

4. 機械学習の基礎

17t4055g 築地俊平

4.機械学習の基礎

本章で取り上げる内容

- 分類や回帰にとどまらない機械学習の形式
- 機械学習モデルの正式な評価手続き
- ディープラーニング用のデータの準備
- 特徴エンジニアリング
- 過学習の対処
- 機械学習の問題に取り組むための一般的なフレームワーク

4.1 機械学習の4つの手法

- これまでの例

- 二値分類、他クラス分類、スカラー回帰

- これらはすべて教師あり学習 (supervised learning)

教師あり学習

- 訓練の入力値と目的値の関係を学習する

4.1 機械学習の4つの手法

4つの手法

- 教師あり学習
- 教師なし学習
- 自己学習
- 強化学習

4.1 機械学習の4つの手法

- 教師あり学習

→一連のサンプルに基づいて、入力データを既知の目的地に写像するための学習

目的値はアノテーション(annotation)と呼ばれる

光学文字認識や音声認識、画像分類、言語翻訳などのアプリケーションはほぼ例外なく、教師あり学習

4.1 機械学習の4つの手法

- 教師あり学習

ほとんどは分類と回帰で構成されているが、変わった種類もある

シーケンス生成

→与えられた画像を説明するキャプションを予測する。

構文木予測

→与えられた文章から構文木への分解を予測

4.1 機械学習の4つの手法

物体検出

→与えられた画像内の特定のオブジェクトの周りに教会を描く。
分類問題や同時分類/回帰問題として表現される。

画像分割

→与えられた画像の特定のオブジェクトをピクセルレベルでマスクする

4.1 機械学習の4つの手法

- 教師なし学習

目的値の助けを借りずに、入力データの重要な変換を見つけ出す。

教師なし学習では、次元削除(dimensionality reduction)とクラスタリング(clustering)のカテゴリがよく知られている。

次元削除→高次元のデータからできるだけ情報を保持できるように
低次元のデータに変換する

クラスタリング→要素を分ける

4.1 機械学習の4つの手法

- 自己学習

アノテーションラベルのない教師あり学習
オートエンコーダ(autoencoder)が一例

教師あり学習、教師なし学習、自己学習の違いは曖昧

4.1 機械学習の4つの手法

- 強化学習

エージェント(agent)がその環境に関する情報を受け取り、何らかの報酬が最大になるような行動の選び方を学習する

囲碁などのゲームで成果を上げてる