Journal of System Management (JSM)
Online ISSN: 2538-1571
Print ISSN: 2322-2301
Vol. 7, No. 2, (26) 2021, pp. 171-203
Research Paper

Mining a Set of Rules for Determining the Waiting Time for Selling Residential Units

Farshid Abdi [1] **Shaghayegh Abolmakarem** [2]

Abstract

Being aware of the waiting time for selling residential units is one of the important issues in the housing sector for the majority of people, especially investors. There are several factors affecting the waiting time for selling residential units. Determining the influential factors on the time period of selling real estates can lead to an informed decision making by real estate consultants, sellers as well as those seeking to buy real estates. Using a real estate database in Iran, the present paper proposes a two-module procedure. The first module deals with implementation of association rule mining. Using the well-known association rule mining techniques namely FP-Growth, several association rules have been extracted which indicate the effective factors on the waiting time for selling residential units. Generated association rules have been evaluated based on metrics such as support, confidence and lift and finally the best rules are selected. The main objective of the second module is to develop a fuzzy inference system which can determine the factors influencing the waiting time for selling residential units from historical data, so that the model can be used to estimate the time it to sell the property for a real estate agency. Several IF-THEN rules are extracted from this module. Extracted rules can be used by real estate agencies as well as buyers and sellers of residential units to make better decisions in their investments. In conclusion section, a number of suggestions for future studies are presented. For example, machine learning algorithms such as neural networks, decision trees, etc. can also be used to predict the duration of residential units' sale.

Keywords *Data Mining, Real Estate Market, FP-Growth Algorithm, Association Rule Mining, Fuzzy Inference System*

Received: 11/05/2021 Accepted: 26/08/2021

[1.2.] Department of Industrial Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran
[2.] Corresponding author: farshidabdi@azad.ac.ir

Introduction

The housing sector is a one of the leading sectors in the economy with a major impact on economic fluctuations by creating economic growth and employment in the construction sector and other related parts of the building sector. Housing is an immovable, durable and costly commodity, and in some cases, it is a source of income, saving money or increasing the income. The housing is a very attractive market due to its unique features, such as high working capital and reliable profits for investment. Banks and other financial institutions also invest heavily in residential real estate markets (Carrillo, 2013). Real estate is an important indicator of economic development. Modern buildings are signs of economic progress. Housing as an economic commodity includes features that distinguish it from other products. On the one hand, it is a consuming product which is one of the vital needs of humans and considered the most expensive necessary product of household. On the other hand, as an immovable enduring product, it is capital goods which is an important investment option in many countries (Zhu et al., 2016) and could be the biggest assets of a family and even attractive for economic corporations. The decision to buy a house is often the most important financial decision in each family requiring enough time and effort to evaluate all available choices and find appropriate criterions to analyze the choices and finally select the best one. Having knowledge about duration of real estates' sale and knowing the effective factors on it, can help both real estate agents and buyers to make right investment decisions. Through such information and having a clear vision about the duration of real estates' sale, people can make the best decision whether to sell the residential and commercial units or invest in real estate market in the coming months. In addition, given the high amount of cash required to purchase a residential unit, this financial decision is of great importance (Carrillo, 2013). A fundamental feature differentiating real estate from financial assets is the illiquidity. The real estate illiquidity can be determined by the required time to sell a property. The required time to sell the property may be affected by various factors and which are not under full control of the seller. So real estate investors are faced with uncertainty about price as well as required time for selling (Cheng, Lin, and Liu, 2008). Understanding the factors influencing the time required for selling residential units can lead to an informed decision making and increased profits. As a result, suppliers and active applicants in this market are trying to optimize their decisions by using a set of factors that influence the selection of residential units in this market.One of the solutions that can be used to make proper use of investment opportunities is to predict the sales time of residential units. Due to multiplicity of factors affecting the length of the sales time of a residential unit, determining the factors affecting this period seems critical.Increased data and the development of storage technology have generated huge amounts of data. In the housing industry, precise data is also being produced by real estate agencies in very large volumes and low prices which is stored in databases. By proper use of the stored knowledge available in the housing market, the customer satisfaction can be increased along with significant economic savings for both customers and suppliers. Data mining is an efficient tool for extracting knowledge from these data. Some of data mining methods include classification, clustering, and associations rule mining (Ngai, Xiu and Chau, 2009). Association rule mining, one of the most important and most common techniques of data mining, was first introduced by Agrawal, Imielinski, and Swami (1993). Association rule mining algorithms extract interesting correlations, frequent patterns, associations or casual structures among sets of items in large databases. In this paper, association rules mining have been used to extract rules to predict the required time for selling real estates.Several rules have been extracted using FP-growth algorithms. The represented rules specify the residential units with sales time less than 30 days. Afterwards, generated rules are examined based on relevant evaluation criteria and ultimately the best rules are selected. Furthermore, fuzzy inference systems (FIS) have provided the means to solve various problems in the field of real estate. In this study, the fuzzy inference system is used to predict waiting time for selling residential units. Based on historical data, the proposed FIS determines the correlations between the factors that influence the waiting time for selling residential units, and provide the opportunity to estimate the selling duration of real estates, using “IF...THE...” rules. The rest of the paper is organized as follows. In Section 2, a literature review on the subject of the application of data mining techniques in real estate industry and application of association rule mining in various industries is presented. In Section 3, the concepts of association rule mining are discussed. In Section 4 proposed methodology of the research is described. In Section 5, the results are represented and discussed. Finally, in Section 6 conclusion of the study is provided.

Background

In previous studies, data mining in real estate has been widely used and each of which has addressed a different aspect of real estate marketing. Bahia (2013) proposed a model for predicting the value of housing using artificial neural networks and applied Cascade Forward Back Propagation (CFBP) and Feed Forward Back Propagation (FFBP) methods for modeling.Banerji and Saxena (2012) analyzed the data set for real estate, including 5821 records and 43 features and applied statistical techniques and association rules to remove the noise and clean the data prior to traditional classification to achieve more accurate results. Du et al. (2014) reviewed the application of Big Data in real estate market in China.Nguyen, and Cripps (2001) compared the two models of Artificial Neural Networks (ANN) and Multiple Regression.Jaen (2002) applied data mining techniques such as neural networks for developing an intelligent model and predicting real estate property values based on a variety of factors. Chiarazzo et al. (2014) proposed a model based on Artificial Neural Network (ANN) and applied it to real estate appraisal. Chica-Olmo (2007) estimated housing location price using kriging methods, isotopic data cokriging, and heterotopic data kriging methods. Juan et al. (2006) presented a system to support decision making in housing customization using a hybrid approach combining case-based reasoning (CBR) and genetic algorithm (GA). Park and Bae (2015) developed a housing price prediction model based on machine learning algorithms to help house seller or a real estate agent make better informed decisions. Li and Chu (2017) applied the radial basis function neural network (RBF) and back propagation neural network (BPF) neural networks to model the correlation function between macro-economic parameters variation and house price index variation.Ghosalkar and Dhage (2018) considered the financial plans and needs of customers and proposed a model to predict house prices using Linear Regression. Neloy et al. (2019) proposed apartment rental price prediction model using a hybrid approach combining ensemble learning and advanced regression techniques. Fabozzi et al. (2019) proposed a model to detect bubbles in real estate markets using stepwise regression, ridge regression, lasso, bridge regression. Yu et al. (2020) constructed a real estate pricing model using integration of data mining technology and machine learning technology. García-Magariño (2020) proposed a model to estimate the missing prices in real-estate market using machine learning and dimensionality reduction methods. Singh et al. (2020) identified the variables affecting housing prices using several machine learning algorithms such as LASSO regression, random forest, and gradient boosting. Dambon et al. (2021) presented a maximum likelihood estimation (MLE) approach for Gaussian process-based spatially varying coefficient (SVC) models with an application to real estate price prediction. Sharma et al. (2020) applied supervised machine learning algorithms to real estate price's forecasting. Kamara et al. (2020) addressed the problem of Days on Market (DOM) prediction in real estate industry and proposed a hybrid deep learning model called HNPD. Convolutional neural network-based attention (CNNA), and bidirectional long-short term memory (BLSTM) applied for feature extraction. The proposed method was compared to several learning algorithms such as Linear Regression, Ridge, Lasso, Location-specific Linear Regression, Decision Tree, Random Forest, and Support Vector Regression

(SVR).Castelli et al. (2020) developed a model to predict the days on market variable by applying several algorithms, in particular, Lasso, Ridge, and Elastic Net regressions and Neural Networks. Accordingly, in previous studies, association rules mining methods have been rarely used in the field of real estate. In this paper, several rules are extracted using the association rule mining techniques, determining the factors affecting the waiting time for selling residential units. In the following, a brief review of previous studies on the application of association rule mining in various industries is represented. Having proper information about duration of real estates' sale and also an accurate prediction about its future, surely can be profitable via a correct investment. In recent years by extending the data mining techniques, the researchers have tried to use such methods in different fields, such as predicting the issues related to real estates and property. Data mining is used to extract the desired information and knowledge from the huge size of the data stored in the database. Data mining has various techniques such as association rule mining for the extraction of data. Association rule mining is a technique used for finding relation between hidden patterns of the large database and also for discovering the Simultaneous occurrence of items and constructing the relations among them (Narvekar and Syed, 2015). Several methods have been proposed in the previous studies. Two of the most common algorithms for discovering the frequent item sets are Apriori (Agrawal and Srikant, 1994) and FP-Growth (Lin, Liao, and Chen, 2011). Frequent pattern growth (FP growth) algorithm is one of the most efficient algorithm in finding out the desired association rules. This algorithm Firstly generates a huge number of conditional FP trees and then finds out the frequency of the frequent item sets to gain the desired association rules. Association rule mining algorithms have been used in various fields and industries. In the following, some researches that have used this algorithm are mentioned. Cheng, Lin, and Leu (2010) utilized a well-known association rule mining algorithm i.e. apriori algorithm in order to extract cause-effect relationships in occupational accidents in the Taiwan construction industry. Aribowo and Cahyana (2015) applied Weighted Itemset Tidset (WIT)–tree methods to extract rules and identify characteristic patterns of prospective lenders. The obtained results of the proposed model can reduce the risk of bad debts. Kuo, Lin, and Shih (2007) proposed a two-stage clustering-association rule mining method. In the first stage, the ant system-based clustering algorithm (ASCA) is applied to group database and association rules mining algorithm is employed to extract the useful rules for each group. Tsai and Chen (2004) have used association rules mining algorithm in the customer databases and transaction databases. They identified all the large item sets from the transaction database and generated association rules from the customer database and the large item sets identified. Chiang (2011) proposed a method for mining association rules of customer value. Dataset of online shopping industry in Taiwan was utilized and Supervised Apriori algorithm was employed with customer values to create association rules. Korczak and Skrzypczak (2012) applied FP-Growth algorithm in discovery of customer patterns. For this purpose, the authors have used a real database including transactions of e-shop customers. Czibula, Marian, and Czibula (2014) addressed the problem of defect prediction and proposed a model based on relational association rules mining for predicting software entities that are likely to be defective. Singh et al. (2014) focused on discovering frequent pattern of website usage from server log files by fetching, processing efficiency, and memory size etc. The authors applied the collected data from server log data and compared the performance of apriori and FP-growth algorithm. Ghousi (2015) applied a well-known association rule mining algorithm i.e. apriori algorithm in order to extract relationships between the several fields of dataset of the work accidents in an Iranian manufacture. Arincy, and Sitanggang (2015) applied FP-growth and Eclat algorithms in forest fire and land. They found association and pattern of hotspot occurrences.Chang et al. (2016) addressed the problem of mining association rules in dynamic huge data set. They compared FP-growth with other algorithm and the result showed significant reduction in execution time of incremental updating frequent item sets. Goel and Goel (2017) designed an algorithm for FP Growth technique by considering the repeating patterns in the frequent item sets. The authors have stated that Implementing FP growth using the tries data structure improve the performance of association rule mining undoubtedly.Rachburee et al. (2018) apllied apriori algorithm and FP-growth to discover association rules mining from maintenance transaction log of ATM maintenance.Feng et al. (2019) proposed an expert recommendation algorithm based on Pearson correlation coefficient and FP-growth. This research, gives an idea about how the interesting patterns are generated from the large databases using association rule mining (ARM) methodologies. In this study ARM have been used to solve one of the most important issues in real estate industry, namely, forecasting the waiting time for selling residential units and specify the characteristics of residential units which their sales time is less than 30 days. Liu Zhang, and Wu (2006) proposed a fuzzy neural network model based on hedonic price theory to estimate the appropriate price level for a new real estate. Król et al. (2007) proposed two fuzzy models i.e. Mamdani-type and Takagi-Sugeno-Kang-type for real estate appraisal. The proposed fuzzy models include seven input variables referring to main attributes of a property being appraised such as 'Distance', 'Front', 'Area', 'Infrastructure', 'Arrangement', 'Neighborhood' and 'Communication'. In order to generate the rule bases for both models, Pittsburgh approach has been applied. Guan, Zurada, Levitan (2008) have used the dataset of houses sales in the United States and proposed an adaptive neuro-fuzzy inference system-based to estimate prices for residential properties. Kusan, Aytekin, and Özdemir (2010) developed a fuzzy logic system for prediction of the selling price of house-building. Zhang and Yang (2012) stated that real estate is a high risk economy, therefore they considered real estate investment risk analysis and proposed a fuzzy comprehensive evaluation method for real estate investment risk assessment. Guan et al. (2014) proposed an adaptive, neuro-fuzzy inference model, and used it for real estate property price prediction.

Trawiński et al. (2014) applied the ensembles of genetic fuzzy systems to build reliable predictive models based on data stream of real estate transactions. Gerek (2014) utilized two different adaptive neuro-fuzzy (ANFIS) approaches for the estimation of house selling price. Considering the fact that initial investment in real estate is of great importance, Mao and Wu (2011) integrated the fuzzy risk assessment of investment income and cost with the calculation of relevant parameters in fuzzy real option pricing model, and achieved more reasonable and realistic results. Lasota et al. (2011) applied evolving Takagi-Sugeno algorithm (eTS) to build models for property valuation on the basis of real-world data. Evolving intelligent systems are self-developing and self-learning systems that combine intelligent systems with the on-line learning algorithms in order to extract knowledge from data. Azadeh, Ziaei, and Moghaddam (2012) presented a hybrid algorithm based on fuzzy linear regression (FLR) and fuzzy cognitive map (FCM) to solve the problem of forecasting and optimization of housing market fluctuations. Sarip, Hafez, and Nasir Daud (2016) proposed a fuzzy least-squares regression-based (FLSR) model to predict the prices of real estates. Febrita et al. (2017) applied fuzzy inference system to extract rules for predicting house prices. Table 1 summarizes the reviewed researches.

Table 1.

Summary of Reviewed Research

Researcher (s)	**Main Objective**	**Applied method**
Banerji & Saxena (2012)	Applied association rules real estate actual data to remove the noise and clean the data prior to traditional classification to achieve more accurate results.	Association rule mining and classification algorithms
Bahia (2013)	Proposed a model to predict the value of housing.	Cascade Forward Back Propagation (CFBP) and Feed Forward Back Propagation (FFBP)
Chiarazzo et al. (2014)	Proposed a model to real estate appraisal.	Artificial Neural Network (ANN)

Researcher (s)	Main Objective	Applied method
Park and Bae (2015)	Developed a housing price prediction model.	C4.5, RIPPER, Bayesian, and AdaBoost
Li and Chu (2017)	Proposed a model to show the correlation function between macro-economic parameters variation and house price index variation.	radial basis function neural network (RBF) and back propagation neural network (BPF) neural networks
Ghosalkar and Dhage (2018)	Proposed a model to predict house prices	Linear Regression
Neloy et al. (2019)	Developed a model to apartment rental price prediction.	A hybrid approach combining ensemble learning and advanced regression techniques.
Fabozzi et al. (2019)	Proposed a model to detect bubbles in real estate markets.	stepwise regression, ridge regression, lasso, bridge regression
Yu et al. (2020)	Constructed a real estate pricing model.	data mining and machine learning
García-Magariño (2020)	Proposed a model to estimate the missing prices in real-estate market.	linear regression, support vector regression, the k-nearest neighbors and a multilayer perceptron neural network
Singh et al. (2020)	Identified the variables affecting housing prices.	LASSO regression, random forest, and gradient boosting
Sharma et al. (2020)	Proposed a model to real estate price's forecasting.	decision tree and random forest
Dambon et al. (2021)	Proposed a model to real estate price prediction.	maximum likelihood estimation (MLE) approach for Gaussian process-based spatially varying coefficient (SVC) models
Kusan, Aytekin, and Özdemir (2010)	Proposed a model to predict the selling price of house-building	fuzzy logic system
Azadeh, Ziaei, and Moghaddam (2012)	forecasting and optimization of housing market fluctuations	fuzzy linear regression (FLR) and fuzzy cognitive map (FCM)
Guan et al. (2014)	Developed a real estate property price prediction	neuro-fuzzy inference model
Gerek (2014)	Proposed a model to estimate house selling price	neuro-fuzzy (ANFIS) approaches

Researcher (s)	Main Objective	Applied method
Sarip, Hafez, and Nasir Daud (2016)	Developed a model to predict the prices of real estates	fuzzy least-squares regression-based (FLSR) model
Febrita et al. (2017)	Proposed a fuzzy inference system to extract rules for predicting house prices.	fuzzy inference system
Kamara et al. (2020)	Days on Market (DOM) prediction in real estate industry	Linear Regression, Ridge, Lasso, Location-specific Linear Regression, Decision Tree, Random Forest, and Support Vector Regression (SVR).
Castelli et al. (2020)	Developed a model to predict the days on market variables.	Lasso, Ridge, and Elastic Net regressions and Neural Networks
This study	Mining a Set of Rules for Determining the Waiting Time for Selling Residential Units	FP-Growth, Fuzzy inference system

FP-growth method is faster than other methods of extracting rule. This algorithm reduces the total number of candidate item sets by generating a condensed version of the database in terms of a FP-tree. The algorithm consists of two steps: Firstly, converting a large database into a compact, Frequent- Pattern tree (FP-tree) structure. Afterwards, creating a useful FP-tree-based frequent pattern mining method."Support" and "confidence" are the two well-known rule evaluation metrics. Support determines the grade and level of the coverage of $(X \Rightarrow Y)$, while confidence determines how frequently items in Y appear in transactions that contain X. Support and confidence formula can be expressed as (1) and (2), respectively. (Doostan, and Chowdhury, 2017)

$$Support(X \Rightarrow Y) = \frac{\sigma(X,Y)}{N} \tag{1}$$

$$Confidence(X \Rightarrow Y) = \frac{\sigma(X,Y)}{\sigma(X)} \tag{2}$$

Where σ is summation notation, and N represents the total number of all transactions. In addition, there are other metrics, one of which, is “lift” value of a rule. The lift value can be stated as (3)

$$Lift(X \rightarrow Y) = \frac{Support(X \Rightarrow Y)}{Support(Y) \times Support(X)} \quad (3)$$

Fuzzy set theory was put forward by Professor Lotfizadeh (1965) to solve the problem of uncertainty. At the kernel of a fuzzy model lies the fuzzy inference engine that runs the implication mechanism. According to fuzzy logic, based on the limited and inaccurate information that exist from a phenomenon, it will be possible to identify and predict the relationship between variables (Alahyari and Pilevari, 2020). The term "verbal variables" refers to variables whose values are verbal expressions instead of numerical values. Inference requires rules that are used to strike a relationship between input variables and the output variables. These rules are defined as "IF… THEN…" and are used to describe a given system based on verbal variables rather than mathematical formulas. The IF section is known as the antecedent and the THEN section is known as the consequent. The number of rules depends on the number of inputs and outputs, as well as the desired behavior of the system (Zadeh, 1965).

Method

The dataset used in this article deals with the information of a real estate database in Iran including 6836 records. It consists of 34 predicting variables and 1 target variable which is about the waiting time for selling a residential unit. Table 2 shows the variables used in the research.

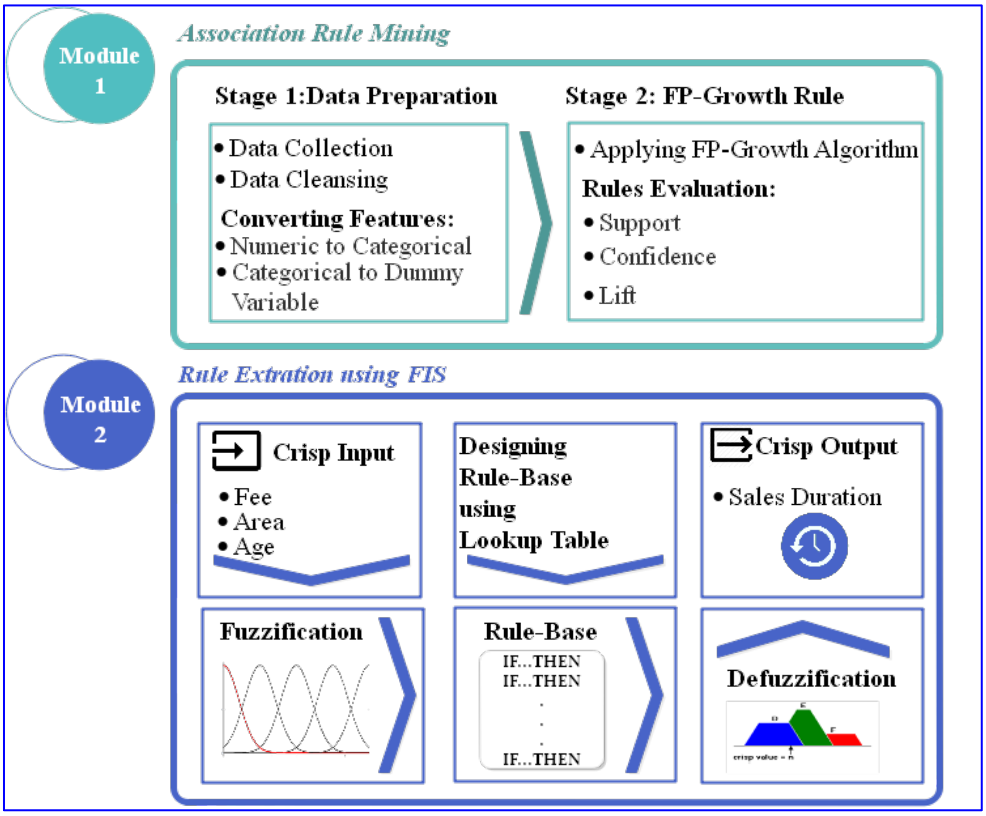
Table 2.
Research Variables

Row	Variables	Description	Type: Range/Category
1	Fee	The price of property	Numeric
2	Owner/ Real estate	The person who registered the property advertisement	Binominal: Owner, Real estate
3	Area	Property area	Numeric
4	Point	The location of apartment	Nominal (1 to 9): Northern, Southern, Eastern, Western, Northern corner, Southern corner, three open sides, Northern-Southern, Eastern - Western
5	Front	The facade of the apartment	Nominal (1 to 9): Brick, Stone, Roman stone, English Stone, ashlar, Granite, Glass, composite, Cement
6	Heating	Type of heating system	Nominal (1 to 6): Split, Gas heaters, Package, Chiller, Central heating radiator, Air Conditioner
7	Cooling	Type of cooling system	Nominal (1 to 6): Water-Cooled Air Conditioner, Gas-Cooled Air Conditioner, Split, Package, Chiller, Air Conditioner
8	Floor	Type of materials used in floor	Nominal (1 to 8): Parquet, Ceramic, Stone, Cement, Granite, Laminate, Carpet, Mosaic
9	Room Floor	Type of materials used in floor of rooms	Nominal (1 to 8): Parquet, Ceramic, Stone, Cement, Granite, Laminate, Carpet, Mosaic
10	Cabinet	Type of Kitchen cabinet in each unit	Nominal (1 to 8): MDF, HDF, High Glass, HPL, NEFF, Metal Cabinet, Wooden Cabinet, Metal-Wooden Cabinet
11	Region	The area where the property is located	Nominal: 1 to 20
12	Bedroom	Number of bedrooms per unit	Nominal: 0,1,2,3,4,11
13	Age	The age of unit	Numeric
14	No Floor	The floor where the property is located	Nominal
15	All Floor	Total number of floors available in the apartment	Nominal
16	Courtyard	Courtyard available in house	Nominal: 1 to 4
17	Unit	The unit number	Nominal: 1to 15
18	Parking	The unit has a parking or not?	Binominal: 1, 0

19	Remote	Parking door has remote control or not?	Binominal: 1, 0
20	Storage	Storage available in selling unit	Binominal: 1, 0
21	Elevator	Elevator available in apartment	Binominal: 1, 0
22	Terrace	Terrace available in selling unit	Binominal: 1, 0
23	Toilet	Toilet Conditions	Binominal: 1, 0
24	Reconstruct	Reconstructing the house if it's too old	Binominal: 1, 0
25	Vacate	Apartment condition on selling	Binominal: 1, 0
26	Unutilized	Unutilized and new unit	Binominal: 1, 0
27	Kitchen	Type of kitchen available in unit	Binominal: 1, 0
28	Grate	Fireplace available in unit	Binominal: 1, 0
29	Fascia	Fascia on walls and roofs	Binominal: 1, 0
30	Paint	Painting the walls and roofs	Binominal: 1, 0
31	Light	Hidden light	Binominal: 1, 0
32	IPhone	Available video door-phone	Binominal: 1, 0
33	Antenna	Central antenna	Binominal: 1, 0
34	Furnish	Equipment like furniture and curtain	Binominal: 1, 0
35	Duration	*Target Variable:* The time taken from ad registration to the sale of the desired property	Numeric

The process of the research is illustrated in Figure 1. As can be seen, the research framework consists of two modules. The first module addresses association rule mining containing two stages. The first stage of module 1, is allocated to data preparation. This stage includes 1- data collection and data cleansing 2- converting numerical variables into categorical factors and 3- Converting categorical factors into dummy variables. Afterwards and in the second stage, a well-known association rule mining algorithm i.e. FP-Growth is utilized to extract appropriate rules in order to determine the waiting time for selling residential units. Evaluation metrics such as support, confidence, and lift are applied to select the best set of rules. Extracted rules can be used by real estate agencies as well as buyers to make better decisions in their investments. The second module addresses constructing fuzzy inference system. This module contains three main parts, including a 1-Fuzzification 2- Designing a rule database using lookup table and fuzzy inference engine, and 3- Deffuzification.

Figure 1.
Research Process



Findings

An essential stage in performing association rule mining for predicting the waiting time for selling residential units is to collect the considerable amount of real estate data. Such data can be gathered via various sources such as Real estate databases. Real-world real estate datasets usually contain a significant amount of useless variables missing values, etc. the data were cleansed to detect and remove useless variables, so the number of continuous variables with zero standard deviations and also discrete variables which more than 90 percent of their values are the same have been removed. After removing useless variables, the number of variables is reduced from 35 to 23. The missing value may also affect the data mining process, therefore missing values were also handled by replacing by average.Real data sets often contain various features that have different types. Some of the features can be continuous, while the others are categorical. In order to develop the process of research, it is necessary to convert all continues features into categorical ones. To implement this task, numerical features can be discretized into different groups and then represented by group number. For instance, to transform the "Fee" into categorical factor, it can be categorized into 5 groups as range [-∞ - 135015000], [135015000 – 175250000], [175250000 – 224425000], [224425000 – 314750000], [314750000 - ∞] and then represented by categorical factors of 1-5, respectively. "Area" variable also can be categorized into 5 groups as range [-∞ - 46.500], [46.500 - 55.500], [55.500 - 64.500], [64.500 - 77.500], and [77.500 - ∞] and then represented by categorical factors of 1-5, respectively. The “Duration” variable is also transform to a nominal variable. Classes 1, 2, and 3 were assigned to records in the data set in which the duration variable was less than thirty days, between thirty to 60 days, and between 60 and 300 days, respectively. At this stage, all predictive attributes whose type is not binominal are converted into a binominal variable. This not only changes the type of selected attributes but also maps all values of these attributes to binominal values i.e. true and false.For example, if a nominal attribute with name 'BedRoom' and possible nominal values '1', '2', '3', '4', '5' and '11' is transformed, the result is a set of five binominal attributes 'BedRoom = 1', 'BedRoom = 2', 'BedRoom = 3', 'BedRoom = 4', 'BedRoom = 5' and 'BedRoom = 11'. Only the value of one of these attributes is true for a specific example, the value of the other attributes is false. Table 3 represents this example.

Table 3.
Changing the Type of Nominal Attributes to a Binominal Type

Row	BedRoom	BedRoom=1	BedRoom=2	BedRoom=3	BedRoom=4	BedRoom=5	BedRoom=11
1	2	False	True	False	False	False	False
2	1	True	False	False	False	False	False
3	3	False	False	True	False	False	False
4	11	False	False	False	False	False	True
5	2	False	True	False	False	False	False
6	1	True	False	False	False	False	False
7	3	False	False	True	False	False	False
8	4	False	False	False	True	False	False
9	2	False	True	False	False	False	False
10	5	False	False	False	False	True	False

In order to perform association rule analysis, the FP-Growth algorithm is utilized. Initially, the value of minimum support and minimum confidence are determined. Minimum support depends on the number of samples and the number of rules to be extracted. The process of association rule mining contains two main stages including Generating frequent item sets and generating rules from frequent item sets. Results of association rule mining are represented in Table 4.

Table 4.
Several Extracted Rules

Row	IF	Then	Support	Confidence	Lift
1	Cabin = MDF, Elevator	Duration= less than 30 days	0.157	0.404	1.569
2	Front = Stone, Elevator	Duration= less than 30 days	0.155	0.405	1.572
3	Cooling = Water-Cooled Air Conditioner, Elevator	Duration= less than 30 days	0.154	0.408	1.583
4	Front = Stone, Cabin = MDF, Elevator	Duration= less than 30 days	0.152	0.409	1.590
5	Cooling = Water-Cooled Air Conditioner, Cabin = MDF, Elevator	Duration= less than 30 days	0.150	0.412	1.602
6	Cooling = Water-Cooled Air Conditioner, Front = Stone, Elevator	Duration= less than 30 days	0.148	0.414	1.608
7	Heating = Central heating radiator, Elevator	Duration= less than 30 days	0.140	0.473	1.837
8	Cooling = Water-Cooled Air Conditioner, Front = Stone, Cabin = MDF, Heating = Central heating radiator, Elevator	Duration= less than 30 days	0.130	0.485	1.883
9	Cabin = MDF, Point = Nourthern, Elevator	Duration= less than 30 days	0.138	0.491	1.908
10	Front = Stone, Iphone, Elevator	Duration= less than 30 days	0.105	0.537	2.087
11	Cabin = MDF, Iphone, Elevator	Duration= less than 30 days	0.106	0.539	2.093
12	Floor = Ceramic, Elevator	Duration= less than 30 days	0.144	0.543	2.107
13	RoomFloor = Ceramic, Elevator	Duration= less than 30 days	0.144	0.540	2.099
14	Cooling = Water-Cooled Air Conditioner, Floor = Ceramic, Elevator	Duration= less than 30 days	0.138	0.543	2.109
15	Front = Stone, RoomFloor = Ceramic, Elevator	Duration= less than 30 days	0.140	0.547	2.125
16	Room Floor = Ceramic, Floor = Ceramic, Elevator	Duration= less than 30 days	0.142	0.549	2.132

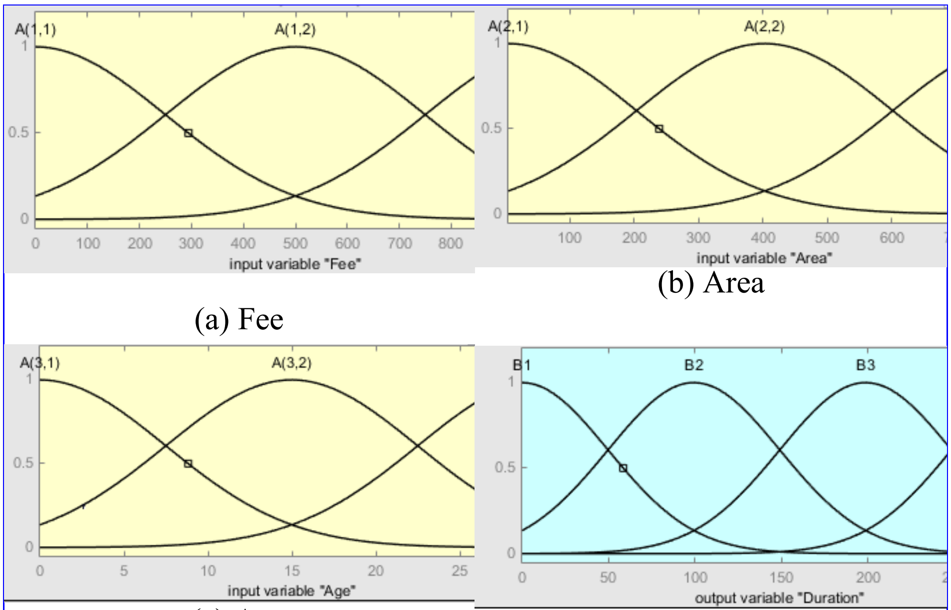
Explaining the first rule, we can say if a residential unit has an MDF cabinet and elevator, then it will be sold in less than 30 days. The support, confidence and lift of this rule are 15%, 40% and 1.569, respectively. According to the second rule, if the facade of the residential unit is made of stone and has an elevator then it will be sold in less than 30 days. The support, confidence, and lift of this rule are 15%, 40% and 1.572, respectively. Rule number eight states that if the residential unit has a Water-Cooled Air Conditioner, a stone façade, MDF cabinet, central heating radiator, and elevator, then it is sold in less than 30 days. The support, confidence and lift of this law are 13%, 48% and 1.883 respectively. The last law (No. 16) states that if the material used for the floor of the building is Ceramic and the building has an elevator, then it will be sold in less than 30 days. The rest of the rules are also described in the same way. A remarkable point in all the extracted rules is the importance of the existence of an elevator in a residential unit. Therefore, the existence of an elevator is of great importance in reducing the required time for selling a residential unit. Having a Water-Cooled Air Conditioner, a stone façade, and MDF cabinet, are also some of the advantages that reduce the waiting time for selling a residential unit. In this section, the results of applying the second module are presented. In this module, construction of fuzzy inference system and extraction of several IF-THEN rules are discussed. “Fee”, “Area”, “Age” and “Duration” variables are used in this module. The input and output factors of the proposed FIS of the paper are presented in Table 5.

Table 5.
Inputs/Outputs of Proposed FIS

Inputs	Fee (F)
	Area (Ar.)
	Age (Ag.)
Output	Duration (D)

In the first stage of the model, four fuzzy sets of membership functions are applied for both inputs and outputs. All inputs are modelled with three membership functions with linguistic terms {Low, Medium, High} and the output is modelled with four linguistic terms including "Low", "Medium", "High" and "very high" as shown in Figure 2.

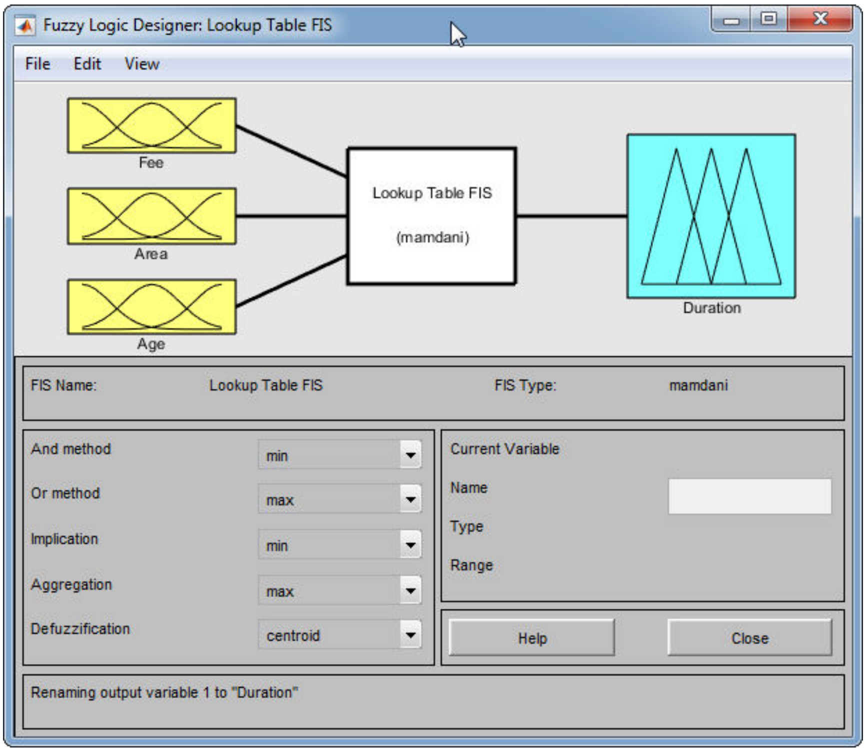
Figure 2.
Defining Gaussian Membership Functions for the Inputs and Output



(a) Fee (b) Area (c) Age (d) Duration

Figure 3, indicates the fuzzy system and the number of inputs and output of the system. Fuzzification transforms the feature value of input variables into proper linguistic fuzzy information. Fuzzy rule base stores the knowledge and rules for deriving the outputs. These rules are expressed in the If–Then format. Mamdani max–min is used for inference mechanism and centroid method is used to defuzzification which is responsible for transforming the fuzzy results into crisp value.

Figure 3.
FIS Designer in Matlab



The linguistic variables and their corresponding values are presented in Table 6.

Table 6.
The Linguistic Variables and their Corresponding Values for each Input and Output Variables

Input Variable	**Linguistic term**	**Corresponding Fuzzy Interval**
Fee (F)	Low	[250 0]
	Medium	[250 500]
	High	[250 1000]
Area (Ar.)	Low	[198.8 12]
	Medium	[198.8 402.5]
	High	[198.8 800]

Input Variable	Linguistic term	Corresponding Fuzzy Interval
Age (Ag.)	Low	[7.5 0]
	Medium	[7.5 15]
	High	[7.5 30]
Duration (D)	Low	[49.67 0]
	Medium	[49.67 99.33]
	High	[49.67 198.7]
	Very High	[49.67 298]

In here, the rule-base of FIS is provided using lookup table. If-then-fuzzy rules are applied to determine the level of waiting time for selling a residential unit is presented in Eq. (1). (Khalili-Damghani et al., 2013)

$$IF\ x_{1j} = F_j\ is\ A_{1j}\ and\ x_{2j} = Ar_j\ \ is\ A_{2j}\ and\ x_{3j} = Ag_j\ \ is\ A_{3j} \quad (1)$$
$$THEN\ D_j = B_j\ (j = 1.2. \ldots . n)$$

Where $A_j = \left(A_{1j}.A_{2j}.A_{3j}.j = 1.2.3\right)\ and\ B_j\ (j = 1.2.3.4)$are fuzzy sets and called the antecedent and consequent of rule j , respectively. Table 7 presents the defined rules.

Table 7.

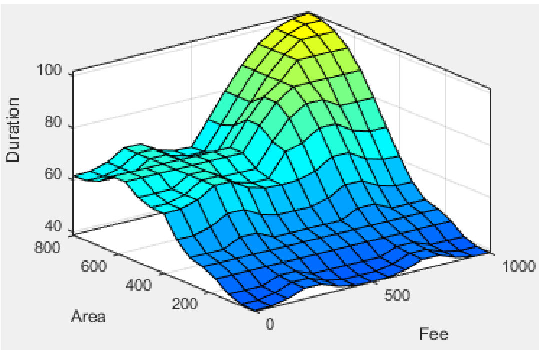
IF – Then Rules Extracted from the Proposed FIS

Row	IF (Antecedence)			THEN (Consequences)
	Fee (F)	Area (Ar.)	Age (Ag.)	Duration (D)
1	Low	Low	Low	Low
2	medium	Low	Low	Low
3	Low	medium	Low	Low
4	medium	medium	Low	Low
5	High	medium	Low	Low
6	Low	High	Low	Low
7	medium	High	Low	Medium
8	High	High	Low	Medium
9	Low	Low	medium	Low
10	medium	Low	medium	Low
11	High	Low	medium	Low
12	Low	medium	medium	Low
13	medium	medium	medium	Low

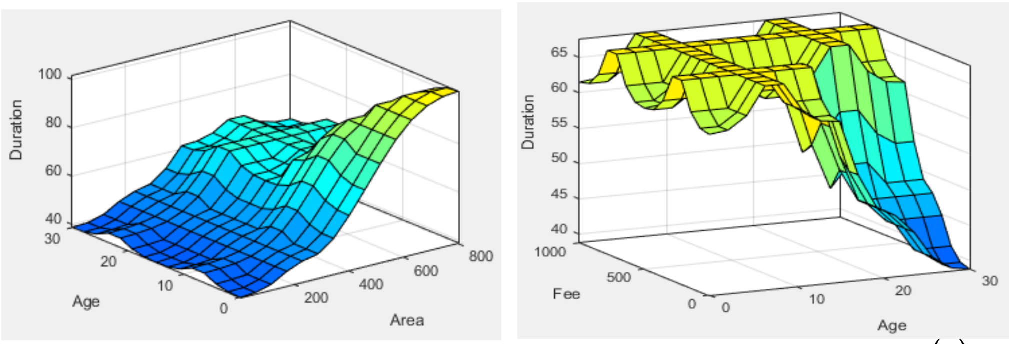
Row	IF (Antecedence)			THEN (Consequences)
	Fee (F)	Area (Ar.)	Age (Ag.)	Duration (D)
14	High	medium	medium	Low
15	Low	High	medium	Low
16	medium	High	medium	Low
17	High	High	medium	Medium
18	High	Low	High	Low
19	Low	medium	High	Low
21	medium	medium	High	Low
22	Low	High	High	Low

In the following, three-dimensional graphs are provided for further interpretation. Figure 4 presented the relationships between inputs and the output. Axis Z denotes the output, and axes X and Y denote the inputs.

Figure 4.
The Surface View of the FIS



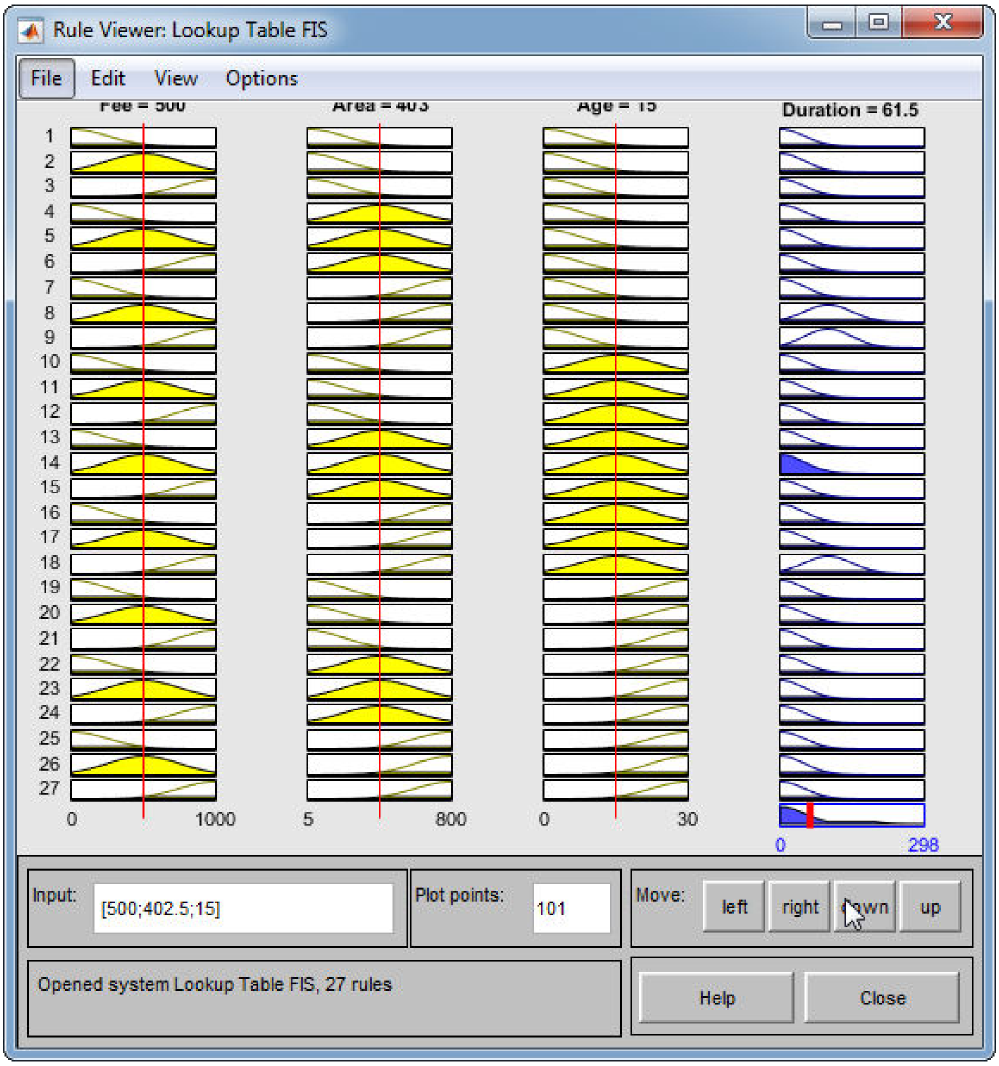
(a)



(b) (c)

Figure 5 illustrates the graphical representation of the rules of a designed fuzzy system, in which the instantaneous output of the system can be controlled and checked by changing the input values at the same time.

Figure 5.
Rule Viewer of the FIS



Discussion and Conclusion

Housing in Iran is one of the areas with numerous problems. In addition to macroeconomic factors, the existence of hardware barriers such as lack of production of some construction materials and software factors are also involved in disruption of housing market. These software factors are a set of factors that are all rooted in the inappropriate use of data in this area. With the development of information technology and the possibility of storing data related to the field of housing, the basis for the application of data mining techniques has been provided. In this research, using the method of association rule mining and also by presenting a fuzzy inference system, the factors affecting the sales time of residential units in the study data set have been investigated. In the present study, Determining the waiting time for selling residential units has been dealt with using the available features of the housing unit. These characteristics refer to variables that are specific to a particular apartment or residential unit. These variables fall into two main subgroups. The first group of variables has a quantitative nature such as total area, number of bedrooms, number of toilets, number of parking lots, etc. The second group refers to the features with qualitative nature, such as the arrangement of the equipment, the condition of the building exterior, heating system, etc. Also, the age of the building is a variable that has been discussed in research both quantitatively and qualitatively. The results of the first module of the research showed that the presence of an elevator has a great impact on the time required for selling a house. This variable is observed in almost all extraction laws. Moreover, having a stone facade, central heating radiator, water-cooled air conditioning, and MDF cabinets have been among the factors affecting the sale of residential units in less than a month. In addition, the average age of construction of the studied residential units, which have been sold in less than a month, is about five years, and the area of these residential units has been about 65 square meters. It should be also noted that in addition to the mentioned features that are related to the building, some distance variables such as distance to work, distance to school, distance to supermarket, distance to the city center, access to the bus station, etc. can have a significant impact on the value of the residential unit and its sales time. Also, environmental variables such as the attractiveness of the district, the quality of the surrounding houses, having attractive landscapes, etc. can be effective. Finally, it should be noted that housing market developments are certainly affected by the macroeconomic structure of each country, and macroeconomic factors such as interest rates, GDP, population, exchange rates, and household income also play a decisive role in the real estate market. Macroeconomic variables are important due to having a significant impact on real estate properties prices and the price itself affects the duration of residential units' sale. Government policies and legislation, including tax incentives, deductions, and subsidies can boost demand for real estate. Also, the other factors impacting real estate supply and demand is seasonality. For example the holiday season and academic year influence the supply and demand of real estate market. Families with children typically would not like to uproot their family in the middle of the academic year and typically will wait until its end. The busiest moving times of the year occur during the summer. People are more likely to refer to the real estate market at the end of the academic year meaning duration of residential units' sale is usually reduced in the summer. Being aware about the waiting time for selling real estates and the effective factors on it, will lead to a conscious decision which could help real estate planners to analyze and predict the situation in the future and select the appropriate solutions leading to increased profits. Accordingly, in this paper, a two-module framework was proposed for analyzing the waiting time for selling residential units. The first module deals with implementation of association rule mining. In this module, a well-known association rule mining algorithm namely FP-Growth was applied. The second module addresses constructing fuzzy inference system. In the second module a fuzzy inference system (FIS) was proposed to estimate the time it takes to sell the property for a real estate agency. The proposed framework was applied in a real case study of real estate agency in Iran and extracted rules which indicated the effective factors on the waiting time for selling residential units. The represented rules of the first module specify the characteristics of residential units which their sales time is less than 30 days. The extracted rules indicated that availability of an elevator has a great importance in reducing the required time for selling a residential unit. Also having a Water-Cooled Air Conditioner, a stone façade, and MDF cabinet, are the features that reduce the time period required for selling a residential unit. Extracted rules were examined based on relevant evaluation criteria such as support, confidence, and lift. In the proposed FIS of the second module, "Fee", "Age", and "Area" are considered as input variables and waiting time for selling residential units is taken as an output of fuzzy inference system. By using the proposed FIS, the level of the sales duration can be calculated using relevant information available for each residential unit. Real estate consultants, buyers and sellers of residential units can easily predict the required time to sell a desired residential unit by using the framework proposed in this study. So they can make better decisions regarding their investments. Identifying the waiting time for selling residential units could help Real estate consultants to do more effective marketing activities. Also, it made it possible for the real estate consultants to devote their time to residential units that are more likely to be sold in a shorter period of time.

References

Alahyari, M., & Pilevari, N. (2020). CO-Active Neuro- Fuzzy Inference System Application in Supply Chain Sustainability Assessment Based on Economic, Social, Environmental, and Governance Pillars, Journal of System Management, 6 (3) 265-287

Agrawal R. & Srikant R. (1994). Fast Algorithms for Mining Association Rules, in: Proceedings of the 20th International Conference on Very Large Data Bases, pp. 487-499

Agrawal, R., Imielinski, T., & Swami, A. N. (1993). Mining association rules between sets of items in large databases. In Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, 207-216.

Aribowo, A.S., & Cahyana N.H. (2015). Feasibility study for banking loan using association rule mining classifier, International Journal of Advances in Intelligent Informatics, 1 (1) 41-47.

Arincy, N., Sitanggang, I.S.: Association rules mining on forest fires data using FP-Growth and ECLAT algorithm. In: 2015 3rd International

Conference on Adaptive and Intelligent Agroindustry (ICAIA), pp. 274–277 (2015)

Azadeh, A., Ziaei, B., & Moghaddam, M. (2012). A hybrid fuzzy regression-fuzzy cognitive mapalgorithm for forecasting and optimization of housing market fluctuations. Expert Systems with Applications, 39(1), 298–315.

Bahia, I.S.H, (2013). A Data Mining Model by Using ANN for Predicting Real Estate Market: Comparative Study, International Journal of Intelligence Science, 3, 162-169

Banerji G., & Saxena K. (2012). An Efficient Classification Algorithm for Real Estate domain, International Journal of Modern Engineering Research (IJMER), 2 (4) 2424-2430

BogdanTrawiński B., Smętek M., Lasota T., & Trawińsk G. (2014) Evaluation of Fuzzy System Ensemble Approach to Predict from a Data Stream. In: Nguyen N.T., Attachoo B., Trawiński B., Somboonviwat K. (eds) Intelligent Information and Database Systems. ACIIDS 2014. Lecture Notes in Computer Science, vol 8398. Springer, Cham.

Carrillo P-E. (2013). To Sell or Not to Sell: Measuring the Heat of the Housing Market, Real estate economics, 41, 310–346.

Castelli, M., Dobreva, M., Henriques, R., & Vanneschi, L. (2020). Predicting days on market to optimize real estate sales strategy. Complexity, 2020. https://doi.org/10.1155/2020/4603190

Chang, H.Y., Lin, J.C., Cheng, M.L., Huang, S.C. (2016) A novel incremental data mining algorithm based on FP-growth for Big Data. In: 2016 International Conference on Networking and Network Applications (NaNA), pp. 375–378.

Cheng C-W, Lin C-C, & Leu S-S. (2010). Use of association rules to explore cause–effect relationships in occupational accidents in the Taiwan construction industry, Safety Science 48, 436–444.

Cheng P, Lin Z., & Liu Y. (2008). A Model of Time-on-Market and Real Estate Price Under Sequential Search with Recall, Real estate economics, 36, 813–843

Chiang W-Y (2011). To mine association rules of customer values via a data mining procedure with improved model: An empirical case study, Expert Systems with Applications 38, 1716–1722.

Chiarazzo V, Caggiania L, & Marinelli M. (2014). A Neural Network based model for real estate price estimation considering environmental quality of property location, Transportation Research Procedia. 3, 810 – 817.

Chica-Olmo, J. (2007). Prediction of housing location price by a multivariate spatial method: Cokriging. Journal of Real Estate Research, 29, 92–114.

Czibula G., Maria Z., & Czibula I.G. (2014). Software defect prediction using relational association rule mining, Information Sciences 264, 260–278.

Dambon, J.A., Sigrist, F., & Furrer, R. (2021). Maximum likelihood estimation of spatially varying coefficient models for large data with an application to real estate price prediction, Spatial Statistics, 41, 100470

Doostan M, & Chowdhury B.H. (2017). Power distribution system fault cause analysis by using association rule mining, Electric Power Systems Research 152, 140–147.

Du D., Li A., & Zhang L. (2014). Survey on the Applications of Big Data in Chinese Real Estate Enterprise, Procedia Computer Science, 30, 24 – 33.

Fabozzi, F.J., Kynigakis, I., Panopoulou, E., Tunaru, R.S. (2019). Detecting Bubbles in the US and UK Real Estate Markets. The Journal of Real Estate Finance and Economics, 60, 469–513. https://doi.org/10.1007/s11146-018-9693-9.

Febrita, R. E., Alfiyatin, A. N. Taufiq, H. & Mahmudy, W. F. (2017). Data-driven fuzzy rule extraction for housing price prediction in Malang, East Java, International Conference on Advanced Computer Science and Information Systems (ICACSIS), Bali, 2017, pp. 351-358, doi: 10.1109/ICACSIS.2017.8355058.

Feng, W., Zhu, Q., Zhuang, J., Shimin, Y.. An expert recommendation algorithm based on Pearson correlation coefficient and FP-growth. Cluster Comput 22, 7401–7412 (2019). https://doi.org/10.1007/s10586-017-1576-y

García-Magariño, I., Medrano, C. & Delgado, J. (2020) Estimation of missing prices in real-estate market agent-based simulations with machine learning and dimensionality reduction methods. Neural Computing and Applications, 32, 2665–2682. https://doi.org/10.1007/s00521-018-3938-7

Gerek I.H. (2014). House selling price assessment using two different adaptive neuro-fuzzy techniques, Automation in Construction, 41, 33–39.

Ghosalkar, N.N., & Dhage, S. N. (2018). Real Estate Value Prediction Using Linear Regression, Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018, pp. 1-5, doi: 10.1109/ICCUBEA.2018.8697639.

Ghousi R. (2015). Applying a decision support system for accident analysis by using data mining approach: A case study on one of the Iranian manufactures, Journal of Industrial and Systems Engineering, 8 (3) 59-76.

Goel M., Goel K. (2017) FP-Growth Implementation Using Tries for Association Rule Mining. In: Deep K. et al. (eds) Proceedings of Sixth International Conference on Soft Computing for Problem Solving. Advances in Intelligent Systems and Computing, vol 547. Springer, Singapore. https://doi.org/10.1007/978-981-10-3325-4_3

Guan J, Zurada J, & Levitan A.S. (2008). An adaptive neuro-fuzzy ınference system based approach to real estate property assessment, Journal of Real Estate Research. 30 (4) 395–421.

Guan, J., Shi, D., Zurada, J. M., & Levitan, A. S. (2014). Analyzing Massive Data Sets: An Adaptive Fuzzy Neural Approach for Prediction, with a Real Estate Illustration. In: Journal of Organizational Computing and Electronic Commerce. 24 (1) 94-112.

Jaen, R. D. (2002) Data Mining: An Empirical Application in Real Estate Valuation. In FLAIRS Conference, pp. 314-317.

Juan, Y. K., Shin, S. G., & Perng, Y. H. (2006). Decision support for housing customization: A hybrid approach using case-based reasoning and genetic algorithm. Expert Systems with Applications, 31, 83–93.

Kamara, A.F., Chen, E., Liu, Q, & Pan, Z. (2020). A hybrid neural network for predicting Days on Market a measure of liquidity in real estate industry, Knowledge-Based Systems, 208, 106417

Khalili-Damghani K., Sadi-Nezhad, S., Hosseinzadeh Lotfi, F., & Tavana, M. (2013) A hybrid fuzzy rule-based multi-criteria framework for sustainable project portfolio selection, Information Sciences, 220, 442-462

Korczak J., Skrzypczak P. (2012) FP-Growth in Discovery of Customer Patterns. In: Aberer K., Damiani E., Dillon T. (eds) Data-Driven Process Discovery and Analysis. SIMPDA 2011. Lecture Notes in Business Information Processing, vol 116. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-34044-4_7

Król, D., Lasota T., Trawiński B., & Trawiński K. (2007) Comparison of Mamdani and TSK Fuzzy Models for Real Estate Appraisal. In: Apolloni B., Howlett R.J., Jain L. (eds) Knowledge-Based Intelligent Information and Engineering Systems. KES 2007. Lecture Notes in Computer Science, vol 4694. Springer, Berlin, Heidelberg.

Kusan, H, Aytekin, O, & Ozdemir I. (2010). The use of fuzzy logic in predicting house selling price, Expert Systems with Applications. 37, 1808–1813.

Kuo R.J., Lin S.Y., & Shih C.W. (2007). Mining association rules through integration of clustering analysis and ant colony system for health insurance database in Taiwan, Expert Systems with Applications, 33, 794–808.

Lasota T., Telec Z., Trawiński B., & Trawińsk K. (2011) Investigation of the eTS Evolving Fuzzy Systems Applied to Real Estate Appraisal, Multiple-Valued Logic and Soft Computing, 17, 229-253.

Lin K-C, Liao I-E, & Chen Z-S. (2011). An improved frequent pattern growth method for mining association rules, Expert Systems with Applications 38, 5154–5161.

Li, L., & Chu, K. (2017). Prediction of real estate price variation based on economic parameters, International Conference on Applied System Innovation (ICASI), 2017, pp. 87-90, doi: 10.1109/ICASI.2017.7988353.

Liu, J., Zhang, G. X. L., & Wu, W. P. (2006). Application of fuzzy neural network for real estate prediction. Lecture Notes in Computer Science, 3973, 1187–1191.

Mao Y., & Wu W. (2011) Fuzzy Real Option Evaluation of Real Estate Project Based on Risk Analysis, Systems Engineering Procedia, 1, 228–235

Narvekar, M., Syed, S.F., An optimized algorithm for association rule mining using FP tree, Procedia Computer Science, 2015, 45(C), pp. 101–110.

Neloy AA, Haque HMS, & Ul Islam MM (2019) Ensemble learning based rental apartment price prediction model by categorical features factoring. North South Univ. Res. study, pp. 350–356, 2019.

Nguyen, N., & Cripps, A. (2001). Predicting housing value: a comparison of multiple regression analysis and artificial neural networks, J. Real Estate Res. 22 (3) 313–336.

Ngai, E.W.T., Xiu, L. & Chau, D.C.K. (2009). Application of data mining techniques in customer relationship management: a literature review and classification, Expert Systems with Applications, 36, 2592–2602

Park B., & Bae J.K. (2015) Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data, Expert Systems with Applications 42 (2015) 2928–2934.

Rachburee N., Arunrerk J., Punlumjeak W. (2018) Failure Part Mining Using an Association Rules Mining by FP-Growth and Apriori Algorithms: Case of ATM Maintenance in Thailand. In: Kim K., Kim H., Baek N. (eds) IT Convergence and Security 2017. Lecture Notes in Electrical Engineering, vol 449. Springer, Singapore. https://doi.org/10.1007/978-981-10-6451-7_3

Sarip, A.G., Hafez, M.B., & MD. Nasir Daud (2016) Application of Fuzzy Regression Model For Real Estate Price Prediction, Malaysian Journal of Computer Science 29(1) 15-27.

Sharma, Nitin & Arora, Yojna & Sharma, Vikas & Gupta, Hardik. (2020). Real Estate Price's Forecasting Through Predictive Modelling. In book: Machine Learning for Predictive Analysis, Proceedings of ICTIS 2020, pp.589-597, 10.1007/978-981-15-7106-0_58.

Singh, A.K., Kumar, A., Maurya, A.K. (2014) An empirical analysis and comparison of apriori and FP-growth algorithm for frequent pattern mining. In: 2014 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), pp. 1599– 1602.

Singh, A., Sharma, A. & Dubey, G. (2020). Big data analytics predicting real estate prices. International Journal of Systems Assurance Engineering and Management, 11, 208–219. https://doi.org/10.1007/s13198-020-00946-3

Tsai P. S.M., & Chen C-M. (2004). Mining interesting association rules from customer databases and transaction databases, Information Systems 29, 685–696

Yu Y., Lu, J., Shen, D., & Chen B. (2020). Research on real estate pricing methods based on data mining and machine learning, Neural Computing and Applications.
https://doi.org/10.1007/s00521-020-05469-3

Zadeh, L.A., Fuzzy sets, Information and Control. 8 (3) (1965). 338–353.

Zhang ML, &Yang Wp. (2012). Fuzzy Comprehensive Evaluation Method Applied in the Real Estate Investment Risks Research, Physics Procedia, 24, 1815 – 1821.

Zheng X., & HuiE.C.M. (2016). Does liquidity affect housing market performance? An empirical study with spatial panel approach, Land Use Policy, 56, 189–196

Zhu H., Xiong H., Tang F., Liu Q, Ge Y., Chen E., & Fu Y. (2016). Days on Market: Measuring Liquidity in Real Estate Markets, In: Proceeding KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 393-402.