



자동가치산정의 정확성 분석 및 투명성 관리 방안

An Analysis of the Accuracy of Automated Valuation Model and Transparency Management Plan

구강모* · 임동준**

Kang Mo Koo · Raymond Chetti

■ Abstract ■

This study provided measures to improve the accuracy and reliability of automated valuation models (AVMs), which have been gradually expanding their scope of use. The US market, which introduced AVMs earlier than other markets, has improved the requirements for AVM fairness, accuracy and reliability through the Dodd-Frank Act, while the South Korean AVM market mostly remains unregulated, leaving AVM firms to regulate themselves. When valuations of domestic AVM firms which are possible to be massively obtained were analyzed by mainstream statistical standards such as median absolute value percentage error and percentage predicted error@10%, levels of accuracy were different in regional types and the same result applied to apartments and officetels. This suggests that even for the same AVM company, the level of accuracy could be different from regional types and that it is necessary to present accuracy standards for each region. In addition, when the accuracy of AVM is stable and validated, it could be used as a means to prevent non-performing loans and Jeonse fraud due to the gap between market and appraised values. Moreover, the scope of the AVM could be expanded to monitoring collateral values in the secondary market of real estate-backed securities and inspecting investment portfolios. To achieve this, it is necessary to establish accuracy standards for them and introduce an third-party, non-profit and independent verification agency.

Keywords: Automated valuation model (AVM), Valuation, Over-valuation, Dodd-Frank Act, AVM validation

* 연세대학교 미래캠퍼스 경제학과 부교수(주저자 · 교신저자) | Associate Professor, Department of Economics, Yonsei University Mirae Campus | First Author · Corresponding Author | kang.koo@yonsei.ac.kr |

** 자이랜드 대표이사 | CEO, XAI Land | ray.chetti@xai.land |

1. 서론

1. 연구의 배경과 목적

국내에 비교적 늦게 도입된 자동화 가치산정 방법(automated valuation model, AVM)은 빅데이터와 알고리즘을 활용하여 쉽고 빠르게 부동산의 가치산정을 실행할 수 있다는 장점이 있다. 국내의 일부 금융회사는 AVM 기술과 사전 대출 심사 서비스를 접목하여 대출 서비스를 보다 편리하게 제공하고 있다. 그러나, 국내 AVM 기업들은 서로 다른 알고리즘을 사용하고 있으며 자료 수집, 처리, 정교화 방법 또한 상이하다. 국내의 현실적인 문제점은 이러한 과정의 투명성 및 결과물에 대한 정확도 검증 기준이 명확하지 않다는 점이다. 금융소비자의 보호를 목표로 2010년 처음 고안된 미국의 Dodd-Frank Act는 2023년 6월, AVM을 활용하는 대출 및 2차 주택담보 대출 시장에 대한 기준을 제시하면서 AVM의 관리 강화를 시사하였다. 국내는 AVM의 관리·감독 및 정확도에 대한 검증 요건이 없는 상황으로, AVM의 정확도 및 신뢰도 향상 방안 논의가 필요한 상황이다. 아울러 AVM의 신뢰도가 확보되면, AVM 활용은 감정가-실거래가 격차 확대와 연결되는 과당 대출 문제를 차단하는 데 도움을 줄 것으로 기대된다.¹⁾²⁾ 감정가-실거래가 격차 문제는

해외에서도 오랫동안 연구되어 왔다. Agarwal et al.(2015)은 실제 가치(true value)보다 높게 감정평가가 발생한 경우, 부실채권으로 연결될 가능성이 증가한다고 주장하였다. Ben-David (2011)는 거래가격의 2%~6%에 이르는 거래 비용을 거래가격에 포함하여 가격을 높게 평가하는 행위가 미국에서 많이 발생했었고, 감정평가사들은 거래 비용을 포함한 가격을 거래가격으로 승인하는 경향이 있다고 보고했다. Conklin et al. (2020)은 감정평가사의 과도한 경쟁이 감정가-실거래가 차이가 크게 나타날 확률을 증가시키며, 감정평가사들의 경쟁이 1 표준편차 증가할 경우 거래가격과 동일한 수준으로 평가한 감정평가의 시장 비율(share of at-price appraisals)³⁾은 3.7% 증가한다고 주장하였다. 이러한 감정가-실거래가 격차는 AVM을 활용하여 좁혀질 수 있다고 전망되나, AVM의 폭넓은 활용을 위해서는 먼저 정확도에 대한 기준 마련이 필요하다. 이에 본 논문은 AVM의 정확도 비교 분석과 AVM 적용 확대 가능성 논의를 연구 목표로 하였다. AVM이 이미 널리 사용되고 있는 미국은 AVM을 활용하는 기업들에게 주기적으로 정확도 검증을 요구하고 있다. 이러한 정확성 기준 및 신뢰도 확보가 어떻게 이루어지고 있는지에 대하여 연구함으로써, 이와 유사한 방안들이 국내에 적용 가능한지를 우선 살펴보고자 한다. 다음으로 대표적인 AVM의

1) 부당대출 관련 금융사고가 증가하고 있으며, 이는 은행뿐 아니라 제2금융권에도 같은 방법으로 악영향을 미치고 있다고 보고되었다(송승희, 2024).

2) 감정가-실거래가 격차 확대와 연계된 과당 대출은 상업용 부동산에도 영향을 미치고 있는 것으로 파악되고 있다(김나경, 2024).

3) Calem et al.(2018)은 거래가격과 동일한 감정평가(at-price appraisal)의 증가는 감정가-실거래가 격차 확대와 연결되며, 대출위험관리 측면에서 정보손실을 일으킨다고 주장한다. 실제 거래가격보다 낮은 가치를 보이는 감정평가서는 거의 존재하지 않으므로, Conklin et al.(2020)은 거래가격과 감정평가 금액이 동일 경우가 증가하는 경우를 감정가 과대평가의 대리변수로 활용하였다.

정확도 판단 방법들을 사용하여 국내 AVM 추정치의 비교 분석을 실시하고자 한다. 마지막으로, 사회적 문제로 떠오르고 있는 과당 대출에 대한 방지책으로서의 AVM 활용 가능성에 대한 논의를 진행하고자 한다.

2. 연구의 범위와 방법

연구 목표에 따라 본 논문은 아래 두 개의 연구 질문을 제시하고 분석을 실시하였다.

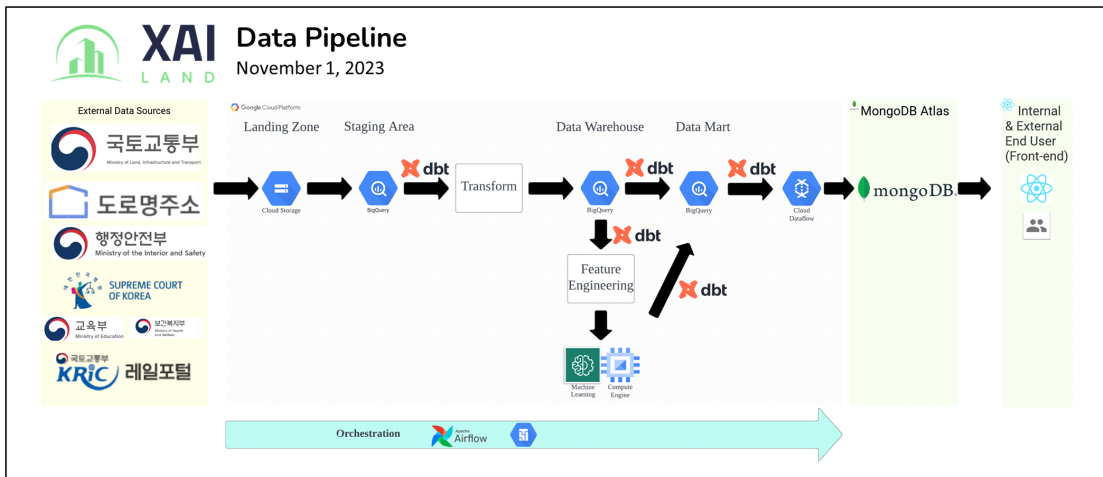
- a) AVM이 활성화된 미국은 AVM의 정확성을 어떻게 관리하고 있는가?
- b) AVM의 정확성 검증방법에 따른 국내 AVM 기업의 추정치 사이에 유의한 차이가 존재하는가?

AVM에 사용되는 알고리즘은 기업별로 다양하며 이러한 알고리즘에 대한 검증이 필요한바, 미국에서의 AVM 규제가 어떻게 전개됐으며, 최근의 동향은 어떠한지를 우선 논의한다. 이를 위하여 2000년대 초반부터 2024년 Dodd-Frank Act의 수정까지 주요 규제 변화를 살펴보았다. 아울러 AVM의 정확도 분석에 주로 사용되는 방법론인 MdAPE(median absolute value percentage error), PPE(percentage predicted error)@10% 등의 방법을 살펴보고 알고리즘별로 추정치의 정확도 차이가 어느 정도 발생하는지 실증적으로 검사하여 AVM 활용방안에 대한 정책적 함의를 도출하였다.

II. Automated Valuation Model 활용에 대한 이론적 논의

부동산과 기술을 의미하는 Property와 Technology를 합성하여 만들어진 프롭테크(PropTech)는 2017년 이후 급속도로 성장하기 시작하였고, 기술을 활용한 부동산 시장의 선진화 및 효율성 증대가 프롭테크의 본질이라고 할 수 있다(구강모, 2022; 허윤경·김성환, 2019). 프롭테크의 응용 분야 중, 기계학습과 빅데이터를 활용하여 부동산의 가치산정을 시행하는 행위를 일반적으로 자동 가치산정이라고 하며, 정주희(2022)는 AVM을 “부동산 정보를 이용하여 개별 부동산에 대한 특정 시점의 시장가치를 자동으로 추정하는 통계 기반의 컴퓨터 프로그램”으로 정의한다. 이는 컴퓨터 기반 대량 감정평가(computer-aided mass appraisal, CAMA)⁴⁾를 시작점으로 하며, 컴퓨터를 사용하여 대량의 감정평가를 진행하는 CAMA와 다르게 AVM은 개별 물건의 가치평가를 대상으로 하는 데 차이점이 있다고 서술한다. 또한 정식 감정평가는 감정평가사의 현장 방문이 수반되는 반면, AVM은 현장 방문이 없이 진행된다는 점에서 정식감정평가의 대체가 될 수 없으며 탁상자문과 유사한 성격을 갖는다고 구분하고 있다. 현장 방문 없이 데이터와 모델에 의존하는 AVM은 <그림 1>에서 보이는 바와 같이, 원자료를 취합하여 각 AVM 활용 기업에서 데이터 처리 과정을 거치고 개별적으로 개발한 모형에 적용하는 것이 일반적 구조다. 국내 가용 데이터 수준을

4) 한국감정원이 자체 개발한 담보 시세 자동산정시스템(e-시세)이 CAMA의 한 예이다.



자료 : XAI Land.

〈그림 1〉 AVM 기업의 데이터 파이프라인 처리 과정

고려하면, 원자료의 수준은 기업별로 크게 상이하
지 않을 것으로 판단된다. 그러므로 효율적이고
효과적인 데이터 처리 과정 및 적합한 분석 모형
선택이 AVM의 정확성과 직결된다고 할 수 있다.

본 연구는 비교적 최근에 제도적 논의가 시작
된 국내 AVM 시장에서 법규적·제도적 보완에 대
한 시사점을 제공할 것으로 기대한다. 박상빈 외
(2022)는 시세 산정 서비스를 감정평가업으로 볼
근거가 약하므로 법적으로 제재할 근거 또한 희박
하다고 주장하였다. 이러한 쟁점은 탁상자문과의
역할 비교에도 적용된다. 정주희(2022)는 탁상
감정과 감정평가의 차이뿐 아니라 탁상자문과
AVM의 차이 등에 대한 논란이 종결되지 않고 있
다고 주장한다. 아울러 금융위원회와 국토교통부
가 AVM에 관련된 직·간접적 정책을 내어놓았으
나, 데이터 구축 및 여러 모형에 대한 연구개발이
중장기적으로 추가 진행되어야 한다고 서술하고
있어 AVM 모델의 보완이 필요하다는 점을 시사

하고 있다. 또한 AVM을 활용한 금융기관의 내부
감사에 대한 가능성을 열어두었으나, 정확성 및
신뢰도의 확보 방안에 대한 논의는 다루어지지 않
고 있는 취약점이 발견된다.

정확도에 대한 평가는 실거래가와와의 차이로 검
증되는 것이 일반적이다. 그러나, 감정평가 금액
과 매매 가격은 현실적으로 동일하게 추정되기는
어렵다. Kok et al.(2017)은 MSCI에서 매년 발
표되는 감정평가 금액과 매매 가격의 차이 보고서
(Private Real Estate Valuation and Sale
Price Comparison)를 근거로 평가금액과 실거
래가격의 차이는 여러 국가에서 공통으로 관찰되
는 현상이라고 주장한다. 그러나, 그 차이가 현격
한 경우는 금융시장에 문제를 초래할 수 있다고
주장한다. Kok et al.(2017)에서 인용된 MSCI 보
고서의 가장 최근 연구는 2022년 1,816건의 거래
에 기반한 보고서이다(MSCI, 2023). 해당 보고서
에 따르면 2021년의 거래가격과 감정평가 가치의

차이는 11.9%를 보여 연구가 시작된 이래 최고치를 기록하였고, 2022년에는 가중평균 절대 차이(weighted-average absolute difference)가 11.1%로 소폭 하락하였다. 이렇듯 여러 나라에서 관찰되는 감정가격과 거래가격 차이는 과대 대출에 따른 부실채권으로 연결될 가능성이 크므로, 감정가-실거래가의 격차와 부동산담보대출의 채무불이행 관계는 학계에서 오랜 기간 연구가 되어왔다. 이러한 문제는 비단 최근의 문제는 아닌 것으로 보인다. 1990년 기준 미국 Fannie Mae의 10년 치 담보대출 매입 기록에 따르면 매입한 대출의 30% 이상이 감정가와 거래가격이 동일하고, 80% 이상은 감정가와 거래가격의 차이가 0%~5% 범위 안에 속하는 것으로 보고되었다(Cho and Megbolugbe, 1996).

결국 Fannie Mae 혹은 Freddie Mac과 같은 GSE(government sponsored enterprise)는 이러한 문제점에 초점을 둔 규제안을 제시하게 되었다. 미국 주택담보대출 2차 시장에서 가장 큰 역할을 하는 GSE는 감정가-실거래가의 폭의 확대를 방지하기 위하여 2009년 5월, HVCC(home valuation code of conduct) 규정을 발표하였으며, 이는 대출 실행에 관여하는 담당자, 모기지 브로커 등이 감정평가사에게 희망하는 감정평가액을 알려주거나, 평가과정에 의견을 표출하는 행위를 금지하는 내용을 포함하였다. Shi and Zhang(2015)은 이 규제의 실행 이후에 GSE가

매입한 대출채권의 부실채권화율이 유의하게 감소한다고 분석하였다. 이러한 감정가-실거래가 차이 확대의 문제는 국내에도 적용된다. 선우영(2023)은 2022년 발생했던 대규모 전세 사기⁵⁾의 원인으로 과대 평가된 감정가를 지목하였다. 특히 ‘무자본 갭투자’에 전형적으로 악용되면서 전세금을 올려받는 사기 사건이 발생하였고, 이는 비교거래사례를 찾기 어려운 빌라시장의 취약점에 기인하는 것으로 알려졌다. 그러므로, 아파트가 아닌 단독, 연립, 다세대 주택 등은 이른바 ‘업감정’과 결탁한 전세 사기의 위험요인이 크게 존재한다고 할 수 있다.

미국의 2차 담보 시장에 관련한 정부의 직접적 규제 이외의 해결 수단으로 프롭테크의 기술 활용을 주장한 연구가 있다. Kok et al.(2017)은 감정가의 차이는 기계학습(machine learning)을 활용한 자동 가치산정 방법을 통해서 좁혀질 수 있다고 주장했다. 이에 관련된 국내 학술연구는 AVM의 소개 및 발전 방향에 집중해왔다. 안지희(2020)는 AVM의 국가별 제도와 실태를 연구하여, 미국과 유럽 각 국가의 활용 정도를 보고하였다. 특히, 미국은 주거용 부동산의 거래가격이 40만 달러 이하인 경우 대주에게 감정평가 보고서 제출 의무가 면제되고, AVM을 사용할 수 있다고 보고한다.⁶⁾ 반면, 유럽의 경우 AVM 결과치는 실제 감정평가로서의 의미보다는 벤치마크, 혹은 사기 방지 수단으로서의 기능이 크다는 점을 강조

5) 백민정(2024)은 전세 보증사고의 피해 금액은 2조 원에 이르며, 이는 주택도시보증공사(HUG)의 재무 상태에 악영향을 끼칠 것으로 전망하였다.

6) 이에 대한 해석은 주의가 필요하다. 40만 달러 이상에 대한 거래는 AVM 활용을 불허한다는 의미로 해석이 가능하며, AVM의 활용에 따른 고가 시장의 불안이 발생할 수 있음을 내포한다. 그러므로, 이에 대한 문제의 방지 차원에서 40만 달러 이하에만 AVM 허용을 하고 있다는 해석이 가능하다.

하고 있다. 이훈(2022)은 국내 여러 AVM 기업들의 가치 추정 결과를 감정평가액과 비교하였으며, 특수 상황을 제외하고는 10%~20% 이내의 가격 격차를 보인다고 보고하고 있다. 그러나 공매와 같은 특수 상황이 존재하는 물건의 경우, AVM의 추정 결과치와 감정평가액과의 차이가 무려 140%까지 높아지는 것으로 나타났다. Matysiak (2017)은 AVM과 실제 감정평가액의 적정 차이 수준은 10% 수준이라고 제안하고 있으나, 저자 또한 서술하듯이 이에 대한 명확한 규칙은 거의 없는 실정이다.

이처럼 현격한 감정가-실거래가의 차이를 줄이는 보완 수단으로 AVM 활용을 추천하는 의견도 있었다. 그러나 AVM의 폭넓은 활용 이전에 신뢰성 확보가 선행되어야 할 것이다. AVM의 신뢰성 확보는 최근 금융시장의 부실·과당 대출과도 직결될 수 있는 문제이다. 부실·과당 대출은 보통 실거래가 보다 높은 감정가가 이면에 존재하게 되며 금융기관에서 내부적으로 이러한 문제를 감지하기는 쉽지 않다는 취약점이 있다. 내부적 감시체계가 명확하지 않은 경우, 동일 부동산에 대한 다수의 감정평가를 진행해야 하는데 이는 금전적·시간적 비용과 연결되며, 현실적으로 모든 대출에 대해 다수의 감정평가를 실행하기는 불가능하다는 한계점이 존재한다.

III. Automated Valuation Model 정확성 검증 수단 및 방법

1. 미국의 Automated Valuation Model 규제 변천 과정⁷⁾

미국의 AVM은 2000년대 초반에 효율성과 비용 절감의 장점 덕분에 부동산 평가 도구로 주목받기 시작하였다. 특히 담보대출 분야에서 활발히 사용되었지만, 초기 단계에서는 관리·감독 체계의 미비로 신뢰성에 대한 의구심이 제기되었다. ASC(The Appraisal Subcommittee)와 FFIEC(The Federal Financial Institutions Examination Council)이 기본 틀을 마련했으나 구체적인 규제는 없는 상황이었다. 2000년대 중반에는 AVM 활용이 증가하며 관리 기관들이 정밀 조사 필요성을 인식하게 되었고, Interagency Appraisal and Evaluation Guidelines에 AVM의 신뢰성과 정확성을 강조하는 규정을 포함하게 되었다. 모든 금융기관은 Supervisory Guidance on Model Risk Management 규정을 준수하여 모델 개발 과정을 문서화하고, 독립적인 전문가가 검증을 거치도록 요구받았다. 2008년 금융위기 이후, Dodd-Frank Act는 금융위기 원인을 해결하고자 2010년에 제정되었으며, 담보가치 산정에 사용되는 AVM의 엄격한 품질관리 기준을 포함하였다. AVM을 활용하는 금융기관은 정확성과 신뢰도를 확보하기 위해 정기적으로 테스트를 실시하고 이를 문서화해야 하며, 대출 실행과 감

7) 본 절은 Federal Financial Institutions Examination Council(2010), Federal Register(2023), Federal Reserve System(2011), U.S. Government Publishing Office(2010)의 내용을 종합적으로 재구성하여 기술하였다.

정평가 기능이 분리되도록 하여 이해 상충을 방지하는 규정을 적용받았다. COVID-19 대유행 중에는 비대면 서비스 수요가 증가에 따른 AVM의 적응력과 회복력에 대한 정책도 마련되었다. 이는 금융기관이 현장 방문 없이 대출을 실행할 수 있도록 AVM과 공공 데이터를 사용한 비대면 평가를 가능하게 하였다. 최근에는 2024년 Quality Control Standard가 도입되어, 확실성 확보, 데이터 보호, 이해 상충 방지, 무작위 샘플링 및 검증 요구, 비차별법 준수 등의 기준을 제시하였다. 비차별법 준수는 특히 다문화 사회로 진입 중인 국내 금융시장에도 중요한 시사점을 제공한다. 이와 관련하여, 미국의 연구에서는 인종에 따라 감정가가 다르게 나타날 수 있음을 지적하고, AVM이 차별적 요소를 줄일 수 있는 방안으로 평가되고 있다(Williamson and Palim, 2022).

이렇듯 미국의 AVM 관련 규제는 지속적으로 변화하였다. 국내 AVM 모델의 정확성 및 신뢰도에 대한 검증은 현재로서는 AVM을 활용하는 기관의 내부적 규율에 따라 자율적으로 통제되고 있다. 국내의 AI 활용 기술에 대한 인증은 한국 인공지능인증센터, KSA 한국표준협회 등의 기관에서 실행하고 있다. 그러나 많은 AVM 기업들이 스타트업 단계의 중·소규모인 점과 고가의 인증 비용을 고려하면, 정기적 인증 유지는 어려운 실정이다. 이는 대형 자본이 일부 기업에 편중될 때 문제가 더욱 심각해질 위험이 있다. 자본력이 뒷받침되는 기업은 AI 인증에 대한 지속적 비용지출이 가능하나, 중소기업의 경우 인증비용 부담이 상

당할 것이기 때문이다. 그러므로, 이러한 비용 부담에 대한 지원방안이 함께 수반될 필요가 있다. 미국의 대표적인 AVM 기업 House Canary(하우스캐너리)는 이러한 부담 및 검증 문제점에 대하여 제3의 비영리 기관이 설립되어 이해 상충관계 없이 AVM 모델을 검증할 필요가 있다고 주장하고 있으며(FDIC, 2023), 평균 퍼센트 오차(mean percent error, MPE), MAPE(mean absolute value percentage error), MdAPE와 같은 통계 및 PPE@5%, PPE@10%, PPE@20% 등을 사용하여 AVM의 표준을 지정하는 것을 제안하였다.

2. Automated Valuation Model 정확성 검증 방법

AVM 기술의 개발은 감정평가 기법과 기술에 근거하고 있으나, 모델을 정확히 특정(specification)하고 조정(calibration)하는 작업이 필요하며, 이러한 일련의 과정은 작업의 영역 설정부터 모델 명시, 모델 조정, 품질확인을 거쳐 실제 적용에 이르기까지 여러 단계를 거치게 된다(IAAO, 2018).⁸⁾ 그러므로, 기업별 정확도 수준은 같을 수 없으며, 이에 대한 통일된 기준이 필요하다. Matysiak (2017)은 감정평가와 AVM 추정금액의 차이를 5%, 10%, 20% 수준에서 검토해야 한다고 주장하지만, 그 근거가 명확히 제시되고 있지 않았고, 검증 방법에 따라 오차 수준이 다르게 나타날 수 있다. 아울러 여러 미국 AVM 기업들의 오차 수준

8) IAAO(2018)는 AVM 개발에 있어 기본 절차를 아래와 같은 아홉 단계로 설명하고 있다. Creation of a scope of work, identification and acquisition of property data, exploration data analysis, stratification, determination of data representativeness, model specification, model calibration, quality assurance, model application and value review가 그 아홉 단계이다.

이 666개 카운티 평균 6% 수준에 머무른다고 보고하고 있으나, 지역별로 편차가 크게 나타나고 있다. 이에 학계 및 AVM 기업들이 주로 활용하는 정확성 및 신뢰도에 평가 방법에 대한 다양한 모델이 제시될 필요가 있다.

기계학습에 대한 검증 방법론으로는 평균 절대 비율 오차(MAPE), 예측오차율(PPE), 중앙값 절대 비율 오차(MdAPE), 절대 비율 오차와 임계값의 차(percent error less than a predetermined threshold, PAPE@ τ %) 등이 제시되었으며, 정확성 추정에 널리 사용되어 왔다(Hu et al., 2019; Lee et al., 2024; Makridakis et al., 1979; Tan et al., 2019; Yrjö, 1976). 대표적 방법론의 공식은 <표 1>과 같다.

이 연구에서는 국내 AVM 활용 기업들의 정확도에 대한 비교를 통하여 기업 간 자동가치산정 결과치의 차이를 보임으로써 산정모델의 기준이

필요함을 설명하는 것이 목적이므로, <표 1>에서 소개된 방법론 중, MdAPE와 PPE@10%를 선택하여 대량 자료 확보가 가능한 두 기업의 자동가치산정 결과치를 비교하도록 한다. 일반적 연구 방법으로는 특정 기간(월 단위)을 선택하여 해당 기간 동안 발생한 총거래의 5%~10%에 해당하는 실거래가격 데이터의 확보 후, AVM 가치 추정을 실시한다. 그 다음 AVM 추정 가치와 실거래가의 차이를 고려한 오차율(percent error)을 계산한다. 오차율 계산은 (식 1)을 활용하며, 수집한 모든 관측치에 대한 오차율 계산을 반복한다(Yrjö, 1976). 단, 오차율 계산은 실거래가와 AVM 추정 가치의 격차를 계산하는 과정이므로 절댓값으로 계산한다.

$$\text{Percent Error} = \frac{\text{Estimated Value}}{\text{Sale Price}} - 1 \quad (\text{식 1})$$

<표 1> AVM의 대표 검증 방법론

구분	공식 및 변수 설명
MAPE	$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{ \hat{Y}_i - Y_i }{Y_i}$
	\hat{Y} : 예측가격; Y_i : 실제 거래 가격; N : 관측
MdAPE	$MdAPE = \text{median}\left(\left \frac{\hat{Y}_i - Y_i}{Y_i}\right \right)$
	median: 중앙값
PPE	$PPE_x = \frac{\text{Len}\left(\frac{ \hat{Y}_i - Y_i }{Y_i} < x\%\right)}{N}$
	Len: x% 기준 이내로 들어온 관측 수 \hat{Y} : 예측가격; Y_i : 실제 거래 가격; N : 관측수

주 : AVM, automated valuation model; MAPE, mean absolute value percentage error; MdAPE, median absolute value percentage error; PPE, percentage predicted error.

다음으로 중앙값 절대 비율 오차(MdAPE)를 선택한다. 중앙값 절대 비율 오차는 위의 단계에서 구해진 오차를 중에서 중앙값을 선택한다(Makridakis et al., 1979).

절대 비율 오차와 임계값의 차(PAPE@ τ %) 방법은 (식 1)에서 구해진 오차율이 일정 임계치 이내에 들어오는 경우의 비율을 계산하는 방식이다. 예를 들어 임계값 τ %를 10%로 설정하는 경우, 전체 관측치에서 오차율이 10% 이내에 들어오는 관측치 수를 계산하게 되고, 이 관측치 수를 전체 관측치 수로 나누어 백분율로 표시하게 된다. PAPE가 높게 나타나는 모델일수록 실거래가격과 AVM의 추정치의 차이가 미미하다고 해석할 수 있다.

IV. 국내 Automated Valuation Model의 비교 분석

1. 국내의 제도적 상황

국내에는 다수의 AVM 기업들이 존재하고 있고, 금융권과의 협업을 통하여 금융소비자들의 대출 사전 조사에 널리 사용되고 있음에도 불구하고, 최초의 관련 정책은 2022년 11월 14일 개정·시행된 「감정평가 실무기준 개정」이 유일하다. 해당 개정안은 「자동가치산정모형 사용 윤리」를 새로이 추가하였으며, 자동가치산정모형의 정의 및 역할에 대하여 명시하고 있다. 이는 자동가치산

정모형이 감정평가의 대체 수단이 될 수 없고, 보조적 수단으로 사용되어야 하며 AVM에 의한 추정 가치는 감정평가액으로 볼 수 없다는 법적 근거를 명확히 하였다는 점에서 그 의의가 있다고 하겠다. 개정 내용에서 불분명하게 명시되어 있는 부분은 「활용 시 유의사항」 부분으로, 감정평가법인 등이 AVM을 활용할 경우 모형의 알고리즘, 데이터 및 범위 적합성, 산출 결과물의 적정성에 대하여 “이해하고 검토”하도록 하고 있다. 현실적으로 알고리즘의 정확성에 대한 판단기준이 존재하지 않고 소비자에 대한 정확성 공시의무가 없는 현재 상황은 AVM 활용 기업들의 내부적 검증을 “유의사항” 수준으로 남겨두었다는 점에서 개선 여지가 있다고 판단한다.⁹⁾ 이에 하우스캐너리가 제안하는 내용과 같이 정부 기관 혹은 제3의 비영리 검증기관을 지정하거나 기업 스스로 정확성 검증 결과를 공시하도록 하여 AVM 알고리즘의 정확성에 대한 소비자에 대한 정보전달이 진행되어야 비로소 국내 AVM의 완전한 정착이 이루어질 것이다.

2. 국내 Automated Valuation Model 비교 분석

이 논문은 AVM을 활용한 추정치의 대량 데이터 수집이 가능한 두 기업의 모델¹⁰⁾을 비교하여 기업 간, 지역 간 오차율이 얼마나 크게 나타날 수 있는지 확인하여 실무에서 정확성 검증에 대한 가이드라인의 필요성을 보이려 한다. 이에 대하여

9) 김성훈(2024)은 금융감독원이 은행들에게 내부통제를 강화함으로써 미흡한 담보가치 평가의 적정성 확인 절차를 보완하라는 지시가 있었다고 보고를 하였으나, 내부통제에 맡긴다는 점에서 「감정평가 실무기준 개정」과 크게 다르지 않다.

2023년 8월 발생한 모든 거래에서 19,995건¹¹⁾의 거래를 무작위 추출하여 각 AVM 모델에서 추정하는 가치와 실거래 가치의 차이를 MdAPE, PPE@10%의 두 기준을 통해 제시한다. 분석 대상은 광역시 및 도(道) 단위로 구분하였고, 형태별로는 아파트와 오피스텔로 구분하여 분석을 진행하였다. <표 2>는 주요 분석 대상인 아파트 거

래의 기초 통계량을 보여준다. 지역별로는 경기도가 가장 큰 비중을 차지하고 있으며, 그 다음으로 서울특별시가 약 12%를 차지하고 있다. 지역별 사이에 아파트의 면적은 큰 차이를 보이지 않으나, 실거래가 및 AVM 추정가치는 지역적 차이가 확연히 나타나고 있다. 추정가치의 AVM 기업 간 차이는 부산광역시의 경우 실거래가의 약 6%

<표 2> 무작위 추출된 아파트 거래의 기초통계량

거래 지역	거래수	면적(㎡)		실거래가(백만 원)		추정가치(A사)		추정가치(B사)	
		평균	표준편차	평균	표준편차	평균	표준편차	평균	표준편차
강원특별자치도	669	71.7	21.4	208.4	108.1	209.7	109.5	209.6	106.1
경기도	5,745	74.7	21.9	461.4	217.3	468.1	211.5	463.4	214.4
경상남도	1,352	73.6	19.7	231.4	124.7	235.7	124.2	231.4	123.3
경상북도	606	71.4	18.6	168.3	98.2	168.3	99.6	167.7	98.0
광주광역시	816	73.2	19.0	270.2	147.0	272.3	144.5	272.1	143.4
대구광역시	1,210	75.5	18.8	291.1	135.0	299.5	133.5	284.9	131.5
대전광역시	748	77.2	21.1	336.0	158.2	340.4	158.6	332.9	155.7
부산광역시	1,183	77.7	21.8	364.9	195.1	380.5	197.6	358.9	188.4
서울특별시	2,387	74.2	21.4	1,012.5	463.3	1021.1	450.5	1002.2	453.6
세종특별자치시	153	77.9	18.4	547.5	173.2	549.4	171.3	538.7	165.5
울산광역시	517	76.4	19.5	302.5	167.2	304.9	164.7	298.3	161.2
인천광역시	1,448	73.8	21.0	361.2	161.9	365.8	155.6	364.3	161.2
전라남도	552	71.9	18.0	157.0	94.1	160.1	96.2	160.8	94.8
전라북도	845	70.7	19.4	188.7	113.2	192.1	114.1	191.8	113.7
제주특별자치도	63	78.3	14.6	461.6	157.9	460.0	153.7	460.2	160.0
충청남도	858	69.7	21.4	193.4	116.8	199.0	116.3	198.6	118.2
충청북도	843	67.6	20.3	189.5	111.6	190.8	110.3	192.5	112.2
전국	19,995	73.9	20.9	412.2	332.8	418.1	330.6	411.4	327.8

10) A사는 앙상블학습기반의 머신러닝모형, 보수 및 비모수 모형, 공간계량분석 등을 사용하여 가치추정치를 도출하며, B사는 XGBoost 하이퍼파라미터를 활용한 모델 및 지리가중회귀분석을 기초로 사용하는 것으로 파악되었다.

11) 이는 2023년 8월 발생한 아파트 거래의 51.9%를 차지하여, 일반적으로 사용되는 5%~10%를 충족시켜 샘플 데이터의 대표성이 충분히 확보되었다.

수준인 2,106만 원의 차이를 보여 가장 큰 격차를 보였다. 반면 강원특별자치도 및 제주특별자치도 등은 20만 원 수준의 차이를 보여 A사와 B사의 추정치가 근접하는 것으로 나타났다.

〈표 3〉은 아파트를 대상으로 한 AVM의 MdAPE를 보여준다. 전국적으로 MdAPE의 평균은 A사가 4.9%, B사가 3.2%로 실거래가와 AVM 추정가치의 절대 차이가 양사 모두 5% 미만으로 나타나 양호한 수준이라고 판단이 된다. 그러나 지역별로 나누어보면 최솟값과 최댓값의 차이는 두 배

가까이 되는 것을 알 수 있다. A사의 경우 최소 MdAPE는 3.6%, 최대 MdAPE는 6.6%를 보였고, B사는 최소 2.2%, 최대 4.9%를 보인다. 이는 지역별로 모델을 다르게 적용하는 것이 일반적이라는 사실을 고려하면 같은 AVM 기업이라 하더라도 지역적으로 정확도는 다르게 나타날 수 있음을 보여준다. 10%를 임계치로 설정하고 PPE@10%를 분석하면 AVM 기업 간 격차는 확연하게 나타난다. A사의 경우 오차율 10% 안에 들어올 확률은 평균 78%, B사의 경우는 87.6%이다. 즉

〈표 3〉 아파트 거래에 대한 MdAPE 분석 결과

거래 지역	MdAPE(%)		PPE@10%(%)		전체 관측 수
	A사	B사	A사	B사	
강원특별자치도	5.6	3.5	75.6	86.5	669
경기도	4.4	2.6	82.1	89.5	5,745
경상남도	5.2	3.1	76.1	89.1	1,352
경상북도	6.6	3.7	67.8	84.7	606
광주광역시	5.4	4.0	74.6	84.2	816
대구광역시	5.5	4.1	74.3	83.9	1,210
대전광역시	5.0	3.7	81.8	88.6	748
부산광역시	6.0	3.7	69.7	85.3	1,183
서울특별시	3.9	2.2	84.7	95.1	2,387
세종특별자치시	3.8	2.6	86.3	94.1	153
울산광역시	5.0	3.3	79.5	86.5	517
인천광역시	4.9	3.0	79.1	88.3	1,448
전라남도	6.3	4.9	67.8	73.7	552
전라북도	6.0	4.6	70.2	80.6	845
제주특별자치도	3.6	2.7	85.7	90.5	63
충청남도	6.0	4.2	71.4	82.2	858
충청북도	5.2	4.2	76.9	83.4	843
전국	4.9	3.2	78.0	87.6	19,995

주 : MdAPE, median absolute value percentage error; PPE, percentage predicted error.

12.4%~22%의 AVM 가치 추정치는 실거래가와
의 차이가 10%를 벗어나게 된다. A사의 경우 세종
특별자치시에서 86.3%를 보여 가장 높은 정확도를
보였고, B사는 서울특별시에서 95.1%의 정확도를
보였다. 그러나 두 기업 모두 최저값은 67.8%,
87.6%를 보임으로써 전국 평균 PPE@10%와의 차
이를 나타내고 있다. 10% 오차범위 바깥에 위치
하는 추정치들을 금액으로 환산하면 그 차이는 상
당하다. <표 4>는 10% 임계치를 넘어서는 표본들

의 오차(과대 평가금액) 분석을 보여준다. 총 과대
평가금액을 총합산하게 되면 A사는 1,735억 원,
B사는 630억 원이 되어, 두 회사의 추정치 차이
는 거의 3배에 이르게 된다. 이 분석은 2023년 8월
한 달 동안 발생한 거래의 무작위 추출 표본이라는
점을 상기하면 AVM 기업 간 연간 과대(혹은 과소)
평가 금액의 총합은 더욱 크게 나타날 것이다.

오피스텔에 대한 MdAPE 분석과 PPE@10%
분석은 표에 별도로 보이지 않았으나, 아래와 같

<표 4> 아파트 시장 PPE@10% 초과 표본 분석

거래 지역	PPE@10% 기준 초과 표본 수		10% 기준 초과 표본에 대한 총과대평가금액(백만 원)	
	A사	B사	A사	B사
강원특별자치도	100	72	2,834.8	1,320.8
경기도	82	449	50,488.2	23,717.2
경상남도	253	93	7,010.6	2,122.0
경상북도	102	49	2,191.5	944.6
광주광역시	138	97	4,571.0	2,781.4
대구광역시	276	72	10,861.8	1,975.7
대전광역시	105	39	5,319.0	1,375.2
부산광역시	315	67	17,534.7	2,214.7
서울특별시	307	55	41,770.8	8,237.5
세종특별자치시	18	4	1,447.5	360.9
울산광역시	79	44	2,804.0	1,037.8
인천광역시	243	139	11,148.3	6,308.8
전라남도	118	114	2,665.2	2,357.2
전라북도	174	122	4,040.0	2,684.4
제주특별자치도	6	3	249.5	78.9
충청남도	197	128	5,361.7	3,275.3
충청북도	114	110	3,191.4	2,233.7
전국	3,472	1,657	173,492.9	63,026.0

주 : PPE, percentage predicted error.

은 결과를 나타냈다. 아파트와 비교해 거래 수가 적고, 지역적으로 오피스텔 거래가 전혀 발생하지 않는 거래 지역이 다수 존재하는 하위시장의 특성상 충분한 정보제공에 한계가 있다는 취약점에도 불구하고, 두 회사 모두 MdAPE, PPE@10%의 결과는 아파트 시장 분석보다 정확성이 크게 하락하지는 않는 것으로 나타났다. A사는 오피스텔의 MdAPE가 6.2%, B사는 3.8%를 보였고, 최솟값은 양사 모두 1.4%를 보여 일부 지역에서는 상당히 높은 정확성을 보여주었다. 그러나, 최댓값은 각각 12.1%, 10.9%로 과거 거래기록 및 데이터 제공 가능 범위의 한계는 존재하는 것으로 보인다. 특히 주목할 점은 AVM 추정 가치가 임계치 10% 안에 들어오는 지역이 상당수 존재한다는 점이다. A사의 경우 광주광역시, 제주특별자치도의 오피스텔 추정 가치가 모두 10% 임계치 안에 들어왔으며, B사는 경상북도, 울산광역시, 제주특별자치도에서 모든 추정치가 10% 임계치를 넘지 않았다. 단, 이러한 높은 정확도는 거래 수가 적은 지역에 해당하는 것으로 나타났으며 지역별 거래량이 늘어나는 경우, PPE@10%도 함께 하락하는 추세를 보였다. 동일 기간 동일 주소지에 대한 AVM 추정 오차에 따른 총과대평가금액은 9.4억~47.2억으로 분석되었으며, 기업 간 과대 평가 금액 차이는 4배가 넘는 것으로 나타났다.

전국 평균 기준으로 판단하면, 아파트와 오피스텔 모두 B사가 더 높은 정확도를 보인다. 아파트의 분석에서는 B사가 최소 4.8%, 최대 16.9% 높은 PPE@10%를 나타내고 있으나, 오피스텔 시

장에서는 A사가 경상남도, 광주광역시에서 더 높은 정확성을 보였다. 이러한 점은 같은 AVM 기업이어도 지역적 모델 정확성이 상이하게 나타날 수 있고, 추정 방법에 대한 지역적 검증을 별도로 진행할 필요가 있음을 의미한다.

V. 연구의 정책적 함의

「감정평가 실무기준 개정」에 기술된 바와 같이 AVM은 감정평가의 효율성 향상을 위한 보조적 수단으로 활용이 되어야 한다. 이러한 보조적 수단으로서의 활용 방안에는 과당 대출 예방책이 포함될 수 있을 것으로 보인다. 전술한 바와 같이 과당 대출은 흔히 실거래가에 비해 과대평가된 감정가와 연결이 되어 있으며, 학술적 논의에 의하면 가격상승이 가파르거나 비교사례를 찾기 어려운 경우 더욱 심하게 나타나고 있다. Williamson and Palim(2022)에서 정의하는 감정가 과대평가는 정식 감정 가치와 AVM 추정 가치의 차이가 10% 이상 발생할 경우로 국한된다.¹²⁾ 국내 「감정평가 실무기준 개정」에 따르면 AVM은 감정평가의 보조수단으로 활용이 가능하다. 그러므로 금융기관의 부실·부당 대출과 연결되어 있는 감정가 과대평가를 사전에 감지할 수 있다면, AVM의 활용은 단순히 내 집 시세의 추정치를 알려주는 데 머무르지 않을 것으로 생각한다. 이와 같은 맥락에서 Baum et al.(2021)은 AVM의 미래 활용 분야에 2차 담보 시장적용을 논하였다. 모기지 저당

12) Diop et al.(2021)은 평균적으로 감정평가 가치는 AVM의 추정 가치보다 5% 높게 나타나며, 연구에 사용된 비교사례의 60%에서 감정평가 가치가 AVM 추정치보다 높게 나타난다고 보고하였다.

증권(mortgage-backed securities) 시장에서 AVM의 활용은 투자자에게 위험 노출을 알려줄 수 있고 포트폴리오의 지속적 모니터링을 위해 대량 평가를 가능하게 한다고 주장한다. 실제 판례에서도 FHFA(Federal Housing Finance Agency)와 Nomura Holding America 등과의 소송에서 AVM에 비해 감정평가액이 과대평가 되었다는 주장에는 15.1% 임계치를 사용하였는데, 이는 전문가인 The Appraisal Institute가 사용했던 10% 임계치와 다르다는 내용이 법원 판결문(Casetext, 2015)에 등장한다. 다른 나라도 10%~15%를 일반적으로 사용하는 것으로 조사되었다.¹³⁾ 이처럼 국내의 AVM 정확성 및 신뢰도 확보가 선행되어야 담보물의 감정평가가 과도하게 평가되고 있는지를 확인하는 검증 수단으로 AVM 활용이 가능하다는 것은 판례 및 학술 논문에서 이미 논의된 사실이다. 신뢰도 확보 이후에는 AVM 기술의 적용이 1차 담보 시장뿐 아니라 2차 시장의 발행 및 포트폴리오 관리 차원까지 확대 가능할 것으로 보인다. 이와 더불어 AVM의 미래 확산 방향은 감정평가가 필요한 대주의 업무 영역에 전반적으로 적용이 가능할 것으로 보인다. 이는 앞서 이야기한 부동산 담보대출과 재융자(refinancing)뿐 아니라, 부동산 투자 및 개발 대출에도 확대 적용이 가능하다는 것을 의미한다. 특히 부동산 개발 업무에서 분양가 산정은 최근 비교 거래를 바탕으로 하고 있어 분양 시점의 적정 분양가 추정 능력이 충분히 확보되지 않고 있

는 것으로 판단된다. 또한 대출 이후 사후 모니터링에도 활용이 가능할 것이다. 부동산 가치는 지속적으로 변하고 있으므로, 대출 실행 이후 담보 물건의 가치 변동 모니터링이 AVM을 통하여 가능할 것으로 예상한다. AVM의 광범위한 활용을 위해서는 대출 및 재융자, 그리고 2차 담보 시장에 관련된 금융기관들로 하여금 AVM 활용 기업들과의 협업 및 제휴를 필수 조건으로 정하여 금융 소비자들에게 편의를 제공할 필요가 있다. 아울러 다수의 AVM 기업의 추정치를 활용하도록 장려하여 모델에 따른 AVM 추정치 차이를 비교할 수 있는 토대가 마련되어야 한다. 이를 위해서는 본 논문의 분석 방법 및 결과를 토대로 각 AVM 기업들의 정확성 검증 및 공개를 필수 조건으로 하는 선결 작업이 필요할 것이다.

IV. 결론 및 연구의 한계

Baum et al.(2021)은 개발국의 경우 30%~70%의 담보대출 심사가 AVM을 활용한다고 보고한다. 이를 비추어 볼 때, 자동가치산정의 활용은 프롭테크의 발전과 더불어 날로 증가하고 있다는 것을 유추할 수 있다. 비교적 일찍 AVM의 활용을 시작한 미국은 Dodd-Frank Act를 시작으로 관련 법령 제도 정비에 관계기관들의 협업을 통해 꾸준히 이루어진 반면, 국내 AVM 관련 제도는 이제 모양을 갖추기 시작하였다. 최근 미국에

13) 영국은 FCA(Financial Conduct Authority), RICS(Royal Institution of Chartered Surveyors)가 기준을 제시하며, 일반적으로 $\pm 15\%$ 의 임계치를 사용한다. 호주의 경우, APRA(The Australian Prudential Regulation Authority)가, 뉴질랜드는 The Reserve Bank of New Zealand가 담당하고 있고 10%~15%의 임계치를 사용한다.

서 개정된 Dodd-Frank Act 관련 규제들은 AVM을 활용하는 금융기관 및 제2차 시장 참여자들에게 공정성 및 정확성 검증에 대한 요구를 명시하고 있다. 그에 반하여, 국내 시장은 정확성에 대한 검증 기준이 정비되어 있지 않으며 내부통제에 맡겨놓은 상황이다. 향후 AVM은 정식 감정평가의 보조수단 뿐 아니라 탁상 감정 및 부실 대출 모니터링 수단, 포트폴리오의 담보 평가 등의 영역까지 확장이 가능하지만, AVM 기술의 정확성과 투명성, 공정성이 확보되어야 국내 시장에 안정적으로 자리 잡을 수 있을 것이다.

본 논문에서 약 20,000건의 거래 사례를 분석한 결과, 국내 AVM 기업의 부동산 추정 가치는 전반적으로 양호한 오차율을 보였으나, 지역별·유형별 차이가 확연하게 나타났다. 이러한 분석 결과는 정확성 검증이 지역적·유형별로 실행될 필요가 있음을 입증하였다. 이에 제도 정비 작업을 통하여 AVM의 정확성에 대한 기준 마련이 필요하며, 스타트업 기업의 형태를 갖는 다수의 AVM 기업의 상황을 고려할 때, 독립적인 제3의 비영리 검증기관 설립이 필요함을 보여준다. 더욱이 AVM 기술이 지속 확장하고 있는 상황 고려 시, 정확성 기준 설정 전담 기관의 필요성도 대두되고 있다.

본 연구는 두 개의 AVM 기업의 자료를 수집하여 실거래가와와의 차이를 비교하였으나, 더 많은 AVM 기업들로부터 자료를 구할 수 없어 비교 대상 및 연구 기간의 확장에 한계가 있다. 또한 이상값으로 판단되는 거래는 데이터 처리 과정에서 누락되는 것이 일반적으로, 데이터 처리 과정에서 모든 거래에 대한 AVM 추정치를 비교 분석할 수 없다는 한계점이 있다. 이러한 한계점은 AVM 정

확도에 대한 기준이 정립되는 경우 해결될 수 있을 것으로 기대한다.

ORCID

구강모 <https://orcid.org/0000-0001-7497-5255>

임동준 <https://orcid.org/0009-0001-9095-3205>

참고문헌

1. 곽상빈, 전경근, 최진희. (2022) 부동산 시세 산정 서비스 관련 감정평가 관계 법규상 쟁점에 관한 소고 *대한부동산학회지*, 40(2), 5-23.
2. 구강모. (2022). *부동산 플랫폼의 생태계 발전 방향*. <https://www.hanaif.re.kr/boardDetail.do?hmpeSeqNo=35158>
3. 김나경. (2024.04.14.). “과당대출 대거 적발되나” 은행, 298조 상업용부동산대출 조사... ‘내부통제 시스템 실패’ 관건. 파이낸셜뉴스. https://www.fnnews.com/news/202404101426590917?cm=news_headline
4. 김성훈. (2024.07.11.). 은행 이사회 의장 소집 급감원 “대출 내부통제 보완해야”. SBS Biz. <https://biz.sbs.co.kr/article/20000181222>
5. 백민정. (2024.05.19.). 전세보증사고 2조 육박... HUG가 대신 갚은 돈 회수율은 17%. 중앙일보. <https://www.joongang.co.kr/article/25250248#home>
6. 선우영. (2023.10.31.). 전세사기로 드러난 감정평가사 만났 “눈 뜨고 코 베이는 세입자”. 프라임경제. <https://www.newsprime.co.kr/news/article/?no=617375>
7. 송승희. (2024.05.25.). 감정평가사에 갑질하고,

- 브로커랑 짬짜미...은행도, 새마을금고도 반복되는 '대출사고', 왜[머니머니]. 헤럴드경제. <https://biz.heraldcorp.com/view.php?ud=20240524050646>
8. 안지희. (2020). *자동차자산정보형(AVM)의 제도와 실태연구*(2020-09). 한국부동산연구원.
9. 이훈. (2022). 프롭테크 업체 자동평가모형(AVM)의 적정성에 대한 연구. *한국퍼실리티매니지먼트학회지*, 17(1), 13-20.
10. 정주희. (2022). *금융기관의 탁상자문 및 AVM 활용에 따른 쟁점 및 과제*(2022-13). 한국부동산연구원.
11. 허윤경, 김성환. (2019). *프롭테크 기업 부동산 산업의 새로운 미래*. 한국건설산업연구원.
12. Agarwal, S., Ben-David, I., & Yao, V. (2015). Collateral valuation and borrower financial constraints: Evidence from the residential real estate market. *Management Science*, 61(9), 2213-2280.
13. Baum, A., Graham, L., & Xiong, Q., (2021). *The future of automated real estate valuations (AVMs)*. <https://www.sbs.ox.ac.uk/sites/default/files/2022-03/FoRE%20AVM%202022.pdf>
14. Ben-David, I. (2011). Financial constraints and inflated home prices during the real estate boom. *American Economic Journal: Applied Economics*, 3(3), 55-87.
15. Calem, P. S., Lambie-Hanson, L., Nakamura, L. I., & Kenney, J. H. (2018). *Appraising home purchase appraisals* (Federal Reserve Bank of Philadelphia Working Paper No. WP 18-28). Federal Reserve Bank of Philadelphia.
16. Casetext. (2015). *Fed. Hous. Fin. Agency v. Nomura Holding Am, Inc.* <https://casetext.com/case/fed-hous-fin-agency-v-nomura-holding-am-inc-3>
17. Cho., M, & Megbolugbe, I. F. (1996). An empirical analysis of property appraisal and mortgage redlining. *The Journal of Real Estate Finance & Economics*, 13, 45-55.
18. Conklin, J., Coulson, N. E., Diop, M., & Le, T. (2020). Competition and appraisal inflation. *The Journal of Real Estate Finance & Economics*, 61, 1-38.
19. Diop, M., Yavas A., & Zhu S. (2021). *Appraisal inflation and private mortgage securitization*. http://abfer.org/media/abfer-events-2021/annual-conference/papers-real-estate/AC21P6009_Appraisal-Inflation-and-Private-Mortgage-Securitization.pdf
20. Federal Deposit Insurance Corporation [FDIC]. (2023). *Re: Quality control standards for automated valuation models*. <https://www.fdic.gov/resources/regulations/federal-register-publications/2023/2023-quality-control-standards-for-automated-valuation-models-3064-ae68-c-003.pdf>
21. Federal Financial Institutions Examination Council. (2010). *Interagency appraisal and evaluation guidelines*. <https://www.fdic.gov/news/news/financial/2010/fil10082a.pdf>
22. Federal Register. (2023). *Quality control standards for automated valuation models*. <https://www.federalregister.gov/documents/2023/06/21/2023-12187/quality-control-standards-for-automated-valuation-models>
23. Federal Reserve System. (2011). *Supervisory guidance on model risk management*. <https://www.federalreserve.gov/supervisionreg/srletters/sr1107a1.pdf>
24. Hu, L., He, S., Han, Z., Xiao, H., Su, S., Weng, M., & Cai, Z. (2019). Monitoring housing rental prices based on social media: An integrated approach of machine-learning algorithms and hedonic modeling to inform equitable housing policies. *Land Use Policy*, 82, 657-673.
25. International Association of Assessing Officers [IAAO]. (2018). Standard on automated valuation models (AVMs). *Journal of Property Tax Assessment & Administration*, 15(2), 67-101.

26. Kok, N., Koponen, E. L., & Adriana Martínez-barbosa, C. (2017). Big data in real estate? From manual appraisal to automated valuation. *The Journal of Portfolio Management*, 43(6), 202–211.
27. Lee, H., Han, H., Pettit, C., Gao, Q., & Shi, V. (2024). Machine learning approach to residential valuation: A convolutional neural network model for geographic variation. *The Annals of Regional Science*, 72, 579–599
28. Makridakis, S., Hibon, M., & Moser, C. (1979). Accuracy of forecasting: An empirical investigation. *Journal of the Royal Statistical Society. Series A*, 142(2), 97–145
29. Matysiak, G. (2017). *Automated valuation model (AVMs): A brave new world?* Wroclaw Conference in Finance 2017, Wrocław, Poland.
30. MSCI. (2023). *Private real estate valuation and sale price comparison 2022*. <https://www.msci.com/documents/10199/0f838bb6-e5dc-5a30-f94c-df71f0b452ac>
31. Shi, L., & Zhang, Y. (2015). Appraisal inflation: Evidence from the 2009 GSE HVCC intervention. *Journal of Housing Economics*, 27, 71–90.
32. Tan, F., Cheng, C., & Wei, Z. (2019). Modeling and elucidation of housing price. *Data Mining and Knowledge Discovery*, 33, 636–662.
33. U.S. Government Publishing Office. (2010). *Dodd-Frank Wall Street reform and consumer protection act*. <https://www.congress.gov/111/plaws/publ203/PLAW-111publ203.pdf>
34. Williamson, J., & Palim, M. (2022). *Appraising the appraisal*. <https://www.fanniemae.com/media/document/pdf/appraising-appraisal.pdf>
35. Yrjö, O. V. (1976). *Relative changes and index numbers*. <https://www.eta.fi/wp-content/uploads/A4.pdf>

논문 접수 일: 2024년 8월 27일

심사(수정)일: 2024년 10월 30일

게재 확정 일: 2024년 11월 15일

국문초록

본 연구는 점차 활용범위를 넓혀가고 있는 자동가치산정기법(automated valuation model, AVM)의 정확성 및 신뢰도 개선 방법을 연구하였다. 일찍이 AVM이 소개된 미국 시장은 Dodd-Frank Act 등을 통하여 AVM의 공정성, 정확성, 신뢰도에 대한 요구사항을 개선해 온 반면, 국내 AVM 시장은 대부분 내부 통제 수준에 머무르고 있다. 대량 정보확보가 가능한 국내 AVM 기업들의 추정치를 median absolute value percentage error, percentage predicted error@10% 기준을 사용하여 비교했을 때, 지역별로 상이한 정확도를 보였으며, 아파트 및 오피스텔의 하위시장에 대해서도 정확도가 다른 수준으로 나타났다. 이는 같은 AVM 기업이라도 지역별로 정확도가 다를 수 있으며, 지역별 정확도 기준 제시가 필요하다는 시사점을 보인다. 아울러, AVM의 정확성이 확보되는 경우, 감정가-실거래가 격차 확대에 따른 부실대출 및 전세 사기 예방을 위한 수단으로 활용할 수 있고, 부동산 담보채권의 2차 시장의 담보가치 모니터링 및 투자 포트폴리오 검사에까지도 활용범위를 넓힐 수 있을 것으로 전망한다. 이를 위해서는 AVM의 정확성 기준 설정 및 독립적인 제3자 비영리 검증기관의 도입이 필요하다.

주제어 : 자동가치산정모형, 가치평가, 담보가치 과대평가, 도드-프랭크법, AVM(automated valuation model) 정확성 검증 방법