Available online at www.CivileJournal.org

Civil Engineering Journal

(E-ISSN: 2476-3055; ISSN: 2676-6957)

Vol. 9, Special Issue, 2023

"Innovative Strategies in Civil Engineering Grand Challenges"

Forecasting the Real Estate Housing Prices Using a Novel Deep Learning Machine Model

Hossam H. Mohamed [1], Ahmed H. Ibrahim [1], Omar A. Hagras [1*]

[1] *Construction Engineering and Utility Department, Faculty of Engineering, Zagazig University, Zagazig, Egypt.*

Received 23 November 2022; Revised 25 January 2023; Accepted 11 February 2023; Published 04 March 2023

Abstract

There is an urgent need to forecast real estate unit prices because the average price of residential real estate is always fluctuating. This paper provides a real estate price prediction model based on supervised regression deep learning with 3 hidden layers, a Relu activation function, 100 neurons, and a Root Mean Square Propagation optimizer (RMS Prop). The model was developed using actual data collected from 28 Egyptian cities between 2014 and 2022. The model can forecast the price of a real estate unit based on 27 different variables. The model is created in two stages: adjusting the parameters to obtain the best ones using a sensitivity k-fold technique, then optimizing the result. 85 percent of the real estate unit data gathered was used in training and developing the model, while the other 15 percent was used in validating and testing. By using a dropout regularization technique of 0.60 on the model layers, the final developed model had a maximum error of 10.58%. After validation, the model had a maximum error of about 9.50%. A graphical user interface (GUI) tool is developed to make use of the final predictive model, which is very simple for real estate developers and decision-makers to use.

Keywords: Residential; Real Estate; Price; Decision Makers; Deep Learning.

1. Introduction

The term "home" refers to the social unit formed by a family or by one or more unrelated individuals living together. A house, on the other hand, is a collection of attributes such as size, quality, and location. House valuation is difficult for a variety of reasons. Each house, as a physical asset, has its own unique location. Furthermore, because a house is a long-term durable good with a long life, houses of varying ages can exist in the same market at the same time. Each house has a distinct set of characteristics, which have an impact on its value. Furthermore, different geographical areas may value different housing characteristics [1].

The status of the economy, interest rates, real income, and changes in population size all have an impact on the housing market. In addition to these demand-side considerations, the available supply will determine property prices. Rising house prices, high rents, and an increased danger of homelessness will result from periods of rising demand and inadequate supply. Many studies were based on estimating the price of real estate units based solely on the cost of the real estate units, and thus these studies lacked realistic simulation due to the lack of an integrated model that includes a set of factors and is most important in determining the price of real estate units in accordance with the real estate market [2].

* Corresponding author: omarhagras10@gmail.com; OAhagras@zu.edu.eg

http://dx.doi.org/10.28991/CEJ-SP2023-09-04

Furthermore, the price of the real estate does not only comprise the price of the property itself but also the price of the land on which it is built. Real estate price forecasting should be based on elements such as the cost of the building itself and its management [2]. By reviewing and analyzing previous studies, various points in dealing with factors influencing the price of real estate units were discovered. Selim [3] analyzed a variety of aspects only connected to the building component of real estate units. Only economic aspects were explored by Borowiecki [4], Rafiei & Adeli [5], and Limsombunc et al. [6]. It is interesting to note that both Khalafallah [7] and Rahman [8] looked at factors that influence the pricing of a real estate unit without taking into consideration the effect of the population or any factors linked to buildings or service systems, but only economic-related factors.

Various models are utilized in forecasting real estate unit prices. Traditional models such as multiple regression and stepwise regression are used. Much research has been conducted on the use of hedonic regression models, such as (Selim [3]; Candas et al. [9]; Malpezzi [10]; Yusof & Ismail [11]; Pow et al. [12]; Gustafsson & Wogenius [13]; Ozgur et al. [14]) using a limited number of factors in their models.

Khalafallah [7] designed an Artificial Neural Networks (ANNs) model to get the ratio between a house's selling and asking prices every 3 months in housing market performance in Orlando, Florida. (Dai et al. [15]; Story and Fry [16]; Butcher et al. [17]) evaluate the performance and the housing price in the United States by applying Artificial Neural Networks (ANNs) using a limited number of different types of factors in their models.

Rafiei & Adeli [5] predict the real estate unit price in the United States by using novel machine learning through deep belief-restricted Boltzmann machines and a unique genetic algorithm. Patil et al. [18] forecast the real estate unit price by using the CatBoost algorithm with Robotic Process Automation for real-time data extraction using a limited number of economic factors only in their deep learning models.

According to the assessment of the literature, most prior research relied mainly on regression models, which have low accuracy. Because the great majority of these models rely on the use of a restricted number of characteristics to forecast real estate unit values. Furthermore, many earlier models rely solely on one type of factor, for example, some models rely solely on economic factors, such as Rafiei & Adeli [5], others rely solely on construction factors, such as Selim [3], and still others rely solely on the nature of the architectural division of a real estate unit, such as Limsombunchai [6].

Cheung et al. [19] conducted a revealed preference analysis using a price gradient approach to assess the impact of pandemic risk on residential real estate markets. When taking into account microlevel housing transaction data in 62 areas from nine districts in Wuhan City from January 2019 to July 2020, hedonic pricing and price gradient models predict a 4.8% and a 5.0–7.0% year-on-year fall in house prices immediately after the pandemic outbreak, respectively. Ho et al. [20] used three machine learning algorithms: SVM, random forest (RF), and gradient boosting machine (GBM). This study has demonstrated that SVM is still a valuable technique in data fitting.

Hu et al. [21] studied the effect of the COVID-19 on the daily housing return in five Australian cities using a daily hedonic housing price index. Mora-Garcia et al. [22] investigated and quantified the effect of the COVID-19 epidemic on home prices in Alicante, Spain. Because of their superior adaptability to the nonlinearities of complex data, such as real estate market data, machine learning algorithms outperform conventional linear models.

This paper contributes to filling many gaps in previous research studies by including the following points:

- Previous studies concentrated on a small number of factors, whereas the current study included 27 different types of factors to develop a deep learning model.
- Many previous studies concentrated on hedonic regression and machine learning techniques, so this study concentrated on using the AI system, specifically the deep learning technique, with the Keras Library, which is the best Python deep learning library and contains the TensorFlow backend, the Theano backend, and the Microsoft Knowledge Toolkit (CNTK) to give more accuracy to the developed model and contribute to the lowest price error and high accuracy of results not previously found in previous studies. As there has been little work on the current situation in Egypt, the paper collected local data for a wide range of time periods and cities in Egypt over an eight-year period and in 28 capital cities under various conditions.
- The development of a graphical user interface to assist real estate developers in more accurately and easily calculating the price of real estate units based on a variety of short-term economic indicators that change every three months, such as inflation rates and bank interest rates.

This study aims to develop a new and comprehensive methodology for calculating new house prices in any city in Egypt at the design stage, the commencement of construction, and existing buildings. using a wide range of factors that differ in nature and classification.

The model is expected to be used by real estate companies and developers to assess the sale market for real estate units before beginning construction on new ones. The expected model has an effective data structure and takes into consideration the measurement of economic indicators every three months, making the model flexible enough to deal with economic events more efficiently. A graphical user interface tool has also been created to help real estate developers and companies precisely forecast real estate unit prices.

2. Research Methodology

The methodology undertaken in this study is summarized in 6 steps as mentioned below:

1- Identifying factors affecting the real estate unit price from previous studies and classify them to groups according to their nature.

2- Identifying the most significant factors that affecting real estate unit price using questionnaire survey.

3- Real estate units' prices data collection and exploratory data analysis.

4- Deep learning model training to get most accurate hyper parameters.

5- Development, validation and testing of the final deep learning model.

6- Develop final model graphical user interface tool.

3. Factors Affecting Real Estate Unit Price

The main factors that affected the price of real estate units and were discussed in the literature were broken down into seven groups based on their nature and classification. These groups included building characteristics, structural and architectural, financial properties, social, governmental, physical and environmental, and economic factors. All of the factors listed in Table 1 have been applied to the Egyptian real estate market in order to identify the most relevant ones that influence the sale price of real estate units.

Table 1. Factors Affecting on Real Estate Price from the Reviewed Literature

Group ID.	Factor ID.	Factor Description	References
		Building Characteristics Factors	
GA	GA1	land size (in square meters)	Rafiei & Adeli (2015), Limsombunchai (2008), and Corsini (2019) [5, 6, 23]
	GA2	House age	Limsombunchai (2008), Khalafallah (2004), and Corsini (2019) [6, 7, 23]
	GA3	Location type (Urban or rural)	Selim (2009), and Candas et al. (2015) [3 ,9]
	GA4	Land value	Selim (2009), and Candas et al. (2015) [3, 9]
	GA5	Gross used area	Selim (2009), Candas et al. (2015) and Corsini (2019) [3, 9, 23]
	GA6	Site/building type	Selim (2009), and Candas et al. (2015) [8, 9]
	GA7	Fire System	Selim (2009), and Candas et al. (2015) [3, 9]
	GA8	Elevator	Selim (2009), and Candas et al. (2015) [3, 9]
	GA9	Distance to the road	Candas et al. (2015) [9]
	GA10	Distance to the hospital	Candas et al. (2015) [9]
	GA11	Distance to the mall	Candas et al. (2015) [9]
	GA12	Distance to the sea	Candas et al. (2015) [9]
		Structural and Architectural Factors	
GB	GB1	Number of bedrooms	Selim (2009), Limsombunchai (2008), and Corsini (2019) [3 ,6, 23]
	GB2	Number of bathrooms	Selim (2009), Limsombunchai (2008), and Corsini (2019) [3 ,6, 23]
	GB3	With or without a basement	Selim (2009), Limsombunchai (2008) [3 ,6]
	GB4	Number of garages	Selim (2009), Limsombunchai (2008), and Corsini (2019) [3 ,6, 23]
	GB5	The type of house (with garden, or without garden)	Limsombunchai (2008) [6]
	GB6	Amenities around the residential areas (such as public facilities)	Limsombunchai (2008) [6]
	GB7	Water system	Corsini (2019) [23]
	GB8	Level of Finishing	Corsini (2019) [23]

GC	**Financial Properties Factors**		
	GC1	Project locality defined in terms of zip codes	Rafiei & Adeli (2015) [5]
	GC2	Preliminary estimated construction cost per m^2 based on the prices at the beginning of the project	Borowiecki (2009) and Rafiei & Adeli (2015) [4, 5]
	GC3	Duration of construction	Rafiei & Adeli (2015) [5]
	GC4	Price of the unit at the beginning of the project per m^2	Rafiei & Adeli (2015), and Khalafallah (2004) [5, 7]
GD	**Social Factors**		
	GD1	Population Trends (Growth, Decline, Stability)	Haider & Miller (2000) [24]
	GD2	Environmental Consciousness	Haider & Miller (2000) [24]
	GD3	Family Composition	Haider & Miller (2000) [24]
	GD4	Security Consciousness	Haider & Miller (2000) [24]
	GD5	Aging of the population	Haider & Miller (2000) [24]
	GD6	Standard Level In the region	Corsini (2019), and Haider & Miller (2000) [23, 24]
	GD7	Evolution of Home Offices	Haider & Miller (2000) [24]
	GD8	The Family and Functional Utility	Haider & Miller (2000) [24]
GE	**Governmental Factors**		
	GE1	Real Estate Taxes and Assessments	Rahman (2008), and Haider & Miller (2000) [8, 24]
	GE2	Labor Issues	Rahman (2008) [8]
	GE3	Quality of Schools	Corsini (2019), and Haider & Miller (2000) [23, 24]
	GE4	Quality of Services	Haider & Miller (2000) [24]
GF	**Physical/Environmental Factors**		
	GF1	Transportation	Haider & Miller (2000) [24]
	GF2	Environmental awareness	Selim (2009), and Haider & Miller (2000) [3, 24]
	GF3	Topography	Rahman (2008) [8]
	GF4	Climate/Weather	Haider & Miller (2000) [24]
GG	**Economic Factors**		
	GG1	The bank interest rate	Borowiecki (2009), Rafiei & Adeli (2015), and Haider & Miller (2000) [4, 5, 24]
	GG2	The inflation rate	Rafiei & Adeli (2015) [5]
	GG3	The economic climate	Rafiei & Adeli (2015), and Rahman (2008) [5, 8]
	GG4	Percentage change in median house price compared to the previous year	Rafiei & Adeli (2015) [5]
	GG5	Average Months a house spends on the market	Rafiei & Adeli (2015) [5]
	GG6	The volume of inventory	Khalafallah (2004) [7]
	GG7	The inventory months' supply	Khalafallah (2004) [7]
	GG8	Gross domestic product	Borowiecki (2009), Rafiei & Adeli (2015), and Jian & Zhang (2012) [4, 5, 25]
	GG9	Private Real Estate Price Index.	Borowiecki (2009), and Rafiei & Adeli (2015) [4, 5]
	GG10	Building services index for a preselected base year	Rafiei & Adeli (2015) [5]
	GG11	Wholesale price index (WPI) of building materials for the base year	Rafiei & Adeli (2015) [5]
	GG12	Cumulative liquidity (millions of Pounds)	Rafiei & Adeli (2015) [5]
	GG13	Private sector investment in new buildings in the city	Rafiei & Adeli (2015), and Rahman (2008) [5, 8]
	GG14	Land price index in the city for the base year	Rafiei & Adeli (2015), and Rahman (2008) [5, 8]
	GG15	Number of loans extended by banks in a quarter	Rafiei & Adeli (2015) [5]
	GG16	Amount of loans extended by banks in a quarter (millions of Pounds)	Rafiei & Adeli (2015) [5]
	GG17	The interest rate for the loan in the quarter	Rafiei & Adeli (2015) [5]
	GG18	The official exchange rate for dollars	Rafiei & Adeli (2015), Ding (2014), and Cao (2003) [5, 26, 27]
	GG19	Nonofficial (street market) exchange rate for dollars (used only in countries with controlled currencies)	Rafiei & Adeli (2015) [5]
	GG20	Consumer price index (CPI) of housing, water, fuel, and power in the base year	Borowiecki (2009), and Rafiei & Adeli (2015) [4, 5]
	GG21	Stock market index	Rafiei & Adeli (2015) [5]
	GG22	Gold price per ounce (pounds)	Rafiei & Adeli (2015) [5]
	GG23	Total Investment Amount in Fixed Assets	Li (2015), Rafiei & Adeli (2015), and Rahman (2008) [2, 5, 8]

4. Questionnaire Survey

This questionnaire was designed for real estate developers, and 92 questionnaires were gathered from two groups of engineers and managers from the most powerful real estate firms. This questionnaire was distributed via interviews with real estate professionals ranging in experience from 8 to 22 years to establish the most relevant factors in the pricing of real estate units in Egypt, as shown in Table 2.

Table 2. Respondents' classification

Experience Years / Experts	Between (8-12) years	Between (12-15) years	Between (15-18) years	Between (18-22) years
Engineers	10	12	9	11
Real estate developers	8	11	14	17
Total	**18**	**23**	**23**	**28**

4.1. Determination of Required Sample Size

Bartlett et al. [28] dealt with this formula to find the sample size to be inferred for an unlimited population:

$$n = \frac{K^2 \times P\,(1-P)}{E^2} \tag{1}$$

where n is the essential sample size for finite population, K-value equals to 1.645 at confidence level equals to 90%, P is the proportion of the population, and E is the acceptable boundary of error = 10% for confidence 90%.

The P range between (0 up to 1.0). By substituting these parameters in Equation 1 to obtain the required sample size (the vital value of P is 0.5), the essential sample size for this study of finite population was calculated to be 68 as the lowest value.

4.2. The Most Influential Factors on the Price of Real Estate Units

Real estate marketing, building, and development specialists were asked to complete a questionnaire that included questions from all seven groups as well as questions about factors. The reliability of this questionnaire was tested, and the SPSS Ver.25 [29] alpha coefficient value for 63 components was 0.847, indicating that it was reliable.

The questionnaire's design was based on assigning a score from 1 to 10 to two basic measures. The first scale represents the frequency of each factor in determining the price of real estate units, while the second scale expresses the importance of this factor if it is taken into account when calculating the price of the real estate unit (Gab-Allah et al. [30]; Mohmad et al. [31]; and El Touny et al. [32]).

Next, after the calculation of the first two metrics, the following three indices are computed: The first index is known as the frequency index, as shown by Equation 2, the second index is known as the importance index, as shown by Equation 3, and the third index is known as the final index, which is equal to the product of the previous indices and is used to assess and identify the most significant factors.

$$Frequency\ index = \frac{\sum_{i=1}^{n} F_i}{a \times n} \tag{2}$$

$$Importance\ index = \frac{\sum_{i=1}^{n} M_i}{a \times n} \tag{3}$$

where $\sum_{i=1}^{n} F_i$ *and* $\sum_{i=1}^{n} M_i$ is Totality of frequency and importance scores of each factor from the respondents. The weighting ranges from 1 (lowest) to 10 (highest), n is total Number of respondents which is constant and equals to 92, and a is upper scale for each measure equals10.

To identify the factors that have the greatest influence on the price of real estate units based on the final index, which is equal to the sum of multiplying the frequency index by the importance index, The relative weight is estimated according to the final index, with a value of 100%, and the rest of the values are assigned according to the final index, with a value of 100%. Factors with a relative weight of greater than 70% are picked as the most essential (Gab-Allah et al. [30]; Mohmad et al. [31]; and El Touny et al. [32]).

Thus, the questionnaire found that 27 factors were most influential in determining the price of real estate units in Egypt, based on a relative weight of greater than 70%, as shown in Table 3.

Table 3. The Most Influential Factors

Factor No.	Suggested Factors	Group Related Factors	Frequency Index	Importance Index	Final Index	Relative Weight
1	The Interest rate	GG	0.7543	0.8239	0.6215	100.00%
2	Level of Finishing	GB	0.7641	0.7880	0.6022	96.89%
3	Number of bedrooms	GB	0.7685	0.7609	0.5847	94.08%
4	The inflation rate	GG	0.7174	0.8141	0.5841	93.97%
5	Consumer price index (CPI)	GG	0.7087	0.8163	0.5785	93.08%
6	Standard Level In the region	GD	0.7174	0.7891	0.5661	91.09%
7	Preliminary estimated construction cost per m^2	GC	0.6989	0.7891	0.5515	88.74%
8	Land value	GA	0.7163	0.7587	0.5435	87.44%
9	Price of the unit at the beginning of the project per m^2	GC	0.7304	0.7424	0.5423	87.25%
10	Average Months a house spends on the market	GG	0.7087	0.7641	0.5415	87.13%
11	Gross used area	GA	0.7076	0.7543	0.5338	85.88%
12	land size (in square meters)	GA	0.7076	0.7511	0.5315	85.51%
13	The economic climate	GG	0.6978	0.7533	0.5256	84.57%
14	Quality of Services	GE	0.7011	0.7370	0.5167	83.13%
15	House age (in months)	GA	0.7022	0.7261	0.5098	82.03%
16	Location type (Urban or rural)	GA	0.6761	0.7370	0.4982	80.17%
17	Population Trends	GD	0.6902	0.7043	0.4862	78.22%
18	Number of bathrooms	GB	0.6500	0.7033	0.4571	73.55%
19	The type of house	GA	0.6848	0.6630	0.4540	73.05%
20	Quality of Schools	GE	0.6163	0.7359	0.4535	72.97%
21	Site/building type	GA	0.6359	0.7120	0.4527	72.84%
22	Elevator	GA	0.6304	0.7130	0.4495	72.33%
23	Duration of construction	GC	0.6489	0.6826	0.4430	71.27%
24	Distance to the road	GA	0.6511	0.6739	0.4388	70.60%
25	Distance to the mall	GA	0.6141	0.7130	0.4379	70.46%
26	Transportation	GF	0.6554	0.6674	0.4374	70.38%
27	Number of garages	GB	0.6587	0.6620	0.4360	70.16%

Table 4 displays a summary list of categories before and after ranking. It clearly illustrates that those nine factors were only considered "building characteristics" factors. Instead of 12 factors, only 5 were considered structural and architectural factors. Instead of eight factors, only three were considered under the financial properties category. Instead of 4 factors, 2 were only considered social factors. In addition, instead of 8 factors, 2 factors were only considered under Governmental Factors. Instead of four factors, only one was taken into account under physical and environmental factors. Furthermore, instead of four factors, Finally, for the category of economic factors, it was shown that the total number of factors decreased from 23 to only 5. Figure 1 depicts the total number of factors influencing real estate unit price in the seven major groups both before and after the questionnaire ranking results. Figure 2 displays the weight of related factors in each group after ranking and questionnaire results.

Table 4. Factors Affecting Real Estate Unit Price Before and After Results of Ranking from Questionnaire

Group No.	Category	Factors in each group before Ranking		Factors in each group after Ranking	
		Sum	Weight	Sum	Weight
1	Building Characteristics	12	19.05%	9	33.33%
2	Structural and Architectural	8	12.70%	5	18.52%
3	Financial Properties	4	6.35%	3	11.11%
4	Social	8	12.70%	2	7.41%
5	Governmental	4	6.35%	2	7.41%
6	Physical/Environmental	4	6.35%	1	3.70%
7	Economic	23	36.51%	5	18.52%
	Total	**63**	**100.00%**	**27**	**100.00%**

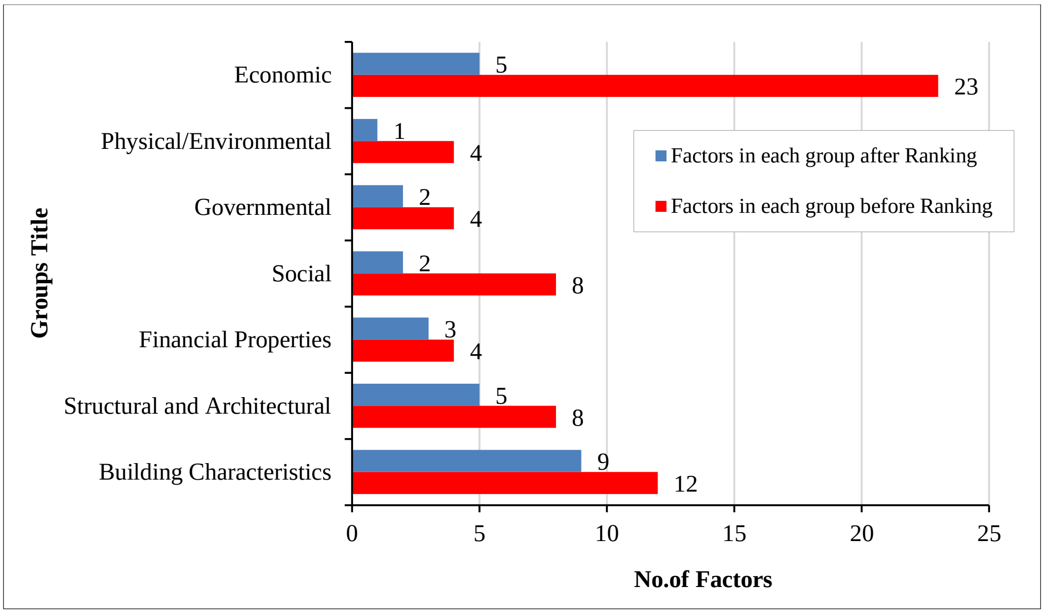

Figure 1. The Total Number of Factors in The Main Groups Before and After Ranking Results from Questionnaire

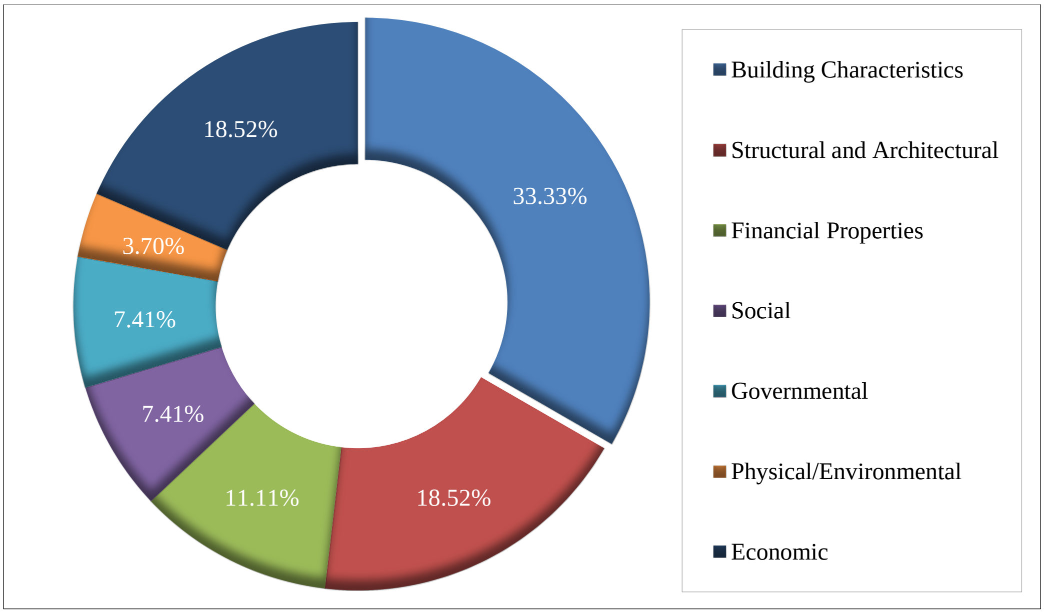

Figure 2. The Weight of Related Factors in Each Group After Questionnaire Results and Ranking

5. Real Estate Units Data Collection

Based on the research methodology, and after identifying the factors that most affect the price prediction of real estate units in Egypt, data on these factors has been collected from 2040 implemented housing units over a period of time between the years 2014 and 2022. Using Boxplot through SPSS Ver. 25 [29], it excluded 50 real estate unit data points that were considered to be outliers, thus becoming data used in the predictive model for 1990. So, Table 5 shows the adjusted R square of the gathered data before and after Boxplot application using the Enter Regression Model. The factors were divided into two groups: the first is the economic factors, which were calculated every three months according to reports issued by the Accountability State Authority of Egypt, as mentioned in Table 6. These factors are F23, F24, and F26. The second set of factors was gathered using information from each housing unit, as shown in Table 6.

Table 5. Data Adjusted R Square

Models Summary				
Model	**R**	**R Square**	**Adjusted R Square**	**Boxplot Application**
1	0.879	0.772	0.769	Before
2	0.898	0.806	0.803	After

Table 6. Form of Data Gathered Sessions

Factor No.	Factor Description	Scale	Unit of Measure	No. of Collected Units
F1	Total Land Size	Between (300- 720)	m^2	1990
F2	Gross Area of unit	Between (65- 240)	m^2	1990
F3	House Age	Between (5- 55)	Months	1990
F4	Land Value	Between (1.1- 9)	Million EGP	1990
F5	Location Type	Rural region	1	165
		Urban region	2	1825
F6	Building Type	Residential	1	1256
		Residential and commercial	2	734
F7	Exist Elevator	No Elevators	1	168
		Elevators Exist	2	1804
F8	Distance from The Mall	Far From the Mall	1	1083
		Near From the Mall	2	907
F9	The Distance from The Main Roads	Far From the Road	1	267
		Near From the Road	2	1723
F10	Number of Bedrooms	Two bedrooms	1	728
		Three bathrooms	2	1126
		more than three	3	134
F11	Number of Bathrooms	One Bathroom	1	670
		Two Bathrooms	2	1063
		More than two	3	257
F12	Number of Garages	Not exist	1	812
		Exists for each unit	2	1178
F13	The Garden Share	Without garden	1	1330
		With garden	2	660
F14	Level of Finishing (Cladding)	Half Finishing	1	533
		Full Finishing	2	850
		Extra Finishing	3	607
F15	The Preliminary estimated construction cost per m^2 based on the prices at the beginning of the project	Between (1300 - 5000)	EGP/m^2	1990
F16	Duration of The Construction	Between (8- 22)	Months	1990
F17	The Price of the unit at the beginning of the project per m^2	Between (1700 - 22000)	EGP/m^2	1990
F18	Population Trends in The City	Stability	1	523
		Decline	2	34
		Growth	3	1433
F19	The Standard Level in The Proposed Region	Normal	1	654
		Medium	2	876
		High	3	460
F20	Quality of Schools	Normal	1	1189
		Well	2	801
F21	Quality of Services	Normal	1	1306
		Well	2	684
F22	Transportation type	Private only	1	154
		Public and private	2	1836
F23	The Interest rate	Between (11- 19.70)	%	1990
F24	The inflation rate	Between (8.20- 33.30)	%	1990
F25	The economic climate	Bad	1	342
		Normal	2	1103
		Well	3	545
		Very Well	4	0
F26	Consumer price index (CPI)	Between (8.40- 32.20)	%	1990
F27	Average Months a house spends on the market	Between (3- 30)	Months	1990
Output	**Real Estate Unit Sold Price**	**Number**	**EGP**	

The average price per square meter of real estate units in many Egyptian cities researched and situated within Greater Cairo is shown in Table 7. The table illustrates the dramatic increase in the price per square meter in 2017, which coincided with the flotation of the Egyptian pound. Additionally, it is noticed that the growth occurred gradually between 2017 and 2022, indicating that 2017 marks the limit for splitting the price into two phases.

Table 7. Average Real Estate Unit Prices in Egyptian Cities

No.	City Name	Average price per square meter per year according to data collected EGP/m^2								
		2014	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**
1	El Haram	2000	2000	2100	2700	2800	3000	3000	3200	3500
2	Badr City	2100	2400	2500	3000	3100	3100	3200	3400	3700
3	Helwan	2500	2800	2900	3600	3600	3700	3800	4000	4400
4	El Jizah District	2600	2900	3000	3700	3700	3800	3900	4100	4600
5	Downtown	2600	2900	3100	3800	3800	3900	4000	4200	4500
6	El Oubour	2700	3100	3200	4000	4000	4100	4200	4400	4900
7	10th of Ramadan	2900	3300	3400	4200	4300	4400	4500	4700	5000
8	Hadayek El Ahram	2900	3300	3400	4200	4300	4400	4500	4700	5100
9	15th of May	2900	3200	3400	4100	4200	4300	4400	4600	5000
10	El Hadabah El Wosta	3100	3500	3600	4500	4500	4600	4700	5000	5600
11	El Koba Gardens	3100	3500	3600	4500	4500	4600	4700	5000	5400
12	6th of October	3200	3600	3700	4600	4600	4700	4800	5100	5600
13	New Heliopolis	3200	3600	3700	4600	4600	4700	4800	5100	5500
14	Mokattam	3400	3800	4000	4900	5000	5100	5200	5400	5700
15	Ain Shams	3400	3800	4000	4900	5000	5100	5200	5500	6000
16	El Abbasiya	3400	3900	4000	5000	5000	5200	5300	5500	6100
17	Shoubra	3500	3900	4100	5000	5100	5200	5300	5600	6000
18	Nasr City	3600	4100	4300	5300	5300	5400	5500	5800	6300
19	El Maadi	3700	4200	4400	5400	5500	5600	5700	6000	6500
20	Heliopolis – Masr El Gedida	4200	4700	4900	6100	6200	6300	6400	6800	7200
21	El Agouza	5100	5700	6000	7300	7400	7600	7800	8200	9000
22	El Sheikh Zayed City	5400	6200	6400	7900	8000	8200	8400	8800	9300
23	New Cairo – Fifth Settlement	5600	6400	6600	8200	8300	8500	8700	9100	9600
24	El Mohandeseen	5700	6400	6700	8200	8300	8500	8700	9200	9500
25	Dokki	6400	7200	7500	9300	9400	9600	9800	10300	10800
26	Manial	6500	7400	7700	9500	9600	9800	10000	10500	11100
27	Garden City	9900	11200	11700	14400	14600	14900	15200	16000	17500
28	El Zamalek	13000	14700	15300	18900	19100	19600	20000	21000	23000

6. Deep Learning Model Training

Deep learning is one of the most evidently complicated fields in which a computational structure is required. The cutting-edge research being conducted will result in numerous unique developments in the future. Pedregosa et al. [33]. Deep learning is a system of machine learning in artificial intelligence areas Mishra and Gupta [34] Deep learning has been employed in a variety of key fields, including signal processing solutions. Yu & Deng [35], medical image recognition Shen et al. [36] Deep learning is also used in applications from industries such as client services. Xu et al. [37], healthcare Miotto et al. [38]. Rafiei & Adeli [39] created a deep learning model to estimate building costs while accounting for economic variables in the construction field.

The purpose of this paper is to develop an innovative and more realistic estimation model for the real estate price problem using an advanced deep learning library developed in the Python programming language called Keras, and the steps of working with the Keras library are shown in Figure 3. The proposed model is a supervised regression deep learning model that is trained on 27 factors as input tensors to produce the predicted output. The model consists of three hidden layers of the Relu activation function to accurately find the relationship that maps the inputs into the output. Depending on Relu's activation function is a trail and-error procedure. The proposed model depends on the adaptive Root Mean Square Propagation (RMSprop) optimizer to optimize the model loss function.

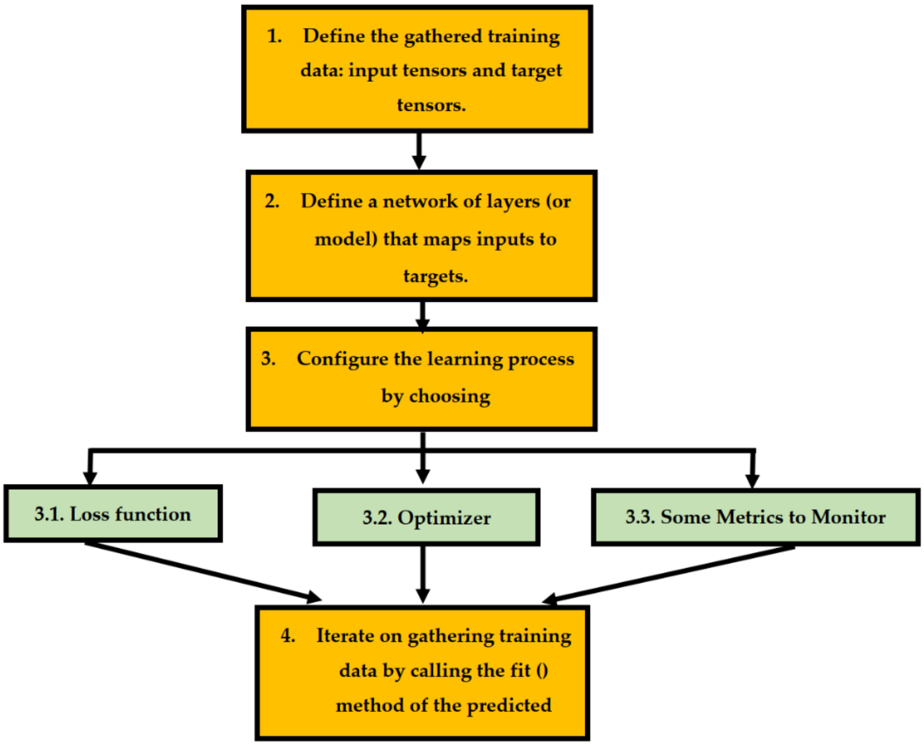

Figure 3. Keras Framework Steps in the deep learning model Chollet [40]

6.1. Data Preparation

The collected dataset was first divided into two subsets: the training dataset, which was used to train the model and calculate its weights, and the test dataset, which was used to validate the model's performance.

After dividing the datasets, there were some features with values in different ranges, so normalization is required to simplify the neural network learning process, such as factors (F4, F15, and F17). To deal with such data, a common best practice is to perform feature-wise normalization: for each feature in the input data (a column in the input data matrix), subtract the mean and then divide by the standard deviation, so that the feature is centered on 0 and has a unit standard deviation.

6.2. Adjusting the Model Parameters Using K-Fold Validation Algorithm

The K-fold cross-validation algorithm is used to validate our model. This algorithm divides the available training data into K partitions, instantiating K identical models, training each one on K – 1 partition, and evaluating the remaining partition. The validation score for the used model is then the average of the K validation scores obtained. Figure 4 shows a 4-fold validation algorithm that is used in validating the proposed deep learning.

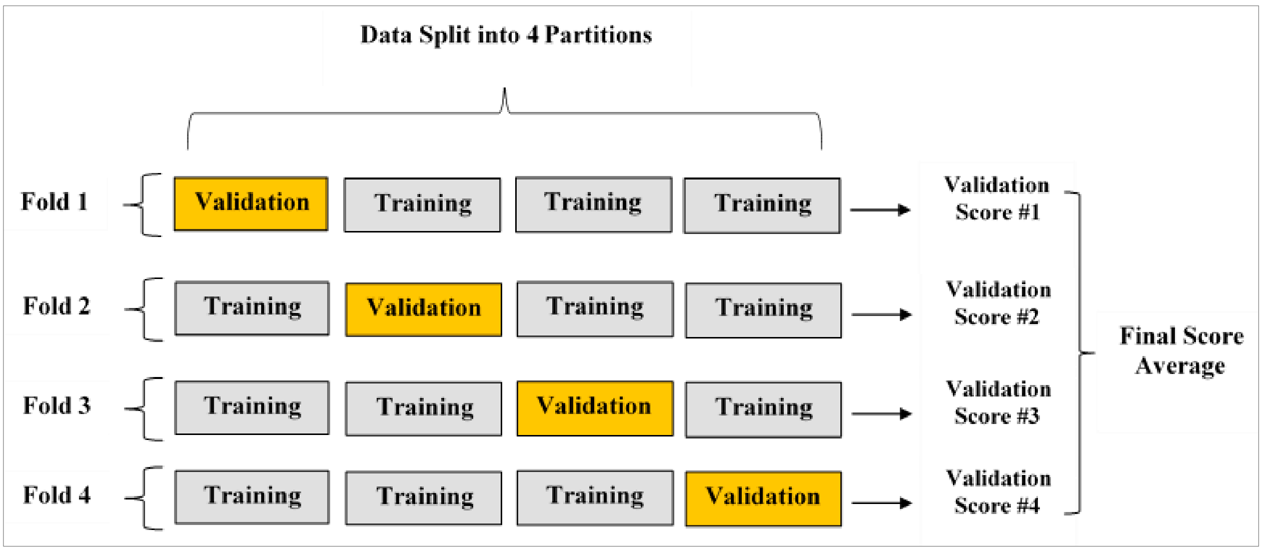

Figure 4. Validation Algorithm in The Deep Learning Model

6.3. Framework Model

As shown in Figure 5, there are two main phases to finding the best model for the current problem: phase one involves adjusting model parameters using the K-fold validation algorithm, and phase two involves minimizing the error function. During this phase, model parameters are constantly created, and the deep neural network model is tested using the k-fold cross-validation technique until the optimal model parameters with the lowest error value are obtained.

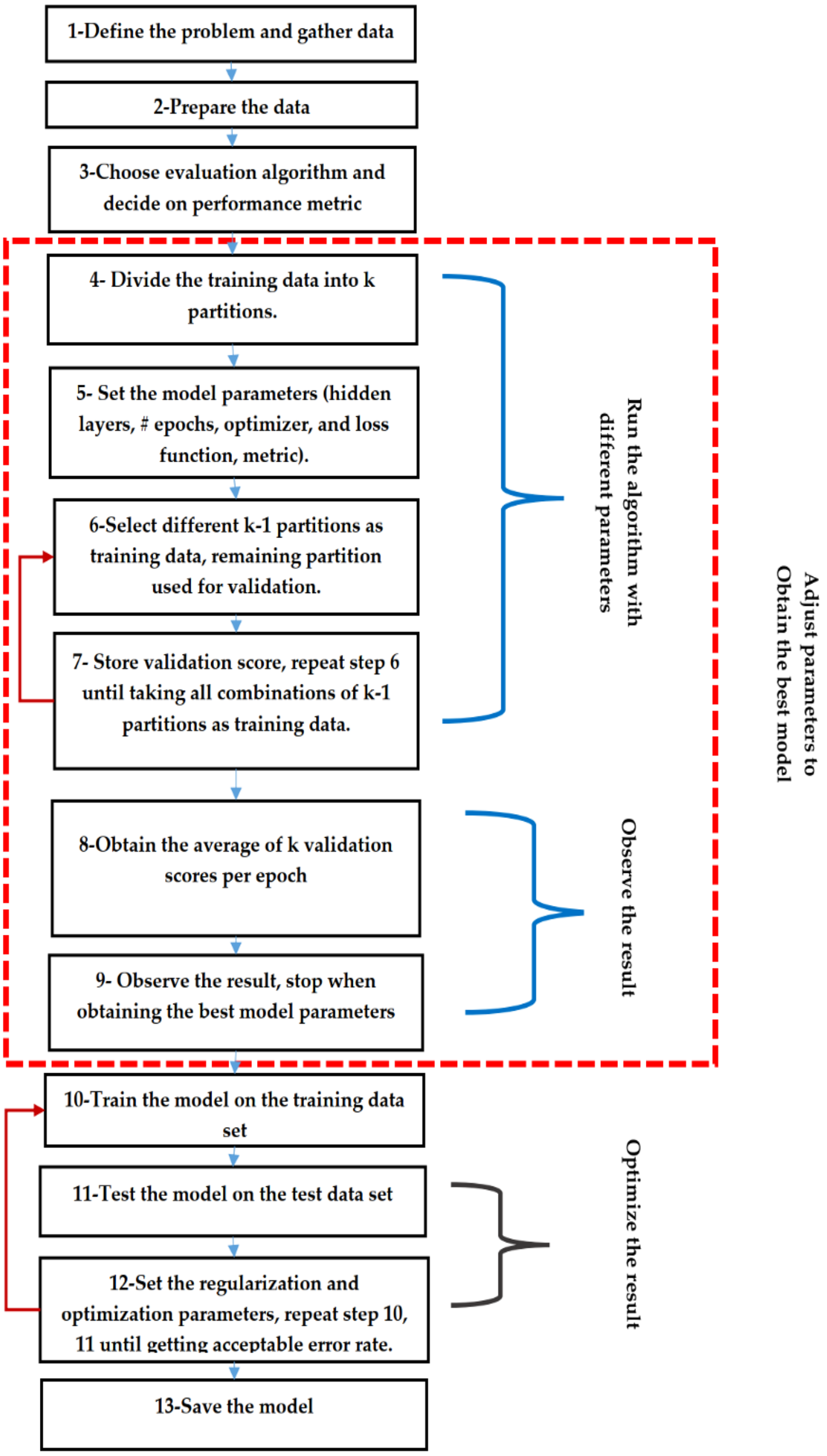

Figure 5. Deep Learning Model Framework

6.4. Different Hyper Parameters to Adjust in Training

The parameters that can be adjusted to get the most accurate model are:

- The number of hidden layers: several layers should be matching the data size and complexity. Taken (2 – 3 – 4) hidden layers will be chosen for the model.

- The activation function: There are many activation functions, such as sigmoid, Tanh, Relu, and soft max.
- The number of neurons: Biological Neurons are the fundamental units of the brain and nervous system. Taken from 50 to 150.
- Learning Rate: Using Adaptive optimizers learn such as RMSProp optimizer; Adam; and Adamax.
- Number of Epochs: Passing all input data through Forward Propagation and Backward Propagation once makes one epoch. Taken from 50 to 500.
- Loss Function Technique: Mean absolute error (MAE) is used for regression. The loss is the mean overseen data of the absolute differences between target and predicted values, where this is calculated for each Fold of 4-Fold. Then calculate the mean for the group of means produced per Fold, which is the basis for evaluating the proposed model, known as the Mean (MAE) Chollet [40].

At the end, Table 8 displays all the Parameters that have been selected for training and also shows the best Parameters in terms of accuracy of results that have been stabilized to develop the final model.

Table 8. Hybrid Parameters Adjustments

	Number of Hidden Layers	Activation Function	Number of Neurons	Learning Rate	Number of Epochs
Parameters Features	2; 3; 4	Sigmoid; Tanh; Relu; and Softmax	50 up to 150	Adam; Adamax; and RMSProp Optimizers	50 up to 500
Best Parameters	3	Relu	100	RMSProp Optimizer	up to 100

6.5. Minimizing the Error Function

This phase includes three steps:

1-Train the model on the entire training data set: after obtaining the model with the best parameters, now the model will be trained on the entire training data set.

2-Test the model on the test data set: now the model is tested to show if the error rate is acceptable.

3-Set the regularization and optimization parameters: this is an optimization step for the performance metrics (MAE in the proposed case). the regularization technique and parameters are continually adjusted, and also test another optimizer to get an acceptable error rate and remove the overfitting and underfitting effects. Figure 6 shows the architecture of the proposed model.

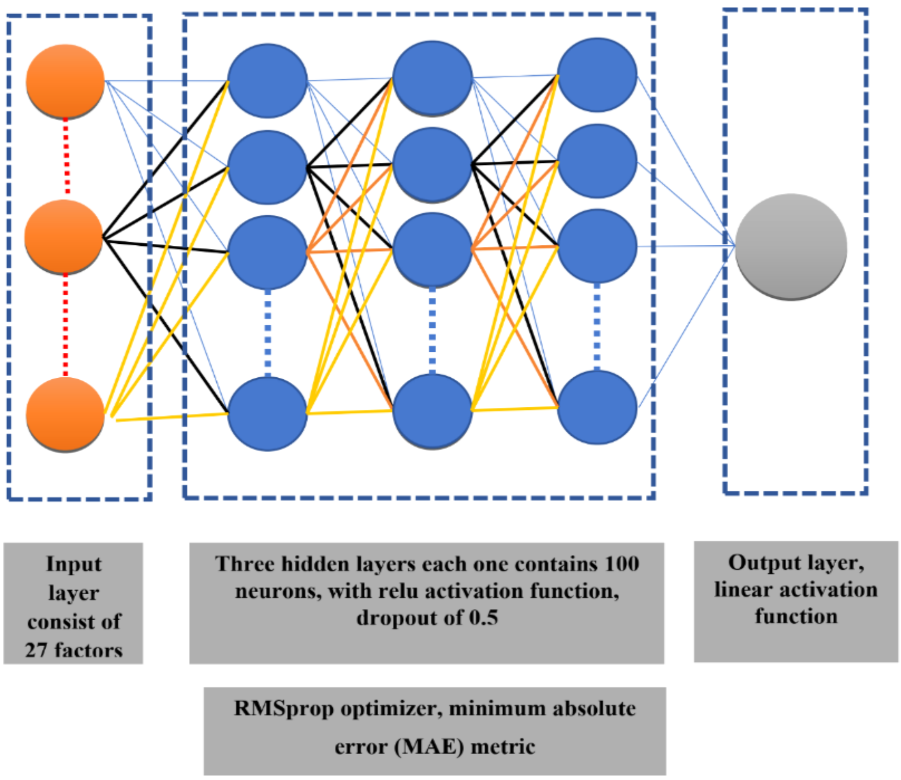

Figure 6. Deep Learning Model Architecture

6.6. Implementation

Python was used to implement the suggested model. All trails have been directed on Google Colaboratory, often known as Google collab, which is a set of free tools built by Google to assist deep-learning researchers and specialists in analysing data more easily, precisely, and rapidly Chollet [40].

7. Development, Testing and Validation of The Final DL Model

To adjust the parameters of the model, the model has been run many times with different parameters. A K-fold validation algorithm was used for validating the result. The result was displayed using the matplotlib library for exploring the performance of the model using different parameters. Running the model with three hidden layers, the Relu activation function, 500 epochs, and the RMSprop optimizer—the most accurate parameters—resulted in a maximum error rate of about 7.27 percent during training, but when tested on never-before-seen test data, the maximum error rate was about 18.18 percent, indicating that regularising the model was necessary.

After performing dropout regularization of 0.6 on the model layers, the error ratio at training was about 10.58%, but at testing and validation, it was about 9.50%. Figure 7 displays the result of running the model after regularization at the training step. Table 9 shows the best two training trails for the model. Table 10 shows the outcomes of testing the two best models, as well as their output on test and validated data.

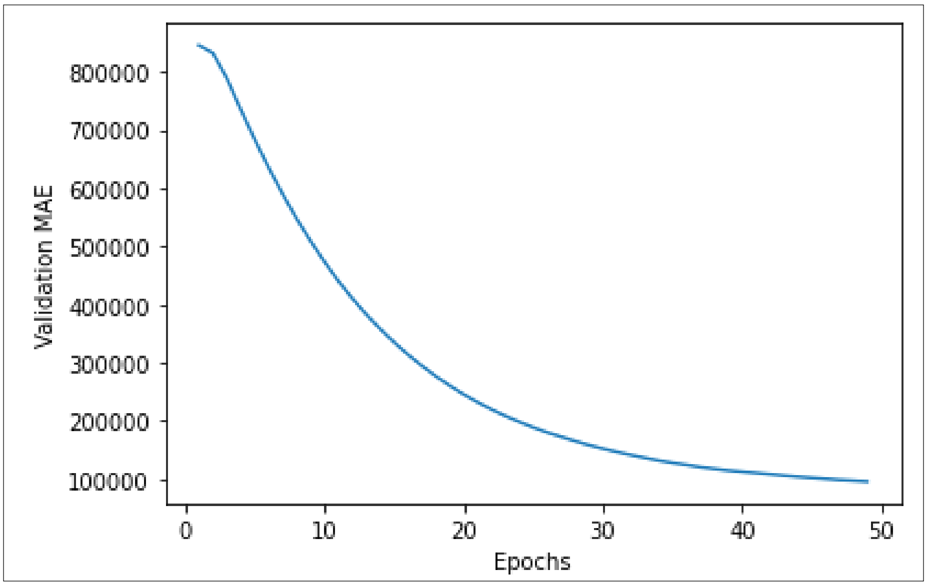

Figure 7. Model Performance after Regularization

Table 9. The Best Two Model After Training

Trail No.	No. of Hidden Layers	No. of Neurons in Each Layer	No. of epochs	Folds No.	Max MAE	Mean MAE	Min MAE	Error%	Activation Functions	The applicable Optimizer	Regularization Technique
4	3	100	500	4	861,495	65,408	54,497	7.27%	Relu, Relu, Relu	RMSProp	No
16	**3**	**100**	**250**	**4**	**878,612**	**95,253**	**81,701**	**10.58%**	**Relu, Relu, Relu**	**RMSProp**	**dropout, 0.6**

Table 10. Testing the Best Two Models

Model No.	No. Hidden Layers	No. epoch	No. Neurons in each layer	Batch Size	Mean MAE	Max. Error	Activation Functions	The applicable Optimizer	Regularization Technique
4	3	100	100	1	163,638	18.18%	Relu, Relu, Relu	RMSProp	No
16	**3**	**100**	**100**	**1**	**85,460**	**9.50%**	**Relu, Relu, Relu**	**RMSProp**	**dropout, 0.6**

Table 11 illustrates the total trails in training the models and Table 12 shows the results of testing these models in Table 11.

Table 11. Training the Model Sample Trails

Trail No.	No. of Hidden Layers	No. of Neurons in Each Layer	No. of epochs	Folds No.	Max MAE	Mean MAE	Min MAE	Error	Activation Functions	The applicable Optimizer	Regularization Technique
1	2	100	50	4	881,252	462,155	149,925	51.35%	Relu, Relu	RMSProp	No
2	3	100	50	4	858,846	107,281	62,542	11.92%	Relu, Relu, Relu	RMSProp	No
3	3	100	100	4	858,675	85,821	61,287	9.54%	Relu, Relu, Relu	RMSProp	No
4	3	100	500	4	861,495	65,408	54,497	7.27%	Relu, Relu, Relu	RMSProp	No
5	3	100	500	4	881,334	881,175	881,017	97.91%	Softmax, Softmax, Softmax	RMSProp	No
6	3	100	500	4	881,292	873,311	865,331	97.03%	Tanh, Tanh, Tanh	RMSProp	No
7	3	100	200	4	733,094	68,182	55,130	7.58%	Relu, Relu, Relu	Adam	No
8	3	100	200	4	879,580	85,617	59,902	9.51%	Relu, Relu, Relu	Adamax	No
9	4	100	50	4	376,532	74,286	57,684	8.25%	Relu, Relu, Relu, Relu	RMSProp	No
10	3	100	50	4	767,814	104,741	71,297	11.64%	Relu, Relu, Relu	RMSProp	dropout, 0.5
11	3	100	50	4	767,295	103,545	65,622	11.50%	Relu, Relu, Relu	RMSProp	dropout, 0.5 l2(0.001)
12	3	100	50	4	833,058	118,193	74,500	13.13%	Relu, Relu, Relu	RMSProp	dropout, 0.6 l2(0.001)
13	3	100	50	4	831,190	115,840	71,172	12.87%	Relu, Relu, Relu	RMSProp	dropout, 0.6 l1_l2(l1=0.001, l2=0.001)
14	3	100	50	4	842,411	122,142	75,006	13.57%	Relu, Relu, Relu	RMSProp	dropout, 0.6 l1_l2(l1=0.001, l2=0.001)
15	3	100	250	4	857,926	126,837	71,419	14.09%	Relu, Relu, Relu	RMSProp	dropout, 0.6 l1_l2(l1=0.001, l2=0.001)
16	**3**	**100**	**250**	**4**	**878,612**	**95,253**	**81,701**	**10.58%**	**Relu, Relu, Relu**	**RMSProp**	**dropout, 0.6**

Table 12. Testing the Final Models

No.	No. Hidden Layers	No. epoch	No. Neurons in each layer	Batch Size	Mean MAE	Max. Error	Activation Functions	The applicable Optimizer	Regularization Technique
1	2	100	100	4	195,352	22.13%	Relu, Relu, Relu	RMSProp	No
2	3	100	100	4	172,451	19.16%	Relu, Relu, Relu	RMSProp	No
3	3	100	100	1	176,262	19.58%	Relu, Relu, Relu	RMSProp	No
4	3	100	100	1	163,638	18.18%	Relu, Relu, Relu	RMSProp	No
5	3	100	100	4	254,826	38.42%	Softmax, Softmax, Softmax	RMSProp	No
6	3	100	100	4	311,236	45.23%	Tanh, Tanh, Tanh	RMSProp	No
7	3	100	100	8	165,742	18.42%	Relu, Relu, Relu	RMSProp	No
8	3	100	100	16	178,017	19.78%	Relu, Relu, Relu	RMSProp	No
9	4	100	100	32	193,513	21.50%	Relu, Relu, Relu, Relu	RMSProp	No
10	3	100	100	1	172,253	19.14%	Relu, Relu, Relu	RMSProp	dropout, 0.5
11	3	100	100	1	157,944	17.55%	Relu, Relu, Relu	RMSProp	dropout, 0.5 l2(0.001)
12	3	100	100	1	160,488	17.83%	Relu, Relu, Relu	RMSProp	dropout, 0.6 l2(0.001)
13	3	100	100	1	155,615	17.29%	Relu, Relu, Relu	RMSProp	dropout, 0.6 l1_l2(l1=0.001, l2=0.001)
14	3	100	100	1	163,704	18.19%	Relu, Relu, Relu	RMSProp	dropout, 0.6 l1_l2(l1=0.001, l2=0.001)
15	3	100	100	1	157,511	13.50%	Relu, Relu, Relu	RMSProp	dropout, 0.6 l1_l2(l1=0.001, l2=0.001)
16	**3**	**100**	**100**	**1**	**85,460**	**9.50%**	**Relu, Relu, Relu**	**RMSProp**	**dropout, 0.6**

To explain the results more clearly, Figure 8 illustrates the final error % of each model in calculating the price per trail from Table 11. Figure 9 also illustrates the error % for each final model test trail found in Table 12**.**

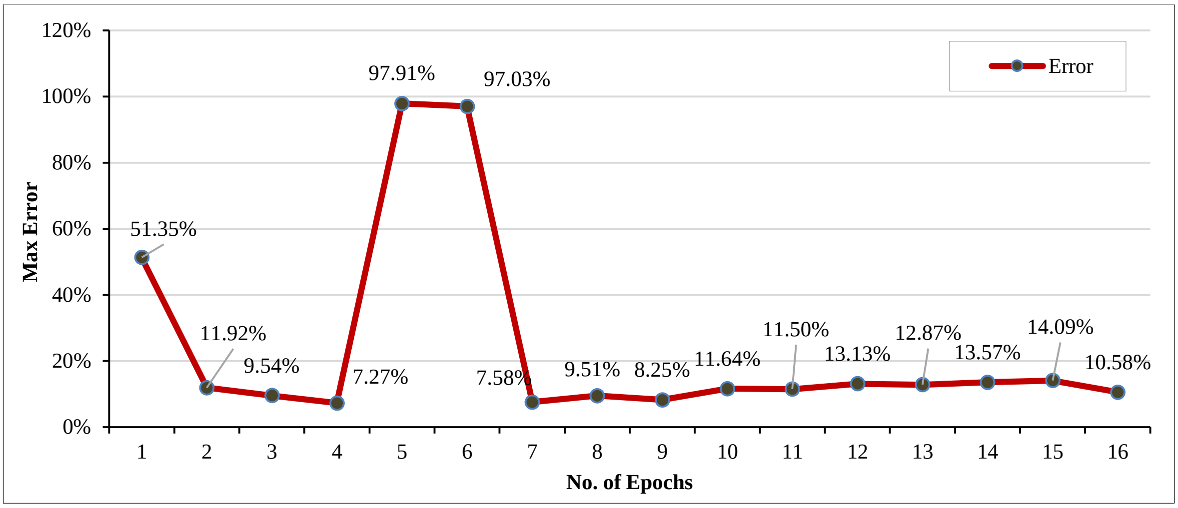

Figure 8. Error% in Each Model Trail

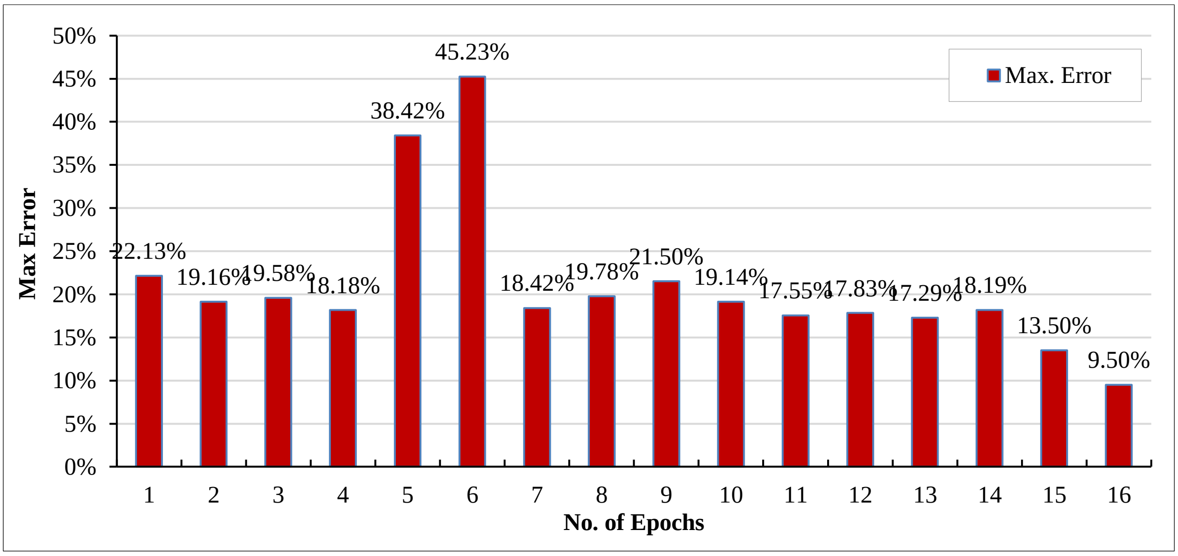

Figure 9. Max. Error% in Each Test Trail

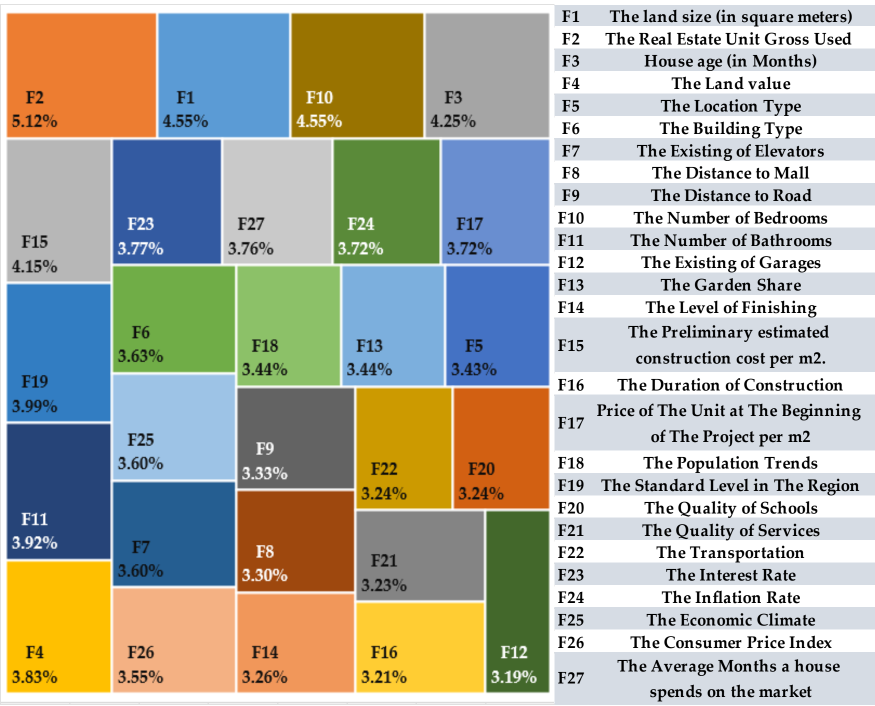

Figure 10. Percentage Impact of Each Factor on the Model

After the development of the final model, it was necessary to know the degree to which each factor had an impact on this model. The real estate unit gross used area was 5.121% larger than the land size (in square meters), 4.547%. The least influential factor was the duration of construction at 3.212%, followed by the existing garages at 3.189%, which is considered to have the least impact on the model. Figure 10 shows the impact of each factor on the final model.

8. Final Model Graphical User Interface Tool Development

To get the benefit of the final model that has been built after a huge number of attempts and to help real estate developers and investors in the construction sector find the right price for the real estate units they build, the final model of the Collab was saved in the Jason format, and the final model was programmed using Python using the PyCharm IDE by calling the model using the Keras library. The GUI work of the model is in the final format, and also the Tkinter library was used.

The GUI model was presented in four split screens, as well as three input screens and the fourth output screen. The first screen displays the factors that are introduced in quantitative form as shown in Figure 11, and the second and third screens display the factors that are introduced qualitatively as shown in Figures 12 and 13. The fourth screen displays the final predicted price of the real estate unit, as shown in Figure 14.

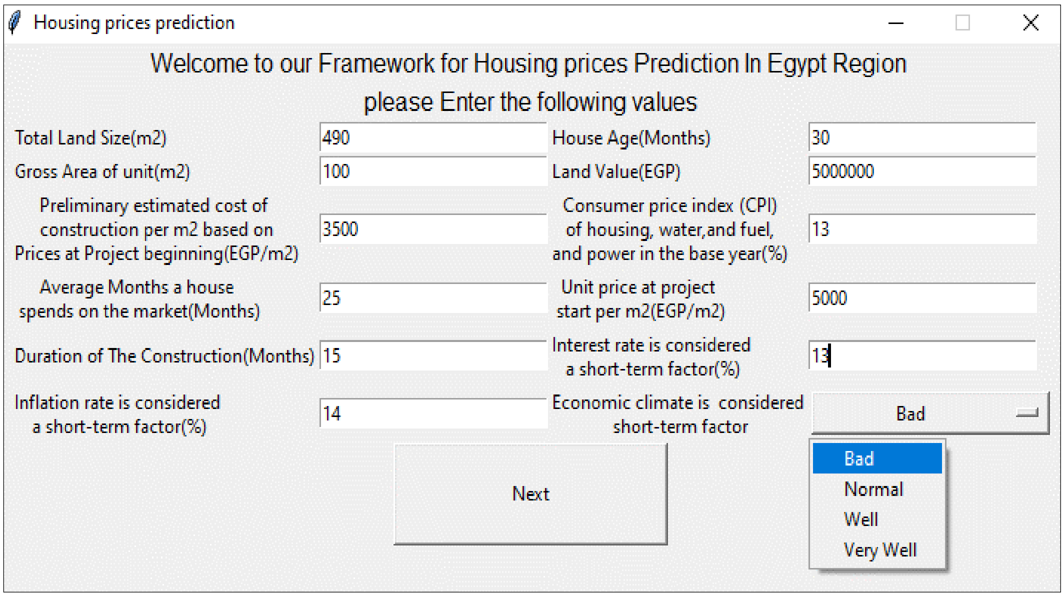

Figure 11. The First Data Input Screen

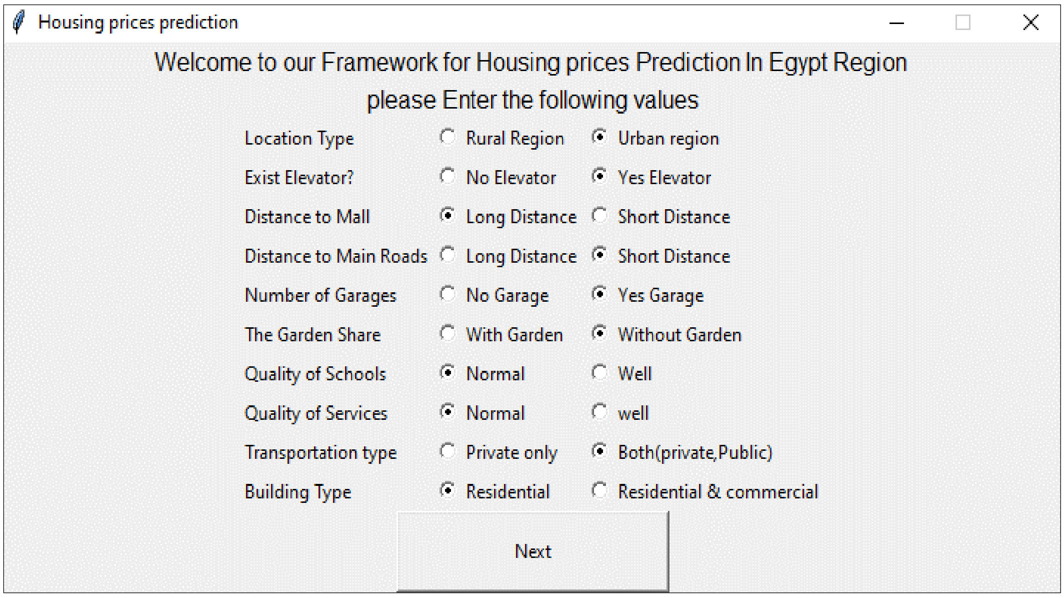

Figure 12. The Second Data Input Screen

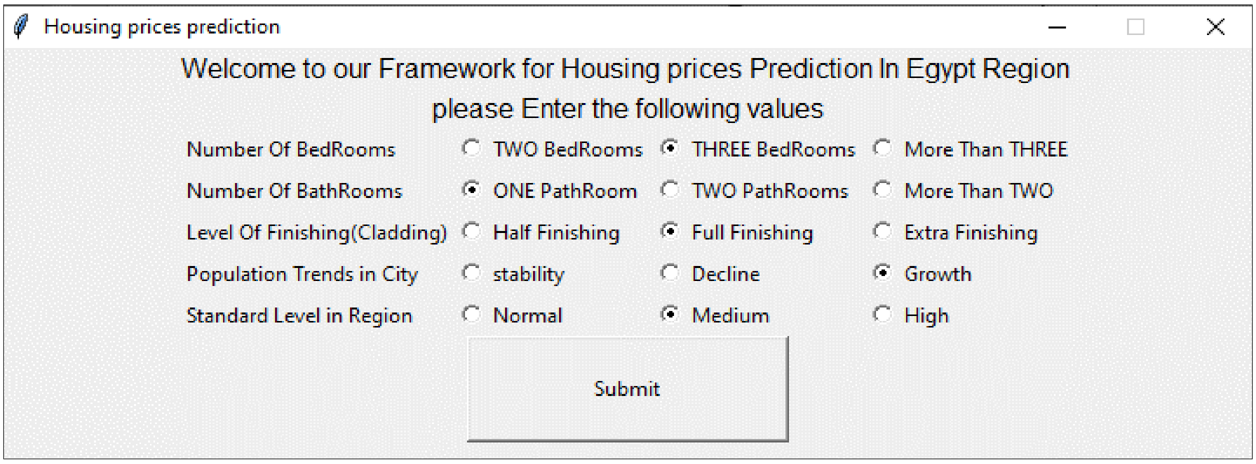

Figure 13. The Third Data Input Screen

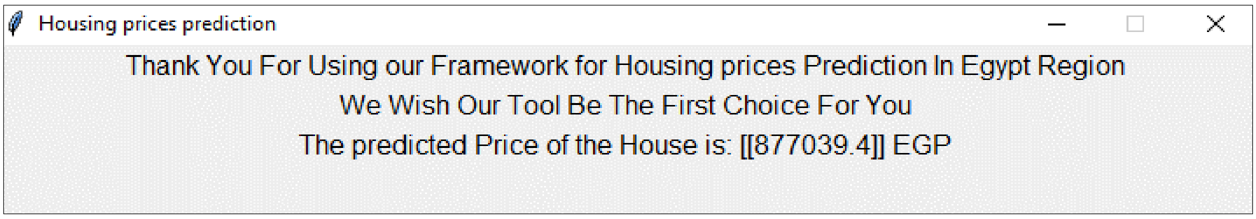

Figure 14. The Predicted Price Screen

9. Conclusion

This paper reviewed a wide range of factors affecting the determination of the price of the real estate units. A questionnaire was prepared to find the most important factors influencing the 27 factors, and the predictive model was built. It also investigated an innovative deep learning model for housing price prediction. The raw data collected from 2040 real estate units in Egypt from 28 cities over 8 years, using Boxplot through the SPSS Ver. 25 program, excluded 50 real estate unit data points that were considered to be outliers, thus becoming data used in the predictive model 1990. 1700 were used in the model's training process, and another 290 in the model's validation using the deep learning technique. The deep learning model was done using the Keras Python software library, which contains the TensorFlow backend, the Theano backend, and the Microsoft Knowledge Toolkit (CNTK) to give more accuracy to the developed model. A GUI tool is created to make use of the final predictive model, which is very simple for the real estate developers. A huge number of attempts to train the model have been tried. Table 11 illustrates the best part of the attempts that have been tried to train the model to find the most accurate parameters to build a model that gives accurate results.

The parameters are the number of hidden layers, the number of neurons, and the type of activation function. Using a range of optimizers, such as RMSProp, Adam, and Adamax, this was done without using regularization. The best parameters were reached, and they were three hidden layers with 100 neurons in each layer, and the best activation function is Relu, so the value of the mean MAE was 65,408 EGP and the maximum error percentage was 7.27%, as shown in Table 9, but when these parameters were used without regularization in the test process, as shown in Table 10, the value of the mean MAE was 163,638 EGP and the maximum error percentage was 18.18%. Regularization was then used using the RMSProp Optimizer, using L1, L2, and L1_L2 regularizing technologies, and changing dropout values between 0.50 and 0.60, as shown in Table 11. The most accurate parameters were reached, and they were 3 hidden layers with 100 neurons in each layer, and the best activation function is Relu. Accordingly, the value of the mean MAE was 95,253 EGP and the maximum error percentage was 10.58% as shown in Table 9, so when these parameters were used in the test process as shown in Table 10, the value of the mean MAE was 85,460 EGP and the maximum error percentage was 9.50%. So now the final model for predicting the price of real estate units in Egypt is ready for use. The new model could either help construction companies deal with construction or not. In a market, developers may decide to suspend the start of construction, depending on the future sale price provided by the model. As a result, the proposed model assists the builder in logically deciding whether or not to build at a given time. As a consequence, the proposed model could be an invaluable tool for construction companies.

10. Declarations

10.1. Author Contributions

Conceptualization, O.A.H., A.H.I., and H.H.M.; methodology, A.H.I., and H.H.M.; software, O.A.H.; validation, O.A.H., A.H.I., and H.H.M.; formal analysis, O.A.H.; investigation, A.S.E.T.; resources, A.H.I.; data curation, H.H.M.; writing—original draft preparation, O.A.H.; writing—review and editing, O.A.H., A.H.I., and H.H.M. All authors have read and agreed to the published version of the manuscript.

10.2. Data Availability Statement

The data presented in this study are available in the article.

10.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

10.4. Conflicts of Interest

The authors declare no conflict of interest.

11. References

[1] Sirmans, G. S., Macpherson, D. A., & Zietz, E. N. (2005). The composition of hedonic pricing models. Journal of Real Estate Literature, 13(1), 3–43. doi:10.1080/10835547.2005.12090154.

[2] Li, L. (2015). Prices and Bubbles: Factors Affecting the Chinese Real Estate Market. Master Thesis, Lund University, Lund Sweden.

[3] Selim, H. (2009). Determinants of house prices in Turkey: Hedonic regression versus artificial neural network. Expert Systems with Applications, 36(2), 2843–2852. doi:10.1016/j.eswa.2008.01.044.

[4] Borowiecki, K. J. (2009). The Determinants of House Prices and Construction: An Empirical Investigation of the Swiss Housing Economy. International Real Estate Review, 12(3), 193–220. doi:10.53383/100112.

[5] Rafiei, M. H., & Adeli, H. (2016). A Novel Machine Learning Model for Estimation of Sale Prices of Real Estate Units. Journal of Construction Engineering and Management, 142(2), 04015066, 1–8. doi:10.1061/(asce)co.1943-7862.0001047.

[6] Limsombunc, V., Gan, C., & Lee, M. (2004). House Price Prediction: Hedonic Price Model vs. Artificial Neural Network. American Journal of Applied Sciences, 1(3), 193–201. doi:10.3844/ajassp.2004.193.201.

[7] Khalafallah, A. (2008). Neural network based model for predicting housing market performance. Tsinghua Science and Technology, 13(S1), 325–328. doi:10.1016/s1007-0214(08)70169-x.

[8] Rahman, M. M. (2008). Australian housing market: causes and effects of rising price. Proceedings of the 37th Australian Conference of Economists (ACE 2008), 28 Sptember-4 October, 2008, Gold Coast, Australia.

[9] Candas, E., Kalkan, S. B., & Yomralioglu, T. (2015). Determining the factors affecting housing prices. FIG Working Week, 17-21 May, 2015, Sofia, Bulgaria.

[10] Malpezzi, S. (2008). Hedonic Pricing Models: A Selective and Applied Review. Housing Economics and Public Policy, 67–89. doi:10.1002/9780470690680.ch5.

[11] Yusof, A., & Ismail, S. (2012). Multiple Regressions in Analyzing House Price Variations. Communications of the IBIMA, 1–9. doi:10.5171/2012.383101.

[12] Pow, N., Janulewicz, E., & Liu, L. (2014). Applied Machine Learning Project 4 Prediction of real estate property prices in Montréal. Course project, COMP-598, Fall/2014, McGill University, Montreal, Canada.

[13] Gustafsson, A. & Wogenius, S. (2014). Modelling apartment prices with the multiple linear regression model. Degree Project in Applied Mathematics and Industrial Economics, Royal Institute of Technology, School of Engineering Sciences, Stockholm, Sweden.

[14] Ozgur, C., Hughes, Z., Rogers, G., & Parveen, S. (2016). Multiple linear regression applications in real estate pricing. International Journal of Mathematics and Statistics Invention (IJMSI), 4(8), 30-59.

[15] Dai, H., Xue, G., & Wang, W. (2014). An Adaptive Wavelet Frame Neural Network Method for Efficient Reliability Analysis. Computer-Aided Civil and Infrastructure Engineering, 29(10), 801–814. doi:10.1111/mice.12117.

[16] Story, B. A., & Fry, G. T. (2014). A structural impairment detection system using competitive arrays of artificial neural networks. Computer-Aided Civil and Infrastructure Engineering, 29(3), 180–190. doi:10.1111/mice.12040.

[17] Butcher, J. B., Day, C. R., Austin, J. C., Haycock, P. W., Verstraeten, D., & Schrauwen, B. (2014). Defect detection in reinforced concrete using random neural architectures. Computer-Aided Civil and Infrastructure Engineering, 29(3), 191–207. doi:10.1111/mice.12039.

[18] Patil, P., Shah, D., Rajput, H. & Chheda, J. (2020). House Price Prediction Using Machine Learning and RPA. International Research Journal of Engineering and Technology, 7(3), 5560-5563.

[19] Cheung, K. S., Yiu, C. Y., & Xiong, C. (2021). Housing Market in the Time of Pandemic: A Price Gradient Analysis from the Covid-19 Epicentre in China. Journal of Risk and Financial Management, 14(3), 108. doi:10.3390/jrfm14030108.

[20] Ho, W. K. O., Tang, B. S., & Wong, S. W. (2021). Predicting property prices with machine learning algorithms. Journal of Property Research, 38(1), 48–70. doi:10.1080/09599916.2020.1832558.

[21] Hu, M. R., Lee, A. D., & Zou, D. (2021). COVID-19 and Housing Prices: Australian Evidence with Daily Hedonic Returns. Finance Research Letters, 43. doi:10.1016/j.frl.2021.101960.

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

[23] Corsini, K. R. (2009). Statistical analysis of residential housing prices in an up and down real estate market: a general framework and study of Cobb County, GA. Ph.D. Thesis, Georgia Institute of Technology, Atlanta, United States.

[24] Haider, M., & Miller, E. J. (2000). Effects of Transportation Infrastructure and Location on Residential Real Estate Values: Application of Spatial Autoregressive Techniques. Transportation Research Record: Journal of the Transportation Research Board, 1722(1), 1–8. doi:10.3141/1722-01.

[25] Jian, D. S., & Zhang, X. W. (2012). The Real Estate Economics. Shanghai: University of Finance and Economics, 23-78.

[26] Ding, J. (2014). An empirical analysis of factors affecting Chinese Real Estate prices. Central China Normal University. Lingnan Journal of Banking, Finance and Economics, 5(1), 1-13.

[27] Cao, Z. L. (2003). The General Theory of Real Estate Economics. Peking University Press, Beijing, China.

[28] Bartlett, J. E., Kotrlik, J. W. K. J. W., & Higgins, C. (2001). Organizational research: Determining appropriate sample size in survey research appropriate sample size in survey research. Information Technology, Learning, and Performance Journal, 19(1), 1-43.

[29] Statistical Package for Social Sciences (SPSS). (2017). Ver. 25.0, IBM Corp, Armonk, New York, United States.

[30] Gab-Allah, A. A., Ibrahim, A. H., & Hagras, O. A. (2015). Predicting the construction duration of building projects using artificial neural networks. International Journal of Applied Management Science, 7(2), 123–141. doi:10.1504/IJAMS.2015.069259.

[31] Mohmad, H. H., Ibrahim, A. H., & El Nagar, H. H. (2016). Pre-Tender cost estimate of agriculture subsurface drainage projects. International Journal of Applied Management Science, 8(4), 271–289. doi:10.1504/IJAMS.2016.080314.

[32] El Touny, A. S., Ibrahim, A. H., & Mohamed, H. H. (2021). An integrated sustainable construction project's critical success factors (Iscsfs). Sustainability (Switzerland), 13(15). doi:10.3390/su13158629.

[33] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. Journal of machine learning research, 12, 2825-2830.

[34] Mishra, C., & Gupta, D. L. (2016). Deep Machine Learning and Neural Networks: An Overview. International Journal of Hybrid Information Technology, 9(11), 401–414. doi:10.14257/ijhit.2016.9.11.34.

[35] Yu, D., & Deng, L. (2011). Deep Learning and Its Applications to Signal and Information Processing [Exploratory DSP. IEEE Signal Processing Magazine, 28(1), 145–154. doi:10.1109/msp.2010.939038.

[36] Shen, D., Wu, G., & Suk, H.-I. (2017). Deep Learning in Medical Image Analysis. Annual Review of Biomedical Engineering, 19(1), 221–248. doi:10.1146/annurev-bioeng-071516-044442.

[37] Xu, A., Liu, Z., Guo, Y., Sinha, V., & Akkiraju, R. (2017). A New Chatbot for Customer Service on social media. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. doi:10.1145/3025453.3025496.

[38] Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2017). Deep learning for healthcare: Review, opportunities and challenges. Briefings in Bioinformatics, 19(6), 1236–1246. doi:10.1093/bib/bbx044.

[39] Rafiei, M. H., & Adeli, H. (2018). Novel Machine-Learning Model for Estimating Construction Costs Considering Economic Variables and Indexes. Journal of Construction Engineering and Management, 144(12), 1–9. doi:10.1061/(asce)co.1943-7862.0001570.

[40] Chollet, F. (2021). Deep learning with Python. Simon and Schuster, Manhattan, New York, United States.