

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**SÜRDÜRÜLEBİLİR HAVA KALİTESİ İÇİN YAPAY ZEKA YÖNTEMLERİ
İLE PARTİKÜLER MADDE TAHMİNİNİN MODELLENMESİ**

YÜKSEK LİSANS TEZİ

Saliha ÇELİKCAN BİLGİN

Meteoroloji Mühendisliği Anabilim Dalı

Atmosfer Bilimleri Programı

ARALIK 2023

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**SÜRDÜRÜLEBİLİR HAVA KALİTESİ İÇİN YAPAY ZEKA YÖNTEMLERİ
İLE PARTİKÜLER MADDE TAHMİNİNİN MODELLENMESİ**

YÜKSEK LİSANS TEZİ

**Salih ÇELİKCAN BİLGİN
(511021003)**

Meteoroloji Mühendisliği Anabilim Dalı

Atmosfer Bilimleri Programı

Tez Danışmanı: Prof. Dr. Hüseyin TOROS

ARALIK 2023

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

**MODELING PARTICULATE MATTER ESTIMATION WITH ARTIFICIAL
INTELLIGENCE METHODS FOR SUSTAINABLE AIR QUALITY**

M.Sc. THESIS

**Saliha ÇELİKCAN BİLGİN
(511021003)**

Department of Meteorology Engineering

Atmospheric Science Programme

Thesis Advisor: Prof. Dr. Hüseyin TOROS

DECEMBER 2023

İTÜ, Lisansüstü Eğitim Enstitüsü'nün 511021003 numaralı Yüksek Lisans Öğrencisi Saliha ÇELİKCAN BİLGİN, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı “SÜRDÜRÜLEBİLİR HAVA KALİTESİ İÇİN YAPAY ZEKA YÖNTEMLERİ İLE PARTİKÜLER MADDE TAHMİNİNİN MODELLENMESİ” başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : **Prof. Dr. Hüseyin TOROS**
İstanbul Teknik Üniversitesi

Jüri Üyeleri : **Prof. Dr. Bülent Oktay AKKOYUNLU**
Marmara Üniversitesi

Dr. Öğr. Üyesi Ceyhan KAHYA
İstanbul Teknik Üniversitesi

Teslim Tarihi : 13 Kasım 2023
Savunma Tarihi : 22 Aralık 2023





Aileme,



ÖNSÖZ

Tez sürecim boyunca kısıtlı vaktini bana ayıran, bilgi birikimini, desteklerini esirgemeyen ve çok değerli katkılarda bulunan hocam Prof. Dr. Hüseyin TOROS'a, akademik tecrübesini her daim benimle paylaşan, beni yönlendiren ve bana yol gösteren eşim Prof. Dr. Turgay Tugay BİLGİN'e, sevgisi ile bana her zaman enerji veren, motivasyon kaynağım sevgili kızım Gökçe Ceyda BİLGİN'e, sonsuz sevgi, güven ve destekleri için anneme ve babama, yarım bıraktığım yüksek lisans eğitimime geri dönmeme vesile olan değerli arkadaşlarım Emine KÜÇÜKOĞLU, Serkan ÖZGÜR, Şiir Deniz ÜNAL ve Halkbank Bursa Bölge Koordinatörlüğü Kredi Tahsis ve Yönetimi Bölümünde birlikte çalıştığım tüm arkadaşlarıma teşekkürlerimi sunarım.

Kasım 2023

Saliha ÇELİKCAN BİLGİN
(Meteoroloji Mühendisi)



İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	vii
İÇİNDEKİLER	ix
KISALTMALAR	xi
SEMBOLLER	xiii
ÇİZELGE LİSTESİ.....	xv
ŞEKİL LİSTESİ.....	xvii
ÖZET.....	xix
SUMMARY	xxiii
1. GİRİŞ.....	1
1.1 Tezin Amacı	5
1.2 Literatür Araştırması	5
1.2.1 Eksik kirletici verisinin tamamlanmasına yönelik çalışmalar.....	6
1.2.2 Hava kirleticilerinin tahminine yönelik çalışmalar	7
1.2.3 Bursa ilini konu alan hava kirliliği çalışmaları	11
2. HAVA KALİTESİNİ ETKİLEYEN FAKTÖRLER	13
2.1 Coğrafik Koşullar	13
2.2 Antropojenik Kaynaklar	13
2.3 Meteorolojik Koşullar	14
3. BURSA İLİ'NİN COĞRAFİK KONUMU, İKLİMİ VE HAVA KALİTESİ	17
3.1 Bursa'nın İklimi ve Coğrafi Koşulları	18
3.2 Bursa'nın Hava Kalitesi	18
4. HAVA KİRLİLİĞİ TAHMİNİNDE YAPAY ZEKA YÖNTEMLERİ	19
4.1 Rastgele Orman (Random Forest).....	19
4.2 Çok Katmanlı Algılayıcı (Multilayer Perceptron)	20
4.3 Gradyan Artırma (Gradient Boosting)	20
4.4 Destek Vektör Makineleri (Support Vector Machines)	20
4.5 Lineer Regresyon (Linear Regression)	21
4.6 Evrişimli Sinir Ağı (Convolutional Neural Network).....	21
4.7 Yinelemeli Yapay Sinir Ağı (Recurrent Neural Network).....	21
4.8 Uzun Kısa Dönem Bellek (Long Short Term Memory)	21
4.9 Karar Ağacı (Decision Tree)	22
5. VERİ ve YÖNTEM	23
5.1 Veri Tanımı	23
5.1.1 Hava kirliliği verileri.....	23
5.1.2 Meteorolojik veriler	24
5.2 Veri Ön İşleme	26
5.3 Yöntem	31
5.3.1 Makine öğrenmesi yöntemleri ile PM _{2.5} ve PM ₁₀ tahmini.....	33

5.3.1.1 Bursa Hava Kalitesi Ölçüm İstasyonu için PM _{2.5} ve PM ₁₀ tahmini ...	33
5.3.1.2 İnegöl, Kestel ve Beyazıt Caddesi istasyonları için PM _{2.5} tahmini....	38
5.3.1.3 Uludağ Üniversitesi ve Kültürpark İstasyonları için PM ₁₀ tahmini ..	44
6. SONUÇ VE ÖNERİLER.....	49
KAYNAKLAR.....	53
ÖZGEÇMİŞ.....	59



KISALTMALAR

WHO	: Dünya Sağlık Örgütü (World Health Organisation)
PM	: Partiküler madde
RF	: Rastgele Orman (Random Forest)
SVM	: Destek Vektör Makineleri (Support Vektör Machines)
LR	: Lineer Regresyon (Linear Regression)
CNN	: Evrişimli Sinir Ağı (Convolutional Neural Network)
LSTM	: Uzun Kısa Dönem Bellek (Long Short Term Memory)
BPNN	: Geri Beslemeli Yapay Sinir Ağı (Back Propagation Neural Network)
GB	: Gradient Boosting
GAM	: Generalized Additive Model
GRU	: Kapı Özyinelemeli Geçitler
HKİ	: Hava Kalitesi İndeksi (Air Quality Index-AQI)
DT	: Karar Ağacı (Decision Tree)
MLP	: Çok Katmanlı Algılayıcı (Multilayer Perceptron)
RBF	: Radyal Tabanlı Fonksiyon (Radial Basis Function)
TÜİK	: Türkiye İstatistik Kurumu
MCP	: Ülkeler arası işbirliği ile hazırlanan yayınlar (Multiple Country Publications)
SCP	: Yazarları aynı ülkeden olan yayınlar (Single Country Publications)
OSB	: Organize Sanayi Bölgesi
ECMWF	: Avrupa Orta Menzilli Hava Tahminleri Merkezi
UHKİA	: Ulusal Hava Kalitesi İzleme Ağı



SEMBOLLER

u₁₀	: 10 metrede rüzgarın u (doğuya doğru) bileşeni
v₁₀	: 10 metrede rüzgarın v (kuzeye doğru) bileşeni
WS	: 10 metrede rüzgar hızı
T_{2m}	: 2 metrede hava sıcaklığı
Td_{2m}	: 2 metrede çiy noktası sıcaklığı
RH	: Bağıl nem



ÇİZELGE LİSTESİ

Sayfa

Çizelge 5.1 : İstasyonlarda ölçümü yapılan kirleticiler.	23
Çizelge 5.2 : Meteorolojik parametreler ve birimleri.	26
Çizelge 5.3 : Bursa istasyonu için $PM_{2.5}$ tahmininde GB ve MLP hata metrikleri. ..	36
Çizelge 5.4 : Bursa istasyonu için PM_{10} tahmininde GB ve MLP hata metrikleri. ...	38
Çizelge 5.5 : GB ve MLP yöntemlerine ait hata metrikleri.	44
Çizelge 5.6 : İnegöl Hava Kalitesi İzleme İstasyonu tam veri seti.	44
Çizelge 5.7 : GB ve MLP yöntemlerine ait hata metrikleri.	48



ŞEKİL LİSTESİ

Sayfa

Şekil 1.1 : Anahtar kelimelere göre yayın sıklığı.	7
Şekil 1.2 : MCP/SCP dağılımı.	8
Şekil 3.1 : Bursa OSB’leri ve Limanları.	17
Şekil 5.1 : Hava kalitesi izleme istasyonların konumu.	24
Şekil 5.2 : Aylık bazda PM _{2.5} eksik veri oranı dağılımı.	27
Şekil 5.3 : Aylık bazda PM ₁₀ eksik veri oranı dağılımı.	27
Şekil 5.4 : Aylık bazda SO ₂ eksik veri oranı dağılımı.	28
Şekil 5.5 : Aylık bazda NO ₂ eksik veri oranı dağılımı.	28
Şekil 5.6 : Aylık bazda O ₃ eksik veri oranı dağılımı.	29
Şekil 5.7 : Kirletici korelasyon matrisi.	30
Şekil 5.8 : XGBoost ile tamamlanmış örnek veri seti.	31
Şekil 5.9 : Partiküler madde tahmini iş akış diyagramı	32
Şekil 5.10 : KNIME Platformu	33
Şekil 5.11 : Bursa istasyonu için Gradient Boosting ile PM _{2.5} tahmin modeli.	34
Şekil 5.12 : Ölçüm verisi ve GB ile tahmin verisinin görsel karşılaştırılması.	35
Şekil 5.13 : Bursa istasyonu için Multilayer Perceptron ile PM _{2.5} tahmin modeli.	35
Şekil 5.14 : Ölçüm verisi ve MLP ile tahmin verisinin görsel karşılaştırılması.	36
Şekil 5.15 : Bursa istasyonu için Gradient Boosting ile PM ₁₀ tahmin modeli.	36
Şekil 5.16 : PM ₁₀ ölçüm verisi ve GB ile tahmin verisinin görsel karşılaştırılması.	37
Şekil 5.17 : Bursa istasyonu için Multilayer Perceptron ile PM ₁₀ tahmin modeli.	37
Şekil 5.18 : PM ₁₀ ölçüm ve MLP ile tahmin verisinin görsel karşılaştırılması.	38
Şekil 5.19 : GB yöntemi ile PM _{2.5} ve PM ₁₀ tahmini modeli.	39
Şekil 5.20 : İnegöl istasyonu için PM ₁₀ ölçüm ve GB ile tahmin verisi.	40
Şekil 5.21 : Kestel PM ₁₀ ölçüm ve GB ile tahmin verisi.	41
Şekil 5.22 : Beyazıt Caddesi PM ₁₀ ölçüm ve GB ile tahmin verisi.	41
Şekil 5.23 : MLP yöntemi ile PM _{2.5} ve PM ₁₀ tahmin modeli.	42
Şekil 5.24 : İnegöl PM ₁₀ ölçüm ve MLP ile tahmin verisi.	42
Şekil 5.25 : Kestel PM ₁₀ ölçüm ve MLP ile tahmin verisi.	43
Şekil 5.26 : Beyazıt Caddesi PM ₁₀ ölçüm ve MLP ile tahmin verisi.	43
Şekil 5.27 : GB yöntemi ile eğitim verisi kullanarak PM ₁₀ ve PM _{2.5} tahmini.	45
Şekil 5.28 : Uludağ Üniversitesi PM _{2.5} ölçüm ve GB ile tahmin verisi.	46
Şekil 5.29 : Kültürpark PM _{2.5} ölçüm ve GB ile tahmin verisi.	46
Şekil 5.30 : MLP yöntemi ile eğitim verisi kullanarak PM ₁₀ ve PM _{2.5} tahmini.	47
Şekil 5.31 : Uludağ Üniversitesi PM _{2.5} ölçüm ve MLP ile tahmin verisi.	47
Şekil 5.32 : Kültürpark PM _{2.5} ölçüm ve MLP ile tahmin verisi.	48



SÜRDÜRÜLEBİLİR HAVA KALİTESİ İÇİN YAPAY ZEKA YÖNTEMLERİ İLE PARTİKÜLER MADDE TAHMİNİNİN MODELLENMESİ

ÖZET

Biyosferde canlı ve cansız tüm varlıkları doğrudan veya dolaylı olarak etkileyen en önemli konulardan biri hava kalitesidir. Hava kalitesi, atmosferde bulunan gazlar, partiküller ve diğer kirleticilerin seviyesini tanımlar. Temiz hava insan sağlığına olumlu etkiler sağlarken, kirli hava solunum rahatsızlıkları ve çevresel sorunlara yol açabilir. Bu nedenle, hava kalitesinin korunması ve iyileştirilmesi önemli bir halk sağlığı ve çevre koruma meselesidir. Hava kalitesi, ölçülmesi ve izlenmesi gereken bir parametre olup, çeşitli kaynaklardan salınan kirleticiler tarafından kolaylıkla etkilenmektedir.

Hava kalitesi dengesinin bozulması, ortam havasındaki çeşitli maddelerin canlı sağlığını olumsuz etkilemeye başladığı sınırın aşıldığı anlamında gelmektedir. Hava kalitesi dengesi, doğal ya da insan kaynaklı nedenlerle bozulabilmektedir. İnsan kaynaklı hava kirliliği, özellikle Sanayi Devrimi ile birlikte dramatik şekilde artarak canlı sağlığını tehdit eder hale gelmiştir.

Hava kirliliği, solunum yolu rahatsızlıklarına, astım, bronşit ve diğer solunum sistemi hastalıklarının artmasına neden olmaktadır. Kirli hava, kalp ve akciğer hastalıkları ile kanser riskini artırmaktadır. Ayrıca, diğer pek çok hastalıkla da ilişkisi tespit edilmiştir. Bozulan hava kalitesi çocuklar, yaşlılar ve kronik sağlık sorunu olan bireyler üzerinde daha belirgin ve ciddi etkilere sahiptir.

Kirliliğin özellikle insan sağlığı üzerindeki olumsuz etkileri, bilim insanlarının dikkatini kirlilik kaynaklarına, etkilerine, etkileştiği parametrelere yönlendirmesine neden olmuştur. Multidisipliner bir konu olarak hava kirliliği, meteorologlar, çevre mühendisleri, şehir ve bölge planlama uzmanları, halk sağlığı uzmanları, veterinerler gibi birbirinden farklı temel eğitimleri olan araştırmacıların ortak ilgi alanı olmuştur.

Bilim insanları ile araştırmacıların bir kısmı hava kirliliğinin olumsuz etkilerine karşı bir uyarı mekanizması da oluşturması düşüncesiyle kirletici parametrelerin çeşitli zaman periyotlarında tahminine yönelmiştir. Başarılı hava kalitesi tahmini, kamu otoritelerinin toplum ve çevre ile ilgili tedbirleri zamanında almasını, halkın erken uyarılmasını, böylelikle kirletici kaynaklı zarar oluşumunun engellenmesini ya da en az düzeyde etkilenilmesini sağlayabilecektir. Bilgisayarların veri işleme kapasitelerinin artması ve ileri veri işleme yöntemlerinin geliştirilmesi, hava kirliliği ve hava kalitesi tahminine yönelik çalışmalarının sayısının hızla artmasına ve başarılı tahmin modellerinin geliştirilmesine katkıda bulunmuştur. Tahmin modelleri temel olarak kimyasal taşınım modelleri ve veri odaklı modeller olarak ikiye ayrılabilir. Kimyasal taşınım modelleri, hava kirliliğinin kaynağından itibaren atmosferdeki hareketini ve yayılmasını izler. Atmosferik koşullar, hava akımları, sıcaklık, nem ve diğer faktörler bu modellerde dikkate alınır. Bu modeller, belirli bir bölgede veya şehirde hava kalitesinin nasıl etkileneceğini tahmin etmek için kullanılır.

Veri odaklı modeller ise gerçek zamanlı verileri kullanarak hava kirliliği seviyelerini tahmin ederler. Hava kalitesi izleme istasyonları ile meteoroloji istasyonlarından alınan gerçek zamanlı veriler bu modelleri besler. Makine öğrenimi ve istatistiksel yöntemler, bu verileri analiz ederek gelecekteki kirlilik seviyelerine yönelik öngöründe bulunur.

Veri odaklı modeller başlığı altında yer alan makine öğrenmesi, derin öğrenme gibi yapay zeka uygulamaları, kirlenici seviyesinin ya da hava kalitesinin kısa, orta ve uzun vadeli tahmininde sıklıkla kullanılan yöntemler haline gelmiştir. Yapay zeka yöntemleri, büyük miktarda veri analizi ve karmaşık desenleri tanıma konusunda oldukça etkilidir. Bu durum hava kirliliği tahminlerinde daha hassas sonuçlar elde etmeyi sağlar. Yapay zeka, meteorolojik veriler, hava kalitesi izleme istasyonlarından gelen veriler ve diğer faktörleri analiz ederek kirlilik seviyelerini tahmin etmekte başarılıdır. Ayrıca gerçek zamanlı veri analizi yapabilme yetenekleri sayesinde hava kirliliği seviyelerini anlık olarak izlemeyi sağlar. Bu şekilde, hızlı müdahale ve risklerin azaltılması için önemli bir avantaj yaratmaktadır. Karar vericilere gelecekteki kirlilik durumlarına göre politika oluşturma konusunda da yardımcı olmaktadır.

Yapay zekaya dayalı tahmin modellerinin bu özellikleri, hava kirliliğine yönelik çalışmalarda sıklıkla tercih edilmelerini sağlamıştır. Bazen bir tahmin modelinde tek başına bir yöntem denenirken, bazen de birden fazla yapay zeka yöntemi birlikte kullanılarak hava kirliliği tahmin başarıları artırılmaya çalışılmıştır.

Bursa, nüfus, kentleşme, sanayileşme, coğrafik konum, topoğrafya ve iklim gibi çeşitli özellikleri birlikte değerlendirildiğinde, hava kalitesi dengesi hassasiyetle takip edilmesi gereken bir şehirdir. Bu nedenle şehir için sürekli, düzenli, sağlıklı veri üreten, geleceğe yönelik bilgi ve gerektiğinde uyarı üretebilecek bir hava kalitesi tahmin yapısının oluşturulması önemlidir. Çalışmada önerilen model, Bursa iline ait verilerle oluşturulmuş olmakla birlikte, hava kalitesi dikkatle takip edilmesi gereken diğer yerleşim birimleri için çalıştırılabilecek ve geliştirilebilecek bir yapı içermektedir.

Bu çalışmanın amacı, meteorolojik parametreler kullanarak partiküler madde (PM₁₀ ya da PM_{2.5}) seviyesinin yapay zeka yöntemleri ile tahmin edilmesini amaçlayan sürdürülebilir bir hava kalitesi modelinin oluşturulmasıdır.

Çalışmada Bursa ili sınırları içerisindeki 6 hava kalitesi izleme istasyonundan elde edilen bir yıllık, saatlik periyotta 5 farklı kirlenici (PM₁₀, PM_{2.5}, NO₂, SO₂, O₃) ölçüm verisi ile meteorolojik veriler kullanılmıştır. İlk aşamada kirlenici veri seti ön işleme tabi tutulmuş ve eksik verileri tamamlanmıştır. İkinci aşamada, PM₁₀ ve PM_{2.5} konsantrasyonları arasındaki yüksek korelasyondan yola çıkılarak bir model geliştirilmiştir. Modelin her iki parametrenin ölçüldüğü istasyondan elde edilen bir haftalık PM₁₀ ve PM_{2.5} tahmin sonuçları başarılı bulunmuştur. Üçüncü aşamada, sadece bir tür partiküler madde (PM₁₀ ya da PM_{2.5}) konsantrasyonu ölçülen beş istasyon için yeni bir tahmin modeli oluşturulmuştur. İstasyonda ölçümü yapılan partiküler madde, ölçümü yapılan diğer kirleniciler ve meteorolojik veriler ile her iki partiküler madde türünün ölçüldüğü istasyona ait modelin eğitim verisi seti, yeni modeldeki makine öğrenmesi modülüne girdi olarak verilmiştir. Elde edilen partiküler madde tahmin sonuçlarının teyit edilmesi için aynı modüle tahmin edilen partiküler madde verisinin kullanıldığı bir süreç daha ilave edilmiştir. Bu süreçte istasyonda ölçümü yapılan partiküler madde için tahmin yapılmaktadır. İstasyonda ölçülen partiküler madde ile modülle tahmin edilmiş hali arasındaki yüksek korelasyon

ölçülmeyen partiküler madde için yapılan tahminin tatmin edici derece başarılı olduğunu göstermiştir.

Nüfus bakımından Türkiye'nin dördüncü büyük şehri ve bir sanayi üssü olan Bursa için oluşturulacak etkin ve sürdürülebilir partiküler madde tahmin modeli, tüm paydaşların hava kalitesinin seyrinin takip edebilmesini, tahminler doğrultusunda olumsuz durum beklentisi oluştuğunda kamu otoritelerinin zamanında gerekli tedbirleri almasını sağlayabilir. Cihaz arızası vb. nedenlerle bir istasyondan partiküler madde (PM_{2.5} ya da PM₁₀) ölçüm bilgisi gelmediğinde, tahmin modeli çalıştırılarak elde edilen sonuç, ölçüm bilgisi yerine değerlendirmeye alınabilir. Ayrıca bir istasyonda iki partiküler maddeden birinin ölçümünün yapılması, diğerinin tahmin modeli ile elde edilmesi ölçüm maliyetlerinin azaltılmasını sağlayabilir.





MODELING PARTICULATE MATTER ESTIMATION WITH ARTIFICIAL INTELLIGENCE METHODS FOR SUSTAINABLE AIR QUALITY

SUMMARY

Air quality is one of the most important issues that directly or indirectly affect all living and non-living things in the biosphere. In the realm of environmental concerns, it stands as a linchpin issue, touching upon public health, ecological stability, and the delicate balance of our shared ecosystem. This research focuses on Bursa, a city that embodies the challenges of urbanization, industrialization and various environmental factors, and presents a forecasting model to contribute to the air quality management of the city.

The pervasive impact of air quality on diverse aspects of life necessitates a profound exploration of its dynamics. When the air quality balance is disturbed, it is understood that various substances in the ambient air begin to adversely affect living health. Air quality balance can be disturbed by natural or human-induced causes. Human-induced air pollution has increased dramatically, especially with the Industrial Revolution. In the wake of rapid industrialization and urban expansion, concerns about air pollution have escalated, posing threats to human health and environmental sustainability.

Air quality, defined by the composition of gases, particles and pollutants in the atmosphere, is a delicate balance against natural and man-made disturbances. Clean air has positive effects on human health, contributing to improved quality of life. Conversely, polluted air can lead to numerous health problems, from respiratory ailments to more serious conditions such as heart disease and cancer.

As industrialization has increased, anthropogenic air pollution has emerged as a significant threat to public health. The adverse impacts of pollution, particularly on vulnerable demographic groups such as children, the elderly and those with chronic health conditions, underscore the urgency of addressing air quality issues. Respiratory diseases, asthma, bronchitis and other ailments are rooted in deteriorating air quality and require a holistic approach to air quality management.

The complex web of air quality problems has led scientists and researchers to adopt a multidisciplinary approach. Meteorologists, environmental engineers, urban planners and public health experts come together in an effort to understand the sources, impacts and parameters that affect air pollution. This collaborative approach underpins comprehensive research and sheds light on the complexities of addressing air quality issues.

Given the dynamic nature of air quality, monitoring and measurement become essential. Factors such as environmental pollution and industrial activities can destabilize and rapidly change the balance of air quality. This necessitates continuous measurement and monitoring, which provides a real-time understanding of changing dynamics and enables timely interventions.

The quest to understand and predict air quality has led to the development of predictive models. Crucial for public health and environmental protection, these models provide

insights into how air quality may evolve over time. Essentially, they serve as tools to predict potential risks, formulate preventive measures and reduce the impact of pollutants on society and the environment.

Basically two main categories of predictive models emerge: chemical transport models and data-driven models. Chemical transport models rigorously track the movement and dispersion of air pollution, taking into account atmospheric conditions, air currents, temperature, humidity and other variables. Data-driven models, on the other hand, use real-time data, including pollutant data from air quality monitoring stations and meteorological data, to predict future pollution levels.

In the field of data-driven models, artificial intelligence (AI) has emerged as a transformative force. Machine learning and deep learning applications under the AI umbrella offer sophisticated tools for short, medium and long-term prediction of pollutant levels. AI's capacity to analyze large datasets, discern complex patterns and provide precise predictions has revolutionized the field of air pollution forecasting.

Machine learning and statistical methods, an integral part of data-driven models, analyze real-time data from air quality monitoring stations and meteorological sources. This analytical skill enables AI to predict future pollution levels with a high degree of accuracy. AI's real-time data analysis capabilities not only facilitate instantaneous monitoring of air pollution levels, but also provide a significant advantage for rapid response and risk mitigation.

The selection of Bursa as a case study is deliberate given its unique combination of characteristics such as population density, urbanization, industrialization, geographical location, topography and climate. As the fourth largest city in Turkey, Bursa presents a combination of challenges and opportunities for air quality management. The need for a dedicated air quality forecasting structure capable of producing continuous, regular and future-oriented information for Bursa is clearly evident.

The main objective of this study is to develop a sustainable air quality model for Bursa, focusing on the prediction of particulate matter (PM₁₀ and PM_{2.5}) levels using artificial intelligence methods based on meteorological parameters. The study utilizes data from six air quality monitoring stations in Bursa, covering a comprehensive one-hour period of one year. The dataset, which includes measurements of five different pollutants (PM₁₀, PM_{2.5}, NO₂, SO₂, O₃) and meteorological data, forms the basis for model development. In this study, meteorological data, including wind speed, temperature at 2 meters, relative humidity, boundary layer height, mean sea level pressure, boundary layer distribution, surface net solar radiation, and total cloud cover, were utilized.

The study proceeds in phases, starting with the pre-processing of the pollutant dataset and the completion of missing data. The study addresses the significant challenge of missing data completion in pollutant datasets, a crucial factor impacting the success of prediction models. Analyzing pollutant data from six stations, issues such as the number and periodicity of missing data, simultaneous missing measurements for the same pollutant in the same station, and missing data for the same pollutant type at different stations were visually and analytically examined. Initially, linear regression successfully completed missing PM_{2.5}, NO₂, and O₃ data at the Uludağ University station. However, when facing simultaneous missing data across all stations, an alternative method was sought. The XGBoost machine learning method, exhibiting success in internal behavior analysis, was applied, proving effective for data completion across all pollutant types.

In the second stage, a model was developed based on the high correlation between PM_{10} and $PM_{2.5}$ concentrations. Firstly, for Bursa station measuring both PM_{10} and $PM_{2.5}$, a model using Gradient Boosting and MultiLayer Perceptron methods was evaluated. The MultiLayer Perceptron-based model proved more successful. The success of the one-week PM_{10} and $PM_{2.5}$ prediction results from the station measuring both parameters is highlighted.

Secondly, for stations measuring only one particulate matter type, a two-step model demonstrated successful predictions. A new forecasting model is introduced for stations measuring only one type of particulate matter concentration. In the model, where only one particulate matter type is measured at the station, the focus is on a model that successfully calculates the other particulate matter using measured pollutant data and meteorological information. This two-step model initially predicts $PM_{2.5}$ for stations measuring only PM_{10} . In the second step, to measure the model's success, the predicted $PM_{2.5}$ data is fed into the module like measurement data, and this time, PM_{10} prediction is made. The PM_{10} measurement data is then compared with the PM_{10} prediction data using error metrics. The evaluation of model results indicates that both Gradient Boosting and MultiLayer Perceptron methods produce results close to the actual values. Although both methods are successful, MultiLayer Perceptron performs better for Inegol and Beyazıt Caddesi stations, while Gradient Boosting is more effective for the Kestel station. The success of the second step of the model confirms the hypothesis that the predicted $PM_{2.5}$, which is not measured at the station but calculated by the model in the first step, is successful since it closely aligns with the measurement value of the predicted PM_{10} . A similar model is applied to Uludağ University and Kültürpark stations where $PM_{2.5}$ measurements are taken but PM_{10} measurements are not. According to the model results, predictions with the MultiLayer Perceptron method are significantly more successful in both stations. Considering all results collectively, the model's prediction performance is high for PM_{10} or $PM_{2.5}$ concentrations not measured at the stations, employing both artificial intelligence methods.

Validation of particulate matter prediction results involves a rigorous process where the predicted particulate matter data is used for additional validation. The high correlation observed between the measured particulate matter and the shape predicted by the module confirms the effectiveness of the prediction model for unmeasured particulate matter.

The implications of the developed model for Bursa are far-reaching. In a city full of life, industrial activities and environmental complexities, a sustainable particulate matter prediction model becomes a beacon for informed decision-making. Stakeholders ranging from public authorities to individuals stand to benefit from a tool that not only comprehensively monitors air quality but also enables timely interventions in response to predicted scenarios. Although the model proposed in this study is based on the data of Bursa, it contains a structure that can be operated and developed for other settlements whose air quality should be carefully monitored.

The results obtained from the developed prediction model can replace the measurement data. Furthermore, the potential of the model to reduce measurement costs by predicting one of the particulate matter species at a station underlines the practicality and cost-effectiveness of the model.

In conclusion, this research presents two high-performance predictive models for air quality in urban environments. The study has produced a conclusion that encourages policy makers, environmental agencies and researchers to embrace the transformative potential of AI in the collective quest for sustainable urban environments.





1. GİRİŞ

Dünya Sağlık Örgütü (WHO), hava kirliliğini atmosferin doğal özelliklerini değiştiren herhangi bir kimyasal, fiziksel veya biyolojik ajan tarafından iç veya dış ortamın kirlenmesi olarak tanımlamaktadır (WHO Air Pollution, 2023). Hava kalitesi ise havadaki bu ajanların varlığının ve yoğunluğunun insan sağlığı, çevre ve ekosistem üzerindeki potansiyel etkilerini sınıflandırmak ve değerlendirmek amacıyla kullanılan bir terimdir.

Hava kalitesi doğal ya da insan kaynaklı olarak bozulabilmektedir. Doğal nedenler arasında volkanik patlamalar, orman yangınları, çöl tozu taşınımı sayılabilir. Ulaşım, endüstriyel ve tarımsal süreçler, evsel ısınma ve pişirme, atık bertarafı insan tarafından meydana getirilen kirlenmenin ana kaynaklarıdır. Havadaki kirlilik sadece kaynağın olduğu sınırlı bir bölgeyi etkileyebileceği gibi, bazen taşınım yoluyla binlerce kilometre uzaklıkta dahi etkisini hissettirebilmektedir. Örneğin, Sahra Çölü'nden taşınan toz, zaman zaman ülkemizin hava kalitesini etkilemektedir. Karaman ili için yapılan bir çalışmada, taşınımın olduğu dönemde şehirdeki partiküler madde miktarının önemli ölçüde arttığı tespit edilmiştir (Sassi ve diğ, 2020).

Hakim rüzgar yönü ve yağışın hangi yönden bölgeye ulaştığı, taşıdığı yükün içeriğini etkilemektedir. Bu nedenle ilgilenilen alanın hangi yönlerden gelen taşınıma maruz kaldığı önemlidir. Bu konuda ülkemizde de çalışmalar yapılmaktadır. Dünyanın sayılı büyük metropollerinden olan İstanbul için yapılan bir çalışmada, şehre taşınan kirleticilerin kaynakları geri yönde yörünge analizi tekniği kullanılarak araştırılmış, taşınımında etkili olan rüzgar yönleri ve yağışların bölgeye taşınım yönü belirlenmiştir (Saylan ve diğ, 2009).

Hava kirletici konsantrasyonları çoğunlukla insan aktiviteleri ile bağlantılı olarak çeşitli zaman periyotlarında değişkenlik gösterebilmektedir. Kışın karışım tabakası yüksekliğinin yaza göre daha sıkı olması ve evsel ısınma ihtiyacı nedeniyle yakıt kullanımı, kirletici konsantrasyonlarının otoriteler tarafından belirlenen sınırları aşmasına neden olmaktadır. Hafta içi endüstri ve trafik nedeniyle kirletici konsantrasyonları yüksek seyrederken, hafta sonu faaliyetlerin azalmasıyla daha düşük

değerler ölçülmektedir. İstanbul'da mekansal ve zamansal olarak PM₁₀ kirletici kaynakları ve meteorolojik faktörlerin kirliliğe etkisini konu alan bir çalışmada, Pazar günleri PM₁₀ seviyesinin incelenen tüm istasyonlarda hafta içi günlere oranla yaklaşık % 2- % 21 daha düşük olduğu belirtilmiştir (Unal ve diğ, 2011).

Türkiye genelinde 2008-2012 arasında kalan dönem için PM₁₀ konsantrasyonlarının mekânsal ve mevsimsel değişimleri incelendiğinde; güneydoğu bölgelerinde ilkbahar ve yaz aylarında yüksek konsantrasyon değerleriyle karşılaşmıştır. Bu durumun Sahra Çölü'nden ve Mezopotamya'dan güneyli rüzgarlarla uzun menzilli toz taşınımına bağlı olabileceği belirtilmiştir. Farklı bölgelerde yer alan şehirlerde kış aylarında evsel ısınmanın neden olduğu yüksek PM₁₀ konsantrasyonu, buna karşılık ısınmanın olmadığı yaz aylarında düşük PM₁₀ konsantrasyonu şeklinde mevsimsel değişim gerçekleşmektedir (Erdun ve diğ, 2015).

Sanayileşme, evsel ısınma gibi araç emisyonları da hava kalitesini etkileyen başka bir insan kaynaklı faktördür. Araç emisyonlarının yarattığı etkileri sadece şehir içi ile kısıtlamak da mümkün değildir. Şehirler arası yollar da hava kalitesi incelemelerine dahil edilen alanlardır. Hava kirliliği modellerinin başarısının artırılması amacıyla, yol haritalarının diğer verilerle birleştirilerek emisyon envanteri oluşturulmasına yönelik çalışma mevcuttur (Toros ve Bağış, 2017).

Bir bölgede hava kirleticisinin varlığı, konsantrasyon düzeyi ve kaynaklarına yönelik yapılacak çalışmalar, kirleticinin olumsuz etkilerinin önlenmesi ya da zamanında gerekli tedbirlerin alınması için önemlidir. Türkiye'nin 16 büyük şehrinin 2010-2012 yılları arasında kirletici seviyelerini istatistiksel olarak inceleyen bir çalışmanın sonuçlarına göre, şehirlerin PM₁₀, NO₂ ve O₃ konsantrasyon seviyelerinin Avrupa Birliği standartlarına uygun olmadığı ortaya çıkmıştır (Toros ve diğ, 2013). Bu çalışmalar kamu otoritelerinin sorumluluk alanlarıyla ilgili durumu değerlendirip, gerekli aksiyonları almaya yönlendirmeleri bakımından çok önemlidir.

Dünyayı 2020 yılından itibaren etkisi altına alan COVID-19 salgını, insan faaliyetlerinin bölgesel ve küresel hava kalitesine olan etkilerinin değerlendirilmesinde araştırmacılara farklı kapılar açmıştır. Salgınla birlikte dış ortamı ilgilendiren aktivitelerin kısıtlanması, atmosferik kompozisyonadaki kirletici miktarlarında önemli değişimler gözlenmesine neden olmuştur. İnsan faktörü minimize edildiğinde hava kalitesindeki iyileşmeyi değerlendiren çok sayıda çalışma

mevcuttur. Ülkemizde de COVID-19 kısıtlamaları döneminde İstanbul, Kocaeli, Bursa, Konya ve Trabzon gibi büyük şehirlerin hava kalitesi değişimini inceleyen çalışmalar yapılmıştır (Kılıç ve diğ, 2020; Hojamkulyyeva ve diğ, 2020; Yiğiter ve diğ, 2020; Shayia ve diğ, 2021; Kara ve diğ, 2020).

Kötü hava kalitesi küresel ölçekte çevresel sağlığı tehdit etmesi bakımından en büyük risk faktörüdür (Lelieveld, 2020). Hava kirliliğinin özellikle insan sağlığı üzerinde yarattığı olumsuz sonuçlar ülkelerin kamu otoritelerini yasal düzenlemeler yapmaya zorlamıştır. Örneğin, Birleşik Krallık'ta 19. Yüzyıl boyunca kömür kullanımından kaynaklı emisyonların kontrol edilmesine yönelik çok sayıda yasal ve idari düzenleme yürürlüğe girmiştir. Etkileri sınırlı olsa da 1847 yılındaki Kasabaları İyileştirme Maddeleri Kanunu (Towns Improvement Clauses Act), 1853'teki Duman Rahatsızlığını Önleme (Metropol) Yasası (Smoke Nuisance Abatement (Metropolis) Act), 1863'teki Alkali Yasası bu düzenlemelere örnektir (Flick, 1980). Endüstriyel işlemlerden kaynaklı zararlı atmosferik emisyonlarla ilgili düzenlemelere yer veren Alkali Yasası, sorunu doğrudan ele alan çığır açıcı bir mevzuat olarak değerlendirilmiştir (Wilde, 2010).

Amerika Birleşik Devletleri'nde 1948 yılında meydana gelen Donora olayı ve diğer çevresel felaketler akabinde, 1955 yılında Kongre, Hava Kirliliği Kontrol Yasasını kabul etmiştir. Bu yasa hava emisyon standartlarının ve diğer kontrol araçlarının uygulanmasını sağlayan 1963, 1970 ve 1990 Temiz Hava Yasaları (Clean Air Act) ile güçlendirilmiş ve ABD hava kalitesinde ve dolayısıyla halk sağlığı konusunda ciddi iyileşmeler sağlanmıştır (Helfand ve diğ, 2001).

İngiltere'de, Büyük Londra Sisi sonrası, 1956'da Temiz Hava Yasası (Clean Air Act) adıyla bir yasal düzenleme yapılmıştır. Bu doğrultuda hava kalitesini düzenleyici standartlar belirlenmiş ve yerel otoritelere dumansız bölgeler oluşturma yetkisi verilmiştir (Longhurst ve diğ, 2016).

Temel rollerinden biri hava kalitesi ve sağlık konularında üye devletlerine yönelik kılavuzlar ve öneriler sunmak olan Dünya Sağlık Örgütü (WHO), ilk kez 1958 yılında hava kirliliğini bir sağlık riski faktörü olarak kabul etmiştir (Lehtomäki ve diğ, 2018). Takip eden yıllarda teknik raporlar hazırlamaya devam eden örgütün 1987 yılında yayınladığı "Avrupa için Hava Kalitesi Rehberleri" raporu, Avrupa'daki hava kalitesi sorunlarını ele alan önemli bir kilometre taşı olarak kabul edilmektedir. Raporda

özetle, belli başlı kirleticilerin sınır seviyeleri belirlenmiş, sağlık üzerine etkilerinden bahsedilmiş ve çeşitli öneriler sunulmuştur. Dünya Sağlık Örgütü, 2021 yılına ait Küresel Hava Kalitesi Rehberi'nde kirleticilerin daha önce belirlenen sınır seviyelerin altındaki konsantrasyonlarda dahi insan sağlığını olumsuz etkilediğine dair güçlü kanıtlar bulunduğunu ifade etmiştir. Bu tespit doğrultusunda kirletici eşik değerleri daha aşağıya çekilmiştir (WHO Global Air Quality Guidelines, 2021).

Ülkemizde hava kirliliğinin etkilerinin azaltılmasına ve önlenmesine yönelik çok sayıda yasal ve idari düzenlemeler yapılmıştır. Bu düzenlemelerin yanı sıra Birleşmiş Milletler Sürdürülebilir Kalkınma Amaçları doğrultusunda hazırlanan T.C. Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığının 2020 yılı Sürdürülebilir Kalkınma Amaçları ve Göstergeleri Raporu'nda hava kirliliğinin azaltılması hususunda farklı hedeflerden bahsedilmektedir. Örneğin ülkemizin “Sağlıklı ve kaliteli yaşamı her yaşta güvence altına almak” başlıklı 3 no.lu Sürdürülebilir Kalkınma Amacı (SKA3) hedeflerinden biri (Hedef 3.9), 2030 yılına kadar hava kirliliğinden kaynaklanan ölüm ve hastalıkların kayda değer miktarda azaltılmasıdır. Raporun “Şehirleri ve insan yerleşimlerini kapsayıcı, güvenli, dayanıklı ve sürdürülebilir kılmak” başlıklı 11 no.lu Sürdürülebilir Kalkınma Amacı (SKA11) doğrultusunda ise 2030 yılına kadar hava kalitesine özel önem verilmesi ve kentlerde kişi başına düşen olumsuz çevresel etkilerin azaltılması hedeflenmiştir (Hedef 11.7). Rapordaki “Sürdürülebilir üretim ve tüketim kalıplarını sağlamak” başlıklı 12 no.lu amaç (SKA12) kapsamında kimyasalların ve tüm atıkların yaşam döngüsü boyunca insan sağlığına ve çevreye olan zararlı etkilerini en aza indirebilmek için diğer ortamlar gibi havaya salınımlarını da önemli miktarda azaltma hedefinden (Hedef 12.4) bahsedilmiştir (Sürdürülebilir Kalkınma Amaçları ve Göstergeleri Raporu, 2020).

Dünya Sağlık Örgütü (WHO) 6 farklı kirletici ($PM_{2.5}$ ve PM_{10} , ozon (O_3), nitrojen dioksit (NO_2), sülfür dioksit (SO_2), karbon monoksit (CO)) için hava kalitesi sınıflandırması yapmıştır. Her bir kirletici için kirleticiye özgü zaman periyotlarında eşik değerler belirlenmiştir. Hava kirleticilerinin eşik değerlerin altında tutulması hedef olmakla birlikte, bu değerlerin hiçbir zaman aşamayacağını iddia etmek de mümkün değildir. Bu nedenle toplumu ve kamu otoritelerini bilgilendirmeye ve tedbir almaya yönlendirecek tahmin ve erken uyarı sistemlerinin kurulması önemlidir.

Bursa ili coğrafik konumu, büyüklüğü, nüfusu ve sanayileşmesi bakımından hava kalitesi dikkatle izlenmesi gereken bir şehirdir. Meteorolojik koşullar, hava kirliliği

olaylarında emisyon kontrolü üzerinde doğrudan bir etkiye sahiptir (Pérez ve diğ, 2021). Bu nedenle ilin kirletici kaynakları ile ilgili çalışılırken meteorolojik parametrelerle birlikte değerlendirme yapılması önemlidir. Bu çalışmada Bursa ili için meteorolojik koşullar doğrultusunda makine öğrenmesi yöntemleri ile partiküler madde seviyesinin tahmini incelenmiştir.

1.1 Tezin Amacı

Hava kirliliğinin özellikle insan sağlığı üzerindeki olumsuz etkileri bilim insanlarını kirliliğin tespiti, etkileri ve eğilimleri konusunda araştırmaya sevk etmiştir. Bu doğrultuda hava kalitesi izleme istasyonlarının sayısı hızla artmış, kimyasal taşınım modelleri ve uydu verileri ile uzaktan algılama yöntemleri geliştirilmiştir. Örneğin, ülkemizde İzmir için multispektral uydu görüntülerinin yansıma değerleri kullanılarak PM₁₀ verilerinin belirlenmesine yönelik çalışma yapılmış, sonuçları tatmin edici ve ileriki çalışmalar için umut verici bulunmuştur (Ozelkan ve diğ, 2015).

Bu yöntemlerin yanı sıra yeni makine öğrenmesi (machine learning), derin öğrenme (deep learning) gibi yapay zeka uygulamaları ile kirletici seviyesinin kısa, orta ve uzun vadeli tahminine yönelik çalışmalar hızlanmıştır. Bu çalışmanın amacı meteorolojik koşullar doğrultusunda Bursa ilinde hava kalitesini etkileyen faktörlerden partiküler madde konsantrasyonlarının tahmin edilmesidir. Bir sanayi şehri olan Bursa'da için oluşturulacak etkin ve sürdürülebilir partiküler madde tahmin modeli tüm paydaşların erken süreçte olumsuz durumlarda gerekli tedbirlerin almasını sağlayabilecektir. Model, partiküler madde seviyesinin etkin şekilde izlenmesi ve kontrolünde faydalı olabilecektir.

1.2 Literatür Araştırması

Hava kalitesinin insan sağlığı ve çevre üzerine olumsuz etkileri bilim insanlarını ve araştırmacıları kirletici parametrelerin çeşitli zaman periyotlarında tahminine yöneltmiştir. Başarılı hava kalitesi tahmini kamu otoritelerinin toplum ve çevre ile ilgili tedbirleri zamanında almasını, halkın erken uyarılmasını, böylelikle kirletici kaynaklı zararlardan en az düzeyde etkilenilmesini sağlayabilecektir. Hava kirliliği/kalitesi tahmini, ileri veri işleme yöntemlerinin gündeme gelmesiyle birlikte

daha az bileşenli istatistiksel tahminden, daha çok değişkeni işleyen yapay zeka uygulamalarına geniş bir yelpazeyi içine alan bir araştırma konusudur.

1.2.1 Eksik kirletici verisinin tamamlanmasına yönelik çalışmalar

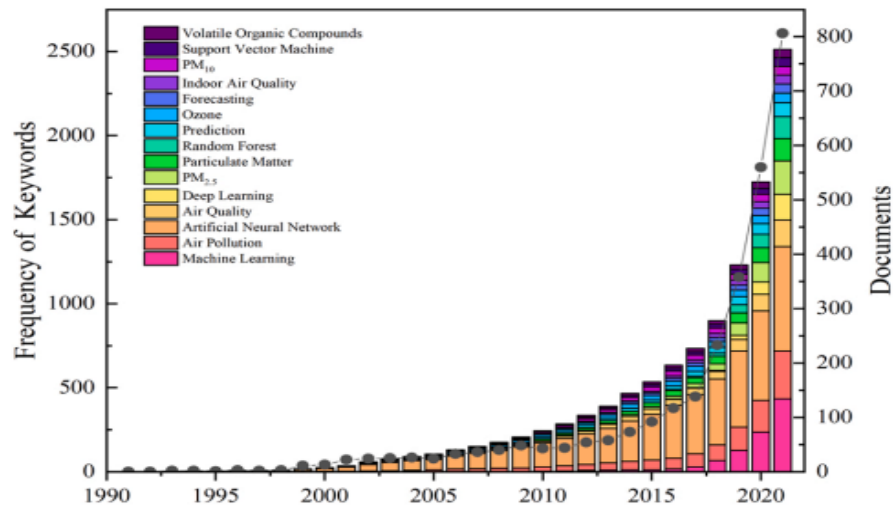
Tahmine yönelik çalışmalarda, araştırmacıları bekleyen en büyük zorluklardan biri, ana girdi konumundaki kirletici ölçüm verilerinde eksikliklerin olmasıdır. Bu nedenle eksik kirletici verilerinin gerçeğe en yakın haliyle tamamlanmasına yönelik çok sayıda yöntem denenmiştir. Yöntemlerden biri, eksik kayıtların tekli ve çoklu imputasyon prosedürleri kullanılarak doldurulmasını ve ardından hava kirliliği seviyesini bulmak için çeşitli sınıflandırıcılardan oluşan bir topluluğun kullanılmasını içeren topluluk-imputasyon-sınıflandırma çerçevelerinin kullanılmasıdır (Narasimhan ve Vanitha 2023). Diğer bir yöntem, izleme istasyonları arasındaki mekansal ve zamansal benzerlikleri göz önünde bulundurarak eksik hava kirletici konsantrasyon verilerini tamamlayan çok değişkenli zaman serisi kümelemesine dayanmaktadır (Alahamade ve diğ, 2021). NARX sinir ağları gibi sinir ağları da hava kirletici sensörlerindeki eksik verileri düzeltmek için kullanılabilmektedir (Calle ve diğ, 2020). Hava kirletici veri setlerindeki eksik değerleri düzeltmek için kullanılan yöntemlerden bir diğeri, düşük sıralı matris tamamlama (LRMC) yöntemidir (Liu ve diğ, 2020). Diğer teknikler arasında ortalama imputasyon, spline interpolasyon, basit hareketli ortalama, üstel ağırlıklı hareketli ortalama ve yapısal zaman serileri ile otoregresif entegre hareketli ortalama modelleri üzerinde Kalman düzeltmesi yer almaktadır (Wijesekara ve Liyanage, 2020). Bir çalışmada, ortalama imputasyon, koşullu ortalama imputasyon, K-En Yakın Komşu imputasyonu, çoklu imputasyon ve Bayesian Temel Bileşen Analizi imputasyonu dahil olmak üzere beş imputasyon yöntemi karşılaştırılmış ve veri setini tamlik, hatalar ve yanlılık açısından iyi bir performansla yeniden yapılandırabildikleri bulunmuştur (Quinteros ve diğ, 2019). Bir başka çalışma, izleme istasyonlarında ve zaman serilerinde eksik kirletici verilerini tahmin etmek için farklı yaklaşımlar deneyerek hava kalitesi değerlendirmesindeki belirsizliği azaltmayı amaçlamıştır (Alahamade ve diğ, 2020). Ek olarak, bir çalışmada önceki ve sonraki verinin ortalamasını alma (Mean Top Bottom ya da Mean Before After), Doğrusal Regresyon, Çoklu İmputasyon ve En Yakın Komşu dahil olmak üzere dört imputasyon yöntemi değerlendirilmiş ve eksik verinin bir üstündeki veri ile bir altındaki verinin ortalamasının alındığı yöntemin (Mean Top Bottom ya da Mean Before After) hava kirletici verilerindeki eksik değerleri doldurmak için en uygun yöntem olduğu

belirtilmiştir (Zakaria ve Noor, 2018). Çin’de yapılan bir çalışmada $PM_{2.5}$ konsantrasyonlarının tahmininde kullanılan uydu aerosol alımlarındaki (aerosol optik derinliği) veri eksikliği sorununu gidermek için XGBoost yöntemi kullanılmıştır. Sonuçlar farklı imputasyon yöntemleri ile karşılaştırılmış, XGBoost eksik veriyi tamamlamada diğer yöntemlere göre başarılı bulunmuştur (Chen ve diğ, 2020). Bu yöntemler, sürekli ortam hava kalitesi izleme istasyonlarından elde edilen hava kirleticisi verilerindeki eksik değerleri doldurmak için kullanılmaktadır. İmputasyon yöntemlerinin performansı, Ortalama Mutlak Hata, Ortalama Karesel Hatanın Kökü, Belirleme Katsayısı ve Uyum İndeksi gibi ölçütler kullanılarak değerlendirilebilmektedir (Hadeed ve diğ, 2020).

1.2.2 Hava kirleticilerinin tahminine yönelik çalışmalar

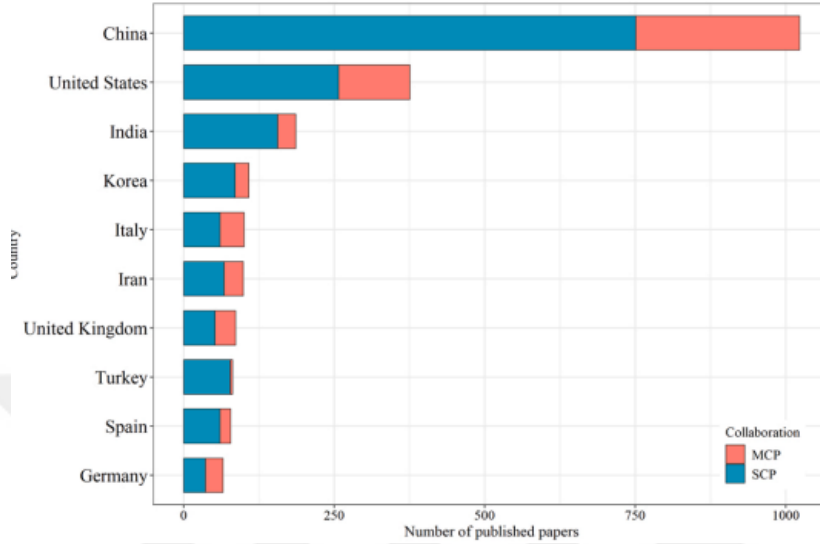
Araştırmacılar tam bir kirleticisi veri seti kullanarak çok farklı yöntemlerle kısa, orta ve uzun vadede hava kirliliği seviyesini ya da hava kalitesi indeksini tahmin etmeye çalışmışlardır. Özellikle bilgisayarların veri işleme kapasitesinin artması, istatistiksel yöntemlerin yanında yapay zeka teknolojilerinin devreye girmesi çok boyutlu veri ile daha hızlı, daha isabetli sonuçlar elde edilmesini mümkün kılmıştır.

Dünya çapında 1990-2021 yılları arasında hava kirliliği ve makine öğrenmesi ile ilgili kavramların anahtar sözcük olarak yer aldığı makaleleri konu alan bir tarama çalışması yapılmıştır. Bu çalışmaya göre, makine öğrenmesi yöntemlerinin kullanıldığı hava kirliliği araştırmalarında Şekil 1.1’de görülebileceği gibi 2017’den sonra neredeyse patlama yaşanmıştır (Li ve diğ, 2023).



Şekil 1.1 : Anahtar kelimelere göre yayın sıklığı (Li ve diğ, 2023).

Aynı çalışmaya göre, Türkiye yayın sayısına göre İngiltere'den sonra 8. sırada yer almaktadır (Şekil 1.2). Şekildeki MCP (Multiple Country Publications) ülkeler arası işbirliği ile hazırlanan yayınları, SCP (Single Country Publications) ise yazarları aynı ülkeden olan yayınları ifade etmektedir.



Şekil 1.2 : MCP/SCP dağılımı (Li ve diğ, 2023).

Hava kirliliği tahmininde yapay sinir ağlarının kullanıldığı ilk araştırma çalışması 1993 yılında yayınlanmıştır. Araştırmada Slovenya'nın en büyük termik santralinin çevresindeki SO₂ kirliliği yapay sinir ağlarıyla tahmin edilmeye çalışılmıştır (Boznar ve diğ, 1993).

Çalışmalarda bazen kirleticilerin biri ya da birkaçının ya da tüm kirleticileri temsil etmek üzere hava kalitesi tahmini hedeflenmektedir. Jojova ve Trivodaliev 2021 yılında Üsküp'te derin öğrenme yöntemleri ile gelecek saat için PM₁₀ tahminini konu alan bir çalışma hazırlamıştır. Evrişimli Sinir Ağı (CNN) ve Uzun Kısa Dönem Bellek (LSTM) ağına dayanan model diğer klasik makine öğrenimi modellerinden (lineer regresyon (LR), rastgele orman (RF), destek vektör makineleri (SVM)) daha iyi performans göstermiştir (Jojova ve Trivodaliev, 2021).

Başka bir çalışmada veri madenciliği yöntemlerini kullanarak hava kirliliğini tahmin etmeye yönelik iki husus tartışılmıştır. İlk aşamada, tahmini etkileyen en önemli özellikler (atmosferik parametreler, mevsim, haftanın günü vb.) tespit edilmiştir. İkinci aşamada, iki farklı yöntem denenmiştir. Birinci yöntemde potansiyel özellik veri seti üç ayrı sinir ağında (çok katmanlı algılayıcı (MLP), radyal tabanlı fonksiyon (RBF), destek vektör makinesi (SVM)) kullanılarak tahmin sonuçları elde edilmiştir. Üç sinir

ağının sonuçları da bir sonraki günün tahmini için ağırlıklı ortalama ve rastgele orman (RF) yöntemi ile birleştirilmiştir. İkinci yöntemde ise rastgele orman (RF) hem tahmin edici hem de bütünleştirici olarak kullanılmıştır. Araştırma sonucunda kirleticiyi etkileyen en önemli özelliklere doğrudan rastgele orman (RF) yönteminin uygulanmasının daha iyi sonuçlar verdiği tespit edilmiştir (Siwek ve Osowski, 2016).

Çin'in Pekin şehri için $PM_{2.5}$ ölçüm verileri ve ERA5 reanaliz veri setinden elde edilen meteoroloji verilerinden rastgele orman (RF) ve geri beslemeli yapay sinir ağı (BPNN) kullanılarak $PM_{2.5}$ konsantrasyonu tahmin edilmeye çalışılmıştır. Her ne kadar her iki yöntem de yüksek $PM_{2.5}$ konsantrasyonları için kötü performans gösterse de, diğer seviyelerde ikisi de konsantrasyonu iyi tahmin edebilmektedir. Her halükarda RF yapay sinir ağı, BPNN'den daha başarılı sonuçlar vermiştir (Wang ve diğ, 2023).

Çin'in Qingdao şehri için yapılan çalışmada ise K-Means kümeleme ve derin sinir ağı kullanarak hava kalitesi tahmini için bir hibrit model önerilmiştir. Modelde, Qingdao şehrinin iki yıllık rüzgar hızı ve yönü, en yüksek ve en düşük sıcaklık gibi meteorolojik verileri ile $PM_{2.5}$, PM_{10} , SO_2 , CO, NO_2 ve O_3 'ten oluşan kirletici verisi kullanılmıştır. Öncelikle meteorolojik veriler K-Means kümeleme algoritması ile mevsimlere göre dört kategoriye ayrılmıştır. Veriler çift yönlü uzun kısa süreli bellek (bidirectional LSTM) ve tam bağlı katman (FC) sinir ağlarından oluşan bir hibrit modelde (BLFC) işlenmiştir. Önerilen hibrit modelin, diğer algoritma modellerine göre daha yüksek hassasiyetle hava kalitesi tahmini yaptığı sonucuna ulaşılmıştır (Ao ve diğ, 2019).

Başka bir makalede, Kaliforniya için destek vektör regresyonu (SVR) kullanılarak kirletici konsantrasyonlarını ve hava kalitesi indeksini (HKİ) tahmin etmek üzere bir model geliştirilmiştir. Çalışmada meteorolojik parametreler olarak sıcaklık, bağıl nem ve rüzgar hızı kullanılmıştır. Sonuçlar, oluşturulan modelle saatlik O_3 , CO ve SO_2 konsantrasyonlarının ve HKİ'nin başarılı şekilde tahmin edildiğini göstermiştir (Castelli ve diğ, 2020).

Kocaeli ve İstanbul'da iki istasyonda sıcaklık, rüzgar hızı, bağıl nem ve hava basıncı bilgileri de kullanarak rastgele orman (RF) yöntemiyle PM_{10} ve CO tahmini denenmiş ve sonuçlar diğer yapay sinir ağlarına göre daha başarılı bulunmuştur (Altınçöp ve Oktay, 2018).

Sarkar ve diğ. (2022), NO, NO_2 , NO_x , SO_2 , CO, O_3 konsantrasyonunun tahmini için sıcaklık, rüzgar hızı, bağıl nem ve güneş radyasyonu bilgileri de girdi olarak kullanan

bir hibrit model geliřtirmiřtir. LSTM-GRU (uzun kısa süreli bellek- geitli yinelenen birim) adı verilen model sonuçları lineer regresyon, K-en yakın komřu, destek vektör makinesi (SVM), uzun kısa süreli bellek (LSTM), geitli yinelenen birim (GRU) modelleri ile karşılařtırılmıřtır. Karşılařtırma sonuçlarına göre hibrit model diğerk modellerle oranla daha başarılı bulunmuřtur.

Tayvan'da $PM_{2.5}$ konsantrasyonunu tahmin etmek için, bir yıl boyunca 77 hava kalitesi izleme istasyonu ve 580 hava durumu istasyonundan alınan veriler Gradient Boosting makine öğrenmesi yöntemiyle oluşturulan modele girdi olarak verilmiřtir. Gemiř $PM_{2.5}$ kayıtları ve komřu meteoroloji istasyonlarının verileri ile eğitilen tahmin modelinin, 24 saatlik tahmin için çoğunlukla iyi sonuç verdiğı belirlenmiřtir (Lee ve diğ, 2020).

Yeni Zelanda'da $PM_{2.5}$ konsantrasyonunda pik olup olmayacağını doğru tahmin etmek için Gradient Boosting yöntemi kullanılmıř ve tahminde %80-90 oranında model doğrulukları bulunmuřtur (Miskell ve diğ, 2019).

Çin'in Lanzhou kentinde 2013-2020 yılları arasında günlük hava kirletici konsantrasyonu ve meteorolojik parametreler kullanılarak yapılan çalışmada, Gradient Boosting yöntemiyle oluşturulan modelin günlük $PM_{2.5}$ konsantrasyonunu tahmin etmek, Generalized Additive Model (GAM) olarak adlandırılan yöntemle göre daha hassas olduğı, GAM'ın ise uzun vadeli trend analizi için daha uygun olduğı görölmüřtür (Cheng ve diğ, 2021).

Çin'in Şanghai şehrinde hava kirliliğı indeksi için meteorolojik tahmin verilerinin de girdi olarak kullanıldığı ve bir sonraki günün ortalama indeks değıerlerinin çıktısını veren MultiLayer Perceptron (MLP) modeli oluşturulmuřtur. Çalışmayı konu alan makalede geliřtirilen algoritmanın insan müdahalesi olmadan çalışması nedeniyle Çin'in farklı meteorolojik kořullara sahip birkaç şehirde kullanıldığı ifade edilmiřtir (Jiang ve diğ, 2004).

Polonya, Krakow'da eşik değıerleri sıklıkla aşan PM_{10} konsantrasyonlarının tahmini için MLP yöntemi kullanılmıř, sonuçlar MLP ağılarının performansının tatmin edici olduğunu göstermiřtir (Pawul ve Śliwka, 2016).

1.2.3 Bursa ilini konu alan hava kirliliği çalışmaları

Türkiye'nin nüfus ve sanayileşme bakımından en büyük şehirlerinden olan Bursa'nın hava kalitesine yönelik farklı disiplinlerden bilim insanları tarafından araştırma çalışmaları yapılmıştır. Garipağaoğlu ve Duman'ın çalışmasında Bursa'nın 1990 ve 2016 yılları arası kirletici konsantrasyonlarındaki değişim incelenmiştir. Çalışmaya göre, evsel ısınmada fosil yakıt kullanımından kaynaklı partiküler madde konsantrasyonu 1990-1999 yılları arasında yüksek seyrederken, yakıt kalitesine yönelik kontroller sayesinde 2000-2008 yılları arasında belirgin ölçüde düşüş gözlenmiştir. Ancak 2008-2016 yılları arasında sanayi bölgelerinin yoğunlaşması, şehirleşmenin endüstri bölgelerini de içine alacak şekilde genişlemesi, motorlu taşıt sayısındaki artış nedeniyle konsantrasyonda yeniden artış tespit edilmiştir. Yakıt kullanımına yönelik tedbirlerin artırılması ile SO₂ konsantrasyonunda yıllar itibarıyla ciddi düşüş gözlenmiştir (Garipağaoğlu ve Duman, 2017). Çalışkan ve diğ (2020), şehrin 2013-2018 yıllarına ait beş yıllık kirletici seviyelerini inceledikleri çalışmalarında, hava kalitesini etkileyen en önemli kirleticilerin PM₁₀ ve NO_x, en önemli kirletici kaynaklarının ise sırasıyla sanayi, ulaşım ve kontrolsüz yakma olduğunu tespit etmişlerdir. Bursa'nın 2019 ve 2020 yılı hava kirliliği verilerini inceleyen bir başka çalışmada, bir istasyonda trafik sıkışıklığı ve/veya kömürle ısınmadan kaynaklı olarak zaman zaman NO₂ değerlerinin yıllık sınırı aştığı, bazı istasyonlarda ise PM_{2.5} ve PM₁₀ konsantrasyonlarının limit değerlerin üzerine çıktığı tespit edilmiştir (Özkayalar ve diğ, 2020).

Bursa ilinde O₃ kirletici konsantrasyonu tahminine yönelik tez çalışmasında iki istasyona ait kirletici verisi (PM_{2.5}, SO₂, NO, NO₂, NO_x) ile meteorolojik parametreler olarak hava sıcaklığı, rüzgar yönü, rüzgar hızı, bağıl nem, hava basıncı verileri kullanılmıştır. Tahmin için rastgele orman (RF), karar ağacı (decision tree-DT), destek vektör makinesi (SVM), K-en yakın komşu ve çok katmanlı algılayıcı (MLP) yöntemleri kullanılmıştır. Sonuçlar değerlendirilmiş ve O₃ konsantrasyonunu tahmin etmede en başarılı yöntemin rastgele orman (RF) ve en başarısız yöntemin ise destek vektör makinesi (SVM) olduğu tespit edilmiştir (Güven, 2022).



2. HAVA KALİTESİNİ ETKİLEYEN FAKTÖRLER

2.1 Coğrafi Koşullar

Coğrafya, hava kalitesini belirlemede kilit bir rol oynamaktadır. Örneğin, sahil bölgeleri, kirleticilerin deniz esintileri tarafından dağılması nedeniyle daha iyi hava kalitesine sahip olma eğilimindedir. Bunun aksine, karasal bölgelerde hava akımını engelleyen yüzey şekilleri nedeniyle kirleticiler belirli bir yerde birikebilmekte, dolayısıyla o bölgede hava kalitesi bozulabilmektedir. Yüksek rakımlı bölgeler, şehirleşme ve sanayileşme faaliyetleri için daha az tercih edilmeleri nedeniyle, hava kirliliğine daha az maruz kalmaktadır. Bununla birlikte, vadilerin meteorolojik faktörlerin etkisiyle kirleticiler maddelerin birikmesine ortam hazırladığı da göz önünde bulundurulmalıdır. Topoğrafik olarak bir çanak özelliği gösteren veya hakim rüzgar yönüne dik doğrultuda uzanan bir oluk biçimindeki çukurluk içinde bulunan yerleşim alanlarında kirlilik daha fazla etkili olmaktadır (Şahin, 1989). Deniz ve okyanuslardan rüzgarla taşınan deniz tuzu ile çöl bölgelerinden taşınan tozlar da doğanın kendine has partikül kaynakları olarak nitelendirilebilirler. Özellikle kuvvetli rüzgarlarla uzun mesafe taşınan çöl tozları hava kirliliğine ve solunması durumunda da önemli sağlık sorunlarına neden olabilmektedir.

2.2 Antropojenik Kaynaklar

Sanayi Devriminden sonra hızla artan endüstriyel imalat, kimyasal üretim ve madencilik gibi faaliyetler hava kalitesinin bozulmasının ana nedenleri olmuştur. Üretim süreçlerinde partiküler madde ve diğer zararlı kirleticilerin atmosfere serbestçe salınmasını engelleyecek tedbirlerin alınmaması önemli sağlık sorunlarına yol açmaktadır. Ulaşım başka bir kirleticiler kaynağıdır. Araçlardan salınan karbon monoksit (CO), azot oksit (NO_x) ve partiküler maddelerin (PM) çevreye olumsuz etkileri bulunmaktadır. Enerji üretimi aşamasında fosil yakıtların yakılması, tarım uygulamaları esnasında amonyak, metan ve diğer kirleticilerin atmosfere salınması hava kalitesini olumsuz yönde etkilemektedir. Şehirleşmenin, trafik, ısınma, endüstrileşme gibi nedenlerle oluşan kirleticiler birikimine yol açacak şekilde gelişmesi

(hava akımlarını engelleyen çarpık ve sıkışık kentleşme) hava kalitesi açısından önemli bir sorundur. Örneğin, Çin'in Hangzhou şehrinin on yıllık kentsel gelişimi (2000-2010) incelenmiştir. Çalışma sonucunda, şehirleşme ile birlikte meteorolojik ortamdaki değişikliklerin (ortalama kentsel rüzgar hızın azalması, kentsel ısı adası etkisinin artması, ortalama kentsel bağıl nemin azalması) kentsel atmosferin difüzyon kapasitesini azalttığı belirlenmiştir. Azalan difüzyon kapasitesinin de ortamdaki kirletici konsantrasyonlarının artmasına neden olduğu ortaya konmuştur (Liu ve diğ, 2015).

2.3 Meteorolojik Koşullar

Kirleticilerin ve meteorolojik koşulların hava kirliliğinin oluşumundaki rolleri arasında temel bir fark bulunmaktadır. Havaya salınan kirletici miktarı tedbirlerle kontrol edilebilirken, meteorolojik koşullar kontrol edilememektedir. Bu nedenle hava kirleticilerinin eşik değerleri aşmamasına yönelik planlama yaparken meteorolojik parametreler mutlaka göz önünde bulundurulmalıdır (Pérez ve diğ, 2020).

Meteorolojik koşullar ile hava kirliliği arasındaki ilişki karmaşıktır. Hava sıcaklığının yüksek olduğu, güneşli bir günde hava kirleticilerinin birikimi daha belirgin olmaktadır. Yüksek rüzgar hızları, kirleticileri dağıtarak belirli bir bölgedeki konsantrasyonlarını azaltıp, hava kalitesini iyileştirebilir. Aksine yüksek hızda esen rüzgar, kirliliğin olmadığı bir bölgeye kirletici taşımaktan da sorumlu olabilir. Düşük rüzgar hızlarında birim havadaki kirleticiyi seyreltecek düzeyde yatay hareket bulunmadığından ortamda kirletici birikebilmektedir. Yağış veya yüksek nem, havada asılı bulunan partiküllere tutunarak, onların çökmesini sağlar. Hava banyosu denilebilecek bu olay havayı temizleyici ve hava kalitesini iyileştirici bir faktör olarak değerlendirilebilir. Sıcaklık terselmesinde (inversiyon) ise sıcak hava tabakası soğuk havayı yüzeye hapsederek dikey karışımı önlemektedir. Bu olay genellikle vadilerde ve havzalarda gece veya sabah erken saatlerde görülür. Kirletici kaynağı yoğun olan yerlerde bu durum ya çok yoğun bir pus halini (smog) ya da yeteri kadar nem varsa sisi oluşturur (Gönençgil ve Sungur, 1997). Londra'da 1952 yılında yüksek basınç sisteminin şehrin üzerinde kaldığı beş gün boyunca kirli havayı dar alana hapseden yoğun sisten etkilenen en az 12 bin kişi hayatını kaybetmiştir. Büyük Londra Sisi olarak kayıtlara geçen olay ölümcül sonuçlarına atfen "Öldüren Sis" olarak da anılmaktadır.

Dikey karışımı tetikleyecek bir olayın bulunmaması durumunda, terselmenin gerçekleştiği tabakada kirletici birikimi ve çökmesi meydana gelecektir. Bu da hem terselmenin şiddetine ve hem de kirleticilerin konsantrasyonuna bağlı olarak hava kalitesini olumsuz yönde etkileyecektir.

Kirletici oluşumunda etkili olan bir başka parametre ise güneş ışınlarıdır. Güneş ışınları troposferik ozon (yer yüzeyi ozonu) gibi ikincil kirleticilerin oluşumuna neden olabildiği gibi, başka kirleticileri de parçalayarak ortamdan uzaklaştırabilmektedir.

Meteorolojik parametrelerin hava kirleticileri ile ilişkisini konu alan ve Çin’de 896 istasyondan alınan veriler üzerinde yapılan bir çalışmada, istasyonların çoğunda kirletici konsantrasyonun rüzgar hızı, yağış ve bağıl nem ile önemli ölçüde negatif korelasyon, diğer yandan atmosferik basınç ile pozitif korelasyon gösterdiği tespit edilmiştir. Çalışmada, enlem arttıkça, sıcaklığın hava kirletici konsantrasyonu üzerindeki etkisinin daha belirgin hale geldiği belirtilmiştir (Liu ve diğ, 2020).

Vietnam’ın Hanoi şehrinde meteorolojik koşulların PM₁₀ konsantrasyonu üzerindeki etkisi araştırılmıştır. Sıcaklık, nem ve rüzgar hızı gibi meteorolojik faktörlerin hem kuru hem de yağışlı mevsimde PM₁₀ konsantrasyonu ile ters orantılı olduğu, PM₁₀ konsantrasyonun az yağış, düşük sıcaklık, sakin rüzgar ve düşük bağıl nemin olduğu kış aylarında daha yüksek olduğu tespit edilmiştir (Dung ve diğ, 2019).

Çin’de 2005-2012 yıllarını kapsayan bir çalışmada ise 67 şehrin meteorolojik parametreleriyle hava kirliliği verileri arasındaki ilişki irdelenmiş, bağıl nem, rüzgar hızı ve sıcaklığın kirletici dağılımı üzerinde baskın faktörler olduğu görülmüştür. Meteorolojik parametrelerin her biri ile hava kalitesi indeksi arasındaki ilişkinin bazı dönemlerde negatif, bazı dönemlerde ise pozitif korelasyon göstermesi mevsimsellik etkisine atfedilmiştir (Song ve diğ, 2019).

Kirleticilerin meteorolojik parametrelerle etkileşimi sonucu ikincil kirletici (aerosol) oluşumu meydana gelebilmektedir. Örneğin, Kore Cumhuriyeti için yapılan bir çalışmada, nemin hava kirliliği sorunu üzerindeki etkisi niceliksel olarak değerlendirilmiş ve ıslak bir ortamda (bağıl nem $\geq$ %60) ikincil inorganik aerosollerin sulu faz reaksiyonun daha aktif hale geldiği, dolayısıyla aerosolün ikincil üretiminde nemin sıcaklık kadar önemli olduğu vurgulanmıştır (Choi ve diğ, 2023).

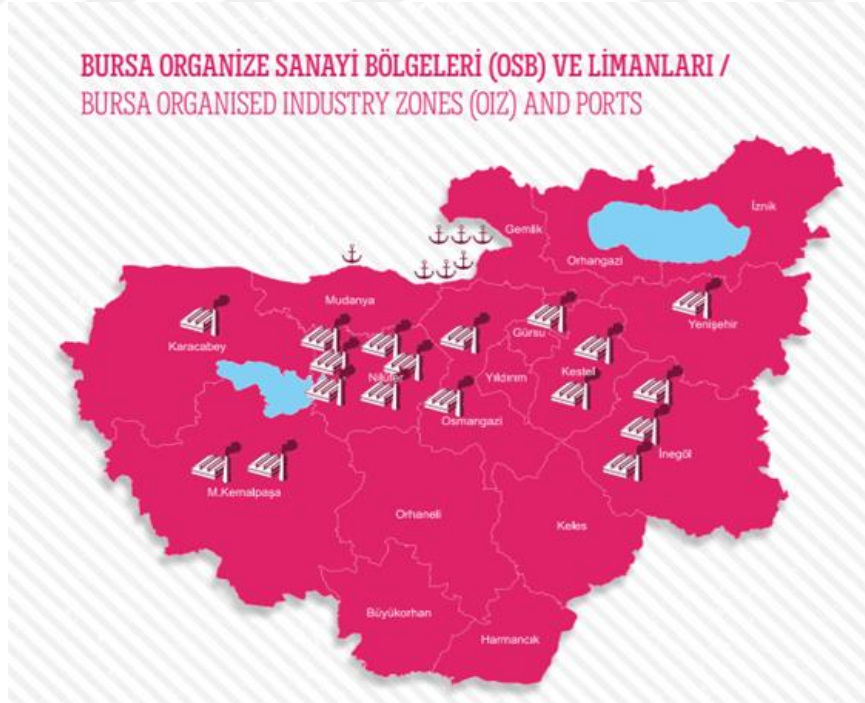
Hava kirliliği çalışmalarında sıcaklık, nem, basınç, rüzgar hızı/yönü gibi temel meteorolojik parametrelerin yanında bu parametrelerin birinin ya da birkaçının ya da

tümünün bir arada etkili olduđu meteorolojik olaylar da irdelenmelidir. Bu meteorolojik olaylar makro ölçekte sinoptik paternler, soğuk cepheler, orta ölçekte deniz ve kara meltemleri, mikro ölçekte atmosferik sınır tabaka (AST) yüksekliğı gibi örneklendirilebilir (Pérez ve diğ, 2020). Çalışmadaki meteorolojik parametrelerden biri olarak seçilen atmosferik sınır tabaka (AST) dünya yüzeyiyle temas halinde olup, yüzeyin özellikleri tarafından etkilenen katmanı ifade etmektedir. AST, hava kirleticilerinin dağılımı ve taşınması, hava kütlelerinin karıştırılması, bulutların ve yağışın oluşumu dahil olmak üzere çeşitli atmosferik süreçlerde kritik bir rol oynar. Düşük AST durumunda, kirleticiler daha dar bir alana hapsolmaktadır. Dolayısıyla birim hacimdeki kirletici konsantrasyonu daha yüksektir. Konuyla ilgili Çin’de yapılan bir çalışmada, yüksek PM_{2.5} konsantrasyonu ve düşük AST gözlemlenen bölgelerde, ikisi arasında doğrusal bir ilişki yerine üstel negatif korelasyon hesaplanmıştır (Lou ve diğ, 2019).

3. BURSA İLİ'NİN COĞRAFİK KONUMU, İKLİMİ VE HAVA KALİTESİ

Bursa ili Marmara Bölgesi'nin güney doğu kısmında yer almakta olup, yüzölçümü 11 bin 027 m²'dir. Denizden yüksekliği 155 metre olan şehrin yaklaşık % 35'i dağlar, % 17'si ovalarla kaplıdır. Doğu-batı şeklinde uzanan dağlar, nüfusun ağırlıklı olarak kuzey kısımdaki ovada yoğunlaşmasına neden olmuştur. Bursa kentinin kuruluş yeri seçimini ve yayılmasını etkileyen iki yer şekli Uludağ ve Bursa Ovası'dır (Şerbetçi, 2017).

Bursa, Türkiye İstatistik Kurumu (TÜİK) 2022 yılı verilerine göre, 3.194.720 kişi ile nüfus bakımından Türkiye'nin dördüncü büyük şehridir (TÜİK, 2023). Stratejik konumu nedeniyle Türkiye'nin önemli sanayi şehirlerinden biri haline gelen Bursa'da Şekil 3.1'de görüldüğü üzere on sekiz Organize Sanayi Bölgesi (OSB), bir Teknoloji Geliştirme Bölgesi ve bir Serbest Bölge bulunmaktadır (Bursa Yatırım Destek Ofisi, 2023).



Şekil 3.1 : Bursa OSB'leri ve Limanları (Bursa Yatırım Destek Ofisi, 2023).

3.1 Bursa'nın İklimi ve Coğrafi Koşulları

Bursa Akdeniz iklimi ile Karadeniz iklimi arasında karma bir iklime sahiptir. Kuzeyde Marmara Denizi'nin etkisiyle daha yumuşak iklim koşulları görülmekte iken güneyde Uludağ'a doğru yükseldikçe daha sert iklimle karşılaşmaktadır (Bacanlı ve Kargı, 2019). Uludağ, Marmara Bölgesi'nin en yüksek dağıdır. Şehrin güneyi ve kuzeydoğusunda yüksekliği 1.000 m'nin üzerinde olan dağlar yer alırken; kuzey, doğu ve batısında ovalık alanlar yer almaktadır (Öncü, 2021). Bursa şehir merkezi, yükseltilerle çevreli konumu nedeniyle, yatay hava akımlarından korunaklı durumdadır (Garipağaoğlu ve Duman, 2017). Bursa ili topraklarının yaklaşık % 35'i genellikle doğu-batı yönlü uzanan sıradağlarla kaplıdır. Bursa'da aynı isimli meteoroloji istasyonundaki ölçümler dikkate alınarak yapılan bir çalışmada genel olarak kuzeyli rüzgarların hakim olduğu, Mart, Nisan ve Mayıs aylarında ise güneyli rüzgarların etkili olduğu belirtilmiştir (Demir, 2020). Uludağ, çoğunlukla cephesel alçak basınç sistemlerinin sıcak sektörlerindeki güneybatılı havanın (Iodosun) dağı aşarak Bursa Ovası'na doğru inerken hızlanmasına ve adyabatik olarak ısınmasına neden olur. Bu ısınmanın sonucunda, Bursa ovasına doğru alçalma sırasında ısınan hava kütlesi fön etkisi yaratır. Bu durum şehirde kışın sıklıkla soba zehirlenmelerine neden olmaktadır (Kadıoğlu, 2012).

3.2 Bursa'nın Hava Kalitesi

Bursa, sanayileşme ve düzensiz, çarpık, hızlı şehirleşmeye karşılık aynı hızda tesis edilemeyen altyapı ve ulaşım hizmetleri, hava kirleticilerinin atmosfere salınımı konusunda bilgi ve özen eksikliği nedeniyle hava kirliliğinden ciddi şekilde etkilenmektedir. Bursa'nın hava kalitesini olumsuz yönde etkileyen en önemli iki kaynak sanayileşme ve ulaşımdır (Çalışkan ve diğ, 2020).

Kış aylarında artış gösteren hava kirliliğinin ise genel olarak sanayileşme ve kentleşmeden kaynaklandığı görülmüştür. Bursa'nın hızlı göç alan bir şehir olması, topografik ve coğrafi yapısı dikkate alınmadan kentleşme ile birlikte hava kirliliğini büyük ölçüde azaltan yeşil alanların hızla yok olması, sağlıklı çevre koşulları oluşturduğu tespit edilmiştir.

4. HAVA KİRLİLİĞİ TAHMİNİNDE YAPAY ZEKA YÖNTEMLERİ

Hava kirliliği tahmininde farklı modeller kullanılmaktadır. İstatistiksel modeller, sayısal modeller ve makine öğrenmesi modelleri çeşitli özellikleri nedeniyle tahmin için tercih edilmektedir. İstatistiksel modeller meteorolojik parametrelerin dinamik davranışlarını dikkate almadığından, sayısal modeller ise genellikle yüksek değişimlerin gerçekleştiği kısa vadeli tahmin için uygun bulunmamaktadır (Srivastava ve diğ, 2018). Yapay zeka başlığı altında yer alan makine öğrenmesi yöntemleri ise, çeşitli veri kaynaklarını işleme, özellik seçme, karmaşık ilişkileri modelleme, gerçek zamanlı izleme, uzamsal analiz yapma ve hava kalitesi tahminlerini yüksek doğrulukla geliştirme yeteneğine sahip tekniklerdir. Bu yöntemler, hava kirliliği seviyelerini tahmin etmek ve izlemek için daha verimli, hızlı ve hassas çözümler sunar. Aşağıda hava kirliliği ve kalitesi tahmininde sıklıkla kullanılan makine öğrenmesi yöntemlerine yer verilmiştir.

4.1 Rastgele Orman (Random Forest)

Rastgele orman yöntemi, hem sınıflandırma hem de regresyon görevleri için kullanılan güçlü ve çok yönlü bir öğrenme yöntemidir. Birden fazla bireysel karar ağacının tahminlerini birleştirerek daha doğru ve güçlü tahminler yapmak için kullanılmaktadır. Karar ağaçlarına dayanan basit, yorumlanabilir bir modeldir. Ormandaki her bir ağaç, verinin rastgele bir alt kümesi ve özelliklerin (niteliklerin) rastgele bir alt kümesi kullanılarak oluşturulur ve modele rastgelelik ve çeşitlilik ekler. Yöntem tek karar ağaçları ile ilişkilendirilen aşırı uyum sorununu azaltmaktadır. Gürültülü verileri ve aykırı değerleri iyi bir şekilde işleyebilir. Rastgele orman yöntemi, orijinal veri kümesinden rastgele seçilmiş örneklerden (tekrarla seçilmiş örnekler) çok sayıda örnek kümesi oluşturur. Karar ağacının her bölünme düğümünde rastgele özelliklerin bir alt kümesi seçilir, bu da ağaçların yüksek derecede korele olmasını engeller.

Rastgele orman yöntemi tahmin yapmak için çoğunluk oylama mekanizmasını kullanır. Her ağaç tahminde bulunur ve en fazla oyu alan sınıf son tahmin haline gelir.

Regresyon görevleri için ise, tüm ağaçların tahminlerini ortalar. Bu yöntem özellik seçimi ve veriyi anlamak için hakim özellikleri belirlemeye yardımcı olur.

Rastgele orman yöntemi yüksek doğruluk ve genelleme performansı ile bilinmektedir.

4.2 Çok Katmanlı Algılayıcı (Multilayer Perceptron)

Çok Katmanlı Algılayıcı (Multi-Layer Perceptron veya MLP), hava kirliliği çalışmalarında en yaygın kullanılan yapay sinir ağı modellerinden biridir. MLP, en az bir gizli katman (hidden layer) içeren ve genellikle sınıflandırma veya regresyon gibi görevler için kullanılan bir yapay sinir ağı türüdür. Her katmanda birçok nöron (perceptron) bulunur ve bu nöronlar veri özelliklerini işler, ağırlıkları öğrenir ve sonucu tahmin etmek için aktivasyon fonksiyonları kullanır. Derin öğrenme ve makine öğrenimi alanlarında çok kullanılan MLP, özellikle karmaşık veri işleme ve desen tanıma görevlerinde etkilidir.

4.3 Gradyan Artırma (Gradient Boosting)

Gradyan artırma, makine öğrenimi alanında yaygın olarak kullanılan ve özellikle sınıflandırma ve regresyon problemlerini çözmek için etkili olan bir topluluk yöntemidir. Bu yöntem, tahminlerini aşamalı olarak iyileştirir. İlk adımda zayıf öğrenici modeli oluşturulur. Ardından hata hesaplanır. Sonraki aşamada, bu hata üzerine bir başka zayıf öğrenici eklenir. Bu şekilde hata giderek azalır. Bir önceki aşamanın hataları en aza indirilmeye çalışılır. Gradyan (eğim) hesaplamaları, her adımda gerçek değerler ile tahminler arasındaki farkları belirler ve bu farkları minimize eder. Gradyan artırma modelin aşırı uyumdan (overfitting) kaçınması için bazı düzenlemeler içerir. Ağaçların derinliği ve öğrenme oranı gibi hiperparametreler kullanılarak modelin performansı ayarlanabilir. Bu yöntem regresyon ve sınıflandırma gibi birçok problemin çözümünde başarılıdır.

4.4 Destek Vektör Makineleri (Support Vector Machines)

Destek Vektör Makineleri (SVM), sınıflandırma ve regresyon görevleri için kullanılan güçlü bir makine öğrenimi yöntemidir. Veri noktalarını farklı sınıflara ayırmak için bu sınıfların veri noktalarını birbirinden ayıran optimal karar sınırını bulmaya çalışırken sınıflar arasındaki marjı maksimize etmeyi amaçlar. SVM, yüksek boyutlu

uzaylarda çalışabilir, aşırı uymaya karşı dirençlidir. Bellek kullanımı açısından verimlidir.

4.5 Lineer Regresyon (Linear Regression)

Lineer regresyon, bağımlı bir değişkenin (sonuç) bağımsız değişkenlerle (neden) doğrusal bir ilişkisinin modellenmesinde kullanılan istatistiksel bir yöntemdir. Bu yöntem, veri noktalarının en iyi uyan bir doğru (veya düzlem) oluşturarak bağımlı değişkenin tahmin edilmesini amaçlar. Lineer regresyon, regresyon analizlerinin temelini oluşturur ve bağımlı değişken ile bağımsız değişkenler arasındaki ilişkiyi anlamak, tahminler yapmak ve gelecekteki değerleri tahmin etmek için kullanılır.

4.6 Evrişimli Sinir Ağı (Convolutional Neural Network)

Evrişimli Sinir Ağı (Convolutional Neural Network veya CNN), özellikle görüntü ve video verileri üzerinde başarılı olan derin öğrenme yöntemlerinden biridir. CNN, verileri evrişim işlemleri ve öğrenme katmanları kullanarak otomatik olarak özelliklerini çıkarır ve bu özellikler üzerinden karmaşık desenleri tanımak ve sınıflandırmak için kullanılır. Bu yöntem, kameralar ve sensörlerle toplanan hava kirliliği seviyelerinin izlenmesi, veri analizi, zaman serileri verilerini işleyerek gelecekteki hava kirliliği seviyelerinin tahmininde faydalı bir araçtır.

4.7 Yinelemeli Yapay Sinir Ağı (Recurrent Neural Network)

Yinelemeli Yapay Sinir Ağı (Recurrent Neural Network veya RNN), sıralı verileri işlemek için tasarlanmış bir makine öğrenimi yöntemidir. RNN, geçmiş zaman adımlarındaki bilgiyi hafızasında saklayarak ve gelecekteki tahminlerde bu bilgiyi kullanarak sıralı verilerdeki bağıntıları modellemeye odaklanır. Bu özelliği sayesinde, dil işleme, zaman serileri analizi, metin üretimi ve çeviri gibi sıralı veri uygulamalarında kullanılır. Uzun Kısa Dönem Bellek (LSTM) ve Kapı Özyinelemeli Geçitler (GRU), ileri RNN türleridir.

4.8 Uzun Kısa Dönem Bellek (Long Short Term Memory)

Uzun Kısa Dönem Bellek (LSTM), özellikle zaman serileri ve sıralı veriler gibi belirli veri türlerini işlemek için kullanılan bir makine öğrenimi yöntemidir. LSTM, önceki

veri noktalarının bağlamını korurken gelecekteki değerleri tahmin etmek için kullanılır. Ağırlıkların dinamik olarak güncellenmesine izin verir, bu da modelin verileri işlerken öğrenme yeteneğini artırır. LSTM içinde bulunan kapı mekanizmaları, hangi bilgilerin belleğe kaydedileceği ve hangi bilgilerin silineceği konusunda karar verir. Bu, modelin gereksiz bilgileri filtrelemesine yardımcı olur. Bu yöntem, metin analizi, dil işleme, zaman serileri tahmini, konuşma tanıma, ve otomatik çeviri gibi birçok uygulamada kullanılır.

4.9 Karar Ağacı (Decision Tree)

Karar Ağacı (Decision Tree) veri madenciliği ve makine öğrenimi alanlarında sınıflandırma ve regresyon görevlerini gerçekleştirmek için kullanılan bir modelleme tekniğidir. Karar ağacı, veri kümesini ağaç benzeri bir yapının içinde temsil eder. Her iç düğüm (node) bir karar noktasını, her dal (branch) bir kararı ve her yaprak (leaf) ise bir sınıf etiketini veya tahmini temsil eder. Karar ağacı modeli, bir veri örneğinin özellikleri üzerinde sırasıyla kararlar alarak sonuç sınıfını veya değerini tahmin etme yeteneğine sahiptir. Başlangıç düğümünden başlayarak, her düğüm veri özelliklerine göre bölünür ve sonuçta yapraklarda sınıflandırma veya regresyon sonuçları elde edilir. Bu yöntem sınıflandırma problemlerinde belirli bir örneğin sınıfını tahmin etmek ve regresyon problemlerinde belirli bir değeri tahmin etmek için kullanılır. Bunun yanı sıra, karar ağaçlarının Random Forest gibi yöntemlerle birleştirilmesi, daha güçlü tahmin modelleri oluşturabilmektedir.

5. VERİ ve YÖNTEM

5.1 Veri Tanımı

Hava kirleticilerin tahminine yönelik geliştirilen modelde iki ayrı kaynaktan alınan verilerin bir kısmı doğrudan, bir kısmı ise ön işleminden geçirilerek kullanılmıştır.

5.1.1 Hava kirliliği verileri

Hava kirleticilerinin saatlik konsantrasyon bilgileri T.C. Çevre, Şehircilik ve İklim Değişikliği Bakanlığına ait Ulusal Hava Kalite İzleme Ağından (UHKİA) elde edilmiştir. İl sınırları içindeki 6 istasyonun 2022 yılı, saatlik PM₁₀, PM_{2.5}, NO₂, SO₂, O₃ verileri çalışmaya dahil edilmiştir. Çizelge 5.1’de X* ile işaretlemiş hücreler, veri kaynağında ölçümünün yapıldığı belirtilmekle birlikte, inceleme döneminde hiçbir verisi bulunmayan kirleticileri göstermektedir.

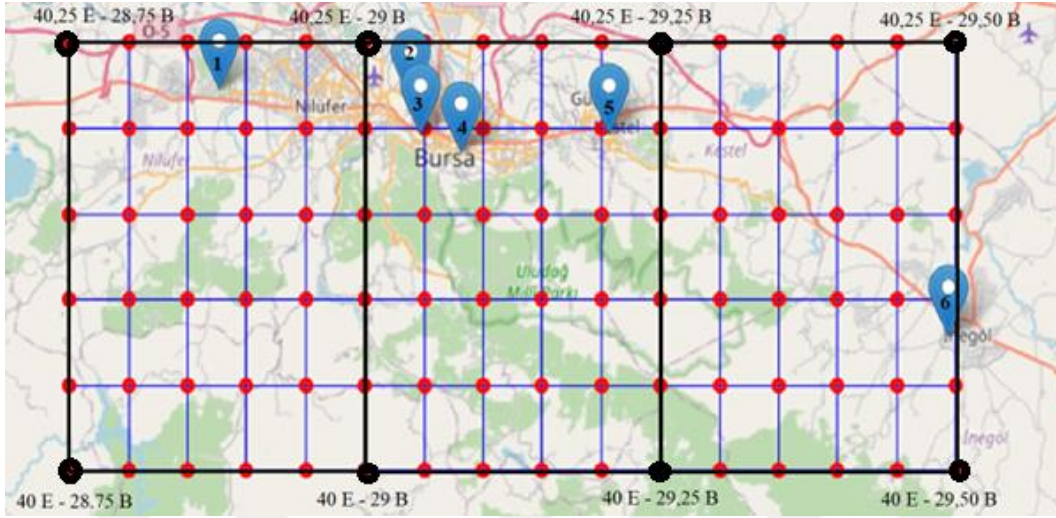
Çizelge 5.1 : İstasyonlarda ölçümü yapılan kirleticiler.

İstasyon Adı	PM ₁₀	PM _{2.5}	NO ₂	SO ₂	O ₃	CO
Uludağ Üniv.		X	X	X	X	
Bursa	X	X		X	X*	
Kültürpark		X	X	X	X	
Beyazıt Cad.	X		X	X		X*
Kestel	X		X	X	X	
İnegöl	X		X	X		
Nilüfer	X	X	X	X	X	X

Veri eksiklikleri nedeniyle modele doğrudan dahil edilmeyen Nilüfer istasyonunun mevcut verileri diğer istasyonlardaki eksiklerin tamamlanması için yöntem araştırması kapsamında kullanılmıştır.

Değerlendirmeye alınan istasyonlar Şekil 5.1’de 0,05° x 0,05°’lik piksellere ayrılmış her biri 0,25° x 0,25° enlem ve boylam boyutlarında üç ızgara sisteminde görselleştirilmiştir. İstasyonların bir ızgara sisteminde ele alınmasının nedeni, meteorolojik parametrelerin elde edildiği ERA5 veri setinin 0,25° x 0,25° mekansal çözünürlüğe sahip veri üretmesidir. İstasyonlara batıdan başlayarak sırayla numara verilmiştir: 1. Uludağ Üniversitesi, 2. Bursa, 3. Kültürpark, 4. Beyazıt Caddesi, 5. Kestel, 6. İnegöl. Şekil 5.1’den de görülebileceği gibi İnegöl istasyonu hariç diğer tüm

istasyonlar doğu batı eksenli ve genel olarak ızgaraların kuzeyine yakın konumlanmıştır.



Şekil 5.1 : Hava kalitesi izleme istasyonlarının konumu.

Hava kalitesi izleme istasyonlarına bulundukları bölgenin özellikleri göz önünde bulundurularak farklı tiplerde cihaz yerleştirilmiştir. Örneğin, Uludağ Üniversitesi, Kültürpark ve Bursa istasyonları ısınma tipi, Beyazıt Caddesi istasyonu trafik tipi, Kestel ve İnegöl istasyonları sanayileşmenin yoğun olduğu bölgelerde bulunduklarından sanayi tipidir.

5.1.2 Meteorolojik veriler

Bu çalışmada kullanılan meteorolojik veriler, Avrupa Birliği'nin bir girişimi olan Copernicus Programı kapsamında Avrupa Orta Menzilli Hava Tahminleri Merkezi (ECMWF) tarafından üretilen ve Copernicus Climate Data Store'dan indirilen ERA5 reanaliz veri setinden elde edilmiştir (Hersbach ve diğ, 2023). ERA5, model verilerini dünyanın dört bir yanından gelen gözlemlerle küresel bir veri kümesinde birleştirmek için fizik yasalarını kullanarak daha fazla tarihsel gözlem verisi kullanmaktadır.

ERA5'in saatlik ve eksiksiz veri üretmesi, meteorolojik parametreler ve hava kirliliği ilişkisi ile ilgili çalışmalar için önemlidir. ERA5 reanaliz veri setlerinin Türkiye'de gerçek zamanlı ölçümlerle uyumluluğunu konu alan bir çalışmada saatlik hava sıcaklığı verileri gerçek zamanlı ölçümlerle karşılaştırılmış ve ERA5 verilerinin güvenli şekilde kullanılabileceği vurgulanmıştır (Yılmaz, 2022).

Bursa il sınırlarındaki hava kalitesi izleme istasyonlarını içine alacak şekilde $0.25^\circ \times 0.25^\circ$ yatay çözünürlükte 3 ayrı ızgara baz alınarak, 2022 yılı saatlik sıcaklık, çiy

noktası sıcaklığı, u ve v bileşenlerinde rüzgar hızı, ortalama deniz seviyesi basıncı, sınır tabaka yüksekliği, sınır tabaka dağılımı, yüzey net güneş radyasyonu, toplam bulut örtüsü bilgileri Copernicus Climate Data Store'dan NetCDF dosya formatında indirilmiştir. Daha sonra veri seti CSV (comma separated value) uzantılı dosyalara dönüştürülmüştür.

Meteorolojik verilerin bazıları üzerinde birim dönüştürme işlemi yapılırken, bazıları başka meteorolojik parametrelerin hesaplanmasında kullanılmıştır. ERA5 reanaliz veri setinden elde edilen sıcaklık ve çiy noktası sıcaklığı bilgileri Kelvin (°K) cinsinden sunulmakta olup, tüm veriler Celcius (°C) birimine çevrilmiştir. İki metre yükseklikteki sıcaklık ve çiy noktası sıcaklığı verisinden bağıl nem hesaplaması yapılmıştır. Hesaplama için meteoroloji ve atmosfer verilerinin analizini ve görselleştirmesini kolaylaştırmak amacıyla tasarlanan MetPy adlı Python kütüphanesinden faydalanılmıştır (May ve diğ, 2022).

MetPy kütüphanesinde kullanılan bağıl nem (RH) fonksiyonun temel aldığı formül 5.1 Bolton (1980) tarafından geliştirilmiştir. Formüle göre:

$$RH(\%) = 100 \cdot \frac{e}{e_s} \quad (5.1)$$

Burada bağıl nem yüzde olarak bağıl nem miktarını, e belirli bir sıcaklık ve basınçta havanın buhar basıncını (hPA), e_s ise belirli bir sıcaklıkta doymuş buhar basıncını (hPA) göstermektedir. Formüldeki aktüel buhar basıncı e , Celcius birim cinsinden çiy noktası sıcaklığı (T_d) kullanılarak formül 5.2'deki gibi hesaplanmaktadır:

$$e = 6.112 \cdot \exp\left(\frac{17.67 \cdot T_d}{T_d + 243.5}\right) \quad (5.2)$$

Doymuş buhar basıncının e_s hesaplanması için ise formül 5.3'te Celcius birim cinsinden sıcaklık bilgisi kullanılmaktadır.

$$e_s = 6.112 \cdot \exp\left(\frac{17.67 \cdot T}{T + 243.5}\right) \quad (5.3)$$

ERA5 reanaliz veri setinde 10 metrede rüzgar hızı iki ayrı bileşen (u: doğuya doğru ve v: kuzeye doğru) şeklinde verilmektedir. Bu iki bileşen MetPy kütüphanesinden Python dilinde rüzgar hızı fonksiyonu kullanılarak tek bir rüzgar hızı verisine dönüştürülmüştür.

$$Rüzgar Hızı(WS) = \sqrt{(u^2 + v^2)} \quad (5.4)$$

Nihai veri setinde yer alan meteorolojik parametrelere Çizelge 5.2’de yer verilmiştir:

Çizelge 5.2 : Meteorolojik parametreler ve birimleri.

Meteorolojik Parametre	Birim
Rüzgar hızı	Metre/saniye (m s ⁻¹)
2 metrede sıcaklık	Celcius (°C)
Bağıl nem	%
Sınır tabaka yüksekliği	Metre (m)
Ortalama deniz seviyesi basıncı	Hektopascal (hPa)
Sınır tabaka dağılımı	Joule/m ² (J m ⁻²)
Yüzey net güneş radyasyonu	Joule/m ² (J m ⁻²)
Toplam bulut örtüsü	-

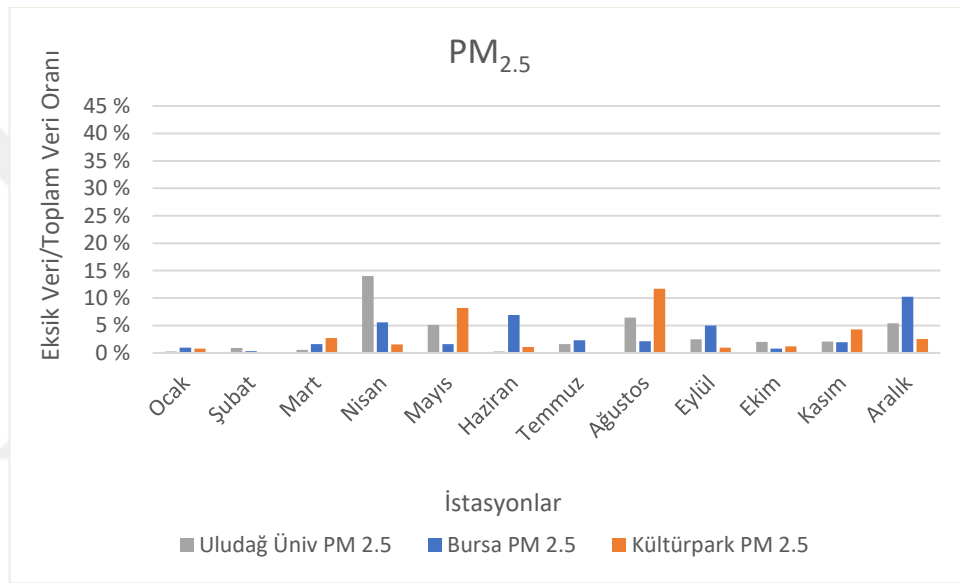
Bu çalışmada yüzey basıncı yerine ortalama deniz seviyesi basıncı verisi kullanılmıştır. Bunun nedeni basıncın yükseklikle hızla değişmesi ve bu durumun dağlık alanlardaki alçak ve yüksek basınçlı hava sistemlerini görmeyi zorlaştırmasıdır.

5.2 Veri Ön İşleme

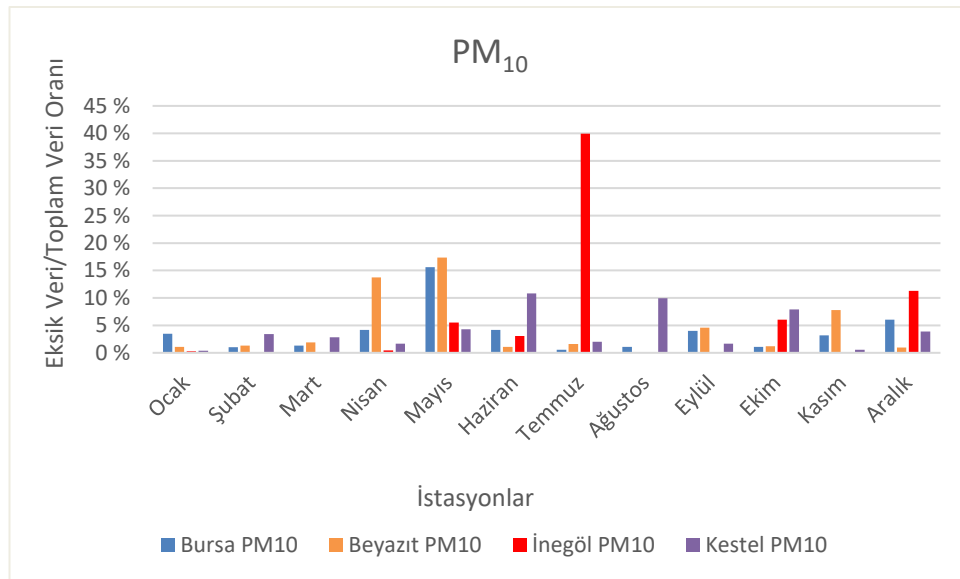
Hava kalitesi veri setleri genellikle çok sayıda eksik veri içermektedir. Kirletici seviyesini ölçen cihazlardaki arızalar, cihazların enerji kaynaklarındaki sorunlar, ölçümlerin aktarıldığı bilgisayar sistemlerindeki hatalar, sensörlerin ölçüm sınırı değerlerini aşan kirletici seviyelerini ölçmemesi gibi nedenlerle veri kaybı ya da hatası oluşmaktadır. Hava kirliliği ile ilgili çalışmalarda, genellikle çalışmanın ilk etabı eksik verilerin tamamlanması işlemidir. Bir zaman serisi olarak elde edilen kirletici verisi ardışık olarak birbiriyle ilişki halindedir. Dolayısıyla zaman serisinin eğilimi seri

içerisinde eksik olan veriyi yerine koymada anlamlı bir kolaylık sağlar. Kirletici veri setlerindeki eksik veriyi tamamlamaya yönelik çok sayıda çalışma yapılmıştır. Ancak hangi yöntem kullanılırsa kullanılsın, kullanılan yöntemle elde edilen verinin hiçbir zaman gerçek veri olarak kabul edilemeyeceği açıktır. Bu nedenle araştırmacılar eksik veri tamamlama yöntemlerine bir “istatistiksel simya” olarak yaklaşılması gerektiğini belirtmektedirler (Junninen ve diğ, 2004).

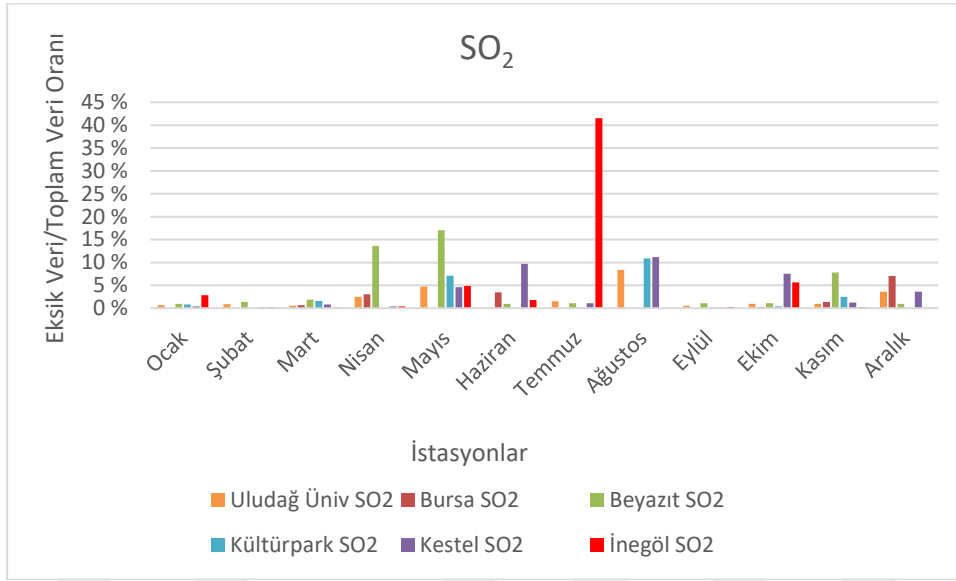
Bursa ilinde incelemeye konu 6 istasyona ait verilerin ön incelemesinde eksik veriler tespit edilmiştir. Kirletici bazında aylık eksik veri/toplam veri oranı Şekil 5.2, Şekil 5.3, Şekil 5.4, Şekil 5.5 ve Şekil 5.6’da gösterilmektedir.



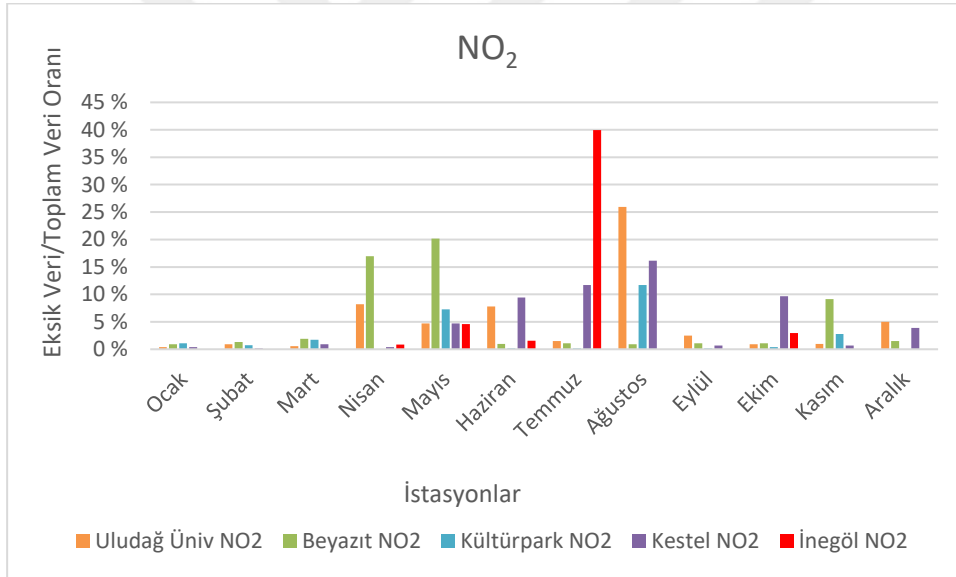
Şekil 5.2 : Aylık bazda PM_{2.5} eksik veri oranı dağılımı.



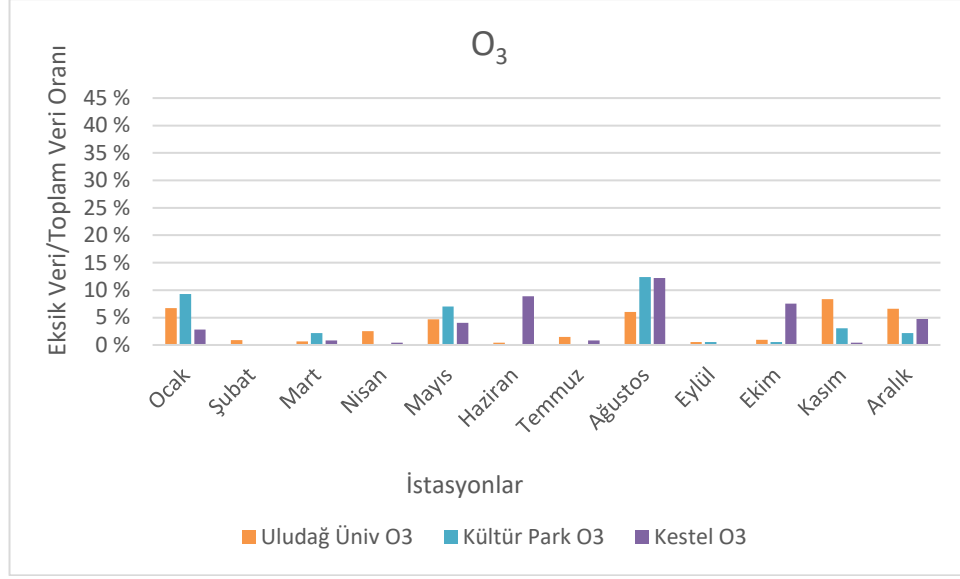
Şekil 5.3 : Aylık bazda PM₁₀ eksik veri oranı dağılımı.



Şekil 5.4 : Aylık bazda SO₂ eksik veri oranı dağılımı.



Şekil 5.5 : Aylık bazda NO₂ eksik veri oranı dağılımı.

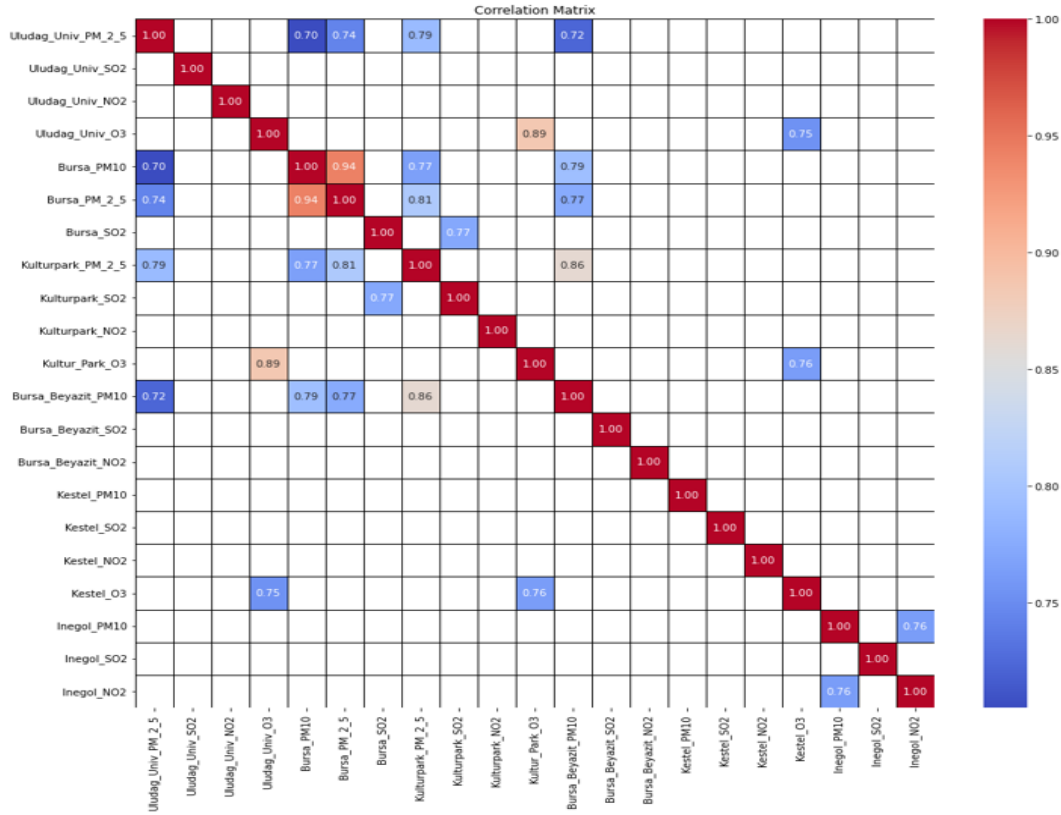


Şekil 5.6 : Aylık bazda O₃ eksik veri oranı dağılımı.

Eksik verilerin tamamlanması için eksik veri oranı, istasyonlar arası korelasyon gibi hususlar değerlendirilerek farklı yöntemler uygulanmıştır. Bu kapsamda çalışmada öncelikle tüm istasyonların kirletici seviyelerindeki eksiklikler görsel olarak incelenmiştir. Birinci aşamada incelemeye konu kirleticinin hiçbir istasyondan ölçüm verisinin alınmadığı saatler araştırılmıştır. İnceleme neticesinde 11.04.2022, 22.09.2022, 13.12.2022 tarihlerinde PM_{2.5} ölçümü yapılan Uludağ Üniversitesi, Bursa ve Kültürpark istasyonlarında ölçümün olmadığı ortak saatler belirlenmiştir. Bu üç istasyonda eksik olmakla birlikte veri ölçümü çok düşük olduğundan çalışmada yer verilmeyen Nilüfer istasyonunun aynı tarih ve saatlerde ölçümünün olduğu görülmüştür. İkinci aşamada Nilüfer istasyonundaki veriler lineer regresyon yöntemi ile aynı ızgarada yer aldığı Uludağ Üniversitesi istasyonundaki eksik verilerin tamamlanmasında kullanılmıştır. Benzer şekilde NO₂ ve O₃ verilerinde de 17.05.2022 ve 18.05.2022 tarihlerinde söz konusu kirleticinin ölçümünün yapıldığı hiçbir istasyonda veri bulunmadığı görülmüş, aynı tarihlerde ölçümü bulunan Nilüfer istasyonundaki veriler kullanılarak Uludağ Üniversitesi istasyonundaki eksik veriler lineer regresyon yöntemi ile tamamlanmıştır. Lineer regresyon, mevcut verilere dayalı doğrusal bir model uydurarak eksik değerleri tahmin etmekte kullanılmaktadır.

Lineer regresyon eksik verisi tamamlanacak kirletici için başka bir istasyonda aynı zamana denk gelen verinin olması durumunda kullanışlıdır. Burada her iki istasyonda aynı kirleticinin mevcut verilerinin birbiriyle korelasyonunun yüksek olması da çok önemlidir. Şekil 5.7’de kirleticilerin istasyonlar arası korelasyonu gösterilmektedir.

Korelasyon matrisinde yalnız % 70 ve üzerinde korelasyona sahip kirleticilere yer verilmiştir.



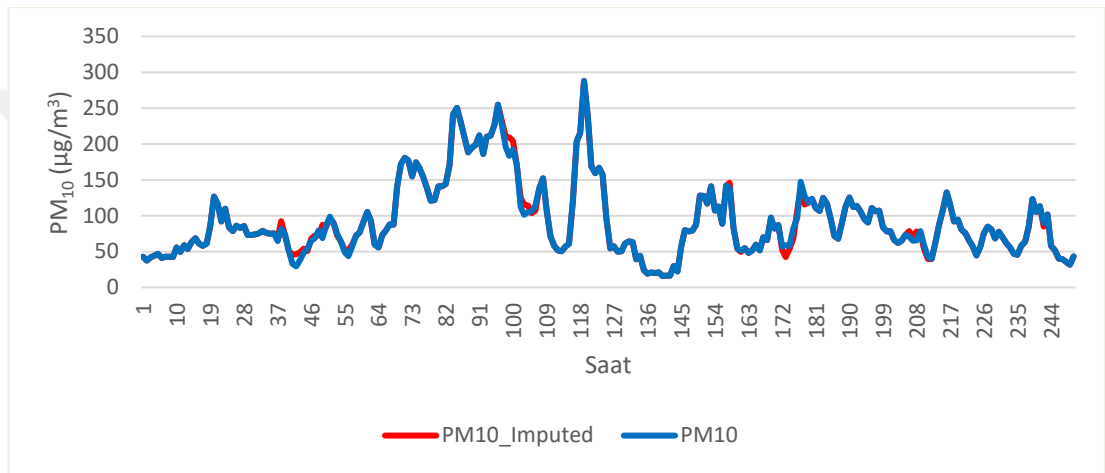
Şekil 5.7 : Kirletici korelasyon matrisi.

Şekil 5.7’den de görüleceği üzere, bazı kirleticiler için istasyonlar arası korelasyon oranına yer verilmemiştir. Örneğin, Uludağ Üniversitesi ve Beyazıt Caddesi istasyonlarında ölçülen SO₂ ve NO₂, Kültürpark istasyonunda ölçülen NO₂, Kestel ve İnegöl istasyonlarında ölçülen PM₁₀, SO₂ ve NO₂’nin veri setlerindeki eksikleri tamamlamak için korelasyonu % 70’in üzerinde olan başka bir istasyon bulunmamaktadır. Bu nedenle lineer regresyon yönteminden vazgeçilmiş ve eksik veri tamamlama için farklı yöntemler incelenmiştir.

Çalışmada eksik kirletici verisini tamamlamak üzere güçlü bir makine öğrenimi algoritması olan XGBoost (Extreme Gradient Boosting) tercih edilmiştir. XGBoost, eksik verileri tahmin etme yeteneği, genel veri işleme yeteneğiyle birleştiğinde eksik veri sorunlarını çözmek için etkili bir seçenek sunmaktadır. XGBoost, Gradyan Artırma (Gradient Boosting) algoritmasının performansı yükseltilmiş versiyonudur. XGBoost eksik verileri modellemek için doğrusal ve doğrusal olmayan ilişkileri bir arada ele alabildiğinden, verilerin karmaşıklığına uygun bir şekilde eksik veri tahminlerinin yapılmasına yardımcı olur. Eksik verilerin imputasyonunda XGBoost

kullanılırken, eksik veri noktalarının tahmin edilmesi ve veri kümesinin daha tutarlı ve eksiksiz hale getirilmesi amaçlanır. Bu, daha sonra veri analizi, modelleme ve tahmin işlemlerinin daha doğru ve güvenilir olmasına yardımcı olmaktadır.

Çalışma için XGBoost yönteminin veri tamamlama başarısını test etmek üzere, Bursa istasyonuna ait 250 saatlik gerçek PM₁₀ veri setinden çeşitli saatlere ait kimisi ardışık olmak üzere 47 saatlik ölçüm bilgisi veri setinden çıkartılmıştır. Python’da XGBoost Regression modeline eksik PM₁₀ verisinin yanında, aynı zaman dilimine ait PM_{2.5} ve SO₂ ölçümleri de girdi olarak verilmiştir. Modelin sonuçları Şekil 5.8’deki gibi görselleştirilmiştir.

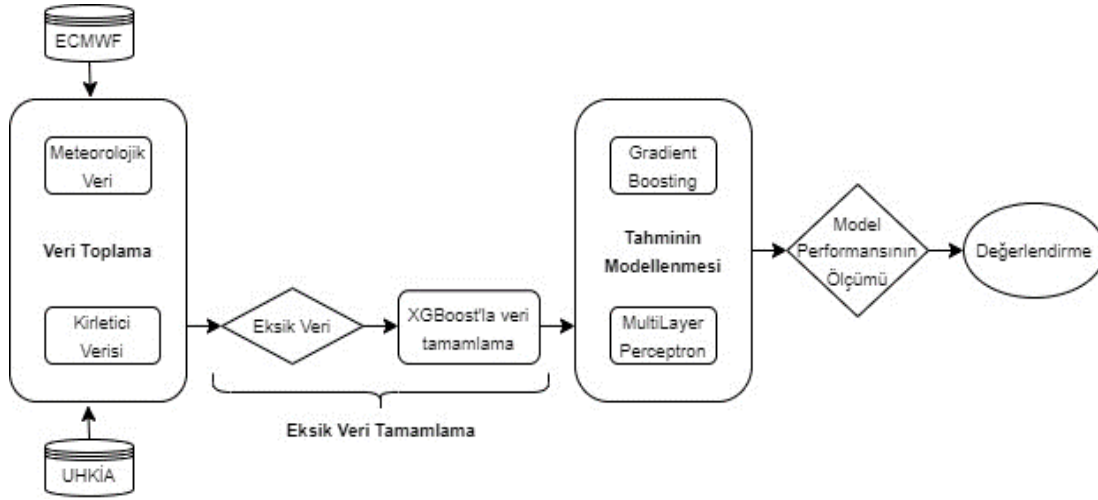


Şekil 5.8 : XGBoost ile tamamlanmış örnek veri seti.

5.3 Yöntem

Bursa ilinde çalışmaya konu edilen altı hava kalitesi istasyonundan sadece üçünde (Uludağ Üniversitesi, Bursa ve Kültürpark) PM_{2.5} ölçümü yapılmaktadır. Buna karşılık istasyonların dördünde (Bursa, Beyazıt Caddesi, Kestel ve İnegöl) PM₁₀ konsantrasyonu ölçülmektedir. Bursa istasyonu hem PM_{2.5} ve hem de PM₁₀ ölçümü yapılan tek istasyondur.

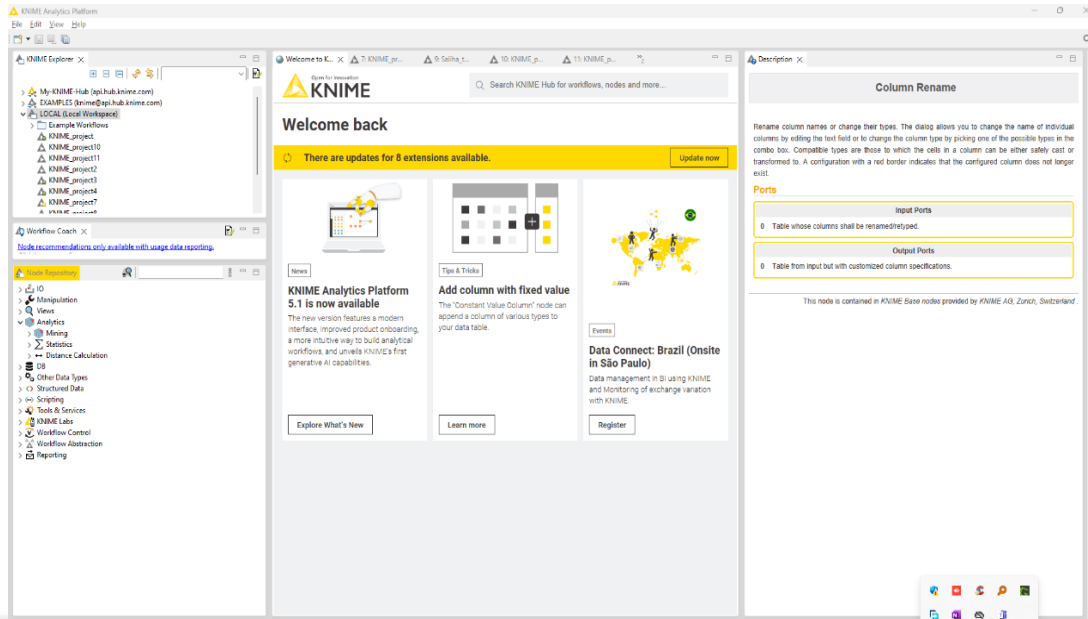
Çalışmada kullanılan iki farklı makine öğrenmesi yöntemiyle birinci aşamada Bursa istasyonu için bir haftalık PM_{2.5} ve PM₁₀ tahmini denenmiştir. İkinci aşamada PM_{2.5} ve PM₁₀ ölçümü yapılmayan istasyonlar için Bursa istasyonu verilerinden oluşturulan eğitim seti kullanılarak tahmin yapılması hedeflenmiştir. Şekil 5.9’da partiküler madde tahmini için hazırlanan akış diyagramına yer verilmiştir.



Şekil 5.9 : Partiküler madde tahmini iş akış diyagramı.

Ulusal Hava Kalitesi İzleme Ağı (UHKİA) ile Avrupa Orta Menzilli Hava Tahminleri Merkezi (ECMWF) tarafından sunulan veriler ön incelemeye tabi tutulmuş, kirlenici verilerindeki eksiklikler bir makine öğrenmesi algoritması ile tamamlanmıştır. Tüm veriler eksiksiz hale getirildikten sonra Gradient Boosting ve Multilayer Perceptron makine öğrenmesi yöntemlerini içeren tahmin modelleri KNIME adı verilen veri madenciliği platformunda çalıştırılmış ve model sonuçları yine bu platform aracılığıyla değerlendirilmiştir.

KNIME (Konstanz Information Miner) veri madenciliği platformu veri analizi, iş akışı yönetimi ve veri madenciliği için kullanılan açık kaynaklı bir platformdur (Berthold ve diğ., 2007). Bu platform, farklı veri kaynaklarından veriyi aktarmada, veriyi ön işleme tabi tutarak analize hazır hale getirmede oldukça kullanışlıdır. Veri madenciliği ve makine öğrenme işlemleri için bir dizi algoritma ve işlevi içerdiğinden, tahmin, sınıflandırma, kümeleme ve diğer veri madenciliği görevlerini gerçekleştirmeyi sağlar. Verileri anlamak için çeşitli grafik ve görselleştirme araçlarını içerir. İş akış yönetimi sayesinde, farklı veri işleme ve analiz adımlarını sıralar ve otomatikleştirir. KNIME, veri bilimi, iş zekası, veri madenciliği ve makine öğrenme projeleri için geniş bir kullanım alanı sunmaktadır.



Şekil 5.10 : KNIME Platformu

5.3.1 Makine öğrenmesi yöntemleri ile $PM_{2.5}$ ve PM_{10} tahmini

Çalışma üç aşamaya ayrılmıştır. Birinci aşamada, hem $PM_{2.5}$ ve hem de PM_{10} ölçümü yapılan tek istasyon olan Bursa istasyonu için kirletici verileri ($PM_{2.5}$, PM_{10} , SO_2) ve meteorolojik parametreler girdi olarak kullanılmış ve iki makine öğrenmesi yöntemi ile bir haftalık (7 gün, 168 saat) $PM_{2.5}$ tahmini yapılmıştır.

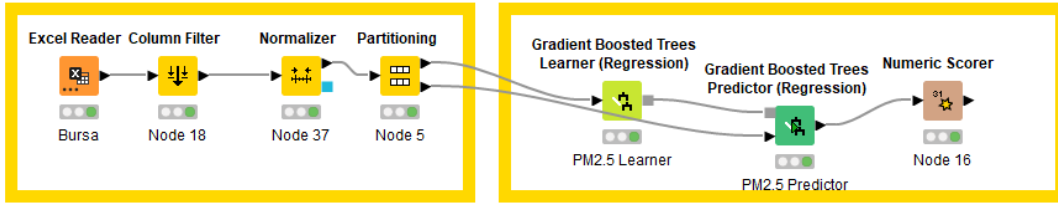
İkinci aşamada Bursa istasyonuna ilişkin $PM_{2.5}$ ve PM_{10} tahmin algoritmaları içinde oluşturulan eğitim verisi setleri, diğer istasyonlarda hangi partiküler madde ölçümü yapılmıyorsa o istasyonun tahmin modeline yerleştirilmiş ve istasyonlar için ölçümü bulunmayan partiküler madde türü için tahmin yapılmıştır.

Üçüncü aşamada ise, tahminlerin tutarlılığını test etmek için bu kez tahmin edilen partiküler madde verisi ile aynı istasyon için istasyonda ölçümü yapılan partiküler madde için tahmin yapılmıştır. Tahmin edilen ile istasyonda ölçülen verinin hata metrikleri aracılığıyla karşılaştırması yapılarak istasyonunda ölçümü yapılmayan partiküler madde tahmininin başarısının ölçülmesi sağlanmıştır.

5.3.1.1 Bursa Hava Kalitesi Ölçüm İstasyonu için $PM_{2.5}$ ve PM_{10} tahmini

Bursa Hava Kalitesi Ölçüm İstasyonundan elde edilen PM_{10} , $PM_{2.5}$ ve SO_2 verileri çalışmada kirletici verisi olarak kullanılmıştır. İstasyonun 2022 yılına ait saatlik kirletici verileri ile aynı yıla ait saatlik meteorolojik veriler öncelikle Şekil 5.11’de

görüldüğü şekilde Gradient Boosting yöntemiyle oluşturulan modele girdi olarak verilmiştir.

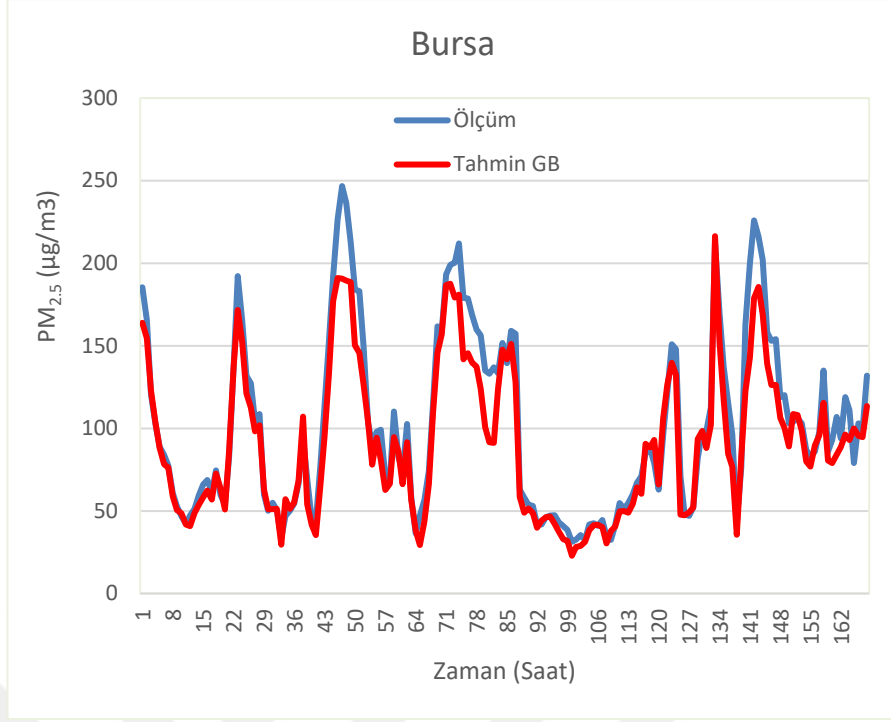


Şekil 5.11 : Bursa istasyonu için Gradient Boosting ile PM_{2.5} tahmin modeli.

KNIME Platformunda kullanılan Gradient Boosting yöntemi Friedman'ın (2001) yöntemi konu aldığı çalışmada verilen algoritmaya göre oluşturulmuştur.

Modelde öncelikle Excel formatındaki veriler içinde enlem/boylam/tarih gibi tahmini doğrudan ilgilendirmeyen sütunlar “Column Filter” düğümü ile ayrılmış, kalan tüm veri seti her bir veri türü kendi içinde olacak şekilde “Normalizer” düğümü ile 0 ile 1 arasına karşı gelecek şekilde bir değere indirgenmiştir. Veri seti, “Partitioning” düğümü ile eğitim verisi (8592) ve tahmin verisi (7 gün karşılığı 168 saat) olarak kullanılmak üzere ikiye ayrılmıştır. Eğitim için ayrılan veriler “Gradient Boosted Trees Learner (Regression)” düğümünde modelin eğitilmesi için kullanılmıştır. Oluşturulan eğitim verisi seti ile tahmin için ayrılan veri “Gradient Boosted Trees Predictor (Regression)” düğümünde kullanılarak model aracılığıyla tahmin gerçekleştirilmiştir. “Numeric Scorer” çeşitli hata metrikleri vererek modelin performansının değerlendirilmesini sağlamaktadır.

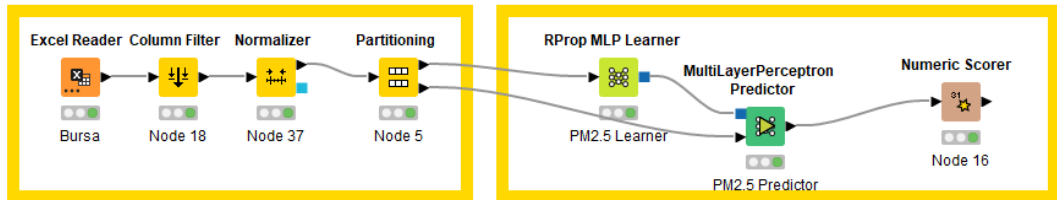
Model sonucunda tahmin edilen PM_{2.5} verisi ile gerçek PM_{2.5} verisi Şekil 5.12'deki gibi görselleştirilmiştir. Grafikte modelin ekstrem değerlerde bir miktar düşük tahmin değerleri üretse de kirleticinin gerçek değerlerine oldukça yakın tahmin yaptığı görülmektedir.



Şekil 5.12 : Ölçüm verisi ve GB ile tahmin verisinin görsel karşılaştırılması.

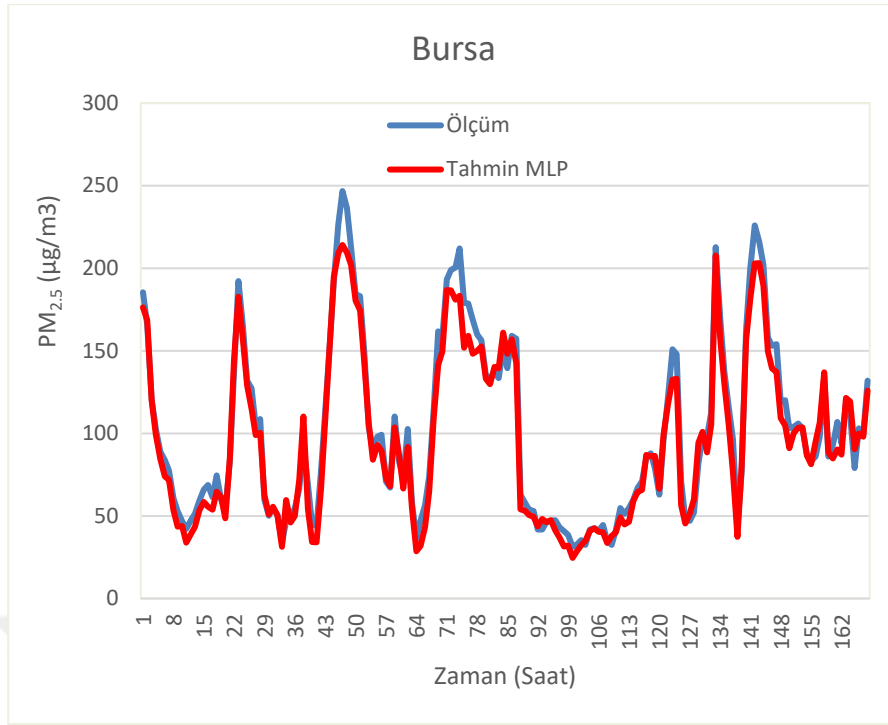
Aynı veriler bu kez MultiLayer Perceptron makine öğrenmesi yöntemi kullanılarak oluşturulan modelde denenmiştir. Platformda kullanılan yöntemin algoritması Riedmiller ve Braun'un (1993) bir makalesine dayanmaktadır. Elastik geri yayılım olarak Türkçe'ye çevrilebilecek RPROP (Resilient Propagation) algoritması çok katmanlı yapay sinir ağlarının eğitiminde yaygın olarak kullanılan bir öğrenme yöntemidir. Bu yöntem, ağırlık çıktıları ile gerçek çıktıları arasındaki farkı hesaplar ve bu farkı ağırlık tüm katmanlarına geriye doğru yayarak ağırlıkların güncellenmesini sağlar. Bu işlem, ağırlık hata fonksiyonunu minimize etmek için tekrarlanır.

Çalışmada kullanılan MultiLayer Perceptron yöntemine dayalı model algoritması Şekil 5.13'deki gibidir.



Şekil 5.13 : Bursa istasyonu için Multilayer Perceptron ile PM_{2.5} tahmin modeli.

MultiLayer Perceptron yöntemi ile tahmin edilen PM_{2.5} ile ölçülen PM_{2.5} verisini karşılaştıran grafiğe de Şekil 5.14'te yer verilmiştir.



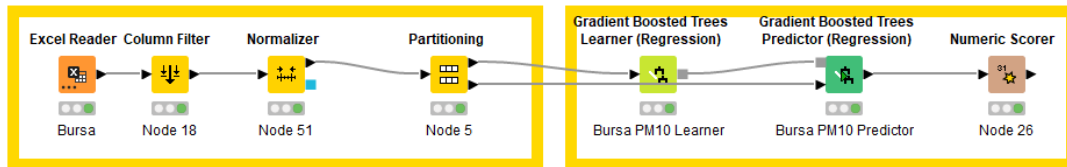
Şekil 5.14 : Ölçüm verisi ve MLP ile tahmin verisinin görsel karşılaştırması.

PM_{2.5} tahmininde her iki yöntemin hata metrikleri ise Çizelge 5.3'te karşılaştırılmıştır. Hata metriklerinden Bursa istasyonu için PM_{2.5} tahmininde MultiLayer Perceptron yönteminin daha başarılı tahmin sonuçları ürettiği görülmektedir.

Çizelge 5.3 : Bursa istasyonu için PM_{2.5} tahmininde GB ve MLP hata metrikleri.

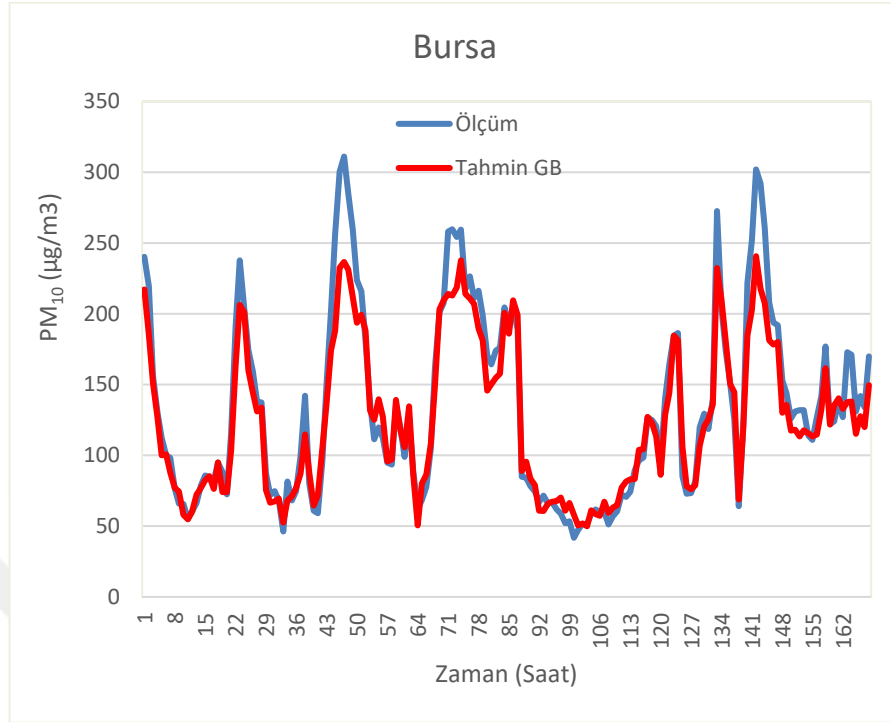
Hata Metriği	GB	MLP
R ²	0,90	0,967
Ortalama Mutlak Hata	0,047	0,031
Ortalama Kare Hata	0,005	0,002
Kök Ortalama Kare Hata	0,068	0,039

Bu kez, Bursa istasyonunda PM₁₀ tahmini için Gradient Boosting yöntemine dayalı Şekil 5.15'teki model oluşturulmuştur.



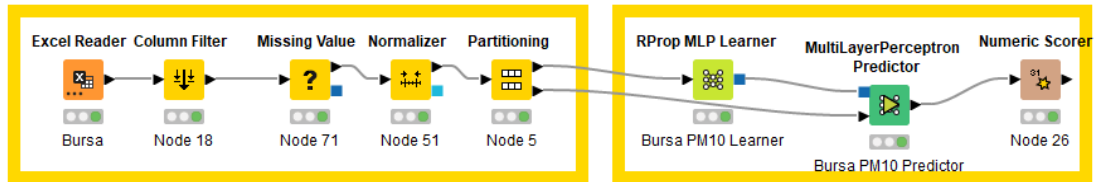
Şekil 5.15 : Bursa istasyonu için Gradient Boosting ile PM₁₀ tahmin modeli.

Şekil 5.16’da Bursa istasyonuna ait tahmin edilen ve ölçülen PM_{10} değerlerini gösteren bir grafik verilmiştir.



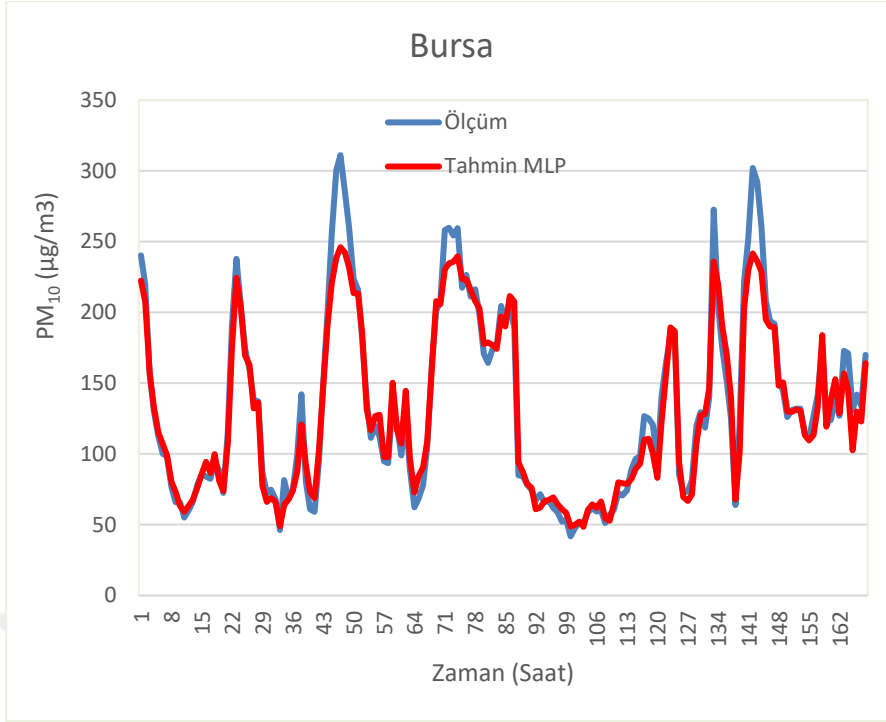
Şekil 5.16 : PM_{10} ölçüm verisi ve GB ile tahmin verisinin görsel karşılaştırması.

Son olarak aynı istasyon ve aynı partiküler madde için bu kez Şekil 5.17’deki MultiLayer Perceptron ile bir model oluşturulmuştur.



Şekil 5.17 : Bursa istasyonu için Multilayer Perceptron ile PM_{10} tahmin modeli.

Model sonuçlarına Şekil 5.18’deki grafikte yer verilmiştir. MultiLayer Perceptron ile tahmin edilen PM_{10} ile ölçüm değerlerinin genelde birbirine çok yakın olduğu, yöntemin yalnız ekstrem konsantrasyonlarda bir miktar daha düşük tahmin değerleri ürettiği görülmektedir.



Şekil 5.18 : PM₁₀ ölçüm ve MLP ile tahmin verisinin görsel karşılaştırması.

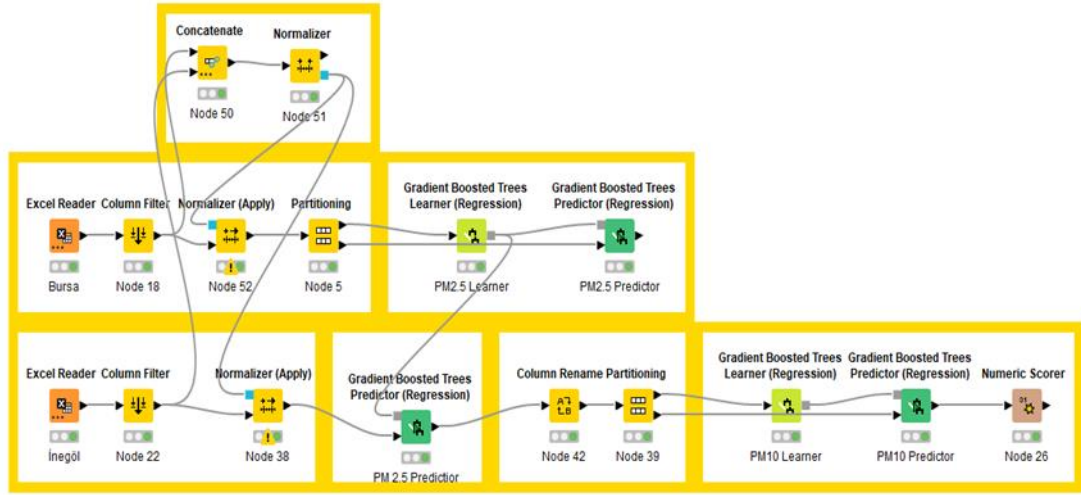
Bursa istasyonu için PM₁₀ tahmininde Gradient Boosting ve MultiLayer Perceptron yöntemlerinin performansı Çizelge 5.4.'de karşılaştırılmıştır. Hata metriklerinden de görüldüğü gibi her iki yöntem de veri setinde PM_{2.5} ve PM₁₀ verisi bulunduğunda başarılı sonuçlar vermektedir. Her halükarda MultiLayer Perceptron daha yüksek tahmin performansı göstermiştir.

Çizelge 5.4 : Bursa istasyonu için PM₁₀ tahmininde GB ve MLP hata metrikleri.

Hata Metriği	GB	MLP
R^2	0,908	0,949
Ortalama Mutlak Hata	0,043	0,032
Ortalama Kare Hata	0,004	0,002
Kök Ortalama Kare Hata	0,064	0,048

5.3.1.2 İnegöl, Kestel ve Beyazıt Caddesi istasyonları için PM_{2.5} tahmini

Bursa ili sınırları içerisinde bulunan istasyonlardan İnegöl, Kestel ve Beyazıt Caddesi adlı hava kalitesi izleme istasyonlarında PM_{2.5} ölçümü yapılmamaktadır. Bu bölümde Bursa istasyonuna ait PM_{2.5} eğitim modu kullanılarak, ölçüm yapılmayan istasyonlar için PM_{2.5} tahmini denenmiştir. Öncelikle Şekil 5.19'da görüldüğü gibi Gradient Boosting yöntemiyle bir model oluşturulmuştur.

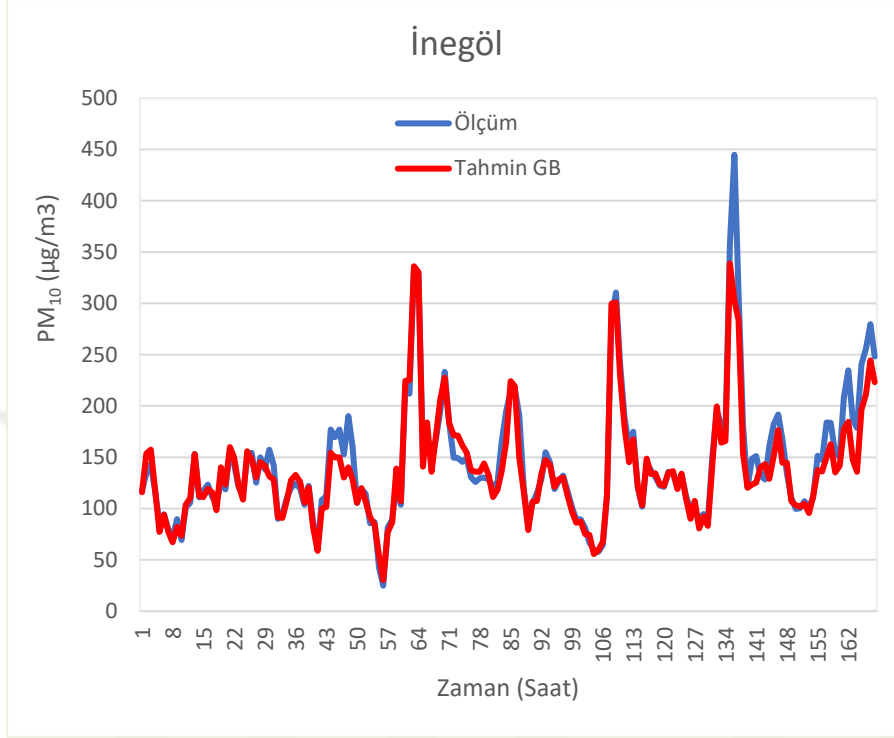


Şekil 5.19 : GB yöntemi ile $PM_{2.5}$ ve PM_{10} tahmini modeli.

Model sonucunda elde edilen verilerin tutarlılığını test etmek için, bu kez tahmin edilen $PM_{2.5}$ verisi, veri setine ölçüm verisi gibi eklenmiş ve ölçümü mevcut olan PM_{10} tahmin edilmiştir. Son aşamada ölçülen ve tahmin edilen PM_{10} verilerinin performansı değerlendirilmiştir. Bu modelle iki ayrı hipoteze cevap aranmıştır. Birinci hipotez ölçüm yapılmayan istasyon için başka bir istasyonun eğitim modunun kullanılması yoluyla $PM_{2.5}$ tahmini başarılıdır. İkinci hipotez ise, $PM_{2.5}$ tahmin verisi de eklenerek aynı istasyon için elde edilen PM_{10} tahmin verisi ile PM_{10} ölçüm verisi yüksek korelasyon gösterirse, istasyonda ölçümü yapılmayan $PM_{2.5}$ tahmini başarılıdır. Şekil 5.19’da Gradient Boosting yöntemi ile Bursa istasyonuna ait $PM_{2.5}$ eğitim modu kullanılarak önce İnegöl istasyonu için $PM_{2.5}$ verisi ve sonra bu veri de eklenerek aynı istasyon için PM_{10} tahmini yapılmıştır. Daha sonra PM_{10} tahmin performansı hata metrikleri ile değerlendirilmiştir.

Platformun görev düğümleri (node) ise sırasıyla şöyle çalıştırılmıştır: Önce eğitim modelinin oluşturulduğu Bursa istasyonu ile kirletici verisi tahmin edilecek istasyonun Excel formatındaki verileri platforma aktarılmıştır. Modellerde Bursa istasyonunda $PM_{2.5}$ (PM_{10} tahmininde PM_{10}) haricinde hangi tür veri bulunuyorsa, diğerlerinde de aynı tür veriler bulunmalıdır. Bu kapsamda İnegöl ve Beyazıt Caddesi istasyonları için model çalıştırılırken NO_2 , Kestel istasyonu için model çalıştırılırken NO_2 ve O_3 veri veri setinden Column Filter düğümünde çıkarılmıştır. İstasyonların verileri birleştirilerek (Concatenate), her iki istasyonun aynı tür verileri normalize (Normalize (Apply)) edilmiştir.

İnegöl Hava Kalitesi İzleme İstasyonu için Gradient Boosting yöntemiyle elde edilen tahminler ile ölçüm verileri Şekil 5.20’de karşılaştırılmıştır.

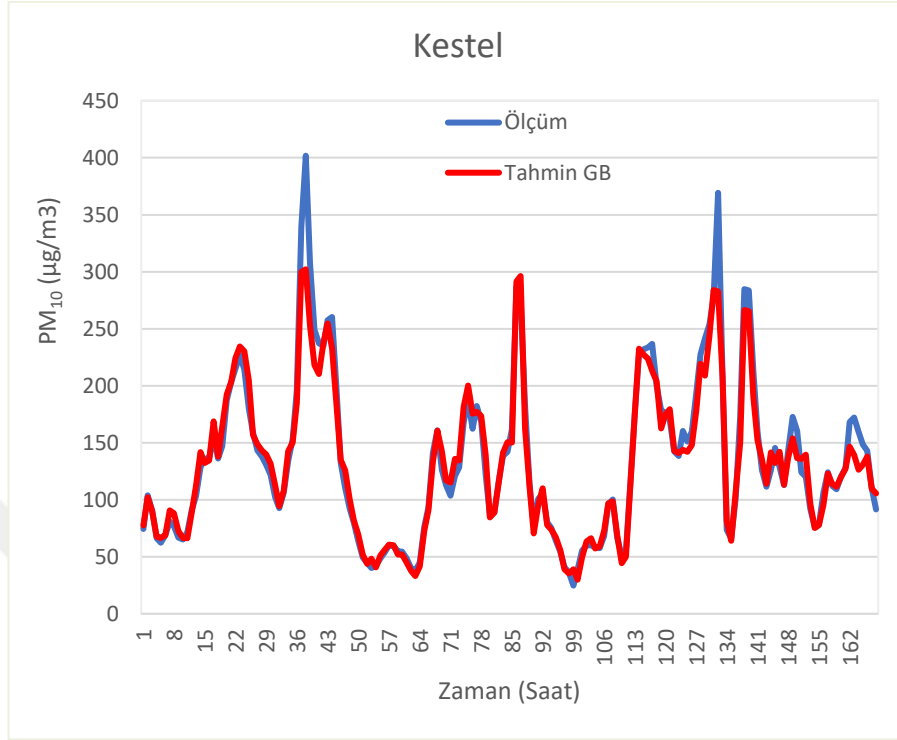


Şekil 5.20 : İnegöl istasyonu için PM₁₀ ölçüm ve GB ile tahmin verisi.

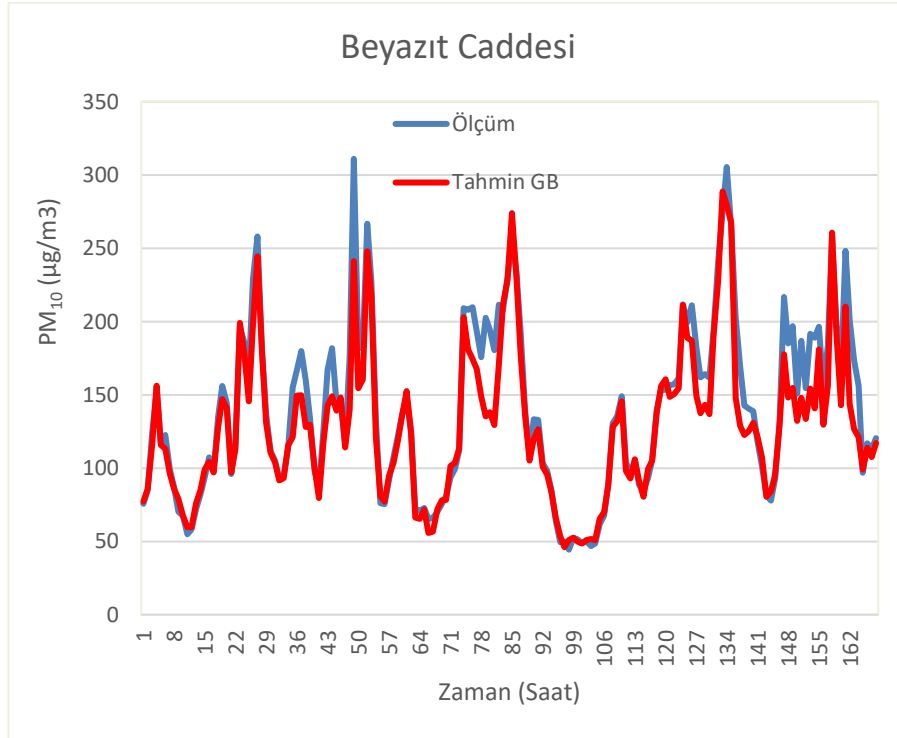
Şekil 5.20’de de görüldüğü gibi Gradient Boosting yöntemi ile tahmin edilen veri istasyonun ölçüm verilerine oldukça yakındır. En büyük sapma kirletici seviyesinin uç değerlerde olduğu ölçümlerde gözlenmektedir. Ancak yöntem genel olarak başarılı tahmin değerleri üretmektedir.

Aynı model bu kez PM_{2.5} ölçümü yapılmayan Kestel ve Beyazıt Caddesi Hava Kalitesi İzleme İstasyonları için kullanılmıştır. Modelde değişen tek husus, modele girdi olarak verilen kirletici verisinin ait olduğu istasyondur. Sırasıyla önce Kestel istasyonuna ait daha sonra da Beyazıt Caddesine ait tüm kirletici verileri ile meteorolojik verileri modele aktarılmıştır.

Şekil 5.21 ve Şekil 5.22’de de sırasıyla Kestel ve Beyazıt Caddesi istasyonlarına ilişkin PM_{10} ölçüm ve tahmin verilerinin grafiklerine yer verilmiştir. Tahmin verisinin, gerçek verilere oldukça sonuçlar ürettiği görülmektedir.

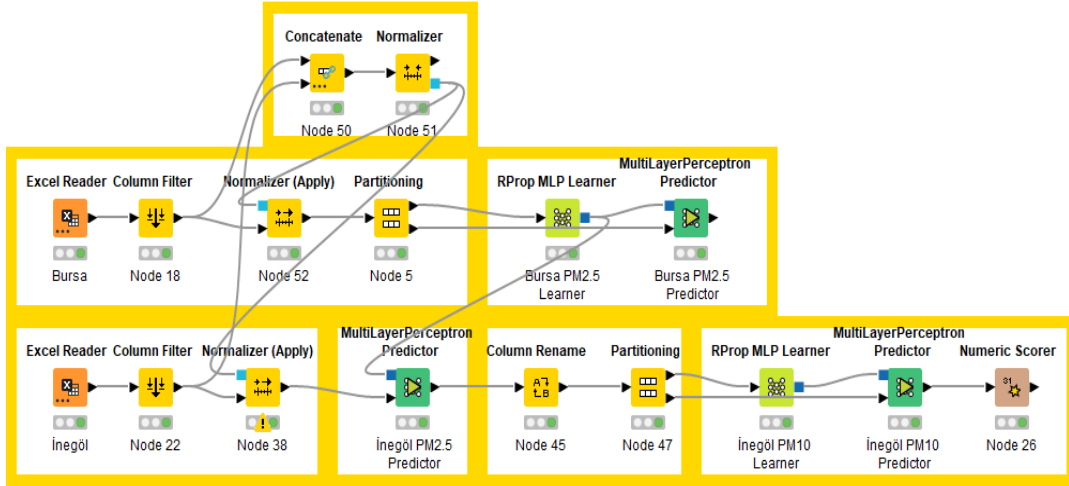


Şekil 5.21 : Kestel PM_{10} ölçüm ve GB ile tahmin verisi.



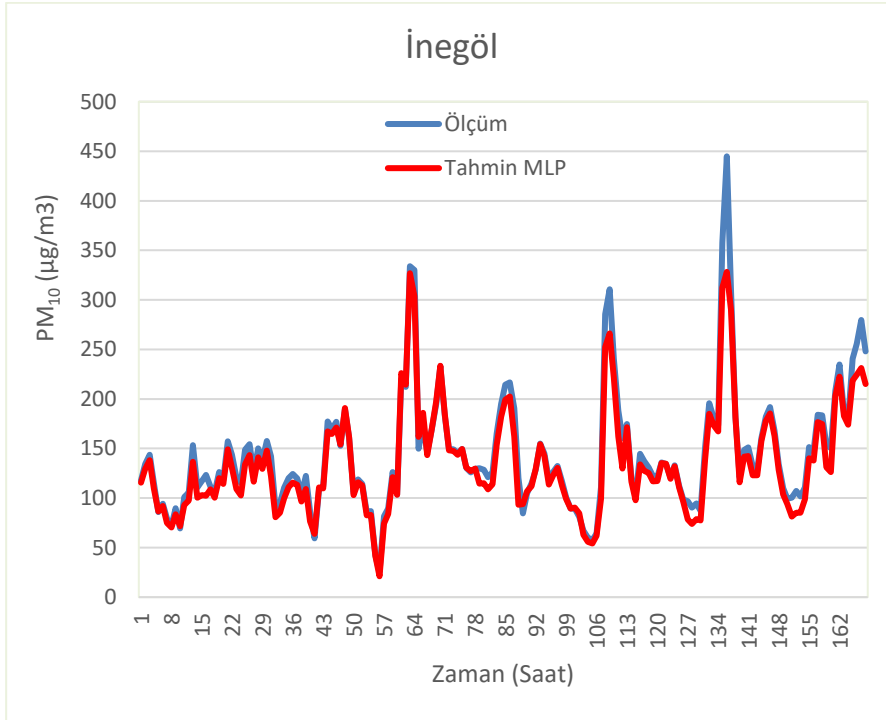
Şekil 5.22 : Beyazıt Caddesi PM_{10} ölçüm ve GB ile tahmin verisi.

Gradient Boosting ile oluşturulan modelde tüm istasyonlar için hesaplama yapıldıktan sonra aynı süreç bu kez MultiLayer Perceptron ile Şekil 5.23'te oluşturulan modelde işletilmiştir.

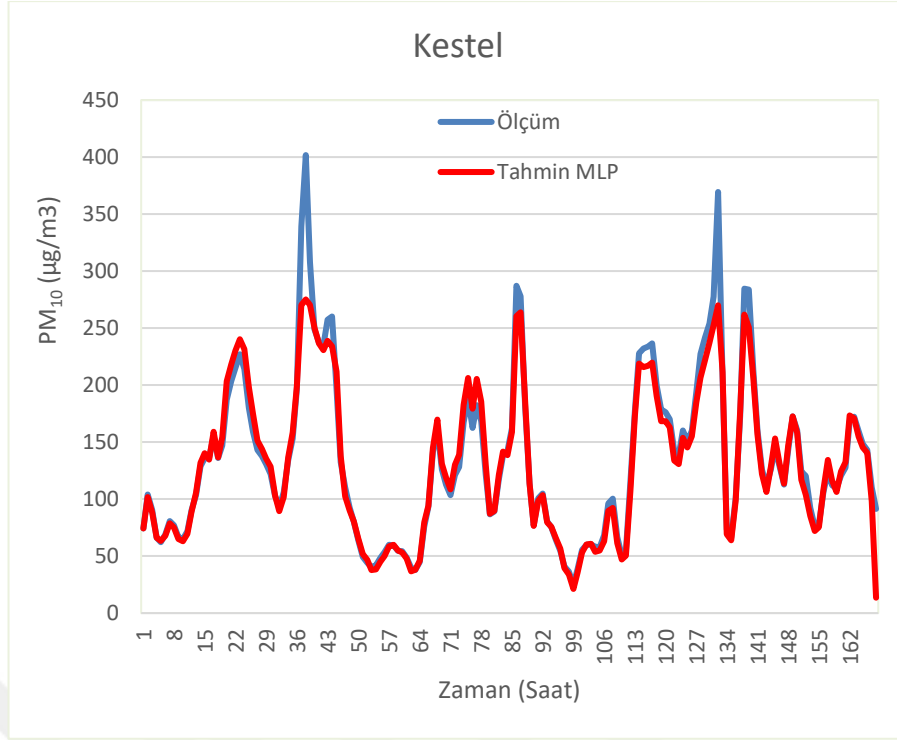


Şekil 5.23 : MLP yöntemi ile $PM_{2.5}$ ve PM_{10} tahmin modeli.

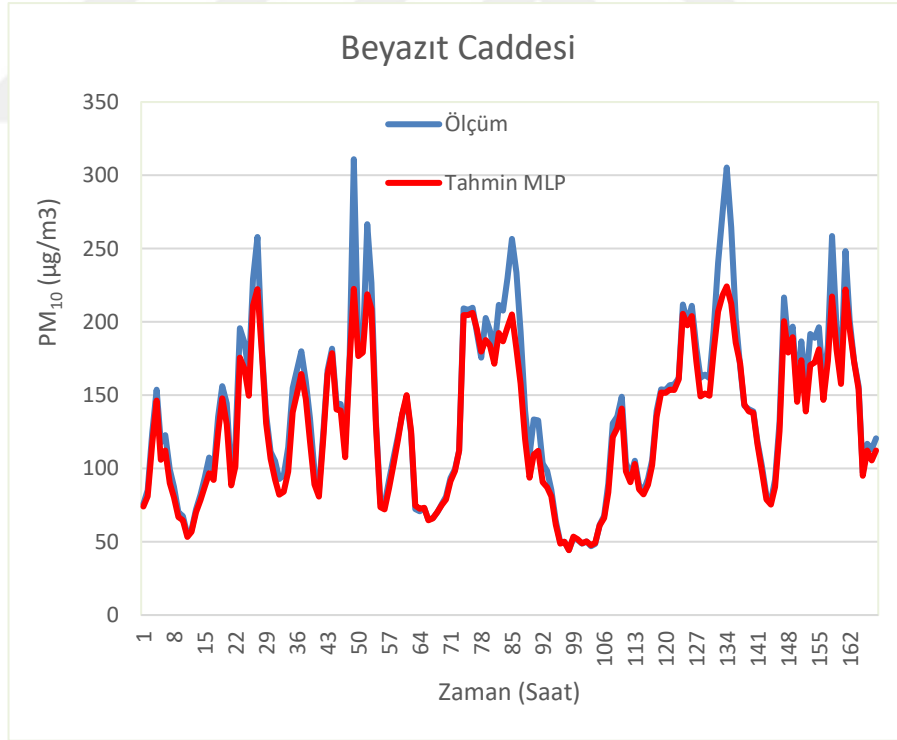
Modelin İnegöl, Kestel ve Beyazıt Caddesi istasyonları için ürettiği sonuçların ölçüm değerleri ile karşılaştırmalı grafikleri Şekil 5.24, Şekil 5.25 ve Şekil 5.26'da verilmiştir.



Şekil 5.24 : İnegöl PM_{10} ölçüm ve MLP ile tahmin verisi.



Şekil 5.25 : Kestel PM₁₀ ölçüm ve MLP ile tahmin verisi.



Şekil 5.26 : Beyazıt Caddesi PM₁₀ ölçüm ve MLP ile tahmin verisi.

Tüm istasyonlar için modellerin hesapladığı hata metrikleri Çizelge 5.5'teki gibidir. Kestel istasyonunda Gradient Boosting, İnegöl ve Beyazıt Caddesi istasyonlarında MultiLayer Perceptron yöntemiyle oluşturulan model daha yüksek performans göstermiştir.

Çizelge 5.5 : GB ve MLP yöntemlerine ait hata metrikleri.

Hata Metriği	İnegöl GB	İnegöl MLP	Kestel GB	Kestel MLP	Beyazıt GB	Beyazıt MLP
R^2	0,904	0,934	0,951	0,94	0,888	0,916
Ortalama Mutlak Hata	0,012	0,01	0,021	0,019	0,04	0,033
Ortalama Kare Hata	0	0	0,001	0,001	0,004	0,003
Kök Ortalama Kare Hata	0,020	0,016	0,035	0,039	0,063	0,054

Yukarıda yer verilen yöntemlerle istasyonlarda ölçülmediği halde oluşturulan modelle tahmin edilen PM_{2.5} verileri de dahil edilerek İnegöl istasyonu için kirletici veri seti örneği Çizelge 5.6'da verilmiştir.

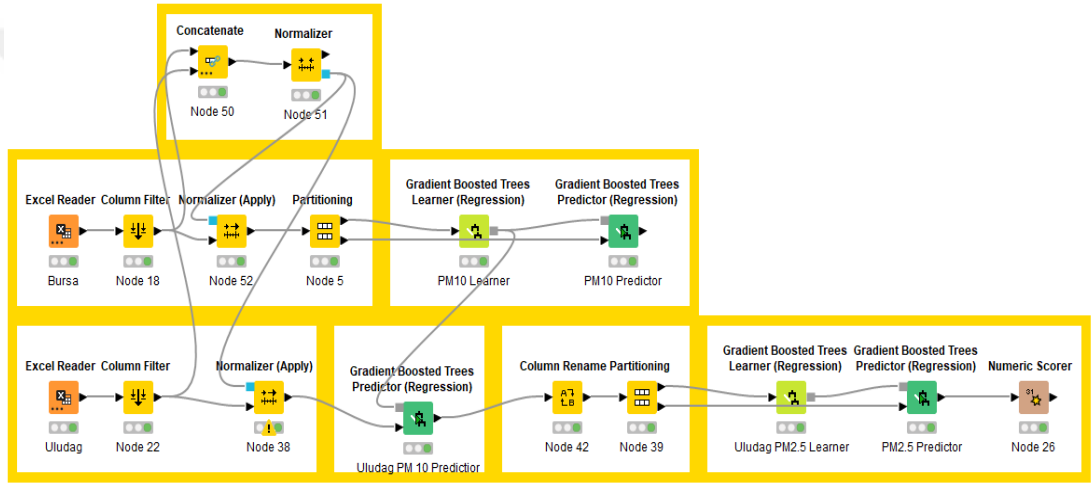
Çizelge 5.6 : İnegöl Hava Kalitesi İzleme İstasyonu tam veri seti.

PM ₁₀	PM _{2.5} GB	PM _{2.5} MLP	NO ₂	SO ₂
117,81	85,66	84,25	41,42	8,36
134,36	108,66	96,21	36,66	6,97
143,68	108,69	102,95	38,90	8,14
115,33	83,36	78,92	40,06	8,01
85,76	53,83	57,22	34,77	8,37
94,23	60,53	66,80	33,23	9,48
79,26	55,34	60,17	43,78	9,96
71,89	46,65	54,93	43,34	11,56
89,78	61,73	69,55	57,50	9,97
69,37	51,78	50,56	51,33	13,63

5.3.1.3 Uludağ Üniversitesi ve Kültürpark İstasyonları için PM₁₀ tahmini

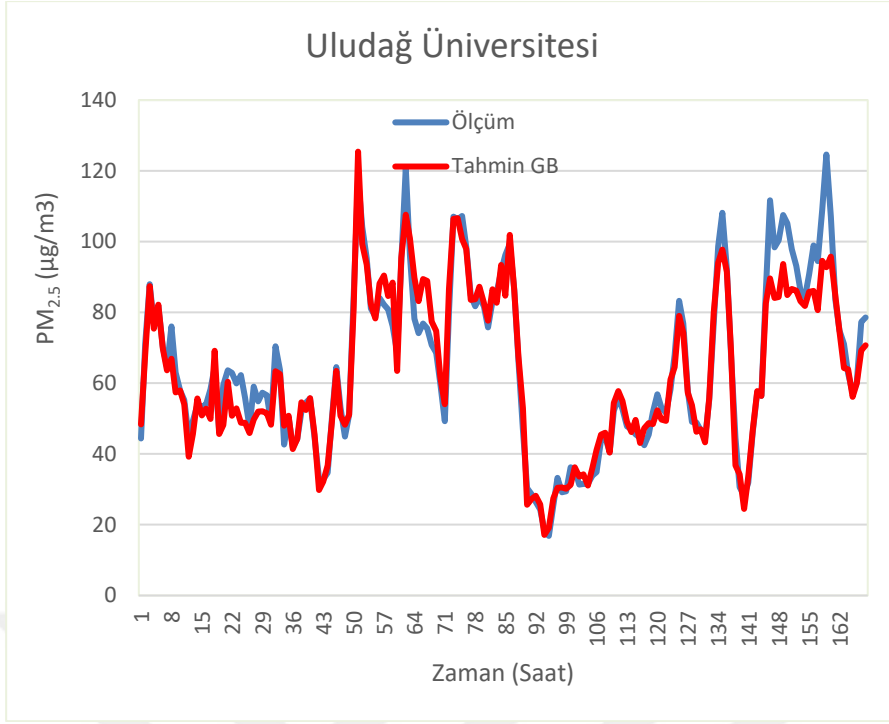
Uludağ Üniversitesi ve Kültürpark adlı hava kalitesi izleme istasyonlarında PM₁₀ ölçümü yapılmamaktadır. Bu bölümde bu kez Bursa istasyonuna ait PM₁₀ eğitim düğümü kullanılarak ölçüm yapılmayan istasyonlar için PM₁₀ tahmini denenmiştir. Bölüm 5.3.1.2'de kullanılan modeller aynı kalmak kaydıyla sadece Learner düğümlerinde PM_{2.5} yerine PM₁₀ seçimi yapılmıştır. Model sonucunda elde edilen verilerin tutarlılığını test etmek için bu kez tahmin edilen PM₁₀ verisi, ölçüm verisi olmayan istasyonlara ait veri setine ölçüm verisi gibi eklenmiş ve geleceğe dönük PM_{2.5} tahmini elde edilmiştir. Son aşamada ölçülen ve tahmin edilen PM_{2.5} verilerinin performansı değerlendirilmiştir. Bölüm 5.3.1.2'de verilenlere benzer şekilde, bu

bölümde de iki ayrı hipotez bulunmaktadır: Birinci hipotez, ölçüm yapılmayan istasyon için başka bir istasyonun PM_{10} eğitim setinin kullanılması yoluyla yapılan $PM_{2.5}$ tahmini başarılıdır. İkinci hipotez ise, modelin ilk aşamasında elde edilen PM_{10} tahmin verisi de eklenerek aynı istasyon için yapılan $PM_{2.5}$ tahmin verisi ile $PM_{2.5}$ ölçüm verisi yüksek korelasyon gösterirse bu PM_{10} tahmininin başarılı olduğunun göstergesidir. Şekil 5.27’de Gradient Boosting yöntemi ile Bursa istasyonuna ait $PM_{2.5}$ eğitim verisi kullanılarak önce Uludağ Üniversitesi istasyonu için PM_{10} tahmin verisi ve sonra bu veri de eklenerek aynı istasyon için $PM_{2.5}$ tahmini yapılmıştır. Daha sonra $PM_{2.5}$ tahmin performansı hata metrikleri ile değerlendirilmiştir. Aynı model Kültürpark istasyonu verilerine de uygulanmıştır.

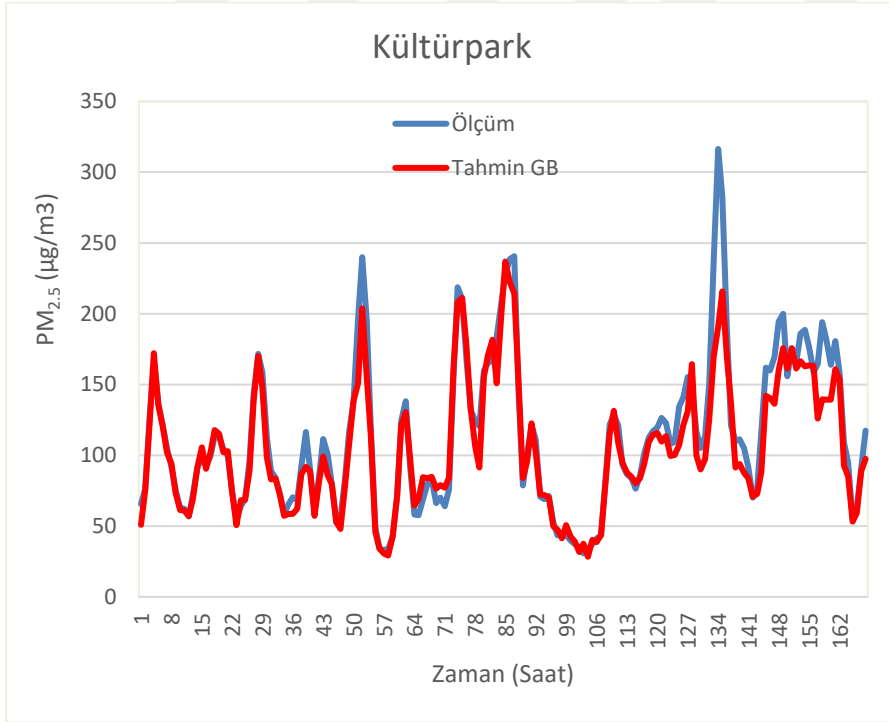


Şekil 5.27 : GB yöntemi ile eğitim verisi kullanarak PM_{10} ve $PM_{2.5}$ tahmini.

Uludağ Üniversitesi ve Kültürpark Hava Kalitesi İzleme İstasyonlarına ait verilerin Gradient Boosting yöntemine dayalı model sonuçları Şekil 5.28 ve Şekil 5.29’daki gibi ölçümlerle karşılaştırılmıştır. Grafiklerden görüldüğü üzere tahmin sonuçları ölçüm değerlerine oldukça yakındır. Genel eğilim gerçek konsantrasyon değerlerinin seyriyle oldukça uyumludur.

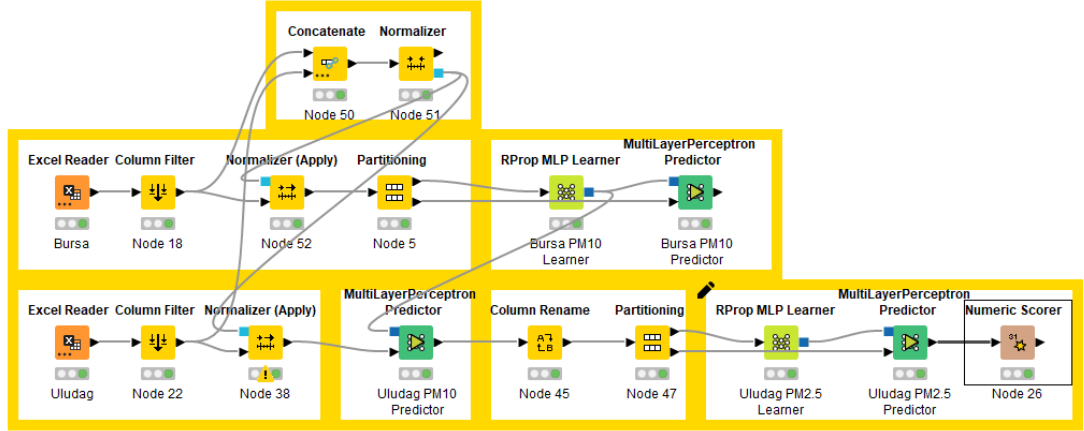


Şekil 5.28 : Uludağ Üniversitesi PM_{2.5} ölçüm ve GB ile tahmin verisi.



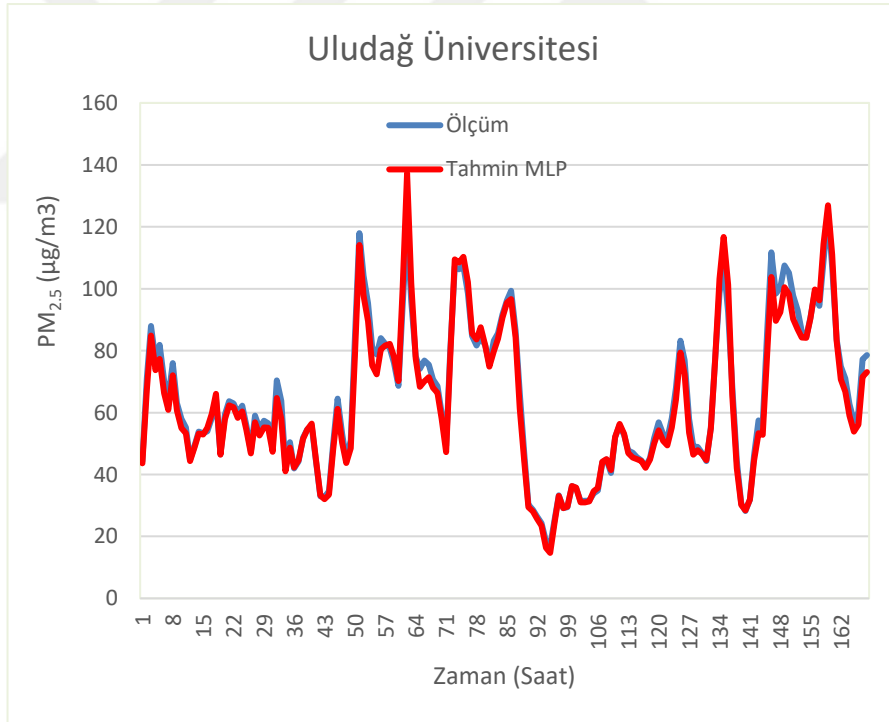
Şekil 5.29 : Kültürpark PM_{2.5} ölçüm ve GB ile tahmin verisi.

Şekil 5.30’da ise aynı istasyonlarda MultiLayer Perceptron yöntemiyle ilk adımda PM_{10} ve ikinci adımda $PM_{2.5}$ tahmini için oluşturulan model gösterilmektedir.

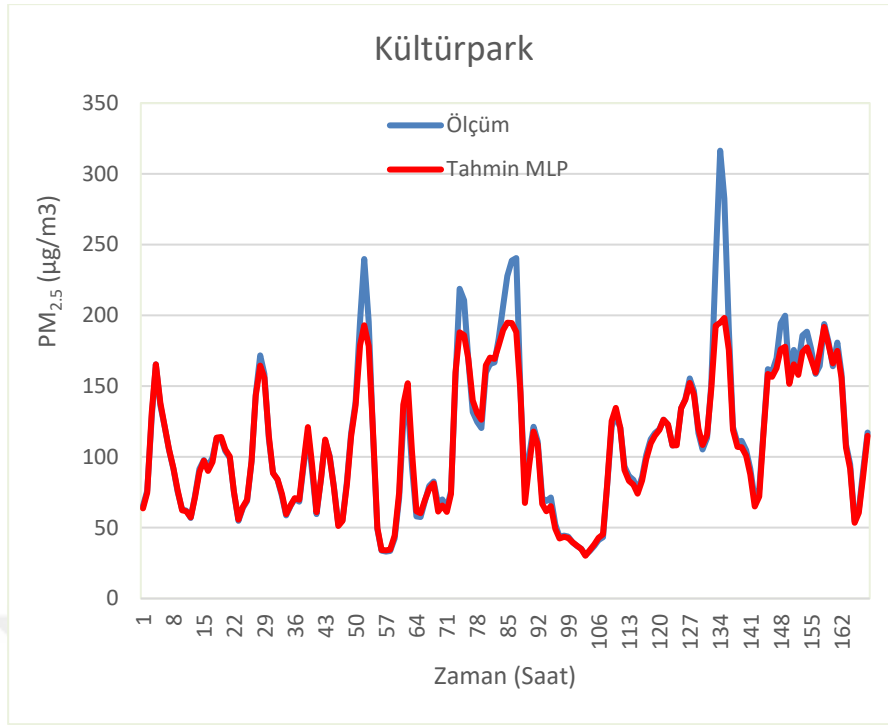


Şekil 5.30 : MLP yöntemi ile eğitim verisi kullanarak PM_{10} ve $PM_{2.5}$ tahmini.

Uludağ Üniversitesi ve Kültürpark istasyonlarına ait MLP yöntemiyle tahmin sonuçları ile ölçüm verileri Şekil 5.31 ve Şekil 5.32’de görsel olarak karşılaştırılmıştır.



Şekil 5.31 : Uludağ Üniversitesi $PM_{2.5}$ ölçüm ve MLP ile tahmin verisi.



Şekil 5.32 : Kültürpark PM_{2.5} ölçüm ve MLP ile tahmin verisi.

Uludağ Üniversitesi ve Kültürpark Hava Kalitesi İzleme İstasyonları için yapılan PM_{2.5} tahminlerinde kullanılan yöntemlerin hata metrikleri Çizelge 5.7’de verilmiştir. Her iki yöntemle de başarılı sonuçlar edilse dahi MultiLayer Perceptron yönteminin kullanıldığı modelin performansı Gradient Boosting yöntemini kullanan modele göre daha yüksektir. MultiLayer Perceptron yönteminin PM_{2.5} tahminlerinin tahmininde daha başarılı olması, modelin ilk adımda ürettiği PM₁₀ tahminlerinin daha isabetli olduğunun bir göstergesi olarak değerlendirilmektedir.

Çizelge 5.7 : GB ve MLP yöntemlerine ait hata metrikleri.

Hata Metriği	Uludağ Ü. GB	Uludağ Ü. MLP	Kültürpark GB	Kültürpark MLP
R ²	0,928	0,979	0,884	0,923
Ortalama Mutlak Hata	0,018	0,001	0,034	0,019
Ortalama Kare Hata	0,001	0	0,003	0,02
Kök Ortalama Kare Hata	0,026	0,014	0,058	0,047

6. SONUÇ VE ÖNERİLER

Bu çalışmada amaç nüfus, kentleşme ve sanayileşme bakımından Türkiye'nin en büyük şehirlerinden biri olan Bursa için sürdürülebilir bir hava kalitesi tahmin modeli oluşturmaktır. Tahmin modelinde kısa vadede solunum yolu rahatsızlıklarına, uzun vadede daha ciddi sağlık problemlerine neden olan partiküler madde konsantrasyonlarının tahminine odaklanılmıştır. İleriki çalışmalarda diğer kirleticiler için benzer konfigürasyonla bir tahmin modeli oluşturulması, akabinde tüm kirletici tahmin modellerinin kombinasyonu ile modellerin bütünleşik bir hale getirilmesi hedeflenmektedir.

Çalışmada ilk aşamada kirletici verilerinde büyük bir sorun oluşturan ve tahmin modellerinin başarısını da doğrudan etkileyen eksik veri tamamlama üzerinde durulmuştur. Altı istasyonda ölçülen tüm kirletici verileri eksik veri sayıları, eksik verilerin dönerselliği, aynı istasyondaki diğer kirleticilerle aynı saatlerde eksik ölçümün bulunup bulunmadığı, farklı istasyonlarda aynı kirletici türü için aynı zaman aralıklarında eksik verinin olup olmadığı gibi hususlar gerek grafiklerle görselleştirilerek, gerekse veri dosyalarında filtreleme yapılarak incelenmiştir. Veri tamamlama işlemlerinde farklı istasyonlarda ölçülen aynı kirletici türlerinin korelasyonu göz önünde bulundurularak, öncelikle Uludağ Üniversitesi istasyonundaki eksik $PM_{2.5}$, NO_2 ve O_3 verileri lineer regresyon yöntemi ile tamamlanmıştır. Veri tamamlama başarılı olmakla birlikte, istasyonların tümünde aynı zaman aralığında eksik veri bulunması durumunda yöntem kullanılamamaktadır. Bu nedenle eksik verilerin tamamlanması için başka bir yöntem arayışına gidilmiştir. Aynı tür veri setinde kendi içindeki davranışları analiz ederek veri tamamlamayan XGBoost makine öğrenmesi yöntemi tam bir veri setinde eksik veri yaratılarak denenmiştir. Sonuçlar başarılı bulunmuş ve tüm kirletici türlerine ait veri setlerinde veri tamamlama için bu yöntem kullanılmıştır.

Çalışmanın ana konusu olan tahmin modeli ise iki hipotez üzerinden oluşturulmuştur. Birinci durumda, her iki partiküler maddenin ($PM_{2.5}$ ve PM_{10}) ölçümünün yapıldığı Bursa istasyonu için geliştirilen modelle bu maddelerin gelecek bir haftalık süre için

başarılı tahmininin yapılabileceği değerlendirilmiştir. Modelde istasyona ait tüm kirletici verileri ve meteorolojik veriler girdi olarak kullanılmıştır. Model hem Gradient Boosting ve hem de çok katmanlı yapay sinir ağı MultiLayer Perceptron yöntemleri ile oluşturulmuştur. Her iki yöntemde de tüm veriler, eğitim verisi ve test verisi olarak ikiye ayrılmış ve eğitim verisi ile model eğitilmiştir. Eğitilen model ile yöntemlerin tahmin düğümleri (node) çalıştırılmıştır. Tahmin ile test verisi dört hata metriği ile karşılaştırılmıştır. Hata metriklerine göre, Bursa istasyonu için MultiLayer Perceptron yöntemine dayalı model, Gradient Boosting yöntemine dayalı modelden daha başarılı olmuştur.

İkinci durumda, istasyonda partiküler madde türlerinden yalnız biri ölçülüyorsa, diğer partiküler maddenin istasyonda ölçülen kirletici verileri ve meteorolojik veriler ile başarılı şekilde hesaplanabileceği bir model üzerinde durulmuştur. İki adımlı modelde öncelikle partiküler madde türlerinden yalnız PM_{10} ölçümü yapılan istasyonlar için $PM_{2.5}$ tahmini yapılmıştır. İkinci adımda modelin başarısını ölçmek için tahmin edilen $PM_{2.5}$ verisi, ölçüm verisi gibi modüle verilmiş ve bu kez PM_{10} tahmini yapılmıştır. PM_{10} ölçüm verisi ile PM_{10} tahmin verisi hata metrikleri ile karşılaştırılmıştır. Model sonuçları değerlendirilmiş, hem Gradient Boosting ve hem de MultiLayer Perceptron yöntemlerinin gerçek değerlere yakın sonuç ürettiği görülmüştür. Her iki yöntem de başarılı olsa da İnegöl ve Beyazıt Caddesi istasyonları için MultiLayer Perceptron, Kestel istasyonu için ise Gradient Boosting daha iyi performans göstermiştir. Modelin ikinci adımı sonucunda tahmin edilen PM_{10} 'un ölçüm değerine yakın olmasının istasyonda ölçülmeyen ancak ilk adım sonucunda model tarafından hesaplanan $PM_{2.5}$ değerinin başarılı olduğu hipotezini doğruladığı değerlendirilmiştir.

Benzer model $PM_{2.5}$ ölçümü yapıldığı halde PM_{10} ölçümü yapılmayan Uludağ Üniversitesi ve Kültürpark istasyonları için çalıştırılmıştır. Model sonuçlarına göre her iki istasyonda da MultiLayer Perceptron yöntemi ile tahminlerin çok daha başarılı olduğu tespit edilmiştir.

Tüm sonuçlar birlikte değerlendirildiğinde, istasyonlarda ölçülmeyen PM_{10} ya da $PM_{2.5}$ konsantrasyonları oluşturulan modelle her iki yapay zeka yönteminin de tahmin performansı yüksek bulunmuştur. Model sonuçlarına göre MultiLayer Perceptron genel olarak daha başarılı sonuçlar üretmiştir. Geliştirilen modül, basit bir bilgisayarda çalıştırılabilir ve günlük hava kalitesi izleme faaliyetlerine dahil edilebilir. Modüle kolaylıkla yeni veri aktarımı yapılabilir, veri seti kolaylıkla güncellenebilir.

Kentsel ortamlarda genelde hava kalitesi izleme istasyonları yetersizdir. Daha fazla istasyonda, ölçülmeyen kirleticiyi modülle tahmin etmek yoluyla maliyet azaltılabilir. Pek çok şehir dünya çapında sürdürülebilir hava kalitesi yönetimi için yoğun çaba sarf etmektedir. Oluşturulan model, kentsel alanlarda daha temiz hava için genel yönetim faaliyetlerine ve stratejilerine entegre edilebilir. Şehirlerdeki hava kirliliği büyük bir endişe kaynağı olduğundan, bu yaklaşım gerçek zamanlı bir uyarı sistemine dahil edilebilir.





KAYNAKLAR

- Alahamade, W., Lake, I., Reeves, C. E., & De La Iglesia, B. (2020).** Clustering imputation for air pollution data. *In International Conference on Hybrid Artificial Intelligence Systems* (pp. 585-597). Cham: Springer International Publishing.
- Alahamade, W., Lake, I., Reeves, C. E., & De La Iglesia, B. (2021).** Evaluation of multi-variate time series clustering for imputation of air pollution data. *Geoscientific Instrumentation, Methods and Data Systems Discussions, 2021*, 1-23.
- Altınçöp, H., & Oktay, A. B. (2018).** Air pollution forecasting with random forest time series analysis. *In 2018 International Conference on Artificial Intelligence and Data Processing (IDAP)* (pp. 1-5). IEEE.
- Ao, D., Cui, Z., & Gu, D. (2019).** Hybrid model of air quality prediction using k-means clustering and deep neural network. *In 2019 Chinese control conference (CCC)* (pp. 8416-8421). IEEE.
- Bacanlı, Ü. G., & Kargı, P. G. (2019).** Uzun ve kısa süreli periyotlarda kuraklık analizi: Bursa örneği. *Doğal Afetler ve Çevre Dergisi, 5*(1), 166-174.
- Berthold, M. R., Cebren, N., Dill, F., Gabriel, T. R., Kötter, T., Meinel, T., ... & Wiswedel, B. (2009).** KNIME-the Konstanz information miner: version 2.0 and beyond. *AcM SIGKDD explorations Newsletter, 11*(1), 26-31.
- Bolton, D. (1980).** The computation of equivalent potential temperature. *Monthly weather review, 108*(7), 1046-1053.
- Boznar, M., Lesjak, M., & Mlakar, P. (1993).** A neural network-based method for short-term predictions of ambient SO₂ concentrations in highly polluted industrial areas of complex terrain. *Atmospheric environment. Part B. urban atmosphere, 27*(2), 221-230.
- Bursa Yatırım Destek Ofisi. (t.y.).** Erişim: 03.10.2023, <https://bursainvest.gov.tr/organize-sanayi-bolgeleri/>
- Calle, M., Orellana, M., & Ortega-Chasi, P. (2020).** Narx neural network for imputation of missing data in air pollution datasets. *In Conference on Information and Communication Technologies of Ecuador* (pp. 226-240). Cham: Springer International Publishing.
- Castelli, M., Clemente, F. M., Popovič, A., Silva, S., & Vanneschi, L. (2020).** A machine learning approach to predict air quality in California. *Complexity, 2020*.
- Chen, Z. Y., Jin, J. Q., Zhang, R., Zhang, T. H., Chen, J. J., Yang, J., Ou, C.Q., & Guo, Y. (2020).** Comparison of different missing-imputation methods for MAIAC (Multiangle Implementation of Atmospheric

Correction) AOD in estimating daily PM_{2.5} levels. *Remote Sensing*, 12(18), 3008.

Cheng, B., Ma, Y., Feng, F., Zhang, Y., Shen, J., Wang, H., Guo, Y. & Cheng, Y. (2021). Influence of weather and air pollution on concentration change of PM_{2.5} using a generalized additive model and gradient boosting machine. *Atmospheric Environment*, 255, 118437.

Choi, W., Ho, C. H., & Kim, K. Y. (2023). Critical contribution of moisture to the air quality deterioration in a warm and humid weather. *Scientific Reports*, 13(1), 4260.

Çalışkan, B., Özengin, N., & Cindoruk, S. S. (2020). Air quality level, emission sources and control strategies in Bursa/Turkey. *Atmospheric Pollution Research*, 11(12), 2182-2189.

Demir, B. A. (2020). *Bursa İlinin Meteorolojik Verilerinin İncelenmesi* (Yüksek lisans tezi). Bursa Uludağ Üniversitesi, Fen Bilimleri Enstitüsü, Bursa.

Dung, N. A., Son, D. H., & Tri, D. Q. (2019). Effect of Meteorological Factors on PM₁₀ Concentration in Hanoi, Vietnam. *Journal of Geoscience and Environment Protection*, 7(11), 138.

Garipağaoğlu, N., & Duman, C. (2017). Bursa kenti hava kalitesinin zaman içerisindeki değişimi. *Marmara Coğrafya Dergisi*, (36), 57-70.

Gönençgil, B., & Sungur, K. A., (1997). Çeşitli İklim Elemanlarının Hava Kirliliği Üzerine Etkileri. *A.Ü. Türkiye Coğ. Araş. Ve Uyg. Merkezi Dergisi*, no.6, 337-345.

Güven, A. (2022). *Hava Kirliliğinin Makine Öğrenmesi Yöntemleriyle Tahmini*. (Yüksek lisans tezi). Bursa Uludağ Üniversitesi, Fen Bilimleri Enstitüsü, Bursa.

Erdun, H., Öztürk, A., Çapraz, Ö., Toros, H., Dursun, S., & Deniz, A. (2015). Spatial variation of PM₁₀ in Turkey. In *7th Atmospheric Sciences Symposium, Istanbul, Turkey* (pp. 311-323).

Flick, C. (1980). The movement for smoke abatement in 19th-century Britain. *Technology and Culture*, 29-50.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.

Hadeed, S. J., O'Rourke, M. K., Burgess, J. L., Harris, R. B., & Canales, R. A. (2020). Imputation methods for addressing missing data in short-term monitoring of air pollutants. *Science of the Total Environment*, 730, 139140.

Helfand, W. H., Lazarus, J., & Theerman, P. (2001). Donora, Pennsylvania: an environmental disaster of the 20th century. *American Journal of Public Health*, 91(4), 553.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D.,

- Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2023):** ERA5 hourly data on pressure levels from 1940 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*, erişim tarihi: 30.08.2023, 03.10.2023.
- Hojamkulyyeva, A., Dursun, Ş., & Toros, H., (2020).** Relationship Between Particulate Air Pollution and Meteorological Variables in Kocaeli. *Journal of Research in Atmospheric Science* , vol.2, no.1, 17-23.
- Jiang, D., Zhang, Y., Hu, X., Zeng, Y., Tan, J., & Shao, D. (2004).** Progress in developing an ANN model for air pollution index forecast. *Atmospheric Environment*, 38(40), 7055-7064.
- Jovova, L., & Trivodaliev, K. (2021).** Air pollution forecasting using CNN-LSTM deep learning model. In *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)* (pp. 1091-1096). IEEE.
- Junninen, H., Niska, H., Tuppurainen, K., Ruuskanen, J., & Kolehmainen, M. (2004).** Methods for imputation of missing values in air quality data sets. *Atmospheric environment*, 38(18), 2895-2907.
- Kadioğlu, M. (2012).** Türkiye’de İklim Değişikliği Risk Yönetimi. *Türkiye’nin İklim Değişikliği II. Ulusal Bildiriminin Hazırlanması Projesi Yayını*, 172 sf. 96.
- Kara, Y., Dursun, Ş., & Toros, H. (2020).** Assessment of Air Pollution in Trabzon During COVID-19 Measures. *Journal of Research in Atmospheric Science*, 2(1).
- Kılıç, Y., Dursun, Ş., & Toros, H. (2020).** Effect of COVID-19 Outbreak on Particulate Matter Pollution in Istanbul City Centre. *Journal of Research in Atmospheric Science*, 2(2).
- Lee, M., Lin, L., Chen, C. Y., Tsao, Y., Yao, T. H., Fei, M. H., & Fang, S. H. (2020).** Forecasting air quality in Taiwan by using machine learning. *Scientific reports*, 10(1), 4153.
- Lehtomäki, H., Korhonen, A., Asikainen, A., Karvosenoja, N., Kupiainen, K., Paunu, V. V., Savolahti, M., Sofiev, M., Palamarchuk, Y., Karppinen, A., Kukkonen, J., & Hänninen, O. (2018).** Health impacts of ambient air pollution in Finland. *International Journal of Environmental Research and Public Health*, 15(4), 736.
- Lelieveld, J. (2020).** Reducing Air Pollution: Avoidable Health Burden. *Health of People, Health of Planet and Our Responsibility*, 1, 105-120.
- Li, Y., Sha, Z., Tang, A., Goulding, K., & Liu, X. (2023).** The application of machine learning to air pollution research: A bibliometric analysis. *Ecotoxicology and Environmental Safety*, 257, 114911.
- Liu, H., Ma, W., Qian, J., Cai, J., Ye, X., Li, J., & Wang, X. (2015).** Effect of urbanization on the urban meteorology and air pollution in Hangzhou. *Journal of Meteorological Research*, 29(6), 950-965.

- Liu, X., Wang, X., Zou, L., Xia, J., & Pang, W. (2020).** Spatial imputation for air pollutants data sets via low rank matrix completion algorithm. *Environment international*, 139, 105713.
- Liu, Y., Zhou, Y., & Lu, J. (2020).** Exploring the relationship between air pollution and meteorological conditions in China under environmental governance. *Scientific reports*, 10(1), 14518.
- Longhurst, J. W. S., Barnes, J. H., Chatterton, T. J., Hayes, E. T., & Williams, W. B. (2016).** Progress with air quality management in the 60 years since the UK clean air act, 1956. Lessons, failures, challenges and opportunities. *International Journal of Sustainable Development and Planning*, 11(4), 491-499.
- Lou, M., Zhou, Q., Jin, J., Peng, Y., Dai, R., Zhang, Y., & Guo, J. (2019).** The linkages between atmospheric boundary layer and PM_{2.5} from different region in China. In *2019 International Conference on Meteorology Observations (ICMO)* (pp. 1-4). IEEE.
- May, R. M., Goebbert, K. H., Thielen, J. E., Leeman, J. R., Camron, M. D., Bruick, Z., Bruning, E. C., Manser, R. P., Arms, S. C., and Marsh, P. T. (2022).** MetPy: A Meteorological Python Library for Data Analysis and Visualization. *Bull. Amer. Meteor.Soc.*, 103, E2273-E2284
- Miskell, G., Pattinson, W., Weissert, L., & Williams, D. (2019).** Forecasting short-term peak concentrations from a network of air quality instruments measuring PM_{2.5} using boosted gradient machine models. *Journal of environmental management*, 242, 56-64.
- Narasimhan, D., & Vanitha, M. (2023).** Machine Learning Approach-based Big Data Imputation Methods for Outdoor Air Quality Forecasting. *Journal of Scientific & Industrial Research*, 82(03), 338-347.
- Ozelkan, E., Karaman, M., Mostamandy, S., Avci, Z. D. U., & Toros, H. (2015).** Derivation of PM₁₀ Levels Using OBRA on Landsat-5TM Images: A Case Study in Izmir, Turkey. *Fresenius Environmental Bulletin*, 24(4), 1585-1596.
- Öncü, S. (2021).** CBS yöntemiyle Bursa'nın bütünleşik doğal tehlike analizi. (Yüksek lisans tezi). Bursa Uludağ Üniversitesi, Fen Bilimleri Enstitüsü, Bursa.
- Özkayalar, E., Dursun, S., & Toros, H. (2020).** Investigation of Air Pollution in Bursa City Centre. *Journal of Research in Atmospheric Science* Vol. 3, No. 1, pp.13-20, June 2021.
- Qiao, Z., Wu, F., Xu, X., Yang, J., & Liu, L. (2019).** Mechanism of spatiotemporal air quality response to meteorological parameters: A national-scale analysis in China. *Sustainability*, 11(14), 3957.
- Quinteros, M. E., Lu, S., Blazquez, C., Cárdenas-R, J. P., Ossa, X., Delgado-Saborit, J. M., ... & Ruiz-Rudolph, P. (2019).** Use of data imputation tools to reconstruct incomplete air quality datasets: A case-study in Temuco, Chile. *Atmospheric environment*, 200, 40-49.

- Pawul, M., & Śliwka, M. (2016).** Application of artificial neural networks for prediction of air pollution levels in environmental monitoring. *Journal of Ecological Engineering*, 17(4), 190-196.
- Pérez, I. A., García, M. Á., Sánchez, M. L., Pardo, N., & Fernández-Duque, B. (2020).** Key points in air pollution meteorology. *International Journal of Environmental Research and Public Health*, 17(22), 8349.
- Riedmiller, M., & Braun, H. (1993, March).** A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In *IEEE international conference on neural networks* (pp. 586-591). IEEE.
- Sarkar, N., Gupta, R., Keserwani, P. K., & Govil, M. C. (2022).** Air Quality Index prediction using an effective hybrid deep learning model. *Environmental Pollution*, 315, 120404.
- Sassi A., Ivanov S., Toros H., Dursun, S. (2020).** The effect of particulate matter pollution of saharan dust over Europe in May-2020: a case study of Karaman City Center, Turkey. *Pakistan Journal of Analytical & Environmental Chemistry*, 21(2), 202-208.
- Saylan, L., Toros, H., & Sen, O. (2009).** Back trajectory analysis of precipitation chemistry in the urban and forest areas of Istanbul, Turkey. *CLEAN–Soil, Air, Water*, 37(2), 132-135.
- Shayia, A., Toros, H., Sağdıç, M., & Dursun, Ş. (2021).** Investigation of the effect of corona virus outbreak measures on atmospheric PM values for Konya city centre. *Journal of International Environmental Application and Science*, 16(2), 53-62.
- Siwek, K., & Osowski, S. (2016).** Data mining methods for prediction of air pollution. *International Journal of Applied Mathematics and Computer Science*, 26(2), 467-478.
- Srivastava, C., Singh, S., & Singh, A. P. (2018).** Estimation of air pollution in Delhi using machine learning techniques. In *2018 International Conference on Computing, Power and Communication Technologies (GUCON)* (pp. 304-309). IEEE.
- Sürdürülebilir Kalkınma Amaçları ve Göstergeleri Raporu, 2020.** T.C. Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı, Ankara
- Şahin, C., (1989).** Hava Kirliliği ve Hava Kirliliğini Etkileyen Doğal Çevre Faktörleri. *Coğrafya Araştırmaları (Atatürk Kültür, Dil ve Tarih Yüksek Kurumu, Coğrafya Bilim ve Uygulama Kolu Yayını)*, Cilt 1, Sayı 1, 25-45.
- Şerbetci, Z.S. (2017).** Bursa'nın Kentsel Gelişim Sürecinde Coğrafyanın Etkileri. (Doktora Tezi). İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul
- Toros, H., & Bağış, S. (2017).** Hava kirlilik modellerinde kullanılacak emisyon envanteri oluşturulması için yaklaşımlar ve İstanbul hava kirliliği dağılımı örneği. *Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 32(2), 1-12.
- Toros, H., Erdun, H., Çapraz, Ö., Betül, Ö. Z. E. R., Daylan, E. B., & Öztürk, A. İ. (2013).** Air pollution and quality levels in metropolitans of Turkey for sustainable life. *Avrupa Bilim ve Teknoloji Dergisi*, 1(1), 12-18.

- Türkiye İstatistik Kurumu, 2023.** TÜİK Haber Bülteni, Erişim: 03.10.2023, <https://data.tuik.gov.tr/Bulten/Index?p=49685>
- Unal, Y. S., Toros, H., Deniz, A., & Incecik, S. (2011).** Influence of meteorological factors and emission sources on spatial and temporal variations of PM₁₀ concentrations in Istanbul metropolitan area. *Atmospheric Environment*, 45(31), 5504-5513.
- Wang, Z., Chen, P., Wang, R., An, Z., & Qiu, L. (2023).** Estimation of PM_{2.5} concentrations with high spatiotemporal resolution in Beijing using the ERA5 dataset and machine learning models. *Advances in Space Research*, 71(8), 3150-3165.
- WHO Global Air Quality Guidelines, 2021.** WHO European Centre for Environment and Health, Germany.
- WHO, 2023.** Air pollution, Erişim: 18 Ocak 2023, https://www.who.int/health-topics/air-pollution#tab=tab_1
- Wijesekara, W. M. L. K. N., & Liyanage, L. (2020).** Comparison of imputation methods for missing values in air pollution data: Case study on Sydney air quality index. In *Advances in Information and Communication: Proceedings of the 2020 Future of Information and Communication Conference (FICC), Volume 2* (pp. 257-269). Springer International Publishing.
- Wilde, M. (2010).** The new directive on ambient air quality and cleaner air for Europe. *Environmental Law Review*, 12(4), 282-290.
- Yılmaz, M. (2022).** Türkiye Üzerinde ERA5 Saatlik Hava Sıcaklığı Verilerinin Doğrulanması. *Doğal Afetler ve Çevre Dergisi*, 8 (2), 207-220.
- Yiğiter, K., & Toros, H. (2020).** Relationship between air pollution and COVID-19 in Bursa, Turkey. *Journal of Research in Atmospheric Science*, 2(2).
- Zakaria, N. A., & Noor, N. M. (2018).** Imputation methods for filling missing data in urban air pollution data form Malaysia. *Urbanism. Arhitectura. Constructii*, 9(2), 159.
- Zhou, X., Tong, W., & Li, L. (2020).** Deep learning spatiotemporal air pollution data in China using data fusion. *Earth Science Informatics*, 13, 859-868.
- Url-1** < <https://bursa.ktb.gov.tr/TR-70229/cografya.html> >, erişim tarihi: 03.10.2023.
- Url-2**<https://unidata.github.io/MetPy/latest/api/generated/metpy.calc.relative_humidity_from_dewpoint.html#>, erişim tarihi: 30.08.2023.
- Url-3**<https://unidata.github.io/MetPy/latest/examples/calculations/Wind_Speed.html#>, erişim tarihi: 30.08.2023.
- Url-4**<https://sim.csb.gov.tr/STN/STN_Report/StationDataDownloadNew#>, erişim tarihi :25.10.2023).
- Url-5**<<https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview>>, erişim tarihi: 03.10.2023.
- Url-6**<<https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form>>, erişim tarihi: 30.08.2023.

ÖZGEÇMİŞ

Ad-Soyad : Saliha ÇELİKCAN BİLGİN

ÖĞRENİM DURUMU:

- **Lisans** : 1998, İstanbul Teknik Üniversitesi, Uçak ve Uzay Bilimleri Fakültesi, Meteoroloji Mühendisliği Bölümü

MESLEKİ DENEYİM VE ÖDÜLLER:

- 1998-2004 yılları arasında Pamukbank T.A.Ş.'de, 2004-2023 yılları arasında T.Halk Bankası A.Ş.'de çalışmıştır.

DİĞER YAYINLAR, SUNUMLAR VE PATENTLER:

- **Çelikcan S.**, Borhan Y. 1998. Artvin İli İçin Rüzgar Enerjisi Potansiyelinin Belirlenmesi. Lisans Bitirme Tezi.
- **Bilgin, S.Ç.**, Toros, H., Bilgin, T.T. (2023). Comparative Assessment of Imputation Techniques for Air Pollution Data in Bursa Province, Turkey. INCOHIS 2023 Autumn.