

生成AIで初めてできるようになったこと,
過去のAIにはできなかったこと

小樽商科大学

佐山公一



授業の目的

生成AIのしくみをイメージできるようにする
(生成AIのメンタルモデルをつくる)。

数学を使わないで、生成AIのソフト的な
しくみを視覚化する。

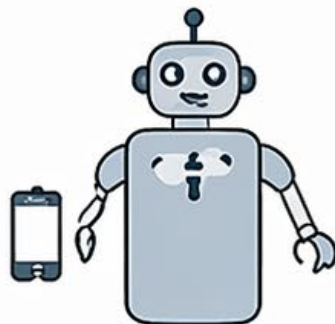
事務職(ホワイトカラー)における 仕事の効率化の歴史

第1段階 クラウドサービス



どこでも作業できる
例：確定申告も
クラウドで完結

第2段階 従来型AI



定型業務を自動化
例：RPA、BIツールで
データ入力・集計

第3段階 生成AI



個人・組織に最適化
例：メール作成、企画書、
採点の自動化

共通の仕事を効率化

個人ごとの仕事を効率化

生成AIと従来型AIの違い

過去のAI (BIツール、RPAなど)



複雑だがプログラムに
書ける定型業務を
自動化し、時間短縮

現在の生成AI (GPT、Geminiなど)



あなたっぽい
メールを作る

あなたの会社に
合う企画書を書く

あなたの視点基準で
レポート評価する

組織の文化や個人の好みを反映し
個人の能力を最大化できる

生成AIが得意なこと

① 個別化・個性化



企業文化や好みの情報の入った
独自データベースを活用

特定の組織や個人にしかできない仕事や
正解が人によって違う仕事を効率化

② 数値と言葉の同時分析



数値データと言葉のデータを同時に分析し、
より正確な未来予測ができる



Web上の膨大な言葉から学習したLLM（大規模言語モデル）
を使うから正確になる

生成AIの強み：個別化

従来のAI (みんな同じ)



- 共通のルールで動く
- 個人の違いは反映できない

生成AI (個人ごとに最適化)



- 組織のルールや個人の好みを学習
- **あなただけのAIになる**

できることの例



自分っぽい
メールを書く



会社独自の
企画書を作る



自分の採点基準で
レポート評価

生成AIの活用で広がる可能性（例：レポート採点）

従来の採点

1つずつ読んで
点数をつける



時間がかかる
採点基準がぶれる



生成AIを使った採点

ワンクリックでAIが採点
→エクセルに出力



効率化
採点基準が一定に

教員はAIの出した
成績を確認



生成AIは「次に来そうな単語」を予測している（例1）

昨日、図書館で本を〇〇。

人は「借りた」だろうと予想

「借りた」 (高い確率)

「返した」 (やや低い)

「読んだ」 (やや低い)

「失くした」 (低い)



高い確率



低い確率

生成AIは「次に来そうな単語」を予測している（例2）

昨日、コンビニで弁当を〇〇。



人は「買った」だろうと予想

「買った」 (高い確率)

「食べた」 (やや低い)

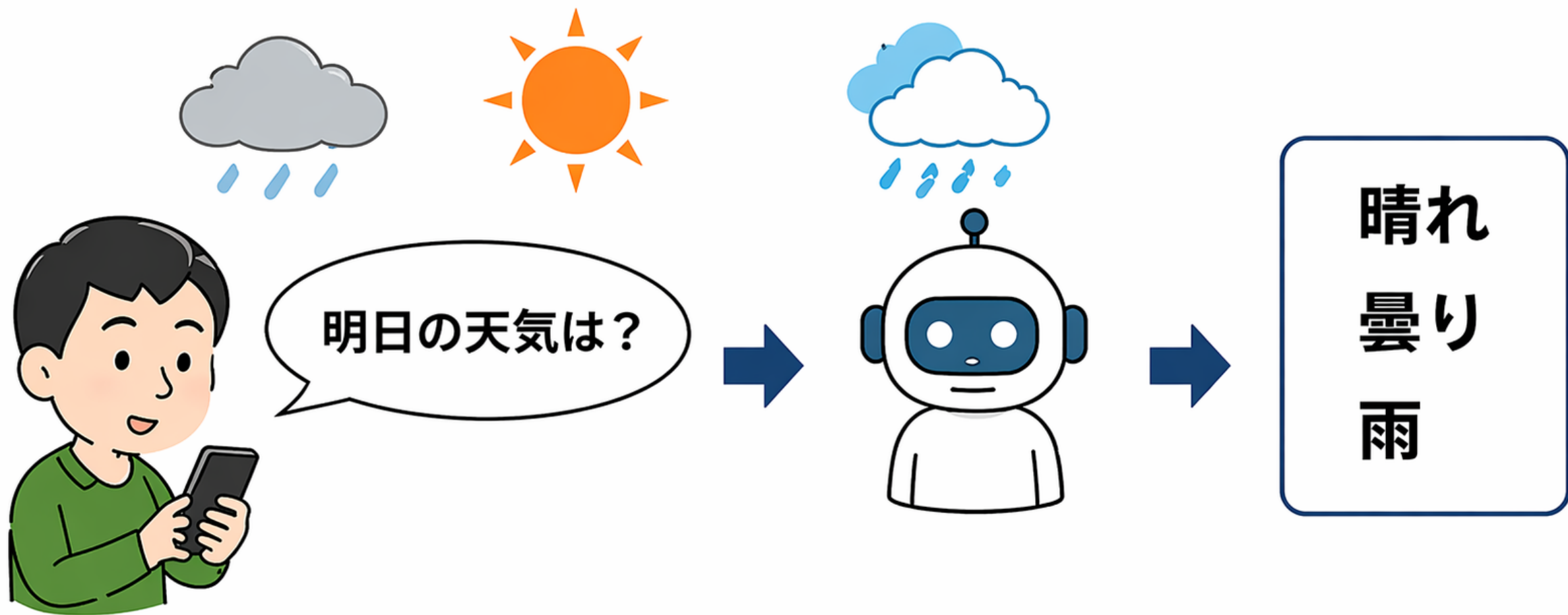
「落とした」 (低い)

高い確率



低い確率

「次に来そうな単語」はその場の状況で決まる



どれを出すかは、ユーザの状況・文脈で決まる

生成AIを理解するには？

2000年代までの過去のAIと認知心理学は共通の考え方をしていた。生成AIと認知心理学・脳科学とは出発点・発想が同じだけ。

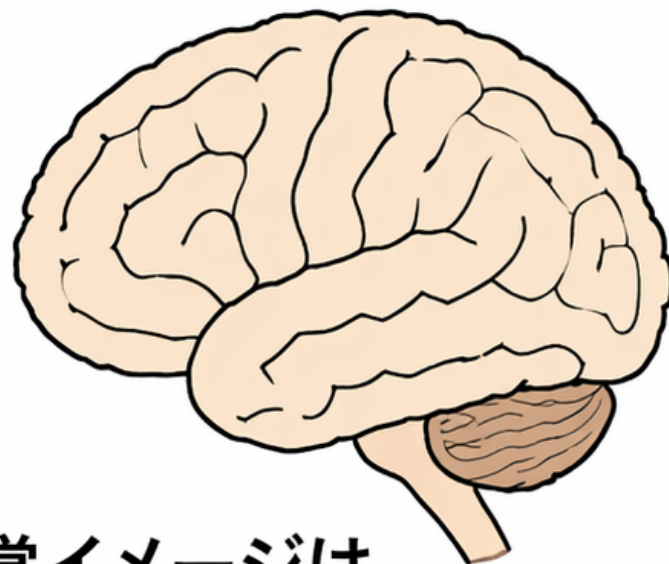
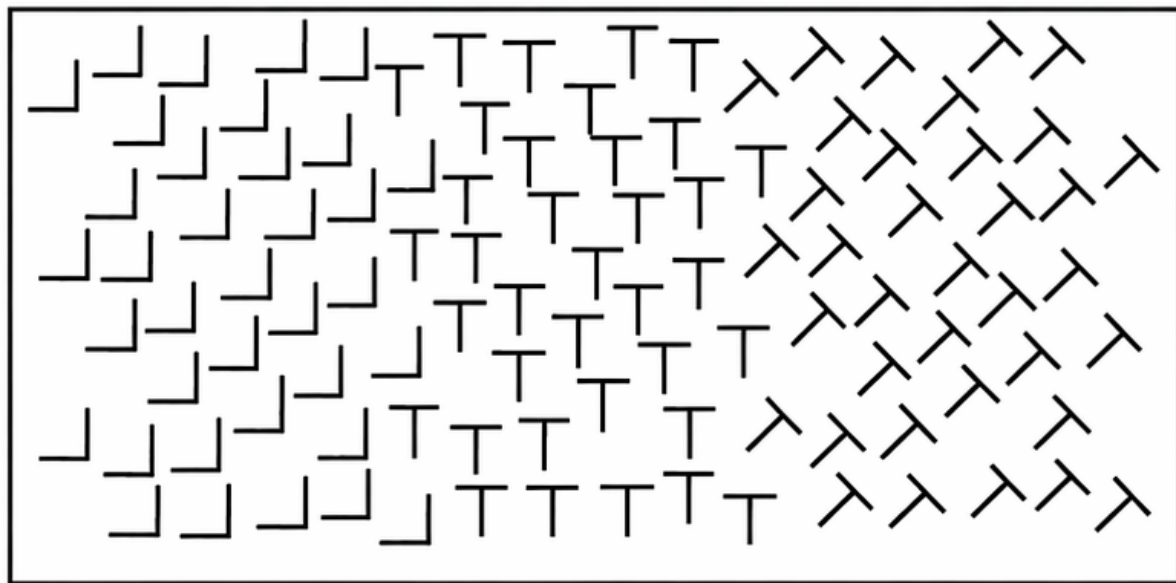
過去のAIの歴史

認知科学
(人間の情報処理)

この2つの関係を理解すると、
生成AIの技術の背景がわかる

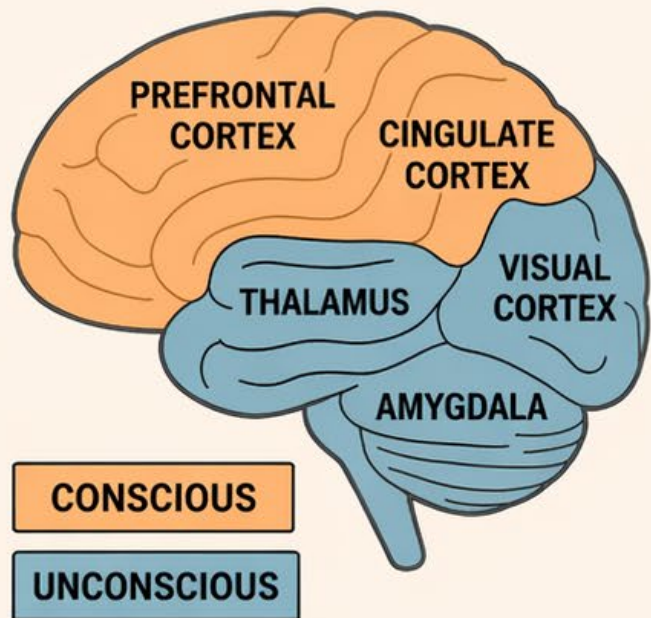
知覚的群化：人は部分に分けて物を見る

どちらの境界が
はっきり見えますか？



視覚イメージは
後頭部の第一視覚野 (V1) に行き、
そこからV2、V3...へ
運ばれる

人の視覚は大脳で処理される



人の視覚は大脳（旧脳ではなく）にある直感的な判断で使う情報は膨大なため、進化の過程で大脳に移された

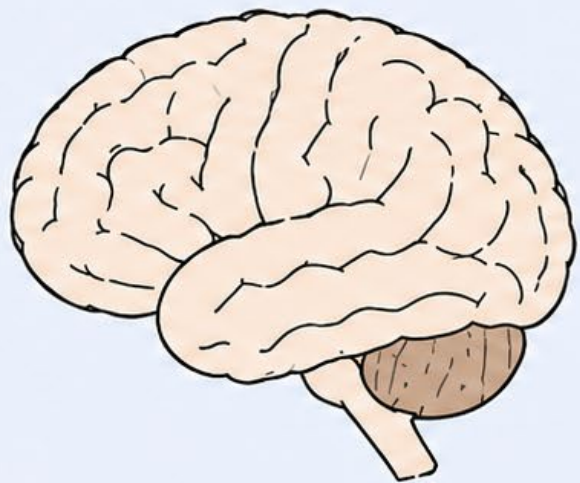


情報の全体像はまだよくわかっていないが、それでも生成AIは人間と同じような出力を出す

学習と推論の違い

人の脳

学習と推論が
混ざっている



生成AI

学習と推論が
明確に分かれている

学習フェーズ

事前学習・追加学習

推論フェーズ

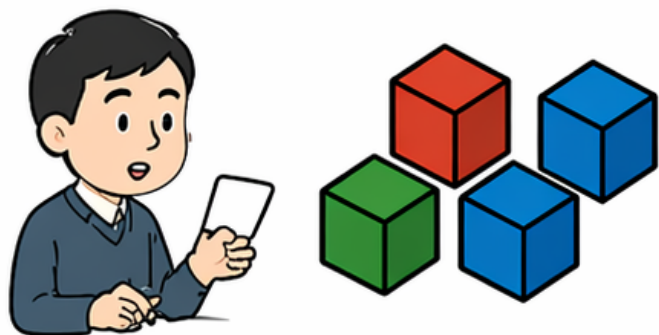
チャットそのもの



生成AIの技術の原点：積み木世界

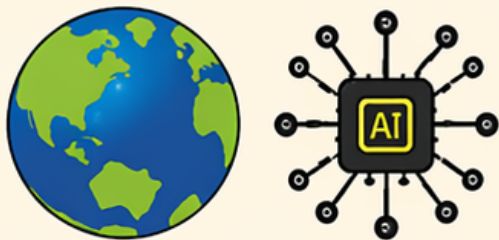
1960年代のAI研究

Terry Winograd (スタンフォード大学)



- 世界を限定する
- 知識源 (情報源) を区別する

現在の生成AI



世界を限定しない
代わりに次に来そうな単語の
確率を予測する

- 1 言葉をn次元ベクトルにする技術 (埋め込み)
数値データと言葉のデータを明確に分析可能に
- 2 ニューラルネットワークの実用化 (2010年代～)
多層NNと事前学習で大規模モデルが実現
- 3 self-attentionの仕組み (2017年～)
重要な部分に注目して予測精度が向上

過去のAIは「人間が教える」ものだった

過去のAI

人間がルール(プログラム)を書く.

例:

「丸い + 目 + 耳 → 猫」

「音の高さがこう → 母音」

全部, 人間が特徴を決めていた.

生成AIは特徴そのものを自分で学ぶ

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.

ヒントンの発想：「特徴も自分で学ばせよう」(自己教師あり学習)

多層ニューラルネットワークによる表現学習 (Representation Learning)

例：「猫」を認識する場合

過去のAI「耳は三角、ヒゲがある…」とあらかじめプログラムの中で定義しておく。

生成AI

1層目で線，2層目で角・曲線，3層目で目や耳

4層目で「あ！これは猫だ」

人間(研究者)が事前に定義しようとしても、定義しきれない特徴が画像や文章の中には無数にある。そういう特徴は認知心理学や脳科学でも何か分っていない。しかし、分かっていない特徴も含めて生成AIはあらゆる特徴を自分で学ぶ。

生成AIの革新的な発明：3つのポイント

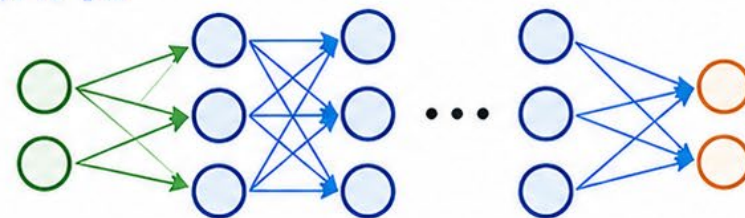
1 言葉をベクトルにする技術（埋め込み）

言葉や文章を数値に変え、
言葉と数値データを同時に
分析できるようになった！



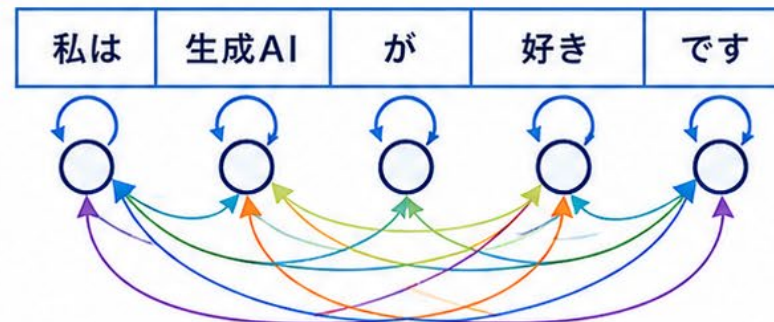
2 多層ニューラルネットワークの実用化

たくさんの層を重ねることで、
複雑なパターンを学習できる
ようになった！



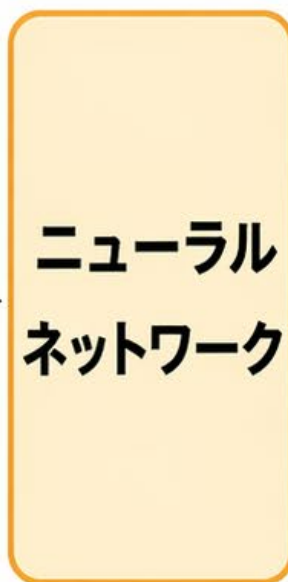
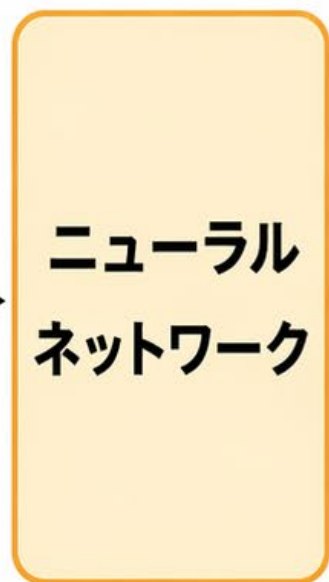
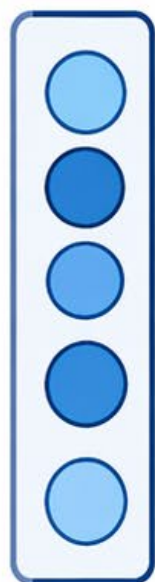
3 self-attentionの発明（2017年）

文脈をより正確に理解し、
次の言葉を高精度で予測
できるようになった！

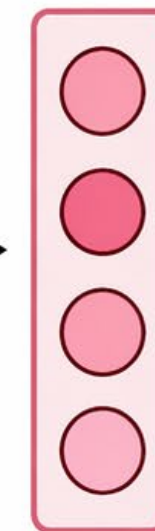


GPTシステムの基本構造

入力 (単語)



出力
(次の単語の確率)



self-attentionと単層ニューラルネットワークを繰り返す

学習フェーズ（事前学習・追加学習）

事前学習（エンジニアが行う）

Web上の膨大なテキスト・画像などを使って学習。

（GPT-4は2023年11月までの情報など）

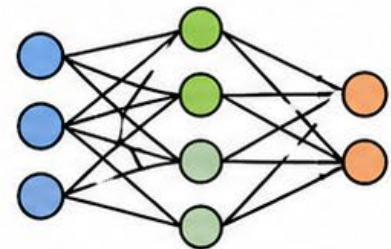
→ LLMの重みを繰り返し計算して学習。



Web上の
データ



テキスト
など



ニューラルネットワーク
(LLMのモデル)



重みには3種類ある

● クエリ (Query)

● キー (Key)

● バリュー (Value)

追加学習（ユーザが行う：ファインチューニング）

独自データを使って、3種類の重みを調整。

必要なときだけ反映された重みを使える。



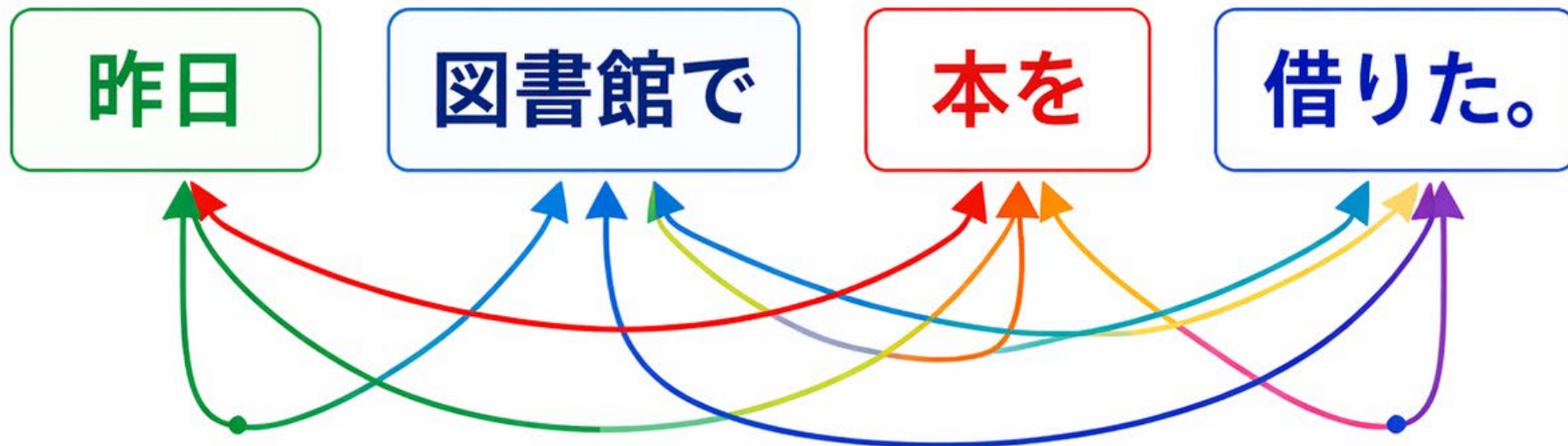
独自データ



self-attention (セルフアテンション) とは？

文章の中のすべての言葉の関係を考えるしくみ

例) 「昨日、図書館で本を借りた。」の場合

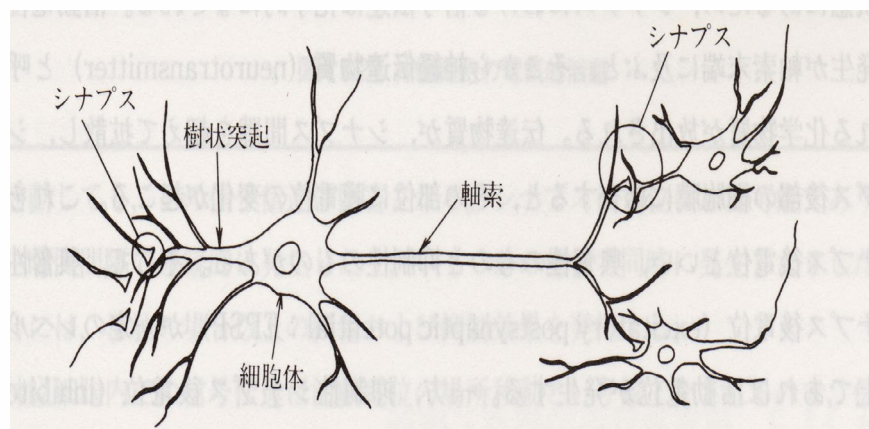


それぞれの言葉が、他のすべての言葉と
どれくらい関係があるかを計算する

ニューラルネットワークとは？

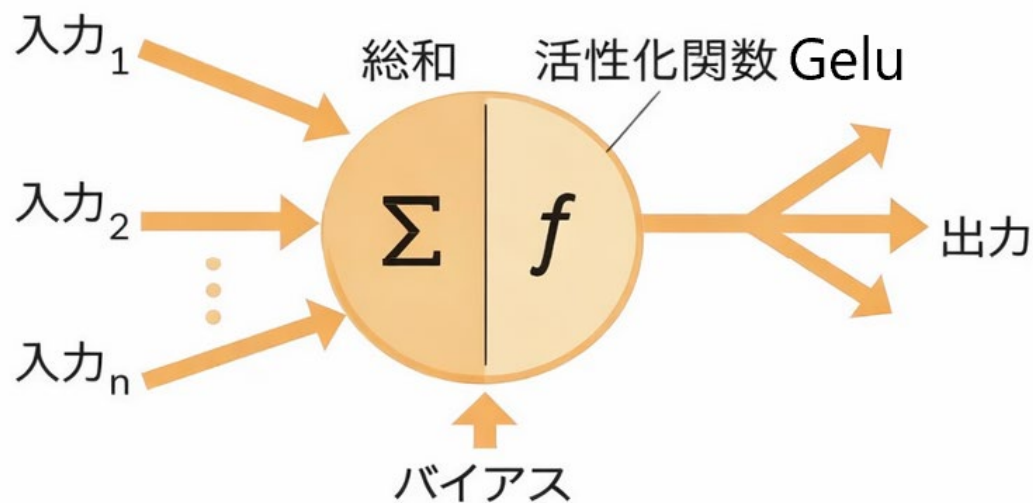
人の脳のしくみをまねた計算モデル

人の脳の神経細胞 (イメージ)



シナプス (つなぎ目) で
情報を伝える

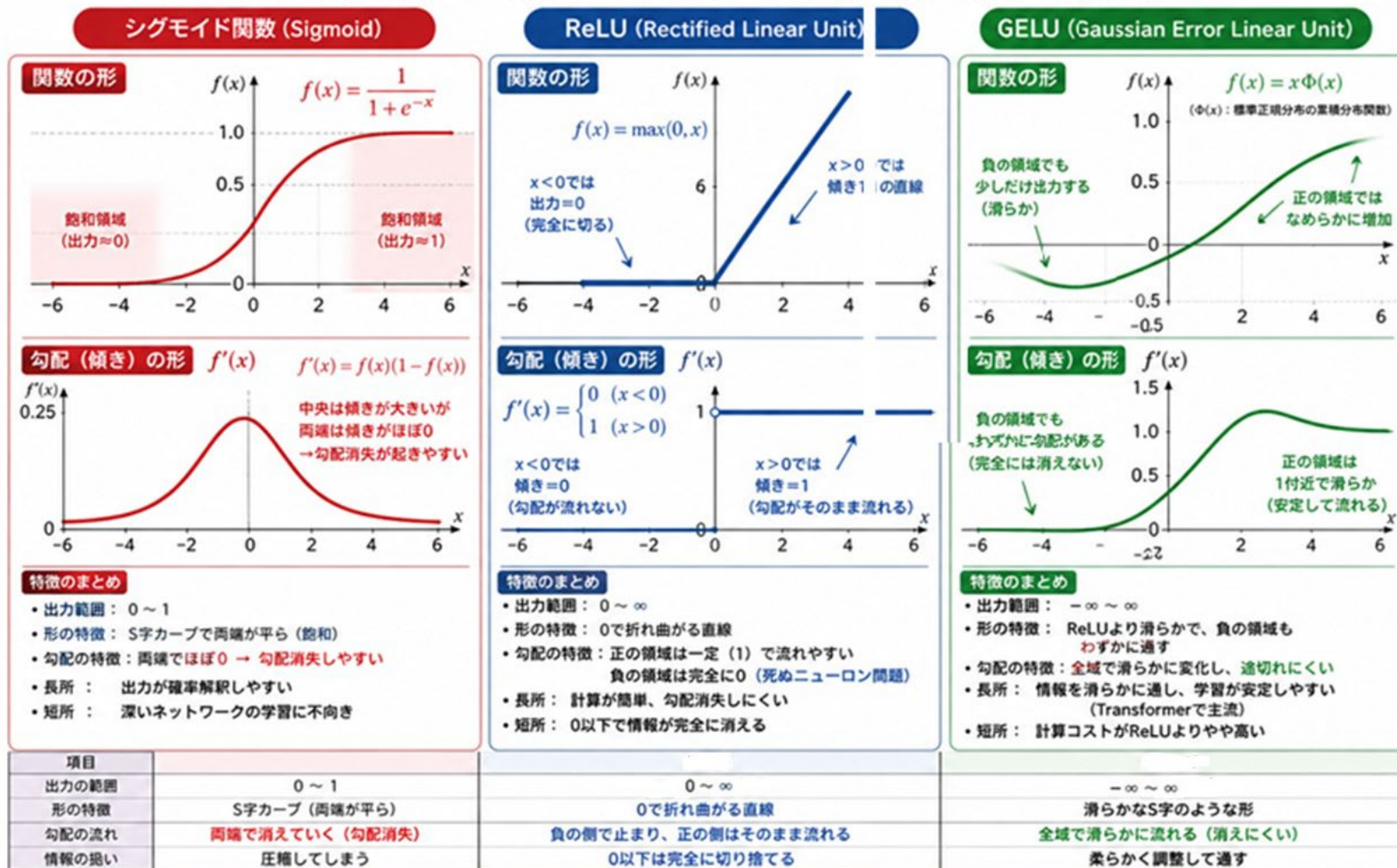
人工ニューロン (イメージ)



たくさん的人工ニューロンをつなげて、
複雑な計算を行う

活性化関数の変遷 Geluには情報の消失がない

シグモイド関数・ReLU・GELU の比較 (形・傾き・特徴の違い)



まとめ: シグモイドは「情報が消えやすい」・ReLUは「0で切る」・GELUは「滑らかに通す」

なぜ中間層が重要なのか？

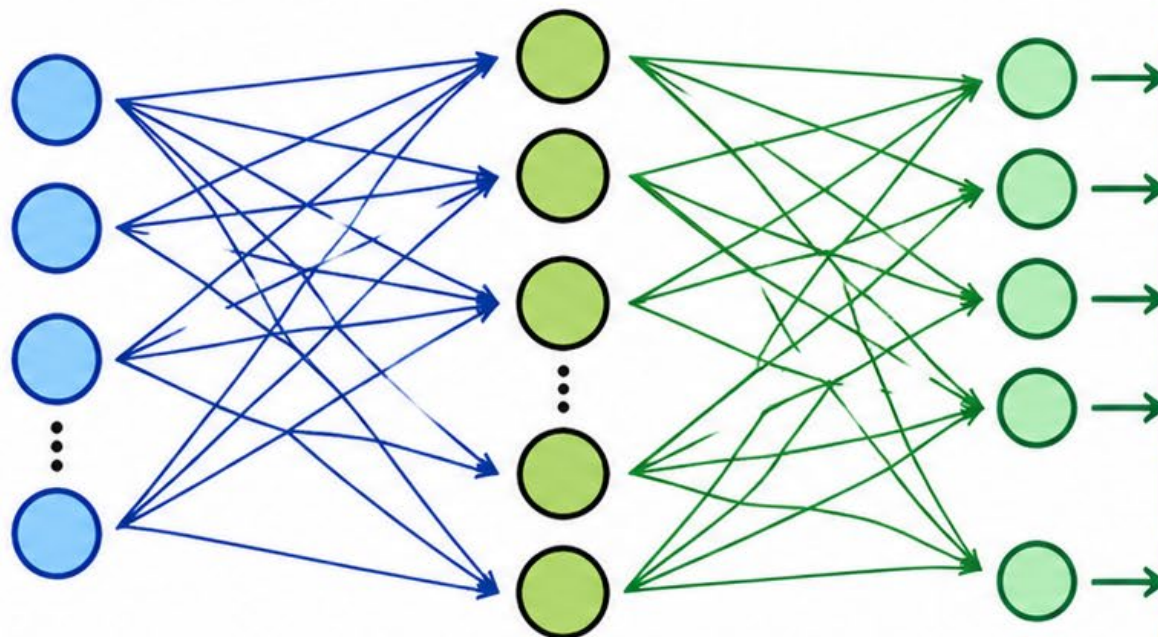
入力層

(数字の画像など)



中間層

(特徴を抽出)



出力層

(3である確率など)

0	→	0.03
1	→	0.05
2	→	0.10
3	→	0.78
		⋮
9	→	0.02

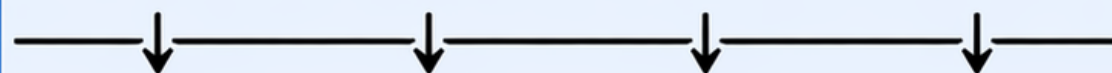
中間層が「**特徴 (部分)**」を見つけることで、
意味を理解できるようになる！

バックプロパゲーション（逆伝播）のイメージ

出力（AIの予測）

最初はすべて同じ

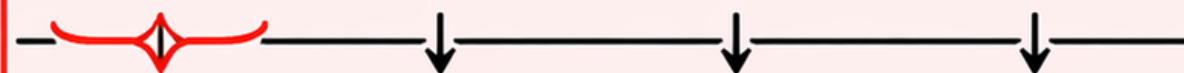
0.25 0.25 0.25 0.25



借りた 返した 読んだ 失くした

期待する出力（正解に近づける）

0.50 0.30 0.15 0.05

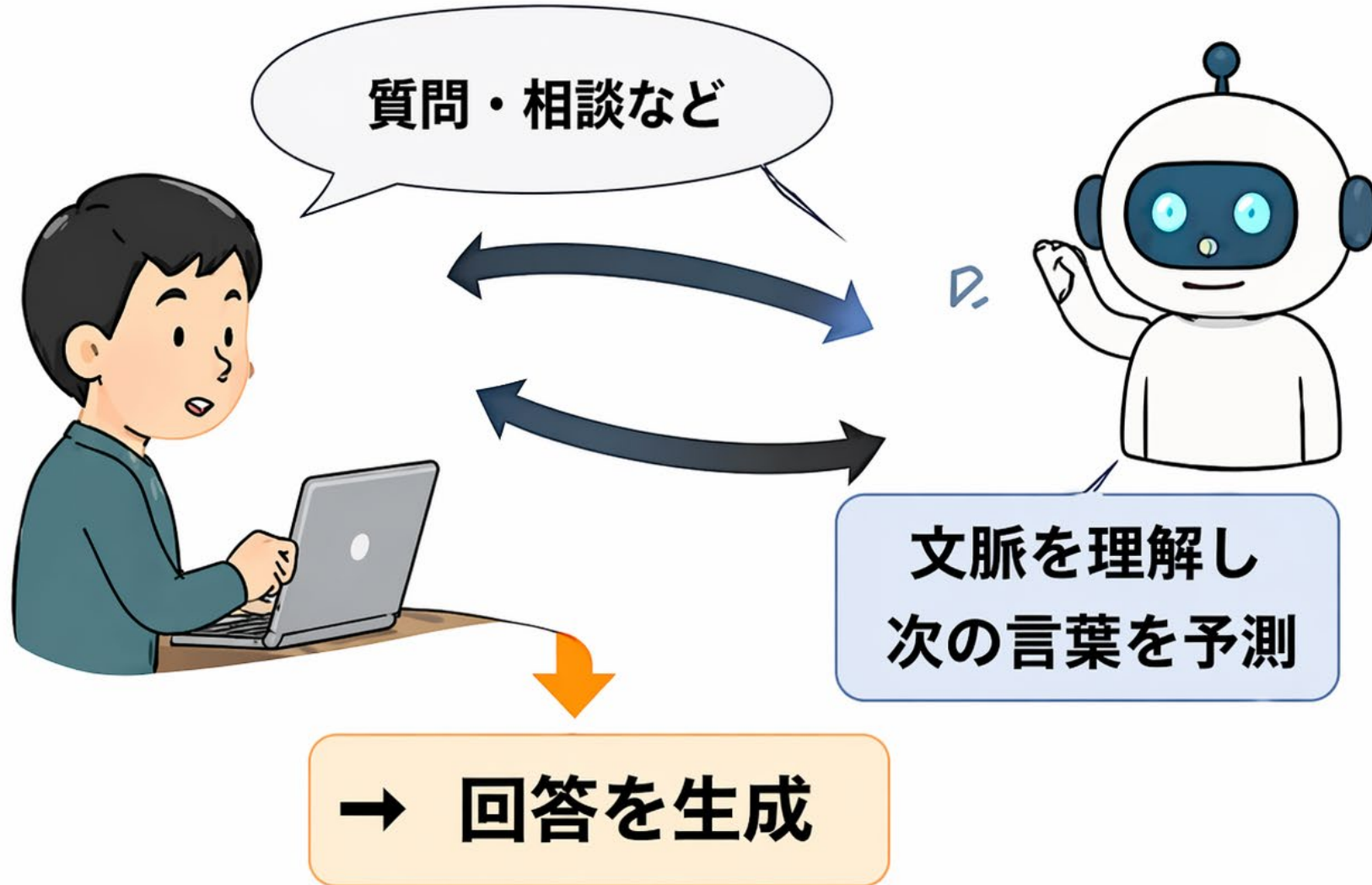


借りた 返した 読んだ 失くした

誤差（ずれ）をもとに、中間層の重みを調整

→ 出力が期待値に近づくように学習（逆伝播）

推論フェーズ：会話そのもの



推論フェーズ（ユーザとの会話）

ユーザの入力（質問）

回答（出力）



会話のオンライン処理（その場での文脈理解）をよくまねできる。

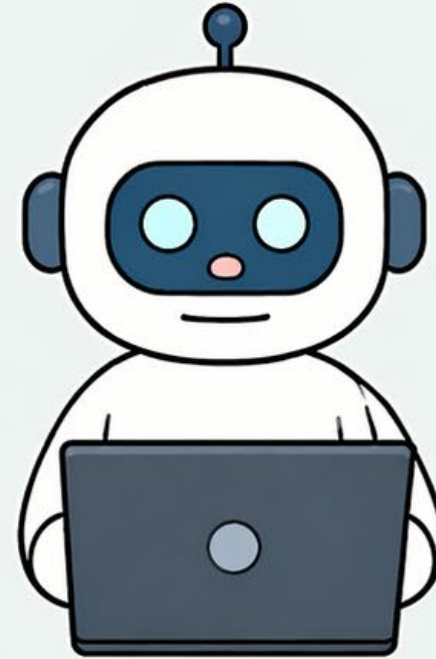
生成AIは超優秀な占い師である

人間の占い師



- ・過去の情報を見る
- ・今の様子を見る
- ・未来を予測する

生成AI



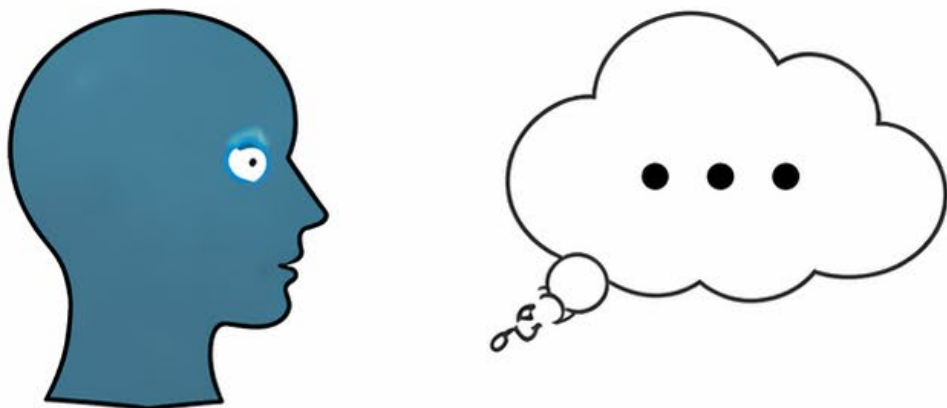
- ・学習データを見る
- ・会話の様子を見る
- ・それっぽい未来を予測

学習と推論が明確に分かれているのが違い

生成AIは超優秀な占い師である(まとめ)

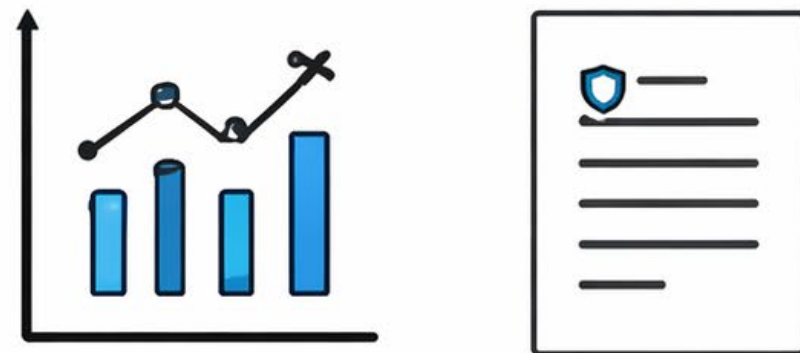
① 次に来る言葉を正確に予測

言葉の理解は確率判断である



② 未来を占う

数値と言葉を同時に分析できる



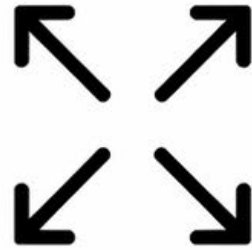
チャット時に**学習済みのLLM**を使う

生成AIのしくみをイメージすることが大切！

仕組みが分かると…



使い方を
工夫できる



応用の幅が
広がる



問題解決の
精度が上がる

生成AIのメンタルモデルを持とう！

授業の目的

生成AIのしくみをイメージで理解できるようにする

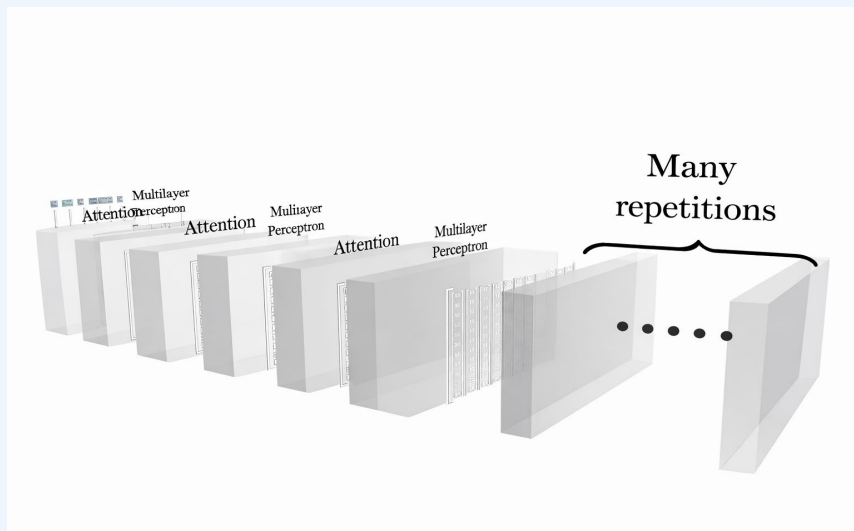


数学を使わないで、
生成AIのソフト的なしくみを
視覚化する

→ 生成AIで個人ができることを最大化する

授業のゴール

生成AIのソフト的なくみを
イメージする



自分だけの
メンタルモデルを持つ



→ 何に使えるかを予想でき、応用の幅が広がる