

1. はじめに

内容

1.1 人工知能・機械学習・深層学習

何が違うか、何ができるか

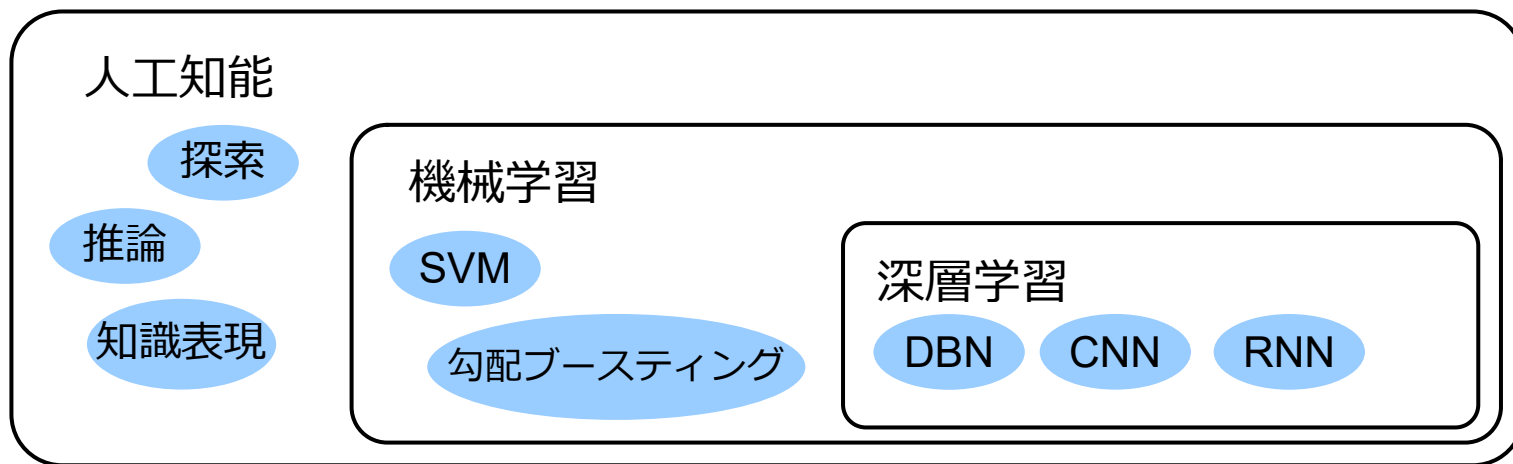
1.2 機械学習とは何か

機械学習の全体像

1.3 機械学習の分類

教師あり学習、教師なし学習、中間的学習

1.1 人工知能・機械学習・深層学習



• 人工知能とは

ある種の

- 現在、人が行っている知的な判断を代わりに行う技術
 - 技術が普及すると人工知能とはみなされなくなる
- 例) 文字認識、顔検出
- 探索・知識表現・推論・機械学習などを含む

1.1 人工知能・機械学習・深層学習

- 機械学習が注目される理由
 - ネットワーク、センサー等の発達によってビッグデータが得られるようになった
 - 計算機の高速化でビッグデータが処理可能になった
- ビッグデータは何に使えるか
 - 有用な知見の獲得
 - 省力化
 - 将来の予測



多様な趣味・嗜好に対応



安心・安全を進化

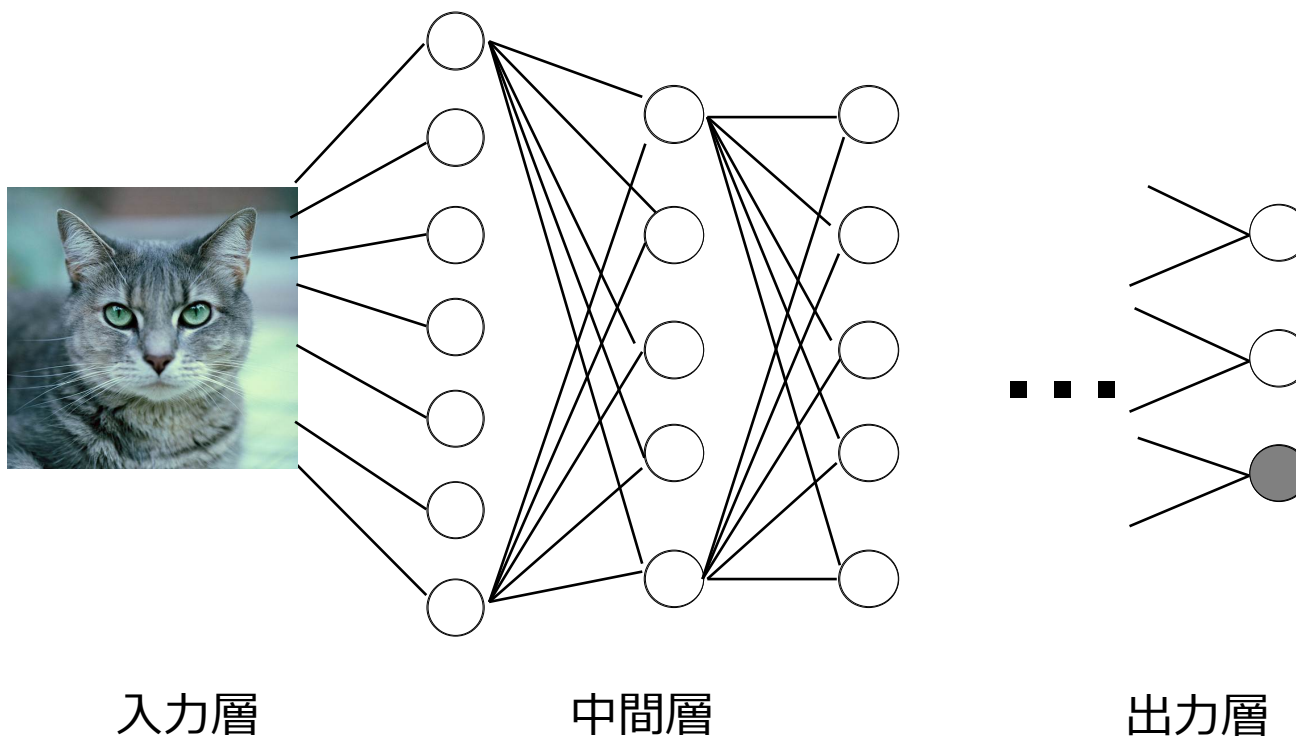


経験や勘を超越

1.1 人工知能・機械学習・深層学習

- 深層学習とは

- 多層に非線形変換を重ねる手法による機械学習
 - 特徴抽出処理も学習対象とすることができる点が特長
 - 問題に適した表現を学習しているという解釈も可能
 - 音声・画像・自然言語の認識・生成で高い性能を示す



1.2 機械学習とは何か

- 機械学習の位置づけ



数値データ

(134.1, 34.6, 12.9)

(135.5, 30.1, 43.0)

...



カテゴリデータ

(パン、ハム)

(パン、牛乳、バター)

...

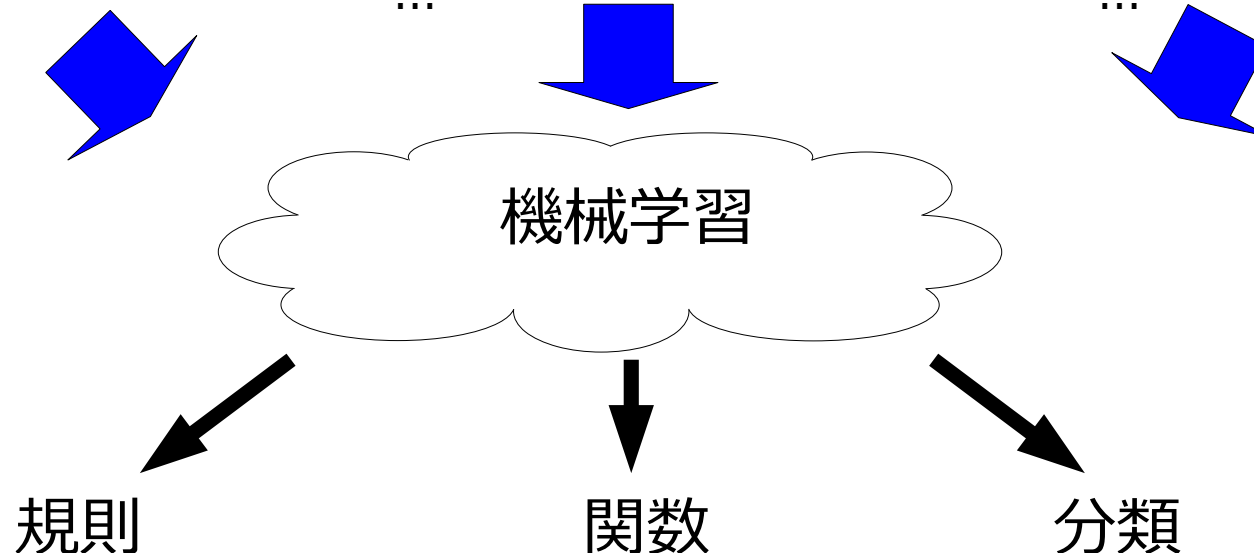


混合したデータ

(男, 28, 178, 75, yes)

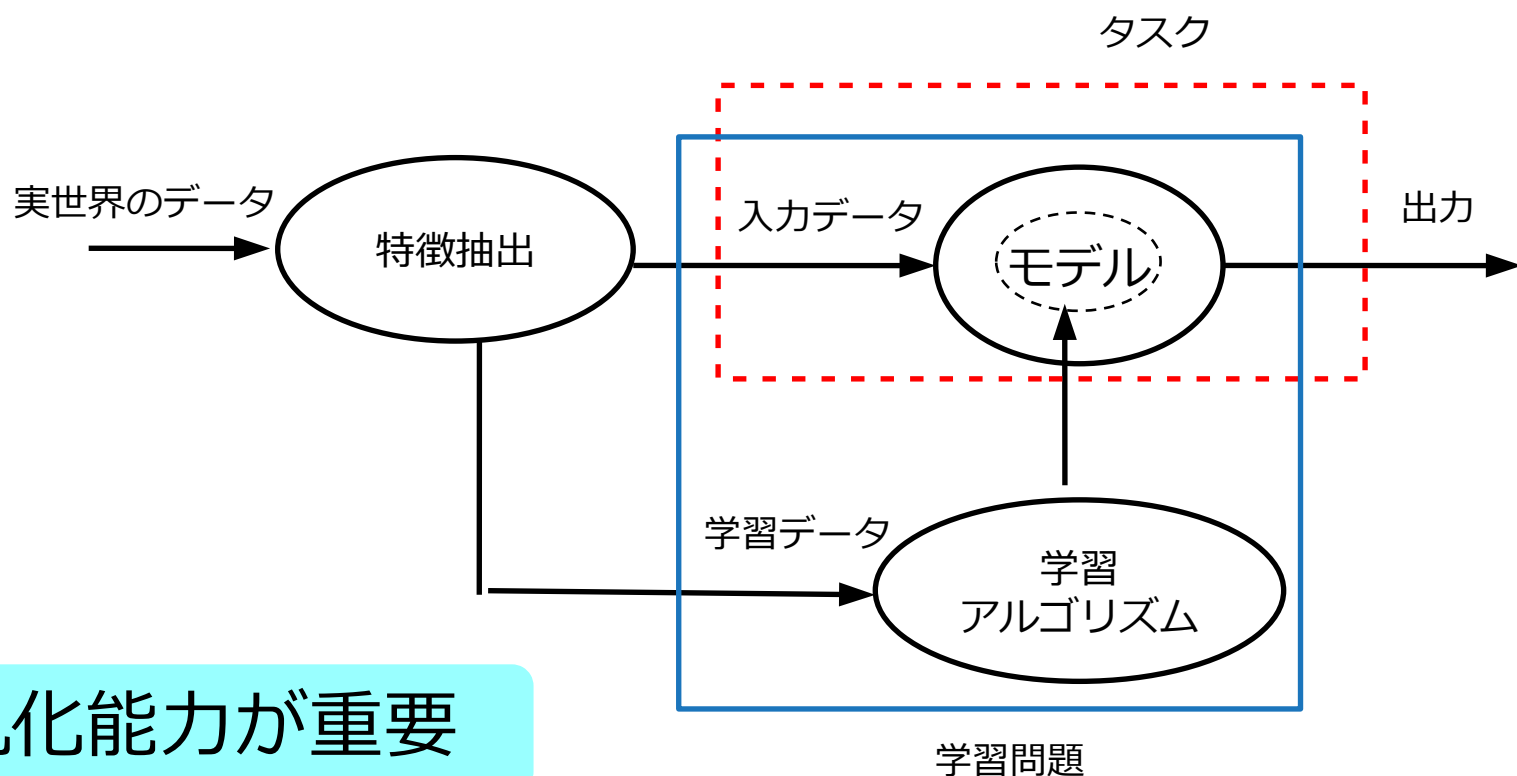
(女, 68, 165, 44, no)

...



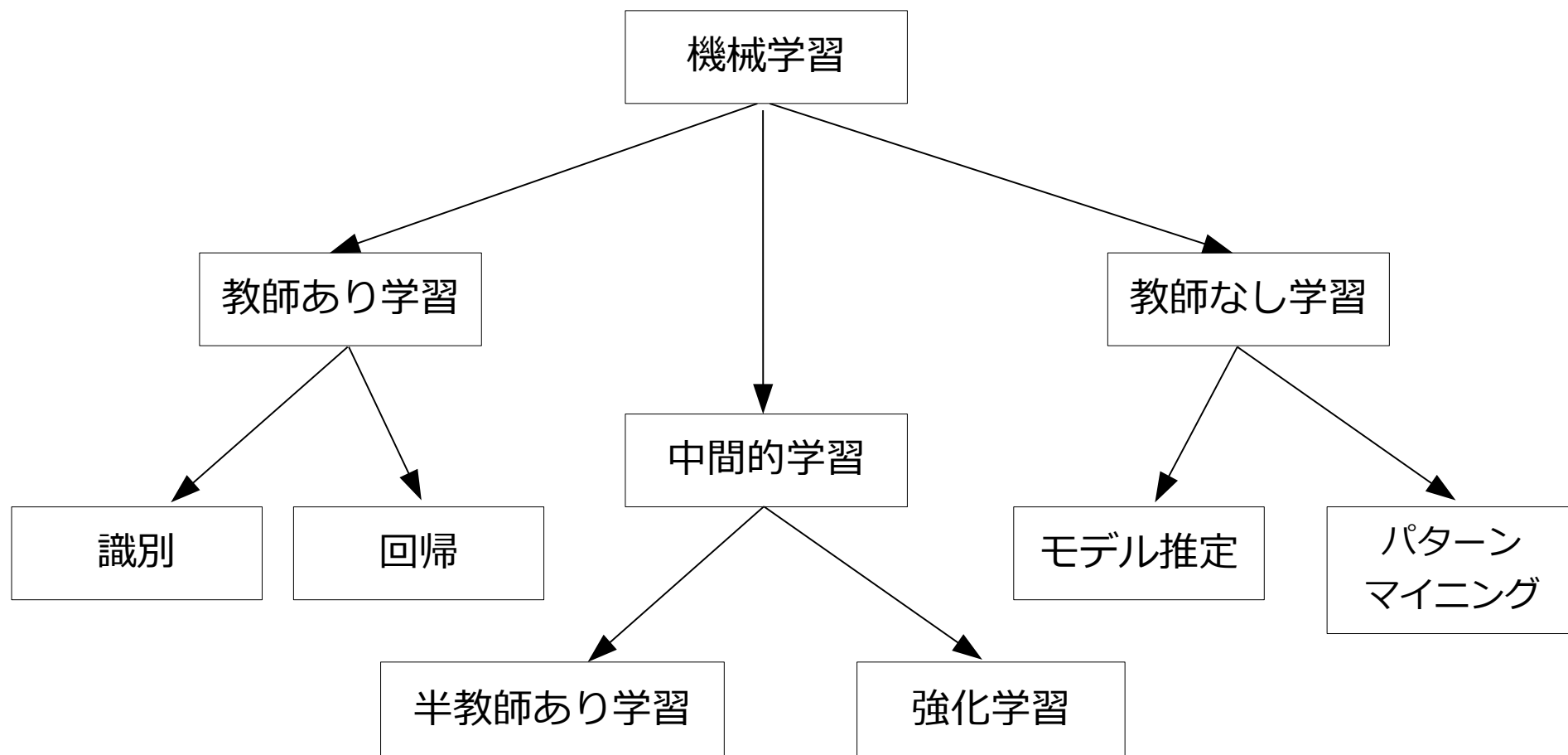
1.2 機械学習とは何か

- 機械学習とは
 - 機械学習は、適切に**タスク**を遂行する適切な**モデル**を、適切な**特徴**から構築すること [Flach 2012]



汎化能力が重要

1.3 機械学習の分類



1.3.1 教師あり学習

- 教師あり学習のデータ

- 特徴ベクトル \mathbf{x} と正解情報 y のペア

$$\{(\mathbf{x}_i, y_i)\}, \quad i = 1 \dots N$$

- 特徴ベクトルは次元数 d の固定長ベクトル

$$\mathbf{x}_i = (x_{i1}, \dots, x_{id})^T$$

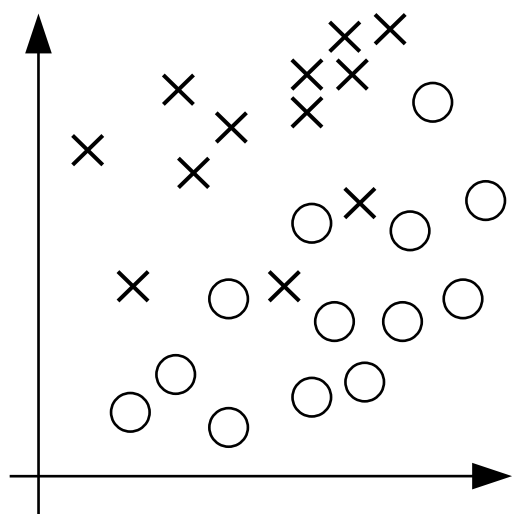
T : 転置記号

- 特徴ベクトルの各要素は数値またはカテゴリ
 - カテゴリデータの例: 性別、職業、天候、 etc.
- 正解情報の型によって問題が分かれる
 - カテゴリデータ: 識別
 - 数値データ: 回帰

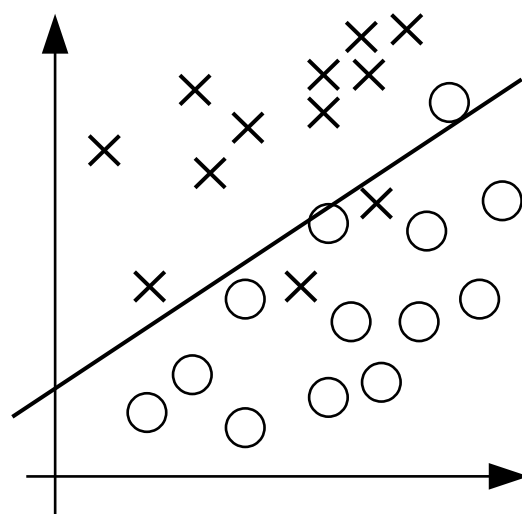
1.3.1 教師あり学習

- 識別

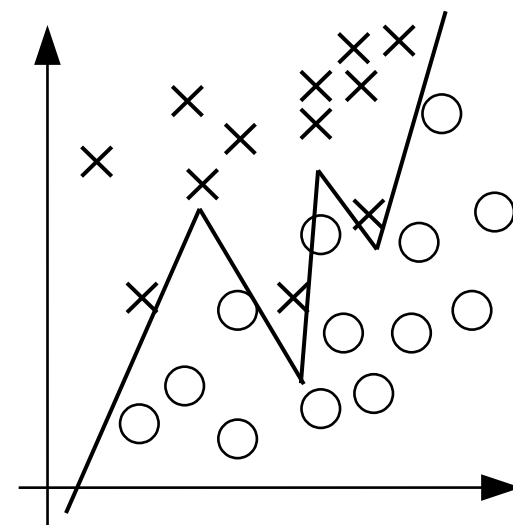
- 正解情報がカテゴリデータ
- 未知データに対する誤りが最小となるような特徴空間上の識別面を求める



(a) 入力が2次元数値ベクトルの識別問題



(b) 学習結果 1

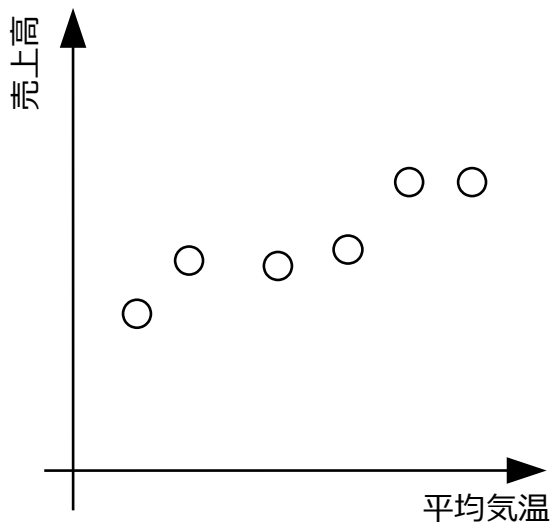


(c) 学習結果 2

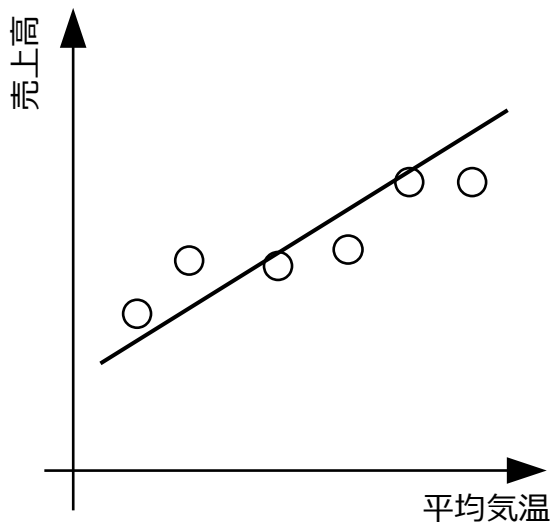
一般化という視点でどちらが適しているか

1.3.1 教師あり学習

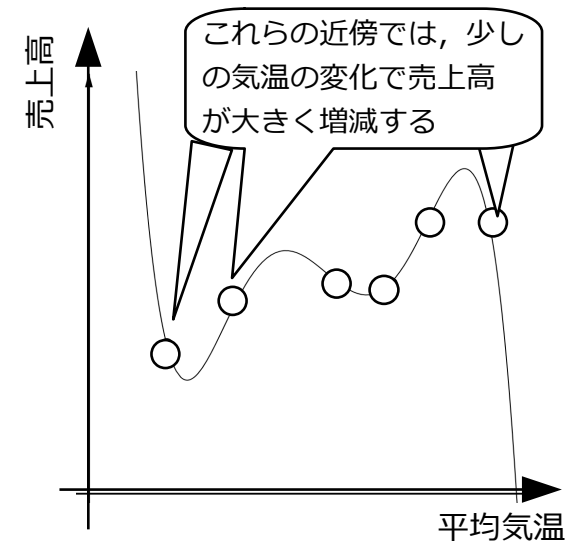
- 回帰
 - 正解情報が数値データ
 - 汎化誤差が最小となるような近似関数を求める



(a) 過去の平均気温と
売上高の関係



(b) 1 次式による回帰



(c) 高次の式による回帰

一般化という視点でどちらが適しているか

1.3.2 教師なし学習

- 教師なし学習のデータ

- 特徴ベクトル \mathbf{x} のみ

$$\{\mathbf{x}_i\}, \quad i = 1 \dots N$$

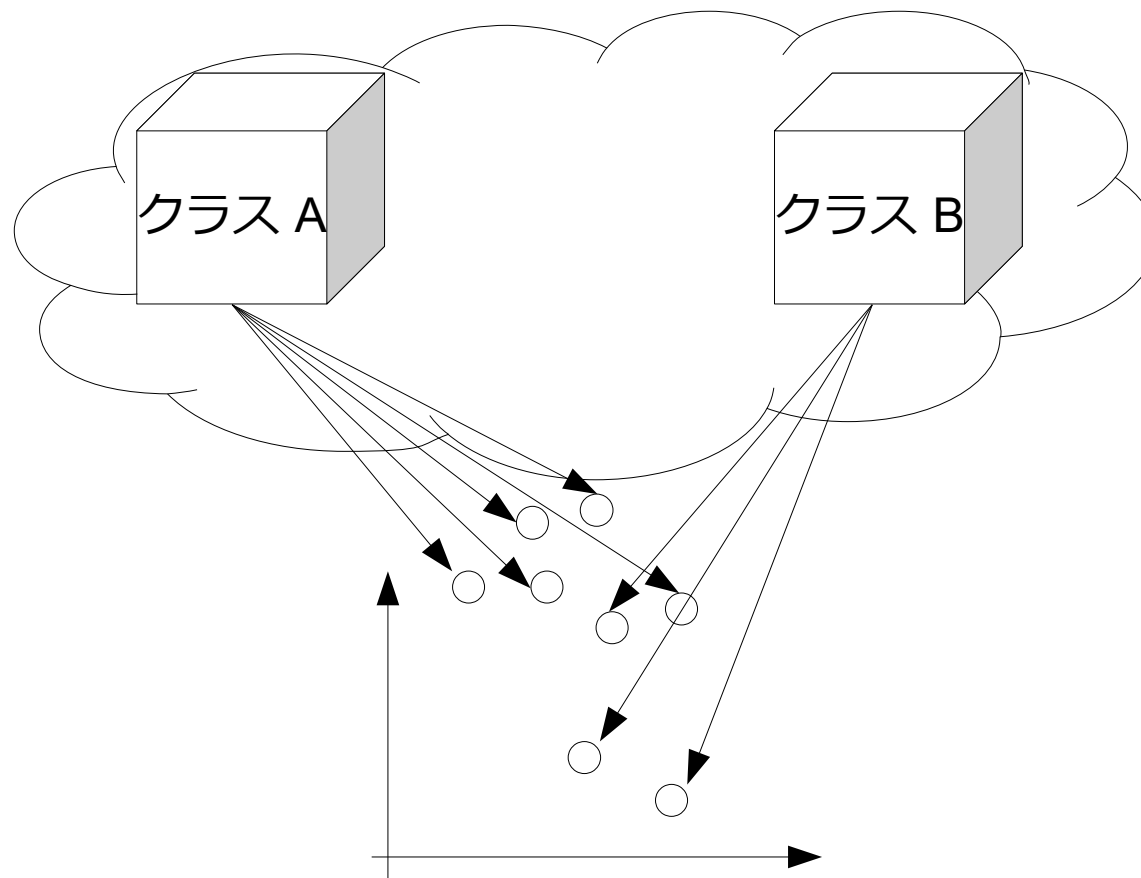
- 特徴ベクトルは次元数 d の固定長ベクトル

$$\mathbf{x}_i = (x_{i1}, \dots, x_{id})^T$$

- 基本的にデータに潜む規則性を学習
- 規則がカバーする範囲によって問題が分かれる
 - データ全体をカバー：モデル推定
 - 頻出する傾向を発見：パターンマイニング

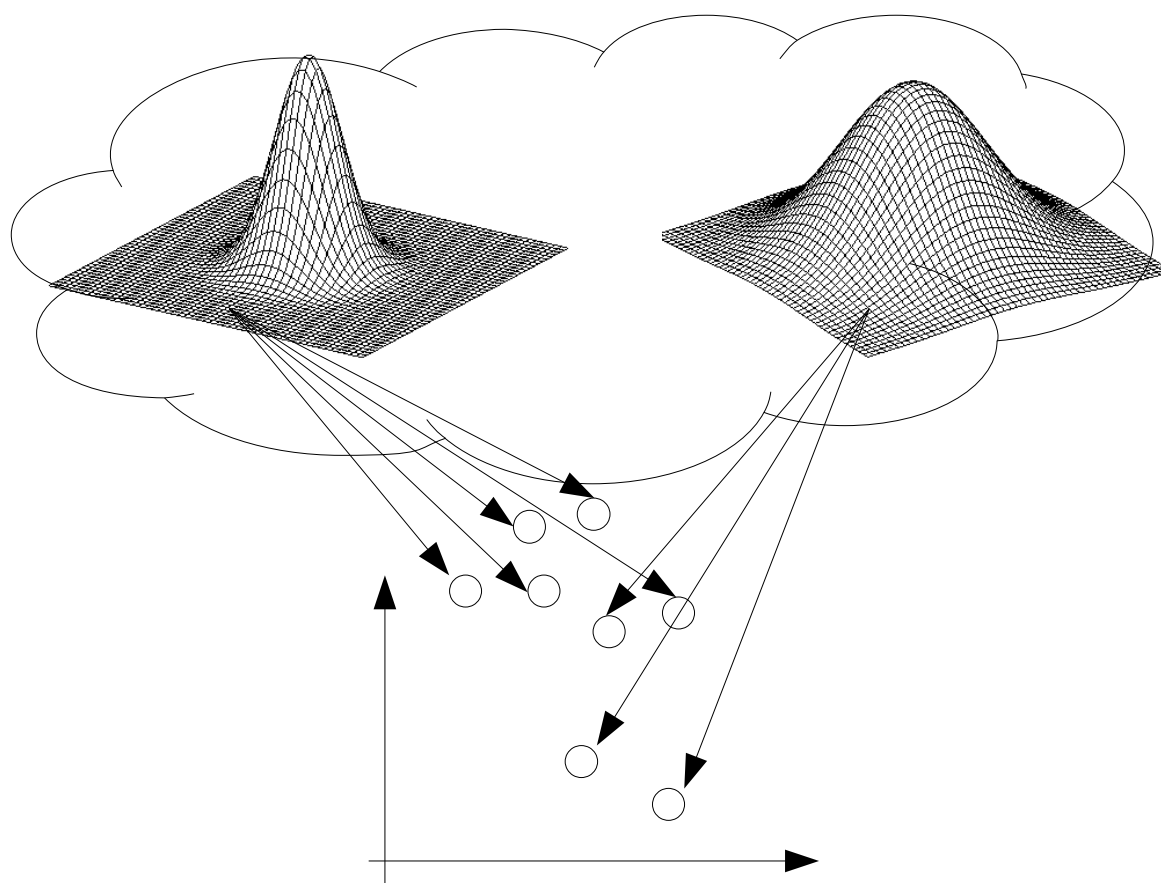
1.3.2 教師なし学習

- モデル推定
 - データを生じさせたクラスを推定
 - 特徴ベクトルは主として数値データ



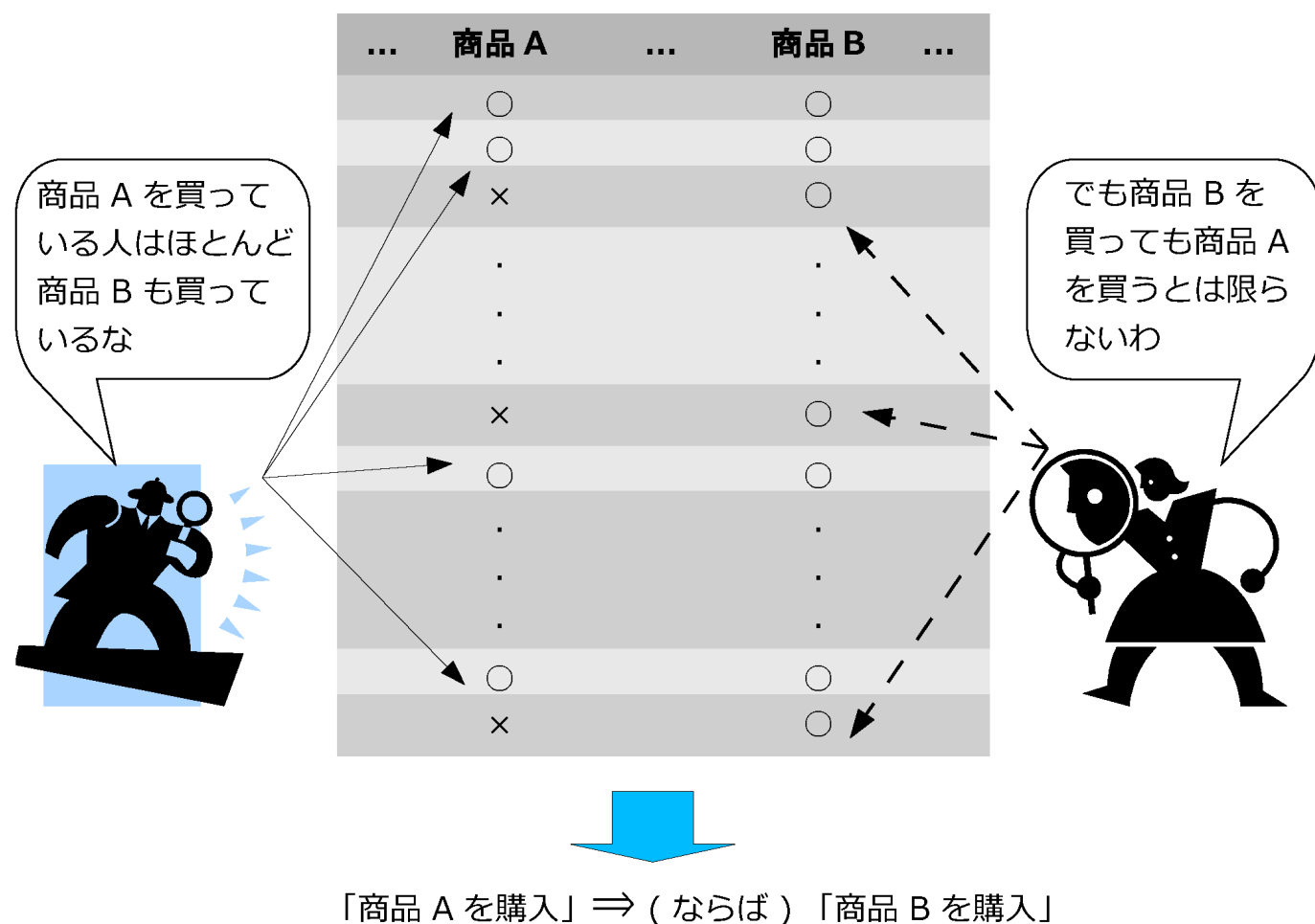
1.3.2 教師なし学習

- モデル推定
 - クラスの分布も推定



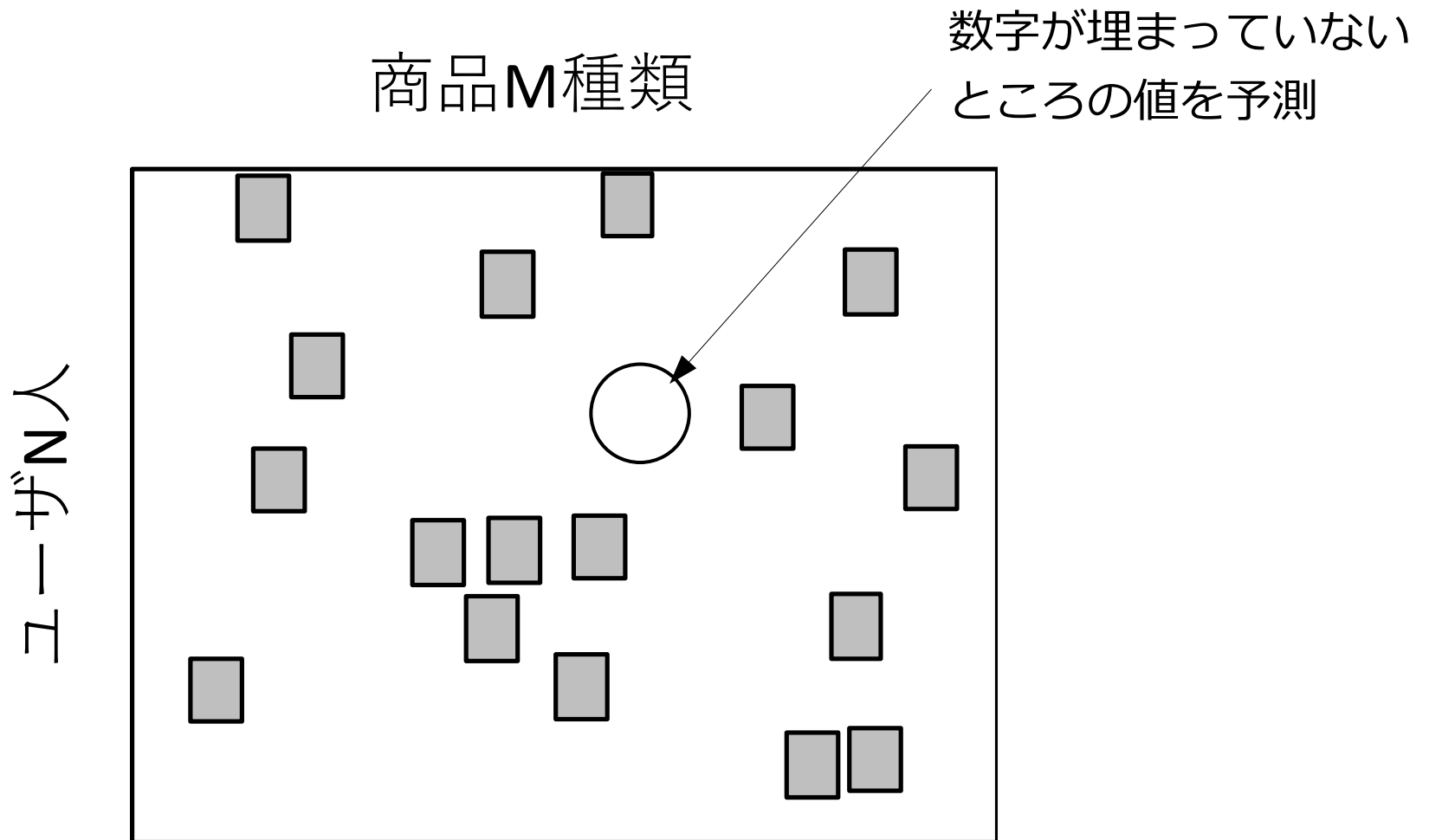
1.3.2 教師なし学習

- パターンマイニング
 - 頻出項目や隠れた規則性を発掘
 - 特徴ベクトルは主としてカテゴリデータ



1.3.2 教師なし学習

- 推薦システム
 - 特徴は表面的には数値、実質カテゴリ

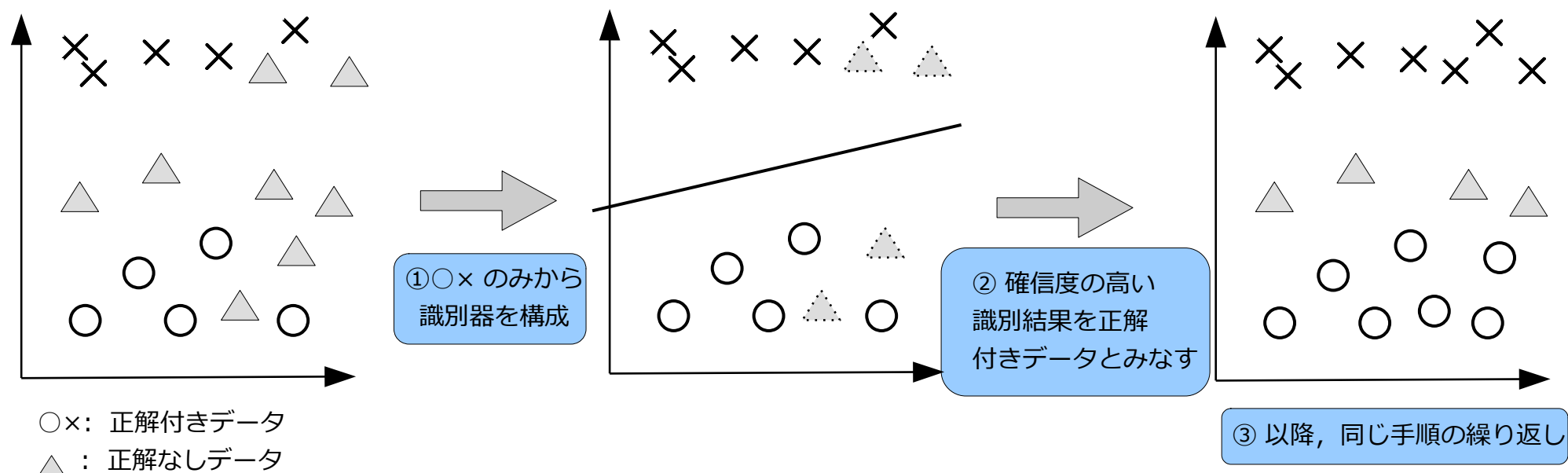


1.3.3 中間的学習

- データが正解付き／なしの組み合わせ
 - 半教師あり学習に適した状況
 - 正解付きの少量のデータ
 - 正解なしの大量のデータ
 - 強化学習
 - 正解情報が、ときどき報酬という形式で与えられる

1.3.3 中間的学習

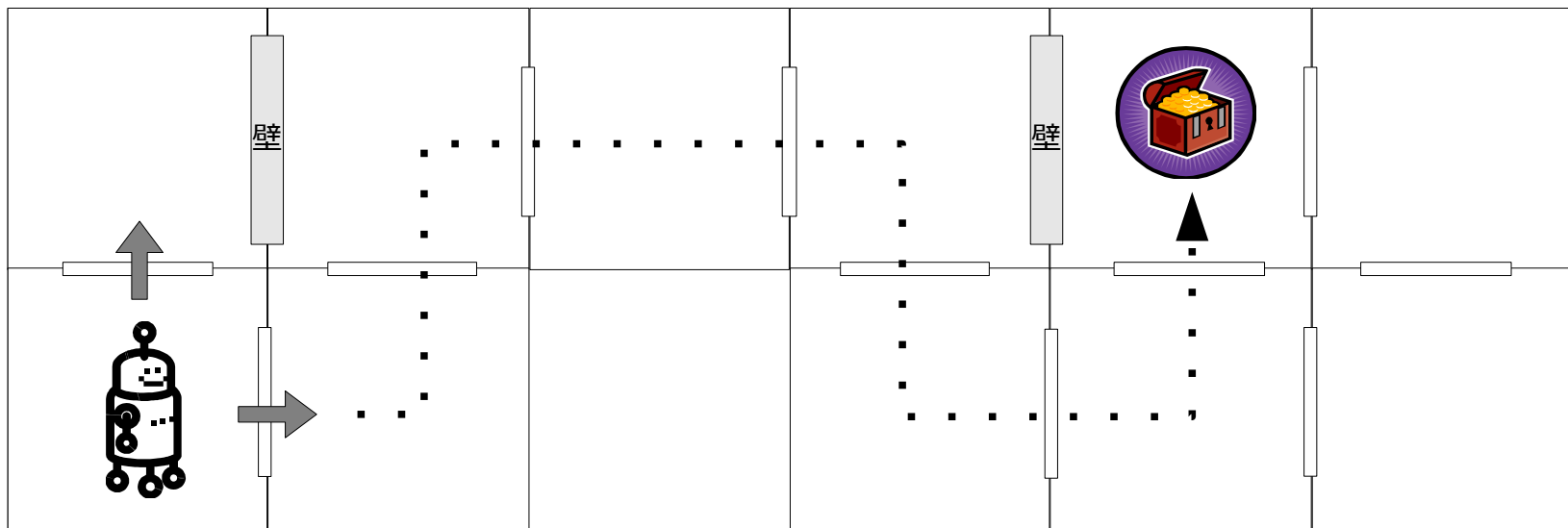
- 半教師あり学習
 - 繰り返しによる学習データの増加



1.3.3 中間的学習

- 強化学習

- 教師信号が、間接的に、ときどき、確率的に与えられる



まとめ

- 人工知能 ⊃ 機械学習 ⊃ 深層学習
- 機械学習とは
 - 適切にタスクを遂行する適切なモデルを、適切な特徴から構築すること
- 機械学習の分類
 - 教師あり・教師なし・中間的
- 推奨資料
 - オンライン講座 Udemy はじめての AI (54 分、無料、要登録)
 - <https://www.udemy.com/google-jp-ai/>
 - 荒木雅弘．マンガでわかる機械学習，オーム社，2018.
 - 谷口忠大．イラストで学ぶ人工知能概論 改訂第 2 版．講談社，2020.