

Novas tecnologias: Oportunidades e Desafios

Inês Dutra

Departamento de Ciência de Computadores
Faculdade de Ciências, Universidade do Porto
CINTESIS@RISE and CRACS INESC TEC



FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO



Agenda



01 Novas tecnologias?

02 Aplicações de IA

03 Oportunidades

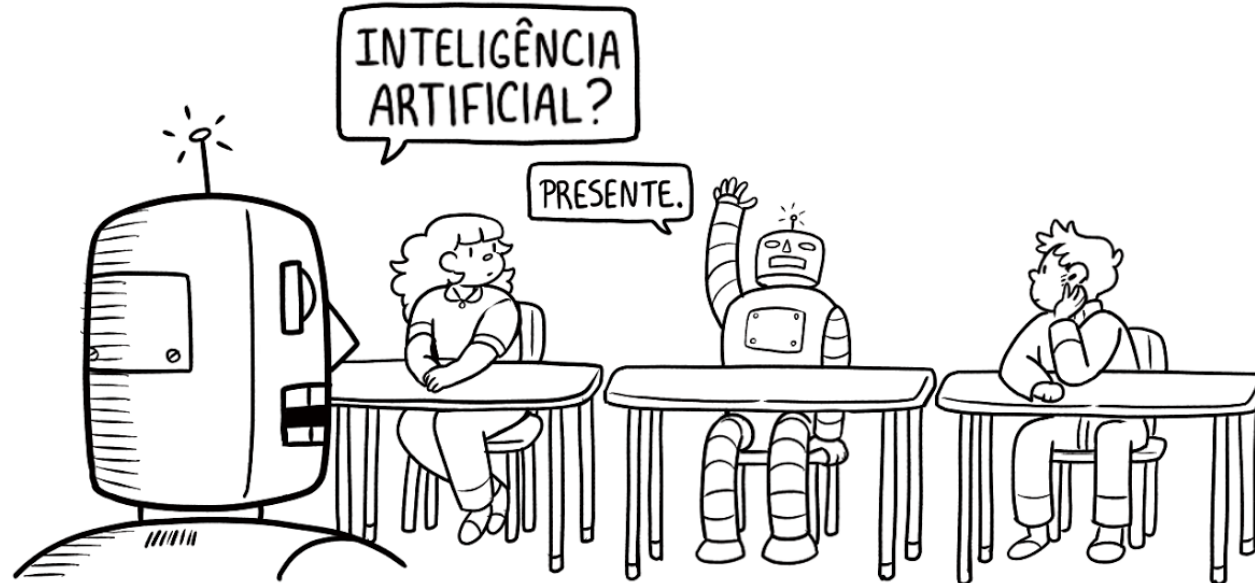
04 Desafios

Novas tecnologias

- Muitas e variadas
 - Internet, realidade aumentada, blockchain, bitcoin etc
- Na atualidade:
 - Todas as baseadas em Inteligência Artificial (IA)
 - Sub-área de IA....
 - Redes Neurais (Neural Networks)
 - Aprendizagem Profunda (Deep Learning)
 - Modelos de linguagem escaláveis (Large Language Models)

Inteligência Artificial

- O que é inteligência artificial (IA)?
- Como ela está a moldar o mundo atual e futuro?



Definição de IA

- A inteligência artificial refere-se a sistemas que ***simulam*** a capacidade humana de pensar, aprender e tomar decisões.



Imagem gerada pelo GPT (ética?)

Categorias

Raciocínio	Ação
“Pensam” como um humano	Agem como um humano
“Pensam” racionalmente	Agem racionalmente

Terminologia?

- Machine Learning?
- Artificial Neural Networks?
- Deep Learning?
- Prediction?
- Large language models?



Inteligência Artificial e Otimização

História da IA

- Marcos importantes:
 - 1956: Termo 'IA' é cunhado
 - 1997: Deep Blue vence o campeão mundial de xadrez
 - 2012: Avanços em redes neuronais profundas
 - 2017: *Transformers*
 - 2018: Modelos de linguagem



Inteligência Artificial (IA)

- Significado atual:
 - Aprendizagem automática
 - *Deep learning* e variações
- Porém, é muito mais:
 - Procura otimizada e eficiente
 - Procura adversarial otimizada e eficiente
 - Otimização, em geral
 - **Raciocínio**
 - **Planeamento**
 - **Ética...**

The Nobel Prize in Physics 2024

John Hopfield

“for foundational discoveries and inventions that enable machine learning with artificial neural networks”



Geoffrey Hinton

“for foundational discoveries and inventions that enable machine learning with artificial neural networks”



<https://www.nobelprize.org/all-nobel-prizes-2024/>

The Nobel Prize in Chemistry 2024

David Baker

“for computational protein design”



Demis Hassabis

“for protein structure prediction”



John Jumper

“for protein structure prediction”



<https://www.nobelprize.org/all-nobel-prizes-2024/>

Large-scale Data is Everywhere!

M
O
T
I
V
A
T
I
O
N

T
O

M
L

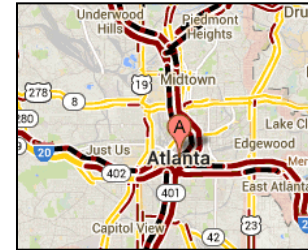
- There has been enormous data growth in both commercial and scientific databases due to advances in data generation and collection technologies
- New mantra
 - Gather whatever data you can whenever and wherever possible.
- Expectations
 - Gathered data will have value either for the purpose collected or for a purpose not envisioned.



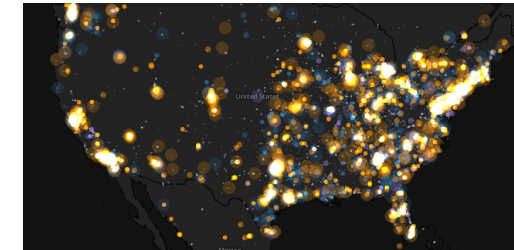
Cyber Security



E-Commerce



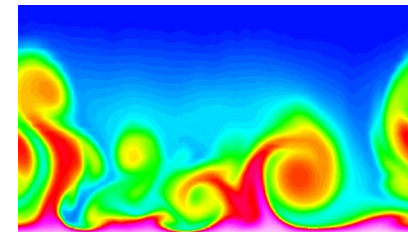
Traffic Patterns



Social Networking: Twitter



Sensor Networks

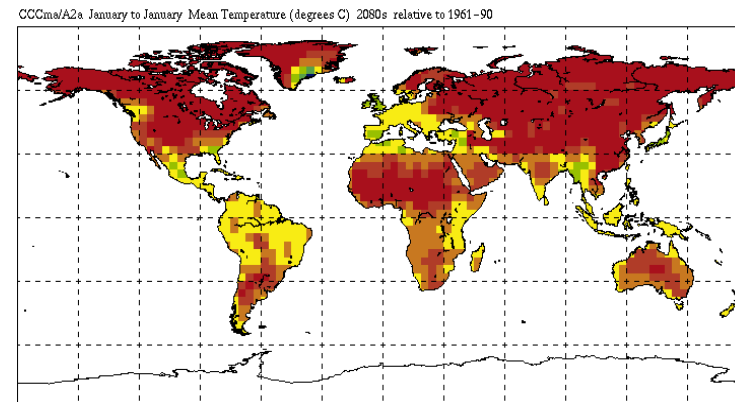


Computational Simulations

Great Opportunities to Solve Society's Major Problems



Improving health care and reducing costs



Predicting the impact of climate change

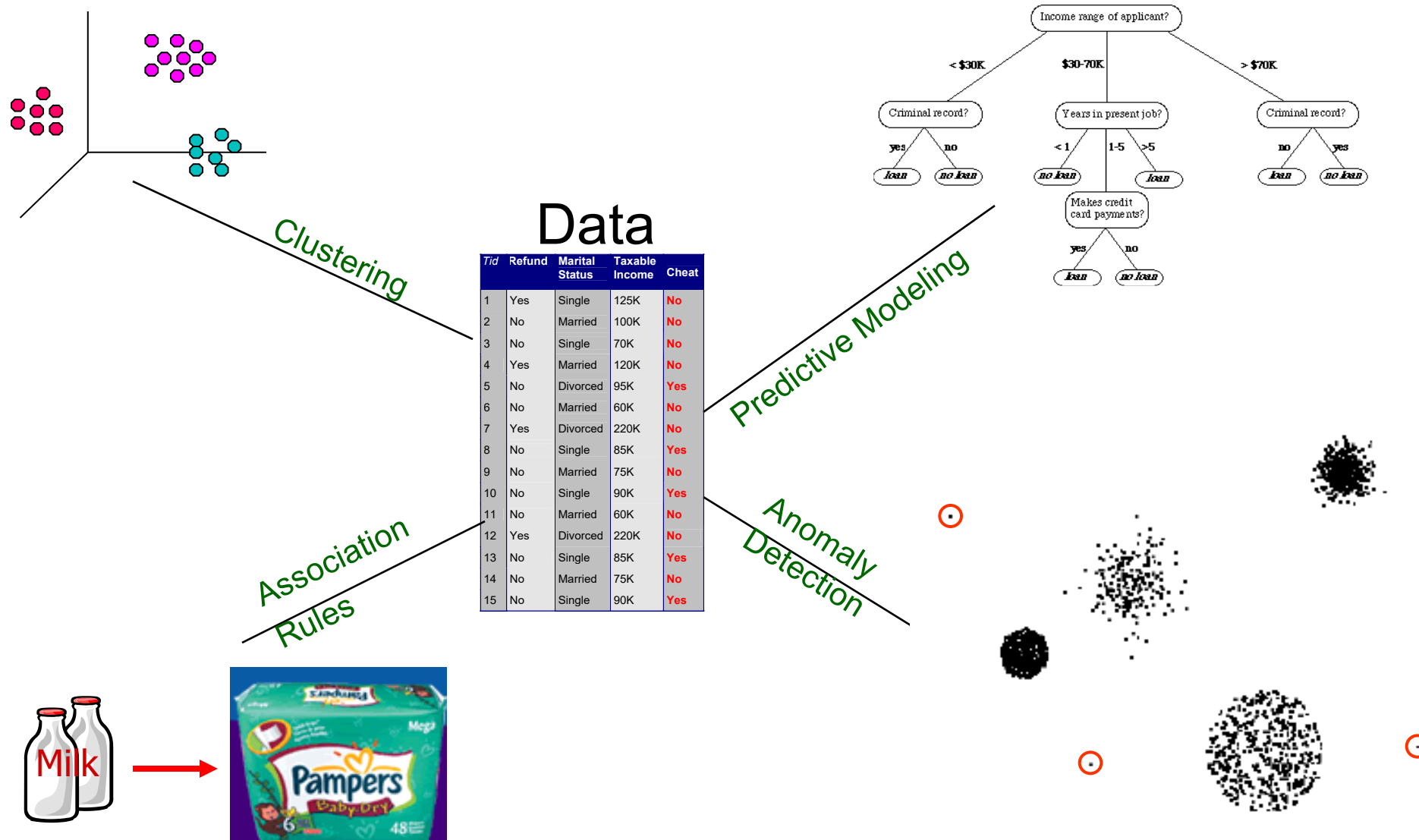


Finding alternative/ green energy sources



Reducing hunger and poverty by increasing agriculture production

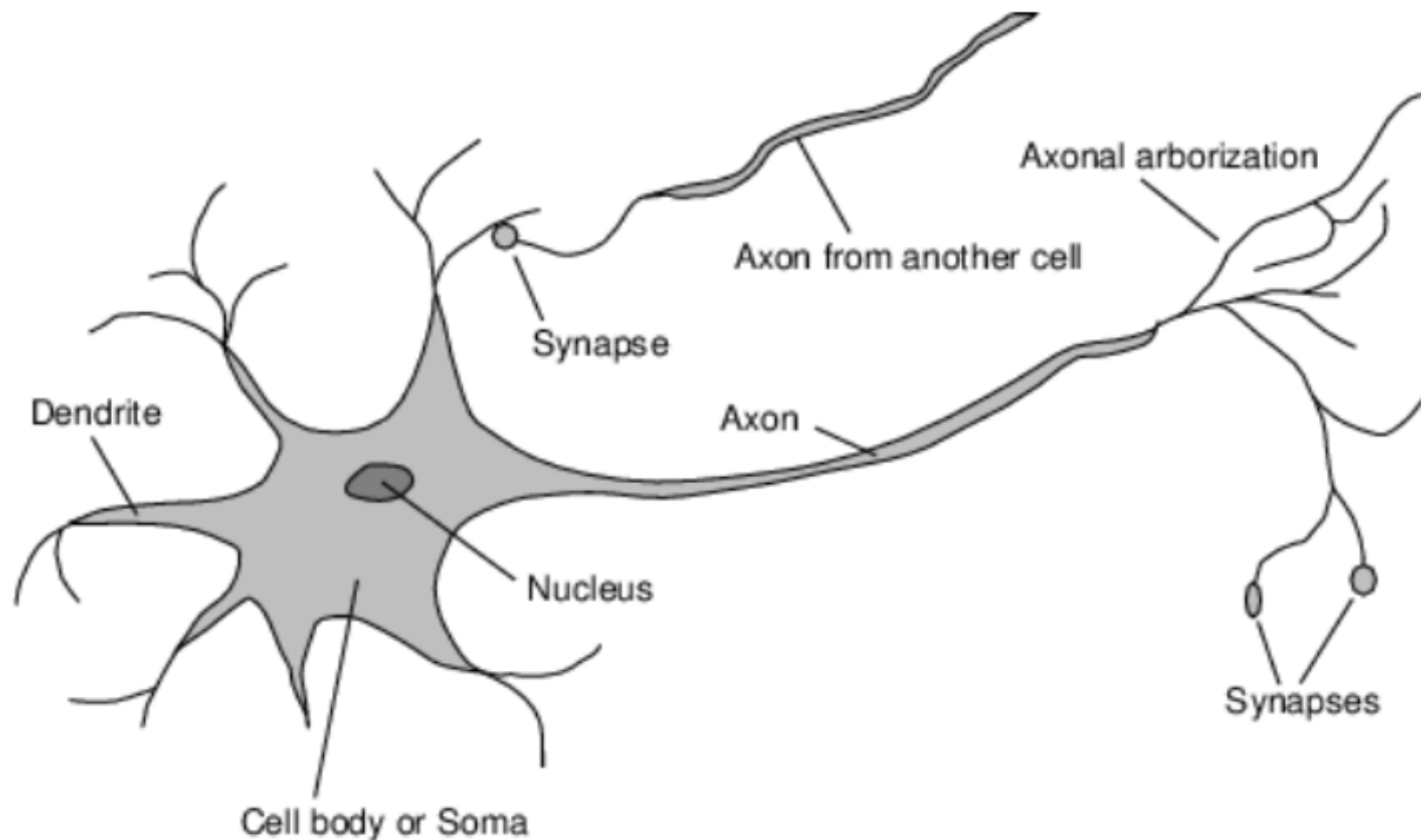
Tarefas de Aprendizagem Automática



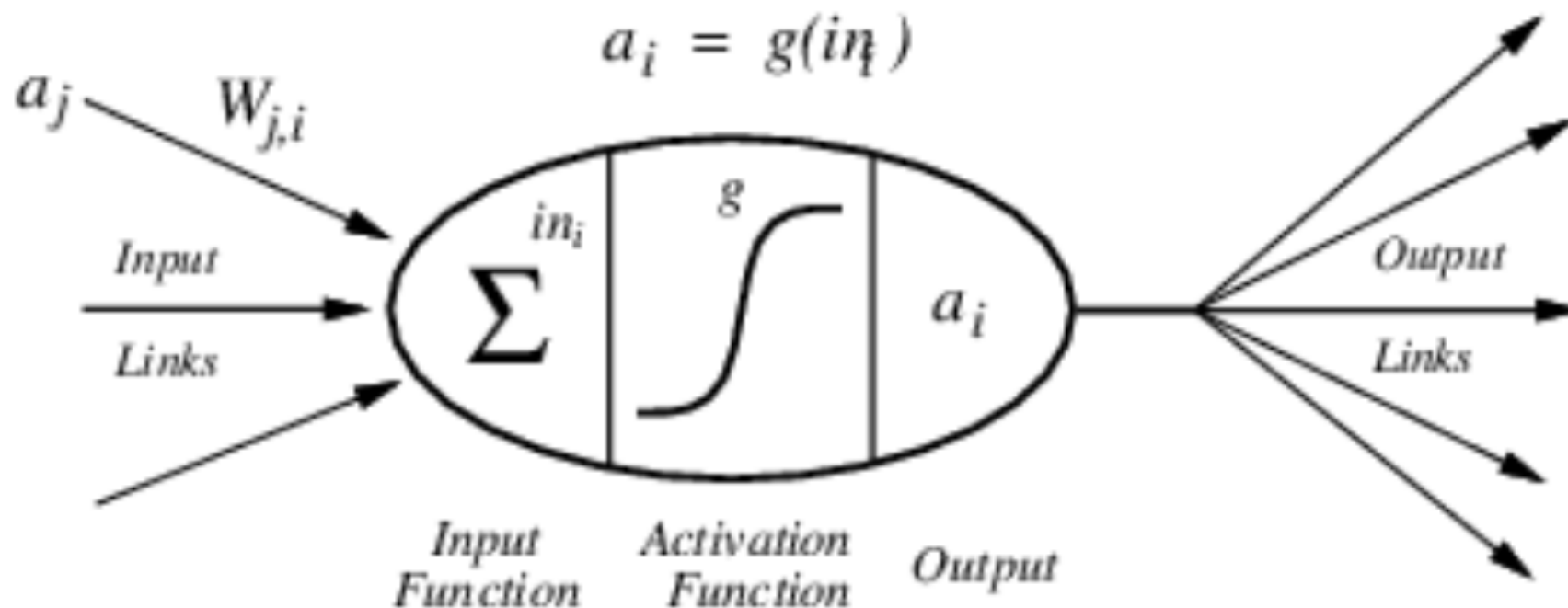
Exemplo de Sistema de IA: Rede Neuronal

- Do ponto de vista computacional: método para representar funções usando redes de elementos aritméticos simples.
- Método pode aprender novas funções através de exemplos.
- Do ponto de vista biológico: Modelo matemático para a operação do cérebro.
- Neurônios: Elementos aritméticos simples.
- Redes Neurais: conjunto de neurônios interligados.

Esquema de um neurônio biológico



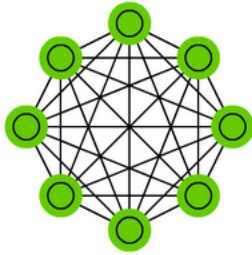
Esquema de um neurônio computacional



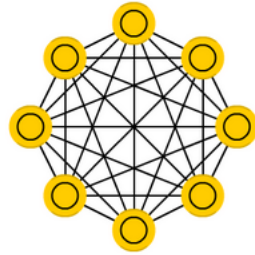
Exemplos de redes neuronais

<https://www.asimovinstitute.org/neural-network-zoo/>

Markov Chain (MC)



Hopfield Network (HN)



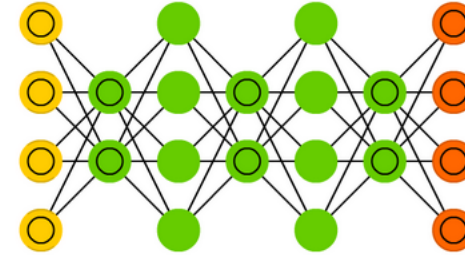
Boltzmann Machine (BM)



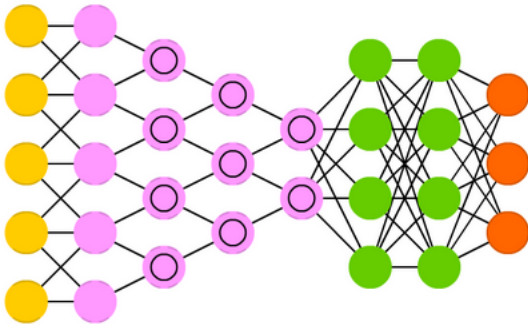
Restricted BM (RBM)



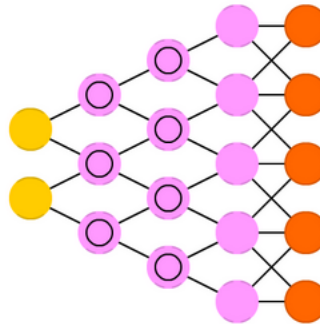
Deep Belief Network (DBN)



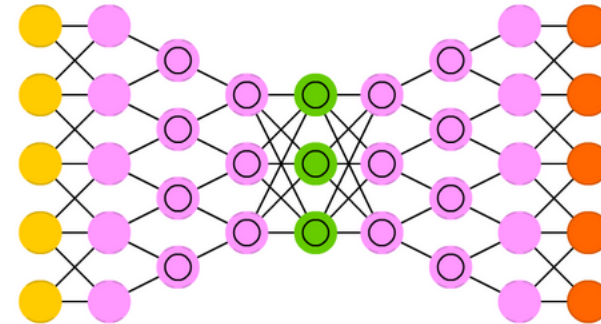
Deep Convolutional Network (DCN)



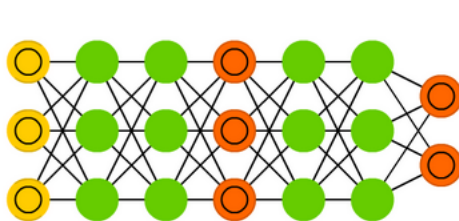
Deconvolutional Network (DN)



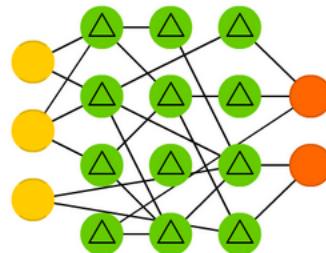
Deep Convolutional Inverse Graphics Network (DCIGN)



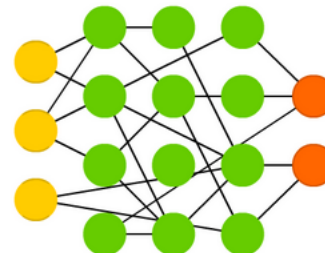
Generative Adversarial Network (GAN)



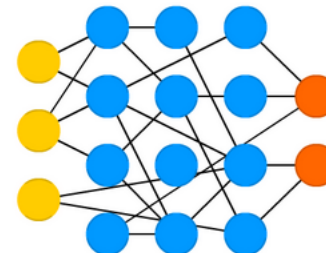
Liquid State Machine (LSM)



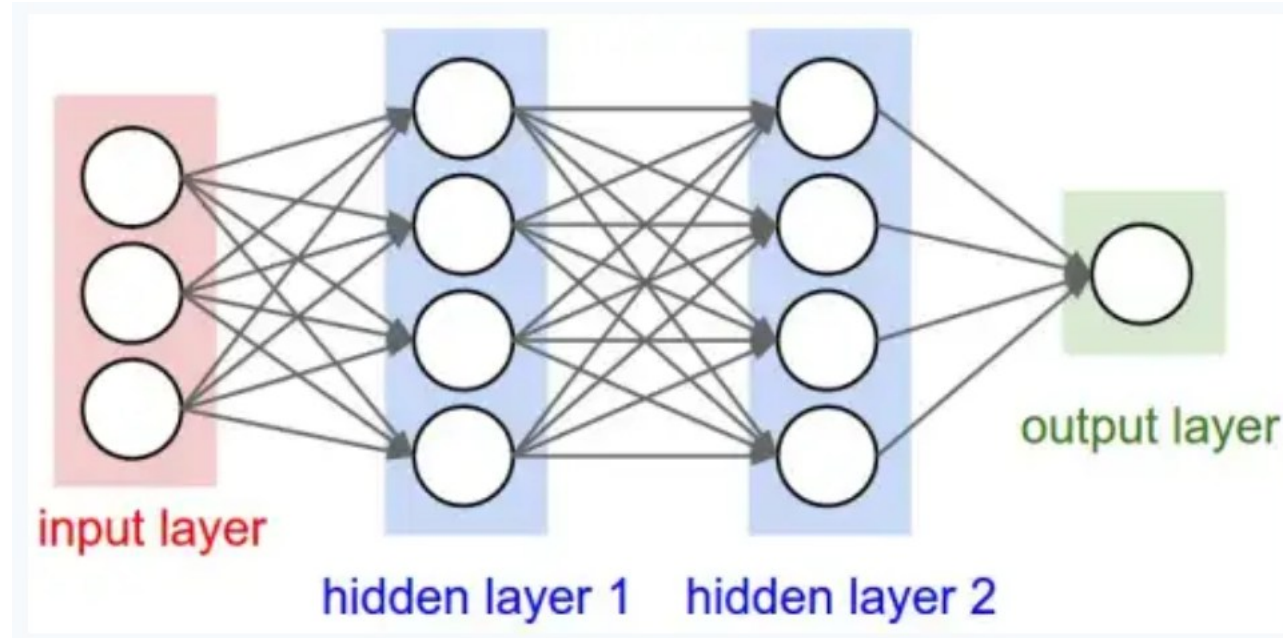
Extreme Learning Machine (ELM)



Echo State Network (ESN)



Arquitetura mais comum: feed forward



- Número de camadas pode variar
- Número de neurónios em cada camada pode variar

Como uma rede neuronal aprende?

- Utilização do algoritmo de *backpropagation*
- Integrado no processo de aprendizagem:

x_i : feature vector

y_i : variável target

(o que queremos aprender)

$\hat{y}$: saída da rede

(valor previsto pela rede)

k : iteração do algoritmo

λ : learning rate

- Initialize weights $w_0, w_1, w_2, \dots, w_d$

- Repeat

- for each training example (x_i, y_i)

- compute $\hat{y} = f(w, x_i)$

- update the weights:

$$w^{(k+1)} = w^{(k)} + \lambda [y_i - f(w^{(k)}, x_i)] x_i$$

- Until stopping criteria condition is met

Cérebro humano x redes neuronais

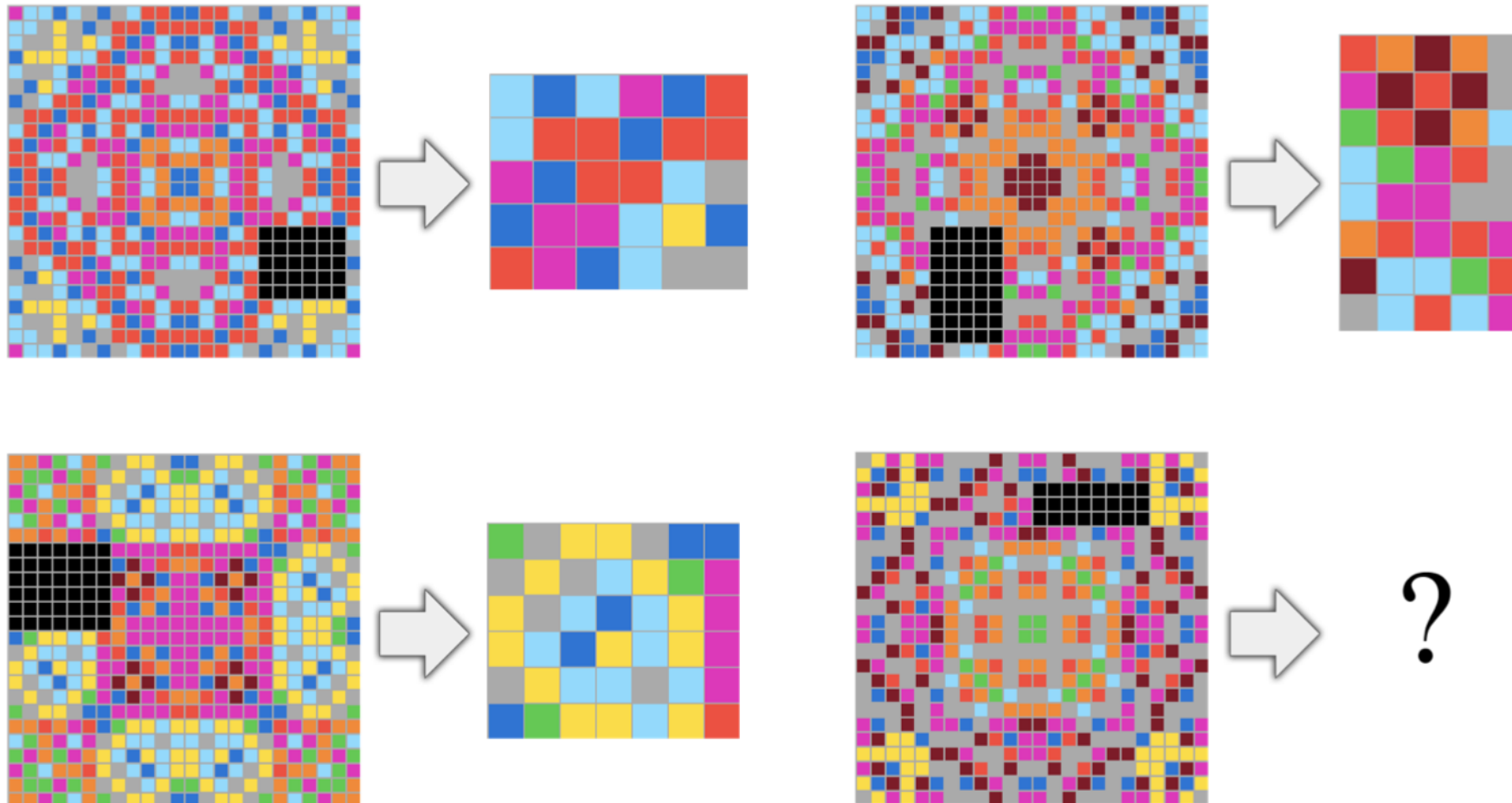
- Cérebro:
 - 86 bilhões de neurônios
 - Aproximadamente 1000 ligações por neurônio (talvez mais)
 - Corresponderia a ~16 Tera parâmetros
- Duas ordens de grandeza maior do que sistemas tais como GPT3

Cérebro humano x Redes neuronais

- Quantas imagens são necessárias para que uma criança aprenda a distinguir cães de gatos?
- E quantas imagens são necessárias para que uma rede neuronal “aprenda” a distinguir cães de gatos?



ARC Challenge



Top solution
produces the
incorrect solutions
almost 80% of the
time

<https://pgpbpadilla.github.io/chollet-arc-challenge>

Desafios

Desafios técnicos da Aprendizagem Automática

- AA procura padrões em dados ou gera novos dados de acordo com alguma distribuição
- Porém:
 - Estes padrões são gerais?
 - Aplicáveis a todos os grupos, idades, etnias, regiões etc?
 - Os dados mostram diversidade? Viés? Qualidade?
 - Dados faltantes? Aleatoriamente? Não aleatoriamente? N/A?
- Variáveis com diferentes tipos?
- Tipos de algoritmos?

Some “pearls”

TABLE 4.3

Causes of Insufficient Data Quality during Setup and Organization of the Registry [38]

Problems at the Central Coordinating Center	Type of Error
Unclear/ambiguous data definitions	Systematic
Unclear data collection guidelines	Systematic
Poor case record form layout	Systematic/random
Poor interface design	Systematic/random
Data overload	Random
Programming errors	Systematic
Problems at the Local Sites	Type of Error
Illegible handwriting in data source	Random
Incompleteness of data source	Systematic
Unsuitable data format in source	Systematic
Data dictionary not available to data collectors	Systematic/random
Lack of motivation	Random
Frequent shift in personnel	Random
Programming errors	Systematic

Tipos diferentes de variáveis

- Alguns algoritmos assumem que todos os dados são numéricos
- Python scikit-learn implementa vários algoritmos de aprendizagem, incluindo árvores de decisão
 - DecisionTreeClassifier não aceita variáveis com valores categóricos

Tipos diferentes de variáveis

- Na área da saúde, dados tendem a ser categóricos
 - Agrupados por idade, peso, género etc.
- Agrupamentos produzidos por algoritmos podem não condizer com os grupos semânticos criados pelos profissionais da saúde
 - Bem vindos ao mundo espetacular dos **embeddings** e **encodings**!

Imputação de dados

- Algumas implementações assumem imputação de dados faltantes
 - Significado?
 - E.g., é correto “inventar” o peso e/ou altura de alguma pessoa?
 - E género?

Desafios com validação e métricas

- Algumas implementações assumem que métricas de desempenho de modelos são calculadas a partir da média das métricas de múltiplas experiências (cross-validation, por exemplo)
 - WEKA usa a media no módulo **experimenter**, mas não no **explorer**

3 folds:

Fold1: precision= $P1=TP1/(TP1+FP1)$

Fold2: precision= $P2=TP2/(TP2+FP2)$

Fold3: precision= $P3=TP3/(TP3+FP3)$

$$P = \frac{P1 + P2 + P3}{3}$$

OU

$$P = \frac{TP1 + TP2 + TP3}{TP1 + TP2 + TP3 + FP1 + FP2 + FP3}$$

Desafios com validação e métricas

Varying Hardware Test (Software Bug)

Processing Unit:Software Environment	0	1	2	3	4
NVIDIA A40:tensorflow_22.03	0.86584	0.86584	0.86584	0.86584	0.86584
NVIDIA A100-40GB:tensorflow_22.03	0.86580	0.86580	0.86580	0.86580	0.86580
Tesla V100-16GB:tensorflow_22.03	0.85203	0.85203	0.85203	0.85203	0.85203
NVIDIA A40:tensorflow_22.06	0.86584	0.86584	0.86584	0.86584	0.86584
NVIDIA A100-40GB:tensorflow_22.06	0.86580	0.86580	0.86580	0.86580	0.86580
Tesla V100-16GB:tensorflow_22.06	0.85203	0.85203	0.85203	0.85203	0.85203
NVIDIA A40:tensorflow_2.8.0-gpu	0.86571	0.86555	0.85680	0.85927	0.86147
NVIDIA A100-40GB:tensorflow_2.8.0-gpu	0.86511	0.84516	0.86360	0.85895	0.86479
Tesla V100-32GB:tensorflow_2.8.0-gpu	0.85983	0.78835	0.85764	0.86031	0.86908
NVIDIA A40:tensorflow_2.9.1-gpu	0.83631	0.86287	0.83891	0.85724	0.85276
NVIDIA A100-40GB:tensorflow_2.9.1-gpu	0.85979	0.86619	0.83156	0.86412	0.84564
Tesla V100-16GB:tensorflow_2.9.1-gpu	0.84447	0.85320	0.85996	0.86511	0.86779
NVIDIA A40:tensorflow_2.9.1-gpu-cuda11.3-cudnn8.2	0.86184	0.86184	0.86184	0.86184	0.86184
NVIDIA A100-40GB:tensorflow_2.9.1-gpu-cuda11.3-cudnn8.2	0.84631	0.84631	0.84631	0.84631	0.84631

Slide source: Kevin Coakley (SDSC, NTNU).

Slide kindly provided by Christine Kirkpatrick, SDSC

Preparação de dados para cross-validation

- Exemplo: casos and controles
- Usualmente, do ponto de vista clínico, como preparar os dados?
 - Caso-controle: matching algorithm
 - Exact matching?
 - Approximate matching?

Por exemplo, como separar exames de pacientes?

Data organization, structuring and sampling

	1	2	3	4	5	6
1	0	1	0	1	1	0
2	1	0	1	0	0	1
3	0	1	0	1	1	0
4	1	0	1	0	0	1
5	0	1	0	1	1	0
6	1	0	1	0	0	1
7	0	1	0	1	1	0
8	1	0	1	0	0	1
9	0	1	0	1	1	0

Arrangement



	6	1	3	2	5	4
4	1	1	1	0	0	0
2	1	1	1	0	0	0
6	1	1	1	0	0	0
8	1	1	1	0	0	0
5	0	0	0	1	1	1
3	0	0	0	1	1	1
9	0	0	0	1	1	1
1	0	0	0	1	1	1
7	0	0	0	1	1	1

Sampling



(a) 8000 points



(b) 2000 points



(c) 500 points

Class label and classification results

- Some software use the order of appearance of labels to calculate metrics for classification
 - WEKA and R use the label seen in the first row of the data table to calculate metrics

V1	V2	...	Class
...	...	...	Yes
...	...	...	No
...	...	...	No

V1	V2	...	Class
...	...	...	No
...	...	...	No
...	...	...	Yes

Dados em formato tabular

- A maior parte dos algoritmos de aprendizagem automática requer que os dados estejam em formato tabular
 - Redundância
 - Introdução de valores nulos
 - Perda de informação

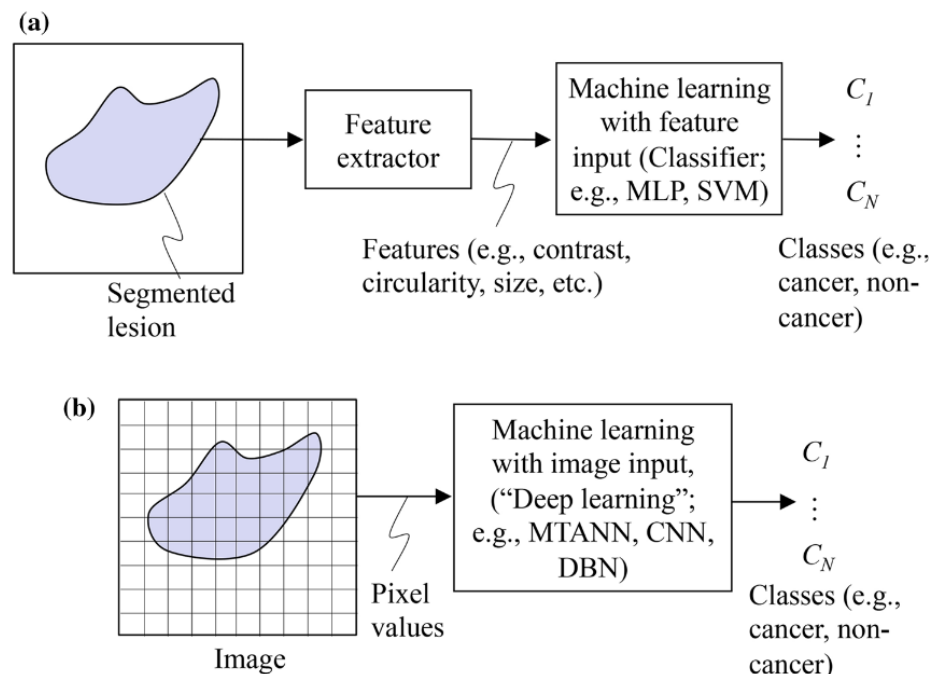
Learning from data is tricky

Statistics	Vs.	Machine learning
Supervised	Vs.	Unsupervised
Induction	Vs.	Deduction
Correlation	Vs.	Causation
Sampling	&	Confidence intervals
Probability	&	Distribution
Deviation	&	Variance
Causation	&	Prediction

Ack: Carlos Somohano, Founder Data Science London, 2013

Tendência atual

- A maior parte dos modelos atuais baseia-se em aprendizagem profunda (deep learning)
- Com foco em imagens e/ou texto
- E.g., diagnóstico



Era pré-*deep learning*

- Estatística
 - Odds ratio
- Regressão
- Sistemas baseados em regras
- Sistemas probabilísticos (e.g., Bayesian nets)
- Árvores de decisão

- Algum grau de interpretabilidade e explicação

Argumentos pro *deep learning*

- Resultados excelentes
 - Boa capacidade de discriminação
 - Boa capacidade de geração

Argumentos contra *deep learning*

- Tempo de treino do modelo
- Falta de interpretabilidade
- Falta de explicação
- Necessidade de um número muito grande de exemplos de treino
 - Alguns domínios (e.g., saúde) apresentam casos raros
 - Alguns domínios não possuem quantidades muito grandes de dados para treino (e.g., saúde novamente 😊)

Performance

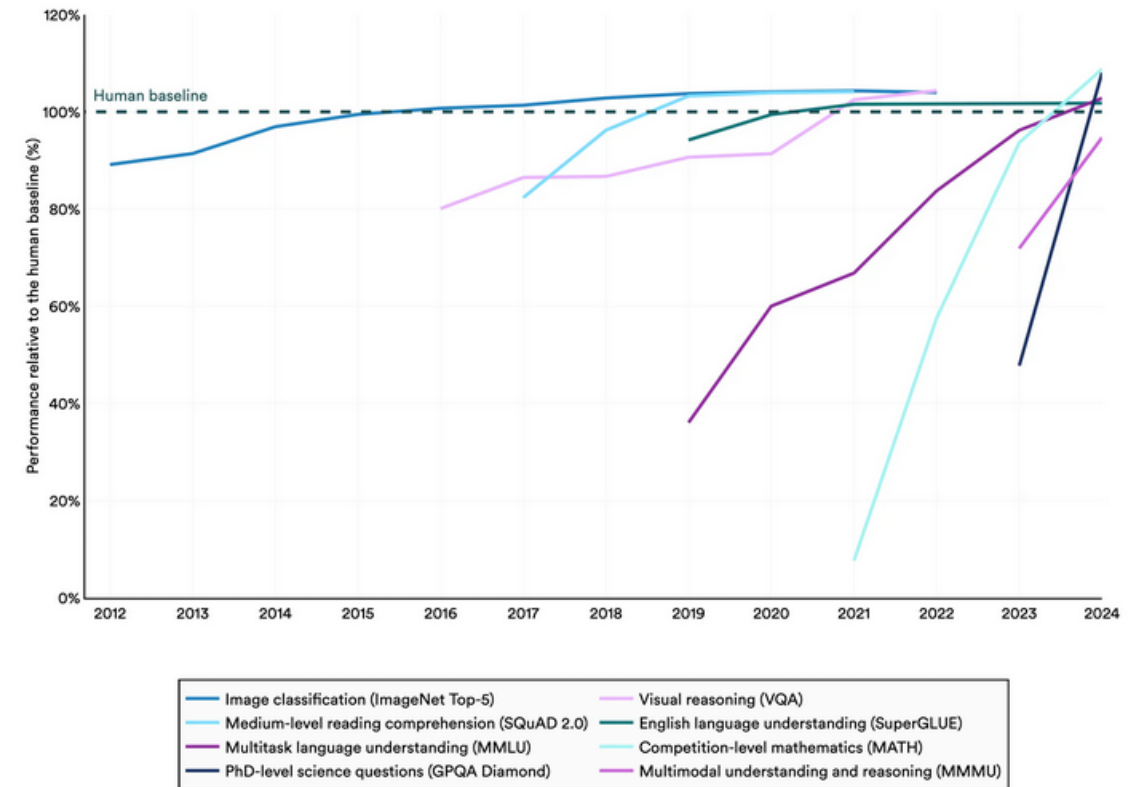
<https://hai.stanford.edu/ai-index/2025-ai-index-report>

1. AI performance on demanding benchmarks continues to improve.

In 2023, researchers introduced new benchmarks—MMMU, GPQA, and SWE-bench—to test the limits of advanced AI systems. Just a year later, performance sharply increased: scores rose by 18.8, 48.9, and 67.3 percentage points on MMMU, GPQA, and SWE-bench, respectively. Beyond benchmarks, AI systems made major strides in generating high-quality video, and in some settings, language model agents even outperformed humans in programming tasks with limited time budgets.

Select AI Index technical performance benchmarks vs. human performance

Source: AI Index, 2025 | Chart: 2025 AI Index report



Momento atual

- Sistemas de IA são “assistentes”
 - Ajudam em tarefas repetitivas
 - Não se cansam
 - Não se estressam
 - Não reclamam 😊
 - Libertam os humanos para tarefas mais intelectuais e de lazer
 - Podem aumentar a produtividade



Site do dia · 28 nov 2024 08:49

Inteligência Artificial ajuda nas compras da Black Friday com identificação de vantagens e desvantagens dos produtos

O site Não sejas Pato quer ajudar os consumidores a fazerem as melhores escolhas na Black Friday e está a usar a Inteligência artificial para dar respostas às dúvidas.

Oportunidades

IA na Educação

- Exemplos:
 - Plataformas como ChatGPT para a aprendizagem
 - Ferramentas adaptativas para o ensino personalizado



IA no Mercado de Trabalho

- Automação de tarefas repetitivas
- Criação de novas profissões, como cientistas de dados e engenheiros de IA



IA na Saúde

- Avanços:
 - Diagnósticos médicos mais precisos
 - Análise de grandes volumes de dados para tratamentos personalizados



IA no Entretenimento

- Personalização de recomendações em plataformas de streaming
- Criação de jogos interativos com IA



IA e Redes Sociais

- Riscos e Benefícios:
 - Conexão de pessoas com interesses similares
 - Algoritmos que podem reforçar bolhas de informação



Capacitação

- Indivíduos podem usar IA para:
 - Criar *startups* inovadoras
 - Desenvolver habilidades digitais avançadas
 - Aumentar a produtividade
 - Concentrar-se em atividades mais intelectuais



Educação Personalizada

- A IA adapta conteúdos para atender às necessidades individuais dos estudantes, melhorando a aprendizagem.



Educação em IA

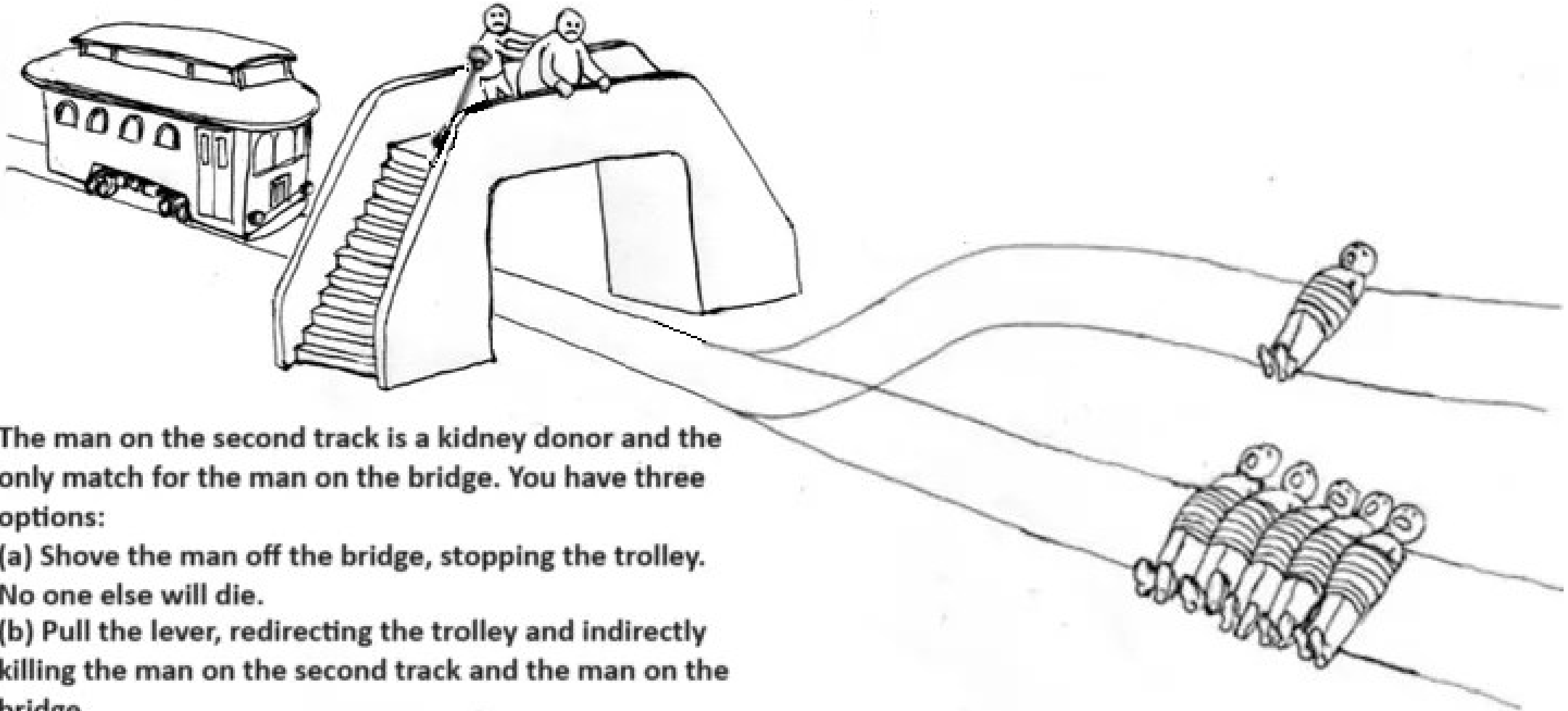
- É importante ensinar sobre inteligência artificial nas escolas para preparar as futuras gerações.
- A educação digital é fundamental para preparar os jovens.
- Eles precisam desenvolver habilidades como pensamento crítico e alfabetização digital para se destacarem e se defenderem no mundo online.

Questões éticas

Problema do comboio (Trolley problem)



Problema similar



The man on the second track is a kidney donor and the only match for the man on the bridge. You have three options:

- (a) Shove the man off the bridge, stopping the trolley. No one else will die.
- (b) Pull the lever, redirecting the trolley and indirectly killing the man on the second track and the man on the bridge.
- (c) Do nothing, letting five people die.

Momento atual

- Potencial de risco
 - Cultura de massa, vieses, pode facilitar a formação de “tribos”
 - Pode tornar o ser humano “preguiçoso”
 - Pode levar as pessoas a confiarem demais nos sistemas e perderem o senso crítico
 - Falta de ética



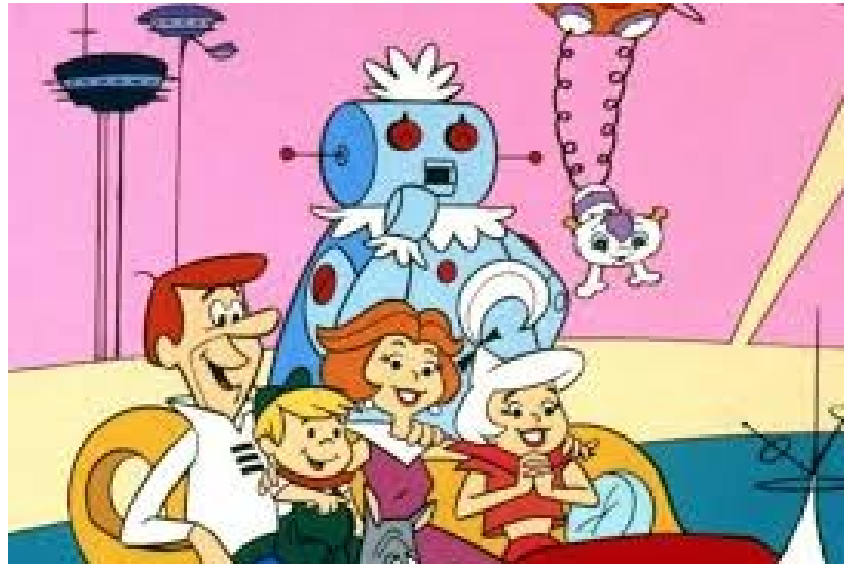
Multimédia · 27 nov 2024 17:58

OpenAI suspende acesso antecipado ao modelo Sora após protesto de artistas

Cerca de 20 artistas divulgaram publicamente o acesso ao modelo de IA Sora que estavam a testar, em forma de protesto, alegando terem sido usados como "marionetas". Em resposta, a OpenAI suspendeu o programa.

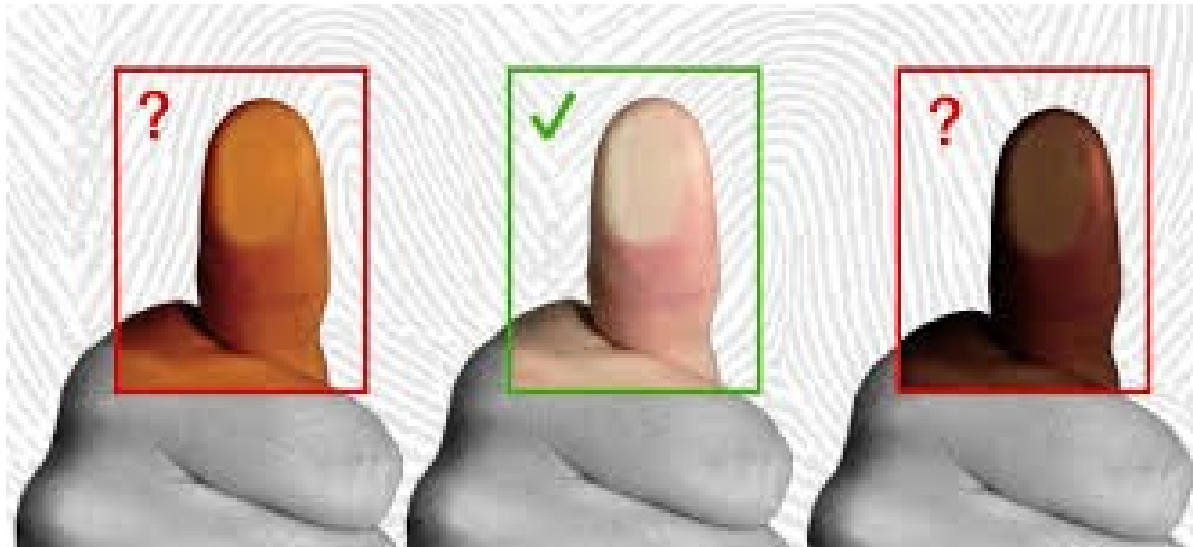
Automação e Desemprego

- Desafios:
 - Substituição de empregos tradicionais por sistemas inteligentes
 - Necessidade de requalificação profissional



Preconceito nos Algoritmos

- Problemas:
 - Vieses em sistemas de IA devido a dados históricos tendenciosos



Privacidade e Ética

- Riscos:
 - Uso indevido de dados pessoais
 - Necessidade de regulação ética



Sky news

Google's AI chatbot Gemini tells user to 'please die' and 'you are a waste of time and resources'

Gemini is supposed to have restrictions that stop it from encouraging or enabling dangerous activities, including suicide, but somehow, it still managed to tell one "thoroughly freaked out" user to "please die".

By Mickey Carroll, science and technology reporter

Cibersegurança

- Desafios:
 - IA usada para ataques cibernéticos
 - Necessidade de defesas inteligentes
- E se a própria IA for alvo de ataque distorcendo seu propósito e resultados?

Resumo dos Desafios

- Principais riscos:
 - Preconceito em algoritmos
 - Perda do senso crítico
 - Privacidade e ética
 - Desemprego em algumas áreas
 - Lacuna digital devido à diferenças socio-econômicas

Estatísticas

- De acordo com o ManpowerGroup Employment Outlook Survey do 3º trimestre de 2024:
 - mais de metade das grandes empresas globais com mais de 5.000 trabalhadores (52%) usam atualmente IA.
 - Os empregadores europeus são mais cautelosos (43%), quando comparados com a média global.
 - Um terço (33%) dos inquiridos, que não são atualmente utilizadores, dizem que a sua organização planeia implementar ferramentas de IA nos próximos três anos

Sistemas de IA...

- não “alucinam”, cometem erros
- não “pensam”, imitam/simulam conversas
- não têm opinião ou características próprias, apenas refletem as opiniões e características de quem os criou
- não têm intenções, apenas produzem resultados
- não desejam manipular e causar danos, mas podem facilitar manipulação, disseminar má informação e causar danos

(tradução livre de um slide apresentado pelo Prof. Baeza-Yates num seminário sobre LLMs na FEUP, em Set de 2024)

Mensagem de John Hopfield

- Ganhador do prémio Nobel de 2024:

“...it's quite conceivable that humanity is just a passing phase in the evolution of intelligence...But then it (AI) can start getting direct experience of the world and learn much faster...”

Em outras palavras, se o ser humano acomodar-se com o avanço das tecnologias de IA, poderá haver um desequilíbrio intelectual entre humanos e máquinas.

Regulamento Inteligência Artificial da UE

- As novas regras proíbem determinadas aplicações de IA que ameaçam os direitos dos cidadãos, incluindo:
 - sistemas de categorização biométrica baseados em características sensíveis
 - recolha não direcionada de imagens faciais da Internet ou de televisão em circuito fechado para criar bases de dados de reconhecimento facial
 - reconhecimento de emoções no local de trabalho e nas escolas
 - classificação social
 - policiamento preditivo (quando se baseia exclusivamente na definição de perfis de uma pessoa ou na avaliação das suas características)
 - IA que manipula o comportamento humano ou explora as vulnerabilidades das pessoas

“Holy grail”: de uma perspectiva de ciência de computadores



Imagem automaticamente
gerada por:
<https://dynabench.org/tasks/adversarial-nibbler/create>

(“a patient being treated by an
artificial intelligence agent”)

Imagem com viés? ;)

Um olhar à frente

- Seria este cenário ideal para qualquer pessoa?
- Estamos muito distantes desta realidade?



Ideal?

- Depende:
 - Do propósito
 - Dos indivíduos
 - Dos métodos

Estamos muito distantes do cenário “ideal”?

- “Far far away” ;)
- Limitações:
 - Adaptabilidade
 - Explicabilidade
 - Flexibilidade
 - Generalização
 - Interdisciplinaridade
 - Sistemas legados



- Faltam ainda **Raciocínio e Planejamento** para uma **tomada de decisão (Decision Making)** adequada



Obrigada!

dutra@fc.up.pt

<http://www.dcc.fc.up.pt/~ines>

