



온라인 기반 부동산 심리지수(Real Estate Sentiment Index)의 개발과 아파트 매매가격 반응 분석 - 머신러닝 예측과 Shapley Additive Explanations 해석을 활용한 접근 -

Development of an Online Real Estate Sentiment Index and Analysis of Apartment Price Responses - A Machine Learning Prediction and Shapley Additive Explanations Interpretation Approach -

박순미* · 이성화**

Soonmi Park · Seonghwa Lee

Abstract

This study develops an online real estate sentiment index (RSI) using unstructured textual data and compares its explanatory power with that of the traditional survey-based index (CSI) in predicting apartment prices in Korea. The analysis employs monthly data from January 2016 to December 2024. Sentiment classification is performed using the KR-FinBert-SC model, built on KoBERT. The RSI is constructed by reclassifying neutral documents based on softmax probabilities and computing the difference between the number of positive and negative documents. The empirical methods employed include univariate regression, Granger causality tests, random forest modeling, Shapley additive explanations (SHAP) analysis, and co-integration-based regression. Results indicate that the RSI significantly outperforms the CSI in both explanatory power ($R^2=0.4359$ vs. 0.0079) and predictive accuracy. Granger causality tests show a statistical relationship significant with a 2-3 month lag. The RSI also exhibited leading behavior at all 16 identified price turning points, including those driven by interest rate shocks. SHAP analysis confirmed the RSI's nonlinear influence and structural contribution to price prediction. Overall, the RSI demonstrates strong potential as an early warning indicator, a behavioral policy tool, and a complement to macroeconomic variables by providing a real-time, emotion-sensitive measure of market sentiment.

Keywords: Real estate sentiment index, Sentiment analysis, Random forest, Shapley additive explanations (SHAP) interpretation, Self-reinforcing mechanism

* 대구대학교 일반대학원 부동산학과 박사과정(주저자) | Ph.D. Student, Department of Real Estate Studies, Graduate School, Daegu University | First Author | soonme1234@naver.com |

** 대구대학교 일반대학원 부동산학과 교수(교신저자) | Professor, Department of Real Estate Studies, Graduate School, Daegu University | Corresponding Author | land@daegu.ac.kr |

1. 서론

1. 연구의 배경과 목적

최근 부동산 시장은 전통적인 경제지표만으로 설명하기 어려운 고변동성을 보이고 있으며, 시장참여자의 심리와 감성적 반응이 가격 형성에 중대한 영향을 미치는 사례가 증가하고 있다. 특히 정부 정책 발표, 언론 보도, 온라인 확산 등 비정형 정보가 심리적 파급력을 증폭시키면서, 기대 심리나 불안심리와 같은 정서적 요인이 실질 가격에 선행하여 반영되는 구조가 강화되고 있다(박재수, 2020; 이현정, 2023).

이에 따라 시장 심리를 계량적으로 파악하려는 다양한 시도가 이어졌으나, 현재까지 활용되고 있는 대부분의 심리지표는 설문조사 기반 지표로서 표본의 대표성 문제, 응답자의 주관 개입, 실시간성 부족이라는 구조적 한계를 안고 있다. 대표적 지표인 소비자심리지수(consumer sentiment index, CSI) 또한 응답자의 주관적 기대에 기반하기 때문에 실시간 시장 반응을 민감하게 반영하기 어려우며, 특히 예측력과 선행성 측면에서 한계를 드러낸다는 지적이 있다(김진석, 2024).

이러한 한계를 보완하기 위해 본 연구는 온라인에서 자생적으로 생성되는 비정형 텍스트 데이터(뉴스·블로그 등)를 감성 분석하여, 실시간으로 시장심리를 수치화한 '부동산 심리지수(real estate sentiment index, RSI)'를 개발하고자 한다. 나아가 본 연구는 RSI가 기존 CSI보다 가격에 대해 더 높은 설명력(R^2), 예측력(root mean squared error, RMSE), 선행성(Granger 인과

성)을 보이는지를 실증적으로 비교 분석함으로써, RSI의 계량적 유용성과 정책적 적용 가능성을 입증하는 것을 주된 연구 목적으로 삼는다.

2. 연구의 차별성

본 연구는 기존 문헌과 다음과 같은 측면에서 명확한 차별성을 가진다.

첫째, 텍스트 데이터를 기반으로 부동산 심리지수(RSI)를 직접 구축하고, 이를 기존 설문조사 기반 CSI와 비교하여 설명력(R^2), 예측력(RMSE), 선행성(Granger 인과성)을 정량적으로 검증하였다.

둘째, random forest와 SHAP(Shapley additive explanations) 해석기법을 통해 RSI가 금리, 물가 등 거시변수와 비교해 높은 설명력과 구조적 기여도를 가진다는 점을 실증적으로 입증하였으며, 금리 충격기 및 가격 변곡점 구간에서 RSI가 선행적으로 반응하는 구조적 특성을 확인하였다.

셋째, 아파트 매매가격지수, 주택담보대출금리, 소비자물가 등 공공통계자료와 뉴스·블로그 등 비정형 감성데이터를 통합하여 분석한 점에서, 공공데이터와 빅데이터를 융합한 실증 분석 사례로서 이론적 정합성, 실증적 설명력, 정책 활용 가능성을 동시에 확보하였다.

3. 논문 구성

본 논문은 다음과 같이 구성된다.

II장에서는 심리지수 개념과 감성 분석, 머신

러닝 기반 설명력 분석에 대한 선행연구를 고찰하고, 본 연구의 차별성과 필요성을 제시한다.

III장에서는 데이터 구성, 변수 정의, RSI 구축 방식 및 통계적 특성을 설명한다.

IV장에서는 RSI와 CSI의 시계열 추이 비교, 회귀분석, Granger 인과성 분석, random forest 및 SHAP 기반 영향력 분석, 정책 반응 시점에서의 RSI의 반응 구조 등을 중심으로 실증 결과를 제시한다.

마지막으로 V장에서는 연구 결과를 요약하고, 정책적 시사점과 함께 RSI의 활용 가능성을 논의하며, 향후 연구 방향을 제시한다.

II. 이론적 배경 및 선행연구 검토

1. 기존 부동산 심리지수의 한계와 대체 필요성

부동산 시장은 불완전 정보, 가격 경직성, 거래 비용 등 구조적 제약 속에서 심리 요인의 영향을 크게 받는다. 이를 반영하기 위해 개발된 소비자심리지수(CSI) 및 주택가격전망지수는 시장 참여자의 기대와 심리를 수치화하는 데 일정한 역할을 해 왔다(김대원·유정석, 2013; 박재수·이재수, 2021; 조태진, 2014). 특히 한국은행과 통계청은 정기적인 설문조사를 통해 소비자의 인식과 기대를 지수화하여 시장 흐름을 간접적으로 파악해 왔다.

그러나 이들 지수는 설문조사 기반이라는 구조적 특성으로 인해 다음과 같은 한계를 지닌다. 첫

째, 응답자의 기대와 실제 시장 행위 간에는 괴리가 존재할 수 있으며, 이는 손실회피 성향, 기준점 의존성 등 비합리적 선택행동을 설명하는 전망이론(prospect theory)에서도 이론적으로 제시된 바 있다(Kahneman and Tversky, 1979).¹⁾ 둘째, 표본 대표성과 응답 편향 문제가 지속적으로 제기되고 있다. 특히 설문 응답의 심리적 왜곡은 감정 기반의 경제 행동이 시장 결과에 실질적 영향을 미친다는 ‘animal spirits’ 개념과도 연결된다(Akerlof and Shiller, 2009).²⁾ 셋째, 조사·공표의 시차로 인해 실시간 시장 변화에 대한 민감한 반응이 어려우며, 거시변수와의 시차 구조 해석에도 제약이 존재한다(유한수·정재호, 2015; 최성수, 2025).

이러한 문제점을 극복하기 위한 대안으로, 감성 분석 기반의 실시간 지수 산출 방식이 주목받고 있다. 특히 비정형 텍스트 데이터를 활용한 부동산 심리지수 개발 시도는 기존 설문 기반 방식의 한계를 극복할 수 있는 유의미한 접근으로 평가된다.

예컨대, 박재수(2020)는 신문·방송 기사를 대상으로 기계학습 기반 감성분석을 수행하여 부동산 감성지수를 개발하였으며, 해당 지수가 아파트 매매가격 예측에서 유의미한 향상 효과를 보였음을 ARIMAX 모형을 통해 입증하였다. 이종민(2018)은 뉴스기사의 토픽 비중과 포털 검색지수를 활용하여 전세가격 예측 모형을 개선하고자 ARIMAX를 적용하였으며, ARIMA 대비 MAE(mean absolute error) 및 RMSE 기준으로 예측력이 최대 약 26%까지 향상됨을 실증적으로 확

1) Prospect theory는 손실보다 손실 회피에 민감한 인간 행동을 설명하는 행동경제학 이론이다.

2) Animal spirits는 경제 주체의 감정, 직관, 신뢰 등이 경제 결정에 영향을 미친다는 개념이다.

인하였다. 이는 시장 참여자의 심리를 반영한 비정형 데이터의 정량화 및 예측모형 활용 가능성을 제시한 연구이다. 이현정(2023)은 경기 순환기(확장기·수축기)에 따라 언론보도에 나타난 부동산 투자심리의 주요 키워드를 텍스트 마이닝 및 토픽 모델링을 통해 분석하였으며, 투자심리 변화가 시기별 실물 시장 흐름(예: 금리, 정책, COVID-19 등)과 밀접하게 연결되어 있음을 확인하였다.

이러한 일련의 연구들은 감성분석 기반 실시간 심리지표의 타당성을 뒷받침하고 있으며, 실시간성, 비구조적 정보 수용, 데이터 확장성 등을 갖춘 텍스트 기반 심리지수에 대한 관심이 확대되고 있으며, 감성분석을 활용한 데이터 기반 심리지표는 기존 소비자심리지수(CSI)의 한계를 보완할 대안으로 주목받고 있다(김수아 외, 2024; 이현정·오윤경, 2024; 주종웅, 2023; Chen et al., 2014; Shayaa et al., 2018; Petropoulos and Siakoulis, 2021). 특히 소셜미디어에 표출된 집단 감정이 자산시장에 체계적 영향을 미친다는 분석도 다수 존재한다(Chen et al., 2014).

2. 비정형 데이터 기반 심리지수와 머신러닝 해석기법

자연어처리(NLP) 기술은 비정형 텍스트로부터 감성 극성을 수치화할 수 있는 도구로 자리매김하고 있다. 특히 BERT 계열의 언어모델은 감성 분류 정확도가 높아 금융 및 경제 분야에서의 응용 가능성이 확인되었으며, KoBERT는 한국어에 특화된 구조로 다양한 국내 실증 연구에서 채택되고 있다(김수아 외, 2024; 박예나·유석중, 2022).

최근에는 뉴스, 블로그, 커뮤니티 등에서 수집된 문서의 제목 또는 본문을 대상으로 감성 분석을 수행하고, 이를 월별로 집계하여 정량화된 지수로 환산하는 시도가 이어지고 있다(김규석, 2025; 박재수·이재수, 2021; 주종웅, 2023; Petropoulos and Siakoulis, 2021; Wojarnik, 2022).

이러한 연구 중 김진석(2024)은 머신러닝 기반의 자동가치평가모형(automated valuation model, AVM)을 구축하고, 인공신경망(artificial neural network, ANN)과 LightGBM을 활용하여 시공간적 분포를 반영한 주택가격지수를 개발하였으며, 해당 지수가 평가 기반 지수 대비 민감도와 정확도가 높고, 소지역 단위에서도 안정적으로 작동함을 실증하였다. 이정현(2023)은 부동산 시장의 고변동성 구간을 대상으로 ARIMA, VAR, LSTM을 비교하여 인공지능 학습모형의 예측력을 검증하였으며, 설명가능 인공지능 기법(SHAP)을 적용해 변수별 영향력을 분석함으로써 예측력과 해석력을 동시에 확보할 수 있는 가능성을 실증하였다. 또한 이승규(2019)는 금리·통화량 등 거시경제변수와 포털 키워드 검색빈도를 활용하여 아파트 매매가격과 전세가격에 미치는 영향을 분석하였으며, 검색빈도가 서울 아파트 전세가격에 유의미한 정(+)의 영향을 주는 등 비정형 데이터가 주택시장에 간접적으로 작용함을 VEC 모형을 통해 실증하였다.

본 연구에서는 KoBERT 기반 모델 중 하나인 KR-FinBert-SC를 활용하여 제목 기반 감성분석을 수행하였으며, 중립 레이블을 logit confidence에 따라 극성으로 재분류한 후, 월별 긍정 문서 수와 부정 문서 수의 차이를 기반으로 부동산 심리

지수(RSI)를 산출하였다. RSI는 (긍정-부정) 문서 수를 지표화한 구조로, 수치가 클수록 긍정 심리가 우세함을 나타낸다.

이러한 방식은 첫째, 대용량의 비정형 데이터를 실시간으로 처리할 수 있으며, 둘째, 수치화된 지표를 통해 계량적 해석과 통계모형 적용이 가능하다는 점에서 기존 CSI 대비 우월한 해석력을 확보할 수 있다(김규석, 2025; 김수아 외, 2024; 양건필·전해정, 2021; Akerlof and Shiller, 2009; Barberis et al., 1998). 특히 단기 충격에 대한 반응 속도, 군집심리의 동조성(co-movement), 정책 발표 이후의 자기강화 경로(self-reinforcing structure) 등을 분석하는 데 강점을 가진다. 이와 유사한 접근으로 김동섭(2020)은 미국 주택담보대출 데이터를 활용하여 인공신경망과 랜덤포레스트 기반의 기계학습 모형을 구축하고, SHAP 및 PDP 등의 설명가능 인공지능 기법을 통해 채무불이행 위험 요인을 분석하였으며, 전통적 계량모형 대비 기계학습 모형의 예측력과 설명가능성 모두에서 우수성을 실증하였다. 이정현(2023) 역시 부동산 가격 예측에 SHAP 분석을 적용함으로써, 정책·금리 등 개별 변수의 영향력을 시계열 구조에서 시각화할 수 있는 방법론적 가능성을 제시한 바 있다.

아울러 RSI의 설명력과 시장 반응 구조를 검증하기 위해, 본 연구는 머신러닝 기반의 random forest 회귀모형과 변수 해석 기법인 SHAP를 병행 적용하였다. Random forest는 변수 간 비선형 관계 및 상호작용을 반영할 수 있어 구조적 예

측에 적합하며, SHAP은 각 변수의 영향력 기여도를 시계열 단위로 분해할 수 있는 설명기법으로, 게임이론 기반의 공정한 기여도 산출 방식을 바탕으로 모델 예측 결과에 대한 해석 가능성을 획기적으로 개선시켰다(이정현, 2023; Hong et al., 2020; Lastrapes, 2002; Kim and Kim, 2022; Wu, 2023).³⁾

이와 같이 본 연구의 접근은 단순한 감성 점수 산출을 넘어서, 심리적 반응의 시간 구조와 경제 변수 간 상호작용을 정량적으로 해석할 수 있는 계량적 기반을 마련하는 데 의의가 있다.

III. 자료 및 변수 구성

1. 자료 수집 및 기간

본 연구는 2016년 1월부터 2024년 12월까지의 월간 데이터를 분석 대상으로 설정하였다. 종속변수는 한국부동산원이 제공하는 '전국 아파트 매매가격지수'를 활용하였으며, 실거래가 기반으로 구축된 본 지수는 시장 가격의 실제 흐름을 안정적으로 반영하는 특징이 있다.

독립변수는 총 세 가지로 구성된다. 첫째, 주택담보대출 신규취급액 기준 금리(int)는 한국은행 경제통계시스템의 월별 대출금리 자료를 활용하였다. 둘째, 소비자물가상승률(cpi_yoy)은 통계청이 제공하는 소비자물가지수(consumer price index, CPI)의 전년 동월 대비 상승률로, 다음의

3) SHAP은 게임이론의 Shapley 값을 기반으로 각 변수의 예측 기여도를 분해하는 설명기법이다.

(식 1)에 따라 계산하였다.

$$CPI_YoY_t = \frac{CPI_t - CPI_{t-12}}{CPI_{t-12}} \times 100 \quad (\text{식 1})$$

셋째, 감성 기반 부동산 심리지수(RSI)는 본 연구자가 수집한 뉴스 및 블로그 데이터를 바탕으로 직접 산출하였다.

2. Real Estate Sentiment Index 산출 방식

본 연구에서 제안한 RSI는 총 226,957건의 부동산 관련 뉴스 및 블로그 데이터를 수집한 후, 각 문서의 제목(title)만을 분석 대상으로 감성 분류를 수행하였다. 감성 분석에는 Hugging Face에서 제공하는 KR-FinBert-SC 모델을 활용하였으며, 각 문서는 긍정, 부정, 중립의 세 감성 범주 중 하나로 분류되었다.

분류 결과 중 중립으로 판별된 문서는 제거하지 않고, 해당 문서에 대해 softmax 기반 긍정 및 부정 확률값 중 더 높은 값을 기준으로 극성(label)을 재분류하였다. 즉, 긍정 확률이 부정보다 높을 경우 '긍정'으로, 그 반대일 경우 '부정'으로 이진화하였다. 이후 월별로 집계된 긍정·부정 문서 수를 기반으로 RSI를 산출하였으며, 기본 공식은 (식 2)와 같다.

$$RSI_t = \text{긍정 문서 수}_t - \text{부정 문서 수}_t \quad (\text{식 2})$$

이 지수는 특정 월에 시장 내 긍정 심리가 부정 심리보다 얼마나 우세한지를 절대 격차로 표현한 값으로, 수치가 클수록 긍정 심리가 강하게 나타

났음을 의미한다. 정규화된 비율이 아닌 실제 감성 격차의 크기를 그대로 반영함으로써, 감성 반응의 강도를 보다 직관적으로 해석할 수 있도록 설계되었다. RSI는 월 단위로 집계되었으며, 각 월별 문서 수는 최소 990건에서 최대 4,000건 이상으로 분포하였다. 월별 문서 수의 차이는 부동산 가격의 안정기보다 부동산 가격의 변동기에 더 많이 작성되기 때문으로 파악되었다. 이는 RSI 지표의 근간이 되는 문서의 작성 빈도가 부동산 시장의 흐름에 적절하게 반응하고 있다는 것으로 주기성이나 계절성과는 큰 관계가 없음을 시사한다. 전체 문서의 평균은 약 2,100건으로, 분석에 필요한 통계적 안정성을 확보하였다. 한편, 감성 지수 설계 과정에서 '중립 문서'를 어떻게 처리할 것인가는 중요한 분석적 판단 요소로 간주된다. 감성 분석 선행연구에서는 중립 문서를 제거하거나, 긍정·부정과 구분하여 별도로 처리하는 사례가 다수 보고되어 왔으며(Kim et al., 2021; Zhang and Liu, 2017), 이처럼 중립의 처리 방식은 분석 목적, 도메인 특성, 데이터 분포 등에 따라 선택적으로 활용되어 왔다. 본 연구는 이러한 선행 사례의 경향을 바탕으로, 중립 문서를 softmax 기반 확률값에 따라 극성이 강한 방향으로 재분류한 방식(rsi_refined), 동일한 재분류 기반에서 긍정-부정 단순 차이를 적용한 방식(posneg_refinde), 그리고 중립 문서를 제거하거나 재분류하지 않고 포함한 방식(rsi_total)의 설명력 차이를 계량적으로 비교하였다. 또한, 설문 기반 주택가격전망지수(housecsi)와의 단변량 회귀 분석도 병행하여 비교하였다. 분석 결과는 <표 1>과 같다.

〈표 1〉 중립 문서 처리 방식에 따른 RSI 설명력 비교 결과

독립변수	회귀계수(β)	표준오차(SE)	t값	p값	R ²	Adj. R ²	F값	F_p값	DW값
rsi_refined	53.25	8.18	6.51	<0.001	0.286	0.279	42.38	<0.001	0.295
posneg_refinde	0.024	0.0026	9.05	<0.001	0.436	0.431	81.9	<0.001	0.455
rsi_total	-53.09	72.32	-0.73	0.465	0.005	-0.004	0.54	0.465	0.016

주 : RSI, real estate sentiment index.

분석 결과, softmax 기반 극성 재분류 방식을 적용한 rsi_refined 및 posneg_refinde는 각각 R²=0.2856, 0.4359로 높은 설명력을 기록하였다. 반면 중립을 그대로 포함한 rsi_total은 R²=0.0051, 설문 기반 housecsi는 R²=0.0079로 분석되었으며, 두 지표 모두 통계적으로 유의하지 않았다.

이는 감성지수 설계 시 중립 문서를 단순 포함하거나 제거하기보다는, 극성 확률 기반의 정교한 재분류 방식이 시장 심리의 실질적 반영에 더 적합함을 시사한다. 본 결과는 RSI의 설계 타당성을 검증하고, 기존 CSI 기반 지표 대비 본 연구의

접근이 갖는 설명력 우위를 정량적으로 입증하는 근거로 기능한다.

3. 변수 정의 및 요약 통계

본 연구에서 분석에 활용한 주요 변수는 아파트 매매가격지수(apt), 부동산 심리지수(RSI), 주택가격전망지수, 주택담보대출금리(int), 소비자물가지수 전년동월비(cpi_yoy)이며, 모두 2016년 1월부터 2024년 12월까지의 월간 자료를 기준으로 구성하였다. 분석에 사용된 변수는 〈표 2〉와 같이 정의된다.

〈표 2〉 변수개요 및 설명

구분	변수명	출처	설명
종속변수	아파트매매가격지수 (apt)	한국부동산원	전국 월별 아파트 매매가격 수준. 기준시점(2021년 6월=100) 정규화된 실물 자산 지수
독립변수	주택담보대출금리 (int)	한국은행 ECOS	예금은행 신규취급액 기준 대출금리. 자금조달 비용 및 통화정책 전달경로 대리 변수
	물가 (cpi_yoy)	통계청	소비자물가지수 전년동월비. 인플레이션 기대 및 실질구매력 반영 변수
	부동산심리지수 (RSI)	자체구축	KR-FinBert-SC 모델기반 긍·부정 문서 비율로 산출된 비정형 정서 지표
비교변수	주택가격전망지수	한국은행 (소비자동향조사)	한국은행의 소비자동향조사에서 제공하는 '주택가격전망지수'를 활용. 이 지수는 응답자의 향후 1년간 주택가격에 대한 기대를 바탕으로 2008년부터 매월 조사되고 있으며, 소비자심리지수(CSI)를 구성하는 하위 항목 중 하나로 향후 부동산 가격에 대한 집단적 기대심리를 정량화한 지표로 활용.

주 : RSI, real estate sentiment index; CSI, consumer sentiment index.

1) 주요 변수의 시계열 추이

〈그림 1〉은 각 변수의 시계열적 움직임을 시각화한 결과로, 전체 분석 기간 동안의 변동 흐름을 비교하여 보여준다.

〈그림 1〉은 주요 변수의 z-score 표준화 시계열을 시각화한 결과로, 아파트 매매가격지수를 중심으로 변수 간 구조적 연동성과 시차 구조를 비교하였다. 아파트 매매가격지수는 2016년부터 2019년까지 완만한 상승세를 유지하다가, 2020년 중반 이후 급등세로 전환되며 2022년 상반기 최고점에 도달한 뒤 하락세로 전환되었다. 이후 2024년까지는 약 93선에서 보합 흐름을 보이며 안정화되는 모습이다. 부동산 심리지수(RSI)는 상승과 하락 전환 시점마다 가격보다 선행하여 움직이는 양상이 반복적으로 나타나며, 심리 반응이 가격 변화에 앞서 발생함을 시사한다. 특히 2020년(팬데믹)과 2022년(금리충격기)

을 전후한 급변 구간에서 RSI는 아파트 가격보다 먼저 꺾이거나 상승하며 구조적 선행성을 보여준다. 반면 소비자동향조사 내 주택가격전망지수 아파트 가격과의 동조성이 낮고, 시계열 전반에서 등락 방향이나 강도의 괴리가 크다. 이는 CSI가 설문 기반 기대심리를 반영하나, 단기 시장 반응을 포착하기에는 구조적 한계가 있음을 보여준다. 주택담보대출금리(int)와 소비자물가지수 전년동월비(cpi_yoy)는 2022년 이후 급등하며 아파트 가격 하락과 유사한 흐름을 보이지만, 시점상 대부분의 변곡점에서는 가격보다 다소 늦게 반응하는 후행적 특성이 나타났다. 이는 금리와 물가가 정책 결정이나 외생 충격에 의해 조절된 후, 일정한 시차를 두고 시장에 영향을 미치는 구조적 속성과 일치한다. 종합적으로, RSI는 다른 변수들과 달리 아파트 가격에 구조적으로 선행하며, 정책 변화나 외부 충격 이전에 민감하게 반응하는



〈그림 1〉 주요 변수의 시계열 추이(2016.01~2024.12)

특징을 보인다. 이러한 시계열 구조는 RSI가 단순한 감성 지표를 넘어 실질적인 선행지표로 기능할 가능성을 시사한다.

2) 주요 변수의 기술통계량 요약

각 변수의 평균, 표준편차, 최소값, 최대값, 왜도, 첨도를 산출한 결과는 <표 3>과 같다.

<표 3>에서 확인되듯, RSI는 표준편차가 180.53으로 비교적 크고, 첨도도 0.44 수준으로 정규분포를 벗어난 분포 특성이 일부 나타난다. 이는 시장 심리가 특정 시점에 급격히 확산되거나 수축되는 특성을 반영한다. 반면 CSI는 평균 104.89, 표준편차 15.51로 상대적으로 안정적이며, 왜도 -0.68과 첨도 0.57로 극단값 발생 가능성이 낮은 분포를 보인다. 금리(int)와 물가(cpi_yoy) 역시 2022년 이후 외생적 정책 충격에 따라 변동성이 일시적으로 확대된 것으로 나타난다.

3) 해석 및 분석 방향

이상의 시계열 추이와 기초통계량은 다음 장에서 수행될 회귀분석, Granger 인과성 검정, random

forest 분석, SHAP 해석 등의 실증분석 기반 데이터 특성을 사전에 점검하는 역할을 한다.

단, 변수 간 우열 판단은 통계적 설명력 비교와 원인성 검정을 통해 IV장에서 정량적으로 도출하며, 본 장에서는 해석적 판단을 유보한다.

4. 정상성 검정 및 공적분 관계

모형의 계량적 타당성을 확보하기 위해 단위근 검정(augmented Dickey-Fuller, ADF) 및 Johansen 공적분 검정을 실시하였다. ADF 검정 결과, 종속변수인 아파트 매매가격지수(apt)를 포함한 주요 변수(int, cpi_yoy, rsi)는 모두 수준에서 비정상(non-stationary) 시계열로 확인되었으며, 이에 따라 변수 간 장기적 균형관계의 존재 여부를 확인하고자 공적분 검정을 수행하였다(<표 4>).

Johansen 공적분 검정 결과, $r \leq 1$ 까지의 귀무가설이 기각되었으며, 이는 분석 대상 변수들 간에 두 개 이상의 공적분 벡터가 존재함을 시사한다(<표 5>).⁴⁾ 이러한 결과는 변수들 간에 장기적

<표 3> 주요 변수의 기술통계량 요약(2016.01~2024.12, 월별 기준)

지표	평균	표준편차	최소값	최대값	왜도	첨도
apt	92.12	6.46	85.80	106.28	1.05	-0.11
RSI	595.64	180.53	258.00	1182.0	0.38	0.44
CSI	104.89	15.51	61.00	132.0	-0.68	0.57
int	3.36	0.67	2.39	4.82	0.41	0.91
cpi_yoy	2.09	1.55	-0.43	6.32	0.81	0.00

주 : RSI, real estate sentiment index; CSI, consumer sentiment index.

4) Johansen 검정은 다변량 시계열 간 장기 균형관계를 파악하는 통계적 검정 방식이다.

〈표 4〉 주요 변수의 정상성 검정 결과(ADF test)

변수명	ADF 통계량	p-value	결론
apt	-2.1250	0.2346	비정상
int	-1.1807	0.6818	비정상
cpi_yoy	-1.1930	0.6766	비정상
rsi	-2.4340	0.1324	비정상

주 : 1) 유의수준 5% 기준. 모든 변수는 수준에서 비정상 시계열로 판단됨.
2) ADF, augmented Dickey-Fuller.

〈표 5〉 Johansen 공적분 검정 결과(Trace 통계량 기준)

Rank(r)	Trace 통계량	5% 임계값	기각 여부
r≤0	65.99	47.85	기각
r≤1	32.28	29.80	기각
r≤2	10.25	15.49	기각 안 됨
r≤3	2.76	3.84	기각 안 됨

주 : Johansen 공적분 검정은 5% 유의수준 기준. r≤1까지 기각 되어 2개 이상의 공적분 관계 존재가 확인됨.

으로 안정적인 균형 관계가 형성되어 있음을 의미하며, 장기 추세와 단기 충격이 혼재된 시계열 구조를 반영한 실증 분석이 필요함을 뒷받침한다. 본 연구에서는 이러한 공적분 구조를 전제로 회귀 분석과 랜덤포레스트 기반의 변수 해석을 병행하여 분석을 수행하였다.

Johansen 공적분 검정을 통해 아파트 매매가격지수(apt), 부동산 심리지수(RSI), 주택담보대출금리(int), 소비자물가지수 전년동월비(cpi_yoy) 간에는 두 개 이상의 공적분 관계가 존재함을 확인하였다. 이는 각 비정상 변수들이 장기적으로 균형 관계를 유지하고 있음을 시사하며, 특히 RSI 또한 거시경제 변수들과 함께 장기 정합성 구조에 포함된다는 점에서 감성 기반 심리지표의 구조적

신뢰성을 뒷받침하는 계량적 근거가 된다.

IV. 실증분석

1. 분석 개요 및 변수 설명

본 장에서는 전국 아파트 매매가격지수(apt)를 종속변수로 설정하고, 독립변수로는 감성 분석 기반 부동산 심리지수(RSI), 설문조사 기반 주택가격전망지수, 주택담보대출금리(int), 소비자물가지수 전년동월비(cpi_yoy)를 포함하여 실증 분석을 수행하였다. 분석에 사용된 자료는 2016년 1월부터 2024년 12월까지의 월간 시계열 데이터로 구성되었으며, 변수의 정의와 수집 경로는 III장에서 제시한 바 있다. 본 실증분석의 목적은 다음의 세 가지로 요약된다. 첫째, 본 연구자가 개발한 RSI와 기존의 설문 기반 소비자심리지수(CSI)의 설명력(R^2), 예측력(RMSE), 선행성(Granger 인과성)을 정량적으로 비교함으로써, 감성 기반 지표의 우수성을 실증적으로 검증한다. 둘째, 머신러닝 기법인 random forest와 설명가능 인공지능 해석기법인 SHAP을 통해 주요 변수들의 영향력, 기여도, 상호작용 구조를 정량적으로 파악한다. 셋째, 금리 급등기(2022년 전후)와 아파트 가격의 국면 전환 구간에서 RSI가 실제로 조기 반응했는지를 시계열적으로 분석함으로써, RSI의 정책 활용 가능성과 조기경보 도구로서의 타당성을 검증한다. 이러한 목적을 달성하기 위해, 실증분석은 다음과 같은 절차로 구성된다.

4.2절에서는 RSI 및 CSI 각각의 회귀모형을 비교하여 설명력과 예측력의 차이를 검토하고,

4.3절에서는 Granger 인과성 검정을 통해 두 지표의 선행성 차이를 분석한다.

4.4절과 4.5절에서는 random forest 및 SHAP을 통해 주요 변수의 상대적 중요도, 영향 방향, 상호작용 구조를 해석한다.

4.6절에서는 공적분 관계를 기반으로 한 회귀 모형의 설명력 및 예측력을 분석하고, 4.7절과 4.8절에서는 금리 충격기 및 아파트 가격 변곡점 시점에서 RSI의 반응 구조를 시계열적으로 분석한다.

이상의 분석은 단일 시점의 설명력에 국한하지 않고, 구조적 반응 양상과 정책적 활용 가능성을 종합적으로 평가하기 위한 계량적 접근이다. 특히, 기존 설문지표의 한계를 보완하는 RSI의 유효성을 실증적으로 검증하고, 정책 당국이 사전 대응 가능한 '심리 선행지표'로서의 활용성을 제시하는 것을 주요 분석의 목표로 삼는다.

2. 회귀모형 분석: Real Estate Sentiment Index vs Consumer Sentiment Index의 설명력 비교

부동산 심리지수(RSI)와 소비자심리지수(CSI)

가 아파트 매매가격지수에 미치는 영향을 단변량 회귀분석으로 비교하였다. 회귀모형은 각각 (식 3), (식 4)와 같이 구성된다.

$$\text{모형 A: } \text{apt}_t = \alpha_0 + \alpha_1 \cdot \text{RSI}_t + \varepsilon_t \quad (\text{식 3})$$

$$\text{모형 B: } \text{apt}_t = \beta_0 + \beta_1 \cdot \text{CSI}_t + \varepsilon_t \quad (\text{식 4})$$

두 지표의 설명력을 비교하기 위해 결정계수(R²), 평균제곱오차(RMSE), 평균절대오차(MAE)를 산출하였으며, 회귀계수의 유의성도 함께 확인하였다. <표 6>은 각 단변량 회귀모형의 분석 결과를 요약한 것이다.

RSI를 설명변수로 한 단변량 회귀모형의 결정계수(R²)는 0.4359로, CSI 모형의 0.0079 대비 압도적으로 높은 설명력을 보였다. 또한 예측오차 지표인 RMSE 및 MAE 역시 RSI 모형에서 각각 4.8274, 3.6377로 CSI 모형보다 현저히 낮게 나타났다. 회귀계수 유의성 또한 뚜렷한 차이를 보인다. RSI의 회귀계수(0.0236)는 유의수준 1%에서 통계적으로 유의(p<0.01)한 반면, CSI의 회귀계수는 음의 값(-0.0369)을 가지며 p값 0.3612로 통계적으로 유의하지 않다.

이러한 결과는 RSI가 단순히 수치상 높은 설명력을 지닐 뿐 아니라, 실제 예측 정확도와 통계적 신뢰성 면에서도 기존의 설문 기반 CSI보다 우수

<표 6> RSI 및 소비자심리지수(CSI) 기반 단변량 회귀 결과

회귀모형	R ²	RMSE	MAE	회귀계수(β)	t값	p값
apt~RSI	0.4359	4.8274	3.6377	0.0236	9.0497	0.0000
apt~CSI	0.0079	6.4019	5.1429	-0.0369	-0.9170	0.3612

주 : RSI, real estate sentiment index; CSI, consumer sentiment index; RMSE, root mean squared error; MAE, mean absolute error.

한 설명변수로 기능함을 시사한다.

〈그림 2〉는 두 모형에서 도출된 예측값과 실제 아파트 매매가격지수 간의 시계열 비교를 시각화한 것이다.

좌측의 RSI 모형은 전체 기간 동안 실제값과 유사한 흐름을 유지하며, 상승과 하락의 전환점에서 도 예측값이 빠르게 반응하고 수렴하는 양상을 보였다. 특히 2020년 중반 이후 급등기와 2022년 하락기에서 RSI는 실제값의 변화를 선형적으로 포착하는 모습을 확인할 수 있다.

반면 우측의 CSI 모형은 예측값의 변동성이 작고, 실제 추세와의 괴리가 지속적으로 발생하였다. 가격이 급변하는 구간에서도 예측값은 완만한 흐름을 유지하며, 전환점을 적시에 반영하지 못했다. 이는 CSI가 시장의 실제 심리 변화보다는 평균적 기대에 머무르기 쉬운 구조적 한계를 지닌 지표임을 시사한다.

이러한 비교는 RSI가 CSI보다 설명력과 예측력에서 모두 우수할 뿐 아니라, 시점별 충격에 대한 반응성도 높다는 점을 보여준다. 실시간 뉴스·

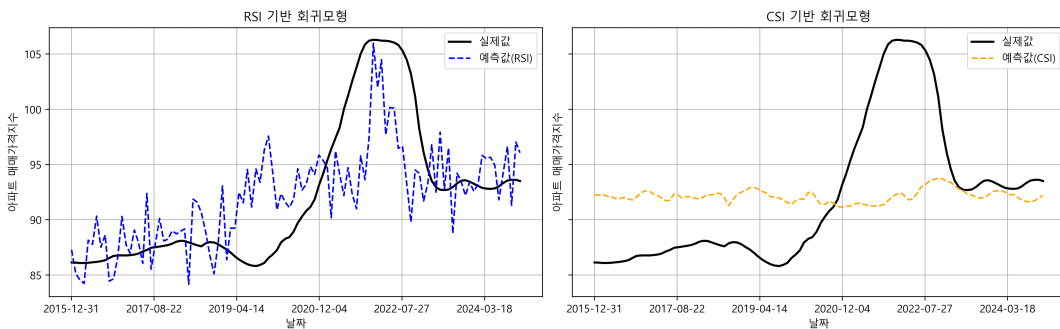
블로그 기반으로 구축된 RSI는 비정형 정보의 변동성과 정책 반응을 민감하게 반영할 수 있는 구조적 강점을 갖는다.

이상의 결과는 RSI가 단순한 설명 변수에 그치지 않고, 시장 반응을 조기 탐지할 수 있는 선행지표로 작용할 가능성을 제시한다. 다음 절에서는 Granger 인과성 검정을 통해 RSI가 아파트 가격의 원인 변수로 기능하는지를 통계적으로 검토한다.

3. Granger 인과성 검정: Real Estate Sentiment Index vs Consumer Sentiment Index의 원인성 비교

RSI와 CSI가 아파트 매매가격지수(apt)에 대해 선행적 영향을 미치는지를 Granger 인과성 검정을 분석하였다. 이는 변수 간 시계열적 정보 전달 구조를 통해 단순 설명력을 넘어선 ‘원인성(leading causality)’ 여부를 계량적으로 판단하는 데 유용한 분석 방법이다.

분석은 시차(lag) 1~3개월 범위에서 실시되었



주 : RSI, real estate sentiment index; CSI, consumer sentiment index.

〈그림 2〉 회귀예측값과 실제값 비교(좌: RSI 모형/우: CSI 모형)

으며, 각 시차별로 F-test 기준 p-value를 추출하여 유의성을 평가하였다. <표 7>은 RSI와 CSI가 아파트 가격에 대해 유의미한 선행 인과성을 갖는지를 비교한 결과이다.

검정 결과, CSI는 모든 시차 구간에서 아파트 가격에 대해 통계적으로 유의한 인과성을 보였다. 시차 1개월에서는 $p=0.0000$ 으로 매우 강한 선행성이 확인되었으며, 2개월($p=0.0001$), 3개월($p=0.0007$)에서도 모두 1% 유의수준에서 통계적으로 유의미하였다. 이는 CSI가 향후 주택가격 변화에 선행하는 기대심리의 구조를 안정적으로 반영하고 있음을 의미한다. 반면, RSI는 시차 1개월에서는 유의하지 않았으나, 2개월 시차에서 $p=0.0016$ 으로 1% 유의수준에서 강한 인과성이 확인되었으며, 3개월 시차에서도 $p=0.0176$ 으로 5% 유의수준에서 통계적으로 유의하였다. 이는 RSI가 단기(1개월) 반응보다는 일정 시차가 누적된 이후(2~3개월)에 아파트 가격에 선행적 영향을 미친다는 점을 보여준다.

이러한 결과는 CSI가 전 구간에서 안정적 선행성을 갖는 반면, RSI는 특정 구간에서만 선택적으로 인과 구조가 작동함을 시사한다. 특히 RSI는 실

시간 정서 반응을 기반으로 설계되었음에도 불구하고, 일정 시차 이후 아파트 가격에 구조적 영향을 미친다는 점에서 단순 설명력을 넘어서는 원인 변수로서의 잠재력을 갖추고 있음을 보여준다.

4. Random Forest 기반 변수 중요도 분석

아파트 매매가격지수(apt)에 영향을 미치는 주요 변수들의 상대적 설명력을 정량적으로 평가하기 위해 Random Forest 모형을 적용하였다. 분석에 사용된 독립변수는 감성 기반 부동산 심리지수(RSI), 주택담보대출금리(int), 소비자물가상승률 전년동월비(cpi_yoy)이며, 변수별 기여도는 random forest의 Gini importance 기준으로 산출하였다. <그림 3>은 각 변수의 설명력 기여비율을 요약한 결과이다. 소비자물가상승률(cpi_yoy)은 전체 설명력 중 73.5%를 차지하며 가장 높은 기여도를 보였고, 그 뒤를 금리(int: 17.7%), 부동산 심리지수(RSI; 8.8%)가 따랐다.

RSI는 전통적인 거시경제 변수인 금리보다 낮은 기여도를 나타냈으며, 전체 설명력의 10% 미

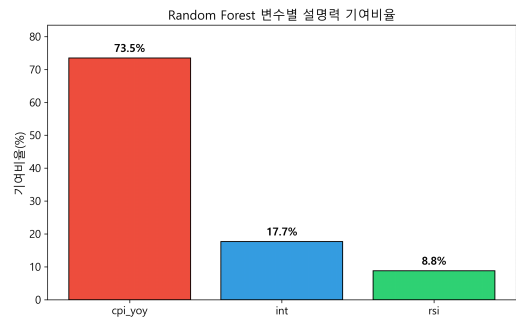
<표 7> Granger 인과성 검정 결과 요약(p-value기준)

시차	RSI → 아파트가격	CSI → 아파트가격
1개월	0.2077 ^{ns}	0.0000 ^{***}
2개월	0.0016 ^{***}	0.0001 ^{***}
3개월	0.0176 ^{**}	0.0007 ^{***}

주 : 1) 각 값은 Granger 인과성 검정의 p값이며, 위첨자는 유의수준을 나타낸다.

2) ^{ns} 유의하지 않음, ^{**} $p<0.05$, ^{***} $p<0.01$.

3) RSI, real estate sentiment index; CSI, consumer sentiment index.



<그림 3> 변수별 설명력 기여비율(Gini importance 기준)

만을 차지하였다.⁵⁾ 이는 비정형 감성 데이터 기반의 심리지수가 일정 수준의 설명력을 갖고는 있으나, 통화정책 변수나 물가 지표 대비 영향력이 제한적임을 의미한다.

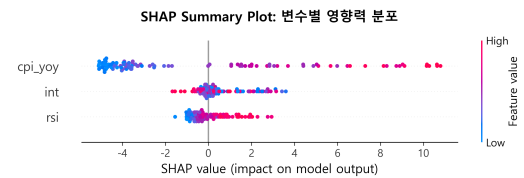
이상의 결과는 다음 절에서 수행할 SHAP 기반 변수 영향력 분석의 기초 자료로 활용되며, 각 변수의 영향 방향, 민감도, 상호작용 구조 등을 보다 정밀하게 해석하는 데 기반이 된다.

5. Shapley Additive Explanations 기반 변수 영향력 분석

본 절에서는 앞서 구축한 random forest 모형의 변수 중요도 결과를 기반으로, 변수별 영향력의 방향성과 민감도를 보다 정밀하게 해석하고자 설명가능 인공지능 기법인 SHAP를 보조적으로 활용하였다. SHAP은 예측 모델 자체가 아닌, 기존 모형에 대한 해석을 위한 도구로서 변수의 작동 구조를 시각화하고 정책적 해석 가능성을 확장하는 목적에 사용되었다. 특히, SHAP은 random forest 모형에서 도출된 변수별 중요도를 기반으로 영향 방향, 민감도, 시계열 구조 등을 시각적으로 분해함으로써, RSI와 가격 간 관계의 비선형성, 구조적 전환점, 완충 효과 등을 정량적으로 해석할 수 있는 기반을 제공한다. 이는 본 논문의 주제인 RSI의 설명력 검증 및 정책적 활용 가능성 해석에 있어 정합성을 유지하면서도 해석력을 보완하는 분석 도구로 기능한다.

분석은 전체 모형($apt \sim RSI + int + cpi_yoy$)을 기준으로 수행하였으며, 변수별 평균 SHAP 절대값을 비교한 결과는 <그림 4>에 요약하였다. 전체적으로 소비자물가상승률(cpi_yoy)이 가장 높은 평균 영향력을 기록하였고, 그다음은 금리(int), 부동산 심리지수(RSI) 순으로 나타났다.

SHAP summary plot의 시각적 분포를 기준으로 변수별 영향 구조를 보다 구체적으로 해석하면 다음과 같다. 첫째, cpi_yoy (소비자물가상승률)는 변수값이 높을수록 예측 아파트 가격에 정(+)의 영향을 미치는 것으로 나타났다. 붉은색 점들이 오른쪽에 집중되어 있으며, 파란색 점은 대체로 왼쪽에 위치하는 구조는 고물가일수록 예측 가격을 밀어올리는 역할을 함을 보여준다. 이는 인플레이션 국면에서 실물자산에 대한 선호가 가격에 반영되는 경향을 시사한다. 둘째, int (주택담보대출금리)는 SHAP 값이 대부분 0에 밀집해 있어 영향력 자체는 미미한 수준이나, 일부 고금리 구간(붉은 점)에서는 예측값을 낮추는 방향으로 작용하는 사례가 관찰된다. 이는 금리 인상이 수요 위축을 통해 가격 하방 압력으로 작용했음을



주 : SHAP, Shapley additive explanations.

<그림 4> SHAP summary plot: 변수별 영향력 분포

5) 본 연구는 random forest 분석 시 시드(seed)를 고정하지 않고 반복 실행하여 변수 중요도를 확인하였으며, RSI는 일부 실행에서는 금리보다 높은 기여도를 나타내기도 하였다. 그러나 재현성과 신뢰도를 고려해 보수적 기준(기여도 8.8%)을 채택하였으며, 관련 해석은 SHAP 분석을 통해 보완하였다.

암시한다. 셋째, RSI(부동산 심리지수)는 SHAP 값 분포가 -2에서 +4 구간에 걸쳐 있으며, 평균 영향력은 크지 않지만 예측 방향성과의 관계는 뚜렷하다. 낮은 RSI 값(파란 점)은 대체로 음의 SHAP 값을 보여 예측 가격 하락에 영향을 미치고, 높은 RSI 값(붉은 점)은 대부분 양의 SHAP 값을 보이며 가격 상승에 기여한다. 특히 0~+4 구간에 붉은 점들이 집중되어 있는 모습은, 심리지수가 높을수록 아파트 가격을 높게 예측하는 경향이 강하다는 점을 명확히 보여준다.

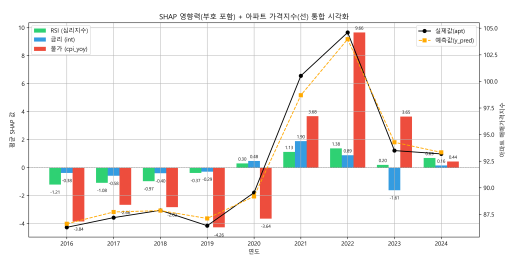
즉, RSI는 값의 크기에 따라 예측 영향 방향이 분명하게 바뀌는 구조를 가지며, 의미 있는 설명력을 지닌 변수로 확인된다. 이는 감성 기반 심리지수가 거시 변수와 함께 사용될 때 예측 모형의 해석력을 높일 수 있음을 시사한다.

한편, 위 결과는 전체 평균 기준이며, 변수의 연도별 구조적 기여도를 파악하기 위해 연도별 SHAP 중요도를 추가 분석하였다.

〈그림 5〉는 2016년부터 2024년까지 연도별 평균 SHAP 절대값을 시각화한 결과로, 아파트 매매가격지수의 실제값 및 예측값과 통합하여 변

수의 시기별 영향 구조를 직관적으로 확인할 수 있도록 구성하였다.

연도별 비교 결과, RSI는 금리(int)보다 더 많은 연도에서 아파트 매매가격 예측값에 강한 영향을 미쳤으며, 특히 2016~2019년에는 두 변수 모두 음의 SHAP 값을 보였지만 RSI의 절댓값이 매년 금리를 상회하였다. 이는 당시 시장에서는 금리보다 심리지표가 예측값 하락을 더 크게 설명하는 변수로 작동했음을 시사한다. 2020년 이후에는 구조적 전환이 나타났다. 금리와 RSI 모두 양의 SHAP 값으로 전환되어 예측값 상승에 기여하였으며, 2021년에는 금리(+1.90)의 영향력이 최고치를 기록했다. 같은 해 RSI도 +1.13으로 상승을 뒷받침했으나, 금리에 비해 영향력은 상대적으로 낮았다. 2022년에는 구조가 역전되어 RSI(+1.38)가 금리(+0.89)를 상회하였다. 이는 금리 수준이 더 높아졌음에도 불구하고, 예측모형 내 한계 기여도는 낮아진 반면, 심리지표가 보다 강한 설명력을 제공했음을 보여준다. 고금리·고물가 환경에서 심리 반응이 핵심적 요인으로 작동했음을 시사한다. 2023년에는 금리(-1.61)가 하방 압력으로 작용한 반면, RSI(+0.20)는 완만한 상승 기여를 유지하며 상반된 방향으로 작용했다. 이는 금리가 예측값을 끌어내리는 시점에도 RSI는 낙관 심리를 일정 수준 유지함으로써 예측값을 완충하는 역할을 했음을 보여준다. 2024년에는 양 변수 모두 영향력이 감소하였으나, RSI(+0.69)는 여전히 금리(+0.16)를 상회하였다. 전반적으로 RSI는 상승기와 하락기 모두에서 지속적인 설명력을 유지하며, 금리와는 다른 작동 경로를 보이는 독립적 설명 변수로 기능하고 있음을 확인할 수 있다.



주 : SHAP, Shapley additive explanations.

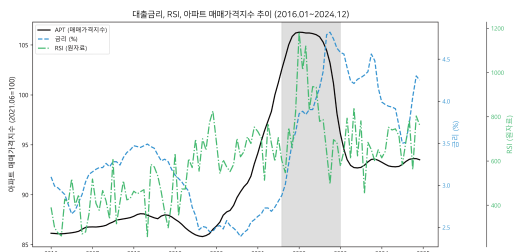
〈그림 5〉 연도별 변수별 평균 SHAP 중요도 (절대값 기준, 2016~2024년)

6. 금리 충격기 내 Real Estate Sentiment Index 반응 분석

주택담보대출 금리(int)의 급격한 인상이 아파트 매매가격에 미친 영향을 분석하고, 이 시기 부동산 심리지수(RSI)가 어떠한 반응을 보였는지를 함께 검토한다. 특히 RSI가 금리 충격에 대해 시장 심리를 얼마나 신속하고 선행적으로 반영하는지를 중심으로, 시차 구조와 예측 가능성을 중점적으로 평가한다.

분석 대상 기간은 2021년 8월부터 2023년 1월까지로 설정하였다. 이 시기는 예금은행 신규취급액 기준 주택담보대출 금리가 2.88%에서 4.58%까지 급등한 대표적인 고금리 충격기로, 자금 조달 비용이 급증하며 주택시장 전반의 매수심리가 위축된 시점이다.

〈그림 6〉은 주택담보대출 금리(int), RSI, 아파트 매매가격지수(apt)의 월별 시계열 변화를 동시에 시각화한 결과이며, 분석 구간은 회색 음영으로 표시하여 금리 충격기의 변동 흐름을 강조하였다. 시계열 흐름에 따르면, 대출금리는 2021년 8월 이후 점진적으로 상승하다가 2022년 중반부



〈그림 6〉 금리 인상기 아파트 가격, 심리지수, 금리의 동시 시계열 변화

터 급등세를 보였으며, 이는 금융 여건 변화가 본격적으로 시장에 영향을 미치기 시작한 시점으로 해석된다.

이 시기 RSI는 금리 상승 초기에는 완만한 상승세를 유지하며 낙관적 기대를 반영하였다. 이는 대출금리 상승에도 불구하고 시장 참여자들이 그 실질적 영향을 과소평가했거나, 자산 가격에 대한 기대심리를 유지하고 있었음을 시사한다. 그러나 2022년 8월을 전후로 금리 상승 폭이 누적되자 RSI는 급격한 하락세로 전환되었고, 이후 2023년 1월까지 연속적인 하락이 이어졌다. 이러한 급격한 심리 반전은 일정 수준 이상의 충격이 가해질 경우 시장 심리가 비선형적으로 반응하는 구조, 즉 감정 반응의 역치효과(threshold effect)가 작동했을 가능성을 시사한다.

한편 아파트 매매가격지수는 같은 시기에는 상대적으로 완만한 조정을 보이다가, RSI 하락 이후 약 2~3개월의 시차를 두고 낙폭이 본격화되는 흐름이 나타났다. 이는 RSI가 시장 심리의 약화를 아파트 가격보다 선행하여 반영하며, 실제 가격 조정의 초기 경고 신호로 기능할 수 있음을 보여준다. 대출금리 상승 → 심리 위축 → 가격 하락이라는 순차적 흐름은 정책 충격이 시장 심리를 매개로 실물 가격에 전달되는 자기강화적 반응 구조가 일부 작동했을 가능성을 암시한다.

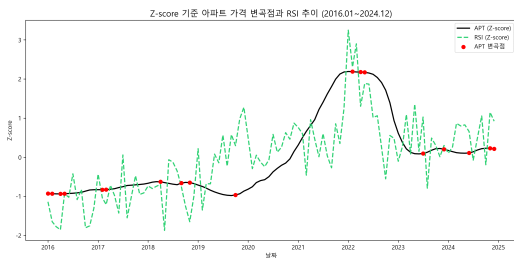
7. 아파트 가격 변곡점과 Real Estate Sentiment Index 추이 분석

아파트 매매가격지수(apt)의 주요 변곡점(turning point)에서 부동산 심리지수(RSI)가 선행적으로

반응하였는지를 실증적으로 분석하였다. 이를 통해 RSI의 시계열 민감성과 조기 경고 지표로서의 가능성을 확인하고자 하였다.

가격 변곡점은 전월 대비 가격지수의 방향성이 일정 기간 유지된 후 반전되는 시점을 기준으로 정의하였으며, 시계열 흐름상 연속된 상승 또는 하락세가 꺾이는 시점을 중심으로 탐지하였다.⁶⁾ 해당 기준에 따라 전체 18개의 변곡점이 도출되었고, 이 중 시장 국면 전환이 뚜렷하게 나타난 주요 변곡점 16개를 중심으로 분석을 수행하였다.

〈그림 7〉은 아파트 매매가격지수와 RSI를 Z-score로 표준화하여 동축상에 시각화한 결과로, 변곡점 시점에서 두 지표의 상대적 움직임과 선후 관계를 명확히 비교할 수 있도록 구성하였다. 분석 결과, 대부분의 변곡점 구간에서 RSI는 아파트 가격지수보다 1~2개월 먼저 방향을 전환하였으며, 특히 2022년 하락 전환기에는 RSI가 급락한 직후 가격지수가 후속적으로 하락하는 시차 구조



주 : RSI, real estate sentiment index.

〈그림 7〉 Z-score 기준 아파트 가격 추세 전환점과 RSI 추이(2016.01~2024.12)

가 명확하게 확인되었다.

이러한 결과는 RSI가 시장 심리의 전환을 조기에 반영하며, 실제 가격 조정에 선행하는 경향이 있음을 실증적으로 보여준다. 특히 RSI는 월 단위로 실시간 산출이 가능하다는 점에서, 전통적인 가격지표 대비 반응 민감도와 시계열 선행성이 높다는 강점을 갖는다. 반복적으로 확인된 이 같은 선행 반응 구조는 RSI가 가격 변곡점을 사전에 탐지할 수 있는 실효성 높은 심리지표임을 시사한다(〈표 8〉).

V. 결론 및 시사점

1. 연구 요약 및 주요 결과 정리

본 연구는 비정형 텍스트 기반 감성분석을 통해 실시간 부동산 심리지수(RSI)를 구축하고, 이를 전국 아파트 매매가격지수에 적용하여 기존의 설문조사 기반 소비자심리지수(CSI)와 비교·검증하였다. 분석에는 회귀분석, Granger 인과성 검정, Random Forest, SHAP 해석, 공적분 기반 회귀 모형, 시계열 비교 및 변곡점 분석 등을 활용하였으며, 주요 결과는 다음과 같다.

첫째, RSI는 가격에 대한 설명력 및 예측력 측면에서 기존 CSI를 압도하였다. 단변량 회귀분석 결과, RSI 모형의 결정계수(R^2)는 0.4359, CSI는 0.0079로, 수치상 55배 이상의 차이를 보였다.

6) 본 연구에서 정의한 '변곡점(turning point)'은 수학적 의미의 변곡점(inflexion point, 이계도함수=0)과는 구별되며, 실증분석 목적상 가격 추세가 상승에서 하락(또는 그 반대)으로 반전된 시점을 의미한다. 이는 SHAP 시계열 분석 결과를 바탕으로 추출된 국면 전환점이며, RSI와 가격 간 시차 구조를 계량적으로 비교하기 위한 지점으로 정의되었다.

〈표 8〉 아파트 가격 추세 전환점 시점과 RSI 반전 시점 간 시차 비교

No.	가격 변곡점 시점	RSI 반전시점	RSI 시차	RSI 선행여부
1	2016-04	2016-02	2개월 전	○
2	2016-05	2016-03	2개월 전	○
3	2017-02	2017-01	1개월 전	○
4	2017-03	2017-01	2개월 전	○
5	2018-04	2018-02	2개월 전	○
6	2018-09	2018-08	1개월 전	○
7	2018-11	2018-10	1개월 전	○
8	2019-10	2019-09	1개월 전	○
9	2022-02	2022-01	1개월 전	○
10	2022-04	2022-02	2개월 전	○
11	2022-05	2022-04	1개월 전	○
12	2023-07	2023-05	2개월 전	○
13	2023-12	2023-11	1개월 전	○
14	2024-06	2024-04	2개월 전	○
15	2024-11	2024-10	1개월 전	○
16	2024-12	2024-11	1개월 전	○

주 : RSI, real estate sentiment index.

예측오차 측면에서도 RSI의 RMSE(4.83)와 MAE(3.64)는 CSI보다 현저히 낮았다. 회귀계수 유의성 역시 RSI는 $p < 0.01$ 수준에서 통계적으로 유의하였으나, CSI는 유의하지 않았다.

둘째, RSI는 아파트 가격에 대한 인과성 구조에서는 제한적이거나, 금리 충격기 및 국면 전환기에는 선행적 반응을 명확히 나타냈다. Granger 인과성 검정 결과, RSI는 시차 1개월에서는 유의하지 않았으나, 시차 2개월($p=0.0016$)과 3개월($p=0.0176$)에서 아파트 가격에 대해 통계적으로 유의한 선행 인과성이 확인되었다. 이는 RSI가 단

기 반응보다는 일정 기간 누적된 심리 변화를 통해 가격 변동을 유도하는 구조임을 시사한다. 이러한 구조는 SHAP 시계열 해석 및 실제 시점 비교 분석 결과와도 일치하였다. 특히 2016년부터 2024년까지 총 16번의 주요 가격 변곡점 모두에서 RSI는 가격보다 1~2개월 앞서 반등하거나 급락하는 움직임을 보였으며, 이는 감성 기반 심리 지수가 시장 전환점에서 구조적으로 선행 반응하는 특성을 갖고 있음을 계량적으로 뒷받침한다.

셋째, RSI는 비선형적 영향 구조와 민감도를 지니며, 거시경제 변수들과 병행할 때도 유효한 설명력을 확보하였다. 본 연구는 이러한 구조적 특성을 보다 정밀하게 해석하고자, 앞서 구축한 random forest 모형의 예측 결과를 기반으로 설명가능 인공지능 기법인 SHAP을 보조적으로 활용하였다. SHAP 시각화 결과, RSI 값이 낮을수록 예측 가격 하락에, 높을수록 상승에 기여하는 경향이 명확히 나타났다. 이러한 결과는 RSI가 단지 설명 변수에 그치지 않고, 시장 심리의 방향성과 민감도를 정량적으로 파악할 수 있는 구조적 지표로 기능할 수 있음을 시사한다. 특히 SHAP은 예측 모델 자체가 아니라, 기존 random forest 모형의 내부 작동 원리를 시각적으로 해석하는 보완적 도구로 활용되었으며, RSI가 비선형적 기여 경로와 심리적 전환 구조를 형성하고 있음을 계량적으로 확인하는 데 기여하였다. 특히 RSI 값이 0을 초과하는 구간에서는 SHAP 값이 양의 방향으로 뚜렷하게 전환되며, 시장의 낙관 심리가 아파트 가격 예측값을 높이는 방향으로 작용하는 구조가 확인되었다. 한편, 연도별 SHAP 평균값 분석에서는 중요한 특징이 나타났다. 가격이 상승한

2020년과 2021년에는 금리가 예측값에 강하게 작용하였으며, 2023년에는 금리가 하방 압력으로 작용한 반면 RSI는 일정 수준의 완충 효과를 보였다. 반대로, 가격 조정기였던 2022년과 2024년에는 RSI가 금리를 상회하는 기여도를 보이며 예측값을 주도하였다. 특히 2023년에는 금리가 가격을 강하게 끌어내린 반면(RF 평균 -1.61), RSI는 소폭이지만 양의 기여도를 유지(+0.20)하며 예측값의 하락을 일부 상쇄하는 완충 효과를 나타냈다. 이처럼 금리는 정책 충격의 직접 반응을 설명하는 변수로, RSI는 시장 참여자의 감정 반응을 반영하여 예측값을 조정하거나 보완하는 변수로 기능한다. 결과적으로 RSI는 특정 시점에 따라 금리를 능가하는 설명력을 발휘하며, 정책 효과를 심리 경로를 통해 증폭 또는 완충시키는 핵심 변수임이 실증적으로 확인되었다. 다만, random forest 기준 Gini importance에서 RSI는 전체 설명력 중 8.8%로 금리(17.7%)나 물가(73.5%) 대비 낮았고, 이는 순수 예측변수로서 RSI의 독립적 영향력이 제한적임을 보여준다. 그러나 공적분 기반 회귀모형에서는 RSI 계수가 0.0134($p < 0.01$)로 유의하며, 장기적인 가격 수준에도 의미 있는 설명력을 갖는 것으로 나타났다.

넷째, RSI는 금리 급등기와 가격 변곡점에서 조기 반응하며, 실질적인 경고 신호로 기능할 수 있음을 실증적으로 확인하였다. 2021년 하반기부터 2023년 초까지의 금리 충격기에서는, 금리 인상 누적 이후 RSI가 2022년 8월을 기점으로 급락한 반면, 아파트 가격은 2~3개월 후 하락세로 전환되는 구조가 나타났다. 이와 같은 흐름은 금리 상승 → 심리 위축 → 가격 하락이라는 시차적

전달 구조를 나타내며, RSI가 초기 반응을 통해 시장 심리 변화의 단서를 제공할 수 있음을 시사한다. 한편, 전체 18개 가격 변곡점 중 16개의 주요 변곡점에서는 RSI가 평균 1.4개월 선행하여 방향을 전환하였다. 이는 RSI가 심리 구조의 변화 경고 및 정책 대응의 실마리를 사전에 제공할 수 있음을 의미한다. 또한 연도별 SHAP 값을 활용한 분석에서는, 예측값에 대한 금리와 RSI의 상대 기여도 변화를 시계열적으로 비교한 결과, 일부 시점에서는 RSI가 금리를 상회하며 예측값의 하락을 완충하는 구조가 확인되었다. 특히 2023년에는 금리가 예측값에 강한 하방 압력으로 작용한 반면, RSI는 소폭이나마 양의 기여를 유지하며 낙폭을 일부 상쇄하였다. 이와 같은 해석은 SHAP을 기존 random forest 모형의 보조 해석 도구로 활용하여, 변수별 영향력의 시기별 구조와 상호작용을 정량적으로 보완 분석한 결과이다.

긍정적 기사 확산 → RSI 상승 → 가격 반등 구조, 혹은 부정적 여론 유도 → RSI 하락 → 과열 진정 등의 정책적 활용 가능성이 존재한다.

이상의 결과는 RSI가 단순한 설명지표를 넘어, 시장 심리를 실시간으로 탐지하고, 조절 가능하며, 가격 변화를 사전 경보할 수 있는 정량지표로 기능할 수 있음을 실증적으로 확인한 것이다.

2. 정책적 시사점

본 연구는 감성 분석 기반으로 구축한 부동산 심리지수(RSI)가 아파트 가격 변화에 선행하여 반응하며, 정책 충격에 대한 감정 경로를 실시간으로 포착할 수 있음을 실증적으로 확인하였다.

특히 RSI는 2022년 금리 급등기와 2016년부터 2024년까지의 주요 가격 변곡점에서 모두 아파트 가격보다 1~2개월 앞서 반응하는 구조적 선형성이 확인되었으며, 이는 기존 CSI나 거시경제 변수로는 포착하기 어려운 시장 심리의 미세한 반전 시그널을 제공할 수 있다는 점에서 정책적으로 다음과 같은 시사점을 갖는다.

첫째, RSI를 기반으로 한 감정 기반 조기경보 체계를 구축할 필요가 있다. 금리 인상이나 대외 충격이 발생한 시점에서 RSI는 가격지수보다 빠르게 급락하거나 반등하는 반응을 나타냈다. 특히 2022년 하반기 금리 충격기와 2024년 상반기 가격 전환기에서는 RSI가 평균 1~2개월 앞서 방향성을 바꾸었으며, 이는 정책 당국이 시장 반전을 사전에 인지할 수 있는 조기경보 지표로서 RSI를 활용할 수 있음을 의미한다. 정부는 RSI가 급락하거나 급등하는 특정 임계값을 사전 경보 기준으로 설정하고, 규제 완화 또는 진입 억제 등의 정책 트리거로 활용하는 전략을 마련할 수 있다.

둘째, RSI는 심리조절형 넛지정책(nudge policy)의 설계 도구로 기능할 수 있다. RSI는 뉴스·블로그·온라인 커뮤니티 등 사회적 감성 반응에 민감하게 반응하는 특성이 있으며, 이는 정책당국이 정책 메시지 전달이나 여론 조정을 통해 RSI 흐름을 간접적으로 조절할 수 있다는 가능성을 시사한다. 예를 들어, 매수심리 회복이 필요한 시점에서는 공급 계획, 금리 안정, 미래 전망 등 긍정적 신호를 강화함으로써 RSI를 인위적으로 상승시킬 수 있고, 반대로 과열 방지 국면에서는 경고성 메시지나 리스크 안내를 통해 RSI를 하향 안정화할 수 있다. 이와 같은 전략은 금리·세금·공급과 같

은 전통적 정책보다 비용은 낮고 반응은 빠른 시장조정 수단으로서 의미를 가진다.

셋째, RSI를 공표 가능한 공식 정책지표로 제도화할 필요가 있다. RSI는 민간 데이터 기반으로 실시간 산출이 가능하고, 월 단위뿐 아니라 주간 단위 집계도 구현 가능하다는 점에서, 향후 한국 부동산원, 국토교통부, 한국은행 등 정책기관이 이를 정례 공표지표로 활용할 수 있는 체계를 마련할 필요가 있다. 특히, RSI가 금리·물가 등 거시지표와 함께 주간 시장 모니터링 지표군으로도 입될 경우, 단기 심리 반응을 실시간으로 감지하고 사전 개입 타이밍을 포착하는 정책 인프라로 기능할 수 있다.

요약하면, 본 연구가 제시한 RSI는 ① 시장 심리의 전환을 조기에 감지하고, ② 정책 메시지를 통해 간접 조절 가능하며, ③ 실시간 공표 및 정책 판단 지표로 활용 가능한 구조를 갖고 있다.

이는 기존 경제지표가 제공하지 못했던 '시장 감정 경로'에 대한 실질적 통제 가능성을 정책적 차원에서 열어주는 새로운 계량도구로서의 기여를 의미한다.

3. 연구의 한계 및 향후 과제

본 연구는 비정형 텍스트 데이터를 활용한 감성 분석 기반 부동산 심리지수(RSI)를 구축하고, 이를 아파트 가격 분석에 적용함으로써 기존 소비자심리지수(CSI) 대비 설명력, 선형성, 정책 활용 가능성 등을 실증적으로 검토하였다. 특히 회귀 모형, Granger 인과성, SHAP, 공적분 회귀모형, 시계열 구조 분석 등을 통해 RSI의 정량적 효과를

다각도로 검증하였다. 그러나 다음과 같은 한계가 존재하며, 이를 보완한 후속 연구를 통해 RSI의 신뢰성과 범용성을 더욱 확장할 수 있다.

첫째, RSI 산출 방식의 정밀화와 구조 고도화가 필요하다. 본 연구는 월간 뉴스·블로그의 제목만을 기반으로 감성 분석을 수행하였으며, (긍정-부정) 방식의 단순 차이값을 지표화하였다. 중립 문서는 logit confidence를 기준으로 극성 재분류했지만, 감성 강도(weight)나 문서 중요도는 반영하지 않았다. 향후에는 softmax confidence를 기반으로 한 감성 강도 가중합, 문서별 주제 중요도를 반영한 TF-IDF, 혹은 attention-weight 기반 감성평가 방식 도입을 통해 지수 설계의 정밀도와 신뢰도 향상이 가능할 것이다.

둘째, 텍스트 중심 RSI를 멀티모달 기반 심리 지수로 확장할 필요가 있다. 시장 심리는 단순 텍스트 외에도 이미지, 영상, 음성 등 다양한 비정형 정보를 통해 형성된다. 특히 유튜브, 인스타그램, 카카오프 등 영상 및 SNS 플랫폼은 감정 전파의 주된 경로로 기능하고 있으며, 댓글, 표정, 말투, 자막 등 감정 신호가 복합적으로 작용한다. 향후 연구에서는 영상 기반 STT 자막, 표정·음성 톤 분석 등을 포함한 멀티모달 감성 분석 기반 RSI를 구축함으로써, 심리의 자기강화 구조와 군집 확산 경로를 계량적으로 해석할 수 있을 것이다.

셋째, 공간적·문화적 확장에 따른 비교연구가 필요하다. 본 연구는 전국 단위 RSI와 아파트 매매가격지수의 관계에 초점을 맞추었으나, 실제 시장 반응은 지역별·계층별·문화적 요인에 따라 이질적으로 나타날 가능성이 존재한다. 향후에는 서울·수도권·광역시 등 지역 단위 RSI를 별도로

구축하거나, 일본·중국 등 이웃 국가의 RSI와 비교함으로써, 정책 수용성과 심리 반응 차이에 따른 가격 전이 구조, 글로벌 심리 파급 경로(cross-border spillover)를 실증적으로 분석하는 후속 연구가 요구된다.

넷째, RSI는 본질적으로 월 단위 시계열 구조를 갖는 지표이므로, 시계열 의존성과 동학(dynamics)을 분석할 수 있는 분석 모형이 필요하며, 향후 연구에서는 RNN이나 transformer와 같은 시계열 딥러닝 기반 모형의 적용도 고려할 수 있다.

다섯째, RSI는 동일한 수준의 점수라도 시기에 따라 상이한 시장 반응을 보일 수 있다. 예컨대, 2020~2022년에는 RSI가 급등과 함께 가격도 상승했지만, 2023년 이후에는 유사한 RSI 수준에서도 시장은 정체된 양상을 나타냈다. 이는 수도권 중심의 뉴스 과반영 현상과 지역별 시장 민감도 차이에서 기인할 수 있으며, RSI 해석 시 시기적·지역적 맥락을 함께 고려할 필요가 있다.

요약하면, 본 연구는 RSI의 설명력과 실시간 경보 가능성을 실증적으로 입증하였으나, 향후에는 ① 지수 산출 방식의 정교화, ② 멀티모달 확장 기반 감성 구조 해석, ③ 공간 및 문화 비교를 통한 일반화 검증, ④ RNN 등 시계열 딥러닝 기반 분석 모형의 도입을 통해 RSI의 이론적·정책적 범용성을 제고할 필요가 있다. ⑤ RSI 해석의 시기적·지역적 맥락을 고려한 정성적 보완 분석을 통해 RSI의 이론적·정책적 범용성을 제고할 필요가 있다.

이는 부동산 시장 분석이 더 이상 경제지표에만 의존할 수 없는 감성 기반 체계로 전환되고 있

음을 반영하며, RSI는 그 전환의 중심에 있는 계량도구로 자리매김할 수 있을 것이다.

ORCID

박순미 <https://orcid.org/0009-0006-3314-1039>

이성화 <https://orcid.org/0009-0000-5967-952X>

참고문헌

1. 김규석. (2025). *아파트 뉴스 심리지수의 개발 및 유용성 검증* [박사학위논문]. 서울대학교.
2. 김대원, 유정석. (2013). 주택가격에 대한 심리적 태도가 주택 매매 거래량에 미치는 영향 분석. *주택연구*, 21(2), 73-92.
3. 김동섭. (2020). *기계학습 모형의 설명가능성에 관한 연구: 미국 주택담보대출 자료를 중심으로* [박사학위논문]. 건국대학교.
4. 김수아, 권미주, 김현희. (2024). 생성 AI 기반 뉴스 감성 분석과 부동산 가격 예측: LSTM과 VAR모델의 적용. *정보처리학회 논문지*, 13(5), 209-216.
5. 김진석. (2024). *머신러닝 기반 아파트 가격지수 연구* [박사학위논문]. 서울대학교.
6. 박예나, 유석종. (2022). 자연어 질의 기반 부동산 환경 검색 시스템 구현. *한국정보기술학회논문지*, 20(11), 23-28.
7. 박재수. (2020). *주택시장 예측을 위한 부동산 감성지수 개발 연구: 뉴스와 방송 빅데이터에 대한 AI 기술 적용* [박사학위논문]. 강원대학교.
8. 박재수, 이재수. (2021). 부동산 감성지수의 주택가격 예측 유용성: 뉴스기사와 방송뉴스 빅데이터 활용 사례. *국토계획*, 56(4), 99-111.
9. 양건필, 전해정. (2021). 딥러닝 기반 언어모델을 이용한 언론과 주택가격의 상관관계 연구. *감정평가학논집*, 20(3), 109-134.
10. 유한수, 정재호. (2015). 주택시장에서의 매매가격 지수와 소비심리지수의 관계. *부동산연구*, 25(4), 49-61.
11. 이승규. (2019). *거시경제변수와 비정형데이터의 검색빈도가 아파트시장에 미치는 영향에 관한 연구* [박사학위논문]. 전주대학교.
12. 이정현. (2023). *인공지능을 이용한 주택가격 변동성 예측 모델 연구: 전통통계모형과 인공지능학습모형 융합을 중심으로* [박사학위논문]. 건국대학교.
13. 이종민. (2018). *비정형 빅데이터의 전세가격 예측 유용성 연구* [박사학위논문]. 강원대학교.
14. 이현정. (2023). *텍스트 마이닝을 이용한 부동산경기 순환기별 부동산 투자심리 특성 연구* [박사학위논문]. 동의대학교.
15. 이현정, 오윤경. (2024). *텍스트 마이닝을 이용한 부동산경기 순환기별 부동산 투자심리 특성 연구*. *토지주택연구*, 15(3), 113-127.
16. 조태진. (2014). 심리지수가 주택시장에 미치는 영향에 관한 연구. *주택연구*, 22(3), 25-48.
17. 주종용. (2023). *부동산정책에 따른 투자심리와 주택 가격 변화에 관한 연구* [박사학위논문]. 서울대학교.
18. 최성수. (2025). *부동산 정책지표의 효용성에 관한 연구* [박사학위논문]. 서울대학교.
19. Akerlof, G. A., & Shiller, R. J. (2009). *Animal spirits: How human psychology drives the economy, and why it matters for global capitalism*. Princeton University Press.
20. Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, 49(3), 307-343.
21. Chen, L., De, P., Hu, Y. J., & Hwang, B. H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *The*

- Review of Financial Studies*, 27(5), 1367–1403.
22. Hong, J., Choi, H., & Kim, W. S. (2020). A house price valuation based on the random forest approach: The mass appraisal of residential property in South Korea. *International Journal of Strategic Property Management*, 24(3), 140–152.
 23. Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–292.
 24. Kim, H., Park, J., & Kim, Y. (2021). Sentiment classification of Korean news articles using BERT. *Journal of Information and Communication*, 32(4), 55–67.
 25. Kim, Y., & Kim, Y. (2022). Explainable heat-related mortality with random forest and SHapley Additive exPlanations (SHAP) models. *Sustainable Cities and Society*, 79, 103677.
 26. Lastrapes, W. D. (2002). The real price of housing and money supply shocks: Time series evidence and theoretical simulations. *Journal of Housing Economics*, 11(1), 40–74.
 27. Petropoulos, A., & Siakoulis, V. (2021). Can central bank speeches predict financial market turbulence? Evidence from an adaptive NLP sentiment index analysis using XGBoost machine learning technique. *Central Bank Review*, 21(4), 141–153.
 28. Shayaa, S., Ainin, S., Jaafar, N. I., Zakaria, S. B., Phoong, S. W., Yeong, W. C., Al-Garadi, M. A., Muhammad, A., & Piprani, A. Z. (2018). Linking consumer confidence index and social media sentiment analysis. *Cogent Business & Management*, 5(1), 1509424.
 29. Wojarnik, G. (2022). Sentiment analysis towards listed companies in the energy sector on the example of the WIG-Energia index on the WSE. *Procedia Computer Science*, 207, 2059–2067.
 30. Wu, L. (2023). Interpretable prediction of heart disease based on random forest and SHAP. In *Eighth International Conference on Electronic Technology and Information Science* (pp. 356–360). SPIE.
 31. Zhang, L., & Liu, B. (2017). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press.

논문 접수일: 2025년 5월 11일

심사(수정)일: 2025년 7월 4일

게재 확정일: 2025년 7월 18일

국문초록

본 연구는 감성 분석 기반 비정형 데이터를 활용하여 온라인 실시간 부동산 심리지수(real estate sentiment index, RSI)를 구축하고, 이를 전국 아파트 매매가격지수와 연계해 기존 설문조사 기반 지표(consumer sentiment index, CSI)와 비교·분석하였다. 분석자료는 2016년 1월부터 2024년 12월까지의 월간 데이터를 활용하였으며, 감성분석은 KoBERT 기반 KR-FinBert-SC 모델을 적용하였다. RSI는 중립 문서를 softmax 확률값 기준으로 극성 재분류한 뒤, 긍·부정 문서 수 차이를 지표화해 산출하였다. 실증분석에는 단변량 회귀, Granger 인과성 검정, random forest, SHAP (Shapley additive explanations), 공격본 회귀모형을 적용하였다. RSI는 CSI 대비 설명력($R^2=0.4359$ vs 0.0079)과 예측력(root mean squared error 및 mean absolute error 기준)에서 우수했고, 시차 2~3개월에 유의한 인과성이 확인되었다. 또한 금리 충격기 및 16개 가격 변곡점 모두에서 RSI가 선행 반응하였다. SHAP 분석에서도 RSI는 낮은 값에서 음의, 높은 값에서 양의 영향력을 보이며 비선형성과 정량적 기여를 입증하였다. 이는 RSI 기반 조기경보체계 구축과 넛지정책 활용, 정책지표 제도화 가능성을 시사한다.

주제어 : 부동산 심리지수(real estate sentiment index, RSI), 감성분석, 랜덤포레스트, SHAP(Shapley additive explanations) 해석, 자기강화 메커니즘