

施策デザインのための 機械学習入門2

廣田雄亮

1章 機械学習実践の ためのフレームワーク

1.1 機械学習の実践に潜む 落とし穴

1.1.1 ビジネスにおける機械学習の実践 1/2

より良い性能を発揮する施策を追い求めて、データやアルゴリズム、機械学習をビジネスに活用する。

施策：KPIの最大化・最小化を目標とするアクション。
□ジック。

KPI：Key Performance Indicator.

例 売上, アクティブユーザ数

性能：施策を導入して得られるKPIの値

1.1.1 ビジネスにおける機械学習の実践 2/2

意図した性能が発揮できず、その原因が分からない



本来踏むべきステップを多く見逃している
無意識に落とし穴に放っている

1.1.2 機械学習の実践に潜む落とし穴 1/2

- 解くべき問題の誤設定
- 過去の施策の傾向に起因するバイアス

といった落とし穴がある

既に説明済み

1.1.2 機械学習の実践に潜む落とし穴 2/2

本来目的変数としたい情報 とあたかも目的変数のように見える表層的なデータの **違い**を見逃すことがある

具体例

「動画配信サービスにおける関連動画推薦」

ユーザが動画に興味を持っているか否か (本来)

≠

動画をクリックするか否か

1.2 機械学習のための フレームワーク

フレームワーク

上流

KPIの設定

データの観測構造のモデル化

解くべき問題の特定

観測データのみを用いた
問題の解決方法を考える

機械学習モデルの学習

施策の導入

下流

1.2.1 KPIの設定

施策の性能を測るための**定量指標 (KPI)** を設定する

例

「できるだけ多くのユーザにサービスを利用してもらう」

⇒ サービスの利用継続年数

「どの定量指標をビジネス目標やサービスの
ビジョンのKPIに設定するか」

最終的なビジネス目標の達成度合いに寄与する

1.2.2 データの観測構造のモデル化 1/2

手元にあるデータがどのような経緯をたどって観測されたのかを**モデル化** (記号, 関係式による表現)する

モデル化する



観測データに潜む**バイアス**に気づき,
バイアスに対応する法則を見出すことができる.

例

「動画配信サービスにおける関連動画推薦」において

- ポジションバイアス

UIの位置 (ポジション) が原因で、分析者が真に追いたい目的変数(ユーザが興味があるか)とクリック数などの表層的に出現する情報との間に発生する乖離

⇒ UIを変更する・対応するモデルを用いる

1.2.3 解くべき問題の特定

解くべき問題を特定する

= モデルパラメータを学習するための
目的関数の定義

機械学習モデルを決める

予測モデル or 意思決定モデル or 分類モデル

1.2.4 観測データを用いて解くべき問題を近似する 1/2

問題を手元の観測データのみを用いて近似する

目的関数を定義した後、性能が良くなるパラメータ ϕ を得たい

しかし、手元のデータ (母集団のサンプル) であるため、母集団の分布が分からない。

直接 ϕ を得ることは現実的には困難

⇒ 教師あり機械学習の場合、経験損失最小化 (ERM) を採用する

経験損失最小化 (Empirical Risk Minimization; ERM)

「手元に観測されている有限個のデータから**推定**して、それを代わりに目的関数とする」

損失関数とデータについて求めた経験損失 (標本誤差) を最小化するようにモデルを学習の基準

例

平均二乗誤差, エントロピー誤差を最小化する

1.2.4 観測データを用いて解くべき問題を近似する 2/2

経験性能を真の性能の代替として採用するには、
一貫性の過程が大前提である。

一貫性の過程：観測データ数 (n) が増えるにしたがって、
経験性能が真の性能に近づいていくと
いう仮定

経験性能：観測データ上での単純な平均により
計算される性能

1.2.5 機械学習モデルを学習する

観測データを用いて近似された目的関数について、
できるだけよい値をとるパラメータ ϕ を取る

前提

- データの観測構造のモデル化
- 解くべき問題の特定

といった上流のステップがクリアされていること

1.2.6 施策を導入する

学習された予測モデルや意思決定モデルを駆使した施策を本番環境に導入する.

重要な視点

「現在稼働している施策が、次に機械学習モデルを学習する際に用いるデータを生成する」

「機械学習モデルを学習しやすいデータを将来のために残しておく」

サンプルコード

実行環境

Python

Jupyter notebook

サンプルコード

https://github.com/ghmagazine/ml_design_book