Article

Not peer-reviewed version

From Feature Selection to Forecasting: A Two-Stage Hybrid Framework for Food Price Prediction Using Economic Indicators in Turkey

Uğur Tahsin Şenel * , Nursal Arıcı , Müslüme Narin , Hüseyin Polat

Posted Date: 13 November 2025

doi: 10.20944/preprints202511.0818.v1

Keywords: food price prediction; economic indicators; relationship and causality; machine learning; inflation and sustainability

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 November 2025 doi:10.20944/preprints202511.0818.v1

Article

From Feature Selection to Forecasting: A Two-Stage Hybrid Framework for Food Price Prediction Using Economic Indicators in Turkey

Uğur Tahsin Şenel [1,*], Nursal Arıcı [2], Müslüme Narin [3] and Hüseyin Polat [4]

[1] Gazi University, Graduate School of Informatics, Department of Management Information Systems, Ankara, Turkiye
[2] Gazi University, Faculty of Applied Sciences, Department of Management Information Systems, Ankara, Turkiye
[3] Ankara Haci Bayram Veli University, Faculty of Economics and Administrative Sciences, Department of Economics, Ankara, Turkiye
[4] Gazi University, Faculty of Technology, Department of Computer Engineering, Ankara, Turkiye
* Correspondence: utsenel@gmail.com

Abstract

This study develops a comprehensive two-stage hybrid framework to forecast food prices in Turkey, addressing inflation prediction challenges critical for sustainable food security in emerging economies. In the first stage, systematic relationship and causality analyses—comprising correlation, ARDL, cointegration, and Granger causality tests—identified ten key predictors from the Turkish Statistical Institute and Central Bank datasets. In the second stage, ten predictive models, including ensemble (Gradient Boosting, Random Forest, SVR), traditional (ARIMA, Linear Regression), and deep learning approaches (LSTM, NARX-RNN, ANFIS), were evaluated using rice prices as a pilot case. Ensemble models demonstrated clear superiority, with Gradient Boosting achieving optimal single-split performance ($R^2 = 0.9990$) and high cross-validation consistency (mean $R^2 = 0.9742 \pm 0.03$). Support Vector Regression ($R^2 = 0.9896 \pm 0.02$) and Random Forest ($R^2 = 0.9811 \pm 0.02$) showed statistically equivalent performance, reinforcing ensemble robustness. NARX-RNN analysis revealed a six-month lag in economic shock transmission, providing a practical policy intervention window. SHAP-based interpretability identified insurance, healthcare, transportation, education, and social protection expenditures as major drivers of food price formation, underscoring Turkey's cross-sector inflation mechanisms. These findings integrate econometric rigor with machine learning transparency, offering practical tools for sustainable inflation management and early warning systems in volatile emerging markets.

Keywords: food price prediction; economic indicators; relationship and causality; machine learning; inflation and sustainability

---

1. Introduction

Food price volatility has emerged as a critical challenge for emerging economies, with significant implications for both macroeconomic stability and social welfare. The interconnected nature of modern economic systems implies that food prices are influenced not only by traditional agricultural factors but also by complex interactions with diverse economic sectors including healthcare, communication services, and financial markets. Understanding and predicting these price movements has become essential for effective economic policy formulation, particularly in emerging economies facing high inflation volatility (Karagöl, 2023). This is increasingly recognized as a cornerstone of economic and social sustainability, where food security underpins equitable development and resilience against systemic shocks.

 doi:10.20944/preprints202511.0818.v1

In Turkey, food price dynamics hold particular significance as the Food and Non-Alcoholic Beverages Consumer Price Index (CPI) group represents approximately 25% of the national CPI basket as of 2024. Recent empirical evidence demonstrates that food price fluctuations create cascading effects throughout economic systems, influencing macroeconomic variables such as economic growth, employment levels, and income distribution patterns (Özçelik & Uslu, 2024). The Turkish Statistical Institute's classification of food prices within the CPI framework underscores the systemic importance of developing accurate forecasting models for these price movements (Eştürk & Albayrak, 2018).

Traditional econometric approaches to food price analysis have predominantly focused on supply-side factors and historical price patterns, often failing to capture the complex cross-sector interdependencies that characterize modern economies. While machine learning techniques offer promising alternatives for capturing non-linear relationships and complex interactions among economic variables (Baumeister & Kilian, 2014), limited research has systematically compared traditional econometric methods with modern machine learning approaches within a unified analytical framework or explored optimal prediction windows for timely policy interventions.

This research addresses these critical gaps by developing a comprehensive two-stage methodology that first systematically identifies key economic drivers through multiple econometric techniques, then evaluates ten different approaches for forecasting accuracy. Using monthly data spanning 2017-2024 from the Central Bank of Turkey and the World Food Programme database, this study provides both theoretical insights into inflation transmission mechanisms and practical tools for economic forecasting. The primary contributions include demonstrating how diverse economic indicators collectively influence food price dynamics, establishing the superiority of ensemble methods for economic forecasting, and providing empirical evidence for optimal prediction frameworks with direct policy implications for emerging economy contexts.

2. Literature Review

The literature on food price prediction using economic indicators and machine learning techniques has expanded significantly in recent years. The reviewed studies can be categorized under two headings: relationship and causality and prediction models.

Table 1 summarizes key findings from the literature review across these categories:

Table 1. Key Findings from Literature Review.

Year	Category	Methodological Approach	Content
2024	Relationship and Causality	The ARDL (Autoregressive Distributed Lag) method was employed.	Özçelik and Uslu (2024) investigate the determinants of food inflation within the Turkish economy. Based on ARDL modeling, the study finds that the Consumer Price Index for Food and Non-Alcoholic Beverages is positively influenced by the Domestic Producer Price Index for Agriculture, Forestry, and Fishing (UFET) and the Consumer Price Index for Electricity, Gas, and Other Fuels (TUFEE), while the Real Effective Exchange Rate based on CPI (REDK) exerts a negative impact. Furthermore, results from the ARDL Error Correction Model, which examines short-term dynamics, indicate that short-term imbalances are corrected in the long run.
2023	Relationship and Causality	Panel Structural Vector Autoregression (PSVAR) technique to assess inflation effects	Fan et al. (2023) explore the relationship between information asset investments and inflation. Utilizing the PSVAR method, they analyze both short- and long-term dynamics. Their findings suggest that low to moderate inflation levels are positively correlated with the market value of R&D firms, whereas high inflation has a negative effect.
2023	Relationship and Causality	Granger causality test to examine the PPI-CPI relationship	Cerveny (2023) investigates the link between Producer Price Index (PPI) and Consumer Price Index (CPI) in the Czech Republic and the Eurozone. Applying the Granger causality test, the study reveals that

doi:10.20944/preprints202511.0818.v1

			PPI influences CPI in the Czech Republic, whereas no such causal relationship is observed in the Eurozone.
2020	Relationship and Causality	Panel cointegration and panel causality tests	Ozpolat (2020) analyzes the causal relationship between CPI and PPI in Central and Eastern European Countries (CEECs), using panel cointegration and panel causality tests. The results indicate a long-term, bidirectional causality between CPI and PPI in these countries.
2018	Relationship and Causality	Econometric methods: DF-GLS unit root test, Johansen and Engle-Granger cointegration approaches, VAR model	Oyeleke and Ojediran (2018) examine the relationship between PPI and CPI in Nigeria using various econometric techniques. The DF-GLS unit root test is applied to assess stationarity, Johansen and Engle-Granger methods are used for long-run cointegration, and a VAR model is employed to analyze interactions. The study concludes that the PPI-CPI relationship in Nigeria does not follow a simple cause-effect pattern and lacks a long-term equilibrium relationship.
2016	Relationship and Causality	Non-parametric regression using the LOESS technique	Akmercan (2016) investigates the relationships among household expenditures, income, and OECD household size data using the LOESS (Locally Estimated Scatterplot Smoothing) non-parametric regression method. Essential consumption items are aggregated into a single expenditure category for analysis.
2016	Relationship and Causality	Comparison of ordered and unordered discrete choice models (LOGIT and PROBIT))	Çelik (2016) analyzes factors influencing household fuel choices for heating in Turkey using TÜİK data. The study compares ordered and unordered discrete choice models, particularly LOGIT and PROBIT variants. Model performance is assessed using OLOGIT, GOLOGIT, PPO, HOLOGIT, AIC, BIC, and MNL statistics to determine the most suitable approach.
2016	Relationship and Causality	Panel data analysis and Dumitrescu-Hurlin panel causality test	Chih-Ping Yu (2016) first applies panel data analysis to explore the general dynamics between CPI and PPI, then uses the Dumitrescu-Hurlin panel causality test for a deeper investigation into the causal nature of this relationship. This dual approach allows for a more nuanced understanding of inconsistencies in CPI-PPI transmission across countries.
2014	Relationship and Causality	Correlation, regression, ANOVA, and coefficient of determination (R^2)	Galodikwe (2014) investigates the PPI-CPI relationship using correlation analysis, regression models, ANOVA, and the coefficient of determination. The findings confirm that PPI indices significantly influence CPI indices.
2001	Relationship and Causality	Limitations of OLS and use of Tobit models	Emeç (2001) examines household consumption expenditures, highlighting the limitations of the Ordinary Least Squares (OLS) method when applied to continuous or ordinal dependent variables across regions. As a solution, Tobit models are suggested, where zero expenditures are bounded at zero, and certain continuous variables are categorized to fit ordered logit models. Results are interpreted in the context of Engel curves.
2022	Relationship and Causality	Combined econometric (ARDL) and machine learning (Support Vector Machine) approach; hybrid model proposed. VIF test used to avoid multicollinearity. Evaluation Metrics: RMSE, MAE, R^2	Ozden (2022) investigates macroeconomic and financial determinants of Turkey's export-import ratio using both econometric and machine learning methods. The ARDL model is applied to monthly data (2010–2021) on normalized GDP, exchange rate, CPI, PPI, crude oil prices, and trade ratio. Trends of each variable are presented. A VIF test confirms no multicollinearity issues. Subsequently, Support Vector Machine (SVM) is used to capture complex patterns. Results from ARDL, SVM, and a hybrid ARDL-SVM model are compared using RMSE, MAE, and R^2. The hybrid model, supported by machine learning, demonstrates superior performance in capturing variable interactions.
2016	Prediction Model	Poisson Quasi Maximum Likelihood estimation; Bootstrap validation test	Balyaner (2016) estimates the number of information technology devices owned by households using the Poisson Quasi Maximum Likelihood (PQML) estimation method. The validity of the model is assessed through bootstrap resampling techniques.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 November 2025 doi:10.20944/preprints202511.0818.v1

2023	Prediction Model	Comparison of Random Forest, Gradient Boosting, SVM, Neural Networks, and AdaBoost Evaluation Metrics: MSE, RMSE, MAE, R^2	Atalan (2023) evaluates economic, social, and environmental factors affecting unit prices of milk in Turkey. Five machine learning algorithms—Random Forest, Gradient Boosting, Support Vector Machine (SVM), Artificial Neural Network, and AdaBoost—are used for price prediction. Performance is assessed using MSE, RMSE, MAE, and R^2. Random Forest yields the best results. Additionally, Random Forest performance is reported across tree counts ranging from 10 to 2000.
2022	Prediction Model	System dynamics model for energy efficiency and resource optimization in the food and beverage industry	Katsumbe (2022) proposes a system dynamics model to optimize energy efficiency and resource use in the food and beverage sector. Separate sub-models are developed for water, electricity, and production lines, with input variables defined for each. Total consumption is formulated and compared against a baseline. The model is used to simulate one-year forecasts.
2022	Prediction Model	SARIMA model for forecasting food and beverage prices in Kenya, accounting for seasonality. Evaluation Metrics: MSE, MAE, MAPE, Theil's U statistic	Wanjuki et al. (2022) propose a model for forecasting food and beverage prices in Kenya. Given seasonal fluctuations, the Seasonal Autoregressive Integrated Moving Average (SARIMA) model is employed. Model accuracy is evaluated using MSE, MAE, MAPE, and Theil's U statistic. High predictive accuracy is achieved, and the model is recommended for short-term price forecasting in the food and beverage sector.
2022	Prediction Model	Multiple regression model implemented in Minitab	Warren et al. (2022) develop a multiple regression model to forecast agave (a key input in tequila production) prices. Variables include rainfall, harvest volume, tequila production, costs, exchange rates, and export volumes. The model, run in Minitab, shows strong predictive performance ($R = 0.86$).
2022	Prediction Model	Comparison of deep learning models: DA-RNN, NARX-RNN, MV-LSTM Evaluation Metrics: RMSE, MAE, MAPE	Ji et al. (2022) investigate deep learning approaches for forecasting wholesale agricultural prices in China. The Dual-Stage Attention-Based Recurrent Neural Network (DA-RNN) outperforms NARX-RNN and MV-LSTM models. Performance is evaluated using RMSE, MAE, and MAPE.
2022	Prediction Model	ARCH and GARCH models for forecasting prices of food items (tomato, garlic, okra, pepper)	Venkateswara et al. (2022) present a regression-based multivariate approach to forecast prices of key food commodities. Emphasizing the importance of price volatility for governments, producers, and consumers, they apply ARCH (Autoregressive Conditional Heteroskedasticity) and GARCH (Generalized ARCH) models. While ARCH generally yields more consistent results, GARCH performs better for certain items.
2022	Prediction Model	Random Forest with three cross-validation techniques: temporal, spatial, spatiotemporal	Kresove and Hess (2022) analyze factors influencing raw milk prices in Russia using 17 variables. Feature selection is performed using Boruta analysis, confirming all variables as relevant. The Random Forest model is tested with three cross-validation strategies: temporal (for time-series), spatial (for geographical), and spatiotemporal (combined). The spatiotemporal approach is found to be the most effective.
2021	Prediction Model	Artificial Neural Networks (ANN) and Multiple Linear Regression for CPI forecasting Software: WEKA	Özcan (2021) examines the influence of macroeconomic variables—External Debt, PPI, USD exchange rate, Exports, Imports, and M2 money supply—on CPI in Turkey using data from 2008 to 2020 (TCMB). Both ANN and Multiple Linear Regression models are implemented in WEKA. The ANN model demonstrates superior predictive accuracy compared to the linear regression model.
2021	Prediction Model	NARXNN model for forecasting food demand	Lutoslawski et al. (2021) employ the Nonlinear Autoregressive Exogenous Neural Network (NARXNN) model to forecast food demand. The study highlights that NARXNN, commonly used in time

			series forecasting, provides more accurate predictions than traditional regression models.
2021	Prediction Model	Backpropagation-trained ANN model for CPI forecasting Software: Zaitun Evaluation: MAPE	Sarangi et al. (2021) aim to forecast the Consumer Food Price Index (CFPI) in India using a machine learning approach. A backpropagation-trained Artificial Neural Network (ANN) is implemented using the Zaitun statistical software. MAPE values are used to validate model accuracy, which is reported to be very high, indicating strong predictive performance.
2020	Prediction Model	ANN, Random Forest, and XGBoost models Evaluation Metrics: R^2, MAE, RMSE	Tosun (2020) forecasts fresh fruit and vegetable imports for OECD countries using data mining and machine learning techniques. ANN, Random Forest, and XGBoost models are applied and compared using R^2, RMSE, and MAE. XGBoost demonstrates the best overall performance.
2020	Prediction Model	Applicability of ANN, SVM, genetic algorithms, and hybrid techniques in stock price forecasting	Strader et al. (2020) conduct a study on stock price forecasting. Their findings suggest that: Artificial Neural Networks (ANN) are best suited for predicting numerical stock index values; Support Vector Machines (SVM) perform well in classification tasks, such as predicting market direction; Hybrid machine learning techniques may overcome limitations of single-method approaches.
2019	Prediction Model	Superiority of ANN over logarithmic regression	Selim and Demirkıran (2019) analyze household budget survey data from TÜİK to identify factors affecting food expenditures and track temporal changes. They develop predictive models using logarithmic regression and Artificial Neural Networks (ANN). Results show that the ANN model outperforms the semi-logarithmic regression model in forecasting accuracy.
2019	Prediction Model	Comparison of ANN, SVM, and ARIMA models	Abidoye et al. (2019) collect data on factors influencing real estate prices in Hong Kong and apply ARIMA, ANN, and SVM models. The models are used for out-of-sample forecasting. The ANN model outperforms both SVM and ARIMA in predictive accuracy.
2018	Prediction Model	ANFIS (Adaptive Neuro-Fuzzy Inference System) combining fuzzy logic and neural networks	Soltani and Pooya (2018) design an AI system to predict the success of new food products. The ANFIS algorithm integrates fuzzy logic and neural networks, processing data from diverse sources such as market research and social media to forecast product performance.
2018	Prediction Model	Evaluation of machine learning as an alternative to statistical methods in time series forecasting	Makridakis et al. (2018) assess machine learning methods as alternatives to traditional statistical approaches in time series forecasting. Eight classical statistical methods and ten machine learning techniques are compared using sMAPE. The results show that statistical methods generally outperform machine learning models. However, the authors note that recent advancements may soon close this gap.

These studies provide valuable insights into the application of various statistical methods and machine learning techniques for analyzing economic data and predicting food prices. They highlight different approaches to feature selection, model development, and evaluation metrics that have guided the current research.

3. Methodology

3.1. Research Framework

This study employs a two-stage analytical framework: Stage 1 systematically identifies economic drivers affecting overall food prices through CPI analysis; Stage 2 validates these drivers through pilot application to rice price prediction, providing a scalable methodology for broader food price forecasting applications.

3.2. Data Collection and Preprocessing

3.2.1. Dataset Creation

The dataset was created by including Dollar Sale, Euro Sale, Gold Bullion Sale prices, and the entire CPI group from the Electronic Data Delivery System (EVDS) of the Central Bank of the Republic of Turkey. The data spans from 2017 to 2024, providing a comprehensive time period for analysis. (EVDS, 2024)

Rice was selected as the pilot case due to its status as a staple food for over half of the world's population (Muthayya et al., 2014) and extensive documentation of its price volatility relationships with macroeconomic indicators (Valera, 2022). Rice prices are obtained from World Food Programme database. The city is selected as Ankara, Turkey to assure data consistency. (World Food Programme Price Database, 2024)

3.2.2. Data Preprocessing

For effective use in analytical applications, Min-Max normalization was applied to the data:

$$X\_normalized = (X - X\_min) / (X\_max - X\_min)$$

This normalization technique is widely adopted in machine learning applications to ensure equal scaling of variables with different ranges (Hastie et al., 2009). Missing value analysis revealed only the "103.POST-SECONDARY AND PRE-UNIVERSITY EDUCATION" CPI group with 106 missing values, which was removed due to minimal expected impact on food products.

3.3. Relationship and Causality Analysis

The relationship and causality analysis phase aims to systematically identify which economic indicators have significant relationships with food prices, following the methodological approaches established in recent econometric literature.

3.3.1. Time Series Analysis and Stationarity Testing

Partial Autocorrelation Function (PACF) analysis was conducted to measure direct effects of the relationship between an observation and its lagged values in the time series (Box et al., 2016). This technique helps determine the appropriate lag structure for autoregressive models by identifying the direct correlation between current and past values:

$$PACF(k) = corr(Xt, Xt\text{-}k \mid Xt\text{-}1, Xt\text{-}2, ..., Xt\text{-}k\text{+}1)$$

Augmented Dickey-Fuller (ADF) Test was used to check for stationarity in the data series, as required for causality testing (Dickey & Fuller, 1979). The test examines the presence of unit roots in time series:

$$\Delta yt = \alpha + \beta t + \gamma yt\text{-}1 + \delta 1\Delta yt\text{-}1 + ... + \delta p\text{-}1\Delta yt\text{-}p\text{+}1 + \varepsilon t$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 November 2025 doi:10.20944/preprints202511.0818.v1

Where the null hypothesis H0: $\gamma = 0$ (unit root exists) versus H1: $\gamma < 0$ (stationary). This test is crucial for ensuring the validity of subsequent causality analyses, as non-stationary series can lead to spurious regression results.

3.3.2. Correlation Analysis

Three different correlation methods were employed to identify variables influencing the food CPI, providing complementary perspectives on variable relationships (Cohen et al., 2003):

Pearson Correlation measures linear relationships between variables:

$$r = \Sigma[(xi - \bar{x})(yi - \bar{y})] / \sqrt{[\Sigma(xi - \bar{x})^2\Sigma(yi - \bar{y})^2]}$$

This parametric measure is most appropriate when variables follow normal distributions and relationships are linear.

Spearman Rank Correlation captures monotonic relationships regardless of linearity:

$$\rho = 1 - (6\Sigma di^2) / [n(n^2 - 1)]$$

This non-parametric approach is robust to outliers and suitable for variables that may not meet normality assumptions.

Kendall's Tau provides robust association measure based on concordant and discordant pairs:

$$\tau = (C - D) / [n(n-1)/2]$$

Where C and D represent concordant and discordant pairs, respectively. This measure is particularly useful for small sample sizes and provides more reliable inference.

3.3.3. Lag Features Analysis

Lag correlation analysis was performed for previous months, creating lagged features for all columns and calculating correlations with the target variable using all three correlation methods. This approach helps identify optimal temporal relationships and leads/lags in the data, which is crucial for understanding how economic indicators influence food prices over time (Baumeister & Kilian, 2014).

3.3.4. Granger Causality Test

After stabilizing the dataset, Granger causality tests were applied to identify causal relationships between variables (Granger, 1969). This approach has been successfully used in food price studies to establish predictive relationships (Cerveny, 2023):

$$yt = \alpha 0 + \Sigma(i=1 \text{ to } p)\ \alpha iyt\text{-}i + \Sigma(i=1 \text{ to } p)\ \beta ixt\text{-}i + \varepsilon t$$

Where rejection of H0: $\beta 1 = \beta 2 = \ldots = \beta p = 0$ indicates x Granger-causes y. This test determines whether past values of economic indicators contain information useful for predicting food prices beyond what is contained in past values of food prices alone.

3.3.5. Autoregressive Distributed Lag (ARDL) Model

The ARDL analysis examined relationships from both short-term and long-term perspectives, following the methodology established by Pesaran et al. (2001). This approach has been particularly effective in food price analysis, as demonstrated by Özçelik and Uslu (2024):

$$yt = \alpha + \Sigma(i=1 \text{ to } p)\ \beta iyt\text{-}i + \Sigma(i=0 \text{ to } q)\ \gamma ixt\text{-}i + \varepsilon t$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 November 2025 doi:10.20944/preprints202511.0818.v1

The model considered 12-month lags for the food CPI and 1-month lags for other items. ARDL models are advantageous because they can handle variables with different orders of integration and provide both short-run dynamics and long-run equilibrium relationships.

3.3.6. Cointegration Test (Engle-Granger)

The cointegration analysis tested for long-term relationships between non-stationary series using the two-step Engle-Granger method (Engle & Granger, 1987). This technique is essential for identifying stable long-run relationships between economic variables, as applied in similar studies examining price relationships (Ozpolat, 2020):

Step 1: $yt = \alpha + \beta xt + ut$ (estimate long-run relationship)

Step 2: Test residuals for stationarity to confirm cointegration

Cointegration testing helps distinguish between spurious and genuine long-term relationships among economic variables.

3.3.7. Random Forest Feature Importance

Random Forest feature importance scores were used to identify variables with the strongest relationship to the target variable, employing both random 80%/20% split and chronological split approaches (Breiman, 2001). This machine learning-based approach complements traditional econometric methods and has shown effectiveness in feature selection for economic forecasting applications.

3.3.8. Attribute Shortlist

According to previous study, we aim to combine all studies best performed items with selecting top two items from each relationship and causality study. Since some top performers might be same so that we have a workable shortlist.

3.4. Predictive Model Development

Based on the relationship analysis results, ten key predictor variables were identified for food price prediction. Multiple prediction models were then developed and compared for their forecasting performance, following the comparative modeling approach established in recent food price prediction literature (Atalan, 2023).

3.4.1. Linear Regression

Linear regression serves as the baseline model for comparison, testing the fundamental assumption of linear relationships between economic indicators and food prices (Hastie et al., 2009):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \varepsilon$$

$$\hat{\beta} = (X'X)^{-1}X'Y$$

Despite its simplicity, linear regression often provides competitive performance in economic forecasting and offers excellent interpretability for policy makers.

3.4.2. Random Forest

Random Forest combines multiple decision trees using bootstrap aggregating to capture non-linear relationships and interactions between variables (Breiman, 2001). This ensemble method has

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 November 2025 doi:10.20944/preprints202511.0818.v1

demonstrated superior performance in agricultural price prediction, as shown by Atalan (2023) for Turkish milk price forecasting:

$$\hat{h}_\beta(x) = \Sigma_{i=1}^{n} w_i(x)y_i$$

$$\hat{h}\_RF(x) = (1/B)\ \Sigma_{\beta=1}^{B}\ \hat{h}_\beta(x)$$

$$VI(X_j) = (1/B)\ \Sigma_{\beta=1}^{B}\ \Sigma_t\ p(t)[\Delta I(t)]I(v(t) = j)$$

Where B is the total number of trees, and VI represents variable importance. The model's ability to provide feature importance rankings makes it valuable for understanding which economic indicators most strongly influence food prices.

3.4.3. Gradient Boosting

Gradient Boosting sequentially combines weak learners to minimize residual errors, building upon the boosting framework developed by Friedman (2001). This method excels at capturing complex patterns in economic data:

$$F(x) = \Sigma_{m=1}^{M}\ \gamma_m h_m(x)$$

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x)$$

$$r_{im} = -[\partial L(y_i, F(x_i))/\partial F(x_i)]\_\{F=F_{m-1}\}$$

Where γm is the step size and rim are pseudo-residuals. The sequential error correction mechanism makes gradient boosting particularly effective for economic time series with complex underlying patterns.

3.4.4. XGBoost

XGBoost implements an optimized gradient boosting framework with regularization techniques to prevent overfitting (Chen & Guestrin, 2016). This algorithm has gained widespread adoption in economic forecasting due to its robust performance:

$$L(\theta) = \Sigma_i\ l(y_i, \hat{y}_i) + \Sigma_k\ \Omega(f_k)$$

$$\Omega(f) = \gamma T + (1/2)\lambda||w||^2$$

$$L^{(t)} = \Sigma_{i=1}^{n}\ [g_i f_t(x_i) + (1/2)h_i f_t^2(x_i)] + \Omega(f_t)$$

Where Ω(f) represents regularization terms. The built-in regularization and efficient implementation make XGBoost suitable for economic datasets with high dimensionality.

3.4.5. Support Vector Regression (SVR)

SVR uses ε-insensitive loss function for robust predictions against outliers, which is particularly valuable in economic data that often contains extreme values (Vapnik, 1995):

$$f(x) = \Sigma_{i=1}^{n}\ (\alpha_i - \alpha_i^*)K(x_i, x) + b$$

$$K(x_i, x_j) = \exp(-\gamma||x_i - x_j||^2)$$

$$\min\ (1/2)||w||^2 + C\ \Sigma_{i=1}^{n}\ (\xi_i + \xi_i^*)$$

The kernel approach allows SVR to capture non-linear relationships while maintaining robustness to outliers, making it suitable for volatile economic indicators.

3.4.6. Long Short-Term Memory (LSTM)

LSTM addresses vanishing gradient problems in RNNs and is designed to capture long-term dependencies in time series data (Hochreiter & Schmidhuber, 1997). Deep learning approaches like LSTM have shown promise in agricultural price prediction (Ji et al., 2022):

$$f_t = \sigma(Wf \cdot [h_{t-1}, x_t] + bf)$$

$$i_t = \sigma(Wi \cdot [h_{t-1}, x_t] + bi)$$

$$\tilde{C}_t = \tanh(WC \cdot [h_{t-1}, x_t] + bC)$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 November 2025 doi:10.20944/preprints202511.0818.v1

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$
$$o_t = \sigma(Wo \cdot [h_{t-1}, x_t] + bo)$$
$$h_t = o_t * \tanh(C_t)$$

The gate mechanisms allow LSTM to selectively remember and forget information, potentially capturing complex temporal patterns in economic relationships.

3.4.7. Artificial Neural Networks (ANN)

ANNs use backpropagation for learning complex non-linear patterns between economic indicators and food prices (Rumelhart et al., 1986):

$$y_j = f(\Sigma_i\, w_{ij}x_i + b_j)$$
$$w_{ij}(t+1) = w_{ij}(t) - \eta\, \partial E/\partial w_{ij}$$

Multi-layer perceptrons can approximate complex non-linear functions, making them suitable for modeling intricate economic relationships.

3.4.8. ARIMA

ARIMA models time series using autoregressive and moving average components, representing the classical approach to time series forecasting (Box et al., 2016). Similar seasonal models have shown success in short-term food price predictions (Wanjuki et al., 2022):

$$\varphi(B)(1-B)^d X_t = \theta(B)\varepsilon_t$$
$$(1-\varphi_1 B-\ldots-\varphi_p B^p)(1-B)^d X_t = (1+\theta_1 B+\ldots+\theta_\varphi B_\varphi)\varepsilon_t$$

While ARIMA focuses solely on temporal patterns, it provides a benchmark for comparing the added value of incorporating economic indicators.

3.4.9. NARX-RNN

NARX networks model nonlinear dynamic systems with exogenous inputs, combining temporal dependencies with external economic factors (Lin et al., 1996):

$$y(t) = f(y(t-1), y(t-2), \ldots, y(t-n_\gamma), u(t-1), u(t-2), \ldots, u(t-n_u))$$
$$h(t) = \tanh(W_{\gamma\gamma}h(t-1) + W_{u\gamma}u(t) + W_{\gamma\gamma}y(t-1) + b)$$
$$y(t) = W_{\gamma\gamma}h(t) + b_\gamma$$

The model was tested with different lag periods (3, 6, 9, and 12 months) to find optimal prediction windows. This approach bridges traditional econometric modeling with neural network capabilities.

3.4.10. ANFIS

ANFIS combines fuzzy logic with neural networks, providing interpretable non-linear modeling (Jang, 1993):

$$x = A,\ Y= B \text{ ise } z = px + qy + r\ ;$$
$$\mu_{ai}(x) = \exp(-(x-c_i)^2/2\sigma_i^2)$$
$$w_i = \mu_{ai}(x) \times \mu_{\beta i}(y)$$
$$\bar{w_i} = w_i/\Sigma_i w_i$$
$$\bar{w_i}f_i = \bar{w_i}(p_i x + q_i y + r_i)$$
$$\text{Output} = \Sigma_i \bar{w_i} f_i$$

Due to architectural constraints, ANFIS was limited to the two highest-ranked features from the feature selection process. The fuzzy rule-based approach offers transparency in decision-making processes.

 doi:10.20944/preprints202511.0818.v1

3.4.11. SHAP-Based Ensemble Interpretability

To enhance model transparency and uncover consensus-driven economic drivers, SHAP (SHapley Additive exPlanations) analysis was applied to the three top-performing models: Gradient Boosting, Random Forest, and Support Vector Regression (SVR). SHAP values provide a game-theoretic approach to explain the output of machine learning models by quantifying each feature's marginal contribution to individual predictions (Lundberg et al., 2020; Lundberg & Lee, 2017). Given the small dataset size ($n = 94$), KernelSHAP was employed for SVR (which lacks native tree-based SHAP support), while TreeSHAP was used for Gradient Boosting and Random Forest for computational efficiency and exact Shapley value estimation. The analysis was conducted on the full test set ($n = 19$) using a representative background sample to ensure stable explanations. Feature importance was derived as the mean absolute SHAP value across all test instances, enabling cross-model comparability and consensus ranking.

3.5. Model Evaluation

3.5.1. Evaluation Metrics

Models were evaluated using an 80-20 train-test split with three performance metrics that provide complementary perspectives on prediction accuracy:

- **Mean Absolute Error (MAE)**:

$$MAE = (1/n)\, \Sigma | y_i - \hat{y}_i |$$

- **Root Mean Squared Error (RMSE)**:

$$RMSE = \sqrt{[(1/n)\, \Sigma(y_i - \hat{y}_i)^2]}$$

- **Coefficient of Determination (R^2)**:

$$R^2 = 1 - (SS_{res} / SS_{tot})$$

These metrics enable comprehensive assessment of model performance, with MAE providing interpretable average error, RMSE penalizing larger deviations, and R^2 indicating explained variance.

3.5.2. Cross-Validation Assessment

To ensure robust model evaluation and detect potential overfitting, K-fold cross-validation is implemented for the top-performing model. K-fold cross-validation partitions the dataset D into k equal folds, where each model f_i is trained on k-1 folds and tested on the remaining fold D_i. The cross-validation score is computed as:

$$CV\_Score = (1/k)\, \Sigma_{i=1}^{k} P_i$$

Where P_i represents the performance score on fold i. This methodology provides robust performance estimates and enables detection of overfitting by assessing model stability across different data partitions (Hastie et al., 2009).

3.5.3. Statistical Significance Test

To assess whether performance differences between top-performing models were statistically meaningful, paired t-tests were conducted using cross-validation scores. For models i and j with CV scores x_i and x_j, the paired t-test evaluates:

$$H_0: \mu_i - \mu_j = 0 \text{ (no difference)}$$

$$H_1: \mu_i - \mu_j \neq 0 \text{ (significant difference)}$$

doi:10.20944/preprints202511.0818.v1

$$t = (\bar{x}_i - \bar{x}_j) / (s^d/\sqrt{n})$$

Where s^d is the standard deviation of paired differences and n is the number of CV folds. This approach provides robust statistical inference for model selection decisions (Hastie et al., 2009).

4. Relationship and Causality Analyses & Results

4.1. Time Series Analysis

Time series analysis was conducted to observe the changes in the 'TP FG J011' (CPI food index) value over the last 10 years. The analysis revealed a significant increase, especially during the pandemic period. However, this increase follows a trend, with a notable difference between the trend until the end of 2021 and the trend afterward. This provides a suitable basis for examining the precursor causes.

4.2. Partial Autocorrelation Function (PACF)

PACF analysis was used to measure the direct effects of the relationship between an observation and its lagged values in the time series. When performing this test for the correlation of the "TP FG J011" column with previous months, it was observed that it only has a direct relationship with the previous month.

4.3. Correlation Analysis

To identify other items that influence the "TP FG J011" dependent variable, three different correlation methods were used: Pearson, Spearman, and Kendall methods. Tables 2, 3, and 4 show the top three items with the highest correlation values for each method.

Table 2. Top 3 Items with Highest Pearson Correlation to Food CPI.

Item Code	Description	Rank	Correlation Value
TP FG J053	053.HOUSEHOLD APPLIANCES	1	0.99891
TP FG J051	051.FURNITURE, FURNISHINGS, CARPETS AND OTHER FLOOR COVERINGS	2	0.998844
TP FG J056	056.GOODS AND SERVICES FOR HOUSEHOLD MAINTENANCE	3	0.997916

As shown in Table 2, household-related CPI components demonstrate the strongest Pearson correlation with food prices, suggesting shared underlying economic factors.

Table 3. Top 3 Items with Highest Spearman Correlation to Food CPI.

Item Code	Description	Rank	Correlation Value
TP FG J127	127.OTHER SERVICES N.E.C.	1	0.99776
TP FG J124	124.SOCIAL PROTECTION	2	0.997718
TP FG J062	062.OUTPATIENT SERVICES	3	0.997687

Table 3 reveals that service-oriented CPI components show strong rank-based correlations with food prices, indicating similar monotonic relationships with underlying economic conditions.

Table 4. Top 3 Items with Highest Kendall Tau Correlation to Food CPI.

Item Code	Description	Rank	Correlation Value
TP FG J127	127.OTHER SERVICES N.E.C.	1	0.970884
TP FG J124	124.SOCIAL PROTECTION	2	0.970381
TP FG J062	062.OUTPATIENT SERVICES	3	0.97034

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 November 2025 doi:10.20944/preprints202511.0818.v1

Kendall's tau correlation results in Table 4 confirm the findings from the Spearman correlation analysis, with service-related CPI items showing the strongest associations with food prices.

The combined correlation results in Table 5 provide a more comprehensive view, highlighting both household maintenance services and non-alcoholic beverages as having the strongest overall correlation with food prices.

Table 5. Combined Correlation Results (Average of Three Methods).

Item Code	Description	Rank	Average Correlation
TP FG J056	056.GOODS AND SERVICES FOR HOUSEHOLD MAINTENANCE	1	0.988253667
TP FG J012	012.NON-ALCOHOLIC BEVERAGES	2	0.988152
TP FG J062	062.OUTPATIENT SERVICES	3	0.988128

4.4. Lag Features Analysis

A lag correlation analysis was performed for previous months. The analysis created lagged features for all columns and calculated correlations with the target variable using all three correlation methods.

As shown in Table 6, a one-month lag of household appliances prices demonstrates the strongest correlation with current food prices, suggesting a potential leading indicator relationship.

Table 6. Items with Highest Lagged Pearson Correlation to Food CPI.

Item Code	Description	Lag Period	Correlation Value
TP FG J053	053.HOUSEHOLD APPLIANCES	1	0.999222306
TP FG J051	051.FURNITURE, FURNISHINGS, CARPETS AND OTHER FLOOR COVERINGS	1	0.998482096
TP FG J012	012.NON-ALCOHOLIC BEVERAGES	1	0.997243669

4.5. Stationarity Test

The Augmented Dickey-Fuller (ADF) test was used to check for stationarity in the data series, which is required for some causality relationship tests. The test showed that all data series were non-stationary, necessitating a data transformation process to achieve stationarity.

4.6. Granger Causality Test

After stabilizing the dataset, Granger causality tests were applied to identify causal relationships between variables. Items with p-values less than 0.05 indicated significant causal relationships with food prices at specific lag months.

4.7. Autoregressive Distributed Lag (ARDL) Model

The ARDL analysis examined the relationship between the dependent variable and multiple independent variables from both short-term and long-term perspectives. The model considered 12-month lags for 'TP FG J011' and 1-month lags for other items.

The ARDL results in Table 7 highlight that healthcare-related CPI components and telecommunication services have significant statistical relationships with food prices, controlling for other factors.

Table 7. Significant Items in ARDL Analysis ($p < 0.05$).

Item Code	Description	Coefficient	Std Error	p-value
TP FG J062.L0	062.OUTPATIENT SERVICES	0.0092	0.004	0.034

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 November 2025 doi:10.20944/preprints202511.0818.v1

TP FG J061.L0	061.MEDICAL PRODUCTS, APPLIANCES AND EQUIPMENT	0.0081	0.003	0.028
TP FG J083.L1	083.TELEPHONE AND TELEFAX SERVICES	0.0079	0.003	0.015

4.8. Cointegration Test (Engle-Granger)

The cointegration analysis tested for long-term relationships between non-stationary series.

The cointegration test results in Table 8 indicate strong long-term relationships between food prices and education, social protection, and transportation costs.

Table 8. Significant Items in Cointegration Test ($p < 0.05$).

Item Code	Description	Cointegration Statistic	p-value
TP FG J105	105.EDUCATION PROGRAMMES OF UNSPECIFIED LEVEL	-5.59518832	0.0000115
TP FG J124	124.SOCIAL PROTECTION	-5.239523044	0.0000583
TP FG J072	072.OPERATION OF PERSONAL TRANSPORT EQUIPMENT	-4.565730017	0.000953

4.9. Random Forest Feature Importance Analysis

Random Forest feature importance scores were used to identify variables with the strongest relationship to the target variable. Two approaches were used: random 80%/20% split and chronological split.

The Random Forest analysis in Table 9 highlights insurance, medical products, and transport services as having the most predictive importance for food prices.

Table 9. Top 3 Features by Importance (Chronological Split).

Item Code	Description	Rank	Importance Score
TP FG J125	125.INSURANCE	1	0.068466
TP FG J061	061.MEDICAL PRODUCTS, APPLIANCES AND EQUIPMENT	2	0.039167
TP FG J073	073.TRANSPORT SERVICES	3	0.036515

4.10. Analysis Results & Attribute Shortlist

The analyses identified ten key factors taking two best performing items from each Relationship and Causality analysis that have the most significant influence on the "TP FG J011 FOOD" CPI item, along with their most influential month lag:

As summarized in Table 10, healthcare-related items, communication services, house items, and education costs emerged as the most influential predictors for food prices in Turkey.

Table 10. Key Predictor Variables for Food Price Prediction.

Item	Description	Method	Delay (Months)
TP FG J053	053. HOUSEHOLD APPLIANCES	Pearson	1
TP FG J051	051. FURNITURE, FIXTURES, CARPETS AND OTHER FLOOR COVERINGS	Pearson	0
TP FG J073	073. TRANSPORTATION SERVICES	Spearman	6
TP FG J127	127. OTHER UNCLASSIFIED SERVICES	Spearman and Kendall Tau	0
TP FG J124	124. SOCIAL PROTECTION	Spearman, Kendall Tau, Cointegration	0
TP FG J062	062. OUTPATIENT SERVICES	Kendall Tau, ARDL	0
TP FG J105	105. EDUCATIONAL PROGRAMS NOT DETERMINED BY LEVEL	Cointegration Test	0
TP FG J061	061. MEDICAL PRODUCTS, INSTRUMENTS AND EQUIPMENT	ARDL, Random Forest	0
TP FG J125	125. INSURANCE	Random Forest	0

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 November 2025 doi:10.20944/preprints202511.0818.v1

TP FG J011	011. FOOD	PACF	1

Based on the optimal lag periods identified in the relationship and causality analysis (Table 10), temporal features were constructed for each predictor. Specifically:

- TP FG J053 (Household Appliances): 1-month lag
- TP FG J073 (Transportation Services): 6-month lag
- TP FG J011 (Food CPI): 1-month lag (based on PACF analysis)
- All other predictors: Current period (0-month lag)

These lagged features were consistently applied across all predictive models except NARX-RNN, which internally optimizes temporal dependencies. For ARIMA, which does not accommodate exogenous variables with mixed lag structures, only the autoregressive component of TP FG J011 was utilized.

5. Prediction Model Development & Results

5.1. Dataset Preparation

Using the identified features from previous analyses, a combined dataset was constructed for the price prediction model. Rice prices from the World Food Programme database were selected as a pilot case due to rice being a staple food item with price volatility that can be analyzed in relation to economic indicators. The dataset included both Turkish Lira (TL) and US Dollar (USD) prices, though only TL-based prices were considered for modeling purposes.

The target variable (Rice TL prices) and features were separated, and Min-Max normalization was applied to ensure that all variables were on the comparable scale. This preprocessing step was particularly important when using economic indicators with different value ranges to prevent any single variable from dominating the model based solely on its scale.

5.2. Model Selection and Implementation

Based on the literature review, several models were selected and implemented to predict rice prices. Each model was evaluated using a train-test split of 80%-20% respectively. The performance was assessed using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and coefficient of determination (R^2). The following models were implemented: Linear Regression, Random Forest, XGBoost, Support Vector Regressor (SVR), Long Short-Term Memory (LSTM), Artificial Neural Network (ANN), Gradient Boosting Regressor, ARIMA, NARX-RNN (with varying lag periods), ANFIS (Adaptive Neuro-Fuzzy Inference System).

5.3. Model Performance Results

The performance metrics for each model are summarized in Table 11.

Table 11. Performance Comparison of Prediction Models.

Rank	Model	MAE	RMSE	R^2	Performance
1	Gradient Boosting	0.2838	0.4229	0.9990	Excellent
2	SVR	0.3743	0.5743	0.9982	Excellent
3	NARX-RNN (6 months)	0.4845	0.6363	0.9988	Excellent
4	Random Forest	0.3690	0.6612	0.9976	Excellent
5	XGBoost	0.3995	0.6659	0.9975	Excellent
6	Linear Regression	0.8147	1.2329	0.9915	Good
7	ANN	1.1036	1.3161	0.9903	Good
8	ANFIS	7.2268	10.7862	0.3497	Poor/Failed
9	ARIMA	9.3682	11.2475	-0.7834	Poor/Failed
10	LSTM	19.8388	21.4423	-5.4817	Poor/Failed

 doi:10.20944/preprints202511.0818.v1

As shown in Table 11, the Gradient Boosting model significantly outperformed other models across all evaluation metrics.

5.4. Model-Specific Findings

5.4.1. Linear Regression

The linear regression model provided a high R^2 score of 0.9915, indicating that a linear model is quite effective in explaining the data. This suggests that rice prices have a linear relationship with the selected features. The model's success demonstrates that economic indicators can produce reasonable short to medium-term predictions using linear relationships.

5.4.2. Random Forest

The Random Forest algorithm produced a high R^2 value of 0.9976, successfully capturing complex relationships in the dataset. Beyond its predictive performance, the model's feature importance capability offered insights into which factors most significantly influence rice prices.

5.4.3. XGBoost

XGBoost delivered strong performance with an R^2 of 0.9975. The gradient boosting technique effectively modeled non-linear relationships between economic indicators and rice prices, demonstrating its ability to capture complex factors affecting price fluctuations.

5.4.4. Support Vector Regressor (SVR)

The SVR model achieved an impressive R^2 value of 0.9982 with relatively low error metrics. These results indicate the model's effectiveness in understanding complex relationships between economic indicators and rice prices, making it one of the top-performing approaches.

5.4.5. LSTM

Contrary to expectations, the LSTM model underperformed significantly with a negative R^2 value (-5.4817), indicating difficulty in explaining the variance in the target variable. This performance issue may be attributed to dataset size limitations, the complexity of the temporal structure, or hyperparameter selection. The results suggest that deep learning models may require larger datasets to show their full potential.

5.4.6. Artificial Neural Network (ANN)

The ANN model achieved a high R^2 value of 0.9903, demonstrating its ability to learn the non-linear and complex relationships between economic indicators and rice prices. However, its relatively higher RMSE compared to other models suggests the presence of some larger deviations.

5.4.7. Gradient Boosting

The Gradient Boosting model outperformed all others with the lowest MAE and RMSE values, and the highest R^2 score (0.9990). This exceptional performance demonstrates the model's ability to capture complex interactions between features and minimize errors through sequential tree-building.

5.4.8. ARIMA

The ARIMA model yielded a negative R^2 value (-0.7834), indicating that the model is not suitable for this dataset. This result suggests that rice prices have complex relationships that cannot be explained by time dimensions alone, highlighting the importance of including multiple factors beyond time in economic data analyses.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 November 2025 doi:10.20944/preprints202511.0818.v1

5.4.9. NARX-RNN

The NARX-RNN model was tested with different lag periods (3, 6, 9, and 12 months) to find the optimal prediction window. The results are presented in Table 12.

Table 12. NARX-RNN Performance with Different Lag Periods.

Lag Period (months)	MAE	RMSE	R^2
3	0.5921	0.7856	0.9968
6	0.4845	0.6363	0.9988
9	0.6733	0.8341	0.9954
12	0.7891	0.9742	0.9932

Table 12 demonstrates that a 6-month lag period provides optimal predictive performance for the NARX-RNN model, suggesting this timeframe best captures the economic relationships affecting rice prices.

5.4.10. ANFIS

The ANFIS model, which combines fuzzy logic with neural networks, showed limited success with an R^2 value of 0.3497. Due to the model's structure, only the two highest-ranked features from the feature selection process were used: household appliances (TP FG J053) and furniture and related items (TP FG J051). The model's relatively poor performance indicates that these two features alone are insufficient for accurate price prediction.

The relatively poor performance of the ANFIS and LSTM models is attributed to the limited data volume and feature dimensionality, which restrict the convergence capability of deep or hybrid architectures. These models typically require larger, more diverse datasets to capture complex non-linear dependencies effectively.

5.4.11. SHAP Interpretability Insights

SHAP Analysis was conducted to detect most important attributes for the prediction models. Top 5 items which have major impact on rice price prediction is shared in the Table 13.

Table 13. SHAP Analysis results for the top 5 items.

Rank	Feature	Description	Avg. SHAP Importance	GB Rank	RF Rank	SVR Rank
1	TP FG J125	125. INSURANCE	2.029	1	4	2
2	TP FG J062	062. OUTPATIENT SERVICES	1.481	2	1	3
3	TP FG J073	073. TRANSPORTATION SERVICES	1.448	3	5	1
4	TP FG J051	051. FURNITURE, FIXTURES, CARPETS	1.207	5	2	5
5	TP FG J061	061. MEDICAL PRODUCTS, APPLIANCES	1.021	7	3	8

SHAP analysis revealed both model-specific and consensus-level insights into the economic determinants of rice prices in Turkey. While individual models assigned varying importance weights, a robust cross-model consensus emerged around four key predictors:

- TP FG J125 (Insurance) consistently ranked among the top-2 features across all three models (avg. rank: 2.03), suggesting that household risk-mitigation expenditures are strongly coupled with food price dynamics.
- TP FG J062 (Outpatient Services) and TP FG J073 (Transportation Services) appeared in the top-3 of at least two models and within the top-5 of all three, reinforcing the role of healthcare accessibility and logistics costs in food inflation transmission.
- TP FG J051 (Furniture & Furnishings) showed high stability, ranking 2nd in Random Forest and 5th in both Gradient Boosting and SVR.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 November 2025 doi:10.20944/preprints202511.0818.v1

Notably, the lagged food CPI (TP FG J011) which is the 7th item having 0.97 average SHAP importance of the SHAP feature list, exhibited moderate influence in SVR and Random Forest but low importance in Gradient Boosting, indicating that exogenous economic indicators may dominate autoregressive effects in ensemble models. The lowest consensus was observed for TP FG J124 (Social Protection) and TP FG J127 (Other Services), which ranked 9th–10th across models.

Spearman correlation analysis of feature rankings revealed moderate agreement between models ($\varrho \approx 0.47$–0.52, $p > 0.1$), suggesting complementary but non-redundant interpretability perspectives. The average SHAP importance across models provides a robust, ensemble-weighted indicator of true economic influence, reducing model-specific bias.

5.5. Model Robustness Assessment

5.5.1. Cross Validation

Given the exceptional performance of the top-ranked Gradient Boosting model ($R^2 = 0.9990$, MAE = 0.2838), 5-fold cross-validation analysis was conducted to assess potential overfitting concerns. Cross-validation results are shown in Table 14.

Table 14. Cross Validation Results for Gradient Boosting.

Performance Metric	5 Fold Cross-Validation Score
Mean R^2 (±SD)	0.9742 ± 0.0324
Mean MAE (±SD)	0.6917 ± 0.4445
Mean RMSE (±SD)	1.6300 ± 1.3980
Original Test R^2	0.9990
CV-Test Difference	0.0248

The moderate variance (standard deviation = 0.0324) indicates acceptable model stability across different data partitions. The difference between the single-split test ($R^2 = 0.9990$) and cross-validation mean ($R^2 = 0.9742$) is 0.0248. The training performance on the original 80% split ($R^2 = 0.9992$) was nearly identical to test performance ($R^2 = 0.9990$), with a minimal gap of 0.0002, confirming excellent generalization without overfitting validating the reliability of the Gradient Boosting approach for food price prediction applications.

Similarly, when we conducted Cross Validation with 5-fold for top 3 performed models, the results shared in Table 15. The results validate top performed models reliability and show their equivalence of representing predictions.

Table 15. Cross Validation Results for Top 3 performed models.

Model	Cross-Validation Score
Gradient Boosting	0.9742 ± 0.0324
SVR	0.9896 ± 0.0168
Random Forest	0.9811 ± 0.0208

Cross-validation means were: SVR (0.9896 ± 0.0168), Random Forest (0.9811 ± 0.0208), and Gradient Boosting (0.9742 ± 0.0324). Although SVR demonstrated the highest cross-validation mean and lowest variance, Gradient Boosting achieved superior single-split test performance ($R^2 = 0.9990$ vs. 0.9982 for SVR). This suggests that while SVR exhibits more consistent performance across data partitions, Gradient Boosting may better capture specific patterns in the test period. Importantly, statistical testing confirms, the statistical analysis confirms that all three methods demonstrate excellent and statistically equivalent performance for rice price prediction.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 November 2025 doi:10.20944/preprints202511.0818.v1

5.5.2. Statistical Significance Test

Paired t-tests based on 5-fold cross-validation scores ($\alpha = 0.05$) were conducted to evaluate whether performance differences between the top three models were statistically significant.

Statistical analysis revealed no significant differences between any pair of models (all $p > 0.05$), indicating statistically equivalent performance despite ranking differences in single-split evaluation. Statistical Significance Test Results shared in Table 16.

Table 16. Statistical Significance Test Results for Top 3 performed models.

Model Comparison	t-statistic	p-value	Interpretation
Gradient Boosting vs SVR	-0.998	0.375	No significant difference
Gradient Boosting vs Random Forest	-1.000	0.374	No significant difference
SVR vs Random Forest	0.986	0.380	No significant difference

6. Discussion

This study presents a novel two-stage machine learning framework that systematically integrates econometric feature selection with predictive modeling to forecast food prices in Turkey, a critical task given the country's high inflation volatility and the substantial weight of food in its CPI basket. The results offer compelling evidence that food price dynamics are not driven solely by agricultural or supply-side factors but are closely intertwined with in cross-sectoral economic interdependencies.

6.1. Feature Selection and Economic Interconnectedness

The relationship and causality analyses revealed a complex network of linkages between food prices and seemingly unrelated sectors. Ten key predictors were identified through a consensus approach that combined Pearson's, Spearman's, and Kendall's correlation analyses, Granger causality, ARDL modeling, cointegration tests, and Random Forest feature importance. Notably, healthcare-related indicators such as Outpatient Services (TP FG J062) and Medical Products (TP FG J061) emerged as statistically significant drivers in both short- and long-run frameworks. This finding challenges conventional inflation models that treat food and health expenditures as independent and instead suggests that cost pressures in the healthcare sector may spill over into food markets, possibly through labor costs, logistics, or household budget reallocation.

Similarly, Transportation Services (TP FG J073) and Insurance (TP FG J125) were consistently ranked among the top predictors across multiple methodologies. The strong influence of transport costs aligns with established economic theory on supply chain transmission, while the prominence of insurance expenditures points to a broader household risk-management behavior that co-moves with food consumption patterns, particularly relevant in high-inflation environments where financial uncertainty is elevated.

Perhaps most striking is the identification of Education Programmes (TP FG J105) and Social Protection (TP FG J124) as cointegrated with food prices, indicating a long-term equilibrium relationship. This implies that structural social expenditures may serve as macroeconomic anchors or leading indicators of consumption stress, offering policymakers early signals of impending food inflation.

6.2. Predictive Performance and Model Robustness

In the forecasting stage, ensemble methods, particularly Gradient Boosting, Support Vector Regression (SVR), and Random Forest, demonstrated exceptional predictive accuracy ($R^2 > 0.997$), significantly outperforming both classical time series models (e.g., ARIMA) and deep learning architectures (e.g., LSTM, NARX-RNN). The failure of ARIMA ($R^2 = -0.78$) underscores the inadequacy of univariate time-series approaches in capturing the multidimensional drivers of food inflation in emerging economies. Likewise, the poor performance of LSTM ($R^2 = -5.48$) highlights the limitations of deep learning models when applied to small, noisy economic datasets without sufficient regularization or domain-informed architecture design.

The NARX-RNN model, however, provided a critical insight: a 6-month lag yielded optimal predictive performance, suggesting that exogenous economic shocks require approximately half a year to propagate through Turkey's supply chains and manifest in consumer food prices. This temporal window has direct policy relevance; it implies that early warning systems based on leading indicators (e.g., transport costs, healthcare inflation) can provide actionable foresight for monetary and fiscal interventions.

Cross-validation and statistical significance testing further validated the robustness of the top models. Although Gradient Boosting achieved the highest point performance ($R^2 = 0.9990$), paired t-tests confirmed that its advantage over SVR and Random Forest was not statistically significant ($p > 0.05$), indicating equivalent predictive reliability across these ensemble methods. This equivalence enhances practical flexibility: policymakers may choose models based on interpretability (Random Forest), robustness to outliers (SVR), or sequential error correction (Gradient Boosting).

6.3. Interpretability and Policy Implications

SHAP-based interpretability analysis bridged the technical and economic narratives by revealing a cross-model consensus on the dominant role of Insurance, Outpatient Services, and Transportation Services in shaping rice price dynamics. The consistent high ranking of insurance expenditures, a proxy for household financial vulnerability, suggests that food inflation in Turkey is not merely a market phenomenon but a reflection of broader socio-economic vulnerability. This insight supports the integration of social safety nets into inflation-mitigation strategies.

Moreover, the moderate correlation ($\rho \approx 0.47$–0.52) between model-specific feature rankings indicates that each algorithm captures complementary aspects of the underlying economic system. The ensemble-weighted SHAP importance thus offers a more holistic and less biased view of true economic influence than any single model could provide.

7. Conclusion

This study bridges econometric rigor with machine learning innovation to address a pressing challenge in emerging economies: accurate and interpretable food price forecasting. By developing a two-stage framework that first identifies causally and statistically relevant economic indicators and then evaluates a diverse set of predictive models, we demonstrate that food inflation in Turkey is deeply intertwined with sectors beyond agriculture, including healthcare, transportation, education, and household financial services.

Our findings yield three key contributions. First, we provide empirical evidence of cross-sectoral inflation transmission mechanisms, revealing that cost pressures in services can significantly influence food prices, a phenomenon underexplored in traditional macroeconomic models. Second, we establish the superiority and statistical equivalence of ensemble machine learning methods (Gradient Boosting, SVR, Random Forest) over both classical econometric and deep learning approaches for small-sample economic forecasting, offering a practical toolkit for central banks and planning agencies. Third, through SHAP-based interpretability and lag-structure analysis, we identify a six-month policy window for intervention and highlight insurance, healthcare, and transportation as high-leverage indicators for early warning systems.

These results carry direct implications for inflation targeting, social protection design, and supply chain resilience in volatile emerging markets. Future research could extend this framework to other food commodities or countries, incorporate real-time data streams, or integrate structural breaks (e.g., currency crises) into the modeling architecture. Nevertheless, the present work lays a robust foundation for data-driven, interdisciplinary approaches to food price stability and macroeconomic resilience.

Author Contributions: Conceptualization, U.T.Ş., N.A., M.N. and H.P.; Methodology, U.T.Ş., N.A., M.N. and H.P.; Formal Analysis, U.T.Ş.; Investigation, U.T.Ş, N.A., M.N. and H.P.; Data Curation, U.T.Ş.; Writing – Original Draft Preparation, U.T.Ş, N.A., M.N. and H.P.; Writing – Review & Editing, U.T.Ş., N.A., M.N. and H.P.; Visualization U.T.Ş, N.A., M.N. and H.P.; Supervision, N.A., M.N. and H.P.; Project Administration, U.T.Ş, N.A., M.N. and H.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. This study did not involve human or animal subjects, nor primary data collection from individuals.

Informed Consent Statement: Not applicable. This study did not involve human participants, personal data, or identifiable information.

Data Availability Statement: The data that support the findings of this study are publicly available from the Central Bank of the Republic of Turkey (Electronic Data Delivery System – EVDS) at https://evds2.tcmb.gov.tr and from the World Food Programme Price Database at https://data.world/wfp/7d7224ed-eff6-421f-9f96-9c8d43905f3c.

Acknowledgments: During: the preparation of this manuscript, the authors used large language models (LLMs) exclusively for improving English language clarity and grammar. These tools were not involved in study design, data analysis, interpretation, or content generation. The authors have reviewed and edited all outputs and take full responsibility for the final content of the publication.

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

References

Abidoye, R. B., Chan, A. P., Abidoye, F. A., & Oshodi, O. S. (2019). Predicting property price index using artificial intelligence techniques: Evidence from Hong Kong. International Journal of Housing Markets and Analysis, 12(6), 1072-1092.

Atalan, A. (2023). Forecasting drinking milk price based on economic, social, and environmental factors using machine learning algorithms. Agribusiness, 39(1), 214-241.

Baumeister, C., & Kilian, L. (2014). Do oil price increases cause higher food prices? Economic Policy, 29(80), 691-747.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). Time series analysis: Forecasting and control (5th ed.). John Wiley & Sons.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.

Cerveny, D. (2023). PPI and CPI: What is the relationship? [Bachelor's thesis, Charles University]. Faculty of Social Sciences.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785-794.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Applied multiple regression/correlation analysis for the behavioral sciences (3rd ed.). Lawrence Erlbaum Associates.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, 74(366), 427-431.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. Econometrica, 55(2), 251-276.

Eştürk, Ö., & Albayrak, N. (2018). Tarım ürünleri-gıda fiyat artışları ve enflasyon arasındaki ilişkinin incelenmesi. Uluslararası İktisadi Ve İdari İncelemeler Dergisi, 147-158.

Fan, X., Xu, Z., Qin, Y., & Škare, M. (2023). Quantifying the short-and long-run impact of inflation-related price volatility on knowledge asset investment. Journal of Business Research, 165, 114048.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29(5), 1189-1232.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37(3), 424-438.

 doi:10.20944/preprints202511.0818.v1

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780.

Jang, J. S. R. (1993). ANFIS: Adaptive-network-based fuzzy inference system. IEEE Transactions on Systems, Man, and Cybernetics, 23(3), 665-685.

Ji, M., Liu, P., Deng, Z., & Wu, Q. (2022). Prediction of national agricultural products wholesale price index in China using deep learning. Progress in Artificial Intelligence, 1-9.

Karagöl, V. (2023). Ekonomik politika belirsizliğinin gıda fiyatlarına etkisi: seçilmiş ülkeler için zamanla değişen nedensellik analizi. İktisat Politikası Araştırmaları Dergisi, 10(2), 409-433.

Kresova, S., & Hess, S. (2022). Identifying the determinants of regional raw milk prices in Russia using machine learning. Agriculture, 12(7), 1006.

Lin, T., Horne, B. G., Tino, P., & Giles, C. L. (1996). Learning long-term dependencies in NARX recurrent neural networks. IEEE Transactions on Neural Networks, 7(6), 1329-1338.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems (NeurIPS), 30.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., ... & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence, 2(1), 56–67.

Lutoslawski, K., Hernes, M., Radomska, J., Hajdas, M., Walaszczyk, E., & Kozina, A. (2021). Food demand prediction using the nonlinear autoregressive exogenous neural network. IEEE Access, 9, 146123-146136.

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. PLOS ONE, 13(3), e0194889.

Oyeleke, O. J., & Ojediran, S. (2018). Exploring the relationship between consumer price index (CPI) and producer price index (PPI) in Nigeria. International Journal of Statistics and Applications, 8(2), 42-46.

Ozpolat, A. (2020). Causal link between consumer prices index and producer prices index: An evidence from central and Eastern European Countries (CEECs). Adam Academy Journal of Social Sciences, 10(2), 319-332.

Özçelik, Ö., & Uslu, N. (2024). Gıda Enflasyonunun Belirleyicileri Üzerine Bir Analiz: Türkiye Örneği. Dumlupınar Üniversitesi Sosyal Bilimler Dergisi(79), 289-309. https://doi.org/10.51290/dpusbe.1391994

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. Journal of Applied Econometrics, 16(3), 289-326.

Qian, J., Dai, B., Wang, B., Zha, Y., & Song, Q. (2022). Traceability in food processing: problems, methods, and performance evaluations—a review. Critical Reviews in Food Science and Nutrition, 62(3), 679-692.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533-536.

Vapnik, V. N. (1995). The nature of statistical learning theory. Springer.

Venkateswara Rao, K., Srilatha, D., Jagan Mohan Reddy, D., Desanamukula, V. S., & Kejela, M. L. (2022). Regression based price prediction of staple food materials using multivariate models. Scientific Programming, 2022, Article 9547039.

Wanjuki, T. M., Wagala, A., & Muriithi, D. K. (2022). Evaluating the predictive ability of seasonal autoregressive integrated moving average (SARIMA) models using food and beverages price index in Kenya. European Journal of Mathematics and Statistics, 3(2), 28-38.

Warren-Vega, W. M., Aguilar-Hernández, D. E., Zárate-Guzmán, A. I., Campos-Rodríguez, A., & Romero-Cano, L. A. (2022). Development of a predictive model for agave prices employing environmental, economic, and social factors: Towards a planned supply chain for agave-tequila industry. Foods, 11(8), 1138.

Xie, M., Ding, L., Xia, Y., Guo, J., Pan, J., & Wang, H. (2021). Does artificial intelligence affect the pattern of skill demand? Evidence from Chinese manufacturing firms. Economic Modelling, 96, 295-309.

Yu, C. P. (2016). Why are there always inconsistent answers to the relation between the PPI and CPI? Re-examination using panel data analysis. International Review of Accounting, Banking & Finance, 8(1), 1-17.

Muthayya, S., Sugimoto, J. D., Montgomery, S., & Maberly, G. F. (2014). An overview of global rice production, supply, trade, and consumption. Annals of the New York Academy of Sciences, 1324(1), 7-14.

 doi:10.20944/preprints202511.0818.v1

Valera, H. G. A. (2022). Is rice price a major source of inflation in the Philippines? A panel data analysis. Applied Economics Letters, 29(16), 1528-1532.

TCMB (EVDS Verinin Merkezi). (2024). https://evds2.tcmb.gov.tr/index.php?/evds/serieMarket

World Food Programme Price Database (2024), https://data.world/wfp/7d7224ed-eff6-421f-9f96-9c8d43905f3c