

説明可能 AI の実現に向けて：地震動指標予測での取り組み例

久保久彦・功刀卓・鈴木亘・木村武志・青井真（防災科学技術研究所）

1. はじめに

ディープラーニングなどの機械学習 (ML) が我々の日常の様々な場面に浸透しつつあり、地球物理学の分野でも様々な場面で活躍の場を広げている (e.g., Kong et al. 2018; Bergen et al. 2019)。他方で、ML、特にディープラーニングに関しては、「ブラックボックス」問題が往々にして指摘されている。医療や金融、軍事などの分野において ML アルゴリズムを実装していく際には、ブラックボックス問題に関連する様々な社会的要請を答えていく必要があり、透明性や説明責任が今後の ML 関連の研究開発ではますます重要になっていくと想定される (e.g., 福島・他 2017)。そのような状況の中で近年人工知能分野において、説明可能 AI (explainable AI, XAI) に関する研究が精力的取り組まれている。XAI は、予測結果や推定結果に至るプロセスが人間によって説明可能になっており、脱ブラックボックス化 (ホワイトボックス化) した ML モデルのことを指す。人の生死に関わる防災分野においても、ML アルゴリズムの将来的な実装は当然想定されるが、その実装の際には XAI であるかどうかの一つのハードルになると考えられる。本研究では、防災分野の XAI を目指す取り組みの一環として、これまで著者らが行ってきた地震動指標の ML 予測モデル (e.g., 久保・他 2018a 人工知能学会; 久保・他 2018b 地震学会; 久保・他 2019 人工知能学会) を題材として、その予測結果および ML モデル自体の解釈を試みる。発表では ML モデルにおける説明変数と目的変数の関係の可視化、予測結果の全出力による予測結果全体の可視化、ML モデルの可視化に関しても触れる予定であるが、このアブストラクトには ML モデルにおける説明変数の重要度に絞って記載する。

2. ランダムフォレスト予測器

久保・他 (2018a) では、地震動指標の一つである地表最大加速度 (PGA) を予測対象とし、震央距離・モーメントマグニチュード (M_w)・震源深さ・予測地点での S 波速度 1400m/s 層の深さ (D_{1400})・予測地点の地下深さ 30m までの平均 S 波速度 (V_{s30}) の 5 つを説明変数とする予測器をランダムフォレスト (RF) によって構築している。RF は、木構造を模したモデルによって回帰・分類を行う決定木を複数用いて、決定木の森を構成し予測を行うアンサンブル学習アルゴリズムである (e.g., Breiman 2001; Zhou 2017)。決定木自体は if 文の集合体であり、ホワイトボックス型の手法である。しかしながら複数の決定木を組み合わせる RF の場合、一つ一つの if 文が最終的な予測結果にどのように効いているのかわかりづらく、ニューラルネットワークほどではないにしてもブラックボックス的な要素を多く含む手法とみなすことができる。特に木構造が深い決定木を多数

用いる場合には、その傾向が顕著となる。

本研究でも久保・他（2018a）と同様のやり方で、RF 予測器を用意した。本研究で用いた RF は、最大深さ 100 の決定木計 100 で構成される。学習データは、久保・他（2019）で用いたデータセットを一部リバイスしたものを利用した。

3. 説明変数の重要度

作成した予測器において「どの説明変数が重要なのか」を知ることは、MLモデルと予測結果を理解する第一歩である。RFおよび決定木では木構造における各説明変数へのアクセス頻度から、各説明変数の重要度を算出できる。以下ではこれをアクセス重要度と呼ぶ。2. のRF予測器におけるアクセス重要度は【震央距離, Mw, 震源深さ, D1400, Vs30】 = [0.40, 0.26, 0.10, 0.12, 0.12] となる。震源距離とMwの値が他に比べて大きく、二つを合わせると全体の三分の二を占めていることがわかる。これは震源からの距離と地震の規模でPGAの大きさがある程度決まることを示しており、これまでの地震学の知見と整合する。

ただしアクセス頻度に基づく重要度の計算はconsistencyを満たさないことが先行研究によって指摘されている(e.g., Lundberg et al. 2018)。そこでPermutation Importance (PI; Breiman 2001; Fisher et al. 2018)を用いた重要度の評価も行った。PIは、ある一つの説明変数に関してデータ配列の順序をシャッフルしたデータを用意し、シャッフルしたことがモデルの精度にどの程度影響を与えるかを評価することで、重要度を求めるものである。ELI5を用いて計算したところ、【震央距離, Mw, 震源深さ, D1400, Vs30】 = [0.9038±0.0083, 0.4836±0.0074, 0.1035±0.0015, 0.1498±0.0020, 0.1372±0.0023] が得られ、PIでもアクセス重要度と同じ傾向が見られた。ただし、震央距離・Mw・その他三つの間の重要度の差がより強調されている。なおPermutation Importanceが負の値だと、その説明変数の順序を変えた場合に予測器のパフォーマンスが向上したことを示すが、今回はそのような変数はなかったことがわかる。

加えてSHAP (SHapley Additive exPlanations; Lundberg and Lee 2017; Lundberg et al. 2018)を用いた評価も行った。SHAPはゲーム理論のShapley valueに基づき、各特徴量が予測にどのような影響を与えたかを測る指標であるが、アクセス重要度・PIと同様の結果が得られた。