ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DOKTORA TEZİ

KİTLESEL DEĞERLEMEDE MAKİNE ÖĞRENME ALGORİTMALARI

Sibel CANAZ SEVGEN

GAYRİMENKUL GELİŞTİRME VE YÖNETİMİ ANABİLİM DALI

ANKARA
2022

Her hakkı saklıdır

ÖZET

Doktora Tezi

KİTLESEL DEĞERLEMEDE MAKİNE ÖĞRENME ALGORİTMALARI

Sibel CANAZ SEVGEN

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
Gayrimenkul Geliştirme ve Yönetimi Anabilim Dalı

Danışman: Doç. Dr. Yeşim TANRIVERMİŞ

Gayrimenkullerin değerlerine vergilemeden, satışa birçok alanda ihtiyaç duyulmaktadır. Gayrimenkullerin tek tek değerlemesi yapılacağı gibi toplu halde de bir model dahilinde değerlemesinin yapılması mümkündür. Kitlesel değerleme, gayrimenkullerin az hata, hızlı ve tarafsız bir şekilde değerlemesine olanak sağlamaktadır. Gayrimenkullerin toplu değerlemesinde son yıllarda makine öğrenme algoritmaları kullanılmaya başlanmış olmakla birlikte makine öğrenme algoritmalarının farklı gayrimenkul türlerinin değerlemesinin de kullanımı sınırlı düzeyde kalmıştır. Bu çalışmada, 5 adet makine öğrenme algoritmaları arsa, arazi, konut ve ticari gayrimenkul olmak üzere 4 farklı gayrimenkul türünde test edilmiştir. Bu amaçla kullanılan makine öğrenme algoritmaları, Yapay Sinir Ağlar (YSA), Rasgele Orman (RO), Çoklu Regresyon Analizi (ÇRA), K-En Yakın Komşuluk (k-nn), Destek Vektör Regresyonu (DVR) şeklindedir. Temel Bileşenler Analizi (TBA) ise çoklu değişkene sahip gayrimenkullerin değişken sayısını azaltılması için kullanılmıştır. Araştırmada Ankara İli Merkez ilçeler için toplanan gayrimenkul verileri ile Amerika Birleşik Devletleri, Ames, Iowa şehrine ait veriler kullanılmıştır. Tez çalışmasının ana amacı gayrimenkullerin toplu değerlemesinde hangi makine öğrenme algoritma türünün daha iyi sonucu verdiğinin bulunması ve gayrimenkul değerini etkileyen değişkenlerden en önemli olanların ortaya konulmasıdır. Böylece Türkiye için bir toplu değerleme sisteminin kullanılabilirliği test edilmiş olacaktır. Çalışmanın diğer bir amacı ise gayrimenkul değerleri ile Ankara İli merkez ilçeleri için gayrimenkul değer haritası oluşturmaktır. Ankara ili ve Ames şehri için elde edilen bulgulara göre konut türü gayrimenkullerde en iyi sonucu RO daha sonra ise sırasıyla YSA, k-nn, ve DR algoritmalarının sonuçlar verdiği görülmüştür. Yine konut türü gayrimenkulün RO algoritmasından elde edilen sonuçlarına göre en çok ısıtma ve önemli yerlere olan mesafelerin değerde etkisi olduğu sonucuna varılmıştır. Arazi, arsa ve ticari tür gayrimenkullerde ise RO ve doğrusal regresyonun en iyi performansa sahip olduğu belirlenmiştir. Elde edilen bulgular değerlendirildiğinde tüm Türkiye'deki il ve ilçeler için gayrimenkul değerleme birimlerinin kurulması ve gayrimenkullere ait değişkenlerine ait veri tabanında oluşturulması önerilmektedir. Böylece bu veriler kullanılarak kitlesel değerleme yapılması ile gayrimenkul satış ve vergilendirme gibi işlemlerinin daha hızlı yapılabilmesine olanak sağlanacaktır.

Ekim 2022, 131 sayfa

Anahtar Kelimeler: Kitlesel değerleme, makine öğrenme, rasgele orman, yapay sinir ağları, gayrimenkul değer haritası.

ABSTRACT

Ph.D. Thesis

MACHINE LEARNING ALGORITHMS IN MASS APPRAISAL

Sibel CANAZ SEVGEN

Ankara University
Graduate School of Natural and Applied Sciences
Department of Real Estate Development and Management

Supervisor: Assoc. Prof. Yeşim TANRIVERMİŞ

Information of real estate value is crucial for many fields such as taxation, sales etc. It is possible to valuate many properties together within a model. Such models allow real estate to be valued with less error, quickly and impartially. In recent years, machine learning algorithms have been used in the mass appraisal of real estates. In this thesis study, 5 machine learning algorithms are used in 4 different real estate types; land plot, land, residential and commercial real estate. Machine learning algorithms used for this purpose are; Artificial Neural Networks (ANN), Random Forest (RO), Multiple Regression Analysis (MRA), K-Nearest Neighborhood (k-nn) and Support Vector Regression (SVR). On the other hand, Principal Component Analysis (PCA) was used to reduce the number of variables for real estate with multiple variables. In order to test the research, real estate data collected for the central districts of Ankara and data from the city of Ames, Iowa were used. The main purpose of this thesis study is to find out which machine learning algorithm gives the best result for mass appraisal of real estates and to reveal the most important variables affecting the real estate value. Thus, the feasibility of creating a mass appraisal system for Turkey will be tested. Another aim for the thesis is to create a real estate value map for the central districts of Ankara by using the real estate values. According to the results obtained for the city of Ankara and Ames, Iowa; it was observed that the best algorithm for mass appraisal is RO followed by the ANN, k-nn, and linear regression algorithms, respectively for residential type real estates. According to the results obtained from the RO algorithm of the residential type real estate, it was concluded that the heating and the distances to the key places had the most effect on the value. Besides RO and linear regression gave the best results in land, land plot and commercial type real estates. As a result of the study, it is recommended to establish real estate valuation units for all municipalities, provinces and other administrative units in Turkey and to create a database of real estate variables. Thus, sales, taxation etc. of real estate would be processed faster by conducting mass appraisal using quality database.

September 2022, 131 pages

Key Words: Mass appraisal, machine learning, random forest, artificial neural network, real estate valuation map.

TEŞEKKÜR

Bu tez çalışması kapsamında, öncelikle Doktora öğrenimim boyunca bu süreçteki tezimin yürütülmesinde bilgisi, tecrübesi, öngörüsü ve pozitif enerjisi ile hep yanımda olan değerli danışman hocam Doç. Dr. Yeşim TANRIVERMİŞ'e en derin ve samimi teşekkürlerimi sunarım. Tez izleme toplantılarımda, bana ayırdıkları vakit, tezime sağladıkları katkı ve görüşleri için çok değerli Tez İzleme Komitesi ve savunma jüri üye hocalarım saygıdeğer Prof. Dr. Oğuz GÜNGÖR ve Doç. Dr. Mehmet TAN'a en içten teşekkürlerimi sunarım. Dahası vakitlerini ayırıp tezimi okuyup görüş bildiren saygı değer doktora savunma jüri üye hocalarım Prof. Dr. Mustafa TOMBUL ile Doç. Dr. Serhat BAŞDOĞAN'a katkılarından dolayı en içten teşekkürlerimi sunarım. Ayrıca doktora tezim boyunca verdiği katkılardan dolayı Anabilim Dalı Başkanımız Prof. Dr. Harun TANRIVERMİŞ'e ve Bölüm tüm öğretim üyelerine teşekkürlerimi sunarım.

Bu tez çalışması 21L0669001 proje numarası ile destekleyen Ankara Üniversitesi Bilimsel Araştırma Projeleri (BAP) birime katkılarından dolayı teşekkür ederim. Ankara iline ait konut verilerini bilimsel araştırma amacıyla paylaşan ENDEKSA şirketine ayrıca teşekkür ederim. Tez dönemim boyunca bana bulundukları destekleri için değerli mesai arkadaşlarım Arş. Gör. Ahmet Hilmi ERCİYES, Mühendis Buket Gülsüm ÜLGER, Arş. Gör. Gülnaz ŞENGÜL GÜNEŞ, Mühendis Mehmet Oğuz ÖNDAĞ, Arş. Gör. Dr. Toygun ATASOY, Arş. Gör. Dr. Abdurrahman TURSUN, Öğr. Gör. Sinan GÜNEŞ, Öğr. Gör. Ümit Gedik, Öğr. Gör. Emir SUNGUROĞLU ve tüm diğer tüm Arş. Gör ve Öğr. Gör arkadaşlarım ile Selim ACAR ve Ahmet ÖNGÖN'e teşekkürlerimi sunarım.

Tez sürecim boyunca bana bu yolculukta en büyük desteği veren sevgili çok değerli Eşim, Eray SEVGEN'e her zaman yanımda olduğu, desteği, görüşleri, sevgisi, saygısı ve anlayışı için sonsuz teşekkür ederim. Son ama en önemli olarak sevgili oğlum Egemen SEVGEN ve yakında dünyaya gelecek olan oğlum Atakan SEVGEN'e varlıkları ile bana hayatta güç verdiği, motive etiği için hayatımın en değerlilerine şükürlerimi sunarım.

Sibel CANAZ SEVGEN

Ankara, Ekim 2022

İÇİNDEKİLER

TEZ ONAYI
ETİK ..... **i**
ÖZET ..... **ii**
ABSTRACT ..... **iii**
TEŞEKKÜR ..... **iv**
KISALTMALAR DİZİNİ ..... **vii**
ŞEKİLLER DİZİNİ ..... **viii**
ÇİZELGELER DİZİNİ ..... **x**
1. GİRİŞ **1**
1.1 Araştırmanın Amacı, Kapsamı ve Yöntemi **2**
1.2 Önceki Çalışmaların Değerlendirilmesi ve Literatür Özeti **4**
2. KİTLESEL DEĞERLEME **16**
2.1 Kitlesel Değerlemede Amaç **17**
2.1.1 Kitlesel değerlemede verilerin toplanması ve saklanması **18**
2.1.2 Coğrafi (mekânsal) veri toplanması ve saklanması **18**
2.1.3 Periyodik yerinde denetimlere alternatif **21**
2.1.4 Değerleme yöntemleri ve yeniden değerleme sıklığı **22**
2.2 Standart üzerine genel değerlendirme **24**
3. MAKİNE ÖĞRENME ALGORİTMALARI **25**
3.1 Rasgele Orman Algoritması **29**
3.1.1 Karar ağacı öğrenmesi **30**
3.1.2 Rasgele orman regresyonu **32**
3.2 Yapay Sinir Ağları Algoritması **33**
3.2.1 Yapay sinir ağlarında aşırı uyumlama ve az uyumlama **37**
3.3 Destek Vektör Regresyonu Algoritması **38**
3.4 K En Yakın Komşuluk Algoritması **42**
3.5 Çoklu Regresyon Analizi **44**
3.6 Temel Bileşenler Analizi **47**
3.7 Makine Öğrenme Algoritmalarında Uyum İyiliği Ölçütleri **48**
3.7.1 Değerin güncellenmesi ve makine öğrenme algoritmalarında eğitim ve test verisi **50**

3.8 Değer Haritaları Oluşturulması 52
4. ÇALIŞMA ALANI ve VERİLER 54
4.1 Çalışma Alanı ve Demografik Yapısı 54
4.2 Çalışmada Kullanılan Veriler 59
4.2.1 Konut türü gayrimenkul verileri 59
4.2.2 Arsa türü gayrimenkul verileri 67
4.2.3 Arazi türü gayrimenkul verileri 69
4.2.4 Ticari türü gayrimenkul verileri 70
4.3 Makine Öğrenme Algoritmalarında Parametre Seçimi 71
5. UYGULAMA SONUÇLARI 73
5.1 TBA Analizi Sonuçları 73
5.2 Konut Türü Gayrimenkul İçin Toplu Değerleme Sonuçları 75
5.3 Arsa Türü Gayrimenkul İçin Toplu Değerleme Sonuçları 86
5.4 Arazi Türü Gayrimenkul İçin Toplu Değerleme Sonuçları 94
5.4.1 Arazi türü gayrimenkul için algoritmaların karşılaştırılması 99
5.5 Ticari Tür Gayrimenkul İçin Toplu Değerleme Sonuçları 101
5.5.1 Ticari türü gayrimenkul için algoritmaların karşılaştırılması 106
5.6 Makine Öğrenme Algoritmaları Sonuçlarının Karşılaştırılması 107
5.7 Konut Türü Uygulama Verisi 2 ve Sonuçları 109
5.8 Gayrimenkul Değer Haritası Sonuçları 111
6. GENEL DEĞERLENDİRME VE SONUÇ 114
KAYNAKLAR 119
EKLER 125
EK 1: 5 katmanlı çapraz doğrulama tüm MÖ algoritmaları ve arsa türü gayrimenkul için sonuçlar 126
EK 2: 5 katmanlı çapraz doğrulama tüm MÖ algoritmaları ve arazi türü gayrimenkul için sonuçlar 127
EK 3: 5 katmanlı çapraz doğrulama tüm MÖ algoritmaları ve ticari türü gayrimenkul için sonuçlar 128
ÖZGEÇMİŞ 129

KISALTMALAR DİZİNİ

ARIMA	Autoregressive Integrated Moving Average
AVM	Alış Veriş Merkezi
CBR	Case Based Reasoning
CBS	Coğrafi Bilgi Sistemleri
ÇRA	Çoklu Regresyon Analizi
DR	Doğrusal Regresyon
DVM	Destek Vektör Makine
DVR	Destek Vektör Regresyonu
IAOO	International Association of Assessing Officers
IDW	Inverse Distance Weighted
HKO	Hata Kareler Ortalaması
k-nn	K-en Yakın Komşuluk
KOKH	Kök Ortalama Kare Hata
LiDAR	Light Detection and Ranging
MÖ	Makine Öğrenme
OMH	Ortalama Mutlak Hata
RBFNN	Radial basis function neural network
RO	Rasgele Orman
SA	Sinir Ağları
SPK	Sermaye Piyasası Kurulu
TLR	Tam-logaritmik Regresyon
TBA	Temel Bileşenler Analizi
TDUB	Türkiye Değerleme Uzmanları Birliği
YLR	Yarı-logaritmik Regresyon
YSA	Yapay Sinir Ağları

ŞEKİLLER DİZİNİ

Şekil 2.1 Gayrimenkul değerleme yöntemleri 16
Şekil 3.1 Yapay zekâ, makine öğrenme ve derin öğrenme ilişkisi 25
Şekil 3.2 Makine öğrenme algoritmalarının çalışma prensibi 26
Şekil 3.3 Makine öğrenme yöntemi genel sınıflandırma 27
Şekil 3.4 Gözetimli öğrenme şematik örneklenmesi 28
Şekil 3.5 Gözetimsiz öğrenme şematik örneklenmesi 29
Şekil 3.6 Karar ağacı algoritmasının ihtiyaç kredisi örneği 31
Şekil 3.7 Karar ağacı algoritmasının bir ağaç örneği 33
Şekil 3.8 Biyolojik beyin sinir hücresi 34
Şekil 3.9 Çok katmanlı bir yapay sinir ağı modeli 35
Şekil 3.10 YSA algoritmasında temel aşamalar 36
Şekil 3.11 YSA algoritmasında görülen az, iyi ve aşırı uyumla gösterimi 38
Şekil 3.12 Doğrusal olarak ayrılabilen iki sınıfın temsili gösterimi 39
Şekil 3.13 Doğrusal olmayan problemlerde uzay dönüşümü 40
Şekil 3.14 Destek Vektör Regresyon ve hata parametresi gösterimi 41
Şekil 3.15 K-nn ile tahmin edilen nokta ve uzaklıklar 42
Şekil 3.16 Minkowski (yeşil), Manhattan (mor) ve Öklid (kırmızı) uzaklığı 44
Şekil 3.17 Basit doğrusal regresyon gösterimi 45
Şekil 3.18 MÖ algoritmalarında verilerin eğitim ve test verisi olarak dağılımı ve çapraz doğrulama 52
Şekil 4.1 Ankara ili resmi il sınırı ile çalışma alanı gösterimi 55
Şekil 4.2 Ankara il geneli son yıllara göre genel nüfus grafiği 56
Şekil 4.3 Ankara İli merkez ilçeler yaş aralığına göre son 4 yılda nüfus dağılımı 57
Şekil 4.4 Ankara İli merkez ilçe 2021 yılı nüfus dağılımı (TUIK,2021), ... 58
Şekil 4.5 Ankara İli merkez ilçeler eğitim seviyeleri 58
Şekil 4.6 Gayrimenkul verilerin çalışma alanı içerisinde mekânsal olarak gösterimi 60
Şekil 4.7 Gayrimenkullerin verilerdeki metre kare değerlerine göre gösterimi 61
Şekil 4.8 Çalışma alanı için Google haritadan çekilen okulların konumu 64
Şekil 4.9 Çalışma alanı içerisindeki önemli yerlerin gösterimi (Çalışma alanı içerisindeki, AVM (a), ibadethane (b), metro durakları (c), okul (d), park (e), üniversite (f)) 65
Şekil 4.10 Çalışma alanı içerisindeki önemli yerlerin gösterimi hastane (g) 66

Şekil 4.11 Önemli yerlerin çalışma alanı içerisinde bir arada gösterimi 66
Şekil 4.12 Arsa türü gayrimenkullerin çalışma alanı içerisindeki dağılımı 68
Şekil 4.13 Ticari türü gayrimenkullerin çalışma alanı içerisindeki dağılımı 70
Şekil 4.14 Izgara arama parametre seçimi gösterimi 72
Şekil 5.1 Konut türü gayrimenkul için korelasyon matrisi 74
Şekil 5.2 Konut türü gayrimenkul değişkenler arası korelasyon 74
Şekil 5.3 Konut türü gayrimenkul için RO algoritması hata histogramı 77
Şekil 5.4 Konut türü gayrimenkul için YSA algoritması hata histogramı 79
Şekil 5.5 Konut türü gayrimenkul için DVR algoritması hata histogramı 80
Şekil 5.6 Konut türü gayrimenkul k-nn algoritması için hata histogramı 82
Şekil 5.7 Konut türü gayrimenkul için MÖ algoritmalarında hata oranı 84
Şekil 5.8 Konut türü gayrimenkul için TBA ve normal değişken hata karşılaştırması 85
Şekil 5.9 Arsa türü gayrimenkul için RO algoritması hata histogramı 87
Şekil 5.10 Arsa türü gayrimenkul için YSA algoritması hata histogramı 89
Şekil 5.11 Arsa türü gayrimenkul için DVR algoritması hata histogramı 90
Şekil 5.12 Arsa türü gayrimenkul için k-nn algoritması hata histogramı 91
Şekil 5.13 Arsa türü gayrimenkul için bilinen ve MÖ algoritmalarıyla elde edilen değerler 93
Şekil 5.14 Arazi türü gayrimenkul için RO algoritması hata histogramı 95
Şekil 5.15 Arazi türü gayrimenkul için YSA algoritması hata histogramı 96
Şekil 5.16 Arazi türü gayrimenkul için DVR algoritması hata histogramı 97
Şekil 5.17 Arazi türü gayrimenkul için k-nn algoritması hata histogramı 98
Şekil 5.18 Arazi türü gayrimenkul için MÖ algoritmalarının kıyaslanması 100
Şekil 5.19 Ticari türü gayrimenkul için RO algoritması hata histogramı 102
Şekil 5.20 Ticari türü gayrimenkul için YSA algoritması hata histogramı 103
Şekil 5.21 Ticari türü gayrimenkul için DVR algoritması hata histogramı 104
Şekil 5.22 Ticari türü gayrimenkul için k-nn algoritması hata histogramı 105
Şekil 5.23 Ticari türü gayrimenkul için MÖ algoritmalarının kıyaslanması 106
Şekil 5.24 Konut 2 verisi için bilinen değerler ile RO algoritması ile tahmin edilen değerler 111
Şekil 5.25 Çalışma alanı için konut türü gayrimenkul değer haritası 113

ÇİZELGELER DİZİNİ

Çizelge 1.1 Gayrimenkul toplu değerleme çalışmaları ve özeti 6
Çizelge 1.2 Toplu değerleme kullanılan ülkeler, kurumları ve yöntemler 14
Çizelge 3.1 Makine öğrenme prensibinin basit bir problem ile açıklanması 27
Çizelge 3.2 İhtiyaç kredisi için banka soru ve cevapları 31
Çizelge 3.3 Biyolojik ve yapay sinir hücresi karşılıkları 35
Çizelge 3.4 Sık kullanılan YSA aktivasyon fonksiyonları 37
Çizelge 4.1 Konut türü gayrimenkule ait değişkenler 61
Çizelge 4.2 Isıtma Tipi için kategorik değişkenlerin açıklamaları 62
Çizelge 4.3 Ames, Iowa şehri konut verileri için değişkenler 63
Çizelge 4.4 Konut türü gayrimenkul için örnek veri seti 67
Çizelge 4.5 Arsa türü gayrimenkule ait değişkenler 68
Çizelge 4.6 Arazi türü gayrimenkule ait değişkenler 69
Çizelge 4.7 Ticari türü gayrimenkule ait değişkenler 71
Çizelge 5.1 Konut türü gayrimenkul için MÖ algoritmaları hiper parametre seçimi 75
Çizelge 5.2 Konut türü gayrimenkul için RO algoritması sonuçları 76
Çizelge 5.3 Konut türü gayrimenkul için YSA algoritması sonuçları 78
Çizelge 5.4 Konut türü gayrimenkul için DVR algoritması sonuçları 80
Çizelge 5.5 Konut türü gayrimenkul için k-nn algoritması sonuçları 81
Çizelge 5.6 Konut türü gayrimenkul için doğrusal regresyon algoritması sonuçları 83
Çizelge 5.7 Konut türü gayrimenkul için algoritmaların karşılaştırılması 84
Çizelge 5.8 Konut türü gayrimenkul için TBA ve normal değişken karşılaştırması 85
Çizelge 5.9 Arsa türü gayrimenkul için MÖ algoritmaları hiper parametre seçimi 86
Çizelge 5.10 Arsa türü gayrimenkul için RO algoritması sonuçları 87
Çizelge 5.11 Arsa türü gayrimenkul için YSA algoritması sonuçları 88
Çizelge 5.12 Arsa türü gayrimenkul için DVR algoritması sonuçları 89
Çizelge 5.13 Arsa türü gayrimenkul için k-nn algoritması sonuçları 90
Çizelge 5.14 Arsa türü gayrimenkul için doğrusal regresyon algoritması sonuçları 92
Çizelge 5.15 Arsa türü gayrimenkul için MÖ algoritmalarının kıyaslanması 93
Çizelge 5.16 Arazi türü gayrimenkul için MÖ algoritmaları hiper parametre seçimi 94
Çizelge 5.17 Arazi türü gayrimenkul için RO algoritması sonuçları 95

Çizelge 5.18 Arazi türü gayrimenkul için YSA algoritması sonuçları 96
Çizelge 5.19 Arazi türü gayrimenkul için DVR algoritması sonuçları 97
Çizelge 5.20 Arazi türü gayrimenkul için k-nn algoritması sonuçları 98
Çizelge 5.21 Arazi türü gayrimenkul için ÇRA algoritması sonuçları 99
Çizelge 5.22 Arazi türü gayrimenkul için MÖ algoritmalarının kıyaslanması 100
Çizelge 5.23 Ticari türü gayrimenkul için MÖ algoritmaları hiper parametre seçimi .. 101
Çizelge 5.24 Ticari türü gayrimenkul için RO algoritması sonuçları 102
Çizelge 5.25 Ticari türü gayrimenkul için YSA algoritması sonuçları 103
Çizelge 5.26 Ticari türü gayrimenkul için DVR algoritması sonuçları 104
Çizelge 5.27 Ticari türü gayrimenkul için k-nn algoritması sonuçları 104
Çizelge 5.28 Ticari türü gayrimenkul için regresyon algoritmaları sonuçları 106
Çizelge 5.29 Ticari türü gayrimenkul için MÖ algoritmaları karşılaştırılması 107
Çizelge 5.30 Gayrimenkul türleri için MÖ algoritmalarının toplu kıyaslanması 108
Çizelge 5.31 Konut türü gayrimenkul için MÖ algoritmaları hiper parametre seçimi . 109
Çizelge 5.32 Konut verisi 2 için kalite analiz sonuçları 110

1. GİRİŞ

Son yıllarda, birçok bilim dalında geleneksel yöntemlerle yapılan çalışmalar yerini Makine Öğrenme (MÖ, Machine Learning) diye adlandırılan yöntemlere bırakmıştır. Özellikle, veri boyutu büyüdükçe büyük veri (Big Data) diye adlandırılan yeni araştırma dalı ortaya çıkmış ve büyük boyutlu verilerle ilgili işlemlerde, geleneksel yöntemlerde gözlemlenen hız ve performans düşüklüğü gibi etkiler makine öğrenme algoritmalarının kullanımı artırılarak giderilmeye çalışılmıştır. Makine öğrenme algoritmalarının kullanım alanları oldukça geniştir. Tıptan, mühendisliğe, finanstan, sosyolojiye makine öğrenmenin birçok çeşitli algoritması birçok farklı uygulamalarda kullanılmaktadır. Makine öğrenme basit tanımıyla, herhangi bir uygulamada, bilinen verinin ve sonucu etkileyen değişkenlerin bilgisayara öğretilmesi ve farklı makine öğrenme algoritmalarından birinin kullanılarak geri kalan verinin bir model dahilinde sonucunun otomatik olarak elde edilmesidir.

Makine öğrenme algoritmalarının son yıllarda kullanıldığı alanlardan bir diğeri ise gayrimenkul değerlemesi alanıdır (Kontrimas ve Verikas 2007). Makine öğrenme yöntemi genellikle büyük veri grupları ile ilgilenmekte olup, gayrimenkul değerleme işlemlerinde ise değerleme işlemi tek tek yapılabileceği gibi kitlesel (toplu) halde de yapılabilmektedir. Birçok gayrimenkulün toplu olarak değerlemesi durumu literatürde Kitlesel Değerleme veya daha yaygın olarak Toplu Değerleme olarak adlandırılmaktadır (Erdem 2017a). Diğer bir ifadeyle, toplu değerleme belirli bir tarih aralığında, birçok gayrimenkulün tek tek yerine bir arada yapılan değerleme türüne verilen tanımdır.

Literatür incelendiği zaman gayrimenkullerin toplu değerleme işlemlerinde genellikle Çoklu Regresyon Analizi (ÇRA) yönteminin tercih edildiği görülmekle beraber diğer MÖ algoritmaları da sıklıkla kullanılmıştır. Gayrimenkul toplu değerlemede yaygın olarak kullanılan MÖ algoritmaları; Rasgele Orman (RO), Yapay Sinir Ağları (YSA) ve Destek Vektör Regresyonu (DVR) şeklindedir. Temel Bileşenler Analizi (TBA) ise daha çok, fazla olan değişken sayısı arasında ilişki kurmak ve değişken sayısını azaltmak için kullanılmaktadır. Bilimsel çalışmalarda MÖ algoritmalarının gayrimenkul toplu değerleme işlemlerinde karşılaştırıldığı görülmüştür. Bu çalışmanın ana amacı, 5 adet

farklı MÖ algoritmalarının arsa, arazi ve konut gibi farklı gayrimenkullerin türünde toplu değerleme amacıyla kullanılması ve sonuçlarının analiz edilerek Türkiye'de hangi MÖ algoritmasının daha iyi sonuç verdiğinin bulunmasıdır. Bu amaçla, YSA, DVR, ÇRA, k-en yakın komşuluk (k-nn) ve RO algoritmaları gayrimenkul toplu değerleme için test edilmiştir. Ayrıca, TBA algoritması ise değişken sayısını azaltıp, model başarısı üzerinde TBA algoritmasının etkili olup olmadığını ölçmektedir. Çalışma alanı olarak Ankara İli merkez ilçeler seçilmiş ve gayrimenkule ait değişkenler ve değerleri güvenilir gayrimenkul internet veri portali ile Sermaye Piyasası Kurulu (SPK) lisanslı değerleme firmalarından elde edilmiştir. Ayrıca Türkiye dışında konut türü Amerika Birleşik Devletleri (ABD), Ames, Iowa şehrine ait gayrimenkul ve değişkenleri bedava veri sağlayıcı internet sitesinden elde edilmiş ve sonuçlar farklı bir veri seti ile de test edilerek karşılaştırılmıştır. Çalışma sonuç çıktı ürünü olarak enterpolasyon yöntemi kullanılarak raster tabanlı değer haritası üretilmiştir.

1.1 Araştırmanın Amacı, Kapsamı ve Yöntemi

Türkiye'de gayrimenkul değerleme işlemleri tekil olarak karşılaştırmalı satış analizi, gelir yaklaşımı ve maliyet yöntemi gibi geleneksel yöntemlerle yapılmakta olup, gayrimenkullerin MÖ algoritmaları ile toplu değerlemesi sonucu değerleme işleminin tarafsız ve daha hızlı bir şekilde yapılabildiği bilinmektedir. Bu amaçla bu tez çalışmasında, gayrimenkul toplu değerleme işlemi 5 adet farklı MÖ algoritması kullanılarak farklı gayrimenkul türleri için test ve analiz edilecektir. Tez çalışmasının amaçlarıyla örtüşerek aşağıdaki sorulara cevap bulunması beklenmektedir.

- ✓ Farklı MÖ algoritmaları ile toplu değerlemede genel olarak hangi algoritma en iyi sonucu verecektir?
- ✓ Farklı gayrimenkul tipleri için kullanılacak olan 5 farklı MÖ algoritmalar aynı başarıya mı sahip olacaktır veya hangi gayrimenkul tipinde hangi algoritma daha başarılı olacaktır?
- ✓ Kullanılması planlanan gayrimenkule ait değişkenlerden hangisi (hangileri) gayrimenkullerin değerini en çok etkileyecektir?
- ✓ Türkiye'de gayrimenkul değerleme işlemleri bundan sonra toplu MÖ algoritmalarıyla yapılabilir mi?

Gayrimenkul alanında makine öğrenme algoritmaları az sıklıkta kullanılmakta, özellikle Türkiye'de uygulamada ise neredeyse hiç kullanılmadığı bilinmektedir. Bu amaç ile Ankara İli merkez ilçelerine ait 1,315,675 adet konut, 209 adet arsa, 107 adet arazi ve 109 adet ticari gayrimenkul türü ve bu gayrimenkullere ait farklı değişkenler değerleme raporları ve gayrimenkul ilan sitesi üzerinden elde edilerek kullanılmıştır. ABD, Ames, Iowa şehri için ise 2,353 adet konut verisi ayrıca algoritmaları kıyaslamak için kullanılmıştır. Değişken sayısını artırmak için gayrimenkullerin önemli yerlere olan mesafeleri ayrıca gayrimenkullerin lokasyonu kullanılarak otomatik olarak bir Coğrafi Bilgi Sistemleri (CBS) yazılımı olan ArcGIS 10.2 programı ile hesaplanmıştır. MÖ algoritmaları ise python programlama dili ve kütüphanesi ile çalıştırılmıştır. Sonuçlar farklı uyum iyiliği ölçütleri kullanılarak test edilmiş ve karşılaştırılmıştır.

Tez çalışmasının konusu gayrimenkullerin farklı MÖ algoritmaları kullanılarak toplu değerlemesi ve en başarılı algoritmanın bulunmasıdır. Böylece, vergilendirmeden alım-satım işlemlerine kadar gayrimenkullerin değerleri hızlıca belirlenebilecektir. MÖ algoritmalarının toplu değerlemede test edilmesi amacıyla Ames, Iowa şehri ile Ankara İli merkez ilçelerine ait konut (genellikle apartman dairesi), araz, arsa ve ticari gayrimenkul türleri seçilmiştir. Ankara İli merkez ilçeleri Altındağ, Çankaya, Etimesgut, Gölbaşı, Keçiören, Mamak, Sincan ve Yenimahalle şeklindedir. Çalışma alanı Gölbaşı, Sincan ve Etimesgut ilçelerinin ana merkezlerini diğer ilçelerin ise tamamını kapsayacak şekilde oluşturulmuştur. Tez çalışmasının ilk bölümünde araştırmanın amaç kapsam ve yöntemi belirtilmiş, ikinci bölümde önceki çalışmaların literatür özetine yer verilmiş, üçüncü bölümde toplu değerleme tanımı yapılmış ve toplu değerleme standartları incelenmiş, dördüncü bölümde MÖ algoritmaları açıklanmış, beşinci bölümde çalışma alanı ve kullanılan veriler detaylı olarak gösterilmiş, altıncı bölümde uygulama sonuçları ayrıntılı olarak analiz edilmiş ve değer haritası oluşturulmuş son bölümde ise genel bir değerlendirme yapılmıştır.

1.2 Önceki Çalışmaların Değerlendirilmesi ve Literatür Özeti

Literatürde MÖ algoritmaları kullanılarak toplu değerleme yapan çalışmalar bulunmaktadır. Bu çalışmalardan farklılık gösterenler Çizelge 1.1'de özetlenmiştir. Çalışmaların bir kısmında tek MÖ algoritması kullanılırken bir kısmında ise 2 veya 3 MÖ algoritmasının karşılaştırıldığı görülmüştür. Sıklıkla kullanılan MÖ algoritmalarının başında ÇRA ve YSA olduğu görülmüştür. Araştırmacıların çoğunluğu YSA algoritmasının iyi sonuç verdiğini ortaya koyarken kimi çalışmaların ise ÇRA algoritmasının daha iyi olduğunu vurguladığı görülmüştür.

Literatürde gayrimenkul toplu değerleme için ilk defa (Borst 1991) tarafından bir makine öğrenme algoritması olan YSA algoritmasıyla toplu değerleme çalışması yapılmıştır. Bu çalışmadan sonra, YSA algoritması kullanılarak toplu değerleme çalışması yapan araştırma sayısı artmıştır. Araştırmacılar bilimsel yayınlarında YSA algoritmasının gayrimenkul toplu değerlemesinde iyi sonuçlar verdiğini yaptıkları çalışmalarla gözlemlemişlerdir (Ahmed S vd. 2014, Lam vd. 2008, McCluskey 1996, Morano ve Tajani 2013, Musa vd. 2013, Özkan vd. 2007, Sampathkumar vd. 2015, Saraç 2012, Selim 2009, Tabales vd. 2013, Tay ve Ho 1992, Varma vd. 2018, Wilson vd. 2002, Xin ve Runeson 2004). Buna karşın bazı araştırmacılar ise YSA algoritmasını regresyon yöntemi ile kıyaslamış ve YSA algoritmasının toplu değerleme sonucuna çok fazla bir katkısı olmadığını çalışmalarında gözlemlemişlerdir (Lenk vd. 1997, McCluskey 1996, Worzala vd. 1995). Literatürden anlaşılacağı üzere, YSA algoritmasının toplu değerlemede iyi sonuç verdiğini gözlemleyen araştırmacılar çoğunlukta olsa da iyi sonuç vermediğini söyleyen araştırmacılar da bulunmaktadır. Bunun nedenin kullanılan değişken ve gayrimenkul türü ile mi diğer bir ifadeyle kullanılan veri ile mi ilişkili olup olmadığı literatürde bilinmeyen bir boşluk olarak yer almaktadır.

YSA algoritmasının genellikle toplu değerlemede iyi sonuçlar verdiğini gören bazı araştırmacılar, YSA dışında diğer makine öğrenme algoritmaları da toplu değerleme çalışmaları için son yıllarda denemeye başlamışlardır. Örneğin, 26 değişkenli apartman daireleri için alım-satım işlem kayıtlarından 2006-2017 yılları arasında 16.601

gayrimenkul verileri kullanılmış ve Rasgele Orman (RO) algoritmasının (Breiman 2001) hedonik modeller için yararlı bir tamamlayıcı olabileceği bulunmuştur (Hong vd. 2020). Bu araştırmada sadece apartman dairesi kullanıldığı ve iki algoritmanın karşılaştırıldığı görülmektedir. Sawant vd. (2018) 29,680 satış ve 55 değişken gibi yüksek sayıda bir veri ile RO algoritması kullanarak toplu değerleme yapmış ve RO algoritmasının değerleme işlemlerinde çok iyi sonuçlar verdiğini gözlemlemiş ve bu çalışmada yine tek tip makine öğrenme algoritması kullanıldığı görülmektedir. Başka bir çalışmada ise, bulanık mantık tekniği kullanılarak nominal değerleme sonucunda arazi değer haritaları üretilmiş ve elde ettikleri sonuçlarına göre araştırmacılar bulanık mantık tekniğinin toplu değerleme çalışmaları için uygun olduğunu savunmuşlardır (Derinpınar ve Aydınoğlu 2015). Ravikumar (2017) ise RO, Destek Vektör Makineleri (DVM), YSA ve çoklu regresyon yöntemlerini Amerika Birleşik Devletleri (ABD) emlak satış sitelerinden topladığı 49,980 adet gayrimenkul değerlemesinde kullanmış ve RO algoritmasının DVM ve YSA'ya oranla daha iyi sonuç verdiğini elde ettiği sonuçlarla gözlemlemiştir. Bu çalışmada değişken sayısı 10 adette sınırlı kalmıştır. Afonso vd. (2019) ise 2015'ten 2018'e kadar Brezilya emlak web sitelerinden 1,223,582 ilan toplamış ve bu gayrimenkullere ait 24 değişken ile RO ve Tekrarlayan Sinir Ağları (YYSA, Recurrent Neural Network) algoritmalarını karşılaştırmıştır. Dellstad (2018) RO, YSA ve DVM algoritmalarını toplu değerlemede kullanmıştır; araştırmacı 57,974. veri ve 44 değişken ile İsviçre'de bir çalışma yapmış ve RO algoritmasının en iyi sonuç verdiğini gözlemlemiştir. Bu çalışmada tek tip gayrimenkul ile 3 adet makine öğrenme algoritması kıyaslanmıştır, fakat diğer makine öğrenme algoritmalarının bir arada farklı gayrimenkul tiplerinde sonuçlarının kıyaslanmasına literatürde rastlanmamıştır.

Çizelge 1.1 Gayrimenkul toplu değerleme çalışmaları ve özeti

No	Çalışma Alanı	Yöntem	Gayrimenkul Sayısı	Değişken Sayısı	Sonuç
1	"A neural network approach to residential property appraisal" (Do ve Grudnitski 1992)				
	San Dioge	Sinir Ağları (SA) ve çoklu regresyon	163	9	%2,73 Çoklu regresyon ve %1.31 SA için hata bulunmuş, SA kalsik doğrusal regresyona göre daha iyi sonucuna varılmıştır.
2	"Artificial neural network in property valuation: application framework and research trend" (Abidoye ve Chan 2016)				
	-	-	-	-	Çalışmada 31 YSA ve toplu değerleme ile ilgili makaleler incelenmiştir.
3	"Artificial neural networks: the next modelling/calibration technology for the assessment community" (Borst 1991)				
	Amerika, New England	4 farklı YSA ağ mimarisi kullanılmış	-	-	Gayrimenkul toplu değerlemede YSA'yı ilk kullanan çalışma, YSA'nın gayrimenkul değerlemede güvenilir olduğu sonucuna varılmıştır.
4	"Predicting property price index using artificial intelligence techniques evidence from Hong Kong" (Abidoye vd. 2019)				
	Hong Kong	YSA, DVM, ARIMA (Autoregressive integrated moving average)	127	11	YSA, DVM ve ARIMA'ya göre daha iyi sonuç vermiştir.
5	"House rent estimation in Dhaka city by multi-layer perceptions neural network" (Samsuddin Ahmed vd. 2014)				
	Dhaka	YSA	100	40	Ev kiraları YSA ile tespit edilmiş ve kira belirlemede YSA'nın iyi olduğu bulunmuştur.
6	"Real estate price range prediction using artificial neural network and grey wolf optimizer" (Al-Gbury ve Kurnaz 2020).				
	Sindian Bölgesi, New Taipei	Greywolf optimizer ve YSA	1,2 milyon	-	%98 oranında iyi bir performans elde edilmiştir.

Çizelge 1.1 Gayrimenkul toplu değerleme çalışmaları ve özeti (devam)

7	"Performance of multiple linear regression and nonlinear neural networks and fuzzy logic techniques in modelling house prices" (Amri ve Tularam 2012).				
	Bathurst şehri, Avusturalya	Çoklu Doğrusal Regresyon, YSA ve Bulanık mantık	7,849	13	SA çoklu regresyondan daha iyi sonuç vermiştir.
8	"Combining BP with PSO algorithms in weights optimisation and ANNs training for mass appraisal of properties"(Yacim ve Boshoff 2018).				
	Cape Town, Güney Afrika	YSA ve farklı regresyon türleri	3,242	148	Semi-log, log-log yakın hatalar verirken YSA ve lineer regresyon daha iyi sonuç vermiştir.
9	"Comparison of different machine learning models for mass appraisal of real estate" (Bilgilioğlu ve Yılmaz 2021).				
	Aksaray	YSA, DVM, ACART, RO ve CHAID	1,982	65	Hepsi benzer sonuçlar olsa da YSA daha iyi sonuç vermiştir.
10	"Real estate value at porto alegre city using artificial neural networks" (Cechin vd. 2000).				
	Porto Alegre şehri	Lineer regresyon ve YSA	1,600	20	YSA lineer regresyona göre 3 kat daha iyi sonuç vermiştir.
11	"Advanced property valuation techniques and valuation accuracy: deciphering the artificial neural network technique" (Abidoye vd. 2019)				
	-	-	-	-	Literatür taraması benzeri bir makale ve yapılan çalışmalarda gayrimenkul değerlemesinde YSA'nın iyi sonuçlar çıkardığı görülmüştür.
12	"Artificial intelligence and the mass appraisal of residential apartments" (Tay ve Ho 1992).				
	Singapur	Çoklu Regresyon ve YSA	1,055	10	YSA çoklu regresyona göre daha iyi sonuç çıkarmıştır.
13	"Comparing three machine learning algorithms in the task of appraising commercial real estate" (Dellstad 2018).				
	İsveç	DVM, YSA ve RO	57,974	44	RO diğerlerine göre daha iyi sonuç çıkarmıştır.

Çizelge 1.1 Gayrimenkul toplu değerleme çalışmaları ve özeti (devam)

14	"Daire tipi konutların değerlemesinde yapay sinir ağlarının kullanımı" (Demirel vd. 2016).				
	Talas, Kayseri	YSA	570	12	YSA toplu değerlemede başarılı olduğu görülmüştür.
15	"A Proposal of a mass appraisal system in greece with CAMA system: evaluating GWR and MRA techniques in thessaloniki municipality" (Dimopoulos ve Moulas 2016).				
	Thessaloniki	Çoklu Regresyon ve YSA	2,583	19	YSA toplu değerlemede başarılı çıktığı gözlemlenmiştir.
16	"Modeling property prices using neural network model for Hong Kong" (Ge vd. 2004).				
	Hong Kong	YSA	-	-	Sonuç ile ilgili değişkenler YSA için başarı kriteri olduğu bulunmuştur.
17	"Analyzing massive data sets: an adaptive fuzzy neural approach for prediction, with a real estate illustration" (Guan vd. 2014).				
	Midwestern	5 tabakalı YSA ya benzeyen bir yöntem	16,523	24	YSA'nın farklı modeli olan ANFIS çoklu regresyondan daha iyi sonuç vermiştir.
18	"Effectiveness comparison of the residential property mass appraisal methodologies in the USA" (Lin ve Chen 2011)				
	Amherst kasabası	ÇRA, YSA ve non-parametric regresyon	33,342	82	YSA ve regresyon çeşitleri toplu değerlemede iyi sonuçlar vermiştir.
19	"Analysis of the mass appraisal model by using artificial neural network in kaohsiung city" (Pi-ying 2011).				
	Kaohsiung	Çoklu regresyon, YSA	2,471	9	YSA çoklu regresyona göre daha iyi sonuç çıkarmıştır.
20	"An artificial neural network and entropy model for residential property price forecasting in Hong Kong" (Lam vd. 2009).				
	Hong Kong	Entropy ve YSA birleşimi	4,143	29	Hibrit modelin daha iyi sonuç verdiği gözlemlenmiştir.
21	"High-tech valuation: should artificial neural networks bypass the human valuer?" (Lenk vd. 1997).				
	Fort Collins	Hedonik ve YSA	271	7	YSA ile Hedonik arasında çok fazla fark olmadığı gözlemlenmiştir.

Çizelge 1.1 Gayrimenkul toplu değerleme çalışmaları ve özeti (devam)

22	"House price prediction: hedonic price model vs. artificial neural network" (Limsombunchao 2004).				
	Christchurch, Yeni Zelenda	Hedonik ve YSA	200	8	YSA yöntemi hedonik yönteme göre daha iyi sonuç vermiştir.
23	"The potential of artificial neural networks in mass appraisal: the case revisited" (McCluskey vd. 2012).				
	Lisburn, Kuzey İrlanda	YSA ve birkaç farklı çoklu regresyon	2,694	7	YSA ve regresyon türleri arasında sonuçlarda çok fazla far olmadı bulunmuştur.
24	"Artificial ıntelligence applied to real estate valuation" (Mora-Esperanza 2004).				
	Madrid	YSA	1,193,000	12	YSA iyi sonuç vermiştir.
25	"Bare ownership evaluation. Hedonic price models vs. Artificial neural network" (Morano ve Tajani 2013).				
	Milan	Hedonik ve YSA	85	7	Hedonik yöntemin YSA ya göre toplu değerlemede daha iyi olduğu görülmüştür.
26	"A Neural-CBR system for real property valuation" (Musa vd. 2013).				
	Benin şehri, Nijerya	YSA, CBR (Case Based Reasoning)	30,000	15	YSA CBR'den daha iyi sonuç vermiştir.
27	"Neural networks: the prediction of residential values" (McGreal vd. 1998).				
	Belfast	YSA	1,026	33	Sonuçların %80 oranında gerçek değere yakın çıktığı görülmüştür.
28	"Predicting housing value: a comparison of multiple regression analysis and artificial neural networks" (Nghiep ve Al 2001).				
	Rutherford kasabası, Tennessee	YSA ve çoklu regresyon	3,906	-	YSA daha iyi sonuç veriyor ve özellikle non-lineer durumlarda daha iyi çalıştığı bulunmuştur.
29	"An investigation on the price estimation of resizable real-estates by using artificial neural network and regression methods" (Özkan vd. 2007).				
	Selçuklu, Konya	YSA ve regresyon teknikleri	170	7	YSA'nın %85 ve regresyonun %83 başarısı bulunmuştur.

Çizelge 1.1 Gayrimenkul toplu değerleme çalışmaları ve özeti (devam)

30	"Neural network hedonic pricing models in mass real estate appraisal" (Peterson ve Flanagan 2009).				
	Wake Kasabası, NC	YSA ve regresyon teknikleri	46,467	-	YSA daha iyi sonuç verdiği görülmüştür.
31	"Improving the results of artificial neural network models for residential valuation" (Rossini 1998).				
	Adelaide ve GüneyAvusturalya	Genetik optimizasyon ve YSA hibrityöntemi	200	30	Genetik optimizasyon YSA sonuçlarını ciddi derecede iyileştirdiği bulunmuştur.
32	"Forecasting the land price using statistical and neural network software" (Sampathkumar vd. 2015).				
	Chennai metropolü	SA ve çoklu regresyon	-	11	SA regresyona göre biraz daha iyi sonuç vermiştir.
33	"Determinants of house prices in Turkey: Hedonic regression versus artificial neural network" (Selim 2009).				
	-	Hedonik ve YSA	5,741	46	Türkiye içim YSA toplu değerlemede kullanılabilir sonucuna varılmıştır.
34	"Artificial neural networks for predicting real estate prices" (Tabales vd. 2013).				
	Kuzey İspanya	YSA	130,563	15	Büyük veri setlerinde YSA iyi sonuçlar vermektedir.
35	"An exploration of neural networks and its application to real estate valuation" (Worzala vd. 1995).				
	Fort collins	SA ve çoklu regresyon	288	7	YSA daha iyi çalışsa da regresyon ile aralarında çok büyük fark yok olduğu görülmüştür.
36	"A comparison of regression and artificial intelligence methods in a mass appraisal context" (Zurada vd. 2011).				
	Lousiville, Kentucky	Çoklu regresyon, SA, M5P trees, DVM ve RBFNN (farklı bir SA)	222,000	143	Homojen veride regresyon heterojen veride ise diğer yöntemler iyi sonuç vermektedir.
37	"Yapay sinir ağları metodu ile gayrimenkul değerleme" (Saraç 2012).				
	İstanbul (33 farklı ilçe)	YSA	4,000	12	%10-15 civarında doğrulukla kabul edilir seviyede sonuçlar elde edilmiştir.

Çizelge 1.1 Gayrimenkul toplu değerleme çalışmaları ve özeti (devam)

38	"Artificial Neural Network, Quantile and Semi-Log Regression Modelling of Mass Appraisal in Housing" (Torres-Pruñonosa vd. 2021).				
	Katalonya	YSA ve hedonik reg.	188,652	12	YSA sonuçları başarılı bulunmuştur.
39	"YSA ve DVM yöntemleri ile taşınmaz değerlemesi için bir yaklaşım geliştirme" (Bulut Nas 2011)				
	Selçuklu, Konya	YSA, DVM ve ÇRA	205	14	Toplu değerlemede en başarılı DVM, sonra YSA ve regresyon bulunmuştur.
40	"Estimating the performance of random forest versus multiple regression for predicting prices of the apartments" (Čeh vd. 2018).				
	ljubljana, Slovenya	RO, Basit Doğrusal Regresyon, TBA	7,407	35	RO regresyona göre daha iyi sonuç vermiştir, TBA çoklu kollinear sorunu için kullanılmıştır.
41	"Mass appraisal of residential apartments: An application of Random forest for valuation and a CART-based approach for model diagnostics" (Antipov ve Pokryshevskaya 2012).				
	Saint Petersburg	RO, CART ve KNN	2,848	17	RO algoritması CART ve KNN e göre daha iyi sonuç vermiştir.
42	"Housing prices prediction with a deep learning and random forest ensemble" (Afonso vd. 2019).				
	Brezilya	RO ve RNN (Recurrent Neural Networks)	1,223,582	19	RO daha iyi sonuç verdiği görülmüştür, bu yoğunluktaki veri Brezilya emlak web sitelerinden çekilmiştir.
43	"Investigation of random subspace and random forest methods applied to property valuation data" (Lasota vd. 2011).				
	-	RO	5,200	-	RO gayrimenkul değerleme çalışmalarında kullanılabilir olduğu sonucuna varılmıştır.
44	"Variable selection for hedonic model using machine learning approaches: A case study in Onondaga, NY" (Yoo vd. 2012).				
	Onondaga kasabası	RO ve Kubist regresyon	4,469	32	RO algoritması regresyona göre daha başarılıdır.
45	"Using SVR and MRA Methods for Real Estate Valuation in the Smart Cities" (Akar ve Yalpır 2021).				
	Konya	DVR ve ÇRA	319	31	DVR R^2 0.88 Regresyonda ise 0.86 çıkmıştır iki yöntem yakın sonuçlar vermiştir.

Çizelge 1.1 Gayrimenkul toplu değerleme çalışmaları ve özeti (devam)

46	"A SVR based approach on real-estate valuation" (Bulut vd. 2011).				
	Konya	DVR	190	5	Değişken sayısı azaltıldıkça DVR daha yüksek doğrulukta sonuç çıkarmıştır.
47	"The mass appraisal of the real estate by computational intelligence" (Kontrimas ve Verikas 2011).				
	Litvanya	Regresyon ve DVM	100	12	DVM'nin regresyondan daha iyi olduğu görülmüştür.
48	"Support vector machine and entropy based decision support system for property valuation" (Lam vd. 2009).				
	Hong Kong	DVM, Regresyon ve YSA	4,143	29	DVM diğer algoritmalara göre daha iyi sonuç vermiştir.
49	"Prediction of House Unit Price in Taipei City Using Support Vector Regression" (Lee ve Chen 2008).				
	Taipei	DVR ve ANFIS (Adaptive Network Based Fuszzy)	3,336	7	DVR algoritmasını gayrimenkul toplu değerleme çalışmaları için önermişlerdir.
50	"Makine Öğrenme Algoritmalarının Konut Değer Tahmininde Kullanımı: Ankara Gölbaşı Uygulaması" (Savaş 2019).				
	Gölbaşı, Ankara	DVR ve ÇRA	95	42	DVR gayrimenkul değerlemede kullanılabilir olduğu görülüyor
51	"Neural networks support vector machine for mass appraisal of properties" (Yacim ve Boshoff 2020).				
	Cape Town, Güney Afrika	YSA, SVM ve hibrit yöntem	3,325	10	DVM'nin YSA'ya göre biraz daha iyi sonuç verdiği görülmüştür.
52	"Predicting house sale price using fuzzy logic, Artificial Neural Network and K-Nearest Neighbor" (Mukhlishin vd. 2017).				
	-	Bulanık mantık, YSA ve k en yakın komşuluk	-	5	Bulanık mantık algoritması YSA ve k en yakın komşuluk algoritmalarına göre daha iyi çıkmıştır.
53	"Data Mining Techniques for Predicting Real Estate Trends" (Vargason 2019).				
	Salamanca, Madrid	RO, DVM ve k en yakın komşuluk	2,266	70	K en yakın komşuluk algoritması en başarısız sonucu vermiştir.

Çizelge 1.1 Gayrimenkul toplu değerleme çalışmaları ve özeti (devam)

54	"A Recommendation Engine to Estimate Housing Values in Real Estate Property Market" (Ziweritin vd. 2021).				
	Kaggle verisi	K en yakın komşuluk ve logistik regresyon	21,613	6	Gayrimenkul değerleri 3 sınıfa ayrılarak sınıflandırma yapılmış ve k-nn daha iyi sonuç verdiği görülmüştür.
55	"Analysis of different approaches to real estate appraisal. Surgelas" (Surgelas vd. 2019).				
	Riga Latvia	TBA	37	9	TBA gayrimenkulün değerini etkileyen değişken sayısı 4' e indirilmiştir. Daha iyi değerleme sonucu elde edilmiştir.
56	"Reduction of Mass Appraisal Criteria with Principal Component Analysis and Integration to GIS"				
	Konya	TBA, ÇRA	558	41	TBA ile 41 kriter 12 ye indirilmiştir. TBA'nın değerleme kalitesine çok katkısı olmamıştır.

Literatür özet tablosu detaylı olarak incelendiğinde literatürde toplu değerleme çalışmalarında en çok kullanılan MÖ algoritmasının YSA ve ÇRA olduğu ve bunu RO algoritmasının takip ettiği görülmektedir. Gayrimenkul toplu değerleme işlemlerini kullanan yaygın ülke örneklerine bakıldığında genellikle vergi değeri belirleme ve değerleme amaçlı vb. işlemlerde kullanılmak üzere toplu değerlemeyi profesyonel bir şekilde kullanıldığı görülmektedir. Konu ülkelerden en dikkat çekenleri, toplu değerleme yapan kurumları ve hangi yöntemin kullanıldığına Çizelge 1.2'de yer verilmiştir (McCluskey 1996, Walacik vd. 2013).

Çizelge 1.2 Toplu değerleme kullanılan ülkeler, kurumları ve yöntemler

Ülke	Kurum	Yöntem
Avusturalya	Arazi Kurumu	Çoklu regresyon analizi
İsveç	Ulusal Vergi Kurumu	Çoklu regresyon analizi
Kuzey İrlanda	Değerleme ve Arazi Kuruluşu	Çoklu regresyon analizi, YSA ve CSA (Comparable Sale Analysis-Karşılaştırmalı Satış Analizi)
Yeni Zelenda	Yeni Zelenda Değerleme Bürosu	Çoklu regresyon analizi
Singapur	Singapur Değerleme Kurumu	Indexation (Indeaxtion of Existing Assessment- Mevcut Değerleme Endekslenmesi)
Hong Kong	Derecelendirme ve Değerleme Kurumu	Çoklu regresyon analizi ve Indexation (Indeaxtion of Existing Assessment- Mevcut Değerleme Endekslenmesi)
Malezya	Değerleme Bölümü	Expert (Uzman Sistemi)
A.B. D	Değerleme Ofisi	Indexation (Indeaxtion of Existing Assessment- Mevcut Değerleme Endekslenmesi), AEP (Adaptive Estimation Procedure, Uyarlanabilir Tahmin İşlemi) ve Çoklu regresyon analizi, YSA ve CSA (Comparable Sale Analysis- Karşılaştırmalı Satış Analizi)
İngiliz Kolombiyası	BC Değerleme Otoritesi kurumu	Çoklu regresyon analizi
İngiltere	Değerleme Ofisi	Basitleştirilmiş değerleme yaklaşımları
Almanya	-	En küçük kareler yöntemi ile regresyon analizi
Slovenya	Değerleme Ofisi	-
Litvanya	Kamu İhtiyaçları Değerleme Dairesi Başkanlığı	Karşılaştırmalı satış analizi

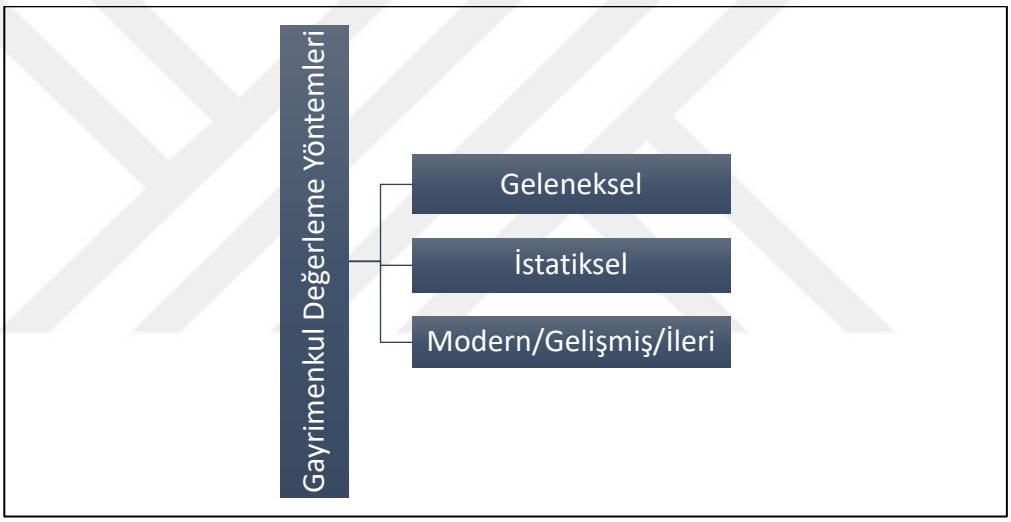
Tez çalışmasını literatürde bulunan gayrimenkul toplu değerleme çalışmalarından ayıran en önemli özellik ise gayrimenkul türü olarak 4 farklı; arsa, arazi, ticari ve konut

verilerinin toplu değerleme amaçlı test edilmesi ve 5 farklı MÖ algoritmasının bu veriler üzerinde denenmesi ve kıyaslanarak en başarılı algoritmanın bulunması ile kapsamlı bir analiz yapılması ve Türkiye için gayrimenkulde toplu değerleme çalışmalarının yapılabilirliğinin araştırılmasıdır. Diğer taraftan yine literatür incelendiğinde, konut türü gayrimenkul veri sayısı genellikle 3-4 haneli olmakla beraber bu tez çalışmasında kullanılan konut türünde yaklaşık 1.4 milyon adet gayrimenkul verisinin kullanılması bu tez çalışmasına özgünlük katan diğer bir nokta olarak dikkat çekmektedir.

Sonuç olarak, literatür özetinden anlaşılacağı üzere, gayrimenkul toplu değerlemede makine öğrenme algoritmaları 1991 yılından beri kullanılmakta olup en çok YSA algoritması çalışmalara konu olmuş ve çoğunlukta araştırmacılar YSA'nın toplu değerleme için iyi sonuçlar verdiğini bazıları ise tam tersini iddia etmiştir. Literatürde genellikle 1-4 arası adet makine öğrenme algoritmaları ve tek tip gayrimenkul çeşidi (genellikle konut) üzerine toplu değerleme çalışmaları yapılmış, fakat geniş kapsamlı birçok makine öğrenme algoritmasını farklı gayrimenkul tipleri (niteliği) için karşılaştıran bir çalışmaya literatürde rastlanmamıştır. Bu durumlar göz önüne alındığında, bu çalışmanın amacı MÖ algoritması (5 adet) kullanılarak farklı gayrimenkul tiplerinde (4 adet) birçok değişken ile test edilmesi ve algoritmaların avantaj ve dezavantajlarının ortaya konarak literatüre kapsamlı bir katkı sağlanmasıdır. Özetle, bu tez çalışmasının ana hedefi, YSA, DVM, RO, ÇRA, k-nn gibi 5 adet farklı algoritma ile literatürde kullanılan tek tip gayrimenkul yerine ticari, konut, arsa ve arazi gibi 4 farklı gayrimenkul tipleri için uygulamak ve sonuçları karşılaştırılarak hangi gayrimenkul niteliğinde hangi makine öğrenme algoritmasının daha iyi sonuç verdiğini kıyaslayarak literatüre kapsamlı bir katkıda bulunmaktır.

2. KİTLESEL DEĞERLEME

Gayrimenkul değerlemesi, bir gayrimenkulün gerçekte ne kadar değerli olduğunu belirleme sürecidir. Özellikle Türkiye'de gayrimenkul değerlemesi halen gayrimenkul bazında tekil olarak yapılmakta, bazı ülkelerde ise gayrimenkullerin bir arada bir model dahilinde toplu değerleme işlemlerini uzun yıllardır gerçekleştirildiği görülmektedir. Gayrimenkul değerlemede farklı yöntemler yer almakta olup ve bu yöntemler geleneksel, istatiksel ve modern/gelişmiş yöntemler olarak gibi üçe ayrılabilmektedir (Şekil 2.1) (Erdem 2017b). Yaygın olarak uygulanan geleneksel yöntemler, karşılaştırmalı satış analizi, gelir yaklaşımı ve maliyet yöntemi yaklaşımıdır.



Şekil 2.1 Gayrimenkul değerleme yöntemleri

Gayrimenkul değerleme yöntemlerinden klasik yöntemler genellikle tekil değerleme işlemlerinde kullanılırken, istatiksel ve modern yöntemler ise daha çok gayrimenkul toplu değerleme işlemlerinde kullanıldığı toplu değerleme yönteminin ise birçok ülkede gayrimenkullerin genellikle vergilendirmesi amaçlı kullanıldığı dikkati çekmektedir.

Gayrimenkullerin toplu değerlemesi için IAOO (International Association of Assessing Officers, Uluslararası Değerleme Görevlileri Birliği) standardı bulunmaktadır. IAOO tarafından Gayrimenkul Toplu Değerleme Standardı (Standard on Mass Appraisal of Real Property) 2017 Temmuz'da onaylanmıştır. Bu standart, Ocak 2012 Gayrimenkullerin

Toplu Değerlendirilmesi Standardının yerini almış ve tam bir revizyon olarak dikkati çekmektedir.

IAOO standardı 14 sayfa, 7 bölümden oluşmakta olup bu bölümlerin başlıkları: Amaç, Giriş, Gayrimenkul Verilerin Toplanması ve Saklanması, Değerleme Yöntemleri, Model Testi, Kalite Güvencesi ve Değer Savunması, Yönetimsel ve Mekânsal Hususlar, Referans Materyaller olarak sıralanmaktadır

2.1 Kitlesel Değerlemede Amaç

Birincil odak, ad valorem[1] vergi amaçları için kitlesel (toplu) değerlendirmedir. Ancak burada tanımlanan ilkeler, ipotek portföy yönetimi gibi başka amaçlar için kullanılan CAMA'lar (veya otomatik değerleme modelleri) ile de ilgili olmalıdır. Standart öncelikle değerlendiricinin, değerlendirme gözetim kurumlarının ve vergi mükelleflerinin ihtiyaçlarını ele almaktadır.

Bu standart, değere ilişkin üç geleneksel yaklaşımın (maliyet, satış karşılaştırması ve gelir) toplu değerleme uygulaması da dahil olmak üzere, mülkteki ücret basit faizinin piyasa değerinde değerlendirilebileceği toplu değerleme prosedürlerini ele almaktadır. Tek gayrimenkul değerlemeleri, kısmi faiz değerlemeleri ve piyasa değeri dışında yapılan değerlemeler bu standardın kapsamı dışındadır. Bu standart, kullanım değeri, sınıflandırma veya değerlendirme artış sınırlamaları gibi yasal kısıtlamalar nedeniyle piyasa değerinden farklı değerlendirilen değerlerin belirlenmesi konusunda rehberlik sağlamayacaktır.

Değerleme amaçlarına yönelik piyasa değeri, genellikle toplu değerleme tekniklerinin uygulanmasıyla belirlenecektir. Toplu değerleme, belirli bir tarih itibariyle ortak veriler, standart yöntemler ve istatistiksel testler kullanılarak bir grup gayrimenkulün değerleme süreci olarak görülmektedir.

[1] advalorem sistem kıymet/değer yani parasal değerler üzerinden vergi matrahının hesaplanması anlamını taşımaktadır.

2.1.1 Kitlesel değerlemede verilerin toplanması ve saklanması

Değerlerin doğruluğu, her şeyden önce gayrimenkulün özelliklerinin ve piyasa verilerinin ulaşılabilirliğine ve doğruluğuna bağlıdır. Gayrimenkulün doğru değerlemesi, doğru, eksiksiz ve güncel gayrimenkul verileri gerektirmektedir (yani, gayrimenkul sahipliği, konum, büyüklük, kullanım, fiziksel özellikler, satış fiyatı, kiralar, maliyetler ve işletme giderleri).

2.1.2 Coğrafi (mekânsal) veri toplanması ve saklanması

Değerleme uzmanları, her bir parsel için benzersiz bir kimlik numarası ile tüm alanını kapsayan doğru, güncel kadastro haritaları (değerleme haritaları, vergi haritaları, parsel sınır haritaları ve gayrimenkul sahipliği haritaları olarak da bilinir) bulundurmalıdır. Bu tür kadastro haritaları, değerleme görevlilerinin hem arazide hem de ofiste tüm parselleri tanımlamasına ve konumlandırmasına izin vermektedir. Haritalar, bir coğrafi bilgi sistemi (CBS) kullanıldığında, toplu değerlendirme sürecinde özellikle değerli hale gelmekte, CBS, satış fiyatlarının, değerlerin, denetim tarihlerinin, iş atamalarının, arazi kullanımlarının ve çok daha fazlasının grafiksel gösterimine izin vermektedir.

Değerleme uzmanı sınıflandırma, değerleme ve diğer amaçlar için yeterli gayrimenkul özellikleri verilerini toplamalı ve muhafaza etmelidir. Herhangi bir yöntemle gayrimenkulün doğru değerlemesi, arazi ve bina özelliklerinin tanımlanmasını gerektirir.

Toplanacak ve saklanacak gayrimenkul özellikleri:

- Söz konusu yerel pazarı etkileyen faktörler
- Kullanılacak değerleme yöntemlerinin gerektirdiği veriler
- Sınıflandırma gereklilikleri ve emlak vergisi politikası
- Diğer resmi ve özel kullanıcıların gereksinimleri
- Her bir gayrimenkul özelliğinin toplanması ve sürdürülmesinin marjinal faydaları ve maliyetleri

Şeklinde olmalıdır.

Bir CAMA sistemi için gayrimenkul özellikleriyle ilgili hangi verilerin toplanacağını ve korunacağını belirlemek, uzun vadeli sonuçları olan çok önemli bir karar olarak görülmektedir. Mevcut verilerin niceliği ve kalitesi gözden geçirilmeli, veriler seyrek ve güvenilmez ise, yeniden tarama gerekmektedir. Mümkün olduğunca güvenilirliği onaylanmış veriler kullanılmasına dikkat edilmelidir.

Aşağıdaki gayrimenkul özellikleri, konut türü gayrimenkul değerlerini tahmin etmede genellikle önemlidir:

- Geliştirme verileri, yaşam alanı, yapı kalitesi veya ana bileşenleri (temel, dış duvar tipi vb.), etkin yaş veya durum, bina tasarımı veya stili, bodrumlar, garajlar, kapalı sundurmalar ve balkonlar dahil ikincil alanlar, banyo ve merkezi klima gibi bina özellikleri, konuk evleri, kayıkhaneler ve ahırlar gibi önemli müstakil yapılar

Arsa/Arazi Verileri

- Gayrimenkul alanı ve mevcut durum (kanalizasyon, su, elektrik)

Mekânsal Veri

- Market/pazar alanı, küçük market/pazar alanı veya mahalle, site olanakları, özellikle manzara ve golf sahası veya su cephesi, olumsuz özellikler, (örneğin, yoğun trafik, havaalanı gürültüsü veya ticari kullanımlara yakınlık).

IAOO standartlarına göre, özellikleri konut türü gayrimenkul değerini tahmin etmede genellikle önemli görülmektedir

Gayrimenkul özellikleri verilerinin toplanması, yeniden değerlendirmenin kritik ve pahalı bir aşaması olarak dikkati çekmektedir. Başarılı bir veri toplama programı, net ve standart kodlamayı ve bir kalite kontrol programı aracılığıyla dikkatli izlemeyi gerektirmektedir. Bir veri toplama kılavuzunun geliştirilmesi ve kullanılması, doğru ve tutarlı veri toplamanın sağlanması için esas olup, veri toplama programı, eksiksiz ve doğru verilerle sonuçlanmalıdır.

İlk veri toplama, veri toplama formatı gibi alt başlıklar IAOO standartta yer verilmiştir. İlk veri toplama için fiziksel incelemenin gerekli olduğu vurgulanmıştır. Veriler hem sahada veri toplamayı hem de verinin bilgisayar sistemine girişini kolaylaştıracak şekilde tasarlanmış önceden belirlenmiş bir formatta toplanmalıdır.

Veri toplama için aşağıdaki doğruluk standartları önerilir:

- Yaşam alanı ve dış duvar yüksekliği gibi sürekli veya alan ölçüm verileri, gerçek boyutların 1 fit (0.3048 m.) veya alanın yüzde 5'i dahilinde doğru olmalıdır. (Bir fit, metrik sistemde yaklaşık 30 santimetreye eşittir.)
- Toplanacak veya doğrulanacak her bir hedef, kategorik veya sayısal veri alanı için kodlanmış girişlerin en az yüzde 95'i doğru olmalıdır. Nesnel, kategorik veya sayısal veri özellikleri, dış duvar malzemesi, eksiksiz banyo sayısı ve su kenarı görünümü gibi nitelikleri içermektedir. Örnek olarak, bir veri toplayıcı 100 özellik için 10 nesnel, kategorik veya ikili veri öğesi yakalarsa, toplam 1000 girişin en az 950'si doğru olmalıdır.
- Toplanan veya doğrulanan her bir öznel kategorik veri alanı için verilerin en az yüzde 90'ında doğru şekilde kodlanmalıdır.
- Belirli doğruluk gereksinimlerinden bağımsız olarak, tutarlı ölçüm önemlidir. Tutarlı ölçümü desteklemek için ulusal, yerel ve bölgesel uygulamaları içeren standartlar mevcuttur. Ölçüm standardı, sürecin bir parçası olarak belgelenmelidir.

Veri doğruluğu standartlarına ulaşıldığından ve sürdürülebilirliğinden emin olmak için bir kalite kontrol programı gereklidir. Bağımsız kalite kontrol denetimleri, veri toplama aşaması başladıktan hemen sonra yapılmalı ve yetkili personel, proje danışmanları, denetim firmaları veya gözetim kurumları tarafından gerçekleştirilebilir. Denetimler, tamamlanmış işlerin rastgele örneklerini eksiksizlik ve doğruluk açısından incelemeli ve doğruluk standartlarına ulaşılıp ulaşılmadığını belirlemek için istatistiksel testlerin tablolarını tutmalıdır. Coğrafi alana, gayrimenkul türüne veya bireysel veri toplayıcıya

göre sınıflandırma, veri hatası kalıplarının tespit edilmesine yardımcı olabilecektir. Kalite kontrol standartlarını karşılamayan veriler yeniden toplanmalıdır.

Tekrar/aynı veri girişini önlemek için, veri toplama formu, veri giriş formu işlevi görebilmelidir. Veri hatası örnekleri, eksik veri kodlarını ve geçersiz karakterleri içerir. Veri değerleri normal aralıkları aştığında da uyarı mesajları oluşturulmalıdır.

2.1.3 Periyodik yerinde denetimlere alternatif

İlk fiziksel denetimlerin zamanında tamamlanması ve etkili bir inşaat izinleri sisteminin veya fiziksel değişiklikleri rutin olarak belirlemeye yönelik diğer yöntemlerin mevcut olması koşuluyla, yetki alanları, bilgisayar destekli bir ofis incelemesi ile saha yeniden denetimlerini desteklemek için bir dizi dijital görüntüleme teknolojisi aracı kullanabilir. Bu görüntüleme araçları:

- Mevcut yüksek çözünürlüklü sokak görüntüleri (kalite derecesinin ve fiziksel durumun doğrulanmasını sağlayan bir inç altı piksel çözünürlüğünde)
- Ortofoto görüntüleri (kentsel/banliyölerde minimum 6 inç piksel çözünürlük ve kırsal alanlarda 12 inç çözünürlük, hızlı büyüyen alanlarda 2 yılda bir veya yavaş büyüyen alanlarda 6-10 yılda bir güncellenir)
- Ölçüm doğrulaması için kullanılabilen düşük seviyeli oblik görüntüleri (dört ana yön, kentsel/banliyölerde minimum 6 inç piksel çözünürlük ve kırsal alanlarda 12 inç piksel çözünürlük, hızlı büyüyen alanlarda 2 yılda bir güncellenir veya 6, yavaş büyüyen alanlarda 10 yıl).
- CAMA sistemindeki bina boyut verilerini coğrafi referanslı görüntülerle veya kaynaklardan (LiDAR [ışın algılama ve mesafe aralığı] gibi) gelen uzaktan algılama verileriyle karşılaştıran değişiklik algılama tekniklerini içerebilir.

Ek olarak, değerleme uzmanları, mahalle durumu, eğilimleri ve gayrimenkul özelliklerindeki değişiklikleri gözlemlemek için yıllık bazda atanan alanları ziyaret etmeleri gerektiği vurgulanmalıdır. Önemli inşaat değişiklikleri tespit edildiğinde, bir

gayrimenkul satıldığında veya bir alan katastrofik hasardan etkilendiğinde yerinde fiziksel inceleme önerilmektedir. İnşaat izinleri düzenli olarak izlenmeli ve önemli değişiklik gösteren gayrimenkuller, çalışma tamamlandığında denetlenme gerektiği vurgulanmalıdır

2.1.4 Değerleme yöntemleri ve yeniden değerleme sıklığı

İster tekil değer belirleme ister kitlesel (toplu) değer belirleme olsun, herhangi bir değerleme, bir model, arz ve talep faktörlerini temsil eden değer ve değişkenler arasındaki ilişkinin kelimelerle bir temsilini veya bir denkleminin kullanılmasıdır. Toplu değerleme modelleri, belirli bir alandaki belirli bir gayrimenkul türü için piyasayı temsil etmeye çalışmaktadır. Toplu değerleme uzmanları öncelikle modeli belirlemeli, yani değeri etkileyen arz ve talep faktörlerini ve gayrimenkul özelliklerini, örneğin yaşam alanının metrekaresini belirlemelidir. Daha sonra modeli kalibre etmeleri, yani seçilen değişkenlerin değer katkısını en iyi temsil eden ayarlamaları veya katsayıları belirlemeleri gerekmektedir. Örneğin, piyasanın yaşam alanının her bir metrekaresine yerleştirdiği para birimi miktarı gibi. Değerleri doğru tahmin eden bir modelin hem spesifikasyonu hem de kalibrasyonu için dikkatli ve kapsamlı bir piyasa analizi yapılması zorunlu görülmektedir. Toplu değerleme modelleri, değere yönelik karşılaştırmalı satış analizi, gelir yaklaşımı ve maliyet yöntemi olmak üzere üç yaklaşımın tümüne uygulanır.

IAOO standardı gayrimenkul türlerini tek aile konut, prefabrik konut, çok aileli konut, ticari ve sınai gayrimenkul, tarım dışı arazi, özel amaçlı gayrimenkuller olmak üzere 7 sınıfa ayırmış ve hangi değerleme yönteminin daha uygun olacağını belirtmişlerdir.

- Tek Aile Konut: karşılaştırmalı satış analizi, kat mülkiyeti dahil tek ailelik konut gayrimenkulleri için en iyi yaklaşımdır.
- Prefabrik konut: Genellikle mobil evler ayrı olarak satın alınır ve mobil ev parkındaki kiralık bir alanda bulunur. Bu durumda en iyi strateji, mobil evleri araziden ayrı modellemektir.

- Çok Aileli Konut: Yeterli satış ve gelir verileri mevcut olduğunda, çok aileli konut değerlemesinde satış karşılaştırma ve gelir yaklaşımları tercih edilir. Bu konu türünün değerlemesinde ÇRA ve ilgili teknikler başarıyla kullanılmıştır.
- Ticari ve Sınai Gayrimenkul: Yeterli gelir verisi mevcutsa, ticari ve sınai mülkiyet değerlemesinde gelir yaklaşımı en uygun yöntemdir. Doğrudan satış karşılaştırma modelleri, yeterli satışın olduğu alanlarında etkili olabilir. Yeterli satış verisi ve gelir verisi bulunmadığında, maliyet yaklaşımı uygulanmalıdır.
- Tarım Dışı Arazi: Tarım dışı arazilerin değerlemesi için satış karşılaştırma yaklaşımı tercih edilir. Diğer bir yaklaşım, arazi ve bina değerlerine ayrıştırılabilen hibrit bir modeldir.
- Tarım Arazileri: Yeterli satış verileri mevcutsa ve tarımsal mülkün piyasa değeri üzerinden değerlemesi yapılacaksa, karşılaştırmalı satış analizi yaklaşımı tercih edilir.
- Özel Amaçlı Gayrimenkuller: Maliyet yaklaşımı, özel amaçlı mülklerin değerlemesinde en uygun olma eğilimindedir.

Emlak Vergisi Politikası Standardının (IAAO 2010) 4.2.2 Bölümü, cari piyasa değerinin tüm gayrimenkullerin yıllık değerlemesi yapılmasını ifade ettiğini belirtilmektedir. Standartta bahsedilen yıllık değerleme, her bir gayrimenkulün her yıl yeniden incelenmesi gerektiği anlamına gelmekte, bunun yerine, modellerin yeniden kalibre edilebilmesi veya gayrimenkul türü, konumu, büyüklüğü ve yaşı gibi kriterlere dayalı olarak uygulanan oran çalışmaları veya diğer piyasa analizlerinden türetilen piyasa ayarlama faktörlerinin uygulanabileceği olarak yorumlanabilmektedir

Sadece fiziksel incelemelerle veri hatalarının düzeltebilmesi gayrimenkul özelliklerinin en az 4 ila 6 yılda bir gözden geçirilmesi ve güncellenmesi gerektiği göz ardı edilmemelidir. Bahsedilen güncellemeler

- Tüm gayrimenkullerin periyodik aralıklarla (yani her 4 ila 6 yılda bir) yeniden incelenmesi,
- Gayrimenkulleri döngüsel olarak yeniden denetlenmesi, ve

- Tüm gayrimenkulleri en az altı yılda bir incelenmesini sağlamaya devam ederken, oran çalışmaları veya diğer hususlar tarafından belirtildiği gibi gayrimenkulleri öncelik bazında yeniden denetlenmesi olmak üzere en az 3 şekilde gerçekleştirilebilmektedir.

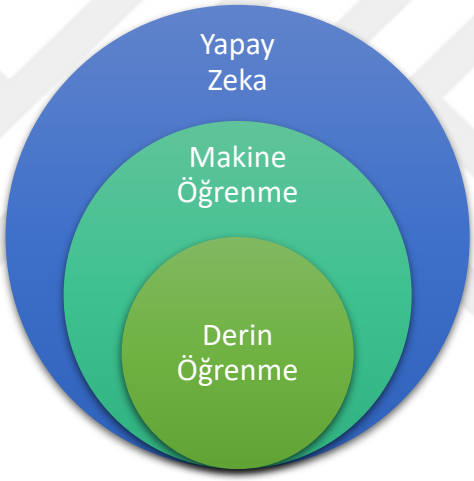
2.2 Standart üzerine genel değerlendirme

Genel olarak IAOO toplu değerleme standartları incelendiği zaman değerleme yöntemlerinde modern yöntemlerden çok bahsedilmediği daha çok klasik yöntemler ve ÇRA modelinde bahsedildiği görülmüştür. Ayrıca gayrimenkulün değerini etkileyen değişkenler ülkelere göre değişiklik göstermektedir ve bunun bir standart haline getirilmesi oldukça güçtür.

Bu standardın bir diğer dikkat çeken noktası ise çok fazla ayrıntının olmamasıdır. Toplamda 13 sayfadan oluşan toplu değerleme standardının önemli kısımları 3.1 bölüm başlığı altında irdelenmiştir.

3. MAKİNE ÖĞRENME ALGORİTMALARI

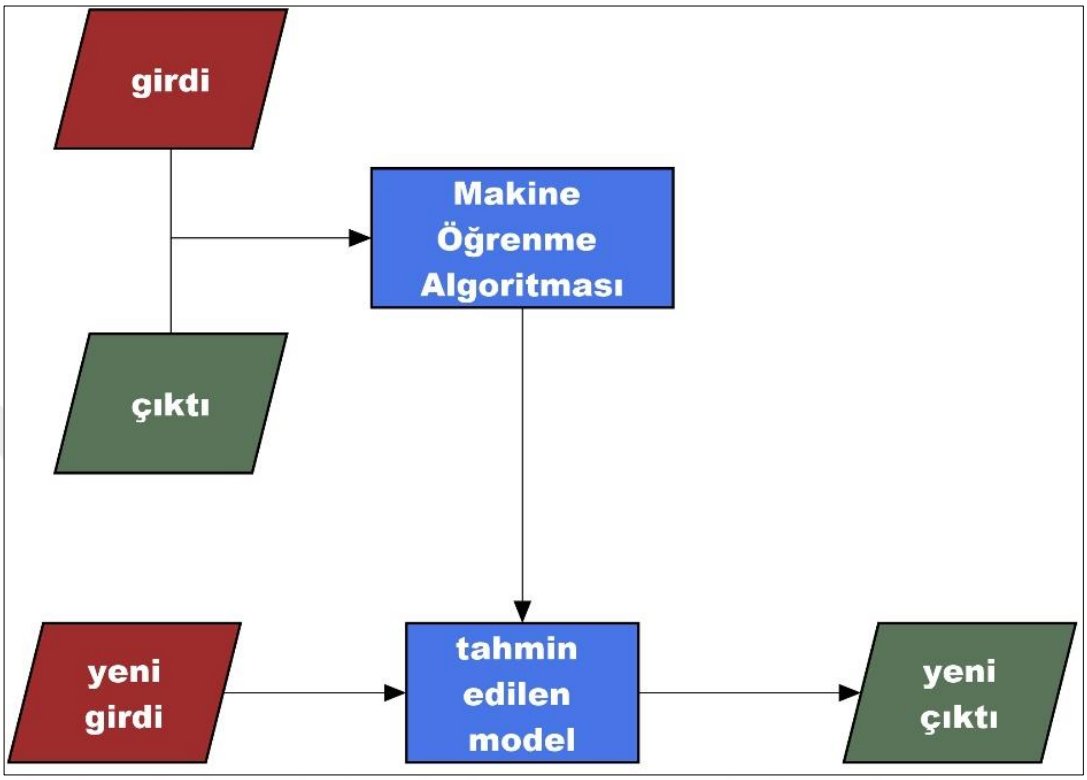
Son 30 yıldır makine öğrenme, yapay zekâ ve derin öğrenme kavramları birçok alanda sıklıkla kullanılan terimler arasına girmiştir. Bu üç terim birbirine benzese de aslında aralarında farklılıklar bulunmaktadır. Yapay zekâ daha kapsayıcı daha üst bir konumda olup, insan davranışlarını taklit eden teknik olarak tanımlanabilir. Makine öğrenme ise adından da anlaşılacağı gibi bir öğrenme süreci sonucu tahminde bulunmasıdır. Derin öğrenme ise sinir ağlarını kullanarak verilerdeki karakteristik özellikleri ortaya çıkaran teknik olarak tanımlanabilecektir. Yapay zekâ, makine öğrenme ve derin öğrenmeyi kapsayan daha geniş bir terim olarak dikkati çekmektedir (Şekil 3.1).



Şekil 3.1 Yapay zekâ, makine öğrenme ve derin öğrenme ilişkisi

Fazla miktarda veriye sahip problemleri tek tek manuel olarak çözmek zaman kaybı olarak görülmektedir. Bunun yerine verilerin bir miktarı kullanarak model oluşturulması ve geri kalan veriler için ise değer tahmini yapılması daha uygun görülmektedir. Bu durumlarda, makine öğrenme devreye girmektedir. Makine öğrenme, bir girdi verisi ile sonuçlarının arasındaki bağlantıyı bir algoritma dahilinde belirleyen yapay zekâ dalıdır. Makine öğrenmesi, bir problemi çözerken, o problem ile alakalı verileri kullanarak, farklı makine öğrenme algoritmaları ile problemin modellenmesi ve geri kalan veriler için problemin çözülmesi işidir. Diğer bir anlatımla makine öğrenme, belirli bir probleme ait

verilerin girdi ve çıktısını kullanarak, problemin modelini tahmin etme ve geri kalan girdilerin için sonucu tahmin etme olarak tanımlanabilir (Şekil 3.2).



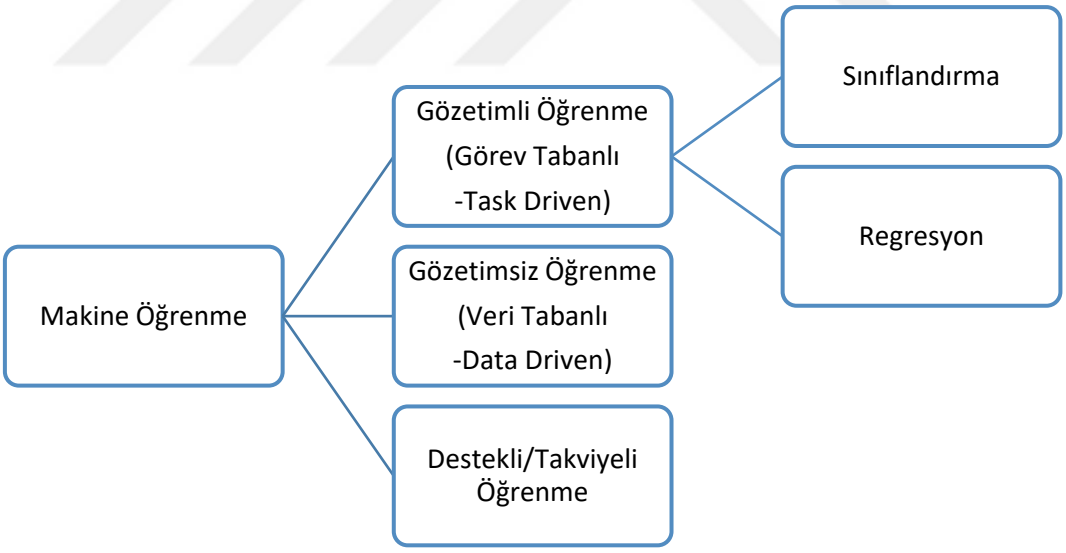
Şekil 3.2 Makine öğrenme algoritmalarının çalışma prensibi

Makine öğrenme, girdi ve çıktı belli ise aradaki modeli belirlemeye dayalı bir sistemdir. Örneğin, bir matematiksel problemde girdiler ve çıktılar belli ise makineye bu veriler öğretilerek, problemin çözümünde geri kalan işlemleri bu öğrenme sonucu yapılması beklenmektedir. Makine öğrenme çalışma prensibi basit bir örnekle açıklanmıştır. Çizelge 3.1'de verilen değerlerin sonucu belli iken problem/model belli değildir. Girdi ve çıktı arasındaki ilişki makine öğrenme algoritması ile çözülebilir. Bu örnekte, Çizelge 3.1'de, en sol sütunda, girdiler ve orta sütünde çıktı yani sonuçlar verilmiştir. Makineye, bu 5 (girdi-çıktı) veri, eğitim verisi olarak öğretilir. Makine seçilen algoritma ile aradaki ilişkiyi (modeli) çözer ve geri kalan veriler için sonuca ulaşılmaktadır.

Çizelge 3.1 Makine öğrenme prensibinin basit bir problem ile açıklanması

Problem (Girdi)	Sonucu (çıktı)	Problem çözümü için işaretler (model)
2 ?$_1$ 8 ?$_2$ 5	2	?$_1$ → + ve ?$_2$ → /
1 ?$_1$ 3 ?$_2$ 4	1	?$_1$ → + ve ?$_2$ → /
3?$_1$ 5 ?$_2$ 2	4	?$_1$ → + ve ?$_2$ → /
1 ?$_1$ 1 ?$_2$ 2	1	?$_1$ → + ve ?$_2$ → /
4 ?$_1$6 ?$_2$ 2	5	?$_1$ → + ve ?$_2$ → /

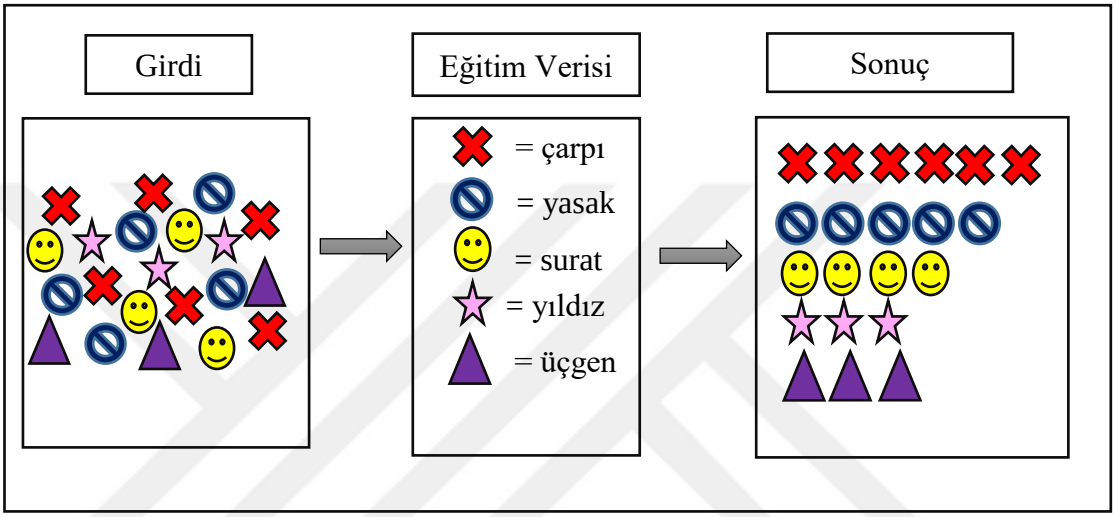
Makine Öğrenme çeşitleri genel olarak 3 ana sınıfa ayrılmakta olup, Şekil 3.3'de makine öğrenme genel sınıflandırılması özetlenmiştir. Makine öğrenme genel olarak Gözetimli Öğrenme, Gözetimsiz Öğrenme ve Destekli/Takviyeli Öğrenme olarak sınıflandırılmaktadır. Gözetimli öğrenme sınıflandırma ve regresyon olarak iki alt gruba ayrılmakta olup, gözetimli öğrenmede görev tabanlı (task driven) yani uygulamaya göre model oluşturulurken, gözetimsiz öğrenme ise veri tabanlı (data driven) yani veriye bağlı olarak model oluşturulmaktadır.



Şekil 3.3 Makine öğrenme yöntemi genel sınıflandırma

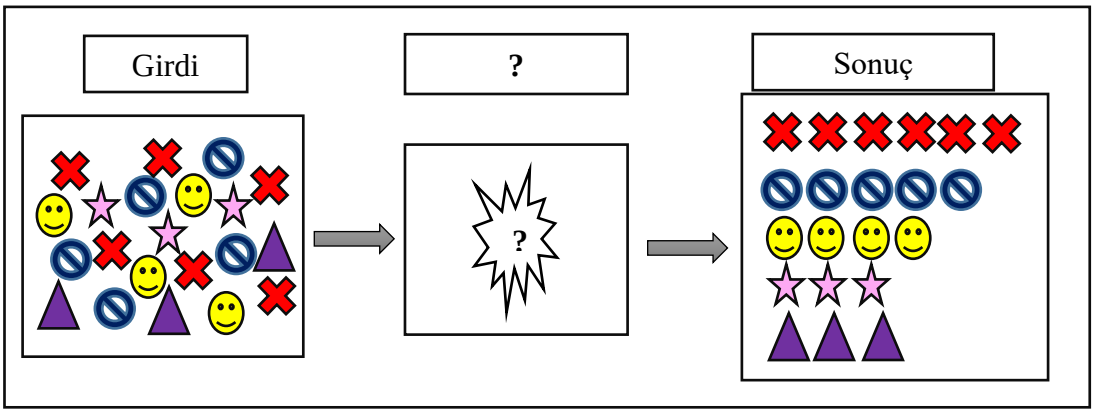
Gözetimli öğrenmede, veriler kullanarak bir model tahmini yapılmaktadır. Bu tür öğrenmede hem girdi hem çıktı algoritmaya verilmektedir. Gözetimli öğrenmede algoritma, sınıflandırma veya eğri uydurma (regresyon) şeklinde olup, bu makine

öğrenme türünde en büyük dezavantaj bir eğitim verisinin (training data) oluşturulması olarak görülmektedir. Çünkü bu eğitim verisi kullanılarak model oluşturulacağı için, eğitim verisinin güvenirliği sonuç ile birebir ilişkili olacaktır. Şekil 3.4'te gözetimli sınıflandırma türüne bir örnek gösterilmektedir. Girdi olarak 6 sınıf bulunmakta ve bu sınıflardan bazıları seçilerek algoritmaya öğretilmekte geri kalan verinin ise öğrenme sonucu oluşan modele göre sınıflandırılması istenmektedir.



Şekil 3.4 Gözetimli öğrenme şematik örneklenmesi

Gözetimsiz öğrenmede ise veri üzerinden tahmin yapılmaktadır. Bu tür öğrenmede eğitim verisi kullanılmamaktadır. Gözetimsiz öğrenmede, veriler gruplanmakta bu şekilde model oluşturulmakta ve diğer veriler en uygun gruba atanmaktadır. Şekil 3.5'de gözetimsiz öğrenme için örnek gösterilmektedir. Makine kendi öğrenip verileri özelliklerine göre sınıflandırarak bir sonuca ulaşmaktadır. Bu tür öğrenmede benzer olayların ayrımı zor olabilmekte ve gözetimli öğrenmeye göre daha hatalı sonuçlar elde edebilmektedir.



Şekil 3.5 Gözetimsiz öğrenme şematik örneklenmesi

Son olarak, destekli/takviyeli öğrenme yöntemi ise, insanın öğrenmesine benzer bir öğrenme tipi olarak dikkat çekmektedir. Makine gözlemler yaparak öğrenmekte ve sonuca ulaşmaktadır. Algoritmaya, beklenen sonuç yerine doğru/yanlış gibi fikir belirtilebilmektedir. Bu grup öğrenme diğer makine öğrenme türlerine kıyasla yeni bir tür olup, genellikle oyun algoritmalarında kullanılmaktadır.

Makine öğrenme algoritmaları kullanarak sonuca bir problemi çözme ve sonuca ulaşma genel olarak verilerin toplanması ve hazırlanması, modeli eğitme, modeli doğrulama ve sonuçların analizi olmak üzere 4 adımdan oluşmaktadır. İlk adımda, problem ile alakalı veriler toplanır gerekirse veri ön işlem uygulanmakta, ikinci adımda verilerin büyük bir kısmı eğitim az bir kısmı test verisi olarak ayrılarak model eğitilir, üçüncü adımda test verisi kullanılarak modelin performansı ölçülmekte ve son adımda ise farklı kalite analiz yöntemleri ile sonuçlar yorumlanmaktadır.

3.1 Rasgele Orman Algoritması

Rasgele orman algoritması karar ağaçlarından üretilmiş son 15 yıldır birçok alanda sıkça kullanılan bir makine öğrenme algoritmasıdır. RO algoritması sınıflandırma ve regresyon işlemleri olarak iki gruba ayrılabilmektedir. İlerleyen alt başlıklarda öncelikle RO regresyonu bir karar ağacı öğrenmesinden yola çıkılarak bulunan bir algoritma

olduğundan karar ağacı (decision tree) algoritmasından bahsedilecek ve sonrasında ise RO algoritması ve test ile eğitim verisi ayrıntılı anlatılacaktır.

3.1.1 Karar ağacı öğrenmesi

Karar ağacı öğrenmesi, bir makine öğrenme türüdür. İsminden de anlaşılacağı üzere yapısı bir ağaca benzeyen çok alt dalları olan bir öğrenmedir. Karar ağacında öncelikle işleme girdi değişkenler ile başlanır ve daha sonra bu girdi ve diğer değişkenlere göre dallar çoğalır. Dalların sonucunda ise yapraklar yani bu durumda kararlar bulunur. Karar ağaçları genelde tek bağımlı değişken ve çok sayıda bağımsız değişkeni olması nedeniyle regresyon modeline benzetilmektedir. Karar ağaçlarının en büyük avantajı sürekli veya kesikli verilerle çalışmaları, eksik veya hatalı verilerde iyi tahmin yapıyor olmalarıdır. Karar ağacı öğrenmesinde çeşitli algoritmalar bulunmakta olup, bu algoritmalardan en yaygın olarak kullanılanları ise Rasgele Orman, Hızlandırılmış Ağaç (Boosted Trees) ve Döndürme Ağacı (Rotation Trees) algoritmalarıdır.

Karar ağacı öğrenmesinde veriler devamlı kategorilere ayrılarak dallanmaktadır. Şekil 3.1'de karar ağacı basit bir örnekle anlatılmış olup, bu örnekle, bir ihtiyaç kredisine başvuracak kişilerin ihtiyaç kredisi alıp/alamayacağı ve alacaksa kredi miktarının belirlenmesi hedeflenmektedir. İhtiyaç kredisine başvuracak kişiye aşağıdaki sorular sorulur:

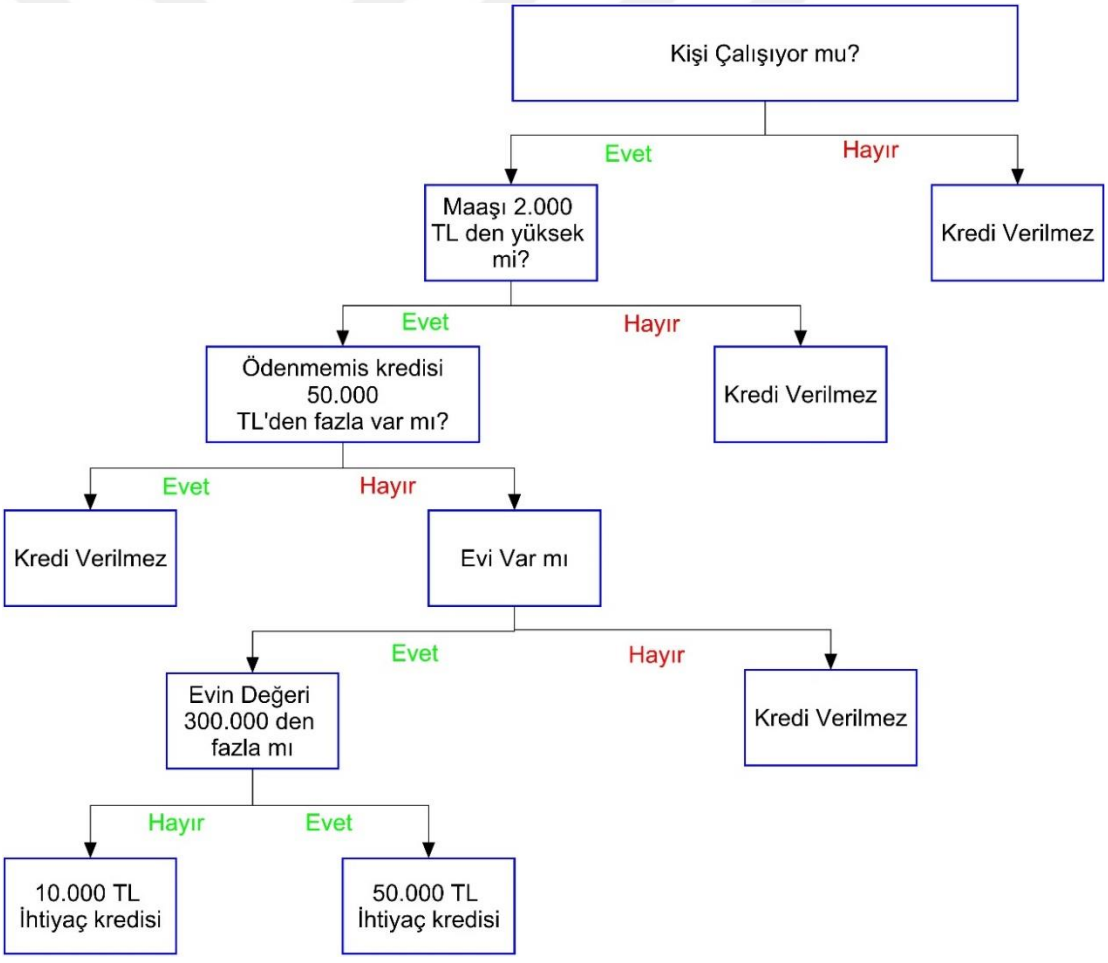
- Kişi çalışıyor mu?
- Maaşı ne kadar? Gelir Belgesi ne kadar?
- Daha önceden çektiği ödenmiş kredisi var mı?
- Sahibi olduğu bir ev var mı? Var ise evin değeri ne kadar? Tapusu var mı?

Yukarıdaki sorulara göre krediye Eray ve Egemen isimli iki kişinin başvurduğu ve bu sorulara verdikleri cevapların Çizelge 3.2'de olduğu gibi varsayılmış ve karar ağaçlarını anlatmak için bu örnek ayrıntılı olarak incelenmiştir.

Çizelge 3.2 İhtiyaç kredisi için banka soru ve cevapları

Kişi	Çalışıyor mu?	Maaşı Ne kadar?	50.000 TL'den fazla ödenmemiş kredisi var mı?	Evi var mı, Var ise değeri ne kadar?
Egemen	Evet	15.000	Hayır	Var, 975.000 TL
Eray	Evet	13.500	Evet	Hayır

Bankanın politikası gereği çalışmayan veya ödenmemiş kredisi 50.000 TL'den fazla olan kişilere ihtiyaç kredisi vermemek yönündedir. Bu durumda, banka tarafından karar ağacı algoritması sonucu, bu iki kişiden Egemen adlı kişiye 10.000 TL'lik ihtiyaç kredisi verilirken, Eray adlı kişiye ihtiyaç kredisi verilecektir. İhtiyaç kredisinin bankadan verilip verilmeyeceğine ve miktarına, Şekil 3.6'da olduğu gibi bir karar ağacı modeli oluşturularak bilgisayar tarafından karar verilebilecektir.

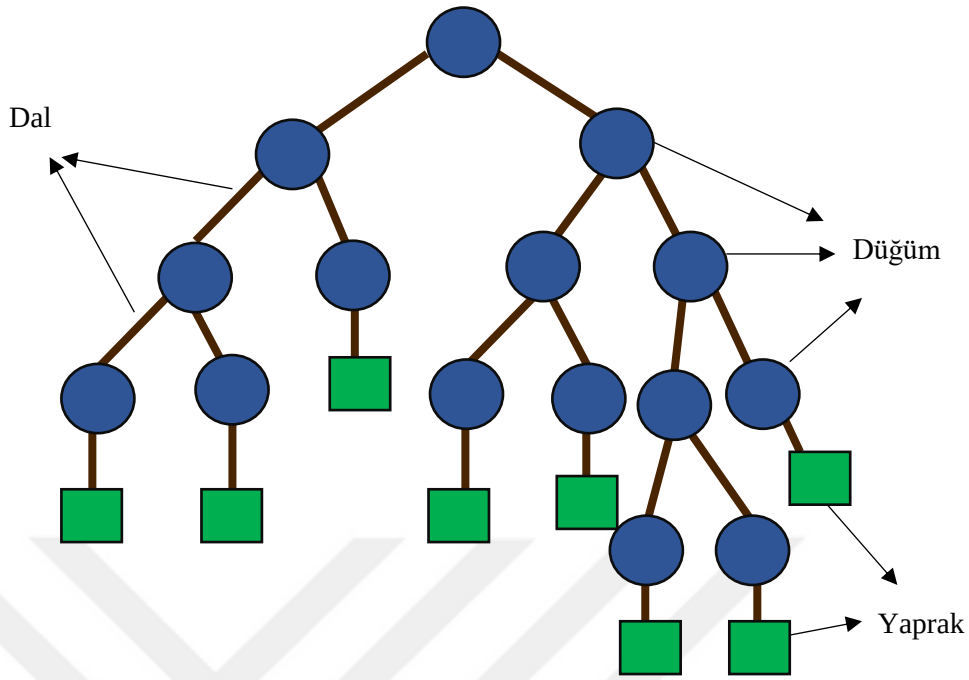


Şekil 3.6 Karar ağacı algoritmasının ihtiyaç kredisi örneği

3.1.2 Rasgele orman regresyonu

RO, önceki bölümlerde de bahsedildiği gibi bir karar ağacı türü olan bir makine öğrenmesi algoritmasıdır. Finans, sağlık bilimleri, harita, elektronik, makine mühendisliği, fizik, biyoloji gibi birçok farklı alanda kullanılmaktadır. RO, 2001 yılında Leo Breiman tarafından bulunmuş bir makine öğrenme yöntemi olup, yine Breiman (1996) tarafından bilinen torbalama (bagging) yönteminin geliştirilmiş şekli olarak literatüre girmiştir. Doğruluğu makine öğrenme içinde en yüksek olan algoritmalardan biri olarak görülmektedir (Breiman ve Cutler 2005). RO, birden fazla karar ağacı üreterek sonuca ulaşmakta, birçok karar ağacı kullanılması nedeniyle *'orman'* algoritması olarak adlandırılmaktadır. RO algoritmasında kullanıcıya ağaç sayısının kaç olacağı (N) ve her bir düğümde kullanılacak değişken sayısı (m) sorulmaktadır. Her bir ağaç eğitim verisinden rasgele seçilen değişkenlerle oluşturulmakta, algoritmaya '*rasgele'* denilmesinin sebebi ise bu durumdan gelmektedir.

Şekil 3.7'de bir karar ağacı yapısı örnek olarak gösterilmektedir. Oluşturulan N adet ağacın sonunda RO algoritması ile sonuçlar elde edilmekte ve veriye ait değişkenlerin ağrılıkları hesaplanarak model oluşturulmaktadır. Sonucu belli olmayan veriler için bu oluşturulan model ile işlem yapılarak sonuca ulaşmaktadır. Her ağaçtan elde edilen sonuçlara bakılarak, RO sınıflandırma makine öğrenmesi türünde, en yüksek olasılıkla elde edilen sonuç; RO regresyon makine öğrenmesi türünde ise her ağaçtan gelen ortalama alınarak, sonuca ulaşılacaktır.



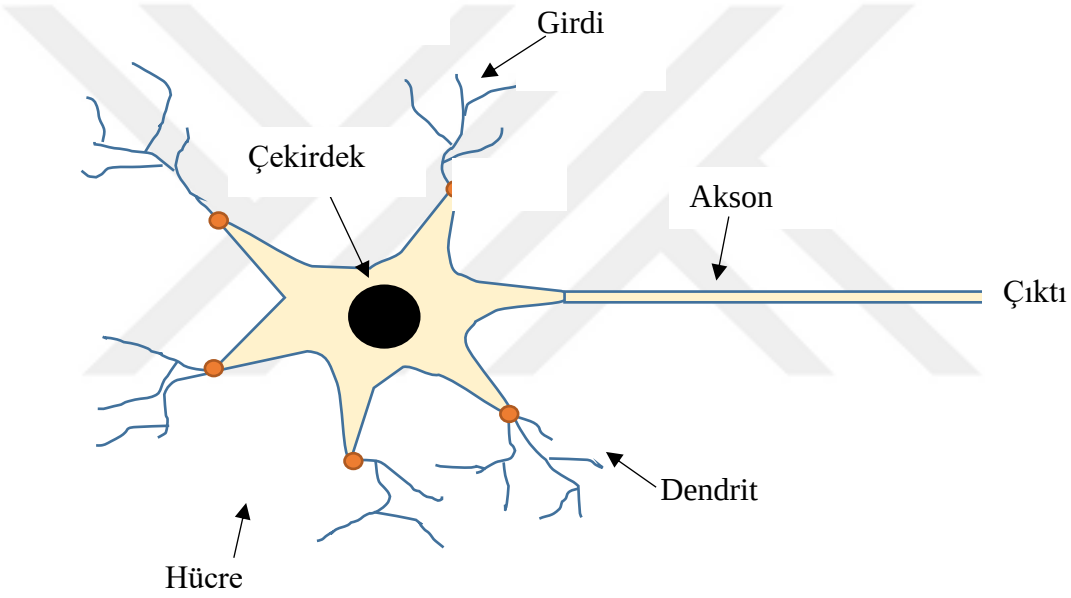
Şekil 3.7 Karar ağacı algoritmasının bir ağaç örneği

RO regresyonu, RO sınıflandırmasının tersine piksel değil obje tabanlıdır. Regresyon algoritmasında, bir sonuç ve bu sonuca bağlı değişkenler makineye verilmekte, makine ise bu sonuçlardan öğrenerek geri kalan test verisi için sonucu bulmaya çalışmaktadır.

3.2 Yapay Sinir Ağları Algoritması

Yapay sinir ağları (YSA) biyolojik insan beyninin yapısından türeyen bir MÖ algoritmasıdır. Birçok MÖ algoritmasında olduğu gibi öğrenme sürecinin matematiksel olarak modellenmektedir. YSA insan beyninin biyolojik birimleri olan nöronların modellenmesi ile başlamaktadır. Günümüzde birçok alanda en çok kullanılan MÖ türünden biridir. YSA algoritmasının kullanıldığı bazı alanlar: veri madenciliği, optik karakter tanıma, robotlarda yön belirleme, haberleşme, nesnelerin interneti, hastalıkların teşhisi, ağaçların sınıflandırılması, gayrimenkullerin değerlemesi vb. gibidir.

YSA algoritmasını daha iyi anlamak için öncelikle Şekil 3.8'de bir örneği gösterilmiş olan biyolojik beyin sinir hücresini incelemek gerekir. Biyolojik sinir hücresi bir gövde, bir akson, dendrit ve çok sayıda sinir hücresinden oluşmaktadır. Sinir hücresi ile sinir ucu arasındaki uzantılar ise akson olarak adlandırılmaktadır. Dendritler, gelen sinyalleri çekirdeğe iletmektedir. Çekirdeğin görevi gelen bütün sinyalleri toplamak ve aksona iletmektir. Toplanan bu sinyaller, akson tarafından işlenerek sinapslara gönderilmektedir. Sinapslar da yeni üretilen sinyalleri diğer sinir hücrelerine iletmektedir. Akson bu sinyalleri işlemekte ve çıktı olarak sunmaktadır. YSA gerçek biyolojik hücrelerin simülasyonu sonucu oluşmaktadır.



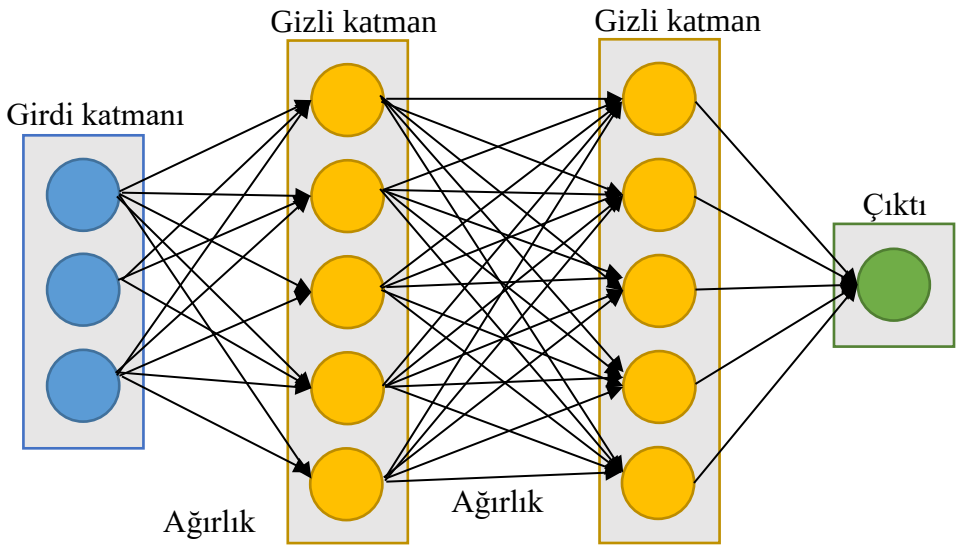
Şekil 3.8 Biyolojik beyin sinir hücresi

Çizelge 3.3'de YSA için biyolojik sinir hücresi ve YSA'daki karşılıkları gösterilmiştir. Biyolojik sinir hücresinde aksonun karşılığı YSA'da problemin çıktısı olarak gerçekleşmektedir. Biyolojik beyin sinir hücresindeki çekirdek aktivasyon fonksiyonuna, dendrit toplama fonksiyonuna, sinapslar ise her bir değişken için YSA algoritmasındaki ağırlıklara karşılık gelmektedir.

Çizelge 3.3 Biyolojik ve yapay sinir hücresi karşılıkları

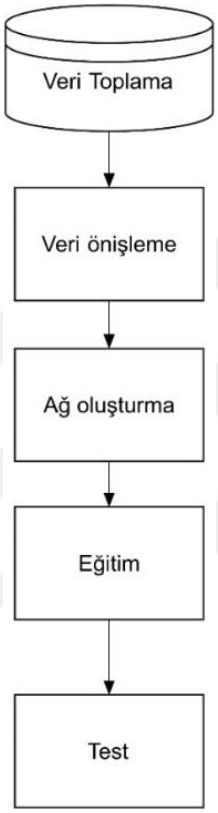
Biyolojik Sinir Hücresi	Yapay Sinir Hücresi
Akson	Çıktı
Dendrit	Toplama fonksiyonu
Çekirdek	Aktivasyon fonksiyonu
Sinaps	Ağırlıklar

Şekil 3.9'da biyolojik sinir hücresinin YSA algoritmasına gösterimi yapılmıştır. YSA algoritması 3 ana katmandan oluşmaktadır bunlar; girdi katmanı, gizli katman(lar) ve çıktı katmanıdır. Gizli katman sayısı birçok çalışmada 2 olarak seçilmektedir. Girdi katmanı ve gizli katmanlar arası ağırlıklandırma yapılır ve tüm değişkenler birbiri ile ilişkilendirilmektedir. Girdi katmandaki nöronların sayısı giriş yani değişken sayısı kadardır. Her nöron bir girişe aittir. Girdiler herhangi bir işleme uğramadan gizli katmana iletilmektedir. Gizli katman ise giriş katmanından aldığı bilgiyi işleyerek bir sonraki katmana iletmektedir. Her gizli katmandaki nöron sayısı değişkenlik göstermektedir. Bu katmandaki nöronların sayısı giriş ve çıkış katmanındaki nöron sayılarından fazla olmaktadır. Çıkış katmanı gizli katmandan gelen bilgiyi işlemekte ve giriş katmanına gelen girdiye uygun olarak üretilen çıktıyı oluşturmaktadır. Ağırlıklar YSA algoritmasının zekasını oluşturmaktadır. YSA algoritmasının öğrenme yeteneği direkt olarak ağrılıklar ile ilişkili görülmektedir.



Şekil 3.9 Çok katmanlı bir yapay sinir ağı modeli

YSA algoritması da birçok MÖ algoritması gibi eğitim verisi kullanılarak öğrenmekte, öğrenme sırasında gerçekleşmiş yani bilinen değerler yer almaktadır. YSA algoritmasının güvenilirliği test verileri ile sınanmaktadır. YSA algoritması genel olarak 5 ana adımda; verilerin toplanması, verinin ön işlenmesi, ağın tasarımı, ağın eğitilmesi ve test aşamaları olmak üzere ayrılabilir (Şekil 3.10).



Şekil 3.10 YSA algoritmasında temel aşamalar

YSA algoritması öncelikle girilen değerleri alır ve seçilen aktivasyon fonksiyonu ile işleme başlamaktadır. Çizelge 3.4'te YSA algoritmasında sıklıkla kullanılan fonksiyon adları ve eşitlikleri gösterilmiştir. Literatürde YSA algoritması ile gayrimenkul değerleme çalışmalarına bakıldığında Sigmoid fonksiyonun en çok tercih edilen aktivasyon fonksiyonu olduğu görülmüştür. Aktivasyon fonksiyonu bir değişkeni farklı bir boyuta taşıyan doğrusal veya doğrusal olmayan bir fonksiyondur. Aktivasyon fonksiyonunun türevinin kolay alınabilmesi eğitim hızını arttıran bir özelliktir.

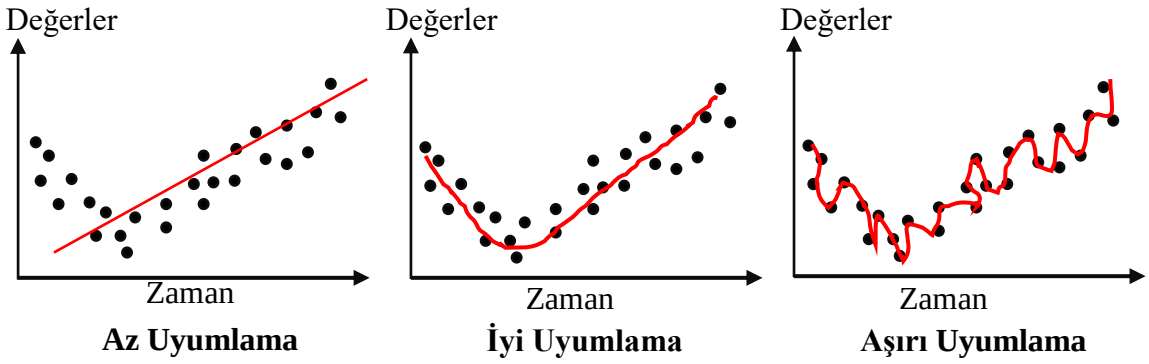
Çizelge 3.4 Sık kullanılan YSA aktivasyon fonksiyonları

No	Fonksiyon adı	Fonksiyon Eşitliği
1	Identity	$f(x) = x$
2	Lojistik	$f(x) = \frac{1}{1 - e^{-x}}$
3	Sigmoid	$f(x) = \frac{1}{1 - e^{x}}$
4	Tanh	$f(x) = \tanh(x/2)$
5	Signum	$f(x) = \begin{matrix} +1 & x > 0 \\ -1 & x < 0 \\ tanımsız & x = 0 \end{matrix}$
6	Hiperbolik	$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$
7	Exponensiyel	$f(x) = e^{-x}$
8	Softmax	$f(x) = \frac{e^{x}}{\sum_i e^{x_i}}$
9	Birim Toplam	$f(x) = \frac{e^{x}}{\sum_i x_i}$
10	Karekök	$f(x) = \sqrt{x}$
11	Sin	$f(x) = \sin(x)$
12	Step	$f(x) = \begin{matrix} 0 & x < 0 \\ 1 & x \geq 0 \end{matrix}$

3.2.1 Yapay sinir ağlarında aşırı uyumlama ve az uyumlama

YSA'da veriyi eğitirken sonuçları etkileyen aşırı uyumlama veya bir diğer adıyla ezberleme (overfitting) ve az uyumlama (underfitting) sorunları oluşabilmektedir. Aşırı uyumlama öğrenme verisini çok fazla detaylı öğrenme durumu olduğundan test verilerinin çıktısının tahmin edilmesinde sorun yaşanmaktadır. Buna karşın az uyumlama ise, modelin çok basit kalması durumunda ortaya çıkmakta, test verisinin çıktısını tahmin etmek yine zorlanmaktadır.

Modelin en iyi şekilde tahmin edilebilmesi için ise öğrenmenin iyi uyumlaması (good fitting) olması gerekmektedir. Şekil 3.11'de bir veri seti üzerinde, aşırı, iyi ve az uyumlama gösterilmektedir. YSA'da az ve aşırı uyumlamayı önlemek için sentetik veri veya gürültü eklenebilecektir.



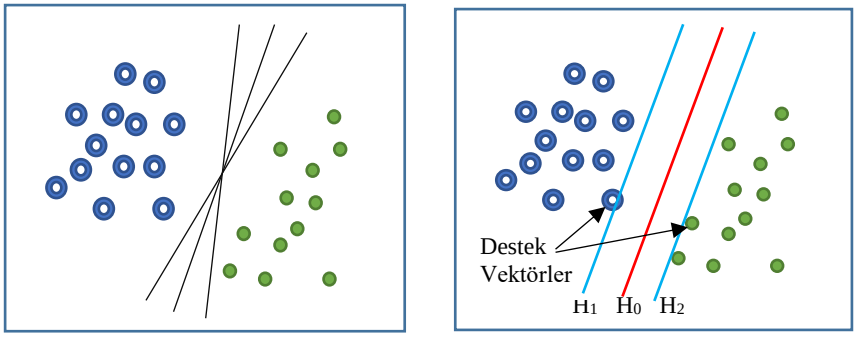
Şekil 3.11 YSA algoritmasında görülen az, iyi ve aşırı uyumla gösterimi

3.3 Destek Vektör Regresyonu Algoritması

DVM algoritması popüler olan bir MÖ algoritmasıdır. DVM algoritması ilk başlarda sınıflandırma için kullanılsa da zamanla regresyon problemleri içinde kullanılmaya başlanmış ve DVR algoritması olarak yaygın kullanılmıştır. DVR algoritması DVM algoritmasının regresyon hali olarak ifade edilebilecektir.

DVM'ler ikili sınıflandırma problemlerini konveks optimizasyon problemleri olarak formüle ederek çözer (Vapnik 1998). DVM algoritması genellikle, örüntü tanıma, nesne tanıma ve metin yazı tanıma işlemlerinde sıklıkla kullanılmaktadır. DVM genel olarak iki farklı sınıfı ayırmak için geliştirilmiş bir algoritmadır. DVM algoritması iki sınıfı birbirinden ayırabilmek için en uygun hiper düzlemi bulmaya çalışır.

Sınıfları ayırmak için en uygun fonksiyonun tahmin edilmesine dayalı bir yöntemdir. Sınıfları ayırmak için kullanılan fonksiyonlar doğrusal olabileceği gibi doğrusal da olmayabilir. Doğrusal çözümler doğrusal olmayanlara göre daha kolay ve basittir. Örneğin Şekil 3.12'de doğrusal bir düzlem ile ayrılabilen sınıflandırma örneği gösterilmiştir.



Şekil 3.12 Doğrusal olarak ayrılabilen iki sınıfın temsili gösterimi

Yukarıdaki şekilde iki sınıflı bir veri görülmektedir. Bu iki veriyi ayırabilen birçok hiper düzlem olabilmekte ancak DVM algoritmasının ana amacı iki veri sınıfına eşit uzaklıkta maksimum boşluklu ve en düşük sınıflama hatalı hiper düzlemi bulmaktır. Bu hiper düzlemlerin diğer adı sınıflandırıcı olarak geçmektedir. H_1 ve H_2 hiper düzlemleri üstünde bulunan ve H_0 hiper düzlemine en yakın noktalara destek vektörler denir. H_1 ve H_2 hiper düzlemlerinin arasındaki uzunluk ise en büyük uzunluk veya marjin (m) dir.

H_0 düzlemi aşağıdaki şekilde ifade edilmektedir.

$$H_0 = wx + b = 0 \tag{3.1}$$

Bu durum da H_1 ve H_2 düzlemler ise aşağıdaki şekilde ifade edilmektedir.

$$H_1 = wx + b = 1 \tag{3.2}$$

$$H_2 = wx + b = -1 \tag{3.3}$$

Bu formüllerde w ağırlık vektörünü, b ise sabit bir sayıyı göstermektedir.

$$wx_i + b \geq 1 \; ise \; y_i = +1 \tag{3.4}$$

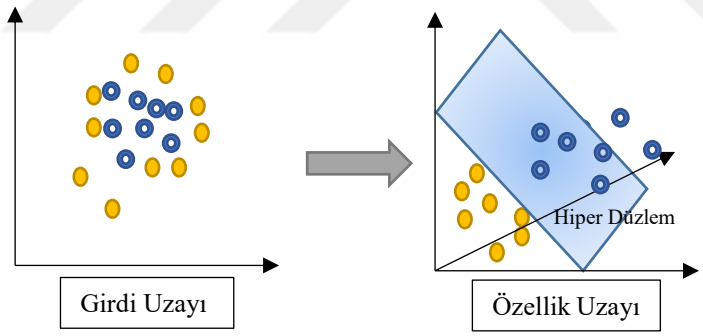
$$wx_i + b \leq -1 \; ise \; y_i = -1 \tag{3.5}$$

DVM'de amaç destek vektörleri ile H_0 hiper düzlem arasındaki dik uzaklık d'nin bulunması ve d uzaklığın en uzak hesaplanmasıdır. Bu durumda, ağrılık vektörü w minimum olması gerekmektedir.

$$d = \frac{|wx_p \mp b|}{\|w\|} \tag{3.6}$$

$$Min: \frac{1}{2}\|w\|^2 \tag{3.7}$$

Son olarak, doğrusal olmayan DVM problemleri için ise daha karmaşık çözümler söz konusudur. Doğrusal olmayan problemlerde yüksek boyutlu uzaya verinin taşınması ile çözüm bulunur. En uygun hiper düzlem bu özellik uzayında bulunabilir. Şekil 3.13'te doğrusal olarak ayrılamayan verilerin farklı boyutlardaki uzaya aktarılması gösterilmiştir. Doğrusal olmayan verilerde DVM yüksek boyutlu uzaya çekirdek fonksiyonları ile dönüştürülür.



Şekil 3.13 Doğrusal olmayan problemlerde uzay dönüşümü

DVM algoritmasının avantajlarını

- Doğrusal olmayan verileri iyi sonuçlar vermekte,
- Hem sınıflandırma hem de regresyon problemlerini çözmekte,
- Verilerde yapılacak küçük değişiklikler, hiper düzlemi ve dolayısıyla DVM'yi büyük ölçüde etkilemez,
- Aykırı değerlere karşı güçlüdür,

- SVR, diğer regresyon tekniklerine kıyasla daha düşük hesaplama gerçekleştirir ve
- Uygulanması kolaydır fakat veri sayısı arttıkça işlem hızı azalması, olarak sıralamak mümkündür.

DVM algoritması genellikle sınıflandırma problemleri için kullanılsa da regresyon problemleri için de kullanılabilmekte bu durum alternatif bir kayıp fonksiyonu verilmesi ile sağlanmaktadır (Gunn 1998). DVR algoritmasında sınıflandırma problemlerinde olduğu gibi aşağıdaki gibi bir hiper düzlem fonksiyonu kullanılır. Buna ek olarak (Ɛ) hedefteki sapma miktarıdır.

$$y = f(x) = w.x + b \tag{3.8}$$

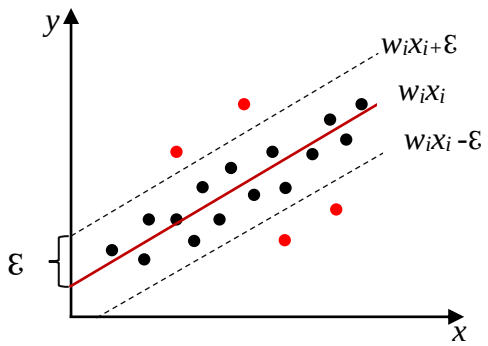
Hata değeri hiper düzlem üzerinde her iki tarafa da aşağıdaki gibi eklenmektedir.

$$y_i - wx_i - b \leq \varepsilon \, ise \tag{3.9}$$

$$-y_i + wx_i + b \leq \varepsilon \tag{3.10}$$

Çözüm için *w* ağırlığı minimum yapmak gerekmektedir.

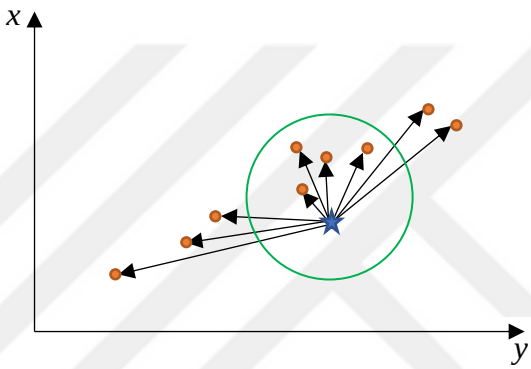
$$min_w = \frac{1}{2}\|w\|^2 \tag{3.11}$$



Şekil 3.14 Destek Vektör Regresyon ve hata parametresi gösterimi

3.4 K En Yakın Komşuluk Algoritması

K en yakın komşuluk algoritması kısaca k-nn algoritması genellikle sınıflandırma problemleri için kullanılsa da regresyon problemleri için de sıklıkla literatürde kullanılmaktadır. K-nn algoritması, herhangi bir yeni veri noktasının değerini tahmin etmek için "değişken benzerliği" (feature similarity) kullanmaktadır. Bilinmeyen değeri olan bu yeni nokta için eğitim kümesindeki noktalara ne kadar yakın olduğuna bağlı olarak bir değer atanmaktadır.



Şekil 3.15 K-nn ile tahmin edilen nokta ve uzaklıklar

K-nn regresyon algoritması aşamasında

1. Tahmini yapılacak nokta ile her bir eğitim noktası arasındaki mesafenin hesaplanması,
2. K tane en yakın noktanın mesafeye göre seçilmesi ve
3. Seçilen veri noktalarının ortalamasının alınmasıyla yeni noktanın tahmin edilmesi, olmak üzere 3 ana adımdan oluşmaktadır.

K-nn algoritmasının en büyük avantajlarından biri diğer MÖ algoritmalarıyla kıyasla daha az parametre ile işlem yapmasıdır. K-nn algoritmasında k değeri ve bir mesafe metriği yeterlidir. En büyük dezavantajı ise daha fazla bellek ve depolama alanına ihtiyaç duyması sayılmaktadır.

K-nn algoritması hesaplama yaparken farklı mesafeleri kullanabilmektedir. K-nn algoritmasında belli başlı sıkça kullanılan mesafe hesaplama yöntemleri aşağıdaki gibidir.

- Öklid Uzaklığı

En yaygın olarak kullanılan uzaklık ölçüsüdür. Aşağıdaki formülü kullanarak, bir nokta (x) ile mevcut bir nokta (y) arasındaki farkların karesinin toplamının karekökü olarak hesaplanmaktadır.

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(y_i - x_i)^2} \quad (3.12)$$

- Manhattan Uzaklığı

Manhattan uzaklığı, iki nokta arasındaki mutlak değeri ölçen bir başka popüler uzaklık ölçüsüdür. Bir kişinin bir adresten diğer adrese, şehir caddesi üzerinden nasıl gidebileceğini gösteren bir ızgara ile yaygın olarak görselleştirildiği için taksi mesafesi veya şehir bloğu mesafesi olarak da adlandırılmaktadır ve aşağıdaki formül kullanılarak hesaplanmaktadır.

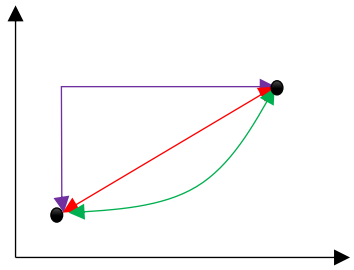
$$d(x,y) = \sum_{i=1}^{n}|y_i - x_i| \quad (3.13)$$

- Minkowski uzaklığı

Minkowski uzaklık ölçüsü, Öklid ve Manhattan uzaklık metriklerinin genelleştirilmiş halidir. Aşağıdaki formüldeki p parametresi, diğer mesafe ölçümlerinin oluşturulmasına izin verir. Öklid mesafesi, p değeri 2'ye eşit olduğunda bu formülle temsil edilir ve Manhattan mesafesi, p değeri 1'e eşit olarak gösterilir.

$$d(x,y) = (\sum_{i=1}^{n}|y_i - x_i|^p)^{1/p} \quad (3.14)$$

Öklid, Manhattan ve minkowski uzaklıkları k-nn MÖ algoritmasında en sık kullanılan mesafeler olup, aralarında fark görsel olarak Şekil 3.16'te gösterilmiştir.



Şekil 3.16 Minkowski (yeşil), Manhattan (mor) ve Öklid (kırmızı) uzaklığı

3.5 Çoklu Regresyon Analizi

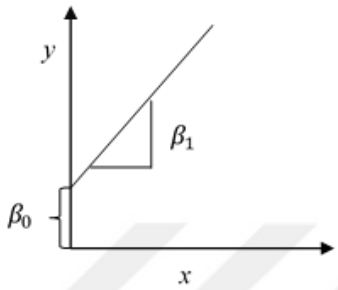
Çoklu regresyon analizi iki veya daha fazla değişken arasındaki ilişkiyi modellemek için kullanılan bir yöntemdir. İstatiksel bir yöntem grubuna da giren regresyon temelde klasik bir makine öğrenme modelidir. Regresyon analizi genel olarak basit ve çoklu regresyon olarak 2'ye ayırmak mümkündür. Çoklu regresyon ise kendi içinde doğrusal, yarı-logaritmik ve tam logaritmik olarak 3'e ayrılabilir. Gayrimenkul değerleme çalışmalarında çoklu regresyon tercih edilmektedir. Yapılan çalışmalar genelde gayrimenkul değer tahminin bir doğrusal olmayan problem olduğunu ortaya koymuştur (Čeh vd. 2018, Kontrimas ve Verikas 2011, Yu ve Wu 2016). Bazı çalışmalarda ise doğrusal regresyon analizi ile değer tahmini yapılmıştır. Bu tez çalışmasında aşağıda anlatılan 3 farklı regresyon tahmini denenmiştir.

- Doğrusal Regresyon Analizi

Basit doğrusal regresyon analizinde bağımlı değişken ile bağımsız değişken arasında bulunan ilişki incelenmektedir. Basit doğrusal regresyonda f(x) fonksiyonunda doğrusal bir bağımlı (y) ve bir bağımsız (x) değişken bulunur. Basit doğrusal model, aşağıdaki denklem kullanılarak ifade edilir.

$$y = \beta_0 + \beta_1 X + \varepsilon \quad (3.15)$$

Bu formülde, β_0 doğrunun y-eksenini kesiştiği yer ve regresyon sabitini, β_1 doğrunun eğimi veya regresyon katsayısını ve ε ise rastgele hata değerini göstermektedir (Şekil 3.17).



Şekil 3.17 Basit doğrusal regresyon gösterimi

ÇRA basit doğrusal regresyonun artırılmış şeklidir. ÇRA iki veya daha fazla değişken ile bir bağımlı değişkenin değerinin tahmini için kullanılmaktadır. Burada tahmin edilmek istenen değişkene bağımlı değişken denirken bağımlı değişkeni tahmin etmek için kullanılan diğer değişkenler ise bağımsız değişken olarak adlandırılmaktadır. ÇRA analizinde tek bir bağımlı değişken ve birden çok bağımsız değişken bulunmak ve aralarındaki ilişki analiz edilerek bir model oluşturulmaktadır.

Çoklu regresyon gayrimenkullerin toplu değerlemesinde sıkça kullanılan bir yaklaşımdır. Çoklu regresyonun terimi ilk olarak Pearson (1908) tarafından kullanılmış olup, çoklu regresyon birden çok bağımsız değişken ile bir bağımlı veya ölçüt değişken arasındaki ilişki hakkında bilgi edinmek için kullanılmaktadır. Örneğin, bir gayrimenkul değerleme uzmanı her bir liste için evin büyüklüğünü (metrekare olarak), yatak odası sayısını, evin yaşı gibi verilere göre ilgili mahalledeki ortalama gayrimenkul değerini tahmin edebilecektir. Bu bilgiler çeşitli gayrimenkuller için derlendikten sonra, evin değerini etkileyen değişkenlerin bir evin satıldığı değer ile nasıl ve ne şekilde ilişkili olduğu bu sayede görülebilecektir.

$$y = a + b_1X_1 + b_2X_2 + \cdots + b_nX_n + \varepsilon \quad (3.16)$$

Bu formülde Y bağımlı değişkendir ve $X_1,\ldots,X_n$ n bağımsız değişkendir. a, $b_1,\ldots,b_n$ ağırlıklarının hesaplanmasında regresyon analizi, bağımsız değişkenler kümesinden bağımlı değişkenin maksimum tahminini sağlayacaktır. Bu işlem genellikle en küçük kareler tahmini ile yapılır. Gayrimenkullerin çoklu regresyona analizi ile tahmininde ise daha spesifik olarak aşağıdaki formül yazılabilir.

$$Değer_i = \beta_0 + \beta_1 Alan_1 + \beta_2 Odasayısı_2 + \cdots + b_nX_n + \varepsilon \quad (3.17)$$

Bu formülde n adet gayrimenkullerin değeri (*Değer*) bağımlı değişken, $\beta_0, \beta_1, \ldots, \beta_n$ ağırlıkları kullanılarak ve *X*'ler ise bağımsız değişkenleri yani gayrimenkulün değerini etkileyen parametreleri göstermektedir.

- Yarı-logaritmik Regresyon Analizi

Doğrusal olmayan modellerde yarı-log diye adlandırılan aşağıdaki formülle gösterilen modeller kullanılmaktadır. Aşağıdaki eşitlikte a sabit terim, b eğim katsayısını ve Ɛ ise hata terimini göstermektedir. Gayrimenkul değeri tahmin edilirken literatürde bazı araştırmacılar doğrusal model kullanılırken bir kısmı ise doğrusal olmayan modelleri denemiştir.

$$Log_y = a + b_1X_1 + b_2X_2 + \cdots + b_nX_n + \varepsilon \quad (3.18)$$

- Tam logaritmik Regresyon Analizi

Logaritmik regresyon, büyüme veya bozulmanın önce hızla hızlandığı ve sonra zamanla yavaşladığı durumları modellemek için kullanılan bir tür regresyondur. Gayrimenkul toplu değerleme işlemlerinde çok sık olmasa da literatürde kullanıldığı görüldüğünden bu

sebeple bu tez çalışmasında da veriler tam logaritmik regresyon ile test edilmiş ve analiz edilmiştir.

$$Log_y = a + b_1 LogX_1 + b_2 LogX_2 + \cdots + b_n LogX_n + \varepsilon \quad (3.19)$$

3.6 Temel Bileşenler Analizi

Temel Bileşenler Analizi (TBA) fazla sayıda değişkene sahip bir veri kümesinde değişkenler arasındaki ilişkiye bakarak değişkenlerin bilgilerini çok fazla kaybetmeden daha az değişkene döndürme diğer bir ifade ile boyut azalma yöntemi olarak bilinmektedir. TBA analizi aşağıda sayılan temel adımlar ile yapılmaktadır.

1. Veri kümesi standartlaştırılır.
2. Veri kümesindeki özellikler için kovaryans matrisini hesaplanır.
3. Kovaryans matrisi için özdeğerleri ve özvektörleri hesaplanır.
4. Özdeğerleri ve bunlara karşılık gelen özvektörleri sıralanır.
5. k özdeğerleri seçilir ve bir özvektör matrisi oluşturulur.
6. Orijinal matrisi dönüştürülür.

MÖ ve optimizasyon algoritmalarının çoğu, genellikle tüm özellikler aynı ölçekte olduğunda daha iyi performans göstermektedir. Bu sebeple TBA algoritmasında ilk adım bir standardizasyon sağlanmasıdır. Veri setini standartlaştırmak için her bir özellik için ortalama ve standart sapmayı hesaplanmalıdır. Bu formülde μ_x değişkenlerin ortalamasını, σ_x varyansı göstermektedir.

$$x_{std}^{i} = \frac{x^{i} - \mu_x}{\sigma_x} \quad (3.20)$$

Bir sonraki adımda ise kovaryans matris hesaplanmaktadır. Kovaryans matris formülü aşağıdaki gibidir. Kovaryans, iki özelliğin birbiriyle nasıl değiştiğini ölçmektedir. Pozitif kovaryans, özelliklerin birlikte arttığını ve azaldığını gösterir. Bunun tam tersi negatif bir

kovaryans ise iki özelliğin zıt yönlerde değiştiğini gösterir. Aşağıdaki formülde k ve j iki değişkeni, göstermektedir.

$$\sigma_{jk} = \frac{1}{n-1}\sum_{i=1}^{n}(x_j^i - \mu_j)(x_k^i - \mu_k) \quad (3.21)$$

Kovaryans matrisi hesaplandıktan sonraki adımda, özdeğerleri ve özvektörler hesaplanmaktadır. Özvektörler, kovaryans matrisinin temel bileşenlerini (maksimum varyansın yönleri) temsil eder. Özdeğerler ise karşılık gelen büyüklükleridir. Karşılık gelen en büyük özdeğere sahip özvektör, maksimum varyansın yönünü temsil etmektedir. Burada λ özdeğerdir.

$$\sum \vartheta = \lambda\vartheta \quad (3.22)$$

Özdeğerler hesaplanıp sıralandıktan sonra, en büyük ilk 2 özvektör seçilir. Son olarak veriler k adet özvektörler çarpılarak dönüştürülen matrise çevrilmektedir.

3.7 Makine Öğrenme Algoritmalarında Uyum İyiliği Ölçütleri

Makine öğrenme algoritmalarının kıyaslanması için uyum iyiliği ölçüt analizi yapılmıştır. Bu karşılaştırma için aşağıda formüller ile gösterilen metrikler her bir MÖ algoritmaları için hesaplanmıştır. Böylece yöntemlerden hangisinin daha başarılı olduğu analiz edilmiştir.

- Hata Kareler Ortalaması (Mean Square Error)

$$HKO = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad (3.23)$$

Bu eşitlikte, HKO, Hata Kareler Ortalaması, *N* test edilen veri sayısını, y_i gözlemlenen değeri ve $\hat{y}_i$ ise MÖ algoritmalı kullanılarak tahmin edilen değeri göstermektedir. HKO

bir diğer tanım ile bilinen ve tahmin edilen değerler arasındaki kare farkının ortalamasını ölçmektedir. Bu formül ile artıkların varyansını ölçmektedir.

- Kök Ortalama Kare Hata (Root Mean Square Error)

$$KOKH = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{3.24}$$

Bu eşitlikte, KOKH, Kök Ortalama Kare Hata, *N* test edilen veri sayısını, y_i gözlemlenen değeri ve $\hat{y}_i$ ise MÖ algoritmaları kullanılarak tahmin edilen değeri göstermektedir. Kök Ortalama Kare Hatası, Ortalama Kare hatasının kareköküdür. Artıkların standart sapmasını ölçmektedir.

- Ortalama Mutlak Hata (Mean Absolute Error)

$$OMH = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{3.25}$$

Bu eşitlikte, OMH, Ortalama Mutlak Hata, *N* test edilen veri sayısını, y_i gözlemlenen değeri ve $\hat{y}_i$ ise MÖ algoritmaları kullanılarak tahmin edilen değeri göstermektedir. Test verisindeki artık değerlerin (residual) ortalamasını ölçmektedir.

- R^2:

Regresyon modellerinin performansını ölçen yaygın olarak bilinen bir uyum iyiliği ölçüt yöntemi olup belirlilik katsayısı olarak da bilinmektedir. Modelin performansını ölçmek için kullanılan bir değerdir.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} \tag{3.26}$$

Bu formülde

y_i gözlem sonucu elde edilmiş değerler,

$\hat{y}$ regresyon eşitliğinden hesaplanmış değerleri ve,

$\bar{y}$ ise verilerin ortalamasını göstermektedir.

- Düzeltilmiş R^2

R^2 bir modeldeki değişkenler tarafından açıklanan bir bağımlı değişken için varyansın oranını hesaplayan istatistiksel bir yöntemdir. Korelasyon bağımsız ve bağımlı değişken arasındaki ilişkiye bakarken, düzeltilmiş R^2 formülünde değişken sayısı daha çok ön plandadır. Bunun yanı sıra modelin kalitesini düzeltilmiş R^2 ile de test edilmektedir.

$$R^2 = 1 - (1 - R^2) * \left[\frac{(n-1)}{n-k-1}\right] \quad (3.27)$$

Bu formülde n veri sayısını, k modelde bağımlı değişken harici olan değişken sayısını göstermektedir.

Belirlilik katsayısı ve determinasyon katsayısının 0 ila 1 arasında olması beklenir. R^2 değerinin 1'e yaklaşması modelinin uygunluğunu ve güvenirliliğini belirtmektedir.

3.7.1 Değerin güncellenmesi ve makine öğrenme algoritmalarında eğitim ve test verisi

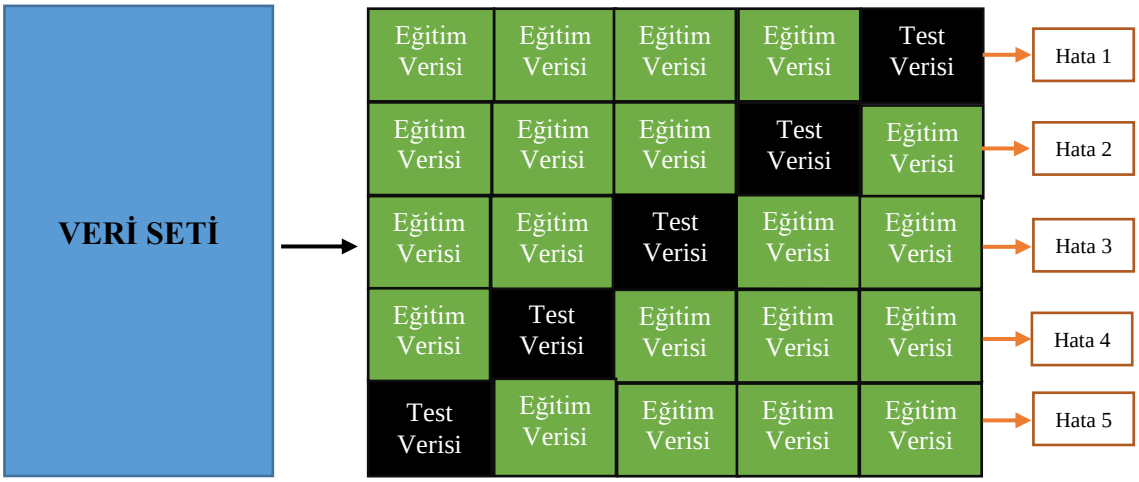
Gayrimenkullerin değeri aşağıdaki formül ile $Yi_{ÜFE}$ kullanılarak bu çalışmanın verilerinin güncellendiği tarih olan Aralık 2021 değerine çekilmiştir. Toplamda 1,315,675 adet konut, 115 adet ticari ve 213 adet arsa türünde gayrimenkule ait veriler bulunmakta ve bu veriler 2017-2021 yılları arasından çekilmiştir. Bu nedenle tüm gayrimenkullerin değerleri Aralık 2021 tarihli güncel değere aşağıdaki formül kullanılarak güncellenmiştir. Bu tez çalışmasında gayrimenkullerin toplam değeri yerine TL/m^2 değeri çıktı değer olarak kullanılmıştır.

$$Güncel\ Değer = \frac{(DeğerxYi_{ÜFE\ (güncel)})}{Yi_{ÜFE\ (eski)}} \quad (4.1)$$

Yukarıdaki eşitlikte Güncel değer; gayrimenkullerin güncelleneceği tarihteki değeri, Değer; gayrimenkulün ilgili tarihteki atanmış değeri, $Yi_{ÜFE(güncel)}$; gayrimenkulün güncelleneceği tarihteki Yi-ÜFE'yi ve $Yi_{ÜFE(eski)}$; gayrimenkulün değeri belirlendiği zamanki Yi-ÜFE'yi göstermektedir (Tanrıvermiş 2017, Tanrıvermiş ve Aliefendioğlu 2008).

Eğitim verisi makine öğrenme algoritmalarında algoritmayı eğitmek için kullanılmaktadır. Ayrıca doğrulama veri seti (validation) bulunmaktadır. Bu veri seti modelin performansını değerlendirmek için kullanılan eğitim veri setinin bir alt veri setidir. Test verisi ise eğitim ve doğrulama seti içinde bulunmayan değerleri içermektedir.

Çapraz doğrulama (cross validation) MÖ algoritmasının kalitesini tahmin etmek için kullanılabilecek bir işlemdir. Çapraz doğrulama sonuçları, en iyi parametre değerleri kümesini seçmek için kullanılabilir. K-kat (k-fold) çapraz doğrulama işlemi, bir veri kümesinde MÖ algoritmasının performansını tahmin etmek için standart bir yöntemdir. K-kat çapraz doğrulama prosedürü, sınırlı bir veri kümesini örtüşmeyen k tane kata böler. Diğer tüm kıvrımlar topluca bir eğitim veri seti olarak kullanılırken, k kıvrımın her birine, geride tutulan bir test seti olarak kullanılma fırsatı verilir. Toplam k model, k tutma test setlerine sığdırılır, değerlendirilir ve ortalama performans rapor edilmektedir (Şekil 3.18).



Şekil 3.18 MÖ algoritmalarında verilerin eğitim ve test verisi olarak dağılımı ve çapraz doğrulama

Bu tez çalışmasında özellikle az veriye sahip arsa, arazi ve ticari tür gayrimenkullerde 5 katlı çapraz doğrulama ile veriler test edilip analiz edilmiştir. Daha açıkça anlatmak gerekirse, verilerde %80 eğitim ve %20 test verisi alınmıştır. Bu verilerden %20 lik ve %80 lik kısmı her seferinde 5 kez değiştirilerek MÖ algoritmaları test edilmiştir. Her bir döngüde hata metrikleri hesaplanmış ve ortalaması alınmıştır (Eşitlik 3.28).

$$Hata = \frac{1}{5}\sum_{i=1}^{5} Hata_i \tag{3.28}$$

5 adet farklı MÖ algoritmaları için python programlama dilinde hazır bulunan kütüphaneler kullanılarak işlemler yapılmıştır. RO, YSA, regresyon, k-nn ve DVR MÖ algoritmaları için hazır kütüphaneler mevcuttur. Bu işlemler python 3.7 programlama dili kullanılmıştır. Python programlama dili için kodlar ise Anaconda geliştirme ortamı 2018.12 versiyon kullanılarak hazırlanmıştır.

3.8 Değer Haritaları Oluşturulması

Harita genel olarak coğrafi (mekânsal) özelliğe bağlı bilgileri içeren grafiklerdir. Harita ile mekânsal yapısal özellikler mekâna bağımlı olarak gösterilmektedir. Genel olarak haritalar içeriklerine göre topografik ve tematik haritalar olmak üzere 2 gruba ayrılabilir.

Topografik haritalar yeryüzündeki doğal ve yapay olguları gösteren haritalardır. Tematik haritalar ise insan ve doğa ile alakalı farklı konuların mekâna bağlı niceliksel özellikleri gösteren haritalardır. Bu haritalara örnek olarak yağış, trafik yoğunluğu, toprak, deprem hatları ve ulaşım vb. haritaları gösterilebilir. Değer haritaları da niceliksel özelliklerin mekâna yansıma olduğundan değer haritaları tematik harita grubuna girmektedir. Haritaların kolay anlaşılması için harita işaretleri, ölçek, lejant vb. harita yapımı sırasında konulması gerekmektedir.

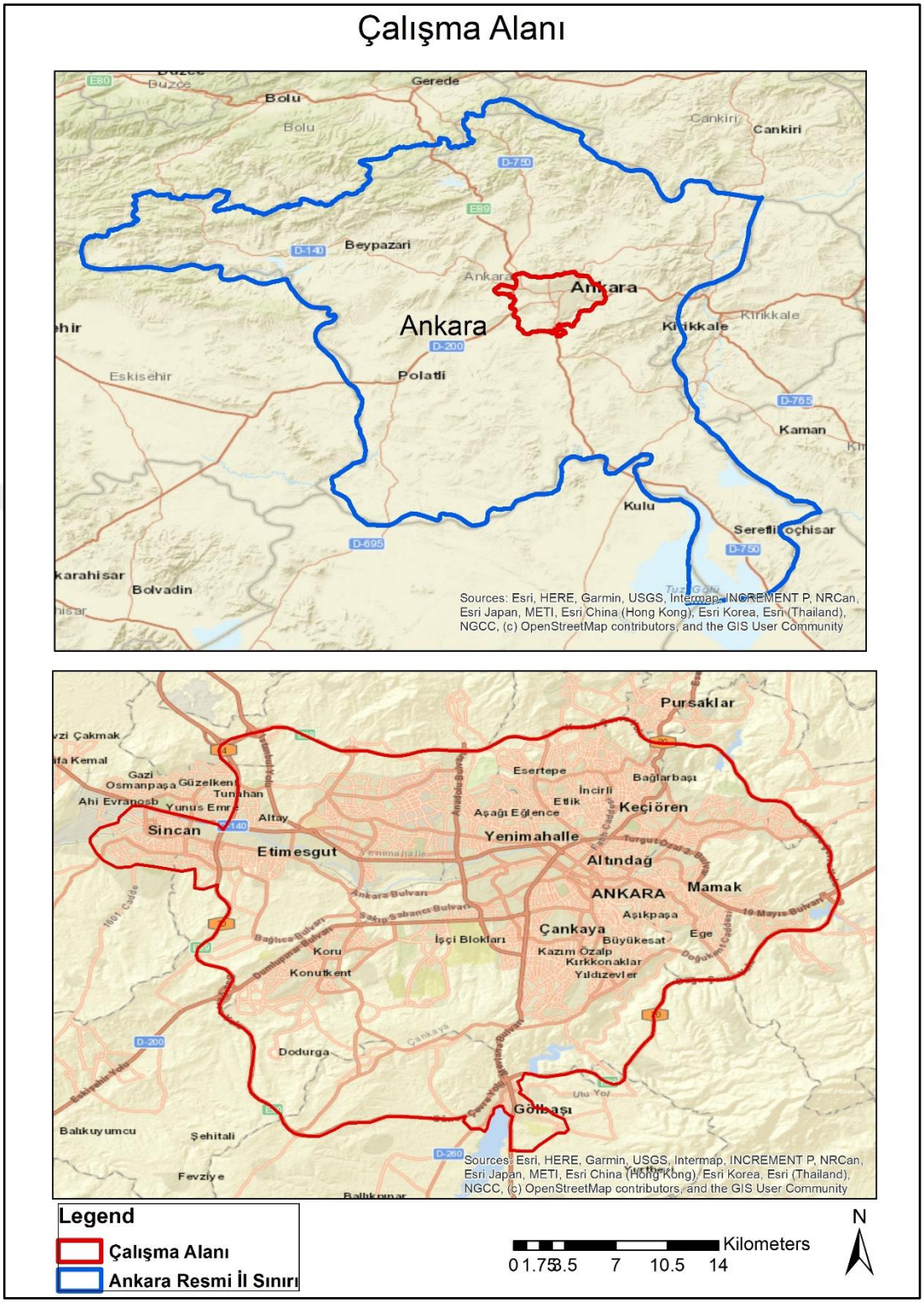
Değer haritaları denildiği zaman akla herhangi bir ürünün mekânsal olarak değerinin harita üzerinde gösterilmesi gelmektedir. Bu tez çalışmasında ise değer haritası; gayrimenkul değerini mekânsal olarak gösteren tematik bir harita olduğu belirtilmelidir. Haritalar piksel bazlı (raster) gösterilebileceği gibi nokta, poligon ve çizgisel gibi özellik türünde de gösterilebilir. Literatürde çalışmalara bakıldığı zaman değer haritaların genellikle piksel bazlı olduğu fakat gayrimenkulün değerini gösterirken gayrimenkulün sınırlarını alıp değer ile gösterildiği haritalarda bulunmaktadır.

Raster tabanlı haritalarda ise tüm pikseller bir değer alacağı için yakındaki gayrimenkul değerleri kullanılarak enterpolasyon yöntemleri ile değer haritası elde edilebilinmektedir. Enterpolasyon yöntemleri çok çeşitli olup bu çalışmada Ters Mesafe Ağırlıklandırma (Inverse Distance Weighting, IDW) yöntemi ile bilinen değerler kullanılarak bilinmeyen değerler hesaplanmış ve raster tabanlı değer haritası çalışma alanı için oluşturulmuştur.

4. ÇALIŞMA ALANI VE VERİLER

4.1 Çalışma Alanı ve Demografik Yapısı

Bu tez çalışmasında çalışma alanı Ankara İli merkez ilçeleri kapsayacak şekilde Otoyol 20 (O-20) bir diğer adıyla Ankara Çevre yolu sınırları olacak şekilde alınmıştır. Ankara İli merkez ilçeleri Altındağ, Çankaya, Etimesgut, Gölbaşı, Keçiören, Mamak, Sincan ve Yenimahalle şeklindedir. Çalışma alanı Gölbaşı, Sincan ve Etimesgut ilçelerinin ana merkezlerini diğer ilçelerin ise tamamını kapsayacak biçimde oluşturulmuştur. Ankara İli resmi idari il sınırı ve çalışma alanın sınırları dikkate alındığında çalışma alanı büyüklüğünün büyük bir alanı yani Ankara'nın merkezini kapsadığı görülmektedir (Şekil 4.1). Çalışma alanının büyüklüğü 699.03 km^2 olarak hesaplanmıştır.



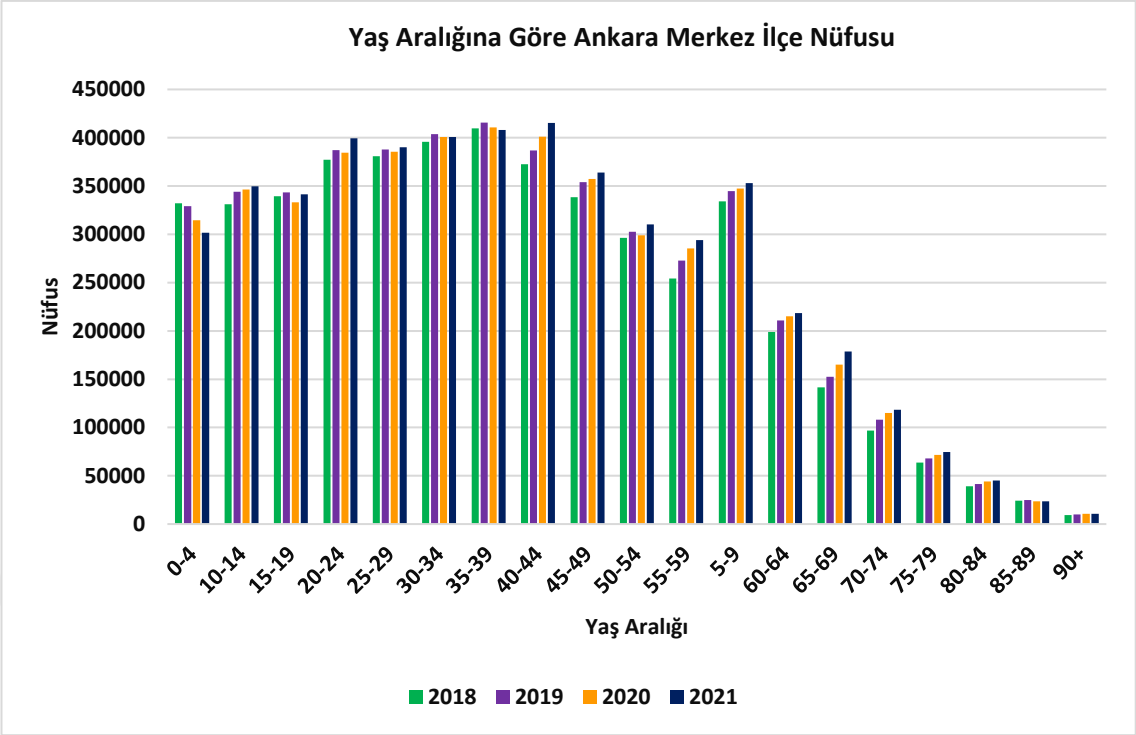
Şekil 4.1 Ankara ili resmi il sınırı ile çalışma alanı gösterimi

Ankara ili tüm ilçeler ile birlikte resmi verilere göre 2021 yılında nüfus sayısı 5.747.325 kişidir. Son yıllara göre nüfus artışı incelendiğinde 2018'den 2019'a %2.39; 2019'dan 2020'ye %0.42, ve 2020'den 2021' %1.46 kadar Ankara ilinde nüfusun arttığı görülmektedir (Şekil 4.2). Verilerden anlaşılacağı üzere Ankara ilinde çok hızlı bir nüfus artışı gözlenmemektedir.



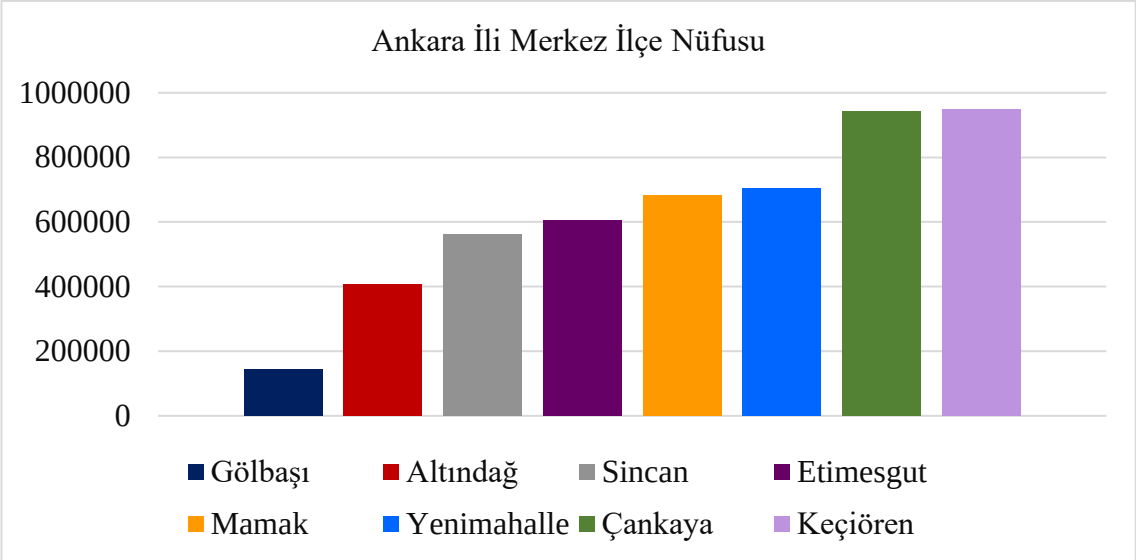
Şekil 4.2 Ankara il geneli son yıllara göre genel nüfus grafiği

Çalışma alanı merkez ilçelerin demografisi ayrıntılı olarak incelendiği zaman, merkez ilçelerde 2021 yılı resmi verilerine göre en çok 40-44 yaş aralığı grubunda kişi bulunduğu görülmektedir. Merkez ilçelerde toplam nüfusun %8.31'i 40-44 yaş aralığındadır. Bu yaş grubunu sırasıyla; 35-39 yaş aralığı ve 30-34 yaş aralığı takip etmektedir. Ankara ili merkez ilçelerde 80 yaş üstü grup ise toplam nüfusun sadece %1.58'ini oluşturmaktadır (Şekil 4.3).



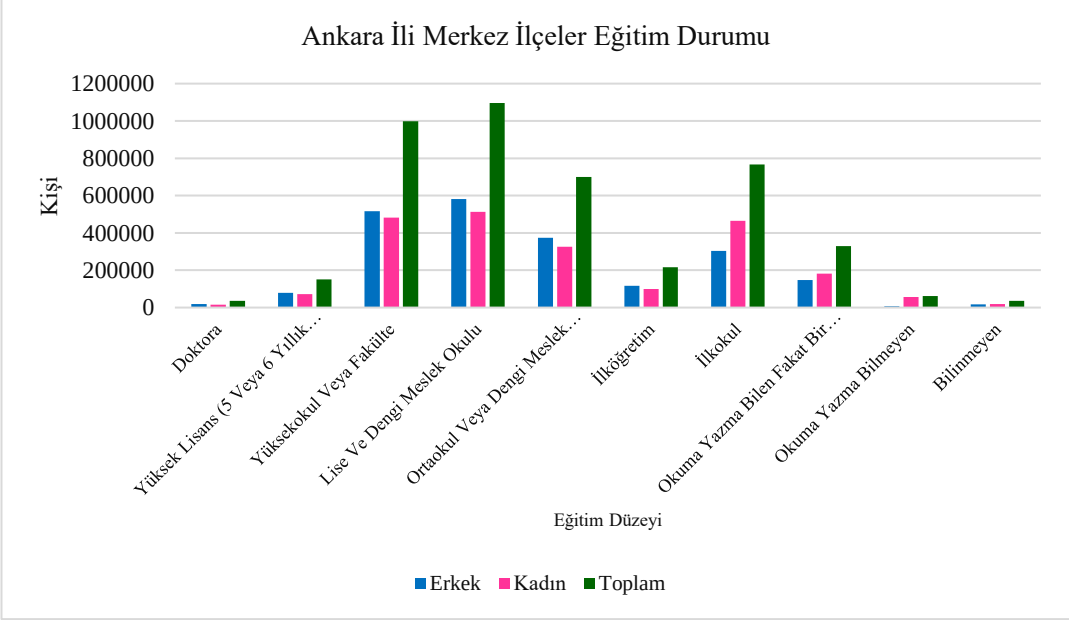
Şekil 4.3 Ankara İli merkez ilçeler yaş aralığına göre son 4 yılda nüfus dağılımı

Ankara ili Merkez ilçelerdeki kadın erkek dağılımı izlendiğinde ise %49,31'i erkek, %50,68'i ise kadınlardan oluştuğu, Ankara ili merkez ilçelerin nüfuslarına ayrı ayrı bakıldığı zaman ise 2021 yılı itibariyle en çok nüfusa sahip ilçenin Keçiören olduğu bu ilçeyi çok az farkla Çankaya'nın takip ettiği görülmektedir, Gölbaşı ilçesi ise en az nüfusa sahip merkez ilçe olarak dikkat çekmektedir (Şekil 4.4).



Şekil 4.4 Ankara İl'i merkez ilçe 2021 yılı nüfus dağılımı (TUIK,2021),

Ankara ili Merkez ilçelerde okur yazarlık ve eğitim durumu incelendiğinde ise okuma yazma bilmeyenlerin oranı %1.42 olduğu görülmektedir. Nüfusun %24,96'sı ise lise veya dengi okul bitirdiği %22,75'inin bir üniversite bitirdiği ve %4.24'ünün doktora veya yüksek lisans yaptığı tespit edilmiştir. Genel olarak Ankara ili Merkez ilçeler için eğitim seviyesinin yüksek olduğu söylenebilmektedir (Şekil 4.5).



Şekil 4.5 Ankara İli merkez ilçeler eğitim seviyeleri

4.2 Çalışmada Kullanılan Veriler

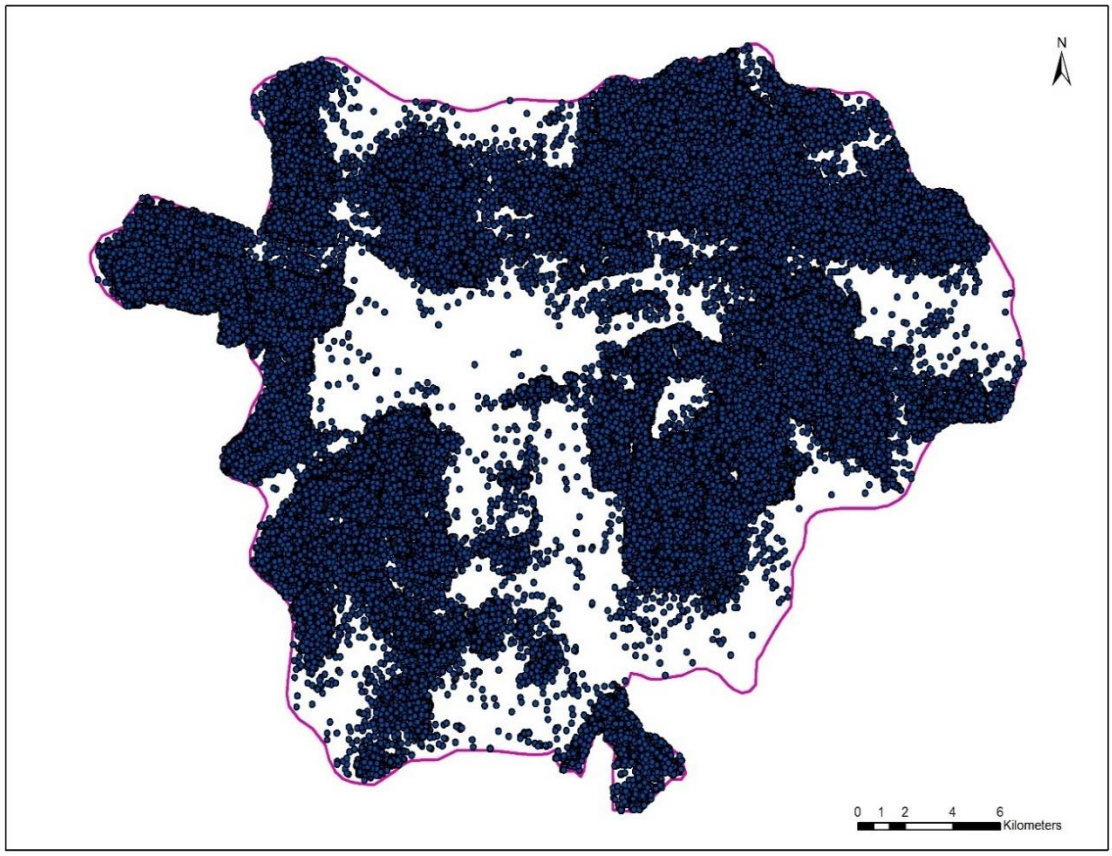
Literatürde birçok araştırmada kullanılan alım-satım değerlerine ait verilerin çoğunlukla ülkelerdeki resmî kurumlardan veya emlak sitelerinden alındığı görülmüştür. Çok az bir çalışmada ise değerleme raporlarının kullanıldığı görülmüştür, bu sayının az olmasının nedeni birçok ülkede gerçek alım-satım değerlerinin resmî kurumlara bildirilmesidir. Türkiye'de ise gayrimenkul alım-satım verileri Tapu Müdürlüklerinde tutulmakta birlikte önemli bir sorun olan gerçek alım-satım değerinin Tapu Müdürlüklerine bildirilmediği ve bu sayede satışı gerçekleştirilen tüm gayrimenkuller için tapuda beyan edilen satış bedelinin %4'ü olarak hesaplanan tapu harcından kaçınıldığı bilinen bir gerçektir. Bu sebeple, bu çalışmada kullanılacak veriler Tapu Müdürlükleri yerine daha gerçekçi değerler olan lisanslı değerleme şirket raporlarından ve ENDEKSA adlı gayrimenkul değerleme üzerine çalışan şirketin veri tabanından çekilmiştir. Bu kapsamda Ankara İl'ine ait son 5 yılın gayrimenkul değerleme raporları ile yine son 5 yılın ENDEKSA şirket konut satış verileri kullanılmıştır. Ayrıca T.C. Merkez Bankası'na tüm değerleme firmaların verileri aktardığı bilindiğinden bu kurumdan veri talep edilmiş fakat olumsuz karşılanmıştır.

Türk Medeni Kanun'unda gayrimenkul terimi yerine taşınmaz denilmiş ve bu kanunun 704'üncü maddesine göre taşınmaz mülkiyetinin konusu arazi ile tapu kütüğünde ayrı sayfaya kaydedilen bağımsız ve sürekli haklar ve kat mülkiyeti kütüğüne kayıtlı bağımsız bölümlerden oluşmaktadır. Bu çalışmada, gayrimenkul tipi olarak arsa, arazi, konut ve ticari gayrimenkuller ele alınarak farklı değişkenleri ile toplu olarak gayrimenkullerin değerlemesi yapılmıştır.

4.2.1 Konut türü gayrimenkul verileri

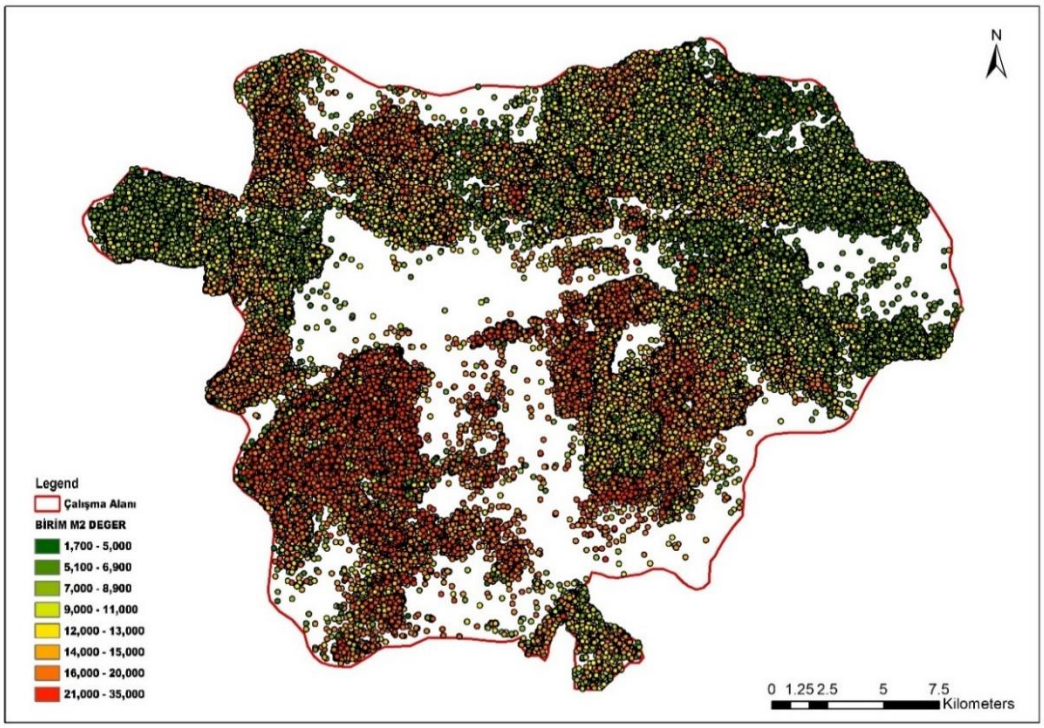
Konut türü gayrimenkul değerleri ve özellikleri Ankara İli için Endeksa şirketi veri tabanından elde edilmiştir. Toplamda 1,315,675 adet konut verisi bulunmaktadır. Veriler nokta özellikte olarak Şekil 4.6'de gösterilmiştir. Şekilden de görüldüğü üzere konut

verileri Ankara ili çalışma alanını neredeyse kapsamaktadır. Bu nedenle bu veriler çok kapsamlı olup sonuçlarının literatüre katkıda bulunması hedeflenmektedir.



Şekil 4.6 Gayrimenkul verilerin çalışma alanı içerisinde mekânsal olarak gösterimi

Çalışmada seçilen örnek alan içerisinde konut türündeki gayrimenkullerin bilinen birim m^2 değerleri yine noktasal özellikle renklendirilerek gösterilmiştir (Şekil 4.7). Yeşil noktalar düşük değere sahip, kırmızı noktalar ise yüksek değere sahip gayrimenkulleri göstermektedir.



Şekil 4.7 Gayrimenkullerin verilerdeki metre kare değerlerine göre gösterimi

Konut türüne ait 22 adet değişken bulunmakta olup bu değişkenler Çizelge 4.1'de gösterilmiştir. Gayrimenkule ait değişkenler nümerik olabileceği gibi kategorik de olabilecektir. Kategorik verilerin işleme konulabilmesi için ise sayısal bir karşılığı olması gerekmektedir.

Çizelge 4.1 Konut türü gayrimenkule ait değişkenler

No	**Değişken Adı**	**Birimi**	**Değişken Veri Türü**
1	Konut alanı	m^2	Nümerik
2	Oturma odası sayısı	Adet	Nümerik
3	Banyo sayısı	Adet	Nümerik
4	Toplam oda sayısı	Adet	Nümerik
5	Konut yaşı	Sayı	Nümerik
6	Asansör sahip olma durumu	0-1 (var-yok)	Kategorik
7	Açık otopark olma durumu	0-1 (var-yok)	Kategorik
8	Kapalı otopark olma durumu	0-1 (var-yok)	Kategorik
9	Kuzey cepheye sahip olma durumu	0-1 (var-yok)	Kategorik
10	Güney cepheye sahip olma durumu	0-1 (var-yok)	Kategorik

Çizelge 4.1 Konut türü gayrimenkule ait değişkenler (devam)

No	Değişken Adı	Birimi	Değişken Veri Türü
11	Doğu cepheye sahip olma durumu	0-1 (var-yok)	Kategorik
12	Batı cepheye sahip olma durumu	0-1 (var-yok)	Kategorik
13	Konutun bulunduğu kat	Sayı	Nümerik
14	Toplam Kat Sayısı	Sayı	Nümerik
15	Park alanlarına olan en yakın mesafe	m.	Nümerik
16	Üniversite yerleşkelerine en yakın mesafe	m.	Nümerik
17	İbadethanelere olan en yakın mesafe	m.	Nümerik
18	Okullara en yakın olan mesafe	m.	Nümerik
19	AVM lere olan en yakın mesafe	m.	Nümerik
20	Metro duraklarına en yakın olan mesafe	m.	Nümerik
21	Hastanelere en yakın olan mesafe	m.	Nümerik
22	Isıtma Tipi	10,20,30,40,50,60,70 ,80, 90,100,110,120, 130,140	Kategorik

Gayrimenkullerin ısıtma tipi için ise 14 farklı durum bulunmakta ve bu ısıtma tipleri Çizelge 4.2'de verilmiştir. Farklı ısıtma tiplerinin gayrimenkulün değere etkisi bulunduğu ve bu verilere erişim olduğu için bu türlerin hepsi toplu değerlemede değişken olarak alınmıştır.

Çizelge 4.2 Isıtma Tipi için kategorik değişkenlerin açıklamaları

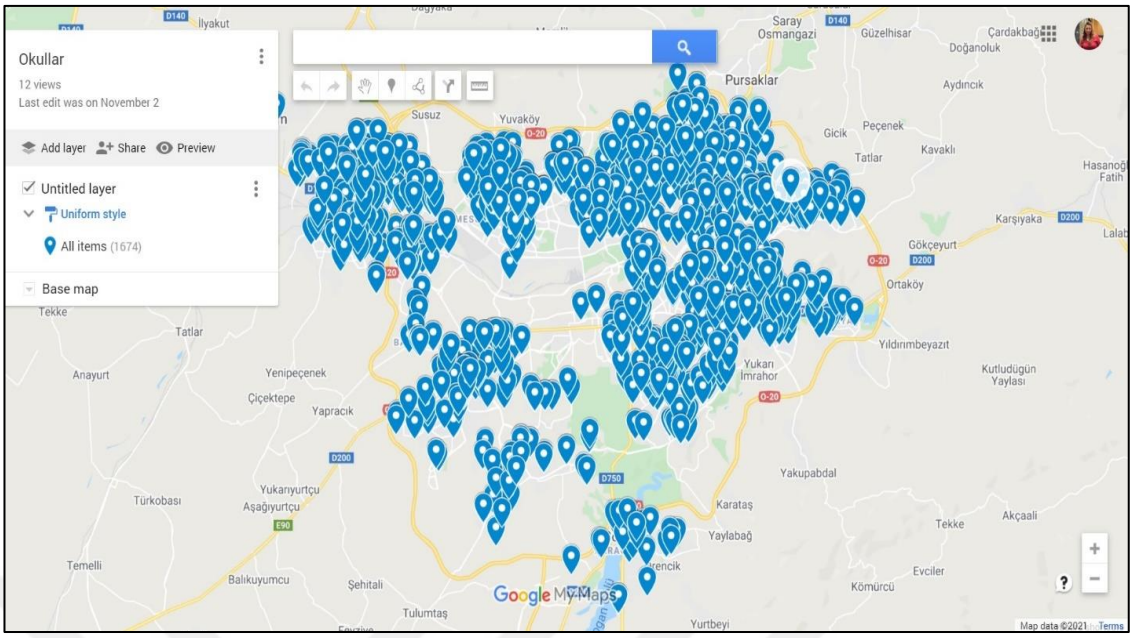
No	Kategorik Sınıf	Isınma Tipi
1	10	Yok
2	20	Soba
3	30	Doğalgaz Sobası
4	40	Kat Kaloriferi
5	50	Doğalgaz Kombi
6	60	Merkezi Sistem
7	70	Merkezi Isı Pay Ölçer
8	80	Yerden Isıtma
9	90	Klima Sistemi
10	100	Jeotermal Isınma
11	110	Güneş Enerjisi
12	120	Diğer
13	130	VRF
14	140	Fancoil

Ankara ili çalışma alanı dışında MÖ algoritmaları ayrı bir şehir ve farklı bir veri ile de denenmek amacıyla ayrıca ABD'den bir şehire ait konut türü gayrimenkul verileri de bu çalışmada test edilmiştir. Ames, Iowa şehrine ait ise 2,353 adet konut verisi ile bu gayrimenkullere ait 47 adet değişken bulunmaktadır (Çizelge 4.3) (Kaggle, 2022). Türkiye'de bulunan konutların aksine Ames, Iowa şehrinde konutların çok daha fazla ve farklı değişkene sahip olduğu görülmektedir. Bu veri setinde toplamda 2,353 adet konut verisi bulunmakta olup bu verilerden 1,881 adeti algoritmaları eğitmek için ve 471 adeti ise test etmek için kullanılmıştır.

Çizelge 4.3 Ames, Iowa şehri konut verileri için değişkenler

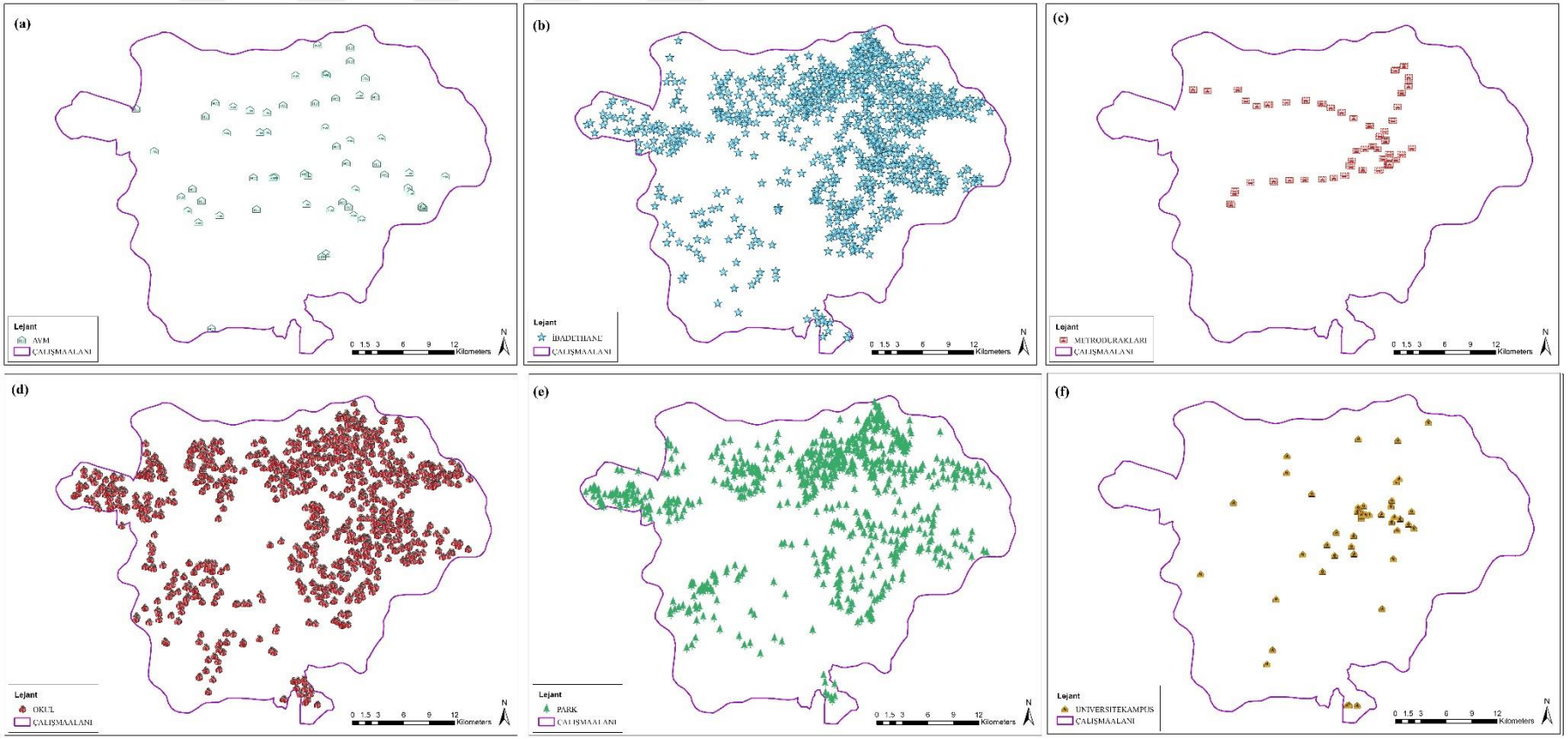
No	**Değişken**	**Değişken veri türü**	**No**	**Değişken**	**Değişken veri türü**
1	Konut satış türü	Kategorik	25	İkinci kat alan	Nümerik
2	Cephe şekli	Kategorik	26	Oturma odası alanı	Nümerik
3	Cephe alanı	Nümerik	27	Birinci kat banyo sayısı	Nümerik
4	Sokak malzemesi	Kategorik	28	Birinci kat yarım banyo sayısı	Nümerik
5	Arsa şekli	Kategorik	29	Toplam banyo sayısı	Nümerik
6	Arsa düzlüğü	Kategorik	30	Toplam yarım banyo sayısı	Nümerik
7	Parsel konumu	Kategorik	31	Yatak odası sayısı	Nümerik
8	Arsa eğimi	Kategorik	32	Mutfak sayısı	Nümerik
9	Konut türü	Kategorik	33	Mutfak durumu	Kategorik
10	Genel kalite	Kategorik	34	Toplam oda sayısı	Nümerik
11	Genel durum	Kategorik	35	Yangın söndürme durumu	Kategorik
12	Konut yaşı	Nümerik	36	Yangın alan kalitesi	Kategorik
13	Çatı türü	Kategorik	37	Garaj yeri	Nümerik
14	Çatı materyali	Kategorik	38	Garaj yapım yılı	Nümerik
15	Konut dış özelliği	Kategorik	39	Garaj araba sayısı	Nümerik
16	Dış özelliklerin durumu	Kategorik	40	Garaj alanı	Nümerik
17	Yapı maddesi	Kategorik	41	Garaj kalite	Kategorik
18	Yapı maddesinin durumu	Kategorik	42	Ahşap alan	Nümerik
19	Bodrum alanı	Nümerik	43	Açık sundurma	Kategorik
20	Isıtma tipi	Kategorik	44	Kapalı sundurma	Kategorik
21	Isıtma durumu	Kategorik	45	Mevsimlik sundurma alanı	Kategorik
22	Merkezi ısıtma	Kategorik	46	Görünen sundurma	Kategorik
23	Elektrik sistemi	Kategorik	47	Havuz alanı	Nümerik
24	Birinci kat alanı	Nümerik			

Bu çalışmada Ankara İli çalışma alanı için, gayrimenkulün değerini etkileyebilecek önemli yerler Google haritalar aracılığı ile çalışma alanı için toplanmıştır. Önemli yerler olarak okul, ibadethane, camii, üniversite yerleşkeleri, metro durakları, AVM ler ile hastaneler belirlenmiştir. Her bir değişken öncelikli olarak Google haritalar üzerinden Şekil 4.8'deki gibi işaretlenmiş ve *.kml dosyası olarak kaydedilmiştir.

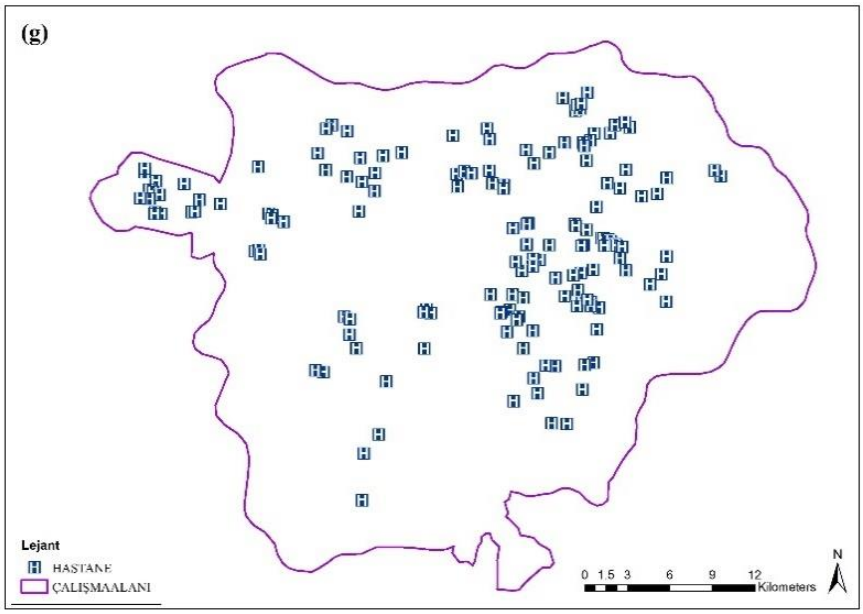


Şekil 4.8 Çalışma alanı için Google haritadan çekilen okulların konumu

Belirlenen 7 farklı önemli yer kriteri için Google haritalardan ayrı ayrı harita oluşturularak veriler çekilmiştir. Sonrasında ise bu veriler ArcMap programına aktarılarak ilerleyen şekillerde gösterilmiştir. ArcMap'e atılan veriler kullanılarak gayrimenkullere olan uzaklıkları bir sonraki aşamada her bir gayrimenkul için ayrı ayrı en yakın mesafe hesaplanmıştır. Google haritalardan çıkarılan önemli yerler ilerleyen şekillerde çalışma alanı içerisinde ayrıca gösterilmiştir (Şekil 4.9-4.10).

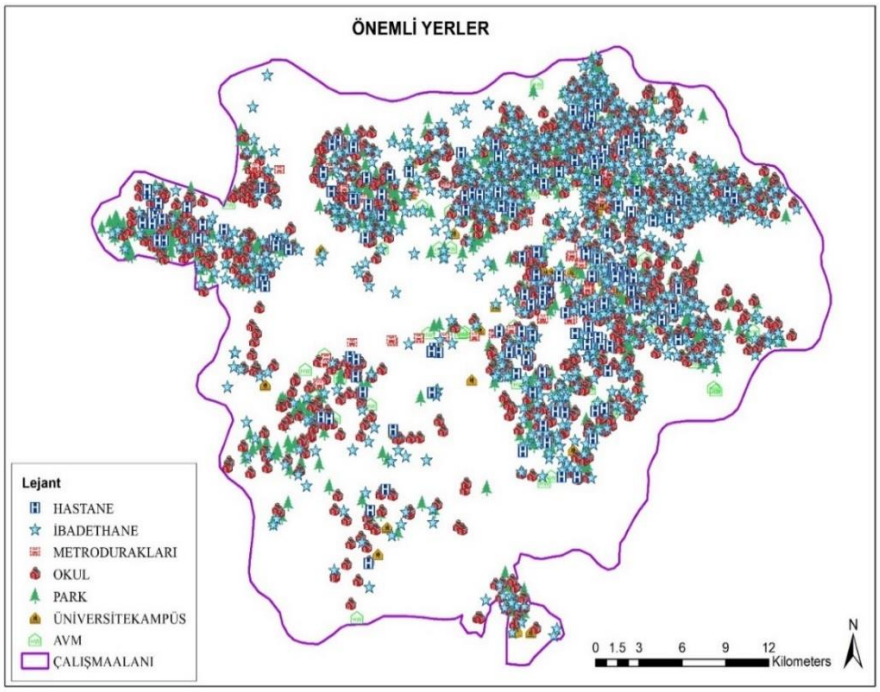


Şekil 4.9 Çalışma alanı içerisindeki önemli yerlerin gösterimi (Çalışma alanı içerisindeki, AVM (a), ibadethane (b), metro durakları (c), okul (d), park (e), üniversite (f))



Şekil 4.10 Çalışma alanı içerisindeki önemli yerlerin gösterimi hastane (g)

Konut türü gayrimenkul için Ankara ili çalışma alanında tüm önemli merkezler bir arada Şekil 4.11'de gösterilmiştir. Tüm bu önemli yerler için en yakın mesafe her bir konut için otomatik olarak CBS programı olan ArcMap ile hesaplanmıştır.



Şekil 4.11 Önemli yerlerin çalışma alanı içerisinde bir arada gösterimi

Yukarıda bahsedilen önemli yerler toplandıktan sonra konut türüne ait gayrimenkullerin değişkenler ve değerlerinin çok küçük bir kısmı örnek olarak verilmiştir (Çizelge 4.4). Örneğin 1 numaralı gayrimenkulün alanı 112 m^2 yaşı 4, oda sayısı 3, oturma oda sayısı 1, banyo sayısı 1, bulunduğu kat 1 ve toplam kat sayısı 6, kuzey cepheye sahip asansörlü ve merkezi sistem kaloriferli ibadethaneye yakın bir konuttur.

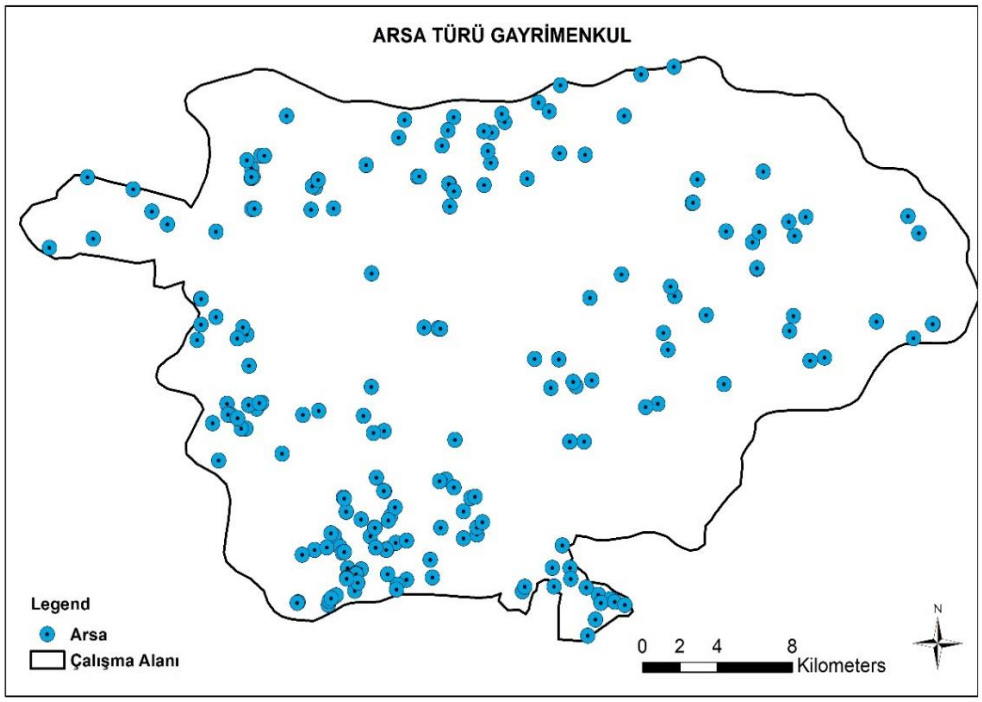
Çizelge 4.4 Konut türü gayrimenkul için örnek veri seti

	ID	1	2	3	4	5	6	7
1	Alan	112	80	142	90	95	165	190
2	Yaş	4	0	8	0	3	8	0
3	Oda sayısı	3	2	3	2	2	4	4
4	Oturma odası	1	1	1	1	1	1	1
5	Banyo sayısı	1	1	1	1	1	2	2
6	Toplam kat sayısı	6	3	6	3	8	7	7
7	Kat sayısı	1	1	4	1	1	6	1
8	Kuzey cephe	0	0	0	0	0	0	1
9	Güney cephe	0	0	1	1	0	0	1
10	Doğu cephe	0	0	1	0	0	0	1
11	Batı cephe	0	0	0	0	0	0	1
12	Asansör	1	1	0	0	0	1	1
13	Kapalı otopark	0	0	0	0	0	1	1
14	Isıtma	60	50	50	50	70	70	60
15	Açık otopark	0	0	0	0	0	0	1
16	$Mesafe_{Üniversite}$ (m)	1,618.38	1,228.90	1,228.84	1,228.84	1,325.34	1,024.8	1,389.21
17	$Mesafe_{Park}$ (m)	1,872.40	1,034.84	1,013.82	1,013.82	1,071.97	393.88	892.36
18	$Mesafe_{Okul}$ (m)	1,302.50	904.94	907.58	907.58	1,004.26	446.56	840.57
19	$Mesafe_{Metro}$ (m)	1,484.14	1,348.98	1,369.74	1,369.74	1,396.72	1361.1	1157.3
20	$Mesafe_{İbadethane}$ (m)	425.20	1,065.29	1,054.07	1,054.07	1,086.48	462.17	832.42
21	$Mesafe_{Hastane}$ (m)	6,836.33	6,477.20	6,465.52	6,465.52	6,375.30	1,252.62	6,223.13
22	$Mesafe_{AVM}$ (m)	6,662.89	6,189.25	6,174.21	6,174.21	6,089.95	6,209.8	5,899.22
	Değer (TL/m^2)	4,278	6,016	3,325	6,155	5,154	6,546	6,431

4.2.2 Arsa türü gayrimenkul verileri

Arsa türü gayrimenkul için SPK tarafından değerleme raporlarından faydalanılmıştır. Konut türü gayrimenkulün aksine diğer tür gayrimenkuller için çok fazla değerleme raporu bulunmamaktadır. Değerleme raporlarından toplam 209 adet gayrimenkul ve bu

gayrimenkullerin 11 adet değişkeni kullanılmıştır. Değişkenler ve türleri Çizelge 4.5'te gayrimenkullerin çalışma alanı içerisindeki dağılımı ise Şekil 4.12'de gösterilmiştir.



Şekil 4.12 Arsa türü gayrimenkullerin çalışma alanı içerisindeki dağılımı

Konut türünün aksine arsa türü gayrimenkullerde daha az değişken bulunması beklenen bir sonuç olup, bu sebeple arsa türü gayrimenkullerde sadece 11 adet değişken ile toplu değerleme yapılmıştır.

Çizelge 4.5 Arsa türü gayrimenkule ait değişkenler

No	Değişken Adı	Birimi	Değişken Veri Türü
1	Arsa alanı	$m.^2$	Nümerik
2	Tapu sayısı	1-2-3	Nümerik
3	Takyidat olup olmama durumu	1-0 (var-yok)	Kategorik
4	Ekspertiz Sonucu	Alıcısı Az, Satılabilir, Satılamaz (1-2-3)	Kategorik
5	Park alanlarına olan en yakın mesafe	m.	Nümerik
6	Üniversite yerleşkelerine en yakın mesafe	m.	Nümerik

Çizelge 4.5 Arsa türü gayrimenkule ait değişkenler (devam)

No	Değişken Adı	Birimi	Değişken Veri Türü
7	İbadethanelere olan en yakın mesafe	m.	Nümerik
8	Okullara en yakın olan mesafe	m.	Nümerik
9	AVM lere olan en yakın mesafe	m.	Nümerik
10	Metro duraklarına olan en yakın mesafe	m.	Nümerik
11	Hastanelere en yakın olan mesafe	m.	Nümerik

4.2.3 Arazi türü gayrimenkul verileri

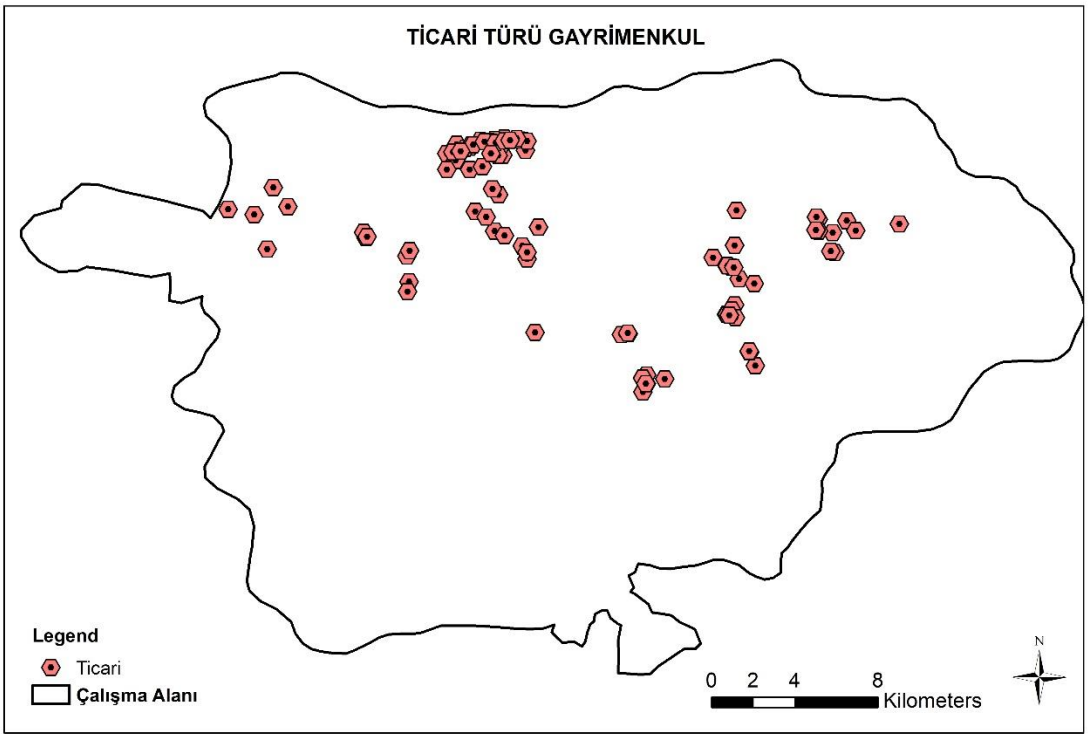
Konut ve arsa türü aksine gayrimenkullerin aksine arazi tür gayrimenkullerde değerleme rapor sayısı oldukça az olduğundan bu çalışmada sadece 107 adet arazi türü gayrimenkul bulunmaktadır. Bu gayrimenkullere ait 12 adet değişken bulunmakta olup bu değişkenler ise Çizelge 4.6'da gösterilmiştir.

Çizelge 4.6 Arazi türü gayrimenkule ait değişkenler

No	Değişken Adı	Birimi	Değişken Veri Türü
1	Arsa alanı	$m.^2$	Nümerik
2	Tapu sayısı	1-2-3	Nümerik
3	Takyidat olup olmama durumu	1-0 (var-yok)	Kategorik
4	Ekspertiz Sonucu	Alıcısı Az, Satılabilir, Satılamaz (1-2-3)	Kategorik
5	Vasfı	Tarla, bağ bahçe, evli tarla (1,23,)	Kategorik
6	Park alanlarına olan en yakın mesafe	m.	Nümerik
7	Üniversite yerleşkelerine en yakın mesafe	m.	Nümerik
8	İbadethanelere olan en yakın mesafe	m.	Nümerik
9	Okullara en yakın olan mesafe	m.	Nümerik
10	AVM lere olan en yakın mesafe	m.	Nümerik
11	Metro duraklarına en yakın olan mesafe	m.	Nümerik
12	Hastanelere en yakın olan mesafe	m.	Nümerik

4.2.4 Ticari türü gayrimenkul verileri

Ticari türü gayrimenkul olarak çalışma alanında 109 adet değerleme raporu hazırlanmış ve bu gayrimenkuller kullanılarak ticari türü gayrimenkul için toplu değerleme yapılmıştır. Ticari gayrimenkullerin çalışma alanı içerisindeki dağılımı Şekil 4.13'te gösterilmiştir.



Şekil 4.13 Ticari türü gayrimenkullerin çalışma alanı içerisindeki dağılımı

Ticari tür gayrimenkullere ait 14 adet değişken bulunmakta olup ve bu değişkenler Çizelge 4.7'de gösterilmiştir. Bu değişkenlerden çoğu nümerik türde olup bazıları kategorik formattadır. Kategorik formattaki veriler de sınıflandırılarak MÖ algoritmaları için kullanıma hazır hale getirilmiştir.

Çizelge 4.7 Ticari türü gayrimenkule ait değişkenler

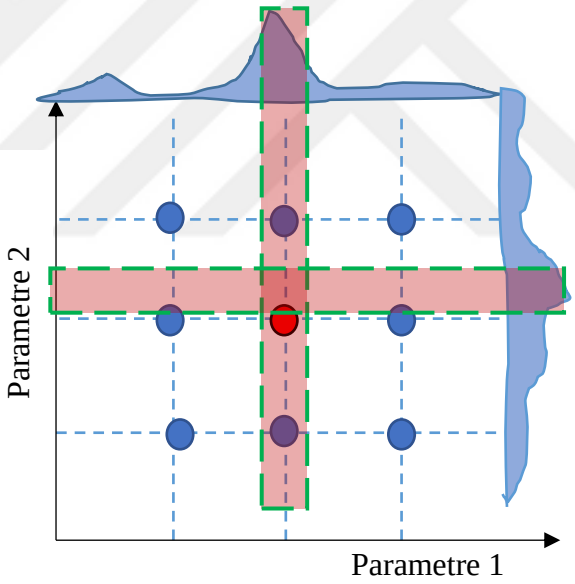
No	Değişken Adı	Birimi	Değişken Veri Türü
1	Alanı	m^2	Nümerik
2	Fiili Kullanım Türü	Büro-Dükkân-Depo-Sanayi Tesisi- Ticari Bina (1-2-3-4-5)	Kategorik
3	Ekspertiz Sonucu	Alıcısı Az, Satılabilir, Satılamaz (1-2-3)	Kategorik
4	Fiili kullanım niteliği	Büro-Depo-Dükkân-küçük sanayi (1-2-3-4)	Kategorik
5	Üniversite yerleşkelerine en yakın mesafe	m.	Nümerik
6	Ana taşınmazın niteliği	Kargır- İşhanı (1-2)	Kategorik
7	Projeye uygun mu	Evet-kısmen-hayır (1-2-3)	Kategorik
8	Takyidat var mı	Evet-hayrı (1-2)	Kategorik
9	Ekspertiz kanaati	Satılmaz-alıcısı az-satılır (1-2-3)	Kategorik
10	İbadethanelere olan en yakın mesafe	m.	Nümerik
11	Okullara en yakın olan mesafe	m.	Nümerik
12	AVM lere olan en yakın mesafe	m.	Nümerik
13	Metro duraklarına en yakın olan mesafe	m.	Nümerik
14	Hastanelere en yakın olan mesafe	m.	Nümerik

4.3 Makine Öğrenme Algoritmalarında Parametre Seçimi

MÖ algoritmalarında kullanılan iki yaygın parametre seçim yöntemi grid search diye anılan ızgara arama ile random search diye adlandırılan rasgele arama yöntemleridir (Lerman 1980). Bu çalışmada daha optimum sonuçlar verdiği bilinen ızgara arama yöntemi kullanılarak parametre kestirimi yapılmıştır (Chen vd. 2022). Rastgele arama, ızgara aramaya benzemekle birlikte ızgaradaki tüm noktaları kullanmak yerine, bu noktaların yalnızca rastgele seçilen bir alt kümesini test etmektedir. Bu alt küme ne kadar küçükse, optimizasyon o kadar hızlı ancak daha az doğru olmaktadır.

Izgara arama verilen model için her hiper parametreyi yineler. Bu arama işleminin sonunda minimum maliyeti veren hiper parametre seti, modelinin optimize edilmiş hiper parametreleri olarak işaretlenmektedir.

Izgara arama, belirli bir model için en uygun parametreleri yapılandırmak için verileri tarama işlemidir. Izgara arama sadece bir model tipi için geçerli değil herhangi bir model için kullanılacak en iyi parametreleri hesaplamak için makine öğrenimi genelinde ızgara arama uygulanabilmektedir. Izgara arama mümkün olan her parametre kombinasyonu üzerinde bir model oluşturmaktadır. Her parametre kombinasyonunu yineler ve her kombinasyon için bir model oluşturarak en başarılı modeli için parametreleri belirlemektedir (Şekil 4.14).



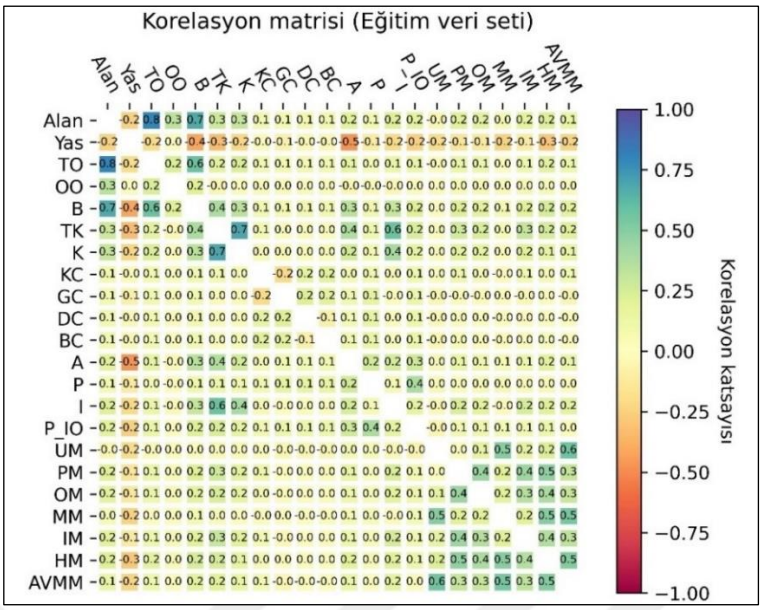
Şekil 4.14 Izgara arama parametre seçimi gösterimi

5. UYGULAMA SONUÇLARI

Konut, ticari, arsa ve arazi verileri türündeki gayrimenkuller için sırayla Rasgele Orman, Yapay Sinir Ağları, Destek Vektör Makineleri, K En Yakın Komşuluk ve Çoklu Regresyon Analizi olmak üzere 5 MÖ yöntemi uygulanmış ve sonuçları ilerleyen alt başlıklarda ayrıntılı olarak incelenmiştir. Temel Bileşen Analiz Yöntemi ise değişken sayısını arasındaki korelasyon ile değişken sayısını azaltarak yeni değişkenler elde edilmek için kullanılmıştır. Bu tez çalışmasında toplu değerleme için karşılaştırılan 5 adet MÖ öğrenme yöntemi sadece konut türü gayrimenkul için hem gayrimenkullere ait 22 değişkenleri kullanılarak hem de TBA ile oluşturulmuş yeni değişkenler kullanılarak hesaplanmıştır. Böylece yöntemler hem birbiriyle kıyaslanmış hem de TBA algoritmasının yöntemler üzerindeki etkisi araştırılmıştır. Diğer gayrimenkul türleri olan arazi, arsa ve ticari türdeki gayrimenkuller için değişken sayısı az olduğu için TBA analizi bu türlerdeki gayrimenkullerde hesaplanmamıştır.

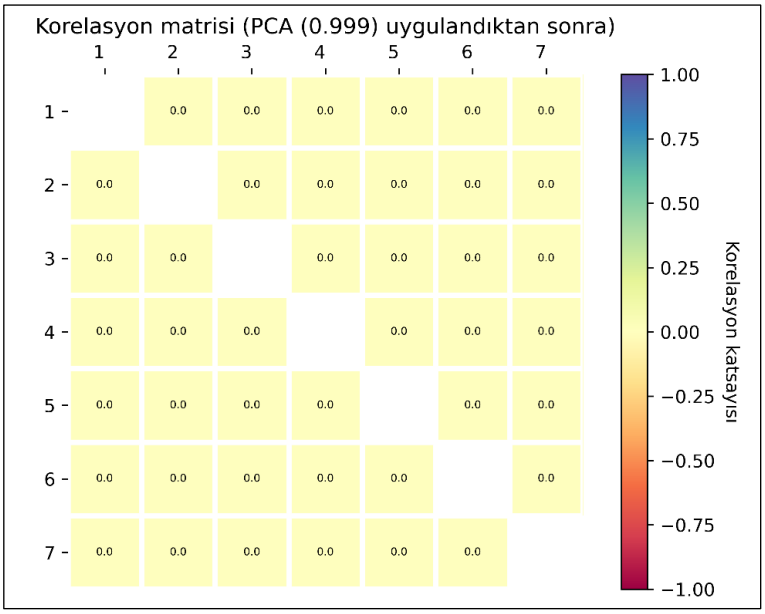
5.1 TBA Analizi Sonuçları

TBA analizi en çok değişken sayısı konut türüne ait gayrimenkulde olduğu için bu tür için yapılmıştır. Diğer gayrimenkul türleri olan arsa, arazi ve ticari için ise TBA analizi değişken sayısının az olması sebebiyle uygulanmamıştır. Konut türü gayrimenkul için toplanan 22 adet değişken TBA algoritması kullanılarak 7'ye düşürülmüş ve MÖ algoritmaları tekrar koşturularak sonuçlar irdelenmiştir (Şekil 5.1 ve Şekil 5.2). Konut türü gayrimenkul değişkenleri arasındaki korelasyon incelendiğinde aralarında çok fazla korelasyon olmadığı gözlemlenmiştir. Burada TO; toplam oda sayısı, OO; oturma oda sayısı, B; banyo sayısı, TK; Toplam kat sayısı, K; kaçıncı kat olduğu, KC; Kuzey cephe, GC; Güney cephe, DC; Doğu cephe, BC; Batı cephe, A; Asansör, P; Park, I; Isıtma, P_IO; Kapalı otopark, UM; Üniversite yerleşkelerine olan mesafe, PM; Parka olan mesafe, OM; Okula olan mesafe, MM; Metroya olan mesafe, IM; İbadethanelere olan mesafe, HM; Hastaneye olan mesafe, AVMM; AVM'ye olan mesafesini göstermektedir.



Şekil 5.1 Konut türü gayrimenkul için korelasyon matrisi

TBA analizinden sonra 22 adet değişken sayısı 7'ye düşmüş böylece değişkenler arasında düşük olan korelasyon tamamen kalkmıştır (Şekil 5.2) ve bu durumun MÖ algoritması üzerindeki etkileri bu çalışma kapsamında araştırılmıştır. TBA analizinin sonucu azalan değişkenler ile toplu değerlemenin daha başarılı sonuç vermediği görülmüş ve ilerleyene bölümlerde ayrıntılı anlatılmıştır.



Şekil 5.2 Konut türü gayrimenkul değişkenler arası korelasyon

5.2 Konut Türü Gayrimenkul İçin Toplu Değerleme Sonuçları

Toplamda 1,315,675 adet konut verisinden %80'i eğitim ve %20'si ise test verisi olarak kullanılmıştır. Yani tüm MÖ algoritmaları için konut türündeki 1,315,675 adet gayrimenkulün 1,052,540 adeti MÖ algoritmalarını test etmek ve modeli oluşturmak için, 263,135 adeti ise test için kullanılmıştır.

Daha önce bahsedildiği üzere MÖ algoritmalarında parametre kestirimi için ızgara arama yöntemi kullanılmış ve bu yöntem sonucu elde edilen parametreler Çizelge 5.1'de gösterilmiştir. Konut türü gayrimenkul için ızgara arama ile denenen ve ızgara arama sonucunda elde edilen optimum parametreler her bir MÖ algoritması için verilmiştir (Çizelge 5.1).

Çizelge 5.1 Konut türü gayrimenkul için MÖ algoritmaları hiper parametre seçimi

Algoritma	***Izgara Arama için Parametreler***	***Kullanılan Parametreler***
YSA	Gizli katman sayısı: 44,66,88 Aktivasyon fonksiyonu: Identity, logistic, tanh, relu Alpha: 0.001,0.0001,0.00001 Öğrenme hızı: 0.01,0.001,0.0001	Gizli katman sayısı: 66 Aktivasyon fonksiyonu: relu Alpha: 0.001 Öğrenme hızı 0.01
K-nn	N- komşuluk: 4,8,12,16,20,24,28,32 Uzaklık: euclidean, manhattan	N-komşuluk:8 Uzaklık: euclidean
DVR	Kernel: linear, poly, rbf, sigmoid Gamma: scale, auto	Kernel: linear Gamma: scale
RO	N_estimators (Ağaç sayısı): 50,100,150,200,250	N_estimators: 200

- *Konut Türü Gayrimenkul İçin Rasgele Orman Algoritması Sonuçları*

RO algoritmasında öncelikle 22 adet değişken ile algoritma test edilmiş sonrasında ise TBA algoritması kullanılarak bu 22 adet değişken sayısı değişkenler arasındaki bağlantılarla 7 adet değişkene indirgenerek tekrar RO algoritması konut türü gayrimenkul için test edilmiştir. 5 adet uyum iyiliği ölçütü ile sonuçlar irdelenmiştir bu ölçütler, R^2, Düzeltilmiş R^2, Hata Kareler Ortalaması (HKO), Kök Ortalama Kare Hata (KOKH) ve

Ortalama Mutlak Hata (OMH) olarak hesaplanmıştır. Sonuçlar Çizelge 5.2'de gösterilmiştir. Sonuçlar irdelendiğinde özellikle R^2'nin oldukça yüksek 1'e yakın olduğu görülmekte ve bu durum modelin başarısının iyi olduğunu göstermektedir. Diğer bir ifadeyle %80 üzerinde test verisinde doğruluk bulunmuş ve bu gayrimenkul değerleme işlemlerinde kabul edilebilir bir değer olarak görülmektedir. Diğer yandan TBA algoritması kullanılarak değişken sayısı azaltılıp tekrar RO algoritması için koşturulduğunda, TBA algoritması ile elde edilen değişkenlerin tüm MÖ algoritmalarında sonucu iyi yönde etkilemediği görülmüştür. Test verisinde TL/m^2 değerleri tüm test verisindeki gayrimenkul için bakıldığında en düşük değerin 833 en yüksek değerin ise yaklaşık 64,000 TL olduğu görülmüştür. Bu durumda OMH değeri incelendiğinde değerin oldukça düşük çıktığı görülmekle beraber daha ayrıntılı değerleme için hata histogramı elde edilmiştir.

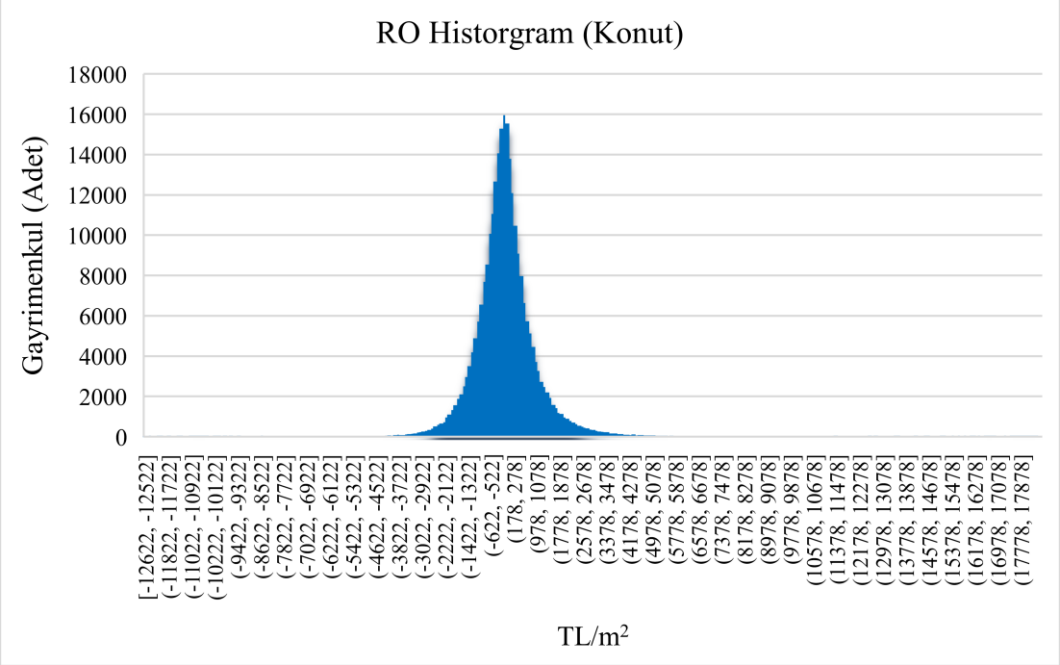
Çizelge 5.2 Konut türü gayrimenkul için RO algoritması sonuçları

	22 adet değişken ile elde edilen sonuçlar	TBA sonucu 7 değişken ile elde edilen sonuçlar
R^2	0.8126	0.5922
Düzeltilmiş R^2	0.8125	0.5921
HKO	1,409,005	3,046,855
KOKH	1,187	1,745
OMH	763	1,233

Sonuçlar ayrıntılı olarak incelendiğinde toplam test için kullanılan değerlerden %54'ünde RO algoritması daha düşük %46'sında ise olması gerekenden daha yüksek değer bulmuştur. Yani algoritma tarafından 263,135 adet konuttan 143,921 adeti düşük değerli, 119,214 adet tahmin edilen değer ise olması beklenen değerden yüksek değerli bulunmuştur.

263,135 adet konut incelendiği zaman en büyük farkın birim m^2 değer üzerinde 18,577 TL en düşük farkın ise 0 TL yani birebir aynı sonucu bulduğu görülmüştür. Ayrıca RO algoritması ile değeri en çok ve en az etkileyen değişkenler belirlenmiştir. RO sonuçlarına göre değeri en çok etkileyen ilk 5 değişken sırasıyla; ısıtma türü, üniversiteye olan mesafe, metro duraklarına olan mesafe, AVM ye olan mesafe ve toplam kat adedir. Konut

türü gayrimenkulde değeri en az etkileyen değişkenler ise kuzey, güney, batı ve doğu cephe olma durumu ile oturma odası sayısıdır. RO algoritması için hata histogramı oluşturulmuş ve hataların çoğunluğunun ortalama 50 TL gibi az bir fark aralığında toplandığı görülmüştür (Şekil 5.3).



Şekil 5.3 Konut türü gayrimenkul için RO algoritması hata histogramı

Test verisindeki sonuçlar incelendiğinde tahmin edilen değer ile bilinen değer arasındaki en fazla farkın 0-500 TL olduğu konut sayısının 129,520 ve en fazla 0-1,000 TL arasında olan konut sayısı ise 198,154 adet olduğu görülmektedir. Yani toplam 263,135 adet konut verisinden %75,11'inden RO algoritması sonucu tahmin edilen değer ile bilinen değer arasında en fazla 1,000 TL fark bulunmakta ve bu durum RO algoritmasının toplu değerlemede kabul edilebilir bir başarısı olduğunu göstermektedir.

- *Konut Türü Gayrimenkul İçin Yapay Sinir Ağları Sonuçları*

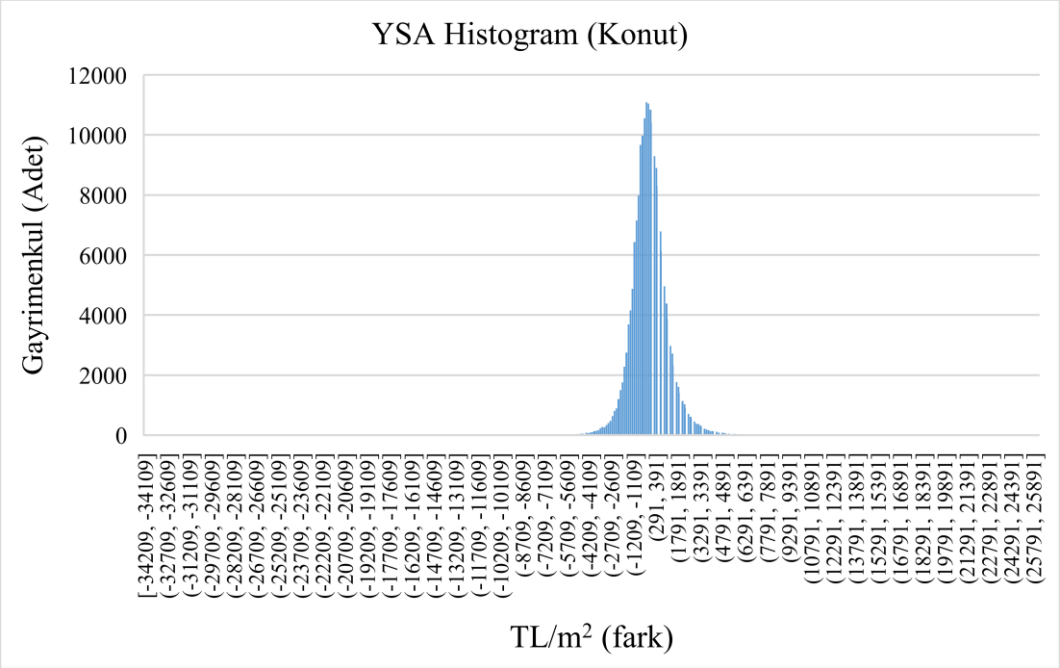
YSA algoritması ile konut türü için toplu gayrimenkul değerleme için denenmiş ve Çizelge 5.3'te elde edilen değerler için uyum iyiliği ölçütleri hesaplanmıştır. R^2 değerinin RO algoritması kadar olmasa da YSA algoritmasının da toplu değerleme modelini iyi

tahmin ettiğini göstermektedir. Ayrıca OMH incelendiği zaman, yaklaşık 900 olduğunu görüyoruz bu değer modelin başarılı sonuçlar verdiğini göstermektedir. Gayrimenkul değerleme işlemlerinde hata oranı %75'e kadar kabul edilebilmektedir. Bu nedenle konut türü büyük veri ile yapılan toplu değerleme sonucunda YSA algoritmasının kullanabilir olduğu ortaya çıkmıştır.

Çizelge 5.3 Konut türü gayrimenkul için YSA algoritması sonuçları

	22 adet değişken ile elde edilen sonuçlar	TBA sonucu 7 değişken ile elde edilen sonuçlar
R^2	0.7617	0.3207
Düzeltilmiş R^2	0.7615	0.3207
HKO	1,791,277	5,074,719
KOKH	1,338	2,252
OMH	922	1,600

Sonuçlar ayrıntılı olarak incelendiğinde test verisi ile tahmin edilen değerlerden %52'sinde YSA algoritması daha düşük %48'inde ise beklenen değerden daha yüksek değer bulduğu görülmüştür. Yani algoritma tarafından 263,135 adet konuttan 136,215 adeti düşük değerli, 126,920 adeti ise yüksek değerli bulunmuştur. 263,135 adet konut incelendiği zaman en büyük farkın birim m^2 değer üzerinde 34,209 TL en düşük farkın ise 0,01 TL yani 1 kuruş olarak bulduğu görülmüştür. Çok yüksek fark çıkan veriler olduğu için hata histogramı oldukça yaygın gözükmektedir. Tahmin edilen değer ile bilinen değer arasındaki farkın yüksek olduğu (10,000 TL) gayrimenkul sayısı ise yalnızca toplam test verisi içinde %0.06'dır.



Şekil 5.4 Konut türü gayrimenkul için YSA algoritması hata histogramı

Test verisindeki sonuçlar incelendiğinde YSA algoritması ile tahmin edilen değer ile bilinen değer arasındaki en fazla farkın 0-500 TL olduğu konut sayısı 103,000 en fazla 0-1,000 TL arasında olan konut sayısı ise 177,402 adettir. Diğer bir ifadeyle, toplam 263,135 adet konut verisinden %67,41'inden YSA algoritması sonucu tahmin edilen değer ile bilinen değer arasında en fazla 1,000 TL fark bulunmaktadır.

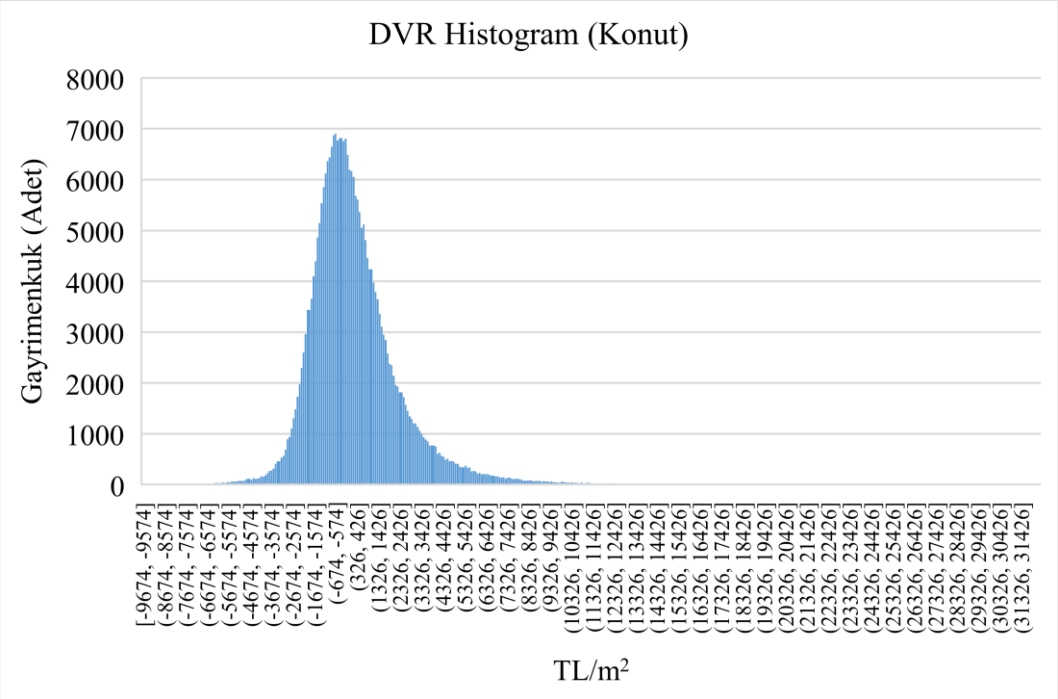
- *Konut Türü Gayrimenkul İçin Destek Vektör Regresyonu Sonuçları*

Konut türü gayrimenkul toplu değerleme için DVR algoritması denenmiş olup Çizelge 5.4'den de görüleceği üzere algoritma sonucunda başarısız sonuç elde edilmiştir. Hem model başarı ölçütü olan R^2 hem de tahmin edilen sonuçlar ile bilinen sonuçların kıyaslanmasının sonucu bulunan HKO, KOKH ve OMH metrikleri DVR algoritmasının konut türü gayrimenkulde başarısız olduğunu göstermektedir.

Çizelge 5.4 Konut türü gayrimenkul için DVR algoritması sonuçları

	22 adet değişken ile elde edilen sonuçlar	TBA sonucu 7 değişken ile elde edilen sonuçlar
R^2	0.2997	0.2815
Düzeltilmiş R^2	0.2997	0.2815
HKO	5,231,685	6,332,585
KOKH	2,287	3,158
OMH	1,532	1,958

Şekil 5.5'te DVR algoritması kullanılarak elde edilen gayrimenkul değerleri ve bilinen değerler arasındaki fark histogram şeklinde çizdirilmiştir. Histogramdan da görüleceği üzere DVR birçok değer için büyük farklar ve hatalı sonuçlar bulmuştur. Özet olarak bu konut verisinde DVR algoritmasının çalışmadığı görülmüştür. Bir diğer DVR algoritması ile not edilmesi gereken durum ise DVR algoritmasının aynı test verisi içinde en uzun hesaplama zamanına sahip olmasıdır.



Şekil 5.5 Konut türü gayrimenkul için DVR algoritması hata histogramı

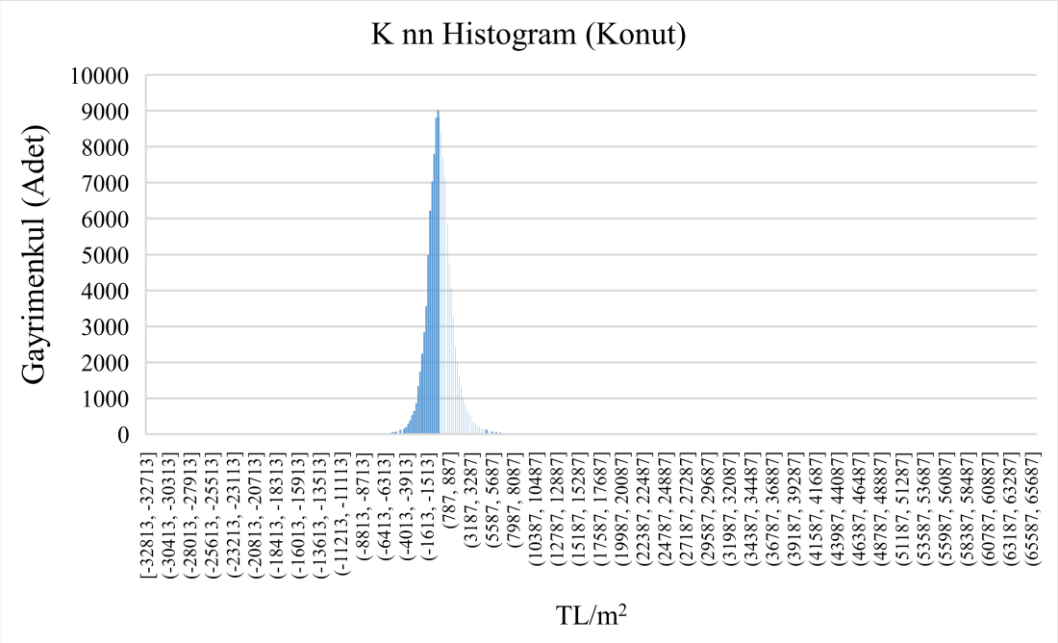
- *Konut Türü Gayrimenkul İçin K En Yakın Komşuluk Algoritması Sonuçları*

K en yakın komşuluk algoritması ile konut türü gayrimenkul için eğitim verisi ile elde edilen model ile test verisi için tahmin edilmiş değerler sonucu uyum iyiliği ölçütleri Çizelge 5.5'te tabloda verilmiştir. Bu sonuçlara göre TBA algoritması sonucu elde edilen değişkenler ile normal sayıdaki değişkenler ile elde edilen değerlerin daha başarısız olduğu gözükse de aralarında çok fazla bir farkın olmadığı gözlemlenmiştir. Genel olarak k-nn algoritması RO ve YSA algoritmalarından sonra konut türü gayrimenkul toplu değerlemede en başarılı sonucu veren 3. algoritma olmuştur. R^2 değeri 0.6594 olarak bulunmuş ve kabul edilmeye yakın bir değer olarak görülmüştür.

Çizelge 5.5 Konut türü gayrimenkul için k-nn algoritması sonuçları

	22 adet değişken ile elde edilen sonuçlar	TBA sonucu 7 değişken ile elde edilen sonuçlar
R^2	0.6594	0.5966
Düzeltilmiş R^2	0.6593	0.5962
HKO	2,544,544	3,013,387
KOKH	1,595	1,735
OMH	1,111	1,222

K-nn algoritması sonuçları ayrıntılı incelendiği zaman, 139,441 adet test verisi beklenen değerden daha düşük değerli ve 123,694 adet test verisi ise daha yüksek değerli olarak bulunmuştur. Şekil 5.6'da ise k-nn algoritması ile bulunan gayrimenkul değerlerinin bilinen ve tahmin edilen arasındaki farkın histogramı gösterilmektedir. 152,507 adet k-nn algoritması ile tahmin edilen konut verisinde ise en fazla fark 1,000 TL olarak gözlemlenmiştir.



Şekil 5.6 Konut türü gayrimenkul k-nn algoritması için hata histogramı

- *Konut Türü Gayrimenkul İçin Çoklu Regresyon Analizi Sonuçları*

ÇRA algoritması için 3 farklı model denenmiş olup bu modeller sırasıyla; doğrusal, yarı-logaritmik ve tam-logaritmik regresyondur. Çizelge 5.6'da 3 farklı model için ÇRA algoritma sonuçları gösterilmiştir. 3 model arasından konut tipi gayrimenkul toplu değerlemesinde en başarılısı yöntem doğrusal regresyon olmakla birlikte yarı-logaritmik model ile aralarında çok fazla fark bulunmamaktadır. Genel olarak ÇRA ile tahmin edilen değerler RO, YSA ve k-nn algoritmalarından sonra en başarılı modeldir. R^2 doğrusal ve yarı-logaritmik regresyon için yaklaşık 0.5 olarak hesaplanmış bu durum ÇRA algoritmalarının konut türünde çok da başarılı olmadığını göstermiştir. Ayrıca TBA sonucu değişken sayısını azaltılması ÇRA algoritmaları için daha başarısız sonuçlar elde edilmesine sebep olmuştur. Son olarak tam-logaritmik model için TBA sonucu elde edilen değişkenlerde negatif değerler olduğu için ve negatif değerlerin logaritması alınmadığı için sonuç elde edilmemiştir.

Çizelge 5.6 Konut türü gayrimenkul için doğrusal regresyon algoritması sonuçları

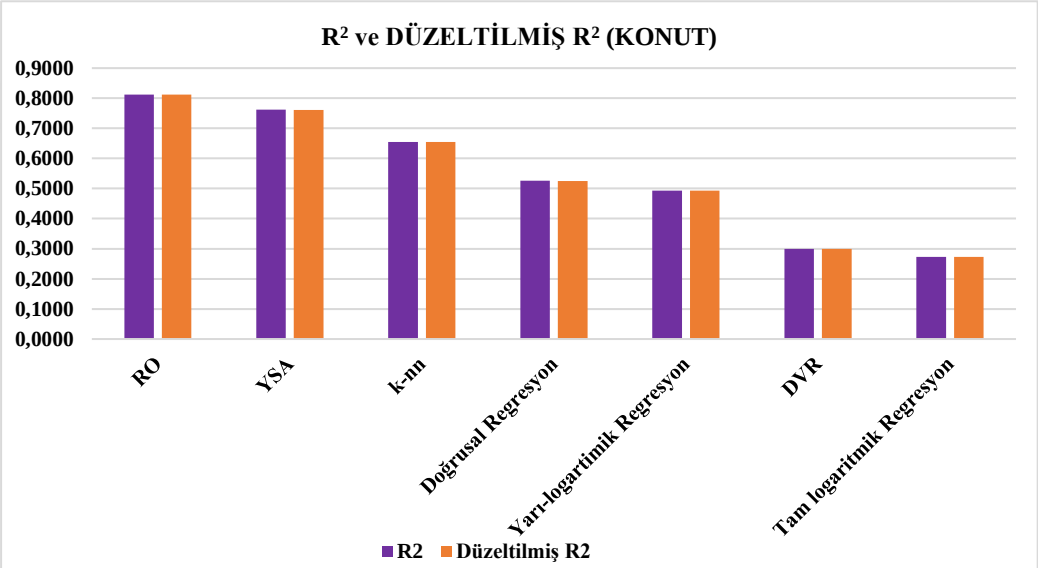
	Yöntem	22 adet değişken ile elde edilen sonuçlar	TBA sonucu 7 değişken ile elde edilen sonuçlar
R^2	Doğrusal	0.5256	0.2739
	Yarı-logaritmik	0.4932	0.2746
	Tam-logaritmik	0.3890	-
Düzeltilmiş R^2	Doğrusal	0.5255	0.2738
	Yarı-logaritmik	0.4932	0.2745
	Tam-logaritmik	0.3890	-
HKO	Doğrusal	3,544,006	5,424,474
	Yarı-logaritmik	3,785,642	5,419,095
	Tam-logaritmik	5,863,194	-
KOKH	Doğrusal	1,882	2,329.05
	Yarı-logaritmik	1,945	2,327
	Tam-logaritmik	2,421	-
OMH	Doğrusal	1,297	1,655
	Yarı-logaritmik	1,321	1,651
	Tam-logaritmik	1,620	-

- *Konut Türü Gayrimenkul İçin MÖ Algoritmalarının Karşılaştırılması*

Konut türü gayrimenkul için 1,315,675 adet konut ve bu gayrimenkullere ait 22 adet değişken ile toplu değerleme uygulaması yapılmıştır. Bu amaçla 5 farklı MÖ algoritması test edilmiştir. Bu algoritmalar için ayrı ayrı kalite analizleri yapılmış ve kalite metriklerinin kıyaslanması Çizelge 5.7'de gösterilmiştir. Bu tablo ve devamına görsel olarak gösterilmiş R^2 metriği incelendiğinde (Şekil 5.7) konut türü gayrimenkul için en başarılı MÖ algoritmasının RO algoritması olduğu ve YSA algoritmasının ise en başarılı 2. algoritma olduğu ve elde edilen modelin güvenilir olduğu görülmüştür. Öte yandan ÇRA ve DVR algoritmalarının bu veri seti ve değişkenler için güvenilir sonuç vermediği tespit edilmiştir. DVR algoritması ise süre olarak model tahmininde en uzun hesaplama süresi olan algoritma olmuştur.

Çizelge 5.7 Konut türü gayrimenkul için algoritmaların karşılaştırılması

	R^2	Düzeltilmiş R^2	HKO	KOKH	OMH
RO	0.8126	0.8125	1,409,005	1187	763
YSA	0.7617	0.7615	1,791,277	1338	922
k-nn	0.6544	0.6543	2,598,203	1611	1122
Doğrusal Regresyon	0.5256	0.5252	3,566,780	1888	1326
Yarı-logaritmik Regresyon	0.4933	0.4932	3,785,642	1945	1321
DVR	0.2997	0.2997	5,231,685	2287	1532
Tam logaritmik Regresyon	0.2734	0.2733	5,863,194	2421	1620



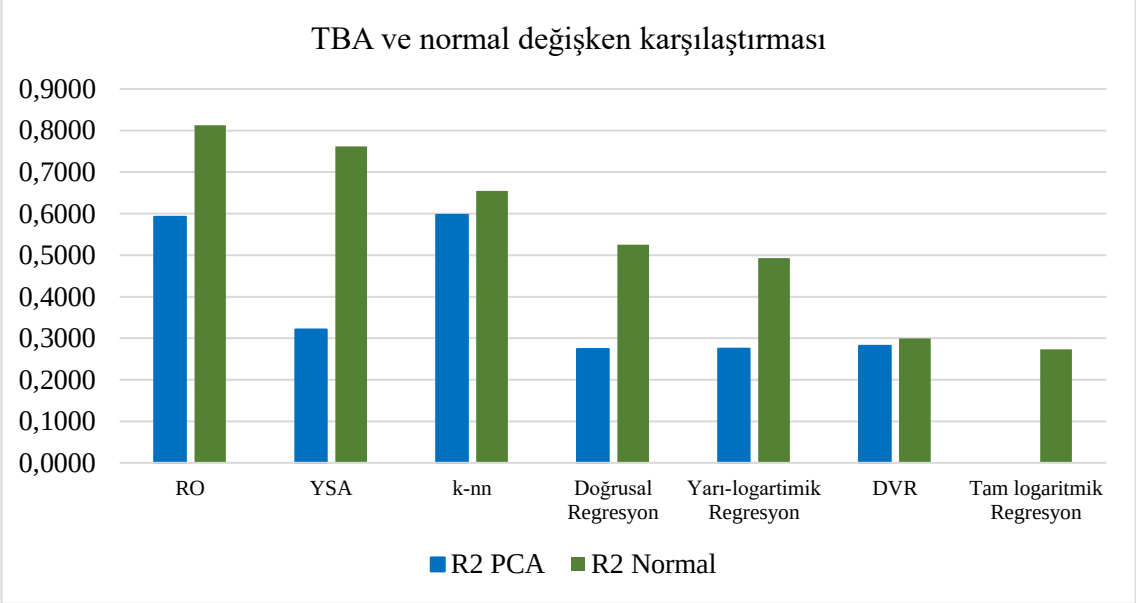
Şekil 5.7 Konut türü gayrimenkul için MÖ algoritmalarında hata oranı

Son olarak Ankara ili konut türü gayrimenkul verileri kullanılarak yapılan toplu değerlemede 22 değişken sayısı TBA algoritması ile azaltılarak tekrar MÖ algoritmalarıyla toplu değerleme yapılmış ve TBA algoritmasının sonuca etkisi incelenmiştir. Çizelge 5.8'de TBA algoritması ile azaltılmış değişken sayısı ve normal değişken sayısı ile MÖ algoritmalarını kullanarak yapılmış toplu değerleme analizi karşılaştırılmalı olarak verilmiştir. Literatürde bazı çalışmalarda TBA algoritması ile değişken sayısını azaltmanın model güvenilirliğini artırdığı çalışmalar bulunsa da bu tez çalışmasında TBA algoritması ile değişken sayısının azaltılmasının modelin başarısını artırmadığı gözlemlenmiştir.

Çizelge 5.8 Konut türü gayrimenkul için TBA ve normal değişken karşılaştırması

		RO	YSA	k-nn	Doğrusal Regresyon	Yarı-logaritmik Regresyon	DVR	Tam logaritmik Regresyon
R2	***TBA***	0.592	0.32	0.596	0.273	0.274	0.281	-
	Normal	0.8126	0.7617	0.6544	0.5256	0.4933	0.2997	0.2734
Düzeltilmiş R2	***TBA***	0.5921	0.3207	0.5966	0.2739	0.2746	0.2815	-
	Normal	0.8125	0.7615	0.6543	0.5252	0.4932	0.2997	0.2733
HKO	***TBA***	3046855	5074719	3013387	5424474	5419095	6332586	-
	Normal	1409005	1791277	2598203	3566780	3785642	5231686	5863194
KOKH	***TBA***	1745	2252	1735	2329	2327	3158	-
	Normal	1187	1338	1611	1888	1945	2287	2421
OMH	***TBA***	1233	1600	1222	1655	1651	1958	-
	Normal	763	922	1122	1326	1321	1532	1122

Son olarak TBA algoritması ile azaltılmış değişken kullanılarak yapılan toplu değerleme model başarısı ile normal değişken ile yapılan toplu değerleme başarısı (R^2) görsel olarak karşılaştırılmıştır (Şekil 5.8). Şekilden de görüleceği üzere TBA genel olarak model başarısını olumsuz etkilemiştir. Tam-logaritmik regresyon da ise TBA algoritmasında negatif değerler olduğu için ve negatif değerlerin logaritması alınamadığı için sonuç elde edilememiştir.



Şekil 5.8 Konut türü gayrimenkul için TBA ve normal değişken hata karşılaştırması

5.3 Arsa Türü Gayrimenkul İçin Toplu Değerleme Sonuçları

Arsa türü gayrimenkullerde toplamda 11 adet değişken ve 209 adet gayrimenkul bulunmaktadır. Arsa türü gayrimenkul için veriler değerleme firmaların tarafından hazırlanan raporlardan temin edilmiştir. Değişken sayısı az olduğu için bu veride TBA analizi kullanılmamıştır. 5 adet MÖ algoritmaları kullanılarak arsa türü gayrimenkullerin değeri toplu olarak test edilmiştir. Konut türünde olduğu gibi arsa türü gayrimenkul için de %80 eğitim ve %20 test verisi seçilmiştir. Yani 209 adet veriden 167'i eğitim ve 42 adeti test için kullanılmıştır. Her defasında test ve eğitim verisi 5 katmanlı çapraz doğrulama ile değiştirilmiş ve uyum iyiliği ölçütleri için sonuçlar 5 test verisinin ortalaması alınarak hesaplanmıştır. 5 katman için tüm sonuçlar ayrıntılı olarak Ek-1'de verilmiştir.

Konut türünde olduğu gibi arsa türü içinde MÖ algoritmalarında kullanılacak parametreler ızgara arama algoritması ile tespit edilmiş ve ön parametreler yani denenen ve ızgara arama algoritması sonucunda elde edilen optimum parametreler her bir MÖ algoritması için verilmiştir (Çizelge 5.9).

Çizelge 5.9 Arsa türü gayrimenkul için MÖ algoritmaları hiper parametre seçimi

Algoritma	***Izgara Arama için Parametreler***	***Kullanılan Parametreler***
YSA	Gizli katman sayısı: 5,11,22	Gizli katman sayısı:22
	Aktivasyon fonksiyonu: Identity, logistic, tanh, relu	Aktivasyon fonksiyonu: relu
	Alpha: 0.01,0.001,0.0001	Alpha: 0.01
	Öğrenme hızı: 0.01,0.001,0.0001	Öğrenme hızı 0.01
K-nn	N- komşuluk: 4,8,12,16,20,24,28,32 Uzaklık: euclidean, manhattan	N-komşuluk:8 Uzaklık: euclidean
DVR	Kernel: linear, poly, rbf, sigmoid Gamma: scale, auto	Kernel: linear Gamma: scale
RO	N_estimators (Ağaç sayısı): 50,100,150,200,250	N_estimators: 200

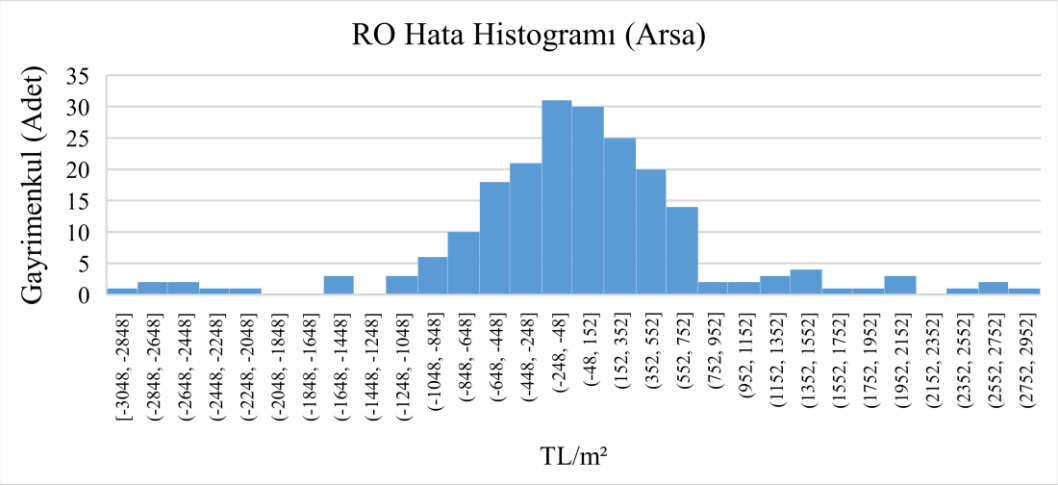
- *Arsa Türü Gayrimenkul İçin Rasgele Orman Algoritması Sonuçları*

RO algoritması ile toplu değerleme sonucu elde edilen değerlerin analizi 5 farklı metrik kullanılarak hesaplanmıştır (Çizelge 5.10). Bu sonuçlara göre RO algoritmasının konut türü gayrimenkulde olduğu gibi arsa türü gayrimenkulde de iyi sonuçlar verdiği sonucu elde edilmiştir. 5 katmanlı çapraz doğrulamasının ortalaması alınarak R^2 değeri 0.8330 bulunmuş olup, bu değer toplu değerleme çalışmaları için modelin başarılı olduğunu göstermektedir.

Çizelge 5.10 Arsa türü gayrimenkul için RO algoritması sonuçları

R^2	0.8330
Düzeltilmiş R^2	0.7708
HKO	724,743
KOKH	840
OMH	571

Test verisi her defasında değişti için tüm veri seti için test sonuçları için bilinen ve tahmin edilen değer arasındaki fark alınarak aşağıdaki histogram oluşturulmuştur. Bu histogram incelendiğinde, çoğunlukla farkın ortalama 0-500 TL arasındaki olduğu görülmektedir. Arsa türü gayrimenkullerin birim fiyatlarının yaklaşık 700 TL ile 9200 TL arasındadır ve yüksek farkların genellikle yüksek değerli gayrimenkullerde gerçekleştiği tespit edilmiştir.



Şekil 5.9 Arsa türü gayrimenkul için RO algoritması hata histogramı

Arsa türü gayrimenkul değerini en çok etkileyen değişkenler incelendiğinde RO algoritmasının bu değişkeni yüzölçümü olarak verdiği sonucuna varılmıştır. Yüzölçümünden sonra arsa değerini en çok etkileyen değişken ise üniversitelere olan mesafe olduğu görülmüştür.

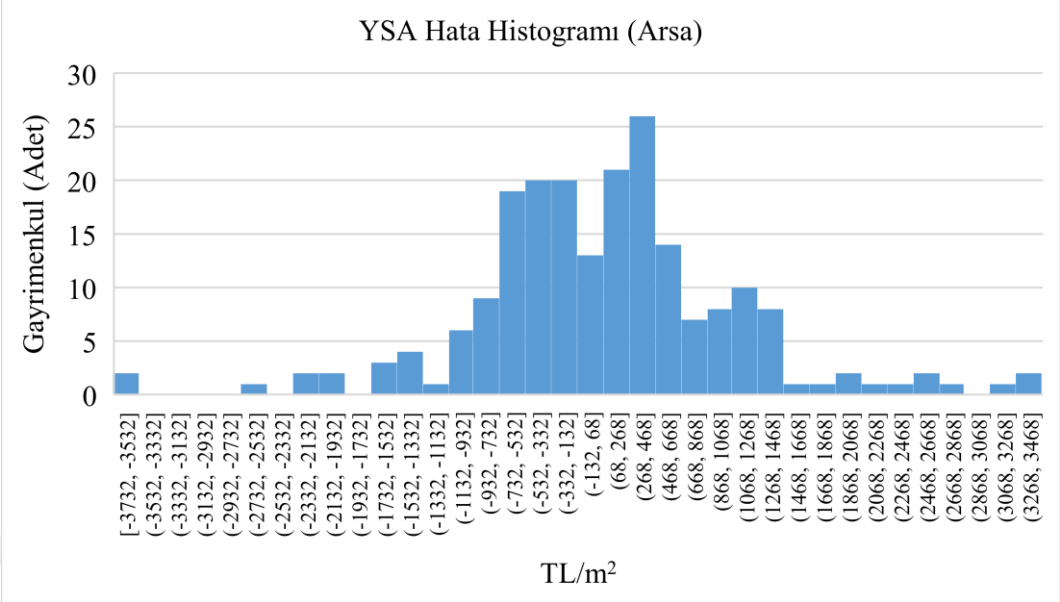
- *Arsa Türü Gayrimenkul İçin Yapay Sinir Ağları Sonuçları*

YSA algoritması kullanılarak tahmin edilmeye çalışılan arsa değerinde YSA algoritmasının da RO algoritması kadar olmasa da başarılı sonuç verdiği görülmüştür (R^2 0.76). YSA algoritması ile elde edilen toplu değerleme sonuç analizinde OMH değerinin metrekare başına 743 TL olduğu ve bu değerin yine arsa türü için kabul edilebilir bir fark olduğu bilinmektedir.

Çizelge 5.11 Arsa türü gayrimenkul için YSA algoritması sonuçları

R^2	0.7612
Düzeltilmiş R^2	0.6725
HKO	1,053,163
KOKH	1,013
OMH	743

YSA algoritması için RO algoritmasında olduğu gibi 5 katlı çapraz sorgulama yapılmış ve bu durum tüm verilerin test verisi olarak kullanılmasına imkân sağlandığı daha önce belirtilmiştir. Bu sebeple arsa türü gayrimenkulde elde olan tüm veriler için bilinen ve tahmin edilen değerler arasındaki fark alınarak aşağıdaki hata histogramı oluşturulmuştur. Bu histograma göre hata aralığının -1000 ile 1000 TL arasında yoğunlaştığı görülmekle birlikte bazı yüksek değerler da göze çarpmaktadır. Bu durumun sebebi yüksek değere sahip arsaların daha çok hata oranı verdiği olarak gözlemlenmiştir.



Şekil 5.10 Arsa türü gayrimenkul için YSA algoritması hata histogramı

- *Arsa Türü Gayrimenkul İçin Destek Vektör Regresyonu Sonuçları*

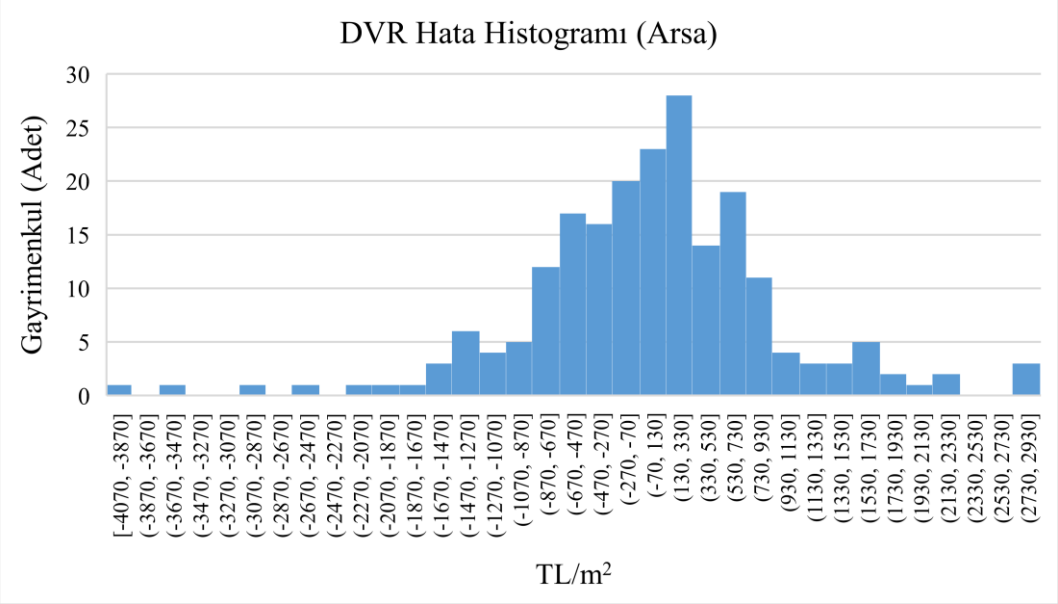
Arsa türü gayrimenkul için DVR algoritması çalıştırılmış ve 5 katlı çapraz sorgulama kalite metriklerinin ortalaması alınarak aşağıdaki çizelgede verilmiştir. Doğrusal regresyon ve RO algoritmalarından sonra arsa türü gayrimenkulde en iyi sonuçları veren MÖ algoritmasının DVR olduğu tespit edilmiştir. R^2 yaklaşık 0.8 çıkmış ve toplu değerleme çalışmaları için bu değer modelin başarılı olduğunu göstermektedir.

Çizelge 5.12 Arsa türü gayrimenkul için DVR algoritması sonuçları

R^2	0.7834
Düzeltilmiş R^2	0.7030
HKO	922,702
KOKH	955
OMH	688

Arsa türü gayrimenkul için toplu değerleme sonuçları için bilinen değer ile DVR ile elde edilen tahmini değerler arasındaki fark alınarak hata histogramı elde edilmiştir. Hata histogramı incelendiği zaman hata yoğunluğunun -900 TL ile 900 TL arasında olduğu

görülmekte ve bu arsa türü gayrimenkul değerinde kabul edilebilir bir sınır olarak varsayılmaktadır.



Şekil 5.11 Arsa türü gayrimenkul için DVR algoritması hata histogramı

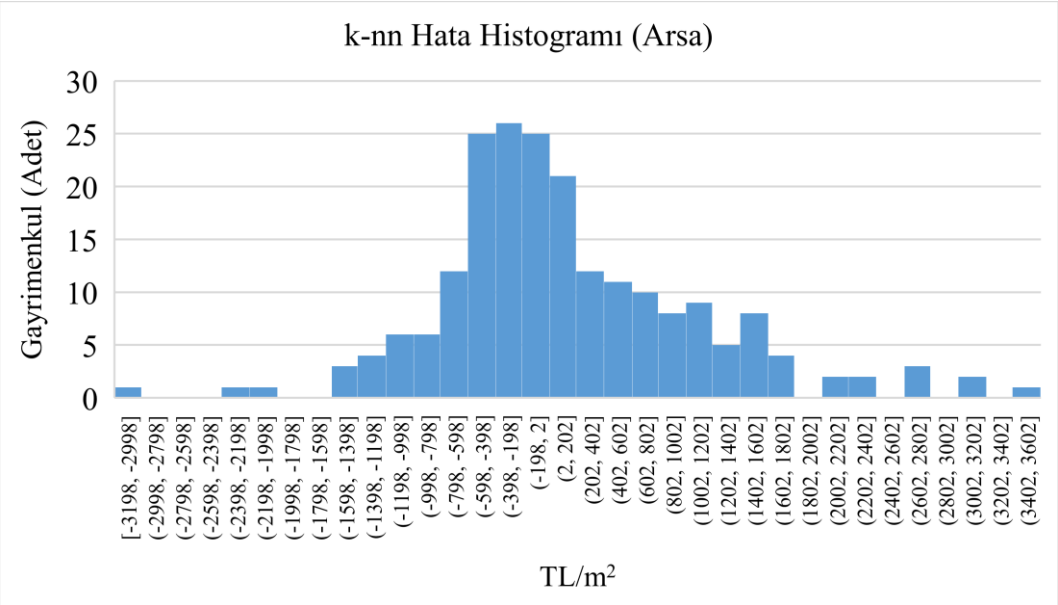
- *Arsa Türü Gayrimenkul İçin K En Yakın Komşuluk Algoritması Sonuçları*

Arsa türü gayrimenkul için k en yakın komşuluk algoritması toplu değerleme sonuçları elde edilmiş ve kalite metrikleri hesaplanmıştır (Çizelge 5.13). Arsa türü gayrimenkul için test verilerinin uyum iyiliği ölçüt sonuçlarına göre özellikle R^2 sonucunun iyi çıktığı görülmektedir. Bu sonuç arsa türü gayrimenkul değerlemesinde k-nn algoritmasının iyi sonuçlar verdiğini göstermektedir.

Çizelge 5.13 Arsa türü gayrimenkul için k-nn algoritması sonuçları

R^2	0.7835
Düzeltilmiş R^2	0.7031
HKO	935,148
KOKH	965
OMH	701

Şekil 5.12'de k-nn algoritması için arsa türü toplu gayrimenkul değerlemede bilinen değer ile model ile elde ede edilen tahmini değer arasındaki farkın histogramı gösterilmektedir. Bazı değerlerde fark çok yüksek çıkmış olarak gözükmekte ve yapılan analize göre bu değerlerin önceki 3 algoritmadaki aynı değerler olduğu tespit edilmiş ve bu durumun bu noktalardaki verilerin güvenilirliği ile ilgili olduğu sonucuna varılabilmektedir.



Şekil 5.12 Arsa türü gayrimenkul için k-nn algoritması hata histogramı

- *Arsa Türü Gayrimenkul İçin Çoklu Regresyon Analizi Sonuçları*

Arsa türü gayrimenkul toplu değerlemesinde konut türünde olduğu gibi doğrusal, yarı-logaritmik ve tam-logaritmik regresyon analizi yapılmış ve 5 katlı çapraz doğrulama ile elde edilen sonuçlar kullanılarak ortalama hata metrikleri hesaplanmıştır (Çizelge 5.14). Hem regresyon analizinde hem de tüm MÖ algoritmalarında arsa türü gayrimenkul için en iyi model RO algoritması ve doğrusal regresyon ile elde edilmiştir. Doğrusal regresyon için R^2 yaklaşık 0.83 OMH ise 588 olarak hesaplanmış bu analizler doğrusal regresyonun arsa türü gayrimenkul için oldukça başarılı olduğunu göstermektedir. Bunun yanı sıra, doğrusal regresyon analizinden sonra tam-logaritmik ve yarı-logaritmik algoritmalar da arsa türü gayrimenkul için başarılı sonuçlar elde etmiştir. Tam-logaritmik model ve yarı-logaritmik model için R^2 değerleri sırasıyla yaklaşık olarak 0.78 ve 0.71 olarak

bulunmuştur. Arsa türü gayrimenkul de genel olarak regresyon analizi başarılı sonuç vermiştir.

Çizelge 5.14 Arsa türü gayrimenkul için doğrusal regresyon algoritması sonuçları

	Yöntem	
R^2	Doğrusal	0.8268
	Yarı-logaritmik	0.7193
	Tam-logaritmik	0.8181
Düzeltilmiş R^2	Doğrusal	0.7623
	Yarı-logaritmik	0.6149
	Tam-logaritmik	0.7504
HKO	Doğrusal	743,593
	Yarı-logaritmik	1,230,613
	Tam-logaritmik	804,008
KOKH	Doğrusal	847
	Yarı-logaritmik	1,106
	Tam-logaritmik	887
OMH	Doğrusal	588
	Yarı-logaritmik	833
	Tam-logaritmik	614

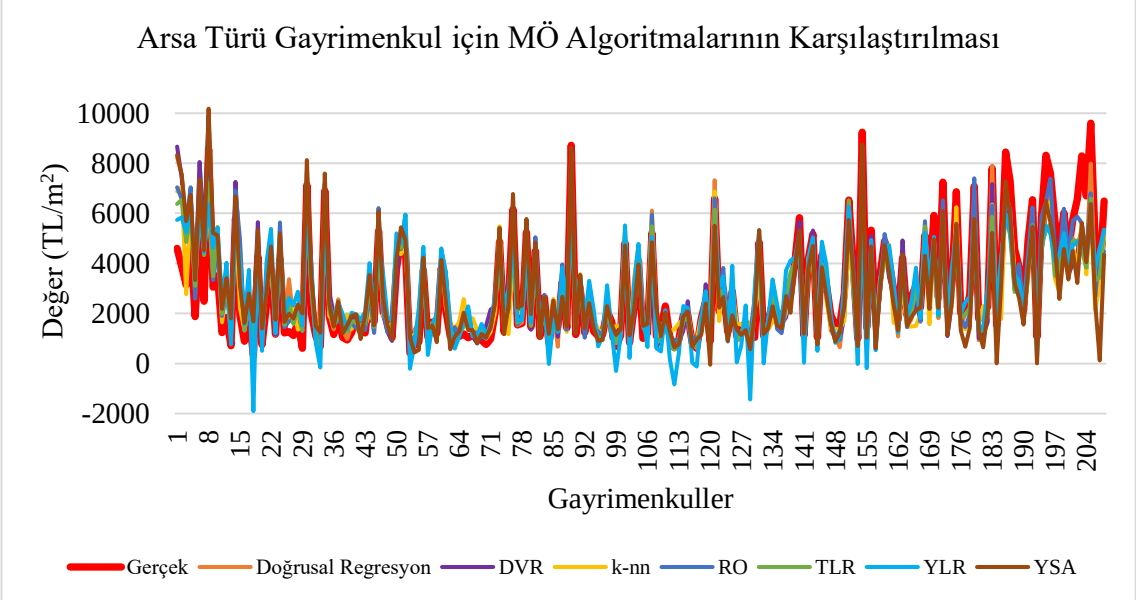
- *Arsa Türü Gayrimenkul İçin Algoritmaların Karşılaştırılması*

Arsa türü gayrimenkul için bu tez çalışmasında kullanılan tüm MÖ algoritmalarının karşılaştırılması 5 farklı uyum iyiliği ölçütü ile verilmiştir (Çizelge 5.15). Çizelge 5.15'de görüleceği üzere RO algoritması ile DR analizi bu veri setinden en iyi sonucu vermiştir. Arsa türü gayrimenkul için veri sayısı konut türü gayrimenkullere göre oldukça az olmakla birlikte ve buna rağmen tüm MÖ algoritmaları kabul edilebilir aralıkta başarılı sonuç vermiştir. YSA algoritması genel olarak daha çok veri ile daha iyi sonuçlar vermekle birlikte az sayıda arsa türü gayrimenkul için de başarılı sayılabilecek sonuç vermiştir.

Çizelge 5.15 Arsa türü gayrimenkul için MÖ algoritmalarının kıyaslanması

	R^2	*Düzeltilmiş* R^2	*HKO*	*KOKH*	*OMH*
RO	0.8330	0.7708	724,743	840	571
DR	0.8268	0.7623	743,593	847	588
Tam-logaritmik Regresyon	0.8181	0.7504	804,008	887	614
k-nn	0.7835	0.7031	935,148	965	701
DVR	0.7834	0.7030	922,702	955	688
YSA	0.7612	0.6725	1,053,163	1,013	743
Yarı-logaritmik Regresyon	0.7193	0.6149	1,230,613	1,106	833

Arsa türü gayrimenkul için bilinen ve MÖ algoritmaları kullanılarak elde edilen değerlerin karşılaştırılması yapılmıştır (Şekil 5.13). 5 katlı çapraz sorgulama sayesinde tüm verinin test edilme imkânı sağlanmıştır. Dikkat çeken bir detay tüm algoritmalar aynı eğriye vermeye yatkındır. Bunun sebebi bazı noktalardaki verilerin sorunlu olmasından olabileceği varsayılmaktadır. Genel olarak tüm algoritmalar bilinen değerlere yakın değerler tahmin etmiştir. Buna rağmen en çok bilinen değerden uzak sonuçlar veren algoritmanın YLR analizi olduğu göze çarpmaktadır.



Şekil 5.13 Arsa türü gayrimenkul için bilinen ve MÖ algoritmalarıyla elde edilen değerler

5.4 Arazi Türü Gayrimenkul İçin Toplu Değerleme Sonuçları

Arazi tür gayrimenkulde Ankara'da çalışma alanında çok fazla arazi bulunmadığı için değerleme rapor sayısı oldukça azdır bu nedenle bu çalışmada sadece 107 adet arazi türü gayrimenkul bulunmaktadır. Bu gayrimenkullere ait 12 adet değişken kullanılmıştır. 107 gayrimenkul verisinden 85 adeti eğitim ve 22 adet gayrimenkul ise test verisi için kullanılmıştır. Her defasında test ve eğitim verisi 5 katmanlı çapraz doğrulama ile değiştirilmiş ve sonuçlardan kalite metrik analizleri için ortalama alınarak gösterilmiştir. Arazi türü gayrimenkul için 5 katman için hesaplanan 5 metrik Ek-1'de verilmiştir.

Çizelge 5.16'da arazi türü gayrimenkul için ızgara arama için denenen ve ızgara arama sonucunda elde edilen optimum parametreler her bir MÖ algoritması için verilmiştir (Çizelge 5.16).

Çizelge 5.16 Arazi türü gayrimenkul için MÖ algoritmaları hiper parametre seçimi

Algoritma	***Izgara Arama için Parametreler***	***Kullanılan Parametreler***
YSA	Gizli katman sayısı: 15,30,45 Aktivasyon fonksiyonu: Identity, logistic, tanh, relu Alpha: 0.01,0.001,0.0001 Öğrenme hızı: 0.01,0.001,0.0001	Gizli katman sayısı:45 Aktivasyon fonksiyonu: logistic Alpha: 0.001 Öğrenme hızı 0.01
K-nn	N- komşuluk: 4,8,12,16,20,24,28,32 Uzaklık: euclidean, manhattan	N-komşuluk:8 Uzaklık: euclidean
DVR	Kernel: linear, poly, rbf, sigmoid Gamma: scale, auto	Kernel: linear Gamma: scale
RO	N_estimators (Ağaç sayısı): 50,100,150,200,250	N_estimators: 100

- *Arazi Türü Gayrimenkul İçin Rasgele Orman Algoritması Sonuçları*

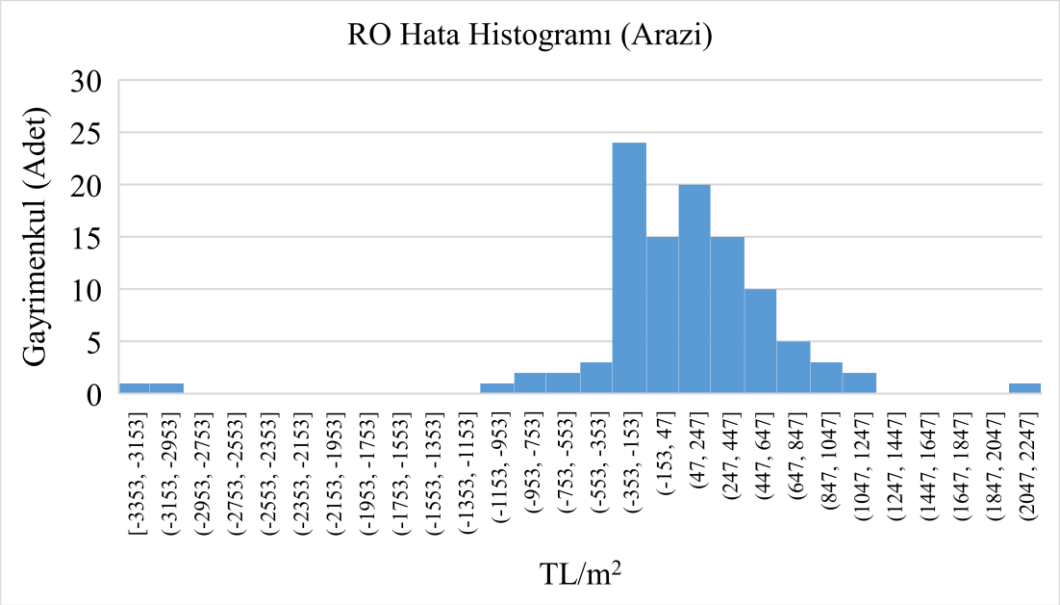
Arazi tür gayrimenkul için de diğer türlerde olduğu gibi 5 MÖ algoritması test edilmiş ve en iyi sonucu RO algoritmasının verdiği görülmüştür. R^2 metriği yaklaşık olarak 0.92 olarak bulunmuştur (Çizelge 5.17). Bu doğruluk bu tez çalışmasında ulaşılan en yüksek ve en başarılı model olmuştur. Veri sayısı göreceli az olmasına rağmen RO algoritması

arazi türü gayrimenkul için başarılı sonuçlar vermiştir. OMH değeri ise 361 TL olarak bulunmuştur. RO algoritması sonucu elde edilen verilere göre arazi türü gayrimenkulde değeri en çok yüzölçümü ve önemli yerlere olan mesafenin etkilediği tespit edilmiştir.

Çizelge 5.17 Arazi türü gayrimenkul için RO algoritması sonuçları

R^2	0.9181
Düzeltilmiş R^2	0.7963
HKO	457,497
KOKH	647
OMH	394

RO algoritmasında 5 katlı çapraz sorgulama sayesinde tüm veri test edilebilmiş ve bilinen ile tahmin edilen değerl arasındaki farkın histogramı çizdirilmiştir (Şekil 5.14). 3 gayrimenkul değeri dışında tüm değerler için farkın -1000 ile 1000 TL arasında olduğu ayrıca çoğunluk farkın -350 ila -150 TL arasında oldgu görülmüştür. Bu değerler bu tez çalışmasında elde edilen en iyi sonuçlar olmuştur. Arazi türü gayrimenkul toplu değerlemede RO algoritması ile %91 doğruluk ile oldukça başarılı sonuç elde edilmiştir.



Şekil 5.14 Arazi türü gayrimenkul için RO algoritması hata histogramı

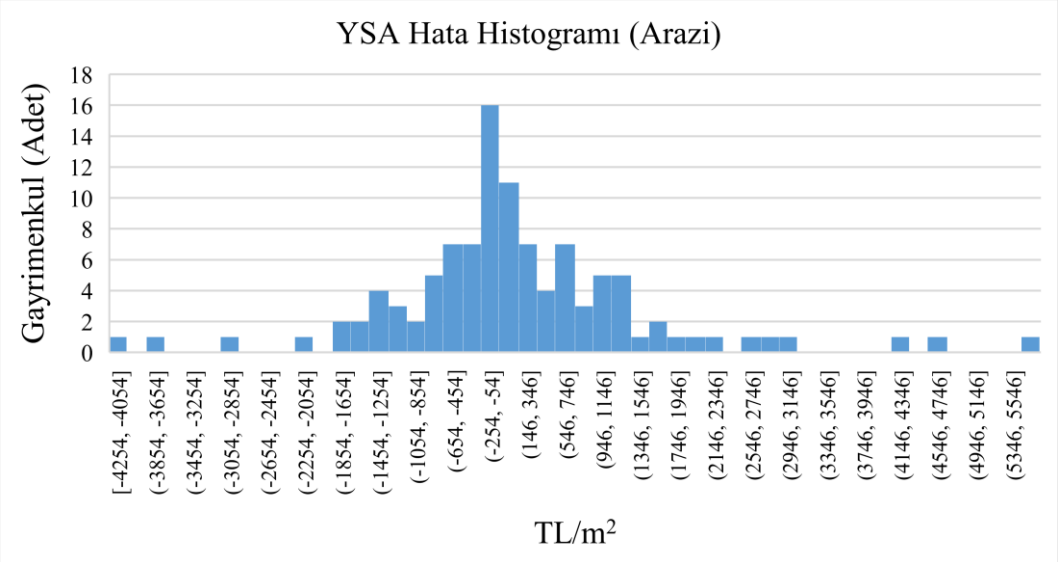
- *Arazi Türü Gayrimenkul İçin Yapay Sinir Ağları Sonuçları*

Arazi türü gayrimenkul için YSA algoritması sonuçları incelendiği zaman sonuçların RO algoritması kadar olmasa da oldukça başarılı sonuç verdiğini, OMH değerinin 612 olduğu görülmektedir (Çizelge 5.18). YSA sonuçları ayrıntılı incelendiğinde birim m^2 değerlerinde tahmin edilen ile bilinen değer arasındaki en küçük farkın yaklaşık 8 TL ve en yüksek farkın ise 1,476 TL olduğu görülmüştür. OMH ise 612 TL olduğu ve sonuçların kabul edilebilir aralıkta olduğu söylenebilecektir.

Çizelge 5.18 Arazi türü gayrimenkul için YSA algoritması sonuçları

R^2	0.8375
Düzeltilmiş R^2	0.5955
HKO	960,131
KOKH	959
OMH	706

Arazi türü gayrimenkul için bilinen değerler ile YSA algoritması ile tahmin edilen değerler arasındaki farkın histogramı çizilerek daha ayrıntılı gösterim sağlanmıştır (Şekil 5.15). YSA algoritmasında histogramın büyük aralıklara yayıldığı farkın çok olduğu gayrimenkul sayısının ise az olduğu görülmüştür.



Şekil 5.15 Arazi türü gayrimenkul için YSA algoritması hata histogramı

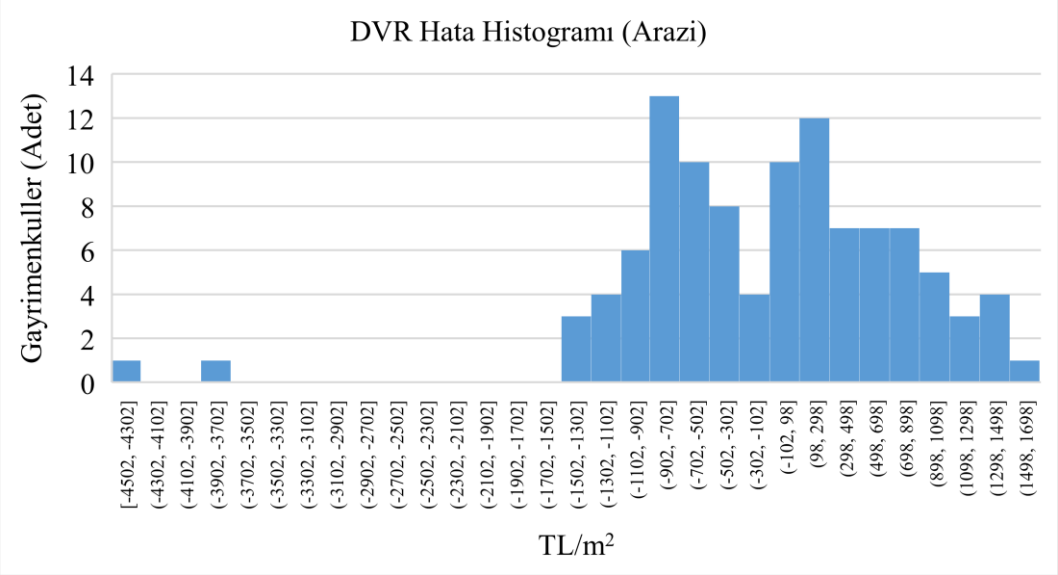
- *Arazi Türü Gayrimenkul İçin Destek Vektör Regresyonu Sonuçları*

DVR algoritması sonuçları incelendiği zaman sonuçların kabul edilebilir düzeyde olduğu ve modelin yaklaşık 0.86 R^2 skoru ile başarılı olduğu analiz edilmiştir (Çizelge 5.19). OMH değerinin 695 TL olduğu ve arazi türü için bu değerlerin değer tarihini alındığı güne göre fazla olmadığı söylenebilecektir.

Çizelge 5.19 Arazi türü gayrimenkul için DVR algoritması sonuçları

R^2	0.8539
Düzeltilmiş R^2	0.6363
HKO	871,660
KOKH	907
OMH	695

Bilinen değerler ile DVR algoritması ile tahmin edilen değerlerin farkı histogram olarak çizdirilmiştir (Şekil 5.16). Bu histogram ayrıntılı incelendiğinde 2 gayrimenkul için bilinen ve tahmin edilen farkın 4500 TL civarlarında olduğu ve bu değerin arazi türü gayrimenkul için kabul edilemez olduğu bilinmektedir. Bu 2 gayrimenkul dışında diğer gayrimenkuller için fark çok fazla gözlemlenmemiştir.



Şekil 5.16 Arazi türü gayrimenkul için DVR algoritması hata histogramı

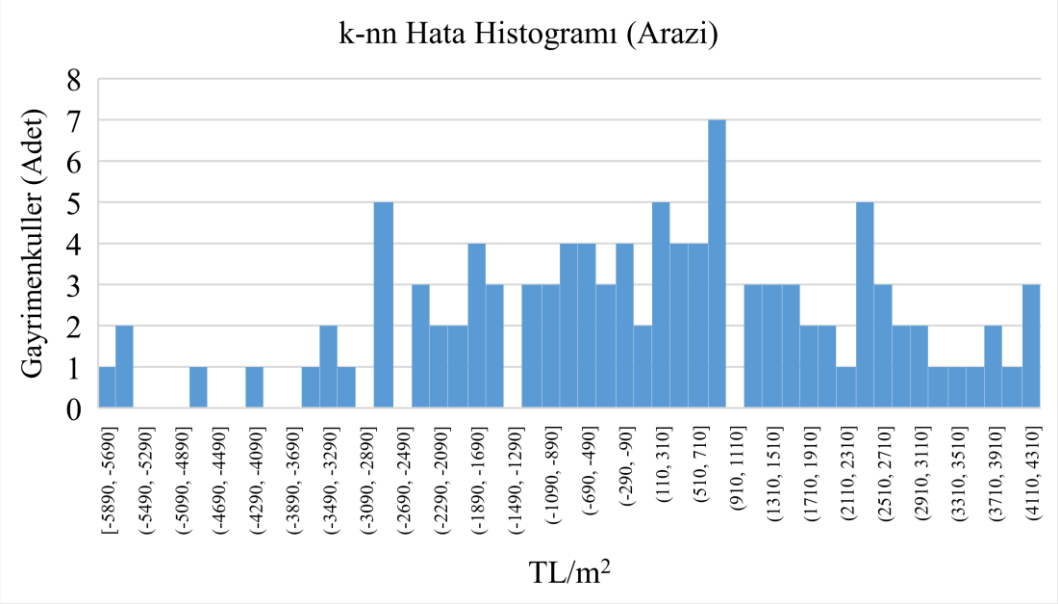
- *Arazi Türü Gayrimenkul İçin K En Yakın Komşuluk Algoritması Sonuçları*

Arazi tür gayrimenkul için k-nn algoritması denenmiş olup uyum iyiliği ölçüt sonuçları hesaplanmıştır (Çizelge 5.20). Sonuçlara bakıldığı zaman k-nn algoritmasının arazi tür gayrimenkul için tamamen başarısız olduğu bu algoritmanın arazi tür gayrimenkul verisinde çalışmadığı görülmüştür.

Çizelge 5.20 Arazi türü gayrimenkul için k-nn algoritması sonuçları

R^2	0.1894
Düzeltilmiş R^2	-1.003
HKO	5,190,478
KOKH	2,251
OMH	1,827

Diğer sonuçlarda olduğu gibi k-nn algoritması için de bilinen ve tahmin edilen değerler arasındaki fark çizdirilmiş ve sonuçta tüm gayrimenkuller için çok fazla fark çıktığı gözlemlenmiştir (Şekil 5.17).



Şekil 5.17 Arazi türü gayrimenkul için k-nn algoritması hata histogramı

- *Arazi Türü Gayrimenkul İçin Çoklu Regresyon Analizi Sonuçları*

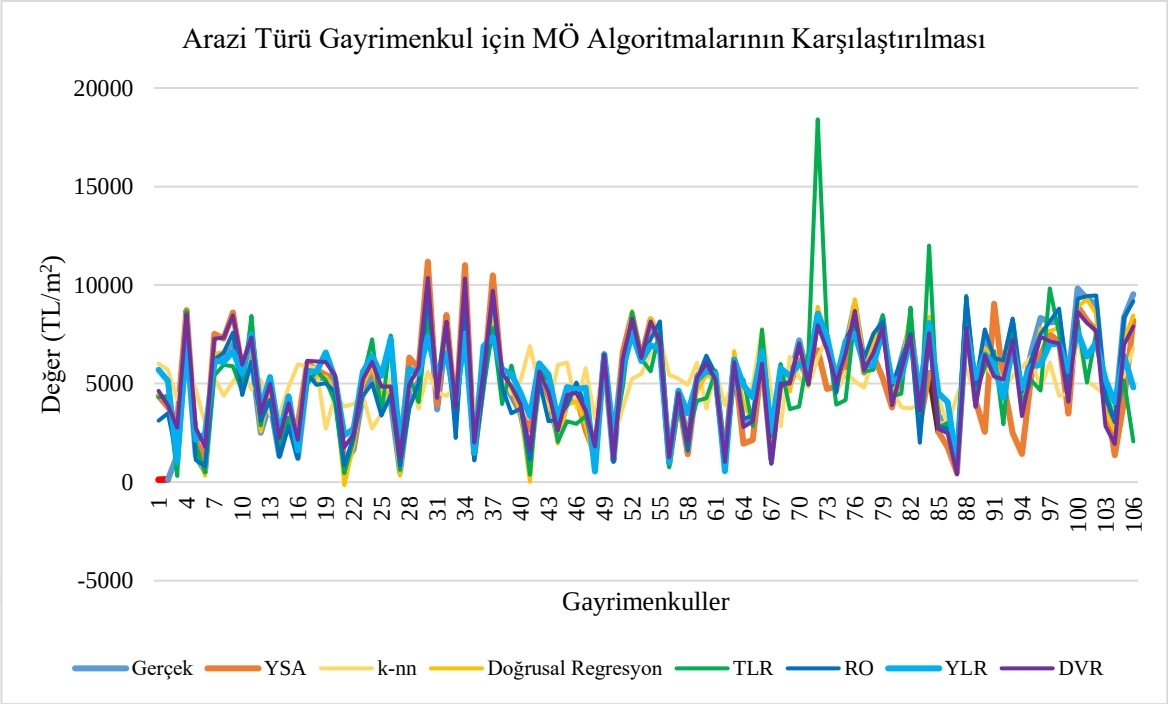
Son olarak arazi türü gayrimenkul için 3 farklı regresyon modeli çalıştırılmış olup uyum iyiliği ölçütleri hesaplanmıştır (Çizelge 5.21). Doğrusal, yarı-logaritmik ve tam-logaritmik regresyon sonuçlarının hepsi başarılı olmak ile beraber en başarı sonucu doğrusal regresyon analizinin verdiği gözlemlenmiştir. Doğrusal regresyon için R^2 değeri 0.8374 bulunurken OMH ise 475 TL olarak bulunmuştur. Yarı logaritmik regresyon analizinin de doğruluğu çok başarısız sayılmamakla birlikte tam-logaritmik regresyon analizinin ise başarısız sonuç verdiği görülmüştür.

Çizelge 5.21 Arazi türü gayrimenkul için ÇRA algoritması sonuçları

	Yöntem	
R^2	Doğrusal	0.8997
	Yarı-logaritmik	0.6384
	Tam-logaritmik	0.4069
Düzeltilmiş R^2	Doğrusal	0.7502
	Yarı-logaritmik	0.1569
	Tam-logaritmik	-0.4755
HKO	Doğrusal	580,426
	Yarı-logaritmik	2,392,909
	Tam-logaritmik	3,109,711
KOKH	Doğrusal	730
	Yarı-logaritmik	1,478
	Tam-logaritmik	1,805
OMH	Doğrusal	475
	Yarı-logaritmik	1,131
	Tam-logaritmik	1,231

5.4.1 Arazi türü gayrimenkul için algoritmaların karşılaştırılması

Son olarak arazi tür gayrimenkullerin için tüm MÖ algoritmaları bilinen ve tahmin edilen değerler kullanılarak bir arada kıyaslanmıştır. Tüm veri seti için bilinen değer ve MÖ algoritmaları kullanılarak elde edilen değerler çizdirilmiştir (Şekil 5.18). Tam-logaritmik regresyon ve k-nn algoritmaları hariç diğer algoritmaların bilinen değere yakın bir çizgisellik verdiği görülmektedir.



Şekil 5.18 Arazi türü gayrimenkul için MÖ algoritmalarının kıyaslanması

Çizelge 5.22'de ise tüm algoritmaların uyum iyiliği ölçüt sonuçları bir arada verilmiştir. K-nn algoritması ve TLR analizi hariç diğer tüm MÖ algoritmaları başarılı sonuç vermiştir. En başarılı sonuçlar sırasıyla RO, doğrusal regresyon, YSA, DVR, tam-logaritmik regresyon ve yarı logaritmik regresyon şeklindedir.

Çizelge 5.22 Arazi türü gayrimenkul için MÖ algoritmalarının kıyaslanması

	R^2	*Düzeltilmiş* R^2	*HKO*	*KOKH*	*OMH*
RO	0.9181	0.7963	457,497	647	394
Doğrusal Regresyon	0.8997	0.7502	580,426	730	475
DVR	0.8539	0.6363	871,660	907	695
YSA	0.8375	0.5955	960,131	959	706
Yarı-logaritmik Regresyon	0.6384	0.1569	2,392,909	1,478	1,131
Tam-logaritmik Regresyon	0.4069	-0.4755	3,109,711	1,805	1,231
k-nn	0.1895	-1.0038	5,190,479	2,252	1,828

5.5 Ticari Tür Gayrimenkul İçin Toplu Değerleme Sonuçları

Ticari türü gayrimenkul olarak çalışma alanında değerleme raporları hazırlanmış ve değişkenleri bilinen 109 adet veriye ulaşılmıştır. Bu verilerden 87 adeti MÖ modellerini eğitmek ve 22 adeti ise test amaçlı kullanılmakla beraber 5 katlı çapraz sorgulama sayesinde her seferinde test ve eğitim verisi değişmiştir. 5 katmanlı çapraz doğrulama ile hesaplanan hata metriklerinin ortalaması alınarak kalite analizi yapılmıştır. 5 katman için tüm sonuçlar ise ayrıntılı olarak Ek-1'de verilmiştir.

Ticari türü gayrimenkul için ızgara arama için denenen ve ızgara arama sonucunda elde edilen optimum parametreler her bir MÖ algoritması için hesaplanmıştır (Çizelge 5.23).

Çizelge 5.23 Ticari türü gayrimenkul için MÖ algoritmaları hiper parametre seçimi

Algoritma	***Izgara Arama için Parametreler***	***Kullanılan Parametreler***
YSA	Gizli katman sayısı: 5,11,22 Aktivasyon fonksiyonu: Identity, logistic, tanh, relu Alpha: 0.01,0.001,0.0001 Öğrenme hızı: 0.01,0.001,0.0001	Gizli katman sayısı:11 Aktivasyon fonksiyonu: relu Alpha: 0.001 Öğrenme hızı 0.01
K-nn	N- komşuluk: 4,8,12,16,20,24,28,32 Uzaklık: euclidean, manhattan	N-komşuluk:8 Uzaklık: euclidean
DVR	Kernel: linear, poly, rbf, sigmoid Gamma: scale, auto	Kernel: linear Gamma: scale
RO	N_estimators (Ağaç sayısı): 50,100,150,200,250	N_estimators: 200

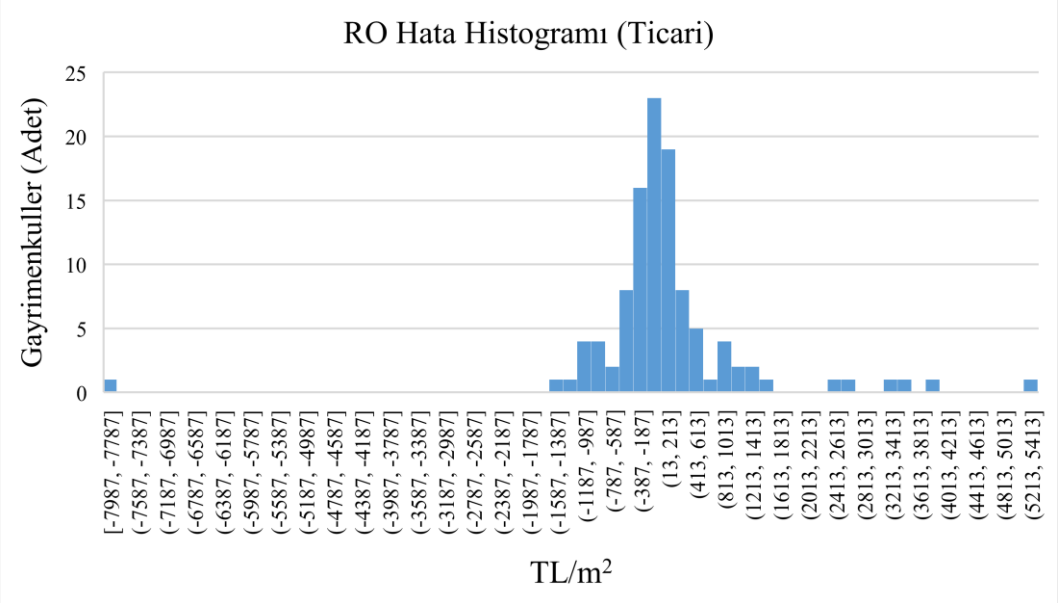
- *Ticari Tür Gayrimenkul İçin Rasgele Orman Algoritması Sonuçları*

RO algoritması ile elde edilen toplu değerleme sonuçları ile uyum iyiliği ölçütleri hesağlanmıştır (Çizelge 5.24). Bu sonuçlara göre RO algoritmasının ticari türü gayrimenkulde de iyi sonuçlar verdiği gözlemlenmiştir. Düzeltişmiş R^2 değerinin daha düşük çıkmasının nedeni ise veri sayısı ile alakalı olup, R^2 değerinin yüksek çıktığı yani sonuçların güvenilir olduğu görülmektedir.

Çizelge 5.24 Ticari türü gayrimenkul için RO algoritması sonuçları

R^2	0.8355
Düzeltilmiş R^2	0.5569
HKO	1,547,916
KOKH	1,080
OMH	628

Ticari tür gayrimenkul için tüm değerler her seferinde test edildiği için tüm veri seti ile bilinen değerler ile tahmine dilen değerler arasındaki fark alınarak hata histogramı oluşturulmuştur (Şekil 5.19). 7 adet uç değer dışında diğer değerler ticari gayrimenkul türünde toplu değerleme için kabul edilebilir seviyede olduğu gözlemlenmiştir.



Şekil 5.19 Ticari türü gayrimenkul için RO algoritması hata histogramı

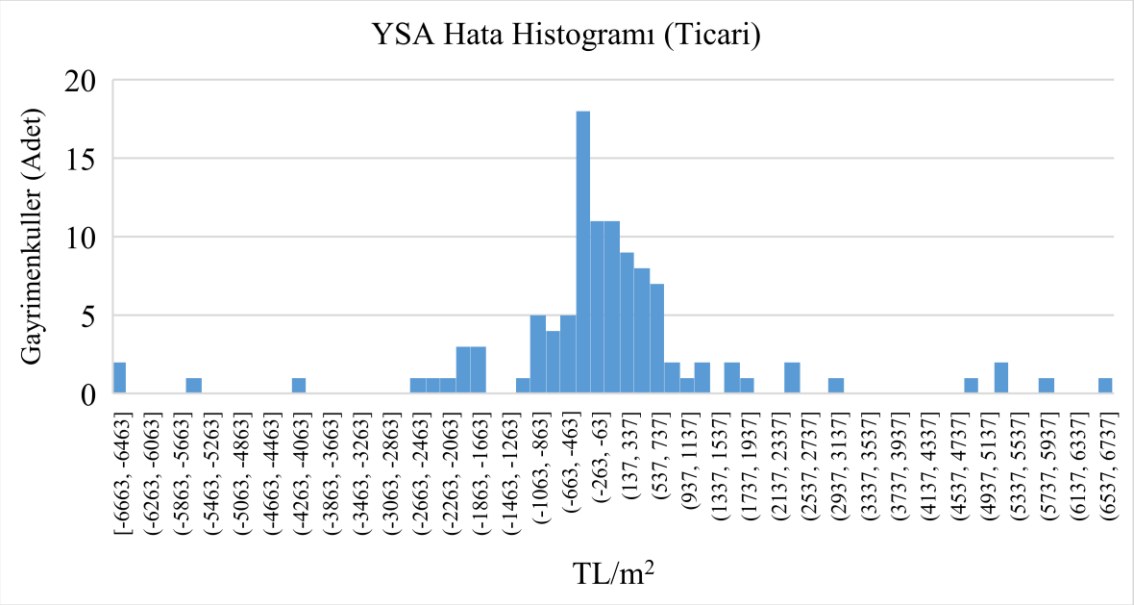
- *Ticari Tür Gayrimenkul İçin Yapay Sinir Ağları Sonuçları*

Ticari tür gayrimenkul için YSA algoritması koşturulmuş olup kalite analiz sonuçlar hesaplanmıştır (Çizelge 5.25). RO algoritması kadar başarılı bir sonuç elde edilmese de YSA algoritması da ticari tür gayrimenkul için güvenilir sonuçlar vermiştir.

Çizelge 5.25 Ticari türü gayrimenkul için YSA algoritması sonuçları

R^2	0.6566
Düzeltilmiş R^2	0.4302
HKO	3,993,499
KOKH	1,869
OMH	1,161

YSA algoritması için hata histogramı incelendiğinde bilinen değerler ile tahmin edilen değerler arasındaki farkın çok dağınık olduğu görülmüştür (Şekil 5.20). RO algoritması hata histogramına kıyasla bu farkların çok olduğu ve modelin bazı gayrimenkuller için oldukça başarısız sonuç verdiği ayrıca gözlemlenebilmiştir.



Şekil 5.20 Ticari türü gayrimenkul için YSA algoritması hata histogramı

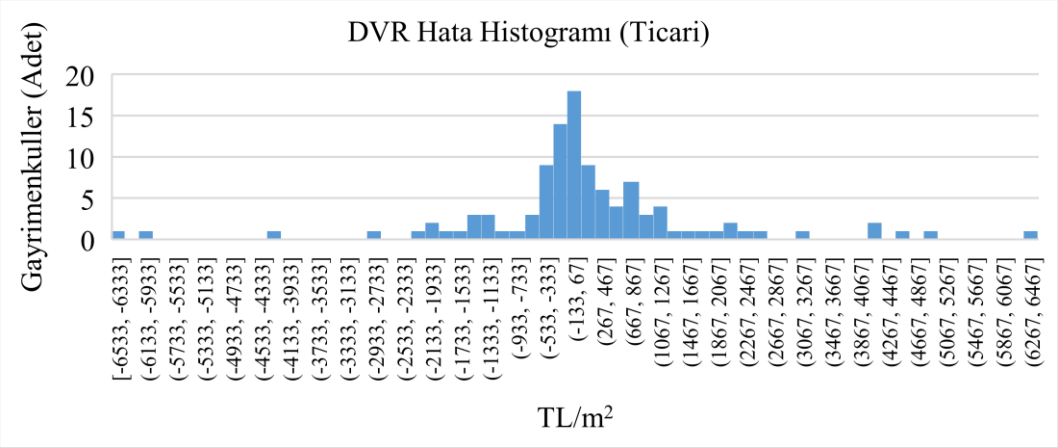
- *Ticari Tür Gayrimenkul İçin Destek Vektör Regresyonu Sonuçları*

DVR algoritması ile ticari tür toplu değerleme işlemi için denenmiş olup, algoritmanın başarısı bu tez çalışmasında hesaplanan 5 uyum iyiliği ölçütü ile analiz edilmiştir. R^2 değeri yaklaşık olarak 0.73 ve OMH ise 1,023 TL olarak hesaplanmıştır ve ticari tür gayrimenkul için bu sonuçlar çok başarılı olmasa da kabul edilebilir düzeydedir.

Çizelge 5.26 Ticari türü gayrimenkul için DVR algoritması sonuçları

R^2	0.7258
Düzeltilmiş R^2	0.3229
HKO	2,934,754
KOKH	1,616
OMH	1,023

DVR algoritması için hata histogramı ayrıntılı olarak incelendiğinde YSA algoritmasında olduğu gibi genel olarak değerlerin kabul edilebilir seviyede bazı değerlerin ise çok yüksek farka sahip olduğu gözlemlenmiştir (Şekil 5.21).



Şekil 5.21 Ticari türü gayrimenkul için DVR algoritması hata histogramı

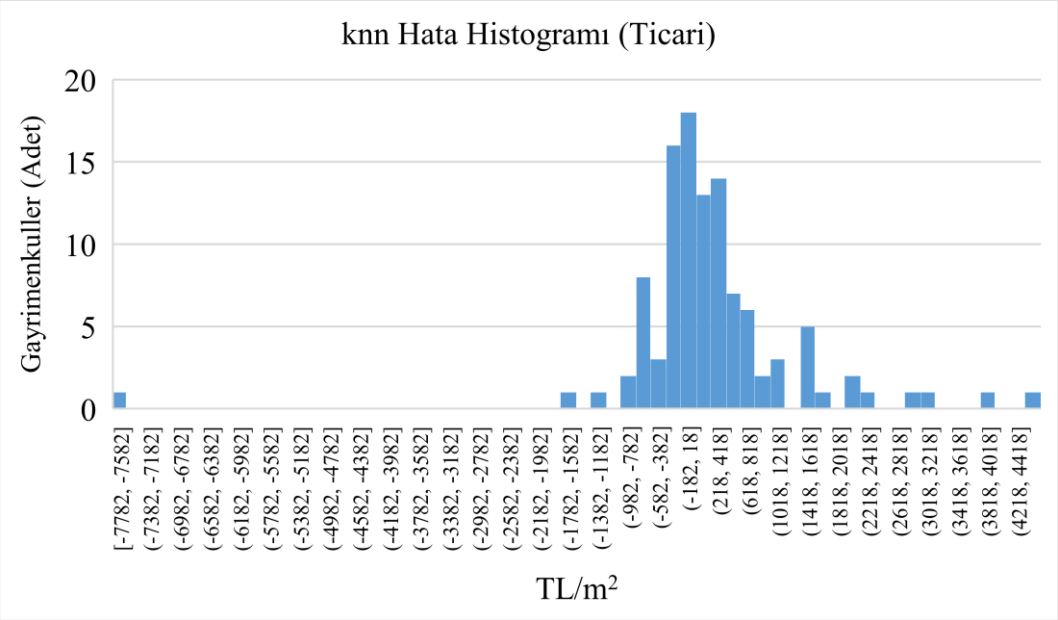
- *Ticari Tür Gayrimenkul İçin K En Yakın Komşuluk Algoritması Sonuçları*

Ticari tür gayrimenkul için k-nn algoritması DVR algoritmasına yakın bir sonuç vermiştir. DVR algoritmasına göre ise OMH değeri biraz daha iyi olarak 806 TL olarak hesaplanmıştır (Çizelge 5.27).

Çizelge 5.27 Ticari türü gayrimenkul için k-nn algoritması sonuçları

R^2	0.7206
Düzeltilmiş R^2	0.2360
HKO	1,987,105
KOKH	1,332
OMH	806

Son olarak k-nn algoritması ile tahmin edilen değerler ile bilinen değerler arasındaki farkın histogramı çizdirilerek sonuçlar daha detaylı incelenmiştir (Şekil 5.22). Bu sonuçlara göre yine ticari tür gayrimenkul için bazı gayrimenkullerde farkın çok yüksek olduğu görülmüştür. Bu gayrimenkul incelendiğinde diğer MÖ algoritmalarında da yüksek fark çıkan gayrimenkuller olduğu gözlemlenmiş ve bu gayrimenkullere ait verilerin sorunlu olduğu sonucuna varılmıştır.



Şekil 5.22 Ticari türü gayrimenkul için k-nn algoritması hata histogramı

- *Ticari Tür Gayrimenkul İçin Çoklu Regresyon Analizi Sonuçları*

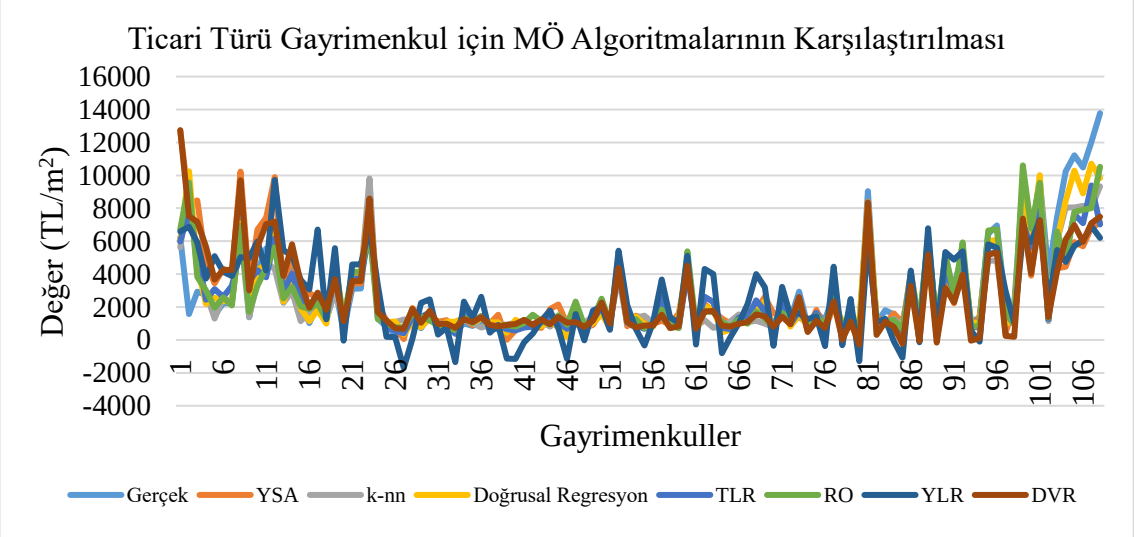
Son olarak ticari tür gayrimenkul için 3 farklı regresyon modeli ile değer sonuçlar denenmiş olup uyum iyiliği ölçütleri hesaplanmıştır (Çizelge 5.28). Bu sonuçlara göre regresyon modelleri içinde en iyi sonucu doğrusal regresyon sonrasında tam-logaritmik algoritması vermiştir. Bunun yanı sıra yarı-logaritmik regresyon analizinin ticari tür gayrimenkul için başarısız olduğu görülmüştür.

Çizelge 5.28 Ticari türü gayrimenkul için regresyon algoritmaları sonuçları

	Yöntem	
R^2	Doğrusal	0.8231
	Yarı-logaritmik	0.1801
	Tam-logaritmik	0.7196
Düzeltilmiş R^2	Doğrusal	0.5171
	Yarı-logaritmik	-1.2887
	Tam-logaritmik	0.2326
HKO	Doğrusal	1,271,166
	Yarı-logaritmik	4,129,219
	Tam-logaritmik	1,985,873
KOKH	Doğrusal	960
	Yarı-logaritmik	1,990
	Tam-logaritmik	1,298
OMH	Doğrusal	608
	Yarı-logaritmik	1,514
	Tam-logaritmik	768

5.5.1 Ticari türü gayrimenkul için algoritmaların karşılaştırılması

Ticari tür gayrimenkul için 5 katlı çapraz sorgulama sayesinde tüm ticari tür verisi kullanılarak bu tez çalışmasında kullanılan tüm MÖ algoritmaları görsel olarak kıyaslanmıştır (Şekil 5.23). Tüm MÖ algoritmaları şekilden de görüleceği üzere iyi sonuç verse de değer arasında en az farkı RO ve doğrusal regresyon algoritmaları vermiştir. En başarısız sonuç ise yarı-logaritmik regresyon analizi ile elde edilmiştir.



Şekil 5.23 Ticari türü gayrimenkul için MÖ algoritmalarının kıyaslanması

Ticari tür gayrimenkul için tüm MÖ algoritmaları için uyum iyiliği ölçütleri karşılaştırılmalı olarak incelenmiştir (Çizelge 5.29). Yarı-logaritmik regresyon dışında bu tez çalışmasında kullanılan tüm MÖ algoritmaları ticari tür gayrimenkul için kabul edilebilir seviyede başarılı olsa da en başarılı algoritma bu tür veri seti için RO ve doğrusal regresyon olarak tespit edilmiştir.

Çizelge 5.29 Ticari türü gayrimenkul için MÖ algoritmaları karşılaştırılması

	R^2	*Düzeltilmiş* R^2	*HKO*	*KOKH*	*OMH*
RO	0.8355	0.5569	1,547,916	1,080	628
Doğrusal Regresyon	0.8231	0.5171	1,271,166	960	608
DVR	0.7258	0.3229	2,934,754	1,616	1,023
k-nn	0.7206	0.2360	1,987,105	1,332	806
Tam-logaritmik Regresyon	0.7196	0.2326	1,985,873	1,298	768
YSA	0.6566	0.4302	3,993,499	1,869	1,161
Yarı-logaritmik Regresyon	0.1801	-1.2888	4,129,219	1,990	1,514

5.6 Makine Öğrenme Algoritmaları Sonuçlarının Karşılaştırılması

Bu başlık altında tüm gayrimenkul türleri için test edilen MÖ algoritmaları bir arada kıyaslanmıştır (Çizelge 5.30). Karşılaştırma analizi ayrıntılı incelendiğinde bu tez çalışmasında kullanılan 4 farklı gayrimenkul türü ve değişkenleri için en iyi toplu değerleme sonuçlarını veren algoritmaların başında tüm gayrimenkul türlerinde RO algoritmasının olduğu tespit edilmiştir. Sonrasında ise konut türü hariç diğer tüm türlerde DR analizinin RO algoritması kadar başarılı olduğu görülmüştür. Sadece çok fazla veriye sahip olan konut türünde YSA algoritmaları RO algoritmasından sonra en başarılı sonucu veren algoritma olmuştur.

Regresyon türleri içinde en başarılı olan doğrusal regresyon olsa da tam-logaritmik regresyon da oldukça başarılı sonuçlar vermiştir. Bunun yanı sıra, yarı-logaritmik regresyon analizi regresyon türleri ile tüm bu tez çalışmasında kullanılan MÖ algoritmaları içinde en başarısız sonucu veren tür olmuştur. Diğer yandan k-nn algoritması da 4 tür gayrimenkul için çok yüksek başarı elde edememiştir. DVR

algoritmasının ise çok fazla veri ve değişkene sahip konut verisi dışında %72'nin üstünde doğruluk sağladığı görülmüştür.

Ticari tür gayrimenkul veri setinde ise yarı-logaritmik regresyon analizi hariç bu tez çalışmasında kullanılan tüm algoritmaların başarılı sonuç verdiği göze çarpmaktadır. Bunun yanı sıra arsa türü gayrimenkulde ise yarı-logaritmik regresyonda dahil olmak üzere tüm MÖ algoritmalarının %70'in üstünde doğruluk verdiği görülmüştür. Arazi türü gayrimenkulde ise RO algoritması ile bu tez çalışmasında elde edilen en yüksek doğruluk sağlanmıştır.

Çizelge 5.30 Gayrimenkul türleri için MÖ algoritmalarının toplu kıyaslanması

	Gayrimenkul Türü	**R^2**	**Düzeltilmiş R^2**	**HKO**	**KOKH**	**OMH**
RO	Arazi	0.9181	0.7963	457,497	647	394
	Arsa	0.8330	0.7708	724,743	840	571
	Ticari	0.8355	0.5569	1,547,916	1080	628
	Konut	0.8126	0.8125	1,409,005	1187	763
YSA	Arazi	0.8375	0.5955	960,131	959	706
	Konut	0.7617	0.7615	1,791,277	1338	923
	Arsa	0.7214	0.6181	1,195,664	1078	781
	Ticari	0.6566	0.4302	3,993,499	1869	1161
k-nn	Arsa	0.7835	0.7031	935148	965	701
	Ticari	0.7206	0.2360	1,987,105	1,332	806
	Konut	0.6544	0.6543	2,598,204	1,612	1,123
	Arazi	0.1895	-1.0038	5,190,479	2,252	1,828
Doğrusal Regresyon	Arazi	0.8997	0.7502	580,426	730	475
	Arsa	0.8268	0.7623	743,593	847	588
	Ticari	0.8231	0.5171	1,271,166	960	608
	Konut	0.5256	0.5252	3,566,780	1,889	1,327
Yarı-logartimik Regresyon	Arsa	0.7193	0.6149	1,230,613	1,106	833
	Arazi	0.6384	0.1569	2,392,909	1,478	1,131
	Konut	0.4933	0.4932	3,785,642	1,946	1,321
	Ticari	0.1801	-1.2888	4,129,219	1,990	1,514
Tam logaritmik Regresyon	Ticari	0.8229	0.5352	1,789,917	1,338	842
	Arsa	0.8181	0.7504	804,008	887	614
	Konut	0.7196	0.2326	1,985,873	1,298	768
	Arazi	0.4069	-0.4755	3,109,711	1,805	1,231
DVR	Arazi	0.8539	0.6363	871,660	907	695
	Arsa	0.7834	0.7030	922,702	955	688
	Ticari	0.7258	0.3229	2,934,754	1,616	1,023
	Konut	0.2997	0.2997	5,231,686	2,287	1,532

5.7 Konut Türü Uygulama Verisi 2 ve Sonuçları

Çalışmada seçilen Ankara ili çalışma alanı dışında Amerika Birleşik Devletleri, Ames, Iowa şehrine ait konut türünde gayrimenkul verileri popüler veri sağlayıcı olan kaggle.com sitesinden indirilmiş ve bu tez çalışmasında kullanılan MÖ algoritmaları bu veri setiyle test edilmiştir. Bu veri setinde toplamda 2,352 adet konut verisi bulunmaktadır ve bu verilerden 1,881 adeti algoritmaları eğitmek için ve 471 adeti ise test etmek için kullanılmıştır. Yani veri setinin %80'i eğitmek ve %20'si ise MÖ algoritmalarını test etmek amacıyla kullanılmıştır.

Çizelge 5.31 Konut türü gayrimenkul için MÖ algoritmaları hiper parametre seçimi

Algoritma	***Izgara Arama için Parametreler***	***Kullanılan Parametreler***
YSA	Gizli katman sayısı: 15,30,45,60 Aktivasyon fonksiyonu: Identity, logistic, tanh, relu Alpha: 0.01,0.001,0.0001 Öğrenme hızı: 0.01,0.001,0.0001	Gizli katman sayısı:45 Aktivasyon fonksiyonu: relu Alpha: 0.001 Öğrenme hızı 0.01
K-nn	N- komşuluk: 4,8,12,16,20,24,28,32 Uzaklık: euclidean, manhattan	N-komşuluk:2 Uzaklık: euclidean
DVR	Kernel: linear, poly, rbf, sigmoid Gamma: scale, auto	Kernel: linear Gamma: scale
RO	N_estimators (Ağaç sayısı): 50,100,150,200,250	N_estimators: 100

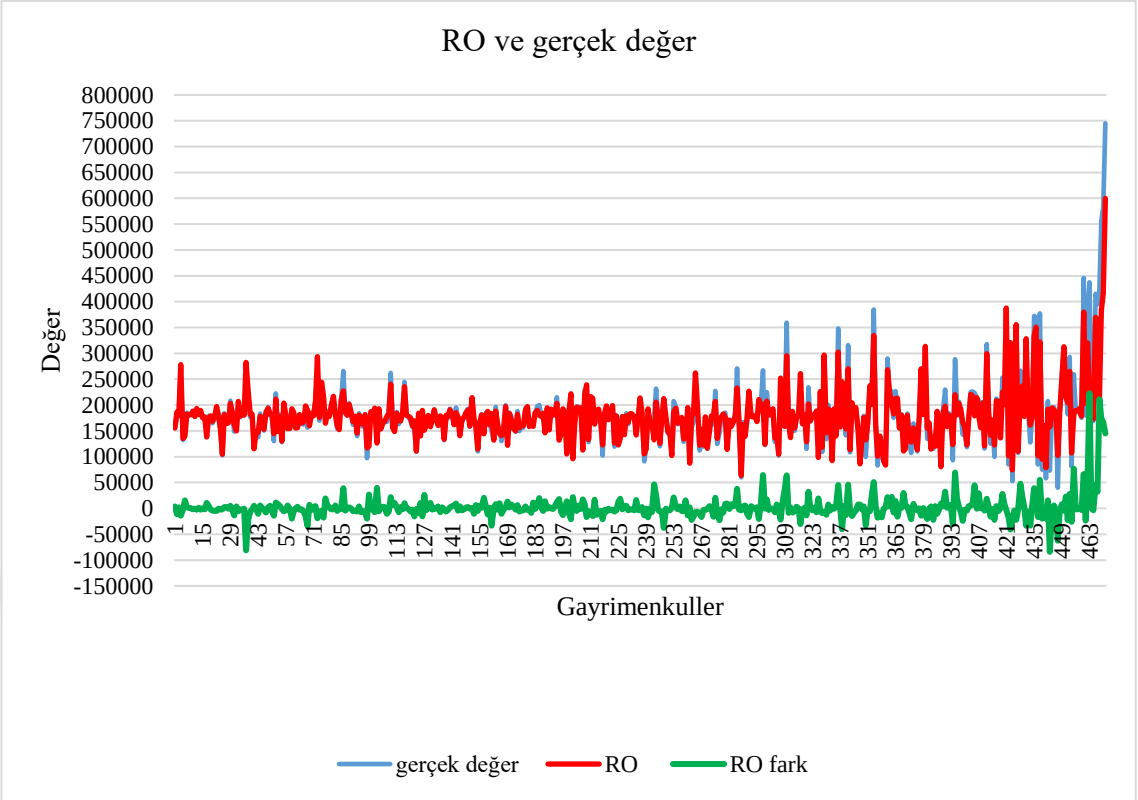
Veri setinde çok fazla değişken olması ve değişkenlerden bazılarının 0 ve negatif değer içermesi nedeniyle regresyon türlerinden sadece doğrusal regresyon bu veri setinde test edilmiş ve uyum iyiliği ölçütleri hesaplanmıştır (Çizelge 5.32). Çizelgeden de görüleceği üzere bu veri setinde RO algoritması oldukça başarılı olmuş ve YSA algoritması da RO algoritması kadar olmasa da başarılı sonuç vermiştir. Öte yandan doğrusal regresyon, k-nn ve DVR algoritmaları bu veri setinde başarısız sonuç vermiştir.

RO algoritması ile elde edilen bilgiler ile değeri en çok etkileyen ilk üç değişkenin sırasıyla konutun genel kalitesi, parselin şekli ve toplam alanı ve değeri en az etkileyen üç değişkenin ise sırasıyla, ısıtma tipi, havuz alanı ve mevsimlik sundurma alanı olduğu tespit edilmiştir.

Çizelge 5.32 Konut verisi 2 için kalite analiz sonuçları

	R^2	Düzeltilmiş R^2	HKO	KOKH	OMH
RO	0.8514	0.8349	637,902,144	25,256	12,021
YSA	0.6923	0.6581	1,320,675,000	36,341	20,963
Doğrusal Regresyon	0.4851	0.4279	2,210,011,995	47,010	29,622
DVR	0.3317	0.2575	2,868,551,276	53,558	27,960
k-nn	0.2580	0.1756	3,184,816,936	56,434	35,338

Şekil 5.24'te konut 2 verisi için bilinen değerler ile RO algoritması ile tahmin edilen değerler karşılaştırmalı olarak gösterilmiştir. Diğer MÖ algoritmaları çok başarılı olmadığı için bu bölümde sadece RO algoritması detaylı olarak incelenmiştir. Bilinen değerler ile tahmin edilen değerler arasındaki farkın konutun değeri arttıkça daha fazla arttığı şekilden de net olarak görülmektedir. Bazı konut verisi aradaki farkın 150.000 doları olduğu ve bu değerin gayrimenkul için çok başarısız bir sonuç olduğunu belirtmekle birlikte bu tarz çok fazla farkın çok olmadığı sadece 5 adet konutta bu fark görülmüştür ve bu verilerin gürültülü/hatalı olabileceği düşünülmektedir. Bu gürültülü veriler eğer test verisinden çıkarılması durumunda bütün verilerde bilinen ve hesaplanan farkın 10,000 dolar altına düştüğü tespit edilmekte ve bu değer kabul edilebilir bir aralıkta olmaktadır. Ortalama farkın ise 12,000 dolar olduğu test verisinde yarısından fazla değerin ise 5,000 dolar gibi bir gayrimenkul için kabul edilebilir bir farka sahip olduğu sonuçların analiz edilmesi ile anlaşılmıştır.

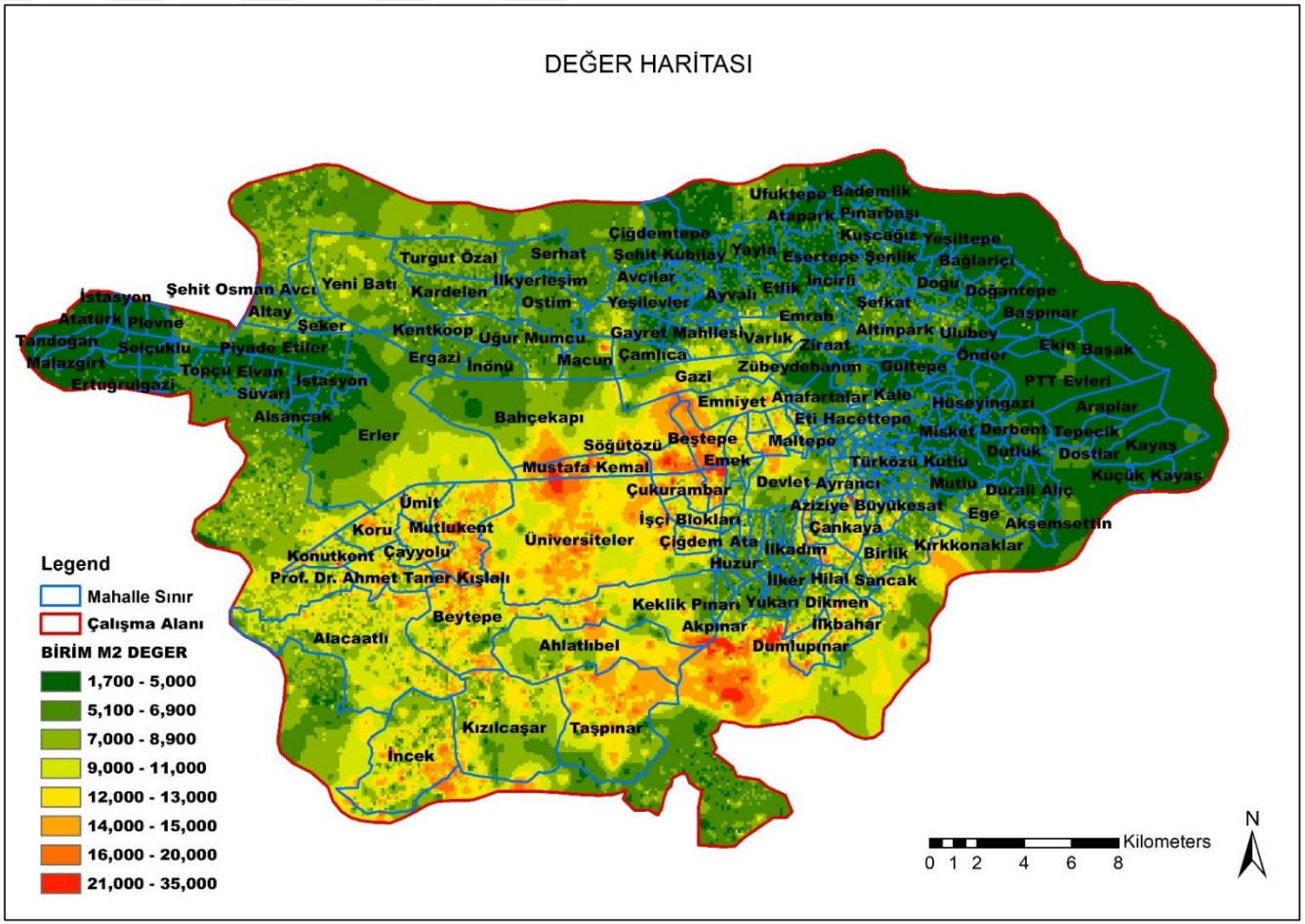


Şekil 5.24 Konut 2 verisi için bilinen değerler ile RO algoritması ile tahmin edilen değerler

5.8 Gayrimenkul Değer Haritası Sonuçları

Çalışmanın son aşamasında gayrimenkul değerleri ile konut türünde Ankara ili çalışma alanı sınırları içerisinde değer haritası oluşturulmuştur. Raster tabanlı haritalarda ise tüm pikseller bir değer alacağı için yakındaki gayrimenkul değerleri kullanılarak enterpolasyon yöntemi ile değer haritası elde edilebilinmektedir. Enterpolasyon yöntemleri çok çeşitli olup bu çalışmada Ters Mesafe Ağırlıklandırma (Inverse Distance Weighting, IDW) yöntemi ile bilinen değerler kullanılarak bilinmeyen değerler hesaplanmış ve raster tabanlı değer haritası çalışma alanı için oluşturulmuştur. IDW çalışma prensibine göre bilinmeyen noktaya en yakın noktalardan daha çok etkilenir yani bir nokta tahmin edilen hücrenin merkezine ne kadar yakınsa ortalama alma işleminde o kadar fazla etkiye veya ağırlığa sahip olmaktadır. Bu şekilde çalışma alanı olan Ankara İli için sadece konut değerinde yeterli ve homojen dağılmış veri bulunması nedeniyle konut gayrimenkul değer haritası oluşturulmuştur.

Şekil 5.25'de tez çalışmasında kullanılan veriler ile konut değeri için IDW enterpolasyon yöntemi kullanılarak raster formatında değer haritası örnek olarak oluşturulmuş olup haritada kırmızı değerler en yüksek TL/m^2 değerini göstermektedir. Mustafa Kemal, Emek, Beştepe mevkilerinde konut değerinin çok yüksek olduğu haritadan görülebilmektedir.



Şekil 5.25 Çalışma alanı için konut türü gayrimenkul değer haritası

6. GENEL DEĞERLENDİRME VE SONUÇ

Geleneksel ve matematiksel yöntemler ile yapılan çalışmalar yerini son yıllarda matematiksel yöntemlerin daha gelişmişi olan ve ileri/gelişmiş yöntemler diye ifade edilen makine öğrenme adlı yönteme bırakmıştır. Teknolojinin ilerlemesi ve depolanabilen veri boyutunun artması ile büyük veri kavramı ortaya çıkmıştır. Tam bu noktada verilerin geleneksel/istatiksel yöntemler ile işlenmesi vakit ve işgücü kaybına neden olmaktadır. Gün geçtikçe büyük veri haline dönüşen verilerden bir tanesi de gayrimenkullere ait verilerdir. Gayrimenkullerin değeri ve bu değeri etkileyecek nicel/nitel değişkenlerin artması ile gayrimenkul alanında da büyük veri oluşmuş, bu veriler ile gayrimenkullerin toplu değerlemesi son yıllarda özellikle MÖ algoritmaları yapılmaktadır.

Gayrimenkullerin toplu değerlemesinde her bir gayrimenkulün değerlemesini geleneksel yöntemlerle bireysel olarak yapmak zaman ve iş gücü kaybı olarak görülmektedir. Bu sebeple son zamanlarda modern yöntemler olan MÖ algoritmaları toplu değerlemelerde kullanılmaya başlanmıştır. Tez çalışmasında farklı makine öğrenme algoritmaları kullanılarak farklı türde gayrimenkullerin toplu değerleme işlemi yapılmıştır. Bu sayede bu MÖ algoritmalardan hangisinin toplu değerlemeye daha uygun olduğu tespit edilmiş ve tüm ülke için MÖ algoritmaları ile toplu değerleme yapılabilirliği için örnek oluşmuştur. Böylece Türkiye'de sorun olan gayrimenkullerin değerleri hızlı bir şekilde bulunabilecek ve birçok alanda kullanılabilecektir.

Literatürde ayrıntılı olarak incelendiği zaman dünyada genellikle konut ve arsa türünde MÖ algoritmalarını kullanarak toplu değerleme yapılan farklı çalışmaların yapıldığı görülmektedir. Bu çalışmalarda kullanılan veri sayısı, değişken sayısı çok fazla olmayıp, özellikle genellikle bir MÖ algoritması tek bir gayrimenkul türünde toplu değerleme amacıyla kullanıldığı gözlemlenmiştir. Tez çalışmasının özgün değerini ise gayrimenkul türü olarak 4 farklı; arsa, arazi, ticari ve konut verilerinin toplu değerleme amaçlı test edilmesi ile 5 adet MÖ algoritmasının bu veriler üzerinde denenmesi ve kıyaslanarak en başarılı algoritmanın bulunması oluşturmaktadır. Diğer taraftan yine literatür

incelendiğinde, konut türü gayrimenkul veri sayısı genellikle 3-4 haneli olmakla beraber bu tez çalışmasında kullanılan konut türünde yaklaşık 1.4 milyon adet gayrimenkul verisinin kullanılması bu tez çalışmasına özgünlük katan diğer bir noktadır.

Tez çalışmasında, kullanılan arsa, arazi, ticari ve konut türü gayrimenkullerde gayrimenkullerin değerini etkileyecek değişken sayısı artırılmak amacıyla değeri etkileyen en önemli kriterlerden biri olan mekânsal bilginin kullanımıyla değişken sayısı artırılmıştır. Bilindiği gibi, mekânsal bilgi deyince akla gelen alan Coğrafi Bilgi Sistemleri' dir. Modern yöntemlerle hem görüntüleme hem işlem açısından CBS ile iç içedir. CBS, mekânsal veriyi toplama, saklama, işleme ve görüntüleme amaçlı birçok farklı alanda kullanılan popüler bir bilim dalıdır. CBS'nin bu özelliğinden yararlanılarak gayrimenkullerin bilinen konumları ile önemli yerlere olan mesafeler ayrıca toplu değerlemede kullanılacak değişken verisi olarak hesaplanmıştır. Son olarak, CBS yardımı ile gayrimenkullerin eğer bilgiler kullanılarak görsel olarak değer haritaları oluşturulmuş ve sonuç ürün alınmıştır.

Tez çalışmasında kullanılan veri türleri, gayrimenkul sayısı ile bu gayrimenkullere ait değişken sayısına bakıldığında konut türünde toplamda yaklaşık 1.4 milyon gayrimenkul ve 22 adet değişkenden, bir diğer konut verisinde 2,353 adet gayrimenkul ve 47 adet değişkenden, arsa türünde, 209 adet gayrimenkul ve 11 adet değişkenden, arazi türünde 107 adet gayrimenkul ve 12 adet değişkenden ve son olarak ticari tür gayrimenkulde 109 adet gayrimenkul ve 14 değişkenden yararlanılmıştır. MÖ algoritmaları için bu verilerden %80'i eğitim ve %20'si test verisi amaçlı kullanılmıştır. Arsa, arazi ve ticari tür gayrimenkullerde veri sayısı az olduğu için 5 katlı çapraz doğrulama ile test ve eğitim verisi her defasında değiştirilmiştir. Böylece bu veri türlerinde her bir veri test verisi olarak kullanılarak modelin güvenilirliği ve uyum iyiliği ölçütleri daha sağlamlaştırılmıştır.

Çalışmada 4 farklı gayrimenkul türü için toplu değerleme işlemi yapmak amacıyla 5 farklı MÖ algoritması test edilmiştir. Bu algoritmalar; Rasgele Orman (RO) Regresyonu, Destek Vektör Regresyonu (DVR), K en yakın komşuluk (k-nn), Yapay Sinir Ağları

(YSA) ve son olarak Çoklu Regresyon Analizi (ÇRA) şeklinde olmakla beraber ÇRA için doğrusal, yarı-logaritmik ve tam-logaritmik olmak üzere 3 farklı model denenmiştir. ÇRA için model sayısının 3'e çıkmasıyla beraber aslında toplamda 7 farklı algoritma ile toplu değerleme işlemi 4 farklı gayrimenkul türünde yapılmış ve sonuçları ayrıntılı olarak test edilmiştir. Bu algoritmalarının yanı sıra Temel Bileşenler Analizi (TBA) ise değişken sayısını ve aralarındaki korelasyonu azaltmak amacıyla ayrıca kullanılmıştır. İşlem süreleri dikkate alındığında DVR algoritmasının en uzun işlem süresi olduğu bunu ise YSA algoritmasının takip ettiği, ÇRA algoritmalarının daha hızlı sonuç verdiği gözlemlenmiştir.

MÖ algoritmalarını kullanılarak yapılan toplu değerleme sonuçlarının ayrıntılı olarak kalite analizleri yapılmış ve incelenmiştir. Gayrimenkullerin değerlemesinde %80 ve üstü doğruluk kabul edilebilir bir sonuç olarak görülmektedir. Analiz sonucunda, 4 farklı gayrimenkul türü ve değişkenleri için en iyi toplu değerleme sonuçlarını veren algoritmaların başında tüm gayrimenkul türlerinde RO algoritmasının olduğu tespit edilmiştir. 4 farklı gayrimenkul türünde de %81'in üstünde RO algoritması ile başarı sağlanmıştır. RO algoritması sonrasında ise, konut türü hariç diğer tüm türlerde Çoklu DR analizinin RO algoritması kadar başarılı olduğu görülmüştür. Sadece çok fazla veriye sahip olan konut türünde YSA algoritmaları RO algoritmasından sonra en başarılı sonucu veren algoritma olduğu bulunmuştur. Yine YSA algoritması genel olarak tüm gayrimenkul türlerinde en az % 65 üzerinde doğruluk sağlanmıştır.

ÇRA türleri içinde en başarılı model doğrusal regresyon olsa da tam-logaritmik regresyon da tüm gayrimenkul türlerinde kabul edilebilir sonuç vermiştir. Bunun yanı sıra, yarı-logaritmik regresyon analizi regresyon türleri ile MÖ algoritmaları içinde en başarısız sonucu veren algoritma olduğu tespit edilmiştir. Diğer yandan k-nn algoritması da 4 tür gayrimenkul için çok yüksek başarı elde edememiştir. DVR algoritmasının ise çok fazla veri ve değişkene sahip konut verisi dışında %72'nin üstünde doğruluk sağladığı görülmüştür.

Gayrimenkulün değerine en çok etki eden değişkenler ise konut türü için ısıtma ve önemli yerlere olan mesafeler, arsa, arazi ve ticari türü gayrimenkul için yüzölçümü olduğu tespit edilmiştir. Önemli yerlere olan mesafeler içinde ise gene olarak 4 tür gayrimenkulde okul, üniversite ve parklara olan mesafelerin gayrimenkulün en çok değerini etkilediği görülmüştür.

Ticari tür gayrimenkul veri setinde ise yarı-logaritmik regresyon analizi hariç bu tez çalışmasında kullanılan tüm algoritmaların başarılı sonuç verdiği göze çarpmaktadır. Bunun yanı sıra arsa türü gayrimenkulde ise yarı-logaritmik regresyonda dahil olmak üzere tüm MÖ algoritmalarının %70'in üstünde doğruluk verdiği görülmüştür. Arazi türü gayrimenkulde ise RO algoritması ile bu tez çalışmasında elde edilen en yüksek doğruluk sağlanmıştır (R^2 =0.9181).

Sonuç olarak, tez çalışması ile Türkiye'ye ait veriler Ankara ili için değerleme firmaları ile güvenilir emlak endeks sitesi olan Endeksa şirketi tarafından temin edilmiş ve Türkiye'de Ankara İlinden yola çıkılarak toplu değerlemenin yapılabilirliği test edilmiştir. Çalışmada gayrimenkul türü olarak 4 farklı; arsa, arazi, ticari ve konut verilerinin toplu değerleme amaçlı test edilmesi ile 5 (4+3) adet MÖ algoritmasının bu veriler üzerinde denenmesi ve kıyaslanarak en başarılı algoritma bulunmuştur. Bu test sonucunda RO ve DR algoritmalarının tüm gayrimenkul türlerinde başarılı olduğu bulunmuştur. Ayrıca gayrimenkullerin sahip olduğu mekânsal bilgi sayesinde bu veriler harita üzerinde yansıtılarak konut türü gayrimenkul için Ankara ili değer haritası oluşturulmuştur. Öte yandan her il ve ilçenin farklı dinamikleri olabileceği bunların gayrimenkulün değerini etkileyen değişken olarak eklenmesi/çıkarılması gerektiği de ayrıca vurgulanmalıdır. Örneğin üniversite olmayan bir il/ilçe için üniversitelere olan mesafe çıkarılmalı varsa daha önemli bir değişken tespit edilerek MÖ algoritmalarıyla toplu değerleme uygulanmalıdır.

Çalışma sonucunda, karşılaşılan zorluklar ise gayrimenkule ait verilere ulaşmanın zorluğudur. Bu sebeple, her il/ilçe belediyelerinin bir taşınmaz değerleme birimi kurulması ve gayrimenkulün değerini etkileyen tüm gerekli değişkenlerin tutulacağı bir

veri tabanı oluşturması önerilmektedir. Gayrimenkule ait veriler kullanılarak toplu değerleme yapılması ve değer haritalarının üretilerek birçok alanda kullanılması bu çalışmanın en önemli vurgusunu oluşturmaktadır.

KAYNAKLAR

Abidoye, R. B. ve Chan, A. P. 2016. Critical determinants of residential property value: professionals' perspective. Journal of Facilities Management.

Abidoye, R. B., Chan, A. P., Abidoye, F. A. ve Oshodi, O. S. 2019. Predicting property price index using artificial intelligence techniques: Evidence from Hong Kong. International Journal of Housing Markets and Analysis, 12(6), 1072-1092.

Afonso, B., Melo, L., Oliveira, W., Sousa, S. ve Berton, L. 2019. Housing prices prediction with a deep learning and random forest ensemble, Anais do XVI Encontro Nacional de Inteligência Artificial e Computacional, Bildiri Özetleri Kitabı.

Ahmed S, Rahman M, I, M. ve S, R. 2014. House rent estimation in Dhaka city by multi layer perceptions neural network. International Journal of u-and e-Service, Science and Technology, 7(4), 287-300.

Akar, A. U. ve Yalpır, S. 2021. Using svr and mra methods for real estate valuation in the smart cities. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences-ISPRS Archives.

Al-Gbury, O. ve Kurnaz, S. 2020. Real estate price range prediction using artificial neural network and grey wolf optimizer, 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Bildiri Özetleri Kitabı.

Amri, S. ve Tularam, G. A. 2012. Performance of multiple linear regression and nonlinear neural networks and fuzzy logic techniques in modelling house prices. Journal of Mathematics and Statistics, 8(4), 419-434.

Antipov, E. A. ve Pokryshevskaya, E. B. 2012. Mass appraisal of residential apartments: An application of Random forest for valuation and a CART-based approach for model diagnostics. Expert Systems with Applications, 39(2), 1772-1778.

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

Borst, R. A. 1991. Artificial neural networks: the next modelling/calibration technology for the assessment community. Property Tax Journal, 10(1), 69-94.

Breiman, L. 1996. Bagging predictors. Machine learning, 24(2), 123-140.

Breiman, L. 2001. Random forests. Machine learning, 45(1), 5-32.

Breiman, L. ve Cutler, A. 2005. Random Forests. Berkeley. In.

Bulut, B., Allahverdi, N., Kahramanli, H. ve Yalpir, S. 2011. A SVR based approach on real-estate valuation. International Journal of Arts & Sciences, 4(15), 323.

Bulut Nas, B. 2011. YSA ve DVM yöntemleri ile taşınmaz değerlemesi için bir yaklaşım geliştirme Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü.

Cechin, A., Souto, A. ve Gonzalez, M. A. 2000. Real estate value at Porto Alegre city using artificial neural networks, Proceedings. Vol. 1. Sixth Brazilian Symposium on Neural Networks, Bildiri Özetleri Kitabı.

Čeh, M., Kilibarda, M., Lisec, A. ve Bajat, B. 2018. Estimating the performance of random forest versus multiple regression for predicting prices of the apartments. ISPRS international journal of geo-information, 7(5), 168.

Chen, H., Zhang, Z., Yin, W., Zhao, C., Wang, F. ve Li, Y. 2022. A study on depth classification of defects by machine learning based on hyper-parameter search. Measurement, 189, 110660.

Dellstad, M. 2018. Comparing three machine learning algorithms in the task of appraising commercial real estate. In.

Demirel, S. B., Reyhan, O., Atasever, Ü. H. ve Hayri, M. 2016. Daire Tipi Konutların Değerlemesinde Yapay Sinir Ağlarının Kullanımı, Uzaktan Algılama CBS Sempozyumu (UZAL-CBS 2016), 5-7 Ekim 2016, Bildiri Özetleri Kitabı. Adana

Derinpınar, M. A. ve Aydınoğlu, A. Ç. 2015. Bulanık Mantık ile Coğrafi Bilgi Teknolojilerini kullanarak taşınmaz değerlemesi. 15. Türkiye Harita Bilimsel ve Teknik Kurultayı, 25, 28.

Dimopoulos, T. ve Moulas, A. 2016. A proposal of a mass appraisal system in Greece with CAMA system: Evaluating GWR and MRA techniques in Thessaloniki Municipality. Open geosciences, 8(1), 675-693.

Do, A. Q. ve Grudnitski, G. 1992. A neural network approach to residential property appraisal. The Real Estate Appraiser, 58(3), 38-45.

Erdem, N. 2017a. Toplu (küme) değerleme uygulama örnekleri ve ülkemiz için öneriler. TMMOB Harita ve Kadastro Mühendisleri Odası, 16, 3-6.

Erdem, N. 2017b. Türkiye için bir taşınmaz değerleme sistemi yaklaşımı. Geomatik, 2(1), 18-39.

Ge, J., Runeson, G., Lam, K. ve Chan, K. 2004. Forecasting Hong Kong Property Prices: Multiple Regression Method vs An Artificial Neural Network Approach, World of Construction Project Management Conference, Bildiri Özetleri Kitabı.

Guan, J., Shi, D., Zurada, J. M. ve Levitan, A. S. 2014. Analyzing massive data sets: an adaptive fuzzy neural approach for prediction, with a real estate illustration. Journal of organizational computing and electronic commerce, 24(1), 94-112.

Gunn, S. R. 1998. Support vector machines for classification and regression. ISIS technical report, 14(1), 5-16.

Hong, J., Choi, H. ve Kim, W.-s. 2020. 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(3), 140-152.

Kontrimas, V. ve Verikas, A. 2007. Neural networks based screening of real estate transactions. Neural Network World, 17(1), 17.

Kontrimas, V. ve Verikas, A. 2011. The mass appraisal of the real estate by computational intelligence. Applied Soft Computing, 11(1), 443-448.

Lam, K. C., Yu, C. ve Lam, C. K. 2009. Support vector machine and entropy based decision support system for property valuation. Journal of Property Research, 26(3), 213-233.

Lam, K. C., Yu, C. ve Lam, K. 2008. An artificial neural network and entropy model for residential property price forecasting in Hong Kong. Journal of Property Research, 25(4), 321-342.

Lasota, T., Łuczak, T. ve Trawiński, B. 2011. Investigation of random subspace and random forest methods applied to property valuation data, International conference on computational collective intelligence, Bildiri Özetleri Kitabı.

Lee, T.-W. ve Chen, K. 2008. Prediction of House Unit Price in Taipei City Using Support Vector Regression.

Lenk, M. M., Worzala, E. M. ve Silva, A. 1997. High-tech valuation: should artificial neural networks bypass the human valuer? Journal of Property Valuation and Investment.

Lerman, P. 1980. Fitting segmented regression models by grid search. Journal of the Royal Statistical Society: Series C (Applied Statistics), 29(1), 77-84.

Limsombunchao, V. 2004. House price prediction: hedonic price model vs. artificial neural network.

Lin, H. ve Chen, K. 2011. Predicting price of Taiwan real estates by neural networks and support vector regression, Proc. of the 15th WSEAS Int. Conf. on Syst, Bildiri Özetleri Kitabı.

McCluskey, W. 1996. Predictive accuracy of machine learning models for the mass appraisal of residential property. New Zealand Valuers Journal, 40-46.

McCluskey, W., Davis, P., Haran, M., McCord, M. ve McIlhatton, D. 2012. The potential of artificial neural networks in mass appraisal: the case revisited. Journal of Financial Management of Property and Construction.

McGreal, S., Adair, A., McBurney, D. ve Patterson, D. 1998. Neural networks: the prediction of residential values. Journal of Property Valuation and Investment.

Mora-Esperanza, J. G. 2004. Artificial intelligence applied to real estate valuation: An example for the appraisal of Madrid. Catastro, April, 1, 255-265.

Morano, P. ve Tajani, F. 2013. Bare ownership evaluation. Hedonic price model vs. artificial neural network. International Journal of Business Intelligence and Data Mining, 8(4), 340-362.

Mukhlishin, M. F., Saputra, R. ve Wibowo, A. 2017. Predicting house sale price using fuzzy logic, Artificial Neural Network and K-Nearest Neighbor, 2017 1st International Conference on Informatics and Computational Sciences (ICICoS), Bildiri Özetleri Kitabı.

Musa, A. G., Daramola, O., Owoloko, A. ve Olugbara, O. 2013. A neural-CBR system for real property valuation. Journal of Emerging Trends in Computing and Information Sciences, 4(8), 611-622.

Nghiep, N. ve Al, C. 2001. Predicting housing value: A comparison of multiple regression analysis and artificial neural networks. Journal of real estate research, 22(3), 313-336.

Özkan, G., Yalpır, Ş. ve Uygunol, O. 2007. An investigation on the price estimation of residable real-estates by using artificial neural network and regression methods, XIIth Applied Stochastic Models and Data Analysis International conference (ASMDA), Chania, Crete, Greece, Bildiri Özetleri Kitabı.

Pearson, K. 1908. On a mathematical theory of determinantal inheritance, from suggestions and notes of the late WFR Weldon. Biometrika, 6(1), 80-93.

Peterson, S. ve Flanagan, A. 2009. Neural network hedonic pricing models in mass real estate appraisal. Journal of real estate research, 31(2), 147-164.

Pi-ying, L. 2011. Analysis of the mass appraisal model by using artificial neural network in Kaohsiung city. Journal of Modern Accounting and Auditing, 7(10), 1080.

Ravikumar, A. S. 2017. Real estate price prediction using machine learning Dublin, National College of Ireland.

Rossini, P. 1998. Improving the results of artificial neural network models for residential valuation, Fourth Annual Pacific-Rim Real Estate Society Conference, Perth, Western Australia, Bildiri Özetleri Kitabı. 19,

Sampathkumar, V., Santhi, M. H. ve Vanjinathan, J. 2015. Forecasting the land price using statistical and neural network software. Procedia Computer Science, 57, 112-121.

Samsuddin Ahmed, M., Rahman, M. ve Islam, S. 2014. House rent estimation in Dhaka city by multi layer perceptions neural network. International Journal of u-and e-Service, Science and Technology, 7(4), 287-300.

Saraç, E. 2012. Yapay sinir ağları metodu ile gayrimenkul değerleme İstanbul Kültür Üniversitesi/Fen Bilimleri Enstitüsü/İnşaat Mühendisliği.

Savaş, B. 2019. Makine öğrenme algoritmalarının konut değer tahmininde kullanımı: Ankara Gölbaşı uygulaması Konya Teknik Üniversitesi.

Sawant, R., Jangid, Y., Tiwari, T., Jain, S. ve Gupta, A. 2018. Comprehensive analysis of housing price prediction in pune using multi-featured random forest approach, 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Bildiri Özetleri Kitabı.

Selim, H. 2009. Determinants of house prices in Turkey: Hedonic regression versus artificial neural network. Expert systems with Applications, 36(2), 2843-2852.

Surgelas, V., Arhipova, I. ve Pukite, V. 2019. Analysis of different approaches to real estate appraisal. Researchfor Rural Devalopment, 1.

Tabales, J. M. N., Caridad, J. M. ve Carmona, F. J. R. 2013. Artificial neural networks for predicting real estate price. Revista de Métodos Cuantitativos para la Economía y la Empresa, 15, 29-44.

Tanrıvermiş, H. 2017. Gayrimenkul değerleme esasları. SPL Sermaye Piyasası Lisanslama Sicil ve Eğitim Kuruluşu, Lisanslama Sınavları Çalışma Kitapları Ders Kodu, 1014.

Tanrıvermiş, H. ve Aliefendioğlu, Y. 2008. Yapı Değerlemesinin Teorik Esasları ve Uygulamaları: Türkiye'de kamulaştırma, Emlak vergisi ve imar düzenlemeleri yönünden bir inceleme. Üçüncü Sektör Kooperatifçilik, 43(4), 30-60.

Tay, D. P. ve Ho, D. K. 1992. Artificial intelligence and the mass appraisal of residential apartments. Journal of Property Valuation and Investment.

Torres-Pruñonosa, J., García-Estévez, P. ve Prado-Román, C. 2021. Artificial neural network, quantile and semi-log regression modelling of mass appraisal in housing. Mathematics, 9(7), 783.

Vapnik, V. 1998. Statistical learning theory new york. NY: Wiley, 1(2), 3.

Vargason, D. 2019. Data Mining Techniques for Predicting Real Estate Trends.

Varma, A., Sarma, A., Doshi, S. ve Nair, R. 2018. House price prediction using machine learning and neural networks, 2018 second international conference on inventive communication and computational technologies (ICICCT), Bildiri Özetleri Kitabı.

Walacik, M., Grover, R. ve Adamuscin, A. 2013. Valuation systems in Poland, Slovakia and the United Kingdom–comparative study. Real Estate Management and Valuation, 21(4), 75-86.

Wilson, I. D., Paris, S. D., Ware, J. A. ve Jenkins, D. H. 2002. Residential property price time series forecasting with neural networks. In Applications and Innovations in Intelligent Systems IX (pp. 17-28). Springer.

Worzala, E., Lenk, M. ve Silva, A. 1995. An exploration of neural networks and its application to real estate valuation. Journal of Real Estate Research, 10(2), 185-201.

Xin, J. G. ve Runeson, G. 2004. Modeling property prices using neural network model for Hong Kong. International Real Estate Review, 7(1), 121-138.

Yacim, J. A. ve Boshoff, D. G. B. 2018. Combining BP with PSO algorithms in weights optimisation and ANNs training for mass appraisal of properties. International journal of housing markets and analysis.

Yacim, J. A. ve Boshoff, D. G. B. 2020. Neural networks support vector machine for mass appraisal of properties. Property Management, 38(2), 241-272.

Yoo, S., Im, J. ve Wagner, J. E. 2012. Variable selection for hedonic model using machine learning approaches: A case study in Onondaga County, NY. Landscape and Urban Planning, 107(3), 293-306.

Yu, H. ve Wu, J. 2016. Real estate price prediction with regression and classification. CS229 (Machine Learning) Final Project Reports.

Ziweritin, S., Ukegbu, C. C., Oyeniran, T. A. ve Ulu, I. O. 2021. A Recommendation Engine to Estimate Housing Values in Real Estate Property Market. Int. J. Sci. Res. in Computer Science and Engineering Vol, 9(1).

Zurada, J., Levitan, A. ve Guan, J. 2011. A comparison of regression and artificial intelligence methods in a mass appraisal context. Journal of real estate research, 33(3), 349-388.

EKLER

EK-1 5 katmanlı çapraz doğrulama tüm MÖ algoritmaları ve arsa gayrimenkul türleri için sonuçlar

EK-2 5 katmanlı çapraz doğrulama tüm MÖ algoritmaları ve arazi gayrimenkul türleri için sonuçlar

EK-3 5 katmanlı çapraz doğrulama tüm MÖ algoritmaları ve ticari gayrimenkul türleri için sonuçlar

EK 1: 5 katmanlı çapraz doğrulama tüm MÖ algoritmaları ve arsa türü gayrimenkul için sonuçlar

Arsa tür gayrimenkul tipi için 5 katlı Çapraz Doğrulama ile MÖ Algoritma Sonuçları						
		R^2	Düzeltilmiş R^2	HKO	KOKH	OMH
YSA	1	0.790027	0.710382	1089260	1043.676	731.8607
	2	0.67333	0.553551	1214863	1102.208	764.5794
	3	0.85294	0.799019	536786.6	732.6572	567.7274
	4	0.717171	0.613467	1482072	1217.404	961.3142
	5	0.772376	0.686036	942834.8	970.9968	687.8052
	ORTALAMA	**0.761169**	**0.672491**	**1053163**	**1013.388**	**742.6574**
RO	1	0.83633	0.774248	849057.2	921.443	594.6216
	2	0.763187	0.676356	880691.4	938.4516	653.7614
	3	0.907649	0.873787	337092.2	580.5965	452.5456
	4	0.873146	0.826633	664734.9	815.3128	504.9719
	5	0.784616	0.702918	892137.1	944.5301	648.3538
	ORTALAMA	**0.832986**	**0.7707884**	**724742.6**	**840.0668**	**570.8509**
DR	1	0.850834	0.794253	773816.6	879.6685	589.236
	2	0.734378	0.636984	987829.4	993.8961	655.4585
	3	0.912595	0.880547	319037.9	564.8344	461.7122
	4	0.881158	0.837583	622750.2	789.1452	587.3866
	5	0.755066	0.66216	1014533	1007.24	645.0855
	ORTALAMA	**0.826806**	**0.7623054**	**743593.4**	**846.9569**	**587.7758**
YLR	1	0.724736	0.620326	1427962	1194.974	879.8601
	2	0.708477	0.601585	1084157	1041.228	831.3877
	3	0.737253	0.640912	959061.5	979.3168	791.8319
	4	0.722473	0.620713	1454289	1205.939	791.8319
	5	0.703628	0.591211	1227594	1107.969	870.383
	ORTALAMA	**0.719313**	**0.6149494**	**1230613**	**1105.885**	**833.0589**
TLR	1	0.788004	0.707592	1099753	1048.691	640.5768
	2	0.783847	0.704591	803857.3	896.5809	623.8756
	3	0.882501	0.839419	428885	654.8931	498.9529
	4	0.835815	0.775614	860355.3	927.5534	647.4011
	5	0.800295	0.724545	827191.3	909.5006	661.4498
	ORTALAMA	**0.818092**	**0.7503522**	**804008.4**	**887.4438**	**614.4512**
DVR	1	0.859663	0.806431	728015.6	853.2383	579.8279
	2	0.67908	0.56141	1193480	1092.465	714.1886
	3	0.814608	0.74663	676706.9	822.622	736.1565
	4	0.810395	0.740874	993559.3	996.7745	742.2325
	5	0.753323	0.659757	1021751	1010.817	667.6056
	ORTALAMA	**0.783414**	**0.7030204**	**922702.5**	**955.1833**	**688.0022**
k-nn	1	0.800041	0.724195	1037307	1018.483	717.0108
	2	0.714714	0.610108	1060962	1030.03	724.8257
	3	0.804302	0.732546	714323.1	845.1764	639.3939
	4	0.817616	0.750741	955724.1	977.6114	709.6619
	5	0.780925	0.697827	907425.6	952.5889	716.3555
	ORTALAMA	**0.78352**	**0.7030834**	**935148.3**	**964.7779**	**701.4496**

EK 2: 5 katmanlı çapraz doğrulama tüm MÖ algoritmaları ve arazi türü gayrimenkul için sonuçlar

Arazi tür gayrimenkul tipi için 5 katlı Çapraz Doğrulama ile MÖ Algoritma Sonuçları						
		R^2	Düzeltilmiş R^2	HKO	KOKH	OMH
YSA	1	0.829465	0.573662	893860.121	945.4418	763.1208
	2	0.759228	0.398069	1265791.16	1125.074	783.3149
	3	0.897289	0.743222	905221.097	951.4311	796.3367
	4	0.947112	0.876594	359818.413	599.8487	436.6472
	5	0.754308	0.38577	1375965.58	1173.016	751.5715
	ORTALAMA	**0.83748**	**0.5954634**	**960131.276**	**958.9622**	**706.1982**
RO	1	0.869668	0.674171	685180.504	827.7563	433.687
	2	0.863748	0.65937	716304.521	846.3478	453.3871
	3	0.984791	0.961977	134043.883	366.12	282.0647
	4	0.969262	0.928278	209123.188	457.2999	380.5953
	5	0.903072	0.75768	542833.204	736.7722	420.1135
	ORTALAMA	**0.918108**	**0.7962952**	**457497.06**	**646.8592**	**393.9695**
DR	1	0.925756	0.814389	389152.376	623.8208	501.4788
	2	0.823885	0.559712	925873.588	962.2233	539.4436
	3	0.953412	0.883531	410588.964	640.7722	501.2909
	4	0.973308	0.937719	181593.955	426.1384	358.3245
	5	0.822347	0.555867	994922.664	997.4581	474.6442
	ORTALAMA	**0.899742**	**0.7502436**	**580426.309**	**730.0825**	**475.0364**
YLR	1	0.817273	0.543181	957766.492	978.6555	785.1429
	2	0.544687	0.138282	2393675.24	1547.151	1170.424
	3	0.420272	-0.44932	5109294.08	2260.375	1811.146
	4	0.834556	0.613963	1125580.11	1060.934	880.7183
	5	0.575344	-0.06164	2378229.41	1542.151	1009.522
	ORTALAMA	**0.638426**	**0.1568932**	**2392909.07**	**1477.853**	**1131.391**
TLR	1	-0.40025	-2.50062	4524808.52	2709.133	1545.655
	2	0.555884	-0.110289	2334808.52	1528.008	1189.583
	3	0.330263	-0.674342	5902568.28	2429.52	1700.972
	4	0.784729	0.497701	1464568.07	1210.193	975.4159
	5	0.76398	0.40995	1321799.5	1149.695	745.8741
	ORTALAMA	**0.406922**	**-0.47552**	**3109710.58**	**1805.31**	**1231.5**
DVR	1	0.873535	0.683838	662864.588	814.165	707.6504
	2	0.8083	0.520751	1007804.78	1003.895	637.6537
	3	0.899062	0.747656	889590.847	943.1812	844.5227
	4	0.954617	0.894107	308754.351	555.6567	450.9497
	5	0.734074	0.335185	1489284.42	1220.362	833.2891
	ORTALAMA	**0.853918**	**0.6363074**	**871659.796**	**907.452**	**694.8131**
k-nn	1	0.380577	-0.548557	3246705.06	1801.862	1394.074
	2	0.149624	-1.125939	4470603.21	2114.38	1734.377
	3	0.078736	-1.303159	8119339.73	2849.446	2489.154
	4	0.325971	-0.572733	4585668.54	2141.417	1739.578
	5	0.012551	-1.468622	5530076.43	2351.611	1780.884
	ORTALAMA	**0.189492**	**-1.003802**	**5190478.6**	**2251.743**	**1827.613**

EK 3: 5 katmanlı çapraz doğrulama tüm MÖ algoritmaları ve ticari türü gayrimenkul için sonuçlar

Ticari tür gayrimenkul tipi için 5 katlı Çapraz Doğrulama ile MÖ Algoritma Sonuçları						
		R^2	Düzeltilmiş R^2	HKO	KOKH	OMH
YSA	1	0.570545	1.630128	6962592	2638.672	1695.311
	2	0.61815	-0.002356	4612334	2147.634	1092.683
	3	0.913988	0.774217	562886.5	750.2576	591.2923
	4	0.607871	-0.029339	5964490	2442.23	1518.776
	5	0.572481	-0.221483	1865193	1365.721	909.3573
	ORTALAMA	**0.656607**	**0.4302334**	**3993499**	**1868.903**	**1161.484**
RO	1	0.830127	0.514648	247929.2	497.9249	360.3498
	2	0.675599	0.148447	3918412	1979.498	907.28
	3	0.924903	0.802871	491452.3	701.0366	543.8361
	4	0.817792	0.521705	2771476	1664.775	902.639
	5	0.928875	0.796785	310307.6	557.0526	426.672
	ORTALAMA	**0.835459**	**0.5568912**	**1547916**	**1080.057**	**628.1554**
DR	1	0.660681	0.030517	495234.8	703.7292	554.9984
	2	0.62468	0.014785	4533457	2129.192	873.7816
	3	0.930908	0.818633	452157.3	672.4264	558.8546
	4	0.959819	0.894525	611175.5	781.7771	627.472
	5	0.939533	0.827238	263805.7	513.6202	423.6292
	ORTALAMA	**0.823124**	**0.5171396**	**1271166**	**960.1489**	**607.7472**
YLR	1	-1.01317	-4.751926	2938220	1714.124	1337.084
	2	0.512175	-0.280539	5892391	2427.425	1654.624
	3	0.715352	0.252798	1862809	1364.848	1222.308
	4	0.611002	-0.02112	5916866	2432.461	1719.56
	5	0.074956	-1.642983	4035811	2008.933	1634.681
	ORTALAMA	**0.180062**	**-1.288754**	**4129219**	**1989.558**	**1513.651**
TLR	1	0.455155	-0.556699	795198.9	891.7393	587.5226
	2	0.60324	-0.041495	4792432	2189.162	1169.188
	3	0.856757	0.623988	937414.8	968.2018	647.8261
	4	0.813706	0.510978	2833632	1683.339	968.8023
	5	0.869193	0.626265	570689.4	755.4399	469.0795
	ORTALAMA	**0.71961**	**0.2326074**	**1985873**	**1297.577**	**768.4837**
DVR	1	0.59866	0.199616	4188558	2046.597	1251.659
	2	0.640141	0.05537	4346708	2084.876	1118.602
	3	0.852071	0.611685	968086.3	983.9138	848.9938
	4	0.709099	0.236386	4424750	2103.509	1242.16
	5	0.829086	0.511674	745669	863.5213	654.2497
	ORTALAMA	**0.725811**	**0.3229462**	**2934754**	**1616.483**	**1023.133**
k-nn	1	0.545482	-0.298624	663367.7	814.4739	627.7124
	2	0.664612	0.119607	4051123	2012.74	998.768
	3	0.797143	0.467499	1327549	1152.193	822.8014
	4	0.803559	0.484343	2987968	1728.574	946.638
	5	0.792447	0.406993	905516.5	951.5863	631.9667
	ORTALAMA	**0.720649**	**0.2359636**	**1987105**	**1331.913**	**805.5773**