Artificial Intelligence

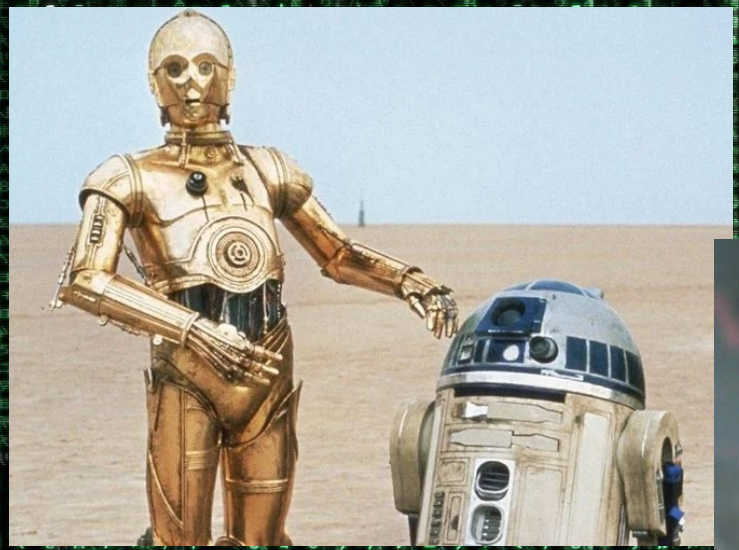
Lecture 12-1

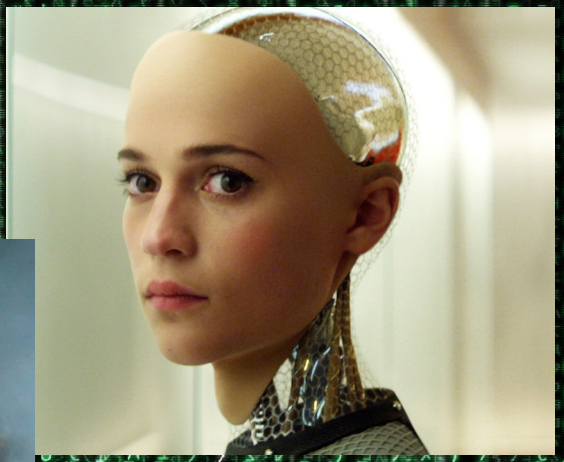
Computers & Society (CPSC 430)

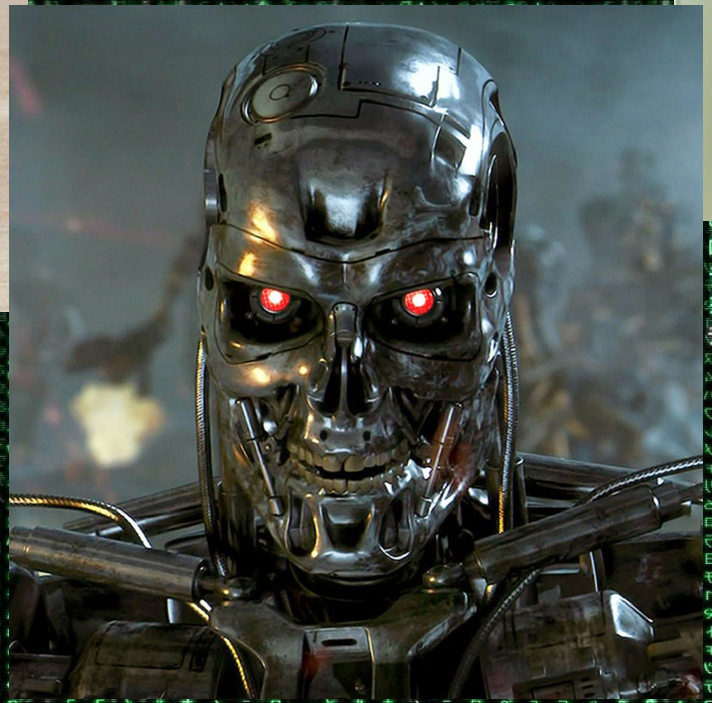
Hedayat Zarkoob
https://www.students.cs.ubc.ca/~cs-430/

Original slides by Kevin Leyton-Brown on Understanding Artificial Intelligence.

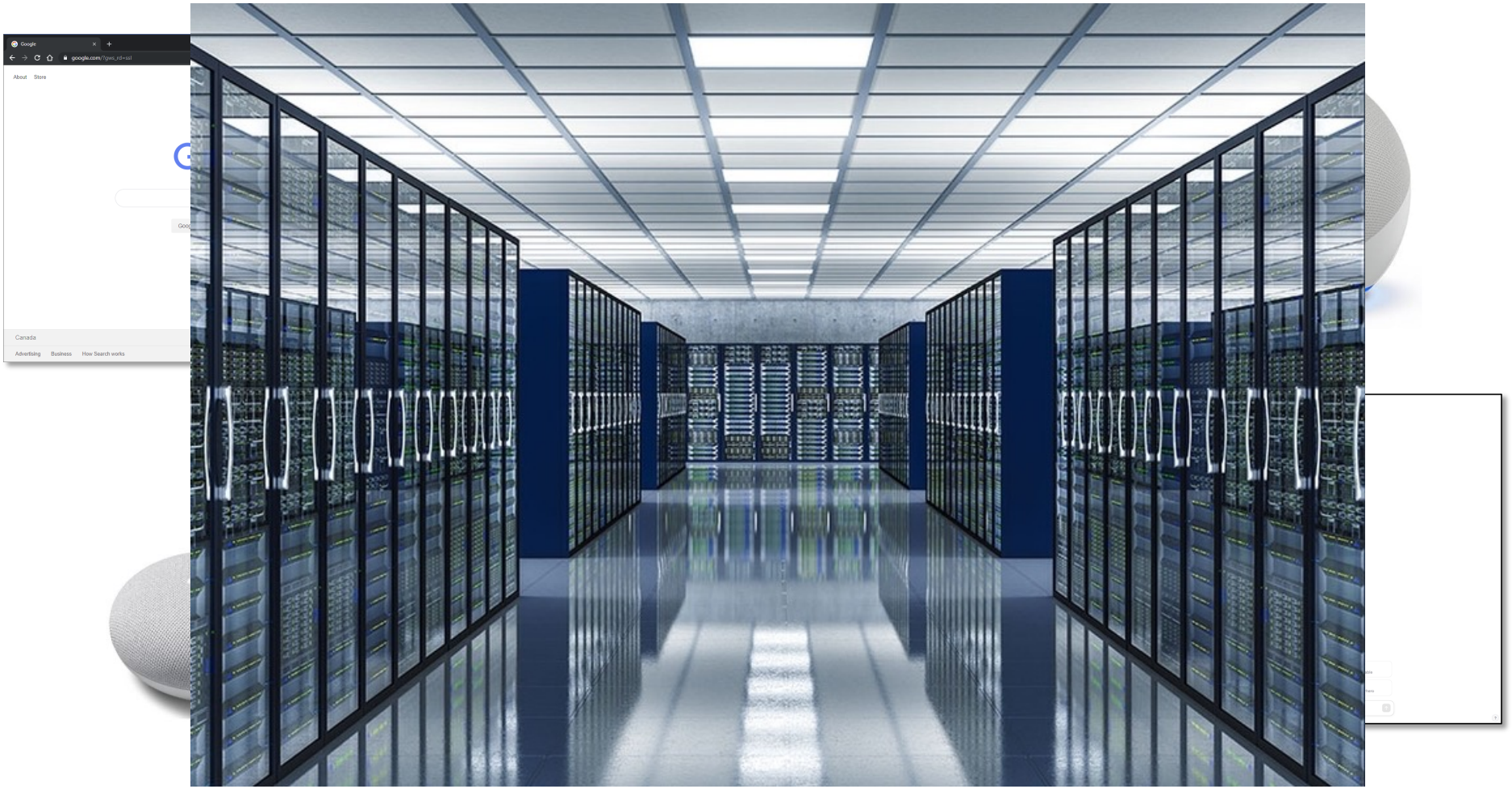
Movies Help Us Think About AI







But most modern AI systems look like…



What is Artificial Intelligence?

AI is the study, design, and development of computational processes to solve problems that **previously required human intelligence**

The "**AI Paradox**": once we become familiar with a technology, we stop considering it AI

“Good Old-Fashioned AI”

- Early AI systems were **explicitly programmed**
  - reasoning systems were based on logic
  - rule-based “expert systems”
  - language systems explicitly modeled grammar
  - vision systems reasoned about optics, geometry
- Many important **conceptual foundations**
- Few **practical successes**
  - systems were brittle in practice
  - dealt poorly with noise, imperfect world models

A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence

August 31, 1955

John McCarthy, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon

1955

Search

- Instead of telling a computer how to solve a problem, tell it how to **recognize a solution & let it experiment**

- Drove many of AI's **early successes**:

1982

ScienceDirect

Outline Get Access Share Export

Artificial Intelligence

Volume 19, Issue 3, November 1982, Pages 279-320

ELSEVIER

A world-championship-level Othello program ☆

Paul S. Rosenbloom

Show more

https://doi.org/10.1016/0004-3702(82)90003-0 Get rights and content

Abstract

Othello is a recent addition to the collection of games that have been examined within artificial intelligence. Advances have been rapid, yielding programs that have reached the level of world-championship play. This article describes the current champion Othello program, IAGO. The work described here includes: (1) a task analysis of Othello; (2) the implementation of a program based on this analysis and state-of-the-art AI game-playing techniques; and (3) an evaluation of the program's performance through games played against other programs and comparisons with expert human play.

1996

AI Magazine Volume 17 Number 1 (1996) (© AAAI)

CHINOOK

The World Man-Machine Checkers Champion

Jonathan Schaeffer, Robert Lake, Paul Lu, and Martin Bryant

■ In 1992, the seemingly unbeatable World Checker Champion Marion Tinsley defended his title against the computer program CHINOOK. After an intense, tightly contested match, Tinsley fought back from behind to win the match by scoring four wins to CHINOOK's two, with 33 draws. This match was the first time in history that a human world champion defended his title against a computer. This article reports on the progress of the checkers (8 3 8 draughts) program CHINOOK since 1992. Two years of research and development on the program culminated in a rematch with Tinsley in August 1994. In this match, after six games (all draws), Tinsley withdrew from the match and the American Checker Federation (ACF), CHINOOK was allowed to play in the 1990 U.S. championship. This biennial event attracts the best players in the world, with the winner earning the right to play a match for the world championship.

CHINOOK came in an undefeated second in the tournament behind the world champion, Marion Tinsley. The four individual games between CHINOOK and Tinsley were drawn. This placing earned CHINOOK the right to challenge Tinsley for the world championship,

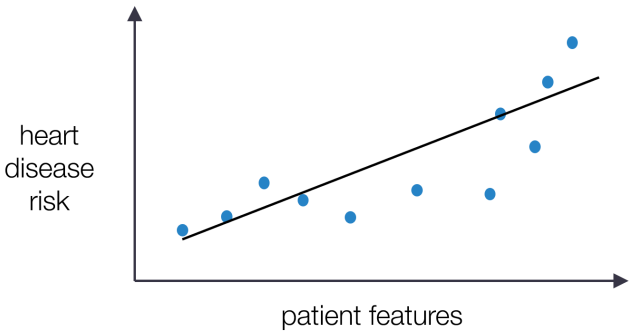
1997

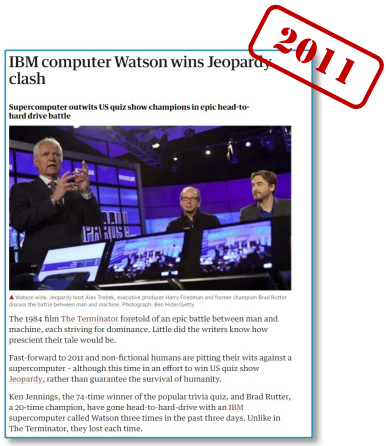
Deep Blue

On May 11, 1997, an IBM computer called IBM ® Deep Blue ® beat the world chess champion after a six-game match: two wins for IBM, one for the champion and three draws. The match lasted several days and received massive media coverage around the world. It was the classic plot line of man vs. machine. Behind the contest, however, was important computer science, pushing forward the ability of computers to handle the

Machine Learning

- Give a computer **examples of a pattern** and ask it to find a **rule**
- x: **features**; y: **labels**
- Example:
  - x = blood pressure, diet, exercise, age, gender
  - y = risk of heart disease




2011

IBM computer Watson wins Jeopardy clash

Supercomputer outwits US quiz show champions in epic head-to-hard drive battle

The 1984 film The Terminator foretold of an epic battle between man and machine, each striving for dominance. Little did the writers know how prescient their tale would be.

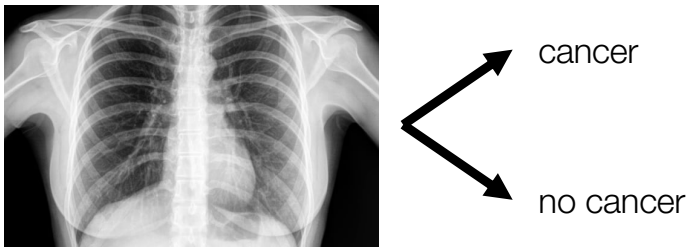
Fast-forward to 2011 and non-fictional humans are pitting their wits against a supercomputer - although this time in an effort to win US quiz show Jeopardy, rather than guarantee the survival of humanity.

Ken Jennings, the 74-time winner of the popular trivia quiz, and Brad Rutter, a 20-time champion, have gone head-to-hard-drive with an IBM supercomputer called Watson three times in the past three days. Unlike in The Terminator, they lost each time.

Deep Learning with Neural Networks

- Get **rid of features**!
  - build machine learning models that take raw inputs like pictures, sound recordings, text, …
- Architecture is loosely analogous to **brains**
- An **old idea** (60s; 80s)
  - Fundamental benefit: scalable model complexity

- Example:
  - x = lung X-ray image
  - y = lung cancer diagnosis



Image, Face Recognition

- Understanding **images and faces** had long been seen as a fundamentally hard AI problem
- Deep learning was a **game changer**

2014

DeepFace: Closing the Gap to Human-Level Performance in Face Verification

Yaniv Taigman Ming Yang Marc'Aurelio Ranzato Lior Wolf

Facebook AI Research
Menlo Park, CA, USA
{yaniv, mingyang, ranzato}@fb.com

Tel Aviv University
Tel Aviv, Israel
wolf@cs.tau.ac.il

Abstract

In modern face recognition, the conventional pipeline consists of four stages: detect ⇒ align ⇒ represent ⇒ classify. We revisit both the alignment step and the representation step by employing explicit 3D face modeling in order to apply a piecewise affine transformation, and derive a face representation from a nine-layer deep neural network. This deep network involves more than 120 million parameters using several locally connected layers without weight sharing, rather than the standard convolutional layers. Thus we trained it on the largest facial dataset to-date, an identity labeled dataset of four million facial images belonging to more than 4,000 identities. The learned representations coupling the accurate model-based alignment with the large facial database generalize remarkably well to faces in unconstrained environments, even with a simple classifier. Our method reaches an accuracy of 97.35% on the Labeled Faces in the Wild (LFW) dataset, reducing the error of the current state of the art by more than 27%, closely approaching human-level performance.

1. Introduction

Face recognition in unconstrained images is at the forefront of the algorithmic perception revolution. The social and cultural implications of face recognition technologies are far reaching, yet the current performance gap in this domain between machines and the human visual system serves as a buffer from having to deal with these implications.

toward tens of thousands of appearance features in other recent systems [5, 7, 2].

The proposed system differs from the majority of contributions in the field in that it uses the deep learning (DL) framework [3, 21] in lieu of well engineered features. DL is especially suitable for dealing with large training sets, with many recent successes in diverse domains such as vision, speech and language modeling. Specifically with faces, the success of the learned net in capturing facial appearance in a robust manner is highly dependent on a very rapid 3D alignment step. The network architecture is based on the assumption that once the alignment is completed, the location of each facial region is fixed at the pixel level. It is therefore possible to learn from the raw pixel RGB values, without any need to apply several layers of convolutions as is done in many other networks [19, 21].

In summary, we make the following contributions : (i) The development of an effective deep neural net (DNN) architecture and learning method that leverage a very large labeled dataset of faces in order to obtain a face representation that generalizes well to other datasets; (ii) An effective facial alignment system based on explicit 3D modeling of faces; and (iii) Advance the state of the art significantly in (1) the Labeled Faces in the Wild benchmark (*LFW*) [18], reaching near human-performance; and (2) the YouTube Faces dataset (*YTF*) [30], decreasing the error rate there by more than 50%.

1.1. Related Work

Big data and deep learning In recent years, a large number of photos have been crawled by search engines, and un-

2015

Microsoft, Google Beat Humans at Image Recognition

Deep learning algorithms compete at ImageNet challenge

By R. Colin Johnson, 02.18.15 14

Share Post Share on Facebook Share on Twitter in

PORTLAND, Ore. -- First computers beat the best of us at chess, then poker, and finally Jeopardy. The next hurdle is image recognition — surely a computer can't do that as well as a human. Check that one off the list, too. Now Microsoft has programmed the first computer to beat the humans at image recognition.

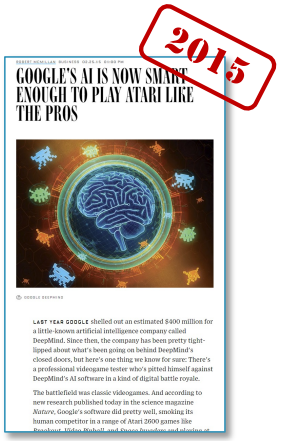
The competition is fierce, with the ImageNet Large Scale Visual Recognition Challenge doing the judging for the 2015 championship on December 17. Between now and then expect to see a stream of papers claiming they have one-upped humans too. For instance, only 5 days after Microsoft announced it had beat the human benchmark of 5.1% errors with a 4.94% error grabbing neural network, Google announced it had one-upped Microsoft by 0.04%.

The top row is a representative of the categories that Microsoft's algorithm found in the database and the image columns below are examples that fit.
(Source: Microsoft)

ImageNet, with hundreds of object categories and millions of example images, has been running the competition since 2010 with about 50 institutions competing, but this is the first year than a computer will take the crown from the best human score. All the contestants are using what today is called deep learning algorithms, which are all derived from various versions of artificial neural networks which mimic the way the human brain works to varying degrees. Most of the contestants freely provide papers describing their

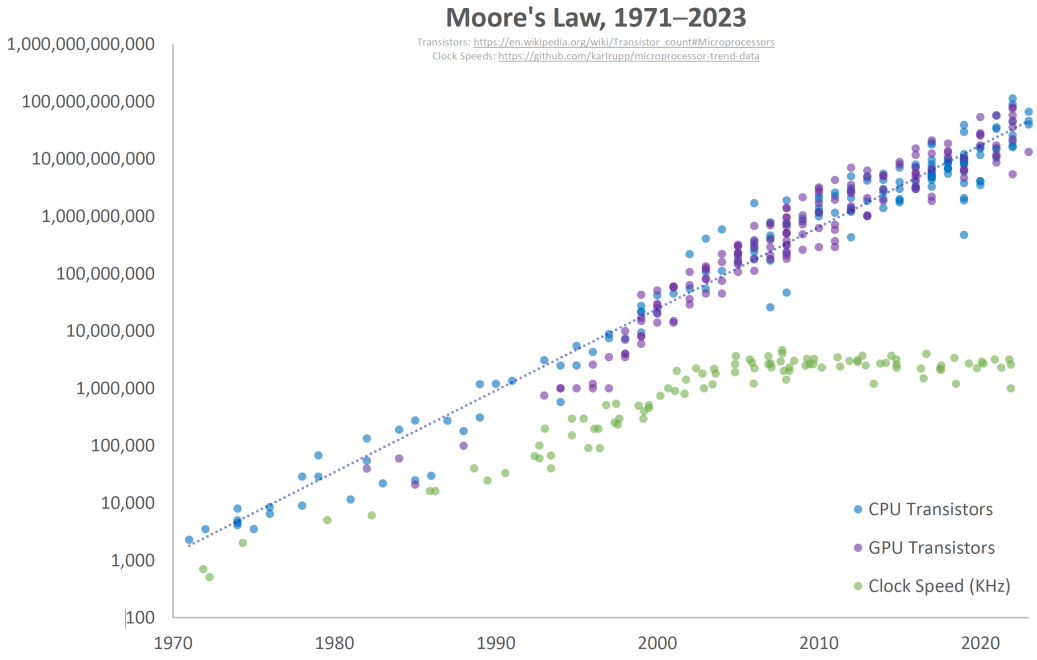
Reinforcement Learning

- Often a complex **sequence of actions** must be taken before reaching a **reward or punishment**
  - RL: an ML approach for such settings
- Example:
  - navigate a maze to reach a goal
  - you need a key to unlock the door
  - quicksand slows you down
- **Foundations** of RL laid in the 80s
- **Breakthroughs** in mid 2010s:
  - state representation using deep learning
  - Monte Carlo Tree Search
  - new policy search algorithms



Why is it happening?

1. Scientific/mathematical **breakthroughs**, esp in machine learning
2. Growth in raw **computing power**



For comparison:

Species	# Neurons in Brain
Fruit Fly	100,000
Cat	1,000,000,000
Chimpanzee	10,000,000,000
Human	100,000,000,000

Self-Supervision and Generative AI

- How can we learn from **huge, unlabeled datasets**?
  - traditional ML needs class labels
  - RL needs rewards
- A really clever idea: turn raw data into **puzzles**

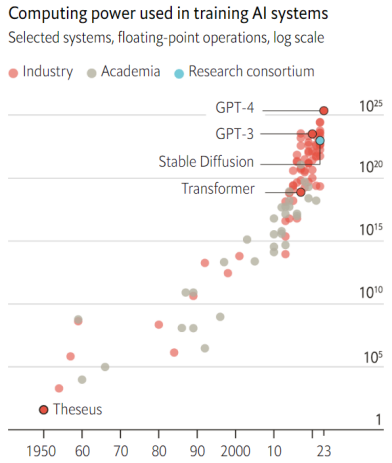
> Stanley Park has a long history. The land was originally used by Indigenous peoples for thousands of years before British Columbia was colonized by the British during the 1858 Fraser Canyon Gold Rush and was one of the first areas to be

Claymation zebra standing in front of a chalkboard holding a yellow book

- The same idea works with **images**
  - add noise and ask the model to remove it
- Eventually, models can **generate** inputs like those they saw during training
  - give the first few lines of an essay and the model will complete it
  - start with random noise and a text description and ask the model to denoise it

Large Language Models (GPT-3, ...)

- Start with a huge corpus of text
  - the entire Internet, filtered for link spam
  - books and news
  - code repositories
  - transcripts of conversations
- Take an absolutely enormous neural net
  - GPT-2: 1.5B parameters
  - GPT-3: 175B parameters
  - GPT-4: 8 * 220B = 1.76T parameters
- Spend literally millions of dollars conducting self-supervised training
- The resulting “foundation model” can be “fine-tuned” to specific tasks
  - e.g., sentiment analysis on Twitter: X = tweet; Y = positive/negative



Absolutely Amazing Performance

You
write a python code that take an image as an input and changes the resolution of it.

ChatGPT
You can use the Python Imaging Library (PIL) or its fork Pillow to resize images. Here's a simple example using Pillow:

python — Copy code

```python
from PIL import Image

def resize_image(input_image_path, output_image_path, new_width, new_height):
    # Open the image file
    image = Image.open(input_image_path)

    # Resize the image
    resized_image = image.resize((new_width, new_height))

    # Save the resized image
    resized_image.save(output_image_path)

# Example usage:
input_image = "input_image.jpg"
output_image = "output_image.jpg"
new_width = 800
new_height = 600

resize_image(input_image, output_image, new_width, new_height)
```

Replace `"input_image.jpg"`, `"output_image.jpg"`, `new_width`, and `new_height` with your desired values. This code will resize the input image to the specified dimensions and save it as the output image.

You
It is possible to program responses into computers that simulate human emotions. For example, when a computer taking on the role of a nurse hears a parent say, " My child has diahrrea," it can respond, " I'm sorry to hear that." Studies have shown that people can develop an emotional bond with machines that appear to demonstrate human feelings such as empathy. write a two paragraph essay in which it is argued that it is wrong to encourage these attachments by programming computers to mimic human emotions using virtue ethics.

ChatGPT
Encouraging attachments to computers programmed to mimic human emotions raises ethical concerns when viewed through the lens of virtue ethics. Virtue ethics prioritizes the development of virtuous character traits, such as honesty, empathy, and authenticity, which are fundamental to ethical behavior and moral flourishing. By programming computers to simulate human emotions, there is a risk of perpetuating deception and undermining the virtue of honesty. Individuals may form emotional bonds with these machines based on false pretenses, as the computers do not genuinely experience emotions. This deception erodes trust and authenticity in human-computer interactions, hindering the cultivation of genuine virtues.

Furthermore, encouraging attachments to emotionally mimetic computers can lead to a distortion of human relationships and a dependency on technology for emotional support. Virtue ethics emphasizes the importance of authentic human connections characterized by empathy, compassion, and mutual understanding. By fostering emotional bonds with machines, individuals may prioritize artificial interactions over genuine human relationships, leading to a degradation of interpersonal virtues. Additionally, dependency on emotionally mimetic computers for emotional support may detract from the cultivation of inner virtues, such as resilience and emotional intelligence, which are essential for moral development and ethical decision-making. In essence, virtue ethics cautions against encouraging attachments to computers programmed to mimic human emotions, as it undermines honesty, authenticity, and genuine human relationships, ultimately hindering the cultivation of virtuous character traits.

Weakness: Hallucinations, Reasoning, Math, …

Lawyers have real bad day in court after citing fake cases made up by ChatGPT

Lawyers fined $5K and lose case after using AI chatbot "gibberish" in filings.

JON BRODKIN - 6/23/2023, 10:32 AM

Enlarge

189

A federal judge tossed a lawsuit and issued a $5,000 fine to the plaintiff's lawyers after they used ChatGPT to research court filings that cited six fake cases invented by the artificial intelligence tool made by OpenAI.

Lawyers Steven Schwartz and Peter LoDuca of the firm Levidow, Levidow, & Oberman "abandoned their responsibilities when they submitted non-existent judicial opinions with fake quotes and citations created by the artificial intelligence tool ChatGPT, then continued to stand by the fake opinions after judicial orders called their existence into question," US District Judge Kevin Castel wrote in an order yesterday. The lawyers, Castel wrote, "advocated for the fake cases and legal arguments" even "after being informed by their adversary's submission that their citations were non-existent and could not be found."

The judge issued one fine of $5,000 to be paid by the two lawyers and their firm under joint and several liability. More embarrassingly for the lawyers, they are required to send letters to six real judges who were "falsely identified as the author of the fake" opinions cited in their legal filings. Castel described the legal analysis in one of the fake cases as "gibberish."

A Brave, New World?

- AI has huge potential **upside**
  - eliminating **drudgery**
  - giving everyone a **personal assistant**
  - new **breakthroughs** in critical sectors:
    - drug discovery; green energy
    - education; transportation; …
- If you're skeptical about **economic growth**, ask yourself whether you'd prefer the standard of living from any previous time in history. Consider:
  - healthcare; entertainment;
  - communications; travel, …
- E.g., whose standard of living would you prefer: **yours or Louis XIV's**?

Ethics of AI

Will a new technology:

- disempower **individuals vs corporations**?
  - ⇒ user modeling; data mining; fostering addictive behaviors; developmental effects on children
- disempower **individuals vs governments**?
  - ⇒ facilitate disinformation (deep fakes; bots masquerading as people; filter bubbles); enable qualitatively new military or security tactics
- take **autonomous actions** in a way that obscures responsibility
  - ⇒ autonomous weapons; self-driving cars; loan approval systems
- disproportionately affect **vulnerable/marginalized groups**
  - ⇒ automated decision making tools trained in ways that may encode existing biases

Concern: Economic Impact

- Many tasks will be automated:
  - Increased mechanization of **routine labor**
  - Automation of lower-end **knowledge work**
  - Industrial revolution was about turning people into big machines
- Some jobs are profoundly AI-proof
  - **human touch:** daycare; eldercare; coach; psychiatrist
  - **regulatory & cultural barriers:** lawyer; doctor; priest
  - **not worth it:** gardener; massage therapist; chef
- Automation is nothing new
  - lump of labour fallacy
  - like other waves of innovation AI will almost certainly make society as a whole **much richer**
  - but, it may also exacerbate **income inequality**
- Ask “Is there still something I’d like a person to do for me for free?” If no:
  - **Post-scarcity** economy?
  - Human labor as a **luxury good**?

Concern: Social Impact

- How will AI technologies **transform society**?
- Will there be a **social backlash** against AI?
  - If so, what will be considered AI?
- We're raising **children** taking for granted many technologies that strike us as magical
- How will **human relationships** change in the presence of always-available social agents?

- As we are increasingly **augmented by AI**, what are our inherent cognitive/emotional/motivational limitations, beyond which augmentation won't help?

Concern: Superhuman Intelligence

- AI systems are increasingly capable of **human-level performance**
- **Superhuman intelligence** isn't such a foreign, scary thing
  - governments, corporations, NGOs exhibit behavior much more sophisticated and complex than that of any individual
- Many important problems need superhuman intelligence; AI can help
  - improved **collective decision making**
  - more efficient use of **scarce resources**
  - addressing **underserved communities**
  - **climate change**; other societal challenges