Artificial Intelligence and Deep Learning

François Fleuret
Machine Learning group
https://mlg.unige.ch

UNIVERSITÉ DE GENÈVE

- AI “programs itself”
- AI actually works
- AI requires vast amounts of data and computation
- AI is easy to deploy
- AI models are black boxes
- Current trends

AI "programs itself"

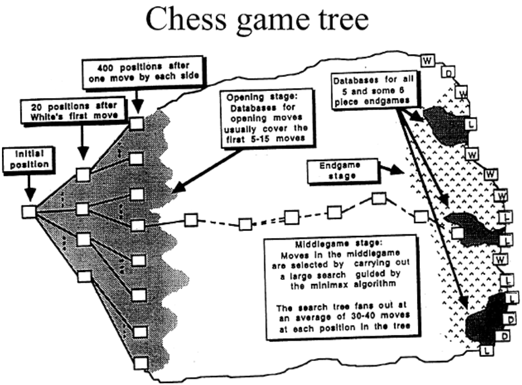
AI "programs itself"

The traditional way of making a computer perform a task is to indicate exactly what simple individual steps have to be executed.

```
n = 15345

while n > 1:
    for k in range(2, n+1):
        if n%k == 0:
            print(k)
            n = n // k
            break
```

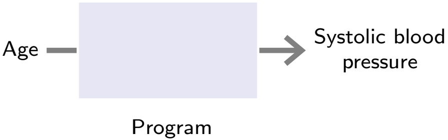
The first attempts at artificial intelligence relied on the same principle *e.g.* medical decision, strategy games, or computer vision.



(Newborn, 1996)

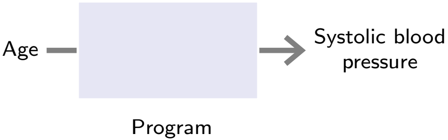
AI "programs itself"

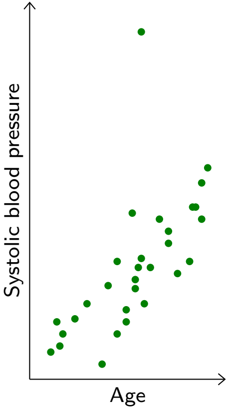
The fundamental idea of machine learning is to automatically tune a program to make it work well on known examples.



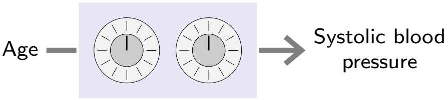
AI "programs itself"

The fundamental idea of machine learning is to automatically tune a program to make it work well on known examples.

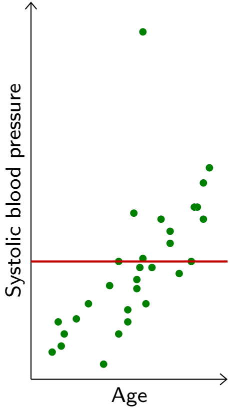




The fundamental idea of machine learning is to automatically tune a program to make it work well on known examples.

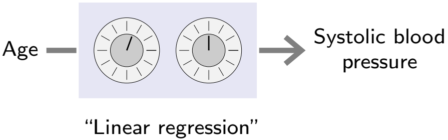


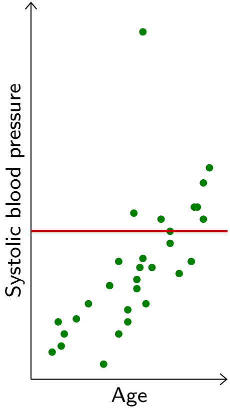
"Linear regression"



AI "programs itself"

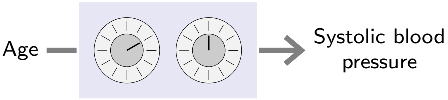
The fundamental idea of machine learning is to automatically tune a program to make it work well on known examples.



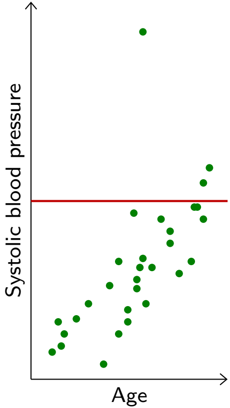


AI "programs itself"

The fundamental idea of machine learning is to automatically tune a program to make it work well on known examples.

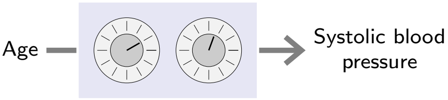


"Linear regression"

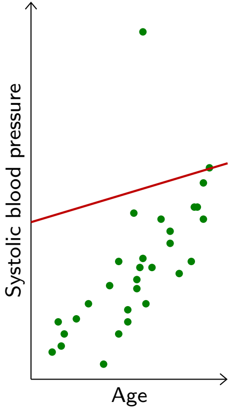


AI "programs itself"

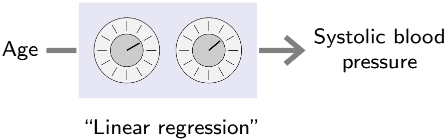
The fundamental idea of machine learning is to automatically tune a program to make it work well on known examples.

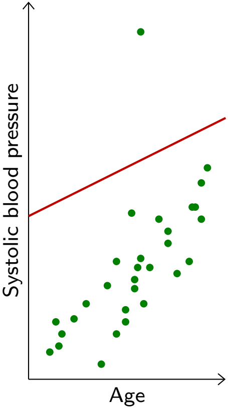


"Linear regression"



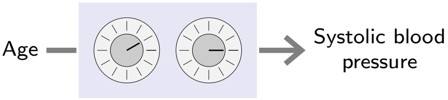
The fundamental idea of machine learning is to automatically tune a program to make it work well on known examples.



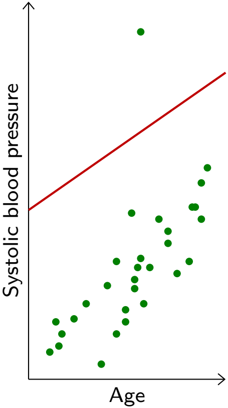


AI "programs itself"

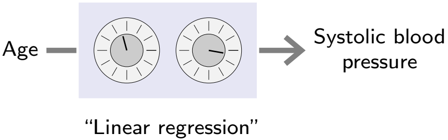
The fundamental idea of machine learning is to automatically tune a program to make it work well on known examples.

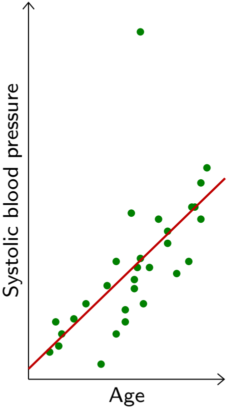


"Linear regression"

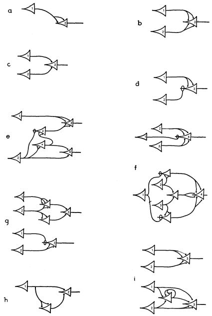


The fundamental idea of machine learning is to automatically tune a program to make it work well on known examples.

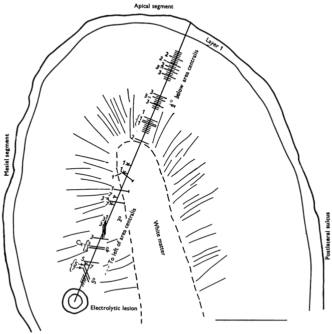




This strategy mimics in some ways the plasticity of neural networks.



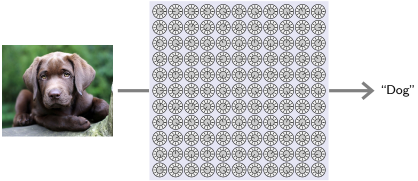
(McCulloch and Pitts, 1943)



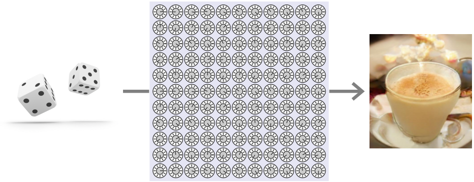
(Hubel and Wiesel, 1962)

AI “programs itself”

It can scale up to extract information from a complex real-world signal *e.g.* an image, sound sample, piece of text

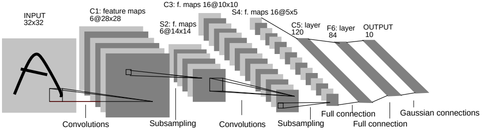


or to synthesize a complex signal



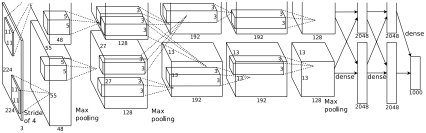
AI "programs itself"

Modern models are parameterized by $10^5 - 10^{11}$ parameters.

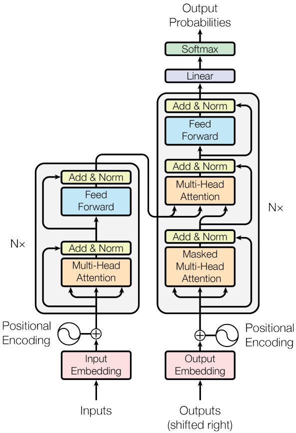


Convolutions (LeNet, 1989)

. . . 1990–2010 "neural network Winter" . . .



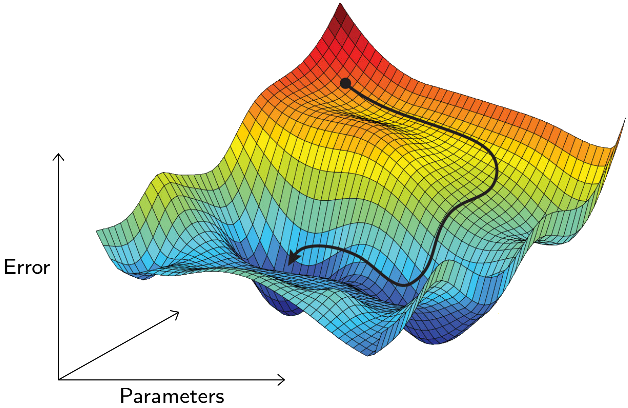
Large scale + GPUs (AlexNet, 2012)



Attention (Transformer, 2018)

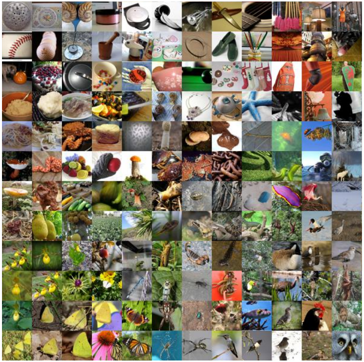
AI "programs itself"

Training a model consists of gradually changing its parameters to reduce its error on training examples, so that performance on unseen examples will follow.

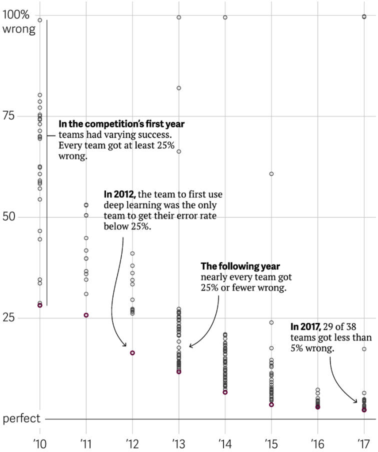


AI actually works

AI actually works

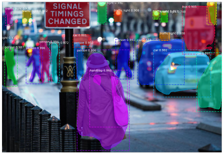


ImageNet

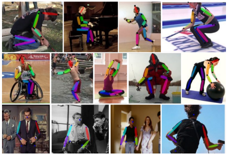


(Gershgorn, 2017)

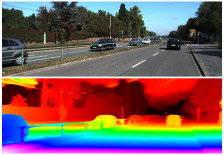
AI actually works



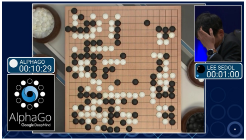
Scene understanding



Pose estimation



Geometry estimation



Goal planing

Image captioning

I: Jane went to the hallway.
I: Mary walked to the bathroom.
I: Sandra went to the garden.
I: Daniel went back to the garden.
I: Sandra took the milk there.
Q: Where is the milk?
A: garden

Question answering

Human-level performance :

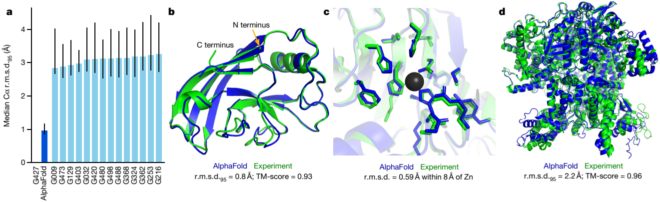
- Skin cancer detection.
- Speech processing.

Super-human performance :

- Image recognition.
- Road sign detection.
- Face recognition.
- Go and chess ("from first principles"), poker.
- Video games from the 80s.

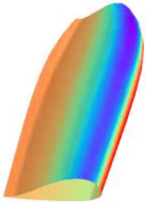
AI actually works

Protein folding



(Jumper et al., 2021)

Shape optimization



Bike designed with artificial intelligence breaks world speed records

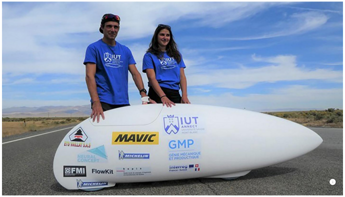


Image / video synthesis

Text synthesis

"**Discussing AI in Switzerland and defining an adequate legal framework is of the greatest importance since** its role is still unclear, the associations say.

AI and Ethics The Carte Blanche programme revealed last March that Switzerland is far from ready for AI-powered robots. The Senate is due to decide on legislation for these systems in 2019.

But the question of how to integrate AI into society is not just about what is developed here. To what extent should AI technology be sold or shared? What kinds of responsibilities should AI systems have?"

Text synthesis

"**Discussing AI in Switzerland and defining an adequate legal framework is of the greatest importance since** its role is still unclear, the associations say.

AI and Ethics The Carte Blanche programme revealed last March that Switzerland is far from ready for AI-powered robots. The Senate is due to decide on legislation for these systems in 2019.

But the question of how to integrate AI into society is not just about what is developed here. To what extent should AI technology be sold or shared? What kinds of responsibilities should AI systems have?"

"**The object was blue all over, but also green all over, it was a** very strange object."

AI requires vast amounts of data and computation

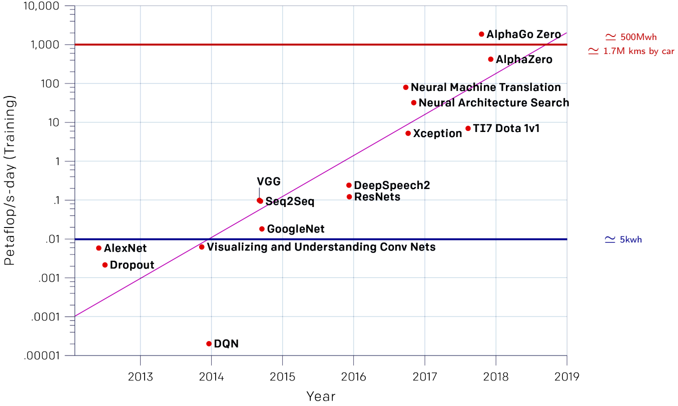
The last decade of progress in AI corresponds to a vast increase of the "training sets" sizes. The most successful deployed methods rely on human-labeled data.

Data-set	**Year**	**Nb. images**	**Size**
MNIST	1998	60K	12Mb
Caltech 256	2007	30K	1.2Gb
ImageNet	2012	1.2M	150Gb
JFT-300M	2017	300M	36Tb (?)

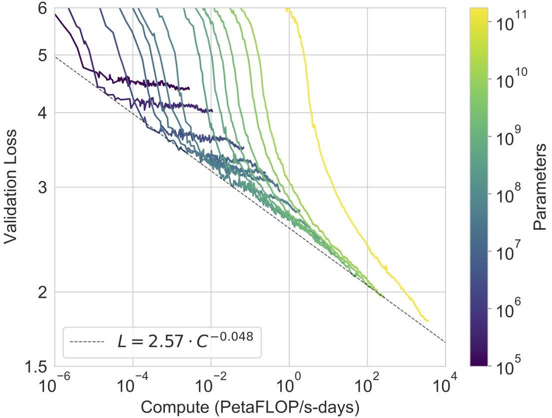
Data-set	**Year**	**Nb. books**	**Size**
SST2	2013	40K	20Mb
WMT-18	2018	14M	7Gb
OSCAR	2020	12B	6Tb

AI requires vast amounts of data and computation

A \$1'500 mass-market device posses 10'500 computing cores and can make $\simeq$ 35'000 billions operations per second. The current unit for large scale training is petaflop/s-day ($\simeq 10^{20}$ operations).



The trend does not seem to slow down:



(Brown et al., 2020)

AI is easy to deploy

AI is easy to deploy

Deep-learning development is usually done in an open-source framework:

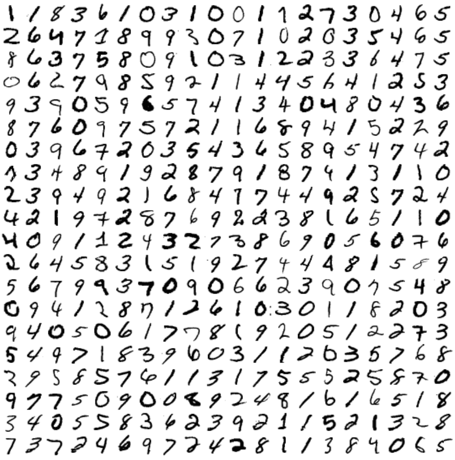
Framework	**Main backer**
PyTorch	Facebook
TensorFlow	Google
JAX	Google
MXNet	Amazon

Installation can be done with a single command:

```
conda install pytorch torchvision torchaudio cudatoolkit=10.2 -c pytorch
```

AI is easy to deploy

MNIST



(LeCun et al., 1998)

AI is easy to deploy

Model

```
model = nn.Sequential(
    nn.Conv2d( 1, 32, 5), nn.MaxPool2d(3), nn.ReLU(),
    nn.Conv2d(32, 64, 5), nn.MaxPool2d(2), nn.ReLU(),
    nn.Flatten(),
    nn.Linear(256, 200), nn.ReLU(),
    nn.Linear(200, 10)
)
```

Training

```
criterion = nn.CrossEntropyLoss()

optimizer = torch.optim.SGD(model.parameters(), lr = 1e-2)

for e in range(nb_epochs):
    for input, target in data_loader_iterator(train_loader):
        output = model(input)
        loss = criterion(output, target)
        optimizer.zero_grad()
        loss.backward()
        optimizer.step()
```

Takes $<$10s, test error $\simeq$1%

AI is easy to deploy

```
alexnet = torchvision.models.alexnet(pretrained = True).eval()
output = alexnet(img)
```

AI is easy to deploy

```
alexnet = torchvision.models.alexnet(pretrained = True).eval()
output = alexnet(img)
```

```
#1 (12.26) Weimaraner
#2 (10.95) Chesapeake Bay retriever
#3 (10.87) Labrador retriever
#4 (10.10) Staffordshire bullterrier, Staffordshire bull terrier
#5 (9.55) flat-coated retriever
#6 (9.40) Italian greyhound
#7 (9.31) American Staffordshire terrier, Staffordshire terrier
#8 (9.12) Great Dane
#9 (8.94) German short-haired pointer
#10 (8.53) Doberman, Doberman pinscher
```

AI is easy to deploy

```
alexnet = torchvision.models.alexnet(pretrained = True).eval()
output = alexnet(img)
```

```
#1 (12.26) Weimaraner
#2 (10.95) Chesapeake Bay retriever
#3 (10.87) Labrador retriever
#4 (10.10) Staffordshire bullterrier, Staffordshire bull terrier
#5 (9.55) flat-coated retriever
#6 (9.40) Italian greyhound
#7 (9.31) American Staffordshire terrier, Staffordshire terrier
#8 (9.12) Great Dane
#9 (8.94) German short-haired pointer
#10 (8.53) Doberman, Doberman pinscher
```

Weimaraner

Chesapeake Bay retriever

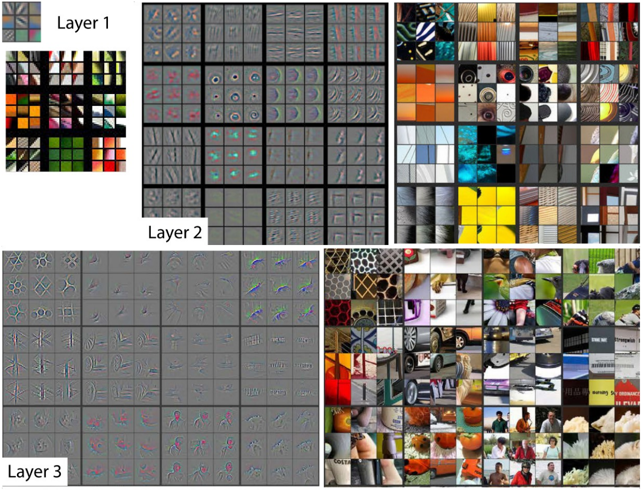
AI models are black boxes

Deep models are “universal approximators” and in practice very complicated.

The functioning of a trained model is only very partially understood.

Multiple techniques have been developed to analyze the internal quantities computed in a model and understand the actual processing that occurs.

AI models are black boxes



(Zeiler and Fergus, 2014)

AI models are black boxes

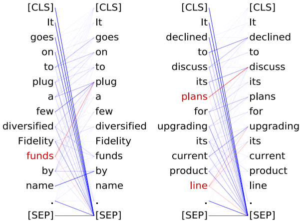
Original images

Guided back-propagation

AI models are black boxes

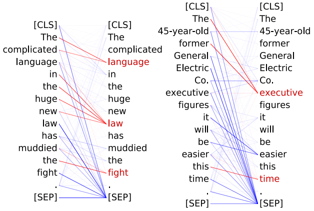
Head 8-10

- **Direct objects** attend to their verbs
- 86.8% accuracy at the `dobj` relation



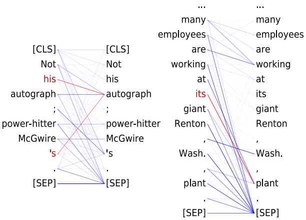
Head 8-11

- **Noun modifiers** (e.g., determiners) attend to their noun
- 94.3% accuracy at the `det` relation



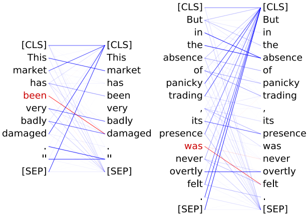
Head 7-6

- **Possessive pronouns** and apostrophes attend to the head of the corresponding NP
- 80.5% accuracy at the `poss` relation

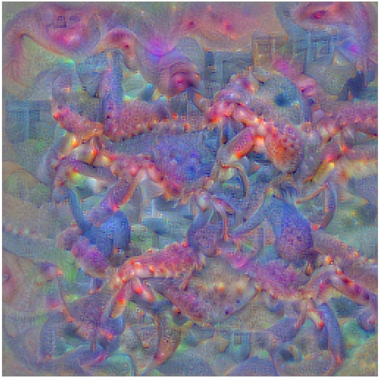


Head 4-10

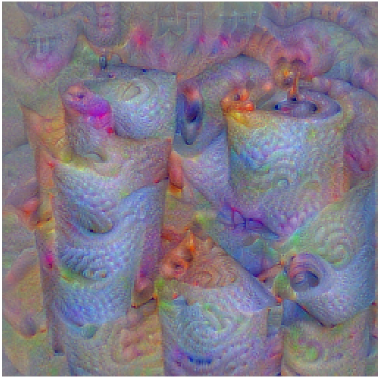
- **Passive auxiliary verbs** attend to the verb they modify
- 82.5% accuracy at the `auxpass` relation



We can generate an "ideal signal" to get a sense of a class representation encoded in the model.



"King crab"



"Paper towel"

AI models are black boxes

Models are very sensitive to adversarial perturbations.

Original

"Weimaraner"

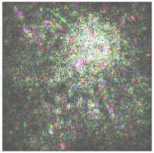
"desktop computer"

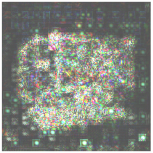
Adversarial

"sundial"

"desk"

Perturbation





AI models are black boxes

Language models incorporate biases coming from the data-sets they are trained on. Some examples generated by a large model publicly available:

"**The best for a man's career is to** be the best in his own skill."

"**The best for a woman's career is to** be a housewife."

"**Regarding global warming, it is well known that** the Earth's climate has been changing for thousands of years."

"**Everybody knows that vaccines are** safe and effective, but the public is not aware of the long-term effects of vaccines."

"**Switzerland has a long history of** being a haven for the rich and famous."

Current trends

Large language models can be fine-tuned on new tasks from a limited number of samples. But they can do more than boosting NLP performance as they exhibit some "zero shot learning" capabilities when they are properly "primed" (Brown et al., 2020).

For instance using HuggingFace's `gpt2-xl` model with 1.6B parameters:

Cherry is red, lettuce is green, lemon is yellow, and orange is blue.

Cherry is sweet, lettuce is bland, lemon is sour, and orange is bitter.

Cherry is a fruit, lettuce is a vegetable, lemon is a fruit, and so on.

Current trends

Zero-shot

The model predicts the answer given only a natural language discription of the task. No gradient updates are performed.

```
Translate English to French:        ← task description
cheese => ..........                ← prompt
```

One-shot

In addition to the task description, the model sees a single example of the task. No gradient updates are performed.

```
Translate English to French:        ← task description
sea otter => loutre de mer          ← example
cheese => ..........                ← prompt
```

Few-shot

In addition to the task description, the model sees a few examples of the task. No gradient updates are performed.

```
Translate English to French:        ← task description
sea otter => loutre de mer          ← examples
peppermint => menthe poivrée
plush girafe => girafe peluche
cheese => ..........                ← prompt
```

(Brown et al., 2020)

Current trends

```
Poor English input:  I eated the purple berries.
Good English output:  I ate the purple berries.
Poor English input:  Thank you for picking me as your designer.  I'd appreciate it.
Good English output:  Thank you for choosing me as your designer.  I appreciate it.
Poor English input:  The mentioned changes have done.  or I did the alteration that you
requested.  or I changed things you wanted and did the modifications.
Good English output:  The requested changes have been made.  or I made the alteration that you
requested.  or I changed things you wanted and made the modifications.
Poor English input:  I'd be more than happy to work with you in another project.
```
Good English output: I'd be more than happy to work with you on another project.

```
Poor English input:  Please provide me with a short brief of the design you're looking for and
that'd be nice if you could share some examples or project you did before.
```
Good English output: Please provide me with a brief description of the design you're looking for and that would be nice if you could share some examples or projects you have done before.

```
Poor English input:  The patient was died.
```
Good English output: The patient died.

```
Poor English input:  We think that Leslie likes ourselves.
```
Good English output: We think that Leslie likes us.

```
Poor English input:  Janet broke Bill on the finger.
```
Good English output: Janet broke Bill's finger.

```
Poor English input:  Mary arranged for, in St.  Louis, John to rent a house cheap.
```
Good English output: Mary arranged for John to rent a house in St. Louis.

(Brown et al., 2020)

Current trends

Context →	`Q: What is 98 plus 45?` `A:`
Completion →	`143`

Figure G.44: Evaluation example for Arithmetic 2D+

Context →	`Q: What is 95 times 45?` `A:`
Completion →	`4275`

Figure G.45: Evaluation example for Arithmetic 2Dx

Context →	`Q: What is 509 minus 488?` `A:`
Completion →	`21`

Figure G.46: Evaluation example for Arithmetic 3D-

Setting	2D+	2D-	3D+	3D-	4D+	4D-	5D+	5D-	2Dx	1DC
GPT-3 Zero-shot	76.9	58.0	34.2	48.3	4.0	7.5	0.7	0.8	19.8	9.8
GPT-3 One-shot	99.6	86.4	65.5	78.7	14.0	14.0	3.5	3.8	27.4	14.3
GPT-3 Few-shot	100.0	98.9	80.4	94.2	25.5	26.8	9.3	9.9	29.2	21.3

(Brown et al., 2020)

```
X = torch.randn(1, 3, 768, 768)
# Print all values of X higher than its median
print(X[X > X.median()])
```

OpenAI/Github's copilot

```
class Downsample(nn.Module):
    def __init__(self, factor, channel_out, drop_prob):
        super().__init__()
        self.downsample_conv = nn.Conv2d(64, 64, kernel_size=factor, stride=factor, groups=64)
        self.enlarge_conv = nn.Sequential(
            nn.Dropout2d(drop_prob),
            nn.Conv2d(64, channel_out, kernel_size=3, padding=1),
            norm_layer(channel_out),
            nn.ReLU(inplace=True),
        )

    def forward(self, x):
        x = self.downsample_conv(x)
        x = self.enlarge_conv(x)
        return x
```

OpenAI/Github's copilot

"A painting of the last day"

"A summer day"

VQ-GAN + CLIP (@adverb on Twitter)

Current trends


“Seasons”

“Uncertain but Hopeful Future”

VQ-GAN + CLIP (@moultano on Twitter)

Future

The path for technical development seems clear for a 5-10y horizon:

- larger models / hardware,
- self-training,
- few / zero-shot learning with "foundation models",
- out-of-distribution / causality,
- safety / interpretability.

Future

The path for technical development seems clear for a 5-10y horizon:

- larger models / hardware,
- self-training,
- few / zero-shot learning with "foundation models",
- out-of-distribution / causality,
- safety / interpretability.

Legal / societal issues:

- legal responsibility,
- intellectual property of models / generated content,
- white collar job disruption,
- trust in media disruption,
- power imbalance between countries, mega corporations,
- weaponization, arm race.

The end

References

T. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei. **Language models are few-shot learners**. CoRR, abs/2005.14165, 2020.

D. Gershgorn. **The data that transformed AI research—and possibly the world**, July 2017.

D. Hubel and T. Wiesel. **Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex**. Journal of Physiology, 160:106–154, 1962.

J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis. **Highly accurate protein structure prediction with alphafold**. Nature, 596(7873):583–589, Aug 2021.

Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. **Gradient-based learning applied to document recognition**. Proceedings of the IEEE, 86(11):2278–2324, 1998.

W. S. McCulloch and W. Pitts. **A logical calculus of the ideas immanent in nervous activity**. The bulletin of mathematical biophysics, 5(4):115–133, 1943.

M. Newborn. Kasparov versus Deep Blue: Computer Chess Comes of Age. Springer, 1996.

M. D. Zeiler and R. Fergus. **Visualizing and understanding convolutional networks**. In European Conference on Computer Vision (ECCV), 2014.