

GA2Mによる予測モデルの構造推定能力に関する研究

1X19C033-9 大森雅仁
指導教員 後藤正幸

1. 研究背景・目的

近年、高い予測精度をもつ様々な機械学習モデルが提案されており、実社会で扱われる表形式データに活用する試みが行われている。しかし、この応用活用を妨げる問題の1つとして、多くの高性能な機械学習モデルは中身がブラックボックスであるため、予測の根拠が明確ではない点が挙げられる。この問題に対する技術として、説明可能なAI(Explainable Artificial Intelligence, XAI)という研究分野がある。この研究分野の方向性として、モデル自体に解釈性を備えさせるアプローチとモデルの予測結果に解釈性を付与するアプローチの2つがある。前者の代表モデルはGeneralized Additive 2 Model(以下、GA2M)[1]であり、後者の代表手法はSHapley Additive exPlanations(以下、SHAP)[2]である。SHAPは、LightGBM[3]のようなブラックボックスモデルに対しても、解釈性を付与することができる。

実データへの応用場面では、両手法を各分野の実データに適用し、検証を繰り返すことで運用しているが、実データは対象問題によって異なる特徴を持つ。前述した2つの手法はモデル構造が大きく異なるため、解釈の方法の以前に一概にどちらか一方が優れているとは言えず、対象データの特徴ごとに有効な手法を選択する必要がある。そこで本研究では、GA2MとLightGBM+SHAPに対して、1)特徴ベクトルと目的変数の関数関係の複雑さ、2)特徴ベクトル内の有効特徴量の割合、3)予測モデルの構造の解釈性の3指標に着目し、手法選択の指針について検討を行う。具体的には、人工データを用いて様々な角度から比較実験を行い、その実験結果から得られる知見に基づき、手法選択の方法について提案を行う。

2. 準備

2.1. GA2M

GA2Mは、表形式データに適用可能なモデルで、一般化加法モデル(Generalized Additive Models)の発展でありながら、解釈性と予測精度を両立したモデルである。GA2Mでは、特徴量 x_i ごとに関数 $f_i(x_i)$ で非線形関係を表現し、さらに有効な2次の交互作用項を検出し、同様に関数 $f_{ij}(x_i, x_j)$ で学習を行うモデルである。GA2Mのモデルを式(1)に示す。 g はリンク関数、 $E[y]$ は期待値、 β_0 は切片である。

$$g(E[y]) = \beta_0 + \sum f_i(x_i) + \sum f_{ij}(x_i, x_j) \quad (1)$$

GA2Mでは単一項目の貢献度をグラフ化、交互作用項の貢献度をヒートマップ化できるため、解釈性が高い。また、交互作用項を考慮しているため、一般化線形モデルなどの手法に比べて、予測精度も高いという特徴を持つ。

2.2. LightGBM

LightGBMは決定木アルゴリズムと勾配ブースティングを組み合わせた手法であり、高い予測精度で知られている。

しかし、学習フェーズで生成した多数の決定木は、それぞれが異なる識別規則を持つため、モデルの構造が非常に複雑であり、人間が理解することは難しいブラックボックスモデルになってしまう。

2.3. SHAP

SHAPはモデルの予測結果に解釈性を付与する手法である。SHAPは、学習済みのモデルに対して、各特徴量が存在する場合と欠損している場合で予測値の変化を比較し、各特徴量の学習済みモデルに対する貢献度を可視化する手法である。これにより、モデル自身が解釈性を備えていないブラックボックスモデルによって導き出された予測結果が、各特徴量の影響をどの程度受けているかを理解することができる。しかし、必ずしもモデルの判断の根拠を完全に明示できるわけではなく、SHAP値のみからモデルの学習に問題がないかを検証することは難しい。また、SHAPでは特徴量ごとに貢献度を算出しているため、交互作用の影響は関連する特徴量の貢献度に分配され、交互作用のみの影響を可視化することができないという問題点も挙げられる。

3. 評価指標

本研究では、実データ応用において重要な3つの指標である、特徴ベクトルと目的変数の関数関係の複雑さ、特徴ベクトル内の有効特徴量の割合、予測モデルの構造の解釈性の観点から評価実験を行う。

3.1. 特徴ベクトルと目的変数の関数関係の複雑さ

実データでは、目的関数の真の構造に高次の交互作用が含まれる可能性が考えられる。そこで、真の構造に含まれる交互作用の最大次数を変化させて実験を行う。この指標に対して、両手法の違いを確認できれば、対象データの複雑さに応じて、適切なモデル選択を行う指針とすることができる。

3.2. 特徴ベクトル内の有効特徴量の割合

実データに対して予測を行う際には、目的関数に関与しない冗長な特徴量が含まれる可能性がある。ここで、冗長な特徴量がどれだけ含まれるかによって有効な手法が変わることが考えられる。そこで、学習データに対して含まれる有効特徴量の割合を変化させて実験を行う。この指標から、冗長なデータに対する頑健性を評価する。

3.3. 予測モデルの構造の解釈性

判断の根拠を求められる分野での実データ分析においては、分析者にとって学習済みモデルの解釈性が重要である。この指標では、モデル自体に解釈性を備えるGA2Mは、モデルの予測結果に解釈性を付与するLightGBM+SHAPよりも、分析者にとって理解しやすいモデルであるかを検証する。

4. 人工データを用いた比較実験

本研究では、前章で示した観点に沿って生成した人工データを用いて、比較実験を行う。

4.1. 人工データの生成条件

人工データの真の構造には、有効特徴量 $x_i \sim N(0, 1^2)$ と有効特徴量同士の組み合わせで構成される交互作用項に対して、それぞれの係数を a_j とした多項式回帰モデルを仮定する。また、関数関係の複雑さは、真の構造に含まれる交互作用項の最大次数を変化させることで決定する。さらに、学習データは、真の構造に含まれる有効特徴量と同じ分布から生成した冗長な特徴量を付け足すことで生成する。すなわち、有効特徴量を固定し、追加する冗長な特徴量の数を操作することで、学習データに含まれる有効特徴量の割合を変化させる。

4.2. 実験条件

モデルの評価指標には RMSE を用い、GA2M と LightGBM+SHAP の予測結果を比較する。試行回数は 10 回、有効特徴量数は 10、サンプル数は 5,000、学習データとテストデータの分割は 7:3 で行う。変化させる実験条件は、真の構造に含まれる交互作用の最大次数 (1~3) と学習データの有効特徴量割合 (0.0~1.0) である。

4.3. 実験結果

交互作用の最大次数と有効特徴量の割合を変化させた場合の実験結果を図 1 に示す。

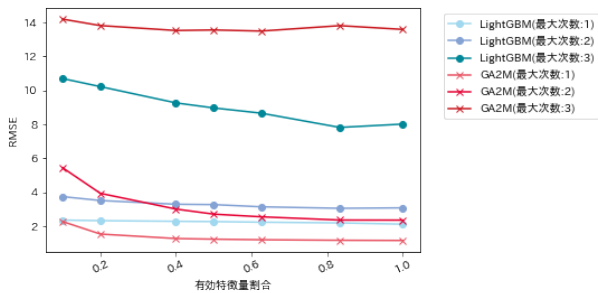


図 1: 交互作用の最大次数と有効特徴量の割合の比較

図 1 が示すように、最大次数が 1 のデータに対しては、GA2M の予測誤差が LightGBM よりも小さい。また、最大次数 2 のデータでは有効特徴量の割合が極めて低い場合のみ、LightGBM よりも予測誤差が大きいことがわかる。これは両手法のモデルの構造による違いだと考えられる。

GA2M と LightGBM+SHAP の解釈性に関する検証結果を図 2, 図 3 に示す。

図 2 より GA2M では、特徴量 x_5 の係数が 5.07、交互作用 $x_5 \times x_6$ の係数が 4.69 と推定されている。これらはそれぞれ、真の構造 $y = \dots + 5x_5 + \dots + 5x_5x_6 + \dots$ の特徴量 x_5 と x_5x_6 の係数と非常に近い値であり、精度よく推定されていることがわかる。また、単一項の貢献度を示したグラフと交互作用項の貢献度を示したヒートマップは、分析者にとって理解しやすい可視化方法になっている。

図 3 に示した LightGBM+SHAP の結果からは、特徴量 x_5 と SHAP 値の明確な関係性を、読み取ることはできない。また、交互作用に特徴量 x_6 を指定することで、特徴量 x_5 と x_6 に交互作用があることを確認できるが、真の構造を推測するには分析者による詳細な分析が必要となる。

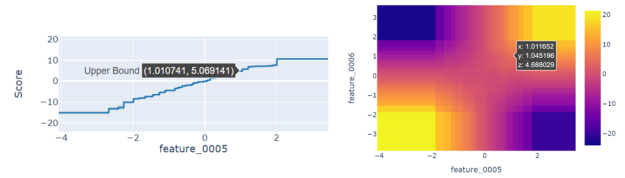


図 2: GA2M による特徴量 x_5 の貢献度を示したグラフと交互作用 $x_5 \times x_6$ の貢献度を示したヒートマップ

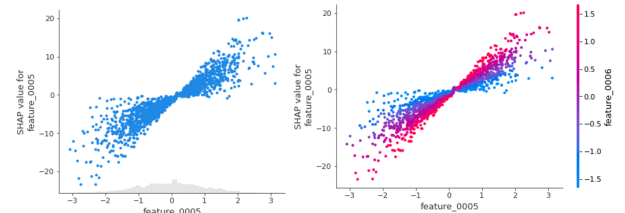


図 3: LightGBM+SHAP による特徴量 x_5 の貢献度と交互作用に x_6 を指定した場合の貢献度

5. 考察

4.3 節の実験結果より、一部の状況においては、GA2M の予測精度は LightGBM を上回っていることがわかる。また、いずれの状況においても GA2M は LightGBM + SHAP よりも高い解釈性を示しており、データの構造推定能力に優れていると考えられる。以上の特徴を表 1 にまとめる。

表 1: 各モデルの特徴

モデル	関数関係の複雑さ	有効特徴量の割合	モデルの解釈性
GA2M	△ (最大次数 2 まで)	△	○
LightGBM + SHAP	○	○	△

6. まとめと今後の課題

本研究では、様々な状況下における GA2M と LightGBM+SHAP の有効性と性質を示し、手法選択の指針を明らかにした。GA2M は、関数関係が複雑なデータや、有効特徴量の割合が低いデータに対して、予測精度が LightGBM に劣ることがわかった。しかしモデルの解釈性については、GA2M は LightGBM+SHAP よりも優れており、医療や人事など、モデルの予測の根拠が重要な分野において有効な手法であると考えられる。

今後の課題としては、本研究にて比較した観点について実データを用いて、検証することなどが挙げられる。

参考文献

- [1] Lou, Y., Caruana, R., Gehrke, J., Hooker, G., "Accurate intelligible models with pairwise interactions," In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 623-631, 2013.
- [2] Lundberg, S. M., Lee, S. I., "A unified approach to interpreting model predictions," *Advances in neural information processing systems* 30, 2017.
- [3] Ke, Guolin, et al., "Lightgbm: A highly efficient gradient boosting decision tree," *Advances in neural information processing systems* 30, 2017.