

自治体保有データとオープンデータを用いた空き家確率推定モデルの構築

東京都市大学 学生会員 ○富田 健人
東京都市大学 正会員 秋山 祐樹

1. 目的

近年、日本では人口減少や高齢化により全国で空き家が増加し続けており、空き家の分布状況を調査することが自治体の重要な業務となっている。しかし、現在の空き家の分布調査の手法は、現地調査（外観目視）が中心となっているため、調査に多大な時間、労力、費用を要することが大きな課題となっている。そこで、自治体の所持しているデータを用いて、空き家の分布を迅速・安価に把握する手法が求められている。近年では様々な統計情報や空間情報を活用して、空き家の分布状況を把握・推定する研究も増えつつあるものの、データの利用上の制約などの課題もある¹⁾²⁾。また、研究成果の社会実装（自治体での利用）を見据えた研究は少ない。

そこで本研究では、群馬県前橋市を対象に自治体が保有するデータと国勢調査等のオープンデータを組み合わせ、空き家の分布推定を行うためのデータベース（以下「空き家データベース」）を構築し、実際の空き家分布情報を教師データとする機械学習により、市全域の空き家の空間分布を推定する手法を開発する。また、同手法の信頼性の検証や空き家と地域特性の関連を調査する。さらに、自治体の空き家担当者へのヒアリングにより、同手法の有用性や改善点を明らかにする。

2. データベースの整備と機械学習モデルの構築

(1) 空き家データベースの整備

まず、表-1 に示すデータを取得し、利用可能な変数を整理した。次に、自治体データに住所に基づいて位置情報を付与した。最後に建物データに対して、各データを空間結合処理によって紐付け、1 つのデータベース（空き家データベース）を整備した。

(2) 機械学習モデルの構築

本研究では、建物ごとの空き家確率（以下「空き家率」）を推定するために、決定木ベースの機械学習モデルである「XGBoost (eXtreme Gradient Boosting)」³⁾を採用した。本手法を採用した理由は、推定精度が高く、欠損値

を扱うことができるためである。学習モデルによって算出された空き家確率が 0.5 以上の場合は空き家、0.5 未満の場合は非空き家と分類した。

3. 結果

(1) 構築モデルの精度評価

表-2 に全データにおける推定結果を示す。正解率（正しく空き家あるいは非空き家と推定された割合）は 89.8%と高い精度となった。また、実際に非空き家のうち、正しく非空き家と予測できた割合（特異度）は 90.6%、実際に空き家のうち正しく空き家と予測できた割合（再現率）は 61.9%となった。

(2) 空き家分布推定結果

図-1 に前橋市全域における 500m メッシュごとの推定空き家率を示す。このように市域全体の指標が把握できることにより、自治体が実際に現地調査を行う際に、重点的かつ早期に空き家調査を実施すべき地域を検討するのに有益な情報となるものと期待される。

(3) 特徴量の重要度と構築モデルの解釈

本研究では構築したモデルに対して、各説明変数が予測値をどのように変化させているのか、ということについての検証を行った。以上の検証には「SHAP

表-1 元データとデータベースに入れた変数の一部

使用したデータ	変数
住宅地図	
水道使用量データ	各年最大使用量 各年の使用比
住民基本台帳	世帯内最高年齢 世帯内最低年齢 世帯人員数 年齢別世帯人員数
国勢調査	雇用形態、世帯構成など
用途地域データ	各建物が立地する地点の用途地域
標高・傾斜度データ	家から駅までの平均直線勾配など
空き家現地調査結果	空き家=1, 非空き家=0

表-2 全建物に対する予測結果（混同行列）

		推定値	
		非空き家	空き家
真値	非空き家	73,595 (TN)	7,651 (FP)
	空き家	872 (FN)	1,415 (TP)

キーワード 空き家、自治体データ、機械学習、SHAP、ユースケース
連絡先 〒158-8557 東京都世田谷区玉堤 1 丁目 28-1
TEL 03-5707-0104

(SHapley Addi-tive exPlanations)」と呼ばれる手法を用いた⁴⁾。図-2に、SHAPを用いて得られた説明変数の重要度上位10項目ごとのビースウォーム図を示した。このように、建物単位で変数のSHAP値が推定できたことにより、特定の地区ごとに空き家確率を高めている変数を把握することが可能となり、地区ごとの空き家対策の政策の指標になる可能性があると考えられる。

4. 自治体担当者へのヒアリング調査

本研究の成果を自治体の空き家関連の業務において利活用していく、すなわち社会実装を目指すためには現場を知る自治体職員の声を取り入れることが重要である。そこで、群馬県前橋市の空き家事業を担当している2つの課にヒアリング調査を実施した。

(1) 建築・住宅課

建築・住宅課は特定空き家の除却等、空き家対策全般を担っている課である。まず、メッシュや町丁字ごとに集計した予測結果は、対策優先地域を絞ることが出来る点で有用性を高められることが分かった。また、空き家周辺の建物の密集度合いが分かることで、近隣住宅への影響の程度を定量化できるというアイデアを得た。

(2) にぎわい商業課

にぎわい商業課は中心市街地の活性化など、空き家の利活用を推進している課である。近年、中心市街地への出店希望者が増加しており、特に賃料・維持費を抑えられるという観点から、空き家への需要も高まっていることが分かった。また、市街地の空き店舗や空き家を毎年調査していることから、本研究で得られた成果がこの調査の代替になりうる可能性があるという意見が聞かれた。

5. まとめ

本研究の手法により、前橋市全域の空き家および非空き家の空間分布を約90%という高精度で推定することが可能となった。また、SHAPを用いることで、地域特性と空き家の関連についても把握可能となった。さらに、自治体へのヒアリングを通して、本研究の有用性や改善点が明らかとなった。

今後の展望として、現地調査結果を保有していない自治体を考慮し、ある自治体で学習したモデルを他の自治体に対してどれほど外挿することが可能か検証を進めていく。また、ヒアリングを別の自治体でも実施し、ユースケース作りのための協議を今後も継続していく。

謝辞：本研究は東京都市大学総合研究所デジタル都市空間情報研究開発ユニットの研究成果の一部である。また、本研究で使用した前橋市の各種公共データや空き家調査データの使用に際し、前橋市未来政策課に多大なるご支援を頂いた。さらに、本研究は東大CSIS共同研究(No.880)の一環として実施した。以上、ここに記して謝意を表したい。

参考文献

- 1) 石河正寛, 松橋啓介, 金森有子, 有賀敏典: 住戸数と世帯数に基づく空き家の詳細地域分布の把握手法. 都市計画論文集, Vol.52, No.3, pp.689-695, 2017.
- 2) 馬場弘樹, 秋山祐樹, 谷内田修: 自治体保有データを活用した空き家の空間分布の将来予測モデル構築—群馬県前橋市を対象として—. 土木学会論文集 D3 (土木計画学), vol.77, No2, pp.62-71, 2021.
- 3) Chen, T. and Guestrin, C.: Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 785-794, 2016.
- 4) Scott, M, L. and Su-In, L. “A Unified Approach to Interpreting Model Predictions,” Proceedings of the 31st International Conference on Neural Information Processing Systems, 4768-4777, 2017.

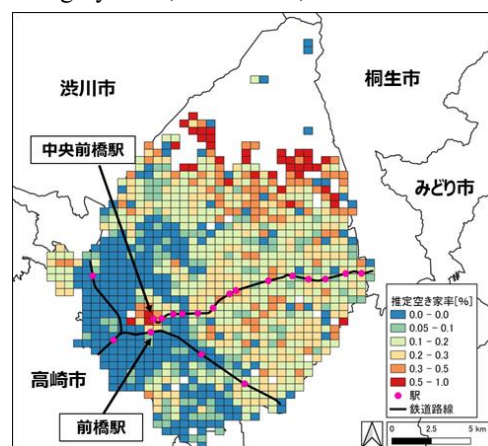


図-1 前橋市全域の推定空き家率 (500m メッシュ単位)

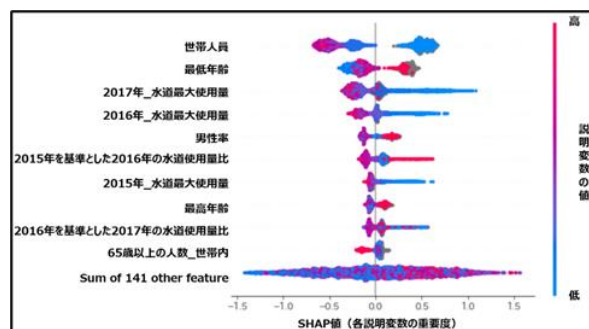


図-2 重要度上位10変数のビースウォーム図