

# 3.1 教師なし学習の種類

欧陽恵子

- データセットの教師なし変換
- クラスタリングアルゴリズム

# データセットの教師なし変換

もとのデータ表現を変換して、人間や他の機械学習アルゴリズムにとって、より分かりやすい新しいデータ表現を作るアルゴリズムのことだ。

# 応用

- 教師なし変換の利用法として最も一般的なのは次元削減だ。
- もう一つの教師なし変換のアプリケーションとして、そのデータを「構成する」部品、もしくは成分を見つけることが挙げられる。

# クラスタリングアルゴリズム

データを似たような要素から構成されるグループに分けるアルゴリズムだ。

クラスタリングとは、分類対象の集合を、内的結合と外的分離が達成されるような部分集合に分割すること。

# 応用

SNSサイトに写真をアップロードすることを考えてみよう。  
SNSサイトは、同じ人物が写っている写真をまとめること  
で整理しようとする。