İstanbul İktisat Dergisi
Istanbul Journal of Economics,
IJE 2025, 2025 (2): 432–452

https://doi.org/10.26650/ISTJECON2025-1707729

Submitted: 28.05.2025
Revision Requested: 19.11.2025
Last Revision Received: 25.11.2025
Accepted: 29.11.2025
Published Online 22.12.2025

İstanbul İktisat Dergisi
Istanbul Journal of Economics

Research Article

Open Access

From Causes to Forecasts: Granger Causality and Machine-Learning Predictions of Housing Sales in Türkiye

Musa Gün [1], Ahmet Akusta [2] & Haydar Karadağ [3]

[1] Recep Tayyip Erdoğan University, Faculty of Economics and Administrative Sciences, Department of Accounting and Finance, Rize, Türkiye
[2] Konya Technical University, Rectorate, Konya, Türkiye
[3] Recep Tayyip Erdoğan University, Faculty of Economics and Administrative Sciences, Department of Economics, Rize, Türkiye

Abstract

This study aims to examine the impact of macro-financial indicators on the housing sales volume in Türkiye and evaluate the forecasting performance of traditional time-series models versus modern machine-learning algorithms. Using monthly data from January 2014 to November 2024, the research uses linear and non-linear Granger causality tests to investigate the lagged effects of 13 economic variables, including mortgage rates, loan volume, housing price indices, inflation, industrial production, consumer confidence, and unemployment. The findings reveal statistically significant causal relationships between housing sales and several indicators, notably interest rates, housing prices, consumer prices, and industrial output. These links imply that the cyclical effects of credit conditions, price expectations, or real-sector activity on housing demand operate with relatively short time lags. The predictive performance was evaluated using six models, including a multilayer perceptron, random forest, polynomial regression, gradient boosting, seasonal LSTM, and SARIMAX. Considering the test MAPE, the neural-network models provide the best predictions, with the multilayer perceptron and the seasonal LSTM obtaining MAPEs of 17.3% and 19.7%, respectively. On the other hand, the SARIMAX and tree-type models have worse generalisation capability. The results demonstrate the value of combining causal analysis with advanced forecasting to capture the dynamics of the housing market. They provide a practical framework for anticipating changes in demand and support the integration of machine-learning tools into economic monitoring and policy evaluation. This dual approach enhances the understanding of how economic conditions propagate through the housing sector and contributes to more informed housing market governance.

Keywords Real Estate Prices · Housing Sales · Machine Learning Algorithms

Jel Codes E00, R32, C88

Citation: Gün, M., Akusta, A. & Karadağ, H. (2025). From causes to forecasts: Granger causality and machine-learning predictions of housing sales in Türkiye. *İstanbul İktisat Dergisi–Istanbul Journal of Economics*, *2025*(2), 432-452. https://doi.org/10.26650/ISTJECON2025-1707729

Corresponding author: Musa Gün musa.gun@erdogan.edu.tr

From Causes to Forecasts: Granger Causality and Machine-Learning Predictions of Housing Sales in Türkiye

Housing markets operate at the nexus of macroeconomic stability, household welfare, and sectoral growth. In many emerging economies, the construction–real-estate nexus is viewed as a strategic lever for employment creation and sustained gross domestic product (GDP) expansion. The development of the housing market not only stimulates investment flows into related sectors but also tends to divert financial and public resources away from other areas of the economy. As the housing sector grows rapidly, it increasingly attracts large volumes of capital and government support. However, this concentration of investment can push housing prices above their fundamental values, ultimately contributing to the emergence of speculative price bubbles.

Türkiye illustrates these dynamics. The direct share of housing in GDP fluctuated between 5% and 8% from 2003 to 2017, while the combined contribution of housing, together with its upstream and downstream industries, is estimated at roughly 30% of GDP (Yardımcı, 2021: 381). Household budget surveys show that about 27 % of private consumption is devoted to housing and rents (Karadağ, 2021: 408). Housing construction is estimated to account for approximately 80% of the total construction activity (Yalçın, 2020: 449). Furthermore, it underpins a broad supply chain that spans iron and steel, cement, glass, ceramics, paint, wood, brick, and tile. A large share of the population depends directly or indirectly on the sector for its livelihood (Yardımcı, 2021: 381); accordingly, rising unemployment can reduce GDP and disturb macro-stability, and Okun's Law suggests that a one-percentage-point increase in unemployment lowers GDP by roughly 2.5 percentage points (Karadağ, 2021: 408).

Recent indicators highlight both the scale and volatility of Türkiye's housing market. The House Price Index (HPI) increased from 12.61 in January 2020 to 155.37 in November 2024, and the New and Non-New House Price indices exhibited similarly rapid post-2020 growth, evidencing a sharp rise in prices. Mortgage-loan volume expanded from 188,770,100 TL in 2019 to 447,865,325 TL in November 2024; pandemic-period data confirm that credit volume continued to widen during this interval. Housing sales volume rose from 113,615 units in January 2020 to 153,014 units in November 2024, although sales movements were not entirely in line with price changes. Mortgage interest rates and sales quantities moved in opposite directions. The composite Uncertainty Index (UI) peaked in 2018 and again in 2020, when economic or political fluctuations were pronounced. The Consumer Price Index (CPI) and the Industrial Production Index (IPI) have trended upward since early 2020, with industrial production displaying a steadier rise. The unemployment rate (UR) followed a cyclical path, increasing during the pandemic and subsequently declining. Two sector-linked measures, the Furniture Manufacturing Index (FMI) and the Durable Consumer Goods Index (DCGI), climbed from 81.39 to 103.26 and from 83.49 to 107.73, respectively, between January 2020 and November 2024.

Given this backdrop, accurate and reliable forecasting of housing sales is indispensable. Sales volumes are shaped by market uncertainty, policy shifts, and technological change (Akusta, 2024:114); precise forecasts help preserve the balance between supply and demand (Selci, 2021:30). Traditional forecasting systems often struggle with seasonality and non-linear dynamics, whereas new-generation techniques such as artificial neural networks learn from examples and solve complex problems more effectively than conventional approaches (Selci, 2021: 22).

Although a growing body of Turkish scholarship analyses house-price movements (for example, Karadağ, 2021; Özçim, 2022), relatively few studies investigate housing sales volume using an extensive macro-financial indicator set or compare classical econometric methods with modern machine-learning approaches. Earlier research confirms that lower policy rates reduce borrowing costs and stimulate demand (Özçim, 2022: 524; Akkaya, 2024: 34), while tighter credit conditions and heightened uncertainty dampen mortgage lending and sales (Kılcı, 2019: 98–99). However, a systematic assessment that couples causality testing with an out-of-sample forecasting comparison remains scarce.

This study aims to address a notable gap in the literature by analysing monthly data spanning from January 2014 to November 2024. The analysis focuses on housing sales volume as the dependent variable, while incorporating a comprehensive set of explanatory variables, including housing interest rates, housing loan volume, the overall house price index, sub-indices for new and non-new houses, the CPI, IPI, DCGI, FMI, UR, UI, and consumer confidence index (CCI). What sets this study apart is its finding that housing sales are significantly influenced by lagged housing interest rates, house price dynamics, and inflation indicators, captured through both linear and non-linear time-dependent relationships. Furthermore, the study evaluates the performance of traditional time-series models against advanced machine learning methods, ultimately favouring the latter due to their stronger empirical results within the given dataset.

Two research questions guide the analysis.

RQ1 (Causality): What are the predominant causal pathways that connect macro-financial indicators to Türkiye's housing-sales volume?

RQ2 (Prediction): Which modelling frameworks, traditional time-series techniques, or state-of-the-art machine-learning algorithms provide the most accurate and reliable forecasts of monthly housing sales?

By combining causal estimation with a forecasting benchmark, the paper makes three principal contributions: (a) it maps how credit conditions, price dynamics, and real-sector indicators transmit to housing demand; (b) it offers practical guidance for analysts monitoring market turning points; and (c) it provides evidence on feedback loops between housing demand and industrial production.

The rest of the study is as follows: In Section 2, we review the empirical and theoretical literature. In Section 3, we outline the study's data and methods, including the use of causality tests and machine-learning architectures. Section 4 presents the empirical findings. Finally, in Section 5, we conclude with policy implications and directions for future research.

Literature Review

The housing market constitutes a multidimensional arena in which economic, social, and behavioural forces intersect. Researchers in both advanced and emerging economies have examined price formation, sales dynamics, mortgage credit, unemployment, macro-financial indicators, and diverse forms of uncertainty. Because developments at the micro and macro levels frequently interact, empirical and theoretical studies situate housing outcomes alongside interest-rate policy, credit volumes, inflation, and labour-market conditions.

A decisive turning point for this literature was the 2008 US subprime mortgage crisis. In its wake, scholars began to investigate more closely how economic-policy uncertainty shapes housing valuations and demand. Antonakakis et al. (2015) employed the Economic Policy Uncertainty (EPU) Index, developed by Baker et al. (2012), to demonstrate that post-crisis US price declines were strongly associated with heightened policy ambiguity. Using a dynamic stochastic general equilibrium model, Huang (2017) showed that a single, temporary uncertainty shock could account for 25 % of the observed surge and 60 % of the subsequent collapse in new-home prices.

Subsequent studies have continued to document the influence of uncertainty. Çepni et al. (2019) found that real-estate-specific uncertainty predicts the tail behaviour of US home-sales growth. In Hangzhou, Zhang et al. (2021) observed that rising uncertainty altered both transaction volumes and investment timing. During the COVID-19 pandemic, Petersen (2024) detected sharp upward revisions—22% in Merced and 14.8% in San José—in Zillow's real estate appraisals, while reporting no statistically significant change for Fresno. Recent machine-learning applications further show that predictive performance remains robust even during shock episodes: Mora-Garcia et al. (2022), for instance, document that machine-learning algorithms preserve accuracy under COVID-19 volatility, while Sobieraj and Metelski (2024) utilise SVM classification with recursive feature selection to identify the most influential market determinants in the United States. These findings collectively show that expectations—and hence realised prices—remain highly sensitive to macroeconomic shocks. In parallel, Truong et al. (2020) highlight that relying solely on house price index-based information is insufficient for housing price prediction, showing that incorporating structural, locational, and socio-economic attributes substantially enhances the performance of the ML model.

Another branch of research explores the interaction between housing and labour market performance. Many studies argue that homeownership can curb worker mobility, thereby increasing unemployment. Branch et al. (2016) employed an adaptive-learning model for the period 1996-2010. They found that home-owners often leveraged their dwellings as collateral to smooth consumption, a mechanism that ultimately reduced unemployment through financial innovation. This relationship illustrates that structural obstacles and social vulnerabilities within the housing market have significant repercussions on the labour market, particularly affecting low-income groups. Riley et al. (2015) documented that low-income US homeowners respond less flexibly than tenants to adverse labour-market shocks. Reed and Ume (2016) contended that widespread ownership may dampen geographic mobility within the "American dream" framework. Within a search-and-matching setting for Texas, Li et al. (2018) calculated that a three-percentage-point rise in unemployment led to a 10.74% decline in house prices and a 5.49% contraction in transactions. Evidence from Central and Eastern Europe points in the same direction: Broulíková et al. (2020) linked postprivatisation increases in ownership to more persistent structural unemployment. In China, Sun (2021) reported a long-run positive and short-run negative relationship between housing prices and unemployment. Taken together, these studies indicate, though do not conclusively prove, that labour-market frictions and housing-market rigidities may reinforce one another under certain conditions.

Beyond labour outcomes, scholars have examined the bidirectional link between prices and sales volumes. Empirical evidence from studies conducted across different countries and periods supports the reciprocal relationship between housing prices and sales, whereby rising prices can stimulate demand, and increased sales activity can, in turn, drive prices upward. In the United States, Wheaton and Lee (2009) detected a two-way Granger causality: higher sales raise prices, yet price hikes can, by swelling inventories, depress future sales. In Finland, Oikarinen (2012) found that prices adjust more sluggishly to demand shocks and are negatively correlated with transaction counts. Asset-deterioration effects provide additional nuance. Harding et al. (2007) showed that depreciation and maintenance costs erode the average capital gains. Semi-parametric models incorporating geographic information enhance predictive accuracy

(Bin, 2004). In contrast, structural models emphasising vacant-housing dynamics indicate that supply shocks raise prices, while demand shocks mainly affect sales (Díaz & Jerez, 2013).

A growing body of research highlights a range of macroeconomic and structural factors—including interest rates, exchange rates, unemployment, inflation, construction costs, economic uncertainty, housing loans, and social policy—as key determinants of housing prices and sales. Empirical studies focusing on Türkiye using time-series and panel evidence reveal that interest rates, exchange rates, unemployment, inflation, and construction costs all influence prices and sales (Akkaya, 2024; Özçim, 2022; Karadağ, 2021; Kılcı, 2019; Atasoy & Tursun, 2022).

Inflationary bursts raise input costs and push prices upward, while periods of elevated uncertainty encourage buyers to postpone commitments and developers to delay projects. Yardımcı (2021) emphasised that uncertainty-induced contractions in construction output directly feed into lower housing sales. Consequently, predictable monetary and fiscal policy frameworks are viewed as vital for market stability in emerging economies such as Türkiye.

Spatial heterogeneity is another research frontier. Geographic Information System (GIS) analyses map regional variations in sales intensity, informing urbanisation, transport and infrastructure policy. Yalçın (2020) underscored the value of such spatial perspectives for Türkiye—but also noted that purely descriptive distributional maps limit causal inference for investors and policymakers. Machine-learning-enhanced spatial models are increasingly being used in this area. Caplin et al. (2008) and Barr et al. (2017) combine spatial hedonics and ML to capture geographic price variation; Baldominos et al. (2018) use spatial clustering to detect profitable real-estate investment opportunities; Tchuente and Nyawa (2022) show that geocoding significantly improves price estimation across French cities; and Rico-Juan and de la Paz (2021) find that ML surpasses hedonic methods in the Spanish market while benefiting from interpretability-driven explainability frameworks. Recent evidence further stresses that heterogeneous spatial features—such as accessibility, density, and land-use configurations—have asymmetric effects on housing values. Chen et al. (2023) demonstrated that ML frameworks are particularly well-suited to capturing these nonlinear and place-specific relationships.

Forecasting techniques have evolved alongside data availability. In recent years, alongside traditional econometric approaches, artificial intelligence-based and advanced forecasting models have been commonly employed to predict the housing demand and sales in Türkiye. This growing trend reflects a broader recognition of the value these modern tools bring to capturing complex market dynamics more accurately. Using data from 2013 to 2019, Selci (2021) developed an approach that leveraged an artificial neural network, incorporating the housing-price index, loan, and exchange rates; the model achieved a high level of predictive accuracy. Akusta (2024) compared ARIMA, LSTM, grey-forecasting, and hybrid approaches, reporting that the hybrid and AI-based specifications achieved powerful predictive results relative to the individual classical models. Complementing these findings, a central stream of international research highlights the comparative performance of various ML algorithms. Adetunji et al. (2022) showed that random forest models achieve high predictive accuracy; Foryś (2022) demonstrated that neural networks outperform classical regressions; and studies by Varma et al. (2018), Viktorovich et al. (2018), and Rafiei and Adeli (2016) further emphasised the superior learning capabilities of neural networks and advanced hybrid models in estimating housing sale prices. Park and Bae (2015) confirmed that ML techniques surpass expert-system-based applications, particularly in large datasets, such as those found in Fairfax County and Virginia. Moreover, Türkiye-specific evidence by Çılgın and Gökçen (2023) and Erkek et al. (2020) shows that decision trees, SVMs, and other ML approaches produce reliable and often superior sales-price forecasts relative to standard econometric models. These insights are consistent with those of Truong et al. (2020), who found that ensemble-based and stacked ML models outperform simpler algorithms, while also noting that complex model architectures may impose substantial time and cost burdens in practical deployment. Similarly, Yazdani (2021) demonstrated that combining deep-learning structures with hedonic modelling significantly boosts accuracy and interpretability, underscoring the advantages of hybrid architectures in capturing nonlinearities and multidimensional interactions.

The housing sector literature typically examines key variables, including housing sales volumes, sales prices, housing loans, interest rates, and the consumer price index. Building on this foundation, unlike much of the existing literature on the housing sector, this study introduces a novel approach. It broadens the explanatory set by incorporating the Industrial Production Index, Durable Consumer Goods Index, Furniture Production Index, unemployment rate, Uncertainty Index, and Consumer Confidence Index for Türkiye. When the economy becomes uncertain and inflation becomes more volatile, people start to lose confidence. In such situations, households often stop directing their savings to productive areas, such as industry or durable goods. Instead, they move their savings into housing. They view housing as both a place to live and a safer investment option. Explicitly modelling these linkages can clarify how confidence shocks propagate through both real-sector output and property markets.

Overall, the literature consistently shows that housing market outcomes are shaped by a complex interplay of macroeconomic, financial, behavioural, spatial, social, and policy-related factors. Housing serves both as a fundamental human necessity and a strategic asset, often viewed as a form of investment and social security. Housing policies have different dimensions. Policies in this area cannot rely only on economic tools such as interest rates or the amount of available credit. They also need to consider factors such as urban planning, population shifts, and the distribution of income across society.

For policymakers, this complexity highlights the need for integrated, evidence-based strategies. Although AI-based forecasting tools, such as neural networks and machine learning models, have improved predictive accuracy, their limited transparency remains a challenge. Nevertheless, their growing use in Türkiye has contributed to more informed housing market decisions. These tools enable the early detection of trends, supporting timely and data-driven actions by both public and private actors. Ultimately, the research underscores that scientifically grounded, interdisciplinary approaches can lead to more effective and sustainable housing policy and market governance.

Methodology

This study investigates the causal relationship between economic indicators and housing sales volume and compares modelling forecasting for housing sales. It consists of two main steps: data analysis and predictive modelling.

In the first stage, the objective is to identify the relationships between the variables. Using the Granger causality model, several lagged-way causal relationships from some economic indicators to housing sales volume are identified. The dataset is the set of monthly values of economic indices and housing volume from January 2014 to November 2024, and it is subjected to heavy pre-processing to make the series stationary and usable for analysis.

The second stage is the development of predictive models to predict the volume of housing sales. Several machine learning and time-series models were trained and evaluated to determine the best approach. These models implement techniques to deal with temporal dependencies, seasonality, and overfitting, which makes them robust models.

Data Analysis

This study relies on a dataset that contains monthly values of 13 economic indices for the period from January 2014 to November 2024, along with the target variable, housing sales volume.

Stationarity Assessment

Granger causality analysis necessitates stationarity in the time series data so that inferences are valid (Papana, 2022). The augmented Dickey-Fuller (ADF) test was applied to verify the stationarity of the series at the 0.05 level. Initially, most variables were non-stationary. We applied various transformations to achieve stationarity, as summarised in Table 1.

Table 1

Stationarity check and differencing levels

Variable	ADF Statistic Before	p-Value Before	Stationary at the Differencing Level	ADF Statistic After	p-Value After
Housing Sales Volume	−2.845	0.052	1	−5.089	0.00001
Housing Interest Rates	−1.039	0.739	1	−6.673	0.00000
Housing Loan Volume (000) TL	1.640	0.998	1	−5.042	0.00002
Housing Price Index	−0.290	0.927	2	−4.529	0.00017
New Housing Price Index	0.560	0.987	2	−5.113	0.00001
Non-new housing price index	−0.412	0.908	1	−5.212	0.00001
CPI	−0.829	0.812	1	−4.789	0.00003
Industrial Production Index	0.421	0.982	1	−5.334	0.00000
Durable Consumer Goods Index	−1.212	0.666	1	−4.987	0.00001
Furniture Manufacturing Index	−0.643	0.863	1	−4.890	0.00001
Unemployment Rate (%)	−1.154	0.698	1	−5.102	0.00001
Uncertainty Index	−0.472	0.894	1	−5.234	0.00001
Consumer confidence index	−0.294	0.926	1	−4.993	0.00001

Granger Causality Testing

Granger causality tests whether past values of a time series variable help predict another variable; the test checks whether one variable changes the other. Based on the past data, predictions can be made about the future, and it is commonly used to understand causal relationships in empirical analysis (Oner, 2022). In this study, the AIC is preferred to determine the optimal lag length for each variable. With the lags, the Granger causality is also examined to demonstrate how various economic indicators influence housing sales volume.

Table 2

Linear Granger Causality Analysis

Variable	Selected Lag	p-Value	Granger-Causal
Housing Interest Rates	4	0.0426	Yes
Housing Price Index	1	0.0254	Yes
Non-new housing price index	1	0.0112	Yes
CPI	1	0.0415	Yes
Industrial Production Index	1	0.0189	Yes
Durable Consumer Goods Index	1	0.0312	Yes
Furniture Manufacturing Index	1	0.0245	Yes
Unemployment Rate (%)	1	0.0478	Yes
Consumer confidence index	1	0.0376	Yes

The findings, summarised in Table 2, show a few important causes. There is a 4-month lag for Granger's cause of housing interest rates, which indicates that changes in interest rates influence housing sales with a 4-month lag. This lag may stem from the time it takes for financial decisions and transactions to adjust to the Variation in borrowing costs. That is, the Housing and Non-New Housing Prices Indices find causality at a lag of 1 month, indicating the immediate effect of price changes on sales volume. This evidence implies that both the buying and selling sides respond promptly to changes in house prices.

In addition to housing market variables, a few other macroeconomic variables also Granger-cause housing market variables with a 1-month lag. These are price levels (CPI), production index, durable goods for consumers, furniture manufacturing, unemployment rate, and consumer confidence index. This finding serves as a reminder of the significant connection between the real estate economy and the broader economy.

Table 3

Nonlinear Granger Causality Analysis

Variable	Selected Lag	p-Value	Granger-Causal
Housing Price Index	3	0.026320	Yes
New Housing Price Index	3	0.014566	Yes
Non-new housing price index	3	0.022840	Yes
CPI	5	0.013768	Yes
Industrial Production Index	4	0.002169	Yes

The possibility of using linear Granger causality to analyse complex stochastic processes with nonlinear dynamics or higher-order statistical properties is limited. A more recent methodology employing conditional independence is increasingly being used (Seth and Principe, 2012) and allows one to infer non-linear Granger causality relationships without making assumptions on the underlying process.

Additional insights were obtained from nonlinear Granger causality analysis (Table 3). Such a way of working, which implements the model described here, is a Python package provided by Rosoł et al. (2022); however, it presents more complex dynamics than these results. For instance, each of the three types of house price indices—composite, new, and non-new—has significant non-linear Granger-causal links under a 3-month lag. This indicates that price volatility interacts with the market of housing sales in a manner that linear models do not entirely explain.

The CPI and IPI also showed nonlinear causality with longer lags of 5 and 4 months, respectively. These findings indicate that broader economic conditions influence housing sales through more intricate, time-dependent mechanisms. The combination of linear and non-linear Granger causality results provides a nuanced understanding of the relationships between economic indicators and the housing market.

Machine Learning Modelling

For predicting the volume of housing sales, several machine learning and deep learning models were used and compared for their performance measures. The workflow encompasses crucial phases, including data preprocessing, model training, hyperparameter tuning, and validation, to ensure reliable and accurate predictions.

Data Preprocessing

The data spans the years 2014–2024 and was obtained from the Central Bank of the Republic of Türkiye, Turkish Statistical Institute, and worlduncertaintyindex.com. The dataset was then partitioned into training (80%) and testing (20%) sets without shuffling to preserve the time-series structure. Blank parts of the dataset were supplemented with interpolations. Employing a common stochastic trend, we propose a linear interpolation method to fill in the missing values in the nonstationary panel data models. The method is an easy and fast algorithm compared to the best approaches typically used in the statistical literature (Kang, 2011). The features are then standardised using the StandardScaler, ensuring that each variable is on the same scale. The number of standard deviations a particular score has from the mean is called the Z-score, also known as the standard score. It is computed based on the average and standard deviation of the statistics set (Santhanakrishnan and Senthooran, 2022).

Model Selection and Training

All machine-learning models are implemented in Python 3.11. Tree-based models and polynomial regression were estimated using the scikit-learn library. The multi-layer perceptron (MLP) and seasonal long short-term memory (LSTM) networks were built with TensorFlow/Keras, and the SARIMAX specification was estimated using the statsmodels package.

We focus on six widely used forecasting models that cover the prominent modelling families. Polynomial regression with ridge regularisation provides a simple linear baseline. Random forest and gradient boosting represent tree-based ensemble methods that can capture nonlinearities and interactions. The MLP and the seasonal LSTM are artificial neural networks, with the former being a feed-forward architecture and the latter a recurrent network designed for sequential data. Finally, SARIMAX serves as a traditional benchmark for time-series analysis.

The data were split into a training set and a test set in chronological order. The first 80% of the monthly observations were used to estimate the models, and the remaining 20% were reserved for out-of-sample evaluation. No shuffling is applied, so the time-series structure of the data is preserved. We trained and tested the following models via hyperparameter tuning and optimisation for the best accuracy:

Table 4
Models

Model	Architecture/Key Parameters
Polynomial Regression	• **Pipeline**: PolynomialFeatures(degree=1) → StandardScaler → Ridge(alpha=100) - **Best Parameters**: 'poly_features__degree': 1, 'regressor__alpha': 100}
Random Forest	• **Best Parameters**: • n_estimators=200 • max_depth=20 • min_samples_split=2 • min_samples_leaf = 2 (Cross Validation)
Gradient Boosting	• **Best Parameters**: • n_estimators=300 • learning_rate=0.01 • max_depth = 3 (Cross Validation)
MLP	• **Layers**: 64 → 32 → 16 → 1 - **Activations**: ReLU - **Regularisation**: L2(0.001) (on selected layers) - **Optimiser**: Adam (learning_rate = 0.0005 or tuned) - **Epochs** ~200 (EarlyStopping, ReduceLROnPlateau)
Seasonal LSTM	• **Architecture**: LSTM(128) → BatchNorm + Dropout(0.2) → LSTM(64) → BatchNorm + Dropout(0.2) → LSTM(32) → Dropout(0.2) → Dense(1) - **Regularisation**: L2(0.001) on the LSTM layers - **Optimiser**: Adam - **Epochs** ~300, batch_size=32 (EarlyStopping, ReduceLROnPlateau) - **Seasonality**: cyclical = 12
SARIMAX	• **Best Parameters**: **Order** = (1, 2, 0), **Seasonal_Order** = (0, 1, 0), **Seasonal_Period** = 12

The hyperparameters (such as learning rates, regularisation, and network complexity) are tuned during the training. Early stopping and learning rate reduction techniques are used to prevent overfitting, especially in neural networks.

Addressing Overfitting

Overfitting is a fundamental problem in supervised machine learning, which leads a model to perform well on the training data but poorly on previously unseen test data (Ying, 2019). There are many ways to prevent overfitting and make the model more robust.

Regularisation is an example of such a strategy, as it penalises larger weights, which stifles model flexibility and compels the model to understand simple data patterns, rather than memorise specific features (Ying, 2019). Thus, we use L2 regularisation on the MLP and seasonal LSTM for the risk reduction of overfitting that tends to constrain the model complexity.

Early stopping observes the validation performance during training and terminates when no further improvement is made, which enables the model to avoid overfitting and retain its generalisation capacity (Ying, 2019). In this work, the concept of early stopping is employed in the training process of the MLP and Seasonal LSTM models, thereby preserving the generalisation capability of the models.

In our setting, the 20% hold-out block that we describe as the “test set” also plays the role of a validation set for the MLP and the Seasonal LSTM. Early stopping and learning-rate reduction are monitored on this block. This helps stabilise training and reduce overfitting, but it also means that the reported test errors for these two models may be slightly optimistic compared to a completely unseen test sample. The training errors, in turn, reflect models that are kept intentionally simple to favour out-of-sample performance.

Cross-validation is another important method in which a model can be evaluated for generalisation. By splitting the dataset into multiple splits for training and validation, the cross-validation allows a robust estimation of how well the model will perform on new data, and even the possibility of detecting overfitting (Montaha et al., 2022). Deeper 5-fold cross-validation is applied on both random forest and gradient boosting to obtain reliable performance estimates and generalise better.

In deep learning models, dropout is often used to reduce overfitting. During training, the method randomly turns off some neurons. By doing this, the model is prevented from relying too heavily on any single group of neurons, which enables it to generalise more effectively (Liu et al., 2023). The seasonal LSTM model includes dropout layers.

In addition, learning-rate reduction techniques help control the training process. When the model's performance on the validation set stops improving, these methods lower the learning rate. This slows the update steps just enough to support stable convergence while also helping limit overfitting (Macêdo et al., 2024). In the next section, we use a learning rate scheduler for the MLP to dynamically adjust the learning rate, incrementally updating the model weights and preventing overfitting.

Feature scaling, also called normalisation, contributes to the stability of the training procedure by normalising attribute ranges (De Amorim et al., 2024). The study employs StandardScaler to standardise input features, ensuring balanced contributions to the models' predictions and enhancing the training stability.

The models exhibit enhanced generalisation, and overfitting is reduced when these approaches are employed (especially for MLP and Seasonal LSTM). We observe this through their strong generalisation performance on training and testing datasets.

Results

Polynomial Regression

The polynomial regression model acts with moderate accuracy as a baseline predictor and does not outperform the more complicated models. Having a train mean absolute percentage error (MAPE) of 18.61%, it is successful enough to follow general patterns in the training data. However, the out-of-sample MAPE of 22.85% denotes a significant deterioration in the predictive performance, illustrating the difficulty faced by the model when confronted with unseen observations. This contrast implies that although the model can explain simple structures, it fails to approximate the sophisticated correlations and variability of housing sales.

The reported train/test MAPE discrepancy agrees with other performance measures. For example, the test set's mean squared error (MSE) of 989,839,714.5 is much larger than the train set's MSE of 765,523,888.5. Likewise, the root mean squared error (RMSE) rises from 27,668.10 in training to 31,461.71 in testing. These scores show that the polynomial regression overfits and does not well generalize compared with the more flexible models.

Table 5

Performance metrics of the polynomial regression model

Metric	Train	Test
MAE	19,408.46	25,696.51
MSE	765,523,888.5	989,839,714.5
RMSE	27,668.10	31,461.71
MAPE	0.1861	0.2285

Figure 1

Predicted as Actual Values for Polynomial Regression

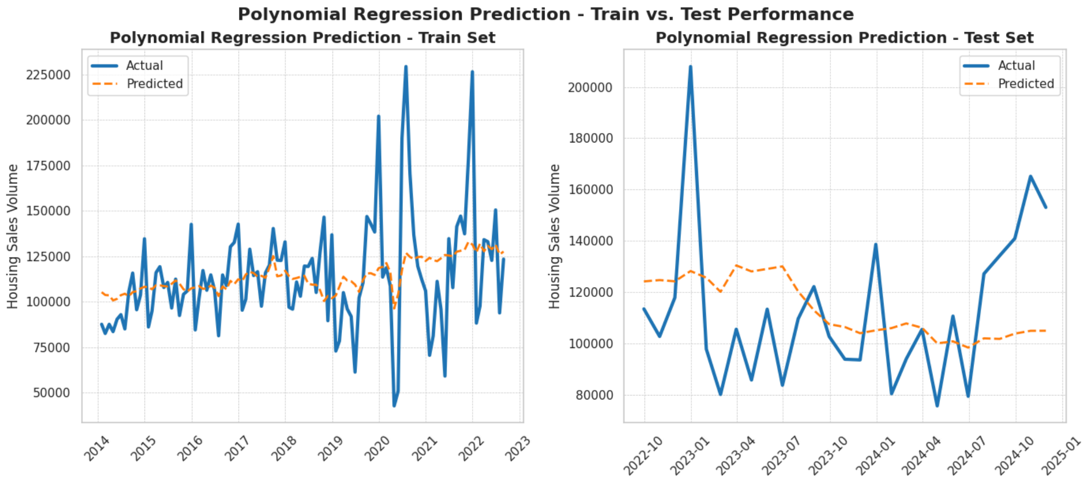


Figure 1 illustrates this weakness visually. The predicted values from the polynomial regression model often deviate substantially from the actual housing sales, especially around sharp peaks and regime shifts. This visual inconsistency aligns with the high RMSE and MAPE values shown in Table 5. These errors imply that a basic polynomial model is unable to track the sharp nonlinear movements and sudden shifts observed in Turkish housing sales. Therefore, we treat polynomial regression mainly as a baseline for comparison rather than as a suitable model for operational forecasting.

Random Forest

The random forest model performs best at capturing the localised patterns and nonlinear relationships in the training set, with an MAPE of 7.97%. Its capability to partition the feature space results in its adaptability to the complexity of the data and provides robust performance during training. However, the strength comes at the expense of generalisation, as indicated by the test set MAPE, which abruptly increases to 30.23%. This difference emphasises the inability of the model to fit the variability and seasonality pattern of housing sales that is present in the new data. The high increase of the test set RMSE (from 12,737.72 to 35,017.70) also highlights the difficulty of detecting the temporal dependencies for the model.

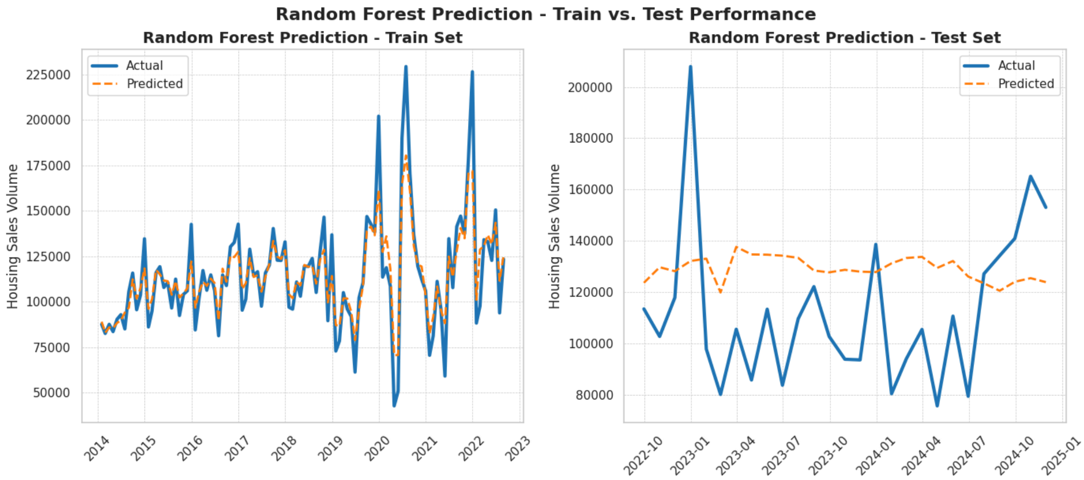
Table 6

Performance metrics of the random forest model

Metric	Train	Test
MAE	8,439.07	30,753.62
MSE	162,249,655.1	1,226,239,469.9
RMSE	12,737.72	35,017.70
MAPE	0.0797	0.3023

Figure 2

Predicted as Actual Values for Random Forest



Gradient Boosting

Gradient Boosting demonstrated moderate improvements over Random Forest, particularly in balancing the training and test errors. The training MAPE of 9.67% indicates that the model efficiently reduced the residual errors during training. On the test set, the MAPE of 29.30% reflects a marginal improvement over Random Forest, although it still reveals challenges in accurately modelling unseen data. The test set RMSE of 35,220.72, which is only slightly better than the RF, shows that the model faced similar struggles in accounting for seasonality and nonlinear trends. Gradient Boosting generalises better than simpler models, such as polynomial regression, but remains constrained by its lack of explicit temporal modelling.

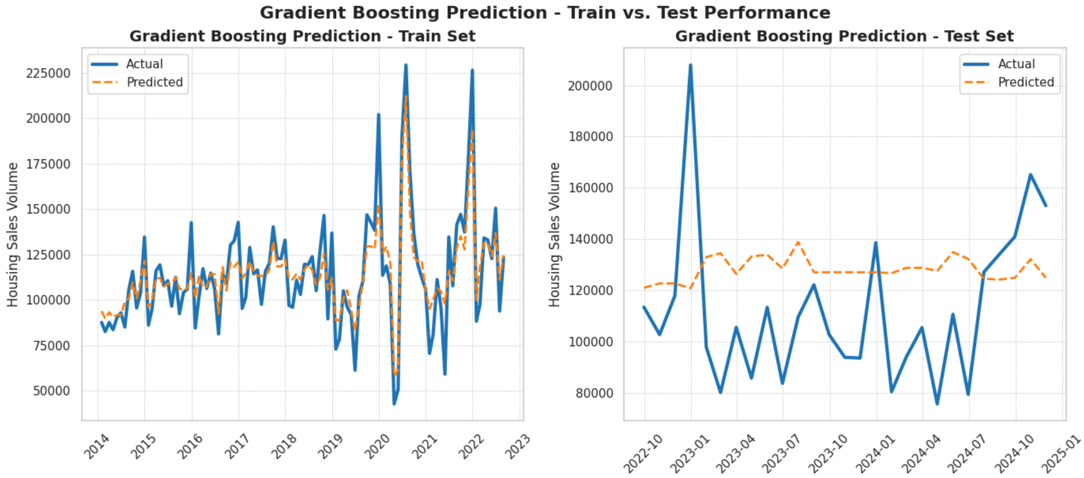
Table 7

Performance Metrics of the Gradient Boosting Model

Metric	Train	Test
MAE	10,236.16	29,755.53
MSE	175,205,650.5	1,240,499,719.9
RMSE	13,236.52	35,220.72
MAPE	0.0967	0.2930

Figure 3

Predicted as Actual Values for Gradient Boosting



Multi-Layer Perceptron (MLP)

A notable and distinctive finding in these results is that the test set error (MAPE: 17.30%) is lower than the training error. While the literature typically anticipates the inverse scenario due to the risk of overfitting, we can largely explain this phenomenon by the structural characteristics of the dataset.

A critical factor is the structural characteristics of the dataset. The training set (January 2014 to September 2022) encompasses periods of extraordinary volatility that disrupted market dynamics, such as the 2018 currency shock and the COVID-19 pandemic (2020). The "noise" inherent in these turbulent periods increased the learning cost for the model. In contrast, the test set (October 2022 to November 2024), despite being situated in a high-inflation environment, represents a period in which housing sales settled into a more consistent macroeconomic trend with more distinct seasonality. Consequently, the model successfully abstracted general patterns from the noisy training data, facilitated by the early stopping mechanism, and applied these learned patterns effectively to the structurally more stable test set. This result is consistent with the MLP’s strong out-of-sample performance. However, it should be viewed as a technical consequence of the training setup rather than as evidence of 'exceptional' generalisation.

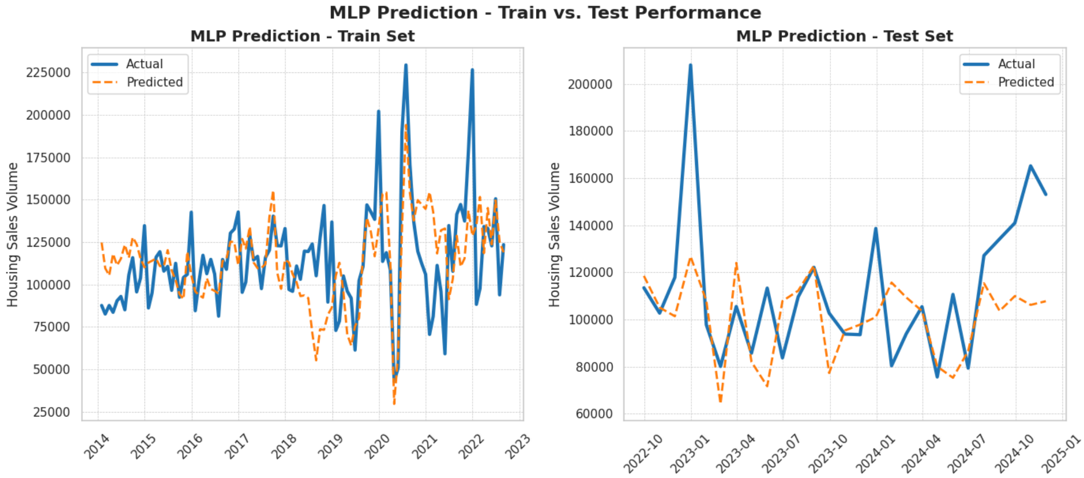
Table 8

Performance metrics of the MLP model

Metric	Train	Test
MAE	23,224.37	21,058.16
MSE	915,277,841.0	829,627,789.1
RMSE	30,253.55	28,803.26
MAPE	0.2171	0.1730

Figure 4

Predicted as Actual Values for MLP



Seasonal LSTM

The Seasonal LSTM model achieves one of the best performances among all models by properly capturing the temporal dependencies and seasonality, and follows the MLP closely in terms of the test MAPE. With a training MAPE of 16.10%, our model is effective in capturing nonlinear trends and seasonal patterns. The test set MAPE of 19.71% demonstrated a strong generalising capability; the model captured the overall seasonal trends more effectively than the benchmark models, such as polynomial regression or random forest.

Despite these strengths, the Seasonal LSTM experienced some difficulty in modelling sharp peaks (test set RMSE of 32,547.84). Our results imply that, although our model is strong enough to identify dynamic and non-linear relationships, additional tuning and/or expansion in the representation space may be needed to mitigate their extreme fluctuations. It models trends and periodic patterns on the training set well; sometimes, however, it cuts off steep peaks. The model predictions are more consistent with long-term seasonal patterns on a test set, and the model outperforms simpler models by capturing more dynamic and nonlinear patterns. Nonetheless, the model struggles with large fluctuations, and further tuning or increasing model complexity is necessary to fully represent the complex dynamics.

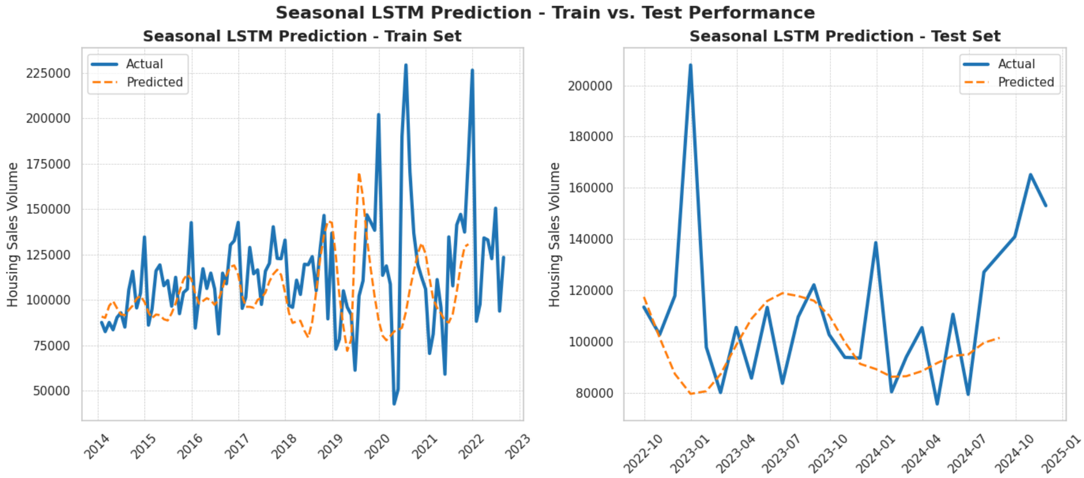
Table 9

Performance metrics of the seasonal LSTM model

Metric	Train	Test
MAE	19,079.23	24,682.42
MSE	628,111,042.4	1,059,361,924.4
RMSE	25,062.14	32,547.84
MAPE	0.1610	0.1971

Figure 5

Predicted as Actual Values for Seasonal LSTM



SARIMAX

The SARIMAX model has a powerful ability to explain seasonality and trend features explicitly, and its MAPE for the training set is 23.18%. It is good at summarising the periodic properties in the training sequences, reflecting a strong capacity for reasoning about the structured temporal dynamics. However, its generalisation ability on unseen data is not as good, as evidenced by the significant increase in MAPE on the test set to 55.63%. This huge jump highlights the model's struggle in handling non-smooth patterns and rapid changes, resulting in less accurate predictions on the test data.

It is more indicative of its instability that the RMSE increases from 33,747.98 in training to 75,996.47 in testing, revealing a problem in the control of the complex dynamics. While SARIMAX is effective at exploiting seasonality in the data, its forecasts on the test set sometimes become unstable, giving off results that may seem unreasonable, negative amounts. These limitations show that the model struggles to generalise when the data contain more complex or irregular time patterns. Even though it can follow well-defined trends, its overall performance remains constrained.

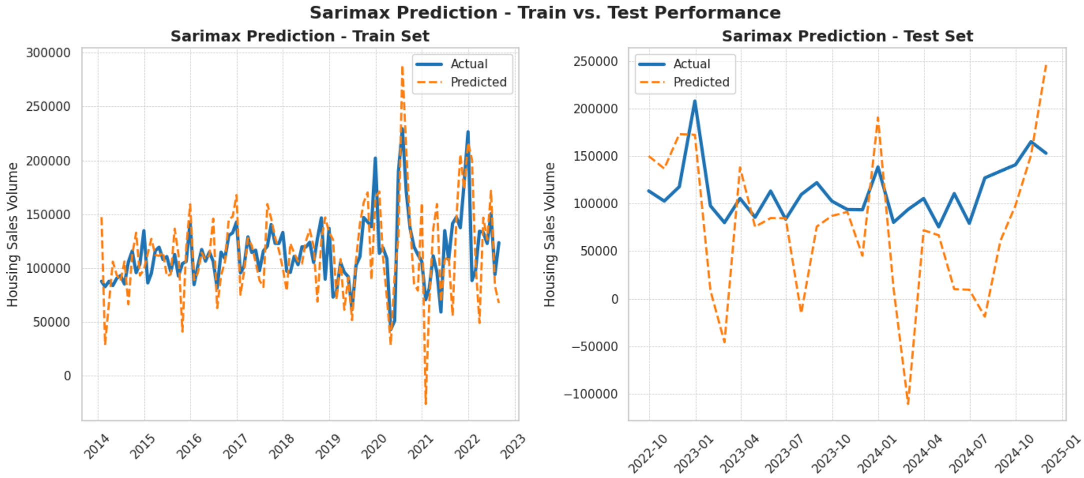
Table 10

Performance metrics of the SARIMAX model

Metric	Train	Test
MAE	24,549.32	59,049.33
MSE	1,138,926,128.3	5,775,464,875.9
RMSE	33,747.98	75,996.47
MAPE	0.2318	0.5563

Figure 6

Predicted as Actual Values for SARIMAX



Conclusion

Predicting housing sales volume is a complex task influenced by non-linear relationships, temporal dependencies, and seasonality. This study examined (a) the causal influence of a broad set of macro-financial indicators on Turkish housing-sales volume (RQ1) and (b) the relative forecasting accuracy of traditional time-series models versus modern machine-learning algorithms (RQ2). Within this analytical framework, the tests' findings provide clear evidence regarding the statistically significant impact of key macro-financial indicators on housing sales.

Linear and non-linear Granger tests show that lagged housing interest rates, the Housing Price Index, consumer prices, and industrial production have statistically significant effects on monthly housing sales. These findings confirm that credit conditions, price expectations, and real-sector activity all influence demand, thus directly addressing RQ1. In the forecasting exercise, the MLP (test MAPE = 17.30 %) and Seasonal LSTM (19.71 %) achieved noticeably lower error rates than the polynomial regression, random forest, gradient boosting, and—in particular—SARIMAX (55.63 %), demonstrating that models capable of learning both non-linearity and sequential structure offer the most reliable out-of-sample guidance (RQ2). In this context, it is essential to assess how these results align with the existing literature.

The central role of credit variables aligns with earlier Turkish evidence on mortgage-rate sensitivity (Karadağ, 2021; Özçim, 2022). Superior neural-network performance aligns with Selci's (2021) and Akusta's (2024) results, while the relatively weak showing of tree ensembles echoes their remarks on the difficulty of those models in handling seasonality.

Accordingly, the policy relevance of these drivers and the practical value of the machine-learning models become even more obvious. Because credit, inflation, and activity indicators emerge as leading drivers, routine monitoring of these series can help analysts and regulators anticipate shifts in housing demand, a need highlighted in the research gap discussion of this paper. Machine-learning forecasts, which outperformed classical benchmarks, therefore constitute a useful complement to existing policy toolkits for evaluating credit risk and real-sector spillovers.

At the same time, the main weakness of the best-performing models is their limited transparency. The MLP and the Seasonal LSTM are high-capacity, nonlinear models, and they do not provide simple, directly interpretable parameters. When these models predict, for example, a fall in housing sales, we cannot clearly state whether this is mainly driven by changes in interest rates, prices, uncertainty, or confidence. This makes it harder to link a single forecast to a specific policy response. In this paper, we therefore use the neural-network models mainly as forecasting tools and rely on the linear and non-linear Granger-causality tests, together with economic reasoning, when we discuss the underlying mechanisms.

Moreover, we enhance the methodological contribution of the study by broadening the set of indicators considered and conducting a comprehensive model comparison. By incorporating industrial production, durable goods, furniture output, unemployment, uncertainty, and consumer confidence indices into both the causality framework and a head-to-head forecasting comparison, the study extends prior work that focused mainly on prices and mortgage variables (Yardımcı, 2021). The joint causal–predictive design offers a more comprehensive map of the Turkish housing market dynamics than either approach in isolation.

Several limitations of the study should be noted. First, the models with the lowest forecast errors (the MLP and the Seasonal LSTM) are also the least transparent, so their internal workings cannot be summarised with simple elasticities or impulse responses. Second, the sample period (January 2014-November 2024) contains major policy shifts and high macro-financial volatility. The performance ranking of the models, and possibly some of the causal relations, may therefore be specific to this period and could look different in calmer times or in other countries. Third, the slightly lower test error of the MLP compared with its training error is partly due to the use of early stopping on the hold-out block and should not be interpreted as evidence of exceptional generalisation. Finally, each modelling approach has its own technical limits: tree-based models have difficulty with strong seasonality, SARIMAX performs poorly in the presence of large structural changes and extreme spikes, and the neural-network results are based on one train–test split and one set of hyperparameters. These issues do not undermine the main conclusions, but they call for some caution when extrapolating the results.

Taken together, these results show that, among all models considered, the MLP is the most effective algorithm for forecasting Turkish housing sales, followed closely by the Seasonal LSTM. At the other end of the spectrum, SARIMAX is clearly the weakest model in our comparison, with the highest test errors across MAE, RMSE, and MAPE.

Finally, based on the study's findings, several directions for future research can be proposed. First, the January 2014-November 2024 sample covers episodes of pronounced macro-financial volatility; additional testing on calmer sub-periods would help assess stability. Second, although neural-network models outperformed their rivals, their internal decision paths remain opaque; future studies could pair explainable AI techniques with the LSTM or MLP architectures used here. Third, hybrid frameworks that combine SARIMAX-style seasonality filters with neural-network residual learner warrant investigation, as they may further improve performance.

Ethics Committee Approval	This article does not contain any studies with animals or human beings performed by any of the authors. For this type of study, formal consent is not required.
Peer Review	Externally peer-reviewed.
Author Contributions	Conception/Design of Study- M.G., A.A., H.K.; Data Acquisition- H.K.; Data Analysis/ Interpretation- M.G., A.A.; Drafting Manuscript- M.G., A.A., H.K.; Critical Revision of Manuscript- M.G., A.A., H.K.; Final Approval and Accountability- M.G., A.A., H.K.
Conflict of Interest	The authors have no conflict of interest to declare.
Grant Support	The authors declared that this study has received no financial support.

Author Details

Musa Gün (Assoc. Prof., PhD.)

[1] Recep Tayyip Erdoğan University, Faculty of Economics and Administrative Sciences, Department of Accounting and Finance, Rize, Türkiye

0000-0002-5020-9342 musa.gun@erdogan.edu.tr

Ahmet Akusta (Lect. Dr.)

[2] Konya Technical University, Rectorate, Konya, Türkiye

0000-0002-5160-3210

Haydar Karadağ (Assoc. Prof., PhD.)

[3] Recep Tayyip Erdoğan University, Faculty of Economics and Administrative Sciences, Department of Economics, Rize, Türkiye

0000-0003-2398-7314

References

Adetunji, A. B., Akande, O. N., Ajala, F. A., Oyewo, O., Akande, Y. F., & Oluwadara, G. (2022). House price prediction using the random forest machine learning technique. Procedia Computer Science, 199, 806-813. doi:10.1016/j.procs.2022.01.100.

Akkaya, M. (2024). Konut fiyat balonu ve konut fiyatını etkileyen faktörlerin analizi: Türkiye uygulaması. Gazi İktisat ve İşletme Dergisi, 10 (1), 33-45. https://doi.org/10.30855/gjeb.2024.10.1.003.

Akusta, A. (2024). Konut Piyasası Trendlerinin Çözümlenmesi: Türkiye'de Toplam Konut Satışlarının Tahmini. Necmettin Erbakan Üniversitesi Siyasal Bilgiler Fakültesi Dergisi, 6 (1), 113-127. doi:10.51124/jneusbf.2024.78.

Antonakakis, N., Gupta, R. André, C. (2015). Dynamic co-movements between economic policy uncertainty and housing market returns. Journal of Real Estate Portfolio Management, 21 (1), 53–60. https://doi.org/10.1080/10835547.2015.12089971.

Atasoy, T. and Tursun, A. (2022). An Analysis of the Housing Market and First Sales of Houses in Turkey Eurasian Business & Economics Journal, 29, 23-40.

Baker, S., Bloom, N. Davis, S. J. (2012). Has Economic Policy Uncertainty Hampered the Recovery? Working Paper, No. 242, The University of Chicago, George J. Stigler Centre for the Study of the Economy and the State, Chicago.

Baldominos, A., Blanco, I., Moreno, A. J., Iturrarte, R., Bernárdez, Ó., & Afonso, C. (2018). Identifying Real Estate Opportunities Using Machine Learning Applied Sciences, 8(11), 2321. doi:10.3390/app8112321.

Barr, J. R., Ellis, E. A., Kassab, A., Redfearn, C. L., Srinivasan, N. N., & Voris, K. B. (2017). Home price index: a machine learning methodology International Journal of Semantic Computing, 11(01), 111-133. doi: 10.1142/S1793351X17500015.

Bin, O. (2004). Prediction comparison of housing sales prices by parametric versus semi-parametric regressions. Journal of Housing Economics, 13 (1), 68-84. doi:10.1016/j.jhe.2004.01.001.

Branch, W. A., Nadeau, N. P. Rocheteau, G. (2016). Financial frictions, the housing market, and unemployment. Journal of Economic Theory, 164, 101-135. https://doi.org/10.1016/j.jet.2015.07.008.

Broulíková, H. M., Huber, P., Montag, J. Sunega, P. (2020). Homeownership, mobility, and unemployment: Evidence from housing privatisation. Journal of Housing Economics, 50, 101728. doi:10.1016/j.jhe.2020.101728.

Caplin, A., Chopra, S., Leahy, J. V., LeCun, Y. and Thampy, T. (2008). Machine learning and the spatial structure of house prices and housing returns Available at: http://dx.doi.org/10.2139/ssrn.1316046.

Central Bank of Republic of Türkiye (CBRT). (2025). Electronic Data Delivery System (EDDS). https://evds2.tcmb.gov.tr/index.php.

Çepi, O., Gupta, R. & Wohar, M. E. (2019). The role of real estate uncertainty in predicting US home sales growth: evidence from a quantiles-based Bayesian model averaging approach Applied Economics, 52 (5), 528–536. doi:10.1080/00036846.2019.1654082.

Chen, Y., Jiao, J., & Farahi, A. (2023). Disparities in affecting factors of housing price: A machine learning approach to the effects of housing status, public transit, and density factors on single-family housing price. Cities, 140, 104432. https://doi.org/10.1016/j.cities.2023.104432.

Çılgın, C., & Gökçen, H. (2023). Machine learning methods for predicting real estate sales prices in Turkey. Revista de la construcción, 22(1), 163-177. https://dx.doi.org/10.7764/rdlc.22.1.163.

De Amorim, L. B. V., Cavalcanti, G. D. C., & Cruz, R. M. O. (2024). Meta-scaler: A meta-learning framework for the selection of scaling techniques. IEEE Transactions on Neural Networks and Learning Systems. doi: 10.1109/TNNLS.2024.3366615.

Díaz, A. and Jerez, B. (2013). House Prices, Sales, and Time on the Market: A Search-Theoretic Framework International Economic Review, 54, 837-872. https://doi.org/10.1111/iere.12019.

Erkek, M. E. H. M. E. T., Çayırlı, K., & Hepsen, A. (2020). Predicting house prices in Turkey using machine learning algorithms Journal of Statistical and Econometric Methods, 9(4), 31-38.

Foryś, I. (2022). Machine learning in-house price analysis: regression models versus neural networks Procedia Computer Science, 207, 435-445. doi:10.1016/j.procs.2022.09.078.

Harding, J. P., Rosenthal, S. S. Sirmans, C. F. (2007). Depreciation of housing capital, maintenance, and house price inflation: Estimates from a repeat sales model. Journal of Urban Economics, 61 (2), 193-217. doi:10.1016/j.jue.2006.07.007.

Huang, X. (Ivy). (2017). Productivity, model uncertainty, and the new home sales price in the U.S. http://dx.doi.org/10.2139/ssrn.3566787.

Kang, W. (2011). Missing-data imputation in nonstationary panel data models. In D. M. Drukker (Ed.), Advances in Econometrics (C. 27, pp. 235-251). Emerald Group Publishing Limited. https://doi.org/10.1108/S0731-9053(2011)000027B007.

Karadağ, H. (2021). Türkiye Ekonomisinde Bankalar Tarafından Verilen Konut Kredileri, Konut Satışları ve İşsizlik Arasındaki İlişki (2010:Q1-2020:Q3). Journal of Social Policy Conferences, 80, 403-422. https://doi.org/10.26650/jspc.2021.80.001609.

Kılcı, E. N. (2019). Konut Kredisi Faiz Oranları ile İpotekli Konut Satışları Arasındaki İlişkinin Analizi; Türkiye Örneği. Turkish Studies - Economics, Finance, Politics, 14 (1), 95-107. https://dx.doi.org/10.7827/TurkishStudies.15033.

Li, G., Wang, P. Zhang, Q. (2018). Market thickness and the impact of unemployment on housing market outcomes Journal of Monetary Economics, 98, 27-49. doi:10.1016/j.jmoneco.2018.04.007.

Liu, Y., Li, Y., Xu, Z., Liu, X., Xie, H. and Zeng, H. (2023). Guided dropout: Improving deep networks without increased computation. Intelligent Automation & Soft Computing, 36(3), 2519-2528. doi:10.32604/iasc.2023.033286.

Macêdo, D., Zanchettin, C., & Ludermir, T. (2024). Sigmoidal learning rate optimiser for deep neural network training using a two-phase adaptation approach. Appl Soft Comput, 167, 112264. https://doi.org/10.1016/j.asoc.2024.112264.

Maciej Rosoł, Marcel Młyńczak, Gerard Cybulski (2022). Granger causality test with nonlinear eural-network-based methods: Python package and simulation study. Computer Methods and Programmes in Biomedicine, 216. https://doi.org/10.1016/j.cmpb.2022.106669.

Montaha S, Azam S, Rafid Akmrh, Islam S, Ghosh P, Jonkman M (2022) A shallow deep learning approach to classify skin cancer using down-scaling method to minimise time and space complexity. PLoS ONE 17(8): e0269826. https://doi.org/10.1371/journal.pone.0269826.

Mora-Garcia, R.-T., Cespedes-Lopez, M.-F., & Perez-Sanchez, V. R. (2022). Housing Price Prediction Using Machine Learning Algorithms in COVID-19 Times. Land, 11(11), 2100. doi:10.3390/land11112100.

Oikarinen, E. (2012). Empirical evidence on the reaction speeds of housing prices and sales to demand shocks Journal of Housing Economics, 21 (1), 41-54. doi:10.1016/j.jhe.2012.01.004.

Oner, S. (2022). The effect of real effective usd/try exchange rate on tourism income: An empirical analysis of Turkey. Journal of Tourism Management Research, 9(2), 97-109. doi:10.18488/31.v9i2.3080.

Özçim, H. (2022). Türkiye'deki konut satışı ile TCMB politika faiz oranı ve konut fiyat endeksi arasındaki ilişkinin analizi. Nevşehir Hacı Bektaş Veli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 12 (1), 523-533. https://doi.org/10.30783/nevsosbilen.1074220.

Papana, A., Kyrtsou, C., Kugiumtzis, D., & Diks, C. (2022). Identification of causal relationships in non-stationary time series with an information measure: Evidence for simulated and financial data. Empirical Economics. doi:10.1007/s00181-022-02275-9.

Park, B. and Bae, J. K. (2015). Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data. Expert systems with applications, 42(6), 2928-2934. doi:10.1016/j.eswa.2014.11.040.

Petersen, A. M. (2024). How much did the pandemic uncertainty affect real-estate speculation? Evidence from the on-market valuation of for-sale versus rental properties. Applied Economics Letters, 1–5. doi:10.1080/13504851.2024.2302898.

Rafiei, M. H. and Adeli, H. (2016). A novel machine learning model for estimating the sale prices of real estate units. Journal of Construction Engineering and Management, 142(2), 04015066. doi:10.1061/(ASCE)CO.1943-7862.0001047.

Reed, R. R. Ume, E. S. (2016). Housing and unemployment: The search for the "American Dream." Journal of Macroeconomics, 48, 72-86. doi:10.1016/j.jmacro.2016.01.001.

Rico-Juan, J. R., & de La Paz, P. T. (2021). Machine learning with explainability or spatial hedonics tools? An analysis of the asking prices in the housing market in Alicante, Spain Expert Systems with Applications, 171, 114590. doi:10.1016/j.eswa.2021.114590.

Riley, S. F., Nguyen, G. Manturuk, K. (2015). House price dynamics, unemployment, and the mobility decisions of low-income home-owners. Housing and the Built Environment., 30, 141–156. https://doi.org/10.1007/s10901-014-9400-y.

Santhanakrishnan, K., & Senthooran, V. (2022). A parallel distributed cluster computing model for Z-score computation regarding Sri Lankan university admissions. In ICARC 2022-2nd International Conference on Advanced Research in Computing: Towards a Digitally Empowered Society (pp. 344–348). doi: 10.1109/ICARC54489.2022.9754093.

Selci, Y. B. (2021). Türkiye'nin Konut Satışı Değerlerinin Yapay Sinir Ağları ile Öngörülmesi. EKOIST Journal of Econometrics and Statistics, 35, 19-32. doi:10.26650/ekoist.2021.35.180033.

Seth, S. and Principe, J. C. (2012). Assessing granger non-causality using nonparametric measure of conditional independence. IEEE Transactions on Neural Networks and Learning Systems, 23(1), 47-59. https://doi.org/10.1109/TNNLS.2011.2178327.

Sobieraj, J., & Metelski, D. (2024). Machine Learning Insights: Exploring Key Factors Influencing the Sale-to-List Ratio—Insights from SVM Classification and Recursive Feature Selection in the US Real Estate Market Buildings, 14(5), 1471. https://doi.org/10.3390/buildings14051471.

Sun, N. (2021). Effects of Housing Price on Unemployment Rate and Stock Market International Journal of Trade, Economics and Finance, 12 (5), 131-137. https://doi.org/10.18178/ijtef.2021.12.5.707.

Tchuente, D., & Nyawa, S. (2022). Real estate price estimation in French cities using geocoding and machine learning Annals of Operations Research, 308(1), 571-608. doi:10.1007/s10479-021-03932-5.

Truong, Q., Nguyen, M., Dang, H., & Mei, B. (2020). Housing price prediction via improved machine learning techniques Procedia Computer Science, 174, 433-442. doi:10.1016/j.procs.2020.06.111.

Turkish Statistical Institute. (2025). https://data.tuik.gov.tr/.

Varma, A., Sarma, A., Doshi, S. and Nair, R. (2018). House price prediction using machine learning and neural networks In 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), 1936-1939. https://10.1109/ICICCT.2018.8473231.

Viktorovich, P. A., Aleksandrovich, P. V., Leopoldovich, K. I. and Vasilevna, P. I. (2018). Predicting sales prices of the houses using regression methods of machine learning. 3rd Russian-Pacific Conference on Computer Technology and Applications (RPC), Vladivostok, Russia, 2018, pp. 1-5, https://10.1109/RPC.2018.8482191.

Wheaton, W. C. Lee, N. J. (2009). The Co-Movement of Housing Sales and Housing Prices: Empirics and Theory (March 1, 2009). MIT Department of Economics Working Paper No. 09-05. http://dx.doi.org/10.2139/ssrn.1353590.

World Uncertainty Index (2025). https://worlduncertaintyindex.com/.

Yalçın, M. (2020). Türkiye Konut Satış İstatistiklerinin CBS Tabanlı Analizleri. Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi, 20 (3), 448-455. doi:10.35414/akufemubid.673133.

Yardımcı, M. C. (2021). Türkiye'de Konut Satışları ve Ekonomik Büyüme Arasındaki İlişkinin Analizi. Ekonomi İşletme Siyaset ve Uluslararası İlişkiler Dergisi, 7(2), 380-391.

Yazdani, M. (2021). Machine learning, deep learning, and hedonic methods for real estate price prediction. arXiv preprint arXiv:2110.07151. https://doi.org/10.48550/arXiv.2110.07151.

Ying, X. (2019). An overview of overfitting and its solutions. Journal of Physics: Conference Series, 1168, 022022. https://doi.org/10.1088/1742-6596/1168/2/022022.

Zhang, L., Shi, D., Chang, X., & Wen, H. (2021). Timing decisions of housing sales and development based on the real option theory. International Journal of Strategic Property Management, 25 (2), 90-101. https://doi.org/10.3846/ijspm.2020.13973.