

レポート第1回

機械学習が自分の研究にどう生かされるか？

情報学専攻経営情報学コース

1930099

服部凌典

私は間取り図を考慮した賃料予測モデルに関する研究を行なっている。

研究概要は次である。

賃貸物件において、属性が全て同じ物件は数少なく、類似した属性を持つ物件でも築年数や階数、間取りなどのどこかに違いがある。

特に、同一の間取りの規格(1K や 2LDK など)であったとしても、各部屋の配置や面積などは物件毎に異なる。また、日本では賃貸物件を探す際に間取り図を見る習慣があり [清田 2017]、多くの物件では間取り図が用意されている。

物件属性は賃料に影響を与えと考えられており [河合 2008, 石島 2010, 福井 2018]、賃料決定の際に利用可能な手法として ヘドニック・アプローチがある。

間取り図の記載内容も物件属性のひとつと見なすこともできるが、これまでのヘドニック・アプローチによる賃料予測の事例では考慮されてきていない。

本研究では、間取り図を用いた賃料の線形回帰モデルを構築し、間取り図を用いない線形回帰モデルと予測精度を比較することで、間取り図の活用が賃料予測に与える影響を明らかにする。前者の線形回帰モデルでは、間取り図から抽出した特徴量を利用するが、特徴抽出器として、主成分分析(PCA: principal component analysis)と畳み込みニューラルネットワーク(CNN: convolutional neural network)の適用を検討する。

このように私は賃料の予測モデルに機械学習を用いて研究している。

機械学習は、観測データを元にパラメータを決定する為、観測データの質と量に予測精度が依存すると考えている。本件は LIFULL HOMES データセットから約 9 万件のデータセットを使用しているため、量に関しては十分であると感じている。

ただし、データセットの中身は各不動産業者が手打ちで記入したものであるため、欠損値や有り得ない値が多く含まれている。

私は、自分の研究に機械学習を用いる際には、データセットの質を担保するデータ処理が最も重要な要素になっていると感じている。

また、賃貸物件の賃料予測の場合、教師データとして使用する真の賃料は、不動産鑑定士によって鑑定された鑑定賃料となっている。そのため、不動産鑑定士によって、同じ物件でも賃料にばらつきが生じる。機械学習を用いる際、この教師データに関するばらつきを説明変数に誤差項加えることで補完するなどの処理が必要になってくる。

また、賃貸物件の家賃に影響を与えている変数に、海外情勢や経済、宗教的習慣などがある。

これらの変数の関係性は複雑で、非線形なものである[DOUGLASB.DIAMOND, 1985]。線形モデルでは、この非線形な部分を表現することができないので、ニューラルネットワークなどの非線形モデルがこれら変数間の非線形性を反映する方法として考えられる。

しかし、ニューラルネットワークなどの非線形モデルでは予測モデルの式が複雑になり、モデルの意味を理解するのが困難になる場合が多い。

今後、私はこの問題を解決した説明性を保ったニューラルネットワークの構築を考えていきたい。

また、本研究では、ある物件の変数を入力して、出力をそれに対応する一つの賃料を出力する点推定モデルを考えている。

そこで、今後の展望として、出力に期待値を出力するモデルを考えたい。

期待値を出力するモデルを考えることで、分散がわかるので、どれくらいの確率で出力結果が正しいのかを判断することができる。

本研究は不動産鑑定士の鑑定支援を目的としているため、予測結果がどれくらい正しいのかを鑑定士に伝えることは重要であると考える。

そこで、今後、ガウス過程回帰モデルやベイズ推定モデルなどを考えていきたい。

ただし、本研究では、9万件のデータと間取り図を考慮しているため、ガウス過程回帰などの計算コストが大きく、現実的に不可能な計算時間が必要になる可能性がある。

さらに、賃貸物件の場合、賃料に影響を与える変数同士が独立であると仮定できない。

そこで、グラフィカルモデルを利用した変数選択をする必要が発生する。そのためにも、不動産に関する知識と経済学的知識が必要となる。

自分の研究に関する機械学習の生かし方は上で挙げたようなものが、どれも単にデータセットを入力するだけでは予測精度やその後の解釈ができないので、この講義を通して、適切なモデルの選択や、予測結果の解釈方法などを学んでいきたい。

参考文献

清田 陽司, 山崎 俊彦, 諏訪 博彦, 清水 千弘:

不動産と AI, 人工知能, Vol.32, No.4, pp.529--535 (2017)

河合 伸治 :

ヘドニック・アプローチによる賃貸住宅価格の価格決定要因の推定 --西武池袋線の賃貸住宅を事例として--, ソシオ サイエンス, Vol.14, pp.49--63 (2008)

石島 博, 谷山 智彦 :

個別性と歪みを考慮した住宅価格分析とパーソナルファイナンスへの応用, ファイナンシャル・プランニング研究, Vol.10, pp.4--17 (2010)

福井 光, 阪井 一仁, 南村 忠敬, 三尾 順一, 木下 明弘, 高田 司郎 :

レインズのニューラルネットワークを用いた不動産価格査定について, 人工知能学会全国大会論文集, Vol.JSAI2018, No.4A2-03, pp.1--4 (2018)

DOUGLAS B. DIAMOND, JR. :

Simultaneity in the Market for Housing Characteristics',
JOURNAL OF URBAN ECONOMICS Vol.17, pp.280--292 (1985)