SOSYAL BİLİMLER DERGİSİ
Journal of Social Sciences
FIRAT ÜNİVERSİTESİ
p-ISSN:1300-9702 e-ISSN: 2149-3243

USE OF MACHINE LEARNING AND DEEP LEARNING METHODS IN HOUSING PRICE INDEX ESTIMATION: AN ANALYSIS ON ANKARA AND ISTANBUL

Konut Fiyat Endeksi Tahmininde Makine Öğrenmesi ve Derin Öğrenme Yöntemlerinin Kullanımı: Ankara ve İstanbul Üzerine Bir Analiz

Ahmed İhsan ŞİMŞEK[1]

[1]*Dr. Öğr. Üyesi, Fırat Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü, Elazığ, aisimsek@firat.edu.tr*, orcid.org/ 0000-0002-2900-3032

Araştırma Makalesi/Research Article

Makale Bilgisi
Geliş/Received:
06.12.2023
Kabul/Accepted:
06.09.2024

DOI:
10.18069/firatsbed.1401213

Anahtar Kelimeler
House Price Forecasting, Time Series, Deep Learning, Machine Learning, Decision Support

ABSTRACT

Factors such as supply chain difficulties, rising energy and oil prices, economic recession and production loss due to the pandemic have increased costs and inflation. All these factors have also seriously affected the construction sector. This study aims to create a deep learning and machine learning focused forecasting system based on Istanbul and Ankara monthly housing price index data for the period of January 2010 to June 2023. The system was created using approximately 13 years of housing interest rates, Consumer Price Index, XGMYO, Monthly Average Dollar and XAU data as the basis of the Istanbul and Ankara Housing Price Index forecasting process. During the research process, different RNN structures (Long and Short Term Memory, Gated Recurrent Unit) and machine learning (Random Forest) structures were tested and the effectiveness of these structures in housing price index forecasting was compared. The performances of the models were evaluated using RMSE, MSE, MAE, MAPE and R2 statistics. According to the results obtained, the method that gave the best performance for both provinces is the RF model. This is followed by LSTM and GRU models, respectively.

Keywords
Konut Fiyat Tahmini, Zaman Serisi, Derin Öğrenme, Makine Öğrenmesi, Karar Destek

ÖZ

Pandeminin etkisiyle tedarik zincirinde yaşanan zorluklar, enerji ve petrol fiyatlarının yükselmesi, ekonomik durgunluk ve üretim kaybı gibi faktörler, maliyetleri ve enflasyonu artırmıştır. Tüm bu etkenler inşaat sektörünü de ciddi biçimde etkilemiştir. Bu çalışma, Ocak 2010 ile Haziran 2023 dönemine ait İstanbul ve Ankara aylık konut fiyat endeksi verilerine dayalı olarak derin öğrenme ve makine öğrenmesi odaklı bir tahmin sistemi oluşturma amacını taşımaktadır. İstanbul ve Ankara Konut Fiyat Endeksi tahminleme işleminin temeli olarak yaklaşık 13 yıllık konut faizleri, Tüketici Fiyat Endeksi, XGMYO, Aylık Ortalama Dolar ve XAU verileri kullanılarak sistem oluşturulmuştur. Araştırma sürecinde, farklı RNN yapıları (Long and Short Term Memory, Gated Recurrent Unit) ve makine öğrenmesi (Random Forest) yapıları denenmiş, bu yapıların konut fiyat endeksi tahminindeki etkinliği karşılaştırılmıştır. Modellerin performansları RMSE, MSE, MAE, MAPE ve R2 istatistikleri kullanılarak değerlendirilmiştir. Elde edilen sonuçlara göre her iki il için en iyi performansı veren yöntem RF modelidir. Sonrasında ise sırasıyla LSTM ve GRU modelleri gelmektedir.

Atıf/Citation: Şimsek, A.İ. (2024). Use of Machine Learning and Deep Learning Methods in Housing Price Index Estimation: An Analysis on Ankara and İstanbul. *Fırat Üniversitesi Sosyal Bilimler Dergisi,* 34, 3, 1345-1353.
Sorumlu yazar/Corresponding author: Ahmed İhsan ŞİMŞEK, aisimsek@firat.edu.tr.

1. Introduction

The Covid-19 pandemic, which started in early 2020, has affected the housing sector as in many other industries. In particular, the need to work from home due to the pandemic has increased the demand for larger residences. At the same time, dense population and city migration contributed to the increase in housing demand. Factors such as difficulties in the supply chain due to the pandemic, rising energy and oil prices, economic stagnation and loss of production have increased costs and inflation. In February 2022, Russia's invasion of Ukraine exacerbated supply chain and transportation problems and contributed to global inflation. All these factors accelerated the rise in construction costs, especially after 2020, while shrinking the housing supply. Against this shrinking supply, rising demand and costs caused housing prices to rise (Çetin, 2022).

Nevertheless, the construction sector in Turkey is a strong contributor to economic growth. High investments in this sector affected the supply-demand balance, positioned the housing market as an advanced investment instrument, and led to significant increases in housing prices. Moreover, the expansion in housing loans has had a significant impact on economic growth by affecting the banking sector. Therefore, forecasting the future of the housing market is of great importance (Akay et al., 2019).

House price indices published monthly by the Central Bank of the Republic of Turkey (CBRT) are considered a critical indicator for monitoring house prices across Turkey. This study analyzes the house price index data, which is frequently used in research on the housing sector and in evaluating the performance of the sector. One of the main purposes of calculating the house price index is to monitor price changes in Turkey's housing market. Therefore, the Central Bank of the Republic of Turkey constructs a nationwide index. When calculating the house price index, all houses are offered for sale and evaluated. In the calculation, price data of all houses are used except for the year of construction (Saraç and Hacıimamoğlu, 2018).

This paper aims to build a deep learning and machine learning-based forecasting system based on monthly house price index data for Istanbul and Ankara for the period 2010:01 - 2023:06. Recurrent neural networks (RNN), which are frequently used in the field of deep learning, are known for their success, especially on sequential data types, such as time series. The system is trained on monthly data and the relevant input data is used as input during this training process. The system was built using approximately 13 years of housing interest rates, Consumer Price Index, XGMYO, Monthly Average Dollar, and XAU data as the basis for the Istanbul and Ankara House Price Index forecasting process. During the research process, different RNN structures (LSTM, Gated Recurrent Unit) and machine learning (Random Forest) structures were tested and their effectiveness in predicting the house price index was compared.

In this framework, the first stage of the study summarizes the related literature, followed by the methodology and the data used. The obtained comparative forecast values are presented. Finally, the study is completed by presenting the interpretation of the results and recommendations.

2. Literature Review

In the literature, deep learning and machine learning methods are frequently used in prediction. Some studies using Random Forest (RF), Long and Short Term Memory (LSTM), and Gated Recurrent Unit (GRU) algorithms as methods for model identification and estimation are summarized below. In 2020, Dutta et al. conducted a study focused on exploring machine learning techniques that surpass traditional time series models when it comes to predicting Bitcoin values. The research employed a range of advanced machine learning forecasting methods to estimate daily Bitcoin values more accurately. The study made use of sequential models like Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) in order to capture complex intra-time interactions and improve Bitcoin price forecasts. Using the Root Mean Square Error (RMSE) measure, multiple methods for predicting the price of bitcoin were evaluated and compared. The results indicate that the utilization of recurrent dropout in conjunction with the Gated Recurrent Unit (GRU) model yields superior performance compared to conventional models. In addition, the application of basic trading strategies in conjunction with the recommended GRU model holds the capacity to yield financial gains through effective knowledge acquisition. In 2021, Wang and Wang proposed an advanced machine-learning technique that demonstrates high precision in predicting future prices of global crude oil futures. In order to tackle this intricate issue, the present study introduced a novel approach by developing a computational framework referred to as the "random deep bidirectional gated recurrent unit neural network." The training approach of this model incorporated the utilization of the 'random inheritance formula' to account for temporal variations in historical data patterns. Furthermore, with the integration of deep pairwise learning in conjunction with the random inheritance technique, the model successfully obtained valuable insights from past data, resulting in enhanced accuracy in prediction. A comparison evaluation of the suggested model was conducted against a number of models, including SVM, GRU, ERNN, LSTM, DBGRUNN, and RIF-GRUNN, using a variety of evaluation criteria including the innovative synchronization assessment approach known as "q-DSCID." The suggested model regularly outperforms other models, as shown by the excellent results in the R2, MAE, TIC, RMSE, and SMAPE measures for forecasting Brent crude oil futures prices. In a research conducted in 2022, Cho et al. created a model to forecast water levels. This work is essential because flood predictions can help to lessen property damage and save lives. The study uses water level data to build a prediction model, which is a crucial indication of flood disasters. For predicting water levels, recurrent neural network models like LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit) are used. Meteorological data, including upstream and downstream water levels, temperature, humidity, and precipitation, make up the first data inputs. Trials with multiple model architectures and input data formats showed that the best results were obtained when the LSTM-GRU-based model and ASOS (Automated Synoptic Observing System) meteorological data were included in the input data. With a mean squared error (MSE) value of 3.92, a Nash-Sutcliffe efficiency coefficient (NSE) value of 0.942, and a mean absolute error (MAE) value of 2.22, the testing findings consistently demonstrated improved performance. Additionally, the test data contained historically high water levels in the research location, hence the largest degree of error in water level forecasts was small at 55.49. In 2023, Song and Choi conducted a study exploring deep learning techniques applicable to various characteristics in the financial sector. This research focuses on the prediction of single and multi-time step closing prices for the DAX, DOW, and S&P500 indices using innovative hybrid models. These models incorporate a combination of ensemble models, gated recurrent unit (GRU)-CNN, convolutional neural network (CNN)-LSTM, and recurrent neural network (RNN)-based methods. Additionally, the study recommends considering the averages of high and low prices for stock market indices. The experimental results highlight the superior performance of these models compared to traditional machine learning approaches. Particularly noteworthy is the achievement of a 48.1% improvement in mean squared error (MSE) and a 40.7% improvement in mean absolute error (MAE) for single time-step predictions. Furthermore, the models exhibit an 81.5% success rate in multi-time-step forecasts.

When the literature is examined, there are various studies on the prediction of house prices using machine and deep learning methods. Hong et al. (2020) compare a classic hedonic pricing model to the features of a Random Forest (RF) method-based housing price forecasting model. Data on apartment transactions in Gangnam, one of South Korea's most developed areas, were used for the study between 2006 and 2017. The study's findings indicate that the Random Forest-based forecasting model is more accurate than the more conventional basic linear regression-based model. It is stated that there is only a 5.5% average percentage difference between the RF forecaster's anticipated and actual market pricing. For the forecaster based on basic linear regression, this variation is about 20%. The likelihood that the price predicted by the RF forecaster would be within 5% of the actual market price is found to be 72%, compared to a chance of roughly 17.5% for regression-based predictions. In 2021, Rawool et al. conducted a study aimed at utilizing machine learning techniques for the estimation of housing values. The research methodology revolves around the prediction of property values through the application of various machine learning methods. These algorithms encompass techniques such as K-means regression, Random Forest Regression, Decision Tree Regression, and Linear Regression. The research findings highlight Random Forest Regression as the most accurate method. The authors contend that adopting this approach instead of engaging a traditional real estate agent can empower consumers in making informed real estate investments. The Random Forest machine learning technique was utilized by Adetunji et al. in 2022 to forecast home values. They underline in their research that the Home Price Index (HPI), a commonly used metric, is insufficient for forecasting changes in home values. This study used a genuine Boston housing dataset with 506 items and 14 characteristics from the UCI Machine Learning Repository. Comparisons between projected and real prices were used to assess the effectiveness of the suggested prediction model, and it was found that the model had a respectable accuracy margin of 5 when predicting actual values. These findings imply that the Random Forest machine learning technique might be a useful tool for forecasting home prices. Different forecasting methods were employed by Tanamal et al. (2023) to forecast

Surabaya real estate values. The research approach makes use of data from real estate agent interviews. Multiple real estate agents were interviewed, and the findings of those interviews were collated. The seventeen most significant factors that influence home prices were determined using the information from these interviews. The study forecasts Surabaya real estate values using the Random Forest machine learning method. The dataset was divided into training and testing data using an 80/20 ratio, and predictions of the outcomes were 88% accurate.

3. Data and Methodology

In this study, monthly data on house price index, house loan, consumer price index (tüfe), XGMYO, USD/TRY and Gold/TRY between January 2010 and June 23 are used. In the study, five different criteria were taken as inputs for the house price index estimation. Housing loan interest rates and loan amounts affect the purchasing power of the house and therefore the demand. This has a direct effect on house prices. The consumer price index reflects the general inflation rate. Inflation is thought to affect housing costs and therefore house prices. The XGMYO index shows the performance of real estate investment trusts. It will reflect the general trends and expectations in the real estate market. The USD/TL exchange rate can show the economic effects of exchange rates and the interest of foreign investors in the housing market in Turkey. Exchange rate fluctuations affect house prices. The gold/TL exchange rate affects the price of gold, which is seen as a safe haven in times of economic uncertainty, and therefore investor behavior. Changes in gold prices will indirectly reflect on the housing market. These data provide a holistic evaluation of economic and financial effects in estimating house prices. The data on house price index, house interest rate, TFP, USD/TRY and Gold/TRY are obtained from the Central Bank's EVDS system. XGMYO data were obtained from www.invensting.com website. In this study, GRU, LSTM and RF models are applied for house price forecasting for both Istanbul and Ankara provinces using the data on housing interest rate, TFP, XGMYO, USD/TRY and Gold/TRY. The Python 3.11 programming language was used in the study to create and manipulate models and manipulate data. Pandas and NumPy were the libraries used for the data manipulation and analysis activities. The Scikit-learn toolkit was utilized to create machine learning models and evaluation measures. Keras and TensorFlow libraries were used to create deep learning models. To meet high-performance computing requirements, analyses were carried out on a system that included an AMD Ryzen 7 processor, an NVIDIA GeForce RTX 2080 graphics processing unit (8 GB GDDR6), 32 GB DDR4 memory, and 1 TB NVMe SSD storage. Using this hardware and software architecture made it easier to train and test crude oil price prediction models in an effective and successful manner.

The performance metrics of the models are calculated using root mean square error (RMSE), mean square error (MSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and coefficient of determination (R^2). The calculation formulas of the error coefficients used are given below in Equation 1-5 respectively.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)}{N}} \tag{1}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{2}$$

$$MAE = \frac{\sum_{i=1}^{n}|y_i - x_i|}{n} \tag{3}$$

$$MAPE = \frac{\sum_{t=1}^{n}\frac{u_t}{Y_t}}{n} * 100 \tag{4}$$

$$R^2 = 1 - \frac{\sum_i(y_i - \hat{y}_i)^2}{\sum_i(y_i - \mu)^2} \tag{5}$$

The min-max normalization process given in Equation (6) was applied to transform the data sets into the 0-1 range.

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{6}$$

3.1. GRU Model

The GRU model can be considered as a derivative of the Recurrent Neural Network (RNN) model. The initial proposition of this model was put forth by Cho et al. in 2014. In order to effectively utilize the historical checkin trajectories of users for the purpose of learning the dependence link between check-ins, certain researchers decide to utilize Recurrent Neural Networks (RNN). Nevertheless, the Recurrent Neural Network (RNN) is constrained by its architecture in terms of its ability to successfully identify the long-range reliance that exists between check-ins (Jozefowicz et al., 2015). Similar to the LSTM system, the GRU utilizes gates units in order to control the internal flow of data. Yet it exhibits a notable distinction in that it lacks distinct cells for memory (Chung et al., 2014). In the GRU approach, unlike the LSTM approach, there are only update(z_t) and reset (r_t) gates. In the setting of a GRU, the update gate assumes a pivotal function in determining the degree to which the input x_t and the previous output h_{t-1} should be propagated to the succeeding cell. Conversely, the reset gate fulfills the function of ascertaining the extent to which the preceding information ought to be disregarded. The current state of memory content enables the effective transfer of relevant information to the next iteration, as determined by the weight W (Gao et al., 2021). The initial step in the GRU approach involves the determination of the update and reset gate. Equations 7 and 8 are employed for this objective.

Update Gate:

$$z_t = \sigma(W_z * [h_{t-1}, x_t)] \tag{7}$$

Reset Gate:

$$r_t = \sigma(W_r * [h_{t-1}, x_t)] \tag{8}$$

After the reset and update gates are defined, the candidate status value $\widetilde{h}_t$ of the GRU is determined. Equation 9 is used for this. Then the final output status h_t is determined using Equation 9.

$$\tilde{h}_t = tanh(W_{\tilde{h}_t} * ([r_t * h_{t-1}, x_t]) \tag{9}$$

$$h_t = (1\text{-}z_t)^* h_{t-1} + z_t {}^*\tilde{h}_t \tag{10}$$

3.2. RF Model (Random Forest)

One potential strategy for improving the accuracy of generalize involves selecting a subset of the input and generating many individual trees. The random-subspace technique was first suggested by Ho (1995) and subsequently developed into the random forest algorithm by Breiman (2001). The random forest model is a learning methodology that falls under the genre of ensemble tree-based methods. The proposed methodology involves generating forecasts by taking the average of the outcomes produced by numerous individual trees. The development of individual trees is predicated upon the utilization of bootstrap samples as opposed to the primary dataset. The method known as bootstrap aggregating, often commonly referred to as bagging, is utilized as a means to address the problem of overfitting (Schonlau & Zou, 2020). Figure 1 shows the the structure of the RF approach.

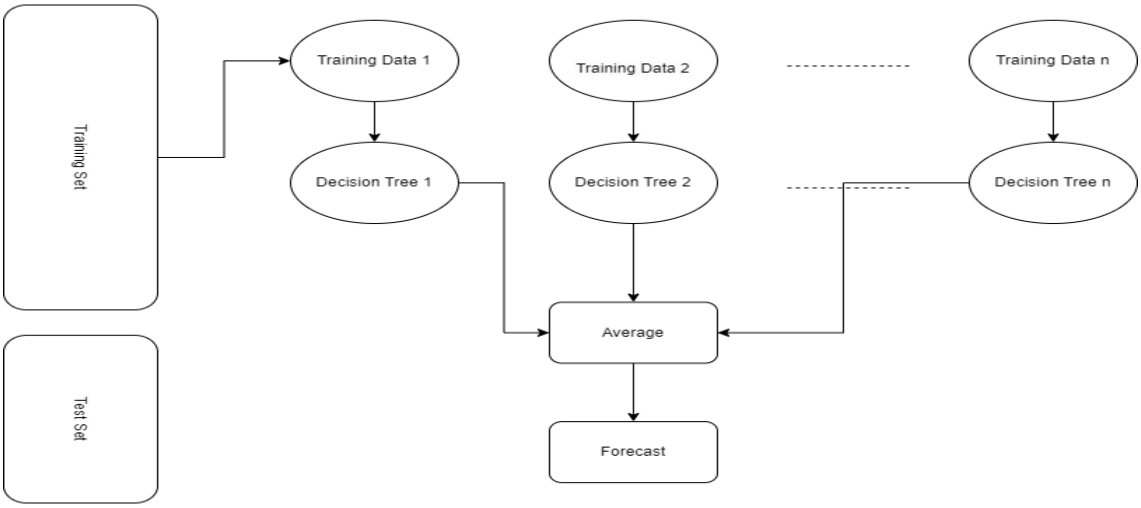

Figure 1. Structure of RF Approach

3.3. LSTM Model (Long Short Term Memory)

The LSTM model was first proposed by Hochreiter and Schmidhuber in 1997. The fundamental RNN structure proves to be inadequate in achieving precise predictions, thus requiring the adoption of the more complex long short-term memory (LSTM) structure within the RNN system. The LSTM unit is a type of recurrent network unit that exhibits exceptional proficiency in retaining information over extended or brief periods of time. The crucial aspect of this capability is in the utilization of recurrent components without the incorporation of an activation function. Consequently, the stored value is not subjected to iterative compression over time, and the gradient or punishment term does not exhibit a tendency to decrease when Backpropagation through time is employed for training (Chen et al., 2017). The stages of the LSTM structure are given in Equation 11-16.

The forget gate of the LSTM algorithm is responsible for accurately discarding the cell state from the previous sequence. The current input of the time series is denoted as x_t , whereas its preceding hidden state is denoted as h_{t-1}.The activation function σ_g is responsible for processing both of these values. The output vector f_t is generated as a result of this processing, and it is directly associated with the forget gate. The expression of this relationship can be represented by Equation 11. In Equation 11, the bias coefficient is denoted as b_f , the forget gates are denoted as U_f and W_f , and the activation function is denoted as σ_g.

Equations 11 and 12 describe the relationship between the current point in the time series input, denoted as x_t , and the hidden state h_{t-1} from the previous time frame. These variables are responsible for determining the values of the coefficients i_t and C'_t within this gate. The calculation of these coefficients utilizes the activation function. The weight coefficients are represented by symbols such as W_i, U_i, W_c, and U_c, whereas the activation function is marked by the initials σ_g and σ_c.

$$i_t = \sigma_g\ (W_i x_t + U_i h_{t-1} + b_i) \qquad (11)$$

$$C'_t = \sigma_c\ (W_c x_t + U_c h_{t-1} + b_c) \qquad (12)$$

In Equation 10, the cell state, represented as C_t, undergoes an update process in which it is obtained by combining the result of the given input gate's output, i_t, and the cell candidate data, C'_t, with the multiplication of the forget gate's output, f_t , and the prior cell state, C_{t-1}.

$$C_t = f_t\ x\ C_{t-1} + i_t\ x\ C'_t \qquad (13)$$

Equation 14 demonstrates the process by which the output vector o_t is generated by the application of the activation function σ_g to the input vectors h_{t-1} and x_t. The bias coefficient, denoted as b_o, together with the weight coefficients of the cell state, represented as W_o and U_o, are associated with the input gate. The output gate's value o_t is then multiplied by the current sequential cell state C_t subsequent to its generation. The output of the last hidden layer is obtained by applying the activation function tanh to the resulting value, as illustrated in Equation 15.

$$o_t = \sigma_g\ (W_o x_t + U_o h_{t-1} + b_o) \qquad (14)$$

$$h_t = o_t * tanh\ (C_t) \qquad (15)$$

4. Results and Discussion

This study employs LSTM, GRU, and RF methodologies to predict property prices. This study employs data on house loan, the consumer price index (TÜFE), XGMYO, USD/TRY exchange rate, and Gold/TRY exchange rate to make predictions about the house price index. The statistics were sourced from the official website of the Central Bank of the Republic of Turkey, namely the Electronic Statistics Delivery System (EVDS), as well as from investing.com. The study utilizes a comprehensive data collection spanning a duration of 162 months, commencing from January 2010 and concluding in June 2023. The performance evaluation of forecasting models commonly employs error statistics such as RMSE, MSE, MAE, MAPE and R2. Table 1 and Table 2 displays the error coefficients derived from the LSTM, GRU, and RF models for Ankara and İstanbul.

Table 1. Statistical results of the models for ankara

Coef.	LSTM	GRU	RF
RMSE	18.79768	42.93883	14.11819
MSE	353.3529	1843.743	199.3234
MAE	11.74884	28.06585	5.640152
MAPE	0.095432	0.21928	0.025373
R2	0.984185	0.917482	0.991079

Table 2. Statistical results of the models for istanbul

Coef.	LSTM	GRU	RF
RMSE	21.32729	24.82598	10.82297
MSE	454.8535	616.3293	117.1366
MAE	16.96418	17.83969	4.556758
MAPE	0.182064	0.175447	0.023618
R2	0.977324	0.969274	0.99416

Table 1 presents the statistical coefficients for both Ankara and Istanbul. Upon analysis of the table, it becomes apparent that the RF model exhibits the lowest RMSE value for both Ankara and Istanbul. RMSE is a quantitative metric used to assess the proximity between predicted values and the corresponding actual values. Better forecasts are typically characterized by lower values of RMSE. The MSE is calculated as the average of the squared differences between the forecasted values and the actual values. A lower MSE is indicative of superior model performance. Hence, it can be concluded that RF algorithm exhibits the lowest MSE value for both urban areas. The MAE parameter is the average of the absolute values of the forecast errors. A smaller MAE is indicative of superior performance of the model. The analysis of the application to both Istanbul and Ankara reveals that the RF model consistently yields the most favorable outcome. The MAPE quantifies the average magnitude of mistakes in a forecast, expressed as a percentage of the actual values. A smaller MAPE is indicative of superior performance of the model. RF model has the lowest MAPE score for both cities. The coefficient of determination, denoted as R2, quantifies the proportion of the variance in the predicted values that can be explained by the variance in the actual values. Models with higher R2 values are indicative of a stronger fit to the data. RF model exhibits the highest R2 value for both cities, indicating its superior performance in elucidating the data. Consequently, it is evident that RF model consistently outperforms other models in terms of all error measures, for both the Istanbul and Ankara datasets. The outcomes of RF model for both Ankara and Istanbul are depicted in Figure 2.

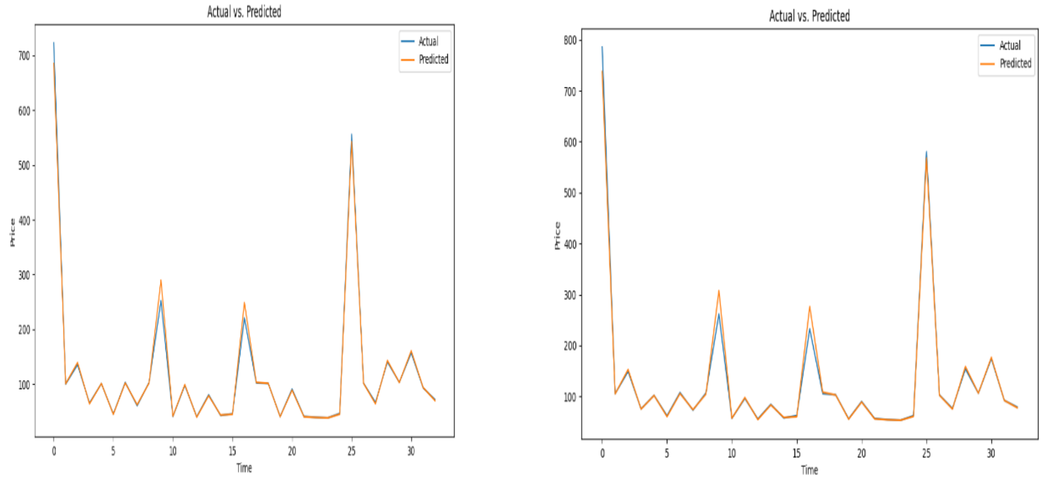

Figure 2. Actual and Predicted Values for RF Model for Ankara and Istanbul

This study utilizes data on house loans, the consumer price index, XGMYO, USD/TRY exchange rate, and Gold/TRY exchange rate to make predictions about the house price index in Ankara and Istanbul. The data undergo processing through the utilization of LSTM, GRU, and RF methodologies. Subsequently, forecasting models are developed for each respective approach, and an evaluation of their respective performances is conducted. Based on the findings, it is evident that the RF technique has superior predicting performance across all performance parameters for both Ankara and Istanbul.

The construction industry in Turkey plays a significant role in promoting economic progress. The findings derived from this study are anticipated to provide substantial benefits to both individual and institutional investors seeking to allocate their resources within the construction sector, which holds great significance for the Turkish economy. This will facilitate the anticipation of the trajectory of the house price index. In future research endeavors, it is recommended to use supplementary variables, such as Gross Domestic Product (GDP), per capita income, and the consumer confidence index, alongside the existing variables employed in this study. Furthermore, an enhanced forecasting performance can be achieved by integrating the LSTM and RF models, which have demonstrated superior outcomes, to construct a hybrid model.

References

Adetunji, A. B., Akande, O. N., Ajala, F. A., Oyewo, O., Akande, Y. F., & Oluwadara, G. (2022). House Price Prediction Using Random Forest Machine Learning Technique. *Procedia Computer Science, 199, 806-813.*

Akay, E. Ç., Topal, K. H., Kizilarslan, S., & Bulbul, H. (2019). Türkiye Konut Fiyat Endeksi Öngörüsü: ARIMA, Rassal Orman Ve Arima-Rassal Orman. *Pressacademia Procedia, 10(1), 7-11.*

Breiman, L. 2001. Random Forests. *Machine Learning 45: 5–32.*

Cho, M., Kim, C., Jung, K., & Jung, H. (2022). Water Level Prediction Model Applying A Long Short-Term Memory (Lstm)–Gated Recurrent Unit (Gru) Method For Flood Prediction. *Water, 14(14), 2221.*

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical Evaluation Of Gated Recurrent Neural Networks On Sequence Modeling. *Proceedings Of The Neural Information Processing Systems Workshop On Deep Learning., 1–9. Http://Arxiv.Org/Abs/1412.3555*

Çetin, D. T. (2022). Antalya-Isparta-Burdur Bölgesi Konut Fiyat Endeksinin Makroekonomik Göstergeler Ve Hisse Senedi Endeksi Kullanılarak Yapay Zekâ İle Tahmini. *Abant Sosyal Bilimler Dergisi, 22(3), 1363-1380.*

Dutta, A., Kumar, S., & Basu, M. (2020). A Gated Recurrent Unit Approach To Bitcoin Price Prediction. *Journal Of Risk And Financial Management, 13(2), 23.*

Ho, T. K. 1995. Random Decision Forests. *In Proceedings Of 3rd International Conference On Document Analysis And Recognition, 278–282. Piscataway, NJ: IEEE.*

Hong, J., Choi, H., & Kim, W. S. (2020). A House Price Valuation Based On The Random Forest Approach: The Mass Appraisal Of Residential Property İn South Korea. *International Journal Of Strategic Property*

Management, 24(3), 140-152.

Jozefowicz, R., Zaremba, W., & Sutskever, I. (2015). An Empirical Exploration Of Recurrent Network Architectures. *32nd International Conference On Machine Learning, ICML 2015, 3, 2332–2340.*

Rawool, A. G., Rogye, D. V., Rane, S. G., & Bharadi, V. A. (2021). House Price Prediction Using Machine Learning. *Int. J. Res. Appl. Sci. Eng. Technol, 9, 686-692.*

Saraç, H., & Hacıimamoğlu, T. U. (2018). Konut Fiyat Endeksi Verilerinin Klasik Ayrıştırma Ve Arıma Yöntemi İle Tahmin Edilmesi. *Journal Of International Social Research, 11(59).*

Schonlau, M., & Zou, R. Y. (2020). The Random Forest Algorithm For Statistical Learning. *Stata Journal, 20(1), 3–29. Https://Doi.Org/10.1177/1536867X20909688*

Song, H., & Choi, H. (2023). Forecasting Stock Market Indices Using The Recurrent Neural Network Based Hybrid Models: CNN-LSTM, GRU-CNN, And Ensemble Models. *Applied Sciences, 13(7), 4644.*

R. Tanamal, N. Minoque, T. Wiradinata, Y. Soekamto and T. Ratih, "House price prediction model using random forest in surabaya city", *TEM Journal, pp. 126-132, Feb. 2023.*

Wang, B., & Wang, J. (2021). Energy Futures Price Prediction And Evaluation Model With Deep Bidirectional Gated Recurrent Unit Neural Network And RIF-Based Algorithm. *Energy, 216, 119299.*

X Chen, L. Wei, J. Xu. House Price Prediction Using LSTM. *Computing Research Repository (Corr). Abs/1709.08432, September 2017*

Etik, Beyan ve Açıklamalar

1. Etik Kurul izni ile ilgili;

☑**.**Bu çalışmanın yazarları, Etik Kurul İznine gerek olmadığını beyan etmektedir.

2. Bu çalışmanın yazarları, araştırma ve yayın etiği ilkelerine uyduklarını kabul etmektedir.

3. Bu çalışmanın yazarları kullanmış oldukları resim, şekil, fotoğraf ve benzeri belgelerin kullanımında tüm sorumlulukları kabul etmektedir.

4. Bu çalışmanın benzerlik raporu bulunmaktadır.