

## VERİ MADENCİLİĞİ İLE BETONARME KONUT BİNALARINDA DEPREM KUVVETİ ANALİZİ

Tarkan KARAÇAY<sup>1\*</sup>, Tolga AÇIKGÖZ<sup>2</sup>

<sup>1</sup> Eskişehir Osmangazi Üniversitesi, Mühendislik Mimarlık Fakültesi, İnşaat Mühendisliği Bölümü, 26480, Eskişehir, ORCID No : <https://orcid.org/0000-0002-4893-5126>

<sup>2</sup> Eskişehir Osmangazi Üniversitesi, Mühendislik Mimarlık Fakültesi, İnşaat Mühendisliği Bölümü, 26480, Eskişehir, ORCID No : <https://orcid.org/0000-0003-2315-6342>

Anahtar Kelimeler	Öz
Veri madenciliği Makine öğrenmesi Deprem Deprem yönetmeliği Betonarme	<i>Bu çalışma perdesiz betonarme konut binalarına etkiyen yatay tasarım deprem yükünün veri madenciliği yöntemleri ile analiz edilmesi ve makine öğrenmesi yöntemleri ile tahminlenmesi konusundadır. Bu amaçla Türkiye Bina Deprem Yönetmeliği'ne (2018) göre eşdeğer deprem yükü yöntemi kullanılarak veri seti oluşturulmuştur. Oluşturulan veri seti üzerinde öznitelik seçimi, uç değerlerin tespit edilmesi ve silinmesi, boyut azaltma gibi veri madenciliği yöntemleri kullanılmış hangi yöntemlerle sonuçların nasıl değiştiği tartışılmıştır. Sonuçlar perdesiz betonarme binalara etkiyen deprem kuvvetinin yapay sinir ağları, rastgele orman ve destek vektör makinesi yöntemleri ile başarılı tahmini için kısa periyot harita spektral ivme katsayısının (<math>S_s</math>) ve bina toplam yüksekliğinin (<math>H_N</math>) gerekmediğini ortaya koymuştur.</i>

## EARTHQUAKE FORCE ANALYSIS FOR REINFORCEMENT RESIDENTIAL BUILDINGS WITH DATA MINING

Keywords	Abstract		
Data mining Machine learning Earthquake Earthquake code Reinforced concrete	<i>This study is about the analysis of horizontal design earthquake loads acting on reinforced concrete residential buildings without shear walls using data mining methods and the prediction of loads using machine learning methods. For this purpose, a data set was created by using the equivalent earthquake load method according to the Building Earthquake Code of Turkey (2018). Data mining methods such as feature selection, detection and removing of outlier values, dimensionality reduction were used on the created data set, and how the results changed with which methods were discussed. The results revealed that short-period spectral acceleration coefficient taken from AFAD map (<math>S_s</math>) and total building height (<math>H_N</math>) are not required for successful prediction of earthquake force acting on reinforced concrete buildings without shear wall with artificial neural networks, random forest and support vector machine methods.</i>		
Araştırma Makalesi	Research Article		
Başvuru Tarihi	: 07.03.2022	Submission Date	: 07.03.2022
Kabul Tarihi	: 17.05.2022	Accepted Date	: 17.05.2022

### 1. Giriş

Türkiye'de nüfusun %86'sı 3. derece ve üstü deprem bölgesinde yaşamaktadır. Türkiye yüz ölçümünün %96'sı ve nüfusun %99'u deprem riski altındadır (Türkoğlu, 2001). Ayrıca tüm dünyada 2020 senesinde büyüklüğü 4'ün üstünde olan 13.654 deprem olmuştur (Cetinkaya, 2021). Bu ve benzeri nedenlerle deprem

konusunda araştırmalar gittikçe artmaktadır. İlgili çalışmalar devam eden paragrafta açıklanmıştır.

Contreras, Wilkinson ve James (2021) deprem sonrası hasar tespiti ve iyileştirme süreçleriyle ilgili verileri toplamak için 39 farklı araştırmayı inceleyerek derleme bir yazı oluşturmuşlardır. Geiß ve Taubenböck (2013) deprem riskinin çok sayıda bileşenin değerlendirilebilmesi için kullanılan uzaktan algılama

\* Sorumlu yazar; e-posta : [mail@tarkankaracay.com](mailto:mail@tarkankaracay.com)



Bu eser, Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) hükümlerine göre açık erişimli bir makaledir.

This is an open access article under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>).

(remote sensing) konusunda bir derleme yazı sunmuşlardır. Yazı ilgili konudaki 251 farklı araştırmayı incelemektedir. Kahandawa, Domingo, Park ve Uma (2018) deprem hasar tahmin sistemleri üzerine bir literatür incelemesi yazmışlardır. Çalışma PACT ve SLAT isimli iki aracın karşılaştırmasını sunmaktadır. Ortega, Vasconcelos, Rodrigues, Correia ve Lourenço (2017) geleneksel mimari eserlerde hasar önleme ve güçlendirme tekniklerinin incelendiği bir literatür taraması sunmuşlardır. Çalışma çok sayıdaki geleneksel mimari yapıda uygulanan tekniklerin açıklamasını ve yazarların yorumlarını barındırmaktadır.

Veri bilimi, veri madenciliği, yapay zekâ ve makine öğrenmesi alanlarının gittikçe popülerleşmesi ile bu alanlardan yararlanılarak yapılan deprem çalışmalarının sayısı da oldukça artmıştır. İlgili çalışmalar devam eden paragrafta açıklanmıştır.

Falcone, Lima ve Martinelli (2020) yapı ve deprem mühendisliğinde yapay zekâ uygulamalarını sunan bir literatür incelemesi kaleme almışlardır. Yazıda 215 farklı araştırma incelenmiş ve açıklanmıştır. Goswami, Chakraborty, Ghosh, Chakrabarti ve Chakraborty (2018) doğal afetlerle mücadele için veri madenciliği tekniklerinin uygulanması üzerine bir literatür taraması yazmışlardır. Yazıda doğal afetlerden toplanan veriler ile tahmin, tespit ve uygun afet yönetimi stratejisi üzerine odaklanılmıştır. Ayrıca farklı veri kaynaklarından elde edilen verilerin afetlerde kullanılabilirliği tartışılmıştır. Otari ve Kulkarni (2012) 1989 ile 2011 yılları arasında yayımlanan doğal felaketlerin veri madenciliği teknikleri ile tahminlenmesi konulu 16 çalışmayı incelemektedirler. Çalışma ek olarak yazarların önerilerini ve görüşlerini içermektedir.

Yapılan literatür taraması sonucunda bu çalışmanın konusu olan tasarım deprem kuvvetlerinin veri madenciliği ya da makine öğrenmesi ile analizi konusunda herhangi bir çalışmaya ulaşılamamıştır.

Türkiye’de yapılara etkiyen deprem yükü hesabı için Türkiye Bina Deprem Yönetmeliği (devam eden bölümlerde “yönetmelik” olarak bahsedilecektir) kullanılmaktadır. Yönetmeliğe göre perdersiz betonarme konut binalarında eşdeğer deprem yükü ile yatay tasarım deprem yükünün hesabı için 5 farklı parametrenin ilgili mühendis tarafından elde edilmesi gerekmektedir (Resmi Gazete, 2018). Bu parametreler sırası ile  $S_s$ ,  $S_1$ ,  $m_t$ ,  $H_N$  ve zemin sınıfıdır. Bu parametreler maddeler halinde aşağıda açıklanmıştır.

- $S_s$  ve  $S_1$  Türkiye Deprem Tehlike Haritası’ndan elde edilebilen değerlerdir.  $S_s$  kısa periyot harita spektral ivme katsayısıdır.  $S_1$  1 saniye periyot için harita spektral ivme katsayısıdır (Bozer, 2020).
- $m_t$  toplam bina kütesidir.

- $H_N$  yönetmelikte madde 3.3.1.3’te açıklanan bina yüksekliğidir (Afet ve Acil Durum Yönetimi Başkanlığı, 2018).
- Zemin sınıfı yönetmelikte ZA, ZB, ZC, ZD, ZE ve ZF olarak sınıflandırılmış parametredir. ZA en iyi zemin sınıfını, ZE en kötü zemin sınıfını belirtir. ZF ise sahaya özel araştırma ve değerlendirme yapılmasını gerektiren zemin sınıfıdır (Aksoylu ve Arslan, 2019).

Yukarıdaki maddelerde açıklanan 5 parametre belirlendikten sonra yönetmelikte açıklanan formüller ve matematiksel yaklaşımlar kullanılarak eşdeğer deprem yükü yöntemi ile yatay tasarım deprem yükü bulunabilmektedir. Fakat yönetmeliğe göre bu 5 parametrenin tamamının bilinmesi halinde deprem kuvveti hesaplanabilmektedir. Bu nedenle aşağıdaki araştırma soruları ortaya çıkmıştır.

- Perdersiz betonarme konut binalarına etkiyen yatay tasarım deprem yükü için ilgili 5 parametreden tamamı kullanılmadan deprem yükleri başarılı bir şekilde tahminlenebilir mi?
- Parametrelerden kaç tanesi ve hangileri kullanılmadan yatay deprem yükleri başarılı bir şekilde tahminlenebilir?
- İlgili tahminlemeler için makine öğrenmesi yöntemleri kullanılabilir mi?

Yukarıdaki araştırma sorularını cevaplayabilmek adına bu çalışma ortaya konmuştur. Çalışma kapsamında öncelikle bir veri seti oluşturulmuş, ardından veri seti üzerinde çeşitli veri madenciliği yöntemleri kullanılarak sonuçlar elde edilmiştir. Elde edilen sonuçlar tartışılmıştır.

Çalışmada 2. bölümde veri setinin oluşturulma aşamaları anlatılmıştır. Bölüm 3’te metodoloji, yapılan denemeler ve elde edilen sayısal değerler sunulmaktadır. Bölüm 4’te elde edilen sonuçlar tartışılmıştır. Bölüm 5 çalışma kapsamında neden perdersiz betonarme konut binaları özelinde araştırma yapıldığını ve çalışmanın gelecekte nasıl genişletilebileceği hakkında yazarların yorumlarını içermektedir.

## 2. Veri Setinin Oluşturulması

Perdersiz betonarme konut binalarında deprem kuvvet hesabı için oluşturulan veri setinin oluşturulma aşamaları bu bölümde, aşağıdaki alt başlıklarda özetlenmiştir. Çalışma kapsamında kullanılan veri seti Wolfram Mathematica ile bu adımlara göre oluşturulmuştur. Veri seti çevrim içi paylaşım platformu olan kaggle.com üzerinde yayınlanmıştır (Karaçay, 2022a).

## 2.1. Harita Spektral İvme Katsayıları Değerlerinin Elde Edilmesi

$S_s$  ve  $S_1$  harita spektral ivme katsayıları Türkiye'deki enlem ve boylam değerlerine göre belirlenmiş ve Resmi Gazete'de tablo halinde paylaşılmıştır. İlgili tablo çalışmayı okuyanlar tarafından istenildiği takdirde kolayca indirilebilmesi için çevrim içi olarak paylaşılmıştır (Karaçay, 2022b).

Veri setinin oluşturulması için tablodan rastgele enlem ve boylam koordinatları seçilmiştir (Mathematica Random fonksiyonu ile). Seçilen koordinatlardaki  $S_s$  ve  $S_1$  değerleri kullanılarak veri setinin oluşturulmasına devam edilmiştir.

## 2.2. Bina Yüksekliği, Ağırlığı ve Zemin Sınıfının Belirlenmesi

Konut bina iç yükseklikleri Planlı Alanlar İmar Yönetmeliği kapsamında en az 2.60 metre olmalıdır (Resmi Gazete, 2017). Bu sayıya göre kat yüksekliği yaklaşık olarak en az 2.80 metre kabul edilebilir. Uygulamada ise zaman zaman kat yükseklikleri 3 metre tercih edilmektedir. Veri seti oluşturulurken 2.80 metreden 16.80 metreye kadar 2.80 metre ve 3 metrenin katları seçilerek toplam bina yüksekliği belirlenmiştir. Diğer bir söylemle 2.80 ve 3.00 değerleri rastgele belirlenen kat sayıları ile çarpılarak bina toplam yüksekliği belirlenmiştir. Ayrıca her bir katın alanının 100 ile 600 m<sup>2</sup> aralığında rastgele değiştiği ve m<sup>2</sup> başına 1 tonluk düşey yük olduğu kabul edilerek bina toplam kütlesi hesaplanmıştır.

Zemin sınıfı değerleri ZA, ZB, ZC, ZD, ZE kategorilerinden rastgele seçilerek oluşturulmuştur.

## 2.3. Deprem Kuvveti Değerlerinin Hesaplanması

Bölüm 2.1 ve 2.2'de anlatılan şekilde  $S_s$ ,  $S_1$ , bina yüksekliği ( $H_N$ ), bina toplam kütlesi  $m_t$  ve zemin sınıfı belirlendikten sonra yönetmeliğe uygun şekilde perdesiz, betonarme konut binaları için tasarım deprem kuvvetleri hesaplanarak veri setine eklenmiştir. Kısa periyot bölgesi için yerel zemin sınıfı ( $F_S$ ) ve 1.0 saniye periyot için yerel zemin etki sınıfı ( $F_1$ ) için yönetmelikteki tablolarda bulunmayan ara değerler 1. dereceden interpolasyon yapılarak belirlenmiştir. Bina doğal hakim titreşim periyodu yönetmelikteki ampirik formül  $T_{pA} = C_t H_N^{3/4}$  ile hesaplanmıştır.

## 2.4. Oluşturulan Veri Setinin Özellikleri

Veri seti 100.000 satırdan ve 6 öznitelikten oluşmaktadır. Oluşturulan veri setinin özellikleri Tablo 1 ve Tablo 2'de sunulmuştur. Oluşturulan veri setinden bir kesit Tablo 3'tedir. Veri setinde VtE tasarım deprem kuvvetinin kısaltması olup hedef (çıkıtı) özniteliktir.

Tablo 1

Veri Setinin Özellikleri

Öznitelikler	En Küçük	Ortalama	En büyük	Standart Sapma
$S_s$	0.134	0.647	2.005	0.329
$S_1$	0.047	0.179	0.536	0.083
$m_t$	100	1142.686	3600	767.534
$H_N$	2.800	9.430	16.800	4.567
VtE	22.233	725.459	5553.776	630.424

Tablo 2

Veri Setindeki Zemin Sınıfı Özniteliğinin Dağılımı

Zemin Sınıfı	ZA	ZB	ZC	ZD	ZE
Veri Sayısı	20014	20177	19721	20119	19952

Tablo 3

Veri Setinden Kesit

ss	s1	mt	hN	zeminSınıfı	vtE
0.575	0.135	1388	11.2	ZC	562.771
0.603	0.159	508	11.2	ZB	129.380
0.394	0.106	513	2.8	ZA	225.795
0.904	0.224	2155	15	ZD	1670.705
0.448	0.126	678	5.6	ZB	230.134

## 3. Metodoloji

Çalışmada izlenen metodoloji bu bölümde ve devam eden alt başlıklarda adım adım anlatılmıştır. İzlenen metodolojide açıklanan adımlar Orange Data Mining ve Weka isimli programlar ile gerçekleştirilmiştir (Demşar ve diğ., 2013; Hall ve diğ., 2009).

### 3.1. Veri Ön İşleme Süreci

#### 3.1.1. Tekrar Eden Verilerin Silinmesi, Eğitim - Test Verilerinin Ayrılması

Oluşturulan veri seti yönetmeliğin istediği girdilerin rastgele seçilmesinden (yönetmeliğe uygun şekilde) olduğu için tekrar eden veri olup olmadığının incelenmesi gereklidir. Yapılan kontrol ile verideki 17 tane satırın birbirinin tamamen aynı olduğu fark edilmiştir. Bu satırlar veri setinden silinmiştir. Böylece veri setinde 99.983 satır kalmıştır.

Makine öğrenmesi yöntemlerinin eğitimi ve testi için oluşturulan verinin ayrılması gerekmektedir. Eğitim ve test verilerinin ayrılması için farklı yüzdeler kullanılabilir. Sıklıkla kullanılan yaklaşım verinin %70 eğitim - %30 test veya %80 eğitim - %20 test şeklinde ayrılmasıdır (Akhgar, Toghraie, Sina ve Afrand, 2019; Alakus ve Turkoglu, 2020; Kadam ve

Vijayumar, 2018; Mia ve Dhar, 2016; Mothilal, Sharma ve Tan, 2020; Wei ve diğ., 2019; Yuan, Jiao, Quddus, Kwon ve Mashuga, 2019).

Bu çalışmada eğitim verisinin büyük olması ile daha başarılı makine öğrenmesi modelleri elde edilebileceği öngörülmüş ve oluşturulan veri seti %80 eğitim - %20 test olarak ayrılmıştır. Test veri seti tamamen rastgele seçilmiştir ve içerisinde eğitim veri setinden veri bulundurmamaktadır.

### 3.1.2. Verinin Anlaşılması

Verideki her bir öznelikliğin birbiri arasındaki ilişkiyi gösteren grafikler çizilmiş ve incelenmiştir. Bu grafikler yazının okunaklı olması için verilmemiş olup bunun yerine korelasyon matrisi (Şekil 1) gösterilmiştir.

	$S_s$	$S_1$	$m_t$	$H_N$	ZA	ZB	ZC	ZD	ZE
$S_s$	1	0.97	0	0	0	0	0	0	0
$S_1$		1	0	0	0	0	0	0	0
$m_t$			1	0.73	0	0.01	0	0	0
$H_N$				1	0	0	0	0	0
ZA					1	-0.25	-0.25	-0.25	-0.25
ZB						1	-0.25	-0.25	-0.25
ZC							1	-0.25	-0.25
ZD								1	-0.25
ZE									1

Şekil 1. Öznelikler Arası Korelasyon Matrisi

Korelasyon matrisi değişkenler arasındaki ilişkinin gücünü ve yönünü belirtir. Korelasyon matrisindeki değerler -1 ile 1 arasında değişir. Değerlerin pozitif olması iki değişkenin değerlerinin birlikte arttığını, negatif olması değişkenlerden birinin değeri artarken diğerinin azaldığını gösterir. Değerler 1'e yaklaştıkça iki değişken arasındaki pozitif ilişkinin miktarının arttığı yorumu yapılabilir. Değerlerin -1'e yaklaşması negatif ilişkinin miktarının artması şeklinde yorumlanabilir. Değerlerin 0 ya da 0'a yakın olması iki değişken arasında ilişki olmadığı şeklinde yorumlanmalıdır (Wikipedia, 2022a, 2022b).

Korelasyon matrisinden görüldüğü üzere  $S_s$  ve  $S_1$  arasında %97 korelasyon olduğu görülmektedir. Ayrıca  $m_t$  ile  $H_N$  arasında %73 korelasyon vardır. Buradan  $S_s$  ile  $S_1$  değerlerinin ve  $m_t$  ile  $H_N$  değerlerinin birlikte arttığı yorumu yapılmalıdır. Diğer bir söylemle veri seti içerisindeki öznelikler olan  $S_s$  ile  $S_1$  arasında ve  $m_t$  ile  $H_N$  arasında güçlü ve pozitif bir ilişki vardır.

### 3.1.3. Kategorik Verilerin Sayısal Dönüştürülmesi ve Normalizasyon

Verideki zemin sınıfı değerleri kategorik verilerdir. Kullanılan makine öğrenmesi algoritmaları ise sayısal veriler ile çalışmaktadır. Bu nedenle kategorik veriler

one hot encoding (yöntemin Türkçe karşılığı bulunamamıştır) ile 0 ve 1 değerlerine çevrilmiştir. Kategorik verilerin 0 ve 1 değerlerine çevrildiği haldeki veri setinin görünümü Tablo 4'tedir.

Tablo 4

Veri Setinden Kesit (One Hot Encoding Uygulandıktan Sonra)

ss	s1	mt	hN	ZA	ZB	ZC	ZD	ZE	vtE
0.575	0.135	1388	11.2	0	0	1	0	0	562.771
0.603	0.159	508	11.2	0	1	0	0	0	129.380
0.394	0.106	513	2.8	1	0	0	0	0	225.795
0.904	0.224	2155	15	0	0	0	1	0	1670.705
0.448	0.126	678	5.6	0	1	0	0	0	230.134

Standardizasyon bir veri setindeki tüm sayısal değerlerin ortalamasının 0, standart sapmasının 1 olacak şekilde dönüştürülmesi işlemidir. Normalizasyon ise tüm sayısal değerlerin belirli bir aralıktaki sayısal değerlere dönüştürülmesi işlemidir. En sık kullanılan normalizasyon yöntemleri tüm verilerin -1 ile 1 arasındaki sayılara dönüştürülmesi veya 0 ile 1 arasındaki sayılara dönüştürülmesidir (Ali ve Faraj, 2014). Makine öğrenmesi algoritmalarında standardizasyon ve normalizasyon çoğunlukla ön işlem adımı olarak kullanılmaktadır. Bunun nedeni bazı makine öğrenmesi algoritmalarının değişken ölçeklerde bulunan verilere karşı duyarlı olması ve bu işlemler uygulanmadığında iyi performans göstermemesidir.

Yapılan testler göstermiştir ki bu veri seti üzerinde yapılacak makine öğrenmesi yöntemlerinde standardizasyon ve [-1,1] normalizasyonu iyi sonuçlar vermemektedir. İlgili sonuçlar Tablo 5'te görülmektedir. Tabloda veri setindeki tüm öznelikler kullanılarak yapılan tahminlerde veriye uygulanan standardizasyon, [-1,1] normalizasyonu ve [0,1] normalizasyonuna göre elde edilen hataların değişimi görülmektedir. Tabloda belirtilen hata ölçümü için kök ortalama kare hatası (root mean square error) kullanılmıştır. Bu hata metriği konusundaki açıklamalar Bölüm 3.2.1'de yapılmıştır.

Tablo 5

Standardizasyon ve Normalizasyon Yöntemlerine Göre Tahmin Hatasının Değişimi

Yöntem	Standar-dizasyon	Normalizasyon [-1,1]	Normalizasyon [0,1]
Destek Vektör Makinesi	140.178	138.697	134.962
Rastgele Orman	21.474	21.470	21.476
Yapay Sinir Ağları	16.337	14.427	13.774

Tablo 5'te görüldüğü üzere rastgele orman yönteminin algoritmanın kendisine haz özelliklerinden kaynaklı olarak verilerin sayısal büyüklüğünden etkilenmemesi sebebi ile sonuçlarda neredeyse hiç değişiklik olmamıştır. Destek vektör makinesi ve yapay sinir ağları ile en iyi performans ise veriler 0 ile 1 arasına normalize edildiğinde alınmıştır. Bu nedenle çalışmanın devamında açıklanan yöntemlerde veriler 0 ile 1 arasına normalize edilmiştir. [0,1] normalizasyonu için veriye uygulanan fonksiyon Denklem (1)'de görülmektedir. Denklemde  $z_i$  herhangi bir öznitelikteki  $i$ 'nci sayısal değer normalize edilmiş hali,  $x_i$  aynı öznitelikteki  $i$ 'nci değeri,  $\min(x)$  ve  $\max(x)$  ise ilgili öznitelikteki en küçük ve en büyük değerlerdir.

$$z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (1)$$

### 3.1.4. Uç Değerlerin Tespiti ve Silinmesi

Eliptik sarmal (elliptic envelope) yöntemi ile uç değer tespiti gauss dağılımından yararlanılarak yapılan uç değer tespit yöntemidir (Azimi, Azhdary Moghaddam ve Hashemi Monfared, 2018). Yöntemde verinin uzaydaki şekli tanımlanmakta ve gauss dağılımına göre verinin dağılımı belirlenmektedir. Genel dağılım eliptik bir şekille sarmalanmaktadır. Sarmalın dışında kalan değerler uç değer kabul edilmektedir. Bu yöntem ile eğitim verisinden %2, %5 ve %10'luk kısımdaki uç değer veriler silinerek eğitilen makine öğrenmesi modellerinde (Bölüm 3.2'de ilgili yöntemler açıklanmıştır) görülmüştür ki: uç değerlerin silinmesi sonucunda test başarısı ciddi oranda düşmüştür. İlgili detaylar Tablo 6'da görülmektedir. Tablodaki yüzdeler eğitim verisinden uç değerlerin % kaçlık kısmının silindiğini göstermektedir. Tablodaki %0 yazan ifade hiçbir uç değer silinmediğini diğer söylemlerle eğitim verisinin aynen kullanıldığını göstermektedir. Tabloda belirtilen hata ölçümü için kök ortalama kare hatası (root mean square error) kullanılmıştır. Bu hata metriği konusundaki açıklamalar Bölüm 3.2.1'de yapılmıştır.

Uç değerlerin silinmesi sonucu test başarı oranlarının ciddi oranda düşmesinin nedeninin şu olduğu düşünülmektedir: Veri setindeki uç değerler aykırı veriler değildir. Bir diğer deyişle uç değerler hatalı, yanlış veya verinin yapısından büyük ölçüde farklılıklar gösteren veriler değildir. Bu nedenle çalışmada ele alınan makine öğrenmesi yöntemlerinin uç değerlerle daha iyi performans göstermesi muhtemeldir.

Eğitim verisi içerisindeki uç değerlerin silinmesinin test başarısını ciddi oranda düşürmesinin sonucunda bu başarısız işlem iptal edilmiş ve uç değerler silinmeden sonraki adıma geçilmiştir.

Tablo 6

Silinen Uç Değerlere Göre Tahmin Hatasının Değişimi

Yöntem	%0	%2	%5	%10
Destek Vektör Makinesi	134.962	606.109	618.327	630.628
Rastgele Orman	21.476	582.617	587.561	589.221
Yapay Sinir Ağları	13.774	573.135	572.598	573.495

### 3.1.5. Boyut Azaltma ile Özniteliklerin Birleştirilmesi

Bölüm 3.1.2'den görüldüğü üzere  $m_t$  ile  $H_N$  arasında %73 korelasyon vardır. Bu iki öznitelik temel bileşenler analizi (principal component analysis: PCA) ile varyansın %86'sı korunacak şekilde birleştirilerek yeni bir öznitelik oluşturulmuş ve veri setine eklenmiştir. Veri setinden  $m_t$  ve  $H_N$  öznitelikleri silinmiştir. Sonrasında makine öğrenmesi modelleri eğitilmiştir. Eğitilen modellerin test başarı oranları ciddi oranda düşmüştür. İlgili sayısal değerler Tablo 7'de görülmektedir. Test başarılarının düşmesinin temel nedenlerinin aşağıdaki maddeler olduğu düşünülmektedir:

- Temel bileşenler analizi iki özniteliği uzayda doğrusal olarak birleştirmektedir. Bu nedenle oluşturulan öznitelik, birleştirilen özniteliklerdeki tüm bilgiyi saklayamamaktadır. Varyansın bir kısmı kaybolmaktadır.
- Temel bileşenler analizi yöntemi, çıktı özniteliklerini hesaba katmayan bir algoritmadır. Sadece girdi öznitelikleri arasındaki ilişkiye göre uygulanan bir yöntemdir. Bu nedenle model performansını kötü etkileyen öznitelikler üretilebilmektedir.

Temel bileşenler analizi uygulandıktan sonra test başarı oranlarının düşmesi nedeniyle yapılan işlemler geri

alınmıştır. Birleştirilerek oluşturulan öznitelik silinmiş,  $m_t$  ve  $H_N$  öznitelikleri veri setine tekrar eklenmiştir.

Tablo 7

Temel Bileşenler Analizi ile Model Performansı

Yöntem	PCA	
	Kullanıldığında	Kullanılmadığında
Destek Vektör Makinesi	299.857	134.962
Rastgele Orman	225.220	21.476
Yapay Sinir Ağları	266.693	13.774

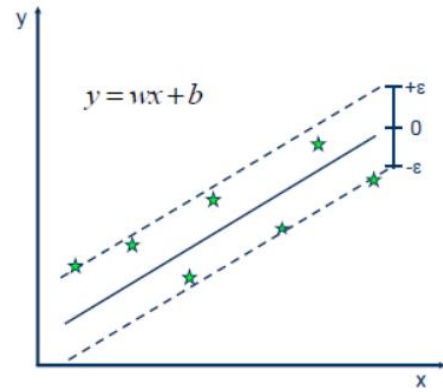
### 3.2. Öznitelik Seçimi (Feature Selection)

Öznitelik seçimi için bir kaba kuvvet yöntemi olan etraflı arama (exhaustive search) kullanılmıştır. Bu yöntemde amaç giriş özniteliklerinin olası tüm kombinasyonlarını denemek ve en iyi sonucu ya da sonuçları veren öznitelik alt kümesini belirlemektir. Bu çalışmada 5 girdi özniteliği olduğu için  $2^5-1=31$  (formüldeki -1, hiçbir özniteliğin seçilmediği durumdan gelmektedir) tane makine öğrenmesi modeli eğitilerek öznitelik seçimi yapılmıştır. Öznitelik seçimi için kullanılan makine öğrenmesi yöntemleri destek vektör makinesi, rastgele orman ve yapay sinir ağlarıdır.

#### 3.2.1. Destek Vektör Makinesi ile Öznitelik Seçimi

Destek vektör makinesi (DVM) çoğunlukla sınıflandırma için kullanılan bir algoritma olsa da regresyon için de kullanılmaktadır. Regresyon ve sınıflandırma için algoritmanın çalışma biçiminde ufak farklılıklar olsa da temel amaç aynıdır. Amaç, hatayı en aza indiren marjin çizgisinin bulunmasıdır. İlgili çizim Şekil 2'de görülmektedir.

Bu çalışmada DVM eğitimi için RBF çekirdek (kernel) kullanılmış ve tüm hiperparametreler optimize edilmiştir.



Şekil 2. Destek Vektör Makinesi Regresyonu (Sayad, t.y.)

Özniteliklerin seçilen tüm kombinasyonu ve ilgili kombinasyona göre elde edilen başarı oranları çalışmanın okunaklı olması için verilmemiştir. Bunun yerine her bir öznitelik sayısı ile en az hata oranının elde edildiği denemeler ve ilgili denemede hangi özniteliğin ya da özniteliklerin kullanıldığı gösterilmiştir. Örneğin 1 özniteliğin girdi olarak kullanıldığı 5 model eğitimi ve testi vardır (Sadece  $S_s$ , sadece  $S_1$ , sadece  $m_t$ , sadece  $H_N$ , sadece zemin sınıfının girdi olarak kullanıldığı durum). Bu modellerden hata oranı en az olanda zemin sınıfı girdi olarak kullanıldı ise sadece buna ait bilgiler verilmiş; Tablo 8, 9 ve 10'da gösterilmiştir.

DVM ile elde edilen sonuçlar Tablo 8'de görülmektedir. Tabloda "1" yazan hücreler ilgili özniteliğin girdi olarak kullanıldığını "0" ise kullanılmadığını göstermektedir. Tablodaki RMSE, Root Mean Square Error'un (kök ortalama kare hatası) kısaltmasıdır. RMSE farklı bilim dallarında veri madenciliği, makine öğrenmesi, derin öğrenme gibi uygulamalarda sayısal verilerin tahminlenme başarısını ölçme amaçlı olarak sıklıkla kullanılan bir metriktir (Bartholomai ve Frieboes, 2018; Bostan, Heuvelink ve Akyurek, 2012; Cao, Kim, Hwang ve Jung, 2018; Chniti, Bakir ve Zaher, 2017; Parbat ve Chakraborty, 2020; Pham, Son, Hoang, Nguyen ve Tien Bui, 2018; Rossel ve Behrens, 2010) Bu nedenle bu çalışmada RMSE metriği ile sonuçlar yorumlanmıştır. RMSE'nin düşük olması ile yapılan tahminlerin iyi olduğu, yüksek olması ile tahminlerin kötü olduğu anlaşılmalıdır. Tablo 8'de  $R^2$  ( $R^2$ ) determinasyon katsayısıdır (Coefficient of Determination). Determinasyon katsayısı gerçek sayısal değer ile tahmin edilen sayısal değer arasındaki yaklaşıklıkla bağlı bir metriktir. Determinasyon katsayısının 1'e yaklaşması ile yapılan tahminin başarılı olduğu, 0'a yaklaşması ile yapılan tahminin başarısız olduğu anlaşılmalıdır.

Tablo 8

## DVM ile Öznitelik Seçimi

Öznitelik Sayısı	S <sub>s</sub>	S <sub>1</sub>	m <sub>t</sub>	H <sub>N</sub>	Zemin Sınıfı	RMSE	R2
1	0	0	0	0	1	536.259	0.289
2	0	0	1	0	1	356.462	0.686
3	0	1	1	0	1	190.242	0.911
4	0	1	1	1	1	125.015	0.961
5	1	1	1	1	1	134.962	0.955

Tablo 8'den görüldüğü üzere 4 özniteliğin girdi olarak kullanılması ile 5 özniteliğe göre daha iyi sonuçlar elde edilmiştir. Bu durum bu algoritma ile bu veri setinde H<sub>N</sub> özniteliğinin kullanılmasının gereksiz olduğunu hatta kullanılmaması gerektiğini göstermektedir. Buna ek olarak S<sub>s</sub> özniteliğinin de kullanılmaması ile iyi başarı oranları elde edilmiştir. S<sub>1</sub>, m<sub>t</sub>, zemin sınıfı özniteliklerinin kullanılması ile gerçek dünyaya uygulanabilir olduğu düşünülen sonuçlar elde edilebilmektedir.

## 3.2.2. Rastgele Orman ile Öznitelik Seçimi

Rastgele orman çok sayıda karar ağacının doğru şekilde birleştirilmesi ile oluşan bir algoritmadır. Hem regresyon hem de sınıflandırma için kullanılabilir. Bu çalışma kapsamında rastgele orman ile xgboost (Chen ve Guestrin, 2016) kullanılmıştır. Xgboost, gradyan destekli olarak hata fonksiyonunu minimize etmeye yarayan bir tekniktir. Bu çalışmada kullanılan rastgele orman 500 karar ağacından oluşmaktadır.

Rastgele orman kullanılarak elde edilen sonuçlar Tablo 9'da görülmektedir. Görüldüğü üzere 4 öznitelik ile eğitilen modellerin test sonuçları 5 öznitelik ile eğitilenlere çok yakındır. Bu sonuçlar bu algoritma ile H<sub>N</sub> özniteliğinin kullanılmasının (gerçek değerlere çok yakın sonuçlar gerekmedikçe) gereksiz olduğunu göstermektedir. Buna ek olarak S<sub>s</sub> özniteliğinin de kullanılmaması ile iyi başarı oranları elde edilmiştir. S<sub>1</sub>, m<sub>t</sub>, zemin sınıfı özniteliklerinin kullanılması ile gerçek dünyaya uygulanabilir sonuçların elde edilebildiği düşünülmektedir.

Tablo 9

## Rastgele Orman ile Öznitelik Seçimi

Öznitelik Sayısı	S <sub>s</sub>	S <sub>1</sub>	m <sub>t</sub>	H <sub>N</sub>	Zemin Sınıfı	RMSE	R2
1	0	0	0	0	1	520.125	0.331
2	0	0	1	0	1	344.35	0.707
3	0	1	1	0	1	133.848	0.956
4	0	1	1	1	1	27.290	0.998
5	1	1	1	1	1	21.476	0.999

## 3.2.3. Yapay Sinir Ağları ile Öznitelik Seçimi

Yapay sinir ağları insanın sinirsel yapısından esinlenilerek oluşturulmuş yapılardır. ax+b şeklindeki doğrusal fonksiyonların aktivasyon fonksiyonları yardımı ile değiştirilmesi sonucu optimum çözüme yakınsamaya çalışırlar. Günümüzde yapay sinir ağları eğitimi için çoğunlukla geri besleme algoritması kullanılır. Geri besleme algoritması türevden yararlanılarak hatanın azaltılması ve ağ ağırlıklarının optimize edilmesini sağlamaktadır.

Bu çalışmada kullanılan yapay sinir ağında 1 gizli katman vardır. Gizli katmanda 100 nöron bulunmaktadır. Ağda ReLu aktivasyon fonksiyonu ve Adam optimizasyon algoritması kullanılmıştır.

Yapay sinir ağları kullanılarak elde edilen sonuçlar Tablo 10'da görülmektedir. Tablodan görüldüğü üzere 4 öznitelik ile eğitilen modellerin test sonuçları 5 öznitelik ile eğitilenlere çok yakındır. Bu sonuçlar bu algoritma ile H<sub>N</sub> özniteliğinin kullanılmasının (gerçek değerlere çok yakın sonuçlar gerekmedikçe) gereksiz olduğunu göstermektedir. Buna ek olarak S<sub>s</sub> özniteliğinin de kullanılmaması ile iyi başarı oranları elde edilmiştir. S<sub>1</sub>, m<sub>t</sub>, zemin sınıfı özniteliklerinin kullanılması ile gerçek dünyaya uygulanabilir sonuçların elde edilebildiği düşünülmektedir.

Tablo 10

## Yapay Sinir Ağları ile Öznitelik Seçimi

Öznitelik Sayısı	S <sub>s</sub>	S <sub>1</sub>	m <sub>t</sub>	H <sub>N</sub>	Zemin Sınıfı	RMSE	R2
1	0	0	0	0	1	520.141	0.331
2	0	0	1	0	1	344.987	0.706
3	0	1	1	0	1	144.935	0.948
4	0	1	1	1	1	23.391	0.998
5	1	1	1	1	1	13.774	0.999

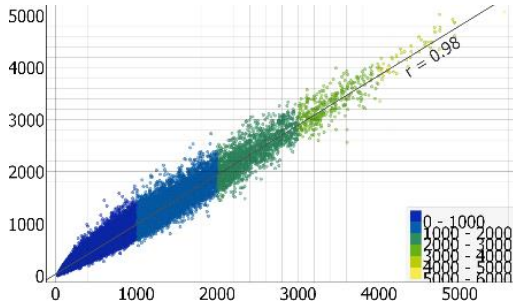
## 3.3. Seçilen Özniteliklerin Değerlendirilmesi

DVM ile yapılan öznitelik seçiminde 4 öznitelik ile 5 özniteliğe göre daha iyi sonuçlar alınmıştır. H<sub>N</sub> özniteliğinin sonuçları kötü etkilediği görülmüştür. Rastgele orman ve yapay sinir ağı ile yapılan uygulamalarda tüm özniteliklerin kullanıldığı denemelerde diğer denemelere göre daha iyi sonuçlar alınmıştır. Fakat 4 öznitelik ile alınan sonuçlar tüm özniteliklerin kullanıldığında alınan sonuçlara çok yakındır. Bu nedenle bu algoritmalarda da H<sub>N</sub> özniteliğinin kullanılmamasında bir yanlışlık olmadığı söylenebilir.

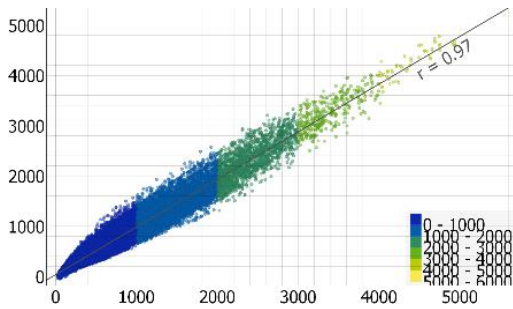
Ayrıca H<sub>N</sub> ve S<sub>s</sub> özniteliklerinin ikisinin de kullanılmadığı uygulamalarda gerçek dünyaya uygulanabilir sonuçların alındığı düşünülmektedir. Bu nedenle gerçek değerlerle bire bir aynı sayılabilecek veya çok yakın tahminlerin elde edilmesi gerekmedikçe

bu çalışmada ele alınan makine öğrenmesi metotları ile  $H_N$  ve  $S_S$  özniteliklerinin kullanılmasına gerek olmadığı yorumu yapılabilir.

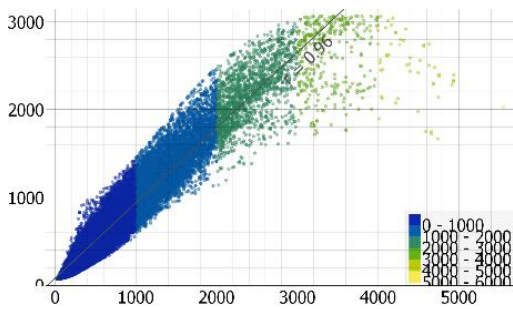
Şekil 3, 4 ve 5  $S_1$ ,  $m_t$  ve zemin sınıfı özniteliklerinin kullanıldığı durumlardaki makine öğrenmesi modellerinin tahmin başarısını göstermektedir. Şekillerde yatay eksenler gerçek deprem kuvveti değerlerini, dikey eksenler model tahminlerini göstermektedir.



Şekil 3. Rastgele Orman Modelinin Tahmin Başarısı



Şekil 4. Yapay Sinir Ağı Modelinin Tahmin Başarısı



Şekil 5. DVM Modelinin Tahmin Başarısı

#### 4. Sonuç

Yapılan literatür taraması sonucunda veri madenciliği, veri bilimi, makine öğrenmesi ve deprem konularında birçok çalışma olduğu görülmüştür. Fakat bu çalışmalarda Türkiye Bina Deprem Yönetmeliği'nin tasarım deprem kuvveti hesabı için istediği girdilerden kaçına ve hangilerine gerek kalmaksızın tasarım deprem kuvvetlerinin elde edilebileceğine veya yaklaşık olarak hesap tahmini yapılabileceğine dair hususlar ele alınmamıştır. Bu nedenle bu çalışma yönetmelik

koşullarının gerekliliğini veri madenciliği, veri bilimi ve makine öğrenmesi perspektifinde tartışmaya açması bakımından literatüre yenilik getirmektedir.

Yapılan farklı denemeler sonucunda makine öğrenmesi tekniklerinin perdesiz, betonarme konut binalarına gelen deprem kuvvetinin tahminlenmesi için kullanışlı araç kutuları olduğu sonucuna varılmıştır. Ortaya çıkan diğer sonuçlar şunlardır:

- Türkiye Bina Deprem Yönetmeliği'nin tasarım deprem kuvvetlerinin hesabı için şart koştuğu  $S_S$  ve  $H_N$  öznitelikleri kullanılmadan tasarım deprem kuvvetleri destek vektör makinesi yöntemi ile  $RMSE = 190.242$  ve  $R^2 = 0.911$ , rastgele orman ile  $RMSE = 133.848$  ve  $R^2 = 0.956$ , yapay sinir ağı ile  $RMSE = 144.935$  ve  $R^2 = 0.948$  olacak şekilde başarıyla tahminlenebilmiştir.
- Bu çalışmada kullanılan eğitim veri seti üzerinde uç değerlerin silinmesi ile test başarıları düşmüştür. Diğer bir söylemle uç değer verilerin model eğitimi zorlaştırmadığı hatta iyi yönde etkilediği sonucuna varılabilir.
- Temel bileşenler analizi (PCA) tekniği ile veri setindeki  $m_t$  ve  $H_N$  özniteliklerinin birleştirilerek yeni öznitelik oluşturulduğu denemelerde  $m_t$  ve  $H_N$  veri setinden silinmiştir. Fakat bu durumda çalışmada ele alınan makine öğrenmesi modellerinin test başarısında ciddi düşüş olmuştur.

Tüm bilim dallarında geçerli olmak üzere; bu zamana kadar yapılan kabuller gözden geçirilmek istendiğinde veya yeni teoriler ortaya konmaya çalışıldığında veri madenciliği yöntemlerinden yararlanılabilir. Veri madenciliği veri içerisindeki saklı bilgileri ortaya çıkarmanın güçlü ve kullanışlı bir yoludur.

#### 5. Gelecekteki Çalışmalar

Bu çalışmada perdesiz betonarme konut binaları özelinde çalışılmıştır. Bunun nedeni yazarlar tarafından Türkiye Bina Deprem Yönetmeliği'ndeki Tablo 3.1 ve Tablo 4.1'deki tüm maddeleri kapsayacak veri setini oluşturmak için gereken kodu yazmanın çok uzun süreceğinin öngörülmesidir. Dilerse başka araştırmacılar bu çalışmanın kapsamını ilgili tablolardaki diğer maddeleri kapsayacak şekilde genişletebilirler.

#### Araştırmacıların Katkısı

Çalışmada Tarkan KARAÇAY veri setinin oluşturulması, literatür taramasının yapılması, sonuçların elde edilmesi ve değerlendirilmesi konularında katkı sağlamıştır. Tolga AÇIKGÖZ makalenin genel kontrollerinin ve düzenlemelerinin yapılması,



çalışmanın gidişatının yönetilmesi ve sonuçların değerlendirilmesi konularında katkı sağlamıştır.

### Çıkar Çatışması

Yazarlar tarafından herhangi bir çıkar çatışması beyan edilmemiştir.

### Kaynaklar

- Akhgar, A., Toghraie, D., Sina, N. ve Afrand, M. (2019). Developing dissimilar artificial neural networks (ANNs) to prediction the thermal conductivity of MWCNT-TiO<sub>2</sub>/Water-ethylene glycol hybrid nanofluid. *Powder Technology*, 355, 602–610. doi: <https://doi.org/10.1016/j.powtec.2019.07.086>
- Aksoylu, C. ve Arslan, M. H. (2019). Çerçeve+Perde Türü Betonarme Binaların Periyod Hesaplarının Tbdy-2019 Yönetmeliğine Göre Ampirik Olarak Değerlendirilmesi. *Uludağ Üniversitesi Mühendislik Fakültesi Dergisi*, 24(3), 365–382. doi: <https://doi.org/10.17482/uumfd.603437>
- Alakus, T. B. ve Turkoglu, I. (2020). Comparison of deep learning approaches to predict COVID-19 infection. *Chaos, Solitons & Fractals*, 140. doi: <https://doi.org/10.1016/j.chaos.2020.110120>
- Ali, P. J. M. ve Faraj, R. H. (2014). Data Normalization and Standardization: A Technical Report. In *Machine Learning Technical Reports* (Vol. 1). doi: <https://doi.org/10.13140/RG.2.2.28948.04489>
- Azimi, S., Azhdary Moghaddam, M. ve Hashemi Monfared, S. A. (2018). Anomaly Detection and Reliability Analysis of Groundwater by Crude Monte Carlo and Importance Sampling Approaches. *Water Resources Management*, 32(14), 4447–4467. doi: <https://doi.org/10.1007/s11269-018-2029-y>
- Bartholomai, J. A. ve Frieboes, H. B. (2018). Lung Cancer Survival Prediction via Machine Learning Regression, Classification, and Statistical Techniques. *2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, 632–637. IEEE. doi: <https://doi.org/10.1109/ISSPIT.2018.8642753>
- Bostan, P. A., Heuvelink, G. B. M. ve Akyurek, S. Z. (2012). Comparison of regression and kriging techniques for mapping the average annual precipitation of Turkey. *International Journal of Applied Earth Observation and Geoinformation*, 19, 115–126. doi: <https://doi.org/10.1016/j.jag.2012.04.010>
- Bozer, A. (2020). Tasarım Spektral İvme Katsayılarının DBYBHY 2007 ve TBDY 2018 Yönetmeliklerine Göre Karşılaştırması. *DÜMF Mühendislik Dergisi*, 11(1), 393–404. doi: <https://doi.org/10.24012/dumf.559965>
- Cao, K., Kim, H., Hwang, C. ve Jung, H. (2018). CNN-LSTM Coupled Model for Prediction of Waterworks Operation Data. *Journal of Information Processing Systems*, 14(6), 1508–1520. doi: <https://doi.org/10.3745/JIPS.02.0104>
- Cetinkaya, I. H. (2021). World saw 13,654 quakes of magnitude 4 or above in 2020. Erişim Adresi (14 Şubat 2022): <https://www.aa.com.tr/en/environment/world-saw-13-654-quakes-of-magnitude-4-or-above-in-2020/2161305>
- Chen, T. ve Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. doi: <https://doi.org/10.1145/2939672.2939785>
- Chniti, G., Bakir, H. ve Zaher, H. (2017). E-commerce Time Series Forecasting using LSTM Neural Network and Support Vector Regression. *Proceedings of the International Conference on Big Data and Internet of Thing - BDIOT2017*, 80–84. New York: ACM Press. doi: <https://doi.org/10.1145/3175684.3175695>
- Contreras, D., Wilkinson, S. ve James, P. (2021). Earthquake Reconnaissance Data Sources, a Literature Review. *Earth*, 2(4), 1006–1037. doi: <https://doi.org/10.3390/earth2040060>
- Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevár, T., Milutinovič, M., ... Zupan, B. (2013). Orange: Data Mining Toolbox in Python. *Journal of Machine Learning Research*, 14, 2349–2353.
- Falcone, R., Lima, C. ve Martinelli, E. (2020). Soft computing techniques in structural and earthquake engineering: a literature review. *Engineering Structures*, 207(November 2019), 110269. doi: <https://doi.org/10.1016/j.engstruct.2020.110269>
- Geiß, C. ve Taubenböck, H. (2013). Remote sensing contributing to assess earthquake risk: from a literature review towards a roadmap. *Natural Hazards*, 68(1), 7–48. doi: <https://doi.org/10.1007/s11069-012-0322-2>
- Goswami, S., Chakraborty, S., Ghosh, S., Chakrabarti, A. ve Chakraborty, B. (2018). A review on application of data mining techniques to combat natural disasters. *Ain Shams Engineering Journal*, 9(3), 365–378. doi: <https://doi.org/10.1016/j.asej.2016.01.012>
- Hall, M., Frank, E., Holmes, G., Pfahringer, B.,

- Reutemann, P. ve Witten, I. H. (2009). The WEKA Data Mining Software: An Update. *ACM SIGKDD Explorations Newsletter*, 11(1), 10–18. doi: <https://doi.org/10.1145/1656274.1656278>
- Kadam, P. ve Vijayumar, S. (2018). Prediction Model: CO2 Emission Using Machine Learning. *2018 3rd International Conference for Convergence in Technology (I2CT)*. IEEE. doi: <https://doi.org/10.1109/I2CT.2018.8529498>
- Kahandawa, K. A. R. V. D., Domingo, N. D., Park, K. S. ve Uma, S. R. (2018). Earthquake damage estimation systems: Literature review. *Procedia Engineering*, 212, 622–628. doi: <https://doi.org/10.1016/j.proeng.2018.01.080>
- Karaçay, T. (2022a). TBDY 2018 - Deprem Kuvveti (Konut Binaları). Erişim Adresi (14 Şubat 2022): [https://www.kaggle.com/tarkankaraay/tbdy-2018-deprem-kuvveti-konut-binalari?select=Veri\\_Seti\\_1.xlsx](https://www.kaggle.com/tarkankaraay/tbdy-2018-deprem-kuvveti-konut-binalari?select=Veri_Seti_1.xlsx)
- Karaçay, T. (2022b). TBDY 2018 - Deprem Kuvveti (Konut Binaları). Erişim Adresi (14 Şubat 2022): [https://www.kaggle.com/tarkankaraay/tbdy-2018-deprem-kuvveti-konut-binalari?select=AFAD\\_ss\\_s1.xlsx](https://www.kaggle.com/tarkankaraay/tbdy-2018-deprem-kuvveti-konut-binalari?select=AFAD_ss_s1.xlsx)
- Mia, M. ve Dhar, N. R. (2016). Response surface and neural network based predictive models of cutting temperature in hard turning. *Journal of Advanced Research*, 7(6), 1035–1044. doi: <https://doi.org/10.1016/j.jare.2016.05.004>
- Mothilal, R. K., Sharma, A. ve Tan, C. (2020). Explaining machine learning classifiers through diverse counterfactual explanations. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 607–617. Barcelona. doi: <https://doi.org/10.1145/3351095.3372850>
- Ortega, J., Vasconcelos, G., Rodrigues, H., Correia, M. ve Lourenço, P. B. (2017). Traditional earthquake resistant techniques for vernacular architecture and local seismic cultures: A literature review. *Journal of Cultural Heritage*, 27, 181–196. doi: <https://doi.org/10.1016/j.culher.2017.02.015>
- Otari, G. V. ve Kulkarni, D. R. V. (2012). A Review of Application of Data Mining in Earthquake Prediction. *International Journal of Computer Science and Information Technologies*, 3(2), 3570–3574. Erişim Adresi: <http://www.ijcsit.com/docs/Volume3/Vol3Issue2/ijcsit2012030258.pdf>
- Parbat, D. ve Chakraborty, M. (2020). A python based support vector regression model for prediction of COVID19 cases in India. *Chaos, Solitons & Fractals*, 138. doi: <https://doi.org/10.1016/j.chaos.2020.109942>
- Pham, B. T., Son, L. H., Hoang, T.-A., Nguyen, D.-M. ve Tien Bui, D. (2018). Prediction of shear strength of soft soil using machine learning methods. *CATENA*, 166, 181–191. doi: <https://doi.org/10.1016/j.catena.2018.04.004>
- Planlı Alanlar Yönetmeliği (2017, 03 Temmuz). Resmi Gazete (Sayı: 30113). Erişim Adresi: <https://www.resmigazete.gov.tr/eskiler/2017/07/20170703-8.htm>
- Rossel, R. A. V. ve Behrens, T. (2010). Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, 158(1–2), 46–54. doi: <https://doi.org/10.1016/j.geoderma.2009.12.025>
- Sayad, S. (b.t.). Support Vector Machine - Regression (SVR). Erişim Adresi (20 Ocak 2022): [https://www.saedsayad.com/support\\_vector\\_machine\\_reg.htm#:~:text=Support Vector Machine can also,the algorithm \(maximal margin\).&text=In the case of regression,already requested from the problem.](https://www.saedsayad.com/support_vector_machine_reg.htm#:~:text=Support Vector Machine can also,the algorithm (maximal margin).&text=In the case of regression,already requested from the problem.)
- Türkiye Bina Deprem Yönetmeliği (2018, 18 Mart). Resmi Gazete (Sayı: 30364). Erişim Adresi: <https://www.resmigazete.gov.tr/eskiler/2018/03/20180318M1-2.htm>
- Türkoğlu, N. (2001). Türkiye'nin Yüzölçümü ve Nüfusunun Deprem Bölgelerine Dağılışı. In *Türkiye Coğrafyası Araştırma ve Uygulama Merkezi Dergisi*.
- Wei, W., Ramalho, O., Malingre, L., Sivanantham, S., Little, J. C. ve Mandin, C. (2019). Machine learning and statistical models for predicting indoor air quality. *Indoor Air*, 29(5), 704–726. doi: <https://doi.org/10.1111/ina.12580>
- Wikipedia. (2022a). Pearson correlation coefficient. Erişim Adresi (17 Nisan 2022): [https://en.wikipedia.org/wiki/Pearson\\_correlation\\_coefficient](https://en.wikipedia.org/wiki/Pearson_correlation_coefficient)
- Wikipedia. (2022b). Spearman's rank correlation coefficient. Erişim Adresi (17 Nisan 2022): [https://en.wikipedia.org/wiki/Spearman%27s\\_rank\\_correlation\\_coefficient](https://en.wikipedia.org/wiki/Spearman%27s_rank_correlation_coefficient)
- Yuan, S., Jiao, Z., Quddus, N., Kwon, J. S.-I. ve Mashuga, C. V. (2019). Developing Quantitative Structure–Property Relationship Models To Predict the Upper Flammability Limit Using Machine Learning. *Industrial & Engineering Chemistry Research*, 58(8), 3531–3537. doi: <https://doi.org/10.1021/acs.iecr.8b05938>