

Elazığ İli için Meteorolojik Ölçüm Verileri Kullanılarak Rüzgar Hızı Tahmini

Wind Speed Prediction Using Meteorological Measurements for Elazığ Province

Serdal POLAT¹ , Nuh ALPASLAN² , İbrahim Rıza HALLAÇ³ 

¹Elektrik-Elektronik Mühendisliği, Bingöl Üniversitesi, Bingöl, Türkiye

²Bilgisayar Mühendisliği, Bingöl Üniversitesi, Bingöl, Türkiye

³Bilgisayar Mühendisliği, Alanya Alaaddin Keykubat Üniversitesi, Antalya, Türkiye

¹spolat@bingol.edu.tr, ²nalpaslan@bingol.edu.tr, ³ibrahim.hallac@alanya.edu.tr

Received: Oct.27, 2023

Accepted: Nov.9, 2023

Published: Dec.20, 2023

Özet: Enerji ihtiyacındaki artış ve çevresel kaygılar nedeniyle yenilenebilir enerji kaynaklarının küresel düzeydeki önemi giderek artmaktadır. Rüzgâr enerjisi, elektrik enerjisi üretiminde son yıllarda giderek daha fazla önem kazanmaktadır. Rüzgâr türbinlerin güvenli işletilmesi için rüzgâr hızı tahmini büyük önem taşımaktadır. Bu çalışmada, Elazığ ilinde farklı bölgelerden elde edilen veriler kullanılarak farklı modellerin rüzgâr hızı tahmin başarıları incelenmiştir. Çalışmada, UKSB, rastgele orman ve XGBoost modelleri kullanılmıştır. Veri seti mevsimsellik ve eğilim bileşenleri MEAY yöntemiyle ayrıştırılmış ve Fourier dönüşümü ile mevsimsel bileşenler belirlenmiştir. Elde edilen sonuçlar, farklı bölgelerde farklı modellerin daha iyi performans gösterdiğini göstermektedir. Sonuçlara göre, Elazığ Merkez, Keban ve Sivrice bölgesinde genelleştirilmiş gradyan artışı ve rastgele orman modelleri en düşük RMSE ve MSE değerlerine sahiptir, bu da bu modellerin bu bölge için daha iyi tahminler yaptığını göstermektedir.

Anahtar kelimeler: Hibrit model, UKSB, Rasgele Orman, genelleştirilmiş gradyan artışı, MEAY, FFT ve rüzgar hızı tahmini.

Abstract: As a result of the increasing energy demand and growing environmental concerns, the global significance of renewable energy resources is steadily rising. Wind energy has been increasingly gaining importance in electricity generation in recent years. The accurate prediction of wind speed is crucial for the safe operation of wind turbines. In this study, wind speed prediction performance of different models was examined using data obtained from various regions in the Elazığ province. LSTM, random forest, and XGBoost models were employed in the study. The dataset was decomposed into seasonal and trend components using the STL method, and seasonal components were determined using Fourier transformation. The results indicate that different models perform better in different regions. According to the findings, XGBoost and random forest models exhibit the lowest RMSE and MSE values in Elazığ, Keban, and Sivrice regions, indicating better predictions for these models in these areas.

Key words: Hybrid model, LSTM, Random Forest, XGBoost, STL, FFT and wind speed prediction.

1. Giriş

Günümüzde enerji ihtiyacının sürekli artış, çevresel kaygılar ve enerji güvenliği endişeleri, yenilenebilir enerji kaynaklarının küresel düzeyde ve ülkemizdeki önemini daha da belirgin hale getirmiştir. Fosil yakıtların sınırlı ve çevreye zararlı oluşu, iklim değişikliği ve çevresel sorunlarla mücadele gerekliliği, yenilenebilir enerji kaynaklarını araştırma, geliştirme ve uygulama konularını öncelikli hale getirmiştir (Ghoniem, 2011). Hem dünya genelinde hem de ülkemizde, güneş, rüzgâr, biyokütle ve jeotermal gibi kaynakların enerji portföyümüzdeki payını artırmak, sürdürülebilir bir gelecek inşa etmek ve ekonomik kalkınmayı desteklemek adına büyük bir öneme sahiptir (Gielen et al., 2019).

Türkiye, coğrafi konumu ve iklim özellikleriyle yenilenebilir enerji kaynaklarının geniş bir yelpazesine sahip olan bir ülke olarak öne çıkmaktadır (Alkan & Albayrak, 2020). Son yıllarda yaşanan hızlı sanayileşme, artan nüfus ve enerji talebi, enerji alanında sürdürülebilir ve çevre dostu çözümlerin arayışını daha da acil hale getirmiştir (Hamzacebi & Es, 2014). Fosil yakıt rezervlerinin sınırlı olması, çevre kirliliği ve iklim değişikliği gibi küresel zorluklarla birlikte, Türkiye'nin de enerji politikalarını yeniden gözden geçirmesine ve yenilenebilir enerji kaynaklarına yönelik yatırımları artırmasına yol açmıştır. Bu çalışmada ülkemizin yenilenebilir enerji potansiyelini, bu kaynakların ülke ekonomisi, enerji güvenliği ve çevre koruma açısından taşıdığı önemi ele alınacak ve Türkiye'nin sürdürülebilir enerji geleceğini şekillendirmedeki rolü incelenecektir.

Elektrik enerjisi üretiminde rüzgâr enerjisi, son yıllarda giderek daha fazla önem kazanan ve sürdürülebilir enerji portföyünün vazgeçilmez bir parçası haline gelen bir kaynaktır. Rüzgâr enerjisi, rüzgâr türbinleri aracılığıyla

rüzgârın kinetik enerjisini mekanik enerjiye dönüştürerek elektrik üretimini sağlar. Bu teknoloji, çevre dostu olması, sınırsız bir kaynak oluşturması ve karbon salınımını minimumda tutması nedeniyle enerji üretimindeki geleceğe yönelik bir çözüm olarak öne çıkmaktadır. Rüzgâr enerjisi üretimi, temiz ve yenilenebilir bir kaynağın kullanılması açısından çevre koruma konusunda büyük avantajlar sunar. Fosil yakıtlara kıyasla rüzgâr enerjisi üretimi, sera gazı emisyonlarını azaltır ve hava kalitesini artırır. Bu durum, iklim değişikliği ile mücadelede etkili bir adım atılmasını sağlar. Ayrıca, rüzgâr enerjisi santrallerinin faaliyete geçirilmesiyle birlikte çevresel tahribatın önüne geçilir ve ekosistemlere olan olumsuz etkiler minimize edilir. Türkiye'nin coğrafi konumu, rüzgâr enerjisi potansiyeli açısından oldukça avantajlıdır. Özellikle Doğu Anadolu bölgesi yüksek ve sürekli esen rüzgârlarıyla rüzgâr enerjisi projeleri için uygun alanlar sunar. Bu bölgelerde kurulan rüzgâr çiftlikleri, ulusal enerji güvenliğini artırırken aynı zamanda ekonomik kalkınmaya da katkı sağlar. Rüzgâr türbinleri, kinetik rüzgâr enerjisini önce mekanik enerjiye daha sonra da elektrik enerjisine dönüştürür. Etkili elektrik üretimi için kesintisiz ve güçlü rüzgâr hızlarına gereksinim duyulmaktadır. Ancak, rüzgâr hızının değişkenliği elektrik üretimi için önemli bir problemdir. Bu problemin çözümü için kısa-dönem rüzgâr hızı tahmini yapmak gerekmektedir. Rüzgâr tarlalarının ve türbinlerin güvenli işletilmesi ve kontrol edilmesi için kısa-dönem rüzgâr hızı tahmini büyük önem taşımaktadır. Bu tahminler, enerji üretimini optimize etmek, türbinlerin ve ekipmanın hasar görmesini önlemek ve güvenli bir çalışma ortamı sağlamak amacıyla kullanılır. Kısa-dönem rüzgâr hızı tahminleri, meteorolojik verilerin analizi ve öngörülerin yapılması süreçleri ile gerçekleştirilir. Kısa-dönem rüzgâr hızı tahminleri, türbinlerin çalışma hızlarını ve yüksekliklerini ayarlamak, enerji üretimini en üst düzeye çıkarmak ve tesis güvenliğini sağlamak için kullanılır. Bu tahminlerin doğruluğu, rüzgâr enerjisi tesislerinin verimliliği ve sürdürülebilirliği açısından kritik bir rol oynar. Derin öğrenme ve makine öğrenmesi teknikleri, rüzgâr hızı tahmininde artan bir ilgi görmektedir. Bu teknikler, rüzgâr enerjisi üretim optimizasyonundan hava taşımacılığına kadar çeşitli alanlarda kullanılmaktadır. Smith & Johnson, derin öğrenme yöntemlerinden biri olan evrişimli sinir ağları (CNN) kullanılarak rüzgâr hızı tahmini gerçekleştirilmiştir. Çalışma, meteorolojik verileri ve rüzgâr hızı zaman serilerini kullanarak tahmin modeli geliştirmiştir. Elde edilen sonuçlar, CNN'nin rüzgâr hızı tahmininde geleneksel yöntemlere kıyasla daha yüksek doğruluk sağladığını göstermektedir. (Du, 2019), rüzgâr enerjisi üretimini optimize etmek amacıyla makine öğrenmesi yöntemlerini kullanmışlardır. Yapay sinir ağları, karar ağaçları ve rastgele orman (RO, random forest) gibi algoritmalar kullanılarak rüzgâr hızı tahmin modelleri geliştirilmiştir. Araştırma, tahmin modellerinin rüzgâr enerjisi üretimi optimizasyonunda etkili olduğunu göstermektedir. (Liu et al., 2018), rüzgâr hızı tahmini için mevcut zaman serisi tahmin yöntemi olan ARIMA ile tekrarlı sinir ağları (RNN) yöntemleri birleştirilerek hibrit bir model geliştirilmiştir. Çalışma, bu modelin rüzgâr hızı tahmininde daha yüksek doğruluk sağladığını ve ARIMA ve RNN modellerine göre daha etkili olduğunu göstermektedir. (Prabha et al., 2019), UKSB (long short-term memory, uzun kısa süreli bellek) sinir ağı kullanarak rüzgâr hızı tahmini yapmıştır. Dalgacık dönüşümü temelli özellik çıkarımı ile tahmin yapma yaklaşımı benimsenmiştir. Elde edilen sonuçlar, UKSB'nin rüzgâr hızı tahmininde etkili olduğunu göstermektedir.

Rüzgâr hızı tahmini literatürde nem, basınç ve sıcaklık gibi farklı parametrelerin işlenmesi veya rüzgâr hızlarının bir zaman serisine dönüştürüp işlenmesi şeklinde 2 farklı türde yapılmaktadır. Bu çalışmada, Elâzığ Meteoroloji Müdürlüğünden alınan Elâzığ ili merkez ve ilçelerinin geçmiş yıllara ait saat başlarında okunan rüzgâr hızı verilerinden günlük ortalama rüzgâr hızı verisi oluşturularak zaman serisi oluşturulmuş, modeller kullanılarak gelecek zamanlara ait günlük rüzgâr hızları tahmin edilmiştir. Geliştirilen model, UKSB, RO ve XGBoost(eXtreme gradient boosting, genelleştirilmiş gradyan artışı) algoritmalarını içermektedir. Önerilen model, rüzgâr enerjisi tesislerinin performansını artırabilir ve enerji üretimini optimize etmeye yardımcı olabilir.

Kullanılan veri seti, Elâzığ ili merkez ve ilçelerinin farklı bölgelerindeki rüzgâr hızı tahminlerinin geliştirilmesi amacıyla kullanılmıştır. Bu veri seti, çeşitli meteorolojik özelliklerin yanı sıra günlük ortalama rüzgâr hızı, günlük ortalama rüzgâr yönü, günlük maksimum ve minimum rüzgâr hızı gibi parametreleri içermektedir. Veri setinin ön işleme aşamasında, MEAY (mevsimsel ve eğilim ayrıştırması, seasonal and trend decomposition using Loess, STL) ve YASGY (yerel ağırlıklı saçılım grafiği yumuşatması) yöntemi kullanılmıştır. Bu yöntem, zaman serisi verilerinin mevsimsel ve eğilim bileşenlerine ayrılmasına olanak sağlamaktadır.

Bu çalışmanın temel amacı, Elâzığ ilindeki farklı bölgelerdeki rüzgâr hızını tahmin etmek için en uygun modeli belirlemek ve bu modelin performansını artırarak daha kesin tahminler yapmaktır. Elde edilen sonuçlar, farklı bölgelerde farklı modellerin daha iyi performans sergileyebileceğini göstermektedir. Bu bulgular, bölgeye özgü modellerin rüzgâr hızı tahmininde etkili olabileceğini ve enerji üretimi planlamalarında kritik bir rol oynayabileceğini vurgulamaktadır. Sonuç olarak, bu çalışma Elâzığ ilindeki farklı bölgelerdeki rüzgâr hızı tahminini optimize etmek için üç model önermekte ve bu modelin rüzgâr enerjisi üretiminde daha güvenilir ve etkili bir yaklaşım sunabileceğini göstermektedir. Çalışmamızı literatürdeki çalışmalardan farklı kılan özellikler aşağıda verilmiştir. Öncelikle, bugüne kadar ki, Elâzığ ili rüzgâr hızı tahmin çalışmalarında Weibull ve Rayleigh dağılımına ait parametre tahminleri kullanılarak yapılmıştır. Bu yönüyle çalışmamız, Elâzığ ilinin rüzgâr hızı tahmininde derin öğrenme modeli kullanan ilk çalışma olmaktadır. Bunun yanında elde edilen veri seti çalışmamızın en büyük yeniliğidir. Ayrıca, UKSB derin öğrenme modeli, rastgele orman ve genelleştirilmiş

gradyan artışı (XGBoost) yöntemleri kullanılarak hangi yapının Elazığ ili için rüzgâr hızı tahmininde daha iyi sonuç vereceğinin araştırması yapılmıştır.

Çalışmanın içeriği şu şekildedir: İkinci bölümde çalışmada kullanılan materyal ve metotlardan bahsedilmiştir. Üçüncü bölümde önerilen yaklaşımdan detaylı bir şekilde bahsedilmiştir. Dördüncü bölümde rüzgar hızı tahminine ilişkin deneysel sonuçlar verilmiştir. Son bölümde, çalışmanın önemi vurgulanmıştır.

2. Materyal

Bu bölümde, Elazığ ilinin farklı bölgeleri (Merkez, Sivrice İlçesi ve Keban İlçesi) içinde gerçekleştirilen rüzgar hızı ölçümlerini içeren veri setinin toplanması, düzenlenmesi ve ön işlem adımları detaylı bir şekilde açıklanacaktır. Bu veri seti, enerji üretimini optimize etmeyi amaçlayan doğru rüzgar hızı tahminlerinin geliştirilmesi amacıyla temel veri kaynağını oluşturmaktadır.

Veri toplama aşaması, saha ölçümlerinden elde edilen ham verilerin toplanması ile başlar. Bu ham veriler, rüzgar hızı değerlerini içermekte olup, her bir veri örneği ayrıca Yıl, Ay, Gün gibi tarih bilgilerini içermektedir. Aynı zamanda günlük ortalama rüzgar hızı, günlük ortalama rüzgar yönü, günlük maksimum rüzgar hızı ve günlük minimum rüzgar hızı gibi çeşitli meteorolojik özellikleri içermektedir. Veri düzenleme aşamasında, toplanan ham veriler tarih bilgileriyle birlikte düzenlenmiştir. Bu düzenleme sayesinde zaman serisi yapısı oluşturulmuş, böylece rüzgar hızı varyasyonlarının mevsimsel ve günlük desenleri analiz edilebilir hale getirilmiştir (Box, 1989). Veri setinin temizlenmesi ve ön işlem aşamaları, istatistiksel analizler ve veri madenciliği yöntemleri kullanılarak gerçekleştirilmiştir. Bu adımlar ile veri setindeki anomaliler, eksik veriler ve aykırı değerleri ele alınmıştır. Ayrıca, veri setinin dengelenmesi ve özellik çıkarımı da bu aşamalarda gerçekleştirilmiştir (Hastie et al., 2009). Bu işlemler sonucunda, tahmin modellerine girdi olarak verilecek öznitelikler elde edilmiştir. Kullanılan veri setine ait örnekler Tablo 1’de görülmektedir.

Bu çalışmada kullanılan veri seti, yukarıda açıklanan adımların tamamlanmasının ardından elde edilmiştir. Elde edilen düzenlenmiş ve işlenmiş veri seti, rüzgar hızı tahmini modellerinin eğitilmesi ve değerlendirilmesi için kullanılacaktır. Bu veri seti, enerji üretimi alanında doğru ve etkili tahminlerin geliştirilmesine katkı sağlayacaktır.

Tablo 1: Elazığ ili merkez için örnek veri seti

Yıl	Ay	Gün	Saat	Dakika	Günlük Ortalama Rüzgâr Hızı	Günlük Ortalama Rüzgâr Yönü	Günlük Maksimum Rüzgâr Hızı	Günlük Minimum Rüzgâr Hızı
2018	1	1	0	0	2.88	150	5.1	1
2018	1	2	0	0	2.84	270.94	5.4	0.9
2018	1	3	0	0	2.2	168.75	4.8	0.9
2019	3	9	0	0	1.79	243.75	2.6	1
2019	3	10	0	0	1.64	241.88	2.6	0.9
2019	3	11	0	0	1.56	209.06	3	0.9
2020	5	3	0	0	2.32	177.19	5.1	1
2020	5	4	0	0	3.88	135.94	6.6	1
2020	5	5	0	0	1.9	174.38	4.4	0.9
2021	8	22	0	0	1.6	266.25	2.9	0.9
2021	8	23	0	0	2.21	198.75	4.3	1
2021	8	24	0	0	2	276.56	4.6	0.9
2022	12	29	0	0	1.39	281.25	2.8	0
2022	12	30	0	0	1.45	249.55	2.5	0
2022	12	31	0	0	1.33	280.13	2.5	0

2.1. Zaman Serisi Analizi için Veri Formatının Düzenlenmesi

Bu bölümde, veri setinin zaman serisi analizi için uygun bir formata dönüştürülmesi süreci ayrıntılı bir şekilde açıklanacaktır. Bu dönüşüm, veri setinin zaman serisi analizi yöntemleriyle daha etkili bir şekilde işlenebilmesini ve anlaşılabilirliğini amaçlamaktadır.

İlk adımda, veri setinde bulunan ayrı ayrı tarih bilgileri (Yıl, Ay, Gün) birleştirilerek tek bir zaman sütunu oluşturulmuştur. Bu birleştirme işlemi, zaman serisi verilerinin süresiz bir yapıdan kesintisiz bir zaman akışını temsil eden homojen bir formata dönüştürülmesini sağlamıştır. Böylece, her bir veri noktası belirli bir zaman noktasına bağlı olarak analiz edilebilmektedir.

Bu düzenleme işlemi sırasında, günlük ortalama rüzgar hızı, günlük ortalama rüzgar yönü, günlük maksimum rüzgar hızı ve günlük minimum rüzgar hızı gibi diğer özellikler de uygun bir formata getirilmiştir. Her bir zaman noktası için bu özellikler düzenlenmiş bir şekilde veri setine entegre edilmiştir. Bu düzenleme sayesinde, zaman serisi analizine tabi tutulacak özelliklerin homojen bir yapıda ve zamana bağlı olarak düzenlenmiş olduğu bir veri formatı oluşturulmuştur.

Veri setinin bu şekilde düzenlenmesi, zaman serisi analizi için önemli bir temel oluşturmaktadır (Brockwell & Davis, 2002). Zaman serisi verilerinin düzenli bir yapıda olması, analiz işlemlerinin daha sağlam ve güvenilir bir şekilde yürütülmesine olanak sağlar. Ayrıca, bu düzenleme işlemi sayesinde veri seti, zaman içindeki değişimleri daha etkili bir şekilde yansıtan ve analiz edilen özelliklerin zaman içindeki eğilimlerini daha iyi görünür kılan bir yapıya kavuşmuş olur.

2.2. Mevsimsellik ve Eğilim Bileşenlerinin MEAY Yöntemi ile Ayrıştırılması

Veri setindeki mevsimsellik ve eğilim bileşenlerini belirlemek amacıyla, MEAY yöntemi kullanılmıştır. MEAY yöntemi, veriyi mevsimsel, eğilim ve düzensiz bileşenlere ayrıştırarak daha ayrıntılı bir analiz imkanı sunmaktadır. MEAY yönteminde öncelikle düzleştirme (smoothing) işlemi uygulanır. Bu aşamada, veri seti $y(t)$, Loess regresyonu kullanılarak düzleştirilir ve düzleştirilmiş veri seti $s(t)$ elde edilir. Sonrasında mevsimsel bileşen tahmini (seasonal component estimation) işlemi uygulanır. Bu aşamada, düzleştirilmiş veriden mevsimsel desenleri çıkarmak amacıyla LOESS regresyonu kullanılarak mevsimsel bileşen $m(t)$ tahmini elde edilir. Sonrasında, düzleştirilmiş veriden mevsimsel bileşen çıkarılarak artıklık bileşeni (residual component) $r(t)$ elde edilir (Cleveland & Grosse, 1991).

Düzleştirme (Smoothing): Bu aşama, LOESS regresyonu kullanılarak veri setinin düzleştirilmesini içerir. Formül şu şekildedir:

$$s(t) = \text{Loess}(y(t))$$

Mevsimsel Bileşen Tahmini (Seasonal Component Estimation): Bu aşama, düzleştirilmiş veriden mevsimsel desenleri çıkarmak amacıyla LOESS regresyonu kullanılır. Formül şu şekildedir:

$$m(t) = \text{Loess}(y(t) - s(t))$$

Burada, her iki durumda da, LOESS regresyonunun formülü şu şekildedir:

$$y_i = g(x_i) + \hat{\epsilon}_i$$

Burada, y_i yanıtın i 'nci ölçümünü, x_i p öngörücünün i 'nci ölçümünü, g regresyon fonksiyonunu ve $\hat{\epsilon}_i$ rastgele bir hatayı temsil eder.

Artıklık Bileşeni (Residual Component): Bu aşama, düzleştirilmiş veriden mevsimsel ve eğilim bileşenlerinin çıkarılması ile elde edilir. Artıklık bileşeni hesaplaması şu şekildedir:

$$r(t) = y(t) - s(t) - m(t)$$

Burada, $r(t)$ artıklık bileşenini, $y(t)$ orijinal zaman serisi verisini, $s(t)$ düzleştirilmiş veriyi ve $m(t)$ mevsimsel bileşen tahminini temsil eder.

MEAY yöntemi, veri setini mevsimsel, eğilim ve artıklık bileşenlerine ayırarak her bir bileşenin daha detaylı analiz edilmesini sağlar. Mevsimsel bileşen, verinin periyodik tekrar eden desenlerini temsil ederken, eğilim bileşeni verinin genel eğilimini gösterir ve artıklık bileşeni ise mevsimsel ve eğilim bileşenlerinden geriye kalan varyasyonu ifade eder. Bu aşamada MEAY yönteminde Fourier dönüşümü kullanılarak veri setinin mevsimsel ve eğilim bileşenlerine nasıl ayrıştırıldığı ayrıntılı bir şekilde açıklanacaktır.

Scipy kütüphanesinin içerdiği hızlı Fourier dönüşümü (FFT) fonksiyonu, bu amaçla kullanışlı bir araçtır. FFT'nin matematiksel formülü aşağıdaki gibidir:

$$Y(k) = \sum_{n=0}^{N-1} (x[n]W_N^{jk}), W_N = e^{-j2\pi/N}, k = 0 \leq k \leq N-1$$

Burada $Y(k)$, k . frekans bileşenini temsil eder. $x(n)$, n . veri noktasını temsil eder. N , veri setinin toplam nokta sayısını temsil eder. j , karmaşık birimdir.

Mevsimsel bileşenin tespit edilmesi için Fourier dönüşümü yöntemi başarıyla kullanılmıştır. Fourier dönüşümü, veri setindeki mevsimsel desenleri tanımlamak ve mevsimsel bileşeni ayırmak için etkili bir matematiksel araçtır. Fourier dönüşümü, veri setindeki dalgasal desenleri frekans bileşenlerine ayırtmak için yaygın olarak kullanılır (Petropoulos et al., 2022). Bu sayede mevsimsel varyasyonların nasıl oluşturulduğunu ve bileşenlerine ayırtıldığı anlamamıza yardımcı olur. Bu yöntemin kullanılmasıyla elde edilen karmaşık sayıların gerçek bileşenleri, mevsimsel bileşenleri temsil eder.

Bu adımdan sonra veri normalizasyon işlemi gerçekleştirilir. Veri setindeki rüzgar hızı değerleri, farklı ölçeklere sahip olduklarından analiz ve işlemler için aynı aralığa getirilmiştir. Bu amaçla Min-Max normalizasyon yöntemi kullanılmış ve veriler 0 ile 1 arasında ölçeklendirilmiştir. Verinin normalleştirilmesi için Scikit-learn kütüphanesinde bulunan MinMaxScaler sınıfı kullanılmıştır. Bu normalizasyon işlemi, farklı ölçeklere sahip verilerin aynı aralığa getirilerek, daha tutarlı ve karşılaştırılabilir sonuçların elde edilmesini sağlamaktadır (James et al., 2013). Uygulama intel UHD 630 8207MB ekran kartı, intel: i5-1050 işlemcili 16384MB RAM'a sahip bilgisayarda çalıştırılmıştır.

3. İlgili Çalışmalar

3.1. UKSB (LSTM-long short-term memory, Uzun Kısa Süreli Bellek)

UKSB, derin öğrenme alanında yaygın olarak kullanılan bir yinelemeli sinir ağı (RNN, recurrent neural network) türüdür. UKSB, özellikle sıralı verileri işlemek için tasarlanmıştır ve uzun vadeli bağımlılıkları yakalamak için geliştirilmiştir. Bu mimari, tipik RNN'lerin karşılaştığı sorunları aşar. UKSB'nin temel özelliği, hücre adı verilen özel bir bellek birimi içermesidir. Bu hücre, bilgiyi saklamak, silmek ve güncellemek için öğrenilebilir gizli birimlerden oluşur. UKSB, özellikle diziler, zaman serileri ve doğal dil işleme (NLP) gibi uygulama alanlarında yaygın olarak kullanılır. Bu sayede, uzun vadeli bağımlılıkları modellemek ve sıralı verilerdeki gizli yapıları daha iyi anlamak mümkün hale gelir. Bu çalışmada, rüzgar hızı tahmininde kullanılmak üzere bir UKSB modeli tanımlanmıştır. UKSB modeli yapısı Şekil 1'de görülmektedir.

Giriş Katmanı: Modelin giriş katmanı, 32 birim içermektedir. Bu katmanda, veri setindeki girdi özelliklerinin işlenmesi ve ön işleme gerçekleştirilir.

Dropout Katmanı: İlk UKSB katmanından sonra %10 dropout oranı ile bir dropout katmanı yer almaktadır. Dropout, ezberlemeyi (overfitting) önlemek ve modelin genelleştirme yeteneğini artırmak için kullanılan bir tekniktir.

İkinci UKSB Katmanı: Modelin ikinci katmanı, 64 birim içermektedir. İkinci bir UKSB katmanının eklenmesi, modelin daha karmaşık örüntüleri öğrenmesine olanak tanır.

Dropout Katmanı: İkinci UKSB katmanından sonra tekrar %10 dropout oranı ile bir dropout katmanı bulunmaktadır.

Çıkış Katmanı: Modelin çıkış katmanı, 1 birim içermektedir. Bu katmanda, tahmin edilen rüzgar hızı değeri üretilir.

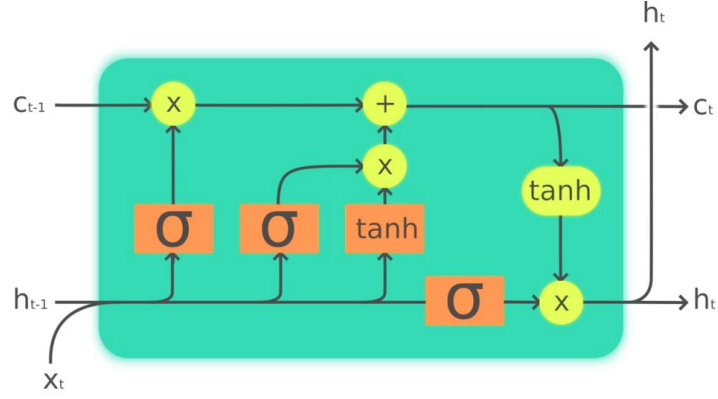
Modelin eğitimi sırasında Adamax optimizyer kullanılarak kayıp (loss) fonksiyonu minimize edilmeye çalışılmıştır. Bu UKSB modeli, zaman serisi verilerinin içerdiği zaman bağımlılıklarını ve örüntüleri öğrenme yeteneği ile rüzgar hızı tahmininde etkili sonuçlar elde etmeyi amaçlar (Hochreiter & Schmidhuber, 1997; Yan et al., 2022).

3.2. Rastgele Orman

RO modeli, ensemble (bir araya getirilmiş) bir öğrenme algoritmasıdır ve birden fazla karar ağacının birleşerek oluşturduğu bir yapıdır (Breiman, 2001). Bu çalışmada, rüzgar hızı tahmininde kullanılmak üzere bir Random Forest modeli tanımlanmıştır.

RO modelinde, toplam 100 karar ağacı bulunmaktadır. Bu ağaçlar, veri setini farklı alt örneklemelerle (bootstrap örnekleme) kullanarak eğitilir ve sonuçlarını bir araya getirerek tahmin yapar. Her bir ağaç, veri setinin farklı özelliklerine göre bağımsız olarak kararlar verir. Sonuç olarak, tüm ağaçların tahminleri bir araya getirilerek genel tahmin üretilir (Brownlee, 2023).

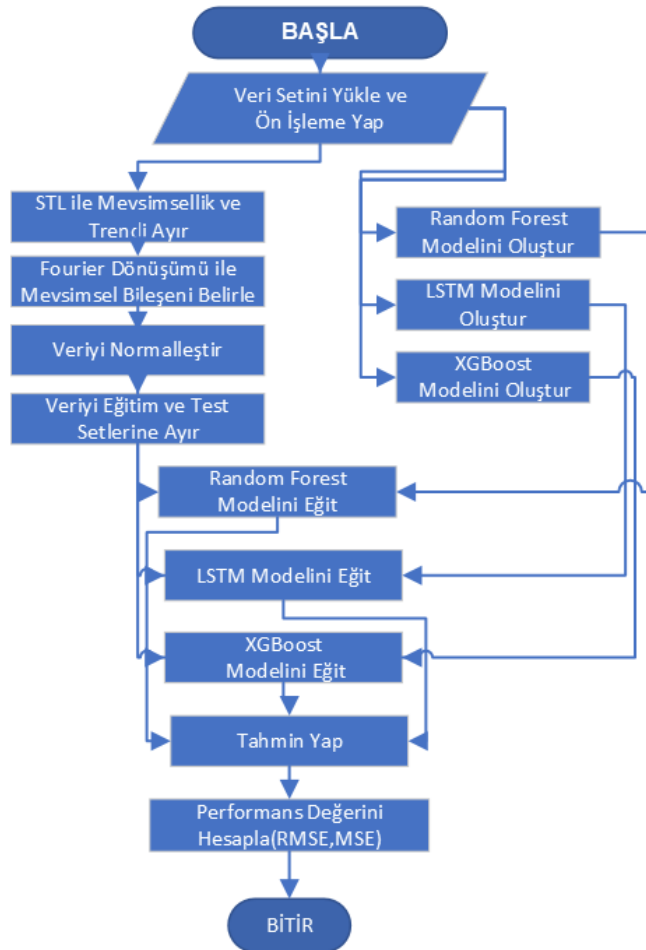
RO modeli, birden fazla ağacın bir araya gelmesiyle oluşturulduğu için ezberlemeyi (overfitting) engellemeye yardımcı olur ve veri setindeki özellikler arasındaki ilişkileri yakalamak için etkili bir yöntemdir (Liaw & Wiener, 2002). Bu sayede, model hem istikrarlı hem de veri setinin özelliklerini doğru bir şekilde anlamada etkili olabilir.



Şekil 1: UKSB mimarisi

3.3. Genelleştirilmiş Gradyan Artışı (XGBoost)

XGBoost, makine öğrenme alanında önemli bir başarı elde etmiş bir ensemble öğrenme algoritmasıdır. Gradient Boosting yöntemlerinin bir türü olan XGBoost, özellikle sınıflandırma ve regresyon problemlerinde yüksek başarı oranları sunar (Chen & Guestrin, 2016). Temel olarak, zayıf öğreniciler olarak adlandırılan basit modelleri (genellikle karar ağaçları) bir araya getirir ve bu modellerin hatalarını düzelterek güçlü bir tahmin modeli oluşturur. Modelin temel özelliği, toplam 100 tahminleyici (estimator) kullanılarak oluşturulmasıdır. Her bir tahminleyici, önceki tahminleyicilerin hatalarını düzelterek ve veri setinin özelliklerini daha iyi yakalayarak tahmin yapar. XGBoost, özellikler arasındaki karmaşıklıkları daha iyi modellemek için gradient boosting algoritmasının ötesine geçer. Bu alanda gerçekleştirdiği yenilikler sayesinde özellik seçimi ve özellik önem sıralaması gibi yöntemlerle veri setindeki önemli değişkenleri belirlemek ve kullanmak için etkili bir araç sunar.



Şekil 2: Çalışmaya ait akış diyagramı

Ayrıca, XGBoost modeli hızlı bir şekilde eğitilebilir ve büyük veri setleriyle başa çıkabilir (XGBoost Documentation, 2023). Bu avantajları sayesinde, rüzgar hızı tahmininde XGBoost'un kullanılması, modelin performansını artırmak ve daha doğru tahminler elde etmek için önemli bir fırsat sunar. Çalışmaya ait akış diyagramı Şekil 2'de görülmektedir.

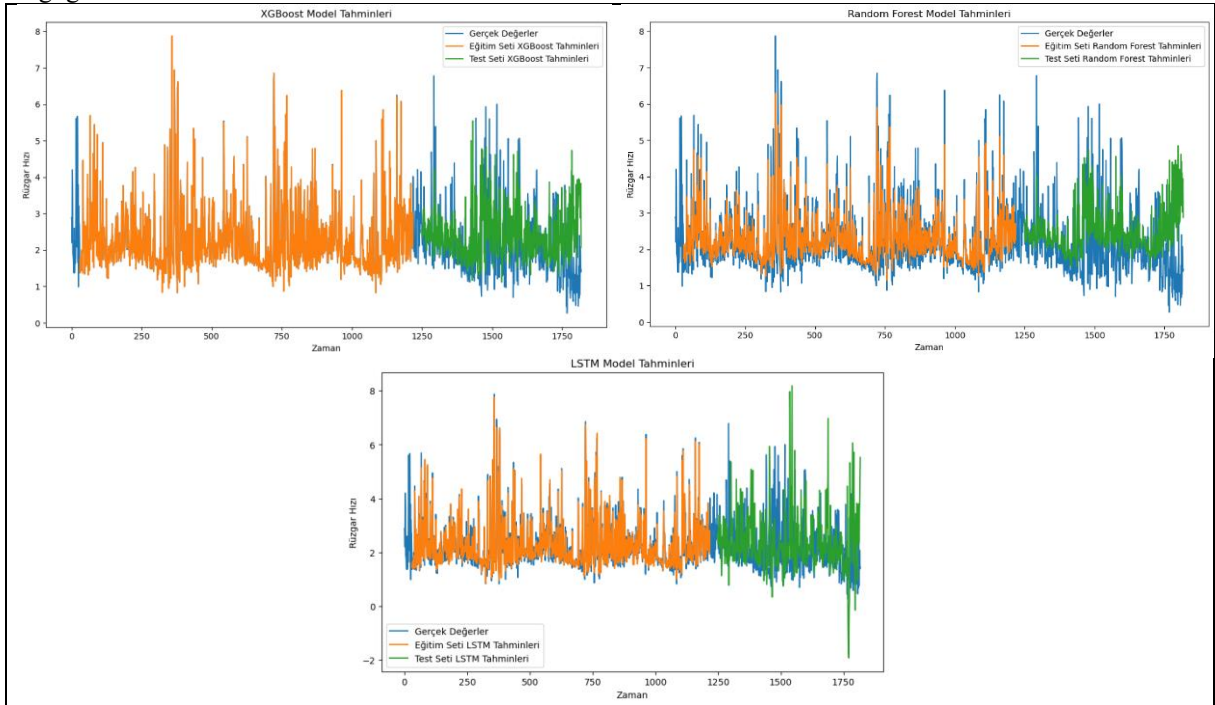
4. Deneysel Sonuçlar

Çalışma kapsamında Elazığ Meteoroloji Müdürlüğünden alınan 01.01.2018 – 31.12.2022 tarih aralığında Elazığ ili merkez ve keban ve sivrice ilçelerinin geçmiş yıllara ait saat başlarında okunan rüzgâr hızı verileri kullanılarak rüzgâr hızı tahmini yapılmıştır. Öncelikle belirtilen bölgelerden elde edilen 5 yılı kapsayacak şekilde her gün saatlik çözünürlükte alınan 5460 adet veri ile çalışılmıştır. Eğitim sonrasında, her bir model için tahminler yapılmış ve performans değerlendirmesi gerçekleştirilmiştir. Tahmin sonuçları gerçek verilerle karşılaştırılarak RMSE (root mean square error, karekök ortalama karesel hata), MSE (mean square error, ortalama karesel hata) ve R-Kare (coefficient of determination, belirlilik katsayısı) değerleri hesaplanmıştır. Verilerin %67'si eğitim ve %33'ü test kümesi olarak kullanılmıştır.

Tablo 2: Elazığ ili merkez için 3 farklı yöntemin test seti performans sonuçları

Yöntem	RMSE	MSE
XGBoost	0.1457	0.0212
RO	0.1484	0.0220
UKSB	0.1961	0.0385

Tablo 2'de, Elazığ ili merkez için XGBoost, RO ve UKSB modellerin test seti için rüzgar tahmini performans sonuçları görülmektedir. Tablo 1'de UKSB modelinin eğitim setinde düşük RMSE (0.0236) ve MSE (0.0006) değerleri elde ettiği görülmektedir. UKSB modeli test setinde ise düşük RMSE ve MSE değerleri elde etmiştir. Rastgele Orman modeli eğitim setinde düşük RMSE (0.0426) ve MSE (0.0018173413) değerleri elde etmiştir. Bunun yanında, RO modeli test setinde düşük RMSE ve MSE değerleri elde etmiştir. XGBoost modelinde eğitim setinde düşük RMSE (0.0011) ve MSE (8.45×10^{-7}) değerleri elde edilmiştir. Bununla birlikte, test setinde oldukça düşük RMSE ve MSE değerleri elde edilmiştir. Elde edilen sonuçlar incelendiğinde 3 yöntemin de birbirine yakın sonuçlar elde ettiği görülmektedir. Bu durum, her üç modelin de yüksek genelleme yeteneğine sahip olduğunu göstermektedir. Bunun yanında, Elazığ ili merkez için XGBoost yönteminin en iyi MSE ve RMSE değerleri elde ettiği görülmektedir.



Şekil 3: Elazığ ili merkez için 5 yıllık, günlük çözünürlükte rüzgar hızı tahmin ve gerçek değerler

Şekil 3'de, Elazığ ili merkez için XGBoost, RO ve UKSB modellerinin eğitim ve test setine ait tahmin başarımları sonuçları görülmektedir. Modelin eğitim setinde tahminleri gerçek değerlere oldukça yakınken, test setinde

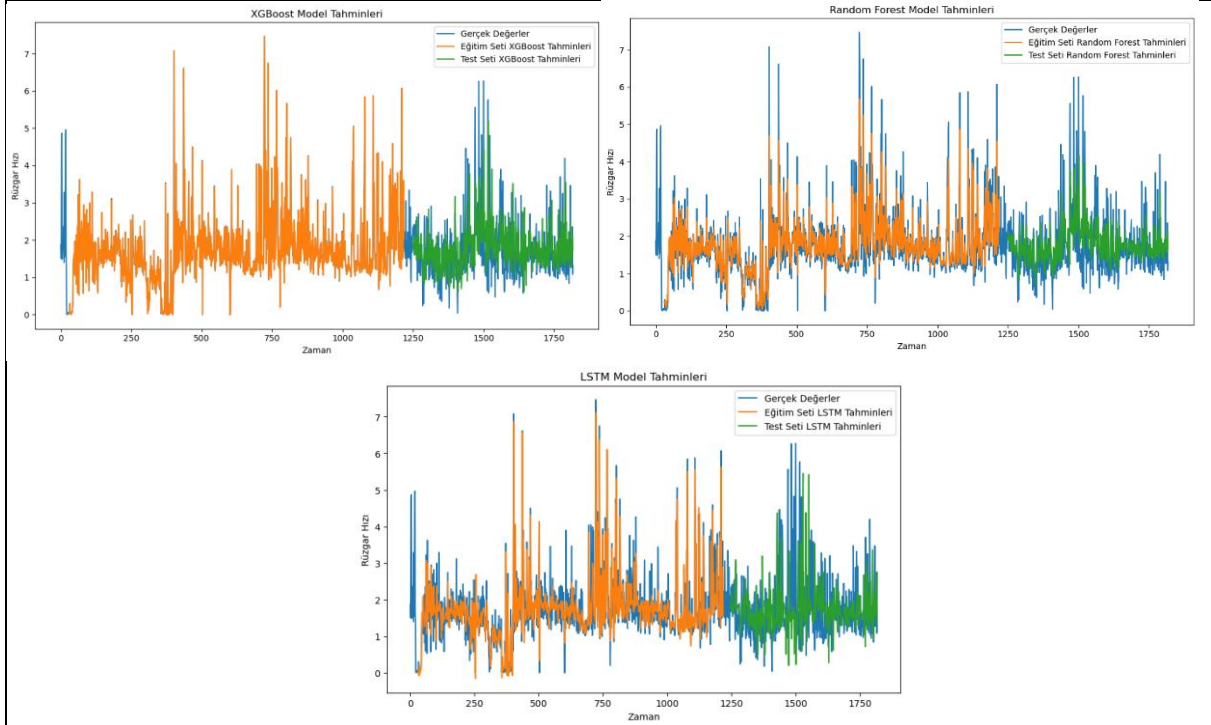
tahminlerin gerçek değerlere göre daha fazla sapma gösterdiği görülmektedir. XGBoost yönteminin diğer iki yönteme göre rüzgar hızı verilerini daha iyi ifade ettiği ve tahmin yaptığı görülmektedir.

XGBoost UKSB'den daha iyi rüzgar hızı tahmini yapmıştır. Bu durumun nedeni aşağıdaki gibi açıklanabilir. XGBoost, bir gradyent boosting algoritmasıdır ve karar ağaçları üzerine kurulmuştur. Bu, verileri düzensiz, gürültülü veya düzensiz desenlere sahipse bile iyi bir şekilde işleyebilir. Rüzgar hızı tahmini gibi zaman serisi verileri bazen belirgin desenlere sahip değildir ve bu nedenle karar ağaçları bu tür verileri işlemekte iyidir. Bunun yanında, XGBoost, öznelik mühendisliğine dayalı olarak modele giriş özelliklerini düzenlemek ve seçmek için kullanılabilir. Bu durum, rüzgar hızı tahmini gibi görevlerde faydalı olabilir. UKSB ise genellikle daha fazla öznelik mühendisliği gerektirmez ve girdi verilerini daha doğrudan kullanır. Ayrıca, UKSB gibi derin öğrenme yöntemleri, genellikle büyük miktarlarda veri gerektirir. XGBoost, daha küçük veri kümeleriyle bile iyi sonuçlar verebilir. Rüzgar hızı tahmini gibi alanlarda daha sınırlı veri olduğu için, XGBoost'un daha iyi bir seçenektir.

Tablo 3: Elazığ ili Sivrice ilçesi için 3 farklı yöntemin test seti performans sonuçları

Yöntem	RMSE	MSE
XGBoost	0.1098	0.0121
RO	0.1011	0.0102
UKSB	0.1390	0.0193

Tablo 3'de, Elazığ ili Sivrice ilçesi için XGBoost, RO ve UKSB modellerin test seti için rüzgar tahmini performans sonuçları görülmektedir. UKSB modelinin eğitim setinde düşük RMSE (0.0495) ve MSE (0.0024) değerleri elde ettiği ve modelin eğitim verilerini düzgün bir şekilde öğrendiği görülmektedir. UKSB modelinin test setinde ise düşük RMSE ve MSE değerleri elde ettiği görülmüştür. Rastgele Orman ve XGBoost yöntemlerinin eğitim setinde sırasıyla RMSE (0.0013) ve MSE (1.77×10^{-6}), RMSE (0.0406) ve MSE (0.0017) değerlerini aldığı görülmüştür. Bununla birlikte, her iki yöntem de test setinde oldukça düşük RMSE ve MSE değerleri elde edilmiştir. Elde edilen sonuçlar incelendiğinde 3 yöntemin de birbirine yakın sonuçlar elde ettiği görülmektedir. Bu durum, her üç modelin de yüksek genelleme yeteneğine sahip olduğunu göstermektedir. Bunun yanında, Elazığ ili Sivrice ilçesi için RO yönteminin en iyi MSE ve RMSE değerleri elde ettiği görülmektedir.



Şekil 4: Elazığ ili Sivrice ilçesi için 1 yıllık, günlük çözünürlükte rüzgar hızı tahmin ve gerçek değerler

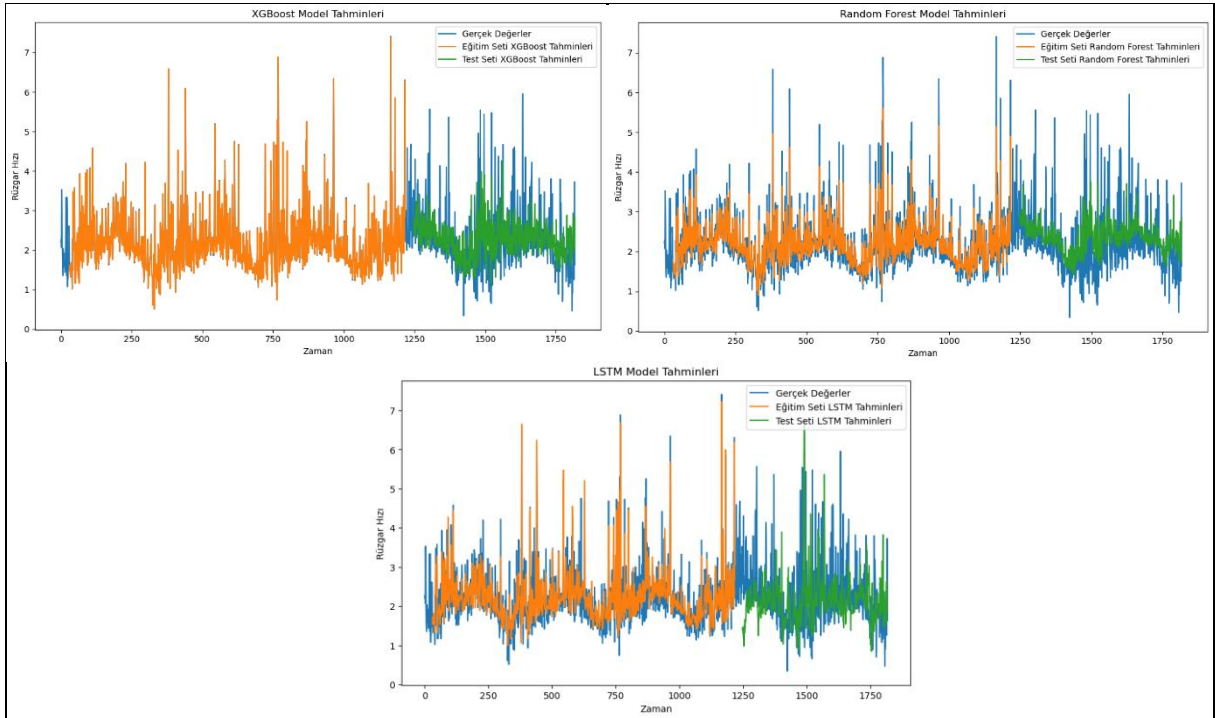
Şekil 4'de, Elazığ ili Sivrice ilçesi için XGBoost, RO ve UKSB modellerinin eğitim ve test setine ait tahmin başarımları görülmektedir. Modelin eğitim setinde tahminleri gerçek değerlere oldukça yakınken, test setinde tahminlerin gerçek değerlere göre daha fazla sapma gösterdiği görülmektedir. RO ve XGBoost yöntemleri diğer UKSB yöntemine göre rüzgar hızı verilerini daha iyi ifade ettiği ve tahmin yaptığı görülmektedir. RO ve

XGBoost yöntemleri benzer sonuçlar elde etmiştir. Bunun yanında UKSB modelinin hem eğitim hem test kümesinde veriyi iyi öğrenemediği ve dolayısıyla yüksek RMSE ve MSE sahip olduğu görülmektedir. XGBoost, düzensiz ve gürültülü verileri işleme konusunda daha iyidir. Rüzgar hızı tahmini gibi zaman serisi verileri bazen belirgin desenlere sahip olmayabilir ve gürültü içerebilir. Bu durumda, karar ağaçları, düzensiz verileri daha iyi işleyebilir. UKSB gibi derin öğrenme modelleri ise daha belirgin yapılar arar ve daha büyük ve düzenli veri kümeleri gerektirebilir.

Tablo 4: Elazığ ili Keban ilçesi için 3 farklı yöntemin test seti performans sonuçları

Yöntem	RMSE	MSE
XGBoost	0.1167	0.0136
RO	0.1105	0.0122
UKSB	0.1448	0.0211

Tablo 4'de, Elazığ ili Keban ilçesi için XGBoost, RO ve UKSB modellerin test seti için rüzgar tahmini performans sonuçları görülmektedir. XGBoost, Rastgele Orman ve UKSB yöntemlerinin eğitim setinde sırasıyla RMSE (0.0014) ve MSE (1.9×10^{-6}), RMSE (0.0384) ve MSE (0.0015) VE RMSE (0.0498) ve MSE (0.0025) değerlerini aldığı ve modellerin eğitim verilerini düzgün bir şekilde öğrendiği görülmektedir. Bununla birlikte, RO ve XGBoost yöntemleri de test setinde oldukça düşük RMSE ve MSE değerleri elde edilmiştir. Sonuç olarak her üç modelin de yüksek genelleme yeteneğine sahip olduğunu görülmektedir. Bunun yanında, Elazığ ili Sivrice ilçesi için RO yönteminin en iyi MSE ve RMSE değerleri elde ettiği görülmektedir.



Şekil 5: Elazığ ili Keban ilçesi için 5 yıllık, günlük çözünürlükte rüzgar hızı tahmin ve gerçek değerler

Şekil 5'de, Elazığ ili Keban ilçesi için XGBoost, RO ve UKSB modellerinin eğitim ve test setine ait tahmin başarımları görülmektedir. Modelin eğitim setinde tahminleri gerçek değerlere oldukça yakinken, test setinde tahminlerin gerçek değerlere göre daha fazla sapma gösterdiği görülmektedir. XGBoost ve RO yöntemlerinin UKSB yöntemine göre rüzgar hızı verilerini daha iyi ifade ettiği ve tahmin yaptığı görülmektedir. Bunun yanında UKSB modelinin hem eğitim hem test kümesinde veriyi iyi öğrenemediği ve dolayısıyla yüksek RMSE ve MSE sahip olduğu görülmektedir.

RO yöntemi UKSB'den daha iyi rüzgar hızı tahmini yapmıştır. Bu durumun nedeni aşağıdaki gibi açıklanabilir. Rastgele orman, bir ensemble öğrenme yöntemi olarak çalışır. Bu, birden fazla karar ağacının sonuçlarını birleştirerek daha güçlü bir tahmin yapma yeteneği sağlar. Zaman serisi verileri genellikle karmaşık yapılar ve gürültüler içerir, bu nedenle birçok ağacın birleştirilmesi, bu zorlukları aşmada yardımcı olabilir. Bunun yanında, UKSB gibi derin öğrenme modelleri, büyük ve karmaşık yapıları öğrenme yeteneğine sahiptir, ancak aynı zamanda veriyi ezberleme (overfitting) riski taşır. Rastgele orman, karar ağaçlarının sınırlı derinlikte olması

ve rastgele özellik seçimi gibi özellikler sayesinde ezberlemeyi azaltabilir. Sonuç olarak, hangi yöntemin rüzgar hızı tahmini gibi bir görev için daha iyi olduğu, veri seti, problem bağlamı ve performans ölçütlerine bağlı olarak değişebilir. Veri setinizi ve gereksinimlerinizi dikkate alarak farklı yöntemleri denemek ve karşılaştırmak en iyi sonuçları elde etmenize yardımcı olacaktır.

5. Sonuç ve Öneriler

Sonuç olarak, Elazığ Merkez için XGBoost modeli en iyi performansı gösterirken, Elazığ Sivrice'de XGBoost ve rastgele orman modelleri benzer düşük hatalara sahiptir; Elazığ Keban'da XGBoost ve rastgele orman modelleri benzer sonuçlar verirken, UKSB modeli daha yüksek hatalara işaret etmektedir; genel olarak daha fazla veri toplama, özellik mühendisliği, hiper parametre ayarı ve sürekli model güncellemeleri, tahmin modelinin performansını artırmak için önemli stratejiler olabilir.

Elazığ'daki farklı bölgeler için yapılan rüzgar hızı tahminlemeleri sonucunda, XGBoost modelinin genel olarak daha iyi performans gösterdiği, ancak bölgeye göre farklı modellerin daha iyi sonuçlar verebileceği anlaşılmaktadır. Bu nedenle, tahmin modelinin seçiminde bölgesel özelliklerin ve veri durumunun dikkate alınması önemlidir ve daha fazla veri toplama, özellik mühendisliği, hiper parametre ayarı ve sürekli model güncellemeleri gibi stratejiler tahmin doğruluğunu artırmak için gereklidir.

Kaynaklar

- Alkan, Ö., & Albayrak, Ö. K. (2020). Ranking of renewable energy sources for regions in Turkey by fuzzy entropy based fuzzy COPRAS and fuzzy MULTIMOORA. *Renewable Energy*, 162, 712–726. <https://doi.org/10.1016/j.renene.2020.08.062>
- Box, G. E. P. (1989). *An Unexpected Route to Time Series*.
- Breiman, L. (2001). Random Forests. *Kluwer Academic Publishers*, 45, 5–32.
- Brockwell, P. J., & Davis, R. A. (2002). Introduction to Time Series and Forecasting. In G. Casella, S. Fienberg, & I. Olkin (Eds.), *Springer Texts in Statistics* (Second Edn).
- Brownlee, J. (2023). *How to Develop a Random Forest Ensemble in Python*. <https://machinelearningmastery.com/random-forest-ensemble-in-python/>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-Aug, 785–794*. <https://doi.org/10.1145/2939672.2939785>
- Du, P. (2019). Ensemble Machine Learning-Based Wind Forecasting to Combine NWP Output with Data from Weather Station. *IEEE Transactions on Sustainable Energy*, 10(4), 2133–2141. <https://doi.org/10.1109/TSTE.2018.2880615>
- Ghoniem, A. F. (2011). Needs, resources and climate change: Clean and efficient conversion technologies. *Progress in Energy and Combustion Science*, 37(1), 15–51. <https://doi.org/10.1016/j.pecs.2010.02.006>
- Gielen, D., Boshell, F., Saygin, D., Bazilian, M. D., Wagner, N., & Gorini, R. (2019). The role of renewable energy in the global energy transformation. *Energy Strategy Reviews*, 24(January), 38–50. <https://doi.org/10.1016/j.esr.2019.01.006>
- Hamzacebi, C., & Es, H. A. (2014). Forecasting the annual electricity consumption of Turkey using an optimized grey model. *Energy*, 70, 165–171. <https://doi.org/10.1016/j.energy.2014.03.105>
- 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>
- Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, 2(3), 18–22.

- Liu, H., Mi, X., & Li, Y. (2018). Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network. *Energy Conversion and Management*, 166, 120–131. <https://doi.org/10.1016/J.ENCONMAN.2018.04.021>
- Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Ben Taieb, S., Bergmeir, C., Bessa, R. J., Bijak, J., Boylan, J. E., Browell, J., Carnevale, C., Castle, J. L., Cirillo, P., Clements, M. P., Cordeiro, C., Cyrino Oliveira, F. L., De Baets, S., Dokumentov, A., ... Ziel, F. (2022). Forecasting: theory and practice. *International Journal of Forecasting*, 38(3), 705–871. <https://doi.org/10.1016/j.ijforecast.2021.11.001>
- Prabha, P. P., Vanitha, V., & Resmi, R. (2019). Wind Speed Forecasting using Long Short Term Memory Networks. *2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies, ICICICT 2019*, 1310–1314. <https://doi.org/10.1109/ICICICT46008.2019.8993279>
- Yan, Y., Wang, X., Ren, F., Shao, Z., & Tian, C. (2022). Wind speed prediction using a hybrid model of EEMD and UKSB considering seasonal features. *Energy Reports*, 8, 8965–8980. <https://doi.org/10.1016/j.egyr.2022.07.007>