

弱学習器間の関係性に着目したアンサンブルモデルの可視化

柏山 美結* 廣川 暢一† 松野 竜太† 佐久間 啓太† 伊藤 貴之*

概要. 機械学習技術を用いた予測モデルは、さまざまな業界で広く採用されている。一方で、運用されている予測モデルは年々複雑化している。複雑なモデルを解釈するためには、予測に使用される訓練データの理解に加えて、モデル自体の構造やモデルに特定のデータを入れた時の振る舞いを把握することも重要である。そこで本研究では、実際に運用段階で頻繁に使用されているアンサンブル学習モデルの学習メカニズムを3次元空間で可視化しモデル構造の理解を支援することで精度劣化原因の特定を促すインタラクション手法を提案する。本報告では、アンサンブル決定木として勾配ブースティング決定木 *GBDT* を採用し、オープンデータセットを用いた実験を通して、提案する可視化手法の有用性を示す。

1 はじめに

現在では、産業など幅広い分野において機械学習技術を用いて開発された予測モデルが頻繁に運用されている。しかし、周辺環境の変化などにより予測精度が劣化することがあるため、運用中も監視することが必要不可欠である。近年、開発されたモデルが巨大・複雑になり、開発者でも全貌を把握するのが困難になりつつある。そこで、機械学習モデルの構造の分析や評価を支援する手法として情報可視化技術が注目されている。Lin hao[1] および Zijie[2] らは学習データの性質やモデルの性能等から複数のモデルが比較できるような可視化システムを提案した。また、Nagasaka ら [3] は、没入型環境における深層学習モデルの可視化システムの提案をしており、3次元空間での可視化の有効性を示している。

そこで本研究では運用段階において広く利用されているアンサンブル学習の勾配ブースティング決定木を対象とし、モデルの構造理解を支援するような3次元空間での可視化の一手法を提案する。また、予測モデルの精度劣化原因を探索するためのインタラクション手法についても議論する。

2 関連研究

機械学習 (ML) モデルの解釈可能性と予測精度を高めることを目的とした ML モデル可視化に関する研究は、既に多く発表されている。Kovalerchuk ら [4] は、機械学習のための新たな2つの決定木可視化手法を提案した。属性間の関係や決定木構造、決定木内のデータフロー等を分析することで、モデルの過度の一般化や過学習を防ぐことに有効である

ことが報告されている。

Chatzimparmpas ら [5] は、多様なアンサンブル学習ベースの ML モデルからの意思決定支援を目的として、可視化分析システムを提案した。各モデルの性能や重要な特徴量、訓練データとテストデータの比較結果を表示し、モデルについての意思決定を行う ML エンジニアとドメイン専門家の共同フローにより、モデルの分析、調整を可能にしている。

予測に使用されるデータの特徴量やモデルの性能結果だけでなく、弱学習器同士の関係性などのようなモデルの学習過程も同時に見ることで、より詳細にモデルの解釈ができると考えられる。また、アンサンブル学習モデルの可視化の多くは2次元の画面上のものが多く、新たに3次元空間での可視化手法を提案する。

3 提案手法

本論文では勾配ブースティング決定木 (GBDT) を対象とする。ここでは、1) 個々の弱学習器の構造・性能に関する情報、2) 弱学習器間の関係性に関する情報、および 3) 異なるデータセット間でのデータフローの差異に関する情報を可視化することを試みる。1) は弱学習器である決定木の葉の数、葉ごとのサンプル数と残差、および葉同士の特徴量類似度とし、また 2) は隣接する決定木の葉に含まれるサンプルの重複度、3) は異なるデータセット間での 2) の差分とした。

3.1 モデルの視覚表現

次に、Australian Credit Approval Dataset を用いて構築した GBDT モデル (LightGBM) を、提案手法により3次元空間上に可視化した結果を示す。描画には Unity2021.3.8 を用いた。シーン内の球体オブジェクトは、弱学習器の各々の葉を示しており、球体の直径および色は葉に含まれるサンプル数およ

Copyright is held by the author(s). This paper is non-refereed and non-archival. Hence it may later appear in any journals, conferences, symposia, etc.

* お茶の水女子大学

† 日本電気株式会社

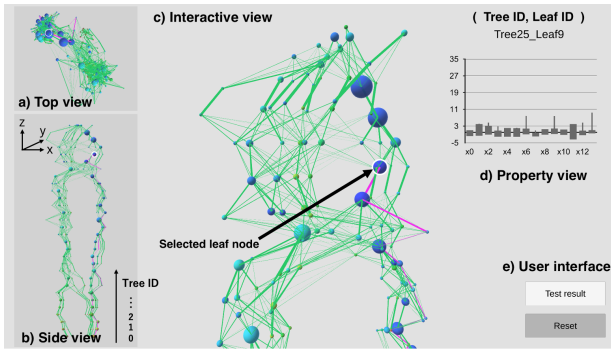


図 1. 可視化手法の概要

びそれらの平均予測誤差をそれぞれ示す. ここでは, 同一の x - y 平面に描画されている球体は一つの弱学習器に含まれる葉を表し, z 軸に沿って複数の弱学習器の葉の情報が描画されている.

球体オブジェクトを接続するリンクは弱学習器間の関係性を示しており, 隣接する層に属する二つの葉に対し, それぞれに含まれるサンプルの重複度に応じてリンクが太くなるように設定した.

また, x - y 平面上における球体の座標は, 1) 同じ弱学習器の葉同士の特徴量類似度, および 2) 隣接する弱学習器間でのサンプルの重複度に基づき決定した.

3.2 インタラクション手法

Unity を用いて実装された提案可視化手法のプロトタイプインタフェースを図 1 に示す. 着目する葉をクリックすることで, 葉内に含まれるサンプルの学習経路をハイライトする. リンクの色をピンク色に変化させることで学習経路を示す. また同時に, 葉内のデータの詳細をグラフ表示する.

Test Result ボタンにより, 図 2 から図 3 のように表示を切り替えることで学習データとテストデータの差分を視覚的に表示する. これらのデータセット間の誤差の大きさの変化は, ノードの色の变化で示される. 赤いノードは学習データに対する誤差の増加を示し, 青いノードは誤差の減少を示す. モデル学習中に存在しなかったデータの軌跡を表すパスは, 赤い破線で描かれる. この機能はモデルの学習結果の全体的な概観を提供し, 異常な葉ノードの識別を容易にする. さらに, モデルの改良と予測誤差の原因分析を支援するために, 個々の木と関連するデータサンプルの詳細な洞察を可能にする.

4 今後の課題

本手法を他のデータセットにも適用させることで, よりモデルの解釈性を向上させるためのインタラクションについて議論を進める. また, 予測モデルの精度監視技術など, MLOps と呼ばれる運用中の課題を解決するための技術開発が進められている

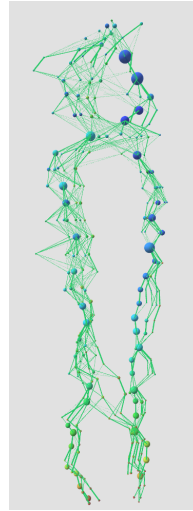


図 2. 学習データによる可視化

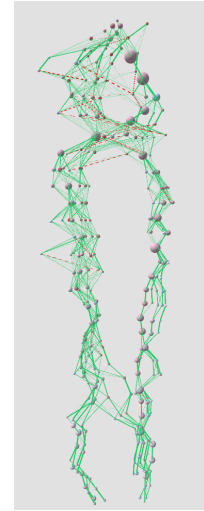


図 3. データセット間の差分表示

[6]. 本研究においても, 運用中の精度劣化原因を探索するような MLOps における技術開発にも適用させたい.

参考文献

- [1] L. Meng, S. van den Elzen, and A. Vilanova, "Modelwise: Interactive model comparison for model diagnosis, improvement and selection," *Computer Graphics Forum*, vol. 41, no. 3, pp. 97–108, 2022.
- [2] Z. J. Wang, C. Zhong, R. Xin, T. Takagi, Z. Chen, D. H. Chau, C. Rudin, and M. Seltzer, "Timbertrek: Exploring and curating sparse decision trees with interactive visualization," in *2022 IEEE Visualization and Visual Analytics (VIS)*, pp. 60–64, 2022.
- [3] H. Nagasaka and M. Izuhara, "Interactive visualization of deep learning models in an immersive environment," 2021.
- [4] B. K. A. Dunn, A. Worland, and S. Wagle, "Interactive decision tree creation and enhancement with complete visualization for explainable modeling," 2023.
- [5] A. Chatzimparmpas, R. M. Martins, and A. Kerren, "Visruler: Visual analytics for extracting decision rules from bagged and boosted decision trees," *Information Visualization*, vol. 22, no. 2, pp. 115–139, 2023.
- [6] D. Sculley, G. Holt, D. Golovin, E. Davydov, T. Phillips, D. Ebner, V. Chaudhary, M. Young, J.-F. Crespo, and D. Dennison, "Hidden technical debt in machine learning systems," in *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2, NIPS'15*, (Cambridge, MA, USA), p. 2503–2511, MIT Press, 2015.