

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**YÜKSEK ÇÖZÜNÜRLÜKLÜ SPOT 7 OPTİK UYDU GÖRÜNTÜLERİNİN
PİKSEL BAZLI SINIFLANDIRILMASINDA MAKİNE ÖĞRENMESİ
YÖNTEMLERİNİN KULLANILMASI VE ANALİZİ**

YÜKSEK LİSANS TEZİ

Çağatay YILDIRIM

Geomatik Mühendisliği Anabilim Dalı

Geomatik Mühendisliği Programı

ŞUBAT 2021

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**YÜKSEK ÇÖZÜNÜRLÜKLÜ SPOT 7 OPTİK UYDU GÖRÜNTÜLERİNİN
PİKSEL BAZLI SINIFLANDIRILMASINDA MAKİNE ÖĞRENMESİ
YÖNTEMLERİNİN KULLANILMASI VE ANALİZİ**

YÜKSEK LİSANS TEZİ

**Çağatay YILDIRIM
(501191616)**

**Geomatik Mühendisliği Anabilim Dalı
Geomatik Mühendisliği Programı**

Tez Danışmanı: Prof. Dr. Hilal Gonca COŞKUN

ŞUBAT 2021

İTÜ, Lisansüstü Eğitim Enstitüsü'nün 501191616 numaralı Yüksek Lisans Öğrencisi Çağatay YILDIRIM, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı “YÜKSEK ÇÖZÜNÜRLÜKLÜ SPOT 7 OPTİK UYDU GÖRÜNTÜLERİNİN PİKSEL BAZLI SINIFLANDIRILMASINDA MAKİNE ÖĞRENMESİ YÖNTEMLERİNİN KULLANILMASI VE ANALİZİ” başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : **Prof. Dr. Hilal Gonca COŞKUN**
İstanbul Teknik Üniversitesi

Jüri Üyeleri : **Prof. Dr. Füsun Balık ŞANLI**
Yıldız Teknik Üniversitesi

Prof. Dr. Hande DEMİREL
İstanbul Teknik Üniversitesi

Teslim Tarihi : 19 Şubat 2021
Savunma Tarihi : 22 Şubat 2021



ÖNSÖZ

Yüksek lisans tez çalışmamda görüş ve önerilerini, sosyal desteğini esirgemeyen, bu çalışmanın ortaya çıkışında ve yürütülmesinde yol gösterici olan değerli danışman hocam Prof.Dr.Hilal Gonca COŞKUN' a,

Yüksek lisans eğitimim ve çalışmam süresince yardımlarını her zaman yanımda hissettiğim, fikir ve görüşleri ile ufkumu zenginleştiren değerli komutanlarım Dr.Öğr.Üyesi Hv.Per.Alb.Orhan GÖZAYDIN'a, Dr.Öğr.Üyesi Hv.Mly.Alb.Mehmet KURUŞÇU'ya ve Dr.Öğr.Üyesi Hv.Müh.Yb.Erhan BERK'e,

İSTANBUL Teknik Üniversitesi Uydu Haberleşme ve Uzaktan Algılama Merkezi (UHUZAM) da görevli Prof.Dr.Elif SERTEL'e ,Yük.Müh.Beril VAROL'a,

Tezimin her aşamasında manevi desteğini esirgemeyen aileme,

Bu zorlu yolda beni sabır ve özveriyle destekleyen ve her zaman yanımda olan sevgili eşim Duygu YILDIRIM' a

En içten duygularımla teşekkürlerimi sunuyorum.

Şubat 2021

Çağatay Yıldırım
Geomatik Mühendisi



İÇİNDEKİLER

Sayfa

ÖNSÖZ
İÇİNDEKİLER	vii
KISALTMALAR	ix
SEMBOLLER	xi
ÇİZELGE LİSTESİ.....	xi
ŞEKİL LİSTESİ.....	xiii
ÖZET	xvii
SUMMARY	xix
1. GİRİŞ	1
1.1. Tezin Amacı	1
1.2. Literatür Araştırması	23
1.3. Tezin Organizasyonu.....	26
2. UZAKTAN ALGILAMA TEKNOLOJİSİ.....	27
2.1. Uzaktan Algılama.....	27
2.2. Elektromanyetik Spektrum.....	30
2.3. Çözünürlük Türleri	33
2.4. Sınıflandırma Yöntemleri.....	35
2.5. Uzaktan Algılamada Veri Elde Etme Yöntemleri.....	37
3. Yöntem ve Materyal.....	39
3.1. Çalışma Alanı ve Uydu Görüntüleri.....	39
3.1.1. Çalışma alanı.....	39
3.1.2. Spot 7 uydusu.....	40
3.1.3. Kullanılan yazılımlar.....	43
3.2. Makine Öğrenmesi Yöntemleri.....	44
3.2.1. Makine Öğrenmesi	44
3.2.2. Destek vektör makineleri yöntemi	46
3.2.2.1. Lineer destek vektör verileri	47
3.2.2.2. Lineer olmayan destek vektör verileri.....	47
3.2.3. Karar Ağaçları Yöntemi.....	48
3.2.4. Rastgele Orman Yöntemi.....	50
3.2.5. XGBOOST Hızlandırma Yöntemi.....	52
3.2.6. K-En Yakın Komşuluk Yöntemi.....	53
3.2.7. Naive Bayes Yöntemi	55
4. Görüntülerin İşlenmesi Ve Sınıflandırılması	56
4.1. Eğitim ve Test Verilerinin Oluşturulması	56
4.2. Sınıflandırma İşlemleri.....	58
4.2.1. Destek Vektör Makineleri İle Sınıflandırma işlemlerinin Analizi.....	58
4.2.2. Karar Ağaçları Yöntemi İle Sınıflandırma işlemlerinin Analizi.....	61
4.2.3. Rastgele Orman Yöntemi İle Sınıflandırma işlemlerinin Analizi.....	63
4.2.4. Xgboost Yöntemi İle Sınıflandırma İşlemlerinin Analizi.....	66
4.2.5. K En Yakın Komşuluk Yöntemi İle Sınıflandırma İşlemlerinin Analizi.	69

4.2.6 Naïve Bayess Yöntemi İle Sınıflandırma İşlemlerinin Analizi.....	72
4.2.7 Adaboost Yöntemi İle Sınıflandırma İşlemlerinin Analizi	75
5. SONUÇ VE ÖNERİLER.....	78
KAYNAKLAR.....	85



KISALTMALAR

- Adaboost** : Adaboost Hızlandırma Algoritması
- DVM** : Destek Vektör Makineleri Algoritması
- KNN** : K-En Yakın Komşuluk Analizi Algoritması
- XGBoost** : XGBoost Hızlandırma Algoritması
- RO** :Rastgele Orman Algoritması
- Naïve Bayes:** Naïve Bayes Algoritması
- NIR** :Yakın Kızılötesi spektral Bant
- MS** : Multispektral
- NDVI** : Normalized Difference Vegetation Index
- PAN** : Pankromatik
- PCA** : Principal Component Analysis (Temel Bileşenler Analizi)
- SAM** : Sayısal Arazi Modeli
- SPOT** : Satellite Pour l'Observation de la Terre
- RGB** :Kırmızı ,Yeşil , Mavi Bantlar



SEMBOLLER

λ : Dalga boyu

c : Işık hızı

E : Enerji

h : Plank sabiti ($6,6260 \times 10^{-34}$ jul/saniye)





ÇİZELGE LİSTESİ

Sayfa

Çizelge 2.1 : Spot-7 uydusu teknik özellikleri.	41
Çizelge 2.2 : Spot-7 Uydusu Bant Özellikleri.	42
Çizelge 3.1 : Destek Vektör Makineleri sınıflandırıcısının test verilerinin sınıflara göre gösterimi.	59
Çizelge 3.2 : Destek Vektör Makineleri Sınıflandırıcısının yakın kızıl ötesi bant dahil edilmiş durum için sınıf bazında doğruluk oranları.	60
Çizelge 3.3 : Karar Ağaçları Sınıflandırıcısının test verilerinin sınıflara göre gösterimi.	64
Çizelge 3.4: Rastgele Orman Sınıflandırıcısının test verilerinin sınıflara göre gösterimi.	65
Çizelge 3.5 : Rastgele Orman Sınıflandırıcısının yakın kızıl ötesi bant dahil edilmiş durum için sınıf bazında doğruluk oranları.	67
Çizelge 3.6 : XGBoost Sınıflandırıcısının test verilerinin sınıflara göre gösterimi. .	68
Çizelge 3.7 : XGBoost Sınıflandırıcısının yakın kızıl ötesi bant dahil edilmiş durum için sınıf bazında doğruluk oranları.	70
Çizelge 3.8 : K-En Yakın Komşuluk Sınıflandırıcısının test verilerinin sınıflara göre gösterimi.	71
Çizelge 3.9 : K-En Yakın Komşuluk Sınıflandırıcısının yakın kızıl ötesi bant dahil edilmiş durum için sınıf bazında doğruluk.	73
Çizelge 3.10 : Naive Bayes Komşuluk Sınıflandırıcısının test verilerinin sınıflara göre gösterimi.	74
Çizelge 3.11 : Naive Bayes Komşuluk yakın kızıl ötesi bant dahil edilmiş durum için sınıf bazında doğruluk.	77
Çizelge 3.12 : Adaboost Sınıflandırıcısının yakın kızıl ötesi bant dahil edilmiş durum için sınıf bazında doğruluk oranları	78
Çizelge 4.1 : Kullanılan tüm algoritmaların sonuç değerleri tablosu	80
Çizelge 4.2 : Kullanılan tüm algoritmaların sınıflara göre sonuçları.	82



ŞEKİL LİSTESİ

Sayfa

Şekil 2.1 : Airbus şirketine ait Spot ve Pleidas Uyduları	28
Şekil 2.2 : Uzaktan algılamanın bileşenleri.	29
Şekil 2.3 : Elektromantetik spektrum.	31
Şekil 2.4 : Mekânsal çözünürlük.	33
Şekil 2.5 : 8,4 ve 2 bitlik görüntü	34
Şekil 2.7 : Uzaktan algılamada piksellerin makine öğrenmesi ile sınıflandırılmasının gösterimi.	34
Şekil 2.8 : Aktif sensörlere örnek LİDAR verisinin görseli.	37
Şekil 4.9 : Pushbroom ve Whisbroom tarayıcı.	38
Şekil 3.1 : Serdivan ilçesine ait 4 bantlı 1.5 mt çözünürlüklü Spot-7 uydu görüntüsü.	39
Şekil 3.2 : Çalışma bölgesinin NIR (Yakın Kızılötesi) bantta görüntüsü.....	43
Şekil 3.3 : Makine öğrenmesi ,derin öğrenme yapay zekanın birbiri ile ilişkisi.	44
Şekil 3.5 : Lineer destek vektörü ile ayrılmış veriler.....	47
Şekil 3.6 : Linner ayrılamayan vektörlerin gösterimi.....	48
Şekil 3.7 : Karar ağacı algoritmasını temsil eden görsel.	49
Şekil 3.8 : Rastgele Orman Algoritması.	50
Şekil 3.9 : Xgboost hızlandırma yöntemi zayıf ve etkin algoritmaların temsili.....	52
Şekil 3.10 : K-En Yakın Komşuluk Yöntemi temsili.....	54
Şekil 3.11 : Naive Bayes yönteminin raslantısal verilerin Bates ağı ile gösterimi....	55
Şekil 4.1 : Eğitim ve test verileri toplanırken alınmış ekran görüntüsü.....	56
Şekil 4.2 : R programında makine öğrenmesi yöntemleri çalıştırılırken alınmış ekran görüntüsü.	57
Şekil 4.3 : Destek Vektör Makineleri Sınıflandırma Sonuç Görüntüsü.	58
Şekil 4.4 : Karar Ağaçları İle Yöntemi İle Sınıflandırma sonuç görüntüsü.	61
Şekil 4.5 : Rastgele Orman Yöntemi İle sınıflandırma sonuç görüntüsü.	63
Şekil 4.6 : Xgboost yöntemi ile sınıflandırma sonuç görüntüsü.....	66
Şekil 4.7 : K-En Yakın Komşuluk yöntemi ile sınıflandırma sonuç görüntüsü.	69
Şekil 4.8 : Naive Bayes algoritması ile sınıflandırma sonuç görüntüsü.....	72
Şekil 4.9 : Adaboost yöntemi ile sınıflandırma sonuç görüntüsü.....	75



YÜKSEK ÇÖZÜNÜRLÜKLÜ SPOT 7 OPTİK UYDU GÖRÜNTÜLERİNİN PİKSEL BAZLI SINIFLANDIRILMASINDA MAKİNE ÖĞRENMESİ YÖNTEMLERİNİN KULLANILMASI VE ANALİZİ

ÖZET

Arazi kullanımının belirlenmesi şehirsal, tarımsal ve ekonomik planlamaların yapılmasında yerel ve ulusal düzeyde ışık tutan önemli bir çalışma alanıdır. Arazi kullanım haritalarının belirlenmesinde kullanılan öncelikli veriler ise uzaktan algılama sistemleri ile elde edilmiş olan verilerdir. Uzaktan algılamada kullanılan verilerin yüksek çözünürlükte olmasına ek olarak bant çeşitliliğinin ve sınıflandırma algoritmasının seçiminin de arazi kullanımının belirlenmesinde önemli rol oynadığı bilinmektedir. Literatürde son zamanlarda kullanılan algoritmaların arazi sınıflarını ayırmadaki gücü, eklenecek olan bantların sınıflandırma etki yüzdesi, algoritmaların çalışma hızları, seçilecek olan verinin niteliği sınıflandırma doğruluğunu etkileyen faktörlerdendir. Seçilecek olan sınıflandırma algoritmalarının avantaj ve dezavantajları bilmek ilerleyen çalışmalarda yol gösterici olacaktır. Çalışmada Spot 7 uydu görüntüleri kullanılarak Sakarya Serdivan ilçesinin arazi kullanımı için 7 sınıf belirlenmiştir. Belirlenen sınıflara ait 150 adet test verisi ve 350 adet eğitim verisi oluşturulmuştur. Oluşturulan aynı eğitim ve test verileri her algoritma için koşulmuştur. Makine öğrenmesi yöntemlerinin çıkarttığı modellere göre her piksel radyometrik değerine göre bir sınıfa atanmıştır. Bu atama sonucunda ortaya çıkan veriler tekrar görsel hale getirilmiş ve ortaya çıkan sonuç görüntüleri arazi sınıflarını göstermiştir. Bu sınıfların ayrılmasında literatürde en çok bilinen Destek vektör makineleri, Xgboost, Adaboost, K en yakın komşuluk analizi, Karar Ağaçları, Naive Bayes ve Rastgele Orman sınıflandırıcıları kullanılmıştır. Ortaya çıkan hata matrisleri, çalışmada kullanılan kırmızı, yeşil, mavi ve yakın kızıl ötesi bantları ile oluşturulan modeller ve sonuç görüntüleri incelenmiş her algoritma için analiz edilmiştir. Elde edilen sonuçların ardından yakın kızılötesi bandın sınıflandırmaya etkisini araştırmak amacı ile yakın kızıl ötesi bant çıkarılıp aynı veriler ile kırmızı, yeşil ve mavi bantlarla

aynı 7 algoritma kullanılarak sınıflandırma işlemi tekrarlanmıştır. Yakın kızıl ötesi bandın da modelleri önemli ölçüde etkilediği gözlenmiştir. Yakın kızıl ötesi bandın sınıflandırmaya etkileri her algoritma için ayrı ayrı değerlendirilmiş ve hata matrisleri yorumlanmıştır. Şehir alanlarının ayrılmasında tüma algoritmada önemli bir artış yakın kızıl ötesi bandı eklendikten sonra gözlenmiştir. Özellikle yakın kızıl ötesi bandın kullandığı sınıflandırmada yeşil alanların ayrılmasında RBG ye göre büyük avantaj sağladığı tespit edilmiştir. Son olarak algoritmaların işlem hızı, avantaj, dejavantajları ve hata matrislerine göre sınıfları ayırma yüzdeleri irdelenmiştir.



USING AND ANALYSIS OF MACHINE LEARNING METHODS IN PIXEL-BASED CLASSIFICATION OF HIGH RESOLUTION SPOT 7 OPTICAL SATELLITE IMAGES

SUMMARY

Determination of land use is an important field of study that sheds light on local and national level in urban, agricultural and economic planning. The primary data used in determining the land use maps are the data obtained by remote sensing systems. In addition to the high resolution of the data used in remote sensing, it is known that the band diversity and selection of the classification algorithm play an important role in determining the land use. The power of the algorithms used recently in the literature in separating the land classes, the classification effect percentage of the bands to be added, the working speed of the algorithms, the quality of the data to be selected are among the factors that affect the classification accuracy. Knowing the advantages and disadvantages of the classification algorithms to be selected will guide in future studies. In the study, 7 classes for land use of Sakarya Serdivan district were determined by using Spot 7 satellite images. 150 test data and 350 training data belonging to the determined classes were created. The same training and test data generated were run for each algorithm. According to the models created by machine learning methods, each pixel is assigned to a class according to its radiometric value. The data obtained as a result of this assignment were made visual again and the resulting images showed the land classes. Support vector machines, Xgboost, Adaboost, K closest neighborhood analysis, Decision Trees, Naive Bayes, and Random Forest classifiers were used in the separation of these classes. The resulting error matrices, models created with the red, green, blue and near infrared bands used in the study, and the result images were analyzed for each algorithm. After the results, the near infrared band was removed in order to investigate the effect of the near infrared band on the classification and the classification process was repeated using the same data and the same 7 algorithms with red, green, and blue bands. It has been observed that the near infrared band also significantly affects the models. The effects

of the near infrared band on the classification were evaluated separately for each algorithm and error matrices were evaluated. A significant increase in the entire algorithm in separating city areas was observed after the near infrared band was added. Especially in the classification used by the near infrared band, it has been determined that it provides a great advantage over RGB in separating green areas. Finally, the processing speed, advantages, disadvantages and percentages of classification according to error matrices of the algorithms are examined.

The use of 1.5 meter high resolution satellite imagery in the classification and literature has significantly influenced the classification accuracy. The situations with and without near infrared band were evaluated separately. The fact that city areas are the most incorrectly evaluated class has affected the overall accuracy for all algorithms. Support Vector Machines algorithm; In addition to being one of the oldest methods, it did not show a high error rate in almost any class and also gave an overall consistent classification result. Decision Trees algorithm; Compared to the support vector machines algorithm in the classification analysis, the city areas were classified quite incorrectly. The Random Forest classification algorithm generally determined all classes consistently. XGboost algorithm; in general it has consistently separated all classes. The XGboost acceleration algorithm has come to the forefront in the classification of urban area pixels, especially thanks to the merging process that improves the overall performance in areas where the decision trees algorithm is weak. K Nearest Neighborhood algorithm; Looking at the error matrix, it was found that it classifies the pixels homogeneously by separating them into clusters. Naive Bayes algorithm; It gave a bad result in city area pixels compared to other algorithms. Adaboost acceleration algorithm; As in the XGboost algorithm, it has achieved more successful results by applying the merging process where the decision tree algorithm is weak.



1. GİRİŞ

1.1 Tezin Amacı

Arazi kullanımının belirlenmesi, imar planlarının oluşturulmasında ve kalkınmaya yönelik planların hayata geçirilmesinde oldukça önem arz eden bir konudur. Arazi kullanımı haritalarının oluşturulması yerel yönetimlere ekonomik faaliyetlerin etkin planlanmasında, endüstriyel gelişimin izlenmesi ve kontrol edilmesi, doğal kaynakların korunması ve geliştirilmesinde önemli avantajlar sağlamaktadır.

Arazi kullanım haritalarının oluşturulması uzaktan algılama verilerinin etkin biçimde kullanımı ve görüntülerin sınıflandırılmasıyla mümkündür. Dijital görüntü sınıflandırması, bir arazi örtüsü sınıfına uzaktan algılama görüntüsünün bir pikselini atama işlemidir. Amaç, her pikseli yalnızca bir sınıfa yerleştirme veya pikseli birçok sınıfla ilişkilendirmektir. Sınıflandırma teknikleri, eğitim sürecine göre veya teorik model oluşturması temelinde kategorize edilebilir. Bu ayrım altında çeşitli sınıflandırma algoritmaları geliştirilmiştir. Örneğin, Maksimum Olabilirlik Sınıflandırıcısı (MLC) denetimli bir parametrik algoritmadır, k-ortalama kümeleme ise denetimsiz bir parametrik algoritmadır [17].

İlerleyen teknolojinin paralelinde uzaktan algıma Sınıflandırma için seçilecek olan yöntem oldukça önemlidir. Sınıflandırma algoritmaları incelendiğinde literatürde birçok yöntemin geliştirildiği ve geliştirilmeye devam ettiği görülmektedir. Bu algoritmaların en başında Destek Vektör makineleri gelmektedir. Bir Destek Vektör Makinesi (SVM), iki sınıf arasındaki aralığı maksimum olan hiper düzlemi bularak sınıflandırma gerçekleştirir. Hiper düzlemi tanımlayan vektörler destek vektörleridir[18].

Karar ağaçları, çok kullanılan kontrollü sınıflandırma algoritmaları kümesidir. Sınıflandırma problemlerinde oldukça iyi performans gösterirler, karar yolunun yorumlanması nispeten kolaydır ve bunları eğitmek için kullanılan algoritma hızlı ve basittir. Karar ağaçları algoritmasının daha yoğun veri sınıfı için kullanılan gelişkin hali rastgele orman algoritmasıdır. Rastgele orman, temelde N sayıda karar ağacının bir birleşimini temsil eder, böylece tahminlerin sağlamlığını artırır[19].

K en yakın komşu diğer adıyla KNN Algoritması, eğitim aşamasında tüm veri setini kullanan basit bir algoritmadır. Görülmeyen bir veri örneği için bir tahmin gerektiğinde, en benzer örnekler için tüm eğitim veri kümesinde arama yapar ve en benzer örneğe sahip veriler sonunda tahmin olarak döndürülür. kNN, benzer öğeleri aradığımız arama motoru uygulamalarında sıklıkla kullanılır[20].

XGBOOST diğer bir adı ile gradyan artırma sınıflandırıcıları, güçlü bir tahmin modeli oluşturmak için birçok zayıf öğrenme modelinin bir araya getirilmesi ile oluşturulmuş bir makine öğrenimi yöntemidir. Karar ağaçları genellikle XGboost yönteminde ön adım olarak uygulanır. XGBOOST algoritması, karmaşık veri kümelerini sınıflandırmadaki başarısı nedeniyle popüler hale gelmiştir ve son zamanlarda uzaktan algılamada ve birçok farklı alanda kullanımına devam edilmektedir[21].

Adaboost, kelime anlamı olarak uyarlanabilir yükseltme anlamına gelmektedir. Bu, her adımda zayıf modelleri yani başta kullanılan referans modeli önceki modellerimizin hatalarına dayanarak güçlendirmeye çalışmak anlamına gelir, böylece daha tutarlı ve eksiksiz bir karma model olurlar. Her modelde problemin iyi uygulandığı bir parçası vardır ama eksiklikleride vardır. Bu yüzden modellerin zayıf kısımları egale edilirken etkin kısımları kullanılarak daha doğru sonuçlar elde edilir[22].

Naive Bayes, Bayes Teoremine dayalı istatistiksel bir sınıflandırma tekniğidir. En basit denetimli öğrenme algoritmalarından biridir. Naive Bayes sınıflandırıcı, hızlı, doğru ve güvenilir bir algoritmadır. Naive Bayes sınıflandırıcıları, büyük veri kümelerinde yüksek doğruluk ve hıza sahiptir[23].

1.2 Literatür Araştırması

Günümüze kadar ki literatür çalışmaları incelendiğinde son yıllarda yapılan arazi örtüsü belirlemek için bir çok sınıflandırma algoritması geliştirilmiştir. Literatürde uzaktan algılama konusu altında çok farklı disiplinlere yönelik çalışmalar yapılmıştır ve yapılmaya devam etmektedir.

İncelenen çalışmalarda birbirinden farklı sınıflandırma yöntemleri ve algoritmaları olduğu görülmüştür. Bu yöntemler denetimli sınıflandırma ve denetimsiz sınıflandırma olarak ayrılabilirdiği gibi aynı zamanda piksel tabanlı ve son zamanlarda üzerinde birçok yeni çalışma yapılan obje tabanlı sınıflandırma yöntemi mevcuttur.

“Performance of different machine learning algorithms on satellite image classification in rural and urban setup” isimli çalışmada kırsal ve kentsel alanlarda meydana gelen farklılıkların daha iyi anlaşılması için makine öğrenmesi algoritmalarının arazi kullanımının belirlenmesi konusundaki rolünü göstermektedir. Sınıflandırmada daha iyi performans sağlayan uygun algoritmaları bulmak için benzer fenolojik aşamalara sahip yüksek çözünürlüklü Landsat-8, Sentinel-2 ve Planet görüntülerini kullanmışlardır. Rastgele orman, Destek Vektör Makinesi (DVM), görüntülerin bireysel ve genel sınıf doğruluğunu değerlendirmek için Landsat-8, Sentinel-2 ve Planet görüntülerine ayrı ayrı uygulanmıştır. 6 görüntüyü sınıflandırmak için iki eğitim verisi seti oluşturmuşlardır. Sentinel-2 görüntüsünün üç görüntü arasında en iyi performansı gösterdiği ve üç farklı algoritma arasında, DVM nispeten daha iyi sonuçlar gösterdiği sonucuna varmışlardır. Ortaya konan sonuçlar planlamacılara, uzaktan algılama bilim adamlarına ve karar vericilerin uygun bir görüntü sınıflandırma algoritması seçmelerine yardımcı olmayı amaçlamıştır [24].

“Vehicle Detection in Satellite Images by Hybrid Deep Convolutional Neural Networks” adındaki çalışmada aktif tarımsal alanları tarımsal vasfını kaybetmiş alanlardan ayırmak için uzaktan algılama uydularından elde edilen düşük bütçeli ve sürdürülebilir bir otonom sistem geliştirmişlerdir. Bu tür tarımsal haritalama sistemi, mahsulün mevsimsel doğası, spektral kanalların karmaşıklığı , bulut ve güneş ışığı gibi

olumsuz koşullar göz önünde bulundurulduğunda uydu görüntü işleme için benzersiz zorluklar ortaya çıkarmışlardır [25].

“Deep Learning Based Feature Selection for Remote Sensing Scene Classification” isimli çalışmalarında Yüksek çözünürlüklü uydu görüntülerinin kullanımıyla, uzaktan algılama yöntemleriyle arazi sınıflandırması üzerine çalışılmıştır. Çalışmalarında son zamanlarda çok kullanılan derin öğrenme tekniği olan yapay sinir ağları algoritmasını kullanmışlardır. 2800 uzaktan algılama görüntüsü kullanılarak yapılandırma hatasını en aza indirerek özellik soyutlamasını başardığını ve daha küçük yapılandırma hatalarına sahip özelliklerin, görüntü temsili için daha fazla öz nitelik barındıracağı sonucuna varmışlardır [26].

“Automatic target detection in satellite images using deep learning ”çalışmalarında isimli çalışmalarında Uydu görüntülerinden yararlanarak otomatik hedef tespiti, hedef nesnenin değişen boyutunu belirlemeyi amaçlamışlardır. Görüntülerin hareketli olması yönü ve arka planı nedeniyle bir sorun olduğundan literatürde bilinen HOG, Gabor özelliği ve Hough dönüşümü gibi geleneksel olarak tasarlanmış özelliklerin, yüksek çözünürlüklü devasa veriler için iyi çalışmayacağını belirtmişlerdir. Bu yüzden uydu görüntülerinden derin öğrenme yöntemini uygulayabilen sağlam ve hesaplama açısından verimli sistemler gerekliliği belirtilmiştir. Çalışmalarında bir sahnedeki hedef ve hedef olmayan nesnelere sınıflandırmak için EdgeBoxes ve Convolutional Neural Network'ü (CNN) kullanan uydu görüntüleri için bir hedef tespit sistemi önermişlerdir. Uydu görüntülerinde hedeflerin uç bilgileri çok belirgin ve öz nitelikler içerdiği ve EdgeBoxes adlı algoritmanın, hedef teklif kümesini filtrelemek için uç bilgilerini kullandığı sonucuna varmışlardır [27].

“Vehicle detection in satellite images by hybrid deep convolutional neural networks” isimli çalışmalarında Uydu görüntülerinde araçlar gibi küçük nesnelere tespitini hedeflemişlerdir. Nesne algılama performansını artırmak için birçok geleneksel yöntem kullanmışlardır. Fakat derin evrişimli sinir ağları eğitim verilerinden zengin özellikleri otomatik olarak öğrenebildiği ve birçok görüntü sınıflandırma veritabanında son teknoloji performans elde etmişlerdir. Derin evrişimli sinir ağları distorsiyona karşı sağlamlık göstermiş olsa da, yalnızca aynı ölçekteki özellikleri çıkarır ve bu nedenle nesnenin büyük ölçekli varyansını tolere etmek için yetersiz olduğu sonucuna varılmıştır. Karşılaştırmalı deneysel sonuçlarda , önerdikleri Derin

evrişimli sinir ağıları araç tespitinde geleneksel yöntemlerden önemli ölçüde daha iyi performans ortaya koyduğunu göstermektedir [28].

“Comparative research on deep learning approaches for airplane detection from very high-resolution satellite images” isimli çalışmalarında vardıkları sonuçlara göre doğruluk değerlendirme sonuçları, Faster R-CNN mimarisinin F1 puanlarına, ortalama hassasiyet (AP) ölçümlerine ve sonuçların görsel incelemesine göre en yüksek doğruluğu sağladığını belirtmişlerdir. YOLO-v3, biraz daha düşük performansla ikinci sırada yer aldığını , ancak doğruluk ve hız arasında dengeli bir denge sağladığını . SSD, algoritmasının ise en düşük algılama performansını sağladığını, ancak nesne yerleştirmede daha iyi olduğunu belirtmişlerdir. Vardıkları sonuçlar aynı zamanda nesne boyutu ve algılama doğruluğu açısından da değerlendirilmiş ve bu da büyük ve orta ölçekli uçakların daha yüksek doğrulukla tespit edildiğini kanıtladığını göstermiştir [29].

1.3 Tezin Organizasyonu

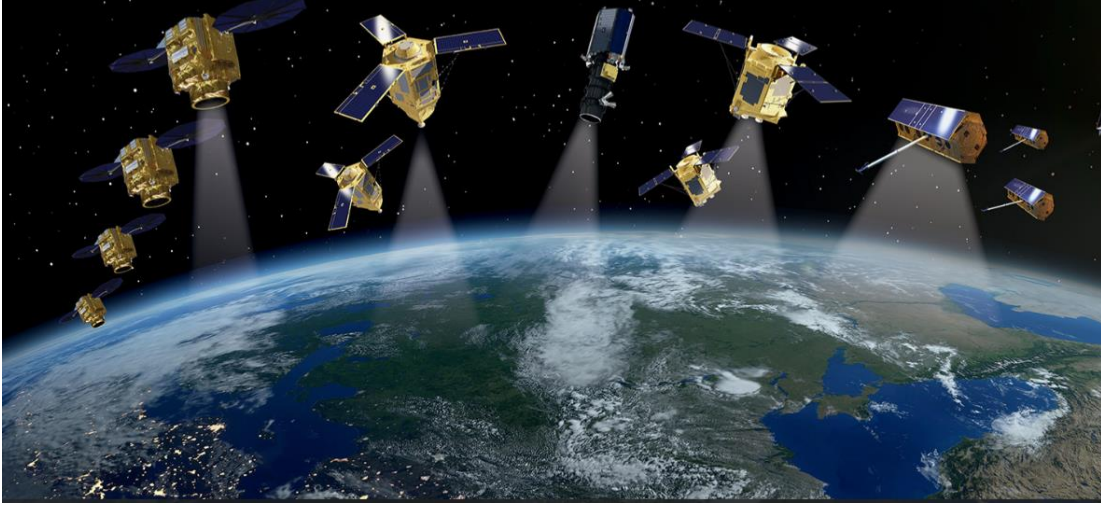
Çalışma 4 ana kısma ayrılmıştır. Çalışmanın ilk kısmında literatür çalışmaları ve çalışmanın amacından bahsedilmiştir. Çalışmanın ikinci kısmında uzaktan algılama teknolojisi ve sınıflandırma yöntemlerinden bahsedilmiştir. Çalışmanın üçüncü kısmında çalışma alanı ve uydu görüntüleri, makine öğrenmesi yöntemlerinin çalışma prensipleri, görüntülerin işlenmesi ve sınıflandırılması yöntemlerinden bahsedilmiştir. Çalışmanın son kısmında sınıflandırma sonuçlarının analizi ve sonuç görüntülerinin yorumlanmasına yer verilmiştir.



2. UZAKTAN ALGILAMA TEKNOLOJİSİ

2.1 Uzaktan Algılama

Uzaktan algılama, nesnelere fiziksel temas kurmadan bilgi toplamak, bu bilgileri analiz etmek ve sonuçlara varmaktır. Uzaktan algılama biliminin analog kameranın icadıyla başladığı düşünülmektedir. Yıllar geçtikçe, icat edilen yeni sensörler neredeyse tüm elektromanyetik spektrumu kapsayacak hale gelmiştir. LANDSAT uydularının fırlatılması, uzaktan algılama çalışmalarına hız katmış ve ülkeler soğuk savaşın etkisiyle uyduları birbiri ardına göndermişlerdir. Örneğin Fransız spot uyduları ve Rusya'nın Sputnik uyduları en bilinen uzay programlarıdır. Uzaktan algılama ile sadece yer üstü değil, aynı zamanda yer altındaki objeler veya uzay hakkında bilgi toplanabilir. Nesnelere ilgili bilgiler amaca yönelik çeşitli sensörler tarafından toplanır. Bilgi elde etmek için kullanılan sensörler, eğer kendileri enerji üretebiliyorlarsa aktif sensörler denir. Aktif sistemlere örnek olarak lidar, radar, sonar gibi uzaktan algılama sistemleri verilebilir. Aktif sensörler güneş ışığına ihtiyaç duymadan çalışabilir, böylece günün hangi saatinde olursa olsun kullanılabilirler[30], [31]. Tarihte ilk uydunun gönderilmesinin amacı meteorolojik gözlemler yapmaktır. Fakat sensör teknolojilerinin ve çeşitliliğinin artması ile uzaktan algılama teknolojisinin kullanım alanı artmıştır. Aynı zamanda özel sensörlere ve amaca yönelik çözünürlüğe göre tasarlanmıştır. Örneğin, varyasyon analizi için bir proje yürütülüyorsa, zamansal çözünürlük ön plandadır. Veya bir bitki ürün çeşidi araştırılırsa, yüksek spektral çözünürlüğe sahip bir sensör tercih edilir. Benzer şekilde, haritacılık amacıyla yerel çözümüme ön planda olmuştur. Spot uydularda yakın kızıl ötesi bant kullanılabilirken, görünür bölgede yüksek hassasiyetle çalışan bir HRV sensörü bulunmaktadır.



Şekil 2.1 : Airbus Şirketine ait Spot ve Pleiadas uyduları [2].

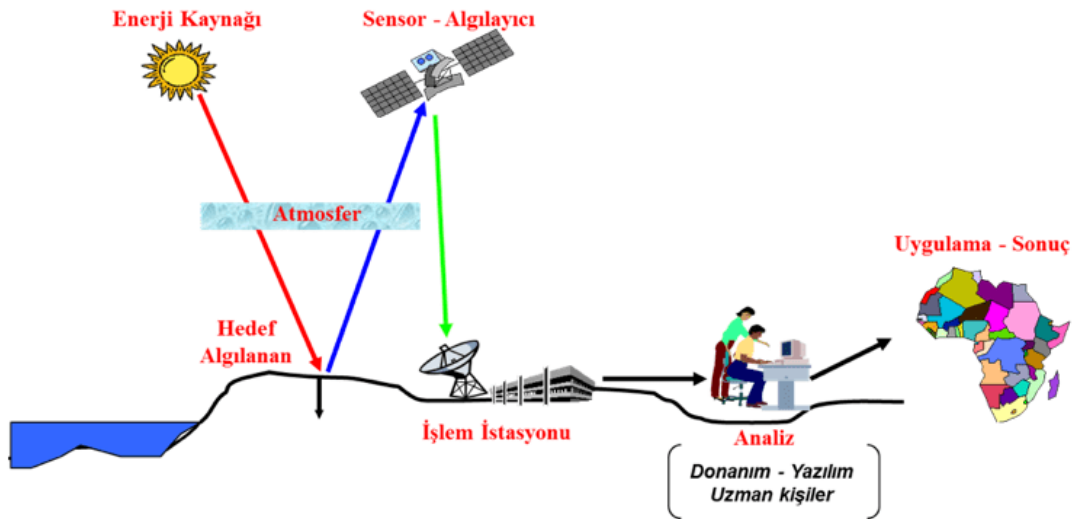
Benzer şekilde Radarsat uydusunda bulunan vegetasyon sensörü sayesinde bitkisel endeks çalışmaları için bir özelleşmiş başka örnektir. Günümüzde bir kısmı keşif ve gözlem, tarımsal alanların izlenmesi, kara ve deniz gözlemleri, iklim takibi, küresel ısınma gibi birçok alanda uzaktan algılama kullanılmaktadır. atmosferik gözlemler, bulutlar, hava tahminleri, doğal afetlerin izlenmesi ve daha birçok konuda uzaktan algılama ilk başvurulan seçenektir[32].

Uzaktan algılamada hedef, bilgilerin maliyet etkin ve zamanında tedarigi aynı zamanda bilgilerin doğruluğu çalışmanın başarısını oluşturan faktörlerdir. Bu yüzden birçok çalışma için uzaktan algılamanın yeryüzü hakkında tutarlı ve düşük maliyetli veri sağlaması önemli bir avantajdır. Ayrıca uzaktan algılama değişim analizi ve sürekli çalışmalar açısından kullanıldığında stratejik anlamda karar verme süreçlerini önemli ölçüde olumlu yönde etkilemektedir. Uzaktan algılama teknolojisinin büyük bir arşiv oluşturma ve ulaşım imkanı olmayan bölgelerden bilgi elde etme imkanı sağlaması önemli avantajlarındandır[33].

Uzaktan algılama teknolojisindeki son gelişmelere bakıldığında uydu sistemlerinin sayısı, sensörlerin farklı tiplerde olması, birçok farklı veriyi elde etme ve işleme olanağı sunmaktadır. Tüm bu teknolojik ilerlemenin sonucu olarak optik uydu verilerinin kullanımının artması ve buna bağlı olarak veri elde etme maliyetlerinin de azalması beklenmektedir [34].

Uzaktan Algılamanın Bileşenleri ise verinin oluşturulmasından işleme kısmına kadar olan tüm adımları ayırdığımızda uzaktan algılamanın bileşenlerini sıralayabiliriz. Bu bileşenler; enerji kaynağı, atmosfer, algılanan cisim, algılayıcı ve son işlemlerdir.

Uzaktan algılama için öncelikli enerji güneşten gelen ışınlardır. Güneşten gelen ışınlar biyolojik unsurların yaşamını sürdürmesini sağladığı gibi aynı zamanda uzaktan algılamanın da yapılabilmesine imkan sağlar. Bu yüzden uzaktan algılamanın temel bileşeni ve ana kaynağı güneştir. elektromanyetik spektrumlar öncelikle atmosferden yeryüzüne ulaşırlar. Güneşten yeryüzüne ulaşana kadar atmosferik etkilerden dolayı bazı değişimler meydana gelir. Bu değişiklikler cisimden tekrar algılayıcıya ulaşana kadar devam eder. Bu yüzden hem kaynaktan cisme hem de cisimden algılayıcıya ulaşana kadar atmosferle etkileşim halinde olduğu için atmosfer uzaktan algılamanın bir diğer bileşenidir[3].



Şekil 2.2 : Uzaktan Algılamanın Bileşenleri [3].

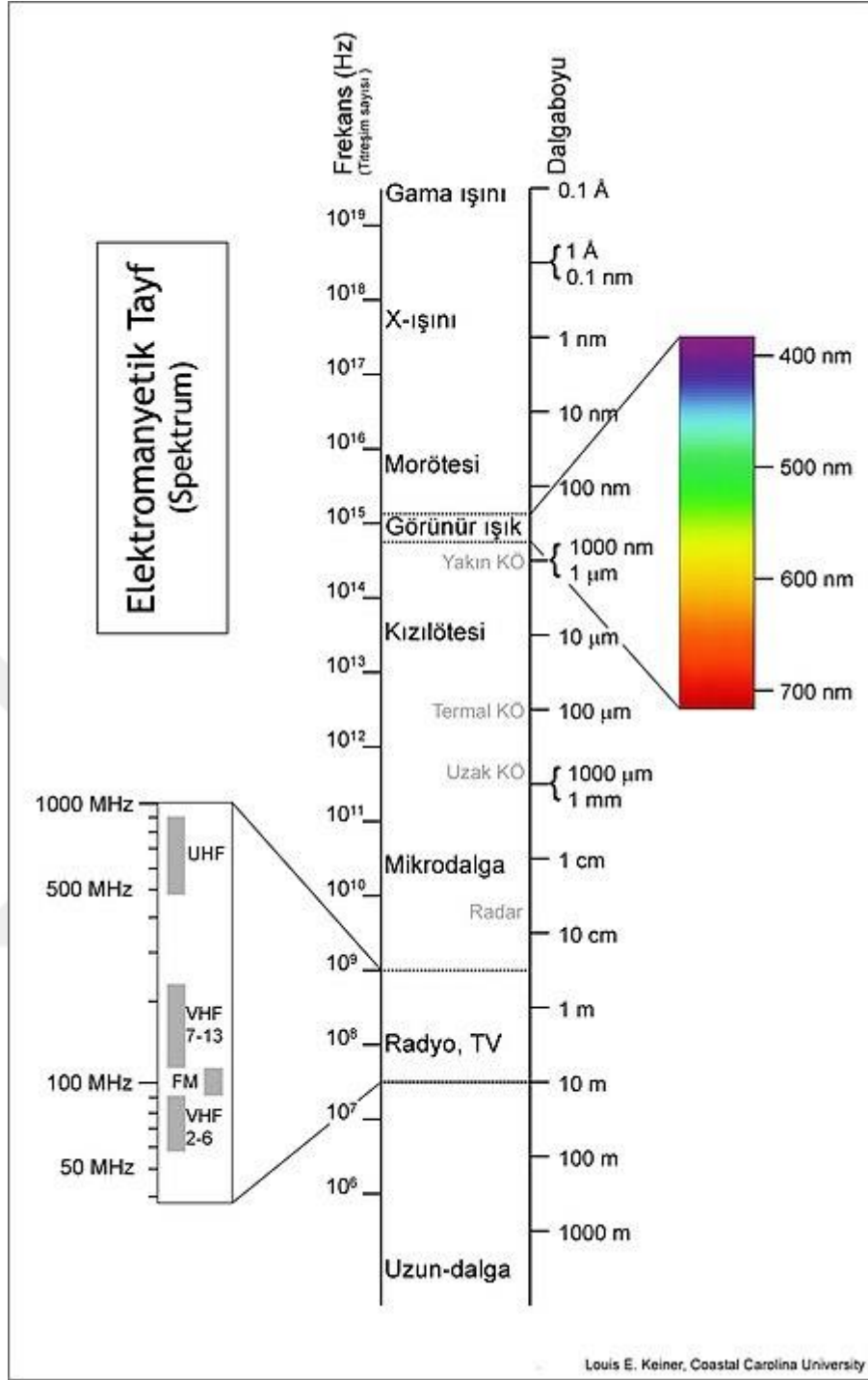
Hedef cisim uzaktan algılamanın bileşenlerindedir. Dünyadaki tüm cisimler bunların fiziksel ve kimyasal yansımaları ve dış görünüşleri ayırt edici özellik olarak uzaktan algılamada kullanılır. Bu ayırt edici özellikler farklı sensörler tarafından kaydedilir. Son olarak uzaktan algılamanın en önemli faktörü sensördür. Sensörler farklı türde elde ettiği verilerle ilgili amaca yönelik seçilebilir. Bunlar termal, optik veya radyo dalgaları şeklinde olabilir [35].

2.2 Elektromanyetik Spektrum

Gamma ışınlarından radyo dalgalarına kadar olan aralıktaki tüm radyasyon değerleri elektromanyetik spektrumda yer alır [36].

Elektromanyetik spektrum ışık ve ısının dalga boylarına göre ayrılmasını sağlar. Tüm elektromanyetik ışınlar boşlukta hareket ettiği varsayılarak elektromanyetik spektrumda sınıflandırılmışlardır. Elektromanyetik spektrumdaki bölgeler dalgaların yayılma hızı, geçirgenlik oranı, tutulma özelliklerine göre sınıflandırılır [37].

Yeryüzündeki cisimlerin karakteristik özelliklerinden biri de farklı spektral yansıtma özelliklerine sahip olmasıdır. Örneğin; tarlada yetişen bir mısır ekildikten toplanana kadar her aşamada farklı spektral yansıtma değerleri göstermektedir. Bu spektral özellikler toprağın içerdiği mineral, su miktarı, toprağın rengi ve güneşin uzaklığı gibi değişkenlerden etkilenmektedir [35].



Şekil 2.3 : Elektromanyetik Spektrum [4].

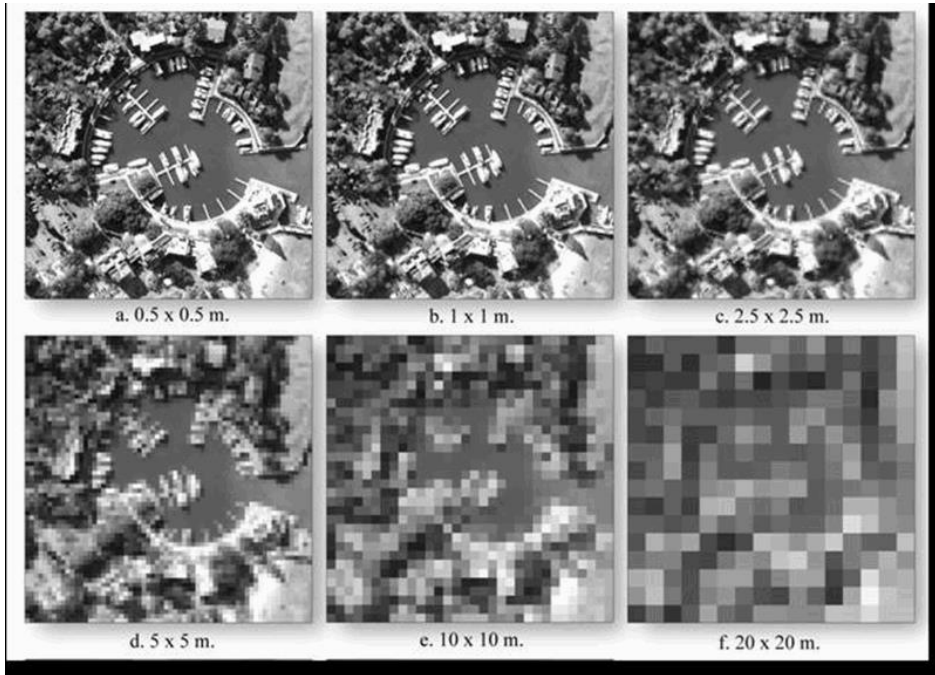
Elektromanyetik spektrumdaki yansımaya çeşitleri tam yansımaya ve dağınık yansımaya olarak ikiye ayrılır.

Tam yansımaya: pürüzsüz yüzeylerde herhangi bir saçılmaya uğramadan teorik olarak gerçekleşir.

Dağınık yansımaya: pürüzsüzlüğün yüksek olduğu yüzeylerde meydana gelir. Enerji kısmen yansıtıldığında kalan enerji diğer yönlerde saçılmaktadır. Cisimlerin yüzey pürüzlülüğü de spektral yansıtma değerini etkilemektedir. Ayrıca hedef cisimlerin yansıtma değerlerinin bilinmesi seçilecek algılayıcının bantlarını belirlemede önemlidir. [38].

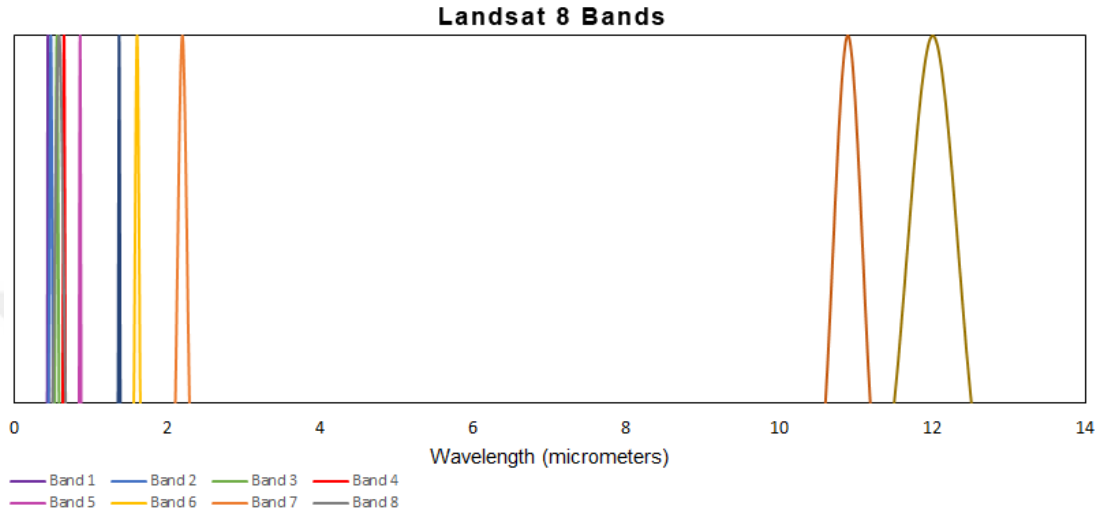
2.3 Çözünürlük Türleri

Uzaysal Çözünürlük: Sensörün zemin örnekleme aralığı olarak ifade edilebilir. 0,5 m uzamsal çözünürlüğe sahip bir sistemde, bir pikselin kenarları 0,5 x 0,5 metre olarak ölçülecektir. Ve bu alan o pikselde bir renk tonu ile ifade edilecektir. Nesne bulunduğu yerden ayırt edilebilecekse bulunduğu bölgeden farklı bir renkte olmalıdır. Örneğin 1 metrelik bir uzaysal çözünürlük, 1 metrelik bir bisikleti bulunduğu yerden ayırt etmek için yeterli olacaktır, ancak bisikletin tam bir görüntüsünü vermeyecektir. Bu nedenle, her pikselin kapladığı alan ne kadar küçükse, o kadar fazla ayrıntı dahil edilir, dolayısıyla uzamsal çözünürlük o kadar yüksek olur [5].



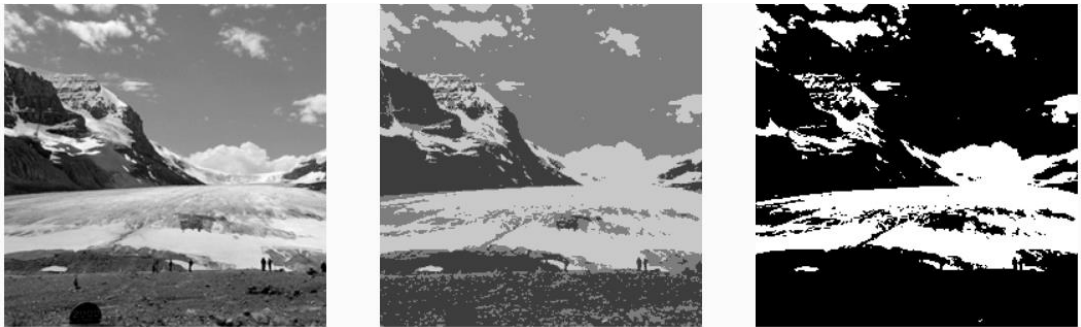
Şekil 2.4 : Mekânsal çözünürlük [5].

Spektral Çözünürlük: Sensörün EMS'de kaydedebileceği belirli bir dalga boyu aralığıdır. Dalga boyu aralığı daraldıkça, spektral çözünürlük artar ve aralık artarsa spektral çözünürlük daha genel olarak ifade edilir. Bir nesnenin spektrum değeri, o nesnenin yaydığı elektromanyetik radyasyondur [39].



Şekil 2.5: Landsat 8 uydusunda bulunan bantların elektromanyetik spektrumda gösterimi [6].

Radyometrik çözünürlük: Bir algılama sisteminin görüntüdeki enerjiyi kaydettiği aralığın genişliği. Örneğin 6 bitlik bir görüntü, 2'den 6'ya kadar 64 parlaklık değerine bölünerek ifade edilir. 16 bitlik bir uydu görüntüsü, 65536 parlaklık değerine bölünerek ifade edilir.



Şekil 2.6: 8 bitlik görüntü 4 bitlik ve 2 bitlik görüntü[6]

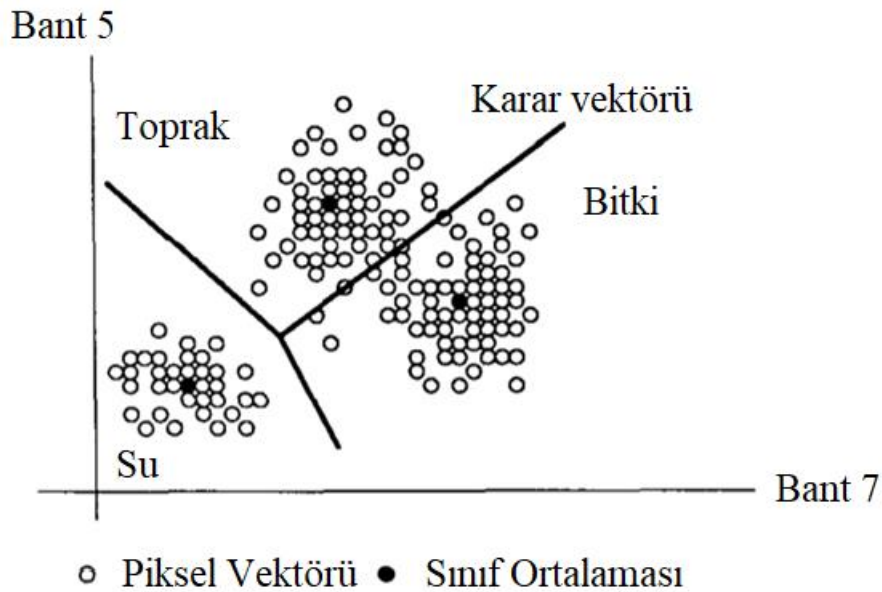
Zamansal çözünürlük: Bir görüntüleme sisteminin görüntüyü aldığı tarihten bir sonraki çekim tarihine kadar geçen süredir. Burada önemli olan görüntüyü aynı açıdan yakalama sıklığıdır. Bir uydu sisteminin yörüngesini değiştirmek mümkün olmamakla birlikte, aynı bölgenin yönlerini değiştirmek farklı bir açıdan elde edilebilir. Ancak bu

görüntü zamansal çözünürlüğü değiştirmez. Sonuç olarak, görüntü aynı açıdan ne kadar az zaman çekilebilirse, zamansal çözünürlük o kadar yüksek olur. [40].

2.5 Sınıflandırma Yöntemleri

En başta hangi sınıftan olduğu bilinmeyen birçok farklı objeyi bir gruba atama işlemine sınıflandırma adı verilir [41]. Ayrıca veri madenciliği de işe yarar verilerin bilgiye dönüşmesi için ayrıştırma yapma işlemidir.

Çok bantlı görüntü sınıflandırma görüntünün ilgili bantlarının ayrı ayrı radyometrik değerlerine göre ayırma işlemidir. Bu sınıflandırma birbirine yakın olan piksel değerlerini gruplayarak çalışır ve her pikselin dahil olduğu sınıf isimlendirilerek nesnelere ayrılmış olur. Bahsedilen kümeleme işlemi gerçekleştirilmek için denetimli sınıflandırma veya denetimsiz sınıflandırma olarak isimlendirilen iki method mevcuttur. Denetimli multispektral sınıflandırmada, operatör kullanılacak sınıfları belirler. Denetimsiz çokbantlı bir sınıflandırma işleminde ise, algoritma kullanılması gereken sınıfları kendisi oluşturur [42].



Şekil 2.7: Uzaktan Algılamada Piksellerin Makine öğrenmesi ile sınıflandırılmasının gösterimi [7].

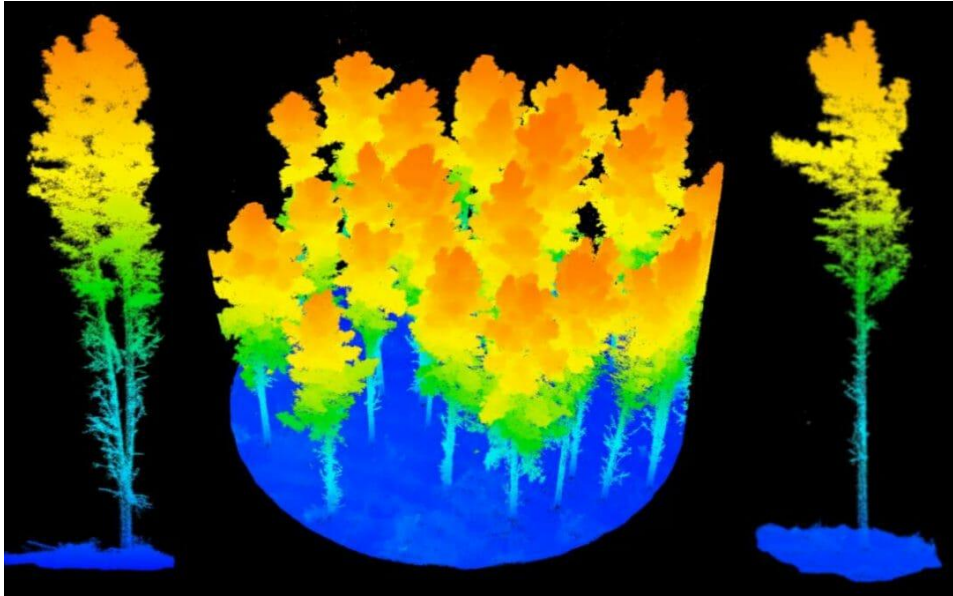
Eğitim verisi seçmek uzun bir çalışma gibi görülmese de denetimsiz sınıflandırma ile karşılaştırıldığında daha yüksek doğruluk ve tutarlı bir sınıflandırma sağladığı için denetimli sınıflandırma uygun görülür [43]. test verileri belirlenerek kontrollü sınıflandırma işlemi başlatılmış olur, Objelerin spektral niteliklerini kayıt altına alınır.

Test verilerinde kayıt altına alınan deęerlere gre dięer piksellerde sınıflandırılır. Burada maximum benzerlik yntemi ile her piksel kendisine en yakın deęerdeki sınıfa atanır.

2.5 Uzaktan Algılamada VERİ Elde Etme Yntemleri

Uzaktan algılamada veri toplama yntemlerini sınıflandırmak iin 3 tr platform vardır. 100 metreye kadar kısa menzilli kara tabanlı platformlar; Orta blge 250 metreye kadar ve uzun blge 1 kilometreye kadar 3 alt bařlık altında incelenebilir. Yer tabanlı platformların rnekleri, kuleye monte edilmiř bir radar veya bir meteoroloji istasyonundaki hava durumu gzlem cihazlarıdır. İkinci platformlar bir uaęa monte edilir ve veri toplanır. Bu uaklar insansız hava araları uaklar, helikopterler, balonlar veya uaklardır. Son olarak, uzayda inřa edilmiř platformlar var. Uzay tabanlı platformlara rnek olarak roketlere, uydulara veya uzay istasyonlarına monte edilen sensrler ile veri elde edilebilir. Uzaktan algılamada ikinci sınıflandırma tr olarak sensre gre bir ayırım yapılabilir. Sensr, bir sensr len bir cihazdır. Mekanik veya elektronik bir cihaz ile gzle grlebilen veya grnmeyen parametreyi elektrik sinyaline dnřtrerek ilgili platformda iřlemeye hazır hale getirir. Sensr eřitleri ařaęıdaki gibi aıklanabilir.

Aktif Sensrler: Radar lidar veya sonar gibi sistemin kendisi tarafından retilen enerjinin yansımısını kaydetmeye dayanır.



řekil 2.8 : Aktif sensrlere rnek LİDAR verisinin grseli [8].

Pasif Sensörler: Doğal bir kaynaktan üretilen enerjinin sistem tarafından kayıt altına alınması esasına dayanır. Uzaktan algılama uydusu ile güneşten gelen ışığı fotoğraflamak buna bir örnektir.

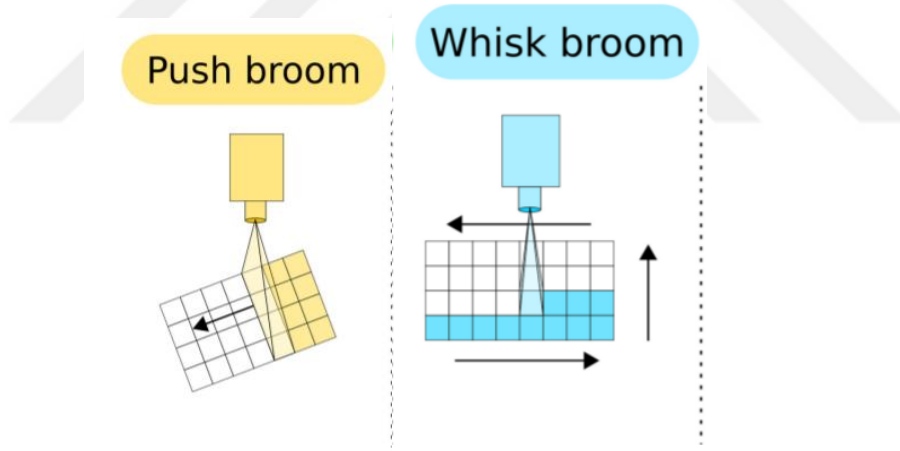
Taramayan sensörler: Bu sensörler, kameralar gibi tüm alanın elektromanyetik enerjisini tek seferde kaydeden sistemlerdir.

Tarama sensörleri: Bilgileri parçalar halinde elde etmek için alanın radyasyon enerjisini kaydederek sonuç verilerini oluşturur. Tarama sensörleri, uzunlamasına çizgilerde veya pikseller temelinde algılayabilir.

Görüntüleme yapmayan sensörler: Bu sensörler verileri görüntü olarak almaz. Zamana bağlı olarak spektral yansımayı veya sayısal değeri kaydederler. Örneğin, termal ölçüm yapan sensörler. **Görüntüleme yapan Sensörler:**

Whiskbrom Tarayıcı: nokta bazında ilerler ve bir tarama aynasına sahiptir.

Pushbrom Tarayıcı: Tarama planı yoktur ve görüntüyü satır olarak kaydederek ilerler.



Şekil 2.9 : Pushbroom ve Whisbroom tarayıcı [9].

3. YÖNTEM VE MATERYAL

3.1 Çalışma Alanı ve Uydu Görüntüleri

3.1.1 Çalışma Alanı

Serdivan ilçesi Marmara bölgesinde Sakarya ilinin batısında yer alan bölgedir. Serdivan ilçesi ulaşım ağlarının yoğun olarak geçtiği bir bölge olması dolayısıyla sanayi, şehirleşme, ormanlık ve yeşil alanların iç içe bulunduğu bir bölgedir. İlçe Adapazarı ovasının uç bölgelerindeki tepeler üzerine kuruludur. Bahsedilen bölgelerin bir kısmı şehirleşme alanıyken diğer kısımlar meskun mahal dışı bölgelerdir. Ayrıca Sapanca Gölü ilçenin sınırları içinde yer almaktadır. 11.385 hektarlık yüzölçümünün 2000 hektarlık kısmı orman ve yeşil alandır. Sakarya ilinin geneline bakıldığında ormanlık alan %44 iken Serdivan ilçesi için %16.8 dir. 17 Ağustos 1999 Gölcük depreminin sonrasında Serdivan ilçesi hızla depremin etkilerini silmiş, sanayileşme ve şehirleşme oranı hızlı bir biçimde artmıştır. Çalışma alanı olarak Serdivan ilçesinin seçilmesinin sebebi bölgenin birçok sanayi, meskun mahal, orman, yeşil alan, tarım alanı ve Sapanca gölünü bünyesinde barındırmasıdır [44].



Şekil 3.1: Serdivan ilçesine ait 4 bantlı 1.5 mt çözünürlüklü Spot-7 Uydu görüntüsü

3.1.2 SPOT 7 UYDUSU

SPOT 7 (Satellite pour l'Observation de la Terre) ticari bir Dünya görüntüleme uydusudur 2014 yılında piyasaya sürüldü ve faaliyete devam etmektedir. 2008 yılında, Fransa Spot uyduları, sürdürülebilir geniş alanlı yüksek çözünürlüklü gözlem hizmetlerine devam etmek için SPOT-6 ve SPOT-7 olarak adlandırılan yeni bir ticari SPOT uydusu planlanmıştır. SPOT serisindeki son uydular, SPOT 6 ve SPOT 7 aynı uydulardır ve 2024'e kadar veri sürekliliğini sağlar. SPOT 7, tüm SPOT görevleri gibi, yüksek çözünürlüklü, geniş alanlı optik görüntüler sağlamayı amaçlamaktadır. Bu, Dünya kaynaklarının bilgi ve yönetimini geliştirmek, iklim bilimi ve oşinografiyi içeren olayları tespit etmek ve tahmin etmek ve insan faaliyetlerini ve doğa olaylarını izlemek için tasarlandı [1].

Öncesinde SPOT-5 görevi tarafından sağlanmakta olan verileri daha hızlı ve yüksek çözünürlükte ve spot 5 uydusunda ki sistematik hataları gidererek Spot 6 ve 7 uyduları planlanmıştır. Birbirini zamansal olarak tamamlamak üzere planlanmış olan SPOT-6 ve Spot-7'nin üretimi, 2009 ortalarında Astrium Services CEO'su Eric Beranger tarafından resmi olarak duyurularak başlatılmıştır. 2011 de gönderilmiş SPOT-4 ve SPOT-5 uydularında mevcut çözünürlük 2,5, 5, 10 ve 20 m olmakla beraber geniş görüntü yelpazesini sunmuştur. Spot 5 in geliştirilmiş versiyonu olan SPOT-6 ve -7 serisinin amacı, son kullanıcılara sürdürülebilir bir operasyonel hizmet sağlamaktır. Teknik olarak SPOT-6 ve Spot-7, Pléiades-1 ve Pléiades 2 programındaki yeni gelişmelerin yanı sıra en son onaylanmış teknolojiyi kullanarak SPOT-5'in tüm deneyimini kullanarak geliştirilmiştir [45].

Yüksek düzeyde kapsama kabiliyetini korunması amacıyla için 60 km' görüntü örnekleme alanını 1,5 m ortometrik görüntü sunması ayrıca daha iyi çözünürlük ve doğal doğal renkli görüntüler elde etmek için mavi bir bant eklenmesi hem geniş kapsama alanlarının toplanmasını hem de ayrıntılı obje gözlemini verimli bir şekilde gerçekleştirmeyi sağlaması ve Spot 5 uydusunun sistematik ve teknik hatalarından arındırılması hedeflenmiştir. Uzaktan algılama için önemli bir sorun olan bulutluluğun önlenmesi ve buluttan arınmış görüntüler elde etmek için mümkün olan günlük 6 adede kadar programlama planlanması önemli geliştirmelerdir. Ayrıca SPOT-6 ve SPOT-7 'nin birbirini takip ederek planlanması durumunda aynı bölge için günlük yenilenen görüntü sağlaması dikkat çeken bir gelişmedir. Spot 6 ve / uyduları için tahmini 10 yıllık kullanım süresi 2023 yılına kadardır. SPOT-6 ve SPOT 7 uzay aracı, ticari

verileri Astrium üyesi Infoterra tarafından sağlanan DLR uydusunun 'nin 1 m çözünürlüklü TerraSAR-X ve TanDEM-X radar uyduları ve yüksek çözünürlüklü Pleiades-1 ve -2 ile birlikte çalışmaktadır. İki Pleiades uzay aracı devlet tarafından finanse edilmesine rağmen Spot Image tarafından ticari kullanım amaçlanmıştır. Bu nedenle, Spot uyduları, iki Yüksek Çözünürlüklü uyu 1,5 m görüntü pankromatik ve iki Çok Yüksek Çözünürlüklü uyu yani 50 cm görüntü aralığı sunan ilk ticari operatör olarak görevini sürdürmektedir [45].

Çizelge 2.1 :[1] Spot-7 Uydusu Teknik Özellikleri

Yörünge Yüksekliği	694 km
Yörünge Tipi	Güneş eşzamanlı
Yörünge eğimi	98.2 °
Yerel ekvator geçiş süresi	alçalan düğümde 10:00 AM
Yörünge Süresi	98.79 dakika
Döngüyü	26 gün tekrarlayın
Aynı anda çalışan SPOT 6 ve SPOT 7	1 gün
Tek uydu ile	1 ila 3 gün arası

New Astrosat Optical Modular Instrument (NAOMI) SPOT 7 üzerindeki iki özdeş NAOMI cihazı, 1,5 m'ye kadar uzaysal çözünürlüklü pankromatik ve multispektral veri sağlar. Aynı anda veya ayrı ayrı olmak üzere iki modda çalışabilirler. Eşzamanlı modunda çalıştırıldıklarında şerit genişliği 120 kilometrekaredir ve bireysel modunda çalıştırıldıklarında 60 kilometrekaredir.

Çizelge 2.2 :Spot-7 Uydusu Bant Özellikleri [1]

Özellik	SPOT-7
Zemin Örneği Mesafesi	Pankromatik: en düşüğe 1.5 m
Multispektral	Nadir 6 m
Şerit Genişliği	60 km
Bakış Alanı	± 30 °
Bantlar PAN	0.45-0.75 µm
Mavi	0.45-0.52 µm
Yeşil	0.53-0.60 µm
Kırmızı	0.62-0.69 µm
NIR (Yakın Kızılötesi)	0.76-0.89 µm

Diğer araçlara bakılacak olursa SPOT programında bitkisel indeksleri oluşturmak ve tarımsal faaliyetleri etkin biçimde tespit etmek amacıyla VEGETATION sensörü SPOT 7 uydusuna eklenmiştir. Bu sensör, 2250 kilometrekarelik görüntü örnekleme alanı sayesinde günlük olarak neredeyse tüm Dünya yüzeyinde veri toplamaktadır. Bu bitkisel sensör dünya yüzeyinden yansıyan ışığı dört spektral bantta (mavi, kırmızı, yakın kızılötesi ve orta kızılötesi bantlarında) veri toplar.

3.1.3 Kullanılan Yazılımlar

Sakarya İlinin Serdivan ilçesine ait 23 Temmuz 2020 tarihinde elde edilmiş uydu görüntüleri alınarak içerisinde bulutluluk oranı en düşük olan görüntüler seçilmiştir. Çalışmada kullanılan Spot-7 uydu görüntüleri 1.5 metre çözünürlükte pankromatik ve 2.5 metre Multispektral çözünürlükte olacak şekilde İTÜ - Uydu Haberleşme ve Uzaktan Algılama UYG-AR Merkezi 'nden alınmıştır.



Şekil 3.2. : Çalışma Bölgesinin NIR (Yakın Kızılötesi) Bantta görüntüsü

Alınan görüntüler ENVI yazılımında Multispektral ve Pankromatik uydu görüntüleri birleştirilmiştir. Ayrıca ENVI programında Görüntüye ait histogram iyileştirme ve kesme işlemi gerçekleştirilmiştir. Görüntü birleştirme, iyileştirme ve kesme işlemlerinin sonrasında ArcGIS yazılımında işlenen görüntü üzerinde şekil dosyaları oluşturularak her sınıfa ait test ve eğitim verileri kaydedilmiştir. Kaydedilen veriler R programına aktarılmıştır. Ardından R programında ilgili kodlar ve kütüphaneler kullanılarak belirlenen algoritmalarla görüntüler sınıflandırılmıştır. Sınıflandırılan görüntüler tekrar ArcGIS yazılımında işlenerek sonuç görüntüleri ortaya çıkartılmıştır.

3.2.Makine Öğrenmesi Yöntemleri

3.2.1 Makine Öğrenmesi

Makine öğrenmesi terimi veri topluluklarından en etkin biçimde yararlanmayı ve bunun günden güne daha yeni ve verimli metotlarını geliştiren önemli bir çalışma alanı haline almıştır. En yeni makine öğrenme teknikleri ile veri bilimcileri ileride mevcut sistemi nelerin beklediğini gerçeğe en yakın tahmin edebilmeye ve değişiklikleri sürekli olarak takip edebilme yeteneğine sahip olmayı istemektedir. Veriler sürekli olarak güncellendiğinden, makine öğrenmesi algoritmaları çözümün sürekli olarak güncel tutulmasını sağlar. Makine öğrenmesi için ancak en etkin ve sürekliliği sağlanan veri kullanırsa, bir sonraki adımı kestirme olanağına sahip olunabilir. Makine öğrenmesi, bir sistemin açık programlama ile değil verileri kullanarak öğrenmesini sağlayan bir yapay zeka anlayışıdır. Ama makine öğrenimi uzun ve gelişmeye sürekli devam eden bir süreçtir [46].



Şekil 3.3 : Makine öğrenmesi ,Derin Öğrenme ve Yapay Zekanın birbiri ile ilişkisi [10].

Makine öğrenmesi metotlarını ön plana çıkaran sebeplerin başında Makine Öğrenimi, büyük hacimli verileri inceleyebilir ve insanların fark edemediği belirli eğilimleri ve baağlantıları keşfedebilir. Örneğin, Amazon gibi bir alışveriş sitesine bakıldığında, kendileriyle ilgili doğru ürünleri, anlaşmaları ve hatırlatıcıları karşılamak için

müşterilerin web sitesi üzerindeki davranışlarını yorumlamaya ve satın alma geçmişlerini etkin şekilde takibine yarar. Makine öğrenmesi ile elde edilen veriler, müşteriye özel reklamları onlara göstermek için kullanır. Makine öğrenimi sayesinde her noktada çalışmaya birebir insan gücü gerekmez. Bu algoritmalar makinelere öğrenme yeteneği kazandırmak anlamına geldiğinden, çıkarımlarda bulunmalarına ve buna ek olarak algoritmaları kendi kendine geliştirmelerine imkan tanır. Bunun başka bir örneği zararlı yazılım tespit uygulamalarıdır; yeni tehditleri fark edildikçe filtrelemeyi öğrenirler. Makine öğrenimi, zararlı içerikleri tanımda oldukça iyidir. Makine öğrenimi algoritmaları deneyim kazandıkça, doğruluk ve verimlilik açısından gelişmeye devam ederler. Bu, daha tutarlı ve işe yarar sonuçlar almalarını sağlar. Eğer bir hava tahmini modeli yapmak için makine öğrenmesi algoritmaları kullanılacaksa sahip olunan veri sayısı arttıkça, algoritma daha tutarlı tahminlerde bulunmayı daha çabuk öğrenecektir [47]. yakın zamanda, çok daha etkili bir şekilde öğrenebilen yapay zekaları gözlemlemiş olacağız. Bu sayede minimum insan kontrolü ile kendi kendine çalışabilen ve hatalarından beslenen yapay zekalar yakın zamanda karşımıza çıkacaktır [48].

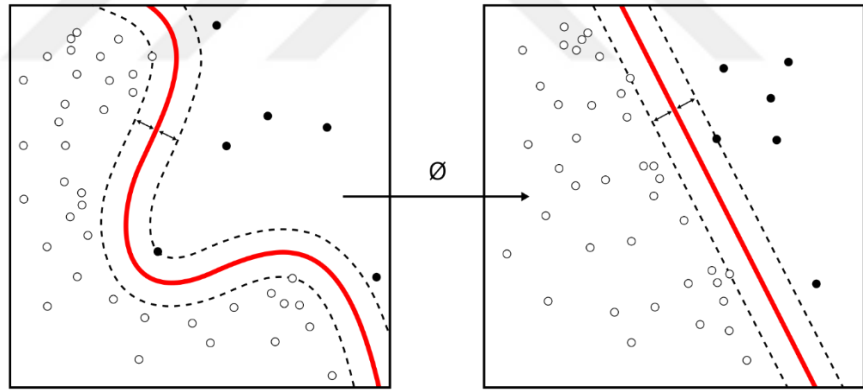
Sınıflandırmada önemli bir konu olan denetimli öğrenme ise, önceki deneyimden veri toplamaya veya bir veri çıktısı oluşturmaya olanak tanır. Önceki deneyimlerini kullanarak performans ölçütlerini geliştirmeye yardımcı olur Denetimli makine öğrenimi, çeşitli gerçek dünya hesaplama problemlerini çözmeye yardımcı olarak elde ettiği veriler sayesinde hayatımıza verim ve planlama konusunda olumlu etkiler göstermektedir. Denetimsiz makine öğrenimi ise verilerdeki her tür ortaya konmamış modeli tespit eder. Denetimsiz yöntemler, kategorize etme için yararlı olabilecek özellikleri bulmanıza yardımcı olur. Gerçek zamanlı olarak gerçekleşir, bu nedenle tüm girdi verileri analiz ederek kullanıcının göremediği detayları gün yüzüne çıkarabilir [49].

3.2.2. Destek Vektör Makineleri

En popüler öğrenme algoritmalarından biri, etkin bir hiper düzlem bularak verileri iki kategoriye ayıran Destek Vektör Makineleri'dir. Etkin hiper düzlem, iki sınıfta birbirine en yakın noktalar olan destek vektörleri kullanılarak oluşturulur. Hiper düzlem ile destek vektörleri arasındaki boşluğa kenar boşluğu denir ve DESTEK VEKTÖR MAKİNELERİ daha iyi ayırma için maksimum sınırı bulmayı amaçlar. [50,

s. 229] Destek Vektör Makineleri en eski yöntemlerden biri olmasına rağmen çoğunlukla iyi sonuçlar verir. Özellik miktarının örneklem miktarından fazla olduğu durum için, çekirdek fonksiyonlarının doğru seçilmemesi sebebiyle aşırı öğrenmeye neden olabilir. Destek Vektör Makineleri, sınıflandırma ve duyarlılık analizi gibi metin sınıflandırma işlevleri içinde oldukça kullanışlıdır. Buna ek olarak elektronik posta otomasyon işlemlerine ek olarak görüntü tanıma, renk tabanlı sınıflandırma, el yazısı rakam tanıma gibi birçok farklı konuda tercih edilir [51].

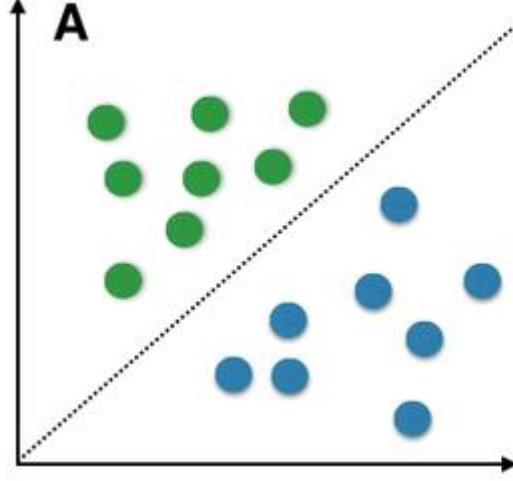
Doğrusal destek vektör makineleri, lineer olarak ayrılabilen veriler için uygundur. Doğrusal destek vektör makineleri bir veri kümesinin tek bir doğrusal fonksiyonu kullanılarak iksi sınıfa ayırır. Bu durumda bu tür veriler lineer olarak ayrılabilir veri olarak adlandırılır ve sınıflandırıcı, lineer destek vektör makineleri sınıflandırıcı olarak adlandırılır. Doğrusal Olmayan destek vektör makineleri, doğrusal olmayan ayrılmış veriler için kullanılır; bu, bir veri kümesinin düz bir çizgi kullanılarak sınıflandırılmaması durumunda, bu tür veriler doğrusal olmayan veri olarak adlandırılır ve kullanılan sınıflandırıcı, doğrusal olmayan destek vektör makineleri sınıflandırıcı olarak adlandırılır [52].



Şekil 3.4 : Destek vektör sınıflandırıcısının gösterimi [53].

3.2.2.1 Lineer Destek Vektör Verileri

Destek vektör makineleri sınıflandırma algoritması temelde eğitim verilerinin seçilmesi ile test verileri üzerinde algoritmanın çalışarak sınıflandırma işlemini gerçekleştirmesidir. belli sayıdaki eğitim verisi ile 2 boyutlu bir grafik şeklinde verilerin ait olduğu sınıflar ve öznitelikleri yer alır. Bu sınıflandırma yapılırken algoritmanın hiperbolik düzlem ile verileri birbirinden ayırmasıdır.[54], [55].



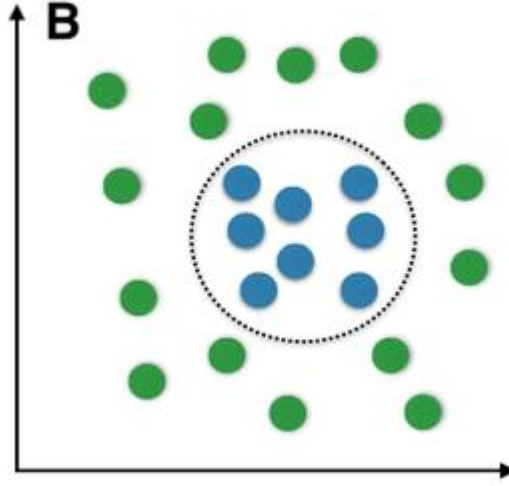
Şekil 3.5 : Lineer destek vektörü ile ayrılmış veriler [11].

3.2.2.2 Lineer Olmayan Destek Vektör Verileri

N Adet farklı özneliğe sahip doğrusal olarak ayrılabilir bir veri kümesi için bir hiper düzlem, temelde veri kümesini iki kümeye ayırmak için kullanılan $(n - 1)$ boyutlu bir alt uzay olacaktır, her küme farklı bir sınıfa ait veri noktaları içerir. Herhangi bir iki değişkenli grafik için iki adet değişkeni temsil eden tek boyutlu bir uzay oluşturulacaktır. Benzer şekilde, 3 boyutla temsil edilen bir küme için, 2 boyutlu ayırıcı bir alt düzlem oluşturulacaktır.

Makine öğreniminde Destek Vektör Makineleri algoritması (DVM), verileri birbirinden ayırt eden bir hiper düzlemi belirleyerek verileri sınıflandırma yöntemi kullanarak olasılıklı olmayan, doğrusal, ikili bir sınıflandırma yöntemidir.[56]

Non-lineer ayrılabilir bir veri setinin bir Destek Vektör Makineleri lineer bir sınıflandırıcı kullanarak sınıflandırmak için bahsedildiği gibi Destek Vektör Makineleri, verileri iki sınıfa ayırmak için öz nitelik sayısından bir eksik sayıda sınıflandırıcıyı öğrenen lineer bir sınıflandırıcıdır. Ayrıca, Non-lineer bir veri kümesinde de kullanılabilir [57].

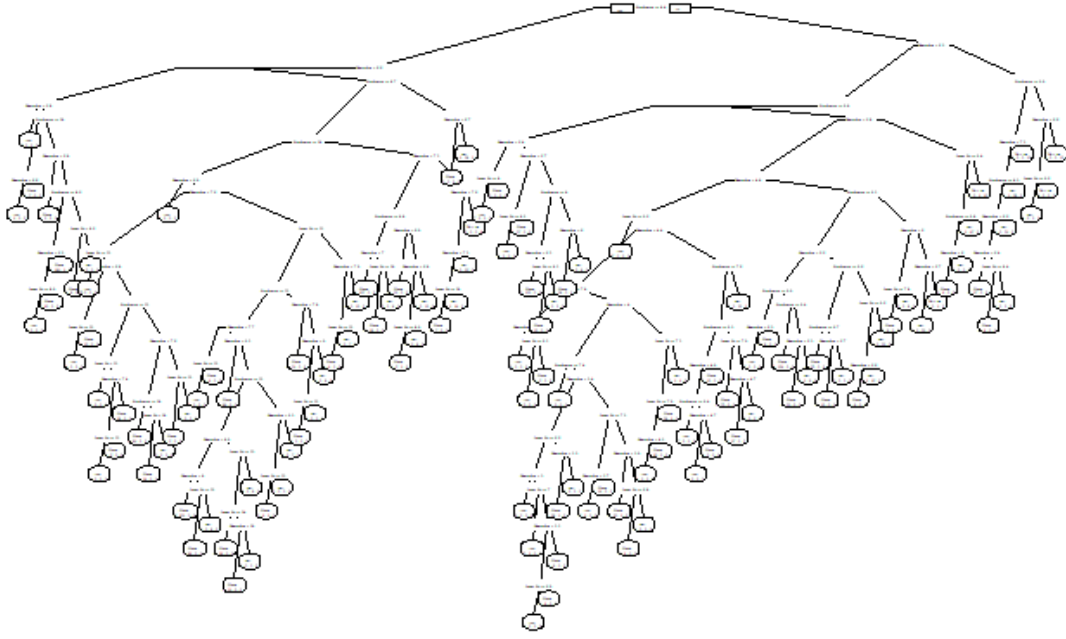


Şekil 3.6 : Linner ayrılamayan vektörlerin gösterimi [11].

3.2.3 Karar Ağaçları

Karar Ağacı algoritması, sınıflandırma ve regresyon problemleri için kullanılabilen denetimli bir makine öğrenimi yöntemidir. Tezde kullanılan en basit yöntem karar ağacı algoritmasıdır. Deneylerde kullanılan diğer yöntemlerle karşılaştırmak için karar ağacı algoritması seçilmiştir [58].

Karar ağaçları algoritması, uzaktan algılama da kullanılan önemli sınıflandırma tekniklerinden birisidir. Literatüde sınıflandırma işlemlerinde oldukça yoğun kullanılmaktadır. Karar ağaçlarında öncelikle bir karar ağacı oluşturulur, daha sonra karar ağacından üretilen kurallar ile veri tabanında bulunan kayıtlar sınıflandırılır. Karar ağaçlarında ID3, C4.5, CART, REP Tree, gibi birçok yöntemler kullanılmıştır. Veriler önce bu karar ağacı algoritmalarında ön işleme tabi tutulur. Algoritma bu verilerden bir karar ağacı oluşturur. Oluşturulan bu karar ağacı ile sınıflandırılmamış verilerin karar ağacı algoritmasına göre oluşturulmuş sınıflarının tespit edilmesi sağlanır [59].



Şekil 3.7 :Karar Ağacı algoritmasını temsil eden görsel [55].

C4-5 algoritması, literatürde yaygın kullanılan karar ağacı algoritmalarından biridir. C4-5 algoritması, test verilerinin özniteliklerini seçiminde ölçüt faktörü olarak bilgi kazancı oranı kullanmakta ve sınıflandırılan her veri seti için, maximum bilgi kazancı oranında olan öznitelik seçilmektedir. C4-5 algoritması, ID3 algoritmasının temellerine dayanan ve ID3 algoritmasının belli dezavantajlarını egale etmiş bir algoritmadır. C4-5 algoritması sürekli ve ayırık veriler için doğru sonuçlar vermektedir. C4-5 algoritmasını, eksik öznitelik değerleri içeren eğitim veri setleri ile çalışabilmektedir. C4-5 diğer algoritmalarından ayıran en önemli özellik Karar ağacı oluşturma sırasında ya da sonrasında bazı veri düğümlerin veya alt ağaçların egale ederek aşırı uygunluk sorununa çözüm üretmektedir. İşte bu sayede eğitim setindeki gürültülü verilerin çıkarılmasını sağlamaktadır [60].

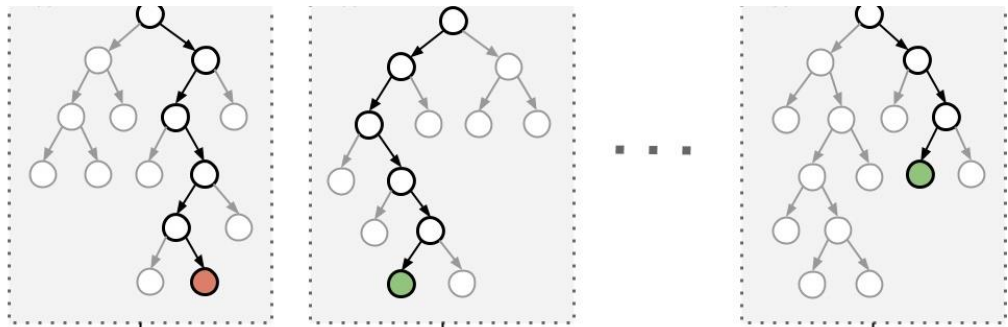
Başka bir karar ağacı yöntemi ise Hoeffding ağaç algoritması olarak bilinen ve kritik veri akışı öğrenme modeli üreten algoritmasıdır. Veri setlerinin uygun olması ve toplam örneklem sayısının yeterli sayıda olması durumunda, normal karar ağaçlarına nazaran daha doğru modeller geliştiren bir algoritmadır. Hoeffding ağaçları algoritması, hoeffding sınırı temeline dayanan bir çeşit geliştirilmiş karar ağacı algoritmasıdır. Yeterli bağımsız gözlemler sağlandığında ,rastgele bir değişkenin

ortalamasını gerçeğe yakın belirleyebilir. Ortalamayı baz alarak ilgili sınıf için yukarısında ve aşağısında sınırlar belirler. Aslında, hoeffding sınırlandırma algoritması $1 - \delta$ olasılığı için, R açıklığında raslantısal verinin n adet yorumlama işlemi sonucundan sonra tahmin edilmiş ortalamadan daha fazla değişmeyeceğini ortaya koyarak algoritmayı sınır dahilinde tutar [61].

3.2.4 Rastgele Orman Algoritması

Rastgele Orman karar ağacı algoritması ağaçlar tarafından oluşturulan sınıflandırmadır. Karar ağaçları, bu Rastgele Orman algoritmasından bağımsız olarak üretilir. Her ağaç için veri kümesine bağlı olarak ikiye ayrılır. Buna ek olarak her dallanmada önitelikler rastgele belirlenir. Bu yüzden bir öz nitelik tekrar tekrar kullanılmış olabilir. Bu yüzden Rastgele Orman sınıflandırma sonuçlarını değerlendiren başka bir algoritma oluşturur. Yani Rastgele Orman algoritması aslında, K karar ağaçlarından meydana gelmiş ve onu sınanan bir algoritmadır [62].

Rastgele Orman sınıflandırıcısı, regresyon problemlerine çözüm arayan birden fazla makine öğrenme algoritmasının bir araya geldiği bütüncül bir algoritmadır. Birleştirdiği algoritmaların negatif yanlarını egale ederek diğer modellerin güçlü olduğu noktaları onların erine kullanarak en etkin sınıflandırmayı sunar. Veri setlerini birbirinden ayırırken, tüm dallanmaların sonuna değer etiketi verir. Bir pikselin veya verinin alacağı en son sınıf etiketi, Rastgele Orman algoritmasında çoğunluk oyu kuramı ile saptanır. Rastgele Orman, özellikle finans sektöründe meydana gelen usulsüzlüklerin tespitinde oldukça başarılı bir performans göstermiştir [62].



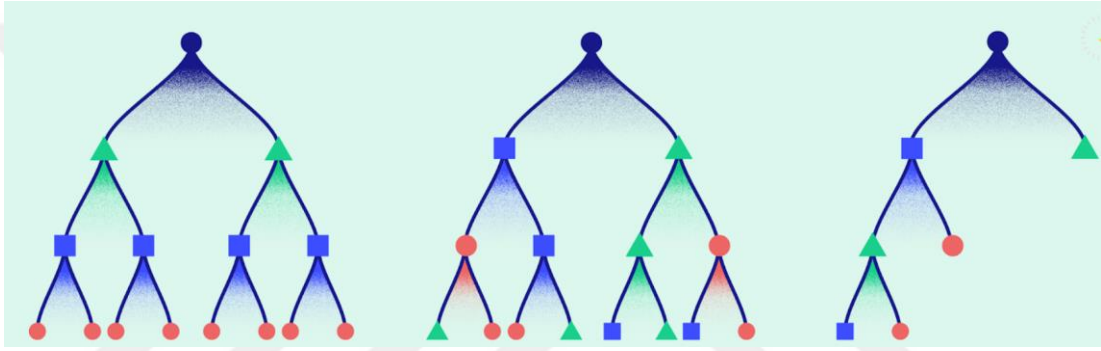
Şekil 3.8 : Rastgele Orman Algoritması [13].

Rastgele orman algoritmasında mevcut dallanmaları yorumlamak ayrıca kullanmak amacıyla, neye dayanarak hesap edildiklerini bilmek Rastgele orman algoritmasını anlamak için önemlidir. Dallanmaların tamamı rastgele ormanların ürettiği ikili veri setlerine göre şekillenir. Hali hazırdaki dallanma için yeni eğitim verileri örneklenerek koşulduğunda ortalama%30 luk bir kısmı bir önceki sınıfın üyesi olmaz. İşte bu dışarıda kalan verileri doğru sınıfa yerleştirmek için, yeni oluşturulan karar ağaçları da doğruluk tahminini elde etmek için kullanılır. Aynı zamanda bu dışarıda kalan veriler sonuç tahmini üretmek için de önem arz eder. Her dallanma üretildikten sonra, bütün veriler dallanmalardan aşağı yönde koşular ardından oluşturulan her yeni durum için yakınlıklar hesaplaması yapılır. Eğer iki durum aynı dallanmayı yaparsa, yakınlık katsayıları artırılır. İşlemlerin ardından, yakınlıkların sayısı dallanma miktarına oranlanarak normalize edilir. Son olarak yakınlıkların sayısı, her sınıftaki eksiklerin tespiti, sınıf özniteliğinin dışında olan verilerin tespiti ve son olarak tespit edilmiş özniteliklere göre sınıfların üretilmesini sağlar [62].

Rastgele Orman algoritması başlıca tercih edildiği dört sektör vardır. Bunlardan ilki Finans ve Bankacılık uygulamalarıdır. Günümüzde kredi puanı belirlenmesinde genellikle Rastgele Orman algoritması kullanılmaktadır. İkinci kullanım alanı Sağlık sektörüdür. Rastgele Orman yöntemi hastalık eğilimleri ve risklerini tespitinde kullanılır. Diğer bir kullanım alanı Arazi Kullanımı sınıflarının tespitinde Rastgele Orman yöntemi araziye ait alınmış uzaktan algılama verisinin içerdiği piksel yansıtma değerlerini farklı bant kombinasyonlarına göre denetimli veya denetimsiz olarak sınıflandırmasıdır. Son olarak Rastgele Orman yönteminin kullanım alanı Pazarlama eğilimlerinin çıkarılmasıdır. Rastgele Orman yönteminin avantajı sınıflandırma ve regresyon işlemlerinin ikisinde de başarılı sonuçlar elde edilmesidir. Yoğun veri setlerini yüksek boyutsallık sayesinde başarılı şekilde modelleyebilir. Rastgele Orman algoritması genel doğruluğu yükseltmenin yanında aşırı uyum sorununun da önüne geçer [52].

3.2.5. Xgboost Hızlandırma Yöntemi

XGBoost algoritması temelde aşırı gradyan artırma metodu kullanan bir yöntemdir. XGBoost kütüphanesi, Tianqi Chen tarafından ilk olarak üretilen ve sonrasında ise literatürdeki çalışmalar ile gelişmeye devam eden gradyan artırma makinelerinin bir uygulamasıdır. XGBoost, model oluşum süratine ve modelin verimliliğini arttırmaya yönelir. Gradyan Artırma, Stokastik Gradyan Arttırma ve Düzenli Gradyan Arttırma gibi farklı algoritmaları sunar. XGBoost algoritması, işlem süresinin ve bellek kaynaklarının verimliliğini yükseltmek için çalıştırılmaktadır. XGBoost Algoritması, veri setlerinde ki eksik değerleri otonom biçimde belirler. Ayrıca, algoritma mevcut bir modeli yeni değerler içinde sürekli eğitilmesine imkan sağlar [63].



Şekil 3.9 : Xgboost hızlandırma yöntemi zayıf ve etkin algoritmaların temsili [14].

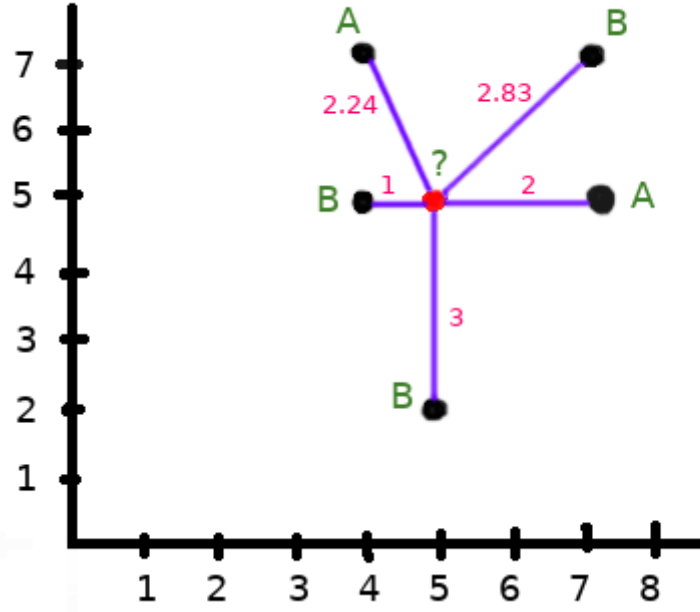
XGBoost'u algoritmasının yoğun olarak kullanılmasının esas sebebi aslında yüksek süratle ve doğrulukla işlem yapabilmesidir. XGBoost algoritması C ++ programlama dilinde kodlanmıştır. Diğer veri sınıflandırma algoritmalarına nazaran daha hızlıdır. Çekirdek XGBoost yönteminin çekirdek fonksiyonu paralelleştirilebilir hale getirilebildiğinden, çekirdek sayısı fazla olan işlemcilerde ve ağ paylaşımı üzerinde aynı anda koşulabilmesi avantaj sağlar., işte bu yüzden büyük veri setleri için eğitimi kolaylaştırır. Bahsedilen özelliğin tutarlı bir şekilde daha iyi performans göstermesi XGBoost algoritmasının en önemli özelliğidir. literatürdeki makine öğrenme algoritmalarına nazaran veri yığınlarının sınıflandırılmasında, denetimli hedef tespit işlemlerinde, kayıp verilerin belirlenmesinde, dallanma kriterlerinin belirlenmesinde ve scikit-learn uyumlu ara yüzünün olmasında ön plana çıkmıştır [64].

3.2.6. K-En Yakın Komşuluk Yöntemi

Knn, eğitim seti yardımıyla veri noktasını belirli bir kategoriye göre sınıflandırmaya çalışan parametrik olmayan denetimli bir öğrenme tekniğidir. Basit bir deyişle, tüm eğitim vakalarının bilgilerini yakalar ve yeni vakaları benzerliğe göre sınıflandırır.

K-en yakın komşular (KNN) algoritması, bir tür denetimli makine öğrenme algoritmasıdır. KNN'nin en temel biçiminde uygulanması son derece kolaydır ve yine de oldukça karmaşık sınıflandırma görevlerini yerine getirir. Özel bir eğitim aşaması olmadığı için yavaş bir öğrenme algoritmasıdır. Bunun yerine, yeni bir veri noktasını veya örneğini sınıflandırırken eğitim için tüm verileri kullanır. KNN parametrik olmayan bir öğrenme algoritmasıdır, yani temeldeki veriler hakkında hiçbir şey varsaymaz. Gerçek dünya verilerinin çoğu gerçekten herhangi bir teorik varsayımı takip etmediğinden, bu son derece kullanışlı bir özelliktir. doğrusal ayrılabilirlik, düzgün dağılım vb. [65].

K En Yakın Komşuluk algoritması, verileri diğer verilerle benzerliklerini tespit ederek kümeler oluşturma temeline dayanır. Model, depolanan kalıplar veya vakalarla tam olarak eşleşmeden, öğrenmede veri modellerini tanımanın bir yolu olarak geliştirilmiştir. Dolayısıyla iki gözlem arasındaki mesafe, birbirini tanımlayan bir kriterdir. Yeni bir gözlem sunulduğunda, modeldeki her bir gözlemden uzaklık hesaplanır. Yeni gözlem, en benzer kategoriye atanır. Bu yöntemi uygularken; öncelikle yeni gözlem ile veri setindeki tüm gözlemler arasındaki mesafe hesaplanır, ardından bu mesafeler büyükten küçüğe sıralanır, sonrasında k en küçük mesafe değerine sahip gözlemler hesaplanır, en sonda ise sınıf değerini bulmak yerine ağırlıklı oylama yöntemi de kullanılabilir. Ayrıca gözlemde en çok yinelenen kategori ve bunun tersi ağırlık olarak kullanılır. Her sınıf için hesaplanan ağırlıklar yardımıyla en yüksek ağırlığa sahip kategori sınıf değeri olarak bulunur. Öğrenme veri setindeki gözlemleri ve test örneğindeki gözlemleri belirleyerek, en yakından test edilen gözlemlerin mesafesini fonksiyon kullanarak en yakın komşuluk algoritması oluşturulur [66].

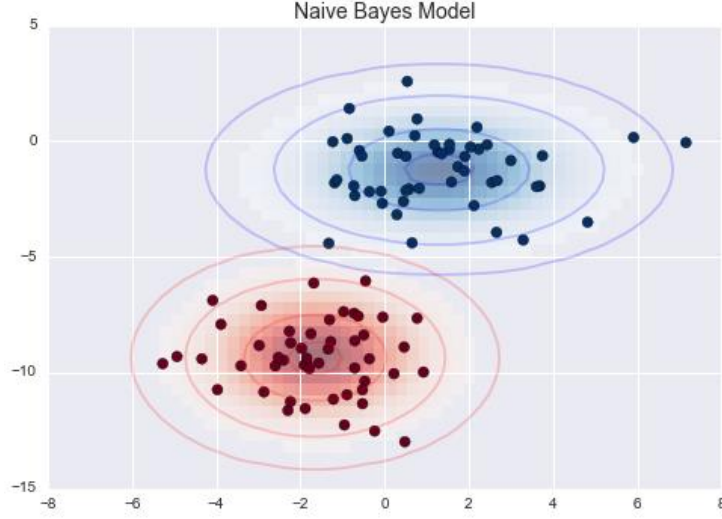


Şekil 3.10 : K-En Yakın Komşuluk Yöntemi temsili [15].

K en yakın komşuluk algoritmasının avantaj ve dezavantajlarına bakıldığında KNN algoritmasının temel teorisi incelendiğinde kontrollü makine öğrenmesi yöntemlerinin en temellerinden biridir. Algoritma seçilen verinin başka verileri uzaklığını hesaplama sistemiyle çalışır. Bu uzaklık Öklid veya Manhattan gibi bir denklem olabildiği gibi bir sabit sayıda olabilir. İşte bu uzaklığı referans alarak en yakın K veri noktalarını seçer. Son adımda, örnek veriyi K en yakın komşulukların en fazla olduğu sınıfa etiketler [65].

3.2.7 Naive Bayes Yöntemi

Naive Bayes temel işlevleri olan bir algoritma olmasına rağmen oldukça etkin ve popüler olarak tercih edilen bir makine öğrenimi yöntemidir. Bayes denkleminin arka planında karar geliştirme yöntemiyle sınıflandırma işlemi gerçekleştiren olasılıksal bir yöntemdir. Ayrıca temel bir Bayes ağı yardımıyla da Naive Bayes algoritması gösterilebilir. Naive Bayes sınıflandırıcıları, özellikle metin sınıflandırması için popülerdir ve zararlı yazılım algılama gibi işlevler için iyi bir çözümdür.[67]



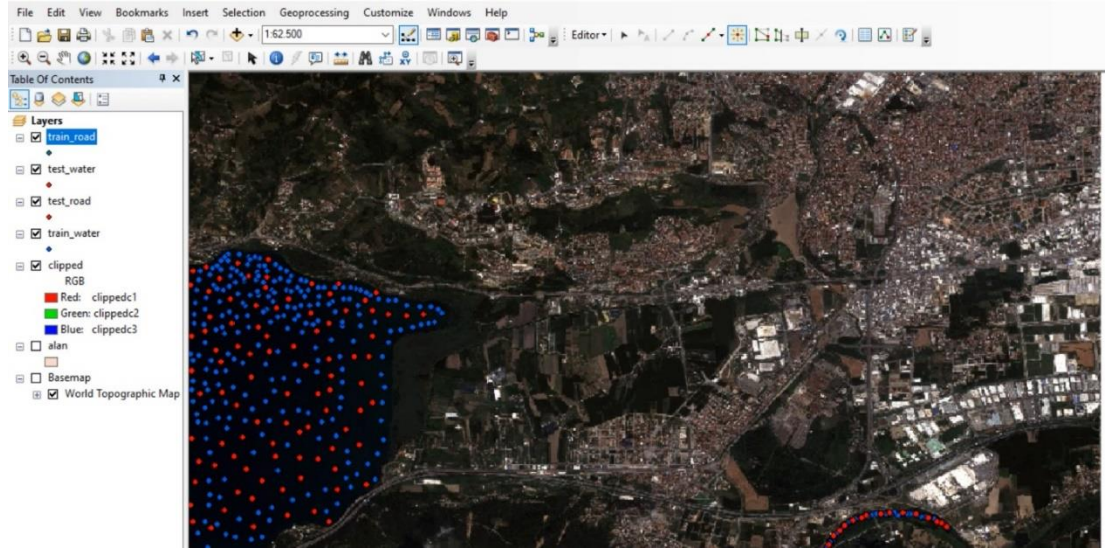
Şekil 3.11: Naive Bayes yönteminin raslantısal verilerin Bates ağı ile gösterimi [16].

Bayes teoremi, Muhterem Bayes tarafından geliştirilmiş en bilinen rastlantısal çıkarım algoritmalarından biridir ve yine de belirli kullanım durumları için son derece iyi performans göstermektedir. Tahminler arasında bağımsızlık teoremi ile Bayes'in algoritmasına dayalı bir sınıflandırma metodudur. Daha açık bir ifadeyle, Naive Bayes algoritması, bir sınıftaki belirli bir özneliğin varlığının başka herhangi bir özneliğin varlığıyla ilgisi olmadığını varsayar. Örneğin, bir meyve kırmızı, yuvarlak ve yaklaşık 3 inç çapındaysa bir elma olarak varsayılabılır. Bahsedilen öznelikler birbirlerine veya diğer özneliklerin varlığına bağlı olsa bile, tüm bu öznelikler bağımsız olarak bu meyvenin bir elma olma olasılığına katkıda bulunur ve bu nedenle saflık (naive) olarak isimlendirilmiştir [68].

4. GÖRÜNTÜLERİN İŞLENMESİ VE SINIFLANDIRILMASI

4.1. Eğitim ve Test Verilerinin Oluşturulması

Öncelikle UHUZAM CRCS' den temin edilen "TIF" uzantısında ki dosyalar ARCMAP programında açılmış ve uygun şekilde kesilerek kullanım alanı düzgün hale getirilmiştir. Ardından 16 bit'lik görüntüde 3 bant yani mavi, yeşil ve kırmızı bantlarında sınıflandırma işlemi tercih edilmiştir. Sonrasında ise Yakın Kızıl ötesi Bandı dahil edilerek işlemler tekrarlanmıştır. İlk olarak nokta bazlı şekil dosyaları WGS-84 koordinat sisteminde her bir sınıf için 150 adet test 350 adet eğitim verisi özenle sınıfı temsil edecek şekilde seçilmiştir. Noktaların seçiminde karışıklığın önüne geçmek adına geo referanslı olarak Google Maps altlığı kullanılmıştır.

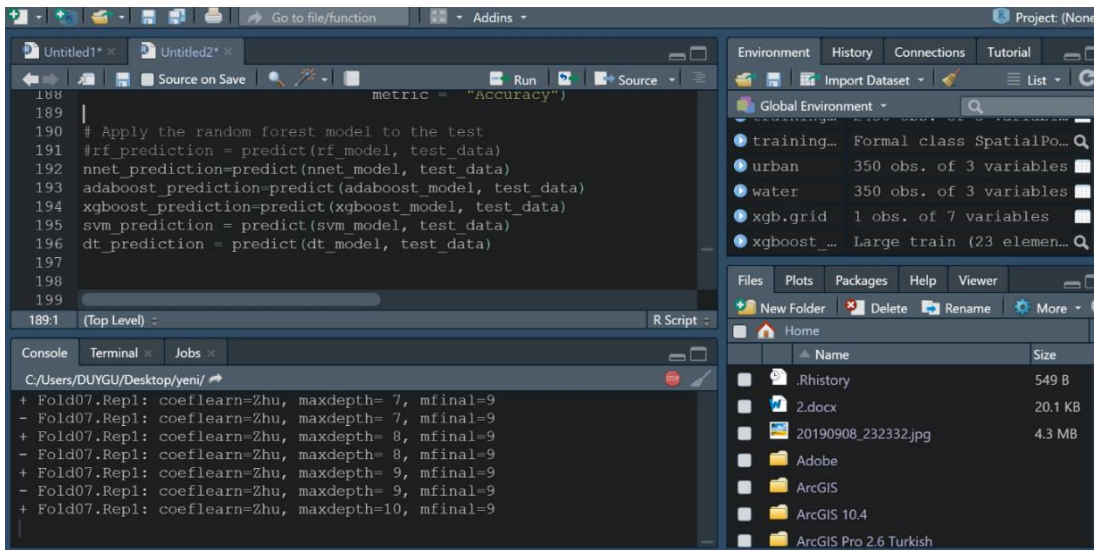


Şekil 4.1 : Eğitim ve Test Verileri toplanırken alınmış ekran görüntüsü

Kullanılan test sınıfları test_Su, test_Yol, test_Orman, test_Şehir Alanı, test_Endüstri, test_culcivated, test_Toprak sınıflarıdır. Aynı şekilde eğitim verisi sınıfları train_Su, train_Yol, train_Orman, train_Şehir Alanı, train_Endüstri train_culcivated,

train_Toprak sınıflarıdır. Test ve eğitim verileri ArcGIS programında shapefile (şekil dosyası) olarak kaydedilmiştir.

Her sınıf için oluşturulan shapefileler R programında sınıflandırılmak amacıyla işlenmiştir. Her R programında sınıflandırılmak amacıyla işlenmiştir. Şekil dosyaları R programında sınıflandırılmak amacıyla işlenmiştir. R programında makine öğrenmesi algoritmalarının etkinliği hem mavi, yeşil, kırmızı bantlarında hem de bu bantlara ek olarak yakın kızılötesi bantta 8 farklı yöntem ile incelenmiştir.



Şekil 4.2: R programında Makine Öğrenmesi Yöntemleri Çalıştırılırken alınmış ekran görüntüsü

Bunlar AdaBoost hızlandırma algoritması, karar ağaçları algoritması, K en yakın komşu algoritması, Naive Bayes algoritması, Yapay sinir ağları yöntemi algoritması, rastgele orman algoritması, destek vektör makineleri algoritması, XGboost hızlandırma yöntemi kullanılmıştır. Belirtilen makine öğrenmesi algoritmalarının R programında koşulmasının ardından sınıflandırma işlemi sonucunda oluşan hata matrisleri ve sonuç görüntüleri ayrıca kaydedilmiştir. Yakın kızıl ötesi bant ile yapılan sınıflandırma işlemi ve mavi, yeşil, kırmızı bantlarla yapılan sınıflandırma karşılaştırılmıştır.

4.2.Sınıflandırma İşlemleri

4.2.1 Destek Vektör Makineleri İle Sınıflandırma İşlemlerinin Analizi

150 adet yeşil alan pikselinin 134 adedi destek vektör makineleri algoritmasıyla doğru şekilde sınıflandırılırken 16 adet test noktası orman pikselleriyle karıştırılmıştır. 150 adet orman pikselinin 144 ü doğru şekilde sınıflandırılırken, 6 adedi yeşil alan pikselleriyle karışmıştır. 150 adet endüstri pikselinin 139 u doğru sınıflandırılırken, 10 adedi yakın yansıtma değeri sebebiyle yol pikselleriyle ve 1 i şehir alanları ile karışmıştır. 150 adet yol pikselinin 149 u doğru şekilde sınıflandırılırken, 1 adedi endüstri pikselleriyle karışmıştır. 150 adet toprak pikselinin 132 si doğru şekilde sınıflandırılırken, 2 adedi yol pikselleriyle, 16 adedi şehir alanı pikselleriyle karışmıştır. 150 adet şehir alanı pikselinin 146 sı doğru şekilde sınıflandırılırken, 2 adedi yol pikselleriyle, 1 adedi toprak pikselleriyle, 5 adedi endüstri pikselleriyle karışmıştır. 150 adet su pikselinin tamamı doğru şekilde sınıflandırılmıştır.



Şekil 4.3: Destek Vektör Makineleri Sınıflandırma Sonuç Görüntüsü

Çizelge 3.1: Destek Vektör Makineleri Sınıflandırıcısının test verilerinin sınıflara göre gösterimi.

DVM(RGB+NIR)	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Yeşil alan	134	6	0	0	0	0	0
Orman	16	144	0	0	0	0	0
Endüstri	0	0	139	1	0	1	0
Yol	0	0	10	149	2	2	0
Toprak	0	0	0	0	132	1	0
Şehir Alanı	0	0	1	0	16	146	0
Su	0	0	0	0	0	0	150
DVM(RGB)	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Yeşil alan	135	8	0	8	0	0	0
Orman	15	142	0	0	0	0	0
Endüstri	0	0	135	0	0	1	0
Yol	0	0	15	141	17	56	0
Toprak	0	0	0	0	121	25	0
Şehir Alanı	0	0	0	1	12	68	0
Su	0	0	0	0	0	0	150

Çizelge 3.2: Destek Vektör Makineleri Sınıflandırıcısının yakın kızıl ötesi bant dahil edilmiş durum için sınıf bazında doğruluk oranları

DVM(NIR+RG B)	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Duyarlılık	0.8933	0.9600	0.9267	0.9933	0.8800	0.9733	1.0000
Özgüllük	0.9933	0.9822	0.9978	0.9844	0.9989	0.9811	1.0000
Pos Ön Değeri	0.9571	0.9000	0.9858	0.9141	0.9925	0.8957	1.0000
Negatif Ön Değer	0.9824	0.9933	0.9879	0.9989	0.9804	0.9955	1.0000
Prevalans	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429
Algılama hızı	0.1276	0.1371	0.1324	0.1419	0.1257	0.1390	0.1429
Algılama Yaygınlığı	0.1333	0.1524	0.1343	0.1552	0.1267	0.1552	0.1429
Dengeli Doğruluk	0.9433	0.9711	0.9622	0.9889	0.9394	0.9772	1.0000

4.2.2. Karar Ağaçları Yöntemi İle Sınıflandırma İşlemleri

150 adet yeşil alan pikselinin 135 adedi destek vektör makineleri algoritmasıyla doğru şekilde sınıflandırılırken 15 adet test noktası orman pikselleriyle karıştırılmıştır. 150 adet orman pikselinin 142 adedi doğru şekilde sınıflandırılırken, 8 adedi yeşil alan pikselleriyle karışmıştır. 150 adet endüstri pikselinin 135 adedi doğru sınıflandırılırken, 15 adedi yakın yansıtma değeri sebebiyle yol pikselleriyle karışmıştır. 150 adet yol pikselinin 141 adedi doğru şekilde sınıflandırılırken, 8 adedi yeşil alan pikselleriyle, 1 adedi şehir alan pikselleriyle karışmıştır. 150 adet toprak pikselinin 121 adedi doğru şekilde sınıflandırılırken, 17 adedi yol pikselleriyle, 12 adedi şehir alanı pikselleriyle karışmıştır. 150 adet şehir alanı pikselinin 68 adedi

dođru Őekilde sınıflandırılırken, 56 adedi yol pikselleriyle, 25 adedi toprak pikselleriyle, 1 adedi endüstri pikselleriyle karışımıřtır. 150 adet su pikselinin tamamı dođru Őekilde sınıflandırılmıřtır.



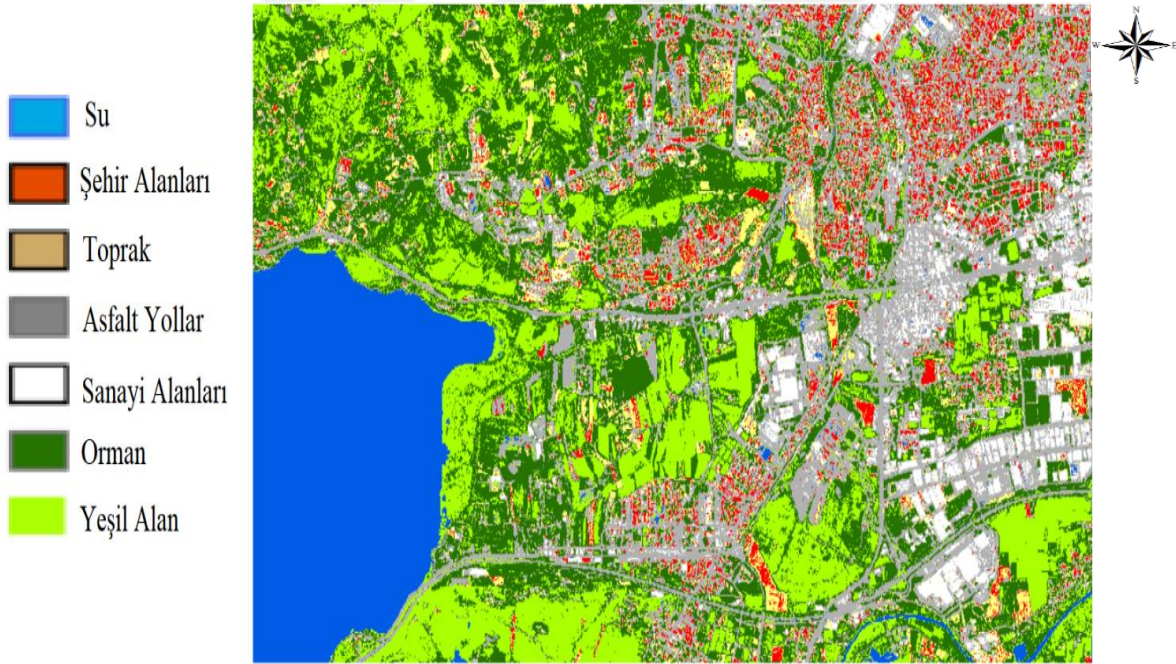
Őekil 4.4: Karar Ađaçları İle Yöntemi İle Sınıflandırma Sonu Görüntüsü

Çizelge 3.3: Karar Ağaçları Sınıflandırıcısının test verilerinin sınıflara göre gösterimi.

Karar Ağaçları Sınıflandırıcısı (RGB+NIR)	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Duyarlılık	0.9000	0.9467	0.9000	0.9400	0.8067	0.45333	1.0000
Özgüllük	0.9822	0.9833	0.9989	0.9022	0.9722	0.98556	1.0000
Pos Ön Değeri	0.8940	0.9045	0.9926	0.6157	0.8288	0.83951	1.0000
Negatif Ön Değer	0.9833	0.9910	0.9836	0.9890	0.9679	0.91538	1.0000
Prevalans	0.1429	0.1429	0.1429	0.1429	0.1429	0.14286	0.1429
Algılama hızı	0.1286	0.1352	0.1286	0.1343	0.1152	0.06476	0.1429
Detection Prevalans	0.1438	0.1495	0.1295	0.2181	0.1390	0.07714	0.1429
Dengeli Doğruluk	0.9411	0.9650	0.9494	0.9211	0.8894	0.71944	1.0000
Karar Ağaçları Sınıflandırıcısı (RGB)	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Yeşil alan	135	8	0	8	0	0	0
Orman	15	142	0	0	0	0	15
Endüstri	0	0	135	0	0	1	0
Yol	0	0	15	136	9	64	0
Toprak	0	0	0	6	104	12	0
Şehir Alanı	0	0	0	0	37	73	0
Su	0	0	0	0	0	0	135

4.2.3 Rastgele Orman Yöntemi İle Sınıflandırma İşleminin analizi

50 adet yeşil alan pikselinin 138 adedi Rastgele orman algoritmasıyla doğru şekilde sınıflandırılırken 12 adet test noktası orman pikselleriyle karıştırılmıştır. 150 adet orman pikselinin 137 adedi doğru şekilde sınıflandırılırken, 9 adedi yeşil alan pikselleriyle, 4 adedi su pikselleriyle karışmıştır. 150 adet endüstri pikselinin 141 adedi doğru sınıflandırılırken, 8 adedi yol pikselleriyle, 1 adedi toprak pikselleriyle karışmıştır. 150 adet yol pikselinin 145 adedi doğru şekilde sınıflandırılırken, 1 adedi endüstri pikselleriyle, 2 adedi toprak pikselleriyle, 2 adedi yeşil alan pikselleriyle karışmıştır. 150 adet toprak pikselinin 134 adedi doğru şekilde sınıflandırılırken, 16 adedi şehir alanı pikselleriyle karışmıştır. 150 adet şehir alanı pikselinin 98 adedi doğru şekilde sınıflandırılırken, 31 adedi yol pikselleriyle, 16 adedi toprak pikselleriyle, 5 adedi endüstri pikselleriyle karışmıştır. 150 adet su pikselinin 142 adedi doğru şekilde sınıflandırılırken 8 adedi orman pikselleriyle karışmıştır.



Şekil 4.5: Rastgele Orman Yöntemi İle Sınıflandırma Sonuç Görüntüsü

Çizelge 3.4: Rastgele Orman Sınıflandırıcısının test verilerinin sınıflara göre gösterimi.

Rasgele Orman (RGB+NIR)	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Yeşil alan	136	8	0	0	0	0	0
Orman	14	142	0	0	0	0	0
Endüstri	0	0	144	0	2	3	0
Yol	0	0	6	149	7	2	0
Toprak	0	0	0	1	138	7	0
Şehir Alanı	0	0	0	0	3	138	0
Su	0	0	0	0	0	0	150
Rasgele Orman (RGB)	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Yeşil alan	138	9	0	2	0	0	0
Orman	12	137	0	0	0	0	8
Endüstri	0	0	141	1	0	1	0
Yol	0	0	8	145	0	2	0
Toprak	0	0	1	2	134	3	0
Şehir Alanı	0	0	0	0	16	144	0
Su	0	4	0	0	0	0	142

Çizelge 3.2: Rastgele Orman Sınıflandırıcısının yakın kızıl ötesi bant dahil edilmiş durum için sınıf bazında doğruluk oranları

Değer	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Duyarlılık	0.9067	0.9467	0.9600	0.9933	0.9200	0.9200	1.0000
Özgüllük	0.9911	0.9844	0.9944	0.9833	0.9911	0.9967	1.0000
Pos Ön Değeri	0.9444	0.9103	0.9664	0.9085	0.9452	0.9787	1.0000
Negatif Ön Değer	0.9845	0.9911	0.9933	0.9989	0.9867	0.9868	1.0000
Prevalans	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429
Algılama hızı	0.1295	0.1352	0.1371	0.1419	0.1314	0.1314	0.1429
Detection Prevalans	0.1371	0.1486	0.1419	0.1562	0.1390	0.1343	0.1429
Dengeli Doğruluk	0.9489	0.9656	0.9772	0.9883	0.9556	0.9583	1.0000

4.2.4 Xgboost Yöntemi İle İle Sınıflandırma İşlemlerinin Analizi

150 adet yeşil alan pikselinin 135 adedi XCBOOST hızlandırma algoritmasıyla doğru şekilde sınıflandırılırken 15 adet test noktası orman pikselleriyle karıştırılmıştır. 150 adet orman pikselinin 143 adedi doğru şekilde sınıflandırılırken, 7 adedi yeşil alan pikselleriyle karışmıştır. 150 adet endüstri pikselinin 145 adedi doğru sınıflandırılırken, 5 adedi yakın yansıtma değeri sebebiyle yol pikselleriyle karışmıştır. 150 adet yol pikselinin 148 adedi doğru şekilde sınıflandırılırken, 1 adedi endüstri pikselleriyle, 1 adedi toprak pikselleriyle karışmıştır. 150 adet toprak pikselinin 137 adedi doğru şekilde sınıflandırılırken, 4 adedi yol pikselleriyle, 9 adedi şehir alanı pikselleriyle karışmıştır. 150 adet şehir alanı pikselinin 142 adedi doğru şekilde sınıflandırılırken, 3 adedi yol pikselleriyle, 2 adedi toprak pikselleriyle, 3 adedi endüstri pikselleriyle karışmıştır. 150 adet su pikselinin tamamı doğru şekilde sınıflandırılmıştır.



Şekil 4.6 : Xgboost Yöntemi İle İle Sınıflandırma Sonuç Görüntüsü

Çizelge 3.6: XGBoost Sınıflandırıcısının test verilerinin sınıflara göre gösterimi.

XGBoost(NIR)	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Yeşil alan	137	11	0	0	0	0	0
Orman	13	139	0	0	0	0	0
Endüstri	0	0	145	3	5	2	0
Yol	0	0	5	147	1	3	0
Toprak	0	0	0	0	131	5	0
Şehir Alanı	0	0	0	0	13	140	0
Su	0	0	0	0	0	0	150

XGBoost	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Yeşil alan	136	9	0	0	0	0	0
Orman	14	138	0	0	0	0	5
Endüstri	0	0	142	1	0	2	0
Yol	0	0	8	148	0	2	0
Toprak	0	0	0	1	144	6	0
Şehir Alanı	0	0	0	0	6	140	0
Su	0	3	0	0	0	0	145

Çizelge 3.7 : XGBoost Sınıflandırıcısının yakın kızıl ötesi bant dahil edilmiş durum için sınıf bazında doğruluk oranları

XGBoost	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Duyarlılık	0.9000	0.9533	0.9667	0.9867	0.9133	0.9467	1.0000
Özgüllük	0.9922	0.9833	0.9956	0.9867	0.9967	0.9900	1.0000
Pos Ön Değeri	0.9507	0.9051	0.9732	0.9250	0.9786	0.9404	1.0000
Negatif Ön Değer	0.9835	0.9922	0.9945	0.9978	0.9857	0.9911	1.0000
Prevalans	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429
Algılama hızı	0.1286	0.1362	0.1381	0.1410	0.1305	0.1352	0.1429
Detection Prevalans	0.1352	0.1505	0.1419	0.1524	0.1333	0.1438	0.1429
Dengeli Doğruluk	0.9461	0.9683	0.9811	0.9867	0.9550	0.9683	1.0000

4.2.5 K-En Yakın Komşuluk Yöntemi İle Sınıflandırma İşlemi

150 adet yeşil alan pikselinin 137 adedi K en yakın komşuluk algoritmasıyla doğru şekilde sınıflandırılırken 13 adedi test noktası orman pikselleriyle karıştırılmıştır. 150 adet orman pikselinin 139 adedi doğru şekilde sınıflandırılırken, 11 adedi yeşil alan pikselleriyle karışmıştır. 150 adet endüstri pikselinin 145 adedi doğru sınıflandırılırken, 5 adedi yakın yansıtma değeri sebebiyle yol pikselleriyle karışmıştır. 150 adet yol pikselinin 147 adedi doğru şekilde sınıflandırılırken, 3 adedi endüstri pikselleriyle karışmıştır. 150 adet toprak pikselinin 131 adedi doğru şekilde sınıflandırılırken, 1 adedi yol pikselleriyle, 13 adedi şehir alanı pikselleriyle, 5 adedi endüstri pikselleriyle karışmıştır. 150 adet şehir alanı pikselinin 140 adedi doğru şekilde sınıflandırılırken, 3 adedi yol pikselleriyle, 5 adedi toprak pikselleriyle, 2 adedi endüstri pikselleriyle karışmıştır. 150 adet su pikselinin tamamı doğru şekilde sınıflandırılmıştır.



Şekil 4.7 : K-En Yakın Komşuluk Yöntemi ile sınıflandırma Sonuç Görüntüsü

Çizelge 3.8: K-En Yakın Komşuluk Sınıflandırıcısının test verilerinin sınıflara göre gösterimi.

K-En Yakın Komşuluk(NIR)	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Yeşil alan	137	11	0	0	0	0	0
Orman	13	139	0	0	0	0	0
Endüstri	0	0	145	3	5	2	0
Yol	0	0	5	147	1	3	0
Toprak	0	0	0	0	131	5	0
Şehir Alanı	0	0	0	0	13	140	0
Su	0	0	0	0	0	0	150
K-En Yakın Komşuluk	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Yeşil alan	136	9	0	0	0	0	0
Orman	14	138	0	0	0	0	5
Endüstri	0	0	142	1	0	2	0
Yol	0	0	8	148	0	2	0
Toprak	0	0	0	1	144	6	0
Şehir Alanı	0	0	0	0	6	140	0
Su	0	3	0	0	0	0	145

Çizelge 3.9 : K-En Yakın Komşuluk Sınıflandırıcısının yakın kızıl ötesi bant dahil edilmiş durum için sınıf bazında doğruluk oranları

Değer	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Duyarlılık	0.9133	0.9267	0.9667	0.9800	0.8733	0.9333	1.0000
Özgüllük	0.9878	0.9856	0.9889	0.9900	0.9944	0.9856	1.0000
Pos Ön Değeri	0.9257	0.9145	0.9355	0.9423	0.9632	0.9150	1.0000
Negatif Ön Değer	0.9856	0.9878	0.9944	0.9966	0.9792	0.9889	1.0000
Prevalans	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429
Algılama hızı	0.1305	0.1324	0.1381	0.1400	0.1248	0.1333	0.1429
Detection Prevalans	0.1410	0.1448	0.1476	0.1486	0.1295	0.1457	0.1429
Dengeli Doğruluk	0.9506	0.9561	0.9778	0.9850	0.9339	0.9594	1.0000

4.2.6 Naive Bayes İle Sınıflandırma İşlemi

150 adet yeşil alan pikselinin 133 adedi Naive bayes algoritmasıyla doğru şekilde sınıflandırılırken 17 adedi test noktası orman pikselleriyle karıştırılmıştır. 150 adet orman pikselinin 141 adedi doğru şekilde sınıflandırılırken, 9 adedi yeşil alan pikselleriyle karışmıştır. 150 adet endüstri pikselinin 139 adedi doğru sınıflandırılırken, 11 adedi yakın yansıtma değeri sebebiyle yol pikselleriyle karışmıştır. 150 adet yol pikselinin 139 adedi doğru şekilde sınıflandırılırken, 1 adedi endüstri pikselleriyle, 1 adedi toprak pikselleriyle, 8 adedi şehir alanı pikselleriyle karışmıştır. 150 adet toprak pikselinin 124 adedi doğru şekilde sınıflandırılırken, 1 adedi yol pikselleriyle, 25 adedi şehir alanı pikselleriyle karışmıştır. 150 adet şehir alanı pikselinin 98 adedi doğru şekilde sınıflandırılırken, 31 adedi yol pikselleriyle, 16 adedi toprak pikselleriyle, 5 adedi endüstri pikselleriyle karışmıştır. 150 adet su pikselinin tamamı doğru şekilde sınıflandırılmıştır.



Şekil 4.8 : Naive Bayes Algoritması ile Sınıflandırma Sonuç Görüntüsü

Çizelge 3.10: Naive Bayes Komşuluk Sınıflandırıcısının test verilerinin sınıflara göre gösterimi.

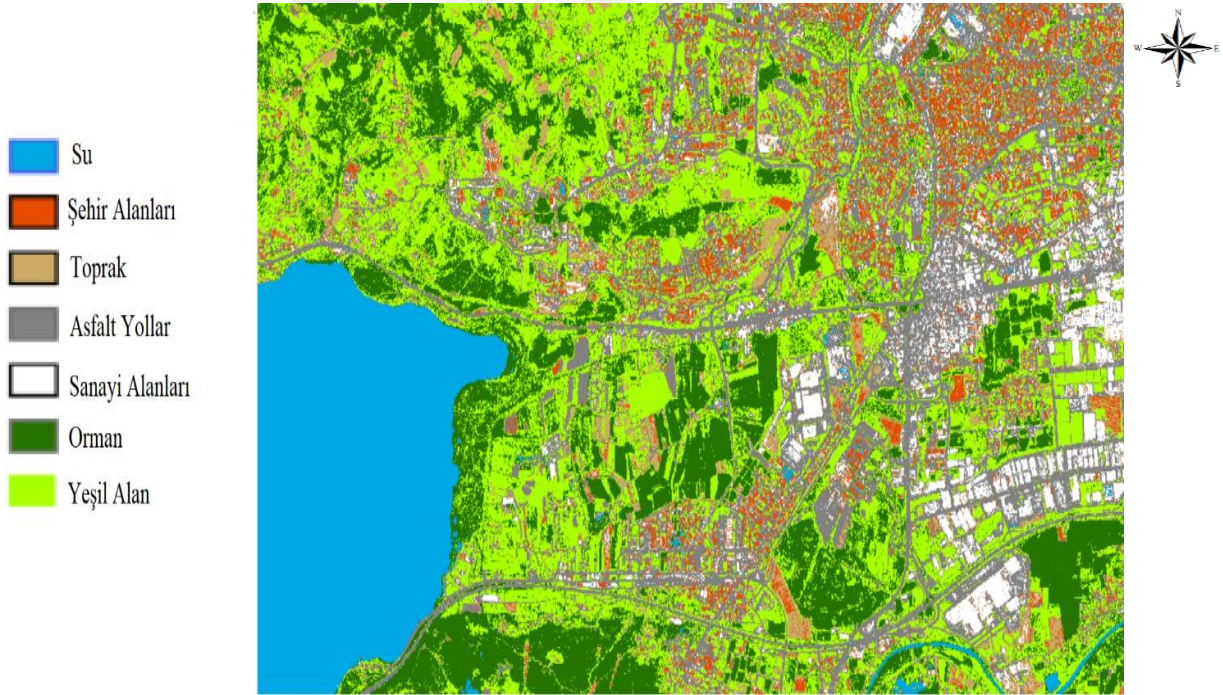
Naive Bayes(NIR+RGB)	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Yeşil alan	133	9	0	0	0	0	0
Orman	17	141	0	0	0	0	0
Endüstri	0	0	139	1	0	5	0
Yol	0	0	11	139	1	31	0
Toprak	0	0	0	2	124	16	0
Şehir Alanı	0	0	0	8	25	98	0
Su	0	0	0	0	0	0	150
Naive Bayes(RGB)	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Yeşil alan	133	10	0	8	2	0	2
Orman	17	140	0	0	0	0	14
Endüstri	0	0	139	1	0	5	0
Yol	0	0	11	100	0	42	0
Toprak	0	0	0	39	117	17	0
Şehir Alanı	0	0	0	2	31	86	0
Su	0	0	0	0	0	0	134

Çizelge 3.11: Naive Bayes Komşuluk Sınıflandırıcısının yakın kıvılcık ötesi bant dahil edilmiş durum için sınıf bazında doğruluk oranları

Naive Bayes	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Duyarlılık	0.8867	0.9400	0.9267	0.9267	0.8267	0.65333	1.0000
Özgüllük	0.9900	0.9811	0.9933	0.9522	0.9800	0.96333	1.0000
Pos Ön Değeri	0.9366	0.8924	0.9586	0.7637	0.8732	0.74809	1.0000
Negatif Ön Değer	0.9813	0.9899	0.9878	0.9873	0.9714	0.94342	1.0000
Prevalans	0.1429	0.1429	0.1429	0.1429	0.1429	0.14286	0.1429
Algılama hızı	0.1267	0.1343	0.1324	0.1324	0.1181	0.09333	0.1429
Detection Prevalans	0.1352	0.1505	0.1381	0.1733	0.1352	0.12476	0.1429
Dengeli Doğruluk	0.9383	0.9606	0.9600	0.9394	0.9033	0.80833	1.0000

4.2.7 Adaboost Yöntemi İle Sınıflandırma İşlemlerinin Analizi ve Sonuçları

150 adet yeşil alan pikselinin 135 adedi Adaboost hızlandırma algoritmasıyla doğru şekilde sınıflandırılırken 15 adedi test noktası orman pikselleriyle karıştırılmıştır. 150 adet orman pikselinin 141 adedi doğru şekilde sınıflandırılırken, 9 adedi yeşil alan pikselleriyle karışmıştır. 150 adet endüstri pikselinin 142 adedi doğru sınıflandırılırken, 7 adedi yakın yansıtma değeri sebebiyle yol pikselleriyle, 1 adedi toprak pikselleriyle karışmıştır. 150 adet yol pikselinin 148 adedi doğru şekilde sınıflandırılırken, 1 adedi endüstri pikselleriyle, 1 adedi şehir alanı pikselleriyle karışmıştır. 150 adet toprak pikselinin 138 adedi doğru şekilde sınıflandırılırken, 4 adedi yol pikselleriyle, 4 adedi şehir alanı pikselleriyle, 4 adedi toprak pikselleriyle karışmıştır. 150 adet şehir alanı pikselinin 141 adedi doğru şekilde sınıflandırılırken, 2 adedi yol pikselleriyle, 4 adedi toprak pikselleriyle, 3 adedi endüstri pikselleriyle karışmıştır. 150 adet su pikselinin tamamı doğru şekilde sınıflandırılmıştır.



Şekil 4.9 : Adaboost Yöntemi ile sınıflandırma Sonuç Görüntüsü

Çizelge 3.12: Naive Bayes Komşuluk Sınıflandırıcısının test verilerinin gösterimi

Adaboost(NIR+RGB)	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Yeşil alan	135	9	0	0	0	0	0
Orman	15	141	0	0	0	0	0
Endüstri	0	0	142	1	4	3	0
Yol	0	0	7	148	4	2	0
Toprak	0	0	1	0	138	4	0
Şehir Alanı	0	0	0	1	4	141	0
Su	0	0	0	0	0	0	150
Adaboost(RGB)	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Yeşil alan	133	8	0	0	0	0	0
Orman	17	136	0	0	0	0	3
Endüstri	0	0	138	2	0	1	0
Yol	0	0	9	147	0	3	0
Toprak	0	0	2	0	141	3	0
Şehir Alanı	0	0	1	1	9	143	0
Su	0	6	0	0	0	0	147

Çizelge 3.12: Adaboost Sınıflandırıcısının yakın kızıl ötesi bant dahil edilmiş durum için sınıf bazında doğruluk oranları

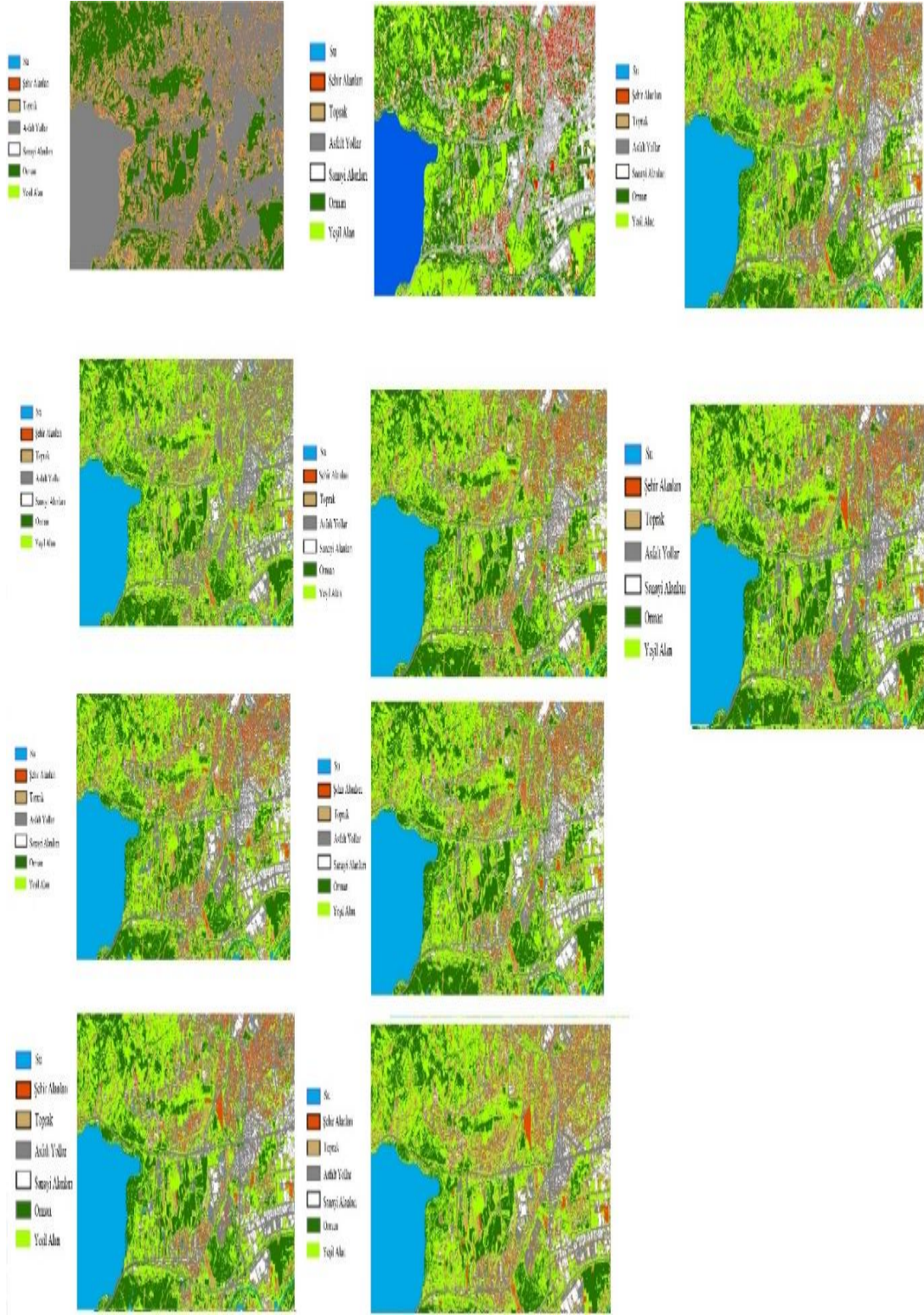
Adaboost(NIR)	Yeşil alan	Orman	Endüstri	Yol	Toprak	Şehir Alanı	Su
Duyarlılık	0.9000	0.9400	0.9467	0.9867	0.9200	0.9400	1.0000
Özgüllük	0.9900	0.9833	0.9911	0.9856	0.9944	0.9944	1.0000
Pos Ön Değeri	0.9375	0.9038	0.9467	0.9193	0.9650	0.9658	1.0000
Negatif Ön Değer	0.9834	0.9899	0.9911	0.9978	0.9868	0.9900	1.0000
Prevalans	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429
Algılama hızı	0.1286	0.1343	0.1352	0.1410	0.1314	0.1343	0.1429
Detection Prevalans	0.1371	0.1486	0.1429	0.1533	0.1362	0.1390	0.1429
Dengeli Doğruluk	0.9450	0.9617	0.9689	0.9861	0.9572	0.9672	1.0000

5. SONUÇ VE ÖNERİLER

Sınıflandırma Sonuçlarının Analiz edildiğinde Destek Vektör Makineleri, Karar Ağaçları, Rastgele Orman, XGboost, K En Yakın Komşuluk, Naive Bayes, Adaboost ve Yapay Sinir Ağları algoritmalarının tümüne ait hata matrisleri ve sonuç görüntüleri incelenmiştir. Su piksellerinin diğer sınıflara göre yüksek yansıtma değerinden kaynaklı olarak tüm algoritmalar tarafından tam doğru olarak ayırt edilmiştir.

Çizelge 4.1 : Kullanılan tüm algoritmaların sonuç değerleri tablosu

Algoritma (NIR+RGB)	Kappa değeri	Güven aralığı	Doğruluk
DVM	0.9378	95% (0.9313, 0.9595)	0.9467
KA	0.8244	95% (0.8264, 0.8706)	0.8495
RO	0.9411	95% (0.9345, 0.962)	0.9495
XGBOOST	0.9444	95% (0.9377, 0.9645)	0.9524
KNN	0.9322	95% (0.926, 0.9553)	0.9419
NAİVE BAYES	0.86	95% (0.8588, 0.899)	0.88
ADABOOST	0.9389	95% (0.9324, 0.9603)	0.9476



Şekil 4.10 : Tüm Algoritmaların sonuç görüntülerinin karşılıklı gösterimi

Destek Vektör Makineleri algoritması; en eski yöntemlerden biri olmasının yanı sıra neredeyse hiçbir sınıfta yüksek hata oranı göstermemiş olup ayrıca genel olarak tutarlı bir sınıflandırma sonucu vermiştir. Destek Vektör Makineleri algoritması, en yüksek yol piksellerini %99 doğrulukla, en düşük toprak piksellerini %87 doğrulukla tespit etmiştir. Ayrıca birbirine en yakın yansıtma değerleri arasında bir hiper düzlem oluşturarak orman piksellerini, yol piksellerini ve şehir alanı piksellerini en doğru şekilde ayırt etmiştir. Yakın kızıl ötesi bandının sınıflandırmaya etkisi incelendiğinde 2 yeşil alan pikselinin, 1 orman pikselinin 8 yol pikselinin 12 yol pikselinin ve 11 şehir pikselinin doğru şekilde sınıflandırılmasını sağladığı tespit edilmiştir. Yakın kızıl ötesi bandın kullanılması belirgin derecede sınıflandırmada doğruluğu arttırmıştır.

Karar Ağaçları algoritması; sınıflandırma analizinde destek vektör makineleri algoritması ile karşılaştırıldığında şehir alanlarını oldukça hatalı biçimde sınıflandırmıştır. Bu noktada karar ağaçları sınıflandırmasında şehir alanlarının farklı radyometrik yansıtma değerlerinden kaynaklı olarak çok küçük gruplara ayrılması sonucu şehir alanları çoğunlukla toprak pikselleriyle karışmıştır. Karar Ağaçları algoritması, en yüksek orman piksellerini %94 doğrulukla, en düşük şehir alanları piksellerini %45 doğrulukla tespit etmiştir.

Destek vektör makineleri algoritması ise ortalama piksel değerlerini kümeleyerek daha tutarlı sonuçlara ulaşmıştır. Yakın kızıl ötesi bandının sınıflandırmaya etkisi incelendiğinde test verilerinde 5 yol pikselini 16 toprak pikselinin ve 11 şehir pikselinin doğru şekilde sınıflandırılmasını sağladığı tespit edilmiştir.

Rastgele Orman sınıflandırma algoritması genel olarak tüm sınıfları tutarlı şekilde belirlemiştir. En kötü sınıflandırma sonucunun şehir alanlarının ayrılmasında göstermiştir. Yakın kızıl ötesi bandının sınıflandırmaya etkisi incelendiğinde test verilerinde 2 yeşil alan pikselinin, 5 orman pikselinin 3 endüstri pikselinin 4 yol ve toprak pikselinin 6 adet şehir alanı pikselinin ve 8 adet su pikselini doğru sınıflandırmayı sağladığı görülmüştür.

XGboost algoritması; genel olarak tüm sınıfları tutarlı biçimde ayırmıştır. XGboost hızlandırma algoritması, karar ağaçları algoritmasının zayıf olduğu alanlarda genel performansı geliştiren birleştirme işlemi sayesinde özellikle dikkat çekici şekilde şehir alanı piksellerinin sınıflandırılmasında ön plana çıkmıştır. XGBOOST algoritması, en

yüksek yol piksellerini %98 doğrulukla, en düşük yeşil alan piksellerini %89 doğrulukla tespit etmiştir. Ayrıca endüstri piksellerini K en yakın komşuluk yöntemiyle beraber aynı oranda en doğru şekilde tespit eden algoritmalarıdır. Yakın kızıl ötesi bandının sınıflandırmaya etkisi incelendiğinde test verilerinde 2 yeşil alan pikselinin, 1 orman pikselinin 8 yol pikselinin 12 yol pikselinin ve 11 şehir pikselinin doğru şekilde sınıflandırılmasını sağladığı tespit edilmiştir.

K En Yakın Komşuluk algoritmasının; hata matrisine bakıldığında genel olarak pikselleri doğru şekilde kümelere ayırarak homojen şekilde sınıflandırdığı tespit edilmiştir. Buna ek olarak tarım alanlarını en doğru biçimde analiz etmiştir. K En Yakın Komşuluk algoritması, en yüksek yol piksellerini %97 doğrulukla, en düşük toprak piksellerini %87 doğrulukla tespit etmiştir. Diğer algoritmalarda olduğu gibi yeşil alan ve orman piksellerini birbirinden ayırt etmede hatalı sonuç vermiştir. Ayrıca diğer algoritmaların aksine toprak piksellerini diğer sınıflarla daha fazla karıştırmıştır. Yakın kızıl ötesi bandının sınıflandırmaya etkisi incelendiğinde test verilerinde 1 yeşil alan pikselinin, 1 orman pikselinin 3 endüstri pikselinin 1 yol pikselinin ve 9 toprak pikselini ve 5 su pikselinin doğru şekilde sınıflandırılmasını sağladığı tespit edilmiştir.

Naive Bayes algoritması; şehir alanı piksellerinde diğer algoritmalara göre kötü bir sonuç vermiştir. Bunun başlıca sebebi olarak her pikselin yansıtma değerlerine göre olasılığı en yüksek olan sınıfa atamasından kaynaklıdır. Naive Bayes algoritması, en yüksek endüstri ve yol piksellerini %92 doğrulukla, en düşük şehir alanı piksellerini %65 doğrulukla tespit etmiştir.

Eğitim verileri 350 adet piksel değil daha az bir piksel sayısı içermiş olsaydı diğer algoritmalara göre daha doğru bir sonuç verebilirdi fakat eğitim verisinin artması karşıt algoritmaların daha etkin çalışmasını sağlamıştır. Yakın kızıl ötesi bandının sınıflandırmaya etkisi incelendiğinde test verilerinde, 1 orman pikselinin 39 yol pikselinin, 7 toprak pikselinin, 12 şehir pikselinin ve 16 su pikselini doğru şekilde sınıflandırılmasını sağladığı tespit edilmiştir.

Adaboost hızlandırma algoritması; XGboost algoritmasında olduğu gibi karar ağaçları algoritmasının zayıf kaldığı yerlerde birleştirme işlemi uygulayarak daha başarılı sonuçlar elde etmiştir. Adaboost algoritması, en yüksek yol piksellerini %98 doğrulukla, en düşük yeşil alan piksellerini %89 doğrulukla tespit etmiştir. Özellikle

toprak piksellerini en doğru ayırt eden algoritmadır. Yakın kızıl ötesi bandının sınıflandırmaya etkisi incelendiğinde test verilerinde, 2 yeşil alan pikselini, 5 orman pikselini, 4 endüstri pikselini, 1 yol pikselini, 3 toprak pikselini, 6 şehir pikselini ve 3 su pikselini doğru şekilde sınıflandırılmasını sağladığı tespit edilmiştir.

Çizelge 4.2 : Kullanılan tüm algoritmaların sınıflara göre sonuçları

Makine Öğrenme Algoritması(NIR+RGB)	Yeşil Alan (%)	ORMAN (%)	ENDÜSTRİ (%)	YOL (%)	TOPRAK (%)	ŞEHİR ALAN (%)	SU (%)
DVM	89,244	95,904	92,574	99,234	87,912	97,236	100
KA	89,91	94,572	89,91	93,906	80,586	45,288	100
RO	89,76	93,72	95,04	98,34	91,08	91,08	100
XGBOOST	89,91	95,238	96,57	98,568	91,242	94,572	100
KNN	91,242	92,574	96,57	97,902	87,246	93,24	100
NAİVE BAYES	88,578	87,246	92,574	92,574	82,584	65,268	100
YSA	0	0	0	0	0	0	100
ADABOOST	89,91	93,906	94,572	98,568	91,908	93,906	100

Uzaktan algılama teknolojisi çoklu disiplinlerde yoğun olarak kullanılmaktadır. Uzaktan algılama teknolojisinin gelişmesi veri elde etme yöntemlerinin çeşitliliğini ve elde edilen verinin kalitesini arttırmış aynı zamanda sınıflandırma yöntemlerinde de ilerlemeyi beraberinde getirmiştir. Günümüzde algılayıcı teknolojilerinin geldiği noktada elektromanyetik spektrumun farklı bölgelerinde çalışma yapılabilmektedir. Çalışmada kullanılan Spot 7 uydu görüntüleri çalışma alanını sınıflandırmada ve arazi kullanımının belirlenmesinde görüntüleri ön işleme tabii tutulmasının ardından yeterli veriyi sağlamıştır. Elde edilmiş olan ham veri ancak doğru sınıflandırma yönteminin kullanılması ile bir sonuç ifade etmiştir. Bu çalışmada olduğu gibi diğer çalışmalarda da sınıflandırma işleminde çıkacak sonuç çalışmayı yapan kişilerin seçmiş olduğu yönteme göre şekillenmektedir.

Eğitim ve test verilerinin atanmasının ardından algoritmaların Destek Vektör Makineleri, Karar Ağaçları, Rastgele Orman, XGboost, K En Yakın Komşuluk, Naive

Bayes, Adaboost ve Yapay Sinir Ağları algoritmalarının hepsi eşit sayıda ve aynı veriler üzerinden koşulmuştur. Sınıflandırma sonuçlarının incelenmesinin ardından en yüksek doğrulukla sınıflandırma sonucunu veren algoritma diğer algoritmaların zayıf olan kısımlarını egale edip aynı zamanda temel işlevlerinden yararlanan XGboost algoritması en doğru sonucu vermiştir. Benzer şekilde çalışan Adaboost hızlandırma algoritması ikinci en iyi sonucu vermiştir. Bu algoritmaya yakın sonuç vererek üçüncü en doğru sonucu çıkaran uzun zamandır literatürde bilinen ve kullanılan Destek Vektör Makineleri algoritmasıdır. Kümeleme yöntemi ile çalışan K en yakın komşuluk yöntemi genel doğruluğa bakıldığında dördüncü en doğru sonucu vermiş olan algoritmadır. Naive Bayes algoritması özellikle şehir alanlarını diğer algoritmalarından daha yanlış sınıflandırdığı için beşinci doğru sonucu vermiştir. Karar Ağaçları algoritması Naive Bayes algoritmasına benzer şekilde şehir alanlarını oldukça hatalı biçimde sınıflandırmıştır. Rastgele Orman algoritması zayıf yönleri itibariyle XGboost, Adaboost ve diğer algoritmaların gerisinde kalmıştır. Yapay Sinir Ağları algoritması ise genel doğruluk itibari ile en kötü doğruluk sınıflandırma sonucunu vermiştir.

Destek vektör makineleri algoritmasının orman piksellerini yeşil alanlardan ayırmada ve Adaboost algoritması toprak piksellerini en iyi şekilde ayırmasının sebebi klorofile hassas olan yakın kızılötesi bandın sınıflandırmasında oldukça başarılı olduğunu göstermektedir. Ayrıca tüm algoritmaların su piksellerini hatasız ayırt etmiş olması çalışmada en iyi sınıflandırılan piksel sınıfı olduğunu ortaya çıkarmıştır. Özellikle Naive Bayes ve Karar ağaçları yöntemlerinin yapay obeleri toprak piksellerinden ayırmada diğer tekniklere göre daha başarısız olması bu tekniklerin kırmızı bandı daha düşük doğrulukla sınıflandırdığını ortaya çıkarmaktadır.

Sonuçlar irdelendiğinde açıkça yakın kızılötesi bandınının su piksellerinin ve tarım alanları ile orman sınıfı arasında ki ayrımın makine öğrenmesi algoritmaları ile sınıflandırmasında doğruluğu kesin biçimde arttırdığı ayrıca arttırdığı tespit edilmiştir. Elde edilen sonuçlar aynı zamanda arazi kullanımının sınıflandırılmasında çalışma yapılacak bölgeye ait uzaktan algılama verilerinin kalitesinin yanında sınıflandırma için seçilecek olan algoritmanın da sonuçları doğrudan etkilediği açıkça göstermiştir. Eğer tarım ve ormancılık amaçlı bir sınıflandırma amaçlanıyorsa yakın kızılötesi bandın ve güncel Adaboost algoritmasının tercih edilmesi olumlu sonuç verecektir.

İlerleyen teknoloji ile birlikte yeni sınıflandırma algoritmalarının geliştirilmesine baęlı olarak arazi kullanımının belirlenmesinde daha yüksek doęruluk elde edileceęi beklenmektedir.





KAYNAKLAR

- [1] “SPOT 7 - Earth Online”.
Url-1<<https://earth.esa.int/eogateway/missions/spot-7> >, (erişim Şub. 16, 2021).
- [2] “Optical and Radar Data”. **Url-2**<<https://www.intelligence-airbusds.com/optical-and-radar-data/> >, (erişim Şub. 09, 2021).
- [3] “Uzaktan Algılama”. **Url-3**<https://cdn-acikogretim.istanbul.edu.tr/auzefcontent/20_21_Guz/uzaktan_algilama/1/index.html> (erişim Şub. 06, 2021).
- [4] “Elektromanyetik spektrum - Vikipedi”.
Url-4<https://tr.wikipedia.org/wiki/Elektromanyetik_spektrum>, (erişim Şub. 09, 2021).
- [5] “Hvkk - ÇÖZÜNÜRLÜK DEĞERİ”. **Url-5**<<https://www.hvkk.tsk.tr/Custom/Hvkk/244>>, (erişim Ara. 21, 2020).
- [6] “Resolution”.
Url-6<https://gsp.humboldt.edu/OLM/Courses/GSP_216_Online/lesson3-1/resolution.html >, (erişim Şub. 09, 2021).
- [7] **F. Wang**, “Fuzzy supervised classification of remote sensing images”, *IEEE Trans. Geosci. Remote Sens.*, c. 28, sy 2, ss. 194-201, Mar. 1990, doi: 10.1109/36.46698.
- [8] **D. Herries**, “Collaborative Forestry Remote Sensing Data Acquisition Campaign”, *Interpine Innovation*, Nis. 03, 2018.
Url-7<<https://interpine.nz/30908/>>, (erişim Şub. 09, 2021).
- [9] **S. Paulus ve A.-K. Mahlein**, “Technical workflows for hyperspectral plant image assessment and processing on the greenhouse and laboratory scale”, *GigaScience*, c. 9, sy 8, s. g1aa090, Ağu. 2020, doi: 10.1093/gigascience/g1aa090.
- [10] Beyaznet, “Derin Öğrenme (Deep Learning) Nedir?”
Url8<https://www.beyaz.net/tr/yazilim/makaleler/derin_ogrenme_deep_learning_nedir.html >, (erişim Şub. 11, 2021).
- [11] *Computational intelligence paradigms in advanced pattern classification*. Berlin: Springer, 2012, s. 179.
- [12] “R’da Karar Ağacı Üzerine Bir Vaka Çalışması: A Case Study on Decision Tree in R”, *Tevfik Bulut*, Ara. 12, 2019.
Url-9<<https://tevfikbulut.com/2019/12/12/karar-agaci-uzerine-bir-vaka-calismasi/> >, (erişim Şub. 11, 2021).
- [13] “Rassal Orman Algoritması”, *yavuz blog*, Eki. 01, 2018.
Url-10<<http://yavuz.github.io/rassal-orman/> >, (erişim Şub. 11, 2021).
- [14] **J. Bhattacharyya**, “Understanding XGBoost Algorithm In Detail”, *Analytics India Magazine*, Kas. 02, 2020.
Url-11<<https://analyticsindiamag.com/xgboost-internal-working-to-make-decision-trees-and-deduce-predictions/> >, (erişim Şub. 11, 2021).
- [15] “KNN (k-nearest neighbours, k-en yakın komşu) Algoritması”.

- Url-12**< <https://mlhtnc.github.io/knn-algorithm.html> >, (erişim Şub. 11, 2021).
- [16] “In Depth: Naive Bayes Classification | Python Data Science Handbook”.
Url-13<<https://jakevdp.github.io/PythonDataScienceHandbook/05.05-naive-bayes.html> >, (erişim Şub. 11, 2021).
- [17] G. World, “Land cover classification from Remote Sensing data”, *Geospatial World*, Ara. 09, 2010.
Url-14<<https://www.geospatialworld.net/article/land-cover-classification-from-remote-sensing-data/>>, (erişim Şub. 17, 2021).
- [18] “Support Vector Machine”.
Url-15<https://www.saedsayad.com/support_vector_machine.htm >, (erişim Şub. 17, 2021).
- [19] **R. Silipo**, “From a Single Decision Tree to a Random Forest”, *DATAVERSITY*, Ağu. 02, 2019. **Url-16**<<https://www.dataversity.net/from-a-single-decision-tree-to-a-random-forest/>>, (erişim Şub. 17, 2021).
- [20] “KNN Algorithm using Python | K Nearest Neighbors Algorithm”, *Edureka*, Tem. 26, 2018.
Url-17<<https://www.edureka.co/blog/k-nearest-neighbors-algorithm/> >, (erişim Şub. 17, 2021).
- [21] “Gradient Boosting Classifiers in Python with Scikit-Learn”, *Stack Abuse*.
Url-18<<https://stackabuse.com/gradient-boosting-classifiers-in-python-with-scikit-learn/>>(erişim Şub. 17, 2021).
- [22] **A. Veronica**, “Understanding Adaboost and Scikit-learn’s algorithm:”, *Medium*, May. 18, 2020.
Url-19<<https://medium.com/datadriveninvestor/understanding-adaboost-and-scikit-learns-algorithm-c8d8af5ace10> >, (erişim Şub. 17, 2021).
- [23] “Naive Bayes Classification using Scikit-learn”, *DataCamp Community*, Ara. 04, 2018.
Url-20<<https://www.datacamp.com/community/tutorials/naive-bayes-scikit-learn>>, (erişim Şub. 17, 2021).
- [24] **A. Rahman vd.**, “Performance of different machine learning algorithms on satellite image classification in rural and urban setup”, *Remote Sens. Appl. Soc. Environ.*, c. 20, s. 100410, Kas. 2020, doi: 10.1016/j.rsase.2020.100410.
- [25] T. T. Nguyen vd., “Monitoring agriculture areas with satellite images and deep learning”, *Appl. Soft Comput.*, c. 95, s. 106565, Eki. 2020, doi: 10.1016/j.asoc.2020.106565.
- [26] **Q. Zou, L. Ni, T. Zhang, ve Q. Wang**, “Deep Learning Based Feature Selection for Remote Sensing Scene Classification”, *IEEE Geosci. Remote Sens. Lett.*, c. 12, sy 11, ss. 2321-2325, Kas. 2015, doi: 10.1109/LGRS.2015.2475299.
- [27] **M. J. Khan, A. Yousaf, N. Javed, S. Nadeem, ve K. Khurshid**, “Automatic target detection in satellite images using deep learning”, *J. Space Technol.*, c. 7, sy 1, ss. 44-49, 2017.
- [28] **X. Chen, S. Xiang, C.-L. Liu, ve C.-H. Pan**, “Vehicle detection in satellite images by hybrid deep convolutional neural networks”, *IEEE Geosci. Remote Sens. Lett.*, c. 11, sy 10, ss. 1797-1801, 2014.
- [29] **U. Alganci, M. Soydas, ve E. Sertel**, “Comparative research on deep learning approaches for airplane detection from very high-resolution satellite images”, *Remote Sens.*, c. 12, sy 3, s. 458, 2020.

- [30] **W. A. Fischer, W. R. Hemphill, ve A. Kover**, “Progress in remote sensing (1972-1976)”, *Photogrammetria*, c. 32, sy 2. s. 40, 1976, doi: 10.1016/0031-8663(76)90013-2.
- [31] **G. K. Moore**, “What is a picture worth? A history of remote sensing”, *Hydrological Sciences Bulletin*, c. 24, sy 4. s. 9, 1979, doi: 10.1080/02626667909491887.
- [32] **L. L. F. Janssen**, *Principles of remote sensing: an introductory textbok*. Enschede: ITC, 2004.
- [33] **T. Kavzođlu ve İ. Çölkesen**, “UZAKTAN ALGILAMA TEKNOLOJİLERİ VE UYGULAMA ALANLARI”.
- [34] **M. Erdoğan ve H. Akdeniz**, “Uzaktan Algılama Amaçlı Uydu Sistemlerindeki Son Gelişmeler”, *Harita Derg.*, sy 132, ss. 11-25, 2004.
- [35] **N. Musaođlu**, “Elektro-optik ve mikrodalga algılayıcılardan elde edilen uydu verilerinden orman alanlarında meşcere tiplerinin ve yetiştirme ortamı birimlerinin belirlenme olanakları”, Thesis, Fen Bilimleri Enstitüsü, 1999.
- [36] “Electromagnetic Spectrum - Introduction”.
Url-21<<https://imagine.gsfc.nasa.gov/science/toolbox/emspectrum1.html> >, (erişim Oca. 31, 2021).
- [37] “electromagnetic spectrum | Definition, Diagram, & Uses | Britannica”.
Url-22< <https://www.britannica.com/science/electromagnetic-spectrum>>, (erişim Oca. 31, 2021).
- [38] “Computer Processing of Remotely-Sensed Images: An Introduction, 3rd Edition | Wiley”, *Wiley.com*. **Url-23**< <https://www.wiley.com/en-us/Computer+Processing+of+Remotely+Sensed+Images%3A+An+Introduction%2C+3rd+Edition-p-9780470021019>>, (erişim Şub. 02, 2021).
- [39] **M. Güven**, “Göktürk-2 uydu görüntüsünün geometrik doğruluğunun araştırılması”, 2016.
- [40] **B. ÖDEVİ ve B. KURT**, “UZAKTAN ALGILAMADA OPTİK UYDU GÖRÜNTÜLERİNİN BİRLEŞTİRİLMESİ”, 2014.
- [41] **R. Yılmaz**, “Türkçe dokümanların sınıflandırılması”, 2013.
- [42] **F. F. Sabins**, “Remote sensing for mineral exploration”, *Ore Geol. Rev.*, c. 14, sy 3-4, ss. 157-183, 1999.
- [43] **P. Mather ve B. Tso**, *Classification methods for remotely sensed data*. CRC press, 2016.
- [44] “Serdivan”. **Url-24**<<http://www.sakarya.gov.tr/serdivan#> (erişim Şub. 16, 2021)>.
- [45] “SPOT-6 and 7 - eoPortal Directory - Satellite Missions”.
Url-25< <https://eoportal.org/web/eoportal/satellite-missions/s/spot-6-7> (erişim Şub. 16, 2021)>.
- [46] **J. P. Mueller ve L. Massaron**, *Machine learning for dummies*. John Wiley & Sons, 2021.
- [47] “Advantages and Disadvantages of Machine Learning Language”, *DataFlair*, Şub. 06, 2018.
Url-26<<https://data-flair.training/blogs/advantages-and-disadvantages-of-machine-learning/>>(erişim Şub. 10, 2021).
- [48] “10 Companies Using Machine Learning in Cool Ways”.
Url-27<<https://www.wordstream.com/blog/ws/2017/07/28/machine-learning-applications> >, (erişim Şub. 10, 2021).
- [49] “Supervised vs Unsupervised Learning: Key Differences”.

- Url-28**< <https://www.guru99.com/supervised-vs-unsupervised-learning.html> >(erişim Şub. 10, 2021).
- [50] **A. Ng**, “CS229 Lecture notes”, *CS229 Lect. Notes*, c. 1, sy 1, ss. 1-3, 2000.
- [51] **Y. Wen, Y. Lu, ve P. Shi**, “Handwritten Bangla numeral recognition system and its application to postal automation”, *Pattern Recognit.*, c. 40, sy 1, ss. 99-107, 2007.
- [52] “Support Vector Machine (SVM) Algorithm - Javatpoint”, **Url-29**<www.javatpoint.com. <https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm> >, (erişim Şub. 10, 2021).
- [53] **A. User:Zirguezi** svg version by, *English: Kernel machines are used to compute non-linearly separable functions into a higher dimension linearly separable function*. 2011.
- [54] **T. Kavzoğlu ve İ. Çölkesen**, “Destek vektör makineleri ile uydu görüntülerinin sınıflandırılmasında kernel fonksiyonlarının etkilerinin incelenmesi”, *Harita Derg.*, c. 144, sy 7, ss. 73-82, 2010.
- [55] **E. E. Osuna**, “Support vector machines: Training and applications”, 1998.
- [56] “ML | Using SVM to perform classification on a non-linear dataset”, *GeeksforGeeks*, Oca. 15, 2019. **Url-30**<<https://www.geeksforgeeks.org/ml-using-svm-to-perform-classification-on-a-non-linear-dataset/>>, (erişim Şub. 18, 2021).
- [57] **İ. Çölkesen**, “Uzaktan algılamada ileri sınıflandırma tekniklerinin karşılaştırılması ve analizi”, *Gebze Üksek Teknol. Entitüsü Jeodezi Ve Fotogram. Mühendisliği Lisans Tezi*, 2009.
- [58] **L. Breiman, J. Friedman, C. J. Stone, ve R. A. Olshen**, *Classification and regression trees*. CRC press, 1984.
- [59] **G. Silahtaroğlu**, “Veri madenciliği”, *Papatya Yayın. İstanbul*, 2008.
- [60] **X. Niuniu ve L. Yuxun**, “Notice of Retraction: Review of decision trees”, 2010, c. 5, ss. 105-109.
- [61] “Data Mining: Practical Machine Learning Tools and Techniques - 3rd Edition”. **Url-31**<<https://www.elsevier.com/books/data-mining-practical-machine-learning-tools-and-techniques/witten/978-0-12-374856-0> >, (erişim Şub. 09, 2021).
- [62] **L. Breiman**, “Random forests”, *Mach. Learn.*, c. 45, sy 1, ss. 5-32, 2001.
- [63] “XGBoost”, *DeepAI*, May. 17, 2019. **Url-32**<<https://deepai.org/machine-learning-glossary-and-terms/xgboost>>, (erişim Şub. 11, 2021).
- [64] “(Tutorial) Learn to use XGBoost in Python”, *DataCamp Community*, Kas. 08, 2019. **Url-33**<<https://www.datacamp.com/community/tutorials/xgboost-in-python>> (erişim Şub. 11, 2021).
- [65] “K-Nearest Neighbors Algorithm in Python and Scikit-Learn”, *Stack Abuse*. **Url-34**< <https://stackabuse.com/k-nearest-neighbors-algorithm-in-python-and-scikit-learn/> >, (erişim Şub. 11, 2021).
- [66] **Z. Yong, L. Youwen, ve X. Shixiong**, “An Improved KNN Text Classification Algorithm Based on Clustering”, *J. Comput.*, c. 4, Mar. 2009, doi: 10.4304/jcp.4.3.230-237.
- [67] **D. Soni**, “Introduction to Naive Bayes Classification”, *Medium*, Tem. 16, **Url-35**<2019. <https://towardsdatascience.com/introduction-to-naive-bayes-classification-4cffabb1ae54> >, (erişim Şub. 11, 2021).
- [68] **J. Shubham**, “Naive Bayes Theorem”, *Medium*, Tem. 18, 2018. **Url-36**< <https://becominghuman.ai/naive-bayes-theorem-d8854a41ea08>>, (erişim Şub. 11, 2021).

ÖZGEÇMİŞ



Ad-Soyad

:Çağatay YILDIRIM

ÖĞRENİM DURUMU:

- **Lisans** : 2018 Ocak, İstanbul Teknik Üniversitesi, İnşaat Fakültesi, Geomatik Mühendisliği Bölümü