



Büyük Sismik Veriler Üzerinde Zaman ve Frekans Tabanlı Özniteliklerin Gerçek Deprem Verilerinin Tespitindeki Etkisi

The Effect of Time and Frequency Based Features on Large Seismic Data in Detecting Real Earthquake Data

Ali Narin , Yunus Emre Erdoğan

Zonguldak Bülent Ecevit Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, Zonguldak, Türkiye

Öz

Depremler, yer kabuğundaki ani hareketler sonucu meydana gelen doğal afetlerdir. Bu olayların hızlı bir şekilde tespit edilmesi, can ve mal kaybının en aza indirilmesi açısından hayati öneme sahiptir. Erken ve doğru tespit, acil müdahale ekiplerinin olay yerine zamanında ulaşmasını sağlayarak halkın güvenliğini artıran önlemlerin hızla alınmasına olanak tanır. Bu çalışmada, sismik sinyaller üzerinden elde edilen özniteliklerin kullanımıyla deprem ve çevresel gürültülerin otomatik olarak ayrıştırılması hedeflenmiştir. Böylece gerçek deprem sinyallerinin tespit süreci hızlandırılmakta ve müdahale süresi kısaltılmaktadır. Sinyaller z-skor normalizasyon yöntemiyle ölçeklendirilmiş ve ardından zaman ve frekans alanlarına ait çeşitli öznitelikler çıkarılmıştır. Zaman alanında ortalama, standart sapma, maksimum, minimum, varyans, çarpıklık ve basıklık gibi öznitelikler kullanılmıştır. Frekans alanında ise tepe frekansı ve ortalama frekans öznitelikleri çıkarılmıştır. Bu öznitelikler sinyallerin daha doğru ve güvenilir şekilde sınıflandırılmasını sağlamaktadır. Çalışmada k-en yakın komşu (k-NN), karar ağaçları (DT) ve toplu torbalanmış karar ağaçları (EBT) algoritmaları kullanılarak sınıflandırma işlemleri gerçekleştirilmiştir. Elde edilen sonuçlara göre k-NN algoritması ile %94,3±0,1, DT algoritması ile %94,8±0,1 ve EBT algoritması ile %95,9±0,1 doğruluk değerleri elde edilmiştir. Bu bulgular deprem sinyallerinin yüksek doğrulukla tespit edilebileceğini göstermektedir. Sunulan çalışma hem akademik araştırmalar hem de pratik uygulamalar açısından önemli bir kaynak olma potansiyeline sahiptir ve depremlerin etkilerinin azaltılmasına yönelik gelecekteki çalışmalara katkı sağlamayı amaçlamaktadır.

Anahtar Kelimeler: Deprem tespiti, karar ağaçları, k-en yakın komşu, öznitelik çıkarma, toplu torbalanmış karar ağaçları.

Abstract

Earthquakes are natural disasters that occur as a result of sudden movements in the Earth's crust. Rapid detection of these events is of vital importance for minimizing loss of life and property. Early and accurate identification enables emergency response teams to arrive on scene in a timely manner and implement measures that enhance public safety. In this study, we aim to automatically distinguish real earthquake signals from environmental noise by using features extracted from seismic recordings, thereby accelerating the detection process and reducing response time. Signals were first scaled using z-score normalization, and then a variety of time- and frequency-domain features were computed. In the time domain, we extracted statistics such as mean, standard deviation, maximum, minimum, variance, skewness, and kurtosis. In the frequency domain, we derived peak frequency and average frequency features. These features improve the accuracy and reliability of signal classification. Classification was performed using k-nearest neighbors (k-NN), decision trees (DT), and ensemble bagged trees (EBT) algorithms. The results show accuracy rates of 94.3±0.1 for k-NN, 94.8±0.1 for DT, and 95.9±0.1 for EBT. These findings demonstrate that earthquake signals can be detected with high accuracy. The presented work has the potential to serve both academic research and practical applications, and aims to contribute to future efforts in mitigating the impacts of earthquakes.

Keywords: Decision trees, earthquake detection, ensemble bagged trees, feature extraction, k-nearest neighbors.

*Sorumlu yazarın e-posta adresi: alinarin@beun.edu.tr

Ali Narin orcid.org/0000-0003-0356-2888

Yunus Emre Erdoğan orcid.org/0000-0003-3677-5564



1. Giriş

Depremler, yeryüzünde geniş çaplı yıkımlara neden olan ve etkileri uzun süre hissedilen doğa olayları arasındadır. Bu olayların ortaya çıkardığı tehlikeleri doğru bir şekilde değerlendirebilmek için, depremlerin oluşturduğu zemin hareketlerinin detaylı bir analizi kritik öneme sahiptir. Sismik dalgaların doğru bir şekilde incelenmesi, bu süreçte başvuru ana yöntemlerden biridir (Avcil vd. 2024). Deprem esnasında yayılan sismik dalgalar, zemin hareketleriyle ilgili önemli bilgiler sağlayarak, yapıların dayanıklılığının yanı sıra altyapının güvenliğinin de değerlendirilmesinde kritik rol oynar. Dolayısıyla, sismik dalgalar sayesinde zemin hareketlerinin karmaşıklığını anlamak ve bu hareketlerin etkilerini analiz etmek mümkün hale gelmektedir (C. Liu vd. 2021).

Türkiye, aktif fay hatları üzerinde bulunan bir ülkedir ve bu nedenle deprem kuşağında yer almaktadır. Bu bağlamda, Düzce'de 12 Kasım 1999'da meydana gelen 7,4 büyüklüğündeki deprem, Türkiye'nin tarihindeki en yıkıcı olaylardan biri olarak kayıtlara geçmiştir (Hearn vd. 2002). Bu büyük felaket sonucunda yaklaşık 18000 bina yıkılmış ve 17000'den fazla can kaybı yaşanmıştır. Bu olay, zemin sıvılaşması ve yapısal zayıflıklar gibi faktörlerin, deprem sonrası ortaya çıkan kayıplar üzerindeki etkisini açıkça göstermiştir. Düzce depreminin ardından, sismik risklerin değerlendirilmesi ve yapı güvenliğinin artırılması amacıyla yeni düzenlemeler yapılmış, inşaat standartları gözden geçirilmiştir. Böylece, Düzce depremi, sismik dalga incelemelerinin ve zemin hareketlerinin analizi konusundaki önemini bir kez daha gözler önüne sermiştir. Benzer bir etki, 6 Şubat 2023'te gerçekleşen 7,7 büyüklüğündeki Kahramanmaraş depreminde de görülmüştür. Bu deprem, Türkiye'nin güneydoğusunda, Kahramanmaraş il merkezine yakın bölgelerde meydana gelmiş ve 20000'den fazla bina yıkılarak yaklaşık 50000 can kaybına neden olmuştur (Yetkin 2024). Kahramanmaraş Depremi, yıkıcı etkileriyle dikkat çekerken birçok insanın evsiz kalmasına ve büyük bir insani krize yol açmıştır. Deprem, çevre illerde de ciddi hasara yol açmış, Gaziantep, Hatay ve Adıyaman gibi şehirlerde de binaların yıkılmasına ve altyapının büyük ölçüde zarar görmesine neden olmuştur. Kahramanmaraş depreminin etkileri sadece fiziksel hasar ile sınırlı kalmamış, aynı zamanda psikolojik etkileri de derin olmuştur. Depremden etkilenen insanlar, yaşadıkları travmalar nedeniyle uzun süreli stres, anksiyete ve depresyon gibi ruhsal sağlık sorunlarıyla karşı karşıya kalmışlardır. Bu durum, afet sonrası rehabilitasyon süreçlerinin önemini vurgulamaktadır. Ayrıca, deprem sonrası yapılan arama-kurtarma çalışmaları, zemin sıvılaşması ve yıkılan

binaların yerleşim alanlarındaki yapı kalitesi gibi faktörlerin zorlukları ile daha da karmaşık hale gelmiştir. Özellikle, bu tür büyük depremler, binaların dayanıklılığı ve yerleşim alanlarının planlanması konusundaki eksikliklerin belirlenmesine yardımcı olmaktadır. Her iki deprem de Türkiye'nin sismik etkinliği ve yapı güvenliğinin önemini vurgulayarak, bu tür olayların önlenmesi için yapılacak bilimsel çalışmaların gerekliliğini ortaya koymaktadır.

Ayrıca, modern teknolojiler ve yapay zekâ uygulamaları, sismik sinyallerin analizine önemli katkılar sağlamaktadır. Yapay zekâ destekli sistemler, büyük miktarda veriyi kısa sürede işleyerek, sinyalin deprem veya gürültü olup olmadığını belirlemede hızlı ve etkili sonuçlar sunmaktadır. Bu tür çalışmalar, yalnızca depremlerin anlaşılmasına değil, aynı zamanda erken uyarı sistemlerinin geliştirilmesine ve depremlerin yaratacağı olası hasarın azaltılmasına da yardımcı olmaktadır. Son yıllarda, bu teknolojilerin depremle ilgili çalışmalarda yoğun olarak kullanılmasıyla birlikte literatürde kayda değer ilerlemeler yaşanmıştır (Essam vd. 2021, Jiao ve Alavi 2020).

Literatürde, yapay zekaya ait farklı yöntemler ve modeller kullanılarak çeşitli başarı oranları elde edilmiştir. Örneğin, Magrini vd. (2020) "Local Earthquakes and Noise DataBase (LEN-DB)" üzerinde 67.939 parametrelili basit bir evrimli sinir ağı (CNN) modeli kullanarak %93,2 gibi yüksek bir doğruluk oranı elde etmişlerdir. Majstorovic vd. (2021) ise, tek istasyonlu üç bileşenli dalga formlarından yararlanarak, uçtan uca iki aşamalı 1 boyutlu CNN modelleri geliştirmiş ve %97 doğruluk oranı ile dikkat çekici sonuçlar ortaya koymuşlardır. Güncel teknolojilerden faydalanan Bhatia vd. (2023) ise nesnelerin interneti kenar (IoT-Edge) bilişim tabanlı bir sistem geliştirmiş, bu sayede %92,52 doğrulukla gerçek zamanlı deprem izleme ve tahminini yapabilmışlardır. Ayrıca, Delastrada vd. (2024) Rastgele Orman (RO) algoritmasıyla %90,78 doğruluk oranı elde ederken, Özkaya vd. (2024) öz nitelik mühendisliği ve ML çoğunluk oyu tekniklerini kullanarak %93,13 doğruluk oranına ulaşmışlardır. Bu farklı çalışmalar, farklı veri setleri, farklı bölgesel odaklar ve farklı sinyal bileşenleri kullanmalarına rağmen genel olarak %90'ın üzerinde doğruluk oranları sunmaktadır.

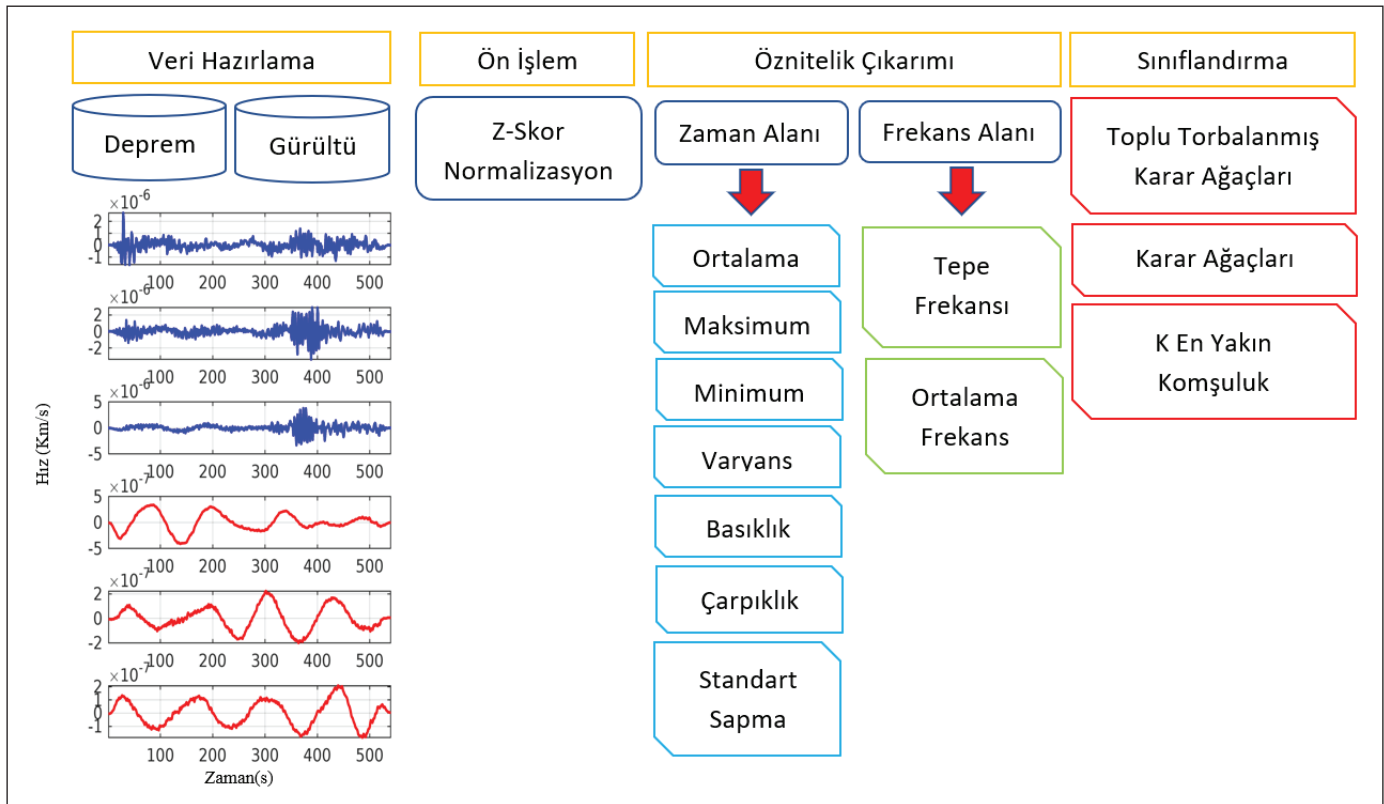
Bu çalışmada, sismik veriler üzerinde sırasıyla ön işleme, öz nitelik çıkarma ve sınıflandırma işlemleri gerçekleştirilmiştir. Çalışmanın genel akış diyagramı Şekil 1'de sunulmuş olup; veri setinin yapısı, uygulanan ön işlem adımları, çıkarılan öz nitelikler, sınıflandırma algoritmalarının değerlendirme kriterleri ve deneysel sonuçlar detaylı olarak açıklanmıştır. Bu çalışma, literatüre birçok açıdan önemli katkılar sağla-

maktadır. Öncelikle hem zaman hem de frekans alanına ait özniteliklerin birlikte kullanılmasıyla sinyallerin genlik ve spektral karakteristiklerinin eş zamanlı olarak değerlendirilmesi sağlanmıştır. Bu bütüncül yaklaşım, yalnızca tek bir alan üzerinden öznitelik çıkarımı yapan önceki çalışmaların birçoğundan daha üstün bir sınıflandırma doğruluğu ortaya koymuştur. Ayrıca, LEN-DB veri seti gibi geniş ölçekli ve gerçekçi bir veri üzerinde yapılan deneyler, yöntemin farklı coğrafi lokasyonlardan gelen yüksek hacimli verilerde etkinliğini kanıtlamış; böylece çalışmanın gerçek dünya uygulamalarına uygunluğu sağlanmıştır. Sınıflandırma algoritmaları arasında yapılan karşılaştırmada, toplu torbalanmış karar ağaçları (EBT) yöntemi hem yüksek doğruluk hem de düşük varyans ile daha kararlı sonuçlar vermiştir. Bu durum, karar destek sistemlerinde EBT yönteminin tercih edilmesini sağlamaktadır. Öte yandan, kullanılan yöntemlerin düşük işlem maliyeti ve hızlı öznitelik çıkarımı, gerçek zamanlı deprem izleme ve erken uyarı sistemlerine entegrasyon potansiyelini güçlendirmiştir. Son olarak, önerilen modelin sadece belirli bir bölgeye değil, farklı sismik ortamlar ve istasyon koşullarına da uyarlanabilir esnek bir yapıya sahip olması, coğrafi bağımsızlık ve genişletilebilirlik açısından önemli bir avantaj olarak öne çıkmaktadır.

2. Gereç ve Yöntemler

2.1. Veri Seti

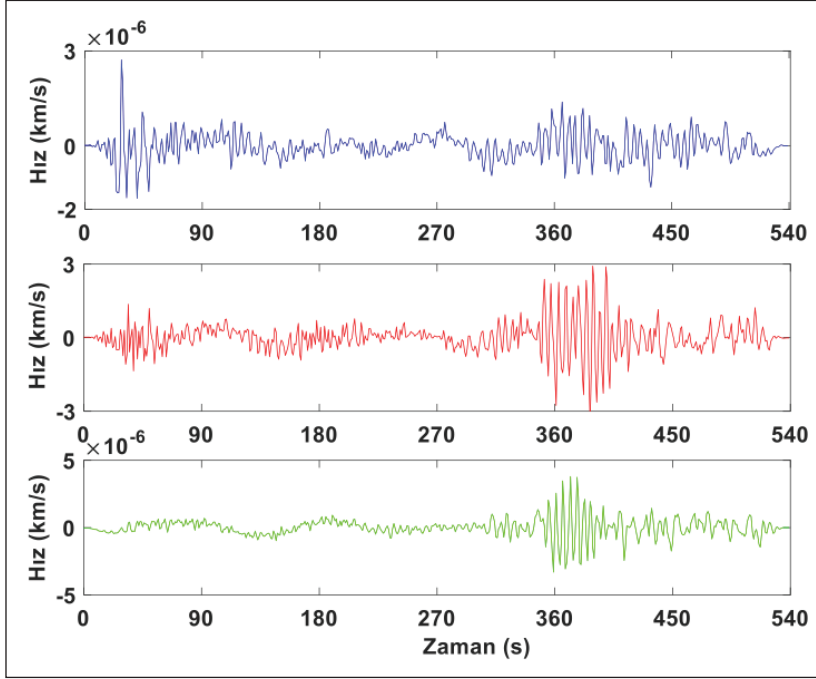
Bu çalışmada “<https://zenodo.org>” adresinde herkese açık olarak sunulan LEN-DB veri kümesi kullanılmıştır (Magrini vd. 2020). Veri kümesi, dünya genelinde konumlandırılmış 1487 geniş bantlı veya çok geniş bantlı alıcıya ait doğu (X), kuzey (Y) ve dikey (Z) bileşenleri boyunca kaydedilen sismogramlardan oluşmaktadır. Her bir deprem olayı için, P-dalgasının beklenen varış zamanından 4 saniye önce başlayan ve toplam 27 s süren bir zaman penceresi oluşturulmuştur. Bu pencere süresi boyunca kaydedilen 3 bileşenli sismogramlar, 20 Hz örnekleme frekansı ile elde edilmiştir. P-dalgasının varış zamanı, IASP91 1-boyutlu hız modeli (Kennett ve Engdahl 1991) ile hesaplanmıştır. Bu sayede, kayıtlar henüz herhangi bir dalga ulaşmadan önce başlamış ve P, S ve yüzey dalgalarını içerecek şekilde tüm sismik aktiviteyi kapsamaktadır. Dalga türlerine özel bir ayrıştırma yapılmamış; sinyaller bütünsel olarak değerlendirilmiştir. Veri seti indirildikten sonra ortalama değerden arındırılmış, trend bileşeni giderilmiş, %5’lik kosinüs taper uygulanmış ve 0,1–5 Hz aralığında bant geçiren bir filtre ile işlenmiştir. Ardından enstrümantal yanıtlar kaldırılarak tüm sinyaller



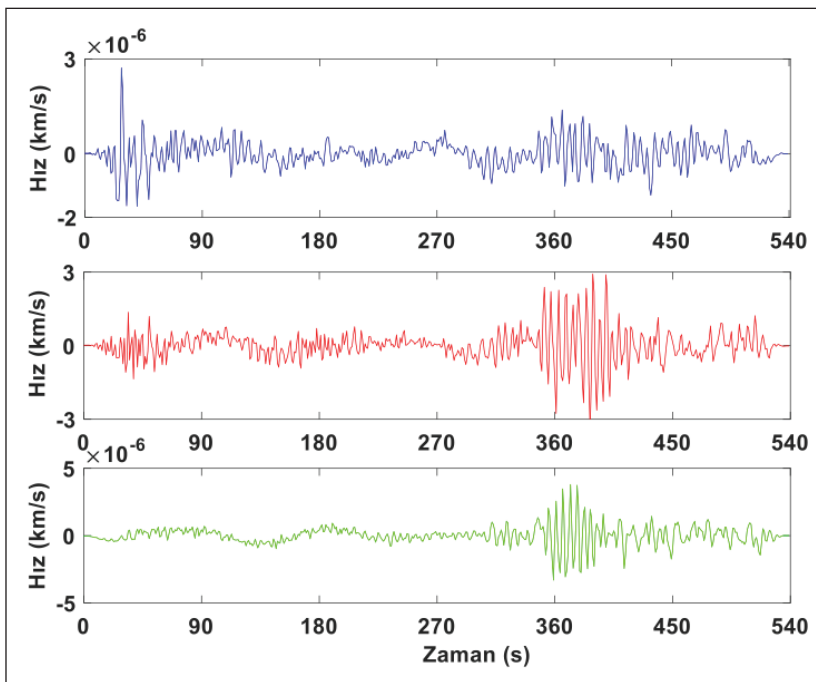
Şekil 1. Deprem sinyallerinin tespiti için önerilen sistemin akış diyagramı.

hız birimine dönüştürülmüştür. Bu ön işleme adımları sonucunda 304.878 yerel depreme ait 1.244.942 adet sismogram elde edilmiştir. Bu sismogramların 629.095'i EQ (deprem), 615.847'si ise AN (gürültü) etiketiyle sınıflandırılmıştır. Her bir sismogram 540 örnek içerecek şekilde yapılandırılmıştır. Veri setinde belirgin bir sınıf dengesizliği bulunmamaktadır.

Etiketlenen bu verilere ait sinyaller Şekil 2 ve Şekil 3'te gösterilmiştir. Bu çalışmada deprem sinyallerinden rastgele 50001 adet ve gürültü sinyallerinden rastgele 50001 adet kullanılmıştır. Her bir veri, 540 örnekten oluşmaktadır. Çizelge 1'de çalışmada kullanılan verilerin sayıları, uzunlukları ve sınıfları detaylı bir şekilde verilmiştir.



Şekil 2. X(Mavi), Y(Kırmızı), Z(Yeşil) yönlerine ait 3 bileşenli deprem sinyali grafiği.



Şekil 3. X(Mavi), Y(Kırmızı), Z(Yeşil) yönlerine ait 3 bileşenli gürültü sinyali grafiği.

Çizelge 1. Çalışmada kullanılan veriler ve özellikleri.

Sınıf	Veri Sayısı	Toplam	Veri Uzunluğu
Deprem	50001	100002	540
Gürültü	50001		540

2.2. Yöntem

2.2.1. Ön İşlem

2.2.1.1. Z-skor normalizasyon

Z-skor normalizasyon, bir veri kümesindeki her gözlemin ortalama ve standart sapmaya göre yeniden ölçeklendirilmesi işlemidir. Bu sayede verilerin ortalaması 0, standart sapması ise 1 olur. Bu yöntem, farklı ölçeklerdeki verileri karşılaştırılabilir hale getirir ve makine öğrenimi algoritmalarının performansını artırır. Matematiksel ifadesi şu şekildedir (Sholeh ve Nurnawati 2024):

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

Burada x, veri setindeki herhangi bir noktayı; μ , veri setinin ortalamasını ve σ , veri setinin standart sapmasını ifade etmektedir.

2.2.2. Öznitelik Çıkarımı

Öznitelik çıkarımı, ham verilerden anlamlı ve ayırt edici özelliklerin elde edilmesi sürecidir; böylece verilerin analiz ve sınıflandırılması kolaylaşır. Bu işlem, verilerin boyutunu azaltarak modelin daha verimli ve doğru sonuçlar üretmesini sağlar. Öznitelik çıkarımı, zaman alanı analizleri, sinyal işleme teknikleri, görüntü işleme ve makine öğrenimi yöntemleri gibi çeşitli yaklaşımlar kullanılarak gerçekleştirilebilir (Singh ve Krishnan 2023). Bu yöntemler sayesinde ham veriler anlamlı hâle getirilerek makine öğrenimi modellerinin performansı artırılır ve eğitim süreçleri hızlandırılır. Deprem tespiti için yapılan bu çalışmada hem zaman alanı hem de frekans alanı özniteliklerine yer verilmiştir. Zaman alanı öznitelikleri arasında sinyalin genel seviyesini temsil eden ortalama, verinin yayılımını ölçen standart sapma, dağılımın ve değişkenliğin derecesini gösteren varyans ile sinyalin uç değerlerini belirleyen maksimum ve minimum değerler bulunmaktadır. Ayrıca çarpıklık, dağılımın simetrik olup olmadığını; basıklık ise dağılımın sivrilmesini analiz etmektedir. Frekans alanı analizinde ise sinyalin en yoğun olduğu frekansı gösteren tepe frekansı ve frekans dağılımının merkezini ifade eden ortalama frekans kullanılmıştır. Bu öznitelikler, deprem sinyalleri ile gürültü arasındaki ayırt edici özellikleri ortaya çıkararak modelin sınıflandırma başarısını önemli ölçüde artırmıştır (Bai vd. 2023). Çizelge 2'de makalede kullanılan tüm öznitelikler detaylı olarak listelenmiştir.

Çizelge 2. Öznitelik çıkarımı

Öznitelikler	Açıklama
Standart Sapma	Deprem sinyallerinin ve gürültü sinyallerinin standart sapması
Ortalama	Deprem sinyallerinin ve gürültü sinyallerinin ortalaması
Maksimum	Deprem sinyallerinin ve gürültü sinyallerinin maksimumu
Minimum	Deprem sinyallerinin ve gürültü sinyallerinin minimumu
Varyans	Deprem sinyallerinin ve gürültü sinyallerinin varyansı
Çarpıklık	Deprem sinyallerinin ve gürültü sinyallerinin çarpıklık değeri
Basıklık	Deprem sinyallerinin ve gürültü sinyallerinin basıklık değeri
Tepe Frekansı	Deprem sinyallerinin ve gürültü sinyallerinin tepe frekansı
Ortalama Frekans	Deprem sinyallerinin ve gürültü sinyallerinin ortalama frekansı

2.2.2.1. Standart sapma

Standart sapma, bir veri kümesindeki değerlerin ortalama-dan ne kadar uzaklaştığını ölçen istatistiksel bir ölçüdür (Cai vd. 2021). Zaman alanında, bir sinyalin dalgalanma miktarını ve gürültü seviyesini belirlemede önemli bir rol oynar. Düşük bir standart sapma, sinyalin zaman içinde stabil olduğunu, yüksek bir standart sapma ise daha büyük varyasyonlar ve dalgalanmalar olduğunu gösterir. Bu öznitelik, sinyal işleme ve analizinde, özellikle sinyalin gürültü içerip içermediğini değerlendirmek için kullanılır. Standart sapma, aşağıdaki formülle hesaplanır:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2} \quad (2)$$

Burada, σ standart sapmayı, x veri noktalarını, μ veri kümesinin ortalamasını ve n veri kümesindeki eleman sayısını ifade eder.

2.2.2.2. Ortalama

Bir veri setindeki değerlerin aritmetik ortalamasıdır. Tüm değerlerin toplamının, veri sayısına bölünmesiyle elde edilir.

$$\text{Ortalama } (\mu) = \frac{\sum_{i=1}^n x_i}{n} \quad (3)$$

Burada x_i veri setindeki her bir değeri, n ise toplam veri sayısını temsil eder.

2.2.2.3. Maksimum

Bir veri setinde bulunan en büyük değerdir.

$$\text{Maksimum} = \max(x_1, x_2, \dots, x_n) \quad (4)$$

Burada x_i veri setindeki her bir değeri, n ise toplam veri sayısını temsil eder.

2.2.2.4. Minimum

Bir veri setinde bulunan en küçük değerdir.

$$\text{Minimum} = \min(x_1, x_2, \dots, x_n) \quad (5)$$

Burada x_i veri setindeki her bir değeri, n ise toplam veri sayısını temsil eder.

2.2.2.5. Varyans

Bir veri setindeki değerlerin, ortalamaya göre ne kadar yayıldığını ölçen bir değerdir. Yüksek varyans, verilerin geniş bir aralığa yayıldığını gösterirken, düşük varyans verilerin ortalama etrafında toplandığını gösterir.

$$\text{Varyans } (\sigma)^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n - 1} \quad (6)$$

Burada x_i veri setindeki her bir değeri, μ ortalamayı ve n ise toplam veri sayısını temsil eder.

2.2.2.6. Çarpıklık

Verilerin simetri durumunu ölçen bir değerdir. Çarpıklık, dağılımın sağa veya sola kaymış olup olmadığını gösterir. Pozitif çarpıklık sağa çarpık, negatif çarpıklık ise sola çarpık olduğunu gösterir.

$$\text{Çarpıklık} = \frac{\sum_{i=1}^n (x_i - \mu)^3}{(n - 1) \cdot \sigma^3} \quad (7)$$

Burada x_i veri setindeki her bir değeri, μ ortalamayı, σ standart sapmayı ve n ise toplam veri sayısını temsil eder.

2.2.2.7. Basıklık

Dağılımın tepe noktasının yüksekliğini ve uçlardaki veri yoğunluğunu ölçer. Yüksek basıklık, dağılımın sivri olduğunu; düşük basıklık ise daha düz olduğunu gösterir.

$$\text{Basıklık} = \frac{\sum_{i=1}^n (x_i - \mu)^4}{(n - 1) \cdot \sigma^4} - 3 \quad (8)$$

Burada x_i veri setindeki her bir değeri, μ ortalamayı, σ standart sapmayı ve n ise toplam veri sayısını temsil eder.

Burada 3 çıkarılarak normal dağılımla kıyaslama yapılır. Pozitif değerler sivri bir dağılımı, negatif değerler ise düz bir dağılımı ifade eder. Bu metrikler, veri analizinde verilerin genel yapısını, dağılımını ve dağılımın şekli hakkında bilgi edinmek için kullanılır.

2.2.2.8. Tepe frekansı

Tepe frekansı, bir sinyalin frekans spektrumunda en yüksek genliğe sahip frekans olarak tanımlanır (Haraguchi vd. 2023). Bu öznelik, sinyalin enerji dağılımını analiz etmek ve sinyalin ana karakteristik frekansını belirlemek için kritik bir öneme sahiptir. Tepe frekansı, bir sinyalin temel bileşeni ve frekans özelliklerini anlamaya yardımcı olur. Özellikle iletişim sistemleri, müzik analizi ve sinyal sınıflandırma uygulamalarında kullanılır. Tepe frekansı, aşağıdaki formülle hesaplanır:

$$f_{\text{peak}} = \text{argmax}(|X(f)|) \quad (9)$$

Burada f_{peak} tepe frekansını, $X(f)$ sinyalin Fourier dönüşümünden elde edilen frekans spektrumunu, $|X(f)|$ frekans spektrumundaki genliği ifade eder.

2.2.2.9. Ortalama frekans

Ortalama frekans, bir sinyalin frekans spektrumunda ağırlıklı ortalama frekansını ifade eder. Bu metrik, sinyalin enerji dağılımının hangi frekans aralığında yoğunlaştığını gösterir. Özellikle sinyallerin spektral analizinde ve biyomedikal sinyallerin analizinde kullanılır. Ortalama frekans aşağıdaki formülle hesaplanır:

$$\text{Ortalama Frekans} = \frac{\sum_{i=1}^n f_i \cdot P(f_i)}{\sum_{i=1}^n P(f_i)} \quad (10)$$

Burada f_i i-inci frekans bileşenini, $P(f_i)$ i-inci frekansta bulunan güç spektral yoğunluk değerini ve n frekans bileşenlerinin toplam sayısını ifade eder. Hesaplama sırasında, sinyalin Fourier dönüşümü alınır ve güç spektral yoğunluğu elde edilir. Her bir frekans bileşeninin ağırlıklı ortalaması, o frekansta bulunan güçle çarpılarak hesaplanır. Sonuç olarak, sinyalde en çok enerjinin yoğunlaştığı frekans değeri elde edilir. Ortalama frekans, özellikle sinyallerin enerji içeriklerini ve frekans aralığındaki dağılımını anlamak için kullanılan önemli bir spektral ölçüttür.

2.2.3. Sınıflandırma Algoritmaları

2.2.3.1. K en yakın komşu (k-NN)

k-NN algoritması hem istatistiksel analizlerde hem de veri madenciliğinde sıkça tercih edilen sezgisel ve etkili bir sınıflandırma yöntemidir. Bu yaklaşım, etiketlenmemiş bir örneğin hangi kategoriye ait olduğunu belirlemek için örneğin öznitelik uzayında konumlandırır ve mevcut etiketli verilerle karşılaştırır. Karşılaştırma genellikle veri noktaları arasındaki mesafelerin hesaplanmasıyla yapılır; bu amaçla en yaygın olarak Öklidyen mesafe ölçütü kullanılır. Algoritma, sınıflandırılacak örneğin eğitim kümesindeki en yakın k

komşunun sınıf etiketlerine göre değerlendirir. Örneğin, $k = 1$ seçildiğinde örnek yalnızca kendisine en yakın veri noktasının sınıfına; $k = 5$ olduğunda ise beş komşu arasındaki çoğunluk sınıfına atanır (Erdoğan ve Narin 2022).

2.2.3.2. Karar ağaçları (DT)

DT veriyi dallanarak sınıflandıran hiyerarşik yapılardır. Ağaç, üzerinde henüz karar uygulanmamış kök düğümden başlar. Her ara düğüm; veri kümesindeki belirli bir özelliğe dayalı olarak bir karar kuralı uygular ve örnekleri alt dallara ayırır. Bu işlem, veri alt grupları saf hâle gelene veya başka bir kriter karşılanana kadar devam eder. Bölünme sona erdiğinde ulaşılan düğümlere yaprak düğümler denir; her yaparak, genellikle o sınıfa ait olasılık ya da sınıf etiketini içerir. Sınıflandırma sürecinde örnek, kök düğümden başlayarak her karar noktasında ilgili özelliğe göre uygun dala yönlendirilir ve nihayetinde bir yaprağa ulaşarak sınıfı belirlenir (Yapıcı vd. 2024).

2.2.3.3. Toplu torbalanmış karar ağaçları (EBT)

DT sınıflandırma ve regresyon problemlerinde yaygın olarak kullanılan, veri setlerindeki örüntüleri keşfetmek amacıyla hiyerarşik bir yapı sunan denetimli bir öğrenme tekniğidir. DT modelleri, özniteliklerin belirli bir dizi koşula göre dallanarak sınıflandırma gerçekleştirmesine dayanır (Bansal vd. 2022). DT her bir dalda veri noktalarını belirli özniteliklere göre ayırarak kararlar alır. Bu süreçte, hangi özelliğin dallanma için kullanılacağı önem taşır ve genellikle en iyi ayrımı sağlayan öznitelik tercih edilir. Ayrımın kalitesini değerlendirmek için entropi kullanılır. Entropi, veri kümesindeki belirsizliği ya da düzensizliği temsil eder. DT’de entropi değeri düşük olan dallanma noktaları daha homojen gruplar oluşturur, bu sayede sınıflandırma daha net yapılabilir. Bu nedenle, entropi hesaplaması DT’nin hangi özelliğe göre bölüneceğinin belirlenmesinde kilit bir rol oynar. Entropi şu formül ile hesaplanır:

$$\text{Entropi}(S) = - \sum_{i=1}^C p_i * \log_2(p_i) \quad (11)$$

Burada p_i ifadesi i sınıfının olasılığını C ise sınıf sayısını temsil eder. Entropi(S), veri kümesindeki belirsizliği veya düzensizliği ölçen bir metriktir. S ise sınıflandırılacak tüm örneklerin bulunduğu ana veri setini ifade eder. Entropi, düğümlerde bilgi kazancını maksimize ederek DT’nin dallanmasına yardımcı olur. Bir diğer önemli ölçüt ise bilgi kazancıdır. Bilgi kazancı, bir özelliğin veriyi ne kadar iyi böldüğünü gösterir ve aşağıdaki formülle ifade edilir:

$$\text{Bilgi Kazancı}(A) = \text{Entropi}(S) - \sum_{j=1}^v \left| \frac{S_j}{S} \right| * \text{Entropi}(S_j) \quad (12)$$

Burada v ifadesi öznitelik A ’ya göre oluşan alt düğüm sayısını, Entropi (S_j) j ’nci alt düğümün entropisini, S_j j ’nci alt düğümdeki örnek sayısını ve S de toplam örnek sayısını belirtmektedir.

Topluluk öğrenme, birden fazla öğrenme algoritmasını bir araya getirerek sınıflandırma ve tahmin performansını iyileştiren gelişmiş bir makine öğrenimi tekniğidir. Bu yöntem DT, destek vektör makineleri ve yapay sinir ağları gibi çeşitli öğrenme modellerinin birlikte kullanılmasını içermektedir. Bu kapsamda, EBT yöntemi, torbalama tekniğini uygulayarak çok sayıda DT’yi bir araya getirir (Ali vd. 2022). Her bir DT, eğitim veri kümesinden rastgele seçilen alt kümeler üzerinde bağımsız olarak eğitilir. Bu süreçte, eğitim veri kümesi farklı örneklerle yeniden örneklendirilerek her bir ağaca farklı bir alt küme verilir. Bağımsız olarak eğitilen bu DT’nin tahminleri daha sonra çoğunluk oylaması yöntemiyle birleştirilir. Böylelikle, tek bir modelin sınırlamalarından kaynaklanan varyans azalır ve genel model performansı artırılır. RO algoritması, EBT yönteminin iyi bilinen ve yaygın olarak kullanılan bir uygulamasıdır. Bu algoritma, her bir DT’yi eğitmek için veri kümesinden rastgele örneklem seçmenin yanı sıra, özellikler arasından da rastgele bir alt küme seçerek modelin çeşitliliğini daha da artırır. RO algoritması bu şekilde hem aşırı uyum riskini azaltır hem de modelin genelleme yeteneğini geliştirir. Farklı öznitelik kombinasyonları ve örnekler üzerinde eğitim alan ağaçlar, daha dayanıklı ve doğru tahminler yapma kapasitesine sahip olur. Bu yöntem, çoğu veri setinde yüksek performans gösterir ve sınıflandırma problemlerinde yaygın olarak tercih edilmektedir.

2.2.4. Performans Ölçütleri

Performans ölçütü, kullanılan sınıflandırma sisteminin verileri ne derece doğru sınıflandırabildiğini ifade eden temel bir değerlendirme terimidir. Bu ölçütlerde, en önemli faktörlerden biri, kullanılan verilerin eğitim ve test kümelerine nasıl ayrıldığıdır. Zira, eğitim ve test kümelerinin birbirlerinden ne ölçüde ayrıldığı, modelin performans sonuçlarını doğrudan etkileyebilmektedir. Literatürde, veri kümelerinin ayrılmasında sıklıkla “ayrık-tutma”, “k-katlı” ve “birini-dışarda-bırak” çapraz doğrulama teknikleri tercih edilmektedir (Kaliappan vd. 2023, Kelter, 2021, Giarno vd. 2020). Bu yöntemler arasında, k-katlı çapraz doğrulama en yaygın olarak kullanılan yaklaşımdır. K-katlı yönteminde, veriler k adet alt kümeye ayrılır. Bu k parçadan biri test işlemi için ayrılırken, kalan $k-1$ parça eğitim işlemi için kullanılır. Bu süreç k defa tekrar edilerek, Doğru Pozitif (TP), Doğru Negatif (TN), Yanlış Pozitif (FP) ve Yanlış Negatif (FN)

değerleri hesaplanır (Çizelge 3). Bu çalışmada ise, Şekil 4'te gösterildiği üzere 5-katlı çapraz doğrulama yöntemi kullanılarak modelin doğruluğu test edilmiştir.

Çizelge 3. Karışıklık matrisi.

		Tahmin Edilen Sınıf	
		Negatif	Pozitif
Gerçek Sınıf	Negatif	TN Doğru Negatif	FP Yanlış Pozitif
	Pozitif	FN Yanlış Negatif	TP Doğru Pozitif

Çizelge 3'te gösterildiği üzere, aslında deprem sinyali olan ve algoritmanın doğru şekilde deprem sinyali olarak tanımladığı durumlar TP olarak ifade edilmektedir. Buna karşılık, gerçek deprem sinyalleri arasından algoritmanın yanlışlıkla gürültü sinyali olarak belirlediği örnekler FN olarak adlandırılır. Diğer yandan, gerçek gürültü sinyalleri arasında algoritmanın doğru bir şekilde gürültü sinyali olarak sınıflandırdığı durumlar TN olarak tanımlanırken, aslında gürültü sinyali olan fakat algoritma tarafından hatalı olarak deprem sinyali olarak değerlendirilen örnekler ise FP olarak adlandırılır. Bu çalışmada, doğruluk (accuracy, Acc), duyarlılık (recall, Rec), özgüllük (specificity, Spe), kesinlik (precision,

Pre) ve F1-skoru (F1-score, F1) gibi performans metrikleri hesaplanmaktadır (Shaheed vd. 2022, Kaur vd. 2022, Zhang vd. 2021).

$$Acc = \frac{TP + TN}{TP + FN + FP + TN} \quad (13)$$

$$Rec = \frac{TP}{TP + FN} \quad (14)$$

$$Spe = \frac{TN}{TN + FP} \quad (15)$$

$$Pre = \frac{TP}{TP + FP} \quad (16)$$

$$F_1 = \frac{2 * Pre * Rec}{Pre + Rec} \quad (17)$$

Bu performans ölçütleri, sınıflandırma modellerinin başarısını detaylı şekilde değerlendirmeye olanak tanır. Acc, modelin genel doğruluğunu ortaya koyarken; Rec, modelin doğru pozitifleri tespit etme kapasitesini gösterir. Spe ise negatif sınıfları doğru bir şekilde ayırt etme oranını ifade eder. Pre, pozitif sınıflandırmaların doğruluk derecesini ölçerken, F1 ise Rec ve Pre arasındaki dengeyi sağlamada önemli bir rol oynar (Tasci vd. 2022, Erdoğan ve Narin 2021). Bu metriklerin birlikte kullanılması, modelin performansına dair kapsamlı bir değerlendirme sunar.



Şekil 4. 5-katlı çapraz doğrulama yönteminin grafiksel gösterimi.

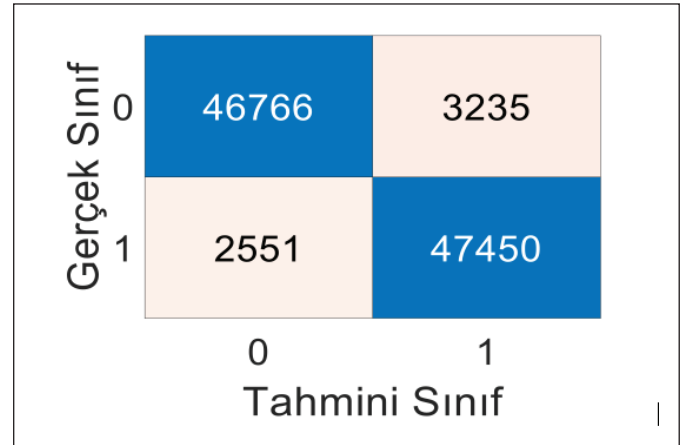
3. Bulgular ve Tartışma

Bu çalışmada gerçekleştirilen tüm deneyler ve analizler, MATLAB 2020 yazılımı kullanılarak titizlikle gerçekleştirilmiştir. Başlangıçta, her iki sınıftan toplam 600 binin üzerinde 3 bileşenli sinyal mevcutken, veri işleme süresini optimize etmek ve hesaplama yükünü minimize etmek amacıyla, 100.002 sinyal içeren daha yönetilebilir bir veri setine geçiş yapılmıştır. Bu stratejik adım, veri seti boyutunun azaltılmasıyla birlikte, analiz sürecinin verimliliğini ve hızını artırmayı hedeflemektedir. Elde edilen sinyallerin 50.001'i gürültü ve diğer 50.001'i deprem sinyallerinden oluşmakta olup, bu eşit dağılım, 3 bileşenli sinyallerin dengeli bir şekilde temsil edilmesini sağlamaktadır. Veri setinin bu dengeli yapısı, aynı zamanda sınıflandırma algoritmalarının genel performansını optimize etmekte ve modelin daha yüksek doğruluk oranlarına ulaşmasını desteklemektedir. Verilerin daha iyi sınıflandırılabilmesi için öncelikle z-skor normalizasyon uygulanmış ve ardından öznitelik çıkarımı gerçekleştirilmiştir. Ayrıca, öznitelik çıkarımı aşamasında, sinyallerden elde edilen temel bileşenler ve önemli özellikler seçilmiş, böylece modelin performansını artırmak amacıyla gereksiz veya alakasız verilerin etkisi azaltılmıştır. Kullanılan öznitelikler arasında standart sapma, ortalama, maksimum, minimum, varyans, çarpıklık, basıklık, tepe frekansı ve ortalama frekans bulunmaktadır. Literatürde, benzer çalışmalarda öznitelik çıkarımının ve normalizasyonun model performansı üzerindeki olumlu etkileri vurgulanmıştır (Tam vd. 2016, Gonzalez vd. 2012, Liu vd. 2021). Bu çalışmalar, verilerin ön işleme aşamalarının, sınıflandırma algoritmalarının başarısını doğrudan etkilediğini göstermektedir.

Bu çalışmada sınıflandırma aşamasında üç farklı makine öğrenimi algoritması olan k-NN, DT ve EBT kullanılmıştır. Modeller, 5-katlı çapraz doğrulama yöntemi ile değerlendirilmiştir. Bu kapsamda, veri kümesi her yinelemede eğitim ve test olarak ayrılmış; eğitim veri seti üzerinde model eğitimi gerçekleştirilmiş ve performans ölçümleri test veri seti üzerinden elde edilmiştir. Her bir algoritma için 10 kez tekrarlanan 5-katlı çapraz doğrulama işlemi sonucunda elde edilen test verisine ait ortalama performans değerleri değerlendirmeye alınmıştır. Karşılaştırmalı analiz sonucun-

da, EBT algoritması en yüksek sınıflandırma performansını göstermiştir. EBT algoritması ile, sismik sinyallerin sınıflandırılmasında %95,9±0,1 Acc, %95,4±0,2 Rec, %96,4±0,2 Spe, %96,4±0,2 Pre ve %95,9±0,1 F1 değerleri elde edilmiştir. DT algoritması ile %94,8±0,1 Acc, %94,2±0,3 Rec, %95,4±0,4 Spe, %95,3±0,4 Pre ve %94,8±0,1 F1 performans değerleri elde edilirken; k-NN algoritması ile %94,3±0,1 Acc, %95,7±0,1 Rec, %92,8±0,2 Spe, %93±0,2 Pre ve %94,4±0,1 F1 performans değerlerine ulaşılmıştır. Sınıflandırma algoritmalarının test verisine ait ortalama performans metrikleri Çizelge 4'te özetlenmiştir.

k-NN modeli ile yapılan sınıflandırmada 50.001 deprem sinyalinin 47450'si doğru şekilde deprem, 2551'i ise yanlışlıkla gürültü olarak sınıflandırılmıştır. Benzer şekilde, 50.001 gürültü sinyalinden 46766'sı doğru şekilde gürültü, 3235'i ise yanlışlıkla deprem olarak tespit edilmiştir. Bu detaylı sonuçlar Şekil 5'te sunulan karışıklık matrisinde gösterilmiştir.



Şekil 5. k-NN algoritması için en yüksek başarıma ait karışıklık matrisi.

DT modeli ile yapılan sınıflandırmada 50.001 deprem sinyalinin 46998'i doğru şekilde deprem, 3003'ü ise yanlışlıkla gürültü olarak sınıflandırılmıştır. Benzer şekilde, 50.001 gürültü sinyalinden 47721'i doğru şekilde gürültü, 2280'i ise yanlışlıkla deprem olarak tespit edilmiştir. Bu detaylı sonuçlar Şekil 6'da sunulan karışıklık matrisinde gösterilmiştir.

Çizelge 4. Sınıflandırma algoritmalarına ait performans sonuçları.

Algoritma	Acc (%)	Rec (%)	Spe (%)	Pre (%)	F1 (%)
EBT	95,9±0,1	95,4±0,2	96,4±0,2	96,4±0,2	95,9±0,1
DT	94,8±0,1	94,2±0,3	95,4±0,4	95,3±0,4	94,8±0,1
k-NN	94,3±0,1	95,7±0,1	92,8±0,2	93±0,2	94,4±0,1

Gerçek Sınıf	Tahmini Sınıf	
	0	1
0	47721	2280
1	3003	46998

Şekil 6. DT algoritması için en yüksek başarıma ait karışıklık matrisi.

Son olarak EBT modeli ile yapılan sınıflandırmada 50.001 deprem sinyalinin 47.526'sı doğru şekilde deprem, 2.475'i ise yanlışlıkla gürültü olarak sınıflandırılmıştır. Benzer şekilde, 50.001 gürültü sinyalinden 48.100'ü doğru şekilde gürültü, 1.901'i ise yanlışlıkla deprem olarak tespit edilmiştir. Bu detaylı sonuçlar Şekil 7'de sunulan karışıklık matrisinde gösterilmiştir. Kapsamlı analizler, EBT algoritmasının deprem ve gürültü sinyallerinin ayrıştırılmasında en yüksek doğruluğu sunduğunu ortaya koymuştur.

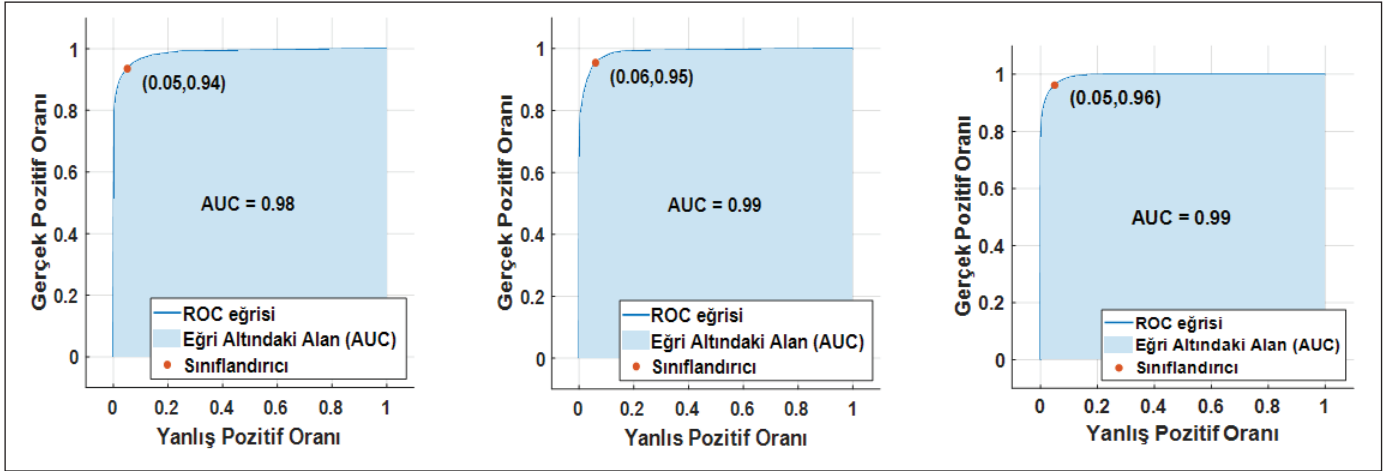
Gerçek Sınıf	Tahmini Sınıf	
	0	1
0	48100	1901
1	2475	47526

Şekil 7. EBT algoritması için en yüksek başarıma ait karışıklık matrisi.

Sınıflandırma performansını daha kapsamlı değerlendirmek amacıyla, üç farklı algoritmanın Alıcı İşletim Karakteristiği (ROC) eğrileri çizilmiş ve Eğri Altındaki Alan (AUC) değerleri hesaplanarak Şekil 8'de gösterilmiştir. Sonuçlar, EBT algoritmasının 0,99 AUC değeri ile en yüksek ayırt edici güce sahip olduğunu göstermektedir. Bu değer, EBT modelinin

sismik sinyallerin deprem ve gürültü olarak doğru şekilde sınıflandırılmasında oldukça başarılı olduğunu doğrulamaktadır. DT algoritması 0,99 AUC ile yüksek performansını korurken, k-NN algoritmasının AUC değeri 0,98 olarak elde edilmiştir. Üç model arasındaki AUC farkları küçük olmakla birlikte, EBT algoritmasının hem Acc hem de ROC eğrisi alanı açısından üstünlüğü, bu yöntemin sınıflandırma problemlerinde özellikle karmaşık verilerle çalışırken daha etkili olduğunu göstermektedir. Bu bulgular, Çizelge 4'te sunulan Acc, Rec, Spe, Pre ve F1 gibi diğer performans metrikleriyle de uyumludur ve EBT algoritmasının deprem tespiti için tercih edilmesinin uygunluğunu desteklemektedir.

Literatürde sismik sinyallerin sınıflandırılması ve deprem olaylarının tespiti için çeşitli ML ve derin öğrenme yaklaşımları geliştirilmiştir. Bu çalışmalar, farklı veri kümeleri ve modelleme stratejileriyle önemli başarılar elde etmiş ve saha uygulamalarına yön verebilecek nitelikte katkılar sunmuştur. Magrini vd. (2020), LEN-DB veri setini kullanarak 67.939 parametrelili, derin katmanlı olmayan ancak optimize edilmiş bir CNN modeli geliştirmiştir. Bu model ile sismik sinyallerin deprem ve gürültü olarak ayrıştırılmasında %93,2 doğruluk oranına ulaşılmıştır. Yöntemleri, düşük model karmaşıklığına rağmen yüksek başarı sunması açısından dikkat çekicidir. Majstorović vd. (2021), uçtan uca bir yaklaşım benimseyerek bir boyutlu CNN tabanlı iki aşamalı bir model geliştirmiştir. Modelin ilk aşaması, veri ön işleme ve öznitelik çıkarımı üzerine odaklanırken, ikinci aşama sınıflandırma görevini yerine getirmiştir. Bu yapı, özellikle gerçek zamanlı uygulamalarda işlem verimliliğini artırmış ve %97 doğruluk oranıyla literatürdeki en yüksek başarılarından birini elde etmiştir. Bhatia vd. (2023), IoT-Edge bilişim altyapısını kullanarak düşük gecikmeli, gerçek zamanlı bir deprem izleme ve tahmin sistemi geliştirmiştir. Modelde Naive Bayes ve Uyarlanabilir Ağ Tabanlı Bulanık Çıkarım Sistemi (ANFIS) yöntemleri birlikte kullanılmıştır. Bu sayede hem açıklanabilirlik sağlanmış hem de sistemin uç cihazlarda çalışabilirliği artırılmıştır. Çalışma, %92,52 doğruluk oranı elde ederek edge bilişimin deprem tahmininde uygulanabilirliğini ortaya koymuştur. Delastrada vd. (2024), klasik makine öğrenimi algoritmalarından biri olan RO ile deprem ve gürültü sinyallerini sınıflandırmıştır. DT topluluğuna dayalı bu model, özellikle özniteliklerin önem derecelerini belirlemede güçlüdür. %90,78 doğruluk oranı elde edilen bu çalışma, özellikle eğitim süresinin kısa olması ve modelin kolay yorumlanabilirliği açısından önem arz etmektedir. Özkaya vd. (2024) ise "Butterfly Pattern (BFPat)" adlı bir öznitelik mühendisliği yaklaşımını geliştirerek, çeşitli ML algoritmalarının çoğunluk oyu ile birleştirildiği bir toplu sınıflandır-



Şekil 8. Sınıflandırma yöntemlerine ait ROC eğrileri: A) k-NN B) DT C) EBT.

Çizelge 5. Literatüre ait performans sonuçları.

Yazarlar	Yıl	Yöntem / Model	Veri Tabanı	Acc (%)
Magrini vd.	2020	67.939 parametrelili basit CNN	LEN-DB	93,2
Majstorović vd.	2021	1D CNN ile uçtan uca iki aşamalı model	AQULO	97,0
Bhatia vd.	2023	IoT-Edge + Naive Bayes + ANFIS	STEAD	92,5
Delastrada vd.	2024	RO	BMKG	90,8
Ozkaya vd.	2024	Öznitelik mühendisliği (BFPat) + ML (Çoğunluk Oyu)	STEAD	93,1
Bu çalışma	2025	Öznitelik Çıkarımı + k-NN	LEN-DB	94,3±0,1
Bu çalışma	2025	Öznitelik Çıkarımı + DT	LEN-DB	94,8±0,1
Bu çalışma	2025	Öznitelik Çıkarımı + EBT	LEN-DB	95,9±0,1

ma sistemi tasarlamışlardır. BFPat yöntemi ile çıkarılan öz-nitelikler, sinyalin yapısal özelliklerini daha iyi temsil etmiş ve bu durum sınıflandırma performansına olumlu katkı sağlamıştır. Model, %93,13 doğruluk oranı elde ederek literatürdeki birçok çalışmayla rekabet edebilir düzeye ulaşmıştır. Literatüre ait sonuçlar Çizelge 5'te verilmiştir.

Bu çalışmaların her biri, farklı model yapıları, veri ön işleme teknikleri ve değerlendirme stratejileri kullanmakla birlikte, ortak amaç deprem ve gürültü sinyallerinin etkili ve güvenilir bir şekilde ayrıştırılmasıdır. Literatürdeki başarılı örneklerle kıyaslandığında, sunulan çalışmada kullanılan EBT modelinin elde ettiği %95,9±0,1 doğruluk oranı, alanındaki mevcut yaklaşımlarla karşılaştırıldığında oldukça rekabetçi ve anlamlı bir başarıdır. Verilerin z-skor normalizasyonu ile standardize edilmesi, modelin eğitim sürecinde öğrenme verimliliğini artırmış ve veriler arasındaki ölçek farklılıklarının etkisini minimize etmiştir. Bu normalizasyon adımı, özellikle farklı kaynaklardan gelen ve çeşitli koşullarda kaydedilen sinyallerin karşılaştırılabilirliğini sağlamak açısından kritik öneme sahiptir. Ayrıca, öz-nitelik çıkarımının titizlikle

gerçekleştirilmesi, modelin karmaşık sinyal yapısını daha iyi anlamasına ve ayırt edici öz-niteliklere odaklanmasına olanak tanımıştır. Çalışmada kullanılan 50.001 deprem ve 50.001 gürültü sinyalinden oluşan dengeli veri seti, modelin genel başarısını güçlendirmiştir. Bu denge, sınıflar arasında öğrenme yanlılığını azaltarak, modelin gerçek dünya uygulamalarında daha tutarlı ve güvenilir sonuçlar üretmesini sağlamaktadır. Ayrıca, bu büyüklükte ve çeşitlilikte bir veri seti, modelin genelleme yeteneğini artırarak, farklı coğrafi bölgelerden ve farklı deprem türlerinden gelen sinyaller üzerinde de etkili performans göstermesine olanak tanımaktadır. EBT algoritması ile elde edilen %95,9±0,1 doğruluk oranı, literatürdeki benzer çalışmalardan daha yüksek performans değerini göstermektedir. Bu sonuç, sadece kullanılan metodolojinin etkinliğini kanıtlamakla kalmayıp, aynı zamanda deprem erken uyarı sistemlerinde ve afet risk yönetiminde kritik öneme sahip güvenilir sinyal ayrıştırma süreçlerinin geliştirilmesine önemli bir katkı sağlamaktadır. Bunun yanı sıra, modelin yüksek performansı, farklı ön işleme teknikleri ve parametre optimizasyonu stratejilerinin dikkatli bir

şekilde seçilmesinden kaynaklanmaktadır. Bu da veri madenciliği ve makine öğrenimi alanlarında uygulanan iyi pratiklerin deprem sinyali ayrıştırma problemlerine başarıyla adapte edilebileceğini göstermektedir. Gelecekte yapılacak çalışmalar, bu yöntemin farklı veri türleri ve gerçek zamanlı uygulamalar için genişletilmesi ve iyileştirilmesi yönünde önemli fırsatlar sunmaktadır. Sonuç olarak, bu çalışma hem teorik hem de pratik anlamda önemli bir adım olup, deprem sinyallerinin otomatik ve doğru tespiti konusunda literatüre değerli bir katkı sağlamaktadır. Elde edilen yüksek doğruluk oranı ve metodolojinin genellenebilir yapısı, bu alandaki diğer araştırmalara ilham verecek ve özellikle erken uyarı sistemlerinde kullanılabilecek yenilikçi yaklaşımların geliştirilmesine zemin hazırlayacaktır.

4. Sonuç ve Öneriler

Bu çalışmanın sunduğu metodoloji, depremlerin daha hassas bir şekilde tespit edilmesine yönelik önemli bir adım atmakta ve sismik analiz alanında kayda değer bir katkı sağlamaktadır. Özellikle, zaman ve frekans alanındaki özniteliklerin kombinasyonu, depremlerin ve gürültü sinyallerinin daha derinlemesine bir analizini mümkün kılmakta, böylece sismik verilerin yorumlanmasında yenilikçi bir bakış açısı sunmaktadır. Çalışmanın kısıtları arasında, toplamda 100.002 veri içeren veri setinin sınırlılığı bulunmaktadır. Ancak, bu veri setinin gelecekte daha da genişletilmesi planlanmaktadır. Gelecek çalışmalarda, daha fazla veri eklenerek, farklı yöntemler kullanılarak öznitelik çıkarımına odaklanılması hedeflenmektedir. Ayrıca, doğrusal olmayan öznitelikler üzerinden daha kapsamlı ölçümler yaparak, modelin performansını artırmak ve daha karmaşık sismik olayları tespit etme yeteneğini geliştirmek amacıyla çalışmalar sürdürülmektedir. Bu yenilikçi yaklaşımlar, sismik verilerin analizi ve yorumlanması konusunda önemli ilerlemeler sağlayabilir ve bu alanda yeni ufuklar açabilir.

Yazar katkısı: Bu çalışmanın tüm aşamalarına (tasarım, veri analizi, yazım ve revizyon dahil) tüm yazarlar eşit düzeyde katkı sağlamıştır.

5. Kaynaklar

Ali, HA., Mohamed, C., Abdelhamid, B., Ourdani, N., El Alami, T. 2022. A comparative evaluation use bagging and boosting ensemble classifiers. 2022 International Conference on Intelligent Systems and Computer Vision (ISCV), 1-6. IEEE. DOI: 10.1109/ISCV54655.2022.9806080

Avcil, F., Işık, E., İzol, R., Büyüksaraç, A., Arkan, E., Arslan, MH., ... Harirchian, E. 2024. Effects of the February 6, 2023, Kahramanmaraş earthquake on structures in Kahramanmaraş city. *Natural Hazards*, 120(3), 2953-2991. DOI: 10.1007/s11069-023-06314-1

Bai, R., Meng, Z., Xu, Q., Fan, F. 2023. Fractional Fourier and time domain recurrence plot fusion combining convolutional neural network for bearing fault diagnosis under variable working conditions. *Reliability Engineering & System Safety*, 232, 109076. DOI: 10.1016/j.ress.2022.109076

Bansal, M., Goyal, A., Choudhary, A. 2022. A comparative analysis of K-nearest neighbor, genetic, support vector machine, decision tree, and long short term memory algorithms in machine learning. *Decision Analytics Journal*, 3, 100071. DOI: 10.1016/j.dajour.2022.100071

Bhatia, M., Ahanger, TA., Manocha, A. 2023. Artificial intelligence based real-time earthquake prediction. *Engineering Applications of Artificial Intelligence*, 120, 105856. DOI: 10.1016/j.engappai.2023.105856

Cai, S., Zhou, J., Pan, J. 2021. Estimating the sample mean and standard deviation from order statistics and sample size in meta-analysis. *Statistical Methods in Medical Research*, 30(12), 2701-2719. DOI: 10.1177/09622802211047

Delastrada, DP., Steffi, S., Hoendarto, G., Tjen, J. 2024. Random Forest Analysis for Predicting the Probability of Earthquake in Indonesia. *Social Science and Humanities Journal*, 9(1), 6295-6304. DOI:10.18535/sshj.v9i01.1574

Erdoğan, YE., Narin, A. 2021. COVID-19 detection with traditional and deep features on cough acoustic signals. *Computers in Biology and Medicine*, 136, 104765. DOI: 10.1016/j.combiomed.2021.104765.

Erdoğan, YE., Narin, A. 2022. Elektrokardiyografi Yardımıyla Hipertansiyonun Otomatik Belirlenmesinde Ampirik Kip Ayrışımının Gürültülü ve Gürültüsüz Sinyaller Üzerindeki Performansının Karşılaştırılması. *El-Cezeri*, 9(2), 788-800. DOI: 10.31202/ecjse.1009456

Essam, Y., Kumar, P., Ahmed, AN., Murti, MA., El-Shafie, A. 2021. Exploring the reliability of different artificial intelligence techniques in predicting earthquake for Malaysia. *Soil Dynamics and Earthquake Engineering*, 147, 106826. DOI: 10.1016/j.soildyn.2021.106826

Giarno, G., Hadi, MP., Suprayogi, S., Murti, SH. 2020. Suitable proportion sample of holdout validation for spatial rainfall interpolation in surrounding the Makassar Strait. *Forum Geografi*, 33(2), 219-232. DOI:10.23917/forgeo.v33i2.8351

Gonzalez, E., Alvarez, L., Mazorra, L. 2012. Normalization and feature extraction on ear images. 2012 IEEE International Carnahan Conference on Security Technology (ICCST), 97-104. IEEE. DOI:10.1109/CCST.2012.6393543

- Haraguchi, T., Emoto, T., Hirayama, T., Imai, Y., Kato, M., Hirano, T. 2023.** Peak-Frequency Histogram Similarity of Bowel Sounds for the Evaluation of Intestinal Conditions. *Applied Sciences*, 13(3), 1405. DOI:10.3390/app13031405
- Hearn, EH., Bürgmann, R., Reilinger, RE. 2002.** Dynamics of Izmit earthquake postseismic deformation and loading of the Duzce earthquake hypocenter. *Bulletin of the Seismological Society of America*, 92(1), 172-193. DOI:10.1785/0120000832
- Jiao, P., Alavi, AH. 2020.** Artificial intelligence in seismology: advent, performance and future trends. *Geoscience Frontiers*, 11(3), 739-744. DOI: 10.1016/j.gsf.2019.10.004
- Kaliappan, J., Bagepalli, AR., Almal, S., Mishra, R., Hu, YC., Srinivasan, K. 2023.** Impact of Cross-validation on Machine Learning models for early detection of intrauterine fetal demise. *Diagnostics*, 13(10), 1692. DOI:10.3390/diagnostics13101692
- Kaur, R., GholamHosseini, H., Sinha, R., Lindén, M. 2022.** Melanoma classification using a novel deep convolutional neural network with dermoscopic images. *Sensors*, 22(3), 1134. DOI: 10.3390/s22031134
- Kelter, R. 2021.** Bayesian model selection in the M-open setting—Approximate posterior inference and subsampling for efficient large-scale leave-one-out cross-validation via the difference estimator. *Journal of Mathematical Psychology*, 100, 102474. DOI: 10.1016/j.jmp.2020.102474
- Kennett, BLN., Engdahl, ER. 1991.** Traveltimes for global earthquake location and phase identification. *Geophysical Journal International*, 105(2), 429-465. DOI: 10.1111/j.1365-246X.1991.tb06724.x
- Liu, C., Fang, D., Zhao, L. 2021.** Reflection on earthquake damage of buildings in 2015 Nepal earthquake and seismic measures for post-earthquake reconstruction. *Structures*, 30, 647-658. Elsevier. DOI: 10.1016/j.istruc.2020.12.089
- Liu, R., Liu, Q., Shi, J., Yu, W., Gong, X., Chen, N., ... Wang, Z. 2021.** Application of a feature extraction and normalization method to improve research evaluation across clinical disciplines. *Annals of Translational Medicine*, 9(20). DOI: 10.21037/atm-21-5046
- Magrini, F., Jozinović, D., Cammarano, F., Micheline, A., Boschi, L. 2020.** Local earthquakes detection: A benchmark dataset of 3-component seismograms built on a global scale. *Artificial Intelligence in Geosciences*, 1, 1-10. DOI: 10.1016/j.aiig.2020.04.001
- Majstorović, J., Giffard-Roisin, S., Poli, P. 2021.** Designing convolutional neural network pipeline for near-fault earthquake catalog extension using single-station waveforms. *Journal of Geophysical Research: Solid Earth*, 126(7), e2020JB021566. DOI:10.1029/2020JB021566
- Ozkaya, SG., Baygin, M., Barua, PD., Tuncer, T., Dogan, S., Chakraborty, S., Acharya, UR. 2024.** An automated earthquake classification model based on a new butterfly pattern using seismic signals. *Expert Systems with Applications*, 238, 122079. DOI: 10.1016/j.eswa.2023.122079
- Shaheed, K., Mao, A., Qureshi, I., Kumar, M., Hussain, S., Ullah, I., Zhang, X. 2022.** DS-CNN: A pre-trained Xception model based on depth-wise separable convolutional neural network for finger vein recognition. *Expert Systems with Applications*, 191, 116288. DOI: 10.1016/j.eswa.2021.116288
- Sholeh, M., Nurnawati, EK. 2024.** Comparison of Z-score, min-max, and no normalization methods using support vector machine algorithm to predict student's timely graduation. *AIP Conference Proceedings*, 3077(1). AIP Publishing. DOI:10.1063/5.0202505
- Singh, AK., Krishnan, S. 2023.** ECG signal feature extraction trends in methods and applications. *BioMedical Engineering OnLine*, 22(1), 22. DOI:10.1186/s12938-023-01075-1
- Tam, A., Barker, J., Rubin, D. 2016.** A method for normalizing pathology images to improve feature extraction for quantitative pathology. *Medical Physics*, 43(1), 528-537. DOI:10.1118/1.4939130
- Tasci, E., Zhuge, Y., Kaur, H., Camphausen, K., Krauze, AV. 2022.** Hierarchical voting-based feature selection and ensemble learning model scheme for glioma grading with clinical and molecular characteristics. *International Journal of Molecular Sciences*, 23(22), 14155. DOI:10.3390/ijms232214155
- Yetkin, M. 2024.** 6 Şubat 2023 Depremleri İçin Kirsaldaki Yiğma Yapıların Performansları Üzerine Bir Saha Araştırması: Nurdağı/Gaziantep Örneği. *Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik Bilimleri Dergisi*, 27(3), 821-837. DOI: 10.17780/ksujes.1430177
- Yapıcı, İŞ., Arslan, RU., Erkamaz, O. 2024.** Kalp yetmezliği tanılı hastaların hayatta kalma tahmininde topluluk makine öğrenme yöntemlerinin performans analizi. *Karaelmas Fen ve Mühendislik Dergisi*, 14(1), 59-69. DOI: 10.7212/karaelmasfen.1429458
- Zhang, Y., Li, H., Du, J., Qin, J., Wang, T., Chen, Y., ... Lei, B. 2021.** 3D multi-attention guided multi-task learning network for automatic gastric tumor segmentation and lymph node classification. *IEEE Transactions on Medical Imaging*, 40(6), 1618-1631. DOI: 10.1109/TMI.2021.3062902