

**ZAMAN SERİSİ YÖNTEMLERİ İLE
DEPREM VERİLERİNİN ANALİZİ**

**EARTHQUAKE DATA ANALYSIS
WITH TIME SERIES METHODS**

KÜBRA SİMAY GÖKMENOĞLU

PROF. DR. CEM KADILAR

Tez Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

İstatistik Anabilim Dalı için Öngördüğü

YÜKSEK LİSANS TEZİ olarak hazırlanmıştır.

Bu tezi; hayatım boyunca sevgisiyle, desteęiyle ve dualarıyla beni her zaman sarıp sarmalayan kıymetlim, anneannem, Ayşe Acar'a ve fiziken yanımda olmasa bile hayatımın her anında varlığını hissettiğim, güçlü duruşu, öğütleri ve sevgisiyle bana yön veren her şeyim, canım, dedem, Yüksel Acar'a, minnet, sevgi ve derin bir özlem ile ithaf ediyorum...

ÖZET

ZAMAN SERİSİ YÖNTEMLERİ İLE DEPREM VERİLERİNİN ANALİZİ

Kübra Simay GÖKMENOĞLU

Yüksek Lisans Tezi

Tez Danışmanı: Prof. Dr. Cem KADILAR

HACETTEPE ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

HAZİRAN 2025

Türkiye, coğrafi konumu itibarıyla birçok aktif fay hattı üzerinde yer almakta ve bu nedenle sık sık çeşitli büyüklüklerde depremlerle karşı karşıya kalmaktadır. Bu depremler hem ekonomik hem de psikolojik açıdan insan yaşamında geri dönülmesi zor etkiler yaratmaktadır.

Bu çalışmada, deprem verilerinin analizi için öncelikle klasik regresyon modelleri kullanılmış, ardından daha gelişmiş yöntemler olan kantil regresyon, makine öğrenmesi tabanlı XGBoost regresyon modeli gibi modeller uygulanmıştır. İstatistiksel olarak, en uygun yöntemin LSTM Yöntemi olduğu sonucuna ulaşılmıştır. Analizler kapsamında deprem büyüklüğü (M_w), derinlik ve episantr koordinatları (enlem-boylam) dikkate alınarak modelleme yapılmıştır.

Çalışmanın uygulama kısmında, özellikle Marmara Bölgesi ve İstanbul ili odaklı analizler gerçekleştirilmiş; İstanbul'un 39 ilçesi için çeşitli derinlik seviyelerinde (10 ve 30 km) deprem büyüklüğü tahminleri üretilmiştir. Elde edilen tahminler görselleştirilerek hem çizgi grafikler

hem de ısı haritaları aracılığıyla sunulmuş ve bölgesel risk dağılımları görsel olarak karşılaştırılmıştır.

Ayrıca, derinlik değişkeninin deprem büyüklüğü üzerindeki etkisi istatistiksel olarak incelenmiş; bu kapsamda doğrusal regresyon analizi uygulanmıştır. Regresyon analizi sonucunda, derinlik ile Mw arasında anlamlı bir ilişki olduğu tespit edilmiş ve derinliğin artmasının tahmini Mw değerini istatistiksel olarak artırdığı sonucuna varılmıştır.

Tüm bu yöntemler sonucunda, depremlerin belirli istatistiksel yapı ve mekansal düzenlilikler gösterdiği ortaya konmuş ve elde edilen modellerle gelecekte yaşanabilecek olası depremlere dair tahminler yapılmıştır. Bu tez çalışması, deprem büyüklüklerinin tahmini için uygun modellerin oluşturulabileceği ve bu modellerden güvenilir bilimsel tahminlerin yapılacağını ortaya koymuştur.

Anahtar Kelimeler: Deprem büyüklüğü tahmini, zaman serisi analizi, regresyon analizi, kantil regresyon, XGBoost, LSTM, makine öğrenmesi, ısı haritası, derinlik analizi, istatistiksel modelleme

ABSTRACT

EARTHQUAKE DATA ANALYSIS WITH TIME SERIES METHODS

Kübra Simay GÖKMENOĞLU

(M. Sc. Thesis)

Supervisor: Prof. Dr. Cem KADILAR

HACETTEPE UNIVERSITY

INSTITUTE OF SCIENCE

JUNE 2025

Turkey is situated on numerous active fault lines due to its unique geographical location, and as a result, it frequently experiences earthquakes of varying magnitudes. These earthquakes have profound and often irreversible effects on human life, both economically and psychologically.

To analyze the earthquake data, the study initially employed classical regression models, followed by the application of more advanced techniques such as quantile regression and the machine learning-based XGBoost regression model. From a statistical standpoint, the LSTM method was determined to be the most suitable among the methods applied. The modeling process considered earthquake magnitude (M_w), depth, and epicenter coordinates (latitude and longitude) as key variables.

In the application phase of the study, analyses were particularly focused on the Marmara Region and the province of Istanbul. For Istanbul's 39 districts, earthquake magnitude predictions were

generated at varying depth levels (10 and 30 km). The resulting predictions were visualized using both line graphs and heat maps, enabling a comparative assessment of regional risk distributions.

Furthermore, the impact of depth on earthquake magnitude was examined statistically through linear regression analysis. The results indicated a statistically significant relationship between depth and M_w , concluding that an increase in depth tends to lead to higher predicted M_w values.

Overall, the findings suggest that earthquakes exhibit certain statistical structures and spatial patterns. The developed models provide scientifically reliable forecasts for potential future earthquakes. This thesis demonstrates that it is possible to construct appropriate models for predicting earthquake magnitudes and that these models can yield robust and credible scientific estimations.

Keywords: Earthquake magnitude prediction, time series analysis, regression analysis, quantile regression, XGBoost, LSTM, machine learning, heatmap, depth analysis, statistical analysis

TEŐEKKÜR

Tezimin hazırlanmasında bilgi, deneyim ve desteęiyle bana yol gösteren, alıőmam boyunca her zaman yapıcı yaklaőımıyla katkı saęlayan deęerli danıőman hocam, sayın Prof. Dr. Cem Kadılar'a,

İhtiyacım olduęunda yardımlarını esirgemeyen sayın hocam Prof. Dr. Gamze Özel Kadılar'a, Bazı verilerinden yararlandıęım, sayın Prof. Dr. Gamze Özel Kadılar yürütücülüęünde gerekleőtirilen 121F208 No'lu "Yapay Zeka ve Olasılıksal Model Tabanlı Deprem Tehlike Haritası" baőlıklı 1001 Projesi ile destek saęlayan TÜBİTAK'a,

Desteklerini benden esirgemeyen canım arkadaşlarıma,

Hayatım boyunca her daim yanımda olan, sevgileriyle beni sarıp sarmalayan, her zorluęumda sabırla yanımda duran, her baőarımda en az benim kadar sevinen canım aileme; özellikle bu tez sürecinde moralimi yüksek tutmamı saęlayan destekleri ve bana duydukları inan sayesinde beni bu günlere getirdikleri, varlıklarıyla bana güç verdikleri, her zaman arkamda durdukları ve koőulsuz sevgileriyle bu süreci daha anlamlı kıldıkları için özellikle annem Gülin Acar'a ve bütün aileme sonsuz teőekkürlerimi sunarım.

İÇİNDEKİLER

ÖZET	i
ABSTRACT	iii
İÇİNDEKİLER.....	vi
ŞEKİLLER DİZİNİ.....	viii
ÇİZELGELER DİZİNİ.....	x
SİMGELER VE KISALTMALAR.....	xi
1. GİRİŞ.....	1
1.1. Deprem Nedir?.....	3
1.2. Deprem Oluş Nedenleri Nelerdir?	4
1.3. Deprem Türleri.....	6
1.3.1. Tektonik Depremler	6
1.3.2. Volkanik Depremler	6
1.3.3. Çöküntü Depremler	6
1.3.4. Deniz Depremleri (Tsunami)	6
1.4. Deprem Fay Türleri	6
1.5. Deprem Parametreleri	7
1.5.1. Odak Noktası (Hiposantr).....	7
1.5.2. Dış Merkez (Episantr)	7
1.5.3. Odak Derinliği.....	7
1.5.4. Eş Şiddet (İzoseit) Eğrileri	8
1.5.5. Şiddet.....	8
1.5.6. Magnitüd.....	8
1.6. Guttenberg-Richer Büyüklük-Frekans Bağlantısı	10
1.7. Türkiye’de Depremler	12
1.8. Deprem Kronolojisi	14
1.9. Deprem Sonuçları	15

2. YÖNTEMLER	17
2.1. Regresyon Analizi.....	17
2.1.1.Literatürde Regresyon Analizinin Önerilişi	18
2.1.2. Regresyon Analizinde Temel Varsayımlar	18
2.1.3. En Küçük Kareler (EKK) Yöntemi	19
2.2. Kantil Regresyon Analizi	22
2.2.1 Kantil Regresyonun Özellikleri	26
2.2.2. Kantil Regresyonun Temel Denklemi	27
2.2.3. Kantil Regresyonun Varsayımları	27
2.3. Makine Öğrenmesi Yöntemleri.....	29
2.3.1 Makine Öğrenimi Türleri	32
2.3.2. Makine Öğrenimli Yazılımlar	34
2.4. Zaman Serileri Analizi	35
2.4.1. ARIMA Modeli	38
2.4.2. LSTM (Long - Short Term Memory, Uzun Kısa Süreli Bellek)	38
2.4.3. Destek Vektör Makinesi (Support Vector Machine).....	41
2.4.4. Karar Ağaçları (Decision Tree).....	42
2.4.5. Rastgele Orman (Random Forest)	43
2.4.6. Yapay Sinir Ağları (Artificial Neural Networks-ANN-YSA).....	44
2.4.7. K- En Yakın Komşu (KNN)	46
2.4.8. XGBoost	47
3. UYGULAMA	48
3.1 İstanbul İlçeleri İçin Uygulama.....	61
4. SONUÇ.....	74
KAYNAKLAR.....	77
ÖZGEÇMİŞ.....	86

ŞEKİLLER DİZİNİ

Şekil 1	Dünya'nın Katmanları	3
Şekil 2	Dünya Deprem Haritası	5
Şekil 3	Türkiye'deki Levhalar ve Fayların Haritası	13
Şekil 4	Türkiye Deprem Kuşakları Haritası	15
Şekil 5	Kantil Regresyon Eğrisi.....	25
Şekil 6	Makine Öğrenimi Şeması	30
Şekil 7	Makine Öğreniminin Çalışma Prensipleri	31
Şekil 8	Makine Öğrenimi Grupları	33
Şekil 9	RNN Sinir Yapısına Ait En Küçük Birim (Hücre)	39
Şekil 10	LSTM Sinir Ağı Yapısına Ait En Küçük Birim (Hücre)	39
Şekil 11	YSA Mimarisi	45
Şekil 12	Ege Bölgesi Fay Haritası	48
Şekil 13	Marmara Bölgesi Fay Haritası	49
Şekil 14	İstanbul İlçeleri için Fay Haritası.....	49
Şekil 15	Zaman Serisindeki Ana Şok Deprem Büyüklükleri	51
Şekil 16	Derinliğin Boxplot Grafiği	52
Şekil 17	Değişkenler Arasındaki Korelasyon Matrisi	53
Şekil 18	Derinlik ve Büyüklük Arasındaki İlişkinin Scatterplot Grafiği	53
Şekil 19	Kantil Regresyon Sonuçları.....	54
Şekil 20	25. Kantil Sonuçları	55
Şekil 21	50. Kantil Sonuçları	56
Şekil 22	75. Kantil Sonuçları.....	56
Şekil 23	Derinlik ve Mw Arasındaki Kantil Regresyon Eğrisi.....	57

Şekil 24	Makine Öğrenmesi Modellerinin Performans Karşılaştırması.....	58
Şekil 25	LSTM ile Zaman Serisi Tahmini-Deprem Büyüklüğü (Mw).....	59
Şekil 26	Makine Öğrenmesi ve LSTM Modellerinin Ortalama Mutlak Hata (MAE) ve Hata Kareler Ortalamasının Karekökü (RMSE) Karşılaştırması.....	60
Şekil 27	İstanbul için 10 km Derinlikte Isı Haritası.....	64
Şekil 28	Derinlik 10 km Olduğunda Tahmin Edilen Deprem Büyüklükleri.....	65
Şekil 29	İstanbul için 30 km Derinlikte Isı Haritası.....	66
Şekil 30	Derinlik 30 km Olduğunda Tahmin Edilen Deprem Büyüklükleri.....	68

ÇİZELGELER DİZİNİ

Çizelge 1: Mercalli Şiddet Ölçeği ile Richter Ölçeği Arasındaki İlişki.....	9
Çizelge 2: Tanımlayıcı İstatistikler.....	50
Çizelge 3: Regresyon Modeli Sonuçları.....	54
Çizelge 4: İstanbul İlçe Koordinatları.....	62
Çizelge 5: İstanbul İlçeleri İçin 10 km Derinlikteki Deprem Tahminleri.....	69
Çizelge 6: İstanbul İlçeleri İçin 30 km Derinlikteki Deprem Tahminleri.....	70
Çizelge 7: İstanbul İlçeleri İçin Basit Doğrusal Regresyon Modeli Sonuçları.....	71

SİMGELER VE KISALTMALAR

Simgeler

m	Hacim Dalgası
M	Yüzey Dalgası
M_d	Süreye Bağlı Büyüklük
MI	Yerel (Lokal) Büyüklük
M_s	Yüzey Dalgası Büyüklüğü
M_b	Cisim Dalgası Büyüklüğü
M_w	Moment Büyüklüğü, Deprem Büyüklüğü
A	Tanecik Titreşim Genliği
T	Periyot (Saniye)
D	Episantr ile Kayıtçı Arasındaki Uzaklık (Derece)
h	Odak Derinliği (Kilometre)
Q(D,h)	Düzeltilme Faktörü
G-R	Guttenberg Richter İlişkisi
N(M)	M Büyüklüğüne Eşit ya da Daha Büyük Deprem Adedi
a	Regresyon Katsayısı
b	Regresyon Katsayısı
\bar{M}	Ortalama Magnitüd Değeri
M_{min}	Minimum Magnitüd Değeri
R^2	Modelin Uyumluluk Göstergesi
n	Örneklem Büyüklüğü
σ^2	Hata Terimi Varyansı
β	Regresyon Katsayısı
τ	İncelenen Kantil Regresyon Değeri ($\tau=0,5$ medyan için)
$P_\tau(u)$	Kantil Hata Fonksiyonu
u	Hata Terimi
p	Otoregresyon Derecesi
d	Fark Alma Derecesi
q	Hareketli Ortalama Derecesi

W_h	Gizli Katman Ağırlığı
W_i	Giriş Katman Ağırlığı
h_{t-1}	Hücrenin Bir Önceki Çıkışı
X_t	Yeni Gelen Giriş
W_f	Ağırlık Değeri
b_f	Yan (bias) Değeri
i_t	Girdi Kapısı
f_t	Unutma Kapısı
o_t	Çıktı Kapısı
c_t	Hücre Durumu
x_t	t Zamandaki Girdi Vektörü
σ	Sigmoid Aktivasyon Fonksiyonu
date	Tarih
time	Zaman
lat	Enlem – Lantitude
long	Boylam – Longitude
depth	Derinlik

Kısaltmalar

MC	Mercalli Cetveli
MSK	Medvedev-Spoheur-Karnik
ML	Richter Ölçeği
KAF	Kuzey Anadolu Fay Hattı
DAF	Doğu Anadolu Fay Hattı
BBK	Bitlis Bindirme Kuşağı
TDBH	Türkiye Deprem Bölgeleri Haritası
AFAD	Afet ve Acil Durum Yönetimi Başkanlığı
EKK	En Küçük Kareler
BLUE	En İyi Doğrusal Yansız Tahmin Edici
LAD	En Küçük Mutlak Sapma Yöntemi
MÖ	Makine Öğrenmesi

SML	Denetimli Öğrenme – Supervised Machine Learning
UML	Denetimsiz Öğrenme – Unsupervised Machine Learning
ARMA	Otoregresif Hareketli Ortalama – Autoregressive Moving Average
ARIMA	Otoregresif Bütünleşik Hareketli Ortalama – Autoregressive Integrated Moving Average
AR	Otoregresyon
MA	Hareketli Ortalama
ES	Üstel Düzleştirme
ANN	Yapay Sinir Ağları
YSA	Yapay Sinir Ağları
RF	Rastgele Orman
SVM	Destek Vektör Makinesi
DVM	Destek Vektör Makinesi
KNN	K En Yakın Komşu
ACF	Otokorelasyon Fonksiyonu – Autocorrelation Function
PCF	Kısmi Otokorelasyon Fonksiyonu – Partial Autocorrelation Function
ARCH	Otoregresif Koşullu Değişen Varyanslılık
GARCH	Genelleştirilmiş Otoregresif Koşullu Değişen Varyanslılık
ECM	Hata Düzeltme Modeli
NN	Derin Öğrenme
MLP	Çok Katmanlı Algılayıcı – Multi Layer Perception
RNN	Tekrarlı Sinir Ağları – Recurrent Neural Network
LSTM	Uzun Kısa Süreli Bellek – Long Short Term Memory
AIC	Akaike Bilgi Kriteri
DVR	Destek Vektör Regresyonu
SVC	Destek Vektör Sınıflandırması – Support Vector Classification
SVR	Destek Vektör Bağlayıcısı – Support Vector Regressor
ÇKA	Çok Katmanlı Algılayıcı
MAE	Ortalama Mutlak Hata
RMSE	Hata Kareler Ortalamasının Karekökü

1. GİRİŞ

Üzerinde yaşamakta olduğumuz Dünyamız, sanıldığı aksine cansız bir oluşum değil, canlı bir oluşumdur. Bu canlılık sebebiyle yer kabuğunun alt katmanlarında hareketler meydana gelerek, kırılmalar oluşmaktadır. Bu kırılmalar sonrasında bir enerji meydana gelmekte ve bu enerji yeryüzünde titreşimlere sebep olmaktadır. İşte bu titreşimlerin dalgalar halinde yayılarak, geçtikleri ortamları ve yeryüzünü sarsma olayı deprem olarak tanımlanmaktadır. Dünyanın varoluşundan beri, belirli bölgelerde sismolojik yönden aktif bulunan coğrafyalarda depremlerin tekrarlı ve ardışık olarak oluştuğu ve bu depremlerin sonucunda da milyonlarca insanın ve canlı yaşamının etkilendiği bilinmektedir (Edemen vd., 2023). Türkiye, jeolojik konumu dolayısıyla dünyada en sık deprem oluş periyoduna sahip ülkelerden birisidir. Ülkemiz Azor Adalarından başlayıp, Güneydoğu Asya'ya kadar uzanan Alp-Himalaya Deprem Kuşağı'nın Doğu Akdeniz Bölgesinde deprenselliğin (sismisitenin) en karmaşık olduğu kesimde yer almaktadır. Bu karmaşıklık, bölgede değişik boyutlarda ve hızlarda levhaların mevcudiyetinden kaynaklanmaktadır. Bölgede deprem oluşumunda egemen rolü Afrika, Arap ve Avrupa levhaları oynamaktadır (İnangu ve Kırbaş,1999). Ülkemiz, jeolojik yapısı itibarıyla yüksek rakımlı, eğimli arazilere sahip olup aktif fay hatları üzerinde yer aldığı için doğal afetlere sıkça maruz kalmaktadır. (Genç, 2007). 2020 yılında gerçekleşen Elazığ depremi (Mw=6.8), Sisam Adası açıklarında meydana gelen Sisam depremi (Mw=6.9) ile 2023 yılında Kahramanmaraş merkezli olarak yaşanan depremler (Mw=7.7 ve Mw=7.6), son yıllarda önemli düzeyde can ve mal kaybına yol açan sismik olaylar arasında yer almaktadır. Geçmiş depremlerde edinilen tecrübeler, bu tür sismik olayların hem maddi hem de manevi açıdan insan yaşamı üzerinde telafisi güç etkilere neden olduğunu ortaya koymaktadır. Ülkemizde meydana gelen depremler ve bunların sonuçları dikkate alındığında, afet yönetimine yönelik kapsamlı çalışmaların yürütülmesi zorunlu bir hâl almıştır. Bu bağlamda, afet yönetimi süreçlerinde gelişmiş bilişim teknolojilerinden, son yıllarda önem kazanan makine öğrenmesi ve yapay zeka gibi ileri analiz tekniklerinden yararlanılması, değerlendirmelerin daha isabetli ve hızlı bir şekilde gerçekleştirilmesine katkı sağlayacaktır (Demirelli, Solak ve Tiryakioğlu, 2023).

Günümüzde deprensellik, jeolojik ve tektonik gözlemlerin yanı sıra istatistiksel yöntemlerle belirlenebilmektedir. İstatistiksel yöntemler, geçmişte oluşan depremlerin yeri, zamanı ve büyüklüğü ile gelecekte olması muhtemel depremlerin yeri, zamanı ve büyüklüğünü hesaplamak, depremlerin oluşturduğu tahribatın, can ve mal kayıplarının önüne geçmek için

önemlidir (Doğruyol, 2020). Depremin olumsuz etkilerini en aza indirebilmek amacıyla bu tez çalışmasında, Ege Bölgesi'nde meydana gelmiş depremlerden oluşturulan deprem verileri kullanılarak, ileride İstanbul'da yaşanması beklenen depremlerin sonucunda ortaya çıkabilecek durumlar istatistiksel yöntemlerden regresyon analizi, zaman serileri analizi, kantil regresyon ve makine öğrenimi yöntemleri kullanılarak en uygun model elde edilmeye çalışılmıştır. Yapılan analizler sonrasında en uygun model olarak XGBoost yöntemi seçilmiştir.

Bu çalışmada, özel olarak İstanbul ilçelerinin ele alınmasının temel nedeni, İstanbul'un ülkemizde sismik risk olarak en yüksek bölgelerden birinde olmasından kaynaklanmaktadır. İstanbul, Kuzey Anadolu Fay Hattının batı kısmına oldukça yakın bir konumda yer almasından ve Kuzey Anadolu Fay Hattının 1939'da Erzincan depreminden başlayarak batıya doğru kademeli bir kırılma sergilemesi sebebiyle büyük depremler üretmiştir ve literatürde bu kırılma zincirinin Marmara Denizinin içerisinden geçerek İstanbul'a kadar ulaşan bir segment ile devam ettiğini ortaya koymakta olduğunu belirtmektedir (Parsons, 2004).

İstanbul'dan geçen bu aktif fay hattı sebebiyle, Marmara Denizi'nin kuzey kesimindeki deniz tabanı segmentlerinde büyük bir gerilim birikim olduğunu ve bunun potansiyel olarak büyük bir deprem üretebileceği şeklinde öngörüler olmaktadır. Bu öngörülere dikkat çeken Kandilli Rasathanesi ve uluslararası sismoloji merkezleri önümüzdeki 30 yıl içerisinde İstanbul'da $M_w \geq 7.0$ büyüklüğünde bir depremin meydana gelme olasılığının yüksek olduğunu raporlamaktadır (Boğaziçi Üniversitesi, 2019).

İstanbul, çeşitli dönemlerde birçok yıkıcı deprem ile sarsılmıştır. 10 Eylül 1509'da "Küçük Kıyamet" olarak adlandırılan deprem yaklaşık olarak 7.2 büyüklüğünde gerçekleşmiştir ve 1000-13000 can kaybına ve ciddi yapısal tahribatlara sebebiyet vermiştir (Ambraseys and Finkel, 1990, 2020). 22 Mayıs 1766'da 7.1 büyüklüğünde gerçekleşen depremde ise yaklaşık 4000 kişi hayatını kaybetmiş, tsunami etkileri gözlemlenmiştir (Ambraseys, 2000). 10 Temmuz 1894'teki 7.0 büyüklüğündeki depremde ise 1349'dan fazla can kaybı yaşanmış ve 1.5 metreye ulaşan tsunami yaşanmıştır (Finkel and Ambraseys, 1997).

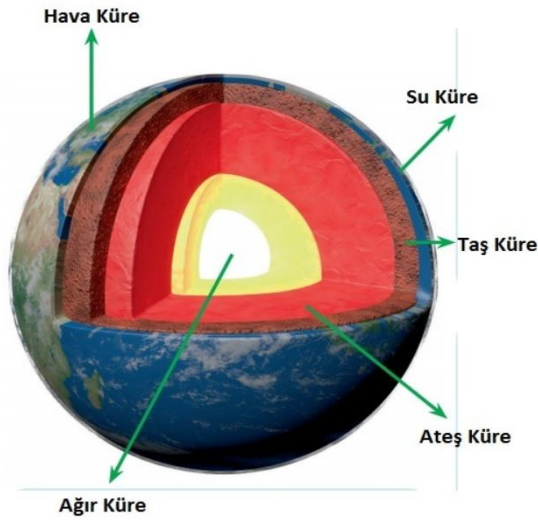
Geçmişte yaşanan depremler göz önünde bulundurulduğunda İstanbul ilçelerinin incelenmesi, bölgesel risk farklılıklarının ve jeolojik kırılma düzeylerinin daha detaylı anlaşılmasına katkı sağlamıştır. Çünkü İstanbul'un ilçelerinin zemin yapısı, yapıların yoğunluğu, yapılaşma şekilleri ve yerel fay hatlarına uzaklıkları bakımından büyüklükleri çeşitlilik göstermektedir. İlçeler düzeyindeki analizler, hem risk haritalarının oluşturulmasında hem de yerel afet

yönetimi stratejilerinin gelişmesine olanak sağlayacaktır. Bu sebeplerle, bu tez çalışmasında İstanbul ilçeleri özelinde yapılan modelleme ile hem mekânsal hem de istatistiksel doğruluğu yüksek tahminler üretilmesi amaçlanmıştır.

Bhatia vd. (2018) belirttiği gibi deprem tahmini, etkilenen bölgenin karşılaştığı ölüm sayısını ve ekonomik kaybı büyük ölçüde en aza indirebilir. Bu düşünce doğrultusunda, elde edeceğimiz en uygun model sayesinde, ilerleyen zamanlarda meydana gelebilecek depremlerde yerine ve büyüklüğüne göre olabilecek hasar tahmin edilebilecek ve buna göre önlemler alınabilecektir.

1.1. Deprem Nedir?

Dünyamız, katmanlardan oluşmaktadır. Bu katmanlar, Şekil 1’de gösterildiği gibi, içten dışa doğru olacak şekilde; “Barisfer” adı verilen “Ağır Küre”, “Manto” adı verilen “Ateş Küre”, “Litosfer” adı verilen “Taş küre”, “Hidrosfer” adı verilen “Su Küre” ve “Atmosfer” adı verilen “Hava Küre” şeklindedir. Litosfer katmanının altında bulunan Manto aslında sert bir katmandır ama Barisfer katmanına doğru inildikçe sıvı hale gelmektedir. Sıvılaştıran kısmın hareketinden kaynaklı taş küre kırılmakta ve parçalara ayrılmaktadır. Bu parçalar “levha” olarak adlandırılmakta ve bu levhaların sınırları da “fay hattı” nı oluşturmaktadır.



Şekil 1: Dünyanın Katmanları (<https://egitim.com/dunyanin-katmanlari>)

İşte Litosferde gerçekleşen bu kırılmalar sebebiyle ani olarak ortaya çıkan titreşimlerin dalgalar halinde yayılarak, yer yüzeyini sarsma olayı “deprem” olarak adlandırılmaktadır.

Sismik aktivite ile kastedilen, depremin meydana geldiği alandaki frekansı, türü ve büyüklüğüdür. Depremler, sismograf denilen bir alet yardımıyla ölçülür (Edemen vd., 2023). Depremlerin nasıl oluştuğunu, deprem dalgalarının yayılmalarını, ölçülmesini, yöntemlerini, deprem kayıtlarının değerlendirilmesi vb. gibi konuları inceleyen bilim dalına “sismoloji” adı verilmektedir (AFAD, t.y.).

1.2. Depremin Oluş Nedenleri Nelerdir?

Yer kabuğunu oluşturan levhaların katmanlardaki hareketlerden kaynaklı olarak, birbirlerine sürtündükleri, birbirlerini ittikleri veya birbirlerinin altına girdikleri levha sınırları dünyada depremlerin olduğu yerler olarak adlandırılmaktadır. Dünyada oluşan depremlerin birçoğu bu levha hareketlerinin gerçekleştiği yerlerde meydana gelmektedir (AFAD, t.y.).

Yurdumuz, Şekil 2’de gösterildiği gibi, dünyanın en etkili deprem kuşaklarından biri olan ve Azor Adası’ndan başlayarak Güneydoğu Asya’ya kadar uzanan Alp-Himalaya deprem kuşağında yer almaktadır. Ülkemizde afet yönetimi ve afetlere ilişkin istatistiksel verilerin sunulduğu AFAD tarafından 2018 yılında gerçekleştirilen çalışmaya göre, afet kaynaklı ölümlerin yaklaşık %60’ı depremlerden kaynaklanmaktadır. Yıllık ortalamalara bakıldığında, yaklaşık 1000 vatandaşımız yaşamını yitirmekte, 2100 kişi yaralanmakta ve 7000 civarında yapı zarar görmektedir (Demirelli, Solak, Tiryakioğlu, 2023). Ancak 6 Şubat 2023 tarihinde meydana gelen Kahramanmaraş merkezli depremin, Mercalli şiddet ölçeğine göre sarsıntılarının şiddeti, ölçeğin en yüksek değeri olan XII (Afetsel) olarak saptandı (Wikipedia, 2023) ve depremler sonucunda Doğu Anadolu Fay Hattı üzerindeki 11 ilde 48 binden fazla insan hayatını kaybetmiş, yarım milyondan fazla bina yıkılmış ya da ağır hasar almış ve önemli maddi kayıplar meydana gelmiştir (Türkiye Cumhuriyeti Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı, Kahramanmaraş ve Hatay Depremleri Raporu, 2023).

Levhaların hareketlerinden kaynaklı olarak, levhalar arasında hareket etmelerini engelleyen bir sürtünme kuvveti bulunmaktadır. Levhaların hareket etmesi için bu sürtünme kuvvetinin giderilmesi gerekmektedir. Bu kuvvet giderildiği zaman levhalar hareket eder, bu hareket çok kısa bir zaman diliminde gerçekleşir ve şok niteliğindedir. Bu şoklar sonucunda uzak mesafelere kadar etki edebilen deprem dalgaları meydana gelmektedir. Bu deprem dalgaları, geçtikleri yerleri sarsar ve merkezinden uzaklaştıkça enerjisini azaltır. Deprem dalgaları bazen gözle görülebilen ve kilometrelerce uzanabilen fay hattı adı verilen kırıkların yer yüzünde oluşmasına sebep olmaktadır.

1.3. Deprem Türleri

Depremler, oluşumlarına göre bazı alt başlıklara ayrılmaktadır (Edemen vd., 2023).

1.3.1. Tektonik Depremler

- Levha hareketleri sonucu çoğunlukla levha sınırlarında oluşur.
- Yeryüzünde gerçekleşen depremlerin %90'a yakını bu grupta olan depremlerdir.

1.3.2. Volkanik Depremler

- Yerin derinliklerindeki ergimiş maddenin yeryüzüne çıkışı sırasında fiziksel ve kimyasal olaylar sonucu oluşmuş gazların yapmış oldukları patlamalar sonucu meydana gelir.
- Japonya ve İtalya'da meydana gelen depremlerin bir kısmı bu gruba girmektedir

1.3.3. Çöküntü Depremler

- Yeraltındaki boşlukların, kömür ocaklarının, tuz ve jipsli arazilerdeki erimenin sonucunda oluşan boşlukların, tavan bloklarını çökertmesi sonucunda oluşur.
- Hissedilme alanları genellikle yereldir.
- Enerjileri azdır.
- Fazla zarar getirmezler.

1.3.4. Deniz Depremleri (Tsunami)

- Odakları deniz dibindedir.
- Denizlerin dibinden bazen kıyılara kadar oluşur ve kıyılarda da büyük hasarlara sebep olmaktadır.
- En çok görüldüğü yer olan Japonya'da, 1896 yılında gerçekleşen tsunamide 30.000 kişi ölmüştür. (Tan, 2021)

1.4. Deprem Fay Türleri

Faylar genelde hareket yönlerine göre isimlendirilmektedir.

Normal Faylar: Kabuğun iraksak sınır gibi uzanmış olduğu alanlarda meydana gelmektedir. 7 büyüklüğünden daha az olan depremler bu kısımlarda meydana gelmektedir.

Ters Faylar: Yakınsak sınır gibi kabuğun kısaldığı kısımlarda meydana gelmektedir. Yakınsak levha sınırları boyunca meydana gelenler genelde en büyük depremlerle veya mega depremlerle ilişkilendirilmektedir. Tümü 8 veya daha büyük büyüklüklere sahip depremlerdir.

Doğrultu Atılımlı Faylar: Fayın iki yakasının birbirini yatay olarak geçtiği dik yapılar şeklinde oluşurlar. Şiddeti 8 büyüklüğünden fazla olan depremler üretebilmektedirler.

Sağ ve Sol Yönlü Doğrultulu Faylar: Fayın oluşturduğu iki ayrı blokun birbirine görece olarak sağa ve sola doğru hareketlerinden meydana gelmektedir.

Eğimli Atılımlı Faylar: Düşey hareketler ile meydana gelen faylardır.

1.5. Deprem Parametreleri

Deprem tahminlerinde kullanılan başlıca parametreler; deprem koordinatları, gerçekleşme zamanı, depremin büyüklüğü ve odak derinliğidir (Guttenberg and Richter, 1954).

1.5.1. Odak Noktası (Hiposantr)

- Yer içinde deprem enerjisinin ortaya çıktığı noktadır.
- Odak noktası veya iç merkez de denmektedir.
- Aslında, enerjinin ortaya çıktığı bir nokta değil, alandır. Pratik uygulamalarda nokta şeklinde kabul edilmektedir.

1.5.2. Dış Merkez (Episantr)

- Yeryüzünde odak noktasına en yakın olan kısımdır.
- Depremin oluştuğunda en çok hasara sebep olduğu ve en kuvvetli hissedildiği noktadır.
- Depremler, şiddetine bağlı olarak çeşitli büyüklüklerde olabilmektedir.

1.5.3. Odak Derinliği

- Deprem enerjisinin açığa çıktığı noktanın yeryüzüne olan en kısa uzaklığıdır.
- Depremler, bu parametrelere göre sınıflandırılmaktadır. (0-70 km sığ depremler, 70-300 km orta derinlikteki depremler, 300 km'den fazla derinlikteki depremler derin depremler)
- Türkiye'de meydana gelen depremler genellikle sığ depremler olarak adlandırılmaktadır.

1.5.4. Eş Şiddet (İzoseit) Eğrileri

- Aynı şiddet ile sarsılan noktaları birbirine bağlayan noktalardır. Bu noktaların tamamlanmasıyla eş şiddet haritaları ortaya çıkmaktadır.
- İki eğrinin arasında kalan alan, depremlerden etkilenme yönüyle şiddet bakımından sınırlandırılmaktadır. Bundan dolayı deprem şiddetleri eş şiddet eğrileri üzerinde değil, alanın içerisinde yer almaktadır.

1.5.5. Şiddet

- Herhangi bir derinlikte olan depremin, yeryüzünde hissedildiği bir noktadaki etkisinin ölçüsü olarak adlandırılmaktadır.
- Etkide, depremin büyüklüğü, odak derinliği, uzaklığı, yapıların depreme karşı göstermiş olduğu dayanıklılığı dahi değişkenlik gösterebilmektedir.
- Depremler, gözlenen etki sonucunda ve uzun yılların vermiş olduğu tecrübelerle bakılarak hazırlanmış olan “Şiddet Cetvelleri” ne göre değerlendirilmektedir. Bu cetveller, her şiddet derecesindeki depremlerin insanlar, yapılar ve araziler üzerinde meydana getirdikleri etkiyi belirlemektedir.
- Bugün kullanılan başlıca şiddet cetvelleri, “Mercalli Cetveli (MC)” ve “Medvedev-Spoheur-Karnik (MSK)” şiddet cetvelleridir. Bu cetveller XII şiddet derecesini kapsamaktadır. Bu cetvellere göre, şiddeti V veya daha küçük depremler genellikle yapılarda hasar meydana getirmezler ve insanların depremleri hissetme şekline göre değerlendirilmektedir.
- Yeryüzünde ölçülmüş en büyük deprem 22 Mayıs 1960 tarihinde Şili Valdivia’da meydana gelen 9,5 şiddetindeki depremdir.

1.5.6. Magnitüd

- Deprem sırasında açığa çıkan enerjinin bir ölçüsü olarak değerlendirilmektedir.
- ABD’de Prof. C. Richter tarafından 1930 yılında bulunan bir yöntemle direkt olarak ölçülemeyen enerjinin aletsel bir ölçüsü olan magnitüd tanımlanmıştır.
- Bugüne kadar olan depremler istatistiksel olarak değerlendirildiğinde, 8,9 büyüklüğünde 31 Ocak 1906’da Colombiya-Ekvator ve 2 Mart 1933’de Sanriku-Japonya depremleri olmuştur.

- Gözlemleri tarafından bildirilen depremin magnitudü deprem enerjisi hakkında bilgi vermez çünkü deprem sığ veya derin odaklı olabilir. Magnitudü aynı olan iki depremden sığ olanı daha çok hasar verirken, derin olanı daha az hasar yapacağından arada bir fark olacaktır. Yine de Richter Ölçeği (magnitudü), depremlerin özelliklerini saptamada çok önemli bir unsur olmaktadır (Şahin, 2023).

Aletsel Magnitudü: Standart bir sismograf ile kaydedilen deprem hareketlerinin maksimum genlik ve periyod değerleri ve alet kalibrasyon fonksiyonlarının kullanılmasıyla yapılan hesaplamalar sonrasında elde edilmektedir. Bazen hacim dalgaları (m) bazen de yüzey dalgaları (M) üzerinden hesaplanmaktadır.

Gözlemsel Magnitudü: Gözlemsel inceleme sonucu elde edilen episantr şiddetinden hesaplanmaktadır. Magnitudü ile şiddet bağlantısı, incelenen bölgeden bölgeye farklılık göstermektedir.

Süreye Bağlı Büyüklük (Md): Depremin, sismometre üzerinde ne kadar uzun süreli bir titreşim oluşturduğunun ölçülmesi ve deprem merkezinin uzaklığı ile ölçeklenmektedir. Bu yöntem genellikle $M < 5,0$ olan küçük ve uzaklık < 300 km olan yakın depremler için kullanılmaktadır.

Yerel (Lokal) Büyüklük (MI): Richter tarafından 1935 yılında depremlerin ölçülmesi için önerilen ilk yöntemdir. Büyüklük $< 6,0$ olan görece küçük ve uzaklık < 700 km olan yakın depremlerin ölçülmesinde kullanılmaktadır.

Yüzey Dalgası Büyüklüğü (Ms): İlk iki yöntemin yetersiz kaldığı alanlarda $M > 6,0$ olan büyük depremleri ölçmek için kullanılmaktadır.

Cisim Dalgası Büyüklüğü (Mb): Yer kabuğunun içinde yayılan dalgaları ölçmek için kullanılmaktadır.

Moment Büyüklüğü (Mw): Diğer ölçeklere göre en güvenilir olan bu yöntem $M > 4,0$ için hesaplanabilmektedir.

Richter Ölçeği (ML): Deprem magnitudünü tanımlamak için kullanılan matematiksel bir formüldür. Dalga genişliğinin logaritması alınarak tanımlanmaktadır.

$$N(M) = a - bM$$

Çizelge 1: Mercalli Şiddet Ölçeği ile Richter Ölçeği Arasındaki İlişki

Şiddet	IV	V	VI	VII	VIII	IX	X	XI	XII
Richter	4	4,5	5,1	5,6	6,2	6,6	7,3	7,8	8,4

Çizelge 1’de verilen Richter Ölçeği dışında 3 yolla daha magnitüd hesaplaması yapılmaktadır. (İşçi, 2008).

1) P ve S Dalgalarından (body-wave magnitude) (Mb):

$$mb = \log_{10} \left(\frac{A}{T} \right) + Q(D, h)$$

A: tanecik titreşim genliği,

T: periyot (saniye)

D: episantr ile kayıtçı arasındaki uzaklık (derece)

h: odak derinliği (kilometre)

Q(D,h): düzeltme faktörü

2) Yüzeysel Dalgalarından (surface-wave magnitude) (Ms):

$$M_s = \log_{10} \left(\frac{A}{T} \right) + 1.66 \log_{10}(D) + 3.30$$

3) Sismik Momentten (moment magnitude) (Mw):

$$M_w = \frac{2}{3} \log_{10}(M_0) - 10.7$$

M_w: sismik moment

1.6. Guttenberg-Richer Büyüklük-Frekans Bağlantısı

Deprem istatistiklerine yönelik çalışmalarda en yaygın biçimde başvurulan ve literatürde iyi bilinen parametrelerden biri, depremlerin büyüklük-frekans dağılımını tanımlayan Gutenberg-Richter (G-R) ilişkisi olarak öne çıkmaktadır. (Guttenberg and Richter, 1944) Gutenberg-Richter ilişkisinde yer alan *b* katsayısı, yalnızca küçük ve büyük ölçekli depremlerin göreceli oranlarını yansıtmakla kalmayıp, aynı zamanda bölgesel gerilme birikimindeki zaman, derinlik ve mekânsal değişimlerle birlikte, ilgili bölgenin sismotektonik karakteristikleri hakkında da önemli bilgiler sunmaktadır. (Öztürk ve Alkan, 2022).

Guttenberg-Richter Bağlantısı:

$$\log N(M) = a - bM$$

şeklinindedir. Burada;

N(M) → *M* büyüklüğüne eşit ya da daha büyük deprem adedi,

$M \rightarrow$ Richter büyüklüğü,

a ve $b \rightarrow$ Regresyon katsayıları,

$a \rightarrow$ incelenen bölgenin alanına, gözlem süresine ve bu alanın deprem üretkenliğine bağlı olarak değişen katsayı,

$b \rightarrow$ bölgenin tektonik özelliklerine bağlı olarak değişim gösterir, sığ depremler için; $0,90 \pm 0,02$, orta ve derin odaklı depremler için $1,2 \pm 0,2$ olarak verilir. (Hiçyılmaz ve Sezer, 2020).

Türkiye için b değeri ise $0,9 \pm 0,2$ değerlerinde verilmektedir.

Gutenberg-Richter ilişkisi kapsamında b değeri, magnitüd-frekans dağılımının eğiminden türetilirken, a parametresi ise bölgedeki sismik aktivite düzeyiyle doğrudan ilişkilidir. a değerinde gözlenen farklılıklar, büyük ölçüde çalışma alanının genişliği, kullanılan deprem kataloğunun süresi ve analiz edilen sismik olay sayısına bağlı olarak değişkenlik göstermektedir. (Öztürk ve Alkan, 2022).

b değeri, dünyanın farklı sismoteknik bölgeleri için $0,3-2,10$ arasında değişim göstermektedir (Utsu, 1971) ve ortalama b değeri $1,0$ civarındadır (Frohlich and Davis, 1993).

Herhangi bir bölgedeki b değerini hesaplamak için birçok yöntem mevcuttur. Fakat b değerinin tahmini için en güçlü ve en yaygın olarak kullanılan yöntem maksimum olasılık yöntemi olup,

$$b = \frac{2,303}{(\bar{M} - M_{min} + 0,05)}$$

denklemleri ile verilir (Aki, 1965). Burada \bar{M} , ortalama magnitüd değeri ve M_{min} ise minimum magnitüd değeridir. $0,05$ değeri ise bir ilişki sabitidir. %95 güven aralığında sismik b değerinin standart sapması, Aki (1965) tarafından önerilen $\pm 1,96b/\sqrt{n}$ denklemi kullanılarak hesaplanabilir. Bu denklemde n parametresi, tahmin yapmada kullanılan depremlerin sayısıdır (Öztürk, 2014).

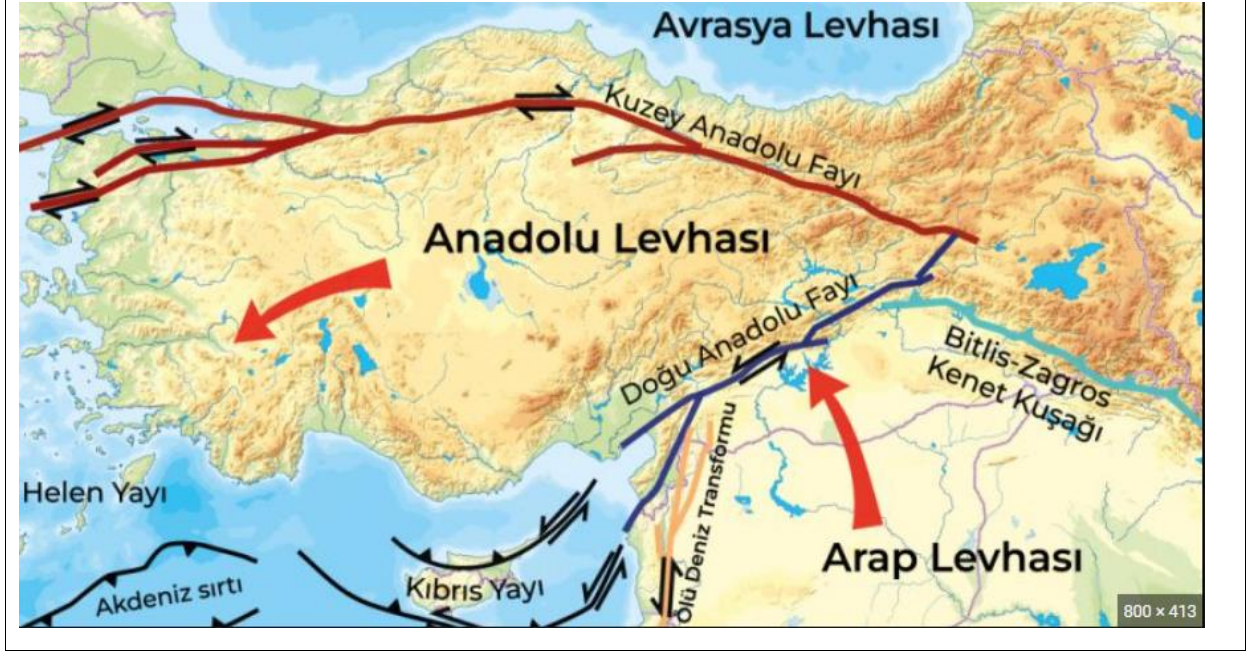
Düşük b değerine sahip bölgelerin, genellikle yüksek gerilme birikimiyle karakterize edilen alanlarla ilişkili olduğu belirtilmektedir. Ayrıca, düşük b değeri ile yüksek Z değeri (anomalik sismisite yoğunluğu) gösteren bölgelerin, gelecekte meydana gelebilecek büyük ölçekli depremler açısından potansiyel olarak en riskli alanlar olduğu değerlendirilmektedir. (Öztürk ve Alkan, 2022).

Weeks et al. (1978), kayaçlar üzerinde yaptıkları deneyler sonrasında b değerinin deprem öncesi azalma gösterdiğini belirtmişlerdir.

1.7. Türkiye’de Depremler

Türkiye, jeodinamik konumu itibarıyla, dünyada en sık deprem üreten bölgelerden biri olarak öne çıkmaktadır. Ülkemiz, Azor Adaları’ndan başlayarak Güneydoğu Asya’ya kadar uzanan ve yüksek sismik aktiviteye sahip Alp-Himalaya Deprem Kuşağı üzerinde yer almakta olup, bu kuşağın Doğu Akdeniz segmentinde sismisitenin en karmaşık ve yoğun olarak gözlemlendiği kesimde konumlanmaktadır. Söz konusu tektonik karmaşıklık, bölgede farklı yönlerde ve hızlarda hareket eden levhaların özellikle Afrika, Arap ve Avrasya levhalarının karşılıklı etkileşiminden kaynaklanmaktadır. Bu levha sınırlarının kesişim noktalarında biriken gerilmeler, Türkiye’nin çok sayıda aktif fay hattına sahip olmasına ve yüksek deprem riski taşımaya neden olmaktadır. (İnangu ve Kırbaş, 1999). Türkiye, topoğrafik olarak yüksek rakımlı ve eğimli arazilerden oluşan bir coğrafi yapıya sahip olup, aynı zamanda çok sayıda aktif fay hattının kesişim noktasında yer almaktadır. Bu özellikleri nedeniyle, ülkemizde başta depremler, heyelanlar ve sel olayları olmak üzere çeşitli doğal afetlerin meydana gelme sıklığı oldukça yüksektir. (Genç, 2007). Deprem Bölgesi Haritasına göre, yurdumuzun %92’sinin deprem bölgeleri içerisinde olduğu, nüfusumuzun %95’inin deprem tehlikesi altında yaşadığı ve ayrıca büyük sanayi merkezlerimizin %98’inin ve barajlarımızın %93’ünün deprem bölgesinde bulunduğu bilinmektedir (Şahin ve Sipahioğlu, 2002). 2020 yılında meydana gelen Elazığ depremi (Mw=6.8), Ege Denizi’nde Sisam Adası açıklarında gerçekleşen 2020 Sisam depremi (Mw=6.9) ile 2023 yılında Kahramanmaraş merkezli olarak yaşanan depremler (Mw=7.7 ve Mw=7.6), son yıllarda ülkemizde ciddi can ve mal kayıplarına yol açan önemli sismik olaylardandır. (Demirelli, vd., 2023).

Ülkemiz, coğrafi konumundan kaynaklı olarak birçok levha üzerindedir. Bu levhalar, Şekil 3’te gösterildiği gibi olmaktadır (İnangu ve Kırbaş, 1999).



Şekil 3: Türkiye'deki Levhalar ve Fayların Haritası
(https://tr.wikipedia.org/wiki/Anadolu_levhası)

Kuzey Anadolu Fayı (KAF): Anadolu ve Avrasya levhaları arasındaki sınırı oluşturan Kuzey Anadolu Fayı (KAF), doğuda Doğu Anadolu Fayı (DAF) ile Karlıova üçlü eklem bölgesinde keşşerek başlamakta ve batıya doğru ilerledikçe geniş açılı bükülmeler ve segment atlamalarıyla uzanmaktadır. Yaklaşık 1000 km uzunluğunda olan KAF, sağ yönlü doğrultu atımlı bir fay zonu olup, tekil ve sürekli bir kırık yapısı değil, farklı uzunluklara ve jeolojik özelliklere sahip çok sayıda aktif fay segmentinden oluşan karmaşık bir fay sistemidir.

Doğu Anadolu Fayı (DAF): Anadolu Levhasının doğu sınırını oluşturan Doğu Anadolu Fayı (DAF) zonu, Karlıova üçlü eklem bölgesinden başlayarak İskenderun Körfezi'ne kadar uzanmakta olup, yaklaşık 400 km uzunluğa sahiptir. Sol yönlü doğrultu atımlı bir karakter sergileyen bu fay, kuzeydoğu-güneybatı doğrultusunda uzanmakta ve yaklaşık 2-3 km genişliğinde bir tektonik zon oluşturmaktadır. Tıpkı Kuzey Anadolu Fayı (KAF) gibi, DAF da farklı uzunluk ve özelliklere sahip çok sayıda fay segmentini bünyesinde barındıran karmaşık bir fay sistemidir.

Bitlis Bindirme Kuşağı (BBK): Doğu Anadolu Fayı (DAF), tarihsel süreçte yüksek sismik aktivite göstermiş ve birçok yıkıcı depreme kaynaklık etmiştir. Ancak 20. yüzyılda göreceli olarak daha düşük bir sismisite sergilemiş olup, bu dönemde yalnızca 1975 yılında meydana gelen ve ciddi yapısal hasarlara yol açan Lice Depremi gibi belirgin bir tahrip edici deprem üretmiştir.

Ülkemizde 1970’li yıllardan itibaren deprem sayılarında belirgin bir artış gözlemlenmektedir. Bu artışın nedenleri arasında, sismik izleme kapasitesinin gelişmesi ve deprem kayıt istasyonu sayısındaki artışın önemli bir rol oynadığı değerlendirilmektedir. Öte yandan, 1990 ile 2010 yılları arasında kaydedilen sismik aktivitedeki belirgin artışın, bu dönem sonrasında nispeten azalma eğilimine girdiği dikkat çekmektedir. (Doğan, 2023). Özellikle son yıllarda meydana gelen depremler, insan yaşamı ve yerleşim alanları üzerinde ciddi düzeyde etkiler oluşturmuştur. Bu durum, deprem verileri üzerinde güvenilir ve doğru tahmin modelleri geliştirmenin önemini daha da artırmaktadır. (Akyüz, 2019). Deprem öngörü modelleri, genellikle sismik olayların meydana geleceği yer, zaman ve büyüklük olmak üzere üç temel bilinmeyen belirlenmeye yönelik olarak geliştirilmektedir. (Doğan ve Demir, 2020).

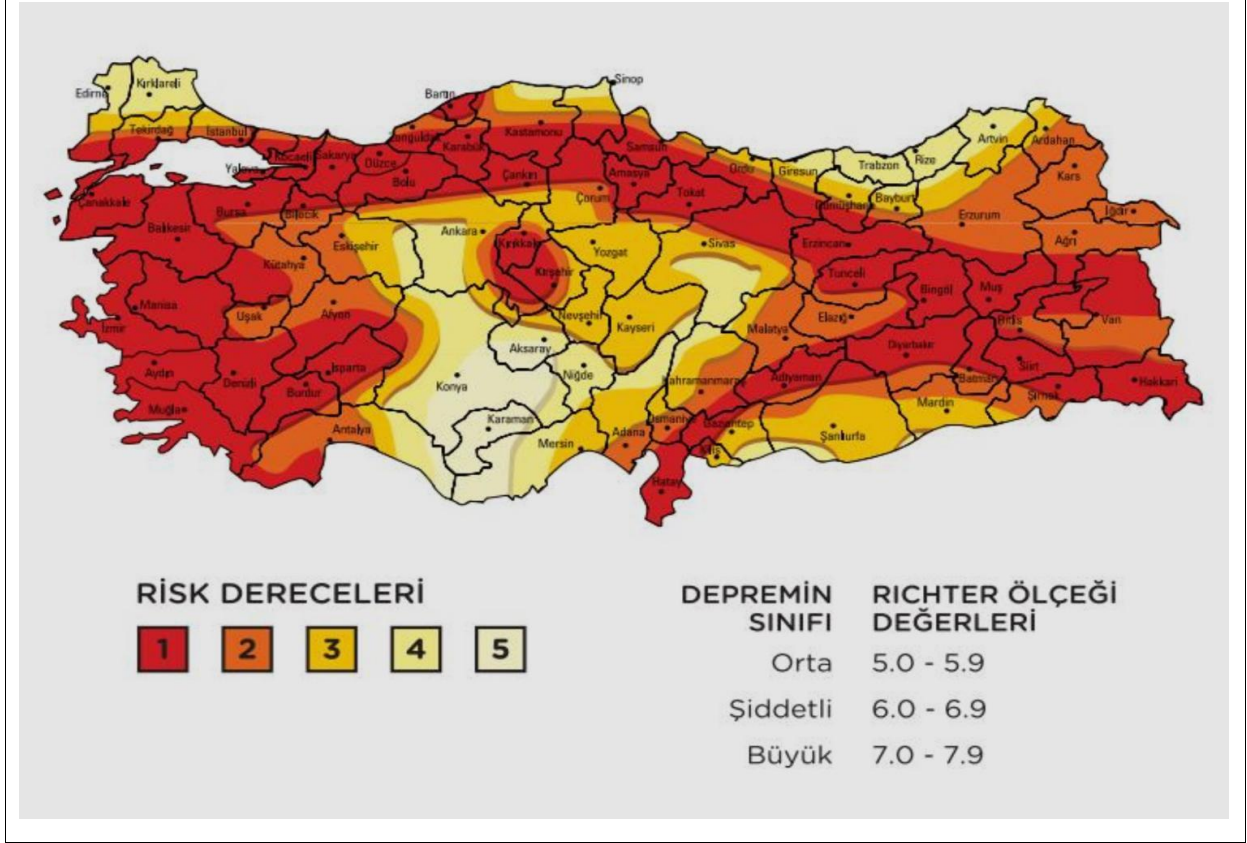
En son 1996 yılında yürürlüğe giren Türkiye Deprem Bölgeleri Haritası (TDBH), Afet ve Acil Durum Yönetimi Başkanlığı (AFAD) Deprem Dairesi Başkanlığı tarafından yenilenmiş, 18 Mart 2018 tarih ve 30364 sayılı (mükerrer) Resmi Gazetede yayınlanmıştır. Yeni harita 1 Ocak 2019 tarihinde yürürlüğe girmiştir. Yeni harita en güncel deprem kaynak parametreleri, deprem katalogları ve yeni nesil matematiksel modeller dikkate alınarak, çok daha fazla ve ayrıntılı veri ile hazırlanmıştır. Yeni haritada bir öncekinden farklı olarak deprem bölgeleri yerine en büyük yer ivmesi değerleri gösterilmiş ve “deprem bölgesi” kavramı ortadan kaldırılmıştır (Doğruyol, 2021).

1.8. Deprem Kronolojisi

Kayıtlı tarihin en yıkıcı depremlerinden birisi, 23 Ocak 1556’da Çin-Şensi’de meydana gelen 1556 Şensi Depremi’dir. 830.000’den fazla insan ölmüştür.

20. yüzyılın en ölümcül depremi ise 1976’da Tangshan Depremidir ve yaklaşık 655.000 kişinin ölümüne sebep olmuştur (Wikipedia, 2023).

Türkiye’deki deprem risk derecelerini gösteren harita ise Şekil 4’te verilmiştir.



Şekil 4: Türkiye Deprem Kuşakları Haritası (AFAD, 2023)

1.9. Deprem Sonuçları

Depremler, şiddetlerine, büyüklüklerine ve etkileme alanlarına göre aşağıdaki gibi bazı sonuçlar ortaya çıkartabilmektedir.

- Sallantı,
- Yeryüzü çatlamaı,
- Heyelan,
- Yangınlar,
- Zemin sıvılaşması,
- Tsunami,
- Seller,
- Gelgit kuvveti,
- İnsana etkileri;
 - Salgın hastalık,
 - Temel ihtiyaçlarda eksiklik,
 - Yaralanma

- Can kayıpları,
- Yüksek sigorta primleri,
- Kritik yapılarda hasar,
- Yollar, köprüler ve toplu taşıma araçlarında hasar,
- Su, elektrik ve gaz hatlarında kesinti,
- İletişim sistemlerinde hasar,

Yakın zamanda ülkemizde gerçekleşen 6 Şubat Kahramanmaraş merkezli 7,8 büyüklüğündeki depremin sonuçları yukarıdaki yazılan sonuçlara en yakın örnek olarak gösterilebilmektedir. Deprem sonrasında 50 bin 783 vatandaşımız hayatını kaybederken, milyonlarca vatandaşımız yaralanmış ve şehirlerdeki binaların çoğu yıkılmıştır. Depremin şiddetinden kaynaklı olarak yollarda büyük kırılmalar meydana gelmiş ve depremden etkilenen şehirlerimize yardımlar bu kırılmalar sebebiyle gecikmiş, yollarda uzun araç kuyrukları oluşmuş ve havaalanlarının zeminlerinin de kırılması sebebiyle havadan yardım gelmesi de zorlaşmıştır.

Depremin kış mevsiminde meydana gelmesi ve sarsıntılar kaynaklı enerji kaynaklarının sağlanmasında kullanılan boruların kırılması sebebiyle insanlar kalacak yer bulamamış, temel ihtiyaçları olan yeme, içme, barınma ve hijyen ihtiyaçlarını giderememişlerdir. Bu yoksunluk sebebiyle hastalıklar ortaya çıkmış, enkazdan etkilenmemiş olsa dahi soğuk hava dolayısıyla hayatını kaybeden vatandaşlarımız olmuştur.

Strateji ve Bütçe Başkanlığı tarafından 18 Mart 2023 tarihinde yayınlanan rapora göre ülkemizin yaşamış olduğu bu felaketin ülkemize maliyeti 2 trilyon lira (103 milyar \$) olmuştur.

2. YÖNTEMLER

Bu bölümde, regresyon analizi ve kantil regresyon analizi gibi klasik yöntemlerden bahsedildikten sonra son yıllarda popüler olan makine öğrenmesi yöntemleri anlatılacaktır.

2.1. Regresyon Analizi

Regresyon analizi, iki ya da daha çok nicel değişken arasındaki ilişkiyi ölçmek için kullanılan bir yöntemdir. Eğer tek bir açıklayıcı değişken kullanılarak analiz yapılıyorsa buna basit regresyon, birden çok açıklayıcı değişken kullanılıyorsa çoklu regresyon analizi olarak isimlendirilir. Regresyon analizi ile değişkenler arasındaki ilişkinin varlığı, eğer var ise, bunun etkisi hakkında bilgi edinilebilir (Gujarati, 2008). Doğrusal regresyonun temel amacı, bağımlı ve bağımsız değişkenler arasındaki doğrusal ilişkiyi modellemek için bir denklem uyarlamak ve bu sayede bağımsız değişkenlerin değerleri verildiğinde, bağımlı değişkenin değeri hakkında güvenilir tahminler yapma olanağı sağlamaktadır (Altman and Krzywinski, 2015).

Regresyon modeli, bağımlı değişken (y) üzerinde bağımsız değişken(ler)in (x_1, x_2, \dots) etkisini inceler ve

$$y = a + bx \quad (2.1)$$

olarak formüle edilir. Değişkenlerden birinin değeri bilindiğinde diğerinin değeri bulunur. Eğer model için bulunan p değeri 0.05'ten küçük ise regresyon katsayısı 0'dan farklıdır, yani iki değişken arasındaki ilişki istatistiksel olarak önemlidir, başka bir deyişle iki değişken arasında doğrusal bir ilişki vardır. Buna karşılık, modelin uyumluluk göstergesi R^2 ile ifade edilir ve R^2 değeri 1'e ne kadar yakınsa model o kadar çok bağımlı değişkeni açıklar (Dawson and Trapp 2001).

Eşitlik (2.1)'de " a " olarak gösterilen sabit terim, regresyon doğrusunun y eksenini kestiği nokta olmaktadır. Buna karşılık " b " olarak gösterilen regresyon katsayısı (eğim) ise analitik yöntemde oransal (proporsiyonel) bir ölçüdür. Yani, bağımsız değişkendeki bir birimlik değişme (artma ya da azalma) olduğunda, bağımlı değişkende meydana gelecek değişim miktarını gösterir (Pagano and Gauvreau, 1993).

2.1.1.Literatürde Regresyon Analizinin Önerilişi

Regresyon analizinin temellerini oluşturan En Küçük Kareler (EKK) yöntemi, ilk olarak 1805 yılında Fransız matematikçi Adrien-Marie Legendre tarafından ortaya konmuştur. Legendre, bu yöntemi 'moindres carrés' (en küçük kareler) terimiyle tanımlamış ve astronomik gözlemler doğrultusunda gök cisimlerinin yörüngelerinin belirlenmesinde kullanmıştır. Kısa bir süre sonra, 1809 yılında Carl Friedrich Gauss, aynı yöntemi daha ayrıntılı biçimde açıklamış ve bu yöntemi 1795'ten itibaren kullanmakta olduğunu öne sürmüştür. Her iki bilim insanı da bu yöntemi, dönemin önemli bilimsel problemlerinden biri olan gezegenlerin ve uyduların Güneş etrafındaki yörüngelerinin belirlenmesinde uygulamıştır. EKK yöntemine ilişkin önemli bir teorik ilerleme ise Gauss'un 1821 yılında yayımladığı bir çalışmada gerçekleşmiş, bu çalışmada Gauss, daha sonra kendi adıyla anılacak olan Gauss-Markov teoreminin bir versiyonunu sunmuştur. Bu teorem, belirli varsayımlar altında EKK tahmin edicilerinin en küçük varyansa sahip, en yansız (BLUE – Best Linear Unbiased Estimator) tahmin ediciler olduğunu ortaya koymaktadır.

Regresyon terimi, 19. yüzyılda İngiliz İstatistikçi Francis Galton tarafından bir biyolojik incelemede kullanılmıştır. Bu incelemenin ana konusu kalıtım olup, aile içinde baba ve annenin boyu ile çocukların boyu arasındaki bağlantıyı araştırmakta ve çocukların boylarının bir nesil içinde eski ata nesillerinin ortalamasına geri döndüklerini, yani bir nesil içinde ortalamaya geri dönüş olduğu inceleme konusu olmuştur. Galton geri dönüş terimi için ilk yazısında İngilizce olarak "reversion" terimini kullanmış olsa da sonradan aynı anlama gelen "regression" sözcüğünü kullanmıştır. Bu çalışmalarında Galton, istatistiksel regresyon kavramını ve yöntemini de geliştirmiştir (Audi, 1996). Udny Yule ve Karl Pearson bu yöntemi daha geniş genel istatistiksel alanlara uygulayıp geliştirmiştir (Yule, 1897). Bu yazılarda bağımlı ve bağımsız değişkenlerin normal dağılım gösterdiği varsayılmaktadır. Bu kısıtlayıcı varsayım R. A. Fisher tarafından 1922 ve 1925 yıllarındaki yayınları ile sadece bağımlı değişkenin koşullu dağılımının normal dağılım olduğu gösterilmiştir (Fisher, 1922).

2.1.2. Regresyon Analizinde Temel Varsayımlar

- Çıkarımsal yöntemler olduğu için kullanılan iki değişkenli örneklemin ya olasılıksal ya da rastgele örneklem olduğu, yani kitleyi çok iyi temsil ettiği bilinmektedir.
- Bağımlı değişken için oluşturulan modelde hata terimi bulunmaktadır. Bu hatanın bir rastgele değişken olması ve ortalamasının sıfır olması gerekir. Sistemik hata da

bulunması mümkündür ancak bu tür hatanın incelemeye alınması regresyon analizi kapsamının dışındadır.

- Bağımsız değişken için bir hata terimi yoktur. Eğer bağımsız değişken içinde hata olduğu varsayılırsa özel bir yöntem şekli olan değişkenler içinde hata modeli teknikleri kullanılarak model kurulmalıdır.
- Hatalar zaman içerisinde ve kendi arasında birbirine bağımlı değildir. Buna, otokorelasyon veya serisel korelasyon sorununun olmaması varsayımı adı verilir.
- Hataların varyansının sabit ve zaman içinde hiç değişmediği varsayılır. Eğer hatalar bu özelliğe sahip olmazsa değişen varyanslılık (heteroskedastisite) sorunu ortaya çıkar. Bu durumda, Box-Cox dönüşümü ya da Ağırlıklı En Küçük Kareler Yöntemi tercih edilebilir.
- Hataların varyans-kovaryans matrisinin çapraz (diagonal) elemanları hataların sabit varyansı olurken matrisin diğer elemanları 0 olur.
- Eğer çoklu regresyon analizi yapılıyorsa, bağımsız değişkenlerin birbirleri ile ilişkili olmamaları gereklidir. Eğer bu varsayım sağlanmıyorsa, bu durumda çoklu bağlantı (multicollinearity) sorunu vardır ve Ridge Regresyon gibi yöntemler tercih edilebilir.
- Hatalar normal dağılım gösterir. Eğer bu hataların normalliği varsayımı sağlanmazsa, genelleştirilmiş doğrusal model tercih edilebilir (Gujarati,2008).

2.1.3. En Küçük Kareler (EKK) Yöntemi

Olağan En Küçük Kareler Yöntemini, Alman Matematikçi Carl Friedrich Gauss'un bulduğu kabul edilir. En Küçük Kareler (EKK) Yöntemi, belirli varsayımlar altında istenilen istatistiksel özellikleri taşıdığından, regresyon analizinde en basit model olarak kabul edilmektedir. Bu yöntemde model,

$$Y_i = \beta_1 + \beta_2 X_i + u_i,$$

şeklinde elde edilmekte olup, tahmin edilen değerleri ise, $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$ biçiminde elde edilmektedir.

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + u_i$$

$$Y_i = \hat{Y}_i + u_i$$

$$u_i = Y_i - \hat{Y}_i$$

$$u_i = Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i$$

$\sum u_i^2 = \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$ toplamının en küçük çıkması gerekmektedir. u_i (mutlak değer olarak) ne kadar büyükse, $\sum u_i^2$ o kadar büyük olur.

En Küçük Kareler Yöntemi $\hat{\beta}_1, \hat{\beta}_2$ değerlerini öyle bir hesaplar ki, verilmiş bir örneklem ya da veri kümesi için, $\sum u_i^2$ en küçük çıkar. Başka bir deyişle, En Küçük Kareler Yöntemi, regresyon modelinde en küçük hataları elde etmemizi sağlar. Bu yöntem, matematiksel olarak aşağıda gösterilmiştir:

$$\begin{aligned}\sum Y_i &= n\hat{\beta}_1 + \hat{\beta}_2 \sum X_i \\ \sum Y_i X_i &= \hat{\beta}_1 \sum X_i + \hat{\beta}_2 \sum X_i^2\end{aligned}$$

Burada, n , örneklem büyüklüğüdür. Bu eşanlı denklemler normal denklemler olarak adlandırılmaktadır. Bu normal denklemleri eşanlı çözersek;

$$\begin{aligned}\hat{\beta}_1 &= \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2} \\ &= \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} \\ &= \frac{\sum x_i y_i}{\sum x_i^2}\end{aligned}$$

Burada, \bar{X} ile \bar{Y} , sırasıyla X ile Y 'nin örneklem ortalamalarıdır, $x_i = X_i - \bar{X}$, $y_i = Y_i - \bar{Y}$ olarak tanımlanmıştır.

$$\begin{aligned}\hat{\beta}_1 &= \frac{\sum X_i^2 \sum Y_i - \sum X_i \sum X_i Y_i}{n \sum X_i^2 - (\sum X_i)^2} \\ &= \bar{Y} - \hat{\beta}_1 \bar{X} \\ \hat{\beta}_2 &= \frac{\sum x_i y_i}{\sum x_i^2} \\ &= \frac{\sum x_i Y_i}{\sum X_i^2 - n\bar{X}^2} \\ &= \frac{\sum X_i y_i}{\sum X_i^2 - n\bar{X}^2}\end{aligned}$$

Bu şekilde elde edilen tahmin ediciler, En Küçük Kareler İlkesinden türetildikleri için En Küçük Kareler Tahmin Ediciler adını alır (Gujarati, 2009). Veriler nasıl türetilmiş olursa olsun, olağan en küçük karelerin kullanımının doğurduğu sonuçlara sağlanan özellikler, sayısal özelliklerdir (Davidson and MacKinnon, 1993).

2.1.3.1 Gauss- Markov Teoremi

Klasik doğrusal regresyon modelinin varsayımları sağlanıyorsa, en küçük kareler tahmin edicileri, yansız (unbiased) ve doğrusal tahmin ediciler arasında en küçük varyanslı olanlardır. Literatürde bu tahmin ediciler, BLUE (Best Linear Unbiased Estimator) tahmin ediciler olarak bilinir. Bu tahmin ediciler;

1. Doğrusaldır, yani regresyon modelindeki bağımlı değişken Y gibi rastgele bir değişkenin doğrusal bir fonksiyonudur.
2. Yansızdır, yani ortalaması ya da beklenen değeri $E(\hat{\beta}_1)$ gerçek değere, (β_1) eşittir.
3. Doğrusal yansız tahmin ediciler içerisinde en küçük varyanslı olanıdır; bu nedenle, en küçük varyanslı yansız bir tahmin edici, yani etkin (efficient) bir tahmin edici diye adlandırılmaktadır.

2.1.3.2. En Küçük Kareler Yöntemi Tahmin Edicilerinin Ortalama ve Varyans

Özellikleri:

1. EKK tahmin edicisi olan $\hat{\beta}$, β 'nin yansız tahmin edicisidir. Yani, tahmin edici gerçek parametre değerini ortalama olarak doğru bir şekilde tahmin etmektedir (Wooldridge, 2013).

$$E[\hat{\beta}] = \beta$$

2. EKK tahmin edicisinin varyans-kovaryans matrisi aşağıda belirtildiği gibidir (Hayashi, 2000).

$$Var(\hat{\beta}) = \sigma^2(X^T X)^{-1}$$

Burada σ^2 , hata teriminin varyansıdır. Varyans-kovaryans matrisi, tahmin edicilerin ne kadar hassas (accuracy) olduğunu belirtir. Daha küçük varyans, daha etkin (efficiency) tahminler anlamına gelmektedir.

- $\hat{\beta}_j$ 'nin varyansı, $Var(\hat{\beta}_j) = \sigma^2[(X^T X)^{-1}]_{jj}$ şeklinde hesaplanmaktadır.
- Kovaryanslar ise $Cov(\hat{\beta}_j, \hat{\beta}_k) = \sigma^2[(X^T X)^{-1}]_{jk}$ şeklinde hesaplanmaktadır.

2.2. Kantil Regresyon Analizi

Kantil Regresyon Yöntemi, Koenker ve Bassett tarafından önerilmiş bir regresyon modelidir. Kantil Regresyon Yöntemi, sağlam (robust) bir regresyon şeklidir. Kantil Regresyon Yönteminde amaç, tahmin edicilerin hatalarının mutlak toplamını minimum etmektir. En Küçük Kareler Yönteminde hataların varyansı konusunda bazı varsayımlar vardır. En Küçük Kareler Yöntemi, aykırı değerlere (ya da uç değerlere) karşı sağlam olmadığı için yapılan tahmini etkilemektedir. Kantil regresyon analizinde uç değerler, modelin eğilimini maskeleyemez; zira araştırmacı, kantil düzeyini verideki uç değerlerin dağılımını dikkate alarak belirler. Bu nedenle, kantil regresyon yöntemi, uç değerlerin varlığına rağmen güvenilir ve tutarlı tahminler sunma kapasitesine sahiptir. Bu özellik sayesinde, Kantil Regresyon Yöntemi, Doğrusal Regresyon Yöntemine alternatif olarak geliştirilmiştir. Bu yöntem, daha sağlam bir yöntem olduğundan regresyon analizindeki bazı varsayımlara gerek duyulmamaktadır. Kantil Regresyon, doğrusal modelde bağımlı değişkenin dağılımının koşullu miktarlarını tahmin etmenin bir yoludur. Kantil Regresyonlar, veri kümelerinin koşullu dağılımındaki değişiklikleri göstermek için kullanışlıdır. Kantil Regresyon Yöntemi, özellikle uç değerlerin olduğu durumlarda kullanışlı bir yöntemdir (Çınar, 2019). Yu vd. (2013), Kantil Regresyon Yöntemi, En Küçük Kareler (EKK) yöntemine kıyasla daha ayrıntılı ve kapsamlı bir istatistiksel model sunmaktadır. Klasik regresyon modellerinin odaklandığı ortalama eğilimlerin ötesine geçerek, Kantil Regresyon modeli, bağımlı değişkenin koşullu dağılımının tamamı üzerinde analiz yapılmasına olanak tanıyarak, yalnızca ortalama etkileri değil, dağılımın farklı kantillerindeki değişken ilişkilerini de modellemeyi mümkün kılmaktadır. Literatürdeki pek çok deneysel çalışmada, basit regresyon analizleri için sıklıkla En Küçük Kareler yönteminin tercih edildiği görülmektedir. Ancak EKK yöntemi yalnızca koşullu ortalamaya odaklanmakta ve dağılımın uç noktalarındaki yapısal farklılıkları göz ardı etmektedir. Buna karşılık, Kantil Regresyon Yöntemi, verinin tüm dağılım yapısı hakkında daha kapsamlı ve açıklayıcı bilgiler sağlayarak, özellikle heterojen yapılar içeren veri setlerinde araştırmacılara güçlü bir analiz aracı sunmaktadır. (Çınar, 2019).

Kantil Regresyon, uç değerlere karşı daha sağlamdır bunun nedeni bağımlı değişken dağılımını çeyreklere ya da parçalara bölebilmemesinden kaynaklanmaktadır (Çınar, 2019). Kantil Regresyon Yöntemi, bağımlı değişkenin koşullu dağılımı boyunca değişkenlerin etkilerini analiz etme olanağı sunarak, özellikle dağılımın farklı dilimlerinde kovaryans yapısının incelenmesine imkân tanımaktadır. (Çağlayan ve Arıkan, 2011).

Kantil Regresyon Yöntemi, fonksiyonel ilişkilerin doğrudan veri setinin koşullu dağılımına bağlı olarak belirlendiği bir yaklaşım sunduğundan, aykırı değerlere karşı yüksek derecede dayanıklıdır. Bu özelliği sayesinde, yanlış fonksiyonel varsayımlardan kaynaklanabilecek model hatalarının önüne geçilmesini sağlar. Araştırmacılar, veri setindeki uç değerlere göre farklı kantil düzeylerinde (örneğin 0.05, 0.25, 0.50, 0.75, 0.95) ayrı ayrı modeller oluşturarak, her bir modelde bağımsız değişkenlerin istatistiksel anlamlılıklarını inceleyip, en uygun sonuç modelini belirleme esnekliğine sahiptir. Bu yönüyle kantil regresyon, araştırmacıya uç değerlerin etkisinden kaçınma ve daha sağlıklı tahminler yapma olanağı tanır. Farklı kantil değerlerine göre kurulan modellerde, kantil düzeyi değiştikçe modelin fonksiyonel formu da değişmekte ve bu durum, koşullu dağılımın çeşitli noktalarındaki yapısal farklılıkları ortaya koymaktadır. Ayrıca, kantil değeri 0.50 ($t=0.5$) olduğunda, Kantil Regresyon En Küçük Mutlak Sapma (LAD) yöntemi ile örtüşmekte ve bu durumda tahminler medyan değerine dayanmaktadır. (Koenker and Basett, 1978).

Kantil Regresyon Yöntemi, her bir kantil düzeyi için tahmin hatasını en aza indiren parametre tahminlerini elde etmeye odaklanır. Parametre tahminlerinde klasik regresyon yöntemlerinin aksine hata terimlerine ilişkin belirli bir dağılım varsayımına ihtiyaç duymaması, bu yöntemi özellikle dağılım yapısı bilinmeyen ya da uç değerler içeren veri setleri için oldukça kullanışlı ve esnek bir araç hâline getirmektedir. (Türkiş, 2015). Bağımlı değişkenin koşullu dağılımının tamamını modelleyerek, dağılımın her bir kantilinde açıklayıcı değişkenlerin etkilerini ayrı ayrı değerlendirme imkânı sunar. Bu yaklaşım, özellikle dağılımın asimetric kuyruk yapılarının da analize dahil edilmesini sağlayarak, geleneksel ortalama odaklı regresyon yöntemlerine kıyasla daha ayrıntılı ve kapsamlı sonuçlar elde edilmesine olanak tanır. (Abadie et al., 2002).

Kantil Regresyon modellerinde, farklı kantil seviyelerine göre regresyon katsayıları ayrı ayrı tahmin edilmekte ve bu sayede bağımlı değişkenin koşullu dağılımının çeşitli noktalarındaki değişken etkileri ortaya konmaktadır. Her bir kantil değeri, modelin fonksiyonel yapısını değiştirdiği için, özellikle dağılımın alt ve üst uçlarında farklı eğilimlerin gözlemlenmesine imkân tanımaktadır. Bu yöntem, uç değerlerin bulunduğu veya dağılımın homojen olmadığı veri setlerinde, klasik ortalama tabanlı regresyon modellerine kıyasla daha güvenilir ve esnek tahminler sunmaktadır. Kantil Regresyon analizinde, tahminlerin hesaplanmasında sıklıkla standart normal dağılımın kantil değerleri referans alınarak parametrik olmayan teknikler kullanılmaktadır. (Çınar, 2019).

Doğrusal regresyonun varsayımları sebebiyle araştırma yapan kişilerin çalışmalarında, uç değerlerin olduğu durumlarda Kantil Regresyon modellerini kullanmalarının faydalı olacağı ve bazı avantajları getireceği belirtilmiştir. Kantil Regresyon avantajları;

- Her bir kantil düzeyinde, bağımlı değişken ile tahmin edilen değerler arasındaki ilişki açık bir biçimde ortaya konulabilir. Bu sayede model, dağılımın farklı bölümlerinde değişkenler arası ilişkiyi ayrıntılı olarak yansıtır.
- Klasik regresyon yöntemlerinin aksine, Kantil Regresyon daha az sayıda istatistiksel varsayıma dayanır; özellikle hata terimlerinin normal dağıldığı varsayımına ihtiyaç duymaz, bu da yöntemi daha esnek ve geniş veri koşullarında uygulanabilir kılar.
- Kantil Regresyon tahmin edicileri, büyük örneklerde tutarlılık ve asimptotik normallik gibi istatistiksel özelliklere sahiptir; bu da yöntemin güvenilirliğini artırır ve parametrik çıkarımlar için uygunluk sağlar (Wang and Cook, 2017).

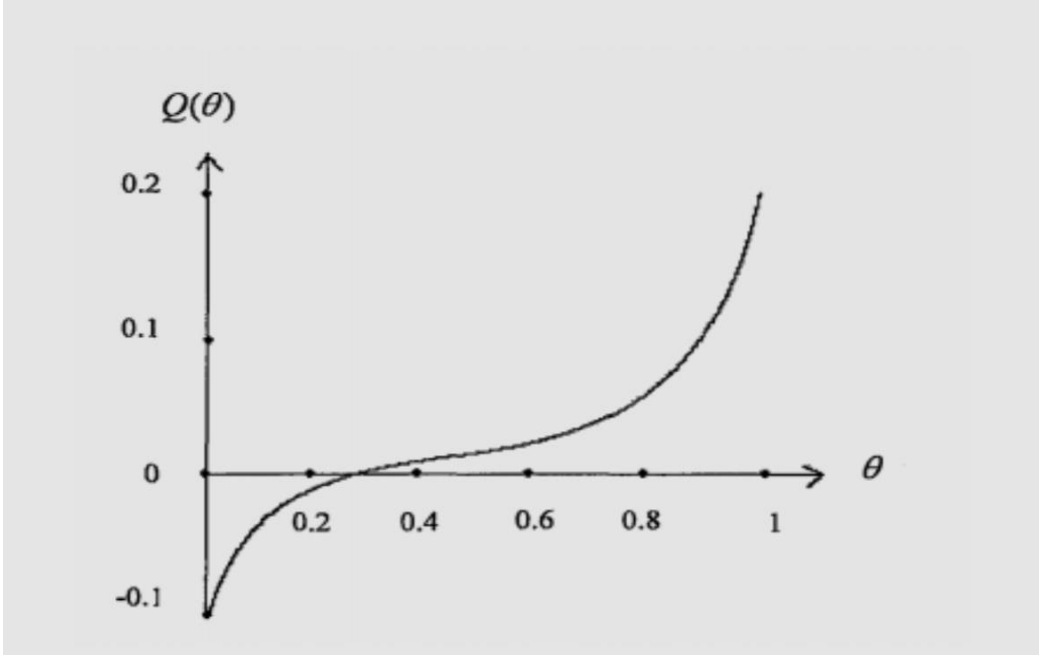
Chen ve Wei (2005), Kantil Regresyon Yönteminin özellikle koşullu kantillerin değişkenlik gösterdiği durumlarda kullanışlı olduğunu belirtmişlerdir (Yavuz ve Aşık, 2017). Kantillere bağlı olarak regresyon katsayıları belirlenmektedir. Kantil Regresyon, klasik regresyon analizinde yer alan hata terimlerinin normal dağılım göstermesi varsayımını gerektirmeyen, dolayısıyla daha sağlam sonuçlar üretebilen bir regresyon tekniği olarak geliştirilmiştir.

Klasik EKK Yöntemi, koşullu ortalamaya bağlı olurken; Kantil Regresyon Yöntemi koşullu kantil fonksiyonuna bağlıdır (Koenker, 2005). Klasik Regresyon Yöntemlerinin aksine hataların dağılımı ve hataların varyanslarının homojenliği varsayımı olmadığından, daha az varsayıma sahiptir. Bu regresyon modelleri uç değerlere ve eğikliğe EKK Yönteminden daha sağlamdır (Saçaklı, 2005).

F dağılım fonksiyonuna sahip Y bağımlı değişkenin, $0 \leq t \leq 1$ için rastgele örnekleme $t \cdot k$ boyutlu tasarım matrisi, b tahmin edilecek katsayı vektörü ve $e_t = y_i - x_i \beta$ hata değeri, t regresyon kantil değeri olmak üzere kurulan

$$\min_{\beta \in R^k} \left[\sum_{t \in \{t: y_t \geq x_t b\}} \tau |y_t - x_t b| + \sum_{t \in \{t: y_t < x_t b\}} (1 - \tau) |y_t - x_t b| \right] \quad (2.2)$$

eşitliği minimum olacak şekilde elde edilmektedir (Koenker and Bassett, 1978).



Şekil 5: Kantil Regresyon Eğrisi (Çınar, 2019)

Uç değerler içeren veri setlerinde, Kantil Regresyon Yöntemi uygulanarak regresyon doğrusu, tahmin hatasını en aza indirecek biçimde araştırmacı tarafından esnek bir şekilde uyarlanabilir. Modelleme sürecinde, Şekil 5'te gösterildiği üzere 0.1'den 0.9'a kadar farklı kantil düzeyleri (örn. 0.1; 0.2; 0.3; ...; 0.9) esas alınarak kantil regresyon eğrileri oluşturulur. Araştırmacı, veri setindeki uç değerlerin dağılımına göre uygun kantil değerini seçerek, her bir kantil için bağımsız regresyon modelleri kurar. Bu modellerin tahmin performansları karşılaştırılarak, en düşük hata ölçütüne sahip olan kantil modeli nihai sonuç modeli olarak tercih edilir. Bu yönüyle kantil regresyon, verideki heterojen yapıları ve uç değer etkilerini dikkate alan güçlü bir modelleme yaklaşımı sunmaktadır (Çınar, 2019).

Değişen varyanslılık (heteroscedasticity) sorunu, çoğunlukla modelde yer alması gereken önemli bir değişkenin dışlanması durumunda ya da veri setinde uç değerlerin bulunması halinde ortaya çıkmaktadır. Bu durum, klasik regresyon analizlerinde kullanılan t ve F testlerinin güvenilirliğini azaltarak istatistiksel anlamlılıklarının sorgulanmasına neden olur. Aynı zamanda tahmin edicilerin 'etkinlik' özelliği zayıflar; yani en düşük varyansa sahip olma koşulu bozulur, varyanslar büyür ve tahmin hataları artar. Özellikle makroekonomik ve mikroekonomik göstergeler, finansal tablolar, ekonometrik veriler ve zaman serileri gibi ekonomik veri setlerinde uç değerlerin bulunması oldukça yaygındır. Bazı veri setlerinde, sabit

varyans varsayımı (homoskedastisite) ile hata terimlerinin normal dağılıma sahip olması gibi klasik regresyon analizinin temel varsayımları sıklıkla ihlal edilmektedir. Bu tür durumlarda, En Küçük Kareler yöntemi hem teorik geçerliliğini hem de tahmin güvenilirliğini kaybedebilmektedir. Bu bağlamda, özellikle uç değerlere karşı dayanıklı ve dağılım varsayımlarına duyarlı yapısı ile öne çıkan Kantil Regresyon Yöntemi literatürde güçlü bir alternatif olarak önerilmektedir. Kantil Regresyon, değişen varyanslılık problemine karşı gösterdiği sağlamlık ve hata terimlerinin belirli bir dağılıma uymasını gerektirmemesi nedeniyle, karmaşık ve yapısal olarak heterojen veri setlerinde güvenilir ve esnek bir modelleme aracı sunmaktadır (Çınar, 2019).

2.2.1 Kantil Regresyonun Özellikleri

Kantil Regresyon Yöntemlerinin özelliklerinden dolayı bazı durumlarda kullanılması tercih edilmektedir.

- Kantiller, bağımlı değişkendeki aykırı değerlere karşı daha sağlamdır (Robust özelliği).
- Hata terimleri normal dağılmadığında, Kantil Regresyon tahmin edicileri EKK tahmin edicilerinden çok daha etkin olabilir (Ocak, 2010).
- Kantil Regresyon Yöntemi, veri setinde mevcut olan değişen varyanslılık yapısının belirlenmesine olanak tanıyarak, varyansın bağımlı değişken üzerindeki etkilerini daha ayrıntılı biçimde analiz etmeyi mümkün kılar.
- Farklı kantil düzeylerinde elde edilen farklı regresyon sonuçları, bağımlı değişkenin koşullu dağılımının çeşitli noktalarında açıklayıcı değişkenlerin etkisinin değişkenlik gösterdiğini ortaya koyar. Bu durum, modelin değişkenler arasındaki heterojen ilişkileri yansıtılma yeteneğinin bir göstergesidir (Ocak, 2010).
- Kantil Regresyon analizinde kantil değeri 0.5 olarak belirlendiğinde, model En Küçük Mutlak Sapma (Least Absolute Deviations – LAD) regresyonuna karşılık gelir. Bu durumda, tahmin süreci medyan değerine odaklanmakta ve mutlak hata terimlerinin toplamını minimize etmeye çalışmaktadır (Altındağ, 2010).
- Kantil Regresyon, farklı kantil düzeyleri için bağımlı değişkenin (y) bağımsız değişkenlere (x) koşullu dağılımını modelleyerek, dağılımın tamamı hakkında kapsamlı bilgi sunar. Bu sayede yalnızca ortalama etki değil, dağılımın alt ve üst uçlarındaki yapısal farklılıklar da analiz edilebilir (Kurtoğlu, 2011).
- En Küçük Kareler regresyon modelinde, hata terimlerinin sabit varyansa sahip olduğu varsayılırken; Kantil Regresyon modelinde bu tür bir varsayım bulunmamaktadır.

Kantil Regresyon, hata terimlerinin deęişen varyanslılık yapısına sahip olmasına izin vererek, daha esnek ve veri yapısına duyarlı bir modelleme imkânı sunmaktadır (Yavuz ve Aşık, 2017).

- Kantil Regresyonun amaç fonksiyonu doęrultusunda elde edilen katsayı tahminleri, bağımlı deęişkenin aşırı (uç) deęerlerine karşı duyarlı deęildir. Bu yönüyle, tahmin edilen katsayılar konum istatistiklerinin sağlam (robust) ölçüleri olarak kabul edilmekte ve klasik regresyon yöntemlerinin uç deęerlerden etkilenme sorununu büyük ölçüde ortadan kaldırmaktadır (Saçaklı, 2005).

2.2.2. Kantil Regresyonun Temel Denklemi

Koenker (2005)'e göre, Kantil Regresyon aşağıdaki optimizasyon problemine dayanmaktadır.

$$\hat{\beta}(\tau) = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^n \rho_{\tau}(y_i - x_i^T \beta)$$

Burada;

y_i : Bağımlı deęişken

x_i^T : Bağımsız deęişken

β : Regresyon katsayısı

τ : İncelenen kantil regresyon deęeri ($\tau=0.5$ medyan için)

$\rho_{\tau}(u)$: Kantil hata fonksiyonu aşağıdaki şekilde tanımlanmaktadır;

$$\rho_{\tau}(u) = \begin{cases} \tau * u & \text{eđer } u \geq 0 \\ (\tau - 1) * u & \text{eđer } u < 0 \end{cases}$$

Kantil Fonksiyonunda $\rho_{\tau}(u)$ terimi, hataların ($u=y_i - x_i^T \beta$) ağırlıklı toplamını minimize etmektedir. Bu fonksiyonda pozitif deęerli hatalar, τ ile çarpılırken, negatif deęerli hatalar ise $(\tau - 1)$ ile çarpılmaktadır (Koenker and Bassett, 1978).

2.2.3. Kantil Regresyonun Varsayımları

1. Kantil Regresyon, $Q(\beta) = \sum_{i=1}^n \rho_{\tau}(y_i - x_i^T \beta)$ amaç fonksiyonunu minimize etmektedir.
2. $\rho_{\tau}(u)$, $u=0$ noktasında süreklidir.

3. $\rho_{\tau}(u)$, u 'nun bir fonksiyonu olarak konveks, yani dış bükeydir.
4. $\rho_{\tau}(u)$ fonksiyonunun dış bükey ve sürekli olması durumu, bir çözümün olduğunun kanıtıdır. Bu çözüm genellikle doğrusal programlama ve iteratif yöntemlerle bulunmaktadır.
5. Kantil Regresyon dış bükey bir optimizasyon problemi olmasından kaynaklı olarak global bir minimum değere sahiptir.
6. Kantil Regresyon, aykırı değerlere karşı daha sağlam bir yöntemdir. Geleneksel regresyonun aksine, dağılımın farklı noktalarını (kantillerini) modelleyebilmektedir.

Ampirik çalışmalarda, sahip olduğu kolaylıklar ve analitik avantajlar nedeniyle En Küçük Kareler Yöntemi genellikle ilk tercih edilen modelleme tekniği olmaktadır. Ancak, bu yöntemin uygulanabilirliği çok sayıda varsayıma bağlıdır. Söz konusu varsayımlardan herhangi birinin ihlal edilmesi durumunda, elde edilen tahminlerin güvenilirliği ciddi biçimde zedelenebilir. Özellikle mikro düzeydeki veri setleriyle gerçekleştirilen analizlerde, hata terimlerinin normal dağılım göstermesi ve sabit varyansa sahip olması gibi temel varsayımlar çoğunlukla sağlanmamaktadır. Bu tür durumlarda, EKK Yönteminin geçerliliğini kaybetmesi nedeniyle, daha sağlam ve esnek alternatif modelleme yaklaşımlarının tercih edilmesi zorunlu hâle gelmektedir (Köseoğulları ve Ulucan, 2022; Çiftçi ve Kangallı, 2015).

Koenker ve Bassett (1978), sabit varyanslılık varsayımının ihlal edildiği durumlarda güvenilir tahminler elde edilebilmesini sağlamak amacıyla Kantil Regresyon Yöntemi'ni geliştirmiştir. En Küçük Kareler Yöntemi, uygulamada bağımlı değişkenin yalnızca ortalama (merkezî eğilim) değerini modellediği ve değişkenler arası ilişkinin sadece bu noktada incelendiği bir yaklaşımdır. Buna karşın, Kantil Regresyon Yöntemi, dağılımın farklı kantil düzeylerine odaklanarak bağımlı değişkenin koşullu dağılımı boyunca daha geniş bir analiz imkânı sunar. Bu yöntem, değişkenlerin tüm dağılımı boyunca eşit derecede ağırlıklandırılmasını sağladığı için sadece merkezî değerleri değil, aynı zamanda uç değerleri ve dağılımın yapısal özelliklerini de dikkate alır. Böylece, değişen varyanslılık sorununu da doğrudan ele alarak klasik regresyon analizlerine kıyasla daha esnek ve bilgi zenginliği yüksek bir çerçeve sunar. Bu bağlamda, Kantil Regresyon Yöntemi, EKK'ya kıyasla istatistiksel açıdan daha avantajlı bir modelleme alternatifi olarak öne çıkmaktadır (Köseoğulları ve Ulucan, 2022; Çiftçi ve Kangallı, 2015; Montenegro, 2001; Hunter and Kenneth, 2000).

Kantil Regresyon Yöntemi uygulanırken kurulan model

$$Z_i = X_i\beta_q + \varepsilon_{qi} \quad \text{Kantil}_q(Z_i|X_i) = X_i\beta_q \quad (2.3)$$

biçimindedir. Eşitlik (2.3)'te yer alan Z_i bağımlı değişken, X_i bağımsız değişkendir. $\text{Kantil}_q(Z_i|X_i)$, X bağımsız değişkenleri bilindiğinde bağımlı olan Z değişkenine q . koşullu kantil fonksiyonunu ifade etmektedir. Modelde yer alan q , 0 ile 1 arasında değerler almaktadır. Böylelikle, modelin farklı noktalarındaki değerleri hakkında bilgi sahibi olunabilmektedir (Köseoğulları ve Ulucan, 2022; Ulucan ve Çiftçi, 2019).

Eşitlik (2.4)'e, Eşitlik (2.3)'ün çözümlenmesi için gerekli olan minimizasyon eşitliği verilmiştir (Cameron and Triverdi, 2009):

$$Q(\beta_q) = \sum_{i=y_i > x'_i\beta} q |y_i - x'_i\beta_q| + \sum_{i=y_i < x'_i\beta} (1 - q) |y_i - x'_i\beta_q| \quad (2.4)$$

Eşitlik (2.4), Kantil Regresyon Yöntemi ile yapılan tahminlerin EKK gibi simetrik değil, asimetrik olduğunu göstermektedir. Burada, q 'nun değeri azaldıkça ağırlık sağ tarafa kayarken, q 'nun değeri arttıkça ağırlık sol tarafa kaymaktadır (Ulucan ve Çiftçi, 2019; Cameron and Triverdi, 2009).

Eşitlik (2.5)'te, oluşturulan son model görülmektedir:

$$Q_{n,t} = \beta_0 + X'_i\beta_1(t) + u_i(t) \quad (2.5)$$

Burada, X'_i modeldeki açıklayıcı değişkenleri, t ise farklı kantil derecelerini (0,25,0,50,0,75,0,90 gibi) temsil etmektedir (Köseoğulları ve Tunca, 2024).

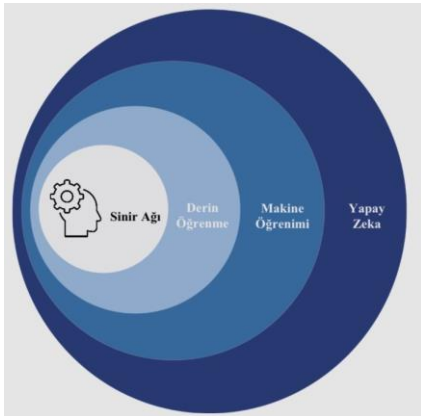
2.3. Makine Öğrenmesi Yöntemleri

Gelişen veri teknolojileri sayesinde büyük ve karmaşık veri kümelerinden anlamlı bilgilere ulaşmak günümüzde daha erişilebilir hâle gelmiştir. Bu doğrultuda, yapay zekâ, makine öğrenmesi ve derin öğrenme gibi ileri analiz yöntemleri kullanılarak söz konusu veriler üzerinden geleceğe yönelik tahminlerde bulunmak mümkün hâle gelmiş ve bu alanlarda yüksek doğruluk düzeyine sahip sonuçlar elde edilmiştir. Ancak, literatürde çok sayıda yöntem, algoritma ve model yapılandırmasının bulunması, araştırmacıların hangi yöntemi seçeceği, hangi model mimarisini kullanacağı ve hiperparametrelerin nasıl belirleneceği konularında çeşitli zorluklarla karşılaşmasına neden olmaktadır. Özellikle model ve hiperparametre seçimi sürecinde, farklı değerlerin sistematik biçimde karşılaştırılması, hem zaman alıcı hem de yüksek hesaplama maliyetine sahip olabilen bir süreçtir. Bu durum, modelleme aşamasında otomasyon, optimizasyon teknikleri (örneğin grid search, random search, Bayesian

optimizasyon) ve çapraz doğrulama gibi yöntemlerin önemini artırmaktadır (Akkurt ve Hasgöl, 2024).

Makine öğrenmesi (MÖ) yöntemleri, günümüzde hızla artan veri hacmi karşısında, bu verilerden anlamlı çıkarımlar yapılmasını mümkün kılan güçlü analiz araçları olarak öne çıkmaktadır. Özellikle büyük veri setlerine uygulandığında, geçmiş gözlemlerden yola çıkarak geleceğe yönelik tahminler yapılmasına olanak tanır. Veri yapılarının giderek daha karmaşık ve yüksek boyutlu hâle gelmesi, klasik analiz yaklaşımlarının yetersiz kalmasına neden olmakta; bu durum, insan yorumunun sınırlarını aşmaktadır. Bu noktada, makine öğrenmesi algoritmaları devreye girerek, karmaşık veri örüntülerini tanımlamakta ve ileriye dönük doğru, hızlı ve uygulanabilir öngörülerde bulunulmasını sağlamaktadır (Türkmenoğlu ve Tantuğ, 2014).

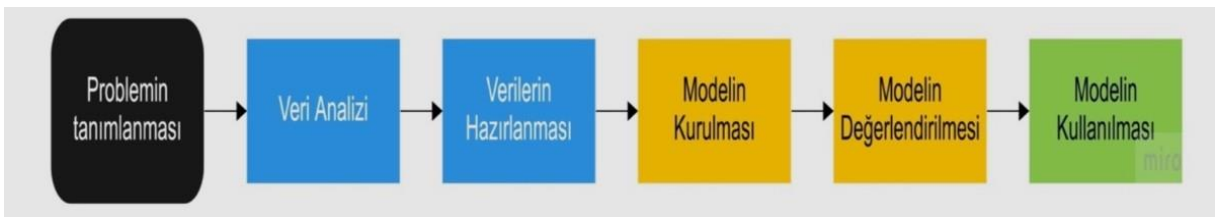
Makine öğrenmesi, yapay zekâ (YZ) alanının bir alt disiplini olarak konumlanmakta olup, derin öğrenme ise makine öğrenmesinin daha özelleşmiş bir alt kümesini oluşturmaktadır. Hiyerarşik bir yaklaşımla ifade edildiğinde, kavramsal sıralama Şekil 6'da gösterildiği üzere en alt düzeyde sinir ağları, ardından derin öğrenme, makine öğrenmesi ve en üstte yapay zekâ şeklinde yapılandırılabilir. Makine öğrenmesi algoritmaları, büyük veri kümeleri içerisindeki örüntüleri, ilişkileri ve korelasyonları tanımlayarak, bu analizlere dayalı olarak tahminlerde bulunmak ve optimal karar mekanizmaları geliştirmek amacıyla eğitilmektedir. Bu süreçte, öğrenme algoritmaları veri içindeki gizli yapıları keşfetmekte ve elde edilen modeller yeni gözlemler üzerinde genelleştirme yapabilme yetisine sahip olmaktadır (SAP, 2021). Makine öğrenimi üzerine yapılan araştırmaların artmasıyla birlikte hesaplamalı öğrenme teorisi, yapay sinir ağları, istatistik ve örüntü tanıma gibi araştırma alanları arasında bağlantı kurulmuş ve bu alanlarla birlikte çalışılabilmiştir. Böylece, makine öğrenmesi teknikleri daha geleneksel problemlerin yanı sıra veri tabanlarında bilgi keşfi, dil işleme ve robot kontrolü gibi yeni problemlere uygulanmaya başlanmıştır (Dietterich, 1997).



Şekil 6: Makine Öğrenimi Şeması (Gökalp, 2022)

Makine öğrenimi kavramı, ilk kez 1959 yılında Arthur Lee Samuel tarafından ortaya atılmıştır. IBM için geliştirilen bir dama oyunu programı üzerinde çalışırken, Samuel bu yaklaşımı 'bilgisayarların açıkça programlanmaksızın öğrenmesini sağlayan bir araştırma alanı' olarak tanımlamıştır. Bu tanım, makine öğreniminin temel amacını, sistemlerin geçmiş deneyimlerden öğrenerek gelecekteki görevlerde performanslarını artırmaları şeklinde özetlemekte ve bu alanın yapay zekâ içindeki konumunu vurgulamaktadır (Mockenhaupt, 2021). Ayodele (2010)'ye göre ise makine öğrenmesi, otomatik olarak öğrenme işlemini deneyimlerden yola çıkarak geliştiren ve gerçekleştiren bilgisayar sistemlerinin geliştirilmesidir.

Makine öğrenmesi kavramının temelleri, 1962 yılında dama uzmanı Robert Nealey'nin IBM tarafından geliştirilen IBM 7094 bilgisayarıyla yaptığı bir karşılaşmada atılmıştır. Nealey, sınırlı bellek kapasitesine sahip bu bilgisayara karşı oyunu kaybetmiştir. Söz konusu bilgisayar programı, dama tahtasındaki taşların olası konumları için bir puanlama sistemi geliştirmiş ve bu puanlamalara dayanarak bir sonraki hamleleri 'minimax algoritması' ile belirleyen bir strateji izlemiştir. Bu olayın ardından Arthur Samuel, programın oyun stratejisini daha da geliştirmek amacıyla çeşitli mekanizmalar tasarlamış ve bu süreçte 'ezberci öğrenme' olarak adlandırdığı bir öğrenme yöntemi geliştirmiştir. Bu yöntemde, program tüm oyun pozisyonlarını hafızasında saklayarak, her bir pozisyona karşılık gelen ödül fonksiyonlarını dikkate alarak gelecekteki hamleleri seçmiştir. Samuel'in bu öncü çalışmaları, 'Makine Öğrenmesi' (Machine Learning) teriminin doğmasına ve bu alanda yeni bir bilimsel yaklaşımın şekillenmesine zemin hazırlamıştır. (Foote, 2019). Makine öğreniminin çalışma prensibi Şekil 7'de verilmiştir.



Şekil 7: Makine Öğreniminin Çalışma Prensibi (Gökalp, 2022)

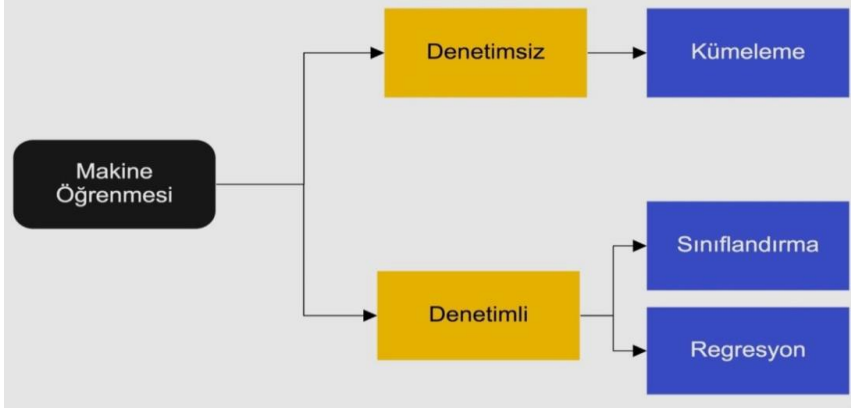
Klasik makine öğrenmesi algoritmaları, büyük ölçüde öznitelik (feature) çıkarımına dayanmaktadır. Bu bağlamda, probleme özgü anlamlı özniteliklerin belirlenmesi genellikle ilgili alandaki uzman bilgisiyle mümkün olmaktadır. Öznitelik mühendisliği sürecinde, yalnızca özniteliklerin oluşturulması değil, aynı zamanda bu öznitelikler arasından en anlamlı olanlarının seçilmesi, modelin eğitimi ve genel performansının artırılması açısından kritik öneme sahiptir. Ancak, çok sayıda öznitelik üretildiğinde, 'boyutun laneti' (curse of

dimensionality) olarak bilinen problem ortaya çıkmaktadır. Bu durum, öznelik sayısındaki artışla birlikte, modelin doğru şekilde öğrenebilmesi için gereken eğitim verisi miktarının üstel olarak artmasına neden olur ve bu da öğrenme sürecini hem hesaplama açısından zorlaştırır hem de modelin genelleme kapasitesini olumsuz etkileyebilir (Verleysen and François, 2005). Eğitim verisinin yetersiz olduğu, buna karşın öznelik sayısının fazla olduğu durumlarda, makine öğrenmesi algoritmaları veri setini genellemek yerine ezberleme eğilimi gösterir; bu durum literatürde 'aşırı öğrenme' (overfitting) olarak tanımlanmaktadır. Aşırı öğrenme, modelin eğitim verisine yüksek uyum göstermesine rağmen, yeni veriler üzerindeki tahmin performansının düşmesine neden olur. Bu durumda, sınıflandırma veya regresyon gibi görevlerde modelin başarısı önemli ölçüde azalır. Klasik makine öğrenmesi algoritmaları bu tür senaryolarda sınırlı kalabilmekte, öznelik çıkarımı ve seçimi genellikle manuel olarak yapılmaktadır. Bu bağlamda, öznelik mühendisliği süreçlerini otomatikleştirme yeteneğine sahip olan derin öğrenme yöntemleri, özellikle karmaşık veri yapılarıyla çalışırken önemli avantajlar sunmaktadır. Ancak, derin öğrenme modellerinin yüksek parametre sayısı nedeniyle etkili şekilde öğrenebilmesi için geniş ve dengeli bir veri setine ihtiyaç duyulmaktadır (Demirezen vd., 2021).

2.3.1 Makine Öğrenimi Türleri

Makine öğrenimi yöntemleri, literatürde genel olarak Şekil 8'de gösterildiği üzere iki temel kategori altında sınıflandırılmaktadır: denetimli öğrenme (supervised learning) ve denetimsiz öğrenme (unsupervised learning). Bu temel ayrım, modelin öğrenme sürecinde hedef değişkenin (etiketli veri) bulunup bulunmamasına dayanmaktadır. Zamanla, bu iki ana kategoriye ek olarak pekiştirmeli öğrenme (reinforcement learning), yarı denetimli öğrenme (semi-supervised learning) ve kendi kendine öğrenme (self-supervised learning) gibi farklı öğrenme paradigmaları da geliştirilmiş ve literatürde yerini almıştır. (Koyuncugil ve Özgülbaş, 2009). Makine öğrenmesinde veri, öğrenme sürecinin temelini oluşturan en kritik unsurlardan biridir. Öğrenme algoritmaları, veri setinde yer alan örüntüleri, ilişkileri ve öznelik yapılarını keşfetmek ve bu bilgilere dayanarak genelleme modeller oluşturmak amacıyla kullanılmaktadır. Bu bağlamda, modelin başarısı doğrudan veri kalitesi, temsil gücü ve öznelik zenginliği ile ilişkilidir (Chao, 2011). Makine öğrenmesi uygulamalarında kullanılan veriler, genel olarak etiketli (labeled) ve etiketsiz (unlabeled) veri olmak üzere iki temel kategoriye ayrılmaktadır. Etiketli veri, denetimli öğrenme algoritmalarını eğitmek amacıyla kullanılırken, etiketsiz veri genellikle önceden eğitilmiş modellerin test edilmesi veya desen keşfi gibi görevlerde kullanılmaktadır. Bu bağlamda, etiketli veri çoğunlukla eğitim seti

(training set), etiketsiz veri ise test seti (test set) veya doğrulama seti (validation set) olarak da adlandırılmaktadır. Veri türlerinin doğru biçimde sınıflandırılması ve kullanımı, modelin genel performansını ve genelleme kapasitesini doğrudan etkilemektedir (Gentleman et al., 2008).



Şekil 8: Makine Öğrenimi Grupları (Gökalp, 2022)

2.3.1.1. Denetimli Öğrenme (SML: Supervised Machine Learning)

Makine öğreniminde en çok kullanılan ve en popüler yöntemdir. Bu yöntemin amacı, sistemin etiketli verileri kullanarak eğitilmesi ile öğrenmenin sağlanmasıdır. Test veri seti, sistemin doğrulanması amacıyla kullanılır. Sistemin doğrulanması aşamasında öğrenme algoritması kategorisi bilinmeyen bir test verisine, eğitim verisinde bulunan çıkışlardan herhangi birini atar (Kotsiantis et al., 2007). Etiketli veriler, makine öğrenmesi sistemlerinin eğitimi sürecinde temel bileşen olarak kullanılır. Her bir giriş verisi (eğitim etiketi) için, sistem tarafından karşılık gelen bir çıktı etiketi öğrenilir. Bu süreçte model, giriş-çıkış çiftleri arasındaki ilişkiyi analiz ederek örüntüleri tanımayı ve genelleme yapmayı öğrenir. Sistem, bu sayede belirli görevlerde (örneğin sınıflandırma veya regresyon) doğru tahminlerde bulunabilecek şekilde eğitilmiş olur. Eğitim süreci genellikle manuel olarak tanımlanmış ve doğruluğu onaylanmış etiketlerle desteklenir; bu da modelin yüksek doğrulukla öğrenmesini ve gerçek dünya verileri üzerinde uygulanabilir sonuçlar üretmesini sağlar (Çoban, 2016). Tahminler yapılırken, regresyon ya da sınıflandırma yöntemi kullanılmaktadır (Gökalp, 2022). Destek Vektör Makinesi, Yapay Sinir Ağları, Lojistik Regresyon, Basit Bayes, Multinom Basit Bayes, k-En Yakın Komşu, Rastgele Orman ve Karar Ağaçları algoritmaları denetimli öğrenme modeli oluşturulurken sıklıkla kullanılan yöntemlerdir (Caruana and Niculescu-Mizil, 2006).

Daha önce yapılan birkaç araştırma, doğrusal olmayan makine öğrenmesi yaklaşımlarının, klasik zaman serileri tekniklerine kıyasla daha etkin tahmin edicilere sahip olduğunu belirtmektedir (Li, Wang and Li, 2015; Menezes and Barreto, 2008; Reddy and Prasad, 2018; Udelhoven, 2012).

2.3.1.2. Denetimsiz Öğrenme (UML: Unsupervised Machine Learning)

Bu modelde sistem eğitilirken etiketsiz veriler kullanılarak modelin öğrenmesi sağlanmaktadır. Denetimsiz öğrenmede amaç, veri setindeki örneklerin çıkışları bilinmediği için tanıma veya sınıflandırma değildir. Genellikle kümeleme, olasılık yoğunluk tahmini, öznelikler arasındaki ilişkilerin bulunması ve boyut indirgeme gibi amaçlarla kullanılmaktadır. Ayrıca, denetimsiz öğrenme algoritması ile elde edilen sonuçlar denetimli öğrenme için de kullanılabilir (Chao, 2011). Parçalayıcı ve hiyerarşik kümeleme algoritmaları ise genellikle denetimsiz öğrenme modeli oluşturulurken kullanılan algoritmalar (Özgür, 2004).

2.3.1.3. Yarı Denetimli Öğrenme (Semi-supervised Learning)

Yarı denetimli öğrenme (semi-supervised learning), denetimli ve denetimsiz öğrenme yaklaşımlarının bir birleşimi olarak yapılandırılmıştır. Bu yöntemde, veri setinin bir bölümü etiketlenmiş iken, geri kalan kısmı etiketsizdir. Yarı denetimli öğrenmenin temel amacı, sınırlı sayıda etiketli veriyle birlikte geniş etiketsiz veri kümesinin birlikte kullanılması yoluyla modelin öğrenme performansını artırmaktır. Literatürde yapılan çalışmalar, etiketsiz verilerin uygun biçimde değerlendirilmesi durumunda, yalnızca etiketli veriye dayalı eğitime kıyasla daha yüksek doğruluk ve genelleme kapasitesi elde edilebileceğini ortaya koymaktadır (Sarıcaoğlu, 2019).

2.3.2. Makine Öğrenimli Yazılımlar

- TensorFlow
- Google Cloud ML Motoru
- Amazon Makine Öğrenimi (AML)
- MonkeyLearn
- BigML
- IBM Watson (Gökalp, 2022).

2.4. Zaman Serileri Analizi

Veriye dayalı karar almanın en önemli parçalarından bir tanesi tahmin sürecidir. Farklı kurumlar, hedefleri doğrultusunda tutarlı kararlar alabilmek için tahmin politikası belirlemeye çalışmaktadır. Tutarlı tahminler, kurum ve kuruluşlar için başarının bir anahtarı durumundadır ve her kurumun varlığını başarılı bir şekilde sürdürebilmesi için gelecekle ilgili tahminlere ihtiyaç duymaktadır. Zaman serisi analizi, literatürde gelecek tahmini için kullanılan bir yöntem olarak kullanılmaktadır. Zaman serisi analizi yöntemlerini tahmin amacıyla kullanan araştırmacılar, tahmin edilen olayların gelecekte de geçmiştekine benzer şekilde gerçekleşeceğini varsaymaktadır (Can, 2009).

Zaman serisi, zamanla ölçülen gözlem değerlerinin yer aldığı serilerdir. Eşit zaman aralıklarında, bir değişkene ait gözlemlerin oluşturduğu seriye zaman serisi denmektedir. Zaman serisi analizi ise, zaman serileri arasındaki anlamlı istatistiksel ilişkilerin çıkartılması için yapılan analizlerdir.

Zaman serisi analizi, mevcut ve geçmişteki verilerin değişim trendlerini göz önünde bulundurarak ileriye yönelik tahminler yapan sistemdir (Sapankevych and Sankar, 2009). Zaman serisi analizi yöntemleri özellikle bilgisayar teknolojisindeki hesaplama yeteneklerinin artması ile birlikte önemli ilerleme kaydetmiştir. Zaman serisi analizi, ardışık olarak elde edilen zamansal verilerin kullanıldığı bir analiz yöntemidir (Box et al., 2015). Bu analiz yönteminde veri kaynakları kronolojik bir düzene göre dizilmektedir. Bu kronoloji bozulmadan makine öğrenmesi ve derin öğrenme algoritmaları kullanılmaktadır. Brownlee (2020) zaman serisinin gözlemler arasına açık bir düzen bağımlılığı eklediğini ve bunun bir zaman boyutu olduğunu; bunu ek boyutun hem bir kısıtlama hem de ek bilgi kaynağı sağlayan bir yapı olduğunu belirtmiştir.

Zaman serisi analizi daha çok doğrusal veriler için çözümler üreten Autoregressive Moving Average (ARMA) modeline verilerin durağan olmayan zamansal farklılıklarını da dikkate alan entegrasyon teriminin eklenmesi ile geliştirilen Autoregressive Integrated Moving Average (ARIMA) (Makridakis and Hibon, 1997) modeli ile gelişmeye başlamıştır. Gerçek hayat verilerinin doğrusal olmama durumları ve makine öğrenmesinin gelişmesi ile birlikte zaman serisi tahmini için birçok farklı makine öğrenmesi algoritması gelişmiştir (Ahmed et al., 2010). Derin öğrenme yöntemleri, her probleme özgü olarak yapılandırılabilen esnek ağ mimarilerine sahip olmaları ve veri kümesi içerisindeki anlamsal ilişkileri otomatik olarak öğrenebilen öznelik çıkarma yetenekleri sayesinde, zaman serisi verilerinin analizinde etkili çözümler sunmaktadır. Özellikle yapay sinir ağlarının (ANN) gelişmiş formları olan

Tekrarlayan Sinir Ağları (RNN), Uzun Kısa Süreli Bellek (LSTM) ağları ve Gated Recurrent Unit (GRU) gibi yapılar, zaman bağımlı verilerin dinamik yapısını modellemek için sıklıkla tercih edilmektedir (Lim and Zohren, 2021).

Zaman serisi analizi, antik çağlardan bu yana çeşitli alanlarda kullanılan köklü bir yöntem olup, günümüzde yatırımcılar, işletmeler ve politika yapımcılar için stratejik öneme sahip bir analiz aracıdır. Matematik, astronomi, demografi ve istatistik gibi birçok disiplin, zaman serisi analizinin hem gelişimine katkı sağlamış hem de bu tekniklerden yoğun biçimde faydalanmıştır. Zaman serisi analizi, zaman boyutuna göre sıralanmış veriler üzerinde matematiksel ve istatistiksel yöntemler uygulanarak örüntülerin, eğilimlerin ve anlamlı içgörülerin keşfedilmesini amaçlar. Zaman serisi tahminleri ise, belirli bir hedef değişkenin geçmiş ve mevcut gözlemlerine dayanarak gelecekteki değerlerinin öngörülmesi sürecini ifade eder. Günümüzde, makine öğrenmesi ve derin öğrenme tekniklerinin gelişimi, geleneksel ekonometrik modellerle yapılan zaman serisi analizlerini hem desteklemekte hem de daha ileri düzeye taşımaktadır. Geleneksel yöntemler arasında Otoregresyon (AR), Hareketli Ortalama (MA), Üstel Düzleştirme (Exponential Smoothing - ES) ve Otoregresif Bütünleşik Hareketli Ortalama (ARIMA) modelleri öne çıkarken; son yıllarda Yapay Sinir Ağları (ANN), Rastgele Ormanlar (RF), Destek Vektör Makineleri (SVM), K-En Yakın Komşular (KNN), AdaBoost ve CatBoost gibi makine öğrenmesi algoritmalarının da zaman serisi tahminlerinde etkin biçimde kullanıldığı görülmektedir. Bu gelişmeler, zaman serisi verilerinden geçmiş kalıpları anlamayı ve geleceğe dair daha isabetli öngörülerde bulunmayı mümkün kılmaktadır (Özdemir ve Çılgın, 2022)

Zaman serisi analizinde önemli kavramlardan birisi durağanlık (stationarity) kavramıdır. Çünkü, analizlerde kullanılan birçok yöntem ancak serinin durağan olması durumunda uygulanabilmektedir. Durağanlık, ortalamada sistematik değişme olmaması, varyansta sistematik değişme olmaması ve periyodik değişmelerin ortadan kalkmış olması şeklinde tanımlanabilmektedir (Akgül,1994). Durağanlık testi için kullanabilecek yöntemler bulunmaktadır.

- Zaman serisi grafiğini incelemek,
- Otokorelasyon Fonksiyonu (Autocorrelation Function-ACF) ve Kısmi Otokorelasyon Fonksiyonu (Partial Autocorrelation Function-PCF) grafiklerini incelemek,
- Birim kök (Dickey-Fuller) Testi yapmak

Durağanlık testlerinin gerçekleşmesi sonucu durağan olmayan bir seriyi durağanlaştırmak için fark alma yöntemi gerçekleştirilir. Durağan olmayıp, fark alma işlemi gerçekleştirildikten sonra durağan hale getirilen serilere uygulanan süreçler otoregresif entegre hareketli ortalama süreci olarak bilinen ARIMA modeline uyar (Box et al., 2015; Canitez ve Savaş, 2022; Çelik, 2013).

Veri teknolojilerindeki hızlı gelişim, zaman serisi analizlerinin daha anlamlı ve güvenilir sonuçlar üretmesini sağlamış; bu durum, zaman serisi analizine olan ilgiyi ve önemini önemli ölçüde artırmıştır. Özellikle büyük veri ve hesaplama gücündeki ilerlemeler sayesinde, geçmiş gözlemlerden yola çıkarak gelecek dönemlere ilişkin daha isabetli tahminlerde bulunmak mümkün hâle gelmiştir. Zaman serisi tahminleri, belirli bir değişkenin geçmiş değerleri temel alınarak, bu değişkenin gelecekteki davranışlarını öngörmeye yönelik yapılan modelleme sürecidir. Bu yönüyle, zaman serisi tahmini; finans, ekonomi, meteoroloji, enerji ve sağlık gibi birçok alanda karar destek sistemlerinin temel bileşeni olarak stratejik bir rol oynamaktadır (Alsharif et al., 2022).

Finansal zaman serisi tahminlerinde, doğrusal regresyon, Otoregresif Hareketli Ortalama (ARMA), Otoregresif Koşullu Değişen Varyans (ARCH), Genelleştirilmiş Otoregresif Koşullu Değişen Varyans (GARCH) ve Hata Düzeltme Modelleri (ECM) gibi klasik istatistiksel yöntemler uzun yıllardır yaygın olarak kullanılmakta olup, günümüzde de önemli yardımcı araçlar olarak varlıklarını sürdürmektedir. Ancak bu geleneksel modeller, veri setinin belirli varsayımları karşılama gerektirmektedir; örneğin, serinin doğrusal yapıda olması, durağanlık özelliği göstermesi ve hata terimlerinin normal dağılıma sahip olması gibi. Bu varsayımlar, finansal zaman serilerinin genellikle doğrusal olmayan ve yapısal kırılmalara açık doğası göz önünde bulundurulduğunda, uzun vadeli tahmin performansını sınırlamakta ve modelin genelleme kapasitesini zayıflatmaktadır. Bu nedenle, doğrusal olmayan ve karmaşık dinamiklere sahip finansal serilerin analizinde daha esnek ve veri odaklı yöntemlerin kullanılmasına yönelik eğilim artmaktadır (Kazem et al., 2013; Yu et al., 2008; He et al., 2012; Niu et al., 2020). Son yıllarda, MÖ yöntemlerinin büyük ölçekli finansal zaman serisinin doğrusal olmayan zaman serisi modelleme kabiliyetleri bu açmazın üstesinden gelmekte ve klasik yöntemlere göre daha iyi performans göstermektedir (Niu et al., 2020; Karasu vd., 2020). MÖ yöntemlerinin içinde Yapay Sinir Ağları (Neural Network, NN)'nin önemli bir yeri vardır ve en sık çalışılan yöntemlerin başında gelmektedir (Crone and Nokopous, 2007; Aras ve Deveci Kocakoş, 2016). Derin öğrenme alt başlığı altında NN'nin diğer türevleri olan Çok Katmanlı Algılayıcı (Multi Layer Perceptron, MLP), Tekrarlı Sinir Ağları (Recurrent Neural Network, RNN), Uzun-Kısa Süreli Bellek (Long-Short Term Memory, LSTM) yine yaygın bir

çalışma alanı bulmuştur. Ancak, NN uygulamalarında gizli katman sayısının, bağlantı aralık değerlerinin ve öğrenme oranının ne olacağı gibi birtakım zorluklar hala geçerliliğini korumaktadır (Kuremoto et al., 2014).

2.4.1. ARIMA Modeli

ARIMA(p,d,q) modeli, otoregresyon (AR) modeli ve hareketli ortalama (MA) modelinin birleşimidir ve I bütünleşme (entegrasyon) anlamına gelmektedir. Burada p , otoregresyon derecesini, d fark alma derecesini, q hareketli ortalama derecesini göstermektedir. AR(p) modeli Eşitlik (2.6)'da gösterildiği gibidir:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + \alpha_t \quad (2.6)$$

Burada, $z_{t-1}, z_{t-2}, \dots, z_{t-p}$ gecikmeli seriler, $\phi_1, \phi_2, \dots, \phi_p$ gecikmeli serilerin katsayıları ve α_t ak gürültü serisini temsil etmektedir. MA(q) modeli Eşitlik (2.7)'de gösterilmektedir:

$$z_t = \alpha_t - \theta_1 \alpha_{t-1} - \theta_2 \alpha_{t-2} - \dots - \theta_q \alpha_{t-q} \quad (2.7)$$

Burada, $\alpha_t, \alpha_{t-1}, \alpha_{t-2}, \dots, \alpha_{t-q}$ ak gürültü gecikmeli hata serileri ve $\theta_1, \theta_2, \dots, \theta_q$ bu serilerin katsayıları olmaktadır. ARMA (p,q) modeli Eşitlik (2.8)'de yer almaktadır:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + \alpha_t - \theta_1 \alpha_{t-1} - \theta_2 \alpha_{t-2} - \dots - \theta_q \alpha_{t-q} \quad (2.8)$$

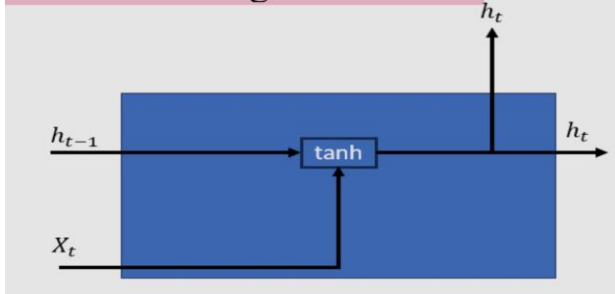
Seri durağansa, yani ortalaması ve varyansı sabitse tahmin edilebilir. ARIMA (p,d,q) modelinde yer alan fark alma parametresi d , seriyi durağan hale getirmek, başka bir deyişle seriyi tahmin edilebilir hale getirmek için yapılan fark işlemi sayısıdır (Box et al., 2015).

ARIMA modellemesinde ortaya çıkan AR ile MA terimlerinin katsayıları ve AIC (Akaike Bilgi Kriteri) değeri modelin seriye uygun olup olmadığını göstermektedir. AR ve MA katsayıları için p değeri 0,05'ten küçük çıkmış ise AR ve MA modeli anlamlıdır. Aksi durumda, AR ve MA modelinin seriye uygun olmadığı anlaşılır. Ayrıca, Eşitlik (2.6)-(2.8)'den anlaşılacağı gibi, modeldeki hata serisi mutlaka ak gürültü serisi olmalıdır. Bunun yanında düşük AIC değerine sahip model ile tüm varsayımlar sağlandığında tercih edilen model olmaktadır (Savaş, 2022).

2.4.2. LSTM (Long - Short Term Memory, Uzun Kısa Süreli Bellek)

Özellikle doğal dil işleme ve zaman serisi analizi gibi, geçmiş verilere dayalı bağımlılıkların kritik öneme sahip olduğu problemlerde, Özyinelemeli Sinir Ağları (Recurrent Neural Networks – RNN), Evrimsel Sinir Ağları'na (Convolutional Neural Networks – CNN) kıyasla daha başarılı sonuçlar vermektedir. RNN mimarisi, önceki zaman adımındaki nöron katmanının

çıkış değerini, aynı nörona tekrar giriş olarak iletmek suretiyle 'hafıza' işlevi görür. Bu geri besleme mekanizması sayesinde model, zaman içindeki veri bağımlılıklarını öğrenebilir. Şekil 9'da, RNN mimarisine ait temel yapı taşı olan bir RNN hücresi (cell) şematik olarak gösterilmektedir.



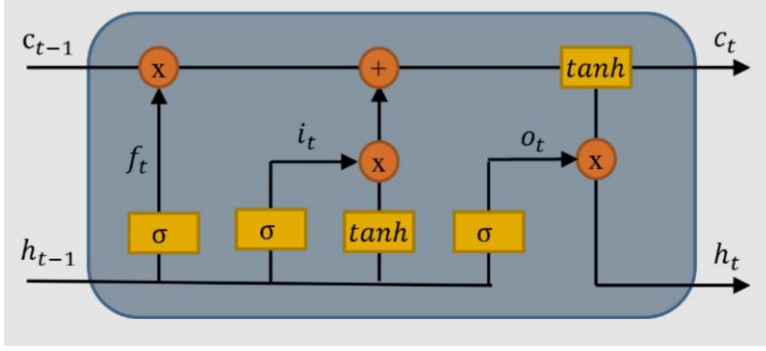
Şekil 9: RNN Sinir Ağı Yapısına Ait En Küçük Birim (Hücre) (Tanışman vd., 2021)

Şekil 9'da, h_{t-1} hücrenin bir önceki çıkışını, X_t yeni gelen girişi ve h_t çıkış kısmını temsil etmektedir. Şeklin matematiksel ifadesi ise Eşitlik (2.9)'da verilmektedir:

$$h_t = \tanh(W_h * h_{t-1} + W_i * X_t) \quad (2.9)$$

Burada, W_h gizli katmanın ağırlığını ve W_i giriş katmanının ağırlığını temsil etmektedir.

Veri setinin boyutunun arttığı ve öznelik sayısının çoğaldığı durumlarda, klasik Özyinelemeli Sinir Ağları (RNN) yapıları, sınırlı hafıza kapasiteleri nedeniyle yetersiz kalmakta ve özellikle uzun bağımlılıkları öğrenme sürecinde 'gradyan sönmesi' (vanishing gradient) gibi problemlerle karşılaşmaktadır. Bu sorun, bilginin ağırlık katmanlarında kaybolmasına ve modelin uzun vadeli ilişkileri öğrenememesine neden olur. Bu temel sınırlılığı aşmak amacıyla, 1997 yılında Sepp Hochreiter ve Jürgen Schmidhuber tarafından Uzun Kısa Süreli Bellek (Long Short-Term Memory – LSTM) ağı geliştirilmiştir. LSTM yapıları, temel olarak RNN mimarisine benzese de, bilgi akışını daha etkin biçimde yönetebilmek için giriş, çıkış ve unutma kapıları (forget gate) gibi özel katmanlar içerir. Bu kapılar, hücre durumunun güncellenmesini kontrol ederek hem kısa hem de uzun dönemli bağımlılıkların öğrenilmesini mümkün kılar. Şekil 10'da, LSTM mimarisine ait temel yapı ayrıntılı biçimde sunulmaktadır.



Şekil 10: LSTM Sinir Ağı Yapısına Ait En Küçük Birim (Hücre) (Tanışman vd., 2021)

RNN’de olduğu gibi giriş ve çıkışlar aynıdır. X_t giriş bilgisi ve h_t çıkış bilgisidir. C_t ve C_{t-1} giriş ve çıkışlar aslında LSTM’i RNN’den ayıran en temel parçadır. Bu hücre yolu ile bilgi taşınmaktadır. Bu sayede, ağ bir hafıza edinmektedir. σ , $[0,1]$ arasında değer alan sigmoid fonksiyonu temsil etmektedir. Eşitlik (2.10) ile hücrenin ilk kısmındaki aktivasyon işlemi ile geçmişten gelen hangi bilgilerin hafızada tutulup tutulmayacağına karar verilir. Bu kısma “unutma kapısı (forget gate)” denilmektedir.

$$forgetgate = \sigma(W_f(h_{t-1}, x_t) + b_f) \quad (2.10)$$

Burada, W_f , ağırlık değerlerini, b_f yan (bias) değerlerini temsil etmektedir. Giriş ve çıkış kapıları ise RNN yapısına benzerlik göstermektedir (Çevik vd., 2023).

Hücre durumu (cell state), LSTM mimarisinde zaman adımları boyunca bilgi taşıyan ve ağ içindeki veri akışının sürekliliğini sağlayan temel yapıdır. Bu yapı, geçmişten gelen bilgilerin uygun biçimde filtrelenerek güncellenmesini ve böylece uzun vadeli bağımlılıkların etkili bir şekilde korunmasını mümkün kılar (Tuncer, 2022). LSTM hücresinde bilgi akışı, unutma kapısı, giriş kapısı ve çıktı kapısı olmak üzere üç temel mekanizma ile düzenlenir. Unutma kapısı, önceki hücre durumunu ve gizli katman çıktısını dikkate alarak, geçmiş bilgilerin hangilerinin korunup hangilerinin silineceğine sigmoid aktivasyon fonksiyonu aracılığıyla karar verir; burada elde edilen $[0,1]$ aralığındaki değer, bilginin önem derecesini temsil eder. Giriş kapısı ise yeni bilginin hücre durumuna eklenip eklenmeyeceğini belirler; bu da yine sigmoid fonksiyonu ile değerlendirilir ve ardından oluşturulan yeni aday bilgi tanjant hiperbolik (tanh) fonksiyonuyla dönüştürülür. Hücre durumu, önceki hücre durumu ile unutma kapısından gelen değerlerin çarpımı ve giriş kapısından gelen değer ile yeni aday bilginin çarpımının toplamı ile güncellenir. Son aşamada, çıktı kapısı güncellenmiş hücre durumundan hangi bilginin çıktı olarak üretileceğini belirler; bu işlemde hücre durumu tanh fonksiyonundan geçirilir ve çıktı kapısının sigmoid sonucu ile çarpılarak bir sonraki zaman adımı için gizli

durum üretilir. Bu mekanizmalar sayesinde LSTM ağları, uzun süreli bağımlılıkları etkili şekilde öğrenebilir ve bilgiyi ağ boyunca taşımada klasik RNN'lere kıyasla çok daha başarılı performans sergiler (Yadav et al., 2020).

Girdi kapısı, unutma kapısı, çıktı kapısı, hücre durumu ve gizli durum değerleri sırasıyla i_t, f_t, c_t, o_t ve h_t olsun. x_t, t zamandaki girdi vektörü, σ , sigmoid aktivasyon fonksiyonu ve W, b sırasıyla parametre matrisi ve vektörü olmak üzere LSTM hücresi Eşitlik (2.11)-(2.15) arasında gösterilmiştir:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (2.11)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (2.12)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (2.13)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (2.14)$$

$$h_t = o_t \odot \tanh(c_t) \quad (2.15)$$

2.4.3. Destek Vektör Makinesi (Support Vector Machine)

Destek Vektör Makineleri (Support Vector Machines – SVM), ilk olarak Vapnik (1995) tarafından geliştirilmiş olup, teorik sağlamlığı, genel performans üstünlüğü ve farklı problem türlerine uygulanabilirliği sayesinde makine öğrenmesi alanında kısa sürede geniş bir kullanım alanı bulmuş ve önemli bir popülerite kazanmıştır. (Gunn, 1998). DVM'ler bir mesafe ölçüsü içerecek şekilde (Gunn, 1998) alternatif bir kayıp fonksiyonun tanıtılmasıyla regresyon problemlerine (Smola, 1996) de rahatlıkla uygulanabilmekte ve daha başarılı sonuçlar elde edilebilmektedir. Destek Vektör Regresyonu (Support Vector Regression – SVR), klasik tahmin yöntemlerinden farklı olarak, eğitim setindeki bireysel tahmin hatalarını en aza indirmeyi amaçlamak yerine, modelin genelleme yeteneğini artırmak üzere yapısal risk minimizasyonu (Structural Risk Minimization – SRM) ilkesine dayanmaktadır. Bu yaklaşım, modelin yalnızca eğitim verisine değil, aynı zamanda daha önce karşılaşılmamış yeni verilere karşı da yüksek doğrulukla tahminlerde bulunabilmesini hedefler. SVR'nin bu özelliği, öğrenme sürecinde elde edilen girdi-çıkı ilişkilerini daha etkin biçimde genellebilmesini sağlayarak, aşırı öğrenme (overfitting) riskini azaltır ve yüksek öngörü performansı sunar (Chen and Wang, 2007).

Destek Vektör Makineleri (SVM), makine öğrenmesi alanında hem regresyon hem de sınıflandırma problemlerinde etkin bir biçimde kullanılan güçlü algoritmalarından biridir. SVM

algoritmaları, hem doğrusal (linear) hem de doğrusal olmayan (non-linear) veri yapılarında başarılı sonuçlar sunmakta olup, özellikle doğrusal olmayan problemlerde 'çekirdek hilesi' (kernel trick) yöntemi ile düşük boyutlu verileri daha yüksek boyutlu özellik uzaylarına dönüştürerek ayırt edilebilirliği artırmaktadır. Bu sayede, sınıflar arasındaki farkları en iyi şekilde temsil eden karar sınırları (hyperplane) oluşturularak, sınıflandırma hatası minimize edilir. SVM'nin bu yapısı, yalnızca statik veri kümeleri değil, aynı zamanda zaman serisi verileri üzerinde de uygulanabilir kılmaktadır. Nitekim Stefan Rüping'in 'SVM Kernels for Time Series Analysis' başlıklı çalışmasında, SVM modellerinin özellikle çok değişkenli zaman serilerinde nasıl uygulanabileceği ve çekirdek fonksiyonu (kernel) seçiminin bu bağlamdaki önemi ayrıntılı şekilde ele alınmaktadır. SVM algoritmasının sınıflandırma problemlerine uygulanmasında, Destek Vektör Sınıflandırması (Support Vector Classification – SVC) yapısı kullanılmaktadır. SVC, denetimli öğrenme paradigması çerçevesinde çalışmakta olup, önceden etiketlenmiş veri setleri (labelled data) üzerinden modeli eğiterek sınıflar arasındaki en uygun ayırıcı sınırı (decision boundary) belirlemeyi amaçlamaktadır. Bu yapı sayesinde, öğrenme süreci boyunca model, kullanıcı tarafından sağlanan etiketler yardımıyla doğru sınıflandırma kurallarını öğrenir ve test aşamasında da bu kuralları uygulayarak yüksek doğrulukla sınıflandırma yapabilir. Bu güdümlü öğrenme mekanizması, sınıflar arasındaki marjini maksimize ederek genelleme performansını artırmayı hedeflemektedir. (Jakkula, 2006).

SVM algoritmasının regresyon problemlerine uygulanabilir versiyonu olan Destek Vektör Regresyonu (Support Vector Regression – SVR), tahmin hatalarını en aza indirmeye yönelik olarak tasarlanmış özel bir yapıdır. SVR, klasik regresyondan farklı olarak, yalnızca gözlem hatalarını minimize etmeye değil, aynı zamanda modelin genelleme kapasitesini artırmaya odaklanır. Bu amaçla, deneysel riski ölçmek için bir tolerans sınırı (ϵ -insensitive loss function) içeren özel bir maliyet fonksiyonu kullanılır. Bu maliyet fonksiyonu, belirlenen ϵ aralığında kalan tahmin hatalarını cezalandırmazken, bu sınırın dışındaki sapmaları optimize etmeye çalışır. Böylece SVR, yüksek doğrulukla çalışan ve aşırı öğrenme riskine karşı dayanıklı bir regresyon modeli sunar (Yu et al., 2016).

2.4.4. Karar Ağaçları (Decision Tree)

Karar ağaçları, sınıflandırma ve regresyon problemlerinde yaygın olarak kullanılan, veriye dayalı karar verme süreçlerini temsil eden yapısal modellerdir. Bu modeller, verilerin basit, ikili veya çoklu karar kurallarına göre ayrıştırıldığı, akış şeması benzeri bir hiyerarşik yapı şeklinde organize edilmiştir. Karar ağacı yapısında veriler, her biri belirli bir özelliğe dayanan sorulara verilen yanıtlara göre dallara ayrılır ve bu ayırım işlemi, verinin belirli kriterlere göre alt

gruplara ayrıştırılması şeklinde ilerler. Bu yapıda, tüm veri kümesini temsil eden başlangıç noktası 'kök düğüm' (root node) olarak adlandırılır. Kök düğümden ayrılan her karar noktası 'düğüm' (node) olarak tanımlanırken, bu düğümlerin uç noktalarında yer alan ve artık daha fazla bölünmeyen son düğümler 'yaprak' (leaf) olarak isimlendirilir. Yaprak düğümler, modelin çıktısını (sınıf etiketi veya tahmin değeri) temsil eder ve her yaprak, o bölgeye düşen tüm gözlemler için geçerli olan bir kararı ifade eder. Bu yapı, karar verme sürecini hem görsel hem de yorumlanabilir bir biçimde sunması nedeniyle özellikle yorumlanabilirlik gerektiren uygulamalarda tercih edilmektedir.

2.4.5. Rastgele Orman (Random Forest)

Rastgele Orman (Random Forest – RF) algoritması, Breiman (2001) tarafından bir topluluk (ensemble) öğrenme yöntemi olarak geliştirilmiş olup, sınıflandırma ve regresyon problemlerine yönelik güçlü bir çözüm sunmaktadır. Bu yöntem, birden fazla karar ağacının eğitim verisinin rastgele örneklenmiş alt kümeleri ve öznitelik alt grupları üzerinden bağımsız olarak inşa edilmesine dayanmaktadır. Ardından, tüm ağaçların çıktılarının birleştirilmesiyle (örneğin sınıflandırmada oylama, regresyonda ortalama alma) nihai tahmin gerçekleştirilir. Rastgele Orman, yüksek öngörü doğruluğu ile modelin yorumlanabilirliğini dengeli bir şekilde bir arada sunması, aşırı öğrenmeye karşı dayanıklı yapısı ve öznitelik önem derecelerini belirleyebilme kapasitesi sayesinde günümüzde en popüler makine öğrenmesi yöntemlerinden biri hâline gelmiştir (Qi, 2012). Rastgele Orman, çok çeşitli tahmin problemlerine uygulanabilirliği, ayarlanabilmesi gereken birkaç parametresinin olması, kullanım açısından kolay olması, birçok pratik probleme başarıyla uygulanması ve küçük örneklem boyutları ile başa çıkabilmesi, yüksek boyutlu özellik uzayları ve karmaşık veri yapıları ile başa çıkabilmesi açısından literatürde çok popüler bir yöntemdir (Tyrallis and Papacharalampous, 2017). RF, çok sayıda karar ağacından oluşan ve standart tek bir karar ağacına kıyasla varyansın azalmasıyla sonuçlanan bir topluluk öğrenme tekniğidir (Couronne et al., 2018). Birçok çalışma ile geçerliliğini ispatlamış olan RF algoritması, sınıflandırma ve regresyon görevlerinde oldukça başarılı sonuçlar ortaya koymaktadır. RF sağladığı değişken önem ölçüleri, farklı sınıf ağırlıklandırma, eksik değer atama, görselleştirme, aykırı değer tespiti, denetimsiz öğrenme gibi ek özellikler nedeniyle araştırmacılar için oldukça çekicidir (Cutler et al., 2012). RF, bir x girdi vektörü kullanarak bağımsız olarak $h_k(x)$, ($k=1,2,\dots,K$) adet regresyon ağacı meydana getirir ve model tahmini, oluşturulan ormandaki her bir ağaçtan tahminin ortalaması olarak elde edilmektedir (Seo et al., 2017). RF denklemi, Eşitlik (2.16)'da belirtildiği gibidir:

$$\text{Rastgele Orman Tahmin} = \frac{1}{K} \sum_{k=1}^K h_k(x) \quad (2.16)$$

Eşitlik (2.16)'dan anlaşılacağı üzere, modelin bireysel ağaçlardan gelen tahmin değerlerinin ortalamasını alması nedeniyle örneklem varyansı azalma göstermektedir (Özdemir ve Çılgın, 2022).

Karar ağacı algoritmaları, bazı durumlarda optimal sonuçlar üretememesi ve özellikle küçük değişimlere duyarlı, hassas veri setlerinde yüksek varyans sergilemesi nedeniyle sınırlı performans gösterebilmektedir. Bu tür sınırlamaları aşmak amacıyla geliştirilen Rastgele Orman algoritması, karar ağaçlarının zayıf yönlerini ortadan kaldırmak için topluluk öğrenmesi tekniklerinden biri olan 'bagging' (bootstrap aggregating) metodunu kullanır. Bagging yöntemi, eğitim verisinin farklı alt örneklem kümeleri üzerinde birden çok karar ağacı eğiterek model varyansını azaltır ve tahmin doğruluğunu artırır. Her bir ağaç bağımsız olarak eğitilir ve nihai sonuç, ağaçların çıktılarının birleştirilmesiyle (sınıflandırmada çoğunluk oyu, regresyonda ortalama) elde edilir. Rastgele Orman algoritmasında, hem gözlem örnekleri hem de öznitelikler rastgele seçildiği için modelin çeşitliliği artırılır ve overfitting riski azaltılır. Ancak, bu süreçte çok sayıda bağımsız karar ağacının paralel olarak çalıştırılması, yüksek işlem yüküne ve özellikle donanım açısından yetersiz sistemlerde aşırı ısınma gibi pratik zorluklara yol açabilir. Ayrıca, açıklamada yer alan 'boosting' metoduna atıf, Rastgele Orman yapısından ziyade 'arttırımlı öğrenme' (boosting) algoritmalarına özgüdür; zira Random Forest algoritması sıralı değil, paralel yapıdır ve boosting metodolojisiyle karıştırılmamalıdır (Öngün, 2023).

2.4.6. Yapay Sinir Ağları (Artificial Neural Networks-ANN-YSA)

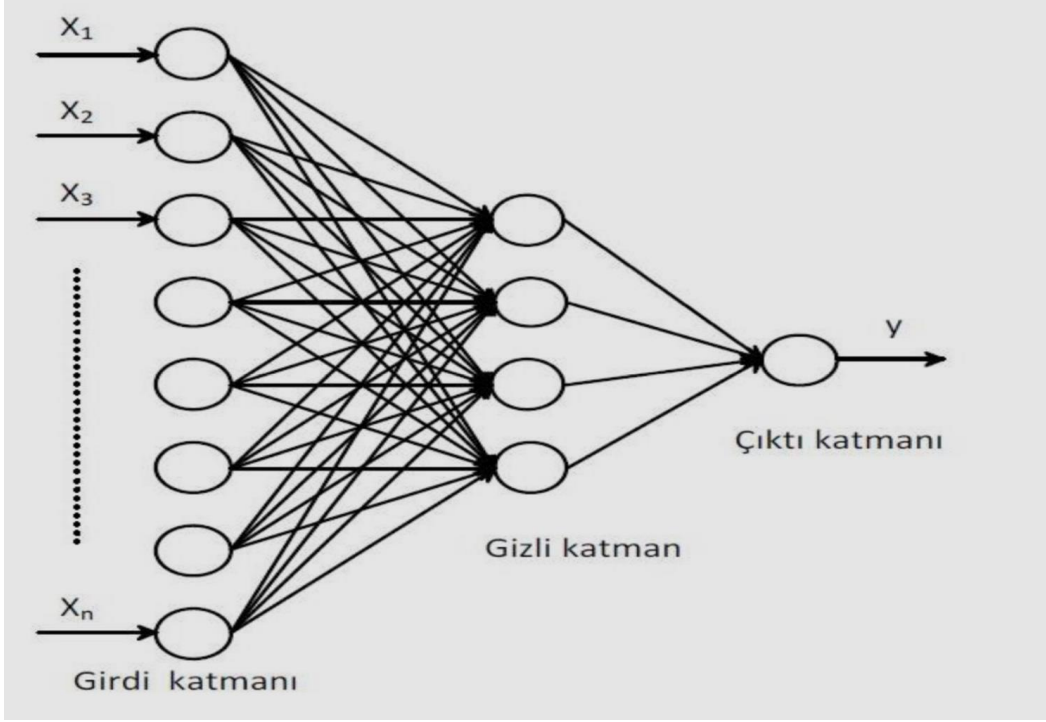
Yapay Sinir Ağları (YSA), insan beynindeki biyolojik sinir hücrelerinin (nöronların) işleyişinden esinlenerek geliştirilmiş, birbirine bağlı yapay nöronlardan oluşan hesaplama modelleridir. Geniş nöronal bağlantı yapısına sahip bu modeller, bilgi işlem süreçlerini katmanlı mimariler aracılığıyla gerçekleştirir. YSA'lar, temelde yarı parametrik regresyon tahmin edicileri olarak işlev görür; yalnızca doğrusal ilişkileri modellemekle kalmaz, aynı zamanda karmaşık ve doğrusal olmayan ilişki örüntülerini de başarılı bir şekilde öğrenip simüle edebilir. Bu özellikleri sayesinde, özellikle klasik istatistiksel yöntemlerin yetersiz kaldığı durumlarda güçlü bir alternatif olarak öne çıkmaktadır (McCluskey et al., 2013). Yapay Sinir Ağları (YSA)'nın temel gücü ve avantajı, modelledikleri veri setlerinden mevcut ilişkileri herhangi bir ön varsayım gerektirmeksizin doğrudan öğrenebilme kapasitelerinde yatmaktadır. Ağ, eğitim süreci boyunca giriş-çıkış ilişkilerini öğrenerek, bu bilgileri ağırlıklar ve aktarım fonksiyonları aracılığıyla içselleştirir. Eğitim tamamlandıktan sonra YSA, bu öğrenilen yapıyı

kullanarak daha önce karşılaşmadığı veriler üzerinde sınıflandırma, zaman serisi analizi, tahmin ve örüntü tanıma gibi çeşitli görevleri başarıyla gerçekleştirebilir (Wesolowski and Suchacz, 2012). Bu ağlar, veri ilişkilerinden öğrenerek ve görünmeyen durumları genelleştirerek model fonksiyonu tahmininde bulunabilir ve doğrusal-doğrusal olmayan fonksiyonları işleyebilirler. Popüler YSA'lardan bir tanesi Çok Katmanlı Algılayıcı (ÇKA)'dır (Taud and Mas, 2018). ÇKA'nın mimarisi veriye uygun ve uygulanacağı göreve göre değişken olmakla birlikte, genel olarak birkaç nöron katmanından oluşmaktadır (Gardner and Dorling, 1998).

Sinir ağı tahmin modellerinin geliştirilmesi ve uygulanması için çeşitli yöntemler ve ağ topolojileri bulunmakla birlikte en yaygın kullanım şekli geri yayılım öğrenme algoritmaları ileri beslemeli sinir ağı topolojisidir (Ali vd., 2017). Bir ÇKA'da model katmanları birbiri ile tamamen bağlıdır, yani katmandaki her bir nöron bir sonraki katmandaki her bir nörona bağlanmıştır. Yapay sinir ağlarında öğrenme süreci, eğitim aşamasında verilen gözlemlere karşılık beklenen çıktının doğru hesaplanabilmesi için her bir bağlantıdaki ağırlıkların hesaplanması ile gerçekleşmektedir. Bir YSA'da öğrenme algoritması temelde verdiğimiz ağın nöral bağlarının ağırlıklarını hesaplama görevinden sorumlu olmaktadır. Bu öğrenme algoritması, çıktı ve beklenen arasındaki hataların karesini en aza indirmek için gradyan arama ile sinir ağını eğitmek için kullanılmaktadır (Özdemir ve Çılgın, 2022).

Sonuç olarak, sinir ağları tekniği, klasik istatistiksel tahmin yöntemlerine göre NDVI görüntülerine dayanan zaman serisi tahmininde yeni bir alternatif olarak uygulanmıştır (Menezes and Barreto, 2008; Stepchenko, 2016).

YSA'ların yapısı Şekil 11'de gösterilmiştir.



Şekil 11: YSA Mimarisi (Ay ve Ekinci, 2022)

Yapay Sinir Ağlarında (YSA), giriş katmanındaki her bir düğüm (nöron), bir sonraki katmanda yer alan düğümlere bağlanarak bilgi akışını başlatır. Bu bağlantılar aracılığıyla, giriş verileri ağ içerisinde iletilir ve işlenir. Ara katmanlar ,diğer adıyla gizli katmanlar, giriş katmanından gelen bilgileri çeşitli ağırlıklar ve aktivasyon fonksiyonları yardımıyla işler ve bu işlenmiş bilgiyi çıkış katmanına aktarır. Çıkış katmanı ise son gizli katmandan aldığı bilgileri birleştirerek nihai tahmini üretir ve bu bilgiyi dış dünyaya aktarır. Bu yapı, verinin katmanlar arasında dönüşerek giderek daha soyut temsillere ulaşmasını sağlar ve ağın öğrenme kapasitesini belirler (Ay ve Ekinci, 2022).

2.4.7. K- En Yakın Komşu (KNN)

Tahmin çalışması kapsamında makine öğrenmesi yöntemlerinden olan k -en yakın komşuluk (KNN) algoritmasından yararlanılır. Bu yöntem, 4 adımda aşağıdaki şekilde gerçekleştirilir:

1. En yakın komşu sayısı olan k belirlenir,
2. Gözlemler arasındaki uzaklıklar belirlenir,
3. Hesaplanan uzaklıklar arasındaki en küçük k tanesi seçilir,
4. Seçilen k tane en küçük uzaklıkların aritmetik ortalaması alınarak tahmin çalışması tamamlanmış olur (Akküçük, 2011).

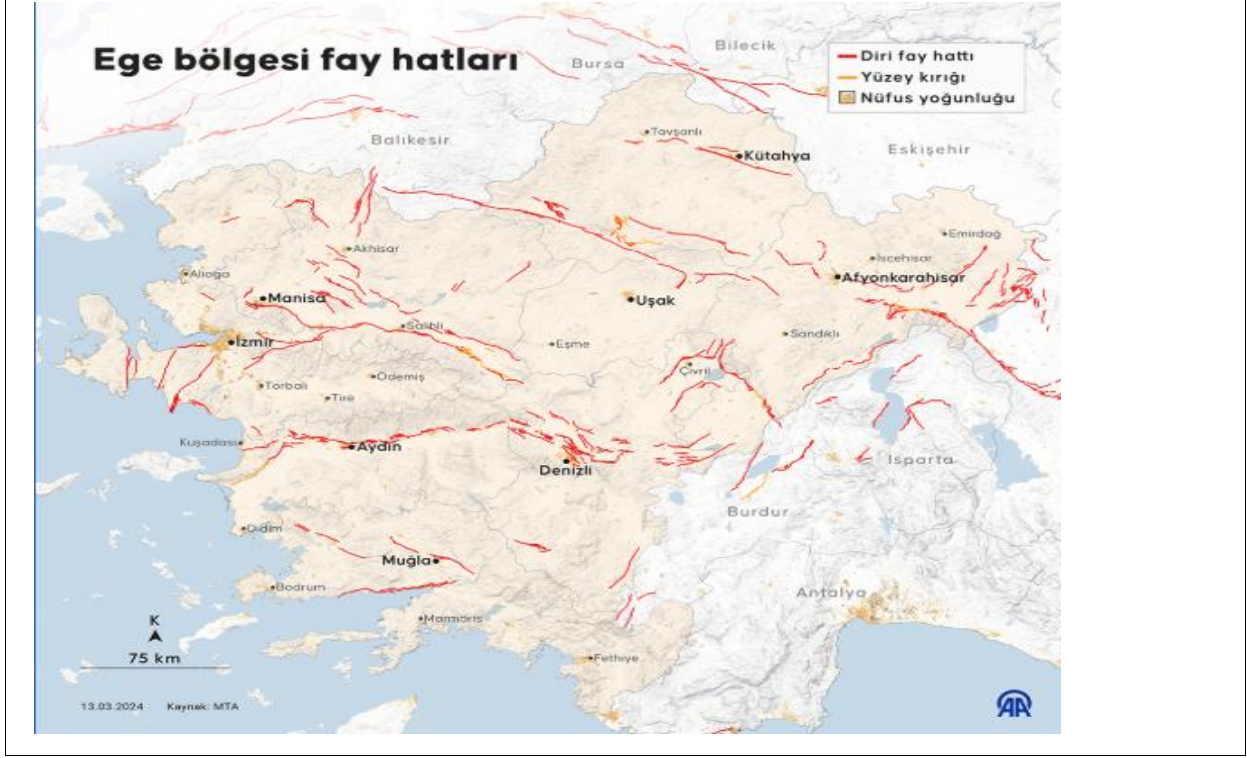
2.4.8. XGBoost

XGBoost karar ağacı temelli bir topluluk öğrenimi algoritmasıdır (Chen et al., 2016). İlgili algoritmanın temel çalışma prensibi, değişkenlere başlangıçta eşit olarak atanan ağırlıkların, modelleme süreci ilerledikçe problem yapısına bağlı olarak dinamik biçimde güncellenmesine dayanmaktadır. Bu yaklaşımda, model birden fazla karar ağacından oluşan bir topluluk (ensemble) yapısı kurar ve bu topluluktan elde edilen çıktılar üzerinden tahminler üretir. Modelin eğitim sürecinde, yanlış sınıflandırılan gözlemlere daha fazla önem verilmesini sağlamak amacıyla bu gözlemlerin ağırlıkları artırılırken, doğru sınıflandırılan gözlemlerin ağırlıkları düşürülür. Bu mekanizma, modelin daha zor ve hataya açık örnekler üzerinde yoğunlaşmasını sağlar. Sonuç olarak, öğrenme süreci boyunca ağaçlar giderek daha uyumlu hale gelir ve sistem, aşırı öğrenme (overfitting) eğilimini azaltan, genelleme performansını ise artıran çeşitli düzenleme (regularization) stratejileri ile desteklenir. Bu yapı, özellikle Boosting temelli algoritmalarda (örneğin AdaBoost, Gradient Boosting) karakteristiktir (Ekinci vd., 2020). Bu özelliğinden dolayı söz konusu yöntem, literatürde sıklıkla 'düzenli artırma' (regularized boosting) tekniği olarak anılmaktadır. Algoritma, doğruluğu artırmaya yönelik çeşitli düzenleme (regularization) mekanizmaları içermekte ve bu sayede modelin genelleme performansını iyileştirmektedir. Ayrıca, paralel işlemeye elverişli yapısı sayesinde büyük veri kümelerinde yüksek işlem verimliliği sağlar. Eksik veri durumlarına karşı da dayanıklı olup, eksik gözlemleri doğrudan işleyebilen standartlaştırılmış bir algoritmik yapıya sahiptir. Ek olarak, yükseltme (boosting) sürecinin her yineleme adımında çapraz doğrulama uygulanarak modelin aşırı öğrenmeye karşı korunması sağlanmakta ve parametre optimizasyonu daha güvenilir biçimde gerçekleştirilmektedir. Bu tür özellikleriyle, XGBoost gibi modern düzenli artırma algoritmaları makine öğrenmesi alanında yaygın bir şekilde kullanılmaktadır (Ay ve Ekinci, 2022).

3. UYGULAMA

Bu bölümde, farklı yöntemler ile deprem büyüklüğü tahmin edilmiş ve en uygun yöntem ile deprem derinliği ele alınarak İstanbul ilçelerine ait deprem büyüklüğü tahminleri elde edilmiştir. Yapılan analizler, Anaconda Navigator içerisinde bulunan Jupyter Notebook üzerinden Python kodları ile oluşturulmuştur.

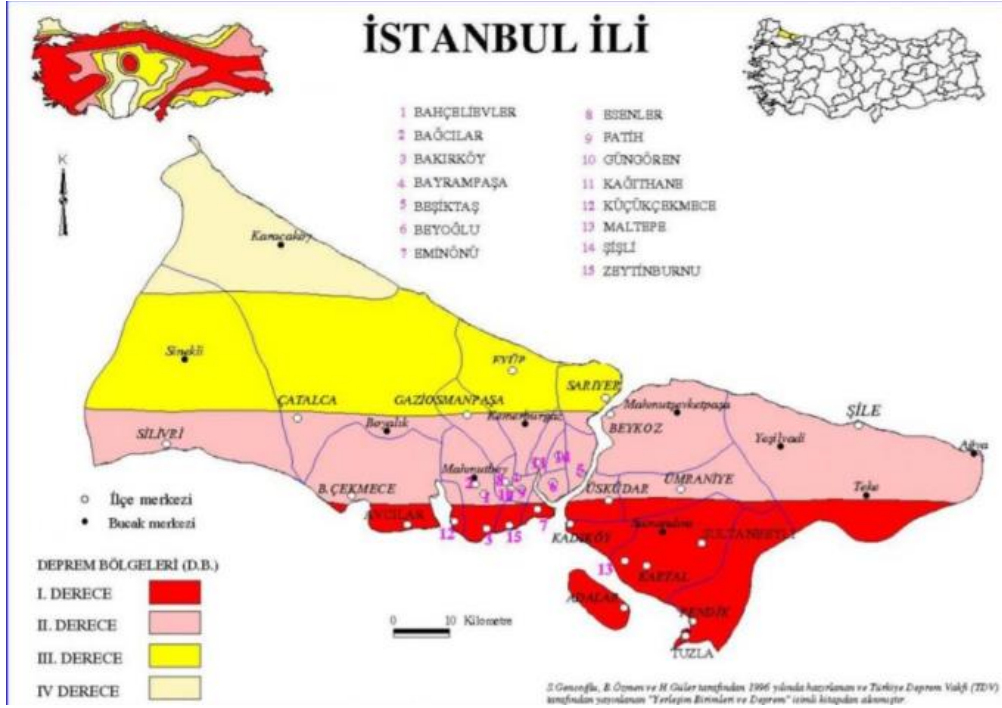
Ele alınan deprem verileri, Gamze Özel Kadılar (2024) tarafından yürütülen 121F208 numaralı Yapay Zeka ve Olasılıksal Model Tabanlı Deprem Tehlike Haritası adlı TÜBİTAK 1001 projesinden elde edilmiştir ve bu veri seti 1900-2021 yılları arasında Ege Bölgesi'nde meydana gelmiş depremlere ait; tarih, zaman, enlem, boylam, derinlik ve büyüklük bilgilerini içermektedir. Şekil 14'te ise sadece İstanbul'daki fay hatları görülmektedir. Deprem kataloğu oluşturulurken AFAD-DDA kataloğu temel katalog olarak ele alınmıştır. Bu veri seti, 1900-2021 yılları arasında meydana gelen Mw 3.0 ve üzeri 16504 depreme ait; tarih (date), zaman (time), enlem (Lat), boylam (Long), büyüklük (Mw), derinlik (depth) bilgilerini içermektedir. Uygulama kapsamında kullanılan veriler, aylık ortalama deprem büyüklükleri elde edilerek zaman serisi analizine uygun hale getirilmiştir.



Şekil 12: Ege Bölgesi Fay Haritası (Demirci, 2024)



Şekil 13: Marmara Bölgesi Fay Haritası (Demirci, 2024)



Şekil 14: İstanbul İlçeleri için Fay Haritası (<https://www.gazetebirlik.com/gundem/istanbulda-fay-hatti-gecen-ilceler-97472/97472>)

İlk olarak Jupyter Notebook üzerinden veri seti yüklendi, bağımsız değişkenler olarak; Depth (km), Lat, Long kullanıldı ve bağımlı değişken olarak Mw alındı. Çizelge 2 incelendiğinde, deprem büyüklüğü (Mw) değişkeninin ortalaması 3.55 olup, veriler 3.0 ile 7.6 arasında dağıldığı görülmektedir. Depremler genellikle sığ odaklı olup, ortalama derinlik 16,6 km olarak belirlenmiştir. Derinlik değerlerinde gözlemlenen standart sapmanın yüksek olması, bazı depremlerin derin odaklı olabileceğini göstermektedir. Deprem büyüklüğü (Mw) değerlerinin 3.0 ile 7.6 arasında sınırlı olması hem jeofiziksel hem de istatistiksel olarak mantıklı olmaktadır. Veri seti genellikle hissedilebilir ve kayda değer depremleri içermekte olup mikro depremler dışlanmıştır. Ayrıca, Ege fay sistemi doğrultusunda meydana gelen depremlerin genellikle 7.5'i geçmediği gözlemlenmiştir. Deprem büyüklüğü (Mw) değerlerinin bu sınırlarda dağılması Gutenberg-Richter yasasına uygun şekilde küçük depremlerin daha sık, büyük depremlerin ise seyrek olduğunu destekler niteliktedir. Analizler, 1900 – 2021 yılları arasında deprem büyüklüğü 3'ten büyük olan toplam 16503 depremin verileri ele alınarak yapılmıştır.

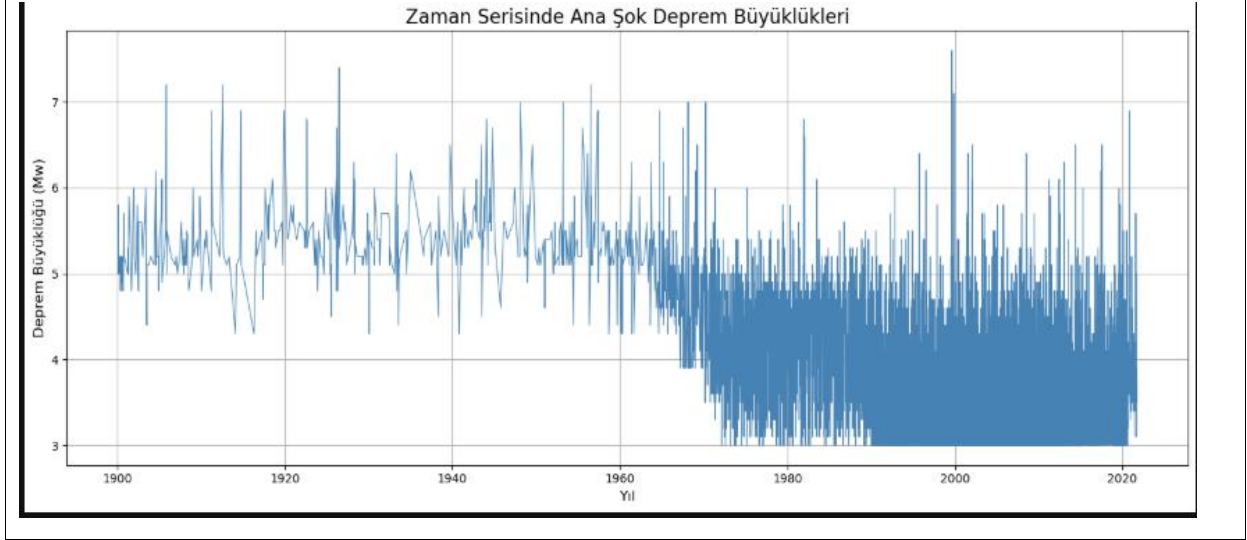
Çizelge 2: Tanımlayıcı İstatistikler

	count	mean	min	25%	50%	75%	max	std
Mw	16503	3.553	3	3.1	3.3	3.8	7.6	0.64
Depth (km)	16503	16.618	0	6.9	10	17	199	21.891
Lat	16503	38.175	34.81	36.77	38.37	39.49	41.83	1.657
Long	16503	28.147	23.759	26.698	27.84	29.35	33.772	1.947

Veri Seti ve Değişken Tanımları:

- 1. Depth (km)-Deprem Derinliği:** Bu değişken, depremin yer kabuğunun ne kadar derininde meydana geldiğini km cinsinden ifade eden değişkendir. Depremler; sığ (0-70 km), orta derinlik (70-300 km) veya derin (300 km'den fazla) şeklinde gerçekleşebilir. Derinliğin artması ile depremin yüzeyde hissedilme şiddeti azalmaktadır. Bu çalışmada kullanılan verilerde genellikle sığ depremler gözlenmiştir.
- 2. Longitude (Long)–Boylam:** Depremin episantrının (yer yüzündeki merkezi) doğu-batı konumunu belirtmektedir. Bu değişken, depremlerin coğrafi dağılımını haritalamada kolaylık sağlaması amacıyla kullanılmaktadır.
- 3. Latitude (Lat)–Enlem:** Depremlerin episantrının kuzey – güney yönündeki konumunu belirtmektedir. Bu değişken genellikle depremlerin coğrafi konum analizlerinde yardımcı olmaktadır.
- 4. Moment Magnitüdü (Mw)–Deprem Büyüklüğü:** Oluşan depremlerin büyüklüğünü gösteren değişkendir. Bu ölçek, deprem sırasında açığa çıkan toplam enerjiyi hesaplamaktadır. Richter ölçeği gibi diğer klasik ölçeklerle kıyasladığımız zaman Mw büyük depremler için daha doğru sonuçlar vermesinden kaynaklı modern sismolojide daha yaygın olarak kullanılmaktadır. Mw değeri logaritmik ölçekli olmasından kaynaklı her bir birim artışın yaklaşık olarak 32 kat fazla enerji anlamına gelmektedir. Bu tez kapsamında Mw değişkeni bağımlı değişken olarak kabul edilmiştir.

Küme Numaraları–All Clusters: Veri setinde yer alan bu değişken, deprem kayıtlarının benzerliklerine göre oluşturulmuş küme atamalarını içermektedir. Veri setinin yapısal çeşitliliğini göstermek amacıyla kullanılmıştır. Bu tez kapsamında söz konusu değişken analiz süreçlerinde kullanılmış, ancak istatistiksel olarak önemli bulunmadığı için İstanbul ilçelerine ait ortalama deprem büyüklükleri tahmininde hesaba katılmamıştır.



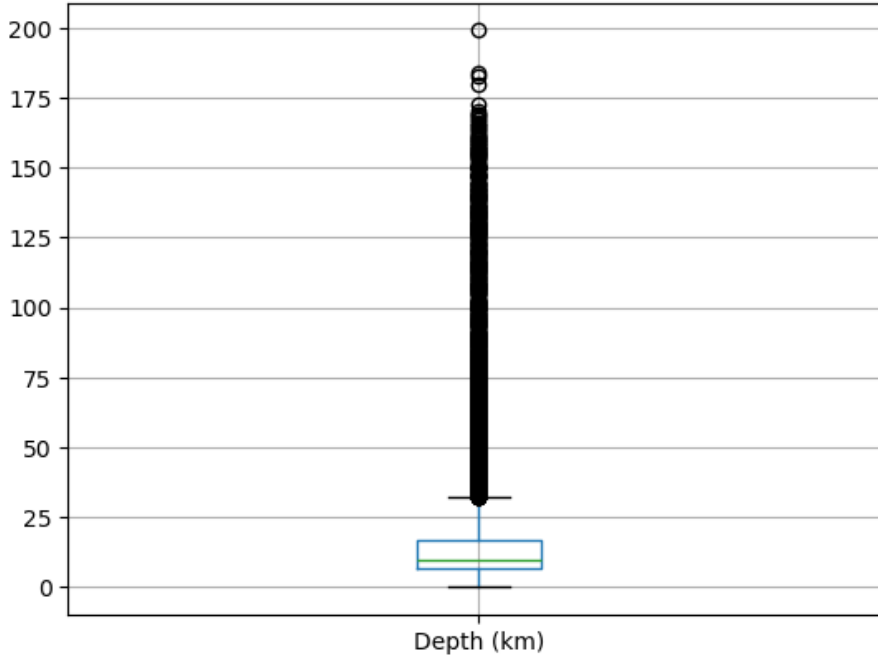
Şekil 15: Zaman Serisindeki Ana Şok Deprem Büyüklükleri

Şekil 15'te ana şoklara ait deprem büyüklüklerinin yıllar içerisindeki değişimi zaman serisi grafiği ile sunulmuştur. Grafikte özellikle 1960 yılları sonrasındaki dönemde gözlem sayısında belirgin bir artış olduğu ancak, bu dönemdeki deprem büyüklüklerinin önceki yıllara göre daha düşük seviyelerde yoğunlaştığı görülmektedir. Bu durum, ölçüm teknolojilerindeki gelişmeler ile daha küçük büyüklükteki depremlerin de kaydedilmesinden kaynaklanıyor olabilir. Ayrıca, zamanla büyüklüklerin azaldığına dair görsel bir eğilim bulunmakta olup, bu yapının istatistiksel olarak test edilmesi amacıyla bir sonraki bölümde doğrusal regresyon ve kantil regresyon analizleri gerçekleştirilmiştir.

Bu çalışmada, deprem büyüklüğünü etkileyen faktörleri belirlemek amacıyla çoklu doğrusal regresyon analizi uygulanmıştır. Bağımlı değişken olarak moment büyüklüğü (Mw), bağımsız değişkenler olarak yıl, derinlik (depth (km)), enlem (lat), boylam (long) ve küme numaraları (all clusters) modele dahil edilmiştir. Daha sonrasında yapılan analizler sonucunda küme numaraları (all clusters) değişkeni istatistiksel olarak anlamsız bulunduğu için modelden çıkartılmış ve modele dahil edilmemiştir.

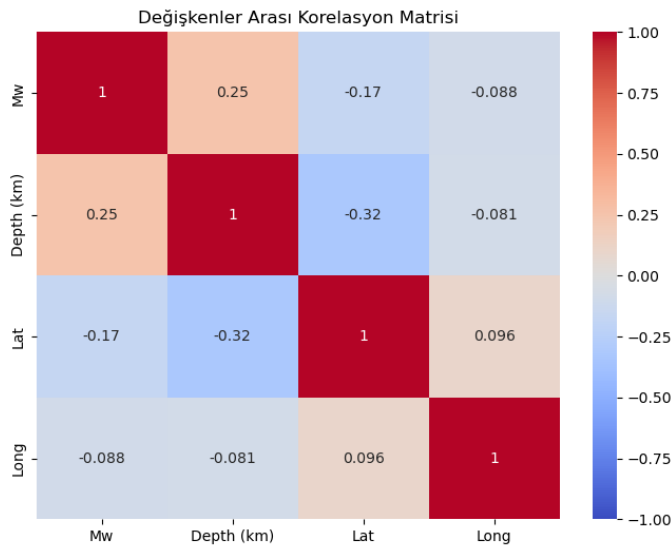
Python ile gerekli kütüphanelerin ve veri setinin yüklenmesi sonrasında all clusters ve datetime sütunları veri setinden çıkartılmıştır ve tanımlayıcı istatistikler tablosu Çizelge 2'de belirtildiği şekilde elde edilmiştir. Veri seti içerisindeki eksik değerlerin kontrolü sağlandıktan sonra veri setinde eksik değere rastlanmadı. Bağımsız değişken olarak kullanılacak Depth (km) değişkeni için aykırı değer varlığını gözlemlemek için bozplot grafiği Şekil 16'da belirtildiği üzere kutunun büyük kısmınının 0-25 km arasında olduğu bunun da depremlerin çoğunun yer kabuğuna yakın yerde yani sığ depremler olduğunu gözlemlemiştir. Aykırı depremlerin ise genellikle

derinlik dağılımına göre çok daha derin depremler olduğunu ve derinliğin sağa çarpık dağıldığı görülmektedir.



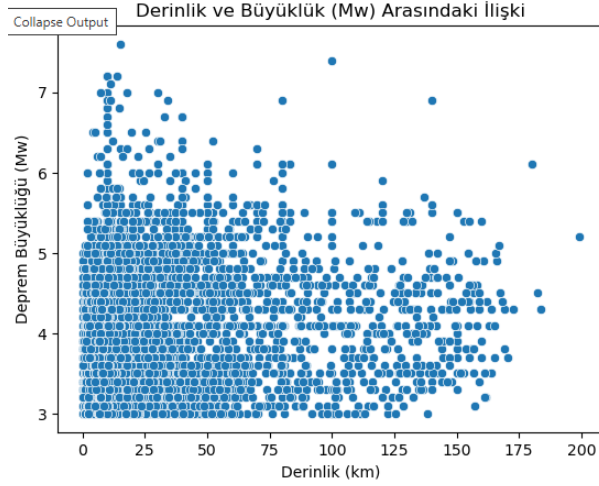
Şekil 16: Derinliğin Boxplot Grafiği

Değişkenler arasındaki ilişkiyi incelemek için Şekil 17’de verilen korelasyon matrisi oluşturuldu ve en anlamlı ilişki 0,25 ile derinlik ve büyüklük arasında bulunmuştur. Yani, derinlik değişkeninin büyüklük değişkenini etkilediği yorumu yapılabilmektedir.



Şekil 17: Değişkenler Arasındaki Korelasyon Matrisi

Derinlik ve büyüklük arasındaki ilişkiyi incelemek için Şekil 18’de verilen scatterplot grafiği çizilmiştir ve derinliğin 0-50 km arasında çok sayıda depreme sebep olduğu ve bu depremlerin çoğunun sığ depremler ile yüzeye yakın gerçekleştiği ve derinliğin artması ile deprem sayısında azalma olduğu söylenebilmektedir.



Şekil 18: Derinlik ve Büyüklük Arasındaki İlişkinin Scatterplot Grafiği

Korelasyon matrisi ve scatterplot grafiklerinin incelenmesi sonucunda bağımlı değişken olarak Mw (deprem büyüklüğü) ve bağımsız değişkenler olarak Depth (km) (derinlik), Lat (enlem) ve Long (boylam) olarak alınmasına karar verilmiştir.

Veri seti %80 eğitim ve %20 test verisi olacak şekilde eğitilmiştir.

Oluşturulan modellerin performanslarını ölçmek amacıyla doğrusal regresyon modeli oluşturulmuştur ve RMSE değeri 0,6148 olarak hesaplanmıştır. Bu değer modelimizin, deprem büyüklüğünü ortalama olarak $\pm 0,61$ birimlik bir hata ile tahmin ettiğini belirtmektedir.

Regresyon modeli sonuçları Çizelge 3’te belirtilmiştir.

Çizelge 3: Regresyon Modeli Sonuçları

```
Model katsayıları (betalar):  
Depth (km): 0.0063  
Lat: -0.0342  
Long: -0.0193  
  
Y-kesişim (Intercept): 5.2957  
R2 (Açıklanan Varyans): 0.0810
```

Çizelge 3 incelendiğinde, derinlik değişkenine ait katsayı $\beta = 0,0063$ olarak hesaplanmıştır ve derinlik değişkenindeki 1 birimlik artışın Mw’yi yaklaşık olarak 0,0063 birim arttırmakta

olduğunu göstermektedir. Lat değişkenine ait katsayı $\beta = -0,0342$ olarak hesaplanmıştır ve enlem değişkenindeki 1 birimlik artışın yani kuzey yönüne doğru çıkılmasının Mw değişkeninde yaklaşık olarak 0,034 birimlik bir azalma gösterdiği söylenebilmektedir. Long değişkenine ait katsayı $\beta = -0,0193$ olarak hesaplanmış ve boylam 1 birim arttığında yani doğuya gidildikçe Mw değişkeninde yaklaşık olarak 0,019 birimlik bir azalma olduğu görülmektedir. $\beta_0 = 5,2957$ olarak hesaplanmış ve hiçbir değişkenin etkisi olmadığında deprem büyüklüğünün 5,3 olduğu söylenebilmektedir. Çizelgedeki R^2 değeri ise 0,081 olarak hesaplanmış ve modelin verideki toplam bilgisinin %8.1'ini açıkladığı söylenebilmektedir.

QuantReg Regression Results						

Dep. Variable:	Mw	Pseudo R-squared:	0.05616			
Model:	QuantReg	Bandwidth:	0.08372			
Method:	Least Squares	Sparsity:	1.004			
Date:	Wed, 18 Jun 2025	No. Observations:	16503			
Time:	10:18:28	Df Residuals:	16499			
		Df Model:	3			

	coef	std err	t	P> t	[0.025	0.975]

const	5.8076	0.108	53.551	0.000	5.595	6.020
Depth (km)	0.0064	0.000	33.905	0.000	0.006	0.007
Lat	-0.0512	0.002	-20.489	0.000	-0.056	-0.046
Long	-0.0219	0.002	-10.832	0.000	-0.026	-0.018

The condition number is large, 1.41e+03. This might indicate that there are strong multicollinearity or other numerical problems.

Şekil 19: Kantil Regresyon Sonuçları

Sonuçlar göz önüne alındığında, $\beta_0 = 5,8076$ olduğunda diğer tüm değişkenlerin 0 olması durumunda beklenen ortalama deprem büyüklüğünün 5.8 olduğunun yorumu yapılabilmektedir. Derinlik değişkeninin katsayısı $\beta = 0,0064$ ile derinlikteki 1 kilometrelik artışın deprem büyüklüğünü ortalama olarak 0,0064 birim arttırdığını, enlem değişkeninin katsayısının $\beta = -0,0512$ ile enlemdeki 1 birimlik artışın yani kuzeye gidildikçe deprem büyüklüğünde 0,0512 birim azalma görüldüğü, boylam değişkeninin katsayısının $\beta = -0,0219$ ile boylamdaki 1 birimlik bir artışın yani doğuya gidildikçe deprem büyüklüğünde 0,0219 birimlik bir azalma olduğu söylenebilir.

Farklı kantil değerleri için (0,25; 0,50; 0,75) model sonuçları Şekil 20-Şekil 22'de verilmiştir.

==== 25. Kantil Sonuçları ====

QuantReg Regression Results

```
=====
Dep. Variable:          Mw      Pseudo R-squared:      0.02788
Model:                 QuantReg Bandwidth:                0.07786
Method:                Least Squares Sparsity:                0.6148
Date:                  Wed, 18 Jun 2025 No. Observations:      16503
Time:                  10:34:33      Df Residuals:          16499
                                           Df Model:              3
=====
```

	coef	std err	t	P> t	[0.025	0.975]
const	4.2412	0.054	78.374	0.000	4.135	4.347
Depth (km)	0.0036	9.94e-05	35.938	0.000	0.003	0.004
Lat	-0.0224	0.001	-17.556	0.000	-0.025	-0.020
Long	-0.0116	0.001	-11.273	0.000	-0.014	-0.010

The condition number is large, 1.41e+03. This might indicate that there are strong multicollinearity or other numerical problems.

Şekil 20: 25. Kantil Sonuçları

==== 50. Kantil Sonuçları ====

QuantReg Regression Results

```
=====
Dep. Variable:          Mw      Pseudo R-squared:      0.05616
Model:                 QuantReg Bandwidth:                0.08372
Method:                Least Squares Sparsity:                1.004
Date:                  Wed, 18 Jun 2025 No. Observations:      16503
Time:                  10:34:33      Df Residuals:          16499
                                           Df Model:              3
=====
```

	coef	std err	t	P> t	[0.025	0.975]
const	5.8076	0.108	53.551	0.000	5.595	6.020
Depth (km)	0.0064	0.000	33.905	0.000	0.006	0.007
Lat	-0.0512	0.002	-20.489	0.000	-0.056	-0.046
Long	-0.0219	0.002	-10.832	0.000	-0.026	-0.018

The condition number is large, 1.41e+03. This might indicate that there are strong multicollinearity or other numerical problems.

Şekil 21: 50. Kantil Sonuçları

===== 75. Kantil Sonuçları =====

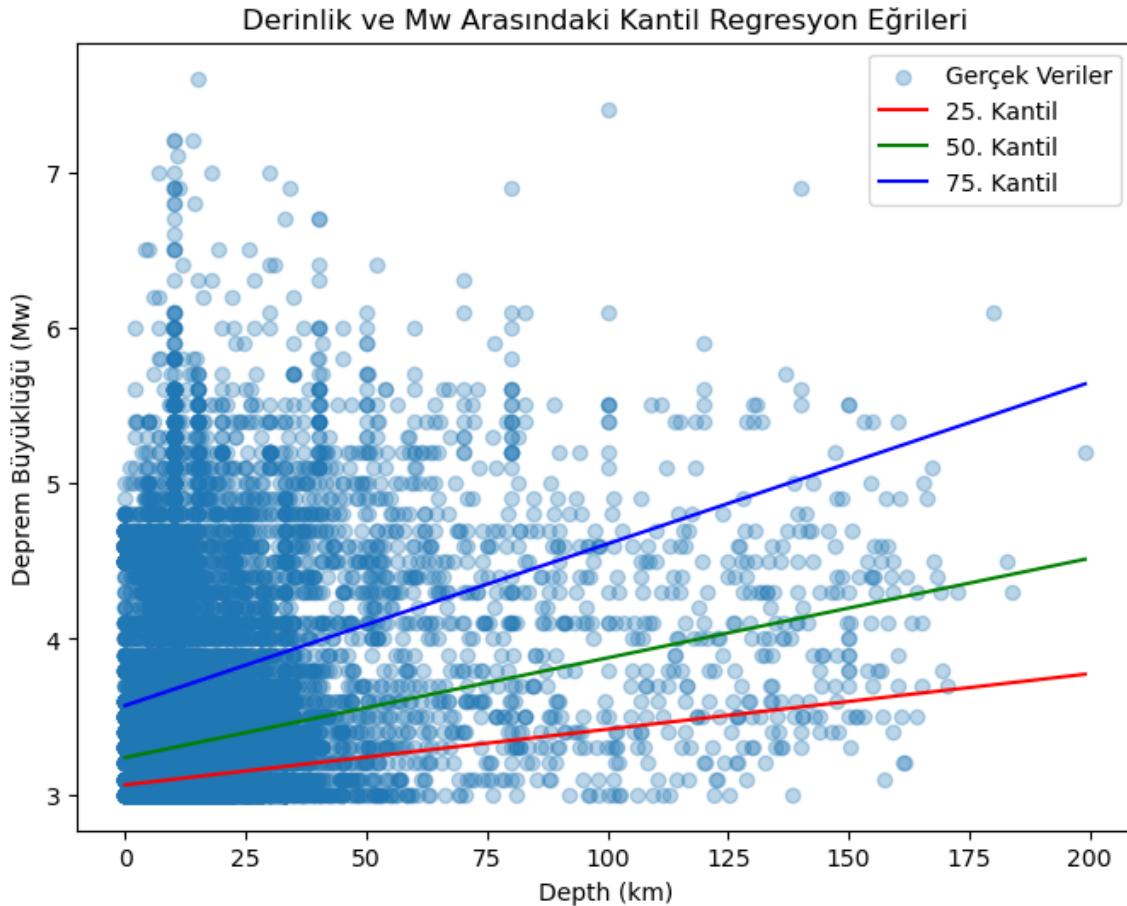
QuantReg Regression Results

```
-----  
Dep. Variable:          Mw      Pseudo R-squared:          0.06496  
Model:                  QuantReg  Bandwidth:                  0.07144  
Method:                 Least Squares  Sparsity:                   2.588  
Date:                   Wed, 18 Jun 2025  No. Observations:         16503  
Time:                   10:34:33      Df Residuals:              16499  
                          Df Model:                3  
-----
```

	coef	std err	t	P> t	[0.025	0.975]
const	6.6396	0.250	26.520	0.000	6.149	7.130
Depth (km)	0.0104	0.000	26.696	0.000	0.010	0.011
Lat	-0.0581	0.006	-10.152	0.000	-0.069	-0.047
Long	-0.0302	0.005	-6.609	0.000	-0.039	-0.021

The condition number is large, 1.41e+03. This might indicate that there are strong multicollinearity or other numerical problems.

Şekil 22: 75. Kantil Sonuçları

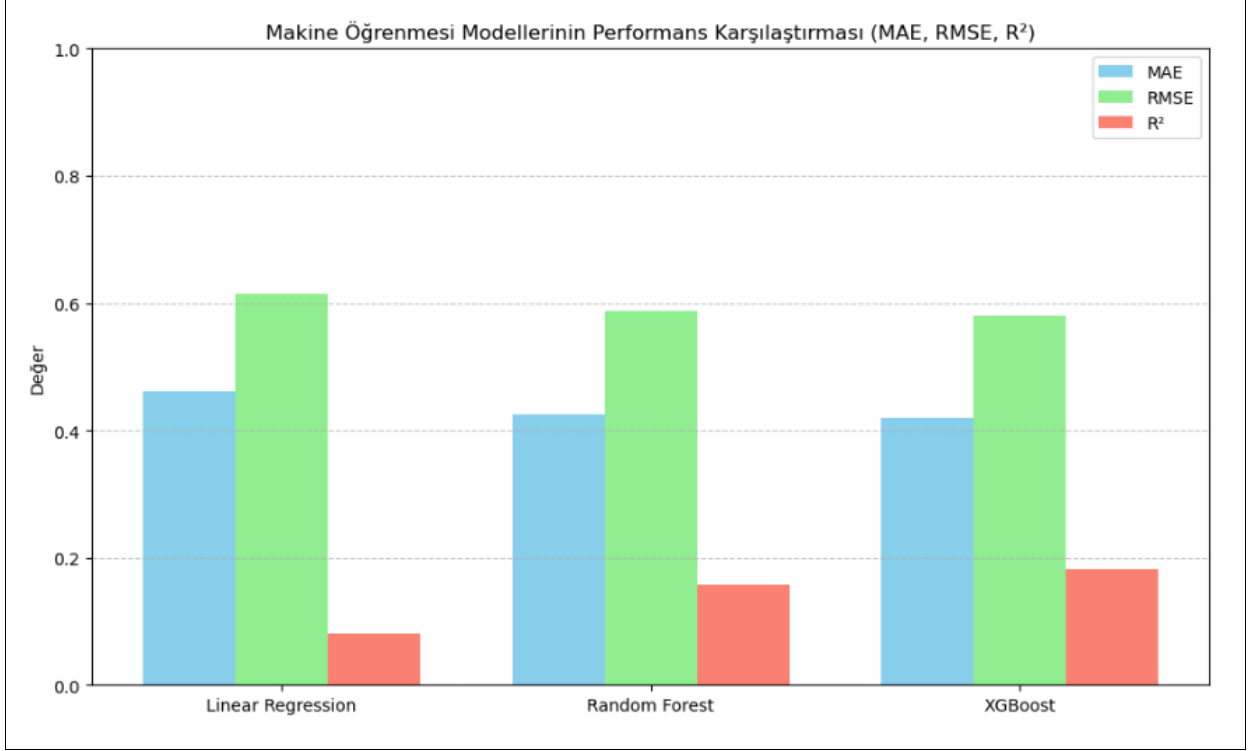


Şekil 23: Derinlik ve Mw Arasındaki Kantil Regresyon Eğrisi

Şekil 23, deprem derinliği ile deprem büyüklüğü arasındaki ilişkiyi farklı kantil düzeylerinde inceleyen kantil regresyon eğrisini göstermektedir. Tüm kantil eğrileri derinlik arttıkça deprem büyüklüğünde hafif bir artış eğilimi gözlemlendiğini ve 50. Kantil yani medyan eğrisinin ortalama ilişkiyi temsil ettiği gözlemlenmiştir. 25. Kantilde daha küçük depremlerin derinlikten daha az etkilendiği düşünülmektedir. 75. Kantilde ise eğrilerin belirgin bir pozitif eğime sahip olduğunu ve büyük depremlerin derinlik artışından fazla etkilenmediğini göstermektedir.

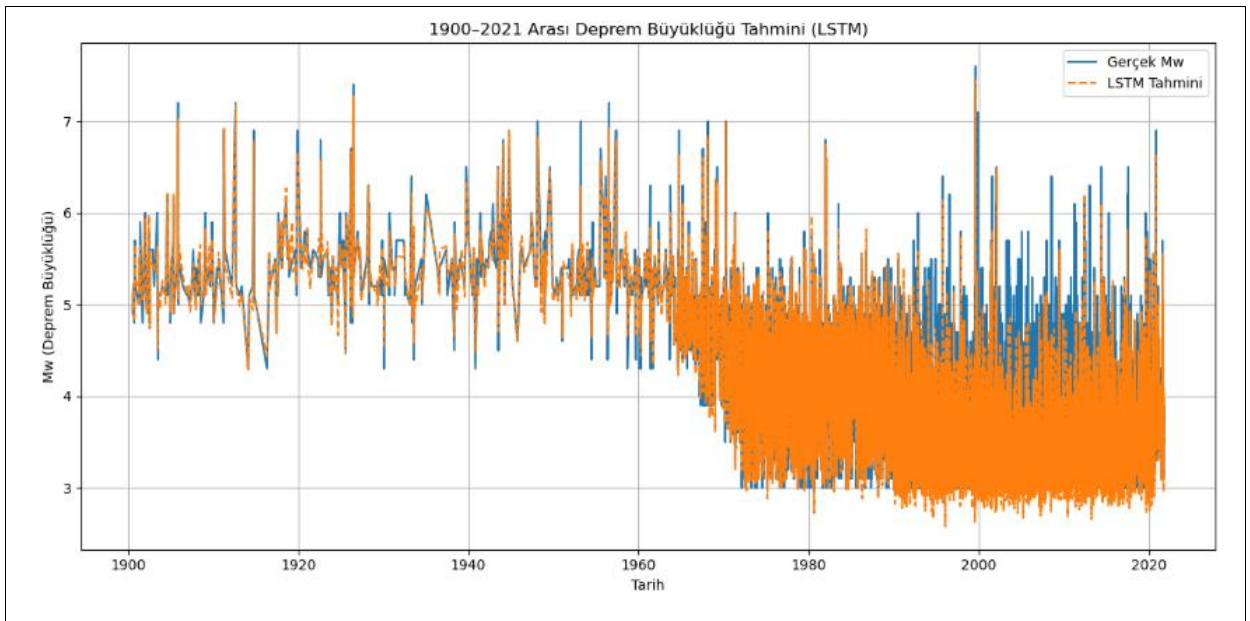
Doğrusal regresyon modeli ile yapılan tahminlerde Ortalama Mutlak Hata (MAE) 0,462 ve Hata Kareler Ortalamasının Karekökü (RMSE) değeri 0,615 olarak tahmin edilmiştir. Bu sonuç, modelin deprem büyüklüğü tahminlerinde ortalama olarak yaklaşık $\pm 0,6$ birimlik bir sapma ile tahminler yaptığını göstermektedir. Elde edilen R^2 değeri 0,081 olup, bağımsız değişkenlerin deprem büyüklüğünün yaklaşık %8'inin açıklandığını göstermektedir. Bu oran, doğrusal regresyon modelinin deprem büyüklüğü üzerindeki etkisini sınırlı bir şekilde yakalayabildiğini, ancak modelin iyileştirilmesi için daha gelişmiş ve karmaşık yaklaşımlar gerektiğini göstermektedir.

Şekil 24'te görüldüğü gibi Random Forest Modeli ile yapılan tahminlerde MAE değeri 0,425, RMSE değeri 0,588 ve R^2 değeri 0,158 olarak elde edilmiştir. Bu sonuçlar, modelin deprem büyüklüğü tahminlerinde ortalama olarak $\pm 0,6$ birimlik bir sapma ile tahminler yaptığını ve modelin deprem büyüklüğündeki varyansın yalnızca %15'ini açıkladığını göstermektedir. Doğrusal regresyon modeline kıyasla, Random Forest Modeli daha düşük hata değerleri ve daha düşük açıklama oranları ile sonuçlanmış olsa bile verinin karmaşıklığını öğrenme konusunda daha başarılı olabildiğini göstermiştir. Bu, Random Forest Modelinin deprem verilerindeki karmaşık yapıyı daha iyi öğrenme potansiyeline sahip olduğunu düşündürmektedir. Makine Öğrenmesi (MÖ) modelleri arasında en başarılı sonuçları XGBoost'un verdiği görülmektedir. Örneğin, 50 ağaçlı XGBoost Modeli için MAE değeri 0,4197, RMSE değeri 0,5801 ve R^2 değeri 0,1819 olarak bulunmuştur. Bu değerler, modelin ortalama tahmin hatasının makul düzeyde olduğunu ancak büyük sapmaların ve açıklanan varyansın sınırlı olduğunu göstermektedir. Ancak modelin ağaç sayısı artırıldığında performansta belirgin bir iyileşme görülmektedir. Doğrusal regresyon modelleri, deprem büyüklüğü gibi karmaşık olayları açıklama konusunda yetersizlik gösterebilmektedir. Ağaç tabanlı makine öğrenmesi modelleri, özellikle Random Forest ve XGBoost gibi algoritmalar bu tür karmaşık ilişkileri daha iyi öğrenerek daha düşük hata ve daha yüksek açıklama başarısı göstermektedir.



Şekil 24: Makine Öğrenmesi Modellerinin Performans Karşılaştırması

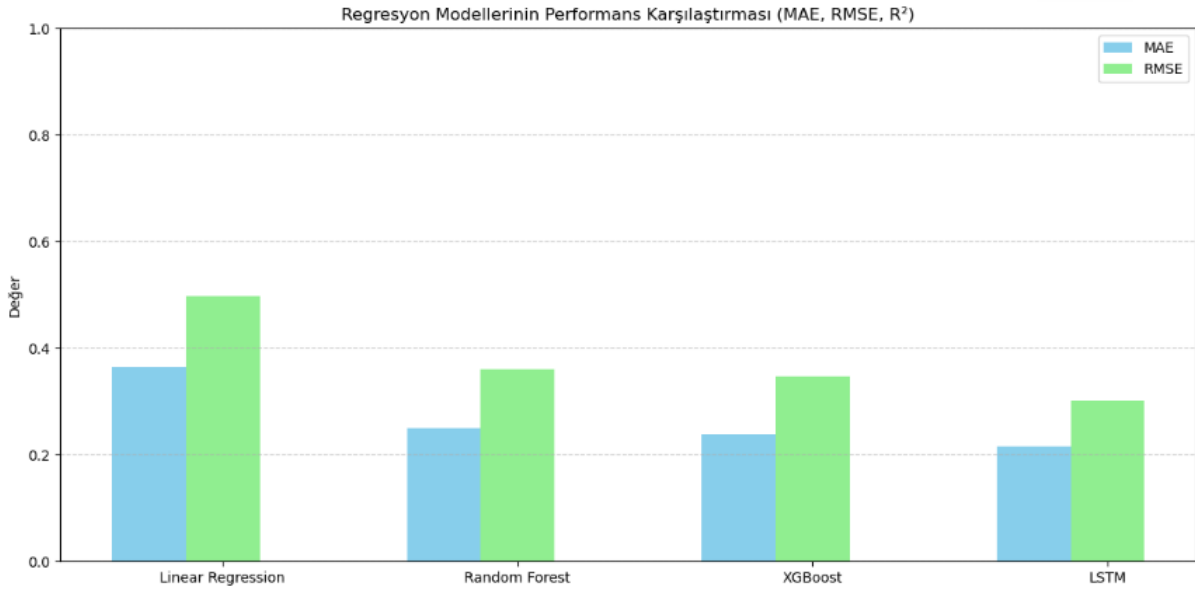
Model eğitimi, 750 epoch boyunca gerçekleştirilmiştir. Eğitim süreci boyunca her epoch modelin öğrenme sürecinde verileri işleyerek kayıp fonksiyonu minimize etmeye çalıştığı bir döngüde her epoch, eğitim verisinin tamamı üzerinden yapılan bir iterasyonu ifade eder ve modelin iyileşmesi için bir adım olmaktadır. Modelin eğitim sürecinde kayıp değeri her epoch sonunda güncellenmiş ve genellikle azalma eğilimi göstermiştir. Bu durum, modelin her bir döngüde öğrenme sürecini iyileştirerek, tahmin doğruluğunu arttırdığını göstermektedir.



Şekil 25: LSTM ile Zaman Serisi Tahmini-Deprem Büyüklüğü (Mw)

Şekil 25'te 1900-2021 yılları arasında meydana gelen depremlerin büyüklüklerini (Mw) ve bu büyüklüklerin LSTM ile tahmin edilen değerleri karşılaştırılmıştır. Mavi çizgiler gerçekleşen deprem ölçümlerini gösterirken, turuncu çizgiler LSTM modelinin tahmin ettiği deprem büyüklüklerini göstermektedir. Şekil 25 incelendiğinde çizgilerin birbirine yakın seyrettiklerini ve LSTM modelinin genel olarak deprem büyüklüklerini doğru tahmin ettiğini ancak bazı noktalarda (1940-1960 yılları arasında) tahminler ve gerçek değerler arasında küçük farklılıklar görülebildiğini göstermektedir. Özellikle son yıllarda, modelin daha tutarlı tahminler yapıldığı görülmektedir.

LSTM modeli ile yapılan zaman serisi tahminlerinde elde edilen performans metrikleri, modelin doğruluğunu değerlendirmek amacıyla kullanılmıştır. MAE=0,2159 değeri ile modelin tahminlerinin ortalama olarak gerçek değerden yaklaşık olarak $\pm 0,22$ birimlik bir sapma gösterdiğini belirtirken RMSE=0,3019 değeri model tahminlerinin genel olarak gerçek değere oldukça yakın olduğunu göstermektedir. LSTM modelinin, zaman bağımlı yapıları öğrenme yeteneği sayesinde elde edilen tahminlerde genel eğilimi başarılı bir şekilde yakaladığı görülmektedir.



Şekil 26: Makine Öğrenmesi ve LSTM Modellerinin Ortalama Mutlak Hata (MAE) ve Hata Kareler Ortalamasının Karekökü (RMSE) Karşılaştırılması

Şekil 26'da dört farklı makine öğrenmesi modelinin deprem büyüklüğündeki performansları üç metrik üzerinden karşılaştırılmıştır. Elde edilen sonuçlara göre, en yüksek MAE ve RMSE değerleri doğrusal regresyon modelinde gözlemlenmektedir. Bu durum doğrusal regresyonun

karmaşık ve doğrusal olmayan örüntüleri yakalama konusunda sınırlı olduğunu göstermektedir. Random Forest ve XGBoost modelleri, hem MAE hem de RMSE değerleri bakımından doğrusal regresyon modeline göre kıyasla daha düşük hata seviyelerine ulaşarak daha başarılı tahminler sunmaktadır. Bu durumda, XGBoost modelinin yaklaşık 0,25 MAE ve 0,35 RMSE değeri ile performans açısından öne çıktığı görülmektedir.

Özellikle LSTM modeli, zaman serisi yapısına uygun olarak çalıştığı için düşük MAE ve RMSE değerleriyle en başarılı model olarak öne çıkmaktadır. Bu bulgu, LSTM'nin geçmiş verilerdeki bağımlılıklarını daha etkili bir şekilde öğrenebildiğini ve bu sayede deprem büyüklüğü gibi zamana bağlı değişkenleri tahmin edebilme konusunda başarı sağladığını göstermektedir.

İstatistiksel olarak değerlendirme yapıldığında hem MAE hem de RMSE değerlerinin azalması, modelin tahmin doğruluğunun arttığını göstermektedir. Bu sonuçlar, makine öğrenmesi tabanlı yöntemlerin, klasik doğrusal modellere kıyasla daha düşük hata oranları ile çalıştığını ve karmaşık veri yapılarıyla daha iyi bir şekilde başa çıkabildiğini ortaya koymaktadır.

3.1 İstanbul İlçeleri İçin Uygulama

Bu tez çalışmasında İstanbul için yapılan analiz, bölgesel farklılıkların ve sismik risklerin daha doğru bir şekilde değerlendirilmesini sağlamayı amaçlamaktadır. İstanbul, büyük bir metropol olup, farklı jeolojik ve sismolojik özelliklere sahiptir. Şehirdeki farklı yapılaşma, yeraltı fay hatları, zemin yapısı ve nüfus yoğunluğu gibi faktörler, deprem riski üzerinde önemli bir etkiye sahiptir. Dolayısıyla, İstanbul genelindeki analiz, çeşitli bölgelerindeki farklı risk seviyelerinin belirlenmesine ve olası bir depremde alınması gereken yerel önlemlerin daha etkin bir şekilde planlanmasına olanak tanır.

İstanbul'un farklı coğrafi ve sosyo-ekonomik yapıları, her bölgesinin depremden alacağı etkiyi farklılaştırmaktadır. Daha yoğun yerleşim alanlarına sahip bölgelerde büyük bir depremde daha fazla can ve mal kaybı riski bulunurken fay hatlarına daha yakın bölgelerde ise deprem büyüklüğünün etkisi çok daha yıkıcı olabilmektedir. Bu sebeple, İstanbul genelinde yapılan tahminler, şehirdeki tüm bölgeler için spesifik ve uygulanabilir veriler sağlayarak, afet yönetimi stratejilerinin daha etkili bir şekilde belirlenmesine katkıda bulunur. Ayrıca, İstanbul'a özgü analizler, şehrin farklı bölgelerinde alınması gereken özel önlemleri, erken uyarı sistemlerini ve afet sonrası müdahale planlarını daha iyi şekillenmesine yardımcı olmaktadır.

Analizler Anaconda Navigator içerisinde bulunan Jupyter Notebook üzerinden Python kodları kullanılarak yapıldı ve ilk olarak 1900-2021 yılları arasında Ege Bölgesi'nde meydana gelen

Mw 3.0 ve üzeri 16504 deprem verisini içeren veri seti üzerinden LSTM modeli tahmini gerçekleştirildi. 750 epoch boyunca eğitilen veri setinde, eğitim kaybı yaklaşık olarak 0,0083 düzeyine ve doğrulama kaybı ise yaklaşık 0,0126 seviyelerine ulaşmıştır. LSTM modeli ile gerçekleştirilen deprem büyüklüğü tahmininde modelin başarı oranını değerlendirmek için MAE ve RMSE değerleri kontrol edilmiştir. MAE=0,3706 birimlik bir sapma değeri ile LSTM modelinin gerçek değerlere oldukça yakın tahminler yaptığını ortaya koymuştur. RMSE=0,5161 olması sonucunda modelin genel olarak tutarlı tahminler yaptığı ve aşırı hata yapma eğiliminde olmadığı görülmüştür.

Analizler Anaconda Navigator içerisinde bulunan Jupyter Notebook üzerinden Python kodları kullanılarak yapıldı ve ilk olarak İstanbul ilçe koordinatları modele dahil edilmek için yeni bir veri seti oluşturuldu. Daha öncesinde oluşturulmuş LSTM modeli dikkate alınarak İstanbul ilçe koordinatları modele eklendi ve 10 km ve 30 km olarak farklı derinliklerde deprem büyüklük tahminleri gerçekleştirildi.

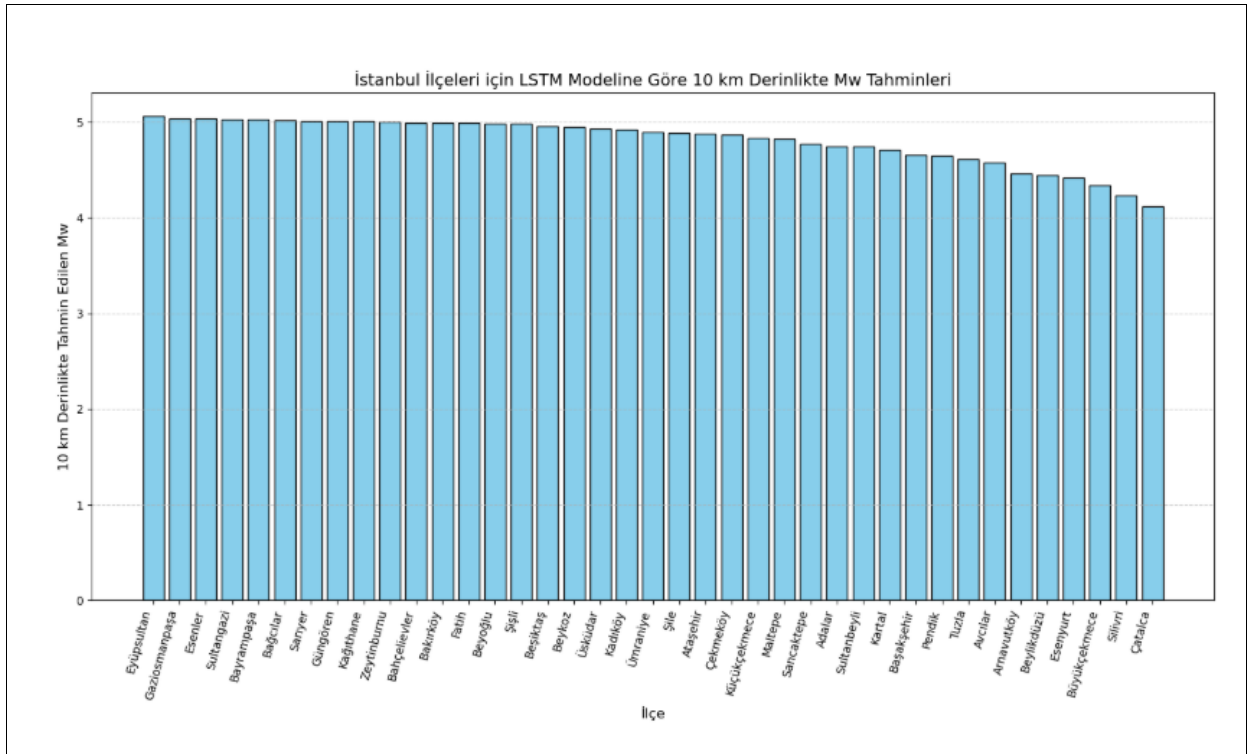
Çizelge 4: İstanbul İlçe Koordinatları

İlçe	Enlem (Lat)	Boylam (Long)
Adalar	40.8688	29.1204
Arnavutköy	41.2000	28.7400
Ataşehir	40.9833	29.1167
Avcılar	40.9761	28.7214
Bağcılar	41.0390	28.8567
Bahçelievler	41.0000	28.8500
Bakırköy	40.9769	28.8532
Başakşehir	41.0931	28.8025
Bayrampaşa	41.0498	28.8996
Beşiktaş	41.0438	29.0094
Beykoz	41.1250	29.0900
Beylikdüzü	40.9833	28.6333
Beyoğlu	41.0386	28.9858
Büyükdere	41.0206	28.5850
Çekmeköy	41.0025	29.1706

Esenler	41.0439	28.8900
Esenyurt	41.0386	28.6750
Eyüpsultan	41.0783	28.9206
Fatih	41.0128	28.9497
Gaziosmanpaşa	41.0794	28.9064
Güngören	41.0214	28.8722
Kadıköy	40.9917	29.0306
Kağıthane	41.0661	28.9747
Kartal	40.8997	29.1781
Küçükçekmece	41.0028	28.7803
Maltepe	40.9369	29.1608
Pendik	40.8742	29.2675
Sancaktepe	40.9875	29.2300
Sarıyer	41.1702	29.0502
Silivri	41.0739	28.2422

Şekil 27, 10 km derinlikte meydana gelebilecek olası depremlerin çok değişkenli LSTM modeli kullanılarak tahmin edilen deprem büyüklüklerinin (M_w), İstanbul ilçeleri üzerindeki mekânsal dağılımı görselleştirilmiştir. Harita, her ilçeye ait enlem ve boylam koordinatları kullanılarak oluşturulan tahminlere dayalı olarak hazırlanmış ısı haritasıdır.

Görseldeki renk yoğunluğu, deprem büyüklüğüne ilişkin tahmini değerleri temsil etmektedir. Mavi tonlar daha düşük büyüklükleri, yeşil-sarı tonlar orta büyüklükleri ve kırmızı tonlar daha yüksek deprem büyüklüklerini göstermektedir. Şekil 27'a göre, İstanbul'un Avrupa Yakası'nda özellikle Bağcılar, Bahçelievler, Bakırköy, Zeytinburnu ve çevresinde deprem büyüklükleri tahminleri diğer ilçelere kıyasla daha yüksek seviyelerde olmaktadır. Bu ilçelerin kırmızı renkte olması modelin bu alanlarda sismik aktivite olarak daha aktif olabileceğini öngörmektedir. Buna karşılık, kentin kuzeyinde ve doğusunda yer alan Sarıyer, Çekmeköy, Sancaktepe, Şile, Arnavutköy gibi ilçelerde tahmin edilen deprem büyüklüğü diğer ilçelere nazaran daha düşük seviyelerde kalmakta ve mavi tonlarla ifade edilmektedir. Bu durum, modelin bu bölgelerde yer sarsıntılarının daha düşük büyüklükte olabileceğini göstermektedir. Harita üzerinden elde edilen sonuçlar, özellikle Marmara Denizi'ne kıyısı olan ilçelerde sismik riskin göreceli olarak yüksek olduğunu ortaya koymakta ve bu alanlarda afet yönetimi açısından öncelikli planlamaların yapılmasının önemini vurgulamaktadır.



Şekil 28: Derinlik 10 km. Olduğunda Tahmin Edilen Deprem Büyüklükleri

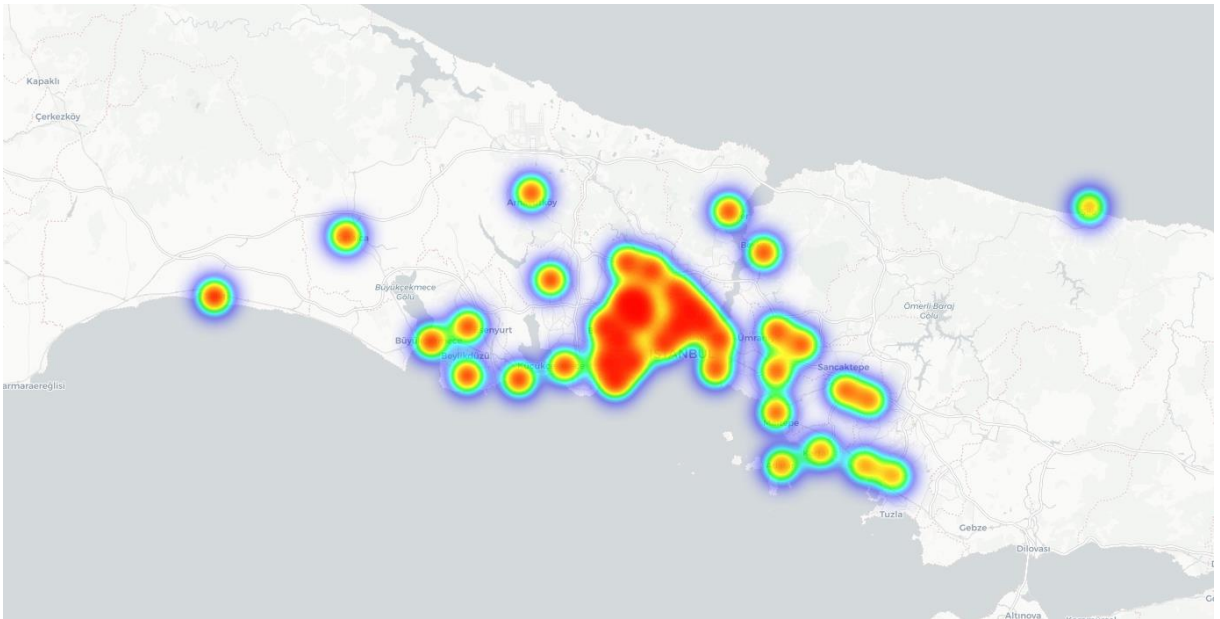
Şekil 28’de yer alan grafik, her bir İstanbul ilçesi için model tarafından tahmin edilen deprem büyüklüğü değerlerini göstermektedir. İlçeler, tahmin edilen Mw değerlerine göre yüksekten düşüğe doğru sıralanmıştır.

En yüksek deprem büyüklüğüne sahip ilçeler: Esenyurt, Gaziosmanpaşa, Beşiktaş, Bayrampaşa, Şişli gibi ilçelerdir. Bu ilçeler genellikle İstanbul’un Avrupa Yakası’nda, yerleşimin yoğun olduğu, zemin sertliğinin yüksek ve yapı stoğunun eski olduğu merkezi alanlarda yer almaktadır. Bu özellikler, modelin bu bölgelerdeki sismik potansiyeli daha yüksek öngörmesine neden olmaktadır.

Orta düzey deprem büyüklüğüne sahip ilçeler: Kadıköy, Maltepe, Üsküdar, Ümraniye, Sancaktepe gibi ilçeler hem Anadolu hem de Avrupa Yakası’nda yer almaktadır. Bu ilçeler, fay hattına doğrudan yakın olmayan, jeolojik yapısı bakımından dengeli olan bölgeler arasında yer almakta ve model tarafından orta seviyede deprem büyüklüğü riskiyle ilişkilendirilmektedir.

En düşük deprem büyüklüğüne sahip ilçeler: Çatalca, Silivri, Büyükçekmece, Adalar, Şile gibi ilçelerdir ve bu ilçeler İstanbul’un kenar bölgelerinde ve özellikle Marmara fay hattına görece olarak uzak konumlarda yer almasından kaynaklı, düşük yapılaşma yoğunluğu ve farklı zemin karakterleri sebebiyle model tarafından daha düşük deprem büyüklüğü ile tahminlenmiştir.

Şekil 29’da, İstanbul ilçeleri için 30 km derinlikte deprem tahminlerini gösteren ısı haritası verilmiştir.



Şekil 29: İstanbul için 30 km Derinlikte Isı Haritası

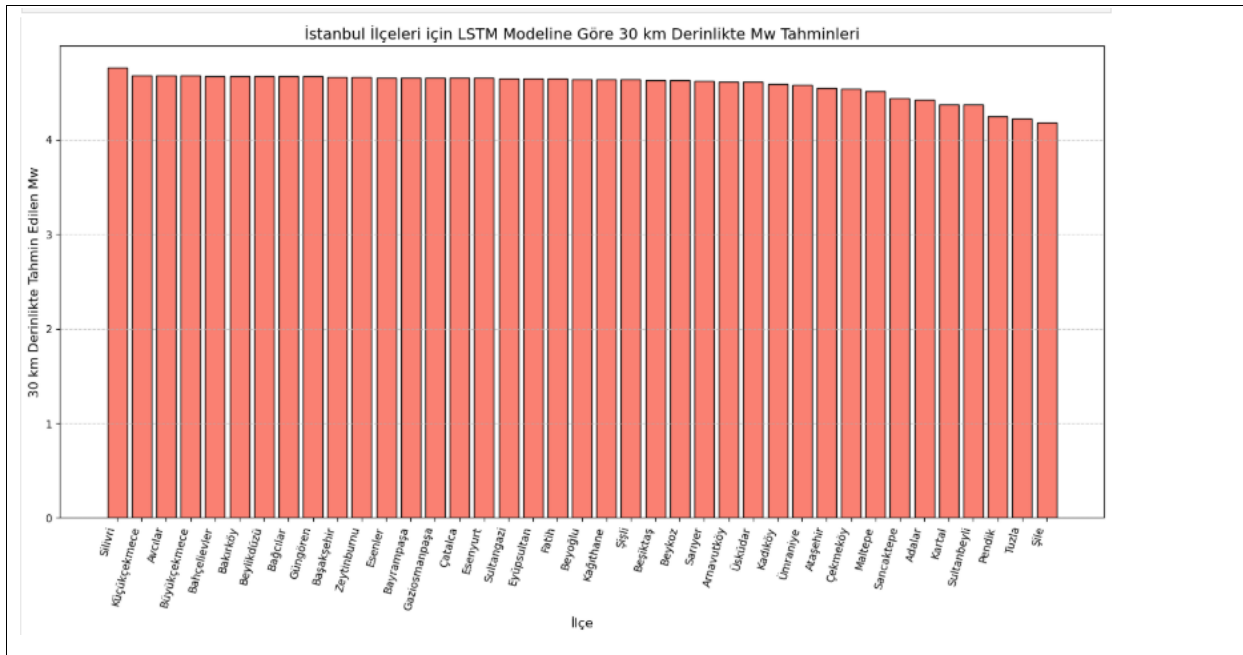
Şekil 29'daki ısı haritası İstanbul ilçelerinde 30 km derinlikte meydana gelebilecek deprem büyüklüklerini LSTM modeliyle tahmin etmiş ve bunu görsel olarak ifade etmiştir.

En yüksek deprem büyüklüğü tahminine sahip ilçeler: haritada kırmızı ile gösterilen alanlar Esenyurt, Gaziosmanpaşa, Beşiktaş, Bayrampaşa ve Şişli ilçeleri olmaktadır. Bu ilçeler genellikle İstanbul'un Avrupa Yakası'nda, yerleşimin yoğun olduğu, zemin sertliğinin yüksek ve yapı stoğunun eski olduğu merkezi alanlarda konumlanmaktadır. Bu fiziksel ve demografik özellikler dolayısıyla, model bu bölgelerde sismik potansiyeli daha yüksek görmektedir.

Orta büyüklükte deprem büyüklüğü tahminine sahip ilçeler: Kadıköy, Maltepe, Üsküdar, Ümraniye ve Sancaktepe gibi ilçeler, hem Anadolu hem de Avrupa Yakası'nda yer alması sebebiyle model tarafından orta seviye deprem büyüklüğü riskiyle ilişkilendirilmektedir. Bu ilçeler, jeolojik yapının görece dengede olması, yapı yoğunluğundaki kontrollülük ve Marmara fay hattına doğrudan yakın olmamaları sebebiyle tahminde etkili olmaktadır.

En düşük deprem büyüklüğü tahminine sahip ilçeler: haritada mavi ile gösterilen Çatalca, Silivri, Büyükçekmece, Adalar ve Şile ilçeleri İstanbul'un kenar bölgelerinde yer almakta ve Marmara fay hattına göre daha uzak konumda olmaları sebebiyle ilçelerdeki düşük yapılaşma yoğunluğu, farklı zemin yapısı ve jeolojik olarak daha sabit özellikler göstermesi nedeniyle model, bu bölgelerdeki deprem büyüklüğünü daha düşük tahmin etmiştir.

Şekil 30'da İstanbul'un tüm ilçeleri için 30 km derinlikte tahmin edilen ortalama deprem büyüklüğü Mw değerleri verilmiştir.



Şekil 30: Derinlik 30 km. Olduğunda Tahmin Edilen Deprem Büyüklükleri

En yüksek deprem büyüklüğü tahminine sahip ilçeler: Silivri, Küçükçekmece, Avcılar, Arnavutköy ve Büyükçekmece gibi ilçeler öne çıkmaktadır. Bu ilçelerin Marmara Denizi kıyısında ve fay hattına yakın bölgelerde yer alması, modelin bu alanlarda sismik aktivite potansiyelini daha yüksek olarak değerlendirmesine neden olmuştur.

Orta düzeyde deprem büyüklüğü tahminine sahip olan ilçeler: Bayrampaşa, Kağıthane, Esenyurt, Eyüpsultan, Pendik ve Beşiktaş gibi yerleşimlerdir. Bu ilçeler hem Avrupa hem de Anadolu Yakası'nda bulunmakta olup, fay hattına olan uzaklıklar bakımından orta seviyede risk taşımaktadır. Bu durum, modelin bu bölgelerdeki jeolojik yapı ve geçmiş sismik aktiviteleri dikkate alarak daha ılımlı tahminler yapmasına yol açmaktadır.

En düşük deprem büyüklüğüne sahip olan ilçeler: Şişli, Tuzla, Çekmeköy, Sultanbeyli ve Adalar yer almaktadır. Bazı ilçelerin yerleşim yoğunluğunun görece düşük olması ve yerel zemin özelliklerinin sismik dalga yayılımını azaltması gibi faktörler, modelin bu bölgeler için düşük Mw değerleri tahmin etmesine neden olmuş olabilmektedir.

Çizelge 5: İstanbul İlçeleri için 10 km Derinlikteki Deprem Tahminleri

	İlçe	10 km Derinlikte Tahmin Edilen Mw
0	Adalar	4.744
1	Arnavutköy	4.465
2	Ataşehir	4.874
3	Avcılar	4.578
4	Bağcılar	5.016
5	Bahçelievler	4.996
6	Bakırköy	4.990
7	Başakşehir	4.656
8	Bayrampaşa	5.027
9	Beşiktaş	4.961
10	Beykoz	4.951
11	Beylikdüzü	4.444
12	Beyoğlu	4.984
13	Büyükçekmece	4.340
14	Çatalca	4.122
15	Çekmeköy	4.871
16	Esenler	5.038
17	Esenyurt	4.420
18	Eyüpsultan	5.059
19	Fatih	4.989
20	Gaziosmanpaşa	5.040
21	Güngören	5.012
22	Kadıköy	4.921
23	Kağıthane	5.009
24	Kartal	4.711
25	Küçükçekmece	4.830
26	Maltepe	4.821
27	Pendik	4.652
28	Sancaktepe	4.774
29	Sarıyer	5.013
30	Şile	4.890
31	Silivri	4.236
32	Şişli	4.984
33	Sultanbeyli	4.741
34	Sultangazi	5.031
35	Tuzla	4.612
36	Ümraniye	4.892
37	Üsküdar	4.934
38	Zeytinburnu	4.999

Çizelge 6'da İstanbul ilçeleri için 30 km derinlikteki Mw tahminleri verilmiştir.

Çizelge 6: İstanbul İlçeleri için 30 km Derinlikteki Deprem Tahminleri

	İlçe	30 km Derinlikte Tahmin Edilen Mw
0	Adalar	4.420
1	Arnavutköy	4.615
2	Ataşehir	4.548
3	Avcılar	4.677
4	Bağcılar	4.667
5	Bahçelievler	4.673
6	Bakırköy	4.668
7	Başakşehir	4.661
8	Bayrampaşa	4.657
9	Beşiktaş	4.631
10	Beykoz	4.628
11	Beylikdüzü	4.668
12	Beyoğlu	4.641
13	Büyükkçekmece	4.675
14	Çatalca	4.652
15	Çekmeköy	4.539
16	Esenler	4.658
17	Esenyurt	4.651
18	Eyüpsultan	4.646
19	Fatih	4.646
20	Gaziosmanpaşa	4.656
21	Güngören	4.667
22	Kadıköy	4.591
23	Kağıthane	4.641
24	Kartal	4.374
25	Küçükçekmece	4.683
26	Maltepe	4.514
27	Pendik	4.249
28	Sancaktepe	4.437
29	Sarıyer	4.620
30	Şile	4.177
31	Silivri	4.758
32	Şişli	4.637
33	Sultanbeyli	4.371
34	Sultangazi	4.647
35	Tuzla	4.219
36	Ümraniye	4.579
37	Üsküdar	4.612
38	Zeytinburnu	4.660

Çizelge 7: İstanbul İlçeleri İçin Basit Doğrusal Regresyon Modeli Sonuçları

Modelin Özelliği	Değer
Bağımlı Değişken	Mw
R^2	0,2535
F-İstatistiği	25,81
Gözlem Sayısı	78
Sabit (cons) Katsayısı	4,9257
Depth Katsayısı	-0,0114
Depth t-istatistiği	-5,08
Depth p-değeri	2,6e-06
F-testi p-değeri	2,6e-06

Derinliğin, Mw üzerindeki etkisini test etmek amacıyla basit doğrusal regresyon modeli oluşturulmuş ve modelin istatistiksel anlamlılığı test edilmiştir. Modelin çıktıları Çizelge 7’de verilmiştir.

Bu model üzerinden hipotez kurulmuştur:

H_0 :Derinliğin Mw üzerinde anlamlı bir etkisi yoktur.

H_1 : Derinliğin Mw üzerinde anlamlı bir etkisi vardır.

İstanbul’un 39 ilçesi için 10 km ve 30 km derinlik düzeylerinde tahmin edilen Mw değerleri kullanılarak gerçekleştirilen basit doğrusal regresyon analizi sonuçları Çizelge 7’de verilmiştir. Çizelge 7 incelendiğinde derinliğin, Mw değişkeni üzerindeki etkisi istatistiksel olarak anlamlı bulunmuştur. Modelin R^2 değeri 0,2535 olarak hesaplanmış olup, bu değer Mw değişkeninin toplam varyansının yaklaşık olarak %25’inin derinlik değişkeni tarafından açıklandığını göstermektedir.

Kullanılan regresyon modelinde F-istatistiđi 25,81 olarak elde edilmiř ve buna karřı gelen p-deđeri 2,6e-06 seviyelerinde bulunmuřtur. Bu sonu, modelin genel anlamda istatistiksel olarak anlamlı olduđunu ve bađımsız deđiřken olarak derinliđin, Mw üzerindeki etkisinin tesadüfi bir sonu olmadığını göstermektedir. Derinliđe ait regresyon katsayısı -0,0114 olarak hesaplanmıř ve negatif iřaretili katsayının derinliđin artmasına bađlı olarak tahmin edilen Mw deđerinde istatistiksel olarak anlamlı bir azalma eđilimi olduđunu göstermiřtir. Derinliđin t-istatistiđi -5,08 olarak hesaplanmıř ve buna bađlı olarak p-deđer 2,6e-06 düzeyinde hesaplanmıřtır. Bu deđer %5 anlamlılık düzeyinin altında olup H_0 hipotezinin reddedilmesini ve H_1 hipotezinin kabul edilmesine olanak tanımaktadır. Yani derinliđin Mw deđiřkeni üzerinde anlamlı bir etkisinin olduđu sylenebilmektedir.

Elde edilen bulgular, İstanbul ilelerinde deprem derinliđinin Mw üzerinde anlamlı ve negatif ynlü bir etkiye sahip olduđunu ortaya koymaktadır. Dolayısıyla İstanbul zelinde deprem riski deđerlendirmelerinde derinlik parametresinin gz ardı edilmemesi gerektiđi ve Mw tahminlerinde derinlik bilgisinin modele dahil edilmesi gerektiđini belirtmektedir.

4. SONUÇ

Dünyamız, sanılanın aksine cansız bir kütle olmayıp, jeodinamik süreçlerin etkisiyle sürekli hareket halindedir. Bu hareketlilik, yer kabuğunun alt kısımlarında meydana gelen tektonik aktiviteler sonucunda kırılmalar oluşturarak enerji birikimine ve bu enerjinin açığa çıkmasıyla yüzeyde titreşimlere neden olmaktadır. Bu titreşimlerin dalgalar halinde yayılarak yeryüzünü sarsması, deprem olarak tanımlanmaktadır.

Türkiye, jeolojik konumu itibariyle Alp-Himalaya Deprem Kuşağı üzerinde yer almakta olup, bu kuşağın Doğu Akdeniz Bölgesi'ndeki en karmaşık yapıya sahip bölümünü oluşturmaktadır. Bu karmaşıklık, çeşitli boyutlarda ve farklı yönlerde hareket eden levhaların bir arada bulunmasından kaynaklanmaktadır. Tarihsel ve istatistiksel veriler, Türkiye'de deprem afetinin ciddi can ve mal kaybına neden olduğunu ortaya koymuş; bu da afet yönetimi ve deprem tahmini çalışmalarının bilimsel olarak ele alınmasını zorunlu kılmıştır.

Bu tez çalışmasında, 1900-2021 yılları arasında Ege Bölgesi'nde meydana gelen $M_w \geq 3.0$ büyüklüğündeki toplam 16504 deprem verisi kullanılarak, öncelikle deprem büyüklüğü (M_w) üzerine etkili faktörler regresyon ve zaman serisi modelleri ile analiz edilmiş, ardından kantil regresyon ve makine öğrenmesi yöntemleri ile tahminleme yapılmıştır. Bu analizler neticesinde en iyi performansı sergileyen LSTM modeli, İstanbul ilçelerine ait M_w tahminlerinin yapılmasında kullanılmıştır.

Analizlerde bağımsız değişkenler olarak derinlik (Depth (km)), enlem (Lat) ve boylam (Long); bağımlı değişken olarak ise deprem büyüklüğü (M_w) kullanılmıştır. M_w değişkeninin ortalaması 3.5 olup veriler 3.0 ile 7.6 arasında dağılım göstermektedir. Ortalama derinlik 16.6 km olup, depremlerin genellikle sığ karakterli olduğu sonucuna ulaşılmıştır. Ayrıca Guttenberg-Richter yasasına paralel olarak, küçük depremlerin sık büyük depremlerin ise seyrek görüldüğü sonucuna varılmıştır.

Zaman serisi grafikleri incelendiğinde, 1960 yılından itibaren deprem kayıtlarında belirgin bir artış olduğu ancak bu dönemde M_w büyüklüklerinin daha düşük değerlerde yoğunlaştığı gözlemlenmiştir.

Çoklu doğrusal regresyon analizinde bağımsız değişkenler olarak derinlik, enlem, boylam ve all clusters değişkenleri modele dahil edilmiş; ancak all clusters değişkeni istatistiksel olarak anlamlı bulunmadığı için İstanbul ilçeleri için yapılan analizlere dahil edilmemiştir.

Bu çalışmada, İstanbul ilçelerine ait deprem büyüklüğü tahmininde kullanılan modellerin performansları değerlendirilmiş ve özellikle derinlik (Depth (km)), deprem büyüklüğü (Mw) değişkeni üzerindeki etkisi incelenmiştir. Çalışmada hem klasik istatistiksel yaklaşımlar hem de LSTM tabanlı derin öğrenme yöntemleri kullanılmıştır.

İstanbul ilçelerine yönelik olarak derinlik değişkeninin Mw üzerindeki etkisini incelemek amacıyla kurulan basit doğrusal regresyon modelinde R^2 değeri 0,2535 olarak hesaplanmış olup, bu değer Mw değişkeninin toplam varyansının yaklaşık olarak %25'inin derinlik değişkeni tarafından açıklandığını göstermektedir. Regresyon katsayısı -0,0114, sabit terim ise 4,9257 olarak hesaplanmıştır. Negatif katsayı işareti, derinlik arttıkça Mw tahminlerinin istatistiksel olarak anlamlı bir şekilde azalma eğilimi gösterdiğini belirtmektedir. Modelin F-istatistiği 25,81 düzeyinde hesaplanmış ve anlamlılık düzeyi $p=2,6e-06$ olarak gerçekleşmiştir. Bu sonuçlar, derinliğin Mw değişkeni üzerindeki etkisinin tesadüfi olmadığını ve istatistiksel olarak anlamlı olduğunu ortaya koymakta ve H_0 hipotezinin reddedilmesini ve H_1 hipotezinin kabul edilmesine olanak tanımaktadır.

Bunun yanı sıra, derin öğrenme tabanlı bir LSTM modeli kullanılarak İstanbul ilçelerinin enlem-boylam bilgileri ve sabit derinlik değerleri ile Mw tahminleri gerçekleştirildi. LSTM modeli, zaman serisine uygun mimarisi sayesinde karmaşık yapıları öğrenebilmiş ve elde edilen hata metrikleri incelendiğinde model performansının güçlü olduğunu belirtmiştir. Modelde MAE=0,4572 ve RMSE=0,6432 olarak hesaplanmıştır. Bu değerler LSTM modelinin İstanbul ilçeleri için deprem büyüklüğü tahmininde kabul edilebilir düzeyde bir doğruluk sağladığını göstermektedir. LSTM modeli sonuçları incelendiğinde, 10 km derinlikte en yüksek Mw tahmin değerlerinin Esenyurt, Gaziosmanpaşa, Beşiktaş, Bayrampaşa ve Şişli gibi merkezi, zemin sertliğinin yüksek olduğu ve yoğun yapılaşmaya sahip ilçelerde olduğu belirlenmiştir.

Öte yandan kantil regresyon analizleri Mw dağılımlarının farklı kantillerinde derinliğin etkisini ayrı ayrı modellemiş ve özellikle üst kantillerde derinliğin artışa bağlı olarak Mw değerlerinde artış eğiliminde olduğu gözlemlenmiştir. Bu bulgu, Mw'nin ortalama düzeyinde derinliğin negatif etkisi olmasına rağmen, yüksek Mw değerlerinde derinliğin pozitif yönde etkili olabileceğine işaret etmektedir. Bu durum, Mw ve derinlik arasındaki ilişkinin homojen ve tekdüze olmadığını, dağılımın farklı noktalarında derinliğin etkisinin yön ve büyüklük açısından değişkenlik gösterebileceğini ortaya koymaktadır.

Doğrusal regresyon analizindeki negatif katsayı, ilçeler genelinde derinlik arttıkça Mw ortalamasının azalma eğiliminde olduğunu gösterirken LSTM modellerinde bazı merkezi ilçelerde derinliğin artışına bağlı olarak Mw değerinde yükselme görülmektedir. Bu farklılık, LSTM modellerinin karmaşık ve doğrusal olmayan örüntüleri öğrenme kabiliyeti sayesinde ilçeler bazında değişken ilişkileri yakalayabilmesine ancak doğrusal regresyonun ortalamaya dayalı tek bir genel eğilimi temsil etmesine bağlanmaktadır.

Sonuç olarak, Mw ve derinlik arasındaki ilişki İstanbul ilçelerinde bölgesel farklılıklar gösterebilmekte ve bu nedenle Mw tahmin modellerinde derinliğin etkisinin sadece ortalama değerlerle değil, kantil ve bölgesel bazda da analiz edilmesi, modelleme doğruluğu ve sismik risk değerlendirmeleri açısından kritik öneme sahiptir.

KAYNAKLAR

- Abadie, A., Angrist, J., & Imbens, G. (2002). Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. *Econometrica*, 70(1), 91–117.
- Afet ve Acil Durum Yönetimi Başkanlığı (AFAD). (t.y.). *Deprem nedir?*. <https://www.afad.gov.tr/deprem-nedir>
- Afet ve Acil Durum Yönetimi Başkanlığı (AFAD). (2023). <https://www.afad.gov.tr>
- Ahmed, N. K., Atiya, A. F., El Gayar, N., & El-Shishiny, H. (2010). An empirical comparison of machine learning models for time series forecasting. *Econometric Reviews*, 29, 594–621. <https://doi.org/10.1080/07474938.2010.481556>
- Akgül, I. (1994). Zaman serisi analizi ve öngörü modelleri. *Öneri*, 1(1), 52–69. <https://doi.org/10.14783/MARUONERI.698511>
- Aki, K. (1965). Maximum likelihood estimate of b in the formula $\log N = a - bM$ and its confidence limits. *Bulletin of the Earthquake Research Institute, University of Tokyo*, 43, 237–239.
- Akkurt, N., & Hasgül, S. (2024). Zaman serisi tahminlemede otomatikleşmiş makine öğrenmesi (AutoML) kütüphanelerinin karşılaştırılması. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 39(3), 1693–1701.
- Akküçük, U. (2011). *Veri madenciliği: Kümeleme ve sınıflama algoritmaları*. Yalın Yayıncılık.
- Akyüz, H. E. (2019). Farklı istatistiksel dağılımlar ile deprem verilerinin modellenmesi ve uygunluklarının karşılaştırılması. *Academic Platform Journal of Engineering and Science*, 7(1), 7–13. <https://dergipark.org.tr/en/download/article-file/592619>
- Ali, Z., Hussain, I., Faisal, M., Nazir, H. M., Hussain, T., Shad, M. Y., & Hussain Gani, S. (2017). Forecasting drought using multilayer perceptron artificial neural network model. *Advances in Meteorology*, 2017, Article ID 5681303. <https://doi.org/10.1155/2017/5681303>
- Alpar R. (2010). Basit Doğrusal Regresyon Çözümlemesi. *Spor, Sağlık ve Eğitim Bilimlerinden Örneklerle Uygulamalı İstatistik ve Geçerlik Güvenirlik*. Detay Yayıncılık, Ankara, 285–304.
- Alsharif, A., Aggarwal, K., Sonia, Kumar, M., & Mishra, A. (2022). Review of ML and AutoML solutions to forecast time-series data. *Archives of Computational Methods in Engineering*, 29(7), 5297–5311. <https://doi.org/10.1007/s11831-021-09670-4>
- Altındağ, İ. (2010). *Kantil regresyon ve bir uygulama*. (Yüksek Lisans Tezi). Selçuk Üniversitesi.
- Altman, N., & Krzywinski, M. (2015). Simple linear regression. *Nature Methods*, 12(11), 999–1000. <https://doi.org/10.1038/nmeth.3627>
- Ambraseys, N. N. (2000). The seismicity of the Sea of Marmara (Turkey) since 1500. *Geophysical Journal International*, 141(3), F1–F11.
- Ambraseys, N. N., & Finkel, C. F. (1990). *The Marmara Sea earthquake of 1509*. University of London, Department of Civil Engineering. Ayrıca: Ambraseys, N. N., & Finkel, C. F. (2001).

The earthquake of 1509 in the Sea of Marmara, Turkey, revisited. *Bulletin of the Seismological Society of America*, 91(6), 1397–1416

Anonim. (2023). *Deprem kronolojisi*. Wikipedi. <https://tr.wikipedia.org/wiki/Deprem> (Erişim tarihi: 1 Eylül 2023)

Anonim. (t.y.). *Deprem nedir?*. <http://www.koeri.boun.edu.tr/sismo/bilgi/deprenedir/index.htm> (Erişim tarihi: 1 Eylül 2023)

Aras, S., & Kocakoş, İ. D. (2016). A new model selection strategy in time series forecasting with artificial neural network: IHTS. *Neurocomputing*, 174, 974–987. <https://doi.org/10.1016/j.neucom.2015.10.019>

Audi, R. (Ed.). (1996). Curve fitting problem. In *The Cambridge Dictionary of Philosophy* (pp. 172–173). Cambridge University Press.

Ay, Ş., & Ekinci, E. (2022). Ozon konsantrasyonlarını modellemek için makine öğrenmesi ve derin öğrenme yöntemlerinin karşılaştırılması. *Zeki Sistemler Teori ve Uygulamaları Dergisi*, 5(2), 106–118.

Ayodele, T. O. (2010). *Machine learning overview*. Intech Open Access Publisher.

Berhich, A., Belouadha, F.-Z., & Kabbaj, M. I. (2023). LSTM tabanlı deprem tahmin modelleri. *ResearchGate*. https://www.researchgate.net/publication/341635348_LSTM-based_Models_for_Earthquake_Prediction

Bhatia, A., Pasari, S., & Mehta, A. (2018). Earthquake forecasting using artificial neural networks. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-5.

Box, G. P. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32 <https://doi.org/10.1023/A:1010933404324>

Brownlee, J. (2020). *Introduction to time series forecasting with Python*. Machine Learning Mastery.

Can, M. (2009) İşletmelerde zaman serileri analizi ile tahmin (Yayımlanmamış doktora tezi). İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü.

Canitez, M. A., & Savaş, S. (2022). Kripto para piyasa değeri tahmini için özellik tabanlı LSTM ve ARIMA karşılaştırması. In *3rd International Conference on Applied Engineering and Natural Science* (pp. 1311–1317). Konya.

Cameron, A. C., & Trivedi, P. K. (2009). *Microeconometrics using Stata*. Stata Press.

Caruana, R., & Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. In *Proceedings of the 23rd International Conference on Machine Learning* (pp. 161–168). ACM.

- Chen, T., & 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. <https://doi.org/10.1145/2939672.2939785>
- Chao, W. L. (2011). Machine Learning Tutorial
- Chen, K. Y., & Wang, C. H. (2007). Support vector regression with genetic algorithms in forecasting tourism demand. *Tourism Management*, 28(1), 215–226. <https://doi.org/10.1016/j.tourman.2005.12.024>
- Chen, C., & Wei, Y. (2005). Computational issues for quantile regression. *The Indian Journal of Statistics*, 67(2), 399–417.
- Cook, D., Wang, W., & Wang, E. (2017). *Package: quankar – Quantile regression outlier diagnostics with K left out analysis* [R package manual]. <https://cran.r-project.org/web/packages/quokar/quokar.pdf> (Erişim tarihi: 22 Eylül 2018)
- Couronné, R., Probst, P., & Boulesteix, A. L. (2018). Random forest versus logistic regression: A large-scale benchmark experiment. *BMC Bioinformatics*, 19(1), 1–14. <https://doi.org/10.1186/s12859-018-2264-5>
- Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random forests. In C. Zhang & Y. Ma (Eds.), *Ensemble machine learning* (pp. 157–175). Springer.
- Crone, S., & Nikolopoulos, K. (2007). Result of the NN3 neural network forecasting competition. In *The 27th International Symposium on Forecasting, Program* (p. 129).
- Çağlayan, E., & Arıkan, E. (2011). Determinants of house prices in Istanbul: A quantile regression approach. *Quality & Quantity*, 45(2), 305–317.
- Çelik, Ş. (2013). Sert kabuklu meyvelerin üretim miktarının Box-Jenkins tekniği ile modellenmesi. *Yüziüncü Yıl University Journal of Agricultural Sciences*, 23(1), 18–30. <https://dergipark.org.tr/en/pub/yyutbd/issue/21973/235937>
- Çevik, M., Cihan, M., Yılmaz, N., Konak, M., Soylu, H., & Ceylan, M. (2023). Yenidoğan laboratuvar bulgularında LSTM tabanlı derin öğrenme ağı ile zaman serileri analizi. *International Journal of Multidisciplinary Studies and Innovative Technologies*, 7(2), 68–73.
- Çınar, U. K. (2019). En küçük kareler regresyonuna alternatif bir yöntem: Kantil regresyon. *Avrasya Uluslararası Araştırmalar Dergisi*, 7(18), 57–71.
- Çiftçi, C., & Kangallı Uyar, S. G. (2015). Eğitim ve gelir. *Ege Akademik Bakış*, 15(1), 141–152.
- Çoban, Ö. (2016). *Metin sınıflandırma teknikleri ile Türkçe Twitter duygu analizi*. (Yüksek Lisans Tezi). Atatürk Üniversitesi.
- Davidson, R., & MacKinnon, J. G. (1993), Estimation and inference in econometrics. Oxford University Press
- Dawson, B., & Trapp, R. G. (2001). Statistical methods for multiple variables. In *Basic & Clinical Biostatistics* (pp. 236–242). Lange Medical Books/McGraw-Hill Medical Publishing Division.

- Demirelli, E., Solak, H. İ., & Tiryakioğlu, İ. (2023). Makine Öğrenmesi Algoritmaları ile Deprem Katalogları Kullanılarak Deprem Tahmini. *Gümüşhane University Journal of Science*, 13(4), 979–989.
- Demirezen, M. U., Civrizoğlu Buz, A., & Yavanoğlu, U. (2021). Sualtı objelerinin makine öğrenmesi yöntemleri ile tespitinde zaman serisi–görüntü dönüşümü tabanlı yeni yaklaşımlar. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 36(3), 1399–1415.
- Dietterich, T. G. (1997). Machine-learning research. *AI Magazine*, 18(4), 97.
- Doğan, A. (2023). Makine öğrenimi yöntemleri kullanılarak Türkiye'nin kuzeybatısı için deprem tahmini. *Yerbilimleri*, 44(2), 166–178.
- Doğan, A., & Demir, M. (2020). Deprem tahmin modelleri ve uygulamaları. *Afet ve Risk Dergisi*, 5(2), 123–135. <https://dergipark.org.tr/tr/download/article-file/1234567>
- Doğruyol, M. (2020). Bingöl ili deprem risk analizi. *Erzincan Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 13(2), 568–579. <https://doi.org/10.18185/erzifbed.688071>
- Doğruyol, M. (2021). Siirt ili deprem tehlike analizi. *Doğal Afetler ve Çevre Dergisi*, 7(1), 145–158. <https://dacd.artvin.edu.tr/tr/download/article-file/990715>
- Edemen, M., ve arkadaşları. (2023). Deprem nedir? Nasıl oluşur? Türkiye’de oluşmuş depremler ve etkileri nelerdir? Depremlere karşı alınabilecek tedbirler hususunda öneriler. *International Journal of Social and Humanities Sciences Research (JSHR)*, 10(93), 719–734.
- Eğitim.com. (2023). *Dünyanın katmanları ve özellikleri*. <https://egitim.com/dunyanin-katmanlari>
- Ekinci, E., Omurca, İ. S., & Sevim, S. (2020). Improve offensive language detection with ensemble classifiers. *International Journal of Intelligent Systems and Applications in Engineering*, 8(2), 109–115.
- Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A*, 222, 309-368. <https://doi.org/10.1098/rsta.1922.0009>
- Finkel, C. F., & Ambraseys, N. N. (1997). The Marmara Sea Earthquake of 10 July 1894 and its effects on historic buildings. *Anatolia Moderna*, 7(1), 49–65.
- Foote, K. D. (2019). A brief history of machine learning. *Dataiversity*. <https://www.dataversity.net/>
- Frohlich, C., & Davis, S. D. (1993). Teleseismic b-values; Or, much ado about 1.0. *Journal of Geophysical Research: Solid Earth*, 98(B1), 631–644. <https://doi.org/10.1029/92JB01891>
- Gardner, M. W., & Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron): A review of applications in the atmospheric sciences. *Atmospheric Environment*, 32(14–15), 2627–2636. [https://doi.org/10.1016/S1352-2310\(97\)00447-0](https://doi.org/10.1016/S1352-2310(97)00447-0)
- Gentleman, R., Huber, W., & Carey, V. J. (2008). Supervised machine learning. In *Bioconductor case studies* (pp. 121–136). Springer.
- Genç, F. N. (2007). Türkiye’de doğal afetler ve doğal afetlerde risk yönetimi. *Stratejik Araştırmalar Dergisi*, 9(5), 201–226.

- Gökalp, Ö. M. (2022). *Makine öğrenmesi – Machine learning*. Gazi University.
- Gujarati, D. (2008). *Temel ekonometri* (Ü. Şenesen & G. G. Şenesen, Çev.). Literatür Yayınları. (Orijinal eser 2004 yılında yayımlanmıştır)
- Gujarati, D. N., & Porter, D. C. (2012). *Temel ekonometri* (Ü. Şenesen & G. G. Şenesen, Çev., 6. bs.). Literatür Yayıncılık.
- Gunn, S. R. (1998). Support vector machines for classification and regression. *ISIS Technical Report*, 14(1), 5–16.
- Gutenberg, B., & Richter, C. F. (1954). *Seismicity of the Earth and Associated Phenomena*. Princeton University Press.
- Hayashi, F. (2000). *Econometrics*. Princeton University Press.
- He, K., Yu, L., & Lai, K. K. (2012). Crude Oil Price Analysis and Forecasting Using Wavelet Decomposed Ensemble Model, *Energy*, 2012; 46(1):564e74
- Hiçyılmaz, M., & Sezer, S. (2020). Afyonkarahisar ili için Poisson ve Gumbell dağılım modelleri ile deprem tehlike analizi. *Uludağ Üniversitesi Mühendislik Fakültesi Dergisi*, 25(1).
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hunter, D. R., & Lange, K. (2000). Quantile regression via an MM algorithm. *Journal of Computational and Graphical Statistics*, 9(1), 60–77.
- İnangu, A., & Kırbaş, H. (1999). Anadolu levhası üzerinde Kütahya'nın deprem tehlikesi. *Dumlupınar Üniversitesi Fen Bilimleri Dergisi*, 1.
- İşçi, C. (2008). Deprem nedir ve nasıl korunuruz? *Yaşar Üniversitesi Dergisi*, 3(9), 959–983. <https://dergipark.org.tr/tr/download/article-file/179169>
- Jakkula, V. (2006). Tutorial on Support Vector Machine (SVM). *School of EECS, Washington State University*.
- Karasu, S., Altan, A., Bekiros, S., & Ahmad, W. (2020). A new forecasting model with wrapper-based feature selection approach using multi-objective optimization technique for chaotic crude oil time series. *Energy*, 212, 118750. <https://doi.org/10.1016/j.energy.2020.118750>
- Karl Pearson, Yule, G. U., Blanchard, N., & Lee, A. (1903). The law of ancestral heredity. *Biometrika*.
- Kazem, A., Sharifi, E., Hussain, F. K., Saberi, M., & Hussain, O. K. (2013). Support vector machines for time series forecasting. *Neurocomputing*, 55(1–2), 307–319.
- Kılıç, S. (2013). Doğrusal regresyon analizi. *Statistically Speaking*, 3(2), 90–92.
- Koenker, R. (2005). *Quantile Regression*. Cambridge University Press.
- Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica*, 46(1), 33–50.

- Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. *Informatica*, 31(3), 249–268.
- Koyuncugil, A., & Özgülbaş, N. (2009). Veri madenciliği: Tıp ve sağlık hizmetlerinde kullanımı ve uygulamaları. *Bilişim Teknolojileri Dergisi*, 2(2).
- Köseoğulları, B., & Ulucan, H. (2022). Education, income distribution and gender pay gap: The case of Turkey. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 49, 229–244.
- Köseoğulları, B., & Tunca, H. (2024). Türkiye’de cinsiyetlere göre eğitim ve gelir seviyesinin vücut kitle indeksine etkisi: Kantil regresyon yöntemi. *Alanya Akademik Bakış Dergisi*, 8(2), 357–375.
- Kuremoto, T., Kimura, S., Kobayashi, K., & Obayashi, M. (2014). Time series forecasting using a deep belief network with restricted Boltzmann machines. *Neurocomputing*, 137, 47–56. <https://doi.org/10.1016/j.neucom.2013.03.011>
- Kurtoğlu, F. (2011). *Kantil regresyon: Teorisi ve uygulamaları*. (Yüksek Lisans Tezi). Çukurova Üniversitesi.
- Li, D., Wang, S., & Li, D. (2015). Spatial data mining. *Springer*.
- Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: A survey. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2194), 20200209. <https://doi.org/10.1098/rsta.2020.0209>
- Makridakis, S., & Hibon, M. (1997). ARMA models and the Box–Jenkins methodology. *Journal of Forecasting*, 16, 147–163.
- McCluskey, W. J., McCord, M., Davis, P. T., Haran, M., & McIlhatton, D. (2013). Prediction accuracy in mass appraisal: A comparison of modern approaches. *Journal of Property Research*, 30(4), 239–265.
- Menezes Jose Maria, P., & Barreto, G. A. (2008). Long-term time series prediction with the NARX network: An empirical evaluation. *Neurocomputing*, 71(16–18), 3335–3345.
- MIT OpenCourseWare. (n.d.). *Quantile regression*. <https://ocw.mit.edu/> (Erişim tarihi: 10 Ekim 2023)
- Mockenhaupt, A. (2021). Maschinelles Lernen. In *Digitalisierung und Künstliche Intelligenz in der Produktion*. Wiesbaden: Springer Vieweg.
- Montenegro, C. (2001). Wage distribution in Chile: does gender matter? A quantile regression approach. *Policy Research Report on Gender and Development Working Paper No: 20*.
- Niu, T., Wang, J., Lu, H., Yang, W., & Du, P. (2020). Developing a deep learning framework with two-stage feature selection for multivariate financial time series forecasting. *Expert Systems with Applications*, 148, 113237.
- Ocak, B. (2010). *Dik ağırlıklandırmanın M-regresyon üzerine etkisi*. (Yüksek Lisans Tezi). Selçuk Üniversitesi.
- Öngün, M. (2023, Ocak). Çok değişkenli finansal zaman serisi analizinde yapay zekanın kullanımı. *YBS Ansiklopedisi*, 11(1).

Özel Kadılar, G. (Yürütücü). (2024). *Yapay Zeka ve Olasılıksal Model Tabanlı Deprem Tehlike Haritası* (Proje No: 121F208) [1001 TÜBİTAK Araştırma Projesi]. Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK).

Özdemir, M. O., & Çılgın, C. (2022). Buğday fiyatlarının öngörülenmesinde makine öğrenmesi ve zaman serisi tahmin modellerinin performanslarının karşılaştırılması. *Yayınlanmamış makale*.

Özgür, A. (2004). *Supervised and unsupervised machine learning techniques for text document categorization*. (Doktora Tezi). Boğaziçi Üniversitesi.

Öztürk, S., & Alkan, H. (2022). Van Gölü ve civarı (Türkiye) için deprem tahmini ve tehlikesi üzerine istatistiksel bir analiz ve değerlendirme. *Türk Deprem Araştırma Dergisi*, 4(2), 191–209. <https://doi.org/10.46464/tdad.1108730>

Pagano, M., & Gauvreau, K. (1993). Simple linear regression. In *Principles of Biostatistics* (pp. 379–424). Duxbury Press.

Pearson, K., Yule, G. U., Blanchard, N., & Lee, A. (1903). The law of ancestral heredity. *Biometrika*.

Qi, Y. (2012). Random forest for bioinformatics. In C. Zhang & Y. Ma (Eds.), *Ensemble machine learning* (pp. 307–323). Springer.

R.A. Fisher. (1922). The goodness of fit of regression formulae, and the distribution of regression coefficients. *Journal of the Royal Statistical Society*, 85, 597–612.

R.A. Fisher. (1925). *Statistical Methods for Research Workers*.

Reddy, D. S., & Prasad, P. R. C. (2018). Prediction of vegetation dynamics using NDVI time series data and LSTM. *Modeling Earth Systems and Environment*, 4(1), 409–419.

Roger Koenker & Gilbert Bassett. (1978). Regression quantiles. *Econometrica*, 46(1), 33–50.

Russel Davidson & James G. MacKinnon. (1993). *Estimation and Inference in Econometrics*. Oxford University Press.

Rüping, S. (2001). *SVM kernels for time series analysis* (Technical Report No. 2001–43). Universität Dortmund, Sonderforschungsbereich 475 – Komplexitätsreduktion in Multivariaten Datenstrukturen.

SAP. (2021). *Was ist maschinelles Lernen?* <https://www.sap.com/germany/insights/what-is-machine-learning.html>

Saçaklı, İ. (2005). *Kantil regresyon ve alternatif regresyon modelleri ile karşılaştırılması*. (Yüksek Lisans Tezi). Marmara Üniversitesi.

Sapankevych, N. I., & Sankar, R. (2009). Time series prediction using support vector machines: A survey. *IEEE Computational Intelligence Magazine*, 4(2), 24–38. <https://doi.org/10.1109/MCI.2009.932254>

Sarıcaoğlu, C. (2019). *Sözcüksel analizler kullanarak kötü niyetli URL'leri derin öğrenme teknikleri ile tespit etme*. (Yüksek Lisans Tezi). Gazi Üniversitesi.

- Saygılı, R. (2012). Dünya deprem haritası. *TÜBİTAK Bilim ve Teknik*. <http://www.biltek.tubitak.gov.tr>
- Savaş, S. (2022). Zaman serisi analizi yoluyla deneme sınavları üzerinden ortaokul öğrencilerinin LGS başarısının tahmini. <https://www.researchgate.net/publication/366158309>
- Smola, A. J. (1996). Regression estimation with support vector learning machines, Diplomingenieur's Thesis, Technische Universität München
- Şahin, A. (2023). İstanbul'da farklı kat yüksekliklerine sahip bitişik nizam yapılarında çekişme etkisi. *6. Uluslararası Fen Bilimleri ve İnovasyon Kongresi*.
- Şahin, C., & Sipahioğlu, Ş. (2002). *Doğal Afetler ve Türkiye*. Ankara: Gündüz Eğitim ve Yayıncılık.
- Seo, D. K., Kim, Y. H., Eo, Y. D., Park, W. Y., & Park, H. C. (2017). Generation of radiometric, phenological normalized image based on random forest regression for change detection. *Remote Sensing*, 9(11), 1163.
- Smola, A. J., & Schölkopf, B. (1998). On a kernel-based method for pattern recognition, regression, approximation and operator inversion. *Algorithmica*, 22(1), 211–231.
- Stander, J., Yu, K., & Lu, Z. (2013). Ticari serbestleşme–ekonomik büyüme ilişkisini bootstrap kantil regresyon yardımıyla analizi. *İstanbul Üniversitesi İktisat Fakültesi Mecmuası*, 62(2), 283–303.
- Stepchenko, A., 2016, “ NDVI Index Forecasting Using a Layer Recurrent Neural Network Coupled with Stepwise Regression and the PCA” pp 130-135
- Tan, O. (2021). A homogeneous earthquake catalogue for Turkey. *Natural Hazards and Earth System Sciences*, 21, 2059–2073. <https://doi.org/10.5194/nhess-21-2059-2021>
- Tanişman, S., Karcıoğlu, A. A., Uğur, A., & Bulut, H. (2021). LSTM Sinir Ağı ve ARIMA Zaman Serisi Modelleri Kullanılarak Bitcoin Fiyatının Tahminlemesi ve Yöntemlerin Karşılaştırılması, *Avrupa Bilim ve Teknoloji Dergisi Özel Sayısı*, 32, 514-520.
- Taud, H., & Mas, J. F. (2018). Multilayer perceptron (MLP). In M. T. C. Olmedo, M. Paegelow, J. F. Mas, & F. Escobar (Eds.), *Geomatic approaches for modeling land change scenarios* (pp. 451–455). Springer.
- T.C. Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı. (2023). *2023 Kahramanmaraş ve Hatay Depremleri Raporu*. <https://www.sbb.gov.tr/2023-kahramanmaras-ve-hatay-depremleriraporu>
- Tuncer, A. (2022). LSTM metodu kullanarak rüzgar hızının tahmin edilmesi. (Yüksek Lisans Tezi). Yıldız Teknik Üniversitesi.
- Türkiş, C. (2015). *Sağkalım analizinde kantil regresyon ve parametrik regresyon modellerinin performanslarının karşılaştırılması* (Yüksek Lisans Tezi). Adnan Menderes Üniversitesi.
- Türkmenoğlu, C., & Tantuğ, A. C. (2014). Sentiment analysis in Turkish media. In *International Conference on Machine Learning (ICML)*.
- Tyralis, H., & Papacharalampous, G. (2017). Variable selection in time series forecasting using random forest. *Algorithms*, 10(4), 114. <https://doi.org/10.3390/a10040114>

- Udelhoven, T. (2012). Long term data fusion for a dense time series analysis with MODIS and Landsat imagery in an Australian savanna. *Journal of Applied Remote Sensing*, 6(1), 063512. <https://doi.org/10.1117/1.JRS.6.063512>
- Ulucan, H., & Çiftçi, C. (2019). Türkiye’de üniversite eğitiminin ücretler üzerindeki etkisi. *Business and Economic Research Journal*, 10(3), 649–669.
- Utsu, T. (1971). Aftershocks and earthquake statistics (3): Analyses of the distribution of earthquakes in magnitude, time and space with special consideration to clustering characteristics of earthquake occurrence. *Journal of the Faculty of Science, Hokkaido University. Series VII, Geophysics*, 3(5), 379–441.
- Verleysen, M., & François, D. (2005). The curse of dimensionality in data mining and time series prediction. In *Lecture Notes in Computer Science* (Vol. 3512, pp. 758–770). Springer.
- Weeks, J. D., Lockner, D. A., & Byerlee, J. D. (1978). Laboratory observations of seismicity. *Bulletin of the Seismological Society of America*, 68(6), 1601–1611.
- Wesolowski, M., & Suchacz, B. (2012). Artificial neural networks: Theoretical background and pharmaceutical applications: A review. *Journal of AOAC International*, 95(3), 652–668.
- Wikipedi yazarları. (2023). 2023 Kahramanmaraş depremleri. *Wikipedi*. https://tr.wikipedia.org/wiki/2023_Kahramanmara%C5%9F_depremleri (Erişim tarihi: 21 Mayıs 2025)
- Wikipedi yazarları. (2023). 1976 Tangshan Depremi. *Wikipedi*. https://tr.wikipedia.org/wiki/1976_Tangshan_depremi
- Wooldridge, J. M. (2013). *Introductory Econometrics: A Modern Approach* (5. Baskı). Cengage Learning.
- Yadav, A., Jha, C. K., & Sharan, A. (2020). Optimizing LSTM for time series prediction in Indian stock market. *Procedia Computer Science*, 167, 2091–2100.
- Yavuz, A., & Aşık, E. (2017). Kantil regresyon. *Uluslararası Mühendislik Araştırma ve Geliştirme Dergisi*, 9(2), 138–146.
- Yu, K., Chen, C. W. S., Reed, C., & Dunson, D. B. (2013). Bayesian variable selection in quantile regression. *Statistics and Its Interface*, 6(2), 261–274. <https://doi.org/10.4310/SII.2013.v6.n2.a9>
- Yu, L., Wang, S., & Lai, K. K. (2008). Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. *Energy Economics*, 30(5), 2623–2635. <https://doi.org/10.1016/j.eneco.2008.01.007>
- Yu, P. S., Chen, S. T., & Chang, I. F. (2006). Support vector regression for real-time flood stage forecasting. *Journal of Hydrology*, 328(3–4), 704–716. <https://doi.org/10.1016/j.jhydrol.2006.01.021>
- Yule, G.U. (1897). On the Theory of Correlation. *Journal of the Royal Statistical Society*, 60(4), 812–854. <https://doi.org/10.2307/2979746>