

**T.C.
HARRAN ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

YAPAY ZEKA TEKNİKLERİ KULLANARAK MEKANSAL ANALİZLER

Yunus Emre CENGİZ

HARİTA MÜHENDİSLİĞİ ANABİLİM DALI

**ŞANLIURFA
2023**

İÇİNDEKİLER

Sayfa No

ÖZET	i
ABSTRACT	ii
TEŞEKKÜR.....	iii
ŞEKİLLER DİZİNİ.....	iv
ÇİZELGELER DİZİNİ.....	v
1. GİRİŞ	1
2. KURAMSAL TEMELLER.....	3
2.1. Yapay Zekâ.....	3
2.1.1. Muhakeme ve problem çözüme (Arama yöntemleri)	3
2.1.1.1. Kör arama yöntemleri	4
2.1.1.2. Bilinçli arama yöntemleri.....	4
2.1.1.3. Mekânsal bağlamda arama yöntemlerine bakış.....	4
2.1.2. Makine öğrenmesi	5
2.1.2.1. Denetimli öğrenme.....	5
2.1.2.1.1. Tahmin (Regresyon).....	6
2.1.2.1.2. Sınıflandırma	7
2.1.2.2. Denetimsiz öğrenme	8
2.1.2.3. Pekiştirmeye dayalı öğrenme	9
2.1.2.4. Derin öğrenme	10
2.1.2.5. Mekânsal bağlamda makine öğrenmesinin değerlendirilmesi	11
2.1.3. Bilgi temsili dili	12
2.1.4. Robotik	13
2.1.4.1. Mekânsal bağlamda robotik biliminin değerlendirilmesi.....	13
2.1.5. Planlama	14
2.1.5.1. Mekânsal bağlamda planlama disiplininin değerlendirilmesi	14
2.1.6. Doğal dil işleme.....	14
2.1.7. Bilgisayar ile görü	15
2.1.7.1. Mekânsal bağlamda bilgisayar ile görünün değerlendirilmesi	15
3. MATERYAL ve YÖNTEM.....	17
3.1. Problemin Tanımı:	17
3.1.1. Hava kalitesi indeksi.....	17
3.1.2. Hava kirliliği kaynakları	18
3.1.2.1. Doğal kaynaklar	18
3.1.2.2. Yapay kaynaklar	19
3.1.3. Partikül maddeler.....	19
3.2. Problemin Çözümü:	20
3.3. Çalışma Alanının Belirlenmesi:	20
3.4. Bağımlı Değişken	22
3.5. Bağımsız Değişkenler	23
3.5.1. SAVI (Soil Adjusted Vegetation Index):	23
3.5.2. BU (Built-up Index):.....	23
3.5.3. Yer yüzeyi sıcaklığı:	24
3.5.4. Yollara yakınlık:	24
3.5.5. Elektromanyetik bantlar (Bant 2):	24
3.5.6. Elektromanyetik bantlar (Bant 3):	25
3.5.7. Elektromanyetik bantlar (Bant 4):	25

3.5.8. Elektromanyetik bantlar (Bant 5):	25
3.5.9. Elektromanyetik bantlar (Bant 6):	25
3.5.10. Fabrikaya yakınlık:	25
3.6. Elde Edilmesi Gereken Veriler	26
3.6.1. Uygulama alanındaki çok bantlı uydu görüntüsü	26
3.6.2. Uygulama alanındaki yol güzergahları	26
3.6.3. Çalışma alanı vektör verisi	26
3.6.4. Uygulama alanında bulunan hava kalitesi izleme istasyonu koordinatları ve pm10 değerleri	27
3.7. Kullanılan Makine Öğrenmesi Algoritması(Naif Bayes)	28
3.7.1. Bayes teoremi	28
3.7.2. Naif bayes	28
3.7.3. Sayısal örnek.....	30
3.8. Yapılan Uygulamanın Naif Bayes Algoritmasına Uyarlanması	31
3.9. Kullanılan Programlar	32
4. ARAŞTIRMA BULGULARI ve TARTIŞMA	33
4.1. Bağımsız Değişkenlere Dair Haritalar	33
4.2. Makinenin Öğrenmesi İçin Veri Seti Oluşturulması	43
4.3. Oluşturulan Veri Setine Dair Analizler	46
4.4. Oluşturulan Veri Setine Naif Bayes Algoritmasının Uygulanması	54
4.5. Oluşturulan Modelin Değerlendirilmesi	55
4.5.1. Hata matrisi.....	55
4.5.2. ROC(Receiver Operating Characteristic) eğrisi	56
4.6. Tahmin Haritaları	57
4.6.1. Sonuç haritasının incelenmesi	63
5. SONUÇ ve ÖNERİLER.....	66
KAYNAKÇA.....	68

ÖZET

Yüksek Lisans Tezi

YAPAY ZEKA TEKNİKLERİ KULLANARAK MEKANSAL ANALİZLER

Yunus Emre CENGİZ

Harran Üniversitesi
Fen Bilimleri Enstitüsü
Harita Mühendisliği Anabilim Dalı

Danışman: Dr. Öğretim Üyesi Fred Barış Ernst
Yıl: 2023, Sayfa: 70

Gelişen teknolojiye ve bilgisayarlardaki işlem gücünün artmasına bağlı olarak, yaşamın birçok alanında kullanılmaya başlanan yapay zekâ teknikleri, günümüzde çok geniş bir uygulama alanına ve büyük öneme sahiptir. Bu önem, var olan büyük veri ile baş edebilecek önemli araçların başında gelen, yapay zekâ tekniklerine olan ihtiyacı üssel bir şekilde artırmaktadır. İhtiyaca bağlı olarak, açık kaynak kodlu veri ile birlikte, kod paylaşım platformlarının kullanılmasıyla, yapay zekâ tekniklerine dair yeni bakış açılarıyla, yeni yöntemler geliştirilmektedir. Böyle bir ortamda, coğrafi nitelikli veri seti ile bu veri setlerinin, yapay zekâ teknikleri kullanılarak analizi, yapay zekâ dünyasında, pastanın büyük bir kısmını oluşturmaktadır. Bu yüzden, coğrafi nitelikli veri setinin, yapay zekâ tekniklerine uygun ve verimli bir şekilde uygulayabilmek adına, gelişmiş yapay zekâ tekniklerine dair iç dinamikler ile konum ve mekân kavramının iç dinamiklerinin birleştirilmesi suretiyle, verimli şekilde karar alabilen modeller ve sistemler geliştirmenin, günümüzde büyük bir önemi vardır. Buna bağlı olarak, mekânsal problemin tanımlanması, bu mekânsal probleme dair çözüm yolları sunabilecek nitelikte, geleneksel yöntemler ile karşılaştırıldığında zaman ve maliyet açısından daha hesaplı olan uygun yapay zekâ tekniğinin, araştırılarak ortaya konulması gerekmektedir. Buradan yola çıkarak, mekân ve konum alanında uzman olan kişilerin, yapay zekâ kavramını, doğru bir hiyerarşi ve sınıflandırma altında öğrenmesi suretiyle, yapay zekânın konum bazlı uygulamalarında kullanılacak yapay zekâ tekniklerini kullanılmasıyla, günümüz problemlerine çözümler ortaya konulabilir (Cengiz Y. E., Ernst F. B., 2022). Bu çalışmada, öncelikle yapay zekâ tekniklerinin, doğru ve anlaşılır bir hiyerarşi altında, anlaşılır tanımları yapılmıştır. Sonrasında ise, bu tanım içerisinde, mekâna ve konuma dayalı uygulamalarda kullanılacak olan yapay zekâ tekniklerine, mekânsal açıdan genel bir bakış yapılarak, bu tekniklerden biri kullanılarak örnek bir uygulama yapılmıştır.

ANAHTAR KELİMELER: Mekânsal, Yapay Zekâ, Modelleme, Analiz, Coğrafi

ABSTRACT

MSc Thesis

GEOSPATIAL ANALYSES BY USING ARTIFICIAL INTELLIGENCE TECHNIQUES

Yunus Emre CENGİZ

**Harran University
Graduate School of Natural and Applied Sciences
Department of Geomatics Engineering**

**Supervisor: Dr. Faculty Member Fred Barış ERNST
Year: 2023, Page: 70**

Artificial intelligence techniques, which have been used in many areas of life due to the developing technology and the increase in processing power in computers, have a wide application area and great importance today. This importance exponentially increases the need for artificial intelligence techniques, which is one of the important tools that can cope with the existing big data. Depending on the need, new methods are developed with new perspectives on artificial intelligence techniques by using code sharing platforms together with open source data. In such an environment, the geographical data set and the analysis of these data sets using artificial intelligence techniques constitute a large part of the cake in the world of artificial intelligence. Therefore, it is of great importance today to develop models and systems that can make efficient decisions by combining the internal dynamics of advanced artificial intelligence techniques with the internal dynamics of the concept of location and space, in order to apply the geographical data set in an appropriate and efficient manner to artificial intelligence techniques. Accordingly, it is necessary to define the spatial problem, to provide solutions for this spatial problem, and to investigate and reveal the appropriate artificial intelligence technique, which is more economical in terms of time and cost compared to traditional methods. From this point of view, solutions to today's problems can be revealed by using artificial intelligence techniques that can be used in location-based applications of artificial intelligence by learning the concept of artificial intelligence under a correct hierarchy and classification by experts in the field of space and location (Cengiz Y. E., Ernst F. B., 2022). In this study, first of all, clear definitions of artificial intelligence techniques were made under a correct and understandable hierarchy. Afterwards, within this definition, a spatial overview of the artificial intelligence techniques that can be used in applications based on space and location was made, and an example application was made using one of these techniques.

KEY WORDS: artificial intelligence, location, machine learning

TEŞEKKÜR

Tez sürecimde; benden yardımını esirgemeyen danışmanım sayın Dr. Öğretim Üyesi Fred Barış ERNST' e, işyerinden çalışma arkadaşım Harita Mühendisi İbrahim Halil Demir'e, Şanlıurfa Çevre ve Şehircilik İl Müdürlüğü çalışanları, Çevre Mühendisi Umut Öner ve Çevre Mühendisi Mustafa Serhat Turgut'a teşekkür ederim.



ŞEKİLLER DİZİNİ

Sayfa No

Şekil 2.1. Ev fiyatlarının tahmini üzerine yapılmış basit bir regresyon grafiği	6
Şekil 2.2. Kanser türünün kötü huylu ya da iyi huylu olduğunu gösteren sınıflandırma grafiği	7
Şekil 2.3. Denetimli ve denetimsiz öğrenmeyi gösteren grafikler.....	9
Şekil 2.4. Makineye verilen veri miktarına bağlı olarak performansı gösteren grafik	11
Şekil 3.1. Çevre Şehircilik ve İklim Değişikliği Bakanlığı'nın Hava Kalitesi İzleme ağına dair web sitesi görüntüsü.....	21
Şekil 3.2. Çalışma alanı ve çalışma alanında bulunan hava kalitesi izleme istasyonlarını gösteren harita	22
Şekil 3.3. Çalışma alanının bulutlu alanlarının çıkartılmış son halini gösteren harita	27
Şekil 3.4. Bayes Teoremini anlatan şekil	28
Şekil 4.1. 30 metre çözünürlüğe sahip yer yüzeyi sıcaklığı haritası.....	33
Şekil 4.2. 30 metre çözünürlüğe sahip fabrikalara uzaklık haritası	34
Şekil 4.3. 30 metre çözünürlüğe sahip BU indeksi haritası	35
Şekil 4.4. 30 metre çözünürlüğe sahip SAVI indeksi haritası	36
Şekil 4.5. 30 metre çözünürlüğe sahip yollara yakınlık haritası	37
Şekil 4.6. 30 metre çözünürlüğe sahip band2 haritası	38
Şekil 4.7. 30 metre çözünürlüğe sahip band3 haritası	39
Şekil 4.8. 30 metre çözünürlüğe sahip band4 haritası	40
Şekil 4.9. 30 metre çözünürlüğe sahip band5 haritası	41
Şekil 4.10. 30 metre çözünürlüğe sahip band6 haritası	42
Şekil 4.11. Öğrenme verilerinin elde edilmesine esas harita	43
Şekil 4.12. Yanlış öğrenmeye sebep olabilecek verileri analiz etmeye esas harita-1	44
Şekil 4.13. Yanlış öğrenmeye sebep olabilecek verileri analiz etmeye esas harita-2	45
Şekil 4.14. Oluşturulan veri setinde yer yüzeyi sıcaklığının sınıflara göre dağılımı	46
Şekil 4.15. Oluşturulan veri setinde yollara uzaklık değişkeninin sınıflara göre dağılımı	47
Şekil 4.16. Oluşturulan veri setinde fabrikaya yakınlık değişkeninin sınıflara göre dağılımı	48
Şekil 4.17. Oluşturulan veri setinde band2 değişkeninin sınıflara göre dağılımı	49
Şekil 4.18. Oluşturulan veri setinde band3 değişkeninin sınıflara göre dağılımı	50
Şekil 4.19. Oluşturulan veri setinde band4 değişkeninin sınıflara göre dağılımı	51
Şekil 4.20. Oluşturulan veri setinde band5 değişkeninin sınıflara göre dağılımı	52
Şekil 4.21. Oluşturulan veri setinde band6 değişkeninin sınıflara göre dağılımı	53
Şekil 4.22. Bağımsız değişkenlerin bağımlı değişkene olan etki derecesini gösteren grafik	54
Şekil 4.23. Bağımsız değişkenlerin bağımlı değişkene olan etki derecesini gösteren grafik	56
Şekil 4.24. Çalışma alanı PM10 değeri sağlıklı sınıfı tahmin haritası	57
Şekil 4.25. Çalışma alanı PM10 değeri hassas sınıfı tahmin haritası	58
Şekil 4.26. Çalışma alanı PM10 değeri orta sınıfı tahmin haritası	59
Şekil 4.27. PM10 değeri odaklı tahmin haritası	60
Şekil 4.28. PM10 değeri odaklı tahmin haritası (uydu görüntüsü ile karşılaştırılmış şekilde)	61
Şekil 4.29. PM10 değeri odaklı tahmin detay haritası-1 (uydu görüntüsü ile karşılaştırılmış şekilde)	63
Şekil 4.30. PM10 değeri odaklı tahmin detay haritası-2 (uydu görüntüsü ile karşılaştırılmış şekilde)	64
Şekil 4.31. PM10 değeri odaklı tahmin detay haritası-3 (uydu görüntüsü ile karşılaştırılmış şekilde)	65

ÇİZELGELAR DİZİNİ

Sayfa No

Çizelge 3.1. EPA Hava Kalitesi İndeksi.....	18
Çizelge 3.2. Ölçülen PM10 değeri ve sınıflarını gösteren çizelge	23
Çizelge 3.3. Bayes Teoremine örnek teşkil eden veri seti.....	29
Çizelge 4.1. Oluşturulan modelin Hata Matrisi.....	55
Çizelge 4.2. Hata Matrisine göre bulunan değerlendirme ölçütleri.....	56



1. GİRİŞ

Günümüzde, 4. Endüstri devriminin de hız kazanmasıyla birlikte, yapay zekâ tanımını içerisinde barındıran algoritma, makine ve sistemlerine ihtiyaç gün geçtikçe artmaktadır. Yapay zekânın birçok uygulama alanı bulunmaktadır. Bu uygulama alanlarının her birini yapay zekâ tekniği olarak adlandırabiliriz. Basit bir arama algoritmasından, karmaşık optimizasyon problemlerine, evlerimizi temizleyebilen robot süpürge makinelerinden, otonom sürücüsüz araçlara kadar hepsinin temelinde yapay zekâ terminolojisi yatmakta olup, her bir uygulama alanı bir veya birden çok yapay zekâ tekniğini içerisinde barındırmaktadır. Yapay zekâ terminolojisini çok büyük bir ağaç olarak düşünürsek, bu ağacın her bir dalının da ayrı bir yapay zekâ tekniğine tekabül edecek kadar büyük bir uygulama alanı olarak görebiliriz. Dallanıp budaklanmış bu ağacı 7 büyük ana dala ayırabilmek mümkündür. Muhakeme ve problem çözme, makine öğrenmesi, bilgi temsili dili, robotik, planlama, doğal dil işleme, bilgisayar ile görü (Mirjalili, S., Tim Buchalkas Learning Acedemy, 2021). Bu çalışmada, bu yedi ana disiplinin neler olduğundan ve ne gibi uygulamalarda kullanıldığından kısaca bahsedilip, bunlarla ilgili mekânsal uygulamalardan bahsedilecektir. Sonrasında bahsedilen uygulamalara örnek teşkil edecek, mekânsal bir problemin çözümüne yönelik, makine öğrenmesi algoritmalarından biri olan, Naif Bayes sınıflandırma algoritması kullanılarak, Denizli ili, merkez ilçelerini kapsayan alanda, hava kirliliği tahmin haritası oluşturulacaktır. Literatürde, makine öğrenmesi algoritması kullanılarak yapılan hava kirliliği tahmini yapan birçok çalışma bulunmaktadır. Bu çalışmayı diğer çalışmalardan ayıran fark ise, yapay zekâ tekniklerini bütüncül şekilde ele alıp, bunları mekânsal açıdan nasıl değerlendirilebileceği ortaya koymasının yanında, hava kirliliği gibi bir problemin ya da başka bir olgu ya da problemin, makine öğrenmesi algoritması kullanarak, hem makinenin öğrenmesi sürecinde, görsel olarak oluşturulan haritalar ve grafikler eşliğinde nasıl kararlar alınması gerektiği, hem de makine öğrenmesi algoritmasının sonuç ürününün yine görsel olarak harita ve diğer altlıklar eşliğinde değerlendirilip nasıl düzeltmeler yapılması gerektiğini aşama aşama anlatmasıdır.

Bu çalışmadaki asıl amaç, eldeki yapay zekâ tekniklerinin barındırdığı teorinin, günümüzde büyük ehemmiyete sahip olan mekânsal problemlere nasıl uygulanıp, etkili tahmin haritalarının nasıl oluşturulacağına dair bir rehber, bu alanda ilgili kişilere yeni bakış açıları oluşturup; çeşitli problemlerin çözümüne ilham kaynağı olmaktır.



2. KURAMSAL TEMELLER

2.1. Yapay Zekâ

Yapay zekânın keşfinden, günümüze kadar, yapay zekâ ile ilgili birçok tanımlama yapılmıştır. Bunlardan birkaçı şu şekildedir. Genellikle zekâ gerektirdiği düşünülen davranış sergileyen bilgisayarlı bir sistem (NSTC, 2016). İnsan tarafından yapıldığında, zekâ gerektirecek şeyleri yapan makineleri yapma bilimi (Raphael, B. 1976). Özellikle zeki bilgisayar programları olmak üzere, zeki makineler yapma bilimi ve mühendisliği(Standford University, 2007).

Yapay zekâ, içerisinde bulundurduğu çok sayıda teknik bakımından, çok büyük bir kavramdır. Bu kavram, 7 ana kısımda incelenebilir: Muhakeme ve Problem Çözme (Arama Yöntemleri), Makine Öğrenmesi, Bilgi Temsili Dili, Robotik, Planlama, Doğal Dil İşleme, Bilgisayar İle Görü (Mirjalili, S., Tim Buchalkas Learning Acedemy, 2021).

2.1.1. Muhakeme ve problem çözme (Arama yöntemleri)

Arama yöntemleri olarak da adlandırabileceğimiz bu disiplin, yapay zekânın temelini dayandığı disiplin olması açısından çok önemlidir. Herhangi bir probleme dair olabilecek optimum çözümü, geliştirilen arama algoritması ile bulabilen arama yöntemleri, yapay zekanın diğer ana dallarında sıkça kullanılmaktadır (Mirjalili, S., Tim Buchalkas Learning Acedemy, 2021). Yapay zekâ teknikleri ile iş yapacak birinin, ilk önce arama algoritmalarına dair bilgi sahibi olması gerekmektedir. Arama yöntemleri, belli bir probleme dair olabilecek tüm muhtemel çözümler arasından en optimumu seçme işini yaparlar. Arama yöntemleri iki kısımda incelenebilir: Kör arama yöntemleri, bilinçli arama yöntemleri.

2.1.1.1. Kör arama yöntemleri

Bu yöntem, bir probleme dair muhtemel çözümlerin bulunduğu arama uzayındaki, tüm çözümlere tek tek bakar ve en iyi çözümü seçer. Bu yöntem ile oluşturulan algoritmalar, probleme dair en iyi çözümü her zaman bulabilirken, çok büyük arama uzaylarının bulunduğu günümüz problemlerinde, zaman ve maliyet açısından verimsiz bir yöntem olmaktadır. Kör arama algoritmalarına örnek olarak; kaba kuvvet algoritması, ağaçlarda sıg öncelikli arama algoritması, derin öncelikli arama algoritması verilebilir.

2.1.1.2. Bilinçli arama yöntemleri

Bu tip arama yöntemleri, arama uzayının tüm muhtemel çözümlere bakılmasına zaman ve maliyetin elvermediği durumlarda, çözüm yolu sunan algoritmalarıdır. Kör arama yöntemlerindeki gibi, en iyi çözümü bulmayı garanti etmese de, kullandığı algoritma ile, kısa zamanda makul olabilecek çözümler sunmaktadır. Örnek olarak; insanların gen yapısını örnek alarak optimizasyon yapan genetik algoritma ve doğadaki canlıları taklit edebilen sezgisel algoritmalarından olan, karınca kolonisi algoritması verilebilir.

Evreni incelediğimizde, her zaman en kısa yolu seçtiğini görebiliriz. Bu tip arama yöntemleri de, evrendeki canlıların, oluşumların yaptıklarını taklit ederek oluşturulan algoritmalarından olduğunu görebiliriz.

2.1.1.3. Mekânsal bağlamda arama yöntemlerine bakış

Günümüz problemlerine bakıldığı zaman, mekân ve konumdan bağımsız problemler yok denecek kadar azdır. Yapay zekâ tekniklerinin temel disiplini olan arama yöntemleri de mekânsal birçok probleme çözüm bulmaktadır. Örnek vermek gerekirse, karıncaların yuvalarından yemek kaynağına olan en kısa yolu bulurken uygulamış oldukları yöntemi, matematiksel olarak ifade eden karınca kolonisi

algoritması kullanılarak, ülkemizin herhangi bir ilinden başlayarak, tüm illeri dolaşmak suretiyle tekrar yola çıktığı şehre dönen en kısa yolu bulma ya da birbiri ile kesişen onlarca, hatta yüzlerce güzergahın bulunduğu bir fabrikada ürün yapımı ve paketlenmesini optimum zamanda yapma amacını gerçekleştirebiliriz.

2.1.2. Makine öğrenmesi

Günümüzde en popüler olan yapay zekâ tekniği olduğunu söyleyebiliriz. Mekânsal yapay zekâ uygulamalarının en geniş olduğu ve en popüler teknik olduğu için, bu bölümde daha fazla şeyden bahsedilecektir.

Makine öğrenmesi, matematiksel modeller aracılığıyla, bilgisayara verilen verilerden öğrenme süreci olarak ifade edilebilir. Makine öğrenmesinde, veri içerisindeki belli düzenleri anlayabilmek adına, matematiksel modelleri ifade eden algoritmalar kullanılmaktadır. Makine öğrenmesi ile yapılan tahmin ya da sınıflandırma işleminde model oluşturabilmek için veri içerisindeki bahsedilen belli düzen ya da kalıplar kullanılır. Nasıl ki insanoğlu daha fazla araştırma yaptığında ya da bir şeyi tecrübe ettiğinde geliyorsa, bilgisayara verilen veri miktarı ve bilgisayarın tecrübe etmesi olarak ifade edebileceğimiz iterasyon miktarı arttıkça, makine öğrenmesi de daha iyi sonuçlar vermektedir.

Makine öğrenmesi, makineye verilen verilerle öğrendiği için, problemin sahip olduğu parametrelerden dolayı, belirlenen bir çözümün kodlanmasının verimli olmadığı durumlarda çok etkili bir seçenektir. Makine öğrenmesi; denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeye dayalı öğrenme olmak üzere üç kısımda incelenebilir.

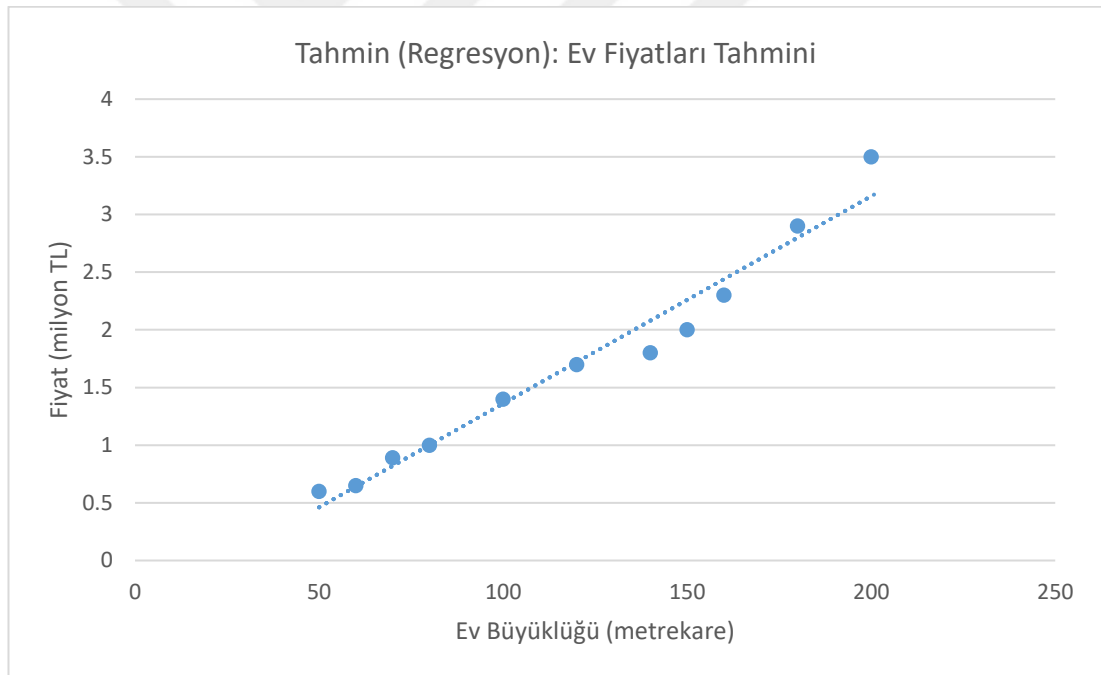
2.1.2.1. Denetimli öğrenme

Bilgisayara verilen verilerin etiketlerle verildiği denetimli öğrenme işinde, veriler bir öğretmen gibi makineyi eğiterek tahmin ya da sınıflandırma yapar. Bu tip öğrenmenin en önemli karakteristik özelliği, öğrenme algoritmasına, öğrenmesi için

örnekler verilmesidir. Verilen örneklerden yola çıkarak girdi ile çıktı arasındaki ilişkiyi öğrenip, sonrasında verilen herhangi bir girdiye ilişkin çıktı ürünü verir (Ng, 2022). Denetimli öğrenmeyi de, 2 alt kısımda inceleyebiliriz: Tahmin(regresyon) ve sınıflandırma algoritmaları.

2.1.2.1.1. Tahmin (Regresyon)

Modellenmeye çalışılan bağımlı değişken ile, bağımlı değişkene etki eden faktörler olan bağımsız değişken arasındaki ilişkiyi tahmin etmede kullanılır. Daha detaylı şekilde, bağımlı değişkenin bağımsız değişkenlerdeki değişimlere göre nasıl değiştiğine odaklanır. Anı zamanda, değişkenler arasındaki gelecekteki ilişkinin modellenmesinde de rol oynar (Bulutistan, 2021).



Şekil 2.1. Ev fiyatlarının tahmini üzerine yapılmış basit bir regresyon grafiği

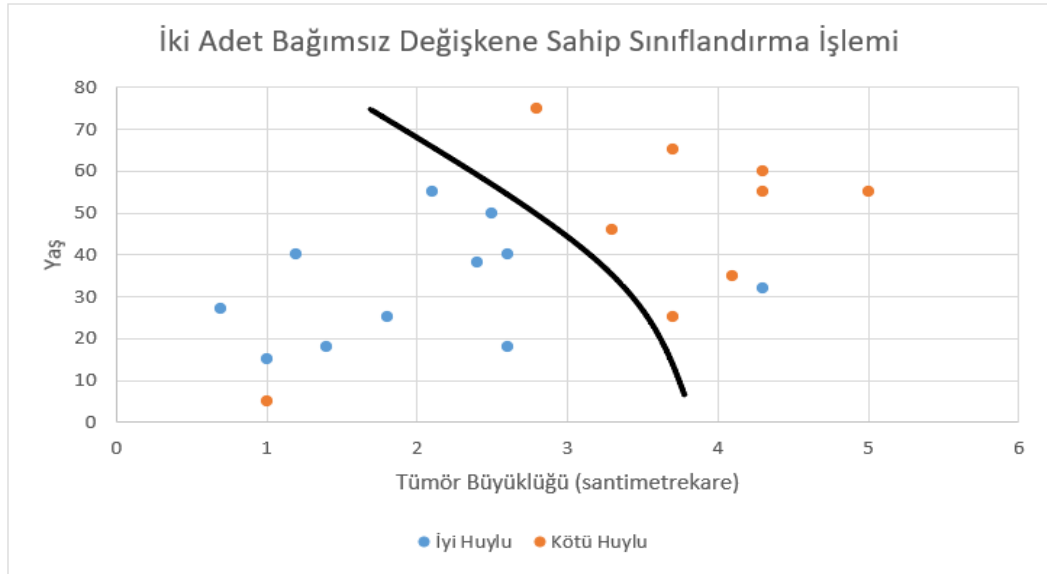
Şekil 2.1’de, ev fiyatlarının tahmini üzerine bir çalışma yapılmış olup, çalışmada ev fiyatına etki eden tek faktör olarak evin büyüklüğü alınmıştır. Bu durumda, bağımlı değişken, modellenmeye çalışılan “ev fiyatı”, bağımsız değişken ise “ev büyüklüğü”dür. Şekil 2.1’deki her bir mavi nokta işareti, bağımsız değişken ve bağımlı değişken

değeri bilinen örnekleri temsil etmektedir. Kesikli mavi doğru ise, örnekleri temsil eden optimum doğrudur. Aslında baktığımız zaman, makineye verilen örneklerden sonra, makine verilerden öğrenerek lineer bir fonksiyon şeklinde ifade edilebilen bir model ortaya çıkarmıştır.

Denetimli öğrenme yöntemini kullanan birçok tahmin algoritması mevcuttur. Bunlardan en çok kullanılan, önemli birkaç tanesini saymak gerekirse; doğrusal regresyon, polinom regresyon, destek vektör makineleri, karar ağacı ve rassal orman gibi algoritmalar sayılabilir.

2.1.2.1.2. Sınıflandırma

Sınıflandırma regresyondan farklı olarak, tahmin edilen sonuç ürünü belirli sayıda kategorik değerden oluşur.



Şekil 2.2. Kanser türünün kötü huylu ya da iyi huylu olduğunu gösteren sınıflandırma grafiği

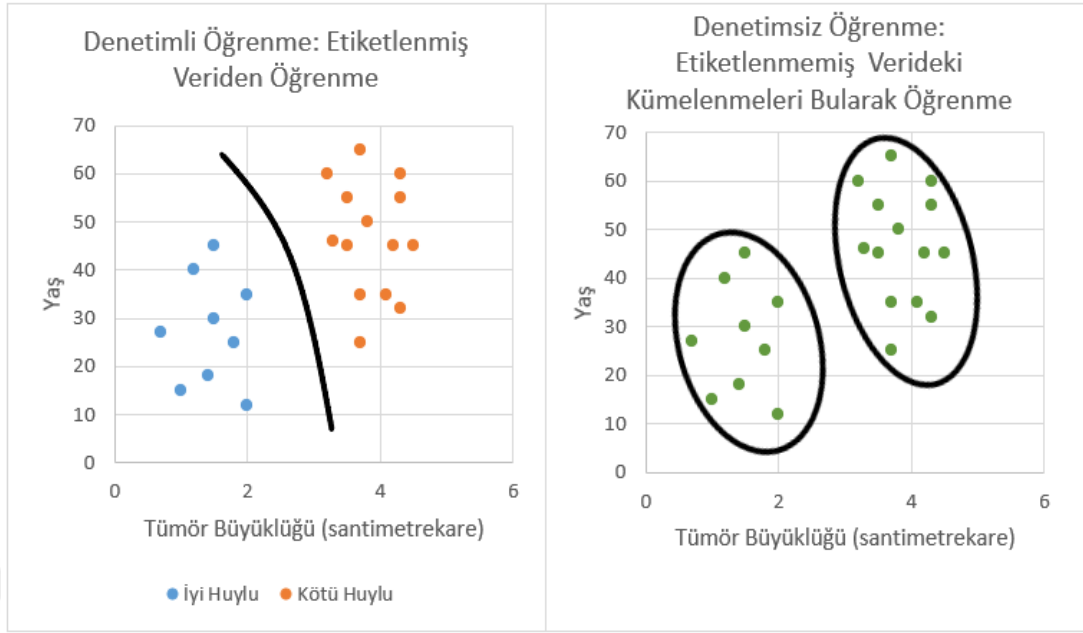
Şekil 2.2’de, kanser türünün iyi huylu ya da kötü huylu olması üzerine bir çalışma yapılmış olup, çalışmada kanser türüne etki eden faktörler olarak, yaş ve tümör büyüklüğü alınmıştır. Bu durumda, bağımlı değişken, tahmin edilmeye çalışılan,

kategorik veri olan “kötü huylu”, ya da “iyi huylu” olma durumu, bağımsız değişkenler ise “yaş” ve “tümör büyüklüğü” dür. Şekil 2’deki her bir mavi nokta işareti, iyi huylu kanser örneklerini, her bir turuncu nokta işareti ise, kötü huylu kanser örneklerini temsil etmektedir. Siyah eğri ise, verilen örnekleri kötü huylu ve iyi huylu olarak iki sınıfa ayıran optimum çizgiyi ifade etmektedir. Aslında baktığımız zaman, makineye verilen örneklerden sonra, makine, verilerden öğrenerek, iki sınıfı birbirinden ayıran bir fonksiyon şeklinde ifade edilebilen bir model ortaya çıkarmıştır.

Denetimli öğrenme yöntemini kullanan birçok sınıflandırma algoritması mevcuttur. Bunlardan en çok kullanılan, önemli birkaç tanesini saymak gerekirse; lojistik regresyon, destek vektör makinesi, naif bayes, karar ağacı ve rassal orman gibi algoritmalar sayılabilir.

2.1.2.2. Denetimsiz öğrenme

Bu teknikte, denetimli öğrenmedeki gibi, bilgisayara verilen veriler etiketlenmeden üzerinde herhangi bir etiketleme yapılmadan, veri içerisindeki belli düzen ya da kalıplar belirlenir. Yani, hazırlanan örnekleri makineye, etiketlenmeden vererek, makinenin verideki örüntüyü bularak buna göre veriyi sınıflandırma işi diyebiliriz.



Şekil 2.3. Denetimli ve denetimsiz öğrenmeyi gösteren grafikler

Şekil 2.3'te de görüldüğü üzere, sol taraftaki grafik, yaş ve tümör büyüklüğü bağımsız değişkenlerine göre, denetimli öğrenme yöntemi kullanarak iyi ya da kötü huylu kanser olma durumunu modelleyen grafiği gösterirken, sağ taraftaki grafik, yaş ve tümör bağımsız değişkenlerine göre, etiketleme yapılmadan verilen örneklere, denetimsiz öğrenme yöntemini kullanarak, verideki örüntüyü, kalıbı ya da kümelenmeyi modelleyen grafiği göstermektedir. Denetimsiz öğrenme algoritmalarını, kümeleme algoritmaları olarak da adlandırabiliriz.

Denetimsiz öğrenmeye ilişkin, literatürde birçok algoritma mevcuttur. Bu tip algoritmalara örnek olarak; K-ortalama, hiyerarşik kümeleme algoritmalarını söyleyebiliriz (Candan, 2021).

2.1.2.3. Pekiştirmeye dayalı öğrenme

Bilgisayar programları olan araçlar, bu tip öğrenmede, insan operatörün yerini alarak, bir geri bildirim döngüsünü temel alması suretiyle sonucun belirlenmesine yardımcı olur (Escape Velocity Labs, 2022). Bu tip öğrenmede bilgisayar programı olan, satranç oyunu, fabrikalarda ürünleri oluşturan robot kollar ya da video oyun

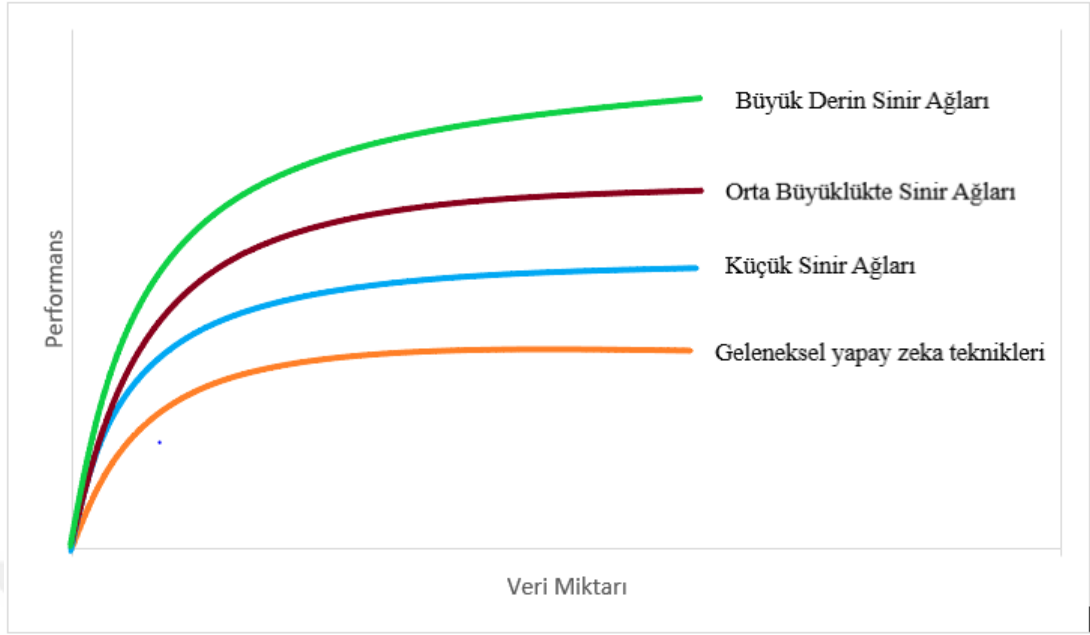
motoru gibi araçlara, yapmaları gereken görevlere göre algoritma entegre edilir ve bu algoritma kötü bir algoritma olsa dahi, zamanla bu araçlar tecrübeye dayalı öğrenme ile yapmaları gereken görevi istenen şekilde yapabilecek seviyeye evrilir.

Pekiştirmeye dayalı öğrenmeye ilişkin, literatürde birçok algoritma mevcuttur. Bu tip algoritmalara örnek olarak, üst güven sınırı ve Thompson örnekleme algoritmalarını söyleyebiliriz.

2.1.2.4. Derin öğrenme

Makine öğrenmesi disiplininin bir alt dalı olarak düşünebileceğimiz derin öğrenme, günümüzde en çok kullanılan ve verimliliği en yüksek olan, denetimli ve denetimsiz öğrenme yöntemlerini de içerisinde barındırdığından dolayı, bu makine öğrenmesi tekniği ayrı bir başlık altında incelenmiştir. Derin öğrenme disiplininin temellerini “sinir ağları” oluşturmaktadır.

Sinir ağları, insan beyninin çalışmasını taklit etmeye çalışan bir algoritma olarak ortaya çıkmıştır. Sinir ağlarındaki çalışmalar 1950'lerde başlamıştır ve bir süreliğine gözden düşmüştür. Daha sonra 1980'lerde ve 1990'ların başında yeniden popülerlik kazanmışlardır ve o zamanlar postaları yönlendirmek için, posta kodlarını okumak ve el yazısı çeklerdeki dolar rakamlarını okumak için bile, kullanılan el yazısı, rakam tanıma gibi bazı uygulamalarda muazzam bir çıkış göstermişlerdir. Ama sonra 1990'ların sonunda tekrar gözden düşüp, bilgisayarlardaki işleme gücünün artmasıyla birlikte, 2005 yılı civarında bir canlanma yaşamış ve derin öğrenme ile markalaşmıştır (Ng, 2022). O günden günümüze kadar ise, popülerliği ve uygulama alanları git gide artmaktadır.



Şekil 2.4. Makineye verilen veri miktarına bağlı olarak performansı gösteren grafik

Günümüzde, baş edilmesi gereken devasa boyuttaki veri miktarı (büyük veri) göz önünde bulundurulduğunda, Şekil 2.4'te de görüldüğü üzere, en etkili yöntemin sinir ağları olduğu görülmektedir.

2.1.2.5. Mekânsal bağlamda makine öğrenmesinin değerlendirilmesi

Bahsedildiği üzere, makine öğrenmesi, uygulama alanı çok büyük bir alandır. Aynı zamanda, mekânsal problemlerin çözümü için de çokça kullanılan bir yöntemdir. Orman yangınlarından, suç analizine; trafik problemlerinden, heyelan, erozyon, sel risk haritalarının oluşturulmasına, düşündükçe örneklerin çoğaltılabileceği günümüzde, karşılaştığımız birçok problem, mekân ve konumu içerisinde barındırmaktadır. Makine öğrenmesi o kadar geniş bir alana hitap etmektedir ki, bu problemlerin hepsinde makul çözümler sunabilmektedir.

Orman yangınlarından bahsedecek olursak, orman yangınlarının sıkça gerçekleştiği bir bölgeyi baz aldığımızı düşündüğümüzde; orman yangınlarına etki eden faktörler araştırılıp, söz konusu bölgede irili ufaklı gerçekleşen orman

yangınlarının konum bilgisi ve diğer araştırılan faktörlere dair, mekânsal bilgi ile niceliksel verilerin toplanıp makineye verilmesi durumunda, makine, bölgedeki her bir metrekaredeki yangın olma olasılığını bize verecektir.

Bu şekilde anlatıldığında, gerekli verileri verdikten sonra, bilgisayar bizim yerimize her şeyi hallediyor gibi gözükebilir; fakat hem optimizasyon öncesi verilerin hazırlanması kısmında, o bölgenin, dünyadaki herhangi bir konumu belirlememize yarayan belli bir projeksiyon ve datum ile ilgili yangın verileri alınıp, yangına etki eden faktörlere dair o bölgeyi kapsayan sayısal haritaların yine aynı projeksiyon ve datum kullanılarak, eğer ki elde edilen verinin projeksiyon ve datum bilgisi farklı ise bu verinin projede kullanılan projeksiyon ve datum bilgisine sahip olacak şekilde dönüştürülerek, bu sayısal haritaların karşılaştırılması ve bu haritaların belli işlemlerden geçirilerek, o bölgedeki herhangi bir konumun bahse konu faktöre dair niceliksel verisinin elde edilebileceği formata dönüştürülmesi gerekmektedir. Optimizasyon sırasında ve sonrasında ise, kullanılacak makine öğrenmesi algoritmasının iç dinamiklerinin bilinmesi, sahip olduğumuz problemin kullanılacak makine öğrenmesi algoritmasına uygun olacağına sonucuna varılması, optimizasyon sonrasında ise, elde edilen verinin doğruluğunun kontrol edilmesi ve istenilen seviyede ise ona göre devam edilmesi gerekmektedir.

Bahsedilen örnek, dünyanın herhangi bir yerinde herhangi bir probleme, bilgi ve veri paylaşımının her geçen gün arttığı günümüz şartlarında, yeterli veri olduğu sürece uygulanabilir.

2.1.3. Bilgi temsili dili

Temel olarak, akıllı bir programın; inançlarının, niyetlerinin ve yargılarının otomatik akıl yürütme için uygun şekilde nasıl ifade edilebileceğinin bir çalışmasıdır. Bilgi temsili dili, bir bilgisayarın anlaması ve daha sonra bu bilgiyi insanlarla doğal dilde iletişim kurmak gibi karmaşık gerçek yaşam problemlerini çözmek için kullanması için gerçek dünyadan gelen bilgileri temsil eder. Yapay zekada bilgi

temsili, yalnızca veriyi bir veri tabanında depolamakla ilgili değildir, bir makinenin bu bilgiden öğrenmesini ve bir insan gibi akıllıca davranmasını sağlar (Sayantini, 2022).

Bilgi temsili dili uygulamalarına örnek olarak, sohbet robotu ve sanal asistan uygulamaları verilebilir (Mirjalili, S., Tim Buchalkas Learning Acedemy, 2021).

2.1.4. Robotik

Robotik temel olarak bir robotun sensörlerden verileri almasına, verileri işlemesine ve kararlar almasına ve nihayetinde, aktüatörler ve motorlarla istenilen görevi yerine getirmeyi sağlayan alandır. Bu alan, yapay zekanın en geniş çaplı ve popüler uygulama alanıdır (Mirjalili, S., Tim Buchalkas Learning Acedemy, 2021).

Örnek olarak, fabrikalarda mekanik parçaları birleştirme ya da ürün paketleme yapabilen robot kollar, evlerde temizlik yapabilen robot süpürge makineleri ve otonom sürücüsüz araçlar verilebilir.

Aslına bakıldığında zaman, robotik ile diğer yapay zekâ uygulamaları birbiriyle iç içedir. Robotu harekete geçiren yine içindeki yapay zekâ algoritmaları olduğunu düşünüldüğünde, robotik alanını yapay zekâ algoritmalarının vücut bulmuş hali olarak düşünebiliriz. Aynı zamanda, robotik bilimini bir tür pekiştirmeye dayalı öğrenme yöntemi olarak ifade edebiliriz.

2.1.4.1. Mekânsal bağlamda robotik biliminin değerlendirilmesi

Robotik bilimini, mekân ve konumdan bağımsız olarak düşünemeyiz. Belli bir görevi yapmak için tasarlanmış robotun, görevini doğru şekilde yerine getirebilmesi için, kendi konumunu ve etrafındaki nesnelere konumunu algılayıp, ona göre hareket etmesi gerekmektedir. Herhangi bir robot, etrafındaki nesnelere makine öğrenmesi algoritma türü olan konvansiyonel sinir ağları yardımıyla tanıyarak, ona göre yine robota entegre edilmiş makine öğrenmesi algoritmaları ile görevini yerine getirebilir.

Otonom sürücüsüz araçların çalışma prensibinden yola çıkılırsa; sürücüsüz aracın, makinesi, yapay zekâsı, motoru, bunların hepsi çok önemli, ama en az onlar kadar önemli olan bir şey daha vardır. Sürücüsüz araç, bir yerden bir yere kendi kendine gitmek durumundadır, yolunu, güzergâhını belirlemek için bir haritaya ihtiyacı olur, aynı zamanda, hareket ettiği her an etrafında ne oluyor, ne bitiyor, hepsinin haritasını üretmesi gerekmektedir(Caner Güney, *YouTube* (11 Temmuz 2021), 00:30:43-00:31:23).

2.1.5. Planlama

Belirli bir amaca ulaşmak için, robot veya bilgisayar programı tarafından gerçekleştirilen karar verme sürecini ifade eder. Örneğin, bir roket planladığınızda, bir robot için bir yol planladığınızda, işte tam da bu noktada bir algoritma geliştirmeniz gerekiyor ve insanlar bu alanda bu tür problemler üzerinde çalışıyorlar (Mirjalili, S., Tim Buchalkas Learning Acedemy, 2021).

Yine bakıldığı zaman, bu disiplin de robotik, makine öğrenmesi ve diğer başka yapay zekâ tekniklerinin birleşimi olarak ifade edebiliriz.

2.1.5.1.Mekânsal bağlamda planlama disiplininin değerlendirilmesi

Planlama disiplini, yukarıda da bahsedildiği gibi, robotik, makine öğrenmesi ve diğer başka yapay zekâ tekniklerinin birleşimi olarak düşünebileceğimiz için, diğer bölümlerde yapılan mekânsal bağlamda değerlendirmeler bu bölümdeki değerlendirmede de geçerlidir.

2.1.6. Doğal dil işleme

Bilgisayarların konuşulan dili anlaması, işlemesi, yorum yapması ve aynı zamanda cümle üretebilmesi işidir (Noyan, 2019). Doğal dil işleme, her ne kadar ayrı bir yapay zekâ tekniği altında işlenmiş olsa da, aslına bakılırsa, makineye verilen örneklerden öğrenerek işlem yapması açısından, bir makine öğrenmesi işidir.

2.1.7. Bilgisayar ile görü

Bilgisayarla görü, resim video ya da görsel herhangi bir unsuru, bu unsurlar içerisindeki pikselleri kullanarak anlamlı bilgi üretilmesi ve buna göre işlem yapılması sürecidir.

Yine bakıldığı zaman, bilgisayar ile görü disiplini de yapay zekanın ayrı bir başlığı olarak incelense de, aslında bilgisayar ile görü işlemi makine öğrenmesinin bir alt dalı olan derin öğrenme tekniklerinden evrişimli sinir ağları disiplini kullanılarak yapılması açısından bir makine öğrenmesi işidir. Evrişimli sinir ağları, herhangi bir görüntüdeki her bir pikseli, makine öğrenmesi kısmında bahsettiğimiz bağımsız birer değişken olarak ele alır ona göre çözümleme yapar.

2.1.7.1. Mekânsal bağlamda bilgisayar ile görünün değerlendirilmesi

Bilgisayar ile görünün, mekânsal problemlere çözüm bulan yapay zekâ uygulamalarında çok büyük bir rolü bulunmaktadır. Mekânsal problemlerin çözümünde, bilgisayar ile görü sıkça kullanılmaktadır.

Günümüzde, sosyal, doğal afet, sağlık, şehir planlaması, iklim değişimi ve daha birçok alanda, çözülmesi gereken çok miktarda problem le karşılaşırız. Bunların hepsi mekân ile ilgilidir. Bu problemlerin çözümünde, dünyanın herhangi bir yerinde neler oluyor bitiyor, bunları anlayarak erken önlemler alabilmemiz ya da aynı problemin bir daha gerçekleşmemesi adına, mekânsal analizler yapabilmemiz için, farklı uydu programları tarafından yayınlanan uydu görüntülerini, mekânsal yapay zekâ uygulamalarında altlık olarak kullanmak gerekmektedir. Mekânsal analizler yapılabilmesi için, bazen, altlık olarak uydu görüntüsünden (raster veri) çok daha fazla şeye ihtiyaç vardır. İlgili raster verideki analize konu nesnelere, vektör veriye dönüştürülüp, bu nesnelere ait öznelik(detay) bilgilerinin atanması gerekmektedir. Bu işi geleneksel yöntemlerle de yapmak mümkündür; fakat çok büyük alanlarda yapılan, analize konu yüzlerce, belki binlerce nesneyi barındıran çalışmalarda geleneksel yöntemler kullanıldığında, bu uzun zaman ve pahalı maliyete yol

açmaktadır. Tam da bu kısımda; bilgisayar ile görü tekniği devreye girebilmektedir. Vektör veriye dönüştürülmek istenen nesneye dair yeterli miktarda örnek bilgisayara verilip, öğrenme algoritması ile öğretim süreci de gerçekleştirildiğinde, yüzlerce, binlerce nesneyi otomatik olarak vektör veriye çevirme işlemi yapılabilir. Gerekli doğruluk analizleri yapıldıktan sonra, bu vektör verilere dair ilgili öznelik verilerinin de atanması ile birlikte, çeşitli mekânsal analizler yapmak mümkündür. Bu analizler, basit mekânsal analizler de olabilir, çalışma yapılan bölgedeki çözülmek istenen bir probleme dair bağımsız değişkenlerin belirlenerek, çalışma bölgesindeki herhangi bir konumundaki bağımsız değişkene dair niceliksel değerlerin belirlenmesiyle birlikte makine öğrenmesi disiplini kullanılarak gerçekleştirilen mekânsal analizler de olabilir.



3. MATERYAL ve YÖNTEM

3.1. Problemin Tanımı:

Ülkemizde, Çevre Şehircilik ve İklim Değişikliği Bakanlığı tarafından, hava kirliliğinin doğru bir şekilde ölçülmesini sağlamak için, tüm illerimizde, hava kalitesi ölçüm istasyonları kurulmuştur. Fakat bu istasyonlar kurulduğu yerdeki hava kirliliği değerini vermektedir. Bu ise hava kirliliği istasyonu bulunmayan yer ve bölgelerdeki hava kirliliği değerlerinin bulunması problemini ortaya çıkarıyor.

Problemin çözümüne geçmeden önce, hava kirliliğine dair bilinmesi gereken önemli kavramlardan bahsedilecektir.

3.1.1. Hava kalitesi indeksi

Hava kirliliğinin insan sağlığı açısından önemli etkileri olmasından dolayı, hava kalitesi konusuna tüm dünya tarafından büyük önem verilmektedir. Hava kirliliği problemini çözmek ve stratejiler belirlemek adına, bilimsel topluluğun ve ilgili otoritelerin atmosferik kirletici konsantrasyonlarını izlemek ve bunları analiz etmek konusuna odaklanmışlardır (Kyrkilis vd., 2007). Hava kalitesinin durumu ifade edilirken, kolayca anlaşılır bir sınıflama sistemi kullanılmalıdır. Hava Kalitesi İndeksi (HKİ) ismindeki bu sınıflama sistemi ile birlikte, hava kalitesi durumunu; iyi, orta, kötü, tehlikeli vs. şeklinde sınıflandırma yapılmaktadır. Bizim ülkemizde de, dünyanın pek çok ülkesinde indeks hesaplanmada kullanılan yöntem ve kriterler, kullanılmaktadır.

Bir bölgedeki hava kalitesi durumunun karakterize edilmesi adına, ülkelerin kendi sınır değerleri ile, kirlilik sınıflandırılmasının yapıldığı indeks, Hava Kalitesi İndeksi (HKİ) (Air Quality Index/AQI) olarak adlandırılmaktadır. İndeks, belli kategorilerde farklı tanım ve renkler ile ifade edilmektedir ve ölçümü yapılan her kirletici unsur için ayrı ayrı düzenlenmektedir (Yavuz, 2010).

Ulusal Hava Kalitesi İndeksi, aynı zamanda EPA Hava Kalitesi İndeksi, ulusal mevzuatımız doğrultusunda ve sınır değerlerimize uyarlanarak oluşturulmuştur. Beş temel kirletici için hava kalitesi indeksleri hesaplanmaktadır. Bunlar; partikül maddeler (PM10), karbon monoksit (CO2), kükürt dioksit (SO), azot dioksit (NO2) ve ozon (O3) dur.

Çizelge 3.1. EPA Hava Kalitesi İndeksi

Hava Kalitesi İndeksi	Seviye	Renk	Anlamı
0-50	İyi	Yeşil	Hava kalitesi memnun edici ve hava kirliliği az riskli veya hiç risk teşkil etmiyor.
51-100	Orta	Sarı	Hava kalitesi uygun, fakat alışılmadık şekilde hava kirliliğine hassas olan çok az sayıda insan için, bazı kirleticiler açısından orta düzeyde sağlık endişesi oluşabilir
101-150	Hassas	Turuncu	Hassas gruplar için sağlık etkileri oluşabilir. Genel olarak kamunun etkilenmesi olası değildir.
151-200	Sağlıksız	Kırmızı	Herkes sağlık etkileri yaşamaya başlayabilir, hassas gruplar için ciddi sağlık etkileri söz konusu olabilir.
201-300	Kötü	Mor	Sağlık açısından acil durum oluşturabilir. Nüfusun tamamının etkilenme olasılığı yüksektir.
301-500	Tehlikeli	Kahverengi	Sağlık alarmı: Herkes daha ciddi sağlık etkileri ile karşılaşabilir.

3.1.2. Hava kirliliği kaynakları

Hava kirliliği kaynaklarına göre 2 kısımdan oluşur: Doğal Kaynaklar, Yapay Kaynaklar.

3.1.2.1. Doğal kaynaklar

Doğada gerçekleşen bazı doğal olayları ile, havayı kirleten maddeler ortaya çıkabilmektedir. Böyle ortaya çıkan kirletici unsurlar, atmosferde uzun süre kalmazlar. Bunlar,

- Yanardağ faaliyetleri
- Orman yangınları
- Çöl tozlarına maruziyet
- Hayvan türlerinin ve bitki örtüsünün bozulmasıdır.

3.1.2.2. Yapay kaynaklar

İnsanoğlunun yaptığı faaliyetler sonucunda, doğanın dengesi olumsuz etkilenmektedir. Hava kirliliğinde durumunda, suni kaynaklardan dolayı meydana gelen kirlilik daha önemlidir. Bunun sebebi, günümüzde insanları en çok alakadar eden, büyük yerleşim merkezleri ile sanayi alanlarındaki hava kirliliğidir. Bu kirlilikte, çoğunlukla insan faaliyetleri sonucu meydana gelir. Buna neden olan kaynaklar şu şekilde sıralanabilir.

- Isınma
- Ulaşım
- Sanayi

İnsana bağlı faaliyetler nedeniyle oluşan bu kirlilik, ilgili bölgenin endüstriyel gelişimi, nüfus durumu, şehirleşme durumu gibi durumlara bağlı olarak değişim gösterir. Bunun dışında; meteorolojik faktörler, konum ve topografik durum, plansız kentleşme ve yeşil alanların yetersiz miktarda olması ve kullanılan yakıtların kaliteleri, yapay kaynaklardan dolayı oluşan kirliliği etkileyen faktörlerdir.

3.1.3. Partikül maddeler

Havada askıda kalmış katı parçacıklar partikül madde olarak isimlendirilir. Partikül maddeler doğal veya antropojenik kaynaklar nedeniyle atmosfere sağlanabilmesinin yanında, atmosferdeki reaksiyonlar ile de oluşabilir. Deniz spreyi mineral ve çöl tozları gibi unsurlar doğal kaynaklı hava kirliliğini oluştururken, sanayi faaliyetleri, güç santralleri ve yanma prosesleri de başlıca antropojenik kaynaklardır.(Fenger 2009). Partiküllerin çapı aynı zamanda sağlığa etkisini de belirlemektedir. Çapı 10 milimetreden küçük botutta olan kaba partiküller, PM 10

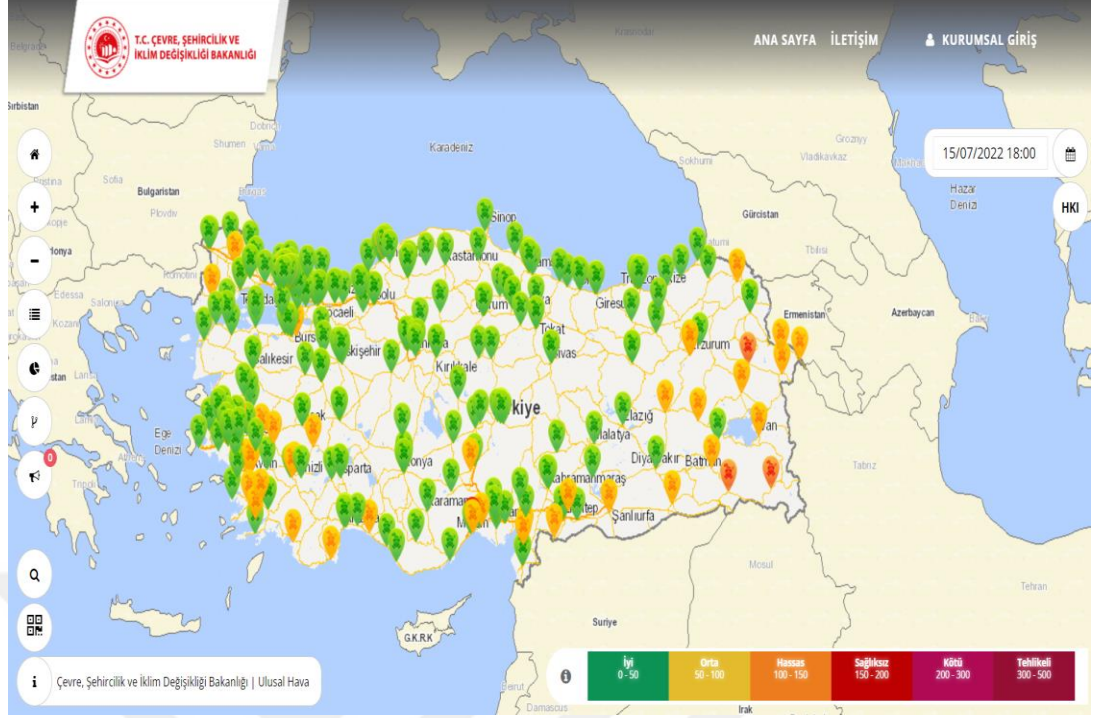
vücut boğaz bölgesinde ve aynı zamanda üst solunum yollarında birikim gösterirken, partikül çapı 2 buçuk milimetre boyutundan küçük olan ince partikülleri ise, akciğerlerde alveollere kadar ulaşmak suretiyle kana karışabilmektedir (Kampa and Castanas, 2008). Partikül maddelere maruziyet sonrası, hastane başvurularında artışlar gözlenmekte(Cruz et al., 2015), vücudun boğaz bölgesinde tahriş astım ile KOAH gibi solunum sistemi hastalıklarının yanında (Kravchenko et al., 2014), kardiyovasküler hastalıklar(Neas,2000) ortaya çıkmaktadır. Bahsedilenden daha uzun süreli maruziyetlerde ise, akciğer kanseri ya da ölüm meydana gelmektedir (Ciazzo et al.,2013). Partiküler maddeler aynı zamanda ekosistemin yapısını değiştirmekte, iklim değişikliğine neden olmakla beraber, görüş seviyesinde düşüklüğe de neden olmaktadır (Fuzzi et al.,2015).

3.2. Problemin Çözümü:

Problemin tanımı başlığında bahsedilen probleme çözüm olarak, hava kalitesi indeksi ölçümü yapan istasyonlardaki PM10 verileri göz önünde bulundurularak, makine öğrenmesinin sınıflandırma tekniklerinden biri olan Naif Bayes algoritması kullanılarak, istasyon bulunmayan alanlardaki PM10 hava kalitesi indeksini de tahmin eden, çalışma alanını kapsayan hava kirliliği haritası çıkarılması, bahsedilen problemin çözümü olarak adlandırılabilir.

3.3. Çalışma Alanının Belirlenmesi:

Çalışma alanının belirlenmesi için, Çevre Şehircilik ve İklim Değişikliği Bakanlığı tarafından, Sürekli İzleme Merkezi Ulusal Hava Kalitesi İzleme Ağını web üzerinden Coğrafi Bilgi Sistemi olarak sunulan internet sitesinde inceleme yapıldı.



Şekil 3.1. Çevre Şehircilik ve İklim Değişikliği Bakanlığı'nın Hava Kalitesi İzleme ağına dair web sitesi görüntüsü

Çalışma alanı belirlenirken, yapılan çalışma bir sınıflandırma işlemi olduğundan, çalışma alanının birden fazla hava kalitesi izleme istasyonuna sahip olması, bunun yanında da, bu istasyonlardan ölçülen PM10 değerlerinin hava kalitesi iyi, orta, hassas, sağlıksız, kötü ve tehlikeli sınıflarından en az iki tane sınıfı barındırmasına dikkat edildi.

Yapılan inceleme sonucunda, çalışma alanı olarak Denizli İlinin merkez ilçeleri, Merkezefendi ve Pamukkale ilçelerini kapsayan alan belirlendi.

belirlenmiştir. Çalışma alanında 4 adet istasyon bulunup, bu istasyonların PM10 değerlerinin sınıfları Çizelge 3.2’de gösterildiği şekilde Orta, Hassas ve Sağlıksız olarak belirlenmiştir.

Çizelge 3.2. Ölçülen PM10 değeri ve sınıflarını gösteren çizelge

No	İstasyon	PM10	Seviye
1	Merkezefendi	75,23	Orta
2	Bayramyeri	108,47	Hassas
3	Trafik	121,42	Hassas
4	Sümer	169,47	Sağlıksız

Uygulamada kullanılan bağımlı değişken, “Seviye” değişkeni olup, bu değişken 3 farklı değere sahiptir. Yani uygulamada yapılacak sınıflandırmada 3 adet sınıf vardır.

3.5. Bağımsız Değişkenler

3.5.1. SAVI (Soil Adjusted Vegetation Index):

Landsat Toprak Ayarlı Bitki Örtüsü İndeksi (SAVI), toprak parlaklığını en aza indirmesinin yanında, bitki örtüsünden gelen verileri vurgular (TDT Türkiye, 2021). Söz konusu indeks aşağıdaki eşitlik ile bulunur.

Landsat 8-9 uydu programlarında;

$$SAVI = ((\text{Bant 5} - \text{Bant 4}) / (\text{Bant 5} + \text{Bant 4} + 0.5)) * (1.5) \quad (3.1)$$

3.5.2. BU (Built-up Index):

Uzaktan algılama yöntemi ile elde edilen görüntüden hesaplanabilen BU indeksi, görüntü üzerindeki yerleşim alanlarını ortaya çıkarmada kullanılmaktadır. Aşağıdaki eşitlik ile bulunur.

$$BU = ((\text{Bant 6} - \text{Bant 5}) / (\text{Bant 6} + \text{Bant 5})) - ((\text{Bant 5} - \text{Bant 4}) / (\text{Bant 5} + \text{Bant 4})) \quad (3.2)$$

3.5.3. Yer yüzeyi sıcaklığı:

Landsat-8'in termal kızılötesi bantları, arazi yüzey sıcaklığını tahmin etmeye olanak sağlar. Hesaplama için Landsat görüntüsünün 10 bandı kullanılır. Aşağıdaki şekilde hesaplanabilir (Vasylenko, 2017).

$$LST(^{\circ}C)=Bt/[1+(w*Bt/p)*\ln(e)]-273.15 \quad (3.3)$$

Bt: Kelvin cinsinden parlaklık sıcaklığı

w: yayılan ışığın dalga boyu

p: $h*c/s'$ dir, burada h Planc sabitidir, c ışık hızıdır, s Boltzmann sabitidir.

e: elektromanyetik emisyonudur. Kentsel alanlarda aşağıdaki şekilde hesaplanır.

$$e=0.017*PV+0.963$$

$$PV=[(NDVI-NDVImin)/(NDVImax-NDVImin)] \quad (3.4)$$

3.5.4. Yollara yakınlık:

Yollarda yakılan yakıtlar düşünüldüğü zaman, yollara yakınlık değeri de bağımsız değişken olarak ele alınmıştır.

3.5.5. Elektromanyetik bantlar (Bant 2):

Uzaktan algılama yöntemiyle, görünür ve görünmeyen birçok bantta uydu görüntüleri elde edilmektedir. Bu bantların her biri görüntüye ait bölgedeki farklı nitelikleri vurgulayabilmektedir.

Bant 2'nin işlevinden bahsedecek olursak, bant 2, bitki örtüsünden toprağı ayırt etmede ve iğne yapraklı bitki örtüsünden işlevini yitirmiş yaprak döken kısımları vurgulama işe yarar.

2. banttaki uydu görüntüsünü kullanarak, zayıf bitki örtüsünü ve bitki örtüsü gibi görünen yeşil alanları anlayıp, hava kirliliğini büyük etkisi olan yeşil alanları daha verimli şekilde anlaşılabilir.

3.5.6. Elektromanyetik bantlar (Bant 3):

Canlı bitki örtüsünü vurgulamada işlevsel olduğundan, yapılan çalışmada uygun görüntüsünün 3. Banttaki görünüşü de bağımsız bir değişken olarak ele alınmıştır.

3.5.7. Elektromanyetik bantlar (Bant 4):

Bitki örtüsüne meyilli alanları vurguladığından çalışmaya dahil edilmiştir.

3.5.8. Elektromanyetik bantlar (Bant 5):

Uydu görüntüsündeki canlı organizma kütlelerini vurgulayabilen banttır. Canlı organizmalar hava kirliliği açısından büyük öneme sahiptirler.

3.5.9. Elektromanyetik bantlar (Bant 6):

Toprak ya da bitki örtüsündeki nem miktarını vurgulamaktadır. Yüksek nemli hava havadaki kirli unsurları, havada asılı şekilde tutabildiğinden nem önemli bir bağımsız değişken olarak ele alınmıştır.

3.5.10. Fabrikaya yakınlık:

Fabrika ve işletmelerin bacalarından yayılan duman da önemli bir hava kirliliği kaynağı olduğu için, fabrikalara yakınlık değeri de bağımsız değişken olarak ele alınmıştır.

3.6. Elde Edilmesi Gereken Veriler

3.6.1. Uygulama alanındaki çok bantlı uydu görüntüsü

Uygulama alanına denk gelen çok bantlı uydu görüntüleri, USGS Earth Explorer sitesinden bulundu. Veri seti olarak, veri üzerinde düzeltmelerin yapılmış hali olan, Landsat 8-9 OLI/TIRS Collection 2 Level 1 veri formatı seçildi. Bu uydu görüntülerinden, en az bulut oranı seçeneği seçilerek, 2021 yılının kasım ayına ait bir görüntü seçildi. Sonrasında, seçilen görüntü, Earth Resources Observation And Science (EROS) Center internet sitesinden indirildi.

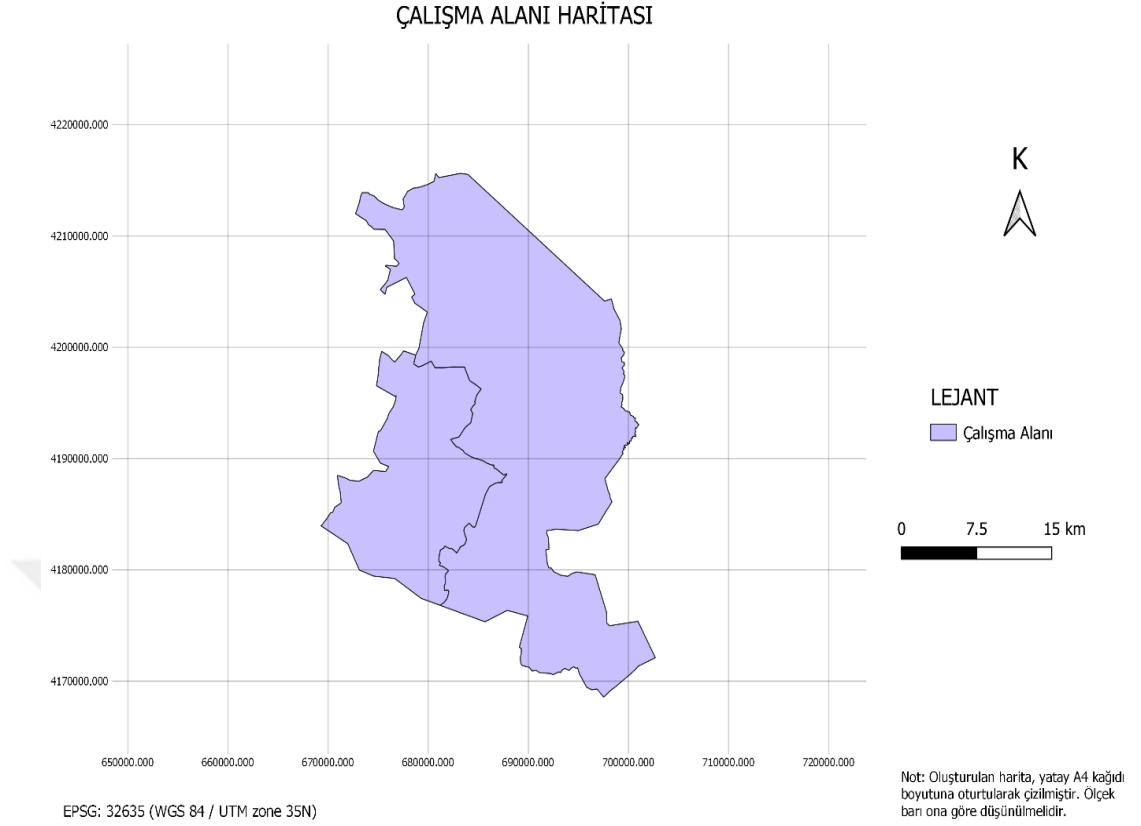
3.6.2. Uygulama alanındaki yol güzergahları

QGIS yazılımında, “quickOSM” eklentisi yardımıyla, çalışma alanındaki yollar vektör veri olarak indirilip, elde edildi. Elde edilen tüm yol güzergahlarından önem derecesi düşük yollar ayıklandı.

3.6.3. Çalışma alanı vektör verisi

Çalışma alanını kapsayan iki ilçenin sınırları poligon olarak Open Street Map verisinden elde edildi.

Yapılan analizde altlık olarak kullanılan uydu görüntüsü, doğal renklerde görüntülenerek, analizi olumsuz etkileyebilecek bulut oluşumunun olup olmadığı incelendi. Çalışma alanının bazı kesimlerinde bulut oluşumuna rastlandı. Bu kesimler çalışma alanından çıkartıldı.



Şekil 3.3. Çalışma alanının bulutlu alanlarının çıkartılmış son halini gösteren harita

3.6.4. Uygulama alanında bulunan hava kalitesi izleme istasyonu koordinatları ve pm10 değerleri

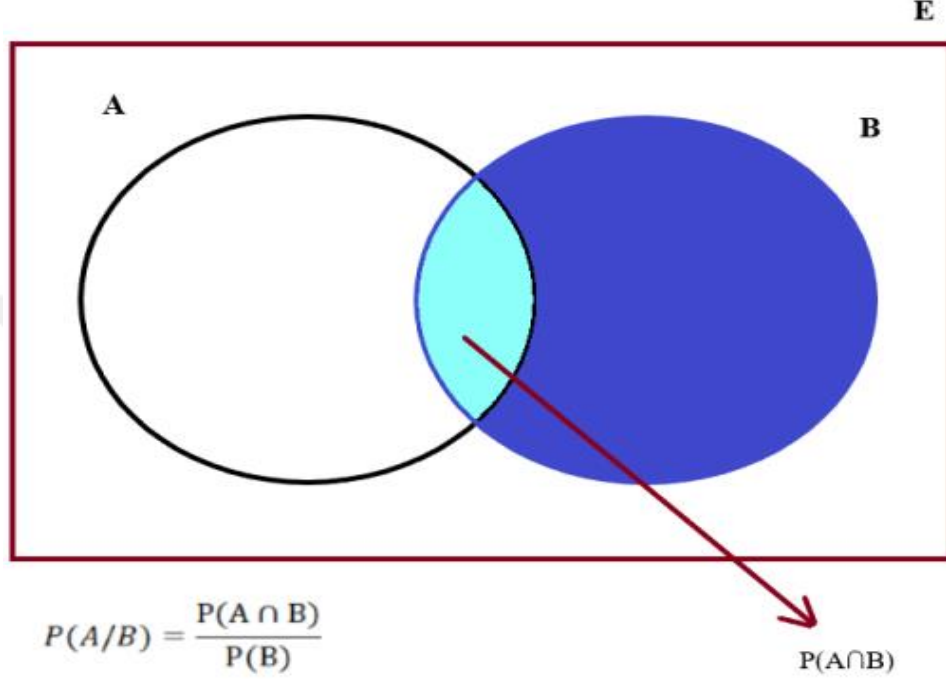
Çevre Şehircilik ve İklim Değişikliği Bakanlığı'nın Sürekli İzleme Merkezi internet sitesinden koordinatlar elde edilmiş olup, yine aynı sitenin veri bankası bölümünden ilgili PM10 verileri indirilerek elde edilmiştir.

3.6.5. Uygulama alanındaki fabrika verileri

Uygulama sahasındaki fabrika ya da hava kirliliği unsuru oluşturabilecek yapılar uydu görüntüsünden noktasal olarak belirlenip vektör veri formatına dönüştürüldü. Bu veri çalışmada kullanıldı.

3.7. Kullanılan Makine Öğrenmesi Algoritması(Naif Bayes)

3.7.1. Bayes teoremi



Şekil 3.4. Bayes Teoremini anlatan şekil

Bayes teoremini Şekil 3.4'te gösterildiği şekilde, kümeler olarak ifade edecek olursak, “ $P(A/B)$ ” olarak gösterilen B olayının olması durumunda A durumunun olması ihtimalinin, $P(A \cap B)$ olarak gösterilen, hem A hem de B'nin olma ihtimalinin, B'nin olması ihtimaline bölümüne eşit olmasıdır. Öncelikle B olayı gerçekleşiyor, sonrasında bu B olayının üzerinde A olayının gerçekleşme ihtimali olarak da ifade edilebilir (Şeker, 2020).

3.7.2. Naif bayes

$$P(Y/X) = \frac{P(X/Y)P(Y)}{P(X)} \quad (3.5)$$

$P(Y/X)$ = X olayının gerçekleşmesi durumunda Y olayının gerçekleşmesi ihtimali

$P(Y)$ =Y olayının gerçekleşme ihtimali

$P(X)$ =X olayının gerçekleşme ihtimali

Bayes teoreminin sağladığı en önemli imkan, koşulların yer değiştirilebiliyor olmasıdır. Yani, X gerçekleştikten sonra, Y'nin gerçekleşme ihtimali ile, Y gerçekleştikten sonra X'in gerçekleşme ihtimali arasında bir bağlantı kurulmasını sağlar.

Bu, gerçek dünyada bir sınıflandırma problemine dönüştürüldüğü zaman;

Çizelge 3.3. Bayes Teoremine örnek teşkil eden veri seti

Yaş	Gelir	Öğrenci	Kredi puanı	Bilgisayar alma durumu
<=30	Yüksek	Hayır	iyi	Hayır
<=30	Yüksek	Hayır	çok iyi	Hayır
31...40	Yüksek	Hayır	iyi	Evet
>40	Orta	Hayır	iyi	Evet
>40	Düşük	Evet	iyi	Evet
>40	Düşük	Evet	çok iyi	Hayır
31...40	Düşük	Evet	çok iyi	Evet
<=30	Orta	Hayır	iyi	Hayır
<=30	Düşük	Evet	iyi	Evet
>40	Orta	Evet	iyi	Evet
<=30	Orta	Evet	çok iyi	Evet
31...40	Orta	Hayır	çok iyi	Evet
31...40	Yüksek	Evet	iyi	Evet
>40	Orta	Hayır	çok iyi	Hayır

Sınıflar (bağımlı değişken):

Y1=Bilgisayar alma durumu: "Evet"

Y2= Bilgisayar alma durumu: "hayır"

Sonucu tahmin edilecek veri (bağımsız değişken):

X= yaş, gelir, öğrenci, kredi puanı

Yukarıdaki veri seti doğrultusunda, yaşı 30'un altında olanların kaçısı bilgisayar almıştır veya öğrenci olanların kaçısı bilgisayar almıştır, hesaplanabilir.

Aslında, Bayes teoremi kullanılarak, bilgisayar almayı etkileyen olayların herhangi birisinin bilgisayar alma üzerindeki etkisi ölçülebilir. Bu şekilde bağımsız değişken değerleri(X) verilen bir verinin, bağımlı değişken değeri(Y1,Y2) değeri tahmin edilebilir hale gelmiş olur.

3.7.3. Sayısal örnek

Çizelge 3.3'te verilen veri setine göre, Naif Bayes algoritması ile sınıflandırma yapıldığı bir durumda; bağımsız değişken değerleri $X = \text{yaş} \leq 30$, $\text{gelir} = \text{orta}$, $\text{öğrenci} = \text{evet}$, $\text{kredi puanı} = \text{iyi}$ olan örnek verinin bilgisayar alma durumunu hesaplanacak olursa;

$$P(Y1):P(\text{bilgisayar alma durumu} = \text{"evet"}) = 9/14 = 0.643$$

$$P(Y2):P(\text{bilgisayar alma durumu} = \text{"hayır"}) = 5/14 = 0.357$$

Veri setindeki toplam örnek sayısı 14, bilgisayar alma durumu evet olan 9, hayır olan 5 adet örnek olduğundan dolayı bağımlı değişken ihtimalleri yukarıdaki gibi olmaktadır. Her bir değişken için $P(X/Y)$ değeri hesaplanır.

$$P(\text{yaş} = \text{"<=30"} / \text{bilgisayar alma durumu} = \text{"evet"}) = 2/9 = 0.222$$

$$P(\text{yaş} = \text{"<=30"} / \text{bilgisayar alma durumu} = \text{"hayır"}) = 3/5 = 0.6$$

$$P(\text{gelir} = \text{"orta"} / \text{bilgisayar alma durumu} = \text{"evet"}) = 4/9 = 0.444$$

$$P(\text{gelir} = \text{"orta"} / \text{bilgisayar alma durumu} = \text{"hayır"}) = 2/5 = 0.4$$

$$P(\text{öğrenci} = \text{"evet"} / \text{bilgisayar alma durumu} = \text{"evet"}) = 6/9 = 0.667$$

$$P(\text{öğrenci} = \text{"evet"} / \text{bilgisayar alma durumu} = \text{"hayır"}) = 1/5 = 0.2$$

$$P(\text{kredi puanı} = \text{"iyi"} / \text{bilgisayar alma durumu} = \text{"evet"}) = 6/9 = 0.667$$

$$P(\text{öğrenci} = \text{"evet"} / \text{bilgisayar alma durumu} = \text{"hayır"}) = 2/5 = 0.4$$

$$P(X/Y):P(X/\text{bilgisayar alma durumu} = \text{"evet"}) = 0.222 \times 0.444 \times 0.667 \times 0.667 = 0.044$$

$$P(X/\text{bilgisayar alma durumu} = \text{"hayır"}) = 0.6 \times 0.4 \times 0.2 \times 0.4 = 0.019$$

$$P(X/Y) * P(Y):P(X/\text{bilgisayar alma durumu} = \text{"evet"}) * P(\text{bilgisayar alma durumu} = \text{"evet"}) = 0.028$$

$$P(X/\text{bilgisayar alma durumu} = \text{"hayır"}) * P(\text{bilgisayar alma durumu} = \text{"hayır"}) = 0.007$$

Sonuçta da görüldüğü üzere bilgisayar alma durumunun “evet” olan miktarın, toplam miktara bölünmesi sonucu “0.8” dir. Yani, belirlenmesi istenen verinin %80 oranında “evet” olduğu ortaya çıktığından, verilen bağımsız değişken değerlerine sahip verinin bilgisayar alma durumu da “evet” denilir (Şeker, 2020).

3.8. Yapılan Uygulamanın Naif Bayes Algoritmasına Uyarlanması

Bağımlı değişken(Y): Seviye=”Orta”, “Hassas”, “Sağlıksız”

Bağımsız değişkenler(X):band2, band3, band4, band5, band6, yer yüzeyi sıcaklığı, SAVI, BU indeksi, yola uzaklık değerleri, fabrikaya olan uzaklık değerleri

Çalışma alanında 30 metre çözünürlüğündeki uydu görüntüsü kullanılmıştır. Bu yüzden çalışma alanındaki tüm bağımsız değişkenler 30 metre çözünürlüğündeki raster veri formatında elde edilmiş ya da 30 metre çözünürlüğünde raster veri formatında oluşturulmuştur. Çalışma alanındaki her bir piksel bir örnek veri setini temsil etmekte ve her bir pikselin 10 adet bağımsız değişkene dair değeri bulunmaktadır. Bunun dışında makine öğrenmesi algoritmasının öğrenmesi için çalışma alanında bulunan hava kalitesi izleme istasyonlarının 850 metre uzaklığa kadar olan her bir pikselin bağımlı değişken değeri(Seviye) o hava kalitesi izleme istasyonunda ölçülen seviye ile aynı seviyede olduğu düşünüldü; fakat bu 850 metre yarıçapında yapılan analizi olumsuz etkileyebilecek durumlara bakılarak, makine öğrenme sürecinde, makineye verilen tutarsız örneklerden dolayı yanlış öğrenme durumuna yol açmamak için, makinenin öğrenmesi için makineye verilen veri setinde bazı düzenlemeler yapıldı. İleriki bölümlerde bu durum detaylı olarak açıklanacaktır.

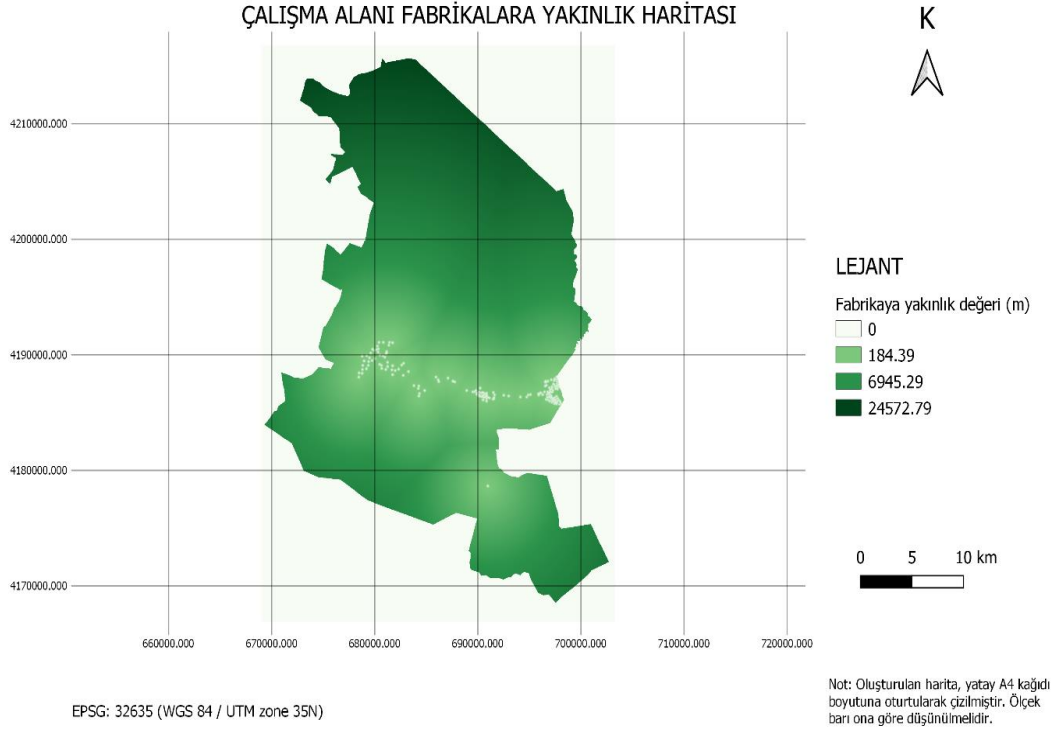
Makinenin öğrenmesi için oluşturulan bağımlı ve bağımsız değişken değerleri belli olan, Çizelge 3.2’deki veri formatında verilerin oluşturulması ile çalışma alanı içerisindeki, hava kalite seviyesi belli olmayan ama bağımsız değişken değerleri belli olan herhangi bir pikselin hava kalitesi seviyesi, yukarıdaki sayısal örnekte hesaplandığı gibi hesaplanabilir.

3.9. Kullanılan Programlar

Söz konusu çalışmada, mekânsal analiz, mekânsal işlemler, görselleştirme haritalama için QGIS yazılımı, yine mekânsal analiz, mekânsal işlem, görselleştirme, veri setini öğrenme ve test veri seti olarak ayırmada, makine öğrenme algoritması Naif Bayes kullanarak öğrenme ve tahmin işleri için R programlama dili kullanılmıştır.



çözünürlüğündeki her bir pikselin sıcaklık değeri aralığını göstermektedir. Harita incelendiğinde, özellikle çalışma alanının güney kısmının sıcaklık değerinde, belirgin bir düşüklük olduğu görülmektedir.

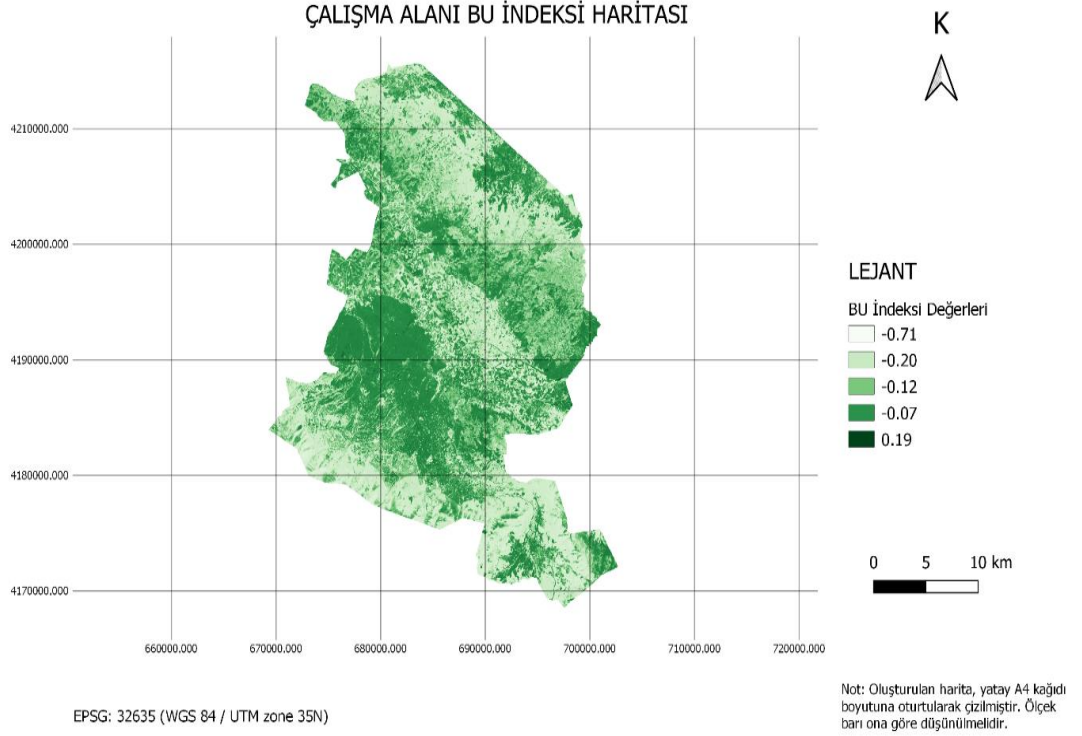


Şekil 4.2. 30 metre çözünürlüğe sahip fabrikalara uzaklık haritası

Şekil 4.2'deki harita, uydu görüntüsünden, nokta verisi olarak sayısallaştırılan fabrika ve işletmeler verisinin, QGIS yazılımı yardımı ile 30 metre çözünürlüğündeki raster formata dönüştürülüp, sonrasında, çalışma alanını kapsayacak şekilde, bu raster formattaki veriyi baz alarak, yine QGIS yazılımı yardımıyla yakınlık analizi yapılması suretiyle oluşturulmuştur.

Çalışma alanındaki her bir pikselin, çok önemli bir hava kirliliği kaynağı olan fabrikalara yakınlık değeri aralığını gösteren, Şekil 4.2'deki harita, aynı zamanda,

fabrikalara yakınlık değeri “0” olan, fabrikaların ya da işletmelerin (haritada nokta şeklinde parlak kısımlar) çalışma alanındaki dağılımını göstermektedir.

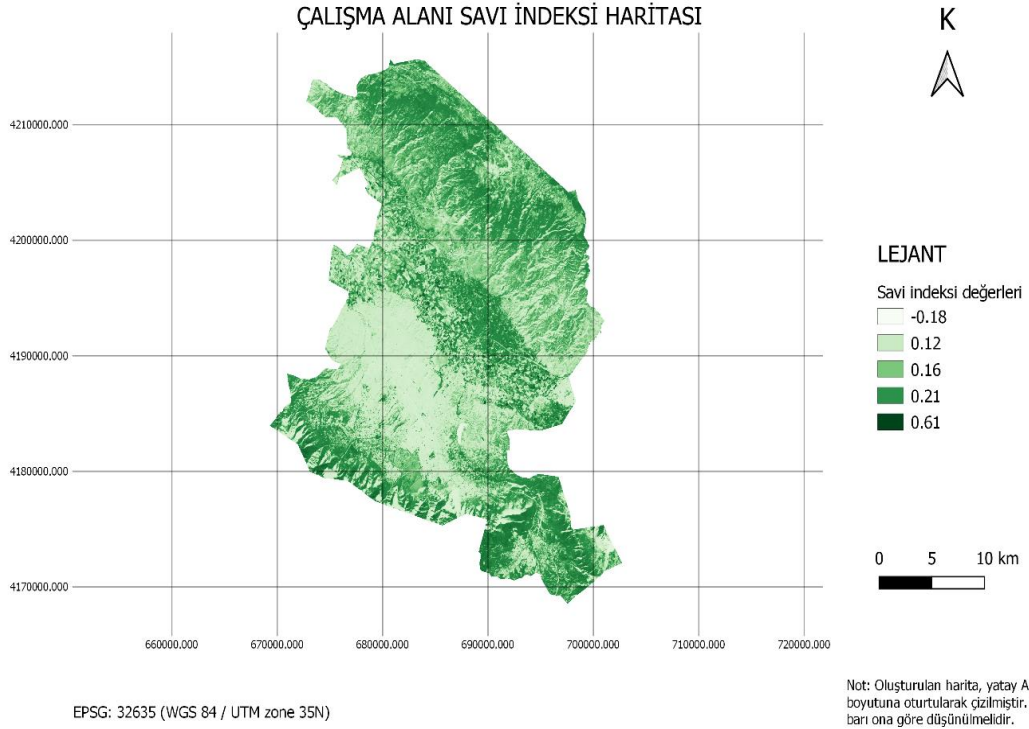


Şekil 4.3. 30 metre çözünürlüğe sahip BU indeksi haritası

Şekil 4.3'teki harita, üçüncü bölümde anlatılan, BU (Built-up) indeksini bulmaya yarayan eşitlik yardımıyla, indirilen uydu görüntüsündeki Band4, Band5 ve Band6 görüntüleri kullanılarak, QGIS yazılımında bulunan “raster hesaplayıcısı” aracıyla, ilgili eşitliğe göre, her bir pikseline denk gelen BU (Built-up) indeks değeri bulunup, bunun, sadece çalışma sınırının içerisindeki veriler gözükecek şekilde kırılmasıyla oluşturulmuştur.

Şekil 4.3'teki harita, yerleşik alanları vurgulayan, çalışma alanına dair BU (Built-up) indeks haritasını göstermektedir. Koyu renklere sahip piksellerin bulunduğu bölgeler, o bölgenin yerleşik alan olduğunu vurgularken, koyudan açık renklere doğru gidildikçe, yerleşik alanların azaldığını vurgulamaktadır.

Söz konusu harita, uydu görüntüsü ile karşılaştırılarak incelendiğinde; baskın koyu renkle gözüken bölgelerin şehrin yerleştiği bölge ve fabrika ve işletmelerin konuşlandığı bölge olduğu görülmüştür. Bu da BU (Built-up) indeksinin doğru şekilde hesaplandığını ortaya koymaktadır. Bu şekilde uydu görüntüsü ya da buna benzer yardımcı araçlarla elde edilen bağımsız değişken haritaların üst üste karşılaştırılıp incelenmesi ile, uygulama alanındaki her bir pikselin bağımsız değişkenlere dair hesaplanan değerlerinin doğruluğu kontrol edilebilir ve edilmelidir.



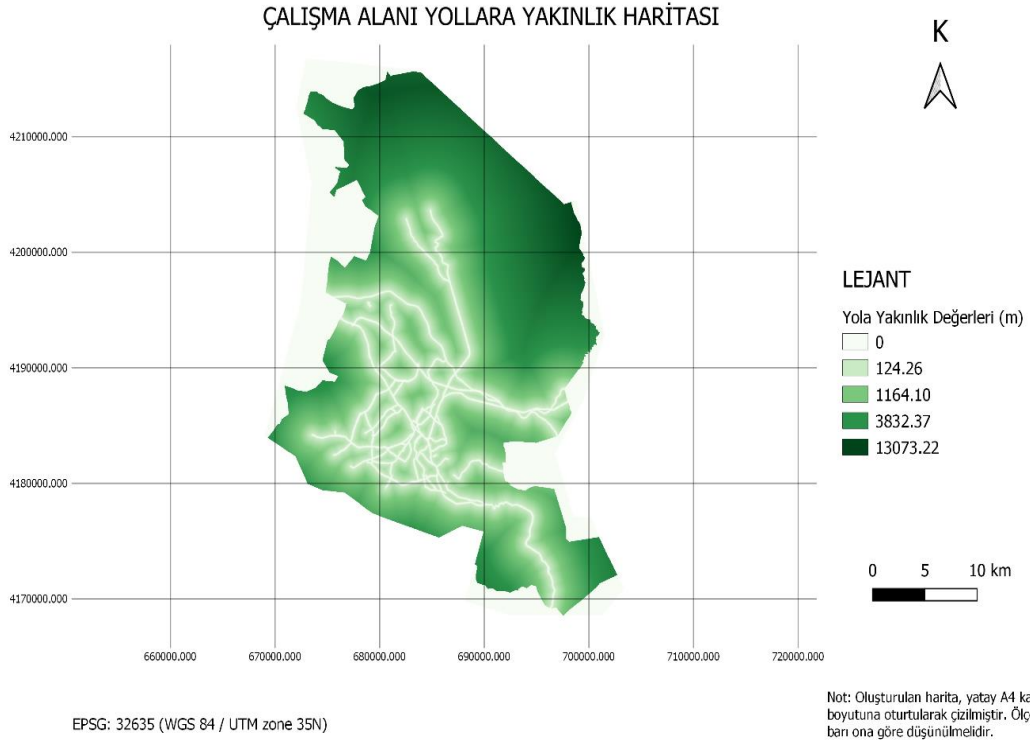
Şekil 4.4. 30 metre çözünürlüğe sahip SAVI indeksi haritası

Şekil 4.4'teki harita, üçüncü bölümde anlatılan, SAVI (Soil-Adjusted Vegetation Index) değerini bulmaya yarayan eşitlik yardımıyla, indirilen uydu görüntüsündeki Band4 ve Band5 görüntüleri kullanılarak, QGIS yazılımında bulunan "raster hesaplayıcısı" aracıyla, ilgili eşitliğe göre, her bir pikseline denk gelen SAVI

(Soil-Adjusted Vegetation Index) değerinin bulunup, bunun, sadece çalışma sınırının içerisindeki veriler gözükecek şekilde kırılmasıyla oluşturulmuştur.

Şekil 4.4'te, çalışma alanındaki, 30 metre çözünürlüğündeki her bir pikselin sahip olduğu, bitki örtüsü alanlarını ortaya çıkaran SAVI (Soil-Adjusted Vegetation Index) değer aralığını gösteren harita gösterilmektedir.

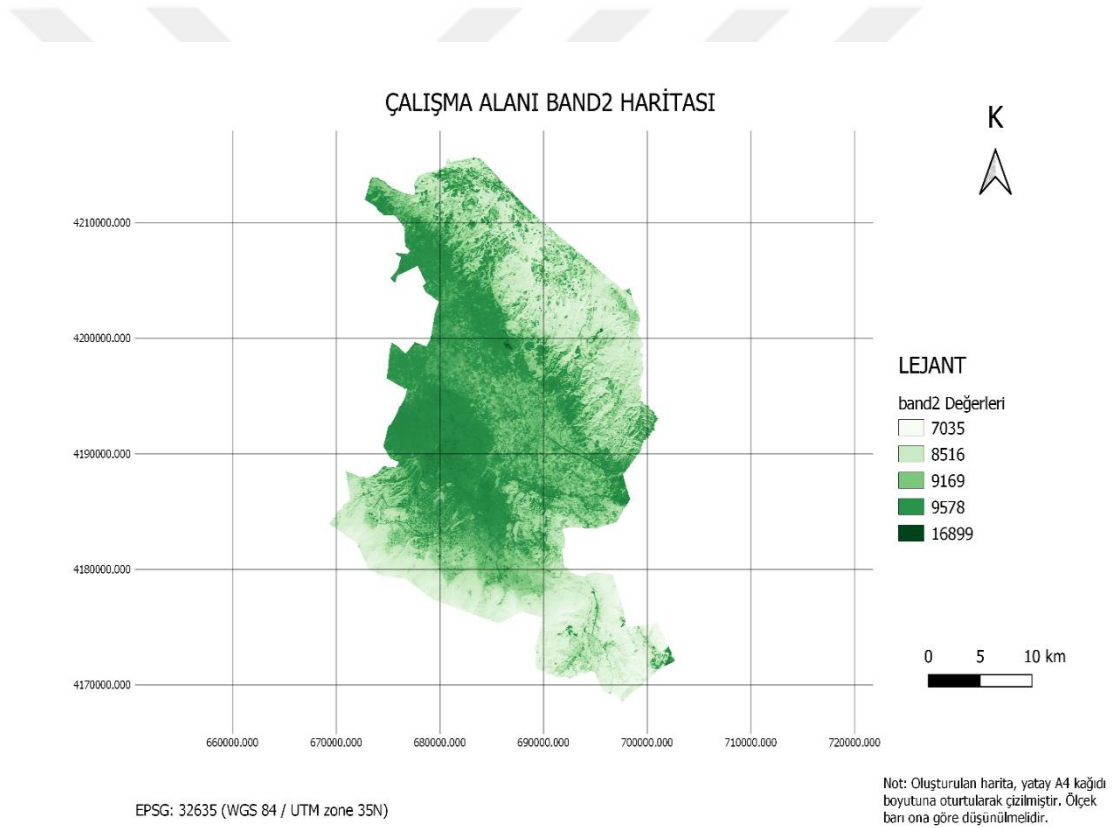
Haritada, SAVI (Soil-Adjusted Vegetation Index) değerinin yüksek olduğu koyu renkli alanlarda bitki örtüsü yoğunluğu varken, koyu yeşilden açık yeşile gidildikçe bitki örtüsü yoğunluğu azalmaktadır. Bu durumun da, söz konusu harita ile uydu görüntüsünün karşılaştırılması suretiyle teyidi yapılmıştır.



Şekil 4.5. 30 metre çözünürlüğe sahip yollara yakınlık haritası

Şekil 4.5'teki harita, Open Street Map vektör verisinden çekilen vektör formattaki yollar verisinin, QGIS yazılımı yardımı ile 30 metre çözünürlüğündeki raster formata dönüştürülüp, sonrasında, çalışma alanını kapsayacak şekilde, bu raster formattaki yol verisini baz alarak, yine QGIS yazılımı yardımıyla yakınlık analizi yapılması suretiyle oluşturulmuştur.

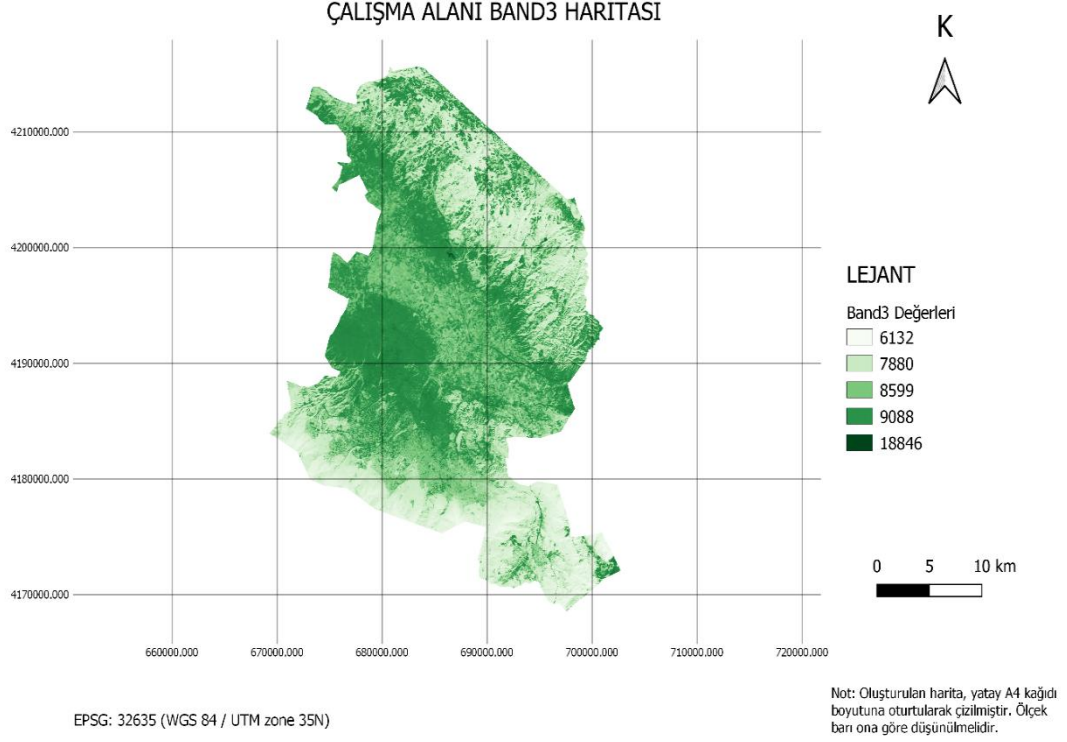
Çalışma alanındaki her bir pikselin, önemli bir hava kirliliği kaynağı olduğu düşünülen, yollara yakınlık değeri aralığını gösteren, Şekil 4.5'teki harita, aynı zamanda, yollara yakınlık değeri "0" olan, önem derecesi yüksek yolların (haritada çizgi şeklindeki parlak kısımlar) çalışma alanındaki dağılımını göstermektedir.



Şekil 4.6. 30 metre çözünürlüğe sahip band2 haritası

Şekil 4.6'daki harita, ilgili internet sitesinden indirilen uydu görüntüsündeki band2'ye ait raster formattaki verinin, sadece çalışma alanındaki veriler görünecek şekilde QGIS yazılımı kullanılarak kırılması suretiyle oluşturulmuştur.

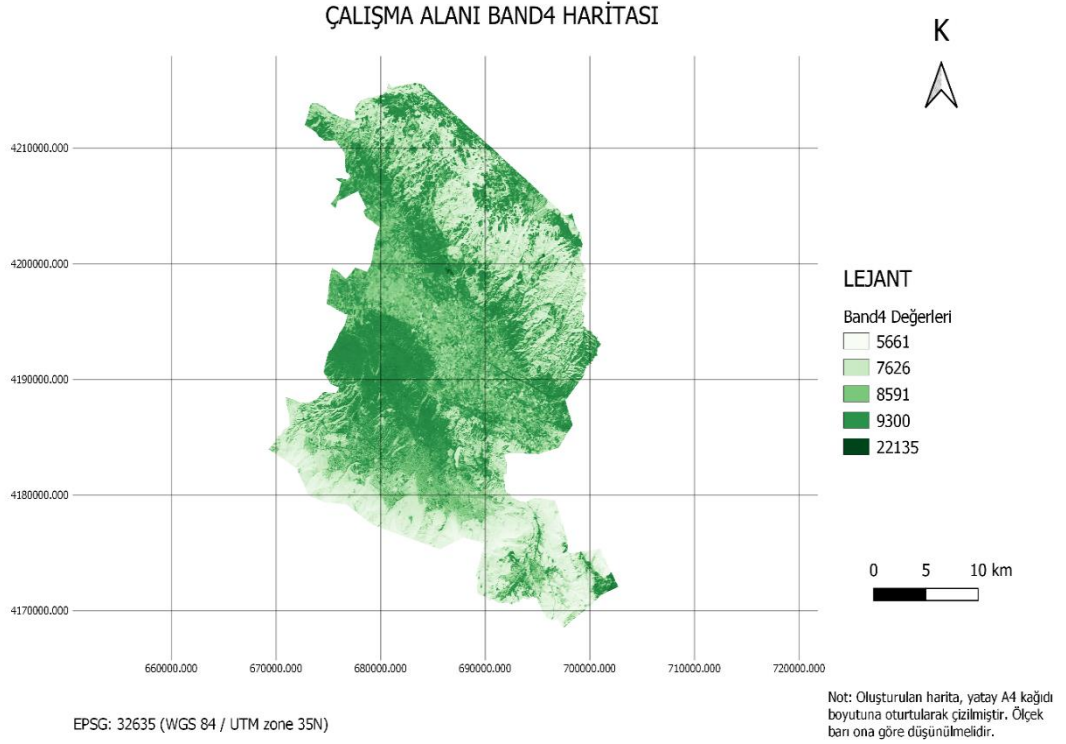
Şekil 4.6'daki harita, çalışma alanındaki, 30 metre çözünürlüğündeki her bir pikselin, band2 değer aralığını göstermektedir.



Şekil 4.7. 30 metre çözünürlüğe sahip band3 haritası

Şekil 4.7'deki harita, ilgili internet sitesinden indirilen uydu görüntüsündeki, band3'e ait raster formattaki verinin, sadece çalışma alanındaki veriler görünecek şekilde QGIS yazılımı kullanılarak kırılması suretiyle oluşturulmuştur.

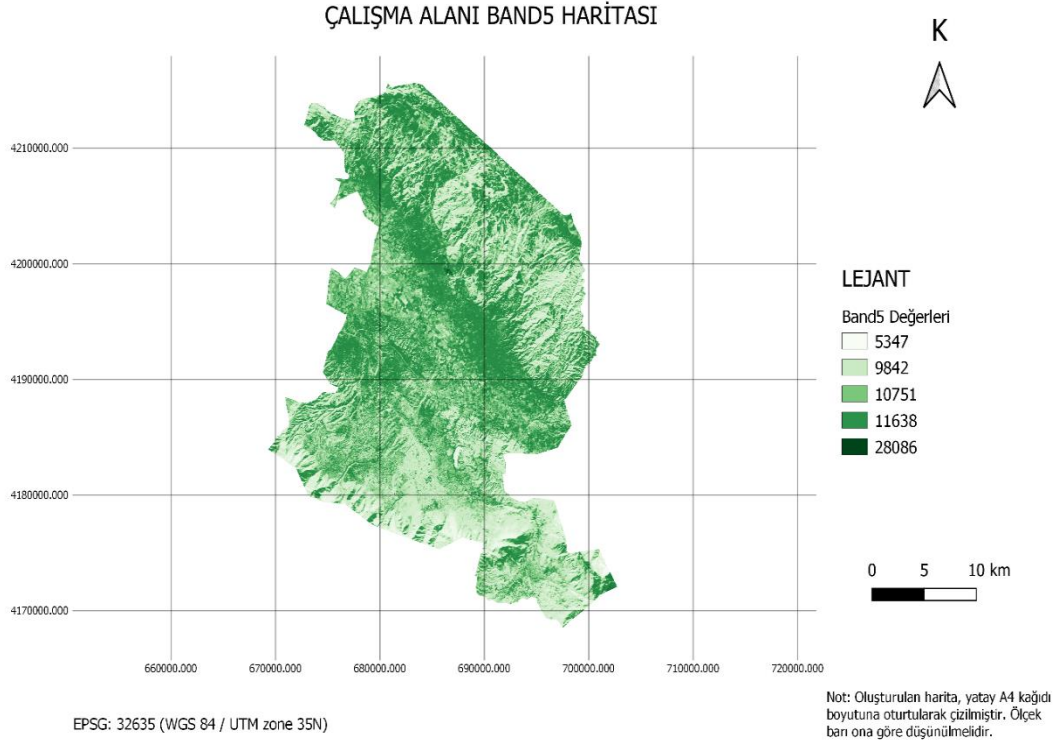
Şekil 4.7'deki harita, çalışma alanındaki, 30 metre çözünürlüğündeki her bir pikselin, band3 değer aralığını göstermektedir.



Şekil 4. 8. 30 metre çözünürlüğe sahip band4 haritası

Şekil 4.8'deki harita, ilgili internet sitesinden indirilen uydu görüntüsündeki, band4'e ait raster formattaki verinin, sadece çalışma alanındaki veriler görünecek şekilde QGIS yazılımı kullanılarak kırılması suretiyle oluşturulmuştur.

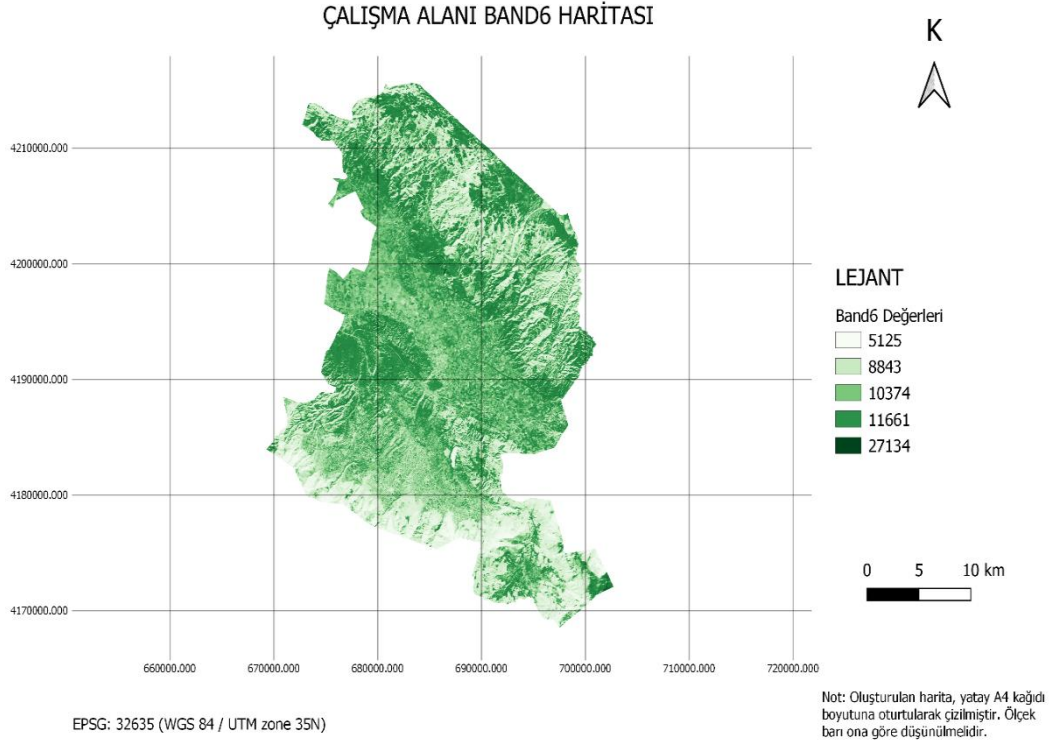
Şekil 4.8'deki harita, çalışma alanındaki, 30 metre çözünürlüğündeki her bir pikselin, band4 değer aralığını göstermektedir.



Şekil 4.9. 30 metre çözünürlüğe sahip band5 haritası

Şekil 4.9'daki harita, ilgili internet sitesinden indirilen uydu görüntüsündeki, band5'e ait raster formattaki verinin, sadece çalışma alanındaki veriler görünecek şekilde QGIS yazılımı kullanılarak kırılması suretiyle oluşturulmuştur.

Şekil 4.9'daki harita, çalışma alanındaki, 30 metre çözünürlüğündeki her bir pikselin, band5 değer aralığını göstermektedir.



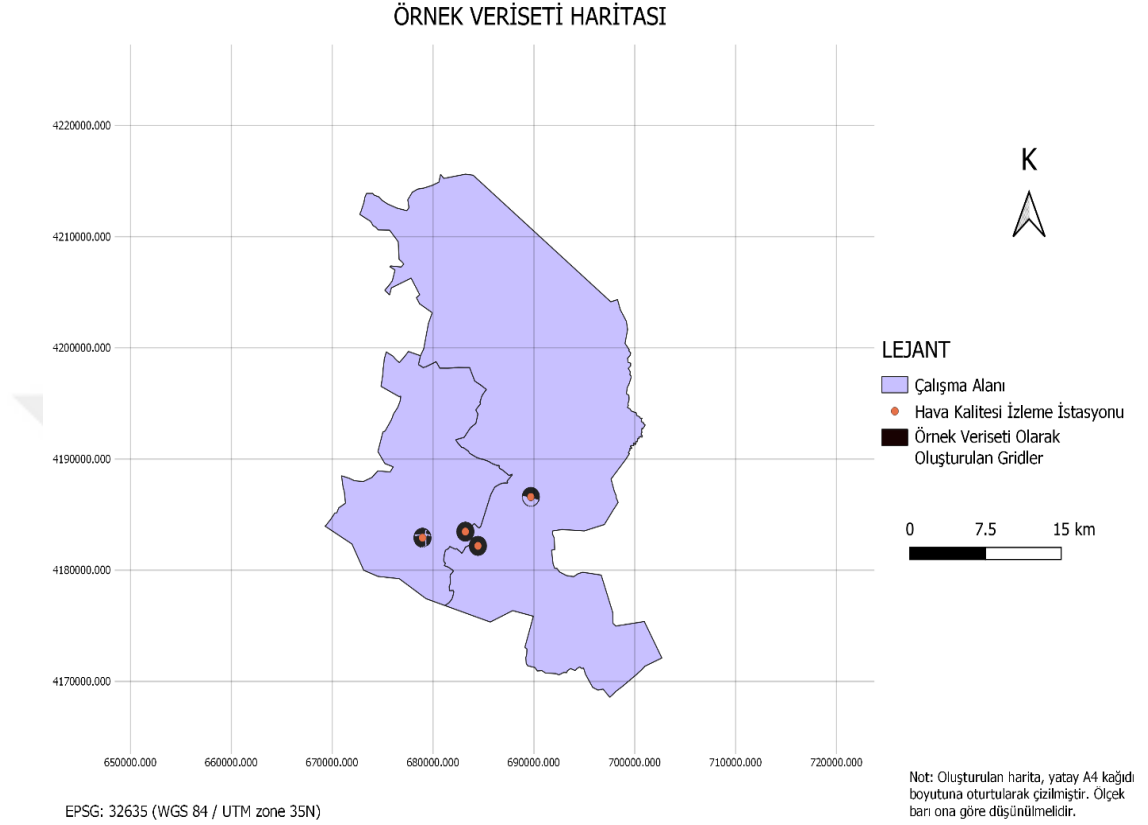
Şekil 4.10. 30 metre çözünürlüğe sahip band6 haritası

Şekil 4.10'daki harita, ilgili internet sitesinden indirilen uydu görüntüsündeki, band6'ya ait raster formattaki verinin, sadece çalışma alanındaki veriler görünecek şekilde QGIS yazılımı kullanılarak kırılması suretiyle oluşturulmuştur.

Şekil 4.10'daki harita, çalışma alanındaki, 30 metre çözünürlüğündeki her bir pikselin, band6 değer aralığını göstermektedir.

Yapılan çalışmada, elde edilen haritalardaki piksel boyutu 30 metre olduğundan her bir 30 metrelik alan bir birim veriyi temsil etmektedir. Bu her bir birim verinin, oluşturulan bağımsız değişken haritalarında da görüldüğü gibi, her bir bağımsız değişken için bir değeri vardır.

4.2. Makinenin Öğrenmesi İçin Veri Seti Oluşturulması

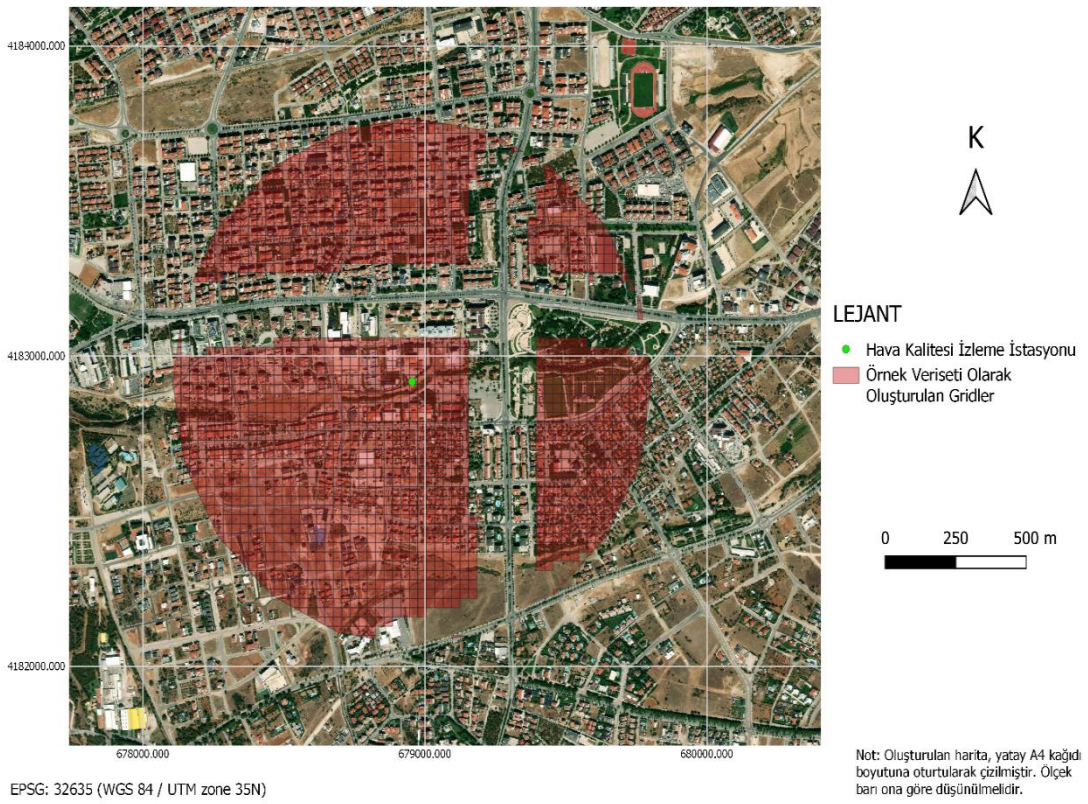


Şekil 4.11. Öğrenme verilerinin elde edilmesine esas harita

Yapılan uygulamada, makinenin öğrenebilmesi için, uygulama alanında her bir piksele denk gelen verinin, bağımlı ve bağımsız değişken değerlerinin belli olduğu veri grubunun makineye verilmesi gerekmektedir. Bunun için uygulama alanındaki hava izleme istasyonlarında ölçülen değerler mevcuttur. Fakat uygulama sahasında 4 adet istasyon olduğu için bu da 4 adet öğrenme verisi demektir. Bu veri seti makinenin öğrenmesi için yeterli değildir. Bu yüzden, daha fazla örnek oluşturabilmek için, uygulama alanındaki hava kalitesi izleme istasyonlarından 850 metre uzaklıktaki her alandaki hava kalitesi seviyesi o istasyonda ölçülen değer olarak düşünüldü. Bunun için QGIS yazılımında her bir istasyondan 850 metre mesafede tampon analizi yapılarak, bu 850 metre mesafedeki alanlarda 30 metrelik gridler oluşturuldu. 850 metrelik mesafe, istasyonların konumlarının (örnek veri seti için oluşturulan dairelerin

kesişmemesi gerekir) ve istasyonların bulunduğu bölgelerde uydu görüntüsünün incelenmesi sonucunda belirlenmiştir. Her bir grid, 1 adet örnek veriyi temsil etmektedir; fakat “Yapılan Uygulamanın Naif Bayes Algoritmasına Uyarlanması” bölümünde de bahsedildiği gibi, makinenin öğrenme sürecinde yanlış öğrenmesine mahal vermemek adına, oluşturulan gridler ile uygulama alanındaki uydu görüntüsü üst üste çakıştırılarak, yanlış öğrenmeye neden olabilecek örnek veriler, veri setinden silindi.

ÖRNEK VERİSETİ DETAY HARİTASI-1



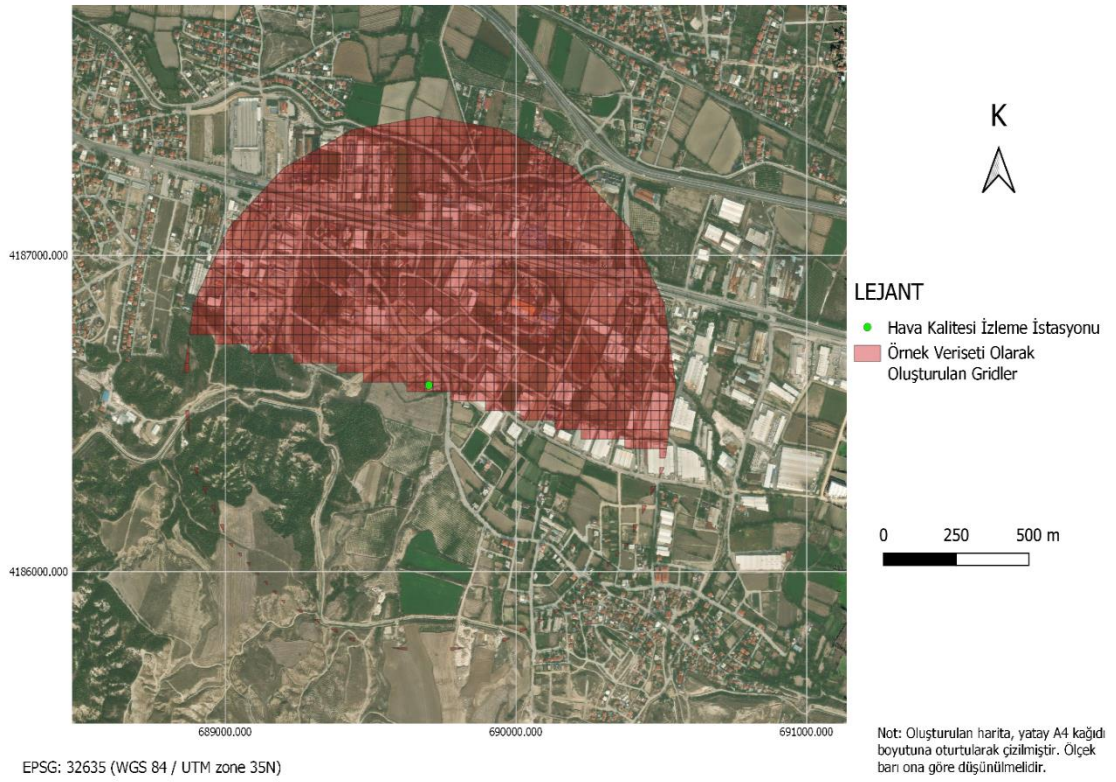
Şekil 4.12. Yanlış öğrenmeye sebep olabilecek verileri analiz etmeye esas harita-1

Şekil 4.12’de görüldüğü üzere, öğrenme verisi oluşturulması için belirlenen 850 metrelik yarıçapta oluşturulan tüm verilerin kullanılması, Şekil 4.12’deki istasyon için konuşacak olursak, makinenin yanlış öğrenip, yanlış tahmin yapmasına sebep olacaktır. Bunun sebebi, söz konusu görüntüde de görüldüğü üzere, bahse konu istasyonun 850 metre yarıçapında, normalde ilgili istasyon yola çok yakın değilken, yolun hemen üzerinde ve yolun çok yakınında öğrenme verilerinin oluşmasına sebep

olmaktadır. Bilindiği gibi, işlek yol hatlarına çok yakın olan alanlarda, hava kirliliği oranı daha fazladır. Yola çok yakın olmayan istasyondaki örnek veri ile, yola çok yakın ya da yolun hemen üzerinde bulunan örnek verinin, aynı hava kirliliği derecesine sahip olduğunu düşünerek işlem yapmak, yanlış analize yol açacaktır.

Bunu önlemek adına, daha önce elde edilmiş yol güzergâhları verisinden faydalanılarak, oluşturulan örnek verilerden yol güzergâhlarına 120 metre mesafede bulunan tüm veriler, öğrenme verisinden silinmiştir.

ÖRNEK VERİSETİ DETAY HARİTASI-2



Şekil 4.13. Yanlış öğrenmeye sebep olabilecek verileri analiz etmeye esas harita-2

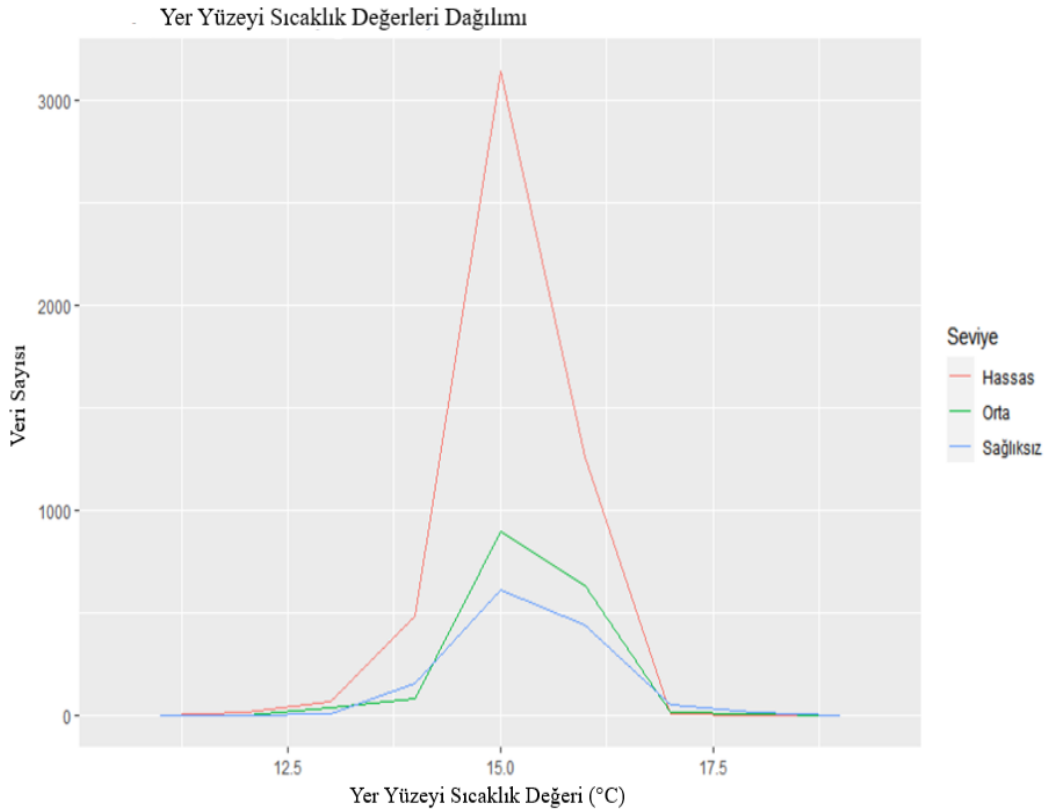
Şekil 4.13'de görülen istasyondan türetilen verilerin bir kısmının, görüldüğü şekilde silinmesinin sebebi ise şudur. İlgili izleme istasyonunun, uydu görüntüsüne bakıldığı zaman, fabrika ve işletmelerin yoğunlukta olduğu bir bölgede bulunduğu anlaşılmaktadır. Fabrika ve işletmelerin, genel itibarıyla istasyonun kuzey tarafında bulunduğu görülmektedir. Yine makinenin yanlış öğrenmesine engel olmak amacıyla,

şekil 4.13'deki şekilde, fabrikaların bulunmadığı bölümdeki öğrenme verileri silinmiştir.

Oluşturulan gridlerin bağımsız değişken değerleri, bir Coğrafi Bilgi Sistemi yazılımı olan “SAGA” yazılımındaki “grid statistics for polygons” aracının kullanılması suretiyle, bağımsız değişken haritalarından çekilmiştir. Oluşturulan gridlere dair bağımsız değişken değerleri ve bağımlı değişken değerleri birleştirilerek, makineye öğrenmesi için gerekli veri seti oluşturulmuştur.

4.3. Oluşturulan Veri Setine Dair Analizler

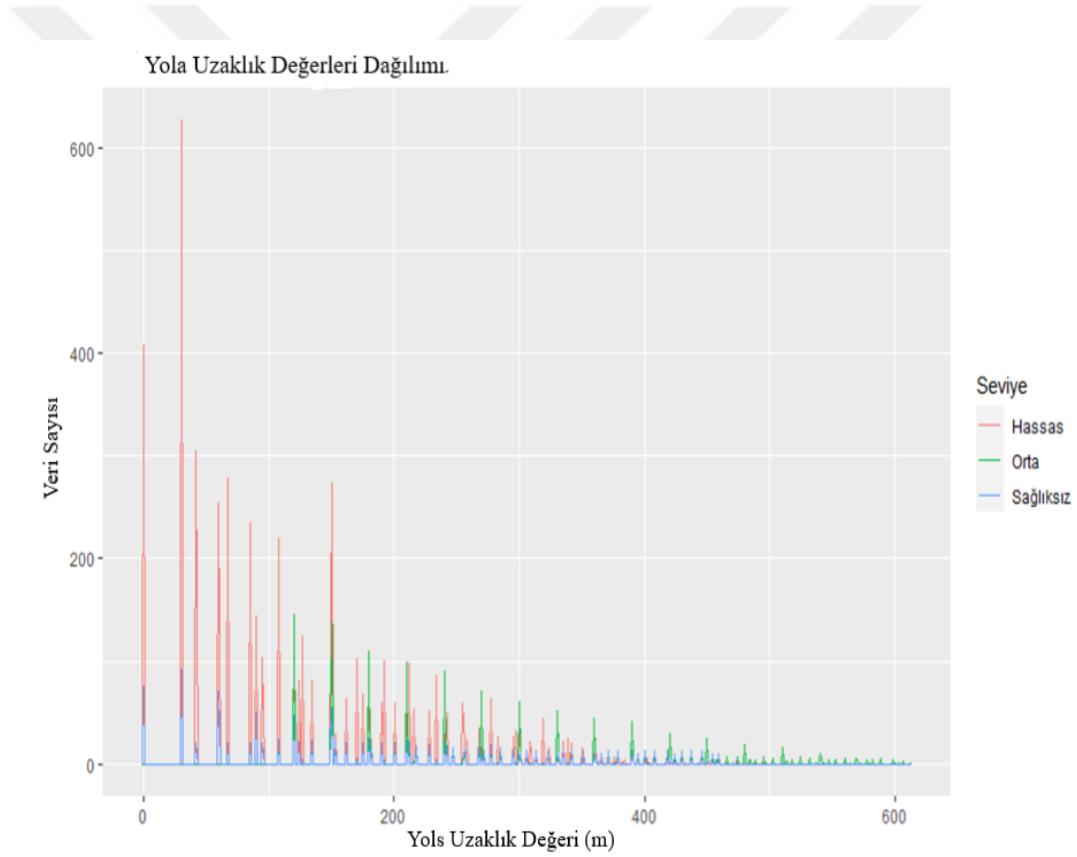
Makineye öğrenmesi için oluşturulan veri setini daha iyi analiz edip, gerekirse bu veri setini daha iyi duruma getirecek güncellemelerin yapılmasını sağlayabilecek grafikler, açıklamaları ile birlikte, bu bölümde sunulmuştur.



Şekil 4.14. Oluşturulan veri setinde yer yüzeyi sıcaklığının sınıflara göre dağılımı

Şekil 4.14'teki grafik, R programlama dilinde, veriler arasındaki ilişkileri anlamaya yardımcı olan ggplot2 isimindeki grafik kütüphanesi kullanılarak oluşturulmuştur.

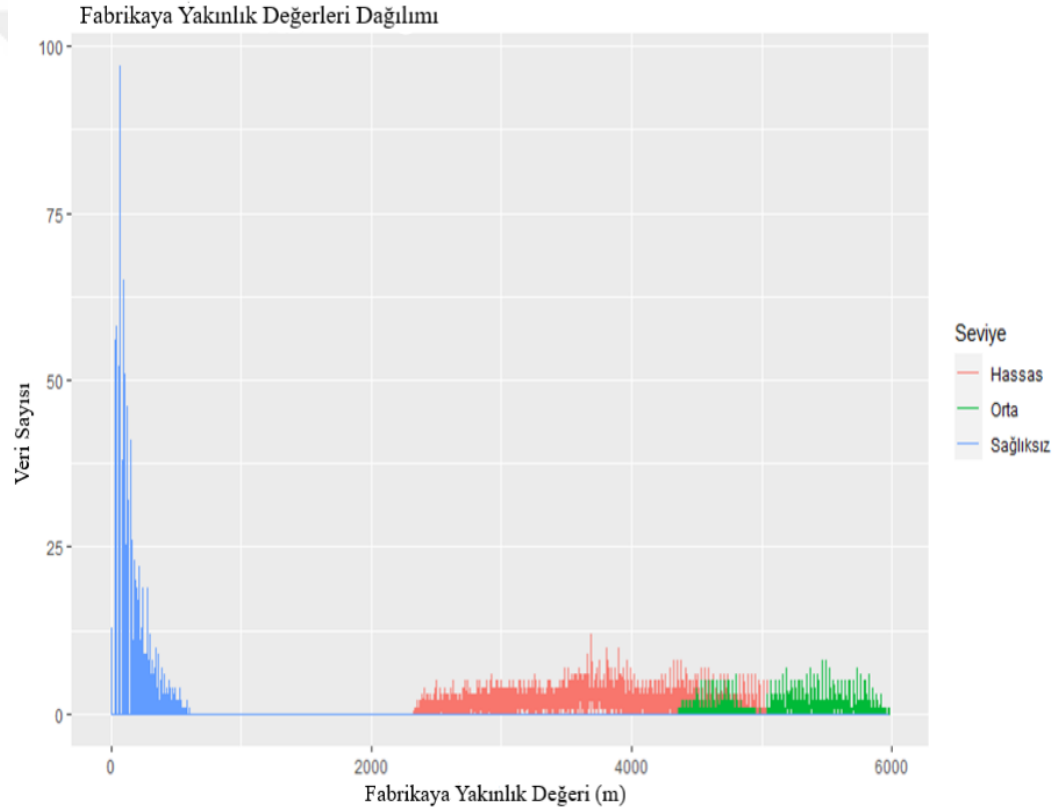
Makineye verilen öğrenme verilerini daha iyi anlayabilmek adına çizdirilen Şekil 4.14'teki grafik incelendiğinde, öğrenme verilerinin büyük kısmının, bağımlı değişkeni "hassas" sınıfı olan verilerden oluştuğu ve grafikteki sınıflara ait çizgi formlarına bakıldığı zaman, her bir sınıfa ait sıcaklık verilerinin ortalama olarak aynı değer aralıklarında olduğu görülmektedir. Böyle bir durum ise "yer yüzeyi sıcaklığı" değişkeninin, önem derecesinin yüksek bir bağımsız değişken olmadığını gösterir.



Şekil 4.15. Oluşturulan veri setinde yollara uzaklık değişkeninin sınıflara göre dağılımı

Şekil 4.15'teki grafik, R programlama dilinde, veriler arasındaki ilişkileri anlamaya yardımcı olan ggplot2 isimindeki grafik kütüphanesi kullanılarak oluşturulmuştur.

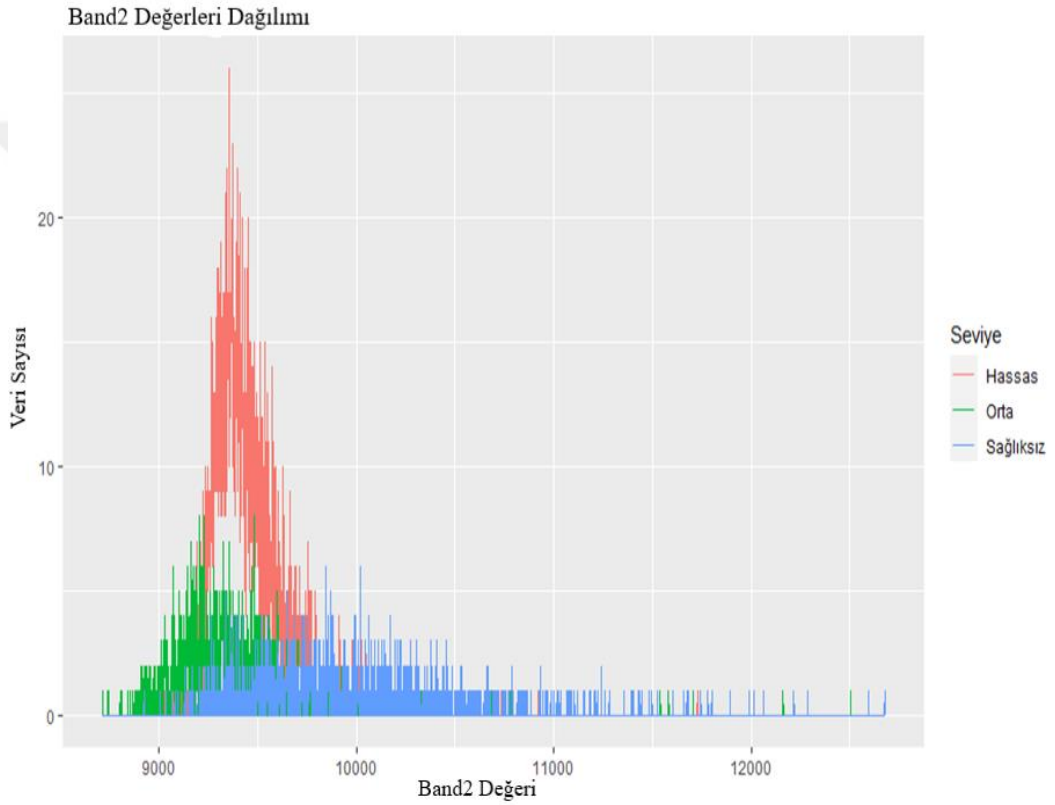
Makineye verilen öğrenme verilerini daha iyi anlayabilmek adına çizdirilen Şekil 24'teki grafik incelendiğinde, yine öğrenme verilerinin büyük kısmının, bağımlı değişkeni “hassas” sınıfı olan verilerden oluştuğu ve yola uzaklık değeri arttıkça, genel olarak “orta” sınıfa ait veri yoğunluğunun arttığı görülmektedir. Bu durum, yoldan uzaklığın arttıkça, hava kirliliğinin azaldığını göstermektedir. Grafiğe bakıldığında, yola uzaklık değeri düşük olan yerlerde, neden “sağlıksız” sınıfına ait veri yoğunluğu değil de, “hassas” sınıfa ait veri yoğunluğu gözüktüğü sorgulanabilir ve sorgulanmalıdır. Bunun sebebi, öğrenme veri setinde “hassas” sınıfa ait verinin, “sağlıksız” sınıfa ait veriden çok daha fazla olmasıdır.



Şekil 4.16. Oluşturulan veri setinde fabrikaya yakınlık değişkeninin sınıflara göre dağılımı

Şekil 4.16'daki grafik, R programlama dilinde, veriler arasındaki ilişkileri anlamaya yardımcı olan ggplot2 isimindeki grafik kütüphanesi kullanılarak oluşturulmuştur

Makineye verilen öğrenme verilerini daha iyi anlayabilmek adına çizdirilen Şekil 4.16'daki grafik incelendiğinde, fabrika ve işletmelere yakın olan tüm verilerin “sağlıksız” sınıfına ait olduğu, en uzak mesafede olan verilerin “orta” sınıfa ait olduğu, en uzak mesafede olan verilere nispeten daha yakında olan verilerin ise genel olarak “hassas” sınıfına ait olduğu görülmektedir. Bu durum, “fabrika ve işletmelere yakınlık” bağımsız değişkeninin, belirgin şekilde önem derecesi yüksek bir değişken olduğunu göstermektedir.

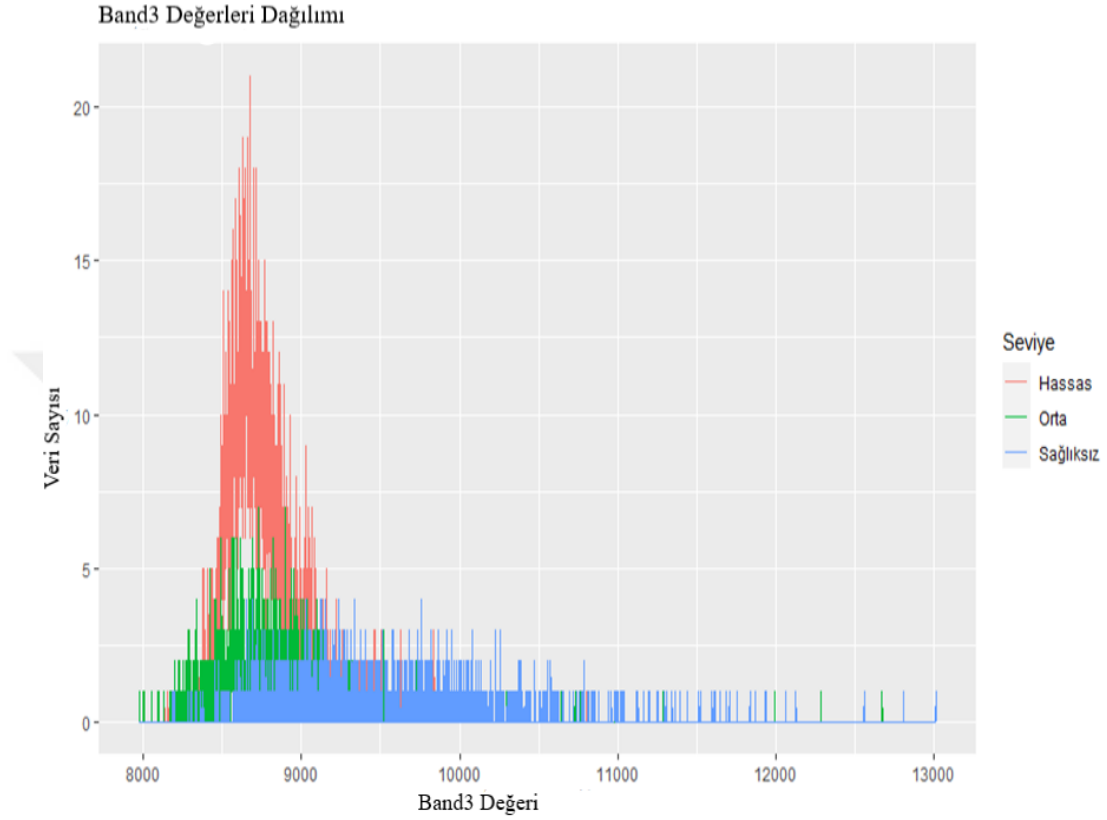


Şekil 4.17. Oluşturulan veri setinde band2 değişkeninin sınıflara göre dağılımı

Şekil 4.17'deki grafik, R programlama dilinde, veriler arasındaki ilişkileri anlamaya yardımcı olan ggplot2 isimindeki grafik kütüphanesi kullanılarak oluşturulmuştur.

Şekil 4.17'deki grafik incelendiğinde, farklı sınıflara ait verilerin band2 değerlerinin genel olarak 9000 ile 11000 değerleri arasında toplandığı; fakat buna

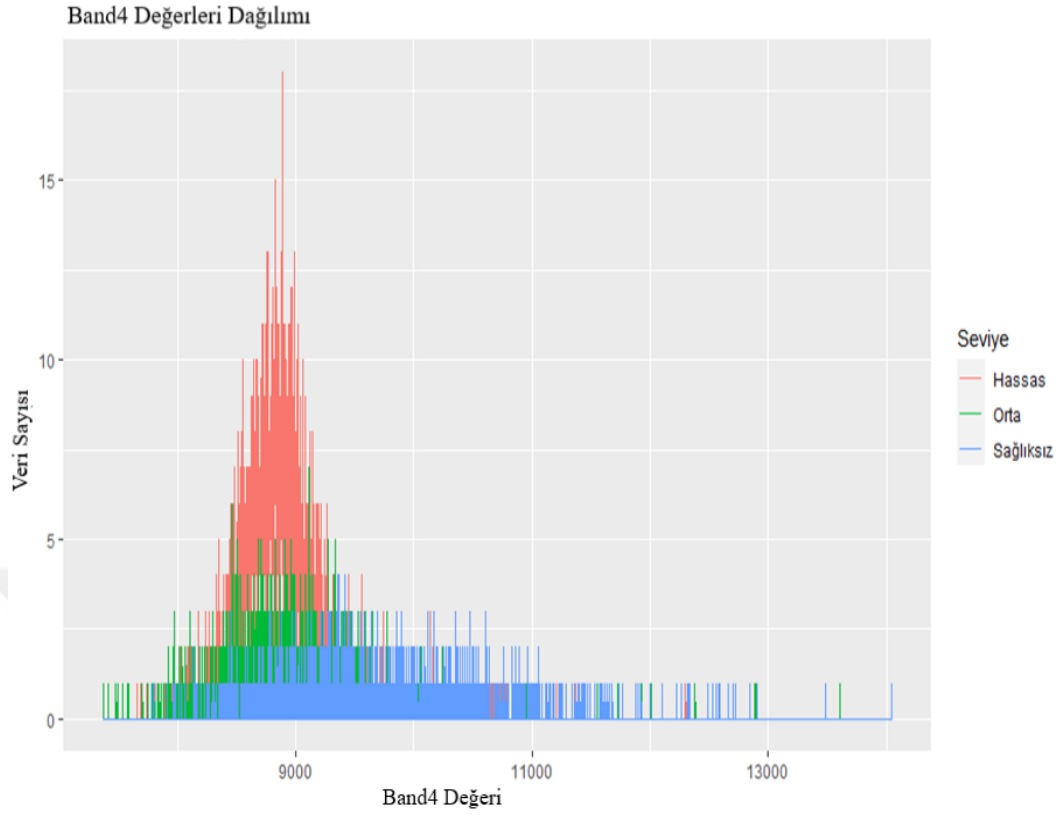
rağmen, farklı sınıflara ait verilerin toplandığı band2 değerleri arasında keskin bir ayırım olmasa da, yapılan çalışmaya yardımcı olabilecek ölçüde bir ayırım olduğu gözlenmektedir.



Şekil 4.18. Oluşturulan veri setinde band3 değişkeninin sınıflara göre dağılımı

Şekil 4.18'deki grafik, R programlama dilinde, veriler arasındaki ilişkileri anlamaya yardımcı olan ggplot2 isimindeki grafik kütüphanesi kullanılarak oluşturulmuştur.

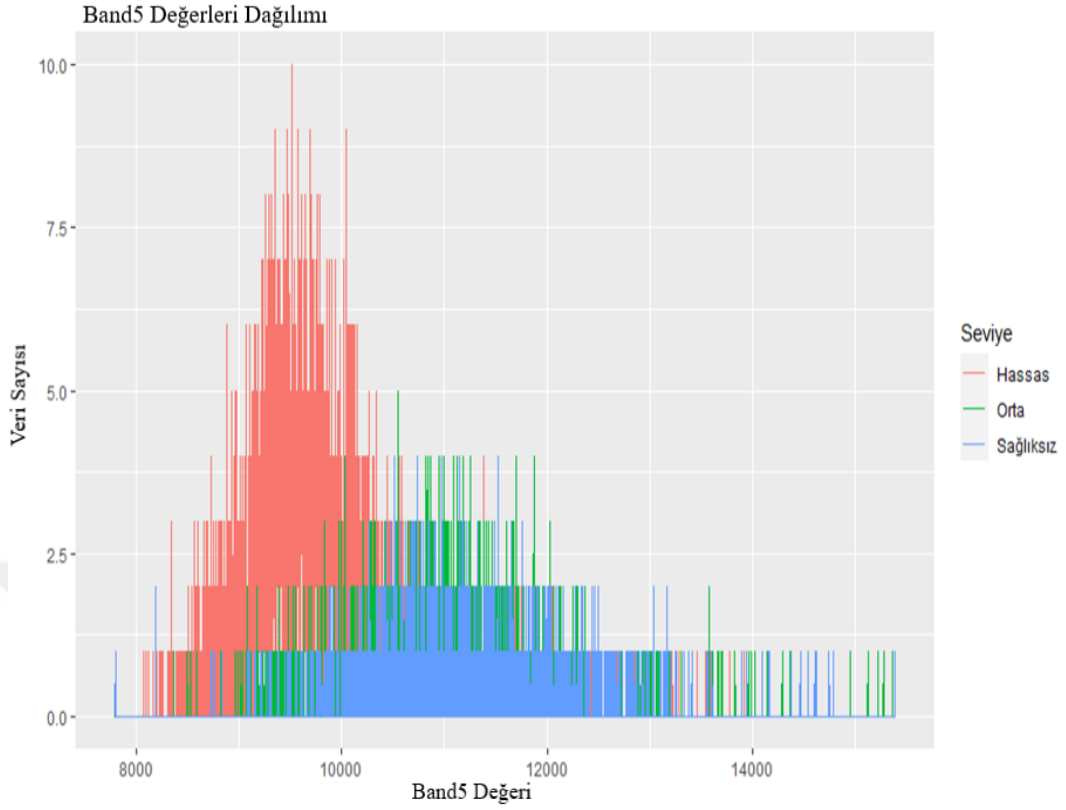
Şekil 4.18'deki grafik incelendiğinde, farklı sınıflara ait verilerin, toplandığı band3 değerleri arasında, yapılan çalışmaya yardımcı olabilecek ölçüde bir ayırım olduğu gözlenmektedir.



Şekil 4.19. Oluşturulan veri setinde band4 değişkeninin sınıflara göre dağılımı

Şekil 4.19'daki grafik, R programlama dilinde, veriler arasındaki ilişkileri anlamaya yardımcı olan ggplot2 isimindeki grafik kütüphanesi kullanılarak oluşturulmuştur.

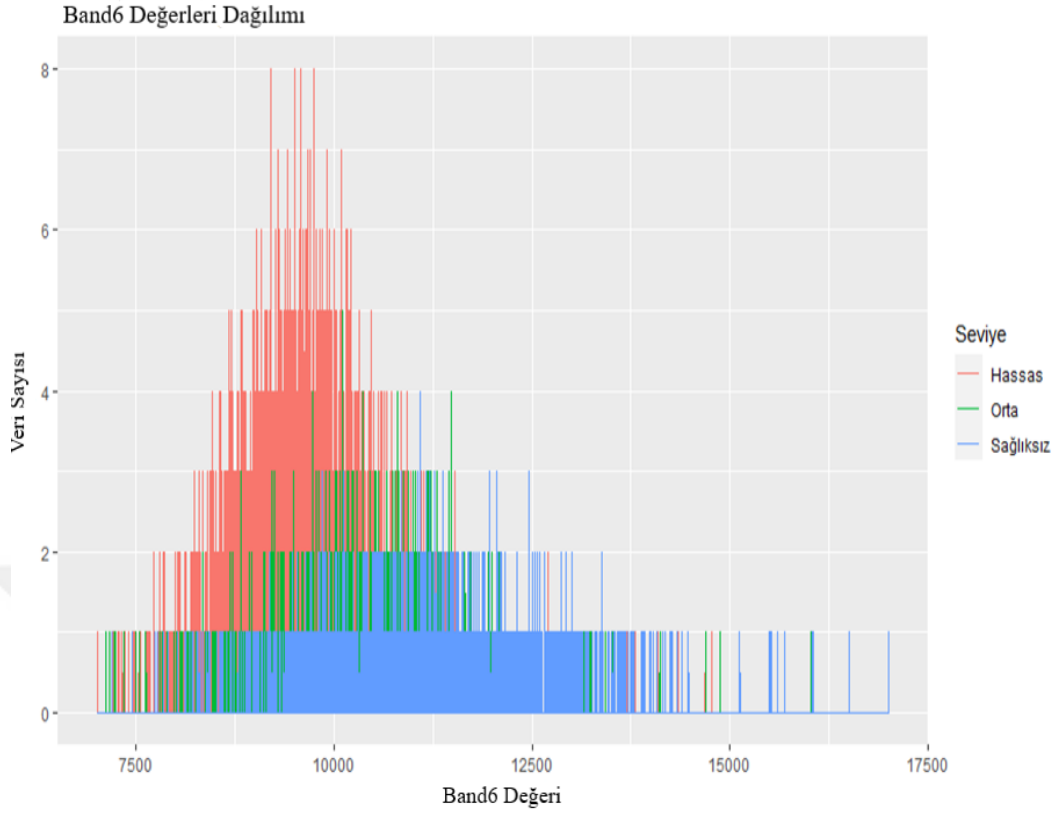
Şekil 4.19'deki grafik incelendiğinde, farklı sınıflara ait verilerin, toplandığı band4 değerleri arasında da, yapılan çalışmaya yardımcı olabilecek ölçüde bir ayrım olduğu gözlenmektedir.



Şekil 4.20. Oluşturulan veri setinde band5 değişkeninin sınıflara göre dağılımı

Şekil 4.20'deki grafik, R programlama dilinde, veriler arasındaki ilişkileri anlamaya yardımcı olan ggplot2 isimindeki grafik kütüphanesi kullanılarak oluşturulmuştur.

Şekil 4.20'deki grafik incelendiğinde, farklı sınıflara ait verilerin, toplandığı band5 değerleri arasında, band2, band3 ve band4 grafiklerindeki gibi daha belirgin bir ayırım söz konusu olduğu gözlemlenmektedir. Bu durum da, band5 değişkeninin, band2, band3 ve band4 değişkenlerine göre önem derecesinin daha yüksek olduğunu göstermektedir.



Şekil 4.21. Oluşturulan veri setinde band6 değişkeninin sınıflara göre dağılımı

Şekil 4.21'deki grafik, R programlama dilinde, veriler arasındaki ilişkileri anlamaya yardımcı olan ggplot2 isimindeki grafik kütüphanesi kullanılarak oluşturulmuştur.

Şekil 4.21'deki grafik incelendiğinde, farklı sınıflara ait verilerin, toplandığı band6 değerleri arasında da, band2, band3 ve band4 grafiklerindeki gibi daha belirgin bir ayırım söz konusu olduğu gözlemlenmektedir.

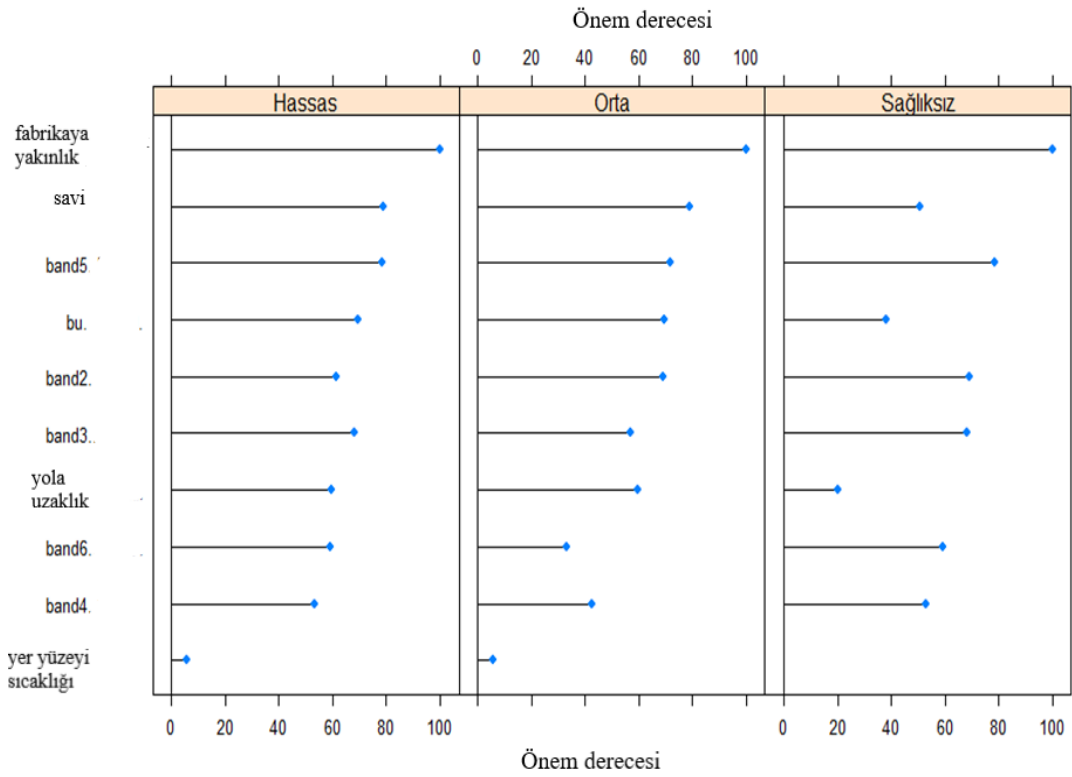
Veri setini daha iyi anlayabilmek için çizdirilen grafiklerde, sınıflar arasında belli ayrımlar yapılabilmesi, belirlenen bağımsız değişkenin modellenmeye çalışılan bağımlı değişken için önemli bir faktör olduğunu gösterir. Sınıflar arası belli ayrımlar yapılması zor ise bu bağımsız değişkenin ise, iyi bir faktör olmadığı, belki de bağımsız değişkenlerden bu faktörü çıkarmamız gerektiği anlamına gelir.

4.4. Oluşturulan Veri Setine Naif Bayes Algoritmasının Uygulanması

Öncelikle, Naif Bayes algoritmasının veri setine uygulanması için, R programlama dili ile yazılan Caret kütüphanesi kullanıldı.

Makinenin öğrenmesi için, bağımlı ve bağımsız değişken değerleri belirlenerek oluşturulan örnek veri seti, %65 öğrenme ve %35 test verisi olarak bölündü. Yani, bu demektir ki, makine verinin %65 ini modeli eğitmek için, Bayes teoreminde bahsedilen koşullu olasılık değerlerini hesaplayıp, bu olasılıkların modellenmeye çalışan bağımlı değişken ile, bağımsız değişkenler arasındaki ilişkiyi modellemek için kullanıp, geri kalan %35'lik kısmını ise, eğitilen modeli test etmede kullanacaktır.

Makine öğrenmesi algoritması oluşturulan veri setine uygulanarak, eğitim modeli oluşturulmuştur. Oluşturulan modelden her bir bağımsız değişkenin, modellenmek istenen bağımlı değişkene etki derecesini gösteren grafikler yazdırıldı.



Şekil 4.22. Bağımsız değişkenlerin bağımlı değişkene olan etki derecesini gösteren grafik

Şekil 4.22’de, her bir hava kirliliği sınıfı bazında, oluşturulacak tahmin haritasını etki eden değişkenlerin, verilen öğrenme veri setine göre ortaya çıkan önem dereceleri gösterilmektedir.

Görüldüğü üzere, önem derecesi en yüksek değişken fabrikaya yakınlık değişkeni olurken, önem derecesi en düşük değişkenin ise yer yüzeyi sıcaklığı değişkeni olduğu görülmektedir.

4.5. Oluşturulan Modelin Değerlendirilmesi

4.5.1. Hata matrisi

Çizelge 4.1. Oluşturulan modelin Hata Matrisi

	Hassas(Tahmin)	Orta(Tahmin)	Sağlıksız(Tahmin)	Toplam
Hassas(Gerçek)	1616	71	1	1688
Orta(Gerçek)	62	499	10	571
Sağlıksız(Gerçek)	32	9	470	511
Toplam	1710	579	481	2770

Çizelge 4.1’de, kırmızı renkle yazılmış değerler denk geldikleri sınıflarda doğru tahmin edilmiş veri sayısını göstermektedir.

Toplam örnek sayısı:2770

Doğru tahmin edilen örnek sayısı:2585

Doğruluk: $2585/2770=0,9332$

Bu modelimizin toplam doğruluğunu vermektedir. Ama bu tek başına yeterli değildir. Tahmin edilen sınıflara göre doğruluk analizlerinin de yapılması gerekir.

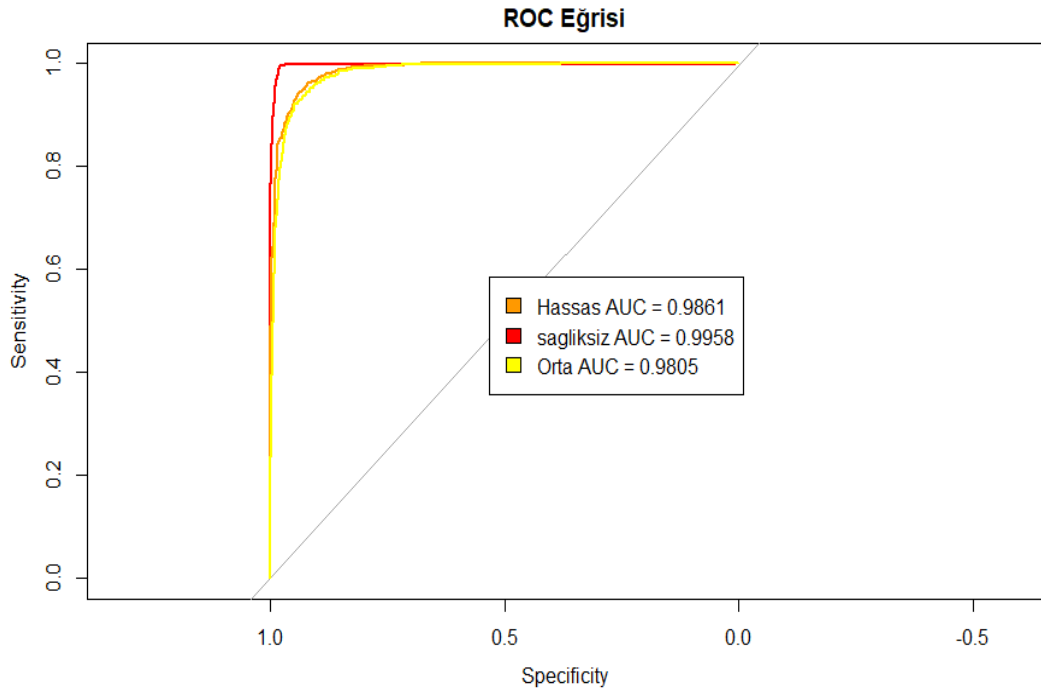
Çizelge 4.2. Hata Matrisine göre bulunan değerlendirme ölçütleri

	Hassas(Tahmin)	Orta(Tahmin)	Sağlıksız(Tahmin)
Duyarlılık	0.945	0.8618	0.9771
Spesifilik	0.9321	0.9671	0.9821
Pozitif tahmin değeri	0.9573	0.8739	0.9198
Negatif tahmin değeri	0.9131	0.9636	0.9951
Yaygınlık	0.6173	0.2090	0.1736
Bulma Oranı	0.5834	0.1801	0.1697
Bulma Yaygınlığı	0.6094	0.2061	0.1845
Dengelenmiş Doğruluk	0.9386	0.9145	0.9796

Çizelge 4.2’de de görüldüğü üzere, her bir sınıfa ait dengelenmiş doğruluğun da en az %91 oranında olduğu görülmektedir.

4.5.2. ROC(Receiver Operating Characteristic) eğrisi

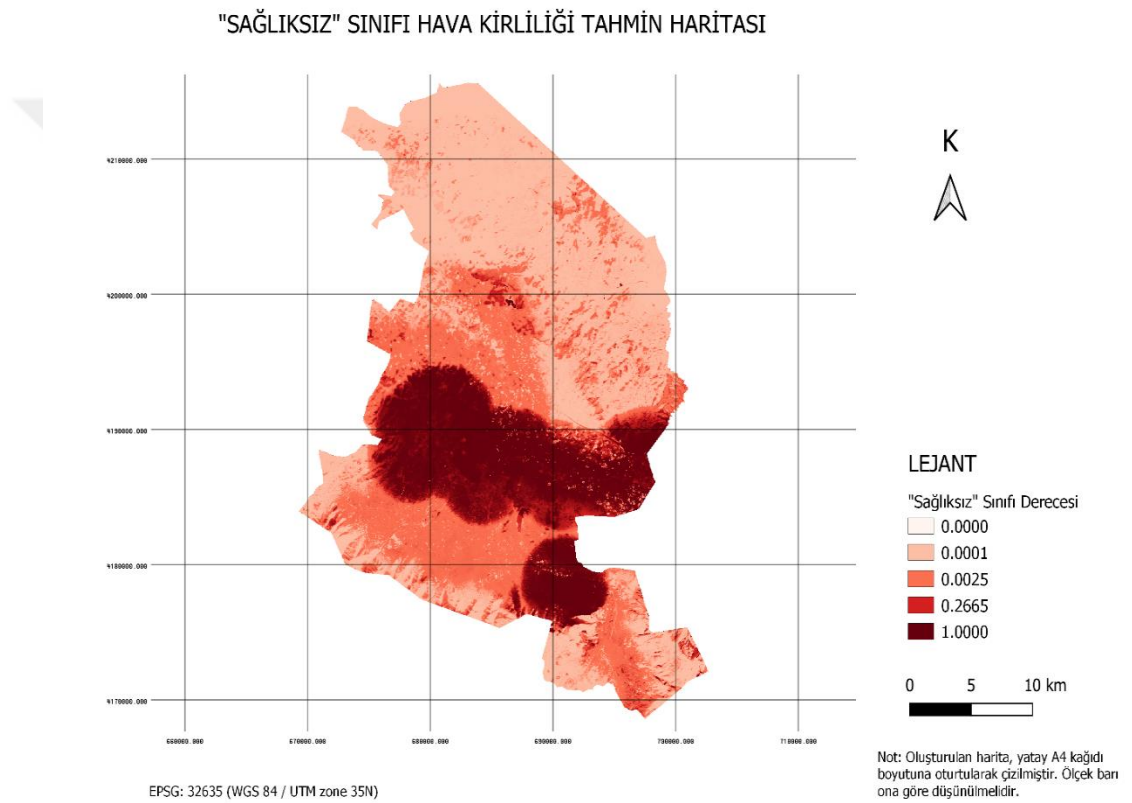
Alıcı İşletim Karakteristiği olarak çevirisi yapılabilecek olan ROC(Receiver Operating Characteristic) eğrisi, oluşturulan modelin görsel olarak değerlendirilmesine imkan sağlamaktadır.



Şekil 4.23. Bağımsız değişkenlerin bağımlı değişkene olan etki derecesini gösteren grafik

Şekil 4.23'teki kırmızı çizgi, "sağlıksız" sınıfını, turuncu çizgi, "hassas" sınıfını, sarı çizgi ise, "orta" sınıfını göstermektedir. Grafikte 1'e 1 noktası en iyi modeli temsil ederken, herhangi bir model bu noktaya ne kadar yakın olursa, o kadar iyi bir model olmaktadır. Yine grafikteki AUC (Eğri Altındaki Alan) değerleri ise, 1'e 1 noktasına hangi oranda yaklaşıldığını göstermektedir.

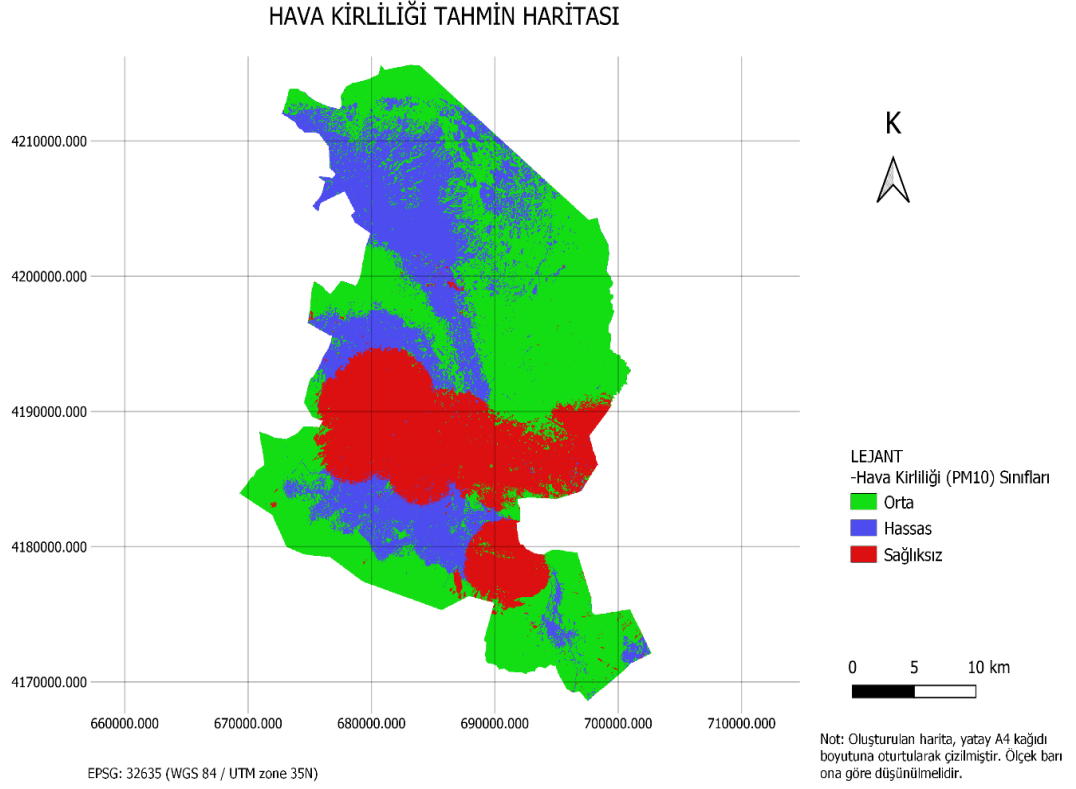
4.6. Tahmin Haritaları



Şekil 4.24. Çalışma alanı PM10 değeri sağlıksız sınıfı tahmin haritası

Şekil 4.24'teki harita, R programlama dilinde, bir veri kümesinde çoklu değişkenlerin bir arada gösterimini sağlayan "ssplot" fonksiyonu kullanılarak, sonrasında da QGIS programı yardımıyla görselleştirme ve haritalama işlemlerinin yapılması ile oluşturulmuştur. Bu harita da aslında, bir nevi, tahmin edilmeye çalışılan bağımlı değişkene (hava kirliliği seviyesi) etki eden faktörler olan bağımsız değişkenlerin, bağımlı değişken seviyesi olan "sağlıksız" sınıfı cinsinden gösterimidir.

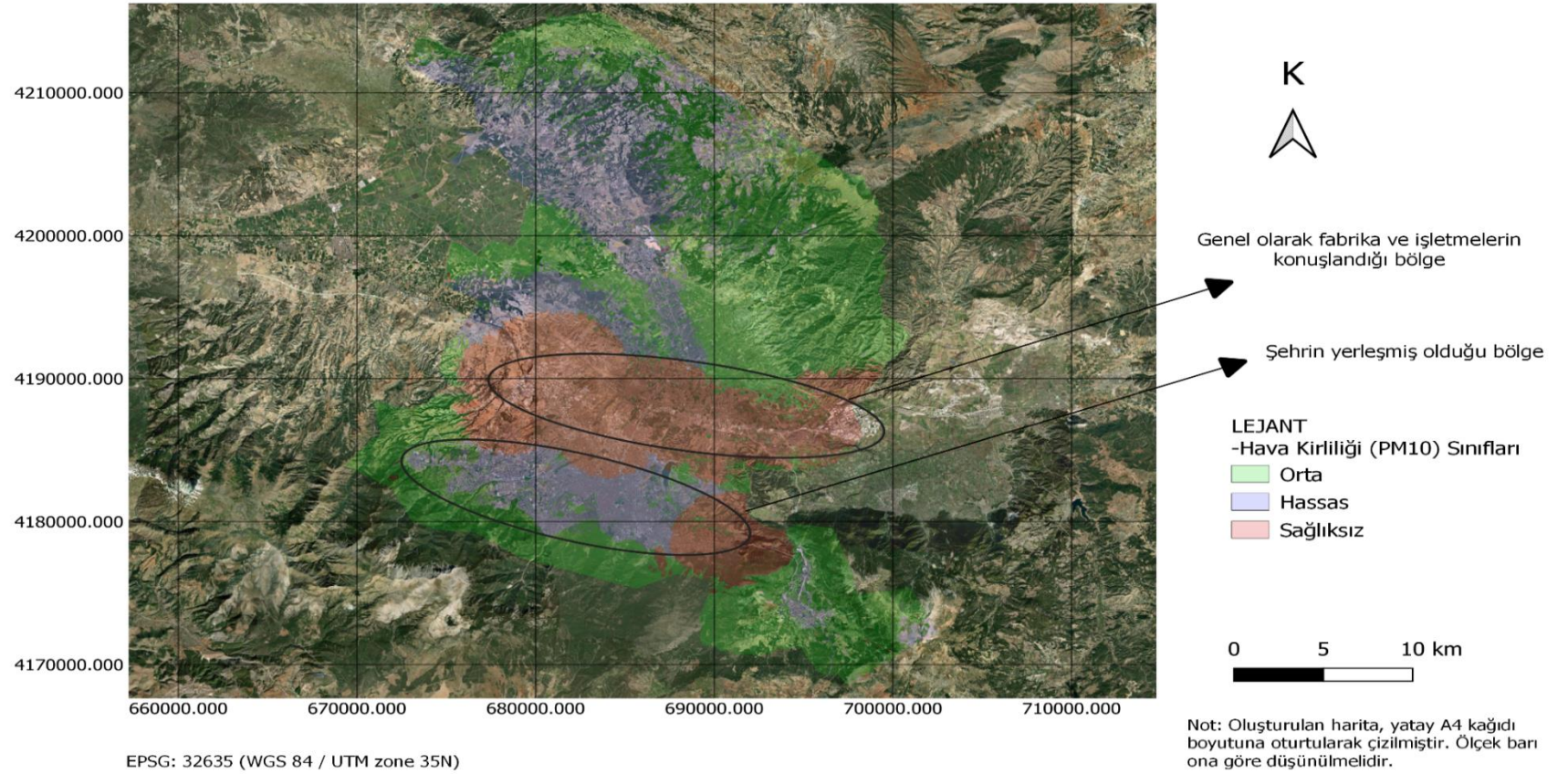
hangi sınıfa ait olduğunu göstermektense, her bir pikselin yüzdelik olarak “orta” sınıfına aitlik derecesini göstermesi açısından önemlidir.



Şekil 4.27. PM10 değeri odaklı tahmin haritası

Yapılan çalışma sonucunda elde edilen sonuç haritası, Şekil 4.27’de görüldüğü gibidir. Sonuç haritası elde edildikten sonra süreç son bulmamaktadır. Sonuç haritasının çeşitli araçlarla analiz edilerek, harita üzerinde, elimizdeki teorik bilgiler doğrultusunda, olmamasına kesin gözüyle bakılan garip sonuçlar var mı, sonuç ürünündeki sınıflandırma istenilen tutarlılığa sahip mi, incelemesinin yapılması gerekmektedir. Yapılan inceleme neticesinde, böyle bir duruma rastlanırsa, bu makinenin verilen verilere göre yanlış bir öğrenme yaptığı, makinenin öğrenmesi için verilen veri setinin, “*Makinenin Öğrenmesi İçin Veri Seti Oluşturması*” bölümünde anlatıldığı gibi gözden geçirilerek, veri seti ile ilgili, makinenin yanlış öğrenmesine mahal vermeyecek şekilde güncelleme yapılması ya da makine öğrenmesi sürecinde kullanılan parametrelerin gözden geçirilmesi gerekmektedir.

HAVA KİRLİLİĞİ TAHMİN HARİTASI



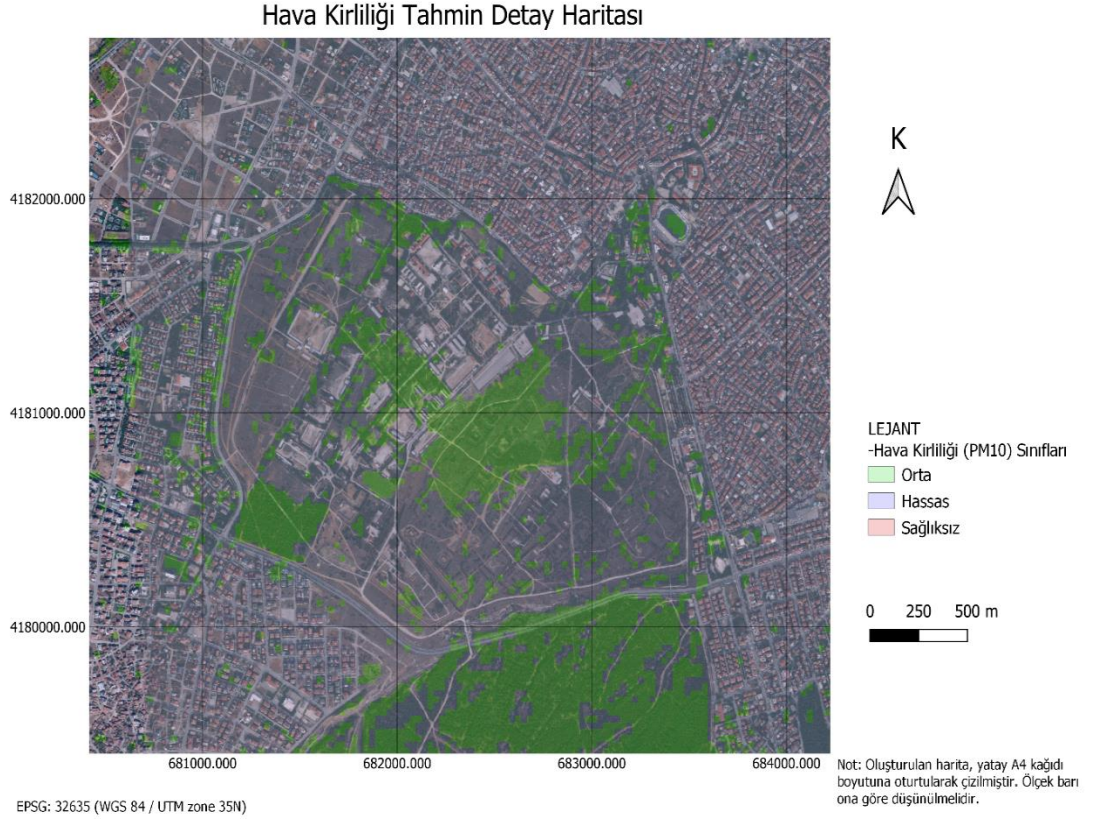
Şekil 4.28. PM10 değeri odaklı tahmin haritası (uydu görüntüsü ile çakıştırılmış şekilde)

Elde edilen sonuç haritasını daha iyi anlayabilmek ve ifade edebilmek adına, Şekil 4.28’de sonuç haritasının uydu görüntüsü ile karşılaştırılmış hali gösterilmiştir. Aynı zamanda, daha iyi fikir verebilmesi adına, söz konusu çalışma alanında bahsedilmesi gerekli bulunan, şehrin yerleşmiş olduğu bölge ile fabrika ve işletmelerin genel olarak konuşlandığı bölge harita üzerinde gösterilmiştir. Bu iki bölge dışında kalan alanlar, çalışma alanındaki kırsal bölgelerdir.

Şekil 4.28’de gösterimi yapılan sonuç haritasında da görüldüğü gibi; çok önemli bir hava kirliliği kaynağı olan fabrika ve işletmelerin bulunduğu bölgenin genel itibariyle “sağlıksız” sınıfında tahmininin yapıldığı, şehrin yerleşmiş olduğu bölgenin ise genel itibariyle “hassas” sınıfı olarak, ufak bazı lokal bölgelerin “orta” sınıf olarak tahmin edildiği, geriye kalan kırsal bölgelerin ise, genel olarak ormanlık ve bitki örtüsüne sahip bölgelerin “orta” sınıf olarak, bitki örtüsüne sahip olmayan kırsal bölgelerin “hassas” sınıf olarak tahmin edildiği görülmektedir.

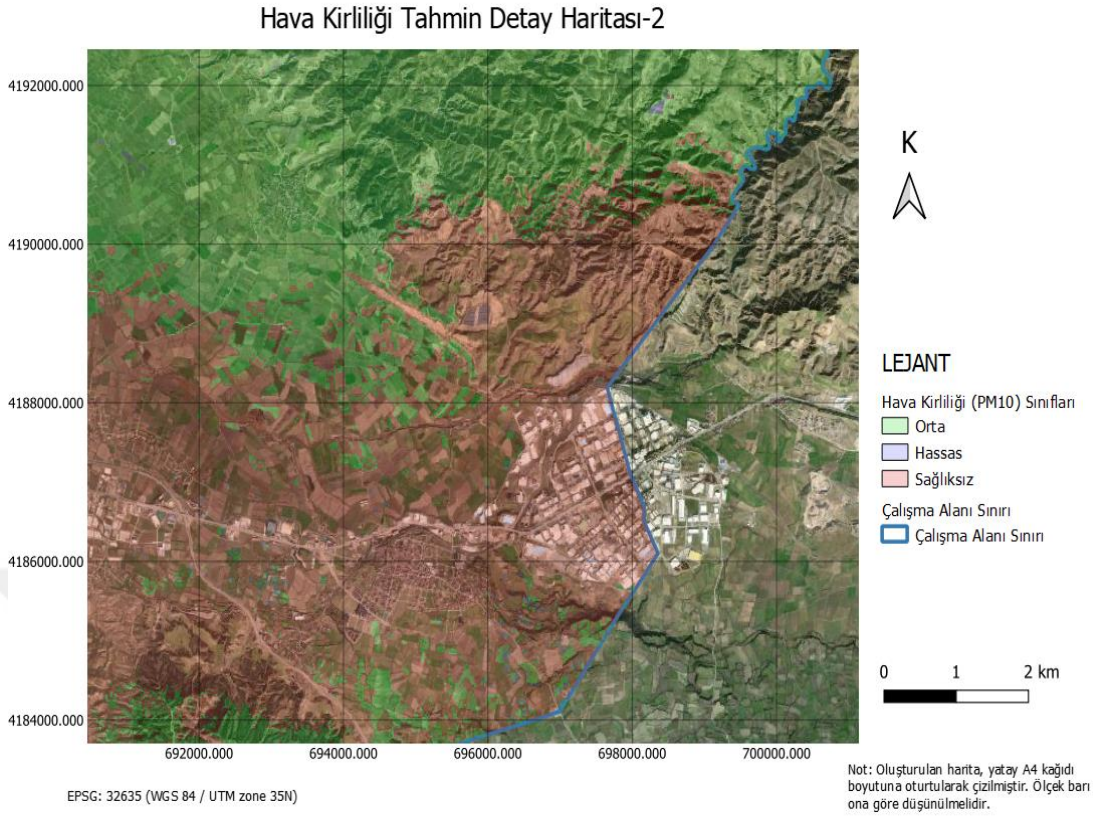
4.6.1. Sonuç haritasının incelenmesi

Bu bölümde, QGIS programı kullanılarak, uydu görüntüsü ile sonuç haritası üst üste çakıştırılmak suretiyle inceleme yapılmıştır. Sonuç haritasının belli bölümleri ele alınarak, yapılan tahmine dair analizler yapılmıştır.



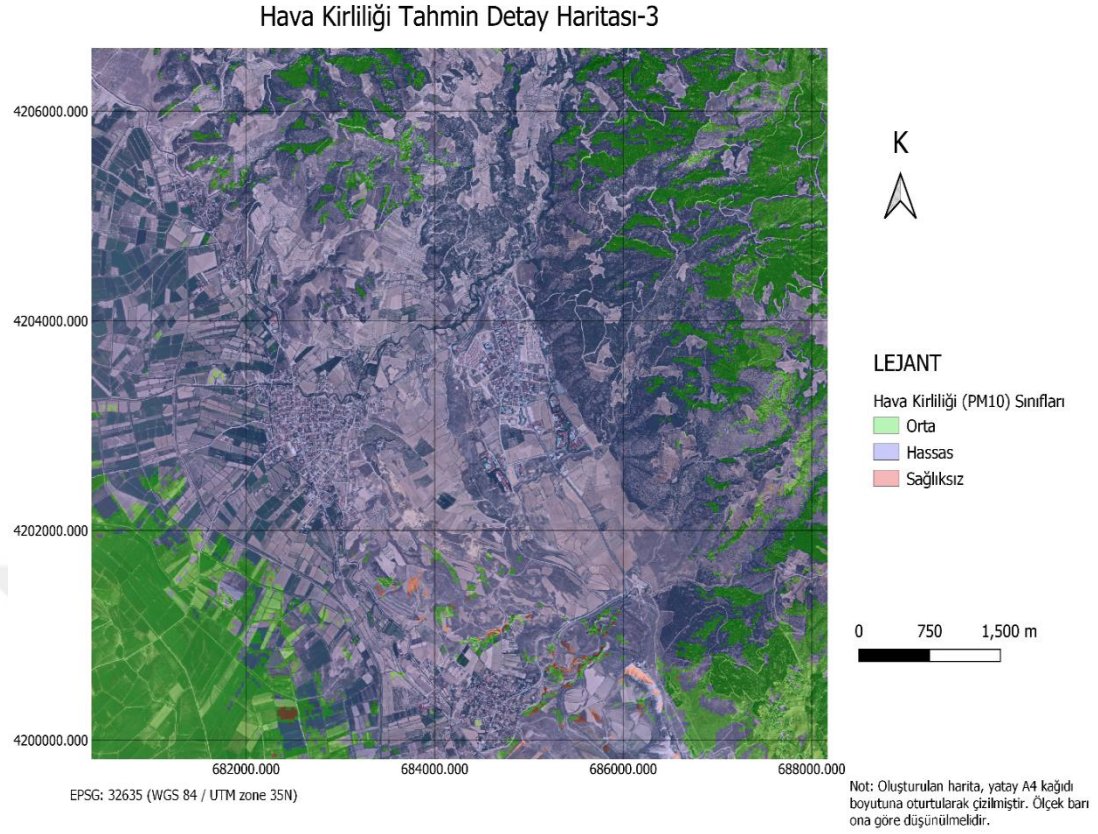
Şekil 4.29. PM10 değeri odaklı tahmin detay haritası-1 (uydu görüntüsü ile çakıştırılmış şekilde)

Şekil 4.29’da çalışma alanında şehrin yerleşmiş olduğu bölgenin bir kısmını gösteren harita görülmektedir. Haritada da görüldüğü gibi şehir bölgesinin bu kısmının genel itibariyle “hassas” olarak tahmin edildiği, bazı bölgelerin “orta” sınıf olarak tahmin edildiği görülmektedir. “Orta” sınıf olarak tahmin edilen bölgelerin haritadaki uydu görüntüsünden de görüldüğü üzere bitki örtüsünün kaplı alanlar olduğu görülmektedir. Buradan hareketle makinenin hava kirliliği için büyük öneme sahip bitki örtüsü faktörünü doğru şekilde öğrenip tahmin yaptığı anlaşılmaktadır.



Şekil 4.30. PM10 değeri odaklı tahmin detay haritası-2 (uydu görüntüsü ile çakıştırılmış şekilde)

Şekil 4.30’da görülen harita, çalışma alanındaki organize sanayi bölgesine (şekildeki haritanın güneydoğusundaki bölge) denk gelen kısmı göstermektedir. Görüldüğü üzere, organize sanayi bölgesinin, bu bölgedeki anayol aksının ve çevrelerinin “sağlıksız” sınıfında tahmin edildiği, bu bölgelerden belli bir mesafe uzaklıktaki kırsal alanların ise, uydu görüntüsünden de görüldüğü gibi, hava kirliliği yayan işletmelerden, bölgedeki ana yol aksından uzaklaşmanın ve bitki örtüsünün de etkisi ile “orta” sınıf olarak tahmin edildiği görülmektedir. Sonuç haritasının bu kesiminden, yapılan tahminde, bağımsız değişken olarak değerlendirilen, fabrikaya yakınlık, yollara uzaklık ve bitki örtüsü ile ilgili diğer değişkenlerin etkilerinin doğru şekilde öğrenilip tahmin yapıldığı görülmektedir.



Şekil 4.31. PM10 değeri odaklı tahmin detay haritası-3 (uydu görüntüsü ile çakıştırılmış şekilde)

Şekil 4.31’de gösterilen alan, köy yerleşkelerinin, ormanlık alanların ve tarlaların bulunduğu, çalışma alanındaki kırsal alanların bir kısmını temsil etmektedir. Şekil 41’de de görüldüğü üzere, kırsal alandaki köy yerleşkeleri ve etrafının “hassas” sınıf olarak, köy yerleşkelerinden uzakta ve bitki örtüsünün bulunduğu alanların “orta” sınıf olarak tahmin edildiği görülmektedir. Bu şekilde tahmin yapılması, yine makinenin, BU(Built-up) indeksi değişkeninin ve bitki örtüsü ile ilgili bağımsız değişkenlerin tahmin edilmek istenen duruma etkisini güzel şekilde öğrenip tahmin yaptığı anlaşılacaktır.

5. SONUÇ ve ÖNERİLER

Bu araştırmada, dijitalleşen dünyanın etkisiyle elde edilen büyük verinin analiz edilmesi ve efektif karar mekanizmalarının oluşturulmasında çok büyük rol oynayan, yapay zekâ terminolojisinden, bu terminolojinin doğru bir hiyerarşi ve sınıflandırma altında yapay zekâ tekniklerinden, bu tekniklerin ne gibi uygulamalarda kullanılabileceği incelenmiştir. Karşımıza çıkan problemlerin hemen hemen hepsinin içerisinde mekân ve konum olgularının yer aldığı, bu problemlerin ne gibi problemler olduğu, bu problemlerin çözümünde ne gibi yapay zekâ tekniklerinin ne şekilde kullanılabileceğinden bahsedilmiştir. Aynı zamanda çalışmada bahsedilen yapay zekâ tekniklerinin, birbiriyle çok sıkı bir ilişki içerisinde olduğu, birçok yapay zekâ tekniğinin gerçekleştirilmesinde diğer birçok yapay zekâ tekniğinin kullanıldığından bahsedilmiştir. Sonrasında ise, yapay zekanın günümüzde belki de en önemli dalı ve uygulama alanı en geniş olan makine öğrenmesi algoritması kullanılarak mekânsal bir probleme dair çözüm ortaya konulup, uygulamanın tüm aşamaları detaylı şekilde açıklanmıştır.

Yapılan uygulamada elde edilen hava kirliliği tahmin haritası baz alınarak, şehirlerde imar planlarının hangi bölgelerde yapılması gerektiği, hangi alanlarda yeşil alanların oluşturulabileceği, hangi alanlarda fabrika kurulmasının daha doğru olabileceği gibi kararlar alınabilir.

Dışarıda çözülmeyi bekleyen yüzlerce mekânsal problem bulunmaktadır. Artık veriye ulaşmanın da zor olmadığı bu çağda bu bir fırsat olarak görülmeli, problemlere bağlı olarak literatürde bulunan yapay zekâ teknikleri kullanılıp, bu tekniklerin doğruluk analizleri yapılması suretiyle problemlere dair optimum çözümler bulunmalıdır.

Dünya üzerinde herhangi bir bölgenin, arazi yapısı ya da demografik özelliklerden dolayı sahip olduğu dezavantajlar yapay zekâ teknikleri kullanılarak minimum seviyeye indirilebilir. Örnek vermek gerekirse; heyelana müsait bir bölgede;

o alana ait uydu görüntüsü indirilip, bu görüntü, eğer istenilen çözünürlüğe sahip değilse derin öğrenme ile çözünürlük artırılabilir. Sonrasında istenilen çözünürlükte elde edilen görüntüdeki yapıları, yine derin öğrenme ile vektör veriyi dönüştürülebilir. Aynı bölgedeki sayısal yükseklik modeli ve bir bölgedeki heyelan fenomenine etki eden diğer faktörler öznitelik verisi olarak kullanılıp, bunları ağırlıklandırıp makine öğrenmesi disiplini olan regresyon işlemine tabi tuttuğumuz zaman, bu bölgedeki heyelan risk haritası ortaya çıkarılabilir. Bu harita ile, uydu görüntüsünden derin öğrenme tekniği ile vektör veriye dönüştürülen yapı haritasının üst üste çakıştırılması suretiyle heyelan riskine sahip yapılar tespit edilebilir. Bu ve bu gibi örnekler çoğaltılabilir.



- VASYLENKO, Y., 2017. Erişim adresi <https://rpubs.com/geka/UrbanWarm>
- ŞEKER, E. S., 2020. Erişim adresi https://www.udemy.com/course/makine-ogrenmesi/learn/lecture/9969304?components=add_to_cart%2Cavailable_coupons%2Cbuy_button%2Cbuy_for_team%2Ccacheable_buy_button%2Ccacheable_deal_badge%2Ccacheable_discount_expiration%2Ccacheable_price_text%2Ccacheable_purchase_text%2Ccurated_for_ufb_notice_context%2Cdeal_badge%2Cdiscount_expiration%2Cgift_this_course%2Cincentives%2Cinstructor_links%2Clifetime_access_context%2Cmoney_back_guarantee%2Cprice_text%2Cpurchase_tabs_context%2Cpurchase%2Crecommendation%2Credeem_coupon%2Csidebar_container%2Cpurchase_body_container#notes
- CENGİZ Y. E. ve ERNST F. B., 2022. Mekânsal Bağlamda Yapay Zekâ Tekniklerine Genel Bir Bakış
- GÜNEY, CANER, 2020. “Harita/Geomatik Mühendisliğinin Geleceği ve Mekansal Zeka”. *YouTube*. Yayın Tarihi 11 Temmuz 2020. https://www.youtube.com/watch?v=lkpGgQZ1GFA&list=PL-5IMFi1nsGippTM1yzvdbJNeAtm_JvqH&index=7