



전세보증금 반환보증 가입의 동태적 결정요인 및 구조적 변화 분석*

- 계량경제모형과 머신러닝 모델의 비교를 중심으로 -

Dynamic Determinants and Structural Changes of Jeonse Deposit Return Guarantee Subscriptions - A Comparative Analysis of Econometric and Machine Learning Models -

이항재**

Hang-Jae Lee

Abstract

As the risk of unreturned Jeonse deposits has materialized due to recent interest rate hikes and the spread of Jeonse fraud, subscriptions to the 'Jeonse Deposit Return Guarantee (hereafter referred to as the "Jeonse Guarantee")'—a key housing safety net for tenants—have surged explosively. This study empirically analyzes the determinants and structural changes in the Jeonse Guarantee using time-series data from January 2018 to December 2024, utilizing the traditional vector error correction model (VECM), non-linear eXtreme gradient boosting (XGBoost) model, the Prophet model for structural break detection, and long short-term memory (LSTM) model specialized for time-series forecasting. The main findings are as follows. First, the VECM analysis identified the Jeonse-to-sales price ratio and Real Estate Consumer Sentiment Index as key drivers of Jeonse Guarantee demand, while the error correction term (ECT) confirmed the market's resilience in restoring long-term equilibrium. Second, XGBoost and SHapley Additive exPlanations (SHAP) analysis highlighted a non-linear surge in the predictive contribution to subscription demand when the Jeonse-to-sales price ratio and price income ratio (PIR) breached specific thresholds. Third, the Prophet model verified the 'Villa King Incident (Nov 2022)' as the most potent trigger for the Jeonse Guarantee surge and the '126% rule (May 2023)' as a structural breaking point causing a downward trend, thereby visualizing the impact of policy and external shocks. Fourth, the deep learning-based LSTM model demonstrated superior predictive performance with a mean

* 본 연구는 연구자 개인의 의견이며, 소속기관인 한국주택금융공사의 공식견해와는 무관함.

** 한국주택금융공사 주택금융연구원 대리 | Assistant Manager (Research Administration), Housing Finance Research Institute, Korea Housing Finance Corporation | 2034@hf.go.kr |

absolute percentage error (MAPE) of 5.31%, significantly outperforming the linear VECM. This study statistically proves that the Jeonse Guarantee demand exhibits dynamic, nonlinear changes driven by policy interventions and market factors, which are difficult to capture using traditional linear models. In conclusion, to ensure the sustainability of the Jeonse Guarantee system, this study provides foundational academic data and policy directions for discussing the potential introduction of a realistic adjustment of the premium rate system (utilizing XGBoost), an AI, and deep learning-based early warning system (utilizing LSTM).

Keywords: Jeonse deposit return guarantee, Vector error correction model, eXtreme gradient boosting, Long short-term memory, 126% rule

1. 서론

1. 연구의 배경과 목적

한국 주택시장의 독특한 임대차 제도인 전세는 주택 구매 자금 조달과 임차인의 주거비 절감 기능을 수행하며 발전해 왔다. 그러나 2020년 이후 부동산 확대와 급격한 금리 인상 등 거시경제 및 정책 환경의 변화는 전세 제도의 본질적 리스크인 '보증금 미반환 위험'을 수면 위로 부상시켰다. 특히 2022년 하반기 '빌라왕 사태'로 대변되는 대규모 전세 사기 및 강통전세의 확산은 임차인들의 극심한 불안감을 조성하며, 전세보증금 반환보증(이하 반환보증) 가입의 폭발적인 증가로 이어졌다. 보증기관의 대위변제 금액이 급증함에 따라, 전세 시장의 안정과 보증기관의 재정 건전성 관리를 위한 정교한 반환보증 가입 수요 분석 모델의 필요성이 대두되었다.

기존의 부동산 시장 연구는 주로 매매가격이나 전세가격 자체의 등락 요인을 규명하는 데 집중되어 왔다. 그러나 반환보증 가입이라는 임차인의 위험 회피적 행태가 거시경제 및 주택시장 내생 변수에 어떻게 반응하는지에 대한 시계열적 연구

는 상대적으로 부족하였다. 특히 기존 계량경제학적 접근법(벡터오차수정 모형[vector error correction model, VECM] 등)은 변수 간의 선형적(linear) 관계와 장기 균형을 설명하는 데 탁월하나, 강력한 사회·정책적 충격(전세 사기 이슈, 126% 룰 등)이 발생하거나 시장 심리가 비선형적으로 급변하는 구간에서는 예측 성능의 한계를 드러낼 수 있다. 따라서 전통적인 선형 모형의 한계를 보완하고, 데이터에 내재된 복잡한 패턴을 보다 유연하게 학습할 수 있는 머신러닝(machine learning, ML) 및 딥러닝(deep learning, DL) 기반의 분석 방법 도입을 검토할 필요가 있었다.

이에 본 연구는 기존 계량경제 모형과 ML 및 DL 모델의 예측 성능을 비교·분석하고, 최적의 모델을 통해 반환보증 가입의 동태적 결정요인을 규명하고자 한다. 특히, 전통적인 선형 모형(VECM)으로 포착하기 어려웠던 반환보증 수요의 비선형적 증감 현상과 정책 충격에 따른 구조적 단절(structural break)을 실증함으로써, 지속 가능한 반환보증 제도를 위한 학술적 기초자료와 정책적 방향성을 제시할 것이다.

2. 연구의 범위와 방법

반환보증 가입은 주택시장의 불확실성과 임차인의 보증금 미반환 위험 인식에 따라 결정된다. 통상적으로 전세가율이 상승하거나 전세대출 금리¹⁾가 변동할 경우, 임차인의 주거비 부담과 리스크 회피 성향이 변화하여 보증 가입 수요에 직접적인 영향을 미치게 된다. 따라서 반환보증 가입 실적(건수)²⁾과 거시경제 지표 및 주택시장 내생 변수 간의 동태적 관계를 분석하는 것은 반환보증 가입 수요를 분석하는 타당하며 합리적인 방법이라 할 수 있다.

특히 전세가율과 부동산 소비심리지수, 소득대비 부동산 가격 비율(price income ratio, PIR) 등의 변수는 주택 전세시장의 보증금 미반환 위험도를 선행 또는 동행하여 반영하는 지표이기에 반환보증 수요에 대해 유의미한 영향성이 존재할 것으로 예상된다. 선정된 변수들이 반환보증 가입에 영향을 미치는 시차는 변수의 성격에 따라 달리 나타날 것이며, 그 영향력 또한 시장 국면에 따라 비선형적으로 작용할 것이기에 다양한 모델링을 통한 실증분석이 요구된다.

단순한 경제 지표만으로 분석할 경우 급변하는 정책 환경과 사회적 충격을 반영하지 못해 시장 상황을 곡해할 가능성이 존재한다. 이에 본 연구는 2018년 1월부터 2024년 12월까지의 주택도시보증공사(Korea Housing & Urban Guarantee Corporation, HUG) 반환보증 가입 건수를 종

속변수로 설정하고, 전세가율·부동산 소비심리지수·COFIX 금리, 실업률, PIR을 독립변수로 하는 월별 시계열 데이터를 구축하였다. 나아가 분석의 정교함을 제고하기 위해 2020년 6·17 및 7·10 부동산 대책(갭투자 방지)과 2020년 8월 임대차 3법 시행, 2022년 10월 ‘빌라왕’ 사태로 인한 전세 사기 위기 가시화(손하늘, 2022), 그리고 2023년 5월 반환보증 126% 룰 시행 등 전세시장의 구조적 변화를 야기한 주요 정책 및 사회적 사건을 정책 더미 변수로 설정하여 모형에 반영하였다.

연구는 크게 세 단계로 진행된다. 첫째, 데이터의 안정성을 확인하기 위한 ADF 검정(argumented Dickey-Fuller test, 단위근 검정)과 요한센 공적분 검정(Johansen cointegration test)을 선행하고, 이를 바탕으로 전통적 계량 모형인 VECM을 구축한다. 둘째, ML(XGBoost[extreme gradient boosting]) 및 DL(LSTM[long short-term memory]), Prophet 모델을 학습시키고 하이퍼파라미터 튜닝을 통해 최적의 예측 성능을 도출한다. 셋째, 모델 간 예측 오차(RMSE[root mean squared error], MAPE[mean absolute percentage error])를 비교하여 최적 모델을 선정하고, SHAP(shapley additive explanations) 분석을 XGBoost 모델에 적용하여 변수 중요도와 의존성을 시각화함으로써 ML의 분석결과에 대한 설명을 더함으로써 반환보증 수요 결정요인에 대한 실증적 함의를 도출한다.

1) 대다수의 금융권에서 COFIX(cost of funds index) 금리를 전세대출의 기준금리로 채택하여 운용 중이다.

2) 본 연구는 반환보증 가입 금액이 아닌 가입 건수를 분석 대상으로 한정하였다. 가입 금액의 경우 주택 가격 변동에 따라 수치가 왜곡될 가능성이 크기 때문이다. 따라서 주택 가격의 등락 효과를 배제하고, 리스크를 인지한 임차인들의 가입 의사결정 빈도와 그 확산 정도를 파악하기 위해 가입 건수를 종속변수로 채택하였다.

3. 이론적 배경

1) 전세 제도의 구조와 보증금 미반환 리스크

한국 주택시장의 고유한 임대차 방식인 전세 제도는 임대인에게는 무이자 레버리지를 활용한 부동산 투자의 기회를, 임차인에게는 주거 비용 절감 및 자산 축적을 위한 주거 사다리의 기능을 제공하며 유지되어 왔다. 이러한 전세 제도는 임대인과 임차인 개인 간의 신용을 바탕으로 전세보증금이라는 채권 채무 관계가 형성된다는 점에서 사금융적 성격을 띤다. 그러나 전세 보증금을 승계하여 주택을 매수하는 이른바 갭투자 방식은 주택 가격 상승기에는 임대인 입장에서 부동산 투자 수익을 극대화할 수 있으나, 주택 가격 하락기에는 주택 처분 가격이 전세보증금보다 낮아지는 깡통전세 위험을 구조적으로 내포하고 있다. 특히 임대인의 유동성 위기는 곧바로 임차인 입장에서 보증금 미반환 위험으로 전이되며, 이는 개별 가계의 리스크를 넘어 사회적 시스템 리스크로 확산될 가능성을 가진다.

2) 반환보증의 기능과 시장 메커니즘

반환보증은 임대인이 임차인에게 보증금을 반환하지 못할 경우, 보증기관(HUG, HF[Korea Housing Finance Corporation], SGI[Seoul Guarantee Insurance] 등)이 이를 대신 변제하고 추후 임대인에게 구상권을 청구하는 공적 보험의 성격을 가진 상품이다. 이는 임차인의 재산을 보호하는 안전장치인 동시에, 개인 간 거래에 내재되어 있는 전세 제도의 신용 위험을 제도권 금융 시스템으로 흡수하는 역할을 한다. 선행 연

구에 따르면, 반환보증 수요는 시장의 위험 인식과 밀접한 관계를 가진다. 전세가격이 상승하여 깡통전세의 위험이 가시화될수록 임차인의 반환보증 가입 유인은 증가한다. 또한, 시장 금리(COFIX 금리 등)는 임차인의 전세자금 대출 비용을 결정하는 핵심 변수이다. 시장금리 상승은 임차인 입장에서 전세자금 대출 비용 증가 요인으로 전세 수요를 위축시켜 반환보증 가입 감소 유인이 되며, 전세 수요 위축은 전세 가격 하락을 유발하는 이중적인 특성을 보이기도 한다. 이론상 전세가격이 하락하면 깡통전세의 위험이 낮아져 반환보증의 가입 유인이 감소할 수도 있지만, 동시에 시장 금리 인상에 따른 전세가격 하락은 임대인의 유동성 부족 및 역전세 위험을 가중시킨다. 이는 임차인들에게 전세보증금 미반환 공포를 자극하여, 반환보증 가입의 필요성을 높이는 비선형적이고 복합적인 메커니즘으로 작동한다. 즉, 특정 거시경제 변수는 반환보증 수요에 일방향으로 작용한다기보다는 여러 변수와 시장 심리, 정책 영향들이 복합적으로 작용하는 비선형적 메커니즘을 내재하고 있다.

4. 선행연구 검토

1) 전세 시장 리스크 및 반환보증 수요에 관한 기존 계량 연구

전세 시장의 메커니즘과 리스크를 규명하기 위한 기존의 계량 연구는 주로 거시경제 변수와 주택·전세가격 간의 관계 분석에서 출발하여, 최근에는 전세보증사고 및 정책 효과 분석으로 확장되어 왔다. 초기 연구인 이영수(2010)는 VECM을

이용하여 주택가격과 전세가격 간에 장기적 균형 관계가 성립함을 밝혔으며, 전세가격의 변화가 주택가격에 영향을 미친다는 점을 실증하였다.

최근 전세보증금 미반환 문제가 심각한 사회 문제로 대두되면서, 실제 보증사고 데이터를 활용한 실증 연구 또한 활발히 진행되었다. 김진유(2022)는 주택도시보증공사의 대위변제 자료를 바탕으로 회귀분석을 실시하여, 전세가율이 높은 고위험 전세가 전세사고로 이어질 가능성이 높음을 규명하고 연립·다세대주택 밀집 지역의 취약성을 지적하였다. 안성연·이상엽(2025)은 이항 로짓모형을 활용하여 전세보증금 미반환 사고에 영향을 미치는 요인을 분석하였으며, 임대인의 다주택 여부, 부채비율, 특정 주택 유형(오피스텔, 연립다세대) 등이 사고 발생 확률을 높이는 주요인임을 밝혔다. 또한, 거시적 관점에서 박진백 외(2023)는 가계금융복지조사 데이터를 활용해 주택 매매가격 하락 시 전세 레버리지(갭투자)로 인한 보증금 미반환 위험 가구 수를 추정하며 선제적 리스크 대응의 필요성을 강조하였다.

더 나아가 반환보증 제도와 전세 시장의 공간적·정책적 구조를 계량적으로 분석한 연구들도 등장하였다. 권혁·성현곤(2025)은 다수준 교차분류 순서형 로지스틱 모형을 통해 반환보증 가입 요건의 완화가 정보 비대칭성과 도덕적 해이를 유발하여 고위험 전세 발생 확률을 증가시켰음을 증명하였다. 이지연·이정란(2026)은 이중차분(difference in difference, DID) 로지스틱 회귀분석을 적용해 반환보증 가입 요건 강화 제도가 비아파트 및 중저가 주택 시장에서 전세의 월세화 현상을 구조적으로 가속화했음을 확인하였다. 김성민·우명제

(2026)는 공간패널모형을 이용하여 전세보증사고가 수도권 내에서 뚜렷한 공간적 군집(high-high)을 형성하며, 인접 지역으로의 공간 누출 효과를 지니고 있음을 실증하였다.

2) 부동산 시장의 머신러닝(Machine Learning) 적용 연구

부동산 시장 분석에 있어 기존의 선형회귀모형은 변수들 간의 관계가 비선형적일 경우 계수 추정치가 과대 또는 과소 추정되어 예측의 정확성을 저하시킬 수 있다는 한계가 지속적으로 지적되어 왔다. 이를 극복하기 위해 최근 부동산 연구에서는 ML 및 DL 기법을 도입하여 데이터의 비선형성을 반영하고 예측력을 높이려는 시도가 활발히 이루어지고 있다. 김승현 외(2022)는 서울시 아파트 실거래가 패널 데이터를 바탕으로 ML 기법인 랜덤 포레스트(random forest)와 다변량 적응 회귀 스플라인(multivariate adaptive regression splines, MARS)을 패널고정효과모형과 비교 분석하였으며, ML 모형이 기존 선형 모형보다 비선형 모델링을 통해 더욱 우수한 예측 정확성을 보임을 실증하였다.

3) 선행연구의 한계 및 본 연구의 차별성

기존의 연구 성과들에도 불구하고, 방법론적 측면에서 연구의 확장 가능성은 여전히 존재한다. 다수의 기존 계량 분석 연구들은 전세가율, 금리, 정책 변수 등이 전세 사고 및 임차인의 주거 불안에 미치는 평균적인 영향력을 규명하여 주택금융 연구의 기반을 마련하였다. 그러나 반환보증 등 주택금융 정책 수요 관련 연구에 대한 비선형

적 모형 분석이나, 특정 위기 국면에서 나타나는 전세 시장 참여자들의 심리적 쏠림 현상 등을 반영하는 분석모형에 대한 연구의 필요성은 여전히 존재한다. 또한, ML을 활용한 최신 부동산 연구들은 주로 주택 가격 및 지수 예측을 중심으로 진행되어 왔으며, 최근 전세 위험 추정에 ML이 일부 도입되기도 하였으나 이를 주택금융 정책 수요인 '반환보증 가입 수요'라는 행태적 영역으로 확장한 실증 연구는 미비하였다.

이에 본 연구는 ML(XGBoost), DL(LSTM), 시계열 구조모형(Prophet)을 '반환보증 가입 수요'라는 행태적 영역에 새롭게 적용하고자 한다. 이를 통해 기존 선행연구들이 포착한 거시 변수의 선형적 방향성을 넘어서, 전세가율과 PIR 등이 특정 위험 수위(임계치)를 돌파할 때 반환보증 가입 수요에 대한 예측 기여도가 급격히 증폭되는 시장 참여자들의 비선형적 쏠림 현상과 특정 정책 변화에 따른 구조적 단절을 실증적으로 확인하였다. 이는 기존 연구의 성과를 보완하고 주택 임대차 시장의 동태적 리스크 관리에 대한 새로운 시사점을 제공한다는 점에서 차별화된 학술적 의의를 지닌다.

II. 연구모형

본 연구는 반환보증 수요의 동태적 변화를 가장 정교하게 예측하기 위해, 전통적인 계량경제학적 접근과 최신 ML 및 DL 기법의 예측 성능을 체계적으로 비교·검증하는 연구방법을 채택한다. 즉, 단일 모형에 의존하는 대신 각기 다른 강점을 지닌 다양한 분석 모델들을 독립적으로 구축하여 분석 성능을 비교하고, 각 모델별 특성 및 시사점을 제공한다.

연구의 절차는 시계열 데이터의 안정성 확보를 위한 기초 통계 검정을 시작으로, 선형 관계 규명을 위한 VECM 구축, 비선형 패턴 학습을 위한 XGBoost 구축, 그리고 시계열 특성 및 정책 변화를 정교하게 포착하기 위한 Prophet 및 LSTM 모델링의 순서로 진행된다.³⁾

1. 기초 통계 검정 및 벤치마크 모형: 벡터오차 수정모형

통상적으로 거시경제 및 부동산 시계열 데이터는 시간의 흐름에 따라 평균과 분산이 변하는 비정상성을 띠는 경우가 많으므로, 먼저 분산 안정화와 변수의 탄력성 해석을 위해 자연로그 변환을 수행하였다. 그러나 로그 변환만으로는 불안정한 추세를 제거할 수 없어 허구적 회귀 오류가 발생

3) 4가지 분석 모형 중 전통적 계량 모형인 VECM(vector error correction model)과 최적화 알고리즘 기반의 Prophet 모형은 결정론적 산출물을 제공하는 반면, 머신러닝(machine learning, ML) 기반의 XGBoost(eXtreme gradient boosting)와 딥러닝(deep learning, DL) 기반의 LSTM(long short-term memory) 모형은 알고리즘 특성상 학습 과정에서 무작위성(randomness)이 개입된다. 이에 본 연구는 모형 분석 결과의 객관성을 더하기 위해 XGBoost와 LSTM 모형에서 별도의 기준을 마련하였다. XGBoost와 LSTM 모두 중심극한정리에 근거하여 모델 수행 시 통계적 유의성이 확보되는 30회 반복 수행을 거치도록 하였으며, 최종 성능 평가에는 이 다중 수행을 통해 도출된 평균값을 성능 지표로 채택하여 알고리즘의 무작위성에 따른 편향을 통제함으로써 분석의 강건성을 확보하였다.

할 수 있다. 이를 방지하고자 ADF 검정을 통해 시계열의 안정성을 진단하였다. 검정 결과 불안정 시계열은 1차 차분으로 통계적 안정성을 확보하였으며, 이후 개별 변수 간의 장기 균형 관계를 규명하기 위해 요한센 공적분 검정을 실시하여 VECM 모형의 적합성을 타진하였다.

VECM 모형은 종속변수와 5개의 독립변수를 포함해 총 6개의 다변수 모형으로 구축되었다. VECM은 변수 간의 높은 상관성을 유의미한 공적분 관계로 모형화하고, 단기 동태식에는 추세가 제거된 1차 차분 값을 사용하므로 다중공선성 (multicollinearity)⁴⁾ 문제가 상당 부분 완화된다.

이를 바탕으로 투입된 6개 변수 간에 장기적인 균형 관계가 존재한다는 가정하에, 다음과 같은 (식 1)을 설정하여 향후 ML 등 후속 모형과의 비교 분석을 위한 기준점(benchmark)으로 삼는다.

$$\Delta Y_t = \alpha\beta^T Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \epsilon_t \quad (\text{식 1})$$

여기서 Y_t 는 6개의 내생변수에 자연로그를 취한 (6×1) 벡터이고, Δ 는 차분 연산자(1-L)를 나타낸다. 우변의 첫 번째 항 $\alpha\beta^T Y_{t-1}$ 은 오차 수정 항으로서 변수 간의 장기 균형 관계를 나타내며, 두 번째 항 Γ_i 는 변수 간의 단기적인 동태적 관계를 설명하는 계수 행렬이다. ϵ_t 는 백색 잡음 과정을 따르는 오차항이다. VECM은 변수 간의 선형적 인과관계를 명확히 보여주는 장점이 있으나,

급변하는 시장의 비선형성을 포착하는 데 한계가 있으므로 이를 ML 모형 등과 비교하는 준거 모형으로 삼는다. 단, 후속 모형과의 공정한 성능 비교를 위해 예측 성능 평가 시에는 로그 변환된 예측 값을 지수 변환(역변환)하여 원자료 스케일인 반환보증 가입 건수로 환원한 후, RMSE 및 MAPE 지표를 산출하여 객관성을 확보하였다.

2. 머신러닝 분석 방법론: eXtreme Gradient Boosting

본 연구는 거시경제 변수와 전세 시장 내생 변수 간의 복잡한 비선형적 상호작용을 정교하게 학습하기 위해, 부스팅(boosting) 알고리즘을 기반으로 한 XGBoost를 ML 모형으로 채택하였다.

1) 모형의 선정 이유 및 장점

다양한 ML 모형 중 XGBoost를 최종 선정한 이유는 다음과 같다.

첫째, 비선형성 학습 능력이다. 반환보증 수요는 특정 시계열 시점에서 급격히 변화하는 비선형적 특성을 보인다. XGBoost는 경사 하강법을 통해 이전 트리가 범한 예측 오차를 순차적으로 줄여나가는 직렬적 학습 구조를 취하고 있어, 선형 모형이 포착하기 어려운 데이터의 복잡한 패턴과 임계치를 효과적으로 학습한다.

둘째, 과적합 제어 능력이다. 시계열 데이터 분

4) 본 연구에 투입된 독립변수 간의 다중공선성 정도를 실증적으로 재확인하기 위해, 단기 동태식에 투입되는 1차 차분(differenced) 변수들과 원자료 데이터를 대상으로 분산팽창지수(variance inflation factor, VIF) 검정을 추가 실시하였다. 검정 결과, 1차 차분한 5개 독립변수의 평균 VIF는 1.15로 나타났으며, 최대 VIF 역시 '부동산 소비심리지수' 변수의 1.34로 산출되었다. 원자료 데이터 5개 독립변수 또한 평균 VIF는 2.35, 최대 VIF는 'COFIX 금리' 3.36 수준으로 도출되어 통상적인 다중공선성 판단 기준인 10(또는 5)을 하회함을 확인하였다.

석 시 발생할 수 있는 과적합 문제를 방지하기 위해, XGBoost는 모형의 복잡도에 페널티를 부여하는 L1(lasso) 및 L2(ridge) 규제 항을 목적함수에 포함하고 있다. 이는 학습 데이터에 대한 편향을 줄이는 동시에 분산을 억제하여, 미래 시점인 테스트 데이터에 대해서도 강건한 예측 성능을 확보하게 한다.

2) 한계점 및 Shapley Additive Explanations 기법을 통한 보완

일반적으로 XGBoost와 같은 앙상블 모형은 높은 예측 정확도를 보장하는 대신, 개별 변수가 결과에 미치는 영향을 직관적으로 이해하기 어려운 블랙박스(black-box)의 한계를 가진다. 즉, 모형이 “반환보증 가입이 증가할 것”이라고 예측할 수는 있어도, “구체적으로 어떤 변수가 얼마나 영향을 미쳤는지”를 설명하는 데에는 제약이 있다.

이에 본 연구는 XGBoost의 예측 결과를 단순히 제시하는 데 그치지 않고, 후속 분석으로 SHAP 기법을 도입하여 연구의 깊이를 더하였다. SHAP 분석을 통해 블랙박스 모형 내부를 시각화함으로써, 각 변수의 기여도와 영향의 방향성(양/음)을 정량적으로 규명하고 정책적 시사점을 도출하는 등 설명 가능한 인공지능(explainable artificial intelligence, XAI) 분석을 수행하였다.

3. 시계열 특화 및 딥러닝 모형: Prophet (Additive Model)과 Long Short-Term Memory

본 연구는 시계열 분석에 특화된 Prophet과

DL 모델인 LSTM을 활용하여, 앞서 제시한 벤치마크 모형인 계량경제학적 모형(VECM) 및 ML(XGBoost) 접근법의 한계를 보완하고자 한다. 즉, 단순한 과거 데이터 패턴의 학습을 넘어, 정책 변화에 따른 추세의 구조적 전환과 시계열 데이터 고유의 장기 의존성을 정교하게 포착하기 위해 각 특성에 최적화된 방법론을 선별적으로 적용하였다.

1) Prophet(Additive Model)

Prophet은 복잡한 시계열 데이터를 트렌드, 계절성, 그리고 휴일 효과의 합으로 분해하여 해석하는 가법 모형이다(Taylor and Letham, 2018). 본 연구가 채택한 기본 수식은 다음 (식 2)와 같다.

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \quad (\text{식 2})$$

여기서 $g(t)$ 는 비주기적인 장기 추세 변화를 나타내며, $s(t)$ 는 월별 시계열 데이터 내에 존재하는 주기적 계절성을 의미한다. 또한 $h(t)$ 는 터미 변수로 설정한 주요 부동산 정책 변동과 같은 데이터 외적인 구조적 이벤트 효과를 의미한다. 특히 본 연구는 Prophet의 변화점 식별 기능을 핵심 분석 도구로 활용한다. 구체적인 예시로, 2023년 5월 시행된 ‘반환보증 가입 요건 강화(126% 룰)’ 등 주요 정책 시행 시점을 사전에 모형 내에서 구조적 변화점으로 지정함으로써, 해당 정책 충격이 반환보증 가입 수요 추세의 기울기를 통계적으로 유의미하게 변화시켰는지를 시각적·정량적으로 규명한다.

2) Long Short-Term Memory

본 연구의 LSTM 네트워크 설계는 전체 시계열 데이터 표본이 84개월로 제한적인 상황임을 고려하여, 모델의 복잡도와 학습 데이터 확보 간의 최적 균형을 찾는 데 주안점을 두었다. 구체적인 아키텍처와 그 설정 근거는 다음과 같다.

첫째, 데이터 전처리 단계에서 min-max 정규화를 적용하였다. LSTM은 입력 데이터의 스케일 차이에 민감하므로, 학습 효율성을 높이고 안정적인 수렴을 유도하기 위해 모든 입력 변수를 0과 1 사이의 값으로 변환하였다.

둘째, 입력 시퀀스의 길이(window size)는 VECM의 최적 시차(lag=2)와 동일하게 2개월로 설정하여, 두 모형의 통제 조건을 일치시켰다.

셋째, 은닉층은 전체 시계열 데이터 표본이 84개월로 제한적인 상황임을 고려하여, 모형의 과적합을 방지하고자 32개의 유닛을 가진 단일 LSTM 층(1-layer)으로 간결화하여 소규모 데이터셋에 최적화된 학습 용량을 확보하였다.

넷째, 과적합을 추가로 제어하기 위해 dropout 비율은 0.2로 설정하여 정보 손실을 막으면서도 일반화 성능을 높였다.

다섯째, 모델 최적화를 위해 Adam optimizer와 MSE(mean squared error) 손실 함수를 채택하였으며, 소규모 표본에서 조기 종료 기법 적용 시 발생할 수 있는 학습 부족(underfitting) 현상을 방지하기 위해 충분한 에포크(epoch=300)를 고정적으로 할당하여 안정적인 수렴을 유도하였다.

마지막으로, 역변환 과정을 거쳐 복원된 예측값으로 평균 RMSE 및 MAPE를 산출하여 분석의

객관성과 강건성을 극대화하고, VECM 등 타 모형과의 객관적인 성능 비교를 가능하게 하였다.

III. 연구모형 실증분석 결과

1. 데이터의 출처 및 기초 분석

1) 변수 선정 및 데이터 구축

반환보증 가입의 동태적 결정요인을 분석하기 위해 <표 1>과 같은 데이터를 분석에 활용하였다.

2) 주요 정책 변수의 설정 및 배경

본 연구의 차별점은 2020년 이후 급격하게 전개된 부동산 정책 변화와 시장 충격 사건을 더미 변수로 모형화하여 분석에 반영했다는 점이다. 시계열 데이터의 구조적 변화를 야기할 수 있는 다음과 같은 주요 사건을 더미변수로 설정하여 분석에 포함하였다.

(1) 2020년 부동산 대책(6·17 및 7·10)

갭투자 억제와 보유세 부담 강화로 인해 임대인의 현금 선호(월세 및 반전세 전환)를 촉진하며, 전세 물량 감소와 전월세 시장의 구조적 변화를 야기하였다.

(2) 임대차 3법 시행(2020년 8월)

전세 매물 잠김과 이중 가격 현상을 유발해 가격 변동성을 급격히 확대시켰으며, 임차인의 보증금 미반환 리스크 인식을 고조시킨 구조적 전환 점이 되었다.

〈표 1〉 데이터 및 출처

데이터명 (변수)	설명	출처
반환보증 가입 건수 (종속변수)	전세보증금 반환보증 신규 가입 건수. 임차인이 전세보증금 미반환 리스크를 어떻게 인지하고 반응하는지를 나타내는 직접적인 지표로 활용됨.	주택도시보증공사(HUG)
전세가율 (독립변수)	주택종합 매매가격 대비 전세가격 비율(평균값 사용, 단위 %).	한국부동산원 전국주택가격 동향조사
부동산 소비심리지수 (독립변수)	전국 종합, 부동산 시장 참여자들의 심리를 지수화(0~200)한 값. 100을 초과하면 가격 상승 및 거래 증가 기대, 100 미만이면 하락 기대를 의미함.	국토연구원
PIR (독립변수)	(소득대비 부동산 가격 비율) 주택 구매 능력 지표. PIR이 상승하면 부동산 매수의 진입 장벽이 높아졌다고 판단됨.	한국주택금융공사 주택금융 통계시스템
COFIX 금리 (독립변수)	(신규취급액 기준) 전세자금 대출의 기준이 되는 기회비용 및 금융비용.	은행연합회
실업률 (독립변수)	경제활동인구 대비 실업자 비율. 실업률 상승은 가계 소득 감소와 임대인의 유동성 위기를 가져오므로, 거시적 충격 요인으로 활용함.	국가데이터처, 경제활동인구조사

주 : 1) 데이터 기간은 2018. 01.~2024. 12.이며, 월별 자료로 가공하였다.
 2) 종속변수인 반환보증 가입 건수는 본 연구의 목적상, 경제 지표와 정책 변화가 임차인의 반환보증 가입 수요 총량에 미치는 영향을 거시적으로 분석하고자 한 데 있으므로 별도로 주택유형을 구분하지 않았다.
 3) HUG, Korea Housing & Urban Guarantee Corporation; PIR, price income ratio; COFIX, cost of funds index.

(3) 빌라왕 사태(2022년 10월)

대규모 전세사기로 무자본 갭투자의 위험이 현실화되면서 전세 제도에 대한 신뢰를 붕괴시켰으며, 반환보증 가입을 선택에서 필수로 전환시킨 핵심 트리거로 작용했다.

(4) 126% 룰 시행(2023년 5월)

보증기관의 리스크 관리를 위한 가입 요건 강화 조치였으나, 역설적으로 비아파트 시장의 보증 가입 거절 사례를 급증시키며 가입 건수의 구조적 하락을 유발한 구조적 단절점이 되었다.

반환보증 가입 건수(종속변수)는 2018년 1월 4,461건에서 2023년 3월 31,158건으로 약 7배 폭증하며 매우 큰 변동성을 보였다. 이는 전세 사기 및 깡통전세 위험 고조에 따른 임차인들의 강력한 위험 회피 성향을 대변하며, 2023년 하반기 이후에는 ‘126% 룰’ 도입과 전세의 월세화 현상 등에 의해 다소 조정 국면에 진입했다.

전세가율은 평균 64.89, 표준편차 1.47로 표면적인 수치 변동 폭은 작아 보인다. 그러나, 전세 가율은 단 1~2%p의 미세한 변화만으로도 전세 시장 전반에 막대한 충격을 가하는 핵심 변수이다. 이는 부동산 거래의 자금 규모와 갭투자 특유의 레버리지 구조가 맞물려 있기 때문이며, 특히 전세가율이 특정 임계점을 돌파할 경우 깡통전세에 대한 시장의 공포 심리를 자극하여 연쇄적인 보증 사고 위험과 반환보증 가입 수요 급증을 야

2. 기초통계량 분석 및 시계열 추이 해석

2018년 1월부터 2024년 12월까지의 데이터에 대한 기초통계 분석 결과는 〈표 2〉와 같다.

〈표 2〉 주요 변수의 기초통계량 분석 결과

변수명	단위	평균(Mean)	표준편차(Std)	최솟값(Min)	최댓값(Max)
반환보증 가입 건수	건	17,809	6,329	4,461	31,158
전세기율	%	64.89	1.47	63.00	68.60
COFIX 금리	%	2.22	1.10	0.80	4.34
부동산 소비심리지수	index	105.44	12.18	78.00	131.70
PIR	배	6.48	0.64	4.84	7.43
실업률	%	3.38	0.78	1.90	5.70

주 : COFIX, cost of funds index; PIR, price income ratio.

기할 수 있다.

COFIX 금리는 2020년 0.8%대에서 2023년 4%대 이상으로 급등하였으며, 단기간 내의 급격한 변동성은 임대인과 임차인 양측의 자금 조달 리스크를 증폭시키며, 전세 시장 전반의 구조적 불안을 야기하는 요인으로 작용할 수 있다.

부동산 소비심리지수는 호황기(130 상회)를 지나 2022년 하반기 ‘빌라왕 사태’ 및 금리 인상 충격으로 100 이하로 급락했으며, 시장 심리에 따라 반환보증 가입 수요가 간접적으로 영향을 받을 수 있다.

PIR은 평균 6.48에서 최대 7.43까지 치솟았으며, 높아진 주택 매매 부담이 전세 수요의 변화를 야기했을 것으로 추측할 수 있다. 예를 들면, PIR 상승(높아진 주택 매매 부담)은 주택 매매 수요가 전세 수요로 전환되도록 하는 요인이 될 수 있으며(김경환 외, 2025), 전세 거래의 규모가 확대된다면 전세의 파생수요인 반환보증의 수요 증가로 이어질 수 있다.

실업률은 분석 기간 내 평균 3.38%에서 최대 5.70%를 기록하였다. 실업률 충격은 가계의 소

득 감소와 신용 경색을 유발하여 임차인의 전세 자금 조달 능력을 약화시키고 임대인의 보증금 상환 능력을 약화시킬 수 있다. 이는 곧 전세 수요의 위축과 전세가격 하락으로 이어져 임대인의 보증금 미반환 리스크를 증폭시키는 시장 불안을 야기하는 경로로 작용할 수 있다.

1) Augmented Dickey-Fuller Test 결과

다변량 시계열 모형인 VECM을 구축하기에 앞서, 분석에 사용된 변수들의 통계적 안정성을 확인하고 허구적 회귀 문제를 방지하기 위해 ADF 검정을 실시하였다. 〈표 3〉의 분석 결과에 따르면, 원자료 상태에서 ‘실업률’을 제외한 5개 변수의 검정 통계량 p-value가 5% 유의수준(0.05)을 초과하여 ‘시계열 자료에 단위근이 존재한다’는 귀무가설을 기각하지 못하였다. 반면, ‘실업률’ 변수는 원자료 상태에서도 p-value가 0.0111로 도출되어 귀무가설을 기각하고 이미 통계적으로 안정적인 영차 적분(I(0)) 시계열임이 확인되었다. 원자료 상태에서 단위근이 존재했던 5개 변수들은 1차 차분 후 모두 p-value가 유의수준 미만

〈표 3〉 주요 변수의 ADF 검정 결과

변수(로그 변환)	원자료(Level)		1차 차분(1st diff.)		판정 (5% 유의수준)
	t-Statistic	p-Value	t-Statistic	p-Value	
반환보증 가입 건수	-1.5423	0.8142	-6.3818	0.0000	I(1)
전세가율	-0.9702	0.9480	-3.7379	0.0036	I(1)
COFIX 금리	-1.7051	0.7486	-3.2173	0.0190	I(1)
부동산 소비 심리지수	-2.0706	0.5625	-4.8898	0.0000	I(1)
PIR	-2.8789	0.1695	-5.0884	0.0000	I(1)
실업률	-3.9293	0.0111	-4.5806	0.0001	I(0), I(1)

주 : 1) 본 ADF 단위근 검정은 84개월이라는 시계열 표본의 특성과 이후 구축될 다변량 모형(VECM) 간의 방법론적 통일성을 확보하기 위해, 검정의 최대 시차를 2개월(lag=2)로 통제하여 수행하였다.
 2) 각 시계열 데이터가 지닌 본연의 추세를 왜곡 없이 반영하기 위해 원자료(level) 변수 검정에는 상수항과 추세항을 모두 포함하였으며, 1차 차분을 통해 수학적으로 추세가 제거된 변수에 대해서는 상수항만을 포함하여 검정력의 누수를 방지하였다.
 3) 귀무가설(H_0)은 '시계열 자료에 단위근이 존재한다(즉, 데이터가 비정상적이고 불안정하다)'이다. 산출된 검정 통계량의 p-value가 유의수준 5%(0.05) 미만일 경우 귀무가설을 기각하고 해당 시계열이 안정적(stationary)이라고 판정한다.
 4) 판정량의 'I(0)'(영차 적분)는 차분을 하지 않은 원자료 상태에서 이미 통계적으로 안정적인 시계열임을 뜻하며, 'I(1)'(일차 적분)은 원자료 상태에서는 불안정하나 1차 차분을 수행한 후 안정화되는 시계열을 의미한다.
 5) ADF, augmented Dickey-Fuller; Diff., difference; COFIX, cost of funds index; PIR, price income ratio.

으로 도출되어 안정성을 확보하였다.

결과적으로, 본 연구에 투입된 모든 변수들이 1차 차분 후 안정적인 시계열 특성(I(1))을 공유하고 있음이 통계적으로 확인되었다. 이는 비정상 시계열 변수들 간의 장기적 선형 결합을 분석하는 요한센 공적분 검정 및 VECM 모형 구축을 위한 계량경제학적 전제조건을 충족하는 결과이다.⁵⁾

2) 요한센(Johansen) 공적분 검정 결과

검정 결과, 〈표 4〉와 같이 Trace 통계량 기준으로 '공적분 벡터가 없다($r=0$)' 및 '1개 이하이다($r \leq 1$)'라는 귀무가설이 모두 5% 유의수준에서 기각된 반면, ' $r \leq 2$ ' 가설은 기각되지 않아 변수

간에 2개의 장기 균형 관계가 존재함이 확인되었다. 이는 개별 불안정 시계열들이 결합하여 강력한 장기 균형 경로를 형성함을 시사한다.

통계적으로는 2개의 공적분 벡터(rank=2)가 도출되었으나, 다중 공적분 벡터를 모두 반영할 경우 모형이 과도하게 복잡해지고 경제학적 해석의 직관성을 잃을 우려가 있다. 이에 본 연구는 모형의 간명성을 확보하고자 가장 설명력이 높은 첫 번째 공적분 벡터(rank=1)만을 제한적으로 적용하였다. 추가적인 강건성 검정 결과, rank=2를 적용했을 때와 주요 변수들의 부호에는 유의미한 차이가 없음을 확인하였으며, 이를 통해 변수 간의 지배적인 장기 균형과 단기 조정 과정을 타당

5) I(0)와 I(1) 변수가 혼재될 경우 ARDL(autoregressive distributed lag) 모형이 권장될 수 있으나, 본 연구에서는 다수의 거시 변수 간의 동태적 내생성을 동시에 통제하고 시스템 전반의 장기 균형(공적분)을 직관적으로 도출하기 위해 요한센 공적분 검정(Johansen cointegration test) 통과를 전제로 VECM을 채택하였다.

〈표 4〉 요한센 공적분 검정 결과

귀무가설 (H ₀)	대립가설 (H ₁)	Trace 통계량	0.05 임계값	판정 (Decision)
r=0 (없음)	r ≥ 1	133.16	95.75	기각 (Reject)
r ≤ 1 (1개 이하)	r ≥ 2	77.81	69.82	기각 (Reject)
r ≤ 2 (2개 이하)	r ≥ 3	43.43	47.85	채택 (Accept)

주 : 1) 변수는 로그 변환된 시계열을 사용함. 이는 '반환보증 가입 건수(만 단위)'와 '금리(퍼센트 단위)' 간의 극심한 스케일 차이로 인한 이분산성 문제를 완화하고, 도출된 계수를 경제학적 탄력성 개념으로 일관성 있게 해석하기 위한 방법론적 조치임.
2) Trace 통계량이 임계값보다 크면 귀무가설을 기각함. 검정 결과 2개의 공적분 벡터가 존재하는 것으로 나타남.

〈표 5〉 적정 시차(lag length) 선정 기준표

시차(Lag)	AIC	SC(BIC)	HQ	판정
0	-27.52	-27.34	-27.45	-
1	-39.66	-38.42*	-39.16	SC 기준 최적
2	-40.19	-37.89	-39.27*	(HQ 기준 최적, 본 연구 채택)
3	-40.24*	-36.87	-38.89	AIC 기준 최적

주 : 1) *는 각 정보 기준에서 최적(최솟값)으로 선정된 시차를 의미함.
2) AIC, Akaike information criterion; SC, Schwartz criterion; BIC, Bayesian information criterion; HQ, Hannan-Quinn.

성 있게 설명하고자 한다.

3. 선형 동태적 분석: 공적분 및 Vector Error Correction Model 모형

1) 적정 시차(Lag Length)의 결정

VAR(vector autoregression) 계열 모형의 구축에 앞서, 분산을 안정화한(자연로그 변환) 내생변수들을 대상으로 정보기준을 통해 적정 시차를 검토하였다. 검토 결과, 〈표 5〉와 같이 각 정보기준이 지시하는 최적 시차가 상이하게 도출되었다. 정보기준만으로는 단일한 최적 시차를 특정하기 어려운 바, 본 연구는 통계적 지표 간의 균형(HQ[Hannan-Quinn] 기준 부합)과 경제적 현실성 등을 종합적으로 고려하여 시차 2(lag=2)를 최종 분석 시차로 설정하였다. 시차 2(lag=2)를 채택한 구체적인 이유는 다음과 같다.

첫째, 반환보증 가입 제도의 요건에 따른 구조

적 시차 발생이다. 통상적으로 임차인이 거시경제적 위험(전세가율 상승 등)을 인지하고 전세계약을 체결하더라도, 「전세임대주택 전세보증금반환보증 규정」(주택도시보증공사, 2020)에 따르면 실제 반환보증 가입은 임차인이 '대항력과 확정일자'를 모두 갖춘 이후에만 가능하다. 주택임대차보호법상 대항력은 주택의 인도(입주)와 전입신고를 마쳐야 발생하는데, 주택시장의 관행상 전세 계약 체결일과 실제 입주일 사이에는 통상 1~2개월의 시차가 존재한다. 따라서 임차인의 '위험 인지 및 의사결정(계약 시점)'과 본 연구의 종속변수인 '반환보증 가입(대항력 요건 충족 이후)' 사이에는 구조적으로 1~2개월의 시차가 발생한다.

둘째, 통계적 모형의 간명성 확보이다. 〈표 5〉에 나타난 바와 같이, 가장 긴 시차를 지지하는 시차 3(AIC[Akaike information criterion]: -40.24)과 본 연구가 채택한 시차 2(AIC: -40.19) 간의 정보기준 값 차이는 불과 0.05로 근소하다. 시차를 늘릴수록 추정해야 할 파라미터가 급증하여 소

표본 분석 환경에서 자유도 손실 및 과적합 우려가 커지게 된다. 응용 시계열 계량경제학의 원칙에 따르면, 정보기준은 시차 설정 시 유용한 지표이나, 이것이 데이터 생성 과정 및 제도적 현실에 대한 고려를 맹목적으로 대체해서는 안 된다(Enders, 2014). 따라서 모형의 간명성을 유지하며 현실적인 작동 원리를 가장 잘 반영할 수 있는 시차 2를 적용하는 것이 타당하다.

2) 벡터오차수정모형 추정 및 그랜저 인과관계(Granger Causality) 분석

변수 간의 장기적 균형 관계와 단기 동태적 조정 과정을 확인하기 위해 VECM을 추정하였다. 추정된 장기 균형 방정식은 (식 3)과 같다. 데이터의 분산 안정화를 위해 로그 변환된 자료를 사용하였으므로, 도출된 계수 값은 변수의 탄력성, 즉 퍼센트(%) 변화율에 따른 가입 건수의 증감률로 해석된다.

$$\begin{aligned} \ln(\text{joinCount})_t = & -17.858 \times \ln(\text{Jeonse})_t \\ & + 0.599 \times \ln(\text{COFIX})_t \\ & + 3.525 \times \ln(\text{Unemp})_t \\ & + 1.458 \times \ln(\text{PIR})_t \\ & - 0.692 \times \ln(\text{Sent})_t \\ & + 80.305(\text{Constant}) \end{aligned} \quad (\text{식 3})$$

주 : *joinCount* : 반환보증가입 건수
Jeonse : 전세가율
COFIX : 신규취급액 기준 *COFIX* 금리
Unemp : 실업률
PIR : 소득 대비 부동산 가격비율
Sent : 부동산 소비심리지수
Constant : 상수항

단기 동태식에서 도출된 오차수정항(error correction term, ECT)의 계수는 -0.0674로 추정되었다. 이는 전월의 실제 가입 건수가 장기 균형 수준을 이탈했을 경우, 매월 약 6.74%의 속도로 불균형을 해소하며 장기 균형으로 수렴해 나감을 의미한다. 장기 균형식의 계수들을 살펴보면, COFIX 금리(+0.599)와 실업률(+3.525) 등은 종속변수와 양(+)의 상관성을, 전세가율(-17.858)은 음(-)의 상관성을 띠는 것으로 나타났다.⁶⁾

다음으로, 각 독립 변수가 반환보증 가입 수요에 미치는 예측적 선행성을 통계적으로 엄밀하게 검증하기 위해 VECM 기반의 그랜저 인과관계 검정을 추가로 수행하였으며, 그 결과는 <표 6>과 같다.

검정 결과, '전세가율'이 본 연구의 적정 시차(lag=2)의 동태적 흐름 안에서 '반환보증 가입 건수'의 증감을 예측하는 데 있어 통계적으로 유의미한 선행성을 가지는 것으로 확인되었다. 이는 전세가율의 변동과 같은 전세시장의 핵심 리스크 요인이 시차를 두고 임차인들의 위험 회피 행동(반환보증 가입)을 촉발하는 선행 지표로 작동함을 시사한다.

반면, 부동산 소비심리지수(p=0.103)는 10% 유의수준에 근접한 한계적 유의성(marginal significance)을 보여 심리적 요인의 단기 선행 가능성을 시사하였으나, 그 외 변수들(COFIX 금리, PIR, 실업률)은 반환보증 가입 건수를 그랜저 인과하지 않는 것으로 나타났다. 해당 변수들은 VECM 모형 내에서 종속변수와 장기적인 균형

6) 이러한 VECM의 계수 추정치는 모형 내에 설정된 변수들 간의 장기적인 공적분 관계와 동조성을 보여주는 것일 뿐, 이것만으로 특정 독립 변수가 반환보증 가입 수요를 구조적·직접적으로 견인한다는 인과적 결론을 담보하지는 않는다.

〈표 6〉 VECM 기반 그래저 인과관계(Granger causality) 검정 결과

종속변수 (Y)	독립변수 (X)	카이제곱 통계량(χ^2)	p-Value	판정 (Decision)
반환보증 가입 건수 (증감분 $\Delta \ln Y$)	전세가율	4.7831	0.0287	기각 (Reject)
	COFIX 금리	2.0364	0.1536	채택 (Accept)
	부동산 소비심리지수	2.6525	0.1034	채택 (Accept)
	PIR	0.1591	0.6900	채택 (Accept)
	실업률	0.0005	0.9814	채택 (Accept)

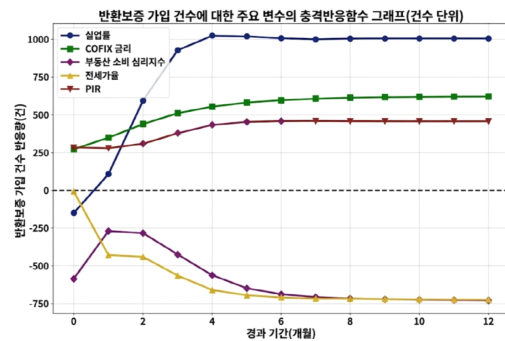
- 주 : 1) 모든 변수를 자연로그 변환 후 1차 차분한 정상 시계열을 바탕으로 본 검정을 수행함.
 2) 귀무가설(H_0)은 '독립변수는 종속변수를 그래저 인과하지 않는다'임.
 3) 유의수준 5% 기준에서 유의확률이 0.05 미만(p -value < 0.05)일 경우, 귀무가설을 '기각'하여 두 변수 간에 유의미한 그래저 인과관계가 존재함을 의미함.
 4) 종속변수와 독립변수의 역방향 검증 결과 모든 변수에서 그래저 인과관계가 존재하지 않음을 확인함.
 5) VECM, vector error correction model; COFIX, cost of funds index; PIR, price income ratio.

관계(동조 현상)를 형성하며 함께 움직일 수는 있으나, 반환보증 가입 수요의 변동을 시계열적으로 선행하여 예측하는 설명력을 갖추었다고 보기에는 어렵다고 할 수 있다. 즉, 해당 변수들은 전세가율 및 부동산 소비심리지수 등 핵심 지표들과 결합하여 전세 시장의 전반적인 리스크 환경을 조성하는 배경 요인이라고 제한적으로 해석하는 것이 타당하다.

3) 충격반응함수(Impulse Response Function) 분석

본 연구는 VECM 모형 기반의 충격반응함수 (impulse response function, IRF)를 도출하여 특정 충격이 반환보증 가입 건수에 미치는 동태적 반응을 12개월간 추적하였으며, 앞선 그래저 인과관계 검정과 연계하여 해석하였다. 반환보증 가입 건수에 대한 주요 변수의 충격반응함수 그래프는 〈그림 1〉과 같으며, 주요 분석 결과는 다음과 같다.

첫째, 예측적 선행성이 확인된 전세가율 충격은 반환보증 가입 건수에 유의미한 음(-)의 반응을 유발하였다. 이는 2022~2023년 매매가격 대비 전세가격 하락으로 수치상 전세가율은 하락했



- 주 : 모든 변수(전세가율, COFIX[cost of funds index] 금리, 부동산 소비 심리지수, PIR[price income ratio], 실업률)가 반환보증 가입 건수에 미치는 영향을 한눈에 볼 수 있는 그래프이다. 검정색 점선(0)은 기준선을 나타내며, 기준선보다 위에 있으면 반환보증 가입 증가(+) 요인, 아래에 있으면 감소(-) 요인이다.

〈그림 1〉 반환보증 가입 건수에 대한 주요 변수의 충격반응함수 그래프

7) 유의수준은 5%로 설정하는 것이 일반적이나, 10% 유의수준($p < 0.10$) 역시 전세시장의 복잡한 마찰요인(2년 계약 등)와 정보 탐색 비용 등 현실적 제약을 반영한다면 통계적 유의성을 띠는 것으로 인정받을 수 있음을 고려하였다.

음에도, 오히려 깡통전세 및 역전세 공포 심리가 자극되어 가입 수요가 폭증했던 비전형적인 시장의 동태성을 반영한다.

둘째, 한계적 유의성($p=0.103$)을 보이며 단기 선행 가능성을 시사했던 '부동산 소비심리지수' 충격 또한 가입 건수에 음(-)의 동태적 반응을 나타냈다. 부동산 시장에 대한 소비자들의 불안 심리가 가중될수록(심리지수 하락), 보증금 미반환에 대한 공포가 확산되어 반환보증 가입 수요를 자극하는 기제가 실재함을 보여준다.

셋째, 그랜저 인과성이 관찰되지 않은 변수들(PIR, COFIX 금리, 실업률)의 충격은 반환보증 가입을 직접 견인하는 인과적 동인이라기보다는, 시스템 내의 구조적 동조 현상으로 해석하는 것이 타당하다. 즉, 해당 변수들은 전반적인 거시경제 지표와 연동되어 주거비 부담이나 시장의 불안 심리를 반영할 뿐, 독립적으로 반환보증 가입 수요를 유발하는 변수로 보기는 어렵다.

4) 예측오차 분산분해(Forecast Error Variance Decomposition) 분석

충격반응함수 분석을 보완하여, 반환보증 가입 건수의 예측 오차 분산이 시스템 내 각 변수의 충

격에 의해 설명되는 상대적 비중(기여도)을 파악하고자 예측오차 분산분해를 실시하였다. 분석결과(〈표 7〉), 반환보증 가입 수요의 결정 요인은 단기와 장기에 따라 구조적인 차이를 보였다.

첫째, 장단기 전반적으로 반환보증 가입 건수 자체의 추세적 관성이 지배적이었다. 이는 반환보증 시장이 거시경제의 즉각적인 변화보다는, 과거에 형성된 시장의 불안 심리와 추세에 강하게 의존하는 자기회귀적 특성을 띠는 시사한다.

둘째, 독립 변수 중에서는 한계적 유의성이 포착되었던 부동산 소비심리지수가 제1의 동인으로 확인되었다. 이는 부동산 소비심리의 불안이 가중될수록, 보증금 미반환에 대한 공포가 확산되어 반환보증 가입 수요를 자극하는 기제가 실재함을 보였던 그랜저 인과관계 분석의 결과를 재차 입증하는 결과이다.

셋째, 12개월 시점 기준 실업률(21.3%)과 전세가율(8.6%)이 그 뒤를 이었다. 단기 그랜저 인과성이 가장 강력했던 전세가율은 즉각적인 위험 회피를 촉발한 후 장기 기여도가 8.6%(12개월)로 비교적 제한적인 반면, 단기 그랜저 인과성이 관찰되지 않았던 실업률은 12개월 차에 21.3%의 설명력을 보였다. 이는 고용 충격이 반환보증 수

〈표 7〉 반환보증 가입 건수에 대한 예측오차 분산분해(variance decomposition) 결과(단위: %)

경과 기간 (Month)	S.E.	반환보증 가입 건수	부동산 소비심리지수	전세가율	실업률	PIR	COFIX 금리	합계
1개월	0.0934	83.89	13.18	0.00	0.44	1.15	1.34	100.0
3개월	0.1397	74.00	11.28	4.24	5.86	2.46	2.15	100.0
6개월	0.2277	49.66	17.09	7.02	19.23	4.08	2.93	100.0
12개월	0.3641	37.90	23.93	8.64	21.32	4.77	3.44	100.0

주 : PIR, price income ratio; COFIX, cost of funds index.

요 자극에 즉각적인 방어책 역할을 하지는 않으나, 시차를 두고 전세 시장의 전반적인 불안에 영향을 미치는 기저 요인임으로 해석할 수 있다. 앞서 그랜저 인과성이 관찰되지 않은 COFIX 금리(3.4%)와 PIR(4.8%)의 설명력은 상대적으로 낮게 도출되었다.

4. 비선형 머신러닝 기반 분석: eXtreme Gradient Boosting와 Shapley Additive Explanations

선형 모형인 VECM은 급변하는 시장의 비선형 상호작용과 임계치 효과를 포착하는 데 한계가 있다. 이에 본 연구는 ML 양상블 기법인 XGBoost를 구축하고, 예측 결과를 XAI 형태로 해석하기 위해 SHAP 변수 중요도 분석을 수행하였다.

1) 모형 구축 및 하이퍼파라미터 최적화

본 연구는 2018년 1월부터 2023년 12월까지를 학습 데이터로, 2024년 1월부터 2024년 12월까지를 테스트 데이터로 설정하여 시계열 예측 모형을 구축하였다. 또한 모형 간 공정한 성능 비교를 위해 VECM(lag=2) 및 LSTM(window size=2)과 동일하게 슬라이딩 윈도우(sliding window) 기법을 적용하여 과거 2개월치의 지연 변수를 생성 및 학습시켰다.

모형의 과적합 방지와 성능 극대화를 위해 시계열 교차 검증(TimeSeriesSplit 3-Fold) 기반의 GridSearchCV를 수행하였다. 그 결과, 최적 파라미터로 학습률(learning rate) 0.1, 트리의 최대 깊이(max depth) 3, subsample 비율 0.8, 의사결정 나무 개수(n_estimators) 100개를 도

출하여 최종 예측 모형을 설정하였다.

또한, 알고리즘 고유의 무작위성에 따른 결과 편차를 통제하기 위해, 중심극한정리에 근거하여 무작위 시드(random seed) 기반의 30회 반복 수행을 실시하였으며, 여기서 도출된 평균값을 최종 성능 지표로 채택하여 분석의 강건성을 확보하였다.

2) 예측 성능 평가(Performance Evaluation)

실증 분석 결과, <표 8>과 같이 ML 모형(XGBoost)이 모든 평가지표에서 벤치마크 모형인 VECM을 압도하는 성과를 보였다. 구체적으로 MAPE 기준, VECM은 15.77%의 오차율을 기록한 반면, XGBoost는 7.63%를 기록하여 예측 오차를 절반 이하(약 51.62% 개선)로 대폭 감소시켰다.

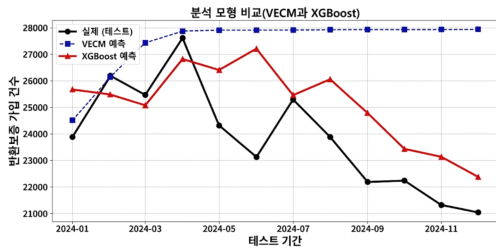
<그림 2>의 시각화 결과에서도 이러한 차이는 명확히 드러난다. VECM(푸른색 점선)은 2024년 초반부터 실제 반환보증 가입 건수(검정색 실선)와의 괴리가 커지며 시장의 변곡점을 제대로

<표 8> VECM과 XGboost 예측 성능 비교 결과(test set 기준)

모형(Model)	RMSE	MAE	MAPE(%)
VECM (Benchmark)	4279.3	3571.9	15.77
XGBoost (ML)	2122.18	1784.25	7.63
성능 개선률(%)	+50.41	+50.05	+51.62

주 : 1) RMSE, MAE, MAPE는 값이 작을수록 예측 성능이 우수함을 의미함.

2) RMSE, root mean squared error; MAE, mean absolute error; MAPE, mean absolute percentage error; VECM, vector error correction model; XGBoost, eXtreme gradient boosting; ML, machine learning.



주 : VECM, vector error correction model; XGBoost, eXtreme gradient boosting.

〈그림 2〉 VECM vs XGBoost 모델 예측 결과 비교

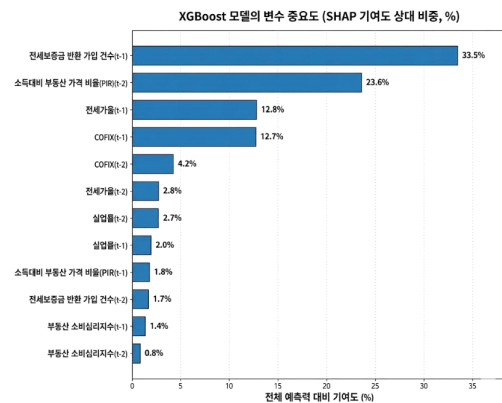
포착하지 못하는 과소 예측 경향을 보였다. 이에 반해 XGBoost(붉은색 실선)는 실제 데이터의 미세한 등락 패턴과 추세를 비교적 정교하게 추종하는 모습을 보였다.

이는 반환보증 가입 행태에 정책 변화(126%를 등)나 시장 심리의 급격한 쓸림 등 선형적 수식으로는 포착하기 어려운 복합적인 비선형 상호작용이 깊게 내재되어 있음을 시사한다.

3) Shapley Additive Explanations 기반 변수 중요도 및 영향력 분석

XGBoost 모델의 예측 기여도를 정량화하기 위해 SHAP 변수 중요도 분석을 수행하였으며, 특히 앞서 적용한 슬라이딩 윈도우 기법을 통해 도출된 과거 2개월(t-1, t-2)의 시차 변수들이 미치는 동태적 영향력을 추적하였다. 머신러닝 모델의 변수 중요도 분석 결과의 시각화는 〈그림 3〉과 같으며, 주요 분석 결과는 다음과 같다.

첫째, '전월(t-1) 반환보증 가입 건수'가 전체의 약 33.5%를 차지하며 1위를 기록했다. 이는 가입 수요가 거시 변수보다 전월 가입 추세에 강하게 의존하는 자기회귀적 관성을 가짐을 시사하



〈그림 3〉 머신러닝 모델의 변수 중요도 분석 (SHAP[shapley additive explanations] 기반)

며, 앞선 분산분해 결과(관성 효과 지배적)와도 논리적으로 일치한다.

둘째, 'PIR(t-2)'이 23.6%의 중요도로 2위를 기록했다. 직전 월(t-1)보다 2개월 전(t-2)의 중요도가 더 높게 나타난 것은, 주택 가격 변화로 인한 매매 부담이 전세 시장으로 전이되어 실제 수요에 반영되기까지 약 2개월의 시차가 존재함을 의미한다.

셋째, '전세가율(t-1)'이 12.8%의 중요도로 3위 변수로 식별되었다. 이는 시장 참여자들이 직전 월의 전세가율 변화를 강동전세 위험의 즉각적인 단기 요인으로 민감하게 인식하고 있음을 보여준다.

넷째, 'COFIX 금리(t-1)'가 12.7%의 중요도로 4위를 차지했으며, 2개월 전의 'COFIX 금리(t-2)' 역시 4.2%의 기여도로 5위에 위치했다. 두 시차의 기여도를 합치면 약 16.9%에 달해 금리의 영향력이 결코 보조적이지 않음을 알 수 있다. 이외의 실업률 및 부동산 소비심리지수 등은 비중이

3% 미만으로 상대적 기여도가 낮아 보조적인 기여 요인으로 작용하고 있음으로 확인되었다.8)

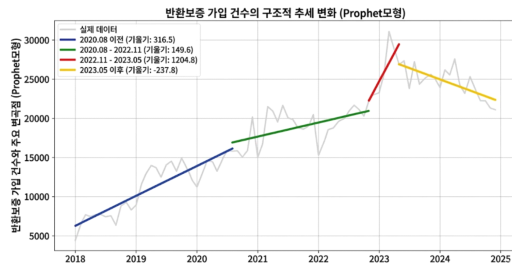
5. 시계열 특화 모형 및 딥러닝 기반 분석: Prophet 및 Long Short-Term Memory

전통적 시계열 모형(VECM)과 ML(XGBoost)에 이어, 시계열 데이터의 구조적 특성과 정밀 예측을 위해 특화 모형인 Prophet과 DL 기반의 LSTM 모형을 적용하였다.

1) Prophet 모형을 이용한 구조적 변화점 식별 및 정책 효과 분석

본 연구에서는 시계열의 구조적 변화점 식별에 탁월한 Prophet 모형을 활용하여, 주요 부동산 정책 도입과 반환보증 가입 추세 간의 연관성을 분석하였다. 분석 결과, <그림 4>와 같이 정책 및 사회적 이슈와 밀접하게 반응하는 세 개의 주요 구조적 변곡점이 확인되었다.

첫째, 2020년 8월(임대차 3법 시행 직후): 이전 기간 대비 상승폭이 다소 둔화되며 완만한 상승세(기울기: +149.6)로 전환되는 패턴이 확인되었다. 이는 계약갱신청구권 등으로 인한 전세 매물 잠김(거래절벽) 현상 등으로 인해 반환보증 가입의 선결 조건인 전세 거래 규모 자체가 감소했고, 그 결과 반환보증 가입 건수의 증가 속도가 둔화된 것으로 해석된다.



<그림 4> Prophet 모형 결과 및 주요 변곡점

둘째, 2022년 11월(빌라왕 사태 본격화): 추세 기울기가 +1,204.8로 폭발적으로 급증하는 강력한 구조적 단절이 발생하였다. 이는 전세 사기 공포가 시장을 완전히 지배하면서, 임차인들의 반환보증 가입이 선택이 아닌 필수 생존 전략으로 바뀌며 가입이 폭발적으로 쇄도했음을 통계적으로 입증한다.9)

셋째, 2023년 5월(126% 룰 시행): 가파르던 상승 추세가 꺾이며 하향 안정화(기울기: -237.8)로 반전되는 변곡점이 확인되었다. 이는 '126% 룰'이라는 정책적 변화가 시장의 과열된 상승 관성을 인위적으로 억제하고, 반환보증 가입 가능 대상을 물리적으로 축소시키는 강력한 구조적 변화 요인으로 작용했음을 시각적으로 뒷받침한다.

2) Long Short-Term Memory 기반 딥러닝 예측 분석

LSTM 모델의 핵심 하이퍼파라미터인 입력 시퀀스 길이(window size)는 2개월로 설정하였

8) VECM 분산분해에서는 선형적 추세를 반영하는 '부동산 소비심리지수'가, XGBoost에서는 비선형적 임계치 효과를 지닌 'PIR'이 핵심 변수로 부각되었다. 이는 분석 결과의 모순이 아니라 모형 간 메커니즘(선형성 vs 비선형성) 차이에 기인한 상호보완적 발견이다. 즉, 본 연구가 계량 모형과 ML 모형을 동시에 비교·분석한 당위성과 차별성을 실증적으로 보여주는 대목이다.

9) 실제 전세사기가 이슈화되고 빌라왕 사망 사건이 보도된 것은 2022년 9~10월이나, 전세사기 사태가 전세시장의 반응과 실질적인 반환보증 가입 폭증으로 반영되기까지 약 1~2개월의 시차가 발생하여 2022년 11월에 구조적 단절점으로 나타났다.

다.10) 분석 결과, <표 9> 및 <그림 5>와 같이 LSTM 모델은 모든 평가지표에서 VECM을 압도하였다.

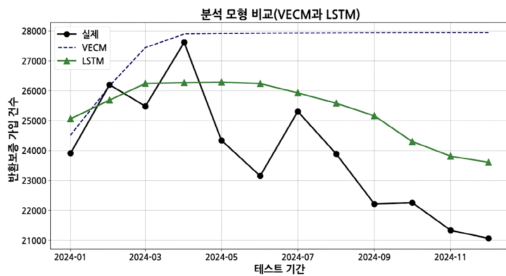
이러한 성능 우위는 LSTM 고유의 '게이트 매커니즘'에 기인한다고 추론할 수 있다. 이 매커니즘이 과거의 불필요한 노이즈(오래된 추세)는 효율적으로 제거하고, 최근의 구조적 변화(빌라왕 사태 및 126% 룰 등)가 반영된 최신 정보에는 높은 가중치를 부여하여 예측에 반영했기 때문이다. 결과적으로, 본 연구는 복잡한 비선형성이 내재된 반환보증 수요 예측에 있어, LSTM 모델의 우수한 측면을 확인하였다.

<표 9> VECM 및 LSTM 모델 예측 성능 비교(2024년 test set 기준)

모형(Model)	RMSE	MAE	MAPE(%)
VECM (Benchmark)	4,279.3	3,571.9	15.77
LSTM (Deep learning)	1,429.24	1,248.29	5.31
성능 개선율(%)	+66.60	+65.05	+66.33

주 : 1) 데이터 전처리는 min-max normalization을 적용하였으며, 평가는 역변환된 원자료 기준으로 수행함.

2) VECM, vector error correction model; LSTM, long short-term memory; RMSE, root mean squared error; MAE, mean absolute error; MAPE, mean absolute percentage error; XGBoost, eXtreme gradient boosting; ML, machine learning.



주 : 1) 파란 점선(VECM)은 검은 실선(실제값)과 크게 이탈하며 방향성을 상실한 반면, 초록색 실선(LSTM)은 검은 실선(실제값)의 하락 반전 추세를 비교적 잘 따라가고 있음을 확인할 수 있다.

2) VECM, vector error correction model; LSTM, long short-term memory.

<그림 5> VECM과 LSTM 모델의 예측 경로 비교

올적으로 제거하고, 최근의 구조적 변화(빌라왕 사태 및 126% 룰 등)가 반영된 최신 정보에는 높은 가중치를 부여하여 예측에 반영했기 때문이다. 결과적으로, 본 연구는 복잡한 비선형성이 내재된 반환보증 수요 예측에 있어, LSTM 모델의 우수한 측면을 확인하였다.

IV. 결론

1. 분석 결과 요약 및 모형별 비교

본 연구는 반환보증 가입의 결정요인을 규명하기 위해 전통적 계량 모형(VECM)부터 최신 DL (LSTM) 기법까지 단계적인 실증 분석을 수행하였다. 각 모형의 분석 목적과 주요 결과는 <표 10>과 같다.

2. 연구의 주요 발견

본 연구의 실증분석 결과, 반환보증 가입은 단순한 선형적 증감이 아닌 거시경제, 부동산 정책 및 사회 이슈에 의해 변화하는 구조적·동태적 과정을 따름을 확인하였다. 주요 발견은 다음과 같다.

첫째, VECM 분석 결과, 전세가율과 부동산 소비심리지수 변동이 반환보증 가입 수요 증감을 유의미하게 선행하였다. 이는 깡통전세 위험이나 부동산 소비심리 악화가 시차를 두고 임차인의 실질적 위험 회피 행동을 견인함을 시사한다.

10) 모형 간 예측 성능 비교의 공정성을 확보하기 위해 VECM에서 최적 시차(lag=2)를 적용한 점, XGBoost에서 슬라이딩 윈도우(sliding window) 기법을 적용하여 과거 2개월치의 지연 변수를 생성 및 학습시킨 점을 고려하였다.

〈표 10〉 연구 모형별 분석 목적, 주요 실증 결과 및 정책적 함의 요약

구분	VECM	Prophet	XGBoost	LSTM
분석 목적	변수 간의 장기 균형 및 인과관계 규명	사회이슈·정책효과 분석 및 구조적 변화점 실증	비선형 상호작용 포착 및 변수 중요도 식별	시계열 패턴 학습을 통한 정밀 예측
예측 성능 (MAPE)	15.77% (선형 모형의 한계)	- (구조 분석 특화)	7.63% (VECM 대비 대폭 개선)	5.31% (가장 정밀한 예측력)
성능 우위 원인	(벤치마크 모형)	정책 충격 시점의 추세 변화 포착에 탁월	복잡한 비선형 패턴과 임계치 효과를 학습하여 급변동 구간 정교 추종	과거 노이즈는 버리고 최신 구조적 변화를 효율적으로 기억(Gate 메커니즘)
정책적 활용 방안	[기초 분석] 거시경제 충격이 시장에 미치는 파급 경로 파악	[정책 모니터링] 부동산 정책 도입 시 시장의 반응(변곡점) 진단	[보증료율 체계 개선] 리스크 임계치를 반영한 보증료율 차등화 등	[조기 경보 시스템] 경제지표를 학습시켜 보증사고 위험 조기 예측

주 : VECM, vector error correction model; XGBoost, eXtreme gradient boosting; LSTM, long short-term memory; MAPE, mean absolute percentage error.

둘째, XGBoost 및 SHAP 분석으로 반환보증 가입 수요의 비선형성 및 임계치 효과를 확인했다. 반환보증 가입 수요는 특정 위험 수준을 넘어설 때 급증하는 ‘비선형적 임계치 효과’를 보였으며, 전월의 가입 추세(관성)에 가장 큰 영향을 받는 가운데 1~2개월 전의 전세가율과 PIR이 가장 민감한 지표로 작용하는 것으로 나타났다.

셋째, Prophet 모형을 통해 부동산 정책 및 사회적 이슈 등 외부 충격에 따른 시장의 구조적 단절을 규명했다. 2022년 10월 본격화된 ‘빌라왕 사태’는 반환보증 수요를 폭증시킨 가장 강력한 요인이었다. 반면, 2023년 5월 시행된 ‘126% 룰’은 가입 대상을 축소시키며 과열된 상승 관성을 억제한 결정적 단절 요인으로 확인되었다.

넷째, LSTM 모형은 최신 구조적 변화(전세사기 이슈 등)를 즉각 반영하는 고유의 게이트 메커니즘을 통해, 기존 선형 모형(VECM) 대비 66% 이상 개선된 가장 뛰어난 예측 성능(MAPE 5.31%)을 기록했다. 이는 향후 급변하는 전세 시장에서 AI (artificial intelligence) 및 DL 기법을 활용한

보증 사고 선제 감지 및 차등 보증료율 산정 등의 정책 도구로 활용될 수 있는 가능성을 시사한다.

3. 연구의 학술적 의의 및 정책적 시사점

1) 연구의 학술적 의의

기존 연구들이 선형 및 ML 모형을 주로 주택 가격 예측에 활용한 반면, 본 연구는 이를 주택금융 정책 수요인 ‘반환보증 가입’의 동태적 비선형성 분석 및 행태적 영역으로 확장 분석한 데에 의의가 있다.

특히, 전통적 계량 모형(VECM 등)으로 포착하기 어려웠던 시장 참여자들의 비선형적 임계 효과와 정책 변화(126% 룰 등)가 유발하는 시계열의 구조적 단절 현상을 ML 및 DL 기법을 통해 정량적으로 규명하였다. 이는 전세 시장 분석의 방법론적 한계를 보완하고, 향후 부동산 정책의 파급 효과를 다각적으로 분석할 수 있는 새로운 틀을 제공한다는 점에서 차별화된 통찰을 지닌다.

2) 정책적 시사점

분석 결과를 바탕으로, 반환보증 제도의 지속 가능성과 전세 시장의 안정성을 제고하기 위한 정책적 시사점을 다음과 같이 고찰해 보았다.

첫째, 향후 ‘리스크 기반의 차등적 보증료율’ 산정 기준을 논의함에 있어 본 연구의 분석 프레임워크가 기초 자료로 활용될 수 있다. 현재 보증기관의 획일적인 보증료율 체계는 전세시장의 개별 위험을 보증료에 제대로 반영하지 못하고 있다. 본 연구의 XGBoost 및 SHAP 분석 결과, PIR과 전세가율이 특정 임계치를 초과하는 구간에서 보증기관의 리스크가 급격히 증가하는 비선형적 패턴이 확인되었다. 향후 후속 연구를 통해 고위험 물건에 대해서는 할증된 요율을 적용하되 저위험 물건은 인하하는 등 ‘임계치 기반의 적정 보증료율 산정 모델’ 등의 개발을 고민해 볼 수 있다. 이는 보증기관의 재무 건전성을 확보하는 주요 기제가 될 수 있을 것이다.

둘째, AI 및 DL 기반의 ‘조기 경보 시스템’ 도입의 가능성을 검토해 볼 수 있다. 본 연구에서 LSTM 모델은 기존 계량 모형대비 2배 이상의 예측 정확도(MAPE 5.31%)를 입증하였다. 이는 시장의 구조적 변화가 잦은 전세 시장에서 과거 데이터에 의존하는 선형적 예측은 한계가 있음을 의미한다. 정책 당국은 보증기관 내에 AI 및 DL 기반의 모니터링 체계를 통해, 거시경제 충격 지표(금리, 실업률 등)나 전세가율 변동에 따른 반환보증 사고 위험을 선제적으로 파악하는 기반을 마련할 수 있다.

셋째, ‘126% 룰’의 보완책 마련을 검토해 볼 수 있다. Prophet 분석을 통해 확인된 바와 같이,

2023년 5월 반환보증 가입 요건 강화(126% 룰) 이후 반환보증 수요 추세는 구조적인 단절 국면으로 전환되었다. 이는 획일적인 보증 가입 요건 강화가 리스크 관리에는 기여했을지라도, 시장 전체의 반환보증 가입 접근성을 위축시키는 부작용을 동반하였을 가능성이 높음을 시사한다. 따라서 반환보증 가입 강화 요건을 일괄적으로 적용하기보다는, 지역별·주택 유형별 가격 데이터 및 피보증인·임대인 데이터 등을 폭넓게 연동하여 가입 한도를 조정하는 등 정책 보완을 고려해볼 수 있다.

4. 연구의 한계 및 향후 과제

본 연구는 전통적 계량 모형과 ML, DL을 비교 분석하여 유의미한 시사점을 도출하였으나, 다음과 같은 한계를 지닌다.

첫째, 시계열 데이터의 제약이다. 분석 기간이 총 84개월(2018.01.~2024.12.)로 제한되어 복잡한 ML 모형의 과적합 우려를 완전히 배제하기 어려우며, 시계열이 확보된 HUG 데이터만 활용하여 타 보증기관(HF, SGI 등)과의 미세한 제도적 차이를 반영하지 못했다.

둘째, 분석 결과의 거시적 일반화 문제이다. 제시된 정책 제언은 단기적 예측 성능에 기반하므로, 이를 구조적인 거시 정책의 절대적 근거로 직결하기보다는 장기적 안목의 신중한 접근이 요구된다.

향후 충분한 데이터 축적과 다양한 거시 통제 변수 추가를 통해 모형의 장기적 안정성을 엄밀히 검증하는 후속 연구가 필요하다. 결론적으로, 반

환보증 제도는 단순한 주거 복지 차원을 넘어 주택 금융 시스템의 리스크 관리 수단으로 고도화될 필요성이 있으며, 본 연구에서 제시한 데이터 기반의 기술적 분석과 예측 분석이 풍부한 기초자료가 되길 기대한다.

ORCID

이항재 <https://orcid.org/0009-0005-9226-3397>

참고문헌

1. 권혁, 성현근. (2025). 전세보증금반환보증 가입 요건의 변화가 고위험 전세 발생 확률에 미친 영향: 서울시 연립, 다세대주택 시장을 중심으로. *국토계획*, 60(3), 101-112.
2. 김경환, 손재영, 정의철. (2025). *부동산경제학*(제4판). 건국대학교 출판부.
3. 김성민, 우명제. (2026). 전세보증사고의 공간적 분포 특성과 영향 요인 분석. *주택도시금융연구*, 11(1), 45-60.
4. 김승현, 김원혁, 이윤수. (2022). 머신러닝과 패널 고정효과를 활용한 아파트 실거래가 예측. *주택연구*, 30(1), 43-69.
5. 김진유. (2022). 고위험 전세와 전세보증금 미반환 위험의 상관관계 분석: 서울시 전세보증사고를 중심으로. *부동산학연구*, 28(4), 55-69.
6. 박진백, 김지혜, 권건우. (2023). 전세 레버리지 리스크 추정과 정책대응 방안. *국토정책 Brief*, 23.
7. 손하늘. (2022.10.13.). 수백 채 전세 사기 혐의로 수사받던 '빌라왕'··호텔서 숨진 채 발견. *MBC뉴스* <https://n.news.naver.com/mnews/article/214/0001228212?sid=102>
8. 안선영, 이상엽. (2025). 전세보증금 미반환에 영향을 미치는 주요요인 연구: 수도권지역 전세보증 사고를 중심으로. *주택금융연구*, 9(2), 47-68.
9. 이영수. (2010). 주택가격과 전세가격: VECM 분석. *부동산학연구*, 16(4), 21-32.
10. 이지연, 이정란. (2026). 전세보증금 반환보증 제도 변화가 주택 임대차 시장 구조에 미친 영향: 서울시 임대차 실거래 자료를 중심으로. *주택도시금융연구*, 11(1), 61-79.
11. 주택도시보증공사. (2020). *전세임대주택 전세보증금 반환보증규정*. 주택도시보증공사.
12. Enders, W. (2014). *Applied econometric time series* (4th ed.). John Wiley & Son.
13. Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1), 37-45.

논문 접수일: 2026년 2월 1일

심사(수정)일: 2026년 3월 27일

게재 확정일: 2026년 4월 17일

국문초록

최근 금리 인상과 전세 사기 확산으로 전세보증금 미반환 리스크가 대두되며 '전세보증금 반환보증(이하 반환보증)' 가입이 급증했다. 본 연구는 2018년 1월부터 2024년 12월까지의 시계열 데이터를 바탕으로 반환보증 가입 결정요인과 구조적 변화를 분석했다. 이를 위해 전통적 선형 모형인 VECM(vector error correction model), 비선형성을 포착하는 XGBoost(eXtreme gradient boosting), 시계열 예측에 특화된 Prophet, 딥러닝 기반의 LSTM(long short-term memory) 모델을 종합적으로 적용하였다. 주요 분석 결과는 다음과 같다. 첫째, VECM 분석을 통해 전세가율과 부동산 소비심리지수가 수요를 견인하는 핵심 동인임을 확인했다. 둘째, XGBoost와 SHAP(shapley additive explanations) 분석으로 전세가율 및 PIR(price income ratio)이 특정 임계치를 초과할 때 반환보증 가입 수요에 대한 예측기여도가 비선형적으로 급증하는 '임계치 효과'를 확인했다. 셋째, Prophet 모형은 '빌라왕 사태(2022.11.)'가 반환보증 가입 급증의 트리거, '126% 룰(2023.05.)'이 반환보증 가입 감소의 구조적 단절점임을 실증했다. 넷째, LSTM은 MAPE(mean absolute percentage error) 5.31%를 기록해 기존 선형 모형 대비 뛰어난 예측력을 입증했다. 본 연구는 반환보증 가입의 동태적 비선형성을 실증하여 기존 선형 모형의 한계를 보완했다는 의의가 있다. 결론적으로 반환보증 제도의 지속 가능성을 위해 리스크 기반의 보증료율 현실화 및 딥러닝 기반 조기 경보 시스템 도입 가능성 논의 등을 위한 학술적 기초자료 및 정책적 방향성을 제시한다.

주제어 : 전세보증금 반환보증, 벡터오차수정모형, 머신러닝, 딥러닝, 반환보증 126% 룰