Introductory Seminar on Artificial Intelligence and Machine Learning

Emanuele Ledda, Cagliari Digital Lab 2024 - Day 1

About Us

Pattern Recognition
and Applications Lab

Emanuele Ledda

May 2024

Introductory Seminar on Artificial Intelligence and Machine Learning

Daniele Angioni

July 2024

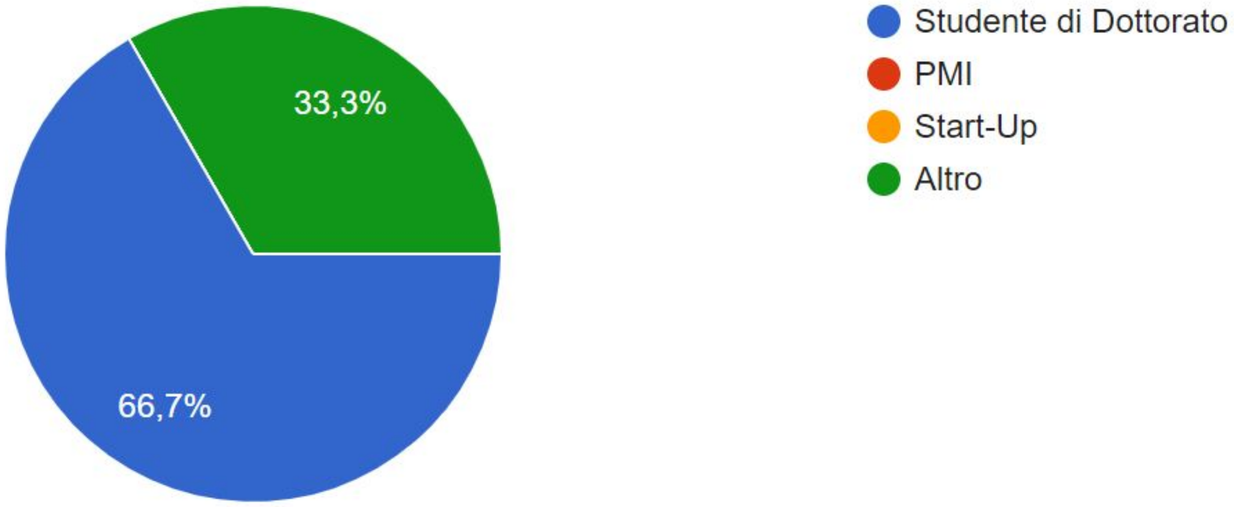
Introductory Seminar on PyTorch for Deep Learning

Sara Concas

September 2024

Introductory Seminar on Computer Vision

About You



Prerequisite

- Knowing at least one **programming language** (preferably **python**)
- Understanding the basic concepts of **Linear Algebra**
- Bring a **PC** 

Seminar Objectives

- Become aware of the evolution of **Artificial Intelligence** over the years and how the central paradigm shifted from logical systems to statistical models until reaching modern **Machine Learning.**

- Learn how the most paramount **Shallow Learning** algorithms work, focusing on **Classification Algorithms**.

- Learn the functioning of **artificial neural networks**, from the perceptron to **deep neural networks**.

- Understand the fundamental concepts that led to the most famous modern **foundation models** (e.g., GPT 4.0, Sora)

- Understand the main critical issues behind **AI Ethics** and **Trustworthy AI** today, both from a technical and a regulatory perspective (**AI Act**, **European Guidelines for Trustworthy AI**).

Seminar Objectives - Technical Python Aspects

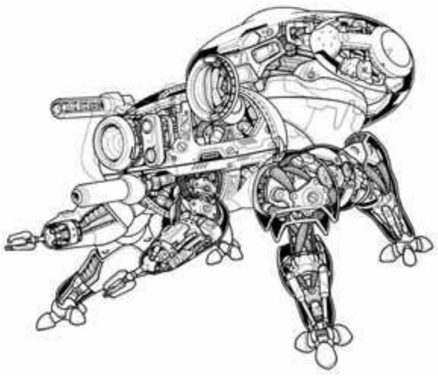
- Learning **numpy** essentials for managing tensor manipulation, essential for every ML task
- Basic data management skills with **pandas**
- General knowledge on using **SciKit-Learn** classifiers

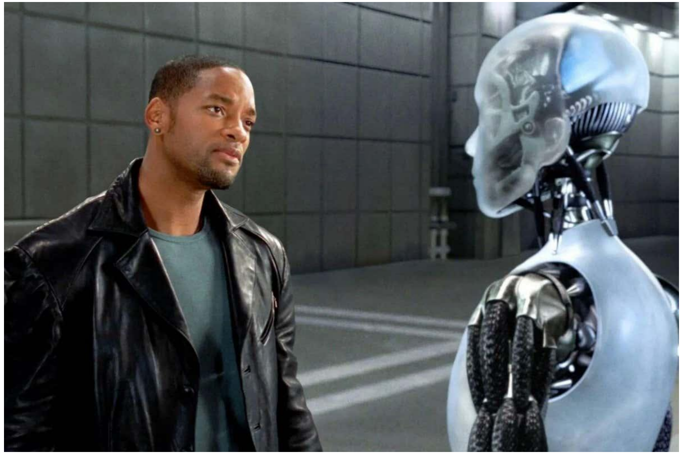
Introduction to Artificial Intelligence and Machine Learning

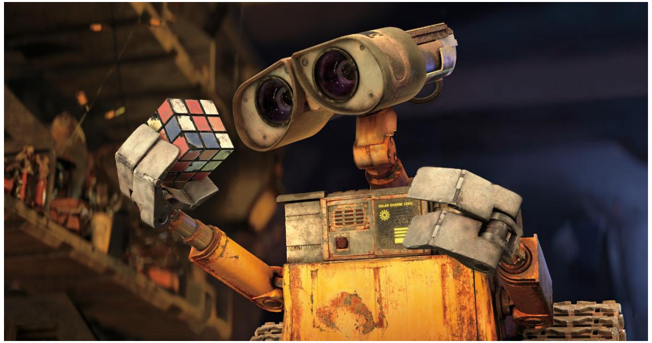
DLAB
CTE • CAGLIARI

Artificial Intelligence

Imagining Artificial Intelligence



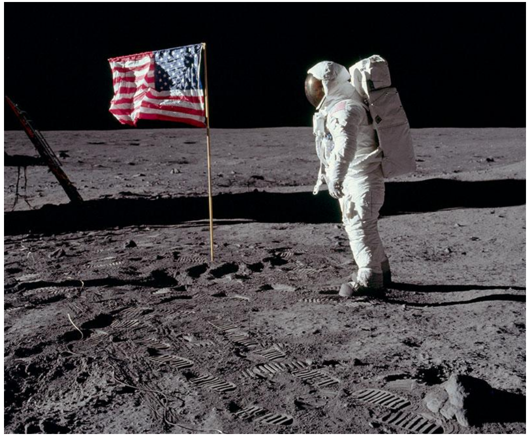


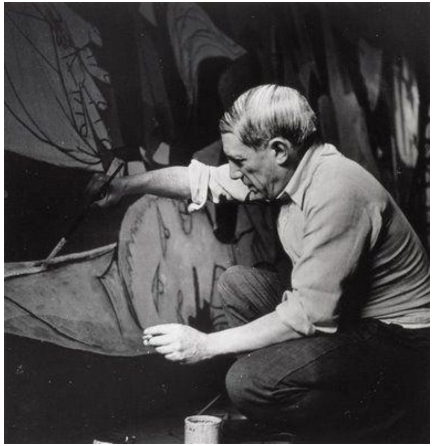


What is Intelligence?

- What should an intelligent agent be?
    - Should it be **similar to human beings**?
    - Should it be able to **play chess**?
    - Should it be able to express **artistic ideas**?

Human Intelligence can be this...





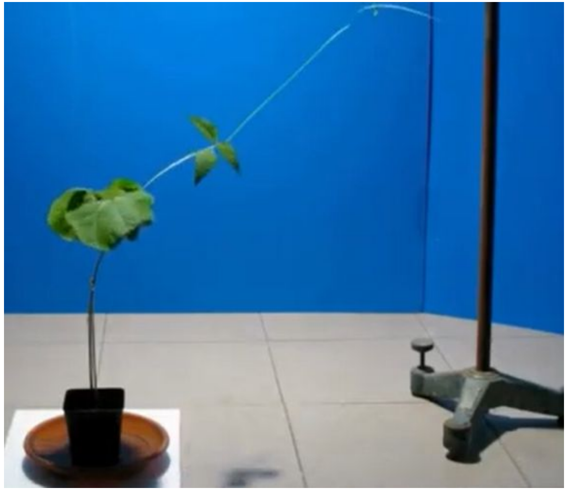
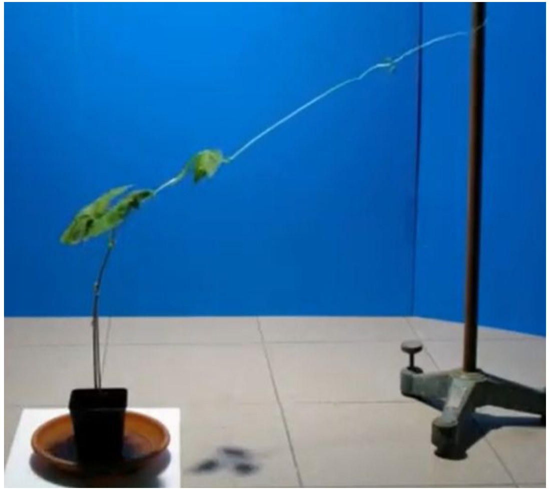
… But also this!

- But intelligence is **not a human privilege**…

Animal intelligence



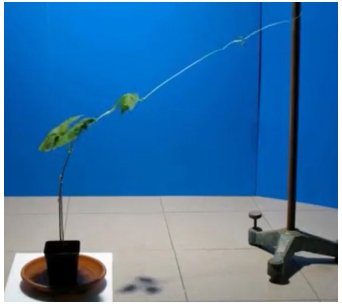
Plant Intelligence

So, what is Intelligence?

- Intelligence does **NOT** mean to be similar to humans

- Intelligence is the ability of a system to **act appropriately** in an **uncertain environment**, where "appropriately" means "**maximizing the success probability.**"



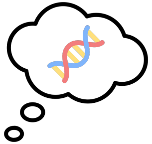
Climbing = Gain more light

Dancing = Impress a partner

Cooking (originally) = Optimize the nutrients

Agent, Action, Environment and Sensory Input

- **act appropriately** in an **uncertain environment**, for **maximizing the success probability.**



Objective:
saving his genes

Sensory Input: See a female of the same species

Agent: “Parotia Vieillot”

Azione: Dance

Environment: Forest

Learning

Sensory Input: Perceiving the light from above

Agent: Wild Bean Plant

Action: Go up!

Feedback: Struggling

Learning

Sensory Input: Perceiving the light from above

Agent: Wild Bean Plant

Action: Climb the pole!

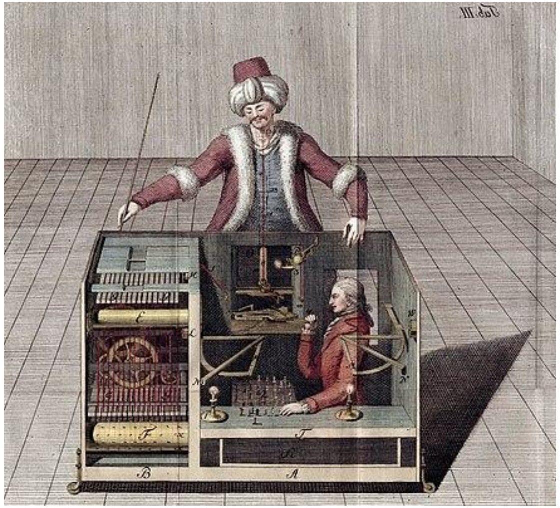
Feedback: Easier, gain more light!

What is…
“Artificial” Intelligence?

History of AI: Good Old Fashioned AI

Early Investigations on “Intelligent” Machines

- “Mimic” human behaviours



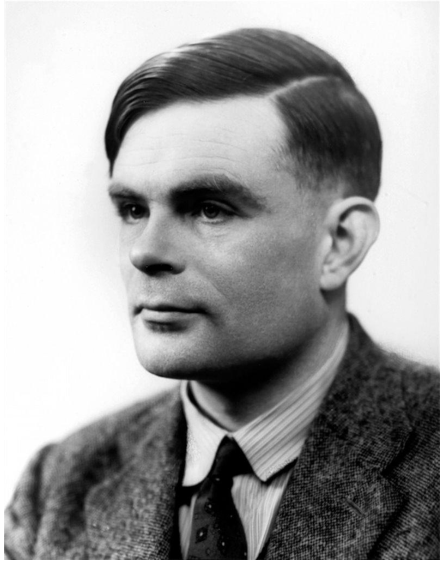
Artificial Intelligence and Computers

*"The Analytical Engine has no pretensions whatever to originate anything. **It can do whatever we know how to order it to perform**. It can follow analysis, but it has no power of anticipating any … relations or truths. Its province is to assist us in making available what we are already acquainted with"*

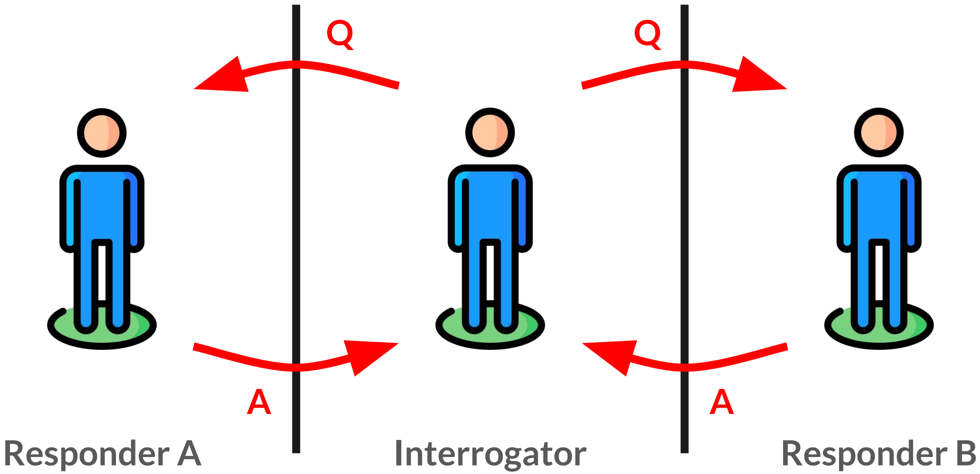
Ada Lovelace

Artificial Intelligence and Computers

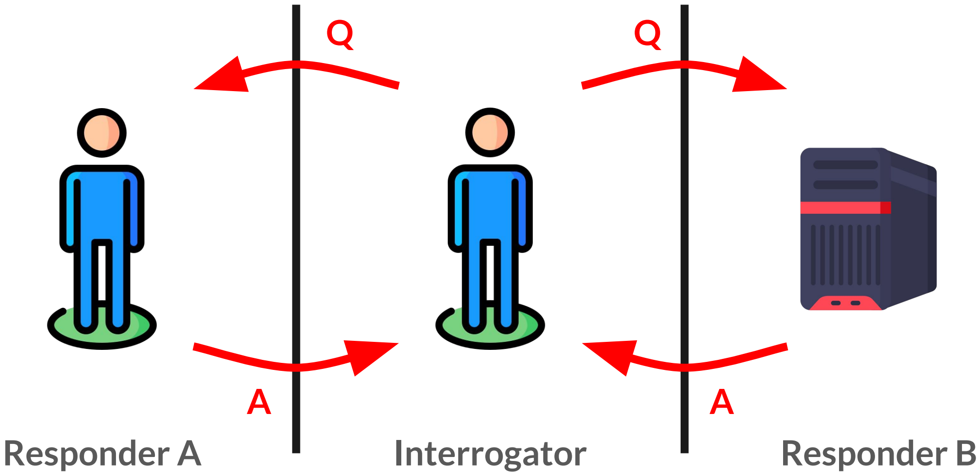
- Early Investigations on the nature of computation
- **Turing Machine**
- **Turing Test**



Turing Test



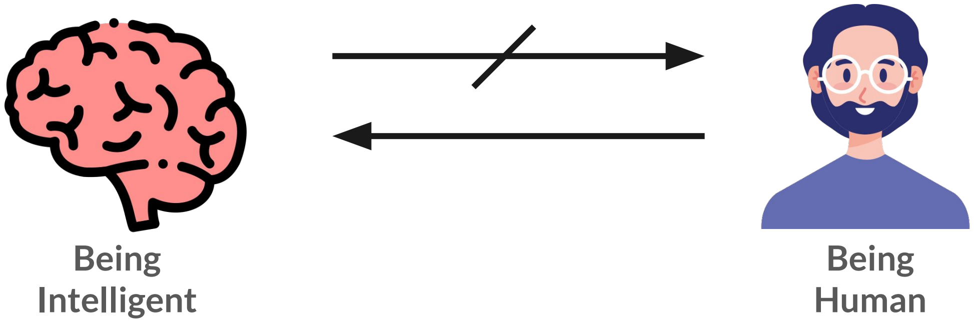
Turing Test



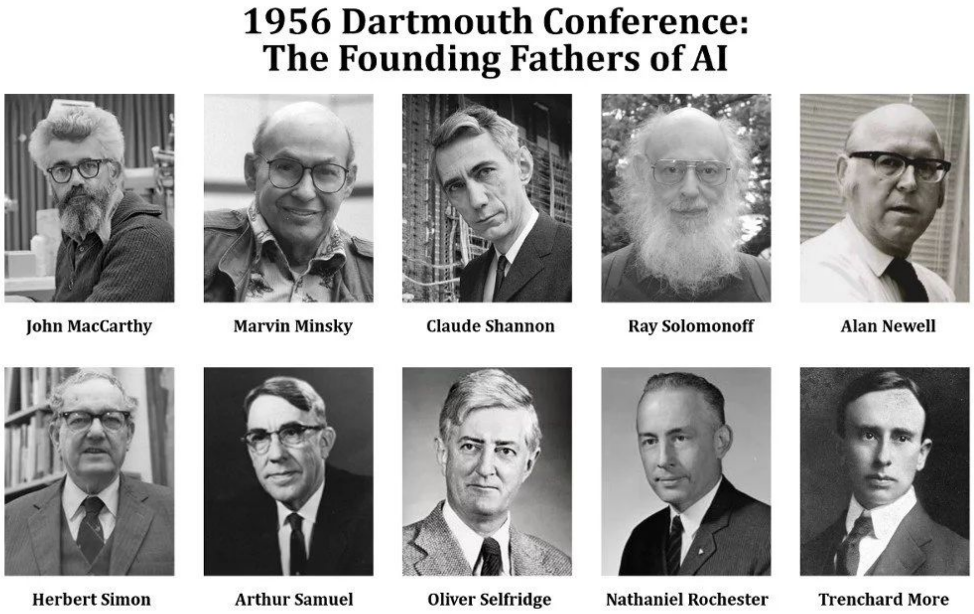
Who's the human and who's the machine?

Turing Test: is it reliable?

- To be intelligent does not necessarily mean to be “human”...
- But humans are intelligent



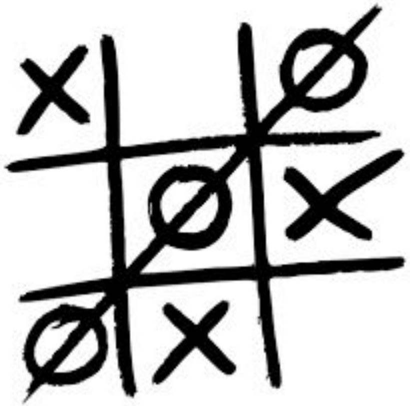
The birth of Artificial Intelligence



AI Early Explorations: 1950s - 1960s

- Goals: identifying specific tasks that require intelligence, and figuring out how to get machines to do them
- High-Level human abilities
  - Reasoning
  - Understand Natural Language
  - Understand Images
- Low-Level abilities:
  - Recognize speech
  - Distinguishing objects in images
  - Reading cursive script

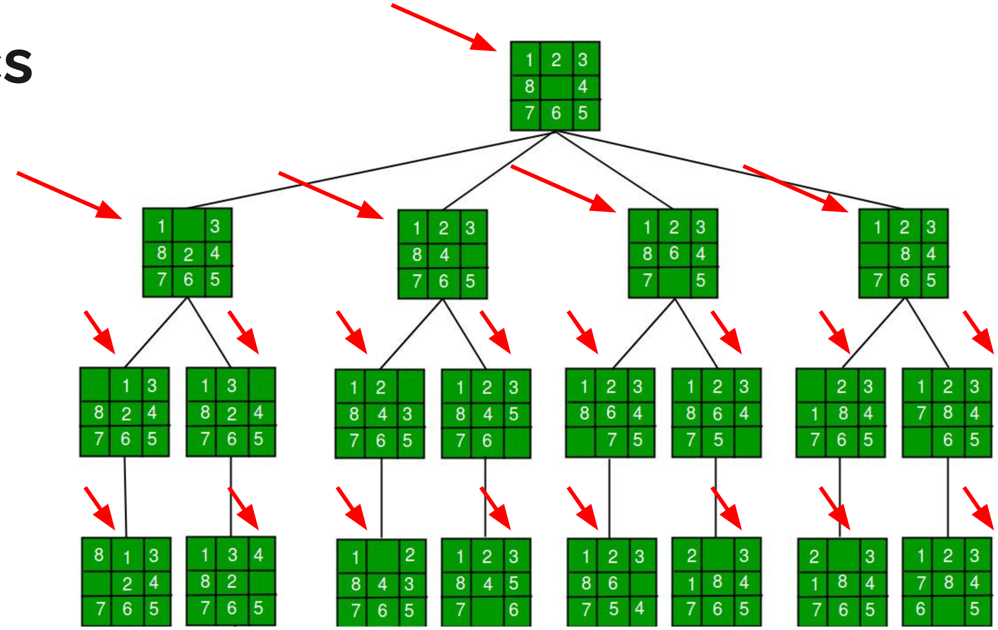
Starting point: Simple Games and Toy Problems



Symbols and Heuristics

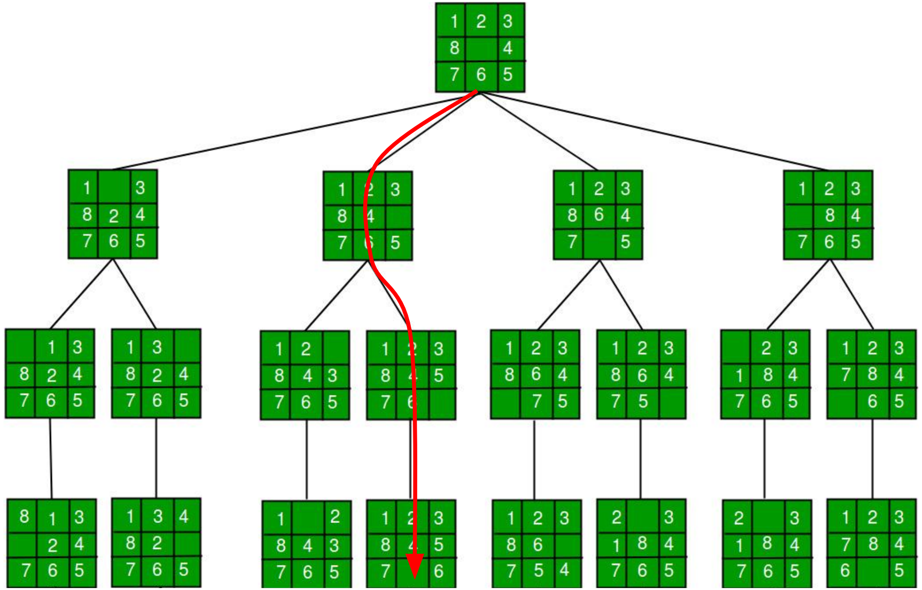
Symbols and Heuristics

- Knowledge Representation: **Lists of Symbols**



Symbols and Heuristics

- Knowledge Representation: **Lists of Symbols**
- Search Methods: **Search tree** and **Heuristics**



Natural Language Processing

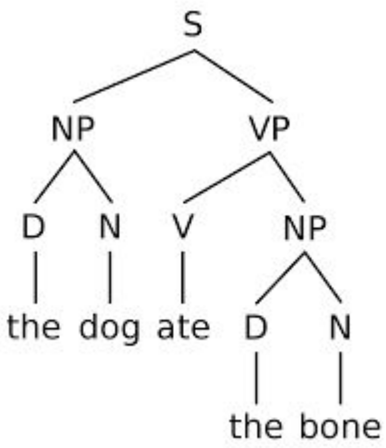
- Understanding, generating and translating natural language

- Difficult problem:
    - Morphology
    - Syntax
    - Semantics
    - Pragmatics

- Parse Tree

Natural Language Processing

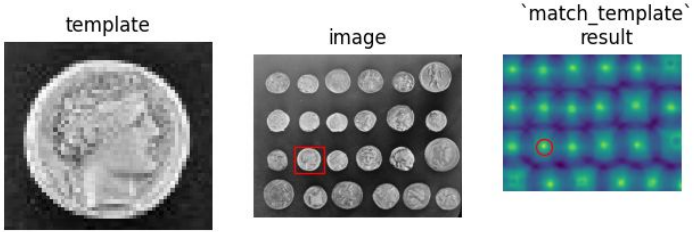
"The dog ate the bone"



- S for **sentence** (Top level)
- NP for **noun** phrase
- VP for **verb phrase** (which serves as the predicate)
- V for **verb**
- D for **determiner**
- N for **noun**

But not only Symbolic...

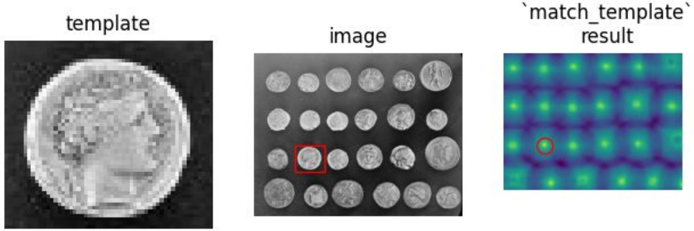
- Pattern Recognition
  - Template Matching
  - Feature extraction

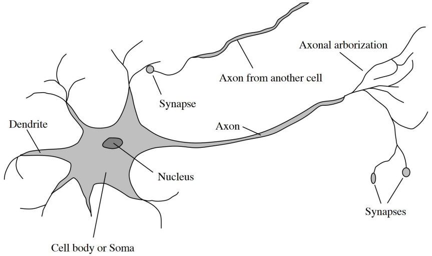


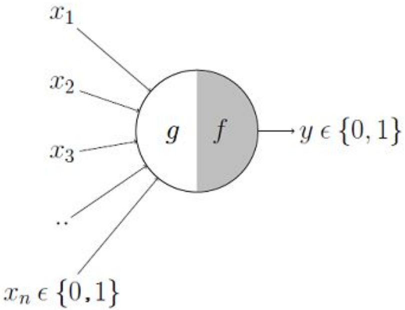
But not only Symbolic...

- Pattern Recognition
  - Template Matching
  - Feature extraction

- McCulloc and Pitts' perceptron
  - Learning adjusting connection weights







AI Expansion: 1960s - 1980s

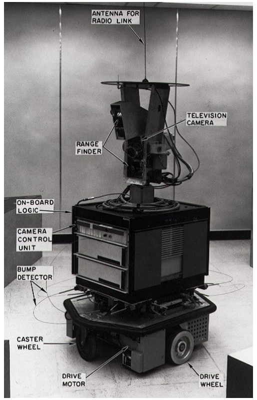
- From toy problems to real world commercial applications
  - Computer Vision
  - Mobile Robots
  - Game Playing
  - Speech Recognition
  - NLP
  - Knowledge Representation

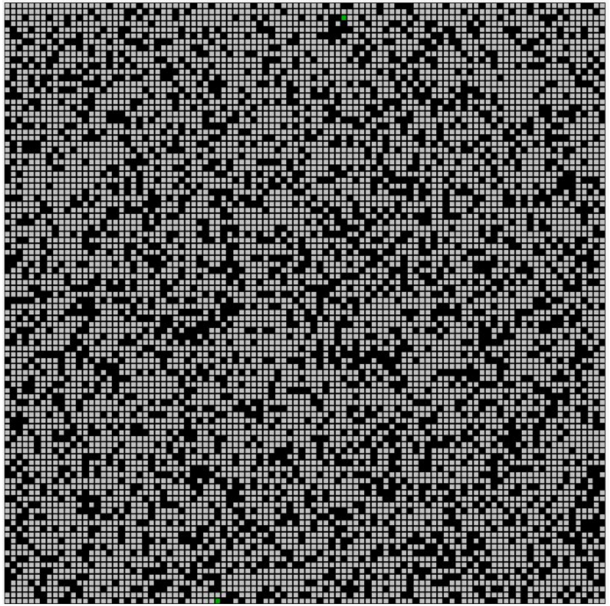
AI Expansion: 1960s - 1980s - NLP

```
  Welcome to
                    EEEEEE  LL      IIII   ZZZZZZ   AAAAA
                    EE      LL       II       ZZ   AA   AA
                    EEEEE   LL       II     ZZZ    AAAAAAA
                    EE      LL       II    ZZ      AA   AA
                    EEEEEE  LLLLLL  IIII ZZZZZZ    AA   AA

  Eliza is a mock Rogerian psychotherapist.
  The original program was described by Joseph Weizenbaum in 1966.
  This implementation by Norbert Landsteiner 2005.


ELIZA: Is something troubling you ?
YOU:   Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:   They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU:   Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU:   He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU:   It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU:
```

AI Expansion: 1960s - 1980s - Mobile Robot

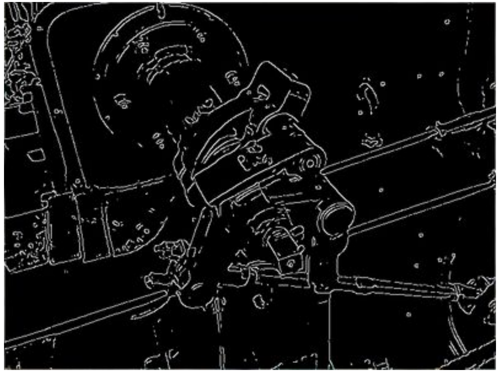
- Route Finding
  - A* Algorithm

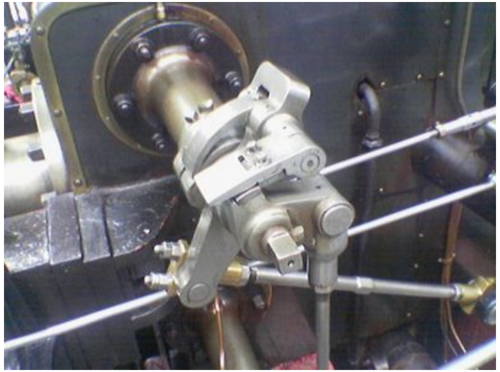


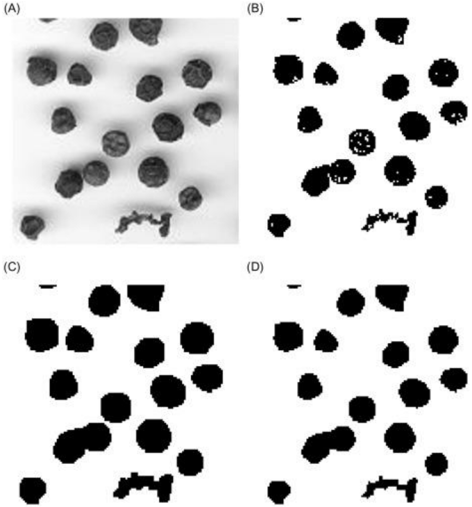


AI Expansion: 1960s - 1980s

- Computer Vision







Expert Systems

- The ***Good Old Fashioned AI (GOFAI)***
- Solving **domain-specific problems** by embedding **expert knowledge** in the form of **logical rules**
- Many success in specific tasks
- But limited achievements for “**general AI**”

Edward Feigenbaum

History of AI: AI Winter and the era of Big Data

AI Winter

- Real-world tasks turned out to require much more “intelligence” than that achievable by heuristic search and symbolic processing:
    - Computational Complexity
    - Background Knowledge

- Drop of interest in AI: reduced goals and less fundings

Statistical Approaches

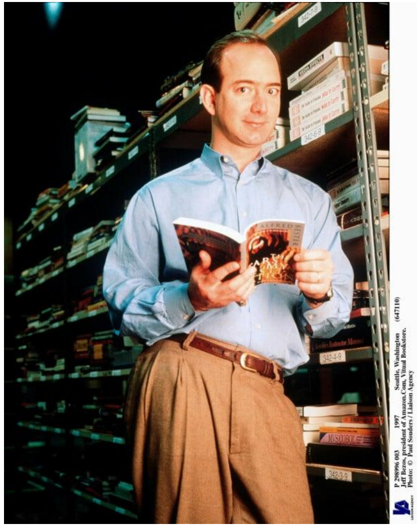
"Every time I fire a linguist, the performance of the speech recognizer goes up"

Frederick Jelinek, 1988

1980s - 1990s: from logic to statistics

- Statistical approaches seemed to overcome logical approaches
- New results in many fields
- Invented the BackPropagation Algorithm (1987)
- CIFAR10 Released (famous dataset for image classification)
- Raising Enthusiasm for statistical and inference based-AI

1990s - 2000s Data and Hardware



"Dearest Amabot,

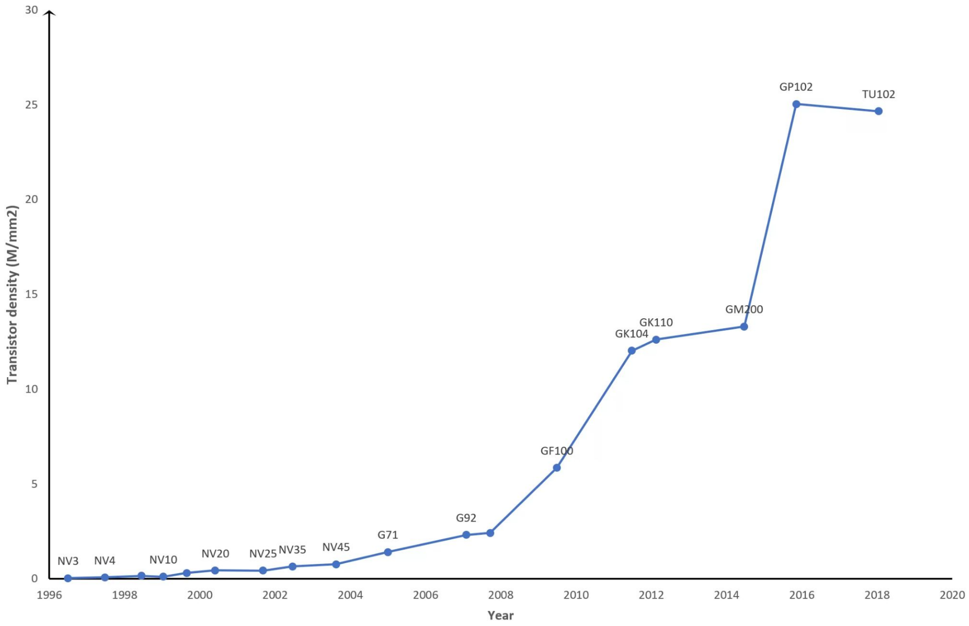
If you only had a heart to absorb our hatred...
Thanks for nothing, you jury-rigged rust bucket.
The gorgeous messiness of flesh and blood with prevail!"

Seattle Weakly, 1999

1990s - 2000s Data and Hardware

- The era of data slowly begins
- Statistical methods can finally exploit large sets of data in the internet

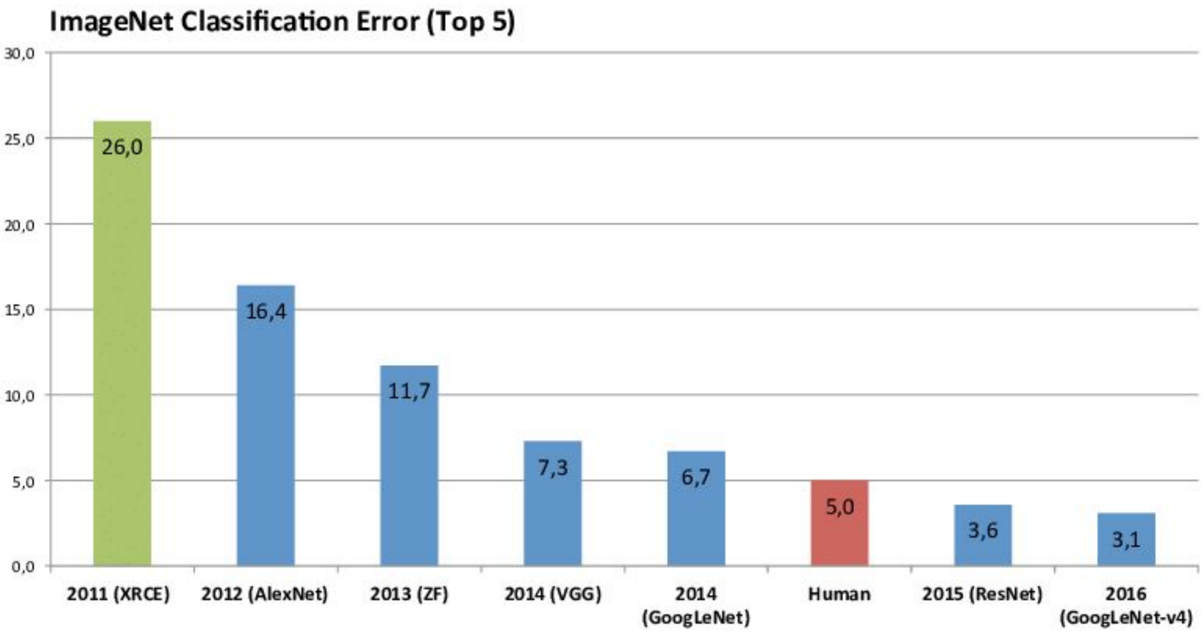
1990s - 2000s Data and Hardware



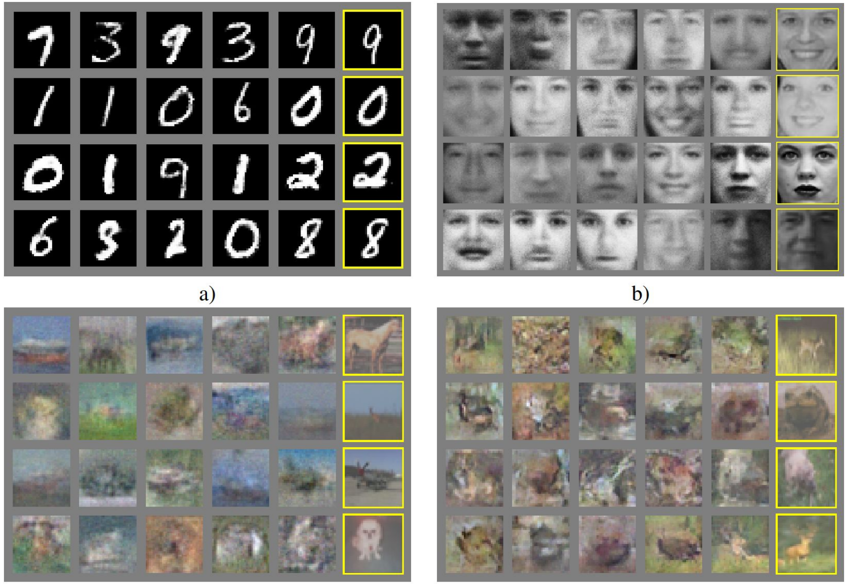
- GPUs become progressively faster in computing data
- The combination of large data sets and powerful machines gives a boost to inference-based AI

ImageNet Dataset

Image Classification Performances



2014 - Generative Adversarial Networks



a) b) c) d)

Generative Adversarial Nets

Ian J. Goodfellow, Jean Pouget-Abadie,* Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair,† Aaron Courville, Yoshua Bengio‡
Département d'informatique et de recherche opérationnelle
Université de Montréal
Montréal, QC H3C 3J7

Abstract

We propose a new framework for estimating generative models via an adversarial process, in which we simultaneously train two models: a generative model G that captures the data distribution, and a discriminative model D that estimates the probability that a sample came from the training data rather than G. The training procedure for G is to maximize the probability of D making a mistake. This framework corresponds to a minimax two-player game. In the space of arbitrary functions G and D, a unique solution exists, with G recovering the training data distribution and D equal to $\frac{1}{2}$ everywhere. In the case where G and D are defined by multilayer perceptrons, the entire system can be trained with backpropagation. There is no need for any Markov chains or unrolled approximate inference networks during either training or generation of samples. Experiments demonstrate the potential of the framework through qualitative and quantitative evaluation of the generated samples.

1 Introduction

The promise of deep learning is to discover rich, hierarchical models [2] that represent probability distributions over the kinds of data encountered in artificial intelligence applications, such as natural images, audio waveforms containing speech, and symbols in natural language corpora. So far, the most striking successes in deep learning have involved discriminative models, usually those that map a high-dimensional, rich sensory input to a class label [14, 22]. These striking successes have primarily been based on the backpropagation and dropout algorithms, using piecewise linear units [19, 9, 10] which have a particularly well-behaved gradient . Deep *generative* models have had less of an impact, due to the difficulty of approximating many intractable probabilistic computations that arise in maximum likelihood estimation and related strategies, and due to difficulty of leveraging the benefits of piecewise linear units in the generative context. We propose a new generative model estimation procedure that sidesteps these difficulties. [1]

In the proposed *adversarial nets* framework, the generative model is pitted against an adversary: a discriminative model that learns to determine whether a sample is from the model distribution or the data distribution. The generative model can be thought of as analogous to a team of counterfeiters, trying to produce fake currency and use it without detection, while the discriminative model is analogous to the police, trying to detect the counterfeit currency. Competition in this game drives both teams to improve their methods until the counterfeits are indistiguishable from the genuine articles.

*Jean Pouget-Abadie is visiting Université de Montréal from Ecole Polytechnique.
†Sherjil Ozair is visiting Université de Montréal from Indian Institute of Technology Delhi
‡Yoshua Bengio is a CIFAR Senior Fellow.
[1] All code and hyperparameters available at http://www.github.com/goodfeli/adversarial

2017 - Transformer Architectures (GPT)

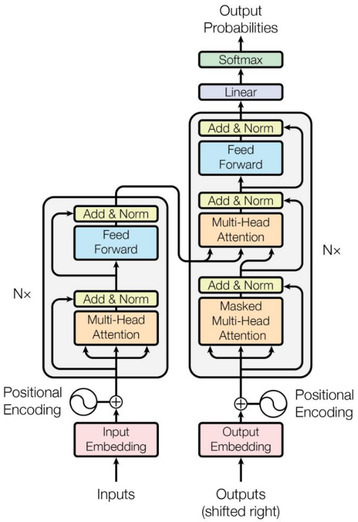
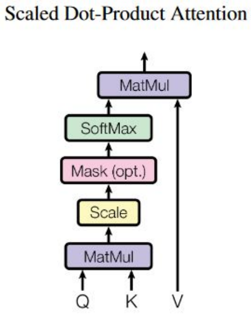
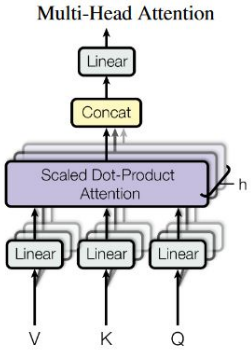


Figure 1: The Transformer - model architecture.





Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com

Noam Shazeer* Google Brain noam@google.com

Niki Parmar* Google Research nikip@google.com

Jakob Uszkoreit* Google Research usz@google.com

Llion Jones* Google Research llion@google.com

Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu

Łukasz Kaiser* Google Brain lukaszkaiser@google.com

Illia Polosukhin* ‡ illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

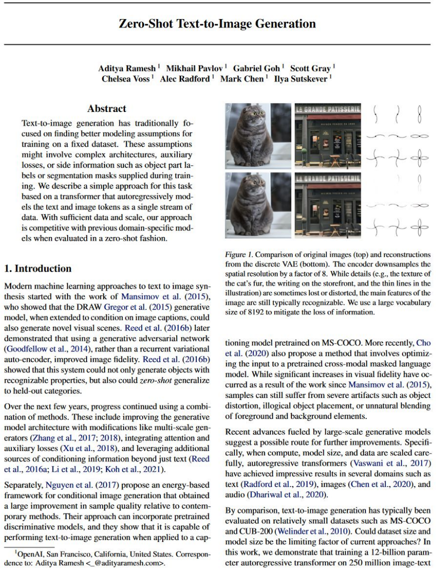
*Equal contribution. Listing order is random. Jakob proposed replacing RNNs with self-attention and started the effort to evaluate this idea. Ashish, with Illia, designed and implemented the first Transformer models and has been crucially involved in every aspect of this work. Noam proposed scaled dot-product attention, multi-head attention and the parameter-free position representation and became the other person involved in nearly every detail. Niki designed, implemented, tuned and evaluated countless model variants in our original codebase and tensor2tensor. Llion also experimented with novel model variants, was responsible for our initial codebase, and efficient inference and visualizations. Lukasz and Aidan spent countless long days designing various parts of and implementing tensor2tensor, replacing our earlier codebase, greatly improving results and massively accelerating our research.

†Work performed while at Google Brain.

‡Work performed while at Google Research.

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

2021 - Zero-Shot Text-to-Image Generation



Zero-Shot Text-to-Image Generation

Aditya Ramesh [1] Mikhail Pavlov [1] Gabriel Goh [1] Scott Gray [1]
Chelsea Voss [1] Alec Radford [1] Mark Chen [1] Ilya Sutskever [1]

Abstract

Text-to-image generation has traditionally focused on finding better modeling assumptions for training on a fixed dataset. These assumptions might involve complex architectures, auxiliary losses, or side information such as object part labels or segmentation masks supplied during training. We describe a simple approach for this task based on a transformer that autoregressively models the text and image tokens as a single stream of data. With sufficient data and scale, our approach is competitive with previous domain-specific models when evaluated in a zero-shot fashion.

Figure 1. Comparison of original images (top) and reconstructions from the discrete VAE (bottom). The encoder downsamples the spatial resolution by a factor of 8. While details (e.g., the texture of the cat's fur, the writing on the storefront, and the thin lines in the illustration) are sometimes lost or distorted, the main features of the image are still typically recognizable. We use a large vocabulary size of 8192 to mitigate the loss of information.

1. Introduction

Modern machine learning approaches to text to image synthesis started with the work of Mansimov et al. (2015), who showed that the DRAW Gregor et al. (2015) generative model, when extended to condition on image captions, could also generate novel visual scenes. Reed et al. (2016b) later demonstrated that using a generative adversarial network (Goodfellow et al., 2014), rather than a recurrent variational auto-encoder, improved image fidelity. Reed et al. (2016b) showed that this system could not only generate objects with recognizable properties, but also could *zero-shot* generalize to held-out categories.

Over the next few years, progress continued using a combination of methods. These include improving the generative model architecture with modifications like multi-scale generators (Zhang et al., 2017; 2018), integrating attention and auxiliary losses (Xu et al., 2018), and leveraging additional sources of conditioning information beyond just text (Reed et al., 2016a; Li et al., 2019; Koh et al., 2021).

Separately, Nguyen et al. (2017) propose an energy-based framework for conditional image generation that obtained a large improvement in sample quality relative to contemporary methods. Their approach can incorporate pretrained discriminative models, and they show that it is capable of performing text-to-image generation when applied to a captioning model pretrained on MS-COCO. More recently, Cho et al. (2020) also propose a method that involves optimizing the input to a pretrained cross-modal masked language model. While significant increases in visual fidelity have occurred as a result of the work since Mansimov et al. (2015), samples can still suffer from severe artifacts such as object distortion, illogical object placement, or unnatural blending of foreground and background elements.

Recent advances fueled by large-scale generative models suggest a possible route for further improvements. Specifically, when compute, model size, and data are scaled carefully, autoregressive transformers (Vaswani et al., 2017) have achieved impressive results in several domains such as text (Radford et al., 2019), images (Chen et al., 2020), and audio (Dhariwal et al., 2020).

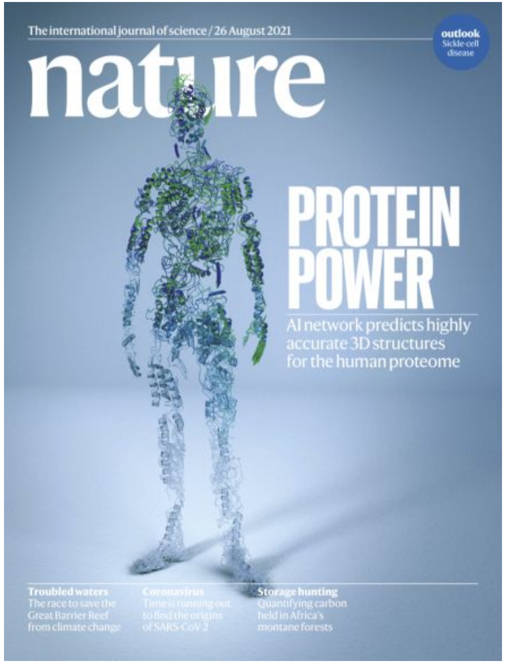
By comparison, text-to-image generation has typically been evaluated on relatively small datasets such as MS-COCO and CUB-200 (Welinder et al., 2010). Could dataset size and model size be the limiting factor of current approaches? In this work, we demonstrate that training a 12-billion parameter autoregressive transformer on 250 million image-text

[1]OpenAI, San Francisco, California, United States. Correspondence to: Aditya Ramesh <_@adityaramesh.com>.

AI Today

AI Today


The international journal of science / 6 October 2022
nature
MATRIX GAMES
Deep reinforcement learning opens route to faster algorithms for matrix multiplication
Protecting Peru
Can technology help Indigenous groups preserve the Amazon?
Invisible touch
How marine clouds are affected by aerosols emitted from shipping
Preferential practice
US universities favour prestige in faculty hiring and retention


The international journal of science / 26 August 2021
outlook
Sickle-cell disease
nature
PROTEIN POWER
AI network predicts highly accurate 3D structures for the human proteome
Troubled waters
The race to save the Great Barrier Reef from climate change
Coronavirus
Time is running out to find the origins of SARS-CoV-2
Storage hunting
Quantifying carbon held in Africa's montane forests


The international journal of science / 31 August 2023
nature
DRONE RACER
AI pilot beats human champions in aerial contest
Offset agreement
Overhaul pricing of carbon credits to help fund climate projects
Dining companions
Corals devour algal partners when food supplies run low
Sentence synthesis
Brain implants show promise in rendering speech from thoughts

AI Today

You
Genera una breve poesia ermetica sui culurgiones

ChatGPT
Tra le pieghe di pasta, sussurri nascosti,
Segreti di terra in bocconi avvolti.
Culurgiones, custodi di antiche storie,
In silenzio raccontano di memorie.

Sapori avvolti in mistero, sogni d'orzo e grano,
Nascosti tra i solchi di un mare lontano.
Nel loro abbraccio, il tempo si placa,
E l'anima s'immerge in dolce placida brama.

Saggi custodi di tradizioni antiche,
Racconti di terre lontane, di ricche fatiche.
In un boccone di culurgiones svelati,
Segreti dell'anima, sogni incantati.

You
Traducila in sardo

ChatGPT
Sotto a sos pliegues de pasta, sussurros iscondidos,
Segretos de terra in morsi abrazados.
Culurgiones, custodes de is stòrias antichas,
In su silenzio contan de memòrias.

Sabores abrazados in misteriu, sognos de orzo e trigu,
Iscondidos entre is sulas de unu mari lontanu.
In su su so abrazu, su tempus si tranquillat,
E s'anima si ischiattat in dolce brama placida.

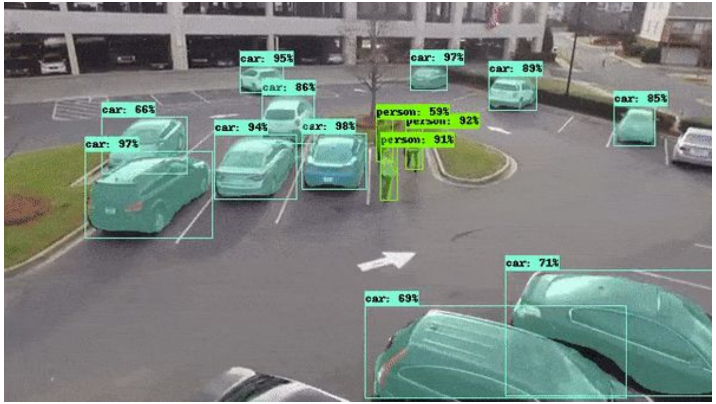
Saggios custodes de traditziones antichas,
Stòrias de terras lontanas, de riques fatigas.
In unu morsu de culurgiones isvelados,
Segretos de s'anima, sognos incantados.

You
ora raccontala usando solo emoji

ChatGPT
🍝🤫🌾🌊🕵️📜🥟🔒🧠🤐🔍🍴✨📖

AI Today

AI Today


car: 95%
car: 97%
car: 89%
car: 86%
car: 85%
car: 66%
person: 59%
person: 92%
car: 94%
car: 98%
person: 91%
car: 97%
car: 71%
car: 69%