

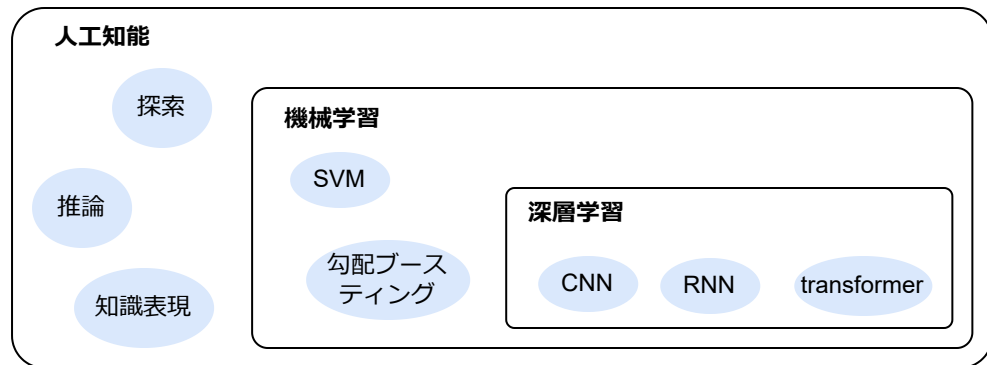
# はじめに

## 本章の内容

- 人工知能・機械学習・深層学習
  - 何が違うか、何ができるか
- 機械学習の分類
  - 教師あり学習、教師なし学習、中間的学習
- 近年の人工知能の話題

# 人工知能・機械学習・深層学習

- 人工知能・機械学習・深層学習の関係
  - 人工知能 ⊃ 機械学習 ⊃ 深層学習



# 人工知能・機械学習・深層学習

- 人工知能の定義
  - **現在**、人が行っている知的な判断を代わりに行う技術
    - 事例：自動運転、サービスロボット、医療補助
    - 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
- 人工知能の要素技術
  - 探索・知識表現・推論・機械学習など

# 人工知能・機械学習・深層学習

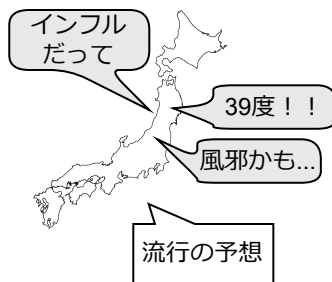
- 機械学習
  - 「知的な判断を行う」技術を、データから規則性を導くことによって実現する方法
  - ビッグデータが利用可能になったことが背景



多様なニーズに対応



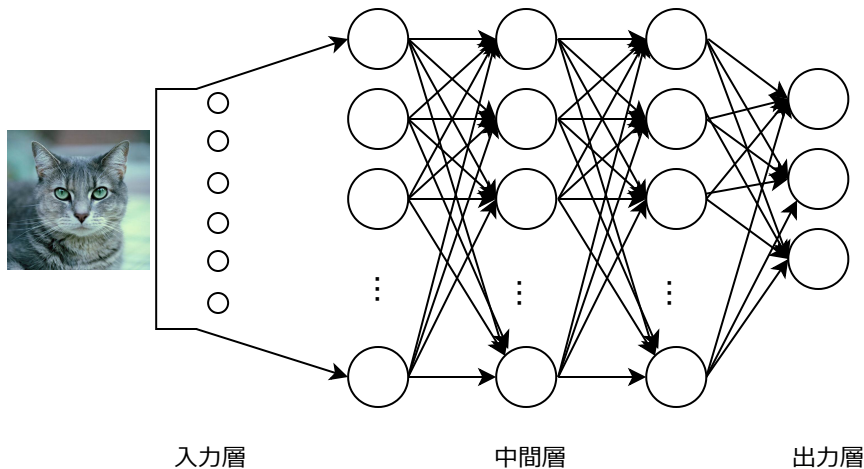
安心・安全を進化



経験や勘を超越

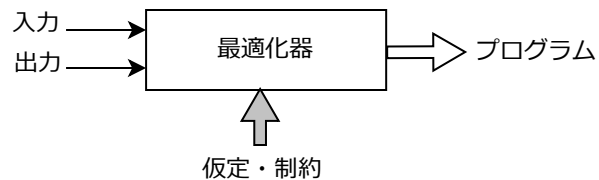
# 人工知能・機械学習・深層学習

- 深層学習
  - 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
  - 特に音声・画像・自然言語の認識・生成で高い性能を示す



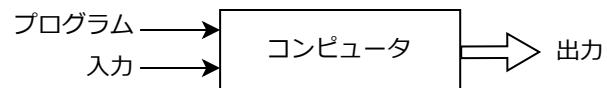
## 1.2 機械学習とは何か

- 機械学習のさまざまな側面
  - 入出力から見た機械学習
  - 最適化問題としての機械学習
  - オープンな問題としての機械学習



## 1.2 機械学習とは何か

- 入出力から見た機械学習の定義
  - 普通の情報処理システム／従来の人工知能システム
    - 入力から出力を得るプログラムを人手で記述
  - 機械学習
    - 入力から出力を予測するプログラムを自動で生成

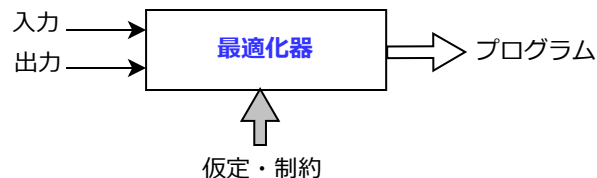


## 1.2 機械学習とは何か

- 最適化問題としての機械学習の定義
  - プログラムとして、入力  $\mathbf{x}$  から出力  $\hat{y}$  を求める数理モデル  $f$  を設定

$$\hat{y} = f(\mathbf{x}; \theta)$$

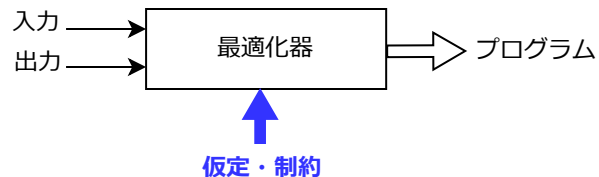
- ただし  $\theta$  はモデルのパラメータ
- 正解  $y$  と出力  $\hat{y}$  から定義される損失関数  $L(y, \hat{y})$  の値が最小となるように、パラメータ  $\theta$  を最適化





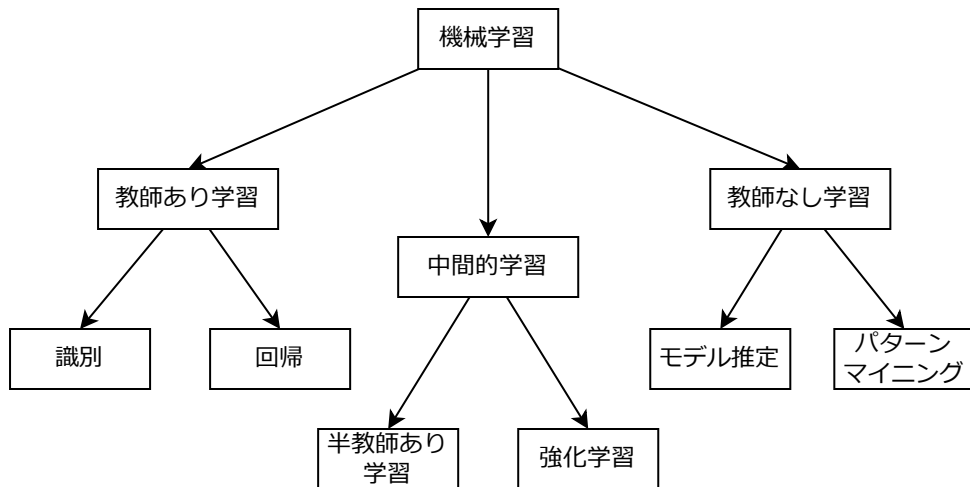
## 1.2 機械学習とは何か

- オープンな問題としての機械学習の定義
  - 今後入ってくる未知の入力に対して、正しい出力を得るための数理モデルを、仮定・制約に基づいて決める
  - 仮定の例
    - 入力の微少な変化に対して、出力は大きくは変化しない
  - 制約の例
    - 出力は社会的偏見を反映してはならない



# 1.3 機械学習の分類

- データへの正解の有無や出力の型で分類できる



## 1.3.1 教師あり学習

- 教師あり学習のデータ
  - 入力ベクトル  $\boldsymbol{x}$  と正解情報  $y$  のペア

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

- 入力ベクトルは次元数  $d$  の固定長ベクトル

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

- 入力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、居住地、天候、etc.
- 正解情報の型によって問題が分かれる
  - カテゴリ: 識別
  - 連続値: 回帰

# 1.3.1 教師あり学習

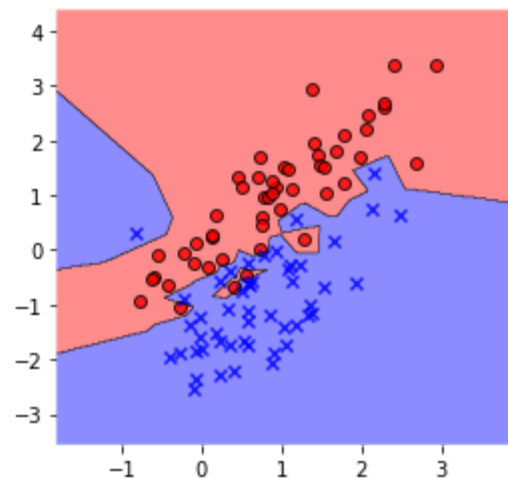
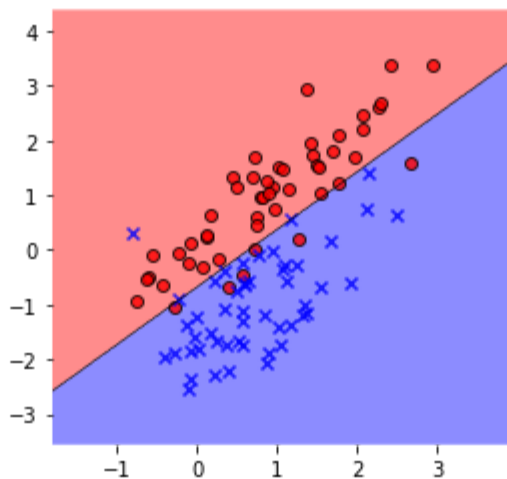
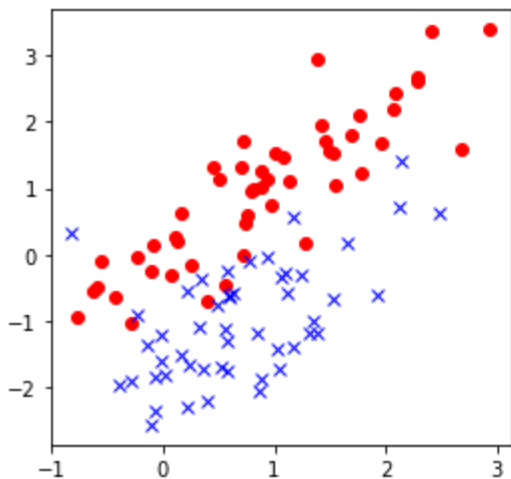
- 識別

- 正解情報がカテゴリ

- 例) 感染の判定: 陽性, 陰性

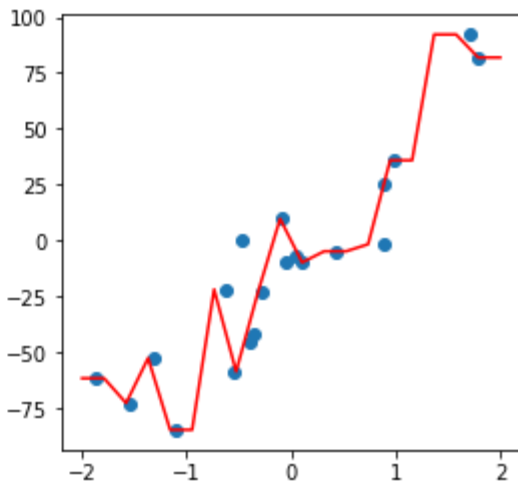
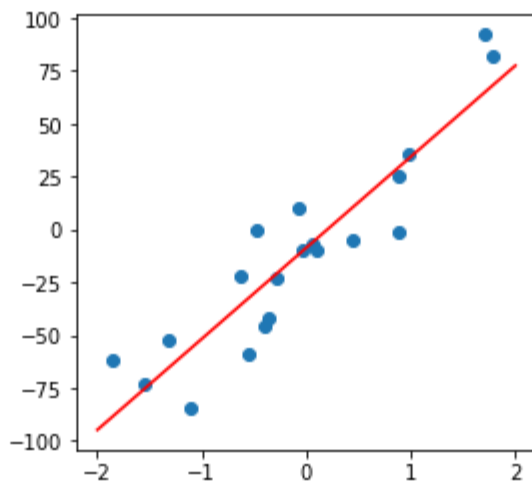
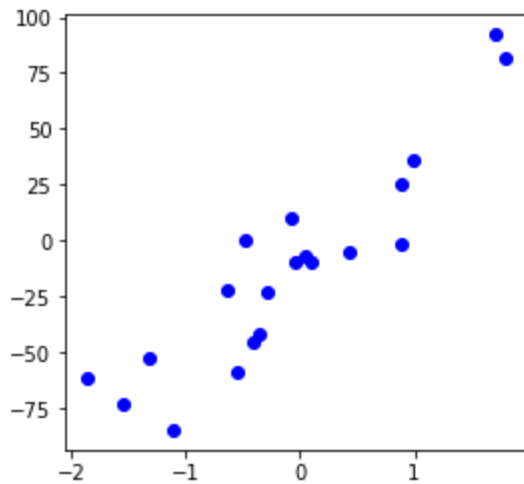
- 未知データに対する誤りが最小となるような入力空間上の識別面を求める

- どちらの識別面が未知データに対してうまく識別できそうか



## 1.3.1 教師あり学習

- 回帰
  - 正解情報が連続値
  - 汎化誤差が最小となるような近似関数を求める
    - どちらの関数が未知データに対してうまく予測できそうか



## 1.3.2 教師なし学習

- 教師なし学習のデータ
  - 入力ベクトル  $\boldsymbol{x}$  のみ

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

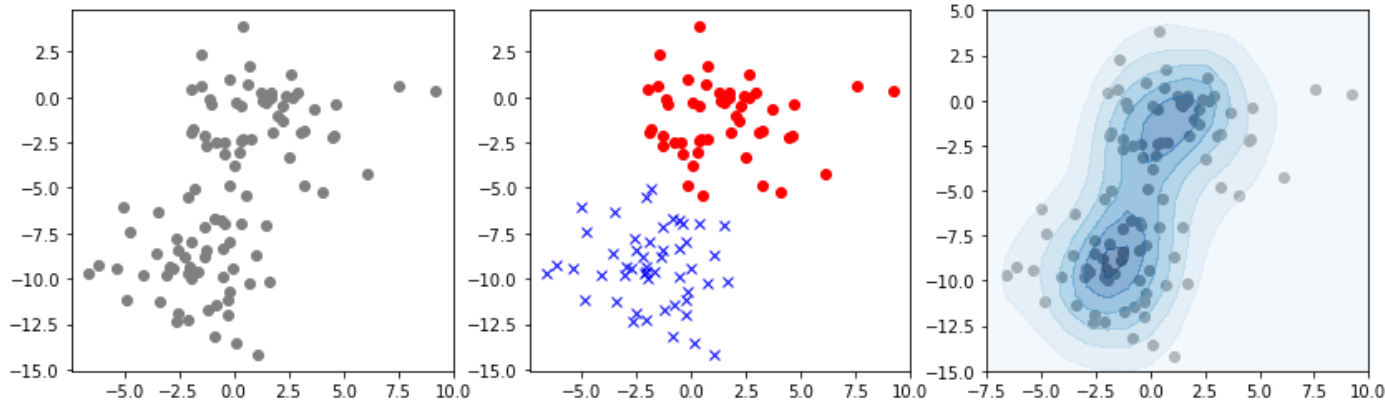
- 入力ベクトルは次元数  $d$  の固定長ベクトル

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

- 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

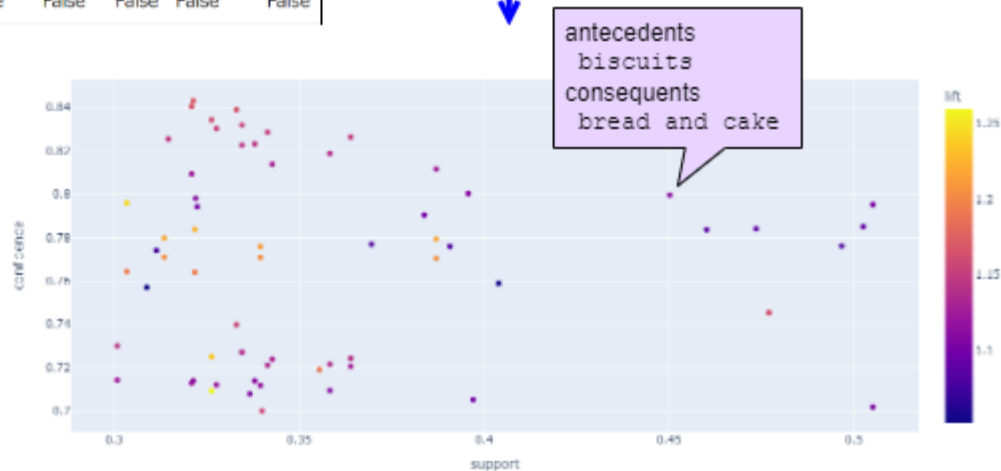
- モデル推定
  - 入力ベクトルは主として数値データ
  - クラスタリング: データをまとまりに分割する
    - データを生じさせたクラスを推定
- モデル推定
  - クラスの確率分布を推定



## 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
  - 入力ベクトルは主としてカテゴリデータ

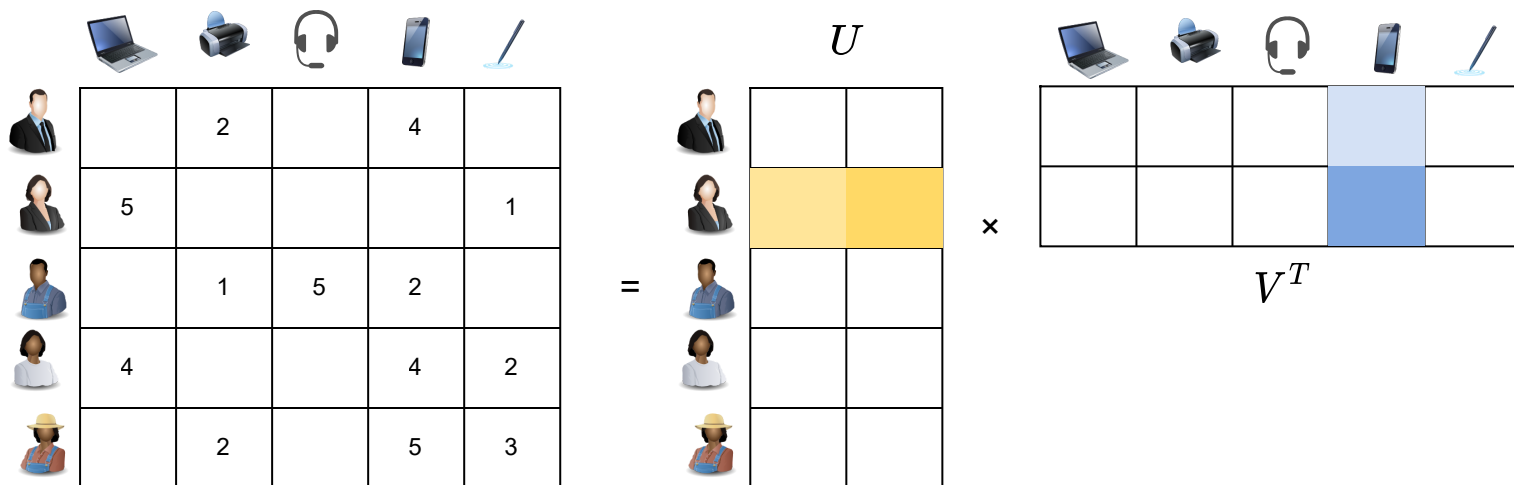
|     | grocery<br>misc | baby<br>needs | bread<br>and<br>cake | baking<br>needs | coupons | juice-<br>sat-<br>cord-<br>ms | tea   | biscuits |
|-----|-----------------|---------------|----------------------|-----------------|---------|-------------------------------|-------|----------|
| 0   | False           | True          | True                 | True            | False   | True                          | False | True     |
| 1   | False           | False         | False                | False           | False   | False                         | False | False    |
| 2   | False           | False         | True                 | True            |         |                               |       |          |
| 3   | False           | False         | True                 | True            |         |                               |       |          |
| 4   | False           | False         | True                 | True            |         |                               |       |          |
| ... | ...             | ...           | ...                  | ...             | ...     | ...                           | ...   | ...      |





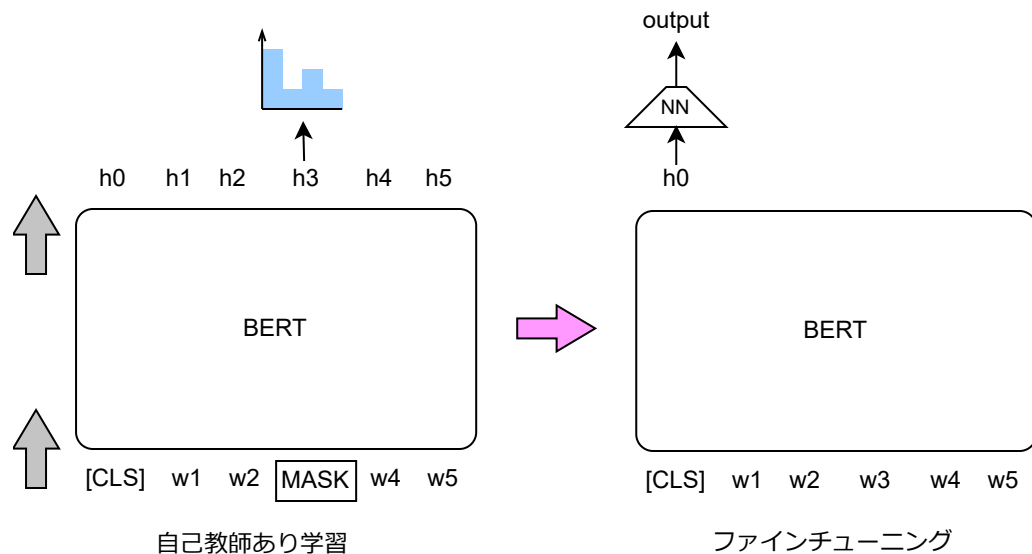
## 1.3.2 教師なし学習

- 推薦システム
  - 入力は表面的には数値、実質カテゴリ



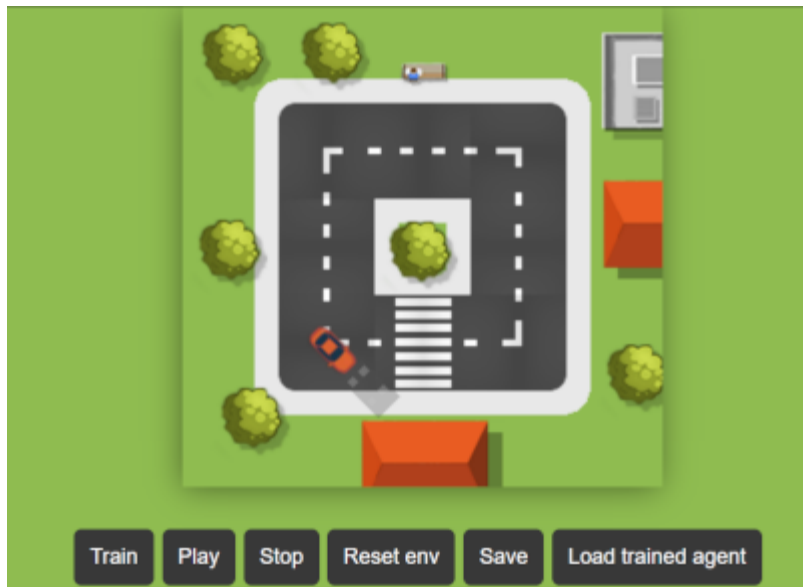
# 1.3.3 中間的学習

- 半教師あり学習: データが正解付き／なしの組み合わせ
- 自己教師あり学習
  - 一部の入力信号を隠して、それを復元するタスクで表現学習
  - 表現学習の後、少量のタスクデータでファインチューニング



## 1.3.3 中間的学習

- 強化学習
  - 遷移する状態における最適な行為を学習
  - 正解情報が間接的・確率的に与えられる



- <https://metacar.scottpletcher.guru/>

# 1.3 機械学習の分類

## まとめ

- 人工知能 ⊃ 機械学習 ⊃ 深層学習
- 機械学習とは
  - あらかじめ設定された仮定・制約を満たす数理モデルのパラメータを、入力と出力がペアになったデータ(あるいは入力だけのデータ)を用いて最適化する方法
- 機械学習の分類
  - 教師あり・教師なし・中間的

## 推奨資料

- 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.