buildings

MDPI

Article

Interpretable Machine Learning for Classifying Expansion and Deceleration Regimes in the U.S. Housing Market Using Construction Cost and Supply Indicators

Minsoo Baek [1,*] and Jung-Hyun Lee [2]

[1] Department of Construction Management, Kennesaw State University, Marietta, GA 30060, USA
[2] Department of Architecture and Construction Management, SUNY Farmingdale State College, Farmingdale, NY 11735, USA; leej15@farmingdale.edu
* Correspondence: mbaek@kennesaw.edu

Abstract

Housing market research has traditionally emphasized forecasting continuous price levels, often overlooking the discrete regime shifts that more directly capture cyclical risk and market turning points relevant to construction planning and investment decisions. This study develops an interpretable machine learning framework to forecast monthly U.S. housing market expansion and deceleration regimes one month ahead by integrating macroeconomic, financial, and construction-related indicators with national housing price data spanning January 1993 through March 2025. Feature selection and hyperparameter tuning are conducted entirely within the training sample using time-series cross-validation, ensuring that all reported performance metrics reflect genuine out-of-sample generalization. Recursive feature elimination combined with variance inflation factor screening yields a compact, seven-variable predictor set, with no macro-financial variable contributing an incremental discriminatory signal. Five classification models spanning linear and tree-based ensemble families are benchmarked under a strict temporal 80%:20% train–test split. Tree-based ensemble models consistently outperform the linear baseline, with Random Forest achieving the highest holdout AUC (0.932) and the most consistent deceleration detection across nested cross-validation folds. To ensure methodological transparency, explainable AI techniques, including SHapley Additive exPlanations, partial dependence, and individual conditional expectation analyses, are employed to interpret both global and local predictive associations underlying regime classification. Construction-related indicators, particularly Construction Put in Place and Building Permits at short lag horizons, emerge as the dominant supply-side predictive signals, outperforming macro-financial variables in regime discrimination by a margin of 0.184 in training CV AUC. By shifting the analytical focus from price forecasting to one-month-ahead regime prediction and integrating predictive accuracy with economic interpretability, this study provides a transparent and scalable framework for monitoring housing market cycles with direct applications to construction risk management, procurement timing, and project planning.

Keywords: housing price index; expansion and deceleration regimes; machine learning; explainable AI; SHAP; construction market dynamics; feature selection

Check for updates

Academic Editor: Jorge Lopes

Received: 30 June 2026
Revised: 20 July 2026
Accepted: 27 July 2026
Published: 28 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The housing market is a cornerstone of economic activity, serving as both a primary asset for households and a major driver of national wealth [1]. Its dynamics are closely

intertwined with the broader economy, influencing consumer spending, investment behavior, and financial stability. In the United States (U.S.), housing-related activity accounts for approximately 18% of gross domestic product (GDP), reflecting the combined effects of residential investment and housing-driven consumption [2]. As a result, fluctuations in housing markets extend well beyond real estate transactions, directly affecting household wealth, budget constraints, and financial security, with broader implications for macroeconomic stability [3]. For construction firms, developers, and material suppliers, these fluctuations translate directly into project pipelines, procurement costs, labor demand, and capital exposure, making the identification of cyclical market states a practical operational necessity rather than an academic abstraction [4].

Historically, periods of heightened housing market volatility have both reflected and triggered broader financial disruptions, underscoring the importance of identifying cyclical regime shifts rather than focusing solely on average price trends [5]. Despite the centrality of expansion and deceleration dynamics in housing markets, existing approaches to modeling regime shifts have largely relied on econometric frameworks, such as Markov-switching models [6], threshold autoregressive models [7], and smooth-transition models [8]. These methods have made substantial contributions by formally characterizing regime dependence and nonlinear dynamics within parsimonious, theory-consistent structures. However, their practical application often requires strong distributional assumptions, a limited number of regimes, and predefined functional forms, which restrict flexibility when modeling complex interactions among high-dimensional macroeconomic, housing, and construction-related variables. Moreover, these frameworks rarely incorporate construction-sector indicators, such as material cost indices, permitting activity, or inventory pipelines, despite the central role of supply-side dynamics in housing cycles.

Recent advances in machine learning offer a complementary, data-driven alternative that relaxes these structural constraints. Tree-based ensemble models have been shown to effectively capture nonlinearities, threshold effects, and higher-order interactions in housing and financial markets, particularly when relationships are driven by multiple interacting predictors and short-term dynamics [9–11]. While machine learning approaches do not replace econometric regime-switching models, they provide a flexible framework for identifying regime patterns and assessing cyclical risk without pre-specifying the mechanisms of regime transitions. Importantly, integrating Explainable Artificial Intelligence (XAI) techniques, such as SHapley Additive exPlanations (SHAP), enables the recovery of economically interpretable insights from these models, addressing concerns about transparency and theoretical coherence [12,13]. To date, however, both machine learning and econometric applications in housing research have concentrated almost exclusively on property-level price prediction or macro-level price forecasting; the classification of market-level cyclical regimes using construction-sector supply indicators remains largely unexplored.

Building on the regime-switching literature, this study develops an interpretable machine learning framework to forecast U.S. housing market expansion and deceleration regimes one month ahead by integrating macroeconomic, financial, and construction-related indicators with national housing price data. Feature selection and hyperparameter tuning are performed entirely on the training sample, using time-series cross-validation to ensure genuine out-of-sample evaluation. The study makes three contributions. First, it reframes housing market analysis from continuous price forecasting to one-month-ahead regime prediction, an operationalization more directly aligned with the discrete, state-dependent decisions faced by construction stakeholders regarding project timing, procurement, and cyclical risk management. Second, it provides a systematic comparison of construction-sector and macro-financial indicators as competing predictor sets, employing recursive feature elimination, multicollinearity screening, and feature-block ablation to

assess their respective regime-discriminatory contributions. Third, by integrating SHAP, partial dependence, and individual conditional expectation analyses within a rigorous temporal evaluation framework, it delivers transparent, economically grounded explanations of regime predictive signals that support practical decision-making for construction managers and policymakers. The remainder of the paper reviews the relevant literature, presents the data and analytical framework, reports the empirical results, and discusses implications, limitations, and directions for future research.

2. Literature Review

2.1. Housing Market Expansion and Deceleration Regimes

Expansion and deceleration regimes in housing markets refer to alternating phases of rapid price appreciation followed by stagnation or decline, reflecting cyclical dynamics that extend beyond short-term fluctuations. The literature employs a variety of approaches to identify these regimes. Agnello and Schuknecht (2011) examined 18 OECD countries from 1970 to 2007 and defined expansion and deceleration as significant deviations of real house prices from long-run trends, showing that downturns tend to be sharper when preceded by strong upswings [4]. Goodhart and Hofmann (2008) used a Markov-switching model across 17 countries to demonstrate the presence of distinct price regimes linked to broader macroeconomic conditions [14]. Wu et al. (2012) studied Chinese cities and identified persistent overvaluation in major urban markets by benchmarking price-to-rent ratios against fundamentals [15]. Tu et al. (2017) applied regime-switching analysis in the Netherlands, concluding that conventional house price models often fail to capture the dynamics of expansion and deceleration in regulated markets where affordability, tax regimes, and supply constraints are central drivers [16]. More recent work by Schellekens and Yasseri (2021) and Martínez (2025) underscores that mortgage debt growth and household lending capacity are effective at anticipating transitions into expansion phases and forecasting the severity of subsequent decelerations [17,18].

2.2. Economic Drivers of Housing Market Dynamics

Credit availability and financing costs are among the most consistently identified drivers of housing market cyclicality. Across 18 industrialized countries, evidence indicates that controlling credit growth, adjusting interest rates, and strengthening regulatory frameworks during expansion can mitigate the severity of subsequent downturns [4]. Cross-country evidence similarly documents synchronized housing price cycles across advanced economies, with common macro-financial conditions amplifying domestic price dynamics and increasing the likelihood of concurrent boom-and-bust episodes [19]. At the national level, mortgage lending, interest rates, income, and the money supply jointly drive both long- and short-run price movements, as demonstrated by cointegration analysis of the United Kingdom from 1983 to 2011 [20]. Capital inflows operate through a related channel: foreign investment raised house prices in England and Wales, particularly for high-value properties, while reducing affordability and long-term vacancies [21]. In emerging markets, fiscal and demographic forces complement credit conditions, with GDP, population growth, and property taxation identified as key determinants in Malaysia [22].

2.3. Housing Market Fundamentals

The flow of new construction and the availability of housing inventory govern near-term supply elasticity and price pressures, and their exclusion from regime models represents a significant gap in the literature. Co-integration and vector error correction analysis of Turkish cities from 2010 to 2020 found that mortgage credit volume strongly influenced demand, while permitting activity, interest rates, and construction costs shaped supply

responses [23]. In Lithuania, new supply, interest rates, inflation, mortgage conditions, and construction costs ranked among the most influential drivers of housing market fluctuations from 2005 to 2015 [24]. A long-run dataset spanning advanced economies from 1870 to 2012 confirmed that GDP, population growth, and construction costs consistently drive housing price cycles across institutional contexts [25]. In Sweden, land prices and building policies determine the extent of cost pass-through, even when interest rates, income, and population dominate price formation [1]. At the micro level, physical and locational characteristics, including infrastructure quality, environmental conditions, home age, size, and coastal proximity, exert strong cross-sectional valuation effects [26,27], though these factors evolve slowly and contribute less to explaining cyclical regime transitions at the market level.

2.4. Methodological Advancements in Housing Market Analysis

Machine learning has expanded the housing analysis toolkit by enabling flexible modeling of nonlinear relationships and high-dimensional sets of predictors. Reviews of the field indicate that regression and hedonic models remain dominant while flexible learners are comparatively underused despite their accuracy advantages [9]. The benchmarking evidence is consistent: ensemble methods outperform linear baselines in housing price prediction across diverse settings, including Boston [28], Bengaluru [29], Chicago [11], and Ames [10]. The Chicago study is particularly relevant because it paired XGBoost prediction with SHAP-based interpretation, identifying square footage, tax burden, and room counts as dominant valuation drivers, an early demonstration that explainability techniques can recover economically meaningful structure from ensemble models [11]. Researchers have also incorporated behavioral signals: online search intensity predicts subsequent price growth in U.S. cities [30], and search-based indices capture latent demand that anticipates price changes through market frictions [31].

2.5. Research Gaps

Synthesizing this literature, three gaps motivate the present study. First, despite the demonstrated importance of expansion and deceleration dynamics, the overwhelming majority of housing studies forecast continuous price levels or growth rates rather than discrete regime states, an objective misaligned with the state-dependent decisions faced by construction firms, developers, and policymakers. The regime identification studies that do exist rely on econometric frameworks whose distributional assumptions and prespecified functional forms limit their capacity to exploit high-dimensional, interacting predictors. Second, construction-sector indicators, including material cost indices, permitting activity, and inventory pipelines, are largely absent from both econometric regime studies and machine learning housing applications, despite the supply-side literature's consistent identification of these variables as central to cyclical behavior. Third, machine learning has been applied almost exclusively to property-level price prediction; regime classification using these methods has rarely been attempted, and no study has systematically benchmarked machine learning classifiers for this task with construction-related indicators under temporally rigorous evaluation. This study addresses these gaps by developing an interpretable machine learning framework for classifying U.S. housing market expansion and deceleration regimes, integrating construction cost and supply indicators alongside macroeconomic and financial variables, and applying explainable artificial intelligence (AI) to recover the economic predictive signals of regime transitions.

3. Materials and Methods

This study integrates regime identification, feature engineering, supervised classification, and model interpretability to analyze cyclical dynamics in the U.S. housing market. Figure 1 presents the overall methodological flow. Monthly price change data derived from the Federal Housing Finance Agency (FHFA) House Price Index (HPI), together with macroeconomic, financial, and construction-related indicators, are used to classify housing market conditions into expansion and deceleration regimes. Lagged features spanning 1 to 3 months are constructed to capture short-term adjustment dynamics, and a temporally ordered train–test split preserves the data's chronological structure to prevent information leakage. Feature selection and multicollinearity control are performed using Recursive Feature Elimination (RFE) and Variance Inflation Factors (VIFs). Five classification models, Logistic Regression, Random Forest (RF), Gradient Boosting (GB), eXtreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM), are benchmarked using accuracy and the Area Under the Receiver Operating Characteristic Curve (ROC-AUC). The classifier achieving the strongest out-of-sample performance is then interpreted using SHAP values, partial dependence (PD) plots, and individual conditional expectation (ICE) analyses to assess global and local feature contributions.

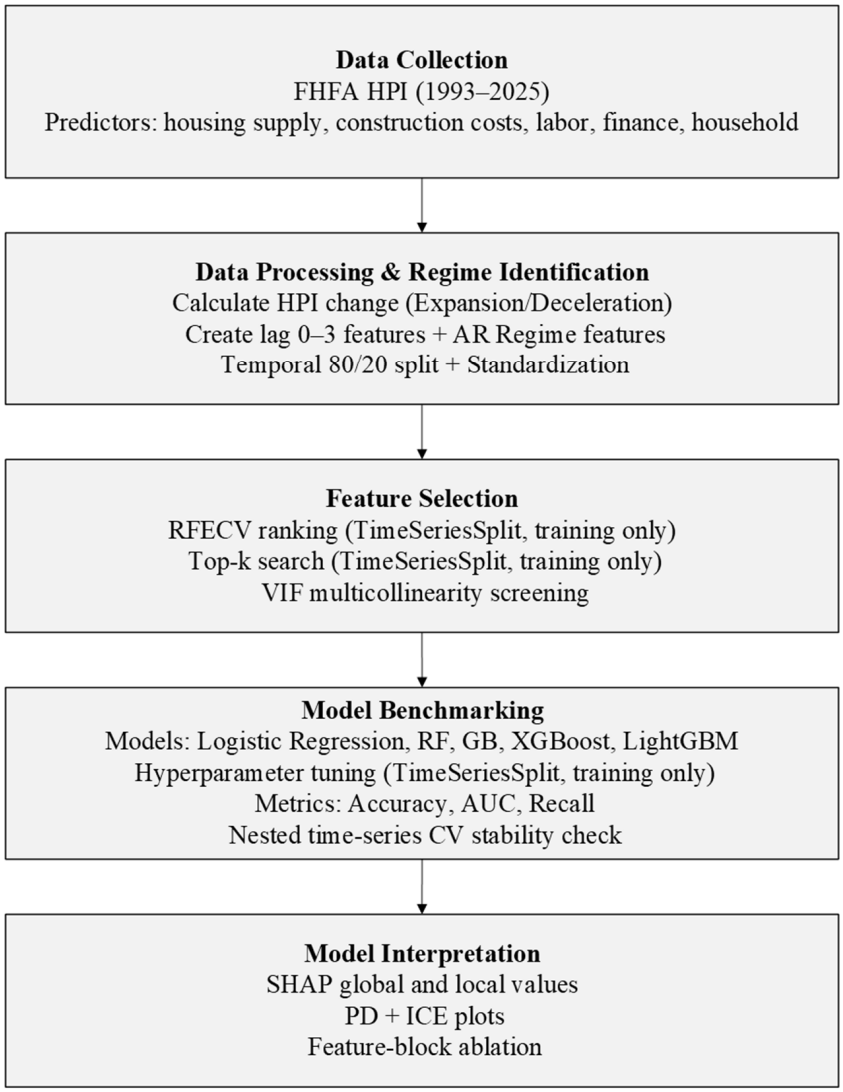


Figure 1. Flowchart of Research Methodology.

3.1. Scope and Data Selection

This study focuses on the U.S. national housing market and employs the FHFA purchase-only HPI as the primary measure of housing price dynamics. The national-level scope is intentionally selected to align with the study's objective of developing and validating a regime-classification framework driven by macroeconomic, financial, and construction-related indicators that operate predominantly at the aggregate level. National housing dynamics are particularly relevant to construction management and policy analysis due to integrated mortgage markets, nationally synchronized monetary policy, and supply chains for construction materials and labor that span regional boundaries.

Substantial spatial heterogeneity characterizes U.S. housing markets, with cyclical behavior, supply constraints, and price volatility varying considerably across states and metropolitan areas. Incorporating sub-national variation would require a fundamentally different modeling framework to account for localized regulatory environments, demographic conditions, and land-use constraints. Such extensions fall outside the scope of the present study and are reserved for future research. The national-level analysis is positioned as a foundational step for regime identification, upon which regional and metropolitan analyses can be systematically developed.

The FHFA purchase-only HPI is selected for several methodological advantages relevant to regime-based analysis. The index is constructed using a weighted repeat-sales methodology that controls for quality changes in housing units over time, reducing compositional bias. The purchase-only specification excludes refinancing transactions, which may introduce volatility unrelated to underlying housing market conditions. The FHFA HPI also provides a long, consistent monthly time series with broad national coverage, making it well-suited for machine learning applications that rely on temporal continuity and stable regime identification.

Alternative indices, such as the S&P/Case–Shiller and Zillow Home Value Index, differ in scope, construction, and coverage. Case–Shiller indices focus primarily on selected metropolitan areas and therefore offer limited national representativeness, whereas Zillow indices rely on proprietary automated valuation models that may be revised algorithmically over time. In contrast, the FHFA HPI offers transparent methodology, comprehensive coverage, and measurement stability, which are essential for identifying cyclical expansion and deceleration regimes rather than short-term price fluctuations.

3.2. Research Objective, Questions, and Hypotheses

The primary objective of this study is to examine whether macroeconomic and construction-related indicators can be used to classify expansion and deceleration regimes in the U.S. housing market using machine learning techniques, and to evaluate the interpretability of such models in explaining cyclical housing dynamics.

To achieve this objective, the study is guided by the following research questions:

RQ1. *Can macroeconomic and construction-related indicators accurately classify expansion and deceleration regimes in U.S. housing prices using machine learning models?*

RQ2. *Do tree-based ensemble machine learning models outperform traditional linear classifiers in predicting housing market regime transitions under nonlinear and interaction-driven dynamics?*

RQ3. *Which macroeconomic, housing supply, and construction cost variables exert the greatest influence on regime classification, and how can their effects be interpreted using explainable AI methods?*

Based on prior research on housing cycles, construction market dynamics, and machine learning applications, the following hypotheses are formulated:

H1. *Tree-based ensemble machine learning models achieve higher predictive accuracy in classifying housing market expansion and deceleration regimes than linear classification models.*

H2. *Construction-related indicators, including construction material costs and housing supply measures, exhibit stronger predictive influence on housing market regime transitions than macro-financial indicators alone.*

H3. *SHAP-based explainability analysis recovers stable, economically meaningful importance rankings for the key signals of housing market regime classification, with construction-related indicators ranking consistently higher than macro-financial variables across cross-validation folds.*

3.3. Data Collection and Processing

Monthly percentage changes in the U.S. HPI (HPI_Change) were used to characterize cyclical housing market conditions and to classify observations into two regimes: expansion and deceleration. The classification threshold was set at the median of HPI_Change over the full study period, partitioning months of above-median price momentum into the expansion regime and months of below-median or negative momentum into the deceleration regime. This median-based partition is consistent with data-driven approaches to regime classification employed in the housing-cycle literature [18,19] and ensures a balanced representation of cyclical states without imposing an arbitrary fixed threshold. The median-based partition captures the cyclical momentum dimension of housing market regimes rather than the level dimension. During the 1993–2025 study period, U.S. house prices appreciated broadly, making sustained above-trend momentum the economically meaningful definition of an expansion rather than absolute price growth. This approach is directly analogous to business-cycle dating methods that define recessions as periods of below-trend growth rather than absolute output decline [32] and to housing-cycle identification approaches that partition price growth relative to long-run trends [4]. All analyses were conducted in Python (v3.11) using the following libraries: scikit-learn (v1.3.0), XGBoost (v1.7.6), LightGBM (v4.0.0), SHAP (v0.42.1), pandas, and NumPy.

Based on this classification, expansion regimes accounted for 207 months (53.6%) with an average monthly growth rate of 0.49%, while deceleration regimes comprised 179 months (46.4%) with an average monthly growth rate of 0.24%. Both regimes exhibit positive average growth, reflecting the broadly appreciating trajectory of U.S. house prices over the 1993–2025 study period; the deceleration regime captures periods of materially weaker momentum rather than outright price decline, consistent with the cyclical deceleration phases documented in long-run housing datasets [25]. These regimes were encoded as a binary target variable (Regime_Label), where expansion = 1 and deceleration = 0, and served as the dependent variable for a one-month-ahead forecasting design in which predictor information does not overlap the period being classified [19,32].

The national FHFA HPI, constructed using a weighted repeat-sales methodology, was employed to ensure consistency over time and to control for quality differences across housing units [33,34]. The analysis draws on monthly national-level data spanning January 1993 through March 2025, yielding 386 monthly observations. Following first-differencing to compute month-over-month HPI change and the introduction of 1-, 2-, and 3-month lag structures, the effective analytical sample comprises 383 observations beginning in May 1993. Although housing markets are inherently local, this study adopts a national framework to capture system-wide dynamics relevant to construction management and policy, including integrated capital markets, nationwide supply chains, and synchronized movements in construction materials and labor costs. This aggregation facilitates robust regime detection and model validation, while regional and metropolitan extensions are explicitly identified as an important avenue for future research.

To capture short-term adjustment dynamics associated with construction activity, material costs, and financing conditions, lag structures of one, two, and three months were constructed for all indicators. Throughout this study, the notation $t-1$, $t-2$, and $t-3$ denotes the value of a given indicator lagged by one, two, and three months, respectively, reflecting the time elapsed between the predictor observation and the month being classified. To ensure a genuine forecasting design, predictors are constructed at lags t, $t-1$, $t-2$, and $t-3$ relative to the forecast origin month, while the classification target corresponds to the regime state at $t+1$, ensuring that no predictor shares a reference month with the label it forecasts. Unlike long-horizon price forecasting, regime shifts often emerge from near-term shocks and feedback effects rather than slow-moving fundamentals, and prior studies in financial and housing time-series modeling similarly emphasize short-lag structures for directional or regime-based prediction [35,36]. Observations with missing values introduced by lag construction were removed. A strict temporal 80%:20% train–test split was implemented, with the training set comprising the earlier portion of the sample (May 1993–September 2018, $n = 305$) and the test set comprising the most recent observations (October 2018–February 2025, $n = 77$), ensuring that all predictions are generated strictly out-of-sample. The training set was used exclusively for feature selection, hyperparameter tuning, and model fitting, while the test set was held out entirely for final performance evaluation.

A comprehensive set of macroeconomic, financial, and construction-related indicators was assembled to capture the multiple dimensions that influence U.S. housing market dynamics (Table 1). Housing supply conditions were represented by Building Permits and New Privately Owned Houses for Sale, while the Value of Construction Put in Place captured overall construction-sector activity. Construction input costs and material price pressures were measured using the Consumer Price Index (CPI) for Single-Family Houses Under Construction and the suite of Engineering News-Record (ENR) cost indices, including the Construction Cost Index (CCI), Building Cost Index (BCI), and Materials Price Index (MPI). Labor market conditions were captured through the ENR Skilled Labor Index (SLI), the ENR Common Labor Index (CLI), and the Unemployment Level, reflecting wage pressures and employment conditions that affect both housing demand and construction costs.

Table 1. Input Variables Used in Housing Price Index Analysis.

Category	Variable	Definition	Units	Source
Supply & Construction Activity	Building Permits	Number of new residential building permits issued in the U.S., a forward-looking housing supply indicator	Number of permits	U.S. Census Bureau
	New Privately Owned Houses for Sale	Number of new privately owned homes listed for sale, proxy for housing inventory	Number of homes	U.S. Census Bureau
	Value of Construction Put in Place (Total Construction)	Total monthly value of construction placed, captures overall construction industry activity	USD (million /billion)	U.S. Census Bureau

Table 1. *Cont.*

Category	Variable	Definition	Units	Source
Labor Market & Wages	ENR Skilled Labor Index (SLI)	Tracks wage rate changes for skilled trades (e.g., bricklayers, carpenters, ironworkers), including fringe benefits	Index (base = 100)	ENR
	ENR Common Labor Index (CLI)	Tracks wage rate changes for unskilled (common) labor, including fringe benefits	Index (base = 100)	ENR
	Unemployment Level	Total number of unemployed individuals in the U.S. labor force	Thousands of persons	Bureau of Labor Statistics (BLS)
Construction Costs & Materials	ENR Construction Cost Index (CCI)	Tracks cost changes in construction inputs (unskilled labor, cement, structural steel, lumber) across 20 U.S. cities	Index (base = 100)	ENR
	ENR Building Cost Index (BCI)	Similar to ENR CCI but incorporates skilled labor instead of common labor	Index (base = 100)	ENR
	ENR Materials Price Index (MPI)	Measures weighted price changes in key building materials, including steel, cement, and lumber	Index (base = 100)	ENR
	CPI for Single-Family Houses Under Construction	Consumer Price Index tracking price changes for single-family houses under construction	Index (base = 100)	BLS
Monetary & Credit Conditions	30-Year Fixed-Rate Mortgage Average	Average interest rate on 30-year fixed-rate mortgages	Percent (%)	Federal Reserve Bulletin
	Federal Funds Effective Rate	Interest rate banks charge each other for overnight loans, benchmark set by the U.S. Federal Reserve	Percent (%)	Federal Reserve System
Macroeconomic Indicators	CPI for All Urban Consumers (CPI-U)	Broad consumer price index measuring average price changes over time for goods and services purchased by urban households	Index (base = 100)	BLS
Household Finances	Personal Saving Rate	Percentage of disposable personal income that households save, indicator of financial posture	Percent (%)	Bureau of Economic Analysis

Broader macroeconomic and financial conditions were represented by the Consumer Price Index for All Urban Consumers (CPI-U), the 30-Year Fixed-Rate Mortgage Average, and the Federal Funds Effective Rate, which jointly capture inflationary pressures, borrow-

ing costs, and monetary policy stance. The Personal Saving Rate was included to reflect households' financial capacity and potential participation in the housing market. Together, these indicators provide a balanced representation of supply, demand, cost, and macro-financial forces shaping national housing price regimes. Figure 2 illustrates the national HPI series and its month-over-month percentage change, highlighting long-term growth trends alongside major episodes of volatility, including the 2007–2009 global financial crisis and the 2020–2022 pandemic-era housing market surge.

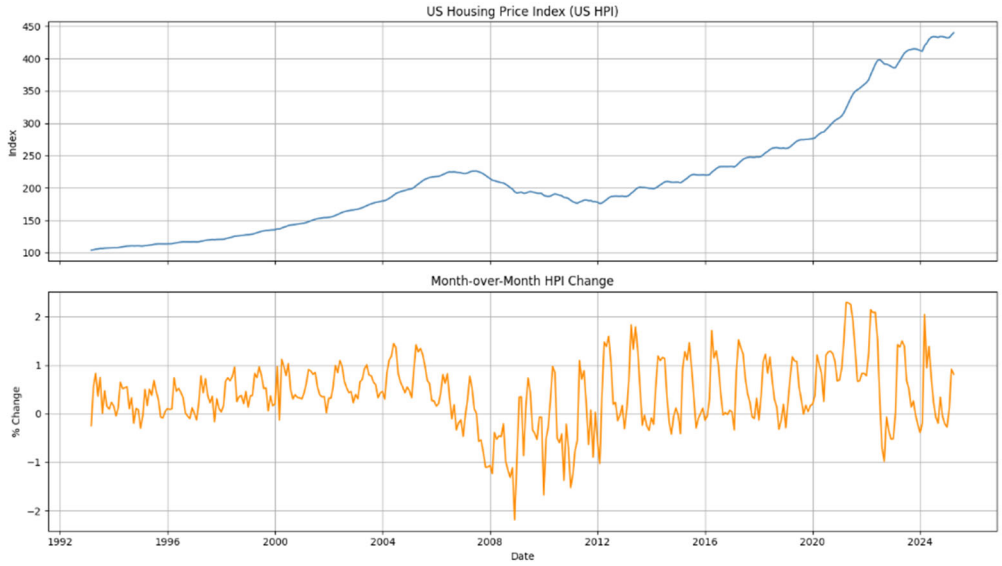


Figure 2. U.S. Housing Price Index and Month-over-Month Percentage Change (January 1993–March 2025).

3.4. Feature Selection and Multicollinearity Control

RFE and VIF screening are employed as model-agnostic preprocessing steps to reduce dimensionality, improve stability, and enhance interpretability rather than to optimize any specific classifier. Although tree-based ensemble methods are less sensitive to multicollinearity than linear models, correlated predictors can still inflate the variance of feature-importance measures and complicate the economic interpretation of model outputs. The combined RFE-VIF procedure, therefore, identifies a compact, non-redundant feature set that supports consistent and interpretable comparisons across all benchmarked classifiers.

RFE iteratively removes the least important variables based on model-assigned importance scores, establishing a stable predictor ranking and ensuring that weak contributors are excluded early to improve parsimony [37]. A top-k search (k = 1 to 30) was then conducted to identify the subset size that best balanced predictive performance and model stability. Candidate subset sizes were evaluated using five-fold TimeSeriesSplit cross-validation on the training set only, with each fold expanding the training window chronologically to prevent future information from leaking into feature selection. The subset achieving the highest training cross-validated ROC-AUC was selected as the optimal feature count [38,39].

The retained predictors were subsequently screened for multicollinearity using VIF. VIF measures the degree to which each predictor's variance is inflated by its correlation with the remaining predictors, calculated as:

$$VIF_j = \frac{1}{1 - R_j^2} \quad (1)$$

where R_j^2 is the coefficient of determination from the regressing predictor X_j on all other predictors. Values above 10 are widely considered indicative of problematic multicollinear-

ity and warrant variable removal [40,41]. This screening step was particularly important to ensure that retained predictors represent distinct economic predictive signals of regime dynamics rather than redundant proxies for the same underlying process, which would otherwise destabilize feature-importance rankings across cross-validation folds. By combining RFE with VIF screening, the study produced a compact, non-redundant, and interpretable set of predictors that provided a stable foundation for benchmarking the five classification models.

3.5. Model Benchmarking

To evaluate the ability of different modeling approaches to classify expansion and deceleration housing market regimes, this study benchmarked five classification algorithms: Logistic Regression, Random Forest, Gradient Boosting, XGBoost, and LightGBM. These models were selected not simply to compare algorithmic performance, but to examine how differing assumptions about linearity, feature interactions, and regularization affect regime prediction in a housing market shaped by construction costs, supply pipeline dynamics, and macro-financial conditions. Logistic Regression serves as a linear baseline, providing an interpretable benchmark against which the gains from nonlinear ensemble methods can be quantified. The four tree-based ensemble models collectively represent the state of practice in tabular classification and span a range of ensemble strategies, from independent bagging in Random Forest to sequential residual correction in Gradient Boosting, XGBoost, and LightGBM.

Expansion and deceleration regimes reflect nonlinear, threshold-driven behavior rather than a smooth price evolution. Housing markets often respond asymmetrically to shocks: increases in construction material prices may trigger rapid expansion conditions when supply is constrained, while elevated inventories or tightening credit conditions may precipitate abrupt transitions into deceleration phases. Capturing such dynamics requires models capable of learning interaction effects, lagged responses, and regime-specific sensitivities, capabilities that extend beyond the linear log-odds structure of traditional econometric classifiers and motivate the inclusion of progressively more flexible ensemble architectures in the benchmark suite.

All models were trained on the final feature set obtained from the feature selection process. Hyperparameter tuning focused on controlling model complexity and mitigating overfitting, given the temporally structured nature of regime observations. The parameters governing tree depth, number of estimators, learning rate, and regularization strength were optimized using cross-validation on the training data, with the tuned configurations held constant across all subsequent evaluation steps. Model performance was assessed on the held-out test set using classification accuracy and ROC-AUC, enabling a robust and consistent comparison of predictive accuracy and out-of-sample generalization across all five classifiers.

3.5.1. Logistic Regression (Baseline)

Logistic Regression was employed as a baseline model, serving as a conventional, fully interpretable approach to regime classification. In the context of housing markets, logistic models have been widely used to estimate the probability of expansionary versus contractionary phases based on macroeconomic indicators [42–44]. The model specifies the log-odds of an expansion regime as a linear function of the predictors:

$$Logit(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \sum_{i=1}^{m} \beta_i x_i \quad (2)$$

where p is the predicted probability of an expansion regime, x_i represents the retained predictor variables, and β_i is the estimated coefficient [40,44]. Positive coefficients increase the predicted likelihood of an expansion, while negative coefficients are associated with deceleration conditions, offering direct interpretability of how supply, cost, and macro-financial variables contribute to regime classification. For example, a positive coefficient on lagged construction material costs would indicate that rising input prices are associated with higher expansion probability, while a negative coefficient on housing inventory would signal that elevated supply exerts downward pressure on regime momentum.

To improve generalization and reduce coefficient instability arising from correlated predictors, L2 regularization was applied, with the penalty strength (λ) selected via cross-validation on the training data [45]. Despite this regularization, Logistic Regression remains constrained by its linear log-odds structure, which limits its ability to capture nonlinear responses, threshold effects, and higher-order interactions that are characteristic of housing and construction market dynamics. These limitations motivate the inclusion of more flexible, nonlinear ensemble methods in the benchmark suite.

3.5.2. Tree-Based Ensemble Methods

Tree-based ensemble methods were applied to better reflect the nonlinear, interaction-driven nature of housing market expansion and deceleration regimes. Decision trees recursively partition the predictor space, allowing the models to learn regime boundaries that depend on combinations of variables rather than additive linear effects. This property is particularly important in housing markets, where regime transitions often arise from the joint influence of construction cost pressures, supply constraints, and macro-financial conditions.

Random Forest aggregates predictions from multiple decision trees built on bootstrap samples and random subsets of predictors. In the context of housing regime classification, this approach improves robustness and reduces variance by averaging across many alternative representations of market dynamics [28,46]. However, because Random Forest averages across trees, it may smooth sharp regime boundaries, potentially limiting its sensitivity to abrupt transitions triggered by sudden changes in construction costs or inventory levels.

Gradient Boosting addresses this limitation by building trees sequentially rather than independently, with each successive tree focusing on correcting the misclassifications of prior iterations. This learning strategy is well-suited to regime detection, where observations near transition points, such as the onset of an expansion following sustained cost escalation, are particularly informative. By iteratively refining decision boundaries, Gradient Boosting improves the model's ability to detect subtle early signals of regime change [47,48].

XGBoost extends Gradient Boosting by incorporating explicit regularization and parallel computation, improving generalization and computational efficiency. In housing market applications, XGBoost has demonstrated strong performance in capturing nonlinear relationships between prices, supply, and macroeconomic variables while controlling overfitting [10,11,49,50]. LightGBM further enhances the Gradient Boosting framework through a leaf-wise tree growth strategy and histogram-based split finding. Rather than expanding trees level by level, LightGBM focuses growth on regions of the predictor space that yield the greatest reduction in classification error. This property allows the model to more precisely capture threshold-driven regime dynamics, such as sharp increases in expansion probability once lagged construction material costs or permitting activity exceed critical levels. These characteristics make LightGBM particularly well-suited for identifying

housing market regimes driven by nonlinear responses and lagged construction-sector signals [51,52].

Across all four ensemble models, hyperparameters, including learning rate, tree depth, number of estimators, and regularization terms, were optimized via cross-validation on the training data. Early stopping was applied where available to prevent overfitting and ensure stable out-of-sample performance.

3.6. Model Interpretation

To assess whether model performance and feature importance rankings are stable across diverse historical sub-periods rather than artifacts of a single train–test split, a nested time-series cross-validation analysis was conducted. The full sample was divided into five sequential, non-overlapping outer folds using TimeSeriesSplit, with each fold comprising a training window and a held-out test window that always follows the training period chronologically. Within each outer fold, the complete pipeline, including RFECV feature selection, VIF screening, and hyperparameter tuning, was repeated from scratch using only that fold's training data, with an inner TimeSeriesSplit applied for all internal cross-validation steps. Each tuned model was then evaluated exactly once on that fold's held-out test window, which played no role in feature selection, tuning, or threshold selection. This nested structure ensures that each fold's results are fully independent and leakage-free, providing five replications of the complete analysis across the pre-crisis expansion, the global financial crisis, the post-crisis recovery, the low-rate expansion, and the pandemic and rate-shock correction periods. The mean and standard deviation of holdout AUC and Recall across the five folds are reported as measures of cross-episode stability.

While flexible ensemble models can capture complex nonlinear and interaction-driven dynamics, their flexibility can obscure direct economic interpretation. To address this trade-off and avoid treating the models as black boxes, this study integrates SHAP with partial dependence (PD) and individual conditional expectation (ICE) analyses to recover economically meaningful insights from complex classifiers [12,13,53].

Using SHAP with TreeExplainer, each model prediction is decomposed into additive contributions from individual predictors:

$$f(x) = \phi_0 + \sum_{i=1}^{m} \phi_i \tag{3}$$

where $f(x)$ is the predicted probability of an expansion regime for a given month, ϕ_0 is the baseline expectation representing the average expansion likelihood across the sample, and ϕ_i represents the marginal contribution of predictor i to that individual prediction [12]. This framework enables both global interpretations, identifying the most influential construction and macroeconomic signals of regime classification across the full test set, and local interpretations, explaining specific regime outcomes for individual months. Cross-validated SHAP stability analysis further assesses whether importance rankings are consistent across training folds or sensitive to particular historical episodes, distinguishing robust signals from estimation artifacts. It is important to note that SHAP values quantify the statistical association between each predictor and the model's classification output and do not establish causal relationships between construction-sector conditions and housing market regimes. All interpretations of SHAP results should therefore be understood as predictive associations consistent with economic theory rather than as evidence of causal direction.

PD and ICE plots complement SHAP by illustrating how predicted regime probabilities respond to changes in individual predictors across their observed ranges, revealing nonlinear threshold effects and heterogeneous sensitivities that aggregate importance measures cannot capture [13,54]. Together, these interpretability tools ensure that predictive

accuracy is paired with transparent, economically grounded explanations, aligning the analysis with the needs of construction managers and policymakers seeking actionable signals of cyclical regime transitions [55]. To ensure reproducibility, all stochastic elements, including model initialization and SHAP computation, were controlled using fixed random seeds.

4. Results

4.1. Feature Selection

RFE with logistic regression was used to rank predictors by their contribution to the classification of expansion and deceleration regimes. As shown in Table 2, the top-ranked predictors are led by the seven variables jointly selected at the optimal subset size: Building Permits (contemporaneous and one-month lag), Value of Construction Put in Place for Total Construction (two- and three-month lags), CPI for Single-Family Houses Under Construction (one-month lag), and two autoregressive regime-persistence features capturing the realized regime state at the forecast origin and one month prior.

Table 2. RFE-Based Ranking of Top 20 Predictors for Expansion and Deceleration Regime Classification.

Rank	Feature
1	Building Permits
2	Building Permits (t − 1)
3	Value of Construction Put in Place for Total Construction (t − 2)
4	Value of Construction Put in Place for Total Construction (t − 3)
5	CPI for Single-Family Houses Under Construction (t − 1)
6	Realized Regime State (t)
7	Realized Regime State (t − 1)
8	ENR CLI
9	ENR BCI
10	ENR CCI
11	Building Permits (t − 2)
12	Consumer Price Index for All Urban Consumers (t − 1)
13	Consumer Price Index for All Urban Consumers
14	Value of Construction Put in Place for Total Construction (t − 1)
15	Value of Construction Put in Place for Total Construction
16	ENR CCI (t − 2)
17	New Privately Owned Houses for Sale (t − 3)
18	Personal Saving Rate (t − 3)
19	30-Year Fixed-Rate Mortgage Average (t − 1)
20	Consumer Price Index for All Urban Consumers (t − 2)

Candidate subset sizes were evaluated using five-fold time-series cross-validation on the training set only, with each fold's validation set following its training set chronologically to prevent future information from leaking into feature selection. As shown in Figure 3, cross-validated ROC-AUC peaked at seven retained features (AUC = 0.825), indicating that a subset of seven variables provided the optimal balance between predictive performance and model parsimony.

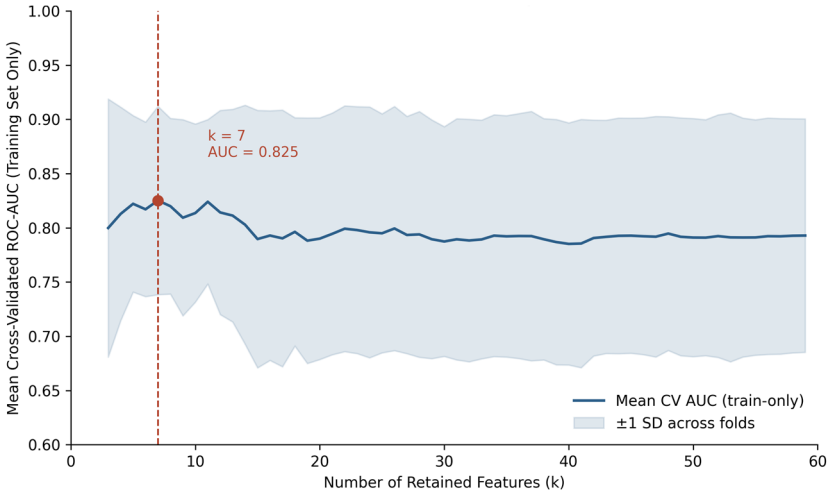


Figure 3. Cross-Validated ROC-AUC by Number of Retained Features.

VIF analysis was performed to check for multicollinearity among the seven RFECV-selected predictors. Unlike the pre-revision analysis, in which several construction and cost indicators exhibited near-perfect collinearity with their own lags (VIF exceeding 800 in one case) because raw indicator levels were lagged directly, the current analysis first transforms all base indicators to period-over-period percentage change, or first difference, for indicators already expressed as rates, before lagging. This transformation removes the adjacent-lag redundancy that previously drove extreme VIF values: all seven RFECV-selected predictors already met the 10 threshold without requiring any iterative removal, with VIF values ranging from 1.30 to 3.13 (Table 3).

Table 3. VIF of RFECV-Selected Predictors (Threshold = 10).

Variable	VIF
Realized Regime State (t − 1)	3.13
Realized Regime State (t)	2.63
Value of Construction Put in Place for Total Construction (t − 3)	2.06
Value of Construction Put in Place for Total Construction (t − 2)	1.93
Building Permits	1.40
CPI for Single-Family Houses Under Construction (t − 1)	1.38
Building Permits (t − 1)	1.30

Both Building Permits and its one-month lag were retained, reflecting distinct but complementary aspects of housing supply dynamics: contemporaneous permitting activity captures current market conditions, while the one-month lag captures the delayed transmission of permitting decisions into subsequent supply and price movements. Similarly, both the two- and three-month lags of Value of Construction Put in Place for Total Construction were retained, capturing the cumulative effect of construction spending over a longer window rather than a single lag. The two autoregressive regime-persistence features, the realized regime state at the forecast origin and one month prior, were also retained, consistent with the strong regime persistence documented in the walk-forward validation. Together, the retained variables represent construction activity, construction cost, and regime-persistence indicators, offering a parsimonious and interpretable input space for model development.

4.2. Model Development Results

Five classification algorithms were trained on the final seven-variable feature set and evaluated on both training and test datasets (see Table 4). Random Forest achieved the highest test AUC (0.9319) and tied with XGBoost for the highest test accuracy (0.8701), indicating strong discrimination between expansion and deceleration regimes with balanced precision (0.8571) and recall (0.9333). XGBoost delivered nearly identical performance, matching Random Forest in accuracy, F1 (0.8936), precision, and recall, with a slightly lower AUC (0.9194). Gradient Boosting exhibited more uneven behavior, combining high precision (0.9333) with substantially lower recall (0.6222) and accuracy (0.7532), suggesting under-identification of one regime despite confident positive predictions. The baseline logistic regression model, while offering interpretability advantages, achieved moderate predictive performance (test accuracy = 0.8442; test AUC = 0.9111) with balanced precision and recall (0.8667 each). LightGBM showed more modest results in this evaluation (test accuracy = 0.8052; test AUC = 0.9014), with lower recall (0.8000) relative to the other ensemble methods. Overall, the ensemble tree-based models, particularly Random Forest and XGBoost, outperformed the baseline, offering the best balance of predictive accuracy, discrimination, and generalization, making them the preferred models for interpretation and further analysis.

Table 4. Results of Model Comparisons.

Model	AUC	Accuracy	F1	Precision	Recall
Random Forest	0.932	0.87	0.894	0.857	0.933
Gradient Boosting	0.927	0.753	0.747	0.933	0.622
XGBoost	0.919	0.87	0.894	0.857	0.933
Logistic Regression	0.911	0.844	0.867	0.867	0.867
LightGBM	0.901	0.805	0.828	0.857	0.8

To complement the holdout evaluation and address concerns about performance reliability, two additional validation analyses were conducted. First, walk-forward validation using an expanding window, beginning with 60% of the sample and advancing in 12-month steps, yielded a model AUC of 0.893 and accuracy of 0.838, compared to a persistence baseline accuracy of 0.844. The model's walk-forward AUC of 0.893 confirms genuine predictive skill beyond naive regime persistence. Second, nested time-series cross-validation across five non-overlapping historical windows (Table 5) shows Random Forest achieving a mean holdout AUC of 0.836 (SD = 0.104) and mean Recall of 0.823 (SD = 0.078), the most consistent performance across all models, confirming that the holdout result reflects a generalizable regime-discriminatory signal rather than period-specific overfitting.

Figure 4 presents an example decision tree extracted from the Random Forest model at a maximum depth of three levels, illustrating the threshold-based decision logic underlying the ensemble classifier. The root node splits on AR Regime (lag 0), reflecting the dominant role of prior regime state in one-month-ahead prediction. Subsequent splits incorporate Construction Put in Place and Building Permits lags, demonstrating how the model combines regime persistence with construction supply-side signals to partition observations into expansion and deceleration classifications. This single tree is shown for interpretability purposes and is representative of the ensemble logic, though individual trees within the Random Forest may vary in structure.

Table 5. Nested Time-Series Cross-Validation Stability Results (5 Outer Folds).

Model	Mean AUC	SD AUC	Mean Accuracy	Mean F1	Mean Recall	SD Recall
Random Forest	0.836	0.104	0.752	0.767	0.823	0.078
Gradient Boosting	0.834	0.111	0.683	0.694	0.741	0.231
Logistic Regression	0.832	0.109	0.702	0.693	0.713	0.245
XGBoost	0.825	0.106	0.724	0.733	0.786	0.142
LightGBM	0.814	0.127	0.692	0.699	0.714	0.166

Note. Outer fold test periods: Fold 0 (1998–2004), Fold 1 (2004–2009), Fold 2 (2009–2014), Fold 3 (2014–2019), Fold 4 (2019–2025).

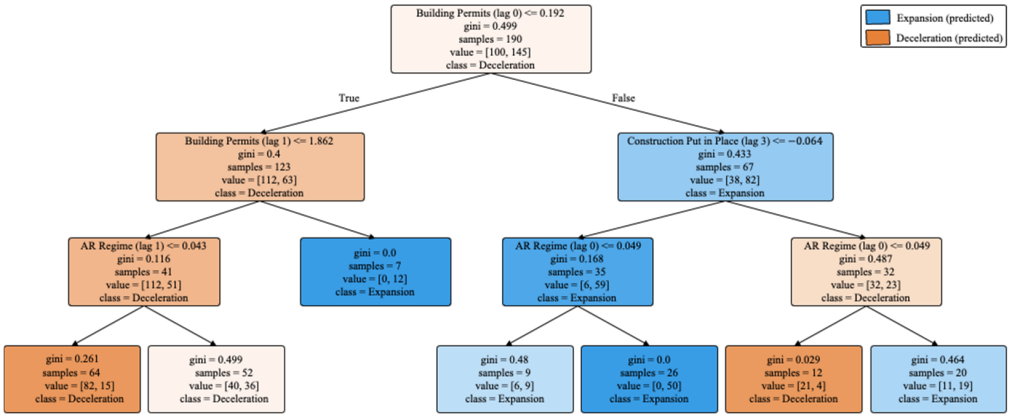


Figure 4. Example Decision Tree Classifier (Maximum Depth = 3).

4.3. Feature Interpretation

SHAP-based interpretation of the Random Forest model was applied to assess the direction and magnitude of each predictor's contribution to expansion and deceleration regime classification at the one-month-ahead prediction horizon. Figure 5 illustrates the distribution of SHAP values across all test-set observations, with point color indicating whether the underlying feature value was high (red) or low (blue), and Table 6 summarizes mean absolute SHAP values for all seven selected predictors.

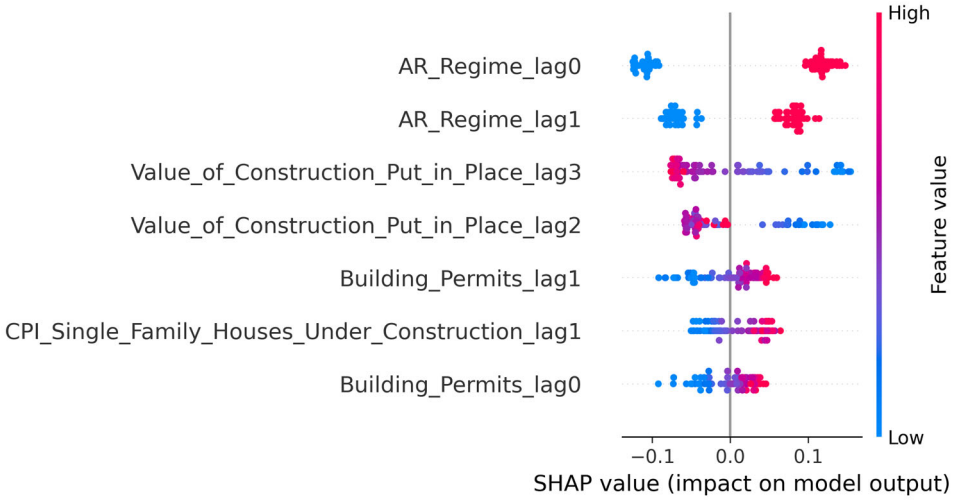


Figure 5. SHAP Beeswarm Plot for Random Forest Expansion and Deceleration Regime Classification.

Table 6. SHAP Feature Importance of Random Forest Model (Expansion and Deceleration Regime Classification, t + 1 Prediction).

Rank	Feature	Category	Mean |SHAP|
1	AR Regime (lag 0)	Autoregressive	0.115
2	AR Regime (lag 1)	Autoregressive	0.0761
3	Construction Put in Place (lag 3)	Construction	0.0638
4	Construction Put in Place (lag 2)	Construction	0.0535
5	Building Permits (lag 1)	Construction	0.0329
6	CPI Single-Family Construction (lag 1)	Construction	0.0319
7	Building Permits (lag 0)	Construction	0.0253

The two autoregressive regime features dominated the importance ranking. AR Regime (lag 0) showed the strongest predictive association (mean |SHAP| = 0.115), with prior expansion states strongly associated with predicted expansion at t + 1 and prior deceleration states associated with deceleration predictions. AR Regime (lag 1) ranked second (mean |SHAP| = 0.076), confirming that regime persistence is the single most powerful predictive signal, regimes tend to continue into the next month, and the model captures this accurately. Among construction-sector predictors, Construction Put in Place (lag 3) ranked third (mean |SHAP| = 0.064), with higher construction activity three months prior associated with higher predicted expansion probability, consistent with the delayed propagation of construction pipeline momentum into housing price dynamics. Construction Put in Place (lag 2) contributed similarly (mean |SHAP| = 0.053), confirming that the two-to-three-month lag window is the most informative supply-side horizon. Building Permits (lag 1) ranked fifth (mean |SHAP| = 0.033), reflecting the forward-looking nature of permit issuance as a housing supply signal. CPI for Single-Family Houses Under Construction (lag 1) ranked sixth (mean |SHAP| = 0.032), with higher construction cost pressures positively associated with expansion probability, consistent with cost escalation during strong-demand conditions. Building Permits (lag 0) contributed the smallest signal (mean |SHAP| = 0.025). Together, the five construction-sector predictors demonstrate that supply-side dynamics carry meaningful regime-discriminatory signals beyond pure regime persistence, providing economically grounded information about upcoming expansion and deceleration conditions that complement the autoregressive signal. The complete absence of macro-financial variables from the selected feature set confirms that construction activity and regime persistence together provide a sufficient discriminatory signal for one-month-ahead prediction, consistent with Hypothesis 2.

Figure 6 presents partial dependence and ICE plots for all seven selected predictors, ordered by mean absolute SHAP importance. AR Regime (lag 0) and AR Regime (lag 1) exhibited sharp, near-binary relationships with predicted expansion probability, prior expansion states sharply increased t + 1 expansion probability while prior deceleration states produced the opposite, with narrow ICE spread confirming highly uniform behavior across observations. Construction Put in Place at lags 3 and 2 showed positive relationships with moderate ICE dispersion, reflecting variability in how construction momentum translates across historical episodes. Building Permits (lag 1) exhibited a positive directional pattern with wider ICE spread, suggesting context-dependent effects depending on prevailing market conditions. CPI for Single-Family Construction (lag 1) showed a broadly positive relationship consistent with cost escalation under strong demand conditions. Building Permits (lag 0) showed the most dispersed ICE curves, indicating interaction effects with

other features rather than a consistent independent contribution. Together, these patterns confirm that regime persistence and construction supply-side dynamics are the dominant predictive signals of one-month-ahead regime classification.

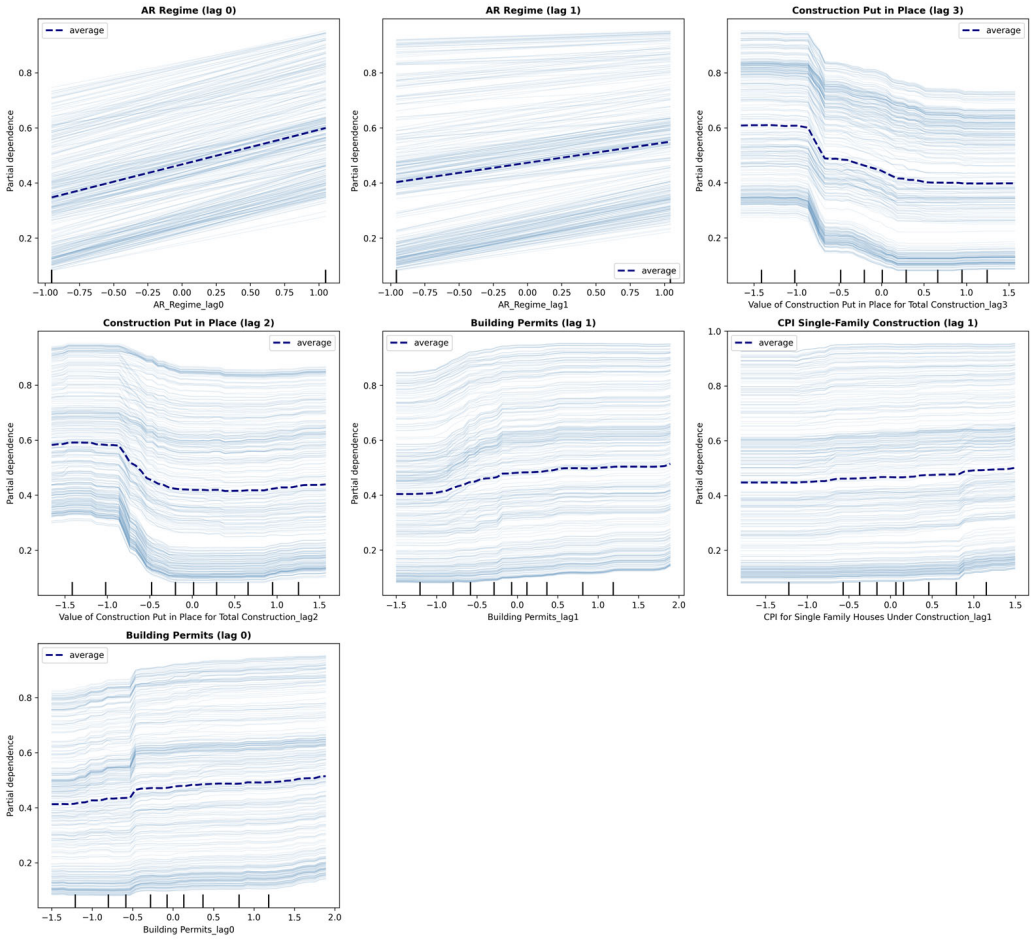


Figure 6. Partial Dependence and ICE Curves for All Seven Selected Predictors.

To directly test whether construction-sector indicators provide greater incremental predictive value than macro-financial variables independently of SHAP rankings, a feature-block ablation analysis was conducted using Random Forest across four specifications (Table 7). The full set, including autoregressive features, achieved the highest holdout AUC (0.921) and training CV AUC (0.855), confirming that regime persistence carries the strongest incremental signal. Removing the AR features reduced holdout AUC to 0.863, demonstrating their contribution while confirming that economic indicators alone still provide meaningful discrimination. The construction-only block achieved a training CV AUC of 0.849, comparable to the full non-AR set (0.847), while the macro-financial-only block achieved only 0.665. This 0.184 gap in training CV AUC provides direct model-based evidence supporting Hypothesis 2. The higher holdout AUC for the macro-financial block (0.885 versus 0.814 for construction-only) should be interpreted cautiously, as it reflects differences in feature set size rather than genuine predictor quality. The training CV AUC, which is the leakage-free metric, consistently favors construction indicators.

Table 7. Feature-Block Ablation Results (Random Forest).

Feature Subset	*n* Features	Train CV AUC	Holdout AUC	Holdout Recall
Full Set with AR Features	59	0.855	0.921	0.778
Full Set without AR Features	56	0.847	0.863	0.644
Construction Only	36	0.849	0.814	0.600
Macro-Financial Only	20	0.665	0.885	0.844

5. Discussion

5.1. Critical Contributions of the Study

This study advances housing market analytics by introducing a regime-based classification framework that transforms continuous housing price movements into discrete expansion and deceleration states and generates genuine one-month-ahead forecasts using information available prior to the classified period. By shifting the analytical focus from price levels to cyclical regime forecasting, the approach more directly reflects the state-dependent decisions faced by construction firms, developers, and policymakers. The corrected pipeline, with feature selection and hyperparameter tuning performed entirely on the training sample using time-series cross-validation, ensures that the reported performance reflects genuine out-of-sample generalization rather than test-set optimization.

Benchmarking results demonstrate that tree-based ensemble models, particularly Random Forest, outperform the logistic regression baseline in capturing nonlinear relationships, threshold effects, and short-term temporal dependencies among construction-related variables and autoregressive regime signals. Random Forest achieved the highest holdout AUC (0.932) and the most consistent deceleration detection across nested cross-validation folds (mean Recall = 0.823, SD = 0.078), confirming the superiority of ensembles, consistent with prior research on the advantages of ensemble learning in complex economic systems [56,57]. Integrating SHAP, partial dependence, and ICE analyses enhances methodological transparency, enabling predictive results to be interpreted in economically meaningful terms rather than treated as black-box outputs.

Substantively, the study makes three contributions. First, it demonstrates that a seven-variable feature set, comprising five construction-sector indicators and two autoregressive regime-persistence features, is sufficient for one-month-ahead housing market regime forecasting, with no macro-financial variable contributing an incremental discriminatory signal. Second, it provides direct model-based evidence through feature-block ablation that construction indicators achieve a training CV AUC of 0.849 compared to 0.665 for macro-financial indicators alone, a gap of 0.184, independently confirming the supply-side dominance finding. Third, the autoregressive regime features reveal that regime persistence is the single strongest predictive signal, with Construction Put in Place and Building Permits contributing complementary supply-side momentum beyond pure persistence.

5.2. Implications for the Construction Industry and Stakeholders

The findings carry several important implications for construction management and related stakeholders. Consistent with prior studies [22,23], the results confirm that building permits and construction activity are critical supply-side predictive signals of housing market regime dynamics. This study extends existing work by demonstrating that these indicators function as genuine one-month-ahead forecasting signals through their lagged

effects. Past permitting activity and construction output provide actionable information about the regime state that will prevail in the following month, supporting their use as early indicators for procurement timing, project initiation decisions, and cyclical exposure management.

The dominance of Construction Put in Place at lags 2 and 3 reflects the delayed propagation of construction pipeline momentum into housing market conditions. Construction activity initiated in a given month takes two to three months to translate into observable supply-side effects, as materials are ordered, labor is deployed, and units move toward completion, creating a measurable lag between construction momentum and housing price regime dynamics. This supply-chain transmission delay is precisely what the model captures, and it explains why two- to three-month-lagged construction activity outperforms both contemporaneous measures and macro-financial indicators in regime discrimination.

The role of Building Permits as a complementary signal reflects the forward-looking nature of permit issuance. Permits signal developer intentions to add supply, and elevated permitting activity one month prior is associated with expansion regime predictions, consistent with the interpretation that active development pipelines sustain price momentum. CPI for Single-Family Houses Under Construction captures construction cost pressures, with higher costs positively associated with expansion probability, consistent with cost escalation occurring predominantly under strong demand conditions when construction activity is elevated, and materials are bid upward.

The dominance of autoregressive regime features in the SHAP ranking reveals that housing market regimes are highly persistent. Once a market enters an expansion or deceleration state, it tends to remain in that state for multiple consecutive months. From a construction management perspective, this persistence implies that regime transitions are relatively infrequent and that current market conditions are themselves a strong predictor of near-term conditions. The construction indicators contribute incremental predictive value on top of this persistence signal, identifying the supply-side conditions under which regime transitions are more or less likely to occur. Together, these findings support the use of AI-driven regime classification as a strategic decision-support tool for timing project initiation, managing procurement risk, and anticipating cyclical exposure.

5.3. Interpreting Regime Transitions Through Explainable AI

The SHAP-based interpretability results provide additional insight into how economic mechanisms manifest in one-month-ahead regime forecasting. The two-tier importance structure, with autoregressive persistence features dominating and construction indicators contributing complementary supply-side signals, reflects a fundamental property of housing market dynamics: regime states are sticky, and the most reliable near-term forecast is that current conditions will persist. The construction indicators quantify how much the supply-side economic environment modifies the persistence signal, increasing or decreasing the probability of expansion relative to what pure regime momentum alone would predict.

The directional patterns observed in Figure 6 are consistent with the economic mechanisms described above. High AR Regime values, indicating a prior expansion state, sharply increase the predicted expansion probability at t + 1, while prior deceleration states produce the opposite effect. The narrow ICE spread for both AR Regime features confirms that the persistence mechanism operates consistently regardless of broader market conditions. Construction Put in Place at lags 2 and 3 shows positive relationships with moderate ICE dispersion, reflecting variability in how construction momentum translates into regime predictions across different historical episodes. Building Permits (lag 1) shows a wider ICE spread, indicating that the effect of permitting on predicted regime depends on the broader

market context in which permits are issued. Building Permits (lag 1) shows a wider ICE spread, indicating that the effect of permitting on predicted regime depends on the broader market context in which permits are issued.

The nested time-series cross-validation results in Table 5 provide additional evidence that these findings are stable across diverse historical sub-periods. Construction Put in Place appears as a selected predictor in all five historical windows, confirming that construction activity is the most consistently informative supply-side regime signal regardless of which market episode dominates the estimation window. No macro-financial variable enters the selected feature set in any fold, providing cross-period confirmation that interest rates, unemployment, and savings conditions do not contribute incremental regime-discriminatory signal at the one-month-ahead horizon once construction dynamics and regime persistence are accounted for. Future research could formally extend this framework by computing episode-specific SHAP contributions to compare how the relative importance of construction activity, permitting, and cost pressures evolves across structurally distinct market phases, further enhancing the transparency and practical utility of AI-driven housing market regime forecasting.

6. Conclusions

This study examined expansion and deceleration regimes in the U.S. housing market using an interpretable machine learning framework that integrates construction-related indicators and autoregressive regime persistence features to generate one-month-ahead regime forecasts. Rather than forecasting continuous house price levels, the analysis focused on identifying discrete cyclical regime states, which more directly reflect market conditions relevant to construction planning, risk management, and policy decision-making. Housing market regimes were defined from observed monthly changes in the FHFA HPI, using a median-based classification consistent with retrospective regime-identification approaches in the housing-cycle literature. Feature selection and hyperparameter tuning were conducted entirely within the training sample using time-series cross-validation, ensuring that all reported performance metrics reflect genuine out-of-sample generalization.

Across the evaluated models, tree-based ensemble methods consistently outperformed the linear baseline, with Random Forest achieving the strongest holdout AUC (0.932) and the most consistent deceleration detection across nested cross-validation folds (mean Recall = 0.823, SD = 0.078). These results confirm Hypothesis 1 by demonstrating that tree-based classifiers provide meaningfully higher discriminatory performance than the linear baseline, and reinforce the value of ensemble learning for modeling housing market dynamics that exhibit asymmetric responses and short-term regime transitions rather than smooth price evolution.

The application of explainable AI techniques played a central role in bridging predictive accuracy and economic interpretability. SHAP analysis, together with partial dependence and ICE analyses, revealed a clear two-tier importance structure: autoregressive regime persistence features dominated the ranking, reflecting the strong month-to-month continuity of housing market regime states, while construction-sector indicators, particularly Construction Put in Place at lags 2 and 3 and Building Permits, contributed complementary supply-side signal capturing the delayed propagation of construction pipeline momentum into housing price dynamics. The feature-block ablation confirmed that construction indicators achieve a training CV AUC of 0.849 compared to 0.665 for macro-financial indicators alone, providing direct model-based evidence for Hypothesis 2. No macro-financial variable entered the selected feature set under any specification, indicating that construction activity and regime persistence together provide a sufficient discriminatory signal for one-month-ahead regime prediction.

Several limitations define the scope of the present study. The analysis relies on national-level data and does not explicitly model regional or metropolitan heterogeneity in housing market dynamics. While the national framework is well-suited for capturing system-wide conditions relevant to integrated capital markets and construction supply chains, substantial spatial heterogeneity characterizes U.S. housing markets, with cyclical behavior, supply constraints, and price volatility varying considerably across states and metropolitan areas. National aggregation may therefore mask regional variation in regime timing and intensity, and extending the approach to sub-national markets requires a fundamentally different modeling framework accounting for localized regulatory environments, demographic conditions, and land-use constraints; such extensions represent an important direction for future research.

The study period from 1993 to 2025 encompasses several structurally distinct market episodes, including the 2020–2022 pandemic-era surge, driven by demand displacement and supply-chain disruptions rather than by conventional housing market fundamentals. The inclusion of this anomalous episode may affect the generalizability of model weights to future periods with different underlying mechanisms, and practitioners should exercise caution when applying the framework to conditions that differ substantially from those observed during the training period. Additionally, the analytical sample of 383 monthly observations spans a limited number of major downturn cycles, constraining the model's exposure to diverse historical regime-transition mechanisms. To empirically assess subsample sensitivity, nested time-series cross-validation was conducted across five non-overlapping historical windows spanning the pre-crisis expansion, the global financial crisis, the post-crisis recovery, the low-rate expansion, and the pandemic and rate-shock correction periods. Random Forest achieved consistent performance across all five episodes, and Construction Put in Place appeared as a selected predictor in every fold, confirming that the core findings are not driven by any single historical episode. Future work should nonetheless extend the framework to longer historical series or cross-country data to further strengthen robustness and generalizability across a wider range of downturn mechanisms.

The regime definition is based on the full-sample distribution of HPI_Change, which characterizes historical market states retrospectively; future work could explore real-time monitoring frameworks that account for publication lags and data revision cycles to support operational early-warning applications. Future research could also evaluate forecasting horizons beyond $t + 1$ to assess how far in advance construction-related indicators retain predictive value, and explore hybrid frameworks that combine econometric regime-detection models, such as Markov-switching specifications, with machine learning classifiers to better integrate structural inference with predictive performance.

Overall, this study contributes methodologically by demonstrating how AI-driven classification and explainable machine learning can be applied to housing market regime forecasting with rigorous temporal validation, and substantively by identifying construction-sector indicators and regime persistence as the dominant signals of one-month-ahead cyclical regime prediction. By emphasizing transparency, feature interpretability, and economic grounding, the proposed framework provides a foundation for future research at finer spatial scales and supports the use of AI-based tools for monitoring and managing cyclical risk in housing and construction markets.

Author Contributions: Conceptualization, M.B. and J.-H.L.; methodology, M.B.; validation, M.B. and J.-H.L.; formal analysis, M.B.; investigation, M.B.; resources, M.B. and J.-H.L.; data curation, M.B. and J.-H.L.; writing—original draft preparation, M.B. and J.-H.L.; writing—review and editing, M.B. and J.-H.L.; visualization, M.B.; supervision, M.B.; project administration, M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data used in this study are sourced from publicly available repositories. U.S. House Price Index and macroeconomic indicators were obtained from the Federal Reserve Economic Data (FRED) repository (https://fred.stlouisfed.org). Engineering News-Record (ENR) cost indices were obtained from ENR (https://www.enr.com). The analysis code is available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

U.S.	United States
GDP	gross domestic product
SHAP	SHapley Additive exPlanations
ANN	artificial neural network
SVM	support vector machine
XGBoost	eXtreme Gradient Boosting
RMSE	root mean square error
SVR	support vector regression
FHFA	Federal Housing Finance Agency
HPI	House Price Index
RFE	Recursive Feature Elimination
VIF	Variance Inflation Factor
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve
ICE	Individual Conditional Expectation
GB	Gradient Boosting
LightGBM	Light Gradient Boosting Machine
RF	Random Forest
PD	partial dependence
CPI	Consumer Price Index
ENR	Engineering News-Record
CCI	Construction Cost Index
BCI	Building Cost Index
MPI	Materials Price Index
SLI	Skilled Labor Index
CLI	Common Labor Index
BLS	Bureau of Labor Statistics
XAI	Explainable Artificial Intelligence

References

1. Engerstam, M. Housing Market Dynamics in Sweden: Understanding the Role of Economic and Institutional Factors in Shaping Housing Prices and Supply. *J. Hous. Econ.* **2023**, *45*, 101–115.
2. National Association of Home Builders. Housing's Contribution to Gross Domestic Product. Available online: https://www.nahb.org/news-and-economics/housing-economics/housings-economic-impact/housings-contribution-to-gross-domestic-product (accessed on 31 January 2026).
3. Holly, S.; Pesaran, M.H.; Yamagata, T. A spatio-temporal model of house prices in the USA. *J. Econom.* **2010**, *158*, 160–173. [CrossRef]
4. Agnello, L.; Schuknecht, L. Booms and busts in housing markets: Determinants and implications. *J. Hous. Econ.* **2011**, *20*, 171–190. [CrossRef]
5. Ang, A.; Timmermann, A. Regime changes and financial markets. *Annu. Rev. Financ. Econ.* **2012**, *4*, 313–337. [CrossRef]
6. Hamilton, J.D. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econom. J. Econom. Soc.* **1989**, *57*, 357–384. [CrossRef]

7. Teräsvirta, T. Threshold models. In *Macroeconometrics and Time Series Analysis*; Palgrave Macmillan: London, UK, 2010; pp. 308–316. [CrossRef] [PubMed]
8. Dijk, D.V.; Teräsvirta, T.; Franses, P.H. Smooth transition autoregressive models—A survey of recent developments. *Econom. Rev.* **2002**, *21*, 1–47. [CrossRef]
9. Bello, M.Z.; Sanni, M.L.; Mohammed, J.K. Conventional methods in housing market analysis: A review of literature. *Balt. J. Real Estate Econ. Constr. Manag.* **2020**, *8*, 227–241. [CrossRef]
10. Sharma, H.; Harsora, H.; Ogunleye, B. An Optimal House Price Prediction Algorithm: XGBoost. *Analytics* **2024**, *3*, 30–45. [CrossRef]
11. Xu, K.; Nguyen, H. Predicting housing prices and analyzing real estate market in the Chicago suburbs using Machine Learning. *arXiv* **2022**, arXiv:2210.06261. [CrossRef]
12. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems*; NeurIPS: San Diego, CA, USA, 2017; Volume 30. [CrossRef]
13. Molnar, C. *Interpretable Machine Learning*; Lulu.com: Durham, NC, USA, 2020.
14. Goodhart, C.; Hofmann, B. House prices, money, credit, and the macroeconomy. *Oxf. Rev. Econ. Policy* **2008**, *24*, 180–205. [CrossRef]
15. Wu, J.; Gyourko, J.; Deng, Y. Evaluating conditions in major Chinese housing markets. *Reg. Sci. Urban Econ.* **2012**, *42*, 531–543. [CrossRef]
16. Tu, Q.; de Haan, J.; Boelhouwer, P. The mismatch between conventional house price modeling and regulated markets: Insights from The Netherlands. *J. Hous. Built Environ.* **2017**, *32*, 599–619. [CrossRef] [PubMed]
17. Schellekens, M.; Yasseri, T. Credit Crunch: The Role of Household Lending Capacity in the Dutch Housing Boom and Bust 1995–2018. *arXiv* **2021**, arXiv:2101.00913. [CrossRef]
18. Martínez, C.C. Dating housing booms fueled by credit: A Markov switching approach. *J. Financ. Stab.* **2025**, *78*, 101412. [CrossRef]
19. Igan, M.D.; Loungani, M.P. *Global Housing Cycles*; International Monetary Fund: Washington, DC, USA, 2012.
20. White, M. Cyclical and structural change in the UK housing market. *J. Eur. Real Estate Res.* **2015**, *8*, 85–103. [CrossRef]
21. Sá, F. The effect of foreign investors on local housing markets: Evidence from the UK. *J. Econ. Geogr.* **2024**, *25*, 329–349. [CrossRef]
22. San Ong, T. Factors affecting the price of housing in Malaysia. *J. Emerg. Issues Econ. Financ. Bank.* **2013**, *1*, 414–429. [CrossRef]
23. Kader, S.A.; Zayed, N.M.; Faisal-E-Alam, M.; Salah Uddin, M.; Nitsenko, V.; Klius, Y. Factors Affecting Demand and Supply in the Housing Market: A Study on Three Major Cities in Turkey. *Computation* **2022**, *10*, 196. [CrossRef]
24. Tupenaite, L.; Kanapeckiene, L.; Naimaviciene, J. Determinants of housing market fluctuations: Case study of Lithuania. *Procedia Eng.* **2017**, *172*, 1169–1175. [CrossRef]
25. Knoll, K.; Schularick, M.; Steger, T. No price like home: Global house prices, 1870–2012. *Am. Econ. Rev.* **2015**, *105*, 351–355. [CrossRef]
26. Nam, P.P.; Phuong, T.T. Factors affecting commercial housing prices in Vietnam. *Int. J. Hous. Mark. Anal.* **2022**, *15*, 1019–1032. [CrossRef]
27. Yu, J. A Multivariate Regression Analysis of Factors Influencing California Housing Prices. In *Proceedings of the International Conference on Mathematics and Machine Learning*; ACM: New York, NY, USA, 2023; pp. 165–169. [CrossRef]
28. Adetunji, A.B.; Akande, O.N.; Ajala, F.A.; Oyewo, O.; Akande, Y.F.; Oluwadara, G. House Price Prediction using Random Forest Machine Learning Technique. *Procedia Comput. Sci.* **2022**, *199*, 806–813. [CrossRef]
29. Manasa, K.; Gupta, S.; Narahari, A. House Price Prediction Using Machine Learning Algorithms: A Case Study of Bengaluru. *Int. J. Adv. Sci. Technol.* **2020**, *29*, 12345–12356.
30. Beracha, E.; Wintoki, M.B. Forecasting Residential Real Estate Price Changes from Online Search Activity. *J. Real Estate Res.* **2013**, *35*, 283–312. [CrossRef]
31. Møller, S.V.; Pedersen, T.; Montes Schütte, E.C.; Timmermann, A. Search and predictability of prices in the housing market. *Manag. Sci.* **2024**, *70*, 415–438. [CrossRef]
32. Reinhart, C.M.; Rogoff, K.S. *This Time Is Different: Eight Centuries of Financial Folly*; Princeton University Press: Princeton, NJ, USA, 2009. [CrossRef]
33. Doerner, W.M.; Leventis, A.V. Distressed sales and the FHFA house price index. *J. Hous. Res.* **2015**, *24*, 127–146. [CrossRef]
34. Leventis, A. Revisiting the differences between the OFHEO and S&P/Case-Shiller house price indexes: New explanations. In *Research Paper*; Office of Federal Housing Enterprise Oversight: Washington, DC, USA, 2008.
35. Moews, B.; Herrmann, J.M.; Ibikunle, G. Lagged correlation-based deep learning for directional trend change prediction in financial time series. *Expert Syst. Appl.* **2019**, *120*, 197–206. [CrossRef]
36. Zhang, Y.; Rahman, A.; Miller, E. Longitudinal modelling of housing prices with machine learning and temporal regression. *Int. J. Hous. Mark. Anal.* **2023**, *16*, 693–715. [CrossRef]
37. Zhang, X.; Chen, S.; Xue, J.; Wang, N.; Xiao, Y.; Chen, Q.; Hong, Y.; Zhou, Y.; Teng, H.; Hu, B.; et al. Improving model parsimony and accuracy by modified greedy feature selection in digital soil mapping. *Geoderma* **2023**, *432*, 116383. [CrossRef]

38. Huang, J.; Ling, C.X. Using AUC and accuracy in evaluating learning algorithms. *IEEE Trans. Knowl. Data Eng.* **2005**, *17*, 299–310. [CrossRef]
39. Eusebi, P. Diagnostic accuracy measures. *Cerebrovasc. Dis.* **2013**, *36*, 267–272. [CrossRef] [PubMed]
40. Washington, S.; Karlaftis, M.G.; Mannering, F.; Anastasopoulos, P. *Statistical and Econometric Methods for Transportation Data Analysis*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2020.
41. O'brien, R.M. A caution regarding rules of thumb for variance inflation factors. *Qual. Quant.* **2007**, *41*, 673–690. [CrossRef]
42. He, K.; He, C. Housing price analysis using linear regression and logistic regression: A comprehensive explanation using melbourne real estate data. In *2021 IEEE International Conference on Computing (ICOCO)*; IEEE: Piscataway, NJ, USA, 2021; pp. 241–246. [CrossRef]
43. Hosmer, D.W., Jr.; Lemeshow, S.; Sturdivant, R.X. *Applied Logistic Regression*; John Wiley & Sons: Hoboken, NJ, USA, 2013.
44. Saefudin, M.R.; Putri, M.R.; Hadi, A.; Wijayanto, H.; Irmawati, B. Significant Features for House Price Prediction Using Machine Learning. In *2024 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT)*; IEEE: Piscataway, NJ, USA, 2024; pp. 659–664. [CrossRef]
45. Ng, A.Y. Feature selection, *L1* vs. *L2* regularization, and rotational invariance. In *Proceedings of the Twenty-First International Conference on Machine Learning*; ACM: New York, NY, USA, 2004; p. 78. [CrossRef]
46. Wang, C.; Wu, H. A new machine learning approach to house price estimation. *New Trends Math. Sci.* **2018**, *6*, 165–171. [CrossRef]
47. Almaslukh, B. A gradient boosting method for effective prediction of housing prices in complex real estate systems. In *2020 International Conference on Technologies and Applications of Artificial Intelligence (TAAI)*; IEEE: Piscataway, NJ, USA, 2020; pp. 217–222. [CrossRef]
48. Jiang, F.F. Gradient Boosting Decision Tree for House Price Prediction with Google Trends. *J. Econ. Manag. Sci.* **2025**, *8*, 117. [CrossRef]
49. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; Association for Computing Machinery: New York, NY, USA, 2016; pp. 785–794. [CrossRef]
50. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. LightGBM: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems*; NeurIPS: San Diego, CA, USA, 2017; Volume 30, pp. 3146–3154.
51. Li, T.; Akiyama, T.; Wei, L. Constructing a highly accurate price prediction model in real estate investment using LightGBM. In *2021 IEEE 4th International Conference on Multimedia Information Processing and Retrieval (MIPR)*; IEEE: Piscataway, NJ, USA, 2021; pp. 273–276. [CrossRef]
52. John, L.M.; Shinde, R.; Shaikh, S.; Shar, D. Predicting House Prices using Machine Learning and LightGBM. In *Proceedings of the 7th International Conference on Innovations and Research in Technology and Engineering (ICIRTE-2022)*; VPPCOE & VA: Mumbai, India, 2022. [CrossRef]
53. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [CrossRef] [PubMed]
54. Goldstein, A.; Kapelner, A.; Bleich, J.; Pitkin, E. Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *J. Comput. Graph. Stat.* **2015**, *24*, 44–65. [CrossRef]
55. Trindade Neves, F.; Aparicio, M.; de Castro Neto, M. The impacts of open data and eXplainable AI on real estate price predictions in smart cities. *Appl. Sci.* **2024**, *14*, 2209. [CrossRef]
56. Kansal, M.; Singh, P.; Shukla, S.; Srivastava, S. A comparative study of machine learning models for house price prediction and analysis in smart cities. In *International Conference on Electronic Governance with Emerging Technologies*; Springer Nature: Cham, Switzerland, 2023; pp. 168–184. [CrossRef]
57. Aniobi, D.E.; Ochuba, C.O.; Nguideen, S.B. House Price Prediction: Comparative Analysis of Regression-Based Machine Learning Algorithms. *Int. J. Res. Appl. Sci. Eng. Technol.* **2023**, *11*, 1550–1557. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.