

情報リテラシー(第11回 後期)ハンドアウト

機械学習の基礎—3つの学習方法—

1. 今日のねらい

- 機械学習の全体像を理解する
 - 教師あり・教師なし・強化学習の違いを説明できる
 - それぞれの具体例を知る
-

2. 機械学習とは

定義: データから学習し、予測や判断を行う技術

従来との違い: 従来はルールを人間が書く → 機械学習はデータからルールを学習

3. 機械学習の3つの分類

教師あり学習: 正解ラベル付きデータで学習 → 予測

教師なし学習: 正解ラベルなしでパターン発見 → グループ化

強化学習: 試行錯誤で最適な行動を学習 → 報酬で改善

4. 教師あり学習

データ: 説明変数 + 正解ラベル

目的: 新しいデータの予測

タイプ: 回帰（数値予測）、分類（カテゴリ予測）

例: スパム判定、画像認識、株価予測、本墨打30本以上予測

デシジョンツリー: 質問を繰り返して分類。視覚的で判断理由が明確

5. 教師なし学習

データ: 説明変数のみ（正解ラベルなし）

目的: パターン発見

クラスタリング: 似ているものをグループ化

例: 顧客セグメント、異常検知、レコメンド

K-means法: K個のグループを作り、最も近いグループに割り当て、中心を再計算

6. 強化学習

定義: 試行錯誤で最適な行動を学習

仕組み: 行動 → 観察 → 報酬 → 学習 → 繰り返し

例: AlphaGo（囲碁AI）、ロボット制御、自動運転

7. 3つの比較

	教師あり	教師なし	強化学習
データ	正解あり	正解なし	報酬あり
目的	予測	パターン発見	最適行動
例	スパム判定	顧客分類	ゲームAI

使い分け: 予測→教師あり／パターン発見→教師なし／行動最適化→強化学習

8. 今日のColab実習

Part 1: デシジョンツリー (打率+打点 → 本塁打30本以上予測)

Part 2: クラスタリング成功例 (打率+打点 → 3グループ)

Part 3: クラスタリング失敗例 (打率+出塁率 → 3グループ)

Part 4: 比較 (なぜ結果が違う?)

9. 変数選択の重要性

成功例: 打率+打点 → グループ1が100%完璧に30本以上発見

失敗例: 打率+出塁率 → グループ1が90%程度

理由: 打率と出塁率は似た変数 (相関0.875) → 新情報が少ない

教訓: 相関が強く、異なる情報を持つ変数を組み合わせる

10. まとめ (キーワード)

機械学習／教師あり学習／教師なし学習／強化学習／回帰／分類／デシジョンツリー／クラスタリング／K-means法／変数選択／相関係数