

機械学習による土砂災害発生予測モデルに対する説明可能な AI の適用

鹿児島大学 正会員 ○伊藤 真一, 酒匂 一成
 鹿児島大学 学生会員 笛田 泰成, 藤門 裕武

1. はじめに

近年, 土木分野において AI 技術の導入が期待されているが, 深層学習などを用いて構築された AI モデルは, 予測性能は高いものの, モデルの内部構造がブラックボックスになる傾向がある。技術者の判断が地域住民の生命や財産に関わる土木や防災の分野において, AI 技術をより導入していくためには, AI モデルのホワイトボックス化は重要な課題である。AI モデルが導き出した予測結果を説明するための方法として, 近年, 説明可能な AI (Explainable AI, 以下 XAI)¹⁾が注目されている。本研究では, 豪雨時の土砂災害の素因と誘因を学習させたメッシュ単位での土砂災害発生予測モデルに対して, XAI を適用した場合の効果について明らかにすることを目的とする。

2. XAI の手法

XAI による説明は, 大域説明と局所説明に大別される。大域説明は, AI モデルの全体的な振る舞いを理解することを目的としており, どの特徴量(パラメータ)が効いているかなどを把握するための方法である。局所説明は, 個別の予測結果の判断理由を理解することを目的としており, ある一つの入力に対してなぜその結果を出力したかという説明をすることができる。本研究では, 大域説明の方法として Permutation importance²⁾, 局所説明の方法として LIME³⁾を用いた。Permutation importance は, 確認したい特徴量の値をランダムに並び替えてモデルに与えることで, その特徴量が重要であるかを判断する手法である。LIME は, 説明対象とするデータの近傍データをランダムに発生させ, それらの近傍データを用いて線形回帰モデルのような近似モデルを構築し, その近似モデルによって局所説明を行う方法である。これらの手法の詳細は参考文献¹⁾を参照されたい。

3. 適用データの作成とモデル構築

平成 26 年 8 月豪雨によって多数の土砂災害が発生した標準地域 3 次メッシュ(以下 1km メッシュ)400 個分の範囲を分析対象⁴⁾とした。土砂災害が発生した 1km メッシュが 69 個である。DEM データと解析雨量を用いて, 各 1km メッシュの崩壊 or 非崩壊を推定するための 6 つの特徴量(傾斜の平均, 傾斜の標準偏差, ラプラシアン_平均, ラプラシアン_標準偏差, 時間雨量, 土壌雨量指数)を算出した。各特徴量の算出方法は, 既往の文献⁴⁾を参照されたい。また, 合計 400 個のデータは 300 個の学習データと 100 個のテストデータにランダムに分割した。モデル構築のために用いた機械学習のアルゴリズムはランダムフォレスト²⁾である。表-1 はランダムフォレストによって構築されたモデルに対して学習データとテストデータを適用した場合の混同行列を示している。構築されたモデルは, 学習データは正確に再現できているものの, テストデータに対しては合計 8 個の予測の誤りがある。これに加えて, モデルの内部がブラックボックスであるため, 構築されたモデルを土砂災害発生予測モデルとして実際に活用するのは難しい。次章では, このモデルに対して XAI を適用することで, モデルの説明を試みるとともに, 土砂災害発生予測モデルに対する XAI の効果について考察する。

4. XAI による構築されたモデルの説明

まず, 構築されたモデルの大域説明を試みた。図-1 は Permutation importance によって得られた各特徴量の重要度を示している。この結果から, 土壌雨量指数, ラプラシアンの標準偏差, 時間雨量の順で影響が大きく, ラプラシアンの平均の影響が最も小さいことがわかる。豪雨時の土砂

表-1 混同行列
(a) 学習データ

		予測	
		非崩壊	崩壊
実績	非崩壊	249	0
	崩壊	0	51

(b) テストデータ

		予測	
		非崩壊	崩壊
実績	非崩壊	79	3
	崩壊	5	13

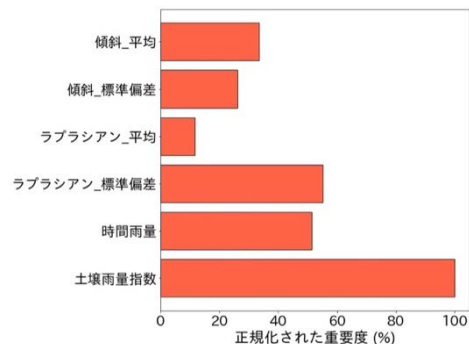


図-1 Permutation importance の結果

キーワード 説明可能な AI, 機械学習, 土砂災害, 降雨

連絡先 〒890-0065 鹿児島県鹿児島市郡元 1-21-40 鹿児島大学工学部海洋土木工学科 TEL : 099-285-8482

災害を対象としているため、土壌雨量指数や時間雨量の影響が大きいことは容易に想定できる。ラプシアンに関しては、標準偏差が大きければ1kmメッシュ内での尾根地形や谷地形が発達していることを意味しているため、ラプシアン標準偏差は土砂災害の発生に影響すると考えられる。以上のことから、構築されたモデルは、土砂災害発生に影響が大きい特徴量を重要だと判断できていることがわかる。このようなモデル構築前に想定したモデルの内部構造と実際のモデルが重要視しているポイントを比較できる点が、大域説明の利点の一つである。

次に、個別のデータに対する局所説明を試みた。ここでは、構築されたモデルが予測を誤った8つの事例の中で代表的な2つの事例について局所説明を行う。表-2は対象とするデータの特徴量、正解ラベル、LIMEによる予測値を示している。まず、事例1について、図-2はこの1kmメッシュの標高の分布図を示しており、図-3はLIMEによって算出された特徴量の条件ごとの影響を示している。この結果から、構築されたモデルは、この1kmメッシュは急斜面が多く谷地形が発達しているため素因的には危険だが、時間雨量が39mm/hと少ないため非崩壊と判断したことがわかる。また、LIMEの予測値が0.43であることから、崩壊する可能性もあると考えながら非崩壊と推定しているともいえる。このような個別のデータに対する予測の根拠を説明できるのが局所説明である。図-4は事例2の1kmメッシュの標高の分布図を示しており、図-5は特徴量の条件ごとの影響を示している。事例2では激しい雨が降ったことは危険だと判断しているが、平地が多く傾斜の平均も低いため非崩壊と判断していることがわかる。この結果は、構築されたモデルよりも局所的な急斜面を考慮できないデータの与え方に問題があったと考えられる。このような現状の入力データの問題点を抽出できる点も局所説明の利点の一つである。

5. おわりに

本研究では、豪雨時の土砂災害の素因と誘因を学習されたメッシュ単位での土砂災害発生予測モデルに対して、XAIを適用した場合の効果について考察した。その結果、XAIを用いることで、構築されたモデルに対する影響が大きい特徴量を把握できること、個別データに対する予測根拠を説明できること、現状の入力データの問題点を抽出できることなどが明らかになり、XAIは機械学習によって構築されたモデルの信頼性を評価するための手法として有効であることがわかった。

謝辞：本研究は科学研究費補助金(若手研究，課題番号：20K14828)の援助を受けて行われました。

参考文献：1) 大坪直樹ほか：XAI(説明可能なAI)そのとき人工知能はどう考えたか？，リックテレコム，

2021. 2) Breiman, L.: Random Forests, Machine Learning, Vol.45, pp.5-32, 2001.

3) Ribeiro, M. T. et al.: "Why Should I Trust You?" Explaining the Predictions of Any Classifier, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp.1135-1144, 2016.

4) 伊藤真一ほか：機械学習を用いた1kmメッシュごとの斜面崩壊に対する危険度評価，地盤工学会誌，Vol.66, No.9, pp.8-11, 2018.

表-2 対象データの特徴量，正解ラベル，LIMEによる予測値

事例番号	特徴量						正解ラベル y	予測値 LIME
	Ang_ave	Ang_std	Lap_ave	Lap_std	Hourly	SWI		
1	35.01	9.41	4.1.E-04	0.05	39.00	223.87	1	0.43
2	8.12	10.03	4.8.E-04	0.01	90.00	303.80	1	0.23

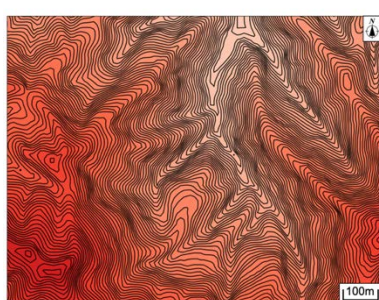


図-2 標高の分布図(事例1)

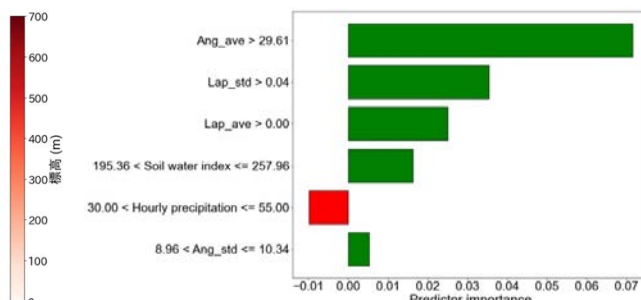


図-3 LIMEの結果(事例1)

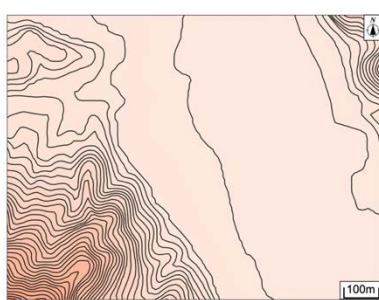


図-4 標高の分布図(事例2)

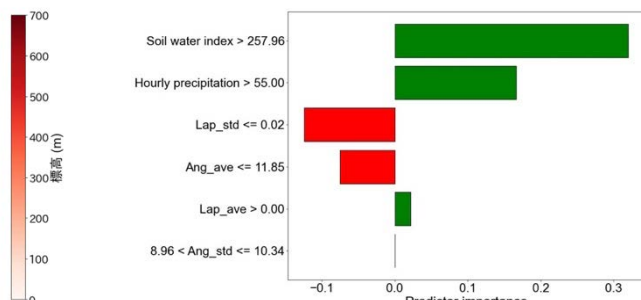


図-5 LIMEの結果(事例2)