T.C.
ISTANBUL AYDIN UNIVERSITY
INSTITUTE OF GRADUATE STUDIES

COMPARING THE HEDONIC MODEL VS. ARTIFICIAL NURAL NETWORK IN HOUSING PRICE PREDICTION IN ISTANBUL, TURKEY

MASTER'S THESIS

Mohammed Hussein Ibadi

Department of Business
Business Administration Program

JUNE, 2021

T.C.
ISTANBUL AYDIN UNIVERSITY
INSTITUTE OF GRADUATE STUDIES

COMPARING THE HEDONIC MODEL VS. ARTIFICIAL NEURAL NETWORK IN HOUSING PRICE PREDICTION IN ISTANBUL, TURKEY

MASTER'S THESIS

Mohammed Hussein Ibadi
(Y1712.130060)

Department of Business
Business Administration Program

Thesis Advisor: Asst. Prof. Doçent Dr. Salvatore Joseph Terregrossa

JUNE, 2021

ONAY FORMU

DECLARATION

I hereby declare with the respect that the study “Combining Hedonic Model Forecasts With Artificial Neural Network Forecasts Of Housing Prices In Istanbul, Turkey”, which I submitted as a Master thesis, is written without any assistance in violation of scientific ethics and traditions in all the processes from the project phase to the conclusion of the thesis and that the works I have benefited are from those shown in the Bibliography. (27/03/2020)

MOHAMMED HUSSEIN IBADI

FOREWORD

To the ones who made everything that I am now my Father and Mother, to my light at the end of the tunnel my sister Noor may you have a blessed happy life, to my beloved brother and sisters, my family and friends thank you all for the help and support, and for the souls who left us early. Last but not least, to my advisor Mr. Terregrossa, thank you for all the knowledge and efforts, this work could not be done without your continuous help and encouragement.

June,2021 Mohammed Hussein Ibadi

COMBINING HEDONIC MODEL FORECASTS WITH ARTIFICIAL NEURAL NETWORK FORECASTS OF HOUSING PRICES IN ISTANBUL, TURKEY

ABSTRACT

BACKGROUND: In forecasting housing-unit prices, the conventional hedonic model is the most common used method, and the newer artificial neural network (ANN) models also have been used in forecasting various, different economic and financial variables. More recently, combining forecasts models have been introduced to enhance forecasting accuracy. The current study aims in developing a more accurate forecasting model by combining the hedonic and neural network forecasts of housing-unit prices.

MATERIAL AND METHODS: A total of 100 apartments in Istanbul, Turkey were included in the study. Housing-unit characteristics were taken including the price, the geographical location, the land size, the age of apartments, the number of bedrooms and bathrooms, and the floor within the building. First, the hedonic and ANN models are applied, and their forecasts are compared to detect the better model. Second, combining models are generated by combining the forecasts of hedonic and ANN model using different sets of forecast weights, generated by restricted and unrestricted, weighted least squares (WLS) regression technique, respectively. Average absolute forecast error (MAFE) of each model is calculated, and the average difference in MAFE among all pairs of models are compared and tested, and the superior model is the one with the lowest average absolute forecasting error (MAFE).

RESULTS: The study finds that between the ANN- and hedonic models, the ANN model performs better. However, the ANN model was outperformed by the combination forecast formed with restricted WLS estimated regression coefficients as component forecast weights. In all, seven combining models were generated using different methods in calculating the forecast weights. The unrestricted combining models were outperformed by the ANN model, while the constrained combining models performed better than ANN model. The study finds that the restricted combining models have the lowest MAFEs and are considered as the superior models.

CONCLUSION: The present study successfully generates combining forecasts models from the housing units' forecasts of the hedonic and neural network models. The study finds the combining forecasts model formed with weights generated by constrained WLS regressions generally perform the best out of all other forecasts' models. Our study demonstrates that combining forecasts can improve estimations of housing units' prices in Istanbul, Turkey.

Keywords: Forecasting, hedonic, artificial neural network, combining model, Istanbul, Turkey.

HEDONİK MODEL TAHMİNLERİNİN YAPAY SİNİR AĞLARI İLE BİRLEŞTİRİLMESİ İSTANBUL, TÜRKİYE KONUT FİYAT TAHMİNLERi

ÖZET

ARKA PLAN: Konut birim fiyatlarının tahmininde, geleneksel hedonik model en yaygın kullanılan yöntemdir ve daha yeni yapay sinir ağı (YSA) modelleri de çeşitli, farklı ekonomik ve finansal değişkenlerin tahmininde kullanılmıştır. Daha yakın zamanlarda, tahmin doğruluğunu artırmak için tahmin modellerini birleştirmek tanıtıldı. Mevcut çalışma, konut birim fiyatlarının hedonik ve sinir ağı tahminlerini birleştirerek daha doğru bir tahmin modeli geliştirmeyi amaçlamaktadır.

GEREÇ VE YÖNTEM: Çalışmaya İstanbul ilinde toplam 100 daire dahil edilmiştir. Fiyat, coğrafi konum, arazi büyüklüğü, dairelerin yaşı, yatak odası ve banyo sayısı ve bina içindeki zemin dahil olmak üzere konut-birim özellikleri alınmıştır. İlk olarak, hedonik ve YSA modelleri uygulanmakta ve daha iyi modeli bulmak için tahminleri karşılaştırılmaktadır. İkinci olarak, birleştirici modeller, sırasıyla kısıtlı ve kısıtsız, ağırlıklı en küçük kareler (WLS) regresyon tekniği ile oluşturulan farklı tahmin ağırlıkları setleri kullanılarak hedonik ve ANN modelinin tahminlerinin birleştirilmesiyle oluşturulur. Her modelin ortalama mutlak tahmin hatası (MAFE) hesaplanır ve tüm model çiftleri arasındaki MAFE'deki ortalama fark karşılaştırılır ve test edilir ve üstün model, en düşük ortalama mutlak tahmin hatasına (MAFE) sahip olan modeldir.

SONUÇLAR: Çalışma, YSA- ve hedonik modeller arasında YSA modelinin daha iyi performans gösterdiğini bulmuştur. Bununla birlikte, ANN modeli, bileşen tahmin ağırlıkları olarak kısıtlı WLS tahmini regresyon katsayıları ile oluşturulan kombinasyon tahmininden daha iyi performans göstermiştir. Toplamda, tahminağırlıklarının hesaplanmasında farklı yöntemler kullanılarak yedi birleştirme modeli oluşturulmuştur. Sınırsız birleştirme modelleri, YSA modelinden daha iyi performans gösterdi; kısıtlı birleştirme modelleri YSA modelinden daha iyi performans gösterdi. Çalışma, kısıtlı birleştirme modellerinin en düşük MAFE'lere sahip olduğunu ve üstün modeller olarak kabul edildiğini buluyor.

SONUÇ: Bu çalışma, konut birimlerinin hedonik ve sinir ağı modellerinin tahminlerinden birleştirerek tahmin modellerini başarıyla üretmektedir. Çalışma, kısıtlı WLS regresyonları tarafından oluşturulan ağırlıklarla oluşturulan birleştirici tahmin modelinin genellikle diğer tüm tahmin modellerinden en iyi performansı gösterdiğini bulmuştur. Çalışmamız, tahminleri birleştirmenin İstanbul, Türkiye'deki konut fiyatları tahminlerini iyileştirebileceğini göstermektedir.

Anahtar Kelimeler: Tahmin, hedonik, yapay sinir ağı, birleştirme modeli, İstanbul, Türkiye.

TABLE OF CONTENTS

ABBREVIATIONS

ANN : Artificial neural network

WLS : Weighted least squares

OLS : Ordinary least squares

LMS : Least median of squares

GLS : Generalized least squares

EM : Expectation maximization

GIS : Geographic information system

AQI : Air quality index

SPSS : Statistical package for social scinces

MAFE : Mean absolute forecast error

SD : Standard deviation

TL : Turkish lira

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

A. Study Topic

Housing may be considered as one of the most essential aspects of an individual's life, on a number of levels. Housing plays a major role in regard to the stability and quality of life. Two, a housing unit may also serve as a very important asset in an individual's portfolio of wealth. Providing an accurate prediction on housing unit prices is not only beneficial for the current and future housing unit owners but also for the investors and other real estate market participants (Frew and Jud, 2003). In the past, the housing market was based upon two main principles that will define the value of the property; the first one, a consideration of the initial purchase price of the real estate, and the second was the real estate's selling price. In other words, a consideration of a preceding capital gain or loss as a determining factor of current actual value, lacking for an accepted standard and a certification process (Limsombunchai, Gan and Lee, 2004). Recently, as a result of an expanding real estate markets, the need has arisen for an accurate model of properties price prediction. A more accurate price prediction model will better allow homeowners and potential real estate investors to ascertain both current and future real estate values. Also, having an accurate price forecasting model will enhance the efficiency of predicting the real value of the real estate for the future investment purposes of the real estate market (Limsombunchai et al., 2004; Calhoun, 2003). Recently, many researches have been done focusing the estates' values and aiming to improve the price predictive models by emphasizing the features of properties along with the factors affecting the future expected value like the geographical site, the environment and the housing quality (Schulz and Werwats, 2004; Limsombunchai et al., 2004).

B. Price Prediction Models

A number of various models have been used in order to generate price forecasts. Of these models, the hedonic model has been the most frequently used to predict the properties' prices. Newer models, such as the artificial neural network (ANN), have been used recently in a limited number of studies, with some of these studies equating the performance of neural network model with other models. More recently, combining forecast models has been applied in order to improve forecast accuracy. The contribution of the present analysis is to improve forecast accuracy with the implementation of a combination forecast model. Specifically, we combine hedonic model forecasts with ANN model forecasts of Istanbul housing prices, in an effort to generate forecasts that are more accurate than either component forecast.

1. Hedonic Price Theory

The hedonic term can be described as a dimension in which a consumer can sense the values that are associated with feelings, pleasures and emotions (Limsombunchai et al., 2004). The hedonic theory supposes that goods values including houses can be considered as a collection of specific attributes or characteristics (Griliches, 1971). Also, Rosen (1974) stated that the prices usually reveal the differences in quality depending on how the characteristics of the goods are customized according to the customers' desires.

The origin of the hedonic price theory from the work of Lancaster in 1966, when he proposed that commodities are considered as inputs for the consuming activity which end up in a set of characteristics. Hedonic price theory has been used in estimation of many marketed goods, as a sum of individual goods which cannot be sold separately in the market. The amount of the features associated with these goods is defined as a set of hedonic prices (Rosen, 1974). The advantage of the hedonic models is that they have the ability to control the properties' characteristics, thus permitting to distinguish the impact of marginal change in one of these attributes from the actual property appreciation (Calhoun, 2001). Therefore, the hedonic price can be illustrated as the additional cost when purchasing a house with a slightly better characteristic. These parameters measure the proportional change in prices caused by proportional changes in characteristics (Boardman, Greenberg, Vining and Weimer 2001). The potential values of the property's characteristics are achieved by recognizing the function of the hedonic price according to each characteristic (McMillan, Reid and Gillen, 1980).

Since then, hedonic price theory has been used in valuation of many aspects including residential properties (Witte, Sumka and Erekson, 1979; McMillan et al., 1980; Blomquist and Worley, 1981; Milon, Jonathan and Mulkey, 1984), agricultural commodities (Ethridge and Davis, 1982; Brorsen, Grant and Rister, 1984; Wilson, 1984), and wildlife related resources (Pope and Stoll, 1985; Messonnier and Luzar, 1990). Other applications of the hedonic price theory had included the prediction of environmental improvements' benefits (Freeman, 1979; and McMillan et al., 1980; Blomquist and Worley, 1981).

The hedonic model has been recognized and used in many price prediction settings. However, many issues can affect its performance like the heteroscedasticity, multicollinearity, interactions of the independent variables, outlier data points and non-linearity (Limsombunchai et al., 2004). Therefore, Lenk, Worzala and Silva (1997) and Owen and Howard (1998) suggested the artificial neural network (ANN), as an alternative model that can solve many of these issues.

2. Artificial Neural Network Theory

The concept of ANN models has arisen from the *universal approximation concept*, employed by Hornik, Stinchcombe, and White (1989). Artificial neural networks are able to identify and closely approximate explicitly unknown functional forms, with a high degree of accuracy (Selim, 2009). The concept of universal approximation has led to the use of neural networks in general as non-linear statistical methods that are flexible (i.e., non/semi-parametric, model-free) regression technique not requiring a prior specific theory to work on (Kauko, 2003; Curry, Morgan and Silver, 2002).

Artificial neural network is considered as an artificial intelligence model, and its original design was made to resemble the learning process of the human brain (Limsombunchai, 2004; Morano and Tanjani, 2013). The system of ANN is a complex which consists of a group of primary units, namely the neurons, that is joined in netting structures consisted of interconnecting layers. The complexity of the neural network's structure is depending on the total neurons number and the existing connections (Morano and Tanjani, 2013).

Artificial neural network has three main layers: the input data layer (characteristics of the property), the hidden layer or layers (referred as the "black box”), and the output layer (house price forecast), (Limsombunchai et al., 2004). Neural network is considered as an "interconnected network" consisting of artificial neurons that have the ability to adjust the units' connections weights and strength according to the data externally provided (Stanley, Alastair, Dylan and Patteron, 1998). Within the neural network matrix, every neuron has connecting units to some of its neighbors. The sum of theses weighted input connections will be transformed by a transfer function into output.

3. Previous Literature

Since housing is considered as an essential part of the life quality in any society, therefore, evaluating and comparing different price predicting models have become an area of interest of a large number of researchers around the world. To achieve these objectives, most of these researches were carried out using the hedonic models based on multiple regression analysis. These models are basically considered as appropriate straightforward predictors of the relation between a house price and its different characteristics (Kauko, 2003).

Many empirical researchers specifically have studied the hedonic price model and searched the best methods of its application, along with studying the factors that affect its results. Among these researchers, Adair, Mcgreal, Smyth, Cooper, and Ryley (2000) examined the factors affecting the residential properties in the urban areas of Belfast. Adair and Colleagues found a relative importance of the property characteristics, the property's accessibility and the socio-economic factors on the its price. Furthermore, Frew and Wilson (2000) applied the hedonic model to assess the impact of the property’s location to its value, in Portland, Oregon, and have found a significant relationship between them.

Janssen, So¨derberg and Zhou (2001) suggested that data observations from the real estate markets often contain outliers that usually affect price estimations by having a large influence on least squares estimates. To overcome this issue, they used a robust method in predicting the relationships between the price and income for apartment buildings, by comparing the performance of ordinary least squares (OLS) regression and the robust least median of squares (LMS) regression. Their study found that LMS regression identifies outliers in the data and gives more accurate estimates than OLS, and recommended employing the robust methods in price prediction for more reliable estimates.

Meese and Wallace (2003) studied the impact of the marketing fundamentals on the properties' prices in Paris, France over the period 1986 to 1992, using two prediction methods. The first estimation method was the traditional two-step method where the housing unit price index is predicted and then the predicted index was used to structure the model. The second method was applying the Kalman filter which permits the simultaneous prediction of a dynamic hedonic price model along with the price index and the structured model for the housing prices. The findings from their study showed a successful two estimations strategy and suggested that implementing the traditional methods is relatively easier and can outweigh the small efficiency gained by the simultaneous estimator.

Stevenson (2004) re-investigated the heteroscedasticity issue in hedonic house price models by using data from Boston, United States that has an average age of properties. The study results largely supported the findings in previous literature with the evidence of heteroscedasticity in respect to the house's age. The study found that the generalized least squares (GLS) iterative correction (specified in case of age), will lead to heteroscedasticity elimination at both the aggregate and disaggregate levels.

Schluz and Werwatz (2004) have investigated an empirical price prediction model which extends beyond the traditional hedonic regression model. The model includes applying the observable house characteristics like age of the house and number of bedrooms and bathrooms, as well as unobserved components like mortgage, inflation rates and building permissions. They derived a new equation by combining Kalman filter which estimates the observable common components with the expectation-maximization (EM) algorithm to estimate the unobserved components. The study found that the estimated hedonic coefficients were plausible in sign and magnitude, and this model can be specifically useful in estimating the unobserved house characteristics.

Bin (2004) used a semi-parametric hedonic regression in generating price forecasts and compared its performance with the performances of conventional parametric models in residential areas in North Carolina, the United States. Geographic Information Systems (GIS) data are used in this study for the locational characteristics of the properties. The study found that, in both the in-sample and out-of-sample price forecasts, the performance of the semi-parametric model outperforms the one of the parametric models. This finding indicates that the semi-parametric model can be used as a useful tool for housing prices prediction.

Furthermore, Filho and Bin (2005) modeled a non-parametric hedonic regression model for housing prices prediction. Forecasting was generated by a back-fitting procedure combined with a local polynomial predictor in order to avert the issues of an unrestricted non-parametric forecast. They used a novel plug-in method in choosing the bandwidths, as this method can decrease the mean square error in the regression analysis. Then, they compared the findings from the non-parametric model to another parametric models, and found that the non-parametric model has the superior performance.

Fan, Ong and Koh (2006) applied the hedonic-based decision tree approach to study the relation between the house price and its characteristics. They study used data from the Singapore resale public housing market, and found that this approach is a successful technique.

With all these previous studies applying the hedonic modellings in price prediction, issues have been faced with these techniques like nonlinearity, outliers, fuzziness, discontinuity, spatial and other kind of dependence between observations (Kauko, 2003). Therefore, some plausible alternatives were introduced. One of them is the use of artificial neural networks, which are considered as suitable models to deal with these issues. It was suggested that this alternative predicting model can be useful in the appraisal practice (Worzala, Lenk and Silva, 1995).

In recent years, many researches have been conducted in order to evaluate the accuracy of the ANN model in forecasts and to compare its performance with other forecasting models.

One of these researches, Din, Hoesli and Bender (2001) compared various real estate forecasting models including the artificial neural network model with the standard hedonic linear regression model. The study showed that the neural network models, which are non-linear *per se*, show generally similar price forecasts to the traditional hedonic model.

Kauko, Hooimeijer, and Hakfoort (2002) evaluated ANN modeling in the valuation of housing market in Helsinki, Finland. The study demonstrated that it is possible to detect different dimensions of housing market by using uncovering patterns in the data set. Also, the study successfully applied the two classification techniques of the neural network model; the learning vector quantization and the self-organizing map in the housing price forecasting.

Curry, Morgan and Silver (2002) investigated the potential of applying a neural network approach to the analysis of hedonic linear regressions**,** where the price is dependent on the quality characteristics. The results of the study found a relatively marginal improvement on linear formulations, and that ANN can be considered as a useful tool of specification testing. Hence, their results supported the linear formulation as an adequate approximation.

There was a number of studies which compared the predictive power of the more traditional hedonic models with the newer artificial neural network models. One of these studies, a study by Kauko (2003) which studied and compared the pros and cons of both the neural network and the hedonic models in evaluating properties. The review showed some examples, like the effect of the environment, the property location and other factors on the housing price. Kauko (2003) stated that the neural network can be considered as a relevant tool for mass appraisal, and it can be applicable in situations where a large number of sites or houses or sites have to be valued quickly and within a predefined range of error tolerance.

Limsombunchai, Gan and Lee (2004) applied and compared the hedonic and ANN models by generating housing price forecasts in Christchurch, New Zealand. House characteristics including the house type, age of the house, its size, number of bedrooms, number of bathrooms, number of garages and geographical location were all considered. Their study showed that ANN models outperformed the hedonic price models, and it can overcome the data patterns issues related to the hedonic model.

Selim (2009) also finds that an ANN model does a better job of relating a data input set of housing attributes to housing price in Turkey, compared to a hedonic regression model.

Peterson and Flanagan (2009) also compared the property value prediction performance of the neural network model with the hedonic linear regression model, employing a sizeable US data set covering 1999-2005. They found that ANN performance was better than the hedonic by generating a significantly lower price forecast errors in the out-of-sample forecasts, and that the multi-layered artificial neural networks are capable of modeling complex nonlinearities. The study agreed with the previous studies in that ANN is a better alternative to hedonic models because the parameter evaluation in neural network does not rely on the regressor matrix rank.

Morano and Tajani (2013) used two forecasting models in order to obtain an effective tool for estimating market value of the bare ownership in Italy and to compare the respective performances of these models. One of the prediction models used was based on the hedonic prices theory and the other model was the artificial neural networks (ANN). They found that choosing the best estimation method depends on the main purpose of the valuation. If the purpose is the exclusive prices forecasts of the bare ownership, then ANN model is the best option. While the hedonic price model should be used if the purpose is to both prices prediction and investigating the factors contributing to their formation, as the hedonic models has a simpler approach in achieving these outcomes.

C. Combination Forecasting Model

As is widely known, the central idea of a combining model is to form a weighted average of forecasts generated by two or more different models. The objective is to create a forecast that is more accurate than any of the individual component model forecasts. The initial and conventional approach is to combine one or more structural model forecasts with one or more time-series model forecasts of a given forecast variable. The idea is to offer a *structural* explanation of the variance of the forecast variable, along with a *time series* explanation of that part of the variance that cannot be explained by the structural model [or models] (see, for example, Bischoff, 1989).

With regard to the present analysis, of particular interest is the Terregrossa (1999) study, which pivots from the conventional approach (described above) by first hypothesizing and then demonstrating that combining forecasts generated by two (or more) individual models may be effective if each model contributes independent information with regard to movement of the forecast variable, regardless of the types of model employed. In other words, a successful combining model does not need to exclusively combine a structural model forecast with a time series model forecast. The present analysis follows this tack in forming combination forecasts of housing prices. (This approach has been successfully applied by others, including Gupta, Kabundi and Miller (2011), as indicated below.)

A test of independent information may be achieved with an *in sample* regression of realized values against component-model forecasts of the target variable. If the regression coefficients are all nonzero and statistically significant, then this would imply that each of the individual component forecasts contain independent information. In this case, a weighted average combination forecast can be formed with the estimated *in sample* regression coefficients serving as *component forecast weights* (see, for example, Terregrossa, 1999).

Early work in the area of combination forecasting was in regard to macroeconomic variables. For examples, see Bischoff (1989), Fair and Shiller (1990), Moreno and López (2007).

A large number of previous studies have experimented with combination forecasts regarding financial variables. See for example, Guerard (1987), Newbold, Zumwalt, and Kannan (1987), Lobo (1991, 1992), Terregrossa (1999, 2005), Loh (2005), and Kumar and Patel (2010).

Lately, the combining model technique has been to be applied to other arenas with different types of forecast variables. For example, Wu, Zhou, Chen, and Ye (2015) successfully applied the combination forecasting method in improving hydrological operational predications (i.e. flood forecasting). Also, the combining method has been applied in the prediction of energy consumption (Liu, Moreno and Garcíac, 2016), in solar radiation forecasting (Heng, Wang, Xiao and Lu, 2017), in tourism demand forecasting (Jun, Yuyan, Lingyu and Peng, 2018), in electrical load forecasting (Wang, Wang and Xu, 2019), in forecasting wind speed (Liu, Zhang, Chen and Wang, 2018; Niu and Wang, 2019; Liu, Jiang, Zhang and Niu, 2020) and in forecasting the air quality index (AQI) (Song and Fu, 2020).

To date, the combination forecasting method has also been applied in a limited number of empirical studies that utilize real estate pricing models. For example, Bradley, Gordon and Mcmanus (2003); Fleming and Kuo (2007); Drought and McDonald (2011); Gupta, Kabundi and Miller (2011); and Cabrera, Wang and Yang (2011) have each experimented with some type of combining model to forecast real estate entity values.

The present analysis differs from the other studies cited above (that form combination forecasts of housing prices), in that we explore the issue of applying constraints on the linear regression model to generate weighted average combinations of *housing price* forecasts.

D. Study Objectives

The objectives of our study are:

- To investigate the factors affecting the housing unit price, including the house age, the land size, number of bedrooms and bathrooms and the geographical location.
- To apply the hedonic regression and the artificial neural network models to generate housing units' price forecasts.
- To compare the predictive power of the hedonic regression model with the artificial neural network model.
- To improve the forecast accuracy by generating a combining model from averaging both the weighted hedonic and the artificial neural network forecasts.
- To compare the predictive power of the combining model with both the forecasts of hedonic and ANN in order to suggest the most accurate method for the future predictions of house prices.

II. MATERIAL AND METHODS

A. Study Area and Data Sources

The study covered the real estate housing market in the European area of Istanbul, Turkey. A sample of 100 housing units in Istanbul were randomly selected through retrieving through various real estate websites on the internet. The data collection was taken place from the period of 1st of May 2019 to 15th of July 2019. The following real estate websites were used to collect the study data:

- Istanbul Real Estate.
- Istanbul Property World.
- Property Turkey.
- Turkey Expert.
- Istanbul Homes.
- Turkey Homes.

The study sample was distributed between the residential areas and the city centre and other parts of the city. Due to the fact that majority of the houses in Istanbul are in the form of residential buildings, the study only included apartments in the data collection. For each apartment, the following information were taken; the sale price in Turkish Lira (each 1 USD = 5.71 TL in 15th of July 2019), the geographical location, the land size in square meter, the number of bedrooms and bathrooms, the property's age in years, and the apartment floor within the building.

B. The Study Sample

According to the standard analytical practice from the study of Limsombunchai et al. (2004), the study sample was divided randomly into two sets; the ‘‘estimation set’’ (*in sample*), and the ‘‘forecasting set’’ (*out of sample*) as known in regression analysis literature. Or the ‘‘training set’’ (*in sample*) and the ‘‘production set’’ (*out of sample*) as known in neural network literature. The (in sample) set contains 80% of the test data and the (out of sample) set contains the remaining 20% of data.

Data in the in sample is used to test the effects of house characteristics on its price and to generate housing units’ price forecasts. While data in the out of sample is used to determine the model accuracy by calculating forecast errors (discussed later).

C. Data Analysis

This study evaluated both the Hedonic and the Artificial Neural Network models. The housing unit price was the dependant variable and the other house characteristics including the geographical location, land size, age of the house, number of bedrooms and bathrooms and the building floor, were the independent variables. In previous literature (Selim, 2009; Limsombunchai et al., 2004; Halvorsen and Palmquist, 1980), the semi-logarithmic form was the most commonly used functional form of the hedonic model. This form fits the data very well, and it is preferred to be used. The resulted estimated coefficients can be interpreted as being proportional to the property’s price that is directly correlated to its characteristics. Thus, in this study, we used the natural logarithm of the housing unit price as the dependant variable.

All the information obtained were noted down in a Microsoft Excel sheath, and the data statistical analysis in this study were performed using the Statistical Package for the Social Sciences (SPSS) software (IBM, Version 24) for Microsoft Windows. The housing unit characteristics were represented as numerical values in the statistical analysis software. For example, giving numerical codes (such as 1, 2, 3…etc.) for the properties' locations, where each number represented a particular neighbourhood in Istanbul. And this was applied for the rest of the house characteristics. All the information concerning the details of the codes for the housing unit characteristics were saved in a separate sheath for further reference and evaluation in the results and discussion sections. *P*-values of less than 0.05 are considered as statistically significant.

1. The hedonic method

The hedonic model is carried out first by performing the linear regression analysis on a random 80% of the housing units (*in sample*). The housing units' log prices (the dependant variable) are regressed against their characteristics (the independent variables).

Implicitly, the model for the hedonic price function (*f*) is generated as follow:

$$Price = f\,(\mathrm{L, S, BD, BA, A, FL})$$

Equation 1

where,

(L) = the location

(S) = the land size in square meters (m^2)

(BD) = number of bedrooms

(BA) = number of bathrooms

(A) = house age in years

(FL) = apartment floor within the building.

The results of the regression analysis on the (in sample) data will generate estimated coefficients for each of the house characteristics which will identify the relative contribution of the housing unit characteristics to its price. Then, these coefficient estimates will be used to generate the forecasts for the remaining 20% housing units (*out of sample*) forecasts, as follows;

$$\mathrm{P_y} = \beta_0 + \beta_1 \mathrm{L} + \beta_2 S + \beta_3\, \mathrm{BD} + \beta_4 \mathrm{BA} + \beta_5 \mathrm{A} + \beta_6 \mathrm{FL}$$

Equation 2

Where,

(P_y) is the predicted price

(β_0) is the constant coefficient

($\beta_1 - \beta_6$) are the estimated coefficients for each house attributes

The predicted (out of sample) forecasts will be used to investigate the performance of the hedonic model by calculating the mean absolute error and compare it to the other predicting models' performances.

The relationship between a housing unit's price and its characteristics is also tested by the hedonic analysis. A correlation coefficients matrix between the housing units' prices and their characteristics is generated. For example, the estimated coefficient testing between the housing unit size and its price, or between the number of bedrooms and the housing unit price, and so on for the rest of the variables.

In the present analysis, as a result of using various variables, the issue of heteroscedasticity may arise. Heteroscedasticity (absence of homoscedasticity) is defined as data that has unequal variance with a non-constant spread of dots along the regression line (Pinder, 2017).

In a number of studies regarding the hedonic housing models, the property's age has been found to be the primary cause of heteroscedasticity (Goodman and Thibodeau, 1995 & 1997; Stevenson, 2004). Fletcher, Gallimore and Mangan (2000) found that the property's external area also causes heteroscedasticity. The presence of heteroscedasticity can affect the regression analysis results by producing high errors (Stevenson, 2004).

To overcome this issue, a test for the presence of heteroscedasticity is run by plotting the standardized predicted value (X axis) against the residual value (Y axis). The resulted (residual plot) graph will show whether the data are spreading along the regression line (homoscedastic data) or not (heteroscedastic data).

If the heteroscedasticity presence is confirmed, the Weighted Least Square (WLS) technique will be used in the regression analysis instead of the Ordinary Least Square (OLS) technique, since OLS can produce high standard errors because of the heteroscedasticity of the variables (Stevenson, 2004).

2. The ANN method

For the artificial neural network (ANN) model, relative contribution factors (for each of the housing-unit explanatory variables) to the housing unit price were identified using the multilayer perception neural network analysis with one hidden layer on a random 80% of the studied sample (in sample). The information resulted from the (in sample) analysis will be used to generate the housing forecasts in the (out of sample) data.

The neural network application method is a similar process to the hedonic price model, where the logarithm house price is the dependant variable and the housing-unit set of attributes and characteristics (location; size; house age; number of bedrooms and bathrooms; building floor) were the explanatory variables. However, the neural network model differs from the hedonic model in that, for a specific input (set of attributes and characteristics), an output (predicted house price) is directly generated from the model (see Limsombunchai et al., 2004). Afterward, a comparison between ANN price forecasts with the actual housing unit price, and with the hedonic forecasts to investigate the ANN model performance.

3. Combining forecasts method

First, a test of independent information in the hedonic and ANN models is done by regressing the actual, realized housing prices (from the *in sample* values) against the housing price forecasts (*in sample*) generated separately by the hedonic and the artificial neural network models, in the following fashion:

$$\mathrm{A} = \alpha + \beta(\mathrm{hedonic}) + \gamma(\mathrm{ANN})$$

Equation 3

where,

A = actual (realized) housing unit price (*in sample*);

α = constant term;

β = coefficient estimate of the hedonic model;

γ = coefficient estimate of the ANN model;

Hedonic refers to *in sample* hedonic forecasts;

ANN refers to *in sample* ANN forecasts

If the regression coefficients (β and γ) are both nonzero and statistically significant, then this would indicate that both models contain independent information and can be useful in generating combining forecasts. Indeed, this is the case in the present analysis. The estimated coefficients from the *in sample* regression analysis are then used as weights to form weighted average combinations of the *out of sample* forecasts generated by the *hedonic* and ANN models.

Component forecast weights are generated in the present study by using *weighted least squares* (WLS), to overcome the issue of heteroscedasticity. In a similar vein as the studies of Guerard (1987), Lobo (1991) and Terregrossa (2005), both restricted and unrestricted regressions are run as follows: With a constant term; without a constant term; with the regression coefficients unconstrained; with the regression coefficients constrained to sum to one. In this way, four different versions of the combining model are generated: i) With an estimated intercept term included, and the regression coefficients unconstrained; ii) Without an estimated intercept term included, and the estimated regression parameters unrestricted; iii) With an estimated intercept term included, and the estimated regression parameters restricted to add up to one; iv) Without an estimated intercept term included, and the estimated regression parameters restricted to add up to one. Each set of estimated regression parameters serve to form a weighted average of the component model forecasts.

The intercept term is included in methods i and iii, respectively, to capture exogenous or unexpected macroeconomic disturbances that may impact housing unit prices.

The idea behind the restricted regression technique is to create greater efficiency of the *in sample* estimated regression coefficients, to enhance the *out of sample* forecast accuracy.

Weighted averages of the component model forecasts (hedonic and ANN), estimated with *out of sample* data, are formed as follows:

$$F_c = w_1\,(\text{hedonic}) + w_2\,(\text{ANN})$$

Equation 4

Where,

F_c: weighted average (combination) forecast;

Hedonic: the hedonic model component forecast;

ANN: the artificial neural network (ANN) component model forecast;

w_1, w_2 = the proportional weights which are the estimated regression parameters from the independent information test regression.

Then, to provide context, combinations are formed using simple weighted averages as follows:

- $w_1 = 0.5$, and $w_2 = 0.5$;
- $w_1 = 0.25$, and $w_2 = 0.75$;
- $w_1 = 0.75$, and $w_2 = 0.75$.

Finally, all these seven combining models' forecasts will be generated and compared to each other to determine the superior method.

4. Comparing the models' performances

To determine superior performance, *mean absolute forecasting errors* (MAFE) of each model (component- and combining-) are calculated and compared. MAFE is calculated by first measuring the *absolute value* of the difference between the actual housing unit price and the predicted housing price. The average (mean) of these absolute values is the MAFE:

$$\mathrm{MAFE} = \frac{1}{n} \sum \left| \mathrm{P} - \mathrm{P_y} \right|$$

Equation 5

Where *P* is the actual (*realized*) *housing unit* price, P_y is the predicted price and *n* is data number.

The *Wilcoxon signed rank test* is employed as nonparametric exam of the average difference between *mean forecast errors* for all pairs of the estimated models (*hedonic, ANN, and combining* models) and across all *housing units* in the *out of sample* data set.

III. RESULTS

A. Descriptive analysis

A total of 100 apartments information from 19 different neighbourhoods in Istanbul, were included in the data analysis, see Table 1 for details of the residential areas.

The mean ± standard deviation (SD) of the house price was 83570802.13 ± 23627704.05 Turkish Lira (TL), ranged from 110000 to 2161438408 TL. The mean land size ± SD was 116.92 ± 8.3 square meter (m2), ranged from 42 to 677 m2. And for the age of the houses, the mean age ± SD was 2.51± years (range: 0 – 20 years), see Table 2.

For number of bedrooms and bathrooms, majority of the apartment have one bedroom (38%) and one bathroom (73%), see Table 3 for more details.

Table 1 Locations of the 100 apartments sample in Istanbul, Turkey.

Location	*Frequency*	*Percent*
Sisli	9	9%
Beyoglu	10	10%
Bosfor	2	2%
Tarlabaci	1	1%
Cihangir	2	2%
Avcilar	10	10%
Beylikdozo	12	12%
Basaksehir	5	5%
Beyukcekmece	6	6%
Esenyurt	18	18%
Bakirkoy	3	3%
Fatih	8	8%
Kavakli	3	3%
Kucukcemece	3	3%
Zencirlikoy	2	2%
Zeytinburnu	3	3%
Pendik	1	1%
Maslak	1	1%
Euyp	1	1%
Total	100	100%

Table 2 descriptive analysis of the characteristics of 100 apartments in Istanbul, Turkey.

Variables	*Mean*	*Standard deviation*	*Range*
House price (TL)	83570802.13	236277040.5	110000 -2161438408
Land size (m^2)	116.92	82.999	42 – 677
Age of building (years)	2.51	4.984	0 – 20
Floor	5.73	5.059	-1 – 25

Table 3 Number of bedrooms and bathrooms in a sample of 100 apartments in Istanbul, Turkey.

Bedrooms	Frequency	Percent	Bathrooms	Frequency	Percent
1	38	38%	1	73	73%
2	31	31%	2	23	23%
3	21	21%	3	3	3%
4	4	4%	4	0	0
5	3	3%	5	1	1%
6	2	2%	6	0	0
7	1	1%	7	0	0
Total	100	100.0	Total	100	100%

B. Hedonic Regression Analysis

Figure 1 shows that the sample data in the present study has unequal variance (heteroscedastic data) as the dots in the scatterplot are not spreading along the regression line.

Thus, the Weighted Least Square (WLS) technique is used in the regression analysis instead of the Ordinary Least Square (OLS) technique due to the presence of heteroscedasticity in the sample data.

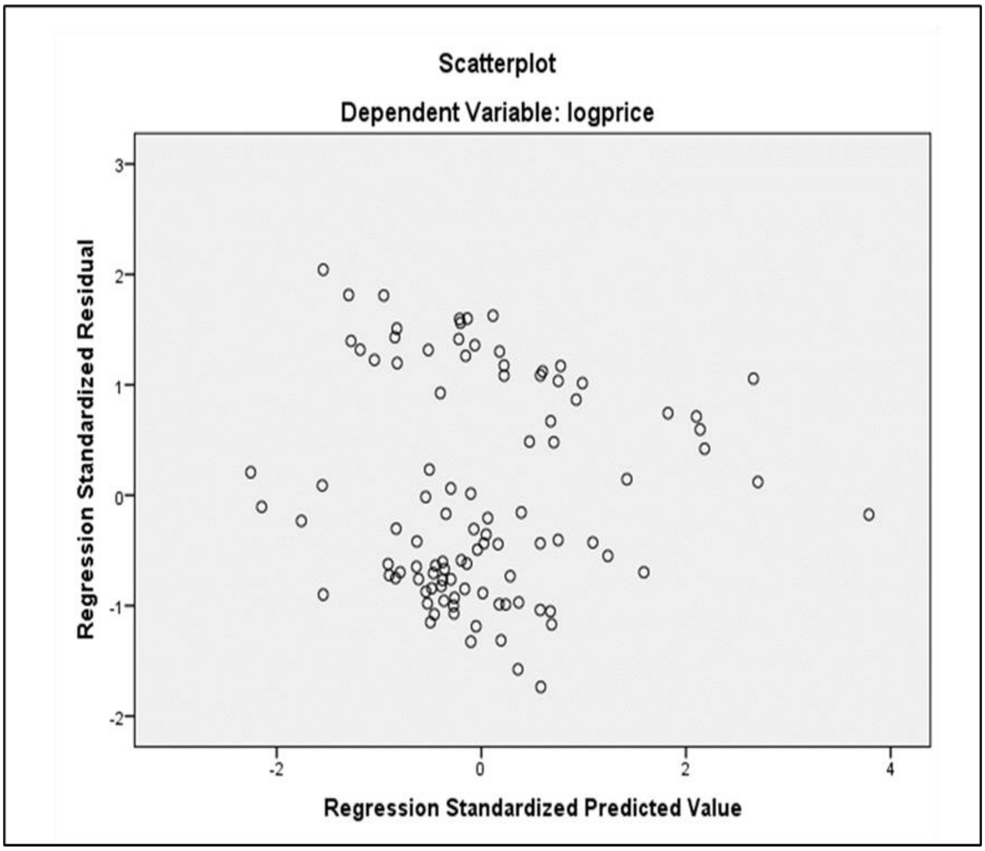


Figure 1 Residual plot of the study data shows that data are not spreading along the regression line (heteroscedastic data).

The correlation coefficients from equation 1 using the linear regression analysis are shown in Table 4. Majority of the coefficients are statistically significant.

The correlation coefficients show that larger houses with more bedrooms and bathrooms have higher prices, with correlation coefficients of 0.274, 0.462 and 0.375 respectively. Also, apartments in higher floors within the building are priced higher, the correlation coefficient is 0.172. The results also showed high correlation coefficients between bedrooms and bathrooms numbers (1.00), the bedrooms' number and the land size (0.799) and between the number of bathrooms and the land size (0.742). However, these high coefficients are neither necessary nor important in determining the houses prices. The correlation coefficient between the house price and its location is 0.096. This low coefficient indicates a low degree of linear relation.

On the other hand, the correlation coefficient between the house price and the house age has a negative sign (-0.109). This indicates a negative relation between a house and its age, in which an older house has a lower price and vice versa.

The hedonic pricing model results using Weight least square (WLS) analysis on the estimation set (In sample) in Table 5 shows positive coefficients for the housing unit's location, the number of bedrooms and the floor (0.027, 0637, 0.421 and 0.046) respectively. While the age of the housing unit and its land size shows negative coefficients (-0.034 and -0.007) respectively.

Table 9 shows the forecasts results from the hedonic model in comparison to the actual price and the ANN model forecasts in (out of sample) set.

Table 4 Correlation coefficient matrix

	Price	Location	Size	Bedrooms	Bathrooms	Age	Floor
Price	1.00	.096	.274**	.462**	.375**	-.109	.172
Location	.096	1.000	-.182	-.156	-.265**	-.230*	-.037
Size	.274**	-.182	1.000	.799**	.724**	-.059	-.062
Bedrooms	.462**	-.156	.799**	1.00	.727**	-.040	-.082
Bathrooms	.375**	-.265**	.724**	.727**	1.00	-.076	-.047
Age	-.109	-.230*	-.059	-.040	-.076	1.00	-.115
Floor	.172	-.037	-.062	-.082	-.047	-.115	1.000

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Table 5 The Hedonic regression model Weight least square analysis (In sample).

Variables	Coefficient	t-value	p-value
Constant (C)	5.207	16.684	0.000**
Location (L)	0.027	1.167	0.247
Size (S)	-0.007	-6.953	0.000**
Bedrooms (BD)	0.637	5.530	0.000**
Bathrooms (BA)	0.421	6.620	0.000**
Age (A)	-0.034	-1.605	0.113
Floor (F)	0.046	3.237	0.002*
n = 80			

*Coefficient is significant at the 0.05 level.

* Coefficient is significant at the 0.01 level.

C. Artificial neural network analysis

The neural network analysis by the multilayer perception with one hidden layer was done on the (in sample) set. The log price of the houses was used as the dependant variable in order to eliminate data skewness and outliers.

The relative contribution factors of the artificial neural network analysis (the relative importance of inputs) are shown in Table 6 and Figure 2. The neural network results demonstrate that the land size plays a large role in the house price determination more than the other factors, relative importance (100%) and relative contribution (0.258). Also, the number of bedrooms and the building floor have high impacts on the house price compared to the other house characteristics, relative importance (81.8% and 62%) and relative contributions (0.209 and 0.160), respectively.

Table 9 shows the housing units' forecasts results from the ANN model in comparison to the actual price and the hedonic model forecasts in (out of sample) set.

Table 6 Neural network relative contribution factors.

Factors	Relative contribution	Importance
Location	0.112	43.5%
Land size	0.258	100%
Bedrooms	0.209	81.1%
Bathrooms	0.124	48%
Age	0.138	53.4%
Floor	0.160	62%
n =80		

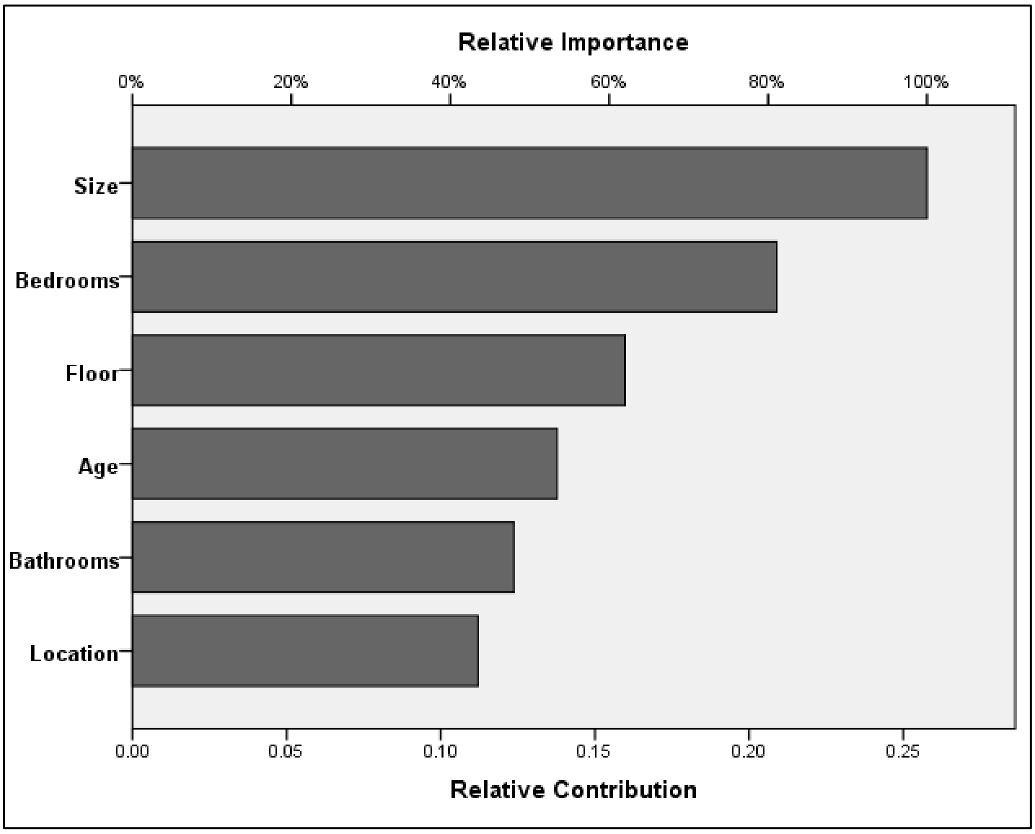


Figure 2 Relative contribution and importance of house characteristics on the house price by ANN.

D. The combining Models

Tests for the independent information were done to estimate the regression coefficients (which serve as forecast weights) from the (in sample) forecasts, in four different methods (as discussed previously in section 2.3.3).

The four methods of weighted-least-squares (WLS) regression tests result in the following coefficients, see Table 7:

First, the unrestricted WLS regression with the constant term generates positive β and γ coefficients, (β = 0.415) and (γ = 0.780). However, the constant term coefficient was negative (-1.267).

Second, the unrestricted WLS regression with the constant term suppressed also results in positive β and γ parameters, (β = 0.459) and (γ = 1.067).

Third, the WLS estimated regression inclusive of an intercept term and with the estimated regression parameters restricted to add up to one (Model 3) generates positive parameters (β = 0.028) and (γ = 0.972).

Finally, the WLS regression without an intercept term and with the estimated regression parameters restricted to add up to one also generates positive parameters (β = 0.006) and (γ = 0.994).

The combination forecasting model are then generated by forming a weighted average of the hedonic and ANN forecasts according to equation 4.

Seven sets of weights are used in turn to generate seven combining forecasts models. Models 1, 2, 3 and 4 employ forecast weights generated by WLS regressions; whereas Models 5, 6 and 7 are formed using simple weighted averages.

Table 8 shows the forecasts in log form from the (out of sample) that are generated by the combining models.

Table 7 Weights results from WLS regression (in sample) in four different methods according to the following equation;

$$A = \alpha + \beta(hedonic) + \gamma(ANN)$$

	α	B	γ
Unrestricted WLS			
Estimated coefficients	-1.267	0.415	0.780
standard error	0.273	0.070	0.086
t-statistic	-4.636**	5.953**	9.095**
Unrestricted WLS with the constant suppressed			
Estimated coefficients	NC	0.459	1.067
standard error		0.07	0.001
t-statistic		5.910**	1309.55**
Constrained WLS			
Estimated coefficients	-0.001	0.028	0.972
standard error	0.595	0.108	0.081
t-statistic	-9.33*	5.019*	3718.515*
Restricted WLS with the constant suppressed			
Estimated coefficients	NC	0.006	0.994
standard error		0.05	0.077
t-statistic		4.2*	12.011**

* test significant at 0.05 level.

** test significant at 0.000 level.

E. Models comparison

1. Hedonic vs ANN forecasts

First, comparing the hedonic regression model and the artificial neural network model performances in forecasting the housing units' prices; the predicted prices generated by both models and the actual prices are shown in Table 9 and illustrated graphically in Figure 3. As seen from the table and the figures, the forecasts from the ANN are closer to the actual sale prices than those of the hedonic model.

Furthermore, the hedonic and neural network forecasts in the (out of sample set) are compared in term of the mean absolute forecast error (MAFE). The results show that the artificial neural network forecasts' MAFE (0.157945) lower is than the hedonic model's MAFE (0.740237), see Table 10 and Table 11.

Wilcoxon signed rank test of the mean difference between the average absolute forecast errors of the hedonic and ANN models is statistically significant (p = 0.002).

Table 8 Predicted housing units' prices in (log form) by hedonic and ANN and models (Out of sample forecasts).

House	Actual prices	Hedonic forecasts	ANN forecasts
1	6.373923	6.386624	6.245102
2	6.965999	6.95126	6.814697
3	6.452075	6.454495	6.145102
4	5.60206	6.696627	5.519844
5	8.149426	6.894628	8.185957
6	8.417989	7.143494	8.356686
7	8.223672	7.440532	7.577866
8	8.191488	8.535972	7.885387
9	8.518454	7.787051	8.237441
10	7.934877	7.028458	7.95461
11	8.031787	6.330481	7.558797
12	7.568906	7.70411	7.771727
13	7.554666	7.670106	7.650467
14	7.679604	6.994265	7.604218
15	5.816573	6.767045	5.830473
16	5.852785	7.284653	5.879501
17	5.895754	6.298843	5.881533
18	6.213757	7.658636	6.204415
19	5.716003	6.145794	5.552907
20	5.39794	6.502776	5.509376

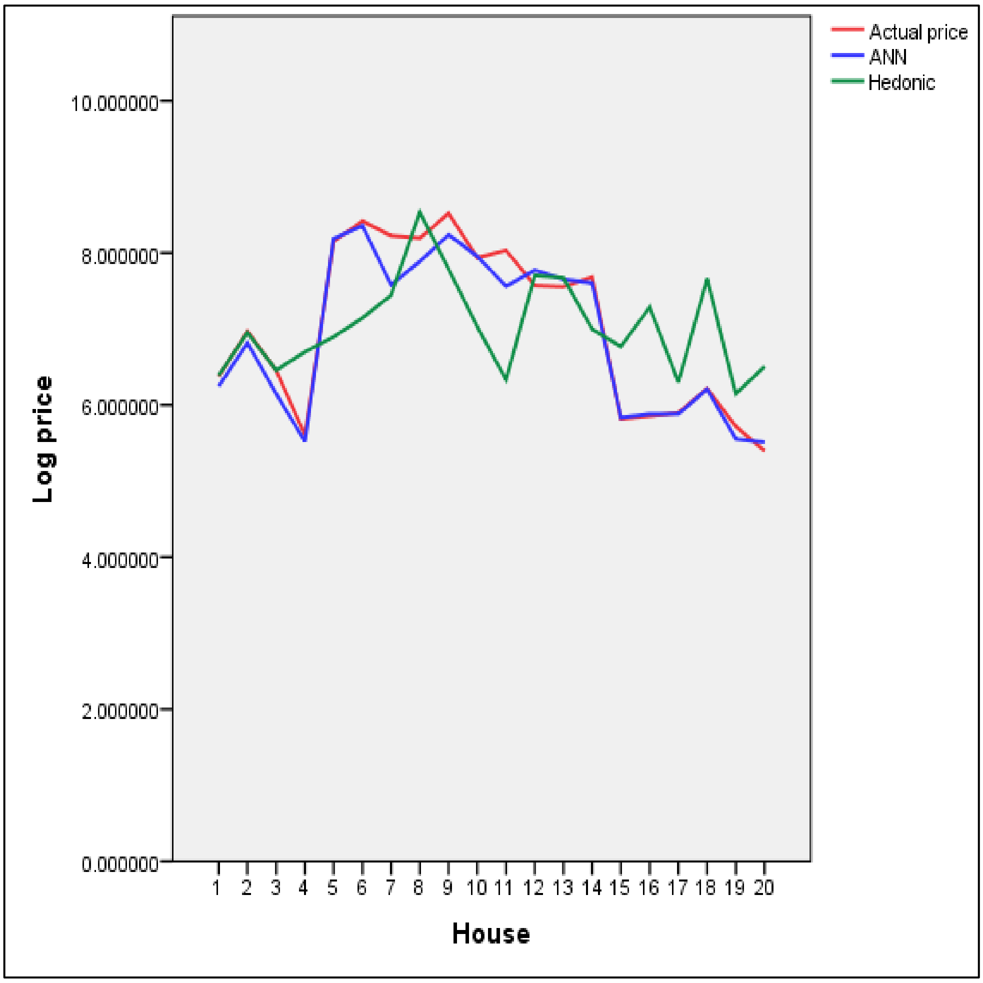


Figure 3 Comparison between actual and predicted house prices obtained from the hedonic and ANN in log form (Out of sample forecast).

Note: the forecasts by the ANN model are very close to the actual prices.

2. Combining models' performances

As mentioned above, the artificial neural network forecasts are found to be superior to the hedonic model forecasts. So that the performance of the ANN model is compared to each of the seven combining models.

The housing units' prices forecasts generated by these combining models are compared graphically in Figure 4.

Also, the mean forecast error (MAFE) of each model is calculated, and the average variation in mean forecast error (MAFE) between hedonic and ANN forecasts, and all the combining forecasts, respectively is calculated and compared, see Table 10 and Table 11.

The constrained combining model (Model 4) shows the least MAFE (0.019785) followed by Model 3 (MAFE = 0.026338), out of compared to the rest of combination forecasting models. The MAFEs of the unrestricted Model 1 is (0.197057) and for Model 2 (0.270502). While the combining models (Model 5, Model 6 and Model 7) with the following weights combination; (0.5 and 0.5, 0.25 and 0.75, 0.75 and 0.25) have MAFEs of (0.369436, 0.184036, 0.554837) respectively.

The comparison between neural network mean absolute error and those of the combining models shows that the constrained Model 3 and Model 4 perform better than the ANN model and the rest of the unrestricted models. While ANN model performs better than Model 1, Model 2, Model 5, Model 6 and Model 7.

Wilcoxon signed rank tests for the mean difference between MAFEs show statistically significant difference (lower MAFEs) of the Model 3 and Model 4 compared to ANN model (p = 0.001 and 0.00) respectively. While Wilcoxon signed rank tests of the mean difference between hedonic and the combining model MAFEs show that hedonic forecasts have statistically significant higher MAFEs than Models 1, 3, 4, 5,6 and 7 respectively (p = 0.00).

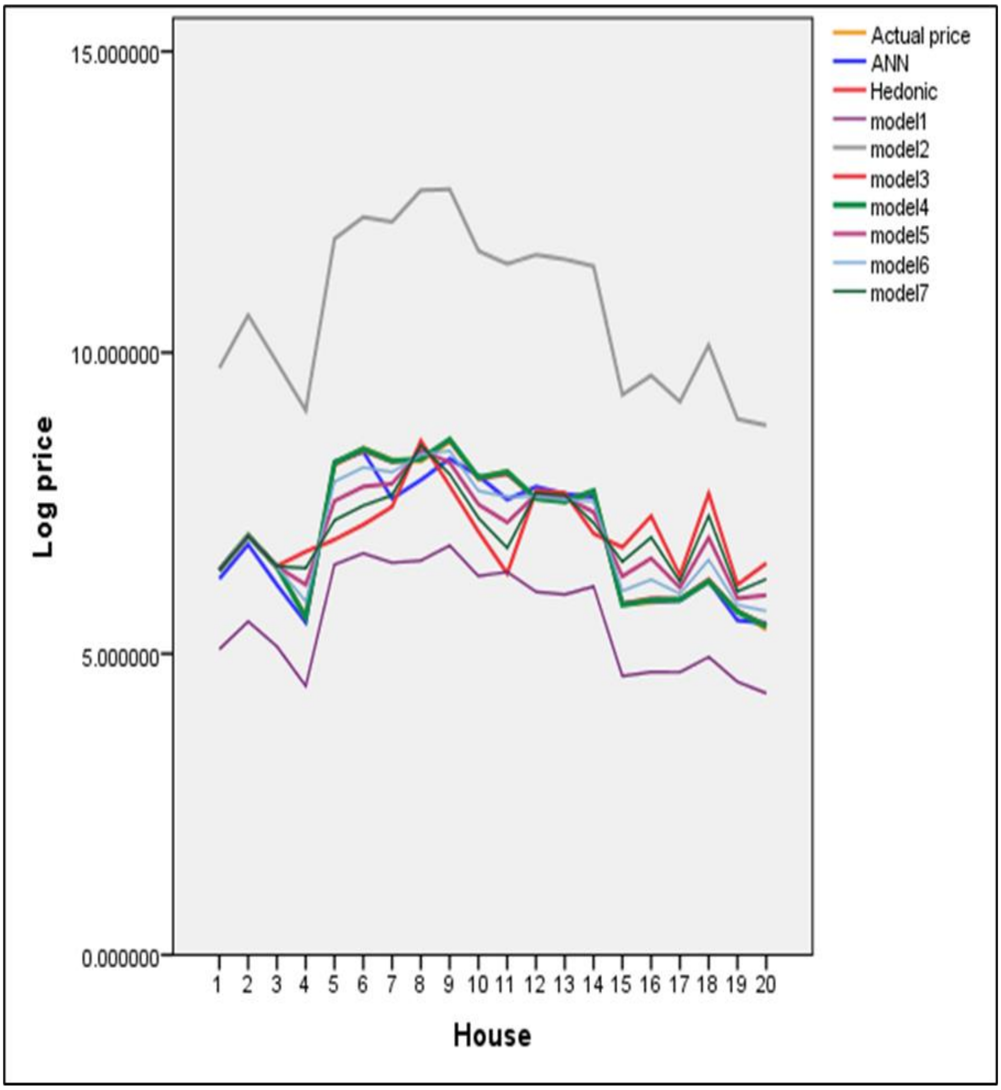


Figure 4: Comparison of the actual housing prices and the forecasts generated by the ANN models and the seven combining forecasts models.

Note: model 3 and model 4 forecasts are very close to the actual price.

Table 9 Mean Absolute Forecast Error (MAFE) summary (in percentage).

Forecast model	*MAFE*
Model A	0.740237
Model B	0.157945
Model 1	0.197057
Model 2	0.270502
Model 3	0.026338
Model 4	0.019785
Model 5	0.369436
Model 6	0.184036
Model 7	0.554837

Notes:

Model A: hedonic forecasting method;

Model B: neural network forecasting method;

Model 1: weighted average formed with estimated parameters of unrestricted WLS regression including an intercept term;

Model 2: weighted average formed with estimated parameters of unrestricted WLS regression without an intercept term;

Model 3: weighted average formed with estimated parameters of restricted WLS regression including an intercept term;

Model 4: weighted average formed with estimated parameters of restricted WLS regression without an intercept term;

Model 5: combining model generated by the following weights [0.5 (Hedonic) + 0.5 (ANN)];

Model 6: combining model generated by the following weights [0.25 (Hedonic) + 0.75 (ANN)];

Model 7: combining model generated by the following weights [0.75 (Hedonic) + 0.25 (ANN)].

Table 10: Average variation in Mean Forecast Error (MAFE) among pairs of models

Forecast model	*Average variation in MAFE*
Model A – Model B	0.582292*
Model 1 – Model 3	0.170719**
Model 2 – Model 4	0.250717**
Model A – Model 1	0.54318*
Model A – Model 2	0.469735*
Model A – Model 3	0.713899**
Model A – Model 4	0.720452**
Model A – Model 5	0.370801**
Model A – Model 6	0.556201**
Model A – Model 7	0.18540**
Model B – Model 1	- 0.039112**
Model B – Model 2	-0.112557**
Model B – Model 3	0.131607*
Model B – Model 4	0.13816**
Model B – Model 5	-0.211491*
Model B – Model 6	-0.026091
Model B – Model 7	-0.396892**

Model A: hedonic forecasting method;

Model B: neural network forecasting method;

Model 1: weighted average formed with estimated parameters of unrestricted WLS regression including an intercept term;

Model 2: weighted average formed with estimated parameters of unrestricted WLS regression without an intercept term;

Model 3: weighted average formed with estimated parameters of restricted WLS regression including an intercept term;

Model 4: weighted average formed with estimated parameters of restricted WLS regression without an intercept term;

Model 5: combining model generated by the following weights [0.5 (Hedonic) + 0.5 (ANN)];

Model 6: combining model generated by the following weights [0.25 (Hedonic) + 0.75 (ANN)];

Model 7: combining model generated by the following weights [0.75 (Hedonic) + 0.25 (ANN)].

* Wilcoxon Signed rank test significant at 0.05 level.

** Wilcoxon Signed rank test significant at 0.000 level.

IV. **DISCUSSON**

A. General Discussion

This study was carried out on a sample of 100 apartments in Istanbul, Turkey to evaluate the performances of the hedonic regression and the artificial neural network models in predicting housing units' prices. Also, the current study generated a combining model from averaging the weighted forecasts of both the hedonic and ANN in order, in an attempt to produce more accurate forecasts.

In this section, the results obtained from our study will be discussed.

B. Effects of House Characteristics on its Price

The estimated results obtained from the correlation test of the house price and its characteristics in the hedonic regression test demonstrate that larger apartments with more bedrooms have higher prices. Also, apartments in higher floors within the building are priced higher. These findings came in agreement with the previous work by Limsombunchai and colleagues (2004). However, the location of the houses did not correlate significantly with their prices. An explanation to this finding could be due to the random process of data collection where some areas in Istanbul have more representation in the data analysis than other areas.

On the other hand, the study finds a negative relation between a house and its age, in which older houses have lower prices than newer houses and vice versa. Again this finding agrees with the finding in Limsombunchai (2004) study.

Furthermore, the relative importance of housing unit characteristics on its price results from the artificial neural network (ANN) analysis shows that the size of the housing unit and the number of bedrooms have great influence on its price. While the location of the housing unit has the lowest importance on the price.

C. Hedonic and ANN Models

In term of comparing the performances between the hedonic and the neural network models, we find that the artificial neural network forecasts of the house prices are superior to the forecasts generated by the hedonic regression model. ANN forecasts are very close to the actual house prices with lower MAFE than the hedonic forecasts.

The Wilcoxon signed rank test shows a statistically significant negative difference between the mean absolute errors of the hedonic and ANN forecasting models. This implies a superior forecasting by the ANN model.

This finding was expected as it came in agreement with the previous works by Kauko (2003), Limsombunchai et al. (2004), Selim (2009), Peterson and Flanagan (2009) and Morano and Tajani (2013) in which they all found that the performances of ANN models in prediction house prices are superior to the hedonic price predictions.

D. The Combining Models

The results of our study show that the combining models with the estimated WLS regression parameters restricted to add up one, with and without an intercept term, respectively (Model 3 and Model 4) were successful in enhancing forecast accuracy. Both the restrained models (Model 3 and Model 4) forecasts statistically outperform the ANN model’s forecasts.

Our study also finds that the unrestricted combining forecasts models with and without constant term (Model 1 and Model 2) failed to enhance the forecast accuracy. The forecasts generated by the ANN model has significantly lower MAFE than both the unrestricted models’ forecasts. These empirical findings are in stark contrast with the findings of both the Guerard (1987) and Lobo (1991) studies, which found that unrestricted-regression models performed best in leading to a combining model with superior forecast-accuracy (over each of the individual, component model forecasts in those studies).

Lastly, combining models of simple weighted averages (Model 5, Model 6, Model 7) generated by the following sets of weights: 0.50 and 0.50; 025 and 0.75; 0.75 and 0.25), were not successful in increasing forecast accuracy. These models show higher MAFEs compared to the ANN forecasts and other combining models' forecasts, see Table 10.

The empirical findings of the present study are supportive of the results of Terregrossa (2005), and thus counter to the results of both the Guerard (1987) and Lobo (1991) studies, in that the combination model constructed by estimated parameters of restricted WLS regression including an intercept term generally performed best, in leading to enhanced forecast accuracy over both of the individual, component model (ANN and hedonic) forecasts of housing prices in Istanbul.

E. Summary and Conclusions

This study generates forecasts of housing-unit prices, with employment of data from a cross-sectional random-sample of apartment dwellings in Istanbul. The models that are implemented include the conventional hedonic model, and the newer artificial neural network (ANN) model. Consistent with the findings of previous studies, our results indicate that the ANN model has greater accuracy than the hedonic model, in forecasting housing unit prices. However, our study develops a forecasting model that we demonstrate to have greater accuracy than the ANN model, by forming weighted- combinations of housing-unit price forecasts generated separately by the hedonic and ANN models. Our empirical analysis finds that the combining model the combination model constructed by estimated parameters of restricted WLS regression including an intercept term generally performed best, in leading to enhanced forecast accuracy over both of the individual, component model (ANN and hedonic) forecasts of housing prices in Istanbul; and generally outperformed all other forecast-models tested in our study. In the process, our study demonstrates that a constrained, linear combining model can improve ex ante estimates of housing-unit prices.

F. Study Limitations

Mainly, there is the issue of limited data availability for a study of housing prices in Istanbul. We were forced to painstakingly, manually collect data from limited areas in Istanbul, which led to a smaller data set than we would have preferred.

Additionally, the limited data availability forced us to undertake a cross-sectional empirical analysis, only. We would have preferred a multi-year data set to undertake a time-series empirical analysis, to see how well our empirical findings hold up over time.

G. Future Work

Future work may be done where a larger, multi-year sample size is taken; along with inclusion of different housing unit types, such as houses, apartments and studios.

Also, different forecasting models may be used in generating the combining model; other than, or in conjunction with, the artificial neural network and the hedonic models.

VI. REFERENCES

BOOKS

ARMSTRONG, J S. (2001). **Principles of forecasting: a handbook for researchers and practitioners,** Kluwer Academic Publishing, pages 417-439.

BOARDMAN A E, Greenberg, D H, Vining, A R & Weiner, D L. (2002**). Cost-benefit analysis: concepts and practice**. 2nd edition. Upper Saddle River, N.J.; [Great Britain]: Prentice Hall. pp: 349-352.

FREEMAN, M A. (1979). **The Benefits of Environmental Improvement**. Washington, D.C.: Resources for the Future.

LANGLOIS, J H and Roggman, L A. (1990), **"Attractive faces are only average," Psychological Science**, 1 (March), 115-121.

ARTICLES

ADAIR, A, Mcgreal, S, Smyth, A, Cooper, J, & Ryley, T. (2000**).** House prices and accessibility: The testing of relationships within the Belfast urban area. **Housing Studies** 15(5), 699–716.

ALFIATYN, A N, Febrita, R E, Taufiq, H & Mahmudy, W F. (2017). Modeling House Price Prediction using Regression Analysis and Particle Swarm Optimization Case Study: Malang, East Java, Indonesia. **International Journal of Advanced Computer Science and Applications (IJACSA)** 8(10).

BIN, O A. (2004). Prediction comparison of housing sales prices by parametric versus semi-parametric regressions. **Journal of Housing Economics**. pp. 68-84.

BISCHOFF, C W. (1989). The combination of macroeconomic forecasts**. Journal of Forecasting, 8**, 293–314.

BLOMQUEST, G, Worley, L. (1981). Hedonic Prices, Demand for Urban Housing Amenities and Benefit Estimates. **J. Urban Econ.,** 9: 212 - 221.

BRADLEY M, Gordon J, Mcmanus D (2003) Method for combining house price forecasts. **United States Patent** 6609109

BRORSEN, B W, Grant W R and Rister, M E. (1984). A Hedonic Price Model for Rough Rice Bid/Acceptance Markets. **American J. Agri. Econ.,** 66: 156 - 163.

CABRERE J, Wang T, Yang J (2011) Linear and nonlinear predictability of international securitized real estate returns: A reality check. **Journal of Real Estate Research** 33: 565-594

CALHUN, C A. (2003). Property Valuation Models and House Price Indexes for the Provinces of Thailand: 1992 – 2000**. Housing Finance International**. 17: 31 – 41.

CALHUN, C A. (2001). Property Valuation Methods and Data in The United States. Housing. **Finance International,** 16: 12 - 23.

CURRY, B, Morgan, P, & Silver, M. (2002). Neural networks and nonlinear statistical methods: An application to the modelling of price quality relationships. **Computers & Operations Research**, 29, 951–969.

DIN, A, Hoesli, M, Bender, A. (2001). Environmental variables and real estate prices. **Urban Studies,** 38(11), 1989–2000.

DROUGHT S, McDonald C (2011) Forecasting house price inflation: A model combination approach. Discussion Paper Series 32**., Reserve Bank of New Zealand.**

ETHRIDGE, D E, Davis, B. (1982). Hedonic Price Estimation for Commodities: An Application to Cotton. **Western J. Agri. Econ.,** 7: 293 - 300.

FAIR, R, Shiller, R. (1990). Comparing information in forecasts from time series models and econometric models. **American Economic Review**, 80(3), 375–89.

FAN, G, Ong, Z S E, Koh, H C. (2006). Determinants of house price: A decision tree approach. **Urban Studies,** 43 (12), pp. 2301-2315.

FILHO, C M, Bin, O. (2005). Estimation of hedonic price functions via additive nonparametric regression. **Empirical Economics**, 30 (1), pp. 93-114.

FLEMING M, Kuo C L (2007) Method and apparatus for predicting and reporting a real estate value based on a weighted average of predicted values. **United States Patent** 7305328.

FLETCHER, M, Gallimore, P, Mangan, J. (2000a). Heteroscedasticity in hedonic house price models**. J. Property Res.** 17, 93–108.

FREW, J, Jud, G D. (2003). Estimating the Value of Apartment Buildings. **The J. Real Estate Res**., 25: 77 - 86.

GOODMAN, A, Thibodeau, T. (1995). Dwelling age heteroscedasticity in hedonic house price equations**. J. Housing Res.** 6, 25–42.

GOODMAN, A, Thibodeau, T. (1997). Dwelling age heteroscedasticity in hedonic house price equations: an extension**. J. Housing Res**. 8, 299–317.

GRANGER, C and Ramanathan, R. (1984) Improved methods of combining forecasts, **Journal of Forecasting**, 3, 197–204.

GRILICHES, Z. ed. (1971). Price Indexes and Quality Changes: Studies in New Methods of Measurement. **Cambridge, Mass: Harvard U. Press.**

GUERARD, J B. (1987). Linear constraints, robust weighing and efficient composite modeling, **Journal of Forecasting**, 6, 193–9.

GUPTA R, Kabundi A, Miller S (2011) Forecasting the US real house price index: Struc-tural and non-structural models with and without fundamentals. **Economic Modelling** 28: 2013-2021.

Halvorsen, R and Palmquist, R. (1980). The Interpretation of Dummy Variables in Semilogarithmic Equations. **The American Economic Review**, 70 (3), 474-475.

HASHEM, S. (1997). Optimal linear combinations of neural networks. **Neural Networks,** 10 (4), pp. 599-614.

HENG, J, Wang, J, Xiao, L and Lu, H. (2017). Research and application of a combined model based on frequent pattern growth algorithm and multi-objective optimization for solar radiation forecasting. **Applied Energy**, 208; 845-866.

HORNIK, K, Stinchcombe, M and White, H. (1989). Multi-layer feedforward networks are universal approximators**. Neural Networks**, 2(5), 359–366.

JANSSEN, C, So¨derberg, B and Zhou, J. (2001). Robust estimation of hedonic models of price and income for investment property. **Journal of Property Investment & Finance,** 19(4), 342–360.

JIANG, C, Zhang, J and Song, F. (2014). Selecting Single Model in Combination Forecasting Based on Cointegration Test and Encompassing Test. **Scientific World Journal**, doi: 10.1155/2014/621917.

JUN, W, Yuyan, L, Lingyu, T and Peng, G. (2018). Modeling a combined forecast algorithm based on sequence patterns and near characteristics: An application for tourism demand forecasting. **Chaos, Solitons & Fractals** 108; 136-147.

KAUKO, T. (2003). On current neural network applications involving spatial modelling of property prices. **Journal of Housing and the Built Environment,** 18(2), 159–181.

KAUKO, T, Hooimeijer, P and Hakfoort, J. (2002). Capturing housing market segmentation: An alternative approach based on neural network modelling. **Housing Studies,** 17(6), 875–894.

KUMAR, M and Patel, N R. (2010) Using clustering to improve sales forecasts in retail merchandising. *Ann Oper Res* 174, 33–46.

LANCASTER, K J. (1966). A New Approach to Consumer Theory. **J. Political Economy,** 4:132 – 157.

LANDAJO, M, Bilbao, C and Bilbao, A. (2012). Nonparametric neural network modeling of hedonic prices in the housing market. **Empire Econ** 42, 987–1009.

LENK, M M, Worzala, E M and Silva, A. (1997). High-tech Valuation: Should Artificial Neural Networks Bypass the Human Valuer? **J. Property Valuation & Investment**, 15: 8 – 26.

LIMSOMBUNCHAI, V, Gan, C and Lee, M. (2004). House Price Prediction: Hedonic Price Model vs. Artificial Neural Network. **American Journal of Applied Sciences**, 1 (3): 193-201.

LIU, X, Moreno, B and Garcíac, S. (2016). A grey neural network and input-output combined forecasting model. Primary energy consumption forecasts in Spanish economic sectors**. Energy**, 115(1); 1042-1054

LIU Y, Zhang S, Chen X and Wang J. (2018). Artificial combined model based on hybrid nonlinear neural network models and statistics linear models-research and application for wind speed forecasting. **Sustainability (Switzerland).**

LIU, Z, Jiang, P, Zhang, L and Niu, X. (2020). Combined Forecasting Model for Time series: Application to Short-term Wind Speed Forecasting. **Applied Energy** 259, 1-25.

LOBO, G. (1991) Alternative methods of combining security analysts' and statistical forecasts of annual corporate earnings, International **Journal of Forecasting**, 7, 57–63.

LOBO, G. (1992) Analysis and comparison of financial analysts', time-series and combined forecasts of annual earnings, **Journal of Business Research**, 24, 269–80.

LOH, E Y L. (2005) Profiting from moving averages and time-series forecasts: Asian-Pacific evidence, The Asia Pacific **Journal of Economics & Business**, 9(1), 62-82, 89.

MCMILLAN, M L, Reid, B G and Gillen, D W. (1980). An Extension of the Hedonic Approach for Estimating the Value of Quiet**. Land Economics**, 56: 315 - 328.

MEESE, R and Wallace, N. (2003) House price dynamics and market fundamentals. **The Parisian housing market Urban Studies**, 40 (5–6), pp. 1027-1045.

MESSONNIER, M L and Luzar, E J. (1990). A Hedonic Analysis of Private Hunting Land Attributes Using an Alternative Functional Form. Southern **J. Agri. Econ.**, 232: 129 - 135.

MILON, J W, Jonathan G and Mulkey, D. (1984). Hedonic Amenity Valuation and Functional Form Specification. **Land Economics**, 60:378- 387.

MORANO, P and Tajani, F. (2013). Bare ownership evaluation. Hedonic price model vs. artificial neural network *Int.* ***J. Business Intelligence and Data Mining****, 8-4.*

MORENO, B and López, A.J. (2007) Combining economic forecasts through information measures, **Applied Economics Letters**, 14:12, 899-903, DOI: 10.1080/13504850600689964.

NEWBOLD, P and Granger, C. (1974) Experience with forecasting univariate time series and the combination of forecasts. **Journal of the Royal Statistical Society,** Series A, 137, 131–46.

NEWBOLD, P, Zumwalt, J and Kannan, S. (1987) Combining forecasts to improve earnings per share prediction: an examination of electric utilities, **International Journal of Forecasting**, 3, 229–38.

NIU, X and Wang, J. (2019). A combined model based on data preprocessing strategy and multiobjective optimization algorithm for short-term wind speed forecasting**. Appl Energy**, 241: 519–39.

OWEN, C and Howard, J. (1998). Estimation Realisation Price (ERP) by Neural Networks: Forecasting Commercial Property Values. **J. Property Valuation & Investment**, 16: 71 – 86.

PETERSON, S and Flanagan, A B. (2009). Neural Network Hedonic Pricing Models in Mass Real Estate Appraisal. Journal of Real Estate Research, **American Real Estate Society** 31(2), 147-164.

POPE, C A and Stoll, J R. (1985). The Market Value of Ingress Rights for White-Tailed Deer Hunting in Texas. Southern **J. Agri. Econ.,** 17:177 - 182.

ROSEN, S. (1974). Hedonic Prices and Implicit Markets. Product Differentiation in Pure Competition, **J. Political Econ.,** 82: 34 – 55.

SCHULZ, R and Werwatz, A. (2004). A State Space Model for Berlin House Prices: Estimation and Economic Interpretation. **The Journal of Real Estate Finance and Economics**, 28(1). Pp. 37–57.

SELIM, H. (2009). Determinants of house prices in Turkey: Hedonic regression versus artificial neural network. **Expert Systems with Applications**, 36: 2843–2852.

SHONKWILER J S and Reynolds, J E. (1986). A note on the use of hedonic price models in the analysis of land prices at the urban fringe. **Land Economics**, 62(1).

SONG, C and Fu, X. (2020). Research on different weight combination in air quality forecasting models (Revision). Journal of Cleaner Production.

STANLEY, M, Alastair, A, Dylan, M and Patterson, D. (1998). Neural Networks: The Prediction of Residential Values. **J. Property Valuation & Investment**, 16: 57 – 70.

STEVENSON, S. (2004). New empirical evidence on heteroscedasticity in hedonic housing models**. Journal of Housing Economics**, 13, pp. 136-153.

TERREGROSSA, S J. (2005). On the efficacy of constraints on the linear combination forecast model. **Applied Economics Letters**, 12, pp. 19–28.

TERREGROSSA, S J. (1999). Combining analysts' forecasts with causal model forecasts of earnings growth. **Applied Financial Economics**, 9, pp. 143-153.

WANG, R, Wang, J and Xu, Y. (2019). A novel combined model based on hybrid optimization algorithm for electrical load forecasting**. Applied Soft Computing,** 82; 105548

WILSON, W W. (1984). Hedonic Prices in the Malting Barley Market. Western **J. Agri. Econ.,** 9:29 - 40.

WORZALA, E, Lenk, M and Silva, A. (1995) An Exploration of Neural Networks and Its Application to Real Estate Valuation, **The Journal of Real Estate Research**, 10(2), 185–201.

ELECTRONIC SOURCES

Turkey Expert. www.turkeyexpert.com. [Accessed on 20th, 21st &30th June 2019].

Turkey Homes. www.turkeyhomes.com. [Accessed from 1st to 9th July 2019].

Property Turkey. www.propertyturkey.com. [Accessed from 29th May to 10th June 2019].

Istanbul Homes. www.istanbulhomes.com. [Accessed from 3rd to 6th June 2019].

Istanbul Property World. www.istanbulpropertyworld.com. [Accessed from 19th to 22nd May 2019].

Istanbul Real Estate. www.istanbulrealestate.com. [Accessed on 15th, 16th and 17th May 2019].

DISSERTATIONS

PINDER, J P. (2017). Introduction to Business Analytics using Simulation. Elsevier.

PAGOURTZI E, Assimakopoulos V and Litsa A (2005) Theta model forecasts of quarterly and monthly dwelling prices in the UK. Briengs in Real Estate Finance 5: 75-105

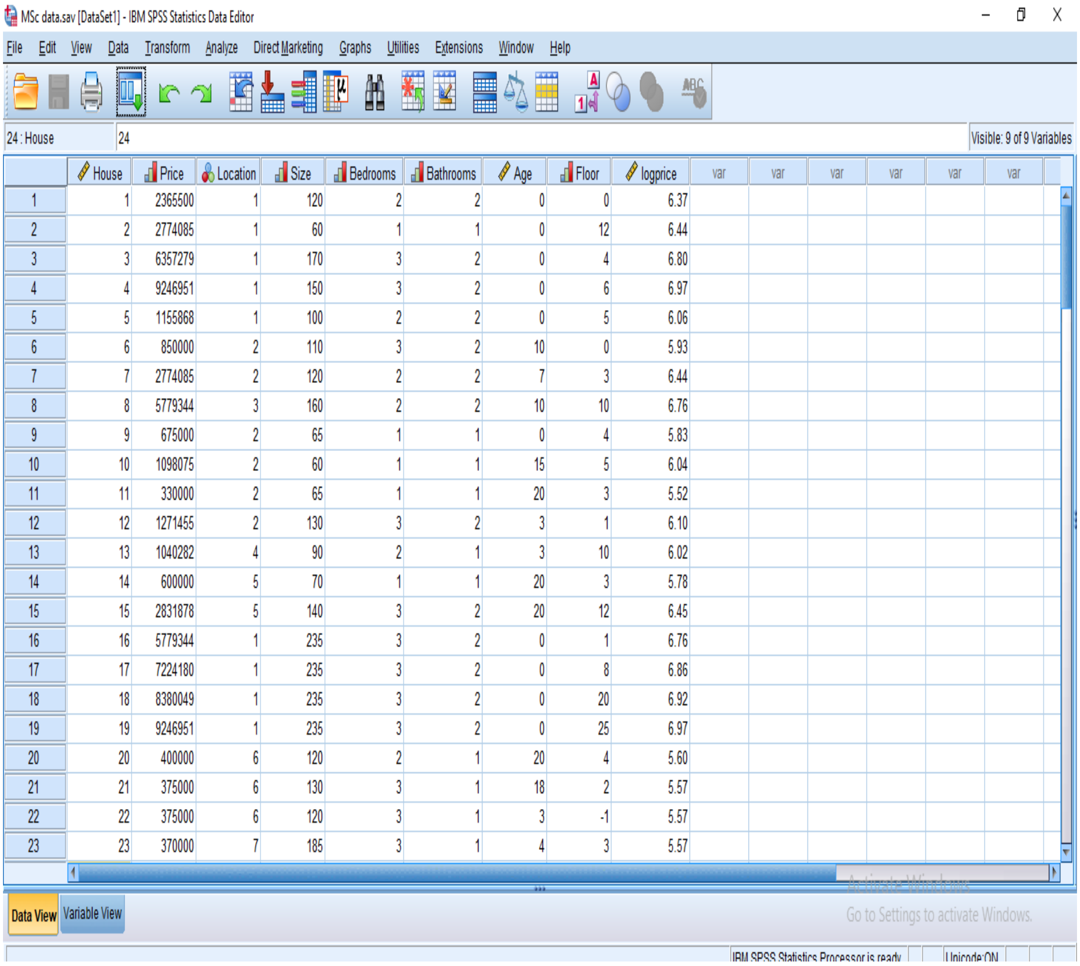
OTHER SOURCES

WITTE, A D, Sumka, H J and Erekson, H. (1979). An Estimate of a Structural Hedonic Price Model of the Housing Market: An Application of Rosen's Theory of Implicit Markets. Econometrica, 47:1151 - 1173.

ZHANG, G P and Berardi, V L. (2000). Combining multiple neural networks for time series forecasting, in: **Proceedings of the Decision Science Institute Annual Meeting**, 2000,pp. 966–968.

WU, J, Zhou, J, Chen, L and Ye, L. (2015) Coupling forecast methods of multiple rainfall–runoff models for improving the precision of hydrological forecasting, **Water Resources Management,** 29:5091–5108 DOI: 10.1007/s11269-015-1106-8.

APPENDICES


MSc data.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Extensions Window Help

24 : House 24 Visible: 9 of 9 Variables

	House	Price	Location	Size	Bedrooms	Bathrooms	Age	Floor	logprice	var	var	var	var	var	var
1	1	2365500	1	120	2	2	0	0	6.37						
2	2	2774085	1	60	1	1	0	12	6.44						
3	3	6357279	1	170	3	2	0	4	6.80						
4	4	9246951	1	150	3	2	0	6	6.97						
5	5	1155868	1	100	2	2	0	5	6.06						
6	6	850000	2	110	3	2	10	0	5.93						
7	7	2774085	2	120	2	2	7	3	6.44						
8	8	5779344	3	160	2	2	10	10	6.76						
9	9	675000	2	65	1	1	0	4	5.83						
10	10	1098075	2	60	1	1	15	5	6.04						
11	11	330000	2	65	1	1	20	3	5.52						
12	12	1271455	2	130	3	2	3	1	6.10						
13	13	1040282	4	90	2	1	3	10	6.02						
14	14	600000	5	70	1	1	20	3	5.78						
15	15	2831878	5	140	3	2	20	12	6.45						
16	16	5779344	1	235	3	2	0	1	6.76						
17	17	7224180	1	235	3	2	0	8	6.86						
18	18	8380049	1	235	3	2	0	20	6.92						
19	19	9246951	1	235	3	2	0	25	6.97						
20	20	400000	6	120	2	1	20	4	5.60						
21	21	375000	6	130	3	1	18	2	5.57						
22	22	375000	6	120	3	1	3	-1	5.57						
23	23	370000	7	185	3	1	4	3	5.57						

Data View Variable View

Go to Settings to activate Windows.

IBM SPSS Statistics Processor is ready Unicode:ON

MSc data.sav [DataSet1] - IBM SPSS Statistics Data Editor

	House	Price	Location	Size	Bedrooms	Bathrooms	Age	Floor	logprice
24	24	478194338	7	300	5	5	0	1	8.68
25	25	190082249	8	100	2	1	0	1	8.28
26	26	237901683	8	110	3	1	0	6	8.38
27	27	267788829	8	135	4	2	0	8	8.43
28	28	141067329	9	90	2	1	1	3	8.15
29	29	186495791	9	125	3	1	1	4	8.27
30	30	255833970	9	160	5	2	1	7	8.41
31	31	167368018	7	80	2	1	0	9	8.22
32	32	261811400	7	105	3	1	0	2	8.42
33	33	381359984	7	120	4	2	0	5	8.58
34	34	504495026	7	150	6	2	0	6	8.70
35	35	225946824	2	65	1	1	2	11	8.35
36	36	62165263	10	65	1	1	1	2	7.79
37	37	100420811	10	90	2	1	1	2	8.00
38	38	185300306	10	130	3	2	1	2	8.27
39	39	76511094	7	70	1	1	0	1	7.88
40	40	131503443	11	55	1	1	1	16	8.12
41	41	167368018	11	140	2	1	1	14	8.22
42	42	71729150	6	85	2	1	1	12	7.86
43	43	71729150	6	55	1	1	0	3	7.86
44	44	106398240	6	75	2	1	0	3	8.03
45	45	155413159	6	75	2	1	0	13	8.19
46	46	131503443	12	65	1	1	1	1	8.12

Data View Variable View

MSc data.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Extensions Window Help

24 : House 24 Visible: 9 of 9 Variables

	House	Price	Location	Size	Bedrooms	Bathrooms	Age	Floor	logprice	var	var	var	var	var	var
47	47	203232593	12	85	2	1	1	1	8.31						
48	48	362232211	12	140	4	2	2	7	8.56						
49	49	329954093	12	65	1	1	0	15	8.52						
50	50	2161438408	12	340	6	3	1	14	9.33						
51	51	86074980	6	86	1	1	5	2	7.93						
52	52	107593726	9	125	2	1	16	4	8.03						
53	53	37060061	7	85	1	1	6	10	7.57						
54	54	131503443	6	184	2	1	0	12	8.12						
55	55	31082631	10	75	1	1	4	2	7.49						
56	56	33473603	10	85	1	1	3	1	7.52						
57	57	29887146	10	70	1	1	2	1	7.48						
58	58	35864575	10	90	2	1	2	12	7.55						
59	59	33473603	10	70	1	1	1	9	7.52						
60	60	47819433	10	110	3	1	3	6	7.68						
61	61	53796863	10	155	4	2	3	4	7.73						
62	62	735300	13	114	2	1	0	3	5.87						
63	63	564300	13	153	2	2	0	2	5.75						
64	64	5044500	9	232	3	2	0	0	6.70						
65	65	416100	10	71	1	1	0	6	5.62						
66	66	655500	14	61	1	1	0	8	5.82						
67	67	4751300	9	370	5	3	0	0	6.68						
68	68	1248300	11	50	1	1	0	9	6.10						
69	69	427500	8	50	1	1	0	5	5.63						

Data View Variable View

Activate Windows
Go to Settings to activate Windows.

MSc data.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Extensions Window Help

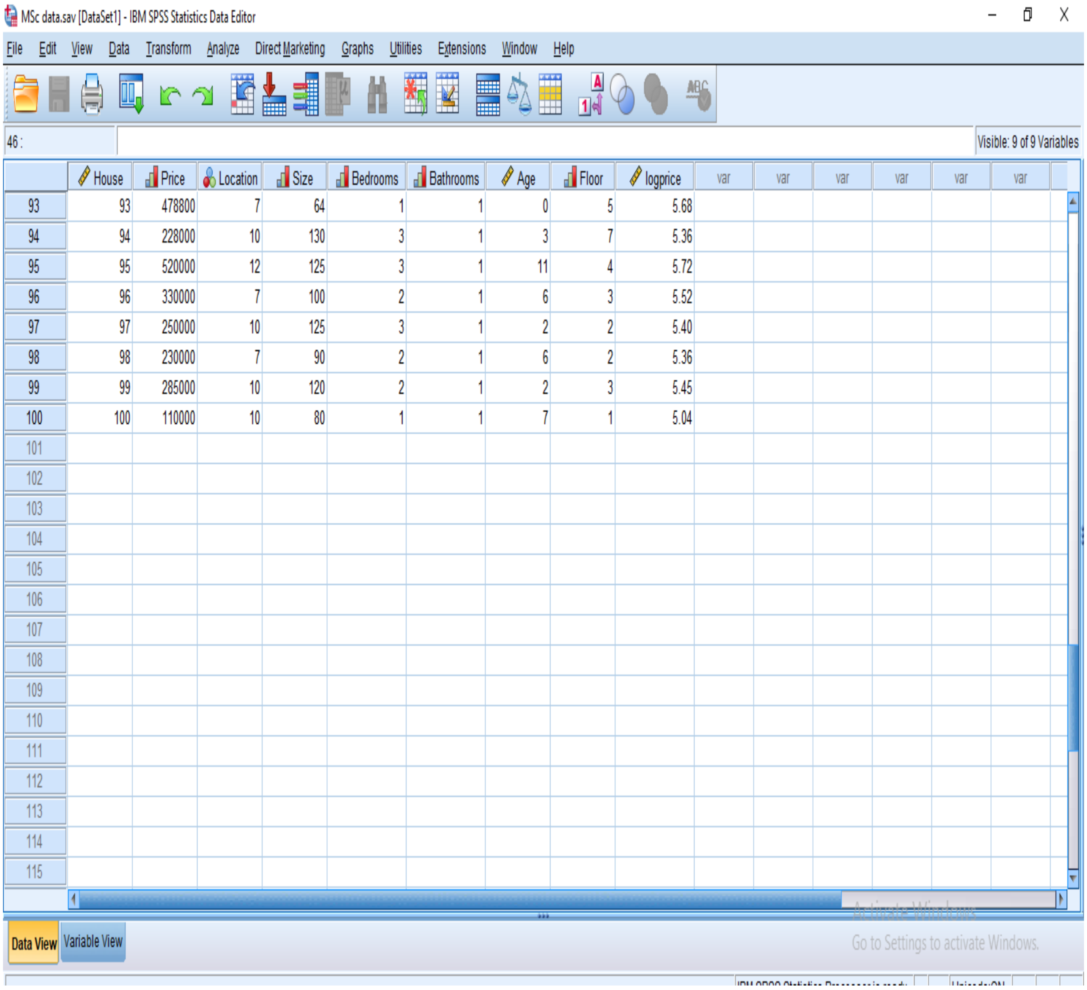
46 :

Visible: 9 of 9 Variables

	House	Price	Location	Size	Bedrooms	Bathrooms	Age	Floor	logprice	var	var	var	var	var	var
70	70	1094400	16	148	2	1	0	1	6.04						
71	71	1077300	16	54	1	1	0	4	6.03						
72	72	456000	14	53	1	1	0	1	5.66						
73	73	381900	17	66	1	1	0	3	5.58						
74	74	712500	6	121	2	1	0	4	5.85						
75	75	467400	14	65	1	1	0	1	5.67						
76	76	843600	13	106	2	1	0	1	5.93						
77	77	250800	10	48	1	1	0	9	5.40						
78	78	1248300	15	62	1	1	0	11	6.10						
79	79	815100	18	42	1	1	0	12	5.91						
80	80	370500	19	61	1	1	0	14	5.57						
81	81	980400	3	46	1	1	0	18	5.99						
82	82	786600	16	89	2	1	0	7	5.90						
83	83	1915200	2	102	2	1	0	2	6.28						
84	84	3083700	7	677	7	3	0	0	6.49						
85	85	1635900	8	115	2	1	0	10	6.21						
86	86	2935500	15	73	1	1	0	16	6.47						
87	87	627000	12	48	1	1	0	2	5.80						
88	88	957600	2	63	1	1	0	5	5.98						
89	89	433200	10	58	1	1	0	8	5.64						
90	90	495900	10	64	1	1	0	4	5.70						
91	91	815100	12	124	2	1	0	2	5.91						
92	92	4246500	2	157	2	2	0	6	6.63						

Data View Variable View

Go to Settings to activate Windows.



	House	Price	Location	Size	Bedrooms	Bathrooms	Age	Floor	logprice	var	var	var	var	var	var
93	93	478800	7	64	1	1	0	5	5.68						
94	94	228000	10	130	3	1	3	7	5.36						
95	95	520000	12	125	3	1	11	4	5.72						
96	96	330000	7	100	2	1	6	3	5.52						
97	97	250000	10	125	3	1	2	2	5.40						
98	98	230000	7	90	2	1	6	2	5.36						
99	99	285000	10	120	2	1	2	3	5.45						
100	100	110000	10	80	1	1	7	1	5.04						
101															
102															
103															
104															
105															
106															
107															
108															
109															
110															
111															
112															
113															
114															
115															

Appendix 2: Descriptive Analysis

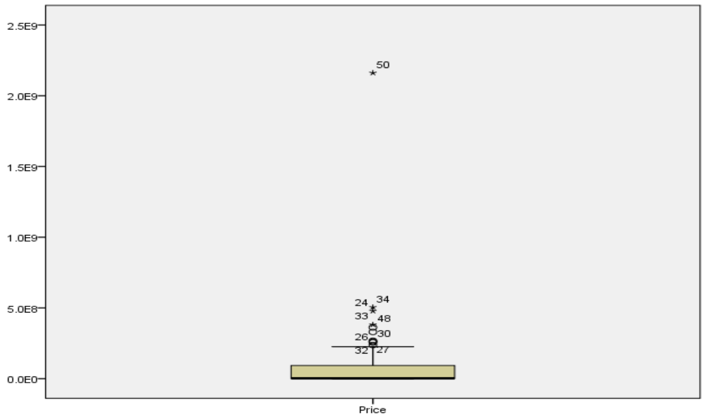
Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Price	100	110000	2161438408	83570802.13	236277040.500
Age	100	0	20	2.51	4.984
Floor	100	-1	25	5.73	5.059
Valid N (listwise)	100				

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Price	.362	100	.000	.347	100	.000

a. Lilliefors Significance Correction



Descriptives

			Statistic	Std. Error
Price	Mean		83570802.13	23627704.050
	95% Confidence Interval for Mean	Lower Bound	36688311.22	
		Upper Bound	130453293.00	
	5% Trimmed Mean		49647682.73	
	Median		2883689.00	
	Variance		55826839880000000.000	
	Std. Deviation		236277040.500	
	Minimum		110000	
	Maximum		2E+9	
	Range		2161328408	
	Interquartile Range		96200228	
	Skewness		7.163	.241
	Kurtosis		61.376	.478

Location

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Sisli	9	9.0	9.0	9.0
	Beyoglu	10	10.0	10.0	19.0
	Bosfor	2	2.0	2.0	21.0
	Tarlabaci	1	1.0	1.0	22.0
	Cihangir	2	2.0	2.0	24.0
	Avcilar	10	10.0	10.0	34.0
	Beylikdozo	12	12.0	12.0	46.0
	Basaksehir	5	5.0	5.0	51.0
	Beyukcekmece	6	6.0	6.0	57.0
	Esenyurt	18	18.0	18.0	75.0
	Bakirkoy	3	3.0	3.0	78.0
	Fatih	8	8.0	8.0	86.0
	Kavakli	3	3.0	3.0	89.0
	Kucukcemece	3	3.0	3.0	92.0
	Zencirlikoy	2	2.0	2.0	94.0
	Zeytinburnu	3	3.0	3.0	97.0
	Pendik	1	1.0	1.0	98.0
	Maslak	1	1.0	1.0	99.0
	Euyp	1	1.0	1.0	100.0
	Total	100	100.0	100.0	

Appendix 3: Hedonic regression analysis

Correlations

		Price	Location	Size	Bedrooms	Bathrooms	Age	Floor
Price	Pearson Correlation	1	.096	.289**	.494**	.407**	-.101	.172
	Sig. (2-tailed)		.398	.009	.000	.000	.375	.127
	N	80	80	80	80	80	80	80
Location	Pearson Correlation	.096	1	-.173	-.152	-.223*	-.285*	-.096
	Sig. (2-tailed)	.398		.126	.178	.047	.011	.398
	N	80	80	80	80	80	80	80
Size	Pearson Correlation	.289**	-.173	1	.809**	.738**	-.101	-.046
	Sig. (2-tailed)	.009	.126		.000	.000	.375	.683
	N	80	80	80	80	80	80	80
Bedrooms	Pearson Correlation	.494**	-.152	.809**	1	.759**	-.084	-.061
	Sig. (2-tailed)	.000	.178	.000		.000	.458	.593
	N	80	80	80	80	80	80	80
Bathrooms	Pearson Correlation	.407**	-.223*	.738**	.759**	1	-.096	-.026
	Sig. (2-tailed)	.000	.047	.000	.000		.397	.821
	N	80	80	80	80	80	80	80
Age	Pearson Correlation	-.101	-.285*	-.101	-.084	-.096	1	-.171
	Sig. (2-tailed)	.375	.011	.375	.458	.397		.130
	N	80	80	80	80	80	80	80
Floor	Pearson Correlation	.172	-.096	-.046	-.061	-.026	-.171	1
	Sig. (2-tailed)	.127	.398	.683	.593	.821	.130	
	N	80	80	80	80	80	80	80

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Model Summary[b,c]

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.946[a]	.895	.886	1.17509	.895	103.547	6	73	.000

a. Predictors: (Constant), Floor, Location, Age, Size, Bathrooms, Bedrooms

b. Dependent Variable: logprice

c. Weighted Least Squares Regression - Weighted by weight

Coefficients[a,b]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	5.207	.312		16.684	.000
	Location	.027	.023	.049	1.167	.247
	Size	-.007	.001	-.828	-6.953	.000
	Bedrooms	.637	.115	.777	5.530	.000
	Bathrooms	.421	.064	.603	6.620	.000
	Age	-.034	.021	-.076	-1.605	.113
	Floor	.046	.014	.163	3.237	.002

a. Dependent Variable: logprice

b. Weighted Least Squares Regression - Weighted by weight

Appendix 4: ANN analysis

Network Information

Input Layer	Covariates	1	Location
		2	Size
		3	Bedrooms
		4	Bathrooms
		5	Age
		6	Floor
	Number of Units[a]		6
	Rescaling Method for Covariates		Standardized
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1[a]		4
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	logprice
	Number of Units		1
	Rescaling Method for Scale Dependents		Standardized
	Activation Function		Identity
	Error Function		Sum of Squares

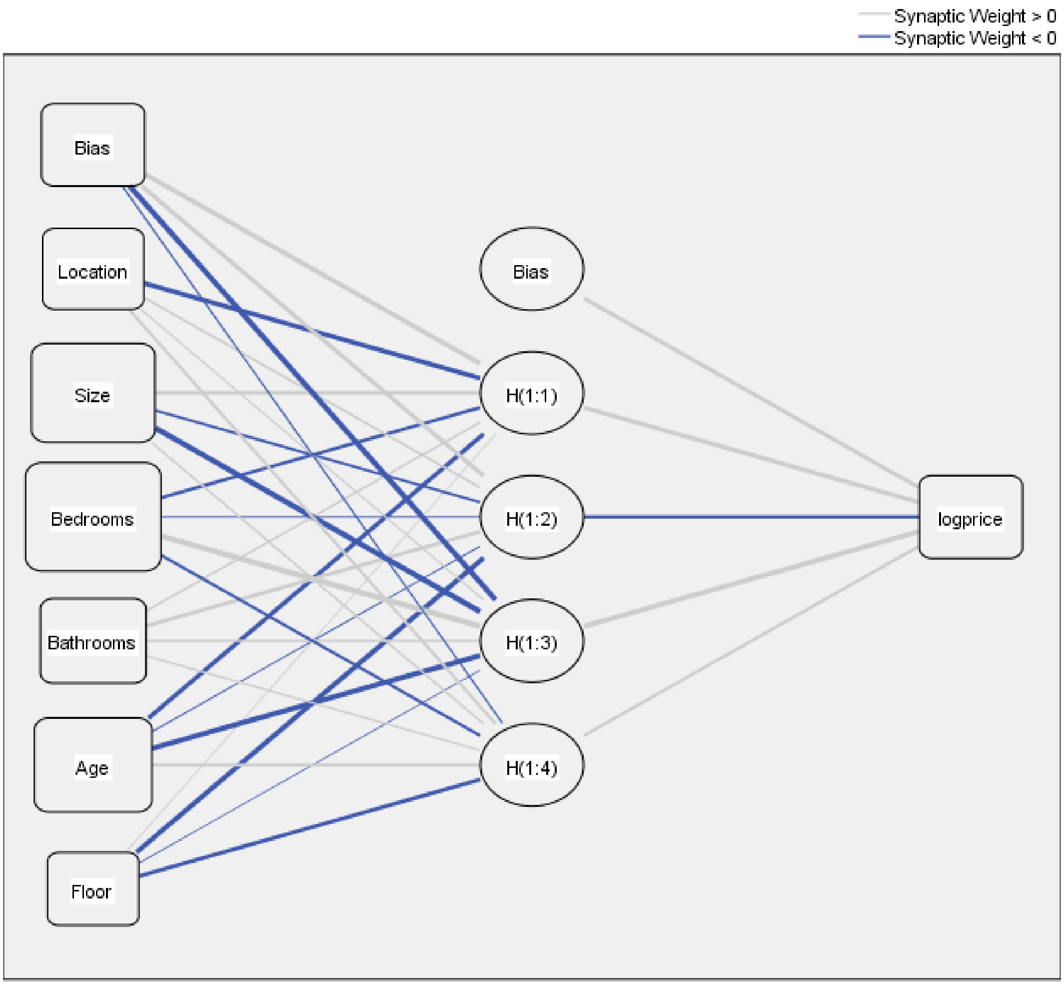
a. Excluding the bias unit

Case Processing Summary

		N	Percent
Sample	Training	80	80.0%
	Testing	20	20.0%
Valid		100	100.0%
Excluded		0	
Total		100	

Independent Variable Importance

	Importance	Normalized Importance
Location	.112	43.5%
Size	.258	100.0%
Bedrooms	.209	81.1%
Bathrooms	.124	48.0%
Age	.138	53.4%
Floor	.160	62.0%



Hidden layer activation function: Hyperbolic tangent

Output layer activation function: Identity

Parameter Estimates

Predictor		Predicted Hidden Layer 1 H(1:1)	H(1:2)	H(1:3)	H(1:4)	Output Layer logprice
Input Layer	(Bias)	1.009	.688	-1.835	-.109	
	Location	-.838	.202	.151	.449	
	Size	.544	-.301	-1.883	.170	
	Bedrooms	-.470	-.070	2.295	-.446	
	Bathrooms	.178	.559	.247	.160	
	Age	-.700	-.032	-1.505	.326	
	Floor	.011	-1.014	-.005	-.572	
Hidden Layer 1	(Bias)					.622
	H(1:1)					.748
	H(1:2)					-.548
	H(1:3)					1.202
	H(1:4)					.450

Appendix 5: Mean absolute error

Test Statistics[a]

	MAE_ANN - MAE_Hedonic
Z	-3.173-[b]
Asymp. Sig. (2-tailed)	.002

a. Wilcoxon Signed Ranks Test
b. Based on positive ranks.

Test Statistics[a]

	MAE3 - MAE1	MAE4 - MAE2
Z	-3.920-[b]	-3.920-[c]
Asymp. Sig. (2-tailed)	.000	.000

a. Wilcoxon Signed Ranks Test
b. Based on positive ranks.
c. Based on negative ranks.

Test Statistics[a]

	MAE1 - MAE_ANN	MAE2 - MAE_ANN	MAE3 - MAE_ANN	MAE4 - MAE_ANN	MAE5 - MAE_ANN	MAE6 - MAE_ANN	MAE7 - MAE_ANN
Z	-3.920-[b]	-3.920-[b]	-3.323-[c]	-3.733-[c]	-2.277-[b]	-.709-[b]	-2.763-[b]
Asymp. Sig. (2-tailed)	.000	.000	.001	.000	.023	.478	.006

a. Wilcoxon Signed Ranks Test
b. Based on negative ranks.
c. Based on positive ranks.

Test Statistics[a]

	MAE1 - MAE_Hedonic	MAE2 - MAE_Hedonic	MAE3 - MAE_Hedonic	MAE4 - MAE_Hedonic	MAE5 - MAE_Hedonic	MAE6 - MAE_Hedonic	MAE7 - MAE_Hedonic
Z	-3.360[c] -	-3.920-[b]	-3.845-	-3.845-[c]	-3.883-[c]	-3.845-[c]	-3.920-[c]
Asymp. Sig. (2-tailed)	.001	.000	.000	.000	.000	.000	.000

a. Wilcoxon Signed Ranks Test
b. Based on negative ranks.
c. Based on positive ranks.

Appendix 6: Combination forecasts weight calculation (WLS) regression

1. Unrestricted regression with constant term

Coefficients[a,b]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	T	Sig.
1	(Constant)	-1.267-	.273		-4.636-	.000
	Hedonic	.4-+][=15	.070	.393	5.953	.000
	ANN	.780	.086	.600	9.095	.000

a. Dependent Variable: logprice

b. Weighted Least Squares Regression - Weighted by weight

2. Unrestricted regression with constant term supressed

Coefficients[a,b,c]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	ANN	1.067	.001	1.000	1309.547	.000

a. Dependent Variable: logprice

b. Linear Regression through the Origin

c. Weighted Least Squares Regression - Weighted by weight

Excluded Variables[a,b,c]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics Tolerance
1	Hedonic	.459[d]	5.910	.000	.556	6.760E-5

a. Dependent Variable: logprice

b. Linear Regression through the Origin

c. Weighted Least Squares Regression - Weighted by weight

d. Predictors in the Model: ANN

3. Restricted (WLS) regression with constant term

Parameter Estimates

Parameter	Estimate	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
A	.028	.108	-.187-	.244
B	.972	.081	.810	1.134
C	-.001-	.595	-1.187-	1.185

Restricted (WLS) regression with constant term supressed

Parameter Estimates

Parameter	Estimate	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
A	.006	.050	-.132-	.188
B	.994	.077	.813	1.131

RESUME

Mohammad Hussain Abbady

An enthusiastic and hardworking Business Economics graduate with excellent English language and sales and marketing ability,

Seek to join a team of professionals committed to impove.

Focused and highly motivated, with interpersonal and communication skills. Confident and at ease in any working setting and quick to adapt and respond to challenges and always willing to learn and open to changes.

Key skills

Excellent vocabulary, grammar and command on spoken and written English.

Computer knowledge : highly competent in use of Microsoft Office programs specialist package Sage.

Excellent email and internet research skills.

Office programs;(Excel, Word, PowerPoint) and video editing software.
Organizational skills: Adaptability, Time management, Documentation.

Work experience

Sep 2015_ Apr2017 Adwaa Alkut, Sales Representative, Sales manager.

Feb 2018_ Mar2020 Trade, real estate.

Aug2020_ Present Wahaj Alkahlaa, Operator.

Academic qualifications

2017_2020 Master of Business Administration (M.B.A). Istanbul Aydin University/Turkey

Combining Hedonic model forecasts with Artificial neural network forecasts of housing prices in Istanbul, TURKEY.

2010_2014 B.Sc. Business Economic Al-Nahrain University / Iraq

Dissertation: The Effect of Advertisement on The Consumer's Behavior.