

ILP を用いた人口減少都市における引越の傾向分析

大谷紀子研究室

1132155 土屋竜佑

1. 研究の背景・目的

近年、人口減少に悩む都市が増加している。人口減少都市では住民の流出により都市の収入が減少するため、少ない支出でより多くの経済効果が得られることが望ましい。したがって人口減少都市において、住民が都心部から外へと引っ越すことは望ましくない。引越をする住民のルールを知ることで、理想的な引越をしない住民の望む政策をとることが可能である。先行研究[1]では、富山市を対象地域としてアンケート調査を実施し、転居履歴を取得している。富山市は「公共交通を軸としたコンパクトなまちづくり」の実現を対象政策としており、郊外に住んでいる世帯が少しでも都心部へ移住するよう取り組んでいる。「まちなか居住助成金」は取組の一部であり、都心部へ移住する世帯に対して助成金を交付している。アンケート調査から得られた 5,089 世帯のデータのうち、過去に富山市内から富山市内への転居経験がある世帯および転居経験のない世帯のデータで、重大な欠損値がないデータを抽出して決定木で分析している。結果、「仮にバス停が家の近くにできたら買い物する場所が変わり、延べ床面積が 250 m²で世帯規模が 3.6 人より多く、交通が不便でないことが転居理由になる」という駅勢圏外から駅勢圏外へ引っ越す世帯のルールが得られ、公共交通機関と連携した買い物利便性の高い空間を提供することで、住み替えの可能性が少なからずあることがわかっている。

本研究では、人口減少都市における住民の引越ルールを帰納論理プログラミング (ILP ; Inductive Logic Programming) を用いて抽出することで、ILP が決定木における分析よりも有用な引越ルールを発見することを目指す。

2. 帰納論理プログラミング

ILP とは、述語論理を表現言語として帰納推論を行い、目標概念を得るための機械学習の手法である。ILP は観測されている事例から一般化された概念を抽出する。ILP では述語論理を用いるため、命題論理を用いる決定木と違い、背景知識と呼ばれる事例に関する知識を与えることが可能である。ILP は複数の表で表現される問題を扱える性質から、決定木より広範囲な問題を対象とできる。目標概念に対して正しい事例を正例、正しくない事例を負例と呼ぶ。機械学習における帰納推論には決定木を構築する手法もあるが、本研究では ILP を採用し、Progol という ILP の代表的なシステムを使用する。

3. 適用手法

Progol による学習を行う際、正負例・背景知識などの問題に関する記述には Prolog を用いる。対象となるデータ間に成り立つべき関係を記述する点が、C 言語などの手続き型プログラミング言語とは異なる。ILP は連続値の扱いが苦手なため、連続値を離散化させることが必要となる。富山市のアンケートでは、年齢や住宅の床面積などの情報は連続値で表されているが、閾値を用いて連続値を区切り、離

散値へと置き換えることによって ILP の適用を可能にする。

表 1 に住民の年齢を離散化したクラスと閾値を記す。最寄り駅からの距離が 2km 未満の世帯を駅勢圏内、2km 以上の世帯を駅勢圏外と定義した。転居に関して、A:「駅勢圏内⇒駅勢圏内」、B:「駅勢圏内⇒駅勢圏外」、C:「駅勢圏外⇒駅勢圏内」、D:「駅勢圏外⇒駅勢圏外」、E:「転居なし」の 5 クラスを設定し、クラス A と C を「望ましい転居」、クラス B と D を「望ましくない転居」とする。目標概念は「望ましくない転居」を行う世帯のアンケートデータを正例、「望ましい転居」を行う世帯と「転居なし」世帯のアンケートデータを負例として与える。背景知識には年齢・性別・職業・交通手段・自分専用車の有無などを含む「世帯属性」や、住宅の広さ・道路交通利便性が高いなどを含む現在と前の住宅の「居住選択理由」、住宅に対する不満な点を表す「不満点」、家業・結婚・経済的理由などを含む「転居理由」などのデータを与える。

表 1：年齢の離散化

クラス	閾値 (単位：歳)
old	65 以上
advanced	51 以上 65 未満
middle	36 以上 51 未満
young	23 以上 36 未満
us	19 以上 23 未満
hss	16 以上 19 未満
jhs	13 以上 16 未満
ess	6 以上 13 未満
childhood	6 未満

4. 評価実験

ILP を用いた分析の有用性を検証するために、評価実験を行った。富山市のアンケートデータのうち、駅勢圏の「外⇒外」と「中⇒外」へ転居しているサンプルを用いて「望ましくない転居」を行う世帯における引越ルールを求め再現率と適合率を求めた。実際に「望ましくない転居」を行った世帯のデータのうち ILP で得られた「望ましくない転居」を行う世帯における引越ルールが当てはまるデータの占める割合を示したものが再現率、ILP で得られた「望ましくない転居」を行う世帯における引越ルールに当てはまるデータのうち実際に「望ましい転居」を行った世帯の占める割合を示したものが適合率である。パラメータ設定ごとに求められた再現率と適合率を表 2 に示す。

表 2：パラメータ設定ごとの再現率と適合率

パラメータ設定	1	2	3	4	5	6	7	8	9	10
再現率(%)	37.0	57.6	67.4	68.0	70.9	57.3	57.0	56.0	78.2	75.3
適合率(%)	100.0	80.5	100.0	99.1	83.0	97.3	97.3	97.3	87.3	85.3

5. 考察

評価実験において、ILP が人口減少都市の引越ルールの抽出手法としてある程度有用であることが判明した。連続値データの離散化によるクラス分けによって実験結果は著しく変わる。また、ILP に入力するパラメータ設定によって抽出されるルールは大幅に変わる。閾値やパラメータの設定を改善することにより、より有用な引越ルールの抽出が可能になると考えられる。

参考文献

- [1] 福岡裕介, 大谷紀子, 杉木直, 宮本和明, “人口減少都市における策手段の戦略的選択と世帯セグメンテーション”, 第 48 回土木計画学研究発表会論文集, 226, 2013.