

A I 解析入門

渡辺英治（TSBセンター・AI解析室）

2023年9月14日15日

講師自己紹介

渡辺英治（わたなべえいじ）

専門：視覚、神経科学

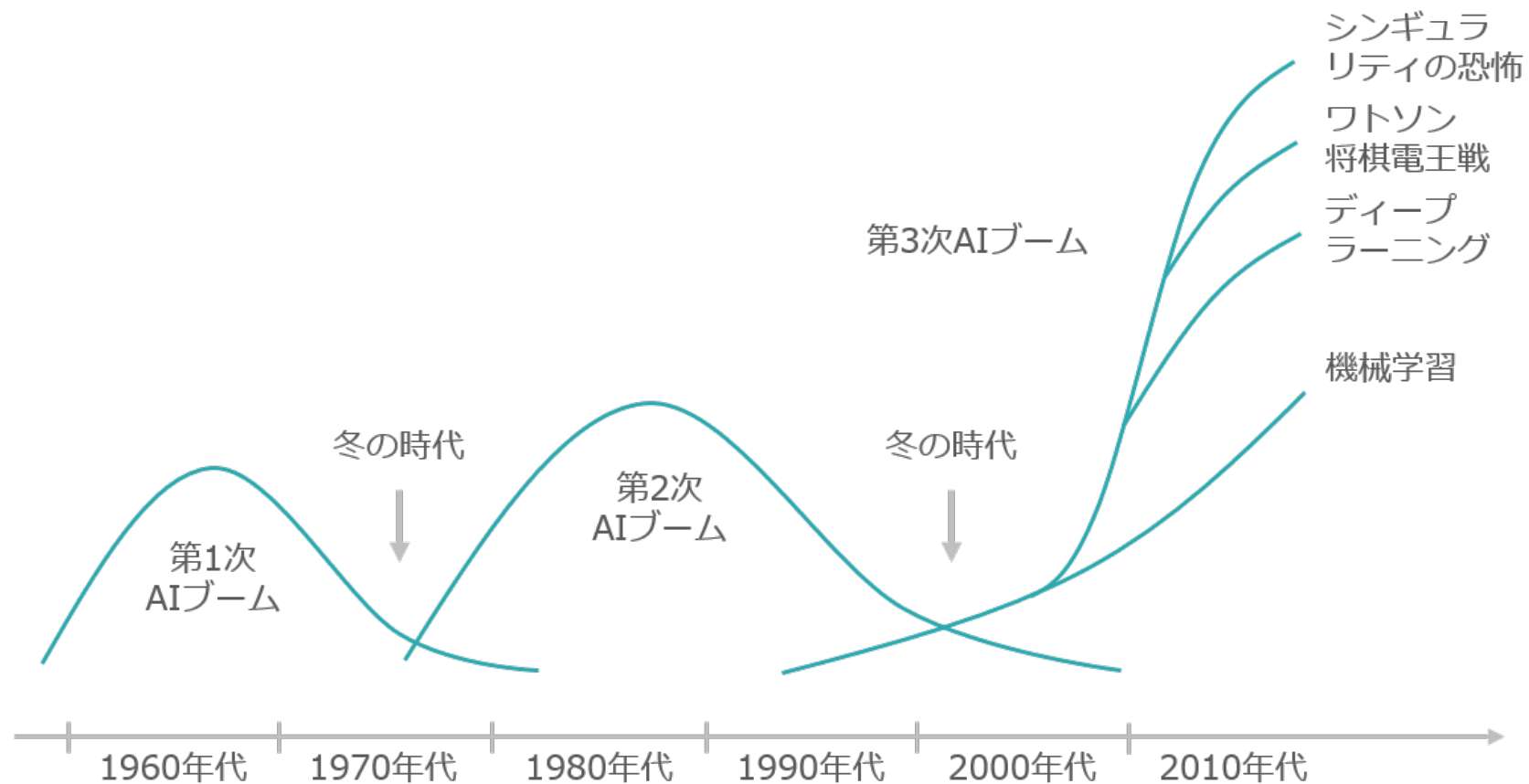
視覚をAIによってモデル化して理解する
（構成論的アプローチ）



講義の概要

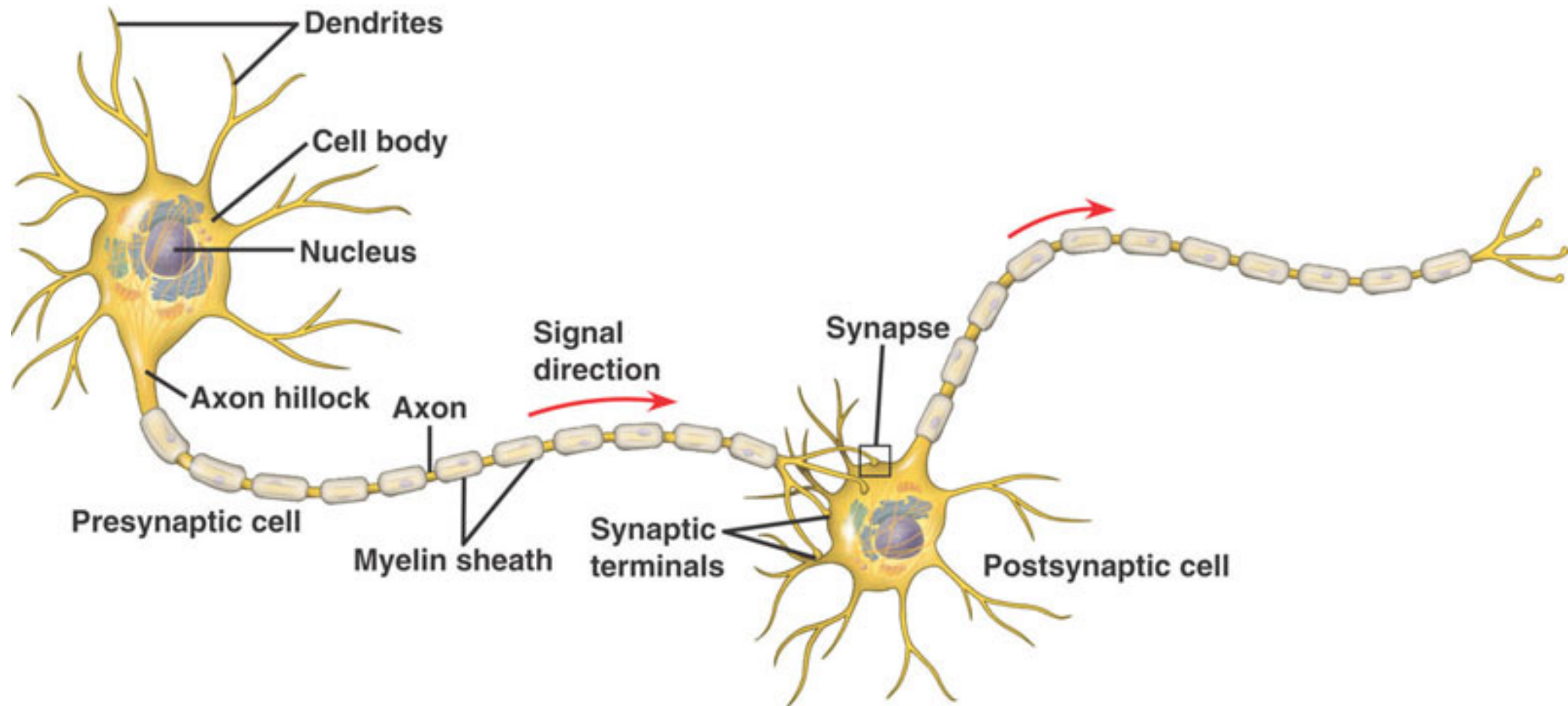
1. A I の歴史
2. 深層学習の要素解説
3. A I の種類概説
4. ChatGPTデモ
5. シングルセル解析デモ
6. A I の動作環境概説
7. A I の勉強方法
8. A I の未来

AIの歴史



松尾豊 『人工知能は人間を超えるのか ディープラーニングの先にあるもの』から引用

A I の歴史



neurons

A I の歴史

The frontiers who proposed the first artificial neurons

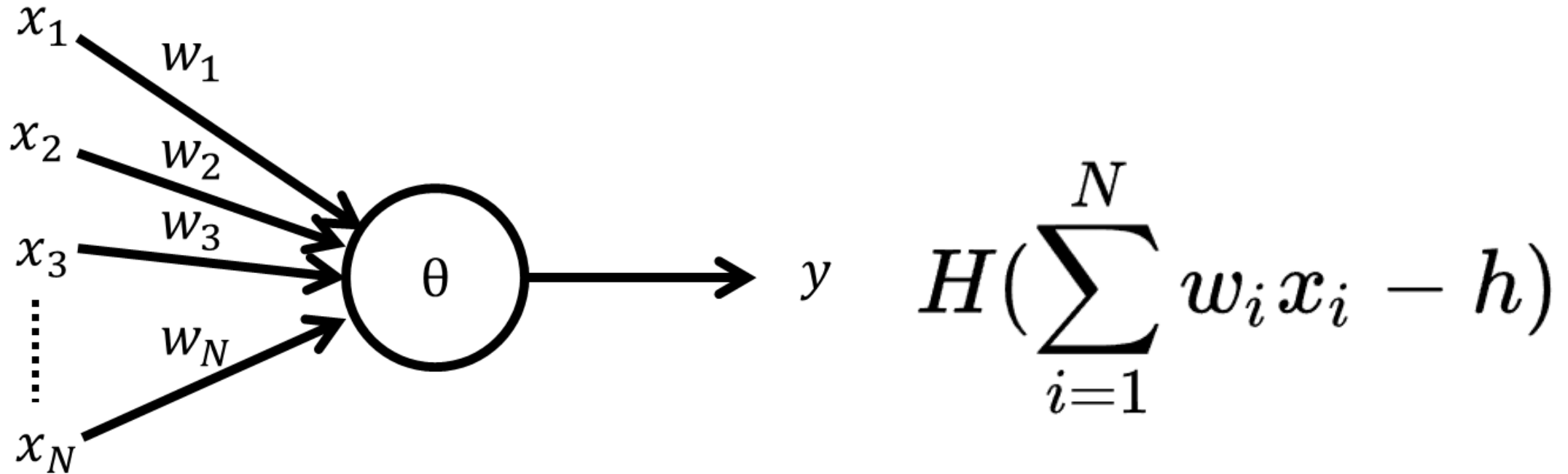


Warren S. McCulloch
Neurophysiologist

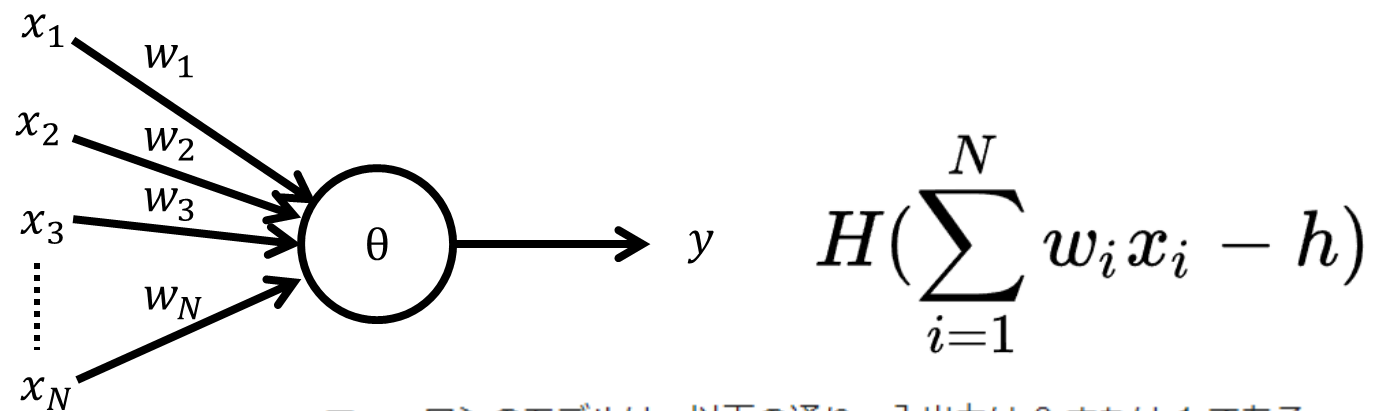
Walter Pitts
Logician

A I の歴史

**“Formal Neuron” is a simplified model of neuron.
It can be represented by a mathematical formula.**



McCulloch, W. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics, 7:115 - 133.



ニューロンのモデルは、以下の通り。入出力は 0 または 1 である。

- w : 重みづけ (実数)
- x : 入力信号 (0 または 1)
- h : しきい値 (実数)
- H : **ヘヴィサイドの階段関数** (出力は 0 または 1)

$$H\left(\sum_{i=1}^N w_i x_i - h\right)$$

実例としては、以下の通り。XOR は3層、他は2層である。

AND

$$H(x_1 + x_2 - 1.5)$$

OR

$$H(x_1 + x_2 - 0.5)$$

NOT

$$H(-x_1 + 0.5)$$

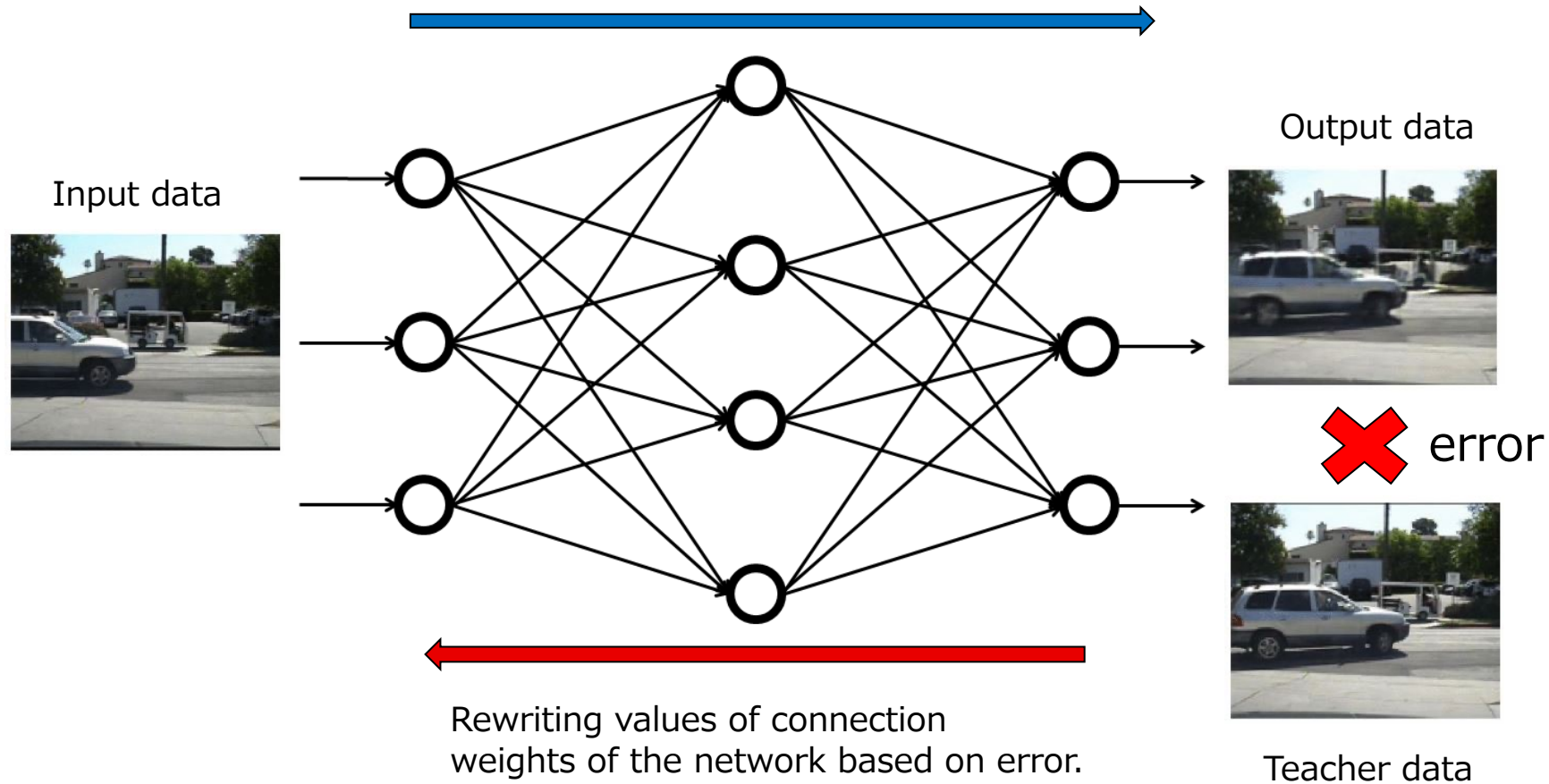
XOR

$$H(x_1 + x_2 - 2H(x_1 + x_2 - 1.5) - 0.5)$$

A I の歴史

Deep neural networks

The networks learn connection weights from teacher data.



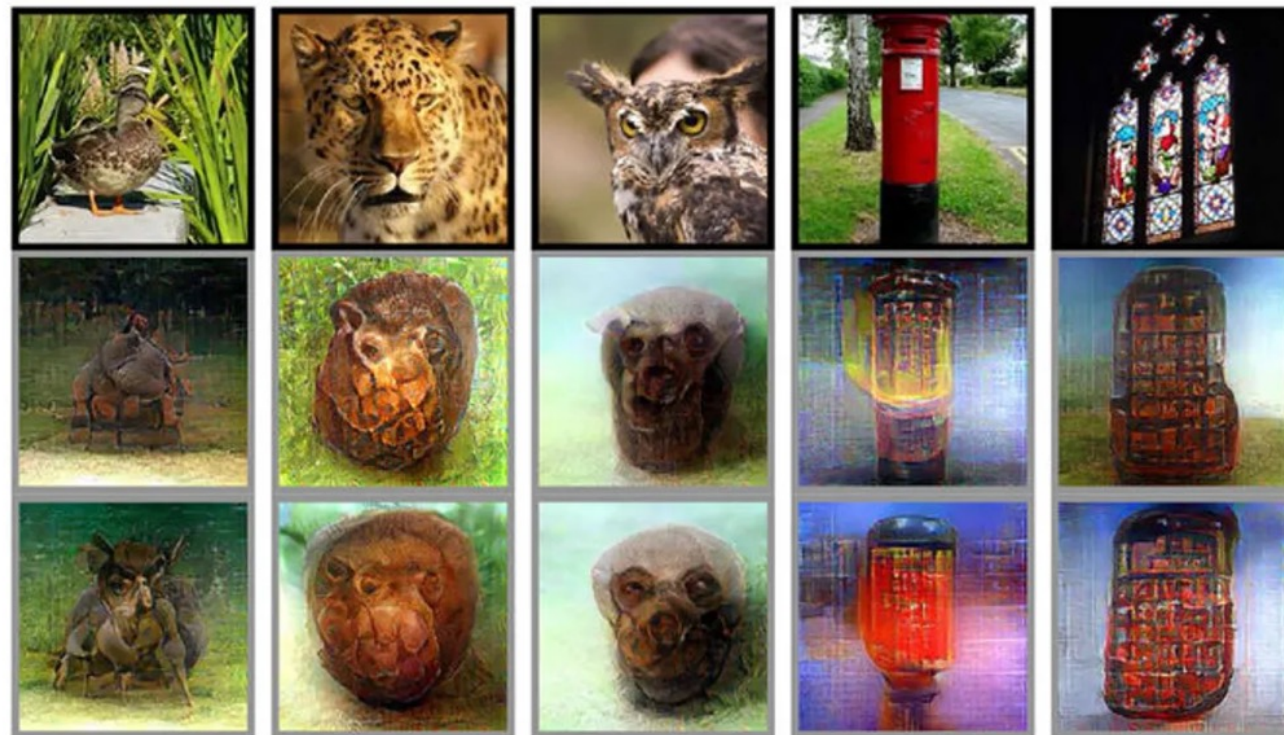
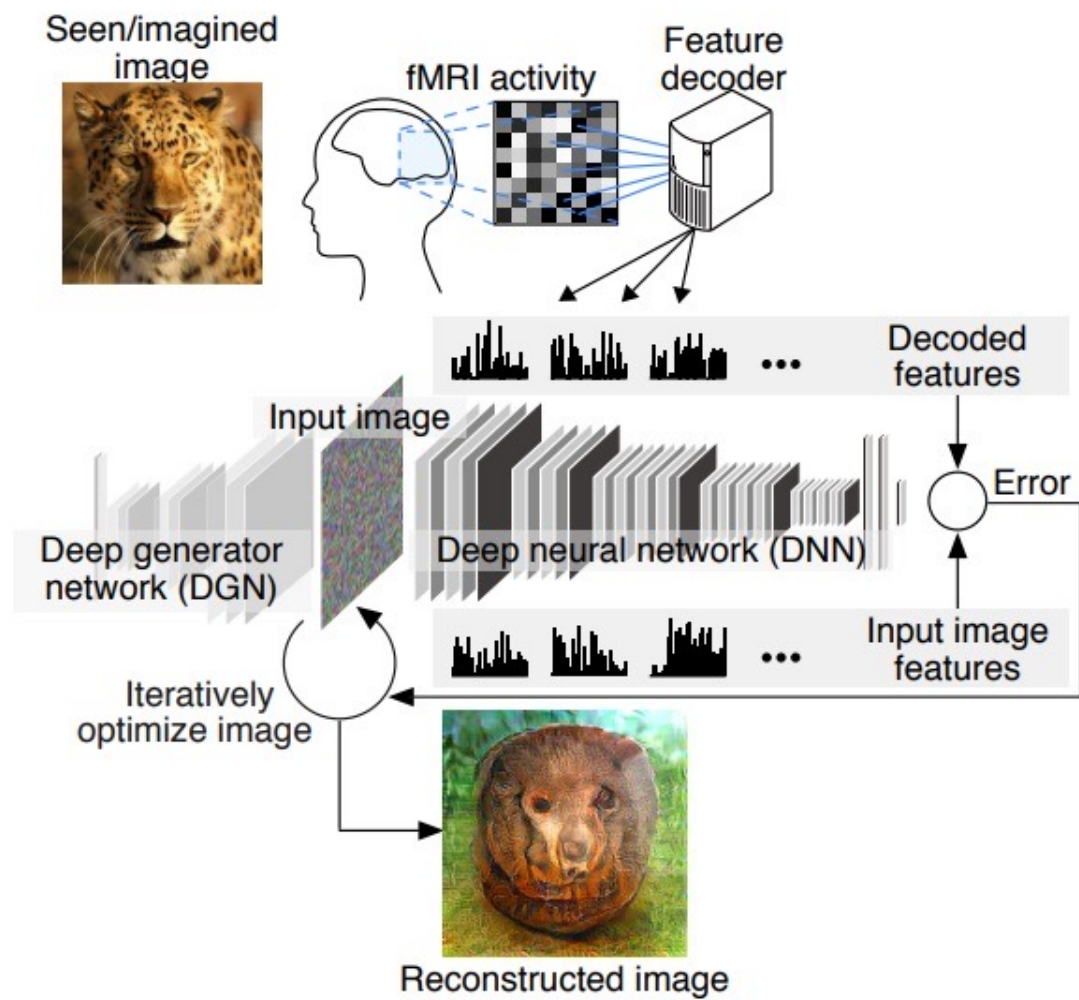
Backpropagation: Rumelhart, Hinton & Williams (1986)

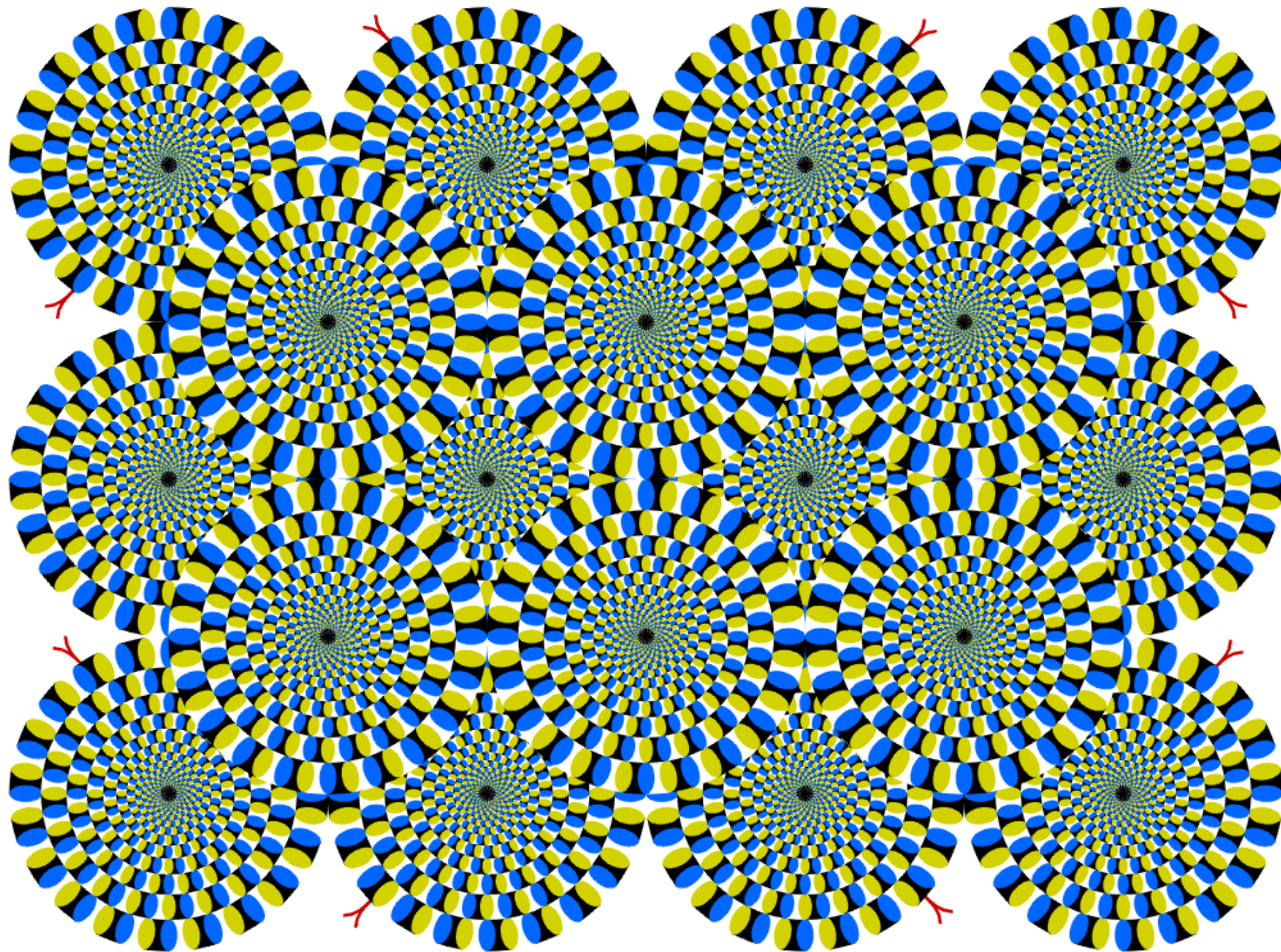
A I の歴史

2012年の物体の認識率を競うILSVRCにおける、GPU利用による大規模ディープラーニング(ジェフリー・ヒントンの研究チームがAlex-netで出場した)の大幅な躍進、同年のGoogleによるディープラーニングを用いたYouTube画像からの猫の認識成功の発表により、世界各国において再び人工知能研究に注目が集まり始めた。この社会現象は第3次人工知能ブームと呼ばれる。その後、ディープラーニングの研究の加速と急速な普及を受けて、レイ・カーツワイルが2005年に提唱していた技術的特異点という概念は、急速に世界中の識者の注目を集め始めた。

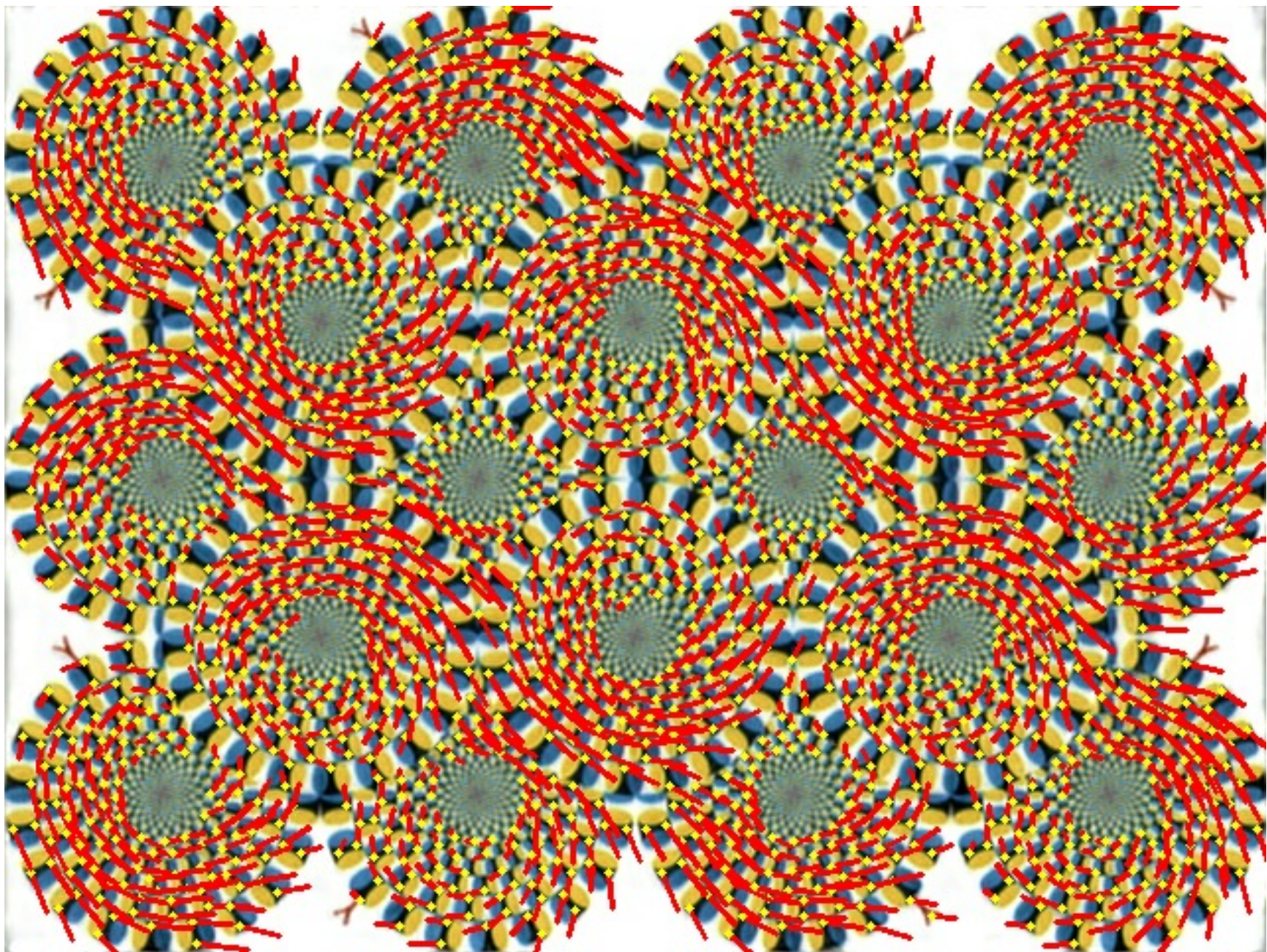


A I の歴史





Rotating Snake Illusion by A.Kitaoka



Watanabe Lab

深層学習の要素解説

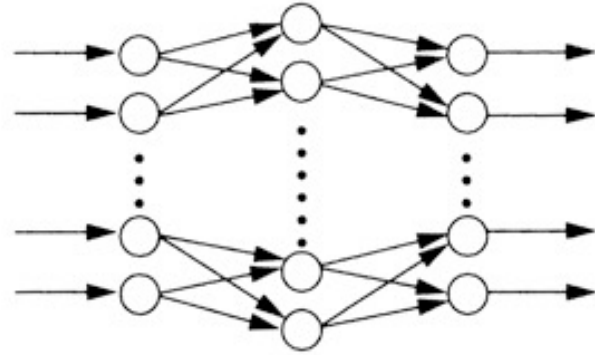
1. 学習ルール
2. 目的関数
3. 構造
4. 環境（学習データ）

A I の種類概要

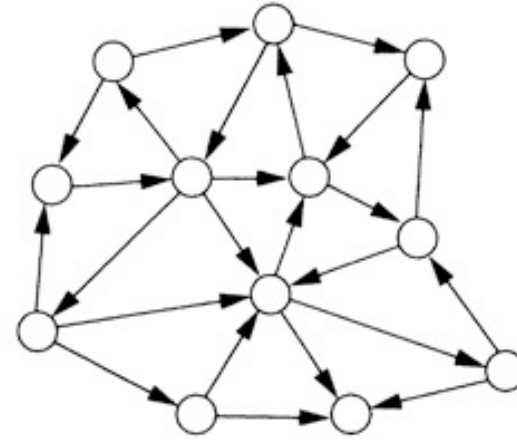
宿題：

深層学習で使用するニューラルネットワークには様々な種類がありますが、このうち畳み込みニューラルネットワークとトランスフォーマーについて自分なりに理解をしてみてください。畳み込みニューラルネットワークだけでも大丈夫です。

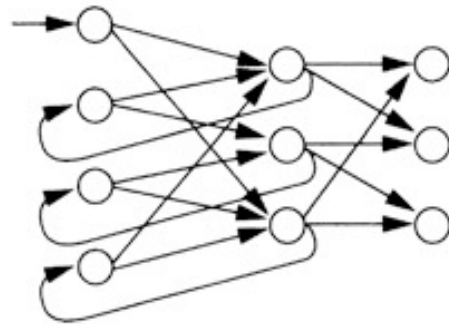
A I の種類概要



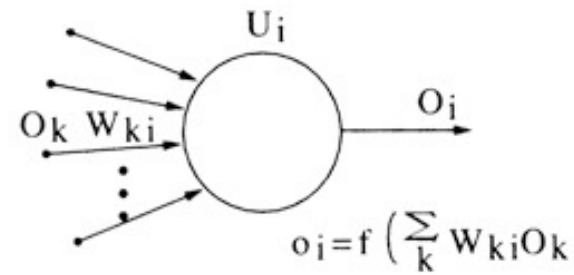
階層型ニューラルネットワーク



相互結合型ニューラルネットワーク



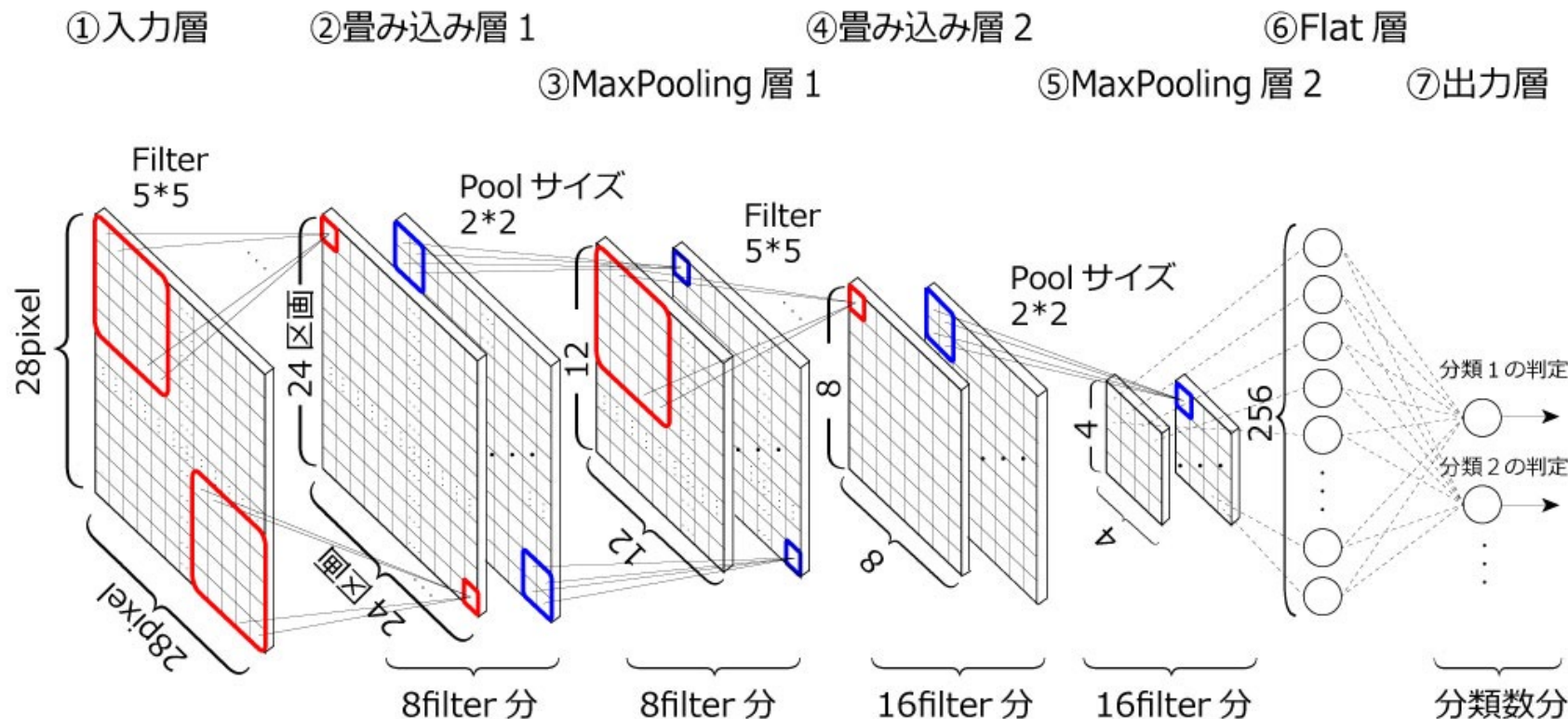
リカレントニューラルネットワーク



ニューロンモデル

A I の種類概要

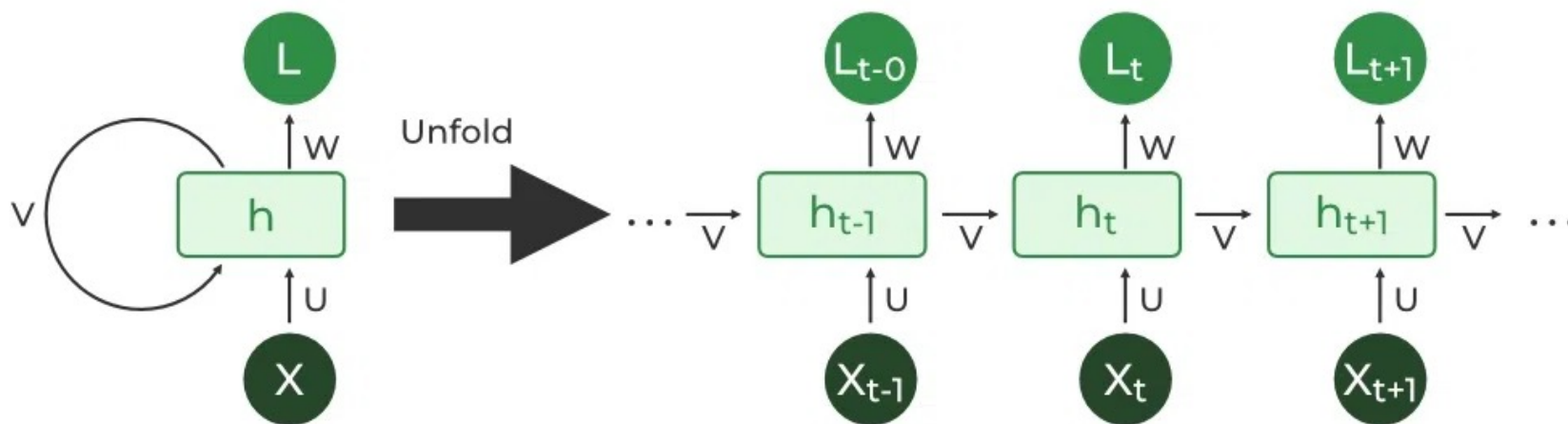
CNN



	①入力	②畳み込み	③MaxPooling	④畳み込み	⑤MaxPooling	⑥Flat	⑦出力
出力サイズ	28*28*1	24*24*8	12*12*8	8*8*16	4*4*16	256 (4*4*16)	分類数
調整：重み	-	200 (5*5*8)	-	3200 (8*(5*5*16))	-	-	256 * 分類数
調整：閾値	-	8	-	16	-	-	分類数

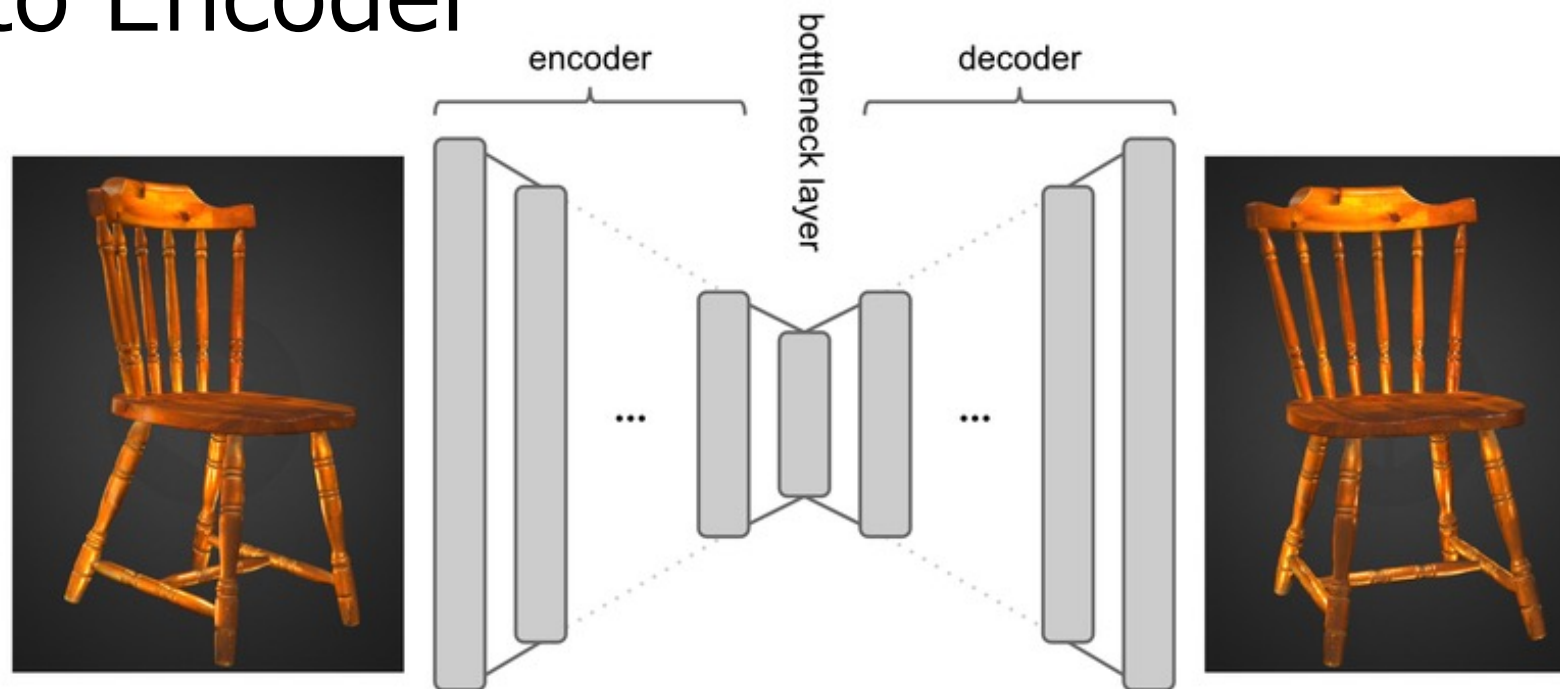
A I の種類概要

RNN

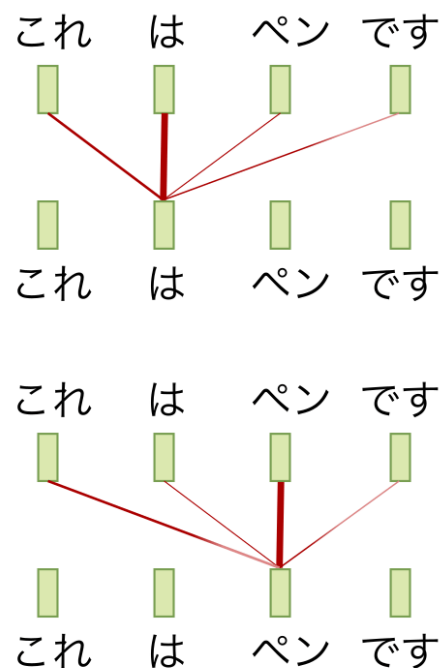


A I の種類概要

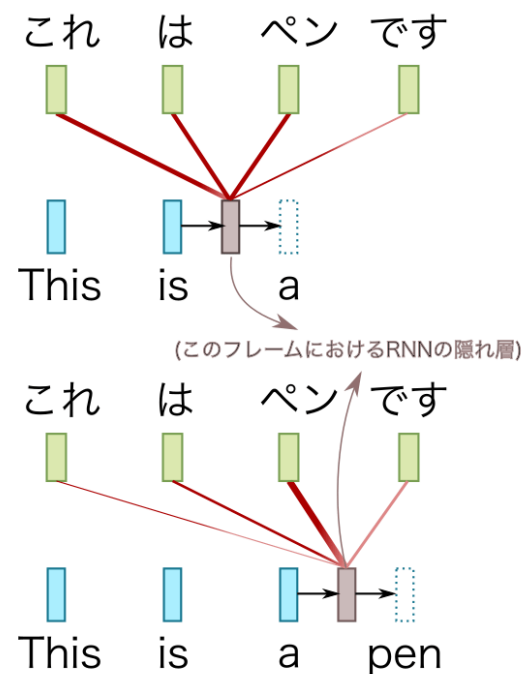
Auto Encoder



A I の種類概要



(a) 自己アテンション (系列内) :
自分以外の全単語との関連度を学習
(周囲中心に系列内のコンテキスト)



(b) 相互アテンション (系列対系列間) :
seq2seqのsoft attention. 次フレーム予測時の,
Decoder全単語との関連度を学習.

トランスフォーマー(Self-Attention)

ChatGPTについて（デモ）

詳細は次の講義にて

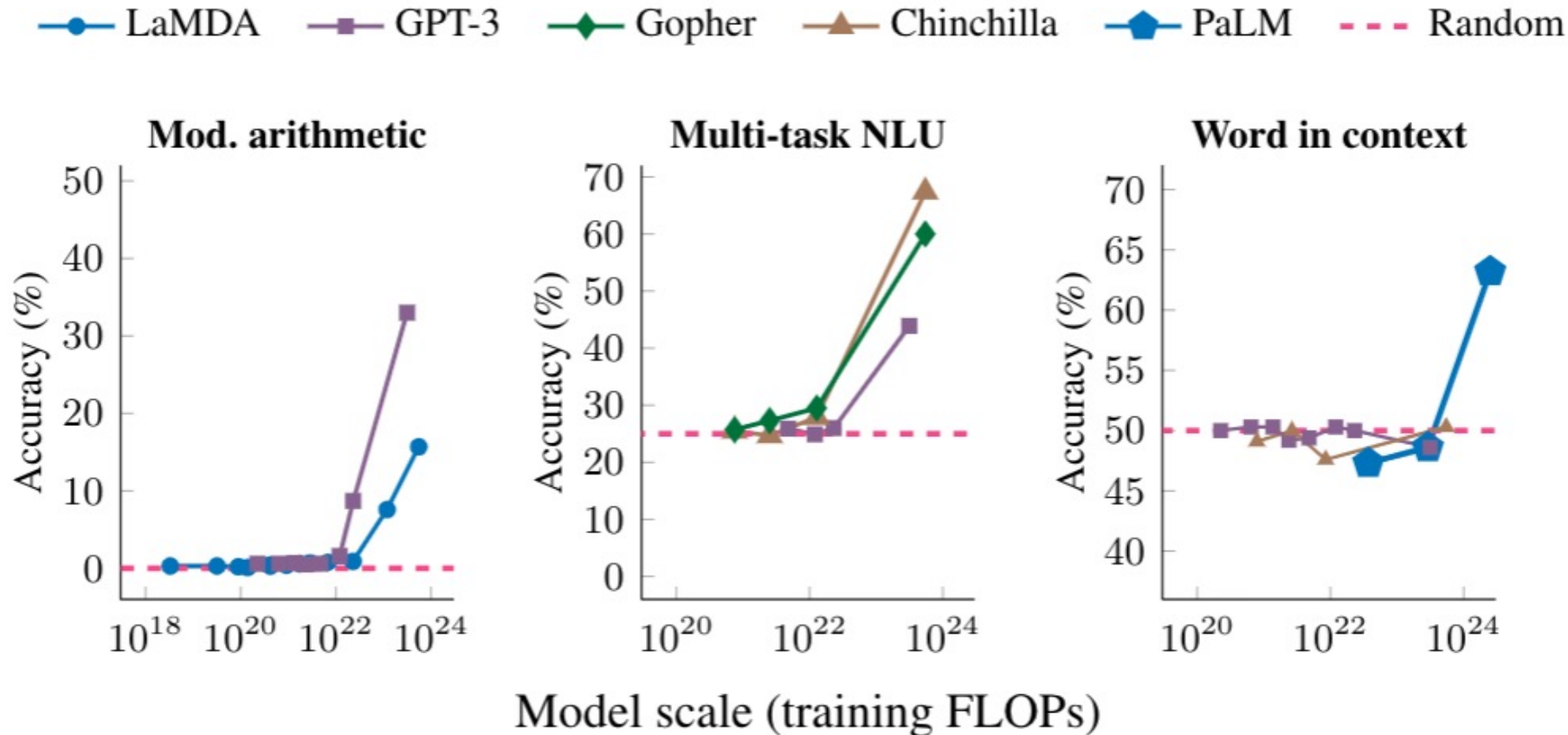
ChatGPT（GPT-3.5 & GPT-4）

非常に膨大なテキストデータから自然な言語の並び方を学習した A I である。

その成り立ちから、最も精度が高いのは短めのコーディング、コードレビューやバグフィクス、スタイルを変えた翻訳、要約、質疑応答が得意な分野。APIも公開されているため、応用がかなり進んでいる。

プログラムを作ろう！

ChatGPTについて

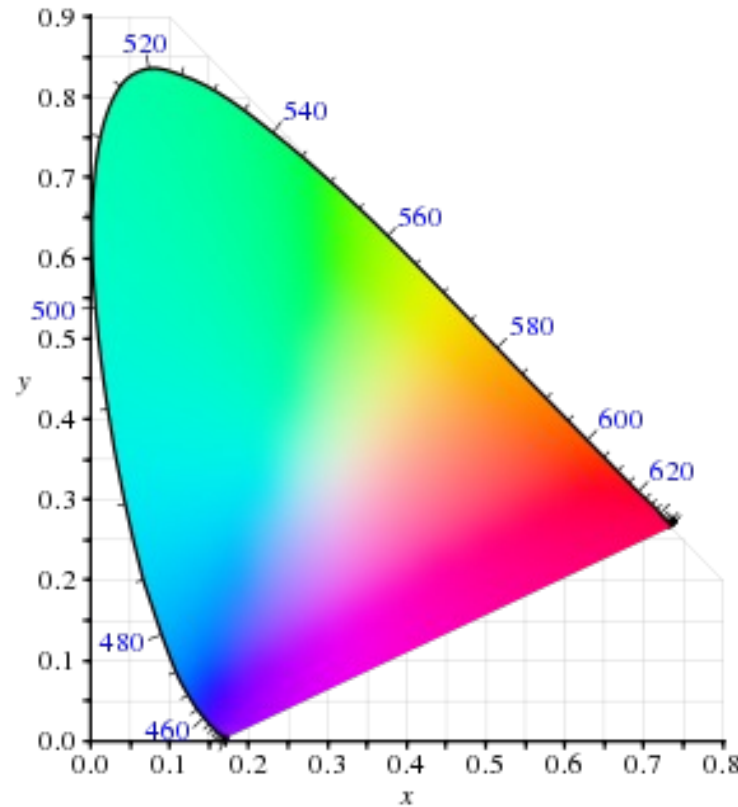


The ability to perform multi-step arithmetic (left), succeed on college-level exams (middle), and identify the intended meaning of a word in context (right) all emerge only for models of sufficiently large scale. The models shown include LaMDA, GPT-3, Gopher, Chinchilla, and PaLM. FLOPs = number of floating-point operations.

複数ステップの計算、大学レベルの試験、文脈の言葉の意味を読み取る精度

<https://ai.googleblog.com/2022/11/characterizing-emergent-phenomena-in.html>

ChatGPTについて

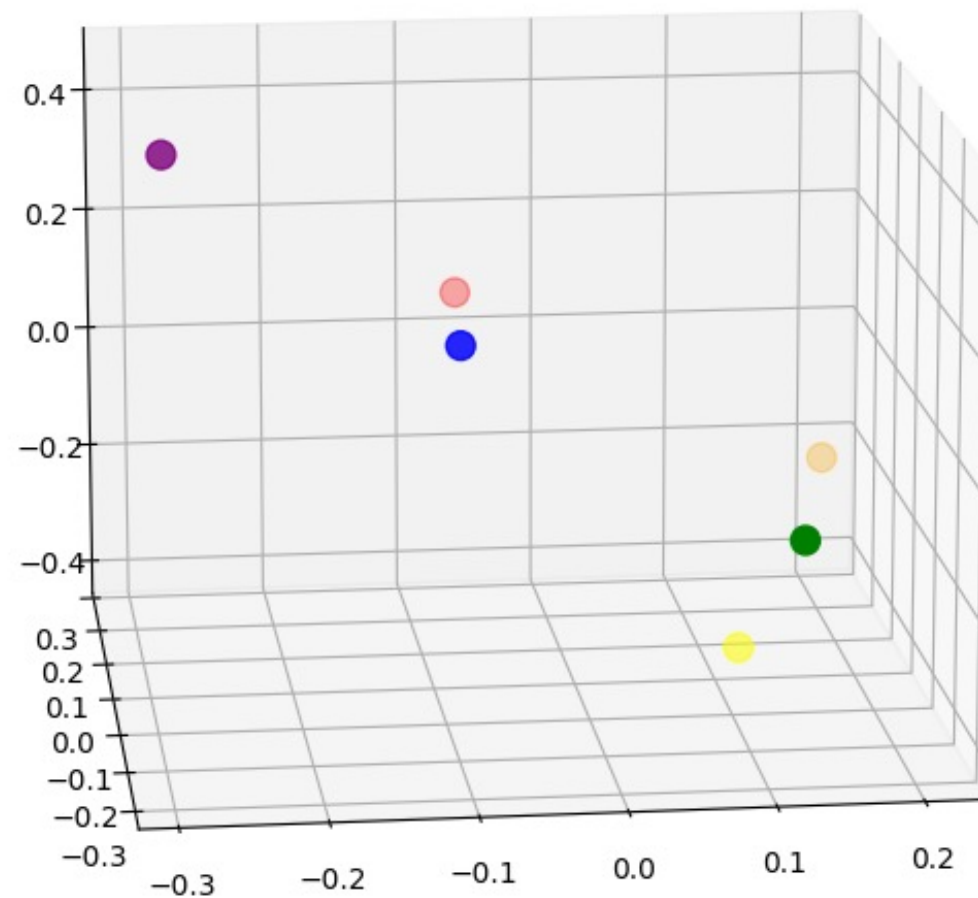
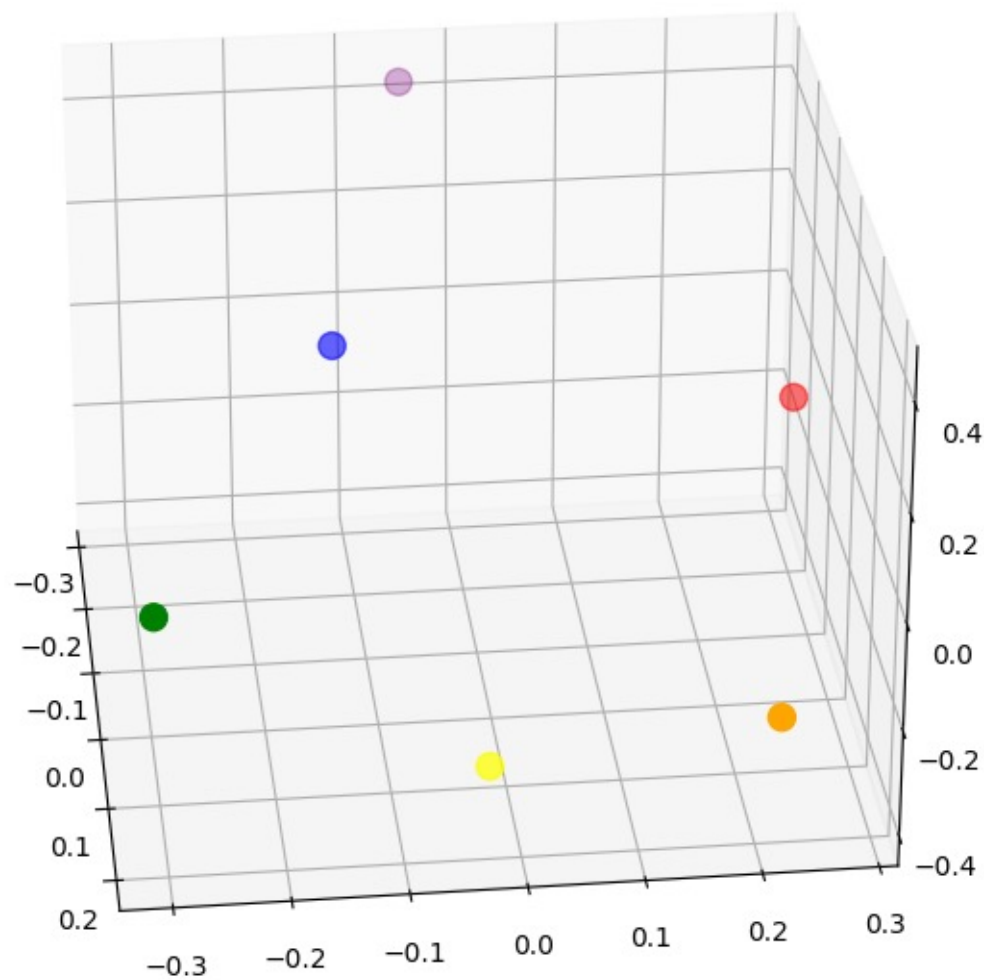


CIE 1931 color space

Smith, Thomas; Guild, John (1931–32). "The C.I.E. colorimetric standards and their use". Transactions of the Optical Society 33 (3): 73–134. doi:10.1088/1475-4878/33/3/301.

ChatGPTについて

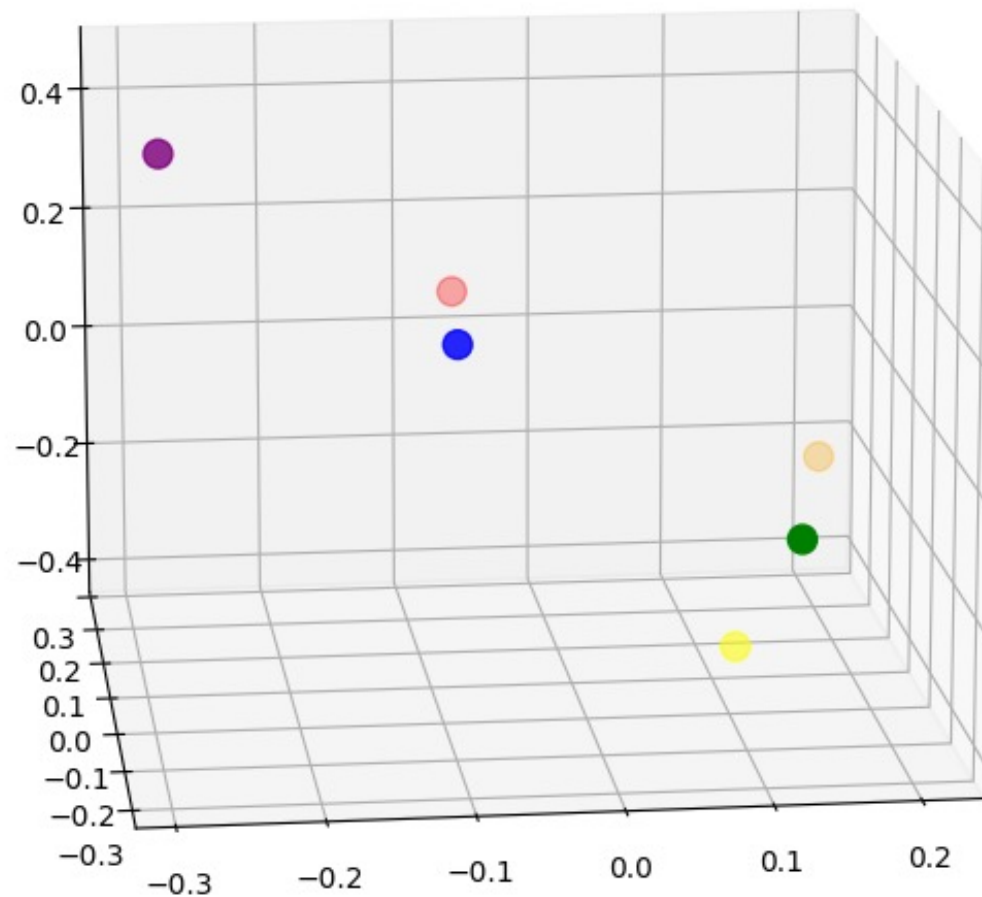
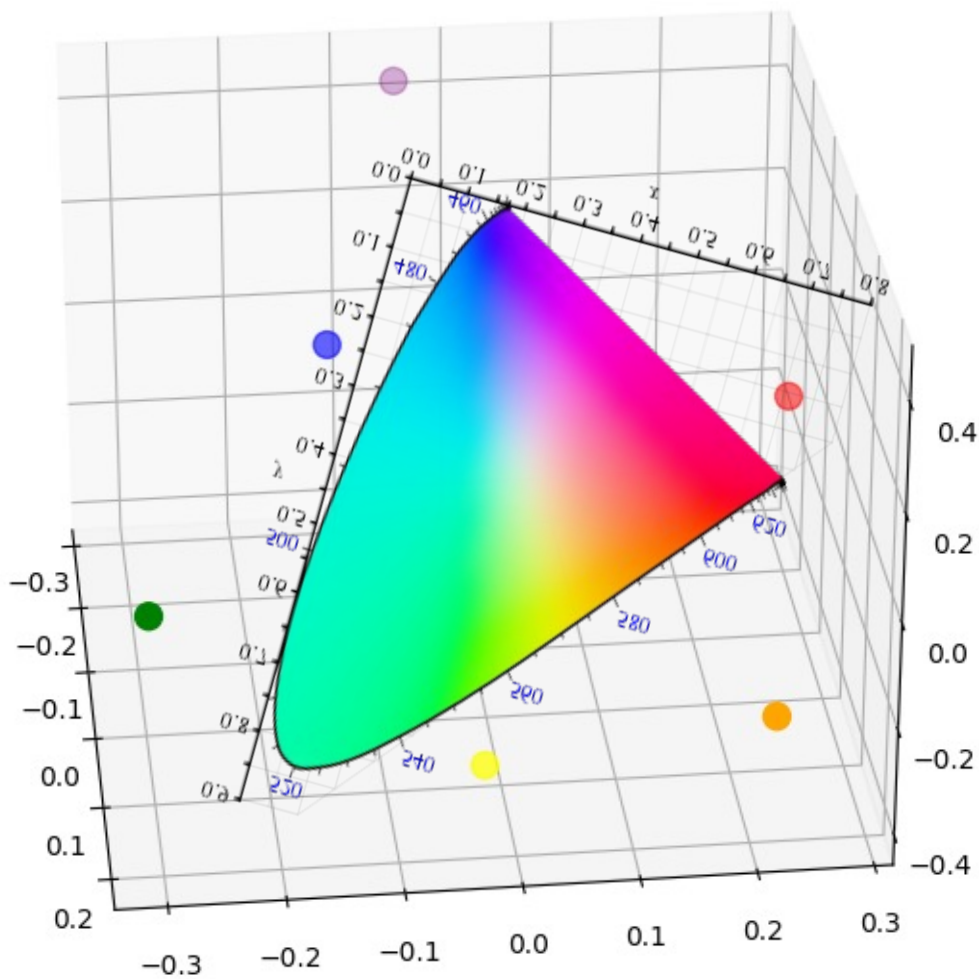
GPT3 Color Space (Multi Dimensional Scaling)



A sample question for GPT3: Red is more like blue or purple?

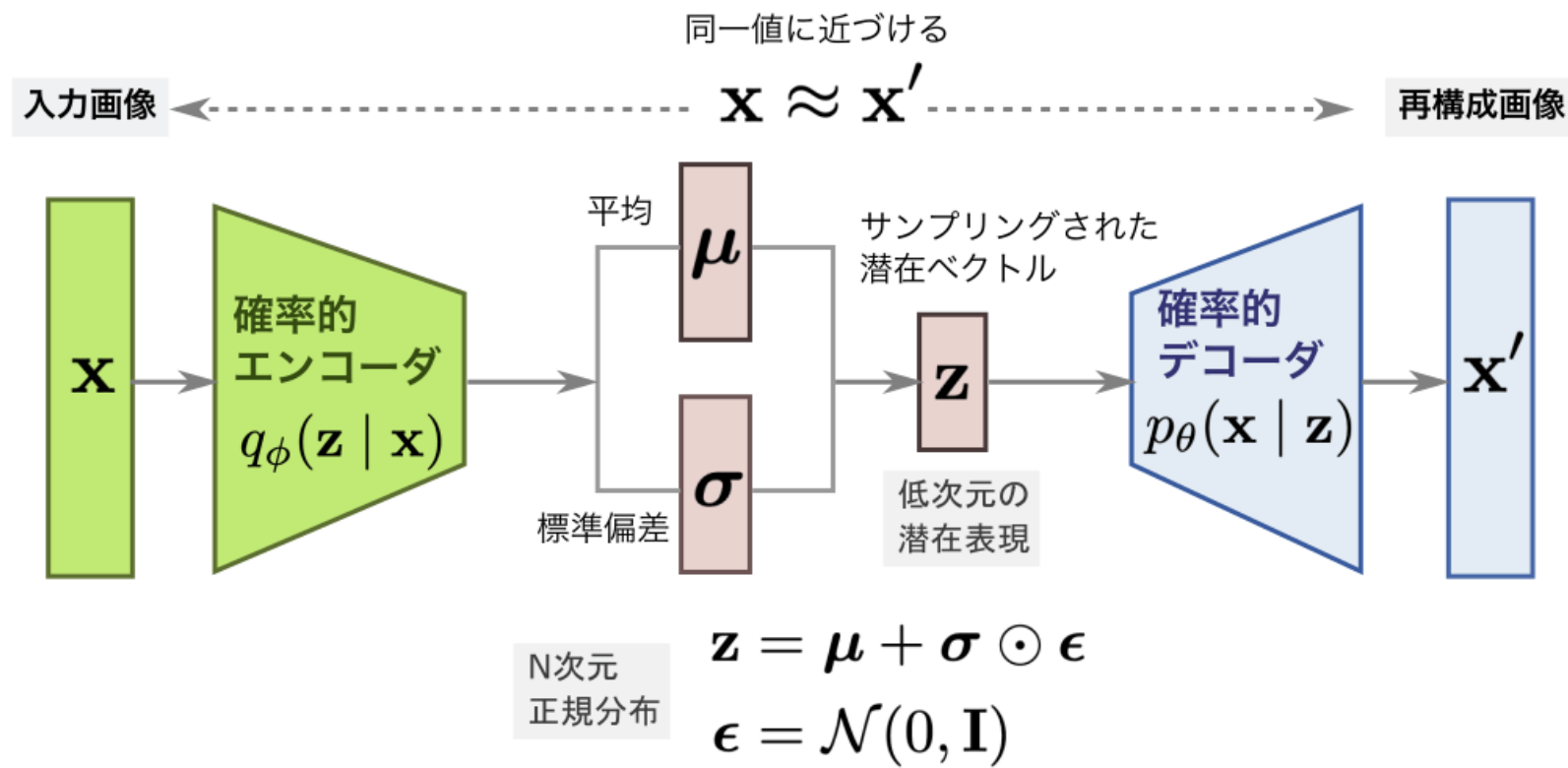
ChatGPTについて

GPT3 Color Space (Multi Dimensional Scaling)



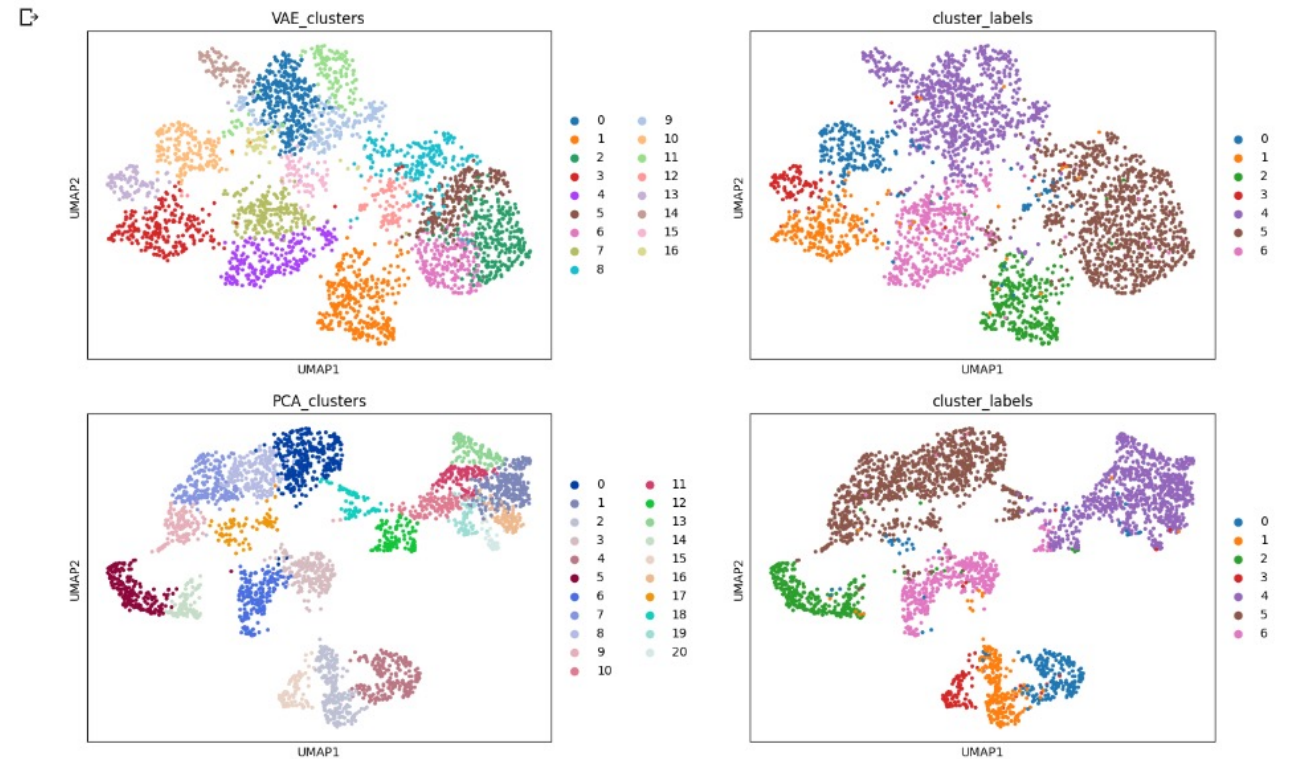
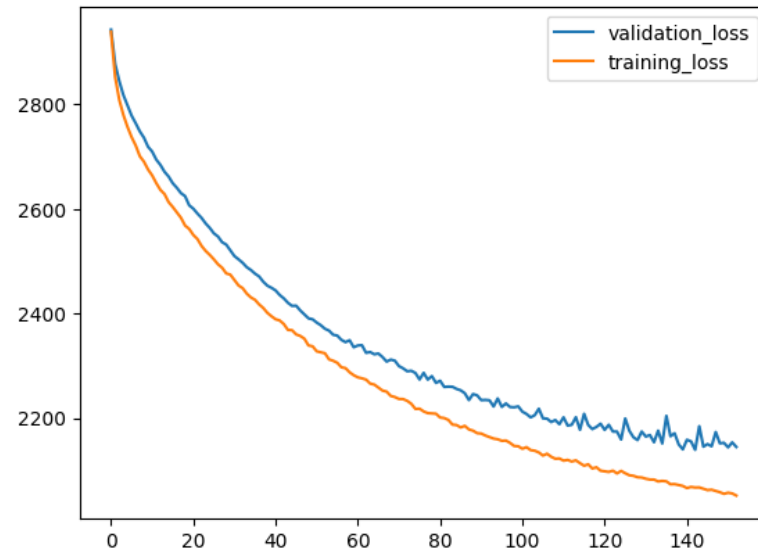
A sample question for GPT3: Red is more like blue or purple?

シングルセル解析とVAE (Google Colabでのデモ)



シングルセル解析とVAE (Google Colabでのデモ)

```
Start!  
validation loss at epoch 0 is 2942.68  
validation loss at epoch 20 is 2600.37  
validation loss at epoch 40 is 2444.00  
validation loss at epoch 60 is 2338.73  
validation loss at epoch 80 is 2271.24  
validation loss at epoch 100 is 2212.16  
validation loss at epoch 120 is 2177.37  
validation loss at epoch 140 is 2158.46  
Early Stopping at 152 epoch  
Done! validation_loss:2139.59
```



ARI score : clusters_label vs PCA_clusters 0.344
ARI score : clusters_label vs VAE_clusters 0.387

A I の動作環境概説

Google Colab

<https://colab.research.google.com/>

自然科学研究機構・計算科学研究センター

<https://ccportal.ims.ac.jp/>

ローカルサーバー

<https://doraemonkokoro.blogspot.com/2021/>

A I の勉強の仕方

A I の業界は、arXiv & github & SNSを活用した徹底したオープン作戦によって超高速での発展を可能にした。

変化が非常に激しいので、これまでの生物分野における情報の取り方では追いつかない。

学会 & 論文 & 教科書からarXiv & github & SNSへ

A I の未来

いずれコードも学習もいらない時代がくるでしょう。
「やりたいことがある人」にとって最高の時代になる。

それまで巨大な変化に対応しながら進むしかない。
使っている人と使っていない人では大きな差がつく。
好き嫌いで使う使わないの判断はしないで欲しい。

「計算資源」と「倫理」の問題によって、
抑制がかかるステージが来るかもしれません。
しかし、それをA Iの限界と勘違いしてはいけません。
A Iの発展は続くと考えたほうが自然です。

A I 解析入門
おわり