	llama-3-70b-Instruct	1207	+2/-3	147550	Meta	Llama 3 Community	2023/12

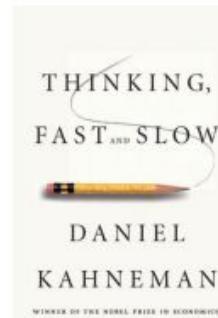




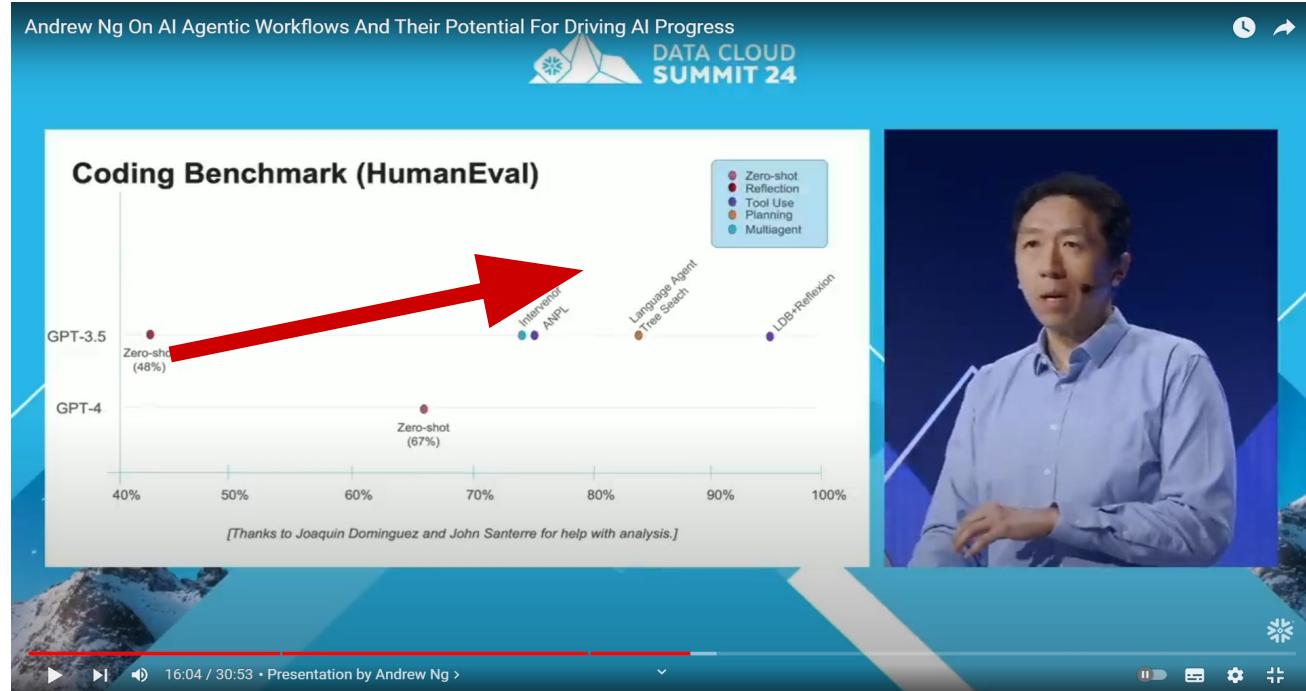
$$2 + 2 =$$

A maior parte dos modelos de I.A.
(incluindo LLMs) são análogas ao
sistema 1.

$$17 \times 24 =$$



Agentic A.I.



ChatDev: Communicative Agents for Software Development

Chen Qian* Wei Liu* Hongzhang Liu* Nuo Chen* Yufan Dang*
Jiahao Li* Cheng Yang* Weize Chen* Yusheng Su* Xin Cong*
Juyuan Xu* Dahai Li* Zhiyuan Liu** Maosong Sun**
*Tsinghua University *The University of Sydney **BUPT *Modelbest Inc.
qianc62@gmail.com liuzy@tsinghua.edu.cn sms@tsinghua.edu.cn

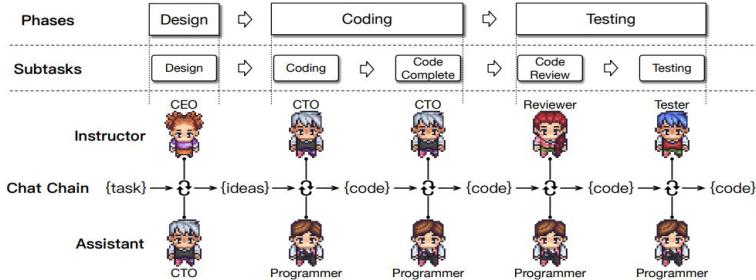


Figure 1: ChatDev, a chat-powered software development framework, integrates LLM agents with various social roles, working autonomously to develop comprehensive solutions via multi-agent collaboration.



“Just as 100 years ago electricity transformed industry after industry, AI will now do the same.” – Andrew Ng

Andrew Ng is Founder & CEO of Landing AI, Founder of deeplearning.ai, Co-Chairman and Co-Founder of Coursera, and is currently an Adjunct Professor at Stanford University. He was also Chief Scientist at Baidu Inc., and Founder & Lead for the Google Brain Project



Marcelo Filipak

mfilipak@hotmail.com

<https://www.linkedin.com/in/marcelo-filipak-97386710/>

<https://github.com/mfilipak>

I.A. NA PRÁTICA

cases em automação comercial

OBRIGADO!

