

勾配ブースティング決定木を用いた損傷原因および補修工法推定の可能性検証

大日本コンサルタント正会員 ○龍田 斉
テクノハイウェイ 正会員 永見 武司
金沢大学 正会員 近田 康夫

大日本コンサルタント	フェロー	横山 広
金沢大学	フェロー	梶谷 浩
富山県金沢河川国道事務所	非会員	山下 忠男

1. 背景と目的

平成 26 年度の道路法改正により，国内の橋梁は 5 年ごとに定期点検が義務付けられ，その点検結果に基づいた損傷対応として補修・補強が実施されている．しかし，地方自治体の多くは橋梁維持管理に必要な専門知識を有する技術者が確保できていない．加えて，少子高齢化により橋梁維持管理に従事する労働力が減少傾向にあることから，ICT 技術の活用による生産性向上が求められている．

一方、データ分析のコンペティションでは機械学習技術の 1 種である勾配ブースティング決定木 (Gradient Boosting Decision Tree, 以下 GBDT) が精度・計算速度ともに優れており、よく利用されている。本研究では、地方自治体に所属する道路管理者の補修工法選定の意思決定補助を目的として、橋梁管理システムによって記録された橋梁管理カルテ情報から損傷原因および補修工法の推定に GBDT が活用できるか検証した。

2. 概要

検証は①学習用データ作成と、②GBDT モデル選定③チューニング④学習結果の分析の4ステップに大別される。以下に各ステップの内容を示す。

3. 学習用データ作成

GBDT 用の学習データは、国土交通省金沢河川国道事務所が管理する 453 橋の 10 年分の橋梁管理カルテデータを用いた。橋梁管理カルテデータには橋梁の諸元の他に、損傷原因、補修工法が記載されている¹⁾。表-1 に橋梁管理カルテデータの主なデータ項目を示す。これらのデータから、GBDT で推定する目的変数を対策区分判定、原因（推定）、原因（確定）、補修補強工法として部材単位で抽出した。GBDT の学習に用いる説明変数として諸元を部材単位で抽出した。なお、表-1 の各目的変数に併記する数字は抽出件数である。

4. GBDT モデル選定

データ分析コンペティションで使用される GBDT のモデルは複数存在する。今回の検証では代表的な GBDT モデルである XGBoost²⁾、LightGBM³⁾、Catboost⁴⁾の精度を確認した。モデルの目的変数は損傷原因（推定）とした。モデルの学習に設定するパラメータは Catboost 以外、各モデルのデフォルト値を利用した。Catboost は GPU に対応しているため、GPU を利用するオプションを有効にし、GPU を 4 ユニット使用した。検証に用いるデータは 5 つのグループに分割し、4 つのグループを学習に、残る 1 つのグループを精度の検証に用いた。これを 5 回繰り返し、精度と計算時間の平均を算出した結果を表-2 に示す。計算は産業技術総合研究所が構築・運用する AI 研究開発用のクラウドサービスである「AI 橋渡しクラウド」の計算資源タイプ rt_F を使用した。表から、全モデルとも精度が 95%以上であることが確認できる。今回は検証した 3 モデルの中で精度および計算速度が最も早い LightGBM に絞って後続の検討を実施する。

表-1 橋梁管理カルテのデータ項目

説明変数(諸元)	目的変数(診断結果)
橋種、交通量、橋長、橋面積、構造形式、防水工有無、等級、設計活荷重、設計震度、平面形狀、平面線形、縦断勾配、迂回路有無、交差物名称、塩害地域区分、幅員構成、添架物、塗装仕様、点検施設、排水施設など	損傷原因(確定): 1795件 損傷原因(推定): 15164件 対策区分判定: 28032件 補修補強工法: 873件

表-2 各モデルの精度と計算時間

モデル	検証結果	
	精度[%]	時間[秒]
XGBoost	95.748	86.6
LightGBM	96.190	28.6
Catboost	96.069	40.0

キーワード 機械学習, GBDT, 維持管理, 橋梁点検, 橋梁管理カルテ

連絡先 〒330-6011 さいたま市中央区新都心 11-2 TEL048-615-2225

5. チューニング

LightGBM のハイパーパラメータ最適化については、ハイパーパラメータ自動最適化フレームワークの Optuna の拡張機能である LightGBM Tuner を利用した。チューニング前後の精度変化を表-3 に示す。表からチューニングによる精度向上は 1%未満であることが確認できる。補修補強工法の高精度は、レコード数が少なく、今回の対象データに特化しており、他の地域や事務所のデータでは同様の精度が得られない可能性がある。

6. 学習結果の分析

学習後のモデルから出力した説明変数の重要度上位 20 位までを表-4 に示す。重要度は目的変数を分類する際の当該説明変数の寄与度で、当該説明変数の分岐によって目的関数の損失を減少させた度合を表している。表から、推定原因（推定）については、床版厚さや交通量、施設完成年度など疲労や材料劣化との関連が高い変数が確認できる。一方で縦断勾配や主桁本数などの形状に関連する変数も上位に来ており、特定の形状に起因する損傷が存在する可能性がある。対策区分判定および推定原因（確定）についても同種の傾向にある。補修補強工法については、具体的な部材形式や路線名など橋梁を特定する情報が目立つ。

7. まとめ

本研究では、GBDT を活用した橋梁点検カルテデータから損傷原因や補修工法等の推定について可能性を検証した。得られた知見を以下に示す。

- (1)GBDT はいずれのモデルも橋梁管理カルテデータから高い精度で損傷原因や対策区分を推定可能であることを確認した。
- (2)重要度の大きい説明変数は、劣化要因と直接関連するものと形状に関連するものとに区分された。

謝辞

本稿は北陸・道路メンテナンス会議に設置された「道路橋の維持管理における各構成部材の限界状態ならびに AI 技術の活用に関する検討ワーキンググループ」の研究成果をまとめたものである。ここに関係された皆様に謝意を表します。

参考文献

1) 国土交通省道路局：橋梁の維持管理の体系と橋梁管理カルテ作成要領（案）,2004.

2)Tianqi C. and Carlos G.：XGBoost: A Scalable Tree Boosting System, In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), pp. 785-794, 2016.

3)Guolin K. et al.：LightGBM: a highly efficient gradient boosting decision tree. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17), pp. 3149-3157, 2017.

4)Prokhorenkova L. et al.：CatBoost: unbiased boosting with categorical features, Advances in neural information processing systems, pp. 6638-6648, 2018.

表-3 チューニング後の精度変化

目的変数	チューニング前後の精度[%]	
	実施前	実施後
損傷原因(確定)	94.467	94.526
損傷原因(推定)	96.205	96.214
対策区分判定	91.398	91.609
補修補強工法	100.000	-

表-4 重要度の大きい説明変数

順位	損傷原因(推定)	補修補強工法
1	材料(コンクリート)	材料(コンクリート)
2	支間長	材料(鋼)
3	大型車混入率	施設完成年度
4	施設完成年度	橋面積
5	縦断勾配	縦断勾配
6	部材種別(堅壁)	部材種別(堅壁)
7	床版厚さ	起点側支承種類(支承板支承)
8	橋長	床版厚さ
9	部材種別(主桁)	設計震度
10	昼夜率	下部工形式(半重力式橋台2基)
11	主桁本数	現旧区分(現道)
12	橋面積	下部工形式(重力式橋台2基,壁式橋脚 (RC) 2基)
13	交通量	通行制限(自動車通行不能)
14	部材種別(横桁)	路線(159号)
15	有効幅員	交通量(昼夜24時間)
16	車道有効面積	大型車混入率
17	設計震度	主桁本数
18	交通量(昼夜24時間)	下部工形式(逆T式橋台,T型橋脚 (RC) 2基)
19	材料区分(PC)	上部工形式(連結PCプレテンT桁橋)
20	塩害地域区分(BS)	中塗り塗料(長油性フタル酸樹脂中塗り塗料)