



Covid-19 Salgın Sürecinde Hava Kalitesi Tahmini: Zonguldak Örneği

Büşra Duygu ÇELİK^{*,a} , Nursal ARICI^a 

^{a,*} Gazi Üniversitesi Uygulamalı Bilimler Fakültesi Yönetim Bilişim Sistemleri Bölümü, 06010, Ankara, Türkiye.

MAKALE BİLGİSİ

Alınma: 07.06.2021
Kabul: 13.08.2021

Anahtar Kelimeler:
Covid-19,hava
kirliliği,hava kalitesi
tahmini,makine
öğrenmesi

ÖZ

Hava kirleticilerinin insan sağlığı üzerinde basit etkilerden erken ölüme kadar ciddi etkilere neden olduğu bilinmektedir. Özellikle solunum problemleri, akciğer rahatsızlıkları ve zatürre gibi birçok rahatsızlığa sebep olmaktadır. 30 Aralık 2019 itibarıyla dünyayı etkisi altına alan Covid-19 salgını bir solunum yolu hastalığıdır ve hava yoluyla bulaşmaktadır. Günümüzde devam eden salgın sürecinde hava kalitesinin tahmin edilmesi ve buna yönelik tedbirlerin alınması Covid-19 gibi hava yoluyla bulaşan hastalıkların yayılma hızını etkilemesi açısından önem taşımaktadır. Çalışma, Covid-19 salgını öncesi ve Covid-19 salgın sürecinde ölçülen çeşitli kirletici konsantrasyonlarını ele alarak makine öğrenmesi yöntemleriyle hava kalitesini tahmin etmektedir. Çalışmada kullanılan veri seti, hava kirliliğinin yüksek olduğu, gelişmiş sanayiye sahip Zonguldak iline ait kirletici konsantrasyonlarından oluşmaktadır. Veriler, Çevre ve Şehircilik Bakanlığı (ÇSB) Hava İzleme İstasyonları web sitesinden sağlanmıştır. Tahmin başarısı yüksek beş farklı makine öğrenmesi yöntemi kullanılmıştır. Çalışma sonucunda en iyi başarı 0.016 (2019 veri seti) ve 0.021 (2020 veri seti) RMSE değerleri ile Karar Ağacı algoritmasında elde edilmiştir. Çalışmada Naive Bayes algoritması en düşük başarıya sahiptir. Deneysel sonuçlar, önerilen modelin hava kalitesini tespit etmek için verimli bir şekilde kullanılabileceğini göstermektedir.

<https://dx.doi.org/10.30855/gmbd.2021.03.05>

Air Quality Forecast in the Covid-19 Outbreak Process: Zonguldak Case

ARTICLE INFO

Received: 07.06.2021
Accepted: 13.08.2021

Keywords:
Covid-19,air
pollution,air quality
prediction,machine
learning

ABSTRACT

Air pollutants are known to cause severe effects on human health, from simple effects to premature death. It causes many ailments, especially respiratory problems, lung diseases, and pneumonia. As of December 30, 2019, the covid-19 outbreak affecting the world is a respiratory disease and is transmitted by air. In the current epidemic process, it is essential to predict air quality and take measures to affect the rate of spread of air-borne diseases such as Covid-19. The study predicts air quality through machine learning methods, considering the various concentrations of pollutants measured before the covid-19 outbreak and during the covid-19 outbreak. The data set used in the study consists of the pollutant concentrations of Zonguldak province, which has high air pollution and developed industry. The data was obtained from the Ministry of Environment and Urban Planning (MoLS) weather monitoring stations website. Five different machine learning methods with high predictive success were used. As a result of the study, the best hit was achieved in the decision tree algorithm with Rmse values of 0.016 (2019 dataset) and 0.021 (2020 dataset). The Naive Bayesian algorithm has the lowest success in the study. Experimental results suggest that the proposed model could be used efficiently to detect air quality.

<https://dx.doi.org/10.30855/gmbd.2021.03.05>

*Sorumlu yazar: nursal@gazi.edu.tr

To cite this article: B.D Çelik ve N. Arıcı, "Air Quality Forecast in the Covid-19 Outbreak Process: Zonguldak Case", *Gazi Journal of Engineering Sciences*, vol.7, no.3, pp. 222-232, 2021. doi:10.30855/gmbd.2021.03.05

1.GİRİŞ (INTRODUCTION)

Bu Hava kirliliği; atmosferdeki gaz, su buharı, duman, toz şeklindeki kirleticilerin, bütün canlılara zarar verecek seviyeye erişmesidir [1]. Özellikle sanayinin gelişmiş olduğu bölgelerde, hava kirliliği yaşam kalitesi için önemli bir etkidir. Hızla büyüyerek gelişen sanayi, çoğalan nüfus ve bunların sonucunda artan kentleşme hava kirliliğini arttırmıştır. Havada bulunan kirleticiler astım, zatürre, kalp krizi, bronşit ve diğer solunum problemleri gibi kronik hastalıkları tetikler veya kötüleştirir [2].

2019 yılının aralık ayında Çin'in Wuhan şehrinde ilk defa görülen Covid-19 vakasından sonra dünyaya yayılan salgının Türkiye'deki ilk vakası 11 Mart'ta tespit edilmiştir. Covid-19 tespit edilen vakada en sık solunum problemleri, akciğerlerde yoğun hasar ve hastalığın ilerlemesi durumunda zatürre görülmektedir. Ayrıca kronik rahatsızlığı olan vakalarda kalp krizi ve ölüm riski daha da artmaktadır. Bilim insanları hava kalitesinin düşük olduğu bölgelerde yaşayan insanlarda Covid-19 vakalarının hastalığı daha şiddetli geçirip geçirmeyeceğini araştırmaya başlamışlardır [3]. Hava kirliliği yoğun olan bölgelerde yaşayanların diğer insanlara göre hastalanmaya daha müsait olduğunu ve kirletici partiküllerin virüsün bulaşmasında kolaylaştırıcı etkiye sahip olduğunu belirtmişlerdir [4]. Ayrıca, hava kirliliğinin yüksek olduğu şehirlerde Covid-19 vakalarının olması durumunda alınan önlemlerin artırılması gerektiği, aksi takdirde yüksek ölüm oranları ile karşılaşılacağı bildirilmiştir [3]. Ülkemizde alınan pandemi tedbirleri kapsamında 4 Nisan 2020'den itibaren otuz büyük şehir ve Zonguldak'a seyahat ve sokağa çıkma kısıtlamasına başlanmıştır. Zonguldak büyük şehir olmamakla birlikte kısıtlama kapsamına alınmasının nedeni hava kirliliğinin en yoğun olduğu bölgelerden biri olup vaka sayısı bakımından da riskli grubunda yer almasıdır. Bu gerekçeyle çalışmamızda Zonguldak ili hava kalitesi değerlendirilmesi ele alınmıştır.

Bu çalışmanın amacı, Zonguldak ilindeki Covid-19 salgını öncesi ve esnasındaki hava kalitesi değişimlerini incelemek ve bu amaç doğrultusunda kirletici parametreleri makine öğrenmesi yöntemleriyle tahmin etmektir. Zonguldak ilinde gözlemlenen PM₁₀, SO₂, NO₂, O₃ ve CO hava kirleticilerinin konsantrasyonları alınarak [5] salgın öncesi ve salgın süreci olmak üzere iki ayrı veri seti oluşturulmuştur. Bu veri setleri üzerinde tahmin başarısı yüksek olan Destek Vektör Makinesi (SVM), K-En Yakın Komşu (KNN), Karar Ağacı (DT), Rastgele Orman (RF), ve Naive Bayes (NB) makine

öğrenmesi algoritmaları uygulanmıştır. Çalışmada kullanılan verileri bölme işleminde öncelikle %80 eğitim ve %20 test, ardından %70 eğitim ve %30 test verisi alınarak model ayrı ayrı denenmiştir. En son aşırı öğrenmeyi engellemek ve daha yüksek başarı için %67'si eğitim ve %33'ü test verisi olarak kullanılmıştır. Her bir yöntemin tahmin başarı performansı doğruluk, RMSE ve R² kriterlerine göre kıyaslanmıştır. Çalışmanın sonraki bölümleri şu şekilde yapılandırılmıştır.

İkinci bölümünde, literatürdeki Covid-19 salgını ile hava kirliliği arasındaki ilişkiyi ve hava kalitesi tahmini için kullanılan makine öğrenmesi yöntemlerini konu alan çalışmalar incelenmiştir. Üçüncü bölümde çalışmada kullanılan veri seti ile tahminleme modelinde kullanılan makine öğrenmesi yöntemleri hakkında bilgi verilmiştir. Dördüncü bölümde Covid-19 öncesi ve Covid-19 sürecindeki veriler değerlendirilmiş ve uygulanan modelin doğruluk değerleri karşılaştırılmıştır. Son bölümde ise, tahminleme sonucu elde edilen çıktılar yorumlanmıştır.

2.ÖNCEKİ ÇALIŞMALAR (PREVIOUS STUDIES)

İnsan sağlığını ciddi anlamda tehdit eden ve birçok hükümeti önlem alamaya iten Covid-19 salgını, bulaştığı kişilerde bağışıklık sisteminin zayıf olması durumunda kişi, hastalığı daha ağır atatabilmektedir. Hava kalitesinin düşük olması kişilerde bağışıklık sisteminin zayıflatacak birçok hastalığa neden olabilmektedir. Bu yüzden birçok araştırmacı, hava kalitesini tahmin edebilmek amacıyla yaptıkları çalışmaların yanında Covid-19'un hava kalitesiyle ilişkisini de incelemektedirler.

Zhu ve ark., Çin'de bulunan 120 şehirdeki PM₁₀, PM_{2.5}, SO₂, CO, NO₂ ve O₃ hava kirletici parametreleri ile günlük tespit edilen Covid-19 vakalarını değerlendirmeye almış ve bu veriler üzerinde genelleştirilmiş katkı modeli uygulayarak Covid-19 ile hava kirliliği arasındaki ilişkiyi araştırmışlardır. Bu çalışmanın sonucunda hava kirliliği ve Covid-19 salgını arasında anlamlı pozitif güçlü bir ilişki olduğu saptanmıştır [6].

Rodríguez-Urrego ve Rodríguez-Urrego, karantina başlangıcından önce ve sonra ölçülen, Dünya Sağlık Örgütü'ne göre dünyanın en kirlili 50 başkentindeki PM_{2.5} partikül maddesinin emisyon davranışını ele almış ve Covid-19 sürecinde hava kalitesi incelemiştir. Çalışmanın sonucuna göre karantina boyunca çoğu şehrin kirletici konsantrasyonlarında ciddi bir düşüş olduğu görülmüştür. Ayrıca seçilen

şehirlerden çoğunun hava kalitesi olumlu etkilenmiştir [7].

Başka bir çalışmada Brandt ve arkadaşları, Covid-19 ölümleri üzerinde hava kirliliğinin etkisini incelemişler ve nüfus yoğunluğunun yüksek olduğu yerleşim yerlerinde hava kirliliğinin daha yoğun olduğunu ve bu yerlerdeki hava kirliliğinin Covid-19 ölümlerinin üzerinde etkili olduğunu vurgulamıştır [8].

Hava kalitesinin değişimini izleme, tahmin etme ve gerekli tedbirlerin erken alınması konuları Covid-19 salgın sürecinde daha da önem taşımaktadır. Özellikle hava kalitesinin doğru tahmin edilmesi, ekolojik denge ve insan sağlığı üzerinde önemli bir rol oynamaktadır. Son yıllarda araştırmacılar klasik istatistik yöntemlerin dışında tahmin başarısı yüksek olan makine öğrenmesi yöntemleri ile hava kalitesini tahmin etmeye çalışmışlardır.

Doreswamy ve arkadaşları, havadaki PM₁₀ ve PM_{2.5} konsantrasyonunu tahmin etmeye yönelik makine öğrenmesi tahmin modellerini kullanmış ve bu modelleri mevcut geleneksel modellerle karşılaştırılmıştır. Elde ettikleri 0,9336 R² ve 0.1302 RMSE performans değerleriyle, makine öğrenmesi modelleri tahmin performansında geleneksel modellere göre daha iyi performans göstermiştir [9].

Başka bir çalışmada Kang ve arkadaşları, Çin'in Shenzhen şehrinden alınan veriler kullanılarak hava kalitesi haritası görselleştirilmiştir. Ayrıca Yapay Sinir Ağı, Genetik Algoritma, Rastgele Orman, Karar Ağacı, Derin İnanç Ağı (Deep Belief Network) ve Destek Vektör Makinesi kullanılan çalışmalarda oluşturulan modellerin çeşitli avantajları ve dezavantajları anlatılmıştır. Karşılaştırılan çalışmalardan en başarılı olan Karar Ağacı Algoritması, %89,46 doğruluk oranına sahiptir [10].

Masood ve arkadaşları, 2016-2018 arasındaki iki yıllık döneme karşılık gelen çeşitli meteorolojik ve kirletici parametrelerin girdileri üzerine SVM ve YSA gibi farklı modeller oluşturulmuştur. PM_{2.5} tahmini için modellerin performans değerlendirmesi yapılmış ve sonuçlar tartışılmıştır. YSA ve SVM için ortalama hata karesi değerleri sırasıyla 0.0191 ve 0.0314 bulunmuştur [11].

Deters ve arkadaşları, araştırma makalelerinde ince partikül madde veya PM_{2.5} nedeniyle milyonlarca erken ölüme neden olan dış mekân kirliliği araştırmasını önermişlerdir. PM_{2.5} konsantrasyonunu tahmin etmek için altı yıllık meteorolojik ve kirlilik

veri analizine dayalı bir makine öğrenmesi yaklaşımı geliştirmişlerdir [12].

Bali çalışmasında, Hindistan hava kalitesinin tahmin ve analizini makine öğrenmesi algoritmalarından SVM, RF, NB kullanarak modellemiştir. En yüksek başarıyı %99,1 ile rastgele orman algoritmasında elde etmiştir [13].

3. MATERYAL VE YÖNTEM (MATERIAL AND METHOD)

3.1. Zonguldak İli Hava Kalitesi (Air Quality of Zonguldak City)

Son yıllarda hava kalitesi oldukça kötüleşen illerimizden biri de Zonguldak'tır [5]. Türkiye Çevre Sorunları ve Öncelikleri Envanteri Değerlendirme Raporu'nda Zonguldak ilinin en önemli çevre sorununun hava kirliliği olduğu belirtilmiştir. Ayrıca Zonguldak, Türkiye İstatistik Kurumu'nun Çevre İstatistiklerine bakıldığında partikül madde kirliliği bakımından Türkiye'nin havası en kirli şehirleri arasındadır [14]. Zonguldak'ta hava kirliliğine neden olan başlıca kirleticiler genellikle sanayileşme, soğuk havadaki evsel ısınma ve trafik kaynaklıdır. Özellikle sanayileşmenin oldukça yoğun olduğu Batı Karadeniz bölgesinde bulunan ilimiz, taş kömürü madenciliği, enerji ve metalürji sektörü sanayisinin büyük bir bölümünü oluşturmaktadır. Enerji sektöründe yer alan Çatalağzı Termik Santrali ülkemizdeki ilk termik santral olup katı yakıt olarak taş kömürü kullanmaktadır [5].

Zonguldak'ta görülen hava kirleticiler partiküller ve gazlar şeklinde iki alt gruba ayrılır. Partikül maddelerden PM₁₀, çoğunlukla fabrika faaliyetleri, inşaatlar, trafik, termik santraller ve evsel ısınma kaynaklıdır. Hava yoluyla taşınan bu kirletici türü öncelikle astım gibi solunum rahatsızlıkları, kalp damar hastalıkları, akciğer rahatsızlıklarına sebep olarak erken ölüm riski taşımaktadır. PM₁₀'nun yanı sıra izleme istasyonunda ölçülen hava kirletici parametreleri kısaca açıklanmıştır [15]:

Renksiz, yanıcı olmayan ve gaz halde bulunan kükürt dioksit (SO₂), genellikle ısınma, endüstriyel faaliyetler ve trafik bölgelerinde oluşan bir kirletici türüdür. Diğer kirleticilerle birleştiğinde damlalar veya katı partiküller halinde uzun mesafeler taşınabilir. Uzun süre maruz kalınması halinde solunum yolu rahatsızlıklarına, kalp krizi riskinin artmasına neden olmaktadır [15].

Azot oksit (NO_x) emisyonları insanların sağlığı üzerinde yarattığı etki bakımından en zararlı kirletici

olup trafik ve endüstriyel tesisler başlıca kaynağıdır. Kısa süre maruz kalındığında bile solunum sistemi üzerinde ağır hasarlara yol açan NO_x emisyonları kronik hastalarda daha da büyük tahribatlara sebep olmaktadır. Uzun süreli maruz kalınması halinde ise bölgedeki solunum yolu rahatsızlıklarının arttığı gözlemlenmiştir [15].

Karbon monoksit (CO) gazı tam yanmamış olan yakıtlar sonucu ortaya çıkan kokusuz ve renksiz bir gazdır. Kış mevsiminde en yüksek değere ulaşan CO' in ana kaynağı trafikte oluşan egzoz gazları ve sigara dumanıdır. Kandaki hemoglobine bağlanarak hücrelere giden oksijen miktarının azalmasına sebep olur. CO gazına maruz kalan insanlarda kalp ve solunum rahatsızlıkları görülebilir [15].

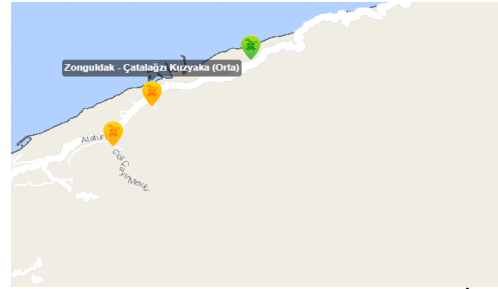
Ozon (O₃), doğrudan ortam havasına karışmayan, renksiz ve kokusuz bir gaz olup karmaşık kimyasal tepkimeler sonucunda oluşur. Maruz kalma süresine göre tahribat oranı artar ve en yüksek risk grubu çocuklar ve kronik hastalardır [15].

3.2. Veri Seti (Dataset)

Veri seti, Çevre ve Şehircilik Bakanlığı (ÇŞB) Hava İzleme İstasyonları web sitesinden sağlanmıştır [16]. Bu web sitesinde Türkiye'deki tüm illerin çeşitli bölgelerinden 7x24 ölçülen SO₂, NO₂, NO_x, CO, O₃, PM₁₀ ve PM_{2,5} kirletici parametrelere ait veri setleri yayımlanmaktadır. Bu çalışmada veriler, Zonguldak ili Çatalağzı beldesinde yer alan hava izleme istasyonundan alınmıştır (Şekil 1).

Termik santralin yer aldığı Çatalağzı-Kuzyaka izleme istasyonundan, salgın öncesi 01.01.2019-31.12.2019 ve salgın sürecine ait 11.03.2020-

31.12.2020 tarih aralığına ait yapılan ölçüm verileri alınmıştır. Çalışmada kullanılan veri seti, 15.578 veriden oluşmaktadır. 8.737 veri 2019 yılına, 6841 veri de 2020 yılına aittir. Türkiye'de resmi kayıtlarda salgın başlangıcı mart ayı kabul edildiğinden 2020 yılına ait ocak ve şubat ayları değerlendirmeye dahil edilmemiştir. Seçilen istasyonda SO₂, NO₂, O₃, CO ve PM₁₀ kirletici parametreleri ölçülmektedir.



Şekil 1. Zonguldak-Çatalağzı Kuzyaka Hava İzleme İstasyonu (Zonguldak-Çatalağzı Kuzyaka Weather Monitoring Station)

Çalışmada tahmin edilmek istenen bağımlı değişken Hava Kalite İndeksi (HKİ), çevremizdeki hava kalitesinin sağlığımız üzerindeki etkisinin ne olduğunu anlamamıza yarayan ölçektir. Bu indeks, havadaki kirleticilerin konsantrasyonlarına göre hava kalitesini iyi, orta, hassas, sağlıklı, kötü ve tehlikeli olmak üzere beş farklı seviyede ortaya koymaktadır. Ulusal Hava Kalite İndeksi kesme noktaları Tablo 1'de verilmiştir. Ayrıca hava kirliliğine karşı hassas olan insanların sağlıklarını korumaya yardımcı olan bir faktördür. HKİ, çevresel faktörlerin insan sağlığına hangi oranda etki yaptığını göstererek, sağlık açısından tehdit oluşturan durumları önlemek için kullanılmaktadır.

Tablo 1. Ulusal hava kalite indeksi kesme noktaları (National air quality index breakpoints) [5]

İndeks	HKİ	SO ₂ [µg/m ³]	NO ₂ [µg/m ³]	CO[µg/m ³]	O ₃ [µg/m ³]	PM ₁₀ [µg/m ³]
		1 Sa. Ort.	1 Sa. Ort.	8 Sa. Ort.	8 Sa. Ort.	24 Sa. Ort.
İyi	0- 50	0-100	0-100	0-5500	0-120	0-50
Orta	51- 100	101-250	101-200	5501-10000	121-160	51-100
Hassas	101- 150	251-500	201-500	10001-16000	161-180	101-260
Sağlıksız	151- 200	501-850	501-1000	16001-24000	181-240	261-400
Kötü	201 - 300	851-1100	1001-2000	24001-32000	241-700	401-520
Tehlikeli	301- 500	>1101	>2001	>32001	>701	>521

Çalışmamızda hava kirliliği parametreleri SO₂, NO₂, CO, O₃, PM₁₀ (µg/m³) giriş parametreleri olarak ele alınmıştır ve bağımlı değişken olan HKİ değeri tahmin edilmiştir. Veri ön işleme aşamasında, öncelikle aykırı değerler tespit edilerek yerine ortalama değer atanmıştır. Ardından eksik veriler tespit edilmiştir. 2020 yılı temmuz ayına ait

ölçümlerde veri bulunmadığından temmuz ayı değerlendirmeye dahil edilmemiştir. Eksik veri olan diğer aylarda kirleticilerin ortalama değerleri kullanılarak tamamlanmıştır. Bir sonraki aşamada verilerin ortak bir ölçekte değerlendirilebilmesi için (1) nolu eşitlikteki min-max normalizasyonu kullanılmıştır.

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

x' = Normalize edilmiş veri

x_i = Girdi değeri

x_{min} = Girdi seti içerisinde yer alan en küçük sayı

x_{max} = Girdi seti içerisinde yer alan en büyük sayı

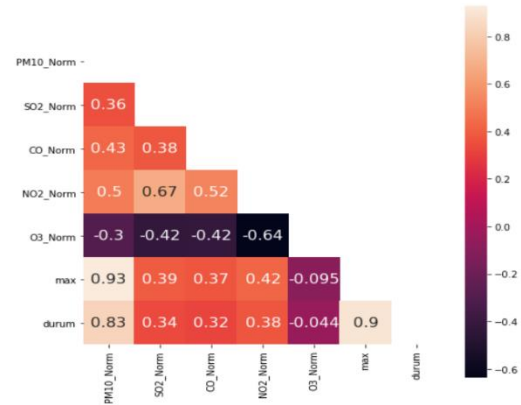
Normalize edilen veriler günlük olarak kıyaslandığında en yüksek ölçülen değer 'max' değeri ile belirtilmiştir. 'max' değeri ölçüm yapılan günün HKİ' sine eşittir. HKİ değeri için 1'den 6'ya kadar 'Durum' etiketi oluşturulmuştur. İlgili etiketlerin değer aralıkları tablo 2'de verilmiştir.

Tablo 2. HKİ aralığına göre durum etiketi dağılımı
(Status label distribution by AQI range)

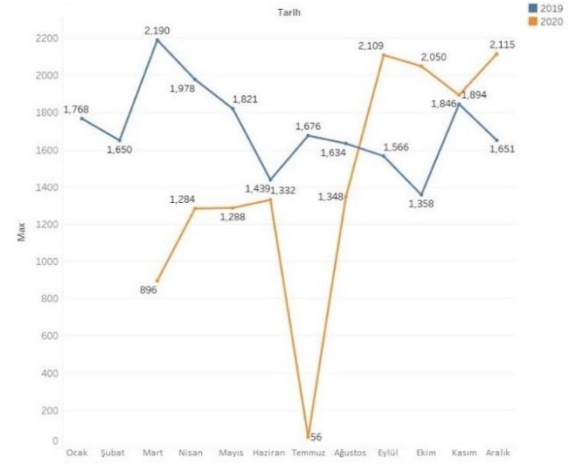
İndeks	HKİ	Durum
İyi	0 – 50	1
Orta	51 – 100	2
Hassas	101 – 150	3
Sağlıksız	151 – 200	4
Kötü	201 – 300	5
Tehlikeli	301 – 500	6

Daha sonra normalize edilmiş parametreler kullanılarak korelasyon matrisi oluşturulmuş ve parametreler arasındaki ilişki incelenmiştir. Şekil 2'de havayı kirleten ve veri setinde yer alan parametreler ile hava kalitesi arasındaki ilişki güçleri [0,0, 1,0] aralığında gösterilmiştir. Şekil incelendiğinde PM₁₀ parametresinin hava kalitesi indeksi ile 0,93'lük pozitif güçlü bir korelasyona sahip olduğu görülmektedir. Hava Kalite İndeksi ile PM₁₀ haricindeki diğer parametreler arasındaki ilişki güçleri sırasıyla NO₂, SO₂, CO ve O₃'a aittir.

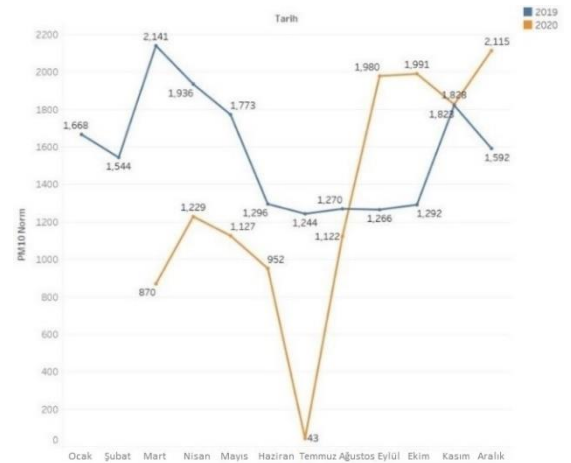
2019-2020 günlük ölçümlenen kirletici konsantrasyonlarından oluşan veri seti incelendiğinde PM₁₀ ve Hava Kalite İndeksi arasındaki ilişki Şekil 3 ve Şekil 4'teki gibidir. Grafiklerde görüldüğü üzere PM₁₀ artış azalış eğrisi hava kalite eğrisi ile birçok noktada paralellik göstermektedir. Bundan dolayı PM₁₀ parametresiyle hava kalitesi arasında pozitif güçlü bir ilişki olduğu çıkarımı yapılabilir.



Şekil 2. Hava Kirleticilerin Korelasyon Matrisi
(Correlation Matrix of Air Pollutants)



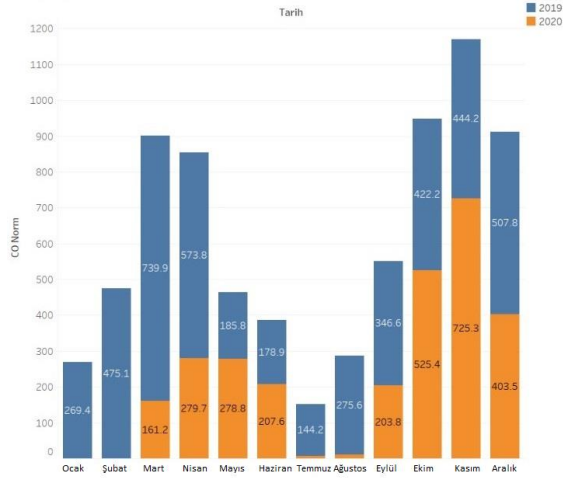
Şekil 3. HKİ artış azalış eğrisi
(AQI increase and decrease curve)



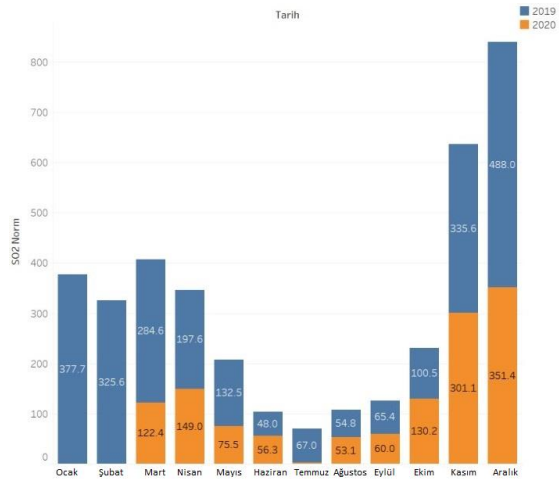
Şekil 4. PM₁₀ artış azalış eğrisi
(PM₁₀ increase and decrease curve)

Covid-19 salgını öncesi ve süresince ölçülen kirletici konsantrasyonlarının aylık dağılımı Şekil 5 ve Şekil 6'da görselleştirilip kıyaslanmıştır. Çalışmada

2020 yılına ait salgın süreci ele alındığından Ocak ve Şubat ayları değerlendirmeye dahil edilmemiştir.



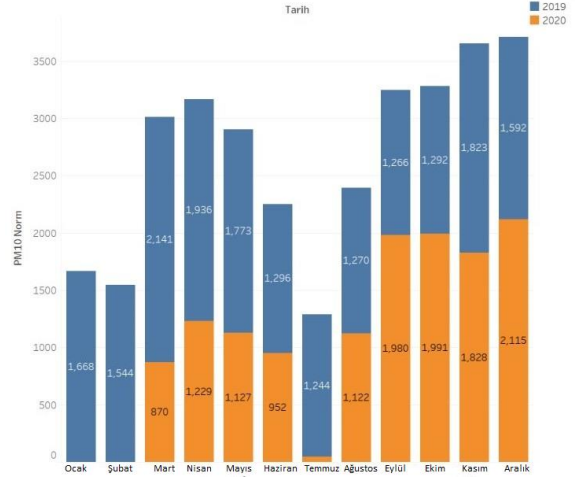
Şekil 5. CO Parametresinin Yıl Bazlı Dağılımı
(Year-Based Distribution of CO Parameter)



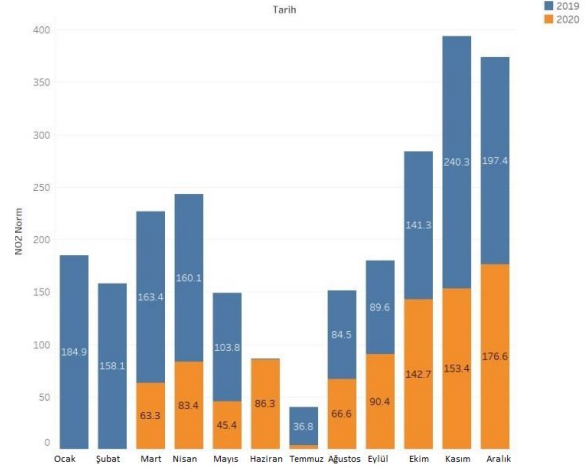
Şekil 6. SO₂ Parametresinin Yıl Bazlı Dağılımı
(Year-Based Distribution of SO₂ Parameter)

Şekil 5 ve Şekil 6'daki grafikler incelendiğinde, Mart ayında başlayan karantina ile CO ve SO₂ kirleticilerinin ölçümlenen değerlerinde ciddi bir düşüş görülmektedir ve bu düşüşler DSÖ'nün çalışmalarıyla da desteklenmektedir [17][18].

Şekil 7 ve Şekil 8'de PM₁₀ ve NO₂ hava kirletici parametrelerinin yıllara göre dağılımları görselleştirilmiştir. Bu grafiklerde de CO ve SO₂ parametrelerinde görüldüğü gibi 2020 yılında ciddi düşüşler görülmektedir.

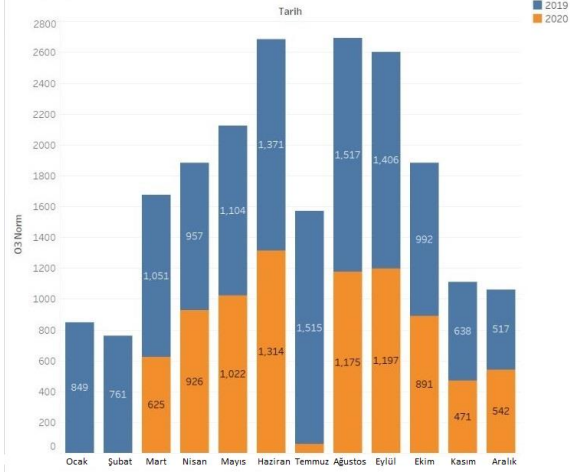


Şekil 7. PM₁₀ Parametresinin Yıl Bazlı Dağılımı
(Year-Based Distribution of PM₁₀ Parameter)



Şekil 8. NO₂ Parametresinin Yıl Bazlı Dağılımı
(Year-Based Distribution of NO₂ Parameter)

Son olarak Şekil 9'da O₃ hava kirletici parametresi yıllara göre dağılımı görselleştirilmiştir. Şekilde görüldüğü üzere O₃ hava kirletici parametresinde de düşüş olduğu gözlemlenmiş fakat düşüşün önceki dört parametrede olduğu kadar kesin olmadığı görülmektedir. Karantina sürecinde genel olarak tüm kirletici parametrelerde azalma tespit edilmiştir. Ayrıca haziran ayında tekrar normalleşme sürecine girilmesi ile kirletici konsantrasyonlarının arttığı görülmektedir.



Şekil 9. O₃ Parametresinin Yıl Bazlı Dağılımı
(Year-Based Distribution of O₃ Parameter)

Modelin kurulması ve değerlendirilmesi aşamasında Python temelli bir kütüphane olan Scikit-learn kütüphanesi kullanılmıştır. Makine öğrenmesi yöntemlerinden KNN, SVM, RF, KA ve NB ile hava kalitesi modellenmiş ve başarı oranları karşılaştırılmıştır.

3.3. Makine Öğrenmesi Algoritmaları (Machine Learning Algorithms)

Günümüzde gelişen teknolojiyle beraber verinin elle işlenmesi ve analizinin yapılması daha zor hale gelmiştir. Yapısal olarak öğrenebilen yapay zekâ algoritmalarının bir alt alanı olan makine öğrenmesinin temel amacı veriler üzerinde başarılı tahminler elde etmektir. Makine öğrenmesi algoritmalarıyla hava kalitesi tahmin sistemi için kullanılan algoritmanın seçimi çok önemlidir. Yapılan sistematik inceleme, 2017'den bu yana, atmosfer kirliliğini tahmin etmek için makine öğrenmesi tekniklerinin uygulanmasının önemli ölçüde arttığını ortaya koymaktadır [19]. Çalışmada, başarı oranlarının yüksek olması sebebiyle kullanılan makine öğrenmesi yöntemleri aşağıda açıklanmıştır. Ayrıca her yonteme ait 2019 ve 2020 yılları hata matrisleri ve doğruluk oranları da ek olarak verilmiştir.

3.3.1. Destek vektör makinesi (Support Vector Machine-SVM)

İki ve daha fazla boyuttaki verileri doğru, düzlem veya hiper düzlem yardımıyla birbirinden ayırarak sınıflandırma ve regresyon analizi yapan makine öğrenmesi algoritmasıdır. Bu algoritmanın amacı verilen veri noktalarını olabildiğince en iyi şekilde ayırmaktır. Küçük ve orta ölçekli verilerde daha başarılı olduğu gözlemlenen SVM algoritması, hafızayı verimli kılan karar fonksiyonunda eğitim noktalarının alt kümesini kullanır [20]. 2019 ve 2020

yıllarına ait veri setlerine uygulanan SVM algoritması sonucunda elde edilen hata matrisleri Şekil 10 ve Şekil 11' de verilmiştir.

```
[[1202  13  0  0  0  0]
 [  6 1432  3  0  0  0]
 [  0  1 189  0  0  0]
 [  0  0  8 19  2  0]
 [  0  0  0  0  8  0]
 [  0  0  0  0  0  1]]
```

Support-Vector-Machine Accuracy 0.9885575589459085
Support-Vector-Machine Rmse: 0.1069693463291776
Support-Vector-Machine R2: 0.9746791055795522

Şekil 10. 2019 hata matrisi (confusion matrix-2019)

```
[[1286  0  0  0  0  0]
 [  23 802  2  0  0  0]
 [  0  4 106  0  0  0]
 [  0  0  3 24  1  0]
 [  0  0  0  1  5  0]
 [  0  0  0  0  0  1]]
```

Support-Vector-Machine Accuracy 0.9849424269264836
Support-Vector-Machine Rmse: 0.12270930312537998
Support-Vector-Machine R2: 0.9674352709627065

Şekil 11. 2020 hata matrisi (confusion matrix-2020)

3.3.2. Rastgele orman (Random forest -RF)

Hem sınıflandırma hem de regresyon amaçlı kullanılabilen denetimli bir makine öğrenme algoritmasıdır. Genellikle sınıflandırma görevleri için tercih edilir. Birden çok karar ağacından oluşan Rastgele orman algoritması veri ögeleri üzerinde rastgele bir orman oluşturur ve her biri için tahmin yapar. En son olarak da aralarındaki en iyi çözümü seçer. Bu sayede Rastgele orman, sonuçların ortalamasını alarak aşırı öğrenmeyi azaltır ve birbirinden bağımsız parametrelerle farklı karar ağaçları üzerinden değerlendirmeye aldığı için başarısı oldukça yüksektir [20]. 2019 ve 2020 yıllarına ait veri setlerine uygulanan RF algoritması sonucunda elde edilen hata matrisleri Şekil 12 ve Şekil 13' te verilmiştir.

```
[[1215  0  0  0  0  0]
 [  0 1440  1  0  0  0]
 [  0  0 190  0  0  0]
 [  0  0  5 24  0  0]
 [  0  0  1  0  7  0]
 [  0  0  0  0  0  1]]
```

Random-Forest Accuracy: 0.9975728155339806
Random-Forest Rmse: 0.0588846871438385
Random-Forest R2: 0.9923270016907734

Şekil 12. 2019 hata matrisi (confusion matrix-2019)

```
[[1286  0  0  0  0  0]
 [  1 825  1  0  0  0]
 [  0  0 110  0  0  0]
 [  0  0  0 28  0  0]
 [  0  0  0  1  5  0]
 [  0  0  0  0  1  0]]
Random-Forest Accuracy: 0.9982285208148804
Random-Forest Rmse: 0.04208894373965181
Random-Forest R2: 0.9961688554073772
```

Şekil 13. 2020 hata matrisi (*confusion matrix-2020*)

3.3.3. K en yakın komşu (*K-nearest neighbour*)

K En Yakın Komşu algoritması, sınıflandırma ve regresyon problemlerinin çözümünde kullanılan basit, oldukça yaygın ve denetlenen bir makine öğrenmesi algoritmasıdır. KNN algoritmasında ilk aşamada veriler etiketlenir ve etiketlenen bu verilerden bir eğitim kümesi hazırlanır. Ardından k adet sınıf merkezi tanımlandıktan sonra bir uzaklık fonksiyonu belirlenir. Daha önce karşılaşılmamış bir veri ile yüz yüze gelindiğinde belirlenen uzaklık algoritması kullanılarak karşılaşılan verinin eğitim kümesindeki verilerle olan mesafesi ölçülür ve aralarında uzaklığı en kısa olan k adet veri eğitim kümesinden seçilerek sınıflama kümesi oluşturulur. En son aşamada, seçilen yeni verinin sınıfı, sınıflama kümesinin en sık içerdiği sınıf olarak tanımlanır ve oluşturulan model sona erdirilir [21]. 2019 ve 2020 yıllarına ait veri setlerine uygulanan KNN algoritması sonucunda elde edilen hata matrisleri Şekil 14 ve Şekil 15' te verilmiştir.

```
[8737 rows x 6 columns]
[[1191  24  0  0  0  0]
 [ 19 1418  4  0  0  0]
 [  0  8 182  0  0  0]
 [  0  0  8 21  0  0]
 [  0  0  1  2  5  0]
 [  0  0  0  0  1  0]]
KNN Accuracy: 0.9767683772538142
KNN Rmse: 0.15579423821243896
KNN R2: 0.9462890118354137
```

Şekil 14. 2019 hata matrisi (*confusion matrix-2019*)

```
[6841 rows x 6 columns]
[[1282  4  0  0  0  0]
 [ 20 804  3  0  0  0]
 [  0  2 108  0  0  0]
 [  0  0  3 23  2  0]
 [  0  0  0  0  6  0]
 [  0  0  0  0  1  0]]
KNN Accuracy: 0.9844995571302038
KNN Rmse: 0.1245007745750856
KNN R2: 0.9664774848145509
```

Şekil 15. 2020 hata matrisi (*confusion matrix-2020*)

3.3.4. Karar ağacı (*Decision tree-DT*)

Karar Ağacı algoritması, popüler ve en basit makine öğrenmesi algoritması olup bir ağaç yapısı türünde sınıflandırma ve regresyon modelleri oluşturur. Algoritmanın amacı, önceki eğitim verilerinden çıkarılan karar kurallarını öğrenerek nihai sonuç değişkeninin sınıfını tahmin etmek için bir model oluşturmaktır. Karar Ağacı, düğüm ve dallardan meydana gelmektedir. Her bir düğüm bir özelliği temsil etmektedir ve her dal ise bir kararı temsil etmektedir [22]. 2019 ve 2020 yıllarına ait veri setlerine uygulanan DT algoritması sonucunda elde edilen hata matrisleri Şekil 16 ve Şekil 17' de verilmiştir.

```
[8737 rows x 6 columns]
[[1215  0  0  0  0  0]
 [  0 1440  1  0  0  0]
 [  0  0 190  0  0  0]
 [  0  0  0 29  0  0]
 [  0  0  0  0  8  0]
 [  0  0  0  0  0  1]]
Decision-Tree Accuracy: 0.9996532593619972
Decision-Tree Rmse: 0.018620973068096465
Decision-Tree R2: 0.9992327001690774
```

Şekil 16. 2019 hata matrisi (*confusion matrix-2019*)

```
[6841 rows x 6 columns]
[[1286  0  0  0  0  0]
 [  0 826  1  0  0  0]
 [  0  0 110  0  0  0]
 [  0  0  0 28  0  0]
 [  0  0  0  0  6  0]
 [  0  0  0  0  0  1]]
Decision-Tree Accuracy: 0.9995571302037201
Decision-Tree Rmse: 0.021044471869825904
Decision-Tree R2: 0.9990422138518443
```

Şekil 17. 2020 hata matrisi (*confusion matrix-2020*)

3.3.5. Naif bayes (*Naive bayes-NB*)

Bayes teoreminde bulunan giriş değişkenlerinin birbirinden bağımsız olduğu bir denetimli bir makine öğrenmesi algoritmasıdır. Tahmin yapmak için her sınıfa ait her bir özelliğin olasılıklarını kullanan tahmine dayalı analiz için basit ve güçlü bir algoritmadır. NB sınıflandırıcısının oluşturulması kolaydır ve büyük veri kümeleri içinde kullanışlıdır. NB sınıflandırıcısının çalışma mantığı; sisteme belirli miktarda eğitim verisi girilmesi ile başlar. Eğitim verilerinin bir sınıfa ait olması gerekmektedir. Eğitim verileri üzerinde gerçekleştirilen olasılık işlemleri ile sisteme girilen test verileri işlenir. Bu işlem daha önce elde edilen olasılık değerlerine göre yapılır ve daha sonra verilen test verilerinin sınıfı tespit edilir. Eğitim verisi sayısı ne kadar fazlaysa, test verilerinin gerçek kategorisini tespit etmek o kadar doğru olur [22]. 2019

ve 2020 yıllarına ait veri setlerine uygulanan NB algoritması sonucunda elde edilen hata matrisleri Şekil 18 ve Şekil 19' da verilmiştir.

```
[8737 rows x 6 columns]
[[1151  64  0  0  0  0]
 [ 93 1285  62  0  1  0]
 [  0  5 182  3  0  0]
 [  0  0  2 26  1  0]
 [  0  0  0  2  6  0]
 [  0  0  0  0  0  1]]
Naive-Bayes Accuracy: 0.9192094313453537
Naive Bayes Rmse: 0.2890752389234825
Naive Bayes R2: 0.8150807407476385
```

Şekil 18. 2019 hata matrisi (confusion matrix-2019)

```
[6841 rows x 6 columns]
[[1229  57  0  0  0  0]
 [ 107 690  30  0  0  0]
 [  0  1 106  3  0  0]
 [  0  0  3 20  5  0]
 [  0  0  0  0  6  0]
 [  0  0  0  0  0  1]]
Naive-Bayes Accuracy: 0.9087688219663419
Naive Bayes Rmse: 0.3020449933928025
Naive Bayes R2: 0.8026960534799281
```

Şekil 19. 2020 hata matrisi (confusion matrix-2020)

4. TARTIŞMA VE BULGULAR (DISCUSSION AND FINDINGS)

Çalışmada verilerin %67'si eğitim, %33'ü test verisi olarak kullanılmıştır. Oluşturulan tahmin modeli iki veri setine de uygulanmış, modelin performans değerlendirmesinde doğruluk, RMSE ve R^2 kullanılmıştır [23], [24]. Sonuçlar detaylı bir şekilde Tablo 4 ve Tablo 5 de verilmiştir.

Tablo 4. Doğruluk değerlerinin kıyaslanması-2019(Comparison of accuracy values-2019)

Yöntemler	KNN	NB	RF	DT	SVM
Doğruluk	0.976	0.919	0.997	0.999	0.988
RMSE	0.155	0.289	0.047	0.016	0.106
R^2	0.946	0.815	0.994	0.994	0.974

Çalışmada kullanılan 2019 yılı veri setine ait doğruluk değerlerinin kıyaslamasına bakıldığında en düşük RMSE(0,016) ve en yüksek doğruluk oranıyla(%99,9) en başarılı algoritma Karar Ağacı'dır. Karar Ağacı'nı RMSE=0,047 ve %99,7 doğruluk oranıyla Rastgele Orman algoritması takip etmektedir. En düşük başarı ise RMSE=0,289 ve %91,9 başarı oranıyla Naif Bayes algoritmasına aittir.

Tablo 5. Doğruluk değerlerinin kıyaslanması-2020 (Comparison of accuracy values-2020)

Yöntemler	KNN	NB	RF	DT	SVM
Doğruluk	0.996	0.908	0.998	0.999	0.984
RMSE	0.124	0.302	0.042	0.021	0.122
R^2	0.966	0.802	0.996	0.999	0.974

Çalışmada kullanılan 2020 yılı veri setine ait doğruluk değerlerinin kıyaslamasına bakıldığında yine en düşük RMSE (0,021) ve en yüksek doğruluk oranıyla(%99,9) en başarılı algoritma Karar Ağacı olmuştur. Karar Ağacı'nı sırasıyla RF, KNN, SVM, NB takip etmektedir. En başarısız algoritma ise RMSE=0,302 ve %90,8 başarı oranıyla Naif Bayes algoritması olmuştur.

İki farklı veri seti üzerinde makine öğrenmesi algoritmalarından KNN, SVM, RF, NB ve DT uygulanmış ve sonuçlar kıyaslandığında Karar Ağaçları ve Rastgele orman algoritmaları yüksek başarı oranına sahiptir. Literatürde yapılan çalışmalar ile oluşturulan modelin sonuçları Tablo 6'da kıyaslanmıştır.

Tablo 6. Çalışmaların doğruluk değerlerinin kıyaslanması (Comparison of accuracy values of studies)

Kaynak	Kullanılan Model	Değerlendirme Kriteri ve Performansı
[9]	Gradient Boosting Regressor, RF, KNN, DT	$GBR_{RMSE}=0,9336$ $GBR_{R2}=0,1302$
[10]	DT	$DT_{Doğruluk}=\%89,46$
[11]	YSA, SVM	$YSA_{MAE}=0,0191$ $SVM_{MAE}=0,0314$
[13]	RF, NB, SVM	$RF_{Doğruluk}=\%99,1$
Çalışmada	KNN, NB, RF, DT, SVM	$DT_{Doğruluk}=\%99,9$

Tablo 6 incelendiğinde Karar Ağacı ve Rastgele Orman ile gerçekleştirilen modellerin diğer modellere göre daha başarılı sonuçlar ortaya koyduğu görülmektedir.

5.SONUÇ (RESULTS)

Hava kirliliği akciğer kanseri, astım, solunum yetmezliği gibi birçok hastalığın yayılmasında önemli bir çevresel risk faktörüdür. Günümüzde devam eden salgın sürecinde hava kalitesinin tahmin edilmesi ve buna yönelik tedbirlerin alınması hastalıkların yayılma hızını etkilemesi açısından önem taşımaktadır. Bu çalışma, geçmiş hava kalitesi verileri ve Covid-19 salgın sürecindeki günlerde ölçümlenen PM_{10} , SO_2 , CO , O_3 ve NO_2 kirletici

konsantrasyonlarını ele almış ve başarılı bir tahmin için hava kalitesini makine öğrenme modelleriyle analiz etmiştir. Uygulanan algoritmalar arasında 0,289(2020 veri seti) ve 0,302(2020 veri seti) RMSE değerleri ile başarısı en düşük olan Naive Bayes algoritmasıdır. En yüksek başarı 0,016(2019 veri seti) ve 0,021(2020 veri seti) RMSE değeriyle Karar Ağacı algoritmasında elde edilmiştir.

Tüm dünyada hava kirlenmeye uzun süre maruz kalma solunum yolu hastalıklarında olduğu gibi Covid-19 virüsünün yayılmasında ve ölümlerle sonuçlanmasında önemli bir etken olarak karşımıza çıkmaktadır. Hava kalitesinin doğru tahmini ve iyileştirilmesi salgın sürecinin uzamamasında ve hastalığı hafif geçirmede önemli bir rol oynayabileceği üzerinde durulmuştur. Bütün dünya ülkeleri için Covid-19 salgını, kriz anında ve sonrasında insan sağlığını korumak adına yeniden hava kirliliği ile mücadelenin gözden geçirilmesinin önemini ortaya çıkarmıştır.

TEŞEKKÜR (ACKNOWLEDGMENT)

Bu çalışma, Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) 2210-A Yurt İçi Genel Yüksek Lisans Burs Programı kapsamında desteklenmiştir.

ÇIKAR ÇATIŞMASI BİLDİRİMİ (CONFLICT OF INTEREST STATEMENT)

Yazarlar tarafından herhangi bir çıkar çatışması bildirilmemiştir.

KAYNAKLAR (REFERENCES)

- [1] İ. Atacak, N. Arıcı, and D. Güner, “Modelling and evaluating air quality with fuzzy logic algorithm-Ankara- Cebeci sample”, *Int. J. Intell. Syst. Appl. Eng.*, vol. 5, no. 4, pp. 263–268, 2018. doi: 10.1039/b000000x.
- [2] P. Samani, C. García-Velásquez, P. Fleury, ve Y. Van Der Meer, “The impact of the covid-19 outbreak on climate change and air quality: four country case studies”, *Glob. Sustain.*, vol. 4, pp. 1–43, Oca. 2021. doi: 10.1017/sus.2021.4.
- [3] T. Şimşek ve V. Yılmaz, “Hava kalitesi ile covid-19 arasındaki ilişki: G-7 ülkeleri üzerine bir uygulama”, *J. Turkish Stud.*, vol. 15, no. 4, pp. 1353–1366, 2020. doi: 10.7827/turkishstudies.43883.
- [4] E. Conticini, B. Frediani, ve D. Caro, “Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in northern Italy?”, *Environ. Pollut.*, vol. 261, no. 114465, 2020. doi: 10.1016/j.envpol.2020.114465.
- [5] Zonguldak Çevre ve Şehircilik İl Müdürlüğü, “Zonguldak ili 2018 yılı çevre durum raporu”, 2019. [Çevrimiçi]. Available at: https://webdosya.csb.gov.tr/db/ced/icerikler/zongudak_2018-cdr_son-20190926093732.pdf. [Accessed: Jun. 12, 2021]
- [6] Y. Zhu, J. Xie, F. Huang, ve L. Cao, “Association between short-term exposure to air pollution and covid-19 infection: evidence from China”, *Sci. Total Environ.*, vol. 727, no. 138704, 2020. doi: 10.1016/j.scitotenv.2020.138704.
- [7] D. Rodríguez-Urrego ve L. Rodríguez-Urrego, “Air quality during the covid-19: PM2.5 analysis in the 50 most polluted capital cities in the world”, *Environmental Pollution*, vol. 266. Elsevier Ltd, Kas. 01, 2020. doi: 10.1016/j.envpol.2020.115042.
- [8] E. B. Brandt, A. F. Beck, ve T. B. Mersha, “Air pollution, racial disparities, and covid-19 mortality”, *Journal of Allergy and Clinical Immunology*, vol. 146, no. 1, pp. 61–63, 2020. doi: 10.1016/j.jaci.2020.04.035.
- [9] H. Doreswamy, K. Harishkumar, Y. Km, ve İ. M. Gad, “Forecasting air pollution particulate matter (PM2.5) using machine learning regression models”, *Procedia Computer Science*, vol. 171, pp. 2057–2066, 2020. doi: 10.1016/j.procs.2020.04.221.
- [10] G. K. Kang, J. Z. Gao, S. Chiao, S. Lu, ve G. Xie, “Air quality prediction: big data and machine learning approaches”, *Int. J. Environ. Sci. Dev.*, vol. 9, no. 1, pp. 8–16, 2018. doi: 10.18178/ijesd.2018.9.1.1066.
- [11] A. Masood ve K. Ahmad, “A model for particulate matter (PM2.5) prediction for Delhi based on machine learning approaches”, *Procedia Computer Science*, vol. 167, pp. 2101–2110, 2020. doi: 10.1016/j.procs.2020.03.258.
- [12] J. K. Deters, R. Zalakeviciute, M. Gonzalez, ve Y. Rybarczyk, “Modeling PM2.5 urban pollution using machine learning and selected meteorological parameters”, *J. Electr. Comput. Eng.*, pp. 1–14, Haz. 2017. doi: 10.1155/2017/5106045.

[13] S. Bali ve M. N. Sengar, “Indian air quality prediction and analysis using machine learning”, *J. Eng. Sci.*, vol. 11, no. 5, pp. 554–557, 2020.

[14] Y. Yıldırım, Ö. Zeydan, ve E. Karakavuz, “Kentleşme ve hava kalitesi açısından ilimiz Zonguldak”, Zonguldak Kent Sempozyumu, ss. 81–89, 2011.

[15] Zonguldak Çevre ve Şehircilik İl Müdürlüğü, “Zonguldak 2019 yılı çevre durum raporu”, 2020. [Çevrimiçi]. Available at: https://webdosya.csb.gov.tr/db/ced/icerikler/2019_zonguldak_cdr-20200914150210.pdf [Erişim tarihi: May. 21, 2021].

[16] Çevre ve Şehircilik Bakanlığı(ÇŞB), “Hava kalitesi istasyon veri indirme T.C. çevre ve şehircilik bakanlığı”, 2021. [Çevrimiçi].https://sim.csb.gov.tr/STN/STN_Report/StationDataDownloadNew [Erişim tarihi: Mar. 10, 2021].

[17] DSÖ, “Science in 5 - episode 9 - hava kirliliği ve covid-19”, 2021. [Çevrimiçi]. Available: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/media-resources/science-in-5/episode-9---air-pollution-covid-19> [Accessed: Jan. 01, 2021].

[18] DSÖ, “Koronavirüs hastalığı (covid-19): iklim değişikliği”, 2021. [Çevrimiçi]. <https://www.who.int/news-room/q-a-detail/coronavirus-disease-covid-19-climate-change> [Erişim tarihi: Ocak 01, 2021].

[19] A. Masih, “Machine learning algorithms in air quality modeling”, *Glob. J. Environ. Sci. Manag.*, vol. 5, no. 4, pp. 515–534, 2019. doi: 10.22034/gjesm.2019.04.10.

[20] J. Huo, T. Shi, ve J. Chang, “Comparison of random forest and SVM for electrical short-term load forecast with different data sources”, in *7th IEEE International Conference on Software Engineering and Service Science, Beijing, China, Jul. 2016*, pp. 1077–1080. doi: 10.1109/ICSESS.2016.7883252.

[21] Z. Yao ve W. L. Ruzzo, “A regression-based k nearest neighbor algorithm for gene function prediction from heterogenous data”, *BMC Bioinformatics*, vol. 7, no. 1, pp. 1–11, Mar. 2006. doi: 10.1186/1471-2105-7-s1-s11.

[22] A. Dey, “Machine learning algorithms: a review”, *Int. J. Comput. Sci. Inf. Technol.*, vol. 7, no. 3, pp. 1174–1179, 2016.

[23] S. Çınaroğlu, “Sağlık harcamasının tahmininde makine öğrenmesi regresyon yöntemlerinin karşılaştırılması”, *Uludağ Üniversitesi Mühendislik Fakültesi Derg.*, c. 22, sayı 2, ss. 179–199, 2017. doi: 10.17482/uumfd.338805.

[24] S. Karasu, A. Altan, Z. Sarac, ve R. Hacıoğlu, “Prediction of bitcoin prices with machine learning methods using time series data”, in *26th IEEE Signal Processing and Communications Applications Conference (SIU), Izmir, Turkey, Jul. 2018*, pp. 1–4. doi: 10.1109/SIU.2018.8404760.

This is an open access article under the CC-BY license
(<https://creativecommons.org/licenses/by/4.0/>)

