

텍스트 마이닝을 이용한 부동산경기 순환기별 부동산 투자심리 특성 연구

A Study on the Characteristics of Real Estate Investment Sentiment by Real Estate Business Cycle Using Text Mining

이현정* · 오윤경**

Hyun-Jeong Lee* · Yun Kyung Oh**

Abstract

This study explores shifts in real estate investment sentiment using media reports from 2012 to 2022, segmenting the market dynamics into three distinct cycles based on housing and land transaction indices. Leveraging 54 BigKinds media sources, we investigate 3,387 headlines and 8,544 body texts using LDA topic modeling. The results show that the first cycle (2012-2015) centered on apartment pre-sales, where policy changes influenced sentiment but did not consistently affect investment decisions. The second cycle (2016-2018) was characterized by interest rate hikes and rising property prices in Seoul, resulting in significant fluctuations in transaction volumes. The third cycle (2019-2022) encompassed the effects of COVID-19, market instability, and policy failures, leading to distorted and weakened investment sentiment. Each cycle demonstrated that policies, interest rates, and economic events significantly shaped investor sentiment, as reflected in media reports.

Keywords: Real Estate Investment Sentiment, Media Reports, Topic Modeling, Market Instability, Investment Demand

1. 서론

1.1 연구의 배경 및 목적

부동산경기는 정책적 요인, 경제 상황 변화, 투자 심리 등 다양한 요인에 의해 영향을 받는다. 특히 투자심리는 언론보도를 통해 직접적으로 기사화되며, 일부 언론사는 투자심리를 자극하거나 우려를 표현하여 시장에 큰 영향을 미친다. 이로 인해 정확한 부동산경기 예측의 필요성이 더욱 부각되고 있다.

한편, 프롭테크 산업과 같은 첨단 정보 기술이 부

동산 분야에 접목되면서 정보 획득에 대한 관심이 확산되고 있다. 이러한 흐름 속에 2021년 국토교통부는 양질의 부동산 데이터를 개방하기 위한 노력의 일환으로 부동산 빅데이터 플랫폼 구축사업을 진행하고 있다. 이렇듯 공공을 비롯한 민간 부문에서도 빅데이터를 이용해 부동산 시장을 분석하고 예측하는 플랫폼 개발이 활발하게 구축되고 있다.

그런데, 이러한 기술은 부동산 시장 예측을 위한 정보로 다양한 지표와 함께 경험적 판단 보고서에 활용되지만, 부동산 시장은 부동산성으로 인해 정보

*동의대학교 부동산개발경영학 부교수(주저자: lhj1846@deu.ac.kr)

**동의대학교 부동산개발경영학 조교수(교신저자: oyk@deu.ac.kr)

비대칭성이 심각하다. 이러한 상황에서 언론 매체는 일반 투자자들에게 중요한 정보 출처가 될 수 있다. 이에 따라 투자심리에 영향을 미치는 언론보도의 패턴 분석은 중요하며 빅데이터 시대에는 대량의 부동산 정보에서 의미 있는 패턴을 발견하는 것은 중요한 과제가 된다. 언론 보도는 부동산 기대심리와 공포심리 등 투자자의 심리에 큰 영향을 미치는 매개체로서 이를 분석하는 것은 부동산 시장을 이해하기 위해 필수적인 과정이다.

본 연구는 '부동산 투자심리' 주제어를 중심으로 부동산경기 순환국면에 따른 언론기사의 변화 특성을 분석하고자 하였다. 우선 부동산경기 순환국면별로 언론보도 기사 제목과 본문의 보도내용 추세와 경향을 파악하여 핵심 이슈를 도출하였다. 즉, 뉴스 빅데이터를 텍스트 마이닝 기법을 통해 분석한 후 시기별 특성을 확인하였다. 둘째, 부동산 경기별 정부 정책 대응에 대한 심리적 반응이 언론기사에 어떻게 반영되는지를 분석하여, 정부의 정책 대응 근거와 투자자들의 투자 패턴 이해를 확인하였다. 셋째, 언론보도 기사 빅데이터를 통해 부동산 투자심리의 의미 있는 패턴을 도출하고, 당시 사회적 투자 인식을 이해하며, 순환국면별 공통점과 차이점을 비교하여 합리적인 투자 행동을 예측하고자 하였다.

이를 통해 부동산 시장의 정책적 이슈와 투자자 심리에 대한 이해를 높이고, 투자자들이 합리적인 의사결정을 할 수 있도록 돕는 것을 목표로 하였다.

1.2 연구의 범위 및 방법

본 연구는 '부동산 투자심리'를 주제로, 다양한 매체에 보도된 부동산 언론 기사를 통해 부동산경기 순환국면에 따라 나타난 변화 특성을 분석하였다.

연구의 시간적 범위는 2012년 1월부터 2022년 12월까지로 설정하여 전국 주택 거래량과 토지 거래량에 따라 부동산경기를 세분화하였다. 2012년부터 2015년까지 제1순환기, 2016년부터 2018년까지

제2순환기, 2019년부터 2022년까지 제3순환기로 나누어 분석을 진행하였다.

공간적 범위는 전국을 대상으로 하였으며 전국일간지, 경제일간지, 지역일간지, 방송사, 전문지를 포함한 다양한 언론매체의 보도 기사를 분석하였다.

내용적 범위는 '부동산 투자심리'를 주제로 한 언론기사로, 기사 제목과 본문을 중심으로 분석하였다. 빅데이터를 활용한 연구에서는 검색어 선정기준은 매우 중요하며, 검색어에 의해 수집되는 빅데이터가 달라지게 되면, 각 문서에 내포되어 있는 단어의 출현 빈도, 클러스터링 결과도 크게 달라진다. 따라서 본문 분석의 경우 면밀한 분석을 위해 부동산 투자심리와 함께 '전망' 키워드를 추가하여 내용적 범위에 포함하였다.

연구방법은 먼저, 빅카인즈의 언론보도 기사를 분석 대상으로 하여 '부동산 투자심리' 검색어로 수집된 기사 데이터를 활용하였다. 수집된 기사를 대상으로 형태소 분석과 데이터 전처리 작업을 실시하여, 불필요한 정보를 제거하고 분석에 적합한 형태로 정제하였다. 다음으로 각 순환국면별 단어 빈도와 역문서 빈도를 통해 중요도가 높은 단어를 선정하였다.

정제된 데이터는 NetMiner 4.5를 활용하여 토픽 모델링 분석을 실시하였다. LDA 기법을 적용하여 각 순환기별로 최적의 토픽수를 도출하고, 각 토픽의 주요 키워드와 출현 확률을 분석하였다. 이를 통해 부동산경기 순환기별로 언론보도 기사에서 나타난 투자심리의 변화를 파악하고, 시계열 변화를 분석하였다. 이를 바탕으로 본 연구의 절차는 Table 1의 순서대로 진행하였다.

2. 선행연구 검토

투자는 미래의 수입과 현재의 지출을 교환하는 것으로, 현재 소비에 대한 희생으로 미래에 발생할 수 있는 수익을 기대하는 것이라 할 수 있다. 이러한 투

Table 1. Research Procedure

Stage		Content
1. Analysis of previous research	Real Estate Investment Sentiment	Theoretical Review on Real Estate Investment and Investment Sentiment
	Text Mining Analysis	Theoretical Review on Real Estate Using Text Mining
↓		
2. Data Collection	Media Reports	Collection of News Big Data Using BigKinds
	Data Filtering	Removal of Duplicate Articles and Articles Irrelevant to the Research Purpose Final Compilation of Media Articles for Analysis
↓		
3. Data Cleaning	Morphological Analysis	Morphological Analysis and Noun Extraction
	Data Preprocessing	Creation of Dictionaries for Defined Terms, Synonyms, and Exclusion Words
↓		
4. Data Analysis and Validation	Keyword Frequency Analysis	Comparative Analysis of Key Keyword Frequencies
	Topic Modeling Analysis	Analysis of Major Keywords and Topic Proportions by Period Comparative Analysis of Time Series Changes for Each Topic by Period
	Expert Validation	Validation of Topic Modeling Analysis Results and Topic Names
↓		
5. Compilation of Results	Writing Analysis Results and Conclusions	Summary, Synthesis, and Implications of Analysis Results

자행위는 예측하기 쉽지 않은 미래 수익을 담보로 하고 있어 현재의 안정적이고 합리적인 소비를 감수해야 하는 위험이 따르게 된다(배기주, 2022).

한편, 부동산 투자특성은 입지, 시설, 환경 등 부동산 자체 특성과 수익성, 환금성 등의 투자 특성으로 일반 투자의 속성과 구별되는데(오철환, 2020), 다른 유형의 투자보다 현재의 소비를 장기간에 걸쳐 훨씬 더 억제해야 하는 만큼 투자자들의 보상 심리와 기대수익도 비례하여 커질 수밖에 없다.

김승환(2017)은 수익형 부동산 투자에서 제도환경과 경제환경이 중요한 요인임을 밝혔으며 권재홍

외(2020)는 재무적 관점과 입지 특성이 투자 의사결정에 중요하다고 분석하였다. 조한국(2021)은 서울 오피스빌딩 투자의 의사결정 요인을 분석을 통해 입지와 건물규모가 중요한 영향을 미친다고 결론지었다. 이렇듯 부동산 투자에서 정책, 경제환경, 입지 등 다양한 요인들의 역할에 대한 연구가 진행되었다.

투자심리는 곧 투자자의 심리이며 시장에서 미래 가격에 대한 투자자의 태도라 할 수 있다. 최기홍(2017)은 지나친 낙관적 과신 또는 비관적인 심리로 인하여 발생하는 결정들은 투자자로부터 비합리적인 수요를 일으킬 수 있기 때문에 투자자 심리는 자

산가격 결정의 중요한 요소로 보았다.

부동산 투자심리에 관한 연구는 다양한 요인들이 투자자의 심리와 의사결정에 어떻게 영향을 미치는지를 중심으로 진행되었다.

조경준(2014)은 주택 투자심리 지표를 직접 계산하여 글로벌 금융위기 이후 주택가격을 선행하는 투자심리에 대한 영향을 분석하였고, 이는 주택시장 변화를 심리적 관점에서 이해하는 데 중요한 시사점을 제공하였다. 허강수(2014)는 실버타운 소비자의 심리적 가치와 선호 속성을 연구하여, 소비자의 심리적 특성이 입지 선택에 중요한 영향을 미친다는 점을 밝혔으며, 김영식(2015)은 부동산 투자심리지수를 개발하여, 투자심리가 부동산 경기의 선행지표로 사용될 수 있음을 제안하였다.

김원식(2019)은 투자자의 자기확신과 균중심리가 투자성과 및 채투자 의도에 중요한 영향을 미친다는 점을 실증적으로 분석하였고, 김동환(2022)은 주거용 부동산 가격과 투자심리의 관계를 실증분석하여 맞춤형 투자자 심리지수 제시를 통해 투자심리지수의 부동산 시장 적용 효용성을 제고하였으며, 심재희(2022)는 주거만족과 부동산 투자심리가 주택구매의도에 미치는 영향을 통해 신규주택 분양의도와 기존주택구입 의도를 구분해 주택구매의도의 투자심리를 살펴보았다.

이와 같은 연구들을 통해 부동산 투자심리가 시장 변화, 정책, 경제적 요인과 밀접하게 연관되어 있음을 알 수 있으며, 투자자의 심리가 부동산 투자 결정에 중요한 역할을 할 수 있다는 경향성 또한 파악할 수 있다.

한편 빅데이터의 다양한 분석과 관련해 비정형 데이터의 급격한 증가로 데이터 마이닝, 텍스트 마이닝 등의 기법들이 주목받고 있다. 데이터 마이닝은 많은 자료로부터 내용 속에 숨겨진 연관성을 파악하여 새로운 가치의 정보를 만드는 분석법으로 다양한 의사결정 과정에 적용할 수 있다. 텍스트 마이

닝은 정보와의 상호 연계성을 파악하여 군집화시키거나 분류 또는 요약함으로써 빅데이터에 숨겨진 수많은 의미 있는 정보를 발견하는 분석법이다. 이는 인간 언어를 컴퓨터가 인식하여 자연어처리(NLP: Natural Language Processing) 기술에 기반한 비정형 텍스트 데이터의 유용한 정보를 추출하거나 가공하는 것을 목적으로 하고 있다.

텍스트 마이닝 분석 기법 중 대표적인 토픽 모델링은 다양한 문서 집합 속에 내재된 토픽, 즉 주제를 파악하고자 할 때 쓰이며 문서에 담긴 의미를 더 구체적이고 명확하게 보여줄 수 있다는 것이 토픽 모델링을 사용하는 이유라 할 수 있다(박상연 외, 2022).

부동산 분야와 관련된 연구에서도 텍스트 마이닝 분석들이 많이 사용되고 있다. 박종문 외(2018)은 도시재생에 대한 국제연구동향과 국내 정책결정자 담화를 토픽 모델링을 통해 분석하였으며, 박근혜 정부와 문재인 정부의 도시재생 정책에서 지역 및 환경적 지속가능성 고려가 미흡하다는 점을 명확히 밝혔다.

장몽현·김한수(2019)는 아파트 가격과 관련된 인터넷 뉴스 기사를 텍스트 마이닝과 토픽 모델링 기법으로 분석하였다. 2017년부터 2018년까지 데이터를 통해 재건축, 하락 등의 키워드가 아파트 매매가격 변동에 미치는 영향을 구체적으로 파악, 지역별 키워드와 매매가격의 상관관계를 도출하였다.

이태용(2020)은 투자심리가 부동산 시장과 주식 시장에 미치는 영향을 텍스트 마이닝, 감성분석, 벡터오차수정 모형을 통해 분석하였다. 부동산과 주식 시장의 감성지수를 통해 두 시장 간의 심리적 상호작용과 그 영향력을 구체적으로 밝히며, 부동산 시장의 긍정심리가 부동산 시장에 미치는 복합적 영향을 명확히 제시하였다.

김규리 외(2022)는 트위터 데이터를 수집하여 부동산 정책에 대한 대중의 감성 여론을 다이나믹 토픽 모델 분석으로 연구하였다. 2008년부터 2021년

까지 데이터를 통해 부동산 정책에 대한 긍정, 부정 의견의 변화를 구체적으로 파악하고 기초자료로 제시하고자 하였다.

따라서 다음과 같은 선행연구와의 차별성을 통해 부동산 투자심리 관련 연구를 보완하고자 하였다.

첫째, 부동산 투자 관련 연구에 처음으로 텍스트 마이닝을 활용한 빅데이터 분석을 시도했다. 기존 연구에서는 주로 텍스트 데이터 분석과 감성지수 도출을 통해 계량적 분석을 수행했으나, 본 연구는 부동산 투자심리 관련 언론보도 기사를 중심으로 토픽 모델링 분석을 사용했다.

둘째, 토픽 모델링 분석법을 활용하여 시기별 도출된 토픽의 비중과 시계열적 변화를 분석했다. 이를 통해 특정 시기에 어떤 토픽이 중요하게 다뤄졌는지 파악하고, 부동산 시장의 세부적 특성을 반영했다.

셋째, 연구의 시간적 범위를 2012년부터 2022년까지로 설정하고, 부동산 경기의 확장기와 수축기로 분석 기간을 세분화했다. 이를 통해 각 순환기별 부동산 투자심리에 대한 언론보도 기사의 변화 특성을 비교 분석했다.

언론보도 기사는 부동산 시장에 대한 전문가의 견해, 시장 반응, 정부 정책 발표, 각종 지표 등 다양한 정보를 포함하며, 실시간으로 최신 정보를 제공한다는 점에서 그 분석이 유의미하다. 본 연구는 이러한 언론보도의 변화 특성을 파악하여 보다 정교한 부동산 시장 분석 기반을 마련하고자 하였다.

3. 분석모형 및 자료

3.1 분석자료 구성

부동산 투자심리 관련 언론보도 기사는 한국언론진흥재단의 빅카인즈(BIGKinds) 빅데이터 시스템을 활용하여 수집하였다.

언론보도 기사 분석 대상은 11개의 중앙지와 28개의 지역지, 5개의 방송사, 2개의 전문지로서, 지역

전체 및 정치, 경제, 사회, 문화, 국제, 지역, 스포츠, IT 과학 통합분야를 모두 포함하여 총 54개의 언론매체를 최종 분석 대상으로 선정하였다.

부동산 투자심리에 관한 언론보도 기사 제목 분석을 위한 자료 수집을 위한 검색어로 ‘부동산 투자심리’로 하였으며 본문내용 분석은 ‘부동산 투자심리 AND 전망’으로 입력 후 검색하였다. 빅데이터를 활용한 연구에서 검색어 선정기준은 매우 중요하다. 검색어에 의해 수집되는 빅데이터가 달라지게 되면 각 문서에 내포되어 있는 단어의 출현 빈도, 클러스터링 결과도 크게 달라지기 때문에 ‘심리’ 키워드와 함께 ‘전망’ 키워드를 추가하여 검색하였다.

언론보도 기사의 경우 객관적 입장에서 작성되는 텍스트이지만 작성자가 제3자로서 사건이나 상황을 서술하며 특정 인터뷰나 의견 등을 인용하는 내용이 들어갈 수 있기 때문에 ‘심리’가 구체화되어 있는 본문내용을 분석하는 것은 필수적 수행과정이라 할 수 있다.

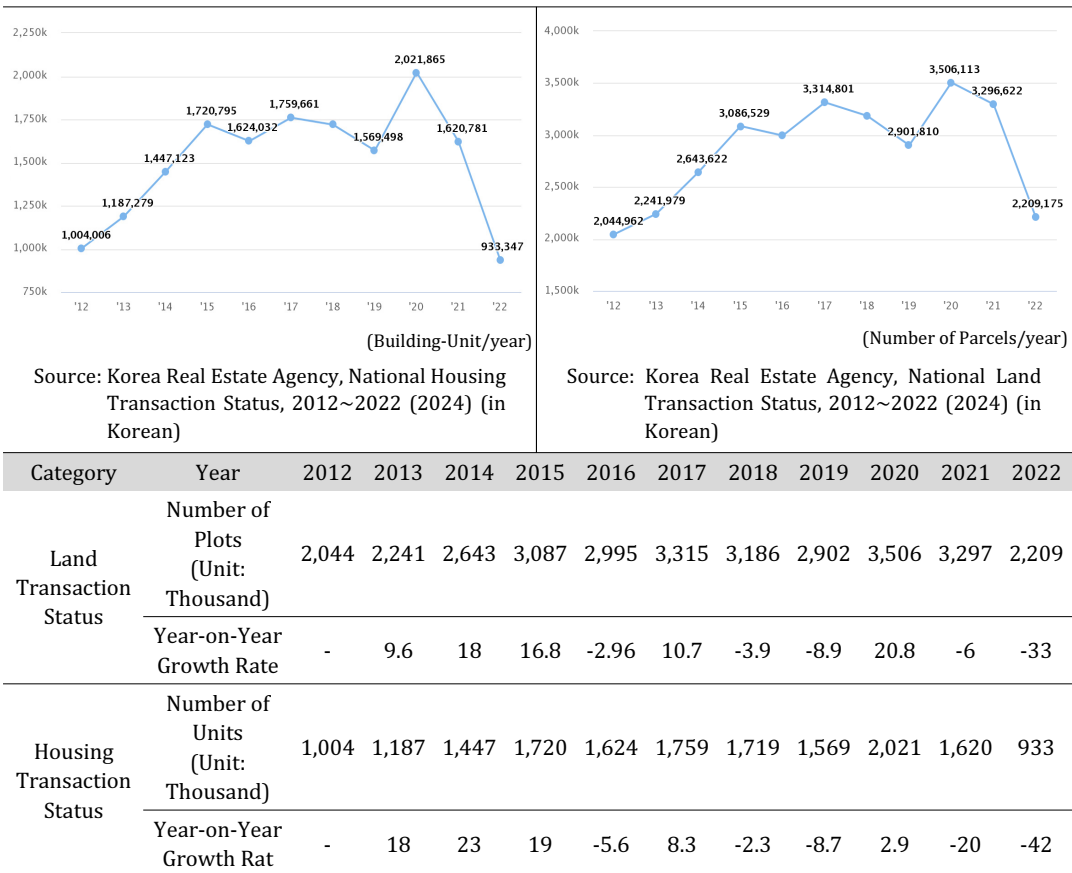
빅카인즈를 통해 수집한 최초 분석자료는 Table 2와 같이 총 18,734건의 언론기사이며, 이 가운데 중복기사와 광고, 단순 행사안내, 인물동정, 인사, 부고 등의 예외 기사를 제외한 최종 분석자료는 11,931건의 기사이다. 본 연구에서 수집한 11,931건의 언론 기사는 빅데이터 분석을 하기에 충분한 데이터양이라고 판단되어 분석을 실시하였다.

한편 선행연구를 참고하여 부동산경기를 인식하기 위한 지표로 전국 주택 거래량과 토지 거래량을 사용하였다. 이를 바탕으로 저점에서 정점에 도달하는 확장기, 정점에서 저점까지의 수축기로 나누었다. 확장기와 수축기를 같이 포함하고 있는 기간을 1개의 순환기로 하여, 2012년부터 코로나 팬데믹 이슈가 발생한 시기를 포함해 2022년까지 전체 기간을 제1순환기, 제2순환기, 제3순환기로 세분해 정리하였다. 2012년부터 2015년까지를 부동산경기 제1순환기, 2016년부터 2018년까지를 부동산경기 제2

Table 2. Number of Analyzed News Articles Related to Real Estate Investor Sentiment

Category		Real Estate Market 1st Cycle	Real Estate Market 2nd Cycle	Real Estate Market 3rd Cycle	Total
		[2012~2015]	[2016~2018]	[2019~2022]	
Original Collected Articles	Title	1,602	1,528	2,768	5,898
	Content	3,948	3,175	5,713	12,836
Excluded Articles	Title	715	497	1,299	2,511
	Content	1,115	997	2,180	4,292
Analyzed Articles	Title	887	1,031	1,469	3,387
	Content	2,833	2,178	3,533	8,544

Table 3. Real Estate Transaction Status and Year-on-Year Growth Rate



순환기, 2019년부터 2022년까지 부동산경기 제3순환기로 하였다. 해당 연도의 전국 주택 거래량과 토지 거래량의 변화 그래프와 전년대비 증감율은

Table 3과 같으며, 이를 근거로 살펴보았다.

본 연구에서는 자료 정제를 위해 빅카인즈를 활용하여 형태소 분석(morphological analysis)과 데이

터 전처리(preprocessing)를 실시하였다. 텍스트의 형태소 분석은 텍스트를 분석하기 위해 의미의 최소 단위인 형태소를 추출하는 작업이며, 데이터 전처리는 텍스트에서 분석에 불필요한 어구(예: 조사, 특수부호, 이메일 주소 등)를 제거하고 동일한 의미를 지니지만 다양하게 표현되는 용어를 정규화시키는 작업이다(박대민 외, 2015). 뉴스 빅데이터 시스템 빅카인즈는 형태소 분석과 데이터 전처리를 위해 Structural SVM(Support Vector Machine) 알고리즘을 활용하여 기사 텍스트에서 모든 명사 키워드를 자동추출하고 불용어를 제거한 후 사용자에게 엑셀 파일로 제공한다.

Structural SVM 알고리즘은 텍스트의 자연어 처리를 위한 기계학습 알고리즘으로, 한국어의 서술어 인식·분류에서 97.13%의 성능을 보인다(이창기의, 2014).

3.2 LDA기법을 적용한 자동 토픽분석

본 연구에서는 NetMiner 4.5를 활용하여 언론기사 텍스트에서 추출된 키워드를 행렬로 생성하여 LDA 분석 및 시각화를 진행하였으며, 특히 NetMiner의 Semantic Network 패키지와 연동된 ‘Topic Modeling’ 플러그인을 활용하여 LDA 기법을 적용한 자동 토픽 분석을 수행하였다. ‘Topic Modeling’ 플러그인은 빅데이터 분석에 널리 활용되는 프로그래밍 언어 파이썬(Python)으로 스크립트를 제작하여 기존의 복잡하고 반복적인 토픽 분석 프로세스를 일괄적으로 처리할 수 있게 제작되었다. 또한 NetMiner를 활용하여 언론기사 텍스트에서 추출된 키워드를 행렬로 생성하여 LDA 분석 및 시각화를 진행하였다.

토픽 분석은 토픽의 수를 미리 지정하여 분석을 수행하며, 본 연구에서는 과다한 언론기사 빅데이터 분석결과로 인한 해석의 한계를 극복하기 위해, 해석 가능성과 적합성을 기준으로 토픽의 수를 결정하

였다. 토픽 모델링의 최적화된 파라미터 값과 토픽 수를 결정하기 위해 NetMiner의 실루엣 계수를 통한 검증 실시하였다. 실루엣 계수는 토픽 내의 유사성과 토픽 간의 차이를 정량적으로 측정할 수 있는 지수로 -1에서 +1까지의 범위로, +1에 근접할수록 토픽 모델링이 최적화되었다는 것을 의미한다(Papadimitriou et al., 2020). 실루엣 계수의 계산은 모든 데이터 포인트 i 에 대해, $a(i)$ 는 같은 클러스터 내의 다른 모든 데이터 포인트와의 평균 거리로 계산하는데 이는 클러스터 내의 응집력을 측정하게 된다.

$$s(i) = \frac{(b(i) - a(i))}{(\max(a(i), b(i)))}$$

주: $b(i) - a(i)$ 는 두 군집 간 거리값이며, 이 값을 정규화 하기 위해 $\max(a(i), b(i))$ 값으로 나누어준다. 따라서 i 번째 데이터 포인트의 실루엣 계수 값 $s(i)$ 는 위와 같은 식으로 나타낼 수 있다.

$b(i)$ 는 i 와 가장 가까운 클러스터에 있는 모든 데이터 포인트와의 평균 거리로 계산하며, 이는 클러스터 간의 분리도를 측정하게 된다. 실루엣 계수를 활용하여 토픽 모델링의 최적화를 검증한 선행연구의 설정값을 반영하여, 알파 값을 0.1, 0.01, 0.05 베타 값을 0.01, 샘플링 반복횟수를 1,000회 반복하도록 설정하여 5개에서 7개가 가장 유의미하고 최적화된 토픽의 결과를 얻었으며, 토픽분석의 결과는 기사의 토픽 분포(topic distribution over document)와 토픽의 단어 분포(word distribution over topic)로 나타난다. 기사의 토픽 분포를 이용하여 가장 출현확률이 높은 토픽으로 분류하고, 출현확률(Probability)이 높은 단어 순으로 키워드를 추출하여 토픽별로 키워드를 정리하고 토픽명을 선정한다 다음 각 토픽이 차지하는 비중 및 시계열로 나타내었다.

4. 부동산 투자심리에 관한 텍스트 마이닝 분석

4.1 부동산경기 제1순환기 분석

4.1.1 언론기사 제목

부동산 투자심리 제목과 관련된 분석대상 언론기사 총 3,387건 중 제1순환기에 해당하는 887건의 언론기사 빅데이터를 대상으로 토픽 모델링 분석을 실시한 결과, Table 4와 같이 7개의 토픽이 도출되었으며, 토픽별 보도건수와 비중은 [토픽7] 155건, 17.47%, [토픽4] 147건, 16.57%, [토픽5] 136건, 15.33%, [토픽3] 122건, 13.75%, [토픽2] 119건, 13.41%, [토픽6] 109건, 12.28%, [토픽1] 99건, 11.16%의 순으로 나타났다. 제1순환기 토픽별 시계열 변화는 Fig. 1과 같다.

4.1.2 언론기사 본문

부동산 투자심리 본문과 관련된 분석대상 언론기사 총 8,544건 중 제1순환기에 해당하는 2,833건의 언론기사 빅데이터를 대상으로 토픽 모델링 분석을 실시한 결과, Table 5와 같이 5개의 토픽이 도출되었

Table 4. Number and Proportion of News Articles by Topic in the 1st Cycle

Category	Topic Name	Number of Articles	Proportion (Unit: %)
Topic7	Real Estate Policy and Apartment Market	155	17.47
Topic4	Effectiveness and Limitations of Real Estate Policy	147	16.57
Topic5	Market Fear and Expectations	136	15.33
Topic3	Trend of Decreasing Subscription Rates	122	13.75
Topic2	Interest Rate Cut Measures	119	13.41
Topic6	Discussions on Easing Real Estate Regulations	109	12.28
Topic1	Abolishment of Capital Gains Tax Increase and Easing of Loan Regulations	99	11.16

으며, 토픽별 보도 건수와 비중은 [토픽4] 931건, 32.86%, [토픽1] 665건, 23.47%, [토픽2] 569건, 20.08%, [토픽3] 381건, 13.45%, [토픽5] 287건,

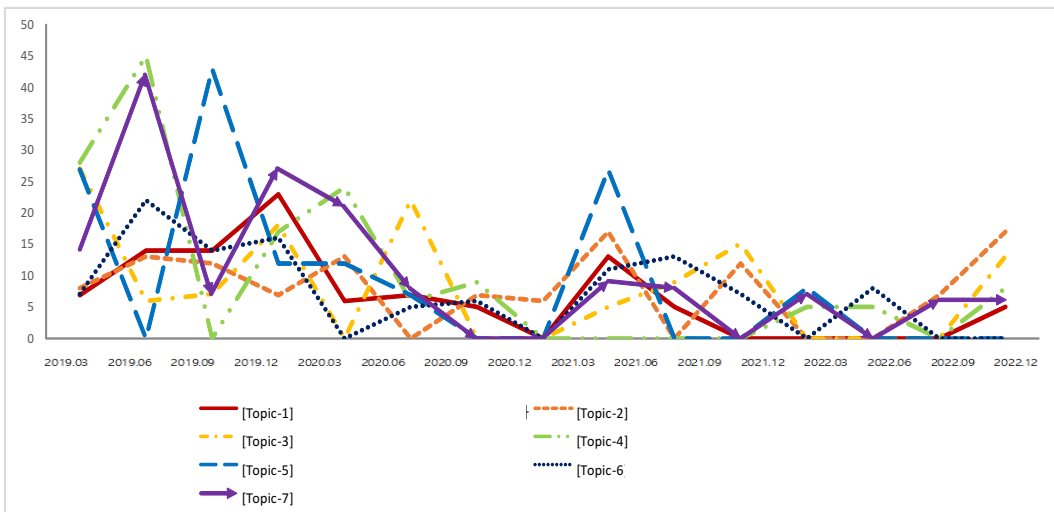


Fig. 1. Time Series Changes in News Article Titles by Topic in the 1st Cycle

Table 5. Number and Proportion of News Articles by Topic in the 1st Cycle Contents

Category	Topic Name	Number of Articles	Proportion (Unit: %)
Topic4	Surge in Jeonse Demand and Changes in Housing Investment Demand	931	32.86
Topic1	Real Estate Activation Policies such as Abolishment of Capital Gains Tax Increase and Interest Rate Cuts, and Investor Sentiment	665	23.47
Topic2	Increase in Demand for Income-Generating Real Estate Investments and Decrease in Rental Housing Investments	569	20.08
Topic3	Undersubscription Phenomenon and Decreased Real Estate Investment Expectations	381	13.45
Topic5	Rising Purchase Trend in Gangnam Area Due to Easing of Real Estate Regulations	287	10.13

10.13% 순으로 나타났다. 제1순환기 토픽별 시계열 변화는 Fig. 2와 같이 나타났다.

4.2 부동산경기 제2순환기 분석

4.2.1 언론기사 제목

부동산 투자심리 제목과 관련된 분석대상 언론기사 총 3,387건 중 제2순환기에 해당하는 1,031건의 언론기사 빅데이터를 대상으로 토픽 모델링 분석을 실시한 결과, Table 6과 같이 5개의 토픽이 도출되었으며, 토픽별 보도 건수와 비중은 [토픽2] 233건, 22.6%, [토픽1] 219건, 21.24%, [토픽3] 202건, 19.6%, [토픽 5] 200건, 19.4%, [토픽4] 177건, 17.16%, 순으로 나타났다. 제2순환기 토픽별 시계열 변화는 Fig. 3과 같다.

4.2.2 언론기사 본문

부동산 투자심리 본문과 관련된 분석대상 언론기사 총 8,544건 중 제2순환기에 해당하는 2,178건의 언론기사 빅데이터를 대상으로 토픽 모델링 분석을

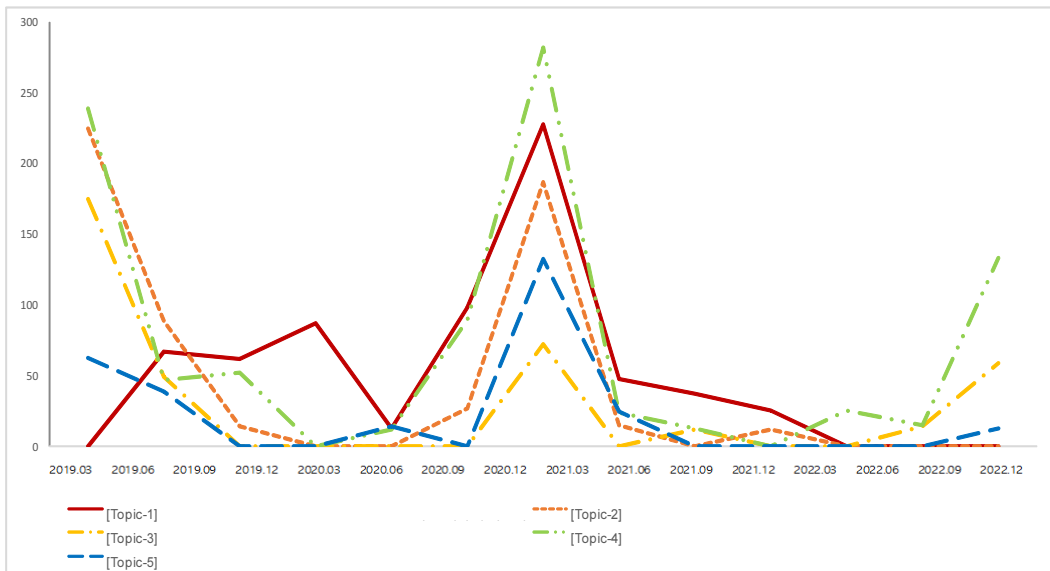


Fig. 2. Time Series Changes in News Article Contents by Topic in the 1st Cycle

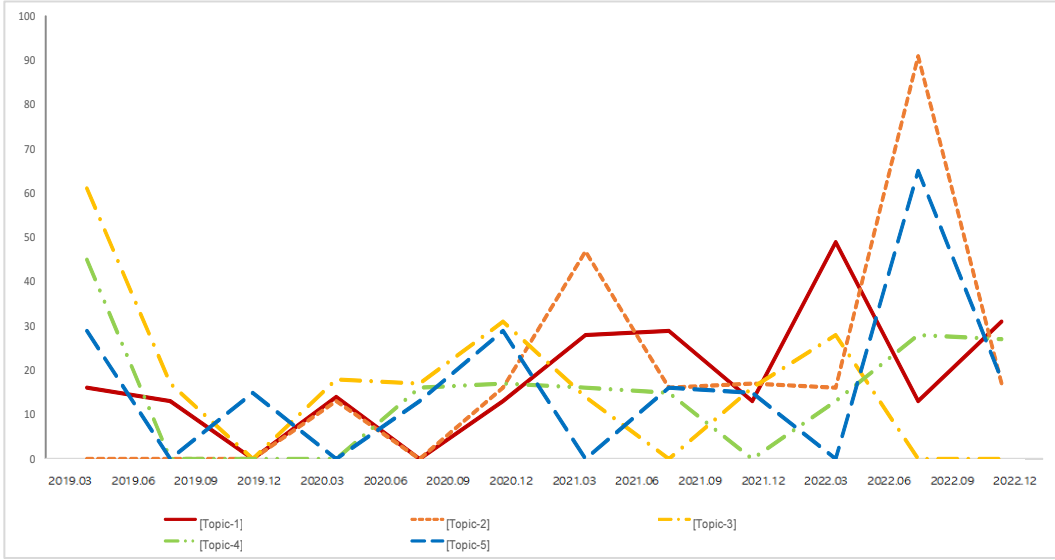


Fig. 3. Time Series Changes in News Article Titles by Topic in the 2nd Cycle

Table 6. Number and Proportion of News Articles by Topic in the 2nd Cycle Titles

Category	Topic Name	Number of Articles	Proportion (Unit: %)
Topic2	Capital Region Real Estate After Government Policy Announcements	233	22.6
Topic1	Interest Rate Hikes and Real Estate Market Contraction	219	21.24
Topic3	House Price Increase Forecast and Investor Sentiment	202	19.6
Topic5	Real Estate Overheating and Market Instability	200	19.4
Topic4	Regulatory Policies and Housing Market Contraction	177	17.16

Table 7. Number and Proportion of News Articles by Topic in the 2nd Cycle Contents

Category	Topic Name	Number of Articles	Proportion (Unit: %)
Topic3	Interest Rate Fluctuations and Differences in Regional Real Estate Investment Sentiment	1,008	46.28
Topic1	Real Estate Regulatory Policies and Market Stabilization Debate	449	20.62
Topic2	Volatility in the Pre-Sale Market and Investor Sentiment in Gangnam Redevelopment	302	13.87
Topic5	Interest in Income-Generating Real Estate and Investment Market Trends	235	10.79
Topic4	Increase in Land Investment and the Psychological Impact of Real Estate Booms	184	8.45

실시한 결과, Table 7과 같이 5개의 토픽이 도출되었으며, 토픽별 보도 기사건수 및 비중은 [토픽3] 1,008건, 46.28%, [토픽1] 449건, 20.62%, [토픽2] 302건, 13.87%, [토픽5] 235건, 10.79%, [토픽4]

184건, 8.45% 순으로 나타났다. 제2순환기 토픽별 시계열 변화는 Fig. 4와 같이 나타났다.

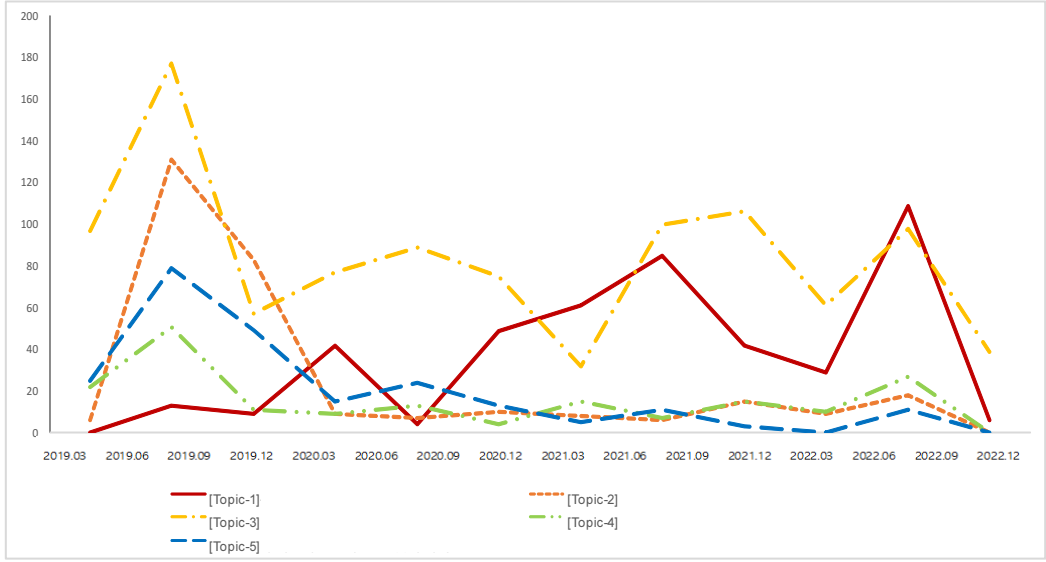


Fig. 4. Time Series Changes in News Article Contents by Topic in the 2nd Cycle

4.3 부동산경기 제3순환기 분석

4.3.1 언론기사 제목

부동산 투자심리 제목과 관련된 분석대상 언론기사 총 3,387건 중 제3순환기에 해당하는 1,469건의 언론보도기사 빅데이터를 대상으로 토픽 모델링 분석을 실시한 결과, Table 8과 같이 6개의 토픽이 도출되었으며, 토픽별 기사건수와 보도 비중은 [토픽5] 472건, 32.13%, [토픽1] 322건, 21.91%, [토픽3] 220건, 14.97%, [토픽6] 161건 10.95%, [토픽4] 150건, 10.21%, [토픽2] 144건, 9.8% 순으로 나타났다. 제3순환기 토픽별 시계열 변화는 Fig. 5와 같다.

4.3.2 언론기사 본문

부동산 투자심리 본문과 관련된 분석대상 언론기사 총 8,544건 중 제3순환기에 해당하는 3,533건의 언론기사 빅데이터를 대상으로 토픽 모델링 분석을 실시한 결과, Table 9와 같이 5개의 토픽이 도출되었으며, 토픽별 보도건수 및 비중은 [토픽2] 1,590건, 45%, [토픽3] 543건, 15.37%, [토픽1] 536건, 15.17%

Table 8. Number and Proportion of News Articles by Topic in the 3rd Cycle Titles

Category	Topic Name	Number of Articles	Proportion (Unit: %)
Topic5	Changes in Purchasing Trends Due to Interest Rates and Regulations	472	32.13
Topic1	Decline in the Real Estate Market	322	21.91
Topic3	Failures and Limitations of Real Estate Policies	220	14.97
Topic6	Movements in the Gangnam Real Estate Market	161	10.95
Topic4	Market Turbulence and Policy Effects	150	10.21
Topic2	Government Regulations and Reactions of Multi-Homeowners	144	9.8

[토픽5] 452건, 12.79%, [토픽4] 412건, 11.66%의 순으로 나타났다. 제3순환기 토픽별 시계열 변화는 Fig. 6과 같이 나타났다.

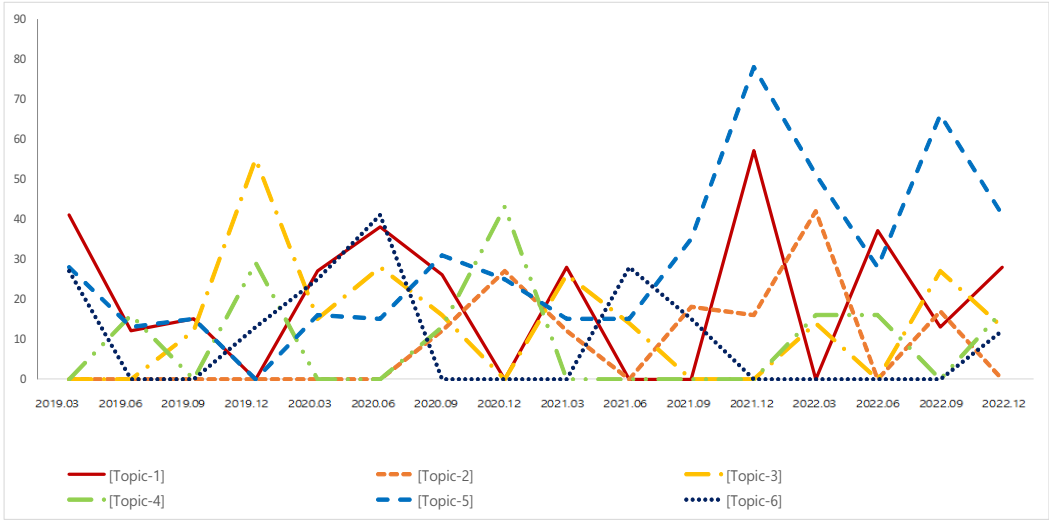


Fig. 5. Time Series Changes in News Article Titles by Topic in the 3rd Cycle

Table 9. Number and Proportion of News Articles by Topic in the 3rd Cycle Contents

Category	Topic Name	Number of Articles	Proportion (Unit: %)
Topic2	Impact of COVID-19 and Interest Rate Changes on Housing Consumer Sentiment in Capital and Regional Areas	1,590	45
Topic3	Overheating of Subscription Markets and Surge in Jeonse Prices in Non-Regulated Areas of the Capital Region	543	15.37
Topic1	Surge in Commercial Real Estate Investment and Reduction in Corporate Investment	536	15.17
Topic5	Expectations of Easing Reconstruction Regulations and Market Uncertainty Due to Government Policies	452	12.79
Topic4	Reactions of Multi-Homeowners to Property and Capital Gains Taxes and Intensified Market Anxiety	412	11.66

5. 분석결과 요약 및 시사점

5.1 분석결과 요약

제1순환기는 부동산 정책과 시장에 대한 변화와 우려에 대한 토픽이 중심을 이루고 있다. 2012년에는 유럽발 재정위기와 국내 가계부채 부담 증가 등으로 주택 매수심리가 극도로 위축된 가운데 5·10 대책 등 정부 부양책에도 불구하고 부동산 투자 우호적 정책의 효과에 대한 의심이나 우려감이 토픽에서 많이 나타나고 있다.

정책이 시행되었을 때의 불안감과 효과에 대한 기대감이 공존할 경우 가장 많은 언론기사가 제목으로 보도되었다. 이후 부동산 대책이 매매에 적용될 경우에는 오히려 투자심리의 방향이 정해지게 되고 투자 심리보다는 부동산 경제주체의 행위에 초점을 둔 기사 제목이 높은 비중으로 등장했다.

또한 부동산 투자심리를 반영하는 언론보도는 우려감, 불안감을 나타내고 있기도 하지만 이러한 투자심리가 정책시행 전후와 상관없이 부동산 투자 방향의 급격한 변화로 보도되지는 않는다는 점이 특징이었다.

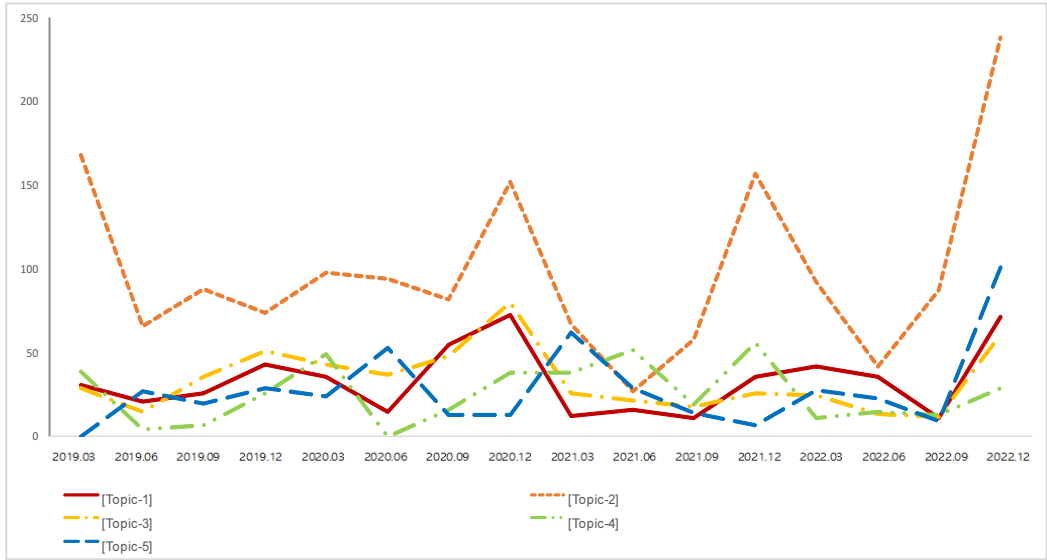


Fig. 6. Time Series Changes in News Article Contents by Topic in the 3rd Cycle

제2순환기는 수도권 부동산과 같은 투자 관심 지역에 대한 비중과 금리 인상과 부동산 시장의 위축, 집값 상승을 기대하는 투자 심리, 부동산 과열에 대한 경계심을 나타낸 토픽이 중심을 이루고 있다. 경기 보합특징의 제2순환기는 규제 활성화 방안보다 투자가치를 더 높일 수 있는 지역이나 투자 대상에 대한 관심과 같은 투자심리에 대한 보도경향이 더 큰 것으로 나타났다. 또한 높아진 집값에 대한 정부의 신혼부부 주거지원이나 임대주택 세제혜택 등 주거 복지와 관련한 정책도 시행되었지만 부동산 투자와 관련한 투자자 집단은 이러한 복지적 정책에는 별로 관심이 안보였고, 투자심리에 영향을 주지 못한 것으로 나타났다.

제3순환기에도 금리와 규제에 대한 관심은 여전히 높고, 부동산 하락세와 정책실패, 시장혼란과 관련한 토픽이 중심을 이루고 있다. 2019년 이후 보합특징을 보인 부동산 시장이 단기 확장되고 그 후 급격히 수축되는 특징이 토픽 속에서 드러났다. 제3순환기는 2019년의 종합규제 대책, 민간택지 분양가상한제 적용, 2020년의 금융규제, 다주택자 대상 종부세 증과세율 인상 등의 과열을 줄이기 위한 정부의 신규

정책에 대한 관심보다는 COVID-19의 부동산 경기에의 영향과 더불어 정부 정책에 대한 비판과 실망에 대한 투자심리가 더 많이 보도된 것으로 나타났다.

5.2 시사점

‘부동산 투자심리’ 주제어를 중심으로 부동산경기 순환국면에 따라 보도된 부동산 언론기사의 변화 특성에 대한 실증분석 결과를 토대로 나온 시사점은 다음과 같다.

첫째, 부동산 투자심리와 관련한 언론보도 내용 추세와 경향에 대한 부동산경기 순환국면별 변화 특성을 파악하였다.

제1순환기는 정책과 관련된 투자심리 내용의 언론보도가 특징으로 나타났다. 제2순환기는 투자가치를 높일 수 있는 지역이나 투자 대상에 대한 관심의 투자심리에 대한 내용이 보도경향으로 나타났다. 그리고 제3순환기는 COVID-19의 부동산 경기에 대한 영향, 정부 정책에 대한 비판과 실망의 투자심리 보도가 특징으로 나타났다.

둘째, 부동산경기 순환국면별 정책대응에 대한 심

리적 반응의 언론기사 반영 양상을 파악하였다.

부동산 정책 변화가 있는 경우 정책의 시장에 대한 효과나 영향이 발생하기 전 우려와 긴장감을 나타내는 기사가 많이 보도되었다. 본 연구를 통해 정책과 관련한 보도와 투자심리의 경향성을 파악할 수 있었다. 부동산 투자자는 실제 부동산경기와 본 연구결과의 경험적이고 구체적인 내용을 향후 투자판단에서 합리적 의사결정을 할 수 있는 자료로 활용할 수 있을 것이다.

셋째, 부동산 투자심리의 부동산경기 순환국면별 패턴 도출을 통한 합리적인 투자 행동을 예측하고 파악하였다.

전체 분석기간을 통한 언론기사의 공통점은 부동산경기과 투자심리에 미치는 영향요인 중 금리에 관련된 기사가 많이 보도되었다는 것이다. 부동산 투자자에게 금리 변동은 투자심리를 크게 좌우하는 행태로 나타나며, 부동산경기 확장국면이지만 투자자가 투자결정을 하게 되기까지 심리변화의 시간차에 의해 투자 타이밍이 어긋나기도 하였다. 결론적으로 금리는 부동산 투자심리에 상당한 영향을 주고 있으며 중요한 기사내용의 토픽임을 알 수 있었다.

본 연구에서 높은 비중으로 뽑힌 토픽들은 투자상품으로서 부동산을 객관적으로 판단할 수 있는 기준과 이에 따른 투자심리의 영향관계에 관한 내용보다는 투자 열풍, 정부 규제, 또는 개발호재 등과 같은 언론기사 보도를 통한 투자심리의 반응을 보여주었다는 것이다. 투자자는 투자 결정시 금리나 정책에 관한 투자심리 변화와 같은 언론보도를 접하더라도 부동산경기의 확장기인지 수축기인지 등 순환국면에 대한 판단 후 투자를 고려해야 할 것이다.

참고문헌

- 권재홍·이재원·이상엽(2020), “주거용 부동산 개발을 위한 투자자 관점에 따른 의사결정 요인에 관한 연구”, 「한국건설관리학회 논문집」, 21(5): 29~37.
- 김규리·오찬희·주영준(2022), “부동산 정책 관련 트위터 게시물 분석을 통한 대중 여론 이해”, 「한국문헌정보학회지」, 56(3): 47~72.
- 김동환(2022), “국내 주거용 부동산시장과 투자자심리의 관계 연구”, 박사학위논문, 서울벤처대학교.
- 김승환(2017), “수익형 부동산 투자가 은퇴 후 삶의 질에 미치는 영향”, 박사학위논문, 한양대학교.
- 김영식(2015), “부동산 투자심리지수 개발에 관한 연구”, 박사학위논문, 동의대학교.
- 김원식(2019), “부동산 투자자의 심리요인이 투자성과 및 재투자 의도에 미치는 영향”, 박사학위논문, 동의대학교.
- 박대민·김선호·백영민(2015), 「뉴스 빅데이터 분석 시스템 연구」, 서울: 한국언론진흥재단.
- 박상언·강주영·정석찬(2022), “텍스트 마이닝 기초”, 「파이썬 텍스트 마이닝 완벽 가이드」, 경기: 위키북스.
- 박종문·김지혜·윤순진(2018), “도시재생에 대한 국제 연구 동향과 국내 정책결정자 담화 분석=의미 연결망 분석과 토픽”, 「도시 행정 학보」, 31(2): 31~59.
- 배기주(2022), “부동산 간접투자상품 유형별 재투자 의도에 관한 연구: 수요자 관점을 중심으로”, 박사학위논문, 서울벤처대학교.
- 심재희(2022), “주거만족과 부동산투자심리가 주택구매의도에 미치는 영향에 관한 연구”, 박사학위논문, 서울벤처대학교.
- 오철환(2020), “정부의 부동산 정책과 시장상황이 부동산 투자에 미치는 영향: 부산지역 부동산시장을 중심으로”, 박사학위논문, 강남대학교.
- 이창기·임수중·김현기(2014), “Structural SVM 기반의 한국어 의미역 결정”, 「정보과학회논문지」, 42(2): 220~226.
- 이태용(2020), “투자심리가 부동산시장과 주식시장에 미치는 영향: 비정형데이터와 벡터오차수정모형”, 박사학위논문, 강남대학교.
- 장몽현·김한수(2019), “텍스트 마이닝을 활용한 주택가격 변동에 관한 연구”, 「한국주거학회논문집」, 30(2): 35~42.
- 조경준(2014), “주택시장과 주택금융에 대한 동태적 분석”, 박사학위논문, 한성대학교.
- 조한국(2021), “상업용 부동산 투자의 의사결정 요인

- 에 관한 연구”, 박사학위논문, 부산대학교.
18. 최기홍(2017), “투자자 심리가 한국 주식시장에 미치는 영향”, 박사학위논문, 부산대학교.
 19. 허강수(2014), “소비자의 심리적 가치에 따른 실버타운 선택”, 박사학위논문, 강원대학교.
 20. Papadimitriou, C. H., P. Raghavan, H. Tamaki and S. Vempala (2000), “Latent Semantic Indexing: A Probabilistic Analysis”, *Journal of Computer and System Sciences*, 2(61):217~235.
 21. 한국언론진흥재단, <https://www.bigkinds.or.kr>.
 22. 한국부동산원, “부동산통계정보시스템”, 2024.8.1 읽음. https://www.reb.or.kr/r-one/portal/stat/easyStatPage/A_2024_00642.do.
Korea Real Estate Board, “Real Estate Statistics Information System”, Accessed August 1, 2024. https://www.reb.or.kr/r-one/portal/stat/easyStatPage/A_2024_00642.do. (in Korean)

요약

부동산 시장에서 언론매체는 투자정보의 주요 출처로, 투자심리 변화에 큰 영향을 미친다. 본 연구는 2012년 1월부터 2022년 12월까지 전국 주택 및 토지 거래량 지표를 기준으로 부동산경기를 3개의 순환기로 나누어 분석하였다. 빅카인즈에서 제공하는 54개의 언론매체로부터 ‘부동산 투자심리’ 검색어로 수집된 기사 중 제목 3,387건, 본문 8,544건을 대상으로 주요 단어를 빈도 분석하고, LDA 기법을 적용하여 토픽 모델링 분석을 수행하였다. 분석 결과, 제1순환기(2012-2015)에는 아파트 분양시장 관련 기사가 주를 이루며, 실수요와 투자 수요의 기대에 대한 보도가 많았다. 정책 변화에 대한 기대가 투자심리에 영향을 미쳤으나, 실제 투자로 이어지지 않는 경우도 많았다. 제2순환기(2016-2018)에는 금리 인상 및 집값 상승에 대한 기사가 많았으며, 수도권 및 강남지역의 부동산 가격 상승 기대가 주요 투자심리를 형성했다. 금리 변화와 연계된 부동산 거래량 변화가 많이 보도되었다. 제3순환기(2019-2022)에는 COVID-19 위기와 부동산 시장의 불안과 양극화, 정부 정책의 실패로 인한 투자심리의 왜곡과 쏠림 현상을 비판하는 내용이 많았다. 부동산 시장의 혼란과 가격 하락, 실질적인 거래 감소가 언론에 많이 보도되었으며, 이는 투자심리의 위축으로 이어졌다. 요약하면, 각 순환기마다 부동산 정책, 금리, 경제적 사건들이 투자심리에 중요한 영향을 미쳤으며, 언론보도는 이러한 심리 변화를 반영하고 있는 것으로 나타났다.

주제어: 부동산 투자심리, 언론보도, 토픽 모델링, 시장 불안정성, 투자 수요