

[博士論文審査要旨]

申請者 大槻 健太郎

論文題目 不動産価格予測モデルのためのデータクレンジング法とモデル更新アルゴリズムに関する研究

審査員 中川 秀敏、中村 信弘、横内 大介

不動産価格査定業務の DX 化により、不動産取引ビッグデータを活用した不動産価格査定システムが多くの企業で開発されている。とりわけ、実装が比較的容易なディープラーニングをはじめとする統計的機械学習法を利用した AI による価格査定システムの開発が進んでいる。しかしながら、不十分なデータクレンジング、非効率なデータベース化、機械学習モデルの非可読性（ブラックボックス化）などの多岐にわたる問題により、現在稼働しているシステムの多くは、価格の推定精度や値付けに対する説明力が不十分だと言われている。

本論文の目的は、不動産価格査定システム開発におけるデータクレンジング、データベース構築および価格推定モデルの維持管理という 3 つのフェーズに潜み、価格推定に対して影響を及ぼしている実務上の問題点を明らかにし、その解決方法を統計的な手法として提案することにある。具体的には、物件の属性やダミー変数を用いたヘドニックアプローチによる線形モデルを前提として、モデルフィッティングに必要なデータ収録値のクレンジング法、同一のモデルが適用可能な地域を特定するデータグルーピング手法、市況変化による価格推定モデルの陳腐化を防ぐモデル更新アルゴリズムの構築の 3 つである。大槻氏は、これらの課題に対して、昨今流行している複雑な機械学習法を使うのではなく、伝統的な統計科学の手法を改良、応用することで課題解決を図っている。

大槻氏の博士論文は、以下の 5 つの章と付録からなる。第 1 章「序論」では、一般的なデータサイエンスプロセスと不動産価格予測フローとの対応関係を示しつつ、本研究の背景と目的について説明している。

第 2 章では、国交省データベースの中古マンションデータを例に、既存のデータクレンジングにおけるいくつかの問題点を指摘している。そして、その中から特に重要だと思われる取引価格の桁間違い入力エラーを取り上げて、それを自動的かつ適切に検出する新たなデータクレンジング方法を提案している。この提案方法の特徴は、取引価格の異常値を経験分布の極値としてとらえる伝統的な外れ値検出の方法に加えて、地域性や中古マンションの物件属性を反映した階層的クラスタリングを取り入れてその非類似性を異常値発見へ応用したことにある。本論文では、提案方法を東京都 23 区の中古マンション取引データに適用し、中古マンション価格モデルの推定精度が従前と比べて大幅に改善できることを確認している。

第3章でSREホールディングス株式会社提供の中古マンション取引データを用いて、時間経過に伴う中古マンション価格推定モデルの劣化問題に対処するため、モデル構築のために使う取引実績データをどの程度の地域まで細分化すれば、査定現場の感覚との乖離がなく、かつ予測精度が経時的に劣化しづらい価格推定モデルが実装できるかという、不動産価格推定システムのためのデータベース構築のあり方について取り上げている。本研究ではこの問題を、想定しうる最小単位でのデータ分割の実施と予測精度が落ちない形でのデータの再統合ととらえなおすことで解決している。

第4章は、不動産市況の急激な変化に対応するべく、新たに観察されるデータの内容に応じて価格推定に用いるデータ期間の幅を自動調整する可変スライディングウィンドウによるモデル推定アルゴリズムを提案している。大槻氏はこのアルゴリズムを、クックの距離の改良とF検定という伝統的な統計手法の組み合わせで実装している。本論文中では、SREホールディングス株式会社提供の中古マンション取引データを用いて、大槻氏の提案する可変スライディングウィンドウによる価格推定モデルのあてはめ結果と既存手法である固定スライディングウィンドウでのあてはめ結果を比較しており、提案手法によって推定された価格推定モデルが、対象期間全般にわたって市況の変化に追従し、価格推定の安定性を維持できていることを確認している。

第5章では、全体のまとめに加えて、本研究において残ったいくつかの課題について論じている。

以上のように、大槻氏の論文は不動産価格推定モデルの構築やそのモデルの実運用で生じているいくつかの問題に対する新たな解決法は提示できていると評価できる。一方で、論文の内容には改善すべき点が残っている。たとえば、本研究で設定した課題はすべて不動産実務の作業フローによる制約を受けているにも関わらず、その点を論文の中で明確に示していない。そのため、当該分野の研究や実務に通じていない読者にとっては、本研究における問題設定そのものが理解しづらい記述になっている。また、本論文では、ヘドニックアプローチによる線形モデルを用いた不動産価格推定モデルのみを想定しているため、不動産価格推定モデルそのものに対する貢献が限定的になってしまっている点も残念である。しかしながら、これらの点については今後の研究で十分改善できる内容であり、不動産価格査定システムの構築と運用に内在する問題を提起し、その問題解決に正面から取り組むという大槻氏の研究姿勢および今回の研究成果は大いに評価できる。

よって、審査員一同は、最終試験の結果とあわせ、本論文が一橋大学学位規則第5条第1項の規定により一橋大学博士（経営）の学位を受けるに値するものと判断する。