IT資格道場

# 生成AIパスポート 学習用テキスト

予備知識ゼロから合格ラインへ —— 読んで理解する1冊目

無料サンプル(第1章まで)



GUGA 公式「生成AIパスポートの試験出題範囲（2026年2月試験以降）」改訂日 2025-10-01 準拠
2026年7月版

## 本書のご利用にあたって

- 本書は「IT資格道場」(www.shikaku-doujou.com)の購入者限定テキストのうち、学習用テキストの第1章までを収録した**無料サンプル**です。どなたでもご覧いただけます。
- 無料サンプルであっても著作物です。本書の転載・改変しての再配布・二次販売を禁じます。
- 本書を AI モデルの学習データとして利用すること(学習目的の入力・データセット化を含む)を禁じます。
- 製品版(学習用テキスト・頻出度別ノート・暗記用ノートの3点)は、ご購入者を識別できる透かし入りのPDFとして提供されます。
- 本PDFはスマートフォン・タブレットでの閲覧に合わせたA5版面・文字サイズで組版しています。



不正な配布を見つけた場合は info@shikaku-doujou.com までご連絡ください。

# 目次

## 3点セットの中の位置づけ

このシリーズは3つのテキストで構成されています。

- **学習用**: これだけで問題が解けるようになる本編
- **頻出度別**: テキストを読んだら次はこれ。頻出順に絞った問題と解説で「解ける」に橋渡しします。試験直前の最終チェックにも
- **暗記用**: 覚えるべき要点だけを一問一答に凝縮。空き時間の反復と用語の再確認用

**使い方**: ① 学習用を読む → ② 頻出度別で解き方を掴む → ③ 問題演習で腕試し → ④ 頻出度別に戻って再確認(試験直前もここ)。暗記用は空き時間や用語の再確認に。

**このテキストの位置づけ**: これだけで問題が解けるようになる本編です

## はじめに —— この教材の使い方

このテキストは、**一般社団法人 生成AI活用普及協会(GUGA)** が実施する **生成AIパスポート試験** の合格を目的に作られています。

この試験が問うのは、モデルを作れることでも、プログラムを書けることでもありません。問われるのは、**生成AIという道具の輪郭を正しくつかみ、仕事や生活の中でどう使い、どこで止め、何に気をつけるかを判断できること**です。技術の章と、法律・リテラシーの章と、プロンプトの実践の章が同じ試験の中に並んでいるのは、そのためです。**どれかひとつだけでは合格ラインに届かない設計**になっています。

本書は **GUGA が公開しているシラバス(試験出題範囲)の項目を体系化した教材**であり、実際に出題された問題を再現したものではありません。**本書は GUGA の公認教材ではなく、IT資格道場が独自に書き下ろしたオリジナルの教材です。**

**受験前提の知識は要りません。**プログラミングの経験も、数学の知識も不要です。本書は数式をひとつも使わずに書いてあります。専門用語は、はじめて出てきた場所で必ず日本語の言い換えを添えました。

## 試験の基本情報

| 項目 | 内容 |
|---|---|
| 試験名 | 生成AIパスポート試験 |
| 主催 | 一般社団法人 生成AI活用普及協会(GUGA) |
| 開催形式 | **オンラインでの実施(IBT方式)** —— 自宅などのパソコンから受験します |
| 試験時間 | **60分間** |
| 問題数 | **60問** |
| 出題方法 | **四肢択一式(一部複数選択を含む)** |
| 出題範囲 | シラバス(試験出題範囲)より出題 |
| 受験資格 | **制限なし** |
| 受験費用 | **11,000円**(税込) / 学生は **5,500円**(税込) |
| 結果発表 | 受験期間の終了後、1か月以内に会員ページで通知 |
| 資格の有効期間 | **無期限**。合格者には合格証書とオープンバッジが発行されます |

**受験料・試験時間・出題数・シラバスは改定されることがあります。**お申し込みの前に、GUGA 公式サイトで最新の情報をご確認ください。

**GUGA は合格基準(何点で合格か)を公表していません。**したがって本書は「何割取れば受かる」という数字を書きません。書けば、推測を事実の顔で並べることになるからです。目指すのは、**シラバスの全項目について「聞いたことがある」ではなく「説明できる」状態**にすることです。

### 受験の時期と、シラバスの版

生成AIパスポート試験は **年5回(2月・4月・6月・8月・10月)** 開催されます。各回とも、**受験期間はその月のあいだ**で、申込はその**前々月の1日から前月末まで**です。たとえば10月試験なら、申込期間が8月1日から9月30日まで、受験期間が10月1日から10月31日までです。**申込期間と受験期間がずれている**点に注意してください。受験したい月の前月末までに申し込みを済ませておく必要があります。

もうひとつ、この試験に固有の事情があります。**シラバスが年1回以上のペースで改訂される**ことです。生成AIは動きの速い分野なので、出題範囲そのものが更新されていきます。本書が対象としているのは、**2026年2月試験から適用されているシラバス(改訂日 2025年10月1日)** です。この改訂では、第2章に ChatGPT の新しいモデル群が、第3章に RAG(検索拡張生成)と AIエージェント が、第4章に AI新法 が、それぞれ新たに加わりました。**新しく加わった項目は、出題側が「そこを問いたい」と考えて足した箇所**です。本書はこの3か所を特に厚く書いています。

なお、すでに合格して資格を持っている人向けには、改訂のたびに**資格更新テスト**が実施されます。合格の有効期間そのものは無期限ですが、知識を最新に保つための仕組みが別に用意されている、という位置づけです。

### 受験環境について —— 当日になって慌てないために

IBT方式(インターネット経由で受験する方式)なので、受験環境は自分で整える必要があります。公式が示している条件のうち、**うっかりすると当日に困るもの**を挙げておきます。

- **カフェ・公園・レストランなどの公共スペースでは受験できません。**試験中、自分以外の人が入ってこない場所であることが条件です
- **デュアルディスプレイ・複数モニターは使用不可**です。ふだん2画面で作業している人は、事前に1画面の状態を作っておいてください
- **机の上に物を置けません。**ハンカチ・ティッシュなどの衛生用品と、受験に配慮が必要な方の必要物を除き、**筆記具も禁止**です。計算やメモを取りながら解く前提では組み立てられません
- **翻訳ツール(Google Translate など)を有効にしたままだと試験画面が正常に動かないおそれ**があります。受験開始前に無効にしておきます
- 通信は安定した回線を用意します。**開始後の途中退出はできません**

この「筆記具が使えない」という条件は、学習の設計にも効いてきます。**メモを取って整理しながら解く**やり方は本番では使えないので、ふだんの演習から**頭の中だけで選択肢を絞る**練習をしておくと安全です。

### 時間配分 —— 「調べながら解く」はできない

**60分で60問**ということは、**1問あたりちょうど1分**です。設問を読み、四つの選択肢を読み、選ぶまでが1分。見直しの時間まで含めれば、実際に使えるのは1問50秒ほどでしょう。

オンライン受験なので手元で調べられるように思えるかもしれませんが、**調べる時間は実質的にありません。**1問でも検索に手を出すと、その分だけ後半が削られます。目指すのは、**用語を見た瞬間に、それがどの章のどの話かが反射**

**で出る状態**です。本書の各章末に「試験ポイント」を置いているのは、この反射を作るためです。

## 出題範囲の全体像 —— シラバスの5章と本書の章

| シラバスの章 | 分野 | 本書 | 一言でいうと |
|---|---|---|---|
| 第1章 | AI(人工知能) | 第1章 | AI とは何か。仕組み・種類・歴史 |
| 第2章 | 生成AI(ジェネレーティブAI) | 第2章 | 生成モデルの系譜と ChatGPT |
| 第3章 | 現在の生成AIの動向 | 第3章 | できること・ディープフェイク・RAG・AIエージェント |
| 第4章 | 情報リテラシー・基本理念とAI社会原則 | 第4章 | セキュリティ・個人情報・権利・原則・AI新法 |
| 第5章 | テキスト生成AIのプロンプト制作と実例 | 第5章 | プロンプトの基礎と、仕事での使い方 |

**本書の章の並びは、シラバスの並びとまったく同じです。** 章番号・節の順序を公式にそろえてあるので、シラバスを手元に置いて対照しながら読めます。

**GUGA は、章ごとの出題比率を公表していません。** 公式ページにもシラバス本文にも記載がないため、本書は「この章が何パーセント出る」という数字を書きません。

代わりに根拠にしたのは、**シラバスに列挙されている学習項目そのものの厚み**です。第4章(情報リテラシーと法・原則)と第5章(プロンプトの実践)は学習項目が突出して多く、第2章の ChatGPT も同様です。本書はこの3か所を厚く

書きました。**分量の配分そのものが、学習時間の配分の提案**だと受け取ってください。

## 全設問に効く判断軸 —— 「3つの問い」

この試験の難しさは、内容が難解なことではありません。**似た言葉が大量に並び、どれがどれだか分からなくなること**です。教師なし学習と半教師あり学習、VAE と GAN、転移学習とファインチューニング、要配慮個人情報と機微(センシティブ)情報、肖像権とパブリシティ権 —— どれも「聞いたことはある」で止まると、四つの選択肢の前で手が止まります。

そこで本書は、**どの章の設問にも当てられる3つの問い**を用意し、全5章の章末「試験ポイント」まで一貫して使います。

| 問い | 何を切り分けるか | 分かれ方 |
|---|---|---|
| **問い1: どの層の話か** | 設問が立っている場所 | **技術層**(第1・2章。仕組みがどう動くか) / **動向層**(第3章。いま何ができ、何が起きているか) / **規範層**(第4章。守るべき線は何か) / **実践層**(第5章。どう指示を書くか) |
| **問い2: 学習の話か、利用の話か** | 時間軸のどこか | **学習時**(モデルを作っている最中) / **利用時**(できたモデルを使っている最中) |
| **問い3: 何を入力に、何を出力するか** | 手法・サービスそのものの同定 | 入力と出力の組が違えば、それは別のものです |

**問い1** は、同じ言葉が層をまたいで出るときに効きます。「プライバシー」は、規範層では個人情報保護法の話(何が個人情報にあたるか)、動向層では利用

時の注意の話(入力した内容がどう扱われるか)です。同じ語でも、求められている答えが違います。

**問い2** は、技術と法律の両方に刺さります。技術側では、**過学習・正則化・ドロップアウト・転移学習は学習時の話、ハルシネーション・プロンプトの工夫は利用時の話**です。法律側でも、**学習のためにデータを集める場面**と、**生成された出力が誰かの権利を侵害する場面**は別の論点として問われます。この線を引くだけで選択肢の半分が消えることがあります。

**問い3** は、モデルとサービスの暗記量を減らします。GAN は「二つのネットワークを競わせて本物らしいものを作る」、VAE は「入力をいったん圧縮してから復元する」、RAG は「外部の文書を検索して、その内容を材料に答えを作る」。**入力と出力の組で覚えれば、名前を忘れても機能から引き当てられます。**

各章末の「試験ポイント」では、この3つの問いに引き付けた形で要点をまとめます。

## 用語の表記について

本書は、**同じものを指す複数の呼び方を、初めて出てきた場所で両方書きます。** 検索拡張生成(RAG)、大規模言語モデル(LLM)、深層偽造(ディープフェイク)、過学習(オーバーフィッティング)、技術的特異点(シンギュラリティ) —— といった具合です。

理由は単純で、**本番の設問がどちらの表記で来るか分からない**からです。片方だけを覚えていると、もう片方が出た瞬間に知らない用語に見えてしまいます。読みにくく感じるかもしれませんが、**両方を結びつけて覚えることそのものが対策**だと思ってください。

### 4ステップ学習法

購入者限定テキストは3冊で1セットです。

- **STEP 1(理解する)**: 本書「学習用テキスト」を通読し、5章の地図と3つの問いを頭に入れる
- **STEP 2(解き方を掴む)**: 「頻出度別ノート」で頻出度の高い論点を解いて、解き方を掴む
- **STEP 3(腕試し)**: 本サイトの問題演習で、四肢択一の形式に慣れる
- **STEP 4(再確認)**: 「頻出度別ノート」に戻って総ざらい。試験直前もここへ戻る

「暗記用テキスト」は本線には入れません。通勤時間や休憩時間など、**空き時間の反復と用語の再確認**に並走させてください。

それでは、第1章から始めます。

# 第1章 AI（人工知能） [1.1/1.2/1.3/1.4/1.5]

生成AIの話は、いきなり「ChatGPT の使い方」から始めたくなります。けれどこの試験は、第1章に **AI（人工知能）そのもの** を置いています。理由は単純で、**生成AIは AI の一種**だからです。土台にある「AI とは何か」「AI はどうやって賢くなるのか」を飛ばすと、後の章に出てくる大規模言語モデル（LLM）やハルシネーションの話が、丸暗記の対象になってしまいます。

この章で扱うのは、**AI の定義**、**AI に知能をもたらす仕組み**、**AI の種類**、**AI の歴史**、そして **シンギュラリティ（技術的特異点）** です。とくに厚く読んでほしいのが「AI に知能をもたらす仕組み」で、機械学習・ニューラルネットワーク・ディープラーニングという、この試験の背骨にあたる用語がまとめて出てきます。

読み方のコツを先に1つ。**この章の設問は、似た言葉の取り違えを突いてきます。**「教師なし学習」と「半教師あり学習」、「転移学習」と「ファインチューニング」、「シンギュラリティ」と「AI効果」、「ヴァーナー・ヴィンジ」と「レイ・カーツワイル」。どれも「聞いたことはある」で止まると、四択の前で手が止まります。各項目の最後に **誤りやすい対比** を置いてありますので、そこを重点的に読んでください。

## 1-1 AI（人工知能）とは何か [1.1]

### AI には、研究者が合意した定義が存在しない

まず、いちばん意外なところから始めます。**AI（人工知能、Artificial Intelligence）には、研究者の間で合意された唯一の定義がありません。**

「人間の知的な振る舞いを機械で再現しようとする技術」といった言い方は広く使われます。総務省や経済産業省の資料でも、おおむねこの線で説明されま

す。ただし、この言い方には逃げ道があります。**「知的とは何か」がそもそも定まっていない**からです。人間の知能そのものが解明されていない以上、それを「再現する技術」の輪郭もぼやけます。結果として、研究者ごとに少しずつ違う定義を掲げている、という状態が今も続いています。

試験対策としては、この「定まっていない」こと自体が問われます。**「AI の定義は国際標準化機関によって一意に定められている」「学会が公式定義を制定している」といった選択肢は、それだけで誤り**です。逆に「研究者によって定義が異なる」「明確な定義は存在しない」という趣旨の記述は正しい側に来ます。

そのうえで、実務的にはこう捉えておくと迷いません。**AI とは、これまで人間の頭脳が担っていた「認識する・判断する・予測する・作り出す」といった知的な処理を、コンピュータに行わせる技術の総称**です。ここでいう「総称」が要点で、AI は特定の1つの手法の名前ではありません。後で出てくるルールベースも機械学習もディープラーニングも、すべて AI という大きな傘の下に入ります。

### AI・機械学習・ディープラーニングは入れ子になっている

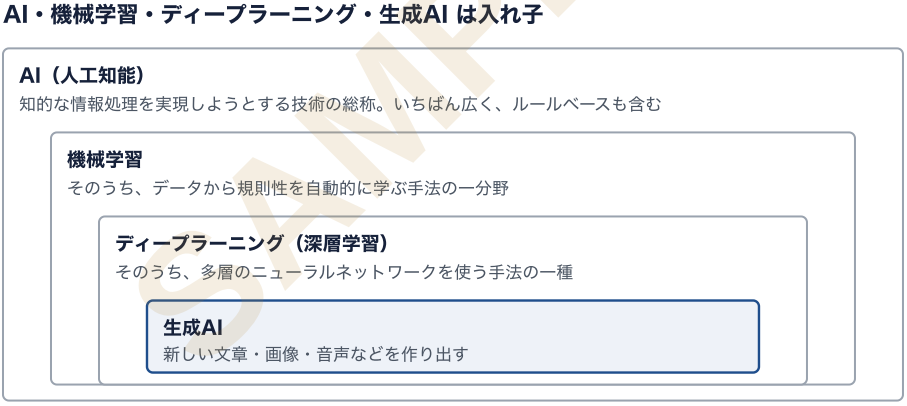
この3語は、**横に並ぶ別物ではなく、入れ子（包含関係）** です。ここを最初に固定してください。

- **AI（人工知能）** —— 知的な情報処理を実現しようとする技術の**総称。いちばん広い**。人がルールを書く方式（ルールベース）も含む
- **機械学習** —— そのうち、**データから規則性を自動的に学ぶ**手法の一分野
- **ディープラーニング（深層学習）** —— そのうち、**多層のニューラルネットワークを使う**手法の一種

つまり **AI ⊃ 機械学習 ⊃ ディープラーニング** です。投資でたとえるなら、ポートフォリオ全体（AI）・その中の株式（機械学習）・さらにその中の1銘柄（ディープラーニング）という関係です。

この向きを逆にした選択肢 ——「ディープラーニングがいちばん広い概念である」「機械学習は AI とは別の独立した領域である」「三者は互いに重ならない」—— は、内容を細かく検討するまでもなく切れます。**包含の向きだけで正誤が決まる設問**は、実際によく出ます。

なお、生成AI もこの入れ子の中にあります。**生成AI は、ディープラーニングを使って新しい文章・画像・音声などを作り出す AI** であり、AI の外側にある別ジャンルではありません。




AI の中に機械学習があり、その中にディープラーニングがある。この向きを逆にした記述は、内容を検討するまでもなく切れる。生成AI は AI の外にある別ジャンルではない。

図 1-1　AI・機械学習・ディープラーニング・生成AI は入れ子

## AI とロボットは別のもの

定義が定まらない言葉だからこそ、**隣にある概念との境目**を引いておく価値があります。いちばん問われるのが **AI とロボットの区別** です。

| 概念 | 中身 | AI との関係 |
| --- | --- | --- |
| **AI（人工知能）** | 認識・判断・予測・生成といった**知的な情報処理を担うソフトウェアの技術** | 物理的な身体を必要としない |
| **ロボット** | モーター・関節・アーム・センサーといった**駆動部を備えた物理的な機械** | AI を積んでいるロボットもあれば、積んでいないロボットもある |

要点は2つです。

1. **AI に身体は必須ではない。** ChatGPT のようなチャットサービス、迷惑メールの判別、クレジットカードの不正検知 —— どれも物理的な身体を持ちませんが、れっきとした AI です。
2. **ロボットに AI は必須ではない。** 工場で同じ動きを繰り返す産業用ロボットの多くは、あらかじめ人が書いた手順どおりに動くだけで、データから学習してはいません。これは「AI を搭載していないロボット」です。

**具体例で3通りを並べておきます。**

- **AI を積んだロボット** —— 部屋の形を学習して効率よく回る掃除ロボット、周囲を認識して動く協働ロボット
- **AI を積んでいないロボット** —— 決められた軌道を正確に繰り返すだけの産業用ロボット、遠隔操作で人が動かすロボット
- **身体を持たない AI** —— 対話型の生成AI、クレジットカードの不正検知、翻訳サービス、迷惑メールの判別

**似た紛らわしさが、事務作業の自動化ツールとの間にもあります。** 画面の操作手順を記録して繰り返す **RPA（ロボティック・プロセス・オートメーション）** は、名前に「ロボット」と付きますが、**物理的な機械ではなく、パソコンの中で**

**決められた手順を再現するソフトウェア**です。**あらかじめ決めた手順どおりに動くだけで学習しない**という点では、AI ではなく自動化に当たります。ただし、**AI と組み合わせて「読み取りは AI、その後の入力作業は RPA」と役割分担する使い方は一般的**です。**名前に惑わされず、「学習しているか」「身体があるか」の2点で切り分けてください。**

**判定は、2つの問いだけで足ります。「物理的な身体があるか」**と**「判断しているか（データから学んで、状況に応じて出力を変えているか）」**です。この2つを掛け合わせると、次の4通りに整理できます。

| | 判断している | 判断していない（決めた手順の繰り返し） |
|---|---|---|
| **身体がある** | **AI を積んだロボット**（掃除ロボット、周囲を認識する協働ロボット） | **AI ではないロボット**（決められた軌道を繰り返す産業用ロボット、自動ドア、自動販売機） |
| **身体がない** | **AI**（対話型の生成AI、不正検知、翻訳、迷惑メールの判別） | **単なる自動化**（毎朝7時に自動送信されるメール、RPA） |

**この表の右下と左下を混同しないことが、いちばんの要点です。**どちらも身体がありませんが、**判断しているかどうかで、AI と自動化に分かれます。**そして、**左上と右上を混同しないこと。**どちらもロボットですが、**AI を積んでいるとは限りません。**

**言い換えれば、「体が無くても AI は AI」であり、「判断しない自動機械は、体があっても AI ではない」**ということです。**ロボットは器（ハードウェア）、AI は頭脳（ソフトウェア）**と整理すると混ざりません。掃除ロボットのように、ロボットという器に AI という頭脳が載っている例もありますし、AI が入っていない自動ドアや自動販売機のような機械もあります。

**AI とロボット ―― 「身体があるか」「判断しているか」の2軸**

| | 判断している<br>データから学び、状況で出力を変える | 判断していない<br>決めた手順の繰り返し |
|---|---|---|
| **身体がある** | **AI を積んだロボット**<br>掃除ロボット、<br>周囲を認識する協働ロボット | **AI ではないロボット**<br>決められた軌道を繰り返す<br>産業用ロボット、自動ドア |
| **身体がない** | **AI**<br>対話型の生成AI、不正検知、<br>翻訳、迷惑メールの判別 | **単なる自動化**<br>毎朝7時に自動送信される<br>メール、RPA |

ロボットは器（ハードウェア）、AI は頭脳（ソフトウェア）。体が無くても AI は AI で、
判断しない自動機械は、体があっても AI ではない。

図 1-2　AI とロボット ―― 「身体があるか」「判断しているか」の2軸

似た論点として、**単純な自動化（オートメーション）との区別**もあります。あらかじめ決めた手順を決めたとおりに繰り返すだけの仕組みは、**学習も適応もしない**という点で AI と区別して論じられます。「毎朝7時に自動でメールを送る」仕組みは自動化であって、AI とは呼びません。ただし後で見る「AI の4つのレベル」では、便宜上これをレベル1として並べます。**並べて論じることと、同じものだと言うことは違います。**

## AI の研究 ―― どこから始まり、何を目指してきたか

**「Artificial Intelligence（人工知能）」という言葉が生まれたのは、1956年のダートマス会議（Dartmouth Conference）です。**アメリカのダートマス大学で開かれた研究者の集まりで、**ジョン・マッカーシー**がこの言葉を提案し、そのまま学問分野の名前になりました。**AI の誕生の場として、試験でほぼ確実に問われる固有名詞**です。

この会議には、開催に先立って提案書（研究計画書）が出されていました。そこに書かれていたのは、**「学習をはじめとする知能のあらゆる側面は、原理的に精密に記述でき、それを機械にシミュレートさせられる」**という前提です。**知**

**能は記述できる、記述できるなら機械で再現できる** —— この考え方が、その後の研究の出発点になりました。

会議の場では、**アレン・ニューウェル**と**ハーバート・サイモン**が、世界初の AI プログラムとされる **ロジック・セオリスト（Logic Theorist）** を披露しています。これは数学の定理を自動的に証明するプログラムでした。**対話プログラムでも、画像認識でも、チェスのソフトでもなく、定理証明のプログラムだった**という点が、名前と中身の対応として問われます。

**この会議の位置づけで問われるのは3点**です。**開催年が1956年であること**、**「Artificial Intelligence」という語をジョン・マッカーシーが提案したこと**、そして **ここが第一次AIブームの起点になったこと**。年号を1946年や1966年に、提案者を別の研究者に差し替えた選択肢が並びますので、**「1956年・ダートマス・マッカーシー」の3点セット**で記憶してください。

AI の研究は、この会議以降、大きく2つの方向を行き来してきました。

- **人間が持つ知識やルールを、機械に書き写して動かす方向**（後述するルールベース、エキスパートシステム）
- **人間の脳の仕組みをまねて、データから自分で学ばせる方向**（ニューラルネットワーク、機械学習、ディープラーニング）

第一次・第二次の AI ブームは前者が主役でした。現在の第三次ブームと生成 AI は、後者の系統にあります。**どちらか一方が「正しい AI」なのではなく、時期によって主役が入れ替わってきた**と捉えてください。歴史の詳細は 1-7 で扱います。

## AI にできることを、種類で押さえる

AI の定義は定まらなくても、**いま AI が実際にこなしている仕事**は種類で整理できます。この整理は、後の章で生成AI のサービスを分類するときにも効きま

す。

| 種類 | 何をするか | 身近な例 |
| --- | --- | --- |
| **認識** | 画像・音声・文字を読み取って、それが何かを判定する | 顔認証、音声入力、書類の読み取り |
| **予測** | 過去のデータから、これから起きることや未知の値を見積もる | 需要予測、故障の予知、不正取引の検知 |
| **最適化・判断** | 多数の選択肢の中から、条件に合う組み合わせを選ぶ | 配送ルートの決定、シフトの組み立て |
| **生成** | 学習した内容をもとに、**新しい文章・画像・音声・動画を作り出す** | 文章の作成、画像の生成、要約や翻訳 |

**生成AI が担うのは、いちばん下の「生成」**です。ただし **生成AI が生成しかできないわけではありません。**文章を読んで要点をまとめる（認識に近い働き）、内容を分類する、といった仕事もこなします。逆に、**画像から不良品を見つける検査 AI のように、生成をまったく行わない AI も現役で広く使われています。**「いまの AI ＝生成AI」と置き換えないでください。

## AI が得意なこと、苦手なこと

定義よりも実務で効くのが、この見取り図です。

**得意なこと。大量のデータから統計的な傾向を見つけること、同じ判断を疲れずに大量に繰り返すこと、人間には見つけられない細かな規則性を拾うこと。**

**苦手なこと。学習したデータにない状況へ対応すること、なぜその答えになったかを説明すること、少ない例から本質をつかむこと、**そして **正しさの保証**で

す。1-4 で見るとおり、AI の出力は「最も確からしい候補」であって、正解の保証ではありません。

この得手不得手が、**AI に任せる仕事と、人が判断すべき仕事を分ける基準**になります。第4章で扱う AI の利活用の原則も、根はここにあります。

### AI の研究が扱ってきた分野

「AI の研究」と一口に言っても、扱われてきた分野は複数あります。**どれも AI という傘の下にある別々の研究領域**です。

- **探索・推論** —— 選択肢を展開して答えへの道筋を見つける。第一次AIブームの主役（1-7）
- **知識表現** —— 人間の知識を機械が扱える形で書き表す。第二次AIブームの主役（1-7）
- **機械学習** —— データから規則性を学ぶ。第三次AIブームの主役（1-2、1-3）
- **自然言語処理（NLP）** —— 人間が使う言葉を機械に扱わせる。翻訳・要約・対話が含まれ、**大規模言語モデル（LLM）はこの分野の成果**
- **画像認識・音声認識** —— 画像や音を数値として解析し、何が写っているか・何と言っているかを判定する（1-4）

**この5つは、時代ごとに主役が入れ替わってきただけで、後から出たものが前のものを消したわけではありません。**現在の生成AI は、機械学習と自然言語処理の系統から生まれています。

**では、AI の研究とは結局「何を研究している分野」なのでしょうか。**一言でいえば、**「人間が知能を使ってやっていることを、機械にやらせるにはどうすればよいか」を調べる分野**です。ダートマス会議の提案書にあった前提 ——「**知能**

**のあらゆる側面は精密に記述でき、記述できるなら機械で再現できる**」――が、そのまま分野の問いになっています。

ここで注意したいのは、**AI の研究は「人間の脳をそのまま再現する研究」ではない**という点です。**目指しているのは、仕組みが同じであることではなく、結果として同じことができること**です。飛行機は鳥の羽ばたきを再現していませんが、飛ぶという目的は果たしています。**AI も同じで、人間と同じやり方でなくてよい** ―― ニューラルネットワークが「脳をまねた」と説明されるときも、それは**発想の出どころが脳だという意味であって、脳の再現ではありません**（1-4）。

**「AI の研究とは、人間の脳を完全に再現することを目的とした分野である」という選択肢は誤り**です。あわせて、**AI の研究は1つの技術の研究ではありません。** 上に挙げた5つのように、**扱う対象も手法も異なる複数の分野の集まり**であり、そのため**「AI とは何か」という定義そのものが1つに定まらない**のです。分野の広さと、定義の定まらなさは、つながっています。

### 定義が定まらないもう1つの理由 ―― AI効果

AI の定義がいつまでも定まらない理由として、**AI効果（AI effect）** という現象がよく挙げられます。**ある仕組みの原理が分かってしまうと、人はそれを「これは知能ではない、ただの計算だ」と見なすようになる**という現象です。

かつては文字認識も、将棋のソフトも、経路探索も、AI の代表例として語られていました。ところが仕組みが理解され、身近になると「あんなものは AI ではない」と言われるようになります。**AI と呼ばれる範囲が、常に「まだ解けていないもの」の側へ逃げていく**ため、定義が後から追いかける形になるわけです。

AI効果はシンギュラリティの節（1-8）のキーワードでもありますので、そこで改めて扱います。ここでは「**定義が定まらないことと AI効果はつながっている**」という点だけ押さえてください。

### 「定義が無い」ことの、実務上の意味

定義が定まらないという話は、学術的な余談ではありません。**「AI 搭載」という表示が、何も保証しない**ことを意味するからです。

同じ「AI 搭載」でも、中身は **人が書いた条件分岐だけの製品（1-6 のレベル1に近いもの）** から、**大量のデータで学習させたディープラーニングの製品（レベル4）** まで幅があります。**表示だけでは、どちらなのか区別が付きません。** 導入を検討する側が確かめるべきは「AI かどうか」ではなく、**「何のデータで、何を学習し、どういう条件でどれくらいの精度が出るのか」**です。

**この章でレベル分けや学び方の違いを学ぶ実務的な意味は、まさにここにあります。** 言葉ではなく中身で判断できるようになるための道具立てだ、と考えてください。

### 誤りやすい対比 [1.1]

- **AI の定義は研究者ごとに異なり、合意された唯一の定義はない。**「公式に統一された定義がある」は誤り
- **AI はソフトウェアの技術で、身体を必要としない。ロボットは駆動部を持つ物理的な機械。**「AI とロボットは同じものの呼び分け」は誤り
- **AI を搭載していないロボットも、身体を持たない AI も、どちらも存在する**
- **「Artificial Intelligence」という語が生まれたのは1956年のダートマス会議。提案者はジョン・マッカーシー**
- **AI ⊃ 機械学習 ⊃ ディープラーニング。**この向きを逆にした記述は誤り
- **AI の研究には探索・推論、知識表現、機械学習、自然言語処理（NLP）、画像認識・音声認識といった分野がある。**後から出た分野が前の分野を消したわけではない

- **生成AI は AI の一種で、AI の外にある別ジャンルではない。**逆に、生成を行わない AI（画像検査など）も広く使われている
- **AI効果は「原理が分かると知能と見なされなくなる」現象。**AI の性能が上がる現象のことではない

## 1-2 知能をもたらす2つの仕組み —— ルールベースと機械学習 [1.2]

### AI が「賢く見える」やり方は、大きく2つしかない

AI が知的に振る舞うための仕組みは、大きく **2つの系統** に分かれます。この2分割が、第1章でいちばん重要な骨組みです。

| 仕組み | 誰が規則を作るか | ひとことで言うと |
|---|---|---|
| **ルールベース** | **人間**が、あらかじめ規則を書く | 「もし〜ならば〜せよ」を書き並べる |
| **機械学習** | **機械**が、データから規則を見つける | 大量の例を見せて、規則を自分で作らせる |

**「規則を書くのが人か、機械か」** —— 判別の軸はこれ1本です。設問で迷ったら、まずこの問いに戻ってください。

### ルールベースとは

**ルールベース（rule-based）** とは、**人間があらかじめ「こういう条件のときは、こう処理する」という規則（ルール）を書いておき、AI がその規則に従って判断する仕組み**です。「もし体温が37.5度以上なら、発熱と判定する」「もし件名に特定の語が含まれていたら、迷惑メールに分類する」といった **if〜then 形式の条件分岐の集まり**だと考えてください。

ルールベースの長所は3つあります。

1. **判断の根拠が完全に説明できる。** どのルールが働いてその結論になったかを、人間が1行ずつ追えます。医療・金融・法務のように「なぜその判断か」を説明する義務がある領域では、いまでも重要な性質です。
2. **少ないデータで動く。** 学習用のデータを集める必要がありません。専門家が規則を書けば、その日から動きます。
3. **意図しない答えを出しにくい。** 書いていない振る舞いはしません。

短所も明確です。

1. **規則を書く手間が膨大になる。** 現実の問題は例外だらけで、規則の数が爆発的に増えます。
2. **保守が難しくなる。** ルールが増えるほど、互いに矛盾したり、どれを直せばよいか分からなくなったりします。
3. **書かれていない状況に対応できない。** 想定外の入力が来ると、何も判断できないか、見当違いの判断をします。

この短所が現実の壁になったのが、1-7 で扱う **第二次AIブームのエキスパートシステム** です。専門家の知識をルールとして書き写す方式は、書き写す作業そのものが終わらないという理由で行き詰まりました。

**なお、ルールベースは「古い方式だから使われていない」わけではありません。** 現在も、条件が明確で説明責任が重い場面では現役です。生成AI のサービスでも、**入力に危険な語が含まれていないかをルールで弾いてから、モデルに渡す**といった組み合わせ方が普通に使われます。「ルールベースは完全に置き換えられた」という趣旨の選択肢は誤りです。

## 機械学習とは

**機械学習（machine learning）**とは、**人間が規則を書く代わりに、大量のデータをコンピュータに与え、そこに含まれる規則性やパターンをコンピュータ自身に見つけさせる仕組み**です。

迷惑メールの判別で考えると違いがはっきりします。ルールベースなら「件名に『当選』が入っていたら迷惑メール」と人が書きます。機械学習なら、**迷惑メール1万通と通常メール1万通を見せて、「どういう特徴を持つメールが迷惑メールなのか」を機械に見つけさせます。**人は「当選という語に注目せよ」とは教えません。データの中から、機械が自分で手がかりを拾います。

機械学習の長所は、**人間が言葉にできない規則も扱える**ことと、**新しいデータを追加して学習し直せば性能を改善できる**ことです。短所は、**大量のデータが必要**なこと、**なぜその判断をしたのかを人間が説明しにくい**こと（この性質は「ブラックボックス問題」と呼ばれます）、そして **データに偏りがあれば、その偏りごと学んでしまう**ことです。

## 2つの仕組みは、どちらかを選ぶものではない

「ルールベースは古く、機械学習が新しい」と単純化すると設問を落とします。**実務では両方を組み合わせるのが普通**です。

| 場面 | 向いている仕組み | 理由 |
|---|---|---|
| 法令や社内規程で条件が明文化されている審査 | **ルールベース** | 規則がすでに文章として存在し、判断根拠の説明が求められる |
| 手書き文字の読み取り、写真からの製品検査 | **機械学習** | 「どういう形なら正しいか」を人が言葉で書き切れない |

| 場面 | 向いている仕組み | 理由 |
| --- | --- | --- |
| 与信審査や不正検知 | **両方の組み合わせ** | 機械学習で疑わしい取引を絞り込み、最終判定は明文化されたルールで行う |

**分かれ目は「規則を人が言葉にできるか」です。**言葉にできるならルールベースのほうが速く、安く、説明できます。言葉にできないなら機械学習の出番になります。生成AIのサービスでも、**入力を機械学習のモデルに渡す前に、危険な語をルールで弾く**といった二段構えが取られます。

### 学習済みモデル —— 学習の「成果物」

機械学習の流れは、**学習フェーズ**と**推論フェーズ**の2段階に分かれます。

- **学習フェーズ**：大量のデータを読み込ませ、規則性を数値の形（後述する「重み」）に落とし込む。時間も計算資源も、ここで大量に使う
- **推論フェーズ**：学習の結果を使って、新しいデータに対して答えを出す。1件ごとの処理は速い

この **学習フェーズの成果物として得られる、規則性を保持したプログラムのかたまり**を **学習済みモデル（trained model）** と呼びます。学習済みモデルは、**学習に使った元データそのものを内部に丸ごと保管しているわけではありません。**データから抽出された規則性が、大量の数値として保存されている状態です。

学習済みモデルという考え方が重要なのは、**一度作れば繰り返し使えて、他人に配布でき、別の目的にも流用できる**からです。生成AIで使われる大規模言語モデル（LLM）も、膨大なテキストで学習させた学習済みモデルにほかなりません。この「流用」の話が、1-5 で扱う **転移学習** につながります。

学習済みモデルには、実務上の注意点もあります。**学習に使ったデータの時点で世界が止まっている**ことです。学習後に起きた出来事や、新しく登場した製品のことは知りません。**モデルが「知らないこと」を知らないまま、もっともらしく答えてしまう**問題は、生成AI の章で扱うハルシネーションにつながります。また、**学習データに偏りがあれば、その偏りごと学習済みモデルに固定されます。** 学習済みモデルは中立な道具ではなく、**与えられたデータの写し**である、という見方をしておいてください。

### 機械学習の考え方 —— 統計と何が違うのか

機械学習の中身は、突き詰めると**「入力と出力の対応関係を、データに合うように調整していく」**ことです。数式は使わずに言えば、**たくさんのつまみが付いた機械があり、出てくる答えが正解に近づくように、つまみを少しずつ回していく**イメージになります。このつまみが後述の **重み** です。

ここで押さえておきたいのが、**機械学習が目指しているのは「学習に使ったデータを正確に言い当てること」ではない**という点です。目指すのは **未知のデータに対して正しく答えられること** で、これを **汎化（はんか）** と呼びます。学習データだけに合わせ込みすぎて未知のデータで外すようになる失敗が、1-5 で扱う **過学習（オーバーフィッティング）** です。

**「学習データでの正答率が100%だから良いモデルだ」は誤り**です。むしろ、それは過学習の疑いを示す危険信号として読みます。

### 機械学習を業務に持ち込むときの流れ

仕組みの理解を実務につなぐために、**機械学習を使うときの一連の流れ**も押さえておきます。

1. **解きたい問題を決める** —— 何を当てたいのか（カテゴリか、数値か）、どれくらいの精度なら役に立つのかを先に決める

2. **データを集めて整える** —— 欠けている値を埋める、表記を揃える、正解ラベルを付ける。**工数の大半はここに掛かります**
3. **学習させる** —— 手法を選び、重みを調整させる
4. **取り分けたデータで評価する** —— 学習に使っていないデータで実力を測る
5. **業務に組み込み、様子を見ながら学習し直す** —— 世の中が変われば、モデルの前提も古くなります

**この5段階のうち、AI らしい処理は3番だけ**です。**うまくいくかどうかは、2番のデータの質でほぼ決まります。**「良いモデルを選べば精度が出る」ではなく、「良いデータがあって初めて手法の差が出る」というのが実際のところです。

### ノーフリーランチ定理 —— 万能の手法は存在しない

機械学習の手法は数多くあり、後で見るように学習の枠組みも複数あります。では「いちばん強い手法」はどれか、という問いに答えるのが **ノーフリーランチ定理（No Free Lunch Theorem）** です。

**ノーフリーランチ定理とは、「あらゆる問題に対して他より優れた、万能の手法は存在しない」ことを示した定理**です。ある手法が特定の種類の問題で優れた成績を出すなら、別の種類の問題では他の手法に劣る —— そういう釣り合いが必ず生じる、という主張です。名前は「ただ飯はない（無料の昼食はない）」という英語の慣用句から来ています。

実務上の意味は明快です。**問題の性質・データの量・説明のしやすさ・かかる費用に応じて、そのつど手法を選ぶ必要がある。**「とりあえずディープラーニングを使えば最善」という発想は、この定理に反します。生成AI が広く使われる今でも、**表形式の少量データを扱う予測では、より単純な手法のほうが精度も説明性も高いことが珍しくありません。**

**誤りやすいのは、この定理を「どの手法も性能は同じ」という意味に読むこと**です。そうではありません。**個別の問題で見れば手法の優劣ははっきり付きます。**「すべての問題を平均すれば差がなくなる＝万能な1つは選べない」という主張です。また、**データを増やせば万能になる、という定理でもありません。**

### 誤りやすい対比 [1.2]

- **ルールベース＝人が規則を書く／機械学習＝機械がデータから規則を見つける。**分かれ目は「規則を作るのが誰か」
- **ルールベースは判断の根拠を説明しやすいが、規則の作成と保守の負担が大きい。**「データが大量に必要」なのは機械学習のほう
- **学習済みモデルは、学習で得た規則性を保持したもの。**学習に使った元データそのものを内部に保管しているわけではない
- **ノーフリーランチ定理＝あらゆる問題に万能な手法は存在しない。**「どの手法も性能が同じ」という意味ではない

## 1-3 機械学習の手法 —— 4つの学び方 [1.2]

### 「何を与えるか」で4つに分かれる

機械学習の手法は、**学習のときに機械へ何を与えるか**で分類します。この軸で切ると、次の4つになります。**「どんな計算をするか」ではなく「何を与えるか」で分かれる**、という点がそのまま設問になります。

| 学び方 | 与えるもの | 解く問題の例 |
| --- | --- | --- |
| **教師あり学習** | 入力と**正解ラベル**の組 | 分類（カテゴリを当てる）／回帰（数値を当てる） |

| 学び方 | 与えるもの | 解く問題の例 |
| --- | --- | --- |
| **教師なし学習** | **正解ラベルのないデータだけ** | クラスタリング（似たもの同士をまとめる）／次元削減 |
| **強化学習** | 行動に対する**報酬** | 試行錯誤で、報酬の合計が最大になる行動の方針を見つける |
| **半教師あり学習** | **少量のラベル付きデータ＋大量のラベルなしデータ** | ラベル付けの手間を減らしながら分類の精度を上げる |

「正解ラベル」とは、**そのデータの答えを人間が書き添えたもの**です。猫の写真に「猫」と書き添える、過去の物件データに「実際の成約価格」を書き添える——この書き添え作業を **アノテーション** と呼び、人手と費用がかかります。**ラベルが手に入るかどうかで、選べる学び方が決まる**というのが実務の実感です。

## 教師あり学習

**教師あり学習（supervised learning）**は、**入力と正解ラベルの組を大量に与え、入力から正解を導く対応関係を学ばせる**方式です。「問題集と解答」をセットで渡して勉強させるイメージですが、実際に起きているのは、**予測と正解のズレ（誤差）が小さくなるように内部の数値を調整していく**ことです。

解く問題は大きく2種類に分かれます。

- **分類（classification）**：**あらかじめ決まったカテゴリのどれに当たるか**を当てる。迷惑メールか通常メールか、この画像は犬か猫か、この申込は不正か正常か
- **回帰（regression）**：**連続する数値**を当てる。来月の売上、住宅の価格、明日の気温

教師あり学習が使われる場面を、分類と回帰の両方から並べておきます。**片方の例だけで覚えると、もう片方が出たときに迷います。**

| 用途 | 分類か回帰か | 与える正解ラベル |
|---|---|---|
| 迷惑メールの判別 | 分類 | 「迷惑」「通常」の区分 |
| 医療画像から病変の有無を判定 | 分類 | 医師が付けた診断 |
| 中古車の価格見積もり | 回帰 | 実際の取引価格 |
| 来月の来店客数の予測 | 回帰 | 過去の実績値 |

**分類と回帰の違いは、出力がカテゴリか数値か**です。「気温を当てるのは分類である」といった選択肢は誤りになります。

**分類は、さらに2つに分かれます。**どちらも教師あり学習の分類ですが、**選択肢の数が違います。**

- **二値分類** —— **答えが2つのどちらか。**迷惑メールか通常メールか、不良品か良品か、退会するかしないか
- **多クラス分類** —— **答えが3つ以上の中のどれか。**手書き数字が0〜9のどれか、問い合わせ文が「請求」「故障」「解約」のどれか

**「分類＝2択」と覚えると誤ります。**カテゴリが3つ以上でも分類です。**分かれ目は選択肢の数ではなく、答えがカテゴリか数値か**です。

もう1つ、実務で効く理解を添えます。**分類の答えは、多くの場合「このカテゴリだ」と言い切っているのではなく、「このカテゴリである見込みがどれくらい高いか」という度合いで出てきます。**迷惑メールの判定であれば、「迷惑である度合いが◯割」という値が出て、**どこから上を迷惑と扱うかの境目は、使う側が決めます。**境目を厳しくすれば見逃しは減りますが、通常のメールを迷惑と誤

って弾く数が増えます。**逆に境目を緩めれば、誤って弾く数は減りますが、見逃しは増えます。境目の置き方は精度の問題ではなく、どちらの間違いをより避けたいかという運用の判断**です —— ここは、AIの判定結果をそのまま鵜呑みにしない、という実務の考え方につながります。

なお、**回帰と分類は、同じデータから両方作れることがあります。**中古車のデータからは、「価格そのものを当てる」（回帰）ことも、「50万円以上か未満かを当てる」（分類）こともできます。**問題の立て方が変われば、同じデータでも分類にも回帰にもなる** ——「このデータは回帰用だ」と決まっているわけではありません。

教師あり学習は、**成果が測りやすく、実務での利用がいちばん多い**方式です。一方で、**正解ラベルを人手で用意する必要がある**ことが最大の制約になります。

### 教師なし学習 —— クラスタリングと次元削減

**教師なし学習（unsupervised learning）**は、**正解ラベルを与えず、データだけを渡して、そこに潜む構造やパターンを見つけさせる**方式です。「答えのない資料の束を渡して、似たもの同士に整理させる」形になります。

代表的な使い道が2つあり、**両方とも教師なし学習の中に位置づけられる**ことがそのまま出題されます。

**クラスタリング（clustering）**は、**データを似たもの同士のグループ（クラスタ）に自動的に分ける**手法です。顧客の購買履歴から「よく似た買い方をする客層」をいくつかのまとまりに分ける、といった使い方をします。ここで重要なのは、**そのグループが何を意味するかを機械は教えてくれない**という点です。「グループA」「グループB」と分かれるだけで、「これは節約志向の層だ」と意味づけるのは人間の仕事です。

**クラスタリングと分類の違い**は頻出です。**分類（教師あり）は、あらかじめ決まったカテゴリに割り当てる。クラスタリング（教師なし）は、カテゴリ自体をデータから作り出す。**「あらかじめ決めた区分があるかどうか」で切り分けてください。

**次元削減（dimensionality reduction）**は、**データが持つ特徴の数（次元）を、重要な情報をできるだけ保ったまま減らす**手法です。1人の顧客について100項目のデータがあるとき、それを2〜3項目の合成された指標にまとめ直す、というような処理を指します。

次元削減の目的は3つあります。**計算を軽くすること、グラフに描いて人間が目で見られるようにすること**、そして **特徴が多すぎることで起きる精度の低下を防ぐこと**です。3つめは直感に反しますが、**特徴量は多ければ多いほど良いわけではありません。**項目数が増えるほどデータは互いに離れていき、規則性を見つけるのに必要なデータ量が跳ね上がるためです。

教師なし学習には、**異常検知**という使い道もあります。**正常なデータだけを大量に学習させ、そこから大きく外れたものを異常として検出する**やり方です。不良品の検出、機械の故障の予兆、不正アクセスの検知などで使われます。**「異常のラベルが付いたデータはほとんど集まらない」**という現実に対する答えになっている点が要点です。ただし、**異常検知は教師なし学習だけで行うものではありません。**十分な異常データが集まる場面では、教師あり学習の分類問題として解くほうが精度が出ます。

**クラスタリングと次元削減は、どちらも教師なし学習の代表例**です。「次元削減は教師あり学習の一種である」「クラスタリングには正解ラベルが必要である」という選択肢は誤りになります。

**2つは、目的がまったく違います。**並べて押さえてください。

| | クラスタリング | 次元削減 |
|---|---|---|
| **何を減らす／まとめるのか** | **データの件数**をグループにまとめる（縦方向） | **項目（特徴）の数**を減らす（横方向） |
| **何のためにあるか** | **手元のデータに、どんなまとまりが潜んでいるかを知るため。**人間が事前に区分を用意できない場面で使う | **扱いやすくするため。**計算を軽くする／グラフに描いて目で見る／項目が多すぎることによる精度低下を防ぐ |
| **結果として得られるもの** | 「どのデータがどのグループに属するか」 | 「元の情報をできるだけ保った、少ない項目のデータ」 |
| **人間に残る仕事** | **各グループが何を意味するかの解釈** | 減らした後の項目が何を表すかの解釈 |

**覚え方は「縦か横か」です。**顧客100人×100項目の表があるとき、**クラスタリングは100人を数グループに分け、次元削減は100項目を数項目にまとめます。**どちらも「減らす」という言葉で説明されるので混ざりますが、**減らしている方向が違う**と押さえれば取り違えません。

**この2つは、組み合わせて使われることもあります。**項目が多すぎるデータは、そのままではグループ分けもうまくいきません。**先に次元削減で項目を絞り、それからクラスタリングにかける**——実務ではよくある順序です。**「どちらか一方を選ぶもの」ではない**という点も、あわせて押さえておいてください。

### 強化学習

**強化学習（reinforcement learning）**は、**正解を与える代わりに、行動の結果に対して報酬（点数）を与え、報酬の合計が最大になる行動の方針を試行錯誤で身につけさせる**方式です。

登場する要素は、**行動する主体（エージェント）**、**その相手となる環境**、**エージェントが選ぶ行動**、**行動の結果として変わる状態**、そして **良し悪しを伝える報

**酬** です。ゲームで高得点を取る、ロボットに歩き方を覚えさせる、広告の出し分けを最適化する —— いずれも「正解の1手」を人が示せない代わりに、結果の良し悪しは点数で示せる問題です。

強化学習の特徴は、**1回ごとの報酬ではなく、将来にわたる報酬の合計を最大にしようとする**点です。目先の点数を捨てて後で大きく取る、という判断ができるのはこのためです。囲碁の AI が、その場では損に見える手を打つことがあるのは、この性質の表れです。

強化学習が向くのは、**「1手ごとの正解は誰にも言えないが、最終的な結果の良し悪しは点数で表せる」問題**です。囲碁や将棋、ロボットの歩行制御、在庫の発注量の調整、広告の出し分けなどが当てはまります。**逆に、正解がはっきり分かっていて手元にラベルもある問題に強化学習を使う理由はありません。**そこは教師あり学習の領域です。

強化学習には固有の難しさもあります。**試行錯誤の回数が大量に要る**こと、**報酬の設計を誤ると、こちらの意図しない抜け道を学習してしまう**ことです。「掃除した面積」だけを報酬にすると、ゴミを撒いてから掃除する振る舞いを覚えてしまう —— そういう失敗が実際に起きます。**報酬をどう設計するかが人間の仕事として残る**点は、押さえておく価値があります。

**教師あり学習との違い**を明確にしておきます。**教師あり学習は「この入力に対する正解はこれだ」と1件ずつ教えます。強化学習は正解を教えず、「今の一連の行動の結果は何点だった」とだけ伝えます。** 与えているのが**正解ラベルか報酬か**、ここが分かれ目です。

違いを、表でもう一段開いておきます。**ここは毎回問われるところです。**

| | 教師あり学習 | 強化学習 |
|---|---|---|
| **人が与えるもの** | **1件ごとの正解ラベル** | **結果に対する報酬（点数）だけ** |
| **正解は示されるか** | **示される。**「この入力の答えはこれ」と直接教える | **示されない。**良かったか悪かったかしか返らない |
| **データの集め方** | **あらかじめ集めておく**（過去のデータに人が答えを書き添える） | **自分で動いてみて集める**（試行錯誤しながら経験を積む） |
| **1回の判断か、続く判断か** | 基本は**1件ごとに独立した判断** | **一連の行動のつながり**を評価する（目先を捨てて後で取る判断ができる） |
| **向く場面** | 過去の答えが手元にある問題 | **1手ごとの正解を誰も言えないが、結果の良し悪しは点数にできる問題** |

**「正解が与えられるのではなく、結果の良し悪しだけが返る」**—— 強化学習を一言で説明するなら、ここです。**「強化学習も正解ラベルを使う」という選択肢は誤り**になります。

## 半教師あり学習

**半教師あり学習（semi-supervised learning）**は、**少量のラベル付きデータと、大量のラベルなしデータを組み合わせて学習する**方式です。

なぜこの方式が要るのかというと、現実には**「データはいくらでもあるが、ラベルを付ける人手と費用がない」**という状況が多いからです。医療画像であれば、画像自体は病院に大量にありますが、1枚ずつ医師が診断名を書き添えるのは大変です。そこで、**少数のラベル付きデータでいったんモデルを作り、**

**そのモデルにラベルなしデータを判定させ、確信度の高い判定を仮の正解として学習に加える**、といったやり方を取ります。

半教師あり学習が使われる場面は、医療画像のほかにもあります。**膨大な音声データのうち、書き起こしができているのはごく一部**という音声認識の場面、**大量の問い合わせ文のうち、分類済みは数百件だけ**というカスタマーサポートの場面 —— いずれも「データは多いがラベルが少ない」という同じ形をしています。

**混同しやすいのは教師なし学習との違い**です。**半教師あり学習は、人手のラベルを「少しは」使います。ラベルを一切使わないのは教師なし学習です。**また、**ラベル付きとラベルなしを同数そろえる必要はありません。**ラベルなしのほうが圧倒的に多いのが前提です。画像に限らず、テキストや音声でも使われます。

### 4つの学び方を1本の軸で整理する

覚え方は「**何を与えるか**」の1本で足ります。

- **正解ラベルを全部与える** → 教師あり学習
- **正解ラベルを一切与えない** → 教師なし学習
- **正解ラベルを少しだけ与える** → 半教師あり学習
- **正解ラベルの代わりに報酬を与える** → 強化学習

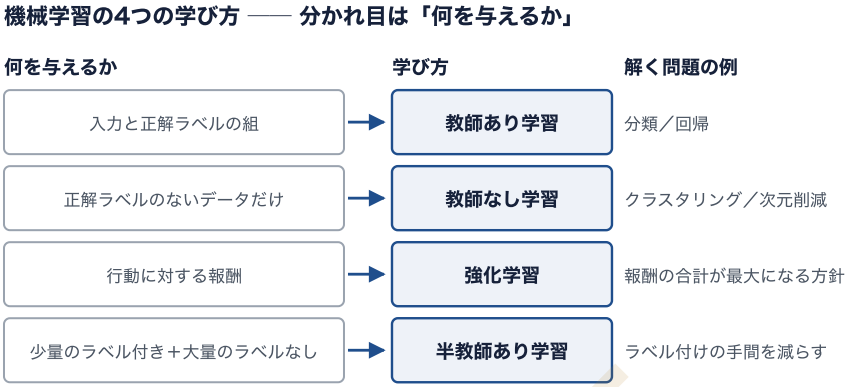



図 1-3　機械学習の4つの学び方 —— 分かれ目は「何を与えるか」

そして、ここまで見た手法のどれが最善かは問題によって変わります —— これが 1-2 で扱った **ノーフリーランチ定理** の実際の意味です。

## 誤りやすい対比 [1.2]

- **教師あり学習＝入力と正解ラベルの組／教師なし学習＝ラベルなしのデータだけ／強化学習＝報酬／半教師あり学習＝少量のラベル＋大量のラベルなし**
- **教師なし学習は正解ラベルを与えない。ラベルを一部だけ与えるのは半教師あり学習**
- **クラスタリングも次元削減も、教師なし学習の代表例**
- **分類（教師あり）はあらかじめ決めたカテゴリに割り当てる／クラスタリング（教師なし）はカテゴリ自体をデータから作る**
- **分類＝カテゴリを当てる／回帰＝連続した数値を当てる**
- **強化学習が与えるのは報酬であって正解ではない**

## 1-4 人間の脳とニューラルネットワーク、そして画像認識 [1.2]

### 人間の脳 —— ニューロンとシナプス

いま主流になっている AI の中身は、**人間の脳の情報伝達の仕組みをまねた構造**です。まず、まねの元になった脳の側を確認します。

**ニューロン（神経細胞）** は、**脳を構成する細胞で、電気的な信号を受け取り、次の細胞へ渡す働きを持つもの**です。人間の脳にはおよそ千数百億個あるとされます。1個のニューロンは、他の多くのニューロンから信号を受け取り、**受け取った信号の合計がある大きさを超えたときにだけ、自分も信号を出します。** この「超えたら出す、超えなければ出さない」という振る舞いを **発火** と呼びます。

**シナプス** は、**ニューロンとニューロンのつなぎ目にあたる部分で、信号を受け渡す接続点**です。ここで押さえるべきは、**シナプスの伝わりやすさは一定ではなく、使われ方によって強くなったり弱くなったりする**という点です。よく一緒に働く経路は伝わりやすくなり、使われない経路は弱まっていきます。**人間が学習して物事を覚えるとき、脳の中で起きているのは、このシナプスの伝わりやすさの調整だ**と考えられています。

**ニューロン＝信号を処理する細胞そのもの／シナプス＝細胞と細胞のつなぎ目。**この対応は、そのまま次の人工ニューロンの説明に写されます。

### 人工ニューロン（ノード）と、情報の重みづけ

**人工ニューロン（artificial neuron）** は、**ニューロンの働きを数値の計算で模したもの**で、**ノード（node）** とも呼ばれます。1つの人工ニューロンがすることは3ステップだけです。

1. 複数の入力を受け取る

2. **それぞれの入力に「重み」を掛けて合計する**
3. 合計の大きさに応じて、次へ渡す値を決める

3ステップめの「合計の大きさに応じて次へ渡す値を決める」部分を、もう少しだけ開いておきます。**脳のニューロンが「合計がある大きさを超えたら発火する」と述べたのと同じで、人工ニューロンにも、合計をそのまま流さずに変換する仕組みがあります。**この変換によって、**ネットワーク全体が「単純な足し算の繰り返し」で終わらず、複雑な対応関係を表せるようになります。**数式に立ち入る必要はありませんが、**「合計をそのまま次へ渡しているのではない」**という点だけは押さえてください。

ここに出てくる **重み（weight）** が、この試験の核心にある言葉です。**重みとは、それぞれの入力をどれくらい重視するかを表す数値**であり、脳でいえば **シナプスの伝わりやすさ**にあたります。**重みが大きい入力ほど結果に強く影響し、小さい入力はほとんど影響しません。**

この「入力ごとに重要度を変えて足し合わせる」処理を **情報の重みづけ** と呼びます。中古車の価格を見積もる場面を思い浮かべてください。走行距離・年式・車種・車体の傷 —— どれも価格に関わりますが、**効き方の強さは同じではありません。**走行距離を重く、車体の細かい傷を軽く見る、といった具合に**入力ごとに重要度の配分を変えること**が情報の重みづけです。

そして、**この重みの値は人間が決めるものではありません。学習によって機械が自分で調整します。**機械学習が「学習している」と言えるのは、まさにこの重みの調整が起きているからです。**「学習済みモデル」の中身は、突き詰めればこの重みの集まり**です。

## ニューラルネットワーク

**ニューラルネットワーク（neural network）**は、**人工ニューロン（ノード）を多数つなぎ合わせ、層状に並べた仕組み**です。基本の形は3種類の層でできています。

- **入力層**：データを受け取る層。画像なら画素の値、テキストなら数値化された単語が入る
- **中間層（隠れ層）**：入力層と出力層の間にある層。ここで情報が段階的に加工される
- **出力層**：最終的な答えを出す層。「犬である確率」「猫である確率」といった形で出る

情報は入力層から出力層へ向かって流れ、**層を通るたびに、重みを掛けて足し合わせる処理が繰り返されます。**1つ1つのノードがしていることは単純ですが、それを大量につなぐと、非常に複雑な対応関係を表せるようになります。

**中間層が1層だけの浅いネットワークにも限界があります。**そこで層を増やしたのが、次のディープラーニングです。

## ディープラーニング（深層学習）

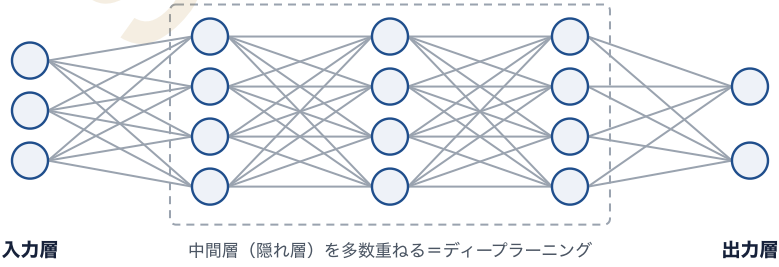
**ディープラーニング（深層学習、deep learning）**とは、**中間層を多数重ねた、層の深いニューラルネットワークを使う機械学習の手法**です。

「深い」以外に何が変わるのかというと、**扱える情報の抽象度が段階的に上がる**点です。画像を例にすると、**入力に近い浅い層では、線や明暗の境目といった単純な特徴**をとらえ、**深い層へ進むほど、目・鼻・耳といった部品、さらには顔全体という具合に、より抽象度の高い特徴**をとらえるようになります。

**層の数え方**にも触れておきます。「中間層が多数ある」といっても、何層からが「深い」という明確な線引きはありません。**中間層が1層だけのものは一般に浅いネットワークとして扱われ、中間層を何層も重ねたものがディープラーニングと呼ばれます。** 現在の大規模なモデルでは、中間層が数十層から百層を超えることも珍しくありません。**「層が深いほど常に高性能」ではない**点も併せて押さえてください。層を深くするほど学習は難しくなり、必要なデータ量も計算量も増えます。深さを増やすこと自体が目的ではありません。

ここが決定的に重要なところです。**従来の機械学習では、「どこに注目すべきか」を人間が設計して機械に渡していました。** 画像なら「輪郭の形」「色の分布」といった見どころを人が定義し、その値を計算して渡す。この人が設計する見どころを **特徴量** と呼びます（1-6 で改めて扱います）。**ディープラーニングは、この特徴量そのものをデータから自分で獲得します。**「輪郭を見ろ」と教えなくても、大量の画像を見るうちに、輪郭に相当する見方を層の中に作り上げてしまう。**人が特徴量を設計する必要がなくなったこと**が、ディープラーニングが起こした最大の変化です。

**ニューラルネットワークの層構造とディープラーニング**




層を通るたびに、重みを掛けて足し合わせる処理が繰り返される。
浅い層は線や明暗の境目、深い層は目・鼻・耳、さらに顔全体をとらえる。

図 1-4　ニューラルネットワークの層構造とディープラーニング

そのぶん代償もあります。**大量のデータと、大きな計算能力が必要**になり、**なぜその答えを出したのかを人間が説明しにくくなります。**ディープラーニングは万能ではなく、少量のデータで説明性が求められる場面では、より単純な手法のほうが適します —— ここでもノーフリーランチ定理が効いています。

### ニューラルネットワークは画像だけのものではない

ニューラルネットワークの応用先は画像に限りません。**扱うデータの性質に合わせて、いくつかの型が使い分けられてきました。**

- **畳み込みニューラルネットワーク（CNN）** —— **画像のように、近い場所どうしに強い関係があるデータ**に向く。小さな窓をずらしながら特徴を拾う
- **回帰型ニューラルネットワーク（RNN）** —— **文章や音声、時系列のように、順番に意味があるデータ**に向く。前の内容を引き継ぎながら順に処理する。**長い並びでは前の内容を保ちにくい**という弱点があり、それを補うために **長・短期記憶（LSTM）** といった改良が作られた
- **Transformer** —— 現在の大規模言語モデル（LLM）や画像生成の中心にある型。**並びの中のどこに注目すべきかを、離れた位置どうしでも直接とらえられる**のが特徴で、文章にも画像にも使われる

**「ニューラルネットワーク＝画像認識」ではありません。**詳細は第2章で扱いますが、**データの性質に応じて型が選ばれる**という見方だけ、ここで持っておいてください。

### AI が画像を認識する仕組み

「AI が画像を見て猫だと分かる」とき、実際には何が起きているのか。順を追うと4段階です。

**第1段階：画像を数値の並びに変える。**コンピュータにとって画像は絵ではなく、**画素（ピクセル）ごとの明るさや色を表す数値の集まり**です。カラー画像であれば、1つの画素につき赤・緑・青の3つの数値が並びます。**AI が扱っているのは最初から最後まで数値であって、人間のように「見て」いるわけではありません。**

**第2段階：近い場所どうしをまとめて、部分的な特徴を取り出す。**画像は、隣り合う画素どうしに強い関係があります。そこで、**画像の小さな範囲を1つのまとまりとして走査し、そこに線があるか、明暗の境目があるかといった手がかりを取り出します。**画像認識でよく使われる **畳み込みニューラルネットワーク（CNN）**は、この「小さな窓をずらしながら特徴を拾う」処理を組み込んだニューラルネットワークです。

**第3段階：層を重ねて、抽象度を上げる。**浅い層で拾った線や境目が、次の層で「丸み」「とがった形」といったまとまりになり、さらに深い層で「目らしきもの」「耳らしきもの」になり、最後に「猫の顔らしい配置」へとまとまっていきます。

**第4段階：出力層で確率として答える。**出力層は「猫である確率92%、犬である確率5%……」という形で答えを出します。**AI は「猫だ」と断定しているのではなく、最も確率が高い候補を提示している**にすぎません。この性質は、後の章で扱う生成AI の誤りの問題ともつながります。

**画像認識が実用化された影響は広い範囲に及んでいます。**顔認証による本人確認、製造ラインでの外観検査、医療画像の読影支援、自動運転での標識や歩行者の検出、書類の文字読み取り —— いずれも同じ仕組みの応用です。同時に、**顔認証のように、精度が上がったからこそプライバシーや差別の問題が生じる**分野でもあります。技術の説明と、社会的な論点が結びつく典型例として、第4章につながります。

ここで1行だけ添えます。**画像に強いのは畳み込みニューラルネットワーク（CNN）ですが、それが唯一の方式ではありません。**前項で見たとおり、文章や音声には回帰型ニューラルネットワーク（RNN）の系統が使われ、現在の生成AIの中心にあるTransformerは画像にも文章にも使われます。**方式は、扱うデータの性質で選ばれます。**

### 誤りやすい対比 [1.2]

- **ニューロン＝信号を処理する神経細胞／シナプス＝ニューロン同士のつなぎ目。**重みが対応するのはシナプスの伝わりやすさ
- **人工ニューロン（ノード）は、入力に重みを掛けて足し合わせる計算単位。**脳細胞を物理的に再現したものではない
- **重みは学習によって機械が調整する。**人間が1つずつ設定するものではない
- **ニューラルネットワーク＝ノードを層状につないだ仕組み／ディープラーニング＝その中間層を多数重ねたもの。**別系統の技術ではなく、後者は前者の一種
- **ディープラーニングの最大の特徴は、特徴量を人が設計せずデータから獲得すること**
- **コンピュータは画像を数値の集まりとして扱う。**出力は断定ではなく確率

## 1-5 AIが自ら学習して改善される仕組みと、過学習・転移学習 [1.2]

### AIが自ら学習して改善される仕組み

ここまでで「重みを調整することが学習だ」と述べました。では、**その調整はどうやって進むのか。**手順は3つの繰り返しです。

1. **予測する。**いまの重みのまま、データを入力して答えを出す
2. **正解と比べて、ズレを測る。**出した答えと正解ラベルの差 —— これを **誤差** と呼びます
3. **誤差が小さくなる向きへ、重みを少しずつ直す。**出力に近い側から入力に近い側へ、各ノードの重みが誤差にどれだけ効いていたかを遡って割り出し、その分だけ調整します

この1〜3を、大量のデータで何万回・何億回と繰り返します。**1回ごとの修正はごくわずかで、少しずつ正解に近づけていく**のが要点です。一度に大きく直そうとすると、行き過ぎて安定しません。

**AI が「自ら学習して改善される」というのは、この繰り返しのこと**であって、AI が自分の意思で勉強しているわけではありません。**誤差という「どれだけ外したか」の情報だけを頼りに、重みという数値が自動的に更新されていく** —— 中身はそれだけです。

大切な補足を1つ。**学習が終わって世に出たあとのモデルは、使っているだけでは勝手に賢くなりません。**学習フェーズと推論フェーズは分かれており、性能を上げるには、新しいデータであらためて学習をやり直す（あるいは追加で学習させる）必要があります。「AI はサービスを使えば使うほど自動的に自己改善する」という趣旨の選択肢は、一般には誤りとして扱われます。

### 学習の成果をどう測るか —— データを分けて使う

過学習の話に入る前に、**学習に使ったデータで成績を測ってはいけない**という原則を押さえます。

機械学習では、手元のデータを **学習に使う分**と、**成績を測るために取り分けておく分**に分けます。取り分けた分は学習中には一切見せず、**モデルにとって**

**初見のデータとして使います。**こうしないと、**「覚えた問題を解かせて満点だった」だけの評価**になり、実際に使ったときの性能が分かりません。

分け方は用途で3つに整理されます。

- **学習（訓練）に使うデータ** —— 重みを調整するために読ませる分
- **調整の確認に使うデータ** —— 学習の途中で成績を確かめ、学習を止める時期や設定を決めるために使う分
- **最終評価に使うデータ** —— 最後に一度だけ、実力を測るために使う分

**この分割があって初めて、次の過学習を見つけられます。**

### 「学習するとき」と「使うとき」を分ける

もう1つ、ここで固めておきたい区別があります。**AIの一生には、性質のまったく違う2つの時期がある**という点です。

- **学習（訓練）のとき** —— データを読ませ、**予測と正解のズレを見て、重みを少しずつ直していく**時期。時間と計算資源を大量に使います
- **推論（実際に使う）のとき** —— **学習が済んで固まった重みを使い、入力に対して答えを出すだけ**の時期。重みは変わりません

**この区別が、そのまま設問になります。**よくある誤解は、**「AIは使えば使うほど、勝手に賢くなっていく」**というものです。**多くの場合、これは誤りです。**学習済みモデルを業務で使っているだけでは、**重みは1ミリも変わりません。**賢くするには、**新しいデータを集めて、あらためて学習をやり直す（再学習する）**必要があります。

**「AIが自ら学習して改善される」という言い方は、この再学習の工程まで含めた表現**だと理解してください。実務では、**使っているうちに現実のほうが変わってしまい、モデルの精度が落ちていく**ことがよく起きます（利用者の好みが変

わる、商品構成が変わる、法律が変わる）。そのため、**定期的に新しいデータで学習し直し、性能を測り直す**という運用が必要になります。**作って終わりではない**というのが、AI を業務に入れるときの実感です。

**1-5 の後半で扱うドロップアウトが「学習時だけの仕組み」だと言われるのも、この区別があるから**です。学習のときと使うときで、**動き方が違う仕組みがある** —— そう押さえておいてください。

### 過学習（オーバーフィッティング）

学習を繰り返せば繰り返すほど良い、とはなりません。ここで出てくるのが **過学習（オーバーフィッティング）** です。

**過学習（オーバーフィッティング）とは、学習に使ったデータに合わせ込みすぎた結果、学習データに対しては非常に高い精度を出すのに、未知の新しいデータに対しては精度が落ちてしまう状態**を指します。

過去問の答えを丸暗記した受験生を思い浮かべてください。過去問なら満点を取りますが、本番の初見の問題には対応できません。**実際に起きているのは、データに含まれるたまたまの偏りやノイズまで規則として覚え込んでしまう**ことです。「この写真の背景がたまたま芝生だった」ことまで猫の特徴として学んでしまえば、室内の猫を猫と見なせなくなります。

**見分け方**は、学習に使わずに取り分けておいたデータで試すことです。**学習データでの成績は良いのに、取り分けたデータでの成績が悪ければ、過学習を疑います。** 逆に、**学習データにすら合っていない状態**は **未学習** と呼ばれ、こちらはモデルが単純すぎるか学習が足りないことを示します。**「精度が出ない＝いつも過学習」ではありません。**

過学習が起きやすいのは、**データの量に対してモデルが複雑すぎるとき、学習を回しすぎたとき、学習データの偏りが大きいとき**です。

**なぜ起きるのかを、もう一段だけ開いておきます。**学習は「誤差を小さくする」方向にひたすら進みます。**モデルに十分な複雑さがあり、学習を回し続ければ、学習データの誤差はどこまでも下げられます。**しかし、学習データの中には、**本当の規則性**と、**そのデータにたまたま含まれていただけの癖（ノイズ）**が混ざっています。**モデルは、その2つを区別できません。**誤差を下げ切ろうとすれば、**たまたまの癖まで規則として覚え込む**——これが過学習の正体です。**「学習しすぎ」が原因であって、「学習が足りない」わけではない**という向きを間違えないでください。

**気づき方も、そのまま出題されます。**学習データと、取り分けておいたデータの、**2つの成績を並べて見る**のが基本です。

| 学習データでの成績 | 未知のデータでの成績 | 状態 | 対処の向き |
|---|---|---|---|
| **高い** | **低い** | **過学習（オーバーフィッティング）** | モデルを単純にする／正則化・ドロップアウト／データを増やす／早期終了 |
| **低い** | **低い** | **未学習（アンダーフィッティング）** | モデルを複雑にする／学習を長く回す／特徴を足す |
| 高い | 高い | 望ましい状態 | — |

**過学習と未学習は、対処の向きが正反対です。**過学習は「抑える」方向、未学習は「増やす」方向。**取り違えると、状況を悪化させます。**「未知のデータで精度が出ない」という記述だけでは、どちらとも決まりません——**学習データでの成績を見て初めて切り分けられる**、というのがここの要点です。

**過学習と未学習 —— 2つの成績を並べて切り分ける**

| 学習データ | 未知のデータ | 状態と、対処の向き |
|---|---|---|
| 高い | 低い | **過学習（オーバーフィッティング）**<br>抑える方向 —— モデルを単純にする／正則化・ドロップアウト／データを増やす／早期終了 |
| 低い | 低い | **未学習（アンダーフィッティング）**<br>増やす方向 —— モデルを複雑にする／学習を長く回す／特徴を足す |
| 高い | 高い | **望ましい状態** |

「未知のデータで精度が出ない」という記述だけでは、どちらとも決まらない。
学習データでの成績を見て初めて切り分けられる。対処の向きは正反対。

図 1-5　過学習と未学習 —— 2つの成績を並べて切り分ける

## 過学習を避ける手法 —— 正則化とドロップアウト

過学習を防ぐ最も素直な方法は、**学習データを増やすこと**です。覚え込めないほど多様な例を与えれば、たまたまの偏りを規則と勘違いしにくくなります。ただし現実にはデータを増やせない場面が多いため、次の工夫が使われます。

**正則化（regularization）**は、**モデルが複雑になりすぎないように、学習の際に「複雑さへの罰則」を加える手法**です。学習は誤差を小さくする方向に進みますが、そこに「重みが大きくなりすぎたら減点」という縛りを足します。すると、**一部の入力だけを極端に重視するような極端な重みづけが抑えられ、全体としてなだらかなモデルになります。**

投資でたとえるなら、**1銘柄への集中投資を防ぐために保有比率の上限を設ける**発想です。当てにいきすぎない縛りを、あらかじめルールとして入れておく。実際には、**重みの大きさに応じた罰則を加える形**が代表的で、罰則の測り方の違いで種類が分かれます。**不要な重みをちょうど0にしやすく、実質的に**

**使う特徴量を絞り込む効き方をするもの**と、**重み全体をまんべんなく小さく抑えるもの**がある、という2系統だけ知っていれば十分です。

**ドロップアウト（dropout）**は、**学習のたびに、ニューラルネットワークのノードの一部をランダムに無効化して学習を進める手法**です。毎回違うノードが休むため、**特定のノードだけに頼りきった構造になりにくく**、結果として未知のデータに強くなります。少しずつ違うメンバーで練習を繰り返すことで、誰か1人に依存しないチームが育つ、というイメージです。

**ドロップアウトで最も問われるのは、これが学習時だけの仕組みだという点**です。**推論のとき（実際に使うとき）は、無効化せずすべてのノードを使います。**「推論時にもノードをランダムに落とす」という選択肢は誤りです。

過学習への対策はこの2つだけではありません。**学習データそのものを増やす、画像を回転させたり明るさを変えたりして学習データを人工的に水増しする（データ拡張）、取り分けたデータの成績が悪化し始めた時点で学習を打ち切る（早期終了）**、といった手も併用されます。**データ拡張（data augmentation）**は、**手元の学習データに加工を施して、水増しする手法**です。画像なら、回転させる・左右反転する・明るさを変える・一部を切り出す、といった加工で1枚を何枚にも増やします。**元データを増やせないときに、実質的なデータ量を増やす**発想です。**早期終了（early stopping）**は、**取り分けたデータでの成績が悪化し始めた時点で、学習を打ち切る**手法です。学習を回しすぎることが過学習の一因なので、行き過ぎる前に手を止めます。投資でいえば、あらかじめ決めた撤退ラインで手を止める発想と同じです。

**「過学習対策＝ドロップアウト」と1つに決め打ちしない**でください。

**正則化とドロップアウトは、名前が並んで出てくるぶん取り違えが起きます。**表で分けてください。

| | 正則化 | ドロップアウト |
|---|---|---|
| **何をするか** | **重みが極端に大きくならないよう、学習に罰則を加えて抑える** | **学習のたびに、ノード（人工ニューロン）の一部をランダムに休ませる** |
| **何を触っているか** | **重みの値**（大きくなりすぎないように縛る） | **ネットワークの経路**（毎回、使う経路を変える） |
| **なぜ過学習が減るのか** | 一部の入力だけを極端に重視する形にならず、**なだらかなモデルになる** | 特定のノードに頼りきった構造にならず、**複数の経路が同じ役割を担えるようになる** |
| **使う場面** | ニューラルネットワークに限らず、**機械学習全般**で使える | **ニューラルネットワーク特有**の手法 |
| **実際に使うとき（推論時）** | 学習が済んだ重みをそのまま使う | **休ませない。すべてのノードを使う** |

**一言でいうと、正則化は「重みの大きさを抑える」、ドロップアウトは「経路をわざと休ませる」**—— 触っている対象が違います。**「正則化はノードを無効化する手法である」「ドロップアウトは重みに罰則を加える手法である」といった、入れ替えた選択肢が典型的な誤り**です。

**なお、この2つは併用できます。**過学習対策は「どれか1つを選ぶ」ものではなく、**データを増やす・データ拡張・正則化・ドロップアウト・早期終了・モデルを単純にする**といった手を、状況に応じて組み合わせます。

### 転移学習

**転移学習（transfer learning）**とは、**ある課題のために大量のデータで学習させた学習済みモデルを、別の課題に流用する手法**です。

なぜこれが強力なのか。ディープラーニングには大量のデータと計算資源が要ります。しかし、**画像を見るための基本的な能力（線を見る、形をとらえる、質感を区別する）は、扱う対象が変わっても共通**です。そこで、**大量の一般的な画像で学習済みのモデルを土台として持ってきて、自分の目的に合わせた出力の部分だけを新しく学習させます。**すると、**手元のデータが少なくても、短時間で実用的な精度に届きます。**

自社の製品の傷を判別したい町工場が、ゼロから画像認識を作るのではなく、公開されている学習済みモデルを土台にして、自社の写真数百枚で仕上げる —— これが転移学習の典型です。

**転移学習とファインチューニングの違い**は、この試験で狙われる対比です。両者はどちらも「学習済みモデルを別の課題に活かす」点では同じですが、**土台の部分を触るかどうかが違います。**

| 手法 | 学習済みモデルの土台部分 | 学習させる範囲 |
|---|---|---|
| **転移学習** | **固定したまま（更新しない）** | 新しく付け足した出力の部分だけ |
| **ファインチューニング** | **固定せず、あわせて調整する** | モデル全体（または広い範囲）を、新しいデータで微調整する |

**固定するか、更新するか** —— ここが分かれ目です。一般に、**手元のデータが少なければ転移学習、ある程度そろえられるならファインチューニング**が向きます。ファインチューニングは表現力の面で有利ですが、**データが少ないまま全体を動かすと過学習しやすい**という難点があります。

生成AI の文脈でも、この考え方はそのまま生きています。大規模言語モデル（LLM）を各社が自社用に仕立て直すときに使われるのが、まさにファインチューニングです。**なお、学習済みモデルの中身を一切変えずに、指示文や参考資**

**料の渡し方だけで振る舞いを変えるやり方もあります。**これは学習ではなく使い方の工夫であり、転移学習ともファインチューニングとも別のものです。

**転移学習には限界もあります。**土台のモデルが学習した領域と、流用先の領域が**かけ離れていると効きません。**一般的な風景写真で学習したモデルは、医療用のレントゲン画像には素直には効きにくい —— 見るべき特徴の性質が違うためです。**「学習済みモデルを持ってくれば、どんな課題でも少ないデータで済む」わけではない**という点は押さえておいてください。ここでも**ノーフリーランチ定理**（1-2）の考え方が効いています。

また、**土台に使う学習済みモデルの素性を確かめる必要**もあります。**学習に使われたデータに偏りがあれば、その偏りは流用先にも引き継がれます。**誰が、どのようなデータで作ったモデルなのかを確認せずに土台に据えるのは、実務上のリスクになります。

### 誤りやすい対比 [1.2]

- **学習＝誤差を見て重みを少しずつ直す繰り返し。**1回で大きく直すのではない
- **過学習＝学習データには高精度だが未知のデータで精度が落ちる状態。**学習データにすら合っていないのは未学習
- **正則化＝複雑さに罰則を加えて、極端な重みづけを抑える／ドロップアウト＝学習のたびにノードをランダムに無効化する**
- **ドロップアウトは学習時だけ。推論時はすべてのノードを使う**
- **転移学習＝土台を固定して出力部分だけ学習／ファインチューニング＝土台も含めて調整する**
- **転移学習が効くのは、手元のデータが少ないとき**

## 1-6 AIの種類 —— 4つのレベルと、弱いAI・強いAI [1.3]

### 特徴量 —— レベル分けを理解する前提

レベル分けの話に入る前に、**特徴量** を押さえます。ここが分かっていないと、レベル3とレベル4の境目が見えません。

**特徴量（feature）とは、データの中で「判断の手がかりになる性質」を数値として取り出したもの**です。中古車の価格を予測するなら、走行距離・年式・排気量・修復歴が特徴量です。手書き文字を判別するなら、線の傾き・曲がり具合・閉じた輪の数などが特徴量になります。**AI は生のデータそのものではなく、特徴量を通して対象を見ている**と考えてください。

問題は、**「何を特徴量にするか」を誰が決めるか**です。

- **従来の機械学習：人間が設計する。** 画像なら「輪郭を見よ」「色の分布を見よ」と人が定義し、その値を計算して機械に渡す。この設計の巧拙が精度を決めてしまい、しかも試行錯誤が要る職人仕事でした
- **ディープラーニング：機械が自ら獲得する。** 大量のデータを見るうちに、有効な見方をネットワークの内部に自分で作り上げる

特徴量について、もう1つ押さえておくことがあります。**何を特徴量に選ぶかで、結果が大きく変わる**という点です。中古車の価格予測に「その日の天気」を入れても精度は上がりませんし、逆に「事故歴」を入れ忘れれば大事な手がかりを落とします。**特徴量が足りなければ精度が出ず、無関係な特徴量が多すぎても精度は落ちます。** 従来の機械学習で「特徴量の設計が職人仕事だ」と言われたのは、この見極めが難しいからです。

**ただし、ディープラーニングでも人間の仕事がゼロになるわけではありません。** どんなデータを集めるか、どういう単位で切り出して見せるか、正解ラベ

ルをどう定義するかは、依然として人間が決めます。**「特徴量の設計」が要らなくなっただけで、「データの設計」は残っている**と捉えてください。

**特徴量を人が設計するのか、機械が自ら学ぶのか。**この一線が、次のレベル3とレベル4を分けます。

なお、**人が特徴量を設計する作業は特徴量エンジニアリングと呼ばれます。**従来の機械学習では、ここに最も時間と経験が費やされました。**ディープラーニングはこの工程を大幅に減らしましたが、代わりに大量のデータと計算資源を要求します。手間をどこに払うかが移っただけ**であり、負担が消えたわけではない、という見方をしておくと実務でぶれません。

## AI の4つのレベル

世の中で「AI」と呼ばれているものは幅が広いため、**中身の高度さで4段階に整理する**見方が使われます。

| レベル | 中身 | 具体例 |
| --- | --- | --- |
| **レベル1** | **単純な制御プログラム。**すべての振る舞いを人が書いており、学習しない | **設定温度に応じて運転を切り替えるエアコン**（温度が上がったら冷房を強める、という動作が人の手で書かれている） |
| **レベル2** | **古典的な AI。**振る舞いのパターンが多く、探索・推論や知識ベースを使って場合分けする | **将棋やチェスのプログラム**（膨大な指し手を探索して選ぶ。ただし規則と評価の枠組みは人が与えている） |
| **レベル3** | **機械学習を取り入れた AI。データから規則を学ぶ。**ただし特徴量は人が設計する | **迷惑メールの判別**（大量のメールから、迷惑メールらしさの規則をデータで学ぶ） |

| レベル | 中身 | 具体例 |
| --- | --- | --- |
| **レベル4** | **ディープラーニングを取り入れたAI。特徴量そのものを自ら学習する** | **画像認識・音声認識・生成AI**（何に注目すべきかを人が教えなくても、データから獲得する） |

**境目は2か所しかありません。**

- **レベル2とレベル3の間：人がルールを書くのか、データから学ぶのか**
- **レベル3とレベル4の間：特徴量を人が設計するのか、機械が自ら学ぶのか**

この2本の線を覚えておけば、具体例が入れ替わっても判定できます。「エアコンの温度制御はレベル3」「画像認識はレベル2」といった選択肢は、この線に照らせば切れます。

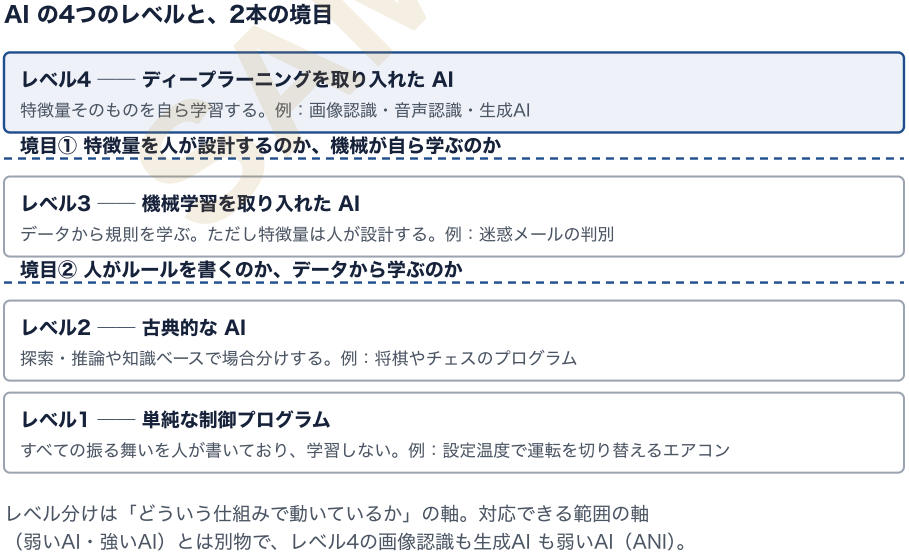



図 1-6　AI の4つのレベルと、2本の境目

**各レベルの具体例を、もう1つずつ添えておきます。**「1例だけ覚えると、別の例が出たときに判定できない」ためです。

- **レベル1（単純な制御）** —— エアコンの温度制御のほか、**設定した時刻に点灯する照明のタイマー**。条件と動作が1対1で人手で書かれている
- **レベル2（古典的なAI）** —— 将棋プログラムのほか、**症状を順に尋ねて診断候補を絞り込む問診システム**。多数の場合分けを探索・推論でたどるが、規則は人が用意している
- **レベル3（機械学習）** —— 迷惑メール判別のほか、**購買履歴から次に買いそうな商品を薦める推薦システム**。データから規則を学ぶが、注目すべき項目は人が設計する
- **レベル4（ディープラーニング）** —— 画像認識・生成AI のほか、**音声から話者の言葉を書き起こす音声認識**。何に注目すべきかまでデータから獲得する

**レベル1について1つ補足します。** レベル1は「学習も適応もしない単純な自動化」であり、1-1 で見たとおり、**厳密には AI と呼ぶべきか議論のある水準**です。それでもこの表に並ぶのは、**世間で「AI 搭載」と宣伝される製品の中に、実際にはレベル1に近いものが混じっているから**です。**レベル分けは「AI という言葉が指す幅の広さ」を示すための整理**であって、上の段ほど偉いという順位表ではありません。

### 弱いAI（ANI）と強いAI（AGI）

レベル分けが「実装の高度さ」の軸だったのに対し、こちらは **「対応できる範囲の広さ」の軸**です。**別の軸なので、混ぜないでください。**

**弱いAI（ANI、Artificial Narrow Intelligence／特化型人工知能）** とは、**あらかじめ決められた特定の課題に限って能力を発揮する AI** です。将棋を指

す、顔を判別する、翻訳する、文章を生成する —— どれも得意分野の中では人間を上回ることがありますが、**その枠の外へは出られません。**将棋のAIに翻訳はできませんし、翻訳のAIに車の運転はできません。

**重要なのは、いま実用化されている AI はすべて弱いAI（ANI）だという点です。**画像認識も、音声アシスタントも、自動運転も、そして ChatGPT のような生成 AI も、この分類では弱いAI に入ります。**「幅広い話題に答えられるから強いAIだ」ではありません。**生成AI は扱える話題こそ広いものの、**人間のように自ら目的を立て、未知の領域へ能力を持ち出して応用することはできない**ため、弱いAI に分類されます。

**強いAI（AGI、Artificial General Intelligence／汎用人工知能）**とは、**人間と同じように、分野を限定せずにさまざまな課題へ柔軟に対応できる AI** を指します。**未知の状況でも、自ら考えて対処し、獲得した知識を別の領域へ持ち出せる**ような知能です。**強いAI は、現時点では実現していません。**実現の可能性や時期については研究者の間でも見解が分かれており、この論点が次の 1-7、1-8 につながります。

**強いAI（AGI）の先を論じる言葉**も、あわせて知られています。**人間の知能をあらゆる面で大きく上回る知能**を指す **人工超知能（ASI、Artificial Super Intelligence）**です。**ANI → AGI → ASI** という順で範囲と能力が広がる整理で語られ、**ASI は 1-8 で扱うシンギュラリティの議論と直結します。**ただし試験で中心的に問われるのは **ANI と AGI の対比**ですので、まずそこを固めてください。

**「弱い」「強い」という呼び方は性能の優劣ではありません。**特化型の AI が人間より速く正確に働く例はいくらでもあります。**分かれ目は、能力が特定の課題に限られるか、分野を越えて汎用に及ぶか**です。「弱いAI は精度が低い AI のことである」という選択肢は誤りになります。

### 2つの軸を混ぜない

まとめると、AI の種類は **2本の別々の軸**で問われます。

- **レベル1〜4 —— どういう仕組みで動いているか**（人がルールを書く／データから学ぶ／特徴量まで学ぶ）
- **弱いAI（ANI）と強いAI（AGI） —— どこまでの範囲に対応できるか**（特定の課題だけ／分野を越えて汎用に）

**レベル4だから強いAI、ということにはなりません。**画像認識も生成AI もレベル4に当たりますが、いずれも弱いAI です。**この掛け違いを狙った選択肢が出ます。**

**生活の中で使われている AI を、この2軸で見てみます。**

| 身近な例 | レベル | ANI / AGI |
|---|---|---|
| スマートフォンの顔認証 | レベル4（ディープラーニング） | 弱いAI（ANI） |
| 迷惑メールの自動振り分け | レベル3（機械学習） | 弱いAI（ANI） |
| 対話型の生成AI | レベル4（ディープラーニング） | 弱いAI（ANI） |
| 設定温度で運転を切り替えるエアコン | レベル1（単純な制御） | AI と呼ぶかに議論がある水準 |

**表の右側がすべて「弱いAI」で埋まる**ことに注目してください。**レベルが上がっても、右側の欄は変わりません。**これが「2軸は別物」ということの意味です。

### 強いAI（AGI）は実現するのか

強いAI（AGI）については、**実現の可能性そのものについて見解が分かれています。**「いまの延長線上で近く到達する」と考える立場もあれば、「規模を大きくするだけでは質の違う知能には届かない」と考える立場もあります。**この試験で求められるのは、どちらが正しいかの判断ではなく、「現時点では実現していない」「見解が分かれている」という事実を正しく押さえること**です。

**現実の見極め方**も1つ添えておきます。あるAIが弱いAIか強いAIかを判定するには、**「学習させていない、まったく別の分野の課題を、自ら方法を考えて解けるか」**を問えば足ります。生成AIは幅広い話題に応答しますが、**それは幅広いデータで学習したからであって、学んでいない領域へ自力で能力を持ち出しているわけではありません。**ここを取り違えると、「生成AIはAGIだ」という誤答に引き込まれます。

### 誤りやすい対比 [1.3]

- **特徴量＝判断の手がかりになる性質を数値化したもの。**レベル3までは人が設計し、レベル4（ディープラーニング）は機械が自ら学ぶ
- **レベル1＝単純な制御（エアコンの温度制御）／レベル2＝探索・推論を使う古典的なAI（将棋プログラム）／レベル3＝機械学習（迷惑メール判別）／レベル4＝ディープラーニング（画像認識・生成AI）**
- **弱いAI（ANI）＝特定の課題に特化／強いAI（AGI）＝分野を越えて汎用に対応**
- **現在実用化されているAIは、生成AIを含めてすべて弱いAI（ANI）。強いAIは未実現**
- **「弱い・強い」は性能の優劣ではなく、対応できる範囲の広さの違い**

- **レベル分けとANI／AGIは別の軸。**レベル4であることは強いAIであることを意味しない

## 1-7 AIの歴史 —— 3度のブームと2度の冬 [1.4]

### 全体の骨格を先に押さえる

AIの歴史は、**3度のブームと、その合間に2度訪れた冬**という形で整理されます。**まずこの数の対応を固定してください。ブームは3度、冬は2度。第三次ブームは現在も続いており、その後に冬は来ていません。**

| 時期 | ブームの中心技術 | できたこと | 冬が来た理由 |
|---|---|---|---|
| **第一次AIブーム**（1950年代後半〜1960年代） | **探索と推論** | 迷路・パズル・定理証明といった、規則の明確な問題を解いた | **単純な問題しか解けなかった。**現実の複雑な問題に歯が立たず、期待が失望に変わった |
| **第二次AIブーム**（1980年代） | **知識表現とエキスパートシステム** | 専門家の知識を書き写し、実用的な助言を出せるようになった | **知識を書き出す作業が終わらなかった。**記述と保守の負担に耐えられず、再び冬へ |
| **第三次AIブーム**（2000年代後半〜現在） | **機械学習とビッグデータ、ディープラーニング** | データから規則も特徴量も自ら学ぶ。画像認識・音声認識・生成AIへ | 現在進行中 |

**「AIの冬（AI winter）」とは、ブームで高まった期待に技術の実力が追いつかず、失望から研究資金や社会の関心が急速にしぼんだ停滞期**を指します。**第一次ブームの後と、第二次ブームの後の2度**訪れました。**冬を招いたのは、事故や規制ではなく、期待と実力の差**です。投資でいえば、業績が追いつかない

まま買われた銘柄が、実態が知られた途端に資金を引き上げられる局面に近いものです。

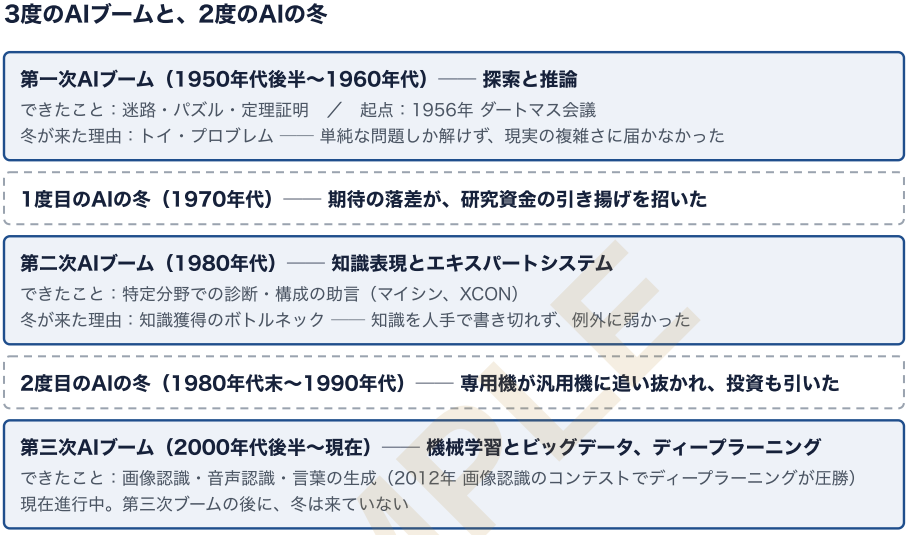



図 1-7　3度のAIブームと、2度のAIの冬

## 第一次AIブーム —— 探索と推論

**始まりは1956年のダートマス会議**です（1-1 参照）。この時期の中心技術は **探索** と **推論** でした。

**探索** とは、**取りうる選択肢を場合分けして枝分かれの形に並べ、目的の状態にたどり着く道筋をしらみつぶしに調べる**やり方です。迷路であれば「この分かれ道を左へ行ったら、次はどうなるか」を順に展開していきます。**推論** は、**与えられた前提と規則から、論理の手続きに従って結論を導く**ことです。「A ならば B」「B ならば C」という規則があるとき、A から C を導く —— その手続きを機械にやらせました。

代表的な成果として、**迷路やハノイの塔といったパズルを解くプログラム、数学の定理を自動で証明するプログラム**（1-1 のロジック・セオリスト）、そして**チェスなどのボードゲームのプログラム**が挙げられます。当時としては驚くべき成果で、「機械が考えている」と受け止められました。

この時期を象徴するもう1つの成果が、**対話プログラムのイライザ（ELIZA）**です。1960年代に作られたもので、**利用者の入力に含まれる語をとらえ、あらかじめ決めた型に当てはめて質問を返す**仕組みでした。「私は疲れています」と入力すると「なぜ疲れているのですか」と返す、といった具合です。**中身は言葉の意味を理解しているわけではなく、パターンに従って言い換えているだけ**でしたが、それでも利用者の中には、本当に理解されていると感じる人が現れました。**この「機械の応答を人が過剰に人間的に受け取ってしまう傾向」はイライザ効果と呼ばれ、生成AI が普及した現在こそ意識すべき論点**になっています。

**イライザは対話プログラム、ロジック・セオリストは定理証明プログラム、マイシンは診断のエキスパートシステム。**名前と中身の対応が問われますので、この3つは取り違えないでください。

しかし、ここに落とし穴がありました。**解けたのは、ルールが明確で、選択肢が限られ、正解がはっきりしている問題ばかりだった**のです。こうした問題は **トイ・プロブレム（おもちゃの問題）**と呼ばれます。現実の問題 —— 病気の診断、経営の意思決定、自然な会話 —— は、**取りうる選択肢が天文学的な数になり、しかも前提条件が明示されていません。**探索の枝は爆発し、計算が終わらなくなりました。

**なぜ探索が行き詰まるのか**を、もう一段だけ説明しておきます。選択肢を1手ごとに枝分かれさせていくと、**手数が伸びるほど枝の数が掛け算で増えていきます。**1手につき10通りの選択肢があるなら、10手先を読むだけで組み合わ

せは天文学的な数になります。当時の計算機では、現実的な時間で調べ尽くすことができませんでした。**「原理的には解ける」ことと「現実の時間で解ける」ことは別だ**という教訓が、ここで得られています。

「実用にならない」と分かると、期待は急速にしぼみます。こうして **1度目のAIの冬**（1970年代）が訪れました。

**冬の実害は、研究資金と人材に及びました。**国や企業が AI 研究への予算を絞り、研究者が別の分野へ移り、「人工知能」という言葉を掲げること自体が避けられる時期すらありました。**技術そのものが否定されたわけではなく、期待の落差が資金の引き揚げを招いた**という構図です。**この構図は、いまの生成AI への期待を考えるうえでも参照されます。**

### 第二次AIブーム —— 知識表現とエキスパートシステム

1980年代、**発想を変えた第二次AIブーム**が来ます。着眼点は明快でした。**第一次で足りなかったのは「知識」だ。ならば、人間の専門家が持つ知識を、機械が使える形で書き写せばよい。**

この考え方から生まれたのが **エキスパートシステム（expert system）** です。**特定分野の専門家の知識を「もし〜ならば〜である」という多数の規則として書き写し、その規則を使って専門家のような助言を出すシステム**を指します。中身は、**規則の集まりである知識ベース**と、**規則を当てはめて結論を導く推論エンジン**に分かれます。1-2 で扱った **ルールベース** の考え方が、この時期に最も大きく花開いたと捉えてください。

代表例として、**血液中の細菌感染症を診断し、抗生物質を提案するマイシン（MYCIN）** が挙げられます。専門医には及ばないものの、**専門外の医師よりは高い正答率**だったと報告され、大きな注目を集めました。また、**コンピュータの受注時に、顧客の要望に応じた部品構成を自動で決める XCON（別名 R1）**

は、企業の業務に導入されて大幅なコスト削減を実現し、**商用のエキスパートシステムの成功例**として知られます。

この時期、知識をどう書き表すか —— **知識表現** の研究も進みました。**概念を点、概念どうしの関係を線で表す意味ネットワーク**や、**概念とその関係を体系的に定義するオントロジー**といった手法が整理されたのもこの流れです。日本でも、**通商産業省（当時）が1982年に第五世代コンピュータプロジェクトという国家プロジェクトを立ち上げ、知識情報処理を目指す新しいコンピュータ技術の開発に取り組みました。**

**第二次ブームには、もう1つの側面があります。**期待の高まりを受けて、**知識処理に特化した専用のコンピュータを開発する企業や、AI を掲げる新興企業が相次いで設立される投資ブーム**が起きました。ところが、**汎用のコンピュータの性能が急速に向上した結果、専用機の優位性が失われ、多くの企業が撤退**しています。**専用のハードウェアは、汎用のハードウェアの進歩に追い抜かれると価値を失う** —— 技術の歴史で繰り返されてきた構図です。

しかし、ここでも壁にぶつかります。**知識獲得のボトルネック**です。専門家の頭の中にある知識は、**本人も言葉にできない「勘」や「経験」を含んでおり、書き出す作業そのものが膨大**でした。しかも書き出せたとしても、**規則の数が増えるほど互いに矛盾し、保守が手に負えなくなります。**さらに、**書かれていない状況には対応できない**という、ルールベースの根本的な限界も残りました。

1980年代の終わりから1990年代にかけて、期待はまたしぼみ、**2度目のAIの冬**が訪れます。

**第二次ブームの終わりごろから第三次にかけて、「知識を集めること」と「意味を理解すること」は別だと分かる出来事**もありました。**2011年、IBM の質問応答システム「ワトソン（Watson）」が、アメリカのクイズ番組で歴代チャンピオンに勝利**しています。ただしワトソンは質問の意味を理解していたわけでは

なく、**大量の知識源から高速に検索し、関連度で答えの候補を絞り込む**仕組みでした。同じ頃、日本では大学入試の突破を目指す研究プロジェクトが進みましたが、**文章の意味を読み取る力の壁**に突き当たっています。**「大量に知っていること」と「意味が分かっていること」は違う** —— この論点は、生成AIのハルシネーションを考えるうえでも生きてきます。

## 第三次AIブーム —— 機械学習とビッグデータ、そしてディープラーニング

**2000年代後半から現在まで続くのが第三次AIブーム**です。中心にあるのは**機械学習**で、後半から**ディープラーニング**が主役になりました。

**発想の転換は「規則は人が書くものではなく、データから学ばせるものだ」という点**にあります。この転換を可能にしたのは、技術だけではありません。**次の3つの条件がそろったこと**が決定的でした。

1. **ビッグデータ** —— インターネットとスマートフォンの普及によって、**文章・画像・音声・行動履歴といった膨大なデータが日々蓄積される**ようになりました。**ビッグデータとは、従来の手法では扱いきれないほど巨大で、種類が多様で、しかも高速に増え続けるデータ**を指します。**機械学習は大量のデータがあって初めて力を発揮する**ため、これは前提条件そのものでした**ビッグデータについて、もう一段だけ補足します。** ビッグデータが語られるとき、しばしば**量・種類・発生の速さ**という3つの側面が挙げられます。**量**は、単純にデータの規模が大きいこと。**種類**は、表形式の数値だけでなく、文章・画像・音声・位置情報・センサーの記録といった、形の揃わないデータが混ざっていること。**発生の速さ**は、それが日々、時に秒単位で増え続けることです。**「大きい」だけがビッグデータの条件ではない**という点が、設問で狙われます。

そして、**ビッグデータそのものが AI なのではありません。**ビッグデータは材料であり、AI は材料を使う技術です。**「ビッグデータとは AI の一種である」といった選択肢は誤り**になります。

2. **計算能力の向上** —— **GPU** をはじめとする、大量の計算を並列にこなせる装置が普及し、現実的な時間で学習を回せるようになりました
3. **手法の進歩** —— ニューラルネットワークを深く重ねる技術的な工夫が積み重なりました

**同じ考え方でも、データと計算資源がなければ動かなかった**という点は出題されやすい視点です。ニューラルネットワークの基本的なアイデア自体は、実は第一次・第二次ブームの時期から存在していました。**眠っていたアイデアが、環境が整って初めて花開いた**というのが第三次ブームの構図です。

**第三次ブームを裾野まで広げたもの**として、**誰でも使えるかたちで公開された開発用の道具（オープンソースのソフトウェア）**と、**計算資源をインターネット越しに借りられるクラウドサービス**の存在も挙げられます。**自前で高価な計算機を買わなくても、必要なときに必要なだけ借りて学習を回せる**ようになったことで、大企業や研究機関以外でも AI の開発に手が届くようになりました。**第二次ブームで専用機が高価だったこと**と対比すると、この違いの大きさが分かります。

**決定的な転換点は2012年**でした。画像認識の精度を競う国際的なコンテストで、**ディープラーニングを用いたチームが、それまでの手法に大差をつけて優勝**したのです。それまで数年かけてわずかずつ改善していた分野で、一気に誤り率が下がったことが世界に衝撃を与えました。ここから、**特徴量を人が設計する時代が終わり、機械が特徴量を獲得する時代**が始まります（1-6 のレベル3からレベル4への移行がこれです）。

その後、画像に続いて音声認識でも人間に匹敵する精度が報告され、2010年代後半には文章を扱う技術が大きく進みました。そして **2022年以降、対話型の生成AI が一般利用者に広く普及**し、AI は研究室の技術から日常の道具へと移りました。生成AI の詳細は次章以降で扱います。

**なお、第三次ブームは「機械学習だけの時代」ではありません。**説明責任が求められる場面ではルールベースが現役ですし、**機械学習とルールベースを組み合わせて使う**のが実務では普通です。「第三次ブームでルールベースは消えた」という趣旨の選択肢は誤りです。

**3度のブームを、1本の流れとして読み直しておきます。**

**第一次は「考え方の手順（探索と推論）を与えれば賢くなるはずだ」と考えた。**足りなかったのは知識でした。**第二次は「知識を与えれば賢くなるはずだ」と考えた。**足りなかったのは、知識を人手で書き切ることの限界を超える方法でした。**第三次は「知識もルールもデータから学ばせればよい」と考え、データと計算資源がそれを可能にした。**

**2度の冬を、理由まで含めて1枚に畳んでおきます。**「いつ冬が来たか」より、**「なぜ冬が来たか」が問われます。**

| | 第一次AIブーム | 第二次AIブーム | 第三次AIブーム |
|---|---|---|---|
| **主役の技術** | **探索・推論** | **知識表現・エキスパートシステム** | **機械学習・ディープラーニング** |
| **できたこと** | パズル・定理証明・ボードゲーム | 特定分野での診断・構成の助言 | 画像認識・音声認識・言葉の生成 |

| | 第一次AIブーム | 第二次AIブーム | 第三次AIブーム |
|---|---|---|---|
| **限界の中身** | **おもちゃの問題（トイ・プロブレム）は解けても、現実の複雑さに届かなかった。**選択肢が増えると枝が爆発し、計算が終わらない。しかも現実の問題は前提条件が明示されていない | **知識を人手で書き続けることの限界。**専門家の勘は言葉にできず、規則が増えるほど互いに矛盾し、保守できなくなる。さらに**書かれていない例外にはまったく対応できない** | （継続中。データの無い領域では学べない／判断の理由を説明しにくい／データの偏りを受け継ぐ） |
| **その後** | **1度目のAIの冬**（1970年代）—— 期待の落差が研究資金の引き揚げを招いた | **2度目のAIの冬**（1980年代末〜1990年代）—— 専用ハードウェアが汎用機に追い抜かれ、投資も引いた | — |

**2つの冬の理由を、一言ずつで言えるようにしてください。第一次は「現実の複雑さに届かなかった」、第二次は「知識を人が書き続けられなかった・例外に弱かった」。**冬の理由を取り違えた選択肢（第一次の理由として知識獲得の限界を挙げる、など）は頻出です。

**そして、第三次を支えたのが何かも押さえてください。ビッグデータとインターネットの普及**です。第一次・第二次と第三次を分けたのは、**新しい発想が生まれたことよりも、データと計算資源という前提が整ったこと**でした。**インターネットとスマートフォンが日々のデータを大量に生み、そのデータを学習の材料にできるようになった** —— 機械学習の考え方自体は以前から存在していたのに第三次まで花開かなかったのは、**材料が無かったから**です。「第三次ブームは、まったく新しい理論の発明によって始まった」という趣旨の選択肢は、この点で不正確です。

**それぞれのブームが、前のブームの行き詰まりへの回答になっている** —— この流れで捉えると、暗記ではなく理解として残ります。同時に、**第三次にも限界がある**ことも見えます。データが無い領域では学べず、判断の理由を説明しにくく、データの偏りをそのまま受け継ぐ。**この限界が、第4章で扱う社会的な論点の出発点**になります。

## 年と出来事の対応

**年号が問われる形に備えて、この章で扱った出来事を並べておきます。**

| 年 | 出来事 |
| --- | --- |
| **1956年** | **ダートマス会議。「Artificial Intelligence」という言葉が生まれる**（第一次AIブームの起点） |
| **1960年代** | 対話プログラム **イライザ（ELIZA）**。探索・推論による問題解決が盛んになる |
| **1980年代** | **エキスパートシステム**が実用化され、第二次AIブーム（マイシン、XCON） |
| **2011年** | **ワトソン（Watson）** がクイズ番組で歴代チャンピオンに勝利 |
| **2012年** | **画像認識のコンテストでディープラーニングが圧勝**。第三次ブームの決定的な転換点 |
| **2022年以降** | **対話型の生成AI が一般に普及** |

**「第一次と第二次のあいだ」「第二次と第三次のあいだ」が、それぞれ AIの冬**にあたります。

### 誤りやすい対比 [1.4]

- **ブームは3度、冬は2度。**第三次ブームの後に冬は来ていない
- **第一次＝探索と推論（迷路・パズル・定理証明）。**行き詰まりの理由は**トイ・プロブレム**（単純な問題しか解けなかった）
- **第二次＝知識表現とエキスパートシステム（マイシン、XCON）。**行き詰まりの理由は**知識獲得のボトルネック**（知識を書き出す作業が終わらない）
- **第三次＝機械学習とビッグデータ、ディープラーニング。**支えたのはデータの増加と計算能力の向上
- **AIの冬を招いたのは、期待と実力の差。**事故や規制が原因ではない
- **ダートマス会議は1956年で、第一次ブームの起点**

## 1-8 シンギュラリティ（技術的特異点） [1.5]

### シンギュラリティとは何か

**シンギュラリティ（技術的特異点、technological singularity）**とは、**AI が人間の知能を超え、自らそれ以上に優れた AI を作り出せるようになることで、技術の進歩の速度が人間には予測も制御もできなくなる時点**を指す概念です。

「AI が人間より賢くなること」だけを指すのではありません。**核心は「自分より優れたものを自分で作れるようになる」という点**にあります。人間より賢い AI が、さらに賢い AI を設計する。その AI がもっと賢い AI を設計する —— この繰り返しが加速すると、進歩の速度が人間の理解を振り切ってしまう。**その振り切る「時点」を指すのがシンギュラリティ**です。

**なぜ「自分より優れたものを作れる」ことが特別なのか。**人間が AI を改良する速度は、人間の能力と時間に縛られます。ところが、**改良する側が AI になれ**

**ば、その縛りが外れます。**改良されたAIは前より優れているので、次の改良はもっと速く進む。これが繰り返されると、進歩の速度そのものが加速していきます。**この「自己改良が繰り返されることで知能が加速度的に高まる」という考え方は知能爆発と呼ばれ、シンギュラリティの議論の土台**になっています。

**知能爆発とシンギュラリティは、混同されやすい別の語**です。**知能爆発が指すのは「自己改良が繰り返される仕組み」そのもの、シンギュラリティが指すのは「その結果として人間に先が読めなくなる時点」**です。**仕組みの側か、時点の側か**で切り分けてください。

**「特異点」という訳語**が使われるのは、それまでの延長線上で先を読むことができなくなる境目、という意味合いからです。

### ヴァーナー・ヴィンジとレイ・カーツワイル —— 取り違えない

この項目でいちばん問われるのが、**2人の人物と、それぞれが何を言ったのか**の対応です。**逆にした選択肢が用意されます。**

| 人物 | 何をした人か |
| --- | --- |
| **ヴァーナー・ヴィンジ（Vernor Vinge）** | 数学者・SF作家。**「技術的特異点」という概念を論じて広めた人物。**1993年の論文「The Coming Technological Singularity」で、**人間を超える知能が生まれれば、人間の時代は終わり、その先は予測できなくなる**と論じた |
| **レイ・カーツワイル（Ray Kurzweil）** | 発明家・未来学者。**「2045年にシンギュラリティが到来する」という具体的な時期の予測**を示し、議論を一般に広めた人物 |

**覚え方は「ヴィンジ＝概念、カーツワイル＝2045年という数字」です。**年号が出てきたらカーツワイル、と結びつけてください。

このカーツワイルの予測から生まれた言葉が **2045年問題** です。**2045年問題とは、シンギュラリティが2045年に到来するとされることから、そのとき社会・雇用・人間の役割はどうなるのかを問う議論**を指します。**具体的には、多くの職業が AI に置き換わるのではないか、AI が人間の制御を離れるのではないか、といった懸念が論じられます。**

ここで大事なのは、**2045年問題は「確定した予定」ではなく、予測を出発点にした議論だ**ということです。**シンギュラリティが本当に到来するのか、するとしていつなのかについては、専門家の間でも見解が大きく分かれています。**「近いうちに必ず起きる」と考える研究者もいれば、「現在の技術の延長線上では起きない」と懐疑的な研究者もいます。**「シンギュラリティの到来は学術的に確定している」という趣旨の選択肢は誤り**です。

### AI効果 —— シンギュラリティと混同しやすい概念

同じ節のキーワードとして **AI効果（AI effect）** が置かれています。**シンギュラリティとはまったく別の概念なので、ここで切り分けます。**

**AI効果とは、AI によって何かが実現されると、人々はその仕組みを理解した途端に「これは知能ではない。単なる自動処理だ」と見なすようになる現象**を指します。

具体例で見ると分かりやすくなります。文字を読み取る技術、迷惑メールの判別、経路の検索、将棋のソフト —— どれも登場した当時は「AI の快挙」として語られました。ところが原理が知られ、日常に溶け込むと、「あんなものは AI ではない」と言われるようになります。**実現される前は知能の証とされ、実現された途端に知能から外される。**この繰り返しが起きるのが AI効果です。

**AI効果は、ビジネスの現場でも実際に起きます。**導入前は「AI で劇的に変わる」と期待され、導入後は「思ったほどではない」と評価が下がる。**その落差の**

**一部は、性能の問題ではなく、動いているものを人が当たり前と見なすようになるという受け止め方の変化**です。**期待の管理そのものが導入の一部だ**と考えておくと、この現象に振り回されずに済みます。

**AI効果がもたらす帰結は2つ**あります。1つは、**AI と呼ばれる範囲が、常に「まだ実現していないもの」の側へ逃げていく**こと。もう1つは、その結果として **AI の定義がいつまでも定まらない**ことです（1-1 で見たとおりです）。

**シンギュラリティと AI効果の違い**を1行で言い切ります。**シンギュラリティは「AI が人間を超え、進歩が予測不能になる時点」。AI効果は「実現した AI を人が知能と見なさなくなる、人間側の受け止め方の現象」です。** 前者は AI の能力の話、後者は人間の評価の話 —— まったく別の軸にあります。**この2つを入れ替えた選択肢が典型的な引っかけ**になります。

**AI効果の例をもう1つ挙げておきます。** かつては「AI の代表例」として語られたカーナビの経路探索は、いまでは誰も AI とは呼びません。**仕組み（探索）は当時のまま**で、変わったのは人間の受け止め方だけです。**技術の中身が変わっていないのに呼び名が変わる** —— これが AI効果の分かりやすい形です。

### 実務ではどう受け止めるか

シンギュラリティの議論は将来の話ですが、この試験でこの項目が置かれている理由は、**AI に対する過大な期待と過小な評価の両方を戒めるため**です。

**過大な期待の側**では、「AI がすべてを解決する」「もうすぐ人間を全面的に超える」という語り方に引きずられます。しかし 1-6 で見たとおり、**現在実用化されている AI はすべて弱いAI（ANI）であり、強いAI（AGI）は実現していません。** シンギュラリティは、その強いAI がさらに先へ進んだ地点の話です。

**過小な評価の側**にあるのが、まさに AI効果です。**「原理が分かったから大したことはない」と切り捨てていると、実際に業務が置き換わっていく変化を見落**

**とします。**

この試験が求めているのは、**過度に恐れず、過度に軽んじず、いま何ができて何ができないかを見極める姿勢**です。歴史が3度のブームと2度の冬を繰り返してきたこと（1-7）も、同じ教訓を別の形で示しています。**期待が実力を追い越したとき、冬が来た**のでした。

### この節で押さえるべき「言えること」と「言えないこと」

シンギュラリティは将来の議論なので、**断定した記述はたいてい誤り**になります。整理しておきます。

**言えること。** シンギュラリティという概念が **ヴァーナー・ヴィンジ**によって論じられ、**レイ・カーツワイル**が **2045年**という時期を予測したこと。その予測をめぐる議論が **2045年問題**と呼ばれること。**専門家の間で見解が分かれていること。** そして、**現時点で強いAI（AGI）は実現していないこと**（1-6）。

**言えないこと。**「シンギュラリティの到来時期は確定している」「AI が人間を超えることは科学的に証明されている」「2045年に AI が人類を支配することが決まっている」―― いずれも**根拠のない断定**であり、選択肢として出れば誤りです。**将来の予測を、確定した事実のように書いた記述には警戒してください。**

### 誤りやすい対比 [1.5]

- **シンギュラリティ（技術的特異点）＝AI が人間の知能を超え、自らより優れた AI を作れるようになって、進歩が予測不能になる時点**
- **ヴァーナー・ヴィンジ＝概念を論じて広めた人／レイ・カーツワイル＝2045年という時期を予測した人。**年号が出たらカーツワイル

- **2045年問題＝カーツワイルの予測を起点にした、社会や雇用への影響をめぐる議論。**確定した予定ではない
- **シンギュラリティの到来については専門家の間でも見解が分かれている**
- **AI効果＝実現すると「これは知能ではない」と見なされる現象。**AI の性能が向上する現象でも、シンギュラリティの別名でもない

## 第1章の試験ポイント [1.1/1.2/1.3/1.4/1.5]

この章で混同しやすい対を、まとめて並べます。**設問はほぼ、この表のどこかを突いてきます。**

### 用語の対比

| 対比される用語 | 違いを分ける一線 |
| --- | --- |
| **AI と ロボット** | AI は知的な情報処理を担う**ソフトウェア**。ロボットは駆動部を持つ**物理的な機械**。AI を積まないロボットも、身体を持たない AI も存在する |
| **AI ／ 機械学習 ／ ディープラーニング** | **AI ⊃ 機械学習 ⊃ ディープラーニング** の入れ子。並列の別物ではない |
| **ルールベース と 機械学習** | **規則を作るのが人間か、機械か。**ルールベースは説明しやすいが保守が重い。機械学習は大量のデータが要る |
| **教師あり学習 と 教師なし学習** | **正解ラベルを与えるか、与えないか** |
| **教師なし学習 と 半教師あり学習** | 教師なしは**ラベルを一切与えない。ラベルを少しだけ与えるのが半教師あり** |
| **教師あり学習 と 強化学習** | 与えるのが**正解ラベルか、報酬か** |

| 対比される用語 | 違いを分ける一線 |
| --- | --- |
| **分類 と 回帰** | 出力が**カテゴリか、連続した数値か**。どちらも教師あり学習 |
| **分類 と クラスタリング** | **あらかじめ決めたカテゴリに割り当てる（分類・教師あり）／カテゴリ自体をデータから作る（クラスタリング・教師なし）** |
| **クラスタリング と 次元削減** | どちらも**教師なし学習**。**似たもの同士をまとめる**のがクラスタリング、**特徴の数を減らす**のが次元削減 |
| **ニューロン と シナプス** | ニューロン＝**神経細胞そのもの**／シナプス＝**細胞どうしのつなぎ目**。重みが対応するのはシナプスの伝わりやすさ |
| **ニューラルネットワーク と ディープラーニング** | ディープラーニングは、**中間層を多数重ねたニューラルネットワーク**を使う手法。別系統ではない |
| **重み と 情報の重みづけ** | 重み＝**入力ごとの重要度を表す数値**そのもの／情報の重みづけ＝**その重みを掛けて足し合わせる処理** |
| **過学習 と 未学習** | 過学習＝**学習データには高精度、未知のデータで精度が落ちる**／未学習＝**学習データにすら合っていない** |
| **正則化 と ドロップアウト** | 正則化＝**複雑さに罰則を加えて極端な重みづけを抑える**／ドロップアウト＝**学習のたびにノードをランダムに無効化する** |
| **転移学習 と ファインチューニング** | **土台を固定して出力部分だけ学習する**のが転移学習／**土台も含めて調整する**のがファインチューニング |
| **レベル3 と レベル4** | **特徴量を人が設計する（レベル3）／機械が自ら学ぶ（レベル4）** |

| 対比される用語 | 違いを分ける一線 |
| --- | --- |
| **弱いAI（ANI）と強いAI（AGI）** | **特定の課題に特化するか、分野を越えて汎用に対応できるか。**性能の優劣ではない |
| **レベル分けとANI／AGI** | **仕組みの高度さの軸**と、**対応範囲の広さの軸**。別の軸なので掛け違えない |
| **シンギュラリティとAI効果** | シンギュラリティ＝**進歩が予測不能になる時点**（AIの能力の話）／AI効果＝**実現すると知能と見なされなくなる現象**（人間の受け止め方の話） |
| **ヴァーナー・ヴィンジとレイ・カーツワイル** | ヴィンジ＝**概念を論じて広めた**／カーツワイル＝**2045年という時期を予測した。**年号が出たらカーツワイル |

## 数と対応で覚えるもの

| 項目 | 覚える内容 |
| --- | --- |
| **ダートマス会議** | **1956年。ジョン・マッカーシーが「Artificial Intelligence」を提案。**AIという言葉の誕生 |
| **AIブームと冬の数** | **ブームは3度、冬は2度。**第三次ブームの後に冬は来ていない |
| **第一次AIブーム** | 中心＝**探索と推論**／代表＝迷路・パズル・定理証明／冬の理由＝**トイ・プロブレム** |
| **第二次AIブーム** | 中心＝**知識表現とエキスパートシステム**／代表＝**マイシン（MYCIN）・XCON**／冬の理由＝**知識獲得のボトルネック** |
| **第三次AIブーム** | 中心＝**機械学習とビッグデータ、ディープラーニング**／支えた条件＝**データの増加と計算能力の向上** |

| 項目 | 覚える内容 |
| --- | --- |
| **AI の4つのレベル** | 1＝単純な制御（エアコンの温度制御）／2＝探索・推論（将棋プログラム）／3＝機械学習（迷惑メール判別）／4＝ディープラーニング（画像認識・生成AI） |
| **2045年問題** | **カーツワイルの予測を起点にした議論。**確定した予定ではない |

## 「よく出る誤り」を先に潰しておく

- **「AI の定義は公式に統一されている」** → 誤り。合意された唯一の定義は存在しない
- **「ディープラーニングは機械学習より広い概念」** → 誤り。入れ子の向きが逆
- **「ノーフリーランチ定理は、どの手法も性能が同じという意味」** → 誤り。**万能な1つが存在しない**という意味
- **「ドロップアウトは推論時にもノードを無効化する」** → 誤り。**学習時だけ**
- **「学習済みモデルは学習データそのものを内部に保管している」** → 誤り。保持しているのは**データから抽出された規則性（重み）**
- **「生成AI は強いAI（AGI）である」** → 誤り。**現在実用化されている AI はすべて弱いAI（ANI）**
- **「AI効果とは AI の性能が向上していく現象」** → 誤り。**知能と見なされなくなる現象**
- **「AIの冬は事故や規制が原因で訪れた」** → 誤り。**期待と実力の差**が原因
- **「学習データでの正答率が100%なら良いモデル」** → 誤り。**過学習の疑いを示す信号**として読む

- **「クラスタリングには正解ラベルが必要」** → 誤り。**教師なし学習**なのでラベルは使わない
- **「次元削減は特徴量を多くする処理」** → 誤り。**重要な情報を保ったまま特徴の数を減らす**処理

---

## サンプルはここまでです

続き(第2章〜巻末)は、生成AIパスポート 問題集のご購入で、頻出度別ノート・暗記用ノートと合わせて3点すべてダウンロードできます。

# 生成AIパスポート 学習用テキスト(無料サンプル)

**版**

2026年7月版(GUGA 公式「生成AIパスポートの試験出題範囲（2026年2月試験以降）」改訂日 2025-10-01 準拠)

**発行**

IT資格道場(https://www.shikaku-doujou.com)

**お問い合わせ**

info@shikaku-doujou.com