# MAKİNE ÖĞRENMESİ YÖNTEMLERİNİN TOPLU GAYRİMENKUL DEĞERLEME ÇALIŞMALARI AÇISINDAN DEĞERLENDİRİLMESİ

# ASSESSMENT OF MACHINE LEARNING METHODS FOR MASS REAL ESTATE APPRAISAL

**SEÇKİN YILMAZER**

**ASSOC. PROF. DR. SULTAN KOCAMAN GÖKÇEOĞLU**

**Supervisor**

Submitted to

Graduate School of Science and Engineering of Hacettepe University

as a Partial Fulfillment to the Requirements

for the Award of the Degree PhD

in Geomatics Engineering

2022

*Yavrum Muhammed Ömer Tuğraya ithafen…*

# ÖZET

**MAKİNE ÖĞRENMESİ YÖNTEMLERİNİN TOPLU GAYRİMENKUL DEĞERLEME ÇALIŞMALARI AÇISINDAN DEĞERLENDİRİLMESİ**

**Seçkin YILMAZER**

**Doktora, Geomatik Mühendisliği Bölümü**

**Tez Danışmanı: Doç. Dr. Sultan KOCAMAN GÖKÇEOĞLU**

**Nisan 2022, 102 sayfa**

Bu çalışmada, toplu gayrimenkul değerleme için makine öğrenmesi yöntemlerinin kullanımı, geniş bir sahaya uygulanan 5 yöntemin etkinliği, tahmin doğrulukları, ve şeffaflığı dikkate alınarak incelenmiştir. Çalışma alanı Türkiye'nin başkenti Ankara Kenti Mamak İlçesi içinde yer almaktadır. Bu tezde kullanılan veriler özenle denetlenmiş ve incelenmiştir ve yüksek kaliteye ve güvenilirliğe sahiptir. Öte yandan, toplu değerlendirme çalışmalarında, makine öğrenmesi yöntemlerinin uygulanabilirliği, elde edilen sonuçların klasik yöntemlere göre hassasiyetleri, güvenilirliği, yorumlanabilirliği ve açıklanabilirliği tartışılmıştır. Elde edilen sonuçlar incelendiğinde, makina öğrenmesi tabanlı gayrimenkul değerleme yöntemlerinin, birçok gayrimenkulü aynı anda ve çok daha hızlı bir şekilde değerleyebileceği, bu sayede de geleneksel değerleme yöntemlerine göre tercih edilebileceği anlaşılmıştır. Değerleme çalışmaları kapsamında karşılaştırılan yöntemlerden, Rastgele Orman en yüksek tahmin hassasiyetini elde etmiş olup, diğer yöntemler sırasıyla yapay sinir ağları, destek vektör makinaları, çoklu regresyon analizi ve uyarlanabilir sinirsel (nöron) bulanık çıkarım sistemi olarak yüksek doğruluk sağlamıştır. Karşılaştırmalı çalışmanın bir başka sonucu ise, doğrusal olmayan makine öğrenmesi yöntemlerinin yorumlanabilirlik ve şeffaflık açısından kullanılabilirliğinin

tartışılması gerektiğidir. Bu şekilde değerleme çalışmasının amacına göre kullanılabilecek yöntemler tez kapsamında detaylı olarak incelenmiş ve yeni çalışmalara katkı sağlamak amacı ile paylaşılmıştır. Ayrıca toplu değerleme çalışmaları alanında sunulan yöntemlerin gayrimenkul değerleme alanında henüz kurumsallaşmamış olan Türkiye'ye olası katkıları ile örnek bir gayrimenkul değerleme sistemi önerisi geliştirilmiştir.

**Anahtar Kelimeler:** Konut Fiyatı, Gayrimenkul Değerleme, Toplu Değerleme, Makine Öğrenmesi

## ABSTRACT

# ASSESSMENT OF MACHINE LEARNING METHODS FOR MASS REAL ESTATE APPRAISAL

**Seçkin YILMAZER**

**Doctor of Philosophy, Department of Geomatic Engineering**

**Supervisor: Assoc.Prof.Dr. Sultan KOCAMAN GÖKÇEOĞLU**

**April 2022, 102 pages**

In this thesis, the use of machine learning (ML) approaches for the purpose of real estate mass appraisal was investigated using five different methods in a large area considering the efficiency, accuracy, and transparency. The study area was located in the Mamak District of Ankara, the capital city of Turkey. The data used in the thesis were inspected and analysed thoroughly, and thus exhibit high quality and reliability. The applicability of the ML methods in the context of mass appraisal is discussed in terms of the accuracy, reliability, interpretability, and the generalization capability. The results were also compared with the conventional appraisal methods. The results obtained here have shown that the ML-based methods can appraise many real estates together at once and rapidly; and thus, they can be preferred over the conventional valuation methods. Among the methods compared here, the Random Forest (RF) provided the highest prediction performance followed by the Artificial Neural Networks (ANN), Support Vector Machines (SVM), Multiple Regression Analysis (MRA), and Adaptive Neuro-Fuzzy Inference System (ANFIS). The stepwise MRA method, which is a transparent and interpretable linear ML method, was preferred as the conventional approach. Another important outcome was that although the models built with the non-linear ML methods yielded high accuracies, their interpretability was lower and thus usability for the valuation purposes may be questionable. In this thesis, the employed methods are explained and investigated in more detail with the aim of contributing to the mass appraisal context. In addition, recommendations on the real estate valuation system were derived based on the study outcomes together with possible contributions of the methods presented in the field of mass valuation studies to Turkey, which has not yet been institutionalized in the field of real estate valuation.

**Keywords:** Housing Price, Real Estate Valuation, Mass Appraisal, Machine Learning.

# TEŞEKKÜR

Bu çalışmanın fikir olarak geliştirilmesinden, sonuçlandırılmasına kadar devam eden yoğun süreçte, her zaman yanımda olan ve desteğini esirgemeyen değerli danışman hocam Doç. Dr. Sultan KOCAMAN GÖKÇEOĞLU'na, çalışmamın daha başarılı olması için vakitlerini ayıran ve tavsiyeleri ile ufkumu genişleten Prof. Dr. Candan GÖKÇEOĞLU ve Prof. Dr. Fatih İŞCAN'a, tez çalışmamın teknik detaylarında fikir ve görüşleri ile destek olan Doç. Dr. Aslı ÖZDARICI OK ve Dr. Öğr. Üyesi Murat DURMAZ'a, çalışmamın en meşakkatli bölümünde ne zaman bir yardıma ihtiyacım olsa yardımlarını esirgemeyen iş arkadaşlarım, Ümit YILDIZ, Tolgahan ÖZDEN ve Hüseyin ÇALIŞ'a, yoğun çalışma saatlerimden ötürü daha az yanlarında olabilmeme rağmen beni sürekli sabır ve tevazu ile karşılayan hayat arkadaşım Ekin Özgür ve değerli yavrum Muhammed Ömer Tuğra'ya, sorgusuz, sualsiz benim her anımda destekçim olan baştacım Anneme teşekkürü bir borç bilirim.

# CONTENTS

# INDEX OF FIGURES

# INDEX OF TABLES

# ABBREVIATIONS

| | |
|---|---|
| ANFIS | Adaptive-network-based Fuzzy Inference System |
| ANN | Artificial Neural Network |
| BP | Backpropagation |
| COD | Coefficient of Determination |
| COV | Coefficient of Variation |
| MARE | Mass Appraisal of Real Estates |
| ML | Machine Learning |
| MRA | Multiple Regression Analysis |
| r | Correlation Coefficient |
| REA | Real Estate Appraisal |
| RF | Random Forest |
| RMSE | Root Mean Square Error |
| SARE | Single Appraisal of Real Estates |
| SVM | Support Vector Machine |

# 1. INTRODUCTION

Real estate valuation is one of the four pillars of land administration systems. In countries with incomplete real estate valuation systems or areas with dynamic real estate markets, mass appraisal has become a viable option for obtaining rapid and reliable results. Although the land valuation has traditionally been performed by experts, and the expert-based systems can be an applied to the valuation of areas with small to medium size, recently the data-driven machine learning (ML) approaches have also been a promising solution for this purpose.

This thesis aimed at analyzing the usability and quality assessment aspects of ML approaches in the context of mass appraisal of real estates (MARE). The remainder of this Section briefly introduces the basic terminology in real estate valuation, explains various scientific aspects of the problem, investigates the current methodological and financial trends in the area, and provides the main goals of this thesis. In addition, the thesis structure is given by the end of this Section.

## 1.1. Real Estate Valuation Terminology

In real estate economic markets, the terminology is highly complex and non-standardized. Many terms related to the value of a real estate are currently in use and sometimes utilized interchangeably, such as market value, fair market value, selling price, commercial value, value in securities, rental value, expropriation value, etc. In this Section, the terms utilized within the content of this study are explained briefly.

*Real Estate* consists of land and the buildings on it, or attached to it, and its natural resources. The term has several definitions depending on the land administration systems in different countries. In Turkey, it is a piece of land with geometrically defined borders, has an economic value, and legally grants property rights to its owner such as to use, to sell or to rent it.

*Market Value* is the sale price of a property that is determined based on its characteristics at a given time in transparent and open market conditions and also upon the will and knowledge of the buyer and the seller.

*Price* is the result of the agreement between a buyer and a seller about a good at a given time without any obligation.

*Predicted Value* is a value that is obtained via an algorithm based on various mathematical approaches such as statistics, ML, etc., and often computed by using a software tool.

*Raw Data* is obtained from different sources such as real estate companies, websites, etc., and is not always from a trusted source. Thus, the raw data is subject to expert analysis or statistical inspections. The raw data is composed of the price information obtained from the real estate transaction and independent variables affecting it.

*Computer-assisted Mass Appraisal (CAMA)* is an automated appraisal system that is commonly used for real estate taxation or a reference / index value for appraisers by automatically evaluating mass property data with the help of algorithms supported by mathematical methods, e.g., statistics, or artificial intelligence (AI) techniques.

## 1.2. Problem Definition

Countries adopt various land management models according to their characteristics, such as the economic level, development status and the existing legal infrastructures. Land administration systems involve administrative and operational components of land management in which land and its resources are planned and managed in a systematic context to support sustainable development; and following the framework land management model applied that adopts optimum management tools for tenure, use, value, and development of the land. Real estate valuation is one of the main tools of land management systems. It has an essential role in several important tasks in financial systems such as taxing, funds and investments, company listings, litigation, expropriation, banking operations. Real estate valuation, which impacts the financial sector and all sectors indirectly at the macro level, affects almost every citizen at the micro-scale. The most common examples can be listed as renting, selling or buying a house, operating farmland, small offices and shop businesses. From this point of view, the fact that real estate valuation studies in Turkey are not institutional can be expressed as a major shortcoming. On the other hand, because real estate valuation studies are not regulated under a framework legislation, a leading institution to audit the activities has not yet been established. Thus, the lack of precise knowledge on the real estate values throughout the country is known as a major problem.

According to the number of real estates appraises, works can be divided into two categories as Single/individual Appraisal of Real Estates (SARE) and MARE [1, 2]. In addition to the institutional and sectoral problems of the Real Estate Appriasal (REA) regardless of SARE and MARE, accessing reliable data is a problem in itself. The problems in MARE can be related to the management and legal infrastructure, accessibility to reliable data, methods, and the human factor. Within the scope of this thesis, it was assumed that the issues based on legal infrastructure and the management are in suitable conditions; and the other problems, such as the scarcity of data and it reliability, and the lack of transparency of the real estate market were solved by using the data obtained from a pilot MARE project carried out in Turkey.

The ML methods has come to the fore with the increase of available data. The ML-based models trained with a small amount of data provide a great opportunity for the MARE studies by predicting the values quickly and reliably. In different studies, the ML methods were used for the comparison of the prediction accuracy or for developing a local property taxation approach [3-5]. Thus, it contributes to forming a more fair and transparent real estate market, and an environment of trust can be created between the citizen and the state. On the other hand, the valuation activities in developing countries have not yet been institutionalized. They are strongly influenced by real estate market actors, thus prevent from the formation of a transparent and reliable real estate market. The real estate markets can be grouped into three different classes such as; buyer-oriented market, seller-oriented market and fair market. In the context of this study, a fair market data and reliable institutional data were gathered to be used in all stages of the model establishment phases. However, in order to reach a fair real estate market several prerequisites must be met that are objective valuation criteria, transparent and reliable public and private institutions, appropriate valuation methods, and experienced and well-trained experts.

Within the scope of this study, the data collected in a pilot valuation study carried out within the General Directorate of Land Registry and Cadastre (GDLRC) of Turkey [6] were used as raw data; and the extreme values and erroneous data were eliminated by using both statistical methods and expert inspection. After creating a reliable data set, it was necessary to consider the indicators that are effective in the real estate market, such as national, regional, and local politics, social characteristics, awareness of regional education facilities, natural hazard susceptibility levels, topographic characteristics, building diversity, etc.  Although various parameters affect the real estate values, they are rarely used in valuation studies as most of them are not measurable and also the interactions between the variables are often non-linear and thus difficult to model. However, the ML-based MARE methods can help in solving these problems, in part, by solving non-linear interactions between variables. For example, it can provide solution-oriented comparison opportunities by determining the importance levels of the variables. Therefore, additional variables that are not frequently used in REA studies can be used in a ML-based MARE study. However, despite the benefits of ML-based mass valuation methods; discussions continue on their limitations in terms of accuracy, generalizability, reliability, transparency and interpretability.

### 1.3. Current Trends in Real Estate Valuation

Land valuation or REA is one of the most important functions of a modern land administration system. Land value is an important information that supports sustainable development with the other functions of land administration that are land use, land planning, and land tenure. Enemark [7] has identified the REA works among very important land management functions, especially for taxation [2]. The REA studies have a greater importance nowadays in a wide range of application areas. In this context, the REA institutionalized in a successful land management system ensures trust and stability in the financial sectors [7]. Furthermore, it helps to regulate the credit system (mortgage), expropriation, land consolidation, sale of public goods, privatization type transactions; enables a reliable REA infrastructure; secures the rights and benefits of investors; and decreases the risks of banking and insurance services. The importance of REA works can be understood in a broader perspective, as the real estates of the countries are financial assets and an indicator of their economic size. In other words, the measurement of the size of the real estate markets in countries and their economic scales have become a great necessity in the rapidly growing and globalizing capitals. While determining the credibility or investability scores of countries, global credit rating agencies observe the total assets, debts of the countries, financial risks, and calculate the scores by taking into account the economic strength of local markets. Global investors consider the reports of the credit rating agencies, and issues such as the reliability and stability of financial markets in their investment preferences.

On the other hand, the REA, one of the important topics of the economical measurement, has importance in preventing informal economy and tax losses. Since it is time consuming and costly to regulate many properties with SARE in real estate-based taxation works, computer-assisted mass appraisal (CAMA) systems are used in many countries [8]. The SARE can also be considered as a conventional appraisal approach similar to the cost, sales comparison, income and hybrid methods [9]. The MARE enables the valuation of multiple immovable assets with the help of statistics, mathematics or other ML-based algorithms. With SARE methods, appraising a large number of real estates can be difficult. In this context, REA activities have been automated, especially in developed countries, or have been carried out by using mass valuation methods.

The MARE is a useful and feasible approach regarding its efficiency in time and cost reduction [10]. Due to the inability of SARE methods to respond to the developing and changing needs, advanced and different types of applications have increasingly been used in MARE studies depending on the scientific and technological developments. Wang and Li reviewed more than one hundred publications on MARE studies [2] and identified 3Itrend terms: artificial intelligence (AI)-based model, GIS-based model, and MIX-based model. The ML methods, a sub-branch of AI, can also be classified in different ways as explained in Chapter 3.2 in this thesis. Although some examples were seen in the 1970s [11, 12], the ML methods have started to be utilized more frequently in MARE studies during the last 30 years with the technological developments [13-15]. In this study, five commonly used ML methods, namely; Random Forests (RF), Support Vector Machine (SVM) Artificial Neural Networks (ANN), and Artificial Neuro-Fuzzy Inference System (ANFIS), Multiple Regression Analysis (MRA) were employed for the mass valuation of 2936 data samples composed of mostly residential properties and with a few commercial ones.

### 1.4. Thesis Goals

In this thesis, the methodological approaches for using the ML methods in MARE studies were compared using a large dataset in a complex urban environment. The study was carried out to serve the following purposes;

- To reveal the usability of ML methods and contribute to the literature in the context of accuracy, reliability, interpretability and generalizability of the selected methods in MARE works.
- To evaluate the institutional structure of the REA works in Turkey and to provide recommendations for the preparation of value maps using ML-based MARE methods.
- To obtain experimental and comparative results on the use of ML methods in a complex urban setting.

It was expected that appraising the real estates using the same dataset with various experimental techniques may lead to new conclusions on the evaluated methods and the mass appraisal problem.

### 1.5. Thesis Structure

This thesis is presented in a total of six chapters as following;

**Chapter 1.** Introduction: the terminology of real estate valuation, the definition of the problem and the current trends in REA studies and the main purpose of the thesis are given.

**Chapter 2.** Background: the literature on the REA and ML-based MARE studies has been analyzed and discussed.

**Chapter 3.** Methodology: the definitions of the pre-processing and ML methods used in the thesis are presented.

**Chapter 4.** Application: the practical uses, and the implementation results of the ML methods, which are generally defined in the third chapter, are given.

**Chapter 5.** Discussions: the outcomes of thesis are discussed for various aspects.

**Chapter 6.** Conclusions and future work: the major outcomes of the thesis and recommendations for future works are presented.

# 2. BACKGROUND

This Chapter is presented in two parts. In the first part, the land valuation approaches are discussed and the historical development of valuation studies in Land Registry and Cadastre activities is explained with examples. In the second part, the use of ML methods in MARE studies is discussed with examples from the literature.

## 2.1. Land Valuation Approaches

The land valuation efforts have initially started with the values recorded during the land registry and cadastre activities based on the idea that arable land in Europe is a source of wealth and should be taxed [7]. The land registry systems can be categorized as as the deeds system and the title system. In a deeds system, only the transaction information of landowners were registered. However, in a title system, the information on the lands and their owners is registered for security purposes. The deeds system was seen first in the Roman Law, and the title system was seen in the Germanic or Anglo Common Law [7]. These registration implementations were spread worldwide and adapted their characters according to the social and cultural types of the countries and the local legal infrastructures, especially in land tax purposes. In the beginning of the 20$^{th}$ century, the cadastral systems and the registries were combined in most of the developed countries, and thus the first steps of the modern multi-purpose cadastre were established.

Today's modern cadastral systems in developed countries have integrated information about land use, land tenure, land planning, and land value as their functions, and their processes are affected by each other. The land valuation must be performed by following scientific principles and the other land administration functions effectively. Since these qualifications can vary according to the subject, purpose, and place of valuation; many countries have differemt legal regulations. Especially in the developed countries, valuation studies have been placed on a legal basis, and the institutions related to the valuation have completed the institutionalization process. In addition to local legal regulations, international organizations provide services for REA and related fields to ensure the reliability and standardization of valuation studies and the transparency of real estate markets. The organizations such as the Royal Institution of Chartered Surveyors (RICS) [9], Uniform Standards of Professional Appraisal Practice (USPAP), International Association of Assessing Officers (IAOO) [16] and International Valuation Standards Council (IVSC) [17] have published standards at certain time intervals to support appraisal works. The International Valuation Standards Commission (IVSC) [17] is an organization that provides services to support international asset valuations in the context of consistency, transparency and trust and to develop standards. It is a public interest institution with guides and standards published in this field. Using such guides in valuation studies contribute to forming a worldwide valuation norm.

When looking at the terminology of land valuation, it is a broad concept. It is defined as the valuation of unstructured lands, zoning parcels and buildings or assets within the scope of those areas. However, the terms housing valuation, zoned land valuation or annex valuation are not used separately in the literature. Still, the same meanings such as land valuation, real estate valuation, and property valuation are used. Nevertheless, when valuing the assets, valuation methods would vary. For instance, the sales comparison approach is generally used when a built house valuation is considered.

In contrast, the income method is preferred when an agricultural land or commercial estate valuation is considered. However, valuation studies differ according to the purpose of the valuation and the number of real estate properties to be valued. These changes can be arranged in local legal regulations. For example, in Turkey, the expropriation legislation, the enforcement of bankruptcy law and the commercial law contain different limitations on valuation works. The SARE methods are carried out with a report prepared by a licensed expert following valuation standards, using one of the traditional valuation methods, such as cost, income, and sales comparison, or hybrid methods for one or a few real estates. If the number of real estates to be valued increases, it becomes both a time-consuming and costly process. The mass valuation methods are used to reduce the cost with the advancement of technology and evaluate a large number of real estate together [18-21]. Statistical techniques supported tax authorities in the early 1950s for mass appraisal [22]. Due to the recent developments in econometrics and computing technologies, the appraisal studies have turned into computerized applications.

Both Shenkel [11] and Rollo [12] described the MARE in the early 1970s using regression analysis. However, the prediction accuracy of MARE works was discussed and questioned for a long time. In the late 1990s, Fraser et al. [21] stated that the manual valuation results outperformed the computerized REA results. Thus, it should not be assumed that any Automatic Valuation Model or MARE application gives better results than the expert analysis. In many applications, MARE results are considered a reference value made available to appraisers [23], or a system that helps to establish a data infrastructure for property tax values [1]. The number of studies have risen in parallel with the increasing importance and the need for REA studies, and have turned into different perspectives with the advancement of technology and science. Statistics, Geographic Information Systems (GIS), ML methods and a mix of them have been applied to the valuation field [2]. Some of these studies were carried out for the assessment of prediction accuracy [24-26], for convenience in taxation studies [11, 27], for building a land re-adjustment works [28], or for evaluating time and cost factors [23, 29, 30]. Moreover, all studies contributed greatly to the progress of valuation works. In addition, the application of ML topics in MARE context are relatively new and continue to develop while keeping their place on the research agenda.

### 2.2. The Use of Machine Learning in Mass Appraisal Studies

The REA studies should be carried out in an unbiased, objective and reliable manner. Unrealistic valuation results provide financially misleading information to stakeholders and reduce trust in governments and authorities [1]. The conventional SARE methods are subjective, biased or un-controlled. In the USA, the traditional CAMA program was used in MARE works, giving authorities a chance to deflect accusations of subjectivity [31].

The ML is a branch of artificial intelligence that combines computer science and statistics, modelling how humans learn and apply what they learn to improve the predictive accuracy of their mathematical models by processing complex and large-scale data [32]. The ML algorithms serve many different branches of science. The ML techniques are frequently used in product development, research and development procedures and process improvement, and more specifically in image and speech recognition [33], medical diagnosis [34], classification accuracy [35], regression [36], prediction accuracy [37] and city planning [38].

In REA studies, the ML algorithms have been increasingly used for 30 years in parallel with the mass valuation needs, especially time and cost saving [15, 39]. With the increasing amount of data and novel ML methods, it is still an active research and application area [40-42]. In the first years of their use, the prediction accuracy was prioritized and compared with traditional or statistical methods over time. However, during the last 10 years, these comparisons were made regarding various perspectives and different fields, such as prediction accuracy [43], interpretability [44], transparency [45] and generalizability [46] of the models. These concepts are also very important in the field of mass valuation. The results of the valuation method used may sometimes affect tax values that concern large masses of the population or court decisions. For this reason, the contribution of the results obtained in comparing the methods to the literature will also make great contributions to real-life applications.

There are many benefits of using ML methods in MARE studies. The most significant features are time saving and hence low-cost predictions. In addition, another important contribution of the ML methods can be the generalization capability after obtaining established MARE models. In the literature, the generalizability of a model has been tested by establishing some automatic forecasting systems especially in the field of health [47, 48], but no study has been carried out on the MARE concept yet. The reusability of a model on a similar or related subject will also provide a great advantage in terms of decreasing processing time and reducing a model building cost and thus the data requirement. In addition to these aspects, the ML studies maintain their place on the agenda with its parts on which intensive academic studies are still carried out, such as interpretability, transparency and reliability of the ML models, which are important features in REA and are affected by many factors. For example, during the model establishment phase, the quality of the collected data and the method selection directly affect these aspects. Unless coming from a known source, the reliability and transparency of the data cannot be assured. On the other hand, the results and procedure cannot be explained if the method and its configuration performed during the estimation phase are unknown. For this reason, it is necessary to pay attention to these issues in building ML-based MARE studies.

From this point of view, there are also some prerequisites and obstacles for building a MARE. The prerequisites are the availability of reliable and high quality input data, the determination of an inclusive and appropriate variable list, responsible and authorized institutions, and legal regulations about appraisal works, etc. The most important features of the ML methods are that they can obtain accurate results with a small amount of data samples. The models can be applied to unseen and related data on new study areas by generalizability or associated sites by transferability. Some of the known issues can be listed as the curse of dimensionality, overtraining and overfitting. When eliminating these issues, aspects such as the correlations of the variables should be tested. The multicollinearity test shows whether the variables contain multicollinearity, and it provides effective results in solving these problems. Moreover, the Cross-Validation (CV) technique is another effective technique in removing overtraining and overfitting curses. However, these issues can still cause the model to produce biased results.

Finally, in solving these issues and making the right choice of the methodology, the dataset should be analyzed carefully and the features of the methods should be taken into account. However, some methods such as stepwise regression and sub-clustering based ANFIS contain natural data reduction techniques. Using data reduction techniques in methods or seperately, the prediction accuracy and the total variance would be decreased. While building a MARE model, the use of expert based data reduction techniques is another option to eliminate the curses. If real estate appraisers have sufficient knowledge and experience in the field, they can perform valuation works and data pre-processing based on expert knowledge. Such expert-based approaches can commonly be used with conventional valuation approaches according to the collected data and the valuation purpose. However, subjectivity and biases may occur in the model results. Moreover, the expert intervention may also face problems, especially high dimensionality, long processing times and of course higher cost. Therefore, in the real-world applications, more automated methods are the first choice in a complex MARE work.

# 3. METHODOLOGY

This Chapter is divided into two parts as a Pre-Processing Methods and the Supervised Machine Learning Methods. In the first part, pre-processing techniques from the literature are discussed and a workflow of the study is given. In the second part, an example of the classification of ML studies is presented, supported by the literature, and general information about the five ML methods used within the scope of the thesis are explained.

## 3.1. Pre-Processing Methods

In this thesis, the ML and MARE techniques are discussed in parallel. The accuracy, precision and reliability are the major principles in both methods. Therefore, it is important and necessary to review all stages of applying ML methods in the field of REA and apply validation and revision processes. Although three main processing phases can be considered in this thesis, the first and most important step is to obtain reliable and accurate data. The second and the third stages are the method selection and the modelling, respectively. If data quality and reliability can not be ensured in the first stage, accurate results cannot be obtained in the subsequent steps. Therefore, both expert inspection and statistical tests were performed to eliminate the issues, noise, and losses assumed to be present in the collected data. In general, the following pre-processing methods can be applied for data cleaning and error elimination;

- Data distribution must be analyzed and the collected data may not comply with the normal distribution due to systematic errors and outlier samples. For linear regression models, a normal distribution is expected. However, in non-linear modelling methods or non-parametric models, a normal distribution of data is not necessarily required.

- Data cleaning techniques such as determination of missing values, identification of duplicate and unusual samples, and outlier detection in the data can be applied.

After a statistical analysis and cleaning, data can be considered more reliable. Some of the data reduction techniques were utilized within the scope of this thesis.

- The amount of data must be analyzed for identifying the suitability of the potential modelling methods. Here, it is aimed to obtain accurate and reliable results with least amount of data and variables as they are bound to cost and labour requirements.

- Variable scaling based on a dependent variable by analyzing statistics of variance ($R^2$ change), and field research as suggested by Guan et al. [49].

- The $R^2$ change can be tested in the model with the F-test. The difference in the F value means that the added variables are useful and improve the forecast. The F-test is also called variable significance analysis as it gives a similar result to the analysis of variance.

- Factor analysis [50], i.e. optimal scaling [51], principal axis factorization [52] and principal component analysis (PCA) can be utilized for dimension reduction as suggested by [53].

- Data clustering for dimension reduction and process optimization [54] such as using the Hierarchical Fuzzy Inference Systems (HFIS) structure [55] or Sub-clustering technique in ANFIS [56].

The data used in this thesis is a combination of the data collected within the scope of the research and the data obtained from the mass valuation pilot project by the General Directorate of Land Registry and Cadastre [57]. The overall methodological workflow of the study is shown in Figure 3.1.

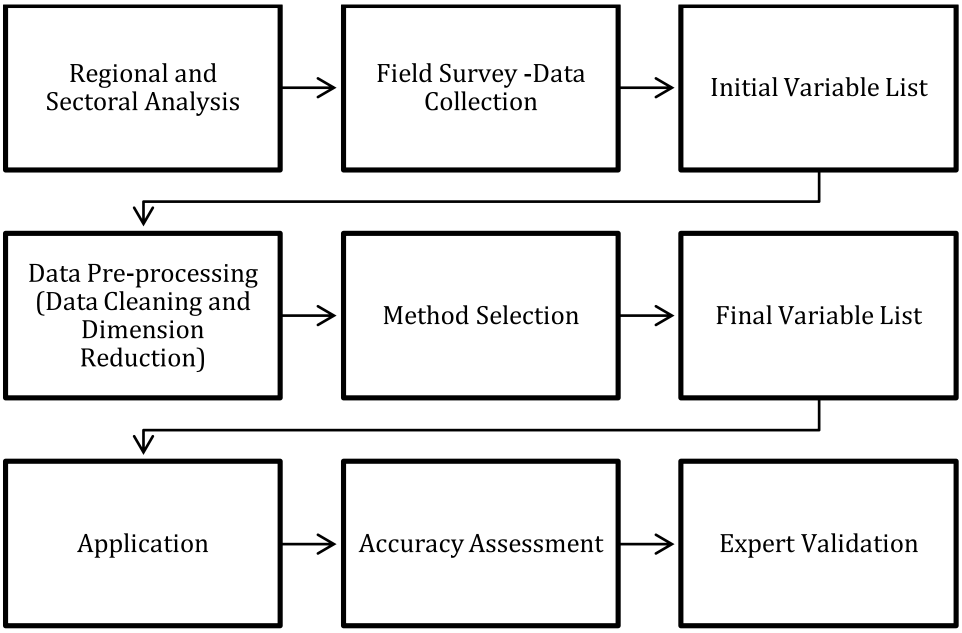



Figure 3.1. The overall methodological workflow of the study.

### 3.2. Supervised Machine Learning Methods

The ML methods are considered as a sub-branch of AI that carry out predictions or support decision-making. The ML methods have been categorized under three [58], four [59], six [60], and eight [61] classes in the literature. The differences come from the interpretation of detailed methodological assumptions; and all classification approaches would be valid for a wide range of application areas. In this study, four categories of commonly used ML applications [59] are considered as shown in Figure 3.2.

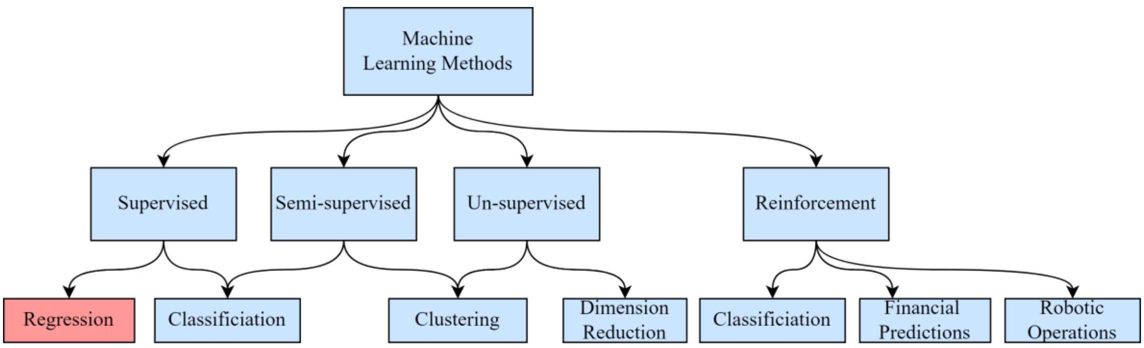



Figure 3.2. Classification of ML algorithms (adapted from [62]).

The supervised methods belong to a sub-category of ML approaches. A supervised ML method induces information from the labelled input data, improves them with advanced statistical techniques, and forms a dataset for prediction or classification. In contrast, the unsupervised ML methods extracts patterns from unlabelled data to make a decision by using techniques such as clustering, dimension reduction etc. In this thesis, the PCA, which is an unsupervised data reduction method, was used to reduce the dimensionality. On the other hand, the semi-supervised methods combine expert information, labelled and unlabelled data.

The reinforcement methods belong to another important area of ML that supports agents taking feasible reactions to maximize reward in commercial circumstances in general. Reinforcement learning does not make decisions by processing information from a given training set. Instead, it makes decisions about how to act in new problem based on the situations it has experienced.

As mentioned above, many other ML types can be included into the abovementioned classes. In this thesis, a major focus was on the supervised ML methods, which can be utilized for both the regression and the classification tasks. Real estate value prediction processes generally require regression-based supervised ML methods. In this thesis, four commonly used regression methods such as ANN, SVM, RF and ANFIS were investigated using Matlab Statistics [63] environment. In addition, the MRA was performed for comparative evaluations. The methods are explained in detail in the following sub-sections.

### 3.2.1. Random Forest

The RF is a supervised ML algorithm that uses an ensemble of decision trees (DT) to classify a dataset or to make predictions [1]. It can handle large datasets and produce accurate classification tasks according to the majority voting of partitioned trees [64]. On the other hand, regression type of ML methods is the preferred method on REA tasks [24, 65, 66]. Two types of RF regression models, bagging [67] and boosted trees [68] have been frequently used in recent literature. The boosted trees create larger tree groups that minimize the errors by adding new trees to DT based on existing erroneous results. The strengthened tree groups built with this approach provide a structure that produces more accurate results. In the bagging method, by partitioning the variables present in the training data set into different predictor tree groups according to their importance levels [69], the trained model is affected as little as possible by the noisy and outlier samples in the data set [70]. As a result, the predictions appraised by taking the average of the predictors obtain more accurate than each tree group. The RF bagging model built for the study is illustrated in Figure 3.3.

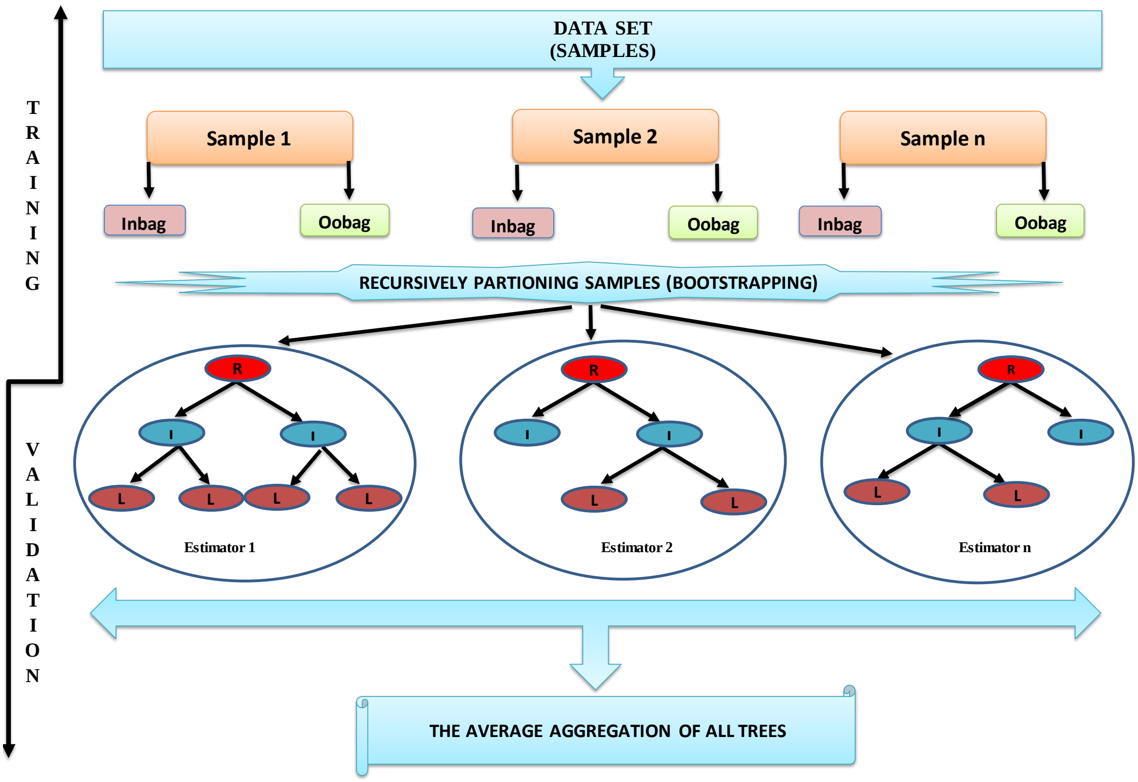



Figure 3.3. The RF bagging method workflow (modified after [1]).

In the bagging-based RF models, the prediction error is measured by out-of-bag (OOB) error [71]. Thanks to the subsampling with the replacement of training data in the bootstrap aggregating model, the input data is divided into two parts such as OOB and in-bag classes. The true prediction error of the model is considered to be in one-to-one correspondence with the value calculated by averaging each prediction error of the estimators as called OOB error. It is a major tool for estimating the prediction error of the bagging method effectively. On the other hand, bootstrap aggregating allows for an easy measure of the bagged tree model error. In this case, some of the samples, namely, OOB samples, could not use in bootstrapped trees. In other words, the OOB samples are not included in the bootstrap samples. In general, 36.8% of the data allocated as the total training dataset constitutes the OOB sample. To explain this fact;

Assuming there are N rows in a training dataset; the probability of not picking a row in a random draw is;

(a) $\frac{N-1}{N}$ ; the probability of not getting a row in a random draw

In a sub-sampling with replacement phase,

(b) $(\frac{N-1}{\mathrm{N}}^{N}$ ; the probability of not selecting N rows

(c) $\lim_{n\to\infty}\left(1-\frac{1}{N}\right)^{N} = e^{-1} = 0.368$

As a result, about 36.8% of the total training dataset can be called as OOB sample and can be used to evaluate the bagging-based RF model.

In contrast, Mitchell [72] mentioned that the OOB error is not a stable indicator. It is affected by the sample and variable sizes and the methodological assumption of bagging defined as subsampling with replacement. So, a high bias occurs in OOB error compared to the model's true prediction error. In addition, they have demonstrated that the bias could be greatly reduced by selecting the same proportion of observations from each group for in-bag samples and subsampling without replacement [72]. Based on the experimental findings and interpretations, a number of data reduction techniques were adopted in the study dataset to obtain an effective training set. A subsampling replacement process was used to select optimal parameters to build the bagging model using a K-fold-cross-validation for efficient training. The experimental results are presented and discussed in the Sections 4 and 5.

### 3.2.2. Artificial Neural Networks

The ANN mimics the decision-making mechanism of the human brain [73]. It is commonly used to figure out complex non-linear problems and produces accurate results in many scientific fields such as prediction [74], classification [75], clustering [76], pattern recognition [77], and simulation [78], etc.

There are many ANN architecture designs in the literature, and Leijnen et al. provided a detailed review of them [79]. The well-known Back Propagation (BP) based multi-layer feedforward (MLP) method is preferred in this study, and a typical architecture is given in Figure 3.4.

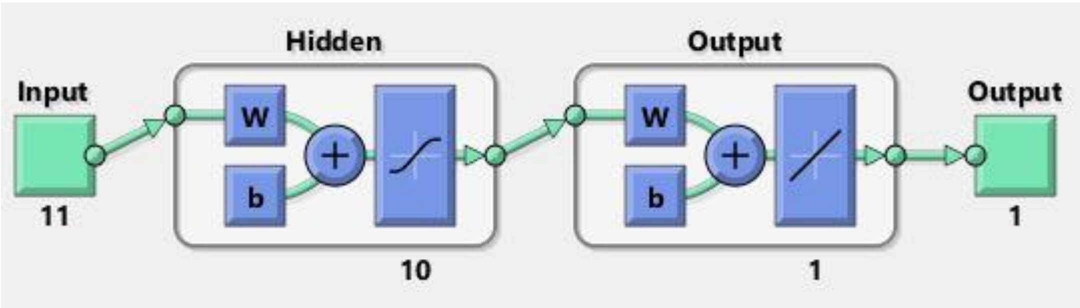



Figure 3.4. Typical Architecture of an ANN [63].

The BP repeatedly adjusts the weights of neurons to minimize the difference between the predicted output vector of the network and the actual output vector [80]. The structure consists of 3 layers such as input data, hidden layer, and output.

As for the working principle of the BP based MLP method given below;

(d) $y_i = \varepsilon \left( \sum_{j=1}^{N} W_{ij} x_j + b_i^j \right)$,

Where ;

- $\varepsilon$ is called activation (or transfer) function,
- $N$ is the number of input neurons,
- $W_{ij}$ the weights, $X_j$ is the $j^{th}$ input neuron, $b_i$ is the bias.

The process can be summarized as;

- Input $X_j$ is multiplied by $W_{ij}$, which is a connection weight between the $i^{th}$ input signal and $j^{th}$ hidden layer signal;
- $b_i$ is summed with the result;
- Output neuron signal $Y_j$ is obtained due to the $j^{th}$ process between the data pairwise $Xi$ and $Yi$.

For the error calculation of the first epoch, the predicted output $Y_i$, and the actual value of the target $Y_j$ are employed as;

(e) $e_j = \frac{1}{2} \sum_{j-1}^{N} (\Upsilon_j - Y_j)$ .

Finally, the error obtained with this epoch will be received for the entire data set. According to the results, the error is minimized by updating the $W$, $\varepsilon$ and $b$ coefficients iteratively.

### 3.2.3. Support Vector Machines

The SVM is a ML algorithm that is frequently used for pattern recognition [81], time-series analysis and prediction [82, 83] and regression tasks [83]. The regression-based SVM can effectively solve both linear problems and also those related to multidimensional input space such as REA data. The ML methods can produce results by classifying linearly separable datasets into different classes, as shown in Figure 3.5.

However, a dataset that cannot be modeled linearly (e.g., see Figure 3.6) must be moved to a high-dimensional plane [84].

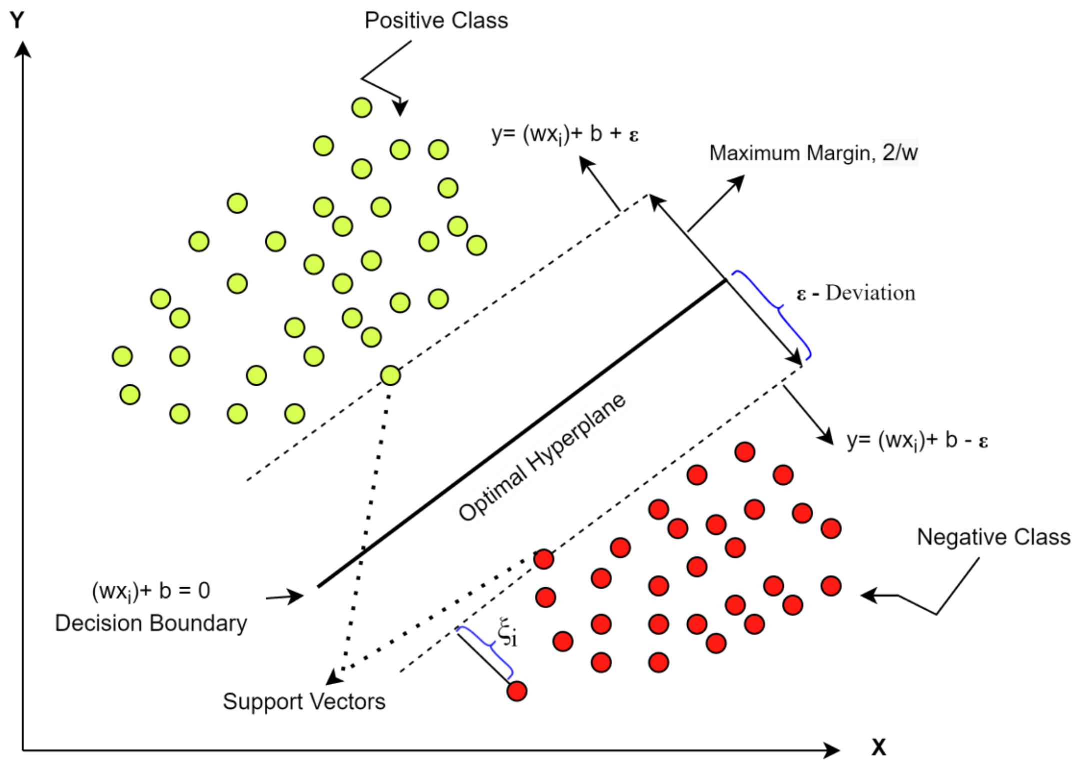



Figure 3.5. Classification of data points into two classes in a linear SVM (adapted from [85]).

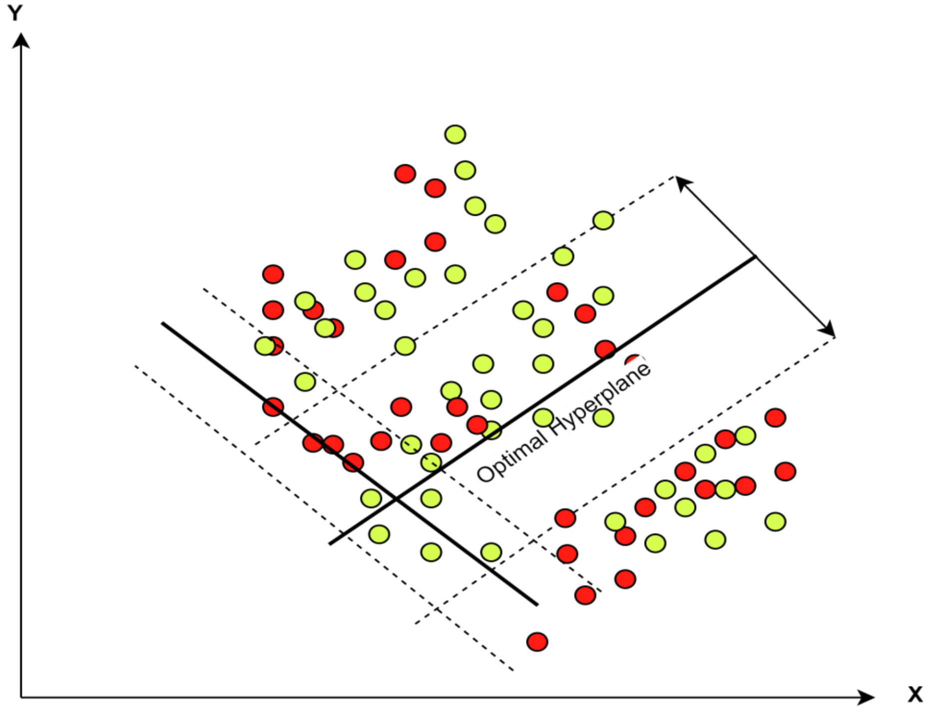



Figure 3.6. A dataset that cannot be classified linearly.

The ML methods carry out the task using a kind of converter called kernel, which is a function used in ML methods to separate the data. The SVM also commonly operates various kernel functions, which moves a low dimensional input space into a higher dimensional space via adding new dimensions based on the problem. The Gaussian Radial Basis Function (GRBF) based SVM works in infinite dimensions. Figure 3.7 illustrates as simple depiction of GRBF for separating high dimensional data.

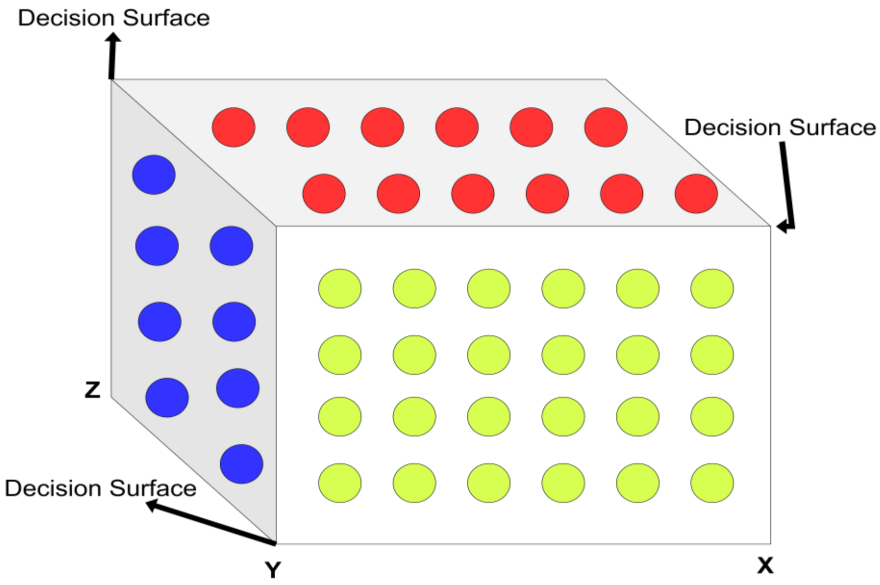



Figure 3.7. Multidimensional input space

Many variables and complex input data increase the importance of selecting kernel functions. According to Noble [86], the best method for kernel selection is trial-and-error. The linear, GRBF [87, 88], hyperbolic tangent [89] and sigmoid [90] kernels are commonly used for SVM in different research problems. However, the most widely used kernel function is the radial basis function (RBF) in the SVM method [91]. In this study, based on the trial-and-error test [86], the RBF that can work with infinite dimensions was also found efficient for the MARE application.

### 3.2.4. ANFIS

The ANFIS is a supervised ML method developed by Jang in 1993 [92] based on the Sugeno fuzzy inference system (FIS). The ANFIS a hybrid method that benefits from both the ANN and FIS (i.e., Fuzzy Logic (FL)) [92]. Here, the basis of the FL theory is sourced from the idea of Fuzzy Set Theory, which was developed by Zadeh in 1965 [93] to deal with the uncertainty in real-world problems. While many prediction and decision-making processes treat input variables as fixed and bivalent (i.e., true or false; or “0” or “1”), in fuzzy set theory, as in human cognitive processes, when evaluating an event, a value is determined in a range between 0 and 1 [94]. The FL can easily handle ambiguous and fuzzy situations by attributing the degree of membership to which an object belongs to a set. Thus, linguistic data can be easily measurable in decision-making processes.

In the FL, for solving a complex problem, the independent variables affecting the dependent variable are specified in membership functions in a rule-based structure. In fuzzy set theory, indefinite proportional representations are often used instead of classical bivalued sets, which do not determine precision and where true and false are proportionally expressed as defined below;

(f) $A(x) = (x, \mu A(x)), x = X)$,

where;

- $x$ is a member of fuzzy set $A$
- $\mu A(x)$ is the membership function
- $X$ is the universe of discourse.

In the FL approach, events are taken with an *if-then* setup. While planning this setup, the most commonly used rule structures are the expert-based Mamdani algorithm [95] and the data-driven Sugeno algorithm, which was developed by Takagi-Sugeno-Kang (TSK) [96]. If the expert has high level of expertise in an area, the Mamdani algorithm is the preferred method; while the Sugeno is preferred if the sample size is sufficiently large.

The Sugeno algorithm uses *if-then* rules that consist of two parts. The first part is precedent (fuzzy), and the second is the consequent, which is a function of the input dataset. The main difference between the Mamdani [95] and the TSK [96] approaches is that the TSK rules have a role in the consequent part, and the Mamdani rules have a linguistic output such as “small, medium, large” [49, 94].

The Sugeno type rule format is demonstrated below (g), and a generic architecture of it is illustrated in Figure 3.8.

(g) *IF x1 is A1 and ... and xk is Ak THEN,* $y_j = p_1 \cdot x_1 + \cdots + p_k \cdot x_k + p_0$ ,

where;

- $x_{1...k}$ are input variables;
- $A_{1...k}$ are the fuzzy set members;
- $p_{1...k}$ are the coefficients;
- $y_j$ is the target.

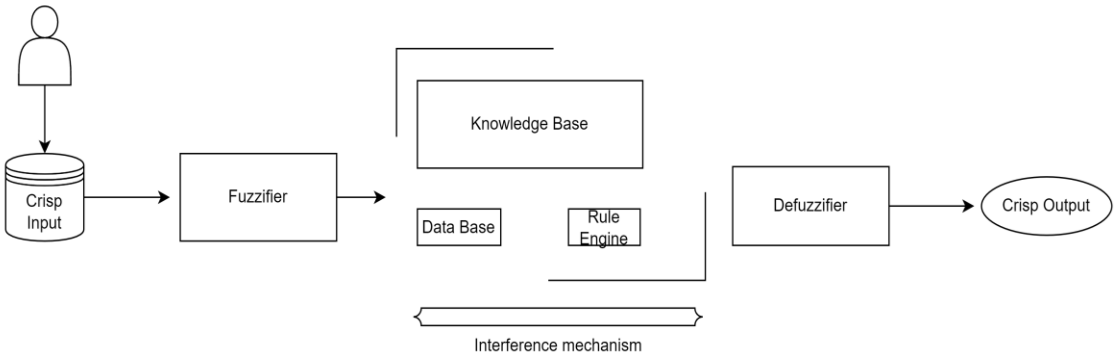



Figure 3.8. The architecture of Sugeno type FIS.

The ANFIS handles the linguistic data through membership functions and uses the rules derived from the ANN to predict the results. Various functions are used to create membership functions, such as triangular, trapezoidal, sigmoid, Gaussian, etc. At the same time, for determining a membership, the type of transaction and the data size [97] are essential to represent the uncertainty in measurements properly [98]. The Gaussian membership function is depicted in Figure 3.9. It was preferred here as it was often found more successful than the other membership functions in different applications.

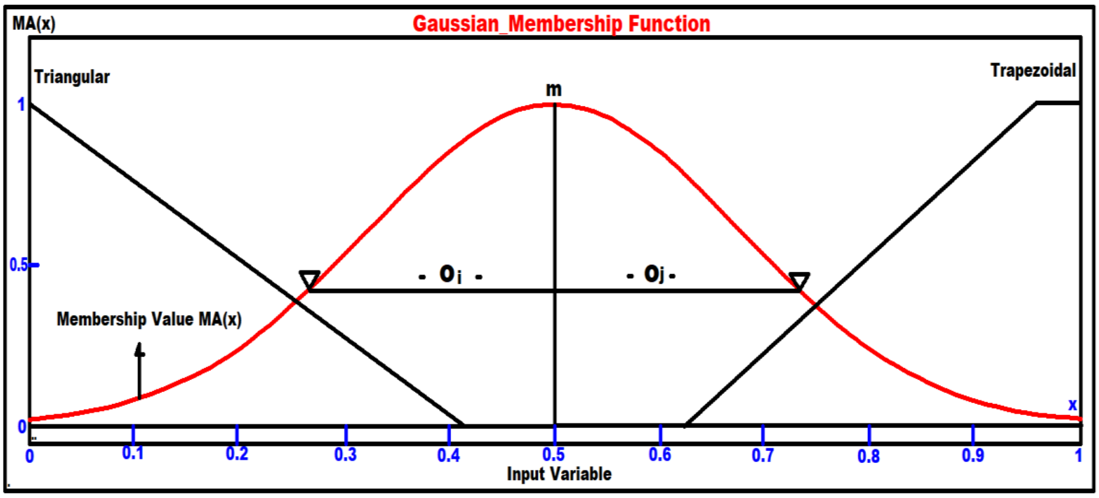



Figure 3.9. The Gaussian membership function and its parameters (adapted from [94]).

(h) $$M_{Gauss} = \frac{1}{e^{\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}} \; ,$$

Where, $m$ denotes the mean value, and the $\sigma$ represents the standard deviation.

### 3.2.5. Multiple Regression Analysis

The MRA is a ststistical technique that reveals the linear or non-linear relationship between dependent and independent variables. In addition, it is defined as a conventional supervised ML algorithm and frequently used in many scientific fields for comparing the ML methods [1, 31, 99, 100]. The Stepwise Regression is a type of MRA method which establishes a linear model as a result of the independent variables in a step by step manner according to the significance of each variable on the dependent variable.

The equation of the linear MRA is;

(i) $$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_j X_{ji} + \varepsilon_i,$$

where;

- $Yi$ = $i^{th}$ observation of the target
- $X_{1..j}$ = Dependent variable
- $\beta_0$ = Intercept
- $\beta_{1..j}$ = Coeeficients
- $\varepsilon$ = Error term.

The Stepwise Regression is frequently preferred due to several features such as being a natural data reduction method that protects the model from the curse of dimensionality [101]. In addition, being a simple and interpretable [102] method is another important feature [1]. Although the method has been often classified as a supervised ML method, the stepwise regression is a kind of semi-supervised method. Additionally, it can also be defined as a semi-expert ML method due to the possibility of expert intervention in the variable selection. Despite the advantages of stepwise based MRA, there are some weaknesses that also affect the model in some significant perspectives, such as relatively low model prediction accuracy and weak representation capacity of high dimensional data. Furthermore, since it is a linear model, it fails to reveal non-linear interactions between variables [1]. Furthermore, pruning the variables decreases the model stability. Therefore, the method may produce unstable results in some datasets.

### 3.3. Performance Assessment Metrics

Several performance metrics were employed from a ML perspective for the quantitative evaluation of the performance of MARE methods with complex and high dimensional data structures. The metrics can be categorized based on their broad use in two different communities. In the first group, the statistical ML metrics such as the root mean square error (RMSE), $R^2$ and the Adjusted $R^2$ (the total variance explained from the model), can be listed. Their descriptipons with formulas (j, k, l) are given below.

- *RMSE:* is a derivative of the mean squared error *(MSE)* of model predictions. The *MSE* sums the squared difference between the predictions and the target values and averages these values. It can be interpreted as the higher the *MSE* value, the more incorrect the model prediction. The *RMSE* is the square root of *MSE*. It was preferred in this study since the errors are measurable in the target.

(j) $$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (\, Actual\ Price_i - \ Estimated\ Price_i)^2}{N}}$$

- *R Squared ($R^2$)*: is an indicator that indicates the proportion of the variance for a target explained by an explanatory variable in a model.

(k) $$R^2 = 1 - \left(\frac{MSE}{\sigma_y{}^2}\right)$$

- *Adjusted R Squared (Adj $R^2$)*: is an $R^2$ indicator that penalizes redundant or ineffective variables by subtracting them from the sample size. It produces results that are more meaningful when the number of variables is more than one.

(l) $$Adj\ R^2 = 1 - (1 - R^2)\left(\frac{N-1}{N-P-1}\right),$$

where;

- *N* is the number of observations,
- *P* is the number of variables,
- *MSE* is the mean square error (square of *RMSE*),
- $\sigma y$ is the standard deviation obtained from the sample set

The second group of quality metrics is sourced from the conventional REA metrics used to evaluate the mass appraisal. The metrics include Coefficient of Dispersion (*COD*), Price-Related Differential (*PRD*) and Coefficient of Variation (*COV*) and Standard Deviation (*SD*). The definitions are given below.

- *COD*: is a widely used measure of appraisal uniformity. It is expressed as the percentage of the average deviation of the ratios from the median. A *COD* value should be in the range of 5-15. A lower *COD* value indicates outliers or non-representative samples [16].

- *PRD*: measures the regressivity or the progressivity of the assessments. The regressive appraisals occur when high-valued real estates are under-appraised relative to low-valued ones. Progressive appraisals indicate the occurrence of the opposite pattern. The *PRD* value should be in the range of 0.98 to 1.03 to indicate vertical equity[16].

- *COV*: is used especially for the MARE context. The *COV* is a measure of relative variability and is defined as the standard deviation ratio to the mean. The higher values have more variances from the mean.

- *SD*: is the average distance of the ratios from the ratio mean [1, 16].

(m) $$R_i = \frac{A_i}{S_i}, \qquad (i = 1, \dots, n)$$

(n) $$COD = \%100 \; x \; \frac{\frac{(\sum_1^n |R_i - \check{R}|)}{n}}{\check{R}}$$

(o) $$PRD = \frac{\bar{R}}{Wt.\bar{R}}$$

(p) $$Wt.\bar{R} = \frac{\sum_1^n A_i}{\sum_1^n S_i}$$

(r) $$COV = \frac{\sigma_y}{Mean} * 100 \,,$$

where;

- $n$ Number of observations

- $A$ Numerator of the *i-th* ratio ($i = 1, \ldots, n$). (Appraisal Value)
- $S_i$ Denominator of the *i-th* ratio ($i = 1, \ldots, n$). (Actual Sale Price)
- $R_i$ The *i-th* ratio ($i = 1, \ldots, n$). (Appraisal Ratio)
- $\check{R}$ Median
- $\bar{\text{Ṛ}}$ Mean
- $Wt.\ \bar{Ṛ}$ Weighted Mean

# 4. APPLICATION

In this Chapter, the study area, the datasets, the pre-processing results and the outputs of the five ML methods used in the thesis are presented.

## 4.1. The Study Dataset and Pre-Processing Results

In this thesis, five data-driven ML approaches were applied for the MARE of several properties in the study area of Mamak district of Ankara, Turkey (Figure 4.1). The area was selected for the variability in property object types and their prices. Mamak is an urban expansion area with diverse socio-economic conditions. Besides the old and newly constructed buildings, it has slum areas and urban transformation projects recently implemented in some parts.

The data used in this study were obtained within the scope of this thesis and also from a mass valuation pilot project carried out by the GDLRC [6]. The pilot project data was analyzed by the real estate appraisers and GDLRC experts, as well as two experts from the World Bank who served as consultant. There were approximately 20,000 samples and more than 50 independent variables collected in the project. The prices were obtained from licensed valuation institutions, through the field work, and from real estate companies. Their reliability was tested by comparing them with the data obtained from official institutions. Different subsets of data were previously analyzed as part of the doctoral research and the results are either published [1] and or submitted for publication [94]. The analysis results of another subset were presented at the international International Federation of Surveyors (FIG) conference [10].

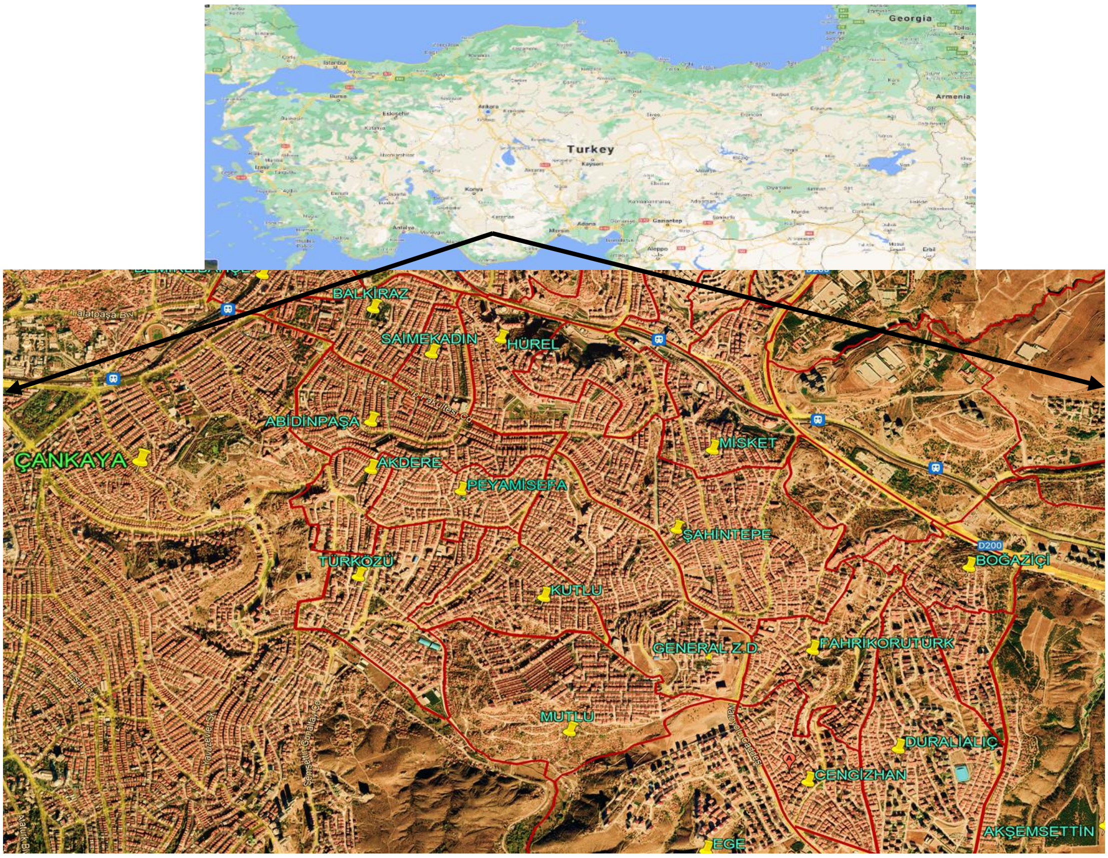



Figure 4.1. The location of the study area with centre coordinates at 39.9243° latitude, 32.9160° longitude at the centre ([94]).

In addition, further data collection efforts were made to obtain additional information, such as court reports, real estate companies, and real estate valuation experts within the scope of doctoral studies. A dataset composed of 4100 samples and 38 variables (Appendix A) was analyzed in the thesis. In this experimental study, the dataset was reduced in terms of both the number of samples and the number of variables to obtain a reliable model for the purpose of automated mass appraisal. After the pre-processing, the number of data samples was reduced to 3936, and the number of variables was reduced to 12. As a result, a dataset consisting of residences and a small number of commercial properties was obtained. The data processing steps are depicted in Figure 4.2.

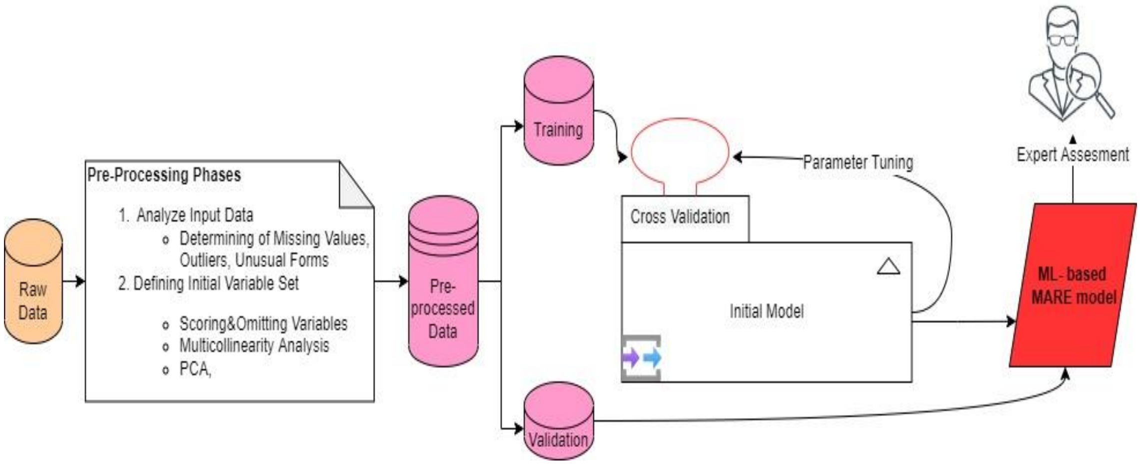
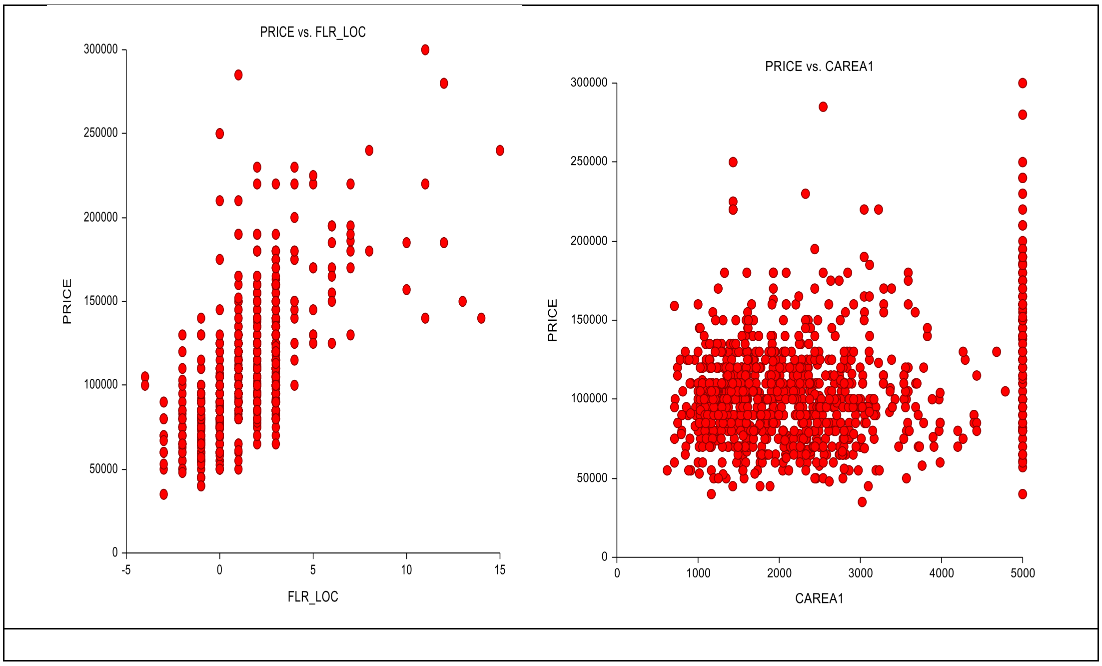



Figure 4.2. The architecture of ML-based MARE model applied here.

The Statistical Variable Importance Analysis (significance analysis) in stepwise-based MRA was performed on the initial variable list. According to the expert analysis, the independent variables insignificant for the dependent variable target (price) or had insufficient significance levels were omitted. The graphs of some of the important variables vs. sales prices are given in Figure 4.3.



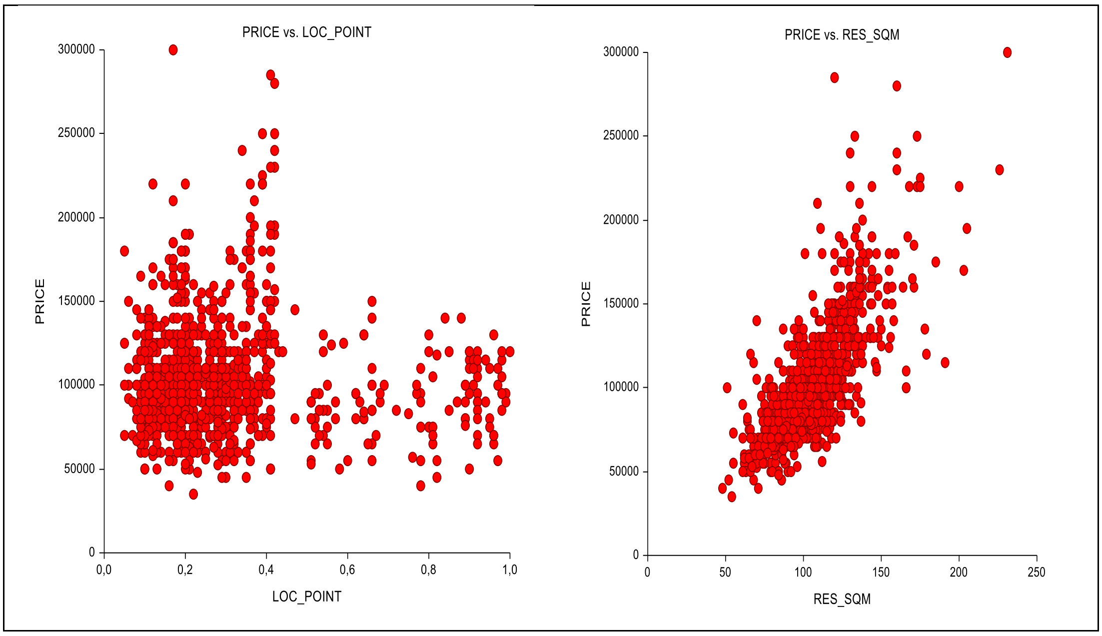



Figure 4.3. Sales Price vs X variable's plots.

In this first data reduction step, variables numbered 12, 28-33 and 36 (see Appendix A) were removed from the initial variable list due to the significance analysis. The most interesting elimination was the year built (YBUILT : 36) variable since it is commonly expected that the building year would be an important variable for appraising residential prices. Although further importance tests were applied, the results were similar to the linear regressive variable importance test. A multicollinearity test was performed on zonal variables in the second step of the preprocessing stage. Another five variables with a correlation ratio greater than 0.7 were omitted from the model, as reflected with red color in Table 4.1.

Table 4.1. Multicollinearity Analysis

| | DIS-SOC | DIS-BUS | DIS-CEN | DIS-COL | DIS-ROD | DIS-CUL | DIS-MET | DIS-HOS | DIS-UNI | DIS-TRA | DIS-HOS | DIS-MAL | DIS-MRK |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DIS-SOC | **1.00** | 0.12 | 0.00 | 0.20 | 0.16 | 0.04 | -0.14 | -0.03 | 0.12 | -0.02 | -0.00 | -0.05 | 0.19 |
| DIS-BUS | 0.12 | **1.00** | 0.59 | -0.11 | **0.86** | 0.26 | 0.53 | 0.30 | **0.96** | 0.44 | **0.57** | 0.07 | -0.04 |
| DIS-CEN | 0.00 | 0.59 | **1.00** | 0.02 | **0.71** | 0.16 | **0.88** | 0.27 | 0.48 | 0.03 | **0.80** | 0.51 | -0.16 |
| DIS-COL | 0.20 | -0.11 | 0.02 | **1.00** | -0.01 | 0.12 | -**0.91** | -0.09 | -**0.88** | 0.07 | 0.11 | 0.21 | 0.13 |
| DIS-ROD | 0.16 | **0.86** | **0.71** | -0.01 | **1.00** | 0.25 | 0.46 | 0.30 | **0.77** | 0.47 | 0.61 | 0.24 | 0.03 |
| DIS-CUL | 0.04 | 0.26 | 0.16 | 0.12 | 0.25 | **1.00** | 0.15 | 0.30 | 0.26 | 0.19 | 0.16 | 0.05 | 0.13 |
| DIS-MET | -0.14 | 0.53 | **0.88** | -**0.91** | 0.46 | 0.15 | **1.00** | 0.26 | 0.45 | 0.00 | **0.77** | 0.47 | -0.20 |
| DIS-HOS | -0.03 | 0.30 | 0.27 | -0.09 | 0.30 | 0.30 | 0.26 | **1.00** | 0.21 | 0.11 | 0.10 | -0.03 | -0.03 |
| DIS-UNI | 0.12 | **0.96** | 0.48 | -**0.88** | **0.77** | 0.26 | 0.45 | 0.21 | **1.00** | 0.45 | 0.60 | 0.13 | 0.05 |
| DIS-TRA | -0.02 | 0.44 | 0.03 | 0.07 | 0.47 | 0.19 | 0.00 | 0.11 | 0.45 | **1.00** | 0.12 | -0.05 | 0.29 |
| DIS-HOS | -0.00 | 0.57 | **0.80** | 0.11 | 0.61 | 0.16 | **0.77** | 0.10 | 0.60 | 0.12 | **1.00** | **0.79** | 0.08 |
| DIS-MAL | -0.05 | 0.07 | 0.51 | 0.21 | 0.24 | 0.05 | 0.47 | -0.03 | 0.13 | -0.05 | **0.79** | **1.00** | 0.25 |
| DIS-MRK | 0.19 | -0.04 | -0.16 | 0.13 | 0.03 | 0.13 | -0.20 | -0.03 | 0.05 | 0.29 | 0.08 | 0.25 | **1.00** |

In the last step, expert-based variable analysis and the PCA were applied under different scenarios. Based on the results;

- The R-SQM and R-OPSQM variables were consolidated because their significance levels were similar.

- A total of eight variables reflecting various zonal characteristics (i.e., DIS-MET, DIS-CUL, DIS-SOC, DIS-HOS, DIS-MAL, DIS-MRK, DIS-TRA, DIS-UNI) were combined into a single variable with the help of PCA. The obtained variable was scored according to the significance level of the variable with the D-STTS variable, but in this scoring D-STTS significance rate was almost "0". As a result, the "location score variable" was derived from the zonal variables as can be seen in Table 4.2.

Table 4.2. The PCA results.

| **Total Variance Explained** | | | | | | |
|---|---|---|---|---|---|---|
| **Component** | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | |
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| **1** | 5.933 | 74.160 | 74.160 | 5.933 | 74.160 | 74.160 |
| **2** | 1.447 | 18.085 | 92.245 | | | |
| **3** | 0.385 | 4.811 | 97.056 | | | |
| **4** | 0.182 | 2.269 | 99.325 | | | |
| **5** | 0.031 | 0.393 | 99.718 | | | |
| **6** | 0.014 | 0.180 | 99.898 | | | |
| **7** | 0.007 | 0.089 | 99.987 | | | |
| **8** | 0.001 | 0.013 | 100.000 | | | |

| Communalities | |
|---|---|
| Location Point | Extraction |
| DisHos | 0.781 |
| DisMark | 0.782 |
| DisMall | 0.435 |
| DisTrain | 0.811 |
| DisMetro | 0.806 |
| DisCul | 0.680 |
| DisSoc | 0.875 |
| DisUni | 0.764 |

After applying the data reduction techniques, the two datasets, such as the initial one consisting of 38 variables and the final one with 12 variables, were compared by using the $R^2$ values from the trial-and-error based linear regression method. As a result, no significant accuracy loss was observed in the model explainability rate. Therefore, the final variable list (Table 4.3) was used in the ML models.

Table 4.3. The final variable list and their descriptions.

| Variable Name | Description | Measure | Role |
|---|---|---|---|
| **NEIGHBOURHOOD** | Neighbourhood | Nominal | None |
| **STREET** | Street | Scale | None |
| **IU-ID** | Individual Unit ID | Scale | None |
| **BLOK** | Blok | Scale | None |
| **PARCEL** | Parcel | Scale | None |
| 1. **NROOMS** | Number of Rooms | Nominal | Input |
| 2. **NBATH** | Number of Baths | Nominal | Input |
| 3. **NIU** | Number of Individual Units | Scale | Input |

| | | | |
|---|---|---|---|
| 4. **NBF** | Number of Building Floor | Nominal | Input |
| 5. **PARLOC** | Location of Building in Parcel | Nominal | Input |
| 6. **FLR_LOC** | Floor Location | Scale | Input |
| 7. **CAREA** | Construction Right (Area) | Scale | Input |
| 8. **ELVTR** | Building Has or Not Elevator | Nominal | Input |
| 9. **CMPLX** | Building is in Campus or Not | Nominal | Input |
| 10. **LOC_GRADE** | Location Grade | Scale | Input |
| 11. **RES_SQM(m²)** | Residential Area | Scale | Input |
| 12. **PRICE (₺)** | PRICE | Scale | Output |

In addition to the dimension reduction, data cleaning was applied to detect outliers and missing data were. As a result, 171 data samples were omitted and the remaining 3929 data samples were utilized as training dataset. A total of 2929 data samples were randomly determined as training data for k-fold cross-validation (CV) technique. The CV is a resampling method used to validate a training dataset with different combinations and to generalize the model for different datasets by finding the most effective result. The statistical analysis results of the training data samples are given in Table 4.4.

Table 4.4. The training data characteristics.

| **Variable** | **Min** | **Max** | **Mean** | **Std. Dev.** |
|---|---|---|---|---|
| NROOMS | 2 | 5 | 3.91 | 0.44 |
| NBATH | 1 | 2 | 1.82 | 0.38 |
| NIU | 1 | 134 | 10.31 | 8.49 |
| NBF | 2 | 36 | 6.77 | 3.126 |
| PAR-LOC | 1 | 2 | 1.44 | 0.49 |
| FLR_LOC | -4 | 15 | 1.5 | 2.107 |

| | | | | |
|---|---|---|---|---|
| CAREA | 173 | 3,285,900 | 63,546 | 232,400 |
| ELVTR | 1 | 2 | 1.59 | 0.49 |
| CMPLX | 1 | 2 | 1.94 | 0.24 |
| LOC_GRADE | 0 | 1 | 0.30 | 0.22 |
| RES_SQM($m^2$) | 40 | 275 | 105.89 | 22.08 |
| PRICE (₺) | 35,000 | 340,000 | 104,651 | 33,730 |

The remaining 1000 data samples were randomly selected for model testing from the whole dataset. Although it was not required to split the test data at the CV stage, this approach was preferred to be able to use the same test set in order to compare the ML models properly. While splitting the dataset as training and test, attention was paid to the distribution of test set. As shown in Table 4.5, the test set was chosen randomly in 32 neighbourhoods (sub-districts). The mean and median values obtained from the appraisal/sale price ratio of each sample derived for each neighborhood are also given in the Table.

Table 4.5. The number of samples and the mean and median values of appraisal/sale (A/S) price ratio obtained from test data in different sub-districts.

| **Neighbourhood** | **Number of samples** | **Median A/S ratio** | **Mean COD A/S ratio** |
|---|---|---|---|
| ABIDINPAŞA | 7 | 0.9971 | 1.0889 |
| AKDERE | 21 | 1.0863 | 1.0578 |
| ALTIAĞAÇ | 2 | 1.1795 | 1.1795 |
| AŞIK VEYSEL | 6 | 1.0336 | 1.0117 |
| BAHÇELER.U. | 6 | 0.9537 | 0.9622 |
| BAŞAK | 18 | 0.9581 | 0.9924 |
| BOĞAZIÇ | 40 | 1.0368 | 1.0796 |
| BOSTANCIK | 6 | 1.1185 | 1.0839 |
| CENGIZHAN | 31 | 1.0616 | 1.0690 |
| ÇAĞLAYAN | 5 | 1.0386 | 1.0431 |
| DERBENT | 7 | 1.0615 | 1.0457 |
| DURALI ALIÇ | 101 | 0.9839 | 1.0077 |
| EGE | 19 | 0.9115 | 0.9263 |
| EKIN | 59 | 1.0170 | 1.0165 |
| FAHRI KOR | 33 | 1.0439 | 1.0580 |

| | | | |
|---|---|---|---|
| GENERAL Z.D. | 92 | 0.9988 | 1.0081 |
| HARMAN | 12 | 1.0324 | 1.1078 |
| HÜREL | 6 | 1.0881 | 1.1052 |
| HÜSEYING | 5 | 0.9603 | 1.0341 |
| KARTALTEPE | 11 | 1.1165 | 1.1480 |
| KAZIM ORBA | 39 | 0.9805 | 0.9939 |
| KÜÇÜK K. | 47 | 1.0325 | 1.0330 |
| MISKET | 51 | 1.0516 | 1.0514 |
| MUTLU | 178 | 0.9923 | 1.0007 |
| PEYAMI SEFA | 54 | 1.0221 | 1.0335 |
| ŞAFAKTEPE | 1 | 0.8608 | 0.8608 |
| ŞAHAP G. | 7 | 1.0140 | 1.0269 |
| ŞAHINTEPE | 77 | 1.0498 | 1.0428 |
| ŞEHIT C.T. | 2 | 0.9804 | 0.9804 |
| ŞIRINTEPE | 15 | 1.0318 | 1.0735 |
| TUZLUÇAYI | 7 | 1.0000 | 1.0277 |
| TÜRKÖZÜ | 35 | 1.0055 | 0.9835 |
| | **Average** | **1.0155** | **1.0249** |

## 4.2. The Multiple Regression Analysis Results

The stepwise linear regression method was chosen as the first method for the comparative study because it can easily handle the variables in the model according to their significance order. The model parameters of the stepwise MRA method was given in Table 4.6.

Table 4.6. The parameterization of the stepwise MRA method.

| **Parameter** | **Value** |
|---|---|
| Learning Rule | Stepwise Linear Regression |
| Input and Output | (8+3) x 1 |
| Number of training samples | 2936 (75 % of all samples) ( k=5 Cross Validation) |

Although eight of the independent variables were selected as the most important variables in the model (Table 4.7), the three variables omitted from the model were added again by adjusting preconditions for the model parameters to use the same number of variables as in the other ML methods.

Table 4.7. The model summary obtained from the stepwise MRA method.

| **Stepwise Training Variable Selection and Model Summary[i]** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | |
| | | | | | R Square Change | F Change | df1 | df2 | Sig. F Change |
| 1 | ,729[a] | 0.531 | 0.531 | 23105.8 | 0.531 | 3312.8 | 1 | 2927 | 0 |
| 2 | ,783[b] | 0.613 | 0.613 | 20986.00 | 0.082 | 622.18 | 1 | 2926 | 0 |
| 3 | ,795[c] | 0.633 | 0.632 | 20457.84 | 0.019 | 154.03 | 1 | 2925 | 0 |
| 4 | ,806[d] | 0.650 | 0.650 | 19966.17 | 0.018 | 146.83 | 1 | 2924 | 0 |
| 5 | ,814[e] | 0.662 | 0.662 | 19622.13 | 0.012 | 104.43 | 1 | 2923 | 0 |
| 6 | ,815[f] | 0.664 | 0.663 | 19577.09 | 0.002 | 14.46 | 1 | 2922 | 0. |
| 7 | ,815[g] | 0.665 | 0.664 | 19553.93 | 0.001 | 7.92 | 1 | 2921 | 0.005 |
| 8 | ,816[h] | 0.665 | 0.665 | 19535.19 | 0.001 | 6.6 | 1 | 2920 | 0.010 |
| h. Predictors: (Constant), RES_SQM, FLR_LOC, NBF, LOC_POINT, ELVTR, PAR-LOC, CMPLX, NROOMS | | | | | | | | | |
| i. Dependent Variable: PRICE | | | | | | | | | |

In the MRA stepwise model, the regression error measurement indicators, i.e., the $R^2$, RMSE and Adj-$R^2$, and the appraisal ratio indicators, i.e., COV, COD, and PRD were utilized to evaluate the model prediction accuracy. The results of the MRA model are shown in Table 4.8, and the detailed results of the appraisal ratio was given in Appendix B – Table B.1.

Table 4.8. Validation results of the MRA based MARE model.

| Method | Ratio Median | Ratio Mean | SD (TRY) | COD | PRD | COV | RMSE (TRY) | $R^2$ | Adj-$R^2$ |
|---|---|---|---|---|---|---|---|---|---|
| MRA | 1.0115 | 1.0290 | 0.161 | 12.28 | 1.024 | 15.19 | 17484.2 | 0.74 | 0.72 |

According to the regression metrics given in Table 4.8, the $R^2$ and Adj-$R^2$ values are comparable but not identical. It shows that the variables added to this model are consistent with the explainability rate of the overall model, but the model could not be declared as a stable model. In addition, when the RMSE values of 17,482.9 and 0.72 $R^2$ values are evaluated together, it was observed that the MRA stepwise model can be used in MARE studies. On the other hand, the model produced results in accordance with the COD and the PRD indicator range recommended by the IAOO [16] in MARE evaluations. The COV parameter is an indicator that can be interpreted more meaningfully than the results obtained by the other methods. In the MARE context, it is a measure of relative variability and is defined as the ratio of the standard deviation to the mean. The higher values have more variance than the mean, which can be interpreted as a decrease in the model predictive accuracy.

### 4.3. The ANFIS Results

The ANFIS model used in this study has a six-layered architecture as shown in Figure 4.4. The model parameters were selected carefully for a successful application of the ANFIS method. The eight tunable parameters are given in Table 4.9.

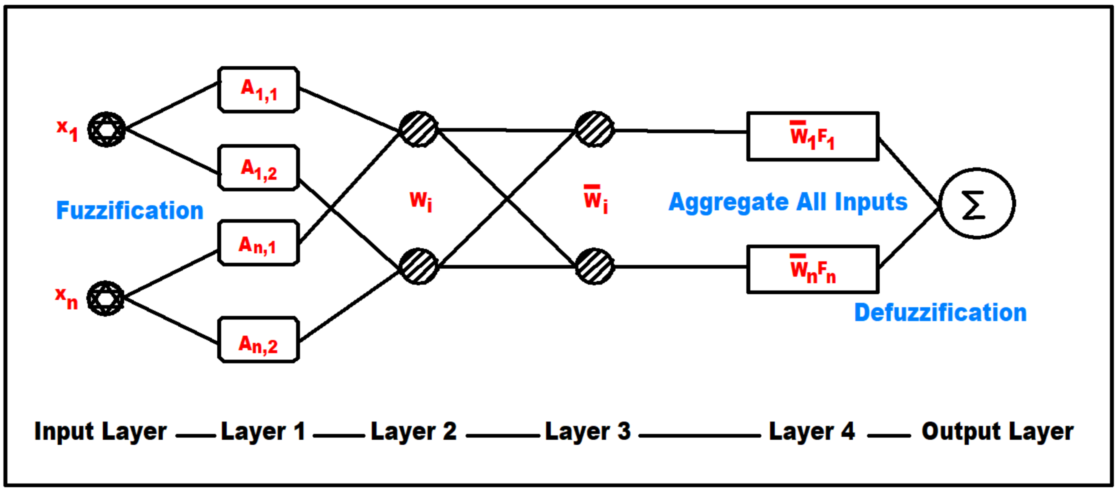



Figure 4.4. The architecture of the ANFIS employed here ([94]).

Table 4.9. The parameterization of the ANFIS method.

| **Parameters** | **Value** |
|---|---|
| Input Membership Function | Gauss |
| Output Membership Function | Linear |
| Learning Rule | Back Propagation |
| Epochs | 50 |
| De-fuzzification Rule | Weighted Average |
| Input and Output | 11 x 1 |
| Number of training samples | 2936 (~75% of all samples) ( k=5 Cross Validation) |
| Number of validation samples | 1000 (~25 % of all samples) |

In the selection of the membership function, the Gaussian Membership Function was chosen as it can handle complex datasets and represent the uncertainty in measurements [98]. In determining the learning rule, the BP method, which is flexible and easy to apply, was chosen as the learning algorithm as it does not require prior network knowledge [80]. The training and validation samples was selected randomly similar to the other methods.

In the ANFIS model, rthe egression error indicators, i.e., $R^2$, RMSE Adj-$R^2$, and the appraisal ratio indicators (COV, COD, and PRD) were utilized to evaluate the model prediction accuracy. The results of the ANFIS model is shown in Table 4.10, and detailed results of the appraisal ratio was given in Appendix C – Table C.1

Table 4.10. Validation results of the ANFIS based MARE model.

| **Method** | **Ratio Median** | **Ratio Mean** | **SD** (TRY) | **COD** | **PRD** | **COV** | **RMSE** (TRY) | **$R^2$** | **Adj-$R^2$** |
|---|---|---|---|---|---|---|---|---|---|
| ANFIS | 1.0128 | 1.0275 | 0.156 | 12.226 | 1.023 | 15.15 | 18,229 | 0.72 | 0.72 |

According to the regression metrics given in Table 4.10, the $R^2$ and Adj-$R^2$ values are the identical, which indicates the consistency of the variables added to the model, and also the explainability rate of the overall model. In addition, when the RMSE values of 18,229 TRY and the $R^2$ value of 0.72 are evaluated together, it was observed that the ANFIS model can also be used in MARE studies. Furthermore, the ANFIS model produced results compliant with the COD and PRD indicator ranges recommended by IAOO [16] in MARE measurements.

### 4.4. The Support Vector Machines Results

In SVM, the linear kernel serves to separate and represent data in two-dimensional input space. However, it could not produce effective results with multidimensional input space as in the MARE data [10]. According to the trial-and-error method for determining kernel function, four different types of kernel-based SVM were implemented and compared such as GRBF, hyperbolic tangent, sigmoid and linear. The GRBF-based SVM outperformed other kernel-based methods, as shown in Table 4.11. Therefore, due to its ability to accommodate nonlinear mapping in high dimensional input space, the GRBF selected as the kernel function. The parametrization of the GRBF-SVM based MARE experimental study is shown in Table 4.12.

Table 4.11. The SVM results obtained from four different kernel functions.

| **Trial-and-Error for Kernel Function** | **RMSE (TRY)** | **Verify** |
|---|---|---|
| Gaussian radial basis function (RBF) | 18,955 | Yes |
| Hyperbolic tangent kernel | 19,773 | No |
| Sigmoid kernel | 22,229 | No |
| Linear Kernel | 22,475 | No |

Table 4.12. The parameterization of the SVM method implemented in this study.

| | |
|---|---|
| Kernel Function | Gaussian – Radial Basis Function |
| Kernel Scale | 4.4 |
| γ | Semi-automatic |
| Box Constraints- C | Semi-automatic |
| Resambling Method | K=5 Cross Validation |
| Number of training samples | 2936 (~75 % of all samples) ( k=5 Cross Validation) |
| Number of validation samples | 1000 (~25 % of all samples) |
| Number of iterations | 15 |

The GRBF-based SVM seeks to find local minima and soft margin that separates all positive and negative samples. The kernel function for the GRBF is given below (s).

(s) $$K(X_I X_J) = exp(-\gamma \left\| X_I X_J \right\|^2, \gamma > 0$$

Where; $\gamma$, the regularization parameter, is determined by CV scales the squared distance. The $X_i$ $X_j$ are support vectors, and $K$ is the kernel function coefficient. $C$ parameter is the penalty parameter of the error term for outliers or unusual samples to conserve in bag stability of the dataset. It is the soft margin cost function parameter that adjusts support vector influences.

On the other hand, the regularization parameter, $\gamma$, has a major effect on the variance and stability of the model. Additionally, $\gamma$ is the free parameter of the GRBF. In the GRBF-based SVM, the optimized gamma values mean a large variance and a wider kernel. In addition, a wider kernel means a simpler model. Otherwise, the large gamma leads to a high bias, low variance, and a narrow kernel. So, each data group in a data set has more influence on the models. Thus, the optimal $C$ and $\gamma$ values need to adjust kernel scale and thus building a stable model, which can capture the complexity of the input data. Lower $C$ values usually lead to more splitting, i.e., the generation of more support vectors, which increases the prediction time. Therefore, it is important to balance the model's ability to represent the data and the prediction time when adjusting the $C$ value.

The difference between each prediction and actual value was minimized during the model building phase by adjusting C and $\gamma$ coefficients in MATLAB [63]. In determining these parameters, different numbers were attained to the module. The iterations stop when the adjusted convergence criterion rate or the value is achieved according to the specified algorithm [10]. The maximum number of iterations was set to 15, since the use of GRBF increases the training time, which may also lead to over-training [10]. As a final step, after the calculating kernel function, the plane that represents the high dimensional data best is determined. In this way, the interactions between the independent and the target variables are also described.

In Figure 4.5, the SVM based MARE results and the actual prices were compared. It can be seen in the Figure that the predictions of samples that are close to the average are stable and successful. On the other hand, the model produces lower accuracy results in predicting high-priced assets that can be seen on the upper-right side of the graph.

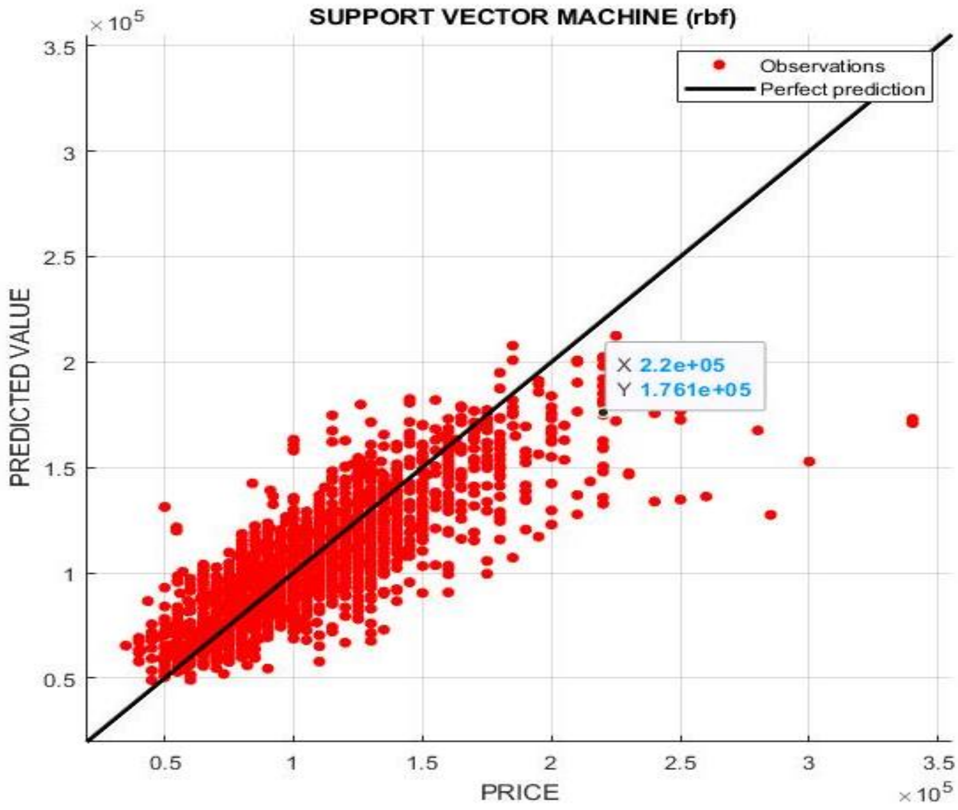



Figure 4.5. The comparison of predicted values and the actual prices of SVM based MARE model.

When evaluating the model prediction accuracy in the SVM model, the $R^2$, RMSE Adj-$R^2$ and the appraisal ratio indicators (COV, COD, and PRD) were utilized. The results of the model are shown in Table 4.13, and the detailed results of the appraisal ratio are given in Appendix D – Table D.1

Table 4.13. The validation results of the SVM-based MARE model.

| **Method** | **Ratio Median** | **Ratio Mean** | **SD (TRY)** | **COD** | **COV** | **PRD** | **RMSE (TRY)** | **$R^2$** | **Adj-$R^2$** |
|---|---|---|---|---|---|---|---|---|---|
| SVM | 1.0009 | 1.0128 | 0.143 | 10.867 | 1.024 | 14.16 | 16,916 | 0.77 | 0.76 |

According to the regression metrics presented in Table 4.13, the $R^2$ and Adj-$R^2$ values were similar. It shows that the additional variables were consistent with the explainability rate of the overall model. In addition, when the RMSE values and $R^2$ values are evaluated together, the SVM method can cope with complex datasets in MARE experimental implementation. Moreover, the SVM model produced results following the COD and the PRD indicator ranges recommended by IAOO [16] in MARE applications.

## 4.5. The Artificial Neural Networks Results

In this study, the effectiveness of the ML methods in the experimental MARE study was measured, and the model was created with the ANN method. Table 4.14 shows the parameters of the ANN model used in this study.

Table 4.14. Parameterization of the ANN method.

| | |
|---|---|
| Learning algorithm | Back Propagation |
| Number of hidden layers | 10 |
| Resampling method | K=5 Cross-Validation - Randomly |
| Number of training samples | 2936 (~75 % of all samples) |
| Number of validation samples | 1000 (~25 % of all samples) |
| Number of iterations | 15 |

The BP is a learning algorithm used in ANN to compute a gradient descent concerning connection weights [80]. The number of hidden layers plays a significant role in the prediction accuracy of the ML models. If the number of hidden layers increases in the ANN model, the model will produce better results [103]. However, increasing the number of hidden layers will increase the data processing time. On the other hand, overfitting problems may also arise. For this reason, it was desired to choose a balanced number of hidden layers, and the number of hidden layers of 10, which is the reference value of the

MATLAB [63] program, was determined by applying the trial-and-error method. Then the model captured the best validation score at the 21$^{st}$ epoch as depicted in Figure 4.6.

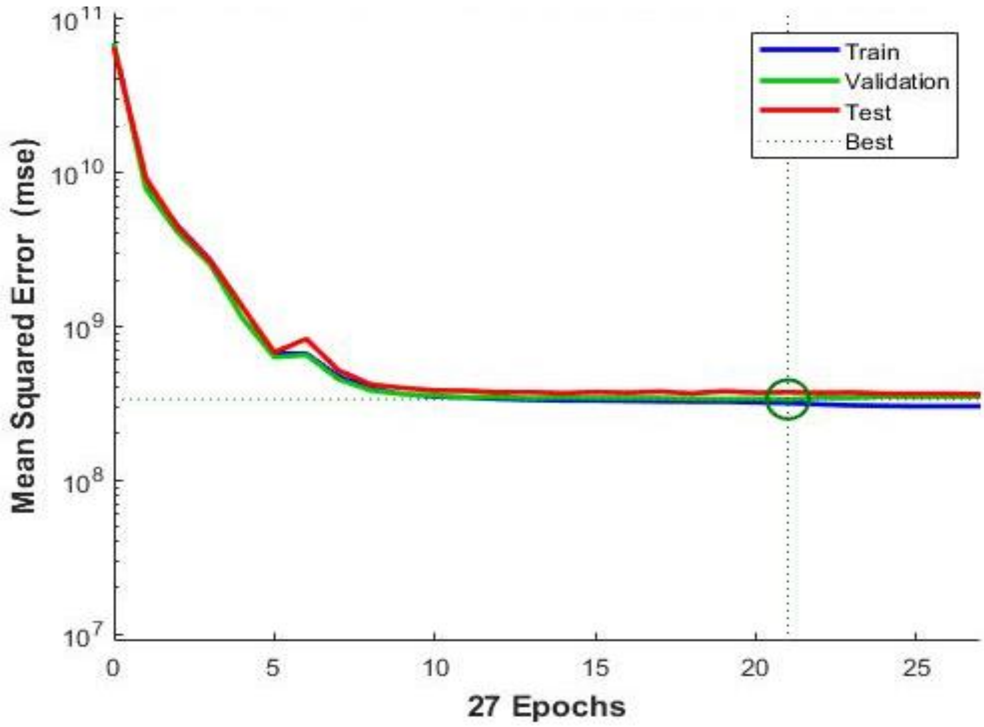



Figure 4.6. Validation performances at different epochs and the optimal value obtained from the ANN .

In Figure 4.7, the ANN-based MARE results and the actual prices were compared. It can be seen in the graph that most of the predictions were compatible with the model, but in the samples in the upper left part of the graph, it can be seen that the results were higher than the actual prices.


Figure 4.7. The comparison of predicted values and the actual prices of the ANN-based MARE model.

The ANN model uses regression measurement indicators to evaluate the model's prediction accuracy; $R^2$, RMSE Adj-$R^2$ and MARE valuation ratio indicators COV, COD and PRD. The results of the model are shown in Table 4.15, and the detailed results regarding the appraisal rate were given in Appendix E – Table E.1.

Table 4.15. Validation results of the ANN-based MARE model

| Method | Ratio Median | Ratio Mean | SD (TRY) | COD | COV | PRD | RMSE (TRY) | $R^2$ | Adj-$R^2$ |
|---|---|---|---|---|---|---|---|---|---|
| ANN | 1.0070 | 1.0189 | 0.131 | 9.947 | 12.86 | 1.023 | 15,279 | 0.81 | 0.80 |

According to the regression metrics given in Table 4.15, the $R^2$ and Adj-$R^2$ values are very close. It shows that the additional variables were consistent with the explainability rate of the overall model. In addition, when the RMSE and $R^2$ values are evaluated together, it can be said that the ANN method can accurately cope with complex datasets in MARE experimental implementations. Moreover, the ANN model results follow the COD and the PRD indicator ranges recommended by IAOO [16] for MARE applications.

## 4.6. The Random Forest Results

The fifth method used for the experimental work on the MARE study was RF. The Table 4.16 shows the parameters of the RF model used in this exploratory study.

Table 4.16. The parameterization of the RF model used in this study.

| Parameters | Value |
|---|---|
| Ensemble Method | Bagging (Bootstrap Aggregating) – Regressive |
| Minimum Leaf Size | 8 |
| Number of Learners/Trees | 100 |
| Combination of Weights | Weighted Average |
| Input and Output | 11 x 1 |
| Number of training samples | ~2936 (75 % of all samples) ( k=5 Cross Validation) |
| Number of validation samples | ~1000 (25 % of all samples) |

In general, the more trees used to build the model, the more accurate the predictions will be obtained. However, since the number of trees will lose its influence after a certain value, it will only increase the computational cost. Therefore, it is necessary to balance the computational cost and to improve the estimates. In the Matlab regression toolbox [63], if the minimum leaf size is reduced to 7 or smaller, the processing time increases as many sub-samples create several estimators. Likewise, if the number of learners is greater than 100, both the transaction complexity and the processing time would increase.

On the other hand, the OOB estimation error [39] determines the number of trees used in the RF model. As mentioned previously, it was used for the prediction accuracy of the RF models. The OOB error graphics were used to select the number of trees in a bagging method by Yilmazer and Kocaman [1], which can be seen in Figure 4.8.

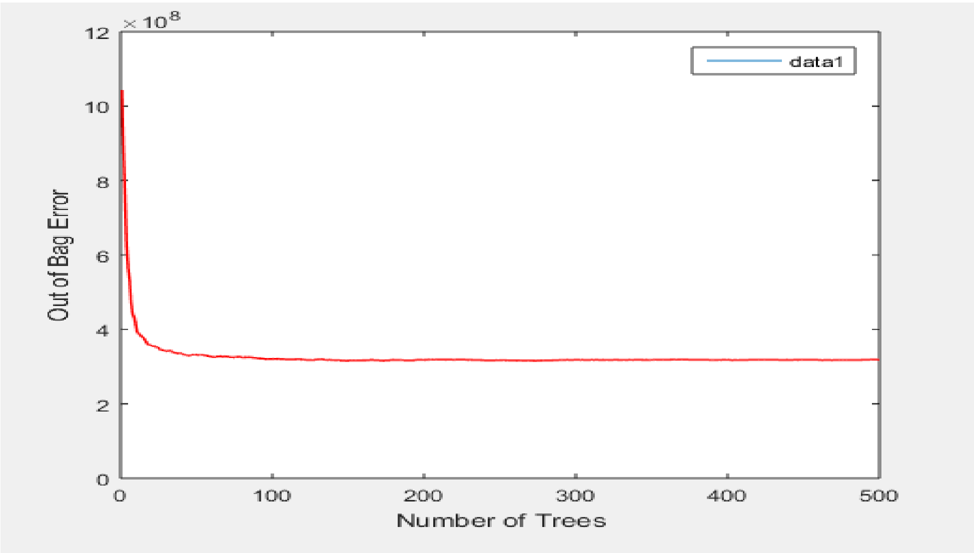



Figure 4.8. Selection of the number of trees in a RF model.

In the same tasks, Oshiro et al. [104] carried out an experimental work on selecting the number of trees used in a bagging method. They declared that while increasing the number of trees in a model, there was no significant change observed in the accuracy assessment results, except the computational expense. In addition, Kertész [105] performed investigations on the number of trees, and declared that if the number of trees increases too much, the model tends to overfit.  Thus, the default settings of the software were adapted using the trial-and-error and OOB error graphic, and the number of trees was set to 100. In the experimental study, while establishing the RF method, the overall dataset was split into two classes, i.e., the training and the validation sets. The CV approach was applied on a set of DT classifiers in various sub-samples of 75% of the whole data set. In this way, the training data was used effectively and increased the prediction accuracy. If the number of leaf sizes was reduced, many sub-samples occurred, which led to an increase in the processing time.

After the RF-based MARE model was created, the prediction was made on the validation data and the results given in Figure 4.9. and in Table 4.17 were obtained. As can be seen in Figure 4.9, the majority of the estimates were close to the best fit line, and the model produced an underestimation only for the high-priced housing predictions.

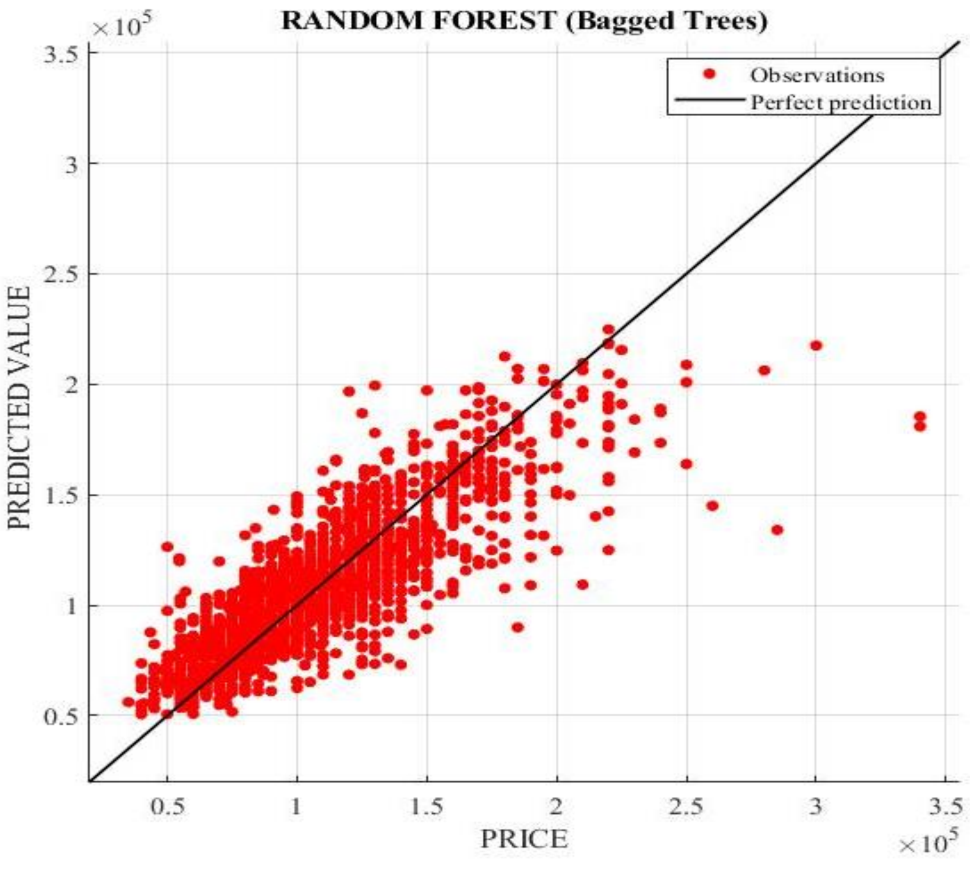



Figure 4.9. The comparison of predicted values and the actual prices of RF-based MARE model.

Table 4.17. The validation results of the RF-based MARE model.

| Method | Ratio Median | Ratio Mean | SD (TRY) | COD | COV | PRD | RMSE (TRY) | $R^2$ | Adj-$R^2$ |
|---|---|---|---|---|---|---|---|---|---|
| RF | 1.0155 | 1.0249 | 0.1255 | 9.5 | 12.2432 | 1.0206 | 14,417 | 0.82 | 0.82 |

According to the regression metrics in Table 4.17, the $R^2$ and Adj-$R^2$ values were similar. It shows that the additional variables were consistent with the explainability rate of the overall model. In addition, when the RMSE values and $R^2$ values are evaluated together, it was observed that the RF method can cope with complex datasets accurately in MARE experimental implementation. Moreover, the RF model produced results following the COD and PRD indicator ranges recommended by IAOO [16] for MARE. The detailed results regarding the appraisal rate are given in Appendix F – Table F.1.

# 5. DISCUSSIONS

This thesis aimed to compare various ML methods, i.e., the RF, ANN, SVM, ANFIS, and MRA, by using a large data set and a broad list of variables for the valuation of mass real estate and contribute to the literature and future studies. The supervised ML methods learn the relationship between the input data given to the model and the target variable with the training module; and produce output based on a series of logical and mathematical functions. In order to correctly evaluate the accuracy results of ML-based mass valuation, it is necessary to assess the valuation process and the input factors together. Often, the reference variable, which has a bias in its nature, could not be the same as actual transactional prices or real market values [18]. On the other hand, the input data cannot encompass all of the influencing independent variables due to their high dimensionality and long processing time. Therefore, these and some other pre-requisites were accepted and omitted to perform a comparative work in MARE.

The dataset used in this study was collected by a GDLRC Pilot Project Team composed of six field experts, three valuation experts, and eight assistant staff. In addition, two international experts also worked as external consultants to analyze the reliability and accuracy of the collected data. The obtained studies and the resulting data sets formed the basis for many different academic studies and theses; thus, the reliability and data quality of the input data used in this study were confirmed. Furthermore, additional checks were carried out within the scope of this thesis.

Here, the model performances were evaluated using several statistical metrics and the conventional MARE measures suggested by IAOO [16]. According to the outcomes, all five models were able to solve both linear and non-linear interactions between the variables. The accuracies obtained from the models show that all methods can be effectively used in MARE studies. However, some methodological approaches have come to the fore based on their characteristics. For example, the RF method has come to the forefront compared to other manners by finding even small interactions between the variables and the target, thanks to the partitioning of the tree feature. However, the RF requires much longer training time than the ANN as it generates a large number of trees and makes decisions on the majority of votes.

The ANN is an algorithm that mimics the decision-making mechanism of the human brain. It produced successful results in this study, as it has a structure that delivers effective results in complex problems with a small amount of training data. When the ANN results are analyzed in the perspective of prediction accuracy, it can be seen that it is close to the RF results and produces better results than the other methods, i.e., the SVM and the ANFIS. In addition, it is the fastest method in terms of training time.

The GRBF-based SVM kernels were used for solving nonlinear problems such as real estate valuation studies in such a higher dimensional input space. According to the prediction accuracy results, the SVM model can also handle the MARE works and provide superior accuracy results than the ANFIS and MRA. However, the training time was much longer than the ANN and RF methods.

The ANFIS, a hybrid method that takes advantage of both the FRBS and the ANN, produced less accurate results when the $R^2$ and RMSE performances are considered and ranked fourth in terms of the SD and COV values. The main reason is that a sub-clustering technique was used for the ANFIS implementation in this study due to the large number of independent variables. As can be seen in this thesis and in [94], the ANFIS has difficulties solving large-scale data due to its rule-based structure. However, it can be stated that it is a more transparent and explainable method compared to the other black box ML methods in terms of determining the range of membership functions and enabling to interfere with the rules during the model-building phase.

Further aspects of the ML methods employed here are discussed in the following sub-sections in detail.

### 5.1. Accuracy and Reliability

A supervised ML-based MARE work consists of two processes such as model building (design) and calibration. The model-building phase consists of three different stages, which are the data collection, the determination of variables, and selection of the ML method. Therefore, all stages should be evaluated for obtaining high accuracy and reliability.

The data collection phase is the first one in model building process that is comparatively more important than the other phases. Many parameters such as the data collection techniques, qualifications of the experts, the distribution, the quality and quantity of the collected data, and the skills and backgrounds of the experts, who perform the data analysis, will affect the data quality and thus the success of the study. On the other hand, if the collected data is not of high quality and reliable, the accuracy of the results is questionable. Furthermore, with an inaccurate dataset, overfitting or other adverse effects may occur.

The variable selection phase is another important stage of the model building process. The effectiveness and significance level of variables greatly affect the accuracy of the predictions and reliability of outcomes.

In the ML-based MARE studies, the method selection is the last step of the model building process. The quality of the collected data and the actual prices greatly influence this step. For example, if the outlier rate in the input data is high, the ANN and the RF models, which are relatively more immune to outliers, will enable producing results that are more accurate. The models created using the RF and the ANN methods provided higher accuracies in the comparative evaluation. Although it is known that the GRBF-based SVM model is successful in non-linear data sets, it could not produce as accurate results as the RF and the ANN here.

On the other hand, the ANFIS method did not yield to accurate results when compared with the other ML methods used in this study. The main reason for this was applying a sub-clustering technique since ANFIS cannot solve high-dimensional data [94]. Based on these results, it could be mentioned that the sub-clustering-based ANFIS could not provide reliable outcomes as the accuracy values differ in various test runs. Several analyses should be performed to determine the quality of available real estate data and assess the extent of data collection or validation required as part of the mass valuation exercise. The accuracy and the reliability are important issues for all organizations, especially for informed decision-making processes like statistics and ML. When obtaining quality data in the mass valuation studies discussed within the scope of this study, there is a need for appraisers who are both competent in the field to be appraised and qualified in the valuation field.

Moreover, the outlier detection tests should be applied for the different data collection techniques at the preliminary inspection stage, and the results should be validated. In this thesis, it was possible to utilize a dataset with a high reliability level, which have already been used in a few pilot applications and scientific studies. After the data collection phase, some statistical tests and expert analyses were applied to the collected data at the pre-processing step, which is the most significant step influencing the accuracy and reliability of the models. The tests include the normality test, data cleaning, data reduction, and lastly (but most significantly) the expert inspection. The tests are shortly described below.

- *Normality Test*: While the P-P Chart is often used to assess the normality, it also compares distributions to determine how well the variables match the target. An example graphic is shown in Figure 5.1. According to the normality assumption test, it can be concluded that the residuals are not distributed normally.

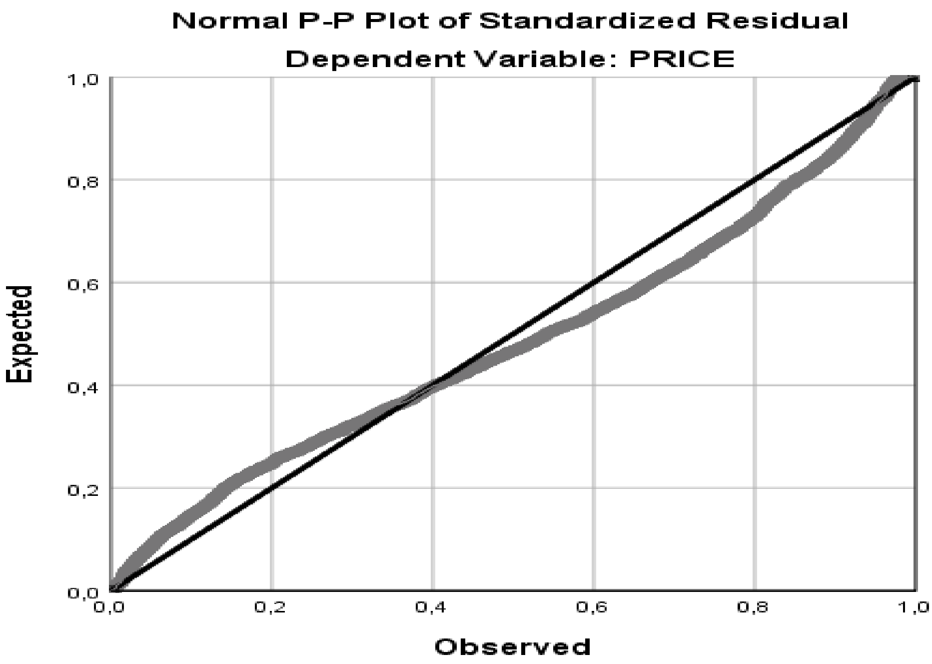



Figure 5.1. The normal P-P plot of regression residuals.

- *Data Cleaning and Reduction*: The data as much as the number of statistical samples calculated in the area to be evaluated is required. In this study, the SVM algorithm needed a long processing time working with 3000 samples. At the same time, the ANFIS method required both a long processing time with the same number of samples and could not produce results when the number of variables are greater than 7. Thus, data completion or deletion operations were performed regarding the missing data. Then the outliers were determined with the percentage analysis. However, some of the outliers were not omitted from the dataset to boost generalization ability after all preliminary data interactions. The PCA analysis was applied to the initial data to reduce the number of the variables. After this step, it was possible to obtain results with the RF, the ANN, and the SVM. However, the ANFIS gets in the curse of dimensionality. In terms of data reduction, the PCA was used to consolidate zonal (distance) variables as shown in Table 4.2. According to the PCA outcomes, the zonal (distance) variables and the D-STTS were consolidated into one component. In the second step of the data reduction, changes in $R^2$ values were inspected while adding new variables. Some of the variables were omitted from the model.

- *Accuracy Inspection*: The systematically described tests suggest to check data quality before applying conventional or advanced ML-based prediction models, but the appraiser does not always use these techniques systematically. Depending on the results of the validation data and the model suitability, additional tests can be performed, or some tests can be ignored. In addition to these, in the context of the study, some MARE tests determined by IAOO were also applied, and the data's suitability, quality, and compatibility were tested.

The supervised ML methods learn the relationship between the input data variables given to the system and the target by training module and produce output data by using logical and mathematical functions. Regression metrics can be used to measure the error in the input-output process. If the initial MARE results are far from the price, the dataset are tested with additional techniques, and if the results indicate inaccurate outcomes, they are omitted from the model.

In this study, in addition to the error metrics used as the quality indicators of the model in statistics and ML models, the ratio indicators were also used to evaluate the quality of the REA. The comparison summary of the statistical assessment and mass appraisal indicators used to measure ML-based MARE validation results are provided in Table 5.1.

Table 5.1. Comparison Results of the Models

| Method | Ratio Median | Ratio Mean | SD | COD | PRD | RMSE (TRY) | $R^2$ | Adj-$R^2$ | COV | Training Time (Sec) | Overall Prediction Accuracy |
|---|---|---|---|---|---|---|---|---|---|---|---|
| RF | 1,015 | 1,025 | 12,600 | 9.5 | 1.021 | 14,417 | 0.82 | 0.82 | 12.24 | 3.34 | 90.29% |
| ANN | 1,007 | 1,019 | 13,100 | 9.947 | 1.023 | 15,279 | 0.81 | 0.8 | 12.86 | < 1 | 89.97% |
| SVM | 1.001 | 1.013 | 14,300 | 10.867 | 1.024 | 16,916 | 0.77 | 0.76 | 14.16 | 7.25 | 89.12% |
| MRA | 1.012 | 1.029 | 16,100 | 12.28 | 1.024 | 17,484 | 0.74 | 0.72 | 15.19 | 42.32 | 87.55% |
| ANFIS | 1.013 | 1.028 | 15,600 | 12.226 | 1.023 | 18,229 | 0.72 | 0.72 | 15.15 | 46.5 | 87.58% |

The $R^2$ results show that the RF model explaining rate and the average deviation is better than the other four methods. The RF and the ANN methods achieved comparable $R^2$ and RMSE values. The comparison of the ANN and RF model results to the target given in Figure 5.2 showed that these two methods yielded close to the each other. In addition, based on the $R^2$ and RMSE values, the SVM was ranked in the third place after the RF and the ANN.

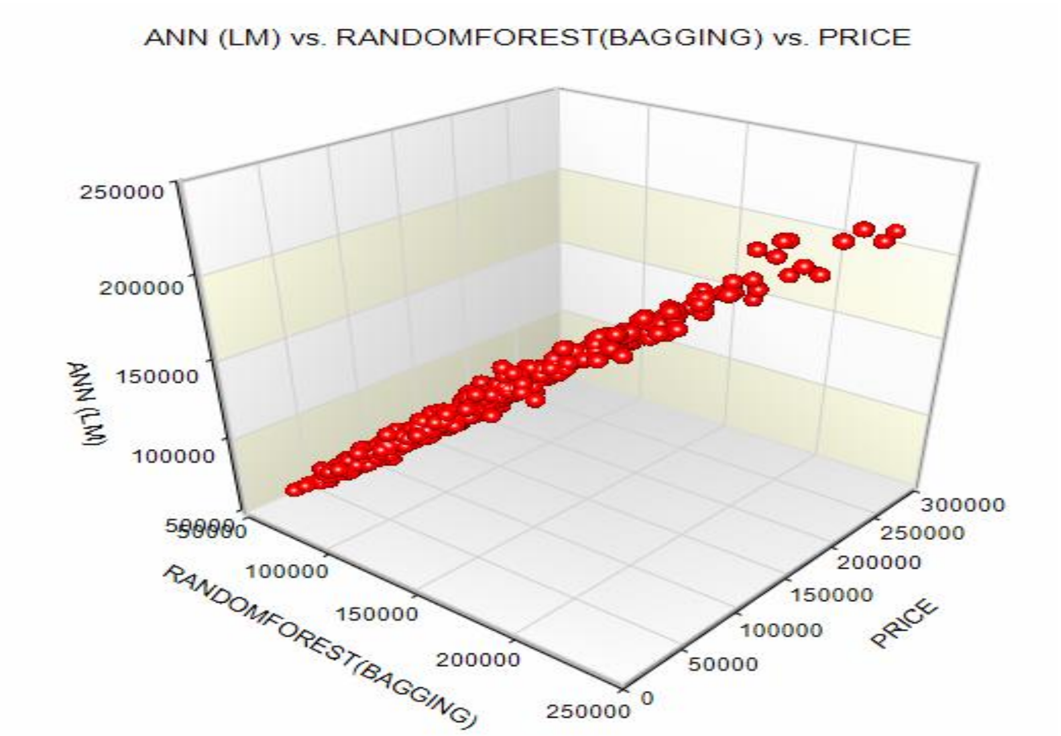



Figure 5.2. The comparison of the ANN and RF model results with respect to the price variable.

A general measure in ratio studies is the appraised value ratio to the sale price (A/S ratio). This simple calculation produces a variety of indicators to describe the set of ratios under valuation, including the median, mean, weighted mean, COD, and PRD among many others. Confidence intervals were provided for the COD and PRD. In this study, the second group of performance measures, the conventional metrics of REA can be listed. Among those, the COD, PRD and COV can be used especially for mass appraisal processes (IAAO, 2013) [16]. The COD is a widely used as a measure of appraisal uniformity. It is expressed as a percentage of the average deviation of the ratios from the median. A COD value should be in a range between 5 to 15. If a COD value is higher than 15, it may indicate outliers or non-representative samples. If lower than 5, the model could not produce satisfactory results. The PRD value should be between 0.98 and 1.03 to indicate vertical equity. The PRD was used to evaluate of relation between prediction levels and actual price levels; and thus it indicates vertical equity. If the ratio is lower than 1.00, it is suggested that the prediction level is progressive and therefore real estates with lower prices were under-appraised relative to high value properties. In the opposite condition, the prediction level is called regressive and high-value real estates were under-appraised relative to the low-value ones. The COV is a measure of relative variability and the standard deviation (SD) ratio to the mean. The higher values have more variances from the mean. The SD is the average distance of the ratios from the ratio mean.

The results obtained from all methods were evaluated based on the performance measures and the ratio mean, ratio median, SD, COD, COV, and PRD results are also presented in Table 5.1. Based on the results, the PRD values obtained from all methods were comparable. On the other hand, the COV, COD, and SD metrics obtained from the RF were superior to all other methods, demonstrating that the RF outperforms the ANN, the SVM, the MRA and the ANFIS. However, all mass appraisal indicators presented here were within meaningful ranges according to the IAAO (2013) standards [16]. Therefore, it can be said that all of the methods can be used in mass appraisal studies efficiently.

Although the ANFIS method uses the neural network infrastructure, it has not been as successful as the other non-linear methods in explaining the model. The most important reason for this is using the sub-clustering method to reduce the data size in the ANFIS model, which remained unsolved due to a large number of variables. As a result, it shows that the ANFIS method is weaker in terms of reliability when working with large data.

As a further analysis, the results of the COV values on the basis of neighborhoods are presented for the ANFIS, MRA, SVM, ANN and RF methods in Figures 5.3, 5.4, 5.5, 5.6, 5.7 respectively. As can be seen in the Figures, the COV results obtained in some of the neighbourhoods were high in all maps, while some others remained to be low, which could be caused by regressive appraisals. According to the PRD ratio, the overall neighbourhood results are greater than 1.00, so predictions could be interpreted as regressive and therefore high-valued real estates were under-appraised relative to the low-valued ones. This may source from regional factors that affected the real estate prices, which should be analyzed further in different sites.

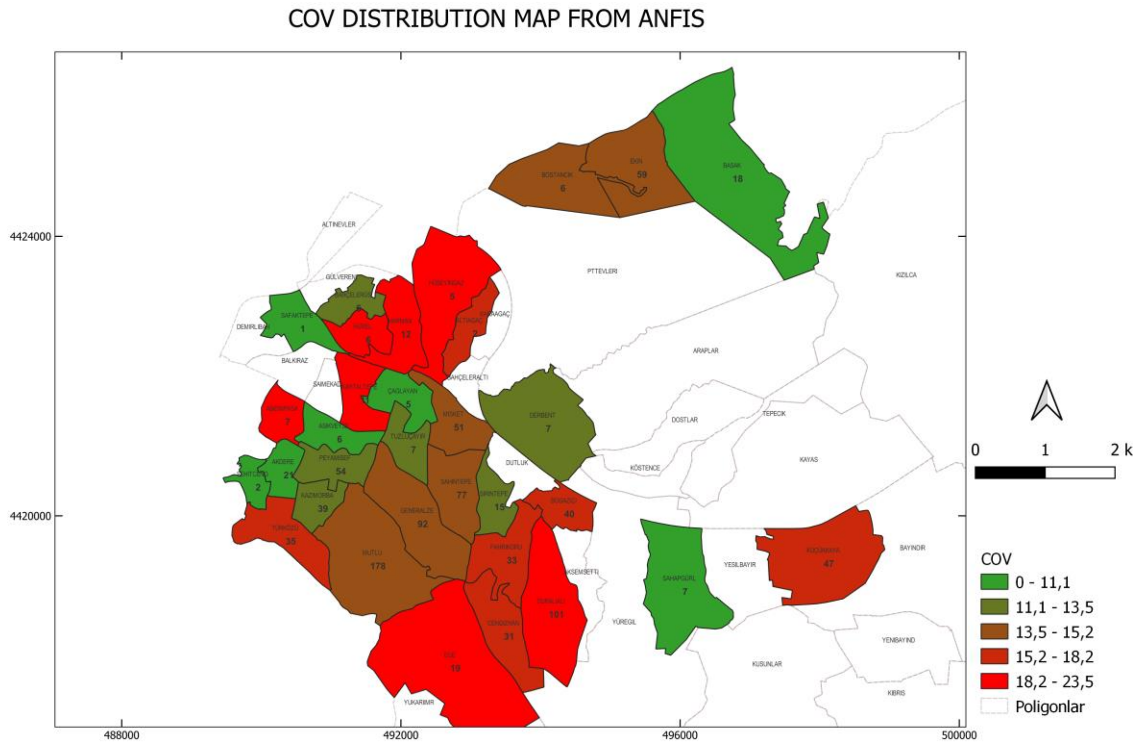



Figure 5.3. The COV distrubution map obtained from the ANFIS.

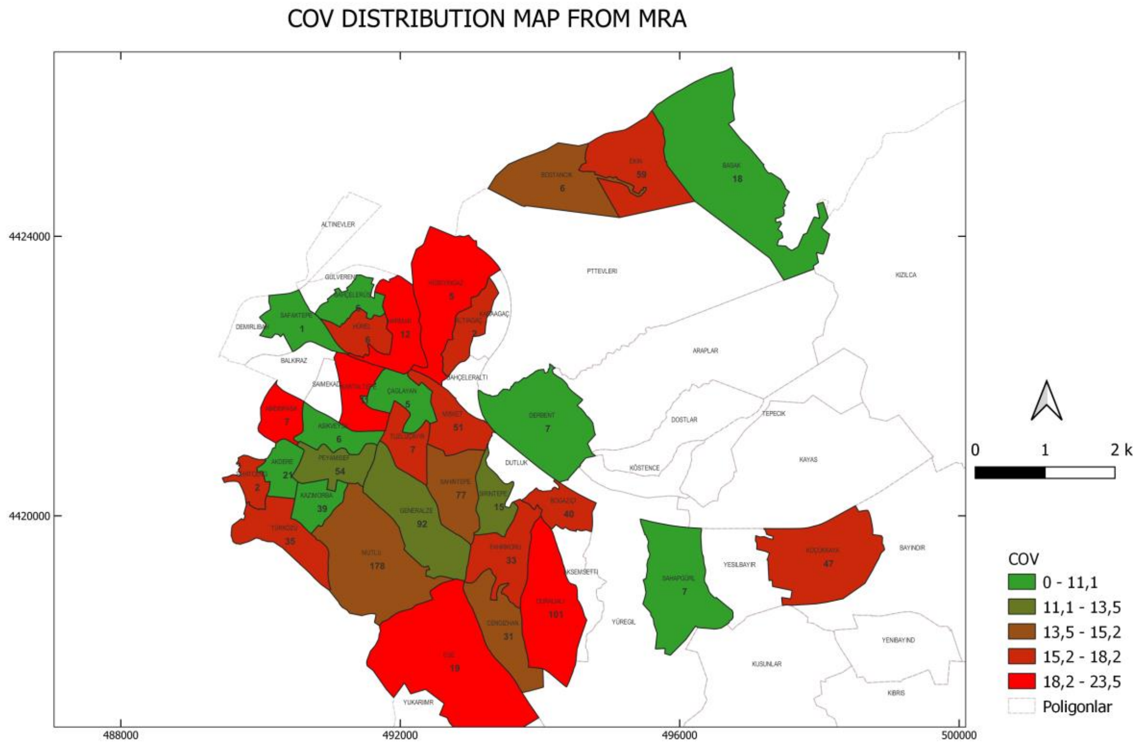



Figure 5.4. The COV distrubution map obtained from the MRA.

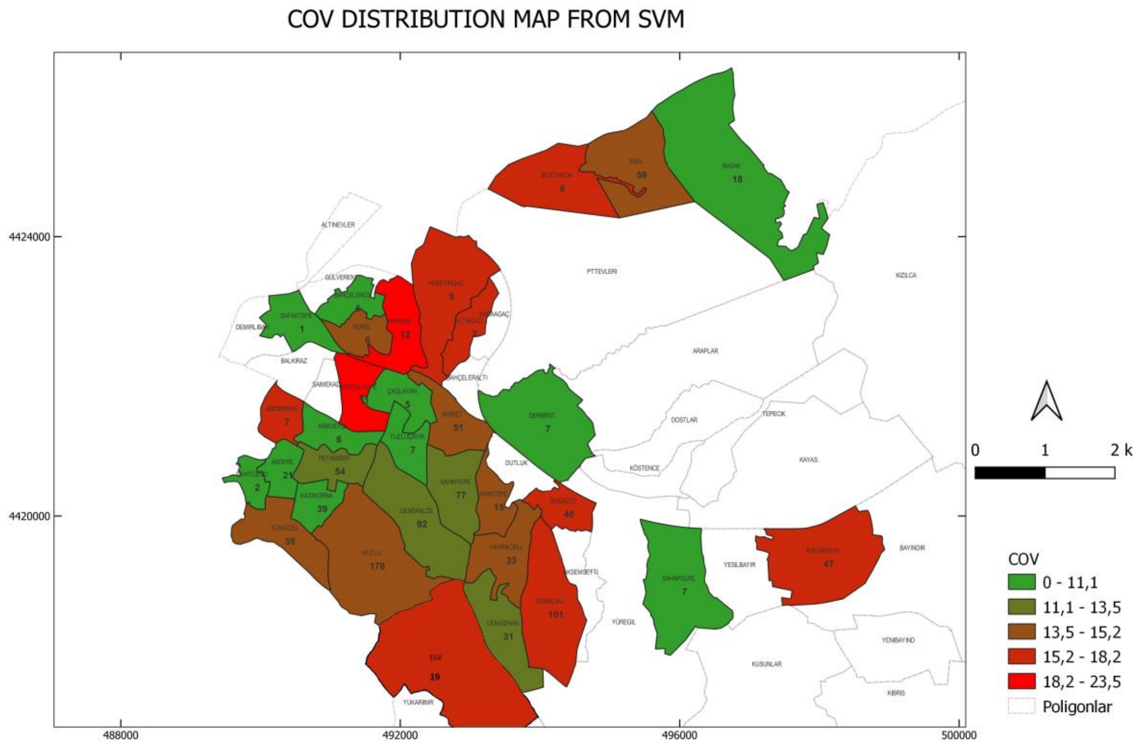



Figure 5.5. The COV distrubution map obtained from the SVM.

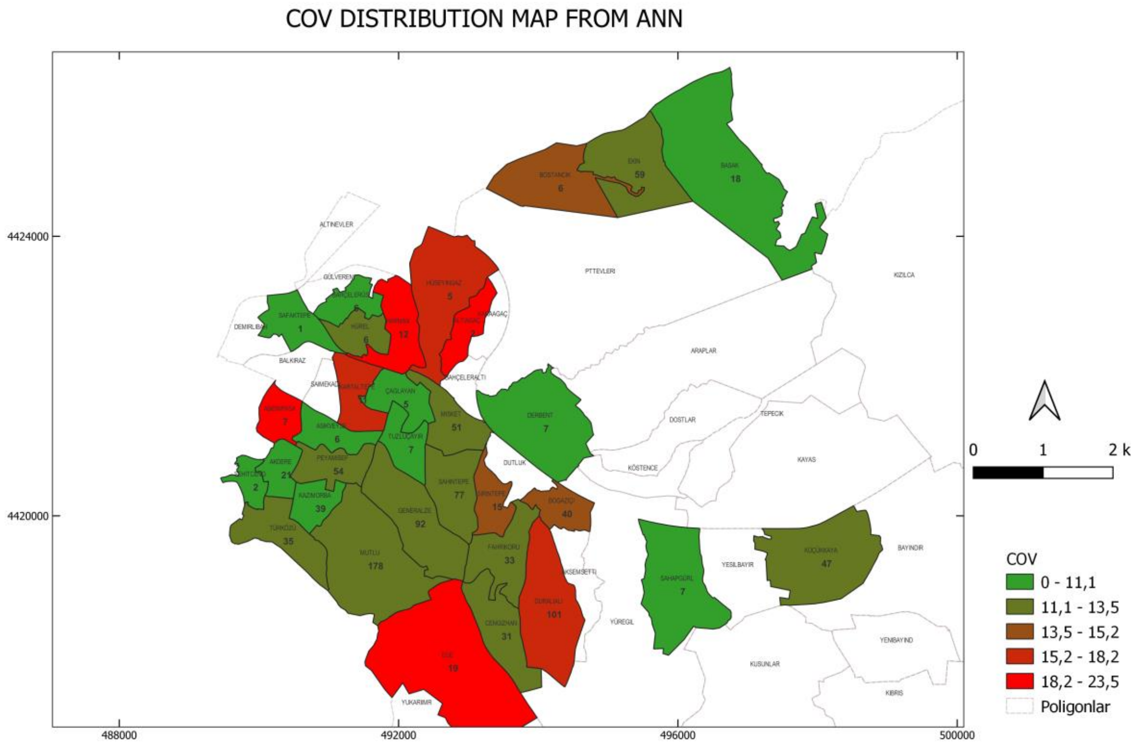



Figure 5.6. The COV distrubution map obtained from the ANN.

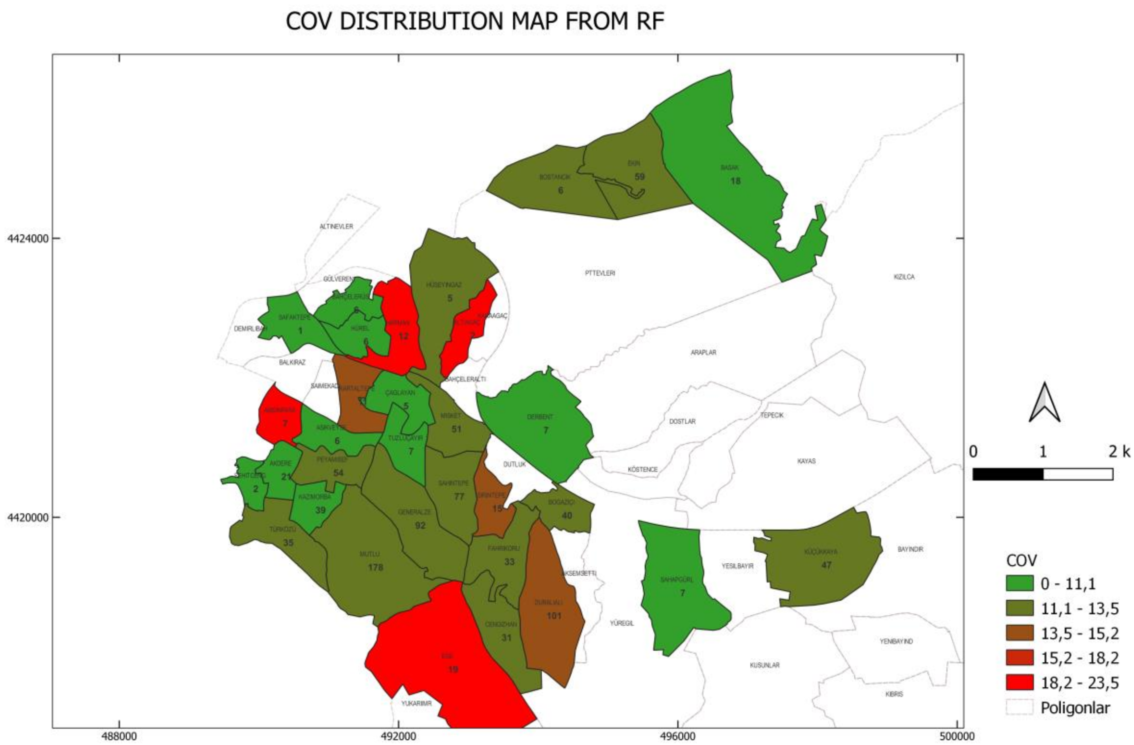



Figure 5.7. The COV distrubution map obtained from the RF.

The COV is a relative SD and widely used for accuracy and quality assessment in financial models. As can be seen in the Figures given in this Section, the RF and ANN provided better COV distributions over the neighborhoods in comparison to the others. The MRA method produced the least accurate results when compared to other methods. It is estimated that this result is caused by the linear method's inability to solve non-linear interactions and the stepwise method's structure that produces a weak model against new data.

### 5.2. Interpretability

To determine the transparency of a ML model, an expert must analyze the configuration and the outputs. In terms of interpretability, the decision-making methods can be classified into three groups as white box, grey box, and black-box. In the literature, the white box models are called as interpretable [106, 107].

Since the supervised ML methods used in this thesis are black-box models except the MRA, thus they are considered as complex in terms of interpretability and transparency. Interpretability is a difficult term to explain. It is challenging to determine the interpretability of events or situations that cannot be related to cause and effect, and contains linguistic expressions. Transparency and interpretation are also frequently studied topics in MARE studies. A wide range of mathematical operations is required when interpreting the supervised ML methods in the MARE context [108]. It indicates how close predicted results produced by the algorithm, which is only possible if there is information about the model operations to the existing inputs to obtain these results.

In the conventional sales comparison appraisal technique, due to the expertise to be made about real estate, the valuation made by two different experts will be different from each other. The critical issue here is that all estimates at a threshold value will be acceptable to the stakeholders, as both experts will explain and interpret their results. However, if experts do not explain the reasons for their prediction and how they came to the conclusion, the results would undoubtedly be controversial and unacceptable. As can be seen, the interpretation issue may be even more important than the results obtained in some cases.

The ANN, SVM, RF, and ANFIS supervised methods produce results from the input data by learning algorithms using many functions with an indefinite number of mathematical or statistical operations. In this context, although partial interpretation can be performed with some analyzes of the results obtained in such black-box methods, such as in the RF, variable importance analysis help to make model results more interpretable. On the other hand, the stepwise MRA is a white box model and can be easily interpreted. The stepwise regression adds or removes variables from the model, respectively. This process is carried out step by step, so the significance of the variables in the model can be clearly determined. In this way, the model shows an explainable and transparent model feature.

Nowadays, the ML models have increasingly been used in businesses and transactions related to the analytical decision-making processes. The non-transparent characteristic of the ML methods is often criticized in the application areas which require analytical or deterministic approaches. Especially in REA, applications such as real estate tax calculation and expropriation require transparency. In a methodological description, the RF, SVM, ANN and ANFIS are black-box methods, and the models and outcomes of these methods could not be defined as transparent. On the other hand, the ANFIS method is also a ML method, but some of the model establishing phases, especially rule generation phase, could be determined and explained with some detailed analysis.

The GRBF-based SVM method carries data from one-dimensional space to high-dimensional space by using the kernel function, thus revealing non-linear interactions between variables. Therefore, the GRBF-based SVM method is a black-box method that is difficult to interpret similar to ANN. In addition, the RF is also a black-box method for data scientists. There is no information or control on what the model does while randomly partitioning the trees. However, the RF is computationally less expensive than the ANN. On the other hand, the interpretability of the RF model is better than the ANN. The RF uses simple decision trees and random selection of predictor trees. If an expert solve the randomness or make an analysis on variable importances, the model could be more interpretable. However, in ANN, the model uses hidden layers and nodes. Therefore, explainability of the entire network functions is very difficult. In this manner, the model interpretability would be more difficult than the other ML models. Finally, stepwise-based MRA model is an easily interpretable and transparent model for appraisers. The results of the model could be handled easily by experts and appraisers.

### 5.3. Generalization Capability

Generalizability in ML is the applicability of an established model from a sample to an entire population or a new population. In other words, it is the ability of the model to handle unseen patterns or datasets. The main issue in ML-based mass appraisal is how it could be adapted to new datasets and how well the model fits the new study area. Generalization is of great importance for the ML methods. One of the main goals in supervised ML methods for MARE is to train the model with the available data and save time and cost by using this model in more extensive or different study areas. In a ML-based MARE study, to ensure model capability to generalize, one should be very careful in selecting variables, and a comprehensive and meaningful set of variables should be used as much as possible. Using an extensive and high quality dataset while creating the model, the prediction accuracy of the model and inclusiveness rate will also increase. In addition, it will show faster and higher performance in new application areas. On the other hand, the transferability of ML models is a related but broader concept than generalizability. Transferability in ML can be defined as acquiring knowledge about a problem and the model itself to solve a new but related issue. For example, in a ML-based mass appraisal model, when zonal parameters are known to cause the curse of multicollinearity, these parameters will not be used in a ML model for identifying new real estate investments. The model with the first experience here will greatly contribute to the creation of subsequent related models.

When the Figures 5.3 - 5.7 in Section 5.1 are evaluated, it can be seen that the RF model achieved more successful COV values in all neighborhoods than the other models. After the RF model, the ANN method produced the most accurate results. Essentially all models produced COV values consistent with statistical and mass appraisal indicators. Therefore, the MRA method produced less reliable outcomes according to the COV and generalization context than others. In addition, the SVM and ANFIS methods produced reliable COV results but they should be improved according to the COV distribution maps. Moreover, despite the ANN-based MARE model provides accurate COV result, it also has some weaknesses about generalization in training period. For instance, ANN generalization performance mostly depends on how and with which data the network is trained. Successful generalization results will be achieved if the training data are created with sufficiently effective and inclusive variables. In this study, the ANN performed well but it may not be produce accurate results when applied to geographically different regions.

# 6. CONCLUSIONS AND FUTURE WORK

This chapter is covered under two parts. In the first part conclusions, the comments about the results obtained related to the whole thesis and the limitations of the study are given. In the second part, suggestions for future studies are presented.

## 6.1. Conclusions

The REA is a complex and challenging process. The target variable (price) is affected by a large number of independent variables. Identifying all the predictors, especially in aggregate valuation, is a time-consuming and labour-intensive processes. Within the scope of this thesis, besides being a difficult working process, some constraints created a obstacles throughout the study. In particular, the fact that the available data is old (from 2012 and 2013) has made it almost impossible to compare with the current real estate prices (as of year 2022), since there was no reliable translater index in Turkey to rectify the real estate prices.

On the other hand, in an article prepared within the scope of this thesis published last year, a data set in which all 38 variables were used in an unprecedented wide range of data was used. Its impressive results were evaluated within the scope of this study. With such a large data set used in the study of Park and Bae, approximately 5000 samples were appraised with 28 variables in Fairfax County, Virginia, USA [109]. In this respect, the studies carried out within the scope of the thesis and, like other wide range of practical case studies, will contribute to the future works to be carried out from now on.

In this study, five ML methods, such as the RF, ANFIS, SVM, ANN and MRA based MARE models, were established in newly developed and unstructured areas. A large part of the Mamak district, which has different areas and includes developed areas and shantytowns, was used for these studies as a test field. Within the scope of the thesis, studies have also been carried out with complete and extensive datasets as published previously. Initially, a comprehensive data set consisting of 36 variables and 4100 samples were employed [1]. The REA studies have not encountered a data variable set with such a wide variety of variables before. Thanks to the width of the data set, the reactions of ML methods compared to the data can be observed more easily [109]. Based on the outcomes of the tests performed here, it can be concluded that that the five ML methods could be efficiently utilized in MARE studies instead of the traditional methods. It would be appropriate to decide which of these methods should be preferred according to the specific features of the methods. As for the inferences from the results obtained, the RF method achieved the highest performance in terms of accuracy and generalization capability. It is also less affected by outliers and missing values. It can also produce accurate results when working with small datasets and is faster than the ANN,SVM, MRA and ANFIS methods.

In addition, the ANN results were very close to the RF method. It is also affected less by outliers and missing values, and obtained more accurate results with small amount of data than the SVM, MRA and ANFIS methods. However, when working with small datasets, the tendency to overfit is greater than the RF method. Although the SVM method is not as successful as RF and ANN in terms of overall accuracy, it has proven to be a method that can be used in REA studies. For the method to produce more successful results, the selection of kernel function is also important. Therefore, it is necessary to test with different datasets and in different study areas as future work.

The ANFIS method exhibited poorer prediction performance and less reliability because it could not produce solutions with large variable data sets, which is the biggest disadvantage, namely the curse of dimensionality. Although it is essentially a rule-based method and can work with linguistic and numerical data, which are of great importance for REA studies, the necessity of using it with the sub-clustering method has adversely affected the results. Although the MRA method achieved better results than the ANFIS according to the mass appraisal indicators, it produced unsuccessful results compared to the other three methods due to its failure in non-linear interactions, and its generalization capability was insufficient.

As a result, it was evaluated that this study, in which five methods are systematically compared and evaluated in a wide application area with a very large data set, will contribute to the literature and shed light on future studies. In addition to all these, in a broad perspective, in countries like Turkey where REA studies have not yet been institutionalized, the ML-based MARE methods will contribute to the creation of large-scale value maps, and indirectly;

- It will be possible to prevent the state's tax losses based on real estate.
- Thanks to a fair tax system, the trust between the citizens and the state will be strengthened.
- Confidence and stability are ensured and the credit system works effectively in financial sectors.
- It is possible to obtain accurate and sufficient guarantees for systemic risks in all banking sectors and financial markets.
- Purchase, loan (mortgage), expropriation, insurance, land consolidation, reform plans, determination of property income and rent, losses and unjust occupation, inheritance division, sale of public properties, privatization and nationalization type of transactions can be carried out easier with MARE.

## 6.2. Future Work

The MARE studies are in general carried out in many countries for real estate taxation. In this sense, the valuation must be transparent and reliable. When considering a regional basis, valuation studies have not yet been carried out in Turkey, apart from a few pilot applications and the housing price index studies carried out by the Turkish Central Bank. Therefore, both property and land taxes are often based on unrealistic values. The most significant effect of this is seen as lost taxes. The ML methods can be effectively used to create a value index map that will be made to form a basis or reference for all real estate value-based institutional works in Turkey.

In conclusion, the ML methods obtained in this study are promising. The results of this thesis, in which the five ML methods were systematically compared and evaluated in MARE studies in a wide range of applications with an extensive data set. The results of the models are very accurate in terms of prediction accuracy. Moreover, due to the data quality and reliability, the results are attractive. They were faster and less costly than the conventional methods in terms of processing time and cost factor. However, except the MRA, the four non-linear ML methods cannot be considered as transparent and interpretable. With the integration of expert methods into ML methods, studies can be more transparent and interpretable and may produce more accurate results.

On the other hand, some investors need a real-time valuation to work in unpredicted situations like economical crises or pandemic effects. In addition, some real estate agencies and property management companies opened their doors in stock markets. Therefore, in stock exchange transactions were needed to fast and accurate valuations and cash turn calculations about the real estate. Therefore, future studies need to investigate allowing expert intervention and online real-time applications to provide different perspectives and conveniences in this MARE area.

# 7. REFERENCES

[1] S. Yilmazer, S. Kocaman, A mass appraisal assessment study using machine learning based on multiple regression and random forest, Land Use Policy, 99 (2020) 104889.

[2] D. Wang, V.J. Li, Mass appraisal models of real estate in the 21st century: A systematic literature review, Sustainability, 11 (2019) 7006.

[3] A. Sangha, Property valuation by machine learning for the Norwegian real estate market, ScienceOpen, 2021.

[4] B. Alarie, A. Niblett, A.H. Yoon, Using machine learning to predict outcomes in tax law, Can. Bus. LJ, 58 (2016) 231.

[5] S.J. Johnson, M.R. Murty, Detecting High Risk Property Taxpayers Using a New Business Intelligence Model: A Case of New York City Property Tax.

[6] M.A.P.P. Team, Land Registry and Cadastre Modernization Sub Project Mass Appraisal Case Study, GDLRC Ankara, 2014.

[7] I. Williamson, S. Enemark, J. Wallace, A. Rajabifard, Land administration for sustainable development, Citeseer2010.

[8] W. McCluskey, W. Deddis, A. Mannis, D. McBurney, R. Borst, Interactive application of computer assisted mass appraisal and geographic information systems, Journal of Property Valuation and Investment, (1997).

[9] R.I.o.C.S. (RICS), November 2021, RICS VALUATION – GLOBAL STANDARDS, Royal Institution of Chartered Surveyors (RICS), www.rics.org, 2021.

[10] S. Yilmazer, S. Kocaman, On the Prediction Performances of SVM and ANN Methods for Mass Appraisal Assessment: A Case Study from Ankara (Turkey) FIG e-Working Week 2021 - Smart Surveyors for Land and Water Management - Challenges in a New Reality FIG, Virtually in the Netherlands, 2021.

[11] W.M. Shenkel, Computer Assisted Assessments: Potentialities and Implications for the Organizational Structure of Property Tax Administration, Proceedings of the Annual Conference on Taxation under the Auspices of the National Tax Association, JSTOR, 1970, pp. 66-86.

[12] G.P. Rollo, Valuation theory and real property assessment, University of British Columbia, 1971.

[13] P. Rossini, Artificial neural networks versus multiple regression in the valuation of residential property, Australian Land Economics Review, 3 (1997) 1-12.

[14] S. McGreal, A. Adair, D. McBurney, D. Patterson, Neural networks: the prediction of residential values, Journal of Property Valuation and Investment, (1998).

[15] D. Demetriou, A spatially based artificial neural network mass valuation model for land consolidation, Environment and Planning B: Planning and Design, 44 (2016).

[16] (IAAO), Standard on Ratio Studies, International Association of Assessing Officers (IAAO), Kansas City, Missouri, USA., 2013.

[17] (IVSC), A submission to the International Accounting Standards Board in respect to an issues paper issued for comment by the IASC Steering Committee on Extractive Industries, International Valuation Standards Committee, 2001.

[18] R.M. Kathmann, Neural networks for the mass appraisal of real estate, Computers, Environment and Urban Systems, 17 (1993) 373-384.

[19] W. McCluskey, W. Deddis, A. Mannis, D. McBurney, R. Borst, Interactive application of computer assisted mass appraisal and geographic information systems, Journal of Property Valuation Investment, 15 (1997) 448-465.

[20] G. Büyük, F.B. Ünel, Comparison of Modern Methods Using the Python Programming Language in Mass Housing Valuation, Advanced Land Management, 1 (2021) 21-26.

[21] R. Fraser, F. Blackwell, Comparables selection and multiple regression in estimating real estate value: an empirical study, Journal of Valuation, (1989).

[22] J.P. Sestric, MASS APPRAISING AND ASSESSING OF METROPOLITAN CITY PROPERTIES, Proceedings of the Annual Conference on Taxation under the Auspices of the National Tax Association, JSTOR, 1953, pp. 190-200.

[23] F. Bergadano, R. Bertilone, D. Paolotti, G. Ruffo, Learning real estate automated valuation models from heterogeneous data sources, arXiv preprint arXiv:1909.00704, (2019).

[24] T. Dimopoulos, H. Tyralis, N.P. Bakas, D. Hadjimitsis, Accuracy measurement of Random Forests and Linear Regression for mass appraisal models that estimate the prices of residential apartments in Nicosia, Cyprus, Advances in Geosciences, 45 (2018) 377-382.

[25] W.J. McCluskey, M. McCord, P. Davis, M. Haran, D.J.J.o.P.R. McIlhatton, Prediction accuracy in mass appraisal: a comparison of modern approaches, 30 (2013) 239-265.

[26] S.S. Bilgilioğlu, H.M. Yılmaz, Comparison of different machine learning models for mass appraisal of real estate, Survey Review, (2021) 1-12.

[27] T. Güneş, Ü. Yildiz, Property valuation and taxation for improving local governance in Turkey, Land Tenure Journal, 15 (2016) 141-160.

[28] D. Demetriou, The assessment of land valuation in land consolidation schemes: The need for a new land valuation framework, Land Use Policy, 54 (2016) 487-498.

[29] M. Robbins, Process for automated real estate valuation, Google Patents, 2001.

[30] J.M. Clapp, A semiparametric method for valuing residential locations: application to automated valuation, The Journal of Real Estate Finance Economics, 27 (2003) 303-320.

[31] J. Zurada, A. Levitan, J. Guan, A comparison of regression and artificial intelligence methods in a mass appraisal context, Journal of real estate research, 33 (2011) 349-388.

[32] E. Alpaydin, Machine learning, MIT Press2021.

[33] M. Schuld, I. Sinayskiy, F. Petruccione, An introduction to quantum machine learning, Contemporary Physics, 56 (2015) 172-185.

[34] I. Kononenko, Machine learning for medical diagnosis: history, state of the art and perspective, Artificial Intelligence in medicine, 23 (2001) 89-109.

[35] S.B. Kotsiantis, I.D. Zaharakis, P.E. Pintelas, Machine learning: a review of classification and combining techniques, Artificial Intelligence Review, 26 (2006) 159-190.

[36] D. Maulud, A.M. Abdulazeez, A review on linear regression comprehensive in machine learning, Journal of Applied Science and Technology Trends, 1 (2020) 140-147.

[37] C. Mair, G. Kadoda, M. Lefley, K. Phalp, C. Schofield, M. Shepperd, S. Webster, An investigation of machine learning based prediction systems, Journal of systems and software, 53 (2000) 23-29.

[38] Ö.F. Ertugrul, Y. Kaya, Smart city planning by estimating energy efficiency of buildings by extreme learning machine, 2016 4th International Istanbul Smart Grid Congress and Fair (ICSG), IEEE, 2016, pp. 1-5.

[39] P. Rossini, Application of artificial neural networks to the valuation of residential property, Third Annual Pacific-Rim Real Estate Society Conference. Palmerston North, New Zealand, Citeseer, 1997.

[40] R. Abidoye, J. Ma, C.L. Lee, Barriers, drivers and prospects of the adoption of artificial intelligence property valuation methods in practice, Pacific Rim Property Research Journal, (2022) 1-18.

[41] J. Torres-Pruñonosa, P. García-Estévez, C. Prado-Román, Artificial Neural Network, Quantile and Semi-Log Regression Modelling of Mass Appraisal in Housing, Mathematics, 9 (2021) 783.

[42] S. Sisman, A.U. Akar, S. Yalpir, The novelty hybrid model development proposal for mass appraisal of real estates in sustainable land management, Survey Review, (2021) 1-20.

[43] P.-F. Pai, W.-C. Wang, Using machine learning models and actual transaction data for predicting real estate prices, Applied Sciences, 10 (2020) 5832.

[44] C. Rudin, Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead, Nature Machine Intelligence, 1 (2019) 206-215.

[45] M. Strobel, Aspects of transparency in machine learning, Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems, 2019, pp. 2449-2451.

[46] X.L. Cai, D.J. Xie, K.H. Madsen, Y.M. Wang, S.A. Bögemann, E.F. Cheung, A. Møller, R.C. Chan, Generalizability of machine learning for classification of

schizophrenia based on resting-state functional MRI data, Human brain mapping, 41 (2020) 172-184.

[47] S. Rose, Machine learning for prediction in electronic health data, JAMA network open, 1 (2018) e181404-e181404.

[48] T.D. Azad, J. Ehresman, A.K. Ahmed, V.E. Staartjes, D. Lubelski, M.N. Stienen, A. Veeravagu, J.K. Ratliff, Fostering reproducibility and generalizability in machine learning for clinical prediction modeling in spine surgery, The Spine Journal, 21 (2021) 1610-1616.

[49] J. Guan, J. Zurada, A. Levitan, An adaptive neuro-fuzzy inference system based approach to real estate property assessment, Journal of Real Estate Research, 30 (2008) 395-422.

[50] P. Paatero, U. Tapper, Analysis of different modes of factor analysis as least squares fit problems, Chemometrics and Intelligent Laboratory Systems, 18 (1993) 183-194.

[51] Y. Takane, F.W. Young, J. de Leeuw, Nonmetric common factor analysis: An alternating least squares method with optimal scaling features, Behaviormetrika, 6 (1979) 45-56.

[52] J.C. De Winter, D. Dodou, Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size, Journal of applied statistics, 39 (2012) 695-710.

[53] H. Shi, Determination of real estate price based on principal component analysis and artificial neural networks, 2009 Second International Conference on Intelligent Computation Technology and Automation, IEEE, 2009, pp. 314-317.

[54] S. Chopra, R. Mitra, V. Kumar, Fuzzy controller: Choosing an appropriate and smallest rule set, International Journal of Computational Cognition, 3 (2005) 73-78.

[55] B. Ozer, B. Mutlu, H. Nefeslioglu, E. Sezer, M. Rouai, A. Dekayir, C. Gokceoglu, On the use of hierarchical fuzzy inference systems (HFIS) in expert-based landslide susceptibility mapping: the central part of the Rif Mountains (Morocco), Bulletin of Engineering Geology and the Environment, 79 (2020) 551-568.

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

[57] (GDLRC), Pilot Mass Appraisal Report of Mamak, General Directorate of Land Registry and Cadastre, Ankara, 2014.

[58] K. Sultan, H. Ali, Z. Zhang, Big data perspective and challenges in next generation networks, Future Internet, 10 (2018) 56.

[59] R. Sengupta, D. Sengupta, A. Kamra, D. Pandey, JOURNAL OF CRITICAL REVIEWS ARTIFICIAL INTELLIGENCE AND QUANTUM COMPUTING FOR A SMARTER WIRELESS NETWORK, 2020.

[60] T.O. Ayodele, Types of machine learning algorithms, New advances in machine learning, 3 (2010) 19-48.

[61] B. Mahesh, Machine Learning Algorithms-A Review, International Journal of Science and Research (IJSR).[Internet], 9 (2020) 381-386.

[62] K. Sundararajan, L. Garg, K. Srinivasan, A. Bashir, J. Kaliappan, G. Ganapathy, S.K. Selvaraj, M. Thiruvadi, A Contemporary Review on Drought Modeling Using Machine Learning Approaches, Computer Modeling in Engineering and Sciences, 128 (2021) 447-487.

[63] MATLAB and Statistics Toolbox The MathWorks, Inc., Natick, Massachusetts, United States, 2020.

[64] C. Nyamekye, S. Kwofie, B. Ghansah, E. Agyapong, L.A. Boamah, Assessing urban growth in Ghana using machine learning and intensity analysis: A case study of the New Juaben Municipality, Land Use Policy, 99 (2020) 105057.

[65] M. Čeh, M. Kilibarda, A. Lisec, B. Bajat, Estimating the performance of random forest versus multiple regression for predicting prices of the apartments, ISPRS international journal of geo-information, 7 (2018) 168.

[66] J. Hong, H. Choi, W.-s. Kim, A house price valuation based on the random forest approach: the mass appraisal of residential property in south korea, International Journal of Strategic Property Management, 24 (2020) 140-152.

[67] P. Kumkar, I. Madan, A. Kale, O. Khanvilkar, A. Khan, Comparison of Ensemble Methods for Real Estate Appraisal, 2018 3rd International Conference on Inventive Computation Technologies (ICICT), IEEE, 2018, pp. 297-300.

[68] T.F. Sing, J.J. Yang, S.M. Yu, Boosted Tree Ensembles for Artificial Intelligence Based Automated Valuation Models (AI-AVM), The Journal of Real Estate Finance and Economics, (2021) 1-26.

[69] G.-Z. Fan, S.E. Ong, H.C. Koh, Determinants of house price: A decision tree approach, Urban Studies, 43 (2006) 2301-2315.

[70] E.A. Antipov, E.B.J.E.S.w.A. Pokryshevskaya, Mass appraisal of residential apartments: An application of Random forest for valuation and a CART-based approach for model diagnostics, 39 (2012) 1772-1778.

[71] L. Breiman, Bagging predictors, Machine learning, 24 (1996) 123-140.

[72] M.W. Mitchell, Bias of the Random Forest out-of-bag (OOB) error for certain input parameters, Open Journal of Statistics, 1 (2011) 205.

[73] F. Rosenblatt, The perceptron: a probabilistic model for information storage and organization in the brain, Psychological review, 65 (1958) 386.

[74] E. Worzala, M. Lenk, A. Silva, An exploration of neural networks and its application to real estate valuation, Journal of Real Estate Research, 10 (1995) 185-201.

[75] M.M. Saritas, A. Yasar, Performance analysis of ANN and Naive Bayes classification algorithm for data classification, International Journal of Intelligent Systems and Applications in Engineering, 7 (2019) 88-91.

[76] P. Monika, D. Venkatesan, DI-ANN clustering algorithm for pruning in MLP neural network, Indian Journal of Science and Technology, 8 (2015) 1.

[77] K. Fukushima, A neural network for visual pattern recognition, Computer, 21 (1988) 65-75.

[78] M.L. Hines, N.T. Carnevale, The NEURON simulation environment, Neural computation, 9 (1997) 1179-1209.

[79] S. Leijnen, F.v. Veen, The neural network zoo, Multidisciplinary Digital Publishing Institute Proceedings, 2020, pp. 9.

[80] D.E. Rumelhart, R. Durbin, R. Golden, Y. Chauvin, Backpropagation: The basic theory, Backpropagation: Theory, architectures and applications, (1995) 1-34.

[81] H. Byun, S.-W. Lee, A survey on pattern recognition applications of support vector machines, International Journal of Pattern Recognition and Artificial Intelligence, 17 (2003) 459-486.

[82] N.I. Sapankevych, R. Sankar, Time series prediction using support vector machines: a survey, IEEE Computational Intelligence Magazine, 4 (2009) 24-38.

[83] K.C. Lam, C. Yu, C.K. Lam, Support vector machine and entropy based decision support system for property valuation, Journal of Property Research, 26 (2009) 213-233.

[84] B. Scholkopf, A.J. Smola, Learning with kernels: support vector machines, regularization, optimization, and beyond, Adaptive Computation and Machine Learning Series2018.

[85] E. Acar, M. Altun, Classification of the Agricultural Crops Using Landsat-8 NDVI Parameters by Support Vector Machine, Balkan Journal of Electrical and Computer Engineering, (2021).

[86] W.S. Noble, What is a support vector machine?, Nature Biotechnology, 24 (2006) 1565-1567.

[87] S. Han, C. Qubo, H. Meng, Parameter selection in SVM with RBF kernel function, World Automation Congress 2012, IEEE, 2012, pp. 1-4.

[88] J.-H. Chen, C.F. Ong, L. Zheng, S.-C. Hsu, Forecasting spatial dynamics of the housing market using support vector machine, International Journal of Strategic Property Management, 21 (2017) 273-283.

[89] M. Sellathurai, S. Haykin, The separability theory of hyperbolic tangent kernels and support vector machines for pattern classification, 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No. 99CH36258), IEEE, 1999, pp. 1021-1024.

[90] H.-T. Lin, C.-J. Lin, A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods, submitted to Neural Computation, 3 (2003) 16.

[91] Y. Tang, W. Guo, J. Gao, Efficient model selection for support vector machine with Gaussian kernel function, 2009 IEEE Symposium on Computational Intelligence and Data Mining, IEEE, 2009, pp. 40-45.

[92] J.-S. Jang, ANFIS: adaptive-network-based fuzzy inference system, IEEE transactions on systems, man, and cybernetics, 23 (1993) 665-685.

[93] L.A.J.I. Zadeh, control, Fuzzy sets, 8 (1965) 338-353.

[94] S. Yilmazer, S. Kocaman, A Comparative Assessment of Adaptive Neuro-Fuzzy Inference System and Random Forest Method for Mass Appraisal of Housing Prices, Manuscript submitted for publication (2022).

[95] E. Mamdani, S.J.I.J.o.M.-M.S. Assilian, An experiment in linguistic synthesis with a fuzzy logic controller, 7 (1975) 1-13.

[96] T. Takagi, M.J.I.t.o.s. Sugeno, man,, cybernetics, Fuzzy identification of systems and its applications to modeling and control, (1985) 116-132.

[97] N. Talpur, M.N.M. Salleh, K. Hussain, An investigation of membership functions on performance of ANFIS for solving classification problems, IOP Conference Series: Materials Science and Engineering, IOP Publishing, 2017, pp. 012103.

[98] V. Kreinovich, C. Quintana, L. Reznik, Gaussian membership functions are most adequate in representing uncertainty in measurements, Proceedings of NAFIPS, 1992, pp. 15-17.

[99] J. Chen, K. de Hoogh, J. Gulliver, B. Hoffmann, O. Hertel, M. Ketzel, M. Bauwelinck, A. Van Donkelaar, U.A. Hvidtfeldt, K. Katsouyanni, A comparison of linear regression, regularization, and machine learning algorithms to develop Europe-wide spatial models of fine particles and nitrogen dioxide, Environment international, 130 (2019) 104934.

[100] S. Kavitha, S. Varuna, R. Ramya, A comparative analysis on linear regression and support vector regression, 2016 online international conference on green engineering and technologies (IC-GET), IEEE, 2016, pp. 1-5.

[101] M.J. Sharma, S.J. Yu, Stepwise regression data envelopment analysis for variable reduction, Applied Mathematics and Computation, 253 (2015) 126-134.

[102] T. Kim, S. Sharda, X. Zhou, R.M. Pendyala, A stepwise interpretable machine learning framework using linear regression (LR) and long short-term memory (LSTM): City-wide demand-side prediction of yellow taxi and for-hire vehicle (FHV) service, Transportation Research Part C: Emerging Technologies, 120 (2020) 102786.

[103] M. Uzair, N. Jamil, Effects of Hidden Layers on the Efficiency of Neural networks, 2020 IEEE 23rd International Multitopic Conference (INMIC), IEEE, 2020, pp. 1-6.

[104] T.M. Oshiro, P.S. Perez, J.A. Baranauskas, How many trees in a random forest?, International workshop on machine learning and data mining in pattern recognition, Springer, 2012, pp. 154-168.

[105] C. Kertész, Rigidity-Based Surface Recognition for a Domestic Legged Robot, IEEE Robotics and Automation Letters Journal with IEEE International Conference on Robotics and Automation presentation, 1 (2016) 309-315.

[106] M. Du, N. Liu, X. Hu, Techniques for interpretable machine learning, Communications of the ACM, 63 (2019) 68-77.

[107] M. Velez, P. Jamshidi, N. Siegmund, S. Apel, C. Kästner, White-box analysis over machine learning: Modeling performance of configurable systems, 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE), IEEE, 2021, pp. 1072-1084.

[108] C. Molnar, Interpretable machine learning, Lulu. com2020.

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

# APPENDICES

**Appendix A – Table A.1** The initial list of variables derived from the datasets and assembled by the expert.

| **Initial Variable List** | **Variable Name** | **Description** | **Measure** | **Role In Model** |
|---|---|---|---|---|
| a | Neighbourhood | Neighbourhood | Nominal | None |
| b | Street | Street | Scale | None |
| c | IU-ID | Individual Unit ID | Scale | None |
| d | Blok | Blok | Scale | None |
| e | Parcel | Parcel | Scale | None |
| 1 | NROOMS | Number of Rooms | Nominal | Input |
| 2 | NBATH | Number of Baths | Nominal | Input |
| 3 | NIU | Number of Individual Units | Scale | Input |
| 4 | NBF | Number of Building Floor | Nominal | Input |
| 5 | PARLOC | Location of Building in Parcel | Nominal | Input |
| 6 | FLR_LOC | Floor Location | Scale | Input |
| 7 | CAREA | Construction Right (Area) | Scale | Input |
| 8 | ELVTR | Building Has or Not Elevator | Nominal | Input |
| 9 | CMPLX | Building is in Campus or Not | Nominal | Input |
| 10 | BALC | 'Balcony' | Nominal | Input |
| 11 | BL | 'Building Licence' | Nominal | Input |
| 12 | D-SLN | 'Dependent Saloon Or Not' | Nominal | Input |
| 13 | D-STTS | 'Developing Status' | Nominal | Input |
| 14 | DIS-MET | 'Distance To Metro' | Scale | Input |
| 15 | DIS-BUS | 'Distance To Bus Stop' | Scale | Input |
| 16 | DIS-CEN | 'Distance To Centre' | Scale | Input |

| | | | | |
|---|---|---|---|---|
| 17 | DIS-COL | Distance To Social' | Scale | Input |
| 18 | DIS-CUL | 'Distance To Cultural Area' | Scale | Input |
| 19 | DIS-GAR | 'Distance To Garbage' | Scale | Input |
| 20 | DIS-HOS | 'Distance To Hospital' | Scale | Input |
| 21 | DIS-ROD | 'Distance To Main Road' | Scale | Input |
| 22 | DIS-MAL | 'Distance To Mall' | Scale | Input |
| 23 | DIS-MRK | 'Distance To Market' | Scale | Input |
| 24 | DIS-TRA | Distance To Train' | Scale | Input |
| 25 | DIS-SCHL | 'Distance To Primary School' | Scale | Input |
| 26 | DIS-SOC | Distance To Social Area' | Scale | Input |
| 27 | DIS-UNI | 'Distance To University' | Scale | Input |
| 28 | FR-GARD | 'Front Garden' | Scale | Input |
| 29 | ROADFR | 'Main Road Frontage or Not' | Nominal | Input |
| 30 | H-MAX | 'Max Construction Height' | Scale | Input |
| 31 | NM-FACA | 'Number Of Facade' | Scale | Input |
| 32 | NM-PARK | 'Number Of Parking Area' | Scale | Input |
| 33 | CL-PARK | 'Parking Or Not' | Nominal | Input |
| 34 | R-SQM | 'Residence Gross Area' | Scale | Input |
| 35 | R-OPSQM | 'Residence Gross Open Space' | Scale | Input |
| 36 | YBUILT | 'Building Year' | Scale | Input |

**Appendix B – Table B.1** The Ratio Results of the MRA Stepwise Model.

| Neighbourhood | Count | Median | Mean | Wtd. Mean | SD | COD | COV | PRD |
|---|---|---|---|---|---|---|---|---|
| ABIDINPASA | 7 | 1.121 | 1.171 | 1.097 | 0.242 | 16.834 | 20.627 | 1.068 |
| AKDERE | 21 | 1.068 | 1.064 | 1.043 | 0.118 | 9.018 | 11.055 | 1.020 |
| ALTIAGAÇ | 2 | 1.148 | 1.148 | 1.128 | 0.206 | 12.687 | 17.942 | 1.019 |
| ASIKVEYSE | 6 | 0.943 | 0.946 | 0.931 | 0.091 | 6.942 | 9.613 | 1.016 |
| BAHÇELERÜS | 6 | 0.977 | 1.000 | 1.007 | 0.077 | 5.476 | 7.734 | 0.993 |
| BASAK | 18 | 0.984 | 0.988 | 0.975 | 0.095 | 7.439 | 9.646 | 1.014 |
| BOGAZIÇI | 40 | 1.127 | 1.153 | 1.129 | 0.181 | 12.193 | 15.684 | 1.021 |
| BOSTANCIK | 6 | 1.170 | 1.144 | 1.123 | 0.169 | 11.914 | 14.749 | 1.019 |
| CENGIZHAN | 31 | 1.041 | 1.045 | 1.041 | 0.155 | 11.684 | 14.862 | 1.003 |
| ÇAGLAYAN | 5 | 1.151 | 1.082 | 1.086 | 0.118 | 7.455 | 10.903 | 0.997 |
| DERBENT | 7 | 1.067 | 1.037 | 1.033 | 0.103 | 7.064 | 9.907 | 1.004 |
| DURALIALI | 101 | 0.995 | 1.010 | 0.966 | 0.195 | 14.715 | 19.284 | 1.046 |
| EGE | 19 | 0.889 | 0.931 | 0.903 | 0.215 | 15.766 | 23.095 | 1.031 |
| EKIN | 59 | 0.940 | 0.960 | 0.946 | 0.156 | 13.151 | 16.228 | 1.015 |
| FAHRIKORU | 33 | 1.046 | 1.074 | 1.040 | 0.169 | 12.237 | 15.697 | 1.033 |
| GENERALZE | 92 | 0.989 | 1.006 | 0.987 | 0.131 | 10.160 | 12.972 | 1.019 |
| HARMAN | 12 | 1.057 | 1.123 | 1.077 | 0.259 | 14.245 | 23.076 | 1.043 |
| HÜREL | 6 | 1.028 | 1.107 | 1.091 | 0.198 | 11.857 | 17.897 | 1.015 |
| HÜSEYINGAZ | 5 | 0.981 | 1.058 | 1.006 | 0.198 | 15.222 | 18.694 | 1.052 |
| KARTALTEPE | 11 | 1.105 | 1.195 | 1.185 | 0.236 | 15.598 | 19.745 | 1.008 |
| KAZIMORBA | 39 | 0.965 | 0.992 | 0.987 | 0.110 | 8.782 | 11.083 | 1.005 |
| KÜÇÜKKAYA | 47 | 1.039 | 1.067 | 1.051 | 0.184 | 13.534 | 17.223 | 1.016 |
| MISKET | 51 | 1.092 | 1.103 | 1.095 | 0.169 | 12.056 | 15.280 | 1.007 |
| MUTLU | 178 | 0.987 | 1.003 | 0.984 | 0.147 | 10.980 | 14.613 | 1.020 |
| PEYAMISEF | 54 | 1.025 | 1.025 | 1.015 | 0.129 | 10.242 | 12.549 | 1.010 |
| SAFAKTEPE | 1 | 0.746 | 0.746 | 0.746 | 0.000 | 0.000 | 0.000 | 1.000 |
| SAHAPGÜRL | 7 | 0.967 | 0.982 | 0.977 | 0.097 | 8.046 | 9.840 | 1.005 |
| SAHINTEPE | 77 | 1.039 | 1.036 | 1.016 | 0.150 | 10.875 | 14.464 | 1.020 |
| SEHITCENG | 2 | 0.896 | 0.896 | 0.888 | 0.137 | 10.807 | 15.284 | 1.009 |
| SIRINTEPE | 15 | 1.054 | 1.066 | 1.046 | 0.131 | 7.897 | 12.286 | 1.020 |
| TUZLUÇAYIR | 7 | 1.054 | 1.030 | 0.982 | 0.181 | 13.070 | 17.602 | 1.049 |
| TÜRKÖZÜ | 35 | 1.000 | 0.986 | 0.964 | 0.153 | 11.752 | 15.527 | 1.022 |
| **Combined** | **1000** | **1.0115** | **1.029** | **1.0047** | **0.16** | **12.28** | **15.19** | **1.0244** |

**Appendix C – Table C.1** The Ratio Results of the ANFIS Model.

| Neighbourhood | Count | Median | Mean | Wtd. Mean | SD | COD | COV | PRD |
|---|---|---|---|---|---|---|---|---|
| ABIDINPAşa | 7 | 1.121 | 1.168 | 1.096 | 0.238 | 16.677 | 20.354 | 1.066 |
| AKDERE | 21 | 1.051 | 1.060 | 1.040 | 0.117 | 8.772 | 11.044 | 1.019 |
| ALTIAĞAÇ | 2 | 1.147 | 1.147 | 1.127 | 0.200 | 12.320 | 17.423 | 1.018 |
| AŞIK VEYSEL | 6 | 0.947 | 0.951 | 0.934 | 0.101 | 7.990 | 10.574 | 1.018 |
| BAHÇELER | 6 | 0.953 | 0.979 | 0.989 | 0.111 | 7.280 | 11.328 | 0.989 |
| BAŞAK | 18 | 0.962 | 0.964 | 0.958 | 0.077 | 6.385 | 8.001 | 1.006 |
| BOĞAZIÇ | 40 | 1.117 | 1.144 | 1.116 | 0.181 | 11.335 | 15.786 | 1.025 |
| BOSTANCIK | 6 | 1.102 | 1.094 | 1.076 | 0.154 | 10.902 | 14.077 | 1.017 |
| CENGIZHAN | 31 | 1.040 | 1.051 | 1.042 | 0.163 | 11.820 | 15.464 | 1.009 |
| ÇAĞLAYAN | 5 | 1.086 | 1.069 | 1.071 | 0.113 | 6.702 | 10.595 | 0.998 |
| DERBENT | 7 | 1.130 | 1.075 | 1.071 | 0.134 | 8.644 | 12.445 | 1.003 |
| DURALI AL. | 101 | 0.974 | 1.005 | 0.960 | 0.190 | 14.531 | 18.909 | 1.047 |
| EGE | 19 | 0.882 | 0.905 | 0.877 | 0.208 | 16.091 | 22.957 | 1.032 |
| EKIN | 59 | 0.956 | 0.956 | 0.941 | 0.136 | 10.922 | 14.244 | 1.017 |
| FAHRI KOR. | 33 | 1.092 | 1.121 | 1.139 | 0.176 | 12.707 | 15.675 | 0.984 |
| GENERAL Z.D. | 92 | 0.992 | 1.015 | 0.997 | 0.142 | 10.368 | 13.983 | 1.019 |
| HARMAN | 12 | 1.055 | 1.127 | 1.081 | 0.264 | 15.759 | 23.460 | 1.042 |
| HÜREL | 6 | 1.057 | 1.118 | 1.099 | 0.205 | 12.500 | 18.367 | 1.017 |
| HÜSEYINGAZI | 5 | 0.958 | 1.061 | 1.008 | 0.206 | 16.263 | 19.409 | 1.053 |
| KARTALTEPE | 11 | 1.080 | 1.186 | 1.177 | 0.233 | 15.740 | 19.662 | 1.007 |
| KAZIM ORB. | 39 | 0.951 | 0.982 | 0.972 | 0.113 | 9.173 | 11.534 | 1.010 |
| KÜÇÜK K. | 47 | 1.066 | 1.064 | 1.045 | 0.183 | 13.172 | 17.165 | 1.018 |
| MISKET | 51 | 1.079 | 1.087 | 1.075 | 0.161 | 11.372 | 14.788 | 1.011 |
| MUTLU | 178 | 0.980 | 1.001 | 0.981 | 0.147 | 11.102 | 14.658 | 1.020 |
| PEYAMI SEFA | 54 | 1.011 | 1.023 | 1.011 | 0.134 | 10.797 | 13.111 | 1.012 |
| ŞAFAKTEPE | 1 | 0.751 | 0.751 | 0.751 | 0.000 | 0.000 | 0.000 | 1.000 |
| ŞAHAP GÜ. | 7 | 0.990 | 1.000 | 0.996 | 0.107 | 8.588 | 10.682 | 1.005 |
| ŞAHINTEPE | 77 | 1.046 | 1.034 | 1.015 | 0.151 | 10.804 | 14.646 | 1.018 |
| ŞEHIT C.T. | 2 | 0.920 | 0.920 | 0.916 | 0.065 | 5.016 | 7.094 | 1.004 |
| ŞIRINTEPE | 15 | 1.046 | 1.076 | 1.053 | 0.138 | 7.645 | 12.830 | 1.021 |
| TUZLUÇAYIR | 7 | 1.013 | 1.007 | 0.975 | 0.114 | 8.523 | 11.282 | 1.033 |
| TÜRKÖZÜ | 35 | 0.989 | 0.986 | 0.963 | 0.153 | 12.072 | 15.471 | 1.024 |
| **COMBINED** | **1000** | **1.013** | **1.028** | **1.0042** | **0.156** | **12.226** | **15.151** | **1.023** |

**Appendix D – Table D.1** The Ratio Results of the SVM Model.

| **Neighbourhood** | **Count** | **Median** | **Mean** | **Wtd. Mean** | **SD** | **COD** | **COV** | **PRD** |
|---|---|---|---|---|---|---|---|---|
| ABIDINPAşa | 7 | 1.044 | 1.127 | 1.070 | 0.197 | 15.151 | 17.497 | 1.053 |
| AKDERE | 21 | 1.047 | 1.057 | 1.042 | 0.107 | 7.949 | 10.157 | 1.015 |
| ALTIAĞAÇ | 2 | 1.181 | 1.181 | 1.161 | 0.200 | 11.946 | 16.895 | 1.017 |
| AŞIK VEYSEL | 6 | 1.000 | 0.974 | 0.962 | 0.071 | 5.442 | 7.270 | 1.013 |
| BAHÇELER | 6 | 0.988 | 0.962 | 0.969 | 0.072 | 5.489 | 7.483 | 0.993 |
| BAŞAK | 18 | 0.976 | 0.999 | 0.992 | 0.069 | 5.183 | 6.952 | 1.007 |
| BOĞAZIÇ | 40 | 1.089 | 1.105 | 1.085 | 0.169 | 11.144 | 15.271 | 1.018 |
| BOSTANCIK | 6 | 1.080 | 1.072 | 1.045 | 0.171 | 12.559 | 15.971 | 1.025 |
| CENGIZHAN | 31 | 1.037 | 1.034 | 1.024 | 0.130 | 9.534 | 12.544 | 1.010 |
| ÇAĞLAYAN | 5 | 1.044 | 1.035 | 1.039 | 0.105 | 7.726 | 10.180 | 0.997 |
| DERBENT | 7 | 1.022 | 1.014 | 1.010 | 0.088 | 6.726 | 8.659 | 1.004 |
| DURALI AL. | 101 | 0.976 | 0.980 | 0.936 | 0.172 | 13.285 | 17.490 | 1.048 |
| EGE | 19 | 0.899 | 0.918 | 0.888 | 0.220 | 16.400 | 24.006 | 1.034 |
| EKIN | 59 | 0.969 | 0.976 | 0.951 | 0.141 | 10.553 | 14.419 | 1.026 |
| FAHRI KOR. | 33 | 1.009 | 1.060 | 1.033 | 0.155 | 10.823 | 14.628 | 1.027 |
| GENERAL Z.D. | 92 | 0.983 | 1.001 | 0.977 | 0.127 | 9.280 | 12.640 | 1.025 |
| HARMAN | 12 | 1.034 | 1.132 | 1.094 | 0.258 | 14.425 | 22.753 | 1.035 |
| HÜREL | 6 | 1.019 | 1.084 | 1.071 | 0.165 | 10.474 | 15.173 | 1.013 |
| HÜSEYINGAZI | 5 | 0.943 | 1.048 | 0.999 | 0.190 | 15.172 | 18.127 | 1.050 |
| KARTALTEPE | 11 | 1.102 | 1.158 | 1.152 | 0.237 | 16.605 | 20.492 | 1.006 |
| KAZIM ORB. | 39 | 0.969 | 0.972 | 0.961 | 0.092 | 7.663 | 9.483 | 1.011 |
| KÜÇÜK K. | 47 | 0.989 | 1.018 | 1.003 | 0.156 | 11.235 | 15.309 | 1.015 |
| MISKET | 51 | 1.047 | 1.067 | 1.053 | 0.161 | 11.571 | 15.062 | 1.014 |
| MUTLU | 178 | 0.982 | 0.991 | 0.976 | 0.137 | 10.133 | 13.835 | 1.016 |
| PEYAMI SEFA | 54 | 1.025 | 1.020 | 1.011 | 0.124 | 9.579 | 12.108 | 1.009 |
| ŞAFAKTEPE | 1 | 0.790 | 0.790 | 0.790 | 0.000 | 0.000 | 0.000 | 1.000 |
| ŞAHAP GÜ. | 7 | 0.959 | 0.964 | 0.961 | 0.063 | 5.218 | 6.551 | 1.003 |
| ŞAHINTEPE | 77 | 1.031 | 1.018 | 1.005 | 0.135 | 9.808 | 13.211 | 1.014 |
| ŞEHIT C.T. | 2 | 0.910 | 0.910 | 0.904 | 0.097 | 7.507 | 10.617 | 1.006 |
| ŞIRINTEPE | 15 | 1.037 | 1.084 | 1.048 | 0.164 | 8.594 | 15.137 | 1.034 |
| TUZLUÇAYIR | 7 | 1.027 | 0.984 | 0.964 | 0.078 | 6.178 | 7.885 | 1.020 |
| TÜRKÖZÜ | 35 | 0.964 | 0.972 | 0.954 | 0.133 | 10.421 | 13.724 | 1.019 |
| **COMBINED** | **1000** | **1.0009** | **1.0128** | **0.9887** | **0.1434** | **10.867** | **14.1599** | **1.0244** |

**Appendix E – Table E.1** The Ratio Results of the ANN Model.

| Neighbourhood | Count | Median | Mean | Wtd Mean | SD | COD | COV | PRD |
|---|---|---|---|---|---|---|---|---|
| ABIDINPAŞA | 7 | 1.021 | 1.108 | 1.050 | 0.203 | 15.489 | 18.331 | 1.055 |
| AKDERE | 21 | 1.059 | 1.057 | 1.043 | 0.097 | 7.150 | 9.134 | 1.013 |
| ALTIAĞAÇ | 2 | 1.181 | 1.181 | 1.158 | 0.225 | 13.451 | 19.023 | 1.020 |
| AŞIK VEYSEL | 6 | 1.017 | 0.993 | 0.982 | 0.068 | 4.867 | 6.807 | 1.011 |
| BAHÇELER | 6 | 0.971 | 0.962 | 0.969 | 0.067 | 5.015 | 6.956 | 0.993 |
| BAŞAK | 18 | 0.976 | 0.996 | 0.988 | 0.072 | 5.578 | 7.245 | 1.007 |
| BOĞAZIÇ | 40 | 1.064 | 1.092 | 1.074 | 0.149 | 9.964 | 13.653 | 1.017 |
| BOSTANCIK | 6 | 1.109 | 1.078 | 1.055 | 0.151 | 10.701 | 14.037 | 1.021 |
| CENGIZHAN | 31 | 1.061 | 1.052 | 1.043 | 0.125 | 9.217 | 11.921 | 1.008 |
| ÇAĞLAYAN | 5 | 1.039 | 1.039 | 1.042 | 0.067 | 4.918 | 6.457 | 0.997 |
| DERBENT | 7 | 1.042 | 1.030 | 1.026 | 0.083 | 6.241 | 8.019 | 1.004 |
| DURALI AL | 101 | 0.981 | 0.994 | 0.955 | 0.157 | 12.349 | 15.740 | 1.041 |
| EGE | 19 | 0.916 | 0.922 | 0.896 | 0.196 | 15.006 | 21.201 | 1.029 |
| EKIN | 59 | 0.993 | 0.996 | 0.974 | 0.130 | 9.747 | 13.018 | 1.023 |
| FAHRI KOR. | 33 | 1.029 | 1.059 | 1.030 | 0.133 | 9.247 | 12.580 | 1.028 |
| GENERAL Z.D. | 92 | 0.998 | 1.005 | 0.983 | 0.117 | 8.507 | 11.641 | 1.022 |
| HARMAN | 12 | 1.033 | 1.120 | 1.086 | 0.235 | 12.752 | 21.002 | 1.031 |
| HÜREL | 6 | 1.054 | 1.095 | 1.084 | 0.134 | 8.971 | 12.278 | 1.010 |
| HÜSEYINGAZI | 5 | 0.952 | 1.041 | 0.999 | 0.160 | 12.631 | 15.370 | 1.042 |
| KARTALTEPE | 11 | 1.129 | 1.153 | 1.146 | 0.196 | 13.266 | 17.033 | 1.006 |
| KAZIM ORBAY | 39 | 0.979 | 0.983 | 0.975 | 0.086 | 6.787 | 8.697 | 1.008 |
| KÜÇÜK K. | 47 | 1.018 | 1.025 | 1.010 | 0.137 | 10.025 | 13.379 | 1.015 |
| MISKET | 51 | 1.065 | 1.059 | 1.044 | 0.139 | 10.051 | 13.091 | 1.014 |
| MUTLU | 178 | 0.988 | 0.996 | 0.981 | 0.127 | 9.202 | 12.761 | 1.016 |
| PEYAMI SEFA | 54 | 1.022 | 1.027 | 1.018 | 0.123 | 9.521 | 11.963 | 1.009 |
| ŞAFAKTEPE | 1 | 0.826 | 0.826 | 0.826 | 0.000 | 0.000 | 0.000 | 1.000 |
| ŞAHAP GÜ. | 7 | 0.987 | 0.996 | 0.992 | 0.072 | 5.845 | 7.275 | 1.004 |
| ŞAHINTEPE | 77 | 1.039 | 1.031 | 1.016 | 0.126 | 8.875 | 12.198 | 1.014 |
| ŞEHIT C.T. | 2 | 0.945 | 0.945 | 0.939 | 0.098 | 7.345 | 10.387 | 1.006 |
| ŞIRINTEPE | 15 | 1.034 | 1.079 | 1.045 | 0.156 | 7.836 | 14.415 | 1.032 |
| TUZLUÇAYIR | 7 | 1.022 | 1.006 | 0.985 | 0.083 | 6.720 | 8.229 | 1.021 |
| TÜRKÖZÜ | 35 | 0.997 | 0.978 | 0.960 | 0.123 | 9.056 | 12.622 | 1.018 |
| **COMBINED** | **1000** | **1.01** | **1.02** | **0.996** | **0.13** | **9.947** | **12.8579** | **1.0225** |

**Appendix F – Table F.1** The Ratio Results of the RF Model.

| Neighbourhood | Count | Median | Mean | Wtd Mean | SD | COD | COV | PRD |
|---|---|---|---|---|---|---|---|---|
| ABIDINPAŞA | 7 | 0.997 | 1.089 | 1.029 | 0.211 | 15.844 | 19.415 | 1.058 |
| AKDERE | 21 | 1.086 | 1.058 | 1.045 | 0.092 | 6.743 | 8.650 | 1.012 |
| ALTIAĞAÇ | 2 | 1.180 | 1.180 | 1.154 | 0.250 | 14.958 | 21.155 | 1.022 |
| AŞIK VEYSEL | 6 | 1.034 | 1.012 | 1.002 | 0.067 | 4.772 | 6.572 | 1.009 |
| BAHÇELER | 6 | 0.954 | 0.962 | 0.968 | 0.074 | 5.657 | 7.691 | 0.994 |
| BAŞAK | 18 | 0.958 | 0.992 | 0.985 | 0.082 | 6.522 | 8.235 | 1.008 |
| BOĞAZIÇ | 40 | 1.037 | 1.080 | 1.063 | 0.136 | 9.189 | 12.615 | 1.016 |
| BOSTANCIK | 6 | 1.119 | 1.084 | 1.065 | 0.135 | 9.094 | 12.417 | 1.017 |
| CENGIZHAN | 31 | 1.062 | 1.069 | 1.063 | 0.129 | 9.548 | 12.027 | 1.006 |
| ÇAĞLAYAN | 5 | 1.039 | 1.043 | 1.045 | 0.031 | 2.147 | 2.973 | 0.998 |
| DERBENT | 7 | 1.062 | 1.046 | 1.042 | 0.082 | 5.774 | 7.847 | 1.004 |
| DURALI AL | 101 | 0.984 | 1.008 | 0.974 | 0.150 | 11.784 | 14.860 | 1.035 |
| EGE | 19 | 0.912 | 0.926 | 0.904 | 0.175 | 14.228 | 18.917 | 1.024 |
| EKIN | 59 | 1.017 | 1.017 | 0.996 | 0.125 | 9.460 | 12.298 | 1.020 |
| FAHRI KOR. | 33 | 1.044 | 1.058 | 1.027 | 0.122 | 8.669 | 11.563 | 1.030 |
| GENERAL Z.D. | 92 | 0.999 | 1.008 | 0.989 | 0.115 | 8.383 | 11.431 | 1.019 |
| HARMAN | 12 | 1.032 | 1.108 | 1.079 | 0.215 | 12.245 | 19.404 | 1.027 |
| HÜREL | 6 | 1.088 | 1.105 | 1.097 | 0.108 | 7.564 | 9.764 | 1.007 |
| HÜSEYINGAZI | 5 | 0.960 | 1.034 | 1.000 | 0.130 | 10.136 | 12.581 | 1.034 |
| KARTALTEPE | 11 | 1.117 | 1.148 | 1.141 | 0.161 | 11.080 | 14.059 | 1.006 |
| KAZIM ORBAY | 39 | 0.981 | 0.994 | 0.989 | 0.089 | 6.520 | 8.943 | 1.005 |
| KÜÇÜK K. | 47 | 1.033 | 1.033 | 1.018 | 0.124 | 9.164 | 12.011 | 1.015 |
| MISKET | 51 | 1.052 | 1.051 | 1.036 | 0.127 | 9.368 | 12.105 | 1.015 |
| MUTLU | 178 | 0.992 | 1.001 | 0.986 | 0.123 | 8.772 | 12.255 | 1.015 |
| PEYAMI SEFA | 54 | 1.022 | 1.034 | 1.025 | 0.128 | 9.574 | 12.425 | 1.008 |
| ŞAFAKTEPE | 1 | 0.861 | 0.861 | 0.861 | 0.000 | 0.000 | 0.000 | 1.000 |
| ŞAHAP GÜ. | 7 | 1.014 | 1.027 | 1.023 | 0.084 | 6.439 | 8.154 | 1.004 |
| ŞAHINTEPE | 77 | 1.050 | 1.043 | 1.028 | 0.124 | 8.737 | 11.866 | 1.014 |
| ŞEHIT C.T. | 2 | 0.980 | 0.980 | 0.974 | 0.100 | 7.193 | 10.173 | 1.006 |
| ŞIRINTEPE | 15 | 1.032 | 1.074 | 1.042 | 0.150 | 7.305 | 13.940 | 1.030 |
| TUZLUÇAYIR | 7 | 1.000 | 1.028 | 1.007 | 0.097 | 7.649 | 9.396 | 1.021 |
| TÜRKÖZÜ | 35 | 1.006 | 0.984 | 0.966 | 0.120 | 8.355 | 12.231 | 1.018 |
| **COMBINED** | **1000** | **1.0155** | **1.0249** | **1.0042** | **0.1255** | **9.5** | **12.2432** | **1.0206** |