electronics

Article

Application of Artificial Intelligence for Predicting Real Estate Prices: The Case of Saudi Arabia

Elham Alzain [1,†], Ali Saleh Alshebami [1,*,†], Theyazn H. H. Aldhyani [1,2,*,†] and Saleh Nagi Alsubari [2]

1 Applied College in Abqaiq, King Faisal University, Al-Ahsa 31982, Saudi Arabia
2 Department of Computer Science, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad 431001, India
* Correspondence: aalshebami@kfu.edu.sa (A.S.A.); taldhyani@kfu.edu.sa (T.H.H.A.)
† The Saudi Investment Bank Chair for Investment Awareness Studies, The Deanship of Scientific Research, The Vice Presidency for Graduate Studies and Scientific Research.

Abstract: The housing market is a crucial economic indicator to which the government must pay special attention because of its impact on the lives of freshly minted city inhabitants. As a guide for government regulation, individual property purchases, third-party evaluation, and understanding how housing prices are distributed geographically may be of great practical use. Therefore, much research has been conducted on how to arrive at a more accurate and efficient way of calculating housing prices in the current market. The goal of this study was to use the artificial neural network (ANN) technique to correctly identify real estate prices. The novelty of the proposed research is to build a prediction model based on ANN for predicting future house prices in Saudi Arabia. The dataset was collected from Aqar in four main Saudi Arabian cities: Riyadh, Jeddah, Dammam, and Al-Khobar. The results showed that the experimental and predicted values were very close. The results of the proposed system were compared with different existing prediction systems, and the developed model achieved high performance. This forecasting system can also help increase investment in the real estate sector. The ANN model could appropriately estimate the housing prices currently available on the market, according to the findings of the assessments of the model. Thus, this study provides a suitable decision support or adaptive suggestion approach for estimating the ideal sales prices of residential properties. This solution is urgently required by both investors and the general population as a whole.

Keywords: artificial intelligence; investment; prediction house timeseries prediction model

Citation: Alzain, E.; Alshebami, A.S.; Aldhyani, T.H.H.; Alsubari, S.N. Application of Artificial Intelligence for Predicting Real Estate Prices: The Case of Saudi Arabia. *Electronics* **2022**, *11*, 3448. https://doi.org/10.3390/electronics11213448

Academic Editor: Amir Mosavi

Received: 27 September 2022
Accepted: 18 October 2022
Published: 25 October 2022

1. Introduction

The housing market, which is a segment of the real estate business, is an essential component of any viable economy. In many countries, the possession of one's own piece of real estate is regarded as a symbol of social standing, and attaining this status is a goal for many young people who are just starting out in their careers. On the other hand, investors are driven to the housing market because they regard property as an investment opportunity rather than a mere commodity [1].

When entering the property market, prospective homeowners and investors alike do so with the expectation of making a profit from further price increases. Property value is directly related to the homeownership rate. Previous studies have discussed the sustainability of high homeownership rates, especially in developing countries. The key element determining the sustainability of the housing market is affordability, with affordable and efficient housing essential to long-term viability. Whether housing is cost-effective also impacts the sustainability of its usage as an investment strategy. It should also be noted that the real estate market is more stable than comparatively volatile financial markets, such as foreign exchanges, interest rates, and the stock market [2].

In the real estate sector, which has become a highly profitable investment source, especially in the last 15 years, profitability is determined by housing prices. Recently, the determination of housing prices has become one of the most important subtopics of the sector. This topic has prompted many market players, from residential investors to real estate investment trusts and from individual investors to government officials, to predict the movement of housing prices, with these players adopting various methods to achieve this [3].

Since the Industrial Revolution, increasing urbanization throughout the world has led to a significant rise in the number of people living in cities, which has resulted in a scarcity of housing. This issue has evolved to involve many different factors over time. In [4], the writers explored the housing issues prevalent in Germany's major cities during the 1980s, when Germany was still a developing country. Throughout this time period, people recognized that there was a need for more social housing and that there was an imbalance in the housing supply due to wealth disparities. These traits have also been mentioned as difficulties that are prevalent in emerging nations that have recently urbanized as a result of population increases. For instance, a significant amount of research has been conducted on the issue of inadequate housing in India. Abhay [5], while presenting the number of houses built in a recent decade (2001–2011) in Delhi, noted that, although the housing supply is stimulated by the housing shortage problem, low-quality housing is prevalent throughout the city. This was mentioned while Abhay discussed the number of houses built in the studied decade in Delhi. According to Kumar [6], a primary factor that contributes to the lack of available housing is a disproportionately large number of outdated homes.

Understanding the trajectory of house prices is beneficial not just for current and prospective homeowners but also for real estate agents, appraisers, assessors, mortgage lenders, brokers, property developers, investors, fund managers, and lawmakers. The intricate relationship among a property's geographical location, surrounding environment, and physical attributes makes it exceedingly difficult to accurately forecast home prices [7].

In Saudi Arabia, the issue in increased housing prices has gradually become the focus of the government, consumers, investors, and academic researchers. House price planning plays a critical role in urban development. Most real estate transactions are based on field investigations and qualitative analyses, which are limited by the factors of both parties. There is no evaluation standard that can reasonably price a house, which leads to unfair transactions to a certain extent [8].

In this paper, we propose a novel house price valuation model based on data obtained from Shenzhen, China. Previously, many models have been used to estimate house prices. For example, Rosen [9] proposed the hedonic model in 1972, which measures property prices using numerous environmental factors. In the previously proposed models, housing prices were considered a function of three main components—location, structural, and neighborhood traits—and approximately linearly related to these in an exponential manner [10]. The worldwide price of oil and crude production fell as a consequence of COVID-19, which reduced Saudi Arabia's ability to generate income from oil exports. Spending on healthcare emergencies and government aid programs further strained an already tight budget. However, real estate in Saudi Arabia started indicating an overall recovery from the decline brought about by COVID-19 in May of that 2019 [11]. The number of new COVID-19 cases has now stabilized at a manageable level, the international environment is showing signs of improvement, and the country's vaccination program is gaining traction, putting Saudi Arabia on course to regain its former stature after a severe slowdown in economic activity in 2020. The success of real estate firms throughout Saudi Arabia is directly attributable to the reduction in property taxes, the filling of manpower shortages, and the establishment of cheap housing plans as part of several stimulus packages.

Traditional methods for resolving this problem have only received a tiny amount of attention in the present body of research, and this is especially true for the housing price projection model. This is despite the fact that this is a significant problem that has to be addressed. The above paragraph mentioned a few of the papers that were discovered.

The vast majority of studies that came before this one tackled the problem of the housing market as classification difficulties, with the intention of developing a classification model as opposed to a prediction model. Therefore, the study fills the available research gap by utilizing artificial neural networks in order to estimate the values of the housing market while also taking into account competitive regression models. Including the predicted goal price binning variable as a feature in the model is one of the ways that we propose to make the ANN-based approach more accurate. This model's prediction errors were much lower when compared to both contemporary machine learning models and traditional models. No previous research has, to the best of our knowledge, taken use of the studied datasets that are accessible through the official application (Aqar) in Saudi Arabia addressing the difficulties of home price forecasting, but we plan to change that.

Therefore, developing prediction models has gained importance in Saudi Arabia, and the government has prioritized the monitoring of the real estate market and property prices in the kingdom. This is with the goal of ensuring the continued financial health of the country over the long term. Because the housing market can have both positive and negative effects in established and emerging countries, it is now one of the most important forces propelling the Saudi Arabian economy. The appropriate distribution of resources is becoming an increasingly important factor in the maintenance of economic stability in Saudi Arabia. What occurs in the housing market might have a substantial influence on how resources are distributed in the economy of Saudi Arabia over the long term, which could have reverberating consequences for the overall economic growth of the nation. The main contributions of the present paper are described below. The aim of the study was to assess the accuracy of real estate price forecasts created by an artificial intelligence model using actual house price data. This evaluation focuses particularly on the accuracy of the proposed prediction model. As a result of rising house prices, several parties in Saudi Arabia, including the government, consumers, investors, and academic researchers, have started paying greater attention to the national housing market. The planning of home prices is an essential component of urban development. Therefore, the proposed model can help investors make decisions about buying houses.

The remaining sections of this work are organized as follows: after the introduction, the background of the study is discussed in Section 2, and the materials and methods are presented in Section 3. Section 4 presents the analytical results of the suggested method for predicting house prices in Saudi Arabia. A discussion of the findings of predicting systems is presented in Section 5. Section 6 presents the research's conclusion.

2. Background of the Study

The use of machine learning (ML), which is a subfield of artificial intelligence (AI), to accurately forecast market prices using past data has proven successful. Previous research has shown that model-based forecasting models provide numerous advantages over conventional forecasting models. For example, these models not only exhibit improved accuracy and precision but also provide results that are almost equivalent to those of the actual world. These advantages set model-based forecasting models apart from their contemporaries, which do not use models as a benchmark for comparison.

Furthermore, considering the spatial heterogeneity of various influencing factors, geographically weighted regression (GWR) methods have been proposed, which allow the coefficients to be changed for different locations [12,13]. According to Tobler's First Law of Geography [9], these techniques may be interpreted as a locally weighted linear regression for each specific region. The coefficients in this model take into consideration the effect of data points located in close proximity to one another. Brunsdon and Fotheringham [14,15] addressed five significant obstacles that GWR must solve to develop a model that is more fulfilling for spatially varying regression relationships. The selection of variables and bandwidth and the spatial autocorrelation of errors are some of the problems that must be overcome. Academics have made a number of attempts to solve these issues. In their study on the association between geographical variation and the use of public transportation in Shenzhen, Tu et al. [16] used GWR. Geng et al. [17] developed a model of the property market in Shenzhen, China, using GWR in 2011. The ordinary least squares (OLS) model was significantly outperformed by the GWR model, improving the R^2 from 0.56 to 0.79, which is a considerable improvement. When modeling rent in Nanjing, Zhang et al. [18] used a mixed GWR technique, which meant that certain factors were weighted according to the locality, while other variables were weighted internationally. This led to good results. Lu et al. [19,20] were able to create a model with improved results on the spatial proximity measurement of London and better estimation performance for house prices by incorporating non-Euclidean distance into GWR. These are two examples of geographic elements that do not obey the standard linear measurement.

Banerjee et al. [21] summarized various ML algorithms to predict trends in urban housing prices. Vineeth et al. [22] utilized ML algorithms to investigate the factors that influence property valuations. Phan et al. [23] used an ML system to provide housing price forecasts for Melbourne, Australia. The researchers based their conclusions on an analysis of previous purchase prices. Nevertheless, when investigating house price forecasting models, it is important to consider a number of varying elements in addition to the model selection. For example, Brueckner et al. [24] conducted research in 1987 to determine how concerns associated with urban planning were classified as macro-policies and how this classification influenced real estate prices. Evans et al. [25] conducted research to investigate the ways in which environmental factors, such as the proximity of a neighborhood's schools and the state of its public infrastructure, may influence property prices. The majority of Malpezzi's [26] research focused on the ways in which successful metropolitan corporate economies influenced property values. Diaz et al. [27] conducted research on real estate prices to evaluate the impact of the proposed rail transit system. Wenjie et al. [28] looked into and studied the accessibility of four characteristics: locations of work, social and recreational possibilities, sites of gathering, and natural areas that provide protection.

Pirogova and colleagues [29] explored the elements that facilitate the coordinated expansion of the market under the effect of digitization. A natural experiment was utilized by Dumeignil et al. [30] in their study to evaluate the impact that international labor mobility has on domestic pricing trends. Chernyshova et al. [31] studied the relationship between supply and demand in order to develop predictions about real estate prices. Real estate pricing is generated under the influence of social, economic, and material factors. Rakhman et al. [32] constructed theoretical accommodations for the particular situation of the real estate market in the Kharkov area, which included an assessment of the structural shifts and price fluctuations in the market. These theoretical accommodations were developed for the Kharkov area. Lee et al. [33] constructed their real estate index prediction model by using three distinct machine learning models, compared their findings, and discovered that the model based on the random forest offered the most accurate forecasts. Using concepts from common law, Nwogugu examined the constitutional law, competition law, and economic psychology that drive the real estate market mechanism [34]. Kang et al. [35] constructed a short-term prediction model of apartment prices utilizing machine learning technologies, and then basing it on the frequency with which keywords were searched for in news items. The model was tried and tested using data that was provided by the researchers. The capacity of the model to produce correct predictions was found to be consistent with expectations through a series of trials. Baillif et al. [36] looked at the situation in western Switzerland and proposed a discrete mixed market characteristic pricing model in order to evaluate and forecast real estate values. Saeed investigated the impact that the density of green space had on property values in the city of Ramadi, using metrics such as the proportion of urban green space, the area of green space, the distance from green space, the amount of time it takes to enter the park and public green space, and the proportion of green space in the city [37]. Kang et al. [38] constructed a price prediction model for future auctions utilizing data from real estate auctions in Seoul by combining a regression model, an artificial neural network, and a genetic algorithm. This model was based on the findings of the auctions. The results of the experiments indicated that the genetic algorithm-based real estate auction price prediction model can be broken down into multiple sub-models using the effective area of the auction appraisal price as the dividing criterion. This results in an even greater increase in the precision of the forecast. From the vantage point of the global economy, Jaymin [39] investigated the effects of the coronavirus epidemic on India's real estate market, as well as the risks and opportunities posed by those involved in the real estate market. She also looked at the risks and opportunities posed by those involved in the real estate market. Luo investigated whether or not there was a link between the price of housing and the amount of money available for higher education [40].

3. Materials and Methods

This study's objective was to develop an artificial neural network (ANN) for estimating housing prices in Saudi Arabia.

3.1. Modeling

The model was developed to gather experimental data that could be used to train the ANN prediction model. Following this, an ANN prediction model was constructed to forecast housing prices in Saudi Arabia. Lastly, the outcomes of the ANN model's predictions were broken down, examined, and assessed using accuracy measures. Figure 1 displays a block diagram of the AI algorithm used to predict housing prices in Saudi Arabia.

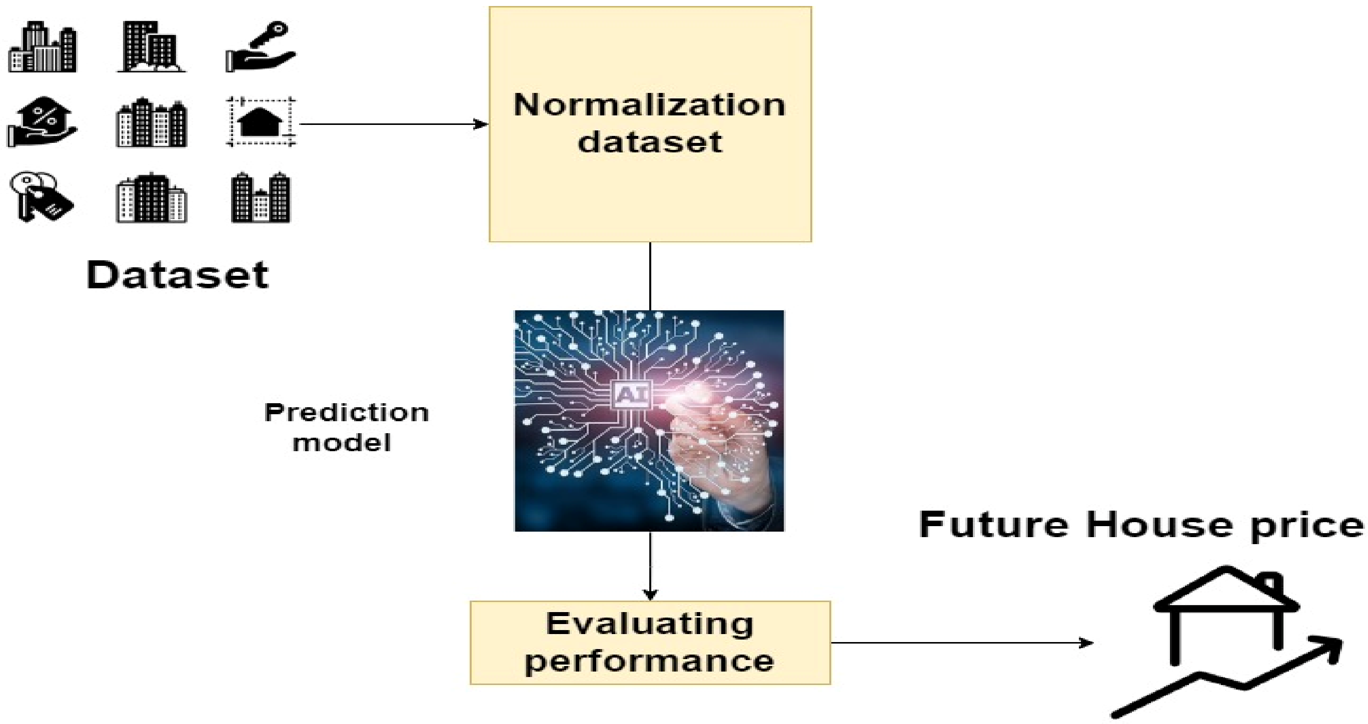

Figure 1. Framework of the ANN.

The objective of predicting housing prices in Saudi Arabia was selected in this study because these prices have been increasing at a fast pace in the last 5 years, spurred by a government plan to boost home ownership by building affordable housing. The authors planned to use 70% of the dataset as the training dataset and 30% as the testing dataset to validate the proposed framework.

The system was developed using Matlab programming with a hardware environment of 8 GB memory and processor Intel (R) Core (TM) i7–4770, operating at 3.20 GHz.

3.1.1. Dataset

The Kaggle dataset includes home pricing information for numerous Saudi Arabian cities, namely Riyadh, Jeddah, Dammam, and Al-Khobar. To help prepare for the future of the housing market in Saudi Arabia, this statistical study attempted to predict property values in the country's major cities. Riyadh, the capital and largest city in Saudi Arabia, has one of the largest metropolitan populations in the Middle East. Thus, it is now home to a diverse range of people, and its population is rapidly expanding. Located in the middle of the eastern coast of the Red Sea, Jeddah is the economic and touristic heart of Saudi Arabia. Dammam is part of a large urban and industrial complex to the northwest of Bahrain Island in the Persian Gulf, together with Khobar, Qatif, and Dhahran. Al-Khobar, together with Dammam and Dhahran, is one of the three largest cities in the Eastern Province. The size of the houses is given in a 3×3 m^2 grid, and their age ranges from 1 to 36 years. The data were collected between 25 and 28 January 2021 at 3 day intervals using the official application Aqar. Table 1 shows the size of the dataset. The features of the data were city, district front, size, property age (1–36 years), bedrooms (1–7), bathrooms (1–5), living rooms (1–5), kitchen (1), garage (1), driver room (1), maid room, furnished roof, pool, front yard, basement, duplex, stairs (1–5), elevator, fireplace, and price.

Table 1. Size of dataset.

City	Data Size (Vector)
Riyadh,	959
Jeddah,	430
Dammam,	656
Al-Khobar	1673

3.1.2. Normalization Method

In the field of statistics, 'normalization' refers to the process of rescaling the data such that they fall inside the interval [0,1]. When data normalization is required, it is often recalculated such that the mean is equal to 0 and the standard deviation is equal to 1 (unit variance). Min–max normalization is often used to scale data between predetermined ranges by applying a linear modification to the original data:

$$z_n = \frac{x - x_{min}}{x_{max} - x_{min}}(New_{max_x} - New_{min_x}), \tag{1}$$

where x_{min} and x_{max} are the minimum and maximum values for scaling the data; 1 is the maximum and 0 is the minimum. The notation New_{min_x} is used to designate the lowest value [0], while the notation New_{max_x} is used to designate the maximum value [1].

3.1.3. ANN Model

Since their introduction in the 1940s, ANNs have been recognized as a class of neural network models [41]. An ANN is a system for parallel information processing that comprises a network of hidden layers of neurons [42–45]. It is a two-layer neuronal framework comprising an input section (where data are fed into the primary predictive model), a hidden layer (where data features are extracted to construct a predictive model), and an output layer. Because it is an adaptive basis function model, the neural network model may take a given basis function and turn it into values for the model parameters when presented with the basis function. This model is referred to as a multilayer perceptron because it is structured in the form of a succession of perceptron layers. Each layer's basis function undergoes an adaptive adjustment when the previous layer's adjustment is applied to the subsequent layer. In the current experiment, the ANN was configured such that the output layer was formed of numerous different output neurons, each of which represents a distinct class and demonstrates the likelihood of belonging to that class. In addition, the error that may have been caused by the ever-changing advertising data search keywords was added to the hidden layer so that it could be corrected. However, the algorithm needs to be made more flexible to account for the fact that the error is caused by the continuous and ongoing changes in relevant search words in real time. The well-known Levenberg–Marquardt backpropagation learning method is an ANN architecture used for feature extraction from the model's inputs by employing the hidden layer. This is to maximize the accuracy of feature extraction. The developed ANN model is presented in Figure 2.

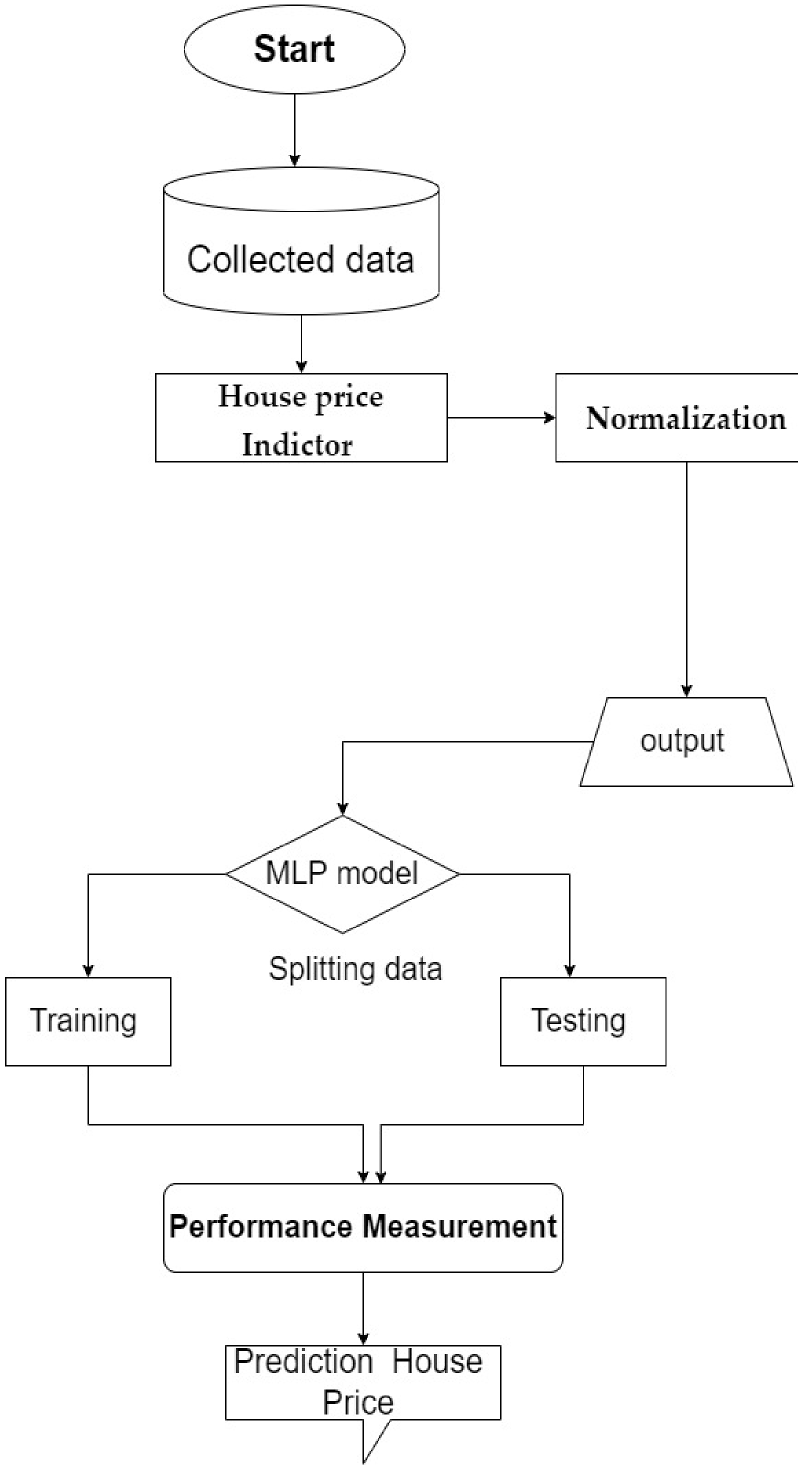

Figure 2. Block diagram of the ANN model for predicting housing prices.

Figure 3 presents a condensed version of the ANN modeling framework, which was previously described. During training, the neurons in each of the hidden layers were connected to those in the next layer using appropriate weights. The sigmoid and linear activation functions, two equations that are frequently adopted in feature extraction and modeling, are investigated in detail here.

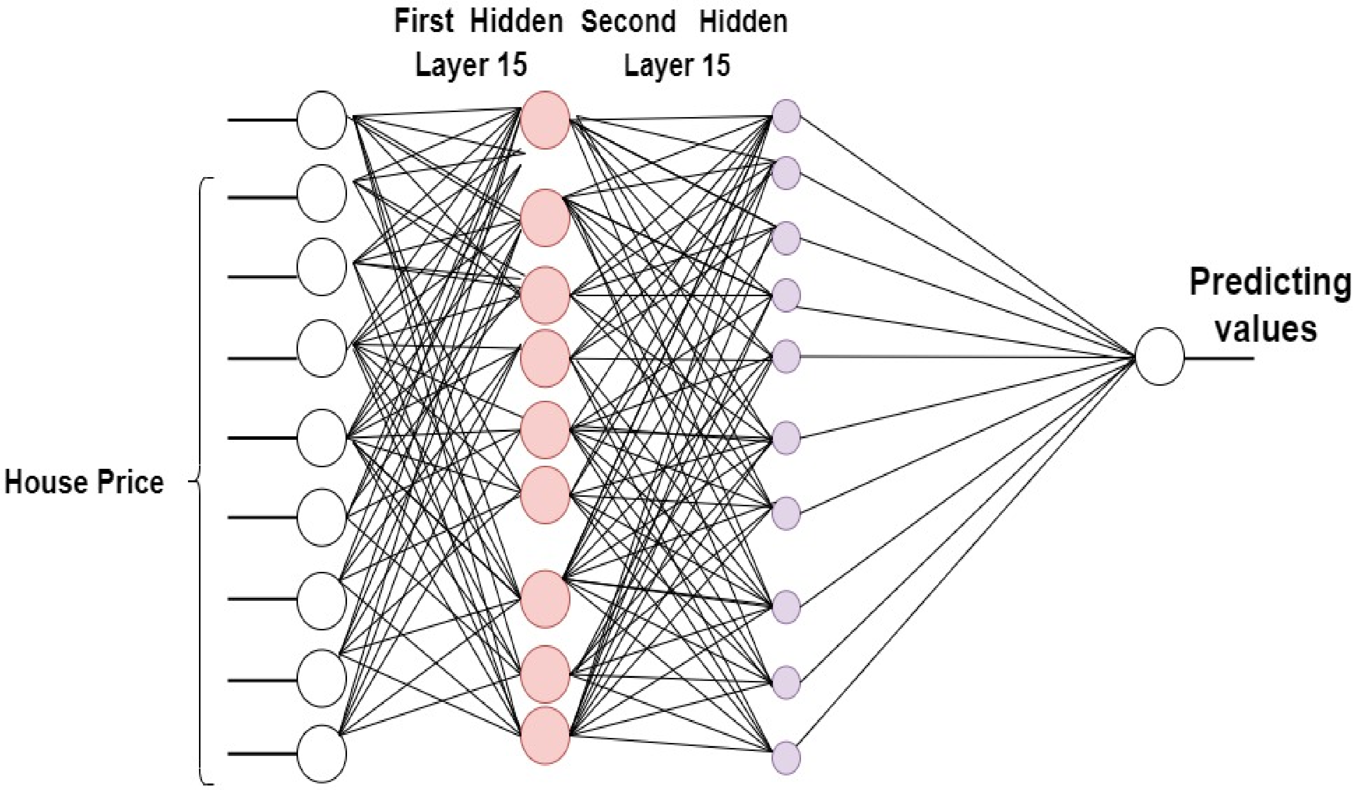

Figure 3. Structure of the ANN model.

In this study, we predict how search advertising bidding systems set their prices using an MLP-based deep learning algorithm. The ANN model parameters are detailed in Table 2. The selected value of ANN is very important for increasing the performance of the model and selecting parameters depending on the type of datasets; we found the appropriate parameter values for our dataset:

$$f(x) = b_2 + w_2 * (f_A(b_1 + w_1 * x), \tag{2}$$

where w_1 and w_2 are the weight values of the neural network (NN), b_1 and b_2 represent the bias of the NN, and f_A is the activation function.

Table 2. Configuration of the ANN model.

Parameter Names	Values of Parameters
Epoch	200
Time	0.00019
Gradient	9.45×10^{-8}
First layer	15
Second layer	15
Delay	1:20

3.1.4. Evaluation Model

The accuracy of the models' estimates of price trends and movement directions was evaluated using various measures from different domains. Error criteria such as mean squared error (MSE), root-mean-squared error (RMSE), normalized root-mean-squared error (NRMSE), and Pearson correlation coefficient error were applied to evaluate the accuracy of the predictions made in the experiments and how well the models performed.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_{i,actual} - y_{i,pred}\right)^2, \tag{3}$$

$$RMSE = \sqrt{\sum_{i=1}^{n}\frac{\left(y_{i,actual} - y_{i,pred}\right)^2}{n}}, \tag{4}$$

$$R\% = \frac{n\left(\sum_{i=1}^{n} y_{i,actual} \times y_{i,\ pred}\right) - \left(\sum_{i=1}^{n} y_{i,actual}\right)\left(\sum_{i=1}^{n} y_{i,\ pred}\right)}{\sqrt{\left[n\left(\sum_{i=1}^{n} y_{i,actual}\right)^2 - \left(\sum_{i=1}^{n} y_{i,actual}\right)^2\right]\left[n\left(\sum_{i=1}^{n} y_{i,pred}\right)^2 - \left(\sum_{i=1}^{n} y_{i,pred}\right)^2\right]}} \times 100, \tag{5}$$

$$\text{NRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{k=1}^{n}\left(y_{i,actual} - y_{i,pred}\right)^2}}{y_{i,pred}}, \tag{6}$$

where $y_{i,actual}$ denotes the observation data of house prices, and $y_{i,pred}$ denotes the predicted values obtained from the ANN model.

4. Results

In this section, the results at the training and testing phase of the ANN model for predicting house prices in the main cities of Saudi Arabia are presented.

4.1. Training Results

Table 3 presents the results of the ANN model for predicting housing prices in the main Saudi Arabian cities of Riyadh, Jeddah, Dammam, and Al-Khobar during the training phase. Seventy percent of the dataset was used as the training dataset for fitting the proposed system before validating the ANN model. The four indicators considered to examine the ANN model in the training phase were MSE, RMSE, NRMSE, and R. The results showed a scoreless prediction error for Jeddah's data (MSE = 0.001318) and the highest correction for Dammam's data (R = 100%).

Table 3. ANN model results in the training phase.

City	MSE	RMSE	NRMSE	R%
Riyadh	0.002826	0.05316	0.33261	81.40
Jeddah	0.001622	0.04028	0.30063	78.26
Dammam	0.001318	0.03630	0.1397	100
Al-Khobar	0.00330	0.01817	0.06994	99.40

4.2. Validation Results

This testing was used to validate the ANN model on an independent dataset to predict the housing prices for the main Saudi Arabian cities with a time interval of 3 days. Table 4 displays the predictions made by the ANN model according to the testing process. The values of the evaluation metrics MSE, RMSE, NRMSE, and R showed that the predicted values of the housing prices were very close to those determined experimentally.

Table 4. ANN model results in the testing phase.

City	MSE	RMSE	NRMSE	R%
Riyadh	0.002826	0.05316	0.3326	83.96
Jeddah	0.003103	0.05570	0.3485	96.49
Dammam	1.1358×10^{-6}	0.001065	0.003210	100
Al-Khobar	6.5545×10^{-6}	0.002560	0.007712	94.98

Figure 4 shows the error histogram for the values predicted while the system was in the training stage. Using the error histogram metrics, we evaluated the difference between the predicted and actual values. These error values might be negative because they represent the degree to which the predicted values differ from the actual values. The histogram errors were obtained as 0.00516, 3.505×10^{-5}, 0.01245, and 0.01661 with 20 bins for Riyadh, Dammam, Jeddah, and Al-Khobar, respectively.

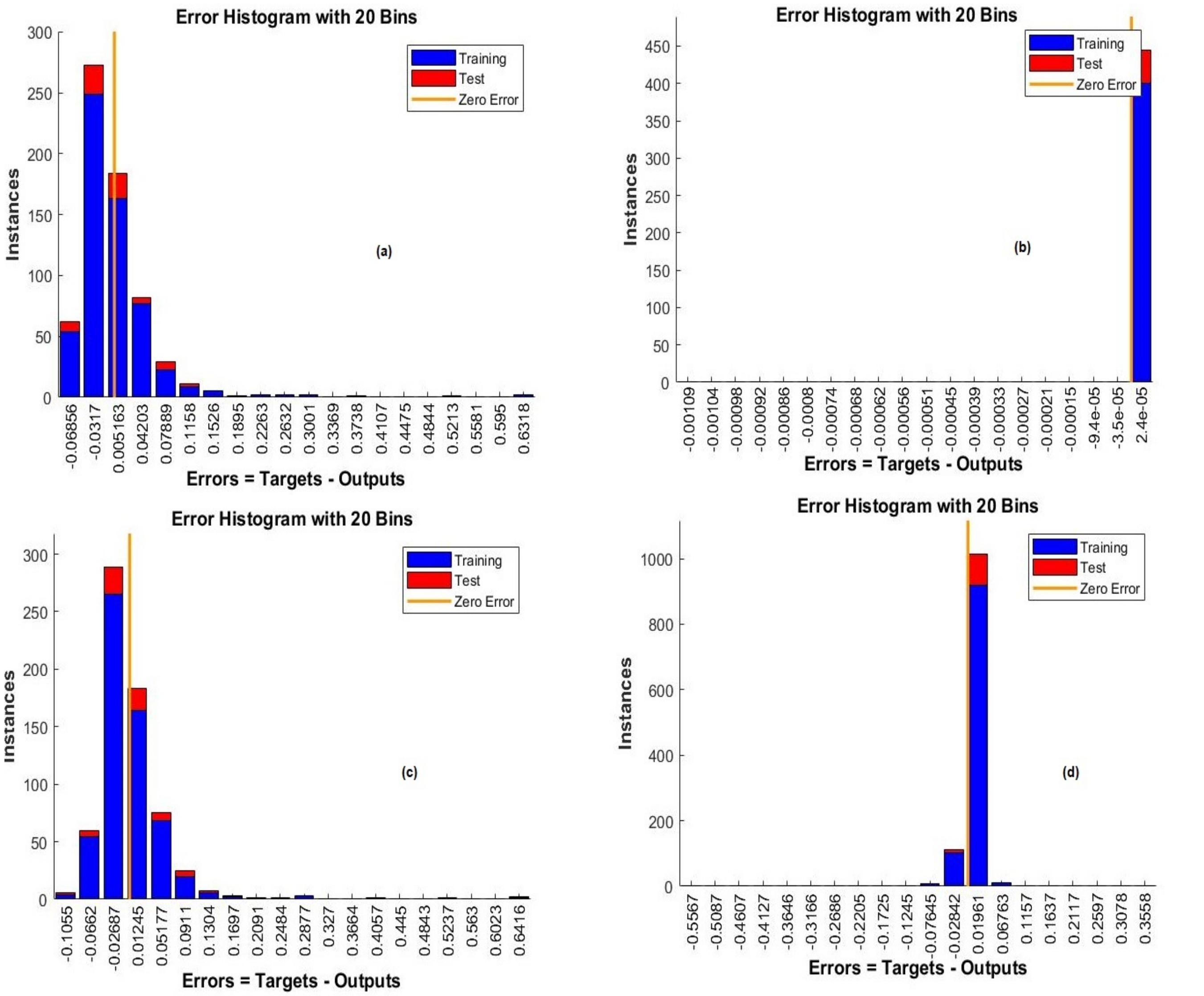

Figure 4. Histogram error of ANN model: (**a**) Riyadh, (**b**) Dammam, (**c**) Jeddah, and (**d**) Al-Khobar.

Figure 5 presents a performance plot during the training and test processes. These plots were created while attempting to predict the experimental values. Testing was terminated when there was a steady rise in the validation error over a period of 173 to 200 epochs. At this instance, MSE was measured and found to have ideal values of 0.00458, 3.2978×10^{-9}, 0.003580, and 0.000397 for Riyadh, Dammam, Jeddah, and Al-Khobar, respectively. This demonstrates the accuracy of the ANN model for forecasting housing prices.

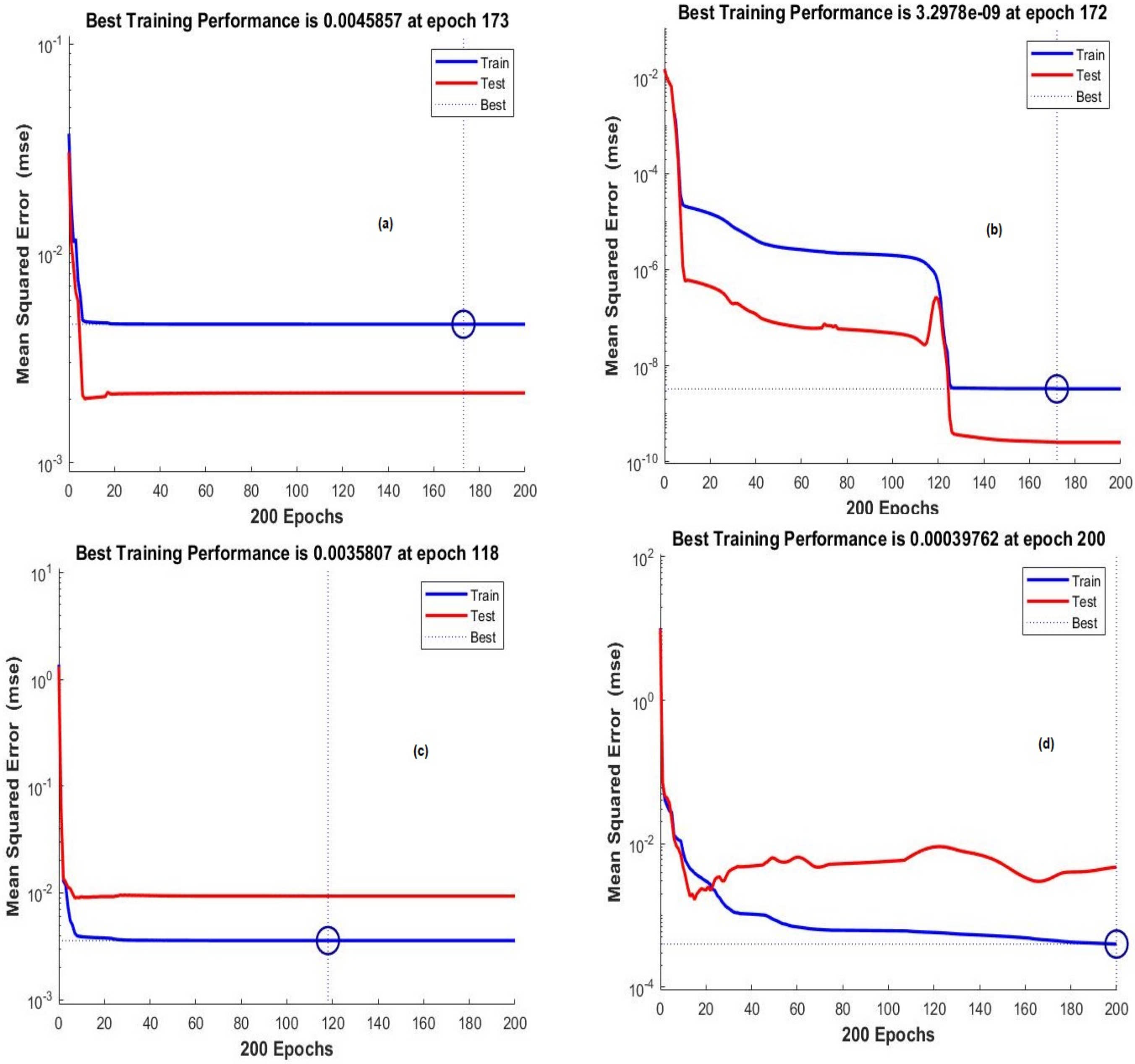

Figure 5. Performance of ANN for predicting housing prices: (**a**) Riyadh, (**b**) Dammam, (**c**) Jeddah, and (**d**) Al-Khobar.

The ANN model was used to forecast Saudi house prices over 60 days, from 29 January 2012 to 30 March 2021. Figure 6 shows the predicted final values of future house prices in the four Saudi cities over the next 60 days.

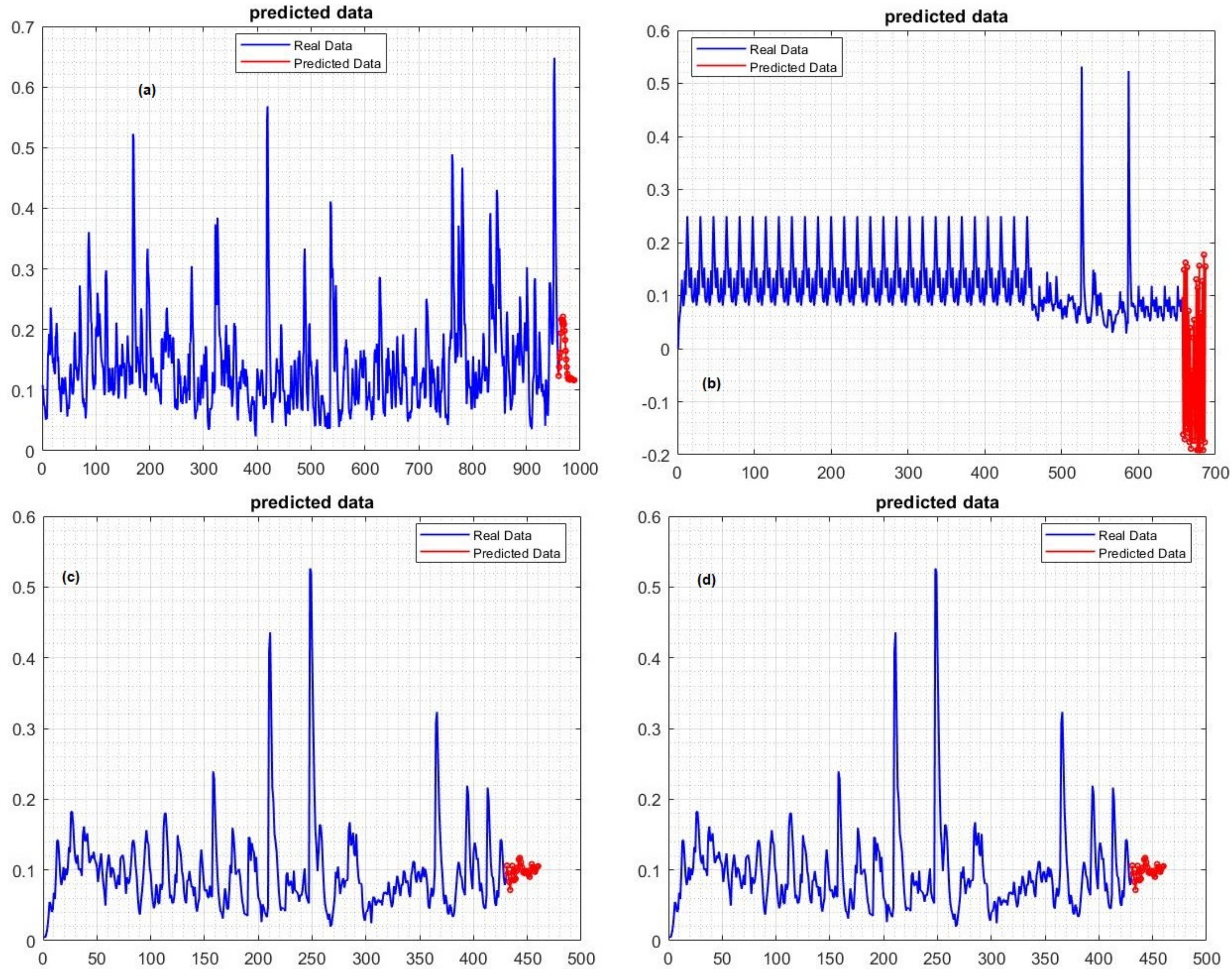

Figure 6. Values forecasted using the ANN model: (**a**) Riyadh, (**b**) Dammam, (**c**) Jeddah, and (**d**) Al-Khobar.

As shown in Figure 7, throughout the training phase, the predicted values (shown on the Y-axis) for house prices and the values of R(%) exactly matched the actual values (shown on the X-axis) in the housing price marketing dataset. The fact that the predicted values were similar to the actual values validates this assertion. Dammam and Al-Khobar had a high R ((%) = 100 and 99.40, respectively), along with exceptionally low MSE and RMSE values. These data demonstrate that the system was able to achieve its objectives.

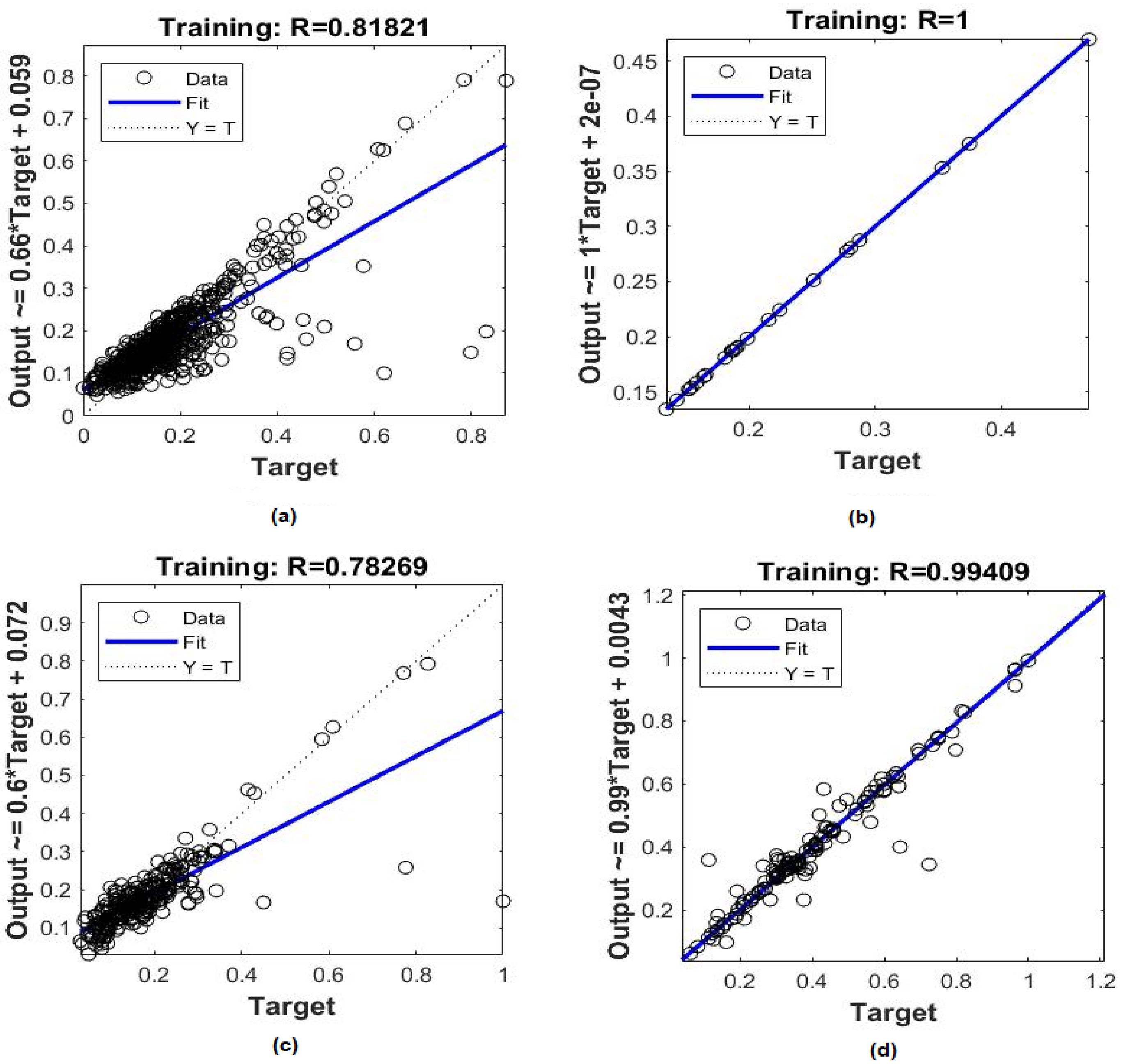

Figure 7. Regression plot of ANN model for predicting house price at the training phase: (**a**) Riyadh, (**b**) Dammam, (**c**) Jeddah, and (**d**) Al-Khobar.

In addition, Figure 8 shows a regression curve depicting the predicted values of housing prices in the testing phase. These graphs use Pearson's correlation to analyze whether the predicted values accurately reflect the actual data. The values indicated on the X-axis represent the actual data, while those indicated on the Y-axis represent the predictions computed by the ANN model. The performance of the ANN model showed a high correlation for predicting the house prices of Dammam and Al-Khobar (R (%) = 100 and R (%) = 96.49, respectively).

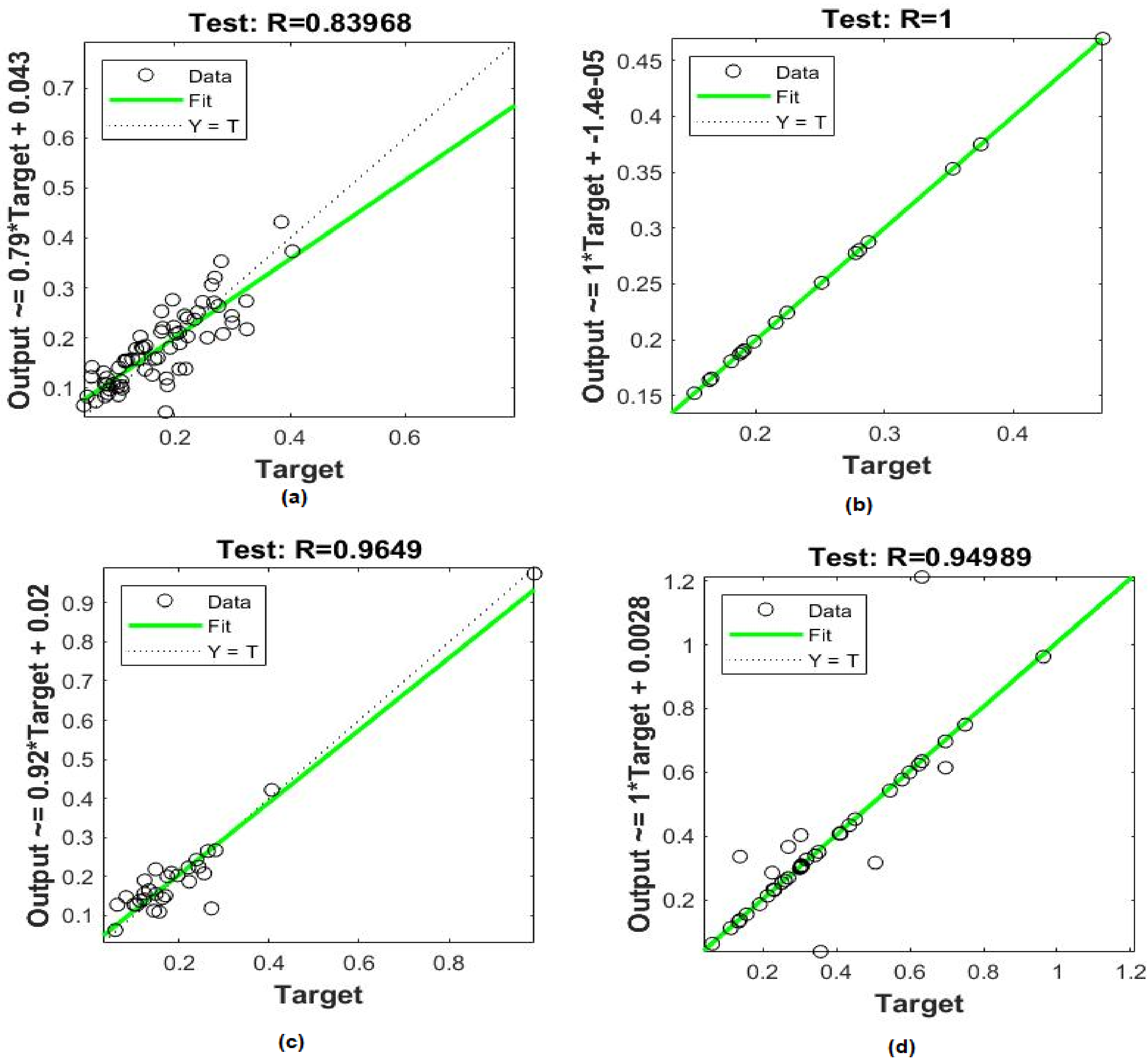

Figure 8. Regression plot of the ANN model for predicting house prices at the testing phase: (**a**) Riyadh, (**b**) Dammam, (**c**) Jeddah, and (**d**) Al-Khobar.

5. Discussion

The real estate business, which involves buying and selling properties, is currently a leading form of business. An effective housing price prediction model plays a crucial role in this business. Such a model would help investors obtain effective house pricing directly from estate owners without involving third-party agents. For this purpose, an ANN model is proposed.

There is no such thing as a healthy national economy that does not include a robust real estate sector. Therefore, both buyers and sellers of real estate, as well as economists, can gain much from monitoring the market and accurately forecasting real estate prices. However, due to the various direct and indirect factors that influence the accuracy of real estate estimates, real estate forecasting is a complex, difficult challenge. Considerable reliance is placed on the information supplied by the housing market in Saudi Arabia for establishing real estate policies and acquiring professional real estate expertise. Thus, the development of accurate models for predicting real estate prices is essential for a broad variety of essential economic services, such as banking, insurance, and urban planning in the main cities of Saudi Arabia.

The provision of housing market data is essential for formulating real estate policies and developing expertise in Saudi Arabia. In this paper, an ANN model was proposed for automatically predicting house prices using innumerable factors, such as area of the property, house site, materials used during construction, property age, number of bedrooms, and garage area. Table 5 shows the comparison results of the ANN model against existing prediction systems. It can be observed that the proposed model achieved high performance.

Table 5. Prediction results of the ANN model against existing systems.

Source:	Model	Region	RMSE
Ref. [46]	ANN	South Korea	102.03
	Random forest	South Korea	80.878
Ref. [47]	Random forest	Chain	0.12980
Ref. [48]	Support vector machine regression	India	0.0912
Proposed model	ANN	Saudi Arabia	0.001065 Testing phase

According to the findings of this study, even in Saudi Arabia, with its robust institutional structure and transparent real estate markets, residential property values are highly sensitive to all aspects of sustainability. This is the case despite Saudi Arabia's position as one of the world's leading economies and a world leader in achieving the Sustainable Development Goals (SDGs). The purpose of this study was to predict the future price of houses with the help of AI. Understanding the fluctuation of prices in real estate will allow us to develop the necessary tools and strategies to predict potential price changes. Understating when and how the trend of prices will change will assist in providing fair prices for houses. This is in line with the social dimension of the SDGs goals, which aims to provide people with reasonable price predictions.

Real estate investments in these Saudi Arabian regions typically appear to be reliable because, over time, property values tend to remain relatively unchanged. Investing one's money in real estate is a smart financial decision. Thus, a credible prediction of real estate prices is an essential component of economic analysis. Because of this, it is essential to conduct research and develop indices for house prices to monitor the movement of the real estate market. House price indicators may be beneficial for mortgage lenders, real estate agents and brokers, investors, banks, landowners, lawmakers, and other businesses and organizations involved in the real estate industry. As a result, a model for predicting house prices and establishing property limits has considerable application potential. AI model design methods have been utilized to develop models that can anticipate newly collected data. This research used an AI model to make accurate price forecasts for real estate transactions in four main cities in Saudi Arabia on the basis of actual transaction data. The empirical outcomes indicate that the use of the housing price attribute in developing an ANN model improves the overall performance of forecasting house prices.

6. Conclusions

The creation of new employment opportunities is a potential contribution made by the real estate sector to the development of the economy. In this scenario, property owners and receivers are closely related. Thus, it is essential to accurately forecast the value of real estate. Home prices are an excellent predictor of the economic health of a country, and homeowners and investors are intensely interested in homes' price ranges. Developing a model to forecast house prices could be highly beneficial for establishing regulations about house usage and estimating future house prices. The ability to predict the future worth of a piece of real estate is significant for various reasons, enabling property owners and brokers to make well-informed decisions and assisting policymakers in setting suitable prices.

In this paper, we aimed to apply an AI model trained on actual data on house prices to produce reliable forecasts of property values. The statistics were collected from the most populous urban regions in Saudi Arabia and contained information on factors such as symmetrical traffic patterns, property ownership, and housing pricing. The prediction capacity of the proposed ANN algorithms proved to be excellent. The dataset was obtained from four major cities of Saudi Arabia: Riyadh, Jeddah, Dammam, and Al-Khobar. The results demonstrated that the ANN model offered the best degree of accuracy in predicting house prices. According to the Pearson correlation metrics, the ANN model achieved high prediction accuracy for house prices in Al-Khobar (R (%) = 100%) and Dammam (R (%) = 99.48), whereas the ANN results were stratified for forecasting pricing values in Riyadh (R (%) = 85.96) and Al-Khobar R ((%) = 94.98%).

This study can assist in the design of solutions for various cities by improving error values. In addition, our investigation can be further developed in various ways. The conclusions of this study can be generalized to serve other forecasting concerns, such as those connected to economic development, oil prices, and stock price indices. The limitations of this study are its ability to handle the outliers that appeared on the regression plot of predicting house prices, an ability that would help improve the prediction system.

However, in the future, improving the accuracy of predictions made by ML would require the inclusion of other data sources, such as user comments on the features of the property, price information from social media, photographs from Google Maps, and economic data.

Author Contributions: Conceptualization, E.A., A.S.A., T.H.H.A. and S.N.A.; methodology T.H.H.A. and S.N.A.; software, T.H.H.A.; validation, E.A., A.S.A. and T.H.H.A.; formal analysis, E.A., A.S.A., T.H.H.A. and S.N.A.; investigation, T.H.H.A. and S.N.A.; resources, A.S.A. and T.H.H.A.; data curation, E.A., A.S.A., T.H.H.A. and S.N.A.; writing—original draft preparation, E.A., A.S.A. and T.H.H.A.; writing—review and editing, E.A., A.S.A., T.H.H.A. and S.N.A.; visualization, A.S.A. and T.H.H.A.; supervision, E.A., A.S.A. and T.H.H.A.; project administration, E.A., A.S.A.; funding acquisition, E.A., A.S.A. and T.H.H.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research and the APC were funded by the Saudi Investment Bank Chair for Investment Awareness Studies, the Deanship of Scientific Research, the Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Al Ahsa, Saudi Arabia (Grant No. 142).

Data Availability Statement: The data presented in this study are available at https://www.kaggle.com/code/fawazalesayi/predicting-saudi-arabia-house-prices (accessed on 18 July 2022).

Acknowledgments: This work was supported by the Saudi Investment Bank Chair for Investment Awareness Studies, the Deanship of Scientific Research, the Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Al Ahsa, Saudi Arabia (Grant No. 142).

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

References

1. Knoll, J.; Groß, R.; Schwanke, A.; Rinn, B.; Schreyer, M. Applying Recommender Approaches to the Real Estate E-Commerce Market. In *International Conference on Innovations for Community Services*; Springer: Cham, Switzerland, 2018; pp. 111–126. [CrossRef]
2. Yu, Y.; Wang, C.; Zhang, L.; Gao, R.; Wang, H. Geographical Proximity Boosted Recommendation Algorithms for Real Estate. In *International Conference on Web Information Systems Engineering*; Springer: Cham, Switzerland, 2018; pp. 61–66.
3. Rehman, F.; Masood, H.; Ul-Hasan, A.; Nawaz, R.; Shafait, F. An Intelligent Context Aware Recommender System for Real Estate. In *Mediterranean Conference on Pattern Recognition and Artificial Intelligence*; Springer: Cham, Switzerland, 2019; pp. 177–191.
4. Hass-Klau, C.H.M. The Housing shortage in Germany's Major Cities. *Built Environ.* **1982**, *8*, 60–70.
5. Abhay, R.K.; Sharma, M. Housing Shortage in a mega city: A spatio-temporal analysis of NCT-Delhi, 2001–2011. *GeoJournal* **2022**. [CrossRef]
6. Kumar, A. Estimating Rural Housing shortage. *Econ. Political Wkly.* **2014**, *49*, 74–79.

7. Durganjali, P.; Pujitha, M.V. House Resale Price Prediction Using Classification Algorithms. In Proceedings of the 6th IEEE International Conference on Smart Structures and Systems, ICSSS 2019, Chennai, India, 14–15 March 2019; pp. 1–4.
8. Qian, D.; Nana, S.; Wei, L. Real estate price prediction based on web search data. *Stat. Res.* **2014**, *31*, 81–88.
9. Rosen, S. Hedonic prices and implicit markets: Product differentiation in pure competition. *J. Political Econ.* **1974**, *82*, 34–55. [CrossRef]
10. Butler, R.V. The specification of hedonic indexes for urban housing. *Land Econ.* **1982**, *58*, 96–108. [CrossRef]
11. Arabicbusiness. Available online: https://www.arabianbusiness.com/money/wealth/money-wealth-real-estate/property-prices-in-saudi-arabia-on-the-rise-riyadh-leads-the-way (accessed on 10 September 2022).
12. Fotheringham, A.S.; Brunsdon, C.; Charlton, M. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*; John Wiley & Sons: Hoboken, NJ, USA, 2003.
13. Brunsdon, C.; Fotheringham, A.S.; Charlton, M.E. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geogr. Anal.* **1996**, *28*, 281–298. [CrossRef]
14. Tobler, W.R. A computer movie simulating urban growth in the Detroit region. *Econ. Geogr.* **1970**, *46*, 234–240. [CrossRef]
15. Brunsdon, C.; Fotheringham, A.S.; Charlton, M. Some notes on parametric significance tests for geographically weighted regression. *J. Reg. Sci.* **1999**, *39*, 497–524. [CrossRef]
16. Tu, W.; Cao, R.; Yue, Y.; Zhou, B.; Li, Q.; Li, Q. Spatial variations in urban public ridership derived from GPS trajectories and smart card data. *J. Transp. Geogr.* **2018**, *69*, 45–57. [CrossRef]
17. Geng, J.; Cao, K.; Yu, L.; Tang, Y. Geographically Weighted Regression model (GWR) based spatial analysis of house price in Shenzhen. In Proceedings of the 2011 19th International Conference on Geoinformatics, Shanghai, China, 24–26 June 2011; pp. 1–5. [CrossRef]
18. Zhang, S.; Wang, L.; Lu, F. Exploring housing rent by mixed geographically weighted regression: A Case study in Nanjing. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 431. [CrossRef]
19. Lu, B.; Charlton, M.; Harris, P.; Fotheringham, A.S. Geographically weighted regression with a non-Euclidean distance metric: A case study using hedonic house price data. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 660–681. [CrossRef]
20. Lu, B.; Charlton, M.; Fotheringhama, A.S. Geographically weighted regression using a non-Euclidean distance metric with a study on London house price data. *Procedia Environ. Sci.* **2011**, *7*, 92–97. [CrossRef]
21. Banerjee, D.; Dutta, S. Predicting the housing price direction using machine learning techniques. In Proceedings of the 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering, Chennai, India, 21–22 September 2017. [CrossRef]
22. Vineeth, N.; Ayyappa, M.; Bharathi, B. House price prediction using machine learning algorithms. In *International Conference on Soft Computing Systems*; Springer: Singapore, 2018; pp. 423–433. [CrossRef]
23. Phan, T.D. Housing price prediction using machine learning algorithms: The case of Melbourne city, Australia. In Proceedings of the International Conference on Machine Learning and Data Engineering, Sydney, Australia, 3–7 December 2018; pp. 8–13. [CrossRef]
24. Brueckner, J.K. Chapter 20 The structure of urban equilibria: A unified treatment of the muth-mills model. *Handb. Reg. Urban Econ.* **1987**, *2*, 821–845. [CrossRef]
25. Evans, R.D.; Rayburn, W. The Effect of school desegregation decisions on single-family housing prices. *J. Real Estate Res.* **1991**, *6*, 107–216.
26. Malpezzi, S. Hedonic Pricing Models: A Selective and Applied Review. In *Housing Economics and Public Policy*; Wiley: Hoboken, NJ, USA, 2002; Volume 10, pp. 67–89. [CrossRef]
27. Diaz, R.B. *Impacts of Rail Transit on Property Values*; Booz Allen & Hamilton Inc.: McLean, VA, USA, 1999.
28. Wenjie, W.; Zhilin, L.; Wenzhong, Z. Evaluation of factors influencing residential land price in Beijing based on structural equation model. *Acta Geogr. Sin.* **2011**, *65*, 676–684.
29. Pirogova, O.; Temnova, N. Dynamics of coworking growth in the real estate market under digitalization. *E3S Web Conf.* **2021**, *244*, 10052. [CrossRef]
30. Dumeignil, C. The impact of cross-border labor mobility on real estate price trends: A natural experiment. *Int. Reg. Sci. Rev.* **2021**, *45*, 108–132. [CrossRef]
31. Chernyshova, M.; Malenkaya, A.; Mezhuyeva, T. Analysis of pricing factors in real estate market. *Interexpo GEO-Sib.* **2019**, *6*, 79–85. [CrossRef]
32. Rakhman, M.S.; Malko, V.V.; Malko, V.V. Structural changes in the real estate market in kharkiv region. *Bus. Inf.* **2019**, *10*, 143–146. [CrossRef]
33. Lee, S.J.M.; Park, H.; Cho, S.H.; Kim, J.H. Comparison of models to forecast real estates index introducing machine learning. *J. Archit. Inst. Korea Struct. Constr.* **2021**, *37*, 191–199.
34. Nwogugu, M. Some constitutional law, competition law and economic psychology issues inherent in some real estate market mechanisms. In *Geopolitical Risk, Sustainability and "Cross-Border Spillovers" in Emerging Markets*; Springer: Cham, Switzerland, 2021; Volume 1, pp. 137–206.
35. Markey-Towler, B. The competition and evolution of ideas in the public sphere: A new foundation for institutional theory. *J. Inst. Econ.* **2019**, *15*, 27–48. [CrossRef]

36. Baillif, M.; de Lapparent, M.; Kazagli, E. A hybrid approach to real estate price definition: A case study in western Switzerland. *Rev. Économique* **2021**, *72*, 1055–1077. [CrossRef]
37. Saeed, A.; Mullahwaish, L. Effect of green areas density on real estate price in Ramadi city. *Int. J. Des. Nat. Ecodyn.* **2020**, *15*, 253–259. [CrossRef]
38. Kang, J.; Lee, H.J.; Jeong, S.H.; Lee, H.S.; Oh, K.J. Developing a forecasting model for real estate auction prices using artificial intelligence. *Sustainability* **2020**, *12*, 2899. [CrossRef]
39. Jaymin, S.R. The effect of covid-19 on the real estate industry in India. *Balt. J. Real Estate Econ. Constr. Manag.* **2021**, *9*, 122–129.
40. Luo, H. Research on the interaction between higher education resource allocation and real estate price. *Open J. Soc. Sci.* **2020**, *8*, 58–68. [CrossRef]
41. McCulloch, W.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [CrossRef]
42. Piccinini, G. The First computational theory of mind and brain: A close look at McCulloch and Pitts's "A logical calculus of ideas immanent in nervous activity". *Synthese* **2004**, *141*, 175–215. [CrossRef]
43. Aldhyani, T.H.H.; Alzahrani, A. Framework for Predicting and Modeling Stock Market Prices Based on Deep Learning Algorithms. *Electronics* **2022**, *11*, 3149. [CrossRef]
44. Al-Adhaileh, M.H.; Aldhyani, T.H.H. Artificial intelligence framework for modeling and predicting crop yield to enhance food security in Saudi Arabia. *PeerJ Comput. Sci.* **2022**, *2022*, e1104. [CrossRef]
45. Ammer, M.A.; Aldhyani, T.H.H. Deep Learning Algorithm to Predict Cryptocurrency Fluctuation Prices: Increasing Investment Awareness. *Electronics* **2022**, *11*, 2349. [CrossRef]
46. Kim, J.; Lee, Y.; Lee, M.-H.; Hong, S.-Y. A Comparative Study of Machine Learning and Spatial Interpolation Methods for Predicting House Prices. *Sustainability* **2022**, *14*, 9056. [CrossRef]
47. Truong, Q.; Nguyen, M.; Dang, H.; Mei, B. Housing Price Prediction via Improved Machine Learning Techniques. *Procedia Comput. Sci.* **2020**, *174*, 433–442. [CrossRef]
48. Manasa, J.; Gupta, R.; Narahari, N.S. Machine Learning based Predicting House Prices using Regression Techniques. In Proceedings of the 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Bangalore, India, 5–7 March 2020; pp. 624–630. [CrossRef]