

脱！ブラックボックス 説明可能なAI

第4回

入力特徴量が多くてもOK！
モデルに依存しない説明手法…PDPlot

切通 恵介

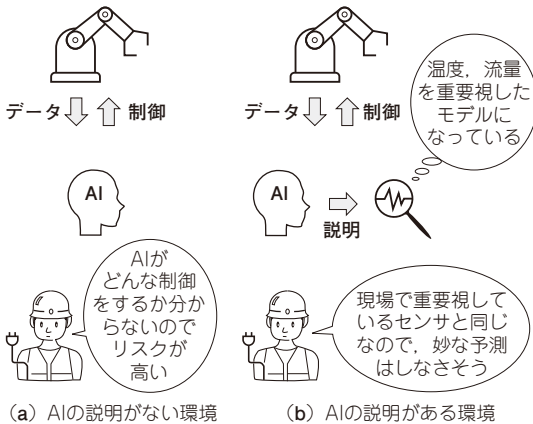
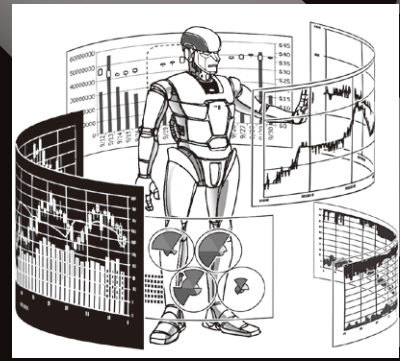


図1 AIの説明性有無による違い

ディープ・ラーニングや機械学習などのAIモデルは、複雑性が高くAIモデルの中で何が起きているのかは、多くの場合、ブラックボックスとして扱われます。そんな中で説明可能なAI (eXplainable AI: XAI) という概念が生まれました(図1)。説明可能とは、AIモデルの説明が得られる、すなわち、AIモデルが出した結果の根拠を示せるということです。結果の根拠とは、モデルへの入力特徴の重要度を数値化したものや、計算過程を可視化したものです。本連載では、AIモデルが出した結果の根拠を計算するアルゴリズムとその実装までを解説します。

● 今回のテーマ

本連載ではAIの説明手法について、Global(大局的)かLocal(局所的)かの軸、Model specific(モデル依存)か Model agnostic(モデル非依存)かの軸で大きく4種類に分けながら紹介します。前回までは、線形回帰モデルの重みなどモデルに依存した説明手法を説明してきました。

今回はモデルに依存しない説明手法であるモデル Partial Dependence Plot (PDPlot) を解説します。この手法は線形回帰、勾配ブースティング、ニューラル・ネットワークなど、どのような機械学習モデルの

表1 モデル依存、モデル非依存な手法の比較

	Model specific (モデル依存) な説明 (モデルの構造に 依存して説明を計算)	Model agnostic (モデル非依存) な説明 (モデルに依存しない ブラックボックスな 説明を計算)
主な 計算方法	モデルの種類によって 変化	モデルの入出力のみ (入力特徴とモデルが出 力する予測値) から計算
メリット	モデルの計算アルゴリズム を紐解くので、より正 確な説明を知ることが できる	どのモデルに対しても同 じ手法を適用できるの で、モデル間の比較が可 能
デメリット	モデルの種類によって計 算される値の意味が変 化するので、モデル間 の比較が難しい	モデル構造は無視する ので、真にモデルが判 断した理由を知ることは 難しい

説明もできることが特徴です。

こんなときにモデル非依存な説明手法 が必要になる

● 複数モデルでの結果を判断したいときに有効

実際のビジネスにおいて機械学習を用いた課題解決を実施する際は、複数種類のモデルを利用し、その精度や説明を比較することがあります。

例えば、天気や降水量、気温などの天気予報のデータから、商品が1週間後にどれくらい売れるかの需要予測をすることを考えます。線形回帰モデルと勾配ブースティング木を用いた場合、精度に対しては平均絶対誤差などの指標を使えば両者の比較はできます。AIの説明手法として、Model specific(モデル依存)な手法(表1)、すなわち線形回帰の重みの可視化と Feature Importance (AI説明手法の1つ、連載第3回で解説)を利用したとします。線形回帰では降水量の重みが最も高く0.5、Feature Importanceは0.2を示しました。そもそも2つの値は全く異なるアルゴリズムで計算されており、定義域も無限、0～1と異なるため比較が難しいという欠点があります。

このような複数モデルを比較する際には、その説明間の違いの考慮した上で比較しなければならず、解釈

第1回 AIモデルが重視している特徴を抽出…線形回帰モデルの場合(2025年3月号)

第2回 決定木回帰モデルの判断が妥当か調べる(2025年5月号)

第3回 どの特徴量が貢献しているかを表す Feature Importance (2025年7月号)