

## 論文・報告

# 機械学習を用いたシステム建築における鉄骨重量の推定

## ～XGBoost 予測モデルとアプリケーションのノーコード開発事例～

### No-Code Development of a ML-based System for Estimating Steel Quantity in Steel Buildings

室田 千春 \*1  
MUROTA Chiharu

山崎 美樹 \*2  
YAMAZAKI Miki

田中 政義 \*3  
TANAKA Masayoshi

金平 徳之 \*4  
KANEHIRA Noriyuki

川田工業(株)建築事業部では、経験豊富な技術者の暗黙知を経験の少ない技術者でも利用可能にすることで構造設計や見積り業務を効率化することが望まれていた。本稿では、過去の設計データで訓練した機械学習モデルを使い、未知の新しい物件の設計条件を入力すると、鉄骨重量合計を予測して出力するシステムを作成した。また、これらの開発にノーコードツールを用いることで内製化し、自社の知識を活かして機械学習モデルの構築・アプリ化までを短期間で実現できた。

本報告では、前半に機械学習の定義や分類、注意すべき点など、一般的な基礎知識を紹介する。後半では実例を通して、ノーコードツールを用いた機械学習モデルの構築とアプリ化、および、その中で行ったハイパーパラメータの最適化について報告する。最後に、今後の課題とともに、正確なデータを蓄積することの重要性について述べる。

キーワード：機械学習、システム建築、XGBoost、データ活用、ノーコード

## 1. はじめに

川田工業(株)建築事業部（以下、建築事業部）では、経験豊富な技術者（以下、ベテラン技術者）の暗黙知を経験の少ない技術者（以下、若手技術者）でも利用可能にすることで構造設計や見積り業務を効率化することが望まれていた。本稿では、過去の設計データで訓練した機械学習モデルを使い、未知の新しい物件の設計条件を入力すると、鉄骨重量合計を予測して出力するシステムを作成した。また、これらの開発にノーコードツールを用いることで内製化し、自社の知識を活かして機械学習モデルの構築・アプリ化までを短期間で実現できた。

本報告では、前半に機械学習の定義や分類、注意すべき点など、一般的な基礎知識を紹介する。後半では実例を通して、ノーコードツールを用いた機械学習モデルの構築とアプリ化、および、その中で行ったハイパーパラメータの最適化について報告する。最後に、今後の課題とともに、正確なデータを蓄積することの重要性について述べる。

## 2. 機械学習の基礎知識

### (1) 機械学習の定義と分類

機械学習は人工知能の一分野であり、明示的なプログラムがなくてもデータから学習して事象をモデル化し、その機械学習モデル(以下、モデル)を使うことで未知の条件に対して予測を行うことを可能にする。一般的には、教師あり学習、教師なし学習、強化学習の3つのカテゴリに分けられる。教師あり学習は、ラベル付きデータに基づいて学習を行うタイプである。一方、教師なし学習では、ラベルのないデータからパターンや関係性を学習する。そして、強化学習では、報酬とペナルティに基づいて意思決定を学習する。

### (2) 機械学習の主要な手法

機械学習には多くの手法が存在し、それぞれが長所と適用範囲を持つ。一般的に用いられる手法には、回帰、分類、クラスタリング、次元削減などを目的とするものがある。また、多層のニューラルネットワークを用いたディープラーニングは、その強力な予測能力から機械学

\*1 川田テクノロジーズ(株)技術研究所 係長

\*2 川田工業(株)建築事業部設計部設計二課

\*3 川田工業(株)建築事業部設計部設計一課

\*4 川田テクノロジーズ(株)技術研究所 所長

習の一部として人気を博している。

### (3) 機械学習の評価指標

モデルの有効性は、様々な評価指標を用いて判断される。分類問題では、精度、適合率、再現率、F1 スコア、受信者操作特性曲線下面積 (AUC-ROC) などが一般的に用いられる。どの指標を選ぶかは、解決しようとする具体的な問題と目標に依存する。

回帰問題の場合、平均絶対誤差 (MAE)、平均二乗誤差 (MSE)、決定係数 ( $R^2$ )、2 乗平均平方根誤差 (RMSE) などの指標が使われる。このうち RMSE は、予測誤差(または残差)の標準偏差である。予測誤差は、実際の値が予測値からどの程度離れているかを示す。わかりやすい例として、単回帰の例を図 1 に示す。

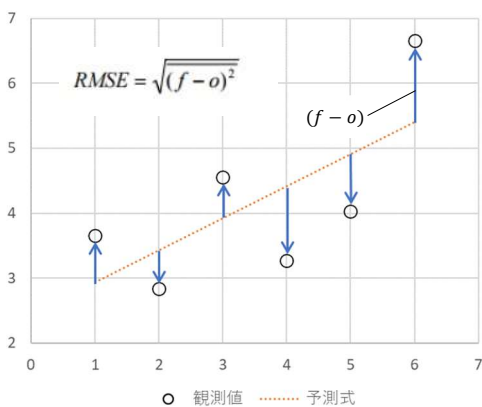


図 1 単回帰の例

### (4) 過学習について

過学習とは、モデルが訓練データのノイズやランダムな変動まで学習してしまうことにより、新しいデータに適応する能力が失われる現象である。この結果、訓練データに対する誤差は小さいが、テストデータや実際の運用時のデータに対する誤差は大きくなる。例として前節の観測値データを使い、次数数の想定の違いによる予測モデルの違いを示す。図 1 では  $y = a + bx$  で表される 1 次式の回帰直線を予測モデルとして描いたが、次数を増やして 3 次式のモデルを想定すると図 2 のようになる。

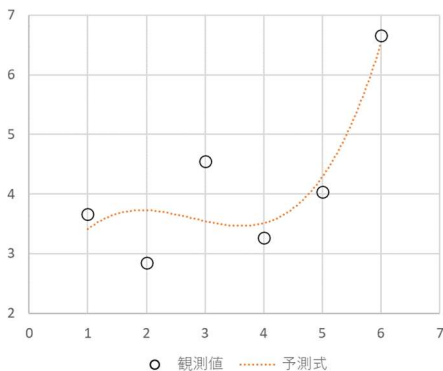


図 2 3 次式による予測モデル

実は、この観測値は 1 次式

$$y = 2 + 0.5x$$

から得られる値  $y$  に対し、分散 2 のばらつきをもつよう逆算して得られた値である。したがって、想定するモデルとしては本来 1 次式とすべきだが、少ないデータしか得られていないこの範囲では、どちらが良いモデルであるか断定することは難しく、一見すると 3 次式の近似の成績が良いようにも見える。しかし図 3 のように、新たなデータを加えると当てはまりが悪く、3 次式のモデルは適切ではなかったことがわかる。

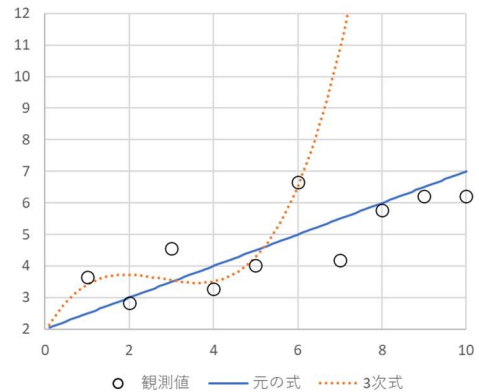


図 3 次数を増やしたことによる過学習の例

この例からわかる通り、過学習を防ぐために訓練データを増やすことは有効である。そのほか、無関係な特徴量や冗長な特徴量を取り除くことでモデルが不要な情報に適合するのを防ぎ、重みづけの項が極端な値を取らないようペナルティを課す「正則化」や、データを訓練データとテストデータに分割して繰り返しモデルを評価する「交差検証」などの手法が存在する。実際にはこれらの手法を複数採用することで、汎化性能と柔軟性のバランスが良く、データの特徴を捉えたモデルを実現する。

### (5) 機械学習モデルによる予測の注意点

前述のように、機械学習による予測とは一般に、既存のデータ内の関係性に、ある種の法則があると仮定し、未知の条件に対してその法則 (モデル) に当てはめた場合に得られる解を予測値として出力することを指す。言い換えれば、過去に似たような条件のデータがない場合や、誤入力などのノイズを多く含む場合、また想定するモデルが実際の現象をうまく説明できていない場合には、実際に近い予測を出すことはできない。また、良いモデルであっても、モデルが出力する値は、実際の値と全く同じにはならない。したがって、機械学習モデルを用いた予測結果を用いて何らかの意思決定を行う場合には、許容されるリスクを勘案し、目的に適した方法を選択して利用することが求められる。

### 3. データ収集と前処理

#### (1) データ前処理による品質改善の重要性と手法

データの質と量はモデルの性能に大きな影響を与える。データ収集の方法は多種多様であり、調査、実験、観察、ビッグデータからの抽出などがあるが、データが収集された後も、そのままモデルの学習に利活用するケースはほとんど無い。収集されたデータには、外れ値や誤入力などのノイズや欠損値が存在したり、スケールが統一されていないなどがある。これらは機械学習に限らず、データ分析をするうえで常に障害となることから、解消する必要がある。また、適用する手法や目的によって、データの変換や統合が必要になることもある。これらの作業をデータ前処理と呼ぶ。

データ前処理には、データクレンジング（誤ったデータの修正や削除、欠損値の補完、外れ値の除去）、データ統合（異なるソースからのデータの統一）、データ変換（データの形式や構造の変更、特徴量のスケールリング、カテゴリ変数の数値化）などが含まれる。これらによりデータの品質が向上することは、モデルの学習効率と性能の向上に大きな影響を与える。

#### (2) 特徴量エンジニアリング

モデルがより対象の特徴を捉えるための訓練データを最適化するために、データの持つ情報を加工することを特徴量エンジニアリングと呼ぶ。例えば長方形の平面形状をもつ建物の X 方向と Y 方向の大きさの情報があった場合、これらを用いて面積を計算することが、X と Y の大きさをそれぞれ単体で特徴量とするよりも、目的関数に対してデータの特徴をより正確に表現することになり、次元削減により良い影響を与える場合がある。

### 4. ノーコードツールの活用

#### (1) ノーコードツールの概要と特徴

ノーコードツールとは、従来のプログラミングの知識やスキルを必要とせず、ビジュアルなインタフェースやドラッグアンドドロップの操作を通じてプログラミングを行う手法やプラットフォームを指す。コーディングや環境構築の高い知識を必要としないため、ソフトウェアの開発を比較的容易に行うことができる。目的によっては、コード記述するよりも迅速にプロトタイピングを行うことができ、データ分析だけでなく自動化などにも利用できる。

後述する事例では、「KNIME Analytics Platform (以下 KNIME)<sup>1)</sup>」を使用し、データ抽出・前処理から機械学習モデル構築、アプリケーション化まで行った。KNIME は、ノーコード環境を提供する一方で、データ解析、機械学習、データ可視化といった高度な機能を持つ。

#### (2) ノーコードツールの利用による開発の容易化と現場技術者の貢献

KNIME のようなノーコードツールを使用することの利点として、プログラミング等の情報処理技術を専門としない現業の技術者が自身の業務知識を活用して、データ分析や機械学習モデルの開発を行えることが挙げられる。取り扱うデータの中には、たとえば建設業に関わるデータであればその分野の専門家でなければその意味を正確に把握できない場合がある。現場の技術者が直接データを分析し、その結果を用いるツールを自身のニーズに合わせて開発することが可能になるため、ソフトウェア技術者に外注する場合よりも、業務のニーズにより密接に、より早く対応することができる可能性がある。

#### (3) ノーコードツールの注意点と課題

一方で、ノーコードであるか否かに関わらず、データ分析や機械学習を扱う場合には、データの取り扱いやアルゴリズムに関する理解が無ければ、適切な結果を得ることはできない。

また、現場の技術者は、システムの内部動作や可能性を完全に理解するのが難しい場合があり、問題発生時のトラブルシューティングやパフォーマンスの最適化に影響を及ぼす可能性がある。更に、ノーコードツールの機能は、一般的にはツール提供元がリリースする機能に依存する。オープンソースである KNIME の場合には拡張機能を開発することが可能で、多くの場合信頼できる開発元がリリースしている拡張機能(以下、エクステンション)を利用することが可能だが、公開されているエクステンションの中に所望の機能を見つけられなかったり、新たにエクステンションやノードを開発したい場合にはより専門的な開発の知識を必要とする。

### 5. 適用事例

本章では、ノーコードツール KNIME を使い、建築事業部が保有する過去の物件データに機械学習を適用した事例を紹介する。この事例ではデータの前処理から機械学習モデルの構築と最適化、試用のためのアプリケーションまでをノーコードで実装した。

#### (1) 開発の経緯

建築事業部において、S 造システム建築「KBS」製品の設計施工は主力事業の一つである。システム建築とは、建物を構成する部材を標準化することにより、提供プロセスをシステム化した建築であり、工場・倉庫・物流施設等に広く用いられている建築工法である。この「KBS」シリーズで建築事業部が特に強みとしているのはコストを抑えた「自由設計」であり、それを可能にしているの

は、豊富な経験に裏打ちされた設計部門の構造設計技術の高さである。しかし、これには 2 つの課題があった。

- ① 最適なプランを提案するため、長い時間を要している見積作成期間を短縮すること
- ② ベテラン技術者と若手技術者の経験の差を減らすこと

これらの課題に対して、ベテラン技術者が直感で判断していることを、アプリを使って支援ができれば、時間の短縮とともに経験知の形式知化のひとつの形を示せると考えた。

具体的には、システム建築の構造設計において、物件の概要情報から鉄骨総重量の概算を推定できれば、工数をかけずに短期間で見積を作ることができ、早期受注や無駄な工数の削減に繋がると考えられる。また、ベテラン技術者が自らの経験からそのような推定を無意識に行うことで設計上の大きなミスを未然に防いでいるとすれば、これによって若手技術者も同じような指標を持つことができるようになる。

## (2) 全体方針

方針として、実績物件の規模やスパンなどの設計上の仕様に関する過去のデータを使って予測モデルを構築し、新しく与えられる物件の設計条件から、建物全体の鉄骨重量合計を出力することを目指す。

目的変数を鉄骨重量合計としたのは、次の理由による。

- ① 鉄骨重量は、システム建築プロジェクトの、経済的・物理的規模を把握するのに重要な指標である
- ② 経験の浅い技術者にとって、その概算を推定することが困難である

また、プログラミングの専門家でない建築系技術者による開発を可能とするために、KNIME を用いることとした。

## (3) 開発手順

### (a) 前処理

まずは既存システムのデータから、訓練およびテストに用いるデータセットを抽出するプログラムを KNIME で作成して、データセットを得た。

前処理として、データセット内の欠損値を確認し、欠損値が 5 割を超える項目を除外した。また、鉄骨重量合計や延床面積などの基本的な項目が欠損しているデータと鉄骨重量が極端に少ないなどの「外れ値」と判断できるデータを除外した。前処理後のデータセットの概要は表 1 の通りである。

表 1 データセット概要

レコード数	539 件
特徴量 (データ項目数)	33
※固有有名詞や社内 ID を除く	

### (b) 機械学習を含むプログラムの動作確認

(a) で準備したデータセットを用いて、学習→予測→精度評価を実施するプログラムを KNIME で作成し、基本的な動作の確認を行うとともに、次項以降に示す特徴量選択およびアルゴリズム選択に向けたベンチマークを得た。この時点では、アルゴリズムは Random Forest を用いた。

### (c) 特徴量選択

機械学習において、目的関数に全く関係のない変数を特徴量に含めると、それがノイズとなって精度が低下することが知られている。Random Forest の結果出力に含まれる特徴量重要度、および、鉄骨重量合計に対する相関を確認し、さらに、設計技術者の建築構造メカニズムに関する専門知識の観点から、関係のない特徴量を除外した。

### (d) 機械学習モデル構築とアルゴリズム選択

準備したデータセットは次のような特徴を持つ。

- ① カテゴリデータ(材料の種類など)と数値データ(スパン、床面積など)が混在している
- ② 非線形なデータパターンを持つ可能性がある
- ③ 欠損値を含む

このことから、KNIME で扱うことのできる手法のうち、決定木を使う 3 つのアルゴリズムである Random Forest, 勾配ブースティング (Gradient Boosting Machine; GBM), XGBoost<sup>2)</sup>を用いてモデルを構築し、交差検証の結果を評価し、比較した。

モデルの評価には R<sup>2</sup>, 予測誤差率の平均と標準偏差を用いた。予測誤差率とは、予測誤差 (予測値 - 正解値) を正解値で除した値である。目的変数のスケール (鉄骨重量合計の大きさ) が異なるデータの予測の集まりを評価するため、予測誤差そのものの大きさではなく、正解値に対する比率を評価することとした。予測誤差率の平均とは、予測誤差率の絶対値の平均である。

各アルゴリズムで構築したモデルの比較を表 2 および図 3 に示す。予測誤差率の分布を図 3 のバイオリンプロットで比較すると、XGBoost では 0 に近い値にデータが集中している一方、RandomForest では広く分布していることがわかる。比較の結果、より予測誤差比率の平均および標準偏差が低く、ばらつきが少ない XGBoost を採用することとした。

表 2 各アルゴリズムで構築したモデルの評価

アルゴリズム	RandomForest	GBM	XGBoost
R <sup>2</sup>	0.863	0.9	0.905
予測誤差率の平均	0.561	0.236	0.223
予測誤差率の標準偏差	1.425	0.299	0.263

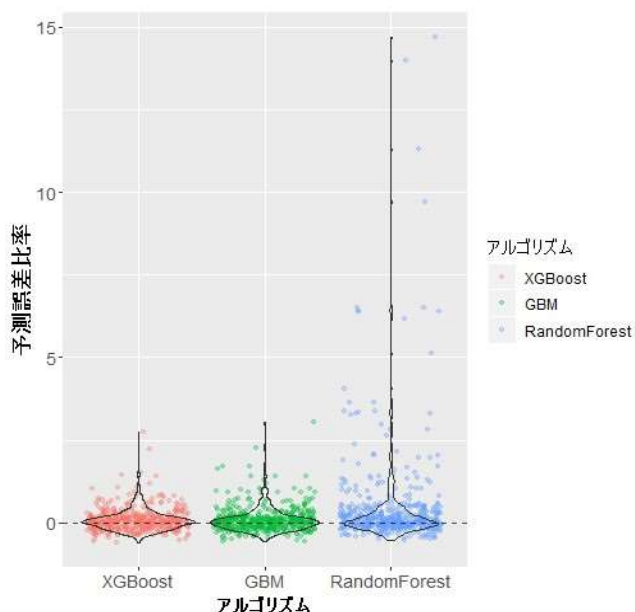


図3 予測誤差比率の分布の比較

採用した XGBoost は、「Extreme Gradient Boosting」を意味する機械学習アルゴリズムであり、勾配ブースティングフレームワークの一つである。勾配ブースティングを基盤としているが、XGBoost は L1 (Lasso 回帰) と L2 (Ridge 回帰) の正則化項を採用しており、モデルの過学習を防止する能力が強化されている。さらに、計算効率の面で XGBoost は内部的に並列処理をサポートしており、大規模なデータセットにおける学習も、GBM と比較して高速に実行できる。

#### (e) ハイパーパラメータの最適化

ハイパーパラメータとは、機械学習モデルの学習前に人手で決めるパラメータのことであり、設定できるハイパーパラメータの数や種類はアルゴリズムによって異なる。ハイパーパラメータを調整することで、精度の向上や過学習の抑制など、モデルの性能向上につなげることができる。ここでは、XGBoost のハイパーパラメータのうち特に重要なものを、ベイズ最適化によって探索した。詳細については次節で述べる。

#### (f) モデルの評価

作成したモデルの評価を行った。詳細は次節で述べる。

#### (g) 利用可能なアプリケーションの開発

新たな設計条件に対してこのモデルを適用した結果を、誰でも簡単に表示できるよう、アプリケーションを開発した。サーバー機能を有する「KNIME Server」を利用することで、設計担当者が実際に使うアプリケーション開発についてもノーコードで完結した。

### (4) 開発結果

本事例では、特徴量に関する建築系の知識をもつ担当

者が開発に直接携わることで、特に特徴量の選択や特徴量エンジニアリングについて効率的に行うことができた。本節では、XGBoost に設定するハイパーパラメータの最適化の結果と、それによって得られたモデルの評価および、アプリケーション化について述べる。

#### (a) ハイパーパラメータの最適化結果

最適な組み合わせを探索する一般的な方法には、グリッドサーチ、ランダムサーチ、ベイズ最適化等がある。グリッドサーチとは、あらかじめ決めた値の組み合わせを総当たりで試行した結果のうち最良のものを選択する方法である。ランダムサーチは、探索する値をランダムに決める点がグリッドサーチと異なる。ベイズ最適化は、試行した結果を考慮して、次に探索する組み合わせを決め、精度が十分に良くなるまでそれを繰り返す方法である。グリッドサーチは一見最も信頼性が高いように思えるが、組み合わせる値の間隔を十分に小さくし、探索点が十分多くないとピークを探索できない一方、探索するパラメータの数が多いと組み合わせ爆発により現実的な時間での計算ができない。ランダムサーチはグリッドサーチよりも筋の良い探索ができる場合もあるが、運任せの側面がある。ベイズ最適化は、試した結果から、より良い結果が期待できる次の組み合わせを予測して試すというサイクルを繰り返して最適な組み合わせを目指す方法で、少ない計算量で最適解に近い答えを得られる利点がある。今回は、ベイズ最適化を採用することとした。

XGBoost には、ほかのアルゴリズムに比較して多くのハイパーパラメータがあるため、影響の大きいパラメータに絞って最適化を行うのが一般的である。今回は表 3 に示すパラメータについて最適化を行うこととした。

ベイズ最適化の注意点として、探索の範囲を適切に設定しないと局所解に留まり、大域的最適解に到達できない可能性がある。パラメータの探索範囲を絞り込むため、それぞれのパラメータを変化させた場合の精度の変化をプロットして確認した(図 4)。精度の指標には予測誤差率の平均を用いた。交差検証における予測精度(validation score)が訓練データを使った予測精度(training score)にできるだけ近くなることが目標となる。training score がほかの範囲よりも悪い(予測誤差率の平均が大きい)範囲は、学習が十分でない「未学習」となっており、それ以外の範囲で validation score が悪化している範囲は「過学習」となっている。これらの範囲は探索範囲から除外することとし、図 4 の赤矢印で示した範囲を探索範囲とした。

これらのパラメータから、最適な組み合わせを探索するため、KNIME のエクステンションに含まれるノードを使い、Parameter Optimization Loop を構成して実行した結果、表 4 に示すハイパーパラメータを得た。

表 3 最適化するハイパーパラメータ

名称	値の取りうる範囲
max_depth	正の整数 (小さいほど保守的)
min_child_weight	0~∞ (大きいほど保守的)
colsample_bytree	0 ~ 1 (小さいほど保守的)
eta	0~1
subsample	0 ~ 1 (小さいほど保守的)
lambda	0~∞ (大きいほど保守的)

表 4 最適化の結果

名称	デフォルト値	最適化後
max_depth	6	7
min_child_weight	1.00	3.93
colsample_bytree	1.00	0.99
eta	0.30	0.11
subsample	1.00	0.69
lambda	1.00	24.75

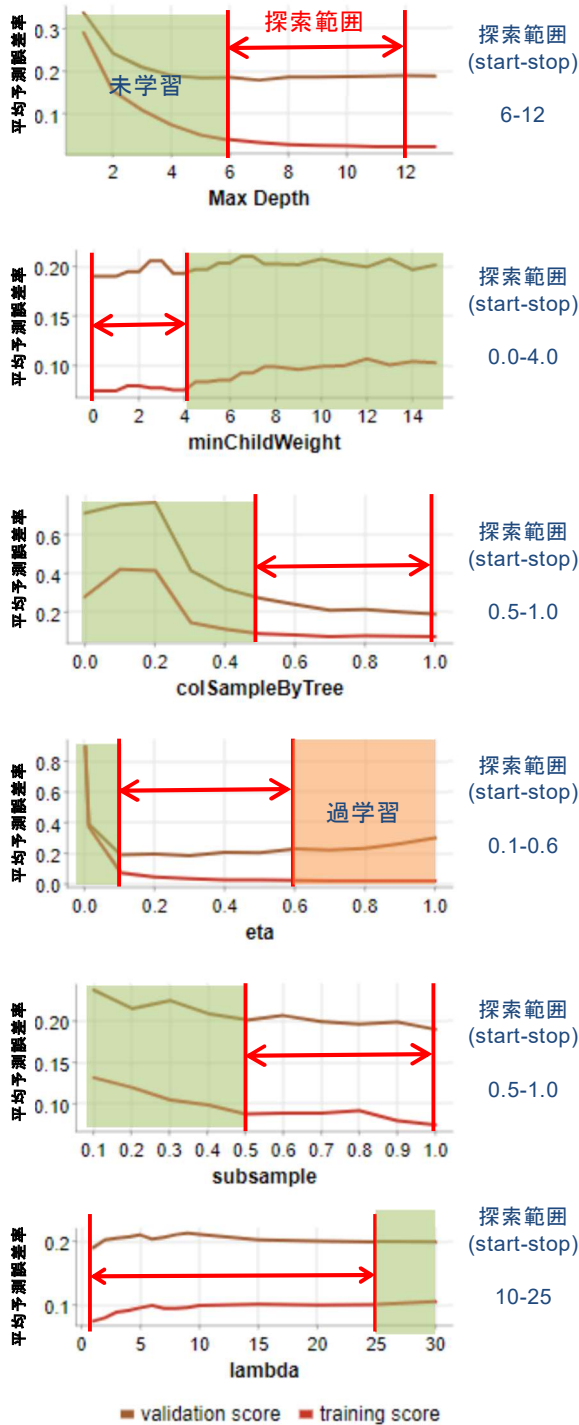


図 4 ハイパーパラメータと探索範囲

(b) モデルの評価

前項で最適化を行ったモデルによる比較を表 5 と図 5 に示す。予測誤差比率の平均は最適化前 0.223 であったのに対し、最適化後は 0.206 に改善し、モデルの当てはまりの良さを示す  $R^2$  および予測誤差比率の標準偏差も改善していることがわかる。

表 5 最適化後のモデルの評価

	最適化前	最適化後
$R^2$	0.905	0.929
予測誤差比率の平均	0.223	0.206
予測誤差比率の標準偏差	0.263	0.202

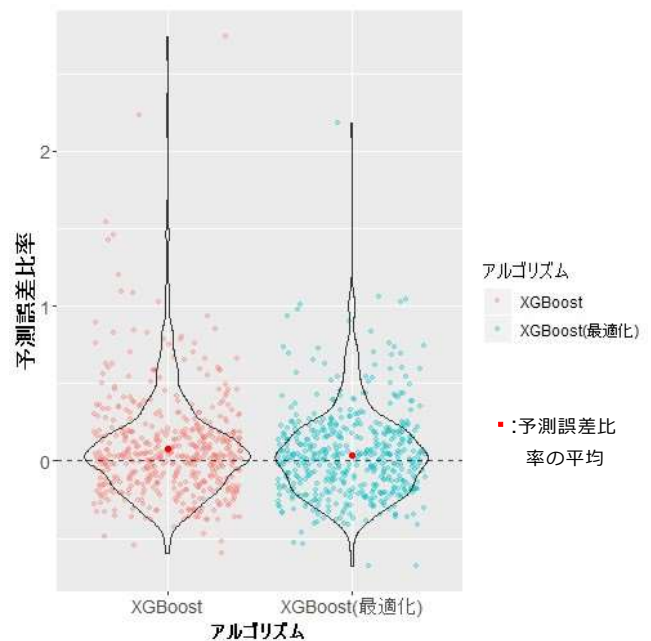


図 5 最適化前と最適化後の予測誤差比率の分布比較

(c) KNIME によるアプリケーション化

構築した機械学習モデルを使った予測を、普段 KNIME を利用したことのない設計技術者も利用できるよう、アプリケーション化した。KNIME Server を使うことで、ブラウザ上で動作する仕組みを構築した条件を入力して鉄骨重量合計を予測するだけでなく、入力した条件と類

似た過去実績の一覧や、スパンなどの条件を変化させた場合の予測結果の変化を表示できるようにし、見積や設計を支援する機能を強化した(図6)。これにより、見積期間を短縮するとともに、若手技術者も過去のデータを容易に利用できるようになった。

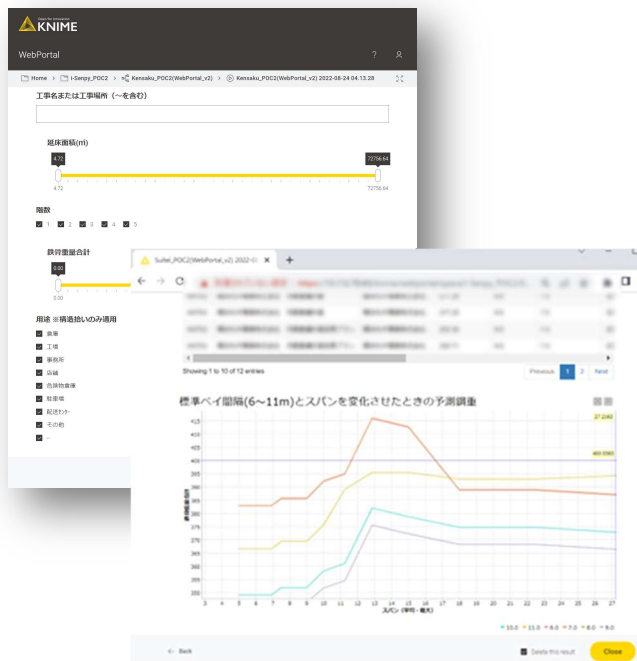


図6 アプリケーションのUIの例

### (5) 適用した結果と今後の課題

作成したアプリケーションを、建築事業部の設計担当者間で試用した結果、次のような意見を得た。

- 鉄骨重量合計の予測値を得られることは、大きなミスを防ぐために一定の効果があり、設計を早く行うための助けになる
- 一方で、実務に耐える高い精度や信頼性があるとはいえず、予測結果が大きな予測誤差を含む可能性も残されている

一定の効果の評価する声がある一方、精度や、算出過程の透明性・説明性についてはまだ課題が残されていることから、このモデルの実務適用は見送られている。実務適用するための課題として以下の事項が挙げられる。

- ① 予測精度を高める必要がある。対策の一つとしては、訓練データの増強が挙げられ、実物件のデータだけでなく、設計シミュレーションなどの訓練データを拡充する方法もある
- ② モデルの説明性向上や、高い精度が期待できる条件の範囲およびその範囲で期待できる精度の評価を行うことが必要である
- ③ 過去の蓄積データの中には、未入力項目や不正確な情報が含まれていた。今後はデータの質を向上させるため、データ入力時に、信頼できる情報

が蓄積できるような仕組みにする

これまで述べてきたように実務運用に向けての課題はあるものの、本開発を通して、手元にある過去のデータから、鉄骨重量合計をある程度予測することは可能であることがわかった。

また、③については、蓄積データの利活用や分析を想定してこなかった経緯により、残すべきデータを蓄積することの重要性を深く認識していなかった背景があると考えられる。今回の取り組みを機に、デジタルトランスフォーメーションにも活かせる重要な気づきを得ることができた。昨今のAIをはじめとする急速な技術的進歩を鑑みれば、過去のデータが今後様々な価値を生み出していくことが考えられる。したがって、データの量と質を高めていくための取り組みは、ますます重要になると考える。

## 6. おわりに

本報告では、機械学習の一般的な知識としての概要を紹介し、さらに鉄骨重量の予測をシステム化することで設計を効率化する機械学習の適用事例について述べた。その中で、データの抽出・前処理から、パラメータの最適化、モデルの評価、アプリケーション化までの流れを紹介した。また、これら一連のデータ分析や開発過程をノーコードで建築系技術者でも完結できることが確認できた。建築系技術者が自身の手で機械学習モデルを構築できたことは、特徴量の合理的かつ迅速な選択などの点で活かされた。

本事例が、業務の効率化や暗黙知の形式知化、プロセスの合理化などの課題解決の参考になれば幸いである。

### 参考文献

- 1) KNIME AG, 「KNIME Analytics Platform」, <https://www.knime.com/knime-analytics-platform>, (2022/1/15 閲覧)
- 2) Chen et al.: XGBoost: A Scalable Tree Boosting System, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp.785–794, 2016.8.